

Capstone 3 - Proposal

By: Richard Broyles

Problem Statement and Background

People who have a vested interest in the New York Stock Exchange (NYSE) have also wanted a way to try to predict the value of a potential stock in their portfolio over time. These same people also wanted a way to invest in a company or stock with a minimal monetary investment and stocks who have a risk value of minimal to low. This project will accomplish both of these goals.

The main objective of this project will be to predict the value of certain penny stocks that will give the maximum return for their investment. A penny stock is defined by the Securities and Exchange Commission (SEC) as a stock that is trading below \$5.00 per share. There are several sectors of the stock market that have these penny stocks such as technology, energy, biotechnology, blockchain, marijuana (recently introduced with changes in marijuana laws in individual states), gold, mining, and psychedelic stocks.

The data for this project will consist of five stocks chosen from different sectors whose data will be obtained through Yahoo! Finance using a Python library that will collect the historical performance of these stocks over the past 10 years. Data cleaning will not be necessary since these records are complete as stipulated by law. EDA will consist of time series analysis of each stock as well as measuring the risk/return of each individual stock.

The main stakeholders of this project will be any investor who wishes to invest into these individual stocks. These investors range from the individual investor to banking executives looking to diversify their stock portfolios to C-level executives who wish to purchase these stocks as a way to pump capital into their company.

Approach

This project is a time series analysis of stock prices that move in a linear direction, depending on the closing price of the stock at the end of each day. Analysis of time series data can be accomplished by using ARIMA (Autoregressive Integrated Moving Average) to convert the non-stationary data of time series to a stationary model so it can be analyzed. The Augmented Dickey-Fuller test (ADF) will be used to determine if each individual series is stationary or not.

Deliverables

There will be three deliverables with this project, all will be uploaded to my Github account (<http://www.github.com/Grodoz/Penny-Stock-Analysis>). They are the following:

- a) The project report, which will explore how this project was tackled and the solution.
- b) A Jupyter notebook which will have all of the code used in this project. The programming language that will be used is Python 3.
- c) A slide deck that will give an overview of the project, the solution, and any charts/graphs that will be relevant to the solution.

