Public Health Monitoring of Behavioural Risk Factors and Mobility in Canada: An IoT-based Big Data Approach

by

Kirti Sundar Sahu

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Public Health and Health Systems

Waterloo, Ontario, Canada, 2022

EXAMINING COMMITTEE MEMBERSHIP

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Sam Liu Associate Professor School of Exercise Science, Physical & Health Education University of Victoria
Supervisor	Dr. Plinio P. Morita Associate Professor School of Public Health Sciences University of Waterloo
Internal Member	Dr. Shannon E. Majowicz Associate Professor School of Public Health Sciences University of Waterloo
Internal Member	Dr. Joel A. Dubin Professor Department of Statistics and Actuarial Sciences School of Public Health Sciences University of Waterloo
Internal-External Member	Dr. Arash Arami Assistant Professor Mechanical and Mechatronics Engineering University of Waterloo

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

STATEMENT OF CONTRIBUTIONS

The two manuscripts presented in this thesis, that have been published, are the work of Kirti Sundar Sahu, in collaboration with his co-authors.

Chapter 5: A New Approach to PASS Indicators: Are Data from IoT Technologies Informative? Is a modified version of the article “Enabling Remote Patient Monitoring Through the Use of Smart Thermostat Data in Canada: Exploratory Study” published in JMIR mhealth and uhealth (November 20, 2020).

As lead author of this chapter, I led the conceptualization of the study design and designed the data collection scripts. This manuscript was co-authored with Plinio P. Morita and Arlene Oetomo. I developed the theoretical framework for the paper in collaboration with my supervisor Plinio P. Morita. I and Arlene Oetomo prepared the ethics application for the study. Arlene Oetomo helped in recruitment of the participants for the study. I have collected, managed, and analyzed data for the pilot study. I wrote the first draft of the manuscript with input from all authors, Arlene Oetomo, Plinio P. Morita helped provide overall direction and planning. All authors contributed to manuscript reading and revision and have approved the submitted version.

Chapter 6: NextGen Public Health Surveillance and the Internet of Things (IoT) is published in the journal of Frontiers in Public Health (December 03, 2021).

This manuscript was co-authored with Shannon E. Majowicz, Joel A. Dubin and Plinio P. Morita, out of which I am the main author. I developed the theoretical framework for the paper in collaboration with my supervisor Plinio P. Morita, I wrote the first draft of the manuscript with input from all authors and Shannon E. Majowicz and Joel A. Dubin helped provide overall direction and planning. All authors contributed to manuscript reading and revision and have approved the submitted version.

ABSTRACT

Background: Despite the presence of robust global public health surveillance mechanisms, the COVID-19 pandemic devastated the world and exposed the weakness of the public healthcare systems. Public health surveillance has improved in recent years as technology evolved to enable the mining of diverse data sources, for example, electronic medical records, social media, to monitor diseases and risk factors. However, the current state of the public health surveillance system depends on traditional (e.g., Canadian Community Health Survey (CCHS), Canadian Health Measures Survey (CHMS)) and modern data sources (e.g., Health insurance registry, Physician billing claims database). While improvement was observed over the past few years, there is still a room for improving the current systems with NextGen data sources (e.g., social media data, Internet of Things data), improved analytical mechanism, reporting, and dissemination of the results to drive improved policymaking at the national and provincial level. With that context, data generated from modern technologies like the Internet of Things (IoT) have demonstrated the potential to bridge the gap and be relevant for public health surveillance. This study explores IoT technologies as potential data sources for public health surveillance and assesses their feasibility with a proof of concept. The objectives of this thesis are to use data from IoT technologies, in this case, a smart thermostat with remote sensors that collect real-time data without additional burden on the users, to measure some of the critical population-level health indicators for Canada and its provinces.

Methods: This exploratory research thesis utilizes an innovative data source (ecobee) and cloud-based analytical infrastructure (Microsoft Azure). The research started with a pilot study to assess the feasibility and validity of ecobee smart thermostat-generated movement sensor data to calculate population-level indicators for physical activity, sedentary behaviour, and sleep parameters for Canada. In the pilot study, eight participants gathered step counts using a commercially available Fitbit wearable as well as sensor activation data from ecobee smart thermostats.

In the second part of the study, a perspective article analyzes the feasibility and utility of IoT data for public health surveillance. In the third part of this study, data from ecobee smart thermostats from the “Donate your Data” program was used to compare the behavioural changes during the COVID-19 pandemic in four provinces of Canada. In the fourth part of the study, data from the “Donate your Data” program was used in conjunction with Google residential mobility data to assess the impact of the work-from-home policy on micro and macro mobility across four provinces of Canada. The study's final part discusses how IoT data can be utilized to improve policy-level decisions and their impact on daily living, with a focus on situations similar to the COVID-19 pandemic.

Results: The Spearman correlation coefficient of the step counts from Fitbit and the number of sensors activated was 0.8 (range 0.78-0.90; n=3292) with statistically significant at $P < .001$ level. The pilot study shows that ecobee sensors data have the potential to generate the population-level health indicators. The indicators generated from IoT data for Canada, Physical Activity, Sleep, and Sedentary Behaviours (PASS) were consistent with values from the PASS indicators developed by the Public Health Agency of Canada.

Following the pilot study, the perspective paper analyzed the possible use of the IoT data from nine critical dimensions: simplicity, flexibility, data quality, acceptability, sensitivity, positive predictive value, representativeness, timeliness, and stability. This paper also described the potential advantages, disadvantages and use cases of IoT data for individual and population-level

health indicators. The results from the pilot study and the viewpoint paper show that IoT can become a future data source to complement traditional public health surveillance systems.

The third part of the study shows a significant change in behaviour in Canada after the COVID-19 pandemic and work-from-home, stay at home and other policy changes. The sleep habits (average bedtime, wake-up time, sleep duration), average in-house and out-of-the-house duration has been calculated for the four major provinces of Canada (Ontario, Quebec, Alberta, and British Columbia). Compared to pre-pandemic time, the average sleep duration and time spent inside the house has been increased significantly whereas bedtime, and wake-up-time got delayed, and average time spent out-of-the-house decreased significantly during COVID-19 pandemic.

The result of the fourth study shows that the in-house mobility (micro-mobility) has been increased after the pandemic related policy changes (e.g., stay-at-home orders, work-from-home policy, emergency declaration). The results were consistent with findings from the Google residential mobility data published by Google. The Pearson correlation coefficient between these datasets was 0.7 (range 0.68-0.8) with statistically significant at $P < .001$ level. The time-series data analysis for ecobee and google residential mobility data highlights the substantial similarities. The potential strength of IoT data has been demonstrated in the chapter in terms of anomaly detection.

Discussion and Conclusion: This research's findings demonstrate that IoT data, in this case, smart thermostats with remote motion sensors, is a viable option to measure population-level health indicators. The impact of the population-level behavioural changes due to the COVID-19 pandemic might be sustained even after policy relaxation and significantly affects physical and mental health. These innovative datasets can strengthen the existing public health surveillance mechanism by providing timely and diverse data to public health officials. These additional data sources can offer surveillance systems with near-real-time health indicators and potentially measure short- and long-term impact policy changes.

ACKNOWLEDGEMENTS

I would like to acknowledge the financial support received from the MITACS Accelerate Canada in the form of a Scholarship and the University of Waterloo in the form of several internal scholarships.

First of all, I would like to express my most sincere appreciation to my supervisor, Dr Plinio Morita, Associate Professor who tremendously encouraged me during my Ph.D. journey. I feel incredibly fortunate to have had an extremely positive student-supervisor relationship that is paramount to the smooth progress and success of completing my doctoral degree. The last four years of working with Dr Morita have been a whirlwind of learning and growing. His encouragement and support for getting involved in various research directions and projects have given me fresh perspectives, lots of opportunities, and academic freedom. Throughout this journey, I gained more confidence to move forward in the direction I wished to pursue. The numerous meetings and discussions with Dr Morita in the past years play a key role in eventually shaping this dissertation. He was my guiding light in more ways than one, providing me with encouragement and patience throughout my dissertation. Throughout my Ph.D. journey, Dr Morita taught me to conquer my weakness and step outside my comfort zone to explore new horizons.

I would also like to extend my deepest gratitude to my committee members, Dr Shannon Majowicz and Dr Joel Dubin. I am very thankful to Dr Majowicz for her continuous support for exploring various research directions related to public health surveillance and staying in my committee since the early days of my Ph.D. journey. I also thank Dr Dubin for providing me with invaluable suggestions for the statistical analysis in my thesis. I truly appreciate having a well-rounded committee that provided me with invaluable insights from multiple aspects. The completion of my dissertation would not have been possible without the support and nurture of my committee.

My heartfelt thanks to my brilliant labmates at UbiLab; Arlene Oetomo, Niloofar Jalali, Dhron Joshi, Frederico Bublitz, Thoko Hanjahanja-Phiri, Laura Fadrique, Raphael Mendonça da Nóbrega, Pedro Miranda, Pedro Elkind Velmovitsky, Shahan Salim, Kang Wang, Amethyst Kuang, Kiemute Oyibo, Dia Rahman, Tatiana Silva Bevilacqua, Rena Yang and Tak Yasunaga.

I would like to thank everyone who volunteered to proofread my dissertation multiple rounds and provided invaluable suggestions on editing. Special mention of Kashish Grover and Arjun Mehra who stood like pillars at the most crucial time and without whom this herculean task could not have cascaded itself into such a refined work. I very much appreciate all their help and invaluable input from many aspects throughout my Ph.D. study.

In addition to above, I would especially like to thank Carol West-Seebeck for her encouragement whenever I reached a milestone; Brian Mills, Daniel Rodgers, and Tracy Taves for their assistance on my questions about the program, degree requirements, and any coordination required throughout my Ph.D. study; Trevor Bain and Brent Clerk for their generous support on any IT-related issues I had ever encountered.

I would like to thank the Writing and Communication Center for providing me support whenever and wherever required. I would also like to thank my team at Residence life, University of Waterloo, especially Navya V Nair, Piyush Garg, and Delores Lanni for all the psychological support during my Ph.D. Journey.

A big and special THANK YOU goes to my wife Bhavna Bharati and daughter Anushka Sahu for their eternal love and emotional support endlessly. To my wife, for always believing in me, finding ways to encourage and support me throughout this journey and being by my side

throughout the toughest times. Thank you, without you, achieving my dreams would not have been possible. Thank you for understanding my decision for switching gears in both career and life, for supporting me in pursuing multiple post-graduate degrees. I would not have reached here without you!

Now it's time to close this chapter and flip to a new one. Obtaining a doctoral degree is just the start of the next chapter of my life. Beginning today, no matter where I go and what I do, I believe that "the dreams bring back all the memories, and the memories bring back you."

DEDICATION

To my wife Bhavna Bharati and daughter Anushka Sahu.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BRFSS	The Behavioural Risk Factor Surveillance System
CCDSS	Canadian Chronic Diseases Surveillance System
CCHS	Canadian Community Health Survey
CCHS RR	Canadian Community Health Survey Rapid Response
CDC	Centers for Disease Control and Prevention
CHMS	Canadian Health Measure Survey
CIHI	Canadian Institute for Health Information
CIHR	Canadian Institutes for Health Research
CUSUM	Cumulative Sum
DCCPGEC	Diabetes Canada Clinical Practice Guidelines Expert Committee
DYD	Donate Your Data
EHR	Electronic Health Record
HBSC	Health Behaviours in School-aged Children
HBSC-Admin	Health Behaviours in School-aged Children Administrator Survey
HDFS	Hadoop Distributed File System
IoT	Internet of Things
ITS	Interrupted Time Series
LSTM	Long Short-Term Memory Model
MERS	Middle East Respiratory Syndrome
mHealth	Mobile Health
PA	Physical Activity
PAM	Physical Activity Monitor
PASS	Physical Activity, Sedentary Behaviour and Sleep
PCA	Principal Component Analysis
PHAC	Public Health Agency of Canada
PHS	Public Health Surveillance
RNN	Recurrent Neural Network
SARIMA	Seasonal Autoregressive Integrated Moving Average
SARS	Severe Acute Respiratory Syndrome
WHO	World Health Organization
ZET	Zero Effort Technology

“If you can't explain it to a six-year-old, you don't understand it yourself.”

— Albert Einstein

Chapter 1 Introduction

1.1 Public Health and Public Health Surveillance

Public health helps protect and improve people's health and their communities by preventing health hazards, injury and disability by keeping the associated risk factors at bay and promoting factors that enhance the quality of life and human health at the population level [1–3]. Over the last five decades, demographic [4] and epidemiological transitions [5,6] led to changes in the global disease scenario and the causal patterns of death, as well as trends in morbidity and mortality [7]. As technology progressed, along with shifting demographic patterns where the proportion of elderly are more than the proportion of children, and epidemiological trends, where there is a triple burden of diseases with existing infectious diseases, growing non-communicable diseases and the problem of multimorbidity, there is a need for near real-time data that provides insights to enable proactive planning and intervention initiatives [8,9]. In current public health systems, surveillance's role in acquiring data is of utmost importance, as better data comprehensiveness will lead to early identification of anomalies and better preparedness to handle this [9]. Traditional public health surveillance methods mainly rely on self-reported data, where health institutions or agents in the community are responsible for every step of the data journey. This multistep process starts with collecting data, followed by the entry and curation of the data. Sometimes, the process of handling this data is entirely manual and involves digitizing hard copies of documents, and there is a chance of a reduction of data quality. This process ends with data analysis and dissemination of the findings through reports and publications [10]. Public health surveillance aims to help public health officials understand present situations and prevent future disease burdens [11].

1.2 Technological Evolution in Public Health Surveillance

As a result of technological evolution, public health surveillance methods have also evolved [12]. Public health surveillance has modernized its processes by utilizing technological innovations, driven by the digitization of outdated paper-based record systems [13]. Monitoring emerging public health threats requires the active use of electronic surveillance systems. Controlling obesity can be supported by digital eHealth technology, for example, a multi-faceted monitoring system of lifestyle behaviours, including food consumption and its patterns; physical activity; fitness and sedentariness; biological, socio-cultural, environmental determinants (e.g., alcohol consumption and smoking as an adolescent) [14]. Other ways eHealth technologies might be applied in public health surveillance influenza prevention using data from social media [15–17] or getting population-level health parameters using data from the wearables [18–20]. Surveillance systems provide public health officials with descriptive information about challenges related to public health issues across the three essential dimensions: when, where, and who is affected [21].

1.3 Big Data and IoT as a Potential Data Source for Public Health Surveillance

The ubiquitous growth of technology helped simplify human life by reducing physical labour and saving time [22]. Novel technologies also aid in quantifying several vital parameters in day-to-day life, which were unimaginable in the past [23]. During this process, large volumes of data are being generated (colloquially termed "big data"), which have the potential to answer complex human community-related questions, for example, outbreak identification, management, investigation and risk communication [24]. The Internet of Things (IoT) is a new technological innovation through which any physical device can be connected to the internet, and communication between different devices is possible in real-time. Such technology helped generate vast volumes of data from which numerous kinds of insights can be investigated, for example, change in human behaviour due to lifestyle change [25]. These have the potential to be used as a data source for public health surveillance, as they provide direct insights into health behaviours [26].

1.4 Motivation

The motivation of this thesis is to explore the use of novel data types to understand population health in Canada, aiming at improving the current public health surveillance systems in Canada. As part of this process, I explore the feasibility and validity of a population-level monitoring platform using real-world data from IoT. Towards these objectives, I have studied the use of population-level analytics while addressing the challenges of current surveillance systems (e.g., delay in the monitoring, lack of real-time information, missing data, recall bias) using novel data sources for supporting public health surveillance.

1.5 Thesis Structure

This thesis is organized into nine chapters, including this introductory chapter. Chapter 2 presents the literature review related to my area of research. Chapter 3 provides the rationale behind this study with specific research questions, goals, and objectives. Chapter 4 explains the methodology, along with a figure describing the thesis. Chapter 5 presents a proof of concept of my work, which was published in JMIR mHealth and uHealth [27]. Chapter 6 corresponds to a perspective paper describing a NextGen Public Health Surveillance System using IoT data, published in Frontiers in Public Health with a special section of Digital Public Health [28]. Chapter 7 presents one of the applications of IoT data for behavioural monitoring in Canada, focusing on understanding the pandemic's impact on time spent in and out-of-the-house and sleep health through a comparison of before and during the COVID-19 pandemic. Chapter 8 elaborates on the second application of IoT data for population-level human mobility measurement and use for public health compared with Google mobility data. Chapter 9 explores policy-level discussions related to the use of IoT and alternative data sources for supporting public health practice, the conclusion and summary of this thesis are provided.

Chapter 2 Literature Review

This literature review focuses on the current evidence available in the domains of public health and health informatics, intending to represent the existing state of evidence on "public health surveillance" from a health informatics perspective. The following keywords were used in isolation and in different combinations, for this literature review: *"population health", "public health", "surveillance", "monitoring", "risk factors", "physical activity", "sleep", "sedentary behaviour", "health informatics", "big data", "machine learning"," deep learning", and "artificial intelligence", "human mobility"*. The search was conducted on the following databases: PubMed, ScienceDirect, Scopus, and Google Scholar. The literature review is limited to articles published in the last two decades, although important articles from historical journal publications were included.

2.1 Public Health

Though the history of "Public Health" can be extended to the 14th century, formally the concept of the scientific study of epidemiology-based public health started in the 19th century when the father of epidemiology, John Snow, linked the spread of infectious disease cholera with the drinking of contaminated water [29]. Public health addresses "society's *interest in assuring conditions in which people can be healthy*," as presented by the Institute of Medicine (US) Committee for the study of the future of public health in 1988 [30,31]. In 1920, Winslow defined public health as "*the science and art of preventing disease, prolonging life, and promoting health through the organized efforts and informed choices of society, organizations, public and private, communities, and individuals*" [32]. Today, public health is defined by a change in scope, the evolution of social systems, and emerging diseases [33]. In the case of infectious diseases, public health might as well be defined as "*a record of successful redefining of the unacceptable*" [34].

New microorganisms like swine flu caused by H1N1 strain of the influenza virus, COVID-19 disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), bioterrorism and similar epidemic infectious diseases [35,36] significantly impact health and mortality. As the health landscape shifts, there is also a transition in its definition [10].

Moreover, infectious diseases, chronic diseases and their impact on healthy wellbeing and lifestyle are becoming a focus of contemporary public health research [35]. The determinants of the global burden of disease can be multifold [37]. Of them, environmental and behavioural factors influence people's health significantly [38]. Globally, there will be an increase in the share of the elderly population (sixty and above), which will eventually lead to more composite health challenges related to ageing [4,7,36]. This change in age structure, inequity in the distribution of resources and utilization of healthcare services will redefine a country's disease burden [39]. This epidemiological transition can cause a double disease burden, leading to a high volume of morbidity and mortality worldwide [38].

2.1.1 Conceptual Framework for Public Health

Figure 1 describes the conceptual framework for public health developed by the Canadian Public Health Association in 2017 to understand the underlying principles that support current public health practice, including but not limited to data sources, structure and implementation of the process [29]. Public health addresses the underlying health determinants with its roots in social justice, attention to human rights and equity, and evidence-informed policy and practice. This framework emphasizes health promotion and population-level surveillance as the cornerstone to prevent diseases, injury, disabilities, and mortality. This field is a result of a multidisciplinary approach. With increasing demand, the role, purpose, and scope of the profession are changing. Utilizing a cyclic path, five distinct components strengthen the foundation of public health, namely, strong backing by evidence, risk assessment, policy, program, and evaluation, respectively. Among the components mentioned above, research and surveillance mechanisms generate evidence to start the cycle.

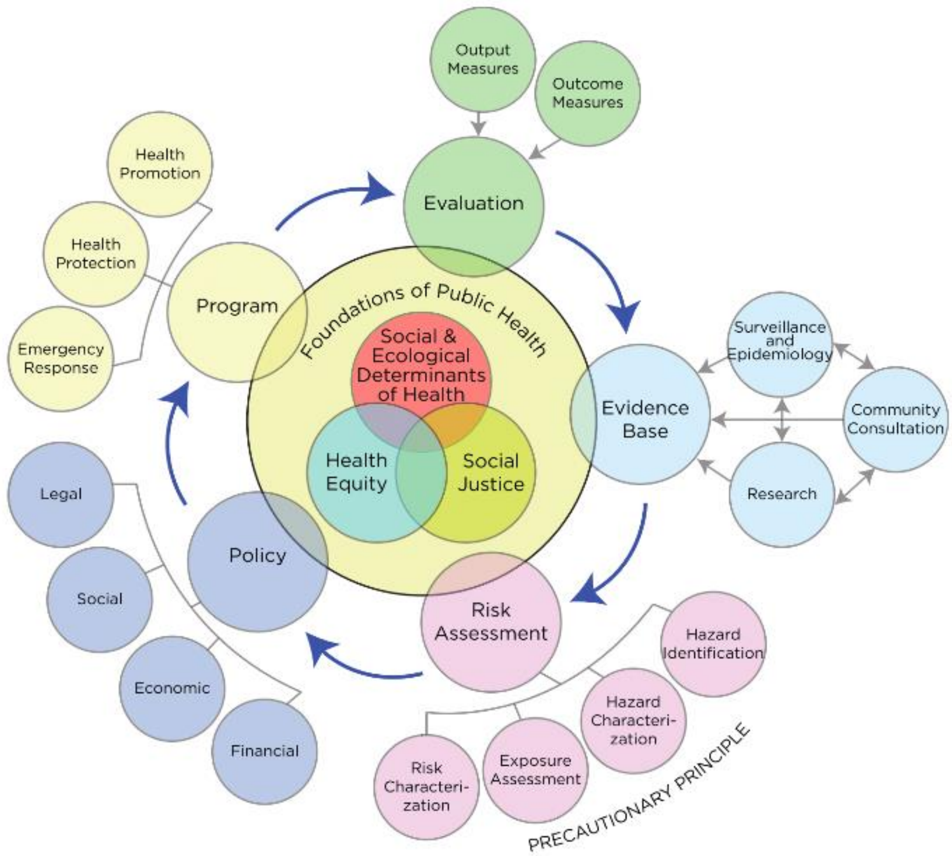


Figure 1. Conceptual framework of public health: image extracted from a Canadian Public Health Association report [29].

The public health profession relies on the robustness, accuracy, and validity of evidence with a scientific mind [40]. At the same time, the validity of existing evidence is strongly related to the research environment. Quantitative, qualitative, mixed-method research, surveillance, epidemiology, and community consultation generate evidence that makes up the framework's base. The interaction between each of the components is strong and interdependent. That means surveillance can be conducted to inform research, and research can be used to focus and strengthen surveillance activities. Overall, the consolidation of those components helps to improve population-based health issues within the general domains of communicable and non-communicable diseases. The research component is dependent on population characteristics, needs, values and preferences, and professional expertise [29].

2.2 Public Health Surveillance

Within the public health system, the role of surveillance is critical [41]. Surveillance of different diseases and their risk factors are amongst the most significant components of the public health surveillance system. Public health surveillance has a long history, which began around 3180 BC in Egypt, with the first recorded epidemic [10]. In his article "*The Past, Present, and Future of Public Health Surveillance*," Choi summarized learning experiences from a review of historical perspectives in the past 5,000 years up to the existing surveillance systems and suggested much-required modification for future mechanisms [10]. The World Health Organization (WHO) defined public health surveillance as "*the continuous, systematic collection, analysis and interpretation of health-related data needed for the planning, implementation, and evaluation of public health practice*" [42]. Surveillance systems provide us with descriptive information about health problems focusing on three key dimensions - when, where and who [21].

Historically, public health surveillance started with identifying the causes of infectious diseases, focusing mainly on epidemics and endemics [10]. With time, public health surveillance has changed its definition and scope [12,41]. Several new methodologies (e.g., telephone-based data collection [43], digital data collection methods [44]), technologies (e.g., smartphones, mobile apps, information technologies), and data sources (e.g., web searches, social media, sensors [45]) are being used to make surveillance systems more effective, efficient and timely. The technology used for public health surveillance has changed the process of data collection, analysis, and interpretation [12,13,44,46,47].

As the world witnessed an epidemiological transition in the last two decades [5], global disease scenarios, patterns in causes of deaths, and trends in morbidity and mortality were also changing [48]. Health is a complex construct, and it depends on multiple indicators [49], including biological, environmental, and social determinants of health. The broader scope of social determinants of health includes individual-level demographics; socioeconomic (e.g., education, occupation) factors as well as environmental (e.g., neighborhood, air quality, infrastructure quality) factors; and factors associated with the healthcare system, like accessibility or affordability [50]. Often, determinants of health also include genetics [51,52], drug use, alcohol consumption [53], and smoking [54].

Healthcare accessibility, for example, has been demonstrated to be affected by race, ethnicity, language, disability, mobility, distance to healthcare services, and the number of healthcare providers present in an area [55]. Health determinants and health outcomes are either directly or indirectly related, making it difficult to measure the influence of any single determinant on an outcome [56,57], which often directs researchers to a multilevel analysis [58]. Therefore, to monitor emerging health threats, the application of information and technology is necessary to collect a wide range of data from a larger population [59,60].

2.2.1 Conceptual Framework for Public Health Surveillance

The Center for Disease Control and Prevention (CDC), in 2012, illustrated the sources of information that can be used to provide an understanding of the health situation of a community [11]. As it is possible to see in Figures 2 and 3, a multidisciplinary strategy including public health surveillance is a critical component of this process. A conceptual framework for public health data feeds is a step to enhance understanding of population health and health risks associated with it. This data feed depends on a variety of inputs, including "public health surveillance," "research studies," "health surveys," "registries of vital events like births and deaths," "medical and laboratory information systems," "environmental monitoring systems," "censuses," and "other data" resources. However, a conceptual framework for public health surveillance examines several similar data systems.

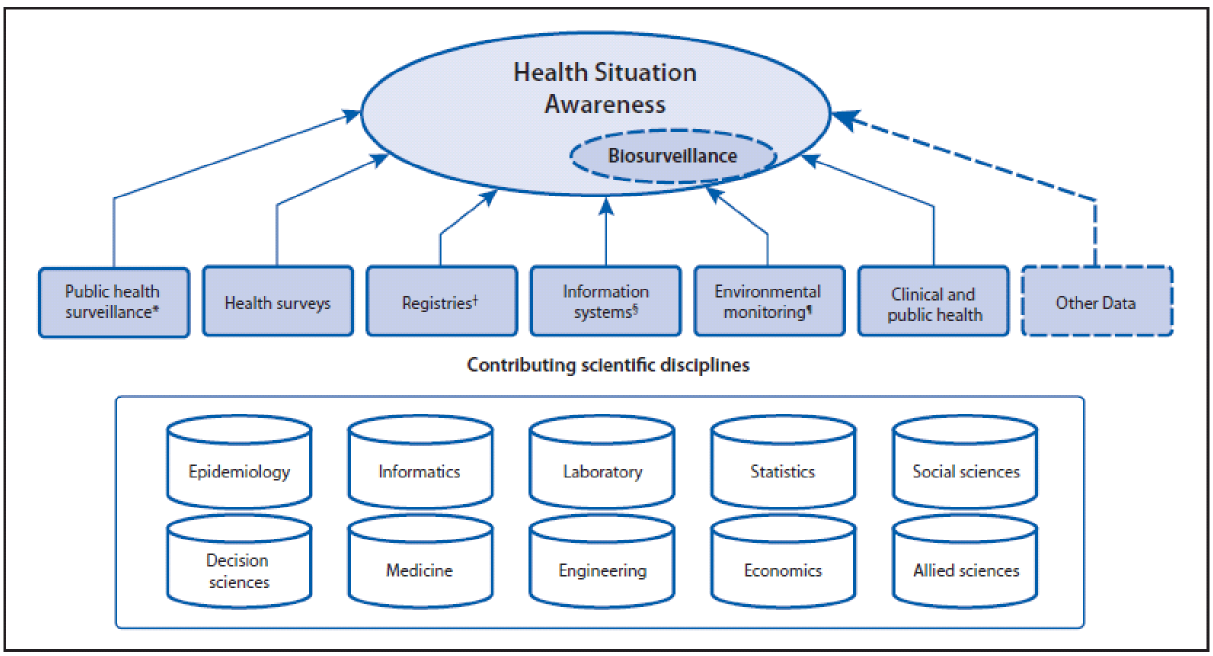


Figure 2. Various data feeds to support health situation awareness: image extracted from Centers for Disease Control and Prevention (CDC) [61].

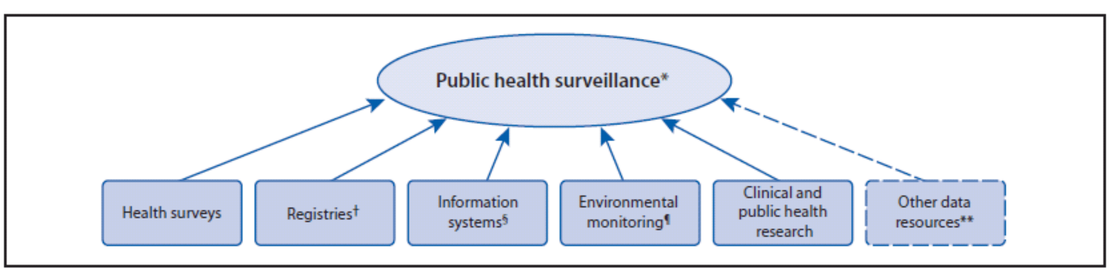


Figure 3. Conceptual framework of public health surveillance: image extracted from Centers for Disease Control and Prevention (CDC) [61].

As the modern public health surveillance mechanism captures multiple data sources, the last component, "Other data resources," still ignites the need for alternative data sources to complement the existing mechanism.

2.2.2 Population-Level Monitoring and Health Indicators

Aggregate health indicators help us by providing insights into population-level health indicators [62]. The population-level analysis helps policymakers bring new policies to bear on the problem of burgeoning chronic diseases. It not only helps to reduce the disease prevalence but also reduces the cost associated with it. As limited resources exist for healthcare services, the most effective and efficient policy has the potential to prevent disease and its complications which can dramatically improve the healthcare system. Identifying and monitoring risk factors for chronic diseases are essential for prevention [63]. According to the "World Health Report 2010," the significant risk factors for chronic diseases include: "tobacco use," "harmful use of alcohol," "raised blood pressure (or hypertension), " "physical inactivity," "increased cholesterol," "overweight/obesity," "unhealthy diet and elevated blood glucose [64,65]." Individual-level risk factors can be classified as follows: "Background risk factors (e.g., age, sex, level of education, and genetic composition)," "Behavioural risk factors (e.g., tobacco use, unhealthy diet, and physical inactivity)," and "Intermediate risk factors (e.g., elevated blood lipids, diabetes, high blood pressure, and overweight/obesity)."

2.3 The Role of Population-Level Monitoring in Pandemics

The literal meaning of the word "Pandemic" is "*all people"* and usually refers to a widespread epidemic of infectious disease throughout a country or a group of continents at the same time [66,67]. For WHO to declare a level six pandemic alert, the highest level alert related to a pandemic, there must be the disease's continued transmission in at least two regions simultaneously [66]. Historically, Spanish influenza in 1918-1920 was the biggest recorded pandemic in history [68]. With each pandemic in place, there was debate upon preventing such large-scale loss of human life using the latest available technologies and scientific methods. Public health monitoring or surveillance systems is in place to prevent, reduce the loss of human lives and protect them from such recurring pandemics. Despite the development of technologies, researchers were not able to eliminate the possibility of pandemics. Within this context, the current pandemic due to COVID-19 (2019-2021) is eye-opening for our contemporary society [69].

2.3.1 COVID-19 Pandemic as a Case Study and its Impact on Society

A new virus named SARS-CoV-2 was first observed in December 2019 in Wuhan, China, with unexplained pneumonia-like symptoms, which was named COVID-19. The cause of this new virus is thought to be associated with civets, bats and pangolins [70]. In the past, a similar kind of viral diseases emerged, named Severe acute respiratory syndrome (SARS) in 2002 and the Middle East respiratory syndrome (MERS) in 2013 [70]. However, after epidemiological and epizootic investigations, involvement of these creatures in diseases transmission has not been validated, and the possibility of an intermediate host remains elusive [70]. Since then, the causative virus has been isolated, sequenced, and was a positive-stranded RNA virus belonging to the *Coronaviridae* family, which was entirely new for humans [70]. Specifically, this virus can transmit from human to human and has a spectrum of severity that spans from asymptomatic to mild illness. The other extreme is severe diseased state and death [70]. Within three months, the disease spread to nearly all the continents and countries across the globe and on May 11, 2020, WHO declared it a pandemic [71]. Since then, COVID-19 has affected more than 1.8 billion people across the globe, and nearly four million deaths have been reported till July 2021 [72].

The effect of the COVID-19 pandemic is not limited to acute illness and deaths. However, it has also transcended into changes in household behaviours, including reduced physical activity and increased sedentary behaviour [73] due to policy-related changes such as introducing work-from-home culture and shutdown of community activities. Performing physical activity (PA) at home during the pandemic is associated with fewer mental health issues [73,74]. To maintain PA routines during COVID-19, motivation and self-interest might be vital to overcome pandemic-related barriers [75].

In August 2020, in its report of public health surveillance for COVID-19, WHO emphasized the role of digital technology and rapid reporting, contact tracing and data management to support and strengthen the existing surveillance capacity of national systems [76].

2.4 Traditional Data Sources for Public Health Surveillance

In 1968, WHO listed ten essential data sources for public health surveillance [77] and regarded data sources as the backbone of surveillance systems. The list includes mortality data, morbidity data case reporting, epidemic reporting, laboratory reporting, individual case reports, epidemic field investigation, household surveys, animal reservoir and vector distribution studies, demographic data, and environmental data. Besides these data sources, other new sources of data are being added to broaden the scope of public health surveillance such as [77]: "hospital and medical care statistics," "general practitioners," "public health laboratory reports," "diseases registries," "drug and biologics utilization and sales data," "absenteeism from school or work," "health and general population surveys" and "newspaper and news broadcasting reports." Interestingly, along with this traditional health and health-related data sources, modern health informatics has also been exploring innovative data sources like the internet [78], web searches [79], social media (Facebook [80], Twitter [81], and Reddit [82]) for health-related problem-solving.

2.5 Challenges with Existing Data Sources for Public Health Surveillance

A gap exists despite using several self-reported and alternative data sources for chronic diseases and their risk factors [83,84]. The critical challenges within existing data sources are the need for enormous resources and funding [85], the extensive time gap between data collection and report writing, the ability to accurately measure the impact of any policy level changes or even short interventions at the individual level [85]. These data sources are not real-time, and the granularity is low [85]. Self-reporting methods often result in recall bias and the potential for variation from actual reality or "*the truth*" [86].

Public health organizations sometimes depend on information obtained via questionnaire-based self-reported surveys, including online surveys, in-person or telephonic interviews, or direct measurements [87]. The usability and benefits of data collected via these methods can be impacted by declining response rates, recall bias, delays between data collection and reporting, and rising data collection costs [86,87]. Broadly, the gap within public health surveillance can be divided into indicator gaps and data gaps [88], but they are interdependent. Indicator gaps show that we do not have enough data to build comprehensive indicators for health [88] and data gaps represent there is room available for the addition of newer datasets, including data from alternative sources.

2.5.1 Need for Alternative Data Sources for Public Health Surveillance

The recent development and uptake of information technologies, computer science, the internet, smartphones, and social media by a large share of the population create large volumes of data [89]. Using these datasets, researchers attempt to build and analyze potential health indicators, as done by Dalton in 2017 for monitoring Google Flu [90], Gomide in 2011 to measure epidemiological indicators for dengue [91]. Wearables are another potential data source. Fitness tracker data (i.e., Fitbit, Garmin) can be used for tracking steps, heart rate, and sleep, which have been previously used in public health where Naghmeh Rezaei, 2021 used Fitbits for population-level studies [20] and Lisa J. Meltzer, 2015 and Dao PD, 2017 used Actigraphs [92,93].

The benefits of these alternative data sources are high granularity [94], the auto-generated and objective (non-self-reported) nature of the data [95], having high validity [96–99], minimal effort [100], and near real-time access to the generated data [9,101–104]. Mobile apps are the ultimate source of big data for physical activity and sedentary behaviour monitoring, and this data source has been previously used in public health [105].

The disadvantages of these data sources are access, quality of the data, source of the data, volume of the data, skills essential to handle the volume and types of data and interoperability within different data sources to name a few [28].

2.6 The Growing role of Technology in Public Health Surveillance

In traditional public health surveillance systems, large amounts of human resources are used to manually collect data from individuals, households, institutions, and communities [41]. Usually, the collected data is compiled manually and is analyzed using various statistical tools [59,106,107]. With improvements in science and technology, the process of surveillance has continuously evolved from a traditional model to a more technology-dependent strategy [10]. The range of transformation is witnessed at the level of sample size calculations [52,108]and goes up to report writing and disseminating information to a broader audience [109,110]. Data collection, entry, and analysis became computerized, report preparation became automated, and data visualization techniques became more sophisticated through upgraded software and technologies, making knowledge translation easy and timely in public health surveillance [10]. The latest trend is to use data collected from the daily activities of individuals using mobile devices and sensors, with minimal disturbance to an individual's daily routine. These large data sources help to assess the population for health indicators in near real-time [24,44]. Evidence highlights how data collected from the IoT has been utilized to monitor epidemics [111] and critical determinants of infectious diseases such as water pollution [112].

Key advantages of this evolution are listed below [113]:

1. Reduction of human effort and time, where computers replace human actions associated with data collection, entry, and cleaning.
2. Latency or time-gap between data collection and report writing has been minimized.
3. Interpretation of the data and results has been improved with new analysis methods and visualization techniques.
4. Dissemination of the findings to a broader public in a short time frame.
5. Reduction of human error throughout the whole process.

At international and global levels, several countries have tried different mechanisms to monitor communicable and chronic diseases. Within the last century, several developed [114–116] and developing countries [117,118] collected information on the design and process of building disease-specific risk factor registries. Additionally, several surveillance systems established the role of clinical, epidemiological, and policy-related information on public health. Extensive data has been collected and made available to the policymakers with attention to an increase in disease burden. Technological and methodological improvement in surveillance systems [44], coupled with an emphasis on cost-effective public health solutions, results in the effective implementation of evidence-based solutions. Integrating data sources from different entities to provide comprehensive solutions is essential as health is a complex phenomenon both at individual and population levels. The use of a "primary key" or "unique individual level identifier" provides a means for linking different digitized databases and files to bring together tailored data in the form of patient demographics, medical history, clinical treatment, and outcomes, which has significantly facilitated the design and application of population-based surveillance studies [24]. Several developing countries are either currently utilizing or considering using a unique patient health identifier [119].

A case of technology-based health indicators monitoring system has been described below to elaborate on the technological and methodological transitions happening in Canada's public health surveillance system.

2.6.1 Technology-based Health Indicator Monitoring in Canada

Over the last five decades, chronic disease surveillance methods have evolved along with Canada's newest and latest technology [120]. In 1969 the first event-based "*National Cancer Incidence Reporting System*" was created [114], followed by the beginning of the patient-oriented "*Canadian Cancer Registry*" after 23 years, in 1992 [114]. Similarly, for diabetes, in the year 1999, the Government of Canada committed to developing a pan Canadian Diabetes Strategy [121]. After almost a decade, in 2009, the "*Canadian Chronic Diseases Surveillance System (CCDSS)*" was initiated, which is a collaborative network of all provincial and territorial surveillance systems for a group of chronic diseases supported by the Public Health Agency of Canada (PHAC) [122]. PHAC is the body that coordinates the whole process of data collection, analysis, and the preparation of reports that help policy makers formulate policies for almost twenty life-threatening chronic diseases broadly divided into cardiovascular diseases, chronic respiratory diseases, mental illnesses, diabetes, musculoskeletal disorders, neurological conditions, and their risk factors, in Canada [123]. This database focuses on capturing nationally comparable data for incidence, prevalence, mortality, complications, comorbidities, and health service utilization (rate of hospitalization, surgery, and other interventions) [124].

Presently, in Canada, the surveillance system for chronic diseases is based on secondary data sources, such as health insurance, administrative databases, and pharmacy databases [120]. The surveillance system includes diabetes, hypertension, cardiovascular diseases, cancer, and respiratory diseases. The system aims to provide early warnings, impact assessment, policy development, policy evaluation, risk assessment, generation of hypotheses for research, recognizing trends, and informing policy and programs for those specific diseases at the federal and provincial levels [125].

Within CCDSS, data processing has four major steps, as described in Figure 4. In the first step, PHAC request data from the provinces and territories of Canada and in return, the provinces and territories collect the required information from administrative health databases using data processing software provided by PHAC to the technical and science committee members. The Technical Committee, comprised of representatives from PHAC and the province/territories (P/Ts), is responsible for overseeing the implementation of the analytic process within each P/T. The members of this Committee participate in the design, development, and operation of SAS analytic code. The Committee also makes recommendations to PHAC on how to maintain and improve the system and interpret the results. The Science Committee, which is comprised of P/T representatives and scientific experts from academia, reviews feasibility studies, reviews and approves methods for constructing chronic disease case definitions and other measures required for ongoing surveillance (e.g., measures of comorbid conditions, health service use and costs), and provides oversight for issues of data quality, and priorities and opportunities for validation activities.

In the second step, the provinces and territories apply definitions to the administrative data to identify chronic disease cases. The data are reconciled internally and with other data sources to ensure consistency and accuracy of the information. In the third step, the output from the "*case definitions*" is processed by incorporating it into registries, including one record per person per fiscal year for each province and territory. In the final step, each province and territories submit aggregate information compiled from PHAC registries. PHAC analyzes the aggregate data and prepares national and province, and territories data products [125].

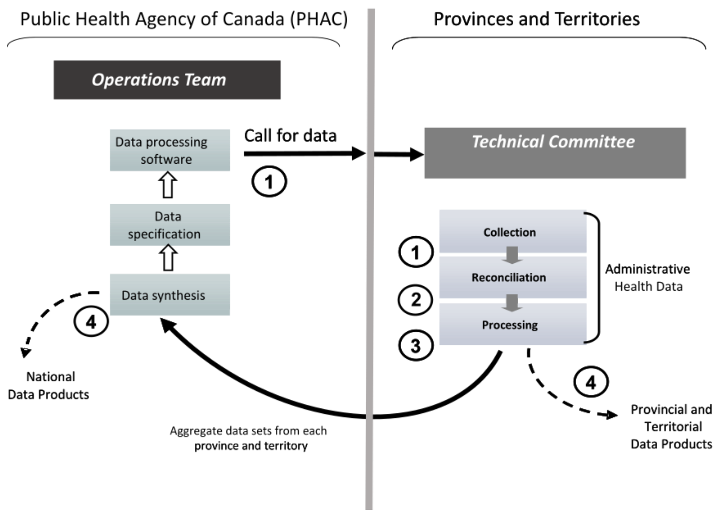


Figure 4. The existing data processing framework of the Canadian Chronic Disease Surveillance System: image extracted from Lix et al. 2018 [120].

Canadian chronic disease indicators are also pan-Canadian resources to understand the burden of chronic diseases and their associated determinants (also known as "*risk factors*") [126]. Determinants of chronic diseases have been grouped into "social and environmental determinants," "maternal and child health risk factors," "behavioural risk and protective factors," "risk conditions," "disease prevention practices," and "health outcomes/status" [126]. Within behavioural, risk and protective factors, the following indicators are listed: *"24-hour movement, Physical activity, sedentary behaviour, sleep, nutrition, chronic stress, alcohol use, smoking, drug use, and the percentage of the population having at least one of four primary chronic disease risk factors: tobacco smoking, physical inactivity, unhealthy eating, harmful alcohol use in the population aged 20 years and more"* [126].

The existing data sources for these chronic diseases include self-reported surveys, administrative data sources from health insurance registries, hospitalization databases, physician billing claims databases, and a prescription drug database. The various data sources for behavioural risk factors include [127]: Canadian Community Health Survey (CCHS) [128]; Canadian Community Health Survey Rapid Response (CCHS RR) [129]; Canadian Health Measures Survey (CHMS) [130]; Health Behaviours in School-aged Children (HBSC) [131]; Health Behaviours in School-aged Children Administrator Survey (HBSC-Admin) [127]; Physical Activity Monitor (PAM) [132].

2.6.2 Gap in the Existing Public Health Surveillance Systems

Despite using modern surveillance systems in Canada, no process to collect real-time information about risk factors for chronic diseases exists. The current data does not match population-level indicators for chronic disease surveillance due to differences in periods for data collection [88], which results in systems that miss prevention opportunities. The United States implemented the Behavioural Risk Factor Surveillance System (BRFSS) system [43] to collect data about behavioural risk factors, which have a high impact on the development of chronic diseases. Unfortunately, there is no comparative work that has been done in Canada. As mentioned on the BRFSS website, Canada has sought technical assistance from BRFSS to develop a similar surveillance system [133]. Similarly, there is enormous potential in adding environmental health surveillance data within the chronic disease surveillance system to complement and enhance the preventive strategy, which is currently missing from the Canadian surveillance systems [124].

The CCDSS has data-related challenges such as heterogeneity in an administrative database across provinces and territories, impacting the estimates' quality and accuracy. Nonetheless, CCDSS relies on a minimum dataset common to all provinces, which may not always satisfy ideal data requirements [84].

2.7 Digital Public Health Surveillance Systems

There is a strong need for digital public health surveillance systems that can embrace the broad domain of big data and newer analytics mechanisms [134].

2.7.1 Big Data

Big data is defined as any dataset having fundamental characteristics of the "5V's," namely volume, variety, veracity, velocity, and value [89,135], as mentioned in Figure 5. Other authors add variability, visualization, venue, vocabulary, vagueness [136], viscosity, volatility, viability, validity [137] to the list above. There are several approaches to collecting, storing, processing, and analyzing big data. Big data includes traditional and semi-structured data from numerous resources such as social media sites, email, documents, sensory data, and millions of web pages [138]. The giant social networking sites like Facebook and Twitter produce data on the scale of terabytes per day [139], and this amount of data is difficult to handle using existing systems.

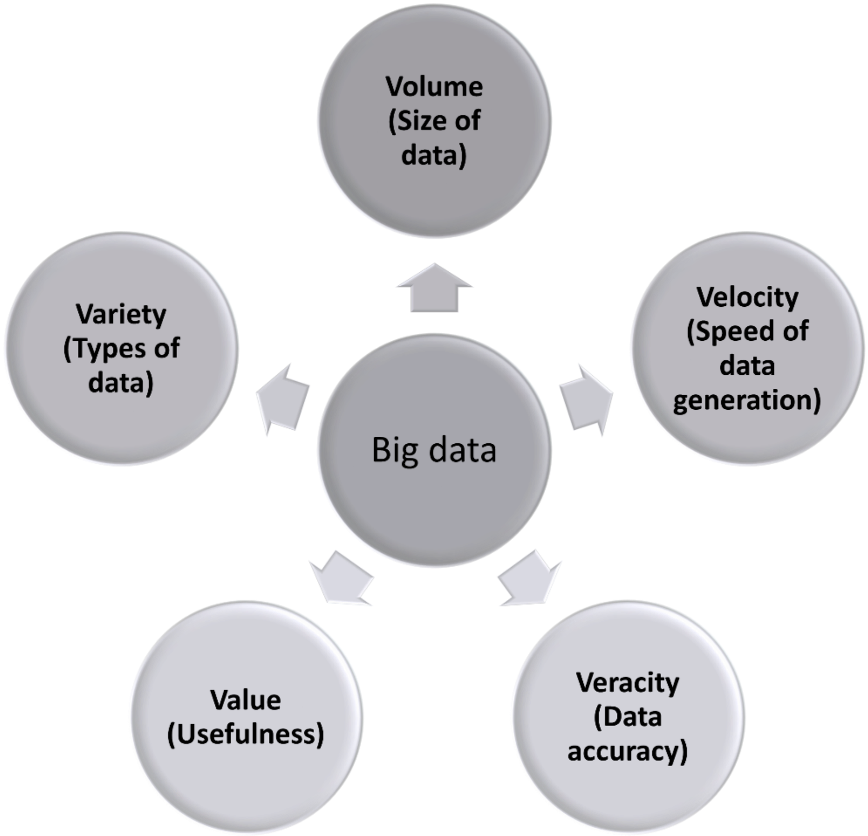


Figure 5. Characteristics of big data.

This data is different in its characteristics; it can be raw, structured, semi-structured, and even unstructured [140,141]. When the data does not fit any specific relational tables or data models, it is called unstructured data, and this kind of data is the fastest-growing data across all other types. Examples of unstructured data types include images, sensors, telemetry, video, documents, log files, and email. Semi-structured data is between these two opposite ends, having a partial component of the structured data and the other component remaining unstructured [140]. Loading and maintaining this amount of data is challenging, especially with the increase in social media usage, which is generally triggered by specific events like federal elections, natural calamities, or even during epidemics or endemics [142]. The term "big data" is a misnomer, as it points out only the size of the data without directing attention to its other properties. Figure 6 describes the standard big data architecture and its various components.

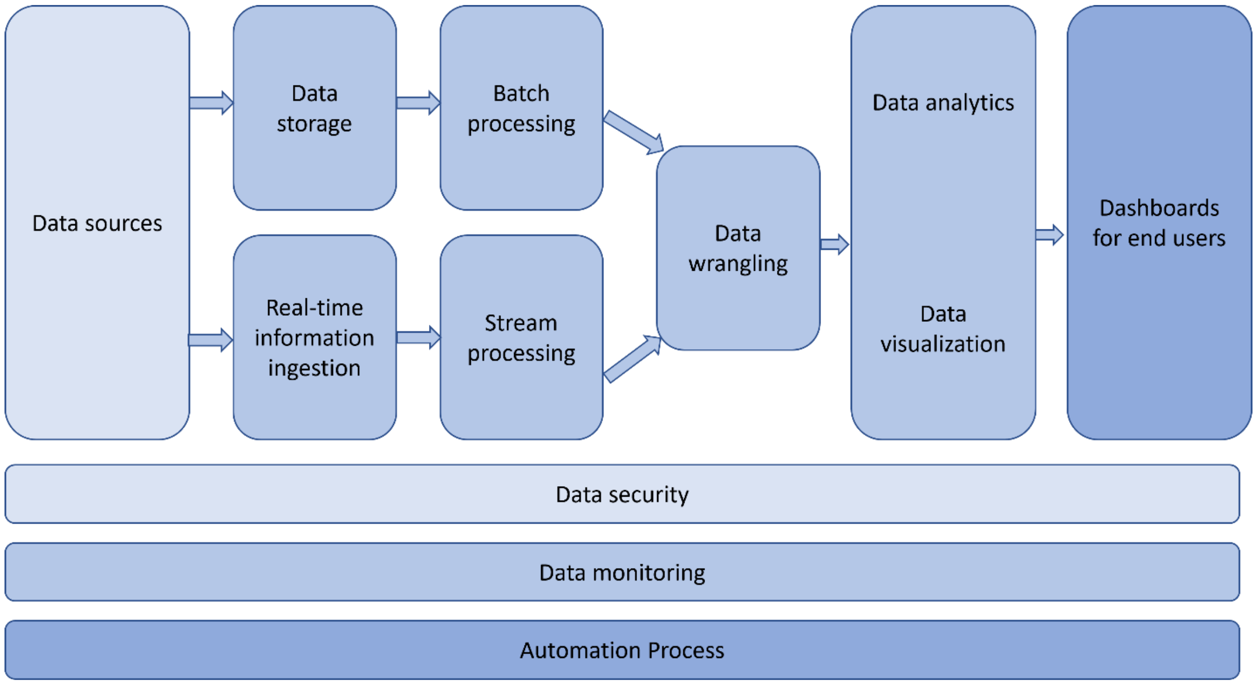


Figure 6. Standard big data architecture.

There is increasing interest in using big data technology to improve data collection and processing in healthcare, including electronic health records, IoT-like sensors, and even mobile health apps [143]. Figure 6 describes the process between big data generation and the application of artificial intelligence to derive meaning from this data. In 2011, the McKinsey Global Institute issued the report "*Big Data: The Next Frontier for Innovation, Competition, and Productivity*," the report notes that big data has the potential to transform five key domains, including public health and health care [144].

2.7.2 Internet of Things (IoT) as a Unique Data Source

IoT is a new technological innovation through which any real-world device can be connected to the internet, and communication between different devices is possible in real-time [145]. With increasing frequency and the ubiquitous presence of sensors, there is increasing potential in data generation. Each device has a unique address that can be tracked and is connected to several other devices to achieve a common objective [146].

Kevin Ashton devised the term "*Internet of Things*" in 1999 for the supply chain management domain [147]. However, within the last few years, the definition of IoT has been broadened to include various purposes, including energy, smart cities, agriculture, transport, and healthcare [148]. Although the definition of 'Things' has changed with time and improvements to technology, the ultimate objective of making a computer sense information without the help of human intervention remains the same. This inclusion of sensors and other products has revolutionized the existing infrastructure and accelerated data collection from the environment [148]. Wireless technologies like Wi-Fi, Bluetooth, RFID, ZigBee, and high-speed internet services are transforming every domain in the world, including healthcare [148]. Connectivity among people, machines, and organizations scaled up after cost reduction and increased manufacturing and availability of these devices worldwide [148].

With the ubiquitous growth of technologies, the number of connected devices exceeded the number of human beings globally in 2011. As of 2019, more than nine billion interconnected devices exist, and it is expected to reach 24 billion by 2020. According to the Groupe Spéciale Mobile Association (GSMA) [149], this amounts to $1.3 trillion revenue opportunities for mobile network operators alone across all areas such as health, automotive, transportation, and consumer electronics. The users range from individual to international level organizations addressing complex and multidimensional issues [150]. Figure 7 describes the domains where IoT has been used.

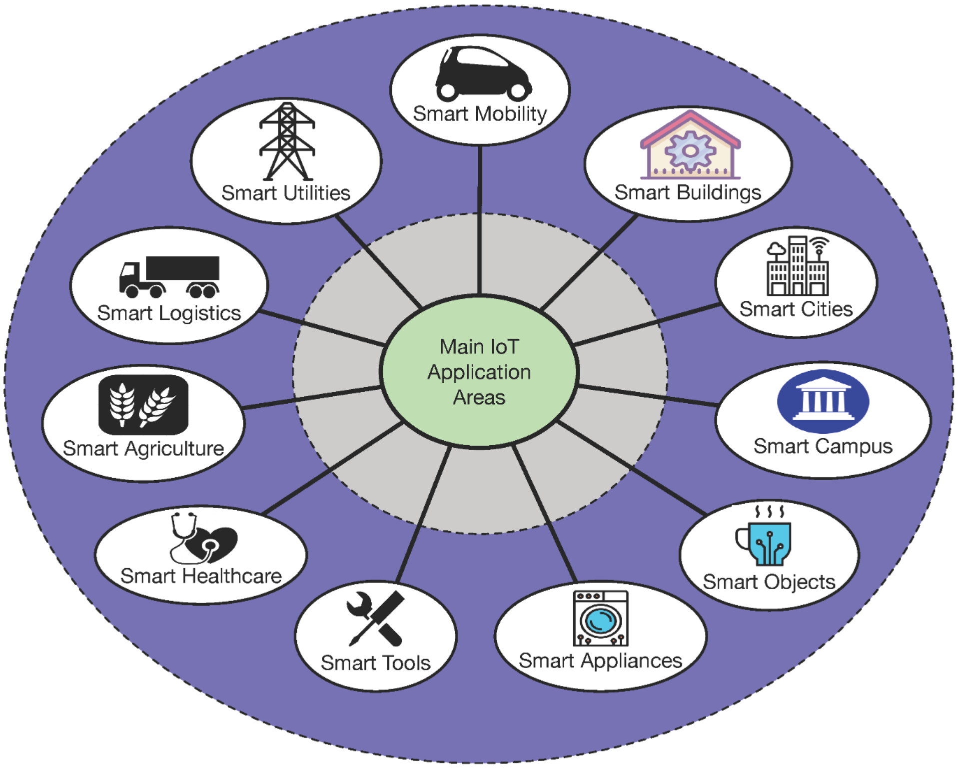


Figure 7. Internet of Things schematic showing the end-users and application areas: image extracted from Fernández-Caramés et al. 2020 [151].

2.8 Application of IoT Data for Public Health

Data from the IoTs can provide insights into different domains of public health [152–163]. The use of many wearables and fitness trackers generates a considerable volume of data, and the value of a single data source is multiplied when fused with other data sources [103,164–166]. When these innovative data sources and analyses are shared on social media platforms, the value of the integration is further augmented [167]. The range of applications of IoT data is vast, for example, elderly care, clinical medical environment, protection of natural resources and many others as listed within these reviews of the application of IoT data for different domains [101,159,168]. IoT is rapidly becoming the next generation of data sources in public health surveillance [161]. This thesis focuses on two critical uses of IoT data: behavioural monitoring and population-level mobility data.

2.8.1 Population-Level Behavioral Monitoring with Special Emphasis on Sleep, Physical Activity and Sedentary Behaviour

Population-level behavioural monitoring of risk factors for chronic diseases has its own importance in public health. Within behavioural monitoring, physical activity, sleep, and sedentary behaviour have been emphasized in the scientific literature. Physical activity is one of the key indicators associated with quality of life and chronic diseases [169,170]. Similarly, sleep deprivation, which directly correlates with chronic diseases [171], is a significant behavioural risk factor. The prevalence of sleep issues is widespread, and sleep duration has decreased across all age groups [172]. Another new risk factor recently added to the group of behavioural factors in sedentary behaviour [173]. The PHAC combined this group of indicators (Physical Activity, Sleep, and Sedentary behaviour) into the PASS indicators framework [127]. The WHO recently published a guideline for PASS for children below five years of age [174,175]. Measuring these critical health indicators at the population level is challenging, and technological solutions can provide real-time updates on these health indicators use. The use of modern data sources incl, using IoT, has been tested in different countries [52,108,176–178], but each of them has its own strengths and challenges.

2.8.1.1 Significance of PASS Indicators on Life Course Perspective

The average duration of sleep, physical activity, and sedentary behaviour have changed in Canada within the last decade. In 2012, Bin *et al.* concluded that average sleep duration was reduced in Canada by approximately 20 minutes from 1986 to 1998 [179]. In contrast, another study in 2017 indicated that average sleep duration increased from 8.1 hours in 1998 to 8.3 hours in 2010, where average screen time (one of the proxy indicators for sedentary behaviour) increased from 140 min in 1998 to 154 min in 2010 [180]. Sleep duration and screen time were positively associated in both periods. The percentage of people sleeping for less than six hours reduced by one percent from 9.6 % to 8.6% in 2010 [180]. However, within the last decade increase in screen time has not influenced the overall duration of sleep.

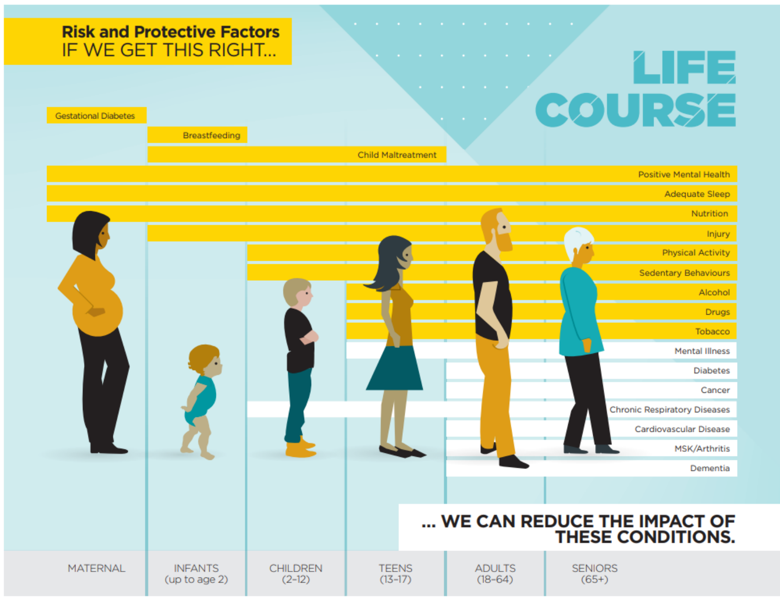


Figure 8. Risk and protective factors for chronic diseases in Canada- life course approach: image extracted from a report of Public Health Agency of Canada [181].

As per the PHAC "*Centre for Chronic Disease Prevention strategic plan 2016-2019*", sleep, physical activity, and sedentary behaviours are the critical indicators that are essential for the prevention of a range of chronic diseases for almost all stages of life [181] (Figure 8).

2.8.2 Population-Level Mobility Measurement

Human mobility data is a modern way to measure population-level activity in society [182–184]. This data can be broadly divided into macro and micro measurements based on the length, duration, and type of travel.

2.8.2.1 Macro-Mobility Measurement

Human macro mobility is defined as any long-distance travel, including travelling with vehicles [183,185,186]. Measurement of this kind of mobility at the population level depends on the type of route, surface, air, and water. Some of the critical sources for this data are flight traffic information, road transportation data and waterways movement information. With restrictions due to the COVID-19 pandemic, all three modes of transportation have been reduced significantly [187].

2.8.2.2 Micro-Mobility Measurement

Micro-level human mobility restricted to in-house movement is often ignored as a non-significant source of physical activity, recently emphasized by the PHAC. As per evidence, the exercise of any intensity has an impact on human health [127]. The daily household-related activity comes under the "light to moderate" level of physical activity [127]. The increasing use of GPS technology in smartphones, wearables, and smart home sensors has increased the generation of micro-mobility data. There is a significant potential for using this type of data for informing public health [183,188,189].

2.9 Data Analysis Frameworks

Many statistical methods have been developed to analyze the health-related data to identify unusual patterns in data series that may result from disease outbreaks [190]. In recent days, newer surveillance systems include datasets that monitor several variables and events of interest [191–193]. As the volume of surveillance data increases, innovative statistical methods are pursued to address multivariate surveillance scenarios [193,194]. Public health surveillance depends on data from several sources, and statistical analysis and methods are essential to generate value from the data. The key indicators from public health data are incidence, prevalence, rate ratio, risk ratio, odds ratio and the association between different variables [21].

Statistical analysis can be (1) descriptive, (2) inferential analysis and (3) predictive or forecasting and modelling. Within descriptive statistics, data aggregation is done using central tendency, dispersions, and variations [195]. The data aggregation can be based on the three fundamental principles of public health surveillance - place, person, and time- factors [195]. Within inferential statistics, a random sample of data is taken from a population to describe and make inferences about the hypothesis[191]. The critical difference between descriptive and inferential analysis is that descriptive statistics use the data to describe the population through numerical calculations, graphs, or tables. In contrast, inferential statistics make inferences and predictions about a population based on a sample of data taken from the population in question [44,195,196]. Predictive forecasting or model building focuses on determining the causation or associations and forecasting future trends [197–200].

The critical algorithms for public health surveillance are outbreak detection, situational awareness, and trend estimation for different health conditions and risk factors [44,195,200] .

The temporal component of the analysis can be extended by using the time series model [201]. A Time Series data is a sequential set of data points, measured typically over successive times [201]. Time-series data analysis explores trend, seasonality, cyclic and irregular characteristics and provides meaningful insights [201,202].

A trend is defined as the "*general tendency of a time series to increase, decrease or stagnate over a long period*" [202]. Seasonal variation is the component that explains "*fluctuations within a year during the season, usually caused by climate and weather conditions, customs, traditional habits, etc.*" [202]. Cyclic variation is the component that describes "*the medium-term changes* *caused by circumstances, which repeat in cycles*" [202]. The duration of a cycle extends over a more extended period. Irregular or random variations in a time series result from "*unpredictable influences, which are not regular and also do not repeat in a particular pattern*." These variations are due to extraordinary events like a pandemic, war, strike, earthquake, flood, revolution, etc. [202].

Time series data for public health has been utilized for model building [203], and time series forecasting methods have been accepted in other research fields, such as infectious disease surveillance [203–207]. Statistical models used for time-series data in public health surveillance are classified into univariate statistical methods, including statistical process control inspired models [208–210], smoothing models [196,197,211], regression methods such as generalized linear models [212], autoregressive models [213], moving average models, ARMA, ARIMA [214–216], SARIMA [217–219] models respectively. The advanced analytics methods include Bayesian models [199,220], Markov models [221], and multivariate analysis, including principal component analysis, multivariate cumulative sum, parallel surveillance and ensemble approach [17,222]. Recently the use of artificial intelligence models such as machine learning [219] and deep learning methods have also been used with time series data for building health surveillance models. The detail about these methods and models are mentioned in Appendix 1. Most surveillance data exhibit strong time trends, cyclic patterns, and other time-dependent effects depending on the aggregation method and underlying population behaviours and environmental factors.

Besides time-series data analysis methods described in Appendix-1, interrupted time series (ITS) analysis can be used where data are measured at multiple time points, i.e., before and after the introduction of an intervention to investigate the impact of that intervention [223]. ITS designs are used to examine the effects of any public health policy interventions when a randomized

controlled trial is not feasible [223]. Also, it can be used to evaluate the effects of policies and population-level interventions retrospectively using administrative databases [223]. One of the advantages of ITS design is that it can account for the pre-intervention trend in estimating the effect of the intervention [224]. Confounders may play a role both before and after interventions, therefore this must be addressed in the model [224]. Understanding non-linear correlations requires discretely graphing exposure and outcome across time before undertaking time series regression analysis [223]. Temporal confounders, such as seasonality and long-term trends, are regularly found and can lead to confounding bias [224]. Time-varying confounders, both measured and unmeasured, are also responsible for causing bias in exposure-outcome relationships [224].
In the case for epidemiological questions related to short-term variation for exposure, a generalized additive model (e.g., log-linear semi-parametric) can be used for regression modeling [225]. Explicit parameters and non-parametric functions can be employed as explanatory variables and model smoothers in this type of model. However, this model fit is ideal, since outcome values are discrete counts of total number events (e.g., mortality, disease) at a specified time point [225].
Artificial intelligence is growing in all domains, including health systems and public health [138]. In the case of public health surveillance, most published research projects utilize social media data as a source of big data, applying machine learning algorithms for model development. Examples of studies related to epidemiological indicators for infectious diseases like the influenza [16] and dengue [91] and other risk factors for health like physical activity [226,227] have been successfully described in the literature using machine learning and deep learning.
Human activity recognition is another required field where typical machine learning and deep learning algorithms have been applied to big datasets. Obinikpo and Kantarci, 2018, conducted a

feasibility study on integrating sensory data generated from different sources and presented insights on the aggregation of heterogeneous datasets with minimal missing data values for future use and found that sensory data generated from wearables are less vulnerable to missing data [228]. In 2017, Willets *et al.* showed that physical activity and sleep behaviours could be classified with 87 % accuracy using data from wrist-worn activity monitors and balanced random forests with hidden Markov models [108]. Trained models can be used to generalize to population-level indicators. The authors also analyzed the seasonal and gender-based variations of physical activity and sleep [108]. We should redefine a typical epidemiological study where machine learning, deep learning, big data, and integration of various data sources have been implemented. Those innovative ideas will be helpful only when the study design is based on some specific theoretical framework augmented with questions related to population health [229].

Machine learning approaches for finding similarities and differences in a set of data or documents include clustering and classification methods [230]. These methods can be used to group products in a catalogue, identify cohorts of similar clients, or group documents by topic or theme [230]. Although both methods have certain similarities, the difference is that classification assigns things to human-labeled classes, whereas clustering identifies similarities between objects and groups them according to common traits that distinguish them from other groups of objects. The groups are known as clusters. The clustering doesn't require an existing labelled data set, but it still seeks to identify groupings and differences in the data. This is called unsupervised learning, as opposed to classification (with labels), which is referred to as supervised learning. Clustering can be done in several ways. Each approach is best suited for a specific data distribution [230]. A brief description of four common approaches follows.

- Centroid-based clustering organizes the data into non-hierarchical clusters. K-means is the most widely used centroid-based clustering algorithm. Centroid-based algorithms are efficient but sensitive to initial conditions and outliers.
- Density-based clustering connects areas of high density into clusters. This allows for arbitrary-shaped distributions as long as dense areas can be connected. These algorithms have difficulty with data of varying densities and high dimensions. Further, by design, these algorithms do not assign outliers to clusters.
- Hierarchical clustering creates a tree of clusters. Hierarchical clustering, not surprisingly, is well suited to hierarchical data, such as taxonomies.
- Distribution-based clustering approach assumes data is composed of distributions, such as Gaussian distributions [231]. As distance from the distribution's center increases, the probability that a point belongs to the distribution decreases.

Chapter 3 Study Rationale and Objectives

3.1 Rationale of the Study

Public health surveillance systems for communicable, non-communicable diseases and their risk factors use real-time data collection and analysis mechanisms under global and national systems [84]. In Canada, the prevalence of non-communicable diseases is increasing, and so are the risk factors. The lifestyles of Canadians are also changing rapidly, and consequently, so are behavioural indicators. As indicated in the literature review in the previous chapter, among different preventable behavioural risk factors, sleep [232–235], physical (in)activity [236] and even sedentary behaviour [237] are critical determinants for health, the impact of which can be short or long term. The public health surveillance system uses indicators like behavioural risk factors to identify, track and mitigate the root cause of the problem. New indicators, data sources and infrastructure for reinforced analytics to capture behavioural and environmental risk factors for public health monitoring are slowly evolving. For policymakers to build policy [238–240] to promote healthy behaviours at a population level, like physical activity or sleep, the existing indicators may be misleading as data is outdated and sparse [241–243]. Existing datasets are not current, and data collection methods include self-reporting and the use of technology to supply data from activity monitors. Activity monitoring is one of the modern ways to measure physical activity levels. As a best-case scenario of the public health surveillance system, the USA started collecting data for these indicators in 1984, routinely using behavioural risk factor surveys. In 2011, the data collection process changed to telephonic mode to monitor the trends and patterns at the population level [43].

In Canada, however, despite several existing small-scale studies that successfully use nontraditional datasets to address timeliness and quality gaps, efforts by the Public Health

Agency of Canada (PHAC) to build a system capable of using these nontraditional datasets to monitor health determinants in mainstream public health surveillance systems have emerged [244]. For a developed country like Canada, where the technology for health is readily available, those gaps can be filled using nontraditional data collection methods such as social media and IoT data, including sensors from smart homes.

The use of smart wearable devices for health monitoring is increasing in Canada [245,246] and globally [247]. Fitbit [248], Garmin [249], and other fitness trackers [250,251] are used by individuals trying to monitor their health. Similarly, mobile apps are also producing vast amounts of data that can be used to monitor indicators such as physical activity [252,253] and sleep [254] at the population level.

Other potential data sources for public health surveillance are smart homes, IoT, and Active Assisted Living systems [255]. These have been used for public health surveillance, as demonstrated by Costa in 2014 for medical diagnostics and intervention [256] and Dalton in 2017 for influenza tracking [90].

“ecobee” is a Canadian smart thermostat manufacturer invested in using their data for research applications [257]. Smart thermostats are wireless internet-connected devices with a ubiquitous presence across most provinces and territories in Canada. ecobee operates a program called "*Donate your Data*" [258], in which any ecobee user can share their anonymized smart thermostat data for research purposes by consenting within the app [259,260]. Also, through existing web Application Programming Interface (API), which is a software intermediary that allows two applications to talk to each other, it is possible to extract pre-consented data from many devices in the community in real-time [257]. These devices are equipped with remote sensors that can monitor room usage through infrared sensors [261], which present an excellent opportunity for in-house activity and sleep monitoring, as demonstrated by other authors in the field, using infrared sensors on a smaller scale [262]. However, this type of data has never been used at such a large scale for public health surveillance. This thesis explored the use of these alternative data sources for sleep, physical activity, and sedentary behaviour monitoring in Canada using IoT technology.

3.2 Overarching Goal

The goal of this thesis is to help drive the use of novel data sources in public health, ultimately providing agencies as the PHAC and the CDC with supporting evidence of the benefits of IoT data for public health research. The results of this research will help policymakers improve programs and help identify the short-term impact of interventions or policy changes.

Through the use of IoT data collected from ecobee thermostats, this project aims to identify how the amount of physical activity and sleep changes over time.

3.3 Objectives of the Research and Research Questions

The objectives and the associated research questions of this thesis were to:

O.1 – Identify the association between wearable data and smart thermostat data through a pilot study.

> RQ1 – What is the association between step data collected by wearable devices and motion sensors activation data from smart thermostat?

O.2 – Evaluate whether the "Donate your Data" (DYD) dataset from ecobee is a possible source of data to measure population-level health indicators for in-house physical activity, sleep patterns and sedentary behaviour (PASS).

> RQ2 – What is the viability of using sensor-based data from the "Donate your Data" (DYD) program to measure in-house physical activity, sleep patterns and sedentary behaviour (PASS) at the population level?

O.3 – Demonstrate the use of data from the Donate your Data (DYD) program, from ecobee, to measure the impact of work-from-home policy during the COVID-19 pandemic for behavioural indicators in Canada as a part of population-level health surveillance programs.

RQ3 – Did the work-from-home policy during the pandemic affected behavioural indicators, such as sleep habits, in-house and the out of home stay duration in Canada?

RQ4 – Do data generated from ecobee smart thermostats have the capacity to measure the impact of the work-from-home policy in Canada?

RQ5- Assuming a difference in sleep, in-house and out of home duration at the household level, are the observed changes also observed within weekdays, week by week and month by month?

O.5 – Use data from Internet of Things (IoT) to monitor population-level micro-mobility and compare it with macro-mobility data.

RQ6 – Is there any association between Google residential mobility data and ecobee Donate your Data?

RQ7 – Do the data collected by ecobee Donate your Data support the exploration of the variability in population-level in-house mobility through the progression of the pandemic?

RQ8 – Is it possible to detect and elicit anomalous behavioural data at the population level, helping identify deviations from stay-at-home policies?

Chapter 4 Organization of the Thesis

The research work for this thesis has been divided into four major components, as described in Figure 9. The first part of my research project is presented in Chapter 5 and involves a pilot study in which I evaluated the feasibility of using smart thermostat IoT data as a data source for health indicators. I have utilized data from the "Donate your Data" program from ecobee, combined with primary data collected from participants using wearables to assess the feasibility of developing population-level indicators for Canadians' physical activity, sleep, and sedentary behaviours. This study established proof that data from IoT could be utilized for population-level health indicators, which has been published in the Journal of Medical Internet Research mHealth and uHealth [27].

The second part of this project, presented in Chapter 6, corresponds to a viewpoint paper discussing the use of the Internet of Things (IoT) as a NextGen data source for public health surveillance. This chapter describes the potential advantages and disadvantages of a modern data source to establish the foundations for potentially implementing IoT as a public health data source in public health surveillance systems. This chapter has been published as a perspective paper in Frontiers in Public Health [28].

The third part of the thesis explores the use of IoT data for two applications: (1) comparing population-level behavioural indicators in Canada before and during the COVID-19 pandemic in Chapter 7, followed by (2) a population-level mobility data assessment in Chapter 8 with particular emphasis on four major provinces, Ontario (ON), Quebec (QC), British Columbia (BC) and Alberta (AB).

My research's fourth and final component Chapter 9 examines policy-level analysis and discusses using this potential data source for public health research and action.

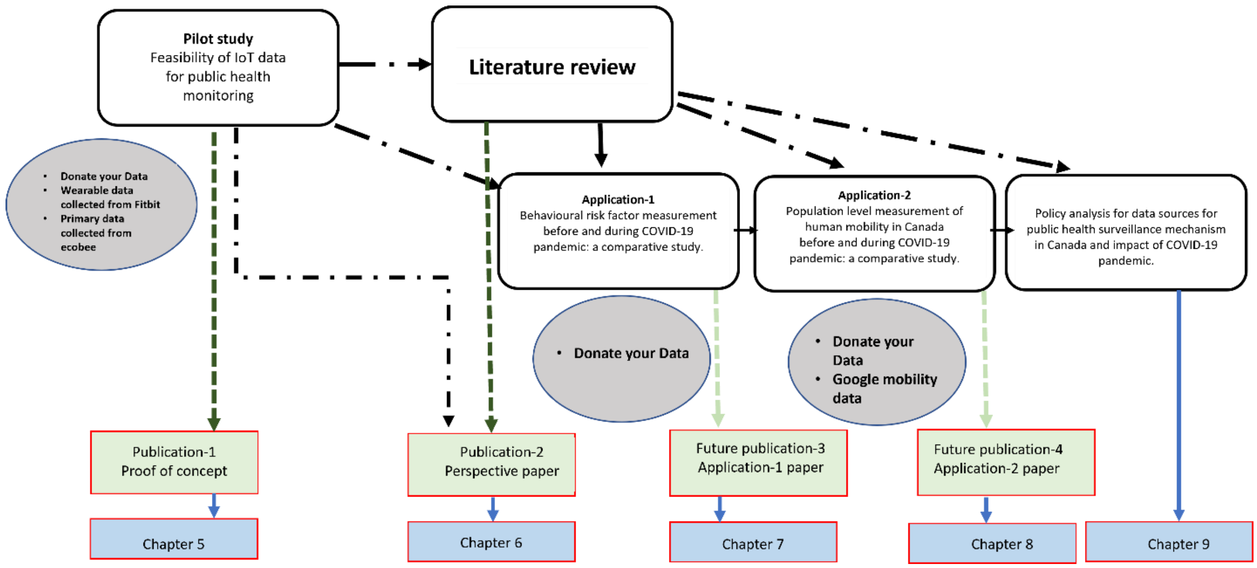


Figure 9. Research work for this thesis.

Chapter 5 A New Approach to PASS Indicators: Are Data from IoT Technologies Informative?

5.1 Preamble

This chapter presents a modified version of a published manuscript presenting the result form the pilot study to assess the use of data generated from the IoT technologies for public health surveillance, including additional analysis and discussions. The manuscript was published in the JMIR mHealth and uHealth journal, on November 20, 2020 [27]. The manuscript provided the foundational results, based on which the chapter 6, 7 and 8 developed.

Citation: Sahu KS, Oetomo A, Morita PP. Enabling remote patient monitoring through the use of smart thermostat data in Canada: exploratory study. JMIR mHealth and uHealth. 2020 Nov 20;8(11):e21016.

5.2 Introduction

In 2017, the MaRS Discovery District, in association with the Public Health Agency of Canada (PHAC) and the Canadian Institute of Health Research (CIHR), hosted a challenge to explore novel mechanisms for monitoring healthy behaviour indicators. Alternatives to traditional survey-based systems were to optimize public health surveillance systems in Canada [244].
As per the description on their website, "*MaRS Discovery District is a not-for-profit corporation and innovation hub in Toronto, Ontario, Canada, dedicated to driving economic and social prosperity by harnessing the full potential of innovation. MaRS works with entrepreneurs and investors to launch and grow companies with broad economic and societal impact. It convenes governments and industry stakeholders to enable widespread adoption in complex markets and systems*" [263]. PHAC was created in 2004 to improve public health capacity and respond effectively to public health issues in Canada [264]. PHAC focuses on preventing diseases (both chronic and infectious), preventing injuries, and responding to public health emergencies and

infectious disease outbreaks [264]. PHAC also promotes good physical and mental health across communities in Canada based on scientific evidence and informed decision-making [264]. The CIHR is the Canadian federal funding agency for health research with a mission to create new scientific knowledge and enable its translation into improved health, more effective health services and products, and a strengthened healthcare system for Canadians [265].

Through this competition, organizations looked for researchers to propose and test creative new types of data and data sources that can be used to measure indicators of the following at the population level:

1. Physical activity (number of steps);
2. Sleep (average number of hours of sleep per night);
3. Sedentary behaviour (average number of hours per day spent sedentary).

These indicators are called PASS [127] (Physical Activity, Sedentary Behaviour and Sleep), which are extracted from the Canadian Health Measures Survey (CHMS) [266], the Canadian Community Housing Survey (CCHS) [128], the General Social Survey [128], and the Physical Activity Monitor (PAM) [267]. These methods use in-person interviews, telephone surveys, and monitoring using an Actical accelerometer (Philips Respironics) for one week [268].

Innovative technologies and data sources, including IoT, mobile health applications, social networking, and other online data sources, offer an opportunity for public health organizations at different levels to access and integrate a more varied range of data into public health surveillance [44]. These datasets can overcome the limitations of current methods of self-reported data collection [8] . They can increase the granularity, diversity and range of data used as part of the analysis process, while also reducing the time lag between data collection and analysis through continuous sampling, near-real-time reporting, augmenting the ability to explore and address new areas of public health [44].

In this part of the thesis, existing technology such as smart Wi-Fi thermostats (ecobee), associated remote thermostat sensors (ecobee), as well as fitness trackers (Fitbit) have been used to evaluate the feasibility of implementing quick, real-time, and more efficient health behaviour data collection. The thermostat sensors were initially designed to monitor motion in the house to maintain a comfortable temperature in the rooms in use. The remote sensors can detect motion in the house and provide real-time, continuous assessment of the patterns of movement between rooms, which can be used to understand health behaviours such as physical activity at home, sleep quality, sleep duration, and sedentary behaviour. This solution can collect granular health behaviour data 24 hours, seven days a week, 365 days a year, without in-person follow-up. Additionally, the project (1) has the capacity to enable long term, longitudinal data to be collected directly within a home setting; (2) it is not dependent on having a physical device carried by study participants; (3) leverages existing technologies already present at home; (4) streamlines health officials' access to data, improving decision making, and policy development by providing real-time data instead of outdated data.

5.2.1 Study Objectives

The goal was to find innovative data types, sources, and methodologies to measure population-level health indicators for Canada's sleep, physical activity, and sedentary behaviour. The objective of this study was to:

1. Find the association between ecobee sensor data and step counting data collected using a Fitbit.

2. Analyze population-level indicators for an average duration of in-house physical activity, sedentary behaviour and sleep in Canada using data from the “Donate your Data” program from ecobee.

5.3 Methodology

This is an exploratory study where I have utilized two data sources, one primary and one secondary, for exploring the abovementioned research questions. In this section, I present the details of the study and more information about the different types of data used.

First, primary data was collected through a pilot study, including data from eight participants. The secondary data source consisted of data collected by ecobee through a data-sharing program known as "*Donate your Data*” [258].

Data were extracted, cleaned, loaded, and analyzed to get the evidence that ecobee sensor data is associated with Fitbit and has validity and feasibility for further analysis. Both the data sources and the process of data cleaning, loading, and analysis have been described in detail below.

5.3.1 Pilot Study

In order to explore the relationship between ecobee sensor activation, motion in the house, and sleep patterns, a pilot study (n = 8) using participants from the University of Waterloo was conducted between September 2017 and December 2017. In the pilot study, we have illustrated the potential of using a smaller number of households and the data that can be collected by leveraging sensors and fitness trackers together to create an algorithm that can be utilized to monitor healthy behaviour at a population level using more extensive datasets.

Eight subjects were recruited for this study (five females and three males), with ages ranging from 25-41 years (six graduate students and two full-time employees, all affiliated with the University of Waterloo and residing in the city of Waterloo).

Throughout the data collection period, the homes were occupied by a single resident wearing a fitness tracker. Data was collected from the ecobee thermostat (and remote sensors) and a Fitbit Zip [248]. Fitbit Zip collects step data from the participants at a minute level; Yangjian *et al.*, in 2016, validated the technology by comparing it with other similar devices [97]. Fitbit Zip has been used for monitoring physical activity, and a high positive correlation was found with the actigraphs instrument, the gold standard for measuring physical activity [269]. Each house was equipped with an ecobee thermostat and had between five and twenty-nine remote sensors installed, depending on the size of the house. Participants wore a Fitbit for about one week to collect step-tracking data in tandem with ecobee's sensor data. Layouts of the homes were obtained, the location of the sensors identified, and we ensured that the sensors had fully covered the house. Additional information about the homes and participant metadata was recorded. Spearman's correlation coefficients test was performed to validate the feasibility of the data and check whether the utilizing the sensor data is a valid method to substitute for fitness trackers for these indicators [270].

Ethics approval for this study was obtained from the University of Waterloo Office of Research Ethics (#31377).

5.3.2 ecobee "*Donate your Data*" Program

Canada has a yearly average temperature range of -1 to +1 °C, with more than 100 days below 0 °C annually [271]. As such, the use of thermostats is ubiquitous across Canada [272]. Thermostats are used to control room temperature, 92% of the households in Canada had at least one thermostat in the house in 2019, and 61% of thermostats in Canada are programmable (where the individual can set the temperature for a specific period) [272]. Heating consumes a significant amount of energy, and to solve this problem; companies started engineering smart thermostats.

Smart thermostats use presence/movement sensors to save energy and maintain the household temperature by reporting the presence or absence of a person within a specific physical location in the house. ecobee [257] and Nest [273] are the two smart thermostat companies with the highest market share for smart thermostats.

The primary advantage of the ecobee dataset is the granularity of the data (i.e., five-minute intervals), where the remote sensor data provides information about the physical presence of individuals throughout the house, which can be used to explore movement of the individual in the house. This data is highly significant as it can provide rich insights into in-home behaviour, while not requiring any additional effort to collect the data. Above all, ecobee is ready to share this data for research. ecobee has a program known as "*Donate your Data*" [258], where participants consent to share their anonymized data (without identification or demographic information) with researchers. Access to this data creates the potential to address some critical public health research questions, as I will explore on this thesis.

The "Donate your Data" dataset has two components: metadata (Appendix 2 and 3) and thermostat data files (Appendix 4 and 5). Metadata includes self-reported information about the location of the house (country, province, and city), the size of the house, the number of floors, the age of the house, the number of sensors, and the number of individuals living in the household. This data is self-reported by the user and completing the fields optional. The thermostat data files include the date and time stamps, external and internal temperatures at 5-minute interval levels, and information about room occupancy (measured through an infrared presence sensor) in binary format (0 for no presence and 1 for presence). The thermostat data files are collected automatically, compiling the data generated by the sensors. While the "*Donate your Data*" program is provided as a data export in .csv files, ecobee offers an open API for additional integration options. Ecobee users enrolled in the DYD program are distributed across almost all the provinces and territories of Canada (see Table 1).

5.3.2.1 ecobee Data Ecosystem

Passive infrared (PIR) sensors are used detect room occupancy, where the thermostat acts as the gateway to relay information to ecobee's cloud server. Each sensor reports its status every five minutes, which results in a longitudinal time series with 288 data points generated per day per household for each sensor. Through the DYD program, ecobee shares this data with researchers *via* multiple .csv files.

The dataset has 48 months of data from 111,297 households across the globe, out of which around 94% (106246) are in North America, with 14077 households specifically in Canada. The distribution at the province and territory level in Canada is described in Table 1.

Table 1. Distribution of sample size across provinces of Canada.

Country	**Province**	**No. of households**
Canada *N*=14077	Alberta	4407
	British Columbia	492
	Manitoba	275
	New Brunswick	43
	Newfoundland and Labrador	24
	Northwest Territories	5
	Nova Scotia	60
	Ontario	7734
	Prince Edward Island	7
	Quebec	767
	Saskatchewan	260
	Yukon	3

The DYD data was loaded into a MySQL database. Data cleaning and preprocessing have been completed using Python scripts to prepare the data for analysis. The second phase of this initiative focused on using the algorithms developed in the proof of concept on a larger dataset secured with ecobee (DYD) [258], providing evidence of the scalability of our platform.

Data from 7,888 households in the USA and 1,302 in Canada were used in this study after data curation, where 556 single occupant homes in the USA and 70 in Canada were included to analyze individual-level indicators. The dataset contained data from January 2015 to March 2017.

Table 2 describes the variables generated for this study. Each new column and the associated indicator were created based on pre-defined criteria. For example, an away label was created when the number of sensors activated was "0" consecutively for at least 24 hours as zero sensors activated means there was no movement within that household, representing an absence of all individuals during that period. Anything other than zero represents that somebody was at home during that time. If one or more individual is at home, then the labels are either sleep, active, or sedentary. Two sub-labels were created for sleep: true sleep and disturbed sleep. For sleep, I have selected the time window from 10:00 pm to 8:00 am, as a previous study in the Canadian population shows more than 50% started sleeping around 10:00 pm and wake up between 7:00 am to 8:00 am on a typical day [274]. If the participant is at home, then depending on the time window and the total number of sensors activated within a period of five minutes, the label can be defined as true sleep, disturbed sleep, sedentary or physically active behaviour as described in Table 2.

Table 2. Definition of each indicator for the study.

Indicators	**Time Window**	**Total Sensors Active**	**Other**
True sleep	True	0	
Disturbed sleep	True	> 0	
Physically active	False	>= 3	
Sedentary behaviour	False	<= 2	
Away	False	0	Consecutively for at least 24 hours
Home	False	NA	If not away

Note: Time window for sleep chosen as 10 pm to 8 am, NA- not applicable

After labelling all data for each timestamp for the complete dataset, descriptive analysis has been done at the national level, followed by stratified analysis at the provincial level.

The metadata has geolocation information, naming the city and has the timestamp for each event that occurred. I have generated several new variables, such as time of the day and type of day, which can help examine human behaviour at a more detailed level. As the movement pattern and average duration of in-house physical activity, sleep, and sedentary behaviour may be associated with location, time of day, and type of day, the dataset provided by ecobee was augmented to determine the significance of these factors. The following additional variables defined according to the algorithms described above were added:

- Home/away
- True sleep/disturbed sleep
- Active/sedentary
- City
- Province
- Country
- Period of day (morning/afternoon/night/evening)
- Weekdays or weekends.

Statistical analysis was performed on the data included mean, standard deviation, and standard error for the "*amount of time*" for away, true sleep, disturbed sleep, active and sedentary behaviour at the national level. The results were also stratified into households with a single individual, households with multiple individuals, and then according to province.

This study has attempted to replicate PASS [127] indicators, improve the data collection process, and develop new indicators for all three categories: sleep, physical activity, and sedentary behaviour. Our proposed method will measure the number of hours of night-time sleep using real-time quantitative data. In contrast, the existing systems use self-reported measures, where participants respond to a survey with the response rounded to the nearest half-hour [275]. When analyzing data from our pilot study, looking for correlations between the data collected through

Fitbits and the number of activated sensors in the house, we have applied the Spearman's correlation test.

5.4 Results

5.4.1 Pilot Study

The results of the Spearman correlation coefficient between the total number of sensors activated in the ecobee thermostat and the number of steps tracked by the fitness tracker (Fitbit) was r = 0.78, N = 3292, *P* <.001. This statistically significant test indicates a strong positive correlation between ecobee's sensor activation, and the number of steps taken by the participants, providing evidence of the potential for using this data as PHAC indicators of physical activity. Figure 10 visualizes the correlation between data from Fitbit and ecobee at the individual level.

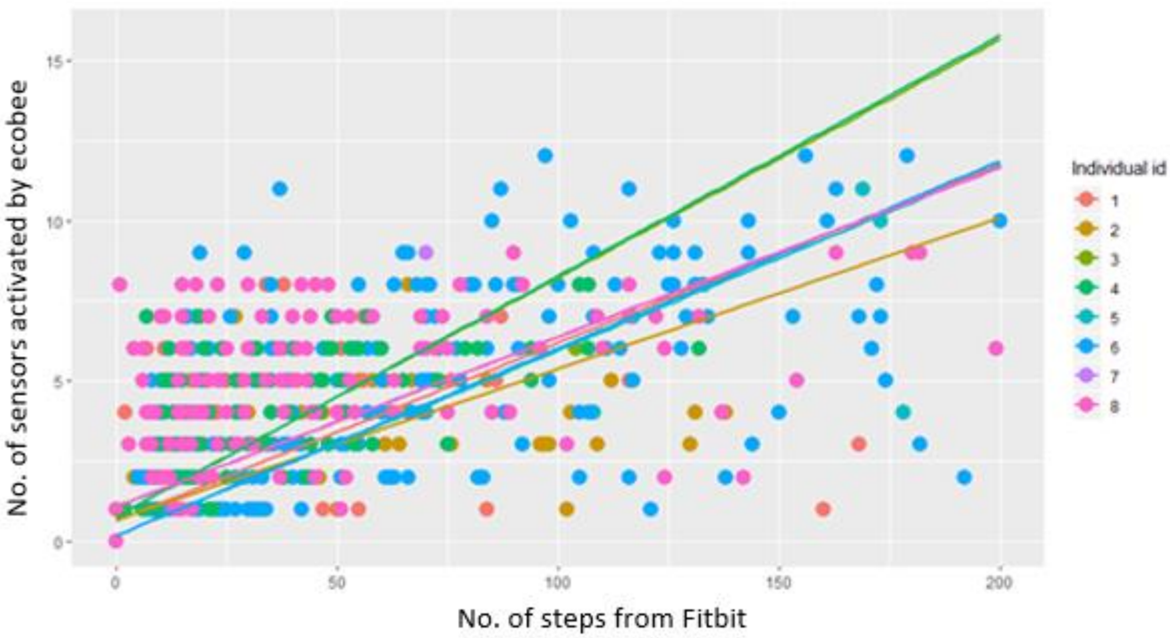


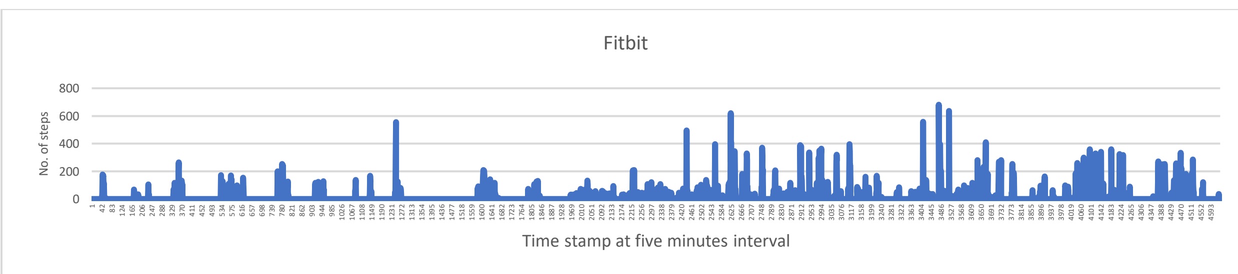
Figure 10. Scatterplot between number of steps from Fitbit with the number of sensors activated in ecobee.

Table 3 describes the Spearman's correlation coefficients [276], stratified by participants (eight participants) involved in the pilot study. The coefficients range from 0.79 to 0.91 and were all statistically significant.

Table 3. Individual-level correlation coefficients between the number of steps from Fitbit with the number of sensors activated in ecobee.

Participant id	Spearman correlation (rho)
1	0.79
2	0.91
3	0.80
4	0.81
5	0.84
6	0.84
7	0.79
8	0.79

Figure 11 represents the association between steps measured through Fitbit and the number of sensors activated on the ecobee thermostat at the same time for the same individual. Whenever the steps were taken were zero, the number of sensors activated was either zero or one, which signifies that the person was sleeping or busy with some sedentary activity.



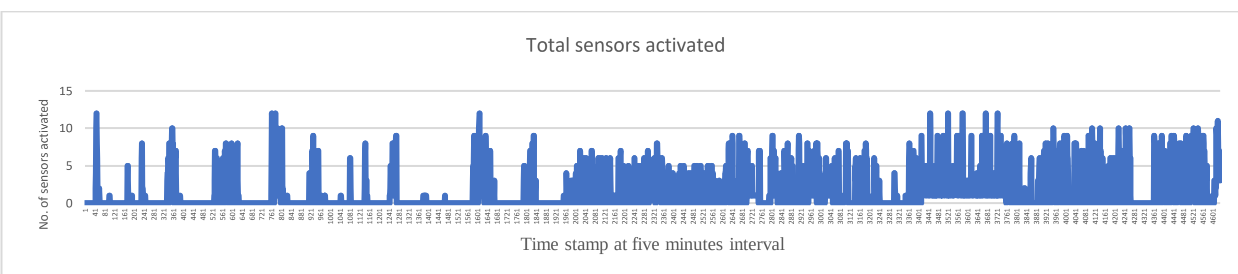


Figure 11. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by all participants.

In contrast, when the number of sensors activated was more than zero, more steps were taken, representing a certain amount of physical activity. This pattern is consistent throughout the data as shown in the Figures 12-19.

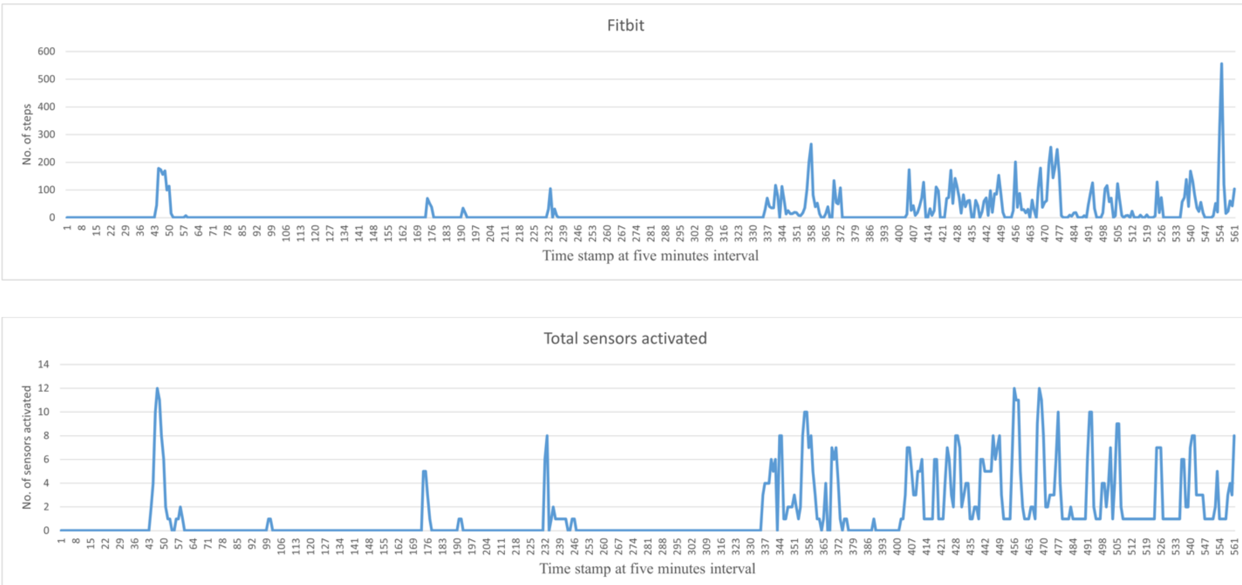


Figure 12. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 1.

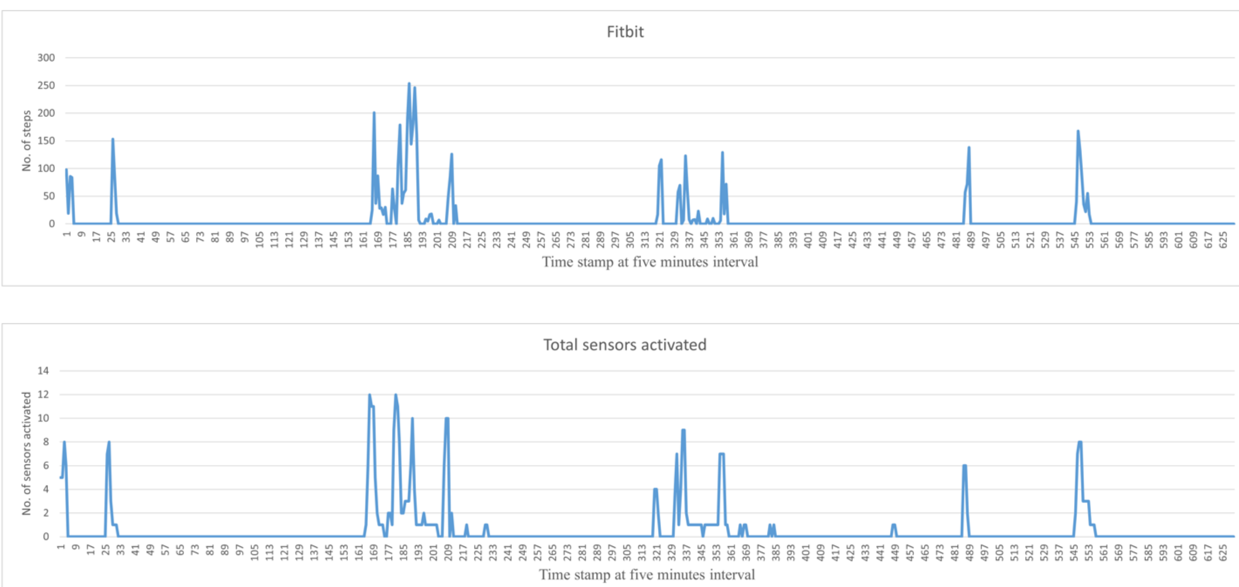


Figure 13. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 2.

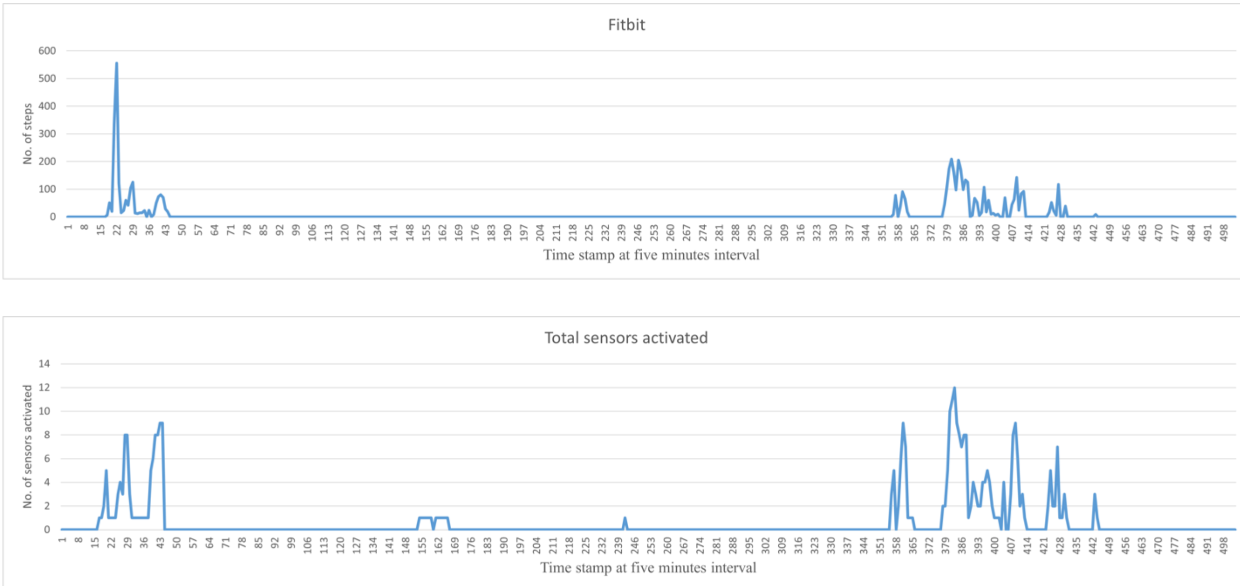


Figure 14. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 3.

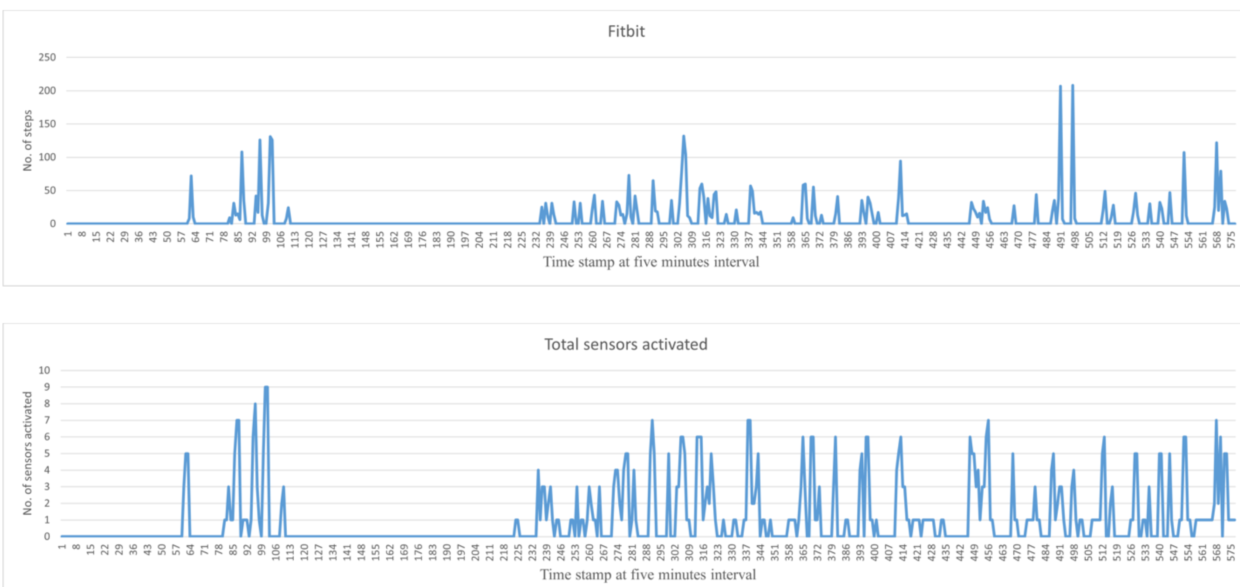


Figure 15. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 4.

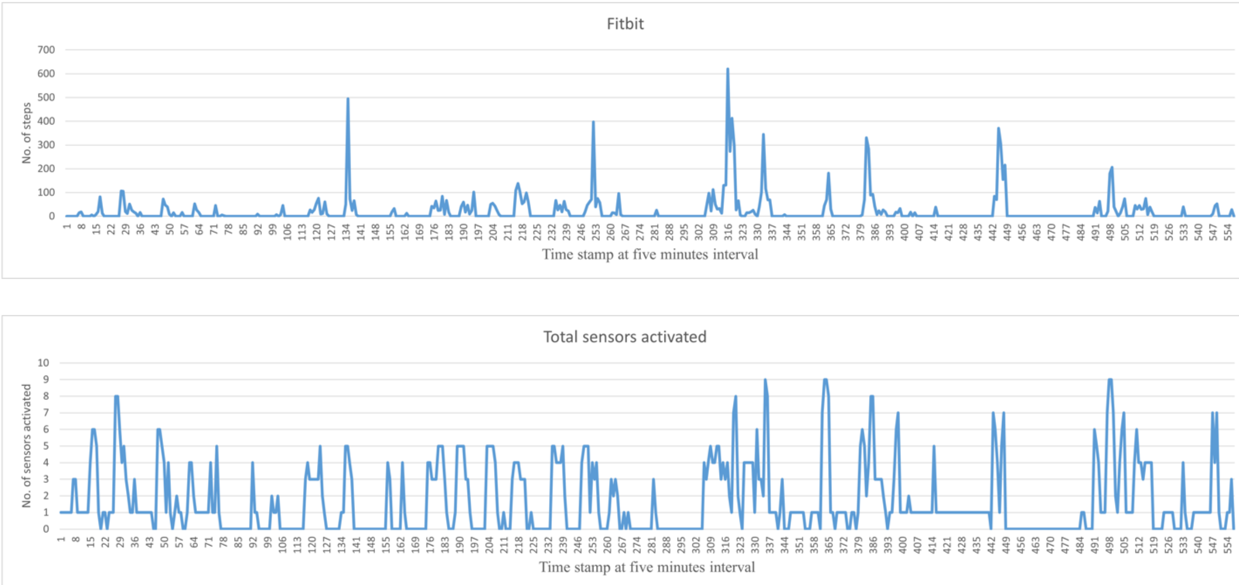


Figure 16. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 5.

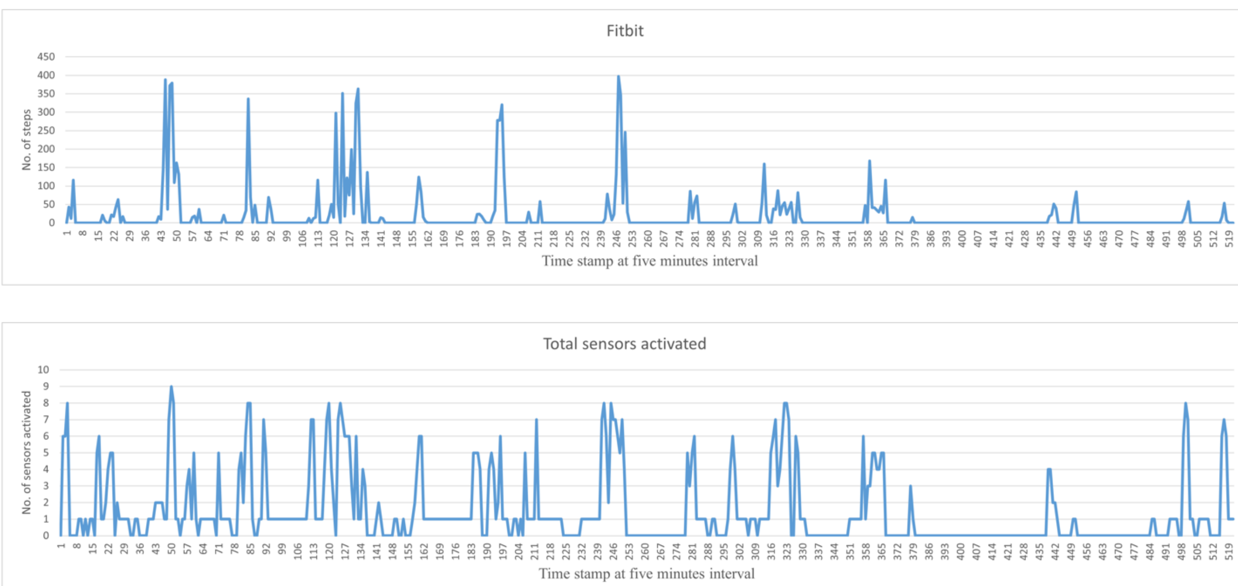


Figure 17. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 6.

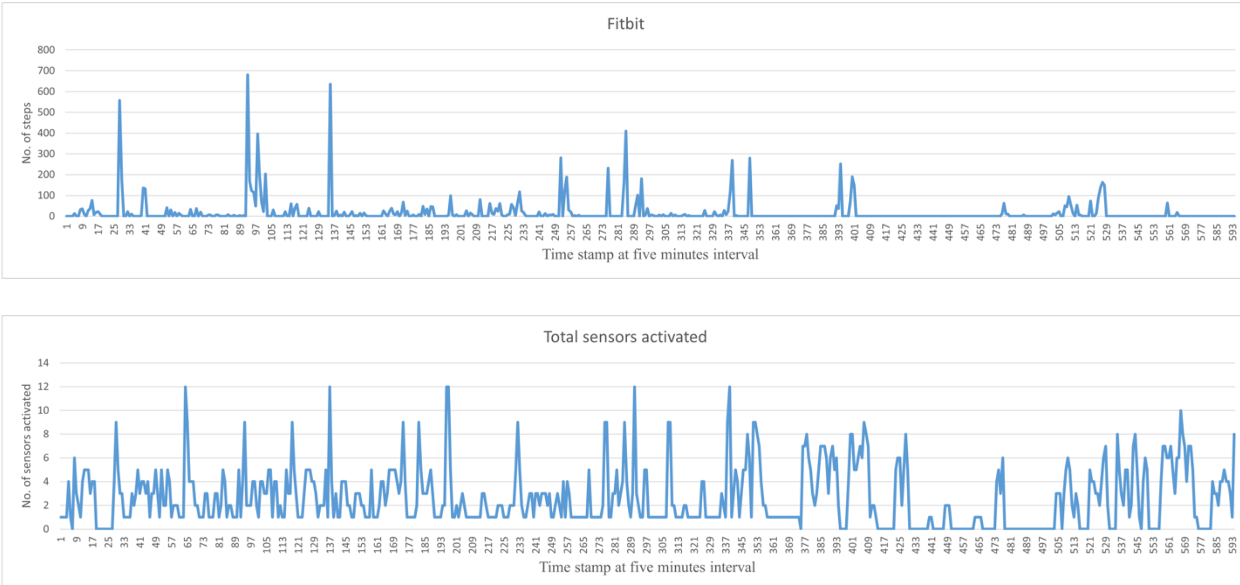


Figure 18. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 7.

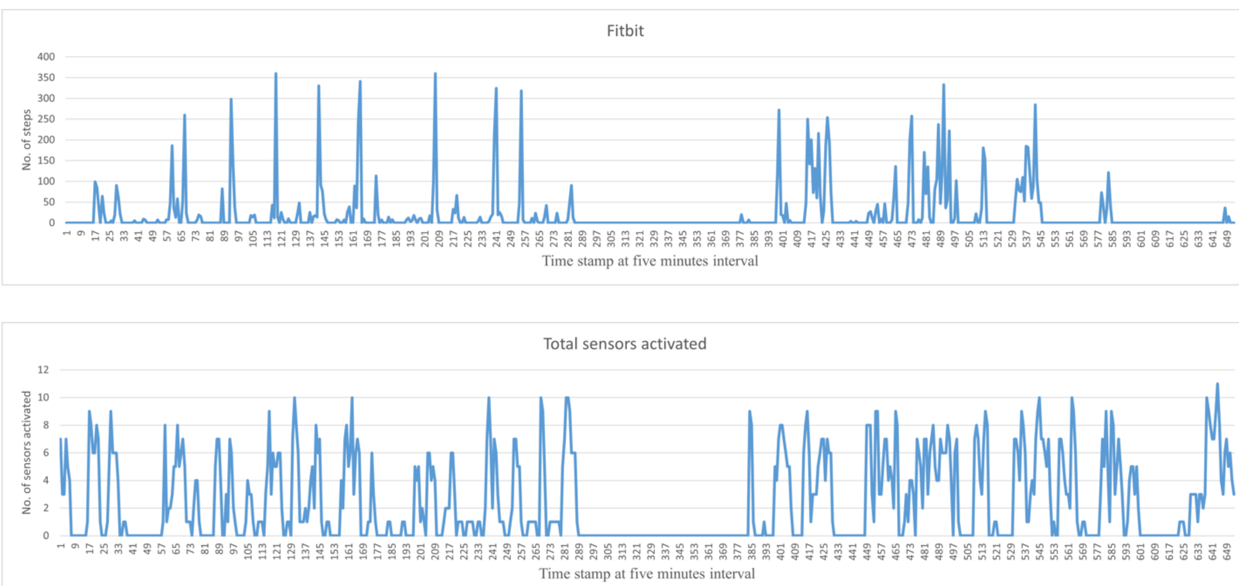


Figure 19. Association between steps data from a. Fitbit (upper) and b. sensor data from ecobee (lower), by participant 8.

Figure 20 combines data from Fitbit and ecobee to visualize the duration of sleep in a household with two individuals. Fitbit collects data at the individual level, whereas ecobee data is at the household level. Sleep labels from ecobee and Fitbit are highly similar, as seen by the colour-coded lines; Subjects 1 and 2 undertook similar activities around the same period.

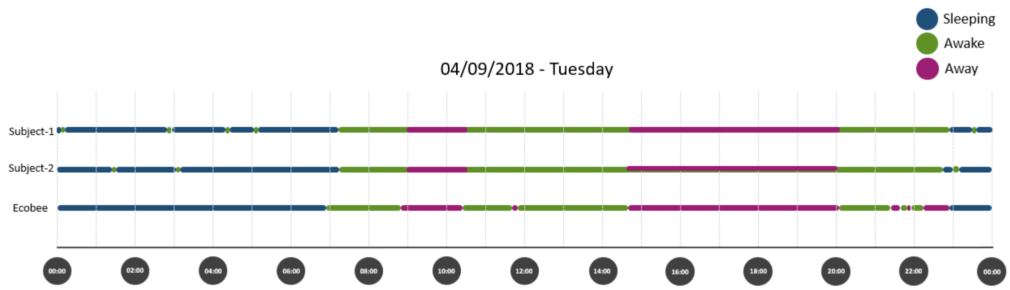


Figure 20. Visualizing sleep from Fitbit and ecobee data.

5.4.2 Population-Level Analysis

Results from the national level descriptive analysis of physical activity, sleep, and sedentary behaviour is described in Table 4. The individual level's average sleep duration was 7.89±0.17 hours, and the household average was 7.71±0.18 hours. Similarly, the average duration of in-house physical activity for single individuals in Canada was found to be around 85±13 minutes per day.

In this study, the in-house physical activity indicator is a proxy. It is based on movement detected by ecobee's sensors in five-minute intervals. The current physical activity indicator from PASS is a self-reported, conscious measure for exercise reported for intervals of at least 10 minutes, rounded to the nearest half-hour [277].

The sedentary behaviour indicator designates a quantitative measure derived from ecobee's data. It is based on participants being home during the day with less than two sensors activated in a 5-

minute interval. The number of sensors represents the amount of activity in the house. When participants move around the house, they trigger different sensors in different rooms. According to our pilot study, individuals triggering a low number of sensors correspond to individuals with lower in-house physical activity. The equivalent PASS indicator is self-reported, consisting of hours rounded to the nearest half-hour. For example, our algorithm treats cleaning the house and moving between rooms as active minutes and only handles periods where the resident is static at home (e.g., in front of the TV or working in the office) as sedentary, which provides a more granular monitoring mechanism for sedentary behaviour.

Table 4. List of indicators and their values for sleep, physical activity, and sedentary behaviour in Canada.

Indicators	**UbiLab[$]**		**PASS[#]**	**Comments**
	Individual N=70	Household N=888	Individual	
Sleep				
Night-time amount of sleep (hours)	7.89±0.17	7.71±0.18	7.2	
Disturbed sleep (hours)	2.10±0.17	2.28±0.18	Non-existing	New indicator that is not currently available as part of the PASS indicators.
Physical activity				
Physical activity in the home (mins per day)	85.2±13	146.4±20	24.1	
Sedentary behaviour				
Sedentary time amount (hours)	4.44±47	5.75±34	9.6	Distinct interpretation of sedentary behaviour
Out of home				
Away period (hours)	8.12±0.57	5.80±41	Non-existing	New indicator that is not currently available as part of the PASS indicators.

[$]Ubiquitous Health technology lab at the University of Waterloo [#] Physical Activity, Sedentary Behaviour and Sleep, Values in the table are Average ± SD, i.e., standard deviation.

Additionally, the developed algorithms can measure the amount of disturbed sleep during the night. A similar indicator of nocturnal environmental noise is currently in development by PHAC and could help to explain interrupted sleep [277]. However, sleep interruptions due to medical conditions (sleep apnea), sleepwalking, or bathroom trips will not be captured in PASS but are captured by ecobee's sensors and our algorithms.

In addition to national-level health indicators, time trend analysis found that the average duration of an individual's sleep has declined within the last three years while the average duration of sedentary behaviour and in-house physical activity increases. Figure 21 demonstrates time trend analysis of sleep, physical activity, and sedentary Behaviour for Canadian households from 2015 to 2017.

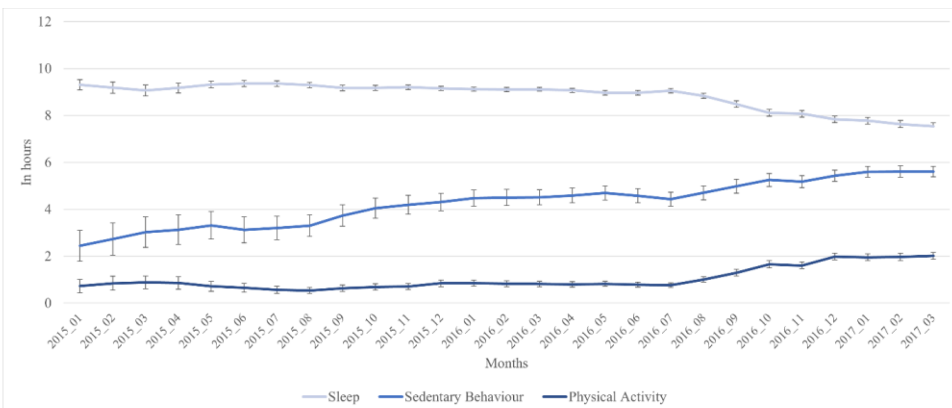


Figure 21. Time trend analysis for sleep, physical activity, and sedentary behaviour for Canadian households.

When exploring the province-level stratification, the result showed significant variation between provinces. Figure 22 visualizes the provincial breakdown of average sleep duration in Canada from 2015 to 2017.

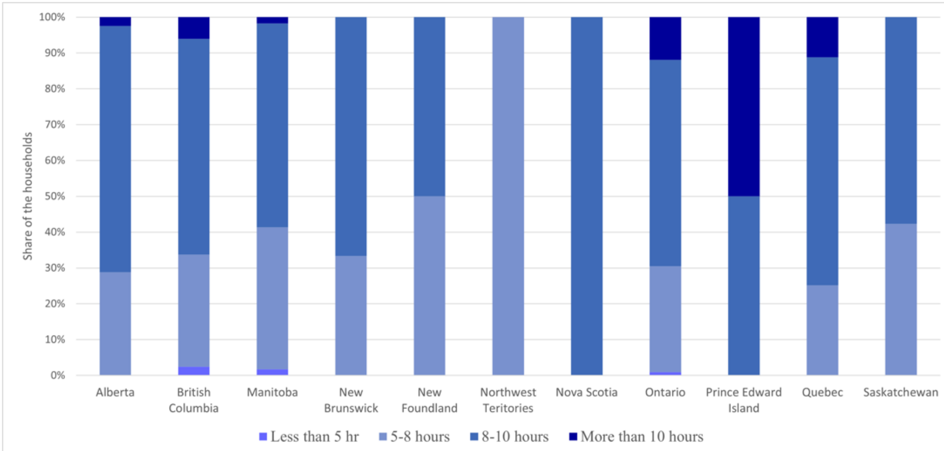


Figure 22. Distribution of average duration of sleep stratified by provinces and territories of Canada.

5.5 Discussion

With the rapidly increasing integration of technological advancements in diverse domains such as retail and commercial sectors, its unification with the medical sector can pave new horizons to improve the quality of existing healthcare systems [278]. For instance, a new field known as mobile health (mHealth) gained increasing attention with the introduction of smartphones and various mobile apps. The use of mHealth can enable users to track their vital health indicators like heart rate, respiratory rate, blood pressure, etc., and can prevent them from getting predisposed to certain chronic diseases [279]. These smart devices also uplift the current medical care delivery by enabling the patients to connect with physicians remotely and improve their quality of life [280].

Smart home technologies are currently expanding into the realm of AI-based virtual assistants, such as Amazon Echo [281], Apple's Siri [282], and Google Home/Nest [283]. In the past, it was not possible to think of technologies that could monitor indoor temperatures using smart devices. However, with the advancements in smart home technologies, this has become a reality. As an added benefit, these data collection systems are based on zero effort technology (ZET) [100], and the data gets collected unobtrusively in real-time. These devices have become more compact and affordable with rapidly evolving technology, providing seamless data collection for health monitoring.

In this context, this study has aimed to leverage existing smart home technology-based data sources to generate population-level health care indicators. Although raw data are difficult to interpret, artificial intelligence (AI) algorithms can be employed to draw meaningful insights from large volumes of data. This analyzed information can be disseminated to end-users using the interactive web-based dashboard, displaying the user's health data in a textual, auditory, and visual manner. These interactive platforms are accessible and cater to the need of users with diverse health backgrounds. Moreover, users can learn about their day-to-day activity patterns and general health behaviours, as well as compare their behaviours with those of their peers or the population average. As an added benefit, these systems provide insights about their health status and recommendations that will improve their health indicators, enhancing users' experience.

In 2017, an estimated 100,000 Canadian households had ecobee thermostats installed in their residences, with even more users in the United States of America, and this number is only expected to increase. With incentives from the provincial government, such as the Ontario Green Fund [284], smart home technology witnessed an increasing interest. Such initiatives would help diversify the sample population and broaden the socio-demographics of the study population while simultaneously reducing the issue of sample bias. This would, in turn, benefit the policymakers to strategize their future plans in the health care sector.

This study is a proof-of-concept demonstrating the feasibility of using alternative data sources for developing population-level health indicators. Time series data from a large population collected without additional effort and from non-health-related sources have been used to develop alternative PASS indicators. From this study, the Spearman Correlation Coefficient between Fitbit and ecobee data sources proves that the association was statistically significant and positively correlated. An increase in the number of steps resulted in a larger number of sensors being activated, which was consistent throughout the different houses included in the study. As previous studies have demonstrated, Fitbit is a reliable source for measuring physical activity [97,269,285]. This study has provided evidence that the same can be observed in ecobee data. The DYD dataset can be a potential source of information for measuring population-level health indicators. Further analysis of the DYD dataset shows that the average sleep duration at the individual level was 7.89 hours and 7.72 hours at the household level, which is very similar to the results of PHAC's published values for Canadians [277]. The real-time monitoring of population health behaviours that do not interfere with the day-to-day activities of individuals that own smart thermostats, which proves to be an important aspect of this study.

5.5.1 Limitations

This study has several limitations that constrain the capabilities of the data collected and possible health behaviour insights. While data granularity is high (data sampled every five minutes) when compared to traditional public health surveillance data, a higher level of granularity would be needed for a system focused on monitoring individual health behaviours. Additionally, while the sensors can track human movements, differentiating between individuals in a multi-occupant household still remains a challenge. In future studies, this limitation could be addressed with the use of additional on-body sensors to distinguish between the occupants of a household. This study also suffers from sampling biases, where the users of these smart home technologies belong to middle-to-higher socio-economic status, as well as not including smaller or remote communities within Canada.

5.6 Conclusion

There is a lot of potential for RPM to expand and leverage commercial technologies. This study is just one example where technology can be used to bring innovative solutions for real use in the realm of health care, especially as it allows the use of technologies that are zero effort and have more than one added benefit. Technologies such as these will be able to advance the fields of RPM and public health surveillance.

5.7 Contributions to the PhD Thesis

This work was published in 2020 and that this served as the foundation for the rest of my thesis. The result of this study motivated me for writing a perspective paper to provide theoretical foundation to my work. The chapter 6 describes the importance of the NextGen data source for public health surveillance. Also, the result this study leads to two broader use cases utilizing a similar data source as standalone and in integration with Google mobility data for micromobility assessment at the residential area. Those two use cases are going to be described in chapter 7 and 8 respectively.

Chapter 6 NextGen Public Health Surveillance and the Internet of Things (IoT)

6.1 Preamble

In this chapter, I present a published manuscript presenting a viewpoint related to the use of IoT technologies for public health surveillance. The manuscript was published in the journal of Frontiers in Public Health, in December 2021 [28]. The manuscript provides an overview of use of alternative modern data sources for public health surveillance with special emphasis on IoT based big data. The manuscript will provide a framework for the development of the rest of the thesis.

Citation: Sahu KS, Majowicz SE, Dubin JA, Morita PP. NextGen Public Health Surveillance and the Internet of Things (IoT). Frontiers in Public Health. 2021 Dec 3:1976.

6.2 Abstract

Recent advances in technology have led to the rise of new-age data sources (e.g., Internet of Things (IoT), wearables, social media, and mobile health). IoT is becoming ubiquitous, and data generation is accelerating globally. Other health research domains have used IoT as a data source, but its potential has not been thoroughly explored and utilized systematically in public health surveillance. This chapter summarizes the existing literature on the use of IoT as a data source for surveillance. It presents the shortcomings of current data sources and how NextGen data sources, including the large-scale applications of IoT, can meet the needs of surveillance. The opportunities and challenges of using these modern data sources in public health surveillance are also explored. These IoT data ecosystems are being generated with minimal effort by the device users and benefit from high granularity, objectivity, and validity. Advances in computing are now bringing IoT-based surveillance into the realm of possibility. The potential advantages of IoT data include high-frequency, high volume, zero effort data collection methods,

with a potential to have syndromic surveillance. In contrast, the critical challenges to mainstream this data source within surveillance systems are the huge volume and variety of data, fusing data from multiple devices to produce a unified result, and the lack of multidisciplinary professionals to understand the domain and analyze the domain data accordingly.

6.3 Introduction

The function of public health systems is to understand and respond to health trends affecting populations [10]. This is achieved through public health surveillance, that is, the ongoing collection and analysis of population health indicators. Traditional surveillance data collection can be cumbersome, expensive, and slow, often relying on paper-based and digitally extracted data sources. Social media and crowdsourcing are data sources that can be leveraged for surveillance data [286,287]. Sources like Twitter, Facebook, Google, and Reddit have been successfully used to explore behaviour and health outcomes [80]. These are now being accepted as potential data sources across several health domains [288].

Another promising data source is the increasing number of devices (e.g., smart home monitors, wearables) and the technology to interconnect them. Internet of Things (IoT) technologies have become mainstream within communities and individual households [289]. Wearables and sensors can track personalized parameters of healthy living, including sleep, physical activity, and sedentary behaviour [290]. These devices can provide insights into population health, disease management, and active assisted living services [291,292]. IoT data has several advantages over traditional surveillance data: high volume and frequency of data collection, data triangulation, real-time availability, and minimal acquisition effort.

Existing literature discusses the potential use of the IoT data sources for different purposes within multiple domains including healthcare. Among healthcare domain, area specific application can be seen for pediatric, geriatrics, chronic disease supervision, private health and fitness management [293,294], but no single study exists to put together the views to utilize the IoT data with specific emphasis on public health surveillance. This chapter summarizes the existing literature on the use of IoT as a data source for surveillance. We discuss the shortcomings of current data sources and how IoT can meet the needs of surveillance. Challenges facing the large-scale application of IoT data to surveillance are also explored.

6.3.1 Public Health Surveillance and Challenges with Existing Data Sources

Public health recommendations focus on the social determinants of health and health equity [295]. Surveillance is the process by which ongoing health data are collected, analyzed, and reported, and it is critical to informing public health services. In 1968, the World Health Organization listed ten essential data sources for surveillance [77] (Figure 23: Traditional data sources) that at the time relied on paper-based data collection and manual data entry. Surveillance capability has evolved enormously alongside advances in technology. It now includes digital data extracted from several sources (Figure 23: Modern data sources), offering reduced processing time, fewer errors, and reduced lag between data collection and its use.

The above said, surveillance data are still often obtained from questionnaire-based surveys (online surveys, in-person or telephone-based interviews [87], and such data collection requires enormous resources and funding [85,86]. Data quality can be compromised by declining response rates [87], recall bias [296], and low granularity of the data [200] as in the traditional data collection system, there is a limited number of subjects provide their inputs. Without complete and comprehensive information, the value of the data reduced. For example, fewer subjects with a smaller *n*, really only impacts the precision of the estimates that come from surveillance. To further explain, the system might not get very precise incidence estimates, which may or may not be a problem depending on the goal of the system. The bigger issue with declining response rates is that they usually do not happen at random, meaning there is a less representative set of results. This is an issue if the factors that lead to making it into surveillance also relate to the issue you are trying to measure with the surveillance system. Current data used for the surveillance have challenges like missing data, under-reporting, inconsistencies, invalid data, illegible handwriting, non-standardization of vocabulary, measurement error, and inappropriate fields [297]. Traditional data sources used in surveillance are often delayed. For example, at least one year is required for getting a Canadian Community Health Survey (CCHS) update. "Public Health Ontario" in Canada affirms interdependent gaps within surveillance, insufficient data to build comprehensive health indicators [88], and an absence of existing mechanisms to capture some of healthcare's vital components.

Current surveillance relies on both prospective and retrospective data collection, analysis, and reporting [9]. The current pandemic has highlighted the essential need for real-time public health surveillance to improve the evidence-based decision-making process [298]. Our evolving knowledge about chronic diseases, their risk factors, and management also demands the modernization of surveillance [9]. Real-time responses to emerging public health threats require real-time and systematic data collection.

6.4 Next-Generation Data Sources for Public Health Surveillance

Researchers have attempted to build and analyze health indicators using innovative data sources [94,299,300]. They are exploring the use of smartphones [301], online searches [302], social media [288], wearables [303], ambient sensors [304], electronic health records (EHRs) [94,305], medical-administrative records [94], and pharmacy sales [299] to broaden the scope of surveillance.

As a source of surveillance data, information technologies are potentially advantageous because their near-universal uptake by a significant portion of the population creates vast quantities and varieties of data [200]. For example, wearable data from six billion nights has been used to understand sleep duration, quality, and change in pattern with time [18,306]. Effective use of big data for surveillance requires innovative analytical methods such as data integration [303] and data visualization [299,307,308]. Big data analytics is becoming mainstream in public health, integrating knowledge and skills from health informatics and biostatistics [309].

6.5 The Internet of Things as a Novel Data Source

The Internet of Things (IoT) is a technological innovation through which devices can communicate with each other in real-time through an internet connection [310]. For example, several household devices are interconnected to achieve a common objective, such as monitoring temperature or motion [310]. Integrated devices can include different sensors, mobile phones, mobile applications, wearable devices, and Radio-Frequency Identification (RFID) tags [310].

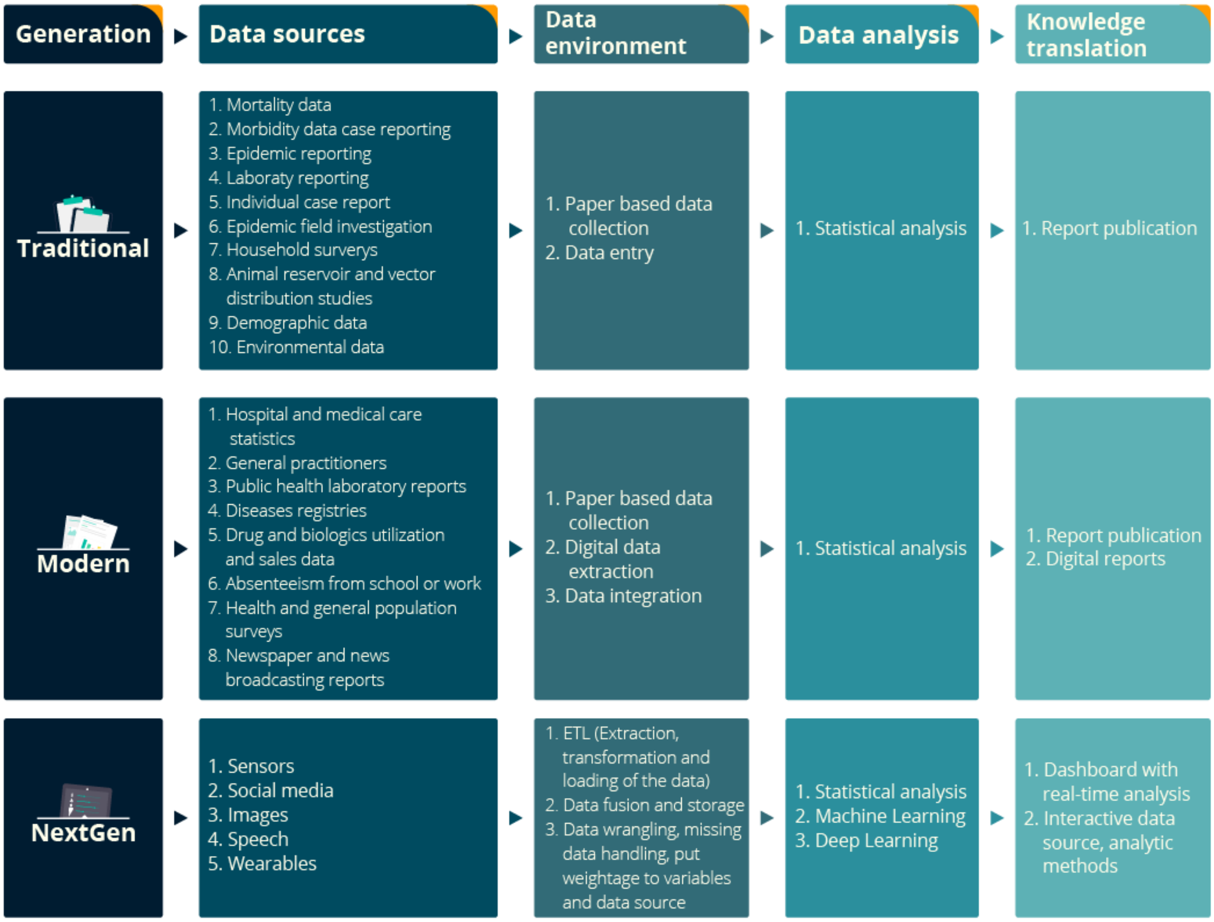


Figure 23. Conceptual framework of NextGen public health surveillance with traditional, modern, and NextGen data sources. Traditional and modern data sources extracted from Declich S, Carter AO 1994 [77].

IoT devices have accelerated data collection [45,292]. Connectivity among people, machines, and organizations increases as device availability and affordability improve [200]. This increase in connectivity is because of the ease of use of the devices, user-friendly designs, and internet speed. These parameters reduced the time gap within communication, broaden the scope of communication by providing different choice, be it audio visual, text, or hybrid of multiple methods. People can interact with the machines and vice versa, which was not possible earlier due to lack of technological progress. In 2011, the number of interconnected devices overtook the actual number of people globally [150]. The potential for data generation is exponential [45]. As the IoT data has already been successfully used in multiple setups to monitor individual health outcomes and report on environmental conditions, some of the best use cases has been described below.

6.5.1 Use of IoT Data to Support Individual Health Outcomes

The management of chronic conditions has traditionally relied on patients interacting with their healthcare providers in person. However, patients spend most of their time outside the clinic. IoT monitoring provides an opportunity to collect real-time health information between patient-healthcare provider interactions.

Smart devices, such as wristbands, with IoT technology have been developed to measure individual physiological data, including physical activity [290,311], sedentary time [25], oxygen saturation [312–314], heart rhythm [312,313], muscle tremors [315], spinal posture [316], brainwaves [317], sleep [153], diet [318,319], electrodermal activity monitoring for sympathetic response [25] and oral health care [320]. With regards to specialized medical care, IoT technology has been used to cater to the need of cardiovascular [87], cardiopulmonary [87] and ophthalmology [321]. With regards to different categories of populations, IoT has been used to help to monitor indicators related to women's health [322], including pregnancy [323], soldiers at the country borders [154], nursing care at the hospitals [324], the elderly population in the long-term-care homes [325], persons with neurological conditions at the rehabilitation center [316], and also for persons with respiratory complaints including asthma [163].

IoT devices have a multipurpose use within the healthcare field, such as their capabilities can range from providing prenatal care to rehabilitation to monitoring seniors or athletes. IoT devices have successfully provided real-time health information on maternal and fetal health between regular appointments [326]. By monitoring vital signs using sensors, IoT platforms have been designed to provide people with diabetes with feedback and notifications to mitigate the risk of complications [152,327,328]. Additionally, wearable devices have been used to detect falls and changes in behavioural activity for seniors living independently [329–332]. Monitoring systems have also been developed to evaluate sports rehabilitation [333–336]. IoT can support individual outcomes by allowing patients to manage their health outside of the clinical setting.

6.5.2 Use of IoT Data to Monitor Environmental Conditions

The IoT can also monitor environmental conditions in areas where we live, work, and play. Monitoring air purification in hospital settings plays a role in mitigating hospital-related infections [337]. Monitoring air quality is already used to quantify climate change impact [338] and has the potential to help mitigate its impact in the future [339]. IoT has been employed to monitor hospital circulating air volume, ozone concentration, temperature, humidity, and leaked ultraviolet intensity [337]. Preventive behaviour like hand washing can also be monitored [340]. Indicators of healthy out-of-the-house environments, such as water pollution and air quality, have been another target of IoT health research [163,341].

6.6 The Internet of Things in Public Health Surveillance

IoT data has been successfully used in other health domains but has not yet been fully used in public health. In response to the pandemic, the 2020 Riyadh Declaration made several recommendations to address the shortcomings in global public health response systems [342]. The Declaration prioritized the need for scalable and sustainable digital health technologies and the adoption of health intelligence [342]. There is a growing interest in using IoT data for building public health indicators at various levels [343–345].

6.6.1 Advantages of IoT in Public Health Surveillance

IoT data have the potential to overcome shortcomings of current surveillance. IoT data sources provide high-frequency data with greater usability, and much of the device infrastructure for surveillance is already in place (i.e., smartphones, wearable technologies, internet access). Currently, worldwide more than three billion smartphone users [346], 722 million users of several kinds of wearable devices [347], and more than 1.2 billion smart-home connected devices exist [348]. IoT data benefits from essential features like high granularity [200], objectivity [303], and validity [44]. These “user-generated data ecosystems” are being generated with minimal effort by the device users and researchers. To date, the monetary cost to participants and researchers is low, suggesting that public health monitoring costs would likewise be minimal [349,350]. Finally, IoT enables near real-time data collection [103]. This can significantly reduce the time gap between health events, data collection, reporting, and intervention.

Here we have assessed IoT's current attributes using the framework for evaluating public health surveillance by Groseclose and colleagues [351], which outlines nine features of surveillance systems to consider (Table 5). As summarized in the table, the major advantages of IoT data sources appear to be high-frequency data collection, the potential to have syndromic surveillance, zero effort data collection method, high volume, and variety of data. The major disadvantages appear to be lack of representativeness within a single data source, private players’ involvement as the data owner, the need for a high technological system to store, clean, and analyze the data, and interoperability. In addition to the above points, data privacy concerns of users are a potential disadvantage of acceptance of this technology from the user point of view [344].

Table 5. Analysis of IoT as a data source for public health surveillance, using Groseclose *et al.*'s 2010# framework for evaluating public health surveillance.

Attributes (Definition)	Features of IoT Data
Simplicity ***"The system's structure and ease of operation. The system should be as simple as possible."*** [351]	• The manufacturer often provides data collection/extraction from users without complex interactions using the Application Programming Interface (APIs). • Easy access to the data, which is often collected by passive sensors, minimizes the burden for the user. • IoT systems rapidly generate large volumes of data in real-time, creating challenges associated with managing, hosting, and analyzing big data. • Diverse types of data are generated: numeric, images, text, or audio. • Collects vast amounts of data from the same individual, often supporting longitudinal analysis.
Flexibility ***"Ability to adapt to changing information needs or technological operating conditions with little additional time, personnel, or allocated funds."*** [351]	• Application Programming Interfaces (APIs) make it easy to adapt to the technology to the end-users being used, type of data, type of database, storage, and security requirements. • New IoT data sources that use APIs can easily be integrated into systems, affording changes in a data structure as technologies evolve. • Changes in case definition can be updated in algorithms rather than requiring changes to data collected since systems can access the raw data. • The system can be automated to generate alert systems without manual effort, which can help public health officials identify potential signals for future outbreaks early.
Data quality ***"Completeness and validity of the data recorded in the system."*** [351]	• IoT data often suffers from missing, inaccurate, and incomplete data. • Wearable sensors that require participants to recharge and remember to interact with the device often have larger volumes of missing data. • Ambient sensors often generate continuous and complete datasets as they are always connected, powered on, and streaming. • Technology development is leading to improved data quality across all IoT sensors.
Acceptability ***"Willingness of persons and organizations to participate in the system."*** [351]	• IoT technologies are pervasive, and in the community, a part of the population is already using those technologies to generate data. • IoT adoption has been accelerating in the last decade and is predicted to be much higher in the near future. • Recent advancements in technology used "skin interfaced sensors" not only to monitor physical activities and vital signs but also keep track of molecular biomarkers of the human body [352] • Users need to agree to share their data, as it has already been collected.
Sensitivity ***"At the level of case reporting: the proportion of cases of a disease or event detected by the system. Ability to detect outbreaks over time and evaluation of surveillance system."*** [351]	• IoT sensors, in most cases, do not focus on the detection of specific diseases such as COVID-19 or influenza but rather on symptoms like fever, abnormal heart rate, or change in gait pattern. • IoT technology is ideal for supporting syndromic surveillance by collecting data about healthy behaviours and health variables in real-time. • IoT technology will collect data often indirectly associated with health and health risk behaviours (e.g., in-house motion data to quantify sleep patterns, phone mobility data used to quantify response to COVID-10 policies). • IoT will provide extensive participant data with a higher likelihood of the presence of events. • The longitudinal nature of the data can detect future anomalies using Artificial Intelligence models within the healthcare sector and send alerts to

	policymakers. The longitudinal and continuity nature of the data will provide richer insights into population behaviours, which increases the likelihood of getting the events of interest.
Positive predictive value ***"The proportion of reported cases that actually have the event under surveillance."*** [351]	• The proportion of the presence of IoT within the community is increasing and predicting the true positive cases will be easier using IoT data by identifying early alerts. • Detecting specific diseases is possible, as technologies such as lab on a chip [353,354] allow for the real-time detection of pathogens and contaminants. • Positive predictive value seems to be in a disadvantageous position with the current IoT data environment, but this might change in the future.
Representativeness ***"Ability to accurately describe the occurrence of a health-related event over time and distribution of the population by place and person."*** [351]	• A large number of participants can provide access to data that were not represented in the traditional data collection method. • IoT technologies are ubiquitous, highly pervasive, and are generating data 24/7. • Data mining from sensors already owned by the population generates a biased sample, with data from the wealthier and more physically active part of the population. • Studies can supplement biased samples by deploying targeted studies to collect data from under-represented subgroups of the population.
Timeliness ***"Reflects the speed between steps in a system."*** [351]	• Data is often collected at high frequencies, often affording access to data in the near real-time. • An increase in the data's granularity and the longitudinal nature of the data can provide richer insights, for instance, faster alerts of anomalies for specific health issues and support the creation of innovative indicators. • In the near future, the IoT data source may become helpful to identify future pandemic and climate-related emergencies. Immediate assessment of the impact of policy changes (for example, "work-from-home" during the pandemic) can be possible using IoT data. • Improvement from traditional data sources where data collection often happens once yearly or less frequently.
Stability ***"Ability to rely on the system for availability and to collect, manage and provide data without failure. Ability to be operational when needed."*** [351]	• Private cloud systems can provide the necessary data security and maintain the users' privacy. • Redundant, always available, more stable public health surveillance platforms/systems can be built using private cloud solutions, having the capacity to collect uninterrupted data without failure. IoT manufacturers and IoT data custodians can deliver such redundant and stable systems for their consumers' everyday use. • The disadvantage of these IoT data manufacturers is ever-changing company environment (for example, corporate and big private entities) might not provide a stable source of data. The alternative source of data should be listed as a backup plan to support and strengthen when required.

[#]*Groseclose SL, German RR, Nsubuga P. "Evaluating Public Health Surveillance," in Principles & Practice of Public Health Surveillance (Oxford University Press). doi:10.1093/acprof:oso/9780195372922.003.0008*

6.6.2 Challenges to using IoT in Public Health Surveillance

The challenge now is how to access and analyze the data being gathered. Some IoT companies create sharable, research-oriented data sources, such as "donate your data" from ecobee, a smart thermostat company in Canada [355]. ecobee's smart home products include motion and temperature sensors, and research teams have access to longitudinal data from thousands of households with a data granularity of five-minute intervals.

Other IoT companies publish studies from their own smart devices using artificial intelligence algorithms for population-level measurements. For example, Fitbit wearables recorded sleep data from over six billion nights of its customers' sleep [18], the most prominent sleep dataset ever collected. Similarly, Oura Health used IoT data gathered from their Oura ring, a wearable sensor that tracks key signals from the human body (sleep, heart rate, skin temperature, physical activity), delivering critical insights to help an individual harness their body's potential daily and also to monitor vital health indicators [356].

Another hurdle is the ability to fuse data from multiple devices to produce a unified result. Several research projects have focused on making IoT data fusion viable in the real world by designing computing infrastructure and data fusion techniques [103,357]. Real-time IoT analysis from multiple health monitoring devices may overwhelm current computational capabilities, such as using multiple devices to monitor each football player's physiological indicators during a game [358]. A distributed computational framework to handle complex computational needs was developed by Higinio *et al.* for health surveillance [358]. The use of each smart devices' computing capabilities effectively shared advanced health monitoring applications [358].

Regarding technical challenges related to IoT, some of the critical issues are energy optimization, hardware compatibility, security, and data connectivity [359]. A recent study by Iwendi *et al.* in 2020 shows that there are certain highly specialized algorithm such as a “hybrid meta-heuristic algorithm” has the potential to optimize the energy consumption of the sensors related to wireless sensor networks [359].

Aberration detection identifies unusual incidents or information trends with possible significance to clinical or public health [219]. Methods for detecting such aberrations have also evolved significantly. Current modelling methods can now analyze individual surveillance data collected from different sources and integrate multiple covariates [360]. The algorithms used for signal recognition have improved over the last decade and are now better equipped to utilize advanced informatics to capture surveillance data aberrations [360] accurately.

In 2018, Faverjon C. and Berezowski J. elaborated on IoT data's utility for aberration detection [360,361]. Two studies have shown that user data from wearables (Fitbit and the Oura ring) could detect early signs of COVID-19 infection [19,102,362–364]. Evidence shows the risk of hospitalization related to COVID-19 can be calculated from self-reported symptoms and predictive physiological signs by combining different health and behavioural data from consumer wearable devices; this may help identify pathological changes weeks before observation using traditional epidemiological monitoring [102,362]. As described in the study using Fitbit wearable, it has the potential to detect almost half of COVID-19 positive cases 24 hours before participants reported the onset of symptoms with 70 percent specificity [19]. Besides joint effort by multiple countries to develop vaccines and potential drugs to prevent and treat COVID-19, skin-integrated and skin interfaced sensors, positioned at optimal locations of the body, might address the ongoing and critical need for objective, continuous, and sensitive tools to detect COVID-19 symptoms early in the general population [352,363]. A research study highlighted a practical approach for managing epidemics using digital technologies with a roadmap to a rapid and universal diagnostic method for the population level detection of several respiratory infections in advance of symptoms [364]. These anomalies could predict future outbreaks [360] and prevent the spread of infectious diseases [365].

6.7 NextGen Public Health Surveillance

The COVID-19 pandemic has revealed a need to strengthen our public health surveillance and response systems. With the availability of public data and advances in collection and analysis, there is an opportunity to strengthen existing surveillance systems by harnessing complementary data sources like IoT-based data [302].

Figure 23 describes the NextGen surveillance systems' conceptual framework. The first layer describes the sources of public health data. The second layer represents the data architecture. Once the data integration process is completed, data manipulation and analysis can be possible using statistics, machine learning, and deep learning algorithms. This process will help discover new public health indicators and advance our understanding of existing disease risk factors.

6.8 Conclusion

Current public health surveillance systems have unique challenges in getting the relevant data at the right time and utilizing those data sources for policy-level decision-making. There is a considerable volume of non-traditional data being self-generated by the public through their ubiquitous use of smart devices. Public health has the potential to utilize the real-time, longitudinal data collected through the Internet of Things (IoT) necessary for health surveillance. Advances in computing are now bringing IoT-based surveillance into the realm of possibility. The advantages of IoT data include high-frequency, high volume, zero effort data collection method, with a potential to have syndromic surveillance.

6.9 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6.10 Funding

This work was supported by Natural Sciences and Engineering Research Council of Canada (RPGIN-2017-05310) and the Ontario Centre of Innovation (33080).

6.11 Acknowledgments

Thanks to all the team members of Ubilab, at the University of Waterloo, for their support.

6.12 NextGen data and Equity, Diversity, and Inclusiveness

Equity is defined as the removal of systemic barriers and biases enabling all individuals to have equal opportunity to access and benefit from the program. To achieve this, everyone involved in the research ecosystem must get a thorough awareness of the systemic challenges that persons from underrepresented groups (e.g., women, persons with disabilities, Indigenous Peoples, racialized minorities, individuals from the LGBTQ2+ community) face and implement effective strategies to overcome them [366].

Diversity is defined as differences in race, colour, place of origin, religion, immigrant and newcomer status, ethnic origin, ability, sex, sexual orientation, gender identity, gender expression and age. To achieve quality in research and training, a diversity of perspectives and life experiences is important [367].

Inclusion is defined as the practice of ensuring that all individuals are valued and respected for their contributions and are equally supported. To achieve research and training excellence, it is essential to ensure that all team members are integrated and supported [367]. The NextGen data revolution presents an exciting frontier to expand public health research, broadening the scope of research and increasing the precision of answers. Despite these advances, scientists must be vigilant against also advancing potential harms toward marginalized communities. Paul Wesson and colleagues, in 2022, provided examples in which NextGen data or big data applications have (unintentionally) disseminated discriminatory practices, while also highlighting opportunities for big data applications to advance equity in public health [366]. The study mentioned big data is framed in the context of the five Vs (volume, velocity, veracity, variety, and value), and the authors of the study proposed a sixth V, virtuosity, which incorporates equity and justice frameworks. The authors present analytic approaches for improving equity using social computational big data, fairness in machine learning algorithms, medical claims data, and data augmentation as examples, emphasizing the biassing influence of data absenteeism and positionality, and concluding with recommendations for incorporating an equity lens into big data research [366].

6.12 Contributions to the PhD Thesis

This publication provides the base for the theoretical understanding about use of non-traditional data sources for the public health surveillance. IoT based big data have the required characteristics as described in the literature to be eligible for public health surveillance.

Ethics approval for this study was not applicable, because this chapter does not contain any studies with human or animal subjects.

Chapter 7 Measuring the Impact of Stay-at-Home Policies in Canada During the COVID-19 Pandemic

7.1 Introduction

7.1.1 Background

The World Health Organization (WHO) declared the outbreak of coronavirus as a global pandemic on March 11, 2020 [71]. None of the countries worldwide were prepared to handle an epidemic of this scale. Most countries witnessed phased lockdowns [110] and total shutdown, drastically impacting their economy [368]. Additionally, this resulted in several new additions to hygiene standards [369] such as physical distancing, frequent handwashing [370], use of hand sanitizers, use of face masks [371], reduced social activities and gatherings as preventive measures to curb the spread of COVID-19 [372].

Apparently, lockdown emerged as an essential tool by the governments to reduce the spread of COVID-19. Taking into account, the implications of the Government guidelines such as work-from-home protocols for several companies and online schooling, these measures had likely have repercussions on adults, children, and youth population [373–375].

Maintaining a physically active lifestyle with the right balance of exercise and rest is crucial for achieving an overall positive state of health, well-being, thereby improving the quality of life of the population irrespective of age, sex, and other sociodemographic indicators [376,377].

Despite several public health measures that attempt to encourage physical activity, it has been found that people of all ages choose sedentary behaviour over an active lifestyle [73,378–380]. However, the pandemic has limited physical activities, making it more difficult to achieve the recommended physical activity targets [381]. Moreover, sleep durations and patterns were also affected during the pandemic restrictions [373,378,381,382] for different age groups. Scientific studies suggest that sleep, sedentary behaviour, and physical activity are associated with a broad spectrum of chronic diseases including diabetes, hypertension, cancer, or mental health problems [383]. Interestingly, a recent study shows that even mild physical activity of daily routine has a positive effect on health [74].

7.1.2 24-Hour Movement Guideline and PASS indicator

Canada developed a 24-Hour Movement Guideline in 2020, for all ages, laying guidance on the ideal amount of physical activity, sedentary behaviour, and sleep for an individual in a day (see Figure 24) [384]. This guideline has the potential to capture movement across different times, environments, and effects of the season. Though the procedure is developed with a crucial ambitious aspect of health, it is challenging to get the required data to measure its impact. This framework requires real-time data collection from a broader range of people [385], which is feasible using data generated from wearables and IoT.

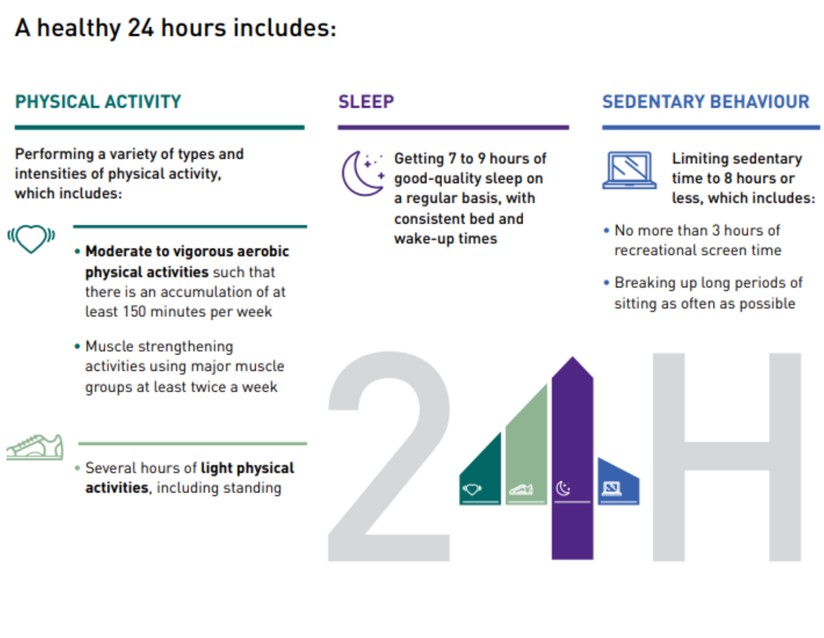


Figure 24. Canada's first 24-hour movement guideline for adults: image extracted from a report of Canadian Society for Exercise Physiology [384].

Canada's Public Health Agency developed the PASS (Physical Activity, Sedentary Behaviour and Sleep) indicator framework to monitor and measure the population health for the 24-hour movement guideline [384]. Within the PASS framework regular collection and reporting of these indicators provide insight into the changes in population-level behaviour and other associated factors that directly or indirectly influence them [127]. A systematic review published in 2018 by Valerie *et al.* concluded that despite having several objective physical activity measurement tools, it is still underutilized in the real world, which would have addressed the challenges of comprehensive and consistent collecting, reporting, and analyzing of physical activity metrics [386]. At the global policy level, the World Health Assembly in 2018 agreed on a global target to decrease physical inactivity by 15% by 2030 and align with the Sustainable Development Goals [387,388]. The action plan includes ensuring regular surveillance and physical activity monitoring [387,388] to achieve this ambitious goal.

7.1.3 Public Health Surveillance and the Use of Technology

Zero-effort technologies are critical in the digital world, and the future of public health surveillance depends on this [100]. Datasets like the ecobee "Donate your Data" program could be integrated with other technologies, datasets, and public health agencies strategies to tackle public health surveillance challenges and improve upon barriers related to traditional data collection methods. For example, when a Remote Sensor (RS) is placed in a household, it can provide insights on occupancy [389] and indicate different household activities, such as physical activity and sleep [390] . The use of RS addresses the challenges of participants' declining engagement, low response rates in surveys and focus groups, and technical barriers to the wearable technology [386]. When an RS is placed in a household, it behaves as a passive data capturing device wherein the individuals remain engaged in their day-to-day activities and their mobility data is collected. In this way, the recall bias also gets eliminated which occurs frequently when the participants need to quantify the amount of sleep and durations of physical activities on a daily basis [386].

Motion data can provide insights into the amount of sleep in the household by using the absence of movement as a proxy indicator of sleep intervals [390]. The lack of activation of motion sensors represents several outcomes depending on context. When there is a brief period of inactivity between two active periods with a regular pattern, it might represent sleep activity. A sustained period of inactive sensor shows either the absence of human individuals at the household due to vacation or out of the house for other reasons [27,391].

7.1.4 Objectives

This study's primary goal is to contrast population behaviours before and during the COVID-19 pandemic using household-level data collected via smart thermostats. As previously discussed, policies in Canada have been implemented to minimize the population's exposure to COVID-19. This study aims to understand household and population-level behaviours, generating insights about how the pandemic has affected the population's lifestyle.

The objectives of this study are to:

(a) Identify household occupancy patterns and variations caused by the COVID-19 pandemic, using the motion and thermostat sensor time-series data from ecobee, and

(b) Determine the impact of policy-level changes during the COVID-19 pandemic such as lockdowns, on household behaviours like sleep parameters (bedtime, wake up time, sleep duration), and time spent in-house and out-of-the-house.

7.2 Methods

This is an exploratory study using data from smart thermostats from four provinces in Canada to measure the changes in lifestyle patterns before and during the COVID-19 pandemic.

Ethics approval for this study was obtained from the University of Waterloo Office of Research Ethics (#31377).

7.2.1 Dataset

"*Donate your Data*" is a program from ecobee, a smart thermostat company in Toronto, Canada. The company shares anonymized data from smart thermostat users across the globe for research purposes. Details about the "Donate your Data" program have been explained in Section 5.2.2.

7.2.2 Data Processing, Cleaning, and Analysis

The "*Donate your Data*" program has two main data tables: metadata and thermostat data. Ecobee shared this data with researchers through Google's BigQuery platform, which is part of the Google Cloud Platform (GCP). The data was transferred from GCP to a Microsoft Azure Gen2 storage space for cleaning and analysis. This approach was preferred due to the secure research environment hosted by the University of Waterloo in Microsoft Azure. As illustrated in Figure 25, the data analysis process leveraged multiple cloud platforms, data analysis software and techniques. Both Python and R programming languages were used to make the process smooth and effective. Time series data wrangling is easier with Python and handling a dataset of seven terabytes requires extensive parallel computing infrastructure. Databricks is a flexible and scalable service within Microsoft Azure where an enormous volume of data can be managed effectively through scalable computational nodes.

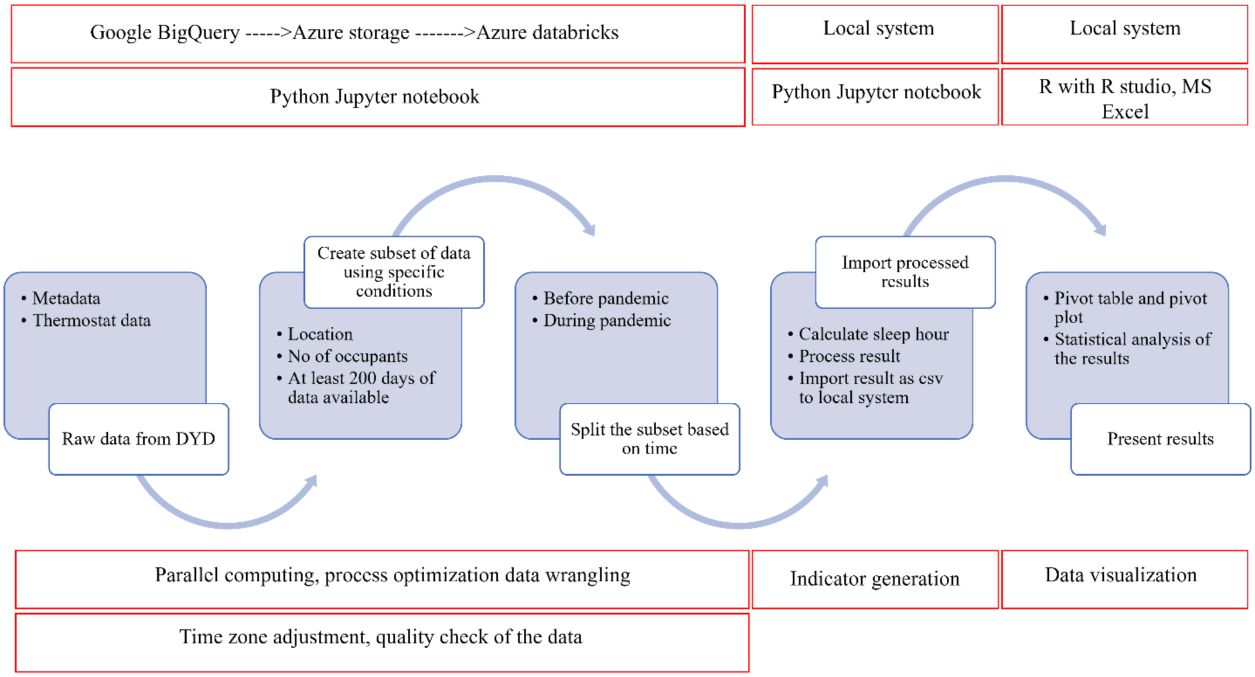


Figure 25. Framework to understand data processing, cleaning, and analysis process for this research.

As described in previous studies, the thermostat data was aggregated from five-minute to 30-minute intervals [389,392]. Similarly, data aggregation for sensor activation has been done at the population level per day to compare the before and during the COVID-19 pandemic. Another study from our research team that I was part of used the following methodology to calculate behavioural indicators [393]. The study has been published on the JMIR mHealth and uHealth. I have leveraged this algorithm in my thesis, expanding the household level analysis to a population level analysis. Based on the time stamp and number of sensors activated, our scripts calculated sleep time, wake up time, sleep duration, time spent in-house, and away time at the household and population level. The text below is a direct extraction from our publication, where I describe the algorithm used.

"The sum of activation of all the sensors for every 30 minutes interval demonstrates the activity in that period. Every 5 minutes of positive activation is given a score of 1. For a 30-minute interval, positive activity was defined as a score greater than or equal to 4 (one sensor active for 4x5 minute interval (20 minutes) or four sensors each active for five minutes). Sums falling below this threshold were considered noise. A binary vector presents a daily record with 48-time slots" [390]. *"To assess the different sleep parameters, each day has been divided into two parts: (1) 12:00 am until 12:00 pm and (2) 12:00 pm to 12:00 am. For every two consecutive days (day 0 and day 1), a sleep cycle was defined as the second part of day 0 combined with the first part of day one (Figure 26).*

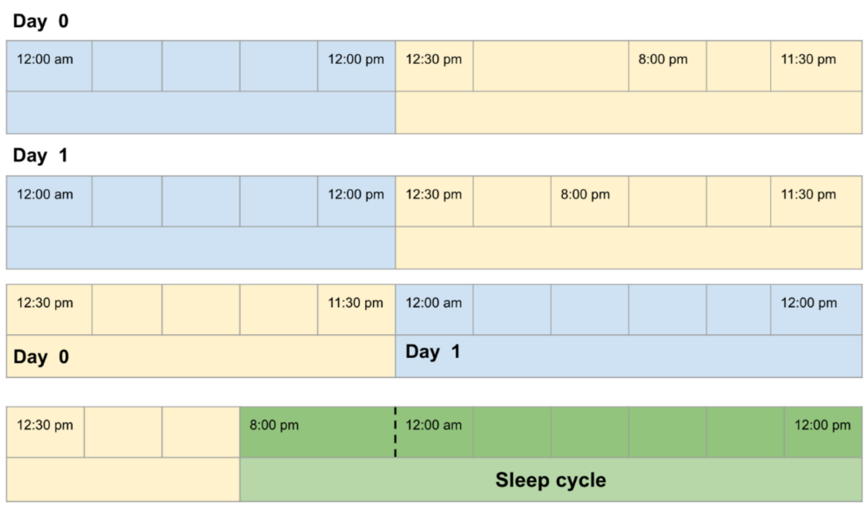


Figure 26. Sleep cycle and division of the day.

The sleep cycle records were segmented into clusters for each household using the Gaussian mixture model. The regular pattern in each cluster was identified by the probability of sensor activation in each time slot (counted as positive motion). For each cluster, the sleep time, wake-up time, and sleep duration of the regular pattern were assessed by hypothesizing that the deactivation (sleep time) occurs before activation (wake-up time). The earliest deactivation *(sleep time) can start from 8:00 pm. The longest interval between two consecutive deactivation and activation times (from 8:00 pm till 12:00 pm) can identify sleep time, wake-up time, and sleep duration. The overall results for each household are defined by the weighted average of the results from all the clusters by considering the cluster weight as: (w_i=number of days in cluster i total number of days). In addition to sleep parameters, we identified the average time spent at home. The 24-hour daily household records were segmented into different clusters. The regular pattern of each cluster identified the duration of activation. The overall result was defined by the weighted average from all the clusters. For each household, we replicated the analysis for the different time scales of the season, weekday, and seasonal weekday as explained below and compared each scale's impact on sleep parameters and time spent at home. The data of each parameter is divided into different time scale subsets. First, the descriptive statistics are identified* [391]*."*

7.2.3 Population Selection

The world's second-largest country (by total area), Canada, has ten provinces and three territories, extending from the Atlantic to the Pacific and northward into the Arctic Ocean. The population of Canada is more than 38 million as of 2021 [394], and the distribution of the population is not uniform across the provinces, as shown in Table 6. I have selected four provinces, Ontario, Alberta, Quebec, and British Columbia, as around 86% of the Canadian population resides within these provinces, as per the Statistics Canada report published in 2021 [394].

Table 6. Population of Canada and provinces for the years 2020 and 2021, along with the corresponding share of the Canadian population.

Geography	**2020**	**2021**	**Share (%)**
Canada	37,979,854	38,131,104	100%
Ontario	14,723,497	14,789,778	38.8%
Quebec	8,572,054	8,585,523	22.5%
British Columbia	5,142,404	5,174,724	13.6%
Alberta	4,417,006	4,444,277	11.7%
Manitoba	1,378,818	1,382,904	3.6%
Saskatchewan	1,179,618	1,179,906	3.1%
Nova Scotia	977,043	982,326	2.6%
New Brunswick	781,024	783,721	2.1%
Newfoundland and Labrador	522,994	520,286	1.4%
Prince Edward Island	159,249	160,536	0.4%
Northwest Territories	45,201	44,991	0.1%
Yukon	41,980	42,596	0.1%
Nunavut	38,966	39,536	0.1%

Source: [394] Statistics Canada. Table 17-10-0009-01 Population estimates, quarterly.

7.2.4 Data Analysis Methods

The thermostat data from the DYD has been divided into two distinct phases based on the declaration of the COVID-19 pandemic at the national level in Canada on March 18, 2021. Any data before that time has been considered before the COVID-19 pandemic, and data points after that time point are assessed as during the COVID-19 pandemic. The starting point for the before COVID-19 data was January 1, 2017, and the endpoint was March 18, 2020. The start date for "during COVID-19 data" was March 19, 2020, and the end date was March 18, 2021. This period was selected as it compasses one year of data.

I have started the data analysis using data visualization techniques to capture 24-hour household behaviour and household-level trends. The 24-hour household activity includes sleep, sedentary behaviour, and physical activity. Visual analysis of the data for selected households has been performed to identify and understand the pattern. I also explored the changes in population behaviour with time during the pandemic. The behavioural health indicators for this study include sleep-time (time period where no activity was observed for at least a period of an hour at night), wake-up time (time period when the first activation of the sensor was observed after a prolonged period of inactivity, i.e., after sleep-time), out-of-the-house time (time period where no activity was observed for at least a period of an hour in the day), and time spent in-house (time period when the sensor gets activated due to various activities performed by individuals while in home).

Descriptive (mean and standard deviations) and inferential statistics for the selected indicators have been calculated. Paired t-test has been used to determine the test of significance for the average difference in sleep duration, out-of-the-house stay duration, and in-house stay duration. A positive difference between during and before the COVID-19 pandemic for any indicator represents more time spent for that indicator, whereas a negative value represents less time spent. A *P*-value lower than 0.05 is considered statistically significant.

7.3 Results

The findings show significant changes at the household and population level for the selected indicators due to pandemic-related policy changes. During the COVID-19 pandemic, people stayed at home for extended periods, and the time away from home was significantly reduced. A year-by-year, month-by-month heatmap visualization shows changes in patterns, intensity, and duration.

7.3.1 Visual Interpretation from Heatmaps

The heatmap from five typical households from the "*Donate your Data*" program have been presented below for the years 2019 and 2020 (before and during the pandemic), as well as the month of March 2020 (more granular visualization) to explore the interpretable difference in patterns between before and during the COVID-19 pandemic restrictions. The plot's x-axis represents a day, and the y-axis represents a time window of 30 minutes intervals. The lighter the colour, the higher is the sensor activation (more sensors activated in the house during that period), and the darker is the colour, the lower is the activation (fewer sensors activated in the house during that period) represented in these Figures 27-31.

7.3.1.1 Year-level Visual Inspection Through Heatmaps (2019 and 2020)

As shown in Figures 27-31, in 2019, the number of sensors activated for each time interval is lesser than in the year 2020. Notably, there is an evident change in the sensor activation immediately after the declaration of the COVID-19 pandemic at the global and national levels. The household-level activity increased dramatically after the declaration of work-from-home policy and stay-at-home order in March 2020. The typical amount of time spent in an office on any given day was replaced by spending the equivalent time at home.

Of the five households, two were from Ontario, and one each from Alberta, Quebec, and British Columbia. The number of individuals within these five households ranges from two to four, and the number of remote sensors ranges from six to eleven. The intensity of the sensor activation depends on the number of people residing within a household, whereas the number of sensors describes the high probability of capturing any activity within the household.

For Household 1 and 3, as there were more than two individuals within the households, even before the pandemic, the intensity of sensors activation was higher than Household 2 and 4 with two individuals.

Characteristic features of Household 1: The heat map in Figure 27 depicts the mobility activity of the household located in the province of British Columbia, with four members and nine remote sensors installed. In Figure 27a, it is evident that the weekdays had less mobility compared to weekdays. This can be ascribed to the reason that people go to work on weekdays out of their homes which contribute to out-of-the-house time. However, on weekends, the residents stay mostly at home, thus increasing the in-house times. This pattern was observed till March 2020 as shown in Figure 27b. Since the declaration of COVID-19 pandemic in March 2020, numerous regulatory changes such as lockdowns and work-from-home policy led to a significant increase in the in-house mobility of individuals, irrespective of the days of the week, because of more in-house time.

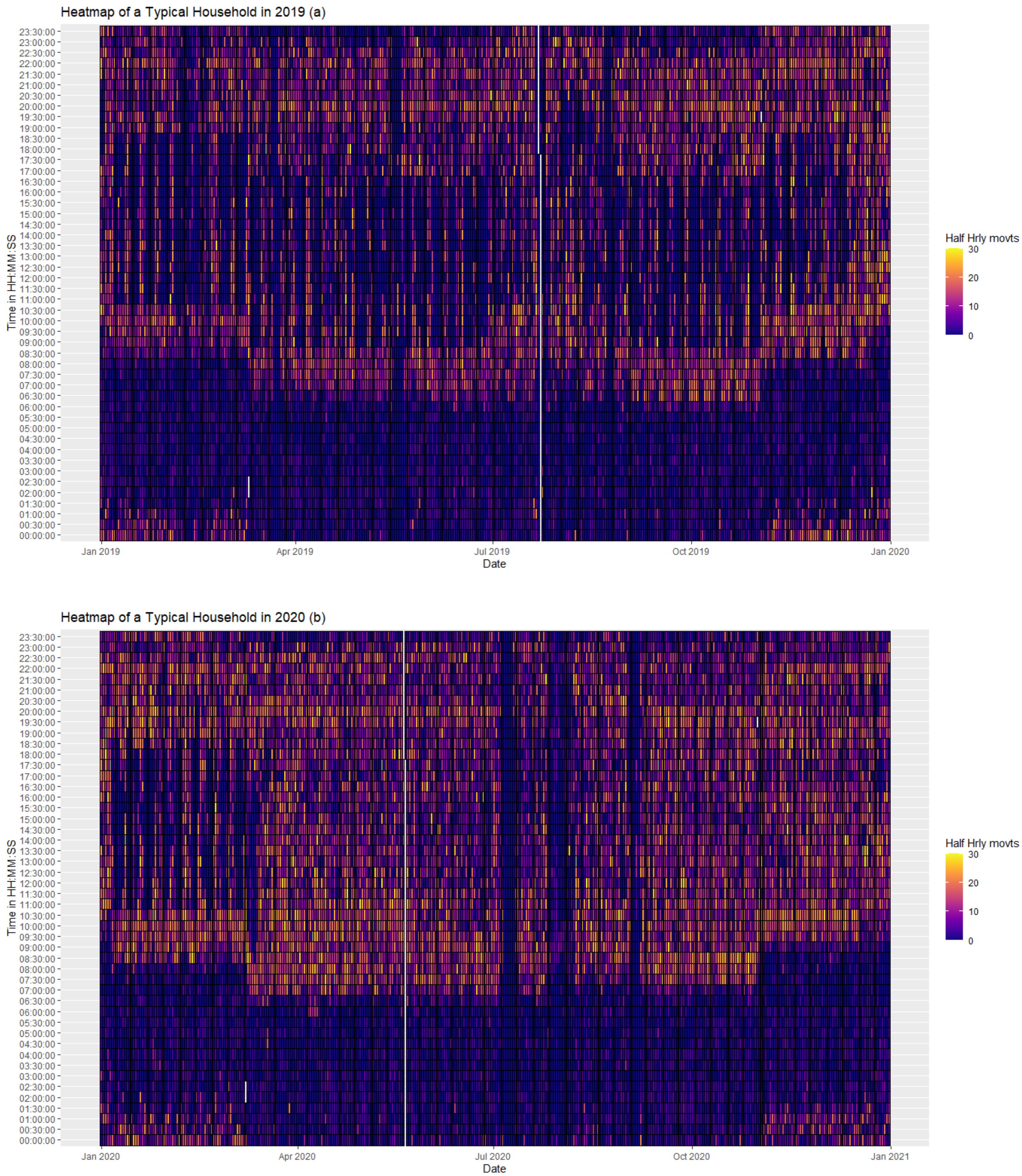


Figure 27. Annual heatmap of a single household sensor activation comparison between a) 2019 and b) 2020 during the pandemic for Household 1.

Characteristic features of Household 2: The heat map in Figure 28 depicts the mobility activity of the household located in the province of Ontario, with two members and six remote sensors installed. Compared to Household 1, there is an overall reduction of mobility before and during the COVID-19 pandemic as the number of residents was less. However, the mobility patterns were like those in Figure 27.

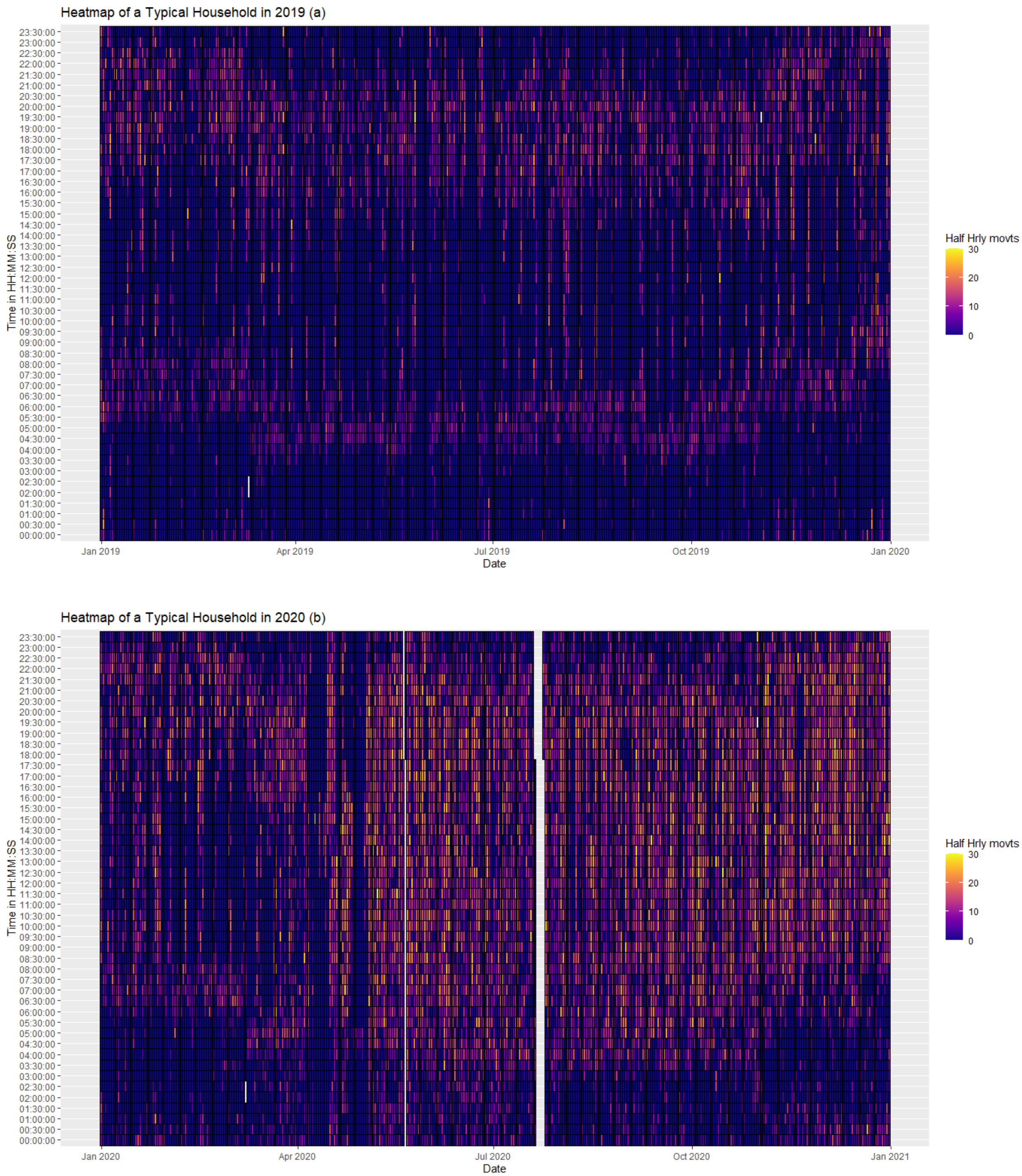


Figure 28. Annual heatmap of a single household sensor activation comparison between a) 2019 and b) 2020 during the pandemic for household 2.

Characteristic features of Household 3: The heat map in Figure 29 depicts the mobility activity of the household located in the province of Ontario, with three members and nine remote sensors installed. Compared to households 1 and 2, the overall in-house mobility is significantly high before and after the COVID-19 pandemic. Additionally, there is no difference in the patterns for weekdays and weekends which can be ascribed to the reason that not all members were going out of home even before the pandemic started.

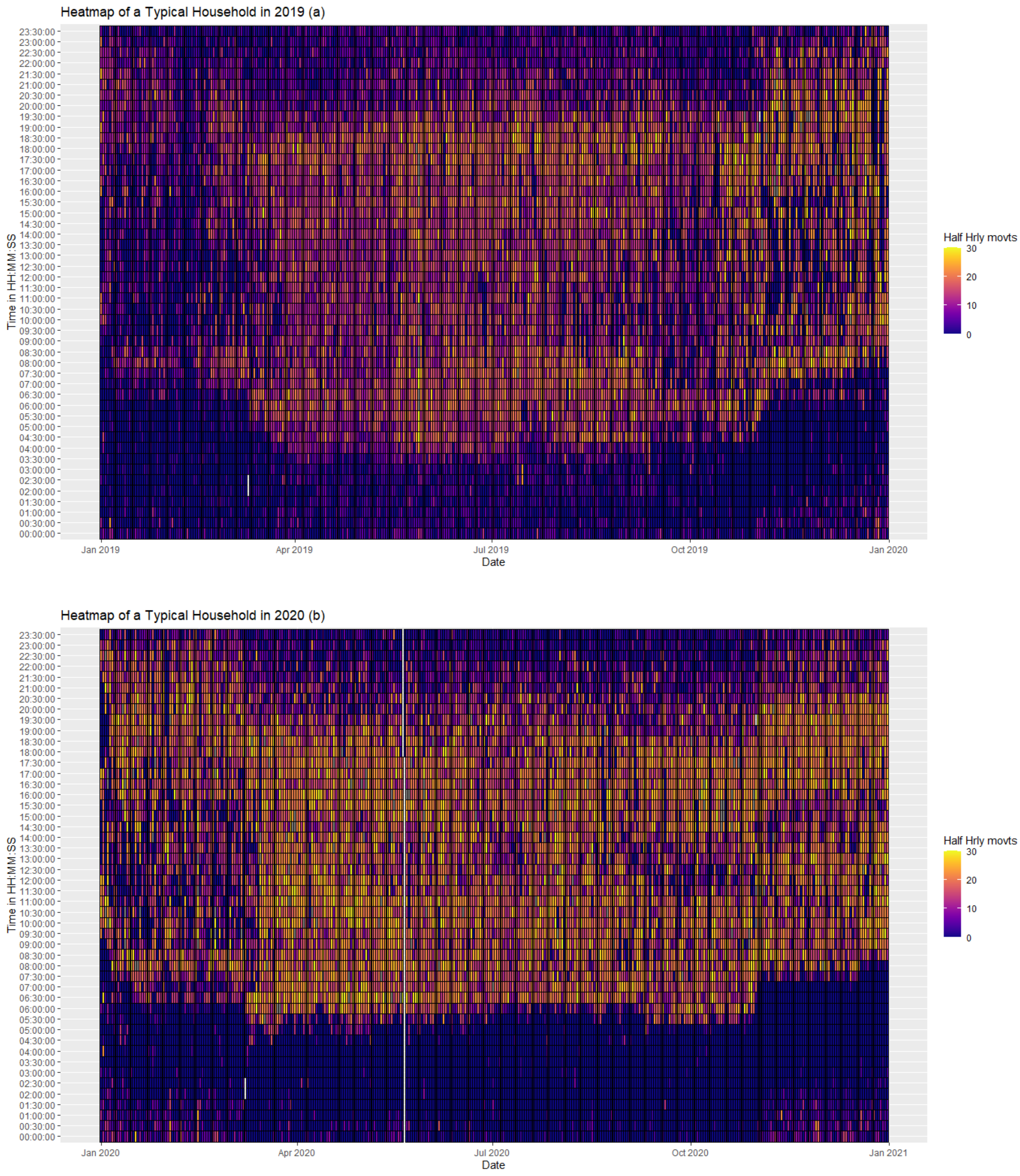
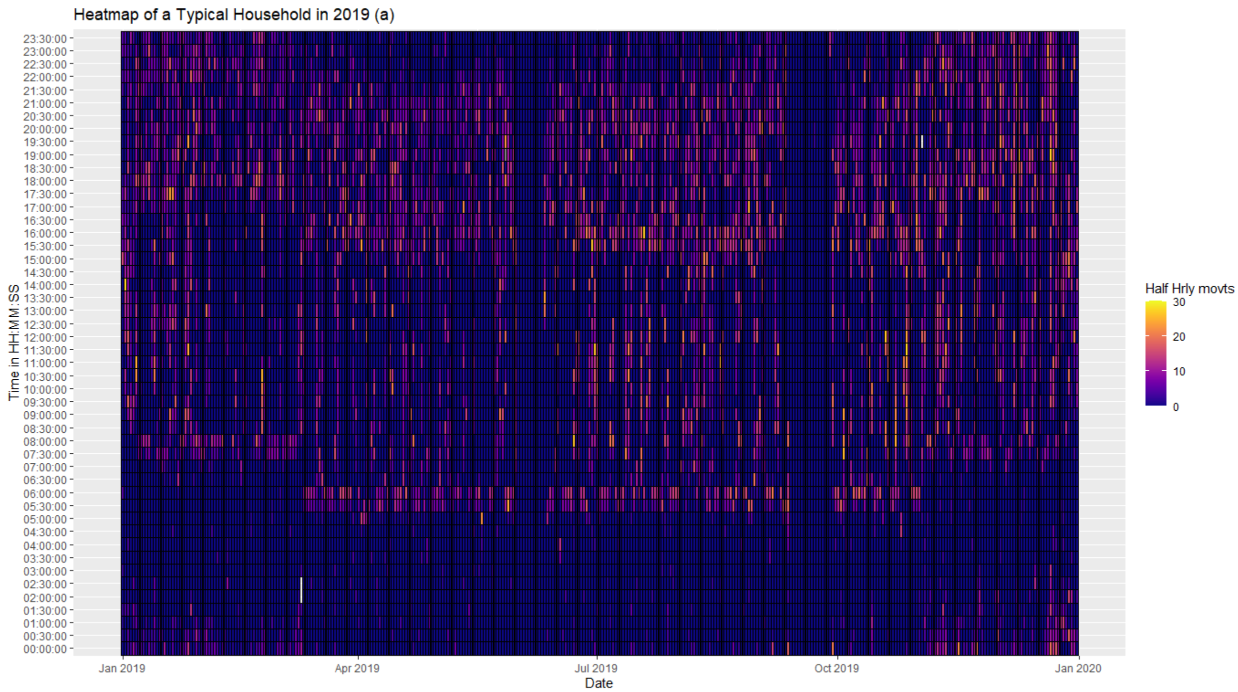


Figure 29. Annual heatmap of a single household sensor activation comparison between a) 2019 and b)2020 during the pandemic for household 3.

Characteristic features of Household 4: The heat map in figure 30 depicts the mobility activity of the household located in the province of Quebec, with two members and eleven remote sensors installed. The mobility patterns in this household are like those in household 2. Strikingly, zero activity periods for a few days observed in June 2019 and September 2019 may be attributed to either vacation time or some other activities related to residents spending time out-of-the-house.



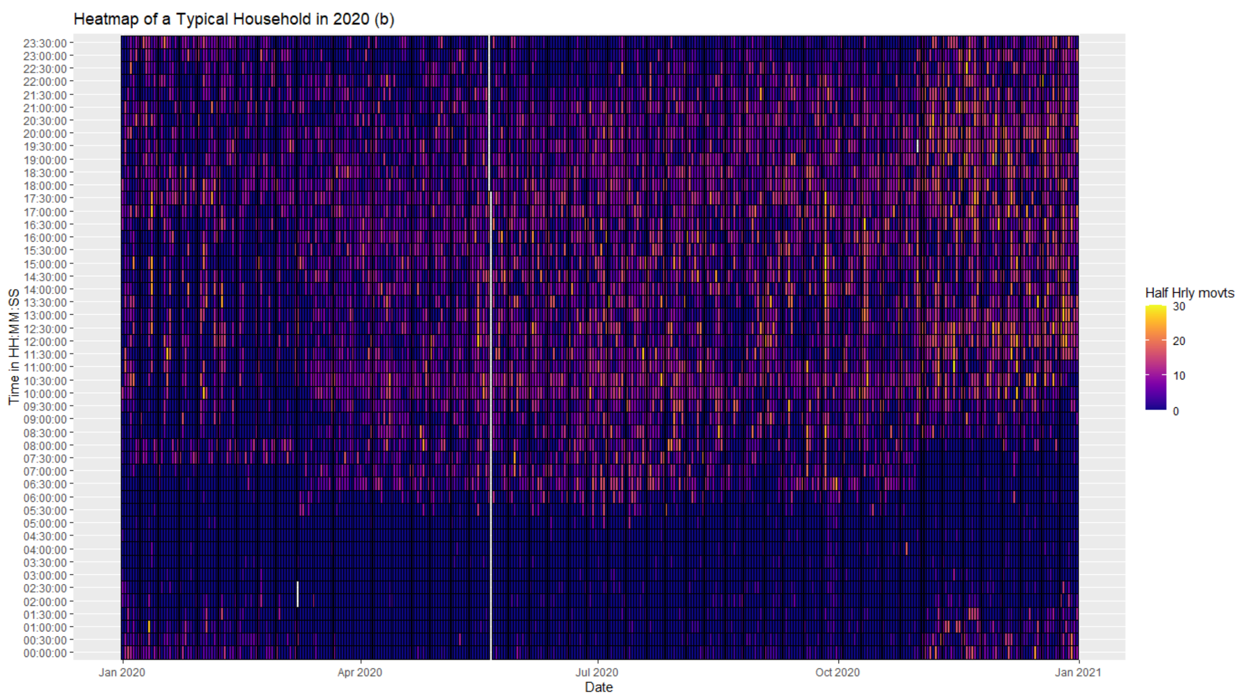


Figure 30. Annual heatmap of a single household sensor activation comparison between a) 2019 and b) 2020 during the pandemic for household 4.

Characteristic features of Household 5: The heat map in Figure 31 depicts the mobility activity of the household located in the province of Alberta, with two members and six remote sensors installed. Strikingly, as evident from the graph, strict wake-up routines were followed throughout the year.

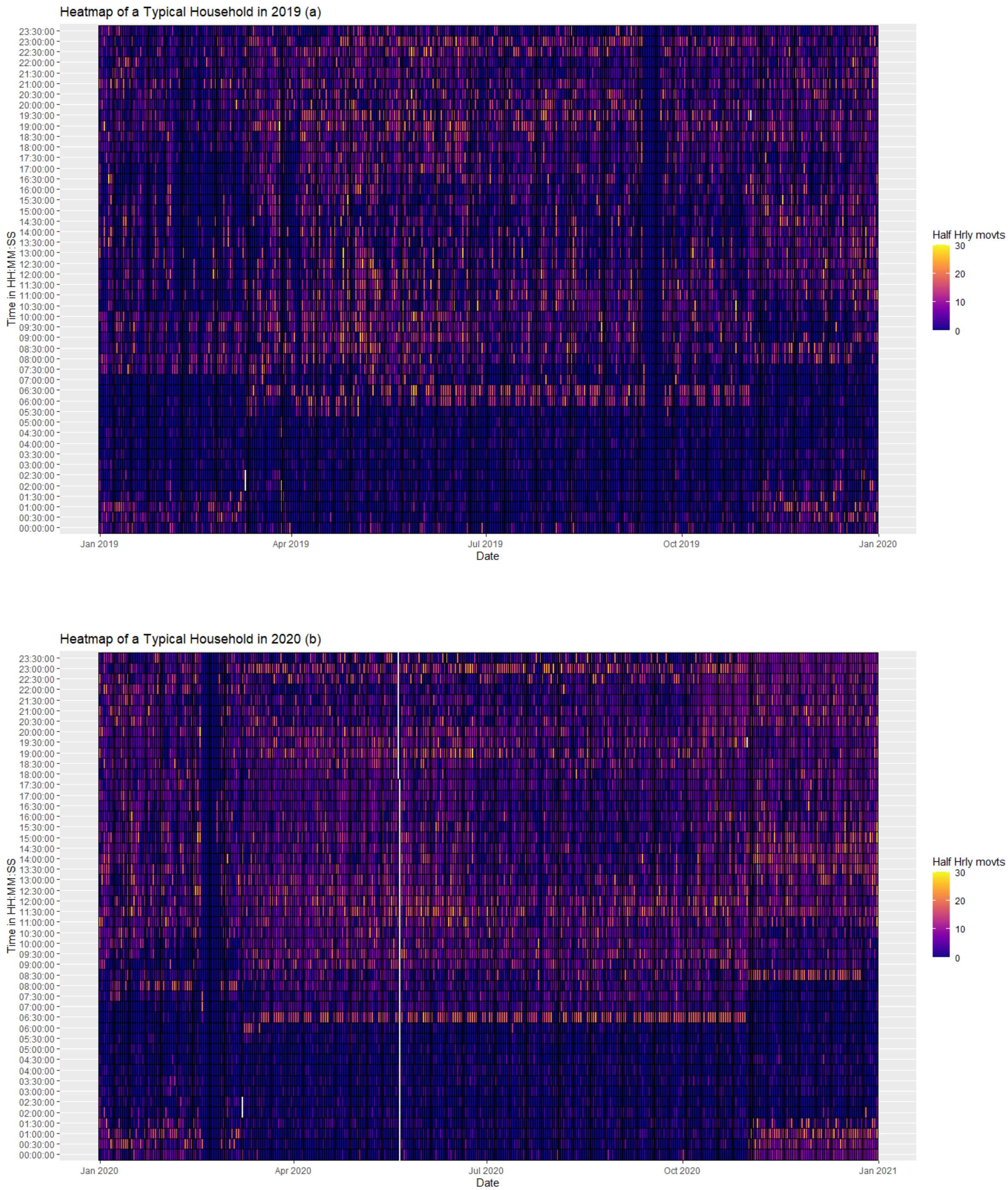


Figure 31. Annual heatmap of a single household sensor activation comparison between a) 2019 and b) 2020 during the pandemic for household 5.

These figures show the difference in sensor activation, which can be a proxy indicator to measure time spent in-house and out of the home, as well as sleep parameters of each observed household. Upon compilation, this data has the potential to provide population-level insights. This would be the first time that such indicators are created based on IoT data, for a larger population.

7.3.1.2 Month-Level Visual Inspection Through Heatmaps (March 2020)

The same data from the five households for the month of March 2020 has been visualized below. Figures 32-36 illustrate the day-level changes in the activation of the sensors during March 2020, depicting the impact of "stay-at-home" order following COVID-19 pandemic declaration. As observed from the heatmaps, the intensity of sensors activation is consistent till March 15, 2020, however, following this date, the intensity of sensors activation changed. Owing to the daylight savings in Canada from March 8, 2020, the data of sensors activation for an hour is missing as represented by the cells colored in white.

Figure 32 shows the changes in Household 1 during March 2020. It is evident that before declaration of the COVID-19 pandemic, the weekdays had less mobility compared to weekends. This can be ascribed to the reason that people go to work on weekdays out of their homes which contribute to out-of-the-house time. However, on weekends, the residents stay mostly at home, thus increasing the in-house times. This pattern was observed till mid-March 2020, followed by increase in mobility irrespective of days of the week.

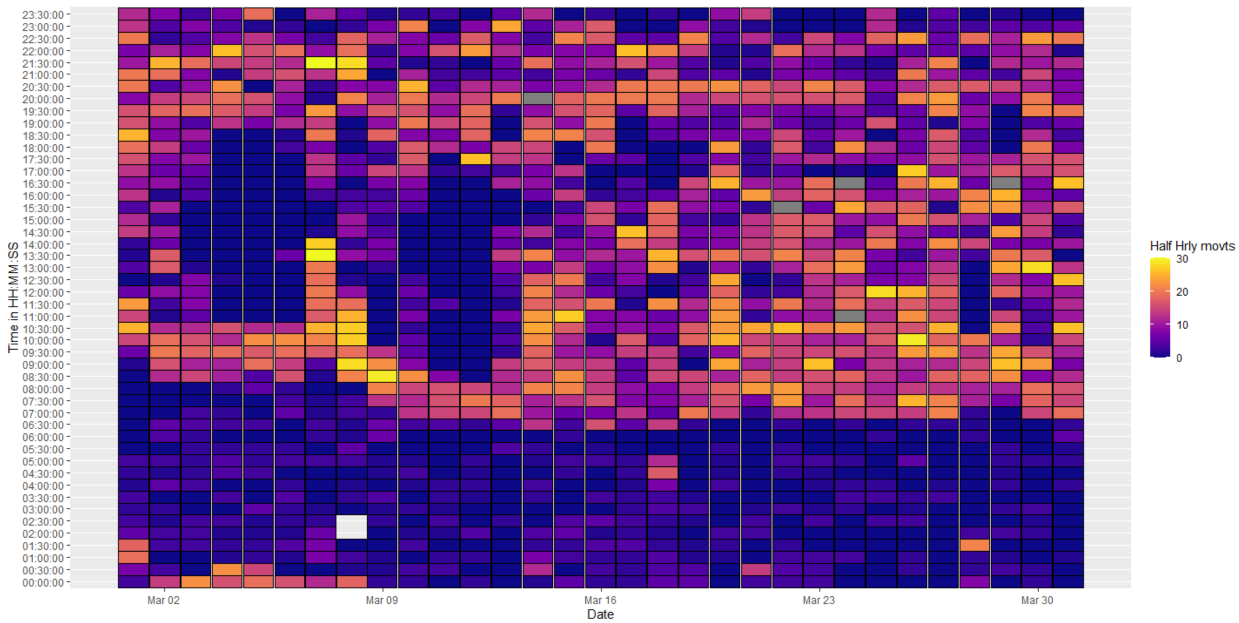


Figure 32. Heatmap of the activity in the Household 1 for March 2020.

Figure 33 shows the changes in Household 2 during March 2020, and like Figure 32, before the declaration of COVID-19 pandemic, weekdays had less mobility compared to weekdays. This pattern was observed till the third week of March 2020, followed by increase in mobility irrespective of days of the week.

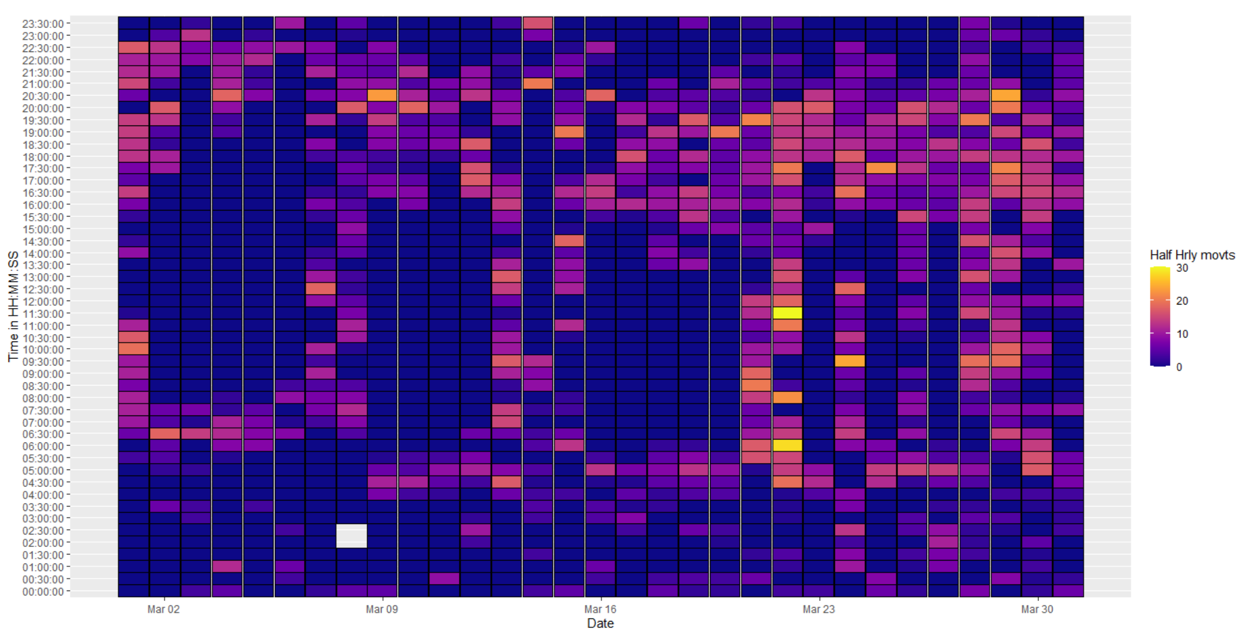


Figure 33. Heatmap of the activity in the Household 2 for March 2020.

Figure 34 shows the changes in Household 3 during March 2020, and the pattern of sensor activation was very different from Households 1 and 2. Even before the declaration of COVID-19 pandemic, the weekdays and weekends had nearly similar mobilities. Following COVID-19 pandemic declaration, the sensors activation increased drastically, indicating higher in-house mobility.

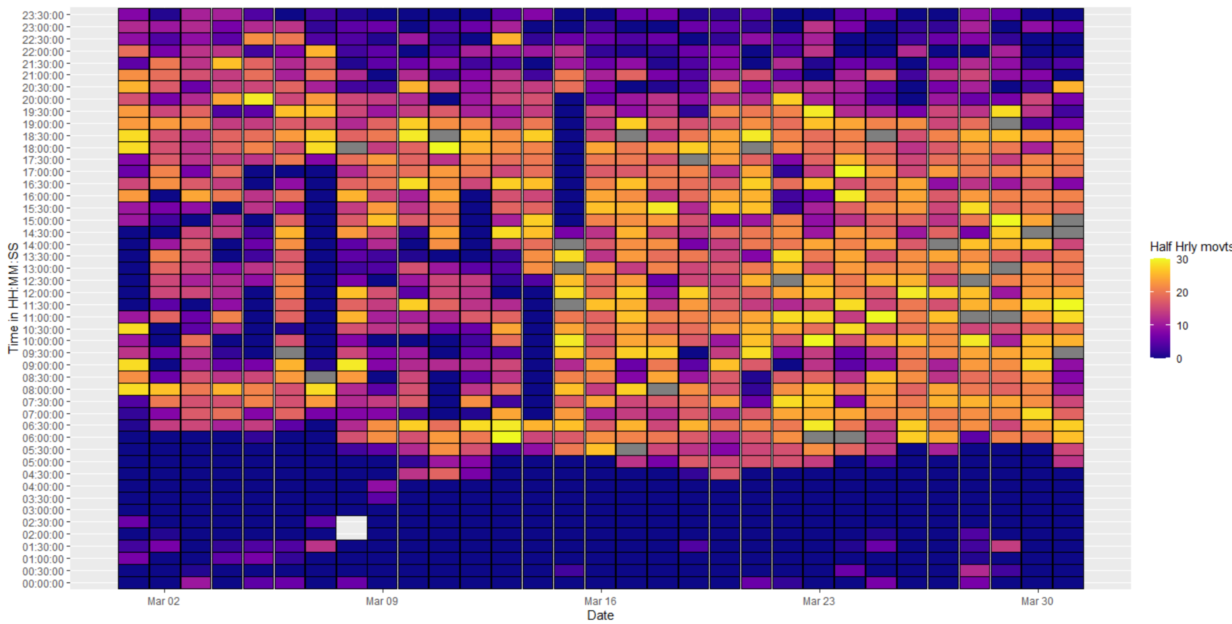


Figure 34. Heatmap of the activity in the Household 3 for March 2020.

Figure 35 shows the changes in Household 4 during March 2020. Before the declaration of COVID-19 pandemic, there was a unique pattern of away time such that there was no sensor activation observed for two days followed by one day higher sensor activation. Following COVID-19 pandemic declaration, there was an increase in in-house mobility irrespective of days of the week.

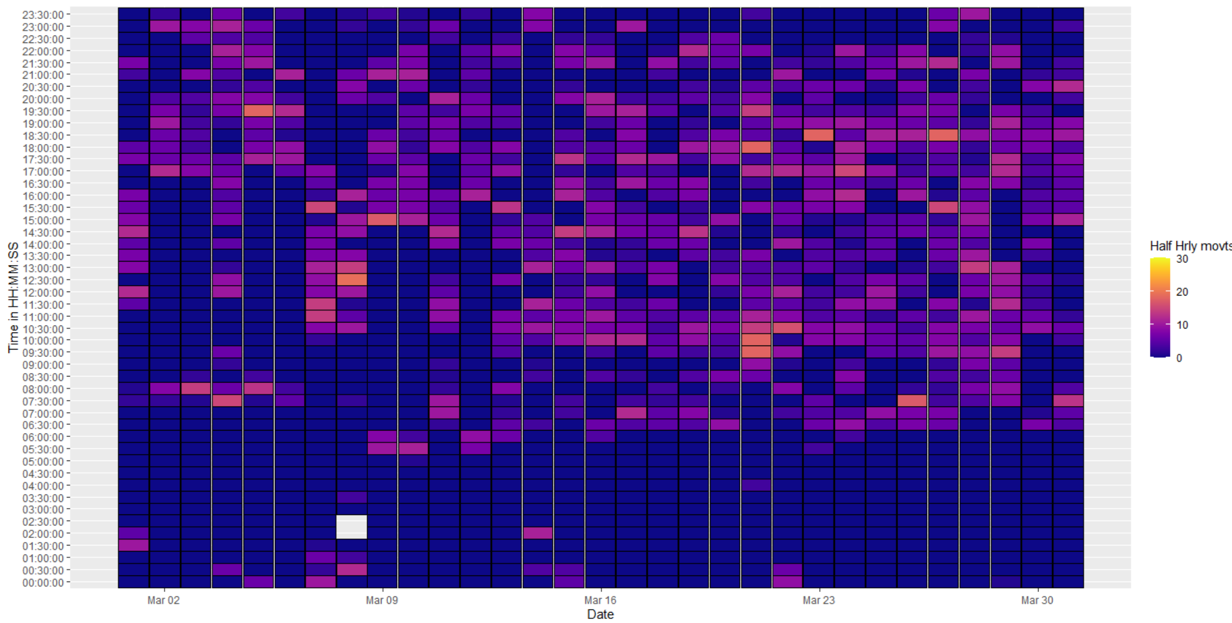


Figure 35. Heatmap of the activity in the Household 4 for March 2020.

Figure 36 shows the changes in Household 5 during March 2020. Strikingly, the weekdays and weekends had nearly similar in-house mobilities before and after the declaration of COVID-19 pandemic.

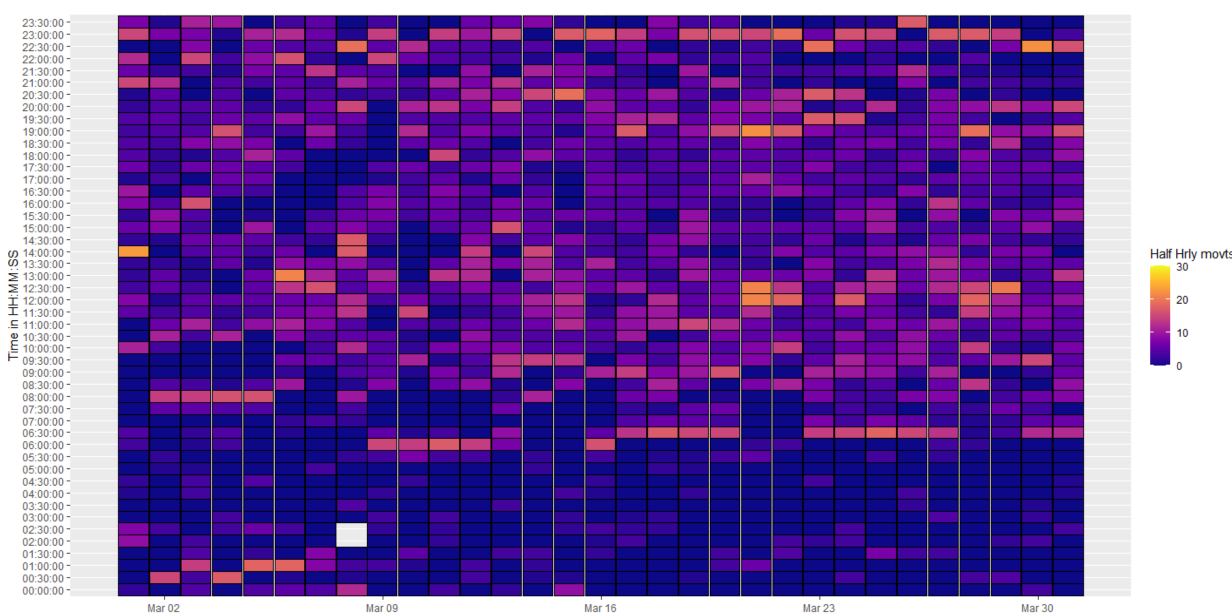


Figure 36. Heatmap of the activity in the household 5 for March 2020.

7.3.2 Weekday level analysis of the motion sensors activation (2019 and 2020) through Box plot

As observed from stratification and aggregation of the data into weekdays and weekends, during the pandemic, there was an overall increase in the total number of sensors activated for all the days of the week which is statistically significant at $P<.001$ level. Moreover, the difference between sensors activation on weekdays and weekends also reduced as shown in Table 7 and Figure 37.

Table 7. Comparison of sensor activation before and during the COVID-19 pandemic for Canada.

	Before Mean±SD	**During Mean±SD**	**Mean Difference**	**95% CI**	**T value**
Monday	0.267±0.043	0.313±0.0523	0.046***	0.036-0.055	9.6173
Tuesday	0.259±0.0424	0.310±0.0517	0.051***	0.042-0.060	10.912
Wednesday	0.255±0.0432	0.310±0.0522	0.055***	0.046-0.064	11.653
Thursday	0.256±0.0404	0.309±0.0522	0.053***	0.044-0.062	11.547
Friday	0.264±0.0427	0.308±0.0536	0.044***	0.035-0.053	9.2417
Saturday	0.286±0.0487	0.312±0.0565	0.026***	0.016-0.037	5.1105
Sunday	0.289±0.0475	0.314±0.0564	0.025***	0.014-0.034	4.763
Total	0.268±0.046	0.311±0.054	0.043***	0.039-0.046	23.04

*** Statistically significant at $P<.001$ level, SD-Standard deviation. CI- Confidence interval

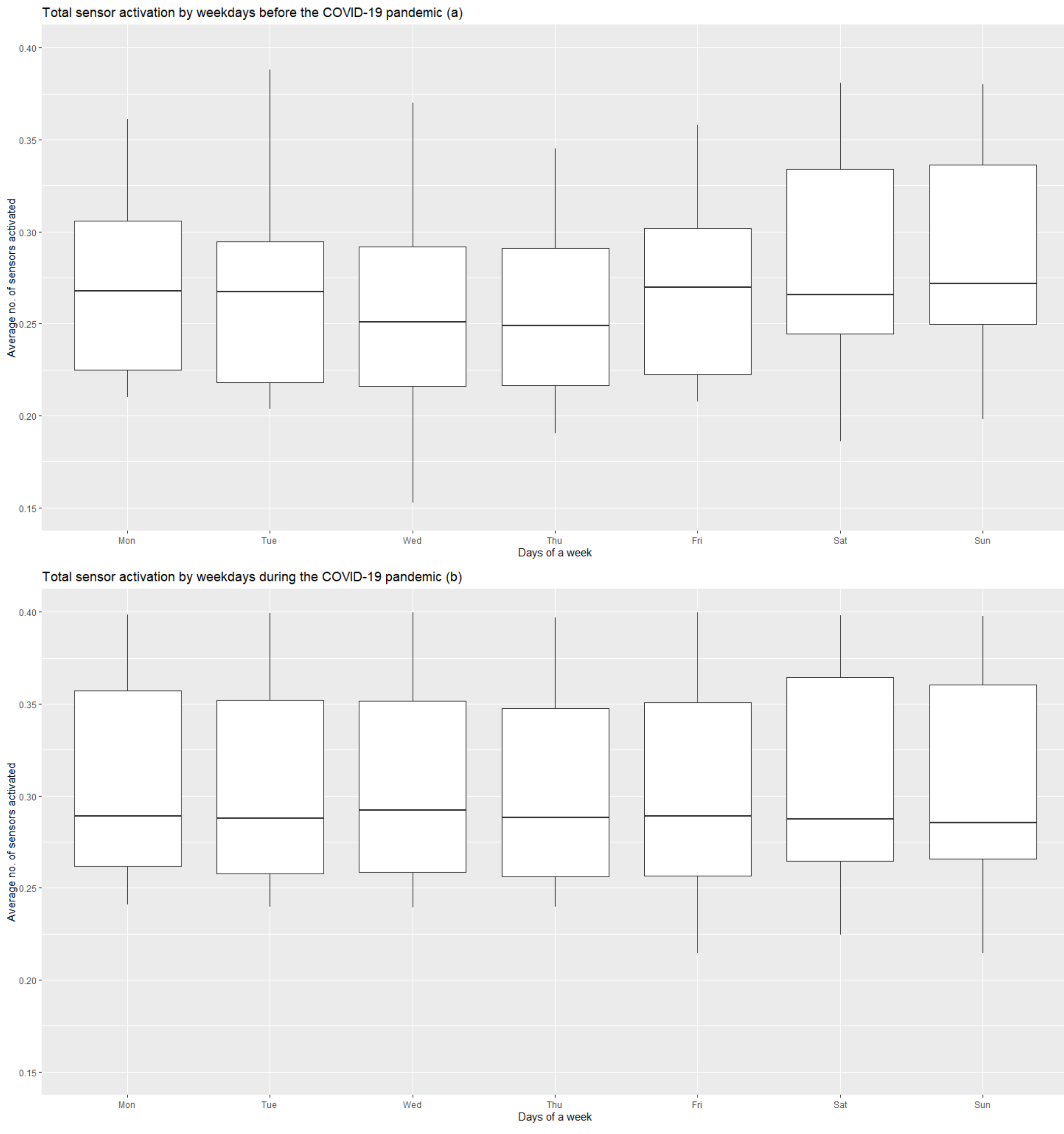


Figure 37. Comparison of total sensor activation between weekdays variation a) before and b) during the COVID-19 pandemic.

7.3.3 Month and Weekday-level analysis of the motion sensors activation (2019 and 2020) through Box plot

Similarly, re-stratifying and aggregating the data into months and weekdays in Figure 38a and b illustrate that before the COVID 19 pandemic, the total number of sensors activation was lower in 2019 for all the months except April. In contrast, during 2020, the increase in sensor activation is consistent for all the months after March.

As shown on the Figure 38a, in 2019, the months of February and March had a wide range of mobility on days of the weeks followed by a regular pattern of sensor activation. In contrast Figure 38b shows, in 2020, the sensor activation increased drastically in April which slowly reduced till the month of September and subsequently the sensor activation increased till December. This “U” shaped pattern in 2020 attributes to the COVID-19 pandemic declaration and stay at home policy, followed by waves of positive cases. The low variability within the days of the week during March 2020 to December 2020 might be because of almost similar pattern of sensor activation across several households.

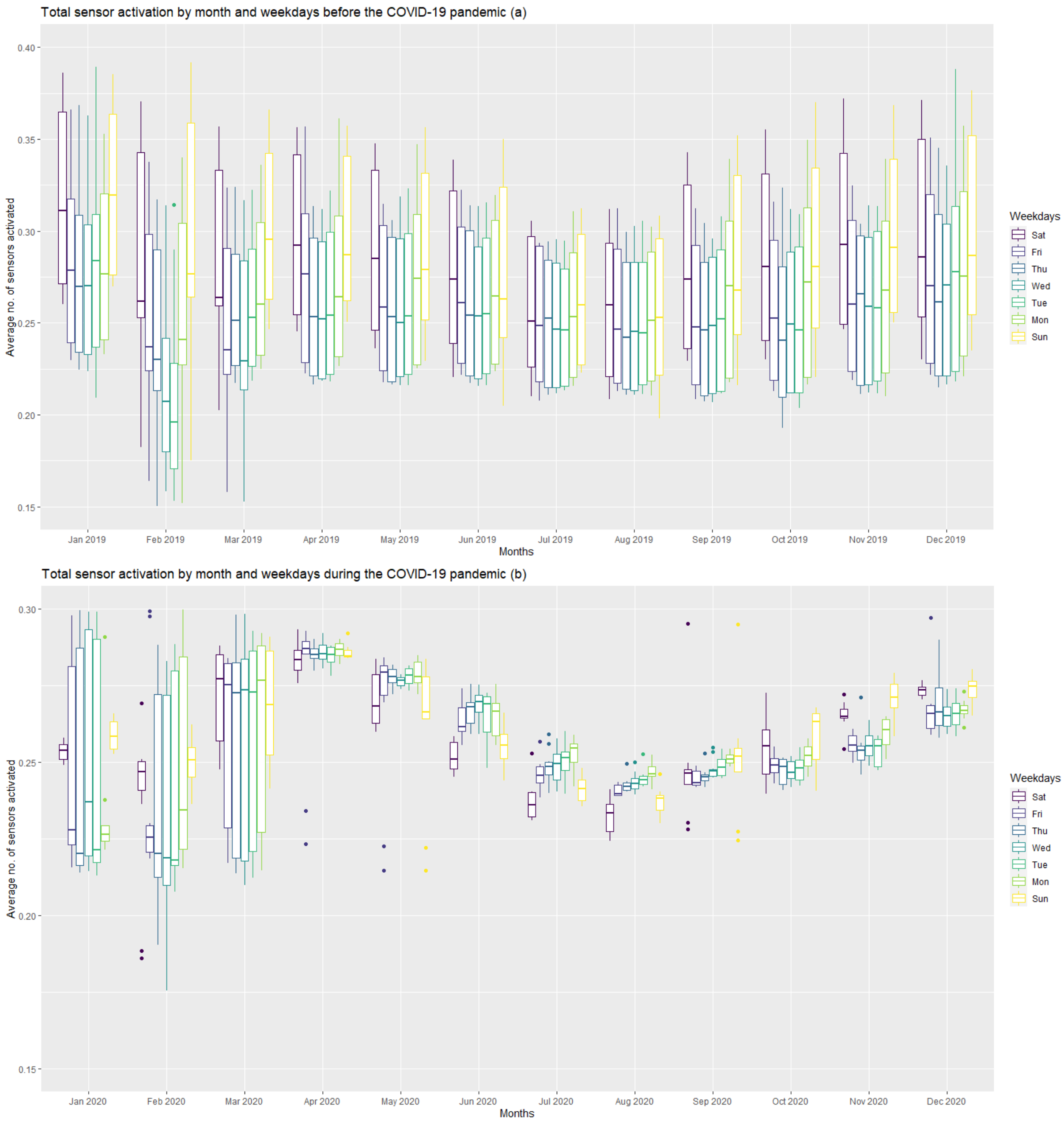


Figure 38. Comparison of total sensor activation among months and weekdays a) before and b) during the COVID-19 pandemic.

7.3.4 Comparison of population-level behavioural indicators before and during COVID-19 pandemic

Further analysis of the sensor activation at the population level in terms of behavioural change is presented in Tables 8 and 9. There is a statistically significant difference in average sleep duration, time spent in-house and out-of-the-house before and during the COVID-19 pandemic in Canada.

Before the pandemic, the average sleep duration was 8.7 ± 2.72 hours (Mean ± SD) whereas, during the pandemic, it was 9.2 ± 2.72 hours. The difference in average sleep duration was 30 minutes which was statistically significant at $P<.001$ level. The difference was more accentuated on Thursday, than on Friday and Saturday, as shown in Table 7. When stratified by weekdays the average sleep duration before the pandemic and during the pandemic on Thursday was 8.6±2.69 hours and 9.4± 2.76 hours, respectively.

For time spent in-house, there is a statistically significant increase in duration of 2.2 hours. Before the pandemic, households in Canada spent 5.1 ± 2.72 hours in-house whereas it increased to 7.3± 2.72 hours during the pandemic.

Before the pandemic, weekdays had less in-house time when compared to Saturdays and Sundays. For instance, people spent 5.0± 2.78 hours on average in-house on a typical Friday and it increased to 5.7± 2.70 hours on Saturday. However, during the pandemic time, the difference between the time spent in-house on weekdays was reduced.

There is a difference of 2.7 hours before and during the pandemic for time spent out-of-the-houses, which was statistically significant at the $P<.001$ level. Before the pandemic, the time spent out-of-the-houses was 10.2 ± 2.72 hours for weekdays whereas the average time spent out-of-the-houses was 7.5 ± 2.72 hours. During the pandemic, the difference of average time spent out-of-the-house between weekdays and weekends was reduced.

The change in sleep duration was further explored. The time to go to bed and wake up time were calculated. During the COVID-19 pandemic, the average difference for bedtime has minimally changed on the weekdays. The average bedtime from 11:15 pm changed to 11:13 pm during the pandemic. However, the change in average wake-up time was found to be increased by 51 minutes. Before the pandemic, the average wake-up time was 5:47 am whereas it changed to 6:38 am during the pandemic in Canada.

These results show an overall change in household behavioural patterns, sleep habits, time spent in-house, and time spent out-of-the-houses, which could be attributed to the policy changes implemented to curb the spread of COVID-19.

Table 8. Comparison of sleep, in-house and out-of-the-house stay duration before and during the COVID-19 pandemic for Canada.

	Before the Pandemic			**During the Pandemic**			**Mean Difference (During-Before)**		
	Sleep duration (hours) Mean±SD	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (hours) Mean±SD**	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (minutes)**	**In-house time (hours)**	**Out-of-the-house time (hours)**
Monday	8.6±2.69	5.1±2.69	10.3±2.69	9.1±2.70	7.4±2.71	7.5±2.71	30 ***	2.3 ***	-2.8 ***
Tuesday	8.6±2.68	4.9±2.68	10.6±2.68	9.1±2.71	7.4±2.72	7.5±2.72	30 ***	2.5 ***	-3.1 ***
Wednesday	8.6±2.65	4.7±2.65	10.8±2.65	9.1±2.71	7.5±2.71	7.4±2.71	30 ***	2.8 ***	-3.4 ***
Thursday	8.6±2.69	4.8±2.69	10.6±2.69	9.4±2.76	7.0±2.76	7.5±2.76	48 ***	2.2 ***	-3.1 ***
Friday	9.0±2.78	5.0±2.78	9.9±2.78	9.3±2.75	6.9±2.75	7.8±2.75	18 ***	1.9 ***	-2.1 ***
Saturday	8.9±2.84	5.8±2.85	9.3±2.85	9.2±2.68	7.4±2.68	7.4±2.68	18 ***	1.6 ***	-1.9 ***
Sunday	8.6±2.70	5.7±2.70	9.7±2.71	9.1±2.72	7.4±2.72	7.5±2.72	30 ***	1.7 ***	-2.2 ***
Total	8.7±2.72	5.1±2.72	10.2±2.72	9.2±2.72	7.3±2.72	7.5±2.72	30 ***	2.2 ***	-2.7 ***

*** Statistically significant at *P*<.001 level, SD-Standard deviation.

Table 9. Comparison of sleep timing before and during the COVID-19 pandemic for Canada.

	Before Pandemic		During Pandemic	
	Bedtime (pm.) Mean±SE	**Wake-up time (am.) Mean±SE**	**Bedtime (pm.) Mean±SE**	**Wake-up time (am.) Mean±SE**
Monday	11:10:03 ± 3.59	05:32:56 ± 0.50	11:09:34 ± 5.49	06:29:56 ± 0.59
Tuesday	11:09:30 ± 4.03	05:34:05 ± 0.49	11:08:09 ± 5.47	06:27:22 ± 1.00
Wednesday	11:12:31± 3.55	05:25:40 ± 0.47	11:09:25 ± 5.47	06:26:27 ± 0.58
Thursday	11:13:50 ± 4.00	05:36:41 ± 0.47	11:13:11 ± 6.26	07:32:41 ± 0.58
Friday	11:27:31± 4.29	06:20:05 ± 0.49	11:23:18 ± 5.50	06:38:34 ± 0.56
Saturday	11:22:52 ± 4.32	06:29:14 ± 0.55	11:23:02 ± 5.44	06:25:17 ± 0.57
Sunday	11:10:09 ± 4.03	05:39:27 ± 0.52	11:09:54 ± 5.46	06:28:30 ± 0.58
Total	**11:15:05 ± 1.34**	**05:47:51 ± 0.19**	**11:13:45 ± 2.13**	**06:38:23 ± 0.22**

SE-Standard error.

7.3.5 Province Specific Findings

Table 10 represents the sample distribution of the households from four selected provinces across Canada. In this study, the province of Ontario comprises more than half of the households, followed by Alberta. Findings from data analysis of each province are explained below to understand the dynamics and association with province-specific policy implications. Our household selection criteria included a minimum of 200 days of data available on the dataset.

Table 10. Number of households selected for each of the provinces.

	Number of houses on the DYD dataset	**Number of houses included in this study**	**The proportion of data by province with national level**
Canada	21690	7930	
Ontario	10968	4495	57%
Alberta	8046	2535	32%
British Columbia	698	345	4%
Quebec	1112	555	7%

7.3.5.1 Ontario

The results from the analysis for households in Ontario show that there was a statistically significant difference in average sleep duration, time spent in-house and out-of-the-house before and during the COVID-19 pandemic (Tables 11 and 12).

The average sleep duration before the pandemic was 9.2 ± 2.86 hours (Mean ± SD) whereas during the pandemic it was 9.4±2.73 hours. The difference in average sleep duration was 12 minutes which was statistically significant at $P<.001$ level. The difference was more accentuated on the weekdays than on the weekends, as shown in Figure 39. When stratified by weekdays, before the pandemic, the difference of sleep duration was higher between a typical Thursday (8.9±2.91 hours) and Friday (9.7± 2.73 hours) whereas during the pandemic, the difference was larger for Wednesday (9.3±2.73 hours) and Thursday (9.7± 2.74 hours).

For time spent in-house, there is a statistically significant duration of 2.5 hours observed before and during the pandemic for Ontario. Before the pandemic, households in Ontario spent 4.9± 3.71 hours in-house whereas it increased to 7.4±4.67 hours during the pandemic. Before the pandemic, weekdays had less in-house time than Saturdays and Sundays within a week. For instance, people spent 4.8±3.61 hours in-house on Friday and 5.5±3.92 hours on Saturday. However, during the pandemic, the difference between the time spent in-house on weekdays was reduced.

There is a statistically significant duration of 2.8 hours for time spent out-of-the-houses before and during the pandemic. Before the pandemic, time spent out-of-the-houses on weekdays was nearly 10 hours. The out-of-the-houses duration on Friday reduced to 9.5 ±3.67 hours compared to 10.5 ±3.98 hours on Thursday. During the pandemic, the average out-of-the-house time difference between weekdays and weekends was reduced.

The increased sleep duration was further explored. The times for going to bed and waking up were calculated. During the COVID-19 pandemic, the average difference for bedtime has minimally changed on the weekdays, whereas the average for Friday is 11.27 pm and for Saturday is 11.24 pm, respectively.

These results show an overall change in Ontario of household behavioural patterns, sleep habits, time spent in-house, and time spent out-of-the-houses, which were likely caused by the policies implemented to curb the spread of COVID-19.

Table 11. Comparison of sleep indicators and in-house stay duration before and during the COVID-19 pandemic for Ontario.

	Before Pandemic			During Pandemic			Mean Difference (During-Before)		
	Sleep duration (hours) Mean±SD	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (hours) Mean±SD**	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (minutes)**	**In-house time (hours)**	**Out-of-the-house time (hours)**
Monday	8.9±2.86	4.7±3.65	10.4±4.02	9.3±2.73	7.5±4.76	7.1±4.22	30 ***	2.8 ***	-2.3 ***
Tuesday	8.9±2.87	4.7±3.61	10.4±3.98	9.3±2.76	7.5±4.76	7.1±4.17	30 ***	2.8 ***	-2.3 ***
Wednesday	8.9±2.84	4.6±3.58	10.5±3.96	9.3±2.73	7.7±4.77	7.1±4.24	30 ***	3.1 ***	-3.4 ***
Thursday	8.9±2.91	4.6±3.56	10.5±3.98	9.7±2.74	7.2±4.64	7.1±4.25	48 ***	2.6 ***	-3.4 ***
Friday	9.7±2.73	4.8±3.61	9.5±3.67	9.5±2.75	7.0±4.48	7.4±3.95	-12 ***	2.2 ***	-2.1 ***
Saturday	9.7±2.81	5.5±3.92	8.8±3.93	9.5±2.64	7.4±4.59	7.1±3.94	-12 ***	1.9 ***	-1.7 ***
Sunday	9.1±2.83	5.4±3.96	9.6±4.19	9.3±2.72	7.5±4.63	7.2±4.06	12 ***	2.1 ***	-2.4 ***
Total	9.2±2.86	4.9±3.71	10.0±4.01	9.4±2.73	7.4±4.67	7.2±4.12	12 ***	2.5 ***	-2.8 ***

*** Statistically significant at *P*<.001 level, SD- Standard deviation.

Table 12. Comparison of sleep timing before and during the COVID-19 pandemic for Ontario.

	Before Pandemic		**During Pandemic**	
	Bedtime (pm.) Mean±SE	**Wake-up time (am.) Mean±SE**	**Bedtime (pm.) Mean±SE**	**Wake-up time (am.) Mean±SE**
Monday	11:00:58±7.21	05:57:08±1.17	11:00:15 ± 8.23	06:51:12±1.21
Tuesday	11:01:26±7.30	06:08:28±1.18	10:59:27 ± 8.23	06:54:27±1.20
Wednesday	11:02:38±7.13	05:48:44±1.16	11:00:53 ± 8.20	06:49:07±1.21
Thursday	11:07:11±7.22	06:05:12±1.17	11:06:22 ± 9.14	08:05:19±1.17
Friday	11:27:34±8.22	07:00:29±1.15	11:17:45 ± 8.25	06:59:56±1.19
Saturday	11:24:45±8.26	07:04:27±1.23	11:17:02 ± 8.15	06:40:15±1.19
Sunday	11:05:04±7.34	06:09:41±1.18	11:00:31 ± 8.21	06:48:51±1.20
Total	**11:09:36±2.54**	**06:18:14±0.29**	**11:05:59 ± 3.12**	**07:01:18±0.30**

SE- Standard error.

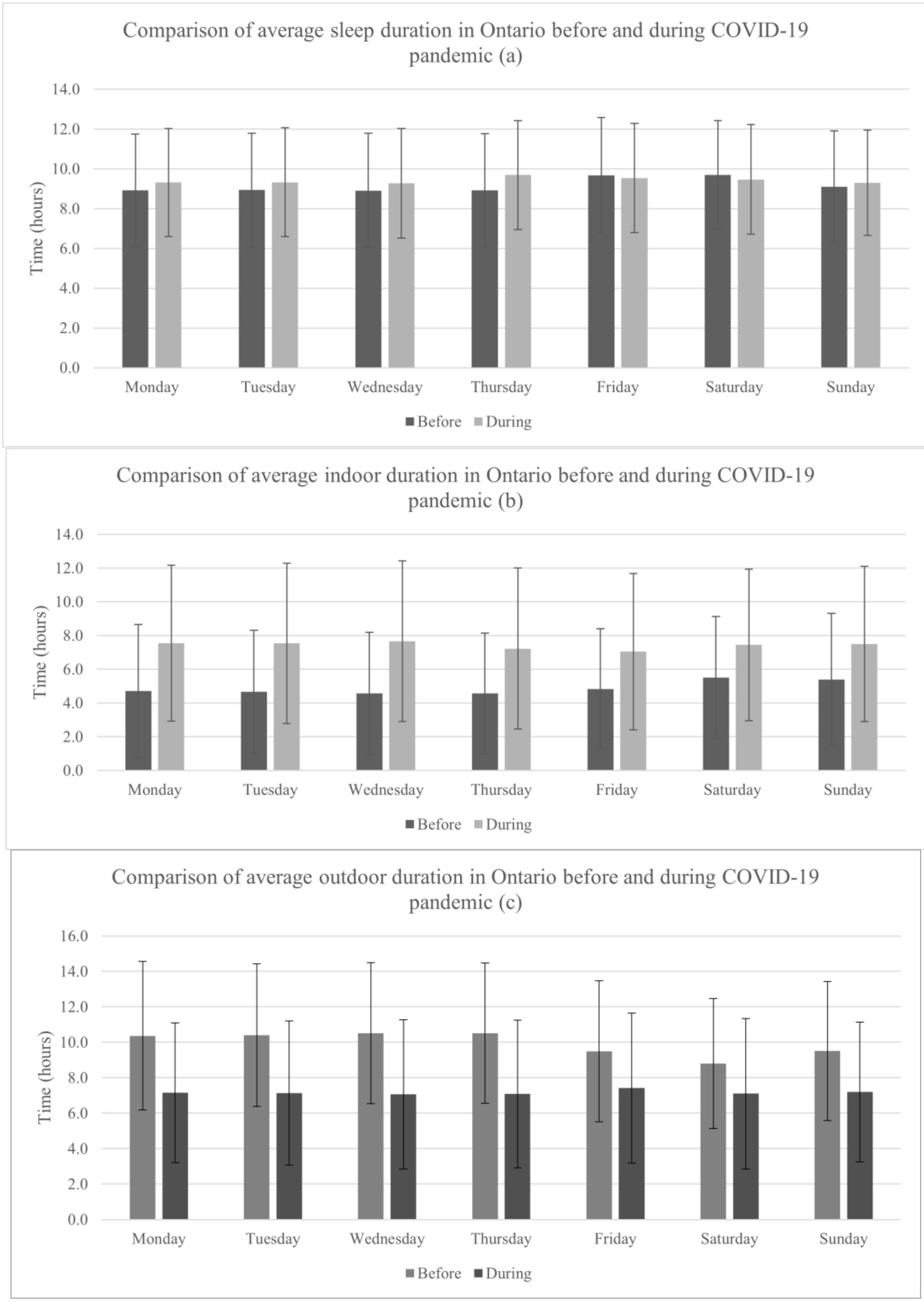


Figure 39. Comparison of average behavioural indicators a) sleep duration, b) in-house duration c) out-of-the-house duration for the province of Ontario before and during COVID-19 pandemic.

7.3.5.2 Quebec

The results from the analysis for households in Quebec show that there is a statistically significant difference in average sleep duration, time spent in-house and out-of-the-house before and during the COVID-19 pandemic (Tables 13 and 14).

The average sleep duration before the pandemic was 9.5 ± 2.75 hours (Mean ± SD) whereas during the pandemic it was 9.6±2.54 hours. The difference in average sleep duration was 6 minutes which was statistically significant at $P<.001$ level. The difference was more accentuated on the weekdays than on the weekends, as shown in Figure 40. When stratified by weekdays, before the pandemic, the difference of sleep duration was higher between a typical Thursday (9.1±2.77 hours) and Friday 10.1± 2.67 hours) whereas during the pandemic, the difference was larger for Wednesday (9.5±2.55 hours) and Thursday (9.8± 2.64 hours).

For time spent in-house, there is a statistically significant duration of 2.9 hours observed before and during the pandemic for Quebec. Before the pandemic, households in Quebec spent 5.2± 3.48 hours in-house whereas it increased to 8.1±4.47 hours during the pandemic. Before the pandemic, weekdays had less in-house time than Saturdays and Sundays within a week. For instance, people spent 4.8±3.61 hours in-house on Friday and 5.5±3.92 hours on Saturday. However, during the pandemic, the difference between the time spent in-house on weekdays was reduced.

There is a statistically significant duration of 2.9 hours for time spent out-of-the-houses before and during the pandemic. Before the pandemic, time spent out-of-the-houses on weekdays was nearly 10 hours. The out-of-the-houses duration on Friday reduced to 8.9 ±3.59 hours compared to 10.2 ±3.81 hours on Thursday. During the pandemic, the average out-of-the-house time difference between weekdays and weekends was reduced.

The increased time spent for the sleep duration was further explored. The time to go to bed and wake up was calculated. During the COVID-19 pandemic, the average difference for bedtime has minimally changed on the weekdays, whereas the average for Friday is 11:25 pm and Saturday is 11:18 pm, respectively.

These results show an overall change in Quebec of household behavioural patterns, sleep habits, time spent in-house, and time spent out-of-the-houses, which were likely caused by the policies implemented to curb the spread of COVID-19.

Table 13. Comparison of sleep indicators and in-house stay duration before and during the COVID-19 pandemic for Quebec.

	Before Pandemic			**During Pandemic**			**Mean Difference (During-Before)**		
	Sleep duration (hours) Mean±SD	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (hours) Mean±SD**	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (minutes)**	**In-house time (hours)**	**Out-of-the-house time (hours)**
Monday	9.2±2.69	4.9±3.34	9.8±3.67	9.4±2.67	8.2±4.57	6.3±4.36	12 ***	3.3 ***	-3.5 ***
Tuesday	9.1±2.68	4.9±3.32	9.9±3.74	9.5±2.54	8.2±4.55	6.3±4.05	24 ***	3.2 ***	-3.6 ***
Wednesday	9.2±2.73	4.8±3.19	10.0±3.67	9.5±2.55	8.3±4.58	6.2±4.13	18 ***	3.5 ***	-3.8 ***
Thursday	9.1±2.77	4.7±3.32	10.2±3.81	9.8±2.64	7.8±4.51	6.4±4.16	42 ***	3.1 ***	-3.8 ***
Friday	10.1±2.67	5.0±3.44	8.9±3.59	9.8±2.36	7.8±4.39	6.5±3.90	-18 ***	2.8 ***	-2.4 ***
Saturday	10.1±2.89	6.1±3.76	7.8±3.99	9.7±2.53	8.3±4.19	6.0±3.85	-24 ***	2.2 ***	-1.8 ***
Sunday	9.4±2.65	5.9±3.73	8.7±3.92	9.6±2.42	8.3±4.45	6.1±3.99	12 ***	2.4 ***	-2.6 ***
Total	9.5±2.75	5.2±3.48	9.4±3.85	9.6±2.54	8.1±4.47	6.3±4.0Six	6 ***	2.9 ***	-2.1 ***

*** Statistically significant at *P*<.001 level, SD-Standard deviation.

Table 14. Comparison of sleep timing before and during the COVID-19 pandemic for Quebec.

	Before Pandemic		**During Pandemic**	
	Bedtime (pm.) Mean±SE	**Wake-up time (am.) Mean±SE**	**Bedtime (pm.) Mean±SE**	**Wake-up time (am.) Mean±SE**
Monday	10:48:59 ± 22.14	06:05:08 ± 3.19	10:50:32 ± 24.51	07:29:05 ± 3.12
Tuesday	10:45:02 ± 22.43	06:18:15 ± 3.28	10:50:14 ± 25.0	07:22:44 ± 3.14
Wednesday	10:50:48 ± 21.53	05:58:54 ± 3.17	10:52:17 ± 24.53	07:27:23 ± 3.04
Thursday	10:53:16 ± 21.31	05:52:06 ± 3.27	10:56:42 ± 26.41	08:46:27 ± 3.07
Friday	11:25:34 ± 24.49	07:13:36 ± 2.53	11:12:41 ± 24.44	07:15:56 ± 3.03
Saturday	11:18:25 ± 26.37	08:46:55 ± 4.07	11:14:52 ± 24.12	06:57:43 ± 3.01
Sunday	10:53:36 ± 23.45	06:43:32 ± 3.23	10:54:15 ± 24.52	06:43:32 ± 2.52
Total	**10:58:58 ± 8.52**	**06:41:03 ± 1.19**	**10:58:41 ± 9.28**	**07:31:47 ± 1.10**

SE- Standard error.

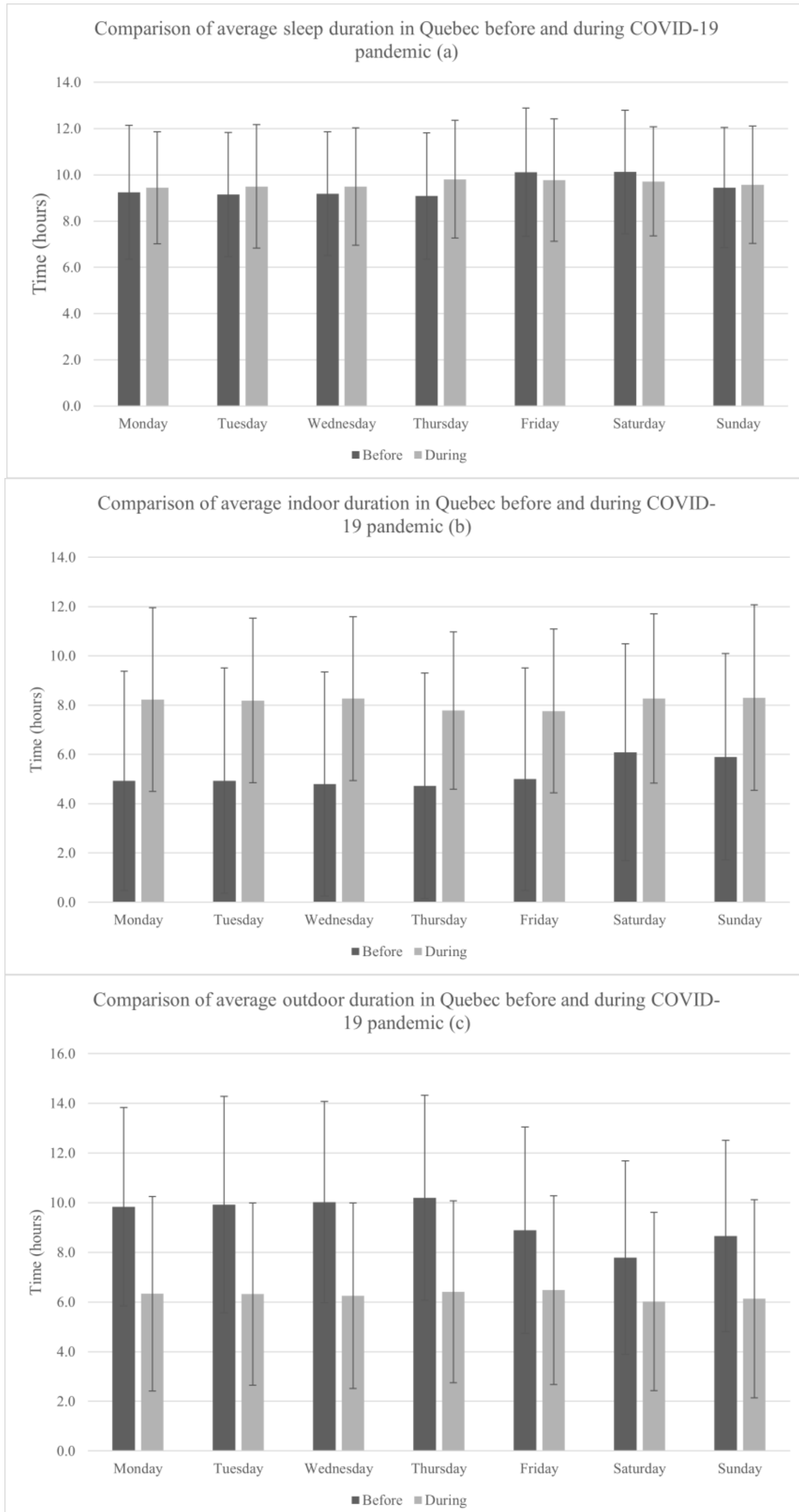


Figure 40. Comparison of average behavioural indicators a) sleep duration, b) in-house duration c) out-of-the-house duration for the province of Quebec before and during COVID-19 pandemic.

7.3.5.3 Alberta

The results from the analysis of households in Alberta show that there was a statistically significant difference in average sleep duration, time spent in-house and out-of-the-house before and during the COVID-19 pandemic (Tables 15 and 16).

The average sleep duration before the pandemic was 8.3 ± 2.57 hours (Mean ± SD) whereas during the pandemic it was 8.8±2.70 hours. The difference in average sleep duration was 30 minutes which was statistically significant at $P<.001$ level. The difference was more accentuated on the weekdays than on the weekends, as shown in Figure 41. When stratified by weekdays, before the pandemic, the difference of sleep duration was higher between a typical Thursday (8.3±48 hours) and Friday 8.5± 2.69 hours) whereas during the pandemic, the difference was higher on Wednesday (8.9±2.68 hours) and Thursday (9.1± 2.76 hours).

For time spent in-house, there is a statistically significant duration of 2.2 hours observed before and during the pandemic for Alberta. Before the pandemic, households in Alberta spent 5.3± 3.66 hours in-house whereas it increased to 6.8±4.55 hours during the pandemic. Before the pandemic, weekdays had less in-house time than Saturdays and Sundays within a week. For instance, people spent 5.3± 3.66 hours in-house on Friday and 6.8±4.55 hours on Saturday. However, during the pandemic time, the difference between the time spent in-house on weekdays was reduced.

There is a statistically significant duration of 2.7 hours for time spent out-of-the-houses before and during the pandemic. Before the pandemic, time spent out-of-the-houses on weekdays was 10.4±3.61 hours whereas the out-of-the-houses duration during the pandemic reduced to 8.3±4.16 hours. During the pandemic, the average out-of-the-house time difference between weekdays and weekends was reduced.

During the COVID-19 pandemic, the average difference in bedtime has minimally changed on weekdays, whereas the average for Friday is 11:27 pm and Saturday is 11:21 pm, respectively. These results show an overall change in Alberta in household behavioural patterns, sleep habits, time spent in-house, and time spent out-of-the-houses, which were likely caused by the policies implemented to curb the spread of COVID-19.

Table 15. Comparison of sleep indicators and in-house stay duration before and during the COVID-19 pandemic for Alberta.

	Before Pandemic			During Pandemic			Mean Difference (During-Before)		
	Sleep duration (hours) Mean±SD	**In-house time (hours)** Mean±SD	**Out-of-the-house time (hours)** Mean±SD	**Sleep duration (hours)** Mean±SD	**In-house time (hours)** Mean±SD	**Out-of-the-house time (hours)** Mean±SD	**Sleep duration (minutes)**	**In-house time (hours)**	**Out-of-the-house time (hours)**
Monday	8.3±2.53	5.3±3.67	10.4±3.61	8.8±2.64	7.0±4.62	8.2±4.20	30 ***	2.3 ***	-2.8 ***
Tuesday	8.3±2.50	5.0±3.51	10.7±3.51	8.8±2.64	6.9±4.60	8.3±4.25	30 ***	2.5 ***	-3.1 ***
Wednesday	8.3±2.47	4.7±3.32	11.0±3.42	8.9±2.68	7.0±4.64	8.1±4.17	30 ***	2.8 ***	-3.4 ***
Thursday	8.3±2.48	4.9±3.44	10.8±3.42	9.1±2.76	6.5±4.50	8.4±4.28	48 ***	2.2 ***	-3.1 ***
Friday	8.5±2.69	5.2±3.59	10.3±3.56	8.9±2.72	6.4±4.36	8.7±3.96	18 ***	1.9 ***	-2.1 ***
Saturday	8.4±2.73	5.9±3.89	9.7±3.73	8.9±2.71	6.9±4.52	8.1±3.94	18 ***	1.6 ***	-1.9 ***
Sunday	8.2±2.55	5.9±3.97	9.9±3.82	8.8±2.75	6.9±4.56	8.3±4.26	30 ***	1.7 ***	-2.2 ***
Total	8.3±2.57	5.3±3.66	10.4±3.61	8.8±2.70	6.8±4.55	8.3±4.16	30 ***	2.2 ***	-2.7 ***

*** Statistically significant at $P<.001$ level, SD-Standard deviation.

Table 16. Comparison of sleep timing before and during the COVID-19 pandemic for Alberta.

	Before Pandemic		**During Pandemic**	
	Bedtime (pm.) Mean±SE	**Wake-up time (am.) Mean±SE**	**Bedtime (pm.) Mean±SE**	**Wake-up time (am.) Mean±SE**
Monday	11:17:01 ± 4.47	05:14:23 ± 1.09	11:25:46 ± 8.50	06:03:23 ± 1.41
Tuesday	11:16:02 ± 4.46	05:07:13 ± 1.08	11:20:36 ± 8.35	05:40:11 ± 1.43
Wednesday	11:19:56 ± 4.40	05:05:51 ± 1.05	11:22:40 ± 8.46	06:01:34 ± 1.38
Thursday	11:19:01 ± 4.47	05:15:57 ± 1.05	11:18:58 ± 9.57	07:03:38 ± 1.45
Friday	11:27:22 ± 5.20	05:51:19 ± 1.09	11:28:35 ± 8.55	06:26:33 ± 1.34
Saturday	11:21:53 ± 5.21	05:55:23 ± 1.18	11:27:05 ± 8.54	06:21:13 ± 1.37
Sunday	11:14:27 ± 4.42	05:12:34 ± 1.14	11:17:54 ± 8.43	06:12:53 ± 1.42
Total	**11:19:20 ± 1.51**	**05:23:04 ± 0.26**	**11:23:04 ± 3.23**	**06:15:31 ± 0.38**

SE- Standard error.

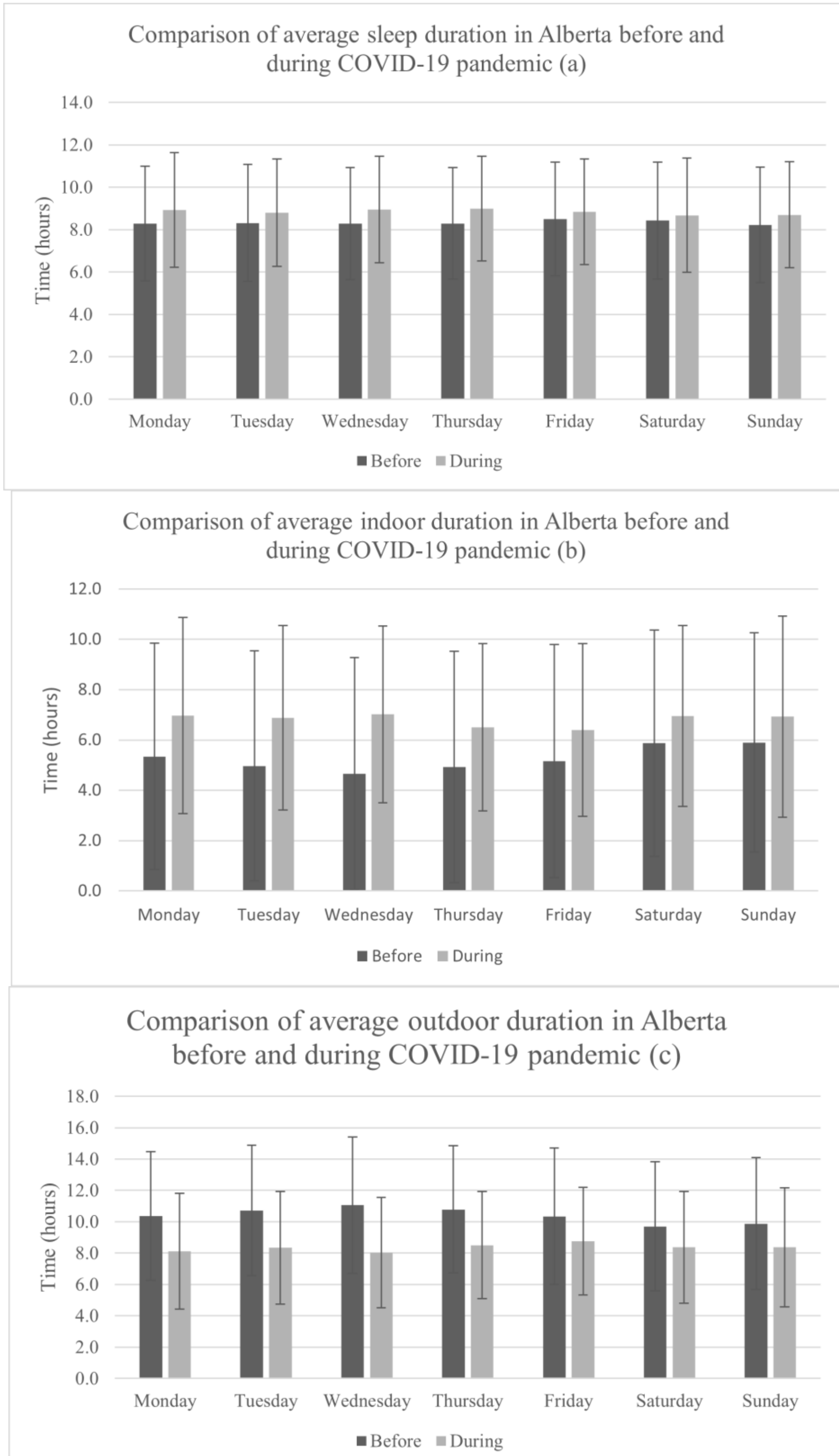


Figure 41. Comparison of average behavioural indicators a) sleep duration, b) in-house duration c) out-of-the-house duration for the province of Alberta before and during COVID-19 pandemic.

7.3.5.4 British Columbia

The results from the analysis for households in British Columbia show that there is a statistically significant difference in average sleep duration, time spent in-house and out-of-the-house before and during the COVID-19 pandemic (Tables 17 and 18).

The average sleep duration before and during the pandemic were remained same as 7.8 ± 2.47 hours (Mean ± SD). When stratified by weekdays, before the pandemic, the difference of sleep duration was higher between a typical Thursday by 18 minutes whereas on Sunday the change in average sleep was -12 minutes.

For time spent in-house, there is a statistically significant duration of 2.2 hours observed before and during the pandemic for British Columbia. Before the pandemic, households in British Columbia spent 5.8± 4.21 hours in-house whereas it increased to 8.0±4.77 hours during the pandemic. Before the pandemic, weekdays had less in-house time than Saturdays and Sundays within a week. For instance, people spent 5.9±4.28 hours in-house on Friday and 6.1±4.31 hours on Saturday. However, during the pandemic, the difference between the time spent in-house on weekdays was reduced.

There is a statistically significant duration of 2.2 hours for time spent out-of-the-houses before and during the pandemic. Before the pandemic, time spent out-of-the-houses on weekdays was 10.4±3.73 hours, whereas during the pandemic the average out-of-the-house duration stands at 8.2±3.86 hours. During the pandemic, the average out-of-the-house time difference between weekdays and weekends was reduced.

The time to go to bed and wake up was calculated and compared for the days of the week. During the COVID-19 pandemic, the average difference for bedtime has minimally changed on the weekdays.

These results show an overall change in British Columbia in household behavioural patterns, sleep habits, time spent in-house, and time spent out-of-the-houses, which were likely caused by the policies implemented to curb the spread of COVID-19.

Table 17. Comparison of sleep indicators and in-house stay duration before and during the COVID-19 pandemic for British Columbia.

	Before Pandemic			**During Pandemic**			**Mean Difference (During-Before)**		
	Sleep duration (hours) Mean±SD	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (hours) Mean±SD**	**In-house time (hours) Mean±SD**	**Out-of-the-house time (hours) Mean±SD**	**Sleep duration (minutes)**	**In-house time (hours)**	**Out-of-the-house time (hours)**
Monday	7.7±2.35	5.6±4.13	10.7±3.67	7.8±2.44	8.2±4.94	8.0±3.87	6 ***	2.6 ***	-2.7 ***
Tuesday	7.7±2.35	5.7±4.11	10.6±3.63	7.7±2.46	8.2±4.81	8.0±3.91	0	2.5 ***	-2.6 ***
Wednesday	7.7±2.35	5.6±4.19	10.7±3.80	7.8±2.35	8.3±4.73	7.9±3.75	6 ***	2.7 ***	-2.6 ***
Thursday	7.7±2.24	5.7±4.18	10.6±3.60	8.0±2.45	7.7±4.87	8.3±4.06	18 ***	2.0 ***	-2.3 ***
Friday	7.8±2.61	5.9±4.28	10.3±3.84	7.8±2.55	7.4±4.63	8.8±3.85	0	1.5 ***	-1.5 ***
Saturday	7.8±2.72	6.1±4.31	10.0±3.76	7.8±2.47	8.0±4.68	8.2±3.64	0	1.6 ***	-1.9 ***
Sunday	8.0±2.64	6.0±4.29	10.1±3.76	7.8±2.49	8.0±4.70	8.1±3.89	-12 ***	2.0 ***	-2.0 ***
Total	7.8±2.47	5.8±4.21	10.4±3.73	7.8±2.46	8.0±4.77	8.2±3.86	0	2.2 ***	-2.2 ***

*** Statistically significant at *P*<.001 level, SD-Standard deviation.

Table 18. Comparison of sleep timing before and during the COVID-19 pandemic for British Columbia.

	Before Pandemic		**During Pandemic**	
	Bedtime (pm.) Mean±SE	**Wake-up time (am.) Mean±SE**	**Bedtime (pm.) Mean±SE**	**Wake-up time (am.) Mean±SE**
Monday	11:35:17 ± 15.09	05:04:41 ± 5.34	11:48:21 ± 15.23	05:05:14 ± 3.52
Tuesday	11:35:50 ± 14.37	05:03:18 ± 4.44	11:38:19 ± 16.30	05:26:26 ± 4.33
Wednesday	11:42:03 ± 16.33	05:30:10 ± 3.55	11:42:12 ± 15.32	05:09:07 ± 4.06
Thursday	11:39:45 ± 15.56	05:18:51 ± 4.13	11:44:24 ± 17.06	05:13:21 ± 3.43
Friday	11:32:48 ± 17.16	05:18:05 ± 5.10	11:43:41 ± 17.13	05:34:36 ± 3.27
Saturday	11:24:58 ± 20.45	06:02:38 ± 6.49	11:43:05 ± 16.03	05:13:01 ± 3.52
Sunday	11:23:38 ± 19.25	05:39:00 ± 6.16	11:50:04 ± 17.12	05:37:07 ± 3.46
Total	**11:33:32 ± 6.30**	**05:25:06 ± 2.00**	**11:44:19 ± 6.12**	**05:19:50 ± 1.28**

SE- Standard error.

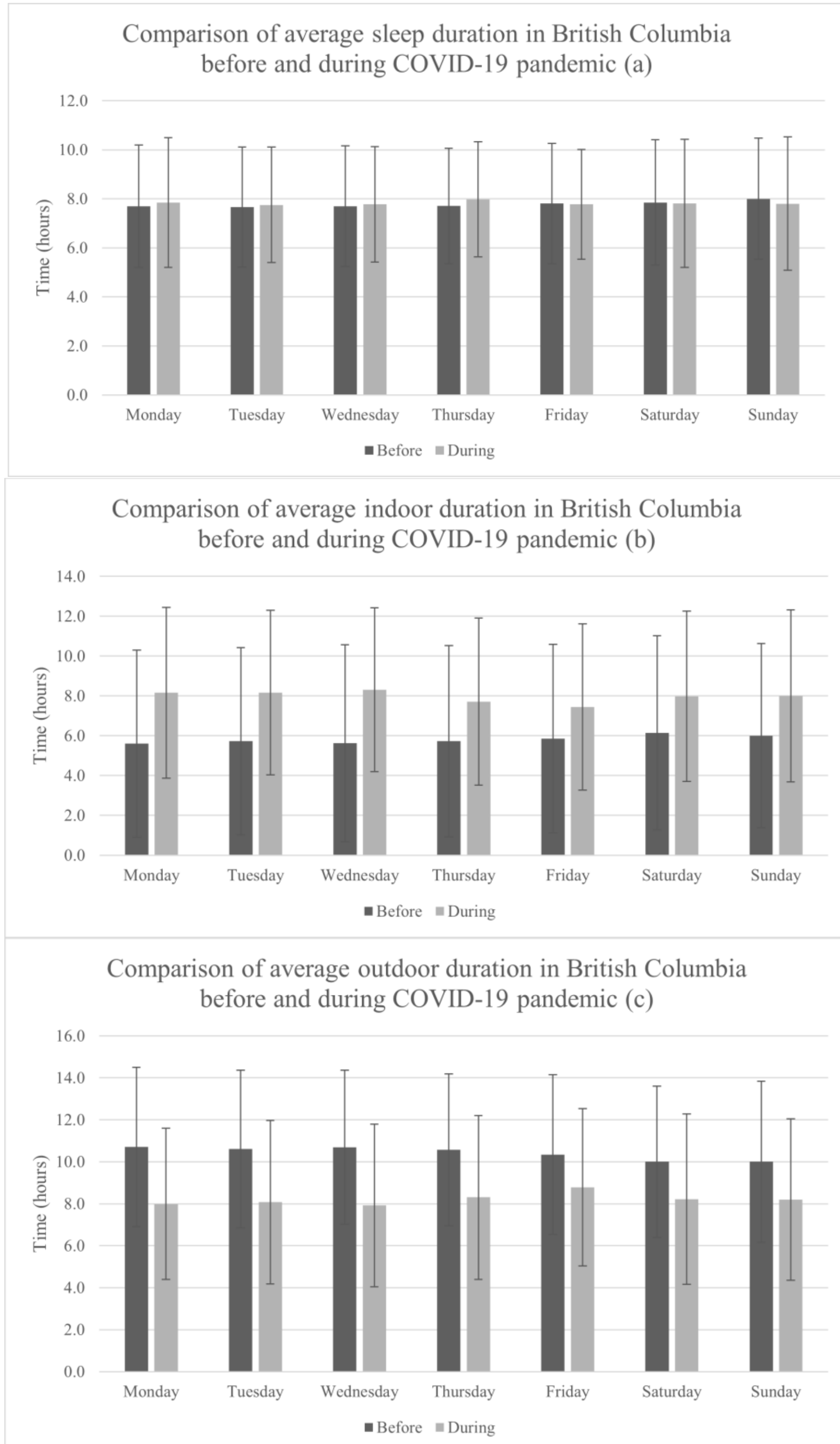


Figure 42. Comparison of average behavioural indicators a) sleep duration, b) in-house duration c) out-of-the-house duration for the province of British Columbia before and during COVID-19 pandemic.

7.4 Discussion

The findings from this study unravel the changes in behaviour at population level across four provinces of Canada before and during the COVID-19 pandemic. The stay-at-home order and work-from-home policies required people to stay in-house for extended periods compared to pre-COVID-19 period. IoT data supports intelligent monitoring of policy-related changes, which can be quickly conducted in a short period requiring less resources. This complete data analysis was performed by one researcher with the support of a software developer, which included the analysis of data from 12,252 households across Canada.

Several national sleep foundations recommend that the amount of sleep for adults should range between 7-9 hours [395]. Our results show that Ontario, Quebec, and Alberta are better positioned than British Columbia for this parameter. During COVID-19, average sleep duration has increased in Canadian provinces, which is accordance to the findings from other countries like Spain [378], the United States [20] and Singapore [381]. Similar to other studies, this study also shows the difference in sleep duration, and sleep patterns vary between weekdays and weekends [381]. Although this study measures important sleep parameters at population level, there remains other unexplored features of healthy sleep to be monitored, like sleep quality. A study completed in 2020 in Canada using a questionnaire-based survey, shows that sleep duration minimally reduced in the initial phase of the pandemic and that the quality of sleep deteriorated [373].

The result of my study shows that the wake-up time changed by around 30 minutes later, and the finding is similar to another study done in 2021 by Robillard et al.[373]. Changes in physical activity, sedentary behaviour and social behaviour occurred soon after the COVID-19 pandemic were declared, and some of these changes differed among those with low and high anxiety [396].

Similarly, another study in Italy in 2021 shows that sleep quality dropped during the COVID-19 pandemic [374]. These changes in the sleep cycle impact the physical and mental health of individuals [397]. Therefore, sleep is a major risk factor for multiple chronic diseases, and lack of sleep has both long-and short-term consequences such as premature mortality [383,397].

Along with sleep duration, the time spent out-of-the-houses is also a critical factor to maintain good health [398,399]. The time spent out-of-the-houses is not only related to the sleep quality and duration [400] but also stimulates the activation of vitamin-D amount in our body, deficiency of which is a potential risk factor for bone and joint health [401,402]. On the other hand, less out-of-the-house time can be related to increased in-house time. More in-house time leads to increase in sedentary behaviour which has several consequences in different age groups [74,379,403]. For example, in children, risk of multiple childhood chronic conditions arises such as childhood obesity, asthma, attention deficit hyperactivity disorders (ADHD). This can lead to long term health effects like, pulmonary, cardiovascular, and mental health problems in adulthood [398]. Richard Louv in 2005 coined the term "nature deficit disorder" wherein he delineates that the time of children spent out-of-the-houses is replaced by electronic media and demanding school schedules [398]. Additionally, the effect of increased sedentary behaviour has negative physical and mental health consequences for adults and aging populations [379,404]. With the prevalence of low behavioural public health indicators, *i.e.*, time spent out-of-the-houses and sleep duration [405], situations like COVID-19 pandemic jeopardize a healthy lifestyle.

In Canada, the prevalence of chronic disease is increasing [64,125,406] and a careful assessment of risk factors of various chronic diseases is essential for predicting the future trends [64]. Recent studies highlight the variability in data collection methods ranging from traditional survey-based methods, wearable-based methods to actigraphy [178,407,408]. Reports suggest a range of applications and technology systems for smart infrastructures and developing 'pandemic-proof' smart communities [409]. Moreover, a cloud-enabled pervasive IoT implementation framework can be adapted and expanded alongside scientific solutions [409]. The use of consumer-based activity trackers and smartphone-based passive sensing [410] has gained increased recognition as a tool for monitoring physical activity and behavioural indicators in free-life conditions in epidemiological studies during the COVID-19 pandemic [411].

The COVID-19 pandemic has impacted society at various levels. Early identification of the effects of changes caused by this pandemic in real-time may help us save lives and reduce the duration of required social isolation. With more people being restricted to stay within their houses and everything being transitioned to virtual platforms, these changes directly, or indirectly affected mental health and sleep patterns. Notably, our study is unique due to the granularity of data used, presence of a large sample size, simplicity of data collection, and flexibility of the dataset.

This new data collection method enables the process to be continuous and allows increased access to households and participants, contributing to the specific condition or data being monitored. With minimal effort, we can answer relevant population health related questions.

In terms of public health surveillance, the near real-time monitoring of these health risk behaviours can help monitor risk factors and build preventive strategies to reduce the impact that precede long-term complications from chronic diseases.

This will also bolster the foundations to build systems to monitor health indicators in near real-time for the critical risk factors for chronic diseases such as diabetes, hypertension, obesity, mental health problems, and cancer[28]. In the long run, it will aid in monitoring and reducing the rising illness burden by quickly measuring the impact of any policy-related changes at the

population level. Leveraging such technology will significantly reduce user burden while enabling remote monitoring, data collection, thus, reducing the research associated costs.
The study's shortcoming was that no adjustment for multiple testing was done in the analysis, and there's always the possibility that some parameter significance is due to random sampling and multiple testing, thus more research is needed to see if these findings are generalizable and replicable.

7.5 Conclusion

Measurement of behavioural risk factors is an essential domain of public health. Pandemics like the COVID-19 have an impact on behavioural indicators, including sleep. Data from IoT technologies have the potential to answer some of the critical questions related to the assessment of policy changes at population level. This study shows the application of smart thermostat-based data, collected from a large number of residents in Canada, to measure sleep duration, sleep time, wake-up time, time spent in and out of the house, for different provinces, during and before the COVID-19 pandemic.

7.6 Innovation

This research is innovative for its (i) population-level comparison of household behaviours, (ii) using novel data source(s) for monitoring healthy behaviours, (iii) measuring new indicators, (iv) using novel tools/solutions to capture data. The methods proposed here will enable access to much larger sample size and increase the generalizability of the results.

Chapter 8 IoT-based Mobility Analysis During a Pandemic

8.1 Introduction

8.1.1 Background

Human mobility data can prove to be an effective tool to unravel and comprehend the complexities of the world and its communities [183], as for example, providing insights to identify the difference in human behaviour [381] and spread of COVID-19 cases along with the pattern of mortality [412,413]. The collection, extraction, and timely analysis of human mobility data can deepen our understanding in health and other different domains. Human mobility data can be classified and obtained at individual [414,415] and population-level [416]. Strikingly, population-level human mobility studies hold the potential to identify several public health and social domain-related issues. The domain of research covers rapid urbanization across the globe, change in lifestyle factors (for example, less physical activity, more sedentary time, and change in sleep pattern), spreading of infectious diseases, the impact of air pollution, effect of climate change on the migration patterns of the population across continents or countries.

The first human mobility studies were completed in the United States of America in 2006, based on data from cell phone calls and proximity from the cellphone towers [417]. Similarly, another research group utilized crowdsourcing techniques to collect population-level human mobility data [182]. In the contemporary world, numerous technological advancements have changed the way data is collected. Currently, there are several methods to gather population-level human mobility data at the global, national, provincial, city, or zip code level granularity. Figure 43 shows technological innovations in tracking human and animal mobilities from the year 1900 to 2020 [186]. In the early 20th century, primitive technologies such as bird markings and small radio transmitters were used to track animal mobility. However, with technological advancements in the early 21st century, communication devices such as smartphones and wearables served as easy means of collecting mobility data. With the passage of time, a sharp increase in the number of such devices led to high volumes of data generation, thus contributing to big data. This will, in turn, help to identify and validate human mobility behaviours. Different parameters such as geolocation and distance travelled can help us classify the type of human mobility data. However, to date, human mobility data has not been much utilized by the public health domain.

Figure 43. Journey of human and animal tracking technologies with a timeline: image extracted from Meekan et al. 2017 [186].

Since the World Health Organization (WHO) declared COVID-19 as an international public health emergency in late January 2020 and subsequently declared it as a pandemic on March 11, 2020, many countries attempted to curb its spread. [418]. Various interventions like social distancing, self-isolation, and personal protective equipment as face masks and face shields have been widely used to limit the spread of coronavirus at community level and protect vulnerable groups [419,420]. In the early stages of the pandemic, the decrease recorded in people's mobility in various digital data sources was closely related to the reduction in the incidence of COVID-19 [412,414,419,421–426]. Most countries sought to implement strict social distancing norms to avoid an overwhelming burden on the healthcare institutions.

There is evidence that automatic movement measures are closely related to the spread of the COVID-19 virus in various countries over time [419]. Considering COVID-19 pandemic as a case study to understand the use of human mobility data in public health, it can be inferred that such data sources are essential and demand further exploration. As highly granular real-time data becomes more readily available, future epidemiological analyses can be supported by such data. Since such population-level vital indicators can be obtained in real-time, mobility data may become an essential forecasting tool. Research involving human mobility tracking data has benefited from the technological advancements such as smartphones and wearable technology (smartwatches and activity trackers) [427]. These innovations have produced a wealth of easily accessible human activity-related health data such as heart rate, time spent in exercise, sleep, and mobile georeferenced data. The use of mobile phones, satellite data or information of flight traffic is used for macro measurement of human mobility or as a proxy [188,198]. These large data sets, equivalent to "big data" in volumes, are now being analyzed to describe human movement patterns, characteristics (such as sleep, stress, and activity), as well as interactions with levels of detail, immediacy, and unprecedented precision. Furthermore, this type of research is first of its kind to characterize human movement patterns on a global scale and can be attributed to big data analysis.

Data obtained from various sources such as smartphones with GPS tracking technologies, texts or photos via photo-sharing platforms like Twitter and Flickr, geolocation-enabled internet posts, public transport cards, and credit card transactions now provide direct access to human location, trajectories, opinions, and interactions [428]. These high-resolution datasets allow researchers to develop and verify models of human movement on different spatial scales. Although the domain of human movement research is relatively new compared to research on animal movement, it has significantly gained the attention of researchers as evident from the large number of publications [427].

8.1.1.1 Population and Facts About Google Map users and ecobee smart thermostat users

Google Map users: Google Maps is the most popular navigation app in the US and Canada. The app surpassed 23 million downloads in 2020. Google Maps has 154.4 million monthly users. In 2018, 67% of all mapping app users relied on Google Maps. Google Maps users contribute 20+ million pieces of information per day. 5 million unique miles of road in Street View, in 3,000+ cities across 45 countries. 20+ petabytes of aerial and Street View imagery combined--the equivalent of 266 years of HD video. Live traffic data in more than 50 countries and more than 600 major cities. Driving directions for 194 countries, spanning 27.9 million miles of road. Schedules for more than one million public transit stops worldwide. Local information for more than 80 million places around the world. There are more than one billion monthly active users of Google Maps services. More than 50% of global Google Maps usage is mobile.

Ecobee users: More than 90% of Canadian households have thermostats, with the majority opting for programmable thermostats. Smart thermostats, often known as IoT devices, are a type of programmable thermostat that may be connected to the internet. The ecobee comes in second place after Nest in terms of smart thermostat market share in Canada. ecobee has a program called DYD, which has a million users, hence the sample size is roughly 172,000+ households across Canada.

Distribution of the Android based smart phone across Canada is widespread and hence the Google map users. When compared to Google, ecobee users are more likely to be younger, tech-savvy individuals and homes with high-speed internet access.

8.1.2 Population selection

This study also analyzed data from the four selected provinces: Ontario, Alberta, Quebec, and British Columbia. These four provinces constitute approximately 86% of the Canadian population, as per the Statistics Canada report published in 2021 [394].

8.1.3 Google mobility report

As global scientific communities including public health officials of various countries respond to COVID-19 pandemic, there is still an ongoing interest to analyze population-level mobility data. With a massive user network of Google Maps, Google has furnished systematic and anonymized information, owing to its significance in making critical decisions to fight COVID-19. The reports show movement trends over time by geographic area across different categories such as retail and recreation, grocery and drug stores, parks, transit stations, workplaces and residences [429]. The information has been passively generated, collected, and now is being made available to researchers and policymakers through Google's open source: 'COVID-19 Community Mobility Reports' [429].

8.1.4 ecobee Donate your Data

In-house mobility has never been captured on a population-level scale before. Herein, I developed a framework to infer the relationship between micro and macro mobilities captured by two distinct data sources at population level. The rationale of this study is to explore, understand and use IoT-based mobility data from ecobee smart thermostats of Canadian households.

This study aimed to find the association between the gold standard mobility data from "Google Mobility Reports" and the Smart thermostat-based data from ecobee. Reportedly, Google's published the data for mobility across the globe has geolocation information at multiple levels of granularity such as, country, province, and region (for example Region of Waterloo which includes three major cities namely Waterloo, Kitchener, and Cambridge), while ecobee's "Donate your Data" program has the geolocation information available with a different level of granularity such as country, province, and city at most granular level (for example city of Waterloo).

8.1.5 Objectives

To assess the suitability of IoT-based ecobee smart thermostat data compared to Google mobility data from selected four provinces of Canada for population level mobility. Furthermore, we explored the seasonality diagnostics patterns in terms of day-by-day, week-by-week, and month-by-month analysis from both the datasets. Next, we sought to find the anomaly detection capacity of both datasets.

8.2 Methods

This study used Google's residential mobility data and ecobee's mobility data curated from "Donate your Data" for one year (February 15, 2020, to February 14, 2021) to find out the association between them. The date range is selected based on the Google's Mobility Report publication dates and I explored one year of data.

Ethics approval for this study was obtained from the University of Waterloo Office of Research Ethics (#31377).

8.2.1 Data Analysis Platforms and Software

ecobee's mobility data preparation from "Donate your Data" was done using Azure Databricks and Jupyter notebook using python [430]. Azure Databricks is a data analytics platform optimized for Microsoft Azure cloud services platform [431] whereas Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. The use of Jupyter notebook includes data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and complex statistical analysis [432]. Data cleaning, analysis and visualization was done in R studio [433] version 1.4.1106 with R software[434], version 4.0.5 and data analysis packages Tidyverse [435] and timetk [436].

8.2.2 Data Sources and Preparation

8.2.2.1 Google Mobility Report

The Google Mobility data [429] is a unique dataset at global level to quantify human mobility. The process of collection of this data initiates from the users accessing their Google accounts in different continents. This utilizes satellite-dependent geolocation service, popularly known as Google Maps.

This service helps the community travel, calculate time to commute and find alternative pathways. During the COVID-19 pandemic, Google released the data for the first time to help researchers guide and measure mobility-related changes. Accuracy of the location and categorization of places varies from region to region. These data points record the total number of visits to specific destinations visited by individuals, as described in Table 18. Google manages these aggregated and anonymized records of users who have turned on the location history setting of Google accounts on their phones and agreed to share this information. These datasets, therefore, may not be representative of the entire population. Additionally, Google has not yet effectively disclosed its precise calculations of these mobility data points publicly. Daily values are compiled across individuals who have enabled their location history and are available for each province in Canada from February 15, 2020 onwards [429].

8.2.2.1.1 Definition of categories of type of place

Table 19. Google mobility data categories and their description as described on the website [429].

Category	**Definition**
"Grocery & pharmacy"	*"Mobility trends for places like grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies."*
"Parks"	*"Mobility trends for places like local parks, national parks, public beaches, marinas, dog parks, plazas, and public gardens."*
"Transit stations"	*"Mobility trends for places like public transport hubs such as subway, bus, and train stations"*
"Retail & recreation"	*"Mobility trends for places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theatres."*
"Residential"	*"Mobility trends for places of residence."*
"Workplaces"	*"Mobility trends for places of work."*

Out of these six categories, we lay emphasis on residential data for this study as ecobee data comes from the household environment and is categorized under the residential category itself.

8.2.2.1.2 Google Data Preparation Protocol as a Reference

Google Maps use aggregated and anonymized data to indicate frequently visited hours of a location to calculate changes in mobility [429]. Daily changes are compared to the baseline value for that particular day of the week. The baseline is the median of the corresponding days in five weeks from January 3 to February 6, 2020 [429]. The data set is trending over several months, with the latest data being represented by the data collected about 2-3 days before the current day because it takes between two to three days to create the data set. The data included in the calculation depends on preferences, connectivity and whether the users meet privacy thresholds. If data quality and privacy thresholds are not met, that data point is ignored. The Google Mobility report has data useful to measure social distancing efforts and access to essential services. The data for Google to calculate these insights is based on data from only the users who have opted for location history in their Google Accounts which makes up their sample. As with all samples, this may or may not represent the correct behaviour of a wider population. Each value of one day is the rate of change in the social mobility category relative to the baseline, showing how the length of visits and stays to various destinations have changed since the pandemic began. However, Google has not released social mobility data from earlier years. A visual inspection is helpful to evaluate whether trends in social mobility correspond with intuition [429].

8.2.2.2 ecobee Mobility Data

ecobee provides researchers with anonymized data from 179,000 households for a period of nearly five years (June 2016 to July 2021). The dataset includes a metadata table with characteristics of the house and Heating, ventilation, and air conditioning system, along with sensor data on a separate table.

8.2.2.2.1 Data Wrangling Process of ecobee Mobility Data

The Donate your Data dataset supplied by ecobee is hosted in Google BigQuery. As our data analytics platform is in Azure Databricks, the first step was to transfer the data from Google’s

Big Query to the Microsoft Azure platform. Once data transfer was completed, the azure storage blob was mounted to the data bricks account. Using pyspark, I loaded the complete thermostat dataset and metadata of size nearly seven terabytes to my environment. I created four subsets of the households from four provinces of Canada. Each of the subsets has Identifier, DateTime and lists of sensors activation. The granularity of the data was five minutes. As the data's time zone did not match geolocation, this was cleaned as described below in detail.

8.2.2.2.1.1 Time Zone Conversion

As the timestamp of the DYD dataset was in UTC format, I converted the time zones in the dataset by locating time zone information from the geolocation of the households in the metadata. Canada and its provinces have different time zones, and Figure 44 describes the distribution of time zones across Canada [437].

I worked on the time zone conversion process with two different approaches. I started with a Python-based automated time zone finder using location information. Though we could extract the time zone name from the location, unfortunately, the output quality was not satisfactory, leading to manual correction of the time zones.

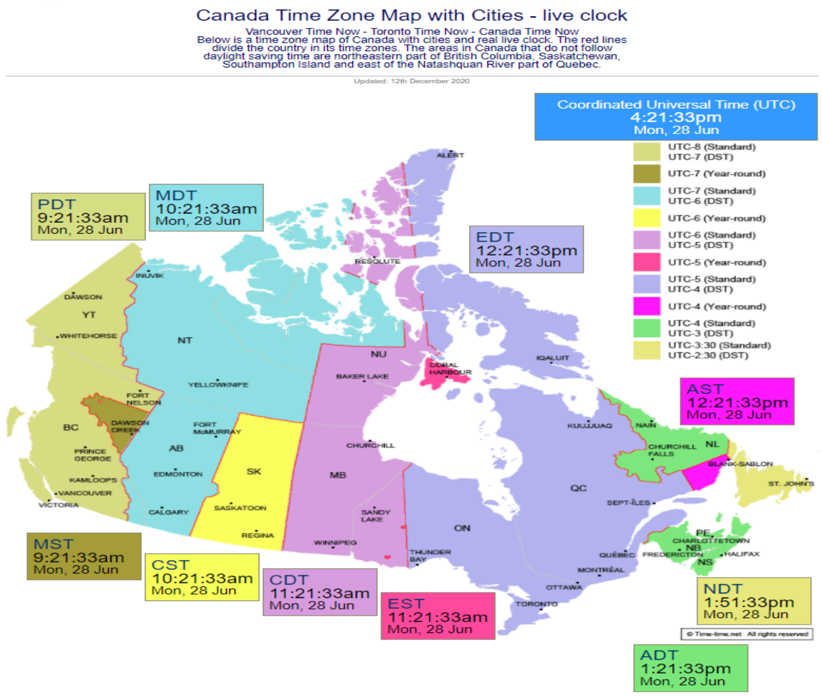


Figure 44. Time zone map of Canada and its provinces: image extracted from time-time.net. [437]

In contrast to the automated time zone extraction technique, the manual method was more efficient. The findings were, in the province of Ontario, six cities have different time zone than EST; namely, Drayton, Kenora, Kenora-Unorganized, Mitchell, Red Lake, Sioux Lookout cities (1% of the DYD dataset for Ontario) have central time zone and the remaining cities fall in the eastern time. The same exercise was repeated for Quebec, and it was found that all the cities belong to . For Alberta, all the cities have the same time zone, that is, mountain standard time, and for British Columbia, 28 households were from different time zones and were excluded from the analysis. Once cleaned for the time zone, the Date Time had been changed. Subsequently, time-series data analysis was performed on the adjusted data on the households included in the study. These numbers are presented in Table 20.

Table 20. Number of households selected for the analysis by province.

	N, before time zone cleaning	N After time zone cleaning	Excluded no of households for time zone diff	The proportion of household data by province (%)
Canada	21690	21690		
Ontario	7145	7134	11	33
Alberta	3989	3989	0	18
British Columbia	449	421	28	3
Quebec	708	708	0	2

8.2.2.2.1.2 Mobility Data Preparation from ecobee Sensors

Using the ecobee mobility data, a baseline value was created as per the Google data preparation method to make the data comparable. The sum of all sensor statuses was calculated to find the total number of sensors activated within each timestamp. Aggregation of daily movement at the province-level has been calculated by averaging the total number of sensors activated within that province for 24 hours. A table was created containing dates and the average number of sensors activated by date, which was saved as a CSV file. R studio and R platform were used for further data analysis.

8.2.2.3 COVID-19 Policy-Related Timeline and Data

The Canadian Government has a dedicated platform to share COVID-19 related information regularly[438]. I have extracted the data for Canada for the period and, as per requirement, selected specific indicators such as the daily number of confirmed COVID-19 cases across Canada, numbers of deaths and timeline of policy changes for Canada and its provinces. To understand the timeline of events in Canada due to the COVID-19 pandemic, I have also extracted the dates and policy level changes from the government platform, as described in the Appendix-6.

8.2.2.3.1 Timeline of COVID-19 for Canada and Provinces

Several policy-level decisions to curb the spread of the COVID-19 pandemic were implemented in Canada and its provinces. As health is considered a provincial matter in Canada, the provincial government can independently implement policies [439], while the federal government retains some control over the nation-wide policy implementation in case of public health emergencies [440]. Therefore, the policy implementation timelines for federal and provincial levels were monitored. Policies related to case management includes contact tracing and self-isolation, closure and openings of academic organizations, non-essential services, and recreation facilities. The Government of Canada recommended the implementation of work from home on March 10, 2020. Similarly, other policy-level decisions like implementing telemedicine or remote patient monitoring; policy related to the health workforce; mandatory mask use; an emergency declaration based on active cases, R-value and death counts; travel restrictions and mass vaccination strategies [438].

In addition to the abovementioned policies, some provinces implemented health education through health promotion campaigns to improve awareness and understanding of COVID-19. Temporary closure of non-essential health services, delay in elective medical procedures, restricting visitors' access to the health facilities, restricting the number of people in social gatherings, implementing pandemic response plans, phase and alert level changes, and the purchase and distribution of vaccines, were among the significant decisions taken to fight COVID-19[438].

8.2.3 Data Analysis Methodology

8.2.3.1 Visual Analysis

Once data from both the sources was ready for all four provinces, a visual inspection was carried out to determine the pattern and association. The time series plot was generated using Google residential and ecobee mobility data for each province. The statistical significance of the association was found using Pearson's correlation and Spearman's correlation [441]. Pearson's correlation measures how two continuous signals co-exist over time and sets the linear relationship from 0 (uncorrelated) to 1 (fully correlated) with a linear relationship of -1, indicating a negative correlation [441]. When using Pearson's correlation, there are two things to consider: a) presence of outliers can distort the correlation estimation results, and b) the data distribution should be uniform across the data range [441]. In general, correlation is a measurement of global synchronization [441]. Correlation does not provide information about the direction between the two signals such as which signal follows the other [442].

8.2.3.2 Seasonal Analysis

When looking at time-series data, seasonality is the fluctuation that occur at specific regular intervals of time which is always less than a year, such as weekly, monthly, and/or quarterly [443]. For some time-series data, it is observed that the series has a seasonal effect, and it is easy to find the season period (e.g., 4 for quarterly data, 12 for monthly data). Seasonality can be visually identified by a pattern that repeats all *k* elements in the series. Seasonality can be caused by various factors, including weather, vacations and holidays, and regular repetition of a specific event at the chronological level, mainly composed of rule-predictable patterns [443]. To determine the seasonality of the data, we ran a seasonal diagnostic test using the timetk [436] package in R, which was specially designed to analyze time-series data. Separate initiatives were implemented to understand the daily, weekly, and monthly seasonality.

One-way ANOVA was performed to check for statistical significance between various season categories [444]. The “timetk” package uses a unique feature called anomaly detection and plotting to find anomalies in data [436]. Visual inspection of these anomalous dates can show the impact of each country’s actual policy change dates and explore the significance of this kind of analysis in the future.

8.2.3.3 Anomaly Diagnostics

Anomaly Diagnostics is a wrapper around group anomaly detection visualization, implementing a two-step process for detecting outliers in a time series [436]. It helps automate the collection of features for time series seasonal analysis. Internal calculations are performed to discover the subrange of features to incorporate the logic: the minimum feature is selected based on the median difference from the continuous timestamp.

8.2.3.3.1 Seasonal and trend decomposition using Loess (STL)

The first step is to conduct a seasonal removal using seasonal and trend decomposition using Loess (STL) method [436,445,446]. There are three components of the time series data, seasonality, trend, and the remainder. The decomposition separates the seasonality and trend components from the observed values and places the remainder to detect anomalies. The user can control two parameters, frequency, and trend. Frequency adjusts the seasonality component removed from the observed value. Trends adjust which graph views are used. Both user frequency and the trend can be specified as a time-based period (e.g., 6 weeks) or a number (e.g., 180) or "automatic." This is used to predetermine the frequency and trend according to the scale of the time series [436].

8.2.3.3.2 Anomaly Detection

When trend and seasonality components are removed, anomaly detection is performed on the remaining data. The anomaly detection method uses inner quartile range (IQR). It takes a distribution and uses the 25% and 75% inner quartile range to establish the distribution of the remainder. Boundaries are determined by default to a factor of 3X above and below the inner quartile range, and any remainders beyond the limits are considered anomalies. The alpha parameter adjusts the 3X factor. Default value for the alpha stands at 0.05. Decrease of the alpha value increases the IQR factor, that controls the limit. Increasing the alpha will make it easier to find outliers. The IQR outlier detection method is used for prediction [436].

Twitter's anomaly detection package [447] used the same outlier detection method. Both Twitter and Forecast have a "time series outlier" method implemented in Business Science's "anomalize" package [448].

8.3 Results

The data from Google mobility dataset and ecobee "Donate Your Data" program for four Canadian provinces was analyzed and the results have been described below.

Within Google mobility data, mobility for grocery and pharmacy increased over the sample period compared to retail and recreational locations. However, from August 2020 onwards, a decline in retail and social mobility was observed. Apparently, the mobility movements for work and transit are significantly correlated, with both of them increasing over time. The highest data value in both variables were observed during the weekends and these values did not decline significantly when compared to pre-pandemic values. The sharp rise and fall of social mobility at parks refer to summertime activities. Initially, residential mobility was high which then declined during the latter part of the sampling period. As ecobee data come from households, it is logical to compare it with residential mobility data. Future analysis will be performed by comparing ecobee mobility data to the residential portion of the Google mobility data.

8.3.1 Association between Google and ecobee mobility data

The association between Google residential mobility data with ecobee mobility was studied across four Canadian provinces. A positive association was evident from the visual inspection between Google mobility data for the residential area and ecobee mobility data for all four provinces. Figure 45-48 shows the mobility trend from February 2020 to March 2021 for Ontario, Alberta, Quebec, and British Columbia.

8.3.1.1 Ontario

Figure 45 shows the trends and patterns from raw data of both Google residential mobility and ecobee mobility for one year. The repeated cyclic pattern in the Google mobility data corresponds to each week.

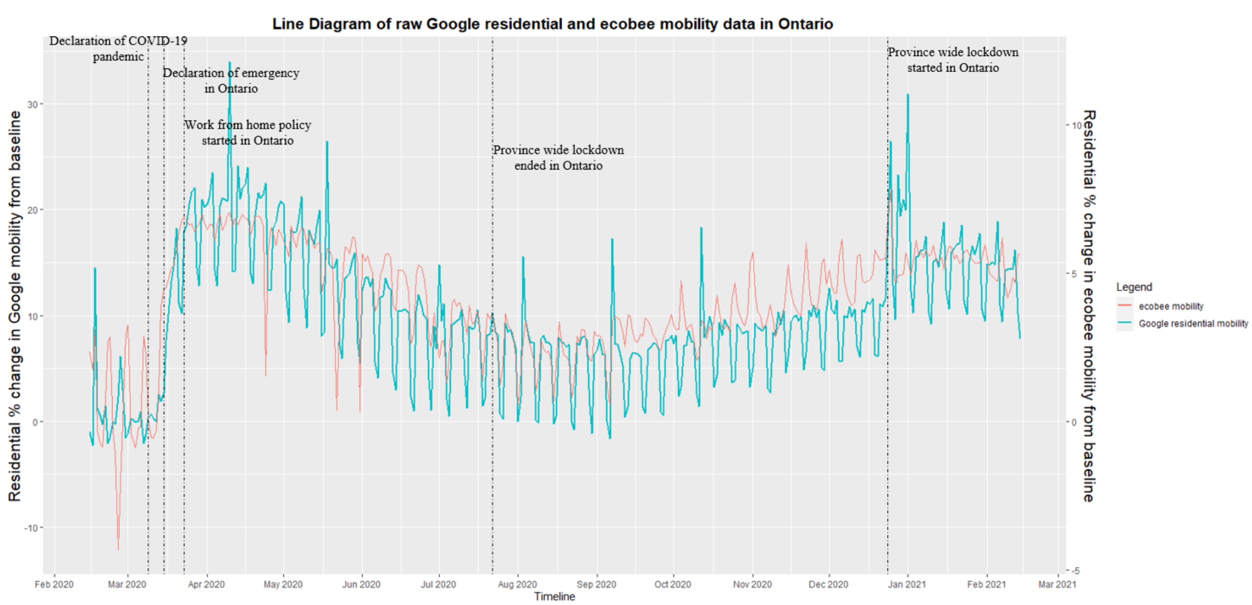


Figure 45. Association between Google residential mobility and ecobee mobility for the province of Ontario, Canada.

To represent the change in in-house and out of house behaviour of the residents, Figure 45 depicts a line plot for Google residential mobility data and ecobee mobility data for one year. The plot shows that in-house behaviour started to change from March 11, 2020, when COVID-19 pandemic was declared whereas there was not much change in the out of house mobility. In contrast, in-house and out of house residential mobility drastically increased when a province-wide emergency was declared in Ontario from March 17, 2020, onwards. On March 25, 2020, when work-from-home policy started in Ontario, the overall residential mobility increased. From March 2020 to July 2020, the trend declined gradually, and on July 24, 2020, when province-wide lockdown ended, the pattern remains consistent till September 2020 which can be inferred from the plot. With the advent of the second wave of the pandemic, another province-wide lockdown was announced on December 26, 2020, which is evident from the increased mobility values for Google and ecobee mobility datasets.

8.3.1.2 Quebec

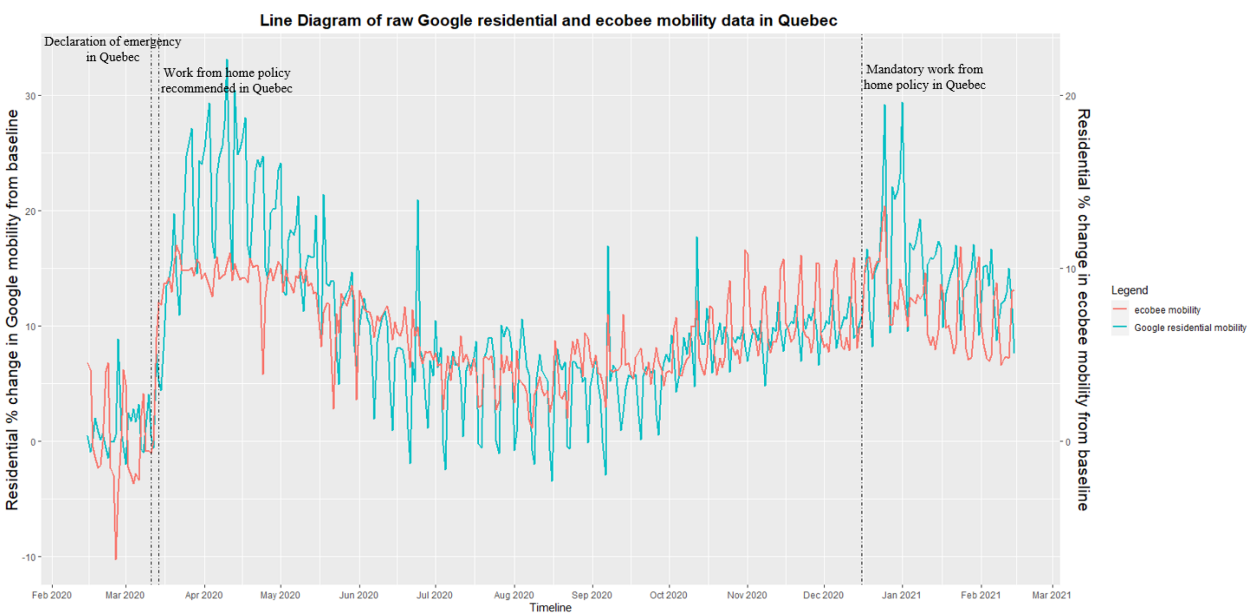


Figure 46. Association between Google residential mobility and ecobee mobility for the province of Quebec, Canada.

The line plot for Quebec Figure 46 shows March 13, 2020, when the declaration of public health emergency was made, in-house behaviour started changing, whereas out of house mobility has not been observed to change much. In contrast, in-house and out-of-the-house mobility increased drastically when work-from-home was recommended in Quebec starting on March 16, 2020. On top of that, when the work-from-home policy was implemented in Quebec on March 25, 2020, it enhanced the overall residential mobility for both datasets. On December 17, 2020, mandatory work-from-home policy was implemented in Quebec, resulting in a sudden increase in in-house mobility.

8.3.1.3 Alberta

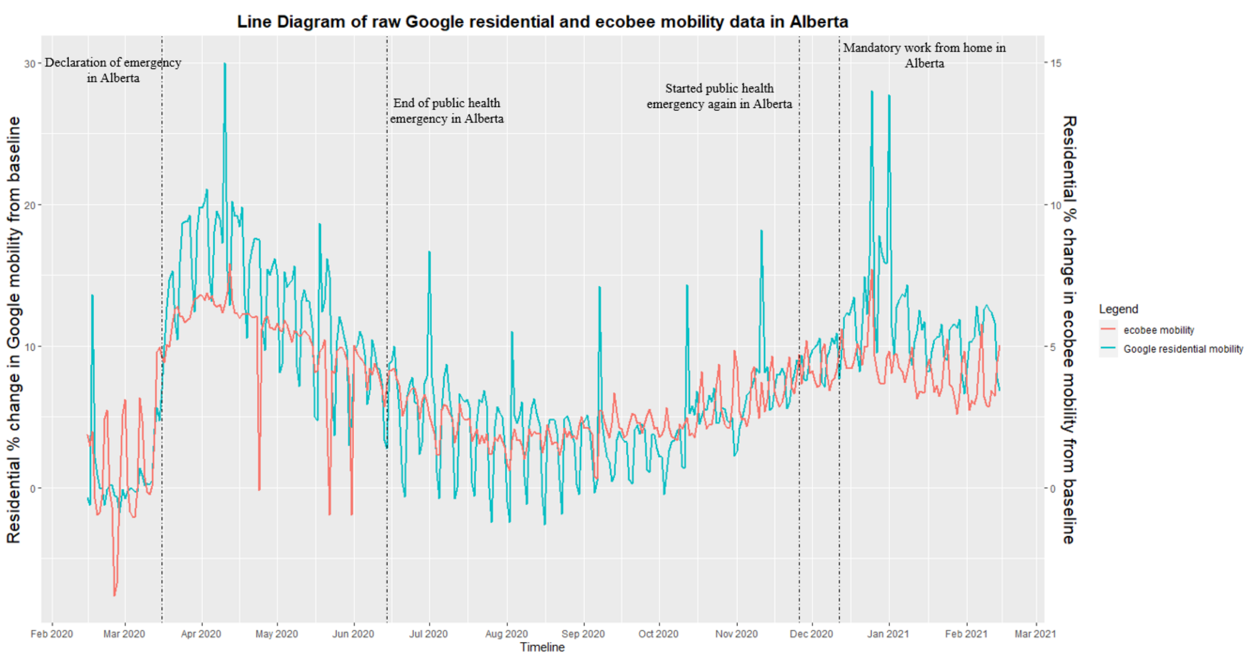


Figure 47. Association between Google residential mobility and ecobee mobility for the province of Alberta, Canada.

The line plot for Alberta in Figure 47 shows that in-house behaviour started changing on March 17, 2020, when a public health emergency declaration was made. In contrast, out-of-the-house mobility has not been observed to have changed too much. In contrast, in-house and out-of-the-house residential mobility increased drastically when Alberta's public health emergency ended on June 15, 2020. On top of that, when the public health emergency was declared again in Alberta on November 27, 2020, it enhanced residential mobility overall. On December 13, 2020, a mandatory work-from-home policy showed a sudden increase in in-house mobility.

8.3.1.4 British Columbia

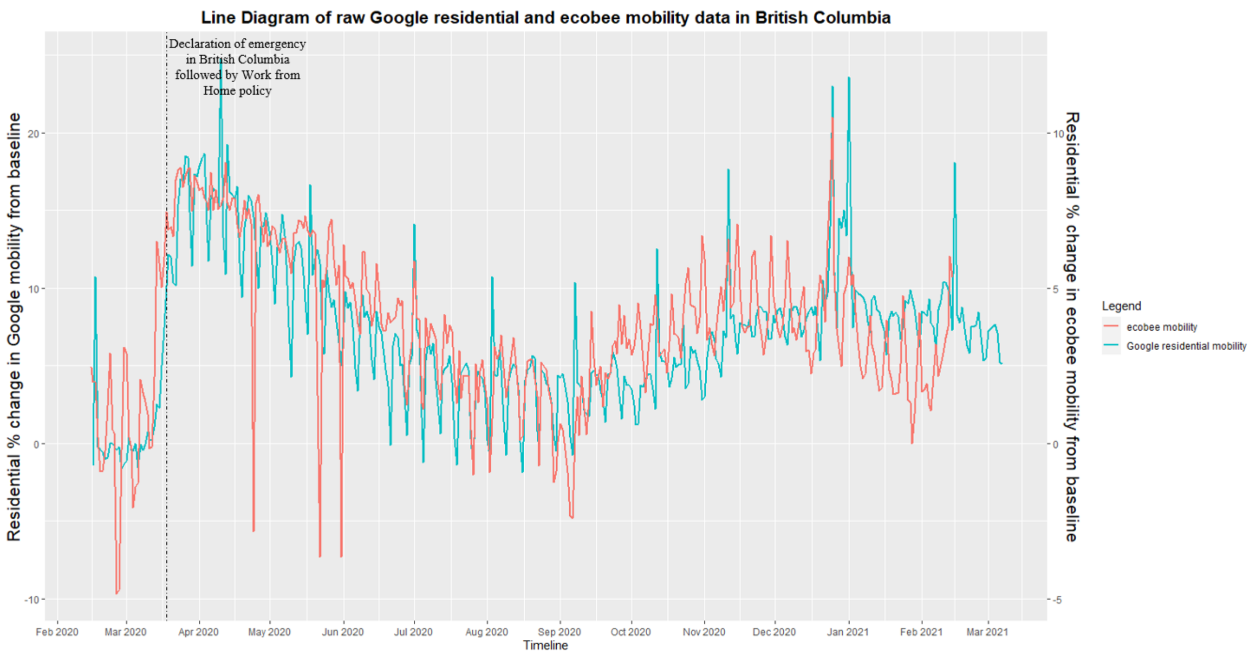


Figure 48. Association between Google residential mobility and ecobee mobility for the province of British Columbia, Canada.

The line plot for British Columbia in Figure 48 shows that in-house behaviour started changing on March 18, 2020, when the public health emergency declaration was made. In contrast, out-of-the-house mobility has not been observed to have changed too much. In contrast, in-house and out-of-the-house residential mobility increased drastically when work-from-home was recommended in British Columbia starting on March 19, 2020. This analysis is limited to Feb 2020 to March 2021. BC additional policies, but most of them in the end of 2021or early 2022.

8.3.1.2 Correlation Test Results for Google Residential Mobility and ecobee Mobility

To understand the statistical significance of the association between Google and ecobee mobility data, Table 21 shows the Pearson's and spearman's correlation test results and 95 percent confidence interval. A statistically significant association between the mobility datasets at the province level was found within the range of 0.67 to 0.73.

Table 21. Correlation between google and ecobee mobility data for Ontario, Alberta, Quebec, and British Columbia.

	No of households	Pearson's product-moment correlation			Spearman's rank correlation
		Correlation coefficient	95 percent confidence interval		
Ontario	7134	0.73	0.67	0.77	0.75
Alberta	3989	0.73	0.69	0.78	0.76
Quebec	708	0.67	0.61	0.73	0.70
British Columbia	421	0.69	0.64	0.74	0.63

Figure 49 depicts the association between Google residential mobility and ecobee mobility for four provinces. The association between these datasets are statistically significant and the trend line shows a linear association between both the datasets.

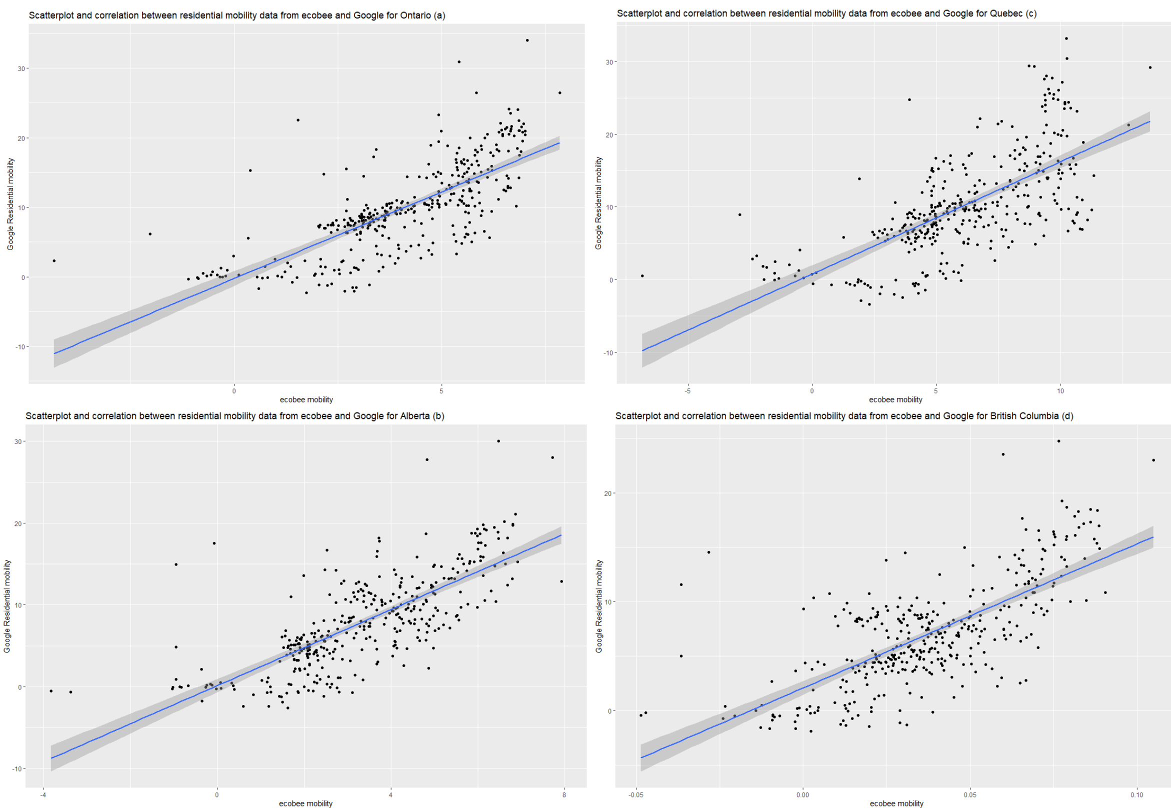


Figure 49. Correlation between residential mobility data from ecobee and Google for the province of a) Ontario b) Alberta c) Quebec d) British Columbia, Canada.

8.3.2 Seasonal Diagnostics of the Google and ecobee Mobility Data

Seasonal diagnostics of the mobility data for each province have been described below. Seasonal diagnostic for the mobility data included three analysis levels: days of the week, month by month, and week by week. The same analysis has been replicated for Google and ecobee mobility datasets to understand granular findings.

8.3.2.1 Days of the Week Variation Analysis

8.3.2.1.1 Ontario

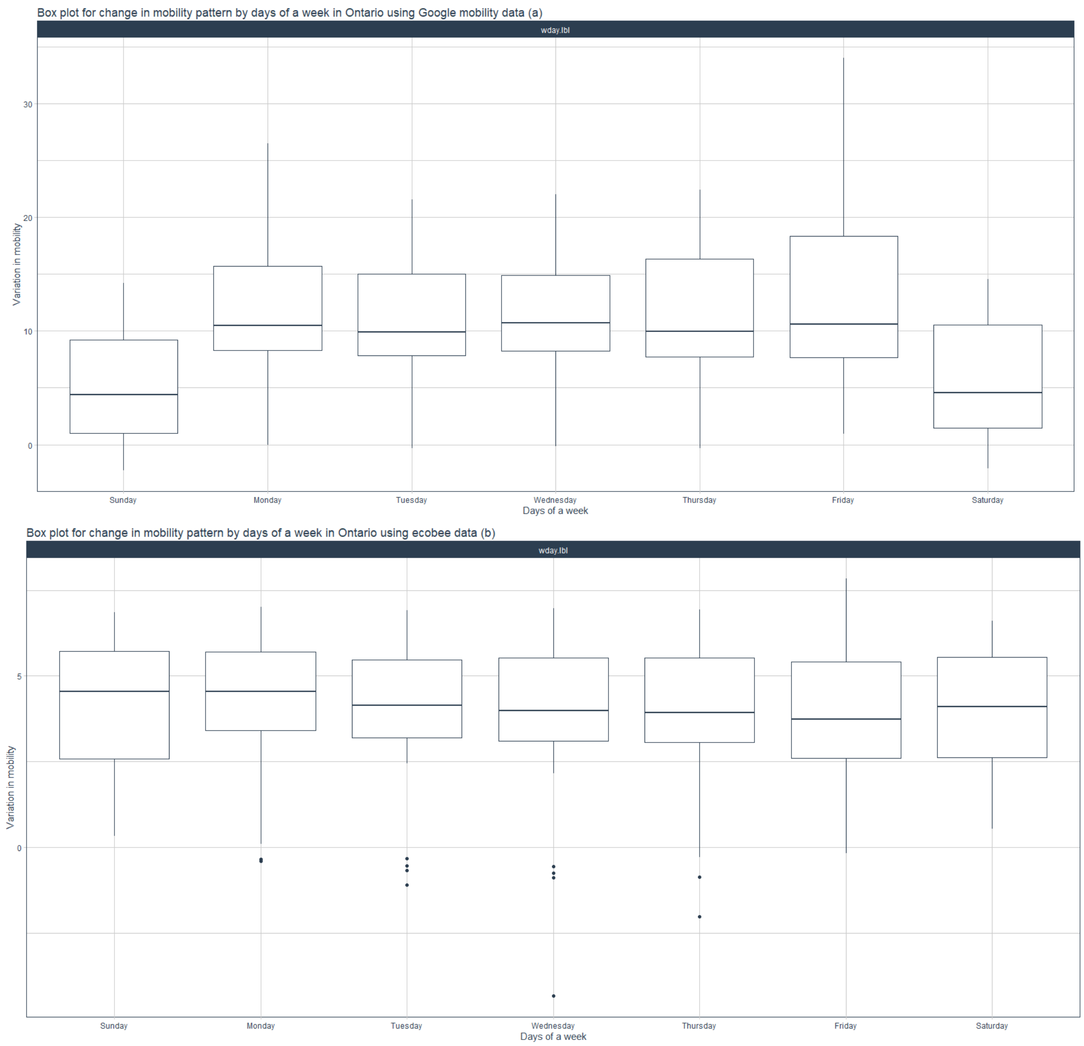


Figure 50. Analysis of the a) Google residential and b) ecobee mobility data among days of a week in Ontario.

The boxplot in Figure 50 shows there is a significant difference in behavioural patterns for Ontario's Google residential mobility dataset between weekends and weekdays. Strikingly, higher mobility was observed on the weekdays compared to weekends. On the contrary, there is no significant difference in the mobility patterns between weekdays and weekends for ecobee's mobility data.

8.3.2.1.2 Quebec

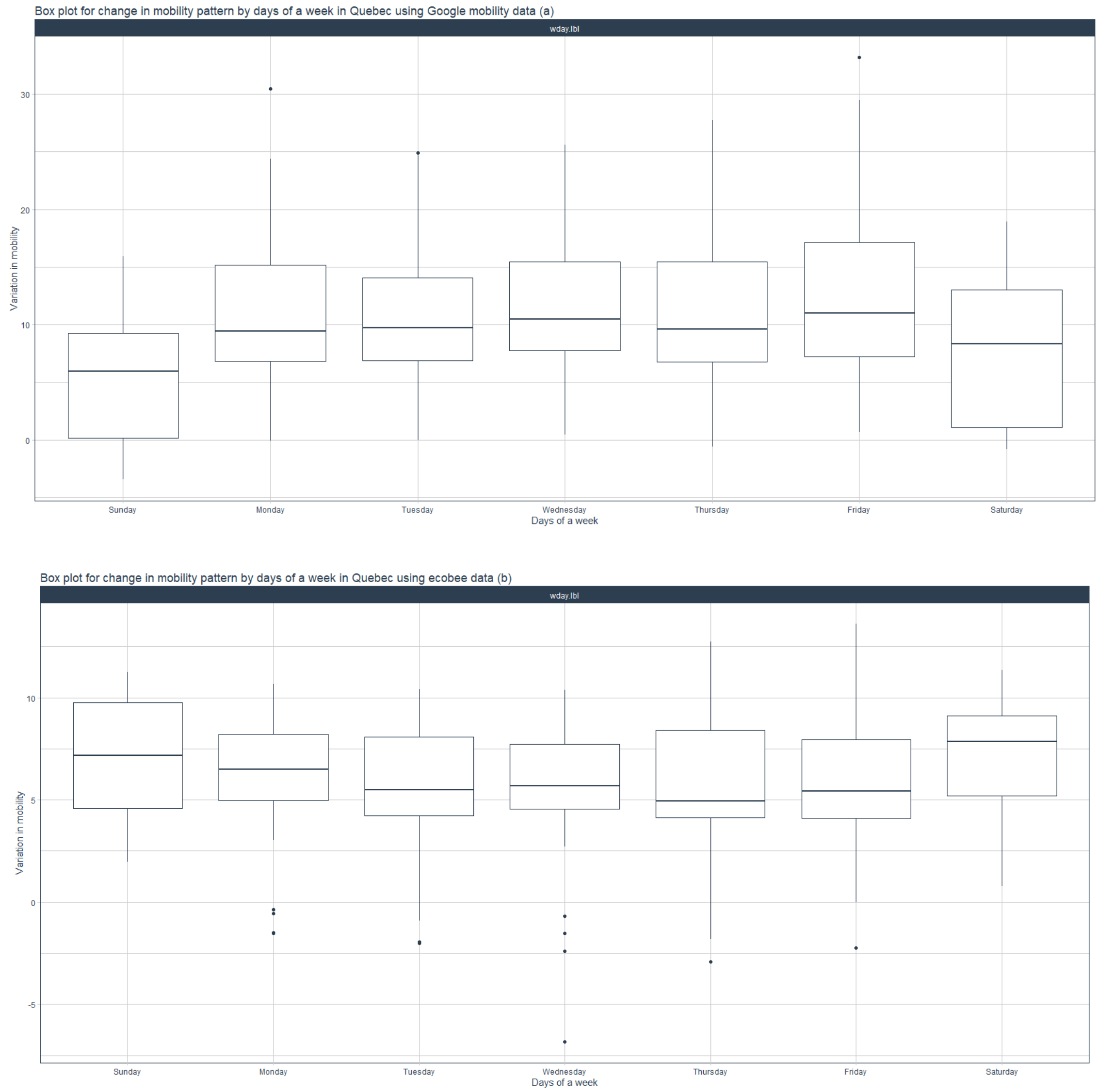


Figure 51. Analysis of the a) Google residential and b) ecobee mobility data among days of a week in Quebec.

Google residential mobility dataset for the province of Quebec Figure 51 shows that mobility behaviour over the weekends is entirely different from that observed on weekdays. Higher mobility on the weekdays out-of-the-house compared to weekends is observed in the plots. In agreement with this observation, ecobee mobility data also shows a statistically significant difference between weekdays and weekends during that period. On the weekends, higher stay time is there compared to weekdays.

8.3.2.1.3 Alberta

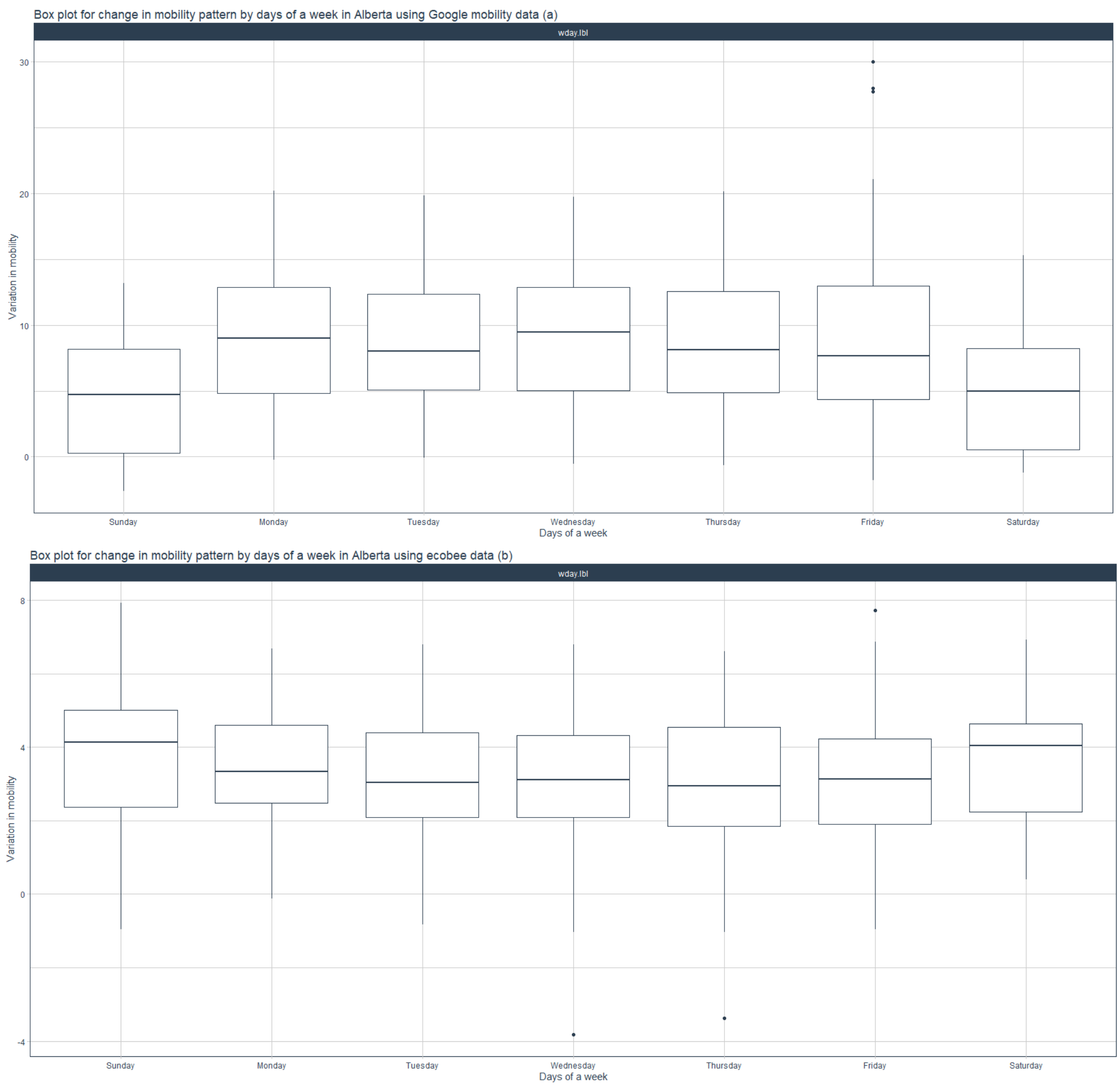


Figure 52. Analysis of the a) Google residential and b) ecobee mobility data among days of a week in Alberta.

Google residential mobility dataset for the province of Alberta Figure 52 shows mobility patterns over the weekends are entirely different from the behaviour observed on the weekdays. Higher mobility on the weekdays on the outside compared to weekends is seen. In contrast, ecobee mobility data shows no statistically significant difference between weekdays and weekends during that period.

8.3.2.1.4 British Columbia

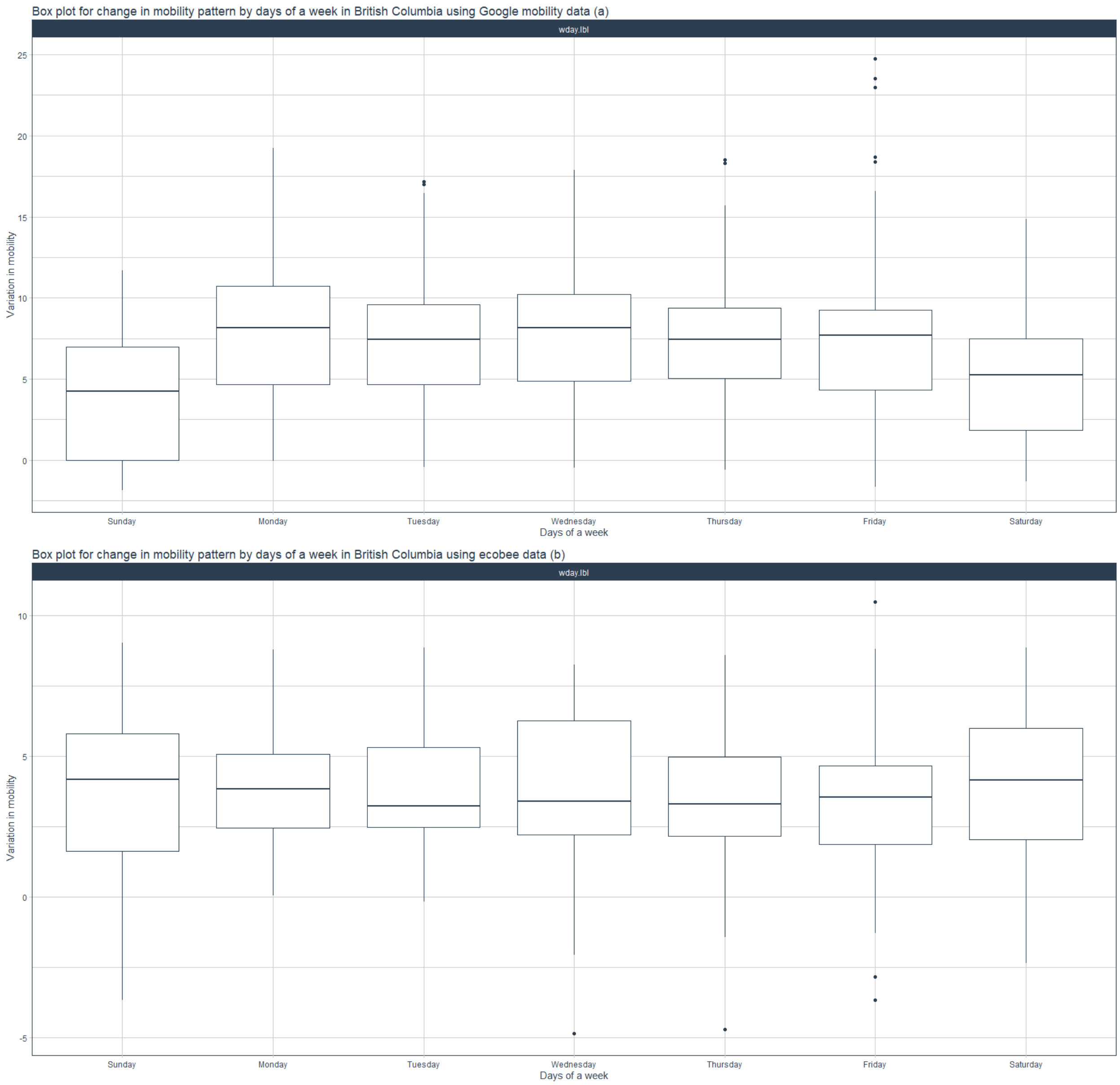


Figure 53. Analysis of the a) Google residential and b) ecobee mobility data among days of a week in British Columbia.

Google residential mobility dataset for the province of British Columbia Figure 53 shows that behaviour over the weekends is entirely different from that observed on the weekdays. Higher mobility on the weekdays on the outside is seen instead of weekends. In contrast, ecobee mobility data shows no statistically significant difference between weekdays and weekends during that period.

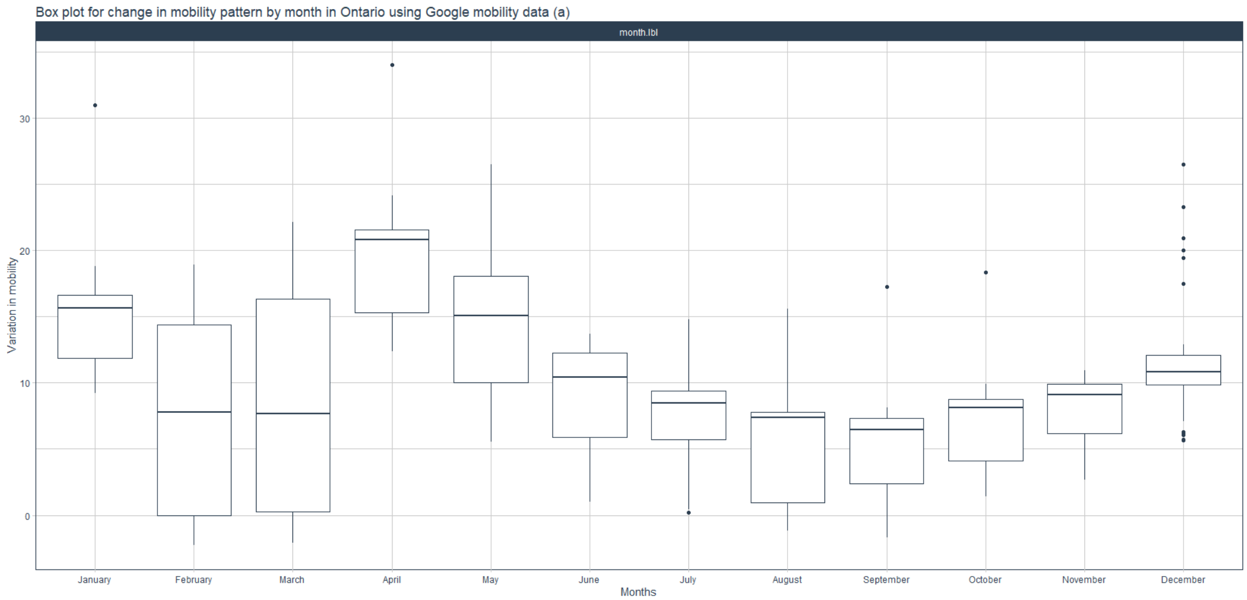
Table 22. The ANOVA test compares the day of the week's impact on Google and ecobee mobility for the four provinces of Canada.

		Google				**Ecobee**			
Province		**df**	**Sum of squares**	**F**	***P*-value**	**df**	**Sum of squares**	**F**	***P*-value**
Ontario	C(Weekday)	6	3467	17.86	<.001	6	10	0.458	0.84
	Residual	358	11579			358	1306		
Quebec	C(Weekday)	6	2426	9.357	<.001	6	130	2.364	0.03
	Residual	358	15471			358	3278		
Alberta	C(Weekday)	6	1456	8.223	<.001	6	22.9	1.123	0.348
	Residual	358	10566			358	1216		
British Columbia	C(Weekday)	6	897	6.955	<.001	6	14.9	0.371	0.897
	Residual	358	7673			358	2404		

One-way ANOVA test results (Table 22) show that day of the week has a statistically significant impact on out-of-the-house mobility for all four provinces (Google). In contrast, no such trend has been found for in-house mobility except Quebec (ecobee).

8.3.2.2 Month by Month Variation Analysis

8.3.2.2.1 Ontario



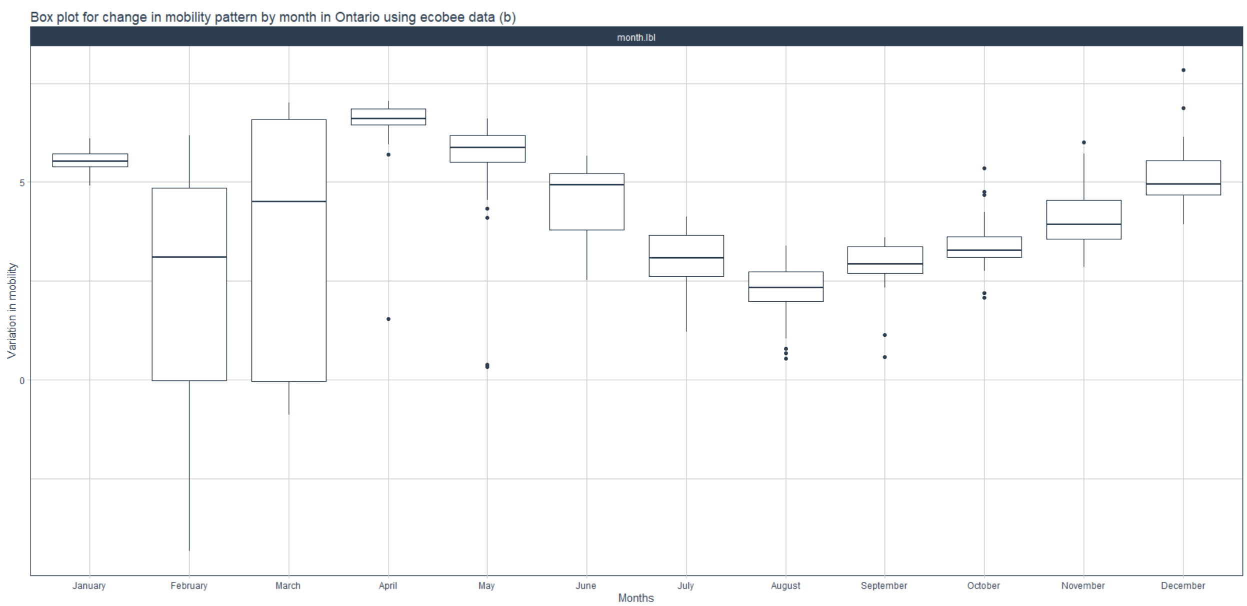


Figure 54. Analysis of the a) Google residential and b) ecobee mobility data among months in Ontario.

As shown in Figure 54, Google residential mobility dataset for the province of Ontario shows enormous variation in mobility for the months of February and March 2020 which later stabilized in April 2020. A higher mobility for April 2020 onwards proves an increase in residential mobility. Interestingly, in contrast to variation observed for weekdays, month by month variation patterns are very similar between Google and ecobee mobility data. The variability within the datasets specifically for ecobee data reduced drastically from the month of April 2020, which may be due to declaration of emergency in the provincial as well as national level due to COVID-19 pandemic.

8.3.2.2.2 Quebec

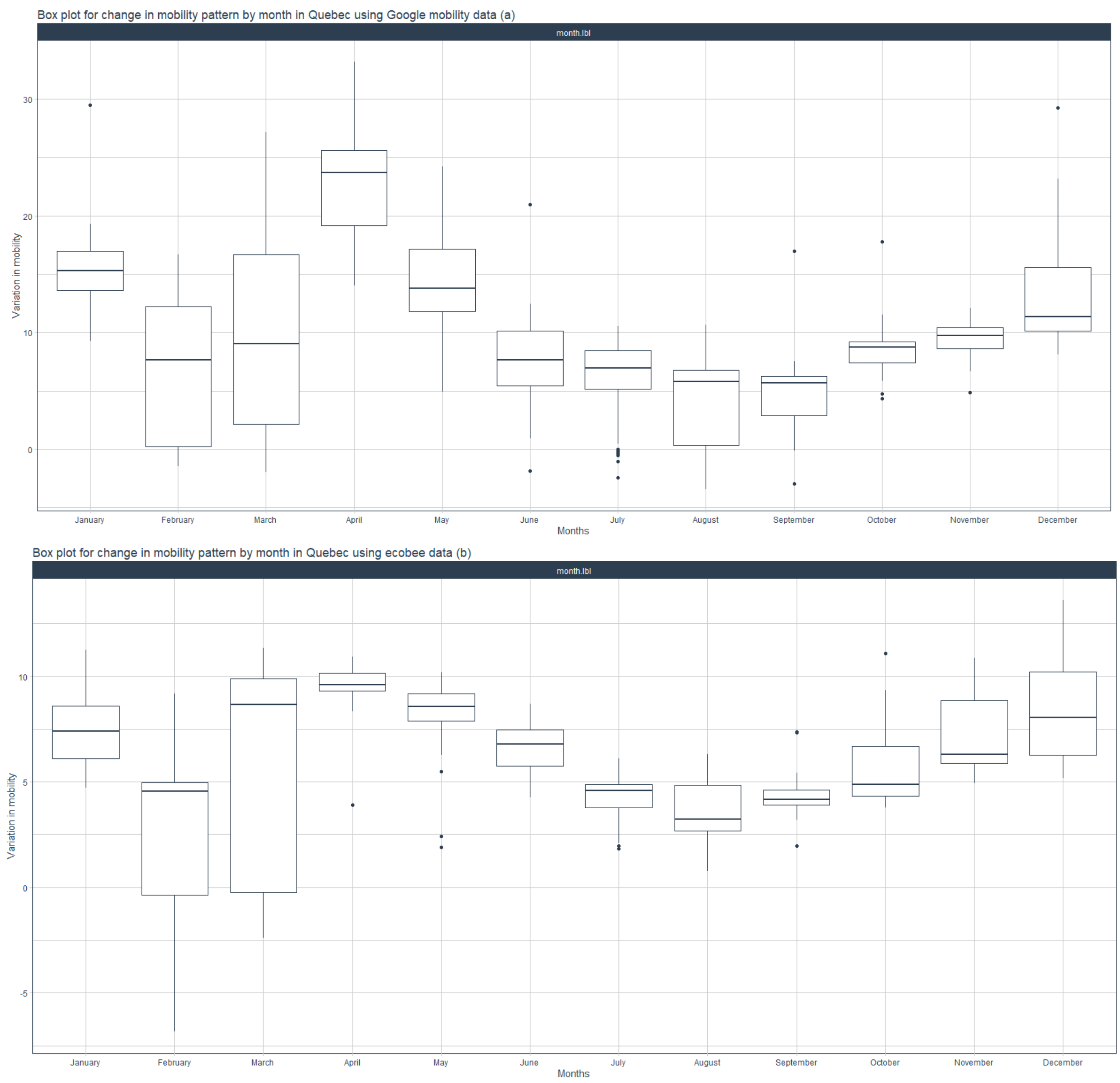
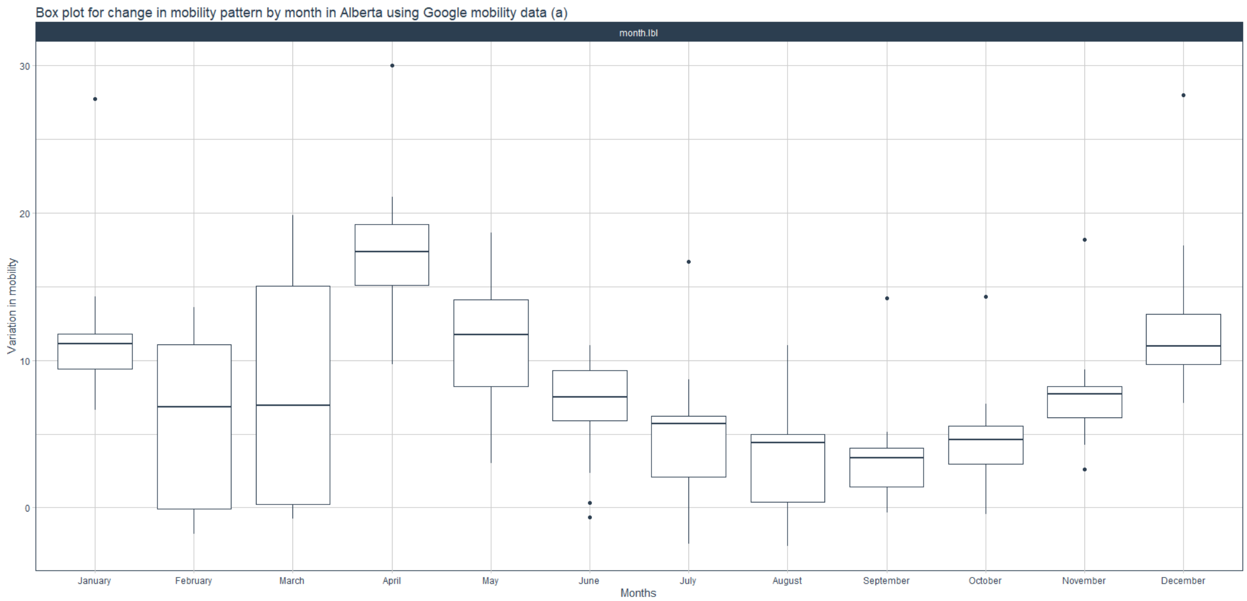


Figure 55. Analysis of the a) Google residential and b) ecobee mobility data among months in Quebec.

As shown in Figure 55, the Google residential mobility dataset for the province of Quebec in February and March 2020 has vast variation in mobility, which stabilized in April with a higher value than the baseline. Higher mobility for April onwards proved an increase in residential mobility compared to other places. Like weekdays variation, month by month pattern was also very similar between Google and ecobee mobility data both show a similarity in the mobility pattern in terms of the month.

8.3.2.2.3 Alberta



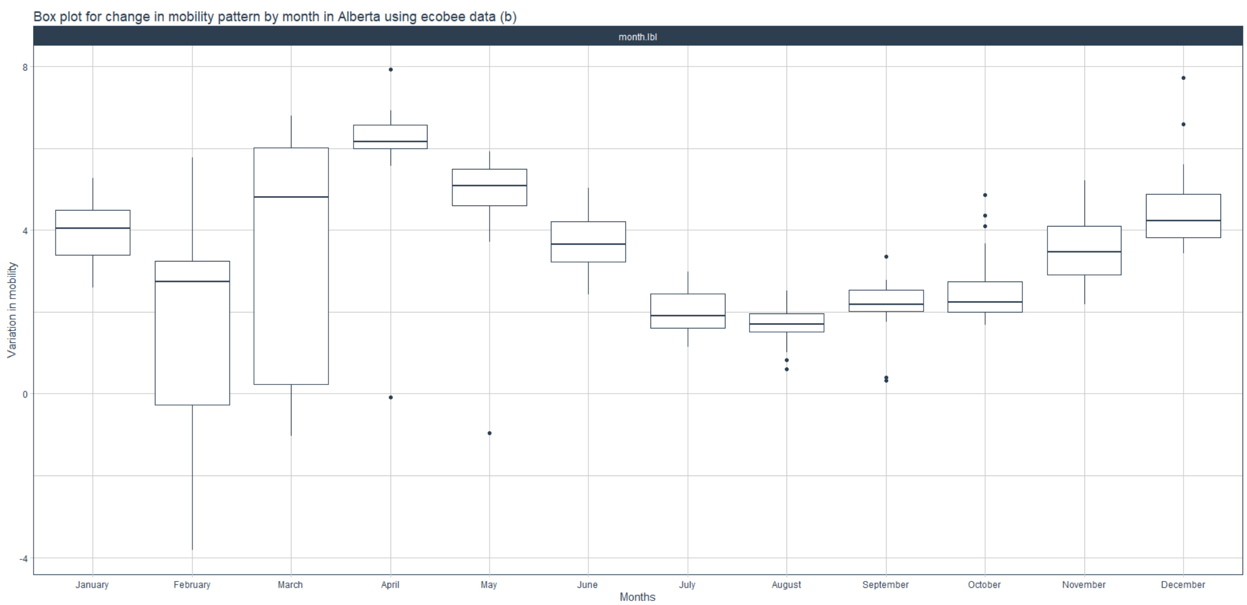


Figure 56. Analysis of the a) Google residential and b) ecobee mobility data among months in Alberta.

Google residential mobility dataset for the province of Alberta shown in Figure 56 shows that February and March 2020 have massive variation in mobility, which stabilized in April with a higher value than the baseline. Higher mobility for April onwards demonstrated an increase in residential mobility in Alberta compared to other places. In contrast, to weekdays variation month by month pattern is very similar to ecobee mobility data shows there is a similarity in the mobility pattern in terms of the month. Like Ontario, the variability of the data has been reduced in April 2020 compared to February and March 2020.

8.3.2.2.4 British Columbia

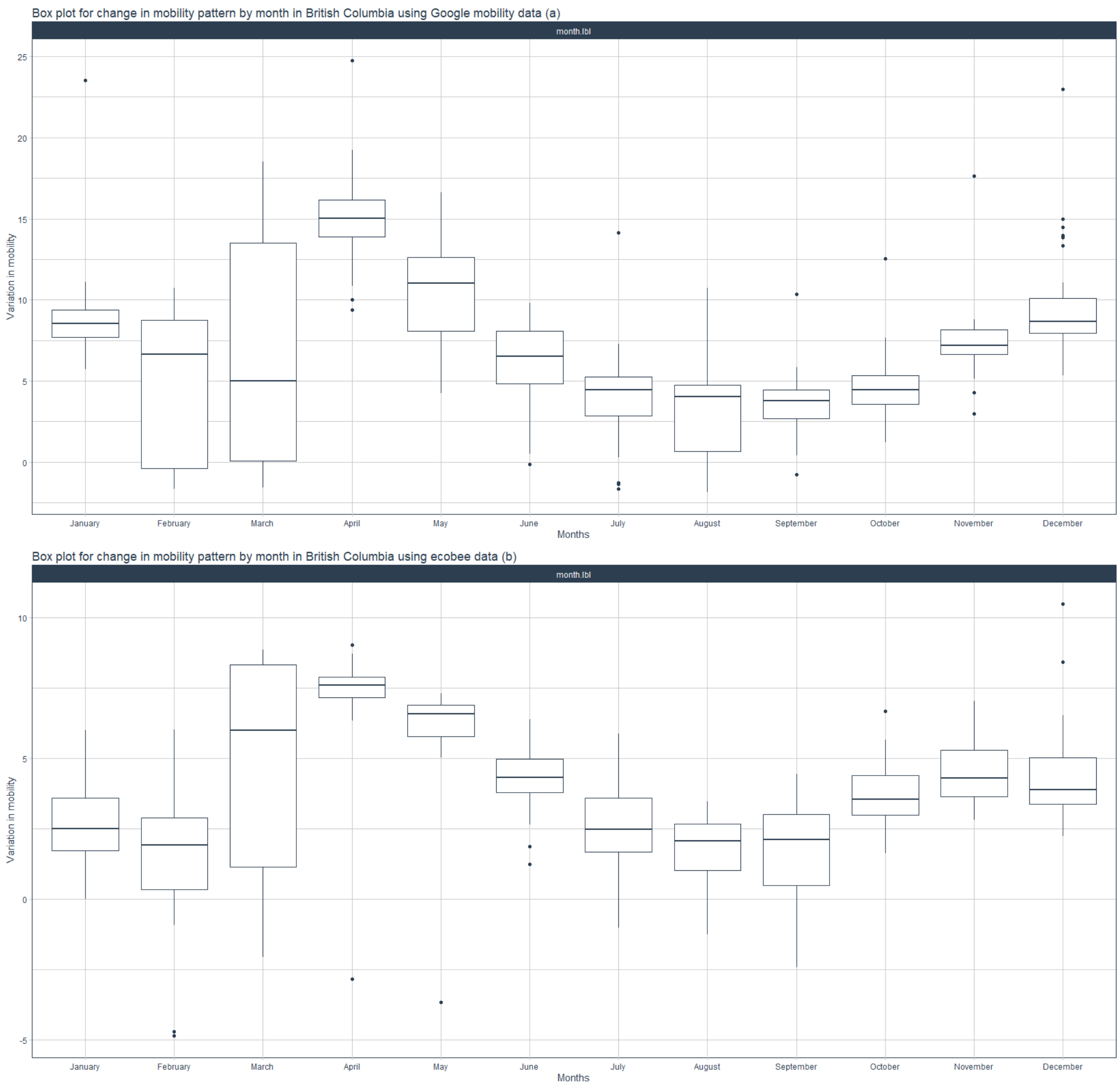


Figure 57. Analysis of the a) Google residential and b) ecobee mobility data among months in British Columbia.

As shown in Figure 57, Google residential mobility dataset for the province of British Columbia shows that February and March have had enormous variation in mobility, which stabilized in April with a higher value than the baseline. Higher mobility trends observed April onwards proved an increase in residential mobility compared to other places. In contrast to weekdays variation, month by month pattern is very similar between google residential mobility and ecobee mobility data.

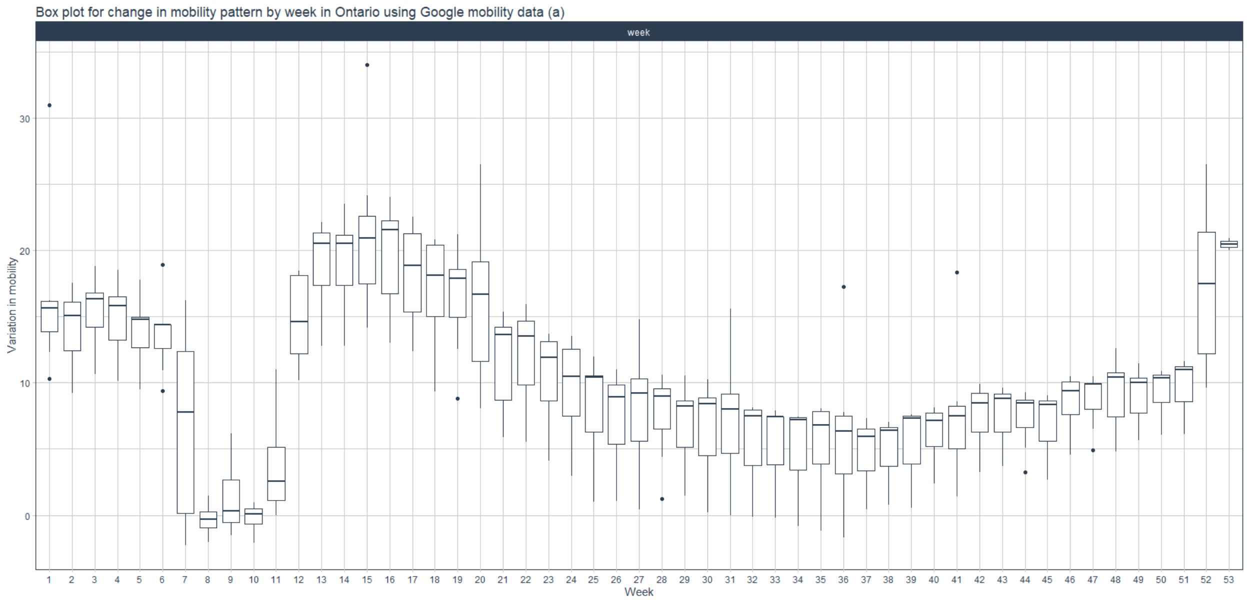
One-way ANOVA test results show a statistically significant difference for out-of-the-house and in-house mobility for all four provinces illustrated in Table 23, when comparing them month-by-month.

Table 23. The ANOVA test compares month by month impact on Google and ecobee mobility for the four provinces of Canada.

		Google				**Ecobee**			
Province		**df**	**Sum of squares**	**F**	***P*-value**	**df**	**Sum of squares**	**F**	***P*-value**
Ontario	C(Month)	11	6383	23.65	<.0001	11	591.4	26.18	<.0001
	Residual	353	8663			353	724.8		
Quebec	C(Month)	11	9541	36.64	<.0001	11	1372	21.61	<.0001
	Residual	353	8357			353	2037		
Alberta	C(Month)	11	5923	31.16	<.0001	11	606.4	30.76	<.0001
	Residual	353	6099			353	632.7		
British Columbia	C(Month)	11	4062	28.83	<.0001	11	1047	24.49	<.0001
	Residual	353	4508			353	1372		

8.3.2.3 Week by Week variation analysis

8.3.2.3.1 Ontario



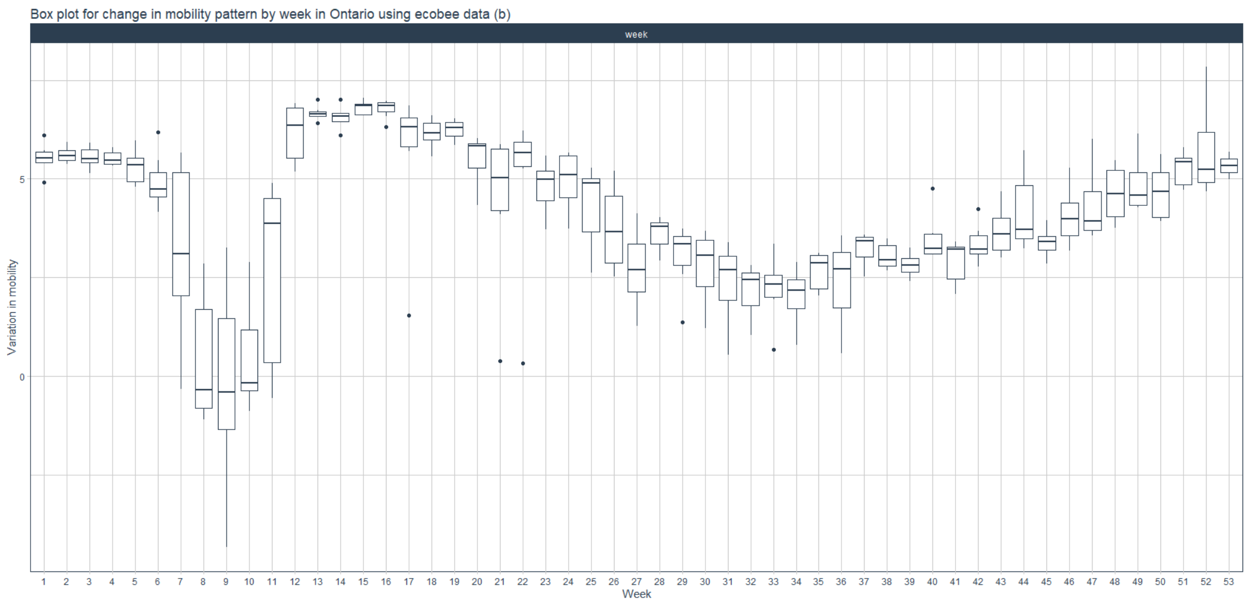
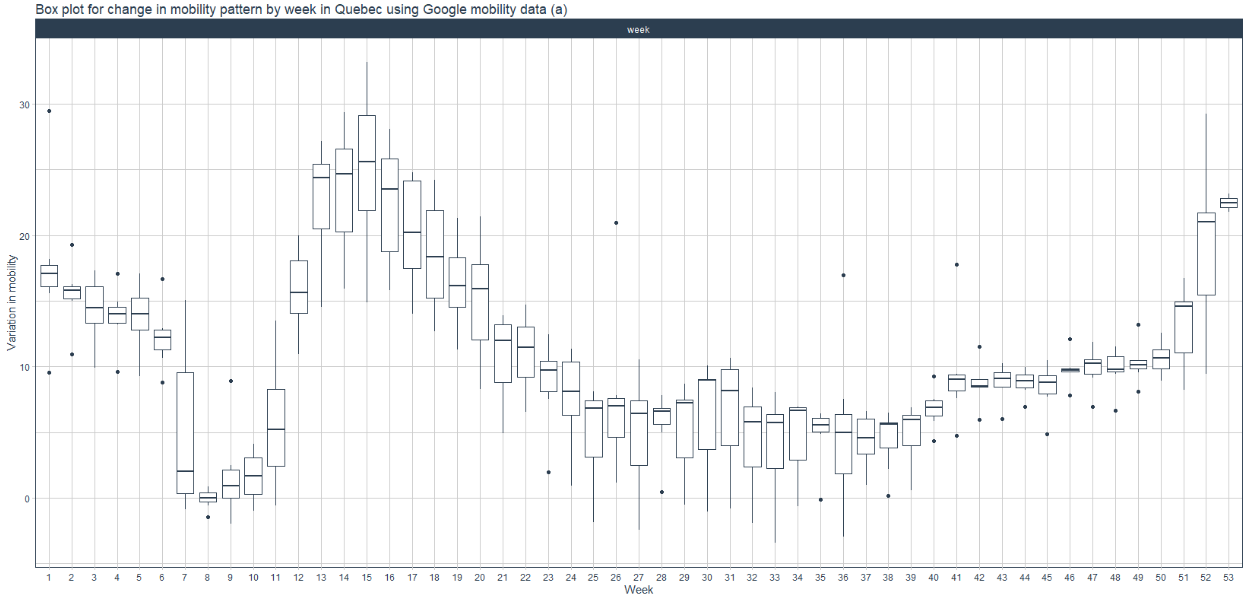


Figure 58. Analysis of the a) Google residential and b) ecobee mobility data among weeks in Ontario.

Google residential mobility dataset for the province of Ontario in Figure 58 shows that from the beginning of February 2020, the initial 5 to 6 weeks were following a similar trend. There is a reduction in mobility levels from the 7th week for both datasets. The beginning few weeks of March and April 2020 showcase larger variations in mobility which later stabilized towards the end of April. Mobility levels observed since April infers an increase in residential activities. Month-by-month variation and week-by-week variation shows both datasets have a similar pattern.

8.3.2.3.2 Quebec



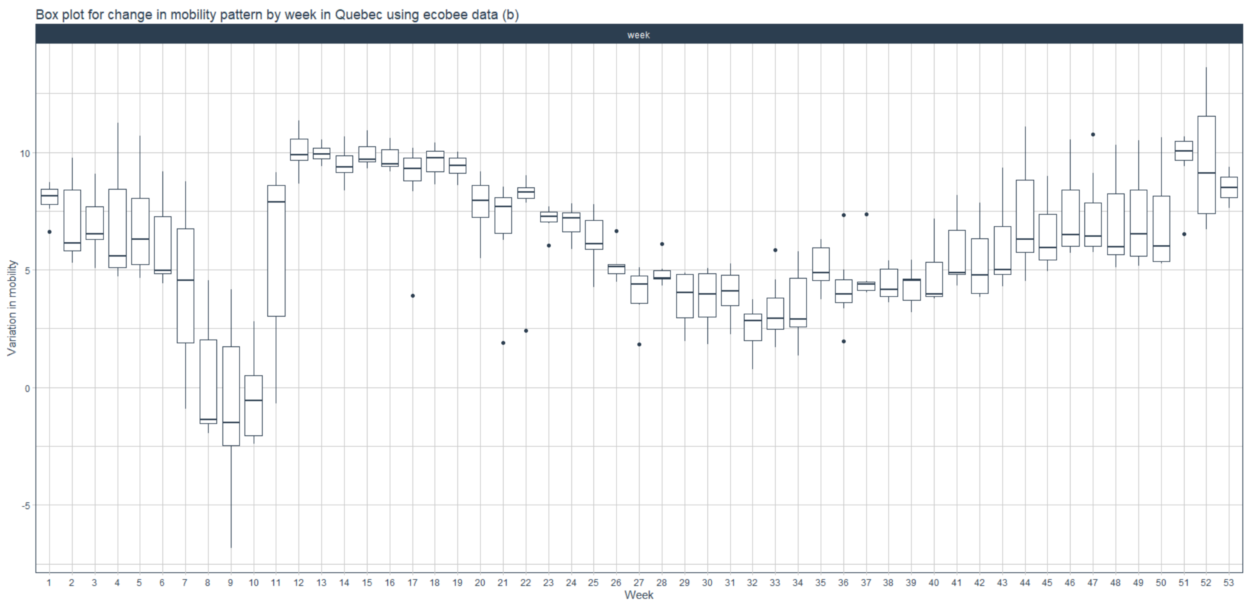
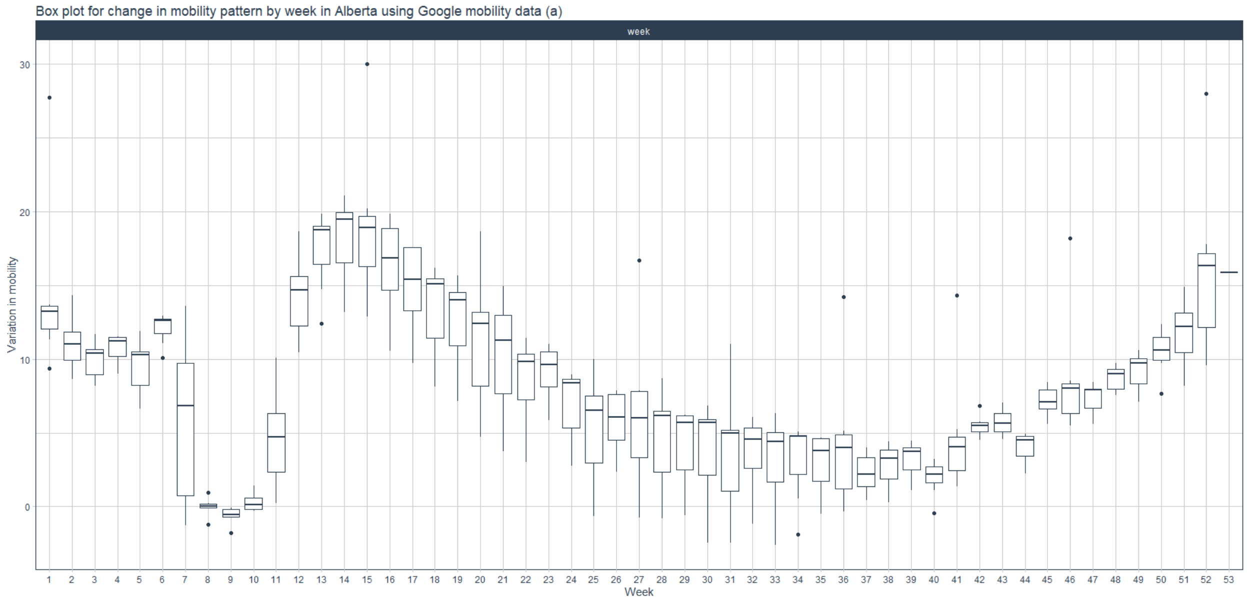


Figure 59. Analysis of the a) Google residential and b) ecobee mobility data among weeks in Quebec.

Google residential mobility dataset for the province of Quebec shown in Figure 59 shows that from the beginning of February 2020, the initial 5-6 weeks were similar. Changes were observed from the seventh week onwards; there is a reduction in the Google residential mobility dataset activity. In contrast, in the ecobee dataset, a lag period of one week before a similar pattern is observed. The weeks of March and April 2020 have enormous variations in mobility, which stabilized at the end of April with a higher value than the baseline. Higher mobility observed from April onwards showed an increase in residential mobility compared to other places. The month-by-month variation and the week-by-week pattern also show that the ecobee mobility data and Google residential mobility data have similar patterns.

8.3.2.3.3 Alberta



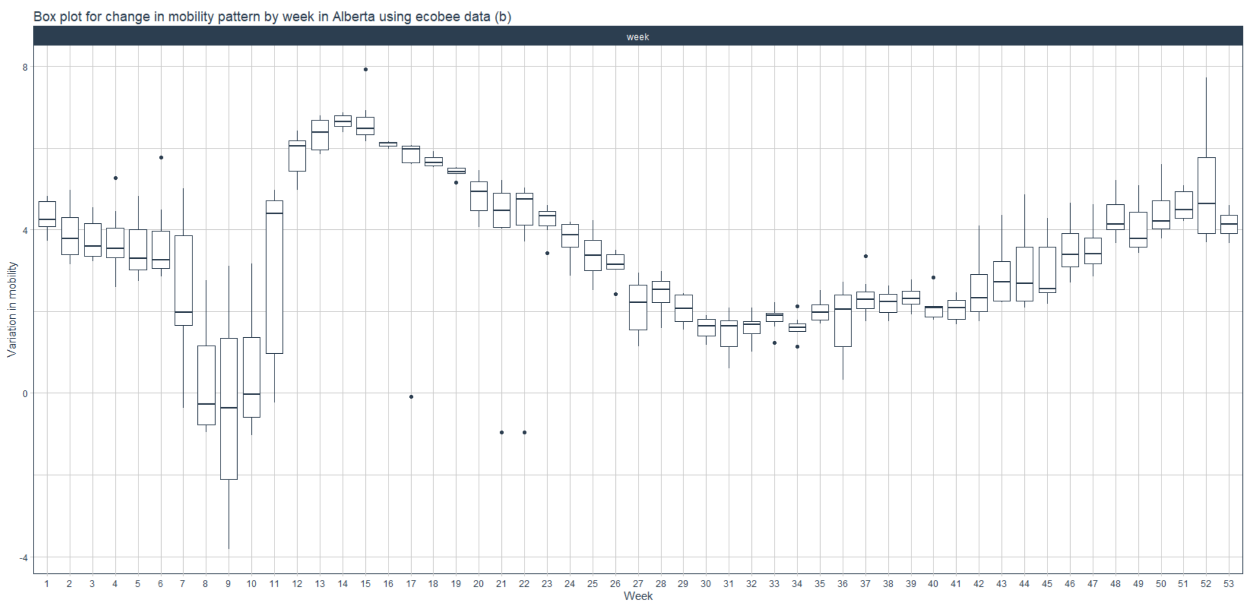


Figure 60. Analysis of the a) Google residential and b) ecobee mobility data among weeks in Alberta.

Google residential mobility dataset for the province of Alberta shown in Figure 60 shows that from the beginning of February 2020, the initial 5-6 weeks were similar. Change is observed from the seventh week onwards; there is a reduction in activity level, both on ecobee and Google residential mobility data. The weeks of March and April 2020 have larger variations in mobility, which stabilized near the end of April with a higher value than the baseline. Higher mobility observed April onwards proved there is an increase in residential mobility compared to other places. The month-by-month variation and week-by-week pattern show that ecobee mobility data and Google residential mobility data have a similar pattern.

8.3.2.3.4 British Columbia

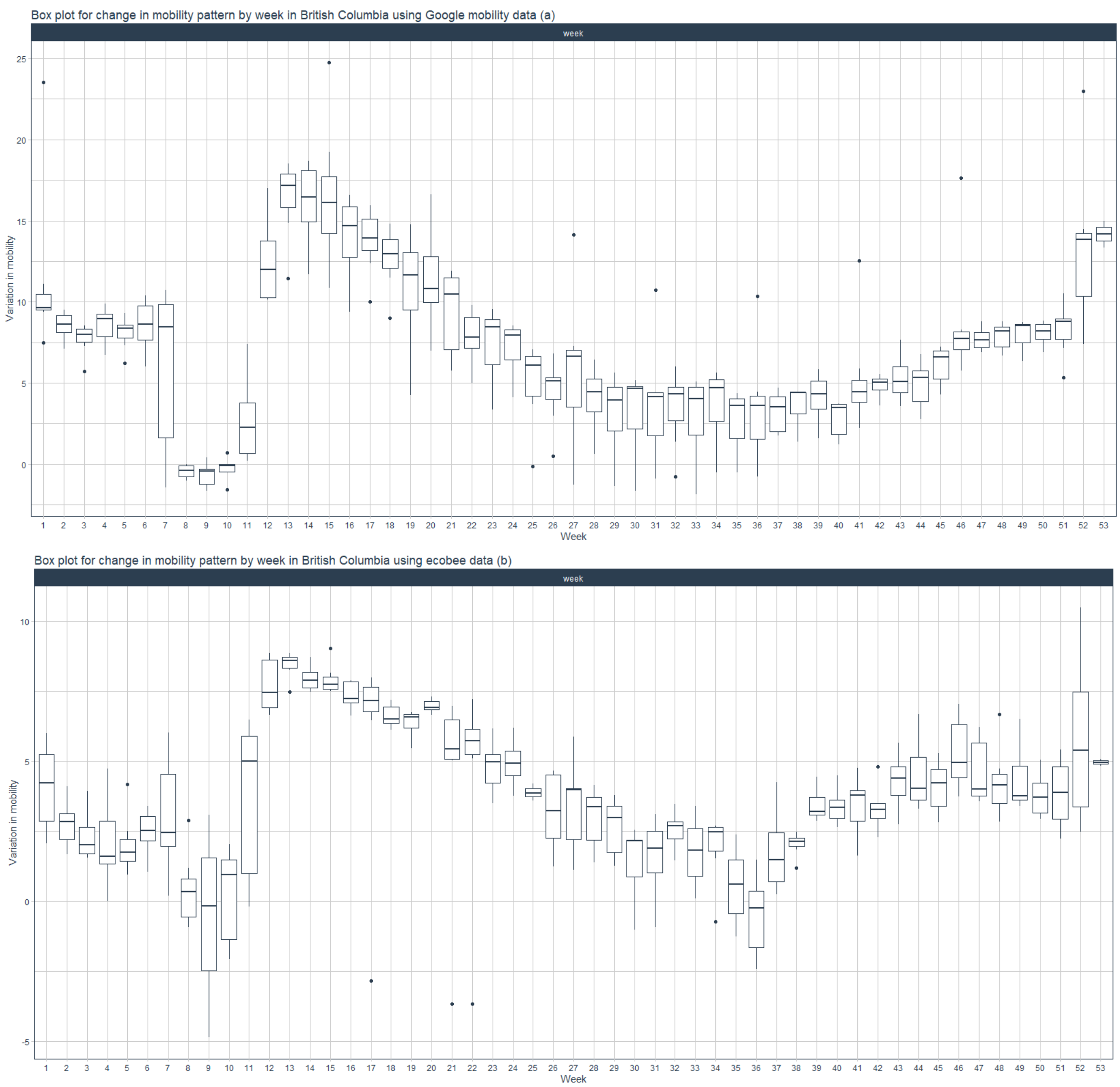


Figure 61. Analysis of the a) Google residential and b) ecobee mobility data among weeks in British Columbia.

As shown in Figure 61, the Google residential mobility dataset for the province of British Columbia shows that from the beginning of February 2020, the initial 5-6 weeks were similar. Changes were observed from the seventh week onwards; there was a reduction in activity level in both the ecobee dataset and Google residential mobility data. The weeks of March and April 2020 had massive variations in mobility, which stabilized near the end of April with a higher value than the baseline. Higher mobility observed April onwards proved there is an increase in residential mobility compared to other places. The month-by-month variation and week-by-week pattern also show ecobee mobility and Google residential mobility datasets have a similar pattern.

The results of ANOVA analysis for the week-by-week mobility pattern on ecobee and Google residential mobility data for four provinces was described in Table 24. A statistically significant difference between weeks was found for both datasets for all four provinces.

Table 24. The ANOVA test comparing week by week impact on Google and ecobee mobility for the four provinces of Canada.

		Google				**Ecobee**			
Province		**df**	**Sum of squares**	**F**	***P*-value**	**df**	**Sum of squares**	**F**	***P*-value**
Ontario	C(Week)	52	10140	12.4	<.0001	52	992.9	18.43	<.0001
	Residual	312	4906			312	323.2		
Quebec	C(Week)	52	13582	18.88	<.0001	52	2375	13.79	<.0001
	Residual	312	4316			312	1034		
Alberta	C(Week)	52	8961	17.57	<.0001	52	927.3	17.84	<.0001
	Residual	312	3061			312	311.9		
British Columbia	C(Week)	52	6558	19.49	<.0001	52	1675.8	13.53	<.0001
	Residual	311	2012			312	742.9		

8.3.3 Anomaly Detection

Anomaly detection is the process of identification of unexpected events, observations that significantly distinct from the typical event. Putting unlabeled data as input and the unsupervised algorithm detects several points in time (here dates). Anomalies occurs very rarely and related to some sort of problem or instances. In this case the findings suggest some policy changes or significant change in mobility pattern. In public health anomalies detection has the capacity to identify abnormal events which are of great importance in terms of preventive measures.

8.3.3.1 Ontario

An anomaly detection analysis of Google residential data for Ontario shown in Figure 62 shows no anomaly within 2020. A similar pattern can be seen in ecobee mobility data.

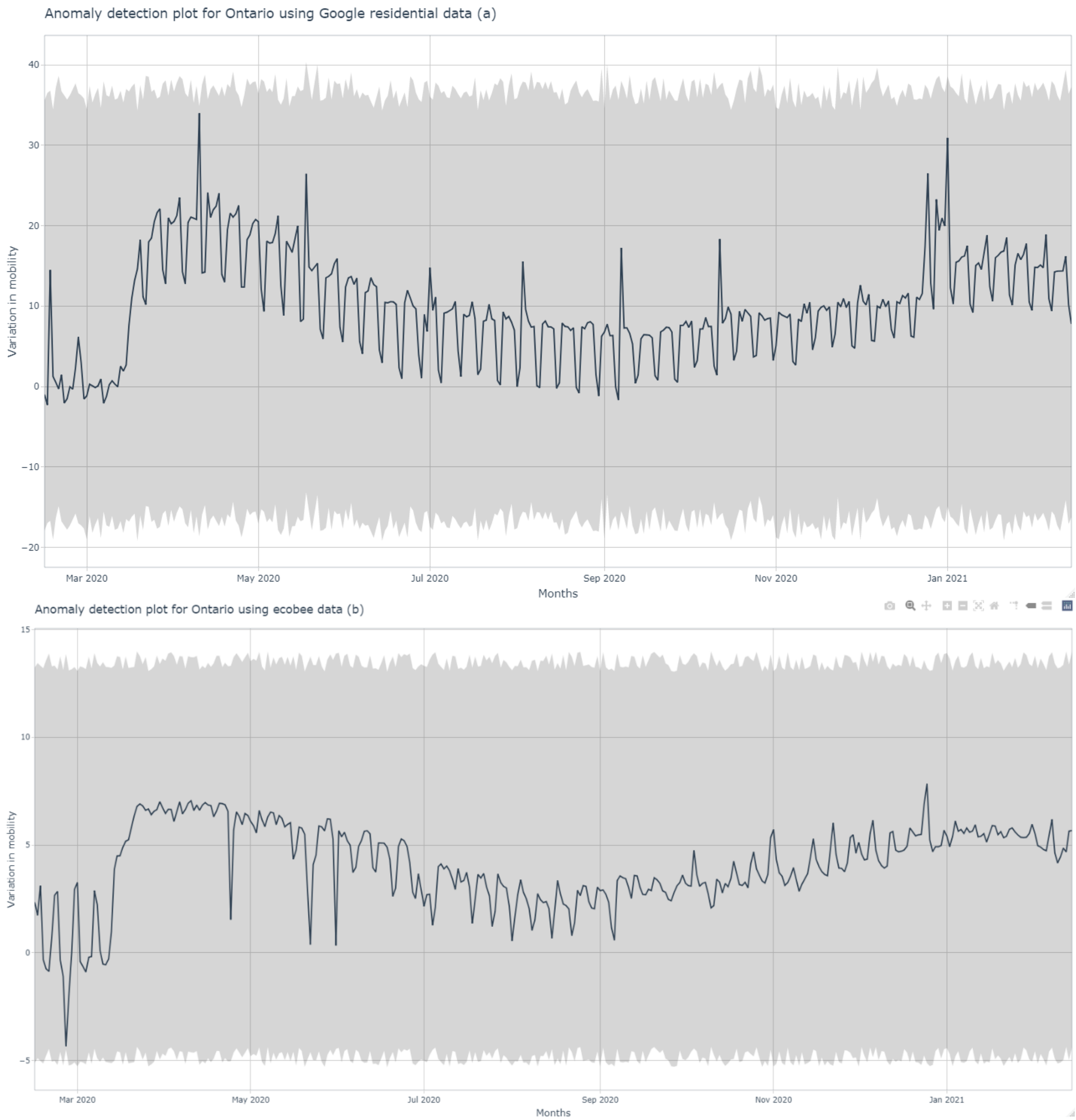


Figure 62. *Anomaly detection plot for the province of Ontario a) Google residential data b) ecobee mobility data.*

8.3.3.2 Quebec

The anomaly detection analysis of Google residential data for Quebec Ontario in Figure 63 shows an anomaly at the beginning period, similar to that of the ecobee data. Notably, this can be related with the dates of COVID-19 related policy changes. Moreover, festive periods like

Christmas and New Year are the other anomalies captured through Google residential data. During these events, abnormal mobility was detected out-of-the-house. However, ecobee could not capture these anomalies because of no significant changes in in-house mobility which can be attributed to policy restrictions. Interestingly, a similar pattern can be seen for both datasets within the last year.

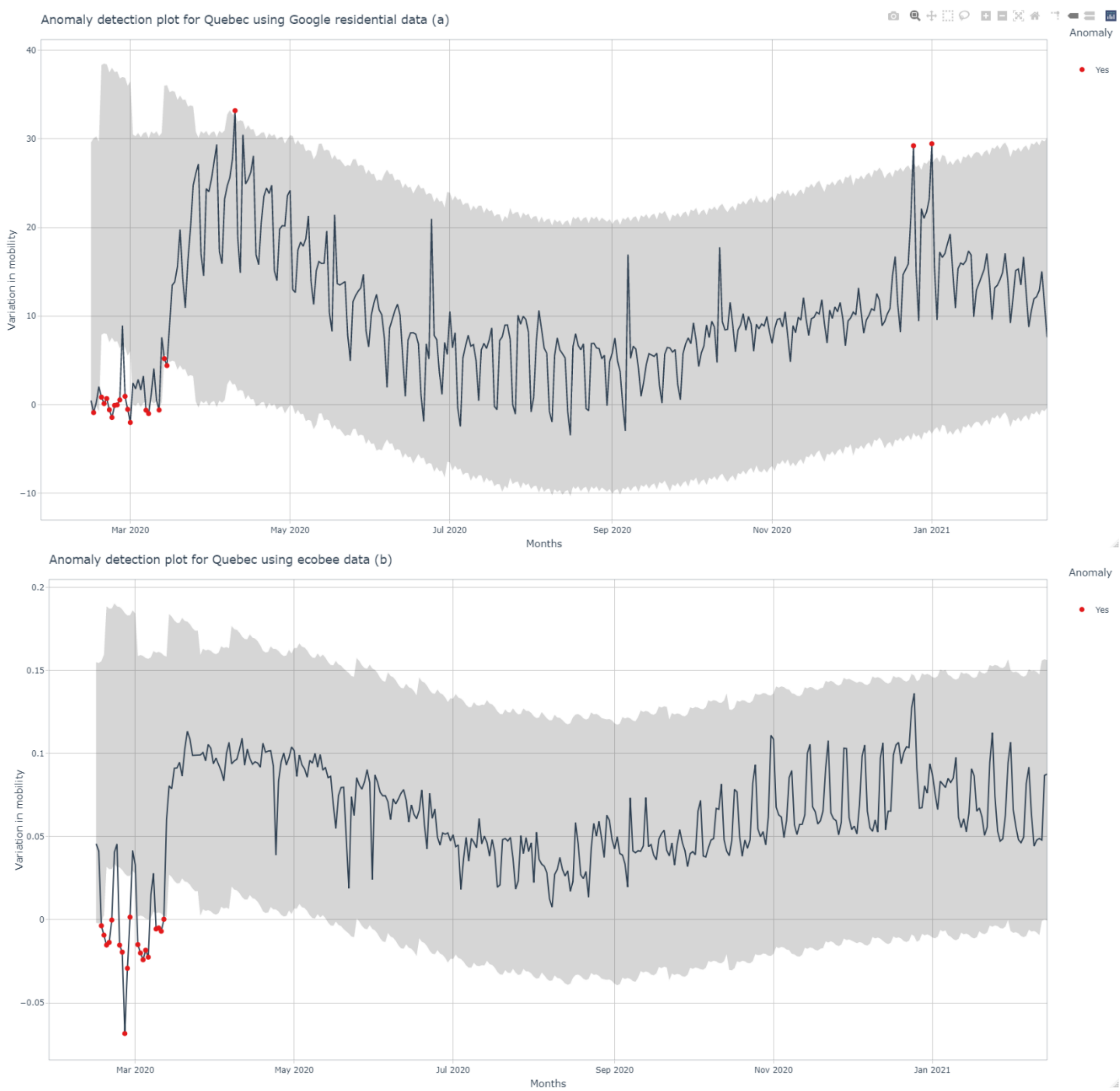


Figure 63. *Anomaly detection plot for the province of Quebec a) Google residential data b) ecobee mobility data.*

8.3.3.3 Alberta

As shown in Figure 64, Anomaly detection analysis of Google residential data for Alberta shows an anomaly at the beginning period, like ecobee data and can be matched with the dates of COVID-19 related policy changes. Festive periods like Christmas and New Year are the other anomalies captured through Google residential data, and ecobee data shows abnormal mobility detected outside and in-house. In addition to those days, ecobee also captured some other anomalies where huge variation in mobility in the in-house happened in May and June 2020 because of phase wise reopening plans.

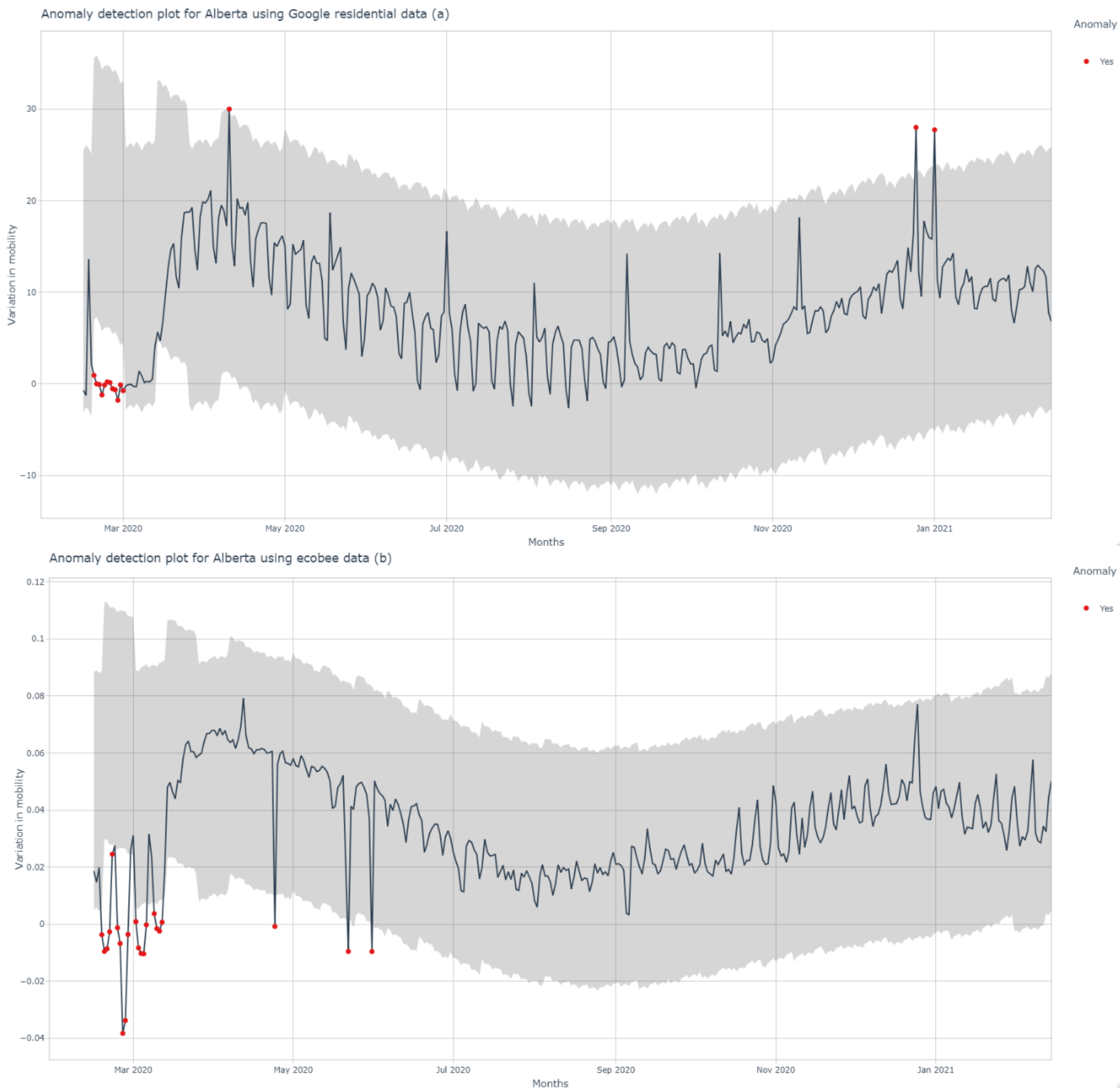


Figure 64. Anomaly detection plot for the province of Alberta a) Google residential data b) ecobee mobility data.

8.3.3.4 British Columbia

An anomaly detection analysis of Google residential data and ecobee data for British Columbia shown in Figure 65 shows an anomaly at the beginning period and can be matched with the dates of COVID-19 related policy changes. April 10, 2020, marked the beginning of using face mask while in the public place and November 11, 2020, social gathering prohibited in the province.

Ecobee data captured the date of reopening in May 2020. Festive periods like Christmas and New Year are the other anomalies captured through Google residential data shows there was abnormal mobility detected outside of the home. Ecobee could only capture Christmas as an anomaly because of no significant change in in-house mobility due to New Year's policy restrictions. A similar pattern can be seen for both datasets within this last year.

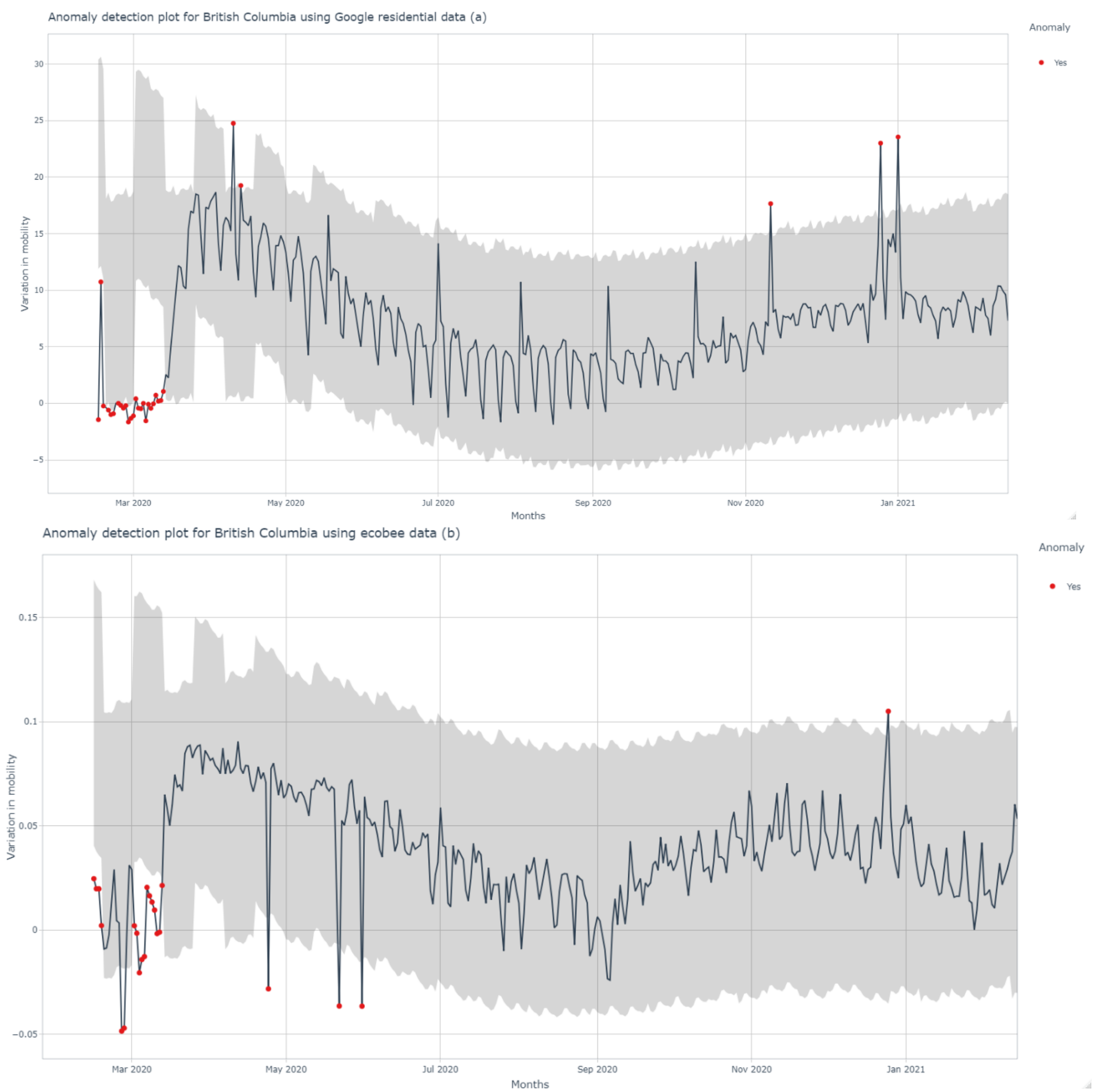


Figure 65. Anomaly detection plot for the province of British Columbia a) Google residential data b) ecobee mobility data.

8.4 Discussion

The main objective of this study was to explore the population-level human mobility data from Google map-based reports and "Donate your Data" from IoT sensors. The goal was to explore the potential use of in-house mobility data from ecobee sensors as an equivalent indicator to out-of-the-house Google mobility data. The key findings of this study are that both datasets have a lot of similarities. When compared to smartphone-based mobility, IoT-based datasets allow researchers to explore micro-level human mobility at home. The association between both datasets has been demonstrated to be statistically significant, with a Pearson's correlation coefficient of 0.7.

The temporal diagnostic analysis illustrates the temporal aspect of human mobility. Temporal diagnostic has been explored at multiple levels, exploring days of the week, month-by-month, and week-by-week respectively. Anomaly detection analysis provides evidence of the capability of both the datasets to find out deviations from the normal pattern. Interestingly, these kinds of analyses help to reinforce the existing public health surveillance mechanisms, bolstering our pandemic preparedness.

8.4.1 Micro vs Macro Human Mobility

Studies have shown that people spend more than 80% of their time in-house, [191] which never been fully explored in terms of mobility [183,449]. In-house mobility data have the capacity to measure changes in people's behaviour [188,450] but are difficult to obtain due to privacy concerns [344,451,452]. Use of smartphones and other wearables is now helping researchers to capture the individual movement within the home, which has been termed as micro-mobility [188]. With the increasing use of internet technology, the data collected by these smartphones are stored and synchronised in real time and being combined with large amounts of geographical data with timestamps [427,428]. This huge volume of the data has a potential to unravel different dimensions of human behavior and lifestyle, such as sleep [373,381], physical activity [410], sedentary behaviours,[73,378,453] and about the movement pattern [226] with a greater detail and granularity. This study shows that there is a statistically significant difference in macro mobility at the population level by days of the week, month-by-month and week-by-week analysis, whereas no statistical difference found in micro mobility found between weekdays for the three provinces, namely Ontario, Alberta, and British Columbia. Findings from Quebec show that the difference between micro mobility across different weekdays is statistically significant. These results indicate that after the implementation of the work-from-home policy and COVID-19 restrictions, the patterns of time spent in-house are almost similar for all days of the week.

When analysed for month-by-month and week-by-week, the findings are uniform for all the provinces. There is a change in mobility pattern between March and April 2020. From May 2020 to September 2020, there is a decline in macro- and micro-mobility across all provinces. The mobility increased from October 2020 to December 2020.

When analysed for month-by-month and week-by-week, the findings are uniform for all the provinces. There is a change in mobility pattern between March and April 2020. From May 2020 to September 2020, there is a decline in macro and micro mobility across all provinces. The mobility increased from October 2020 to December 2020.

Anomaly detection analysis shows the dates of policy changes and its impact on human mobility for Canada and the four provinces. Mobility data analysis captured the dates of emergency declaration, re-opening pattern and the effect of special days such as December 25 (Christmas day) and January 1 (New year eve).

Anomaly detection has its own significance in public health [390,454–457]. Anomaly detection algorithms automatically detect inconsistencies in real time and present the results in meaningful, interactive way. This can be useful for clinical data to find abnormal pattern in the cardiac health [457].

8.4.2 Human mobility as a critical indicator for public health

Individual mobility, an intrinsic property of human behaviour, is a key component in the transmission of respiratory infections like COVID-19. Interestingly, smartphones can act as sensors to capture information about the geolocations of an individual [458,459]. In situations like the COVID-19 pandemic, wherein the prevalence of positive cases were underestimated due to problems, such as low testing facilities and unwillingness of people to get them tested, early detection and identification using real-time mobility data could have been used to track the number of individuals not complying to policy restrictions [458].

Micro mobility at the population level can signify how people move around the house with time. Micro mobility can be used to quantify the activity levels of the population while simultaneously detecting the anomalies within the behaviour. In future, this kind of data source can be incorporated with existing public health surveillance mechanisms to capture real-time data about in-house mobility. Additionally, mobility data analysis in real-time can provide information about the impact of policy changes like work-from-home policy, stay at home order and further assess the compliance towards these policies [423,460,461]. The quantification of human mobility plays a significant role in both infectious and chronic diseases. For infectious diseases, mobility is directly proportional to the rate of spread of the disease whereas dwindling mobility pattern is a predisposition to various chronic diseases.

For both micro and macro human mobility, it is necessary to capture data from multiple sources and integrate them as per the research needs [462]. Although the "Stay at Home" (SAH) and "Work-from-Home" (WFH) strategies are promoted globally, it is still unclear to what extent people have changed their attitudes towards them. Since mobility patterns are closely associated with routine habits and daily chores, the implementation of such restrictive policies can be challenging [187].

8.4.2.1 Individual Mobility vs Population Mobility

Based on the theory of protection motivation, the risks perceived by the transportation sector due to the COVID-19 pandemic have been studied in the literature [187].

The primary psychological model was developed by Ronald Rogers and explains how individuals change their attitudes and behaviours when interpreting and responding to fear appeals and stress stimuli [187]. Moreover, the Health Belief Model is based upon the perceived risk, which is a marker of behavioral change. This, in turn, explains the probability of engagement in health-promoting behavior in response to stimuli. [187]. As individual mobilities constitutes the overall population mobility, it is crucial to understand the behavioural change due to mobility restrictions at the individual level.

8.4.3 IoT as a Potential Data Source for Human Mobility

IoT is a modern passive sensing tool that can quantify both macro and micro human mobility. Changes in lifestyle influence several health-related indicators such as less sleep durations, increase in sedentary behaviours, and slowing down of physical activities. These changes in the daily mobility patterns can be captured through motion sensors for further analyses. Reportedly, several studies have used Wi-Fi signals, mobile phones as a proxy for sensors, Google maps, GPS systems, and even social media-based geotagging of locations to measure human mobility at population levels [463,464]. In this context, thermostat based IoT data can open new opportunities for calculating population level mobility indicators. In future, the concept of motion sensors can be integrated in smart cities to detect overall human mobility from various sources which will, in turn, help to effectively plan and implement different public health strategies [463]. Moreover, the data collected with this strategy can also be used for digital epidemiological analysis [421].

8.4.4 Limitations of the study

As 87% of the population reside in four major provinces of Canada that were analysed in this study, there might be some other attributes that must be factored in while generalizing the results at a nation-wide scale. Moreover, a year of data may not be sufficient to represent the mobility patterns of the whole population. Though ecobee mentioned that these smart sensors cannot be activated from other stimuli by animals, rapid airflow, or other noises, but it may be difficult to completely rule out these probabilities. The absence of demographic information of users restricted our analysis by age, gender, and other socio-demographic features. Therefore, in this study, we limited our analysis to spatiotemporal dimensions.

8.5 Conclusion

In summary, population-level human mobility measurement is relatively new to the scientific community. This component has not been utilized for public health before the COVID-19 pandemic at a large scale. This study explores a new dimension of using IoT-based mobility data at the population level. Similar results can be obtained using Google-map-based mobility report. This kind of data can act as an additional indicator with social determinants of health to understand the complex issues in public health.

Chapter 9 Conclusion

9.1 Summary of the thesis

This thesis aimed answer eight research questions and they are summarized below. Firstly, I explored how researchers and public health officials are using data from Internet of Things (IoT), wearables, and mobile health apps for public health surveillance. This thesis found that data utilization from various NextGen sources is gradually increasing and is being used in various domains of population-level monitoring in small scale-studies. However, these technologies have not been applied yet at national or provincial levels.

Next, I explored how these innovative data sources have been used by Canadian researchers and public health officials. Our findings suggest that modern data sources are widely used for the prevention of chronic diseases, specifically diabetes, cardiovascular diseases, and cancer. In contrast, the use of these datasets for infectious diseases and health monitoring of seniors is still low. The study demonstrated that there is potential growth ahead of us in the use of IoT as a tool for monitoring infections diseases as COVID-19.

Our next research question focused on identifying the characteristics of data that can be used as a potential source for population-level health indicators in public health. Our analysis shows that simplicity, flexibility, sensitivity, validity, high granularity, representativeness, volume, and verity of the data are important characteristics of novel data sources in public health.

Next, I examined the viability of using smart thermostat data (IoT) data, such as data from the "Donate your Data" (DYD) program, for measuring health risk behaviours at population level. This study found a strong positive association between wearable data and smart thermostat data, proving the feasibility of using IoT data for developing health indicators. Our study demonstrated that ecobee's data can be used to calculate health indicators for healthy behaviours such as in-house physical activity, sedentary behaviour, and sleep.

The technical aspects of data wrangling such as preprocessing, data cleaning, and data manipulation before using the DYD dataset for public health surveillance are also of relevance. Importantly, my study shows that several preprocessing steps are required to make data readily usable by researchers and public health officials. These include time zone adjustment, selection of households with pre-determined conditions, and conversion of five-minute data to 30-minutes to make it more meaningful.

Next, I identified a potential method to augment the data collected by ecobee to deliver a system capable of using smart thermostat data to monitor health-risk behaviours. Near-real-time population-level health indicators can be curated from IoT data. Using the COVID-19 pandemic as a case study, I have demonstrated how the work-from-home policy significantly impacted sleep patterns and in-house and out-of-the-house stay durations of people in Canada (Ontario, Quebec, Alberta, and British Columbia). This study suggests that using IoT data for future monitoring of behavioural health risk factors can augment the traditional national and provincial health survey mechanisms.

A completely novel public health indicator called "mobility" played a vital role during the COVID-19 pandemic. In this thesis, I sought to determine how we can leverage the data collected by ecobee to deliver a system that is capable of using smart thermostat data to monitor micro-mobility at the population level. The findings from my second study show that data collected from ecobee’s smart thermostat has the potential to measure population-level micro-mobility (in-house movement) and comparable to Google’s residential mobility data.

The impact of the stay-at-home order, work-from-home policies, and emergency declaration had a significant change in the intensity of in-house and out-of-the-house population-level mobility. I believe that this research study will help policymakers to utilize real-time indicators for assessing behavioural risk factors including physical activity, sleep, sedentary behaviours, and micro-mobility in Canada. This presents a unique way to measure population-level health indicators leveraging the use of alternative data sources, *i.e.*, IoTs. The IoTs capture critical information from the communities leading to big data generation. Insights from the NextGen data sources can influence the policymaking process by providing evidence in near real-time. These modern data sources can lead to completeness of the data and fulfil the indicator gap within existing public health surveillance systems.

In summary, it can help us to understand impact of any policy change in a short period of time. A multidisciplinary approach is to identify and implement modern technological solutions to public health concerns.

9.2 Policy implications

Ideally, a good public policy has the capacity to alter socio-economic, physical, and environmental parameters of a community where people live, learn, work, and spend their life. These policies have impact on the quality of life of the whole community [465]. The ideation, development, and implementation of a public policy often occur in three institutional settings: government, public institutions, and workplaces [465]. The development of public policies follows specific pathways supported by different theories, of which the four most used theories are: 1) Stages Heuristic Model [466] 2) Multiple Streams Framework [467] 3) Advocacy Coalition Framework [468], and 4) Punctuated Equilibrium Theory [469,470]. While in the real world the mechanisms for policy development can vary from theoretical standpoints, ideally, the process of policy formulation generally involves three steps: 1) defining the problem, 2) using evidence to identify solutions and 3) engaging in the political process to influence policy outcomes [465].

9.2.1 Lockdown policies and their impact on a country

Lockdown policies have unique influence on the country's economy; both in short and long run. Each policy has its own merits and demerits [69,471]. For example, prolonged lockdown drastically changed the economy in Canada and GDP growth declined by nearly 10%, followed by a "V-shaped recovery," as shown in Figure 66 [368]. In June 2020, several provincial and federal governments started taking measures to reopen the economy after a closure for more than three months. The opening took place stepwise after careful observation of epidemiological data and trends of the positive cases, death numbers and rate of hospitalization.

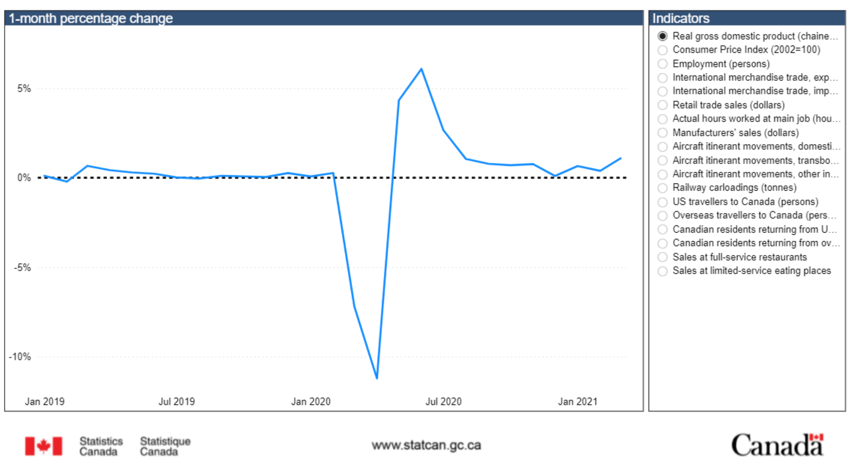


Figure 66 Dashboard to show Canadian economic situation during the COVID-19 pandemic: image extracted from Canadian Public Health Association [472].

COVID-19 pandemic response and recovery goals in Canada helped minimize morbidity and mortality while preventing the social implications of policy changes [472,473]. The Canadian government has taken unprecedented steps to respond to the pandemic that affects all dimensions of life and routine public health activities including surveillance and epidemiology, guideline development, testing, provision of emergency care, border restrictions, quarantine measures, shutting down of schools, private companies, organizations, and others. All these measures had a significant impact on the country's socio-economic features. However, evidence show that these efforts were successful in flattening the curve despite the fact that Canadian economy is expected to decline in GDP growth and experience an increase in the unemployment rate. The pandemic also exposed indicators and barriers in finance and human resource management within the national health organization. Most provincial and territorial health funds along with federal funds are not designed to provide enough support for prevention, protection, and promotion of public health [472]. To prevent this enormous impact on society, it is essential to monitor health and associated innovative indicators in real-time at population level.

9.3 Need of New Generation Monitoring Systems

In Canada, the existing public health monitoring systems for population health are well established, but with its own limitations. There is great potential to enrich the data environment by including NextGen and big data sources as part of the data available to public health scientists in Canada [28]. Several Canadian organizations, including government, ministries, department of health, and research organizations, have to address significant gaps in their data infrastructure to be able include additional data as part of their ecosystem. The following statement by Canadian Institute for Health Information "*Better data, better patient outcomes*" [474] showcases the importance of a more diverse data portfolio, including NextGen data [28].

The use of advanced algorithms as artificial intelligence applied to big datasets in public health requires the implementation of data ecosystems capable of handling the generation, collection, preparation, and integration of these data sources while respecting the data owner's (e.g., research participant, patient, citizen) safety, security and maintaining trust [474].

9.4 Evidence Building with Time as A Critical Factor

The COVID-19 pandemic taught us the value of timely responses [472]. In such a situation, using real-time data to predict future anomalies and prepare society to handle them with utmost care is essential [9]. Using real-time data generated by millions of users of technologies, including mobile phones and the internet is in utmost demand, as it provides us with near-real time information on human behaviour and public health [475].

9.4.1 Use of Real-Time Data

Mobile technology, wearables, and the internet created an environment where innumerable data points are generated from various information sources, which can be directly or indirectly related to health [311,476–478]. Variables of interest like steps count [479–482], vital parameters of human health [483–485], or geolocation [226] are all part of these diverse data sources. Through the integration multiple data sources, researchers can extract a more comprehensive view of population health since large, high-frequency, longitudinal datasets are available from participants [24,183,228]. The observation of these patterns can help us understand the inherent trend of changes within the community. One important source that has been used in other areas of policy research includes the use of social media data [78,343,486–488], which has been demonstrated to have the potential to detect the outbreak of infectious diseases.

9.4.2 Use of Time-Sensitive Data

The data generated at one point reduces or loses its value over time. Use of specific time-sensitive data can create useful insights. Otherwise, the same data would be of no use after a certain period. Innovative data integration mechanisms and analysis can add value using these data points at the right time.

The Canadian Institute for Health Information (CIHI) consulted with health system leaders across Canada and reviewed relevant literature to identify practical actions to systematically improve the use of these data sources in Canada [489]. These consultations demonstrated that data and information contribute towards improving health in innovative ways [489]. A consistent theme emerged from this consultation suggests the need to improve the quality and consistency of health data governance practices across organizations seeking to share data and realize its strategic value [489]. The essential steps to address the existing challenges are reviewing and describing the meaning, completeness, and quality of data assets to optimize their use internally and by trusted partners [489]. Prioritizing to share data, and algorithms to achieve common goals whilst maintaining robust privacy, security controls and harmonizing policies, practices and its standards were pragmatic and valuable [489]. Earning and retaining public trust in an important step in the establishment of proper data governance mechanisms [489].

9.5 Respect for Data

Statistics Canada respects the value of data. As per the mandate of Statistics Canada, each data point has three critical features: privacy of the people to whom the data belongs, security of information throughout the data lifecycle, and confidentiality of information [490]. Therefore, safe, secure, and privacy-preserving data environment needs to be implemented to maintain trust in the system. Technological solutions such as distributed ledgers (e.g., Blockchain, Ethereum) can help implement a secure ecosystem. [46].

In summary, policymakers should use the latest available evidence and technology to inform the development of new policies [491,492]. The evidence-based policymaking process can save more lives, reduce morbidity, and improve the quality of life when compared to the traditional policy making process [40,493]. The COVID-19 pandemic helped identify weak points in the existing health system and there is an urgent need to reframe the policies, addressing the problems within the health sector [69,494]. Interestingly, Next Gen data sources can influence the policymaking process by providing greater insight to near real-time situation of the problem, completeness of the data and indicator gap, and understanding impact of any policy change in a short period of time [28]. IoT as a unique data source has the potential to generate large volumes of data on a short timeframe, compiling critical information directly from communities [28,157]. Proper utilization of those ubiquitous datasets with algorithms based on artificial intelligence may fill the gaps and predict future outbreaks [289].

A multidisciplinary approach to this challenge requires expertise from a diverse team, starting from a solid epidemiological component, allied with expertise in human health, and a solid understanding of the technological solutions used to answer relevant public health questions [177,344,495,496]. Recent literature shows that researchers worldwide use Nextgen data sources in combination with traditional survey data, to bridge the data gap [28]. Several technologies have been used to help public health decision-makers quickly answer specific unique policy gaps for infectious and chronic diseases [491].

There is a need to incorporate evidence to policymaking process to reduce the impact of future pandemics at the scale of COVID-19 and might save numerous lives. Data generated from Internet of Things (IoT) has the potential to create evidence and enrich population-level monitoring systems, which can be used in evidence-based policy making. The access to NextGen data will benefit the policymakers and increase the efficiency and effectiveness of health policy in the future [28].

9.6 Strengths of this Research

This IoT data serves various advantages such as zero effort for data collection, near real-time output, and use of continuously expanding data source. In terms of representativeness, the data is well distributed amongst all the provinces and territories of Canada and provide an overall view of the community. The key findings of this research can be summarised as follows:

- NextGen data sources are helpful for public health surveillance
- Near real-time population level monitoring can be possible using NextGen data sources
- IoT is crucial within the NextGen data source for real-time data
- These unique data sources can be used as a standalone component or in conjunction with traditional public health surveillance systems
- The impact of public health policy can be monitored to measure the intended outcomes to some of the targeted and proxy indicators
- ecobee's data can be helpful to monitor special groups of the population, especially vulnerable individuals including persons with locomotor disability, elderly persons living alone, pregnant women, persons with multiple chronic diseases including cancer, or persons with mental health issues.
- Importantly, the cost-effectiveness of this technology for the public health system plays a critical role in adoption. The use of "user-generated/passive monitoring" can be more cost-effective than traditional surveys. This method of data collection, analysis, and reporting comes with the benefit of low cost. As mentioned above, the data collected for this research comes from the already existing infrastructure; for smart thermostats no special efforts are required from the researchers in terms of purchasing sensors, deploying them for data collection, and training of data collectors.

9.7 Challenges and Limitations

The challenges faced within this thesis can be broadly divided into technical, data analytics and ethics domains. The technical challenges included volume of the data, verity of the data, access to data, and structure of the data. For example, ecobee shares its data through the DYD program to researchers in different formats and changing database, starting from .csv files to Google BigQuery. This, in turn, requires a unique skillset to extract, load, and manipulate datasets along with data processing and cleaning before the statistical analyses can be performed. Owing to the large volumes of data obtained, advanced computing infrastructure and resources are important resources that are needed.

In terms of data analytics domain, the main challenge is the absence of detailed demographic information about the residents. While the data informs about the geographic location, important population characteristics like age, sex, education, occupation, marital status, and socio-economic status are not present in the dataset. Since this dataset was not collected for the purpose of monitoring health, there is an absence of data about health conditions, including chronic diseases. Notably, the data quality and missing values needs to be considered as there is a considerable amount of missing data. The reasons for the missing data in IoT may be attributed to power outages, interruptions in the internet connectivity, and other technical and non-technical issues.

Despite the anonymous nature of data, the concerns of privacy, ethics, and security must also be factored in. As the data source is a private company, there are concerns whether this data will be available in future or not. Various factors such as the relationship between the private company, academic research environment, and government policy needs to be considered to ensure continuous availability of this data.

Additionally, the use of smart thermostats is more commonly observed among individuals with higher socio-economic status compared to others. Therefore, the result from this study needs to be generalized with some considerations regarding the individuals with low-income backgrounds. Furthermore, individuals with higher socio-economic status will have higher probability of working from home compared to others and consequently, their chances of spending time in-house during the pandemic.

Additional limitations of this thesis include: (1) the population distribution of ecobee smart thermostat users may differ from the general population; (2) as the data collection was restricted to in-house movements only, it lacks 24-hour movement data; (3) because the sensors only capture the activity as a binary output, they lack the capacity to assess the type of the activity being performed by the residents; (4) despite having sensors designed and tuned to capture mostly human movement, there is still probability of greater noise or errors in a house with a pet than a house without a pet; (5) the absence of demographic information and health information in the DYD limits the analysis to broad population level rather than stratified by key demographic factors such as age and gender; (6) as there is a time gap between the data generation and providing access to the researchers, the analysis are not real time in nature; (7) health indicators measured in this thesis were proxies to the real indicators for sleep, physical activity, in-house and out-of-the-house mobility pattern; (8) individual level sleep, physical activity, sedentary behaviour can only be analyzed using data from households with single individuals, which limits the applicability of these results.

9.8 Future Research Opportunities

In this thesis, I have presented evidence that IoT will be a critical component of the future public health surveillance. In the coming years, integrating IoT data within existing public health surveillance system will strengthen the evidence-based monitoring mechanism. Providing real time and critical public health insights from the NextGen data sources will help the policy makers to improve the existing policy accordingly.

Adding traditional and modern data sources will help to improve comprehensiveness of the data and indicators. With the combination of time, resources, funding, and data from various sources, Canada can develop a comprehensive a public health monitoring platform. Integrating these data in a secured environment and calculating comprehensive public health indicators will provide our public health system with the necessary responsiveness for dealing with future pandemics. This kind of research can be extended to study health indicators for specific kind of population such as people with one or more chronic conditions, persons with disabilities, other vulnerable segments of the population including older adults and people staying alone and so on.

References

1. Public Health Agency of Canada. The Chief Public Health Officer's report on the state of public health in Canada 2008 – What is public health? Ottawa (ON): 2008.
2. Centers for Disease Control and Prevention (CDC). Introduction to Public Health [Internet]. 2018 [cited 2019 Oct 15];Available from: https://www.cdc.gov/publichealth101/public-health.html
3. John Hopkins Bloomberg School of Public Health. What is Public Health - About - Johns Hopkins Bloomberg School of Public Health [Internet]. 2019 [cited 2019 Sep 4];Available from: https://www.jhsph.edu/about/what-is-public-health/index.html
4. Dyson T. Population and Development: The Demographic Transition. [Internet]. Zed; 2010 [cited 2019 Sep 3]. Available from: http://dx.doi.org/10.1177/146499341001100307
5. Santosa A, Wall S, Fottrel E, Högberg U, Byass P. The development and experience of epidemiological transition theory over four decades: A systematic review. Glob Health Action 2014;7(SUPP.1).
6. Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet [Internet] 1997 [cited 2017 Jul 28];349(9063):1436–42. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9164317
7. Defo BK. Demographic, epidemiological, and health transitions: Are they relevant to population health patterns in Africa? Glob Health Action 2014;7(SUPP.1).
8. Kickbusch I, Piselli D, Agrawal A, Balicer R, Banner O, Adelhardt M, et al. The Lancet and Financial Times Commission on governing health futures 2030: growing up in a digital world. Lancet [Internet] 2021;Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673621018249

9. Chretien JP, Tomich NE, Gaydos JC, Kelley PW. Real-time public health surveillance for emergency preparedness [Internet]. Am. J. Public Health2009 [cited 2019 Oct 15];99(8):1360–3. Available from: /pmc/articles/PMC2707469/?report=abstract
10. Choi BCK. The Past, Present, and Future of Public Health Surveillance. Scientifica (Cairo) [Internet] 2012 [cited 2018 Jun 5];2012:1–26. Available from: http://www.hindawi.com/journals/scientifica/2012/875253/
11. Hall HI, Correa A, Paula W. Yoon, Christopher R. Braden. Lexicon, Definitions, and Conceptual Framework for Public Health Surveillance. Morb Mortal Wkly Rep [Internet] 2012 [cited 2018 Dec 11];61(03):10–4. Available from: https://www.cdc.gov/mmwr/preview/mmwrhtml/su6103a3.htm
12. Choi BCK. What Could Be Future Scenarios?-Lessons from the History of Public Health Surveillance for the Future. AIMS public Heal [Internet] 2015 [cited 2018 Oct 3];2(1):27–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29546093
13. Lombardo JS, Buckeridge DL. Disease surveillance : a public health informatics approach [Internet]. Wiley-Interscience; 2007 [cited 2018 May 16]. Available from: https://www.wiley.com/en-us/Disease+Surveillance%3A+A+Public+Health+Informatics+Approach-p-9780470068120
14. Tabacchi G, Jemni M, Viana JL, Bianco A. Adolescence surveillance system for obesity prevention (ASSO) in EUROPE: A pioneering project to prevent obesity using e-technology [Internet]. In: Censorship, Surveillance, and Privacy: Concepts, Methodologies, Tools, and Applications. IGI Global; 2018 [cited 2019 Sep 4]. page 2088–113.Available from: http://services.igi-

global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-7113-1.ch104

15. Santillana M, Zhang DW, Althouse BM, Ayers JW. What Can Digital Disease Detection Learn from (an External Revision to) Google Flu Trends? Am J Prev Med [Internet] 2014 [cited 2018 Jul 6];47(3):341–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24997572

16. Lazer D, Kennedy R, King G, Vespignani A. The Parable of Google Flu: Traps in Big Data Analysis. Science (80-) [Internet] 2014;343(6176):1203–5. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1248506

17. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining Search, Social Media, and Traditional Data Sources to Improve Influenza Surveillance. PLOS Comput Biol [Internet] 2015 [cited 2018 May 10];11(10):e1004513. Available from: http://dx.plos.org/10.1371/journal.pcbi.1004513

18. Pogue D. What Fitbit discovered from 6 billion nights of sleep data [Internet]. Yahoo2018 [cited 2020 Mar 2];Available from: https://uk.finance.yahoo.com/news/exclusive-fitbits-6-billion-nights-sleep-data-reveals-us-110058417.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cud2Vmb3J1bS5vcmcv&guce_referrer_sig=AQAAABstsMCo-97qDi46Zbs49ibnguljjRq8MoYfmG2AbL91TIL3PrHpIbH0SES9PfTsokd

19. Fitbit. Early Findings from Fitbit COVID-19 Study Suggest Fitbit Devices Can Identify Signs of Disease at Its Earliest Stages - Fitbit Blog [Internet]. 2020 [cited 2020 Sep 10];Available from: https://blog.fitbit.com/early-findings-covid-19-study/

20. Rezaei N, Grandner MA. Changes in sleep duration, timing, and variability during the COVID-19 pandemic: Large-scale Fitbit data from 6 major US cities. Sleep Heal

[Internet] 2021 [cited 2021 Mar 31];7(3):303–13. Available from: https://reader.elsevier.com/reader/sd/pii/S2352721821000292?token=A6EEEF4493621C9255155AE303FB4A6F8F32623DA93E4119FE9E827F8BEBA621A8B05C9AB10930B623D7C568DA6E5EA2&originRegion=us-east-1&originCreation=20210331191121

21. Rothman KJ, Greenland S, Lash TL. Modern epidemiology [Internet]. 3rd ed. Philadelphia: Oxford University Press (OUP); 2008 [cited 2018 May 16]. Available from: https://www.amazon.ca/Studyguide-Epidemiology-Rothman-Kenneth-9780781755641/dp/1619058057

22. Karunanithi M, Zhang Q. An Innovative Technology to Support Independent Living: The Smarter Safer Homes Platform. Stud Health Technol Inform [Internet] 2018 [cited 2018 Aug 20];246:102–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29507263

23. Swan M. Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0. J Sens Actuator Networks 2012, Vol 1, Pages 217-253 [Internet] 2012 [cited 2021 Aug 16];1(3):217–53. Available from: https://www.mdpi.com/2224-2708/1/3/217/htm

24. Kostkova P. A Roadmap to Integrated Digital Public Health Surveillance: the Vision and the Challenges [Internet]. In: International World Wide Web Conference Committee. 2013 [cited 2018 Jun 18]. page 687–93.Available from: http://www.phac-aspc.gc.ca/gphin/

25. Hilty DM, Armstrong CM, Luxton DD, Gentry MT, Krupinski EA. A Scoping Review of Sensors, Wearables, and Remote Monitoring For Behavioral Health: Uses, Outcomes, Clinical Competencies, and Research Directions. J Technol Behav Sci [Internet] 2021;1–36. Available from: https://doi.org/10.1007/s41347-021-00199-2

26. Hamilton JJ, Turner K, Lichtenstein Cone M. Responding to the pandemic: Challenges

with public health surveillance systems and development of a COVID-19 national surveillance case definition to support case-based morbidity surveillance during the early response. J Public Heal Manag Pract [Internet] 2021 [cited 2021 Aug 16];27:S80–6. Available from: https://journals.lww.com/jphmp/Fulltext/2021/01001/Responding_to_the_Pandemic__Challenges_With_Public.14.aspx

27. Sahu KS, Oetomo A, Morita PP. Enabling Remote Patient Monitoring Through the Use of Smart Thermostat Data in Canada: Exploratory Study. JMIR mHealth uHealth [Internet] 2020;8(11):e21016. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33216001

28. Sahu KS, Majowicz SE, Dubin JA, Morita PP. NextGen Public Health Surveillance and the Internet of Things (IoT). Front public Heal [Internet] 2021;9:756675. Available from: https://www.frontiersin.org/articles/10.3389/fpubh.2021.756675/full

29. Canadian Public Health Association. Public Health: A conceptual framework [Internet]. 2017 [cited 2021 May 25]. Available from: www.cpha.ca

30. Institute of Medicine (US). The Future of Public Health [Internet]. National Academies Press (US); 1988 [cited 2019 Jan 24]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25032306

31. Roberts KC, Rao DP, Bennett TL, Loukine L, Jayaraman GC. Prevalence and patterns of chronic disease multimorbidity and associated determinants in Canada. Heal Promot Chronic Dis Prev Canada 2015;35(6):87–94.

32. Winslow C-EA. The untilled fields of Public Health. Science (80-) [Internet] 1920 [cited 2018 Nov 9];51(1306):23–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17838891

33. Rothstein MA. Rethinking the meaning of public health. In: Journal of Law, Medicine and Ethics. Blackwell Publishing Inc.; 2002. page 144–9.

34. Dobson A. Redefining the unacceptable. Aust J Public Health [Internet] 2010 [cited 2019 Oct 15];18(1):9–14. Available from: http://doi.wiley.com/10.1111/j.1753-6405.1994.tb00187.x

35. Rosner D, Fried LP. Traditions, transitions, and transfats: new directions for public health. Public Health Rep [Internet] 2010 [cited 2018 Nov 9];125(1):3–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20402191

36. Choi J, Cho Y, Shim E, Woo H. Web-based infectious disease surveillance systems and public health perspectives: a systematic review. BMC Public Health [Internet] 2016 [cited 2018 May 10];16(1):1238. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27931204

37. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet [Internet] 2020;396(10258):1204–22. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673620309259

38. Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392(10159):1736–88.

39. Atella V, Piano Mortari A, Kopinska J, Belotti F, Lapi F, Cricelli C, et al. Trends in age-related disease burden and healthcare utilization. Aging Cell [Internet] 2019 [cited 2019

Sep 6];18(1):e12861. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30488641

40. Brownson RC, Fielding JE, Green LW. Building Capacity for Evidence-Based Public Health: Reconciling the Pulls of Practice and the Push of Research. Annu Rev Public Health [Internet] 2018 [cited 2021 Sep 13];39(1):27–53. Available from: http://www.annualreviews.org/doi/10.1146/annurev-publhealth-040617-014746

41. Choi BC, Pak AW. Lessons for surveillance in the 21st century: a historical perspective from the past five millennia. Soz Praventivmed [Internet] 2001 [cited 2018 Oct 3];46(6):361–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11851070

42. World Health Organization. Public health surveillance. WHO [Internet] 2012 [cited 2018 Oct 1];Available from: http://www.who.int/topics/public_health_surveillance/en/

43. Centers for Disease Control and Prevention (CDC). About BRFSS [Internet]. 2019 [cited 2019 Feb 7];Available from: https://www.cdc.gov/brfss/about/index.htm

44. Park H-A, Jung H, On J, Park SK, Kang H. Digital Epidemiology: Use of Digital Data Collected for Non-epidemiological Purposes in Epidemiological Studies. Healthc Inform Res [Internet] 2018 [cited 2019 Apr 25];24(4):253. Available from: https://synapse.koreamed.org/DOIx.php?id=10.4258/hir.2018.24.4.253

45. Steele R, Clarke A, Steele R, Clarke A. The Internet of Things and Next-generation Public Health Information Systems. Commun Netw [Internet] 2013 [cited 2019 Sep 6];05(03):4–9. Available from: http://www.scirp.org/journal/doi.aspx?DOI=10.4236/cn.2013.53B1002

46. Bellod Cisneros JL, Cisneros JLB, Aarestrup FM, Lund O. Public Health Surveillance using Decentralized Technologies. Blockchain Healthc Today [Internet] 2018 [cited 2018 Oct 3];1(0). Available from: https://blockchainhealthcaretoday.com/index.php/journal/article/view/17

47. McNabb SN, Ryland P, Sylvester J, Shaikh A. Informatics enables public health surveillance. J Heal Spec [Internet] 2017 [cited 2018 Oct 3];5(2):55. Available from: http://www.thejhs.org/text.asp?2017/5/2/55/205075

48. Murray CJL, Barber RM, Foreman KJ, Ozgoren AA, Abd-Allah F, Abera SF, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. Lancet [Internet] 2015 [cited 2019 Apr 25];386(10009):2145–91. Available from: https://www.sciencedirect.com/science/article/pii/S014067361561340X

49. Jayasinghe S. Conceptualising population health: from mechanistic thinking to complexity science. Emerg Themes Epidemiol [Internet] 2011 [cited 2018 May 16];8(2). Available from: https://ete-online.biomedcentral.com/track/pdf/10.1186/1742-7622-8-2

50. World Health Organization. Closing the gap in a generation: health equity through action on the social determinants of health. Closing gap a Gener Heal equity through action Soc Determ Heal [Internet] 2008 [cited 2019 Feb 22];Available from: https://www.cabdirect.org/cabdirect/abstract/20083249386

51. Martin GM, Bergman A, Barzilai N. Genetic Determinants of Human Health Span and Life Span: Progress and New Opportunities. PLoS Genet [Internet] 2007 [cited 2019 Sep 4];3(7):e125. Available from: https://dx.plos.org/10.1371/journal.pgen.0030125

52. Klimentidis YC, Raichlen DA, Bea J, Garcia DO, Wineinger NE, Mandarino LJ, et al. Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. Int J Obes [Internet] 2018 [cited 2019 Feb 21];42(6):1161–76. Available from:

http://www.nature.com/articles/s41366-018-0120-3

53. Room R, Babor T, Rehm J. Alcohol and public health. Lancet [Internet] 2005 [cited 2019 Sep 4];365(9458):519–30. Available from: https://www.sciencedirect.com/science/article/pii/S0140673605178702?via%3Dihub

54. Eysenck HJ. Smoking, Health & Personality [Internet]. Routledge; 2018 [cited 2019 Sep 4]. Available from: https://www.taylorfrancis.com/books/9781351306478

55. Field KS, Briggs DJ. Socio-economic and locational determinants of accessibility and utilization of primary health-care. Heal Soc Care Community 2001;9(5):294–308.

56. Ahnquist J, Wamala SP, Lindstrom M. Social determinants of health – A question of social or economic capital? Interaction effects of socioeconomic factors on health outcomes. Soc Sci Med [Internet] 2012 [cited 2019 Apr 29];74(6):930–9. Available from: https://www.sciencedirect.com/science/article/pii/S0277953612000238

57. Hankivsky O, Christoffersen A. Intersectionality and the determinants of health: a Canadian perspective. Crit Public Health [Internet] 2008 [cited 2019 Mar 11];18(3):271–83. Available from: http://www.tandfonline.com/doi/abs/10.1080/09581590802294296

58. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. J Epidemiol Community Health [Internet] 2001 [cited 2019 Apr 29];55(2):111–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11154250

59. Thomas G. Savel, Seth Foldy. The Role of Public Health Informatics in Enhancing Public Health Surveillance. MMWR [Internet] 2012 [cited 2018 Dec 11];61(Supplement):21–3. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.259.9548&rep=rep1&type=pdf

#page=22

60. Krishnamurthy RS, Louis ME St. Informatics and the Management of Surveillance Data [Internet]. Oxford Scholarship Online; 2010 [cited 2018 May 16]. Available from: http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195372922.001.0001/acprof-9780195372922-chapter-5

61. Centers for Disease Control. Public Health Surveillance in the United States: Evolution and Challenges* [Internet]. 1988 [cited 2018 Oct 1]. Available from: https://www.cdc.gov/mmwr/preview/mmwrhtml/su6103a2.htm

62. Hancock T, Labonte R, Edwards R. Indicators that Count! Measuring Population Health at the Community Level. Can J Public Heal [Internet] 1999 [cited 2019 Apr 30];90(1):S22–6. Available from: https://link.springer.com/article/10.1007/BF03403574

63. Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? Lancet [Internet] 2005 [cited 2019 Apr 25];366(9496):1578–82. Available from: https://www.sciencedirect.com/science/article/pii/S0140673605673412

64. Public Health Agency of Canada. Chronic Disease Risk Factors - Canada.ca [Internet]. 2015 [cited 2019 Apr 30];Available from: https://www.canada.ca/en/public-health/services/chronic-diseases/chronic-disease-risk-factors.html

65. World Health Organization. Burden: mortality, morbidity and risk factors [Internet]. In: Global status report on noncommunicable diseases 2010. 2011 [cited 2019 Apr 30]. Available from: https://www.who.int/nmh/publications/ncd_report_chapter1.pdf?ua=1

66. Honigsbaum M. Pandemic [Internet]. Lancet2009 [cited 2021 May 13];373(9679):1939. Available from: www.thelancet.com

67. Qiu W, Rutherford S, Mao A, Chu C. The Pandemic and its Impacts. Heal Cult Soc

[Internet] 2017 [cited 2021 May 13];9(0):1–11. Available from: http://hcs.pitt.edu/ojs/index.php/hcs/article/view/221/280

68. Flecknoe D, Charles Wakefield B, Simmons A. Plagues & wars: the 'Spanish Flu' pandemic as a lesson from history. Med Confl Surviv [Internet] 2018;34(2):61–8. Available from: https://www.tandfonline.com/doi/full/10.1080/13623699.2018.1472892

69. Berger L, Berger N, Bosetti V, Gilboa I, Hansen LP, Jarvis C, et al. Rational policymaking during a pandemic. Proc Natl Acad Sci [Internet] 2021 [cited 2021 Sep 13];118(4). Available from: https://www.pnas.org/content/118/4/e2012704118

70. Joint WHO-China Study Team. WHO-convened global study of origins of SARS-CoV-2: China Part [Internet]. 2021 [cited 2021 Jul 6];Available from: https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part

71. World Health Organization. Listings of WHO's response to COVID-19 [Internet]. 2020 [cited 2021 Jul 6];Available from: https://www.who.int/news/item/29-06-2020-covidtimeline

72. World Health Organization. WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data [Internet]. 2021 [cited 2021 Jul 6];Available from: https://covid19.who.int/

73. Stockwell S, Trott M, Tully M, Shin J, Barnett Y, Butler L, et al. Changes in physical activity and sedentary behaviours from before to during the COVID-19 pandemic lockdown: a systematic review. BMJ Open Sport Exerc Med [Internet] 2021 [cited 2021 May 13];7(1):e000960. Available from: http://dx.doi.org/10.1136/bmjsem-2020-000960

74. Schuch FB, Bulzing RA, Meyer J, Vancampfort D, Firth J, Stubbs B, et al. Associations of

moderate to vigorous physical activity and sedentary behavior with depressive and anxiety symptoms in self-isolating people during the COVID-19 pandemic: A cross-sectional survey in Brazil. Psychiatry Res 2020;292.

75. Wolf S, Seiffer B, Zeibig J-M, Welkerling J, Brokmeier L, Atrott B, et al. Is Physical Activity Associated with Less Depression and Anxiety During the COVID-19 Pandemic? A Rapid Systematic Review. Sports Med [Internet] 2021 [cited 2021 May 13];Available from: http://www.ncbi.nlm.nih.gov/pubmed/33886101

76. World Health Organization. Public health surveillance for COVID-19. 2020.

77. Declich S, Carter AO. Public health surveillance: historical origins, methods and evaluation. Bull World Health Organ [Internet] 1994 [cited 2018 Oct 1];72(2):285–304. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8205649

78. VELASCO E, AGHENEZA T, DENECKE K, KIRCHNER G, ECKMANNS T. Social Media and Internet-Based Data in Global Systems for Public Health Surveillance: A Systematic Review. Milbank Q [Internet] 2014 [cited 2018 May 10];92(1):7–33. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24597553

79. Althoff T, Horvitz E, White RW, Zeitzer J. Harnessing the Web for Population-Scale Physiological Sensing [Internet]. In: Proceedings of the 26th International Conference on World Wide Web. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee; 2017 [cited 2020 Nov 10]. page 113–22.Available from: https://dl.acm.org/doi/10.1145/3038912.3052637

80. Gittelman S, Lange V, Gotway Crawford CA, Okoro CA, Lieb E, Dhingra SS, et al. A New Source of Data for Public Health Surveillance: Facebook Likes. J Med Internet Res [Internet] 2015 [cited 2019 Jan 18];17(4):e98. Available from:

http://www.jmir.org/2015/4/e98/

81. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a Tool for Health Research: A Systematic Review. Am J Public Health [Internet] 2017 [cited 2018 Dec 24];107(1):e1–8. Available from: http://ajph.aphapublications.org/doi/10.2105/AJPH.2016.303512

82. Bagroy S, Kumaraguru P, De Choudhury M. A Social Media Based Index of Mental Well-Being in College Campuses [Internet]. New York, New York, USA: ACM Press; 2017 [cited 2018 Mar 20]. Available from: http://dl.acm.org/citation.cfm?doid=3025453.3025909

83. Butler AL, Smith M, Jones W, Adair CE, Vigod S, Kurdyak P, et al. Multi-province epidemiological research using linked administrative data: a case study from Canada. Int J Popul Data Sci [Internet] 2018;3(3). Available from: https://ijpds.org/article/view/443

84. Lix L, Reimer K. The Canadian Chronic Disease Surveillance System: A Distributed Surveillance Model. Online J Public Health Inform [Internet] 2017 [cited 2018 Oct 3];9(1). Available from: http://journals.uic.edu/ojs/index.php/ojphi/article/view/7726

85. Frieden TR, Harold Jaffe DW, Thacker SB, Moolenaar RL, Lee LM, Meyer PA, et al. CDC's Vision for Public Health Surveillance in the 21st Century [Internet]. 2012 [cited 2019 Sep 6]. Available from: https://www.cdc.gov/mmwr/pdf/other/su6103.pdf

86. Rosenman R, Tennekoon V, Hill LG. Measuring bias in self-reported data. Int J Behav Healthc Res [Internet] 2011 [cited 2019 Sep 6];2(4):320–32. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25383095

87. IOM (Institute of Medicine). A Nationwide Framework for Surveillance of Cardiovascular and Chronic Lung Diseases [Internet]. Washington, D.C.: National Academies Press;

2011. Available from: https://www.ncbi.nlm.nih.gov/books/NBK83166/pdf/Bookshelf_NBK83166.pdf

88. Public Health Ontario. Gaps in public health indicators and data in Ontario. Revised ed [Internet]. Toronto,ON: 2016 [cited 2019 Feb 22]. Available from: https://www.publichealthontario.ca/-/media/documents/public-health-indicator-gaps.pdf?la=en

89. Sagiroglu S, Sinanc D. Big data: A review [Internet]. In: 2013 International Conference on Collaboration Technologies and Systems (CTS). IEEE; 2013 [cited 2019 Jan 24]. page 42–7.Available from: http://ieeexplore.ieee.org/document/6567202/

90. Dalton CB. Enablers of innovation in digital public health surveillance: Lessons from flutracking. Int Health [Internet] 2017 [cited 2018 May 11];9(3):145–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28582555

91. Gomide J, Veloso A, Meira W, Almeida V, Benevenuto F, Ferraz F, et al. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter [Internet]. In: Proceedings of the 3rd International Web Science Conference on - WebSci '11. New York, New York, USA: ACM Press; 2011 [cited 2019 Feb 4]. page 1–8.Available from: http://dl.acm.org/citation.cfm?doid=2527031.2527049

92. Meltzer LJ, Hiruma LS, Avis K, Montgomery-Downs H, Valentin J. Comparison of a commercial accelerometer with polysomnography and actigraphy in children and adolescents. Sleep 2015;38(8):1323–30.

93. Dao P, Maskevich S, Jumabhoy R, Drummond S, Stout J. 1163 USING WRIST-WORN ACTIGRAPHIC DEVICES TO MEASURE SLEEP IN PEOPLE WITH HUNTINGTON'S DISEASE. Sleep 2017;40(suppl_1):A434–A434.

94. Seeskin ZH, LeClere F, Ahn J, Williams J. Uses of Alternative Data Sources for Public Health Statistics and Policymaking: Challenges and Opportunities [Internet]. 2018 [cited 2020 Aug 24]. Available from: https://www.norc.org/PDFs/Publications/SeeskinZ_Uses of Alternative Data Sources_2018.pdf

95. Ramirez M, Wu S, Jin H, Ell K, Gross-Schulman S, Myerchin Sklaroff L, et al. Automated Remote Monitoring of Depression: Acceptance Among Low-Income Patients in Diabetes Disease Management. JMIR Ment Heal [Internet] 2016 [cited 2018 Feb 9];3(1):e6. Available from: http://mental.jmir.org/2016/1/e6/

96. Farina N, Lowry RG. The validity of consumer-level activity monitors in healthy older adults in free-living conditions. J Aging Phys Act 2018;26(1):128–35.

97. Huang Y, Xu J, Yu B, Shull PB. Validity of FitBit, Jawbone UP, Nike+ and other wearable devices for level and stair walking. Gait Posture [Internet] 2016 [cited 2019 Sep 11];48:36–41. Available from: https://www.sciencedirect.com/science/article/pii/S096663621630042X

98. Beekman E, Braun SM, Ummels D, van Vijven K, Moser A, Beurskens AJ. Validity, reliability and feasibility of commercially available activity trackers in physical therapy for people with a chronic disease: A study protocol of a mixed methods research. Pilot Feasibility Stud 2017;3(1).

99. Sushames A, Edwards A, Thompson F, McDermott R, Gebel K. Validity and reliability of fitbit flex for step count, moderate to vigorous physical activity and activity energy expenditure. PLoS One 2016;11(9).

100. Boger J, Young V, Hooey J, Jiancaro T, Mihailidis A, Hoey J, et al. Zero Effort Technologies: Considerations, Challenges, and Use in Health, Wellness, and

Rehabilitation,Second Edition. Synth Lect Assist Rehabil Heal Technol [Internet] 2018 [cited 2018 Jun 26];8(1):i–118. Available from: https://slack-files.com/files-pri-safe/T2FBH4M4N-FAG13RBJ5/zet-_j.boger.pdf?c=1530040203-35a51635ccc6e2dcb8cbd10635920bdf60ed16ff

101. Nguyen HH, Mirza F, Naeem MA, Nguyen M. A Review on IoT Healthcare Monitoring Applications and a Vision for Transforming Sensor Data into Real-time Clinical Feedback [Internet]. In: IEEE 21st International Conference on Computer Supported Cooperative Work in Design. 2017 [cited 2021 Aug 16]. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8066704

102. Kogan NE, Clemente L, Liautaud P, Kaashoek J, Link NB, Nguyen AT, et al. An Early Warning Approach to Monitor COVID-19 Activity with Multiple Digital Traces in Near Real-Time. ArXiv [Internet] 2020 [cited 2020 Oct 27];Available from: http://www.ncbi.nlm.nih.gov/pubmed/32676518

103. Akbar A, Kousiouris G, Pervaiz H, Sancho J, Ta-Shma P, Carrez F, et al. Real-time probabilistic data fusion for large-scale IoT applications. IEEE Access 2018;6:10015–27.

104. Rahmat RF, Athmanathan, Syahputra MF, Lydia MS. Real time monitoring system for water pollution in Lake Toba. In: 2016 International Conference on Informatics and Computing, ICIC 2016. Institute of Electrical and Electronics Engineers Inc.; 2017. page 383–8.

105. Hershman SG, Bot BM, Shcherbina A, Doerr M, Moayedi Y, Pavlovic A, et al. Physical activity, sleep and cardiovascular health data for 50,000 individuals from the MyHeart Counts Study. Sci Data [Internet] 2019 [cited 2019 May 7];6(1):24. Available from: http://www.nature.com/articles/s41597-019-0016-7

106. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome C V., et al. Public Health Surveillance: A Tool for Targeting and Monitoring Interventions [Internet]. The International Bank for Reconstruction and Development / The World Bank; 2006 [cited 2018 Jun 8]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21250345

107. Lee LM, Teutsch SM, Thacker SB, St. Louis ME. Principles & Practice of Public Health Surveillance [Internet]. Oxford University Press; 2010 [cited 2019 Jan 18]. Available from: http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195372922.001.0001/acprof-9780195372922

108. Willetts M, Hollowell S, Aslett L, Holmes C, Doherty A. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. Sci Rep [Internet] 2018 [cited 2019 May 7];8(1):7961. Available from: http://www.nature.com/articles/s41598-018-26174-1

109. Cheng CKY, Ip DKM, Cowling BJ, Ho LM, Leung GM, Lau EHY. Digital dashboard design using multiple data streams for disease surveillance with influenza surveillance as an example. J Med Internet Res [Internet] 2011 [cited 2019 Apr 22];13(4):e85. Available from: http://www.jmir.org/2011/4/e85/

110. World Health Organization. WHO Coronavirus (COVID-19) Dashboard: Public Health and Social Measures (PHSM) [Internet]. 2021;Available from: https://covid19.who.int/measures

111. Sareen S, Sood SK, Gupta SK. IoT-based cloud framework to control Ebola virus outbreak. J Ambient Intell Humaniz Comput 2018;9(3):459–76.

112. Shafi U, Mumtaz R, Anwar H, Qamar AM, Khurshid H. Surface Water Pollution

Detection using Internet of Things [Internet]. In: 2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT). IEEE; 2018 [cited 2019 Sep 4]. page 92–6.Available from: https://ieeexplore.ieee.org/document/8551341/

113. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The Benefits Of Health Information Technology: A Review Of The Recent Literature Shows Predominantly Positive Results. Health Aff [Internet] 2011 [cited 2019 Feb 22];30(3):464–71. Available from: http://www.healthaffairs.org/doi/10.1377/hlthaff.2011.0178

114. StatisticsCanada. Canadian Cancer Registry (CCR) [Internet]. 2007 [cited 2019 Apr 25];Available from: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=29336

115. NIH. National Institutes of Health. List of Registries [Internet]. NIH Clin. Res. Trials You2018 [cited 2019 Sep 4];Available from: https://www.nih.gov/health-information/nih-clinical-research-trials-you/list-registries

116. Government of Canada. Immunization coverage and registries [Internet]. Vaccines Immun.2018 [cited 2019 Sep 4];Available from: https://www.canada.ca/en/public-health/services/immunization-coverage-registries.html#a3

117. HCITExpert. Healthcare IT Experts: List of Disease Registries in India [Internet]. 2017 [cited 2019 Sep 4];Available from: https://blog.hcitexpert.com/p/india-disease-registries.html

118. Treat NMD: Neuromuscular network. TREAT-NMD : DMD registry - China [Internet]. 2019 [cited 2019 Sep 4];Available from: http://www.treat-nmd.eu/registry/379/

119. Soman S, Srivastava P, Murthy BK. Unique Health Identifier for India: An algorithm and

feasibility analysis on patient data. In: 2015 17th International Conference on E-Health Networking, Application and Services, HealthCom 2015. Institute of Electrical and Electronics Engineers Inc.; 2015. page 250–5.

120. Lix L, Ayles J, Bartholomew S, Cooke C, Ellison J, Emond V, et al. The Canadian Chronic Disease Surveillance System: A model for collaborative surveillance. Int J Popul Data Sci [Internet] 2018;3(3). Available from: https://ijpds.org/article/view/433

121. Government of Canada. The Canadian Diabetes Strategy: History, Evolution, Moving Forward - Canada.ca [Internet]. 2015 [cited 2019 Apr 25];Available from: https://www.canada.ca/en/public-health/services/chronic-diseases/diabetes/canadian-diabetes-strategy-history-evolution-moving-forward.html

122. Public Health Agency of Canada. Public Health Infobase : Canadian Chronic Disease Surveillance System (CCDSS), CCDP, PHAC [Internet]. 2018 [cited 2018 Jun 8];Available from: https://infobase.phac-aspc.gc.ca/CCDSS-SCSMC/data-tool/?l=eng&DDLV=CDSAMI&DDLM=ASIR&sexB=B&yrB=2012&=10&VIEW=2

123. Government of Canada PHA of C. Public Health Agency of Canada [Internet]. 2021 [cited 2021 Oct 26];Available from: https://www.canada.ca/en/public-health.html

124. Choi B. Chronic Disease Surveillance in Canada [Internet]. 2013 [cited 2018 Jun 8]. Available from: http://www.ncceh.ca/sites/default/files/Surveillance_Workshop_Feb_2013-Choi.pdf

125. Government of Canada. The Canadian Chronic Disease Surveillance System – An Overview - Canada.ca [Internet]. 2018 [cited 2018 Oct 3];Available from: https://www.canada.ca/en/public-health/services/publications/canadian-chronic-disease-surveillance-system-factsheet.html

126. Government of Canada. Public Health Infobase: Canadian Chronic Disease Indicators (CCDI), PHAC [Internet]. 2018 [cited 2019 Apr 25];Available from: https://infobase.phac-aspc.gc.ca/ccdi-imcc/

127. Public Health Agency of Canada. Physical Activity, Sedentary Behaviour and Sleep (PASS) Indicator Framework for surveillance [Internet]. 2019 [cited 2021 Nov 23];Available from: https://www.canada.ca/en/services/health/monitoring-surveillance/physical-activity-sedentary-behaviour-sleep.html

128. Statistics Canada. Surveys and statistical programs - Canadian Community Health Survey - Nutrition (CCHS) [Internet]. Surv. Stat. Programs2007 [cited 2019 May 10];Available from: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=7498

129. StatisticsCanada. Canadian Community Health Survey: Rapid response on barriers to physical activity, 2017 [Internet]. 2017;Available from: https://www150.statcan.gc.ca/n1/daily-quotidien/180626/dq180626c-eng.htm

130. Statistics C. Surveys and statistical programs - Canadian Health Measures Survey (CHMS) [Internet]. 2012 [cited 2018 Jan 13];Available from: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5071

131. Public Health Agency of Canada. Health Behaviour in School-Aged Children - Canada.ca [Internet]. 2016 [cited 2019 Sep 6];Available from: https://www.canada.ca/en/public-health/services/health-promotion/childhood-adolescence/programs-initiatives/school-health/health-behaviour-school-aged-children.html

132. Institute CF and LR. 2014-2015 Physical Activity Monitor | www.cflri.ca [Internet]. 2014 [cited 2019 Oct 19];Available from: https://www.cflri.ca/category/2014-2015-physical-activity-monitor

133. Centers for Disease Control and Prevention (CDC). BRFSS Survey Data and Documentation [Internet]. 2017 [cited 2018 Oct 5];Available from: https://www.cdc.gov/brfss/annual_data/annual_2017.html

134. Shakeri Hossein Abad Z, Kline A, Sultana M, Noaeen M, Nurmambetova E, Lucini F, et al. Digital public health surveillance: a systematic scoping review. npj Digit Med [Internet] 2021;4(1):41. Available from: https://doi.org/10.1038/s41746-021-00407-6

135. Katal A, Wazid M, Goudar RH. Big data: Issues, challenges, tools and Good practices [Internet]. In: 2013 Sixth International Conference on Contemporary Computing (IC3). IEEE; 2013 [cited 2019 Jan 24]. page 404–9.Available from: http://ieeexplore.ieee.org/document/6612229/

136. PanimalarS A, ShreeS V, KathrineA V. The 17 V's Of Big Data. Int Res J Eng Technol [Internet] 2017 [cited 2019 Sep 4];Available from: www.irjet.net

137. Khan N, Alsaqer M, Shah H, Badsha G, Abbasi AA, Salehian S. The 10 Vs, Issues and Challenges of Big Data [Internet]. In: Proceedings of the 2018 International Conference on Big Data and Education - ICBDE '18. New York, New York, USA: ACM Press; 2018 [cited 2019 Sep 4]. page 52–6.Available from: http://dl.acm.org/citation.cfm?doid=3206157.3206166

138. Benke K, Benke G. Artificial Intelligence and Big Data in Public Health. Int J Environ Res Public Health [Internet] 2018 [cited 2019 Sep 4];15(12):2796. Available from: http://www.mdpi.com/1660-4601/15/12/2796

139. Tan W, Blake MB, Saleh I, Dustdar S. Social-Network-Sourced Big Data Analytics. IEEE Internet Comput [Internet] 2013 [cited 2019 Apr 25];17(5):62–9. Available from: http://ieeexplore.ieee.org/document/6596496/

140. Khan N, Yaqoob I, Hashem IAT, Inayat Z, Ali WKM, Alam M, et al. Big data: survey, technologies, opportunities, and challenges. ScientificWorldJournal [Internet] 2014 [cited 2019 Apr 25];2014:712826. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25136682

141. Wang L, Zhan J, Luo C, Zhu Y, Yang Q, He Y, et al. BigDataBench: A big data benchmark suite from internet services [Internet]. In: 2014 IEEE 20th International Symposium on High Performance Computer Architecture (HPCA). IEEE; 2014 [cited 2019 Apr 25]. page 488–99.Available from: http://ieeexplore.ieee.org/document/6835958/

142. Jagadish H V., Gehrke J, Labrinidis A, Papakonstantinou Y, Patel JM, Ramakrishnan R, et al. Big data and its technical challenges. Commun ACM [Internet] 2014 [cited 2019 Apr 25];57(7):86–94. Available from: http://dl.acm.org/citation.cfm?doid=2622628.2611567

143. Shafqat S, Kishwer S, Rasool RU, Qadir J, Amjad T, Ahmad HF. Big data analytics enhanced healthcare systems: a review. J Supercomput [Internet] 2018 [cited 2019 Jan 25];1–46. Available from: http://link.springer.com/10.1007/s11227-017-2222-4

144. Manyika J, Michael C, Brad B, Jacques B, Richard D, Charles R, et al. Big data: The next frontier for innovation, competition, and productivity. McKinsey Glob Inst [Internet] 2011 [cited 2019 Sep 6];(June):156. Available from: www.mckinsey.com/mgi.

145. Atzori L, Iera A, Morabito G. Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm. Ad Hoc Networks [Internet] 2017 [cited 2019 Sep 4];56:122–40. Available from: https://www.sciencedirect.com/science/article/pii/S1570870516303316?via%3Dihub

146. Giusto D, Iera A, Morabito G, Atzori L, editors. The Internet of Things [Internet]. New York, NY: Springer New York; 2010 [cited 2019 Jan 25]. Available from:

http://link.springer.com/10.1007/978-1-4419-1674-7

147. Ashton K. That “Internet of Things” Thing [Internet]. 2010 [cited 2019 Sep 6]. Available from: http://www.rfidjournal.com/article/print/4986

148. Madakam S, Ramaswamy R, Tripathi S. Internet of Things (IoT): A Literature Review. J Comput Commun [Internet] 2015 [cited 2019 Sep 6];03(05):164–73. Available from: http://www.scirp.org/journal/jcchttp://dx.doi.org/10.4236/jcc.2015.35021http://dx.doi.org/10.4236/jcc.2015.35021

149. Mark Page ; Maria Molina. The_Mobile_Economy_2013.pdf.crdownload [Internet]. 2013. Available from: http://www.gsmamobileeconomy.com/GSMA Mobile Economy 2013.pdf

150. Gubbi J, Buyya R, Marusic S, Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions. Futur Gener Comput Syst [Internet] 2013 [cited 2019 Jan 25];29(7):1645–60. Available from: https://www.sciencedirect.com/science/article/pii/S0167739X13000241

151. Fernández-Caramés TM, Fraga-Lamas P. Teaching and Learning IoT Cybersecurity and Vulnerability Assessment with Shodan through Practical Use Cases. Sensors 2020, Vol 20, Page 3048 [Internet] 2020 [cited 2021 Aug 13];20(11):3048. Available from: https://www.mdpi.com/1424-8220/20/11/3048/htm

152. Al-Taee MA, Al-Nuaimy W, Al-Ataby A, Muhsin ZJ, Abood SN. Mobile health platform for diabetes management based on the Internet-of-Things. In: 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, AEECT 2015. Institute of Electrical and Electronics Engineers Inc.; 2015.

153. Surantha N, Kusuma GP, Isa SM. Internet of things for sleep quality monitoring system: A survey. In: Proceedings - 11th 2016 International Conference on Knowledge,

Information and Creativity Support Systems, KICSS 2016. Institute of Electrical and Electronics Engineers Inc.; 2017.

154. Patil N, Iyer B. Health monitoring and tracking system for soldiers using Internet of Things(IoT). In: Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2017. Institute of Electrical and Electronics Engineers Inc.; 2017. page 1347–52.

155. Mathew A, Amreen FSA, Verma A. Smart Disease Surveillance Based on Internet of Things (IoT). Int J Adv Res Comput Commun Eng 2015;4(5).

156. Partha P. Ray. Internet of Things based Physical Activity Monitoring (PAMIoT): An Architectural Framework to Monitor Human Physical Activity [Internet]. In: IEEE CALCON. 2014 [cited 2019 Sep 4]. Available from: http://capsil.org/capsilwiki/index.php/MICAz

157. Panchatcharam P, Vivekanandan S. Internet of things (IoT) in healthcare - Smart health and surveillance, architectures, security analysis and data transfer: A review. Int J Softw Innov [Internet] 2019;7(2):21–40. Available from: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/IJSI.2019040103

158. Chung AE, Jensen RE, Basch EM. Leveraging emerging technologies and the “Internet of Things” to improve the quality of cancer care. J Oncol Pract [Internet] 2016;12(10):863–6. Available from: /pmc/articles/PMC5063188/

159. Tun SYY, Madanian S, Mirza F. Internet of things (IoT) applications for elderly care: a reflective review. Aging Clin Exp Res [Internet] 2021;33(4):855–67. Available from: https://link.springer.com/10.1007/s40520-020-01545-9

160. Dobkin BH. A Rehabilitation-Internet-of-Things in the Home to Augment Motor Skills

and Exercise Training. Neurorehabil Neural Repair 2017;31(3):217–27.

161. Wan J, Gu X, Chen L, Wang J. Internet of Things for Ambient Assisted Living: Challenges and Future Opportunities [Internet]. In: 2017 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). IEEE; 2017 [cited 2018 Jun 27]. page 354–7.Available from: http://ieeexplore.ieee.org/document/8250384/

162. Yoldemir T. Internet Of Things and women's health [Internet]. Climacteric2020;23(5):423–5. Available from: https://www.tandfonline.com/action/journalInformation?journalCode=icmt20

163. Isaac N, Sampath N, Gay V. SAM Smart Asthma Monitoring: Focus on Air Quality Data and Internet of Things (IoT). In: International Symposium on Medical Information and Communication Technology, ISMICT. IEEE Computer Society; 2018.

164. Qi J, Yang P, Newcombe L, Peng X, Yang Y, Zhao Z. An overview of data fusion techniques for Internet of Things enabled physical activity recognition and measure. Inf Fusion 2020;55:269–80.

165. H. Medjahed, D. Istrate, J. Boudy, J.L. Baldinger BD. A pervasive Multi-sensor Data Fusion for Smart Home Healthcare Monitoring [Internet]. In: IEEE Conference on Fuzzy Systems. Taipei, Taiwan: 2011. page 1466–73.Available from: https://www.researchgate.net/publication/221358732_A_pervasive_multi-sensor_data_fusion_for_smart_home_healthcare_monitoring

166. Hoog Antink C, Schulz F, Leonhardt S, Walter M. Motion Artifact Quantification and Sensor Fusion for Unobtrusive Health Monitoring. Sensors (Basel) [Internet] 2017 [cited 2018 Aug 20];18(1):38. Available from: http://www.mdpi.com/1424-8220/18/1/38

167. Moro Visconti R, Larocca A, Marconi M. Big Data-Driven Value Chains and Digital Platforms: From Value Co-Creation to Monetization. SSRN Electron J [Internet] 2017 [cited 2021 Aug 16];Available from: https://papers.ssrn.com/abstract=2903799

168. Hajjaji Y, Boulila W, Farah IR, Romdhani I, Hussain A. Big data and IoT-based applications in smart environments: A systematic review. Comput Sci Rev [Internet] 2021;39:100318. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1574013720304184

169. Leskinen T, Stenholm S, Aalto V, Head J, Kivimäki M, Vahtera J. Physical activity level as a predictor of healthy and chronic disease-free life expectancy between ages 50 and 75. Age Ageing [Internet] 2018 [cited 2019 Apr 30];47(3):423–9. Available from: https://academic.oup.com/ageing/article/47/3/423/4930839

170. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. Lancet [Internet] 2006 [cited 2019 Apr 30];367(9524):1747–57. Available from: https://www.sciencedirect.com/science/article/pii/S0140673606687709

171. von Ruesten A, Weikert C, Fietze I, Boeing H. Association of sleep duration with chronic diseases in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study. PLoS One [Internet] 2012 [cited 2019 Apr 30];7(1):e30972. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22295122

172. Stranges S, Tigbe W, Gómez-Olivé FX, Thorogood M, Kandala N-B. Sleep Problems: An Emerging Global Epidemic? Findings From the INDEPTH WHO-SAGE Study Among More Than 40,000 Older Adults From 8 Countries Across Africa and Asia. Sleep [Internet] 2012 [cited 2019 Apr 30];35(8):1173–81. Available from:

https://academic.oup.com/sleep/article-lookup/doi/10.5665/sleep.2012

173. Gibbs BB, Hergenroeder AL, Katzmarzyk PT, Lee I-M, Jakicic JM. Definition, measurement, and health risks associated with sedentary behavior. Med Sci Sports Exerc [Internet] 2015 [cited 2019 Apr 30];47(6):1295–300. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25222816

174. World Health Organization. Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age [Internet]. 2019 [cited 2019 Apr 30]. Available from: http://www.who.int/iris/handle/10665/311664

175. World Health Organization. To grow up healthy, children need to sit less and play more [Internet]. Saudi Med. J.2019 [cited 2019 Apr 30];40(5):519–20. Available from: https://www.who.int/news-room/detail/24-04-2019-to-grow-up-healthy-children-need-to-sit-less-and-play-more

176. Hamer M, Sharma N, Batty GD. Association of objectively measured physical activity with brain structure: UK Biobank study. J Intern Med [Internet] 2018 [cited 2019 Feb 21];284(4):439–43. Available from: https://journals-scholarsportal-info.proxy.lib.uwaterloo.ca/pdf/09546820/v284i0004/439_aoompawbsubs.xml

177. Dolley S. Big Data's Role in Precision Public Health. Front public Heal [Internet] 2018 [cited 2018 May 10];6:68. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29594091

178. LaMunion SR, Fitzhugh EC, Crouter SE. Challenges and opportunities related to the objective assessment of physical activity within U.S. health surveys. Ann Epidemiol [Internet] 2020 [cited 2020 Apr 23];43:1–10. Available from: https://reader.elsevier.com/reader/sd/pii/S1047279719304326?token=0CC1AC94DD5677

1C62ABFBC15D7D0EE90AEDBF328E5617B84FB4D855375322EEE35A993930BEAA224A27818E81136D0A

179. Bin YS, Marshall NS, Glozier N. Sleeping at the Limits: The Changing Prevalence of Short and Long Sleep Durations in 10 Countries. Am J Epidemiol [Internet] 2013 [cited 2019 May 7];177(8):826–33. Available from: https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kws308

180. Leech JA. Changes in sleep duration and recreational screen time among Canadians, 1998-2010. J Sleep Res [Internet] 2017 [cited 2019 May 7];26(2):202–9. Available from: http://doi.wiley.com/10.1111/jsr.12479

181. Public Health Agency of Canada. Improving health outcomes- A paradigm shift. Centre for chronic disease prevention Strategic plan 2016-2019 [Internet]. Ottawa (ON): 2015 [cited 2019 May 21]. Available from: http://www.phac-aspc.gc.ca/cd-mc/assets/pdf/ccdp-strategic-plan-2016-2019-plan-strategique-cpmc-eng.pdf

182. Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel. Nat 2006 4397075 [Internet] 2006 [cited 2021 Aug 11];439(7075):462–5. Available from: https://www.nature.com/articles/nature04292

183. Wang A;, Zhang A;, Chan EHW;, Shi W;, Zhou X;, Liu ZA, et al. A Review of Human Mobility Research Based on Big Data and Its Implication for Smart City Development. ISPRS Int J Geo-Information [Internet] 2020 [cited 2021 Aug 10];10(1):13. Available from: https://www.mdpi.com/2220-9964/10/1/13

184. Big Data and Mobility Analytics: What can we learn from the way things (and humans!) move? | Open Data Science Conference [Internet]. [cited 2021 Aug 11];Available from: https://odsc.com/speakers/big-data-and-mobility-analytics-what-can-we-learn-from-the-

way-things-and-humans-move/

185. Pirra M, Diana M. Integrating mobility data sources to define and quantify a vehicle-level congestion indicator: an application for the city of Turin. Eur Transp Res Rev 2019;11(1):1–11.

186. Meekan MG, Duarte CM, Fernández-Gracia J, Thums M, Sequeira AMM, Harcourt R, et al. The Ecology of Human Mobility. Trends Ecol Evol 2017;32(3):198–210.

187. Barbieri DM, Lou B, Passavanti M, Hui C, Hoff I, Lessa DA, et al. Impact of COVID-19 pandemic on mobility in ten countries and associated perceived risk for all transport modes. PLoS One [Internet] 2021;16(2):e0245886. Available from: https://dx.plos.org/10.1371/journal.pone.0245886

188. Trivedi A, Silverstein K, Strubell E, Iyyer M, Shenoy P. WiFiMod: Transformer-based Indoor Human Mobility Modeling using Passive Sensing [Internet]. In: ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS) (COMPASS '21),. Virtual Event, Australia. ACM; 2021 [cited 2021 Jul 30]. page 18.Available from: https://doi.org/10.1145/3460112.3471951

189. Williams NE, Thomas TA, Dunbar M, Eagle N, Dobra A. Measures of human mobility using mobile phone records enhanced with GIS data. PLoS One [Internet] 2015;10(7):e0133630. Available from: http://www.hsph.harvard.edu/ess/

190. Brookmeyer R, Stroup DF. Monitoring the Health of Populations: Statistical Principles and Methods for Public Health Surveillance [Internet]. 2005 [cited 2020 Jan 28]. Available from: http://www.oup-usa.org

191. Sonesson C, Bock D. A review and discussion of prospective statistical surveillance in public health. J. R. Stat. Soc. Ser. A Stat. Soc.2003;166(1):5–21.

192. Al-Azazi S, Singer A, Rabbani R, Lix LM. Combining population-based administrative health records and electronic medical records for disease surveillance. BMC Med Inform Decis Mak [Internet] 2019 [cited 2020 Jan 21];19(1):120. Available from: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-019-0845-5

193. Pentland A, Lazer D, Brewer D, Heibeck T. Using reality mining to improve public health and medicine. In: Studies in Health Technology and Informatics. IOS Press; 2009. page 93–102.

194. Wang Y, Holt JB, Xu F, Zhang X, Dooley DP, Lu H, et al. Using 3 health surveys to compare multilevel models for small area estimation for chronic diseases and health behaviors. Prev Chronic Dis 2018;15(11).

195. Zwetsloot IM, Woodall WH. A review of some sampling and aggregation strategies for basic statistical process monitoring. J Qual Technol [Internet] 2019 [cited 2020 Jan 28];1–16. Available from: https://www.tandfonline.com/doi/full/10.1080/00224065.2019.1611354

196. Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: a review. J R Stat Soc A [Internet] 2012 [cited 2020 Jan 21];175(1):49–82. Available from: http://doi.wiley.com/10.1111/j.1467-985X.2011.00714.x

197. Zhang X, Zhang T, Young AA, Li X. Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance Data. PLoS One [Internet] 2014 [cited 2020 Feb 3];9(2):88075. Available from: www.plosone.org

198. Tizzoni M, Bajardi P, Decuyper A, Kon Kam King G, Schneider CM, Blondel V, et al. On the Use of Human Mobility Proxies for Modeling Epidemics. PLoS Comput Biol

[Internet] 2014;10(7):e1003716. Available from: https://dx.plos.org/10.1371/journal.pcbi.1003716

199. van de Schoot R, Depaoli S, King R, Kramer B, Märtens K, Tadesse MG, et al. Bayesian statistics and modelling. Nat Rev Methods Prim 2021 11 [Internet] 2021 [cited 2021 Aug 16];1(1):1–26. Available from: https://www.nature.com/articles/s43586-020-00001-2

200. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C. Big Data for Infectious Disease Surveillance and Modeling. J Infect Dis [Internet] 2016 [cited 2020 Jul 16];214(suppl 4):S375–9. Available from: https://academic.oup.com/jid/article/214/suppl_4/S375/2527914

201. Collins LM. Analysis of Longitudinal Data: The Integration of Theoretical Model, Temporal Design, and Statistical Model. Annu Rev Psychol [Internet] 2006 [cited 2020 Feb 19];57(1):505–28. Available from: http://www.annualreviews.org/doi/10.1146/annurev.psych.57.102904.190146

202. Zhang M. Time Series: Autoregressive models AR, MA, ARMA, ARIMA Overview 1 Introduction of Time Series Categories and Terminologies White Noise and Random Walk Time Series Analysis. 2018.

203. Zhou L, Zhao P, Wu D, Cheng C, Huang H. Time series model for forecasting the number of new admission inpatients. BMC Med Inform Decis Mak [Internet] 2018 [cited 2020 Jan 28];18(1):39. Available from: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-018-0616-8

204. Siriyasatien P, Phumee A, Ongruk P, Jampachaisri K, Kesorn K. Analysis of significant factors for dengue fever incidence prediction. BMC Bioinformatics [Internet] 2016 [cited 2020 Jan 28];17(1):166. Available from:

http://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-016-1034-5

205. Zhou L, Xia J, Yu L, Wang Y, Shi Y, Cai S, et al. Using a Hybrid Model to Forecast the Prevalence of Schistosomiasis in Humans. Int J Environ Res Public Health [Internet] 2016 [cited 2020 Jan 28];13(4):355. Available from: http://www.mdpi.com/1660-4601/13/4/355

206. Wu W, Guo J, An S, Guan P, Ren Y, Xia L, et al. Comparison of two hybrid models for forecasting the incidence of hemorrhagic fever with renal syndrome in Jiangsu Province, China. PLoS One 2015;10(8).

207. Song X, Xiao J, Deng J, Kang Q, Zhang Y, Xu J. Time series analysis of influenza incidence in Chinese provinces from 2004 to 2011. Med (United States) 2016;95(26).

208. Fricker RD, Hegler BL, Dunfee DA. Comparing syndromic surveillance detection methods: EARS' versus a CUSUM-based methodology. Stat Med [Internet] 2008 [cited 2021 Aug 16];27(17):3407–29. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/sim.3197

209. Centers for Disease Control and Prevention (CDC). Initial Evaluation of the Early Aberration Reporting System --- Florida [Internet]. 2005 [cited 2020 Jan 28];Available from: https://www.cdc.gov/Mmwr/preview/mmwrhtml/su5401a20.htm

210. Thor J, Lundberg J, Ask J, Olsson J, Carli C, Härenstam KP, et al. Application of statistical process control in healthcare improvement: Systematic review. Qual. Saf. Heal. Care2007;16(5):387–99.

211. Zeger SL, Irizarry R, Peng RD. ON TIME SERIES ANALYSIS OF PUBLIC HEALTH AND BIOMEDICAL DATA. Annu Rev Public Health [Internet] 2006 [cited 2020 Feb 3];27(1):57–79. Available from: http://www.annualreviews.org/doi/10.1146/annurev.publhealth.26.021304.144517

212. López-Cuadrado T, De Mateo S, Jiménez-Jorge S, Savulescu C, Larrauri A. Influenza-related mortality in Spain, 1999-2005. Gac Sanit 2012;26(4):325–9.

213. Afolabi Ibikunle J. Life Expectancy, Public Health Spending and Economic Growth in Nigeria. Soc Sci [Internet] 2019 [cited 2020 Feb 19];8(6):369. Available from: https://www.researchgate.net/publication/319016619

214. Ramezanian M, Haghdoost AA, Mehrolhassani MH, Abolhallaje M, Dehnavieh R, Najafi B, et al. Forecasting health expenditures in Iran using the ARIMA model (2016-2020). Med J Islam Repub Iran 2019;33(1):1–4.

215. Time Series Analysis using ARIMA Models: An Approach to Forecasting Health Expenditures in USA - Istituto di Economia Internazionale [Internet]. [cited 2020 Jan 28];Available from: http://www.iei1946.it/en/rivista-articolo.php?id=165

216. Wang Y wen, Shen Z zhou, Jiang Y. Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. PLoS One 2018;13(9).

217. Cong, Ren, Xie, Wang. Predicting Seasonal Influenza Based on SARIMA Model, in Mainland China from 2005 to 2018. Int J Environ Res Public Health [Internet] 2019 [cited 2020 Jan 28];16(23):4760. Available from: https://www.mdpi.com/1660-4601/16/23/4760

218. Liu H, Li C, Shao Y, Zhang X, Zhai Z, Wang X, et al. Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011–2019 using the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ETS) models. J Infect Public Health 2020;13(2):287–94.

219. Yuan M, Boston-Fisher N, Luo Y, Verma A, Buckeridge DL. A systematic review of aberration detection algorithms used in public health surveillance. J Biomed Inform [Internet] 2019 [cited 2020 Jan 21];94:103181. Available from:

https://linkinghub.elsevier.com/retrieve/pii/S1532046419300991

220. C Rietbergen TDIKKJKM. Reporting of Bayesian analysis in epidemiologic research should become more transparent. J Clin Epidemiol 2017;86:51-58.e2.

221. Charras-Garrido M, Abrial D, De Go¨er J, Go¨er G, Peyrard N. Classification method for disease risk mapping based on discrete hidden Markov random fields. Biostatistics [Internet] 2012 [cited 2021 Aug 16];13(2):241–55. Available from: https://academic.oup.com/biostatistics/article/13/2/241/263673

222. Hagenbuchner M, Cliff DP, Trost SG, Van Tuc N, Peoples GE. Prediction of activity type in preschool children using machine learning techniques. J Sci Med Sport 2015;18(4):426–31.

223. Lopez Bernal J, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol [Internet] 2016 [cited 2020 Jan 28];dyw098. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyw098

224. Lopez Bernal J, Cummins S, Gasparrini A. The use of controls in interrupted time series studies of public health interventions. Int J Epidemiol [Internet] 2018;47(6):2082–93. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29982445

225. Bender A, Groll A, Scheipl F. A generalized additive model approach to time-to-event analysis. Stat Modelling [Internet] 2018;18(3–4):299–321. Available from: http://journals.sagepub.com/doi/10.1177/1471082X17748083

226. Gordon BA, Bruce L, Benson AC. Physical activity intensity can be accurately monitored by smartphone global positioning system ‘app.’ Eur J Sport Sci [Internet] 2016 [cited 2019 Feb 4];16(5):624–31. Available from:

http://www.tandfonline.com/doi/full/10.1080/17461391.2015.1105299

227. Graham DJ, Hipp JA. Emerging technologies to promote and evaluate physical activity: cutting-edge research and future directions. Front public Heal [Internet] 2014 [cited 2019 Feb 4];2:66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25019066

228. Obinikpo AA, Kantarci B. Big data aggregation in the case of heterogeneity: a feasibility study for digital health. Int J Mach Learn Cybern [Internet] 2019 [cited 2019 Jan 25];10(10):2643–55. Available from: http://link.springer.com/10.1007/s13042-018-00904-3

229. Glymour MM, Bibbins-Domingo K. The Future of Observational Epidemiology: Improving Data and Design to Align With Population Health. Am J Epidemiol [Internet] 2019 [cited 2019 May 14];188(5):836–9. Available from: https://academic.oup.com/aje/article/188/5/836/5379098

230. Ahuja R, Chug A, Gupta S, Ahuja P, Kohli S. Classification and clustering algorithms of machine learning with their applications. Stud Comput Intell [Internet] 2020 [cited 2022 May 13];855:225–48. Available from: https://link.springer.com/chapter/10.1007/978-3-030-28553-1_11

231. Reynolds D. Gaussian Mixture Models [Internet]. In: Encyclopedia of Biometrics. Boston, MA: Springer US; 2015. page 827–32.Available from: http://link.springer.com/10.1007/978-1-4899-7488-4_196

232. Grandner MA, Jackson NJ, Izci-Balserak B, Gallagher RA, Murray-Bachmann R, Williams NJ, et al. Social and Behavioral Determinants of Perceived Insufficient Sleep. Front Neurol [Internet] 2015 [cited 2019 Mar 11];6:112. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26097464

233. Grandner MA, Patel NP, Gehrman PR, Perlis ML, Pack AI. Problems associated with short sleep: bridging the gap between laboratory and epidemiological studies. Sleep Med Rev [Internet] 2010 [cited 2019 Mar 11];14(4):239–47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19896872

234. Institute of Medicine. Sleep Disorders and Sleep Deprivation [Internet]. Washington, D.C.: National Academies Press; 2006 [cited 2019 Mar 11]. Available from: http://www.nap.edu/catalog/11617

235. Medicine I of. Resident Duty Hours [Internet]. Washington, D.C.: National Academies Press; 2009 [cited 2019 Mar 11]. Available from: http://www.nap.edu/catalog/12508

236. Chastin SFM, Craemer M De, Cocker K De, Powell L, Cauwenberg J Van, Dall P, et al. How does light-intensity physical activity associate with adult cardiometabolic health and mortality? Systematic review with meta-analysis of experimental and observational studies. Br J Sport Med [Internet] 2019 [cited 2019 Mar 11];53(6):370–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29695511

237. Bellettiere J, Healy GN, LaMonte MJ, Kerr J, Evenson KR, Rillamas-Sun E, et al. Sedentary Behavior and Prevalent Diabetes in 6,166 Older Women: The Objective Physical Activity and Cardiovascular Health Study. Journals Gerontol Ser A [Internet] 2019 [cited 2019 Mar 11];74(3):387–95. Available from: https://academic.oup.com/biomedgerontology/article/74/3/387/4991878

238. Head BW. Toward More "Evidence-Informed" Policy Making? Public Adm Rev [Internet] 2016 [cited 2019 Sep 4];76(3):472–84. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/puar.12475%4010.1111/%28ISSN%291540-6210.lessons-in-public-administration

239. McCaughey D, Bruning NS. Rationality versus reality: the challenges of evidence-based decision making for health policy makers. Implement Sci [Internet] 2010 [cited 2019 Sep 4];5(1):39. Available from: http://implementationscience.biomedcentral.com/articles/10.1186/1748-5908-5-39

240. Rutter H, Savona N, Glonti K, Bibby J, Cummins S, Finegood DT, et al. The need for a complex systems model of evidence for public health. Lancet (London, England) [Internet] 2017 [cited 2019 Sep 4];390(10112):2602–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28622953

241. Cairney P, Oliver K. Evidence-based policymaking is not like evidence-based medicine, so how far should you go to bridge the divide between evidence and policy? Heal Res Policy Syst [Internet] 2017;15(1):35. Available from: http://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-017-0192-x

242. Oliver K, Innvar S, Lorenc T, Woodman J, Thomas J. A systematic review of barriers to and facilitators of the use of evidence by policymakers. BMC Health Serv Res [Internet] 2014 [cited 2019 Sep 6];14(1):2. Available from: https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-14-2

243. Oliver K, Lorenc T, Innvær S. New directions in evidence-based policy research: A critical analysis of the literature [Internet]. Heal. Res. Policy Syst.2014 [cited 2019 Sep 6];12(1). Available from: http://www.health-policy-systems.com/content/12/1/34

244. MaRS Health. Healthy Behaviour Data Challenge [Internet]. 2017 [cited 2019 Jan 24];Available from: http://www.healthdatachallenge.com/pastchallenge/

245. Paré G, Leaver C, Bourget C. Diffusion of the Digital Health Self-Tracking Movement in Canada: Results of a National Survey. J Med Internet Res [Internet] 2018 [cited 2019 Sep

4];20(5):e177. Available from: http://www.jmir.org/2018/5/e177/

246. Statista. Wearables - Canada | Statista Market Forecast [Internet]. 2019 [cited 2019 Sep 4];Available from: https://www.statista.com/outlook/319/108/wearables/canada#market-users

247. Costello K. Gartner Says Worldwide Wearable Device Sales to Grow 26 Percent in 2019 [Internet]. Gartner, Inc.2018 [cited 2019 Sep 4];Available from: https://www.gartner.com/en/newsroom/press-releases/2018-11-29-gartner-says-worldwide-wearable-device-sales-to-grow-

248. Fitbit. Fitbit Official Site for Activity Trackers & More [Internet]. 2019 [cited 2019 Jan 24];Available from: https://www.fitbit.com/en-ca/home

249. Garmin. Garmin | Canada [Internet]. 2019 [cited 2019 Sep 4];Available from: https://www.garmin.com/en-CA/

250. Mi. Mi Global Home [Internet]. Xioami2018 [cited 2019 Sep 4];Available from: https://www.mi.com/global/miband/#01

251. Russey Cathy. Xiaomi Tops Apple and Fitbit in Q3 2018 | Wearable Technologies [Internet]. 2018 [cited 2019 Sep 4];Available from: https://www.wearable-technologies.com/2018/12/idc-report-xiaomi-tops-apple-and-fitbit-with-21-5-global-wearable-market-share-in-q3-2018/

252. Fukuoka Y, Haskell W, Lin F, Vittinghoff E. Short- and Long-term Effects of a Mobile Phone App in Conjunction With Brief In-Person Counseling on Physical Activity Among Physically Inactive Women. JAMA Netw Open [Internet] 2019 [cited 2019 Sep 4];2(5):e194281. Available from: http://jamanetworkopen.jamanetwork.com/article.aspx?doi=10.1001/jamanetworkopen.20

19.4281

253. Sanders JP, Loveday A, Pearson N, Edwardson C, Yates T, Biddle SJH, et al. Devices for Self-Monitoring Sedentary Time or Physical Activity: A Scoping Review. J Med Internet Res [Internet] 2016 [cited 2019 Sep 4];18(5):e90. Available from: http://www.jmir.org/2016/5/e90/

254. Sahami Shirazi A, Clawson J, Hassanpour Y, Tourian MJ, Schmidt A, Chi EH, et al. Already up? using mobile phones to track & share sleep behavior. Int J Hum Comput Stud [Internet] 2013 [cited 2019 Sep 4];71(9):878–88. Available from: https://www.sciencedirect.com/science/article/pii/S1071581913000244

255. Mittelstadt B. Ethics of the health-related internet of things: a narrative review. Ethics Inf Technol [Internet] 2017 [cited 2019 Sep 6];19(3):157–75. Available from: http://link.springer.com/10.1007/s10676-017-9426-4

256. Costa FF. Big data in biomedicine. Drug Discov Today [Internet] 2014 [cited 2019 Sep 6];19(4):433–40. Available from: https://www.sciencedirect.com/science/article/pii/S1359644613003644?via%3Dihub

257. ecobee. ecobee | Smart Home Technology | [Internet]. 2019 [cited 2018 Oct 26];Available from: https://www.ecobee.com/

258. ecobee. Donate your Data | Smart WiFi Thermostats by ecobee [Internet]. 2019 [cited 2017 Sep 21];Available from: https://www.ecobee.com/donateyourdata/

259. ecobee on the App Store [Internet]. [cited 2021 Oct 26];Available from: https://apps.apple.com/us/app/ecobee/id916985674

260. ecobee – Apps on Google Play [Internet]. [cited 2021 Oct 26];Available from: https://play.google.com/store/apps/details?id=com.ecobee.athenamobile&hl=en_CA&gl=

US

261. Talati S. How to make the most of your CPD. Dent Nurs [Internet] 2015 [cited 2019 Sep 4];11(8):480–1. Available from: https://thesmartcave.com/ecobee3-remote-sensors/

262. Chen Z, Wang YS. Sleep monitoring using an infrared thermal array sensor [Internet]. In: Wang K-W, Sohn H, Huang H, Lynch JP, editors. Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2019. SPIE; 2019 [cited 2019 Sep 4]. page 53.Available from: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10970/2515515/Sleep-monitoring-using-an-infrared-thermal-array-sensor/10.1117/12.2515515.full

263. Marsdd.com. MaRS Discovery District [Internet]. 2013 [cited 2019 May 10];Available from: http://www.marsdd.com/articles/sourcing-candidates/

264. PHAC, Desanto KR. Public Health Agency of Canada - Canada.ca [Internet]. J. Consum. Health Internet2011 [cited 2019 Apr 21];15(2):196–206. Available from: https://www.canada.ca/en/public-health.html

265. Canadian Institutes of Health Research G of C. Welcome to the Canadian Institutes of Health Research - CIHR [Internet]. 2003 [cited 2019 May 10];Available from: http://www.cihr-irsc.gc.ca/e/193.html

266. Canada. S. Canadian Health Measures Survey (CHMS) [Internet]. 2019 [cited 2019 Mar 11];Available from: https://www.statcan.gc.ca/eng/survey/household/5071

267. Canadian Fitness and Lifestyle Research Institute (CFLRI). 2006 Physical Activity Monitor [Internet]. 2006 [cited 2019 May 10];Available from: https://www.cflri.ca/category/2016-2018-physical-activity-monitor

268. Clarke J, Colley R, Janssen I, Tremblay MS. Accelerometer-measured moderate-to-

vigorous physical activity of Canadian adults, 2007 to 2017. Heal reports [Internet] 2019 [cited 2019 Sep 4];30(8):3–10. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31454407

269. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fitbit Zip™ physical activity monitor as a measure of free-living physical activity. BMC Res Notes [Internet] 2014 [cited 2019 Sep 11];7(1):952. Available from: http://bmcresnotes.biomedcentral.com/articles/10.1186/1756-0500-7-952

270. Lee Rodgers J, Nicewander WA. Thirteen Ways to Look at the Correlation Coefficient. Am Stat [Internet] 1988 [cited 2018 Jan 15];42(1):59–66. Available from: http://www.tandfonline.com/doi/abs/10.1080/00031305.1988.10475524

271. Government of Canada. Temperature change in Canada - Canada.ca [Internet]. 2019 [cited 2019 Oct 21];Available from: https://www.canada.ca/en/environment-climate-change/services/environmental-indicators/temperature-change.html

272. Canada S. Use of thermostats, Canada, provinces and census metropolitan areas (CMA) Table 38-10-0049-01 [Internet]. 2021 [cited 2018 Jan 15];Available from: http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=1530060

273. Nest. Nest Learning Thermostat | Programs Itself, Helps Save Energy | Nest [Internet]. 2019 [cited 2019 Jan 24];Available from: https://nest.com/ca/thermostats/nest-learning-thermostat/overview/

274. Chaput J-PP, Janssen I. Sleep duration estimates of Canadian children and adolescents. J Sleep Res [Internet] 2016 [cited 2019 Sep 11];25(5):541–8. Available from: http://doi.wiley.com/10.1111/jsr.12410

275. Karen C. Roberts, Xiaoquan Yao VC, Jean-Philippe Chaput, Ian Janssen MST. Meeting

the Canadian 24-Hour Movement Guidelines for Children and Youth [Internet]. 2017. Available from: https://www150.statcan.gc.ca/n1/en/pub/82-003-x/2017010/article/54875-eng.pdf?st=YSjyK17b

276. Hauke J, Kossowski T. Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data. QUAGEO [Internet] 2011;30(2):87–93. Available from: https://www.sciendo.com/article/10.2478/v10117-011-0021-1

277. Public Health Agency of Canada. Physical Activity, Sedentary Behaviour and Sleep (PASS) Indicators - Public Health Infobase | Public Health Agency of Canada [Internet]. 2017 [cited 2019 Oct 21];Available from: https://health-infobase.canada.ca/pass/

278. Vegesna A, Tran M, Angelaccio M, Arcona S. Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. Telemed e-Health [Internet] 2017 [cited 2017 Jun 21];23(1):3–17. Available from: http://online.liebertpub.com/doi/pdf/10.1089/tmj.2016.0051

279. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T [Internet] 2014;39(5):356–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24883008

280. Brownson RC, Kreuter MW, Arrington BA, True WR. Translating scientific discoveries into public health action: how can schools of public health move us forward? Public Health Rep [Internet] 121(1):97–103. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16416704

281. All-new Echo (4th Gen) | With premium sound, smart home hub, and Alexa | Charcoal : Amazon.ca: Amazon Devices & Accessories [Internet]. [cited 2022 May 19];Available from: https://www.amazon.ca/dp/B07XKF5RM3?tag=googcana-

20&hvadid=467840376851&hvpos=&hvexid=&hvnetw=g&hvrand=9566730313341326564&hvpone=&hvptwo=&hvqmt=e&hvdev=c&ref=pd_sl_4bka701ejv_e&hvtargid=kwd-50712656950

282. Siri - Apple (CA) [Internet]. [cited 2022 May 19];Available from: https://www.apple.com/ca/siri/

283. Google Nest Mini - A Smart Speaker for Any Room - Google Store [Internet]. [cited 2022 May 19];Available from: https://store.google.com/ca/product/google_nest_mini?hl=en-GB

284. Fund GO. Smart Thermostat Rebate [Internet]. 2017 [cited 2018 Jan 13];Available from: https://www.greenon.ca/en/programs/smart-thermostat-rebate?postalcode=N2J2J4&type=Home

285. Kirk MA, Amiri M, Pirbaglou M, Ritvo P. Wearable Technology and Physical Activity Behavior Change in Adults With Chronic Cardiometabolic Disease: A Systematic Review and Meta-Analysis. Am J Heal Promot [Internet] 2019 [cited 2019 Sep 4];33(5):778–91. Available from: http://journals.sagepub.com/doi/10.1177/0890117118816278

286. Wazny K. Applications of crowdsourcing in health: An overview. J Glob Health [Internet] 2018 [cited 2020 Dec 9];8(1). Available from: /pmc/articles/PMC5840433/?report=abstract

287. Kraemer JD, Strasser AA, Lindblom EN, Niaura RS, Mays D. Crowdsourced data collection for public health: A comparison with nationally representative, population tobacco use data. Prev Med (Baltim) 2017;102:93–9.

288. Edo-Osagie O, De La Iglesia B, Lake I, Edeghere O. A scoping review of the use of Twitter for public health research. Comput. Biol. Med.2020;122:103770.

289. Plageras AP, Psannis KE, Ishibashi Y, Kim BG. IoT-based surveillance system for ubiquitous healthcare. In: IECON Proceedings (Industrial Electronics Conference). IEEE Computer Society; 2016. page 6226–30.

290. Qi J, Yang P, Waraich A, Deng Z, Zhao Y, Yang Y. Examining sensor-based physical activity recognition and monitoring for healthcare using Internet of Things: A systematic review. J. Biomed. Inform.2018;87:138–53.

291. Talal M, Zaidan AA, Zaidan BB, Albahri AS, Alamoodi AH, Albahri OS, et al. Smart Home-based IoT for Real-time and Secure Remote Health Monitoring of Triage and Priority System using Body Sensors: Multi-driven Systematic Review [Internet]. J. Med. Syst.2019 [cited 2020 Sep 17];43(3). Available from: https://pubmed.ncbi.nlm.nih.gov/30648217/

292. Asghari P, Rahmani AM, Javadi HHS. Internet of Things applications: A systematic review. Comput Networks 2019;148:241–61.

293. Bhuiyan MN, Rahman MM, Billah MM, Saha D. Internet of Things (IoT): A Review of Its Enabling Technologies in Healthcare Applications, Standards Protocols, Security, and Market Opportunities. IEEE Internet Things J [Internet] 2021;8(13):10474–98. Available from: https://ieeexplore.ieee.org/document/9365708/

294. Riazul Islam SM, Daehan Kwak, Humaun Kabir M, Hossain M, Kyung-Sup Kwak. The Internet of Things for Health Care: A Comprehensive Survey. IEEE Access [Internet] 2015;3:678–708. Available from: http://ieeexplore.ieee.org/document/7113786/

295. DeSalvo KB, O'Carroll PW, Koo D, Auerbach JM, Monroe JA. Public health 3.0: Time for an upgrade [Internet]. Am. J. Public Health2016 [cited 2020 Aug 11];106(4):621–2. Available from: http://ajph.aphapublications.org/doi/10.2105/AJPH.2016.303063

296. Coughlin SS. Recall bias in epidemiologic studies. J Clin Epidemiol [Internet] 1990 [cited 2018 Jan 15];43(1):87–91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2319285

297. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems [Internet]. Int. J. Environ. Res. Public Health2014 [cited 2020 Nov 6];11(5):5170–207. Available from: /pmc/articles/PMC4053886/?report=abstract

298. Buehler JW, Centers for Disease Control and Prevention. CDC's vision for public health surveillance in the 21st century. Introduction. MMWR Suppl [Internet] 2012 [cited 2020 Sep 17];61(3):1–2. Available from: https://pubmed.ncbi.nlm.nih.gov/22832989/

299. Groseclose SL, Buckeridge DL. Public Health Surveillance Systems: Recent Advances in Their Use and Evaluation [Internet]. Annu. Rev. Public Health2017 [cited 2020 Aug 13];38:57–79. Available from: https://pubmed.ncbi.nlm.nih.gov/27992726/

300. Althouse BM, Scarpino S V, Meyers LA, Ayers JW, Bargsten M, Baumbach J, et al. Enhancing disease surveillance with novel data streams: challenges and opportunities. EPJ Data Sci 2015 41 [Internet] 2015 [cited 2021 Jul 26];4(1):1–8. Available from: https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-015-0054-0

301. Navin K, Krishnan MBM, Lavanya S, Shanthini A. A mobile health based smart hybrid epidemic surveillance system to support epidemic control programme in public health informatics. In: IEEE International Conference on IoT and its Applications, ICIOT 2017. Institute of Electrical and Electronics Engineers Inc.; 2017.

302. Mancuso M, Yao X, Otchere D, Rasali D, Clark E, Svenson LW, et al. Proof of concept paper: Non-Traditional data sources for public health surveillance [Internet]. In: DH 2016 - Proceedings of the 2016 Digital Health Conference. New York, New York, USA:

Association for Computing Machinery, Inc; 2016 [cited 2020 May 1]. page 91–2.Available from: http://dl.acm.org/citation.cfm?doid=2896338.2896369

303. LaMunion SR, Fitzhugh EC, Crouter SE. Challenges and opportunities related to the objective assessment of physical activity within U.S. health surveys. Ann. Epidemiol.2020;43:1–10.

304. Lamiae E, Fatiha E, Mohammed B, Hicham GT. Contribution to a Smart Home Design for Medical Surveillance [Internet]. Springer, Cham; 2020 [cited 2020 Dec 15]. page 1080–93.Available from: https://doi.org/10.1007/978-3-030-37629-1_78

305. Huang J, Loschen W. Potential Applications of Emerging Technologies in Disease Surveillance. Online J Public Health Inform [Internet] 2019 [cited 2020 Aug 13];11(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6606178/

306. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband fitbit models in assessing sleep: Systematic review and meta-analysis [Internet]. J. Med. Internet Res.2019 [cited 2020 Mar 2];21(11):e16273. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31778122

307. DeSalvo KB, Wang YC, Harris A, Auerbach J, Koo D, O'Carroll P. Public Health 3.0: A Call to Action for Public Health to Meet the Challenges of the 21st Century. Prev Chronic Dis [Internet] 2017 [cited 2020 Sep 17];14(9):170017. Available from: http://www.cdc.gov/pcd/issues/2017/17_0017.htm

308. Zeng D, Cao Z, Neill DB. Artificial intelligence–enabled public health surveillance—from local detection to global epidemic monitoring and control [Internet]. In: Artificial Intelligence in Medicine. Elsevier; 2021 [cited 2020 Oct 22]. page 437–53.Available from: /pmc/articles/PMC7484813/?report=abstract

309. Chiolero A, Chiolero A, Chiolero A, Chiolero A, Buckeridge D. Glossary for public health surveillance in the age of data science [Internet]. J. Epidemiol. Community Health2020 [cited 2020 Aug 11];74(7):612–6. Available from: http://jech.bmj.com/

310. Atzori L, Iera A, Morabito G. The Internet of Things: A survey. Comput Networks 2010;54(15):2787–805.

311. Subasi A, Radhwan M, Kurdi R, Khateeb K. IoT based mobile healthcare system for human activity recognition. In: 2018 15th Learning and Technology Conference, L and T 2018. Institute of Electrical and Electronics Engineers Inc.; 2018. page 29–34.

312. Marques G, Roque Ferreira C, Pitarma R. A System Based on the Internet of Things for Real-Time Particle Monitoring in Buildings. Int J Environ Res Public Health [Internet] 2018 [cited 2021 Feb 17];15(4):821. Available from: http://www.mdpi.com/1660-4601/15/4/821

313. Li SH, Lin BS, Wang CA, Yang CT, Lin BS. Design of wearable and wireless multi-parameter monitoring system for evaluating cardiopulmonary function. Med Eng Phys 2017;47:144–50.

314. Adiputra RR, Hadiyoso S, Hariyani YS. Internet of Things: Low Cost and Wearable SpO2 Device for Health Monitoring. Int J Electr Comput Eng 2018;8(2):939–45.

315. Chiuchisan I, Geman O. An Approach of a Decision Support and Home Monitoring System for Patients with Neurological Disorders using Internet of Things Concepts. WSEAS Trans Syst 2014;13(Art. 46):460–9.

316. Nam KH, Kim DH, Choi BK, Han IH. Internet of Things, Digital Biomarker, and Artificial Intelligence in Spine: Current and Future Perspectives. Neurospine [Internet] 2019 [cited 2021 Feb 17];16(4):705–11. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/31905461

317. de Oliveira Júnior WG, de Oliveira JM, Munoz R, de Albuquerque VHC. A proposal for Internet of Smart Home Things based on BCI system to aid patients with amyotrophic lateral sclerosis. Neural Comput Appl 2020;32(15):11007–17.

318. Iwendi C, Khan S, Anajemba JH, Bashir AK, Noor F. Realizing an Efficient IoMT-Assisted Patient Diet Recommendation System Through Machine Learning Model. IEEE Access [Internet] 2020;8:28462–74. Available from: https://ieeexplore.ieee.org/document/8964364/

319. Mantey EA, Zhou C, Anajemba JH, Okpalaoguchi IM, Chiadika OD-M. Blockchain-secured Recommender System for Special Need Patients using Deep Learning. Front Public Heal 1AD;0:1256.

320. Salagare S, Prasad R. An Overview of Internet of Dental Things: New Frontier in Advanced Dentistry [Internet]. Wirel. Pers. Commun.2020 [cited 2021 Feb 17];110(3):1345–71. Available from: https://doi.org/10.1007/s11277-019-06790-4

321. Ting DSW, Lin H, Ruamviboonsuk P, Wong TY, Sim DA. Artificial intelligence, the internet of things, and virtual clinics: ophthalmology at the digital translation forefront [Internet]. Lancet Digit. Heal.2020 [cited 2021 Feb 17];2(1):e8–9. Available from: www.thelancet.com/digital-health

322. Yoldemir T. Internet Of Things and women's health [Internet]. Climacteric2020 [cited 2021 Feb 17];23(5):423–5. Available from: https://www.tandfonline.com/action/journalInformation?journalCode=icmt20

323. Saarikko J, Niela-Vilen H, Ekholm E, Hamari L, Azimi I, Liljeberg P, et al. Continuous 7-Month Internet of Things-Based Monitoring of Health Parameters of Pregnant and

Postpartum Women: Prospective Observational Feasibility Study. JMIR Form Res [Internet] 2020 [cited 2020 Sep 6];4(7):e12417. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32706696

324. Mieronkoski R, Azimi I, Rahmani AM, Aantaa R, Terävä V, Liljeberg P, et al. The Internet of Things for basic nursing care—A scoping review. Int. J. Nurs. Stud.2017;69:78–90.

325. Finch M, Griffin K, Pacala JT. Reduced Healthcare Use and Apparent Savings with Passive Home Monitoring Technology: A Pilot Study. J Am Geriatr Soc [Internet] 2017 [cited 2018 Feb 9];65(6):1301–5. Available from: http://doi.wiley.com/10.1111/jgs.14892

326. Saarikko J, Niela-Vilen H, Ekholm E, Hamari L, Azimi I, Liljeberg P, et al. Continuous 7-month internet of things⇓based monitoring of health parameters of pregnant and postpartum women: prospective observational feasibility study [Internet]. JMIR Form. Res.2020 [cited 2021 Feb 17];4(7):e12417. Available from: https://formative.jmir.org/2020/7/e12417

327. Alelyani S, Ibrahim A. Internet-of-Things in telemedicine for diabetes management. In: 2018 15th Learning and Technology Conference, L and T 2018. Institute of Electrical and Electronics Engineers Inc.; 2018. page 20–3.

328. Bruen D, Delaney C, Florea L, Diamond D. Glucose sensing for diabetes monitoring: Recent developments. Sensors (Switzerland)2017;17(8).

329. Ben Hassen H, Dghais W, Hamdi B. An E-health system for monitoring elderly health based on Internet of Things and Fog computing. Heal Inf Sci Syst [Internet] 2019 [cited 2021 Feb 17];7(1):1–9. Available from: https://doi.org/10.1007/s13755-019-0087-z

330. Hosseinzadeh M, Koohpayehzadeh J, Ghafour MY, Ahmed AM, Asghari P, Souri A, et al.

An elderly health monitoring system based on biological and behavioral indicators in internet of things. J Ambient Intell Humaniz Comput [Internet] 2020 [cited 2021 Feb 17];1:3. Available from: https://doi.org/10.1007/s12652-020-02579-7

331. Yacchirema D, De Puga JS, Palau C, Esteve M. Fall detection system for elderly people using IoT and Big Data. In: Procedia Computer Science. Elsevier B.V.; 2018. page 603–10.

332. Dohr A, Modre-Opsrian R, Drobics M, Hayn D, Schreier G. The Internet of Things for Ambient Assisted Living. In: 2010 Seventh International Conference on Information Technology: New Generations. IEEE; 2010.

333. Zhan K. Sports and health big data system based on 5G network and Internet of Things system. Microprocess Microsyst 2021;80:103363.

334. Zhao Y, You Y. Design and data analysis of wearable sports posture measurement system based on Internet of Things. Alexandria Eng J 2021;60(1):691–701.

335. Liu L. Construction of youth public sports service system based on embedded system and wireless IoT. Microprocess Microsyst 2021;83:103984.

336. Jiang Y. Combination of wearable sensors and internet of things and its application in sports rehabilitation. Comput Commun 2020;150:167–76.

337. Yang L, Yao T, Liu G, Sun L, Yang N, Zhang H, et al. Monitoring and Control of Medical Air Disinfection Parameters of Nosocomial Infection System Based on Internet of Things. J Med Syst [Internet] 2019 [cited 2021 Feb 17];43(5):1–7. Available from: https://doi.org/10.1007/s10916-019-1205-9

338. USGCRP. Impacts, Risks, and Adaptation in the United States: The Fourth National Climate Assessment, Volume II [Internet]. U.S. Global Change Research Program,

Washington, DC: 2018. Available from: https://nca2018.globalchange.gov/

339. Múnera D, Diana V, Aguirre J, Gómez NG. IoT-based air quality monitoring systems for smart cities: A systematic mapping study. Int J Electr Comput Eng [Internet] 2021 [cited 2021 Aug 5];11(4):3470–82. Available from: https://investigaciones-pure.udem.edu.co/en/publications/iot-based-air-quality-monitoring-systems-for-smart-cities-a-syste

340. Bal M, Abrishambaf R. A system for monitoring hand hygiene compliance based-on Internet-of-Things. In: Proceedings of the IEEE International Conference on Industrial Technology. Institute of Electrical and Electronics Engineers Inc.; 2017. page 1348–53.

341. Balaji B, Bhattacharya A, Fierro G, Gao J, Gluck J, Hong D, et al. Brick : Metadata schema for portable smart building applications. Appl Energy [Internet] 2018;226:1273–92. Available from: http://www.sciencedirect.com/science/article/pii/S0306261918302162

342. Al Knawy B, Adil M, Crooks G, Rhee K, Bates D, Jokhdar H, et al. The Riyadh Declaration: the role of digital health in fighting pandemics [Internet]. Lancet2020 [cited 2020 Oct 27];0(0). Available from: https://doi.org/10.1016/S0140-6736

343. Aiello AE, Renson A, Zivich PN. Social Media-and Internet-Based Disease Surveillance for Public Health. Annu Rev Public Heal [Internet] 2020 [cited 2020 Apr 24];41:101–18. Available from: https://doi.org/10.1146/annurev-publhealth-

344. Mooney SJ, Pejaver V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. Annu Rev Public Heal Annu Rev Public Heal [Internet] 2018 [cited 2018 May 16];39:95–112. Available from: https://doi.org/10.1146/annurev-publhealth-040617-014208

345. Hung K, Zhang Y-T. Implementation of a WAP-based telemedicine system for patient

monitoring. IEEE Trans Inf Technol Biomed [Internet] 2003 [cited 2018 Feb 9];7(2):101–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12834165

346. Statista. Smartphone users 2020 [Internet]. 2020 [cited 2021 Apr 16];Available from: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

347. Statista. Global connected wearable devices 2016-2022 [Internet]. 2021 [cited 2021 Apr 16];Available from: https://www.statista.com/statistics/487291/global-connected-wearable-devices/

348. Statista. Global smart home market size 2016-2022 [Internet]. 2021 [cited 2021 Apr 16];Available from: https://www.statista.com/statistics/682204/global-smart-home-market-size/

349. Grace SL, Taherzadeh G, Jae Chang IS, Boger J, Arcelus A, Mak S, et al. Perceptions of seniors with heart failure regarding autonomous zero-effort monitoring of physiological parameters in the smart-home environment. Hear Lung J Acute Crit Care [Internet] 2017 [cited 2018 Oct 31];46(4):313–9. Available from: https://www.sciencedirect.com/science/article/pii/S0147956316302266

350. Mihailidis A, Boger J, Hoey J, Jiancaro T. Zero Effort Technologies: Considerations, Challenges, and Use in Health, Wellness, and Rehabilitation. Synth Lect Assist Rehabil Heal Technol [Internet] 2011 [cited 2018 Aug 28];1(2):1–94. Available from: http://www.morganclaypool.com/doi/abs/10.2200/S00380ED1V01Y201108ARH002

351. Groseclose SL, German RR, Nsubuga P. Evaluating Public Health Surveillance. In: Principles & Practice of Public Health Surveillance. Oxford University Press; 2010.

352. Xu C, Yang Y, Gao W. Skin-Interfaced Sensors in Digital Medicine: from Materials to Applications. 2020 [cited 2021 Jul 26];Available from:

https://doi.org/10.1016/j.matt.2020.03.020

353. Nikoleli GP, Siontorou CG, Nikolelis DP, Bratakou S, Karapetis S, Tzamtzis N. Biosensors based on microfluidic devices lab-on-a-chip and microfluidic technology. In: Nanotechnology and Biosensors. Elsevier; 2018. page 375–94.

354. Patra B, Sharma M, Hale W, Utz M. Time-resolved non-invasive metabolomic monitoring of a single cancer spheroid by microfluidic NMR. Sci Rep [Internet] 2021 [cited 2021 Jan 15];11(1):53. Available from: https://doi.org/10.1038/s41598-020-79693-1

355. ecobee. Donate your Data | ecobee | Smart Home Technology [Internet]. 2018 [cited 2019 Jan 24];Available from: https://www.ecobee.com/donateyourdata/

356. Koskimäki H, Kinnunen H, Kurppa T, Röning J. How Do We Sleep - a Case Study of Sleep Duration and Quality Using Data from Oura Ring [Internet]. New York, New York, USA: Association for Computing Machinery, Inc; 2018 [cited 2020 Nov 10]. Available from: https://doi.org/10.1145/3267305.3267697

357. Baloch Z, Shaikh FK, Unar MA. A context-aware data fusion approach for health-IoT. Int J Inf Technol [Internet] 2018 [cited 2020 Dec 15];10(3):241–5. Available from: https://doi.org/10.1007/s41870-018-0116-1

358. Mora H, Gil D, Terol RM, Azorín J, Szymanski J. An IoT-Based Computational Framework for Healthcare Monitoring in Mobile Environments. Sensors [Internet] 2017 [cited 2020 Aug 13];17(10):2302. Available from: http://www.mdpi.com/1424-8220/17/10/2302

359. Iwendi C, Kumar Reddy PM, Reddy TG, Lakshmanna K, Kashif Bashir A, Piran J. A Metaheuristic Optimization Approach for Energy Efficiency in the IoT Networks. [cited 2021 Oct 15];Available from: https://e-space.mmu.ac.uk

360. Faverjon C, Berezowski J. Choosing the best algorithm for event detection based on the intended application: A conceptual framework for syndromic surveillance. J Biomed Inform [Internet] 2018;85:126–35. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1532046418301564

361. Jalali N, Sahu K, Oetomo A, Morita P. Understanding User Behavior Through the Use of Unsupervised Anomaly Detection: Proof of Concept Using Internet of Things Smart Home Thermostat Data for Improving Public Health Surveillance. JMIR mHealth uHealth [Internet] 2020 [cited 2021 Aug 5];8(11). Available from: https://pubmed.ncbi.nlm.nih.gov/33185562/

362. Natarajan A, Su H-W, Heneghan C. Assessment of physiological signs associated with COVID-19 measured using wearable devices. npj Digit Med [Internet] 2020 [cited 2020 Aug 21];3(1):156. Available from: https://doi.org/10.1101/2020.08.14.20175265

363. Jeong H, Rogers JA, Xu S. Continuous on-body sensing for the COVID-19 pandemic: Gaps and opportunities. Sci Adv [Internet] 2020 [cited 2020 Sep 10];6(36). Available from: www.columbia.edu/~jls106/galanti_shaman_ms_supp.pdf.

364. Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Early Detection Of COVID-19 Using A Smartwatch. medRxiv [Internet] 2020 [cited 2020 Sep 10];2020.07.06.20147512. Available from: http://medrxiv.org/content/early/2020/07/07/2020.07.06.20147512.abstract

365. Rahman MS, Peeri NC, Shrestha N, Zaki R, Haque U, Hamid SHA. Defending against the Novel Coronavirus (COVID-19) outbreak: How can the Internet of Things (IoT) help to save the world? Heal Policy Technol [Internet] 2020 [cited 2021 Feb 17];9(2):136–8. Available from: /pmc/articles/PMC7175864/

366. Wesson P, Hswen Y, Valdes G, Stojanovski K, Handley MA. Risks and Opportunities to Ensure Equity in the Application of Big Data Research in Public Health. Annu Rev Public Health [Internet] 2022;43(1):59–78. Available from: https://www.annualreviews.org/doi/10.1146/annurev-publhealth-051920-110928

367. Ibrahim SA, Charlson ME, Neill DB. Big Data Analytics and the Struggle for Equity in Health Care: The Promise and Perils. Heal Equity [Internet] 2020;4(1):99–101. Available from: https://www.liebertpub.com/doi/10.1089/heq.2019.0112

368. StatisticsCanada. Canadian Economic Dashboard and COVID-19 [Internet]. 2021 [cited 2021 Jul 9];Available from: https://www150.statcan.gc.ca/n1/pub/71-607-x/71-607-x2020009-eng.htm

369. Cowling B. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. Lancet Public Heal 2020;5(5):e279–88.

370. Rundle CW, Presley CL, Militello M, Barber C, Powell DL, Jacob SE, et al. Hand hygiene during COVID-19: Recommendations from the American Contact Dermatitis Society. J Am Acad Dermatol [Internet] 2020;83(6):1730–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0190962220322568

371. Karaivanov A, Lu SE, Shigeoka H, Chen C, Pamplona S. Face masks, public policies and slowing the spread of COVID-19: Evidence from Canada. J Health Econ [Internet] 2021;78:102475. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0167629621000606

372. Chinazzi M. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science (80-) 2020;368(6489):395–400.

373. Robillard R, Dion K, Pennestri M, Solomonova E, Lee E, Saad M, et al. Profiles of sleep changes during the COVID-19 pandemic: Demographic, behavioural and psychological factors. J Sleep Res [Internet] 2021;30(1). Available from: https://onlinelibrary.wiley.com/doi/10.1111/jsr.13231

374. Costi S, Paltrinieri S, Bressi B, Fugazzaro S, Giorgi Rossi P, Mazzini E. Poor Sleep during the First Peak of the SARS-CoV-2 Pandemic: A Cross-Sectional Study. Int J Environ Res Public Health [Internet] 2021;18(1):306. Available from: https://www.mdpi.com/1660-4601/18/1/306

375. Kephart JL, Delclòs-Alió X, Bilal U, Sarmiento OL, Barrientos-Gutiérrez T, Ramirez-Zea M, et al. The impact of population mobility on COVID-19 incidence and socioeconomic disparities at the sub-city level in 314 Latin American cities. medRxiv [Internet] 2021 [cited 2021 Jul 8];2021.04.13.21255413. Available from: https://www.medrxiv.org/content/10.1101/2021.04.13.21255413v1

376. Ross R, Chaput JP, Giangregorio LM, Janssen I, Saunders TJ, Kho ME, et al. Canadian 24-Hour Movement Guidelines for Adults aged 18-64 years and Adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. Appl Physiol Nutr Metab [Internet] 2020;45(10):S57--S102. Available from: https://pubmed.ncbi.nlm.nih.gov/33054332/

377. PHAC. 24 Hour Movement Guidelines for Children and Youth: An Integration of Physical Activity, Sedentary Behaviour, and Sleep - Canada.ca [Internet]. 2016 [cited 2020 Feb 14];Available from: https://www.canada.ca/en/public-health/services/health-promotion/healthy-living/physical-activity/24-hour-movement-guidelines-children-youth.html

378. Sañudo B, Fennell C, Sánchez-Oliver AJ. Objectively-assessed physical activity, sedentary behavior, smartphone use, and sleep patterns preand during-COVID-19 quarantine in young adults from Spain. Sustain [Internet] 2020 [cited 2021 May 13];12(15):5890. Available from: www.mdpi.com/journal/sustainability

379. Meyer J, McDowell C, Lansing J, Brower C, Smith L, Tully M, et al. Changes in physical activity and sedentary behavior in response to covid-19 and their associations with mental health in 3052 us adults. Int J Environ Res Public Health [Internet] 2020 [cited 2021 May 13];17(18):1–13. Available from: https://pubmed.ncbi.nlm.nih.gov/32899495/

380. Qin F, Song Y, Nassis GP, Zhao L, Cui S, Lai L, et al. Prevalence of insufficient physical activity, sedentary screen time and emotional well_being during the early days of the 2019 novel coronavirus (COVID-19) outbreak in China: a national cross-sectional study. Lancet [Internet] 2020 [cited 2021 May 13];28. Available from: https://www.ssrn.com/abstract=3566176

381. Ong JL, Lau T, Massar SAA, Chong ZT, Ng BKL, Koek D, et al. COVID-19-related mobility reduction: heterogenous effects on sleep and physical activity rhythms. Sleep [Internet] 2021;44(2). Available from: http://www.ncbi.nlm.nih.gov/pubmed/32918076

382. Dutta K, Mukherjee R, Sen D, Sahu S. Effect of COVID-19 lockdown on sleep behavior and screen exposure time: an observational study among Indian school children. Biol Rhythm Res [Internet] 2020 [cited 2021 May 13];Available from: https://www.tandfonline.com/action/journalInformation?journalCode=nbrr20

383. Chattu V, Manzar M, Kumary S, Burman D, Spence D, Pandi-Perumal S. The Global Problem of Insufficient Sleep and Its Serious Public Health Implications. Healthcare 2018;7(1):1.

384. Canadian Society for Exercise Physiology. Canadian 24-hour movement guidelines [Internet]. 2021 [cited 2021 Jul 7];Available from: https://csepguidelines.ca/

385. Rollo S, Lang JJ, Roberts KC, Bang F, Carson V, Chaput J-P, et al. Health Reports Health Associations with Meeting the Canadian 24-Hour Movement Guidelines for Adults: Results from the Canadian Health Measures Survey. [cited 2022 Jan 24];Available from: https://www.doi.org/10.25318/82-003-x202200100002-eng

386. Silfee VJ, Haughton CF, Jake-Schoffman DE, Lopez-Cepero A, May CN, Sreedhara M, et al. Objective measurement of physical activity outcomes in lifestyle interventions among adults: A systematic review. Prev. Med. Reports2018;11:74–80.

387. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world [Internet]. Geneva, Switzerland: 2018. Available from: https://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf?sequence=1&isAllowed=y

388. World Health Organization. Physical activity factsheet [Internet]. 2020 [cited 2021 Nov 23];Available from: https://www.who.int/news-room/fact-sheets/detail/physical-activity

389. Huchuk B, Sanner S, O'Brien W. Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data. Build Environ [Internet] 2019 [cited 2019 Oct 23];160:106177. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0360132319303877

390. Jalali N, Sahu KS, Oetomo A, Morita PP. Understanding User Behavior Through the Use of Unsupervised Anomaly Detection: Proof of Concept Using Internet of Things Smart Home Thermostat Data for Improving Public Health Surveillance. JMIR mHealth uHealth [Internet] 2020;8(11):e21209. Available from:

https://www.scopus.com/inward/record.uri?eid=2-s2.0-85096202925&doi=10.2196%2F21209&partnerID=40&md5=f8fa87a1d75f42e421f613a1390a36d5

391. Jalali N, Sahu KS, Oetomo A, Morita PP. Utilization of smart thermostat data at the population level to identify sleep parameters and time spent at home (Preprint). JMIR mhealth uhealth [Internet] 2021 [cited 2022 Feb 7];Available from: https://preprints.jmir.org/preprint/28811

392. Kleiminger W, Mattern F, Santini S. Predicting household occupancy for smart heating control: A comparative performance analysis of state-of-the-art approaches. Energy Build [Internet] 2014 [cited 2021 Aug 27];85:493–505. Available from: https://linkinghub.elsevier.com/retrieve/pii/S037877881400783X

393. Jalali N, Sahu KS, Oetomo A, Morita PP. Usability of Smart Home Thermostat to Evaluate the Impact of Weekdays and Seasons on Sleep Patterns and Indoor Stay: Observational Study (Preprint). JMIR mHealth uHealth [Internet] 2021 [cited 2022 Feb 7];Available from: https://preprints.jmir.org/ojs/index.php/preprints/preprint/28811

394. Statistics Canada. Population estimates, quarterly [Internet]. 2021 [cited 2021 Aug 9];Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901

395. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National sleep foundation’s sleep time duration recommendations: Methodology and results summary. Sleep Heal 2015;1(1):40–3.

396. McCormack GR, Doyle-Baker PK, Petersen JA, Ghoneim D. Perceived anxiety and physical activity behaviour changes during the early stages of COVID-19 restrictions in community-dwelling adults in Canada: a cross-sectional study. BMJ Open [Internet]

2021;11(8):e050550. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2021-050550

397. Morin CM, Carrier J, Bastien C, Godbout R. Sleep and circadian rhythm in response to the COVID-19 pandemic. Can J Public Heal [Internet] 2020 [cited 2021 Sep 30];111(5):654–7. Available from: www.sleeponitcanada.ca

398. McCurdy LE, Winterbottom KE, Mehta SS, Roberts JR. Using Nature and Outdoor Activity to Improve Children's Health. Curr Probl Pediatr Adolesc Health Care [Internet] 2010 [cited 2022 Feb 23];40(5):102–17. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1538544210000441

399. Kerr J, Marshall S, Godbole S, Neukam S, Crist K, Wasilenko K, et al. The Relationship between Outdoor Activity and Health in Older Adults Using GPS. Int J Environ Res Public Health [Internet] 2012;9(12):4615–25. Available from: http://www.mdpi.com/1660-4601/9/12/4615

400. Chang VC, Chaput J-P, Roberts KC, Jayaraman G, Do MT. Factors associated with sleep duration across life stages: results from the Canadian Health Measures Survey. Heal Promot Chronic Dis Prev Canada [Internet] 2018 [cited 2020 Feb 14];38(11):404–18. Available from: https://www.canada.ca/en/public-health/services/reports-publications/health-promotion-chronic-disease-prevention-canada-research-policy-practice/vol-38-no-11-2018/factors-sleep-duration-life-stages-canadian-health-measures-survey.html

401. Manferdelli G, La Torre A, Codella R. Outdoor physical activity bears multiple benefits to health and society. J Sports Med Phys Fitness [Internet] 2019;59(5):868–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30650943

402. Papadimitriou DT, Vassaras AK, Holick MF. Association between population vitamin D status and SARS-CoV-2 related serious-critical illness and deaths: An ecological integrative approach. World J Virol [Internet] 2021 [cited 2022 Feb 23];10(3):111. Available from: /pmc/articles/PMC8152454/

403. Thorp AA, Owen N, Neuhaus M, Dunstan DW. Sedentary behaviors and subsequent health outcomes in adults: A systematic review of longitudinal studies, 19962011. Am. J. Prev. Med.2011;41(2):207–15.

404. Kazmi SSH, Hasan K, Talib S, Saxena S. COVID-19 and Lockdwon: A Study on the Impact on Mental Health. SSRN Electron J 2020;

405. Larouche R, Garriguet D, Gunnell KE, Goldfield GS, Tremblay MS. Outdoor time, physical activity, sedentary time, and health indicators at ages 7 to 14: 2012/2013 Canadian Health Measures Survey. Heal reports [Internet] 2016;27(9):3–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27655167

406. Morrison H, Kristjansson E, Scientic A, Mccormack G, Poulin M, Odrowaz-Pieniazek J, et al. Chronic Diseases and Injuries in Canada Centers for Disease Control and Prevention Monitoring chronic diseases in Canada: the Chronic Disease Indicator Framework. [cited 2018 Oct 11];Available from: www.publichealth.gc.ca/cdic

407. Hamari L, Kullberg T, Ruohonen J, Heinonen OJ, Díaz-Rodríguez N, Lilius J, et al. Physical activity among children: Objective measurements using Fitbit One® and ActiGraph. BMC Res Notes 2017;10(1).

408. Freedson PS, Miller K. Objective monitoring of physical activity using motion sensors and heart rate. Res Q Exerc Sport 2000;71:21–9.

409. Gupta D, Bhatt S, Gupta M, Tosun AS. Future Smart Connected Communities to Fight

COVID-19 Outbreak. Internet of Things 2021;13:100342.

410. Simone, Marzolini B, Bulgheroni M. Smartphone-Based Passive Sensing for Behavioral and Physical Monitoring in Free-Life Conditions: Technical Usability Study. JMIR Biomed Eng 2021;6(2)e15417 https//biomedeng.jmir.org/2021/2/e15417 [Internet] 2021 [cited 2021 Jul 8];6(2):e15417. Available from: https://biomedeng.jmir.org/2021/2/e15417

411. André, Johannessen E, Hartvigsen G, Grimsgaard S, Hopstock LA. Consumer-Based Activity Trackers as a Tool for Physical Activity Monitoring in Epidemiological Studies During the COVID-19 Pandemic: Development and Usability Study. JMIR Public Heal Surveill 2021;7(4)e23806 https//publichealth.jmir.org/2021/4/e23806 [Internet] 2021 [cited 2021 Jul 8];7(4):e23806. Available from: https://publichealth.jmir.org/2021/4/e23806

412. Zeng C, Zhang J, Li Z, Sun X, Olatosi B, Weissman S, et al. Spatial-temporal relationship between population mobility and COVID-19 outbreaks in South Carolina: A time series forecasting analysis. medRxiv [Internet] 2021 [cited 2021 Jul 30];2021.01.02.21249119. Available from: https://www.medrxiv.org/content/10.1101/2021.01.02.21249119v2

413. Sulyok M, Walker MD. Mobility and COVID-19 mortality across Scandinavia: A modeling study. Travel Med Infect Dis 2021;41:102039.

414. MARCUS MKLLJS. The impact of COVID-19 restrictions on individual mobility [Internet]. 2020 [cited 2020 May 15];Available from: https://www.bruegel.org/2020/05/the-impact-of-covid-19-restrictions-on-individual-mobility/

415. González MC, Hidalgo CA, Barabási A-L. Understanding individual human mobility patterns. Nature [Internet] 2008;453(7196):779–82. Available from:

https://www.nature.com/articles/nature06958

416. Kang Y, Gao S, Liang Y, Li M, Rao J, Kruse J. Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic. Sci Data 2020 71 [Internet] 2020 [cited 2021 Aug 4];7(1):1–13. Available from: https://www.nature.com/articles/s41597-020-00734-5

417. Eagle N, (Sandy) Pentland A. Reality mining: sensing complex social systems. Pers Ubiquitous Comput 2005 104 [Internet] 2005 [cited 2021 Aug 11];10(4):255–68. Available from: https://link.springer.com/article/10.1007/s00779-005-0046-3

418. World Health Organization. WHO Director-General’s opening remarks at the media briefing on COVID-19 - 11 March 2020 [Internet]. 2020 [cited 2021 May 14];Available from: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

419. Nouvellet P, Bhatia S, Cori A, Ainslie KEC, Baguelin M, Bhatt S, et al. Reduction in mobility and COVID-19 transmission. Nat Commun 2021 121 [Internet] 2021 [cited 2021 Jul 8];12(1):1–9. Available from: https://www.nature.com/articles/s41467-021-21358-2

420. Flaxman S. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 2020;584(7820):257–61.

421. Kishore K, Jaswal V, Verma M, Koushal V. Exploring the Utility of Google Mobility Data During the COVID-19 Pandemic in India: Digital Epidemiological Analysis. JMIR Public Heal Surveill [Internet] 2021 [cited 2021 Aug 9];7(8):e29957. Available from: http://preprints.jmir.org/preprint/29957/accepted

422. Hanvey L. What mobility data is telling us about COVID-19 [Internet]. 2020 [cited 2020 May 15];Available from: https://urbanlogiq.com/the-road-to-recovery-what-mobility-data-

is-telling-us-about-covid-19-and-how-it-can-help-us-plan-economic-recovery/

423. Yilmazkuday H. Stay-at-home works to fight against COVID-19: International evidence from Google mobility data. https://doi.org/101080/1091135920201845903 [Internet] 2021 [cited 2021 Aug 4];31(1–4):210–20. Available from: https://www.tandfonline.com/doi/abs/10.1080/10911359.2020.1845903

424. Badr H. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 2020;20(11):1247–54.

425. Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 2020;368(6490):493–7.

426. Buckee CO, Balsari S, Chan J, Crosas M, Dominici F, Gasser U, et al. Aggregated mobility data could help fight COVID-19. Science (80-.).2020;368(6487):145–6.

427. Thums M, Fernández-Gracia J, Sequeira AMM, Eguíluz VM, Duarte CM, Meekan MG. How Big Data Fast Tracked Human Mobility Research and the Lessons for Animal Movement Ecology. Front Mar Sci [Internet] 2018 [cited 2021 Jul 30];5(FEB):21. Available from: http://journal.frontiersin.org/article/10.3389/fmars.2018.00021/full

428. WuFang-jing, ChenYing-Jun, SouSok-Ian. CoCo: Quantifying Correlations between Mobility Traces using Sensor Data from Smartphones. ACM Trans Internet Things [Internet] 2021 [cited 2021 Aug 4];2(3):1–22. Available from: https://dl.acm.org/doi/abs/10.1145/3457139

429. Google LLC. Google COVID-19 Community Mobility Reports [Internet]. [cited 2020 May 13];Available from: https://www.google.com/covid19/mobility/

430. Python Software Foundation. Welcome to Python [Internet]. 2021 [cited 2021 Aug

9];Available from: https://www.python.org/

431. Microsoft. What is Azure Databricks? [Internet]. 2021 [cited 2021 Aug 9];Available from: https://docs.microsoft.com/en-us/azure/databricks/scenarios/what-is-azure-databricks

432. Jupyter Team. The Jupyter Notebook — Jupyter Notebook 6.4.0 documentation [Internet]. 2015 [cited 2021 Aug 9];Available from: https://jupyter-notebook.readthedocs.io/en/stable/

433. RStudio Team. RStudio: Integrated Development for R [Internet]. PBC, Boston, MA: 2021 [cited 2021 Aug 9]. Available from: https://www.rstudio.com/

434. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: 2021 [cited 2021 Aug 9]. Available from: https://www.r-project.org/

435. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. J Open Source Softw [Internet] 2019 [cited 2021 Aug 9];4(43):1686. Available from: https://joss.theoj.org/papers/10.21105/joss.01686

436. Matt D, Davis V. R package timetk version 2.6.1-A Tool Kit for Working with Time Series in R [Internet]. Comprehensive R Archive Network (CRAN); 2021 [cited 2021 Aug 9]. Available from: https://cran.r-project.org/package=timetk

437. time-time.net. Canada Time Zone Map with provinces with cities [Internet]. 2021 [cited 2021 Jul 7];Available from: http://time-time.net/times/time-zones/usa-canada/canada-time-zone-map.php

438. Canadian Institute for Health Information. COVID-19 Intervention Scan [Internet]. 2021 [cited 2021 Jul 9];Available from: https://www.cihi.ca/en/covid-19-intervention-scan

439. Canada's Health Care System - Canada.ca [Internet]. [cited 2022 Feb 17];Available from: https://www.canada.ca/en/health-canada/services/health-care-system/reports-

publications/health-care-system/canada.html#a4

440. Federal/Provincial/Territorial Public Health Response Plan for Biological Events - Canada.ca [Internet]. [cited 2022 Feb 17];Available from: https://www.canada.ca/en/public-health/services/emergency-preparedness/public-health-response-plan-biological-events.html#s2-3

441. Sedgwick P. Pearson's correlation coefficient. BMJ [Internet] 2012 [cited 2022 Feb 17];345(jul04 1):e4483–e4483. Available from: http://www.bmj.com/permissionsSubscribe:http://www.bmj.com/subscribeBMJ2012;345:e4483doi:10.1136/bmj.e4483

442. Cheong JH. Four ways to quantify synchrony between time series data [Internet]. Towar. Data Sci.2019 [cited 2021 Aug 9];Available from: https://towardsdatascience.com/four-ways-to-quantify-synchrony-between-time-series-data-b99136c4a9c9

443. Nwogu EC, Iwueze IS, Nlebedim VU. Some Tests for Seasonality in Time Series Data. J Mod Appl Stat Methods [Internet] 2016;15(2):382–99. Available from: http://digitalcommons.wayne.edu/jmasm/vol15/iss2/24

444. Larson MG. Analysis of Variance. Circulation [Internet] 2008;117(1):115–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18172051

445. Vallis O, Hochenbaum J, Kejariwal A. A Novel Technique for Long-Term Anomaly Detection in the Cloud [Internet]. In: Proceedings of the 6th USENIX Conference on Hot Topics in Cloud Computing. USA: USENIX Association; 2014 [cited 2021 Aug 11]. Available from: https://dl.acm.org/doi/10.5555/2696535.2696550

446. Robert C, William C, Irma T. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. J Off Stat [Internet] 1990;6(1):3–73. Available from:

https://www.wessa.net/download/stl.pdf

447. twitter. GitHub - twitter/AnomalyDetection: Anomaly Detection with R [Internet]. 2015 [cited 2021 Aug 20];Available from: https://github.com/twitter/AnomalyDetection

448. Dancho M, Vaughan D. Tidy Anomaly Detection [R package anomalize version 0.2.2] [Internet]. 2020 [cited 2021 Aug 20];Available from: https://cran.r-project.org/package=anomalize

449. Asgari F, Gauthier V, Becker M. A survey on Human Mobility and its applications. 2013 [cited 2022 Feb 17];Available from: http://arxiv.org/abs/1307.0814

450. Xiao Y, Yang M, Zhu Z, Yang H, Zhang L, Ghader S. Modeling indoor-level non-pharmaceutical interventions during the COVID-19 pandemic: A pedestrian dynamics-based microscopic simulation approach. Transp Policy 2021;109:12–23.

451. Lin C, Song Z, Song H, Zhou Y, Wang Y, Wu G. Differential Privacy Preserving in Big Data Analytics for Connected Health. J Med Syst [Internet] 2016 [cited 2018 Feb 9];40(4):97. Available from: http://link.springer.com/10.1007/s10916-016-0446-0

452. Syagnik (Sy)Banerjee, Hemphill T, Longstreet P. Wearable devices and healthcare: Data sharing and privacy. Inf Soc [Internet] 2018 [cited 2018 May 16];34(1):49–57. Available from: http://www.tandfonline.com/action/journalInformation?journalCode=utis20

453. Falck RS, Davis JC, Khan KM, Handy TC, Liu-Ambrose T. A Wrinkle in Measuring Time Use for Cognitive Health: How should We Measure Physical Activity, Sedentary Behaviour and Sleep? Am J Lifestyle Med [Internet] 2021;155982762110314. Available from: http://journals.sagepub.com/doi/10.1177/15598276211031495

454. Churová V, Vyškovský R, Maršálová K, Kudláček D, Schwarz D. Anomaly Detection Algorithm for Real-World Data and Evidence in Clinical Research: Implementation,

Evaluation, and Validation Study. JMIR Med informatics [Internet] 2021;9(5):e27172. Available from: http://www.ncbi.nlm.nih.gov/pubmed/33851576

455. Haque S, Rahman M, Aziz S. Sensor Anomaly Detection in Wireless Sensor Networks for Healthcare. Sensors [Internet] 2015 [cited 2018 Feb 9];15(12):8764–86. Available from: http://www.mdpi.com/1424-8220/15/4/8764

456. Saqaeeyan S, Haj Seyyed Javadi H, Amirkhani H. A Novel Probabilistic Hybrid Model to Detect Anomaly in Smart Homes. Comput Model Eng Sci [Internet] 2019;121(3):815–34. Available from: http://tsp.techscience.com//uploads/attached/file/20191219/20191219011745_90859.pdf

457. Šabić E, Keeley D, Henderson B, Nannemann S. Healthcare and anomaly detection: using machine learning to predict anomalies in heart rate data. AI Soc [Internet] 2021;36(1):149–58. Available from: https://link.springer.com/10.1007/s00146-020-00985-1

458. Yechezkel M, Weiss A, Rejwan I, Shahmoon E, Ben-Gal S, Yamin D. Human mobility and poverty as key drivers of COVID-19 transmission and control. BMC Public Health [Internet] 2021 [cited 2021 Aug 24];21(1):596. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-10561-x

459. Charu V, Zeger S, Gog J, Bjørnstad ON, Kissler S, Simonsen L, et al. Human mobility and the spatial transmission of influenza in the United States. PLOS Comput Biol [Internet] 2017 [cited 2021 Aug 24];13(2):e1005382. Available from: https://dx.plos.org/10.1371/journal.pcbi.1005382

460. KARTAL MT, DEPREN Ö, KILIÇ DEPREN S. The relationship between mobility and COVID-19 pandemic: Daily evidence from an emerging country by causality analysis.

Transp Res Interdiscip Perspect 2021;10:100366.

461. Engle S, Stromme J, Zhou A. Staying at Home: Mobility Effects of COVID-19. SSRN Electron J [Internet] 2020 [cited 2020 May 14];Available from: https://voxeu.org/article/staying-home-mobility-effects-covid-19

462. Huang X, Li Z, Jiang Y, Ye X, Deng C, Zhang J, et al. The characteristics of multi-source mobility datasets and how they reveal the luxury nature of social distancing in the U.S. during the COVID-19 pandemic. Int J Digit Earth 2021;14(4):424–42.

463. Rosa L, Silva F, Analide C. Mobile Networks and Internet of Things Infrastructures to Characterize Smart Human Mobility. Smart Cities [Internet] 2021;4(2):894–918. Available from: https://doi.org/10.3390/smartcities4020046

464. Nižetić S, Šolić P, González-de-Artaza DL-I, Patrono L. Internet of Things (IoT): Opportunities, issues and challenges towards a smart and sustainable future. J Clean Prod [Internet] 2020 [cited 2021 Aug 11];274:122877. Available from: /pmc/articles/PMC7368922/

465. Ontario Agency for Health Protection and Promotion (Public Health Ontario), Bray E, Bergeron K, Meserve A. Focus On : Measuring the policy-making process. Toronto,ON: Queen's Printer for Ontario; 2017.

466. Fadlallah R, El-Jardali F, Nomier M, Hemadi N, Arif K, Langlois E V., et al. Using narratives to impact health policy-making: a systematic review. Heal Res Policy Syst [Internet] 2019;17(1):26. Available from: https://health-policy-systems.biomedcentral.com/articles/10.1186/s12961-019-0423-4

467. Herweg N, Zahariadis N, Zohlnhöfer R. The Multiple Streams Framework: Foundations, Refinements, and Empirical Applications. Theor Policy Process [Internet] 2018 [cited

2021 Sep 13];17–53. Available from: https://www.taylorfrancis.com/chapters/edit/10.4324/9780429494284-2/multiple-streams-framework-foundations-refinements-empirical-applications-nicole-herweg-nikolaos-zahariadis-reimut-zohlnhöfer

468. Pierce JJ, Peterson HL, Jones MD, Garrard SP, Vu T. There and Back Again: A Tale of the Advocacy Coalition Framework. Policy Stud J [Internet] 2017;45(S1):S13–46. Available from: https://onlinelibrary.wiley.com/doi/10.1111/psj.12197

469. Boushey G. Punctuated Equilibrium Theory and the Diffusion of Innovations. Policy Stud J [Internet] 2012 [cited 2021 Sep 13];40(1):127–46. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1541-0072.2011.00437.x

470. Baumgartner F, … BJ-T of the policy, 2014 undefined. Punctuated equilibrium theory: Explaining stability and change in public policymaking. books.google.comPaperpile [Internet] [cited 2021 Sep 13];Available from: https://books.google.ca/books?hl=en&lr=&id=MzkGAwAAQBAJ&oi=fnd&pg=PA59&ots=wRNNQhjs2o&sig=5MOnKZzznR0bpqHpzullHI-NXKk

471. Canadian Institute for Health Information. Unintended consequences of COVID-19: Impact on harms caused by substance use. 2021;1–20.

472. Canadian Public Health Association. Review of Canada's Initial Response to the COVID-19 Pandemic [Internet]. 2021 [cited 2021 Sep 12];Available from: https://www.cpha.ca/review-canadas-initial-response-covid-19-pandemic

473. Public Health Agency of Canada, Tam T, Chief Public Health Offcer of Canada. A Vision to Transform Canada's Public Health System: The Chief Public Health Offcer of Canada's Report on the State of Public Health in Canada 2021 [Internet]. Ottawa (ON): 2021 [cited

2022 Jan 17]. Available from: https://www.canada.ca/en/public-health/corporate/publications/chief-public-health-officer-reports-state-public-health-canada/state-public-health-canada-2021.html

474. Canadian Institute for Health Information. Improved stroke data means better care and outcomes for patients [Internet]. 2021 [cited 2021 Jul 9];Available from: https://www.cihi.ca/en/improved-stroke-data-means-better-care-and-outcomes-for-patients

475. Majumder MS, Santillana M, Mekaru SR, McGinnis DP, Khan K, Brownstein JS. Utilizing Nontraditional Data Sources for Near Real-Time Estimation of Transmission Dynamics During the 2015-2016 Colombian Zika Virus Disease Outbreak. JMIR public Heal Surveill [Internet] 2016 [cited 2020 Aug 24];2(1):e30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27251981

476. Cabrita M, Tabak M, Vollenbroek-Hutten M. Mobile technology to monitor physical activity and wellbeing in daily life objective and subjective experiences of older adults. In: ICT4AWE 2017 - Proceedings of the 3rd International Conference on Information and Communication Technologies for Ageing Well and e-Health. SciTePress; 2017. page 164–71.

477. Amorim AB, Pappas E, Simic M, Ferreira ML, Tiedemann A, Jennings M, et al. Integrating Mobile health and Physical Activity to reduce the burden of Chronic low back pain Trial (IMPACT): A pilot trial protocol. BMC Musculoskelet Disord 2016;17(1).

478. Finger F, Genolet T, Mari L, de Magny GC, Manga NM, Rinaldo A, et al. Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks. Proc Natl Acad Sci [Internet] 2016;113(23):6421–6. Available from: http://www.pnas.org/lookup/doi/10.1073/pnas.1522305113

479. Storm FA, Heller BW, Mazzà C. Step detection and activity recognition accuracy of seven physical activity monitors. PLoS One 2015;10(3).

480. An HS, Jones GC, Kang SK, Welk GJ, Lee JM. How valid are wearable physical activity trackers for measuring steps? Eur J Sport Sci 2017;17(3):360–8.

481. Adam Noah J, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. J Med Eng Technol 2013;37(7):456–62.

482. Block VJ, Lizée A, Crabtree-Hartman E, Bevan CJ, Graves JS, Bove R, et al. Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. J Neurol 2017;264(2):316–26.

483. Lee H-C, Jung C-W. Vital Recorder-a free research tool for automatic recording of high-resolution time-synchronised physiological data from multiple anaesthesia devices. Sci Rep [Internet] 2018 [cited 2018 Feb 9];8(1):1527. Available from: http://www.nature.com/articles/s41598-018-20062-4

484. da Costa CA, Pasluosta CF, Eskofier B, da Silva DB, da Rosa Righi R. Internet of Health Things: Toward intelligent vital signs monitoring in hospital wards. Artif Intell Med [Internet] 2018 [cited 2018 Nov 15];89:61–9. Available from: https://www.sciencedirect.com/science/article/pii/S0933365717301367

485. Mahdavi H, Rosell-Ferrer J. In-bed vital signs monitoring system based on unobtrusive magnetic induction method with a concentric planar gradiometer. Physiol Meas [Internet] 2017 [cited 2018 Aug 20];38(6):1226–41. Available from: http://stacks.iop.org/0967-3334/38/i=6/a=1226?key=crossref.914844f0edc5f38bb4c545b4c203ce0c

486. Fung ICHC-HCH, Tse ZTH, Fu K-WW. The use of social media in public health

surveillance. 2015;6(2):3–6.

487. Public Health 2.0: Leveraging social media and technology for public health [Internet]. [cited 2020 Aug 11];Available from: https://www.researchgate.net/publication/266779337_Public_Health_20_Leveraging_social_media_and_technology_for_public_health

488. Social Media As A Public Health Surveillance Tool: Evidence And Prospects. [cited 2018 Jun 7];Available from: http://enterprise.sickweather.com/downloads/SW-SocialMedia_WhitePaper.pdf

489. Canadian Institute for Health Information. CIHI's Health Data and Information Governance and Capability Framework. 2020 [cited 2021 Jul 9];Available from: https://www.cihi.ca/sites/default/files/document/health-data-info-capability-framework-en.pdf

490. Government of Canada SC. Statistics: Power from Data! Confidentiality, privacy and security.

491. Goel V, Expert Advisory Group, Pan-Canadian Health Data Strategy. Building Canada's Health Data Foundation - Expert Advisory Group Report 2 [Internet]. 2021 [cited 2022 Jan 17]. Available from: https://www.canada.ca/en/public-health/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-02-building-canada-health-data-foundation.html

492. Goel V, Expert Advisory Group, Pan-Canadian Health Data Strategy. Pan-Canadian health strategy: Charting a Path toward Ambition- Expert Advisory Group Report 1. 2020.

493. Brownson RC, Chriqui JF, Stamatakis KA. Understanding Evidence-Based Public Health

Policy. Am J Public Health [Internet] 2009 [cited 2021 Sep 13];99(9):1576. Available from: /pmc/articles/PMC2724448/

494. Ontario Human Rights Commission. Policy statement on a human rights-based approach to managing the COVID-19 pandemic | Ontario Human Rights Commission [Internet]. 2021 [cited 2021 Sep 12];Available from: http://www.ohrc.on.ca/en/policy-statement-human-rights-based-approach-managing-covid-19-pandemic

495. Sedig K, Ola O. The Challenge of Big Data in Public Helth: An Opportunity for Visual Analytics. Online J Public Health Inform [Internet] 2014 [cited 2018 Jun 18];5(3):223. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24678376

496. Wong ZSY, Zhou J, Zhang Q. Artificial Intelligence for infectious disease Big Data Analytics. Infect Dis Heal [Internet] 2018 [cited 2018 Dec 24];0(0):44–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2468045118301445

497. Lee D, Rushworth A, Sahu SK. A Bayesian localized conditional autoregressive model for estimating the health effects of air pollution. Biometrics [Internet] 2014 [cited 2020 Feb 19];70(2):419–29. Available from: http://doi.wiley.com/10.1111/biom.12156

498. Wang S, Feng J, Liu G. Application of seasonal time series model in the precipitation forecast. Math Comput Model 2013;58(3–4):677–83.

Appendices

Appendix-1 Statistical and Advanced Methods for Time Series Data Analysis

Model	Examples and Description	Advantages and limitations
Univariate Statistical Methods		
Statistical process control (SPC) inspired model (C1,C2,C3, CUSUM (cumulative sum),W algorithm) [208]	C1 algorithms compute a mean and standard deviation using a short sliding interval of historical data to adjust time trends, cyclic patterns, and other time-dependent effects. C2 algorithm adds a two-day guard-band, and the C3 algorithm sums values of the test statistic over three days in a manner. W algorithms add adjustments for weekdays and weekends to C-algorithms. Cumulative sum (CUSUM)-sequential hypothesis test. [209,210].	C1-C3 algorithms were used widely in Syndromic surveillance systems for early event detection because they are easy to compute and interpret A stationary time series is a prerequisite for applying SPC or SPC-inspired methods. CUSUM and generalized linear models (GLM) often outperform the C1-C3 methods in early outbreak detection performance. Univariate prediction only.
Smoothing Methods	Moving average and exponentially weighted moving average methods. Splines- to address seasonal effects Gaussian kernel smoothing, Quadratic kernel smoothing Loess smoothing	This method can generate accurate forecasts and detect anomalies. Have the capacity to handle trend, seasonality, and day of the week effect. This method cannot incorporate covariates.
Regression method		
Generalized linear models (GLM)	The current observation is a linear function of the previous observation with white noise. Seasonality is usually modelled in a GLM or Serfling model using dummy variables or trigonometric terms. [212]. Adaptive GLM- short historical windows as a sliding baseline.	GLM and Serfling's methods are commonly used for disease surveillance because they produce results that are easy to interpret, they can model temporal and cyclic patterns, and they can exploit external information Outperformed C and W algorithms. These models can handle covariates.
AR Model	The autoregressive model uses the dependent relationship between an observation and some number of lagged observations. There are different modified versions of the AR models that have been used in the public health literature, such as the vector autoregressive model for life expectancy, public health spending and economic growth in Nigeria in 2013 [213] or the Bayesian conditional autoregressive model for estimating health effect of air pollution in 2014 [497].	The benefit of using AR methods is autocorrelation function can be used to tell if there is a lack of randomness. It can forecast any recurring patterns in the data. There must be an autocorrelation coefficient that should not be less than 0.5 for it to be suitable. It can only be used when predicting things related to economics based on the pre-existing time.

		Used when something is considerably affected by social aspects.
MA Model	Moving Average Model uses the dependency between an observation and a residual error from a moving average model applied to lagged observations. Although it looks like a regression model, the difference is that the *weight* is not observable.	Less prone to a lot of false signals Requires maintaining a history of different periods for each forecasted period. Often overlooks complex relationships mentioned in the data. Do not respond to the fluctuation that takes place for a reason, for example, cycles and seasonal impacts.
ARMA Autoregressive Moving Average Model	AR and MA are two widely used linear models that work on stationary time series. Autoregressive and moving average models can be combined to form ARMA models.	If the available data are multivariate and relationships between covariates and outcomes are non-linear, ARMA methods may have poor performance.
ARIMA [214–216] Auto-Regressive Integrated Moving Average	I stand for Integrated and differencing raw observations to make the time series stationary, and I is a preprocessing procedure to "stationarize" time series if needed.	One problem in the ARIMA model is the lack of seasonality
SARIMA [217] Seasonality adjusted Auto-Regressive Integrated Moving Average	A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA models, denoted as ARIMA (p, d, q) (P, D, Q)m, i.e., $\varphi(B^m)\ \varphi(B)\ (1 - B^m)^D\ (1 - B)^d\ X_t = \theta(B^m)\ \theta(B)\ w_t$. where m represents the number of observations per year. The seasonal part of the model consists of terms similar to the non-seasonal components but involves backshifts of the seasonal period [218,219].	It can only extract linear relationships within the time series data [197] Failure in forecasting, especially if the sequence of time series has abnormal changes [498].
Advanced analytics methods		
Bayesian methods [199,220]	PANDA- Population-wide anomaly detection and assessment MCMC- Markov chain Monte Carlo	This has the ability to explicitly model uncertainty from different sources and then propagate the uncertainties through to the results. This method also has the capacity to model data from individuals.
Markov methods	HMRF- hidden Markov random field HMM- Hidden Markov model	All the parameters are easy to interpret, and the model can be adapted easily to different epidemiological situations [221].
Multivariate analysis	PCA-Principal components analysis Multivariate CUSUM Parallel surveillance Ensemble approach	Data from different sources and detect anomalies using the integrated data.
Artificial intelligence-based methods		
Machine learning Methods	Neural network Gradient boosted regression trees	Better prediction performance compared to autoregressive

	LASSO- Least absolute shrinkage and selection operator Support vector regression	models applied to single data sources. Most ML methods make no assumptions about the distribution of the data. However, some practical limitations remain, including the need for a large amount of training data, the risk of over-fitting, and the need for expertise in tuning parameters within these models [219].
Deep Learning Methods	Feedforward neural network (FNN)	Better prediction models Needs complex computing infrastructures

Appendix-2 Schema of the Metadata from the Donate your Data by ecobee

Schema of Metadata
Identifier
Model
UserID
Country
ProvinceState
City
Floor_Area__ft2_
Number_of_Floors
Age_of_Home__years_
Number_of_Occupants
installedCoolStages
installedHeatStages
allowCompWithAux
Has_Electric
Has_a_Heat_Pump
Auxilliary_Heat_Fuel_Type
Number_of_Remote_Sensors
First_Connected

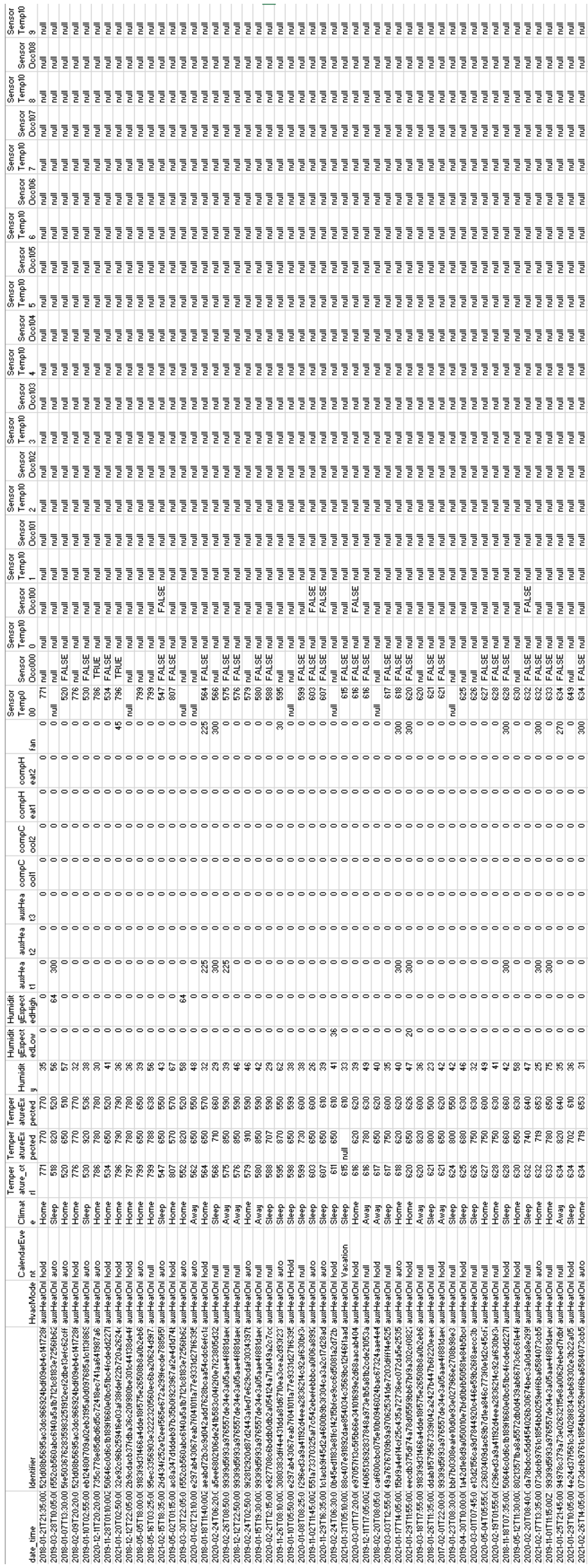
Appendix-3 Snapshot of the Metadata from the Donate your Data by ecobee

Identifier	Model	UserID	Country	Province State	City	Floor_Area__ft2_	Number_of_Floors	Age_of_Home__years_	Number_of_Occupants	installedCoolStages	installedHeatStages	allowCompWithAux	Has_Electric	Has_a_Heat_Pump	Auxilliary_Heat_Fuel_Type	Number_of_Remote_Sensors	First_Connected
7d29d38c4acdf7707d6a518e5422ae2a8fa32f8d	ecobee3Lite	00334d325b56198f5242b29e0799e1b9a639b020	null	null	null	0	null	0	0	0	1	FALSE	FALSE	FALSE	Gas	0	2019-06-21T02:02:01Z
33c989c91d38de97c38b6b38ddc9f7954dbc54f0	ecobee4	003d71b48cbec95b33e1a1fa508adf805a8a62a4	null	null	null	0	null	0	0	1	1	TRUE	FALSE	TRUE	Electric	0	2019-06-27T00:33:44Z
d57be13b6f9923263d099ce9862862ab308369b5	ecobee4	004b082c6f9d8fe3f30746a9a34d1fbc65747377	null	null	null	501	Condomir	29	0	1	1	TRUE	FALSE	TRUE	Electric	1	2019-06-23T15:39:06Z
f25be98df99b22e141c1fbb8dc6a74e5942ca328	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T00:48:06Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
ba4abfeaa972d4e66d45a6c4f67d16d180ebe8f6	ecobee3Lite	004eadaad8e95ca71ea131d24353f75418baf5e0	null	null	null	2501	Detached	19	0	1	1	FALSE	FALSE	FALSE	Gas	2	2019-03-07T01:15:30Z
be3ef32e48140f627814af00adada3c734415d0c	ecobee4	0068ae223dd75a7bfe36138abe3a02b4eaccdc7d	null	null	null	2000	detached	5	0	1	1	TRUE	FALSE	TRUE	Electric	0	2019-06-21T23:45:05Z
517333d9a23ecaf92a46041f683e703a5c57a9d2	ecobee4	0096f38e1b9c5a937ec5622489a690fa28ce94e3	null	null	null	1001	Other	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-25T18:34:00Z
0c0a8d03cd8df18a787acbe6d3fd75efe24b1297	ecobee4	00ab377828ea4b1e4ffbec28b6d3f3b3bd4abc2c	null	null	null	0	null	0	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-07-27T01:33:37Z
bbf558ef0939ca6c3e0233453cdce9f5b362d4a5	ecobee4	00e6704c0b1faa70521fc6032ffffdb250a35eef	null	null	null	0	null	0	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-03-06T17:42:19Z
ab817948163b63bddcfb8396d0d283fc9c36c144	ecobee3Lite	00ee68d71ff1ee0b8d08f813be601115bb6277e7	null	null	null	0	null	0	0	1	1	FALSE	FALSE	TRUE	Electric	1	2019-04-12T22:50:16Z
bee4cbac31f1a05a267d1ebc3ad4c2878840e2d1	ecobee3	0109b9ed538410636f7f3a37295ab52118c03035	null	null	null	1501	Other	29	0	1	1	FALSE	FALSE	TRUE	Electric	1	2019-06-12T19:10:03Z
5aa260570b675344740601de5f12b65d1efcf4fd	ESTWVC	016497a9c794a9ff3d4b89a7eb6a0568ee910c8f	null	null	null	1501	Detached	59	0	0	2	TRUE	FALSE	FALSE	Gas	0	2019-08-29T01:53:30Z
cbe92b1f933ab1a381d30ca5eb5a29cb24d4095d	ecobee4	018e3ae629b1cd0c529b8d76dedef7443828cff8	null	null	null	0	null	0	0	1	1	TRUE	FALSE	FALSE	Gas	1	2019-07-10T02:17:07Z
7e38aa4e79e817b153fe329ff0526be607cf5570	ecobee4	018e3ae629b1cd0c529b8d76dedef7443828cff8	null	null	null	0	null	0	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-07-09T01:38:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
ce108f382c8932ccfdcfe70692df8422a88e7377	ecobee4	01a87a13c7d4cc6347ba185fe469047ab7221eeb	null	null	null	2501	Detached	9	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-06-18T23:21:20Z
81defeaa021065ef468c566a80c6cf9606ba9c1d	ecobee3Lite	01ccf48cfc2a35db894f9526d897db593d427e32	null	null	null	0	null	0	0	1	1	FALSE	FALSE	TRUE	Electric	0	2019-03-20T16:39:36Z
57ce4cdee73a53cf70b9b5d80cebc3ab5354e183	ecobee4	01e0378826647a9f234db5bf619a234085e0187f	null	null	null	0	Other	9	0	1	1	FALSE	FALSE	FALSE	Gas	1	2019-05-20T00:04:31Z
d8b73696195c295437fb3a58fe571851fab5ae25	ecobee3Lite	01f9da0cfa1f781b5f85f73bac6b1b5a21b81b0a	null	null	null	1500	condomin	35	0	1	0	FALSE	FALSE	TRUE	Electric	0	2019-05-13T01:09:29Z
ad6d9681e5411c1a26248c5804564492054f57ce	ecobee4	0211c468fb7f0917961d738aa599f990c7ffecdf	US	IL	null	0	null	0	0	1	1	TRUE	FALSE	FALSE	Gas	0	2019-03-09T01:21:00Z

Appendix-4 Schema of the Thermostat data from the Donate your Data by ecobee

Field name	Type
date_time	TIMESTAMP
Identifier	STRING
HvacMode	STRING
CalendarEvent	STRING
Climate	STRING
Temperature_ctrl	INTEGER
TemperatureExpectedCool	INTEGER
TemperatureExpectedHeat	INTEGER
Humidity	INTEGER
HumidityExpectedLow	INTEGER
HumidityExpectedHigh	INTEGER
auxHeat1	INTEGER
auxHeat2	INTEGER
auxHeat3	INTEGER
compCool1	INTEGER
compCool2	INTEGER
compHeat1	INTEGER
compHeat2	INTEGER
fan	INTEGER
SensorTemp000	INTEGER
SensorOcc000	BOOLEAN
SensorTemp100	INTEGER
SensorOcc100	BOOLEAN
SensorTemp101	INTEGER
SensorOcc101	BOOLEAN
SensorTemp102	INTEGER
SensorOcc102	BOOLEAN
SensorTemp103	INTEGER
SensorOcc103	BOOLEAN
SensorTemp104	INTEGER
SensorOcc104	BOOLEAN
SensorTemp105	INTEGER
SensorOcc105	BOOLEAN
SensorTemp106	INTEGER
SensorOcc106	BOOLEAN
SensorTemp107	INTEGER
SensorOcc107	BOOLEAN
SensorTemp108	INTEGER
SensorOcc108	INTEGER
SensorTemp109	INTEGER

Appendix-5 Snapshot of the Thermostat Data from the Donate your Data by ecobee



Appendix-6 Policy change with regards to COVID-19, a) Canada, b) Ontario, c) Alberta d) Quebec and e) British Columbia

Source: https://www.cihi.ca/en/covid-19-intervention-timeline-in-canada

a) Canada

b) Ontario

Choose a jurisdiction
ONTARIO
Intervention timeline legend
New
Duration
Eased
Choose a start date
JANUARY 1, 2020
Choose an end date
MARCH 22, 2021
Choose categories
CASE MANAGEMENT AND 8 MORE
EXPORT CSV
Daily COVID-19 cases
from January 1
to March 22, 2021
4,000
3,000
2,000
1,000
0
Spring break
Winter break
Postponed Mar
Time window
JAN
APR
JUL
OCT
JAN
Intervention categories
ONTARIO
Case management
Closures/openings
Distancing
Health services
Health workforce
Public information
State of emergency
Travel
Vaccine
CANADA

c) Alberta

Choose a jurisdiction
ALBERTA
Intervention timeline legend
New
Duration
Eased
Choose a start date
JANUARY 1, 2020
Choose an end date
MARCH 22, 2021
Choose categories
CASE MANAGEMENT AND 8 MORE
EXPORT CSV
Spring break
Winter break
Daily COVID-19 cases
from January 1
to March 22, 2021
4,000
2,000
0
Time window
JAN
APR
JUL
OCT
JAN
Intervention categories
ALBERTA
Case management
Closures/openings
Distancing
Health services
Health workforce
Public information
State of emergency
Travel
Vaccine
CANADA

d) Quebec

Choose a jurisdiction
QUEBEC
Intervention timeline legend
New
Duration
Eased
Choose a start date
JANUARY 1, 2020
Choose an end date
MARCH 22, 2021
Choose categories
CASE MANAGEMENT AND 8 MORE
EXPORT CSV
Daily COVID-19 cases
from January 1
to March 22, 2021
Spring break
Winter break
6,000
4,000
2,000
0
Time window
JAN
APR
JUL
OCT
JAN
Intervention categories
QUEBEC
Case management
Closures/openings
Distancing
Health services
Health workforce
Public information
State of emergency
Travel
Vaccine
CANADA

e) British Columbia