

# Predictive Analytics in Stock Markets with Special Reference to BSE Sensex

Rajesh C Jampala, Prasanna Kumar Goda, Srinivasa Rao Dokku

**Abstract**—Predictive analytics in finance is the art and science of using substantial quantities of data to find arrays. An Array can be termed as pattern or movement. Predictive analytics identifies patterns in large data volumes and helps to minimize future uncertainties. Predicting stock market returns is a puzzling task due to the multifaceted nature of the data. The present study is an applied application of the prediction and random walk theory on SENSEX behavior at an advanced level. Stock Market is the most dynamic element in the financial system and will play a crucial role in the progress of any country. The focus is on how much more on how to improve the forecasting models in terms of the performance of indices. The present model shows some commendable results in the prediction modelling reference to Indian stock market (BSE SENSEX). The designed model is also having utility for traders and investors estimating price movements of stocks at near future. Generally, the Fundamental Analysis comprises of evaluating the company's profitability on the basis of its current business environment and financial performance in the future. Technical Analysis includes interpreting the charts and using statistical figures to identify the patterns in the stock market. A number of market indicators are believed to offer signals which are beneficial in anticipating future prices. For this purpose data of BSE Sensex data has been taken (January 2010- September 2018) from bseindia.com. The results exhibit that the Sensex would also gain momentum in the year 2019.

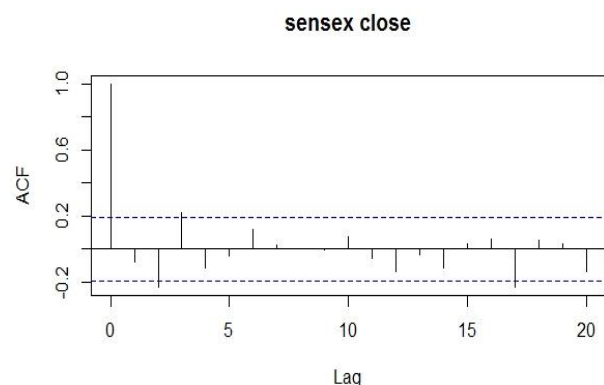
**Key words**— Predictive Analytics, Sensex, Moving Averages, R- Studio

## 1. INTRODUCTION:-

Predictive analytics in finance states about using the past data, artificial intelligence, machine learning and Trend Analysis models to predict what will be the pattern or performance of individual stock or index in the future. Normally, the historical data is used in a model which predicts trends and key patterns. These models are applied to the present data in order to forecast future movements or trends. Predictive analytics identifies different patterns by using large volumes of data and also helps to minimize future uncertainties. The data is transformed into forecasting models from which new observations can be obtained. With the help of statistical techniques and with algorithms, the historical data is been evaluated and the results are used to make predictions in terms of future performance. Strategic decision-making can be facilitated by using the outcomes of the model. Predictive analytics in finance recognizes the unseen patterns so that the managerial team can implement the necessary measures pertaining to the context. Specific forecasts and analyses provide a comprehensive root for

decisions and increase the company's financial competencies. Resources can be used more competently and be made available for other value creation accomplishments. Because of the automation, resources can be saved and the mistakes can be avoided which ultimately reduces the costs. Predicting the price movement of the stock is having its own importance in the arena of capital markets. The very pecuniary nature has drawn attention from various sections of business & academia. However, the moot question of whether the prediction of stock markets is possible or not remains the same. The concept of random walk theory explains that the stock prices are collocated in a random way and it signifies the stock markets are impulsive in nature. Many analysts have designed their own models to predict the stock prices. They have their own applicability with reference to market predictions.

Predictive Analytics deals with voluminous and varied data and uses many mathematical formulae to get the best assessment of a given context. The Predictive analytics may provide the company with a cutting edge with reference to forecasting its financials. Predictive Analytics is considered as a data science that eliminates presumptions while making decisions and also helps in finding the right solutions in the shortest possible time. Predictive analytics considers a wide variety of tools & techniques such as data mining and statistics. Game Theory is also widely used to analyze past & present to predict the upcoming events. Through learning from the rich historical data, predictive analytics offers the strategist or the analyst something outside standard business reports and also in terms of earnings forecasts. These models involve both offline and online data to foresee which investors and traders may buy, respond, click, convert, and cancel.



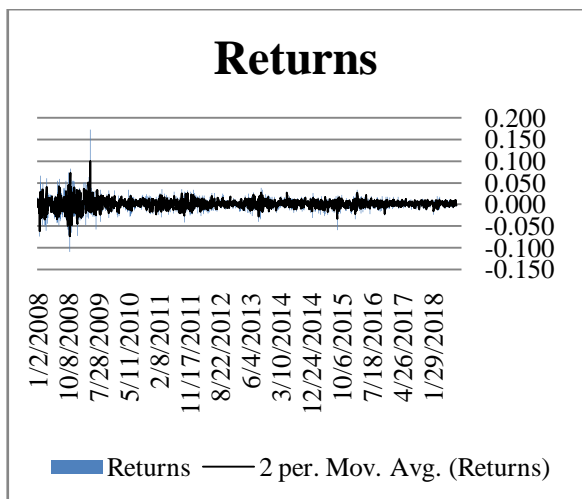
Graph1.1 Depicting the Returns of BSE Sensex from 2008 to 2018

Revised Manuscript Received on April 12, 2019.

Prof. Rajesh C Jampala, P.B. Siddhartha College of Arts & Science, Vijayawada.A.P, India.

Dr. Prasanna Kumar Goda, P.B. Siddhartha College of Arts & Science, Vijayawada.A.P, India.

Dr. SrinivasaRaoDokku, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada.A.P, India.



2. LITERATURE REVIEW

There is an increase in the study of literature on using macroeconomic variables such as index, stock and bond returns, for example, Fama and French (1988), Campbell (1987) and Balvers, Cosimano and McDonald (1990). The present study is also based on the context of predicting stock index movement by using moving averages.

Kishikawa and Tokinaga (2000) applied wavelet transform to predict the short-term feature of stock trends. Kim and Han (2000) applied neural network model by genetic algorithm. Kim and Chun (1998) applied refined probabilistic NN to estimate Index of a stock market index. Pantazopoulos et al. (1998) proposed a neuro-fuzzy model for predicting the prices of individual IBM stock

Chan., Wong., and Lam. used variables relating with technical analysis for the stocks listed in Shanghai Stock Market using Neural Network. Their results supported that the prediction of stock prices is possible with both initialization and algorithm methods.

Jabbariet al. (2001) proposed a model to characterize the time until boundary crossing and related statistics in cellular wireless networks. They recommended a model with terminal movements within a cell by through a discrete random with walk. Lai et al. (2002) have done a study about Kuala Lumpur Stock Exchange Composite Index during January 1977 - December 1999, about the daily returns in relation with trading rules and concluded about non - randomness of consecutive price changes.

Lakshmi and Roy (2012) studied the random walk hypothesis on daily, weekly and monthly returns of six indices in India from January 2000 to October 2009. The results indicated that there is no random movement in different indices. However, Lo and MacKinlay (1988) encountered mixed results when applied the variance ratio test with homoscedasticity and heteroscedasticity.

Kim and Seo (2015) studied the impact of transaction costs on price discovery and market efficiency in the European Union Emission Trading System. They found that transaction costs did impact the behavior and limit the market efficiency and the process of price discovery.

3. DATA AND EMPIRICAL RESULTS

The data sets about closing prices of SENSEX 30 on daily, monthly, quarterly and yearly basis were used in present study. The closing values of BSE SENSEX were collected from January 2010 to September 2018 from website of Bombay Stock Exchange and the analysis of the variables has been calculated by using the formula in R-Studio, Past & Excel software. Moving average method is also been adopted in order to predict the 2019 performance of BSE Sensex30.

In order to predict the seasonality behavior of the Index Ratio to Trend, Method is also been used and the index is been categorized into Quarters. Auto Correlation is the likeness among the observations as a function of the time lag between them.

Table:- 3.1 Descriptive Statistics of BSE SENSEX Yearly, Monthly & Daily data from January 2010 to September 2018

Description	Sensex Yearly	Sensex Monthly	Sensex Daily
Mean	24912.33	24031.80	23900.76
Standard Error	2104.266	595.76	129.713
Median	26117.54	23002.00	22986.12
Mode	#N/A	#N/A	#N/A
Standard Deviation	6312.799	6104.68	6015.945
Sample Variance	39851427	37267113.86	36191597
Kurtosis	-0.9296	-0.87	-0.88994
Skewness	0.153721	0.49	0.48659
Range	18601.91	23190.15	23721.55
Minimum	15454.92	15454.92	15175.08
Maximum	34056.83	38645.07	38896.63
Sum	224211	2523338.59	51410527
Count	9	105.00	2151
Largest(1)	34056.83	38645.07	38896.63
Smallest(1)	15454.92	15454.92	15175.08

Interpretation:-

From the above comparison table of Sensex daily, monthly & yearly closing values it has been inferred that the standard deviation is more in terms of yearly values when compared with the daily and monthly closing values which indicates that yearly values are much volatile than monthly and daily values.

Yearly mean average is more when compared to Sensex monthly & daily values, the standard error of the mean is the measure of the dispersion of sample means around the Sensex daily, monthly and yearly mean values. The standard error values is less in terms of daily values (129.73) of sensex when compared to (2104) of yearly closing values of sensex. The skewness values of the three variables of BSE Sensex are less than 1 and the distribution is symmetrical in nature for the period of study. The count of the variables also we can observe above for all the three segments.



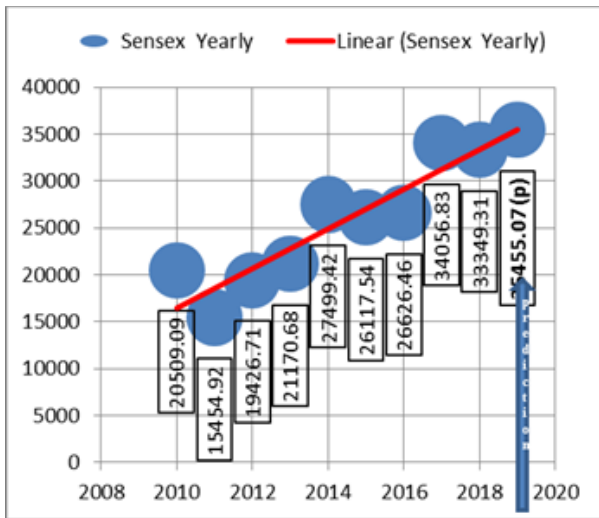
4. DATA ANALYSIS & INTERPRETATION

Forecasting of BSE SENSEX for the year 2019 by using **Moving Average Method** by using yearly Sensex close values in the

Table:- 4.1 BSE Sensex

S.No	Year	Sensex Close Price	
1	2010	20509.09	Actual Value
2	2011	15454.92	Actual Value
3	2012	19426.71	Actual Value
4	2013	21170.68	Actual Value
5	2014	27499.42	Actual Value
6	2015	26117.54	Actual Value
7	2016	26626.46	Actual Value
8	2017	34056.83	Actual Value
9	2018 (September)	33349.31	Actual Value
10	2019	35455.076	Predicted Value

Graph 4.2.1 Depicts the BSE SENSEX and Linear Trend Line



**Interpretation:** - The above table depicts the Prediction value of the variable SENSEX by using the past yearly data from 2010. By basing on the moving average calculations the sensex value would be around 35455.07 in the year 2019(predicted value). Graph 4.2.1 also depicts the value of the predicate of the year 2019 of the BSE Sensex which is depicted by the arrow.

Table:- 4.2 Predictive Analytics using Quarterly data from the year FY 2013 – FY17

Year	Q1	Q2	Q3	Q4
2013	19395.81	19379.77	21170.68	22386.27

2014	25413.78	26630.51	27499.42	27957.49
2015	27780.83	26154.83	26117.54	25341.86
2016	26999.72	27865.96	26626.46	29620.5
2017	30921.61	31283.72	34056.83	32968.68

Prediction using **Ratio to Rend Method** using quarterly closing values of SENSEX from financial year 2013 -14 to 2017-18.

Graph 4.2.2 Depicts the BSE SENSEX and Linear Trend Line

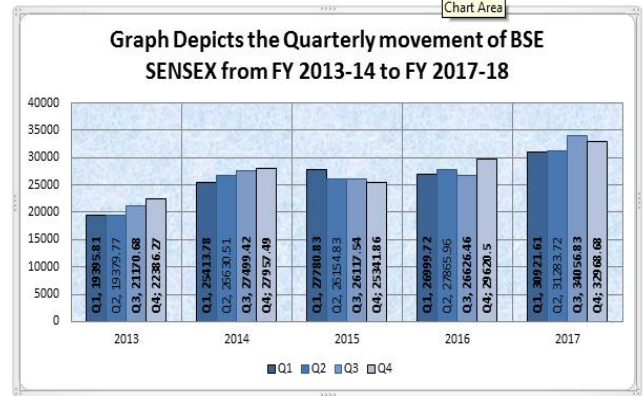


Table 4.2 depicts the quarterly closing prices of the BSE Sensex from the financial year 2013 to 2017 for the purpose of identifying the pattern and trend by using Ratio to Trend Method.

Graph 4.2.2 depicts the Quarterly movement of the Sensex from FY 2013 to FY 2017. From the graph it is been observed that while comparing the performance of all Quarters in the FY 2013, FY 2014 & FY 2016 Q4 yielded the highest value in terms of Sensex performance. In FY 2015, Q1 is the best performer and in the FY 2017 Q3 is the best performer followed by Q4.

Table:- 4.3 Calculation of SeasonalIndices for the BSE SENSEX from FY 2013 to FY 2017

Year	Q1	Q2	Q3	Q4	Yearly Total
2013	19395	19379	21170	22386	82333
2014	25413	26630	27499	27957	107501
2015	27780	26154	26117	25341	105395
2016	26999	27865	26626	29620	111113
2017	30921	31283	34056	32968	129231



Quarterly Avg(Y)	Deviations from 2015 (X)	XY	£X <sup>2</sup>
20583.13	-2	-41166.3	4
26875.30	-1	-26875.3	1
26348.77	0	0	0
27778.16	1	27778.16	1
32307.71	2	64615.42	4
$\sum y = 133893.07$		$\sum xy = 24352.02$	<b>10</b>

Using Lease Squares Method to fit a trend line

$$y = a + bx$$

$$\sum y = Na + b\sum x$$

$$\sum xy = a\sum x + b\sum x^2$$

$$a = \frac{\sum y}{n} = 26778.61 \quad b = \frac{\sum xy}{\sum x^2} = 2435.02$$

$y = 26778.61 + 2435x$	<b>Eq - 1</b>
------------------------	---------------

Table:- 4.4

Year	Deviation Value	Equation of Straight Line Trend	Central Values	Yearly Increment
2013	X = -2	$y = 26778.61 + 2435(-2)$	21908.6 1	
2014	X = -1	$y = 26778.61 + 2435(-1)$	24343.6 1	2435
2015	X = 0	$y = 26778.61 + 2435(0)$	26778.6 1	2435
2016	X = 1	$y = 26778.61 + 2435(1)$	29213.6 1	2435
2017	X = 2	$y = 26778.61 + 2435(2)$	31648.6 1	2435

From the above table it is been observed that the yearly increment of the sample **b = 2435**.

The incremental factor which is the year factor, but the data considered for the analysis is the quarterly data. It has to be converted into Quarterly Incremental value.

From the yearly value to the Quarterly Increment value =  $2435/4 = 608.75$

Table 4.5 depicts the values of calculating the trend values by using central values from 2013 to 2017 of BSE SENSEX as calculated from table 4.4. The quarterly increment value is 608.75, here from the above table by using the equations where the resultant values are the central values. In order to find the trend values for the year 2013

form the central values for Q2 and Q1 central value is been subtracted with the half of the quarterly increment value (21908.61- 304.37) and for Q3 and Q4 quarters the central value is been added with the half of the quarterly increment value (21908.61+304.37). The same procedure is been followed for the remaining years.

Table - 4.5 Calculation of Quarterly Trend Values of BSE SENSEX

Year	Q1	Q2	Central Values	Q3	Q4
2013	21299	21604	21908	22213	22518
2014	23734	24039	24343	24648	24953
2015	26169	26474	26778	27083	27388
2016	28604	28909	29213	29518	29823
2017	31039	31344	31648	31953	32258

Original Values of BSE SENSEX

Year	Q1	Q2	Q3	Q4
2013	19395	19379	21170	22386
2014	25413	26630	27499	27957
2015	27780	26154	26117	25341
2016	26999	27865	26626	29620
2017	30921	31283	34056	32968

Quarterly values as percentage of Trend Values =  $\frac{\text{Original Values}}{\text{Trend Values}} \times 100$





**Table:- 4.6 Quarterly values as percentage of Trend Values**

Year	Q1	Q2	Q3	Q4
2013	91.1	89.7	95.3	99.4
2014	107.1	110.8	111.6	112.0
2015	106.2	98.8	96.4	92.5
2016	94.4	96.4	90.2	99.3
2017	99.6	99.8	106.6	102.2
<b>Total</b>	<b>498.3</b>	<b>495.5</b>	<b>500.1</b>	<b>505.5</b>
<b>Average</b>	<b>99.7</b>	<b>99.1</b>	<b>100.0</b>	<b>101.1</b>

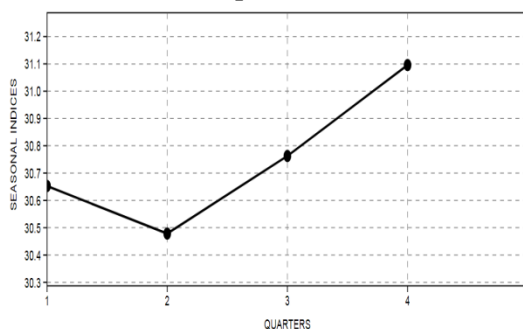
**General Average** =  $\frac{99.661+99.09+100.01+101.10}{4} = 99.713$

**Table :- 4.8 Calculation of Seasonal Indices**

Table :- 4.8 Calculation of Seasonal Indices

	$\frac{99.66}{30.65} \times 100 = 325.13$	$\frac{99.09}{30.47} \times 100 = 325.13$	$\frac{100.01}{30.76} \times 100 = 325.13$	$\frac{101.10}{31.09} \times 100 = 325.13$
<b>Seasonal Indices</b>	30.65265	30.47854	30.76264	31.09566

**Graph:- 4.2.3**



From the table 4.8 and graph 4.2.3 it is clear cut evident that the Q4 generates highest returns when compared to the

other Quarters. The seasonal indices clearly depicts that the BSE SENSEX index has performed well in terms of its value mostly in Quarter 4. Lowest Seasonal performance depicted in Q2 when compared to Q3 and Q1.

## 5. CONCLUSION:-

This study largely aims at and focuses on different statistical techniques used in predictive analytics. Subsequently, predictive analytics is an emerging area of interest to many irrespective of their field of operations. The applications of predictive analytics have provided very useful insights. In this study, it is largely emphasized the importance of the prominent techniques for predicting in terms of moving average and ratio to trend method. The results exhibit that BSE SENSEX would also perform better in the near future and the seasonality depicts that Q4 would yield higher returns when compared to other time periods. The availability of extensive computational approaches and tools of predictive modelling will help the analysts & researchers to find the optimum strategy. The analytical tools facilitate higher transparency and can find and help in analyzing the past data to forecast the future outcomes of an event or its likelihood to occur or its likely occurrence. It also provides useful insights about the hidden nature of data. The present model can be applied by coupling with soft computing techniques to foresee future trends.

## REFERENCE

1. Bollerslev, T. (1986) 'Generalized Autoregressive Conditional Heteroskedasticity', *Journal of Econometrics*, Vol. 31, pp. 307-327.
2. Chitenderu, T.T., Maredza, A. and Sibanda, K. (2014), "The Random Walk Theory And Stock Prices: Evidence From Johannesburg Stock Exchange", *The International Business & Economics Research Journal (Online)*, Vol. 13 no. 6, p.1241
3. Granger, C. W. J., 1988, "Some Recent Developments in A Concept of Causality," *Journal of Econometrics* 39, 199-211.
4. Stockwell I, (2008), *Introduction to Correlation and Regression analysis*, SAS Global Forum, Paper-364, pp. 1-8.
5. Lakshmi, V.D. and B. Roy, 2012. Testing the random walk model in Indian stock markets. *IUP Journal of Applied Finance*, 18(2): 63.
6. Wanzhong Yang. "Granule Based Knowledge Representation for Intra and Inter Transaction Association Mining", *Queensland University of Technology*, July 2009.
7. V.Argiddi ,S.S.Apte "a study of association rule mining in fragmented item-sets for prediction of transactions outcome in stock trading systems" *IJCET-2012*.

## WEBSITES:-

1. <https://www.pwc.ch/en/industry-sectors/financial-services/the-future-of-finance/data-analytics-in-the-financial-function-of-the-future.html>
2. <https://www.bseindia.com/indices/IndexArchiveData.aspx>
3. [www.nseindia.com](http://www.nseindia.com)