Applied Data Science 1- Assignment: Statistics and Trends

Student Name: Jafar Urchintala
Student ID: 23036717
GitHub Repository URL: https://github.com/Jafarurchintala/Assignment-1-ADS_Statisitics-and-Trends.git

Introduction:

In this report, I explore relationships within a financial dataset from Kaggle (https://www.kaggle.com/datasets/paytonfisher/sp-500-companies-with-financial-information/) using Python. The dataset contains information about various companies, including their stock prices, earnings, and other financial metrics. My aim is to uncover any correlations or patterns among these variables to gain insights into the financial performance of these companies.

Dataset Overview:

The dataset consists of several columns, including the company name, stock symbol, stock price, earnings, and other financial metrics. Initially, I read the dataset into a Pandas DataFrame and performed basic data exploration to understand its structure and identify any missing values.

Analysis using Bar Graph:

- **Price/Earnings of Top Companies:** A bar chart was plotted to compare the price/earnings ratio of the top 10 companies. This ratio is a key indicator of a company's valuation relative to its earnings and is commonly used by investors to assess investment opportunities.

The bar chart of price/earnings ratios highlights differences in valuation among the top companies. Lower ratios may indicate undervalued stocks, while higher ratios may suggest overvaluation.

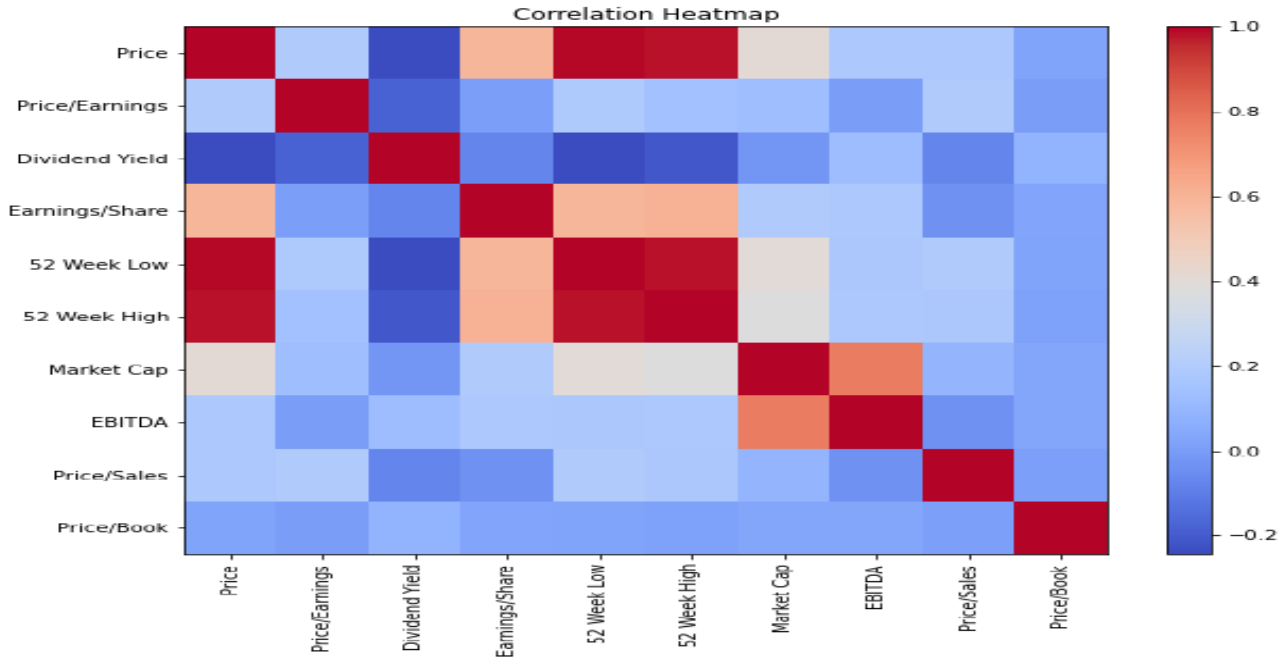
Analysis using Line Graph:

- **Highest and Lowest Prices in the Last 52 Weeks:** I visualize the highest and lowest prices in the last 52 weeks for the top 20 companies in the dataset. This line chart provides insights into the price volatility of these companies over time.

The line chart illustrates the fluctuation in stock prices for the top 20 companies over the past year. Companies with higher volatility may present higher risk but also higher potential returns.

Analysis using Heatmap:

- **Correlation Heatmap:** To explore relationships between different financial metrics, I created a correlation heatmap. This heatmap visualizes the correlation coefficients between various numeric variables in the dataset. Strong correlations (positive or negative) indicate potential relationships between these variables.



The correlation heatmap reveals relationships between different financial metrics. For example, we may observe a positive correlation between certain financial indicators, indicating that they tend to move together.

Conclusion:

In conclusion, my analysis provides valuable insights into the financial performance and relationships within the dataset. Further analysis could involve exploring causal relationships and conducting predictive modelling to forecast future financial metrics.