Title: MM802 - Visualization Mini-project

Date: March 20, 2022

Name: Anjali Bajaj , Sabrina Nasrin

Project Title: Data Visualization Project on Covid-19 Canada

1 Abstract

Coronavirus disease 2019 (COVID-19) is a highly contagious viral illness caused by severe acute respiratory syndrome SARS-CoV-2. It has had a devastating effect on the world's demographics resulting in more than 5.3 million deaths worldwide. It has emerged as the most consequential global health crisis since the era of the influenza pandemic of 1918. Dashboards have been made to keep the count of the cases which are active and recovered cases as well as mortality rates. In this project, a dashboard has been built that will demonstrate the cases in all the provinces of Canada from Jan 2020 to March 2022. This dashboard provides summary of COVID-19 cases across Canada and over time. It Contains detailed data about the spread of the virus over time and in different regions of the country. It also includes breakdowns by age and sex or gender and provides an overview of deaths, testing and recoveries.

2 Introduction

We have developed a multi-page web application to visualize Covid-19 reports of different provinces of Canada. The pages of this application include: **Homepage**, **Cases**, **Mortality**, **Testing**, **Age wise Data**.

This Project takes the view of the following questions:

- Do you want to see the daily reports of Canada?
- Do you want to see the daily/cumulative report of all the provinces and provinces of your choice?
- Do you want to see the the percentage cumulative report of all the provinces?
- Which timeline do you want to choose?
- Do you want to see reports for the entire timeline present in the database?
- Do you want to see the breakdowns by age and sex or gender?

This web application will answer all the questions mentioned above and serve to analyze, visualize, the spread of novel Coronavirus - 2019 (COVID - 19) in Canada. This will also show which province has suffered the most in Canada and show statistics and reports regarding Covid-19.
Datasets that have been used in the project is taken from Canada Covid-19 datasets. This dataset

has been completed from COVID-19 Canada Open Data Working Group. It collects daily time series data on COVID-19 cases, deaths, recoveries, testing, and vaccinations at the health region and province levels. Data are collected from publicly available sources such as government datasets and news releases. The dataset is available in JSON format from API, CSV Format from Covid 19 Canada, and Google Drive. In our project, we have used CSV format and data from 25 January,2020 to 17 March,2022.
There are several Dashboards available for the coronavirus disease which demonstrates confirmed cases, recovered cases, Death rates, and vaccines taken by the entire world. We have followed Canada official website COVID-19 daily epidemiology update and took inspiration from that website.

3 Project Status

Our goal was to make an interactive web application to visualize the statistics and reports of Covid-19 where users can check the reports by giving some inputs. We have successfully reached our goal. Our website is working as we thought.

Anjali Bajaj	**Sabrina Nasrin**
Set up virtual environment for our code implementation	Set up virtual environment for our code implementation
Data Collection	Design and Development of the website
Unit testing and integration	Unit testing and integration
Creation of the final documentation	Creation of the final documentation

4 Development Environment

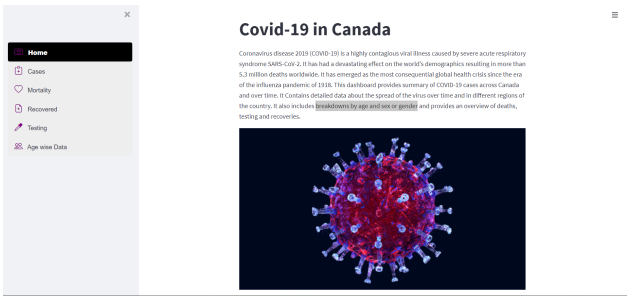
1. **Libraries Used**:
   Streamlit and Plotly Express: Streamlit is an open-source python library for creating and sharing web apps for data science and machine learning projects. Streamlit can seamlessly integrate with other popular python libraries used in Data science such as NumPy, Pandas, Matplotlib, Scikit-learn and many more.
   The Plotly Express Python library is an interactive, open-source plotting library that supports over 40 unique chart types covering a wide range of statistical, financial, geographic, scientific, and 3-dimensional use-cases. The plotly.express module (usually imported as px) contains functions that can create entire figures at once, and is referred to as Plotly Express or PX. Plotly Express is a built-in part of the plotly library, and is the recommended starting point for creating most common figures.

2. **Working with Python**: Codes have been written in python which is compact, easier to debug, and allows for ease of extensibility.

3. **IDE**: Visual Studio Code combines the simplicity of a source code editor with powerful developer tooling, like IntelliSense code completion and debugging.

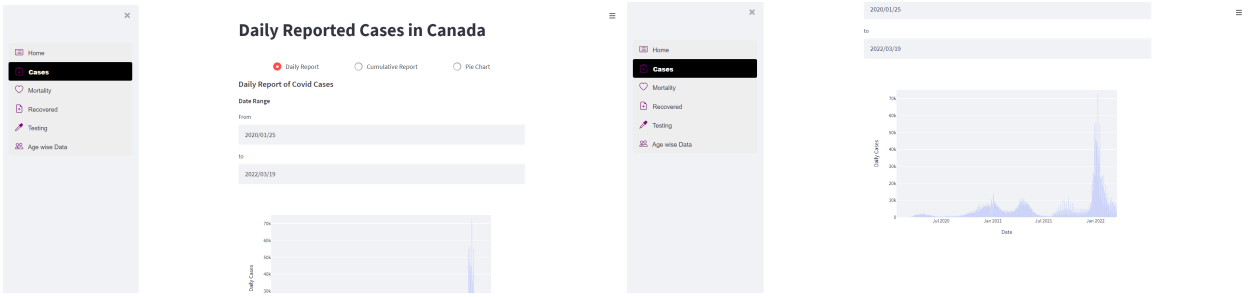
5 Development Work

In our project, we have basically used python streamlit and plotly express libraries. This is the github link of our project. We have described how to use our project in the README section of that repository. We have added some of the output images of our website in this report. The limitation of our website is the pie charts are only showing the cumulative results. They are not made customized with user input and this website is not responsive one.

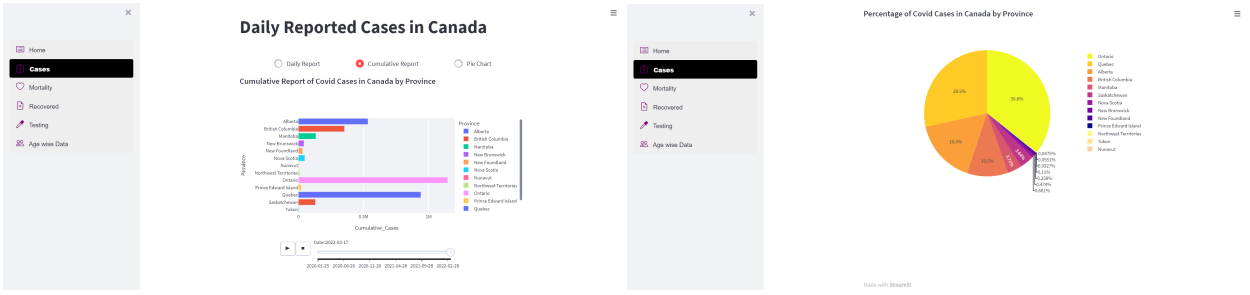


(a)

Figure 1: Homepage

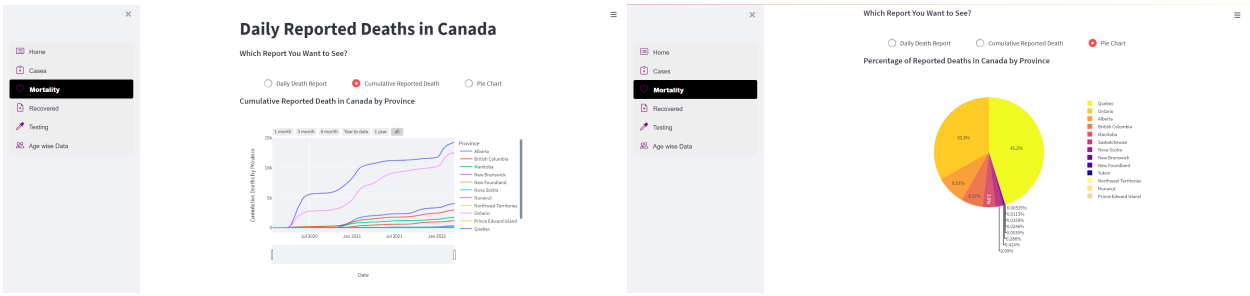


(a) (b)



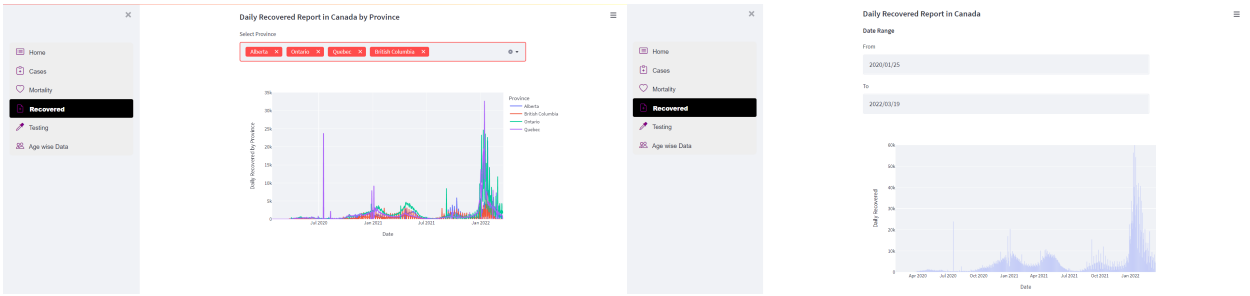
(c) (d)

Figure 2: Cases



(a) (b)

Figure 3: Mortality



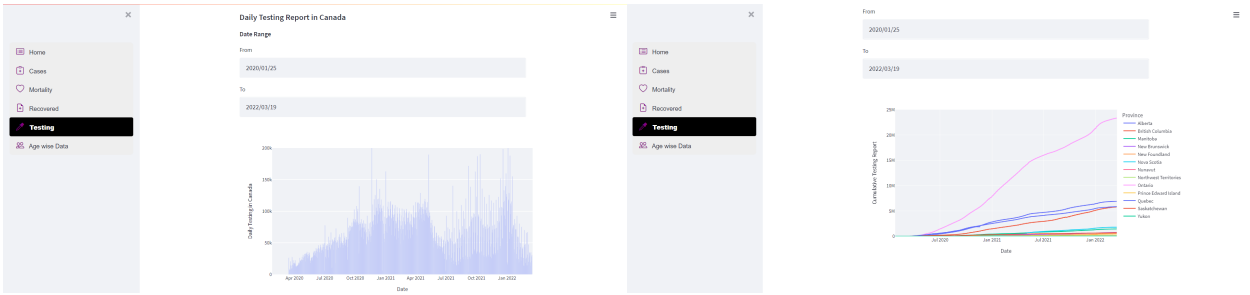
(a) (b)

Figure 4: Recovered

6 Concluding Remarks

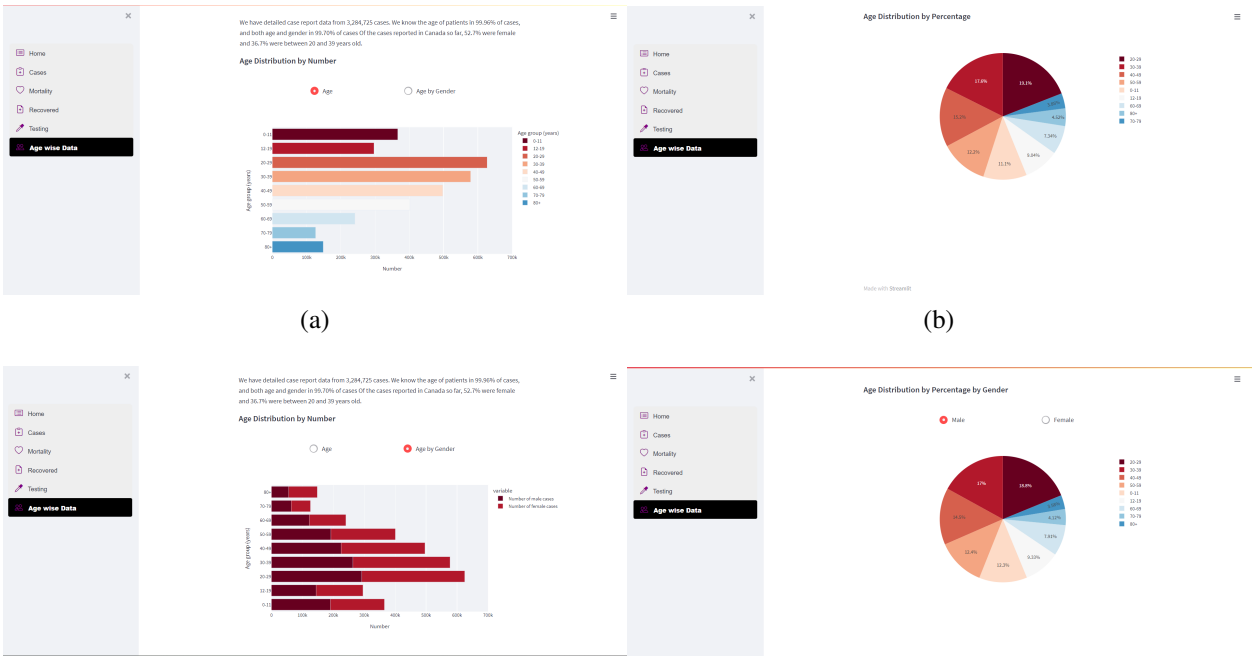
Data Visualization allows us to organize data in a way that's both compelling and easy to digest. It's about representing data in a visual context, such as a chart or a map, to help anyone viewing it better understand the significance of that data.

Our goal was to create an interactive online application that visualized Covid-19 statistics and reports and allowed users to check the reports by providing inputs. Our objective has been accomplished. Our website is functioning as expected. Lessons learned from the project: With the few



(a) (b)

Figure 5: Testing



(a) (b)

(c) (d)

Figure 6: Age wise Data

lines of code, we have created a web application that can help everyone to determine the worst-hit province by Covid and least-hit province. We have learnt about the data visualization libraries and how to work with them.

Possible future extensions to the project work: Future work will include demonstrating the vaccine administrated concerning genders in all the provinces of Canada in the form of different representations, making the website responsive and more interactive to the users.