4/6/2023

Jonathan Theaker

Data Visualization

Table of Contents

**Policies and Procedures3**

GDPR3

Security5

**Excel6**

**Tableau15**

**Reflection20**

**Policies and Procedures – Mindfulness when handling data**

When working with data you need to follow company policy as well as legal requirements.

Following correct procedures for securely storage and usage of data is important so that you do not break the law and use the data in an ethical manner for a specific use. It is also important that you do not leave the personal data you are using in a vulnerable position to be stolen.

Some key issues to keep in mind are as follows:

* GDPR
* Phishing
* Social Engineering
* Ransomware

**GDPR**

The Data Protection Act 1028 is a legal requirement to follow when storing or using personal data about other people. Taken from the government website, the below is a short and very simplified set of rules to follow when using data. Source: UK Government (27/03/23) <https://www.gov.uk/data-protection>

Everyone responsible for using personal data has to follow strict rules called ‘data protection principles’. They must make sure the information is:

* used fairly, lawfully and transparently
* used for specified, explicit purposes
* used in a way that is adequate, relevant and limited to only what is necessary
* accurate and, where necessary, kept up to date
* kept for no longer than is necessary
* handled in a way that ensures appropriate security, including protection against unlawful or unauthorised processing, access, loss, destruction or damage

There is stronger legal protection for more sensitive information, such as:

* race
* ethnic background
* political opinions
* religious beliefs
* trade union membership
* genetics
* biometrics (where used for identification)
* health
* sex life or orientation

There are separate safeguards for personal data relating to criminal convictions and offences.

**Your rights**

Under the Data Protection Act 2018, you have the right to find out what information the government and other organisations store about you. These include the right to:

* be informed about how your data is being used
* access personal data
* have incorrect data updated
* have data erased
* stop or restrict the processing of your data
* data portability (allowing you to get and reuse your data for different services)
* object to how your data is processed in certain circumstances

You also have rights when an organisation is using your personal data for:

* automated decision-making processes (without human involvement)
* profiling, for example to predict your behaviour or interests

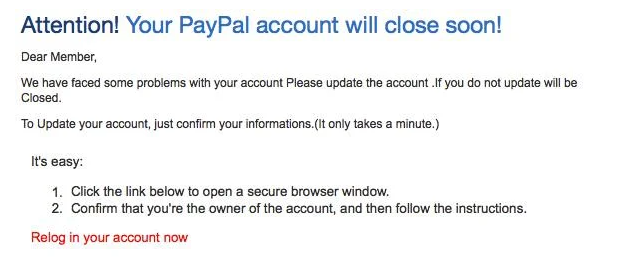
**Security**

You have a legal responsibility to keep other people’s personal data secure so that they are not used by an external party without consent. You must also make sure private data or company secrets are not stolen.

**Phishing and Social Engineering**

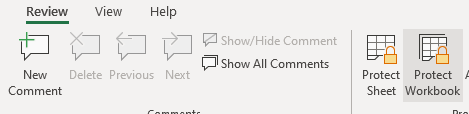
Phishing is a term used to describe when someone is trying to steal your data or infect your computer with malicious software by pretending, they are someone they are not, usually a reputable company. For example, you may receive email from someone posing to be from amazon and asking you to click a link which can put you in danger. As some examples, they may attempt to gain access to your computer, scam you out of money, steal your banking information.

Below is an example of a phishing email posing as PayPal:

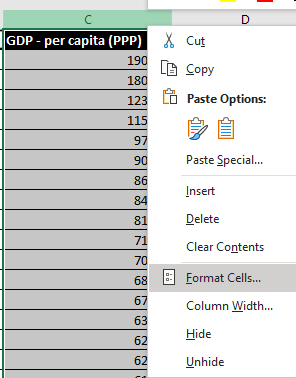


Social Engineering is where a person tries to gain access to confidential information by exploiting human error, judgement and emotions. A human is often easier to crack than a computer. Because of this, it is very important to follow security procedures and not leave yourself open to being tricked by a malicious agent. An example of this is Phishing that uses your trust of a brand name to manipulate you.

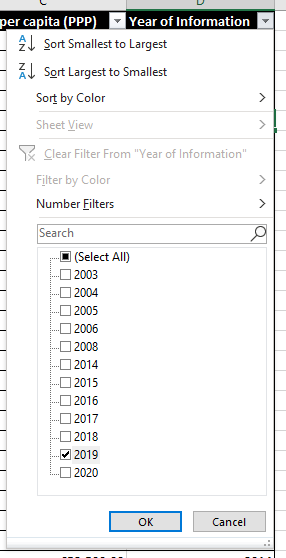
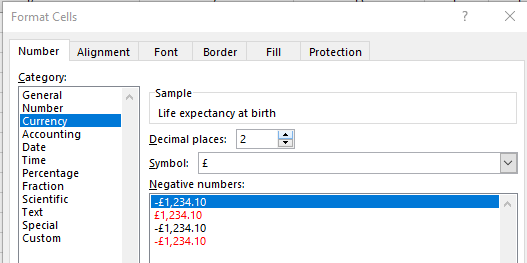
**Excel**



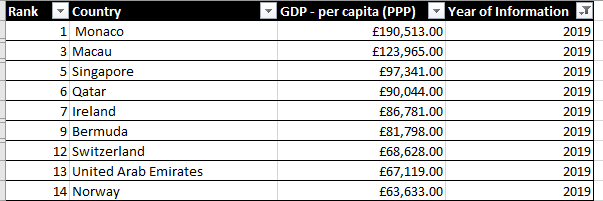
I set a password to protect the workbook. This helps only let others who have a legitimate reason to access the data to be able to do so. I have a duty to protect sensitive information.



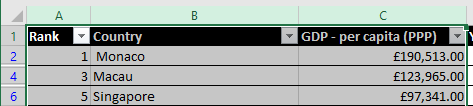
After clicking on Column C, I then formatted the cells to be ‘Currency’ using the ‘£’ symbol. This makes it clearer that the cell contains currency, and of what type. This is important to have correct as each currency has different exchange rates.



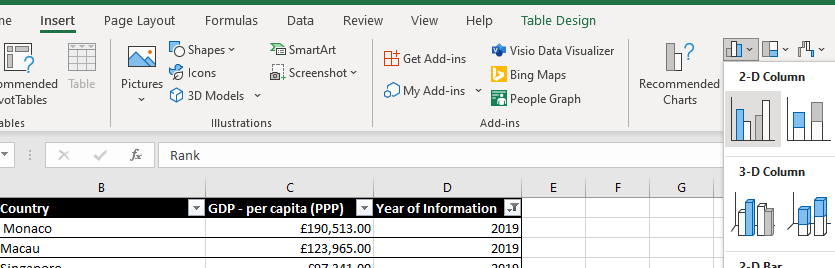
To filter the table to only show data from 2019, I clicked the drop-down menu for the Year of Information column, deselected all options by clicking select all, then selecting just 2019. We will only be using data from 2019.



I then created a chart to display ‘Rank, Country and GDP – per capita (PPP). Since I have only selected data to be shown from the year 2019 the rank will not really be relevant in most cases and will be something to be mindful of, as it still will include all dates in its rankings. We cannot use the data as it is to show the ranking of each country just for the year 2019.



I first selected all columns I will be using.

I then created a 2d bar chart under Insert > Charts > 2d column Chart  


As there are many countries, this chart will have to be incredibly wide to be readable, showing each individual country.

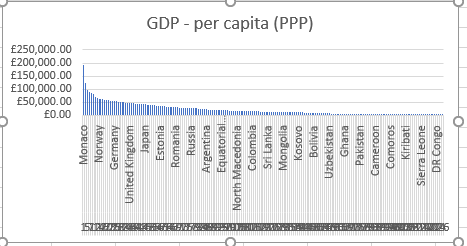
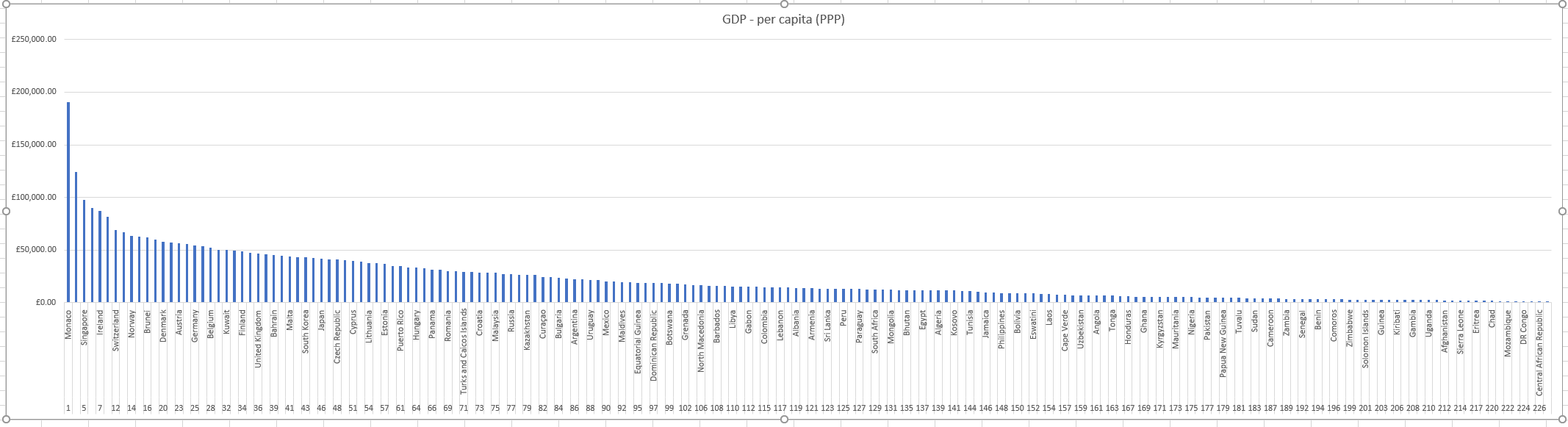


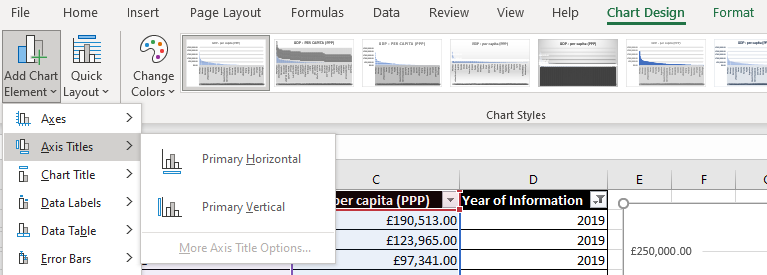
Chart before I widen it by clicking, holding and dragging the chart handles.

After:

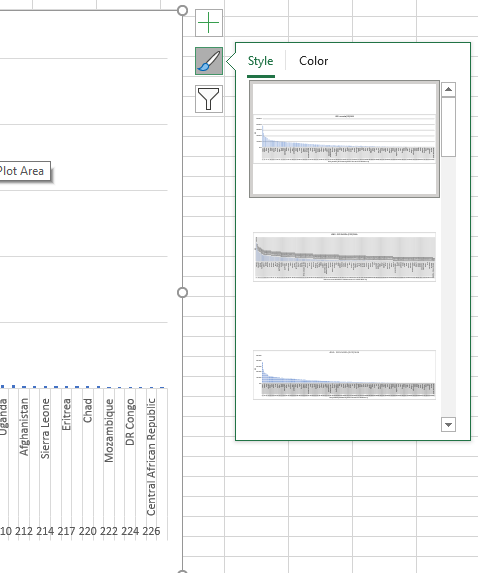


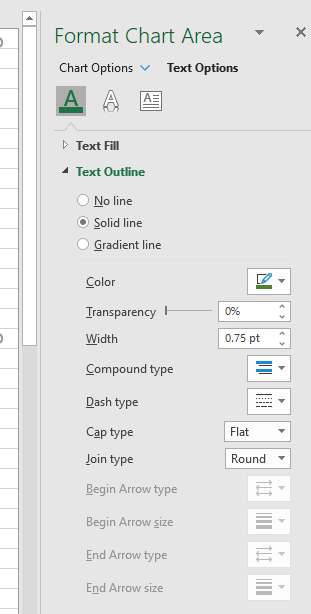
I edited the title by clicking it and including a note that it only shows 2019 data.

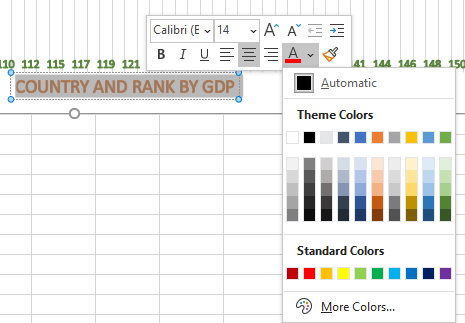
I added X and Y labels to the chart by going to ‘Chart Design’ while I had the chart selected.



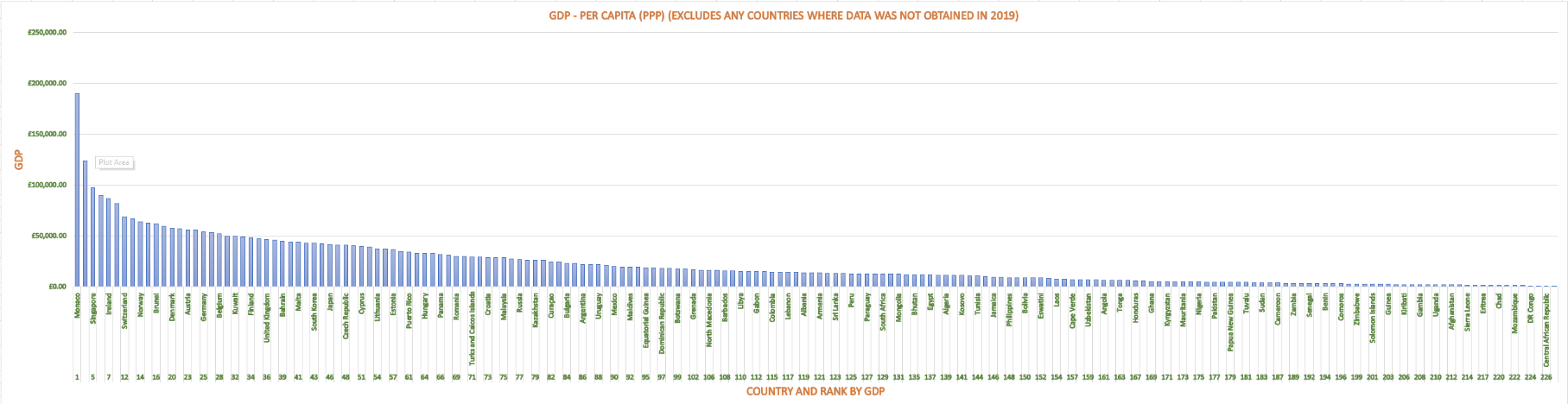
I then changed the style of the chart to one of the pre-sets to make it look more appealing.



I also increased the font size of my titles to make them clearer. Under chart options on the right, text options, I changed to Solid Line and changed the colour to green. I individually changed the titles of the X and Y axis to orange so that it is easier to distinguish the titles from data.



My final chart is below:



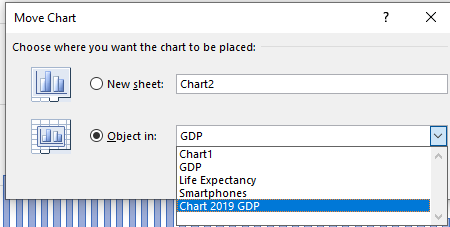
I added a new sheet to separate my chart from the rest of my data by pressing the ‘+’ sign at the bottom of the workbook.



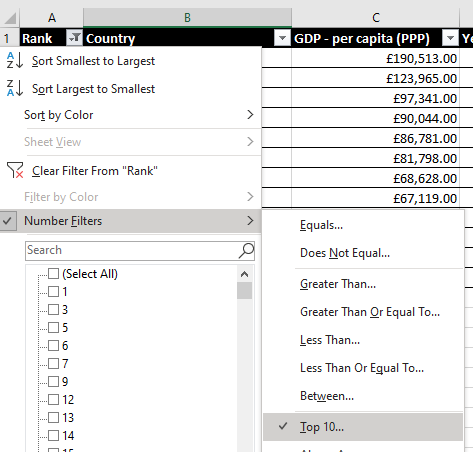
I then double clicked the new sheet to rename it ‘Chart 2019 GDP’, so I can easily tell what is contained on each sheet.

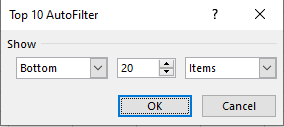


To move the chart here, I right clicked my chart I created on the GDP sheet, and moved it to my new sheet. Alternatively, I could have copy/cut+paste the chart onto the new sheet.

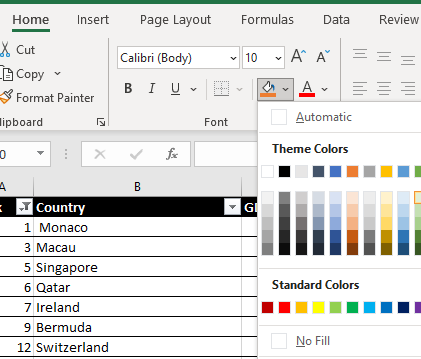


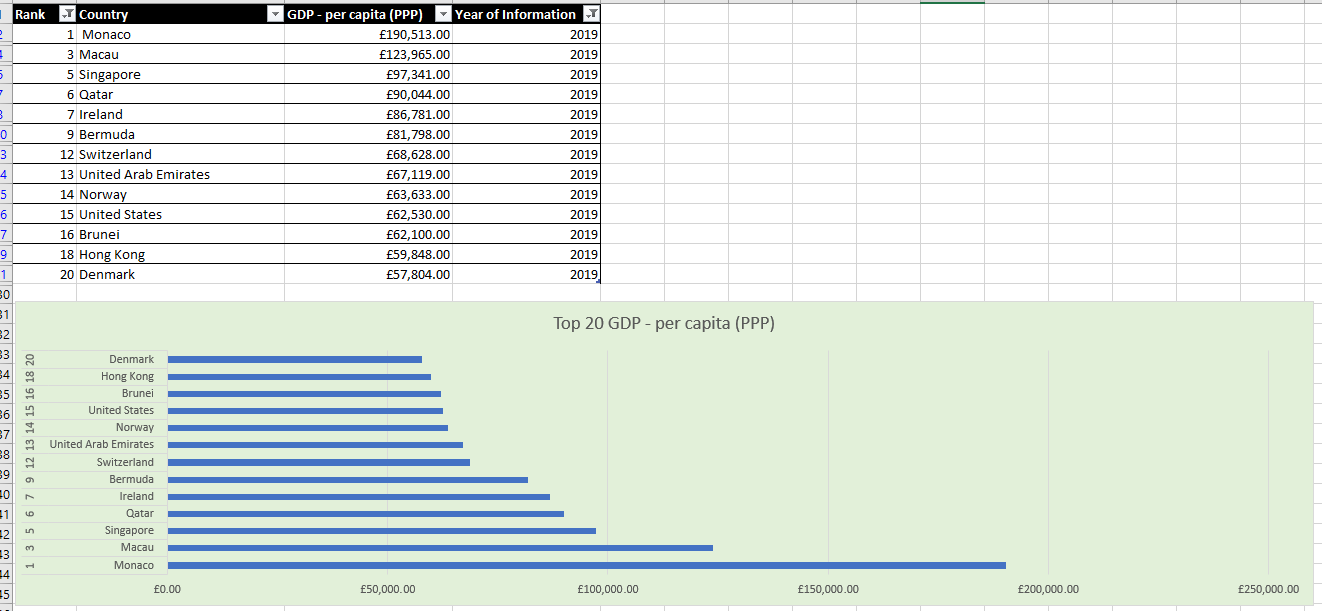
I filtered my data to show only the top 20 countries by rank. (in this case I use the filter for the bottom 20 items, as the lower the number, the higher the ranking)

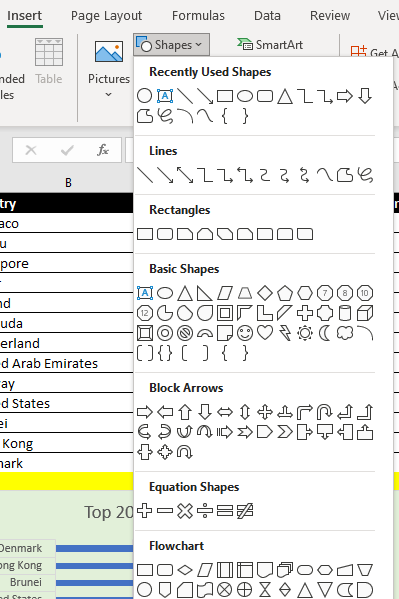




I then created a 2 bar chart to represent the filtered data and added a light green background to the chart.

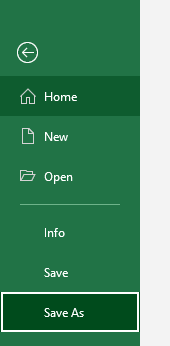


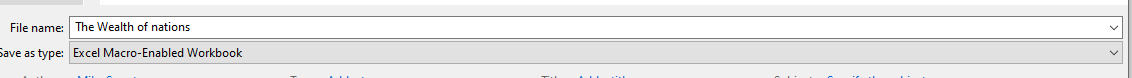


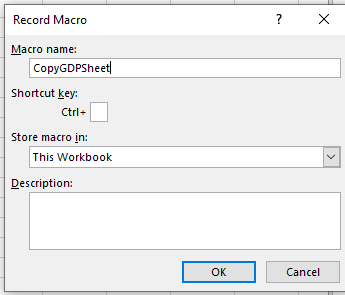


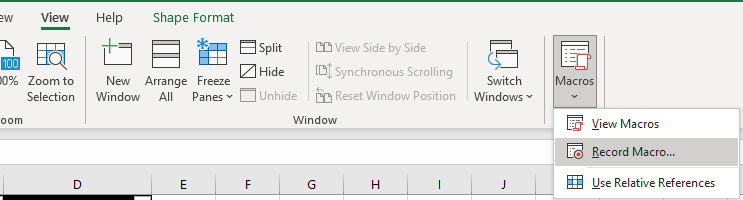
I then created shapes with text I typed inside to which I will apply macros to



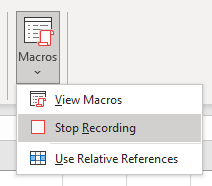
In order to use my macros when I open this file in the future, I need to save as a Macro-Enabled workbook.

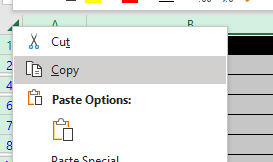


To create a macro, I recorded a macro in the View tab and gave it a name.

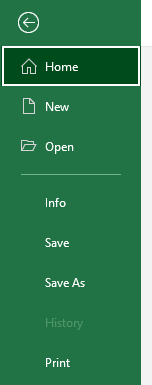


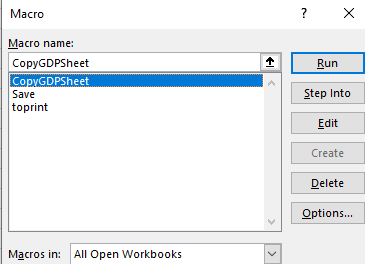
Once I click ok, the macro is now recording any action I take so that it can copy it later. As I want to copy the sheet, I select the top left of the sheet to select it all, right click and copy.

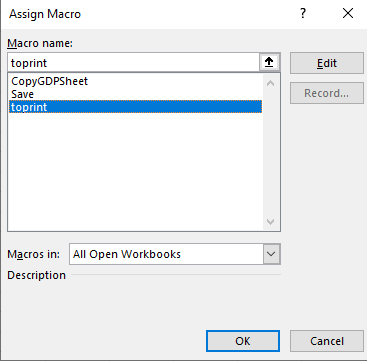


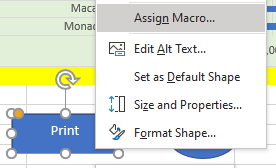


Once this is done, I select the macro drop down menu again and stop the recording.

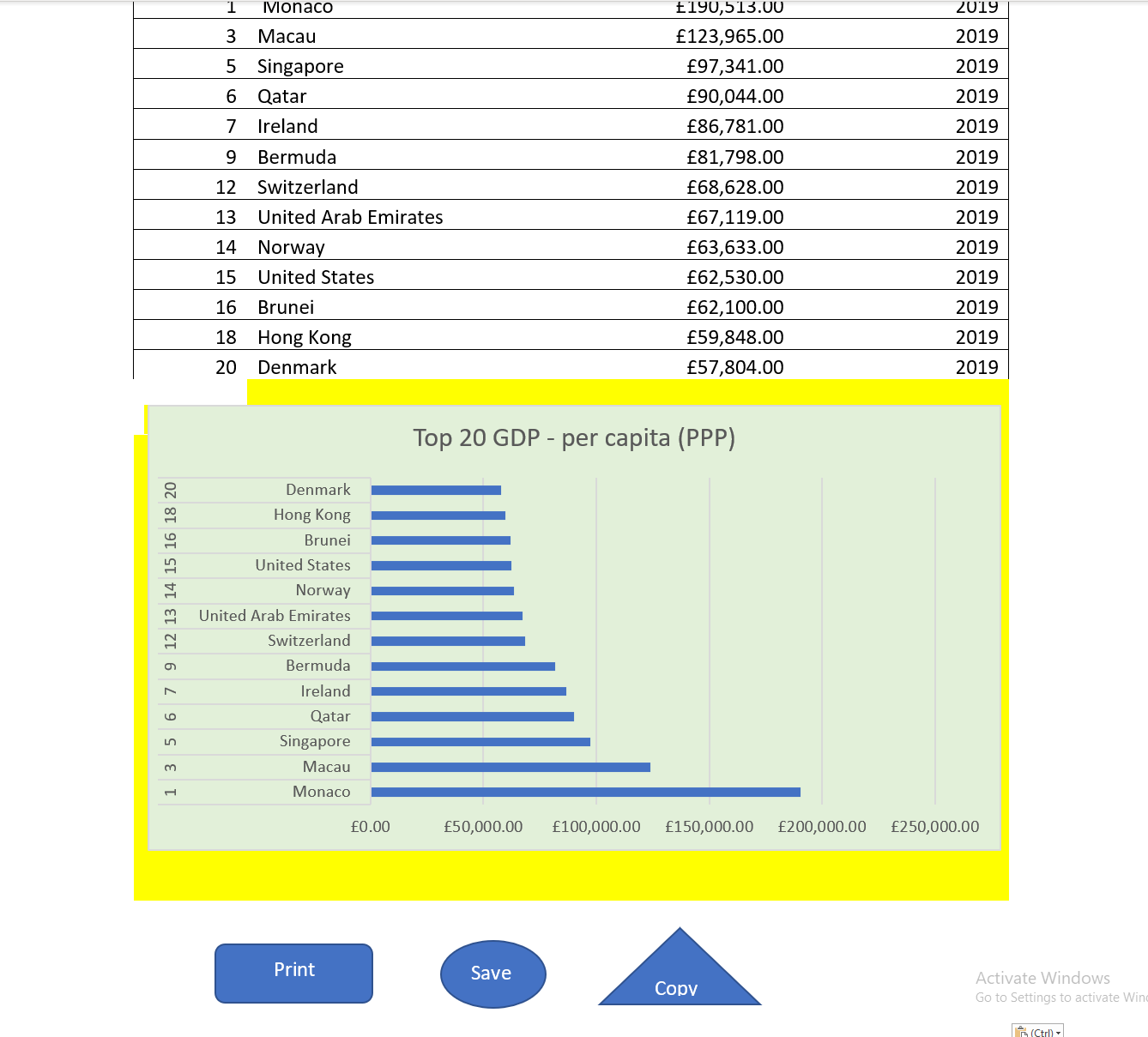
I repeat this process to create Save and Print macros.



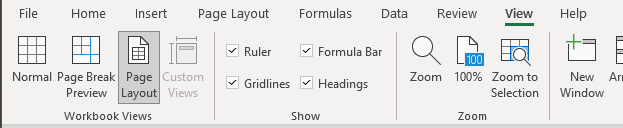
I then assigned the macros to the corresponding buttons



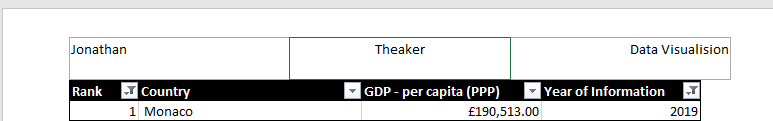
I then used the copy macro to copy the sheet, I then pasted it into a new word document using Ctrl+V.



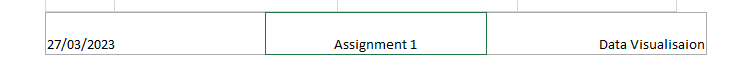
I then went back to my excel work book to add a header and footer to it, under View > Page Layout. Adding text to the boxes that appear at the top and bottom of my page.



Header



Footer

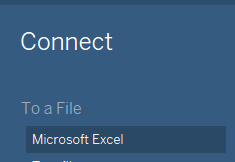
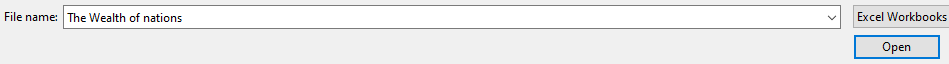


Finished, I then saved my work done on the excel spreadsheet.

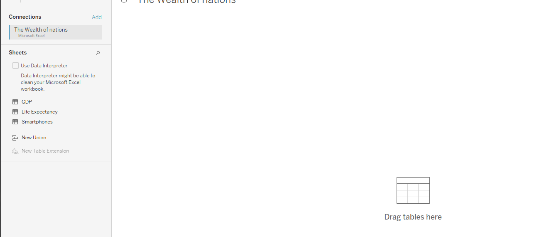
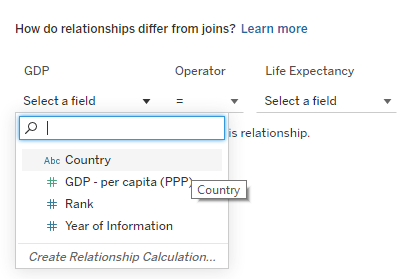


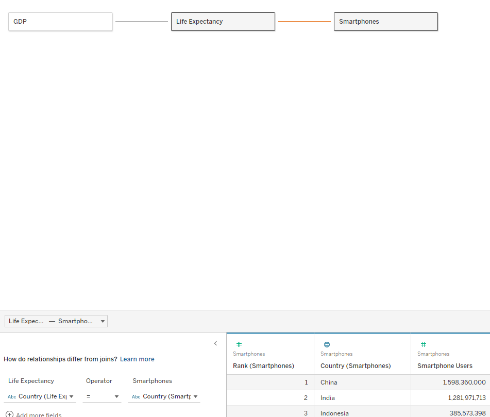
**Tableau Visualisation**

Importing data



Setting relationships. Dragging the sheet’s I wish to use on the left panel to where indicated in the centre, and then setting their relations to ‘country’ as that is common to all of the tables.





Checking the datatypes of all 3 sheets, they have been automatically set to a type I am happy with. I am not changing the dates column to be of the type ‘Date’ because we are only given the year, not the specific day or month.

**Building Dashboard of Charts**

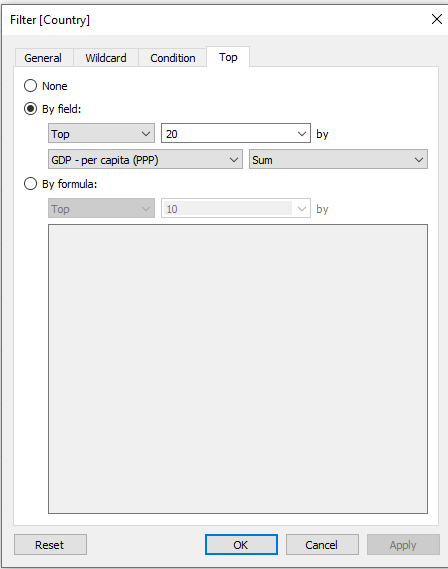
The goal of the dashboard is to show data about the top 20 ranking countries.

The client only wants information from the top 20 highest ranking countries, but they did not specify highest ranked in what subject or anything more specific, so I will start off with top 20 GDP.

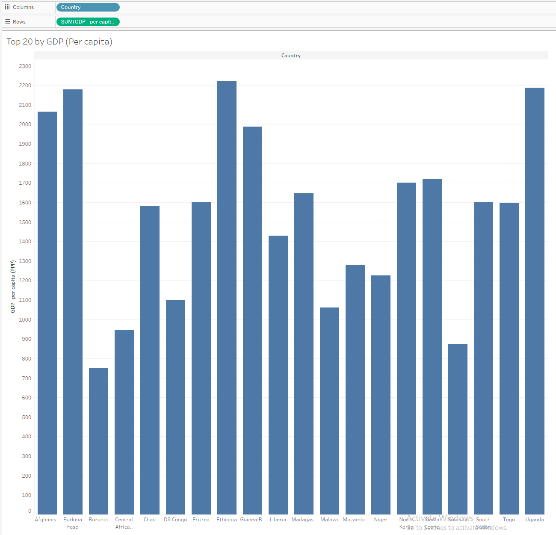
I will make individual visualisations I wish to use on separate sheets before I figure out the best way to present them onto one dashboard/page.

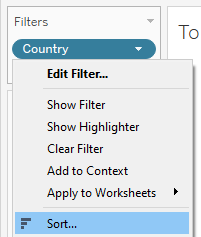


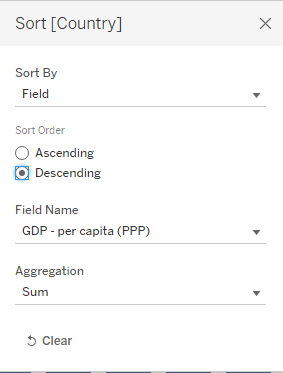
As I only want to work with the top 20 ranked countries, I drag ‘Country’ under ‘GDP’ to the ‘Filter’ selection and set it to show the Top 20 countries by GDP.

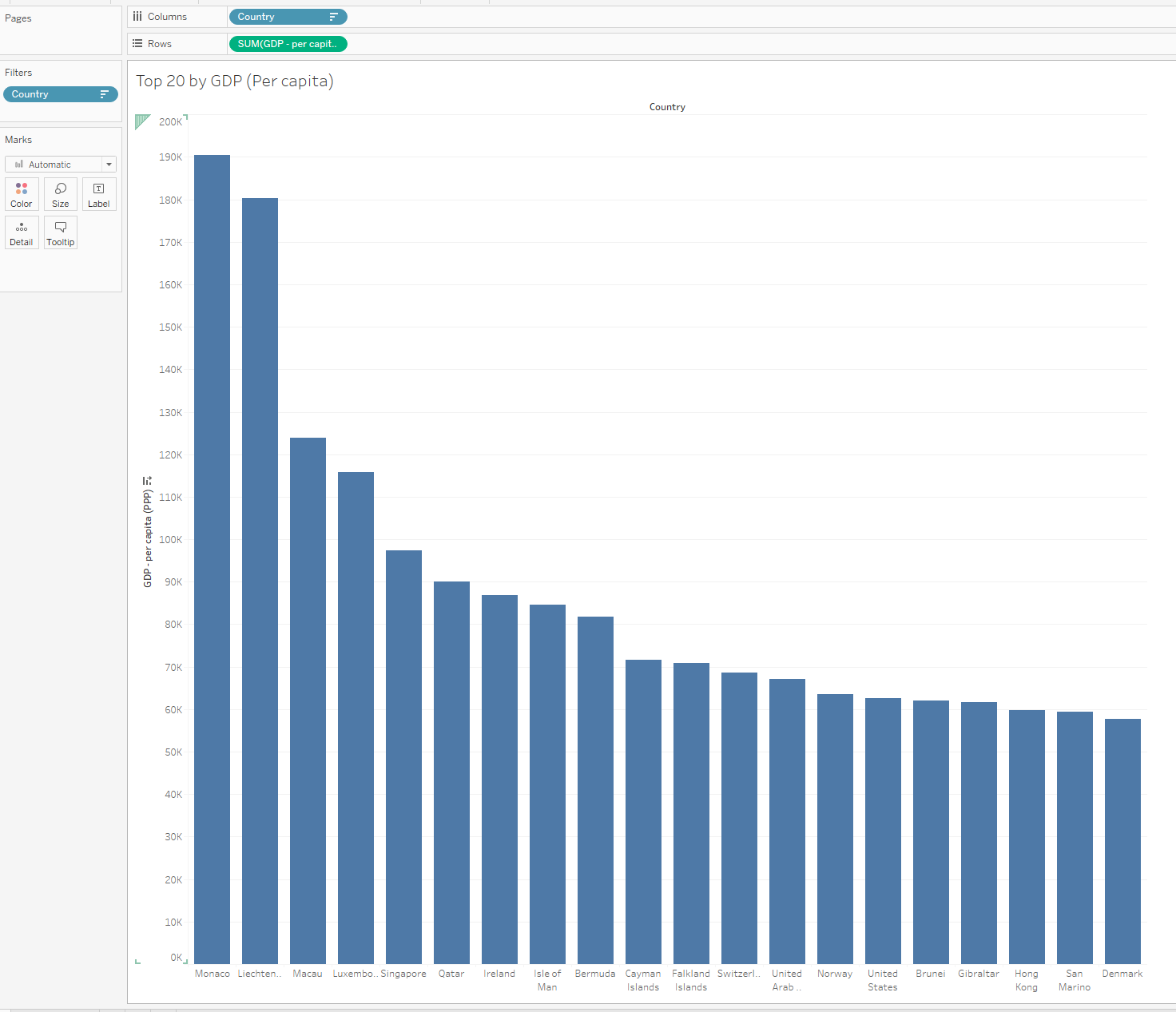


I drag ‘Country’ to the ‘Column’ and ‘GDP – per capita’ to row. I then rename the sheet to reflect what the data is showing.



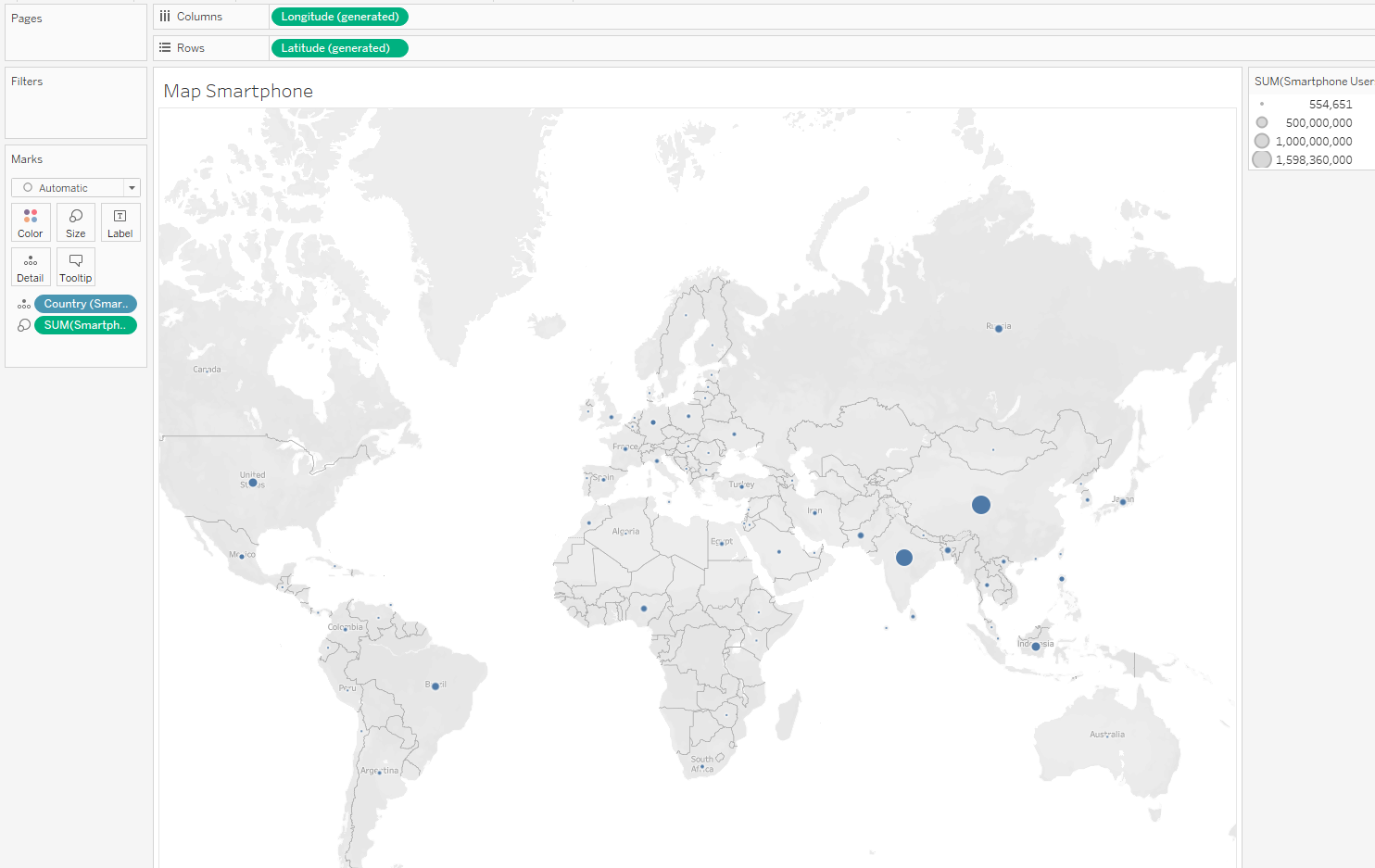
I then sort the data from highest to lowest by country to better visualise the bests performing countries.





I repeated the above process, but with Smartphones replacing GDP on a second sheet.

On a new sheet, I created a hotmap of smartphone usage, dragging the ‘Country’ and ‘Smartphone Users’ form the ‘Smartphone’ table.



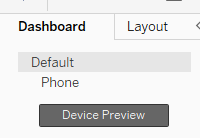
I then did the same as above but for GDP instead of Smartphones. I am not sure if I will use all of these, but having the sheets with the data available to me can be useful if I think of a way to use them or process them further.

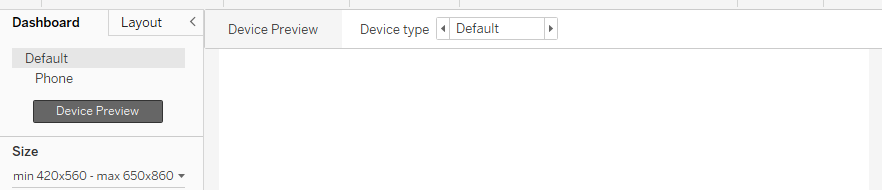
**Dashboard**

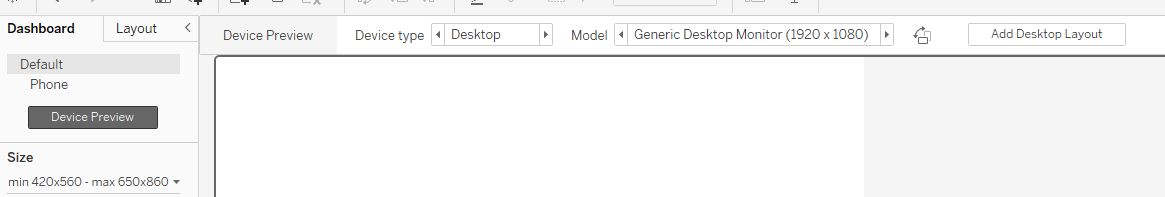
I have to add the dashboard tab at the bottom.



I created a new dashboard layout under ‘Device Preview’ to display for a Generic Desktop Monitor (1920 x 1080) then click add desktop layout. This should fit typical computer screens well, and without first doing this it will be very difficult to work with the impossibly small, default size of a Tableau dashboard that does not seem large enough to fit any kind of device other than a phone.







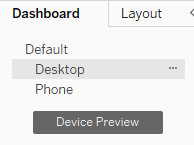
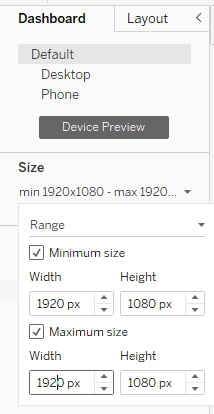
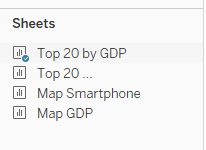
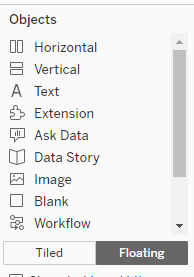


Tableau does not seem to let me drag charts/sheets when on this new device preview layout, so I manually edit the ‘Default’ size to be the same as the typical desktop monitor.

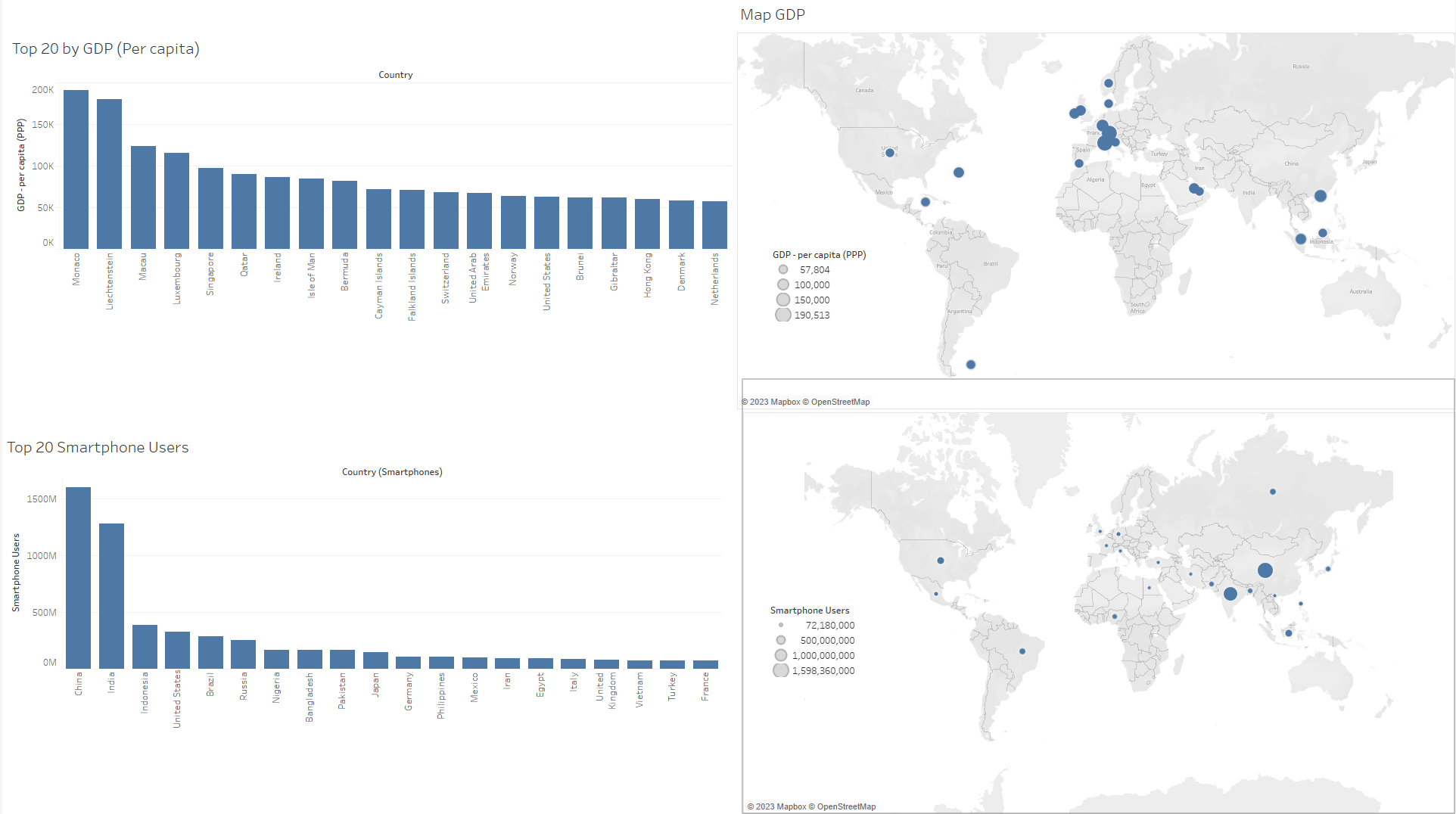


I drag my charts created on other separate sheets onto the dashboard from the left. 

I set objects on the dashboard to be ‘Floating’, this allows me to freely move objects around to fit where I want them to.



After resizing the graphs, I have transferred to the dashboard, the end result was this:



**Reflection**

My main problem when creating my dashboard I wanted a layout reasonably large enough to work with, and one that will fit the display that I would be showing it on. I rectified this by manually typing in the resolution for the dashboard to be displayed on a typical desktop computer screen of 1920x1080, which I believe would be the best size for most cases.

Working out how to apply floating to the entire dashboard was incredibly helpful, instead of doing it for each individual object I could apply to it everything so I could easily move objects around.

There does not seem to be a very strong direct correlation between GDP, Smartphones and life expectancy, so I did not want to create a visualisation relating the two, keeping them separate.

Each country only has data collected for it from a specific year, so filtering any of these out would completely remove the country from the dataset. The gap in years may have changed the values drastically making them poor comparisons, but having to completely remove countries altogether does not seem like an acceptable outcome, as it will not help us draw any conclusions related to ‘Top 20’ countries when we start removing them. Because of this, I have not included any filtering of year the data was acquired even though part of the project was to filter for a specific year.