Business Report

SMDM Project Business Report DSBA

***Sanjay Srinivasan***

***PGP-DSBA Online***

***JULY’ 21 Batch***

***Date: 20-02-2022***

# *INDEX*

|  |  |  |
| --- | --- | --- |
| ***S. No*** | ***Contents*** | ***Page* No** |
| **1.** | **Problem - 1** | 6 |
|  | Summary | 6 |
|  | Introduction | 6 |
|  | Sample Dataset | 6 |
|  | Exploratory Data Analysis | 6 |
|  | Checking for missing values in the dataset | 6 |
|  | 1) Analysis. Read the data as an appropriate Time Series data and plot the data. | 7 |
|  | 2) Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition. | 7 |
|  | 3) Split the data into training and test. The test data should start in 1991. | 12 |
|  | 4. Build all the exponential smoothing models on the training data and evaluate the model using RMSE on the test data. Other models such as regression,naïve forecast models and simple average models. should also be built on the training data and check the performance on the test data using RMSE. | 13 |
|  | 5. Check for the stationarity of the data on which the model is being built on using appropriate statistical tests and also mention the hypothesis for the statistical test. If the data is found to be non-stationary, take appropriate steps to make it stationary. Check the new data for stationarity and comment. Note: Stationarity should be checked at alpha = 0.05. | 25 |
|  | 6. Build an automated version of the ARIMA/SARIMA model in which the parameters are selected using the lowest Akaike Information Criteria (AIC) on the training data and evaluate this model on the test data using RMSE. | 29 |
|  | 7. Build ARIMA/SARIMA models based on the cut-off points of ACF and PACF on the training data and evaluate this model on the test data using RMSE. | 33 |
|  | 8. Build a table (create a data frame) with all the models built along with their corresponding parameters and the respective RMSE values on the test data. | 37 |
|  | 9. Based on the model-building exercise, build the most optimum model(s) on the complete data and predict 12 months into the future with appropriate confidence intervals/bands. | 38 |
|  | 10. Comment on the model thus built and report your findings and suggest the measures that the company should be taking for future sales. | 40 |

***List Of Figures***

|  |  |  |
| --- | --- | --- |
| ***S.No*** | ***Content*** | ***Page No*** |
| 1.1 | Dataset Sample | 6 |
| 1.2 | Datatypes of the variable | 6 |
| 1.3.1 | Check null values | 6 |
| 1.3.2 | Checking NaN values | 6 |
| 1.4 | Initialising Date as index Column | 7 |
| 1.5 | Plotting Rose wine data | 7 |
| 1.6 | Rose wine sales data spread. | 7 |
| 1.7 | Rose wine sales across years | 8 |
| 1.8 | Rose wine sales across months | 8 |
| 1.9 | Monthwise wine sales across years | 9 |
| 1.10 | Monthwise wine sales across years | 9 |
| 1.11 | Decompose data form the original dataset for Rose wines (Additive model) | 10 |
| 1.12 | Trend, Seasonality and residual values after decomposing the original data for Rose wines (Additive model) | 10 |
| 1.13 | Decompose data form the original dataset for Rose wines (Multiplicative model) | 11 |
| 1.14 | Trend, Seasonality and residual values after decomposing the original data for Rose wines (Multiplicative model) | 11 |
| 1.15 | Last 5 values for Training data | 12 |
| 1.16 | First 5 values for testing data | 12 |
| 1.17 | Plotting Training and test dataset for Rose wines | 12 |
| 1.18.1 | Generating numerical time instance for both training and test dataset for Rose wines | 13 |
| 1.18.2 | Loading numerical time into the dataframe | 13 |
| 1.19 | Initializing Linear Regression method | 13 |
| 1.20 | Predicting for training dataset | 13 |
| 1.21 | Predicting for test dataset | 13 |
| 1.22 | Ploting original and predicted train and test datas using linear regression | 14 |
| 1.23 | RMSE and MAPE value Training data | 14 |
| 1.24 | RMSE and MAPE value Test data | 14 |
| 1.25 | Loading RMSE and MAPE value Test data into dataframe | 14 |
| 1.26 | Predicting values for training data using naïve approach | 14 |
| 1.27 | Predicting values for test data using naïve approach | 14 |
| 1.28 | Plotting the predicted values for train and test data using naïve approach | 15 |
| 1.29 | RMSE and MAPE value Training data | 15 |
| 1.30 | RMSE and MAPE value Test data | 15 |
| 1.31 | Loading RMSE and MAPE value Test data into dataframe | 15 |
| 1.32 | Taking mean of Rose wine sales training data | 15 |
| 1.33 | Taking mean of Rose wine sales test data | 16 |
| 1.34 | Plotting the Simple Average data, train and test data | 16 |
| 1.35 | RMSE and MAPE value Training data | 16 |
| 1.36 | RMSE and MAPE value Test data | 16 |
| 1.37 | Loading RMSE and MAPE value Test data into dataframe | 16 |
| 1.38 | training data | 17 |
| 1.39 | Making data from training data for moving average | 17 |
| 1.40 | Plotting Moving average data for rolling 2 point | 17 |
| 1.41 | Plotting Moving average data for rolling 2,4,6,9 point | 17 |
| 1.42 | Plotting Moving average data for rolling 2 point for training and test data | 18 |
| 1.43 | Plotting Moving average data for rolling 4 point for training and test data | 18 |
| 1.44 | Plotting Moving average data for rolling 6,9 point for training and test data | 18 |
| 1.45 | load Moving average data for rolling 2,4,6,9 into the dataframe | 19 |
| 1.46 | RMSE and MAPE Moving average data for rolling 2,4,6,9 | 19 |
| 1.47 | Loading MA data for rolling 2,4,6,9 of RMSE and MAPE value Test data into dataframe | 19 |
| 1.48 | Intializing the Simple Exponential Smoothing | 19 |
| 1.49 | predicting values for the Simple Exponential Smoothing | 20 |
| 1.50 | Plotting predicted values for the Simple Exponential Smoothing | 20 |
| 1.51 | RMSE and MAPE value for the training data using Simple Exponential Smoothing | 20 |
| 1.52 | RMSE and MAPE value for the test data using Simple Exponential Smoothing | 20 |
| 1.53 | RMSE and MAPE value for the test data using Simple Exponential Smoothing | 20 |
| 1.54 | Initializing the Double Exponential Smoothing | 21 |
| 1.55 | Predicting the values using Double Exponential Smoothing | 21 |
| 1.56 | RMSE and MAPE for training data | 21 |
| 1.57 | RMSE and MAPE for test data | 21 |
| 1.58 | RMSE and MAPE for test data into the dataframe | 21 |
| 1.59 | plotting DES predicted output | 22 |
| 1.60 | Finding RMSE and MAPE for different Alpha and beta values | 22 |
| 1.61 | plotting for least RMSE value of different alpha n beta values | 22 |
| 1.62 | Finding the least RMSE and MAPE value from the different ALPHA and BETA values | 23 |
| 1.63 | Initializing the TES | 23 |
| 1.64 | Predicting the values from training data | 23 |
| 1.65 | Predicting the values from test data | 23 |
| 1.66 | Plotting predicted train and test values | 24 |
| 1.67 | RMSE and MAPE values from train data | 24 |
| 1.68 | RMSE and MAPE values from test data | 24 |
| 1.69 | RMSE and MAPE for test data into the dataframe | 24 |
| 1.70 | RMSE and MAPE for different alpha, beta, and gamma values into the dataframe | 24 |
| 1.71 | Finding least RMSE and MAPE for different alpha, beta, and gamma values and loading into the dataframe | 25 |
| 1.72 | Finding the p-value and plotting Rolling mean and standard deviation with original data. | 25 |
| 1.73 | Finding the p-value and plotting Rolling mean and standard deviation with original data after taking 1st difference data. | 26 |
| 1.74 | ACF plot for original data | 26 |
| 1.75 | ACF plot for 1st difference data | 27 |
| 1.76 | PACF plot for original data | 27 |
| 1.77 | PACF plot for 1st difference data | 28 |
| 1.78 | Finding stationarity and p-value for train data | 28 |
| 1.79 | Finding stationarity and p-value for train data after taking 1st difference | 29 |
| 1.80 | Finding combination for the model | 29 |
| 1.81 | Sample AIC value for ARIMA Model | 30 |
| 1.82 | Sorting least AIC value for ARIMA Model | 30 |
| 1.83 | Summary report for ARIMA Model | 30 |
| 1.84 | RMSE and MAPE for ARIMA Model | 30 |
| 1.85 | RMSE and MAPE for ARIMA Model into the dataframe | 31 |
| 1.86 | combination parameters for SARIMA Model | 31 |
| 1.87 | Sample AIC for SARIMA Model | 31 |
| 1.88 | Sample AIC for SARIMA Model | 32 |
| 1.89 | Summary report for SARIMA Model | 32 |
| 1.90 | plotting diagnostic for SARIMA Model | 33 |
| 1.91 | Predicting values for testing data | 33 |
| 1.92 | RMSE and MAPE SARIMA Model | 33 |
| 1.93 | Loading RMSE and MAPE SARIMA Model into dataframe | 34 |
| 1.94 | Plotting 2nd Difference after taking cut-off points in ACF plot | 34 |
| 1.95 | Plotting 2nd Difference after taking cut-off points in PACF plot | 35 |
| 1.96 | Cut-off point summary report for ARIMA model | 35 |
| 1.97 | predicting values for ARIMA model after taking cutoff points | 36 |
| 1.98 | RMSE and MAPE value for cutoff point ARIMA model | 36 |
| 1.99 | Cut-off point for ARIMA model RMSE and MAPE into Dataframe | 36 |
| 1.100 | Cut-off point for SARIMA model summary report | 37 |
| 1.101 | Cut-off point for SARIMA model RMSE and MAPE value | 37 |
| 1.102 | Cut-off point for SARIMA model RMSE and MAPE into Dataframe | 38 |
| 1.103 | Sorting the RMSE value and finding best model | 38 |
| 1.104 | Summary report for SARIMA model | 39 |
| 1.105 | plotting diagnostic report for SARIMA model | 39 |
| 1.106 | plotting diagnostic report for SARIMA model | 39 |
| 1.107 | Predicting the values | 40 |
| 1.108 | RMSE value for future data. | 40 |
| 1.109 | Plotting the future data in graph. | 40 |

# Problem - 1

*Summary*

The data is gathered from ABC Estate wines, for Rose wine sales data from this ABC Estate wine company. An analyst for the company needs to analyze the wine sales in the 19th century and forecast the wine sales for the 20th century.

*Introduction*

The purpose of this exercise is to explore the dataset and make the analyze the wine sales in the 19th Century, based on the sales data we need to forecast for the wine sales data for next 12 months.

*Sample of the dataset:*

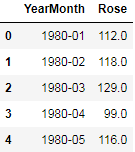


Fig 1.1 Dataset Sample

*Exploratory Data Analysis*

*Let us check the types of variables in the data frame.*



Fig- 1.2. Datatypes of the variable

There are total 187 rows and 2 columns in the dataset. 1 columns are object and 1 columns are int64

*Check for missing values in the dataset:*

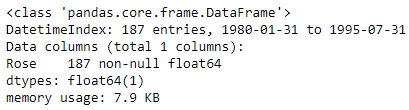


Fig- 1.3.1. Check null values

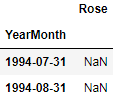


Fig- 1.3.2. Checking NaN values

**1. Analysis.** **Read the data as an appropriate Time Series data and plot the data.**

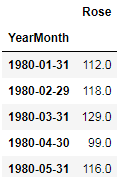


Fig- 1.4. Initialising Date as index Column

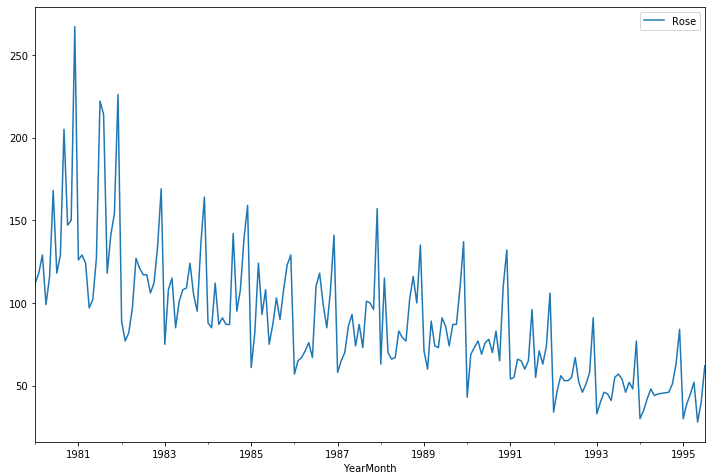


Fig – 1.5 Plotting Rose wine data

The Rose wine sales data has been plotted against the year of sales.

### 2. Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.

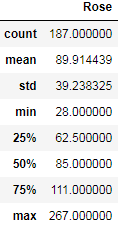


Fig – 1.6 Rose wine sales data spread.

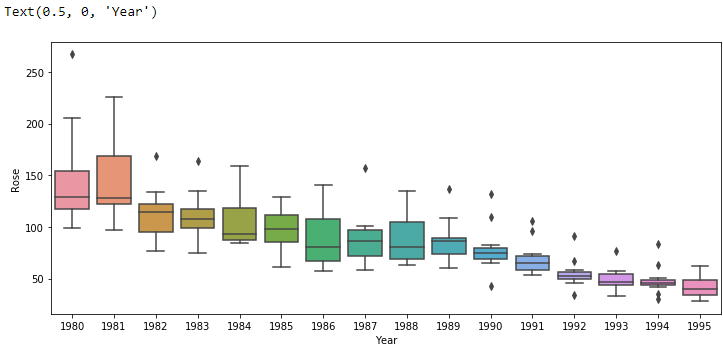


Fig – 1.7 Rose wine sales across years

From the above chart (boxplot), there are outliers present in the data and we can observe that there was good sales record for Rose wine from 1980-1994 and wine sales has been decreased in the year 1995.

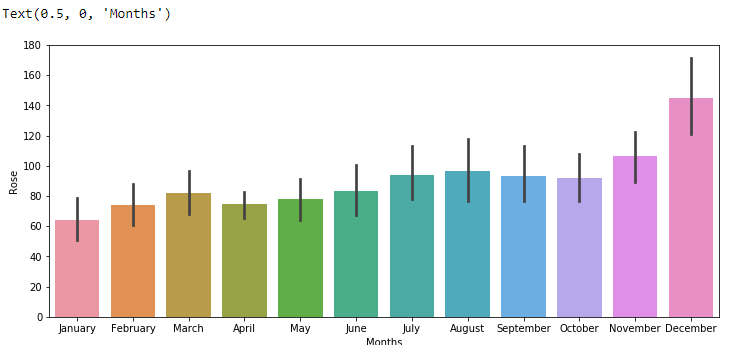


Fig – 1.8 Rose wine sales across months

From the above chart (boxplot), the December month has highest number of Rose wine sales when compared with other months.

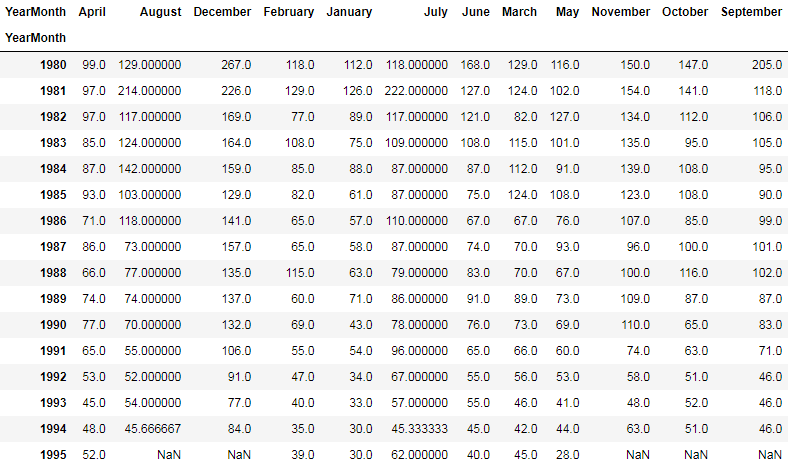


Fig – 1.9 Monthwise wine sales across years

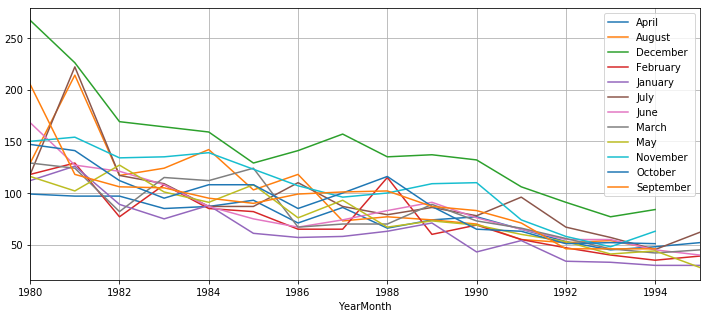


Fig – 1.10 Monthwise wine sales across years

***Additive Model:***

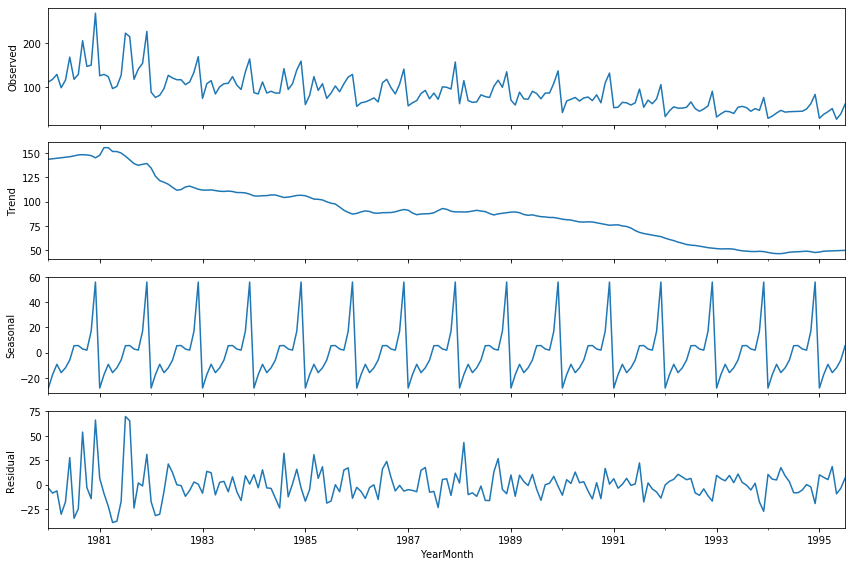


Fig – 1.11 Decompose data form the original dataset for Rose wines (Additive model)

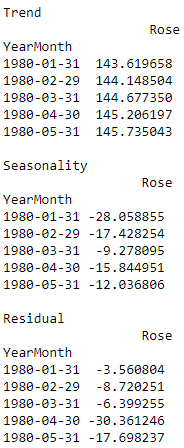


Fig – 1.12 Trend, Seasonality and residual values after decomposing the original data for Rose wines (Additive model)

***Multiplicative Model:***

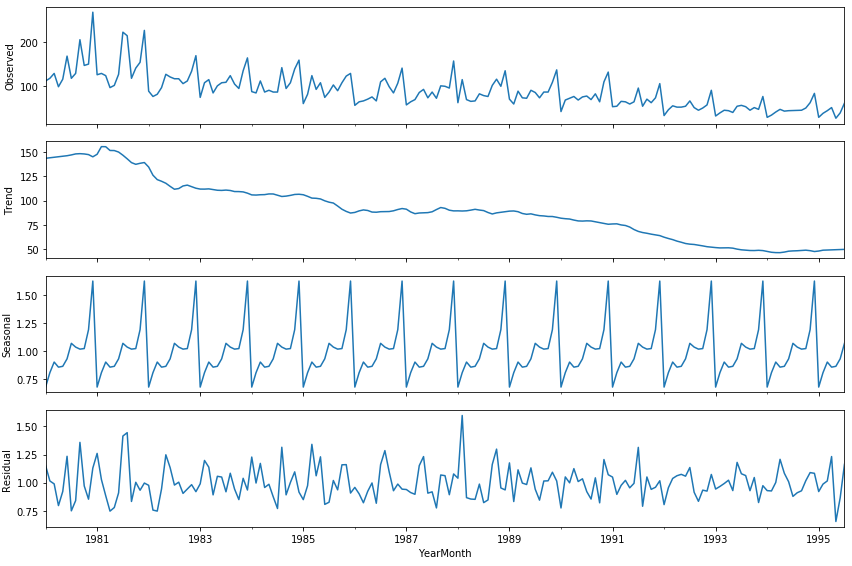


Fig – 1.13 Decompose data form the original dataset for Rose wines (Multiplicative model)

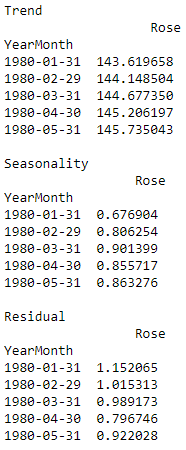


Fig – 1.14 Trend, Seasonality and residual values after decomposing the original data for Rose wines (Multiplicative model)

### 3. Split the data into training and test. The test data should start in 1991

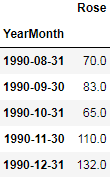


Fig – 1.15 Last 5 values for Training data

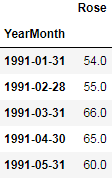


Fig – 1.16 First 5 values for testing data

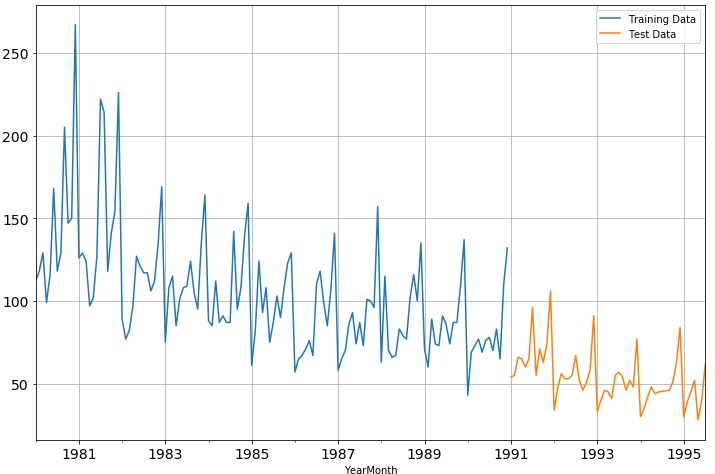


Fig – 1.17 Plotting Training and test dataset for Rose wines

### 4. Build all the exponential smoothing models on the training data and evaluate the model using RMSE on the test data. Other additional models such as regression, naïve forecast models, simple average models, moving average models should also be built on the training data and check the performance on the test data using RMSE.

### *Model 1: Linear Regression:*

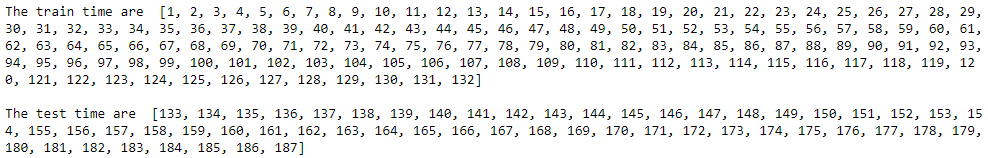


Fig – 1.18.1 Generating numerical time instance for both training and test dataset for Rose wines

We see that we have successfully the generated the numerical time instance order for both the training and test set. Now we will add these values in the training and test set.

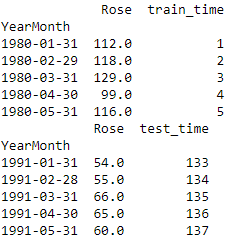


Fig – 1.18.2 Loading numerical time into the dataframe



Fig – 1.19 Initializing Linear Regression method

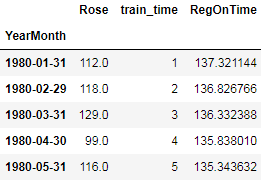


Fig – 1.20 Predicting for training dataset

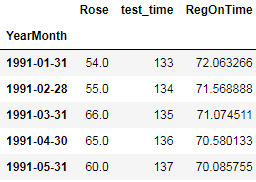


Fig – 1.21 Predicting for test dataset

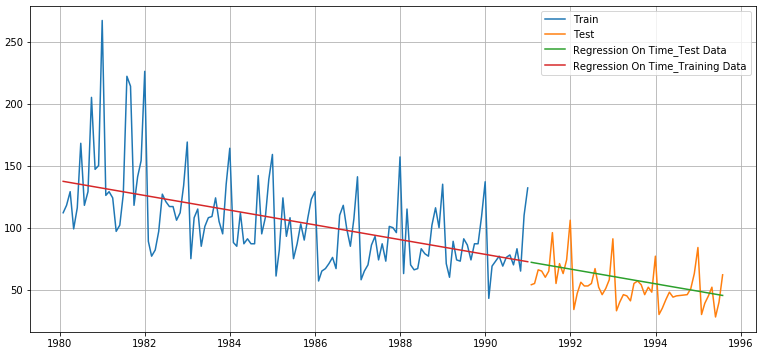


Fig – 1.22 Ploting original and predicted train and test datas using linear regression



Fig – 1.23 RMSE and MAPE value Training data



Fig – 1.24 RMSE and MAPE value Test data



Fig – 1.25 Loading RMSE and MAPE value Test data into dataframe

### *Model - 2: Naive Approach:* (ŷ𝑡+1 = yt)

#### For this particular naive model, we say that the prediction for tomorrow is the same as today and the prediction for day after tomorrow is tomorrow and since the prediction of tomorrow is same as today,therefore the prediction for day after tomorrow is also today.

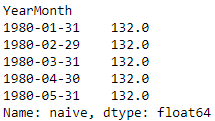


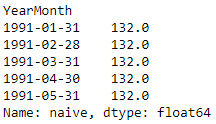
Fig – 1.26 Predicting values for training data using naïve approach  


Fig – 1.27 Predicting values for test data using naïve approach

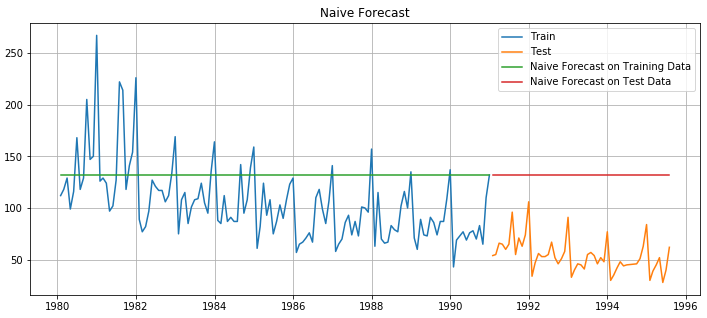


Fig – 1.28 Plotting the predicted values for train and test data using naïve approach



Fig – 1.29 RMSE and MAPE value Training data



Fig – 1.30 RMSE and MAPE value Test data

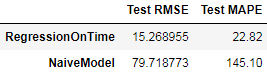


Fig – 1.31 Loading RMSE and MAPE value Test data into dataframe

***Model – 3 Simple Average:***

For this particular simple average method, we will forecast by using the average of the training values

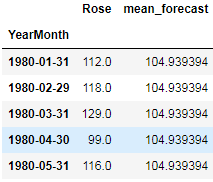


Fig – 1.32 Taking mean of Rose wine sales training data

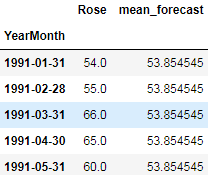


Fig – 1.33 Taking mean of Rose wine sales test data

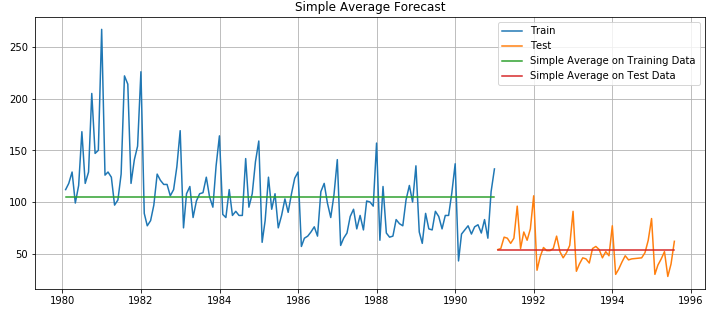


Fig – 1.34 Plotting the Simple Average data, train and test data



Fig – 1.35 RMSE and MAPE value Training data



Fig – 1.36 RMSE and MAPE value Test data

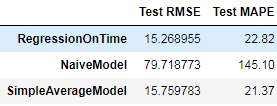


Fig – 1.37 Loading RMSE and MAPE value Test data into dataframe

***Model – 4 Moving Average(MA):***

For the moving average model, we are going to calculate rolling means (or moving averages) for different intervals. The best interval can be determined by the maximum accuracy (or the minimum error) over here.

For Moving Average, we are going to average over the entire data.

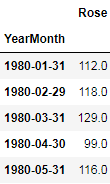


Fig – 1.38 training data

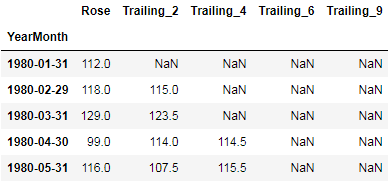


Fig – 1.39 Making data from training data for moving average

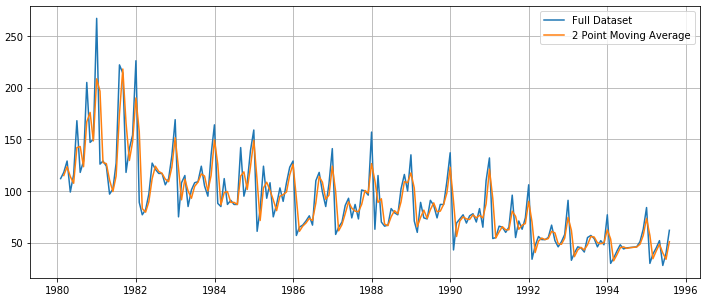


Fig – 1.40 Plotting Moving average data for rolling 2 point

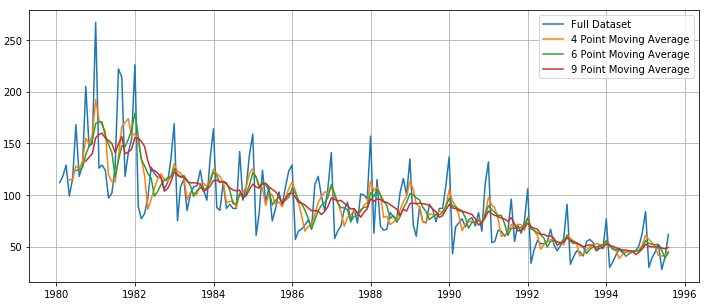


Fig – 1.41 Plotting Moving average data for rolling 2,4,6,9 point

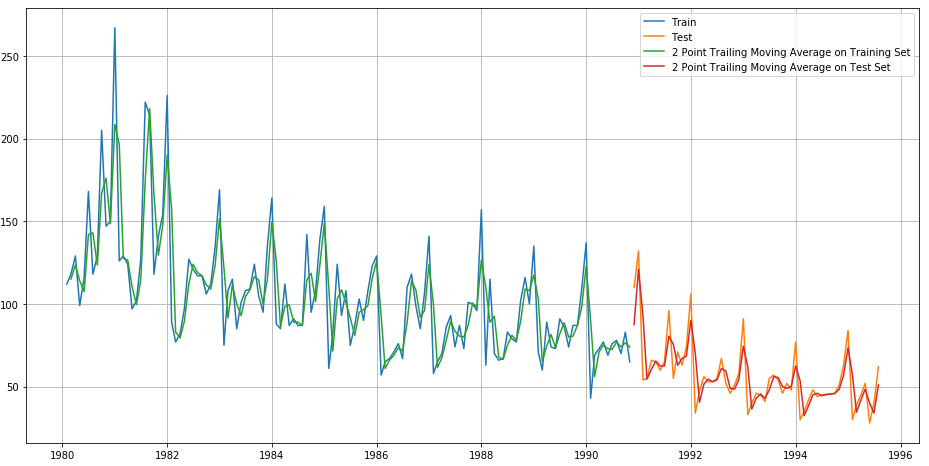


Fig – 1.42 Plotting Moving average data for rolling 2 point for training and test data

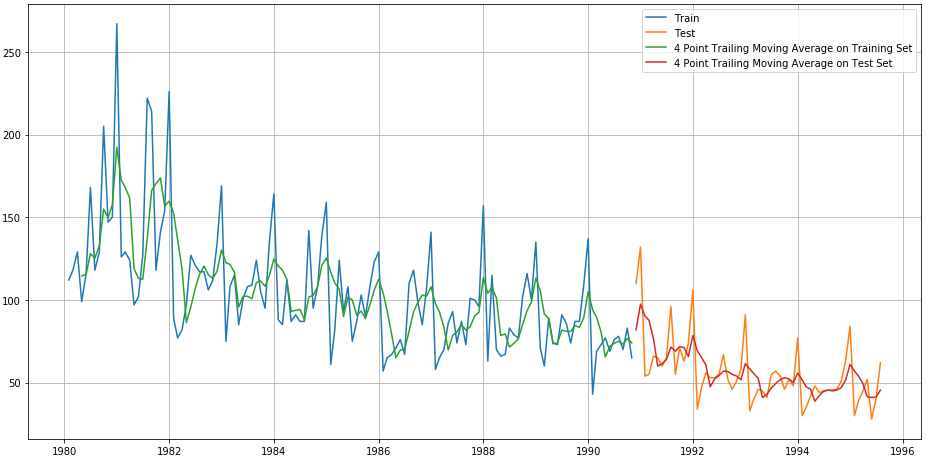


Fig – 1.43 Plotting Moving average data for rolling 4 point for training and test data

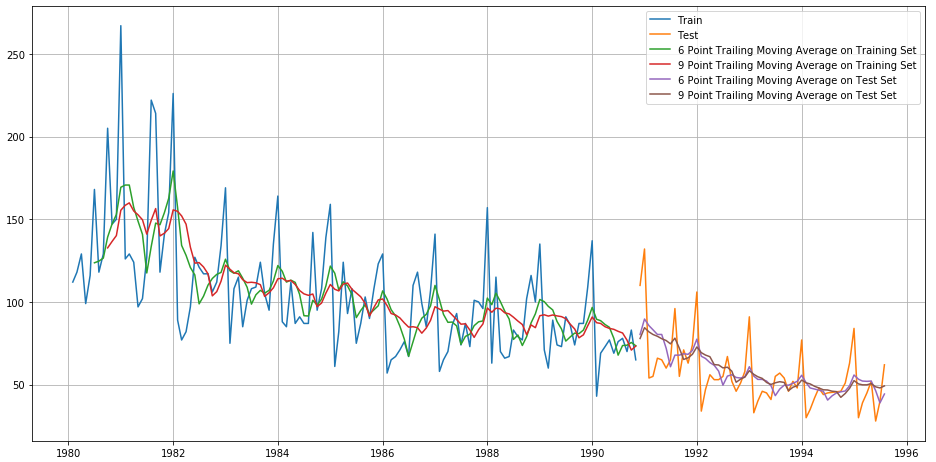


Fig – 1.44 Plotting Moving average data for rolling 6,9 point for training and test data



Fig – 1.45 load Moving average data for rolling 2,4,6,9 into the dataframe

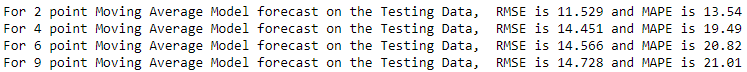


Fig – 1.46 RMSE and MAPE Moving average data for rolling 2,4,6,9

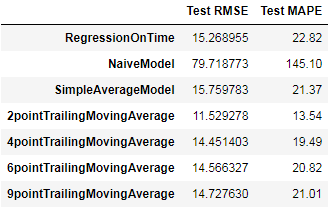


Fig – 1.47 Loading MA data for rolling 2,4,6,9 of RMSE and MAPE value Test data into dataframe

***Model – 5 Simple Exponential Smoothing:***

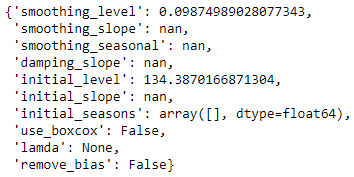
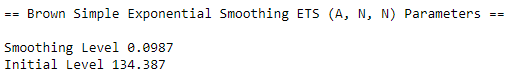


Fig – 1.48 Intializing the Simple Exponential Smoothing



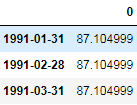


Fig – 1.49 predicting values for the Simple Exponential Smoothing

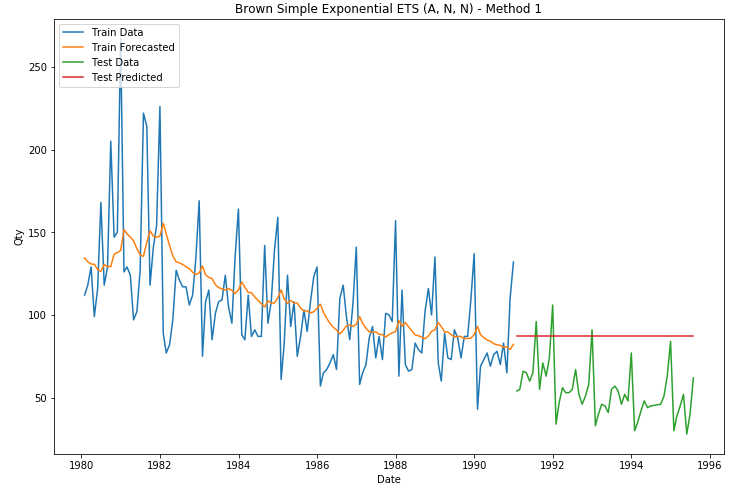


Fig – 1.50 Plotting predicted values for the Simple Exponential Smoothing



Fig – 1.51 RMSE and MAPE value for the training data using Simple Exponential Smoothing



Fig – 1.52 RMSE and MAPE value for the test data using Simple Exponential Smoothing

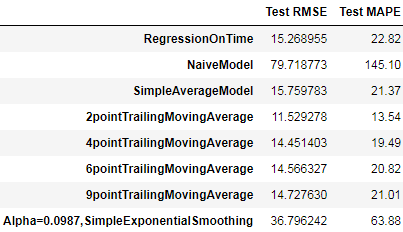


Fig – 1.53 RMSE and MAPE value for the test data using Simple Exponential Smoothing

### *Model 6: Double Exponential Smoothing (Holt's Model):*

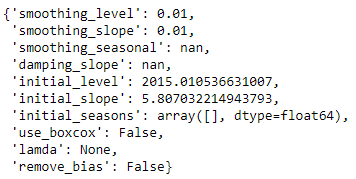
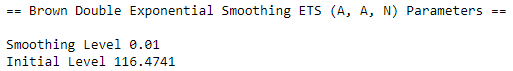


Fig – 1.54 Initializing the Double Exponential Smoothing



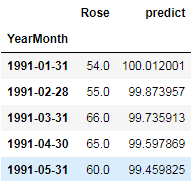


Fig – 1.55 Predicting the values using Double Exponential Smoothing



Fig – 1.56 RMSE and MAPE for training data



Fig – 1.57 RMSE and MAPE for test data

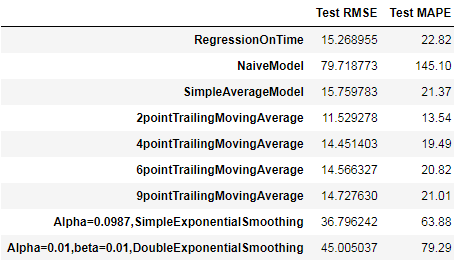


Fig – 1.58 RMSE and MAPE for test data into the dataframe

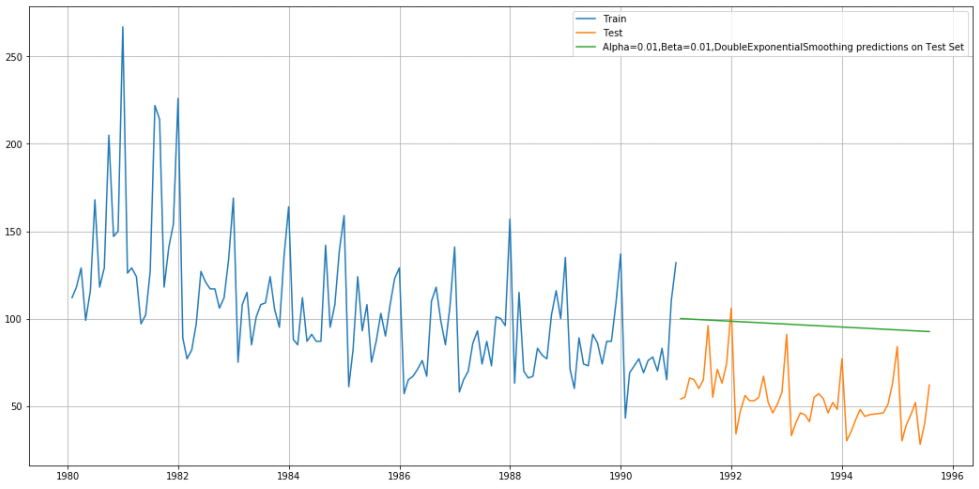


Fig – 1.59 plotting DES predicted output

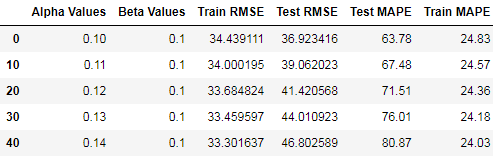


Fig – 1.60 Finding RMSE and MAPE for different Alpha and beta values

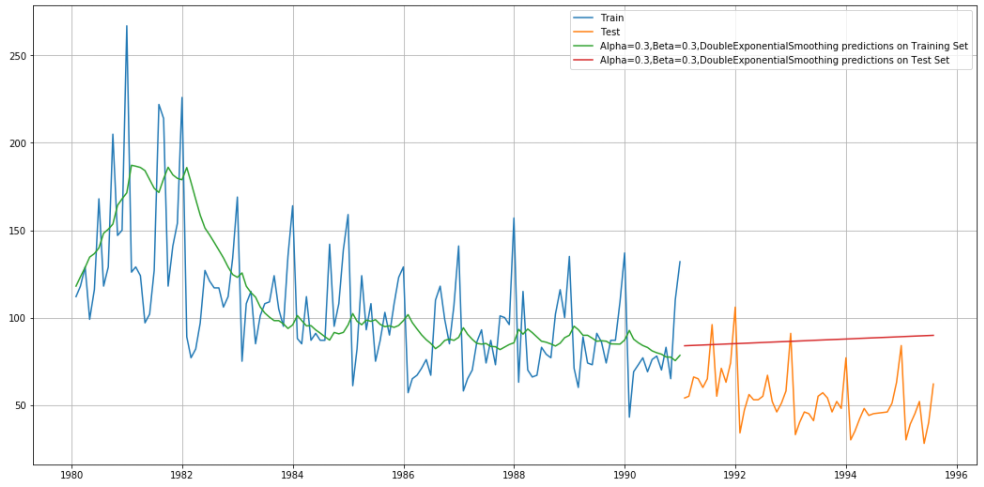


Fig – 1.61 plotting for least RMSE value of different alpha n beta values

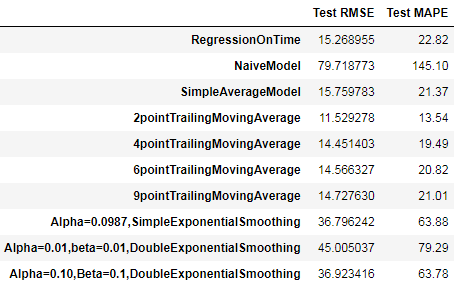


Fig 1.62 Finding the least RMSE and MAPE value from the different ALPHA and BETA values

### *Model 7: Triple Exponential Smoothing (Holt - Winter's Model)*

Three parameters , 𝛽 and 𝛾 are estimated in this model. Level, Trend and Seasonality are accounted for in this model.

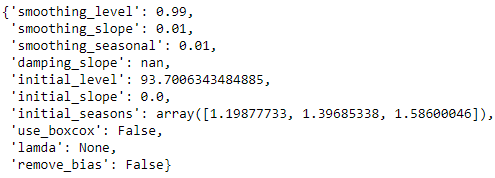


Fig 1.63 Initializing the TES

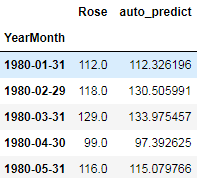


Fig 1.64 Predicting the values from training data

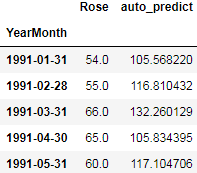


Fig 1.65 Predicting the values from test data

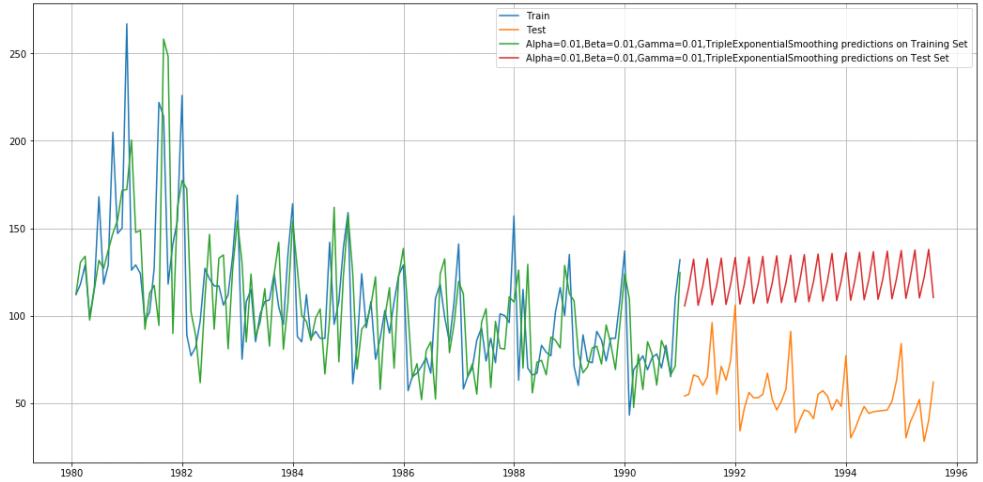


Fig 1.66 Plotting predicted train and test values



Fig 1.67 RMSE and MAPE values from train data



Fig 1.68 RMSE and MAPE values from test data

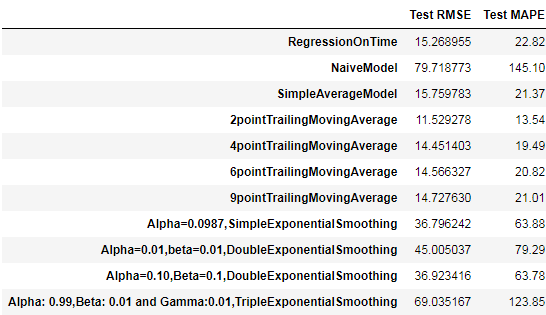


Fig 1.69 RMSE and MAPE for test data into the dataframe

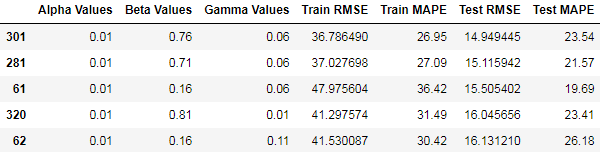


Fig 1.70 RMSE and MAPE for different alpha, beta, and gamma values into the dataframe

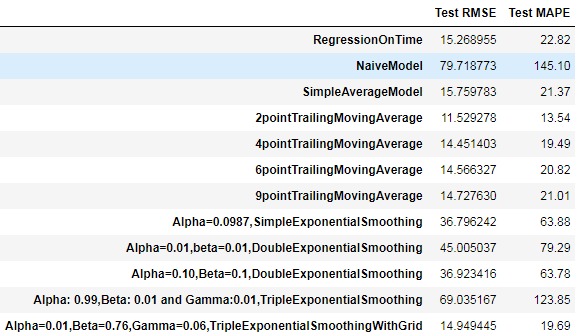


Fig 1.71 Finding least RMSE and MAPE for different alpha, beta, and gamma values and loading into the dataframe

### 5. Check for the stationarity of the data on which the model is being built on using appropriate statistical tests and also mention the hypothesis for the statistical test. If the data is found to be non-stationary, take appropriate steps to make it stationary. Check the new data for stationarity and comment.

***Note:*** Stationarity should be checked at alpha(𝛼) = 0.05.

###### **Null Hypothesis:** p-value > 𝛼 (alpha value) - then the data is not stationary

###### **Alternate Hypothesis:** p-value < 𝛼 (alpha value) - then the data is having stationarity

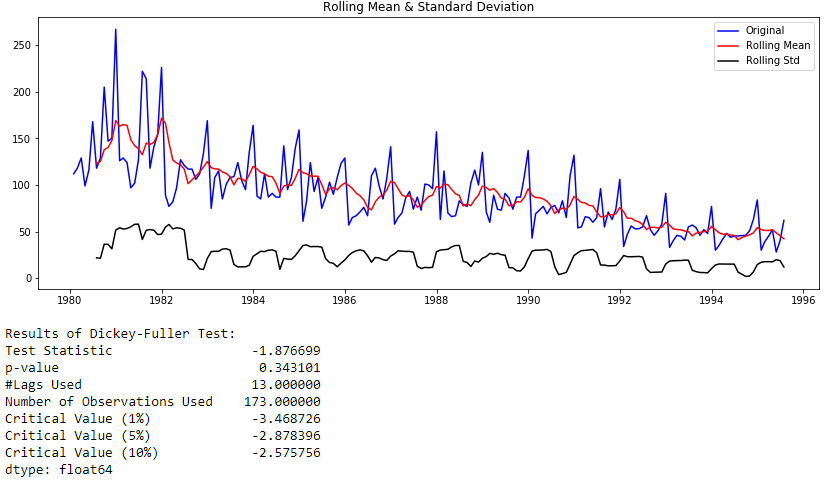


Fig 1.72 Finding the p-value and plotting Rolling mean and standard deviation with original data.

The p-value is greater than the alpha value so the data is not stationarity and alternate hypothesis is rejected. To find the stationarity we need to take the 1st difference and plotting the graph and finding the p-value.

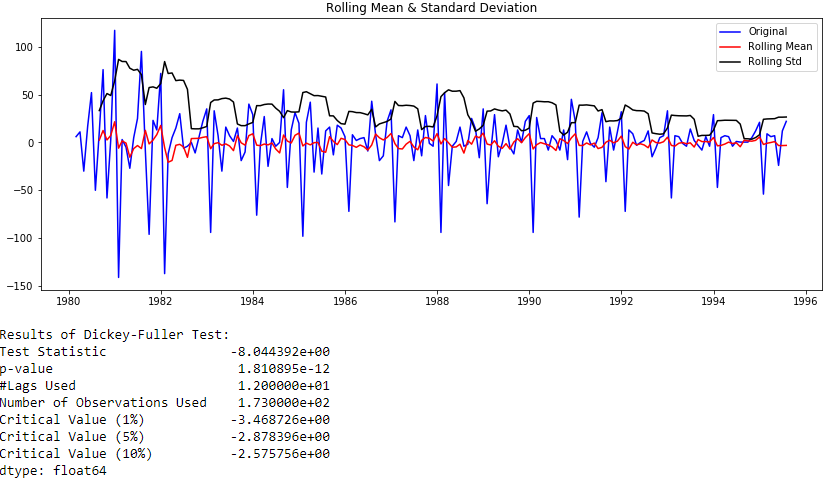


Fig 1.73 Finding the p-value and plotting Rolling mean and standard deviation with original data after taking 1st difference data.

Now p-value is less than alpha value. Therefore null hypothesis is rejected.

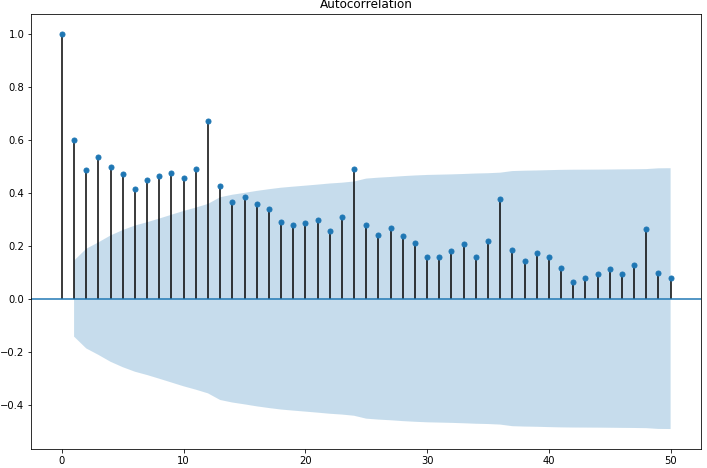


Fig 1.74 ACF plot for original data

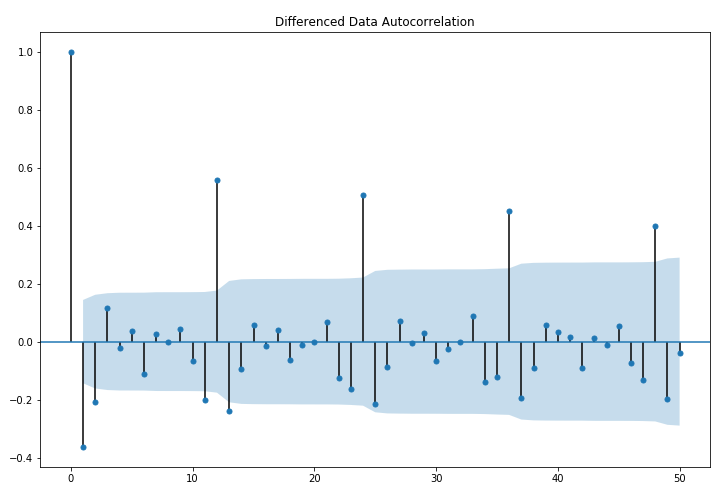


Fig 1.75 ACF plot for 1st difference data

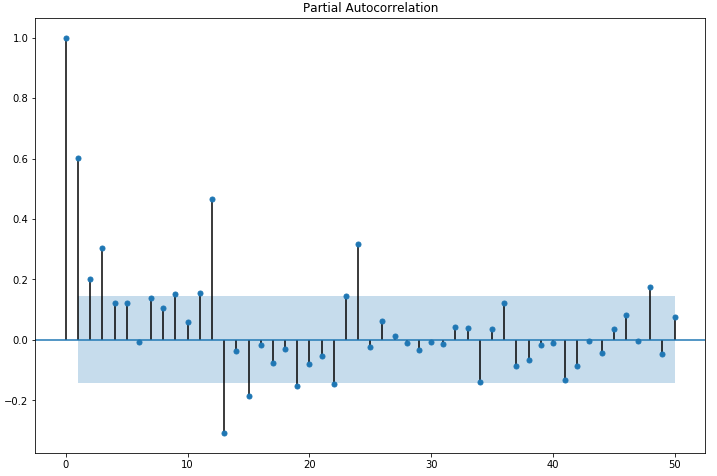


Fig 1.76 PACF plot for original data

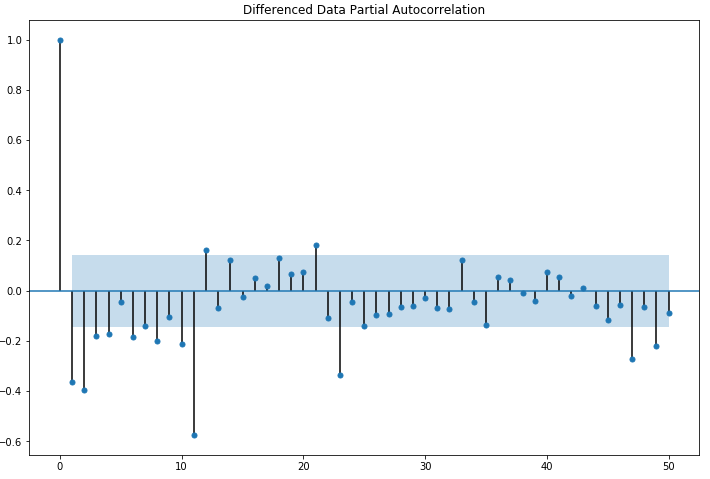


Fig 1.77 PACF plot for 1st difference data

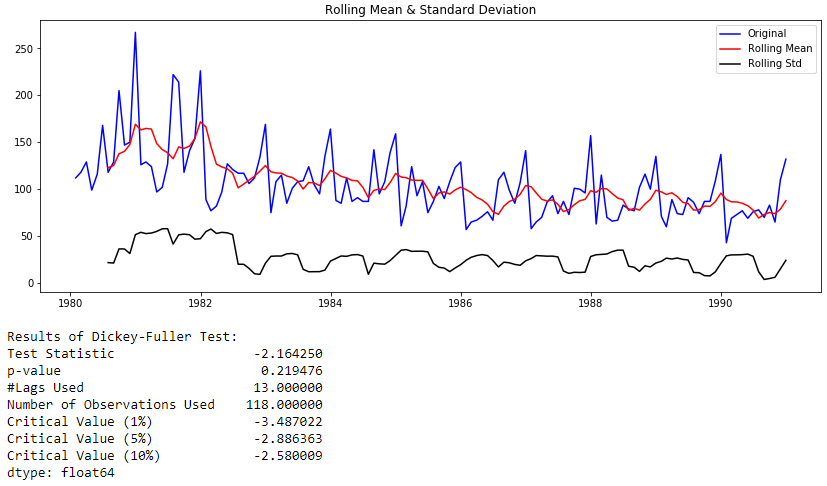


Fig 1.78 Finding stationarity and p-value for train data

The p-value is greater than the alpha value so the data is not stationarity and alternate hypothesis is rejected. To find the stationarity we need to take the 1st difference and plotting the graph and finding the p-value.

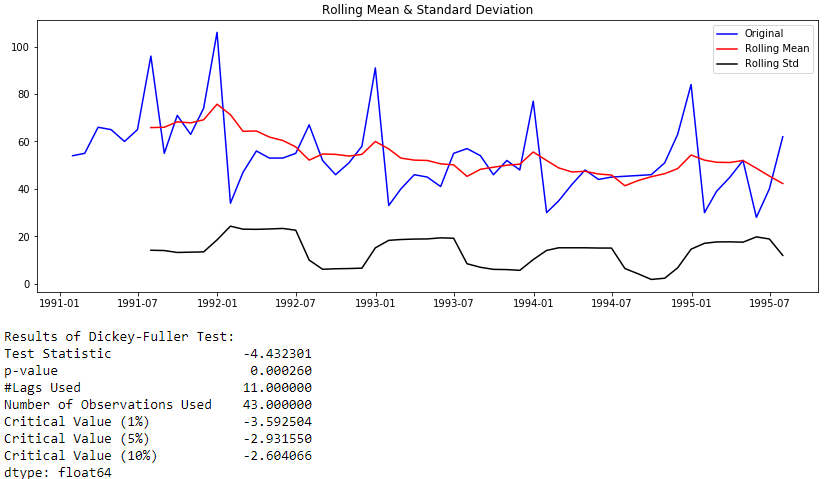


Fig 1.79 Finding stationarity and p-value for train data after taking 1st difference

### 6. Build an automated version of the ARIMA/SARIMA model in which the parameters are selected using the lowest Akaike Information Criteria (AIC) on the training data and evaluate this model on the test data using RMSE.

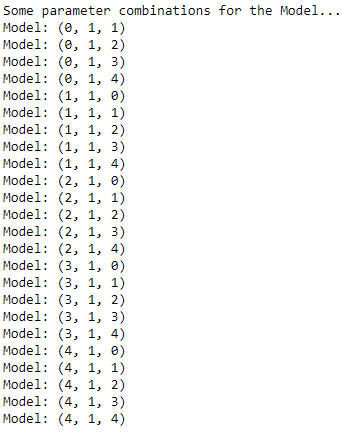


Fig 1.80 Finding combination for the model

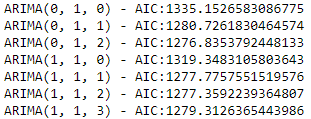


Fig 1.81 Sample AIC value for ARIMA Model

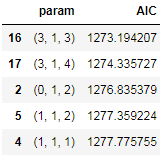


Fig 1.82 Sorting least AIC value for ARIMA Model

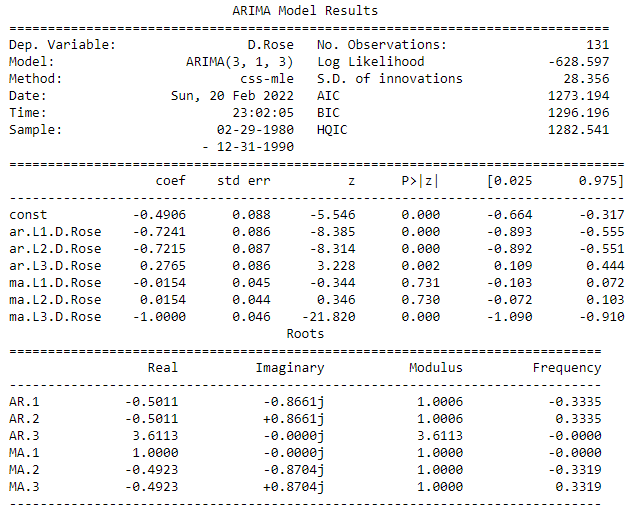


Fig 1.83 Summary report for ARIMA Model



Fig 1.84 RMSE and MAPE for ARIMA Model

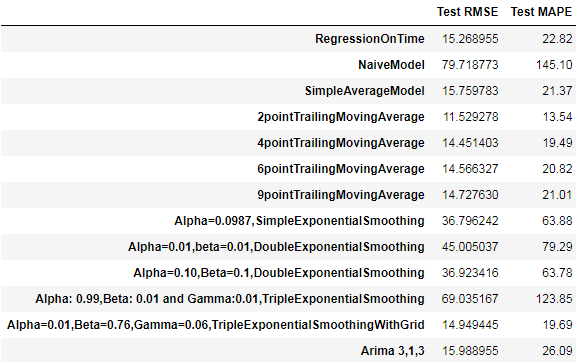


Fig 1.85 RMSE and MAPE for ARIMA Model into the dataframe

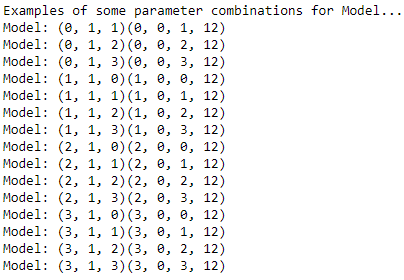


Fig 1.86 combination parameters for SARIMA Model

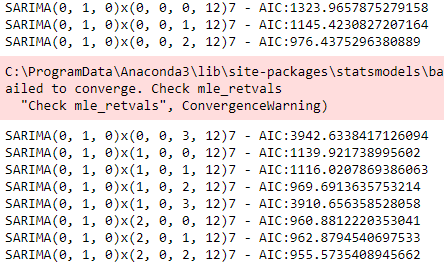


Fig 1.87 Sample AIC for SARIMA Model



Fig 1.88 Sample AIC for SARIMA Model

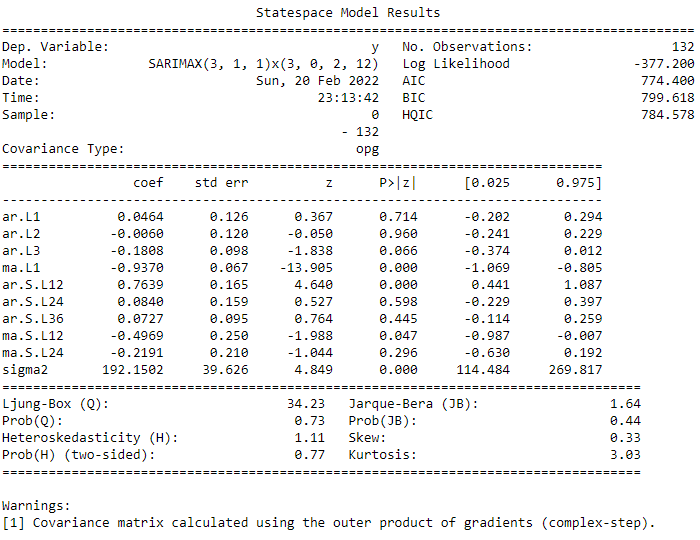


Fig 1.89 Summary report for SARIMA Model

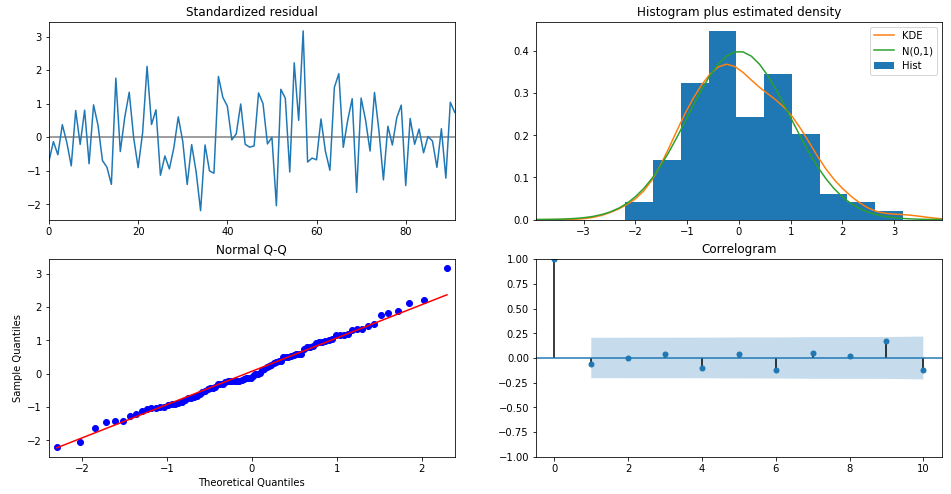


Fig 1.90 plotting diagnostic for SARIMA Model

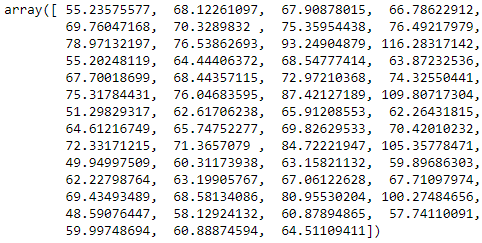


Fig 1.91 Predicting values for testing data



Fig 1.92 RMSE and MAPE SARIMA Model

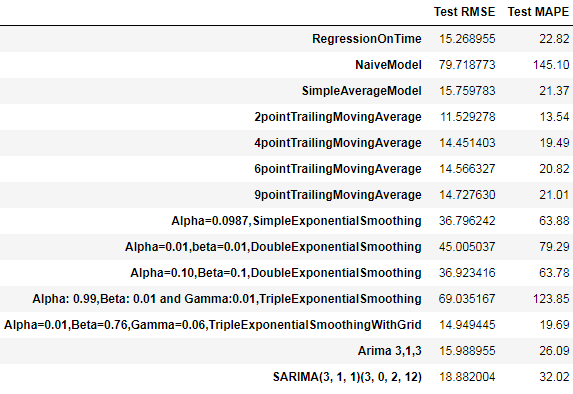


Fig 1.93 Loading RMSE and MAPE SARIMA Model into dataframe

### 7. Build ARIMA/SARIMA models based on the cut-off points of ACF and PACF on the training data and evaluate this model on the test data using RMSE.

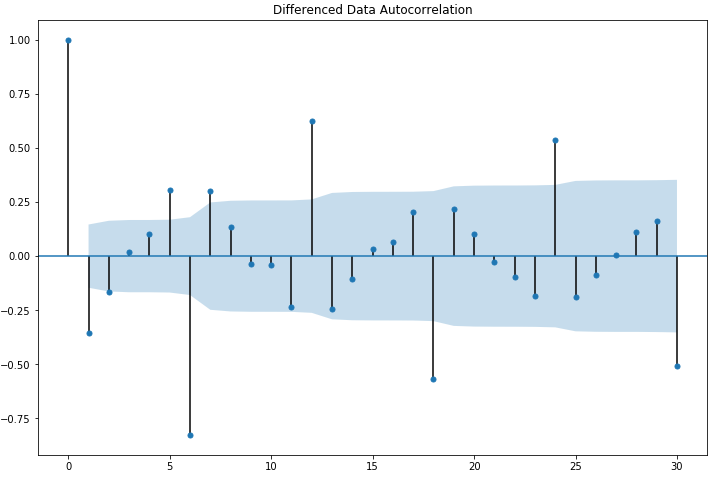


Fig 1.94 Plotting 2nd Difference after taking cut-off points in ACF plot

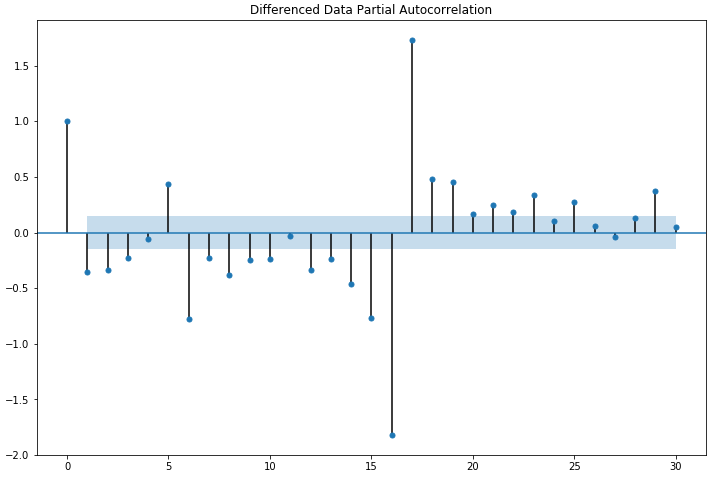


Fig 1.95 Plotting 2nd Difference after taking cut-off points in PACF plot

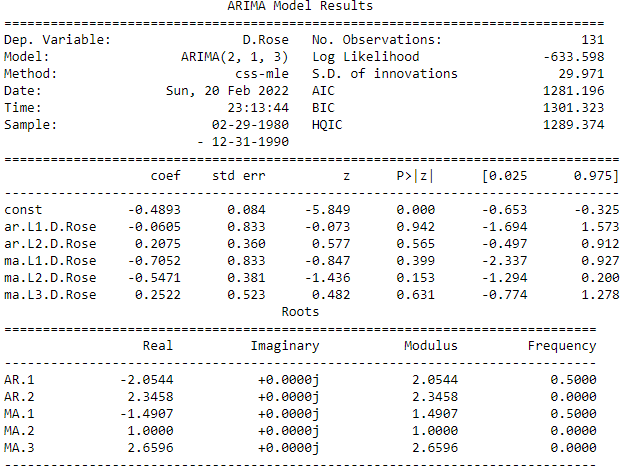


Fig 1.96 Cut-off point summary report for ARIMA model

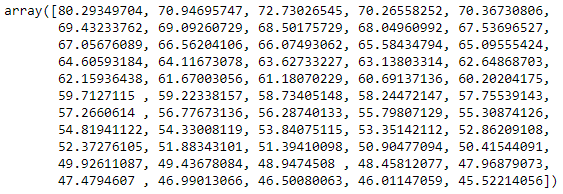


Fig 1.97 predicting values for ARIMA model after taking cutoff points



Fig 1.98 RMSE and MAPE value for cutoff point ARIMA model

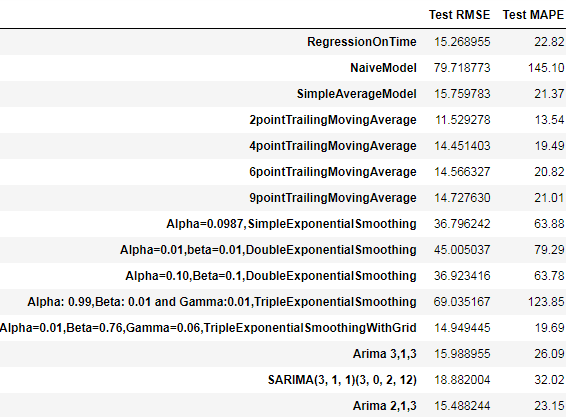


Fig 1.99 Cut-off point for ARIMA model RMSE and MAPE into Dataframe

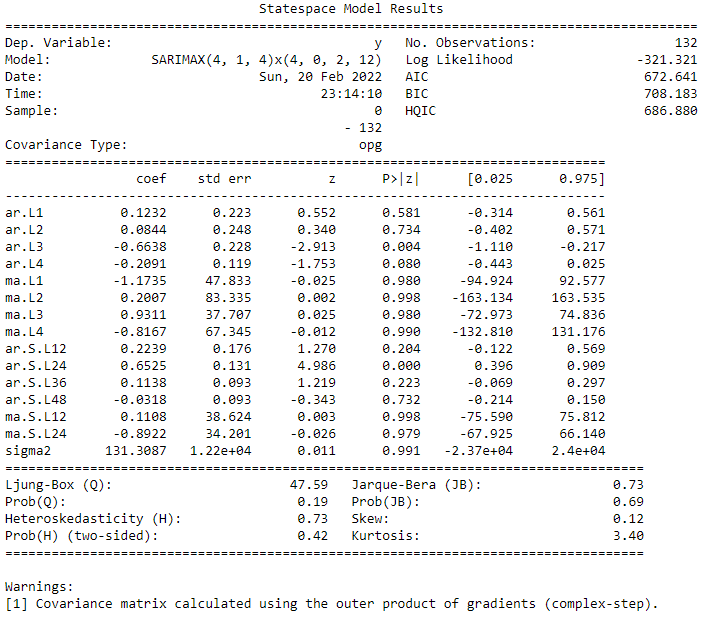


Fig 1.100 Cut-off point for SARIMA model summary report



Fig 1.101 Cut-off point for SARIMA model RMSE and MAPE value

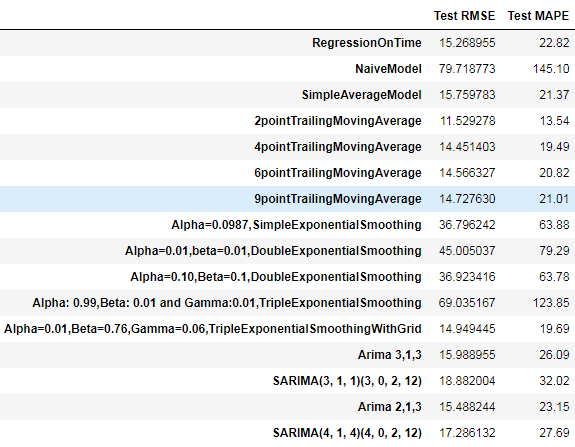


Fig 1.102 Cut-off point for ARIMA model RMSE and MAPE into Dataframe

### 8. Build a table (create a data frame) with all the models built along with their corresponding parameters and the respective RMSE values on the test data.

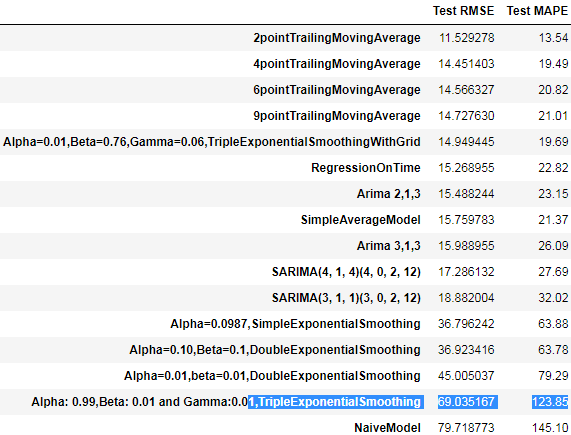


Fig 1.103 Sorting the RMSE value and finding best model

**9. Based on the model-building exercise, build the most optimum model(s) on the complete data and predict 12 months into the future with appropriate confidence intervals/bands.**

Although we are seeing that the best model as per RMSE is the 2pointTrailingMovingAverage, 4pointTrailingMovingAverage, 6pointTrailingMovingAverage, 9pointTrailingMovingAverage as it is giving us the least RMSE value. But the moving average models are actually quite a naive model and assumes that the trend and seasonality components of the time series have already been removed or adjusted for. Hence we will going to choose the second best model which comes out to be Triple Exponential Model. Its RMSE value is very close to the 2point Trailing Moving Average , 4point Trailing Moving Average, 6point Trailing Moving Average, 9point Trailing Moving Average value hence we can choose this model as well.



Fig 1.104 Summary report for SARIMA model

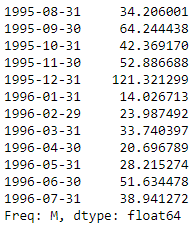


Fig 1.105 plotting diagnostic report for SARIMA model

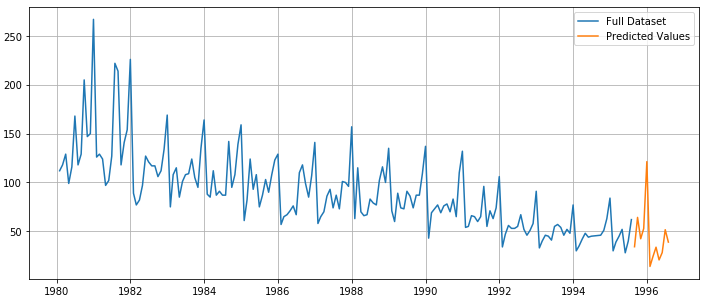


Fig 1.106 plotting diagnostic report for SARIMA model

***Predicting the values for the next 12 months in future***

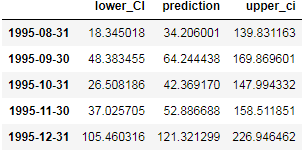


Fig 1.107 Predicting the values



Fig 1.108 RMSE value for future data.

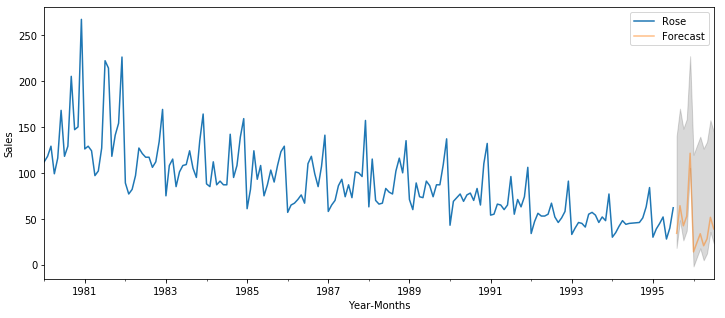


Fig 1.109 Plotting the future data in graph.

### 10. Comment on the model thus built and report your findings and suggest the measures that the company should be taking for future sales.

Triple Exponential Model is performing best in this case giving us the least error

Looking at the bar plot, we can see that on December months the sales are highest. We can use this insights to increase our sales further. We can introduce certain offers in other months to attract more customers.

Year 1981 has the highest sales recorded till data. We can go back to find out the reasons to which pushed the sales so much.

Looking at the prediction, we can say that the sales figure are declining each and every year.The reason behind this may be either deterioration in quality or arrival of any rival product in the market.Hence some important measures have to be taken to identify the reason and tale appropriate measures to increase the trend.