THE WEALTH OF NATIONS

A Data-Driven Analysis

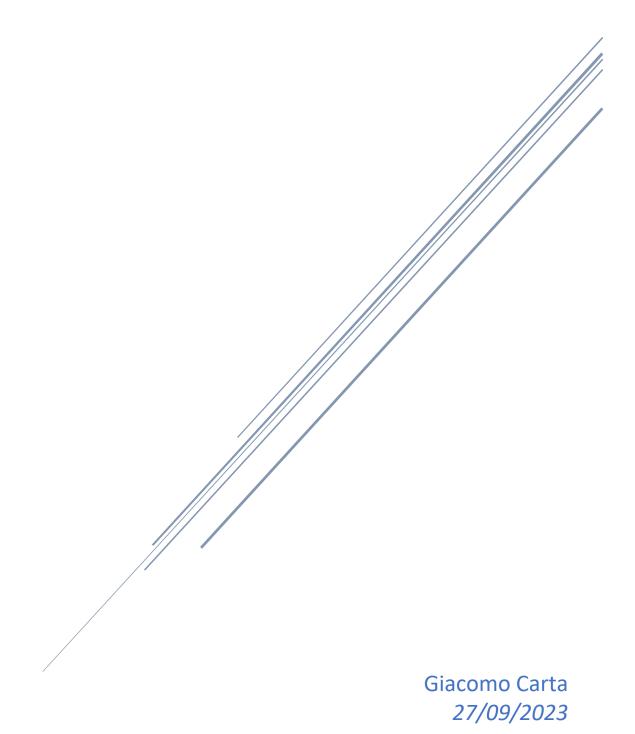


Table of Contents

Policies and Procedures	2
Data Access and Use	2
Data Quality and Governance	2
Data Ethics and Compliance	2
The Wealth of Nations	3
Excel: GDP	3
Setting a password to protect the workbook	3
Highlight column C and change the data to display in British Pound symbol	4
Turn the GDP sheet into a table	5
Filter the table to display only the information for 2019	5
Create and edit a chart containing "Rank", "Country" and "GDP"	6
Sort data by top 20 highest ranking territories	7
Create a bar chart and move it under the table	7
Create the macros "Print", "Save" and "Copy"	8
Adding header and footer	9
Tableau: Data Visualisation	10
First Chart: GDP by Country	12
Second Chart: Top 25%, Above Average, Below Average, Bottom 25%	13
Third Chart: Africa hotspot map	15
Fourth Chart: Bar chart for Africa	16
Fifth Chart: GDP per capita and Life expectancy scatter chart	17
Sixth Chart: Africa's Map Chart by Life expectancy at birth	18
Seventh Chart: Bar chart for Life Expectancy	19
Completing the dashboard: KPIs	20
Building the Dashboard	21
Reflections on "The Wealth of Nation" report	22

Policies and Procedures

Data analysts play a crucial role in the success of organisations at all levels. They can give insights and help companies make informed decisions that drive growth.

As they work with data, including sensitive information about organizations and individuals, they must follow clear policies and procedures to protect the privacy of the data they handle.

Data Access and Use

First, data analysts must only access data that they have been authorised to access. Unauthorised access could lead to a breach of the security system and loss of sensitive data, which could then be used for fraudulent purposes. For the same reason, data must be used only for the purposes for which it was requested and approved.

Data analyst must guarantee that the data will be protected in conformity with the organisation's security procedures.

Data Quality and Governance

Data analysts must regularly clean and validate data to ensure its accuracy and consistency. They must only retain data for the duration necessary to complete the task and the period required by the organisation's policies.

Data Ethics and Compliance

Data analysts must respect the privacy of individuals and organisations when collecting, using and disclosing data. They must comply with all applicable laws and regulations related to data privacy and security. As a result, they must avoid conflicts of interest when working with data.

The Wealth of Nations

The "Wealth of Nations" data set is a collection of information about a comprehensive list of countries. As the name suggests, this information is about the wealth of these economies and seeks to suggest a correlation between economic strength, technological development and life expectancy. It is composed of three sheets: "GDP", "Life Expectancy" and "Smartphones".

Excel: GDP

This is the first of the three sheets that make this dataset. It is a list of 228 territories that have a significant economic activity, even if they are not sovereign states. An example is Hong Kong, a special administrative region of China; although it is not considered to be a country, it has a substantial GDP, even superior to some sovereign states. For this reason, the column "Country" was changed into "Territory", to include both sovereign states and non-sovereign territories.

Rank	Territory	GDP - per capita (PPP)
18	Hong Kong	\$59,848.00
19	San Marino	\$59,439.00
20	Denmark	\$57,804.00
21	Netherlands	\$56,935.00

Hong Kong ranks 18th for GDP, it is higher than 91.67% of territories, which include countries such as San Marino, Denmark and The Netherlands.

Setting a password to protect the workbook

Firstly, a good procedure for data analysts is to protect the workbook. It is a good practice to prevent unauthorised or accidental changes. In certain cases, protecting and hiding worksheets from a workbook is necessary to prevent unauthorised access to sensitive information while still being able to show the end results to the public.

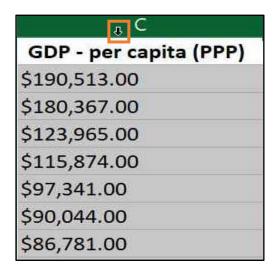
To protect a workbook in Excel:

- Open the workbook
- Click the Review tab
- In the Protect group, click Protect Workbook
- In the dialog box, type in a strong password
- Click OK



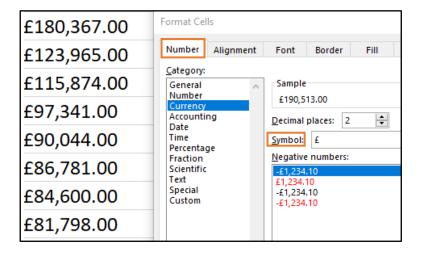
Highlight column C and change the data to display in British Pound symbol

Column C shows GDP for all territories; the amount is currently in USD. To change it to British Pound, we need to select the column first. To do so, hover on the column until the **black arrow** appears, then click on the column:



This will select the whole column C in the sheet, even cells without data. To only select the data within the column, it is possible to simply **drag and select** using the left mouse button; alternatively, a very convenient option is to click on the first cell of the list we want to select, then hold down **Ctrl + Shift**, then press the **Down Arrow** key. This will select only the data under the cell we selected.

Now, to change the currency, just right click in the range and select the "Format Cells..." option. In the window that pops up, select under the **Number** tab the **Currency** option, then change the **Symbol** and select the number of decimal places we want to display:



The **Number** group on the **Home** tab is a very quick and easy option too!



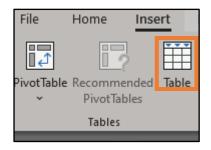
Turn the GDP sheet into a table

Turning the GDP sheet into a table it is a good practice that has many benefits, such as facilitating the insertion of data with automatic table expansion to accommodate the new information. It helps to create structured references, which make it easier to write long, complex formulas. Tables also allow to apply data validation to restrict the types of data that can be entered into the table among a lot of other advantages.

Creating a table is very easy in Excel. First, we need to select the range of cells that we want to make into a table. In this case, it is a matter of clicking on a cell within the GDP cell that has data.

Afterwards, press **Ctrl + A** to select all the cells that have data and are connected to each other.

Finally, press **Ctrl + T**, tick the option if the table has headers and click **OK**. Alternatively, after selecting the range, on the **Insert** tab, in the **Tables** group, select **Table**.



A new tab called **Table Design** will appear, adding many options to select styles to make the table unique.

Filter the table to display only the information for 2019

Filtering the data lets us identify the information that is relevant to our search. In this case, we will display only the information for 2019, by clicking the arrow next to the "Year of Information" column header and select 2019. This will leave us with a list of 188 territories.



On the **Home** tab, in the **Editing** group, it is possible to create and clear all filters by clicking on **Sort & Filter** and selecting one of the options.

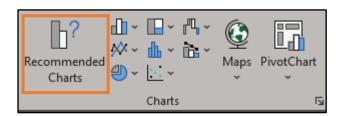
□ 2006	^
□ 2008	
2014	
2015	
2016	п
2017	п
2018	п
✓ 2019	п
2020	V

Create and edit a chart containing "Rank", "Country" and "GDP"

Creating a chart lets us visualise the data, making it more intuitive and easier to read. This can also reveal trends that we might have missed before.

To create a chart in Excel:

- Select the table with Ctrl + A
- Go to the Insert tab
- On the **Charts** group, select the chart



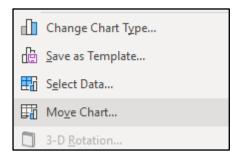
When we click on the chart, a new tab **Chart Design** appears. It will allow us to edit it for our needs.

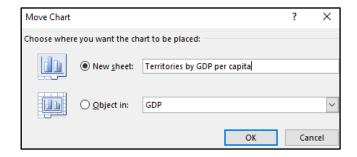
Using the **Chart Design** tab, we will edit the chart to create a Title, label the X and Y axis and make the chart visually pleasing.

To do so:

- In the Chart Layouts group, select Add Chart Element
- On Axis Titles, select both Primary Horizontal and Primary Vertical
- On Chart Title, select Above Chart
- The Axis Titles and Chart Title should now appear on the chart
- By clicking on them, we can rename them
- Finally, on the **Chart Styles** group, we can change colour and design

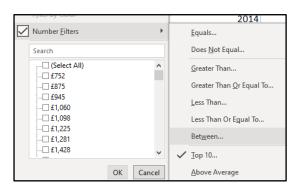
We will now move the chart to a new sheet tab and label with a suitable name, simply by right clicking on the chart and selecting the **Move Chart** option. A window will appear, letting us choose where to move the chart and the name of the new sheet.

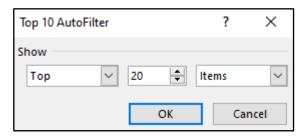




Sort data by top 20 highest ranking territories

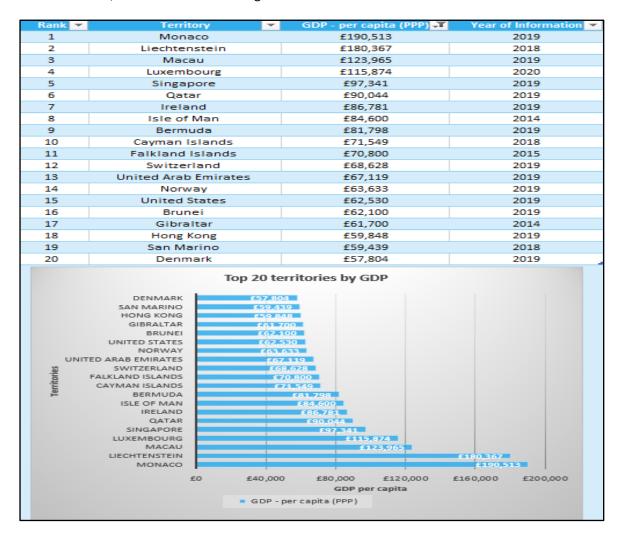
Now, we are going to sort the data by the top 20 highest ranking territories. First, let's remove all the filters by clicking on **Sort & Filter**, then on **Clear**. Second, we will click on the dropdown menu in the **GDP – per capita (PPP)** table header. Here, we will select the **Number Filters** menu and finally we will click on **Top 10**, where we will change the number to show us the top 20.





Create a bar chart and move it under the table

We will use the method used before to create a bar chart, then we will simply drag and drop it under the table. Then, we will colour the background area.



Create the macros "Print", "Save" and "Copy"

Creating macros in Excel has many advantages:



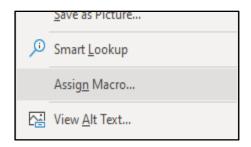
- Improved productivity
- Automation of simple, repetitive tasks and more complex tasks
- Increased consistency
- Filtering data
- Generating reports
- Perform calculations
- Create custom functions

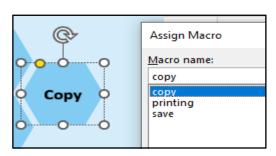
In our case, we will create the macros to automate the functions of print, save and copy.

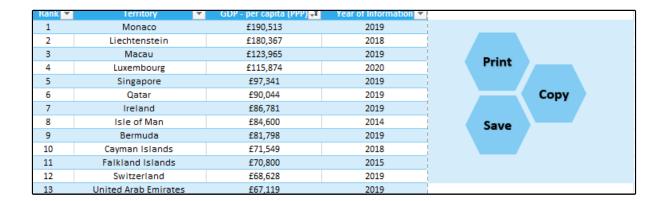
To do so:

- Go to File > Options > Customise Ribbon
- Select **Developer** in Main Tabs and click OK
- Go to the Developer tab
- Click Record Macro
- Perform the task to record (e.g. print)
- Click on Stop Recording

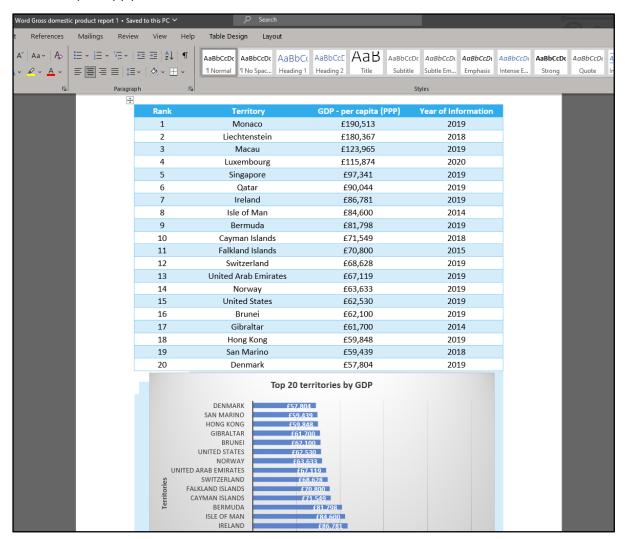
Now that we have created three macros, we will link them to three buttons beside the table and chart. To do so, will right click on the button and select **Assign Macro**. Afterwards, by clicking on our new macro "Copy", we will select the sheet and copy it into a new word document.





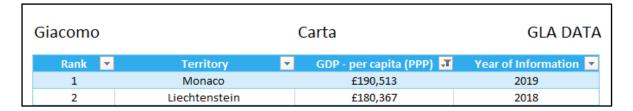


With a simple copy-paste, we created a new word document with the table and chart.



Adding header and footer

To complete the work on our Excel workbook, we will create a header and footer. In the **View** tab, select **Page Layout** under the **Workbook Views** group. This will display the screen with a header and a footer, where we will write our name and date.



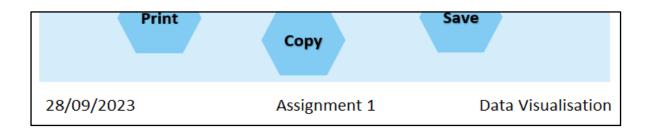
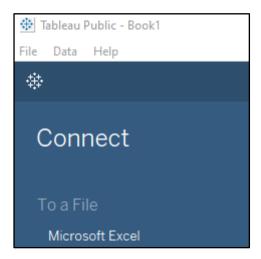


Tableau: Data Visualisation

Tableau is a very useful tool for data analysis, as it allows us to create interactive dashboards and reports in an intuitive way. It is a powerful data visualization tool with a wide range of features. It offers a better way to handle large datasets compared to Excel and a great variety of visualization features.

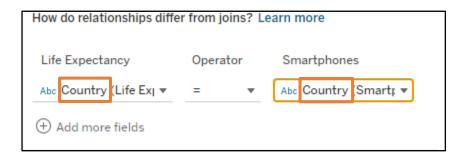
To follow the workflow, first we must import the data into Tableau:



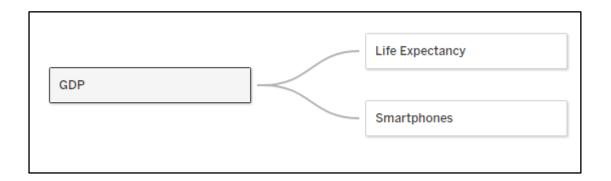
We will select **Microsoft Excel** from the list of possible options, which includes Text file, JSON, Access, PDF, Spatial and Statistical file.



The imported file will show up like this in Tableau. Now we need to create relationships between the three sheets.



We will use the column "Country" as a common denominator to link all three sheets together.



Now, we will make sure that we are working with the correct data types, by looking at the preview.



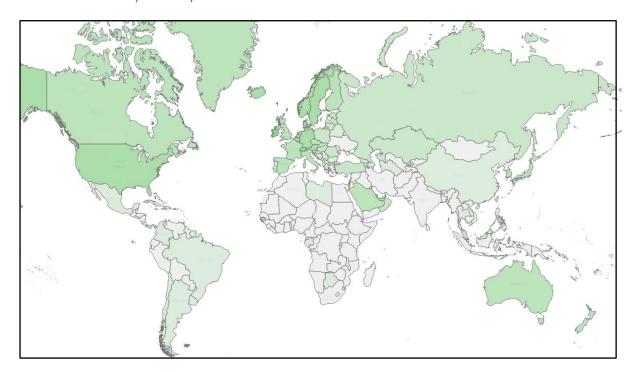
Clicking one of the symbols will allow us to modify them. For example, we can change the format of the "Year of Information" section, from numerical to **Date**.



As we can see, the formatting of the date is such that it shows also day and month and, not having this information, it fills it with 01/01. In this case, it still works as valuable quantitative data, but if needed we can change it to text.

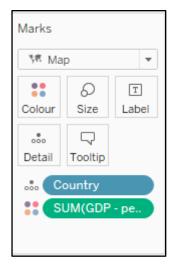
We will now check and clean the data files from the other sheets, then we will be able to start developing charts and visuals.

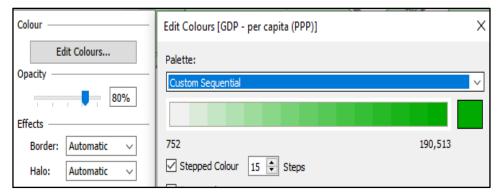
First Chart: GDP by Country



We built this chart by double clicking on **Country** under GDP. This automatically opens the map chart, as it is a list of geographic locations. We then selected the **GDP** – **per capita (PPP)** value and we dragged it to the **Colour** button under the **Marks** group.

Then, by selecting the **Edit Colours** option in under Colour, we changed the palette to **Custom Sequential**, we chose the Green colour with **Stepped Colour** in 15 steps. This created a gradient that shows, the darker the green, the higher the GDP per capita, all in a very intuitive way. Finally, by hovering with the mouse indicator over the single country, the numerical value and the name of the country will be displayed for insight.

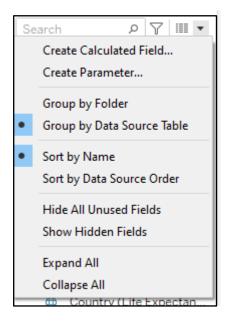




In its visual simplicity, this chart hides a very complex series of steps to create.

The main idea was to highlight countries that, in terms of GDP, are below average, above average, top or bottom 25%.

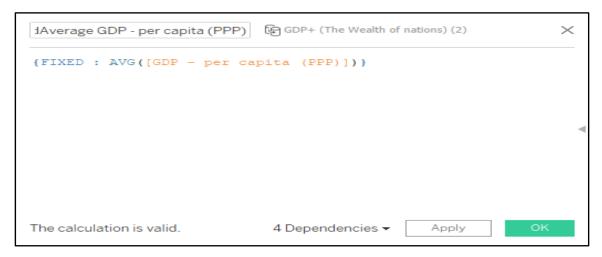
To do so, it was necessary to create three different **calculated fields**, a very powerful feature that enables Tableau users to automate calculations by breaking them in smaller steps.



Top, Bottom, Avg	Count	ries by GD	P
Country =		,	
Danamas	37,101	^	
U.S. Virgin Islands	37,000		
Estonia	36,927		
Guam	35,600		
Sint Maarten	35,342		
Portugal	34,894		
Puerto Rico	34,518		
British Virgin Islands	34,200		
Poland	33,221		
Hungary	32,945		
Slovakia	32,730		
Panama	31,459		
New Caledonia	31,100		
Latvia	30,898		
Romania	29,941		
Greece	29,799		
Turks and Caicos Islands	29,253		
Seychelles	29,223		
Croatia	28,602		
Turkey	28,424		
Malaysia	28,364		
Oman	27,299		
Russia	27,044		
Saint Kitts and Nevis	26,438		
Kazakhstan	26,351		
Trinidad and Tobago	26,176		
Northern Mariana Islands	24,500		
Curação	24,479		
Chile	24,226		
Bulgaria	23,174		

First, we created a calculated field that calculates the average of the GDP per capita for all the countries in the list (we had to exclude the values for "World" and "EU"). To do so, we right clicked on the **Tables** group at the left of the screen and select **Create Calculated Field** as per the picture on the top left.

On the window, we typed the formula, making sure to make a fixed calculation, so that even if we want to filter the data to see a certain country, the average doesn't change based on that country, but it stays like an absolute reference point. With this formula, we can visualise the average total GDP for all the countries. On this we will base our "Above Average", "Below Average", "Top 25%" and "Bottom 25%" parameters.



Next, we need to create a calculated field for the difference from the average. This is to show how far above or below the average each country's GDP per capita is. By subtracting the average GDP per capita from a country's GDP per capita, we will have a positive value for above average, a negative for below average.

```
Difference from Avg GDP+ (The Wealth of nations) (2) X

[GDP - per capita (PPP)] - [FixedAverage GDP - per capita (PPP)]
```

Finally, for the last step, we must create a calculated field that categorises each country based on the value given by the "Difference from Avg" calculated field above. To do so we need to use conditional functions like **IF**:

```
AverageGDP

GDP+ (The Wealth of nations) (2)

IF [Difference from Avg] > 0.25 * [FixedAverage GDP - per capita (PPP)] THEN "Top 25%"

ELSEIF [Difference from Avg] > 0 THEN "Above Average"

ELSEIF [Difference from Avg] < -0.25 * [FixedAverage GDP - per capita (PPP)] THEN "Bottom 25%"

ELSE "Below Average"

END
```

What we are telling Tableau is this: if we subtract the average GDP per capita from the country's GDP per capita, we will have a positive (zero would be considered positive in this case) or a negative value. If the value is greater than 25% than the fixed average GDP per capita, then it should be called "Top 25%". The same principle applies to the other three values.

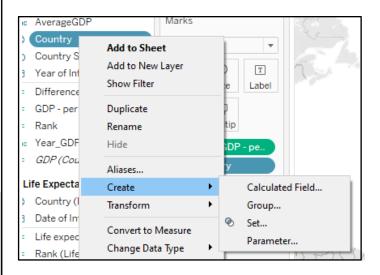
Although this seems a slow and counterintuitive process, Tableau can struggle with long and complex functions, therefore by simplifying and breaking down into steps, it becomes more digestible for the software, and even for us.

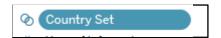
Third Chart: Africa hotspot map

The first world map clearly shows that Africa has a very high percentage of territories below the average GDP per capita. This is confirmed by the second chart that indicates in red a big group of African countries.

Rwanda	2,227
Gambia	2,223
Ethiopia	2,221
Uganda	2,187
Burkina Faso	2,178
Afghanistan	2,065
Guinea-Bissau	1,989
Sierra Leone	1,718
North Korea	1,700
Madagascar	1,647
South Sudan	1,600
Eritrea	1,600
Togo	1,597
Chad	1,580
Liberia	1,428
Mozambique	1,281
Niger	1,225
DR Congo	1,098
Malawi	1,060
Central African Republic	945
Somalia	875
Burundi	752 v

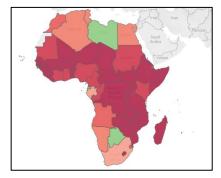
These data make it necessary to further investigate the issue of Africa. For this reason, we decided to create a map that shows the African situation in detail. To do this, we need to create a "set".





To be able to create a map that shows only Africa there are multiple ways. We decided to use the "set" option. To create a set, right clicking on "**Country**" will show the option **Create** where we can select **Set**. This enables us to select all the countries that are in Africa by ticking the boxes in the world country's list.

By doing this, we create a separate field called "Country Set" (as per above) that we can use to make charts where only information for Africa is displayed.

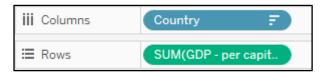


As expected, when highlighting countries with the lowest GDP per capita in red (selecting the "red-green diverging" colour scheme in the **Marks** section), we can observe that most African countries have very low GDP per capita values, with exceptions being Seychelles, Mauritius, Libya, Botswana and Equatorial Guinea. This prompts us to take a closer look at these specific countries.

Fourth Chart: Bar chart for Africa

A map chart is a highly intuitive visual representation that allows us to quickly identify countries that do not follow the same trend as others. In this case, it is easy to observe that Libya and Botswana stand out, with Equatorial Guinea appearing as a small green dot, and Seychelles and Mauritius almost imperceptible. This is why it would be very helpful to complement the map chart with a straightforward bar chart to clearly display the names of the African countries that rank highest in GDP per capita.

To do so we simply select the values we need to use, the **Country** value and the **GDP per capita** value, then we drag and drop them respectively on Columns and Rows.



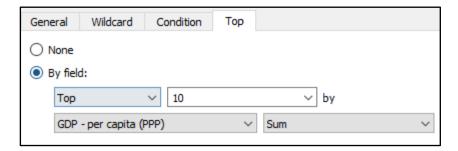
We select the bar chart in the **Show Me** tab.



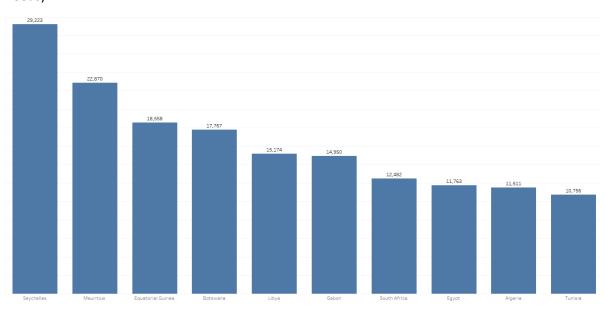
Now, we need to highlight only the African countries, by dragging the value **Country Set** to the Filter section.



This will only show the GDP for the countries in the set. But to have a closer look at the top country, we need to apply another filter. Let's drag and drop the value **Country** in the Filters group, then, after ticking all countries, go to the **Top** tab and choose Top 10 by GDP per capita.



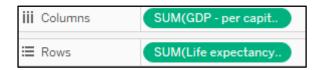
This will be the result (we'll change it to top 20 while building our dashboard to meet the client needs).



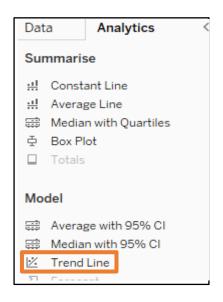
Fifth Chart: GDP per capita and Life expectancy scatter chart

A trend is starting to emerge from the charts. Now, let's explore the potential correlation between GDP per capita and life expectancy, first on global scale, then with a more detailed focus on the situation in Africa.

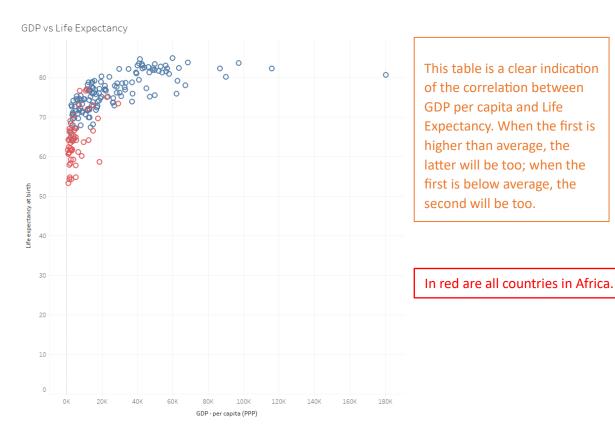
To make a scatter graph, we select two values (or **Measures**), in this case GDP per capita and Life expectancy at birth, and we drag and drop them in Columns and Rows. We then use the Country value as the variable and select the scatter graph from the list. To make it clearer, we can select a **Trend Line** from the **Analytics** tab.





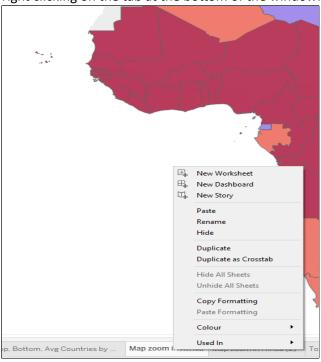


This will be the result:

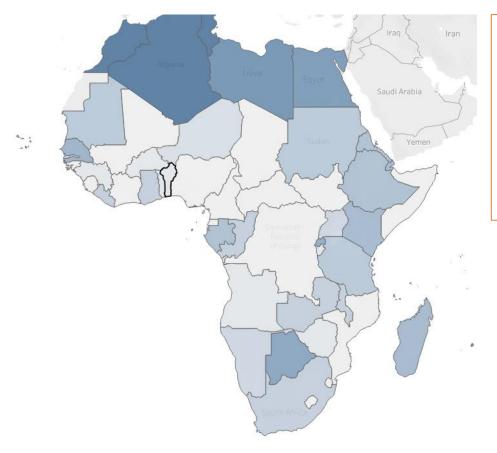


Sixth Chart: Africa's Map Chart by Life expectancy at birth

We can simply duplicate the Worksheet we used before with the map of Africa by GDP per capita by right clicking on the tab at the bottom of the window:



Then, we can simply change the value of GDP per capita (under the Colour section in **Marks**) with "Life expectancy at birth". We will use a different colour palette to show this.



We immediately notice that countries in the central regions of Africa have the lowest Life expectancy at birth, whereas the south of Africa and especially the area of the Mediterranean have the highest.

Seventh Chart: Bar chart for Life Expectancy

As we did before, we will pair the map with a bar chart showing the top 10 countries by Life expectancy at birth. Again, we will duplicate our old bar chart and swap the GDP value by the Life expectancy value. As the value doesn't say much of the world situation, we want to put this data in perspective by adding an average line.

To do so, we need to make a calculated field to create a fixed value for the global life expectancy average.

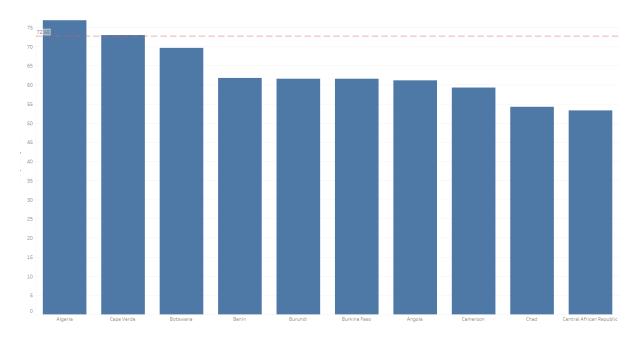
```
{ FIXED : AVG([Life expectancy at birth]) }
```

This will give us an absolute value that will not change based on the chart we are showing. If we don't follow this step, when we try to make an average line for the top 10 countries in Africa, the line will show an average of only the 10 countries selected, and we want to show the difference between the 10 countries in Africa and the world average.

Now, we select the calculated field and we drag it in **Marks**. Finally, we select a **Reference Line** under **Custom** in the **Analytics** tab.



And this will be the result:



It is quite clear that even the top 10 African countries by Life expectancy are mostly below the global average of 72.68. Again, we will change this data to show the top 20 countries to meet the client's needs.

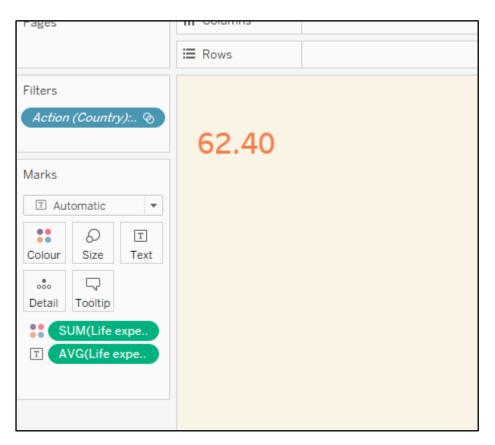
Completing the dashboard: KPIs

A dashboard without KPIs (Key Performance Indicators) will always miss something important, they are essential components as they provide a clear and concise way to showcase data.

We will add three, one for the country's name, one for GDP per capita and one for Life expectancy at birth; afterwards, we will make them interactive.

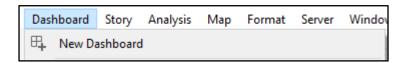
To make a KPI, simply drag a value we want to show and drop it in **Text** under **Marks**. After that, to show which values are above average or below average with colour, we can drag the value again and drop it in **Colour** under **Marks**.

This will be the result. We will do the same for all the values we need for our dashboard.



Building the Dashboard

Now that we have all the charts we need as building blocks, we can create a Dashboard by simply selecting **New Dashboard** from the **Dashboard** tab.



We will find an empty page where we can drag and drop all the charts we need from the dashboard menu.



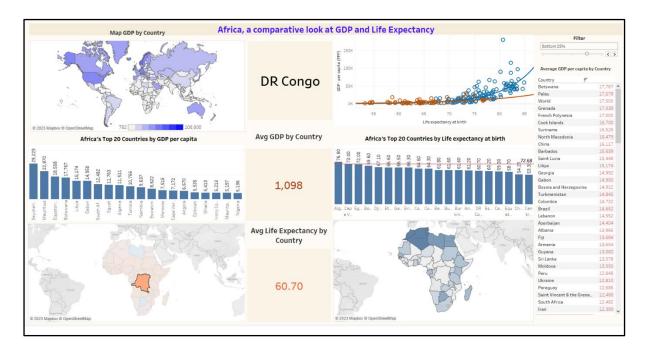
As our client is colourblind, so we must take extra care at the colours we decide to use, and if needed we must go back to the single charts to change them accordingly.

The client also requested that we display only the top 20 countries in some of the charts, so we will make that adjustment.

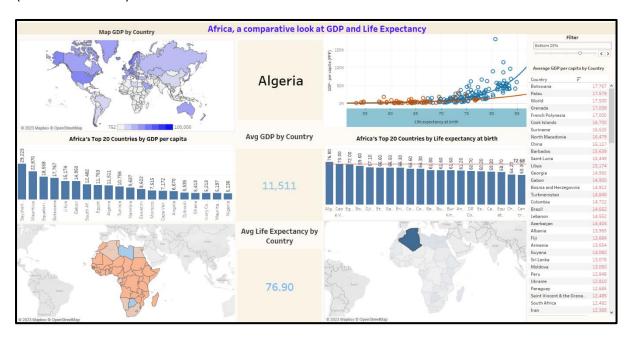
To make the map chart of Africa interact with the KPIs, we will go to **Actions** under the Dashboard tab. There, we will choose our map chart as a source and we will select the KPI charts as target sheets.



And this will be the result:



This is our interactive map, showing data for the Top 20 highest ranking countries in Africa. We built it so that when clicking on a country in either one of the Africa's map charts, we will see information about the name, the average GDP per capita (colour coded) and the average Life expectancy at birth (also colour coded).



Link to Tableau Public project:

https://public.tableau.com/shared/B6Z29KXWB?:display count=n&:origin=viz share link

Reflections on "The Wealth of Nation" report

This project provided an excellent opportunity to apply the skills I acquired during the bootcamp. It allowed me to utilise these techniques to address real-world questions. "The Wealth of Nations," in particular, offered an intriguing perspective on the challenges faced by certain regions of the world, such as poverty and life expectancy.

I chose to focus on Africa, a continent that significantly lags behind global averages in terms of wealth and life expectancy. This project shed light on a reality that is often overlooked and deserves more attention.

The African countries ranking at the top of the list help us understand the underlying issues and make sense of what has been happening in Africa for decades. Libya, Equatorial Guinea, Botswana, Seychelles, and Mauritius all possess elements that have contributed to their relatively high GDP per capita: natural resources like oil (Libya), natural gas (Equatorial Guinea), and diamonds (Botswana); political stability; and thriving tourism industries (Mauritius and Seychelles). These are just a few of the contributing factors. It's important to note that not all African countries have been as fortunate.

The enduring effects of colonialism, political instability, high levels of corruption, inadequate infrastructure, and rapid population growth are among the factors that have negatively impacted the economies of many African nations. As we can see from the charts, economical issues have a big effect in life expectancy and, again, Africa as a whole has an average life expectancy about 11% below the rest of the world, a dreadful number. Hopefully, with time, this gap will narrow, first among African countries and eventually between Africa and the rest of the world.

Regarding the technical aspect of the project, I personally found great satisfaction in exploring various approaches to address the challenges I encountered. It was intriguing to discover that tasks as seemingly simple as displaying different colours for each quarter of a value scale proved to be a complex task when using Tableau, whereas it could have been accomplished in much less time using Excel. Conversely, when it came to visualising and comprehending the data, Excel presented more difficulties compared to the user-friendly and customisable tools provided by Tableau.

In general, I loved that the challenges I faced led me to experiment with different tools and the workarounds eventually proved to be very rewarding. This was especially true with Tableau, as it was a novelty to me. Creating macros in Excel; in Tableau, crafting sets, generating calculated fields from simple to highly intricate, exploring various palettes (particularly those suitable for individuals with colour-blindness), designing interactive dashboards with distinct elements that do not interfere with one another when they are switched or interacted with. It was a great learning experience that motivates me to dig deeper into these incredible softwares.

In the end, the methodical utilisation of both tools (and in the future, the addition of other tools such as PowerBI, SQL, and Python) represents the most effective way to maximise the potential of any dataset.