JustIT Data Bootcamp Assignment 1

report by narmeen Mohammed

due 05.04.2024

2024

**Contents**

Policies and Procedures……………………………………………………..…………. 3

‘The Wealth of Nations’ Excel Task…………..…………………………..…….4-20

Visualisation for the ‘Wealth of Nations’ data……………………………21-28

Reflection……………………………………………………………………………………….29

Policies and Procedures

Complying with policies and procedures as a data analyst is significantly important. This is because part of becoming a data analyst involves handling confidential information on individuals and/or business statistics. This is considered sensitive information and it is therefore crucial that this data remains protected.

Laws and regulations such as the EU General Protection Regulation (GDPR) govern how the data of individuals within the EU may be processed and transferred, and failure to obligate to these protection laws can result in severe sanctions[[1]](#footnote-2).

The way an organisation governs data may vary, however the following models are often adhered to; data quality, data security and compliance, data stewardship and data transparency.[[2]](#footnote-3)  
Data quality is about ensuring the data you have sourced or collected is accurate and reliable. Without this, you cannot have a robust or true analysis of the data.  
Data security and compliance is a process of defining risk levels for data, and creating secure access points, keeping a balance between user interaction whilst also maintaining a matched level of security.  
Data stewardship helps to monitor how individuals and teams use data sources, and leading by example to ensure all of the above; data access and security, and quality of the data source.  
Finally, data transparency; this is also vital as when working with data, data analysts and businesses should be able to easily find out where their data comes from.

It is important to adhere to rules and regulations regarding handling sensitive information, and following laws on GDPR. This is because as a data analyst it helps build trust between yourself and the business you are working with; demonstrating these steps in data protection is good practice and ensures the organisation that they can trust you will secure their data well. Following these procedures of securing datasets with security measures i.e. locking datasets with passwords, also helps prevent fraud and crimes; when confidential data gets in the wrong hands it can jeopardise individuals and/or a business, which can also be costly when dealing with the aftermath of a personal data breach or hacker attack[[3]](#footnote-4).

**‘The Wealth of Nations’ Excel Task**

In this section of my report, I will use ‘The Wealth of Nations’ Excel spreadsheet to answer the questions in the assignment, alongside providing screenshots and annotations of the steps I took to answer the questions.

GDP TASKS

1. Set a password to protect the workbook.

A screenshot of a computer

Description automatically generated

I went to File>Info>Protect Workbook.

A screenshot of a computer

Description automatically generated

After selecting Protect Workbook, I select Encrypt with Password. Following by setting the password as ‘nations’.

A screenshot of a computer screen

Description automatically generated

I then select ‘Ok’, and it prompts me to re-enter the password.

A screenshot of a computer screen

Description automatically generated

The Info now changes to inform the reader a password is required to open the workbook, and it is set to read only.

A yellow sign with black text

Description automatically generated

1. Highlight column C and change the data to display in GBP.

A screenshot of a computer

Description automatically generated

I select column C by clicking on it.

A screenshot of a computer

Description automatically generated

In the ribbon, under the home tab, within the Number field, select the drop down where it is currently on ‘Custom’.

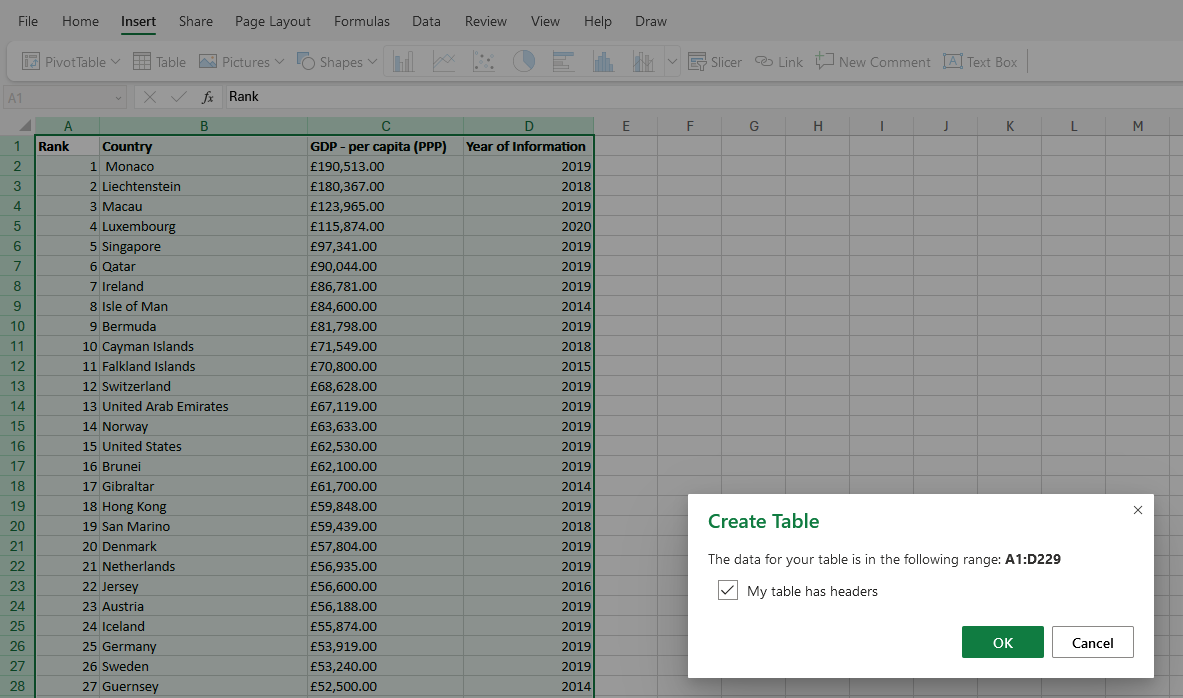
A screenshot of a computer

Description automatically generated I then select ‘Currency: GDP – per capita (PPP)’.

A screenshot of a computer

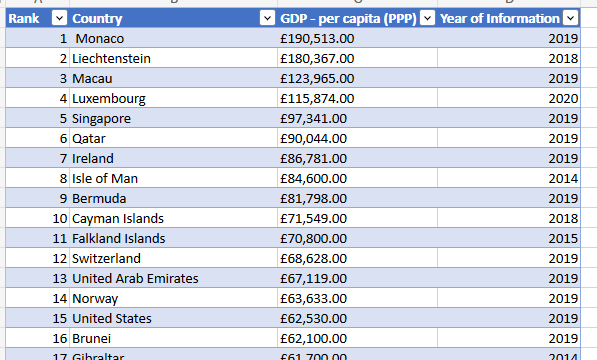
Description automatically generated I can see here it has changed.

1. Turn the GDP sheet into a table.

In this snapshot I demonstrate highlighting the entire contents of the table with CTRL+A, I then click on the ’Insert’ tab>’Table’ to create a standard table.

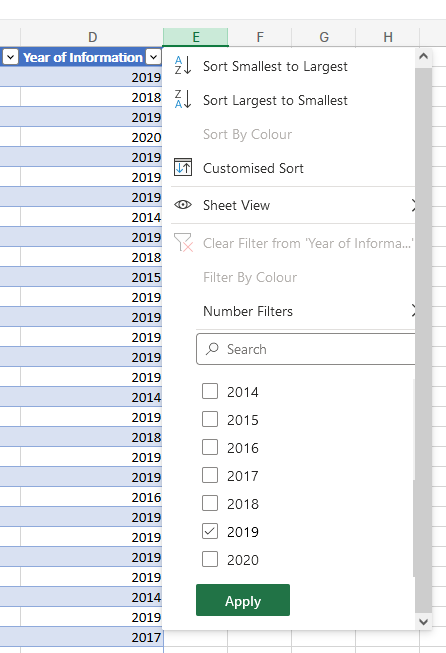
It then prompts me before creating the table with the following box informing me of my selection and if my table has headers, which it does.

I click on ‘Ok’.



I have now created a table.

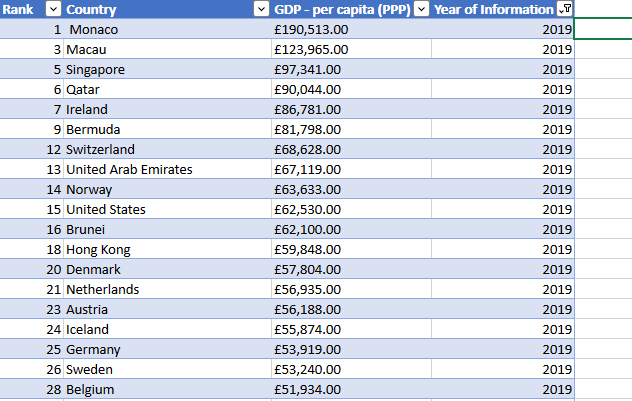
1. Filter the table to display on the information for 2019.



To filter the table to only display information for 2019, I clicked on the dropdown menu next to ‘Year of Information’.

I deselect every year and keep only ‘2019’.

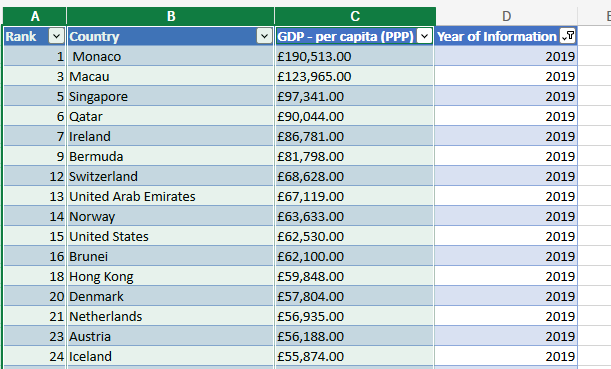
I then click on ‘Apply’.



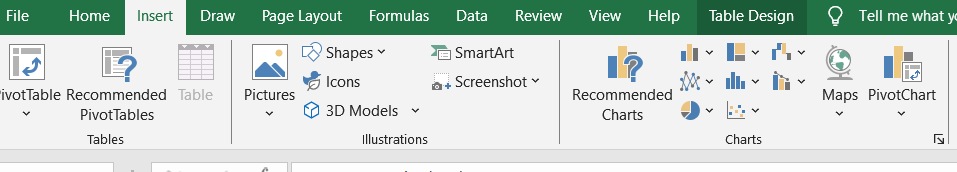
And now the information is filtered.

1. Create a chart that will only display the following data:

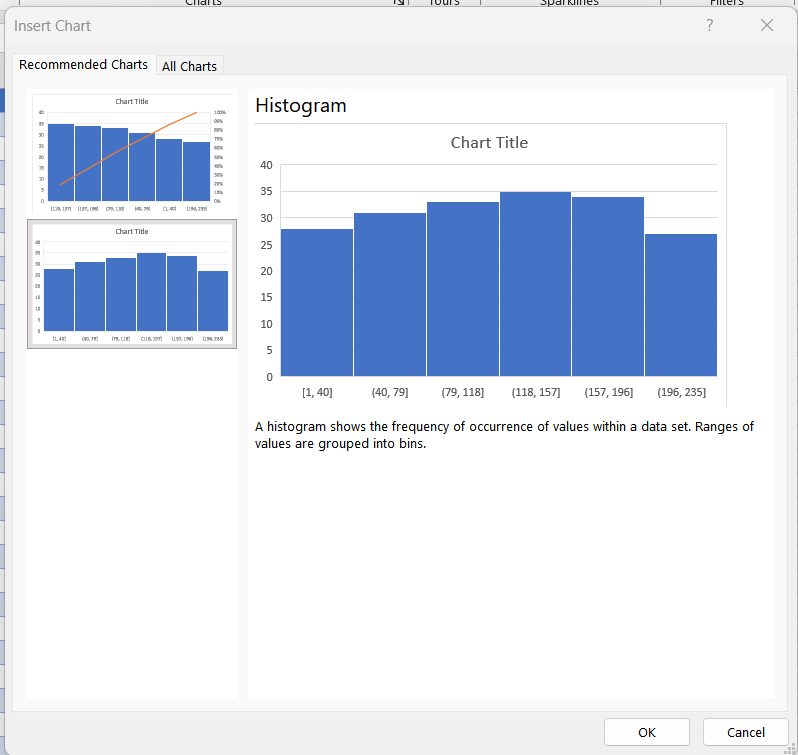
Rank, Country and GDP per capita (PPP).



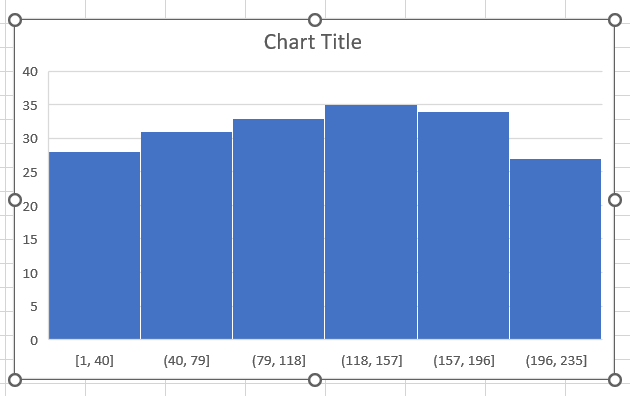
First, I select the necessary columns by clicking on each column I need and holding the CRTL key.



For this project, I go to ‘Insert’>‘Recommended Charts’ for a suitable chart, however another way I can do this is do simply select my own chart style.

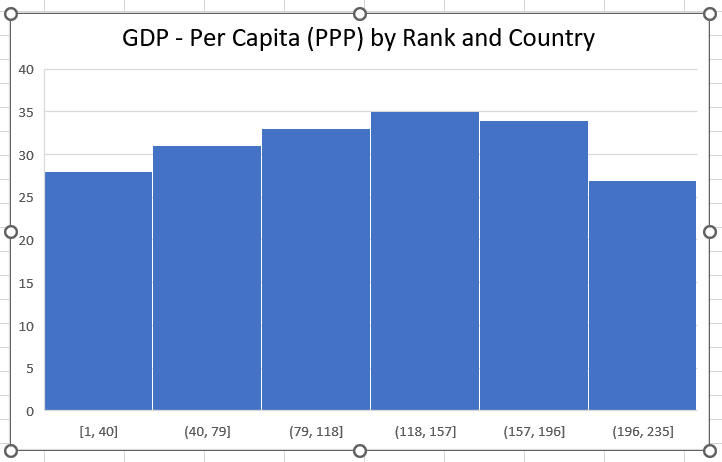


I decide to go with a Histogram chart, and select ‘OK.’



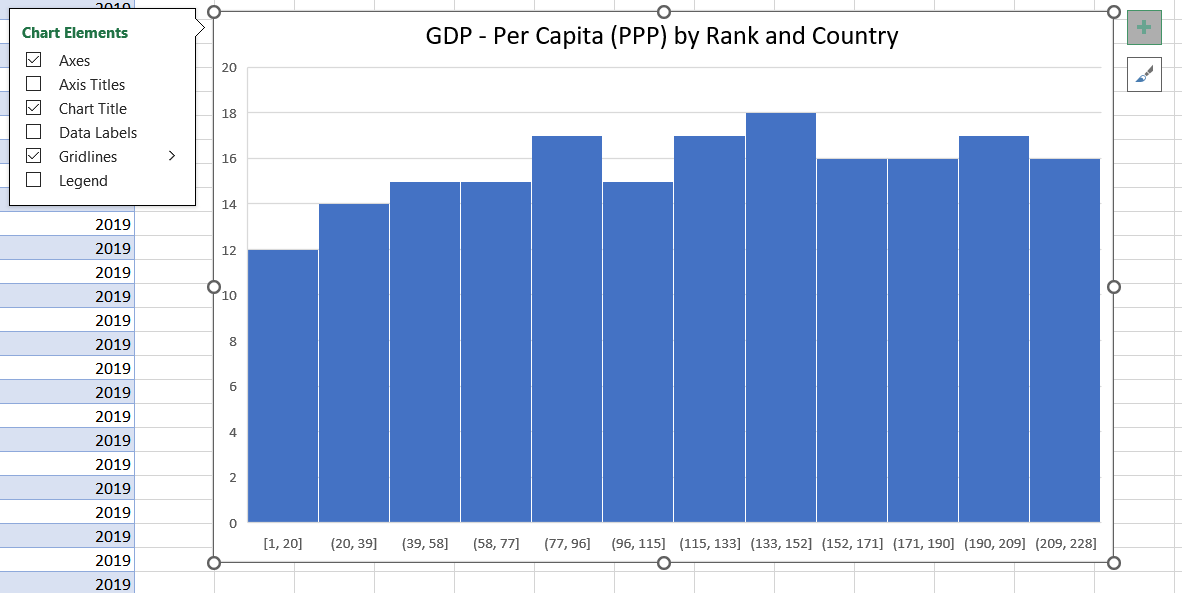
This is the chart I have made (unedited).

1. Edit the chart.
   1. Add a title

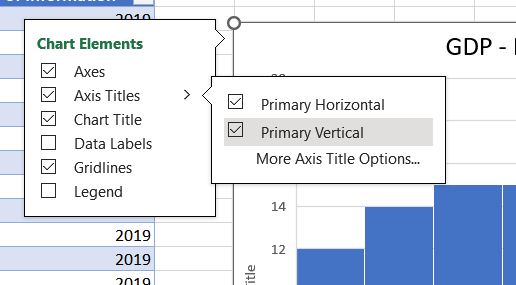


I double clicked the chart title to rename it to ‘GDP – per capita (PPP) by Rank and Country’.  
I then formatted it by changing the size of the font and colour.

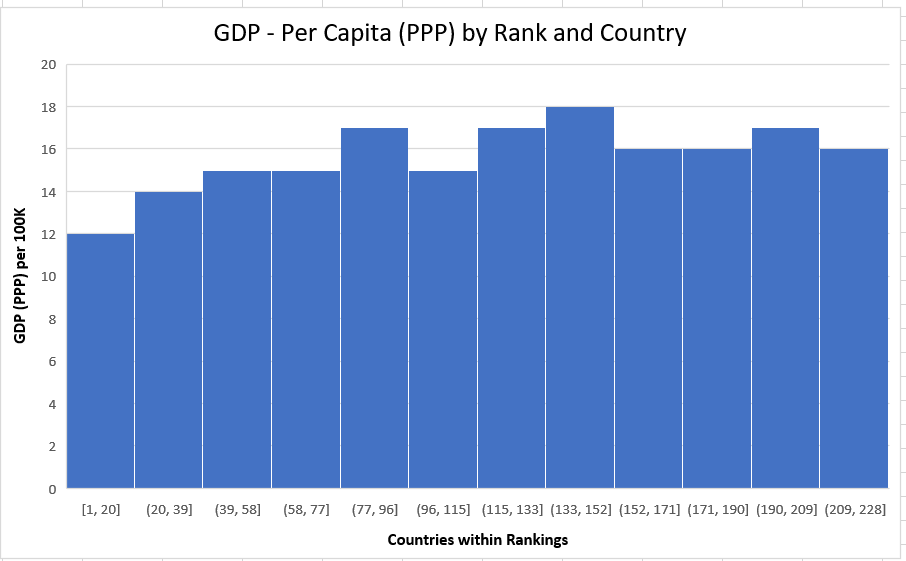
* 1. Add X and Y axis labels



On the chart, I click on the ‘Add’ button on the top right corner of the chart, which brings up the ‘Chart Elements’ menu. I then select ‘Axis Titles’.

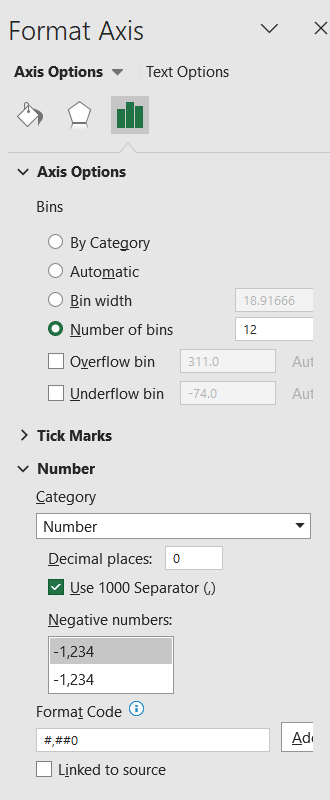


I select ‘Primary Vertical’ and ‘Primary Horizontal’ to create axis titles.

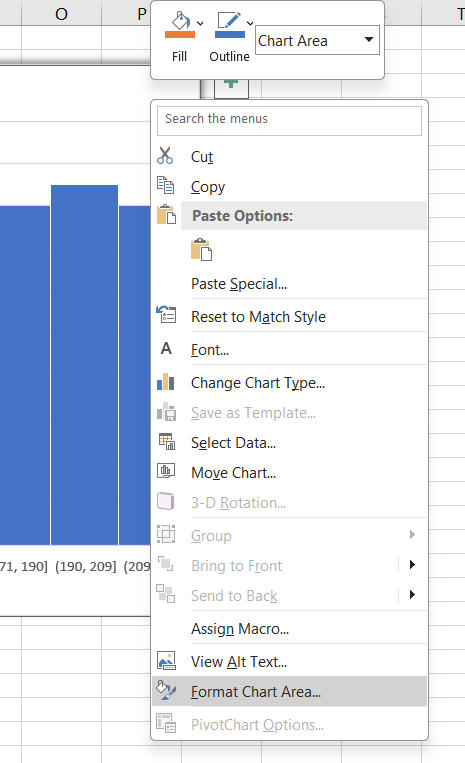


I then double click on the title names to change them, as shown above. I also format the x and y axis titles to be black, bold and one font size bigger.

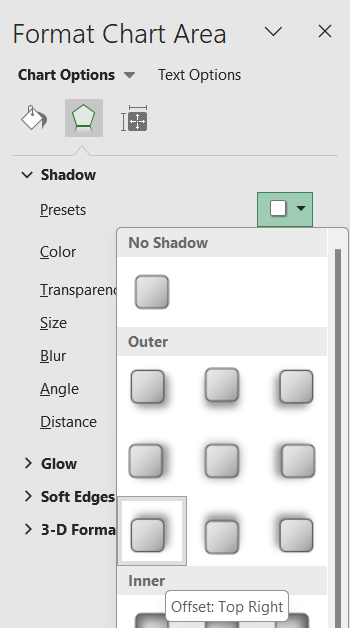
* 1. Make the chart visually pleasing



I did this prior to adding the axis titles,   
however on the X axis I decided to adjust the size of the bins for more clarity and organisation; there was a value of 229 for the ranks therefore I decided if I were to put 23 bins I would have 10 ranks per bin, however visually I think that this may be still uneasy to comprehend, therefore I decided on 12 bins as this organises the bins to a better distribution. I also changed the category to Number, and changed the decimal places to 0 for better visual clarity.



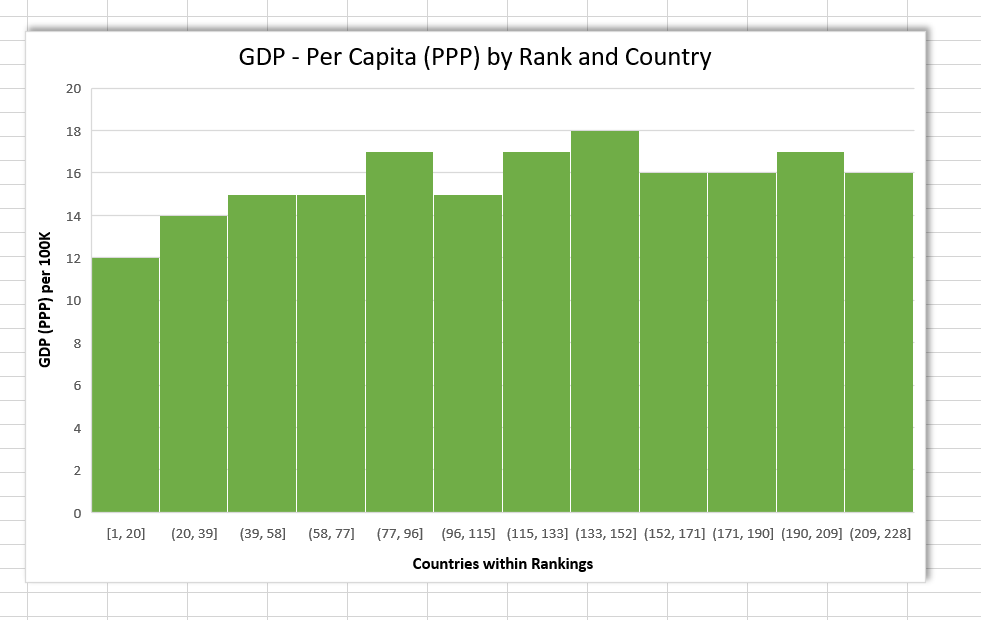
I also reformatted the visuals in terms of shadows and colours on the chart, I did this by right-clicking an empty space on the chart, and selecting ‘Format Chart’.



I select the presents on the ‘Shadow’ menu and select the Offset: Top Right.



To edit the colours of the chart, I selected the brush icon on the corner of the chart and clicked on the ‘Color’ tab, and selected a new colour.

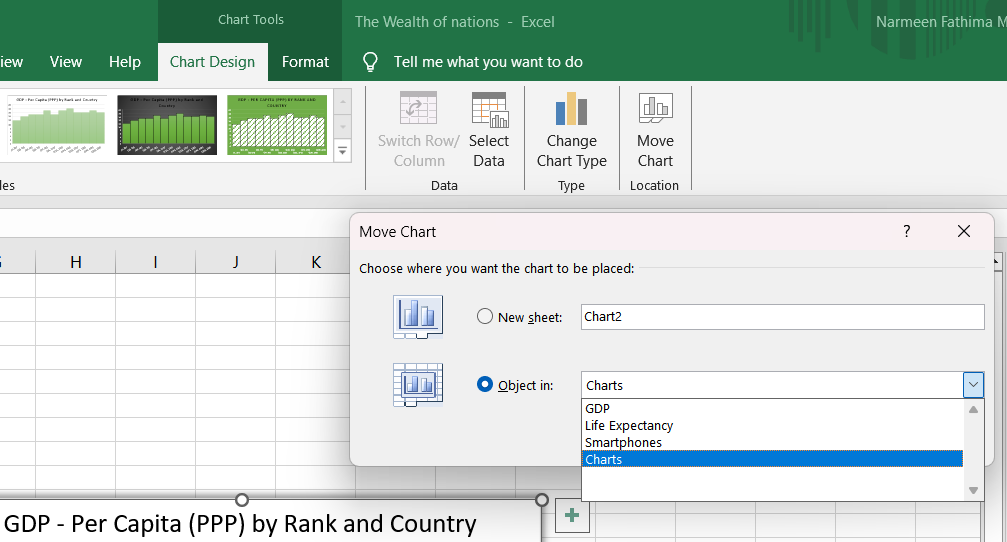


This is the result.

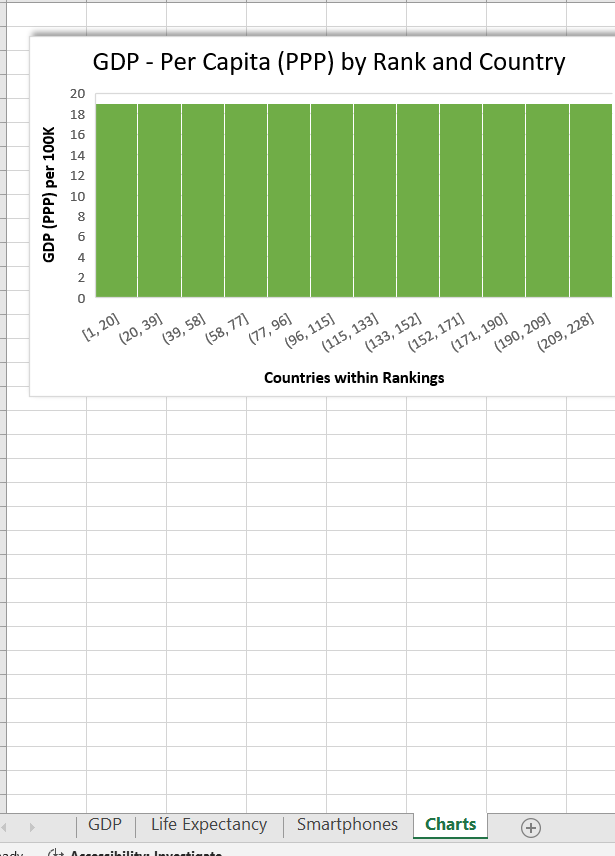
1. Move the chart to a new sheet tab and label with a suitable name.



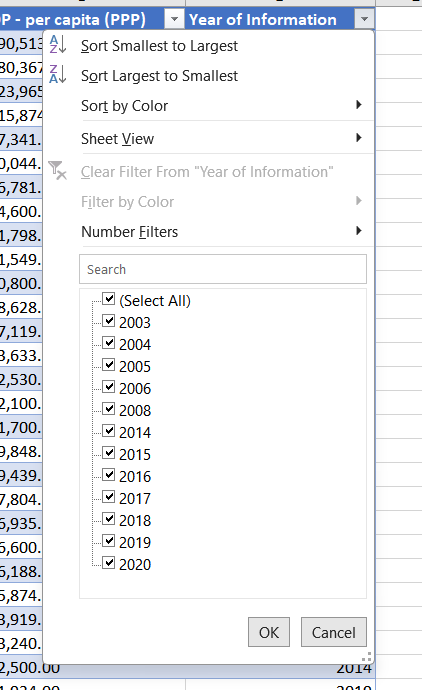
I created a new sheet using the plus sign called ‘Charts’.



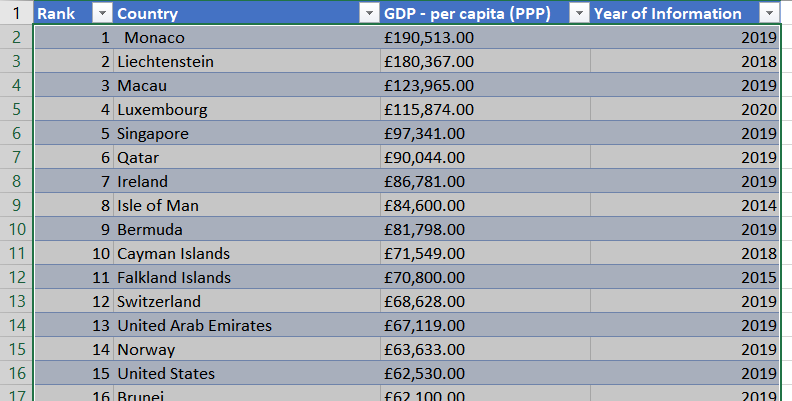
I then clicked on the chart, ‘chart design’ > ‘move chart’ option in the ribbon, selected the sheet I wanted to move it to and clicked ‘Ok’.

The result.

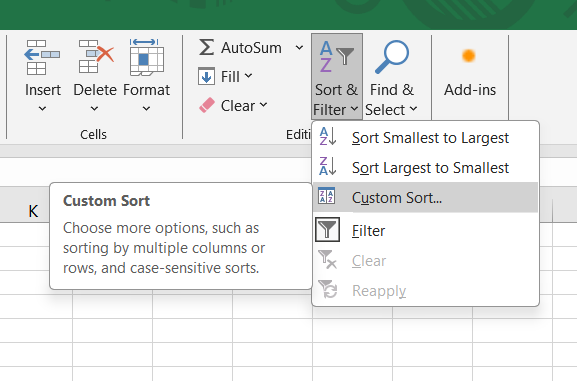
1. Create a sort for the top 20 highest ranking countries.



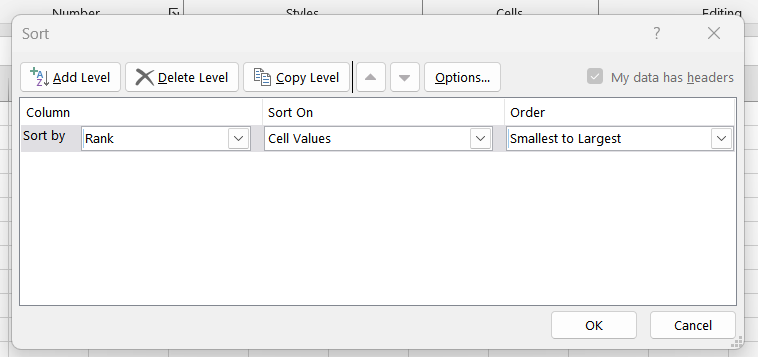
Firstly, I took off the filter for showing results for only 2019 by clicking on the dropdown for ‘Year of Information’, choosing ‘Select All’ and clicking ‘OK’.



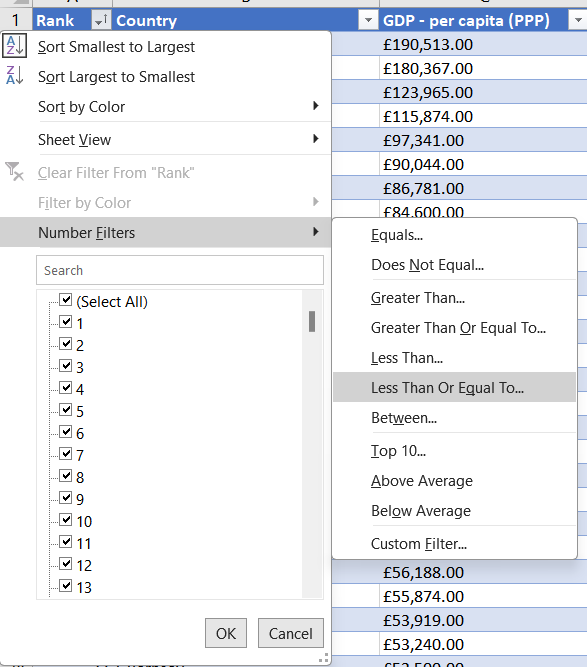
Then I highlighted the values.



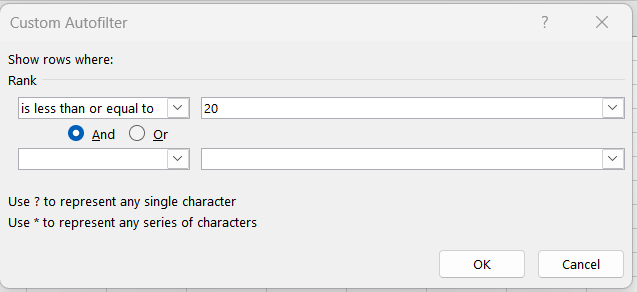
I clicked on ‘Sort & Filter’ in the ribbon>Custom Sort.



I made sure the ranks were already in order to smallest to largest.

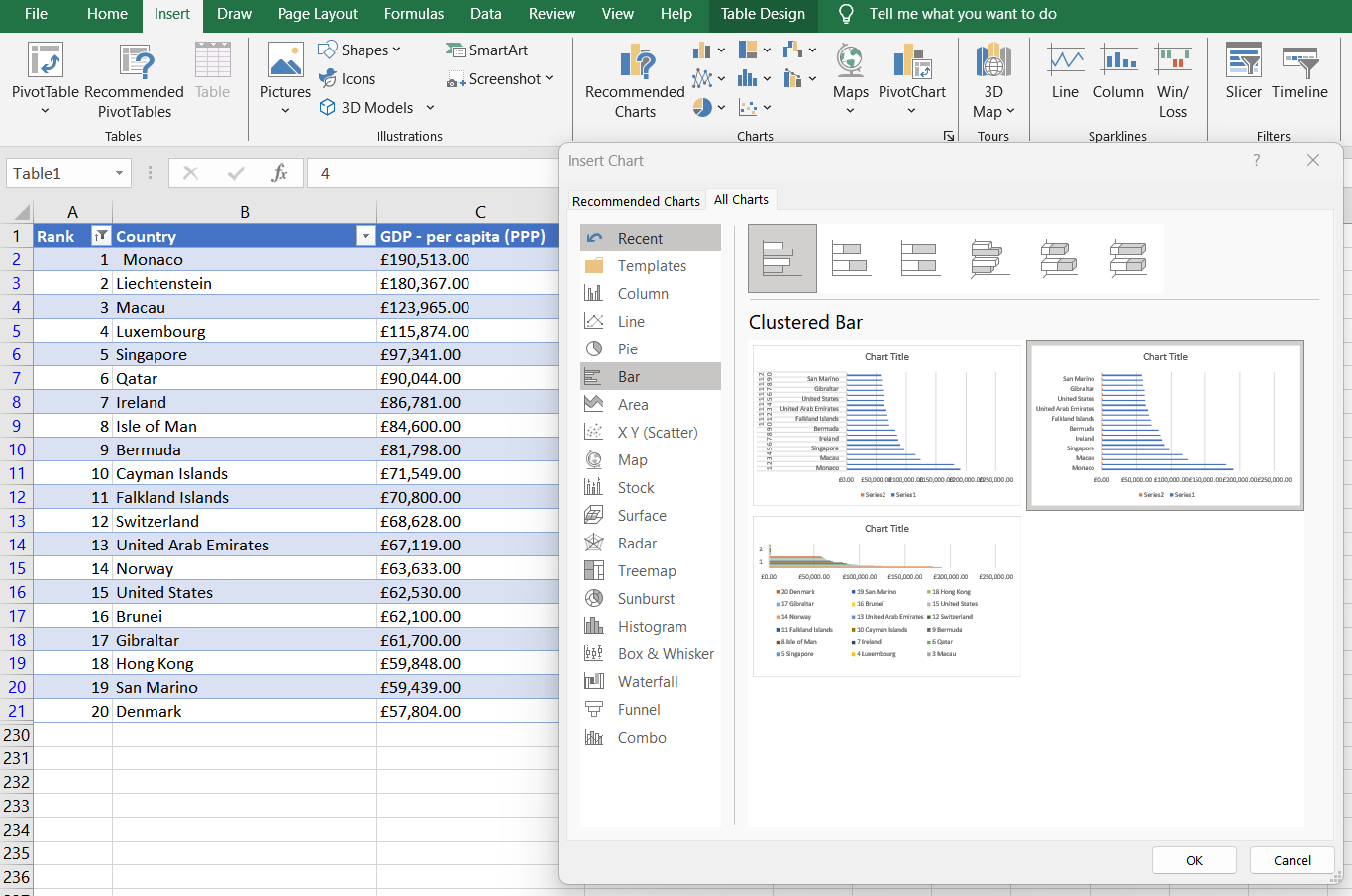


And then I clicked the dropdown on ‘Rank’. Selected ‘Number Filters’, and ‘Less Than Or ‘Equal To…’.



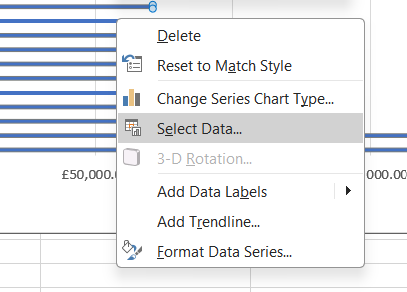
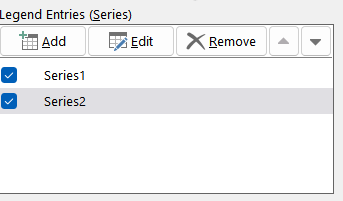
I selected is less than or equal to 20, in order to get the top 20 ranks. And click ‘Ok’.

1. Create a new bar chart to display the 20 highest ranking countries from your sort and move the chart to be underneath the table.

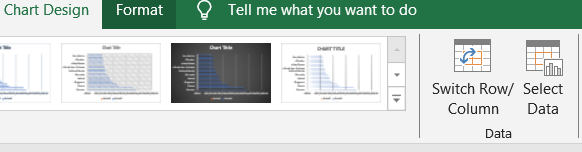


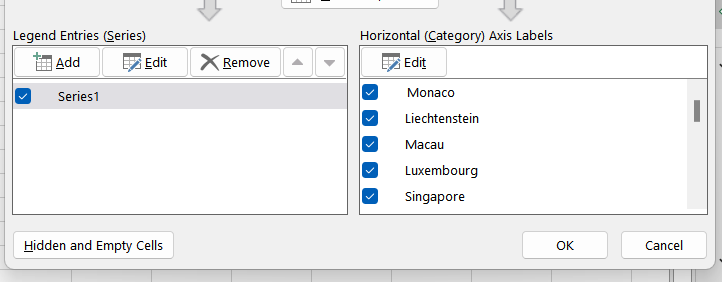
I followed the same steps as previously;  
CTRL+A on table > Insert > Charts > Bar chart > Ok.

I then formatted the chart title by double clicking and renaming.

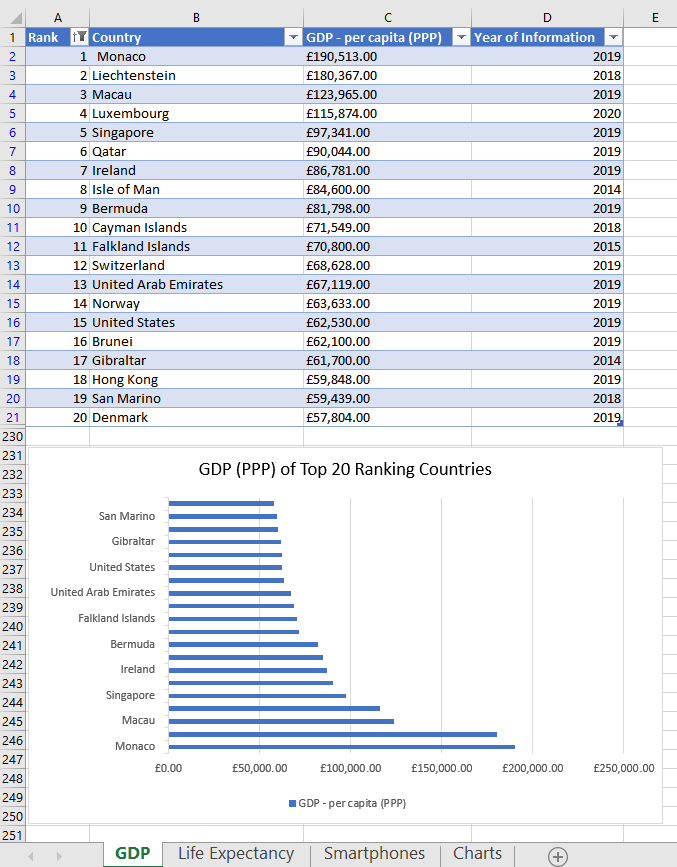
 

I also removed the ‘Year of Information’ data and it wasn’t adding to the chart, by right clicking the data > ‘select data’> removing ‘series 2’ > ‘Ok’.

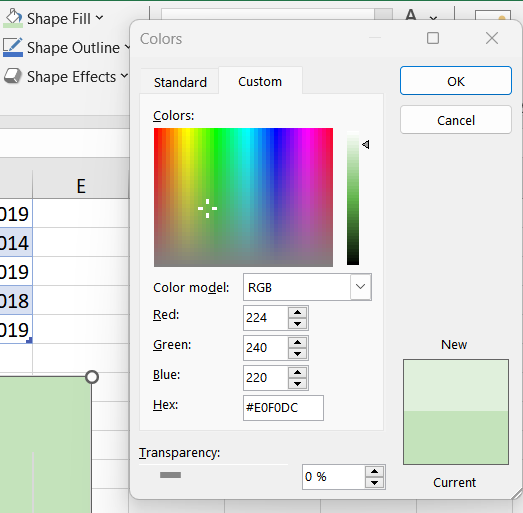




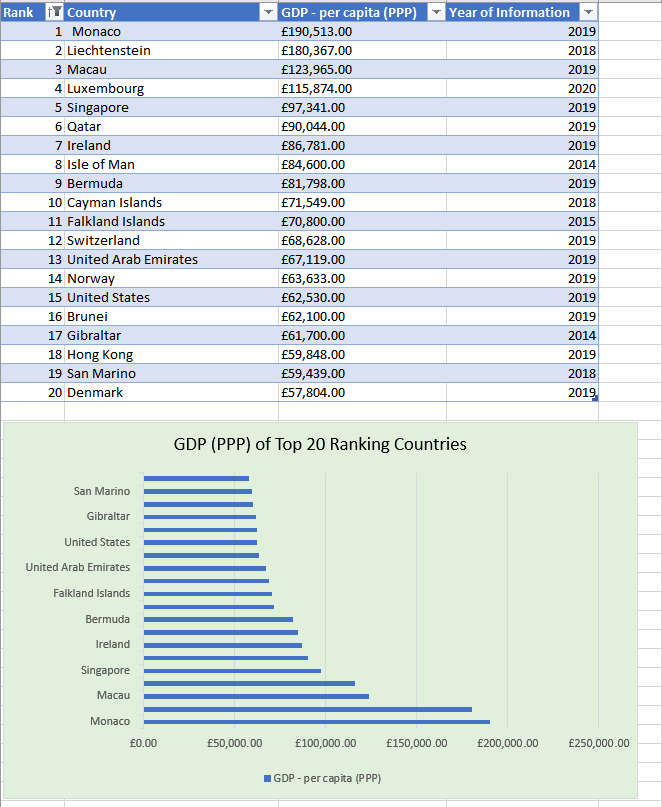
I renamed the legend by going to ‘Chart Design’ > ‘Select Data’ > editing ‘Series1’ to ‘GDP – per capita (PPP)’.

 This is my result.

1. Colour the background by highlighting the area underneath the table as shown below.



I clicked on the chart > went to ‘Shape Fill’ > picked a custom colour > selected ‘Ok’.



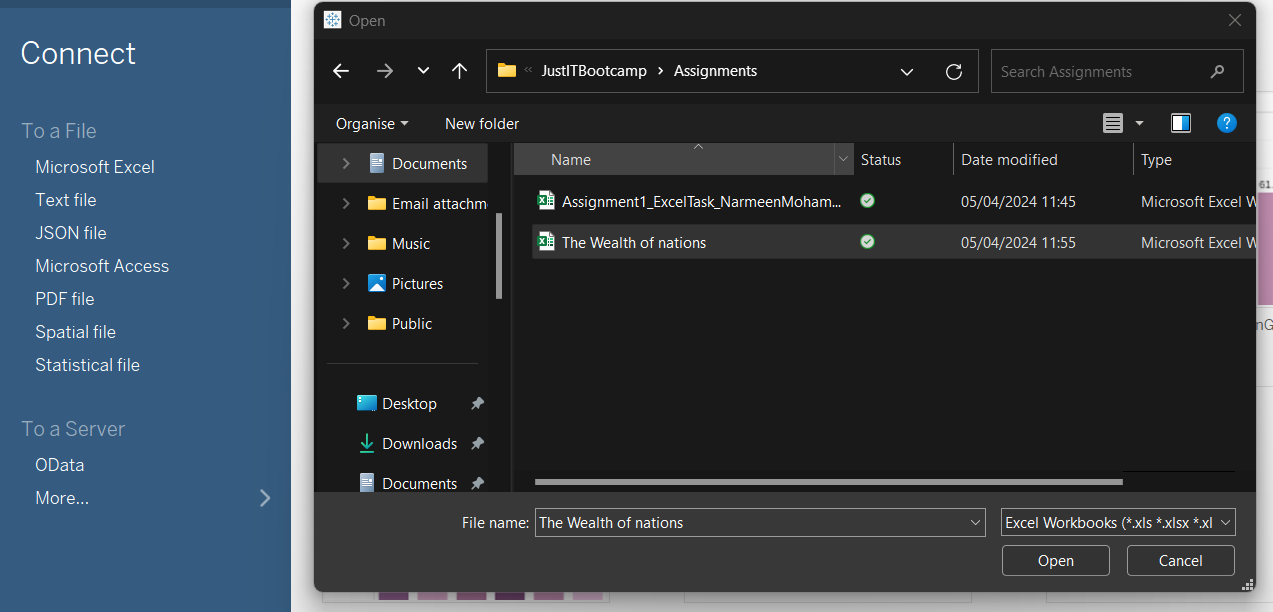
This is my result.

**Visualisations for ‘The Wealth of Nations’ Data**

In this section of my report, I will demonstrate building visualisations using the same excel table, keeping in mind the client requirements:

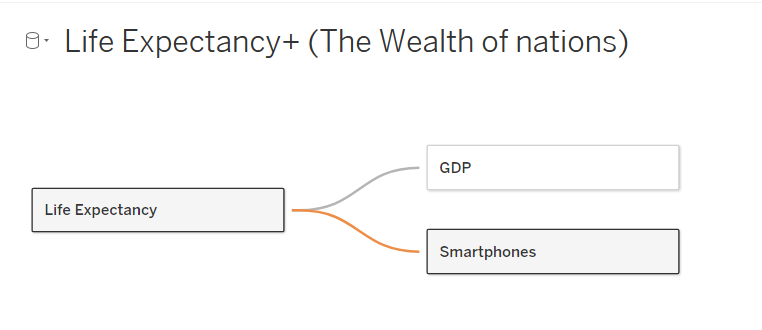
The client is colour blind and requested you to bear this in mind when building your dashboard. The client is only interested in the top 20 highest ranking countries. All your visuals should be for the top 20 highest ranking countries.

1. Import data:

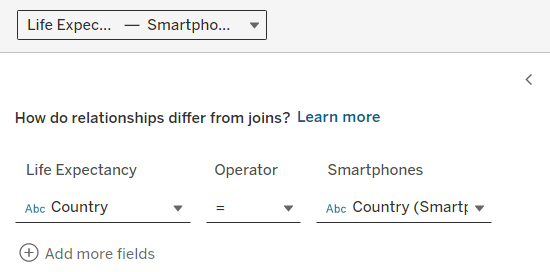


I opened Tableau, connected to an excel file and selected which one I wanted, in this case the ‘Wealth of Nations’.xls

1. Set relationships:



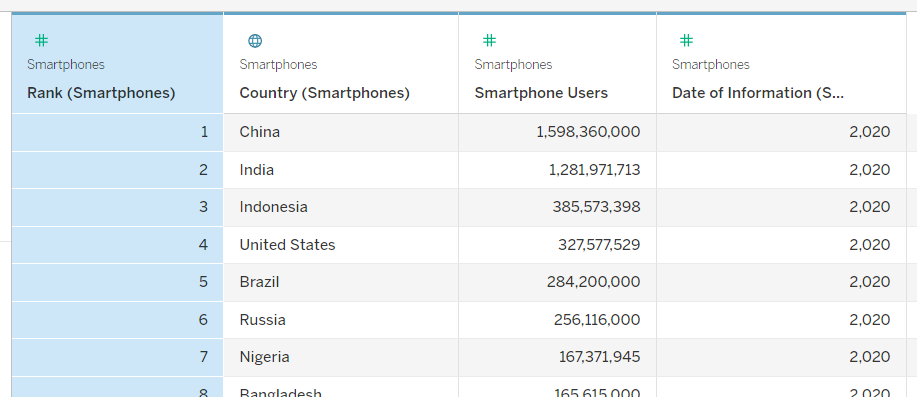
I drag and dropped the sheets and created relationships.



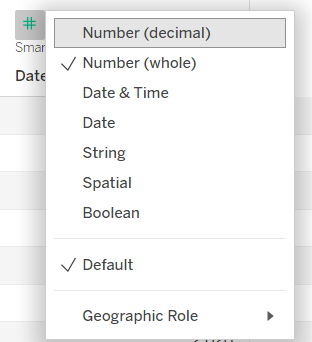
I set the relationships between the three sheets.

They all have one mutual column of data in common; ‘Country’. Therefore, this is the data that links the three sheets of data together.

1. Check data types:

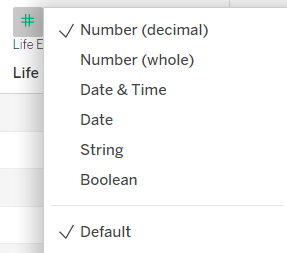


The data types for Smartphones here are as listed; Rank, Smartphone Users and Date of information are Numerical Values and Country is a Geographical Value.



I change the data type for ‘Date of Information’ to Date by clicking on the # symbol which opens the data type menu and selecting date, this however changes the date format to dd/mm/yyyy; I plan to filter just the year later in the process.

I also change the data type for ‘Date of Information’ across the other sheets to Date.



I checked the other sheets for data types.

I notice ‘Life Expectancy at Birth’ in Life Expectancy is Number (decimal), I decide to change it to a whole number as it makes more sense, being an age.

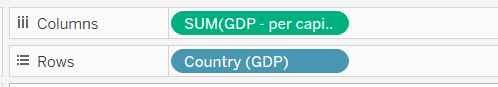
1. Build charts:

I’m going to build a minimum of 4 charts.

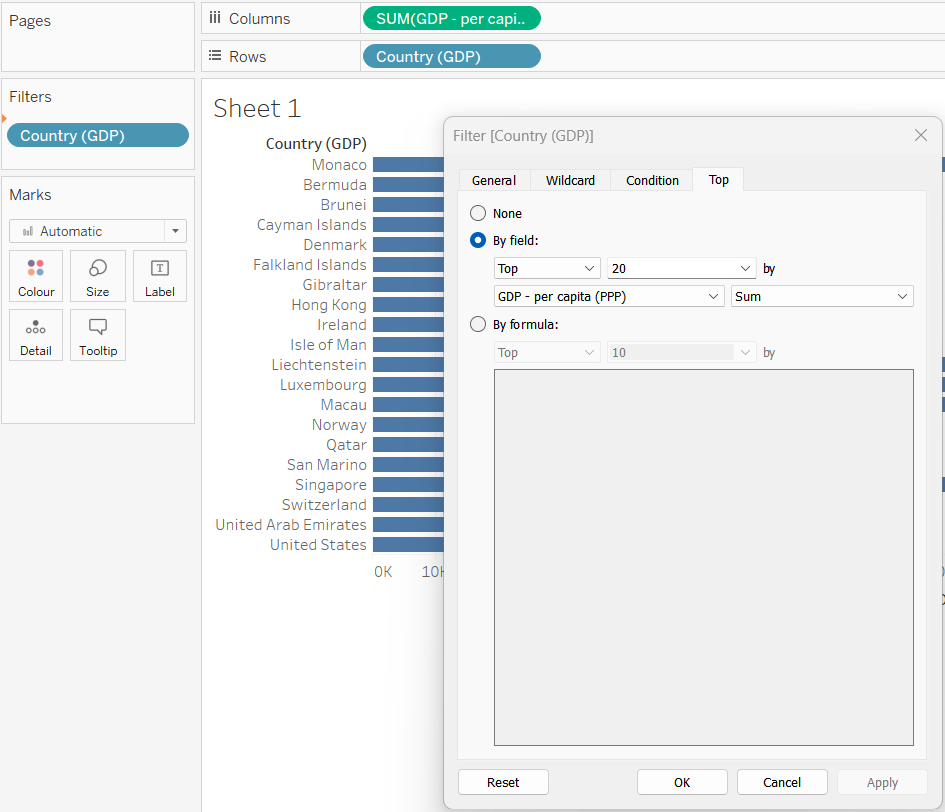
I decided which relationships I’d like to look at first:

* Top 20 GDP countries
* Top 20 smartphone using countries
* Top 20 Countries in Life Expectancy Birth Rates and GDP
* Top 20 Countries in Life Expectancy and Smartphone Usage

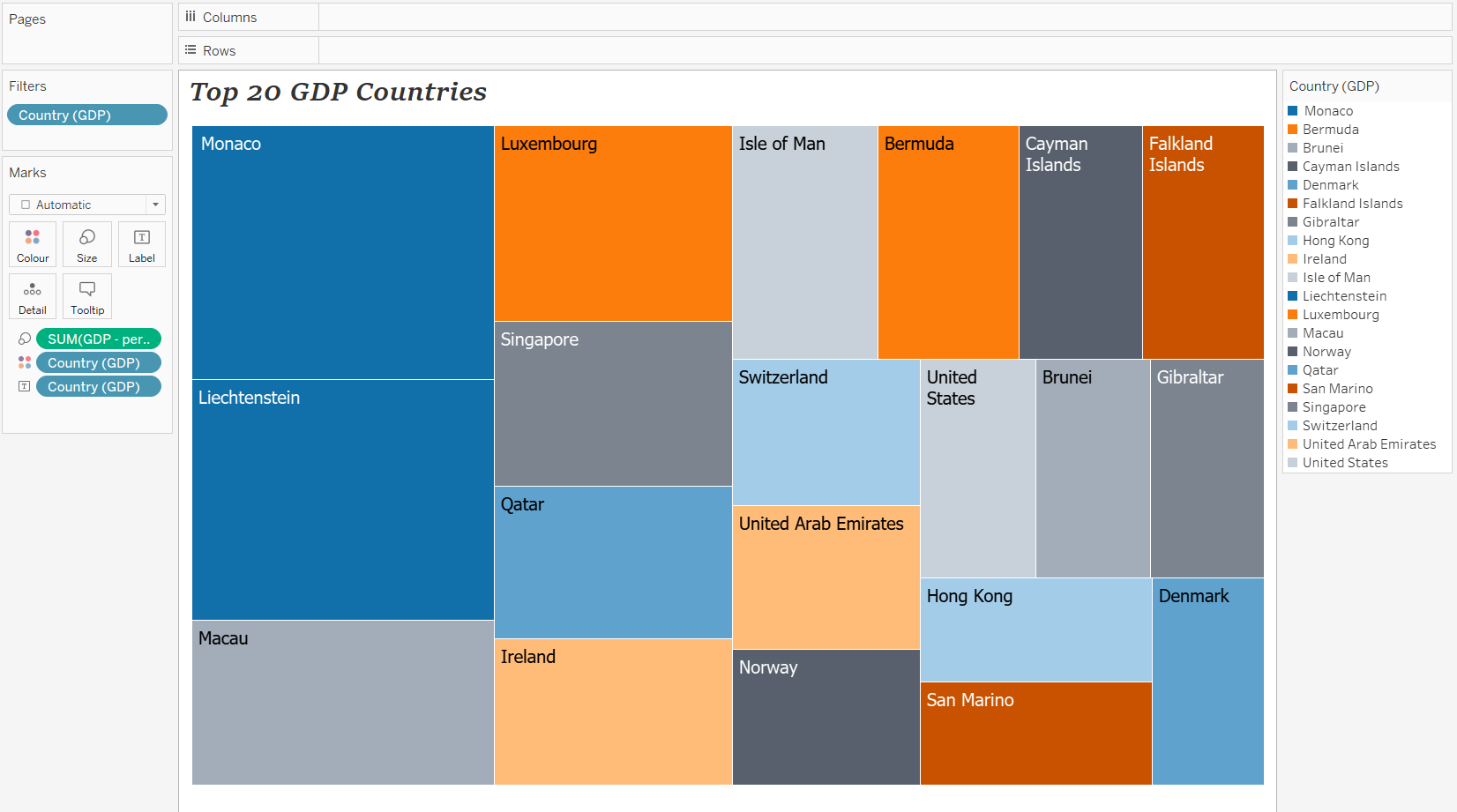
*First visual: TOP 20 COUNTRIES GDP – per capita (PPP)*

**

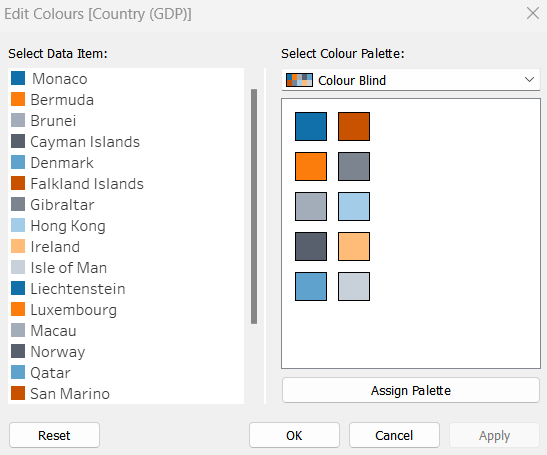
I dragged GDP to columns and Country to rows.



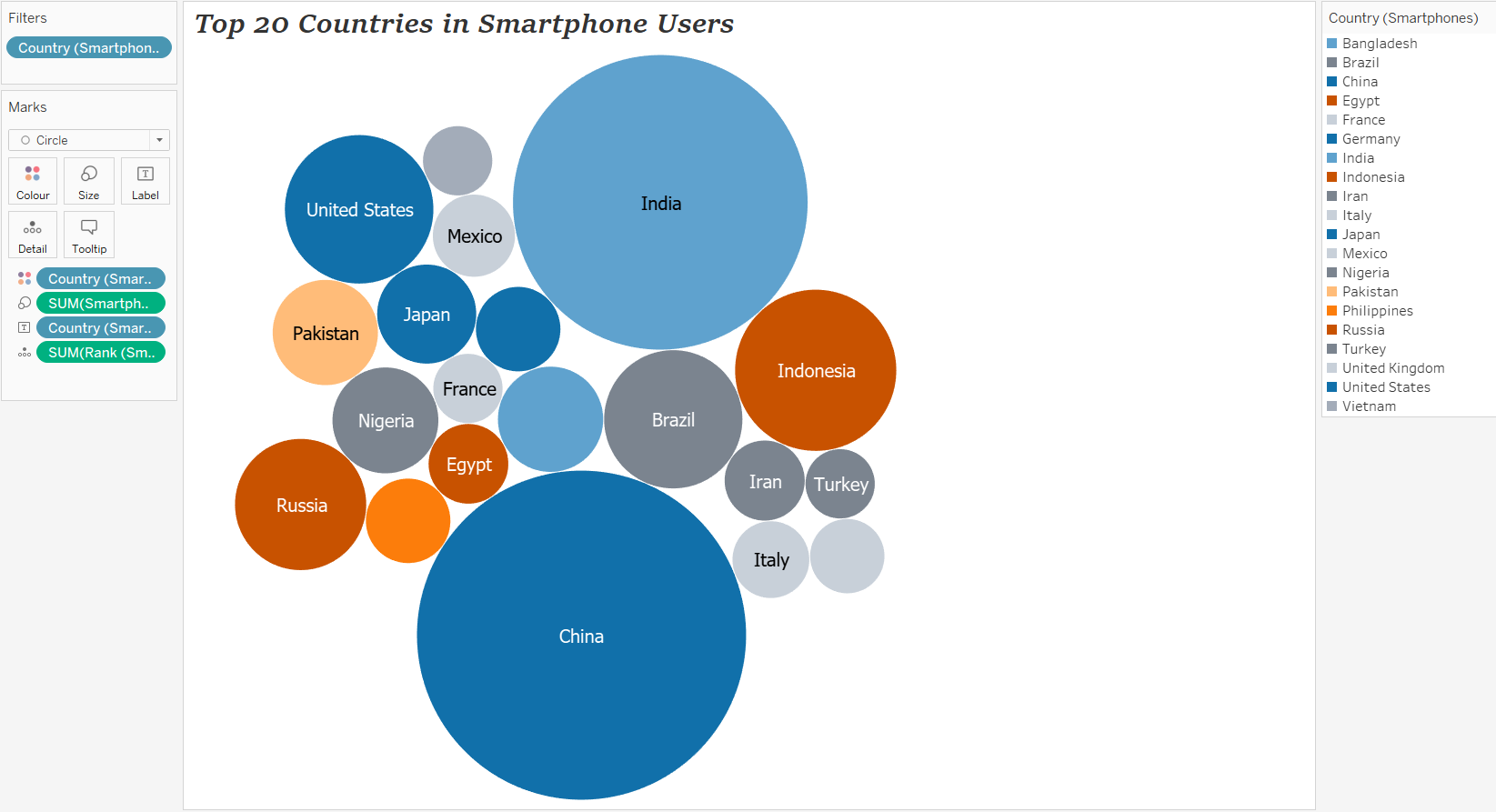
I dragged Country(GDP) to Filters, and in the pop-up window selected the ‘Top’ tab and selected the following options demonstrated in the screenshot. Applied it and selected ‘Ok’.



This is the result of my first visual. I formatted the chart by changing order and visuals; I adjusted colour of the colour palettes to Tableaus built in colour-blind palette, showed ‘mark labels’, put ‘Rank’ in the marks field under Label so the rank is also displayed, I also changed the font to Tahoma.

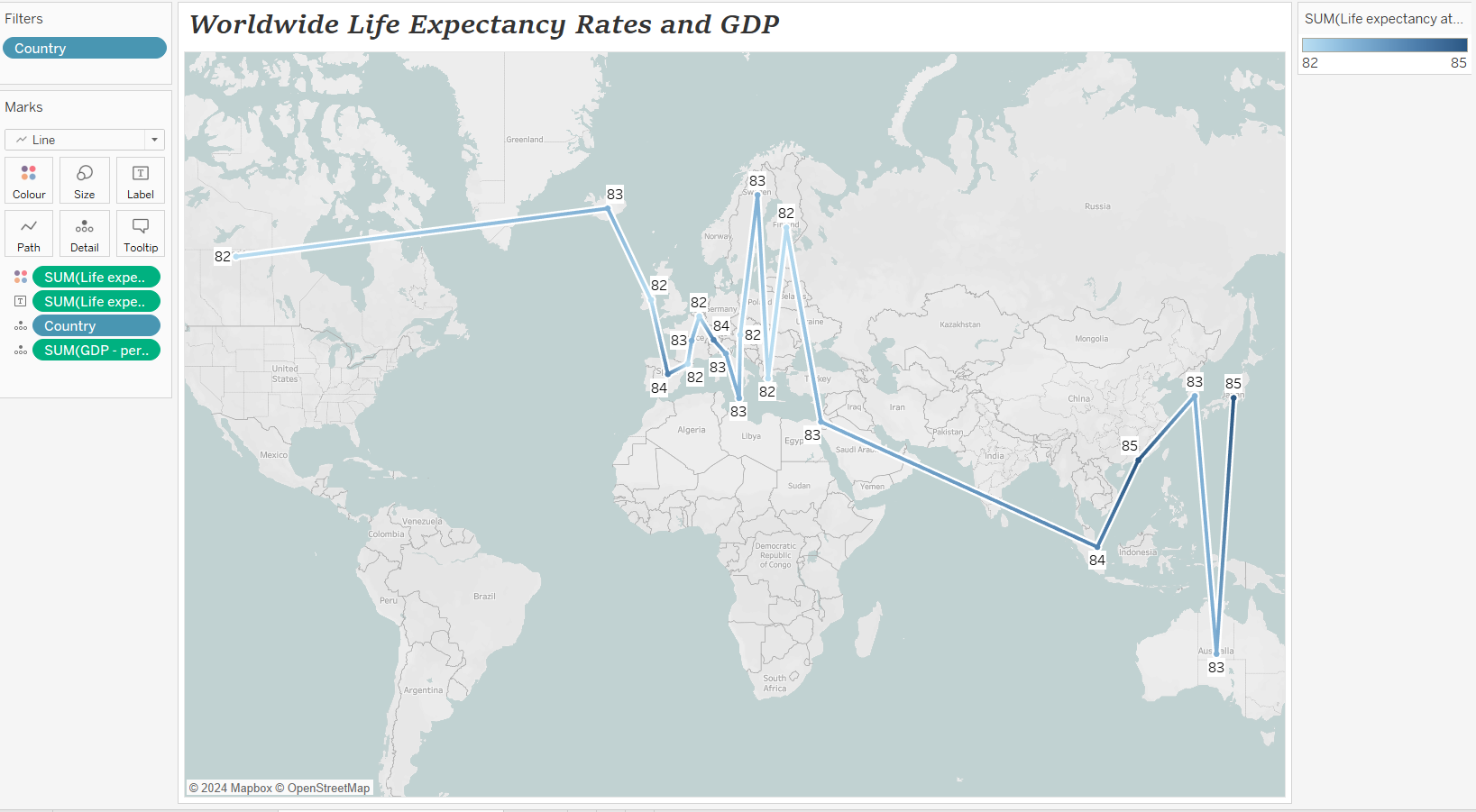


*Second visual: Top 20 Countries In Smartphone Users*



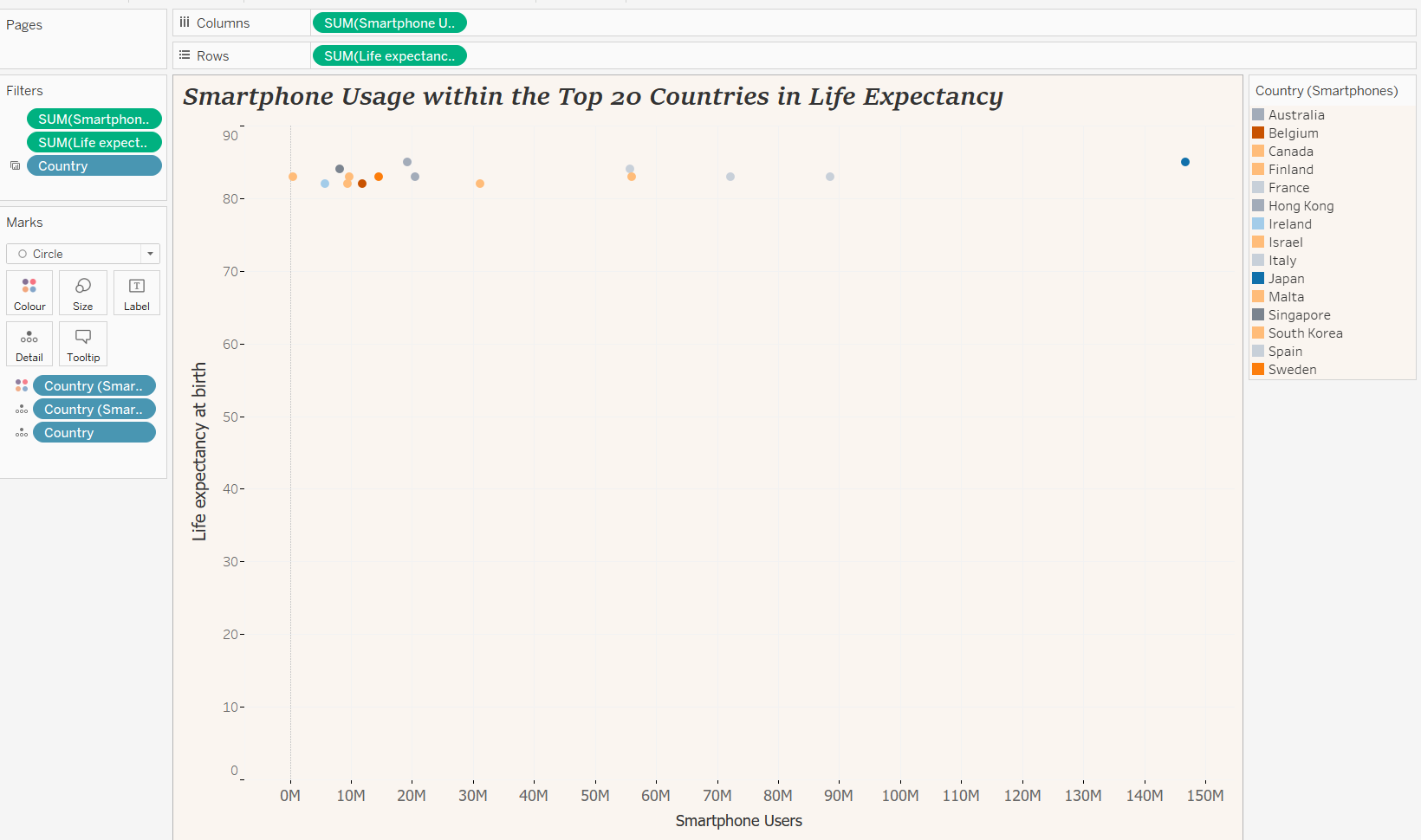
For this visual I took the data from the smartphones sheet on country and smartphone users, I filtered for top 20 countries of smartphone users, changed the style to a bubble chart as I think this told the visual story of the data better, changed the colour scheme to ‘Colour blind’, I added a white border around the bubbles and finally I included rank in the Marks field so the Rank number shows when hovered over.

*Third visual: Related Life Expectancy at Birth*



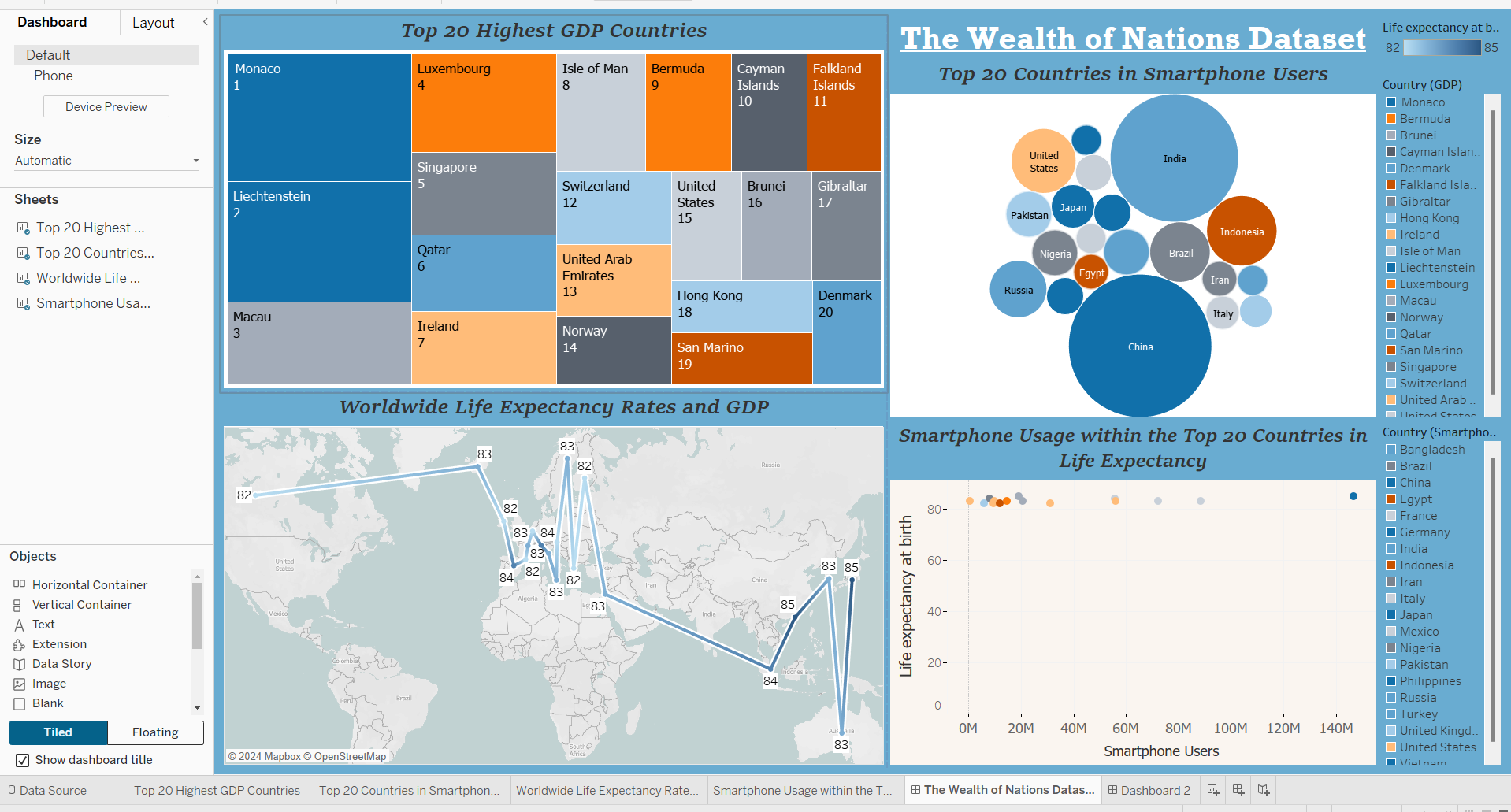
For this visual I filtered the top 20 countries for life expectancy and I chose a map visual. There was one unknown value, which was ‘Arab States’, thus I corrected to United Arab Emirates.   
I edited the map layout to ‘Light’ and then changed the dotting on the map to ‘Line’; I think this represents the regions with higher Life Expectancy rates in a more visually interesting way. Additionally, I included the GDP rates for those countries as well as can be viewed in the Marks field, alongside changing the colour to be suitable for someone with colour-blindness, adding a white ‘halo’ around the line, and putting an opaque white background against the number markings.

*Fourth visual: Smartphone Usage within the Top 20 Countries in Life Expectancy*



For my final visual I looked at the relationship between Life Expectancy and Smartphone users in countries. There were a fair amount of NULL values which I filtered out, I also applied the same filter from the previous worksheet; therefore only the smartphone data of the top 20 countries in life expectancy is shown. I used a scatter plot and formatted it to change the colour of the background to a colour-blind friendly colour, with the dots also adhering to the colour-blind palette. I adjusted the fonts, font sizes, tick marks on the graph, and of course title.

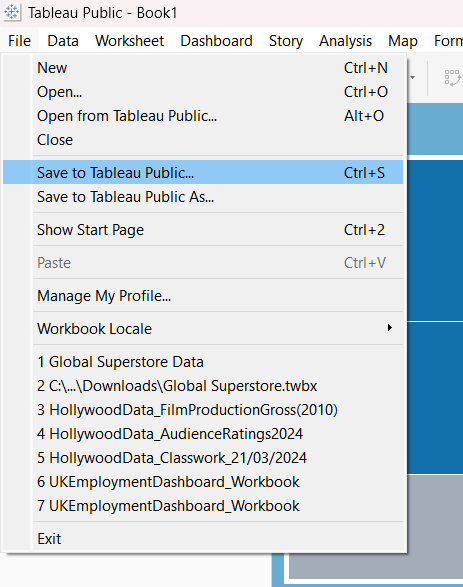
1. Build dashboard:



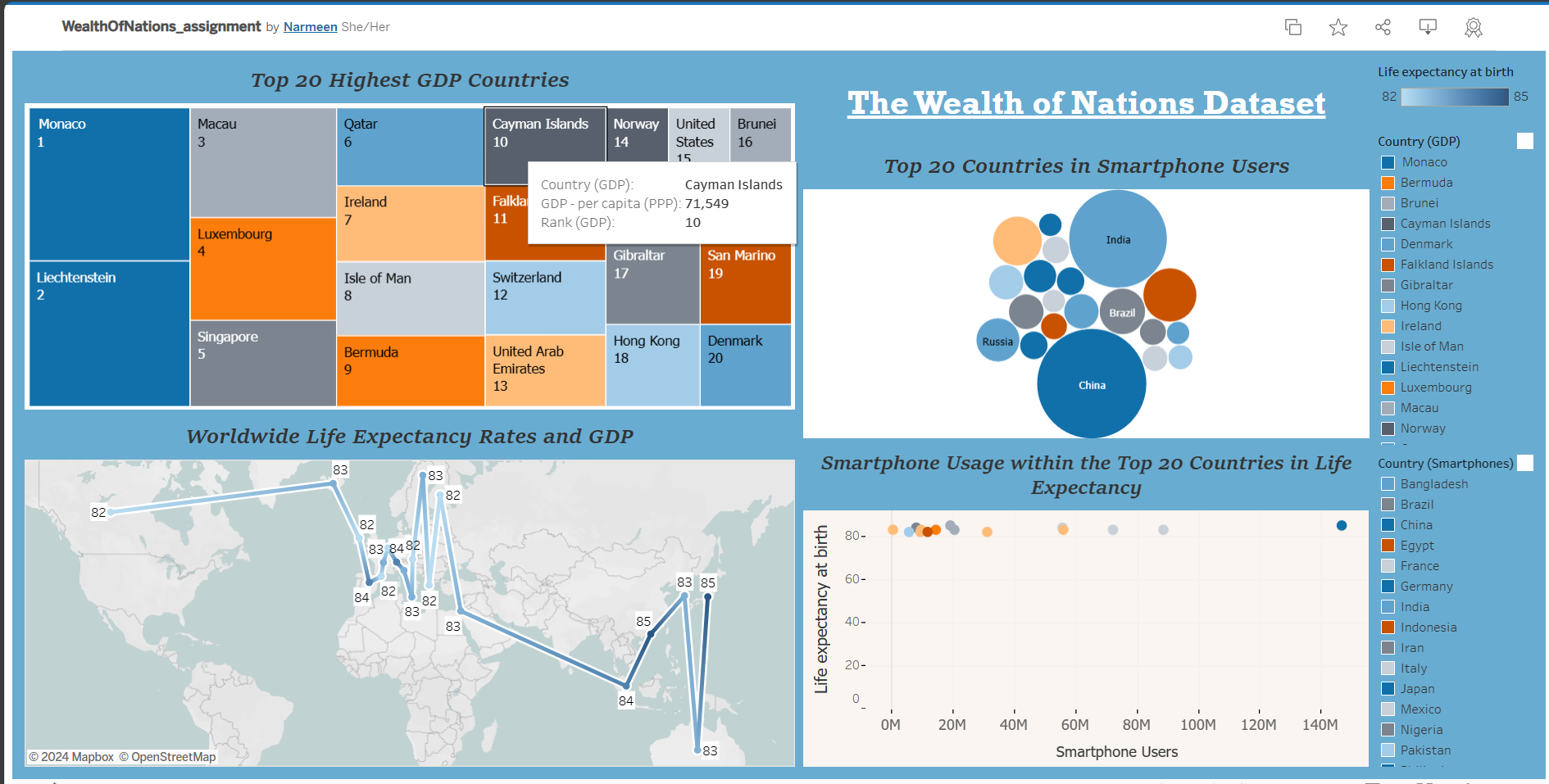
To finalise my assignment, I have created my dashboard!

Incorporating all of my worksheets and adjusting some sizes, fonts, and layout to improve readability and make more visually pleasing. I used a tiled layout as for ease, and included a dashboard title as well. I adjusted the viewing to automatic so the layout isn’t affected much when viewed on various devices, I will also upload to my Tableau page.

By putting all my sheets together in this dashboard, it is easy to see how all the information links across the data, and can be used to tell a data story and provide more insights in the relationships between high GDP, life expectancy rates and smartphone users.



I saved my work.



My work is now available to view on Tableau Public!

[WealthOfNations\_assignment | Tableau Public](https://public.tableau.com/app/profile/narmeen.mohammed/viz/WealthOfNations_assignment/TheWealthofNationsDataset?publish=yes)

**Reflection…**

To reflect and evaluate on my project I will discuss what went well, what I struggled with, what I learnt and areas to develop on.

What went well…  
In terms of what I think went well, applying my newly learnt knowledge to this project showcased to myself the things I have learnt and what I am capable of; inevitably boosting my confidence in Excel and Tableau in a way that I am far less afraid to explore these programs and play around with different features.

What I struggled with…  
Something I found myself struggling with was recalling the things I had learnt during this project; I often found myself having to research/google how to do things I already learnt because I wasn’t too sure on how to do those things. Additionally, there were some more complicated things I wanted to do (such as with the tableau task, I wanted to create some more complex charts). However, this helped me realise my weak points and what I need to work on; for example, the concept of data types, measures, dimensions, continuous and discrete are all things I definitely realised I need to get a better grasp of.

What I learnt…  
As noted above, the challenging parts of the task which I struggled with taught me the areas I need to focus more on. I also learnt how researching/googling how to do things plays a large part in learning how to use programs and tools. Nonetheless, I learnt that I’ve learnt a lot and I am capable enough to keep on learning more.

Areas to develop on…  
I need to work on understanding the logistics of data better -data types, and how I can use my understanding of data types to my advantage. I also need to practice more and continue building up my skills on the programs I am using; to not lose what I have learnt and to build on what I currently know.

End of report.

1. <https://www.consilium.europa.eu/en/policies/data-protection/data-protection-regulation/#:~:text=The%20GDPR%20lists%20the%20rights,his%20or%20her%20personal%20data> [↑](#footnote-ref-2)
2. https://www.tableau.com/learn/articles/data-management-vs-data-governance [↑](#footnote-ref-3)
3. https://www.edpb.europa.eu/sme-data-protection-guide/data-protection-benefits-for-you\_en [↑](#footnote-ref-4)