Executive Summary

For our project, we plan to utilize the skills that we have developed over the previous eleven weeks to create a machine learning model that predicts whether a stock will outperform the S&P 500 on a given day. We plan to use approximately fourteen years of daily stock price data for six companies related to the NAICS Finance and Insurance industry. Our companies consist of two large, two medium, and two small companies based on market capitalization: J.P. Morgan, Goldman Sachs, Discover Financial Services, Synovus, Bank of Hawaii, and New York Community Bank. After accumulating and cleaning our data, we plan to run a logistic regression using the daily stock price data (High, Low, Close, Open, Volume, and percent change of stock price from previous day) to determine whether a stock's percent increase is higher or lower than the percent increase in the value of the S&P 500 index.