Story of Covid-19

Department: Business Analytics MSc

Course: Analytics and Visualization for Managers and Consultants

Course Code: BEMM461

VISUALISATION PROJECT

Name: Kutay Onlen

Student ID: 700059706

Table of Contents

Introduction1

Datasets1

**Minimum Ink Ratio 2**

Nested Model of Validation2

1) Domain3

**2.1) Data3**

**2.2) Task4**

**3) Visual Encoding / Interaction Idiom4**

**4) Algorithm5**

**Self-Evaluation5**

**References6**

**Introduction**

According to World Health Organization (WHO), over 6 million people died from covid-19 until 6th of May 2022 (WHO, 2022). In this period, I believe that many stories are waiting to be discovered. In this visualization project, my purpose is trying to give an opportunity to people to discover some of these stories by themselves in a most comfortable way as possible. I used the story feature from tableau, which people can insert many sheets, dashboards, and animations in it, to increase content that provides more insight into pandemic from WHO dataset. Story has interactive map graph, interactive animation, and dashboard that has bar chart, packed bubbles graph, and area chart. In this report, first, I am going to give insight of the dataset. In the second part, there will be impacts of few academical visualization evaluation methods on the project. At last section of the report, I will reflect my personal critics for the project.

**Datasets**

I used two datasets for this project. First dataset shows cumulative death and covid-19 cases as well as new death and new covid-19 case numbers based on the reporting date. Most of the countries has been reporting in daily basis. However, because of government officials gives just the cumulative numbers of the cases and deaths, WHO subtract the cumulative number of a day from latest cumulative number. **So, when I say daily numbers of new cases and new deaths generally most of the numbers belong to the previous day.**

Text

Description automatically generated

(WHO, 2022)

Second dataset is also from WHO, which is focus on vaccination situation in the countries.

A picture containing text

Description automatically generated

(WHO, 2022)

**Minimum Ink Ratio**

Tufte (1984), argues that a visualization gets better when the ratio of data ink over total ink of the canvas increase. The purpose of this idea, getting rid of from unnecessary amount of ink that has no additional value to the visualization. I tried to use white background in all my dashboards, but in map chart, data points are easy to observe than other options. In addition, I insert many filters, because of the purpose of the project, which is providing different insights in every element of the project to the people.

**Nested Model of Validation:**

This method also referred as Munzner’s 4 Level of Validation. This method evaluates the visualizations in 4 different perspectives and feedback of these perspectives from the audience or focused groups (Munzner, 2009). I did not make the survey or any focused groups study, so in this part of the report I am going to focus on processes of the project.

**1) Domain**

Graphical user interface, text, application, email

Description automatically generatedWhen I think about the audience of this project, I have tried to prepare the visualizations anyone who finish the high school. I have prepared this project to publish it in a news website, so people can spend time on their home and analyze the pandemic through these visualizations. I have also used basic parameters, so everyone can understand the visualizations. However, anyone who wants to dive deep of the visualizations can get much more detailed information.

**2) Data / Task Abstraction**

**2.1) Data**

On the first dashboard, I have prepared graph that shows countries and their cumulative death numbers, percentage of fully vaccinated people, countries’ location and finally I have showed another chart that give the countries’ daily death numbers from the first death that covid-19 caused and until 6th of May 2022. On the second dashboard, I have made a visualization of top10 countries that has lost its citizens due to covid-19. I have used some additional fields that I created from existing dataset such as population, daily case / death ratio. Ratio of case and death throughout the pandemic shows that in every result of the ratio, which is a covid-19 case number, 1 person died in that country. Next data that I showed in this dashboard is percentage of the people who got fully vaccinated based on countries and I have ranked them, so audience can easily see most or least vaccinated countries in a first glance. In addition, I also showed percentage of persons who got at least 1 dose vaccine and percentage of persons who got booster dose. In the 3rd and last visualization of the 3rd dashboard, I showed the number of the total death numbers for per region. Moreover, population of regions also has been inserted. So, audience can also see the covid situations in a regional perspective. In addition to that, I have inserted every necessary explanation in the tooltip, so when the audience hover over the visualizations they can see the explanation.

**2.2) Task**

I have two tasks in this project. First task is using the one of the appropriate graph types according to purpose of story and make it easy to understand by additional tools such as tooltips, filters, and colors. Second task is encouraging audience to interact with the visualizations.

**3) Visual Encoding / Interaction Idiom**

I presented the data in first dashboard in a map graph and mark the countries by blue circles with white borders, which is very easy to realize even small countries on the dark map background. Size of circles are representing cumulative death numbers of the countries. I inserted 4 tooltips, which first 3 of them are name of the country, cumulative death number of the country and percentage of fully vaccinated people. Last tooltip is second graph in this dashboard, which it shows daily numbers as I mentioned in **2.1 Data section**. I used line graph for this visualization as y axis has the number data and x axis has the date data type. This dashboard shows Covid-19 Global Overview. This dashboard is interactive, and audience also can filter the countries according to regions and population range. The second dashboard which is an animation, I used line graph. I wanted to choose animation for most effected top10 countries from covid-19, because following the pattern of lines in animation very effective to understand this visualization. I have showed in this dashboard mark labels which are inserted country name and cumulative deaths. Since we already have cumulative death numbers, I have removed y axis values except 0 and 1M. I also removed the x axis values and move date to the title. Audience can click one or multiple lines to highlight them and they can pick any date to see situation of the most effected 10 countries. At the third and last dashboard the first chart is a bar chart that shows rank of vaccination between countries. Audience can search countries in a single dropdown menu, and they can pick one or multiple of them. Audience can also select countries according to population intervals. Lastly, users also can choose countries according to regions and highlight those countries as a group. Second graph in this dashboard is an area chart of number of total covid-19 death and cases of regions. Volume of areas depend on total cases of the regions and the color depend on total death number for each region. Name of the regions shown on the graph as a mark label. Next to the area chart there is a color palette label for total death number of regions and populations of regions as a text box to compare numbers with the population of regions. Last graph is a packed bubble chart that shows Case / Death Ratio of the countries. Color of the bubbles represents interval of case / death ratio, in addition to that there is a range filter, so audience can choose any range of ratios and display those countries. This dashboard also fully interactive. If the users try any filters or click any element of any graph all of three graphs shows that country, region, or countries that stands between the population range filters. To increase interactivity with the users, I also insert little text boxes that navigate the user to enter interactional relationship with the visualizations in all dashboards.

**Algorithm**

As I mentioned before I used tableau for this project, and I just design the graphs in my head and start to process with preparing the dataset for the visualizations. In visualization process, creating the main frame of the visualizations are not complex at all. However, statistical operations are seemed very limited to compared to the python and R. For example, I could not sum up the population of regions. I was able to select countries that grouped under the regions, and I was seeing the total number which is population of the regions but could not find the way to save them. After a while, I decided to write those totals in a text box and represent in there. Population field is calculated field and I was getting error that was saying “You can not create calculated field from a calculated table”. It was very limited feature of tableau. I solved other small problems in a few clicks, which it was easier as I was keep using to program. Except learning tableau, I finished the project approximately in 6 hours. None of the operations took more than 5 seconds to finished. As a result, tableau creating faster visualizations than python and R in my experience and its less complex, but more limited.

**Self-Evaluation**

I am very satisfied with my first work from tableau. However, it was my one of the first interactive visualization project. So, I am anticipating that I have a lot of fields that I need to improve myself. I would like to create a graph about verbal analysis. For example, graph about type of vaccines that countries use. But I could not diverse the groups of vaccine types for every country. However, I think that, when I consider functions of the visualizations, I have reached every goal except region population and creating graph about the vaccine types.

**References**

**1)** Munzner, T. (2009, October 11). *A nested model for visualization design and validation*. A Nested Model for Visualization Design and Validation. Retrieved March 3, 2022, from https://www.cs.ubc.ca/labs/imager/tr/2009/NestedModel/NestedModel.pdf

**2)** Tufte, E. (1984). *The visual display of quantitative information: 3rd printing*. Graphics Pr.

**3)** World Health Organization. (2022, May 6). WHO Coronavirus (COVID-19) Dashboard. World Health Organization. Retrieved May 06, 2022, from https://covid19.who.int/info/