**SECTION 1 – INTRODUCTION**

**Background:**

Market Capitalisation refers to the total dollar value of all of the outstanding shares of a company. Market researchers use this for two reasons: to find out how much a company is worth and to determine the value of a stock. The following formula is used to calculate the market capitalisation value:

**Market Capitalisation = Outstanding Shares X Stock Price**

In the analysis, the Top 10 companies refer to the 10 companies with the highest market capitalisation values, whereas the Top 500 companies are the 500 companies with the highest market capitalisation values.

**Objectives of this analysis:**

The main objectives of this analysis were to find out the following information:

* The market capitalisations of the Top 10 companies in the world.
* The total percentage of the number of shares for each market capitalisation category.
* The number of companies, in the Top 500, that belong to different stock price ranges.
* The number of companies that belong to each of the G20 countries.

**SECTION 2 – DATASET DETAILS**

**Data source:**

The dataset is called “Top Global Companies by Market Capitalisation”. It was originally created on 29th January 2024 on [www.companiesmarketcap.com](http://www.companiesmarketcap.com) and was made available on Kaggle. The dataset was downloaded from Kaggle into Microsoft Excel for data transformation, analysis and visualisation on 22nd March 2024.

**Variables:**

The dataset contains the following variables:

* Rank (i.e. company ranking)
* Company name
* Company code
* Market capitalisation (in $s)
* Stock price (in $s)
* Origin Flag
* Country (in which company was originally founded in).

**Type of analysis:**

Quantitative analysis.

**Dataset chosen:**

Prior to being uploaded onto Kaggle, the dataset comes from a reputable, reliable and up-to-date source. The dataset contains information for approximately 8,000 companies worldwide.

**Tools:**

Microsoft Excel was used for data transformation, analysis and visualisation.

**SECTION 3 – DATA TRANSFORMATION**

Here is a high-level overview of the data transformation steps taken:

## Added

- 1 column for classifying companies by their Market Capitalisation category.

- Highlighting "Market Capitalisation" figures according to their categories.

- 1 column for showing the number of shares each company has.

## Changed

- "Marketcap" column renamed to "Market Capitalisation".

- "Country Flag" column renamed to "Country Abbreviation".

- "Market Capitalisation" column values converted to currency values with 2 decimal places.

- Rounding all "Stock prices" to 2 decimal places.

- Abbreviated/short-named countries renamed to their respective full names.

## Removed

- "Blanks" from "Country Name".

- "Company Codes" column (not useful for analysis).

- "NA" values from "Market Capitalisation" column.

- "$0 M" values from "Market Capitalisation" column.

- Duplicate company names.

- "NA" values from "Stock Prices".

- Stock price values that are less than $0.01.

- "Origin Flag" column (not useful for analysis).

## Fixed

- Entered "Germany" as country for the company called "Ceconomy".

On the next page, there is an in-depth description of each step of the data cleaning process. This includes assessing data accuracy, completeness and consistency:

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Cleaning Step** | **Description** | **Method** | **Result** |
| **Data Transformation Steps** | | | |
| 1 | Original dataset copied for cleaning | Copy and paste | Dataset pasted in a new Excel worksheet |
| 2 | Converting the dataset to a table |  |  |
| 3 | "Marketcap" renamed to "Market Capitalisation" | Manual renaming | 1 column renamed |
| 4 | "Country Flag" renamed to "Country Abbreviation" | Manual renaming | 1 column renamed |
| 5 | Removing "Blanks" from "Country Name" | Filtering and deleting rows with "Blank" values | 1 row deleted |
| 6 | Removed "Company Codes" as they were not useful for analysis | Manual deletion | 1 column deleted |
| 7 | "NA" values removed from "Market Capitalisation" | Filtering and deleting rows with "NA" values | 7 rows deleted |
| 8 | "$0 M" values removed from "Market Capitalisation" | Filtering and deleting rows with "$0 M" values | 11 rows deleted |
| 9 | "Market Capitalisation" column values converted to currency values with 2 decimal places | Text to Column to separate letters from values |  |
|  | (instead of having "M", "B" or "T") | Text to Column to separate $ sign from amounts |  |
|  |  | Multiplying "M" amounts by 10^6 using "=IFS()" |  |
|  |  | Multiplying "B" amounts by 10^9 using "=IFS()" |  |
|  |  | Multiplying "T" amounts by 10^12 using "=IFS()" |  |
|  |  | Converting all values to $ currency with 2 decimal places |  |
|  |  | Replacing all values in "Market Capitalisation" column with new values | Values in "Market Capitalisation" are full, |
|  |  |  | continuous values with $ sign |
| 10 | Deleted column with "M", "B" and "T" letters | Manual deletion | 1 column removed |
| 11 | Creating column for classifying companies by their Market Capitalisation category | Used Nested "=IFS()" for categorisation (on top of AND() for multiple conditions) | 1 new column created |
| 12 | Removed duplicate company names | Used "Remove Duplicates" tool to remove duplicate company names | 3 duplicate records removed |
| 13 | Highlighting "Market Capitalisation" figures according to their categories | Used "Conditional Formatting" on "Market Capitalisation" column | Fields in "Market Capitalisation" column |
|  |  |  | are highlighted in different colours |
| 14 | Removing "NA" values from "Stock Prices" | Filtering and deleting rows with "NA" values | 2 rows deleted |
| 15 | Removing stock price values that are less than $0.01 | Sorted "Stock Price" in ascending order. |  |
|  |  | Used "Conditional Formatting" to highlight prices less than $0.01. |  |
|  |  | Deleted such values manually. | 4 rows deleted |
| 16 | Rounding all "Stock prices" to 2 decimal places | Used "Find and Replace" to remove $ sign from each value |  |
|  |  | Created a separate column to round each value using "=ROUND()" |  |
|  |  | Renamed new column as "Stock Price" |  |
|  |  | Deleted original column to avoid duplicate columns | 1 new column created with prices to 2 decimal places |
| 17 | Deleted "Origin Flag" column | Manual deletion | 1 column deleted |
| 18 | Filled out "Germany" as country for the company called "Ceconomy" | Used "Filter" to see blank values |  |
|  |  | Carried out online search to find out that company is based in Germany | 1 blank field filled in "Country column" |
| 19 | Ensuring all countries have full names | Used "Filter" to find out countries that do not have full names |  |
|  |  | Used "Find and Replace" to change S. Arabia to "Saudi Arabia" |  |
|  |  | Used "Find and Replace" to change S. Korea to "South Korea" |  |
|  |  | Used "Find and Replace" to change USA to United States of America |  |
|  |  | Used "Find and Replace" to change UAE to United Arab Emirates | All countries now listed with their respective full names |
| 20 | Calculating number of shares each company has | Separate column created |  |
|  |  | Used the following formula: "Number of shares = Market Capitalisation / Stock Price" | 1 new column created |
| **Necessary Data Checks** | | | |
| 21 | Assessing Data Completeness | Using "Filter" to find out if any "NA" or "Blanks" still exist | Dataset is complete |
| 22 | Evaluating Data Accuracy | Compared dataset to most recently available dataset on |  |
|  |  | [www.companiesmarketcap.com](http://www.companiesmarketcap.com/) | Values closely match those that are in the latest dataset |
| 23 | Validating Data Consistency | Checked if all rankings and number of shares are whole numbers |  |
|  |  | Checked if all company names, market capitalisation categories and country names are in string format |  |
|  |  | Checked if all market capitalisation values and stock prices are in $ currency | Data is fully consistent |

**SECTION 4 – RESULTS AND DISCUSSION**

**Results:**

After cleaning the dataset, there remained 7,920 records in the dataset. The data cleaning process only removed 28 records. On the overall, there are over 3,400 mid cap companies, and between 1,400-1,600 of each of the other types of companies.

|  |  |
| --- | --- |
| **Number of each company type** | |
| Number of Large Cap companies | 1571 |
| Number of Mid Cap companies | 3421 |
| Number of Small Cap companies | 1432 |
| Number of Micro Cap companies | 1496 |

The market capitalisation categories are defined as follows:

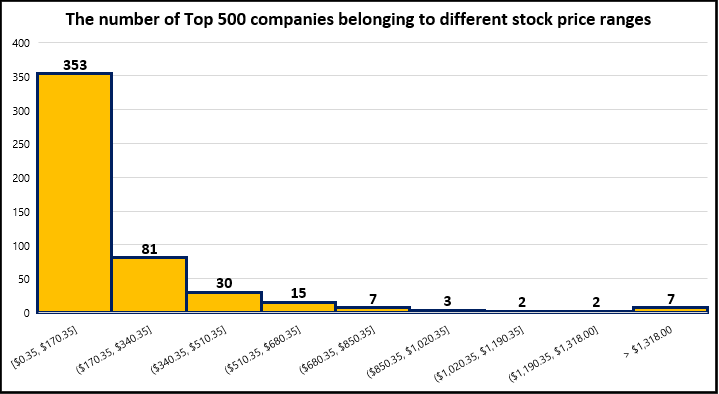
* Large Cap companies are those worth greater than $10 trillion.
* Mid Cap companies are those worth between $1 billion and less than $10 trillion.
* Small Cap companies are those worth between $250 million and less than $1 billion.
* Micro Cap companies are those worth less than $250 million.

The following table shows the company worth (i.e. market capitalisations) and countries of a few well-known global financial organisations:

|  |  |  |
| --- | --- | --- |
|  | **Company Worth** | **Country** |
| Santander | $61,090,000,000.00 | Spain |
| BNP Paribas | $76,880,000,000.00 | France |
| Barclays | $28,710,000,000.00 | United Kingdom |
| Mastercard | $410,650,000,000.00 | United States of America |
| Deutsche Bank | $26,030,000,000.00 | Germany |
| Bank of China | $157,800,000,000.00 | China |

The bar chart on the next page summarises the market capitalisations of the Top 10 companies. As it can be seen, Microsoft is worth the most at over $3 trillion, followed by Apple and Saudi Arabia Aramco (which are both worth between $2 trillion and $3 trillion).

The pie chart shows the total percentage of the number of shares for each market capitalisation category. Large Cap and Mid Cap companies make up most of the market, due to values of 53.49% and 41.33% respectively.

In terms of the stock prices of the Top 500 companies, the following histogram illustrates that over 70% of these companies have stock prices ranging between $0.35 and $170.35. Only 7 of these companies have stock prices above $1,318.00.

The final part of the analysis looked at the number of companies in each of the G20 countries. Here is a pivot table that contains the counts of the number of companies that belong to each of these countries. A slicer was used by only selecting the G20 countries. South Africa, one of the G20 countries, is not listed here as no record exists in the dataset for it.

|  |  |
| --- | --- |
| **Country Names** | **Number of Companies** |
| Russia | 17 |
| Argentina | 19 |
| Mexico | 30 |
| Saudi Arabia | 34 |
| Turkey | 41 |
| Indonesia | 46 |
| Spain | 59 |
| Brazil | 80 |
| Italy | 84 |
| France | 118 |
| Australia | 130 |
| South Korea | 150 |
| Germany | 218 |
| United Kingdom | 257 |
| China | 312 |
| Japan | 346 |
| Canada | 389 |
| India | 563 |
| United States of America | 3,604 |
| **Grand Total** | **6,497** |

From the pivot table, the following bar chart was created. It demonstrates that the United States of America has over 3,600 companies in the dataset. India and Canada come in second and third places respectively. In fact, the United States of America has got six to seven times more companies than India, and almost 10 times more companies than Canada. The bottom three are Russia, Argentina and Mexico.

**Discussion:**

There are certain limitations to the analysis. 11 records were removed from the dataset because they contained “$0M” as a value for market capitalisation. Two records were removed as a result of not knowing their stock prices. A more appropriate approach would have been to search the values of these for 29th January 2024. In this way, the analysis could have been slightly better than it is now.

As well as this, the data is only from approximately 8,000 companies that have some of their market capitalisation information publicly available. There is a possibility that other companies may have this information, but it is not publicly available for data analysis.

Lastly, the data is from the end of January 2024. From that time onwards, the data might have changed. It is recommended to use the latest data on this topic for a more accurate data analysis.

**SECTION 5 – SUMMARY**

The data analysis has revealed the following facts:

* Microsoft tops the list of market capitalisations values, with a worth of over $3 trillion, followed by Apple and Saudi Arabia Aramco (who are worth between $2 trillion and $3 trillion).
* “Large cap” and “Mid cap” companies have almost 95% of the number of shares in total.
* Over 70% of Top 500 companies have a stock price range between $0.35 and $170.35.
* Amongst the G20 countries, The United States of America has over 3,600 companies (which is, by far, the highest).