An analysis of the generated top 20 countries by GDP per capita, highlighting the economic wealth per person within these nations. The data offers insights into global economic distribution and growth, identifying countries that have achieved significant financial prosperity

Analysis of the Top 20 Countries by GDP per Capita:

Economic Insights, Global Wealth Distribution, and Adherence to Data Handling Policies

Bernerditta Nyarko
Date of submission:22/10/24

Table of Contents

Data Privacy and Protection Policies

General Data Protection Regulation (GDPR) (for EU data)

This is a regulation that applies to data processing activities of organizations within the European Union. It governs how personal data can be collected, stored, and used, and provides rights to individuals regarding their personal data.

Non-compliance can lead to significant fines. If the data being used involves any personal identifiers (such as names, financial data, or geographical identifiers), it must be anonymized, or explicit consent must be obtained. The Wealth of Nations data, if dealing with specific people or personal data, may need GDPR adherence to avoid violating individual privacy rights.

California Consumer Privacy Act (CCPA) (for California)

Similar to GDPR, CCPA regulates how data from California residents is handled. It provides consumers with rights over their personal information.

 If "The Wealth of Nations" data contains information on Californians, their data must be handled per CCPA guidelines.

Data Anonymization

In compliance with privacy laws like GDPR and CCPA, sensitive information must be anonymized to prevent identification of individuals.

Anonymization ensures that the dataset cannot be traced back to an individual, protecting personal privacy and reducing legal risks. Analysts need to use anonymized data to avoid legal complications.

When working with **data** like "The Wealth of Nations," several **policies and procedures** need to be adhered to. These policies ensure the ethical, legal, and secure use of data, protecting both the integrity of the data and the rights of individuals or organizations involved. Here's an outline of key policies, why they are essential, and why it is important for a data analyst to be aware of them:

Data Security Policies

Data Encryption

Encryption is a process used to protect data by converting it into a form that is unreadable to anyone who doesn't have the proper authorization to access it. Encrypting data ensures that any sensitive or confidential data is protected from unauthorized access. This could be relevant if "The Wealth of Nations" data contains proprietary or sensitive information about companies or nations.

Data analysts should encrypt data at rest (when stored) and in transit (when being transferred) to prevent data breaches and ensure compliance with data protection standards.

Access Control and User Authentication

Limiting access to the data ensures that only authorized personnel can view or modify it. Analysts need to follow strict access control procedures to ensure that the data is not misused. A way in which this strict access can be enforced through procedure is This practice can be done using passwords on a worksheet as I have done in the data set in excel to prevent unauthorised users from accessing information that they should not have access to.

Failure to adhere to access control policies could result in unauthorized data manipulation, theft, or breaches.

Ethical Use of Data

Data Accuracy

Data analysts are responsible for ensuring that the data used in analysis is accurate and valid. Inaccurate data can lead to incorrect conclusions or biased results. Inaccurate data can be prevented by removing Duplicates, Handling Missing Values, Correcting Inconsistent Data and Dealing with Outliers.

It is crucial to avoid spreading misinformation or biased conclusions, especially in significant studies like "The Wealth of Nations," which might influence economic policies.

Transparency and Accountability

When analysing and reporting data, the process and sources must be transparent. This fosters trust and accountability.

Analysts should ensure that all sources of data are properly cited, and the methodology used in the analysis is transparent to avoid misinterpretation or misuse.

Intellectual Property (IP) and Copyright Policies

Open Data and Licensing Agreements

Some data is open for public use, but it may come with licensing agreements, such as Creative Commons licenses, that specify how the data can be used.

Analysts need to review licensing terms to ensure compliance, especially if the analysis or results will be published or shared with others.

Data Retention and Disposal Policies

Data Retention

Policies regarding how long data can be retained should be followed. Data that is no longer needed should be securely deleted to protect privacy and reduce risks. Failing to securely dispose of outdated data could result in accidental data leaks or misuse.

Compliance with National and International Regulations

Financial Data Regulations

If the "Wealth of Nations" data includes financial information, compliance with financial data regulations, like Sarbanes-Oxley (SOX) in the U.S., is crucial.
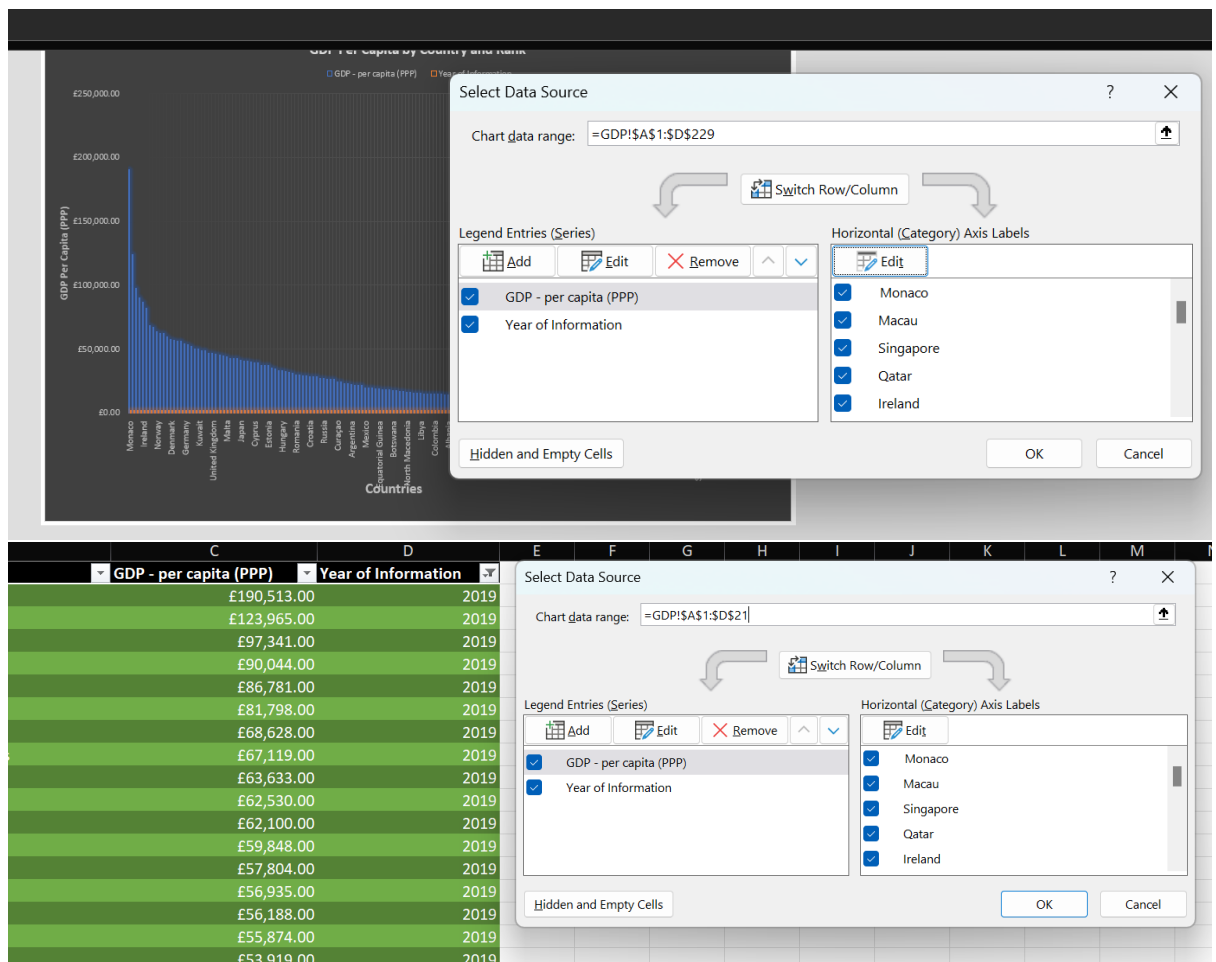
Ensuring that financial data is handled according to regulatory standards prevents issues like fraud or inaccurate reporting, which can have wide-reaching consequences.

The importance of practicing these policies for Data Analysts

As a data analyst, understanding and adhering to these policies is critical for several reasons:

1. Legal Compliance: Failure to comply with data privacy and protection laws like GDPR or CCPA could lead to legal penalties for the analyst and their organization.

2.Data Integrity: By following policies related to data accuracy, ethical use, and security, analysts ensure that their work produces reliable and credible results.

3. Reputation and Trust: Adhering to ethical and transparency standards builds trust with stakeholders, ensuring that analyses and insights are viewed as credible and unbiased.

4. Prevention of Data Breaches: Understanding data security policies helps prevent breaches, safeguarding sensitive data from being exposed or stolen.

Main Body: Process Documentation with Screenshots

I changed the GDP per capita (ppp) column to British pound currency

I sorted the GDP per capita (PPP) column in descending order so that the countries with the highest GDP per capita appeared at the top.

Next, I applied a filter to show only the data for the year 2019,by filter by selected cell's Value and entered 2019 (since that was the only year I was focusing on). This helped me isolate just the data I needed.
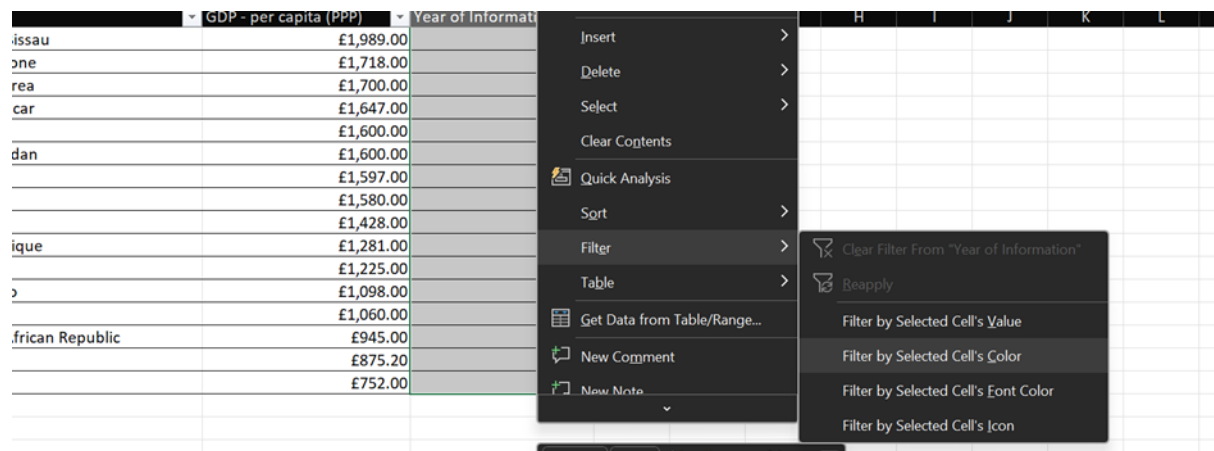
After that, I highlighted the relevant data range, which included the country names and their GDP per capita values for 2019.

In the "Select Data Source" dialog, I chose the data range up to row 21, which made sure my chart would only include the top 20 countries with the highest GDP per capita.

Finally, I clicked "OK" to confirm, and the chart updated to show exactly what I needed: the top 20 countries by GDP per capita for 2019.
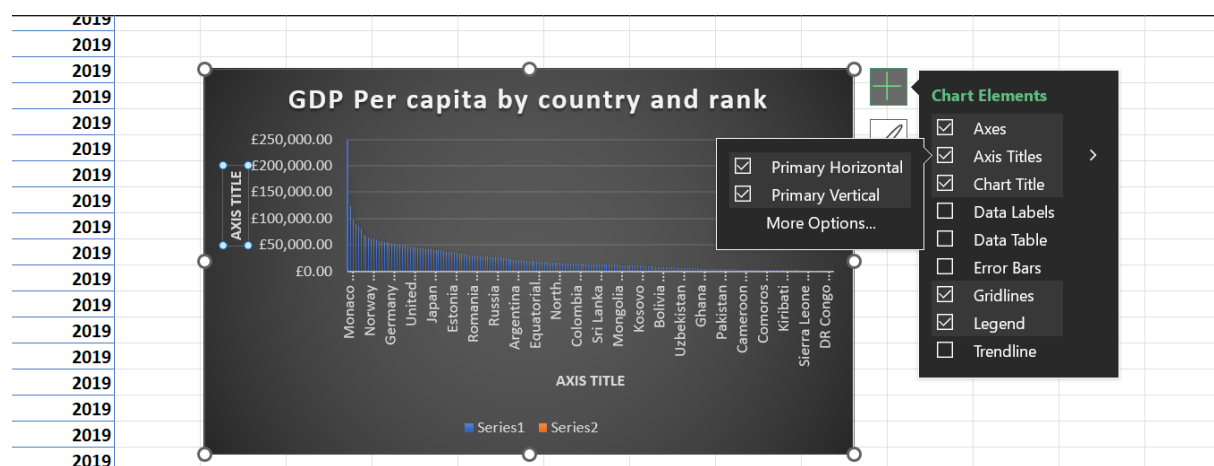
Next I created a chart that will only display the following data 'Rank, Country and GDP - per capita (PPP) and added a title

Here I was making sure the year 2019 was selected .
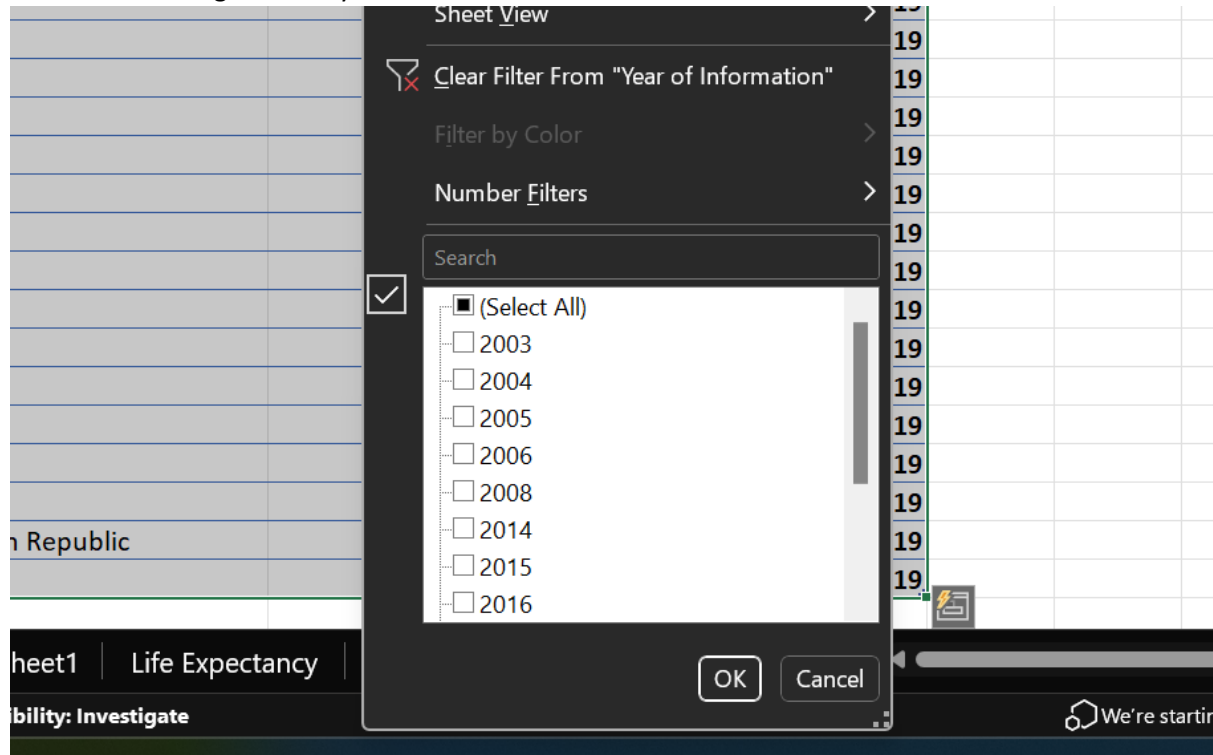

Next

Applying the "Top 20" Filter:

I opened the AutoFilter options for the GDP column (column C: "GDP per capita (PPP)"), as seen by the "Top 10 AutoFilter" dialog that is currently displayed.

In this dialog, I changed the setting to show the top 20 items (instead of the default 10). This filter will ensure that only the top 20 countries with the highest GDP per capita values are shown.
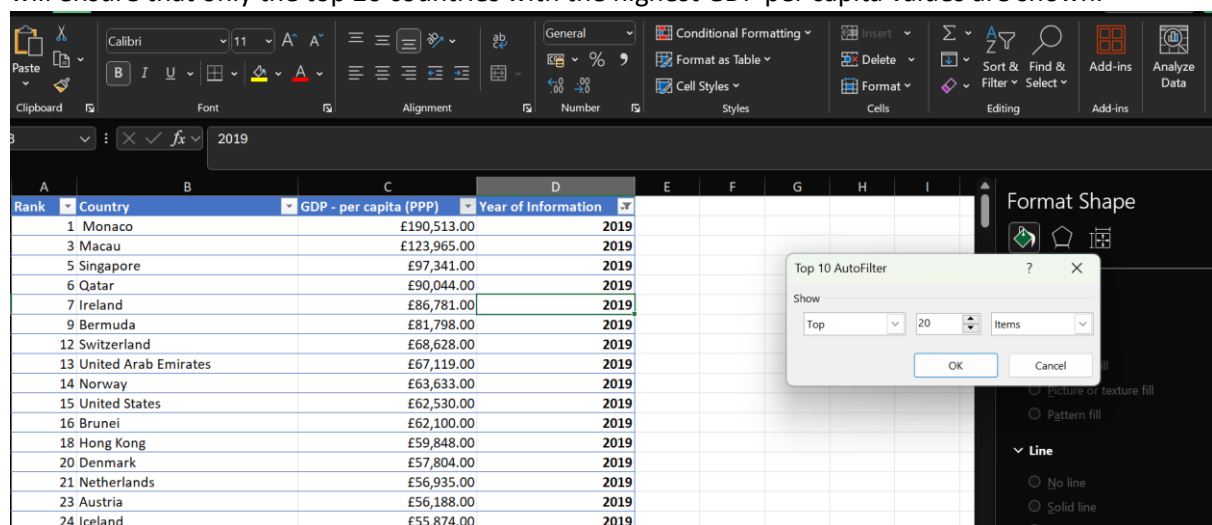
After filtering and formatting the table to the year to 2019 I created an additional chart referenced above and saved and imported to Tableau .With the following excel dataset created four visuals including a tree map, pie chart, Bar chart and a country Map. I used the following filter for each

visual in order to highlight the top 20 countries .
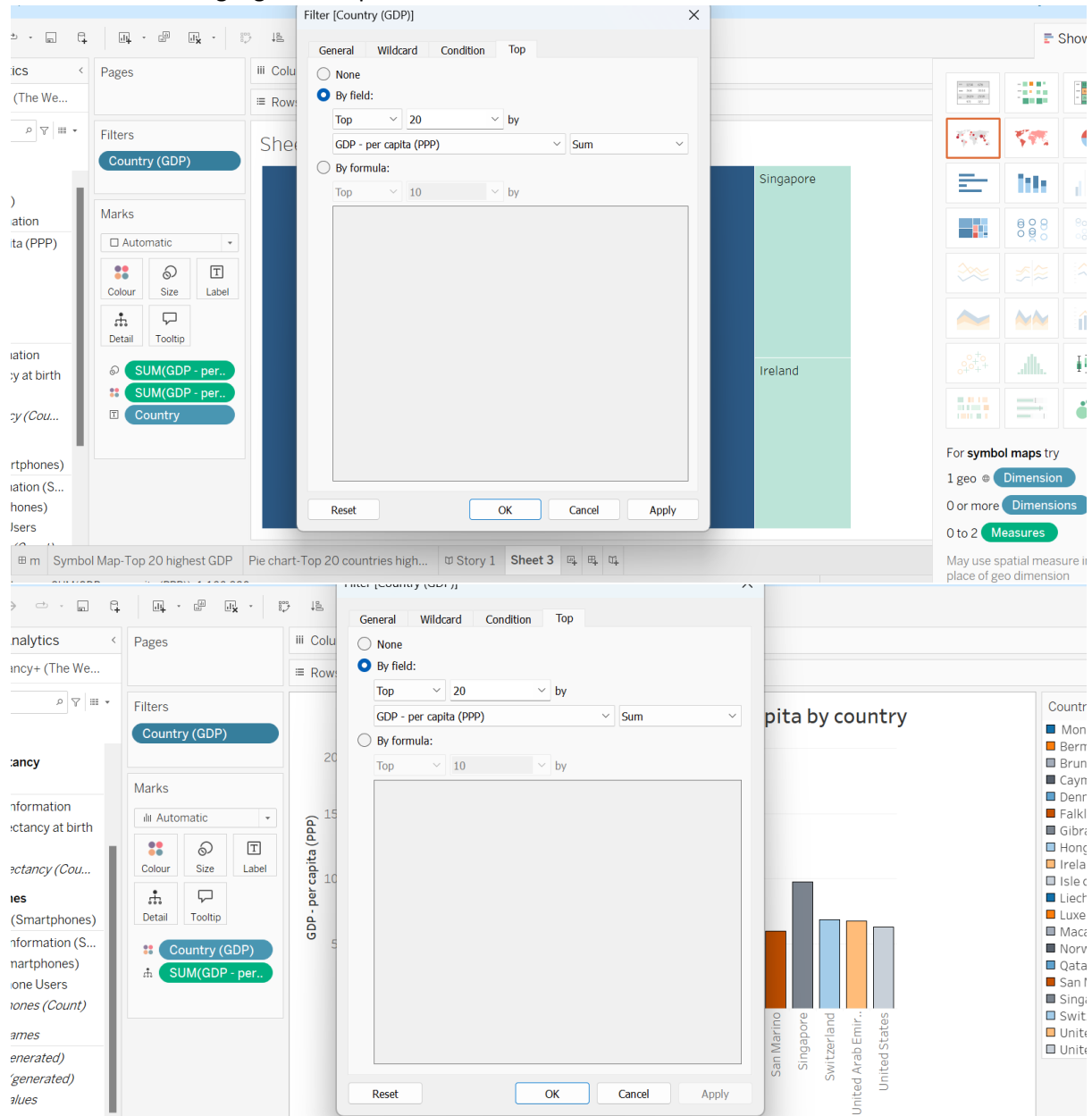
The following four visuals I created were compiled into a dashboard.

Reflective Account: Evaluation of Your Project

The project was not very difficult overall, but I did encounter a few challenges early on. One of the main issues I faced was figuring out how to filter the data to display only the top 20 countries by GDP. Initially, it wasn't clear how to apply the correct filter for this, but once I used Excel's AutoFilter feature and set the filter to show the top 20 items in the GDP column, I was able to resolve that issue smoothly.

When I transferred the data to Tableau, I ran into another challenge. While creating a heat map, certain countries were not appearing correctly, and some were even showing up as "unidentified country." This was confusing at first, and I struggled to understand why this was happening. After some trial and error, I realized that I needed to adjust the marks section in Tableau. By dragging the Country (GDP) field multiple times and dropping it onto the label section, the missing countries finally appeared correctly in the visualization.

What Went Well

The data preparation in Excel went well once I figured out how to filter and sort the data correctly.

The transfer of data into Tableau was straightforward, and I managed to create visualizations without any data loss or formatting issues.

Areas for Development

I need to improve my familiarity with Tableau's various visual elements, especially when using geographical or heat map features. Understanding how to manipulate the marks and labels sections better would have saved me some time.

In will try and practice more advanced filtering techniques and visualizations to make my analysis more effective and detailed in my own time or with the websites on class materials catered towards excel studies.