Part 3. Implementation

To run the implementation, run the " index.html"

Data processing:

The main datasets used are Dataset 1 and Dataset 2, under "asset/data/Big datasets/", and the derived datasets from all files under the "asset/data/". An external dataset is countriesgeo.csv, which contains the longitude and latitude information of each country, is also file under "asset/data/Big datasets/".

Link: https://developers.google.com/public-data/docs/canonical/countries_csv

All the derived datasets are processed using Python. Below are the files names and description of their processing :

1. **combine.csv**: a dataset that result from integrating Dataset 1 and Dataset 2, contain the main quantitative information: cases, deaths, tests, hospitalization. I take the first month of the Dataset 1 (31 Dec 2019 – 31 Jan 2020) including the deaths and cases of everyday of every country, and append it with Dataset 2 , from 1 Feb 2020 to 5 April 2021. Inner join both data on "date".

2. **cases.csv** : daily and cumulative data of new COVID-19 cases reported, everyday from 31st December 2019 till 5th April 2021, worldwide. The data results from using combine.csv, grouping by the date, and summing over the case data.

3. **country_cases_geo.csv** : cumulative data of new COVID-19 cases, hospitalisation, tests of each country reported. The data also have the longitude and latitude data for each country. The data results from using combine.csv integrating with countriesgeo.csv (inner join on "location"), grouping by the countries, and summing over the all the data for each category.

4. **country_deaths_geo.csv** : cumulative data of new COVID-19 deaths, hospitalisation, tests of each country reported. The data also have the longitude and latitude data for each country. The data results from using combine.csv integrating with countriesgeo.csv csv (inner join on "location"), grouping by the countries, and summing over the all the data for each category.

5. **deaths.csv** : daily and cumulative data of new COVID-19 deaths reported, everyday from 31st December 2019 till 5th April 2021, worldwide. The data results from using combine.csv, grouping by the date, and summing over the case data.

6. **hospital.csv** : daily and cumulative data of new COVID-19 hospitalisation and COVID-19 ICU hospitalisation reported, everyday from 31st December 2019 till 5th April 2021, worldwide. The data results from using combine.csv, grouping by the date, and summing over the case data.

7. **test.csv** : daily and cumulative data of new COVID-19 positive test reported, everyday from 31st December 2019 till 5th April 2021, worldwide. The data results from using combine.csv, grouping by the date, and summing over the case data.

8. **owid-covid-geo.csv** : a dataset that result from integrating Dataset 2 and countriesgeo.csv. It provides all the quantitative information: cases, deaths, tests, hospitalization, and also the geographical location (longitude and latitude) of each location. Performed Inner join on "location".

Each dataset is utilise for :

Line/area chart: cases.csv , deaths.csv, hospital.csv, test.csv

Bubble map : owid-covid-geo.csv

Bar chart: country_cases_geo.csv, country_deaths_geo.csv

User interaction:

1. Scroll bar

It is included in the first column and the third column of the visualisation (counting from left to right). It is the simplest level of user interaction included, that aids the visualisation as a whole by reducing the compromisation of the figures size. Without using the scroll bar, the graphs are then needed to reduce in size, resulting harder/ erroneous reading of the details.

Reference link(s):

i.   https://github.com/Grsmto/simplebar
ii.   https://www.youtube.com/watch?v=74eaw_nM5tY&t=183s

2. Dropdown menu selection

It is included in the first column and between the second and third column:

1. In the first column: it allows selection between daily or cumulative data on cases, deaths , tests, hospitalisation on the on the line chart in the first column
2. Between the second and third column: it allows selection between cases or deaths data to be displayed for each country.

Including the menu selection data helps to simplify the visualisation, without having overlay multiple data on top of each other. Therefore, it can easily highly the data to be investigated.

Reference link(s):

i.   https://www.d3-graph-gallery.com/graph/line_select.html

3. Tooltips

There are 3 types of tooltips utilised in this visualisation: line charts, bubblemap, and Bar charts.

Line chart: There are a total of 8 line charts utilised in this visualisation, showing the daily and cumulative data of cases, deaths, hospitalisations and tests. All the line charts are created by appending a line, x-axis and y-axis onto an svg. An area with the same colour as the line are then appended under the line to create the chart. Then, I appended tooltip onto the svg, listening to events touchmove and mousemove, to display information of the data on the selected part: the date of the selected part and the data associated with it.

Bubblemap: The Bubblemap is created by creating a map and appending bubbles/ circles over it. The size of the bubbles scale with the data (cases/ deaths): the higher the number count of cases/ deaths, the bigger the bubble. Each bubble are located on top of the map using, the longitude and latitude information of each country to aid with the positioning. The tooltips I appended onto the bubbles, where when the mouse hover over, the colour of the bubble will change, indicating the selected bubble. After the mouse leave, the colour will then change to a different colour, indicating that bubble has been selected before. As there are many bubbles on top of the map, the tooltip not only aids in visualising the numerical data (cases/deaths), it also helps to reduce the complexity while reading the data by "ticking off" which bubbles has been browsed ("pop" effect).

Bar chart: There are a total of 2 horizontal bar charts utilised in this visualisation, showing the cumulative cases and deaths of COVID-19 up to 5th April of each country. The bar charts are created by appending a rectangle that scales with the cases/ deaths count , x-axis and y-axis onto an svg. The bar charts are arranged in descending order of the cases/deaths counts so the bar chart can also show the ranking in terms of the "riskiness" of the countries. Then, I appended tooltip onto the svg, listening to events mousemove and mouseout, to display information of the data on the selected part. A selected country will show its cumulative deaths/ cases, hospitalisation, and test cases up from 31 Dec 2019 till 5 April 2021. If there is no information regarding a specific field, for example , if no test data for Philippines, it will then display " No data" in the following field. The tooltip helps to give an up-to-date summary of all the data being investigated in the visualisation.

Reference link(s):

    i.   https://www.d3-graph-gallery.com/graph/bubblemap_template.html
    ii.   https://www.d3-graph-gallery.com/graph/bubble_tooltip.html
    iii.   https://bl.ocks.org/larsenmtl/e3b8b7c2ca4787f77d78f58d41c3da91
    iv.   http://www.d3noob.org/2013/01/adding-title-to-your-d3js-graph.html#:~:text=What%20we%20want%20to%20do,append(%22text%22)%20
    v.   https://observablehq.com/@bsaienko/animated-bar-chart-with-tooltip
    vi.   https://www.d3-graph-gallery.com/backgroundmap
    vii.   https://www.d3-graph-gallery.com/graph/barplot_horizontal.html

4. Slider

There are two sliders implemented in the visualisation, each controlling a different map (cases / deaths). However, the building of both sliders is similar , but just appending onto different sets of data. The slider implemented allows the user to control the date between 31 Dec 2019 to 5 April 2021. While sliding through the date, for every date, the bubble will change its size according to the number of cases/ deaths. The slider acts as an important element that helps to investigate the trend of spread of virus over time.

Reference link(s):

    i.   https://observablehq.com/@bradvoracek/d3-simple-slider
    ii.   http://dev.centrogeo.org.mx/viz_desaparecidos/lib/d3.slider/