Data & Data Structures

By: Emma Kaas Andersen

Submission Date: 05/04/2024

Table of Contents

Table of Contents 2

Policies and Procedures 3

GDP Task - Excel 3

- Set a password to protect the workbook 3
- Highlight column C and change the data to display in British Pound Symbol 5
- Turn the GDP sheet into a table 5
- Filter the table to display only the information for 2019 7
- Create a chart that will only display the following data 'Rank, Country and GDP – per capita (PPP)'. 8
- Using your creative skills edit the chart (a. Add a title, b. add X and Y axis label, c. make the chart visually pleasing) 8
- Move the chat to a new sheet tab and label with a suitable name 10
- Create a sort for the top 20 countries 11
- Next create a new bar chart to display the 20 highest ranking countries from your sort and then move the chart to be underneath the table 12
- Colour the background by highlighting the area underneath the table as shown below. Find the add a fill colour icon and select a colour. 13

GDP Task – Tableau 14

- Import Data 14
- Set Relationships 15
- Check Data Types 16

Client Requirements – Tableau 16

- Creating charts 16
- Building the Dashboard 19

Bibliography 20

Policies and Procedures

There are many things that are important to both morally and legally abide to when handling data especially if it is sensitive or personal information. Some of the necessary steps to take when handling data include being in line with data protection law as well as a company's data protection policies which will vary depending on the nature of work at that organization. Organizations will often have their own policies and written procedures to follow when working which will specify more specifically the responsibilities of each person according to their role (ICO, 2023).

It is also important for everyone who is working with data to be aware of different ways that data can be mishandled so that the necessary steps can be taken to keep it safe. It is important to maintain confidentiality by keeping "sensitive information private and secure" (Irwin, 2023). Do not allow anyone who is not authorised to see the data to access it. Marinating the integrity of of the data means guaranteeing that the data remains accurate and protecting it from being corrupted. Finally making sure that the data is available when needed and that it isn't at risk of being lost (for example when the network is down) (Irwin, 2023).

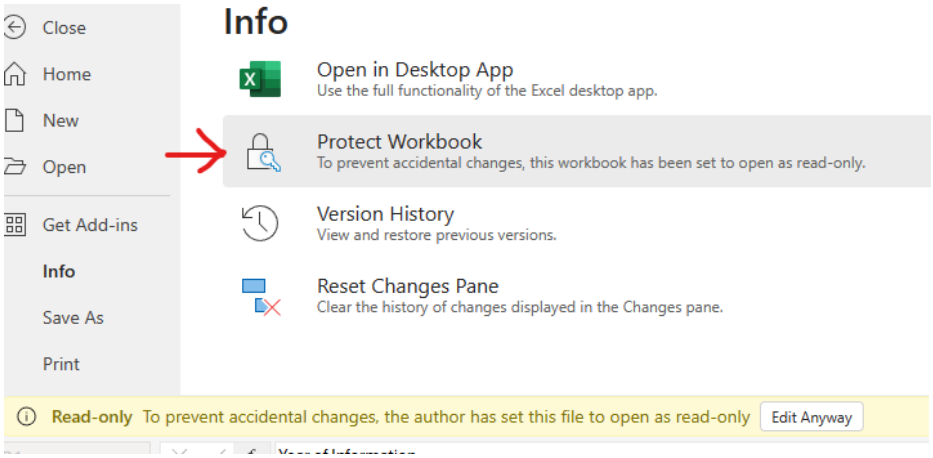
When part of a company it is likely that there is a cyber security specialist who is mainly responsible for keeping information secure, however keeping data secure is the responsibility of everyone regardless of the role. Being aware of and following internal standards and external regulations as a preventative action to avoid any mishandling of data is an obligation for anyone who handles that data (internal standards are good practice and external regulations need to be followed by law) (Agarwal, 2022). Anyone who has access to data has the potential to misuse it, even if they do not do so purposefully, this is why it's important to be aware of data security measures and to follow them.

GDP Task - Excel

Set a password to protect the workbook

It is worth mentioning that password protecting a workbook is not possible using the Excel web version. If you would like to password protect your workbook this is something that can be done in the paid full version app which can be downloaded onto your computer

However, on the online version you can take some steps to protect your workbook!



Protect Workbook

When you click on:

1) File
2) Info
3) Protect Workbook

This will prevent any accidental changes when you open the document. When you close and re-open the document, it will open as 'Read-only', and from there you can click on 'Edit Anyway' to start editing the document again.

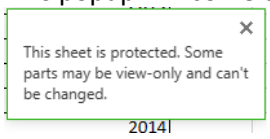
Password Protect Worksheet

You can password protect your worksheet by clicking

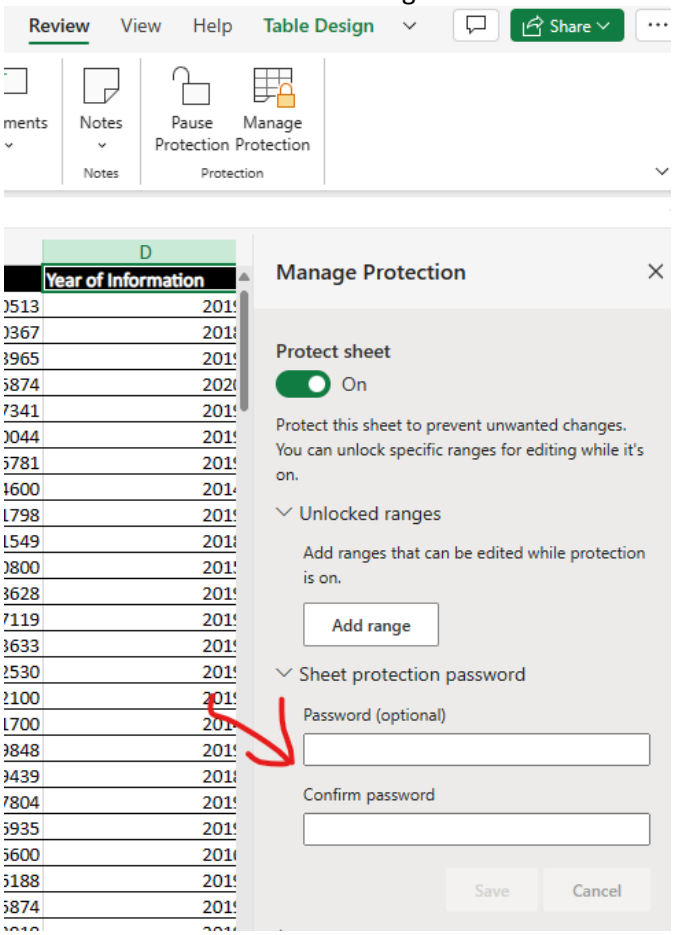
1) Review
2) Manage Protection
3) Click the toggle under 'Protect Sheet' so that it's on
4) Click the arrow next to 'Sheet Protection Password' and in this textbox you can enter in a password to protect your worksheet

Once you've entered a password (make sure it's one you remember as it is not possible to recover the password!) your sheet will have a lock symbol next to it:

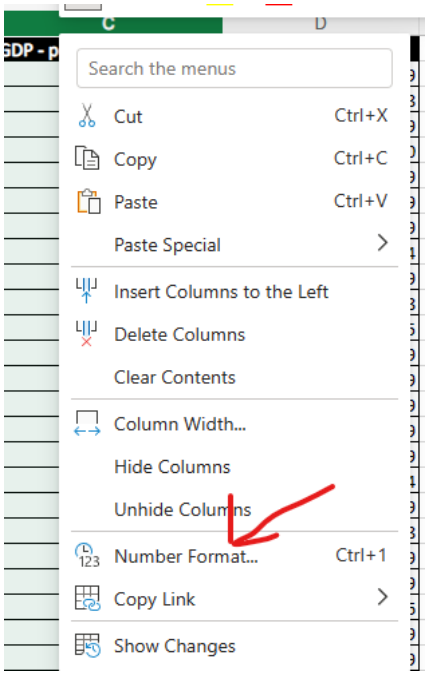
This popup will come up if you try to edit the sheet:



Under the same 'Manage Protections' tab where you set the password is where you can enter your password so that you are able to edit the sheet again.



Highlight column C and change the data to display in British Pound Symbol

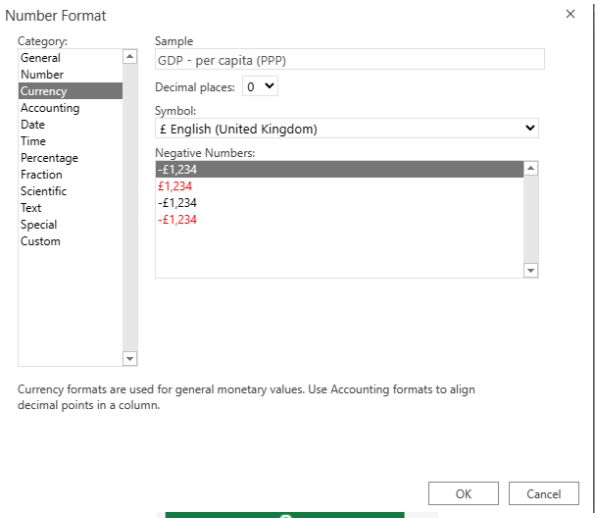


1) Select column C by clicking on it
2) Right-click to bring up the menu
3) Select 'Number Format'

This will bring up the number format menu

Once you have the number format menu you can select the type of data that you have in the column from the categories.

1) Select 'Currency' from the list of categories
2) Click on the dropdown menu under 'Symbol' and select '£ English' which will apply the pound symbol to your data
3) You can also change the decimal places to what works best for your data. For this instance I have picked 0.
4) Click 'OK' when you are happy with your format



Once you have done that the data will have the £ symbol and the number of decimal spaces that were specified.

C
GDP - per capita (£)
£190,513
£180,367
£123,965
£115,874
£97,341
£90,044
£86,781
£84,600
£81,798
£71,549
£70,800
£68,628

Turn the GDP sheet into a table

Data in your worksheet can be turned into a table using pivot tables. This can also be used to isolate certain data on your table and turn them into charts.

Rank	Country	GDP - per capita (£)	Year of Information
1	Monaco	£190,513	2019
2	Liechtenstein	£180,367	2018
3	Macau	£123,965	2019
4	Luxembourg	£115,874	2020
5	Singapore	£97,341	2019
6	Qatar	£90,044	2019
7	Ireland	£86,781	2019
8	Isle of Man	£84,600	2014
9	Bermuda	£81,798	2019
10	Cayman Islands	£71,549	2018
11	Falkland Islands	£70,800	2015
12	Switzerland	£68,628	2019
13	United Arab Emirates	£67,119	2019

Start by highlighting all of your data that you would like to be brought over to the pivot table.
You can do this by selecting the top left box and then pressing ‘shift’ ‘ctrl’ and then clicking the left and then right arrow key. This will select all of your data and it will be highlighted like this.

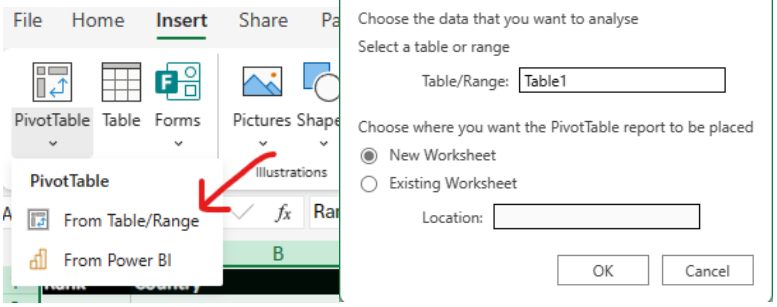
You can then create a pivot table by selecting:

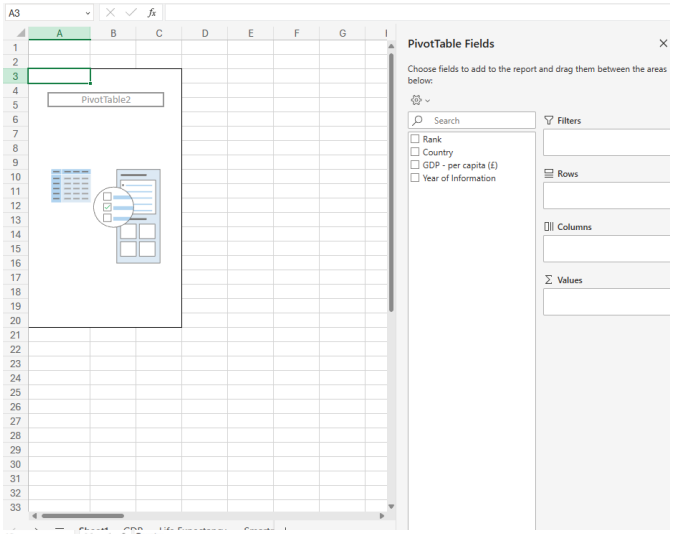
1) Inserts
2) PivotTable
3) From Table/Range (this is the range that has been highlighted)

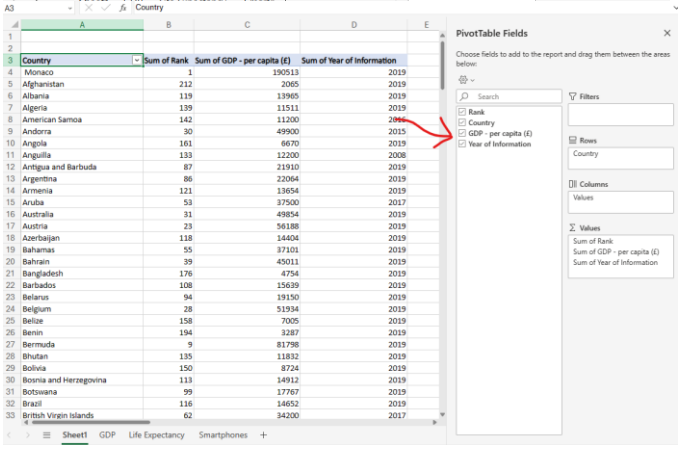
Then a ‘Create PivotTable’ window will pop up.

4) Click on new worksheet
5) OK

This will open a pivot table on a new worksheet







You can find your new worksheet using the tabs at the bottom of the screen.

When you open the tab it will be blank with a menu on the left where you can select pivot table fields.

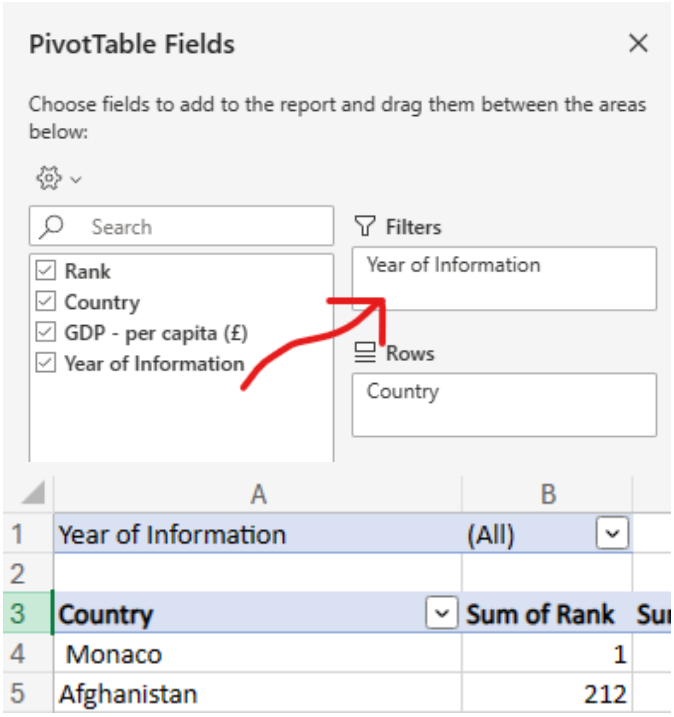
To create a table including all the data from your sheet, click on the checkbox for all the fields and then you will have a table!

Filter the table to display only the information for 2019

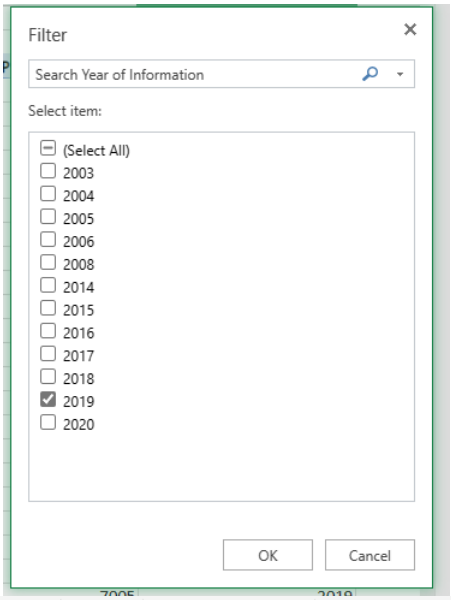
Under your Pivot Table field list you can add a filter so that you are able to filter your data to show what's most helpful.

To only display information from 2019 I first dragged over year of information to the filter box.

There will now show a drop-down menu on your table where you can filter by year of information.



	A	B
1	Year of Information	(All)
2		
3	Country	Sum of Rank
4	Monaco	1
5	Afghanistan	212

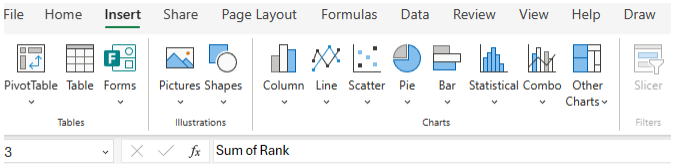


When you click on the dropdown arrow this box will pop up where you can select or deselect the years which you want included in your table. As a default it will select everything so here I have deselected everything except for 2019. Click okay at the bottom of the box when you are happy with your selection.

Now only data from 2019 will show on your table.

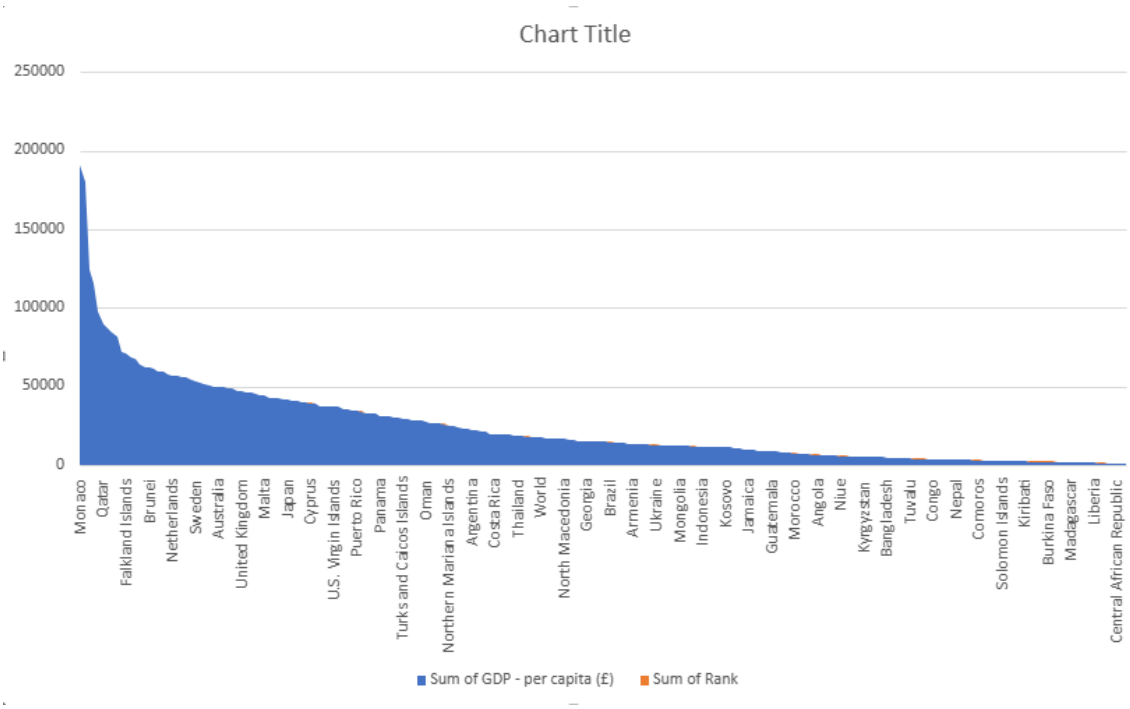
	A	B	C	D
1	Year of Information	2019		
2				
3	Country	Sum of Rank	Sum of GDP - per capita (£)	Sum of Year of Information
4	Monaco	1	190513	2019
5	Afghanistan	212	2065	2019
6	Albania	119	13965	2019
7	Algeria	139	11511	2019
8	Angola	161	6670	2019
9	Antigua and Barbuda	87	21910	2019
10	Argentina	86	22064	2019
11	Armenia	121	13654	2019
12	Australia	31	49854	2019
13	Austria	23	56188	2019
14	Azerbaijan	118	14404	2019
15	Bahamas	55	37101	2019
16	Bahrain	39	45011	2019
17	Bangladesh	176	4754	2019
18	Barbados	108	15639	2019
19	Belarus	94	19150	2019
20	Belgium	28	51934	2019
21	Belize	158	7005	2019
22	Benin	194	3287	2019
23	Bermuda	9	81798	2019
24	Bhutan	135	11832	2019
25	Bolivia	150	8724	2019
26	Bosnia and Herzegovina	113	14912	2019

Create a chart that will only display the following data ‘Rank, Country and GDP – per capita (PPP)’.

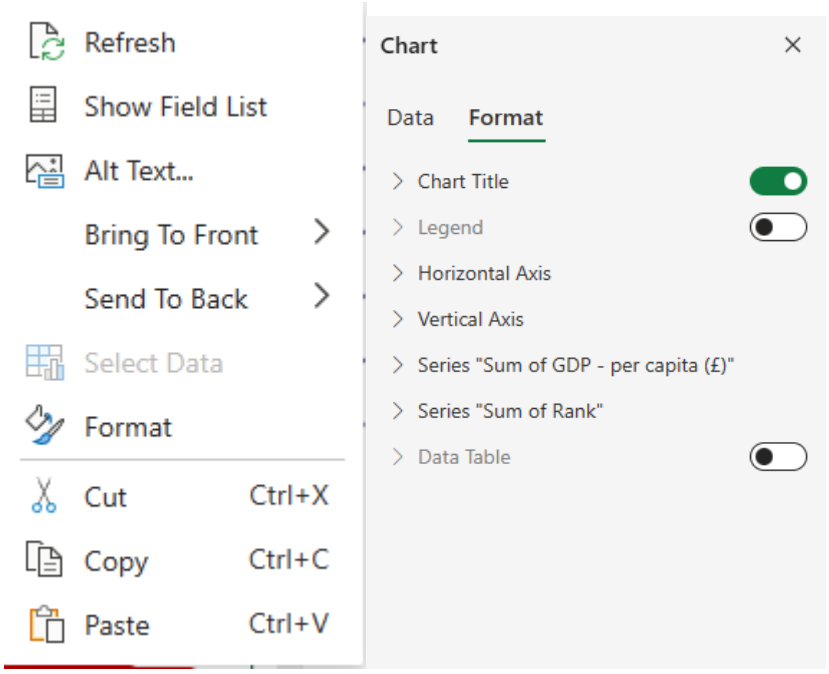


For my chart I went under other charts and selected an area chart as I believe this is the best chart to use for this visualisation.

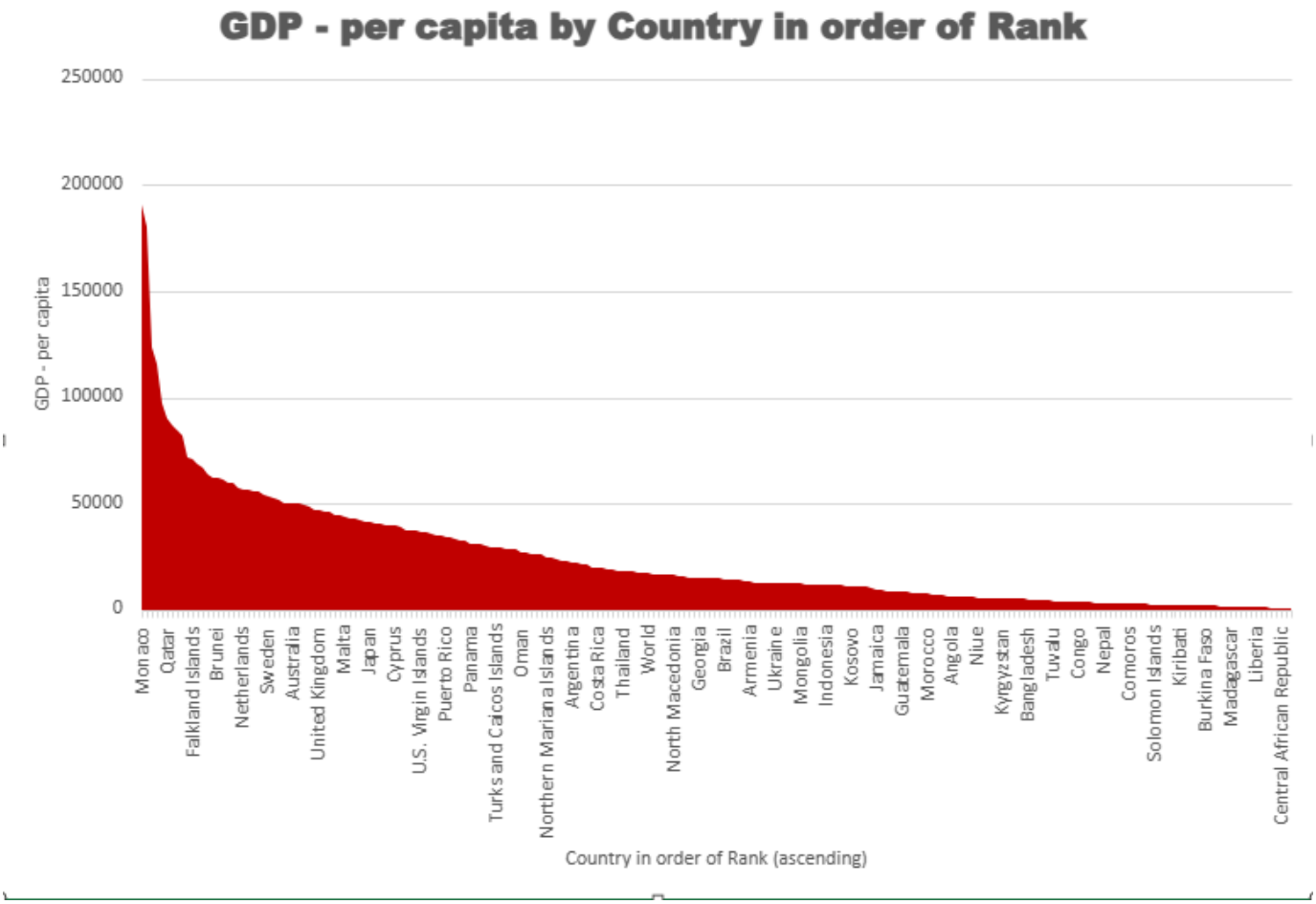
This chart shows the countries on the horizontal axis sorted in order of rank (ascending) and GDP on the vertical axis.



Using your creative skills edit the chart (a. Add a title, b. add X and Y axis label, c. make the chart visually pleasing)

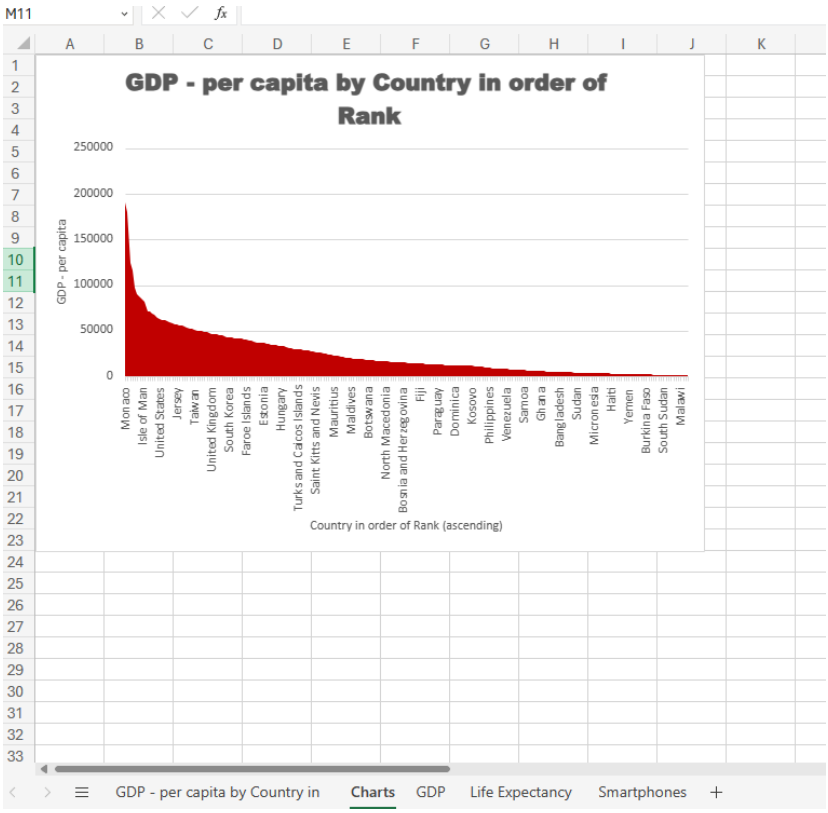


To make the chart more visually appealing and to add more information so that it is clearer, I edited the format of the chart. You can find the options to format the table by right clicking on the chart and clicking on ‘format’. This will bring up a menu of options on the right of the screen where we can edit the chart in each category.



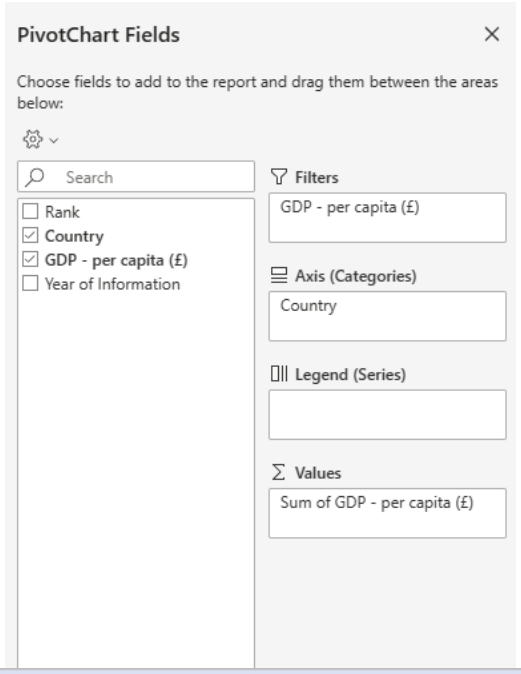
Above is my final chart. I have added a title with a clear and bold font. I have added labels on both axis so that each one is clear, I have also changed the colour of the chart to red

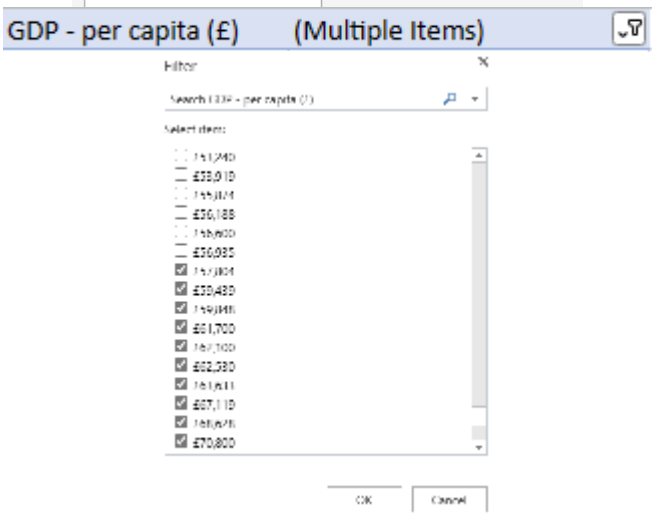
Move the chat to a new sheet tab and label with a suitable name



Now that I am happy with the chart I have moved it to a new sheet. You can add a new sheet by clicking the '+' at the bottom of the page. I copied the chart from the pivot table and added it to this new sheet. I've titled this sheet 'Charts'. Any new charts that I create can be added to this sheet and now I can view all of my charts in one place if I wish to present them.

Create a sort for the top 20 countries



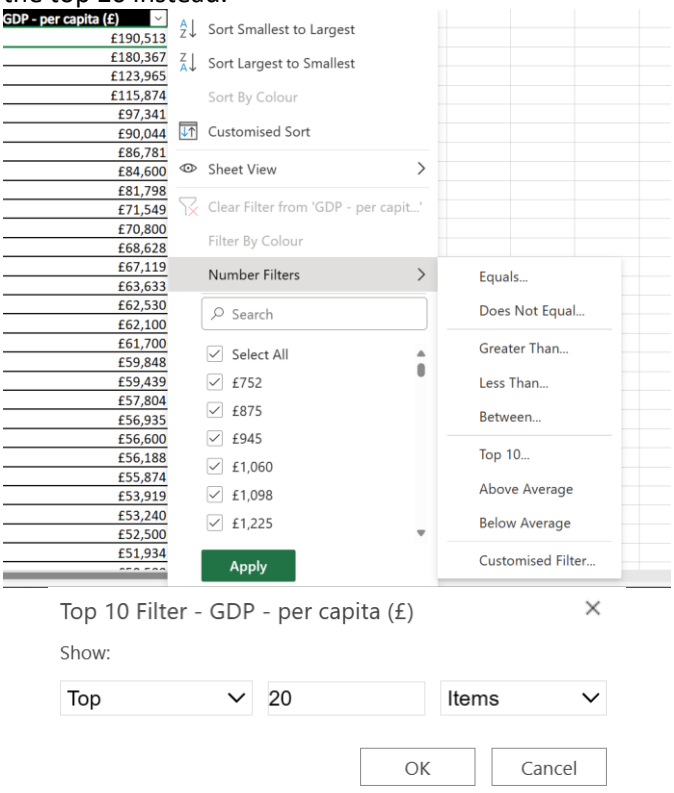


Country	Sum of GDP - per capita (£)
Monaco	190513
Bermuda	81798
Brunei	62100
Cayman Islands	71549
Denmark	57804
Falkland Islands	70800
Gibraltar	61700
Hong Kong	59848
Ireland	86781
Isle of Man	84600
Liechtenstein	180367
Luxembourg	115874
Macau	123965
Norway	63633
Qatar	90044
San Marino	59439
Singapore	97341
Switzerland	68628
United Arab Emirates	67119
United States	62530
Grand Total	**1756433**

To filter the top 20 countries on a pivot table:

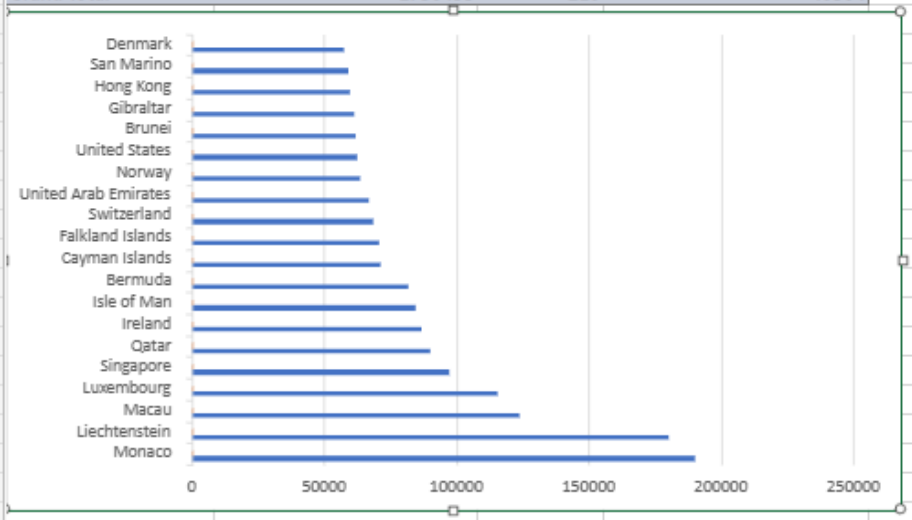
1) Go to your Pivot table field and drag GDP over to Filter
2) An arrow will come up above your table. Click on this arrow
3) A list of the GDPs will come up. Here you can select the top 20. When you are happy click 'OK'
4) Your table is now left with only the data for the countries which are in the top 20 for GDP

NOTE:
On the main GDP sheet you can find the top 20 in an easier way where you can click on the filter arrow for the GDP column. Using the number filter select 'Top 10' and you can edit in the menu that pops up to look at the top 20 instead.



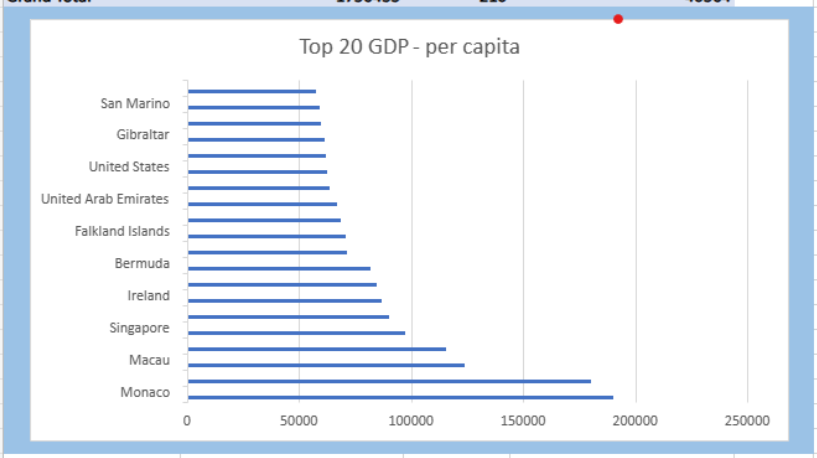
Next create a new bar chart to display the 20 highest ranking countries from your sort and then move the chart to be underneath the table

GDP - per capita (£)	(Multiple Items)		
Country	**Sum of GDP - per capita (£)**	**Sum of Rank**	**Sum of Year of Information**
Monaco	190513	1	2019
Liechtenstein	180367	2	2018
Macau	123965	3	2019
Luxembourg	115874	4	2020
Singapore	97341	5	2019
Qatar	90044	6	2019
Ireland	86781	7	2019
Isle of Man	84600	8	2014
Bermuda	81798	9	2019
Cayman Islands	71549	10	2018
Falkland Islands	70800	11	2015
Switzerland	68628	12	2019
United Arab Emirates	67119	13	2019
Norway	63633	14	2019
United States	62530	15	2019
Brunei	62100	16	2019
Gibraltar	61700	17	2014
Hong Kong	59848	18	2019
San Marino	59439	19	2018
Denmark	57804	20	2019
Grand Total	**1756433**	**210**	**40364**



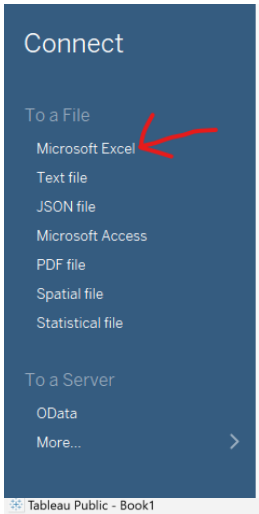
Colour the background by highlighting the area underneath the table as shown below. Find the add a fill colour icon and select a colour.

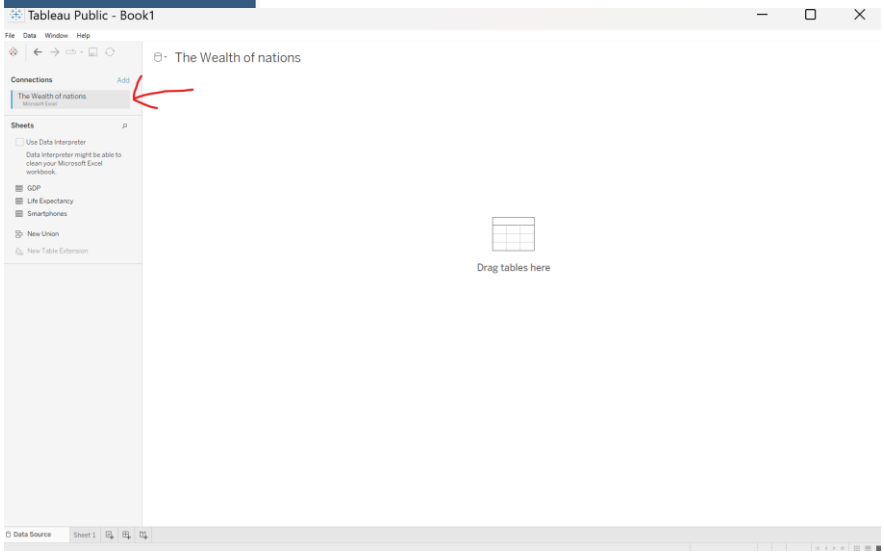
GDP - per capita (£)	(Multiple Items)		
Country	**Sum of GDP - per capita (£)**	**Sum of Rank**	**Sum of Year of Information**
Monaco	190513	1	2019
Liechtenstein	180367	2	2018
Macau	123965	3	2019
Luxembourg	115874	4	2020
Singapore	97341	5	2019
Qatar	90044	6	2019
Ireland	86781	7	2019
Isle of Man	84600	8	2014
Bermuda	81798	9	2019
Cayman Islands	71549	10	2018
Falkland Islands	70800	11	2015
Switzerland	68628	12	2019
United Arab Emirates	67119	13	2019
Norway	63633	14	2019
United States	62530	15	2019
Brunei	62100	16	2019
Gibraltar	61700	17	2014
Hong Kong	59848	18	2019
San Marino	59439	19	2018
Denmark	57804	20	2019
Grand Total	**1756433**	**210**	**40364**



GDP Task – Tableau

Import Data

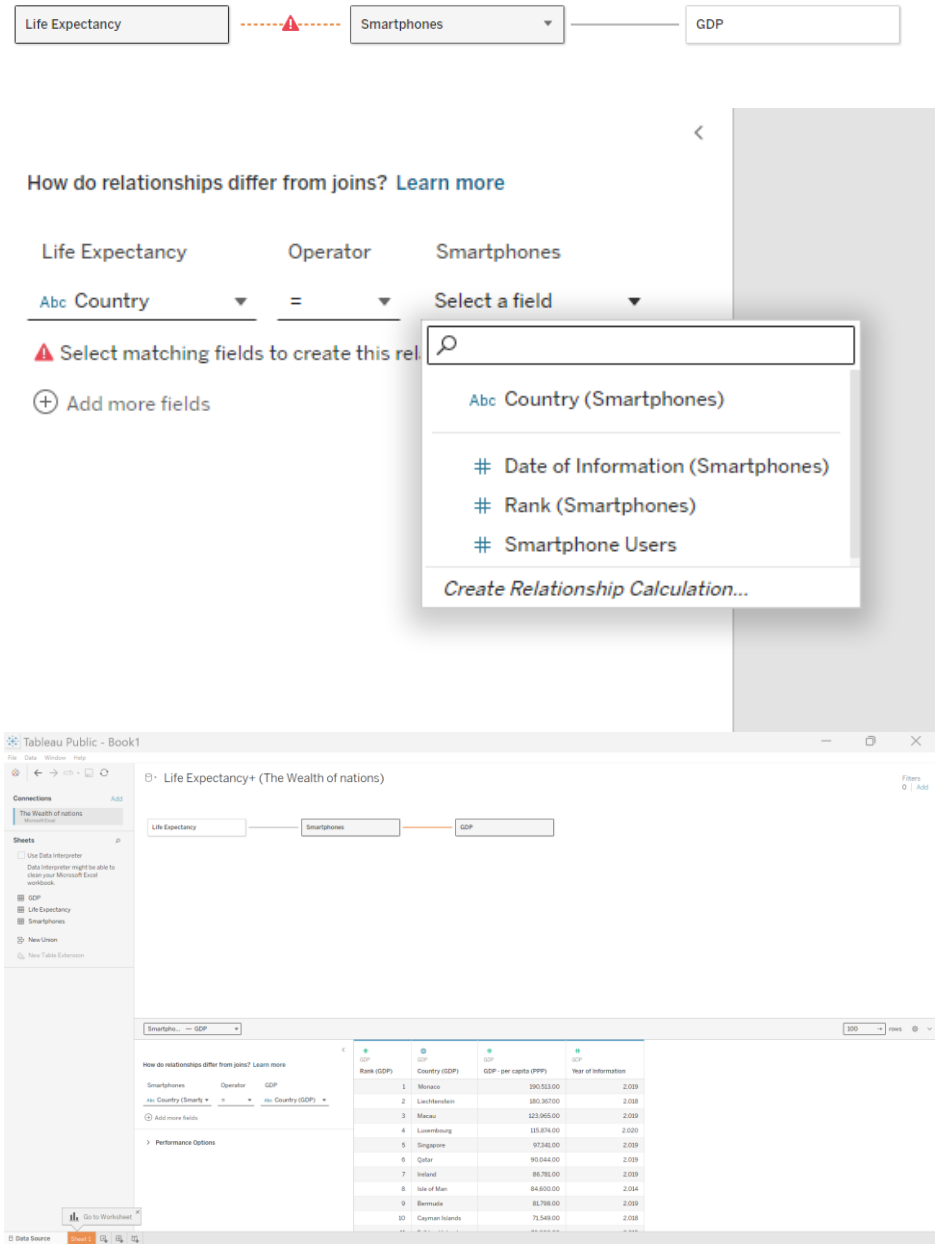




When you open Tableau there is a menu on the left side which gives you the option to import data into Tableau. An Excel file is being used in this case to click on Microsoft Excel and select the file from the computer to import it.

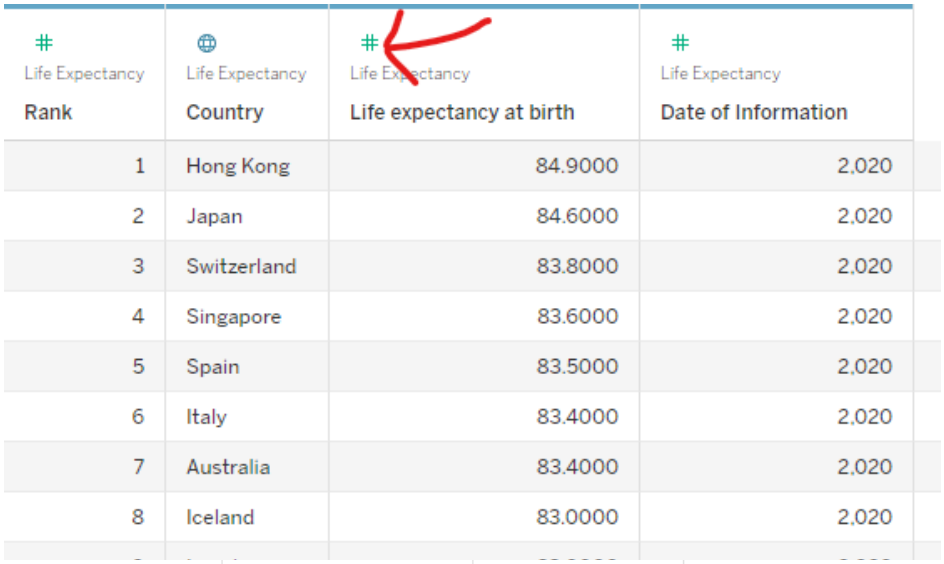
When you have imported the file it should show on the left of your screen under 'connections' and below that a list of sheets which the data can be accessed from.

Set Relationships



You can select the data sheets that you wish to access by draging them into main part of the screen. Once all 3 which are being used have been dragged over it will come up with a warning symbol in-between. Click on the warning symbol and you will be able to select the matching fields. In this case for the 3 datasets 'Country' is a column that all of them have in common. Once that has been selected for each error then the sheets will be able to work together in the visualisations.

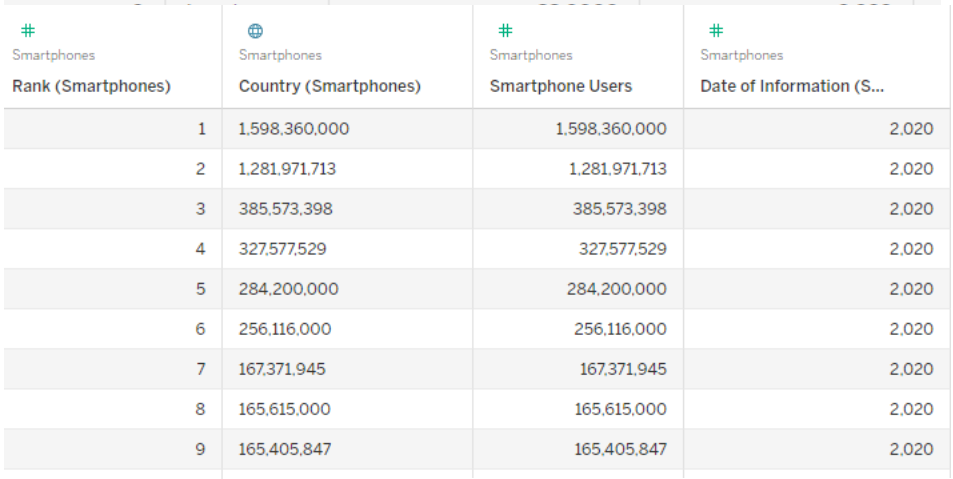
Check Data Types



Life Expectancy Rank	Life Expectancy Country	Life Expectancy Life expectancy at birth	Life Expectancy Date of Information
1	Hong Kong	84.9000	2,020
2	Japan	84.6000	2,020
3	Switzerland	83.8000	2,020
4	Singapore	83.6000	2,020
5	Spain	83.5000	2,020
6	Italy	83.4000	2,020
7	Australia	83.4000	2,020
8	Iceland	83.0000	2,020

The data types can be seen at towards the bottom right of the screen. There will be a table showing the data which Tableau has gotten from the Excel file.

The symbols above each column represent what type of data Tableau sees. Here you can check that it has been recognized as the correct datatype and change it if not by clicking on the symbol and choosing the correct one.



Smartphones Rank (Smartphones)	Smartphones Country (Smartphones)	Smartphones Smartphone Users	Smartphones Date of Information (S...
1	1,598,360,000	1,598,360,000	2,020
2	1,281,971,713	1,281,971,713	2,020
3	385,573,398	385,573,398	2,020
4	327,577,529	327,577,529	2,020
5	284,200,000	284,200,000	2,020
6	256,116,000	256,116,000	2,020
7	167,371,945	167,371,945	2,020
8	165,615,000	165,615,000	2,020
9	165,405,847	165,405,847	2,020

Here is a table showing the different symbols and what datatype each of them represents.

Icon	Data type
Abc	Text (string) values
	Date values
	Date & Time values
#	Numerical values
T|F	Boolean values (relational only)
	Geographic values (used with maps)
	Image role (used with image link URLs)
	Cluster Group (used with Find Clusters in Data)

Source: (help.tableau.com, n.d.)

Client Requirements – Tableau

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.

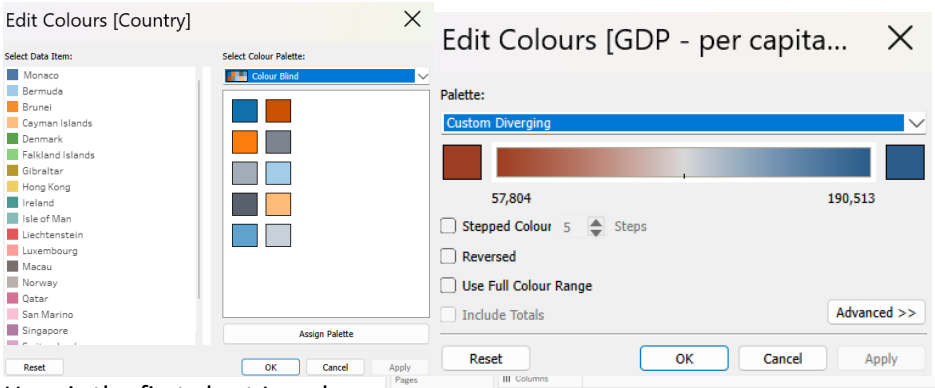
Creating charts

When creating the visualisations according to the client requirements there are a few things to consider. The client mentioned only being interested in the top 20 highest ranking countries. For each of the categories on the spreadsheet (GDP, Life Expectancy and Smartphones) there is a different top 20. Also some of the top 20 in GDP

have no data for the other category, so it doesn't give anything to compare against. In a situation with a client, it would be useful to ask for clarification as to which category of top 20 they wish to see (or is this for all categories) and to fill in any missing data where possible.

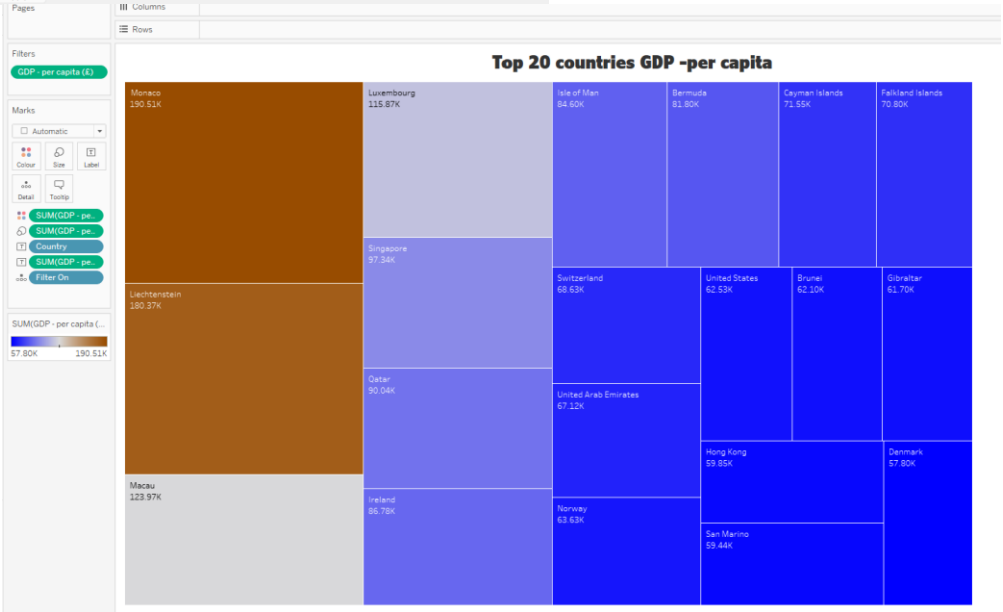
When it comes to catering a visualisation to be accessible for people who are colourblind, there are a few things to consider here as well, especially if being able to tell the difference between the colours is important to fully understanding the chart. One way to cater for this is to avoid making charts which rely on being able to differentiate colour to understand them. Using other tools such as size and labels for example can be accessible and understood the same between those with colour-blindness and those without.

When making use of colour may be useful, there are a few ways of doing this which may be useful to a person who is colourblind. There are several types of colour-blindness so if this is for a specific person it is useful to ask which type of colour-blindness affects them. Otherwise, there are some general rules that you can follow. Red/green colour-blindness is the most common, so avoiding using these colours together can be helpful. Red, green and orange to some people can appear to be brown, so they aren't helpful to differentiate categories. There are some colourblind friendly pallets such as, blue/orange, blue/red and blue/brown which can be seen more easily for those with partial colour-blindness. Otherwise making use of different shades can be a good way to differentiate different data clearly to everyone (Tableau, n.d.).

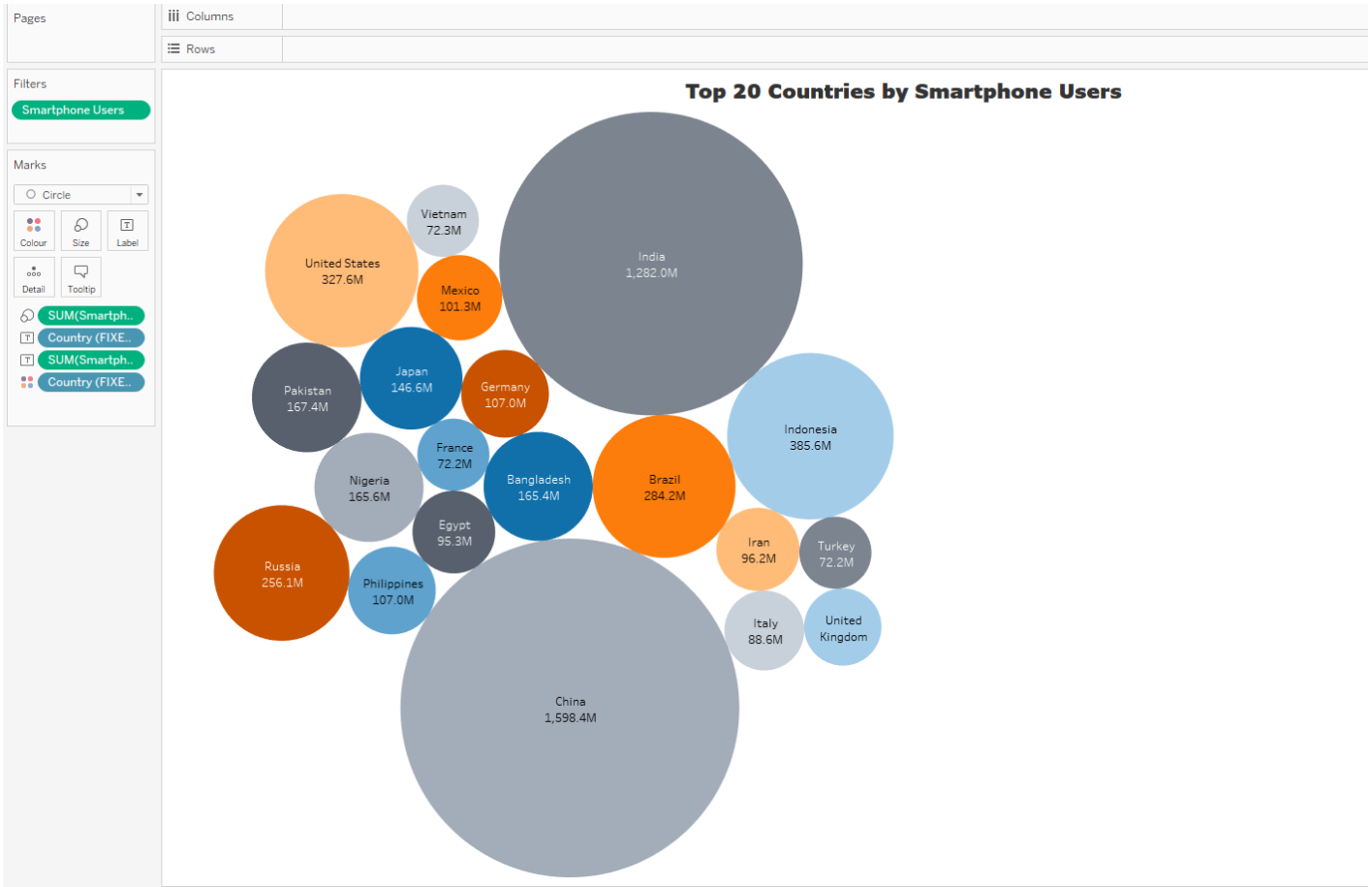


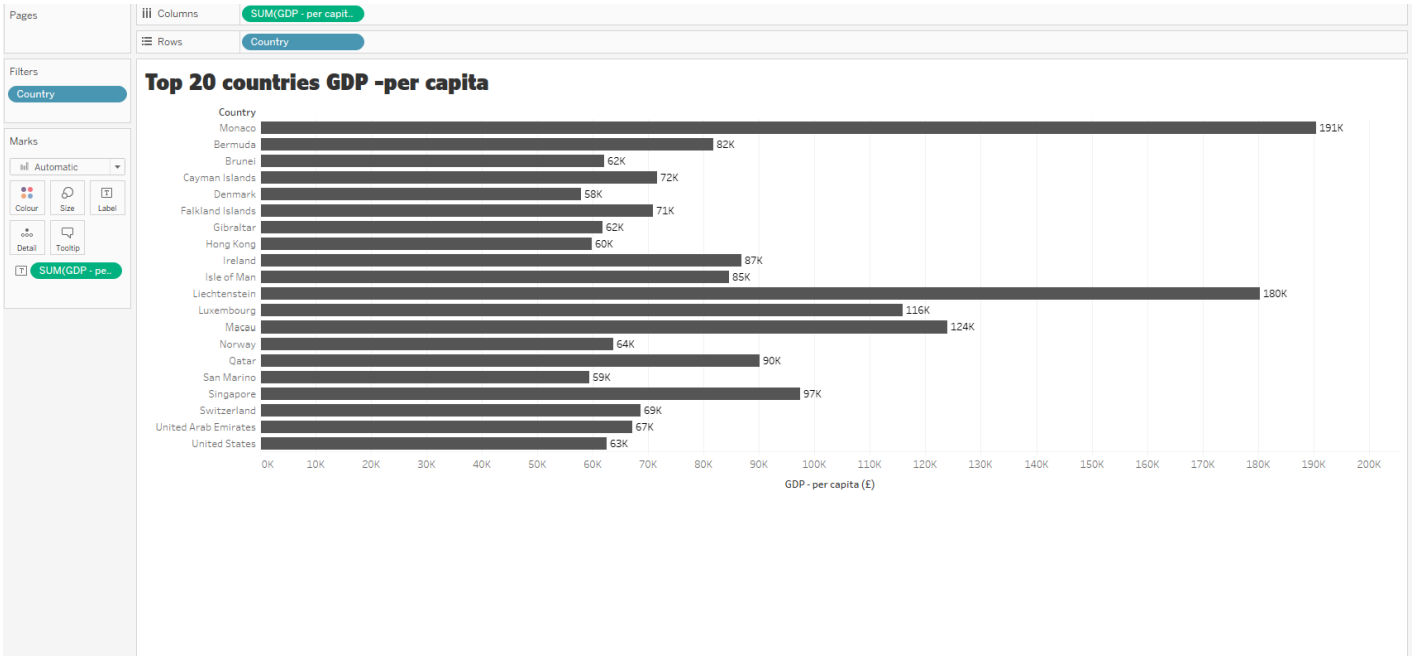
Here are some ways that you can edit colours on Tableau to be more accessible for colour-blindness. On the left I have the option to choose a colour pallet in Tableau which is more accessible. I also have the option to make custom pallets (on the right) where I can choose my own colours.

Here is the first chart I made. I have shown the top 20 countries by GDP – per capita using a tree map and a custom blue/brown colour pallet. Here you can differentiate the countries using the size of the box, the colours and I've added labels stating the GDP as well for extra clarity.



Following I have shown the rest of my Tableau visualisations. Each of them only includes the top 20 countries in their respective categories and making them accessible for colourblind people by making use of size and colourblind friendly colour pallets.

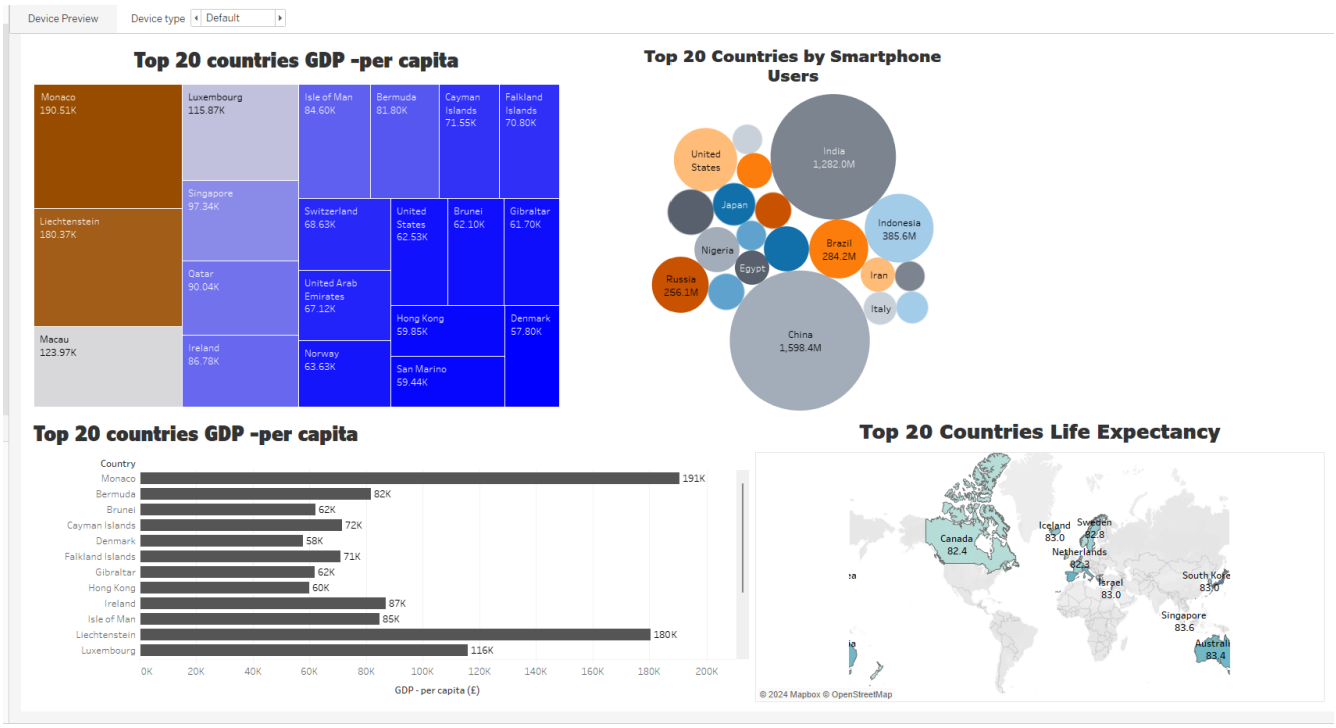




Building the Dashboard

Here is my final dashboard with all of those visualisations put together. You can have a closer look at my dashboard on my Tableau public website below:

https://public.tableau.com/app/profile/emma.kaas.andersen/viz/WealthofNationsTop20/WealthofNationsTop20



Bibliography

Agarwal, A. (2022). *Difference between Standards, Frameworks, Compliance, Laws, Regulations, and Acts*. [online] LinkedIn. Available at: https://www.linkedin.com/pulse/difference-between-standards-frameworks-compliance-laws-agarwal/ [Accessed 4 Apr. 2024].

help.tableau.com. (n.d.). *Data Types*. [online] Available at: https://help.tableau.com/current/pro/desktop/en-us/datafields_typesandroles_datatypes.htm.

ICO (2023). *Policies and procedures*. [online] ico.org.uk. Available at: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/accountability-and-governance/accountability-framework/policies-and-procedures/.

Irwin, L. (2023). *What Is the CIA Triad and Why Is It Important?* [online] IT Governance UK Blog. Available at: https://www.itgovernance.co.uk/blog/what-is-the-cia-triad-and-why-is-it-important.

Tableau. (n.d.). *5 tips on designing colour-blind-friendly visualizations*. [online] Available at: https://www.tableau.com/en-gb/blog/examining-data-viz-rules-dont-use-red-green-together.