Saeid rezaei

CP640 - Machine Learning

Stock market prediction Using time series method & LSTM, Technical

Date

11-DEC-2021

Instructor: Elham Hairpoush

Contents

1. Introduction

A Stock market, equity market or share market is the aggregation of buyers and sellers of stocks, which represents ownership claims of business, Stock market and share price changes based on economy, international reputation, war, and so on. Therefore, Investors are willing to know the future and predict stock market

In this project, I'm going to analyze data related to 30 famous stock market and predict their price and up and down based on price history using analytical model called time series. In addition I will visualize my data in different ways to help portfolio managers convince investor in the better way.

2. Literature Review

2.1 Time Series Modeling

Stock Market prediction has always had a certain appeal for researchers. While numerous scientific attempts have been made, no method has been discovered to accurately predict stock price movement. The difficulty of prediction lies in the complexities of modeling market dynamics. Even with a lack of consistent prediction methods, there have been some mild successes. Stock Market research encapsulates two elemental trading philosophies; Fundamental and Technical approaches. In Fundamental analysis, Stock Market price movements are believed to derive from a security's relative data. Fundamentalists use numeric information such as earnings, ratios, and management effectiveness to determine future forecasts. In Technical analysis, it is believed that market timing is key. Technicians utilize charts and modeling techniques to identify trends in price and volume. These later individuals rely on historical data in order to predict future outcomes. Most existing literature on financial text mining relies on identifying a predefined set of keywords and machine learning techniques. These methods typically assign weights to keywords in proportion to the movement of a share price. These types of analyses have shown a definite, but weak ability to forecast the direction of share prices.

2.2 Moving Average Molding

Technical Analysis Technical analysis is the technique used to identify and utilize patterns of price history of a stock to indicate the future direction of prices. It can also be described as a method of forecasting through which attempts are made to estimate the future behavior of prices based on past observations. Technical analysis emerged in Japan during the 18th century and was in the form of candlestick techniques. Technical analysis has evolved and notable progress that emerged was the "Dow Theory" named after Charles Dow which can be considered the ground model for the modern aspect of technical analysis. The theory was a result of a compilation of 255 Wall Street journal editorials composed by Charles Dow in the 19th century. He presented theories on how to trade stocks profitably and highlighted the

various movements of the stock market. He stated "The market is always to be considered as having three movements, all going on at the same time. The first is the narrow movement from day to day. The second is the short swing, running from two weeks to a month or more; the third is the main movement, covering at least four years in its duration."(Dow, 1900 cited by William Peter Hamilton 1922, p.30). Despite the wide acceptance and utilization of this theoretical philosophy, it has been questioned by several scholars on various grounds. Prings (2002, p.2-3) defines technical analysis as a method “To identify a trend reversal at a relatively early stage and ride on that trend until the weight of the evidence shows or proves that the trend has reversed.” Whilst Murphy (1999, p.1) defines it “as the 10 study of market action, primarily through the use of charts, for the purpose of forecasting future price trends.” There exists an overwhelming skepticism among academicians regarding the use of technical analysis. The reasons of misconceptions include: Technical analysis in academic finance comprises of several variables which are not measureable. The profitability and credibility of technical analysis is discredited by the Efficient Market Hypothesis theory and the Random Walk theory.

The theoretical justification of charting techniques is difficult and the academics are unfamiliar with the jargons used to express them. Principles of Technical Analysis Three market factors are analyzed by studying price movements on the market through technical analysis: open interest, price and volume. The primary factor is price and alterations in other factors are used to confirm the accuracy of a known price trend. Core postulates of technical analysis include: All market actions are discounted in the stock price. This means that any factor that may influence or information (economic, political or psychological) that has potential impact on the price of securities is already incorporated in the price of the stock. The basis for all methods oftechnical analysis is that market prices move in accordance with trends that can be analyzed. This assumption suggests that the price movement due to a trend has two effects. First, the current trend will most likely continue and will not reverse. Second, the current trend will continue until the opposite trend begins. 11 The studies of the past are essential for understanding the future as history repeats itself. Studies of market dynamics and technical analysis are related to human psychology and the core characteristics related to the psychological state of the market are depicted through the graphical price models classified previously. The prevailing sentiment in the market (bullish or bearish) can be recognized and if models have been successful in the past there is reason to believe that they will continue to work in the future as they are based on human psychology which remains unchanged over the years. Types of Technical Analysis Among the several technical analysis techniques, the traditional two forms that analysts have resorted to are: Indicator method and Charting method. The oldest among the two is Charting which includes graphing the historical data over a period of time specified by the analyst. The analysts use their own judgment to forecast future patterns as there is no defined procedure making this method subjective. Consistency can be observed in the Indicator method as discipline is imposed on the analyst. The analyst in this case is required to conform to mathematical rules to predict future patterns. The Indicator method is favored by the economists as compared to the Charting method because it removes the subjectivity issue. Technical rules in the equity market that are studied more than others include: Moving Average, Channel and the Filter rule. When prices rise (fall) by more than a stated percentage from its most recent low (high) the filter rule suggests a buy (sell) signal. A long and short term moving average of historical prices is compared by the moving average rule. If

the short term moving average intersects the long term moving 12 average from below it is a signal to buy whereas if the short term moving average intersects the long term moving average from above it is a signal to sell. The channel suggests a buy (sell) the asset when the price increases (decreases) the maximum (minimum) over the previous n days. Recent innovations in indicators captivated the analysts more than the traditional rules. The newer indicators use mathematical functionality to determine when to buy or sell and include: Rate of Change (ROC), Moving Average Convergence Divergence (MACD), Relative Strength Indicator (RSI) and Exponentially Weighted Moving Average (EWMA).

3. **Questions to be addressed by this research**
   - Find out when the index market or Stock price is dropped and what are the important parameters
   - How is possible to compare several stock market that has same characteristic in same basket, and find out the common impact analyses
   - How long investor should keep their Index, essentially is that possible to find the Stock Movement Average (SMA) and Rate of Capital (ROC)
   - How investor or trade Analyst could find stock is oversold, or find the fraud in trade
   - How to predict the loss and gain based on Long / Short Term Memory (LSTM) model

4. **Data source**
   - Off-line stock price from Yahoo Finance and Google Finance and SP&500(https://quantquote.com/historical-stock-data, https://ca.finance.yahoo.com/)
   - On-line stock price from “Quant mod” package in R
   - Python - matlab panda and nampy lib

5. **Data format and structure**

Data format in my research is not complex; I work with below items in my file

- Stock / Index name (Symbol) , Character , Qualitative
- Stock price date , Date , Quantitative,
- Stock low price , Quantitative, Descriptive
- Stock high price, Quantitative, Descriptive
- Stock adjusted price, Quantitative, Descriptive
- Stock volume, Quantitative, Descriptive

6. **Approach**
   - Merge and cleanup the data set

- Analyze stock market and Index according to 10 years of price history
- Forecast trend market using several method such as time series, technical analyses Simple Movement Analyses (SMA) & LSTM modeling
- Visualize my analyzes with R, Python, and Tableau

7. Users or stakeholders

- Trading experts, business analysts , Investors and Auditors
- New Investor how wants to watch the trade trends and make a decision

8. Initial analyses

8.1. Define Data source

On the first stage I researched to find the most accrued data for Index and Stock market price history, I realized Yahoo Finance and Google Finance has all information that I have required. I have also used R function called "**Quantum**" which gives me ability to get price for every single symbol on-line

8.2 Analyze the input data

For this part of project I used univariate analyses:

After loading the data, I start to review them with subject expert matter and removed columns which I donot need them. Found the relation between low price and close price.

Then, Categorized the data, Symbol or Index is character, price date is consequence date value and other values such as low price, high price, close price, adjusted price are quantitative so consider number. By comparing stock and Index prices with our book of records and finance team I realized "Closeprice" is the best item to use.

At this point the question for business was to find the missing price and how they can value them; Isearched for any missing price below was the rule to consider price is missing:

- Close price is 0 or N/A
- There is no record exist for certain dates. (i.e. 2010-10-10 exist but 2010-11-01), In fact businessexpects to receive price for all Index and Symbols every day or as long as they are not expired

According to above assumption, I searched for 0 or missing records and put the value based on previous day's price.

I also found the range of my values, price date was only business day and Index price was max 6 digit, based on the Index and symbol price range I found the outliers and realized close price

for some date / symbol is not right, In order to fix this issue I used Adjusted price for month end price.

At this point question for business was what type of symbol has this type of issue, during my data analyses I realized mutual fund Company that pays distribution dividend end of each moth does not have accurate closing price hence market uses adjusted price instead.

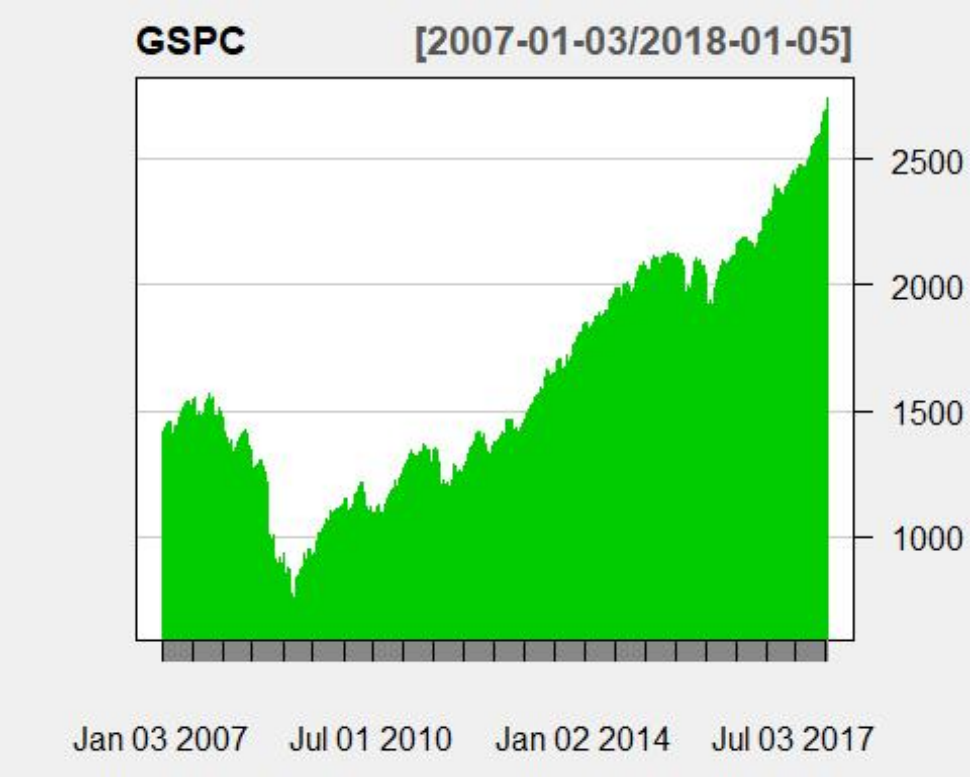
8.3 Exploring the data:

In this stage I focused on Index stock such as IBEX South American Index, GPSC American Index and compared the data in different ranges. I loaded 10 years of data.

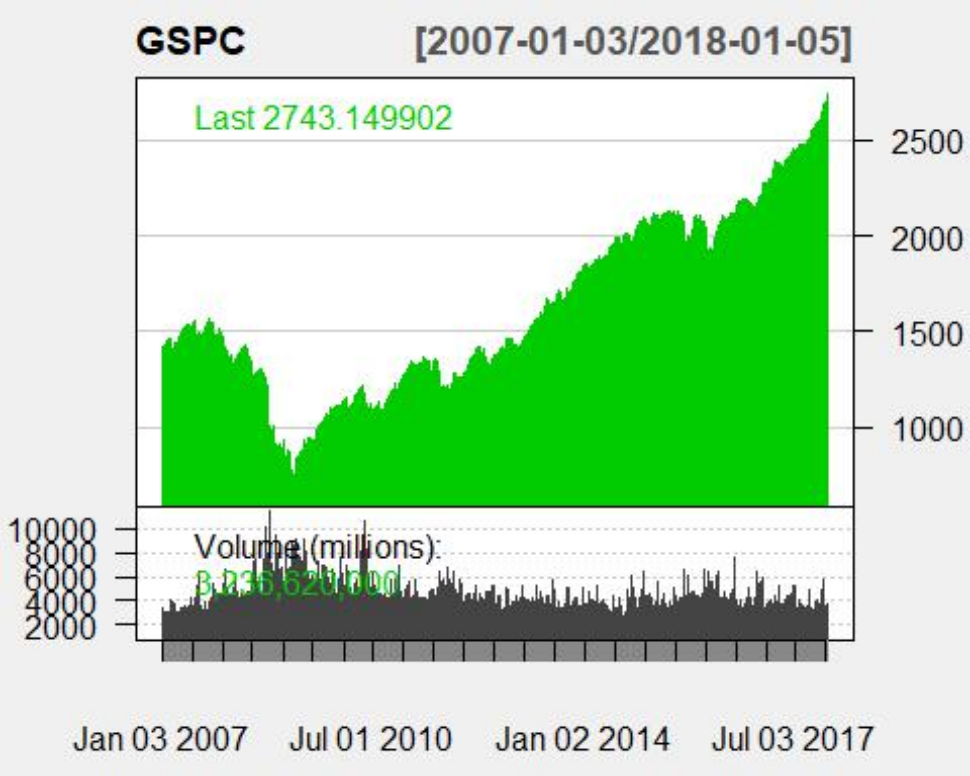
Business wanted to know the trend of above indexes also would like to compare them,

Below chart is showing the trend of GSPC (SP&500), X-vector is showing the data and Y vector is showing the price, I realized price was dropped on 2008 due to real state crisis on USA and it was not recovered till 2010. After 2010 price had organic change till 2016 again dropped due to USA election.

8.4 Visualization and exploring trends



In the below chart, I compared the price changes for the same period in addition to stock volume. I realized when the price was low and recession happened on 2008, price dropped but Index volume increased as investor tried was interested to “sell” their investment and there were no interested to buy.



Trading experts would like to know the trend of stock market and Index for shorter period of time. Below I’m going to explain about trend and few terms of trend data analyses

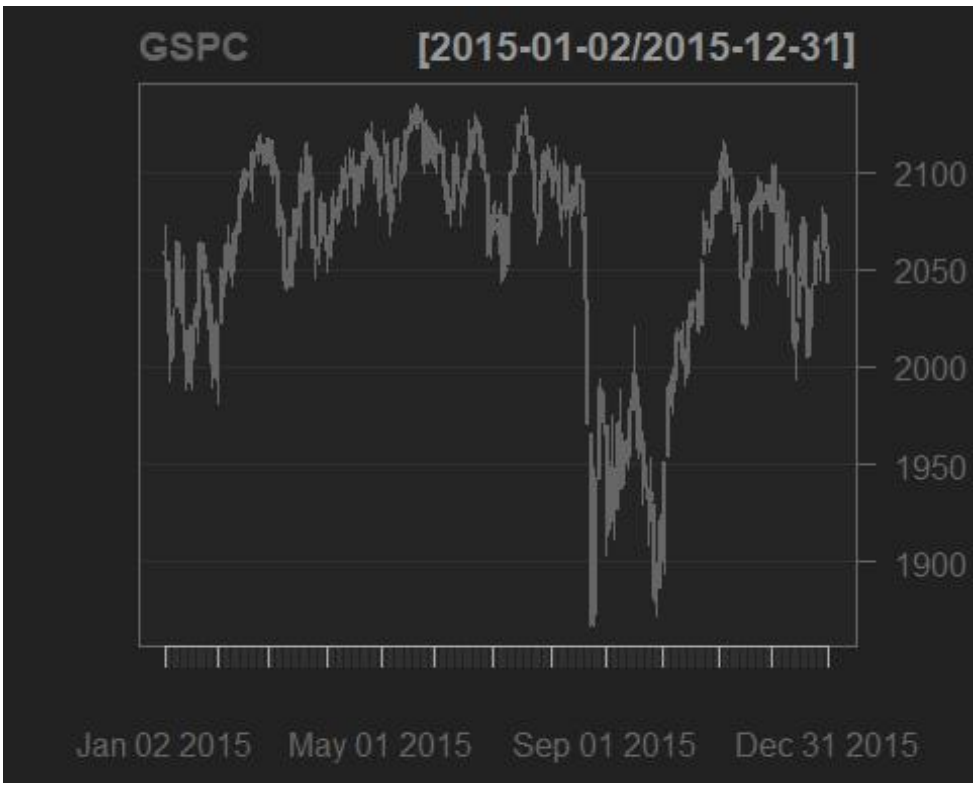
8.4.1 What is a trend?

A trend is the general direction at which the stock is moving. Based on whether the market is bullish, the trends move upwards or downwards. There is specific duration for a movement to be considered a trend, the longer the trend moves (either upwards or downwards), the more noteworthy the trend becomes.

8.4.2 What is trend Analyses?

Share Market Trendor equity market trend analysis is the process of analyzing current trends in order to predict the future trends. Using share market trend analysis, you can attempt to predict if a particular market sector growing now would continue to grow in future. Or, will a market trend in a particular sector start a trend in another. This process of share market trend analysis involves a lot of data, but nobody can predict the trends accurately %100 guarantee.

Below graph is showing GSPC had organic and normalized trends (Upwards / Downwards) in Q1-2015 and Q2-2015, however there was a long term Downwards in Q3. In this period I won't recommend to sell the investment in long period.

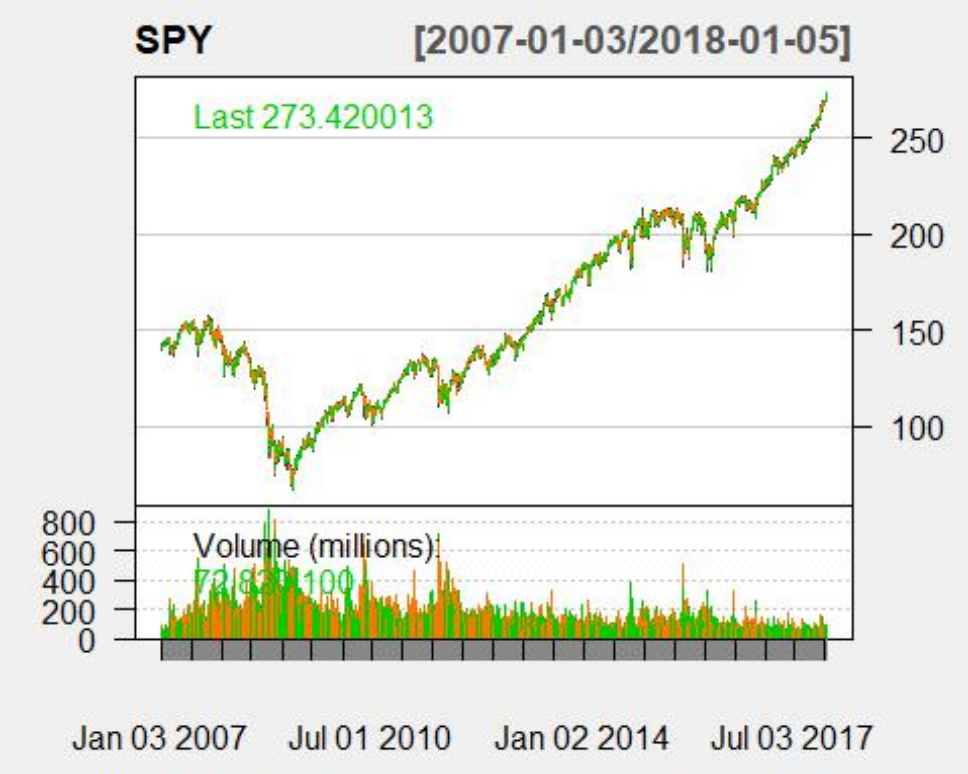

GSPC
[2015-01-02/2015-12-31]
2100
2050
2000
1950
1900
Jan 02 2015
May 01 2015
Sep 01 2015
Dec 31 2015


GSPC
[2017-01-03/2017-01-31]
2300
2290
2280
2270
2260
2250
Jan 03 2017
Jan 11 2017
Jan 20 2017
Jan 30 2017

8.4.3 Share Market 'Trend Analysis'

Share Market Trend Analysis is an aspect of technical analysis that tries to predict the future movements of a stock based on past data. A Share market trend is based on the concept that the past movements are windows to the future trends. There are three main types of share market trends, short-term, intermediate-term, and long-term.

Below graph is showing long-term trend on stock SPY. Green portion is showing when the stock starts to grow and orange part is showing when stock start going down. I realized stock volume was high when the trend is downward for long term.

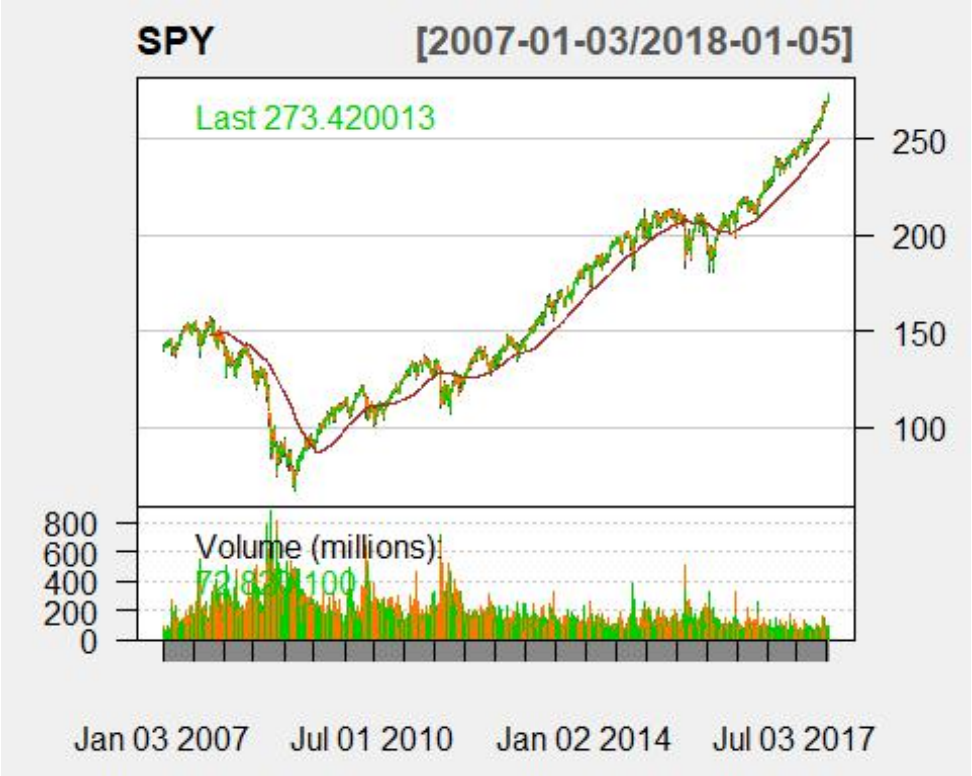


8.4.4 What is Simple Moving Average Modeling (SMA) in stock?

A simple moving average (SMA) is an arithmetic moving average calculated by adding the closing price of the security for a number of time periods and then dividing this total by the number of time periods.

As shown in the chart above, many traders watch for short-term averages to cross above longer-term averages to signal the beginning of an uptrend. Short-term averages can act as

levels of support when the price experiences a pullback. According to below chart investor could find the best time to sell their investment or keep it and predict the gain and loss.

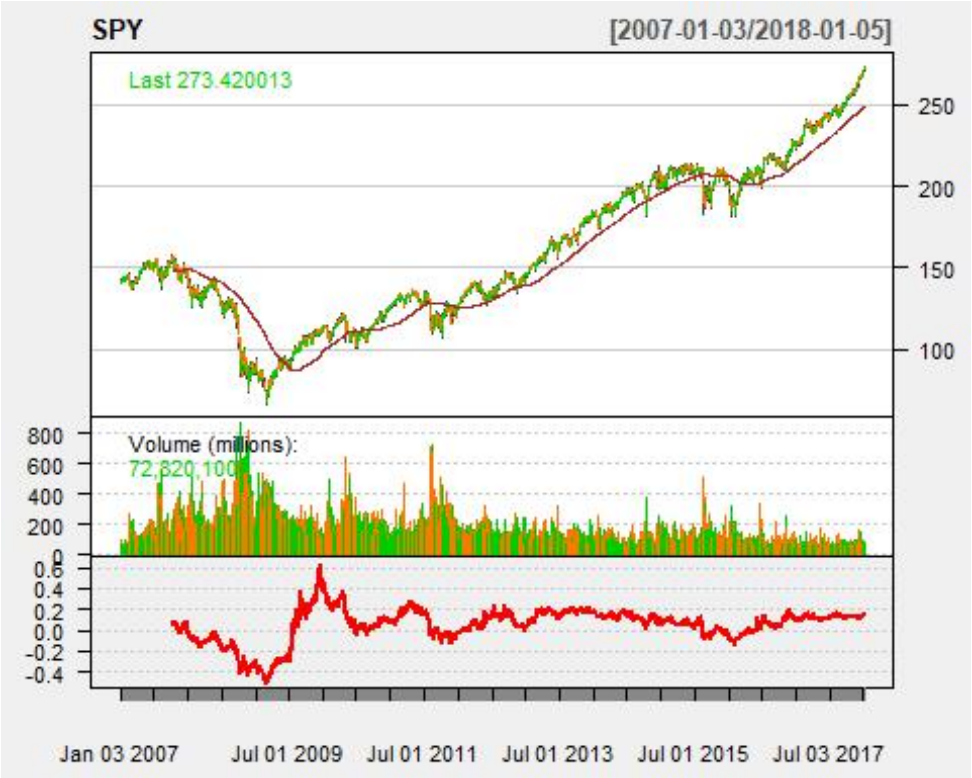


Investor advisor also would like to know, what the ROC is for certain period for one or more Index. Below graphs is showing ROC. But first let's discuss about ROC first.

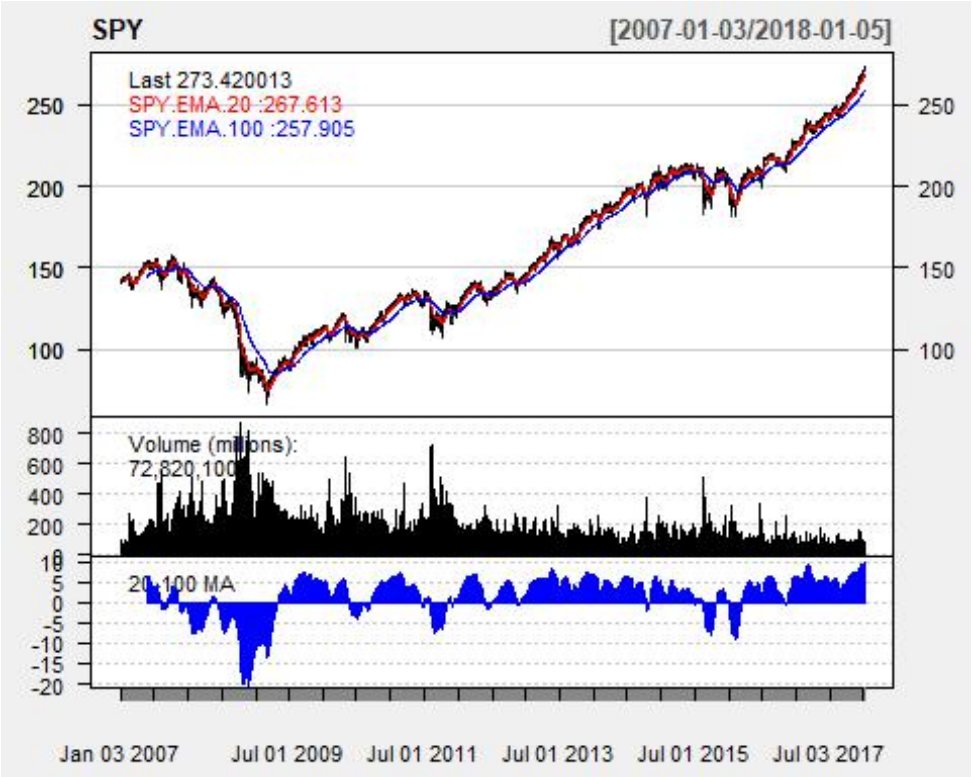
8.4.5 What is Rate of Change (ROC)?

The Rate-of-Change (ROC) indicator, which is also referred to as simply Momentum, is a pure momentum oscillator that measures the percent change in price from one period to the next. The ROC calculation compares the current price with the price "n" periods ago. The plot forms an oscillator that fluctuates above and below the zero line as the Rate-of-Change moves from positive to negative. As a momentum oscillator, ROC signals include centerline crossovers, divergences and overbought-oversold readings. Divergences fail to foreshadow reversals more often than not so this article will forgo a discussion on divergences. Even though centerline crossovers are prone to whipsaw, especially short-term, these crossovers can be used to identify the overall trend. Identifying overbought or oversold extremes comes naturally to the Rate-of-Change oscillator.

```
ROC = [(Close - Close n periods ago) / (Close n periods ago)] * 100
```

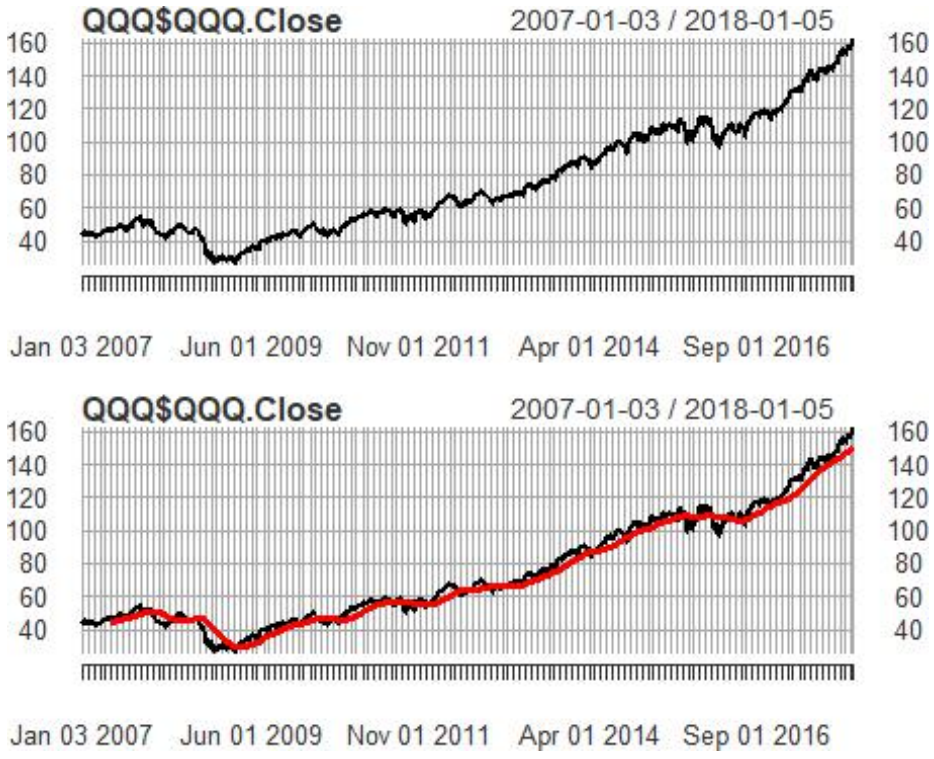


Even though momentum oscillators are best suited for trading ranges or zigzag trends, they can also be used to define the overall direction of the underlying trend. There are approximately 250 trading days in a year. This can be broken down into 125 days per half year, 63 days per quarter and 21 days per month. A trend reversal starts with the shortest timeframe and gradually spreads to the other timeframes. In general, the long-term trend is up when both the 250-day and 125-day Rate-of-Change are positive. This means that prices are higher now than they were 12 and 6 months ago. Long positions taken 6 or 12 months ago would be profitable and buyers would be happy



Below graphs shows Simply Movement Average (SMA), for a few stock. This part of Analysis could help traders and business to find the best time to buy or sell equity, in addition could be used for long-term period investment Analysis

The first plot is the raw QQQ daily closing prices and the second plot, is our smoothed version. Keep in mind that the first 100 days of price data. Can't be used as that is the minimum data we need to create a 100 period average. The issue we have is our new SMA vector contains 2065 entries, while our QQQ market download, has 2165 This should be entries.
easy to understand as it takes 100 entries to calculate an SMA .This is going to make it difficult to overlay our SMA onto the raw market data. One way around this is to buffer our SMA with 100 NA's.

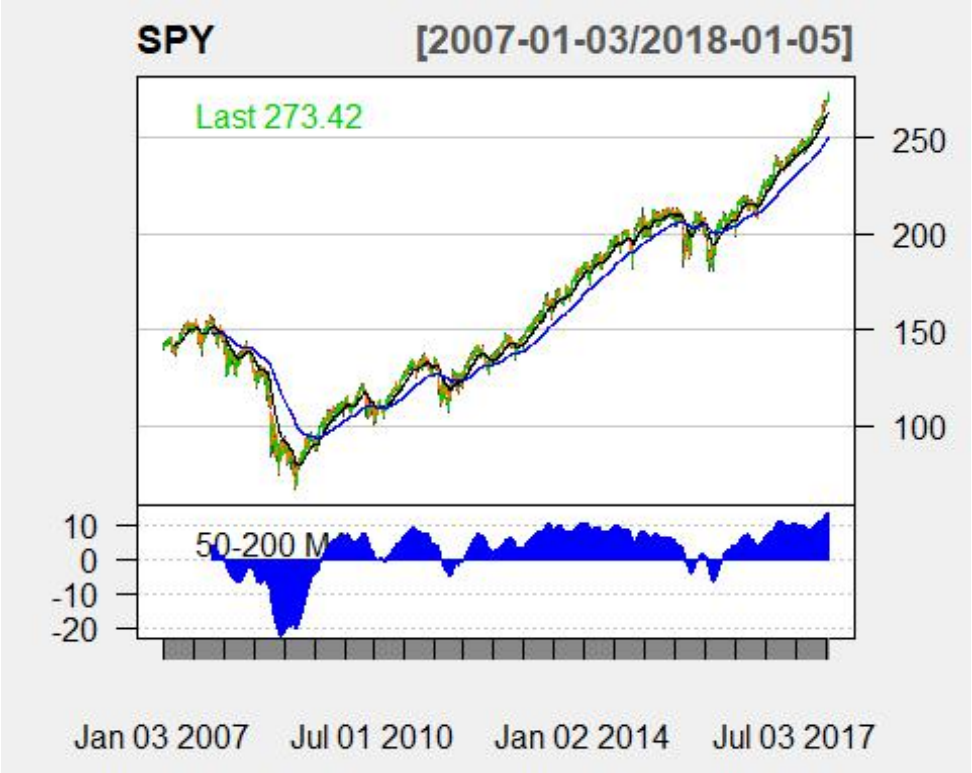


Now, I'm going to show business a comparison graph that compares long- term and short term movement, As I indicated in above chart, 200 days was used to Analyze the Movement Average, Below I'm going to compare SMA for more than one period, combination on short-term and long-term.

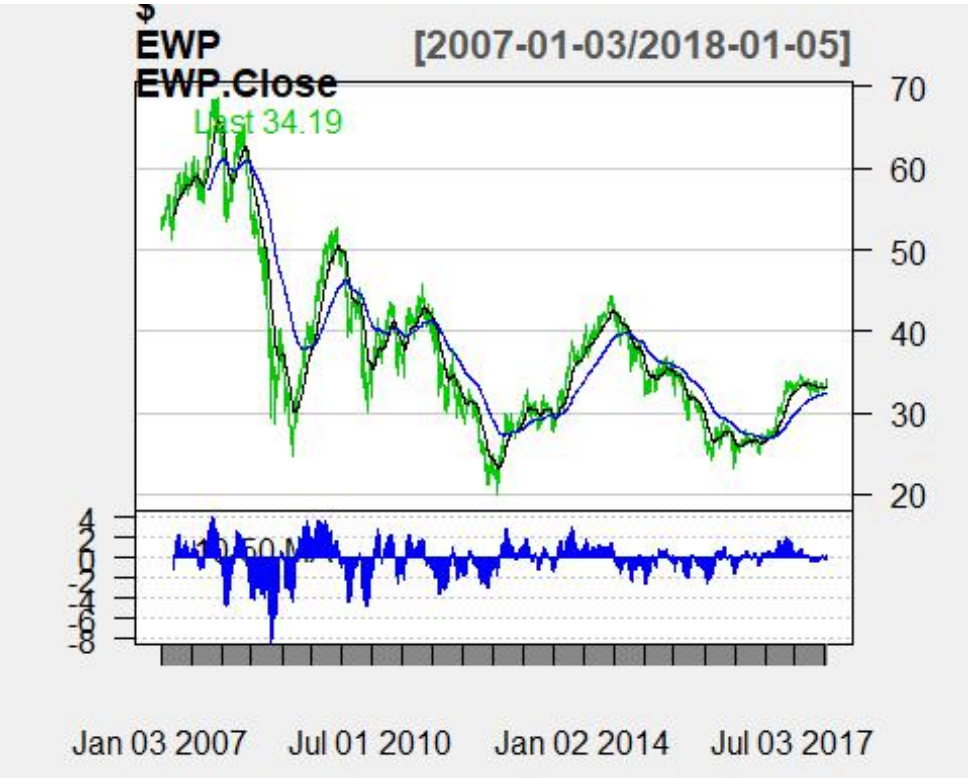

EWP
EWP.Close
[2007-01-03/2018-01-05]
Last 34.19
70
60
50
40
30
20
Jan 03 2007
Jul 01 2010
Jan 02 2014
Jul 03 2017


SPY
[2007-01-03/2018-01-05]
Last 273.42
250
200
150
100
Jan 03 2007
Jul 01 2010
Jan 02 2014
Jul 03 2017

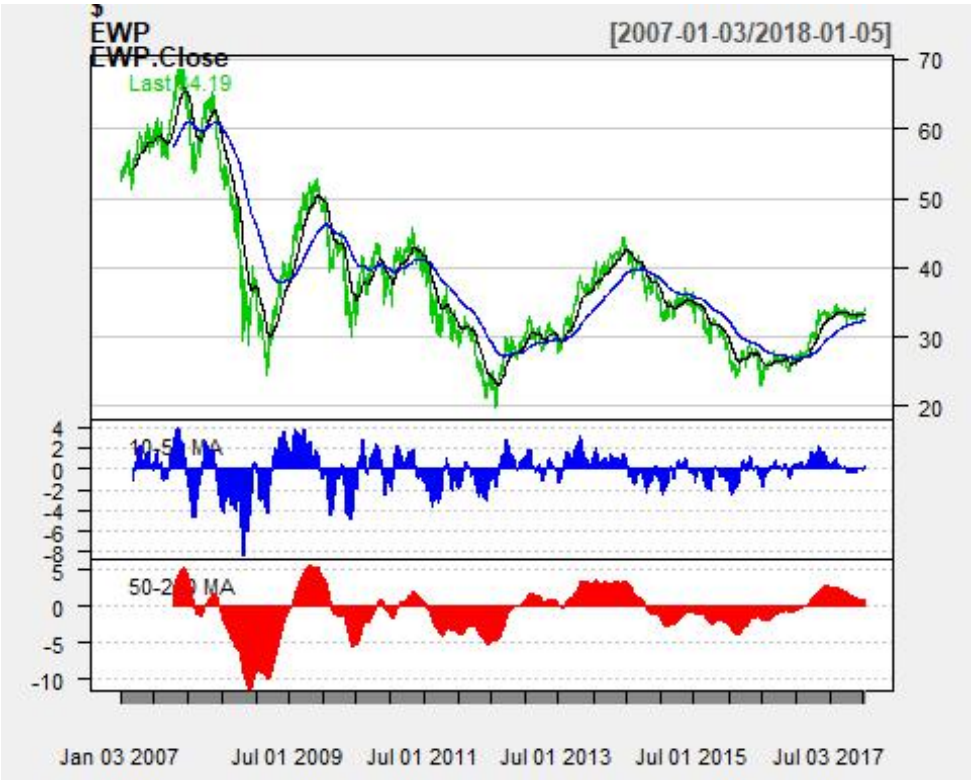
Having two moving averages of different periods removes a lot of the noise. When the fast moving average is above the slow one, the market is moving upwards, and when the fast is below the slow, it is going down. Some traders will look at the crossing of these moving averages to take a directional position.



Everything below Zero - You should not be long - and keep the Index, Holding *everything above Zero - You should not be short - and sell the Index,* Holding below graph is also showing SMA for the period of 10,50,200 days. Which can gives trader analysis to compare short term change versa long-term change, and advise investors to whether keep the holdings or sell it



In below graph I'm going to compare short-term period with mid-term period, and mid-term period with long term period.



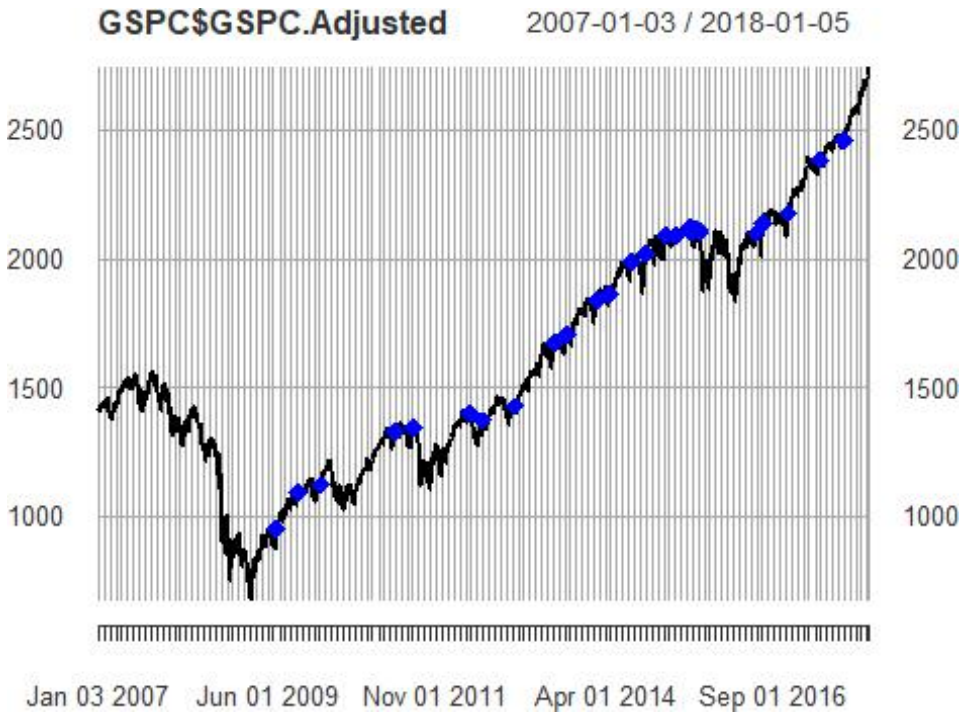
8.4.6 Trading with the trend

Trading with the trend or Trend trading is a trading strategy that attempts to capture gains through the analysis of an asset's momentum in a particular direction. The trend trader enters into a long position when a stock is trending upward (successively higher highs). Conversely, a short position is taken when the stock is in a down trend (successively lower highs).

Trade Analyst can only enter in the direction of the red Slow.Diff indicator, if its above zero you can take long signals, if its below zero, they can take short signals. The Fast.Diff indicator dictates the entries. When the blue line goes from negative to positive, its a long trade (and the slower red Slow.Diff indicator is above zero). Same thing for shorts. This is also referred to as a moving average crossover trading system.

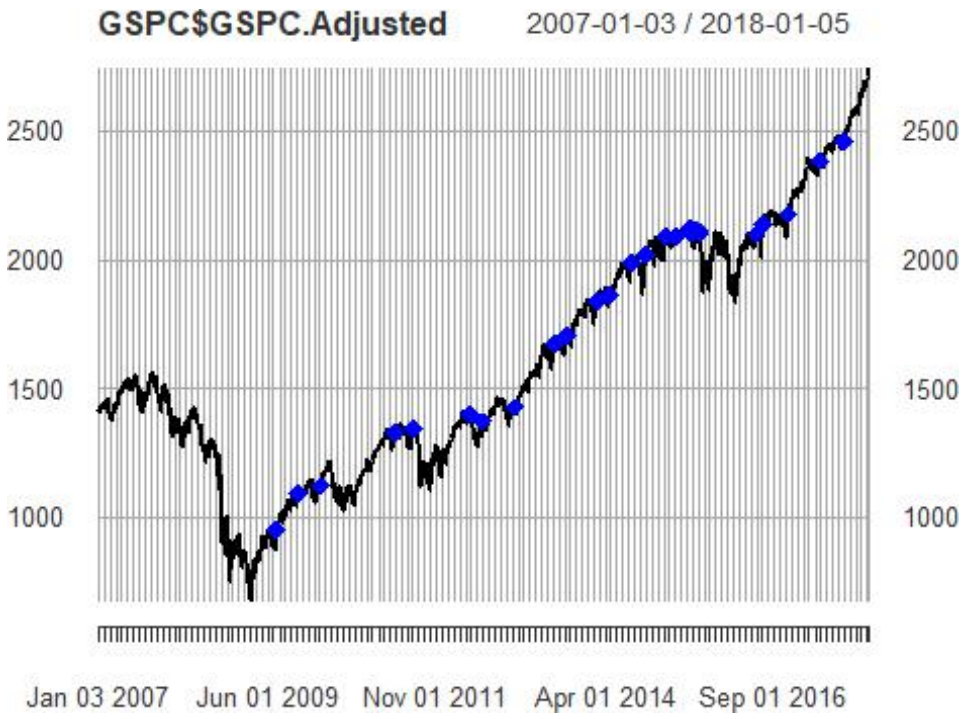
In order to build above logic I had to specify below rules:

if no position: red > 0 and blue-1 < 0 and blue > 0 go long, if long: blue < 0 exit long, if no position: red < 0 and blue-1 > 0 and blue < 0 go short, if short: blue > 0 exit short ,New challenge would to find the blue -1 means, meaning lag of blue, Pre. Price

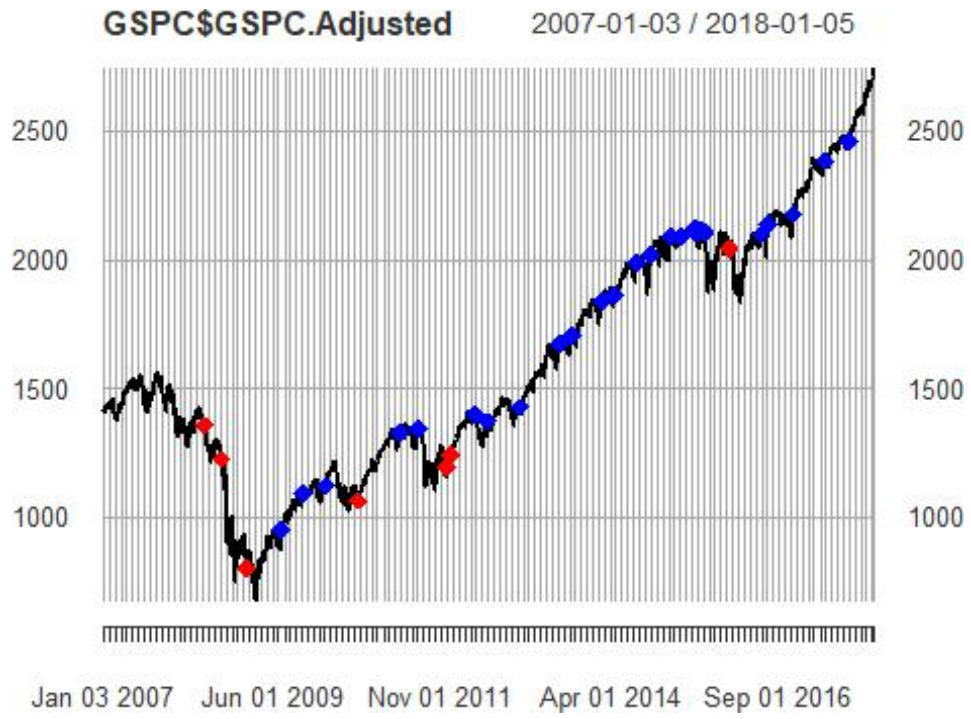


This allows us to compare the values of two different rows on the same row.
We still have our indicator value of today, but we now can compare it with yesterday's value on the same row. Sure, we could have just easily created a loop and run through each value but by doing it this way we stick to vector comparison in its simplest form.

Note: Closing price won't give us best price since company pays dividend / interest and this price is not accrue at the end of the month, hence I have used adjusted price



Below graph is showing trade analyst and experts to do short trading for the period that shows with RED sing and long trade for the position that shows with BLUE. In below example trade is recommended to keep investment for short period in 2008-2010, and keep long term their investment after 2011. In this way traders will be able to set stop loss on their trading system based on below chart. They also will be able to predict future trend based on long term movement.


GSPC$GSPC.Adjusted
2007-01-03 / 2018-01-05
2500
2000
1500
1000
Jan 03 2007
Jun 01 2009
Nov 01 2011
Apr 01 2014
Sep 01 2016

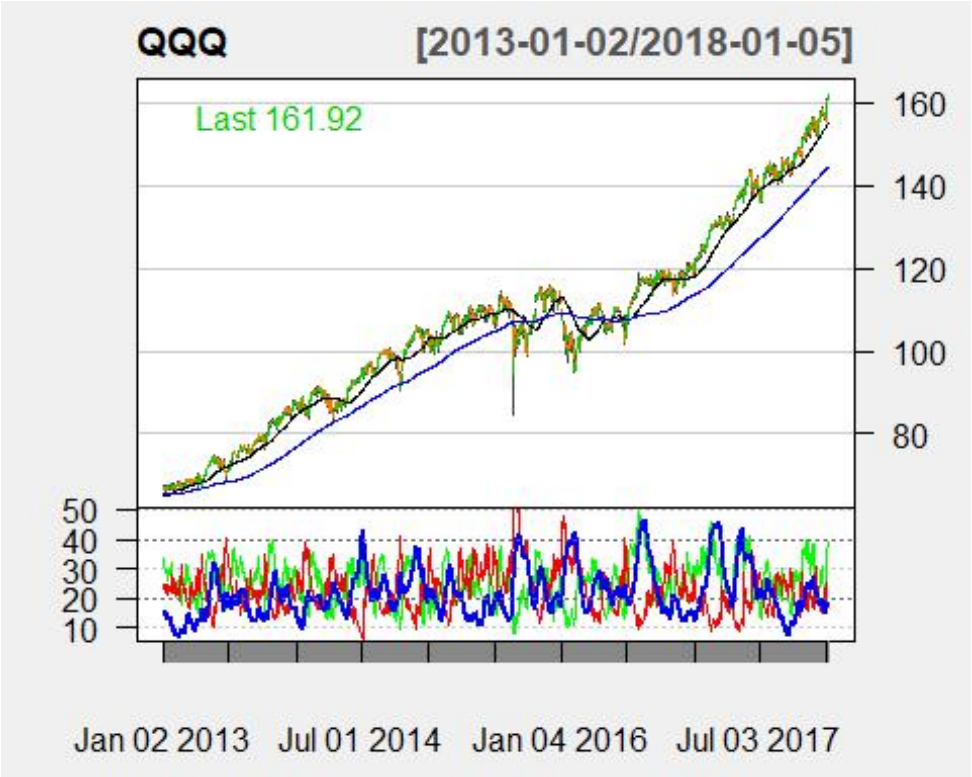

IBEX$IBEX.Adjusted
2007-01-02 / 2018-01-05
14000
12000
10000
8000
6000
Jan 02 2007
Jun 01 2009
Nov 01 2011
Apr 01 2014
Sep 01 2016

8.4.7 Analyze ADX Indicator

The ADX is Welles Wilder's Directional Movement Indicator. It is used by lots of people to determine if the market is trending or range bound.

ADX is plotted as a single line with values ranging from a low of zero to a high of 100. ADX is non-directional; it registers trend strength whether price is trending up or down. The indicator is usually plotted in the same window as the two directional movement indicator (DMI) lines, from which ADX is derived.

In below graph, black line is showing 50 days SMA, and blue line is showing 200 days SMA. This indicator could give trades to identify the market whether is strong or weak. Ten they can plan for their models and portfolios.

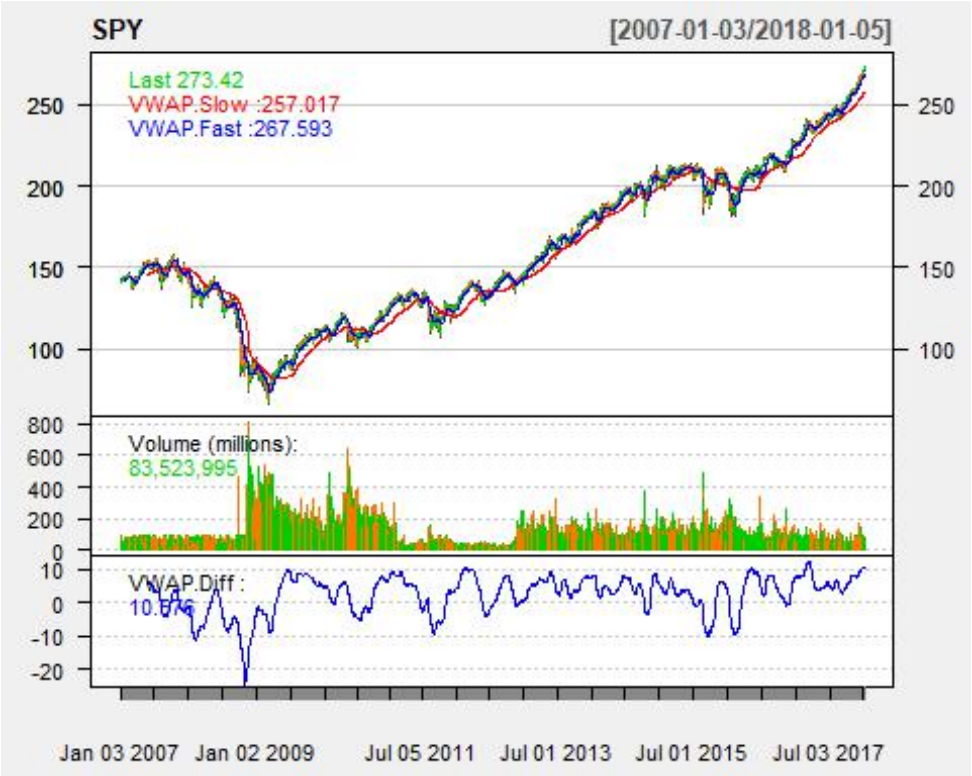


In a nutshell, Welles recommends using the ADX with a 14-day period. When the main blue line is above 20, it is considered a strong, trending market, when it is below; it is considered a weak one. Volume

As this is an introductory course, we're mostly using the closing price but it is important to note that there are a lot of other market variables available. You can design systems with the open price, the high or low, the difference between the open and close, etc. And there is also the volume.

This important indicator. A falling stack on rising volume or a rising stock on falling volume may mean the move is about to reverse. Whatever the reason for abnormal volume, it should be a warning to keep a vigilant eye on the stock.

There are plenty of indicators that include the volume price such as the Volume-weighted average price (VWAP). The VWAP is a guide more than a trading indicator as to where the market is trading compared to the volume adjusted price. It divides dollars traded by volume

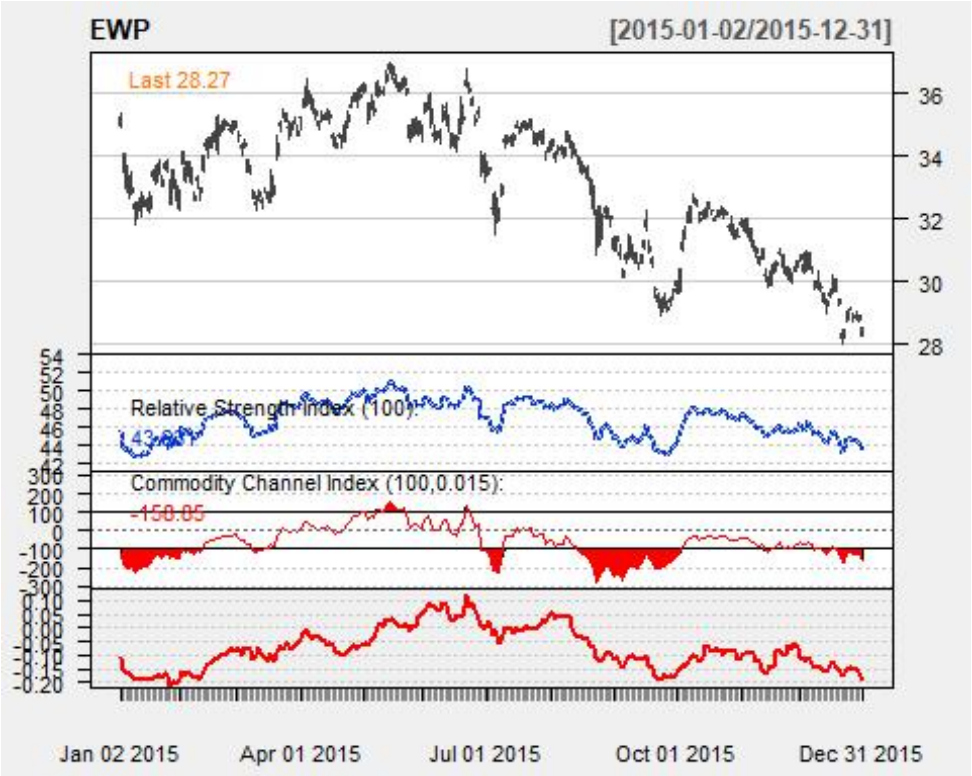


8.4.8 Relative Strength Index (RSI), and Commodity Channel Index (CCI)

Relative Strength Index (RSI), is an momentum indicator that measures movement. Its author, J. Welles Wilder, recommends using a period of 14 and when it is over 70, it is strongly bought (or overbought) and under 30, it is strongly sold (or oversold), This

indicator helps business and advisor to find out when certain index or equity is over sold or had more transaction on it. In addition could help auditors to trace the trading system

Commodity Channel Index (CCI) by Donald Lambert, is a price-derived indicator revolving around 0, where 100 is usually considered overbought and -100, oversold



8.4.9 Basket Analysis

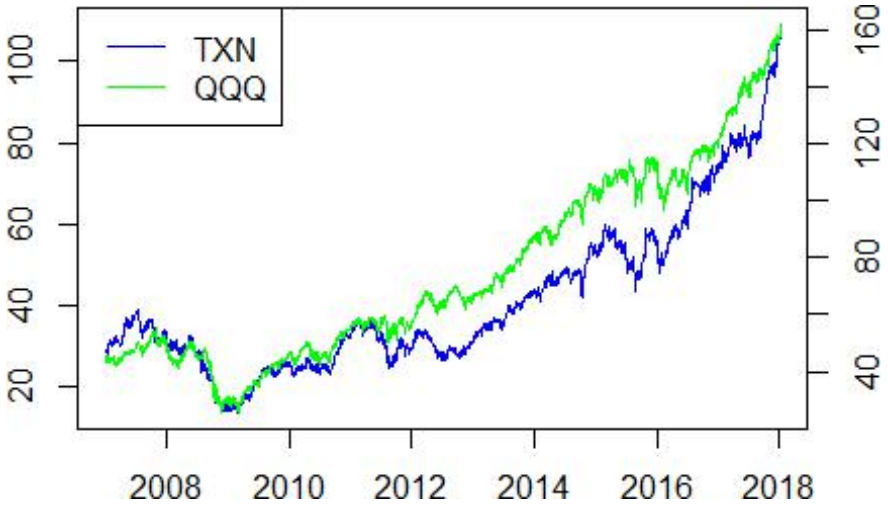
Market Basket Analysis is a modeling technique based upon the theory that if you buy a certain group of items, you are more (or less) likely to buy another group of items. For example, if you are in an English pub and you buy a pint of beer and don't buy a bar meal, you are more likely to buy crisps (US. chips) at the same time than somebody who didn't buy beer.

This feature can help business and trading Analyst to make better portfolio for their clients. For example parameter that impact Information Technology stocks are very similar, hence investor could group more than one IT stock and analyze their trends.

In below graph, I'm going to put *CSCO, INTC, MSFT, YHOO, TXN in the same group.*

In order to have better idea about trend, I have chosen one of above stock that has less movement and compared with other stocks.


MSFT
QQQ
20
40
60
80
40
80
120
160
2008
2010
2012
2014
2016
2018


TXN
QQQ
20
40
60
80
100
40
80
120
160
2008
2010
2012
2014
2016
2018

All the stocks in our basket have followed the QQQ relatively well with the exception of CISCO. The point here is that there may be arbitrage opportunities with stocks that deviate from their group or index but it's important to be cautious. Stocks deviate from their peers for a reason and may want to investigate before jumping in - whether its just a perception or a serious change.

9. Visualization using Tableau

Exploring market trends with Tableau is out of scope for this project, I just want to inform how we can connect to R from Tableau and run our scripts, R has a function call Rserver we need to install that package and call library into our memory. Then you need to go to Tableau menu bar and select “Help>>Setting and Performance >> Managing R Connection

There are 4 built-in functions in Tableau that we can write or R codes.

- SCRIPT_REAL
- SCRIPT_STR
- SCRIPT_INT
- SCRIPT_BOOL

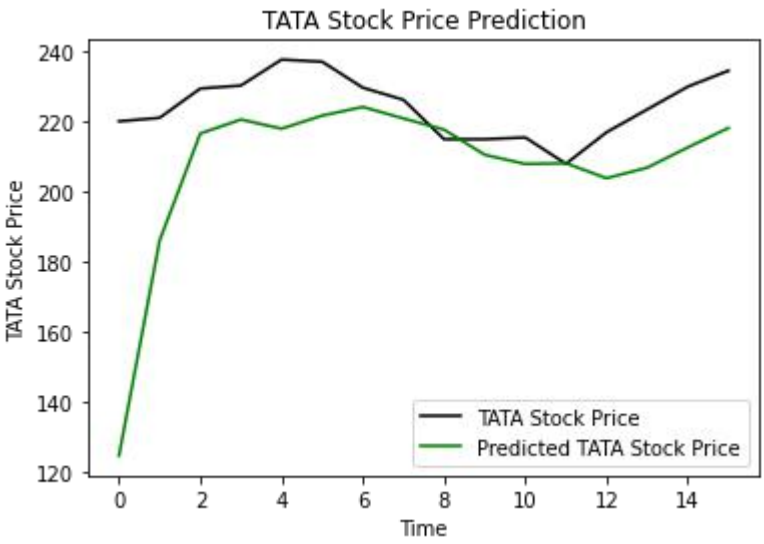
We need to create new Calculation filed and write our function using one of above built-in function in Tableau.

On below chart, I used S&P 500 index price and tried to find the impact of 3 charactersitic on the Index price such as Unemployment rate, Consumer Index , and Oil price. I noticed there are corrolation between Consumer Sentiment and Unemplyment rate, however there is low corelation with Oil price in Sp 500.

Long short-term memory (**LSTM**) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single data points (e.g. images), but also entire sequences of data (such as speech or video inputs).

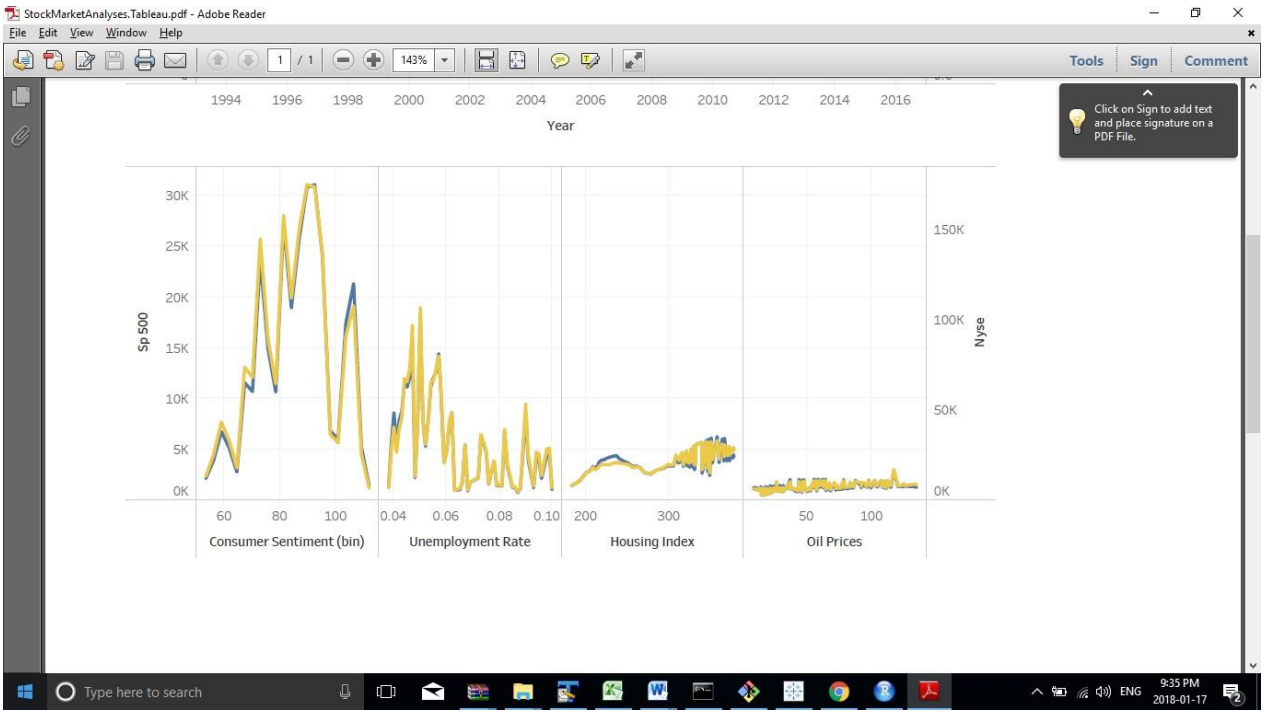
LSTM models are able to store information over a period of time.

LSTMEphoch.log.txt



As you see above, This is not the best prediction, however based on my system limitation I was able to go with 100 ephoc - In stonger PC - result could be better by increasing that number.

From the plot we can see that the real stock price went up while our model also predicted that the price of the stock will go up. This clearly shows how powerful LSTMs are for analyzing time series and sequential data.



10 Conclusion

I attempted to show a large swath of tools, functions, and methodologies in the technical and quantitative analysis field using the R and Python programming language, In addition I tried to address all of questions which was important for my users and stakeholders, there are many factors that could change the stock market trends and it's almost not possible to come up with the product that consider everything and provide prediction, However most of the indicators that I have mentioned on this report could help analysts to manage their future investment must wiser and have efficient decision for their holding.

11 References:

11.1. https://www.investopedia.com/university/technical/techanalysis9.asp

11.2. Fan, J. and Yao, Q. (2003) Nonlinear Time Series: Nonparametric and ParametricMethods, Springer-Verlag, New York, NY

11.3. Weron, R. and Misiorek, A. (2008) Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models. International Journal of Forecasting, 24(4),744–763

11.4. Tong,H. (1990) Non-Linear Time Series: A Dynamical System Approach, ClarendonPress, Oxford, UK.
11.5. http://www.albionresearch.com/data_mining/market_basket.php
11.6. https://www.investopedia.com/articles/trading/07/adx-trend-indicator.asp
11.7. https://www.investopedia.com/terms/t/trendtrading.asp
11.8. https://www.investopedia.com/terms/g/guppy-multiple-moving-average.asp
11.9. http://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:rate_of_ change_roc_and_momentum
11.10. https://www.motilaloswal.com/article.aspx/1067/How-to-analyse-stock-market-trends
11.11. https://en.wikipedia.org/wiki/Average_directional_movement_index
11.12. *Codes are in* GitHub

Saeid Rezaei