Title: Stock data Analysis

Name: Bhushan Suryawanshi

Date: 2020-06-24

**Table of Contents**

[1. Introduction – 1](#__RefHeading___Toc1275_2387457278)

[2. Data 1](#__RefHeading___Toc1277_2387457278)

[2.1 CSV File: 1](#__RefHeading___Toc1279_2387457278)

[2.2 API Data 2](#__RefHeading___Toc1281_2387457278)

[2.3 Screen Scrapping 3](#__RefHeading___Toc1283_2387457278)

[3. Approach 3](#__RefHeading___Toc1285_2387457278)

[4. Summary 4](#__RefHeading___Toc1743_1719100124)

[5. Reference 4](#__RefHeading___Toc1287_2387457278)

# **1. Introduction** –

Stock price analysis based on its open, high, low, and close values is referred to as a Technical analysis of a stock. I am planning to do a stock’s Technical analysis to help investors/traders make data-driven decisions. Also doing chart-based analysis for each stock is time-consuming and if it can be programmed multiple stocks can be analyzed with speed. As an example, I am choosing stock data for the Zillow Group (Z) stock. It is jumped 300% in the last 6 months. Looks like a good candidate to analyze if price rally will sustain for the short/long term.

# 2. Data

## 2.1 **CSV File:**

This will be pulled from Yahoo Finance with values for daily open, close, high, low, and volume. I will be calculating additional features like – Simple moving average for 20 day (SMA20), 50 day (SMA50), 100 day (SMA100), and 200 day (SMA200). The data can be downloaded for required dates hence can show more than 1000 rows if required.

**stock\_open\_price** – Stock’s open price at the start of the trading session.

**stock\_close\_price** – Stock’s closing price at the end of the trading session.

**stock\_high\_price** – The highest price at which stock traded during the day.

**stock\_low\_price** – The lowest price at which stock traded during the day.

**stock\_day\_volume** – Number of shares traded during the day.

Simple moving Average(SMA)[[5](https://www.investopedia.com/terms/s/sma.asp)] – As its name suggest, SMA is average price of a stock calculated over a period. 10 Day SMA is average of closing prices for last 10 days.

**stock\_sma20**: Stock’s 20 day moving average.

**stock\_sma50**: Stock’s 50 day moving average.

**stock\_sma100**: Stock’s 100 day moving average.

**stock\_sma200**: Stock’s 200 day moving average.

Data source link -

<https://finance.yahoo.com/quote/Z/history?p=Z>

## **2.2 API Data**

I am using Alpha Vantage API to get the technical indicators like – Exponential Moving Average (EMA), Double Exponential Moving Average (DEMA) and Bollinger Band for 50 Day, 100 Day, and 200Day.

Exponential Moving Average (EMA)[[1](https://www.investopedia.com/ask/answers/122314/what-exponential-moving-average-ema-formula-and-how-ema-calculated.asp)]: EMA are type of weighted moving average where recent prices of stock are given more weight. Generally 50, 100, and 200 Day exponential moving averages are considered in stock analysis.

**stock\_ema\_50**: 50 day exponential moving average of the stock.

**stock\_ema\_100**: 100 day exponential moving average of the stock.

**stock\_ema\_200**: 200 day exponential moving average of the stock.

Double Exponential Moving Average (DEMA)[[3](https://www.investopedia.com/terms/d/double-exponential-moving-average.asp" \l ":~:text=The DEMA uses two exponential,price is below the average.)]: As its name suggest DEMA uses two exponential moving averages to remove the lag in price of the stock.

**stock\_dema\_50**: 50 day double exponential moving average of the stock.

**stock\_dema\_100**: 100 day double exponential moving average of the stock.

**stock\_dema\_200**: 200 day double exponential moving average of the stock.

Bollinger Bond [[4](https://www.investopedia.com/terms/b/bollingerbands.asp)]: It is a trend line two standard deviations away from simple moving average of a stock price.

**stock\_bb\_50**: Bollinger bond calculated for 50 day period.

**stock\_bb\_100**: Bollinger bond calculated for 100 day period.

**stock\_bb\_200**: Bollinger bond calculated for 200 day period.

Data source link -

<https://www.alphavantage.co/query?function=BALANCE_SHEET&symbol=Z>

## 2.3 Screen Scrapping

I am using **finscreener.com** for the technical indicators of the stock like -Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Pivot points, volume-weighted moving average.

**stock\_rsi:** Relative Strength Index (RSI)[[6](https://www.investopedia.com/terms/r/rsi.asp)] is magnitude of price change to discover overbought or oversold condition.

**stock\_macd:** Moving Average Convergence Divergence (MACD)[[7](https://www.investopedia.com/terms/m/macd.asp)] shows the trend following momentum indicator which shows relationship between two moving averages.

Pivot points[[8](https://www.investopedia.com/terms/p/pivotpoint.asp)]: It is average of high, low, and close price of the stock from previous day. Based on it you can calculate support and resistance of the stock. Different pivot points used for technical analysis are -

Pivot point, three resistance and three support levels.

**stock\_pp**: Pivot point of the stock for the day.

**stock\_r1**: Resistance level 1 of stock calculated using pivot point for the day.

**stock\_r2**: Resistance level 2 of stock calculated using pivot point for the day.

**stock\_r3**: Resistance level 3 of stock calculated using pivot point for the day.

**stock\_s1**: Support level 1 of stock calculated using pivot point for the day.

**stock\_s2**: Support level 2 of stock calculated using pivot point for the day.

**stock\_s3**: Support level 3 of stock calculated using pivot point for the day.

Data Source Link -

https://www.finscreener.com/tittech/technical-analysis?&cp=7627&tframe=4

# 3. Approach

I am collecting data from different sources and there is expected to be difference in the structure of data received from each source. Hence I will need to clean the data first. Data cleaning may require changing data type, confirming there is not duplicate data, the data point being in same measurement units, appropriate column names for the data, etc. Once the data is cleaned it will be combined in single data frame by joining cleaned data frames from each data source. Once its collected in the single data frame you can then apply desired visualizations on the data.

# 4. Summary

It was fun doing this project. I enjoyed the part of getting data from different sources. It is painful at first to find the right sources. I remember during 3rd milestone where I had to take data from screen scrapping from a website. I queried the website and went through different selectors to reach up to data point. But realized that the data is encoded in the java script objects which you cannot read with HTML parser. All my efforts of finding data source went wasted and had to start over again. Luckily I found the required source faster.

The major steps followed in my visualization project are – Collecting data, finding right data point (in case of screen scrapping or API), then cleaning data to ensure that each data point is formatted well, make meaningful renaming of columns, fill missing values wherever required because screen scrapping was not providing all required data points. In case of API I had to work with JSON data and it was easy due to earlier experience from different exercises as well as my professional experience.

Once you have cleaned data from different sources it was time to merge it. But while working through milestone 4 realized that my website for screen scrapping was under maintenance and cannot collect data. This caused problem with rerunning notebook. Then the simple solution I adapted was to put this data in the SQLite database. To have cleaner approach I merged data at each stage before adding to SQLite table.

Once the data is collected I plotted my first graph where I used entire data of 1300 rows. But looking at the graph it was too much to fit on the graph properly and hence reduced to 2 month data. In visualization I first tried to use matplotlib for my stock movement graphs. But you have to do lot of work with it. Matplotlib has another finance visualization library which is of great help – mplfinance.

Once started using this new library it was easy to merge different technical indicators of stock to plot with candlestick or daily closing price of the stock.

I feel like it was much needed exercise for at least those who doesn’t handle data in their day to day professional work. Also it builds confidence on collecting data from different sources and create your visualization story.

# 5. Reference

1. https://www.investopedia.com/ask/answers/122314/what-exponential-moving-average-ema-formula-and-how-ema-calculated.asp

2. https://www.investopedia.com/trading/using-pivot-points-for-predictions/

3. [https://www.investopedia.com/terms/d/double-exponential-moving-average.asp#:~:text=The%20DEMA%20uses%20two%20exponential,price%20is%20below%20the%20average](https://www.investopedia.com/terms/d/double-exponential-moving-average.asp" \l ":~:text=The DEMA uses two exponential,price is below the average).

4. https://www.investopedia.com/terms/b/bollingerbands.asp

5. https://www.investopedia.com/terms/s/sma.asp

6. https://www.investopedia.com/terms/r/rsi.asp

7. https://www.investopedia.com/terms/m/macd.asp

8. https://www.investopedia.com/terms/p/pivotpoint.asp