International Journal of Applied Research 2018; 4(4): 23-30

ISSN Print: 2394-7500
ISSN Online: 2394-5869
Impact Factor: 5.2
IJAR 2018; 4(4): 23-30
www.allresearchjournal.com
Received: 11-02-2018
Accepted: 15-03-2018

Panchal Bhavini V
Research Scholar at Calorx Uni. Ahmedabad, Gujarat, India

Dr. Pradeep J Jha
Research Guide at Calorx Uni. Ahmedabad, Gujarat, India

Correspondence
Dr. Pradeep J Jha
Research Guide at Calorx Uni. Ahmedabad, Gujarat, India

Statistical analysis and ARIMA model

Panchal Bhavini V and Dr. Pradeep J Jha

Abstract
One of the most critical segments in export business which requires both precision management and efficient planning is the export of fruits and vegetables. Any loophole at any stage from the procurement stage till the time it reaches to the foreign buyers may prove substandard and liable to heavy penalty and cancellation of subsequent orders and even marking the exporting corporation on black list. This has long lasting effect on the business as a result, developing multiplier effect, it to some extent, adversely affects the economy of the nation.

Keywords: ports, fruits (apple, orange, mango) export, forecast, ARIMA, ANOVA

Introduction
The basic data, related to export of some preferred fruits like Apples, Oranges and Mangoes, helps us derive averages of the first and second order, short term forecast depending upon ARIMA model, and application of ANOVA. So far as export data is concerned, we have chosen the data from reliable sources available from standard agencies dealing in the export business. The data ranging from 2013 to 2016 inclusive, catered month wise, is related to export from different ports to different countries. The export figure corresponding to each month in a particular year shows the sum of amount of exports made by different buyers during a span of one year.

Averages of first order and to represent them graphically make comparison very simple. At this point the pattern shows the months of the years in which the demand shoots up and the months in which it falls. In addition to this, time series analysis helps identify the trend and also make forecast of the coming months. In order to help prediction to a better degree we have applied ARIMA. This is a better solution to prediction over time-series analysis. Also in order to test the hypothesis of probable equality in the averages for each item (fruit) on the yearly basis ANOVA has been dealt with under certain assumptions.

These entire different notions applied help identify basic features inherent in the data and it also helps draw the future pattern of trend in export business. This process in turn becomes beneficial to the export merchants to plan their business activities at planning, procurement, and production stage.

Assumptions
In order to derive the statistical inferences from the export records of the years 2013 up to 2016, we have the following assumptions.

1. That the reliability of the secondary data is on the sound basis of primary data And it is capable enough to undergo for the data analysis.
2. That the conclusion drawn on the basis of averages of the first and second order also ANOVA techniques, without loss of generality, be applicable to the larger segment of the business falling in ‘ fruits’ category

Working with Data
In the first phase we have the data corresponding to the export of three types of fruits-Apple, Orange and Mango. The data for each correspond to the export quantity (to different countries through different ports of India) spreader over 4 years through each month.

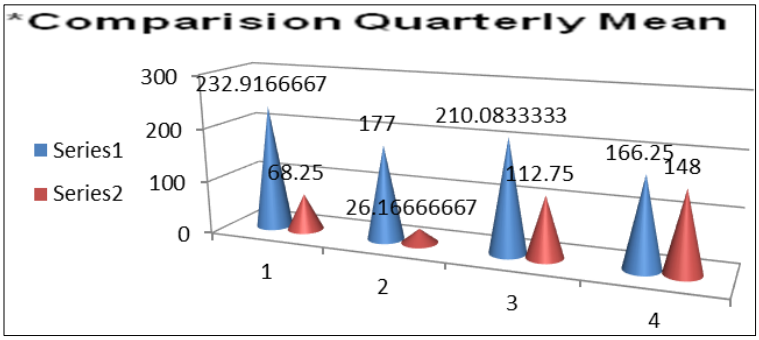
Export data of frozen Fruits from different ports

Table 1: (a) Export of Apples from Bombay air cargo

Month/Year	2013	2014	2015	2016	Avg.	Grand Avg.
	Total no. of items	Total no. of items	Total no. of items	Total no. of items		
January	102	59	161	64	819	
February	63	53	87	90		68.25
March	40	8	39	53		
April	24	5	35	14	314	
May	14	1	19	8		26.16667
June	58	49	25	62		
July	93	99	97	88	1353	
August	114	96	158	95		112.75
September	127	141	137	108		
October	159	156	171	132	1776	
November	132	156	158	102		148
December	181	192	141	96		
Total	1107	1015	1228	912		
Mean	92.25	84.58333	102.3333	76		

Table 1: (b) Export of Apples from Nahva Sheva

Month/Year	2013	2014	2015	2016	Avg.	Grand Avg.
	Total no. of items	Total no. of items	Total no. of items	Total no. of items		
January	216	294	245	154	2795	
February	159	275	269	118		232.9167
March	206	334	373	152		
April	38	247	229	144	2124	
May	189	287	125	157		177
June	102	311	170	125		
July	48	283	151	160	2521	
August	308	336	170	148		210.0833
September	312	296	122	187		
October	125	216	163	163	1995	
November	162	273	144	46		166.25
December	121	302	132	148		
	1986	3454	2293	1702		
Mean	165.5	287.8	191	141.8		



Graph 1

The graph drawn above helps comparison between quarterly averages (for all the four years) of export of Apples scheduled from the two different ports- Bombay air cargo and Nahva sheva.

Conclusion

It may be concluded from the comparison that the general trend of quarterly average from Bombay air cargo shows increasing trend from the second quarter while that from Nahva sheva shows decreasing trend from the second quarter.

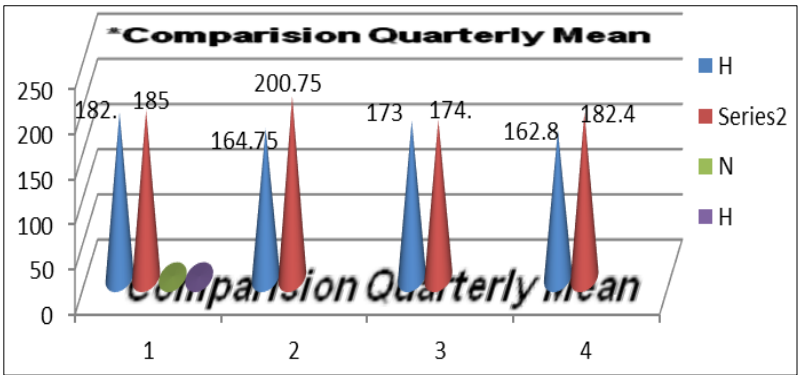
Table 2: (a) Export of oranges from Delhi air cargo

Month/Year	2013	2014	2015	2016	Avg.	Grand Avg.
	Total no. of items	Total no. of items	Total no. of items	Total no. of items		
January	186	153	264	144	2222	
February	191	172	208	129		185.1
March	136	242	191	206		
April	189	272	114	281	2409	
May	165	259	176	227		200.7

June	158	196	180	192		
July	136	193	158	156	2092	
August	185	187	163	194		174.3
September	196	205	133	186		
October	163	162	188	104	2189	
November	208	275	178	193		182.4
December	129	265	138	186		
	2042	2581	2091	2198		
Mean	170.16	215.08	174.25	183.16		

Table 2: (b) Export of oranges from Bombay air cargo

Month/Year	2013	2014	2015	2016	Avg.	Grand Avg.
	Total no. of items	Total no. of items	Total no. of items	Total no. of items		
January	146	183	169	181	2194	
February	178	182	187	183		182.8
March	208	178	165	234		
April	183	157	178	173	1977	
May	179	158	166	149		164.7
June	110	201	164	159		
July	203	159	127	210	2084	
August	179	144	175	203		173.6
September	172	197	144	171		
October	165	172	166	216	1954	
November	183	179	165	107		162.8
December	124	183	180	114		
	2030	2093	1986	2100		
Mean	169.16	174.41	165.5	175		



Graph 2

The graph drawn above helps comparison between quarterly averages (for all the four years) of export of Oranges scheduled from the two different ports- Bombay air cargo and Delhi air cargo

Conclusion

It may be concluded from the comparison that the general trend of quarterly average from Bombay air cargo shows remains steady while that from Delhi air cargo shows increasing trend towards the last quarters.

Arima Model

Introduction

In the second part of this paper we have applied ARIMA model redesigned by Box and Jenkins (1960) [1]. The basic purpose of application is forecasting over an extended period of time. This is also known as extrapolation. We deal with export data of vegetables from different ports over a period of 4 years. (2013, 2014, 2015 2016) Spreaded over all 12 months of each year. ARIMA model is designed through computer programme assisted by statistical mathematics in such a way that it works on error/ discrepancy between prediction and actual values of the variable over a period of time. One can expect ARIMA model to perform more accurate if the chronological data is Spreaded over a long time interval.

Fundamentals of ARIMA: ARIMA model with Box and Jenkins's modified approach has a general form – ARIMA (p, d, q).

Where,

p denotes the order of autoregressive Polynomial

q denotes the degree of differencing and r denotes the order of Moving average Polynomial.

Box and Jenkins's approach autocorrelation and partial autocorrelation functions play an important role in two phases; first it identifies and performs diagnostic checking then it is used to forecast over a coming period.

A simple and easy form of the ARIMA model in the case when p = q = r = 1; i.e. ARIMA (1, 1, 1) is as follows:

$$(1 - B)(1 - \emptyset B) Y_t = (1 - \propto B)\, \varepsilon_t \qquad (1)$$

This can be written as

$Y_t = (1 + \emptyset)\, Y_{t-1} - \emptyset\, Y_{t-2} + \varepsilon_t - \propto \varepsilon_{t-1}$ (2)

Y is variable corresponding to the historical data and the Forecast can be obtained using the equation (2). The extended forecast can be obtained using recurrence relation.

Descriptives

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Delhi Orange Export	48	104	281	185.67	42.604
Bombay Orange Export	48	107	234	170.79	26.087
Valid N (listwise)	48				

Descriptives

V1 = 2013

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Delhi Orange Export	12	129	208	170.17	26.309
Bombay Orange Export	12	110	208	169.17	29.248
Valid N (Listwise)	12				

a. V1 = 2013

V1 = 2014

Descriptive Statistics[a]					
	N	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Delhi Orange Export	12	153	275	215.08	44.970
Bombay Orange Export	12	144	201	174.42	17.048
Valid N (listwise)	12				

a. V1 = 2014

V1 = 2015

Descriptive Statistics[a]					
	N	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Delhi Orange Export	12	114	264	174.25	39.022
Bombay Orange Export	12	127	187	165.50	16.178
Valid N (listwise)	12				

a. V1 = 2015

V1 = 2016

Descriptive Statistics[a]					
	N	Minimum	Maximum	Mean	Std. Deviation
Delhi Orange Export	12	104	281	183.17	46.511
Bombay Orange Export	12	107	234	174.08	37.951
Valid N (listwise)	12				

a. V1 = 2016

One-way ANOVA					
		Sum of Squares	**df**	**Mean Square**	**F**
Delhi Orange Export	Between Groups	14906.167	3	4968.722	3.105
	Within Groups	70404.500	44	1600.102	
	Total	85310.667	47		
Bombay Orange Export	Between Groups	655.417	3	218.472	.307
	Within Groups	31328.500	44	712.011	
	Total	31983.917	47		

ANOVA		
		Sig.
Delhi Orange Export	Between Groups	.036
	Within Groupsz	
	Total	
Bombay Orange Export	Between Groups	.820
	Within Groups	
	Total	

Post Hoc Tests

Multiple Comparisons						
Bonferroni						
Dependent Variable	(I) V1	(J) V1	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval
						Lower Bound
Delhi Orange Export	2013	2014	-44.917	16.330	.052	-90.03
		2015	-4.083	16.330	1.000	-49.20
		2016	-13.000	16.330	1.000	-58.12
	2014	2013	44.917	16.330	.052	-.20
		2015	40.833	16.330	.097	-4.28
		2016	31.917	16.330	.342	-13.20
	2015	2013	4.083	16.330	1.000	-41.03
		2014	-40.833	16.330	.097	-85.95
		2016	-8.917	16.330	1.000	-54.03
	2016	2013	13.000	16.330	1.000	-32.12
		2014	-31.917	16.330	.342	-77.03
		2015	8.917	16.330	1.000	-36.20
Bombay Orange Export	2013	2014	-5.250	10.894	1.000	-35.35
		2015	3.667	10.894	1.000	-26.43
		2016	-4.917	10.894	1.000	-35.01
	2014	2013	5.250	10.894	1.000	-24.85
		2015	8.917	10.894	1.000	-21.18
		2016	.333	10.894	1.000	-29.76
	2015	2013	-3.667	10.894	1.000	-33.76
		2014	-8.917	10.894	1.000	-39.01
		2016	-8.583	10.894	1.000	-38.68
	2016	2013	4.917	10.894	1.000	-25.18
		2014	-.333	10.894	1.000	-30.43
		2015	8.583	10.894	1.000	-21.51

Multiple Comparisons			
Bonferroni			
Dependent Variable	(I) V1	(J) V1	95% Confidence Interval
			Upper Bound
Delhi Orange Export	2013	2014	.20
		2015	41.03
		2016	32.12
	2014	2013	90.03
		2015	85.95
		2016	77.03
	2015	2013	49.20
		2014	4.28
		2016	36.20
	2016	2013	58.12
		2014	13.20
		2015	54.03
Bombay Orange Export	2013	2014	24.85
		2015	33.76
		2016	25.18
	2014	2013	35.35
		2015	39.01
		2016	30.43
	2015	2013	26.43
		2014	21.18
		2016	21.51
	2016	2013	35.01
		2014	29.76
		2015	38.68

Time Series Modeller

Model Description			
			Model Type
Model ID	Delhi Orange Export	Model_1	ARIMA(1,0,0)

Model Summary

Model Fit							
Fit Statistic	Mean	SE	Minimum	Maximum	Percentile		
					5	10	25
Stationary R-squared	.183	.	.183	.183	.183	.183	.183
R-squared	.183	.	.183	.183	.183	.183	.183
RMSE	38.925	.	38.925	38.925	38.925	38.925	38.925
MAPE	16.975	.	16.975	16.975	16.975	16.975	16.975
MaxAPE	78.665	.	78.665	78.665	78.665	78.665	78.665
MAE	29.232	.	29.232	29.232	29.232	29.232	29.232
MaxAE	99.250	.	99.250	99.250	99.250	99.250	99.250
Normalized BIC	7.485	.	7.485	7.485	7.485	7.485	7.485

Model Fit				
Fit Statistic	Percentile			
	50	75	90	95
Stationary R-squared	.183	.183	.183	.183
R-squared	.183	.183	.183	.183
RMSE	38.925	38.925	38.925	38.925
MAPE	16.975	16.975	16.975	16.975
MaxAPE	78.665	78.665	78.665	78.665
MAE	29.232	29.232	29.232	29.232
MaxAE	99.250	99.250	99.250	99.250
Normalized BIC	7.485	7.485	7.485	7.485

Model Statistics					
Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)		
		Stationary R-squared	Statistics	DF	Sig.
Delhi Orange Export-Model_1	0	.183	14.442	17	.636

ARIMA Model Parameters					
					Estimate
Delhi Orange Export-Model_1	Delhi Orange Export	No Transformation	Constant		185.676
			AR	Lag 1	.419

Model Statistics	
Model	Number of Outliers
Delhi Orange Export-Model_1	0

ARIMA Model Parameters					
					SE
Delhi Orange Export-Model_1	Delhi Orange Export	No Transformation	Constant		9.532
			AR	Lag 1	.133

ARIMA Model Parameters					
					t
Delhi Orange Export-Model_1	Delhi Orange Export	No Transformation	Constant		19.479
			AR	Lag 1	3.146

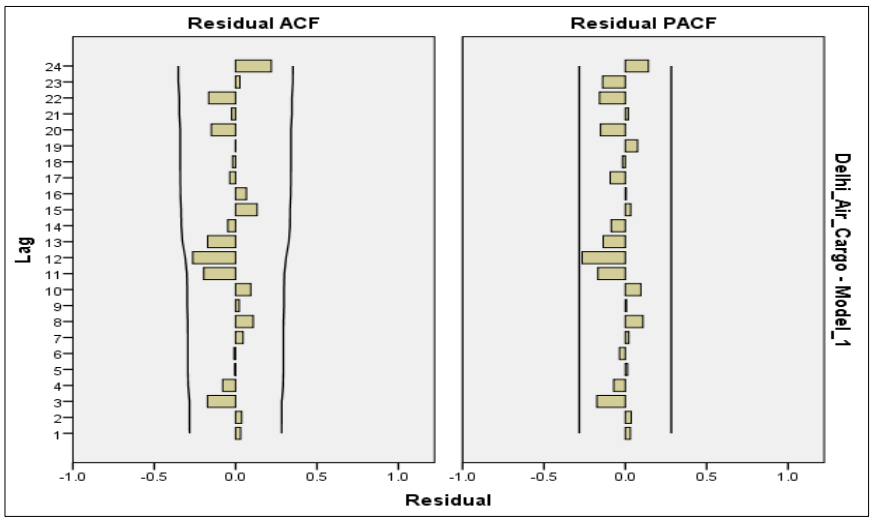
ARIMA Model Parameters					
					Sig.
Delhi Orange Export-Model_1	Delhi Orange Export	No Transformation	Constant		.000
			AR	Lag 1	.003

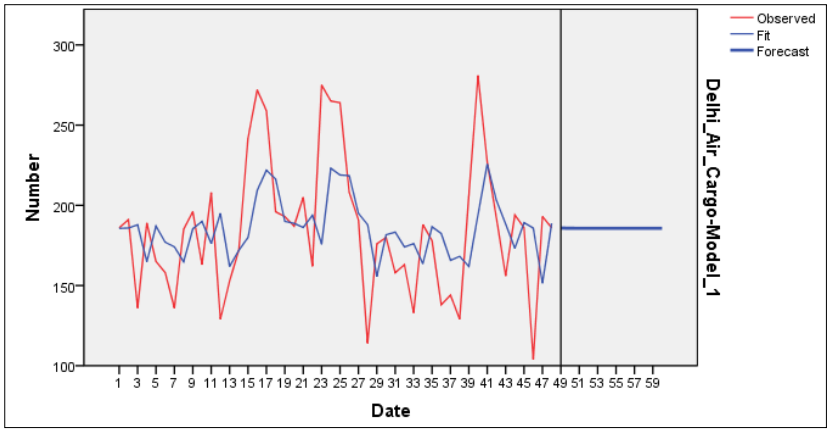
Forecast						
Model		49	50	51	52	53
Delhi Orange Export-Model_1	Forecast	186	186	186	186	186
	UCL	264	271	272	272	272
	LCL	107	101	100	99	99

Forecast						
Model		54	55	56	57	58
Delhi Orange Export-Model_1	Forecast	186	186	186	186	186
	UCL	272	272	272	272	272
	LCL	99	99	99	99	99

Forecast			
Model		59	60
Delhi Orange Export-Model_1	Forecast	186	186
	UCL	272	272
	LCL	99	99

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.





Time Series Modeler

Model Description			
			Model Type
Model ID	Bombay Orange Export	Model_1	ARIMA(0,0,0)

Model Summary

Model Fit						
Fit Statistic	Mean	SE	Minimum	Maximum	Percentile	
					5	10
Stationary R-squared	2.554E-015	.	2.554E-015	2.554E-015	2.554E-015	2.554E-015
R-squared	2.554E-015	.	2.554E-015	2.554E-015	2.554E-015	2.554E-015
RMSE	26.087	.	26.087	26.087	26.087	26.087
MAPE	12.274	.	12.274	12.274	12.274	12.274
MaxAPE	59.618	.	59.618	59.618	59.618	59.618
MAE	19.076	.	19.076	19.076	19.076	19.076
MaxAE	63.792	.	63.792	63.792	63.792	63.792
Normalized BIC	6.603	.	6.603	6.603	6.603	6.603

Model Fit					
Fit Statistic	Percentile				
	25	50	75	90	95
Stationary R-squared	2.554E-015	2.554E-015	2.554E-015	2.554E-015	2.554E-015
R-squared	2.554E-015	2.554E-015	2.554E-015	2.554E-015	2.554E-015
RMSE	26.087	26.087	26.087	26.087	26.087
MAPE	12.274	12.274	12.274	12.274	12.274
MaxAPE	59.618	59.618	59.618	59.618	59.618
MAE	19.076	19.076	19.076	19.076	19.076
MaxAE	63.792	63.792	63.792	63.792	63.792
Normalized BIC	6.603	6.603	6.603	6.603	6.603

Model Statistics				
Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)		
		Stationary R-squared	Statistics	DF	Sig.
Bombay Orange Export-Model_1	0	2.554E-015	13.893	18	.736

Model Statistics	
*Model	Number of Outliers
Bombay Orange Export-Model_1	0

ARIMA Model Parameters				
				Estimate
Bombay Orange Export-Model_1	Bombay Orange Export	No Transformation	Constant	170.792

ARIMA Model Parameters				
				SE
Bombay Orange Export-Model_1	Bombay Orange Export	No Transformation	Constant	3.765

ARIMA Model Parameters				
				t
Bombay Orange Export-Model_1	Bombay Orange Export	No Transformation	Constant	45.360

ARIMA Model Parameters				
				Sig.
Bombay Orange Export-Model_1	Bombay Orange Export	No Transformation	Constant	.000

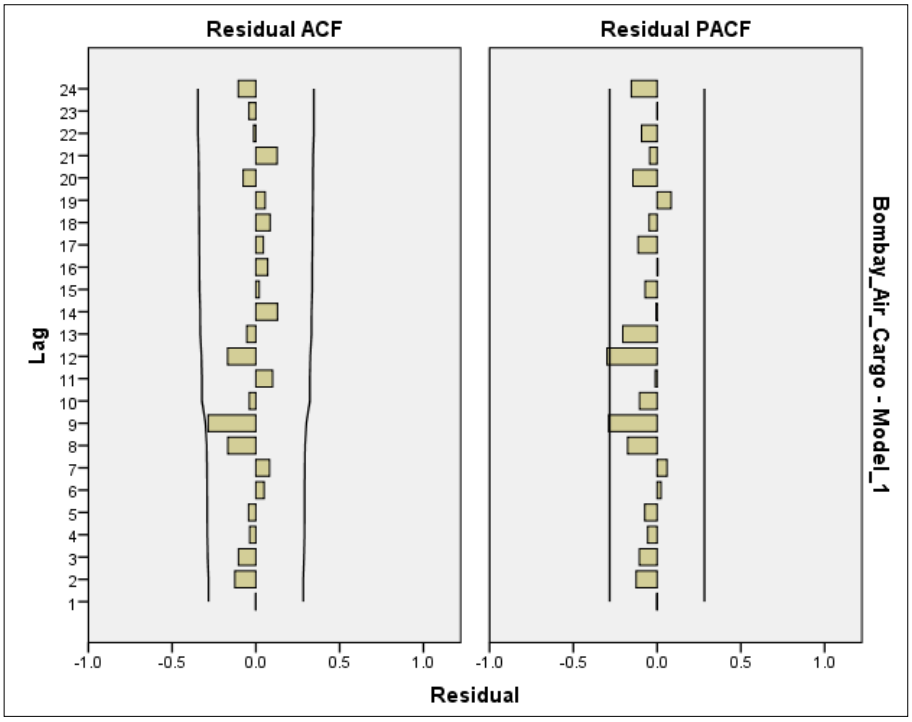
Now, on the basis of application of ARIMA model, we have two important points with us.

1 Results of first and second order averages and those derived from application of ANOVA are shown in above

tables which agree with reality. Also an important feature-forecast, over a period of next 10 months is predicted. We have ARIMA model forecast table shown below.

Forecast						
Model		**49**	**50**	**51**	**52**	**53**
Bombay Orange Export-Model_1	Forecast	171	171	171	171	171
	UCL	223	223	223	223	223
	LCL	118	118	118	118	118

There are two more tables but, close vicinity of data keeps on continuing same results. (We delete the further tables.)



Conclusion

The bifurcated program designed above has revealed important information pertaining to export of fruits and vegetables. Keeping one port fixed for export of two different fruits, calculated averages of first and second order for different years, helps comparison with that of the other port. On the other part application of ARIMA model and its extension gives a forecast which can be interpreted to plan business activities in coming months of the year.

References

1. Box GEP, Jenkins GM. Time series Analysis, Forecasting and Control. Holden Day, San Francisco, Revised Edition, 1960, 1976.
2. Box GEP, Pierce DA. Distribution of Residuals Autocorrelations in Autoregressive Integrated Moving Average Time series Models. J American Statistical Association. 1970; 65:1509-1526
3. Ljung GM, Box GEP. On a Measure of Lack of Fit in Time Series Models. Biometrika. 1978, 65.
4. Marquardt DW. An Algorithm for Least Squares estimation of Non – linear Parameters, SAIM. Applied Mathematics. 1963; 11:431-441.