FIT5147 Data Exploration and Visualisation

Semester, 2, 2022

Data Visualization Project

Name: Tamanna Das

Student ID: 33059047

Tutorial Number: 1

Tutor Name: Bob Dao

Project title: Price of Covid-19 pandemic-Poverty across the globe

Table of Contents Page

I. Introduction----------3

II. Design----------4

III. Implementation----------6

IV. User guide----------8

V. Conclusion----------9

VI. Bibliography----------10

VII. Appendix----------11

Introduction

Problem description:

Performing interactive narrative visualization that communicates some of my finding from the data exploration project.

Project title:

Price of Covid-19 pandemic: Poverty across the globe

Overview:

Through this exercise I want to show my fellow students and professors the effect of COVID-19 pandemic on the global economy across the globe and how did that impact the individual lives and thus increased poverty throughout. My exploration also explains how are the factors that were impacted by the pandemic are correlated to each other.

Questions explored in my visualization:

1. How has the covid_19 pandemic affected the global economy? How are the various factors impacted by the pandemic dependent amongst each other?
2. How did the effect of covid_19 pandemic varied across individuals affecting businesses (all sizes) & increasing poverty across the globe?

Visualisation tools used:

1. R shiny

Intended audience:

My audience are adults in the age group of 21 and above. Through my visualization I want to show my audience the impact that COVID-19 pandemic had on our lives. The themes and the depth of this matter is also mature and relevant to readers who have more life experience.

Design

Sheet 1 brainstorming:

During brainstorming my aim was to how to make my visualization interactive and easy to understand for my audience. I used graphs that were previously explored in my data exploration project to start with. I started with the idea of showing the world map that shows total COVID – 19 cases across the globe with either markers showing the intensity of covid-cases or by a population slider to show that areas with more covid cases tend to have more population.

Then I referred to other charts and tried to figure our how to make those graphs visually more appealing and so I decided to add labels or tooltips to a lot of graphs so that when a reader tries to see the statistics of individual data they can look by simply hovering over the chart. I also decided to use slider inputs for the audience to see the change in statistics with variation in location. I used different types of graphs starting from a scatter plot to a map and bar charts.

Sheet 2:

In my sheet 2 design sheet I decided on a linear reading style so that the intended audience can go through my visualization and understand the effect of Covid-19 step by step. I decided on a fixed page with 2 rows. First row showing a choropleth map so that the audience can understand the variation in covid-19 total cases across the world by simply looking at the intensity of colour variation in different parts of the world. The next graph shows the % of poor in the different countries it was an area chart to see the variation in % of poor across the globe and it was placed next to the map to show that the countries with high cases had more percentage of poor. It was also accompanied with a hovering effect so that the viewers can see the stats for a specific location.

The next row included there graphs a scatter plot with different colour variation indicating different continents and a hovering effect showing a tooltip with the stats of that particular point and a population slider to show how does the stats change with change in population. The next graph element was a pie chart that indicates how many people were actually effect by the pandemic and it was also accompanied by a tooltip. My next element was a line chart it was accompanies by a colour variation of lines to show the difference between working hours that were worked and % of working hours lost due to COVID-19 pandemic. It was also accompanied by a hovering effect to show the stats for a particular location.

Sheet 3:

My sheet 3 of the design process was the same as sheet 2 except that I realised that just showing the stats is not enough it should be accompanied by some explanation of the charts for the intended audience to understand the impact and the story that I wanted to convey.

I decided to put title and explanation of the graphs on top of each chart by keeping the interactive elements as it is. I also decided to remove some graphs like the pie chart and the line chart showing individuals point of view as I believed that my motive to show the audience that the pandemic did affect the global economy and resulted in increase in poverty across the globe was justified from the 4 graphs that I used in my sheet 3.

Sheet 4:

Coming to sheet 4 I realised that showing all my data and visualization in a single tab might confuse my readers and they might get overwhelmed by seeing such huge datasets. So, I decided on using instead a tab function so that the user can shift tabs to see the different visualizations and their related explanation.

In my first tab I showed the overview of the study and what questions I wanted to explore and what I wanted my intended audience to see accompanied by a world map showing the intensity of covid cases across the globe. It was a choropleth map so that the audience can understand the variation in covid-19 total cases across the world by simply looking at the intensity of colour variation in different parts of the world. The second tab showed the variation in percentage of poor across the globe with a hovering effect so that the reader can see the stats of individual locations and the title and explanation or insights that I derived from the graph.

My last two tabs showed the variation of GDP with month and the correlation of GDP and HDI so that the audience can understand how the COVID-19 pandemic affected global economy and the how the factors are correlated. Both these graphs used colour variations to differentiate between continents. The bar graph used a hovering impact to show the data for each individual location.

Sheet 5:

My design sheet number 5 is same as my fourth sheet except the fact that the I realised that the overview and the world map in the same tab becomes a bit untidy and difficult for the reader to understand the overview of the study and hence I decided to place the overview at the starting and the next tab with only the map.

Note: All my design sheets used a linear order style for the reader so that I can explain my story to them instead of them figuring out on their own

IMPLEMENTATION

My final implementation of design in R shiny was based on sheet 5. The libraries used for this was leaflet, plotly, shiny, dplyr, tidyverse,ggplot2 and shinythemes.

My implementation included creating a shiny web app with 5 tabs as discussed in design sheet 5. The first tab shows the overview of the study. The second tab onwards we have our plots.

However, from my design sheets there are few differences that I would like to highlight here:

Firstly, The tab that includes the map is not a choropleth map instead a map with circle markers of red and green, red showing that the particular location has high rate of covid cases, while green showing that the location has low rate of covid cases. I choose this effect because when trying to show the map with population slider the data was not being perfectly represented as I wanted and hence I used a select input slider for the user to choose a particular location and see whether that location had a low rate of covid case or high rate. This change was done keeping in my mind the story that I wanted to depict.

There was also a change in the bar graph that I originally intended to show with a population slider. Here instead of a population slider I used a select input slider of location. By doing this the intended audience can see the change in GDP of a country through 12 months individually and also understand that for each location the GDP decreased evidently.

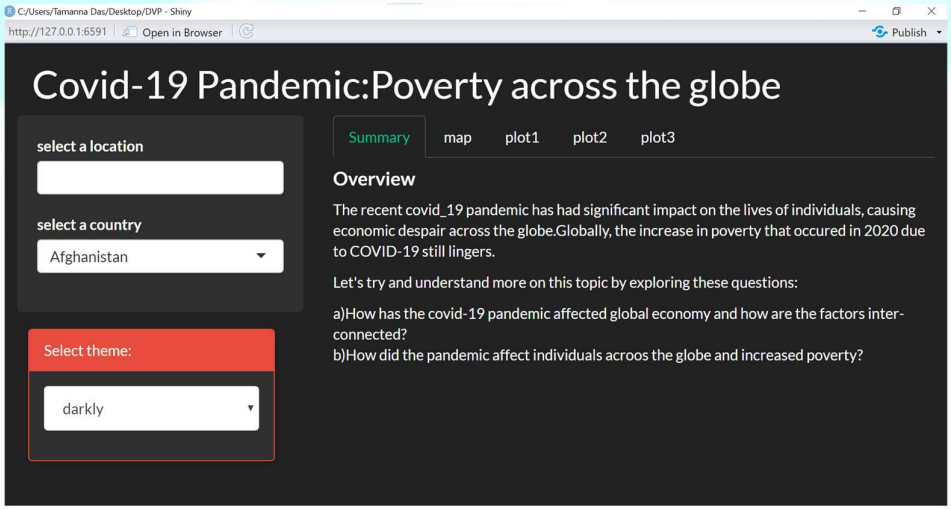
Also, instead of showing an area chart of % of poor I decided to show a bar chart of % of working hours lost during 2020 showcasing and proving my point that the pandemic actually increased poverty across the globe by decreasing the working hours this was also included with a tooltip.

My entire implementation was done in R studio using the libraries already discussed above. For creating a leaflet map with circle markers and putting a select input slider for location. I had to first clean my data that contained country names and total cases and in that I had to also add a column of longitude and latitude by making a left join and then I implemented that data in leaflet to create the required visualization. Also for my employment data I had to perform a split the entire data of employment that was previously used in my exploration process to take out the only percentage working hours lost.

My data visualization process included data wrangling and data modification of my previously used data of exploration process as it had few missing columns which I had to add before starting my visualization and also because my datasets were very huge so, it took considerable amount of time. I had to make advanced used of R studio as I used different kinds of visualization for each plot. For plotting my bar graph which changes as the user inputs a particular location I had to re implement my exploration process and group the data first based on location and summarise the data.

I started of by creating a initial theme or layout and then used tabs for my plots and plotted my graphs in there including different interactivity.

I also used a select theme method wherein the user can change the theme of the web page according to their likeliness.



This is how my final shiny web page looks like.

USER GUIDE

1. The first select input is for the map that is the user can choose multiple locations at one time and see the intensity of covid case that is whether it is high or low based on the legend given.**(fig 1)**
2. The second select input is for plot 3 where in the GDP per month changes with change in location. **(fig 1)**
3. The user can also hover on Plot2 and Plot4 to see the stats of the data.
4. The select theme input is for the user to change theme of the web page based on their choice. **(fig 1)**
5. If a popup is displayed saying updating loaded packages just click on cancel and the app will run smoothly as seen in **fig 2**
6. When importing data please choose heading as **Yes** as seen in **fig 3**

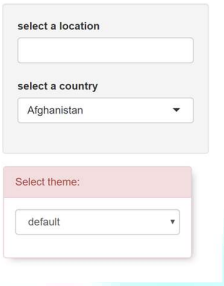


Fig1

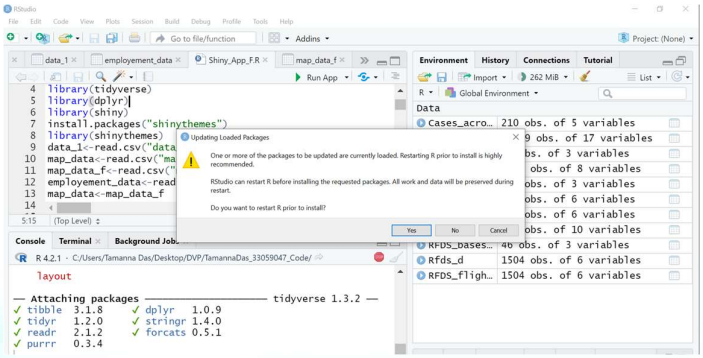


Fig 2

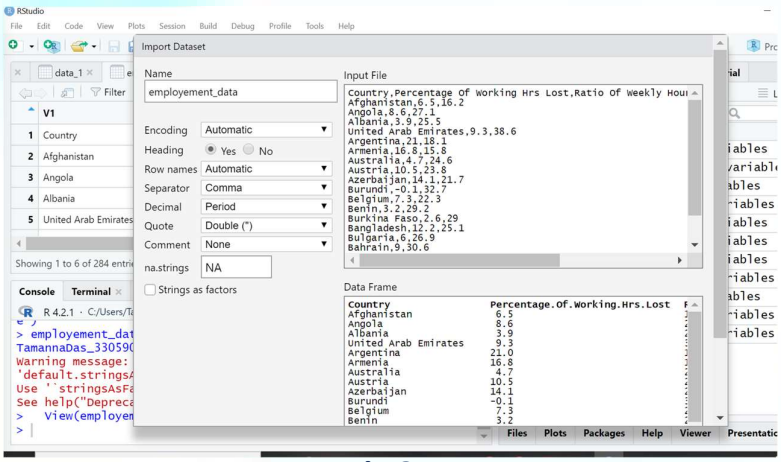


Fig 3

CONCLUSION

The above visualization have led me to prove my point that the covid-19 pandemic has led to increase in poverty across the globe.

Plot 1tab shows us that the GDP and HDI are interlinked with each other and countries with high population has comparatively less HDI and GDP from countries with low population. It also shows a positive correlation amongst GDP and HDI. Plot 3 shows us that from month 1 to month 12 of 2020 the GDP of most of the countries depreciated leading to decrease in global economy and finally we can see from plot 4 that the working hours were lost during covid 19 across the globe indicating increase in poverty.

All these helped me prove my story.

From this project I was able to learn and enhance my skills of implementation of R shiny in a better way. Through this data visualization project, I learnt various new methods of wrangling and cleaning the data. Interconnecting various visualisations from different data sets and answering all the questions.

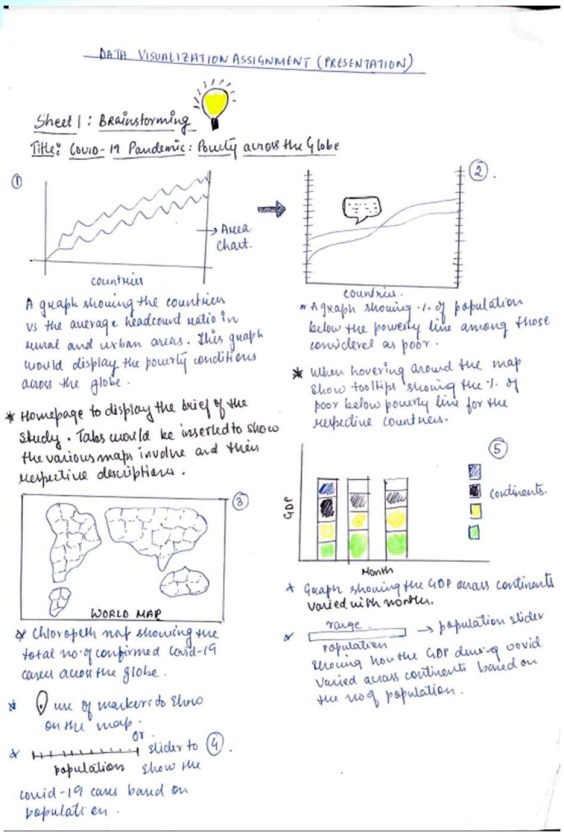
However, I think if I could get more insightful data then I could have more deeply explored and finally visualized my story in a more better way.

BIBLIOGRAPHY

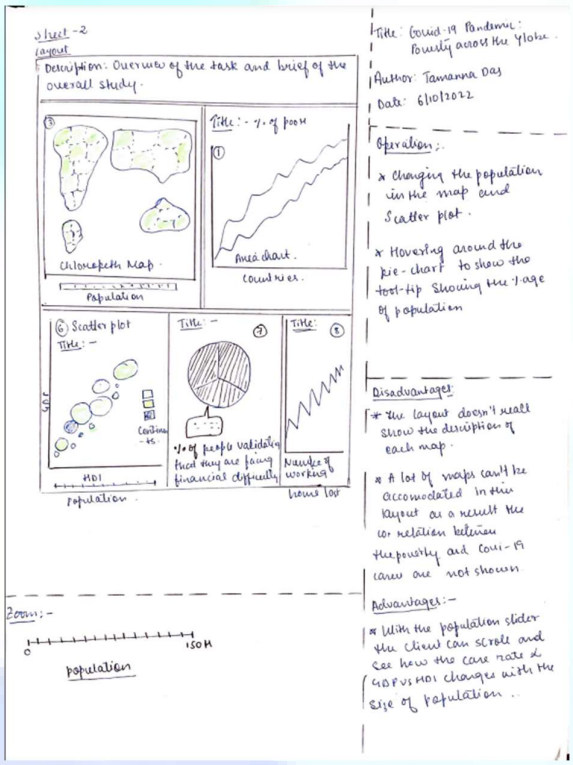
1. https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-022-13698-5- article supporting my visualisation.
2. https://www.rstudio.com/wp-content/uploads/2015/02/data-wrangling-cheatsheet.pdf - use of data wrangling cheat sheet
3. Used R Shiny studio to see various templates - https://shiny.rstudio.com/gallery/
4. Used R studio git hub to see how do we implement the select input method in R shiny - https://github.com/rstudio/rstudio

APPENDIX

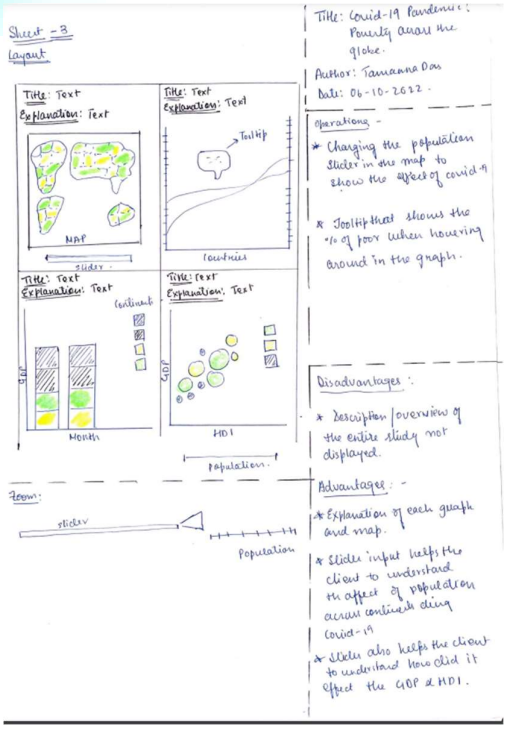
Brain storming sheet:



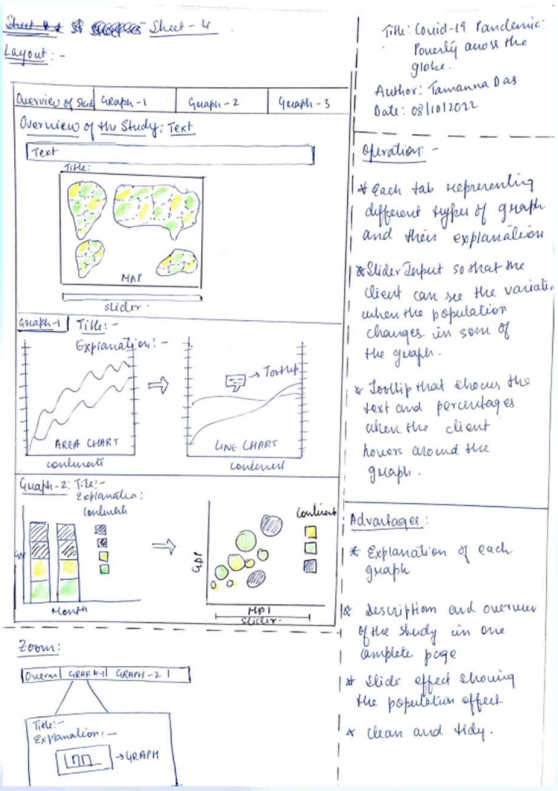
SHEET 2:



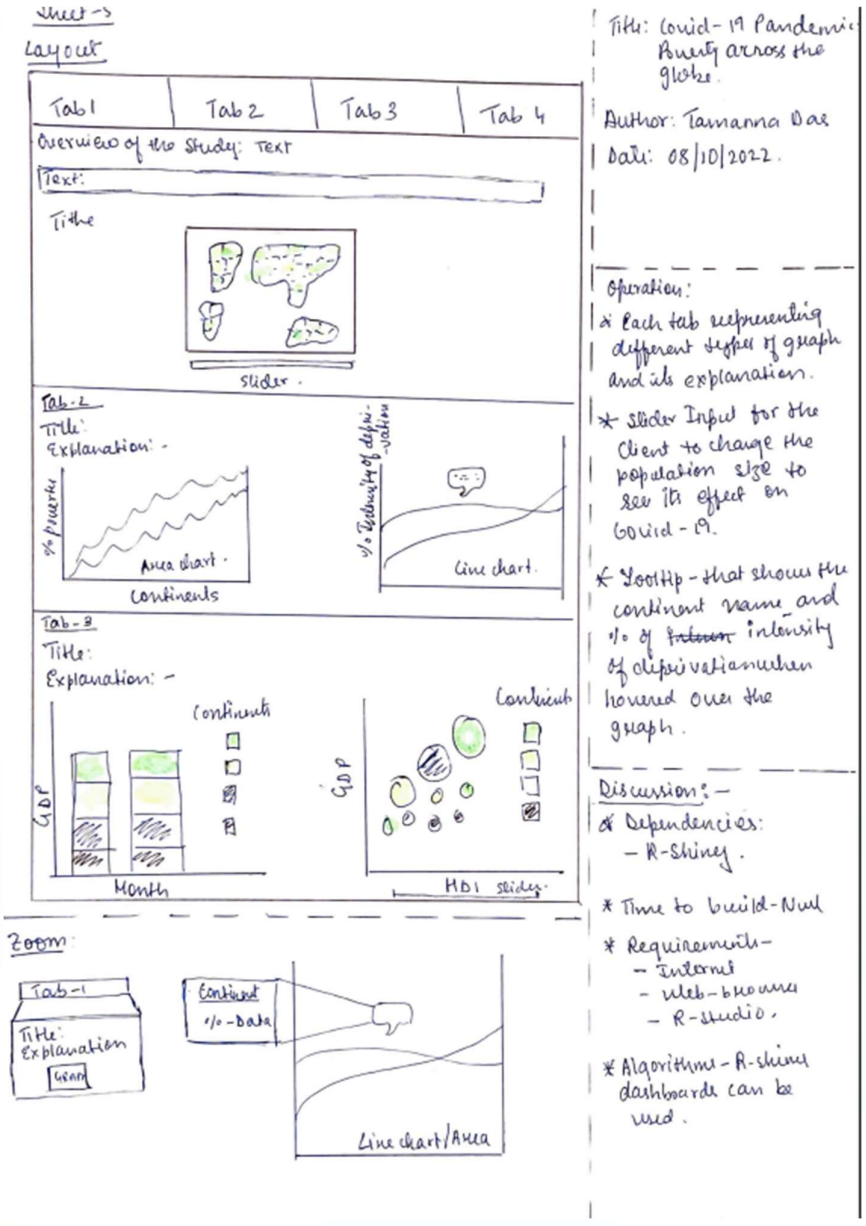
SHEET 3:



SHEET 4:



SHEET 5:


Sheet-5
Layout
Tab 1
Tab 2
Tab 3
Tab 4
Overview of the Study: Text
Text:
Title
Slider.
Tab-2
Title:
Explanation: -
% Poverty
Area chart.
Continents
% Intensity of deprivation
Line chart.
Tab-3
Title:
Explanation: -
Continents
GDP
Month
Continents
GDP
HDI slider.
Zoom:
Tab-1
Title:
Explanation
Continent
% - Data
Line chart/Area
Title: Covid-19 Pandemic Poverty across the globe.
Author: Tamanna Das
Date: 08/10/2022.
Operation:
* Each tab representing different types of graph and its explanation.
* Slider Input for the Client to change the population size to see its effect on Covid-19.
* Tooltip - that shows the continent name and % of intensity of deprivation when hovered over the graph.
Discussion: -
* Dependencies:
- R-Shiny.
* Time to build - Null
* Requirements -
- Internet
- Web-browser
- R-studio.
* Algorithms - R-shiny dashboards can be used.