Apprenticeships | Training | Recruitment

Excel and Tableau Assignment

Suzan Tesfai Yemane

2024-10-24

Contents

1 Data policies and procedures

1.1 Data policies

A data policy is a collection of rules, guidelines, instructions, and preferences designed to ensure that data is collected, stored, accessed, and managed in accordance with the standards set by the governance framework or data rights holders. These policies ensure data accuracy, consistency, and security, while also helping the organization comply with applicable regulations. Additionally, they improve data quality and support better decision-making. A data policy is established by the data rights holder to control access to their shared data within a data space.[1] Data policies typically include:

- Roles and responsibilities for handling data.

- Protocols for data collection, storage, processing, and disposal.

- Standards for data quality, interoperability, and protection.

Before using data, you should consider many things, including [2]:

1. Data quality: Make sure the data is accurate and complete, and check for missing, incorrect, or outlier values.

2. Data sources: Use multiple sources to get a complete picture, and keep track of where the data came from.

3. Data security: Store data securely to prevent loss or corruption, such as by backing it up and using secure storage devices.

4. Data types: Identify the types of data you need, such as primary data collected directly from people, or secondary data from existing reports or research studies.

5. Data preparation: Consider whether you need to change the data, such as to clean it, standardize it, or group it in new ways.

1.2 Data procedures

The data analysis process involves several steps, including [3]:

- Defining objectives: Define the questions and objectives for the analysis.

- Data collection: Use accepted data collection techniques to protect the credibility and reliability of your data.

- Data cleaning: Clean the data, also known as data wrangling, to ensure accurate results.

- Data analysis: Use data analysis tools to find patterns, relationships, and trends

- Data interpretation and visualization: Interpret and visualize the data

- Data storytelling: Tell the story of the data

2 Excel

In this assignment, the 'Wealth of Nations' Excel file is used for illustration. Overview of the Excel sheet see Fig1. .

Rank	Country	GDP - per capita (PPP)	Year of Information
1	Monaco	$190,513.00	2019
2	Liechtenstein	$180,367.00	2018
3	Macau	$123,965.00	2019
4	Luxembourg	$115,874.00	2020
5	Singapore	$97,341.00	2019
6	Qatar	$90,044.00	2019
7	Ireland	$86,781.00	2019
8	Isle of Man	$84,600.00	2014
9	Bermuda	$81,798.00	2019
10	Cayman Islands	$71,549.00	2018
11	Falkland Islands	$70,800.00	2015
12	Switzerland	$68,628.00	2019
13	United Arab Emirates	$67,119.00	2019
14	Norway	$63,633.00	2019
15	United States	$62,530.00	2019
16	Brunei	$62,100.00	2019
17	Gibraltar	$61,700.00	2014
18	Hong Kong	$59,848.00	2019
19	San Marino	$59,439.00	2018
20	Denmark	$57,804.00	2019
21	Netherlands	$56,935.00	2019
22	Jersey	$56,600.00	2016
23	Austria	$56,188.00	2019
24	Iceland	$55,874.00	2019
25	Germany	$53,919.00	2019
26	Sweden	$53,240.00	2019
27	Guernsey	$52,500.00	2014
28	Belgium	$51,934.00	2019
29	Taiwan	$50,500.00	2017
30	Andorra	$49,900.00	2015

GDP   Life Expectancy   Smartphones   +

Figure 1: Overview of Wealth of Nations Excel file.

2.1   Securing workbook

You can set a password to protect your worksheet and ensure that only you or selected individuals can access and modify it. This will secure the worksheet, allowing only authorized users to make changes see Fig2.
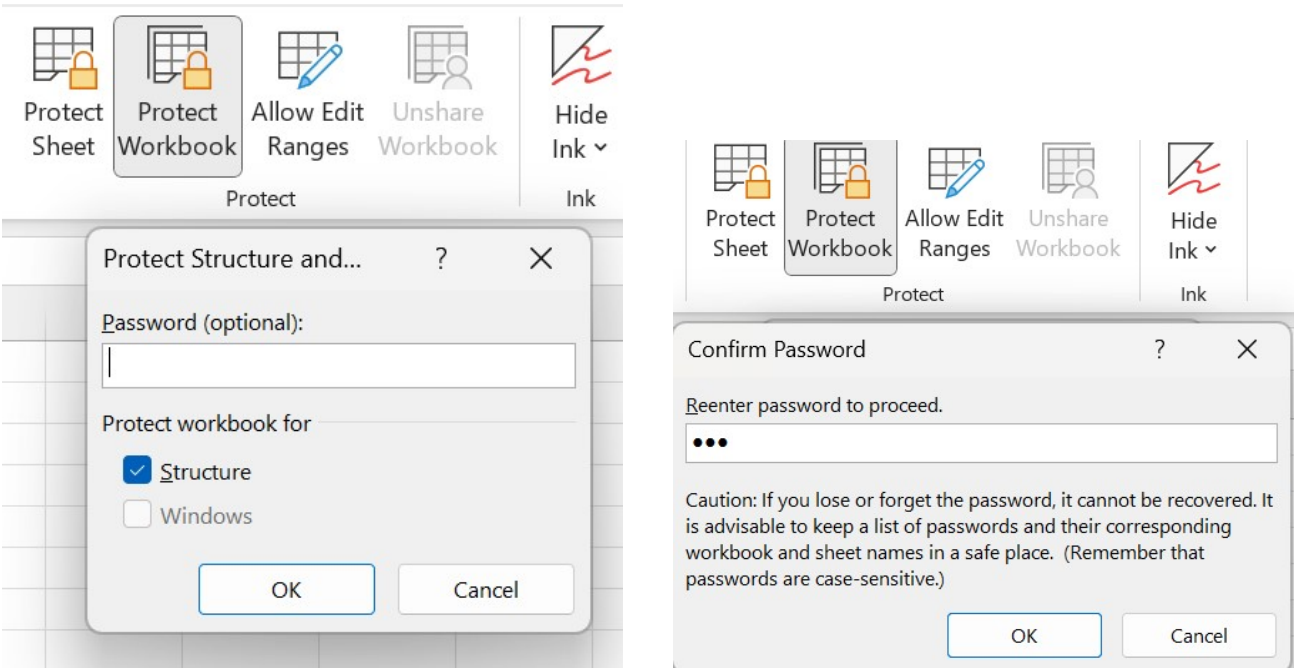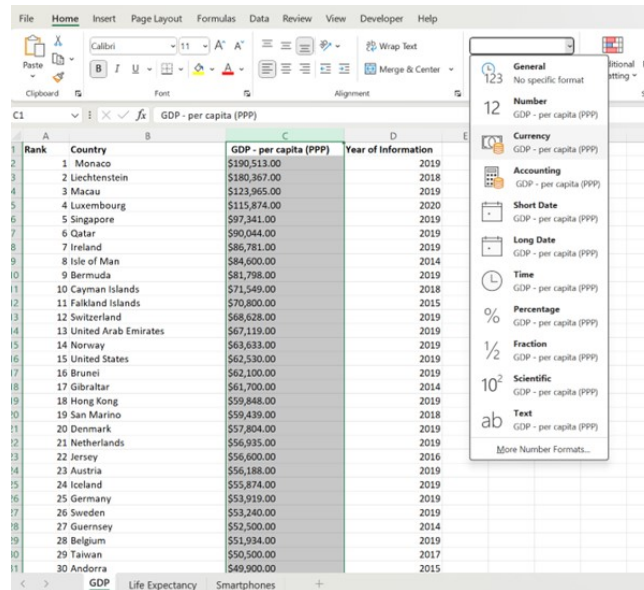

Figure 2: Protecting workbook.

2.2   Changing currency symbol

Select the GDP column and apply the currency format to display the values in GBP. Fig3

Figure 3: Changing currency to British pound symbol.

2.3   Filtering the GDP table

The GDP sheet has been converted into a table and filtered to show data exclusively for the year 2019. Fig4
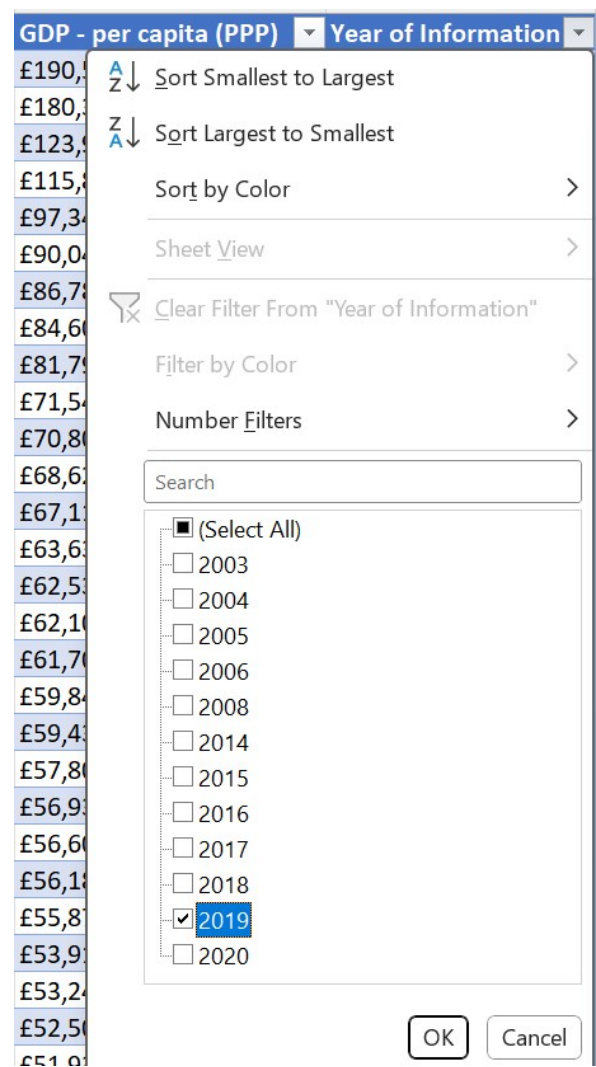
Figure 4: GDP- per capita (PPP).

2.3.1   Creating chart

Creating a chart using the data of âRank, Country and GDP - per capita (PPP). With the appropriate tile, X, and Y axis labels. Fig5
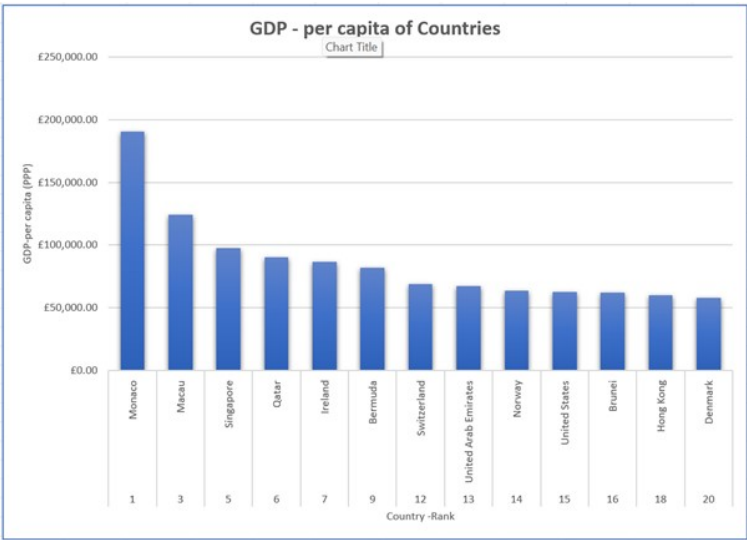

Figure 5: Chart Rank, Country and GDP-per capita (PPP).

2.4   Sorting by Rank

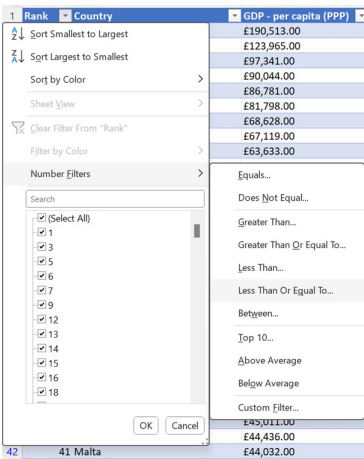Creating a Sort Filter to only display the top 20 countries.Fig6

Figure 6: Menu to filter top 20 countries.

2.5   Creating chart underneath the data with buttons

The below figure is for the top 20 countries of the 2019 year.

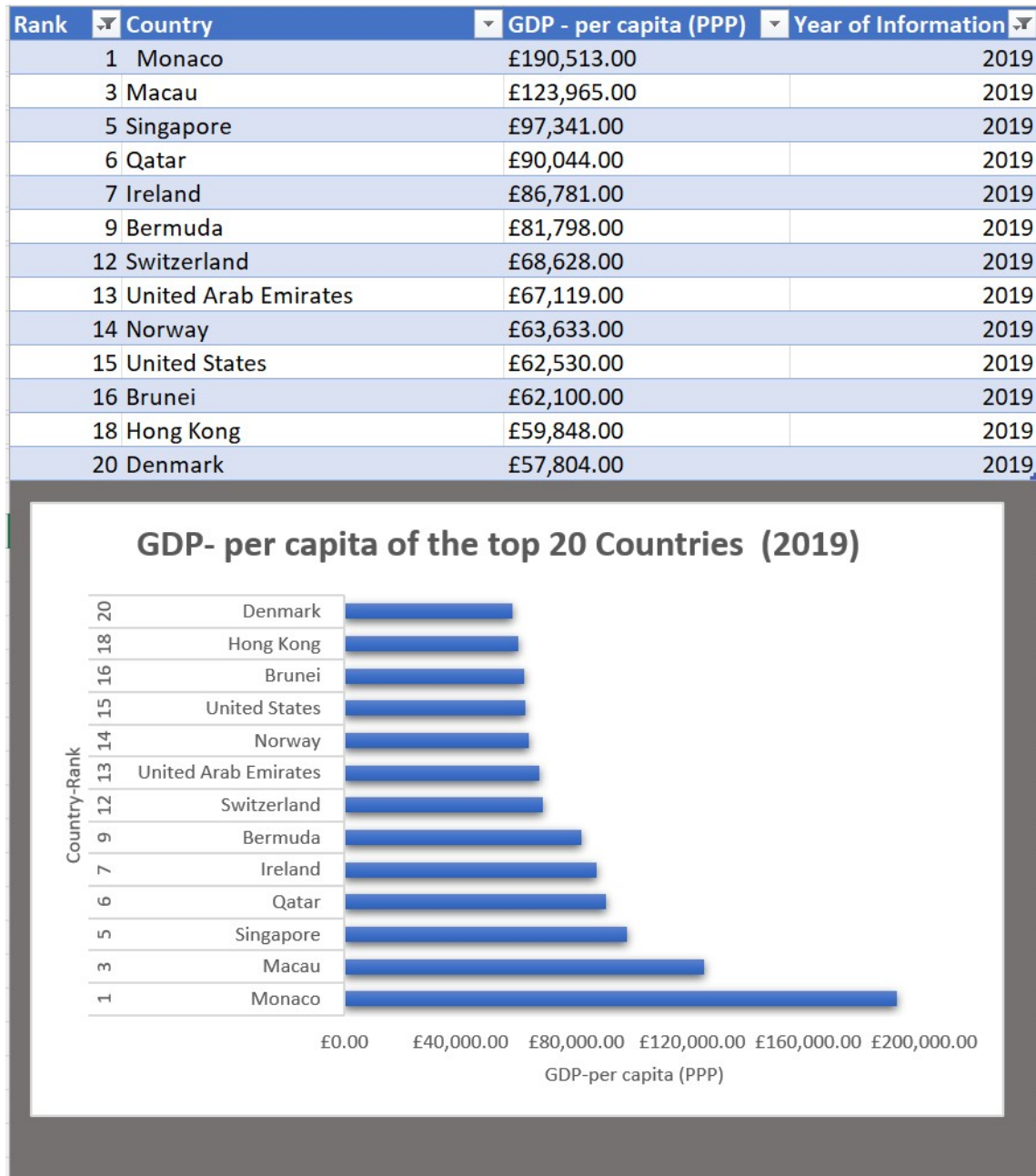| Rank	Country	GDP - per capita (PPP)	Year of Information
1	Monaco	£190,513.00	2019
3	Macau	£123,965.00	2019
5	Singapore	£97,341.00	2019
6	Qatar	£90,044.00	2019
7	Ireland	£86,781.00	2019
9	Bermuda	£81,798.00	2019
12	Switzerland	£68,628.00	2019
13	United Arab Emirates	£67,119.00	2019
14	Norway	£63,633.00	2019
15	United States	£62,530.00	2019
16	Brunei	£62,100.00	2019
18	Hong Kong	£59,848.00	2019
20	Denmark	£57,804.00	2019

Figure 7: Filtered table with chart top 20 countries (2019).

2.6   Macros

A macro in Excel is a set of actions that can be recorded and run to automate repetitive tasks.

2.7   Recording macros

To record a macro select the record macro under the Developer tab. Click the Record Macro button in the Code group see Fig8. Perform the actions you want to automate. Click the Stop Recording button when you are finished.
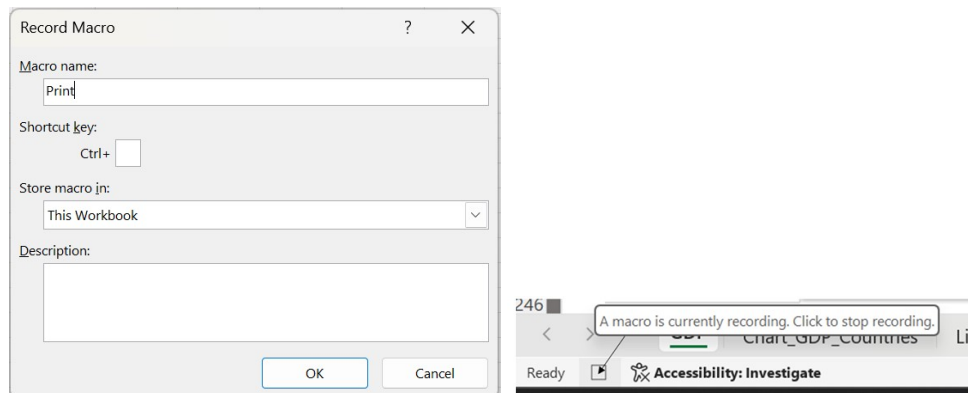

Figure 8: Creating macros.

After recording macros you can assign to the button. In this task, three buttons are first created see Fig 9, then the already saved macros are assigned to them respectively.
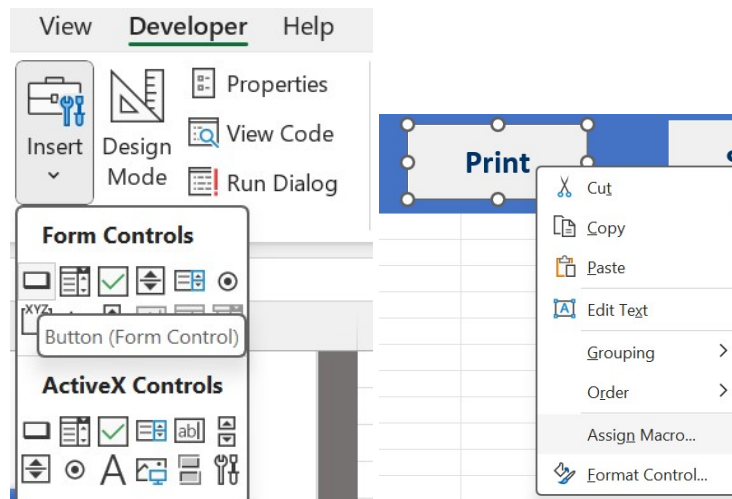

Figure 9: Creating buttons and assigning them to macros.

2.8   Adding header and footer

The header and footer are added to the sheet. Headers and footers are only visible in the Page Layout view and on printed pages. The final view of the GDP sheet after adding macros, header, and footer see Fig10
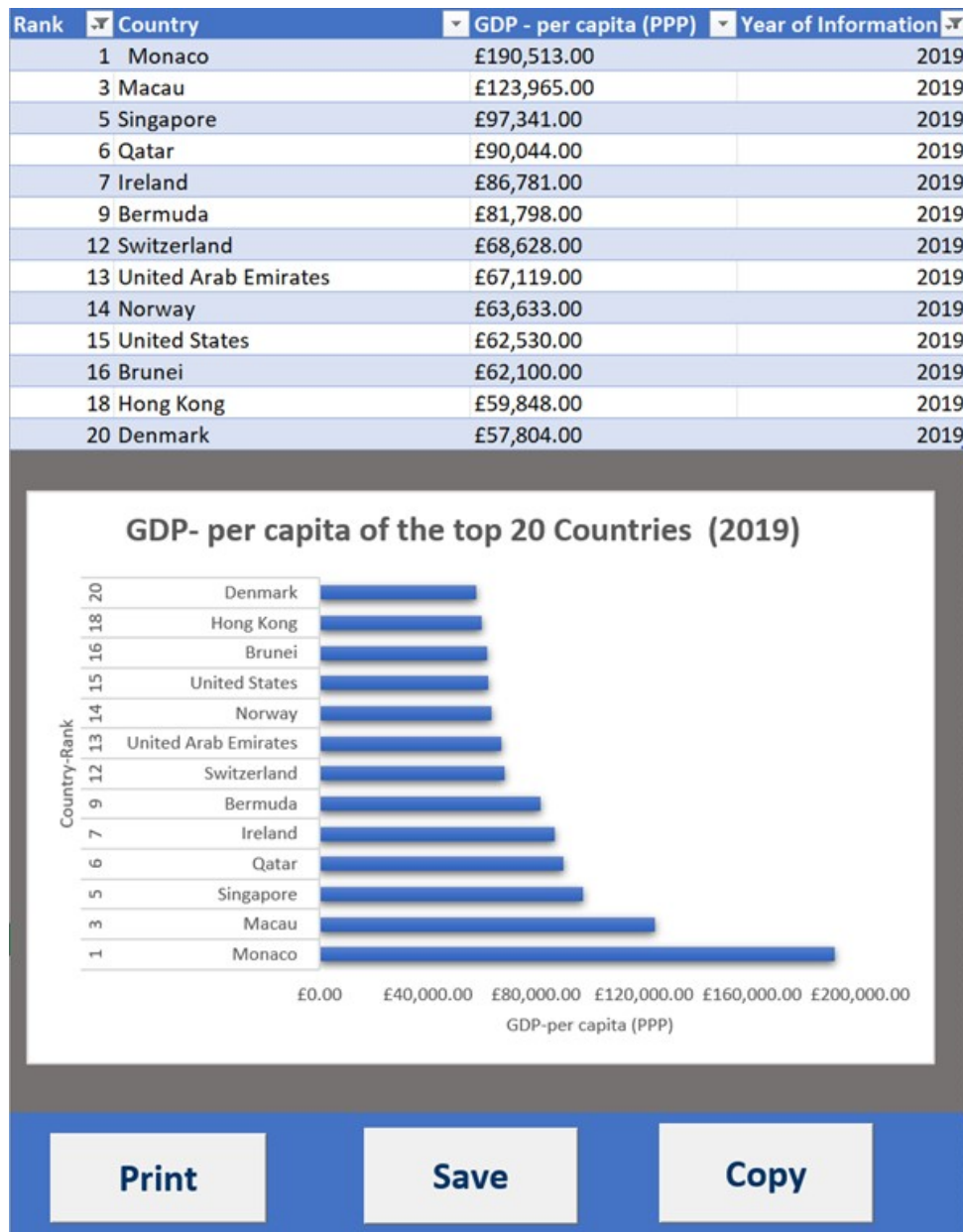

Figure 10: GDP sheet with save, print, and copy buttons.

3   Tableau

The same Wealth of Nations Excel file is used to create the dashboards in Tableau. The Excel file is imported into Tableau, and relationships between the tables are established using "Country" as the common field. see Fig11
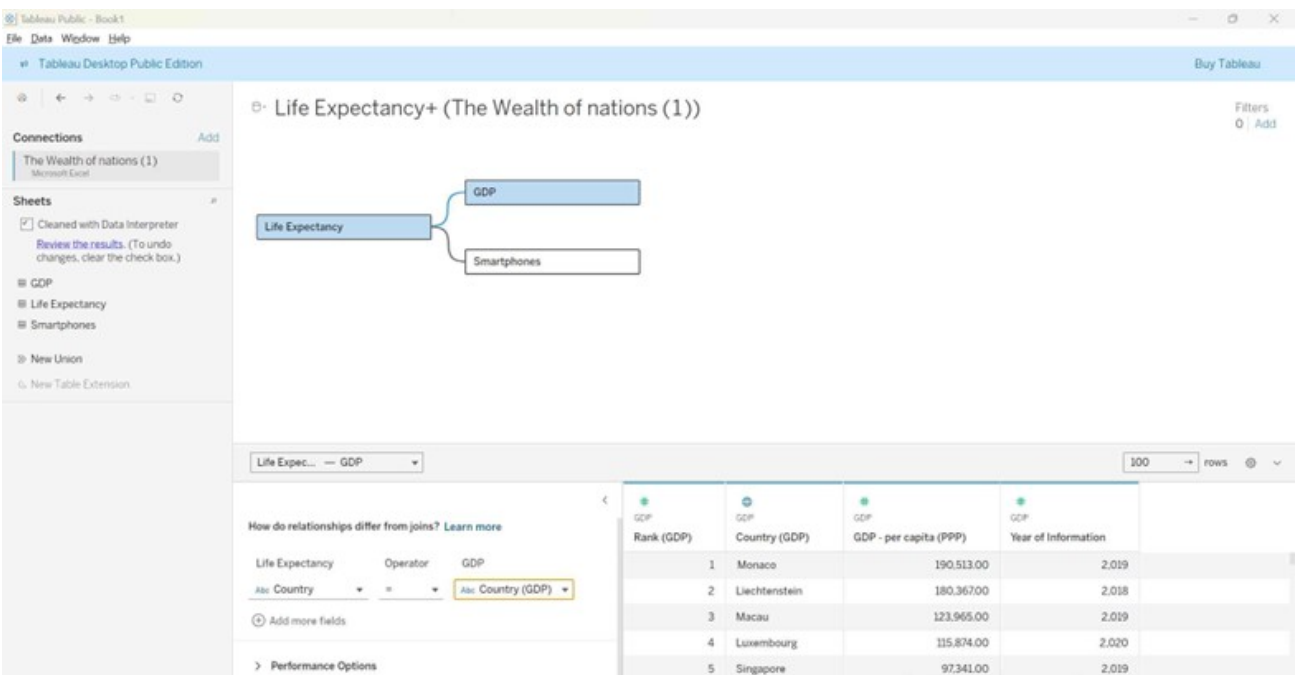

Figure 11: Creating relationship between the tables.

3.1   Dashboard GDP-per capita (PPP)

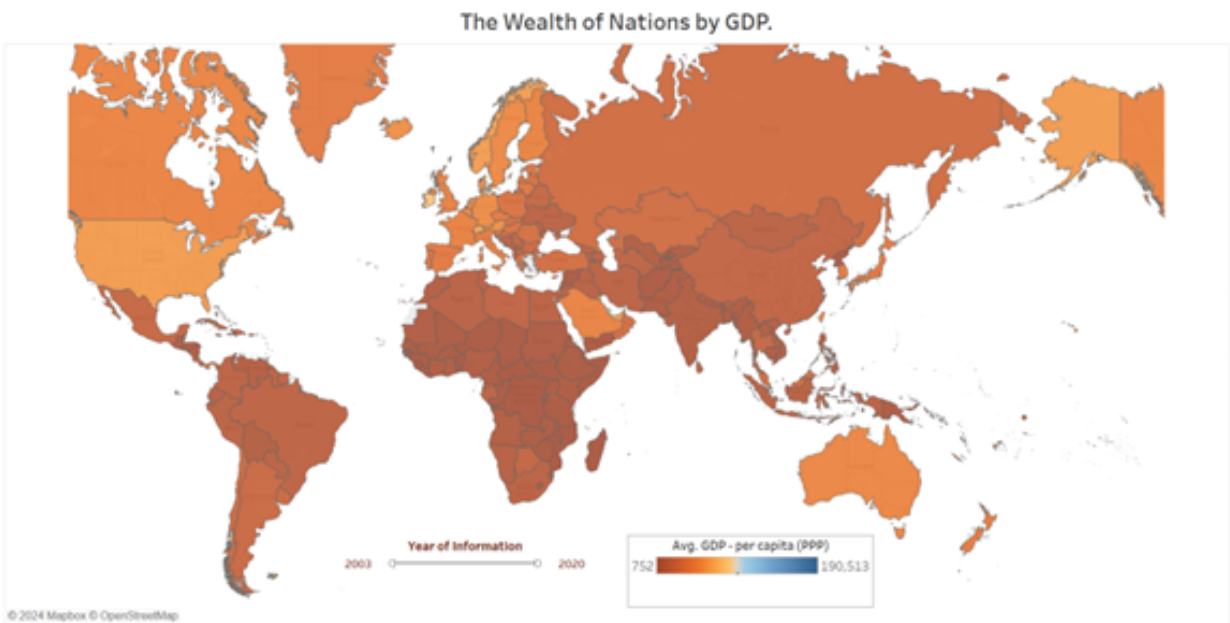Figure 12: Filters and data used to create the sheet.

Figure 13: Wealth of Nations by GDP.

3.2   Top 20 countries

The image below displays the top 20 countries based on the number of smartphone users and their life expectancy at birth. The countries are filtered according to their rank. Fig14
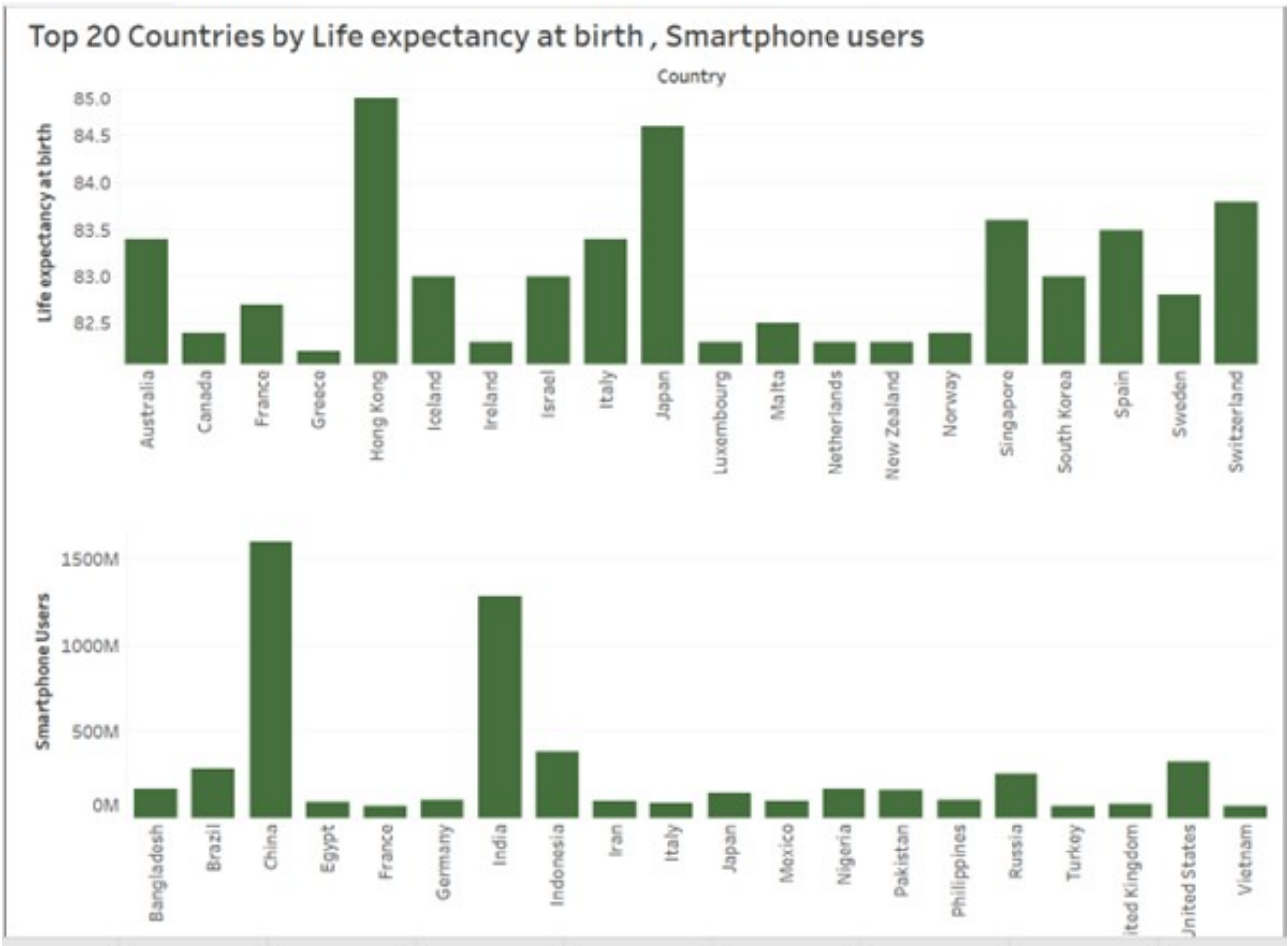
Figure 14: Top 20 countries.

I was also looking to showcase the top ten countries filtered by the ten top smartphone users.
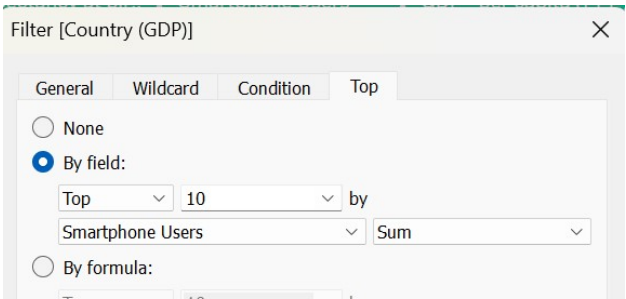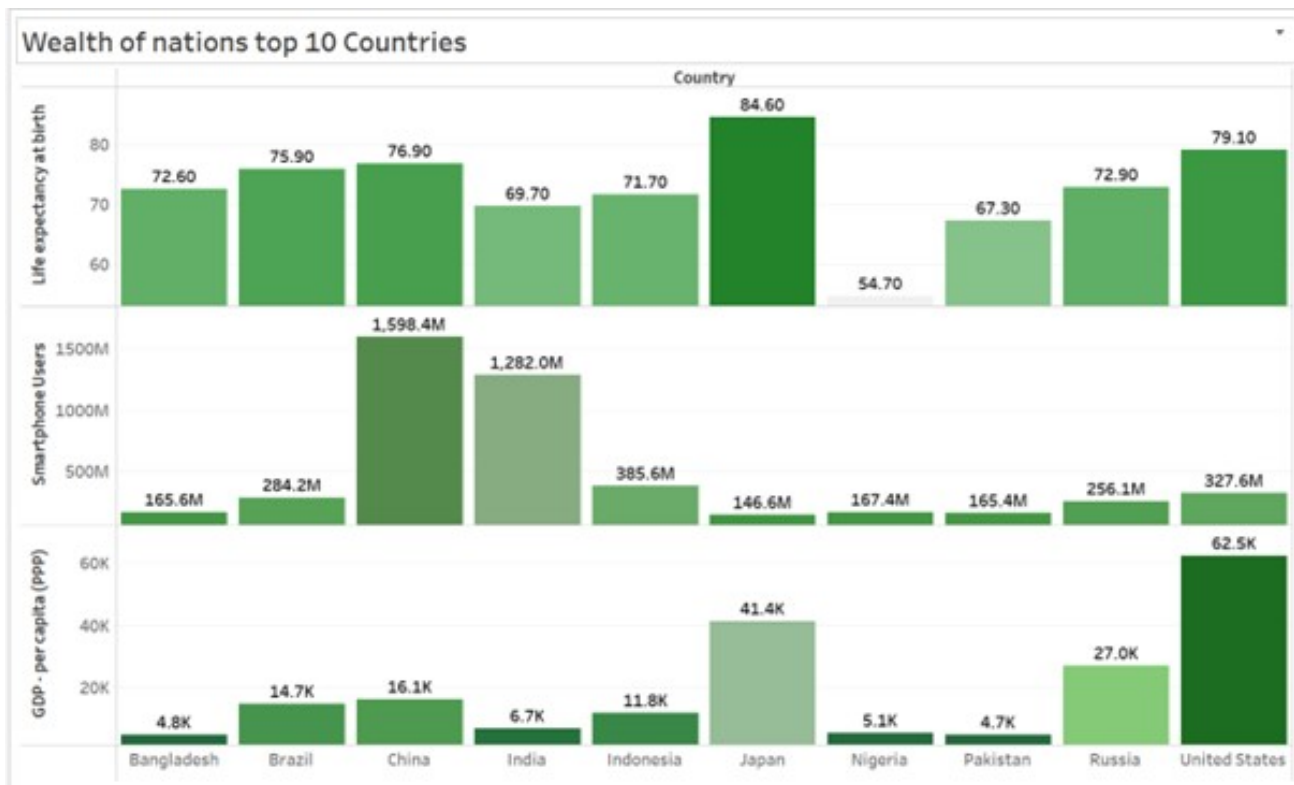

Figure 15: Filters used to display the data.

Figure 16: Top 10 countries filtered by smartphone users.

3.3 Relationships between the tables

The images below aim to illustrate any correlations between GDP, life expectancy at birth, and smartphone users. While there appear to be correlations among the data, it's difficult to draw definitive conclusions due to the data being collected in different years. Fig17

Relationship Life expectancy at birth Vs GDP per capita(PPP)

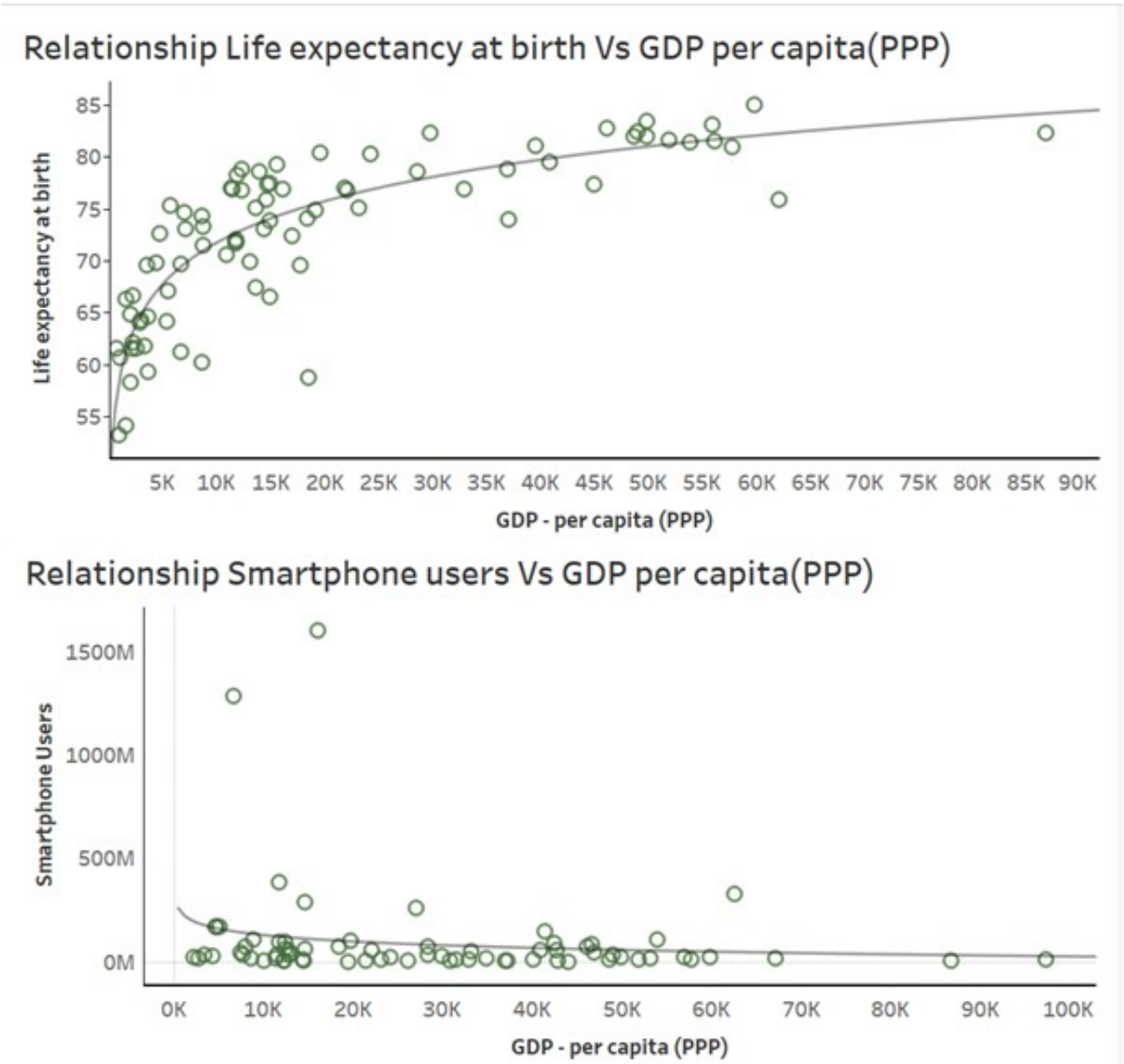## Relationship Smartphone users Vs GDP per capita(PPP)

Figure 17: Relationships between the data.

In all of the above images, I attempted to remove null values by applying filters.
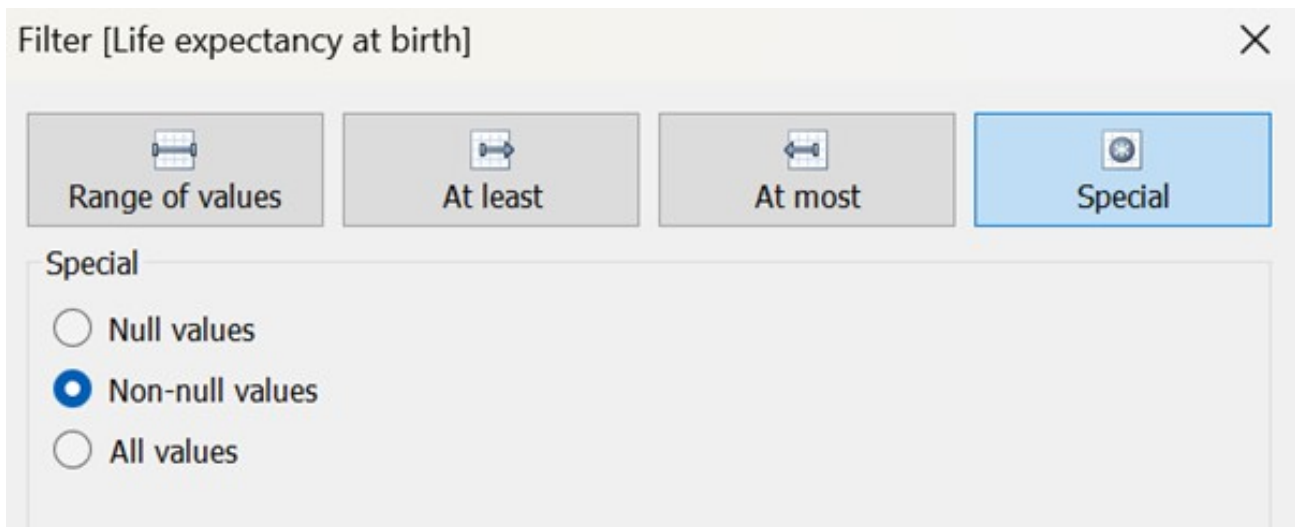
Figure 18: Filter to remove the null values.

Link to the tableau public profile
:https://public.tableau.com/app/profile/suzan.yemane/vizzes

4   Reflective account

This assignment has significantly enhanced my proficiency in Excel and Tableau, allowing me to successfully implement the tasks assigned. I found the steps in Excel clear and user-friendly, enabling me to navigate functions with ease. In contrast, Tableau introduced me to an exciting world of data visualization, which was a refreshing experience. However, I faced challenges when trying to analyze correlations between variables, as the data was collected over different time frames, which limited my ability to draw meaningful conclusions. This experience underscored the importance of understanding the context of data collection.

While I managed to complete the tasks, I realized I need to further improve my skills in both tools. To address this, I plan to dedicate some time to explore advanced functions in Excel, such as pivot tables as well as for Tableau to enhance my visualization techniques. By doing so, I aim to become more proficient in data analysis and visualization, which will better prepare me for future projects.

List of Figures

References

[1] "Data policies and contracts." ://dssc.eu/space/Glossary/176554143/8.+Data+policies+and+contracts.

[2] "Data policies." ://www.tpximpact.com/knowledge-hub/insights/ethical-data-collection.

[3] "The Data Analysis Process." ://www.datacamp.com/blog/what-is-data-analysis-expert-guide.