

S&P 500 Stocks

Udacity EDA course

Overview of Companies under S&P 500

Investigation Overview

In this investigation, I looking to Explore the returns, components and volatility of the S&P 500 index. Also identify high and low performance stocks among the list.

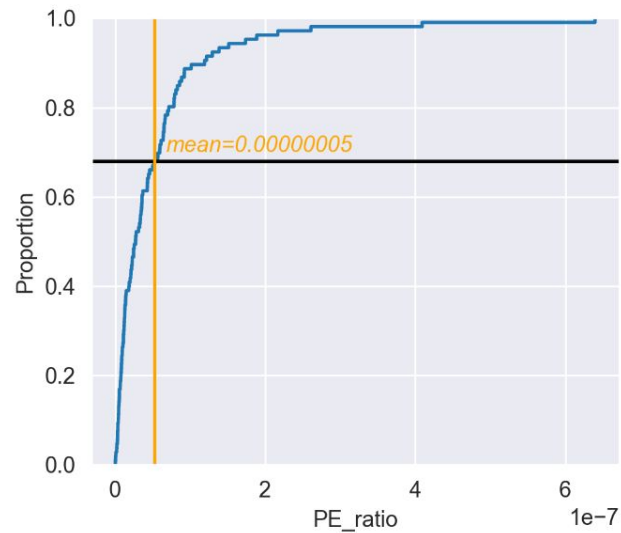
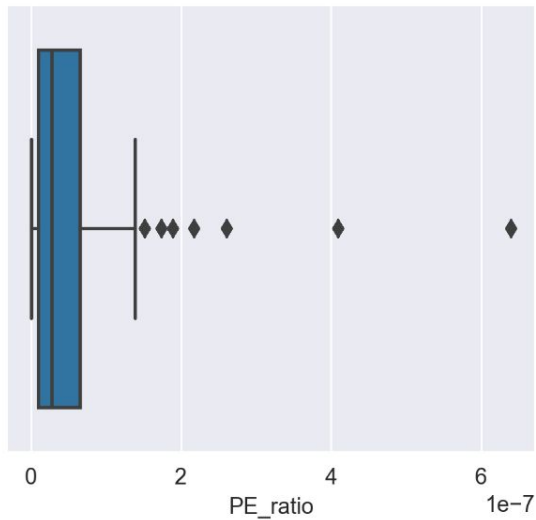
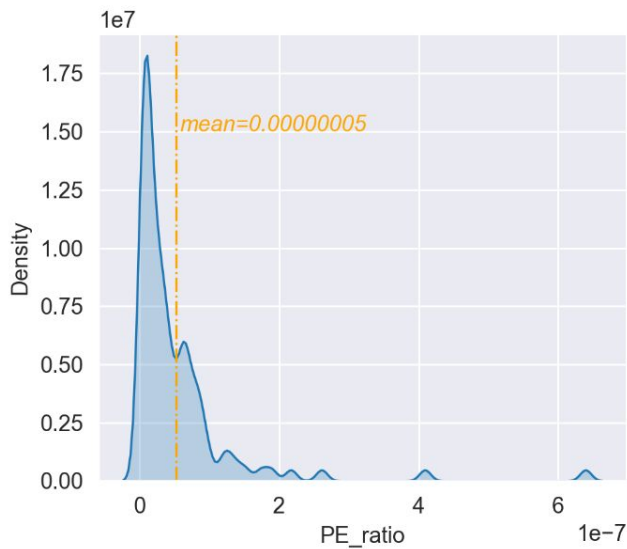
Dataset Overview

The Standard and Poor's 500 or S&P 500 is the most famous financial benchmark in the world.

This stock market index tracks the performance of 500 large companies listed on stock exchanges in the United States. As of December 31, 2020, more than \$5.4 trillion was invested in assets tied to the performance of this index.

Price-to-earnings ratio distribution

PE_ratio Distribution

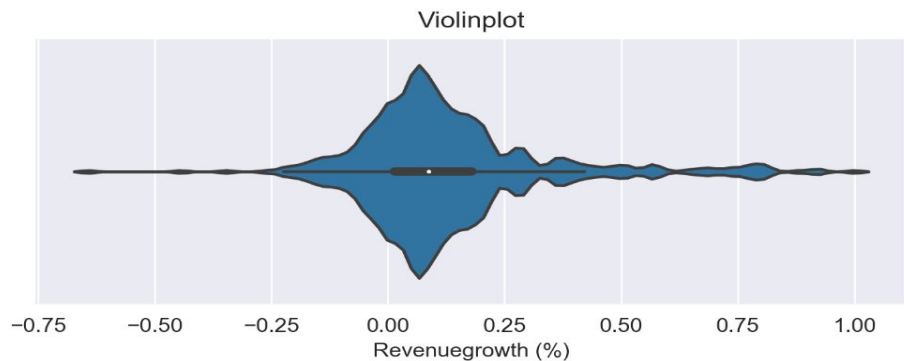
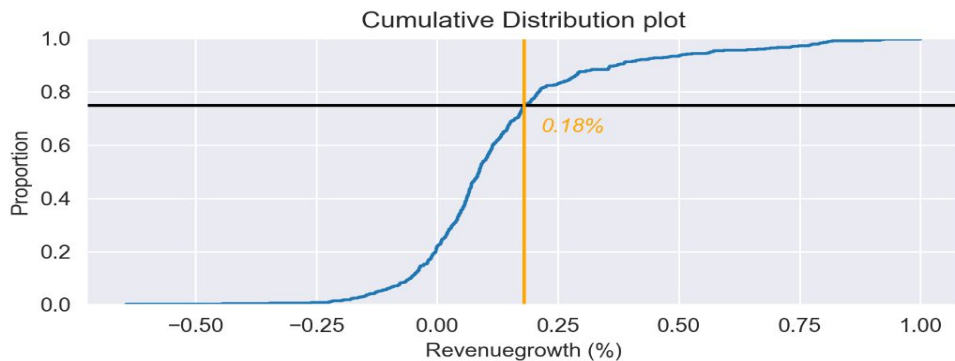
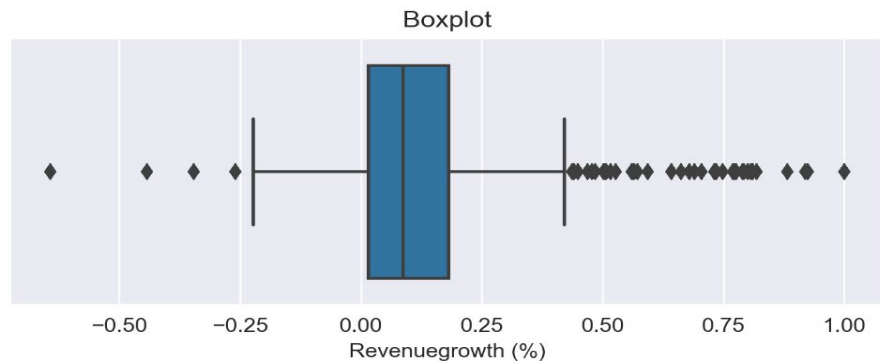
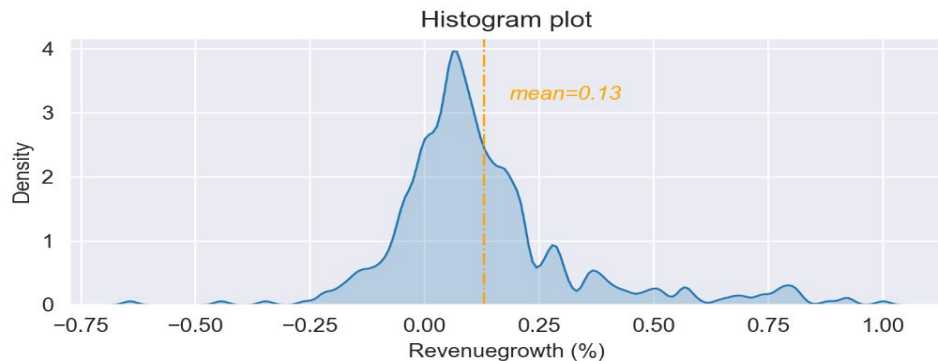


Price-to-earnings ratio distribution

- From the histogram and the box plot we could see that the PE_ratio is right skewed, with many outliers, we could see that the range is from 0 to 0.00000009 (after I removed the outliers above 0.00000009), and the mean value is 0.0000005
- From the ECDF price plot we could see that 68% from the companies are below the mean
- Stocks with a low P/E ratio are generally considered to be undervalued, while those with a high P/E ratio are considered to be overvalued. Therefore, all companies with PE_ratios higher than the mean are considered overvalued

Revenue growth distribution

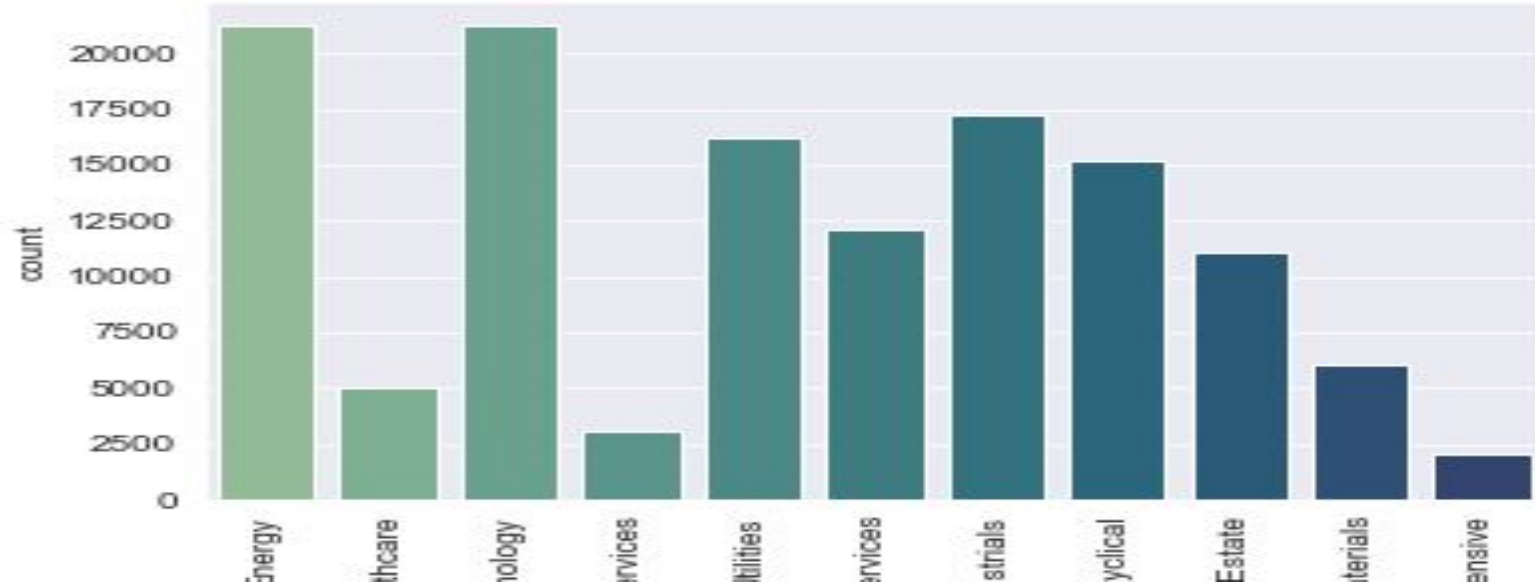
Revenuegrowth Distribution



Revenue growth distribution

- From the histogram and the box plot we could see that the Revenue growth is right skewed, with many outliers, we could see that the range is from -0.65 to 1 (after I removed the outliers above 1), and the mean value is 0.13%
- From the ECDF price plot we could see that 75% of the stocks are below 0.18% Revenue growth
- In this period, companies with revenue growth greater than 0.18% showed better performance than others

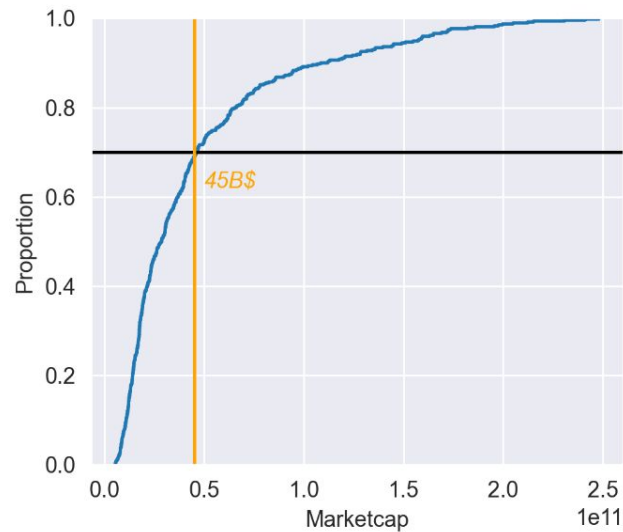
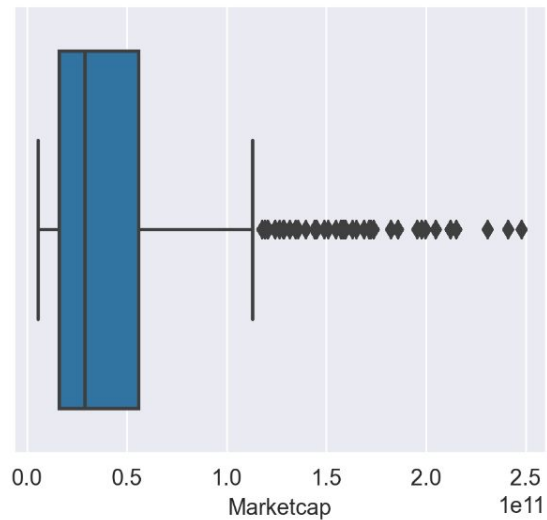
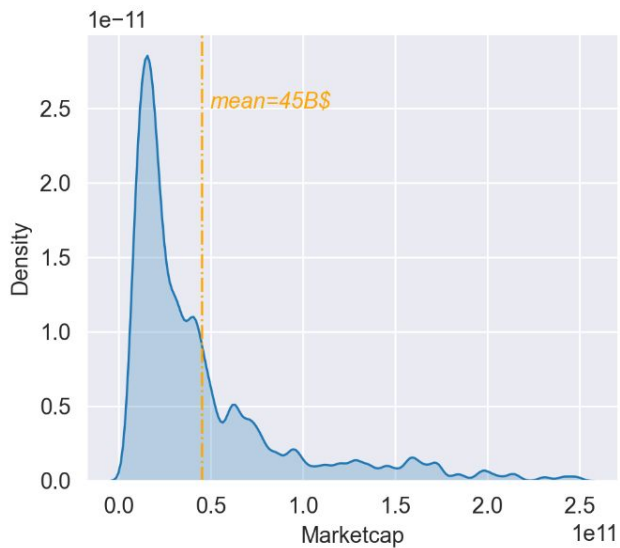
Revenue growth by sector



Technology and energy are the two sectors with the highest Revenue Growth

Marketcap distribution

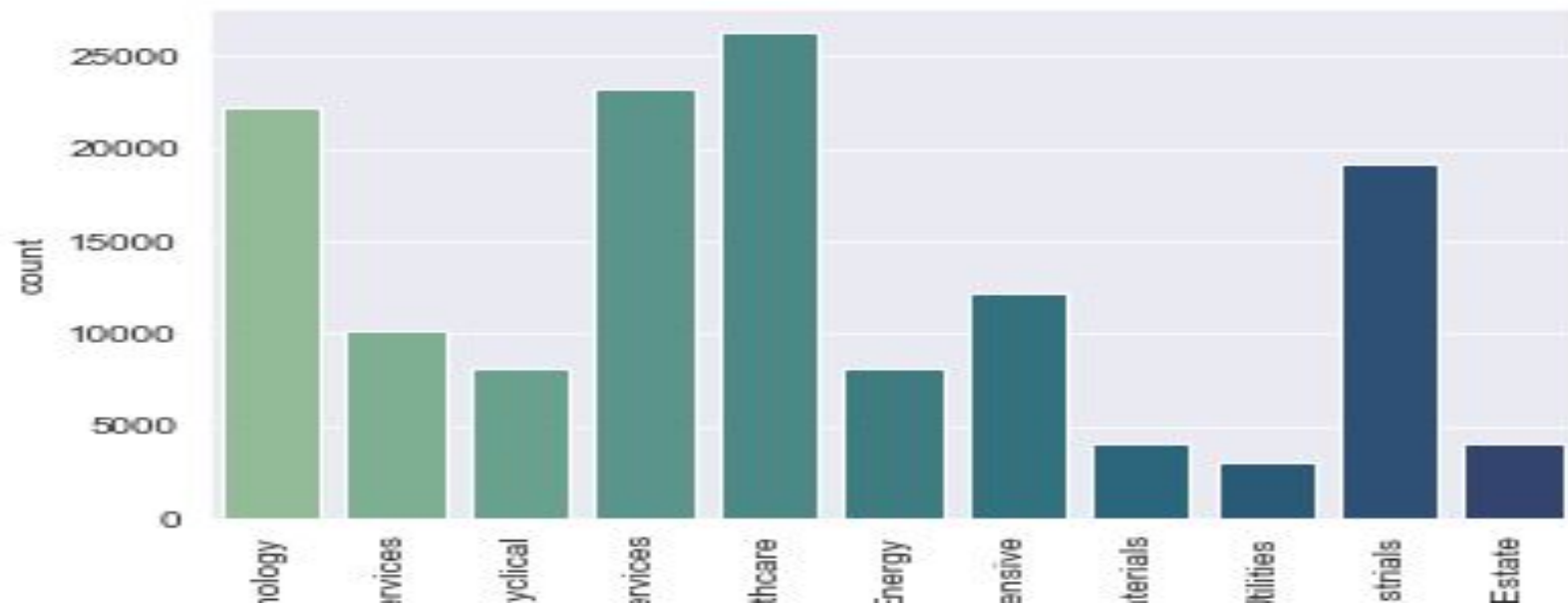
Marketcap Distribution



Marketcap distribution

- From the histogram and the box plot we could see that the Marketcap is right skewed, with many outliers, we could see that the range is from 0 to 250B (after I removed the outliers above 250B), and the mean value is around 45B
- From the ECDF price plot we could see that 70% from the companies are below the 45B Marketcap
- 75% of the companies are up to 55B Marketcap, and the mean is around 45B. That means all the companies with larger marketcap than 55B are considered a large-cap stock

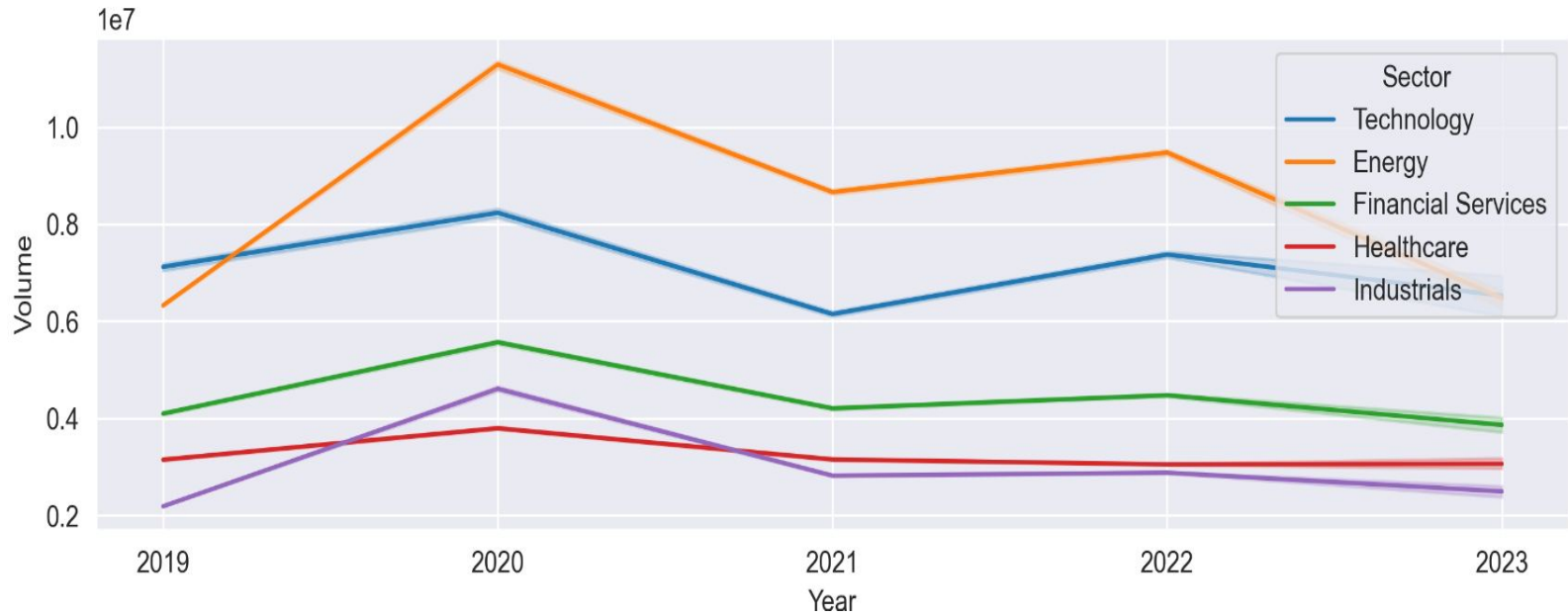
Marketcap by sector



Marketcap by sector/industry

- Technology, Financial Services, Healthcare and Industrials are the sectors with the highest Marketcap
- Consumer Electronics, Software—Infrastructure and Insurance—Diversified are the industries with the highest Marketcap

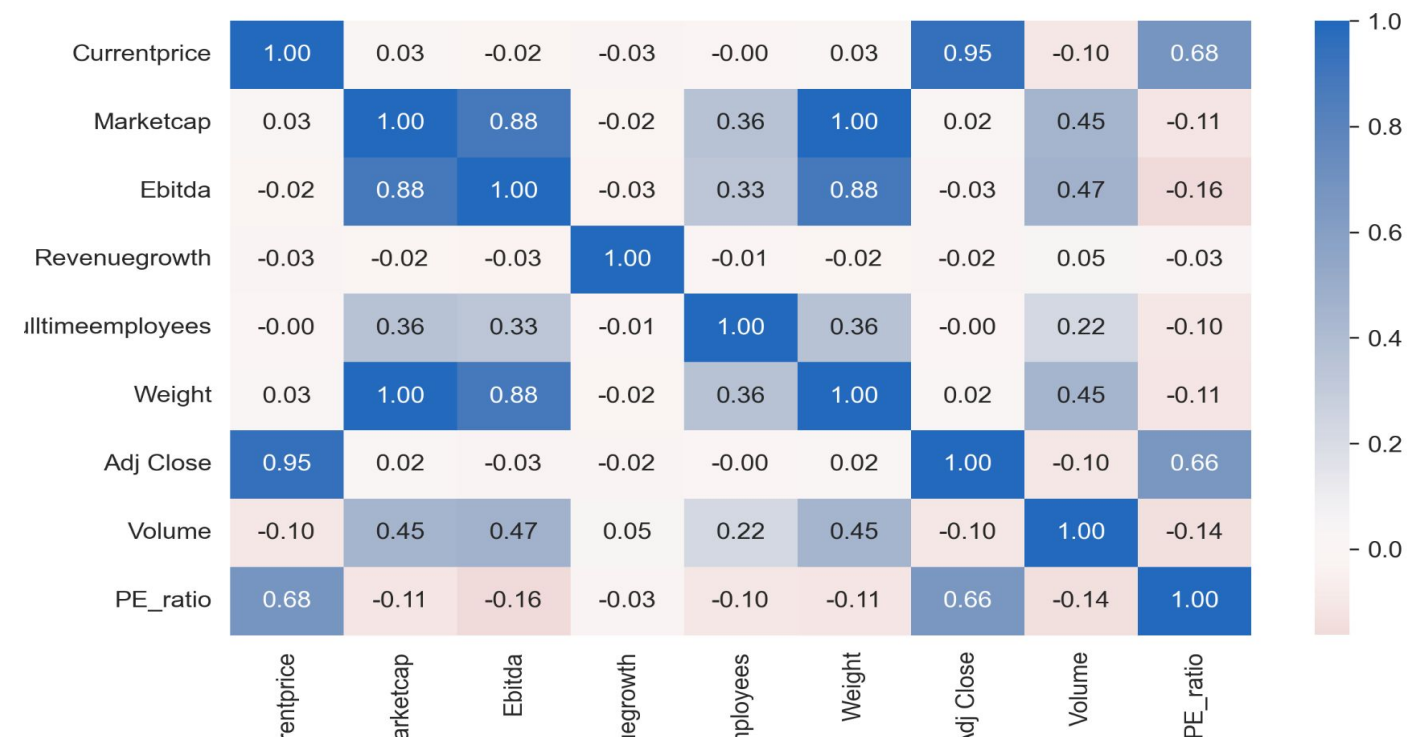
Top sectors Volume by year



Top sectors volume by year

- In 2019, all sectors experienced an increase, with Energy the most
- In 2020, all sectors experienced an decrease (This year the covid-19 pandemic began. It may be the cause, but I don't have any evidence for this in the dataset)
- There will be a good recovery in the Technology and Energy sector in 2021
- All sectors are experiencing sharp declines in 2022, but the energy sector experiences the sharpest decline

Correlation Heatmap



Correlation Heatmap

- A high correlation between EBITDA and market capitalization could mean that the market values the company based on its earnings potential
- A high correlation between EBITDA and weight (percentage of participation on the S&P index) could mean that companies with higher EBITDA are more likely to be included in the S&P index or have a higher weight in the index
- A perfect correlation between market capitalization and weight would mean that the S&P index weight is based solely on market capitalization, and as market capitalization of a company increases, so does its weight in the index
- A high correlation between the adjusted close price and current price of a stock could mean that the current price of the stock is closely related to its historical adjusted close price
- A correlation of 0.68 between the PE ratio and current price of a stock means that as the PE ratio increases, the current stock price also tends to increase, and vice versa

Conclusions

- All companies with PE_ratio higher than 0.0000005 are considered overvalued
- The mean revenue growth is 0.13%, 75% of the stocks are below 0.18%, and stocks above 0.18% showed better performance than others
- Technology and energy are the two sectors with the highest Revenue Growth
- 75% of the companies are up to 55B Marketcap, and the mean is around 45B. That means all the companies with larger marketcap than 55B are considered a large-cap stock
- Technology, Financial Services, Healthcare, and Industrials are the sectors with the highest Marketcap
- Consumer Electronics, Software—Infrastructure, and Insurance—Diversified are the Industries with the highest Marketcap
- In 2019, all sectors experienced an increase, with Energy the most
- In 2020, all sectors experienced a decrease (This year the covid-19 pandemic began. It may be the cause, but I don't have any evidence for this in the dataset)
- All sectors are experiencing sharp declines in 2022, but the energy sector experiences the sharpest decline