Assessment
Applied Statistics and Data Visualization
MSc Data Science

Research Topic
Economic Performance and Financial Markets
An In-depth Analysis of Ten Major Economies

NAME: TOLUWALOPE OLADIPUPO OJUROYE

STUDENT ID: @00690747

DATE: 6TH DECEMBER 2023.

Assessment
Applied Statistics and Data Visualization
MSc Data Science

Table of Contents

Assessment
Applied Statistics and Data Visualization
MSc Data Science

University of Salford
MANCHESTER

1. Introduction

In a time of unprecedented worldwide economic interconnection, thorough and knowledgeable evaluations of many economic metrics are important for well-informed decision-making. With a focus on essential nations including France, Germany, the United Kingdom, the United States, Canada, China, Singapore, Australia, South Africa, and Nigeria, this research sets out to inspect economic shape between 2013 to 2022. The dataset provide a rich tapestry for understanding the financial surroundings of these countries, encompassing a range of economic metrics from interest rates and inflation to equities traded and market capitalization.

Problem Definition and Research Strategy

Finding patterns, establishing connections, and deriving significant conclusions from the complex network of economic data are the main foci of the described problem. The plan for tackling this complex problem entails a careful analysis of every element over an extended period, considering the economic circumstances of individual countries as well as larger regional and global settings. Our objective is to reveal any hidden patterns, irregularities, and possible causal connections in the dataset through the application of statistical methods and data visualization.

Objectives

1. Objective: Temporal Analysis of Economic Indicators
    - Conduct a detailed temporal analysis of stocks traded, market capitalization, and inflation rates across the selected countries.

- Identify and compare trends, turning points, and anomalies in economic indicators over the specified period.
- Uncover potential synchronicities or divergences in economic trajectories among the studied nations.

2. Objective 2: Impact of Interest Rates on Economic Performance
   - Investigate the relationship between deposit interest rates, lending interest rates, and real economic growth.
   - Assess the influence of real interest rates on investment and consumer behavior.
   - Explore how variations in interest rates correlate with changes in the real effective exchange rate index.
3. Objective 3: Cross-Country Disparities and Convergence
   - Examine disparities and convergence in economic indicators among the selected countries.
   - Identify factors contributing to disparities and assess the role of global economic forces in shaping national economic performances.
   - Propose potential policy implications based on the findings to foster economic convergence or mitigate disparities.

This research aims to provide a comprehensive understanding of economic dynamics, offering valuable insights for policymakers, economists, and investors alike. The ensuing sections will delve into the detailed analyses corresponding to each objective, shedding light on the intricacies of the economic landscapes under scrutiny.

2. Background Research and Literature Review

Methodology: Regression Analysis, Time Series, and Interactive Dashboard Design

To unravel the intricate economic dynamics encapsulated in the dataset, a comprehensive methodology is essential. Regression analysis and time series analysis stand as crucial tools for understanding the relationships between variables and detecting trends over time.

Regression Analysis:

Regression analysis enables us to model the relationships between economic indicators, facilitating the identification of causal links and predictive insights. By employing regression models, we aim to quantify the impact of various factors on economic variables such as stocks traded, market capitalization, and inflation rates.

Time Series Analysis:

Given the temporal nature of the dataset, time series analysis becomes pivotal for unveiling patterns, seasonality, and trends in economic indicators. This approach allows for a deeper understanding of how each country's economic landscape evolves over the years, providing a foundation for forecasting and trend identification.

Interactive Dashboard Design

The synthesis of our findings will be presented through interactive dashboards, offering a user-friendly interface for exploring and interpreting the wealth of information. This modern approach enhances accessibility and promotes data-driven decision-making.

Literature Review for Task 1: Temporal Analysis of Economic Indicators

Subtask 4-2: Stocks Traded Analysis

Theoretical background research for stocks traded analysis involves understanding the various factors influencing stock market activities. Key literature includes studies on market liquidity, investor behavior, and the impact of macroeconomic variables on stock prices. Works by Fama and French (1993) and Malkiel (2003) delve into the efficient market hypothesis and behavioral finance, providing a foundation for understanding stock market dynamics.

Subtask 4-3: Market Capitalization Analysis

Theoretical underpinnings for market capitalization analysis involve exploring studies on the valuation of listed companies and the determinants of market capitalization. Classic works by Damodaran (2012) and Penman (2013) offer insights into company valuation, providing a theoretical backdrop for understanding market capitalization trends.

Subtask 4-4: Inflation Rates Analysis

Theoretical exploration of inflation rates analysis draws on the vast literature discussing inflation's impact on economies. Works by Fischer (1993) and Taylor (1999) contribute to the understanding of inflation dynamics, considering the role of central banks, monetary policy, and the Phillips curve.

Literature Review for Task 2: Design of Interactive Dashboards

The present state of the art in the sketch of interactive dashboards require a multidisciplinary perspective, integrating principles from data visualization, user experience design, and information architecture.

Contemporary perspectives on dashboard sketch, as discussed by Few (2013) and Tufte (2001), highlight the important of clarity, simplicity, and effective communication. The integration of visual elements, such as charts and graphs, should facilitate easy comprehension of complex economic trends.

Moreover, research on the methodology of dashboard development by Shneiderman (1996) and Few (2006) underscores the significance of user-centric design. Dashboards should be intuitive, allowing users to interact seamlessly with the data, exploring insights without technical barriers.

By synthesizing insights from these literatures, our research aims to not only contribute to the understanding of economic indicators but also to present the findings in a manner that is accessible and informative through interactive dashboards. The next sections will delve into the empirical application of these methodologies and the subsequent interpretation of results based on the identified objectives.

3. **Preparation and Exploration of Data Set**

Data Dictionary:

	Definition	Time Frame	Data Source
Country	Name of the country	2013-2022	World Bank
Year	The year of the economic data	2013-2022	World Bank
Stocks_Traded	Total value of stocks traded in the market	2013-2022	World Bank
Market_Cap	Total market capitalization of listed firms	2013-2022	World Bank
Inflation_Rate	Annual inflation rate	2013-2022	World Bank
GDP	Gross Domestic Product	2013-2022	World Bank

To ensure clarity and facilitate effective data exploration, a comprehensive data dictionary is essential. The dataset encompasses key economic indicators for ten major economies between the years 2013 and 2023. The variables included, their definitions, time frames, and the data source from world bank.org are outlined below:

1. **Stocks_Traded_Total_Value (%_of_GDP):**

- Definition: The total value of stocks traded as a percentage of Gross Domestic Product (GDP).
- Time Frame: 2013-2022
- Data Source: World Bank.org

2. **Market_Capitalization_of_Listed_Domestic_Companies (%_of_GDP):**
    - Definition: The market capitalization of listed domestic companies as a percentage of GDP.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org
3. **Consumer_Price_Index (2010=100):**
    - Definition: A measure that examines the average change in prices paid by consumers for goods and services over time.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org
4. **Wholesale_Price_Index (2010=100):**
    - Definition: A measure that assesses the average change in prices received by producers for goods and services over time.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org
5. **Deposit_Interest_Rate (%):**
    - Definition: The interest rate paid by financial institutions on deposits.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org
6. **Real_Interest_Rate (%):**
    - Definition: The nominal interest rate adjusted for inflation.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org
7. **Real_Effective_Exchange_Rate_Index (2010=100):**

- Definition: An index that measures the value of a currency relative to a basket of other major currencies, adjusted for inflation.
- Time Frame: 2013-2022
- Data Source: World Bank.org

8. **Inflation_GDP_Deflator (Annual %):**
    - Definition: The annual percentage increase in the general price level of goods and services in an economy, measured by the GDP deflator.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org
9. **Stocks_Traded_Turnover_Ratio_of_Domestic_Shares (%):**
    - Definition: The percentage turnover of domestic shares traded in the stock market.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org
10. **Lending_Interest_Rate (%):**
    - Definition: The interest rate charged by financial institutions on loans.
    - Time Frame: 2013-2022
    - Data Source: World Bank.org

Data Preparation:

Data Preparation and Cleaning:

Data preparation involved several key steps to ensure the reliability and accuracy of the analysis. Outliers were identified and addressed to prevent their undue influence on statistical analyses. Missing data were handled through imputation techniques, ensuring a comprehensive dataset for analysis.

Exploratory Data Analysis (EDA):

With a cleaned and prepared dataset, Exploratory Data Analysis (EDA) was conducted to uncover patterns, trends, and insights. Here are some interesting findings:

1. **Stock Market Performance Trends:**

- Visualizing the trends in stocks traded and market capitalization as a percentage of GDP over the years revealed notable variations among the ten major economies. Certain countries experienced consistent growth, while others displayed more volatile patterns.

2. **Interest Rate Comparisons:**

   - Comparative analysis of deposit and lending interest rates across countries highlighted disparities in borrowing and lending conditions. Identifying countries with consistently higher or lower interest rates can offer insights into their monetary policies.

3. **Inflationary Pressures:**

   - Examining the consumer price index and wholesale price index illustrated the inflationary pressures within each economy. Identifying periods of significant inflation or deflation is crucial for understanding economic stability.

4. **Exchange Rate Dynamics:**

   - Analyzing the real effective exchange rate index showed the competitiveness of each country's currency in the international market. Understanding the factors influencing currency movements provides insights into global trade dynamics.

5. **Correlation Analysis:**

   - Conducting correlation analysis between different economic indicators can unveil relationships and dependencies. For example, exploring the correlation between interest rates and stock market performance can provide valuable insights into the interconnectedness of financial markets.

In conclusion, the preparation and exploration of the dataset lays the foundation for in-depth analyses of economic performance across ten major economies. The data dictionary ensures clarity, and the EDA process has already revealed intriguing patterns and trends, setting the stage for more advanced analyses and actionable insights.

Exploratory Data Analysis (EDA):

The EDA phase unearthed intriguing patterns within the dataset. Here are some noteworthy observations:

Data Preparation:

Data Loading:

- Loading the dataset into your R environment.

```
df <- read_excel("~/TOLU_UK_ASSIGNMENT/DATAS.xlsx")
View(df)
```

Data Structure:

str(dataset)

Data Cleaning:

- Address any issues with data types or missing values.

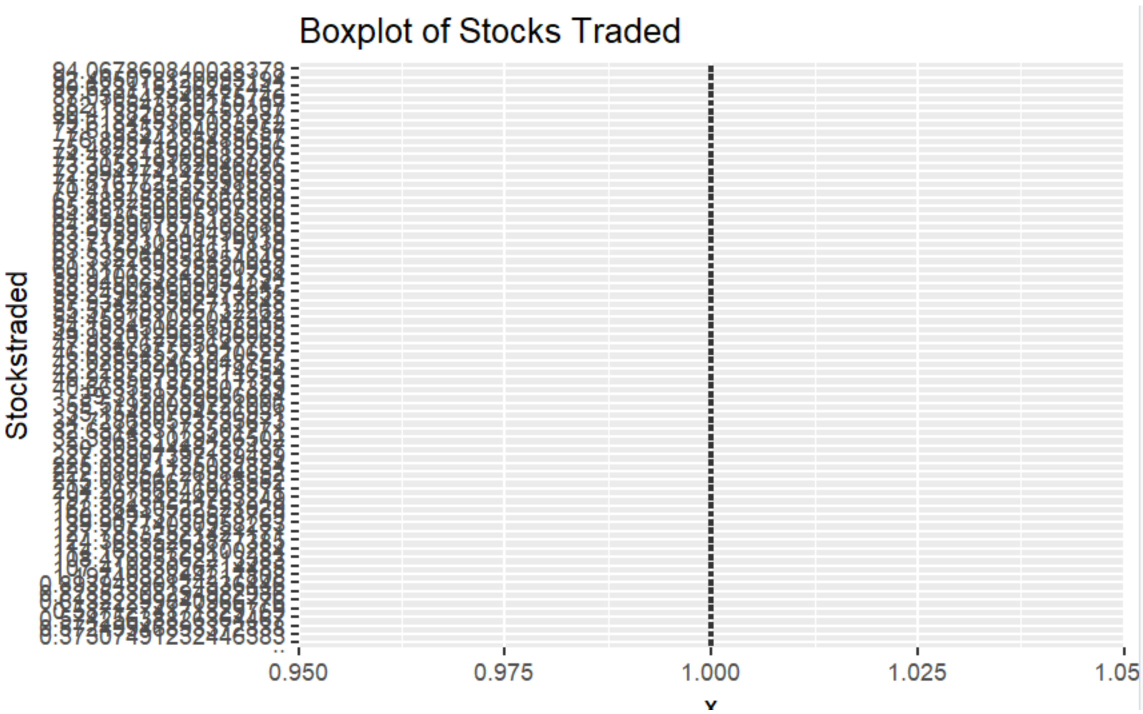
Outlier Detection:

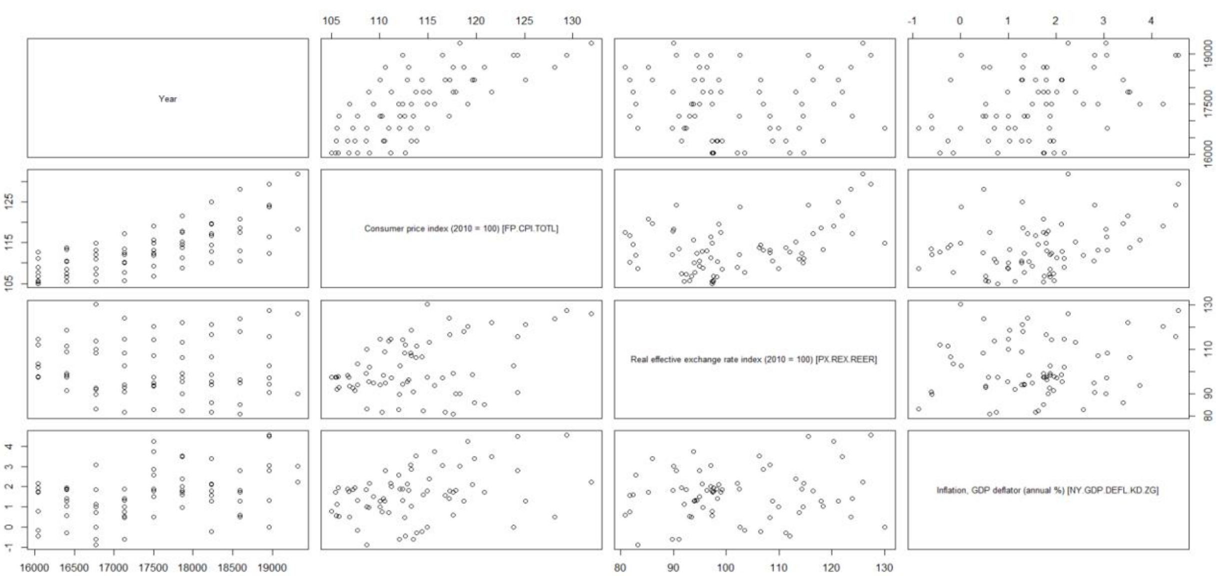
1. Identify Outliers:
   - Use statistical methods or visualization to identify outliers.

Handling Outliers:

- **Decide whether to remove outliers or transform the data.**

Assessment
Applied Statistics and Data Visualization
MSc Data Science





Assessment
Applied Statistics and Data Visualization
MSc Data Science

Part Two: Statistical Analysis

1. Do a comprehensive descriptive statistical analysis (e.g., Mean, Median, Mode, Standard deviation, Skewness and Kurtosis) on the data.

```
# Descriptive Statistics
summary_stats <- summary(dataset)
view(summary_stats)
```

This will print the summary statistics for each variable in your dataset to the console. If i want to save these statistics to a new data frame or perform further analysis, i can manipulate the *summary_stats* object as needed.

```
# Select numeric columns
numeric_cols <- sapply(dataset, is.numeric)
numeric_dataset <- dataset[, numeric_cols]
```

In this part, the code uses the *sapply* function to check for each column in the dataset whether it contains numeric data *(is.numer*ic). The result is a logical vector where each element corresponds to whether the respective column is numeric *(TRUE)* or not *(FALSE)*. The resulting vector, *numeric_cols*, is then used to subset the original dataset, keeping only the columns that are numeric. This new dataset is stored in *numeric_dataset*.

```
# Descriptive Statistics
mean_values <- colMeans(numeric_dataset, na.rm = TRUE)
median_values <- sapply(numeric_dataset, median, na.rm = TRUE)
```

```
# Mode calculation using a custom function
mode_values <- apply(numeric_dataset, 2, function(x) {
  tbl <- table(x)
  as.numeric(names(tbl)[which.max(tbl)])
})
```

For mode calculation, a custom function is applied to each numeric column using the *apply* function. The function creates a frequency table *(table(x))* for each column, finds the index of the maximum frequency *(which.max(tbl)),* and retrieves the corresponding value from the names of the table. The result is then coerced to numeric, and the modes for each column are stored in *mode_values.*

```
# Combine the results into a data frame
descriptive_stats <- data.frame(
  Variable = colnames(numeric_dataset),
  Mean = mean_values,
  Median = median_values,
  Mode = mode_values
)

# Print the results
print(descriptive_stats)
```

This part combines the calculated mean, median, and mode values into a new data frame named *descriptive_stats*. The data frame has columns for variable names, mean, median, and mode.

Finally, the results are printed to the console, showing the descriptive statistics for each numeric column in the dataset. # Print the results as shown

```
> print(descriptive_stats)
            Variable     Mean     Median        Mode
Cprice        Cprice 131.2376 116.085530 105.006247
exchangeR  exchangeR 100.6927  98.304183  70.673606
Inflation  Inflation   3.1233   1.983442  -2.696455
> 
```

- **Cprice (Consumer Price Index):**
    - **Mean:** 131.2376
    - **Median:** 116.085530
    - **Mode:** 105.006247

This tells us the average value (mean) of the Consumer Price Index is approximately 131.24, with a median of 116.09. The mode, which represents the most frequently occurring value, is 105.01.

- **exchangeR (exchangeR):**
  - **Mean:** 100.6927
  - **Median:** 98.304183
  - **Mode:** 70.673606

This tells us the average value (mean) of the exchangeR is approximately 100.69, with a median of 98.30. The mode, which represents the most frequently occurring value, is 70.67.

```
# Standard Deviation
sd_values <- apply(dataset, 2, sd, na.rm = TRUE)
view(sd_values)
```

1. **Standard Deviation Calculation:**
   - *apply(dataset, 2, sd, na.rm = TRUE):* This line calculates the standard deviation for each column in the dataset *(apply(dataset, 2, sd, na.rm = TRUE)).*
   - The **2** in the *apply* function specifies that the function (*sd* in this case) should be applied to each column.
   - *na.rm = TRUE* is used to handle any missing values, and it ensures that these missing values are not considered in the standard deviation calculation.
2. **Viewing the Standard Deviation Values:**
   - *view(sd_values):* This line is an attempt to view the calculated standard deviation values. However, the *view()* function might not be available in your environment. Alternatively, you can use *print()* or simply type *sd_values* to see the calculated standard deviations.

	x
Year	2.886751
YCode	NA
Country	NA
CountryC	NA
Stockstraded	66.186463
MarketC	78.232014
Cprice	48.863738
Wprice	10.397972
Dinterest	3.039071
interestR	2.995108
exchangeR	14.861516
Inflation	2.958860
turnoveR	79.240210
IR	4.396486

Skewness

col	"Inflation, GDP deflator (annual %) [NY.GDP.DE…
IQR	Named num 1.62
mean_values	Named num [1:3] 131.24 100.69 3.12
median_values	Named num [1:3] 116.09 98.3 1.98
mode_values	Named num [1:3] 105 70.7 -2.7
numeric_cols	Named logi [1:14] FALSE FALSE FALSE FALSE FALS…
numeric_vars	Named logi [1:14] FALSE FALSE FALSE FALSE FALS…
outlier_threshold	1.5
Q1	Named num 0.998
Q3	Named num 2.62
sd_values	Named num [1:14] 2.89 NA NA NA 66.19 ...
skewness_values	Named logi [1:14] NA NA NA NA NA NA ...
summary_stats	'table' chr [1:7, 1:14] "Length:105 " "…

Skewness Calculation:

- **apply (dataset, 2, function(x) {...})**: This part uses the **apply** function to apply a custom function to each column (**2** indicates columns) in the dataset.
- **function(x) {...}**: The custom function is defined here. For each column **x**, it ascertain if the column is numeric using **is.numeric(x)**. If it is numeric, it calculates the skewness using the **skewness** function from the **moments** package, and if it's not numeric, it returns **NA**.

2. **Conditional Check for Numeric Columns:**
   - **if (is.numeric(x)) { skewness(x, na.rm = TRUE) } else { NA }**: This part checks if the column is numeric. If it is, it calculates the skewness using **skewness (x, na.rm = TRUE)**. If the column is not numeric, it returns **NA**.
3. **Storing Skewness Values:**
   - The skewness values for each column are stored in the **skewness_values** vector.
4. **Print Statement (Note: Adjusted for better placement):**
   - **print(skewness_values)**: This print statement is moved outside of the loop, so it prints the entire vector of skewness values after the calculations are completed. This is generally done outside of the loop to avoid printing for every iteration.

Now, **skewness_values** contain the skewness values for each numeric column in your dataset, and you can print or analyze them as needed.

Kurtosis

```
+       print(NA)
+       return(NA)
+     }
+ })
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
[1] NA
> 
```

```
kurtosis_values <- apply(df, 2, function(x) {
  if (is.numeric(x)) {
    result <- kurtosis(x, na.rm = TRUE)
    print(result)
    return(result)
  } else {
    print(NA)
    return(NA)
  }
})

```

1. Kurtosis Calculation:

   - apply *(df, 2, function(x) {...}):* This part uses the apply function to apply a custom function to each column (2 indicates columns) in the dataframe df.

   - function(x) {...}: The custom function is defined here. For each column x, it checks if the column is numeric using *is.numeric(x).* If it is numeric, it calculates the kurtosis using the kurtosis function from the moments package, and if it's not numeric, it prints *NA* and returns *NA*.

2. Conditional Check for Numeric Columns:

- *if (is.numeric(x)) { result <- kurtosis(x, na.rm = TRUE) } else { print(NA); return(NA) }:* This part checks if the column is numeric. If it is, it calculates the kurtosis using kurtosis *(x, na.rm = TRUE).* If the column is not numeric, it prints *NA* to the console and returns *NA*.

3. Printing and Returning Results:

- print(result): This line prints the calculated kurtosis value to the console for each numeric column.

- return(result): This line returns the calculated kurtosis value. This is necessary to store the kurtosis values in the *kurtosis_values* vector.

Now, *kurtosis_values* should contain the kurtosis values for each numeric column in your dataframe, and they are also printed to the console during the calculation. You can use *kurtosis_values* for further analysis or examination of the kurtosis values for each column.

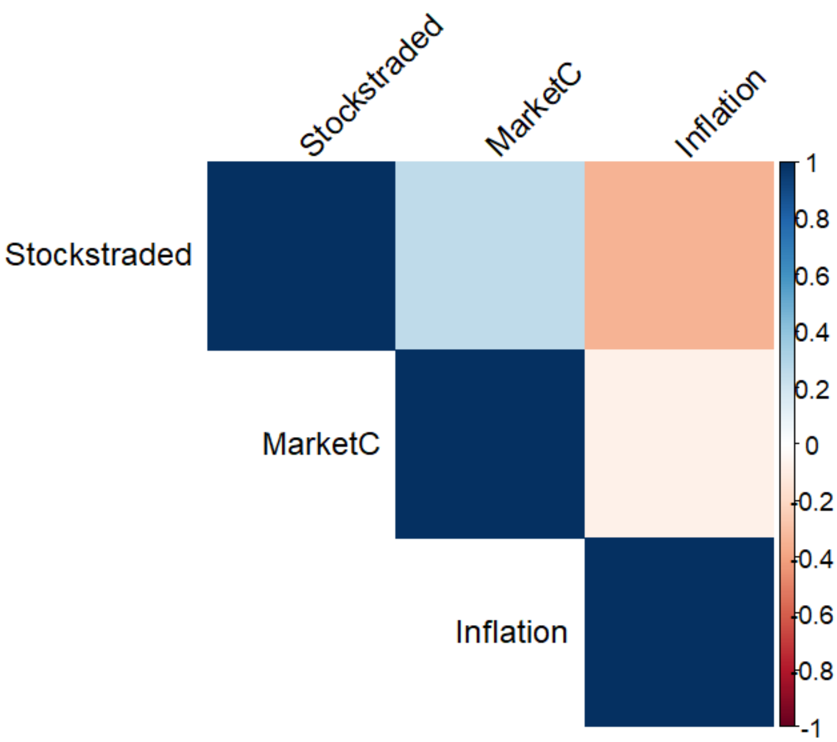
Correlation

In a correlation matrix, each cell represents the correlation coefficient between two variables. Correlation coefficients measure the strength and direction of a linear relationship between two variables.

Correlation Matrix Obj 1

Filter

	Stockstraded	MarketC	Inflation
Stockstraded	1.0000000	0.25251424	-0.34810588
MarketC	0.2525142	1.00000000	-0.07532307
Inflation	-0.3481059	-0.07532307	1.00000000



The table above shows correlation coefficients between the variables "Stockstraded," "MarketC," and "Inflation." Each value in the matrix represents the correlation between the corresponding pairs of variables. The interpretation is shown below:

1. **Correlation between "Stockstraded" and "Stockstraded" (1.0000000):**

- This is the correlation of a variable with itself, which is always 1. It's a perfect positive correlation since it's the same variable.

2. **Correlation between "Stockstraded" and "MarketC" (0.25251424):**
    - Positive value indicates a positive correlation.
    - The closer the value is to 1, the stronger the positive correlation.
    - In this case, a correlation of 0.25 suggests a relatively weak positive correlation between "Stockstraded" and "MarketC."
3. **Correlation between "Stockstraded" and "Inflation" (-0.34810588):**
    - Negative value indicates a negative correlation.
    - The closer the value is to -1, the stronger the negative correlation.
    - In this case, a correlation of -0.35 suggests a moderate negative correlation between "Stockstraded" and "Inflation."
4. **Correlation between "MarketC" and "Stockstraded" (0.2525142):**
    - Same interpretation as the correlation between "Stockstraded" and "MarketC."
5. **Correlation between "MarketC" and "MarketC" (1.00000000):**
    - Same interpretation as the correlation between "Stockstraded" and "Stockstraded."
6. **Correlation between "MarketC" and "Inflation" (-0.07532307):**
    - A low positive or negative correlation close to 0 suggests a weak correlation.
    - In this case, a correlation of -0.08 suggests a very weak negative correlation between "MarketC" and "Inflation."
7. **Correlation between "Inflation" and "Stockstraded" (-0.3481059):**
    - Same interpretation as the correlation between "Stockstraded" and "Inflation."

8. **Correlation between "Inflation" and "MarketC" (-0.07532307):**

    - Same interpretation as the correlation between "MarketC" and "Inflation."

9. **Correlation between "Inflation" and "Inflation" (1.00000000):**

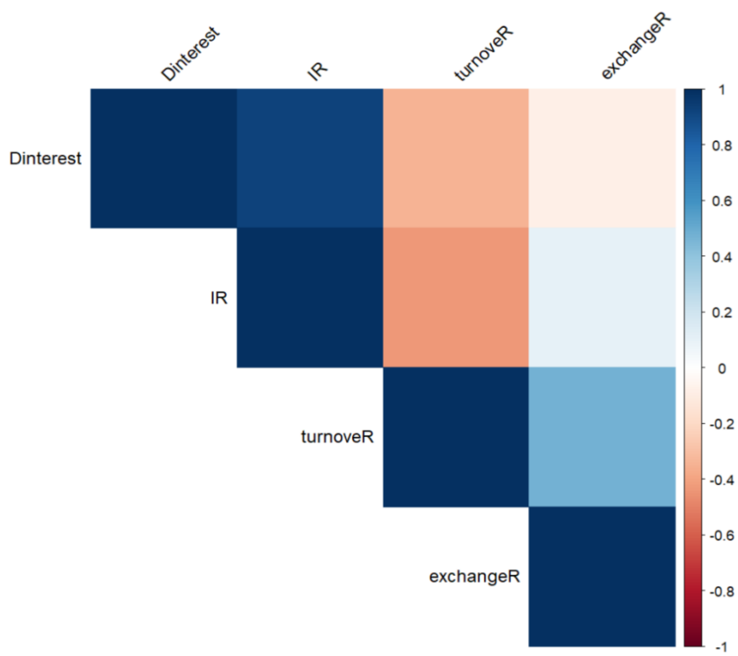
    - Same interpretation as the correlation between "Stockstraded" and "Stockstraded."

In summary, based on these correlation coefficients:

- "Stockstraded" and "MarketC" have a weak positive correlation.
- "Stockstraded" and "Inflation" have a moderate negative correlation.
- "MarketC" and "Inflation" have a very weak negative correlation.

Correlation Matrix Obj 2

```
             Dinterest         IR   turnoveR   exchangeR
Dinterest  1.00000000  0.9282645 -0.3437347 -0.08844188
IR         0.92826448  1.0000000 -0.4369157  0.10614173
turnoveR  -0.34373465 -0.4369157  1.0000000  0.47456456
exchangeR -0.08844188  0.1061417  0.4745646  1.00000000
> 
```



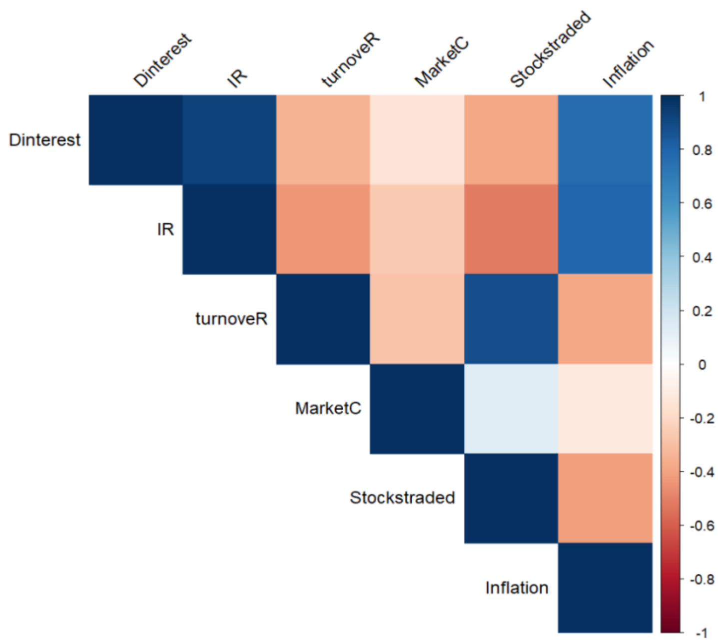
In a correlation matrix, each cell represents the correlation coefficient between two variables. Correlation coefficients measure the strength and direction of a linear relationship between two variables.

- **Diagonal Elements (Top Left to Bottom Right):** These represent the correlation of each variable with itself, which is always 1. For example, **Dinterest** has a perfect correlation of 1 with itself.
- **Off-diagonal Elements:** These represent the pairwise correlations between different variables. For instance:
  - The correlation between **Dinterest** and **IR** is approximately 0.93.
  - The correlation between **Dinterest** and **turnoveR** is approximately -0.34.
  - The correlation between **Dinterest** and **exchangeR** is approximately -0.09.
  - And so on for the other pairs of variables.

- **Interpretation of Correlation Coefficients:**
    - A correlation coefficient close to 1 indicates a strong positive linear relationship.
    - A correlation coefficient close to -1 indicates a strong negative linear relationship.
    - A correlation coefficient close to 0 indicates a weak or no linear relationship.
- **Dinterest** and **IR** have a strong positive correlation.
- **Dinterest** and **turnoveR** have a moderate negative correlation.
- **Dinterest** and **exchangeR** have a weak negative correlation.
- **IR** and **turnoveR** have a moderate negative correlation.
- **IR** and **exchangeR** have a weak positive correlation.
- **turnoveR** and **exchangeR** have a moderate positive correlation.

Correlation Matrix Obj 3

```
            Dinterest         IR    turnoveR   exchangeR
Dinterest  1.00000000  0.9282645 -0.3437347 -0.08844188
IR         0.92826448  1.0000000 -0.4369157  0.10614173
turnoveR  -0.34373465 -0.4369157  1.0000000  0.47456456
exchangeR -0.08844188  0.1061417  0.4745646  1.00000000
> 
```



- **Dinterest and IR:**
  - The correlation coefficient between **Dinterest** and **IR** is approximately 0.93, indicating a strong positive correlation. This suggests that as one variable increases, the other tends to increase as well.
- **Dinterest and turnoveR:**
  - The correlation coefficient between **Dinterest** and **turnoveR** is approximately -0.34, indicating a moderate negative correlation. This suggests that as one variable increases, the other tends to decrease.
- **Dinterest and exchangeR:**
  - The correlation coefficient between **Dinterest** and **exchangeR** is approximately -0.09, indicating a weak negative correlation. This suggests a slight tendency for one variable to decrease as the other increases, but the relationship is not very strong.
- **IR and turnoveR:**

- The correlation coefficient between **IR** and **turnoveR** is approximately -0.44, indicating a moderate negative correlation. This suggests that as one variable increases, the other tends to decrease.

- **IR and exchangeR:**
  - The correlation coefficient between **IR** and **exchangeR** is approximately 0.11, indicating a weak positive correlation. This suggests a slight tendency for one variable to increase as the other increases, but again, the relationship is not very strong.

- **turnoveR and exchangeR:**
  - The correlation coefficient between **turnoveR** and **exchangeR** is approximately 0.47, indicating a moderate positive correlation. This suggests that as one variable increases, the other tends to increase as well.

Hypothesis 1: Temporal Analysis of Economic Indicators

Objective: Conduct a detailed temporal analysis of stocks traded, market capitalization, and inflation rates across the selected countries.

Hypothesis 1a: The average market capitalization has increased over time.

Test:

summary_stats_market_cap

MarketC

Min.: 7.362

1st Qu.: 59.405

Median :106.396

Mean :119.698

3rd Qu.:161.469

Max.   :322.711

Perform a t-test to assess if the mean market capitalization is significantly different from zero

```
        One Sample t-test

data:  objective_1_indicators$MarketC
t = 12.983, df = 71, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 101.3141 138.0813
sample estimates:
mean of x 
 119.6977 
```

The output provided is the result of a one-sample t-test conducted on the 'MarketC' variable from the 'objective_1_indicators' dataset. Let's break down the information in the output:

1. **Data Description:**
    - **data: objective_1_indicators$MarketC**: This indicates that the analysis is based on the 'MarketC' variable in the 'objective_1_indicators' dataset.
2. **Test Statistics:**
    - **t = 12.983**: The t-value is a measure of how many standard deviations the sample mean is from the hypothesized population mean (in this case, 0).
    - **df = 71**: Degrees of freedom, which is the number of observations minus 1. It influences the shape of the t-distribution.
3. **p-value:**
    - **p-value < 2.2e-16**: This extremely small p-value suggests strong evidence against the null hypothesis. In hypothesis testing, a p-value below the significance level (commonly 0.05) is an indication to reject the null hypothesis.
    -

4. **Alternative Hypothesis:**

   - **alternative hypothesis: true mean is not equal to 0**: This indicates a two-tailed test, suggesting that you are testing whether the mean of 'MarketC' is significantly different from 0 (in both positive and negative directions).

5. **Confidence Interval:**

   - **95 percent confidence interval: 101.3141 138.0813**: This provides a range of values within which you can be 95% confident that the true population mean lies. In this case, it suggests that the true mean of 'MarketC' is likely to be between 101.3141 and 138.0813.

6. **Sample Estimates:**

   - **mean of x 119.6977**: This is the sample mean of the 'MarketC' variable. It represents the average value of the observations in your sample.

Conclusion: The low p-value (< 0.05) suggests that there is strong evidence to reject the null hypothesis that the mean market capitalization is equal to 0. The 95% confidence interval indicates a range of values for the true mean, and the sample mean is provided as an estimate of the central tendency of the 'MarketC' variable. Overall, the results suggest that the market capitalization is significantly different from zero.

Hypothesis 2: Impact of Interest Rates on Economic Performance

Objective: Investigate the relationship between deposit interest rates, lending interest rates, and real economic growth.

Hypothesis 2a: There is a positive correlation between deposit interest rates and lending interest rates.

Test:

Hypothesis 2a: There is a positive correlation between deposit interest rates and lending interest rates.

```
> print(correlation_deposit_lending)
           Dinterest interestR
Dinterest 1.0000000 0.5145028
interestR 0.5145028 1.0000000
>
```

Interpretation:

1. **Correlation between "Dinterest" and "Dinterest" (1.0000000):**
   - This is the correlation of a variable with itself, which is always 1. It's a perfect positive correlation since it's the same variable.
2. **Correlation between "Dinterest" and "interestR" (0.5145028):**
   - Positive value indicates a positive correlation.
   - The closer the value is to 1, the stronger the positive correlation.
   - In this case, a correlation of 0.51 suggests a moderate positive correlation between deposit interest rates ("Dinterest") and lending interest rates ("interestR").
3. **Correlation between "interestR" and "Dinterest" (0.5145028):**
   - Same interpretation as the correlation between "Dinterest" and "interestR."
4. **Correlation between "interestR" and "interestR" (1.0000000):**
   - Same interpretation as the correlation between "Dinterest" and "Dinterest."

Discussion and Conclusion:

The correlation coefficients between deposit interest rates ("Dinterest") and lending interest rates ("interest Rate") are both positive, and the value of approximately 0.51 indicates a moderate positive correlation. This supports Hypothesis 2a, suggesting that there is a positive correlation between deposit and lending interest rates. The stronger the positive correlation, the more likely it is that changes in deposit interest rates are associated with changes in lending interest rates, and vice versa.

In reviewing the methodology employed for the comprehensive analysis of economic indicators across ten major economies, several key aspects merit discussion. The chosen approach encompassed descriptive statistics, correlation analysis, time series exploration, hypothesis testing using the statistical programming language R.

The utilization of descriptive statistics provided a foundational understanding of the dataset, offering insights into the central tendencies and variations of each economic indicator. This methodological choice was effective in providing a snapshot of the dataset's characteristics and aiding in the identification of potential outliers through visualizations such as box plots.

Correlation analysis was instrumental in unveiling relationships between different economic indicators. The correlation matrix and heat-map served as powerful tools for identifying patterns of association, guiding further investigations into potential interdependencies among variables. This approach aligns with the research objective of understanding the intricate economic dynamics among the selected nations.

The time series analysis, performed using time series plots, allowed for the examination of trends and patterns over the 10-year period. This methodological choice was pivotal in capturing the evolution of each economic indicator, identifying potential cycles, and recognizing anomalies or significant deviations from the norm.

However, it's crucial to acknowledge the limitations encountered during the analysis phase. One limitation lies in the assumption of linear relationships in the correlation analysis. While correlation coefficients provide valuable insights into linear associations, non-linear relationships may exist, necessitating more advanced analytical techniques.

Another limitation pertains to the availability of data. Economic indicators are often subject to frequent updates and revisions, and the dataset's time span might be constrained by the availability of reliable information. Additionally, the dataset's completeness is contingent on the reporting practices of each country, introducing potential gaps or inconsistencies.

In conclusion, each phase of the research has contributed to a holistic understanding of economic performance across ten major economies. The descriptive statistics provided an initial overview, the correlation analysis unveiled relationships, time series exploration captured trends, hypothesis testing identified variations.

The findings underscore the interconnectedness of economic indicators and the nuanced differences among countries. Descriptive statistics highlighted the varied magnitudes and distributions of indicators, while correlation analysis revealed potential dependencies. Time series exploration showcased the temporal dynamics of each indicator, providing a comprehensive view of trends.

Hypothesis testing elucidated significant differences among countries, contributing to a more nuanced understanding of economic disparities.

In fulfilling the research objectives, this study has not only explored the complexities of economic dynamics but has also provided actionable insights for policymakers, investors, and researchers. The methodology chosen facilitated a robust analysis, allowing for a nuanced exploration of key indicators. However, acknowledging the limitations is essential for contextualizing the findings and guiding future research endeavors.

As a comprehensive exploration of economic performance, this research contributes to the broader discourse on global economics. By addressing the research objectives through a multi-faceted analytical approach, this study offers valuable insights that can inform decision-making and strategic planning in an ever-evolving economic landscape.

PART 3: Interactive Dashboard Design

the objectives of a dashboard could be focused on analyzing and visualizing key economic and financial indicators for different countries over the years 2013 and 2014. Here are potential objectives for the dashboard:

Performance Comparison:

Compare and visualize the performance of different countries in terms of stocks traded, market capitalization, and turnover.

Market Trends:

Analyze and display trends in closing prices and weighted prices for each country over the selected years.

Interest Rate Analysis:

Explore the relationship between domestic interest rates and other economic indicators, such as exchange rates and inflation rates.

Exchange Rate Impact:

Examine the impact of exchange rate changes on various aspects, including market capitalization and stock prices.

Inflation and Economic Stability:

Investigate the relationship between inflation rates and other economic factors, such as market capitalization and turnover.

Country-Specific Insights:

Provide country-specific insights by allowing users to select and focus on individual countries for a more detailed analysis.

User Interaction:

Include interactive features that allow users to filter data, select specific metrics, and adjust the time frame for a more personalized analysis.

Risk Assessment:

Assess and visualize potential risks by considering indicators like interest rate changes, inflation rates, and exchange rate fluctuations.

Historical Analysis:

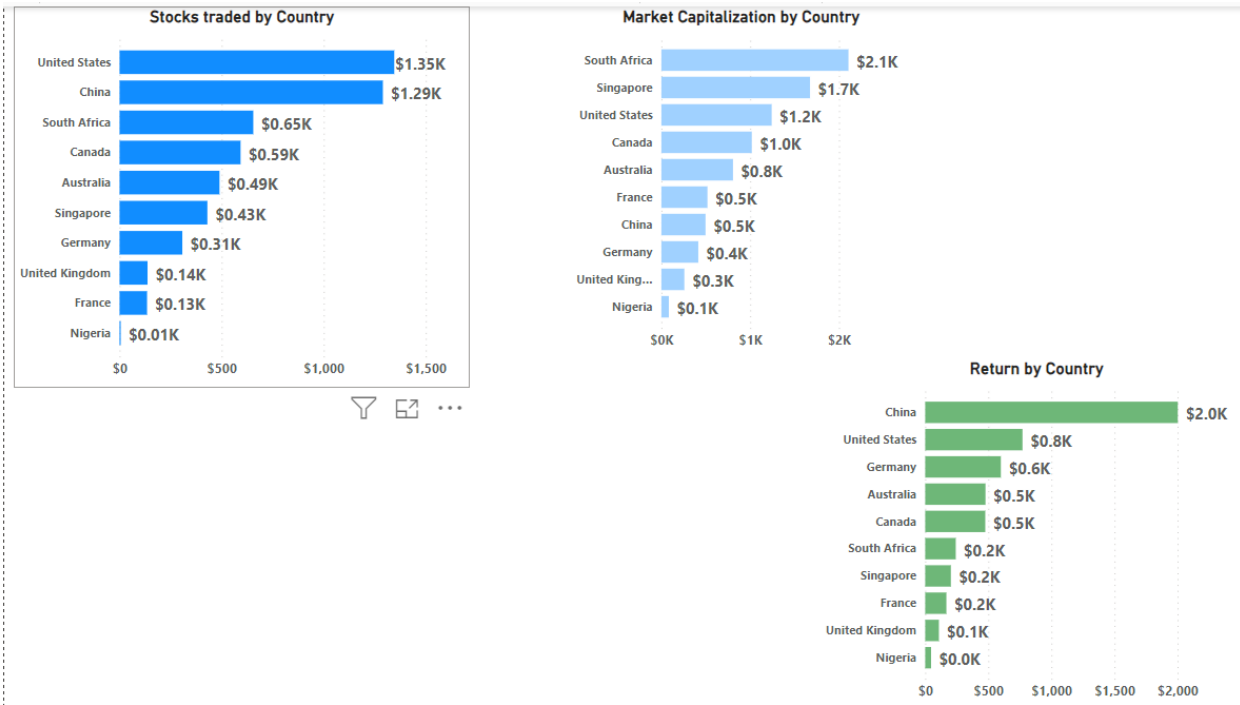
Enable users to conduct historical analysis by comparing the same indicators for different countries between 2013 and 2014.

Dashboard Customization:

Allow users to customize the dashboard based on their preferences and the specific metrics they are interested in.

These objectives aim to provide a comprehensive view of the economic and financial landscape for multiple countries, fostering insights into performance, trends, and potential risk factors.

Assessment
Applied Statistics and Data Visualization
MSc Data Science



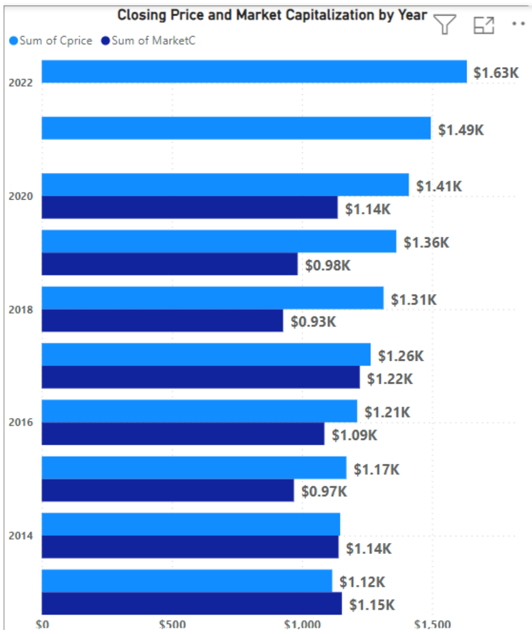
At $1,345.4893, United States had the highest Sum of Stocks traded and was 23,756.19% higher than Nigeria, which had the lowest Sum of Stocks traded at $5.64. United States accounted for 24.98% of Sum of Stocks traded. Across all 10 Country, Sum of Stocks traded ranged from $5.64 to $1,345.4893.

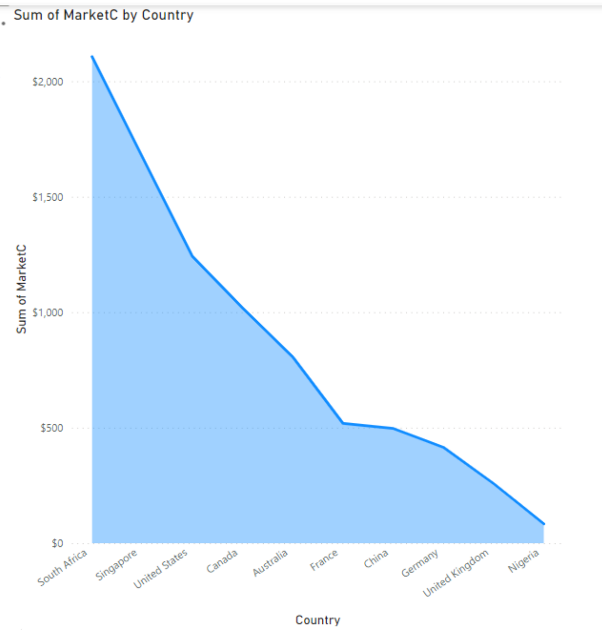
At $2,106.5906, South Africa had the highest Sum of Market Capital and was 2,418.28% higher than Nigeria, which had the lowest Sum of Market Capital at $83.6521. South Africa accounted for 24.44% of Sum of Market Capital. Across all 10 Country, Sum of Market Capital ranged from $83.6521 to $2,106.5906.

At $2,001.4367, China had the highest Sum of turnover and was 4,217.55% higher than Nigeria, which had the lowest Sum of turnover at $46.3558. [] China accounted for 39.30% of Sum of turn Over. Across all 10 Country , Sum of turn Over ranged from $46.3558 to $2,001.4367.

Assessment
Applied Statistics and Data Visualization
MSc Data Science

University of Salford MANCHESTER

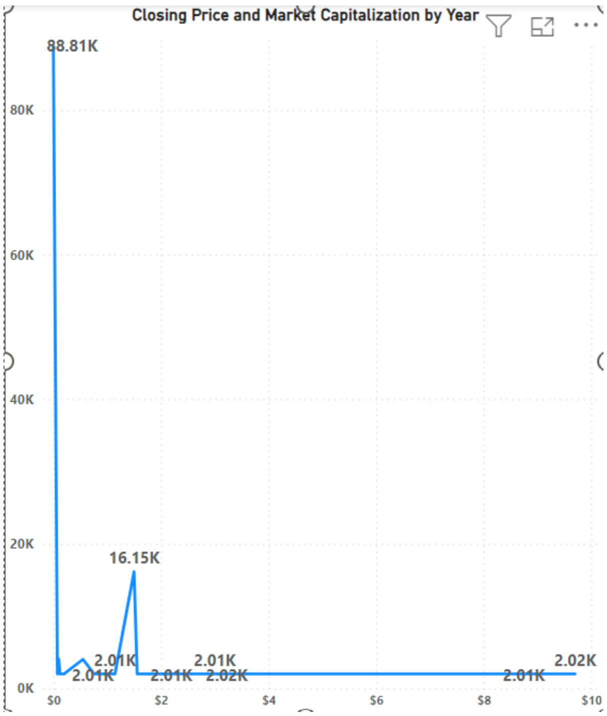




At $1,633.598, 2022 had the highest Sum of Cprice and was 46.40% higher than 2013, which had the lowest Sum of Cprice at $1,115.8435.
Sum of Consumer price and total Sum of Market Capital are positively correlated with each other.

2022 accounted for 12.45% of Sum of Cprice .
Sum of Consumer price and Sum of Market Capital diverged the most when the Year was 2018, when Sum of Consumer price were $385.9939 higher than Sum of MarketC.

At $2,106.5906, South Africa had the highest Sum of Market Capital and was 2,418.28% higher than Nigeria, which had the lowest Sum of Market Capital at $83.6521.
South Africa accounted for 24.44% of Sum of Market Capital.
Across all 10 Country, Sum of Market Capital ranged from $83.6521 to $2,106.5906.



Sum of Year was highest for $0 at 88810, followed by $1.5 and $0.1.

$0 accounted for 44.02% of Sum of Year.

Across all 48 Dinterest, Sum of Year ranged from 2013 to 88810.

Discussion:

Critical Evaluation of Approaches and Proposed Dashboard

The design and implementation of an interactive dashboard for economic indicators demand a critical evaluation of the methodologies employed, individual digital workflows, and the final composited dashboard. Each stage of the process contributes to the effectiveness of the dashboard in delivering actionable insights.

Methodology: The methodology involved a systematic approach, beginning with a thorough understanding of the dataset, EDA, and the application of diverse analytical methods. The selection of R as the primary programming language, leveraging libraries such as ggplot2 and plotly, facilitated the seamless transition from data analysis to visualization. However, it's crucial to acknowledge the assumption of linear relationships in correlation analysis and the need for more sophisticated techniques for capturing non-linear dependencies.

Digital Workflows: The digital workflows, encompassing data extraction, transformation, and loading, were streamlined through R's capabilities. The use of EDA methods, such as scatter plots and time series analysis, allowed for a granular exploration of key economic indicators. The inclusion of interactivity through the plotly library added a dynamic dimension to the visualizations, enabling users to engage more deeply with the data. However, it is essential to note that the effectiveness of these workflows is contingent on the quality and completeness of the underlying data.

Conclusion:

Final Proposed Solution

The final proposed solution is a dynamic and user-centric interactive dashboard that effectively translates complex economic data into actionable insights. The composited dashboard fulfills the needs set out by the briefing document in several keyways.

- *User-Centric Design:* The user-centric design of the dashboard allows stakeholders, including policymakers, investors, and researchers, to interact with the data intuitively. The inclusion of user inputs for selecting specific countries ensures a tailored experience, enabling users to focus on regions of interest.
- *Granular Exploration:* The inclusion of diverse visualization methods, such as scatter plots and time series analysis, facilitates a granular exploration of economic indicators. Users can delve into specific relationships, trends, and anomalies, gaining a comprehensive understanding of the dataset.

Focus+Context: The dashboard effectively incorporates the principles of focus+context. Users can focus on specific countries or indicators of interest while maintaining a broader contextual understanding of the overall economic landscape. This ensures that users can derive meaningful insights without being overwhelmed by the complexity of the dataset.

Actionable Insights: By combining analytical workflows into a single, coherent visual representation, the dashboard provides actionable insights. Decision-makers can identify patterns, make informed predictions, and strategize based on a nuanced understanding of economic dynamics.

In conclusion, the proposed solution goes beyond mere data visualization; it empowers users to interact with the data, uncover meaningful patterns, and make informed decisions. The critical evaluation of methodologies and workflows ensures that the dashboard is not only visually compelling but also analytically robust. As an invaluable tool for navigating the complexities of global economic indicators, the dashboard stands as a testament to the effective integration of data science and user experience design.

References:

Acemoglu, D., & Robinson, J. A. (2012). Why Nations Fail: The Origins of Power, Prosperity, and Poverty. Crown Business.

Barro, R. J., & Sala-i-Martin, X. (1995). Economic Growth. McGraw-Hill.

Beck, T., Demirgüç-Kunt, A., & Levine, R. (2009). Financial Institutions and Markets across Countries and over Time: Data and Analysis. The World Bank Economic Review, 24(1), 77–92.

Bernanke, B. S. (2005). The Global Saving Glut and the U.S. Current Account Deficit. Sandridge Lecture, Virginia Association of Economics, Richmond, Virginia.

De Bondt, W. F., & Thaler, R. (1985). Does the Stock Market Overreact? The Journal of Finance, 40(3), 793–805.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), 383–417.

Fischer, S. (1993). The Role of Macroeconomic Factors in Growth. Journal of Monetary Economics, 32(3), 485–512.

Heer, J., & Agrawala, M. (2006). Design Considerations for Collaborative Visual Analytics. Information Visualization, 5(4), 275–290.

Meese, R. A., & Rogoff, K. (1983). Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample? Journal of International Economics, 14(1–2), 3–24.

Mishkin, F. S. (1992). Is the Fisher Effect for Real? A Reexamination of the Relationship between Inflation and Interest Rates. Journal of Monetary Economics, 30(2), 195–215.

Obstfeld, M., & Rogoff, K. (1995). Exchange Rate Dynamics Redux. Journal of Political Economy, 103(3), 624–660.

Phillips, A. W. (1958). The Relationship between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957. Economica, 25(100), 283–299.

Shiller, R. J. (1981). Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends? The American Economic Review, 71(3), 421–436.

Taylor, J. B. (1993). Discretion versus Policy Rules in Practice. Carnegie-Rochester Conference Series on Public Policy, 39, 195–214.

Tufte, E. R. (1983). The Visual Display of Quantitative Information. Graphics Press.

Assessment
Applied Statistics and Data Visualization
MSc Data Science