

Github link: <https://github.com/izzyjohn/Final-Project.git>

1. The goals for your project (10 points)

- a. The goal of our project was to take three APIs each with covid case data for one country—Canada, the UK, and the US—and compare the data across countries. We wanted to see the differences in hospitalization and ICU rates for the APIs that provided this—Canada and the US. We also wanted to compare the average number of total cases for all three countries. Another goal was to see if we could compare when spikes in new cases occurred and if these were at similar or different times for the different countries.

2. The goals that were achieved (10 points)

- a. We were able to calculate the difference in averages between hospitalization and ICU rates for Canada and the US. We also created visualizations that show the average number of new covid cases per day in each country, the current number of Covid cases in Canada and the US, and the number of deaths in categories that we created for the UK. These visualizations, especially the ones comparing two or more countries in one graph, allow the data to be viewed in a comprehensive way. More observations can be made from these than would be possible with just a large database or table.

3. The problems that you faced (10 points)

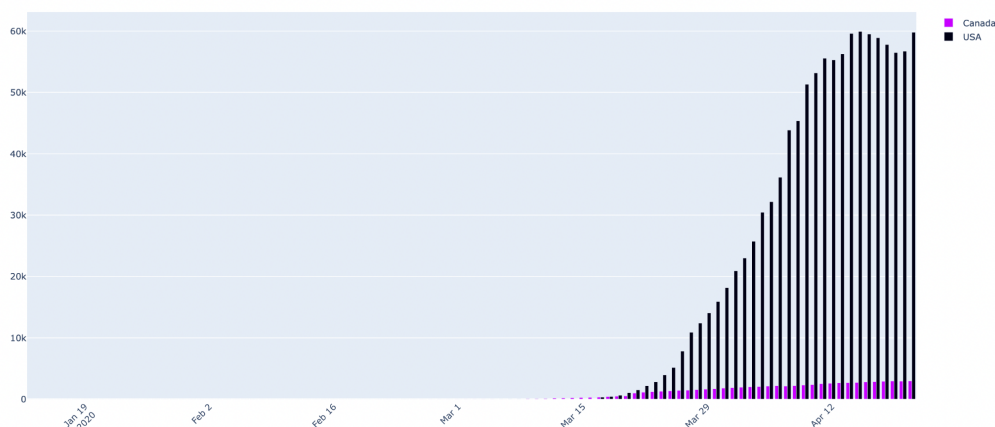
- a. The first problem we encountered was that our original idea proved to be slightly complicated for this project. We were able to come up with a new idea that we could accomplish with free APIs based on our earlier idea. One problem that we faced was with the dates not lining up across the APIs. We were able to fix this to an extent by changing them all to ascending order, starting in 2020, as two of them started in descending order. Another issue was that one API updates their information, so we could not work on that part of our project during certain times.

4. Your file that contains the calculations from the data in the database (10 points)

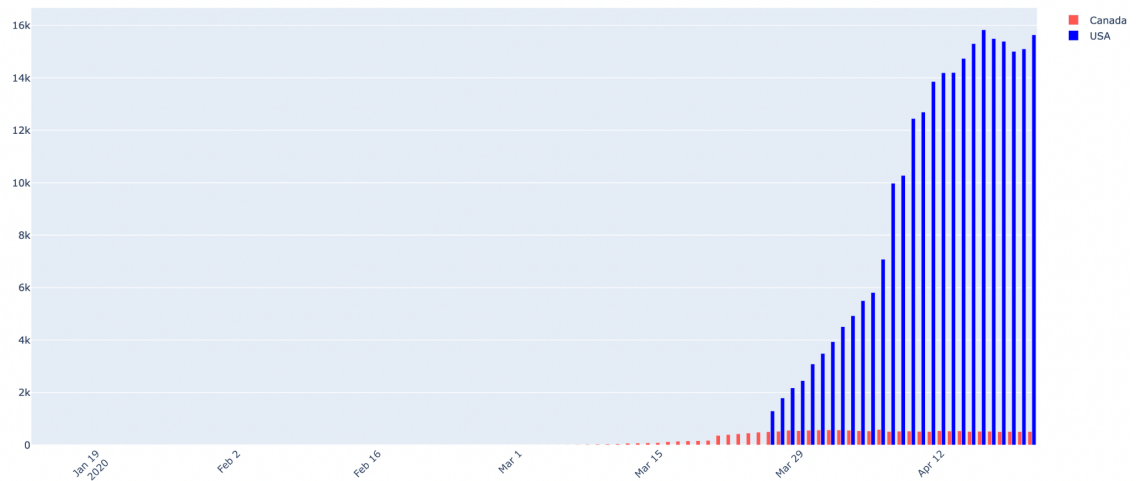
```
Difference between Average Hospitalizations for USA and Canada over 100 days: 12017.364
Difference between Average Number of Patients in the ICU for USA and Canada over 100 days: 2779.614
Average Number of new Covid Cases in the UK over 100 days: 2152.66
Average Number of new Covid Cases in the USA over 100 days: 8089.17
Average Number of new Covid Cases in Canada over 100 days: 594.96
```

5. The visualization that you created (i.e. screen shot or image file) (10 points)

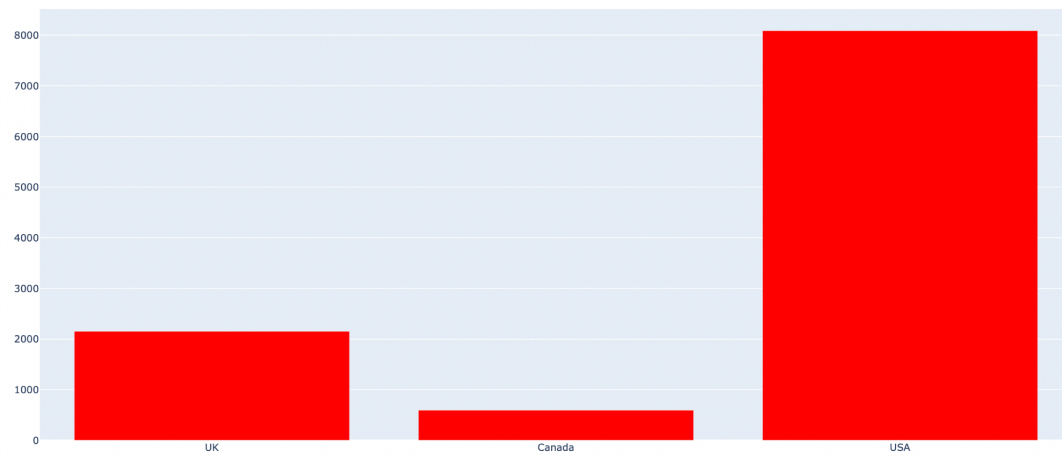
The Number of Hospitalizations Between the US and Canada



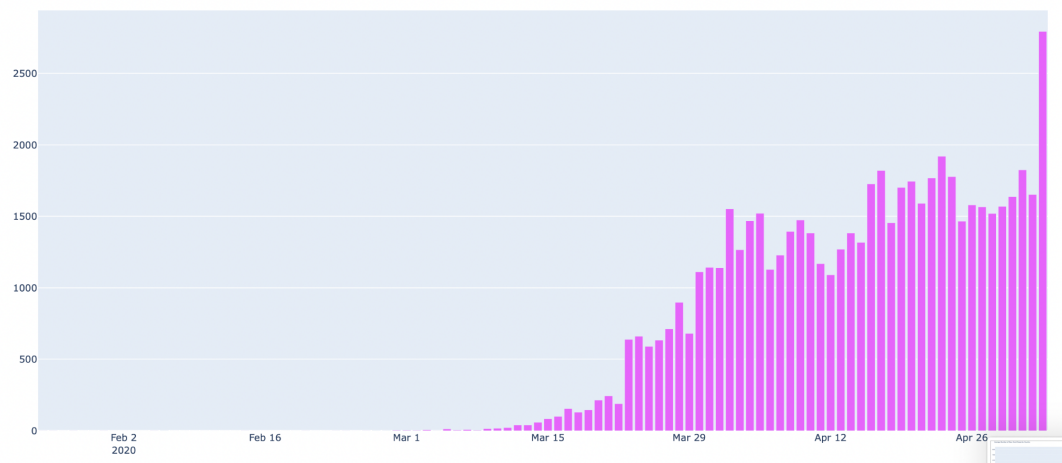
The Number of ICU Patients Between the US and Canada



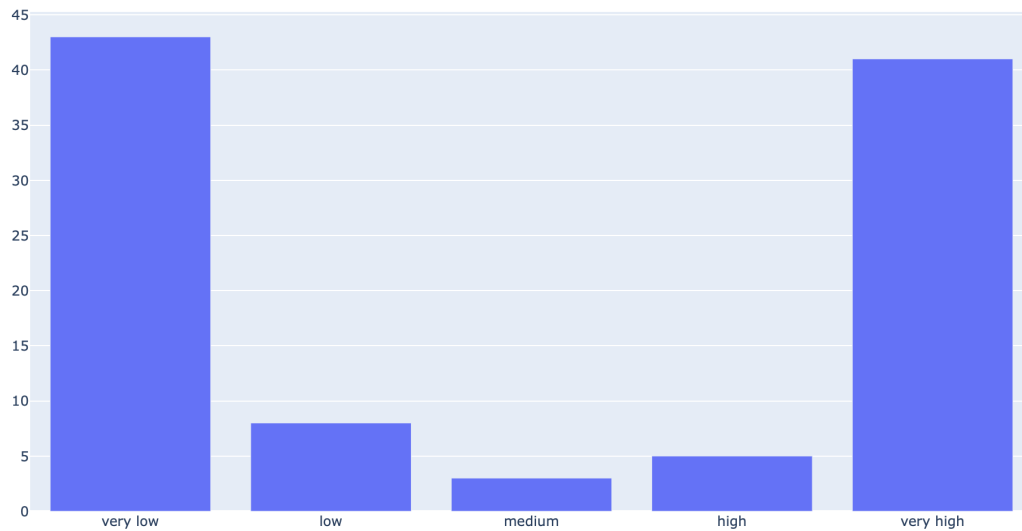
Average Number of New Covid Cases by Country



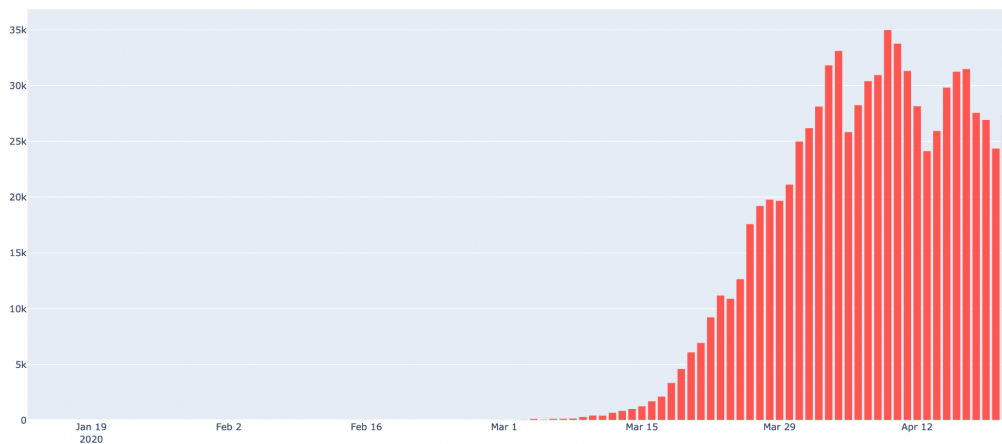
Current Number of Covid Cases in Canada



Number of Days in Each Death Category in UK



Current Number of Covid Cases in USA



6. Instructions for running your code (10 points)

- Running the code will cause six visualizations to open up. After looking at these, you can open the SQL database and see the four tables. You can also open up the text file with the calculations. We built a main function which is called at the end, and in the main function we called each of our other functions.

7. Documentation for each function that you wrote. This includes the input and output for each function (20 points)

- `opendatabase(db_name):`
Takes a filename and opens a database with the name filename.db. It returns conn and cur objects.
 - Input: a string that will be the title of your database
 - Output: curr, conn
- `read_api(url):`
Returns dictionary of information found from api url

- i. Input: the url of an API (type: string)
 - ii. Output: a dictionary of the api information
- c. `uk_data(cur, conn)`:
Reads in information from uk data dictionary and adds information to a table named UK in the database
 - i. Input: `cur, conn`
 - ii. Output: None
- d. `uk_category_table(cur, conn)`:
Creates a table named `death_category` which stores id key that correlates with qualitative information about new deaths daily in the UK
 - i. Input: `cur, conn`
 - ii. Output: None
- e. `canada_data(cur, conn)`:
Reads in information from canada data dictionary and adds information to a table named Canada in the database
 - i. Input: `cur, conn`
 - ii. Output: None
- f. `us_data(cur, conn)`:
Reads in information from US data dictionary and adds information to a table named usa in the database
 - i. Input: `cur, conn`
 - ii. Output: None
- g. `dif_Us_Canada_Average_Icu(cur, conn)`:
Calculates the average number of people in ICU per day in the US and in Canada and then finds the difference between US and Canada's average
 - i. Input: `cur, conn`
 - ii. Output: average difference in the ICU numbers for Canada and the US, type float
- h. `dif_Us_Canada_Average_Hospital(cur, conn)`:
Calculates the average number of people in hospital per day in the US and in Canada and then finds the difference between US and Canada's average
 - i. Input: `cur, conn`
 - ii. Output: average difference in the hospitalization numbers for Canada and the US, type float
- i. `uk_new_cases_average(cur, conn)`:
Calculates the average number of new cases per day in the UK
 - i. Input: `cur, conn`
 - ii. Output: average of new cases in UK, type float
- j. `us_new_cases_average(cur, conn)`:
Calculates the average number of new cases per day in the US
 - i. Input: `cur, conn`
 - ii. Output: average of new cases in US, type float
- k. `canada_new_cases_average(cur, conn)`:
Calculates the average number of new cases per day in Canada
 - i. Input: `cur, conn`

- ii. Output: average of new cases in Canada, type float
- l. write_textfile(filename, cur, conn):

Writes out information for calculation functions to the file with name of passed filename

 - i. Input: filename (string ending in .txt), cur, conn
 - ii. Output: None
- m. visualization_1(cur, conn):

Creates graph based on database that displays the number of current people in hospitalization for Covid-19 in both the US and Canada as time progresses

 - i. Input: cur, conn
 - ii. Output: None
- n. visualization_2(cur, conn):

Creates graph based on database that displays the number of current people in the ICU for Covid-19 in both the US and Canada as time progresses

 - i. Input: cur, conn
 - ii. Output: None
- o. visualization_3(cur, conn):

Creates a bar graph showing the average number of new new cases per day in Canada, the US, and the UK

 - i. Input: cur, conn
 - ii. Output: None
- p. visualization_4(cur, conn):

Creates graph showing the number of current cases of Covid 19 in Canada by day

 - i. Input: cur, conn
 - ii. Output: None
- q. visualization_5(cur, conn):

Creates bar graph shows the amount of days in each number of new deaths category in the UK

 - i. Input: cur, conn
 - ii. Output: None
- r. visualization_6(cur, conn):

Creates graph showing the number of current cases of Covid 19 in the US by day

 - i. Input: cur, conn
 - ii. Output: None

8. You must also clearly document all resources you used. The documentation should be of the following form (20 points)

Date	Issue Description	Location of Resource	Result (did it solve the issue?)
12-04-22	We were struggling to add only 25 entries at a time	Discussion Office Hours	Yes, we figured out how to do this for all our databases
12-07-22	Started receiving an error status 400 and said our data was no longer available	https://coronavirus.data.gov.uk/details/developers-guide	Yes we found out that the data gets updated during this time and that it would be available later that day

		The api documentation	
12-9-22	Creating Plotly Visualizations	Plotly-v8 Presentation Slides	Yes, we were able to create many different visualizations
12-10-22	We were having trouble changing the color of the bars in our Plotly bar graph	Plotly.graph_objects Plotly Bar Charts	Yes, we were able to add colors