IJMSRR E- ISSN - 2349-6746 ISSN -2349-6738

A STUDY ON THE APPLICATION OF CORRELATION AND REGRESSION TO ASSESS THE IMPACT MADE BY THE COMPANIES LISTED UNDER BSE 30 ON THE INDEX

Jahnavi* Pothugunta Krishna Prasad** Dr. M. Venkata Ramanaiah***

*Asst. Prof. Koshys Institute of Management Studies, Bangalore, Karnataka. **Asst. Prof. Koshys Institute of Management Studies, Bangalore, Karnataka. ***Professor, Sri Venkateswara University, Tirupati, Andhra Pradesh.

Abstract

This paper specifies the causal relationship between companies which are listed under BSE 30 index from Jan 2014 to Dec 2014 (day wise data) by using Correlation and Regression. We are finding out which companies have higher impact on the index and which companies have lesser impact on the index. The Predictions of the BSE 30 index are run under three different environments (highly correlated, medium correlated and less correlated environments). Ultimately, we found the environment which is ideal for predictions, based on error measures such as Root Mean Square Error (RMSE), Mean Absolute Deviation (MAD) and Mean Absolute Percentile Error (MAPE). From the analysis, we observed that pharmaceutical companies that are listed under the 30 Index have higher impact on the index. Next automobile industry has moderate impact whereas at last banking industry. Forecasting is used to predict the future value whereas correlation is used in finding the relationship of index with the listed companies under it.

Key words: Regression, Correlation, RMSE, MAD, MAPE.

1.0 Introduction

The foremost aim of this research is to draw out the relative importance of independent variables (causal factors) on the dependent (criterion) variables in a causal relationship. Measuring relative importance of explanatory variables receives much attention in the recent works. Relative importance is itself both a multi-dimensional character and a vague concept that can have many differential meanings. Moreover, the possibility to assess any particular method to find the relative importance for all situations is difficult. The goal is to find how different fields are interdependent on each other in real world applications. The field of interest throughout the study is to find out indirect effect with direct effect from linearly correlated variables which is not possible in multiple regression analysis. So, one extremely significant point about the direct effect and indirect effect deserves special emphasis. The importance of a given independent variable is always a function of the amount of variation in that variable. This is most obvious in the case of regression co-efficient, where we are interested in the amount of change in the dependent variable by a given change in an independent variable.

But the magnitude of a correlation coefficient also depends on the extent of variation in the independent variable from which the indirect effect can be obtained, although the above fact is not recognized in regression analysis. The method of standardizing regression coefficients deals with only the direct effect.

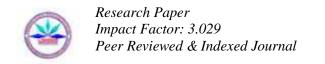
1.1 Types of Forecasting

A forecast is a prediction of some future event or events. Forecast is a quantitative estimate about the likelihood of future events. By extrapolating the models beyond the period over which they are estimated, forecasts about the future events can be made.

In general, there are two broad types of forecasting techniques-qualitative methods and quantitative methods. Qualitative forecasting techniques are often subjective in nature and require judgment on the part of experts.

Qualitative forecasts are often used in situations where there is little or no historical data, which is the base to forecast. An example would be the introduction of a new product for which there is no relevant history. In this situation, the company might use the expert opinion of sales and marketing personnel to subjectively estimate product sales during the new product introduction phase of its life cycle. Sometimes, qualitative forecasting methods make use of marketing tests, surveys of potential customers and experience with the sales performance of other products (both their own and those of competitors). However, some data analysis may be performed; the basis of forecast is subjective judgment.

Quantitative forecasting techniques make formal use of historical data and a forecasting model. The model formally summarizes patterns in the data and expresses a statistical relationship between previous and current values of the variable. Then the model is used to project the patterns of the future data. In other words, the forecasting model is used to extrapolate past and current behavior into the future. There are several types of forecasting models in general use. The three most widely used are regression models, smoothing models and general time series models. We typically think of a forecast as a single number that represents our best estimate of the future value of the variable of interest.



1.2 Uses of Forecasting

Often, there is a time lag between awareness of an event and its occurrence. This lead time is the main reason for need for planning and forecasting. There is no need for planning, if lead time is zero or small. In case, lead time is long, then the outcome of the final event is conditional and depends on identification factors. Hence, forecasting is needed to determine the occurrence of an event so that necessary action can be taken.

There is significant need for planning in management and administration because the lead time for decision making ranges from several years to few days or even to a few hours. Forecasting is an important aid in an effective and more efficient planning. For instance, in the manufacturing business, management must forecast the future demands for its products which inturn decides the requirement of materials, labour and capital to meet the demand.

Forecasting is an art of inventory control system. A firm must anticipate the demands of the items so that supplies may be done accordingly. The management must plan in advance for the inventory required. Forecasting is used to determine the need for supply of items and their requirement. Hence, for an effective and efficient planning, forecasting is an important aid. It is important to note that forecasting must result in present action to improve the future. There are three fundamental steps in forecasting- Collecting and estimating the data, Preparing the forecast and Monitoring the performance of the forecasting system. The important aspect of preparing and implementing a forecast in a particular situation is the initial phase in which the purpose of forecast is established and the necessary data is collected.

1.3 Basic Forecasting Methods

There is a wide variety of forecasting techniques available in literature. These range from the naive methods to highly complex approaches such as neural networks and econometric system of simultaneous equations. The field of forecasting is concerned with approaches to determine future. It is also concerned with proper preservation and use of forecasts. Forecasting is an integral part of the decision making activities of management.

Forecasting is a quantitative estimate about the likelihood of future events. This information is embodied in the form of a single equation model or structural model or multi-equation model or time-series model. By extrapolating the models beyond the period over which they are estimated, we can forecast the future events. The best forecast is the one which yields the forecast error with minimum variance. Basic forecasting methods which are widely used are given below:

1.3.1 Horizontal Models

The aim of horizontal forecasting model is to estimate the average demand for the entries of the past and use the average as the forecast of the demands for the future. The important models are – Naive forecast model, moving average forecasting model, single exponential smoothing, Holt's linear exponential smoothing, Holt-Winter's exponential smoothing approach, Winter's linear and seasonal exponential smoothing and Damped trend exponential smoothing.

1.3.2Trend Models

Many items in an inventory follow demand patterns where the levels either gradually rise or fall at a steady pace from time to time. This is the characteristic of the trend demand pattern where the level at time 't' is of the form, $\mu_t = a+bt$, where 'a' represents the intercept and 'b' is the slope. The trend model used to forecast seeks a straight line fit through the demand entries of the past and project the line forward to forecast the demands for the future time period. The various models are – Double moving average model, Double smoothing model and Single smoothing model with linear trend.

1.4 Regression Models for Forecasting

Simple regression is a special case of multiple regressions and multiple regressions is a special case of econometric models. While multiple regressions involve a simple equation, econometric models include a number of simultaneous regression equations. The main advantage of econometric models lies in their ability to deal with interdependencies. These models can be classified as – Simple regression forecasting technique, multiple regression forecasting technique and Quadratic regression model. The relationship between dependent and independent variables are expressed in mathematical terms. One of the best known and most commonly used causal methods is linear regression and correlation analysis.

In linear regression, one dependent variable is related to one or more independent variables by a linear equation. In a simple linear regression model, the dependent variable (Y) is a function of only one independent variable (X) and the theoretical relationship is linear or a straight line.

IJMSRR E- ISSN - 2349-6746 ISSN -2349-6738

Y = a + b X

Where Y= dependent variable

X = independent variable

a = y-intercept of the line

b = slope of the line

The objective of linear regression analysis is to determine the values of a and b that minimizes the sum of the squared deviations of the actual data points from the graphed straight line. The sample coefficient of correlations measures the direction and strength of the relationship between the independent variable and the dependent variable. The value of correlation varies between -1 to +1. Multiple Regression analysis is a practical extension of the simple regression model. It allows building a model with several independent variables instead of just one variable.

$$Y = a + b_1 X_1 + b_2 X_2 + - - - + b_n X_n$$

Where Y = dependent variable

$$X_1, X_2 - \cdots - X_n =$$
values of independent variable

a = constant

 b_1, b_2, b_3 ----- b_n are coefficients for independent variables

1.5 Criteria for Selecting Forecasting Methods

A common goal in application of forecasting technique is to minimize the errors in forecast. Thus, the efficiency of the forecasting method depends on the accuracy of its forecasts. The accuracy is generally measured in terms of various types of errors involved in it.

1.5.1 Error or Residual:

Error or Residual is defined as the difference between actual time series values (Z_t) and the forecasted values (Z_t) .

$$e_i = Z_t - \hat{Z}_t$$

1.5.2 Root Mean Squared Error (RMSE):

This is the statistic whose value is minimized during the parameter estimation process and it is the statistic that determines the width of the confidence intervals for predictions. The 95% confidence intervals for one-step-ahead forecasts are approximately equal to the point forecast 'plus or minus 2 standard errors' (i.e. plus or minus 2 times the root-mean-squared error).

The root mean squared error can only be compared between models whose errors are measured in the same units. If one model's errors are adjusted for inflation while those of another are not or if one model's errors are in absolute units while another's are in logged units, their error measures cannot be directly compared. In such cases, we have to convert the errors of both models into comparable units before computing the various measures.

The RMSE is given by
$$RMSE = \sqrt{\frac{1}{N} (Z_i - \hat{Z}_i)^2} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} e_i^2}$$

1.5.3 Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error (MAPE) is also often useful for the purpose of reporting because it is expressed in generic percentage terms which will make some kind of sense even to someone who has no idea what constitutes a 'big' error in terms of given units in the data. The MAPE can only be computed with respect to data that are guaranteed to be strictly positive.

MAPE is a very popular measure that corrects the 'canceling out' effects and also keeps into account the different scales at which this measure can be computed and thus can be used to compare different predictions. In general, a MAPE of 10% is considered very good, a MAPE in the range 20% - 30% or even higher is quite common. The MAPE is given by:

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \frac{\left| Z_{t} - \hat{Z}_{t} \right|}{Z_{t}} = \frac{100}{N} \sum_{i=1}^{N} \frac{e_{t}}{Z_{t}}$$

1.5.4 Mean Absolute Error (MAE)

MAE is another popular error measure that corrects the 'canceling out' effects by averaging the absolute value of the errors. MAE is the average over the verification sample of the absolute values of the differences between forecast and the corresponding observation. The MAE is a linear score which means that all the individual differences are weighted equally in the average.

MAE is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. Where a prediction model is to be fitted using a selected performance measure in the sense that the least squares approach is related to the mean squared error, the equivalent for mean absolute error is least absolute deviations.

The mean absolute error is given by:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| Z_{t} - \hat{Z}_{t} \right| = \frac{1}{N} \sum_{i=1}^{N} \left| e_{t} \right|$$

At the end, we should put more weights on the error measures in the estimation period most often the RMSE, but sometimes MAE or MAPE, when comparing among models. A model which fails some of the residual tests or reality checks in only a minor way is probably subject to further improvement, whereas it is the model which flunks such tests in a major way that cannot be considered as a good model.

1.6 The Objectives of the Study

- 1. To identify the companies which are highly correlated and which are least correlated with the Index BSE 30
- 2. To find out the regression equation of dependent variable (BSE 30) with the help of independent variables (listed companies under BSE 30)
- To measure the deviation between the actual and predicted values with the help of RMSE, MAD and MAPE under three environments.

Methodology: Using simple random sampling (lottery method), BSE 30 Index is selected out of various indices listed on BSE. The list of companies which come under BSE 30 Index is as follows:

HDFC	ONGC
HUL	Reliance
Hindalco	SBI
ICICI Bank	Sesa Sterlite
Infosys	Sun Pharma
ITC	Tata Motors
L & T	Tata Power
M&M	Tata Steel
Maruti	TCS
NTPC	Wipro
	HUL Hindalco ICICI Bank Infosys ITC L & T M&M Maruti

The absolute correlation of all the listed companies against the index is given in the below table:

	Highly Correla	ited	Moderately Con	rrelated	Less Correl	ated	
	r > 0.75		r >= 0.5 & <	0.75	r>=0.25 & < 0.5		
	Cipla	0.897	Airtel	0.563	GAIL	0.354	
	Dr. Reddy	0.825	Axis Bank*	0.681	Infosys*	0.379	
	HDFC Bank	0.922	BHEL	0.728	ITC	0.412	
	HDFC	0.891	Bajaj	0.551	NTPC	0.442	
	HUL	0.907	Hero	0.703	Tata Power*	0.283	
	Maruti	0.965	Coal India	0.537	Tata Steel*	0.385	
	Sun Pharma	0.848	ICICI Bank*	0.536			
	Tata Motors	0.928	L & T	0.705			
	TCS	0.787	M&M	0.601			
			Reliance*	0.537			
			SBI*	0.639			
Index			Wipro	0.702			

Note: * - Companies which have negative correlation with the BSE 30 Index.

Detailed Analysis: Highly correlated

Prediction of BSE 30 is found out by using multiple regressions, where the independent variables are Cipla, Dr.Reddy, HDFC Bank, HDFC, HUL, Maruthi, Sun Pharma, Tata Motors and TCS (under highly correlation situation).

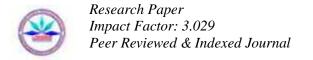
Regression S	Statistics							
Multiple R	0.98469235							
R Square	0.969619023							
Adjusted R Square	0.968484464							
Standard Error	347.2404914							
Observations	251							
	1							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	9	9.27E+08	1.03E+08	854.6217	2.3E-177			
Residual	241	29058806	120576					
Total	250	9.56E+08						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	<i>Upper</i> 95.0%
Intercept	9300.230869	668.0377	13.92171	1E-32	7984.293	10616.17	7984.293	10616.17
Cipla (X ₁)	10.74795714	0.947625	-11.342	3.56E-24	-12.6146	-8.88127	-12.6146	-8.88127
Dr. Reddy(X ₂)	0.330544312	0.173356	1.906737	0.057744	-0.01094	0.672031	-0.01094	0.672031
HDFC Bank(X ₃)	3.094474773	1.132389	2.732696	0.006747	0.863831	5.325118	0.863831	5.325118
HDFC(X ₄)	1.377974527	0.7363	-1.87148	0.062489	-2.82838	0.072431	-2.82838	0.072431
HUL(X ₅)	0.533692493	0.741107	0.720129	0.472143	-0.92618	1.993566	-0.92618	1.993566
Maruti(X ₆)	3.740062687	0.205688	18.18315	4.26E-47	3.334886	4.145239	3.334886	4.145239
Sun Pharma(X ₇)	1.284643567	0.513801	-2.50028	0.013075	-2.29676	-0.27253	-2.29676	-0.27253
Tata Motors(X ₈)	12.81827348	1.369092	9.36261	5.73E-18	10.12136	15.51519	10.12136	15.51519
TCS(X ₉)	1.710573036	0.254937	6.70979	1.38E-10	1.208384	2.212762	1.208384	2.212762

From the above table the regression equation is framed as follows:

 $Y = 9300.23 - 10.74 \ X_1 + 0.33 \ X_2 + 3.09 \ X_3 - 1.37 \ X_4 + 0.533 \ X_5 + 3.74 \ X_6 - 1.28 \ X_7 + 12. \ 81 \ X_8 + 1.71 \ X_9 \\ HIGHLY \ Correlated:$

A	Actual (BSE 30 Index)
P	Predicted (BSE 30 Index)

A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P
28800	28633	28710	28749	29000	29693	27209	27622	28163	28502	26247	26801	26638	26621	26026	25952	25521	25039	23871	23943
29044	28741	28845	28933	29122	29413	27702	27800	28178	28351	26272	26593	26560	26538	25715	25809	25190	24734	23551	23845
29044	28899	29449	29271	29183	29477	27372	27555	28047	28135	26568	26472	26443	26468	25642	25722	25228	24825	22994	23154
28879	28721	29449	29461	29682	29159	27127	27390	27941	27895	26631	26773	26437	26566	25561	25805	25576	25274	22344	22770
28885	28363	29381	29585	29559	29116	26710	26959	28009	28070	26597	26856	26420	26405	25550	25924	25474	25086	22324	22696
		-,,,,,	-,,,,,	-,,,,	-,														
28708	28208	29594	29281	29571	29211	26781	27128	27910	27615	26626	26922	26360	26342	25229	25574	25584	25198	22508	22781
20700	20200	27374	2,201	2/3/1	2/211	20701	2,120	27710	27013	20020	20722	20300	20342	2322)	25574	23364	23170	22300	22701
20517	20102	20.450	20000	20270	20010	27220	27222	27075	27500	26469	26074	26214	26204	25007	25.492	25500	25000	22445	22670
28517	28103	29459	28809	29279	29010	27320	27323	27875	27508	26468	26974	26314	26304	25007	25482	25580	25089	22445	22670



1	I	1	I	1	I	1	I	I	1	I		1	I	I	1	I	1	I	1 1
28504	28533	29220	28974	29279	28892	27351	27489	27869	27468	26745	27098	26421	26649	25024	25231	25396	24755	22404	22494
28260	28587	28747	29265	29006	28869	27602	27538	27916	27495	26776	27323	26391	26391	25373	25344	25020	24788	22418	22625
28260	28502	29008	29208	28889	28992	27831	27411	27860	27330	27207	27346	26103	26233	25445	25456	24806	24567	22466	22507
27057	20227	20005	20216	20705	20504	27797	27.421	27977	27590	27000	26990	25010	26027	25592	25970	24850	24522	22/22	22721
27957	28237	29005	29216	28785	28584	21191	27431	27866	27589	27090	20990	25919	20027	25582	25879	24859	24533	22632	22/21
27976	27796	28975	29294	28262	28294	28119	27624	27346	27214	27112	26889	25881	26027	26100	26240	24685	24544	22688	22974
27459	28070	29231	29142	28122	28368	28458	28077	27098	26908	26631	26409	25519	25816	25962	26065	24217	24252	22877	23087
27439	28070	29231	29142	20122	20300	20430	20077	27098	20908	20031	20409	23319	23610	23902	20003	24217	24232	22011	23087
27458	28055	29462	29086	28076	28306	28563	28138	26881	26555	26493	26052	25329	25556	25824	25983	24234	24647	22758	23106
28112	28448	29320	29183	27347	28207	28443	28255	26753	26803	26817	26184	25589	25840	25841	25881	24556	24667	22765	23151
28162	28567	29136	28860	27426	27948	28444	28231	26851	26952	27061	26455	25665	25918	25516	25670	24550	24636	22629	22955
28192	28603	29136	28757	27585	27800	28560	28475	26787	27029	26996	26597	25908	26107	25414	24868	24717	24798	22277	22717
28261	28647	29095	28483	27458	28236	28694	28198	26576	26602	27057	26559	25723	25923	25100	24908	24693	24700	22485	22722
28470	28828	28805	28345	27275	27884	28439	27719	26430	26273	27265	26651	25481	25542	25063	25097	24374	24262	22629	22725
20.422	******	20.524	******	*****	201	20204	****	24400	*****	25220		****	25242	25244	25120	2.1200	22044	22515	*****
28622	28696	28534	28694	26909	27604	28386	27835	26109	25897	27320	26706	25895	25362	25314	25138	24298	23911	22715	22900
28736	28602	28356	29110	26987	27604	28338	27957	25999	26442	27027	26476	26087	25642	25369	24893	24377	24018	22702	22698
28438	28738	28227	29138	27842	27591	28500	28311	26349	26671	27086	26640	25991	25542	25031	24656	24363	24165	22343	22444
20430	20/30	20221	29138	2/042	2/391	26300	26311	20349	20071	27080	20040	23991	23342	23031	24030	24303	24103	22343	22444
28503	28572	28718	29427	27888	27598	28335	28170	26384	26860	27140	26851	26127	25651	25106	24642	24122	24230	22360	22489
28930	28567	28851	29390	27499	27352	28068	28079	26297	26539	27019	26625	26272	26049	25202	24681	23906	23854	22509	22738
20730	20307	20031	2/3/0	21777	21332	20000	2001)	20271	20337	2/01/	20023	20212	2004)	25202	24001	23700	23034	2230)	22730
28659	28612	28883	29151	27404	27460	28033	28343	26637	27119	26868	26958	26147	25929	25246	24797	23815	23847	22551	22828
																		22446	22803

Moderately correlated:

Prediction of BSE 30 is found out by using multiple regressions, where the independent variables are Airtel, Axis bank, BHEL, Bajaj, Hero, Coal India, ICICI Bank, L&T, M&M, Reliance, SBI and Wipro (under moderately correlation situation)

SUMMARY OUTPUT	
Multiple R	0.98283
R Square	0.965954
Adjusted R Square	0.964238
Standard Error	369.8974
Observations	251

ANOVA:

	df	SS	MS	F	Significance F
Regression	12	9.24E+08	76993014	562.7153	3.90E-167
Residual	238	32564135	136824.1		
Total	250	9.56E+08			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	6271.417	1119.432	5.602323	5.82E-08	4066.158	8476.677	4066.158	8476.68
Airtel(X ₁)	3.500896	1.843621	1.898925	0.058784	-0.131	7.132795	-0.131	7.1328
Axis Bank(X ₂)	-0.85698	0.105865	-8.09503	2.95E-14	-1.06553	-0.64843	-1.06553	-0.6484
BHEL(X ₃)	12.84153	2.110684	6.084062	4.64E-09	8.683525	16.99954	8.683525	16.9995
Bajaj(X4)	1.634372	0.345093	4.73603	3.75E-06	0.954545	2.314199	0.954545	2.3142
Hero(X ₅)	-0.28921	0.314654	-0.91915	0.35895	-0.90908	0.33065	-0.90908	0.33065
Coal India(X ₆)	7.053088	2.163058	3.260702	0.001274	2.791904	11.31427	2.791904	11.3143
ICICI Bank(X7)	-0.1566	0.130969	-1.1957	0.233004	-0.41461	0.101407	-0.41461	0.10141
L & T(X ₈)	4.039346	0.472805	8.543373	1.58E-15	3.10793	4.970763	3.10793	4.97076
M&M(X ₉)	1.850702	0.498613	3.711699	0.000256	0.868443	2.832961	0.868443	2.83296
Reliance(X ₁₀)	-1.45196	0.827427	-1.75479	0.080581	-3.08198	0.178053	-3.08198	0.17805
SBI(X ₁₁)	-0.22529	0.05795	-3.88765	0.000131	-0.33945	-0.11113	-0.33945	-0.1111
Wipro(X12)	8.060167	1.098264	7.339006	3.38E-12	5.896607	10.22373	5.896607	10.2237

Prediction of BSE 30 index under moderately correlated situation is found out by using the regression equation: $Y = 6247.417 + 3.5 \ X_1 - 0.856 \ X_2 + 12.84 \ X_3 + 1.63 \ X_4 - 0.289 \ X_5 + 7.05 \ X_6 - 0.156 \ X_7 + 4.039 \ X_8 + 1.850 \ X_9 - 1.451 \ X_{10} - 0.22 X_{11} + 8.06 \ X_{12}$

A	Actual (BSE 30 Index)
P	Predicted (BSE 30 Index)

A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P
28800	28454	28710	29349	29000	29193	27209	27771	28163	27634	26247	26407	26638	26846	26026	25248	25521	25580	23871	23428
29044	28998	28845	29377	29122	29211	27702	27892	28178	27423	26272	26406	26560	26582	25715	25147	25190	25415	23551	22971
29044	28675	29449	29587	29183	29406	27372	27621	28047	27484	26568	26592	26443	26531	25642	25300	25228	25435	22994	22582
28879	28620	29449	29661	29682	29447	27127	27445	27941	27349	26631	26437	26437	26634	25561	25279	25576	25608	22344	22175
28885	28794	29381	29366	29559	29222	26710	27033	28009	27230	26597	26499	26420	26621	25550	25250	25474	25652	22324	22061
28708	28581	29594	29409	29571	29077	26781	27287	27910	27242	26626	26556	26360	26634	25229	25175	25584	25935	22508	22230
28517	28386	29459	28982	29279	28798	27320	27521	27875	27282	26468	26355	26314	26499	25007	24878	25580	25850	22445	22261
28504	28270	29220	29192	29279	28774	27351	27411	27869	27296	26745	26556	26421	26502	25024	24896	25396	25239	22404	22283
28260	28201	28747	29270	29006	28298	27602	27759	27916	27576	26776	26565	26391	26146	25373	25398	25020	25096	22418	22391
28260	28259	29008	29033	28889	27998	27831	27700	27860	27672	27207	26879	26103	25846	25445	25355	24806	25030	22466	22690
27957	27740	29005	29038	28785	27705	27797	27961	27866	27861	27090	27021	25919	25699	25582	25733	24859	25014	22632	22805
27976	27951	28975	29047	28262	27544	28119	28376	27346	27556	27112	27121	25881	26036	26100	26569	24685	24839	22688	22925
27459	28134	29231	28876	28122	27491	28458	28617	27098	27424	26631	26596	25519	25878	25962	26339	24217	24405	22877	22987
27458	28269	29462	28720	28076	27488	28563	28683	26881	27320	26493	26478	25329	25651	25824	26493	24234	24222	22758	22825
28112	28125	29320	28677	27347	27615	28443	28618	26753	27250	26817	26779	25589	26024	25841	26376	24556	24526	22765	22880
28162	28269	29136	28467	27426	27507	28444	28490	26851	27087	27061	27020	25665	25989	25516	26038	24550	24568	22629	22829
28192	28240	29136	28226	27585	27567	28560	28545	26787	27203	26996	27087	25908	25976	25414	25800	24717	24691	22277	22607
28261	28320	29095	28224	27458	27998	28694	28749	26576	26802	27057	27228	25723	25831	25100	25635	24693	24441	22485	22989
28470	28680	28805	27972	27275	27990	28439	28628	26430	26545	27265	27375	25481	25526	25063	25702	24374	24162	22629	22855
28622	28487	28534	28516	26909	27783	28386	28431	26109	26306	27320	27464	25895	25855	25314	25521	24298	24258	22715	22882
28736	28540	28356	28666	26987	27679	28338	28420	25999	26094	27027	27324	26087	26037	25369	25495	24377	24247	22702	22728
28438	28987	28227	28733	27842	27639	28500	28445	26349	26427	27086	27245	25991	26438	25031	25355	24363	24069	22343	22660
28503	28971	28718	28947	27888	27563	28335	28321	26384	26249	27140	27288	26127	25132	25106	25256	24122	23307	22360	22649
28930	29057	28851	29180	27499	27480	28068	28110	26297	26411	27019	26986	26272	25527	25202	25373	23906	23170	22509	22836
28659	28806	28883	29067	27404	27480	28033	27621	26637	26739	26868	27025	26147	25389	25246	25332	23815	23450	22551	23012
																		22446	22919

Less Correlated:Prediction of BSE 30 is found out by using multiple regressions, where the independent variables are Gail, Infosys, ITC, NTPC, Tata Power, and Tata Steel (under less correlation situation).

SUMMARY OUTPU	JT				
Regression Statistics					
Multiple R	0.905942				
R Square	0.820731				

Adjusted R Square	0.816323							
Standard Error	838.2931							
Observations	251							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	6	7.85E+08	1.31E+08	186.1803	4.36E-88			
Residual	244	1.71E+08	702735.2					
Total	250	9.56E+08						
		Standard				Upper	Lower	Upper
	Coefficients	Error	t Stat	P-value	Lower 95%	95%	95.0%	95.0%
Intercept	1058.341	1717.719	0.616132	0.538382	-2325.11	4441.79	-2325.11	4441.79
$GAIL(X_1)$	23.03355	2.200717	10.46638	2.04E-21	18.69872	27.36838	18.69872	27.36838
$Infosys(X_1)$	-0.48574	0.115161	-4.21786	3.48E-05	-0.71257	-0.2589	-0.71257	-0.2589
$ITC(X_1)$	30.63225	3.923567	7.807245	1.72E-13	22.90387	38.36063	22.90387	38.36063
NTPC(X ₁)	127.8325	7.547093	16.93798	4.47E-43	112.9667	142.6982	112.9667	142.6982
Tata Power(X ₁)	-126.304	13.73459	-9.19603	1.7E-17	-153.357	-99.2503	-153.357	-99.2503
Tata Steel(X ₁)	-1.48444	1.847408	-0.80353	0.422453	-5.12334	2.154461	-5.12334	2.154461

Prediction of BSE 30 index under less correlated situation is determined from the following regression equation $\mathbf{Y} = 1058.341 + 23.033 \ X_1 - 0.485 \ X_2 + 30.63 \ X_3 + 127.83 \ X_4 - 126.304 \ X_5 - 1.48 \ X_6$

A	Actual (BSE 30 Index)
P	Predicted (BSE 30 Index)

A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P	A	P
28800	28992	28710	28659	29000	27506	27209	29344	28163	28432	26247	28017	26638	26278	26026	25568	25521	25557	23871	23968
29044	29174	28845	28881	29122	27392	27702	28297	28178	28327	26272	27595	26560	25902	25715	25309	25190	25209	23551	23906
29044	29225	29449	28443	29183	27173	27372	27530	28047	27941	26568	27317	26443	26219	25642	25229	25228	25086	22994	23007
28879	29477	29449	28906	29682	26810	27127	27637	27941	27639	26631	27499	26437	25813	25561	25155	25576	25919	22344	23014
28885	29644	29381	28934	29559	27595	26710	26949	28009	27566	26597	27364	26420	25376	25550	25241	25474	25534	22324	23015
28708	29280	29594	30041	29571	27967	26781	27145	27910	27575	26626	27362	26360	25521	25229	25217	25584	25631	22508	22907
28517	28808	29459	29965	29279	28750	27320	27737	27875	27673	26468	27429	26314	25629	25007	25427	25580	25857	22445	22690
28504	27661	29220	29195	29279	28294	27351	27716	27869	27338	26745	27206	26421	25826	25024	26034	25396	25922	22404	22404
28260	27559	28747	28972	29006	27813	27602	28132	27916	27710	26776	26919	26391	25758	25373	25886	25020	25117	22418	22869
28260	27776	29008	28495	28889	27768	27831	28342	27860	27812	27207	26818	26103	25945	25445	25955	24806	25418	22466	22686
27957	27168	29005	28663	28785	27643	27797	28662	27866	28762	27090	26414	25919	25595	25582	25563	24859	25452	22632	22507
27976	27117	28975	28245	28262	28112	28119	28958	27346	28458	27112	26286	25881	25007	26100	26066	24685	25146	22688	22864
27459	27247	29231	28565	28122	28139	28458	28907	27098	28083	26631	25990	25519	24057	25962	26327	24217	25391	22877	23594
27458	28014	29462	28472	28076	28290	28563	28777	26881	28090	26493	26061	25329	24706	25824	26079	24234	24681	22758	23634
28112	27515	29320	27890	27347	28935	28443	28134	26753	27865	26817	26034	25589	24784	25841	25946	24556	24330	22765	23679
28162	26852	29136	27970	27426	28414	28444	27735	26851	28000	27061	26405	25665	24724	25516	25372	24550	25380	22629	23485
28192	28396	29136	27868	27585	28347	28560	26915	26787	27990	26996	26251	25908	25007	25414	25619	24717	26050	22277	23348
28261	27991	29095	27526	27458	29077	28694	27080	26576	28114	27057	26235	25723	25162	25100	25422	24693	26284	22485	23132
28470	28620	28805	27706	27275	28916	28439	27116	26430	27284	27265	26763	25481	25076	25063	25428	24374	26304	22629	22993
28622	28140	28534	28023	26909	28802	28386	27339	26109	26887	27320	26585	25895	25672	25314	25977	24298	25335	22715	22890
28736	29092	28356	27897	26987	29108	28338	26718	25999	26845	27027	26738	26087	26078	25369	25827	24377	24939	22702	22936
28438	29132	28227	27518	27842	28858	28500	27344	26349	26533	27086	26483	25991	25573	25031	25121	24363	25436	22343	22646
28503	28391	28718	27927	27888	28492	28335	27879	26384	26615	27140	26125	26127	25701	25106	25678	24122	24655	22360	22657
28930	27681	28851	27149	27499	28434	28068	27452	26297	27154	27019	26751	26272	25438	25202	25512	23906	24850	22509	23000
28659	27762	28883	27349	27404	28464	28033	27829	26637	27718	26868	26336	26147	25642	25246	25381	23815	24463	22551	23011
																		22446	23237

Conclusion

The table below shows the deviation between actual and predicted values from three different environments:

	RMSE	MAPE	MAD
Highly Correlated	340.2530467	0.010614441	283.7288769
Medium Correlated	360.1910477	0.010314296	275.1173962
Less Correlated	826.521055	2.486472224	667.396951

We conclude that the prediction values are closer to the actual values in highly correlated environment, when compared with the other two environments. All the pharmacy companies listed under the BSE 30 have significant relationship with the

IJMSRR E- ISSN - 2349-6746 ISSN -2349-6738

index. Power and Steel companies have the least or minimal impact on the Index and atleast 75% of the manufacturing companies are moderately related.

References

- 1. Blalock (Jr.) H.M. (1961). "Correlation and causality the multivariate case", Social forces, Vol. 39, March, pp. 246-251.
- Blalock H.M. (1961). "Evaluating the relative importance of variables", American sociological review, 26, pp. 866-874
- 3. Boudon, R (1965). "A method of linear causal analysis, Dependence analysis", American sociological review, Vol. 30, pp. 365-374
- 4. Efron, B. (1983). "Estimating the error rate of a prediction rule: improvements on cross validation". Journal of the American Statistical Association, 78, pp. 316-331
- 5. Huang, D.S. (1970). "Regression and econometric methods". John Wiley & Sons. Inc. New York. London, Sydney, Toronto
- 6. Marcoulides, A. George., and Hershberger, L. Scott. (1997). "Multivariate statistical methods: A first course". Lawrence Erlbaum Associates, Mahwah, New Jersey.