SI 206 Final Project Report

Github repository link: https://github.com/Xiao-0614/SI206-XMK-finalproject

1. **The goals for your project (10 points)**

   The goal of our project is to analyze the relationship between vaccination rate and the confirmed covid cases in the majority of the countries across the world. We also explored the differences in the number of confirmed covid cases in each country under different national policies

2. **The goals that were achieved. (10 points)**

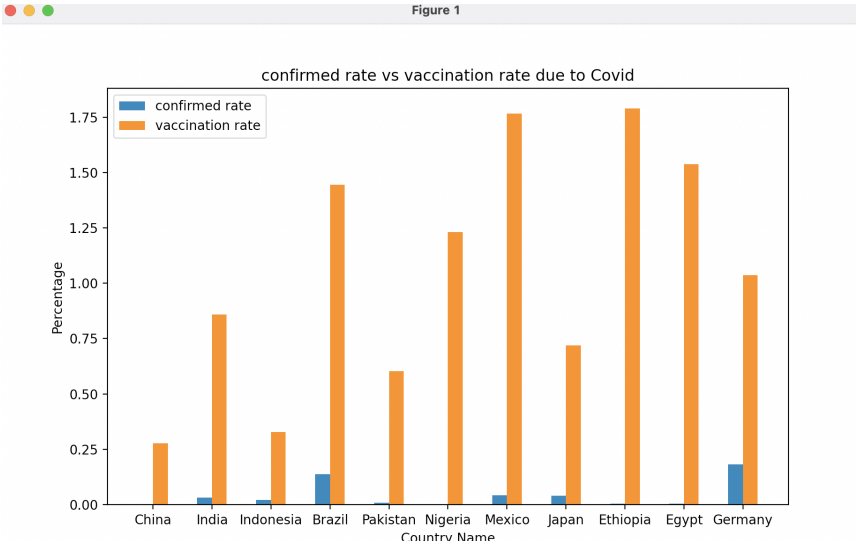
   The goals that were achieved include concluding the relationship between national vaccination rate and confirmed covid cases. There is a negative correlation between national vaccination rate and confirmed covid cases in which higher vaccination rates correlates with lower numbers of confirmed cases.

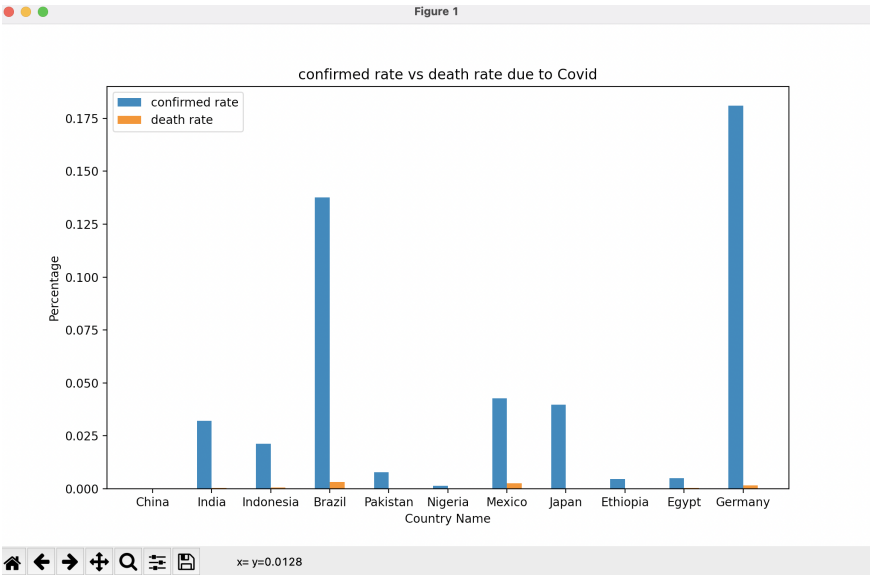
3. **The problems that you faced. (10 points)**

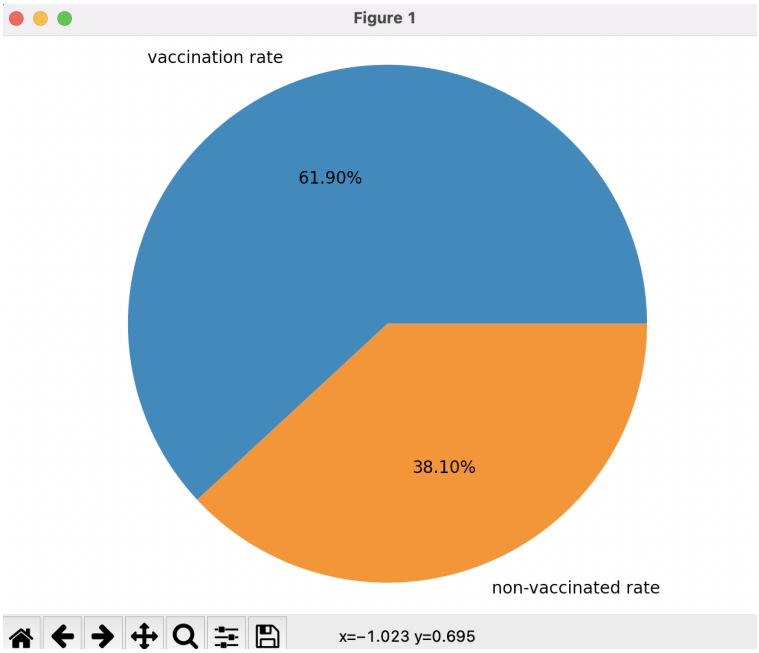
   We had some problems in finding the API, such as it was harder to find the data we wanted. And after we had identified an API and built the code, we found that the API source was deleted, so we couldn't build the table. It took us a long time to debug and find the problem finally.

4. **Your file that contains the calculations from the data in the database (10 points)**
   a. confirmed.txt
   b. deaths.txt
   c. partially-vaccinated.txt
   d. vaccinated.txt

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







6. **Instructions for running your code (10 points)**

   We write everything inside one file. Therefore, you can just run that particular file named final.py

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

a. def getVacData (): input: API source file. It intends to parse a valid JSON string and convert it into a Python dictionary.
b. def getGovData(): input: API source file. It intends to parse a valid JSON string and convert it into a Python dictionary.
c. def build_gov_table(gov_json_data, cur, conn): input is what returned from get function getGovData() mentioned earlier. This function intends to build a table with four columns: country code, confirmed cases, deaths cases, stringency actual.
d. def build_code_table(cur, conn): This function intends to build a table by using beautiful soap to query data from websites. A table was created with a country name corresponding to its three-digit country code.
e. def build_vac_table(vac_json_data, cur, conn): This function intends to create a table with four columns by reading data from API. Those four columns are Text, number of people vaccinated, number of people partially vaccinated, and number of population
f. def calConfirmed_Deathrate(cur, conn): This function intends to calculate the confirmed cases rate and death rate in different countries across the world. For each country, we use two formulas to conclude the results: confirmed cases / total population and death numbers / total population
g. def calVacRate(cur, conn): This function intends to calculate vaccination rate for each country by using the formula: number of people vaccinated / total population for each country.
h. def confirmed_vs_death(data1): Input of this function is confirmed case rate and death rate calculated earlier. This function intends to draw a bar graph in which the x-axis is 10 countries with relatively large populations. Y-axis are confirmed case rate and death rate.
i. def confirmed_vs_vac(data1, data2): Input of this function is confirmed case rate and vaccination rate calculated earlier. This function intends to draw a bar graph in which the x-axis is the top 10 countries with large populations. Y-axis are confirmed case rate and vaccination rate.

8. **You must clearly document all resources you used. The documentation should be of the following form (20 points)**
    a. April 20th Having trouble querying data from website by using beautiful soup https://www.pluralsight.com/guides/extracting-data-html-beautifulsoup Yes, it did solve the problem.
    b. April 22th https://www.geeksforgeeks.org/bar-plot-in-matplotlib/ Having trouble drawing a bar chart with two bars. Yes, it did solve the problem.
    c. https://www.askpython.com/python/examples/pull-data-from-an-api, Having trouble link to the api. Yes, it did solve the problem.
    d. April 24th have trouble draw pie chart, https://matplotlib.org/stable/gallery/pie_and_polar_charts/pie_features.html, Yes, it did solve the problem.
    e. April 24th have trouble draw bar chart (generate x and y axis), https://matplotlib.org/3.5.0/api/_as_gen/matplotlib.pyplot.bar.html, Yes, it did solve the problem.
9. **Conclusion**

    The death rate of getting covid is relatively low while the diagnosis rate is pretty high. There is a negative correlation between vaccination rate and diagnosis rate. Diagnosis rate is also negatively correlated with strictness of government resistance to the covid.