29/03/2024

Anish Mistry

Assignment 1 Data Visualization

Contents

[Policies and Procedures 3](#_Toc162599388)

[Excel 4](#_Toc162599389)

[Part 1 4](#_Toc162599390)

[1. Set a password to protect the workbook 4](#_Toc162599391)

[2. Highlight column C and change the data to display in British Pound symbol 4](#_Toc162599392)

[3. Turn the GDP sheet into a table. 5](#_Toc162599393)

[4. Filter the table to display only the information for 2019 5](#_Toc162599394)

[5. Next create a chart that will only display the following data ‘Rank, Country, and GDP - per capita (PPP). The chart can be anything as long as it is suitable. 6](#_Toc162599395)

[6. Using your creative skills edit the chart 6](#_Toc162599396)

[7. Move the chart to a new sheet tab and label with a suitable name 7](#_Toc162599397)

[8. Create a sort for the top 20 highest ranking counties 7](#_Toc162599398)

[9. 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. 8](#_Toc162599399)

[10. Colour the background by highlighting the area underneath the table. 8](#_Toc162599400)

[Part 2 9](#_Toc162599401)

[1. Create 3 macro buttons, print the sheet, save the file, and copy the sheet. 9](#_Toc162599402)

[2. Using the copy macro, copy the sheet and then paste it into a new word document keeping the formatting. Give the page a title ‘GDP (Gross domestic product)’. 9](#_Toc162599403)

[3. Save your document as ‘Word Gross domestic product report 1’ 10](#_Toc162599404)

[4. Add a header and footer to the table. 10](#_Toc162599405)

[5. In the header enter your name and GLA DATA 1 in the three boxes 10](#_Toc162599406)

[6. In the footer add todays date then Assignment 1 and lastly Data Visualisation. 11](#_Toc162599407)

[7. Return your view to normal 11](#_Toc162599408)

[8. Save your table as ‘Excel Gross domestic product report 1’ 11](#_Toc162599409)

[9. Close your word document only. 11](#_Toc162599410)

[Tableau 13](#_Toc162599411)

[Import Data 13](#_Toc162599412)

[Set Relationships 13](#_Toc162599413)

[Check Data Types 14](#_Toc162599414)

[Build Charts 15](#_Toc162599415)

[Map Charts 15](#_Toc162599416)

[Comparison Charts 15](#_Toc162599417)

[Comparison Graphs 16](#_Toc162599418)

[Scatter Plot – GDP vs Life Expectancy 16](#_Toc162599419)

[Scatter Plots – not useful plots 17](#_Toc162599420)

[Building Dashboards 17](#_Toc162599421)

[Maps Dashboard 17](#_Toc162599422)

[Charts Dashboard 18](#_Toc162599423)

[FINAL DASHBOARD 18](#_Toc162599424)

[Reflective 19](#_Toc162599425)

# Policies and Procedures

When working with sensitive data it is important to follow clear policies and procedures to ensure data is handled responsibly, securely and in compliance with legal requirements such as the General Data Protection Regulation (GDPR).

The UK GDPR lists these seven key principles for data protection: lawfulness, fairness and transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality (security), and accountability. These can be summarized into the data is accurately collected and to be limited to what is necessary and processed for its intended purpose only. The data should be accurate and stored only for the time required securely.

It is important these the GDPR rules are adhered to because people have the right to know about and have control over what information gets collected about them and how it’s further used or shared with. Depending on the context, the data may be enough to identify an individual person, so must be stored securely and only be used for purposes no further than initially set out. A data analyst should be aware of these rules as manipulating, altering, or using the data in non-agreed upon manner could be considered breaking the law.

# Excel

## Part 1

### Set a password to protect the workbook

A screenshot of a computer

Description automatically generated

Password: 123

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

A screenshot of a computer

Description automatically generated

Changed to display GBP (£) and made sure the number format is currency (highlighted)

### Turn the GDP sheet into a table.

A screenshot of a computer

Description automatically generated

### Filter the table to display only the information for 2019

A screenshot of a computer

Description automatically generated

Created a filter macro for the data for 2019 and all years and assigned them to buttons so it is easy to toggle between the data.

### Next create a chart that will only display the following data ‘Rank, Country, and GDP - per capita (PPP). The chart can be anything as long as it is suitable.

A screenshot of a computer

Description automatically generated

Added bar chart with Rank, Country, and GDP

### Using your creative skills edit the chart

1. Add a title
2. Add X and Y axis labels
3. Make the chart visually pleasing

A screenshot of a computer

Description automatically generated

Added a title, labelled both axis, added minor gridlines and readjusted the axis scaling.

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

A screenshot of a computer

Description automatically generated

Moved chart to new sheet and called it ‘GDP Chart’

### Create a sort for the top 20 highest ranking counties

A screenshot of a computer

Description automatically generated

Created a filter macro for the data for the top 20 countries and all countries and assigned them to buttons so it is easy to toggle between the data.

The two filters perform as an AND logic operator, hence there are only 13 countries showing. To perform an OR logic operator for the two filters, a more advanced feature like pivot tables would need to be used, but I felt it was out of the scope for this assignment.

### 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.

A screenshot of a computer

Description automatically generated

Created a new bar chart and placed it underneath the table.

### Colour the background by highlighting the area underneath the table.

A screenshot of a computer

Description automatically generated

Coloured the background to match the table. I adjusted the scale of the chart and added minor gridlines and a legend. I used the SUBTOTAL() formula to show a summary of the information for the visible cells and camera function to have the table stay at the top.

## Part 2

### Create 3 macro buttons, print the sheet, save the file, and copy the sheet.

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

Created three buttons for copy, save and print. I have assigned a copy macro to copy button along with a shortcut: ctrl + shift + c.

### Using the copy macro, copy the sheet and then paste it into a new word document keeping the formatting. Give the page a title ‘GDP (Gross domestic product)’.

Evidence below.

### Save your document as ‘Word Gross domestic product report 1’

A screenshot of a computer

Description automatically generated

### Add a header and footer to the table.

Evidence below

### In the header enter your name and GLA DATA 1 in the three boxes

A screenshot of a computer

Description automatically generated

### In the footer add todays date then Assignment 1 and lastly Data Visualisation.

A screenshot of a computer

Description automatically generated

### Return your view to normal

### Save your table as ‘Excel Gross domestic product report 1’

A screenshot of a computer

Description automatically generated

### Close your word document only.

Document attached with assignment

A screen shot of a computer program

Description automatically generatedA screenshot of a computer

Description automatically generated

I have also assigned macros to the ‘Print’ and ‘Save’ buttons.

# Tableau

## Import Data

A screenshot of a computer

Description automatically generated

Data is imported into Tableau.

## Set Relationships

A screenshot of a computer

Description automatically generated

The relationship between the tables is set as all of them have ‘Country’ as a common column.

## Check Data Types

A screenshot of a phone

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a blue screen

Description automatically generated

Adjusted the data types for date column for the tables all.

## Build Charts

### Map Charts

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

I have made three charts, each showing a map of the top 20 countries for each table. The colours and size both represent the magnitude of the value represented. I have added symbols for each type of data and customised the map to provide better contrast for a colourblind person. There is also a label for each of the top 20 countries.

### Comparison Charts

A screenshot of a graph

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

I have created three charts to show the top 20 countries for each table. There is reference line for each chart showing the global average. There is a label showing the value for each bar. I have changed the axis for the life expectancy to start at 50 years to see a difference more clearly in life expectancy.

### Comparison Graphs

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

These are dual axis graphs for the top 20 countries. Each of the measures has a unique colour consistent across the three graphs.

They do not show any strong corelations for data for the top 20 countries.

### Scatter Plot – GDP vs Life Expectancy

A screen shot of a computer

Description automatically generated

Taking global data into account, this scatter plot shows a clear trend. As GDP per capita (PPP) increases, the life expectancy at birth for a person living in that country increases. A logarithmic trend line gives the lowest r-squared value, hence that is what is used.

I have use two contrasting colours to show the plots and the trend line, so it is easy to see. The graph for only the top 20 countries does not show enough of a trend hence I have used data for all countries.

### Scatter Plots – not useful plots

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

The GDP vs Smartphone users and Smartphone users vs Life Expectancy graphs do not show any corelation, taking into account all the countries.

## Building Dashboards

### Maps Dashboard

A screenshot of a computer

Description automatically generated

I have built this dashboard by creating a parameter which takes the names of the different maps and used it create a calculated field which is used to filter between the three maps. The title is dynamic along with the map.

Charts DashboardA screenshot of a computer

Description automatically generated

I have also created a dashboard for the three bar charts using the same method. Notice the change in title to Top 20 “Smartphone Users” vs Global Average from the GDP chart above.

### FINAL DASHBOARD

A screenshot of a computer

Description automatically generated

Combining the learnings from the previous two dashboards, I made a final dashboard with all the information. There is the map and chart visuals which can be toggled through, and some key stats, such as the maximum, minimum, and average values, highlighted in a short text field on the left. The tooltips for both the visuals contain the name of the county, rank, and its respective value. The visuals can be interacted with by selecting a bar on the graph and the country will be focused on its accompanying map.

# Reflective

The process of filtering the excel table to show the relevant information was very simple, as was creating the graph from the table. I spent a bit of time formatting the graph to make sure that all the information was clearly visible and looked pleasing.

I found creating the copy macro was simple so I went further by also assigning macros to print and save, the former causing trouble as I found the document would be sent to print without showing a preview page unless otherwise stated in the macro.

I also went further by adding a summary table of the visible information. It uses the SUBTOTAL() formula to show the data in the visible cells. I initially created it next to the main table, but as the filters were applied, those rows would also hide alongside the filtered data. I then placed it underneath the main table and used the camera tool to create an image of those cells to be place in a more user-friendly place.

Moving onto Tableau, I initially found it difficult to find a story within the data, but after some research and finding some inspiration online of similar data showcasing data, I had a better understanding of how the data can be presented. I researched how to use page swapping so all the data can be on one dashboard without taking up too much space. I then finished off by combining both all the charts alongside some additional stats to bring the whole dashboard together. The comparisons between the tables were difficult to make due to the number of null values but a trend could be found in the GDP vs Life Expectancy chart, but I had to move away from just looking at the top 20, rather focus on the whole data instead.

To improve the usefulness of the smartphone users data, the populations of the countries could also be added and a calculated field of smartphone users per capita could be made. This could then be compared with GDP per capita to create a potentially meaningful graph as the current table of the magnitude of smartphone users gets skewed because of the comparatively large populations of China and India. I could not find a reliable source of data for the populations of the countries so I did not create this data.