STA2453 Project 2

Realtime COVID-19 Vaccination Status Dashboard

Emily Zhu 1003161393 Siyuan Shang 1005558380

Access the dashboard at: https://emzy.shinyapps.io/STA2453-Project2-main/

1. Introduction

The goal of the Realtime COVID-19 Vaccination Status Dashboard is to allow the general public to monitor the vaccination status of the nation. The dashboard gathers COVID-19 vaccination status information published by the Government of Canada's Public Health Infobase and displays the data in simple, comprehensive graphs, which are described in more detail below. In particular, the dashboard was designed to be easy for users to compare the vaccination progress between different provinces. The "Realtime" in the dashboard's name refers to its ability to update at the same rate as the Public Health Infobase.

2. Data and Methods

2.1 Data Sources

The database uses the following three data sources:

1. https://health-infobase.canada.ca/src/data/covidLive/covid19-download.csv
This data stream contains a range of standard COVID-19 data points, including test counts, infection counts, death counts, and recovery counts by province by day. The data is updated daily. This data source was chosen so that the dashboard provides some basic "colour" on the COVID-19 pandemic.

2. https://health-infobase.canada.ca/src/data/covidLive/vaccination-coverage-map.csv
This data stream contains all vaccination status data, updated on a weekly basis. This data source was chosen for its detailed breakdown by province and age group and for its summary in percentages.

3. https://health-infobase.canada.ca/src/data/covidLive/covid19-epiSummary-casesAfterVaccination.csv
This data stream contains to-date data on total cases, hospitalization counts, and death counts by vaccination status. The data is updated monthly. This data source was chosen for its simplicity as a summary of the effectiveness of the vaccine.

2.2 Data Pipeline

The data is pulled directly from its original url source at the Government of Canada's Public Health Infobase, so that each time the server's script runs, the data is re-pulled along with any updates. This method may not be efficient for larger datasets, but the COVID-19 vaccine status datasets used are not too large.

The data's quality has been checked and is generally sound. Some minor edits are made to the data right after it is pulled, which entail column name changes and value type changes (e.g. ">=99" -> "99"). The data is then funneled into chart rendering functions.

3. Dashboard Framework and Visuals

3.1 Dashboard Framework

The framework used for the dashboard is R shiny. R was chosen for the development language due to its versatility in working with datasets and generating graphical models. The `shiny` and `shinydashboard` packages were chosen for their ease of building interactive dashboards and are the industry standard packages for R web application development.

The dashboard is hosted on a website through shinyapps.io, the most straightforward free option for the developers at the time of initial development. This means that the website is not suited for large flows of web traffic, but should be sufficient for current usage rates and is scalable with paid plans if needed in the future. The website times out after a few minutes of non-usage, forcing the user to refresh the dashboard for data updates.

3.2 Dashboard Visuals and Usage

There are four main visuals on the dashboard:

1. Provincial Summary Map

The Provincial Summary Map allows the user to choose the COVID-19 or vaccination statistic he/she would like to see, and displays a heat map of the statistic by province over a map of Canada. The user can mouse over each province for the specific number to the statistic. Available statistics include *Total Tests*, *Total Cases*, *Total Deaths*, *Total Partially Vaccinated*, *Total Fully Vaccinated*, and *Total Third-Dose Vaccinated*.

This graph design works to tie the statistics with the visual representation of their respective provinces and also functions to aid users in their province selections for the next two visuals. For future development, the map intends to allow users to click a province and display a wider variety of recent 7-day running average statistics.

2. Provincial Comparison Time Series Chart

The Provincial Comparison Time-Series Chart allows the user to choose a vaccination statistic and compare time series charts of the statistic for different provinces. The user can choose to view *At Least 1-Dose Vaccinated* population count or percentage, *Partially Vaccinated* population count or percentage, *Fully Vaccinated* population count or percentage, and

Third-Dose Vaccinated population count or percentage. The user can then choose to show their selected statistic for all of Canada and/or any provinces.

A time-series line graph was chosen here as it provides a clean visual when the user chooses to display and compare multiple provinces' data.

3. Vaccination Percentages Time Series Chart

The Vaccination Percentages Time Series Chart allows users to choose a province and age group and displays the percentages over time of the chosen population who have *At Least 1 Dose*, are *Partially Vaccinated*, are *Fully Vaccinated*, and are *Third-Dose Vaccinated*.

The purpose of this chart is to show when vaccination waves occurred for each age group in each province. A secondary purpose this chart achieves is visualizing the relationship between the four vaccination statuses and showing users that *At Least 1 Dose* = *Partially Vaccinated* + *Fully Vaccinated*.

4. Pie Charts by Vaccination Status

The last section of the dashboard hosts three pie charts that users can mouse over. The pie charts display the percentage of cases, hospitalizations, and deaths by vaccination status. When the user mouses over a pie chart section, the section will display the precise percentage number.

These graphs were chosen to give a straightforward summary of the vaccines' effectiveness and show that the majority of cases, hospitalizations, and deaths are from unvaccinated patients.

4. Conclusion

The Realtime COVID-19 Vaccination Status Dashboard was designed to be simple and straightforward for the general public to use. However, there were lessons learned along the way and improvements that could be made.

Performing analysis on the data and creating informal graphs prior to making the official visuals is an important step that was not realized promptly in this development cycle. This caused some redundancy in the dashboard's data representation, such as the heat map of Canada always having the darkest colour on Ontario since the statistics are by population count instead of percentage, and Ontario always has the largest population no matter which statistic. The other lesson is thinking about what value a piece of information or a visual will bring to the user. In this case, the value-added by the third visual may not be as significant considering the data already provided by the second visual.

Given more time, the problems outlined above would be fixed, and more functionalities for the user would be added, including a time selection option to allow the user to view just the data for their time period of choice. Other additions include more granular age buckets and another visual that allows users to observe correlations between vaccination status and case outcomes with greater interactivity and user choice.