

Data Visualisation- ‘The Wealth of Nations’ Report

Aim: Providing an analysis of ‘The Wealth of Nations’ dataset and use Tableau to create a data visualisation to give an insight of World GDP nowadays.

Hoi Ning Mak
(GLA 4 Data)



Contents

First Task– Policies and Procedures.....	2
Second Task– Excel.....	3
Third Task– Tableau.....	7
Summary Dashboard.....	12
Evaluation.....	13

First Task- Policies and Procedures

'The Wealth of Nations' dataset is structured data as it is presented in an Excel file. It did not include personal or sensitive data and has been collected through surveys anonymously.

Therefore it is classified as **open data** which can be freely used and redistributed by anyone. When using this data, we need to adhere to the **Computer Misuse Act 1990** because it can prevent unlawful access like planting a virus and misuse of information systems. The act can also deal with the issue of modifying data without permission when everyone has access to open data. A user who commits this policy might need to face up to a ten-year prison sentence or an unlimited fine.

Furthermore, the **GDPR** (General Data Protection Regulation) can be applied to open data that aims to protect the rights of individuals and guard against malicious data. Thus, asking for consent before is required in the collection of data.

Lastly, the **Data Protection Act 2018** can also be implemented, where any person who collects, processes or uses personal data should collect and process data fairly and lawfully. It is crucial to ensure transparency and participation in the data. By imposing the policies above, they can help build trust from the public in how data is collected and used.

Working as a data analyst, it is important to comply with these acts because they might work on confidential data (like sensitive data and personal data) that require storing in a secure place while interpreting and analysing them. For this reason, they should also have an ethical responsibility to protect customers' data, defend against breaches and ensure that personal data is not compromised. Hence, data analysts should consider privacy, transparency, regulation and fairness when reviewing data.

Second Task- Excel

I will use the ‘The Wealth of Nations’ data to demonstrate the following Excel tasks.

The screenshot shows a Microsoft Excel spreadsheet titled 'The Wealth of nations - Saved'. The data consists of a table with columns: Rank, Country, GDP - per capita (PPP), and Year of Information. Row 3 is selected, highlighting Liechtenstein. A context menu is open over the data, with the 'Protect Sheet' option being selected. This opens the 'Manage Protection' dialog box on the right side of the screen. The dialog box allows users to add ranges to be unlocked while protection is on, set a sheet protection password, and confirm it. Buttons for 'Save' and 'Cancel' are at the bottom.

(Figure 1)

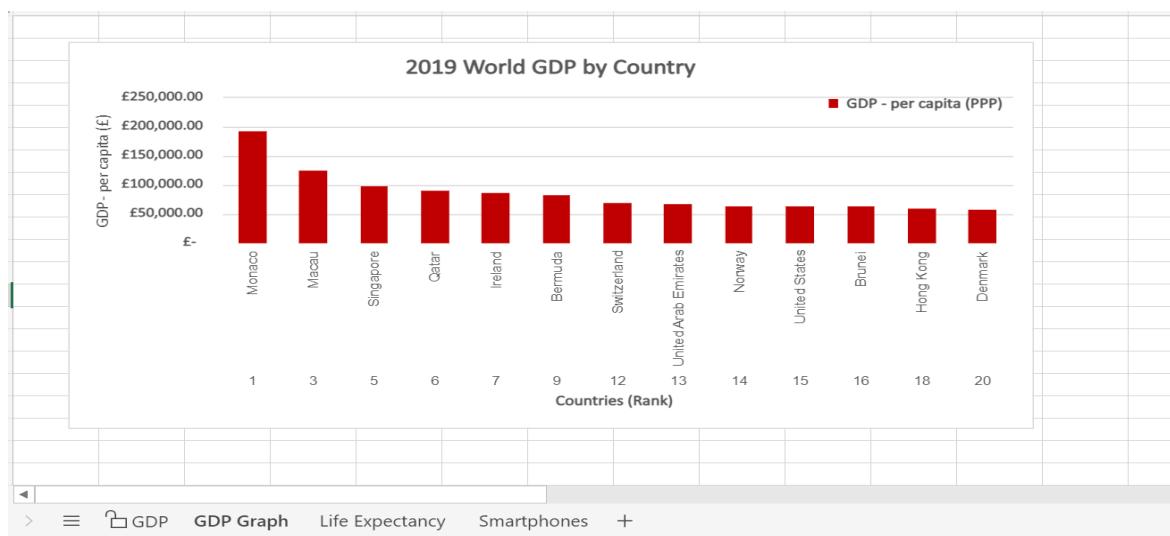
Set a password as ‘GDP’ to protect the workbook so that this can prevent others from accessing data without permission.

The screenshot shows the same Excel spreadsheet as Figure 1. The 'Sort & Filter' dropdown menu is open, specifically the 'Number Filters' section under 'Clear Filter from "Year of Information"'. It lists years from 2014 to 2020, with '2019' checked. Below the menu, the data table is shown with only the 2019 entries visible. The currency symbol in the GDP column has been changed to the British Pound (£).

(Figure 2.1)

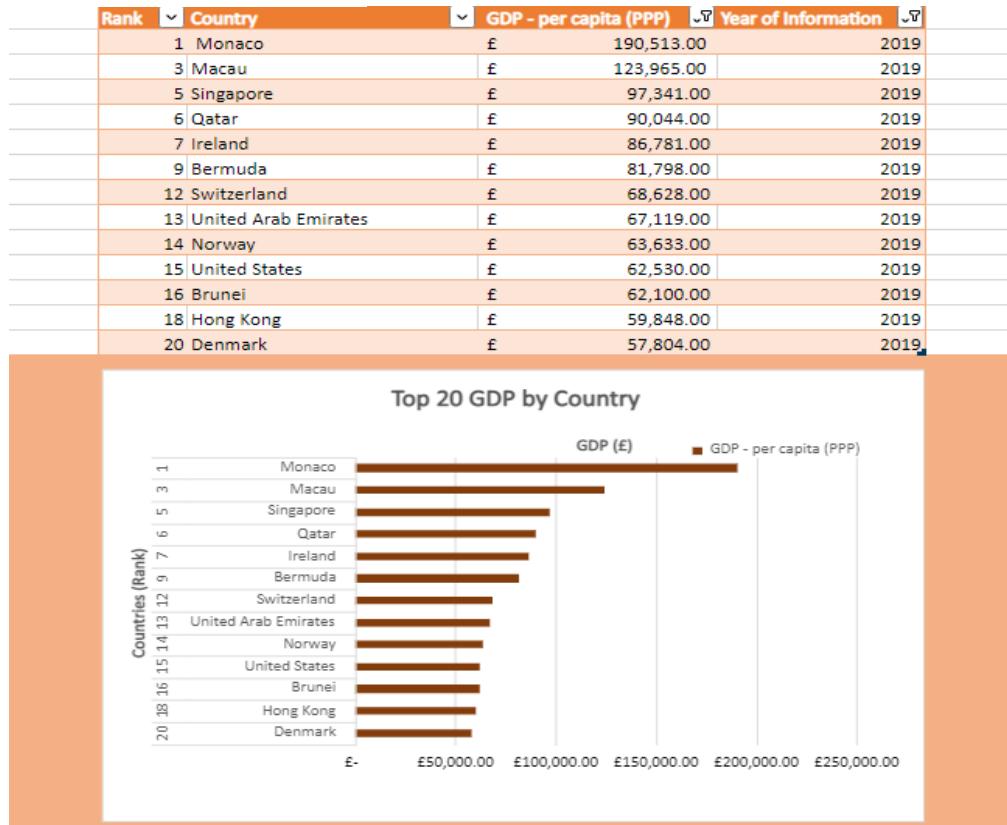
The GDP data turned into a table. The currency sign is changed to the British Pound symbol and filtered year to 2019 (Figure 2.2) so it will only display information in that year.

(Figure 2.2)



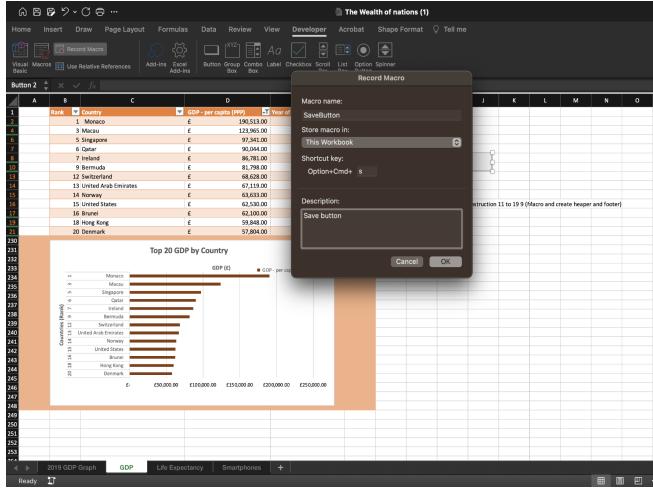
(Figure 3)

Using the above-filtered data, the '2019 World GDP by the Country' bar chart was created in a new sheet tab 'GDP Graph'. The rank in countries was the x-axis and GDP- per capita was the y-axis.

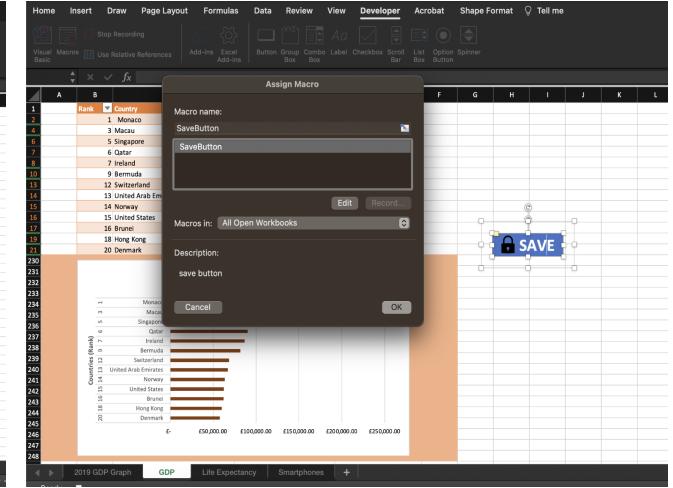


(Figure 4)

The data was sorted for the top 20 highest-ranking countries based on the GDP- per capita. A horizontal bar chart was created in descending order.

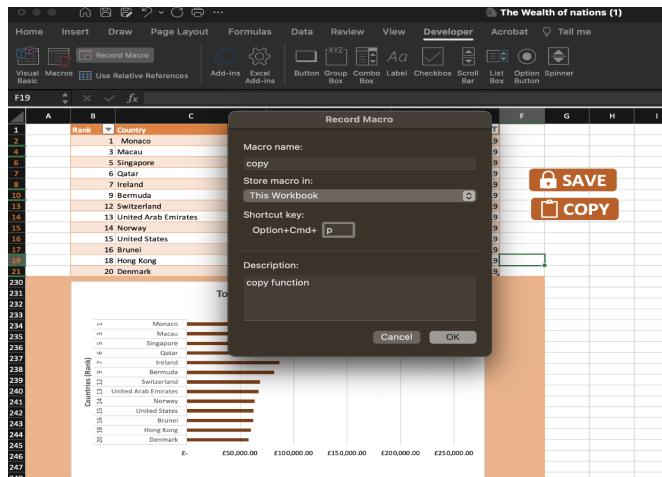


(Figure 5)

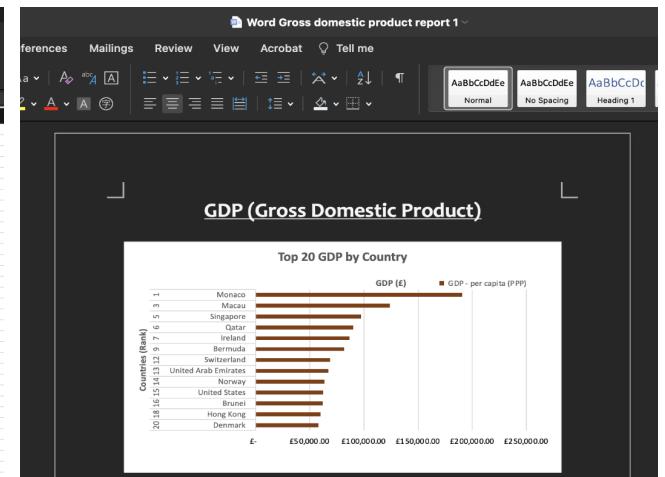


(Figure 6)

I assigned the macro as a 'save' function to the above button so that the spreadsheet will save the work after pressing the button.

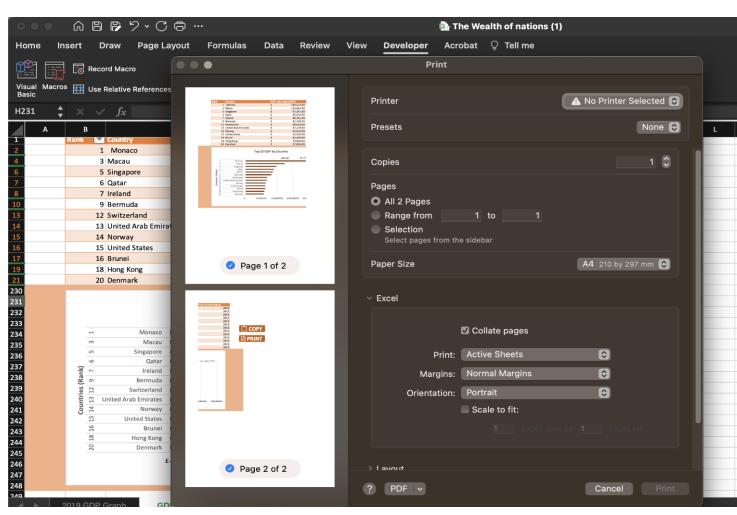
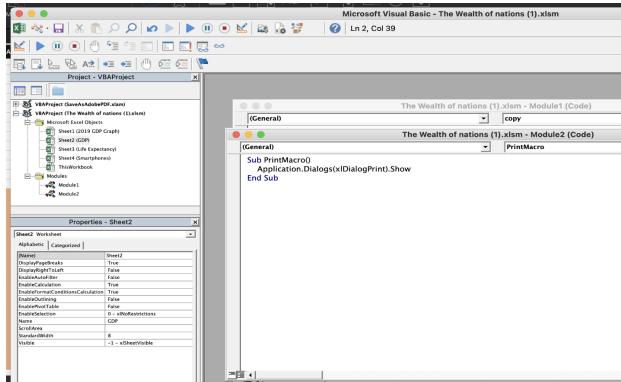
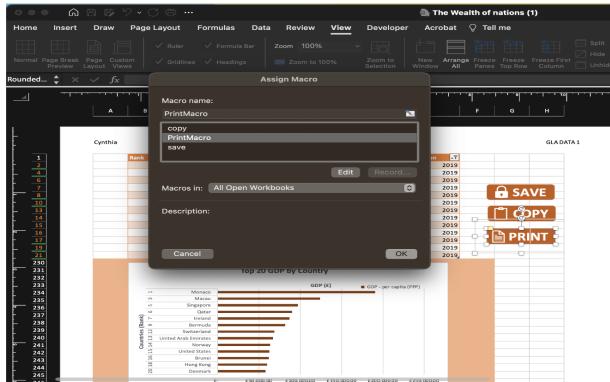


(Figure 7)



(Figure 8)

I highlighted the area to be copied and recorded the action of right-click copy. Then, assigned the macro to the copy button and the macro button was created. Figure 8 shows the area that I copy from Excel- GDP tab to a Word document name 'Word Gross domestic product report 1'.



I used VBA Macros (Figure 10) to create a print function, so the whole page will print when clicking on the 'print' button.

Cynthia	Mak	GLA DATA 1								
<table border="1"> <thead> <tr> <th>Rank</th> <th>Country</th> <th>GDP - per capita (PPP)</th> <th>Year of Information</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Monaco</td> <td>£ 190,513.00</td> <td>2019</td> </tr> </tbody> </table>	Rank	Country	GDP - per capita (PPP)	Year of Information	1	Monaco	£ 190,513.00	2019		
Rank	Country	GDP - per capita (PPP)	Year of Information							
1	Monaco	£ 190,513.00	2019							

(Figure 12)

Excel header and footer added as requested.

14/2/2023

Assignment 1

Data Visualisation

(Figure 13)

Third Task- Tableau

'The Wealth of Nations' data is used to analyse and conduct the result in a dashboard using Tableau.

▫ GDP+ (The Wealth of nations)



(Figure 14)

First, I imported the Excel table 'The Wealth of Nations' into Tableau and set the relationship. As the common column for all of them is the country, they all could be equal.

GDP		4 fields 228 rows				100	→	rows
Name	Type	#	GDP	#	GDP	#	GDP	Year of Information
Fields	Type	Field Name	Physical	Phys...	Rem...	Rank	Country	GDP - per capita (PPP)
						1	Monaco	190,513.00
						2	Liechtenstein	180,367.00
						3	Macau	123,965.00
						4	Luxembourg	115,874.00

(Figure 15)

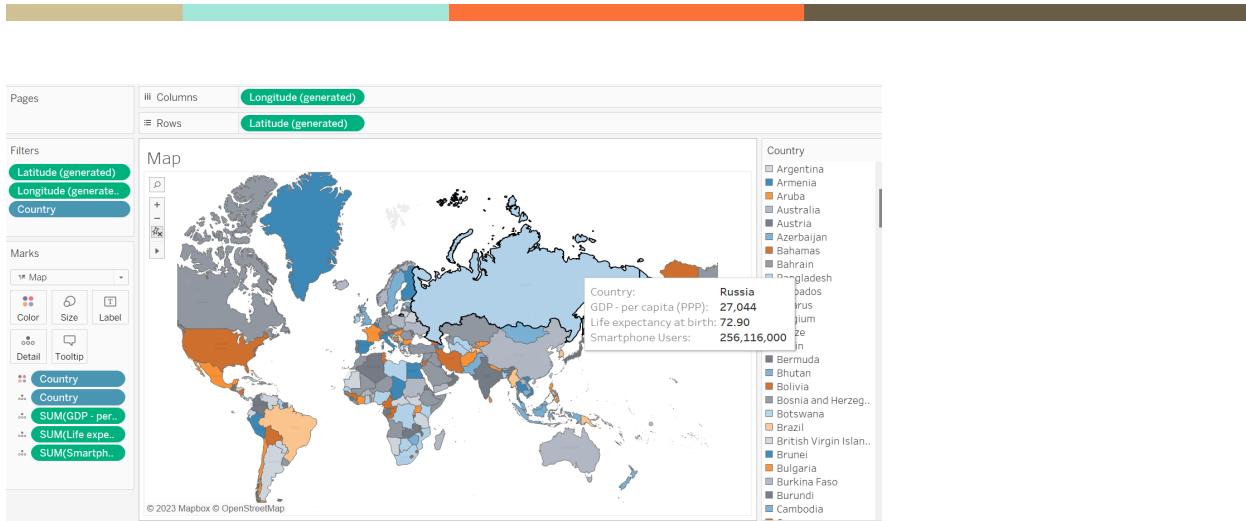
Life Expectancy		4 fields 196 rows				100	→	rows
Name	Type	#	Life Expectancy	#	Life Expectancy	#	Life Expectancy	Date of Inform
Fields	Type	Field Name	Physical	Phys...	Rem...	Rank	Rank (Life Expectancy)	Country (Life Expectanc...
						1	Hong Kong	84.9000
						2	Japan	84.6000
						3	Switzerland	83.8000
						4	Singapore	83.6000

(Figure 16)

Smartphones		4 fields 74 rows				74	→	rows
Name	Type	#	Smartphones	#	Smartphones	#	Smartphones	Date of Information (S...
Fields	Type	Field Name	Physical	Phys...	Rem...	Rank	Rank (Smartphones)	Country (Smartphones)
						1	China	1,598,360.000
						2	India	1,281,971,713
						3	Indonesia	385,573,398
						4	United States	327,577,529

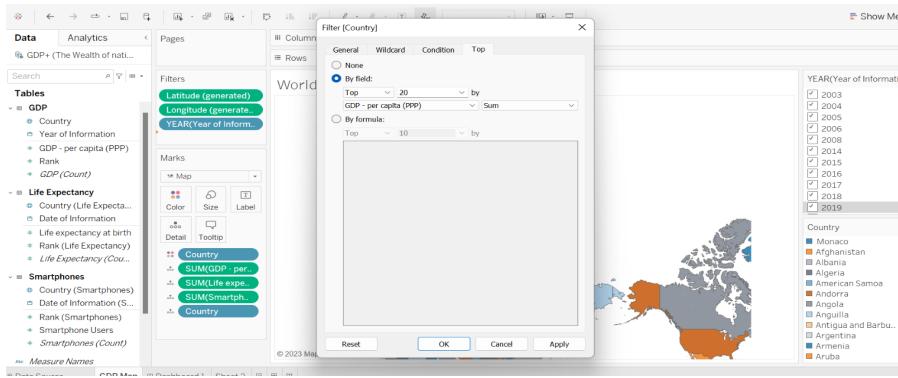
(Figure 17)

Data types are checked before building data visualisation.



(Figure 18)

Map was created based on the GDP data. In the tooltip, the sum of GDP, life expectancy and smartphone users per country is shown.



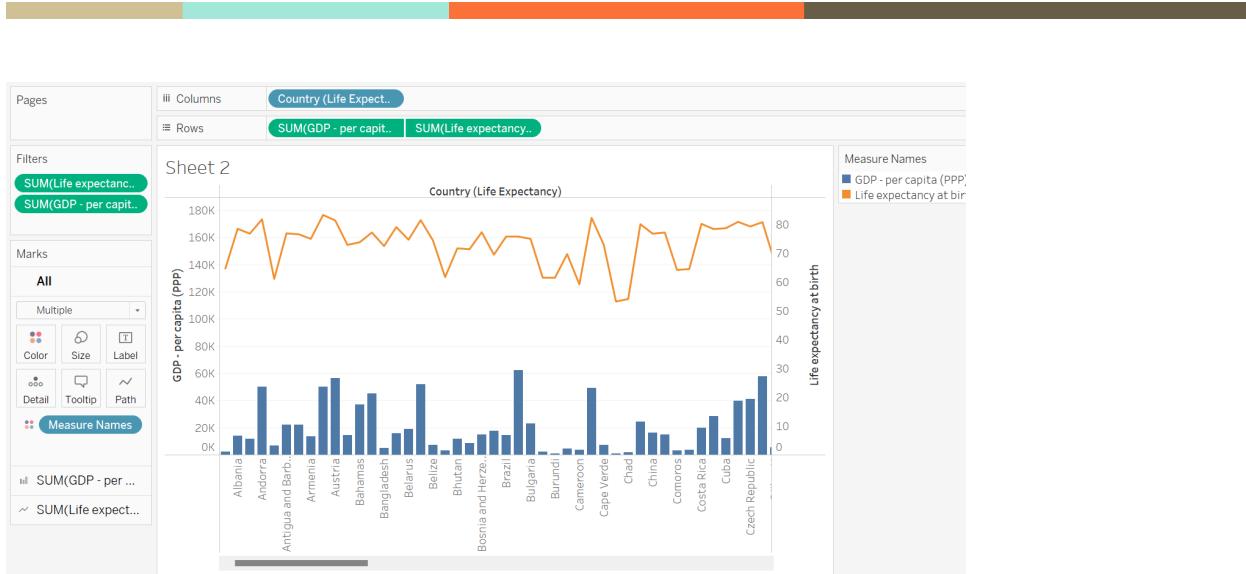
(Figure 19)

Since the client is only interested in the top 20 highest-ranking countries so the map was filtered based on the sum of GDP.



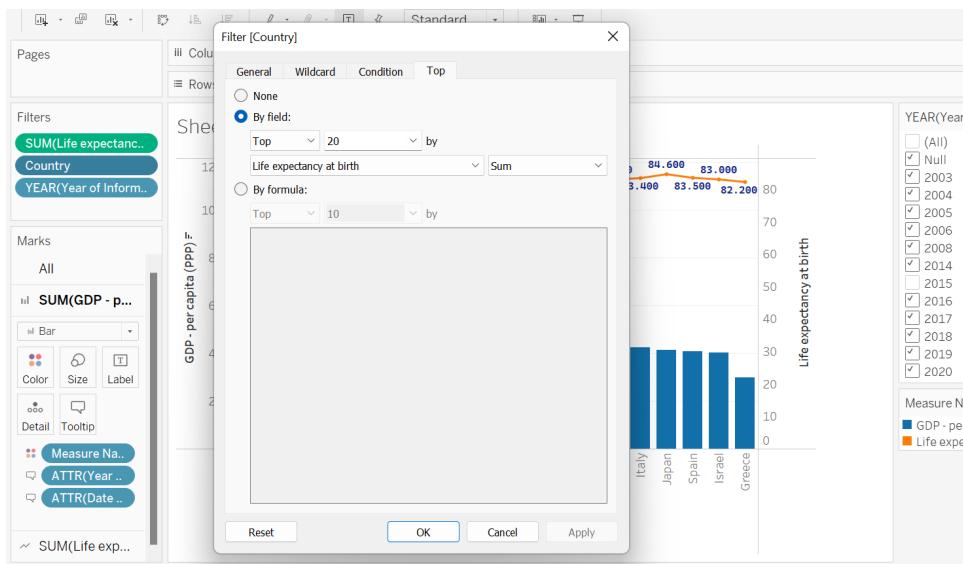
(Figure 20)

I changed the shape into a circle, only the label of the highest and the least sum of GDP will be shown on the map to keep the map clear.



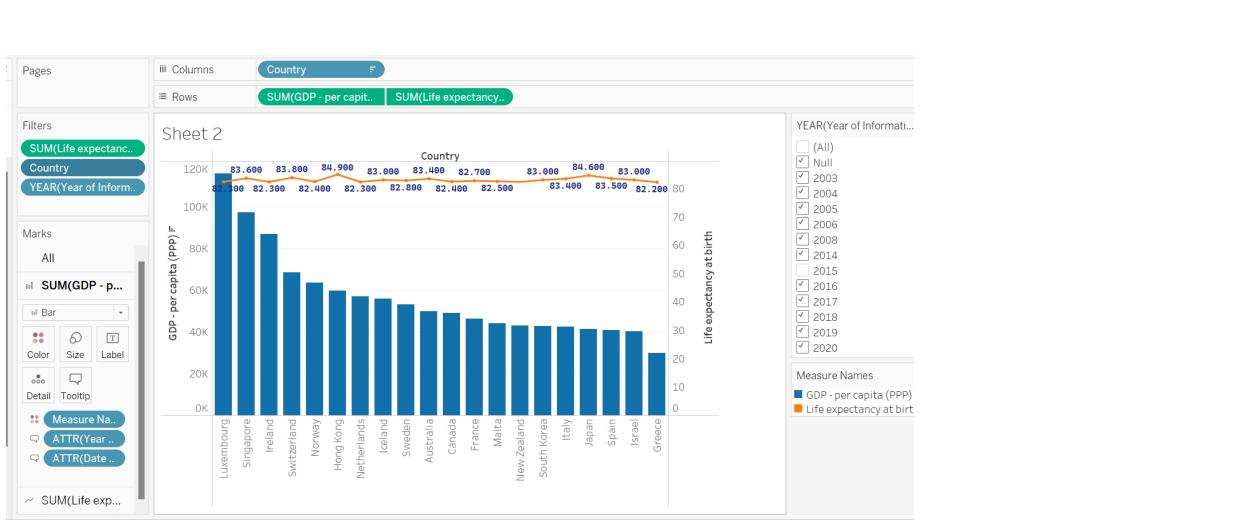
(Figure 21)

A dual-axis chart (line graph and bar chart) is used to show the relationship between GDP and life expectancy.



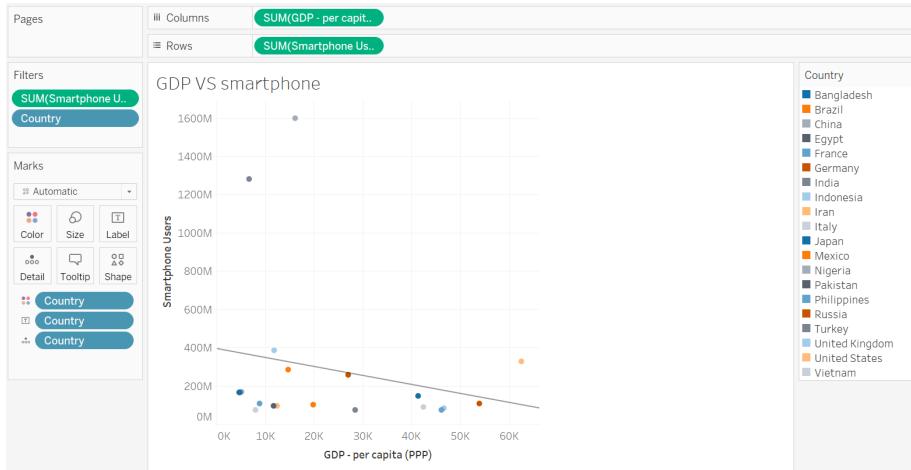
(Figure 22)

I used the country in the life expectancy sheet to filter the top 20 countries.

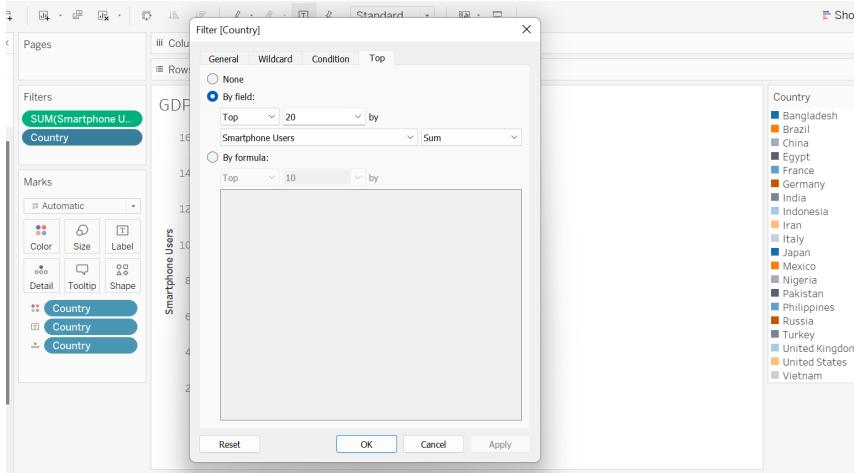


(Figure 23)

The line graph of life expectancy is labelled so that is easier to read and the chart is sorted in descending order.

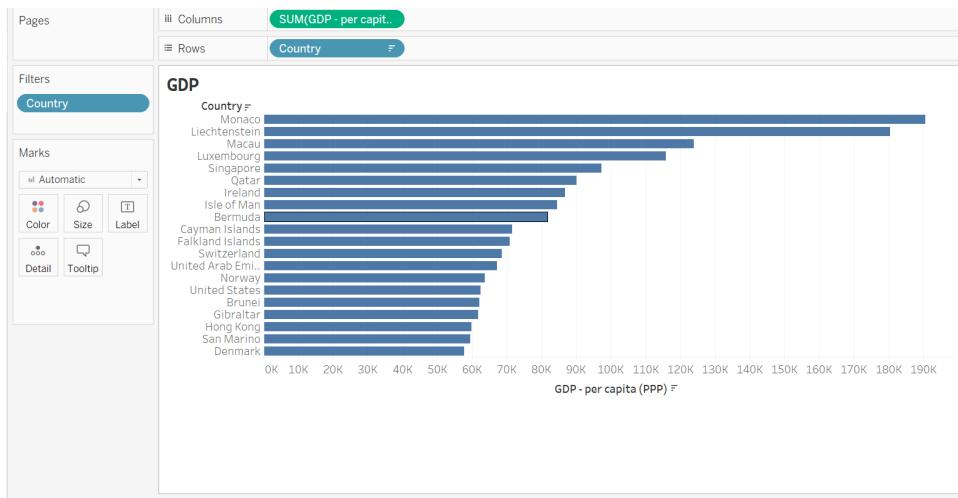


(Figure 24)

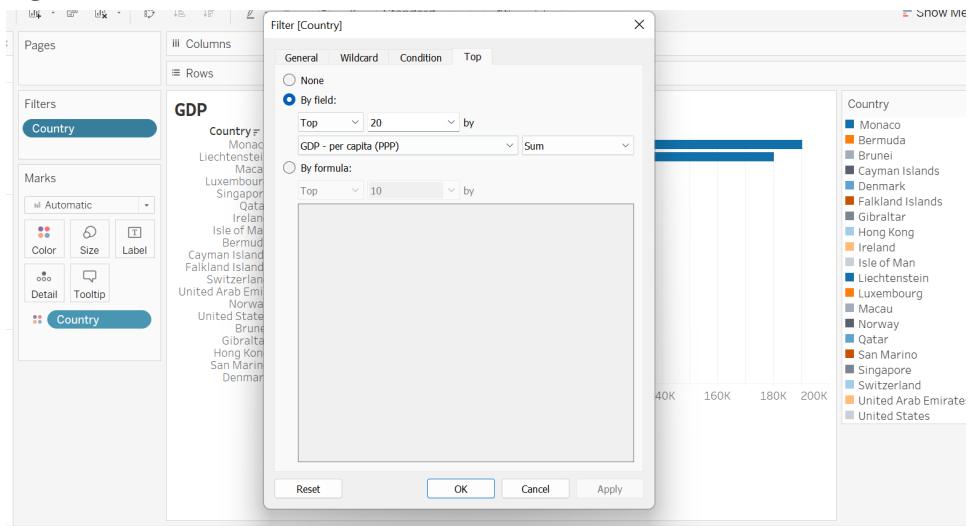


(Figure 25)

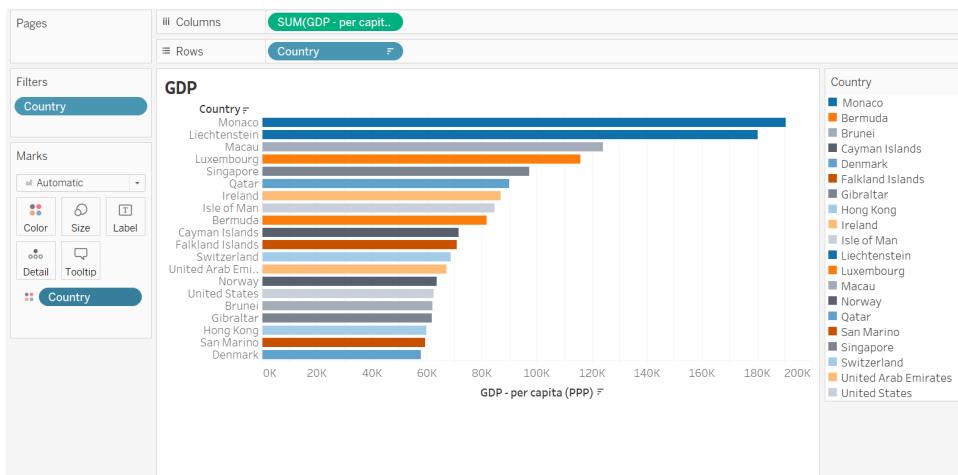
A scatter plot is used to see the correlation between GDP and smartphone users. The top 20 countries are also filtered based on the sum of smartphone users.



(Figure 26)

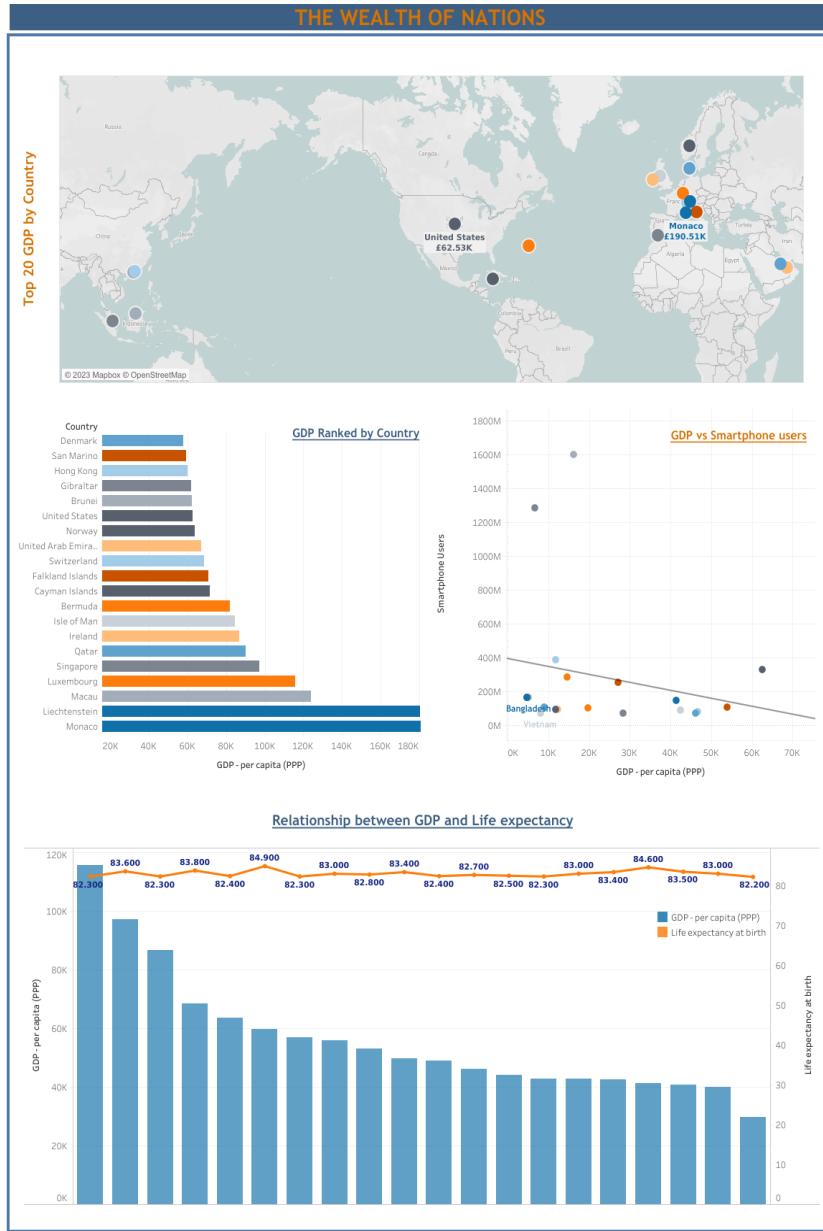


(Figure 27)



(Figure 28)

A horizontal bar chart is created to show the top 20 GDP countries ranked in descending order.



(Figure 29)

Link to the dashboard:

https://public.tableau.com/views/thewealthofnation/Dashboard1?:language=en-GB&:display_count=n&:origin=viz_share_link

Summary Dashboard

To satisfy the client's requirement, the dashboard (Figure 29) is designed for colour blindness and a colour-blind palette was used. Also, 'The Wealth of Nations' is a big dataset which is less likely to present all the counties in one dashboard. Therefore, all visuals are presented in the top 20 highest-ranking countries only.



To conclude the dataset, when we look at the map and bar graph, Monaco has the highest GDP. In the scatter plot, there are no correlations between GDP and smartphone user which suggest the population isn't a factor to be considered when determining the GDP. Whereas, life expectancy at birth can affect the GDP as the top 20 countries by GDP all have a life expectancy above 80.

Evaluation

Overall, the thing that went well in this project was I managed to meet the assignment requirements by following the guidance step-by-step to deliver my work to a good standard. Hence, to make my project clear, I add screenshots with captions for each task so that is easier to catch and keep the reader's attention. I also make further research to learn how to create VBA macros in Excel and charts in Tableau, this helps me to consolidate my technical knowledge.

However, to deliver my project to an outstanding standard next time, I should have better time management by planning a schedule for it. So that I can spend more time on self-study to learn additional techniques, such as creating a dashboard using coding and creating a calculated field to make the analysis more flexible in Tableau. Further, I can set additional objectives for the project to show my ability to go beyond.

After completing the project, I knew the importance to think of how GDPR affect every dataset and the type of policies that need to be adhered to while using open and closed data. Since data privacy fascinated me so I make further research and make suggestions on acts that need to adhere to when handling this dataset. As a data analyst, it is important to consider privacy, transparency, regulation and fairness when reviewing data time.