# SI 206 Project Report

Yiyang Chen, Zhixuan Yang, Jinghan Zhang

04/26/2020

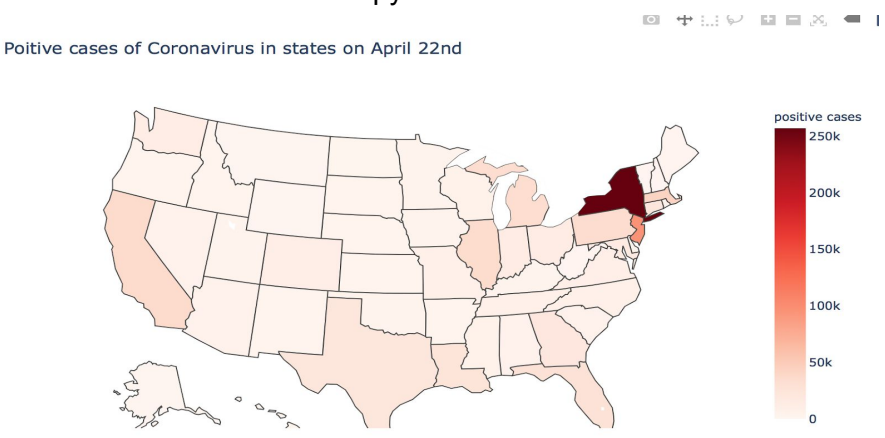
1.Goal of Project: Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus, which has affected millions of people in the world. Through this project, we want to know how the US and other countries are affected by COVID-19 by analyzing the total cases, death rate and other useful data.

2.The goals that were achieved: We successfully gathered and analyzed coronavirus's total cases, death and recovered data for the US and other major countries affected by coronavirus. Our analysis and visualizations provide a brief introduction to the presently terrible situation of the virus in several countries, which will hopefully invoke attention of residents and government on this deadly virus.

3.The problems that you faced: Problem with accessing 20 items each time when running the program. Problem to make change in database shows in SQL. It ends up the database file is in a different folder as the path is only the name of the file. Problem in dealing with how to reset the database once 100 items is stored in the database.

4.Code for calculation: Calculation.py
Output file: CalculationOutput.txt

5.Visualizations
(1)Current US coronavirus cases in states
A.Code: Current US coronavirus cases.py

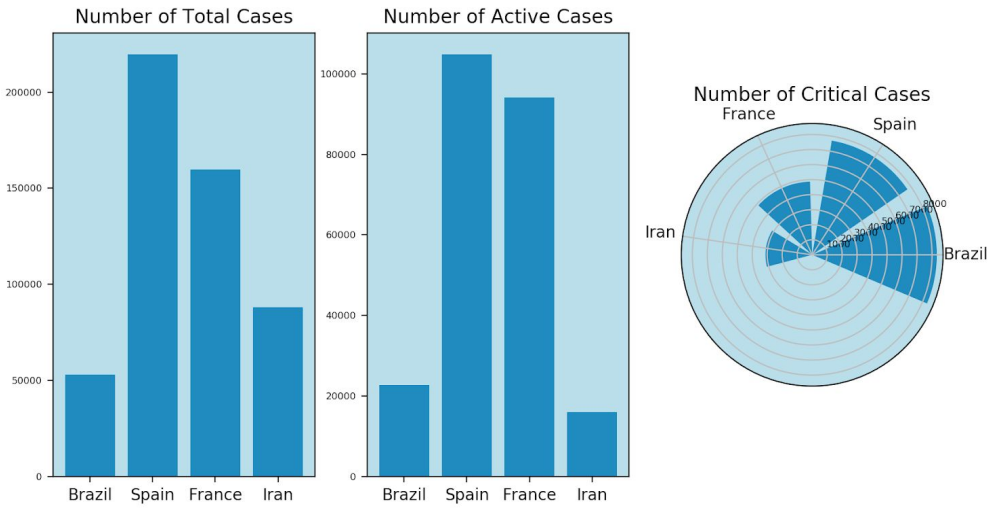



B. Image:
C. Image file: Current US coronavirus cases.webarchive

(2) Other than America, top four countries that has most COVID-19 critical cases

A.Code: plots.py

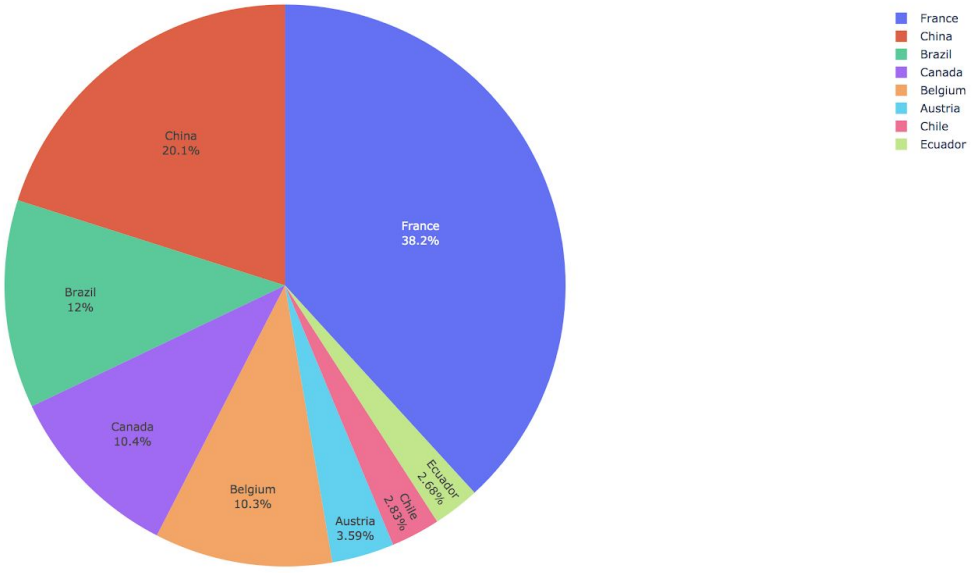
B. Image:




(3) Distribution of COVID-19 confirmed cases among country with more than 10000 confirmed cases

A.Code: graph-anna.py

B. Image:




6.Instructions for our code:

A.Our finalProject.py will gather data from each api and insert data into the database in **finalDB.db**

B.Then we do calculation in **calculation.py** and the result of calculation are in **CalculationOutput.txt**.
C.The code for making current US coronavirus cases is in **Current US coronavirus cases.py**
D.Visualization image of current US coronavirus cases: **Current US coronavirus cases.webarchive**
E.The code for visualization of top four countries that has most COVID-19 critical cases: **plots.py**
F.The code for visualization of Country distribution with more than 10000 cases: **graph-anna.py**
G.The final report: **SI 206 project report.pdf**

7. Documentation for functions:

Function1:

Collect(database) is in finalProject.py, which helps collect data and insert into the database.
Input is a string ("finalDb.db"), which is the name of our database.
There is no return statement in the function, but we write data in the database.

Function2: calc(database, txtfile) is in calculation.py

Input is two strings database and txtfile.
There is no return statement in the function, but we write calculations into calculationOutput.txt.
It selects the data from the database, and does calculation.

8. Documentation for sources

(1)Date: 4/22/2020
Issue Description: current and recent US coronavirus cases
Location of Resources: https://covidtracking.com/api
Results: got data of current and recent US coronavirus cases

(2)Date:4/21/2020
Issue Description: coronavirus cases of various countries in the world
Location of Resources: https://coronavirus-19-api.herokuapp.com/countries
Results: got data from coronavirus situations of different countries

(3)Date: 4/20/2020
Issue Description: The confirmed, death, and recovered COVID-19 cases of all countries in the world
Location of Resources: https://documenter.getpostman.com/view/10808728/SzS8rjbc?version=latest#
Results: got data of confirmed, death, and recovered COVID-19 cases for each country