COVID-19 VACCINATION DRIVE

IAT 814 Final Project Report

Submitted by:
Nirosha Bugatha(Student ID 301398124)
Shreya Sadashiva Shetty(Student ID 301406211)

Overview:

Ever since 2019 December, we all are part of a global pandemic i.e COVID 19. Currently, there are 25 million active cases all over the world. There has been global panic amongst people and newly developed vaccines are a sign of relief. Many pharmaceutical companies are developing/developed vaccines and there is no holistic view of global vaccination progress. There are currently different trackers designed by the COVID-19 vaccine tracker organization and a blueprint developed by WHO which lack an interactive or user-friendly dashboard and displays only the leading vaccine producers but lacks information on the vaccination drive

With a handful of vaccines now authorized for use, we would like to know & compare the vaccination drives in various countries.

Goals:
- Distribution of vaccines in developed, underdeveloped, and developing countries
- Different types of vaccines used in various countries
- Comparison of covid cases w.r.t population of each country and the number of people vaccinated in each country

Intended Audience:
- Government health agencies
- Vaccine development companies
- Any common man interested to see a holistic view of the vaccination drive

Description of underlying Data:

The data has been collected from various sources to create this dashboard. The data is static and mainly contains information on various vaccines, their country of origin, distribution of each amongst various countries, active covid cases in each country with the population, the economy of the countries, and the total people vaccinated. The main data collection was from our world in data and kaggle.

Our approach for data engineering as part of this project included a pipeline consisting of collecting raw data from various sources mentioned below, filtering some of the null values, removing unrelated attributes from the data using python pandas, merging all the relevant data using a common attribute like country and separating other data into different CSV files. We did not find data for the actual share of people vaccinated with different vaccines in each country, hence we had to add some fake values manually to this attribute.
Below are our data sources,

- Covid-19 World Vaccination Progress
  The data is collected from the Our World in data website has country, vaccines, people vaccinated, etc
- Population by country
  The data is collected from the world meter website which has country and population from this dataset
- Latitude and longitude information
 This dataset has all the country's latitude and longitude details.

- Confirmed covid-19 cases

  This website has the confirmed covid cases which we have used in forming our dataset

Visualization Design and its usage:

We have developed a dashboard that uncovers the COVID-19 vaccination drive throughout the world. The design encoding choices we did are: the first one being the color used for vaccines, size denotes the relative percentage of people vaccinated, color saturation uses grayscale to differentiate the Income group types of the country (Developed, Underdeveloped, and Developing).

Dashboard:
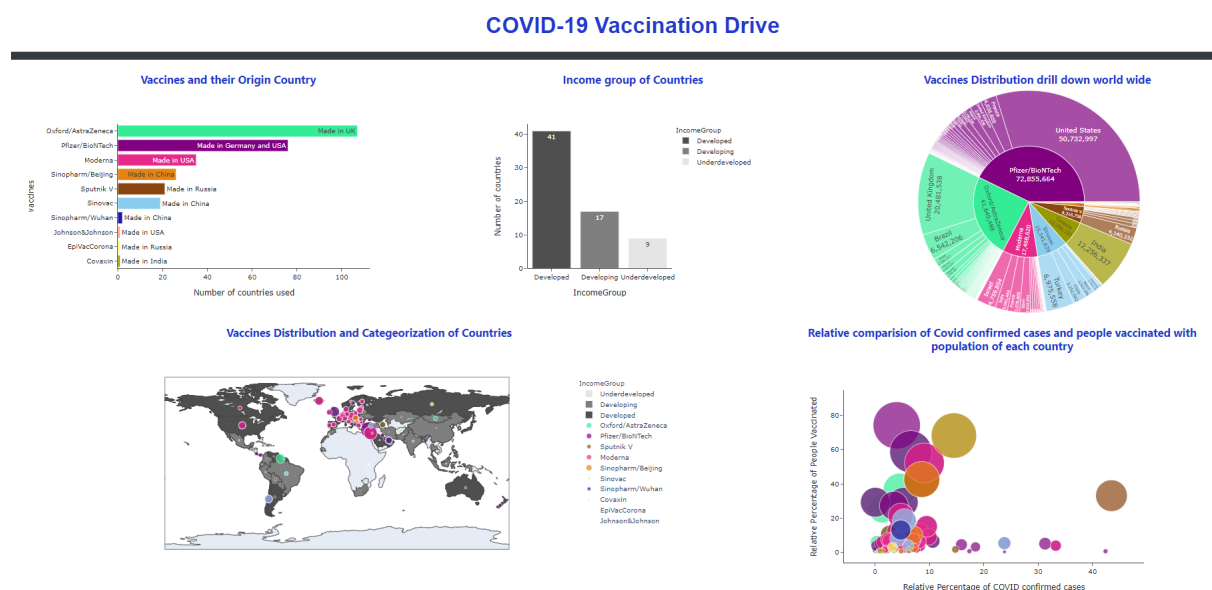
The above is the dashboard of our application- COVID-19 Vaccination Drive. We have chosen 5 charts to show various elements involved like the different vaccines, the share of different vaccines for each country, the economy of the country, vaccination percentage for each country, etc., Let's look at each of the visualization and the rationale behind them,

1. **Bar chart for vaccines**: The top-left visualization is a bar chart. We have chosen this to show all the different vaccines and their country of origin. This bar chart also acts as a legend for vaccines throughout the dashboard. The color encoding of all the visualizations is based on vaccine colors shown in this bar chart.

2. **Bar chart for economy**: The second visualization is a bar chart that shows the three divisions of economic growth of a country i.e., Developed, Developing, and Underdeveloped. Again, this acts as a legend for the user to differentiate a country based on its economy. Since color is already dominant for vaccines, we have chosen a grayscale color saturation to encode this data in our maps.

3. **Hierarchical pie-chart**: The top-right visualization is a hierarchical pie-chart, although plotly calls this a sunburst. This viz shows the various vaccines used by countries all over the world. The parent is a vaccine whereas the children are the respective countries using the vaccine. This helps the user know which vaccines are the most used by countries.

4. **Maps**: The bottom-left visualization is a choropleth map layered with a scatter geo map. This helps in comparing the vaccination rate amongst the countries. The choropleth uses grayscale color saturation to differentiate between developed, developing, and underdeveloped countries whereas the layered scatter map shows the vaccine used in each country with the color encodings for vaccines and size represents the percentage of people vaccinated.

5. **Bubble chart**: The last visualization is a scatter bubble chart that compares the relative covid cases and the percentage of people vaccinated for each country. The color denotes the vaccine used and the size represents the percentage of people vaccinated. The dominant question here is the vaccination rate for countries.

Interactions:

1. **Filter based on clicking on a vaccine legend in the bar chart**:
    The user can click on any of the vaccines on the bar chart and all the other visualizations interact forming a connected view. When the user clicks on 'Moderna' we can see all the other visualizations highlighting only data related to Moderna vaccine color encoded with pink.


COVID-19 Vaccination Drive

2. **Select multiple vaccines from the vaccines bar chart**:
   The user can also select multiple vaccines from the bar chart and see the corresponding comparisons in all the other charts.
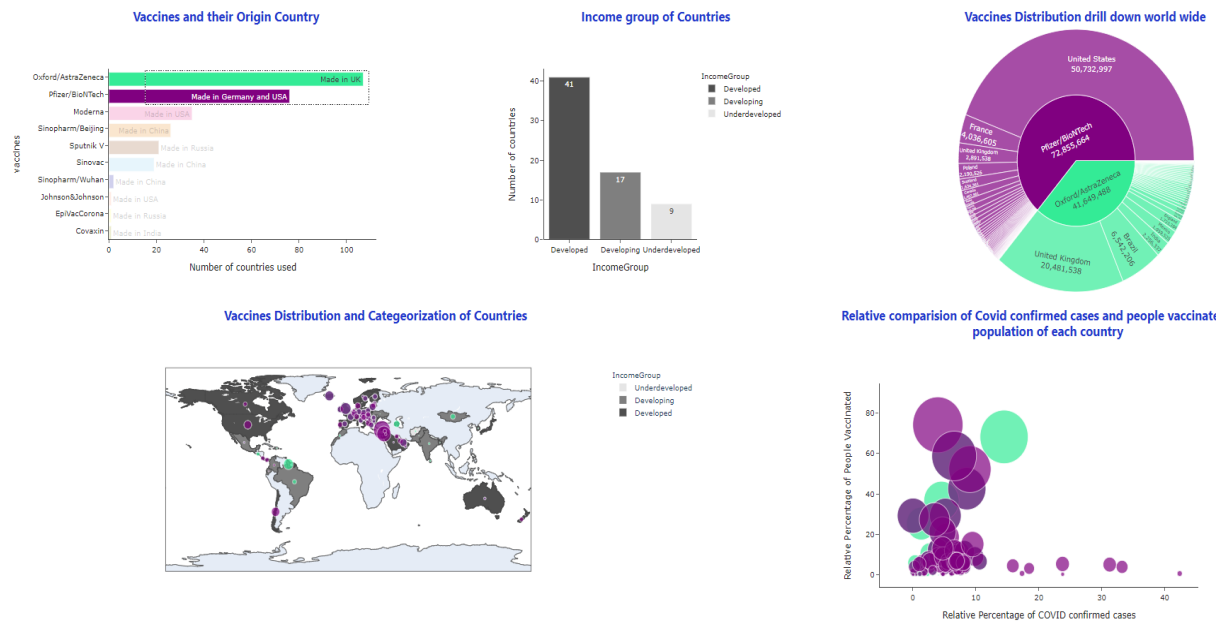

3. **Filter based on the economy bar chart**:
   Users can choose between the different economies by either clicking on each of them or selecting multiple economies at once. This will show all the data about only those countries who come under the selected economies in all other visualizations.
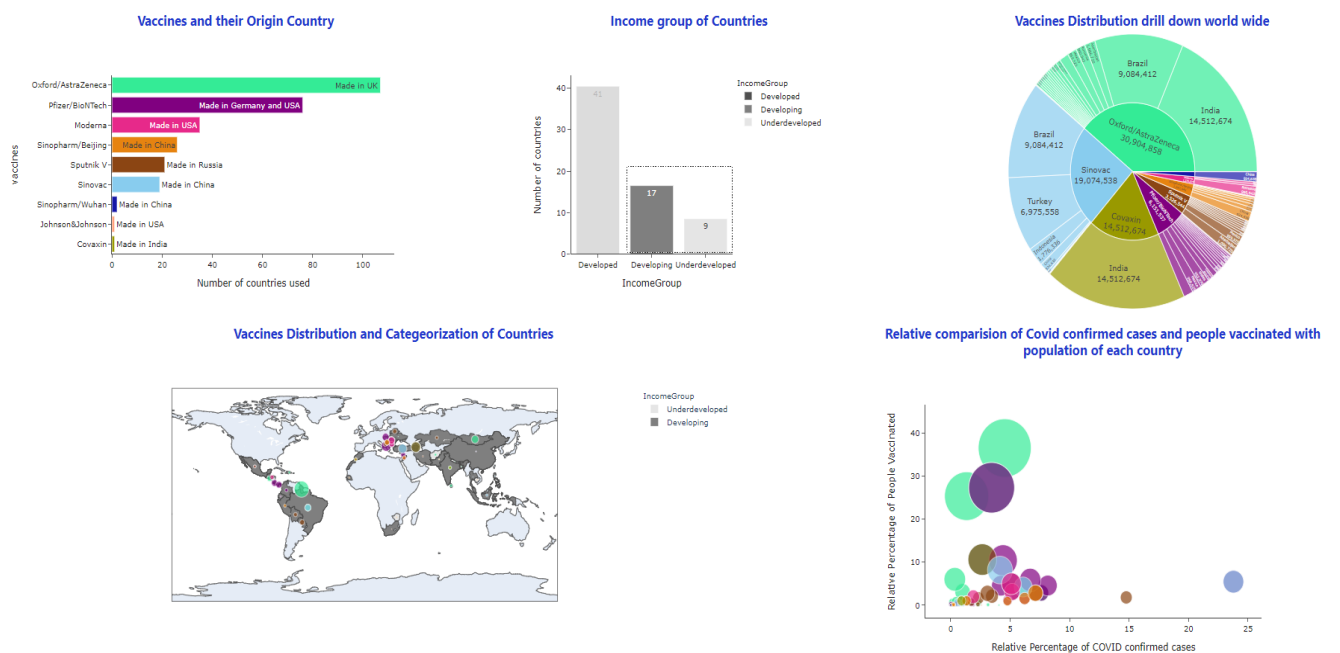
4. **Filter based on vaccine/country from hierarchical pie-chart**:
The users have the flexibility to start from any chart from the visualization. They can select either a country or a vaccine from a hierarchical chart to see the respective results in all other charts.
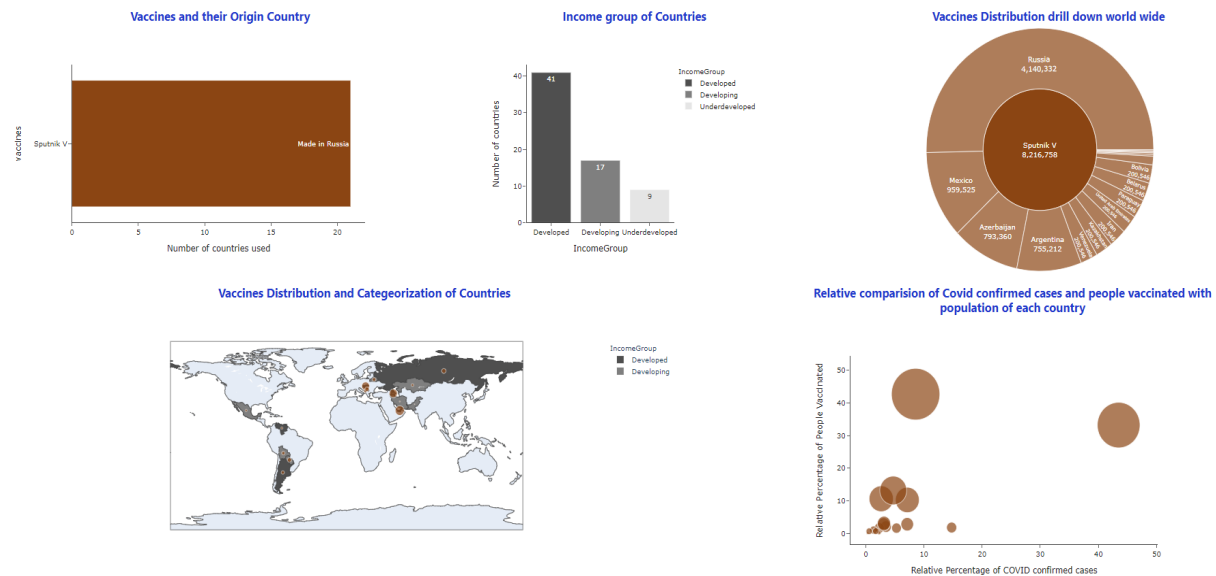
5. **Filter based on the country from the maps**:
Users can click on any country or select multiple countries on the map to view all the information about its vaccination drive being highlighted in all the other charts. This will help in comparing vaccination drives between countries.
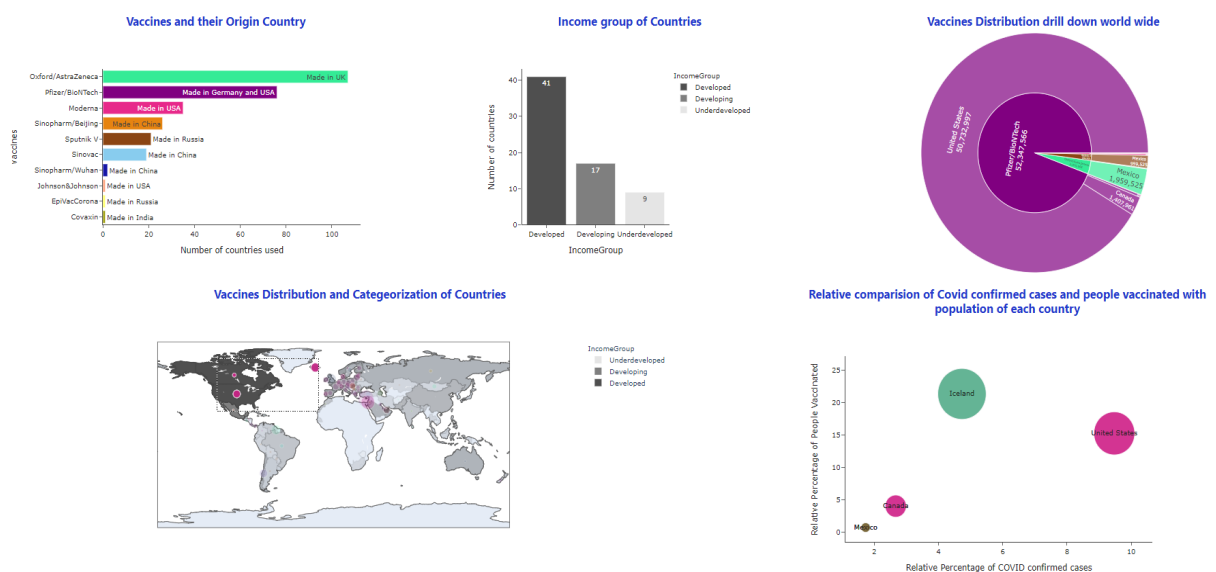
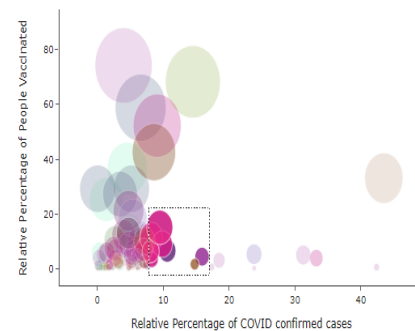6. **Filter based country from bubble scatter plot**:
   Users can either click on a bubble or select multiple and see the respective information in maps being highlighted: the grayscale showing the economy, the bubble color defining the vaccine with size defining the relative vaccination.
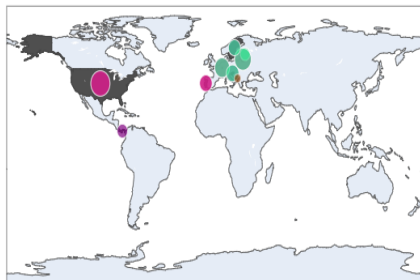
Limitations and Future work:

There were few limitations on the Plotly dash tool such as the axes setting, Size legend, etc., In the future we would want to implement the below-mentioned tasks:

- Times series data visualizations showing the trend of covid cases since 2019 and the trend in the vaccine distribution starting 2021 until now.
- Sankey chart which shows the Supplier and Buyer countries of each vaccine in a color-coded form
- We want to divide the visualizations into different tabs each tab filtering data based on a category either country, month, vaccines, etc
- Update the visualizations with dynamic or live data
- Add more data elements to the current visualizations such as the impact of covid on the economy of the country, job market, immigration, etc

Conclusion:

We have created a COVID-19 vaccination drive dashboard using all the design principles and the knowledge gained during the lecture and the readings of IAT 814. It was great to learn plotly dash as a visualization tool. This dashboard can provide a holistic view of the ongoing vaccination program to the respective audience and insights to questions such as the vaccines used in various countries, the most used vaccine across the globe, vaccination progress in each country, comparison of vaccination drives amongst countries based on their economy, covid cases, etc.,

We hope to improvise the dashboard with dynamic data updates and time-series visualization to show the relative change concerning month/year to make it more insightful.

References

- https://www.kaggle.com/gpreda/covid-world-vaccination-progress
- https://www.kaggle.com/tanuprabhu/population-by-country-2020
- https://developers.google.com/public-data/docs/canonical/countries_csv
- https://ourworldindata.org/covid-cases

Links

Website: **https://covid-19-vaccination-drive.herokuapp.com/**
GITHUB: **https://github.com/Nirosha-Bugatha/COVID-19-Vaccination-Drive**
Video :