Rationale Report

1. Description of Initial Questions

This report will seek answers for some of the key questions as mentioned below in order to analyze the Covid-19 outbreak closely.

- How many total and new corona virus cases are reported weekly for each country?
- How many total and new deaths are reported every week for each country?
- What is the relationship between the life expectancy of the countries with their respective GDP per capita during the outbreak?

2. Description of Visualization Strategies

2.1. Data transformation/Data cleaning

The Covid-19 dataset which I selected for data analysis was so messy having incomplete and irrelevant information due to which I was not able to get accurate answers for the above mentioned questions so in order to prepare the data for analysis I used python's pandas library. I replaced the missing values in each column by the average of that specific column, removed duplicate rows and deleted irrelevant columns from the dataset. Once transforming the raw data into clean and usable form I sorted the columns by date and created a new column for Year and week in order to get weekly cases for each country.

3. Description of Information Visualization Design

3.1. Visual Encodings

Encoding in data visualization means translating the data into a visual element on a chart or map. In my data visualization I used color, size, area, and position as visual elements so every time there is a change in data the size, color, area and position changes according to the data which helps the viewer to analyze the data in a more clear way.

3.2. Data Visualization

Data visualization refers to the techniques used to communicate data and information by encoding it as visual objects such as charts and graphs etc. I used choropleth map and bubble chart in my information visualization.

Choropleth Map

I selected choropleth map to display new and total cases and deaths reported for each country as choropleth maps are a good way to display divided geographical areas or a region for a specific country that are colored or shaded in relation to a data variable which in my case was total cases, total deaths, new cases and new deaths.

Bubble Chart

I selected Bubble plot to show the relation between life expectancy and GDP per capita for each country as the bubble plot is a good way to show the value of two numeric variables on the X and Y axis also it allows to study the value of a third variable which in my case was population for each country using different sizes for the bubbles.

3.3. Interaction

Interaction in a data visualization enables the user to see more details. In my information visualization I used different interaction techniques for both charts. I used a dropdown menu for the choropleth map so every time a viewer selects a data variable from the dropdown menu the map gets updated. The viewer can then explore the map by revolving it, zooming in, zooming out and get more details about specific region by hovering the mouse over a specific region. In addition, I added a slider by which the user will be able to know weekly reported cases and deaths for a specific country. The viewer can also play the button to start animation which will give the idea to the viewer as how the virus is spreading in the world or how many total deaths are reported across the world.

I used checkboxes for bubble chart so every time a viewer selects a continent or number of continents, the bubble chart gets updated according to the countries in the continent to show the relationship. The viewer can then hover the mouse over the bubbles to get more details.