CSE 564 - Visualization and Visual Analytics

Final Project Proposal - Group 59

Piyush Soni (113273375) *Abhishek Bhattad (113277131)*

This proposal is being drafted in response to the final project. For the final project, we have finalised to make a dashboard displaying various attributes related to COVID-19 as the world is still suffering from it and some countries are already witnessing the 2nd wave of it. In the past one year, there has been a huge development in vaccine creation and vaccination drives are being done all over the world at a great pace. We want to explore the impact of COVID-19 on different aspects like pollution, unemployment, etc., and impact of vaccination on COVID-19 infections.

We have gathered a time series data to check the current and previous COVID situation. Below is the description of datasets we have collected.

1. **COVID-19 data:** The world wide dataset is generated by **Our World in Data** on a daily basis. It contains data including but not limited to confirmed cases, deaths, hospitalizations, testing, vaccinations and other related attributes.

   https://github.com/owid/covid-19-data/tree/master/public/data

2. **Unemployment data:** We will try to analyse the impact of covid-19 on unemployment. We know that the unemployment rate was very high in the previous year as the world was halted. How about the current situation as the world has adapted to the new normal?. We have gathered the data from Google trends.

   https://trends.google.com/trends/explore?q=%2Fm%2F07s_c

3. **Pollution data:** As the world came to halt and lockdowns were enforced all over the world to curb the spread of covid-19, there must be less cars, less industries running and thereby less pollution. How about now? The data is being collected from Google trends.

   https://trends.google.com/trends/explore?q=%2Fm%2F066xq

4. **Bitcoin (Cryptocurrency) data:** The market was going through down slash during pandemic. How about the crypto market, particularly bitcoin? We will try and compare about the impact of covid-19 on bitcoin prices then and now.

   https://trends.google.com/trends/explore?q=%2Fm%2F05p0rrx

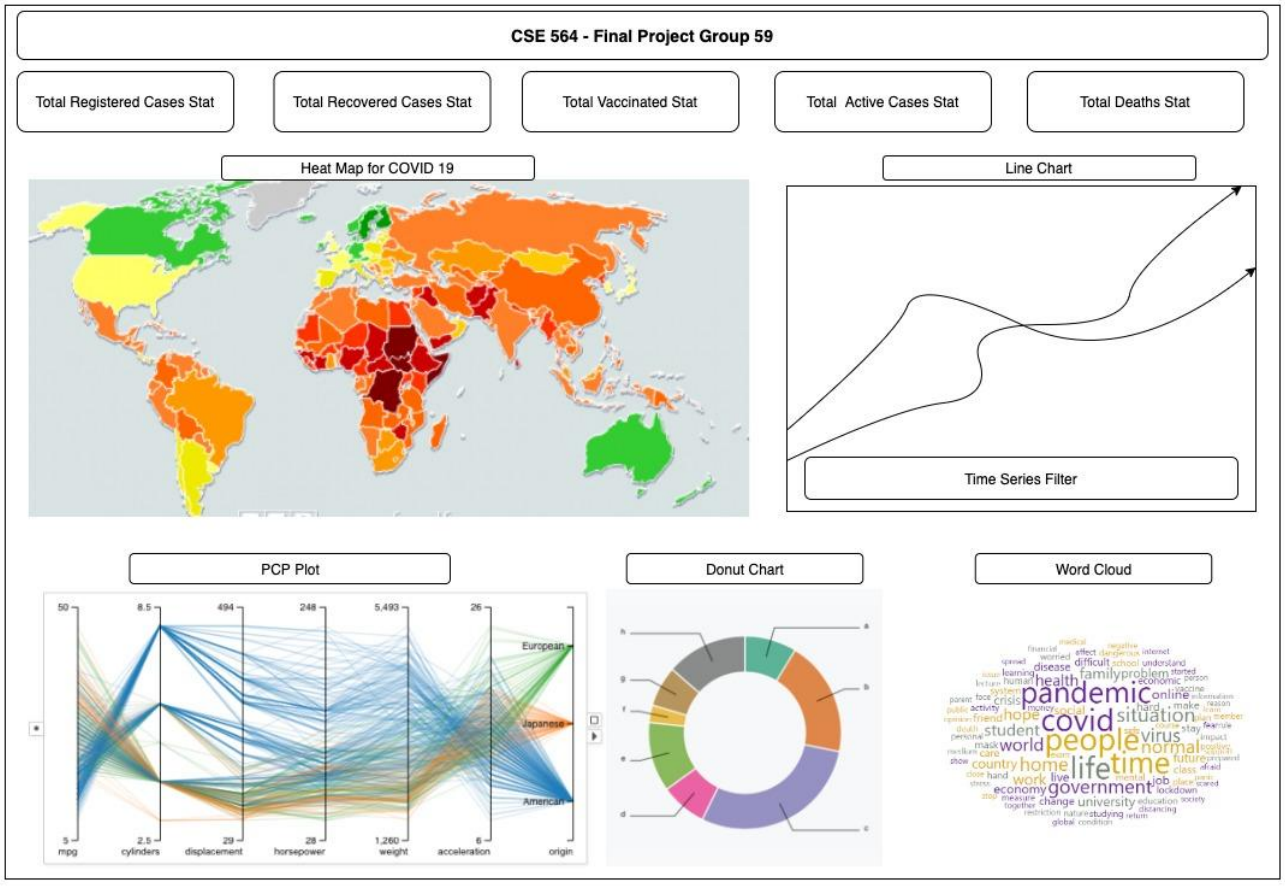
We will merge these data sets on the same time series attribute and do pre-processing of data (null values, downsampling, upsampling, removal of attributes, etc).

Below are the few attributes that we plan to present in the dashboard. We may update attributes as we explore further.

1. date
2. continent
3. location
4. total_cases_per_million
5. total_deaths_per_million
6. hosp_patients_per_million
7. total_tests_per_thousand
8. positive_rate
9. total_vaccinations_per_hundred
10. people_fully_vaccinated_per_hundred
11. population
12. gdp_per_capita
13. female_smokers
14. male_smokers
15. hospital_beds_per_thousand
16. unemployment_rate
17. Pollution_rate
18. Bitcoin_price
19. human_development_index

Dashboard:

Below is the snapshot of the dashboard we plan to make.



Plot 1: **World Heat Map:** based on the number of covid-19 cases around the world, intensity of the heat map will change. More the covid case, more will be the intensity.

Plot 2: **Line Chart (Standard):** line chart will contain time series data. Each line will be one attribute like hospital_beds_per_thousand, positive_rate, etc. In this way multiple attributes can be visualized into a single chart. The slider below the chart can be used to select a particular time-series change.

Plot 3: **Parallel Coordinates Plot (Non-Standard):** PCP is one of the best non-standard plots that removes ambiguity and clearly displays the data without scaling. We will use PCP to analyse correlations between different attributes and do brushing as well.

Plot 4: **Donut Chart (Standard):** It is an upgraded version of pie-chart. We will select attributes to get displayed in the donut chart. While hovering over each, you will be able to see the info about it in the center of the chart. Filtering on other charts will impact the donut chart attributes' values as well.

Plot 5: **World Cloud (Non-Standard):** This chart is a great way to gauge attention of the user to see in order the information (big font corresponds to more intensity). Example, total covid cases in the US across states. If New York has the highest number of cases it will be carrying the highest font.

Header Statistics: We are planning to show some statistics at a glance at the top of the dashboard.

We will be linking plots using filtering, clicking and brushing.

Hypothesis:

Following are our preliminary expectations from the data, there might be more findings as we analyze the data further.

1. Effect of COVID-19 vaccination on the total number of daily infections, expecting lower infections as vaccination increases.
2. Relationship between the GDP of a country and it's vaccination rate, expecting a high vaccination rate for the countries with high GDP.
3. Relationship between hospital beds (medical infrastructure) vs covid deaths, expecting high death for the countries with not good health infrastructure.
4. COVID-19 impact on unemployment or job opportunities, expecting high unemployment during peak of covid wave.
5. Will try to find any pattern in mortality rates among male and female smokers.
6. Bitcoin price analysis during pandemic time, as most of the transactions were moved online.
7. Relationship between GDP and unemployment rate.
8. Handling of COVID-19 by developed vs developing countries.