Group 1 Project 1 Proposal

Analysis of S&P500 Performance Over a Five-Year Period (2013-2018): Investing in the Index Compared with Volume-based Quintiles within the Index

Group 1 has the following members: Nicolas Delatour, Tiago Rocha, and Baruch Frost

The goal of the project is to analyze the performance of a theoretical investment in the S&P500 from 2013-2018 to determine if trading volumes correlated to outsized risk and/or returns. The methodology will include categorizing the companies that comprise the S&P500 during the five-year period in question by trading volume into five quintiles: the 20% with the highest trading volume in the first year of the study, the 20% with the second highest trading volume in the first year of the study, the 20% with the third highest trading volume, the 20% with the second lowest trading volume in the first year of the study, and the 20% with the lowest trading volume in the first year of the study. Then, we’ll follow the returns and drawdowns of these quartiles throughout the following four years, comparing the risk/return profile with the index overall as a baseline.

We will need data on the overall S&P500 index during the study period (2013-2018), as well as data on the individual companies that comprised the S&P500 during that period (the stock price fluctuations and the trading volumes during the first year of the study). We can use the following data from Kaggle as a starting point: <https://www.kaggle.com/datasets/camnugent/sandp500>.

We will explore the following research questions:

1. How did the stocks in each quartile perform over the 5-year period? Did any of the quartiles consistently outperform the others, or did the performance vary from year to year?
2. What were the risks associated with investing in each of the quartiles? Did any quartile have a higher level of risk than the others?

The rough breakdown of tasks is the following:

1. Acquire the data, clean the data, and bring the data into pandas.
2. Categorize the data into the appropriate quartiles based on trading volumes
3. Compare the performance of each quartile with the index as a whole with graphs to highlight the similarities/differences in risk and overall returns.
4. Produce a written analysis based on the findings of the analysis