Tom Ritchie

tomjamesritchie@hotmail.com

Abstract

Policies and Procedure of Data Analysis  
Excel Tasks  
Tableau Tasks

The Wealth of nations

Assignment 1 Data Visualisation

Contents

[Scenario 2](#_Toc175692058)

[Policies and Procedures 2](#_Toc175692059)

[Excel 3](#_Toc175692060)

[Tableau 19](#_Toc175692061)

[References 24](#_Toc175692062)

# Scenario

Data visualisation has become an essential business capability to help transform information into insights that drive meaningful business outcomes and improved experiences. Today, most organizations have accumulated a wealth of data from the different corners of their businesses they are then unable to see how this data can help them make better decisions, making actions, and results.

You have been asked to Look at the data workbook and familiarize yourself with this data. You have also been asked to create a visual report that will show the data in the form of charts and maps using Tableau to the client’s requirements. You will also need to consider data protection and computer misuse policies.

# Policies and Procedures

The Policies and Procedures that Data Analysis professionals must adhere to are:

* Lawfulness, Fairness and Transparency

Lawfulness is self-explanatory and means the controller must have a legal basis to use the data and be compliant with the GDPR and Data Protection Act of 2018.

Fairness means that the use of a person’s data doesn’t negatively affect them or how they’re perceived by others by being misleading or deceptive.

Transparency means that the controller must provide individuals with concise, easily accessible and easy-to-understand information about the existence and purpose of the operation.

* Purpose Limitation

This means the controller should only use individuals’ data for what they have told them it will be used for and not used further in a way that is against the controller's original purpose. Only in certain instances could the data be used in the future i.e. archiving, statistical, scientific or historical research purposes.

* Data Minimisation

This principle requires controllers to only collect the relevant data for their purposes. This results in a ‘Data Protection by Design’ approach as it limits the amount of data that can be lost or stolen in the event of a data breach.

* Accuracy

Very self-explanatory this principle requires that controllers ensure that all individuals’ data is accurate and up to date.

* Storage Limitation

Controllers must only hold onto individuals’ data for no longer than is required for their purposes. It may be stored for longer solely for archiving, statistical, scientific or historical research purposes.

* Integrity and Confidentiality

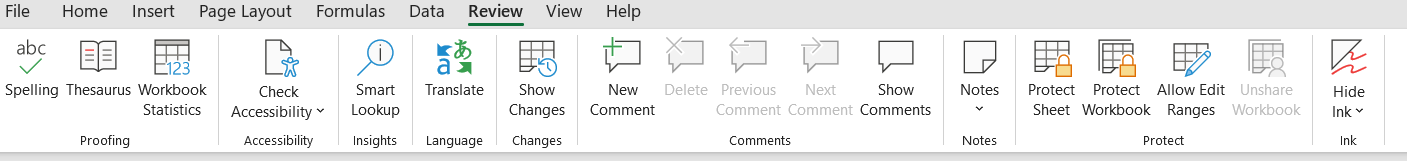
Controllers must ensure their security measures protect against accidental or deliberate harm or loss of personal data. The security measures should cover both cybersecurity and physical and organisational security.

* Accountability

This means that controllers must adhere to all the policies and procedures listed above but also have appropriate records in place to demonstrate this compliance.

# Excel

1. From the ‘Review’ tab I selected ‘Protect Workbook’ in the ‘Protect’ Group.



From the pop up I then added a password to protect the workbook. I then clicked ‘OK’.

A screenshot of a computer

Description automatically generated

I re-entered my password as confirmation. I then clicked ‘OK’.

A screenshot of a computer error

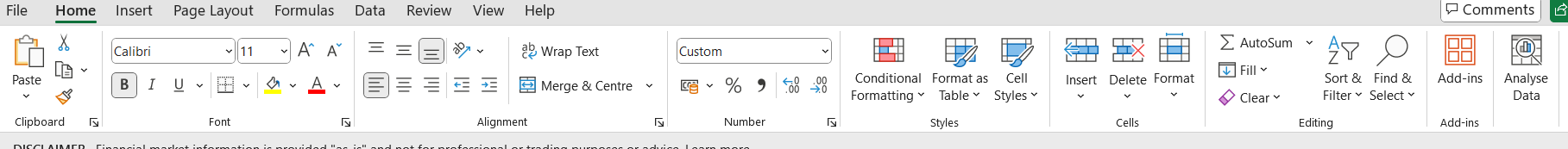
Description automatically generated

1. I highlighted column C (GDP – per capita(PPP)).

A screenshot of a table

Description automatically generated

I navigated to the Home tab, Number group and then clicked the dropdown next to Accounting Number Format.



In the drop-down menu, I selected ‘£ English (United Kingdom)’.

A screenshot of a computer

Description automatically generated

This made the data type of the cells become accounting so I had to select the ‘Accounting’ drop-down in the ‘Number’ group.

A screenshot of a computer

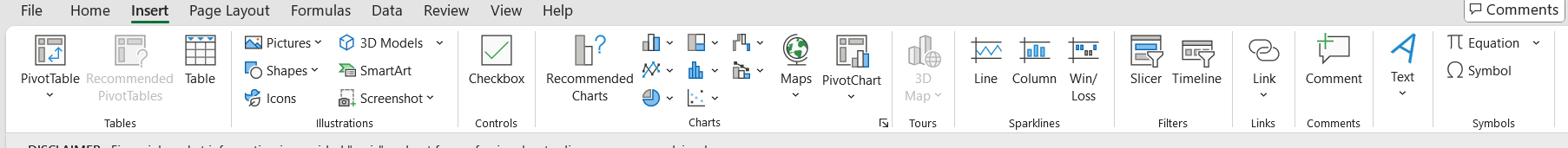
Description automatically generated

From the dropdown, I selected ‘Currency’ to change the data type back to currency.

A screenshot of a computer

Description automatically generated

1. I navigated to the ‘Insert’ tab, ‘Tables’ group and selected ‘Table’ to create a table from my data.



From the pop up I selected all my data by clicking the top left cell of my data, typing a colon (:), and then clicking in the bottom right cell of my data. I then checked the box next to ‘My table has headers’ before clicking ‘OK’.

A screenshot of a computer

Description automatically generated

1. Once my table was created I could go to the drop-down menu next to ‘Year of Information’.

A screenshot of a data

Description automatically generated

From here I unchecked the box next to ‘(Select All)’…

A screenshot of a computer

Description automatically generated

…And checked the box next to ‘2019’. Once I had done this I then clicked ‘OK’.

A screenshot of a computer

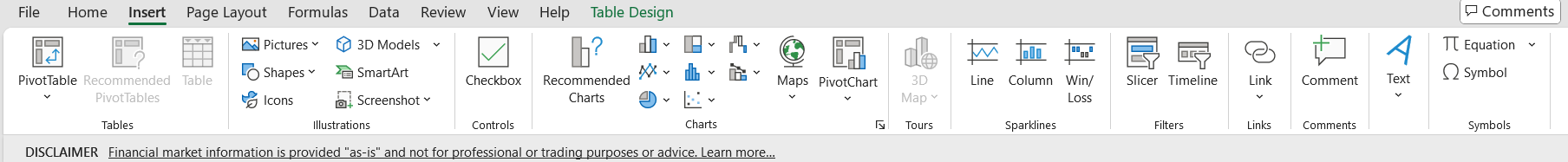
Description automatically generated

1. I selected the required data (‘Rank’, ’Country’, and ‘GDP – per Capita (PPP)’) from my table. I ignored ‘Year of Information’ as my graph didn’t need it.

A screenshot of a data

Description automatically generated

Once the required data was selected I navigated to the ‘Insert’ tab, ‘Charts’ group and in my mind the best way to show this information was in a map chart so I selected to ‘Map’ dropdown.



From the ‘Map’ dropdown I selected ‘Filled Map’.

A screenshot of a computer screen

Description automatically generated

1. I edited the chart by using the drop-down located at the top right of the chart. I then selected a design with clear borders between countries from the ‘Style’ section.

A screenshot of a computer

Description automatically generated

Once I had chosen a design style I clicked on the ‘Colour’ tab and chose a colour that shows the data more clearly.

A screenshot of a chart

Description automatically generated

I then clicked on the legend inside my graph to open the format pop-out on the right side of the window. From here I checked the box called ‘Left’ to move the legend to the left side of my graph. I then unchecked the box called ‘Show the legend without overlapping the chart’ to make the map clearer and sit nicer on the screen.

A screenshot of a computer

Description automatically generated

1. To create a new worksheet I first needed to unlock my Workbook. To do this I clicked on ‘Protect Workbook’ again in the same place as last time.

A screenshot of a computer

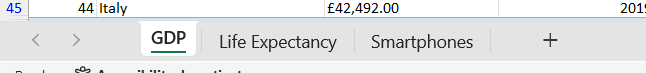
Description automatically generated

When I clicked on the button a pop-up appeared asking for the password. I entered my password here before pressing the ‘OK’ button.

A screenshot of a computer login box

Description automatically generated

Once I had unlocked my workbook I could navigate to the bottom of the page and press the ‘+’ to create a new worksheet.



I then renamed the worksheet for better organisation and to make the data easier to find for someone else who may be in my workbook. To rename it I left-clicked on the worksheet tab at the bottom of the screen and then clicked ‘Rename’.

A screenshot of a computer

Description automatically generated

I named the worksheet ‘GDP Heat Map’ because it is concise for optimum organisation and clarity.

A close up of a sign

Description automatically generated

I moved my finished graph to another sheet using copy and paste but before that, I added a title of ‘GDP Per Capita Heat Map’.

A map of the world

Description automatically generated

1. To sort the data to show the highest-ranking GDP I first had to sort them by GDP from highest to lowest. I started by selecting the drop-down next to ‘GDP – per Capita (PPP)’

A screenshot of a graph

Description automatically generated

I then selected ‘Sort Largest to Smallest’.

A screenshot of a computer

Description automatically generated

I then went back to the same drop-down next to ‘GDP – per Capita (PPP)’

A screenshot of a computer

Description automatically generated

Once in the drop-down, I went to ‘Number Filters’ and selected ‘Top 10…’

A screenshot of a computer

Description automatically generated

From there it opened a pop-up and I changed the number of items to ‘20’ before clicking ‘OK’.

A screenshot of a computer

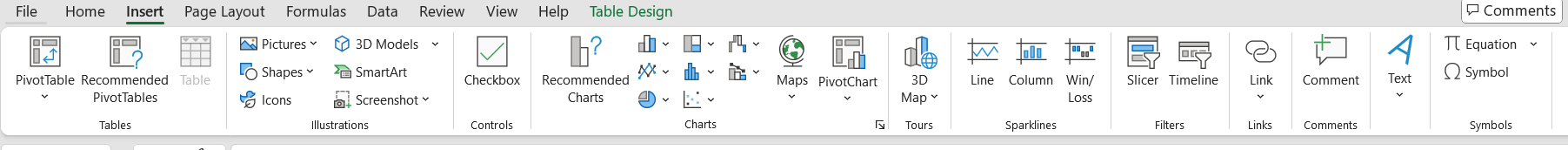
Description automatically generated

1. When creating my second graph I selected all the info I needed from the data, excluding ‘Year of Information’ as it is unnecessary for the graph.

A screenshot of a computer

Description automatically generated

To create the graph I went to ‘Insert’, then selected the ‘Bar Chart’ drop-down



From the drop-down I selected the first version of the ‘2-D Bar’. I then moved it below the table.

A screenshot of a graph

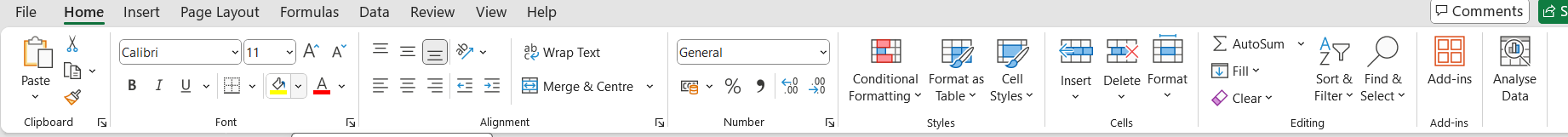
Description automatically generated

1. To make the graph stand out more I selected the area of cells behind it.

A screenshot of a computer

Description automatically generated

I then went to the ‘Home’ tab and selected the drop-down next to ‘Fill Colour’



That opened a pop-up where I selected a light Blue that worked well with my table and graph formatting.

A screenshot of a computer

Description automatically generated

This is the final layout of the data with my Bar Chart.

A screenshot of a computer

Description automatically generated

1. When creating a Macro in Excel you first need to add the ‘Developer’ tab to the Ribbon. You do this by going to ‘File’ and in the bottom left selecting ‘Options’

A screenshot of a phone

Description automatically generated

Once in the Options pop up you select the ‘Customise Ribbon’ tab and from there you tick the box next to ‘Developer’ and then click ‘OK’. Mine has already been ticked because I added it before.

A screenshot of a computer

Description automatically generated

Once the ‘Developer’ tab had been added I navigated there and selected ‘Record Macro’.

A screenshot of a computer

Description automatically generated

Once the pop-up appeared I filled in the ‘Name’, ‘Shortcut key’ and a short ‘Description’ of what the macro will do. I then clicked ‘OK’ to start recording.

A screenshot of a computer

Description automatically generated

When recording the Copy Macro I first selected all the data I wanted the macro to copy, I then left-clicked on the data and selected ‘Copy’ from the menu.

A screenshot of a computer

Description automatically generated

I then pressed ‘Stop Recording’ in the ‘Developer’ tab to stop the recording of the Macro.

A screenshot of a computer

Description automatically generated

I recorded two more Macros for ‘Print’ and ‘Save’ adding the Name, Shortcut and Description as seen below.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

Once I had started recording the Macros for ‘Print’ and ‘Save’ I clicked on the ‘File’ tab and from there pressed ‘Print’ and ‘Save’ respectively before pressing the back button and stopping the recording in the ‘Developer’ tab.

A screenshot of a computer

Description automatically generated

After creating all three Macros I then had to make buttons to execute the macros easily. I navigated to the ‘Developer’ tab and pressed the ‘Insert’ drop down.

A screenshot of a computer

Description automatically generated

From the drop-down menu I selected the ‘Button’ under ‘Form Controls’

A screenshot of a computer

Description automatically generated

I then drew a square button below my graph. Once I had drawn it then the pop-up appeared for assigning a Macro to the button. I selected the Macro I wanted to assign to that button and clicked ‘OK’

A screenshot of a computer

Description automatically generated

I did the previous step another two times for the remaining Macros and renamed the Buttons so it’s clear what they do when clicked. As shown below.

A screenshot of a graph

Description automatically generated

# Tableau

I started by creating relationships between the different tables



I planned to use multiple Worksheets for each table but keep consistency by having the same Worksheets in each dashboard.

I have included an example of each type of worksheet used for each table using the next table as the example. I used three Worksheets for each of the three tables, creating nine Worksheets in total. I will reference the data used for that example.

When creating my first Worksheet I added the ‘Country’ to rows and the average of the ‘GDP per Capita (PPP)’ to columns. I then added a filter for the top 20 ranked countries by filtering ‘Rank’ by the ‘Range of Values’ of 1 to 20 (I will talk about this after the Worksheets). I added colour because the client is colour blind by dragging the average of ‘GDP per Capita (PPP)’ over ‘Colour’ in ‘Marks’ and used the ‘temperature diverging’ gradient as it uses the most colours (Again I will expand on this after the Worksheets). I then ordered the columns by largest to smallest and added a title stating what the table is showing – the ‘Top 20 Countries’

A screen shot of a graph

Description automatically generated

My second Worksheet is a map to show the global distribution of the data. I started by dragging ‘Country (Life Expectancy)’ onto the worksheet to create the map. I then added the average of ‘Life Expectancy’ onto both ‘Colour’ and ‘Size’ in ‘Marks’ so you can easily see the values compared to each other and the relative ranking with the ‘temperature diverging’ gradient. Once again, I filtered the data by ‘Rank (Life Expectancy)’ to show only the top 20 countries. There were some null values, so I filtered them out too. Tableau removed them for me automatically when I clicked the grey box that appeared on the bottom right, this is why ‘Latitude (generated)’ and ‘Longitude (generated)’ are in ‘Filters’. I also added a Title explaining what the map is showing – ‘Global Distribution’.

A map of the world

Description automatically generated

In my third Worksheet, I created a Treemap to use as a filter on my dashboard (I will explain how I did this later). I started by dragging the average of ‘Smartphone Users’ to both ‘Colour’ and ‘Size’ in ‘Marks’, once again adding the ‘temperature diverging’ gradient for consistency throughout my dashboard. However, Tableau auto-created a map visualisation, but I already had this, so I clicked ‘Show Me’ in the top right of the screen and selected ‘Treemaps’. I dragged ‘Country (Smartphone Users)’ onto ‘Label’ in ‘Marks’ so it was clear which country’s data was being shown by each tablet. I then filtered by ‘Rank (Smartphone Users)’ to show the top 20 countries.

A screenshot of a computer

Description automatically generated

This is how I filtered ‘Rank’, ‘Rank (Life Expectancy)’ and ‘Rank (Smartphone Users)’ by selecting the ‘Attribute’ from the bottom of the list of averages. I selected the range of values of 1 to 20 to show the top 20 ranked countries. I did this in all nine of my Worksheets.

A screenshot of a computer

Description automatically generated

This is how I set up my Colour Legend. I used the ‘Temperature Diverging’ gradient as it was a full gradient using multiple colours and not just a smooth transition between two colours which could have been confusing for someone with colour blindness like the client.

At the bottom below ‘<< Advanced’ I have selected the start and end of the legend to be fixed at the total range of the top 20 countries. I only did this for my distribution maps so that when using my Treemaps as filters on the dashboard it would keep the size and colour the same instead of making it red and maximum size and confusing the client.

A screenshot of a computer

Description automatically generated

This is how I set up my Size Legend for both my distribution maps and Treemaps, again I only used the advanced settings for my distribution maps to keep he size the same when applying a filter in the dashboard.

A screenshot of a computer

Description automatically generated

I used my Treemaps as filters in the Dashboard by clicking the visual and selecting the ‘Use as Filter’ button to the top right. I did this in all 3 dashboards.

A screenshot of a computer screen

Description automatically generated

However, I did not want my tables to be filtered by the Treemaps so I selected the table, clicked ‘More Options’ and from the drop-down selected ‘Ignore Actions’. This meant that when selecting a country on the treemap the table was not affected and still showed all 20 countries in order of their rank.

A screen shot of a graph

Description automatically generated A screenshot of a computer

Description automatically generated

Once I had created all three dashboards for ‘GDP’, ‘Life Expectancy’ and ‘smartphone Users’ I created a story and labelled them as ‘Top 20:’ plus their respective names. I then gave my story the Title of ‘Wealth of Nations’ as this is the title of the data used to create the story.

A white rectangle with black text

Description automatically generated

# References

Quick Guide to the Principles of Data Protection by Data Protection Commission -<https://www.dataprotection.ie/sites/default/files/uploads/2019-11/Guidance%20on%20the%20Principles%20of%20Data%20Protection_Oct19.pdf>