**Forecasting the Effect of Lockdown Measures on the Spread of the Covid-19 Infection**

Directed Study Report

Department of Computer Science

The University of Western Ontario

Name: Piyush Adya

Student Number:251235087

Directed Study Supervisor: Mr. Anwar Haque, Dept. of Computer Science

December 2022

Table of Contents

[GLOSSARY: 3](#_Toc122640352)

[ABSTRACT: 4](#_Toc122640353)

[INTRODUCTION: 4](#_Toc122640354)

[RESEARCH OBJECTIVES: 5](#_Toc122640355)

[BACKGROUND AND RELATED WORK: 6](#_Toc122640356)

[CONSIDERATIONS AND DISCLAIMERS: 8](#_Toc122640357)

[METHODOLOGY: 9](#_Toc122640358)

[RESULT: 9](#_Toc122640359)

[DISCUSSION: 15](#_Toc122640360)

[CONCLUSIONS: 16](#_Toc122640361)

[FUTURE WORK AND LESSONS LEARNT: 17](#_Toc122640362)

[ACKNOWLEDGEMENTS: 18](#_Toc122640363)

[REFERENCES: 19](#_Toc122640364)

[APPENDIX A: 21](#_Toc122640365)

[APPENDIX B: 22](#_Toc122640366)

[APPENDIX C: 23](#_Toc122640367)

[APPENDIX D: 24](#_Toc122640368)

[APPENDIX E: 24](#_Toc122640369)

# GLOSSARY:

* Machine Learning (ML): Using an algorithm and statistical models, computer systems may learn from data and adapt. Machine learning will be used in this project to examine and learn from data that has been gathered from the public domain.
* Scikit-learn: A library for open-source data analysis. The ML model for this project was built using scikit-learn in the Python ecosystem.
* Pickle: It is a Python module for serializing and de-serializing Python object structures. This module offers a simple method for saving and importing machine learning models for use.
* MySQL:It is an open-source relational database management system (RDMS) developed by Oracle based on SQL.A database is a structured collection of data. This will be used throughout the project as a database to store and update the covid cases metrics.
* SQLAlchemy:SQLAlchemy is an open-source library that facilitates the communication between Python programs and databases. Throughout this project, this will be used to connect with the MySQL database.
* Active Cases: The number of cases that are still contagious and have not been deemed fully recovered.
* New Cases: The quantity of new infections that have occurred today. In fact, a new case would only be counted once it has been tested and confirmed, however for the sake of this study, the new cases would be counted as soon as they get infected.
* Non-Pharmaceutical Interventions (NPIs): Measures used to manage a public health emergency in the absence of pharmaceutical solutions like a treatment or vaccination.
* Daily New Infection Cases: The number of new infections that have been identified as occurring today. This will serve as the ML model's anticipated value throughout the duration of the project.

# ABSTRACT:

The COVID-19 pandemic profoundly affected the world, causing unprecedented global health and economic disaster. To mitigate the spread of the virus through interpersonal contact, a variety of lockdown strategies with various intensities have been implemented by governments around the world. In particular, types of stringent lockdown measures are enforced to curb the spread of the COVID-19 infection, especially to ease the stress on national healthcare systems. In this research, we integrate the following: (a) characteristics of the various lockdown strategies (i.e., lockdown metrics); (b) assessing the impact of various lockdown strategies/metrics on preventing transmission rates; and (c) data analytics, visual analytics, modelling, and software/tool development. Finally, with an understanding of the interplay between these areas, we design and develop an interactive, user-friendly software/tool with visual analytics features. This planned tool will provide valuable insight to policymakers for developing well-calibrated balanced responses to ongoing COVID-19 pandemic waves and will be adaptable to future pandemics.

# INTRODUCTION:

As of December 4, 2022, the global pandemic known as Coronavirus Disease 2019 (COVID-19) had killed 6.6 million persons and infected over 652.2 million people [1]. Governments and politicians have made several attempts to lessen harm to public health by reducing the spread of illness ever since the World Health Organization labelled COVID-19 a pandemic. Due to the elevated risk of death and high contagiousness of COVID-19, it was even more important to implement certain deliberate policy reforms to reduce the load on healthcare systems while pharmacological therapies were still being researched. The epidemic has generated a lot of attention and research due to its length and the disruption it caused to the normal functioning of the contemporary world. Conveniently, this heightened interest contributes to the vast quantity of information on COVID-19 that is already accessible, particularly in terms of tracking the variety [2] and efficacy [3] of non-pharmaceutical therapies (both of which are emphasized here).

The overall efficiency of various NPIs measurement methods varies. The closure of schools and colleges, as well as limits on social gatherings, do assist to reduce the danger of transmission; however, the closure of the majority of businesses has minimal effect on transmission and would be detrimental to the economy [4]. Businesses are still feeling the consequences of broad lockdowns [5], but when people rush to previously closed venues (such as stores, schools, and concerts) when restrictions are relaxed, the number of cases begins to rise and new NPIs are required. Consequently, the cycle is usually perpetuated by more closures of already vulnerable businesses. Therefore, even though total lockdowns were initially reliable NPIs, they ultimately became harmful and untenable.

This project aims to build a tool capable of properly predicting the impact of different methods on infection rates. This will be accomplished by the use of historical real-world COVID-19 data from [6] and the correlation between infection rates and non-pharmaceutical therapies (NPIs).

There are several distinct types of NPIs [APPENDIX A:], and each one has a unique effect on preventing the spread of the virus. For instance, the likelihood that an infected individual is present in a crowd that is bigger will grow, which in turn will increase the danger of the disease being passed on. Determining the association between these NPIs and infection rates may thus provide a reliable estimate of the daily variance in the number of infection cases that have been reported. In this study, the historical real-world COVID-19 data [6] will be used to train an ML model to forecast the initial daily new infection case value and apply multiple customizable NPIs inputs into the infection cases to forecast future infection cases graph for the policymakers to assist them in balancing their options while developing a strategy against these pandemics. The graph will be provided to them to assist them in developing a strategy against these pandemics.

The interactive and user-friendly tool that will be created because of this study will provide policymakers with a tool that they can use to model and predict the impacts of various NPI policies on infection cases.

# RESEARCH OBJECTIVES:

Goal: The objective of this research is to conduct additional research and collect the most recently available public data so that a machine learning algorithm can be applied to the data to bring the tool closer to the level of a functional prototype. This will allow policymakers to model and anticipate the effects of different policies on infection rates during future pandemics.

O1: Investigate and gather historical real-world COVID-19 data from public domains, then examine the relationship between NPIs and infection rates.

O2: Apply machine learning to learn and evaluate historical real-world COVID-19 data to predict daily new infection cases and link the various lockdown methods with the projected occurrences.

O3: Most of the backend data was kept in a CSV file in the previous prototype, and it is regularly updated. As the dataset will grow, this may have an impact on how well the tool performs. So, to increase the tool's speed and performance, the code needs to be optimized such that we are fetching and modifying the data from database, rather than from CSV file.

O4: Integrate the prototype's front- and back-end results using React, Python and MySQL.

## Significance:

This prototype tool leverages past data from the real-world COVID-19 [6] to estimate daily new infection cases using a polynomial regression machine learning model with a positivity rate as input to the model. The relationship between the infection rate and non-pharmaceutical interventions (NPIs) is then discovered, allowing us to predict future infection cases in an understandable graphical style. This tool's backend has also been enhanced by integrating a MySQL database with the front end to store and retrieve COVID data, which was previously done via a CSV file and makes the program considerably faster. This software is made to give the government and policymakers a quick and simple approach to comprehending how various NPIs tactics may affect the number of infectious cases. Here, it is crucial that the technology be developed to be extremely scalable in case of future pandemics.

# BACKGROUND AND RELATED WORK:

## Useful Tools Against the Pandemic:

* Governments hoped to eliminate COVID-19 through pharmacological interventions, such as a treatment for those who were sick or a vaccine to prevent infections, but they had to look for preventative NPIs while they vaccines were still being produced.
* The COVID-19 Visit Risk Calculator [7], which computes a customized report after asking a series of questions to assist people understand their risk levels, is a fantastic tool created by researchers at Ryerson University for the general public to use when determining whether to visit others. These technologies leverage historical datasets from the actual world to generate predictions.
* The pandemic brought together a large number of researchers to help slow down the spread and improve preparations for pandemics in the future. From a local level to a global level, there are publicly available technologies that may be used to anticipate events like mortality, illness, and more [8].

## Previous Prototype:

* The goal of this research is to provide the government and policymakers with more information than just an estimate of the number of cases; instead, it will consider numerous key NPIs parameters that have a significant impact on reducing viral spread.
* A foundation for the tool's prototype has already been created by three students working under the direction of Western University professor Anwar Haque [9][10][11]. Publicly accessible data have been used to create a prototype, basic modelling has been finished, and some partial data from the public have been acquired.

## Research on the Efficiency of NPIs: -

* The findings of an analysis of NPI effectiveness in 79 different regions revealed that closure in small gatherings, educational institutions, and borders was primarily most effective in halting COVID-19 spread [3].
* The finding of an analysis of NPIs effectiveness on COVID -19 infection rate in 176 countries reveals that the most significant NPI was the restriction and closure of schools, which had a noticeable impact roughly 10 days after introduction. Mass gathering bans, company laws, and limits were found to have a more gradual impact, while social isolation was linked to a delayed impact that began roughly 18 days after introduction [12].
* According to a different study, COVID-19 dissemination is more effectively contained in small groups than it is in large ones [4]. These studies attest to the fact that this plays a significant part in pandemics, and policymakers ought to be able to effectively control this NPIS to slow the viral spread.
* [Nicolas Banholzer el al.](https://www.medrxiv.org/content/10.1101/2020.04.16.20062141v3.full)[13] estimated the relative reduction in the number of new COVID-19 cases assigned to each NPIs. They conducted an empirical estimation of the relationship between NPIs and variations in documented COVID-19 cases which is approximately 1.6 million through April 15, 2020, across n = 20 countries using statistical analysis approach. They assumed that NPIs could only influence the number of new cases with a delay of t0 = 7 days. NIPs are classified into seven categories: School closings, border closures (restricting international travels), event restrictions, gathering prohibitions, venue closures (such as pubs, restaurants, and stores), lockdowns that forbid public mobility without a good reason, and work bans on non-essential company activity. Each country had a different number and timing of NPI implementations.

## This Research’s Unique Contribution

* This research will add a meaningful contribution to the prototype by finding the latest real-world COVID-19 data and removing the unnecessary columns and after that removing the outliers and variants. After that, a clean dataset will be used to train of the ML model that will predict the number of daily covid-19 cases.
* A highly customizable powerful tool that allows policymakers to easily assess the infection impact of NPIs employed during a pandemic based on historical real-world COVID-19 data [6].
* In the prior prototype, the majority of the backend data was stored in a CSV file and was often updated. The performance of the tool may be impacted as the dataset expands. So, to overcome this drawback we can use MySQL Database to fetch and modify our dataset, instead of fetching data from CSV files which will automatically improve the overall performance of this tool.

## Analysis and Research Gap

Analysis:The effects of NPIs on the propagation of COVID-19 has been examined in the research papers [2] and [3]. NPIs' effectiveness at halting virus spread is demonstrated in both articles. When using the tool, it is crucial to carefully analyze how different NPIs may impact infection instances.

Research Gap:This can be used to precisely anticipate infection instances for the tool using the most recent historical real-world dataset of COVID-19 cases. Additionally, analysis of the relationship between NPIs and the prevalence of COVID-19 infections can be incorporated into the program to mimic the effect on infection cases.

# CONSIDERATIONS AND DISCLAIMERS:

* To create this tool, a historical, real-world COVID-19 dataset from the Canadian government's Public Health Infobase will be used. Due to the pandemic's rarity, the model will not be entirely accurate, and several assumptions will need to be made. The Delta and Omicron variants of COVID-19 were the most widely disseminated among the several COVID-19 iterations. Therefore, as the instances fluctuate in the second half of the pandemic, the accuracy of the final model output will be enhanced by removing these outliers from the dataset.
* The Public Health Infobase of the Canadian government's most recent dataset and the addition of additional constraints during data cleaning have enhanced the accuracy of the ML model; the most recent code with all the changes is accessible on GitHub. Most of the project's work is spent gathering data and retrieving it from the database. Nearly all of this report will be devoted to improvements for the extended portion. This report also emphasizes data collecting, retrieving data from the database, and enhancing the precision of the ML model created by an earlier student. It involves broad implementation details as well as high-level tactics and assumptions. Enter the GitHub repository to see the code in greater detail. The back-end server and the front-end UI are both included. To get the website up and running on your local computers, go to the README file that was created. To request access, email [padya2@uwo.ca](mailto:padya2@uwo.ca).

# METHODOLOGY:

As we know the COVID-19 pandemic had multiple variants and each variant impacts differently on the number of covid cases, from which the most widely spread ones are the Delta and Omicron variants. Therefore, it is very important to remove these outliers in the dataset to improve the accuracy of the final model output. So, for this step, we create a python file code [APPENDIX C:] in which python building libraries pandas, NumPy and matplotlib are used so that we can remove the outliers from the dataset by specifying the positivity rate range and cases limit. After creating the final dataset now, we must import the Polynomial regression ML model from the scikit-learn [14] and use the latest COVID-19 dataset [6] to train the model. After training the model, it can be exported by using a pickle which is a python library used to create a byte stream from a Python object and store it in a file. After that, a “.sav” file is created which will be used on our prototype of this project in predicting the daily infection cases based on the positivity rate. The ML model will be imported into the backend server to be integrated into the application of this project, where it will be used to calculate the starting (day 0) day's daily new infection cases based on the input positive rate and other frontend parameters that can be the changed by the user of the application according to their requirement. Users can view and edit the parameters such as opcapacity, Quantity, Average Population, relative risk etc. [APPENDIX E:]. After calculating the day 0 cases, the prototype will anticipate the future days' cases based on the current metric parameters set in the front end, considering the NPIs' correlation with infection rates that are stored in the database table and show output in a graphical format on our prototype web application [APPENDIX D:].

# RESULT:

## Data gathering

The ML model was created using historical, real-world COVID-19 datasets from the Public Health Infobase of the Canadian government, which is accessible to the public [6]. The original dataset can be viewed [here](https://resources-covid19canada.hub.arcgis.com/datasets/covid19canada::provincial-daily-totals/explore) . The original dataset consists of daily new case information for each province. In our case, we need collective data for the whole of Canada. So, to achieve this result I have used different python built-in libraries to modify the original dataset structure. The dataset about the correlation between NPIs and COVID-19 infection rates in percentage can be also viewed [here](https://docs.google.com/spreadsheets/d/1-s81VqyResLLKfOG1XgftHGWyOgiAdv9djSbl1qxNf0/edit#gid=0), which is collected from different research papers [13][15][16][17][18].

## Removing COVID-19 Variants and Outliers

The original real-world COVID-19 dataset includes all the data collected from the start of the pandemic till 14 April 2022 which means that data is inappropriate for our research as it includes the days when you may see a smaller number of cases with respect to the time when a pandemic is on peak due to a new variant arrival. So, for our research, we are only considering the time when the COVID-19 pandemic is at its peak. To achieve this, we are removing Covid-19 variants and outliers.

### Calculation:

* Estimates of the range of these variants indicate that for this project, variants with a positive range of 0.01 to 0.045 and a case count of more than 600 are to be eliminated.
* Additionally, outliers found beyond the interquartile range were eliminated. The interquartile range was calculated for the project using the 1.5 rule.
* The technique and the code to do this was written in Python and are attached in Appendix C.

## CREATE AN MACHINE LEARNING (ML) MODEL

### Model Structure:

We are using polynomial regression with degree 5 for this project with one input and one output value. The model is trained on a subset of the dataset by the government of Canada [6] that includes Daily new cases and positivity rate as a column. The positivity rate was not an original column in the dataset. This column is created by dividing the number of confirmed cases and the number of test columns. The positivity rate columns will act as the input for the model which will predict the Daily new infection cases as an output. Polynomial regression is a regression based on the relationship between the dependent variable (output) and the independent variable (input) up to the nth polynomial degree [14]. A general formula for polynomial regression (Figure: 1) and polynomial regression model flowchart (Figure :2) is shown below:

Chart

Description automatically generated with medium confidence

Figure : Polynomial regression formula

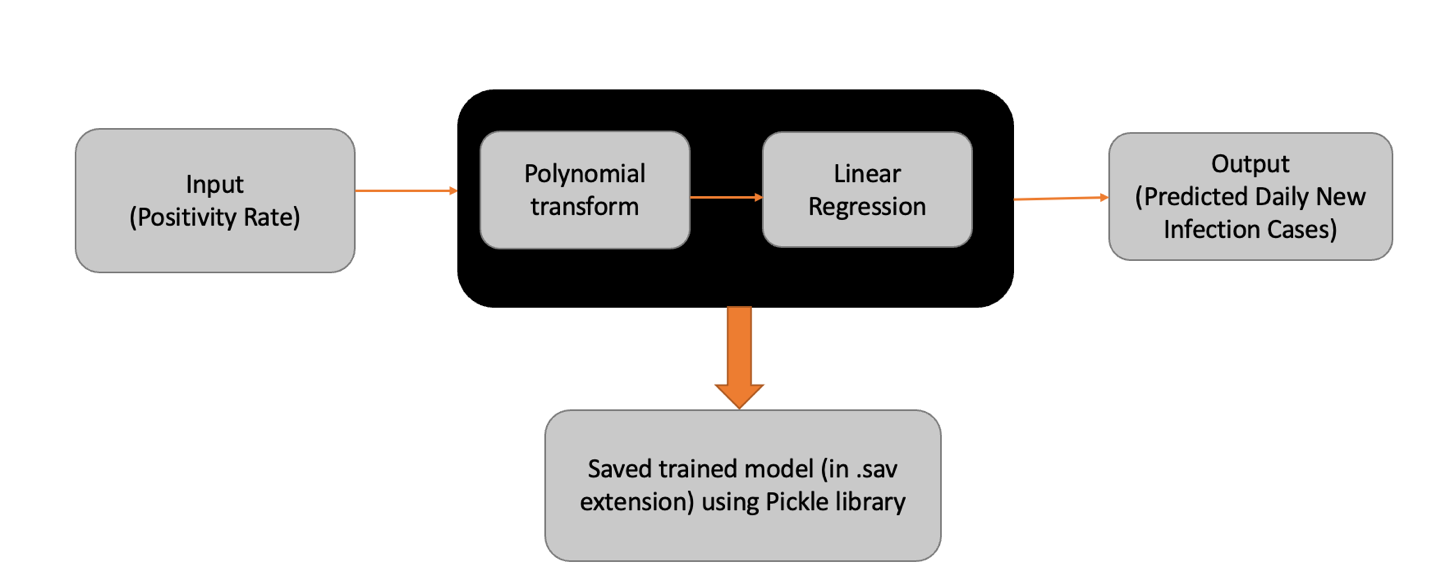


Figure : Polynomial regression model flowchart

### Model Input:

As discussed above that for this research we have added a new column in the existing dataset of COVID-19 [6] which is the positivity rate whose value is equal to the number of confirmed cases/number of tests. The number of confirmed cases and a number of tests are directly used from the original dataset [6].

### Model Output:

The model considers the positive rate as the X value and the number of new infections per day as the Y value. In the model, the X (Positivity rate) value is utilized to train to predict the Y (No of daily new cases) value. For the polynomial regression model, a polynomial degree of 5 is utilized; this number has been tested and offers the best accuracy for the model, which is 90.577% (Figure 3). Before fitting the input into the model, the input was transformed to the polynomial degree [APPENDIX B:].

Chart, scatter chart

Description automatically generated

Figure : This graph shows the result of our ML algorithm whose input value is positivity rates, and it will give us the number of Daily New Cases as an output. The accuracy of this model is 90.577 %.

## CORRELATION BETWEEN VARIOUS NPIS AND COVID-19 INFECTION RISK

As of now, our ML model is trained so our next step is to find various NPIs on which the government had to implement restrictions so that they can control the pandemic and how these NPIs are related to COVID-19 infection risk. As discussed above, matrices data was previously stored in a CSV file which is not an ideal way. So, now data is stored in a more efficient way in MySQL database (Figure: 4). There are a total of 9 columns in this dataset which are id, SubsectionQuantity, AvgPop, OpCapacity, RelRisk, RelChange, SubsectionCount, Business Type. Relchange column is added to represent the correlation between each NPI and infection risks that are collected [13][15][16][17][18]. The number of subsections that fall under a certain NPI category is shown by the SubsectionCount columns. A scale from 0 to 1 is used to express the operating capacity (OpCapacity) of each NPI. Opcapacity 1 implies that no restriction is implemented on that NPI.

For example: The frequency of daily new infection cases will rise by 8%, for instance, when all of Education's subsections OpCapacity values are set to 1.

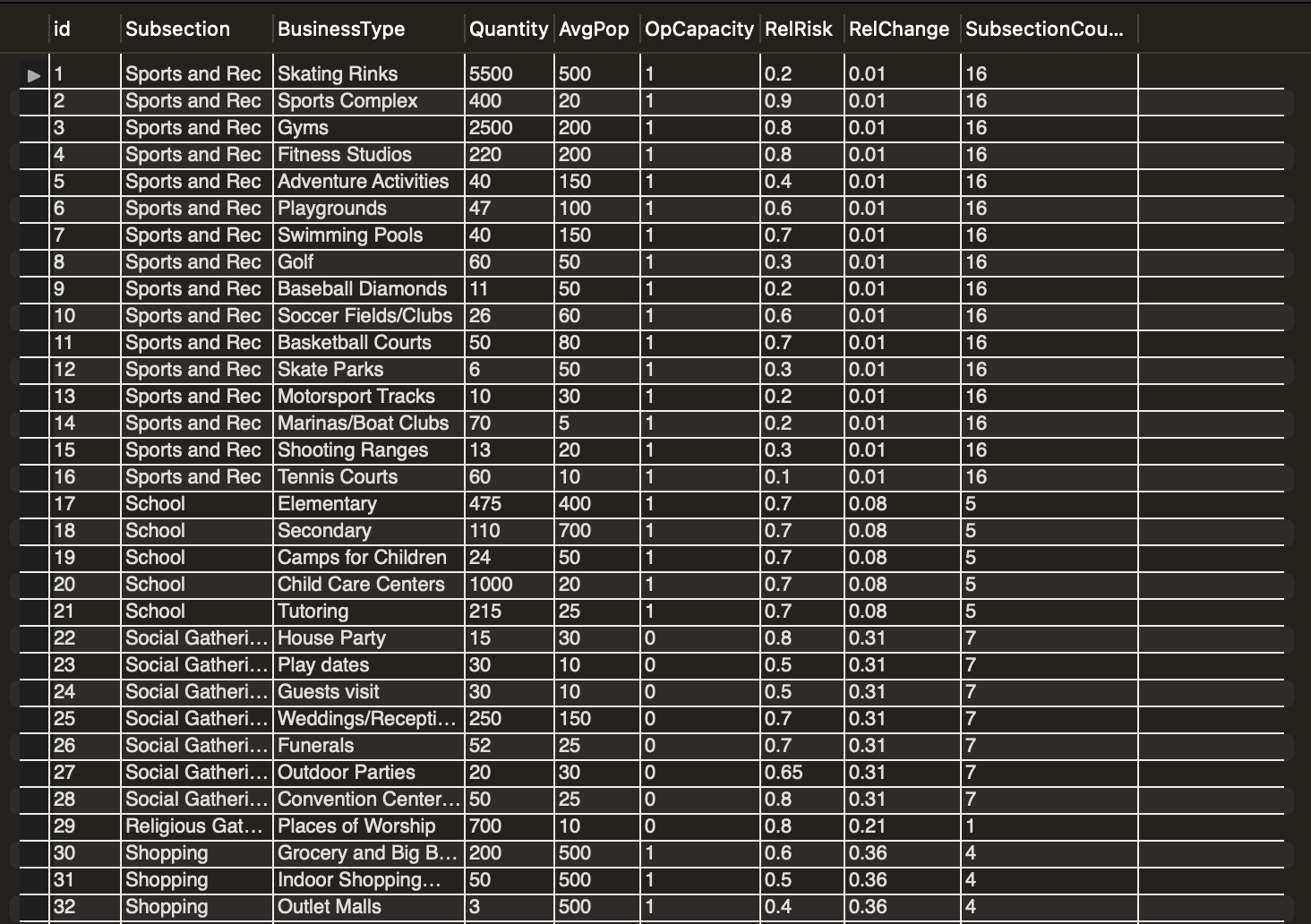


Figure :This screenshot shows the "Sample\_covid\_data" table to stores the covid metrics data in MySQL database.

### Calculation:

Each subpart receives the same amount of each NPI's percentage weight. Tutoring, for instance, falls under the school category and has a RelChange of 0.08 and a SubsectionCount of 5. RelChange divided by SubsectionCount, or 0.016, is the weight that tutoring will have on its impact on daily new infection cases.

### Findings:

The number of cases may be significantly impacted by the data of the relationship between NPIs and COVID-19 infection rates. This functioning prototype can give policymakers a reasonably accurate estimate of the impact of different NPIs and assist them in understanding the crucial balance when making decisions.

### Implementation:

* Most components were created using Google's Material UI React framework. Re-Charts, a free package for React charts, was used to render the line chart showing the cases.
* It's crucial that users can update the four inputs at the top easily because they feature elements that are extremely changing daily. The R-value and positivity rate inputs have something to do with how quickly the pandemic spreads. The number of days entered depends on the level of detail you want to see in the graph. The users' attempt to model the present population of the region is reflected in the population input.
* Each metric's occupational capacity can be changed in the sidebar. The backend server receives a request whenever a user modifies the value to update the database table containing the metrics. The graph will refresh after the user presses "Calculate" to save the adjustments.

# DISCUSSION:

* The daily variation in new infection cases when the Delta and Omnicron variants first appeared had an impact on the dataset's quality and, if disregarded, could have had an impact on the results' quality. Before using the dataset for training, the COVID-19 variation outlier threat was discovered and eliminated.
* This prototype is now based on actual historical data about the COVID-19 pandemic, which is useful for the subject of research. The outcome of this study can be compared to other predictions to assess the precision of various tools.
* This prototype is currently in a usable stage for the area of practice, and it makes use of the relationship between NPIs and actual historical data to forecast the number of cases. By doing so, the decision-makers will have an easy way to balance the effects of the infection and the various lockdown procedures.
* Not simply COVID-19 can be used with this tool. Due to the parameters' flexibility and ability to be easily customized, it can be applied to any pandemic in the future.

# CONCLUSIONS:

The most recent epidemic has made it more challenging for authorities to strike a balance between the public's health and the economy. This study gathered actual historical information regarding COVID-19 cases, identified various NPIs, and examined their relationships with infection rates [O1]. Following the collection of the data, the project's main goal was to train and fit the data into the ML model to forecast daily new infection cases before applying the input NPIs to the predicted cases [O2]. As of right now, the Covid pandemic is still going strong, which means that data on Covid infection cases is growing and is therefore difficult to store in a CSV file. As a result, MySQL will now be used in this project instead of CSV, making it possible for our program to function during health emergencies in the future [O3]. The prototype tool was merged with the NPIs correlation value, and it was set up to function [O4].

As a result, policymakers will have a simple, highly adaptable, and evidence-based instrument at their disposal to help them strike a compromise between NPIs [APPENDIX A] and public health. This application will offer a fresh resource for any impending public health crisis and give policymakers a useful tool for any upcoming unusual conditions. To identify the best answer for the general population, it is crucial to research the effects of various occupancy restrictions on various NPIs, which can have a significant impact on public health.

# FUTURE WORK AND LESSONS LEARNT:

## Future Work

* Currently, this tool is only focussing on the Canada which can be expended by gathering the whole world dataset and then intergrade that dataset to our existing prototype which can eventually forecast the impact the different lockdown measures on the spread of COVID-19 infections over the entire world.

## Lessons Learnt

* A wonderful technique to develop useful applications and tools is to gather and use historical data from the real world to build a model. These data can be used to construct and generate several useful tools such as by using COVID-19 dataset with ML model we can evaluate how differently COVID-19 is impacting on different age groups.
* My research on NPIs and COVID-19 was helpful for this project as well as for understanding the various limitation restrictions for various nations. Outside the purview of the research, you can pick up knowledge and facts that can help you in future projects.

# ACKNOWLEDGEMENTS:

I want to express my gratitude to Dr. Anwar Haque for giving me the chance to work on this important study and for his direction throughout the research process. Additionally, I want to thank Letitia Kwan, Mauli Bhatt, and Dan Lou for outlining their contributions to the project's basis for me. I am appreciative of the opportunity to study in the computer science department at Western University because I have gained information and insight that will help me solve challenges in the real world.

# REFERENCES:

1. By Henrik Pettersson Byron Manley and Sergio Hernandez, CNN Tracking Covid-19’s global spread [Online]. [[CrossRef]](https://www.cnn.com/interactive/2020/health/coronavirus-maps-and-cases/)
2. Zheng, Q., Jones, F.K., Leavitt, S.V. *et al.* HIT-COVID, a global database tracking public health interventions to COVID-19. *Sci Data* 7, 286 (2020). [[CrossRef]](https://doi.org/10.1038/s41597-020-00610-2)
3. Haug, N., Geyrhofer, L., Londei, A. *et al.* Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat Hum Behav* 4, 1303–1312 (2020).[[CrossRef]](https://doi.org/10.1038/s41562-020-01009-0)
4. Jan M. Braune, Sören Mindermann, Mrinank Sharma, David Johnston, John Salvatie, Tomáš Gavenčiak, Anna B. Stephenson, Gavin Leech, George Altman et al., Inferring the effectiveness of government interventions against COVID-19, Journal Article, 2021, Science, eabd9338,371,6531. [[CrossRef]](https://www.science.org/doi/abs/10.1126/science.abd9338)
5. *Economic impacts and recovery related to the pandemic*. (n.d.). Retrieved from Statistic Canada. [[CrossRef]](https://www150.statcan.gc.ca/n1/pub/11-631-x/2020004/s5-eng.htm)
6. Daily provincial totals (n.d.) Retrieved from Statistic Canada. [[CrossRef]](https://resources-covid19canada.hub.arcgis.com/datasets/covid19canada::provincial-daily-totals/explore)
7. Ryerson University. My COVID-19 VISIT RISK. National Institute onAging. Accessed on: November 16, 2021. [[CrossRef]](https://covidvisitrisk.com/)
8. IHME. Covid-19 projections. Institute for Health Metrics and Evaluation. (n.d.), 2020. Accessed on: November 16, 2021. [[CrossRef]](https://covid19.healthdata.org/global?view=vaccinations&tab=trend)
9. Kwan, Letitia. Quantifying the Effectiveness of Pandemic Lockdown Measures: Building a User Interface and Economic Model [Unpublished bachelor’s thesis]. The University of Western Ontario, 2021.
10. Bhatt, Mauli. Modelling the Long-Term Health Impact of COVID-19 Lockdowns [Unpublished bachelor’s thesis]. The University of Western Ontario, 2021.
11. Lou, Dan. Modeling the Impact of Lockdown Measures on Covid-19 Infection Spread [Unpublished thesis]. The University of Western Ontario, 2022.
12. Nader, I.W., Zeilinger, E.L., Jomar, D. *et al.* Onset of effects of non-pharmaceutical interventions on COVID-19 infection rates in 176 countries. *BMC Public Health* 21, 1472 (2021).[[CrossRef]](https://doi.org/10.1186/s12889-021-11530-0)
13. Nicolas Banholzer, Eva van Weenen, Bernhard Kratzwald, Arne Seeliger, Daniel Tschernutter, Pierluigi Bottrighi, Alberto Cenedese, Joan Puig Salles, Werner Vach, Stefan Feuerriegel. [[CrossRef]](https://www.medrxiv.org/content/10.1101/2020.04.16.20062141v3.full)
14. *Linear Models*. (n.d.). Retrieved from Scikit-Learn: [[CrossRef]](https://scikit-learn.org/stable/modules/linear_model.html)
15. Whaley CM, Cantor J, Pera M, Jena AB. Assessing the Association Between Social Gatherings and COVID-19 Risk Using Birthdays. JAMA Intern Med. 2021;181(8):1090–1099. [[CrossRef]](https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2781306)
16. Hongoh V, Maybury D, Levesque J, Fazil A, Otten A, Turgeon P, Waddell L, Ogden NH. Decision analysis support for evaluating transmission risk of COVID-19 in places where people gather. Can Commun Dis Rep 2021;47(11):446–60. [[CrossRef]](https://doi.org/10.14745/ccdr.v47i11a02)
17. Alina S. Schnake-Mahl, Gabriella O’Leary, Pricila H. Mullachery, Vaishnavi Vaidya, Gabrielle Connor, Heather Rollins, Jennifer Kolker, Ana V. Diez Roux, Usama Bilal [[CrossRef]](https://www.medrxiv.org/content/10.1101/2021.04.12.21251656v1.full)
18. Helsingen, L.M., Løberg, M., Refsum, E. *et al.* Covid-19 transmission in fitness centers in Norway - a randomized trial. *BMC Public Health* 21, 2103 (2021). [[CrossRef]](https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-12073-0)
19. Airports Council International, “Assessing Airport Passenger Traffic and COVID-19 Test Positivity Rates in Europe Since July” in November 2020. [Online serial]. [[CrossRef]](https://www.aci-europe.org/downloads/resources/Airport%20traffic%20and%20COVID%20rates_ACI%20EUROPE_11.11.2020.pdf)
20. Chang, S., Pierson, E., Koh, P.W. *et al.* Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* 589, 82–87 (2021). [[CrossRef]](https://doi.org/10.1038/s41586-020-2923-3)
21. Fields, S. (2020 , November 30). Risk of getting COVID at the grocery store is higher in low-income neighborhoods. Retrieved from MarketPlace: [[CrossRef]](https://www.marketplace.org/2020/11/30/risk-of-getting-covid-at-the-grocery-store-is-higher-in-low-income-neighborhoods/)
22. Bajos, N., Jusot, F., Pailhé, A. *et al.* When lockdown policies amplify social inequalities in COVID-19 infections: evidence from a cross-sectional population-based survey in France. *BMC Public Health* 21, 705 (2021). [[CrossRef]](https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-10521-5)
23. Herby, Jonas & Jonung, Lars & Hanke, Steve, 2022. "A Literature Review and Meta-Analysis of the Effects of Lockdowns on COVID-19 Mortality," The Johns Hopkins Institute for Applied Economics, Global Health, and the Study of Business Enterprise. [[CrossRef](https://ideas.repec.org/p/ris/jhisae/0200.html)]
24. Violato C, Violato EM, Violato EM (2021) Impact of the stringency of lockdown measures on covid-19: A theoretical model of a pandemic. PLoS ONE 16(10): e0258205. [[CrossRef]](https://doi.org/10.1371/journal.pone.0258205)
25. Talic S, Shah S, Wild H, Gasevic D, Maharaj A, Ademi Z et al. Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis BMJ 2021; 375 :e068302 [[CrossRef]](https://www.bmj.com/content/375/bmj-2021-068302)
26. Staff Reuters Fact check: Studies show COVID-19 lockdowns have saved lives [Online]. - NOVEMBER 24, 2020. [[CrossRef]](.%20%20https:/www.reuters.com/article/uk-factcheck-lockdowns-idUSKBN2842WS.)
27. Banholzer N, van Weenen E, Lison A, Cenedese A, Seeliger A, Kratzwald B, et al. (2021) Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. PLoS ONE 16(6): e0252827. [[CrossRef]](https://doi.org/10.1371/journal.pone.0252827)
28. Alba Mendez-Brito, Charbel El Bcheraoui, Francisco Pozo-Martin,Systematic review of empirical studies comparing the effectiveness of non-pharmaceutical interventions against COVID-19,Journal of Infection,Volume 83, Issue 3,2021,Pages 281-293,ISSN 0163-4453. [[CrossRef]](http://sciencedirect.com/science/article/pii/S0163445321003169)

# APPENDIX A:

A picture containing calendar

Description automatically generated

Figure 5:This diagram demonstrates how enterprises and industries were broken down into smaller components. Additionally, practically every company that was for occupancy adjustments was included in this collection.

# APPENDIX B:

**Text

Description automatically generated**

**Text

Description automatically generated**

Figure 6: This coding figure shows the techniques to train the ML model with a Polynomial degree of 5

# APPENDIX C:

**Text

Description automatically generated**

Text

Description automatically generated

Figure 7: This coding figure shows the technique for removing the COVID-19 variants and outliers from the dataset.

# APPENDIX D:

Graphical user interface, chart

Description automatically generated

Figure : Main Page of our Prototype web application

# APPENDIX E:

**Graphical user interface, table

Description automatically generated**

Figure : View and Edit Metrics page where user can change the metrics data