Project - Time Series Forecasting

AMAN TANDON

1.	**Read the data as an appropriate Time Series data and plot the data.**	
2.	**Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.**	
3.	**Split the data into training and test. The test data should start in 1991.**	
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.**	
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.**	
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.**	
7.	**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.**	
8.	**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.**	
9.	**Comment on the model thus built and report your findings and suggest the measures that the company should be taking for future sales.**	

Problem:

For this particular assignment, the data of different types of wine sales in the 20th century is to be analysed. Both of these data are from the same company but of different wines. As an analyst in the ABC Estate Wines, you are tasked to analyse and forecast Wine Sales in the 20th century.

Data set for the Problem: Sparkling.csv and Rose.csv.

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

Sparkling data set:

Head

YearMonth	Sparkling
1980-01-01	1686
1980-02-01	1591
1980-03-01	2304
1980-04-01	1712
1980-05-01	1471

Shape

```
1  Sparkling.shape
```

```
(187, 1)
```

Tail

YearMonth	Sparkling
1995-03-01	1897
1995-04-01	1862
1995-05-01	1670
1995-06-01	1688
1995-07-01	2031

5-point summary

	Sparkling
count	187.000000
mean	2402.417112
std	1295.111540
min	1070.000000
25%	1605.000000
50%	1874.000000
75%	2549.000000
max	7242.000000

Rose data set:

Head:

YearMonth	Rose
1980-01-01	112.0
1980-02-01	118.0
1980-03-01	129.0
1980-04-01	99.0
1980-05-01	116.0

Tail:

YearMonth	Rose
1995-03-01	45.0
1995-04-01	52.0
1995-05-01	28.0
1995-06-01	40.0
1995-07-01	62.0

Shape

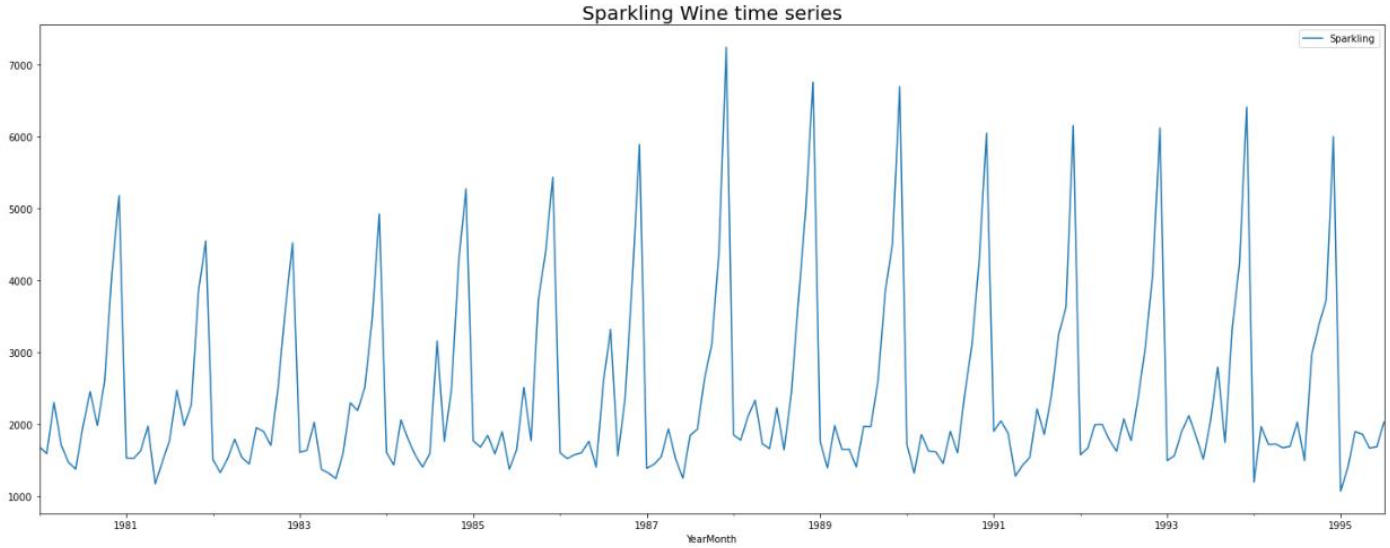
```
1 Rose.shape
```

```
(187, 1)
```

5- point summary

	Rose
count	187.000000
mean	89.914439
std	39.238325
min	28.000000
25%	62.500000
50%	85.000000
75%	111.000000
max	267.000000

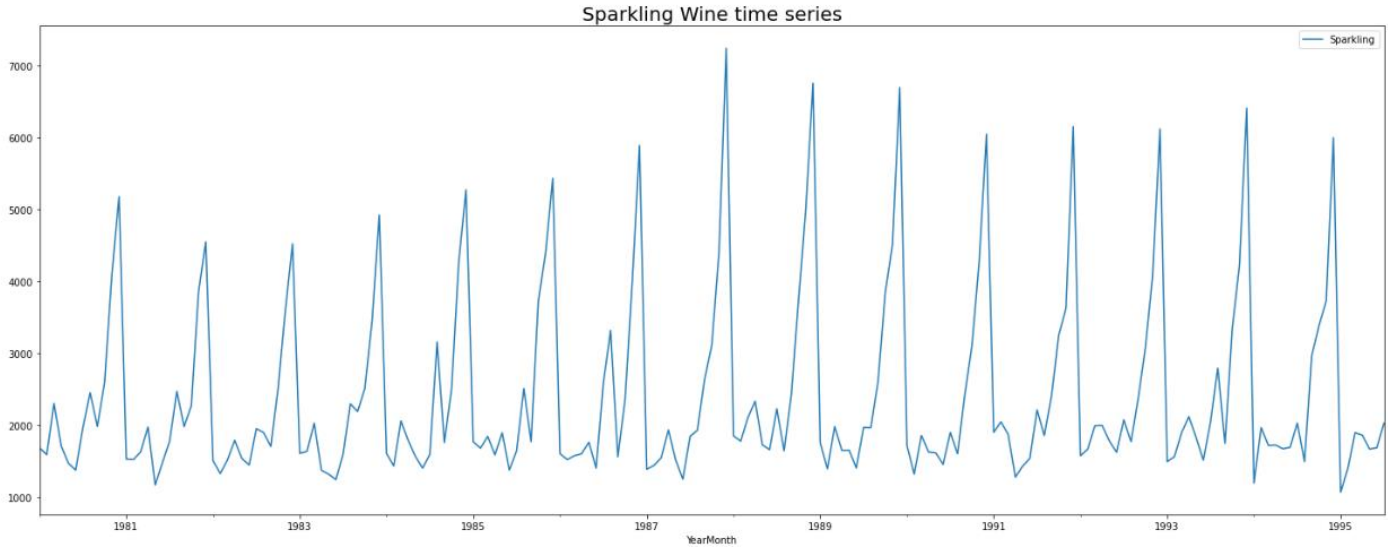
Sparkling wine data set Time Series:



Observations:

- Sparkling wine sales show no much trend in the yearly sale.
- Sparkling wine sales shows seasonality which has yearly pattern.

Rose wine data set Time series:



Observations:

- Rose wine sales shows a decreasing trend in the initial years which stabilizes after few years and again shows a decreasing trend
- Rose wines sales shows seasonality in the data trend and pattern seem to repeat on yearly basis

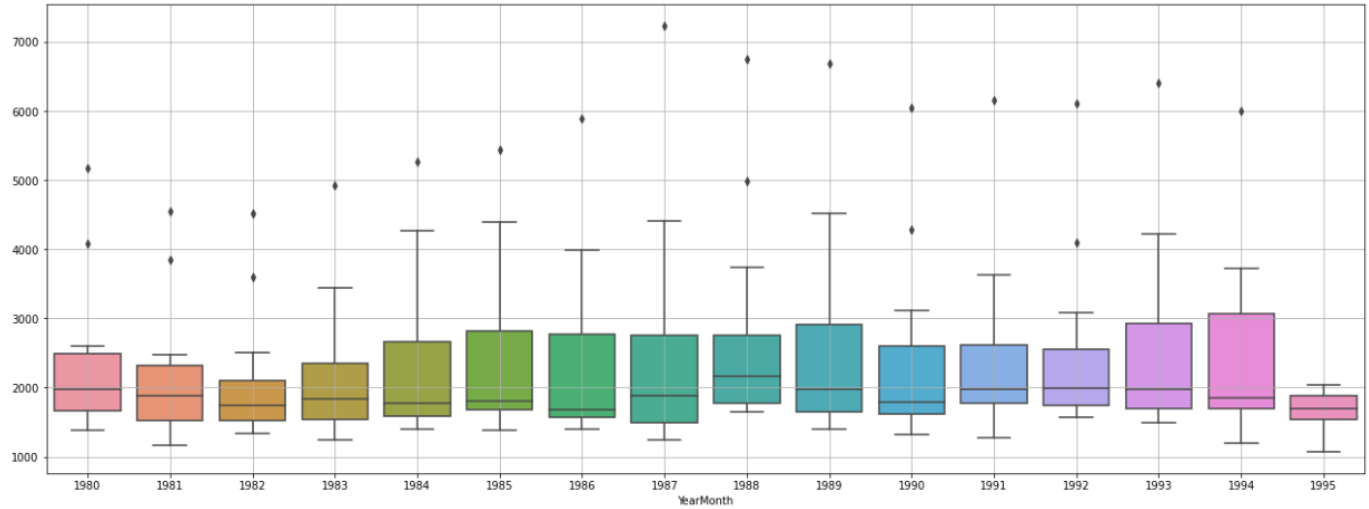
2. Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.

Observations:

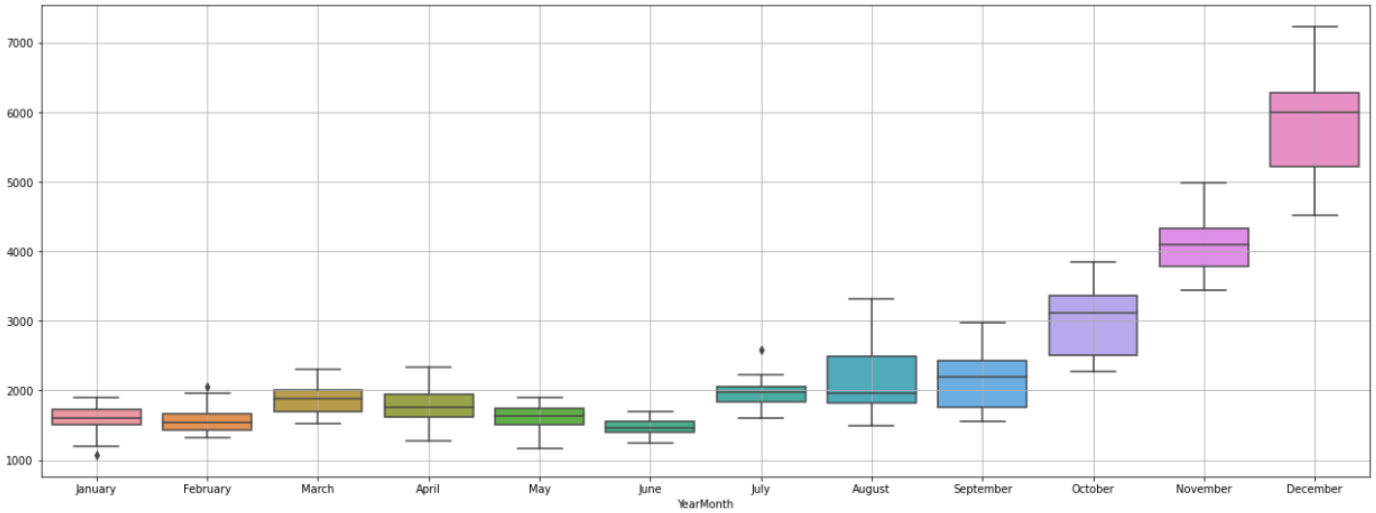
- There are 187 observations in both the data sets which represent the monthly sales of respective wines form the year 1980 to July 1995
- The data has two variables the year/month of sales and the sales for the respective month of the year.
- Mean, min, max values for sparkling wine sales are greater than rose wine sales.

Box Plot of Sparkling Wine:

Yearly:



Monthly:

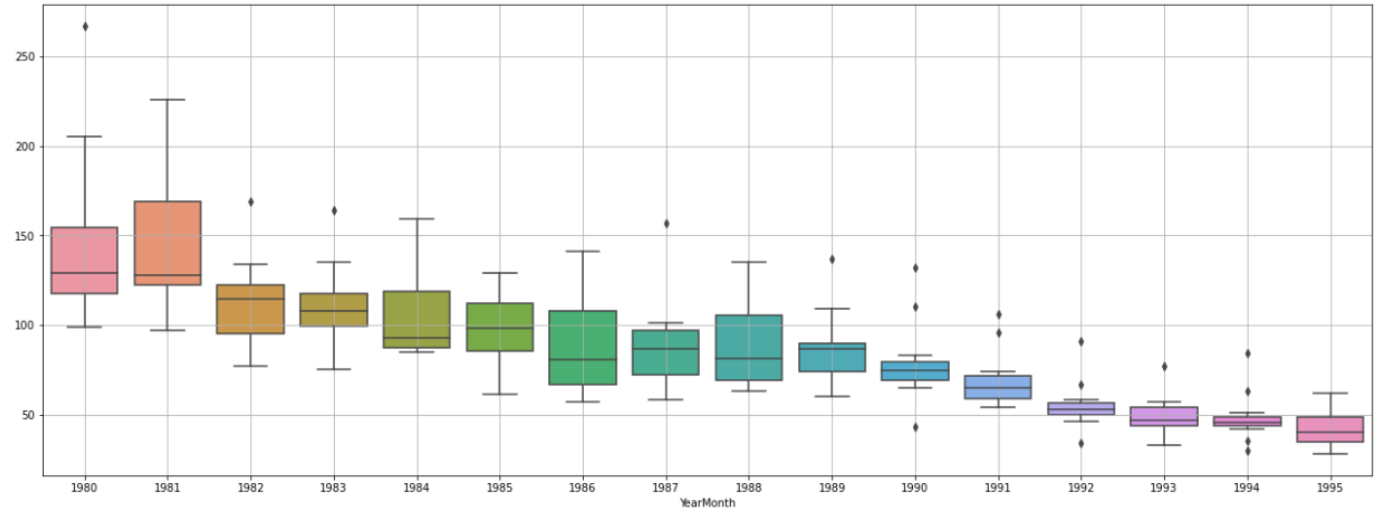


Observations:

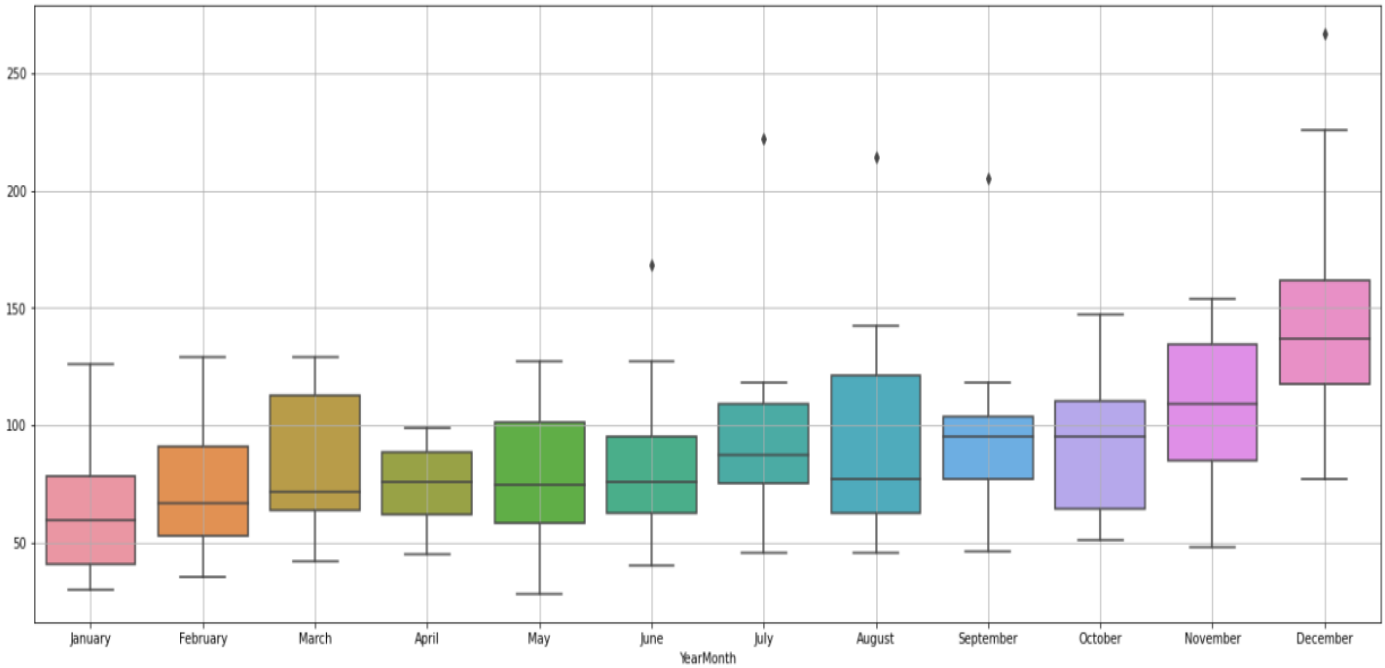
- In agreement with the Time Series plot, the boxplots do not indicate any particular trend.
- The sales of Sparkling wine have some outliers for almost all years except 1995.
- We also observe December month has the highest sales value for Sparkling wine.

Box Plot of Rose Wine:

Yearly:



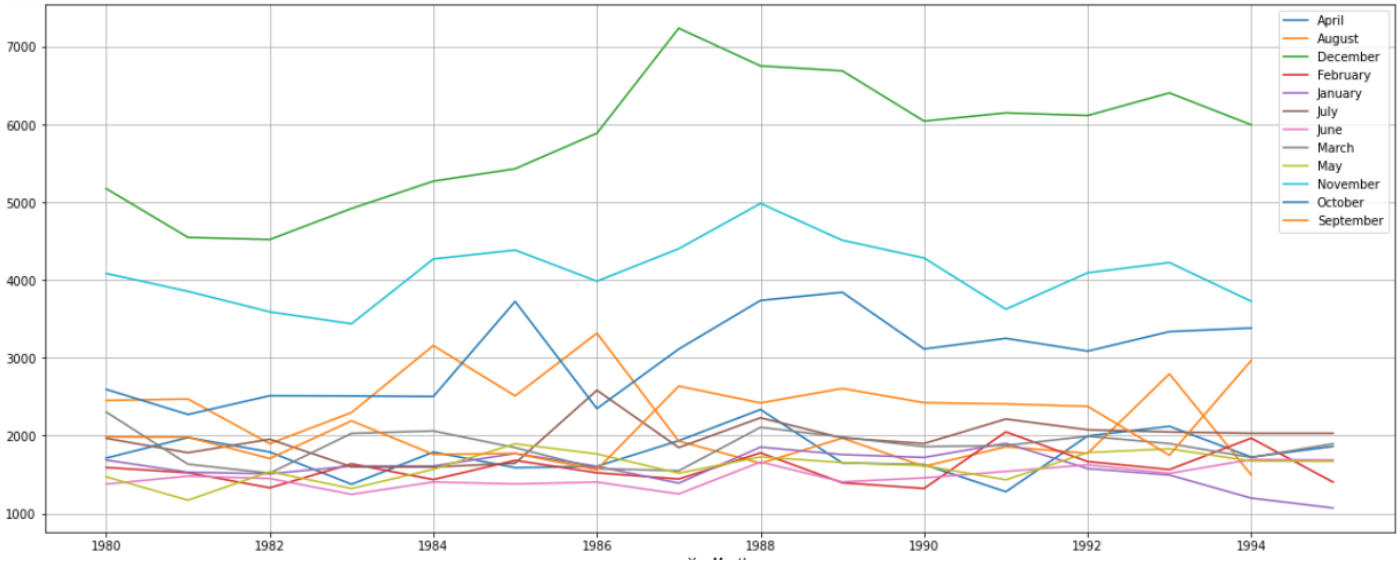
Monthly:



Observations:

- In agreement with the Time Series plot, the year wise boxplots indicate downward trend.
- The sales of Rose wine have some outliers for certain years.
- December seems to have the highest sales of Rose wine and there is also outlier in June, July, August and September months

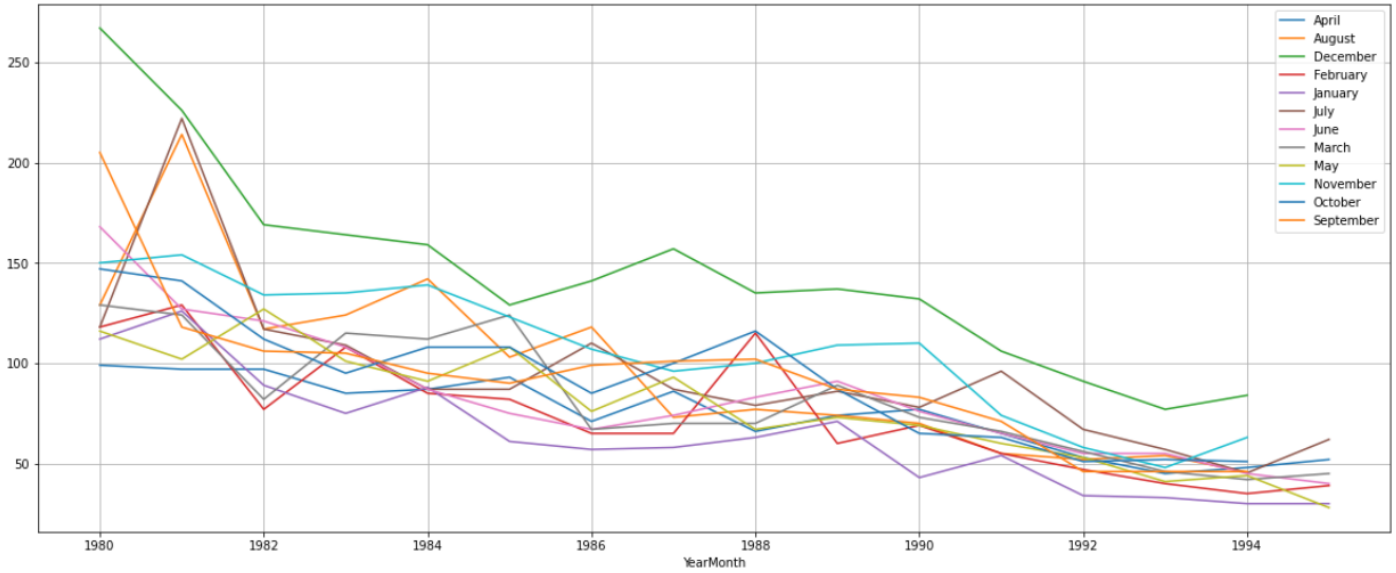
Line Plot of Sparkling Wine:



Observations:

- Maximum sale of sparkling wine is in December and Minimum is in June.

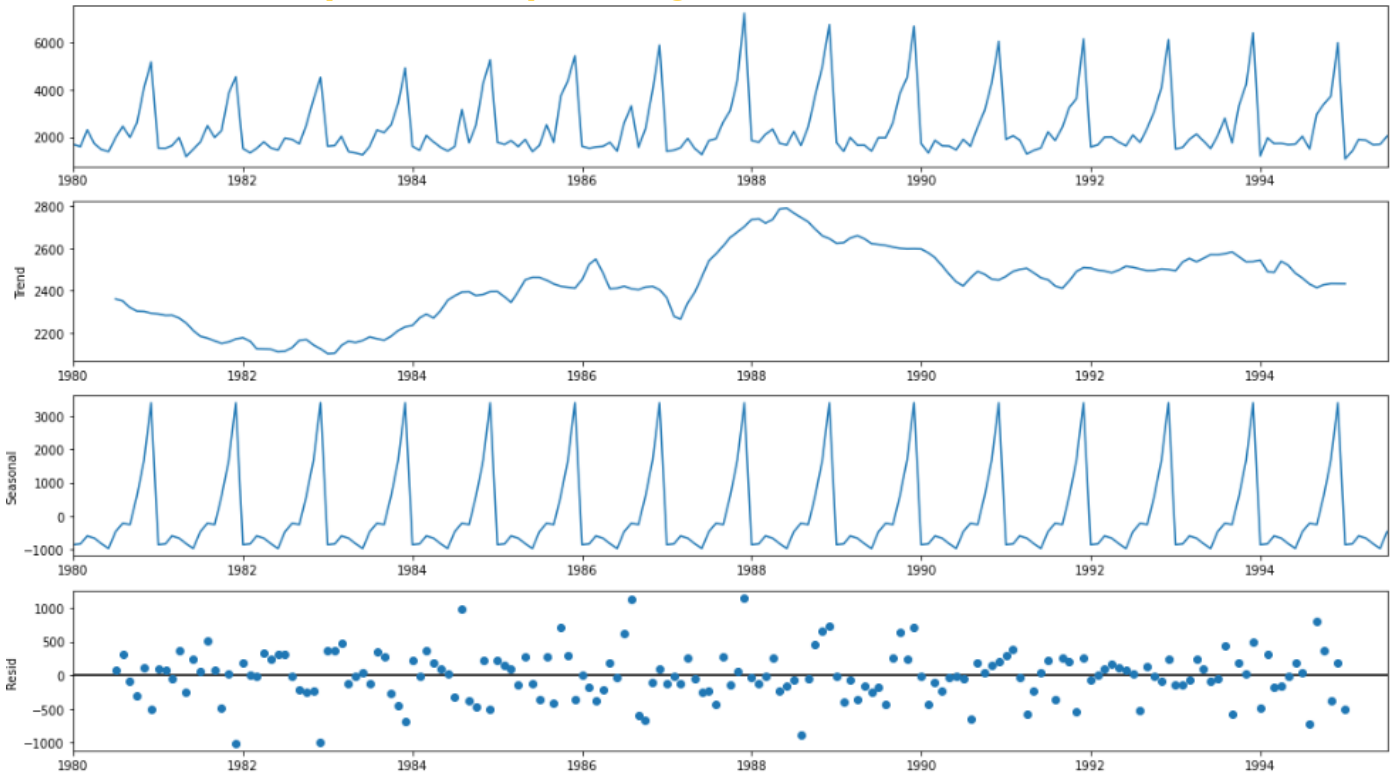
Line Plot of Rose Wine:



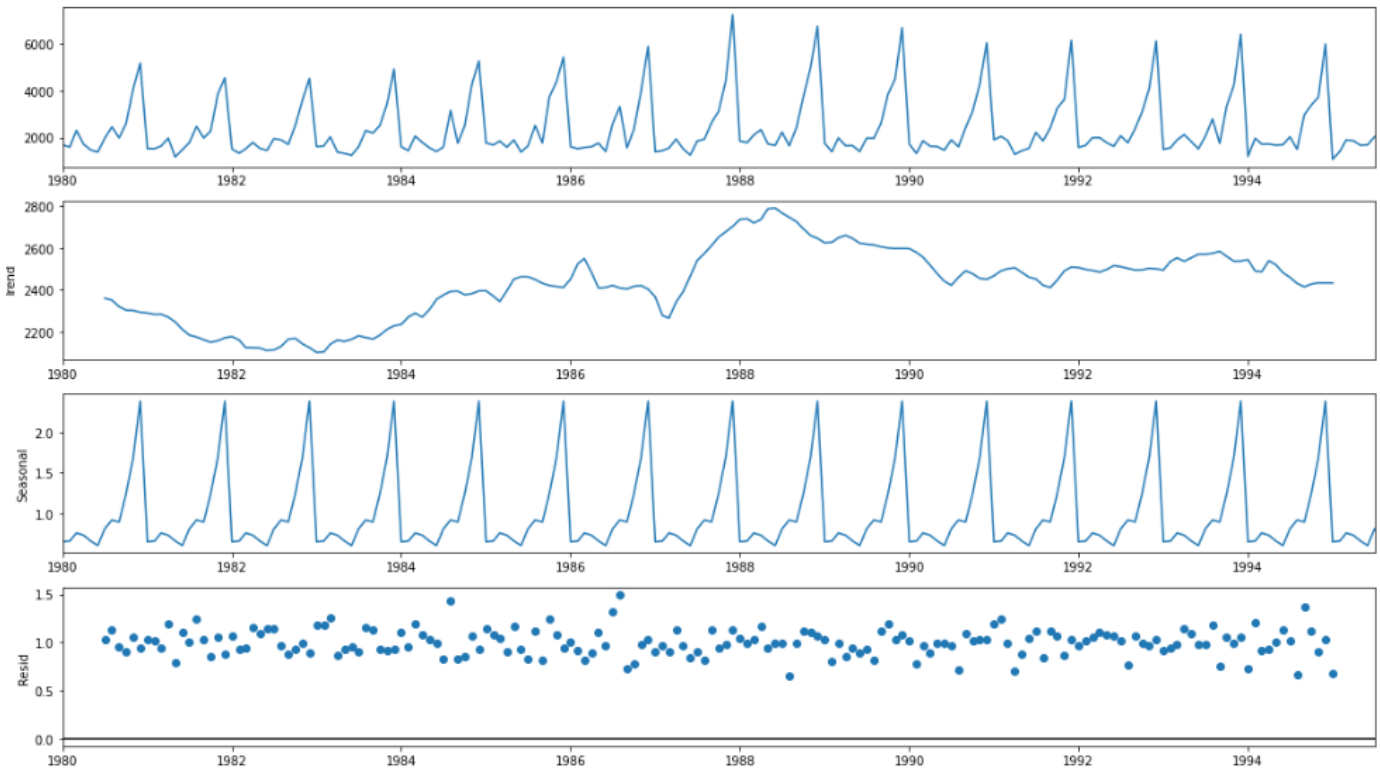
Observations:

- Maximum sale of Rose wine is in December and Minimum is in January.

Additive decomposition sparkling wine sale:



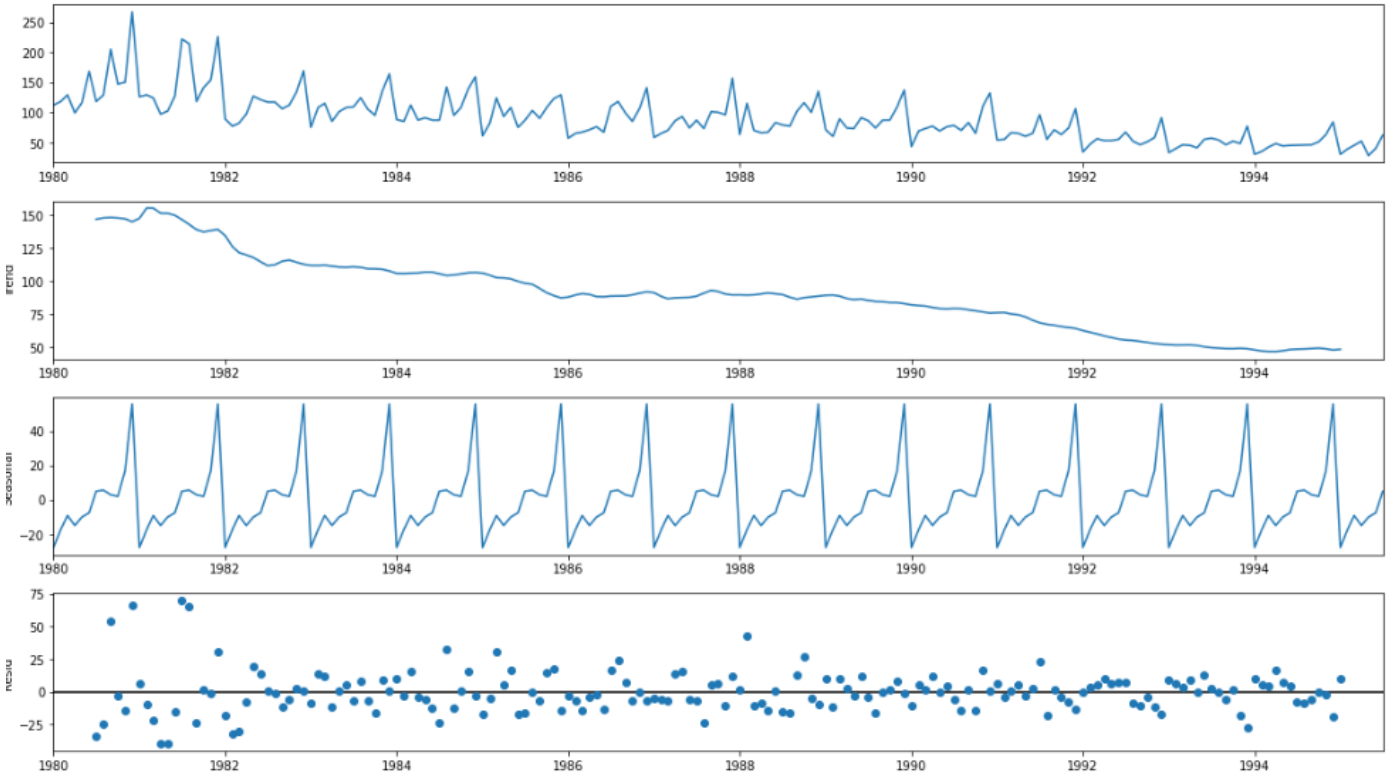
Multiplicative decomposition sparkling wine sale:



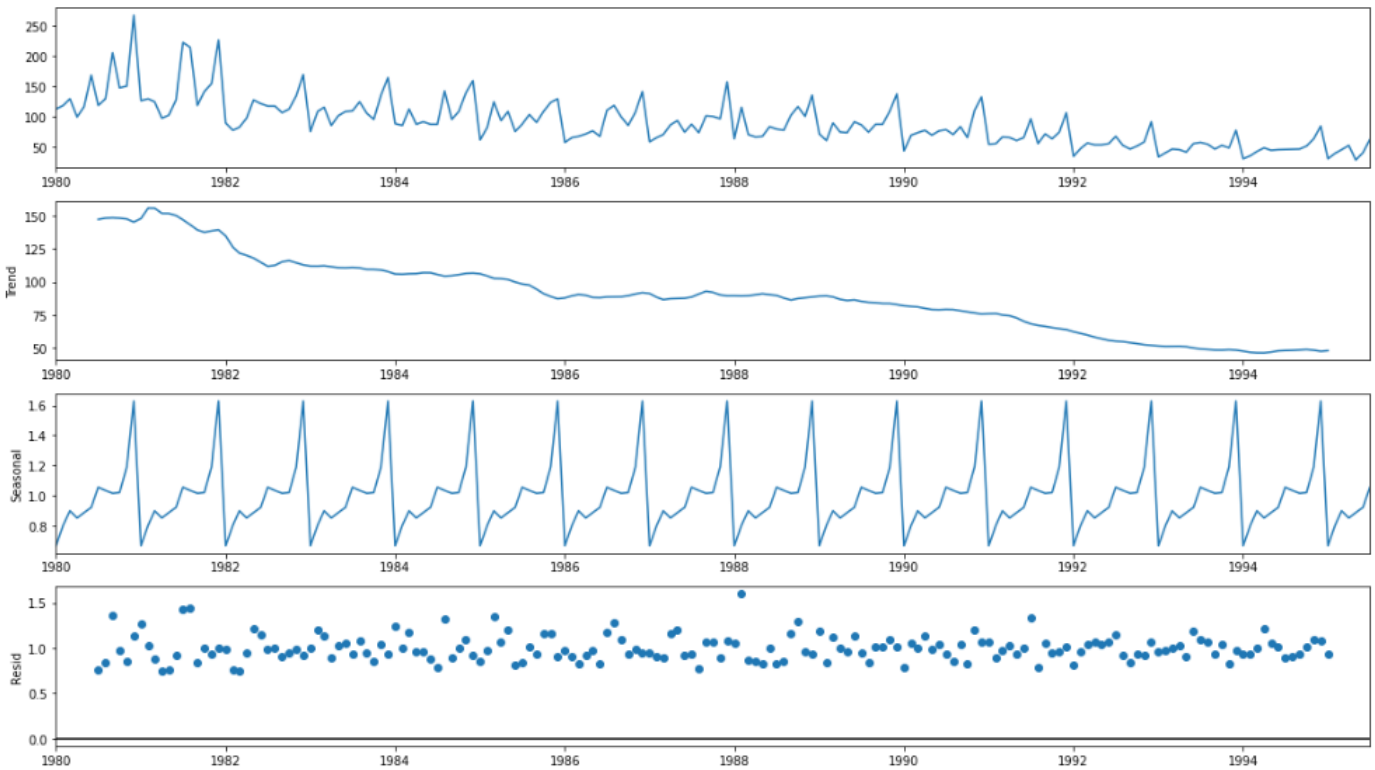
Observations:

- For additive we see the residual values area round 0 and for Multiplicative model we see the residual are around 1.

Additive decomposition Rose wine sale:



Multiplicative decomposition Rose wine sale:



Observations:

- There were two missing values which were interpolated using Linear method.
- For additive we see the residual values area round 0 and for Multiplicative model we see the residual are around 1.

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

Sparkling data set:

Training and Test data-

First few rows of Sparkling Training Data

YearMonth	Sparkling
1980-01-01	1686
1980-02-01	1591
1980-03-01	2304
1980-04-01	1712
1980-05-01	1471

First few rows of Sparkling Test Data

YearMonth	Sparkling
1991-01-01	1902
1991-02-01	2049
1991-03-01	1874
1991-04-01	1279
1991-05-01	1432

Last few rows of Sparkling Training Data

YearMonth	Sparkling
1990-08-01	1605
1990-09-01	2424
1990-10-01	3116
1990-11-01	4286
1990-12-01	6047

Last few rows of Sparkling Test Data

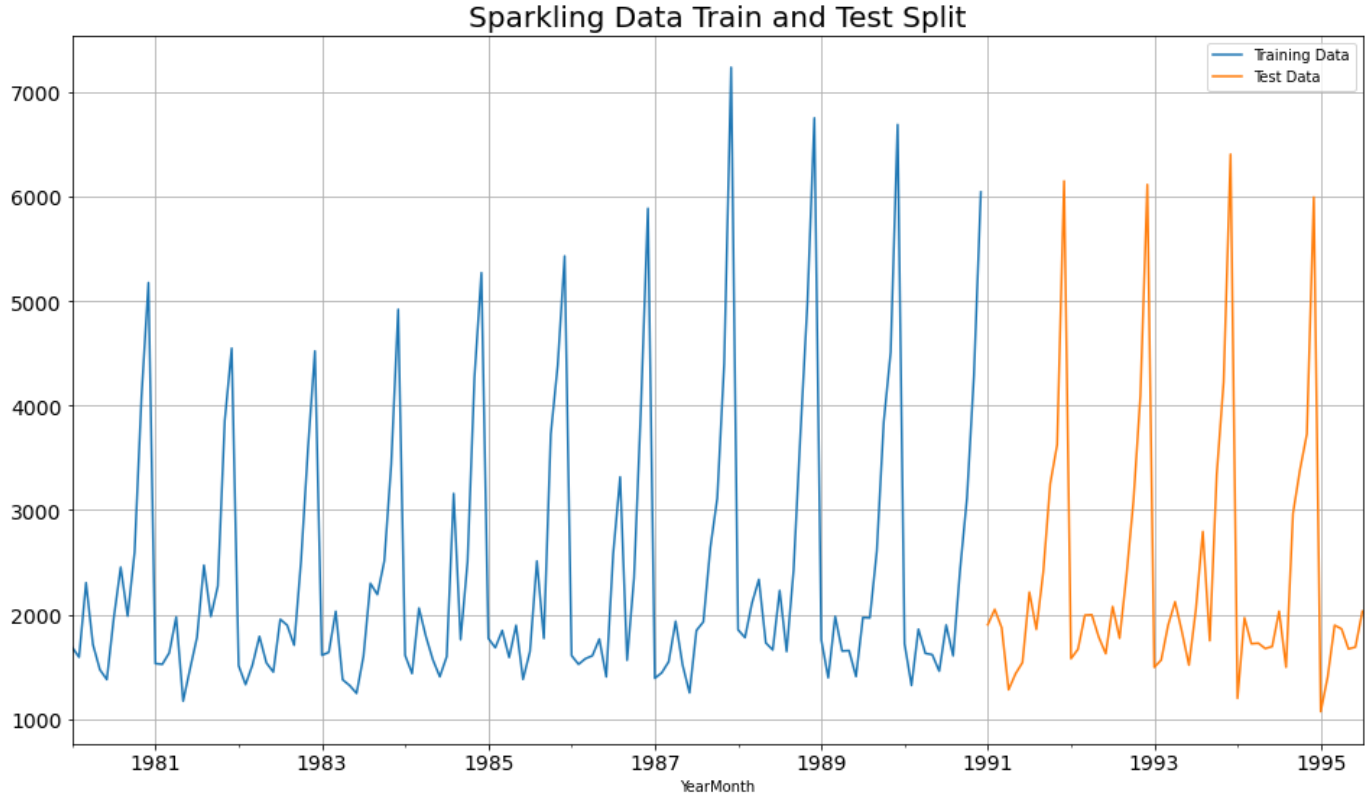
YearMonth	Sparkling
1995-03-01	1897
1995-04-01	1862
1995-05-01	1670
1995-06-01	1688
1995-07-01	2031

Shape of training and test data:

```
1 print(strain.shape)
2 print(stest.shape)
```

```
(132, 1)
(55, 1)
```

SPARKLING DATA TRAIN TEST SPLIT TIME SERIES



Rose Data set:

Training and Test data-

First few rows of Rose Training Data

YearMonth	Rose
1980-01-01	112.0
1980-02-01	118.0
1980-03-01	129.0
1980-04-01	99.0
1980-05-01	116.0

First few rows of Rose Test Data

YearMonth	Rose
1991-01-01	54.0
1991-02-01	55.0
1991-03-01	66.0
1991-04-01	65.0
1991-05-01	60.0

Last few rows of Rose Training Data

YearMonth	Rose
1990-08-01	70.0
1990-09-01	83.0
1990-10-01	65.0
1990-11-01	110.0
1990-12-01	132.0

Last few rows of Rose Test Data

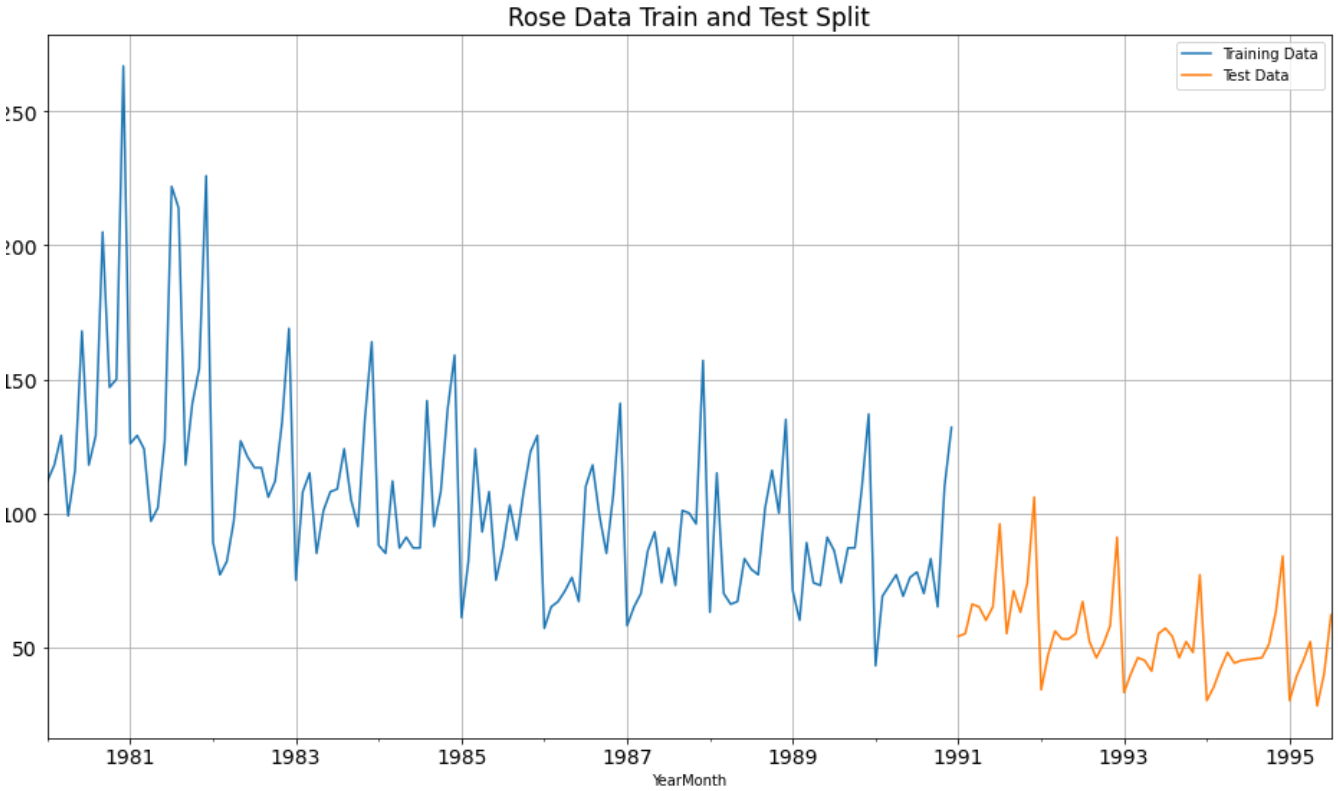
YearMonth	Rose
1995-03-01	45.0
1995-04-01	52.0
1995-05-01	28.0
1995-06-01	40.0
1995-07-01	62.0

Shape of Training and Test data:

```
1 print(rtrain.shape)
2 print(rtest.shape)
```

```
(132, 1)
(55, 1)
```

SPARKLING DATA TRAIN TEST SPLIT TIME SERIES:



Observations:

- The train data of Rose and sparkling wine sales has been split for data up to 1990 and has 132 data points.
- The Test data of Rose and sparkling wine sales has been split for data from 1991 and has 55 data points.

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.

Model 1: Linear Regression

Linear Regression Sparkling:

First few rows of Training Data

YearMonth	Sparkling	Time
1980-01-01	1686	1
1980-02-01	1591	2
1980-03-01	2304	3
1980-04-01	1712	4
1980-05-01	1471	5

First few rows of Test Data

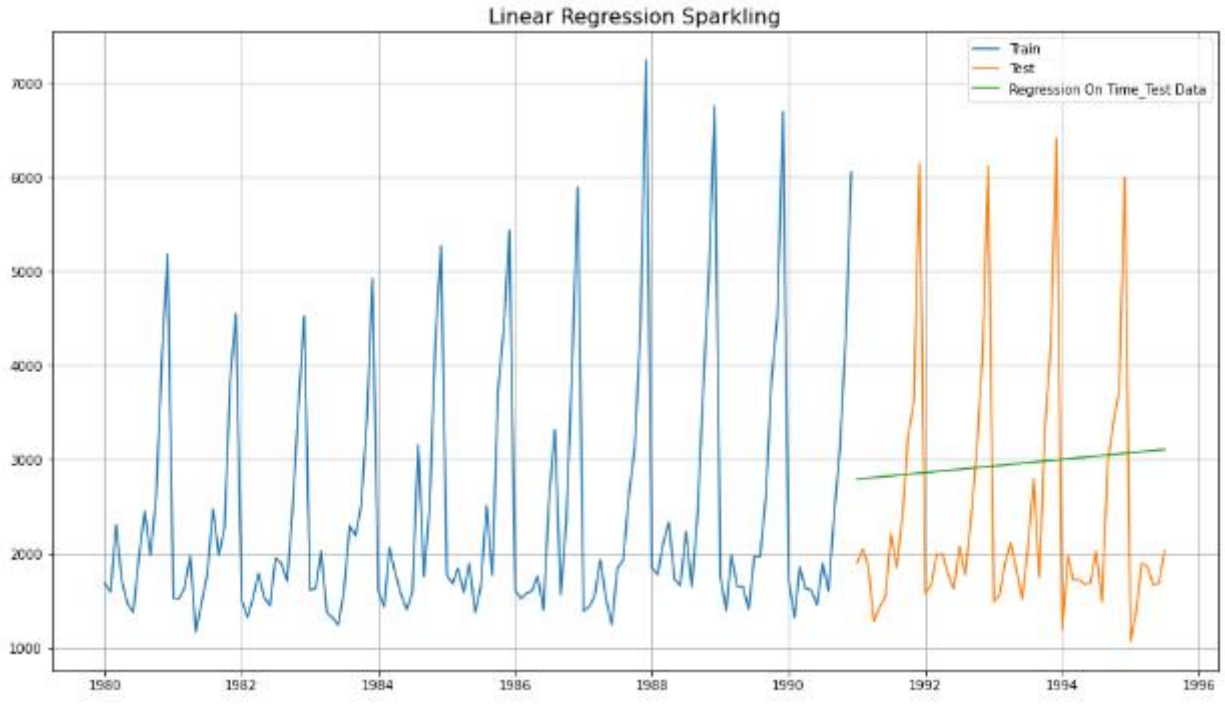
YearMonth	Sparkling	Time
1991-01-01	1902	133
1991-02-01	2049	134
1991-03-01	1874	135
1991-04-01	1279	136
1991-05-01	1432	137

Last few rows of Training Data

YearMonth	Sparkling	Time
1990-08-01	1605	128
1990-09-01	2424	129
1990-10-01	3116	130
1990-11-01	4286	131
1990-12-01	6047	132

Last few rows of Test Data

YearMonth	Sparkling	Time
1995-03-01	1897	183
1995-04-01	1862	184
1995-05-01	1670	185
1995-06-01	1688	186
1995-07-01	2031	187



Linear Regression Rose set:

First few rows of Training Data

YearMonth	Rose	Time
1980-01-01	112.0	1
1980-02-01	118.0	2
1980-03-01	129.0	3
1980-04-01	99.0	4
1980-05-01	116.0	5

First few rows of Test Data

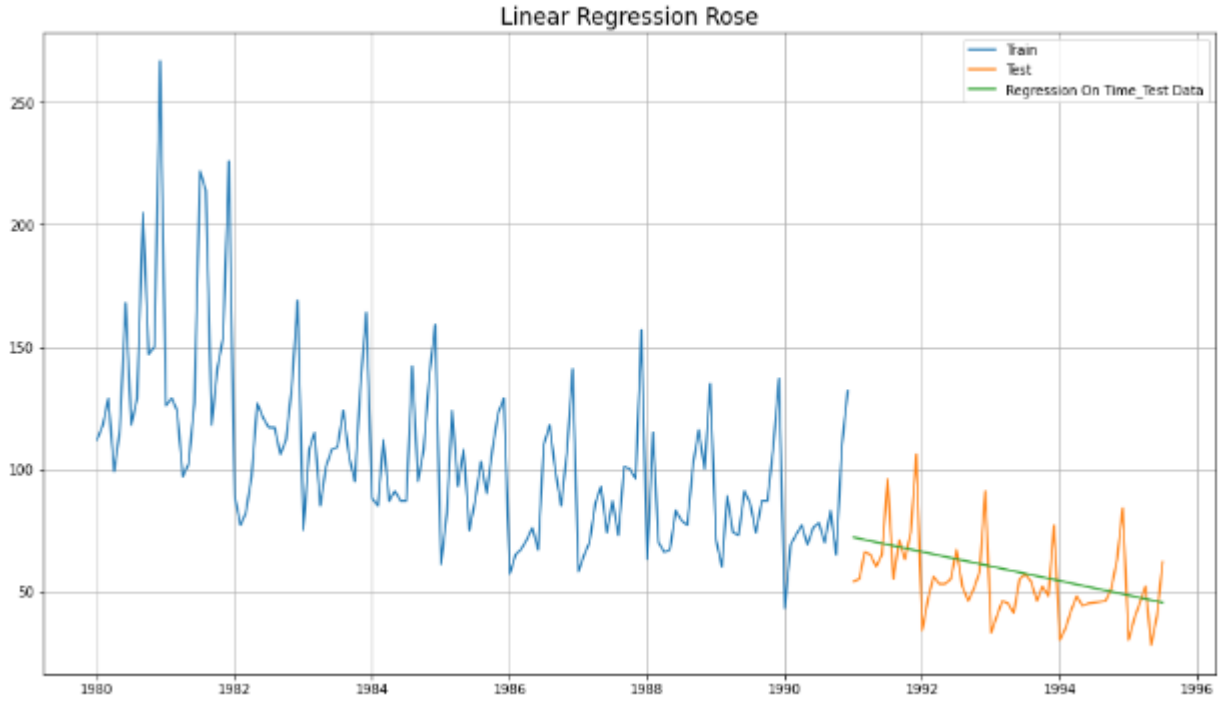
YearMonth	Rose	Time
1991-01-01	54.0	133
1991-02-01	55.0	134
1991-03-01	66.0	135
1991-04-01	65.0	136
1991-05-01	60.0	137

Last few rows of Training Data

YearMonth	Rose	Time
1990-08-01	70.0	128
1990-09-01	83.0	129
1990-10-01	65.0	130
1990-11-01	110.0	131
1990-12-01	132.0	132

Last few rows of Test Data

YearMonth	Rose	Time
1995-03-01	45.0	183
1995-04-01	52.0	184
1995-05-01	28.0	185
1995-06-01	40.0	186
1995-07-01	62.0	187



Model Evaluation:

	Test RMSE Sparkling	Test RMSE Rose
RegressionOnTime	1389.135175	15.268955

Observations:

Sparkling:

- The root means square error for the linear regression model generated = 1389.13
- The line shows a down ward trend for the rose sales whereas it shows an upward trend for the sparkling wine sales.

Rose:

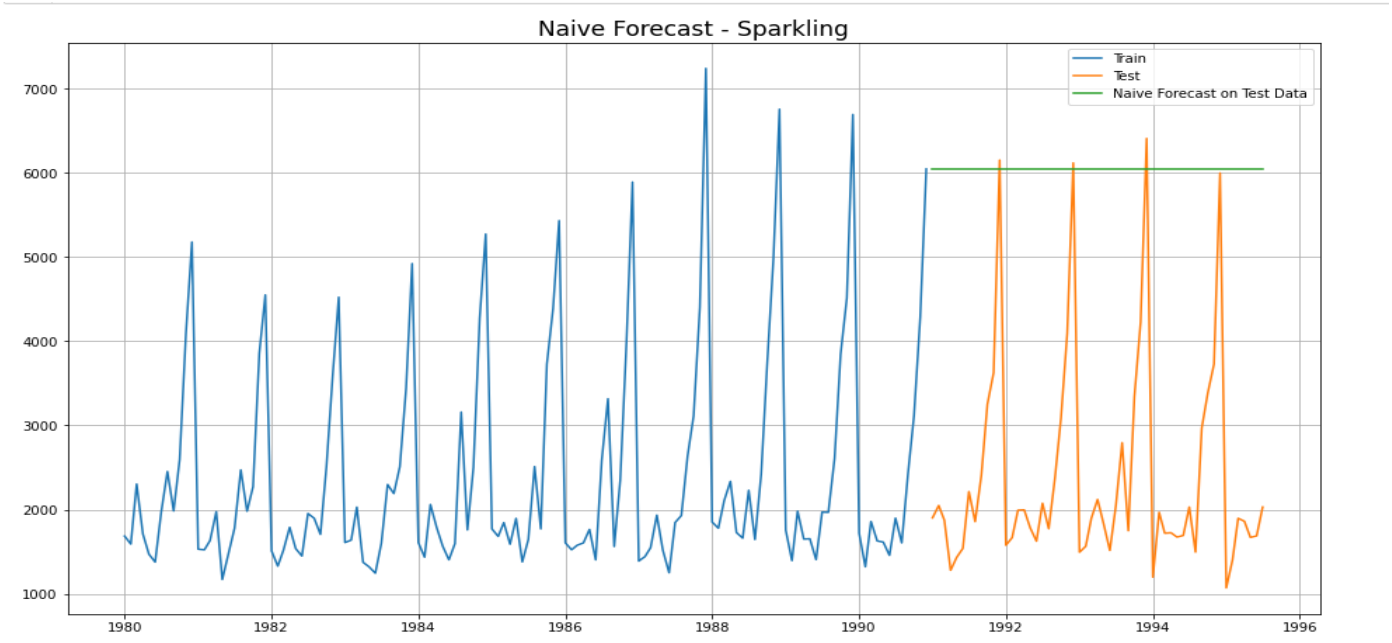
- The root means square error for the linear regression model generated = 15.26.
- The predicted values for the test data using linear regression model are shown as a straight line with slope.

Model 2: Naive Approach

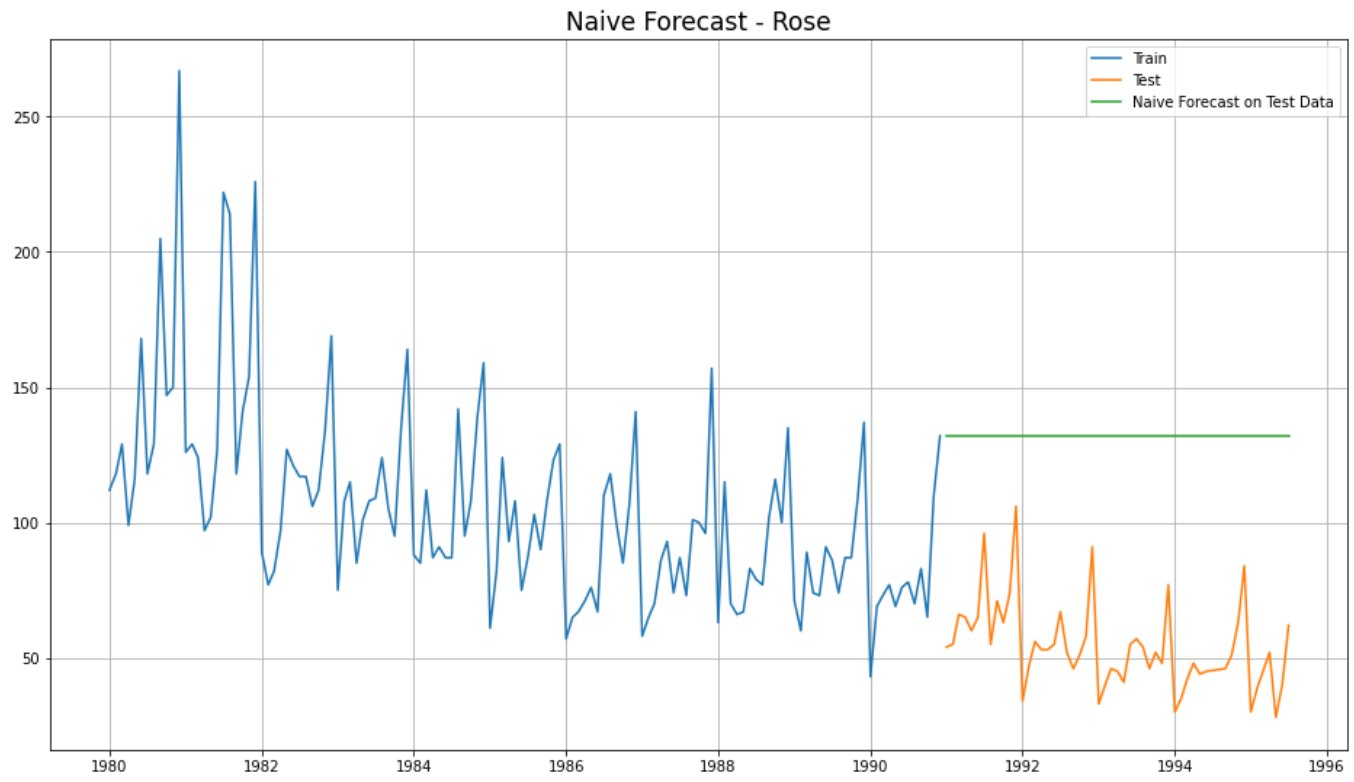
For this particular naive model, we say that the prediction for tomorrow is the same as today and the prediction for day after tomorrow is tomorrow and since the prediction of tomorrow is same as today, therefore the prediction for day after tomorrow is also today.

$\hat{y}\,t+1=yt$

Sparkling:



Rose:



Model Evaluation:

	Test RMSE Sparkling	Test RMSE Rose
NaiveModel	3864.279352	79.718773

Observations:

Sparkling:

- The root means square error for the Naïve Bayes model generated = 3864.27.
- The graph shows Straight

Rose:

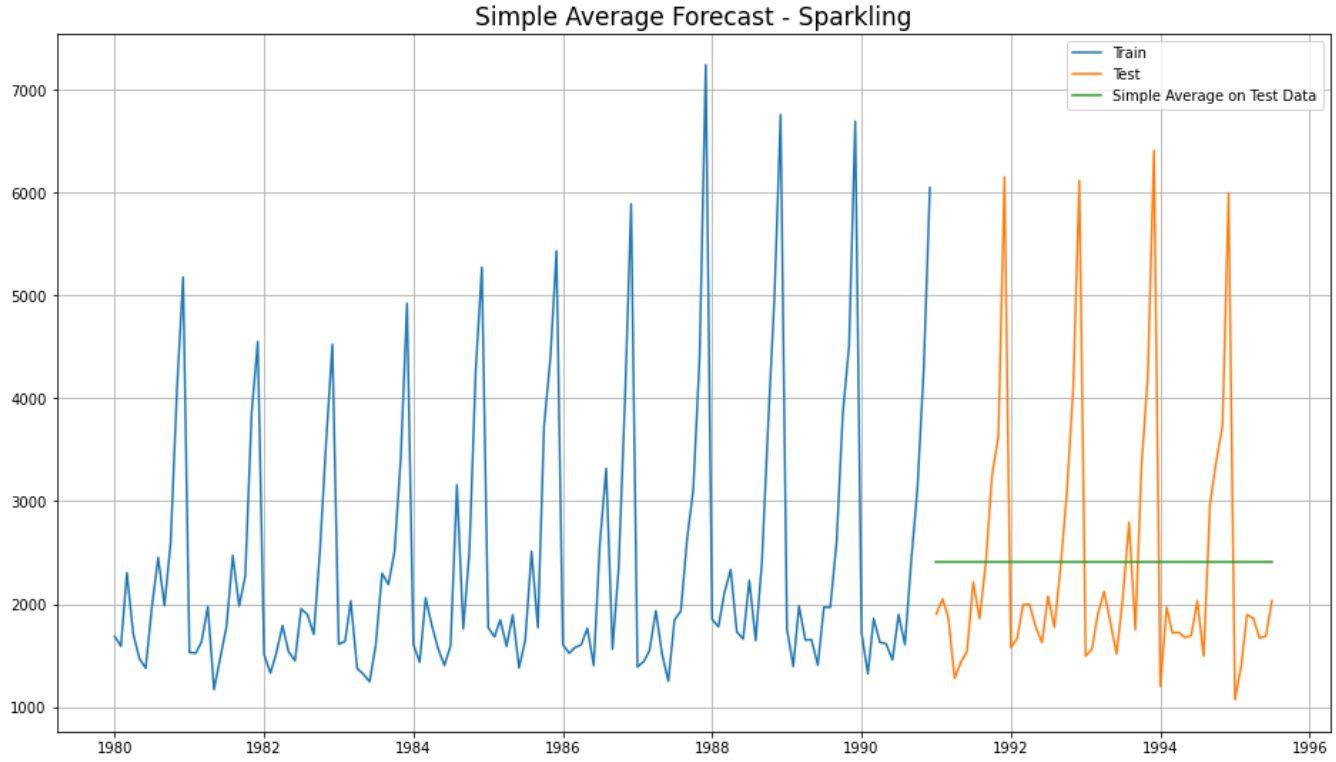
- The root means square error for the linear regression model generated = 79.71
- The graph shows Straight line

RMSE value for naïve model generated for both datasets are much higher than the regression model values.

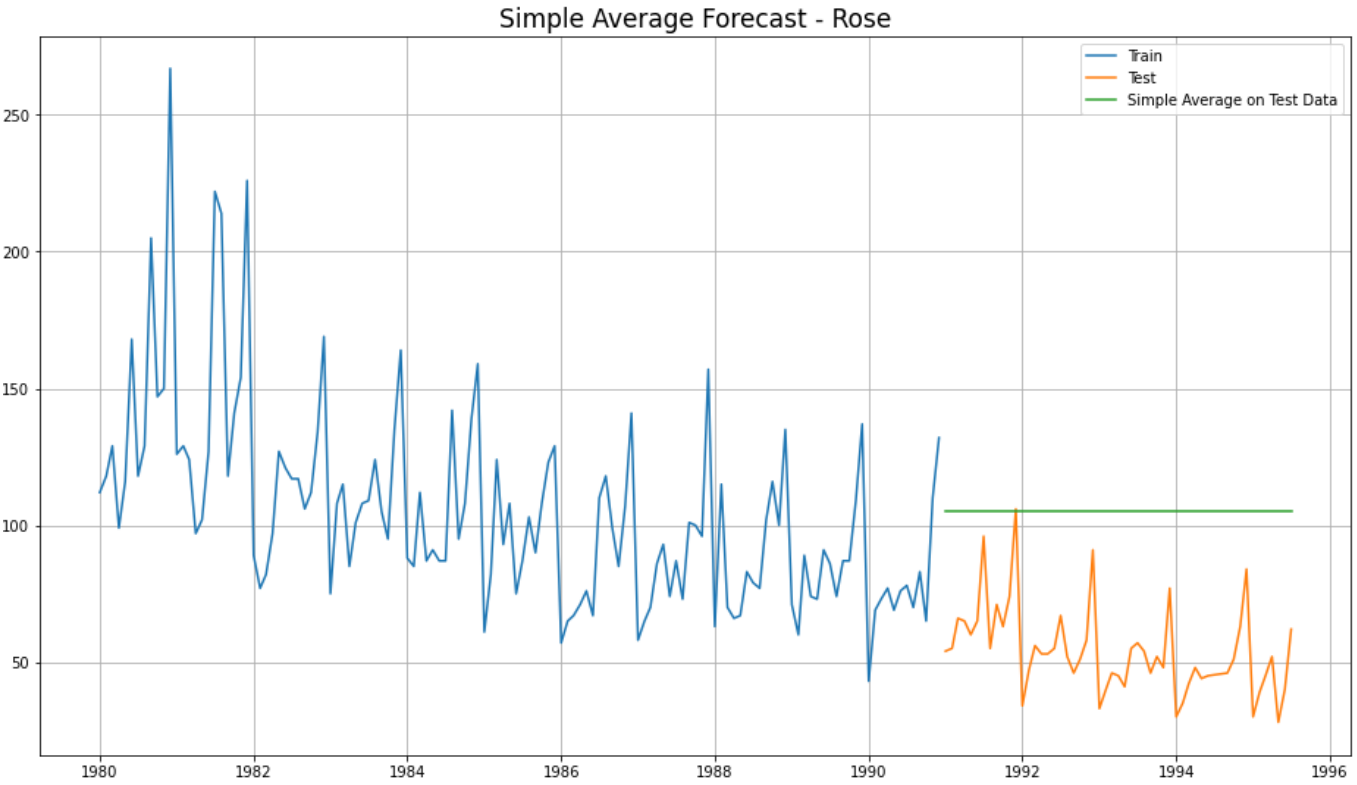
Model 3: Simple Average

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

Sparkling:



Rose:



Model Evaluation:

	Test RMSE Rose	Test RMSE Sparkling
SimpleAverageModel	53.46057	1275.081804

Observations:

Sparkling:

- The root means square error for the Naïve Bayes model generated = 1275.08
- Predicted graph shows a Straight line.

Rose:

- The root means square error for the linear regression model generated = 53.46
- Predicted graph shows a Straight line.

Model 4: Moving Average (MA)

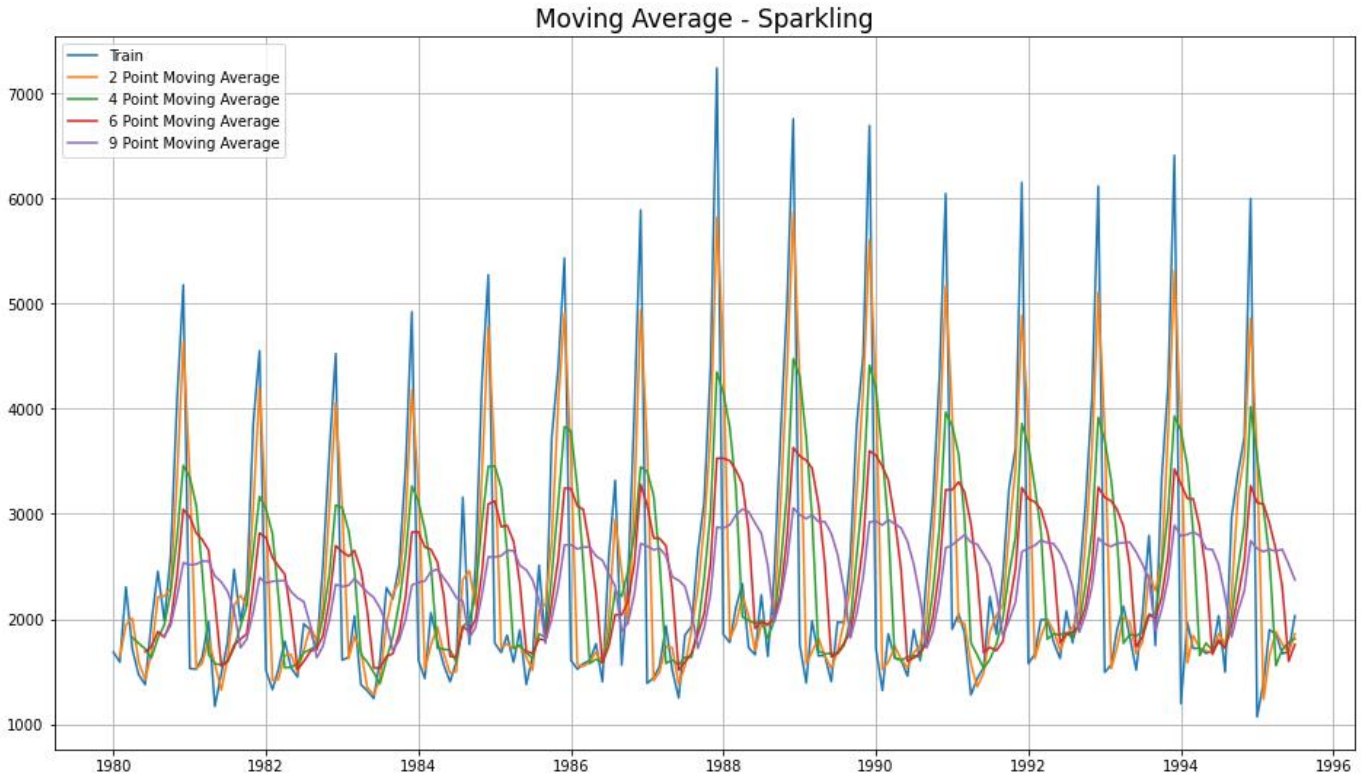
For the moving average model, we are going to calculate rolling means (or moving averages) for different intervals. The best interval can be determined by the maximum accuracy (or the minimum error) over here. For Moving Average, we are going to average over the entire data.

Sparkling:

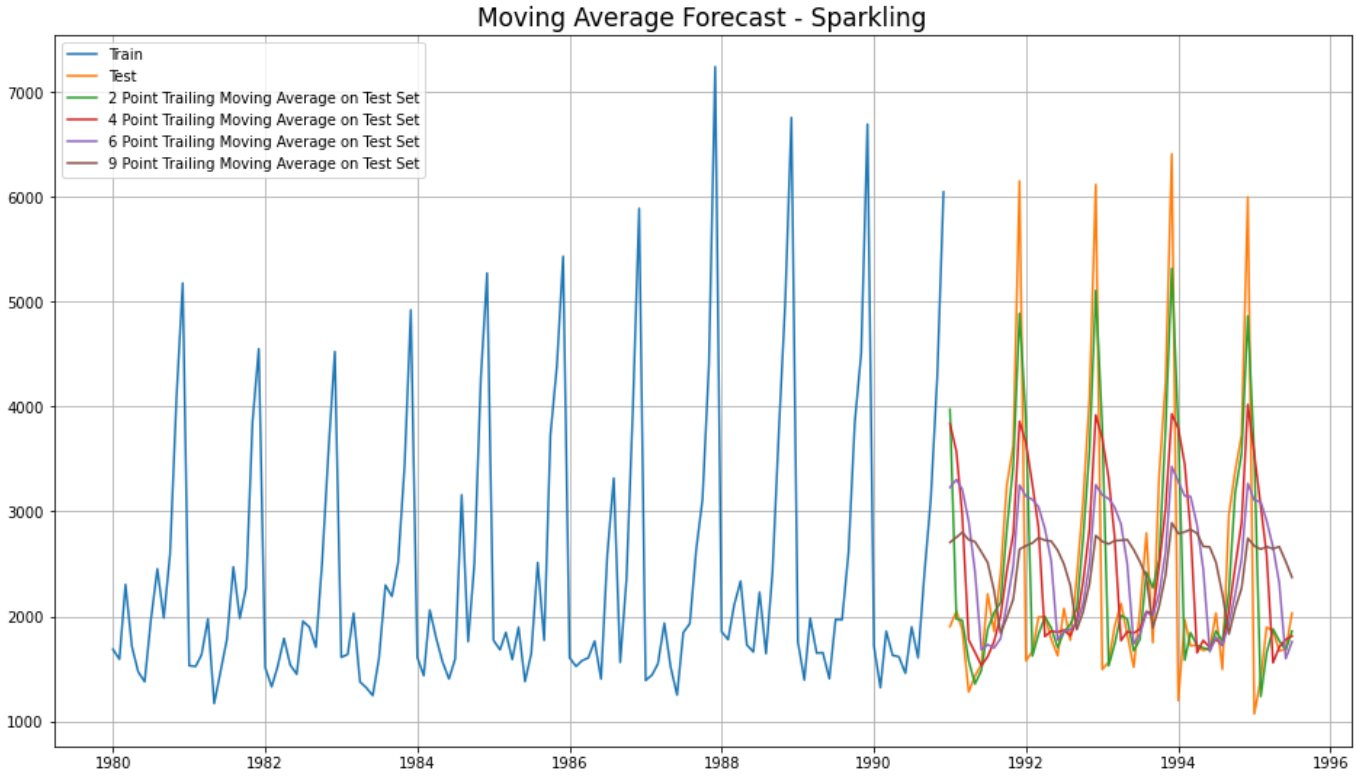
Trailing moving averages:

YearMonth	Sparkling	Trailing_2	Trailing_4	Trailing_6	Trailing_9
1980-01-01	1686	NaN	NaN	NaN	NaN
1980-02-01	1591	1638.5	NaN	NaN	NaN
1980-03-01	2304	1947.5	NaN	NaN	NaN
1980-04-01	1712	2008.0	1823.25	NaN	NaN
1980-05-01	1471	1591.5	1769.50	NaN	NaN

Moving Average – Sparkling



Moving Average Forecast - Sparkling

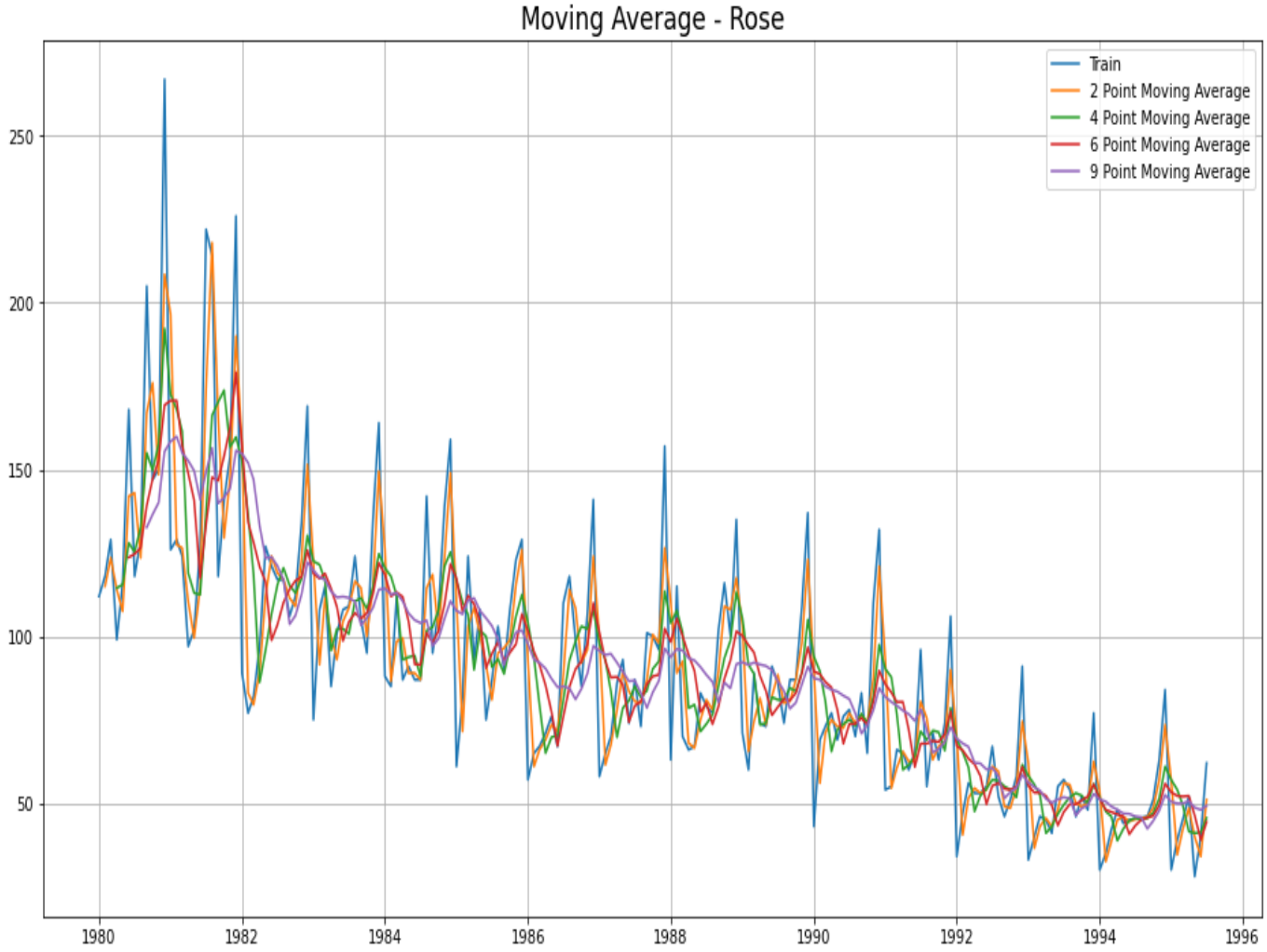


Rose:

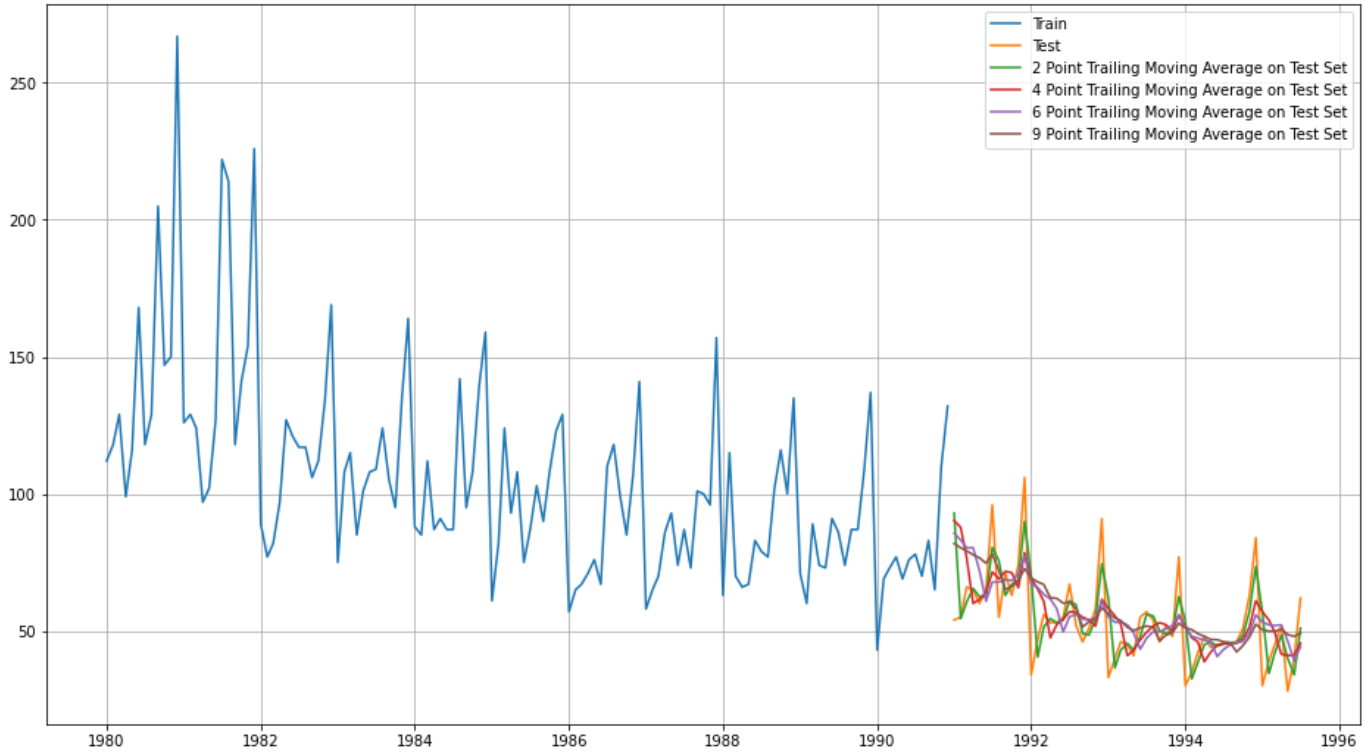
Trailing moving averages:

YearMonth	Rose	Trailing_2	Trailing_4	Trailing_6	Trailing_9
1980-01-01	112.0	NaN	NaN	NaN	NaN
1980-02-01	118.0	115.0	NaN	NaN	NaN
1980-03-01	129.0	123.5	NaN	NaN	NaN
1980-04-01	99.0	114.0	114.5	NaN	NaN
1980-05-01	116.0	107.5	115.5	NaN	NaN

Moving Average Forecast – Rose



Moving Average – Rose



Model Evaluation:

	Test RMSE Sparkling	Test RMSE Rose
2pointTrailingMovingAverage	813.400684	11.529278
4pointTrailingMovingAverage	1156.589694	14.451403
6pointTrailingMovingAverage	1283.927428	14.566327
9pointTrailingMovingAverage	1346.278315	14.727630

Observations:

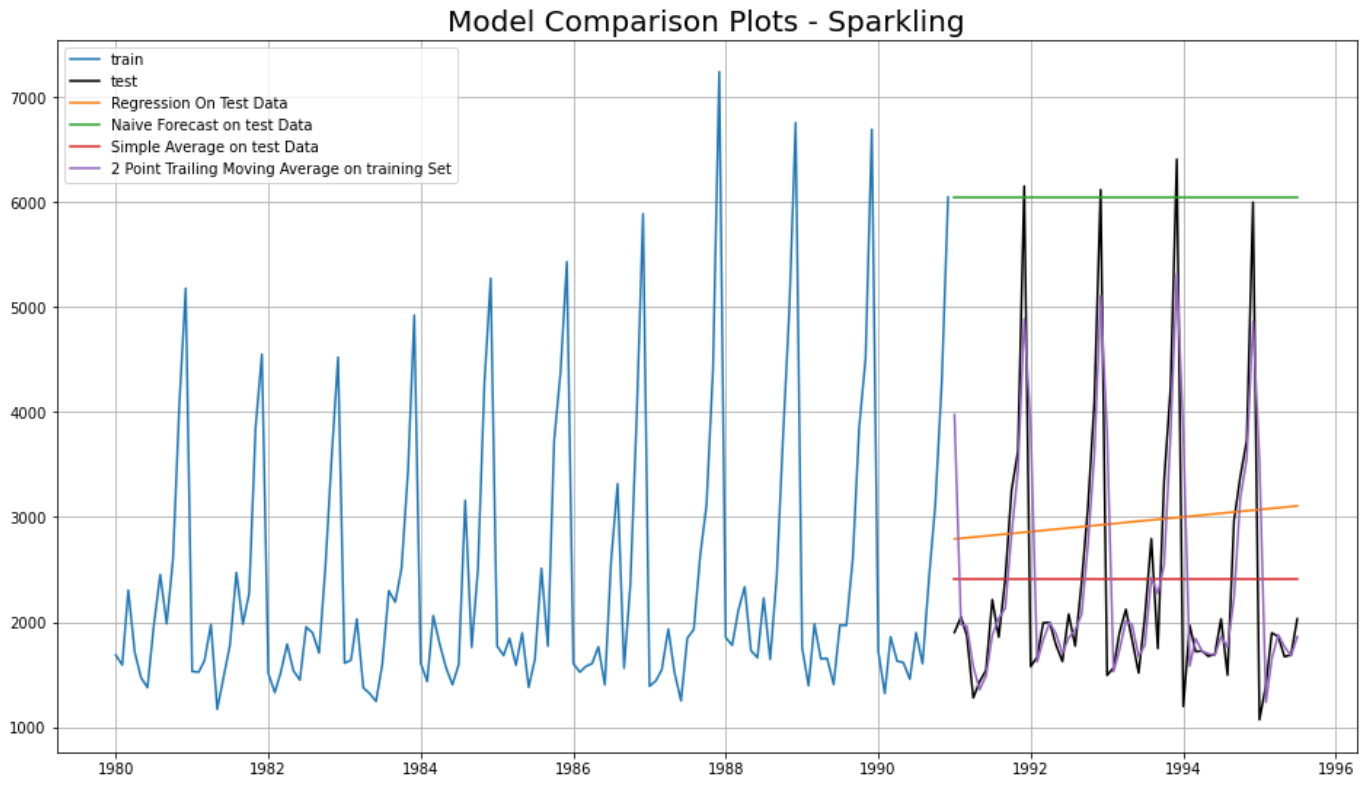
Sparkling:

- For 2 point Moving Average Model forecast on the Training Data, RMSE is 813.40
- For 4 point Moving Average Model forecast on the Training Data, RMSE is 1156.5
- For 6 point Moving Average Model forecast on the Training Data, RMSE is 1283.927
- For 9 point Moving Average Model forecast on the Training Data, RMSE is 1346.27

Rose:

- For 2 point Moving Average Model forecast on the Training Data, RMSE is 11.52
- For 4 point Moving Average Model forecast on the Training Data, RMSE is 14.45
- For 6 point Moving Average Model forecast on the Training Data, RMSE is 14.56
- For 9 point Moving Average Model forecast on the Training Data, RMSE is 114.72

Comparison Model – Sparkling



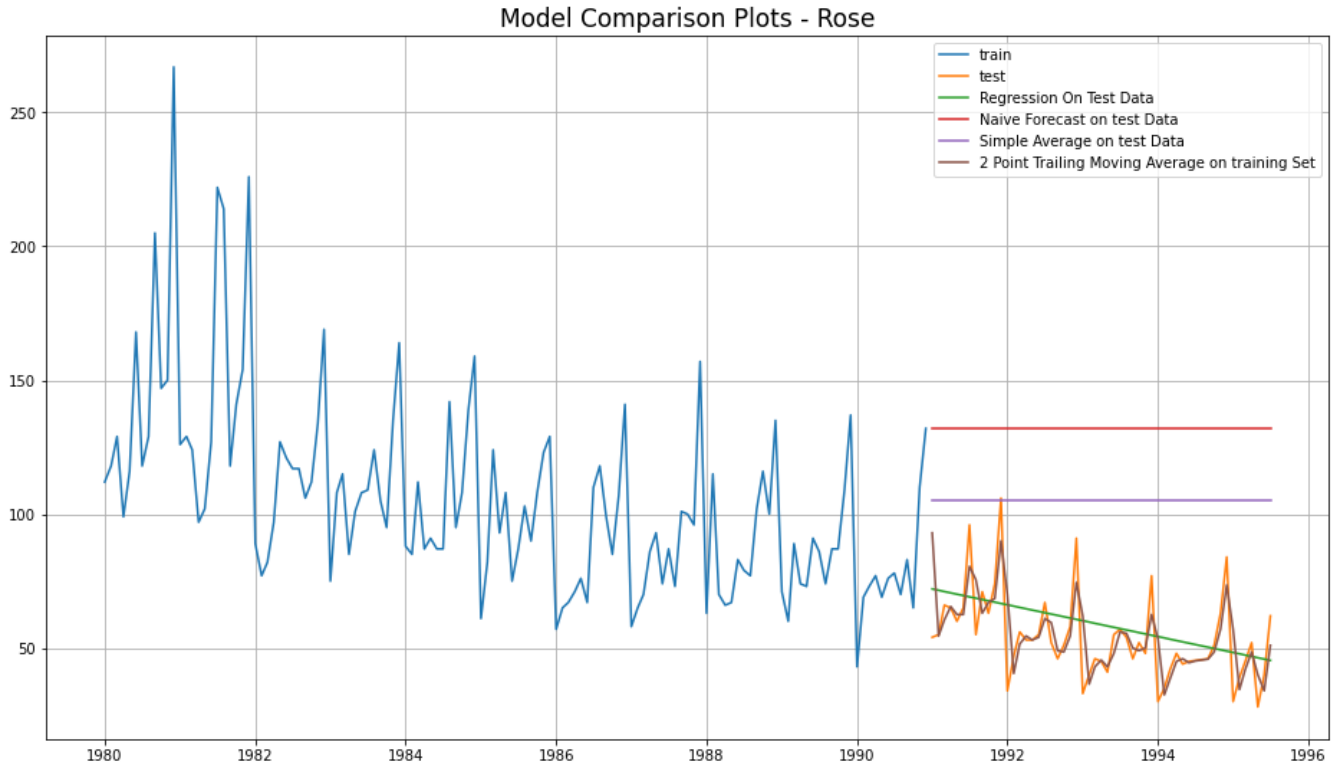
Model Evaluation:

:

	Test RMSE Sparkling	Test RMSE Rose
RegressionOnTime	1389.135175	15.268955
NaiveModel	3864.279352	79.718773
SimpleAverageModel	1275.081804	53.460570
2pointTrailingMovingAverage	813.400684	11.529278
4pointTrailingMovingAverage	1156.589694	14.451403
6pointTrailingMovingAverage	1283.927428	14.566327
9pointTrailingMovingAverage	1346.278315	14.727630

Lowest score is 2 Point Moving average

Comparison Model – Rose



Model Evaluation:

:

	Test RMSE Sparkling	Test RMSE Rose
RegressionOnTime	1389.135175	15.268955
NaiveModel	3864.279352	79.718773
SimpleAverageModel	1275.081804	53.460570
2pointTrailingMovingAverage	813.400684	11.529278
4pointTrailingMovingAverage	1156.589694	14.451403
6pointTrailingMovingAverage	1283.927428	14.566327
9pointTrailingMovingAverage	1346.278315	14.727630

Lowest score is 2 Point Moving average

Method 5: Simple Exponential Smoothing

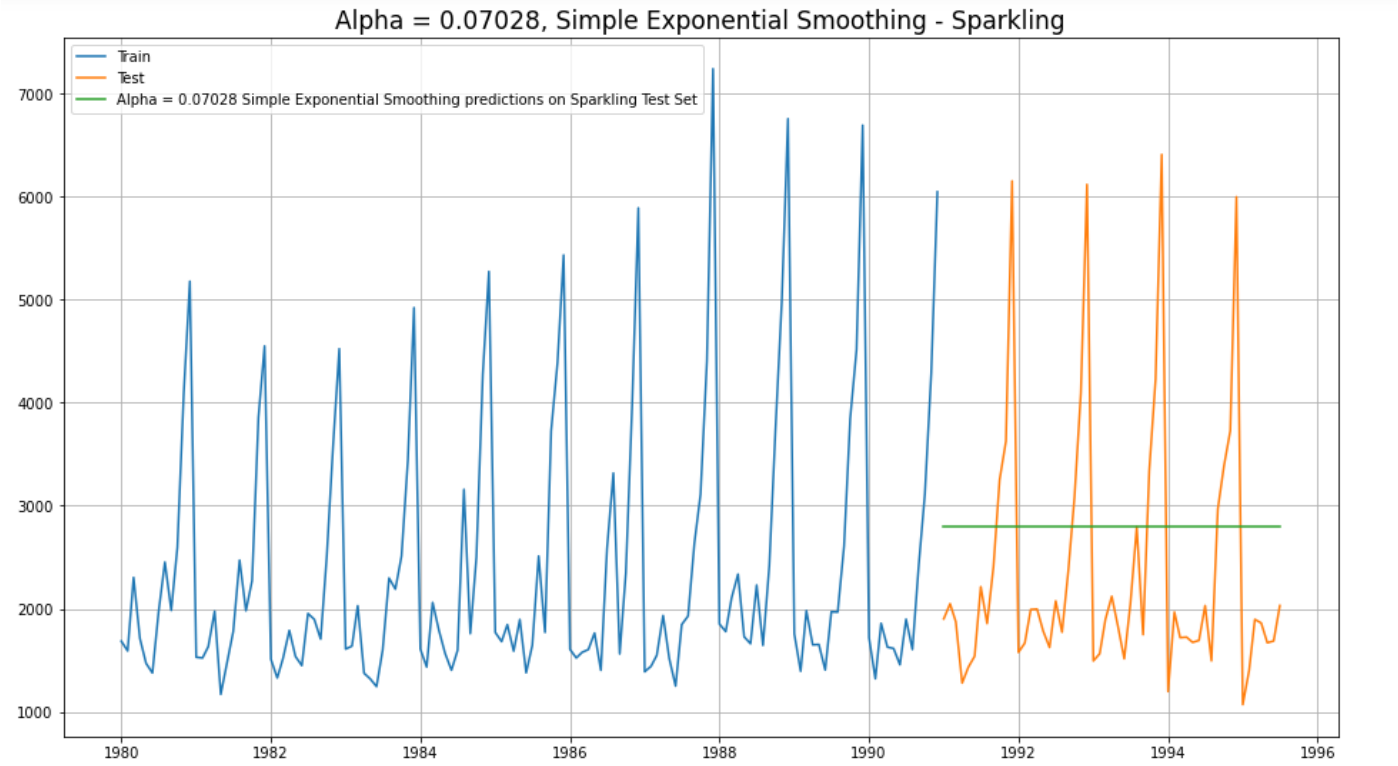
Simple Exponential Smoothing class must be instantiated and passed the training data.

The fit () function is then called providing the fit configuration, the alpha value, smoothing level. If this is omitted or set to None, the model will automatically optimize the value.

Sparkling:

SES - ETS(A, N, N) - Simple Exponential Smoothing with additive errors -

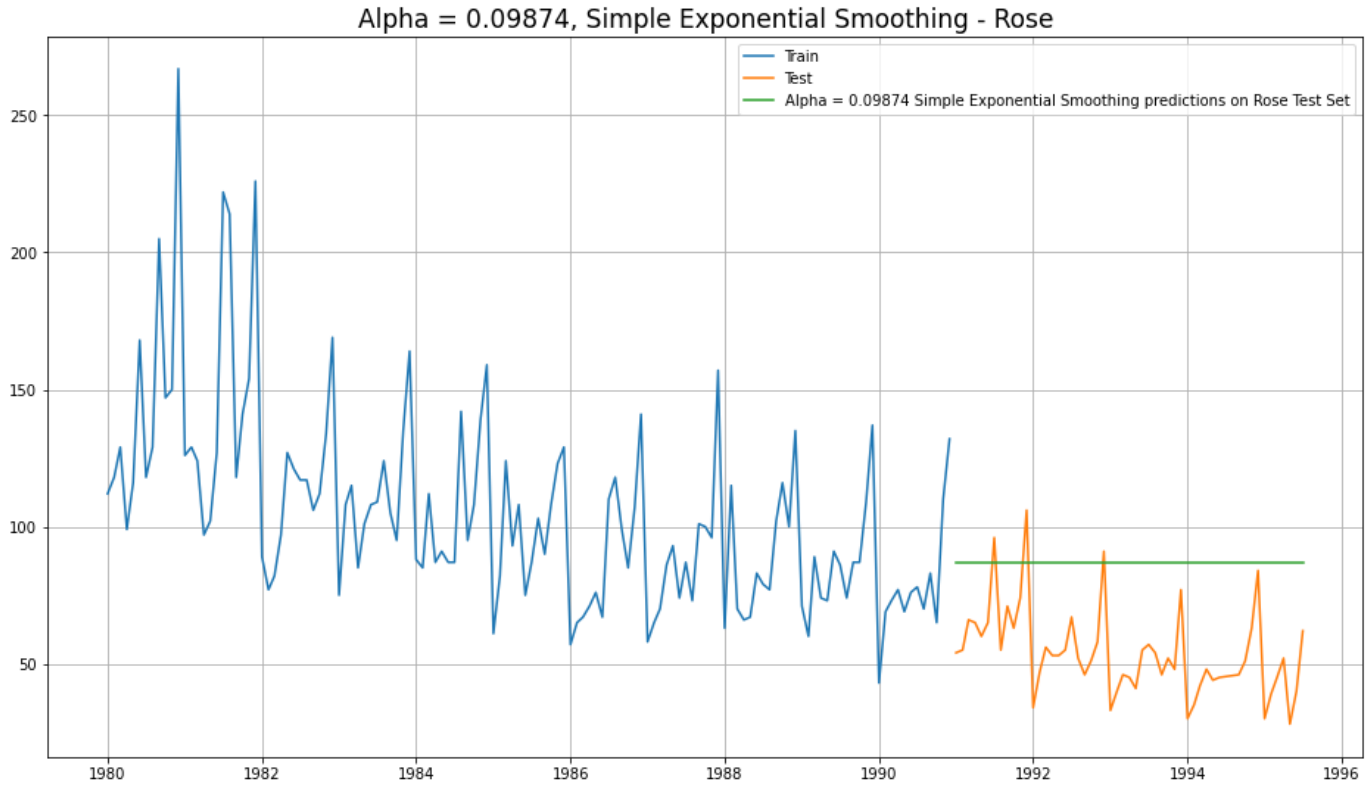
```
{'smoothing_level': 0.07028442075641193,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 1763.8402828521703,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```



Rose:

SES - ETS(A, N, N) - Simple Exponential Smoothing with additive errors

```
{'smoothing_level': 0.09874963957110783,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 134.38708961485827,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```



Model Evaluation:

	Test RMSE Sparkling	Test RMSE Rose
RegressionOnTime	1389.135175	15.268955
NaiveModel	3864.279352	79.718773
SimpleAverageModel	1275.081804	53.460570
2pointTrailingMovingAverage	813.400684	11.529278
4pointTrailingMovingAverage	1156.589694	14.451403
6pointTrailingMovingAverage	1283.927428	14.566327
9pointTrailingMovingAverage	1346.278315	14.727630
Simple Exponential Smoothing	1338.000861	36.796236

Method 6: Double Exponential Smoothing (Holt's Model)

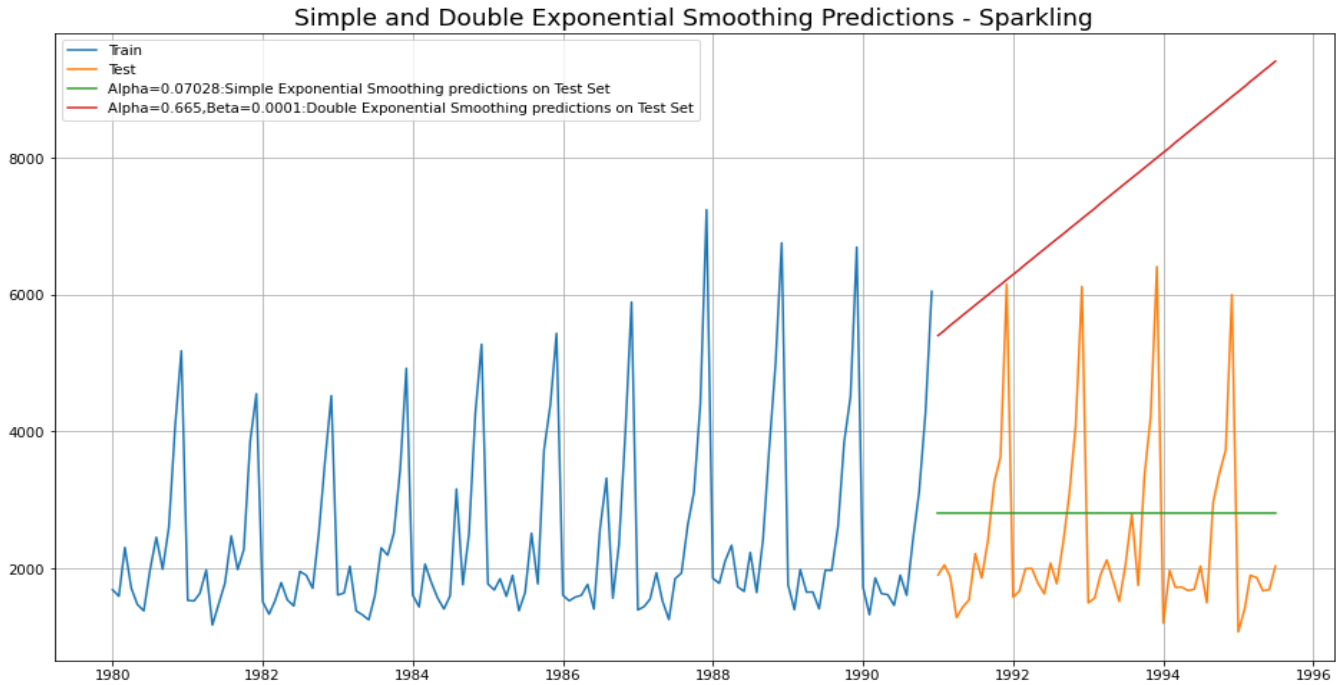
Holt - ETS(A, A, N) - Holt's linear method with additive errors

Double Exponential Smoothing Sparkling:

- One of the drawbacks of the simple exponential smoothing is that the model does not do well in the presence of the trend.
- This model is an extension of SES known as Double Exponential model which estimates two smoothing parameters.
- Applicable when data has Trend but no seasonality.
- Two separate components are considered: Level and Trend.
- Level is the local mean.
- One smoothing parameter α corresponds to the level series
- A second smoothing parameter β corresponds to the trend series

---

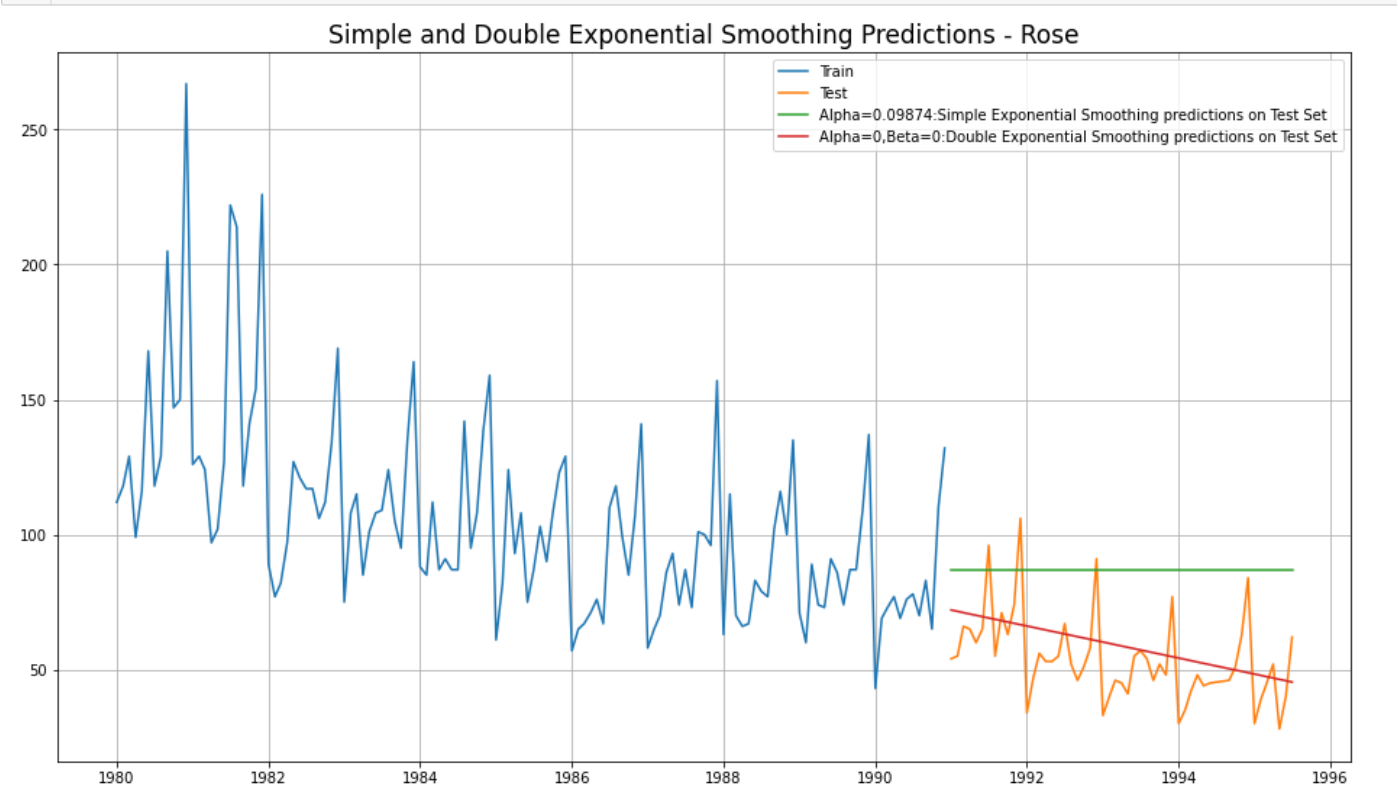
```
Holt model Exponential Smoothing Estimated Parameters :
{'smoothing_level': 0.6649999999999999, 'smoothing_trend': 0.0001, 'smoothing_seasonal': nan, 'damping_trend': nan, 'initial_le
vel': 1502.1999999999991, 'initial_trend': 74.87272727272739, 'initial_seasons': array([], dtype=float64), 'use_boxcox': False,
'lamda': None, 'remove_bias': False}
```



Double Exponential Smoothing Rose:

Holt - ETS(A, A, N) - Holt's linear method with additive errors

```
Holt model  Exponential Smoothing Estimated Parameters :
{'smoothing_level': 1.4901247095597348e-08, 'smoothing_trend': 7.3896641488640725e-09, 'smoothing_seasonal': nan, 'damping_tren
d': nan, 'initial_level': 137.81551313502814, 'initial_trend': -0.4943777717865305, 'initial_seasons': array([], dtype=float6
4), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```



We see that the double exponential smoothing is picking up the trend component along with the level component as well.

Model Evaluation:

	Test RMSE Sparkling	Test RMSE Rose
Double Exponential Smoothing	5291.879833	15.268957

Observations:

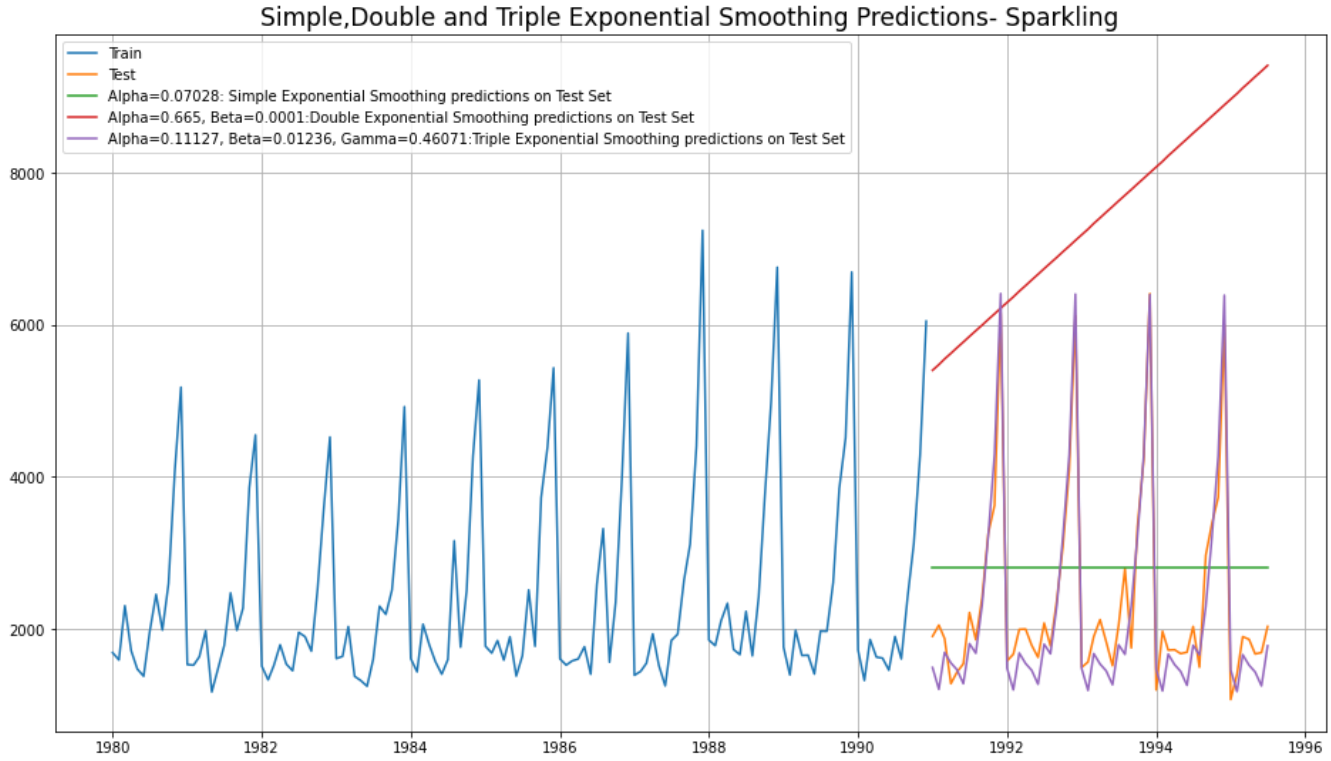
Here, we see that the Double Exponential Smoothing has actually done well when compared to the Simple Exponential Smoothing. This is because of the fact that the Double Exponential Smoothing model has picked up the trend component as well.

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

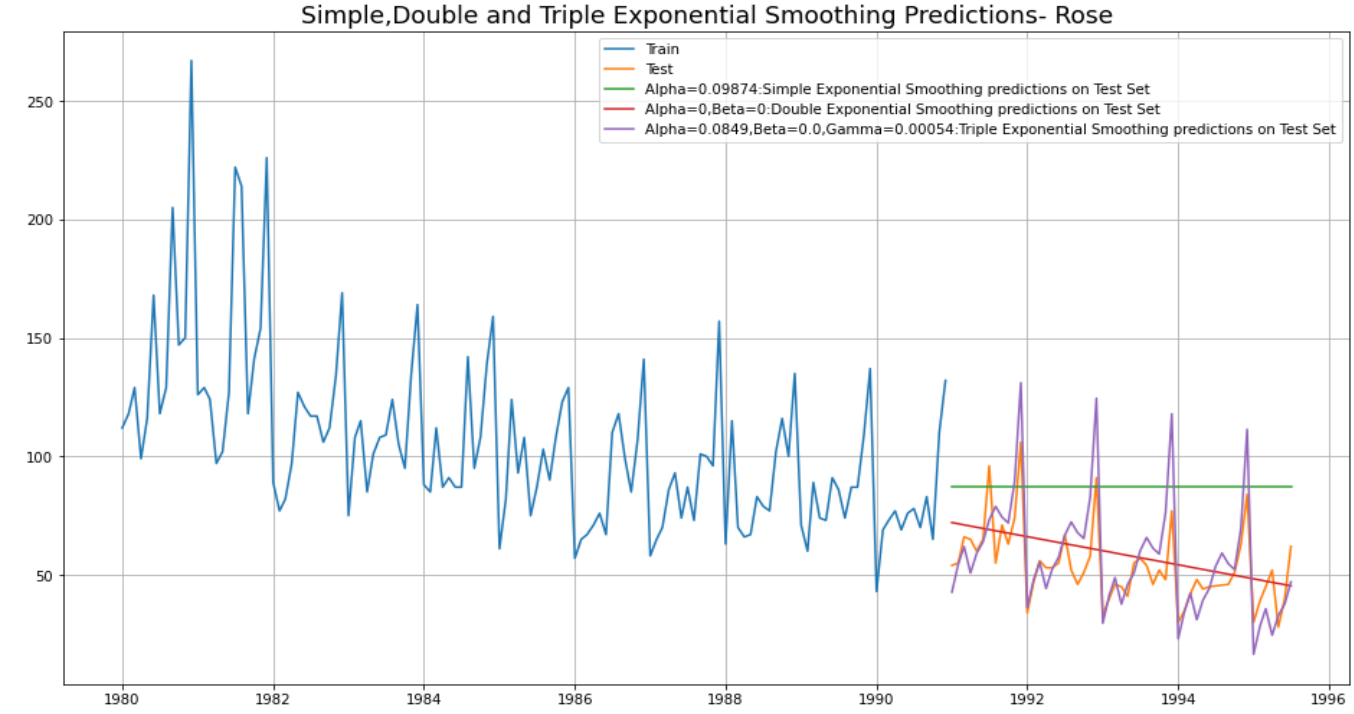
Holt-Winters - ETS(A, A, A) - Holt Winter's linear method with additive errors - Sparkling

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

Triple Exponential Smoothing Sparkling:



Triple Exponential Smoothing Rose:



We see that the Triple Exponential Smoothing is picking up the seasonal component as well.

Model Evaluation:

	Test RMSE Sparkling	Test RMSE Rose
Triple Exponential Smoothing (Additive Season)	378.625883	14.27844

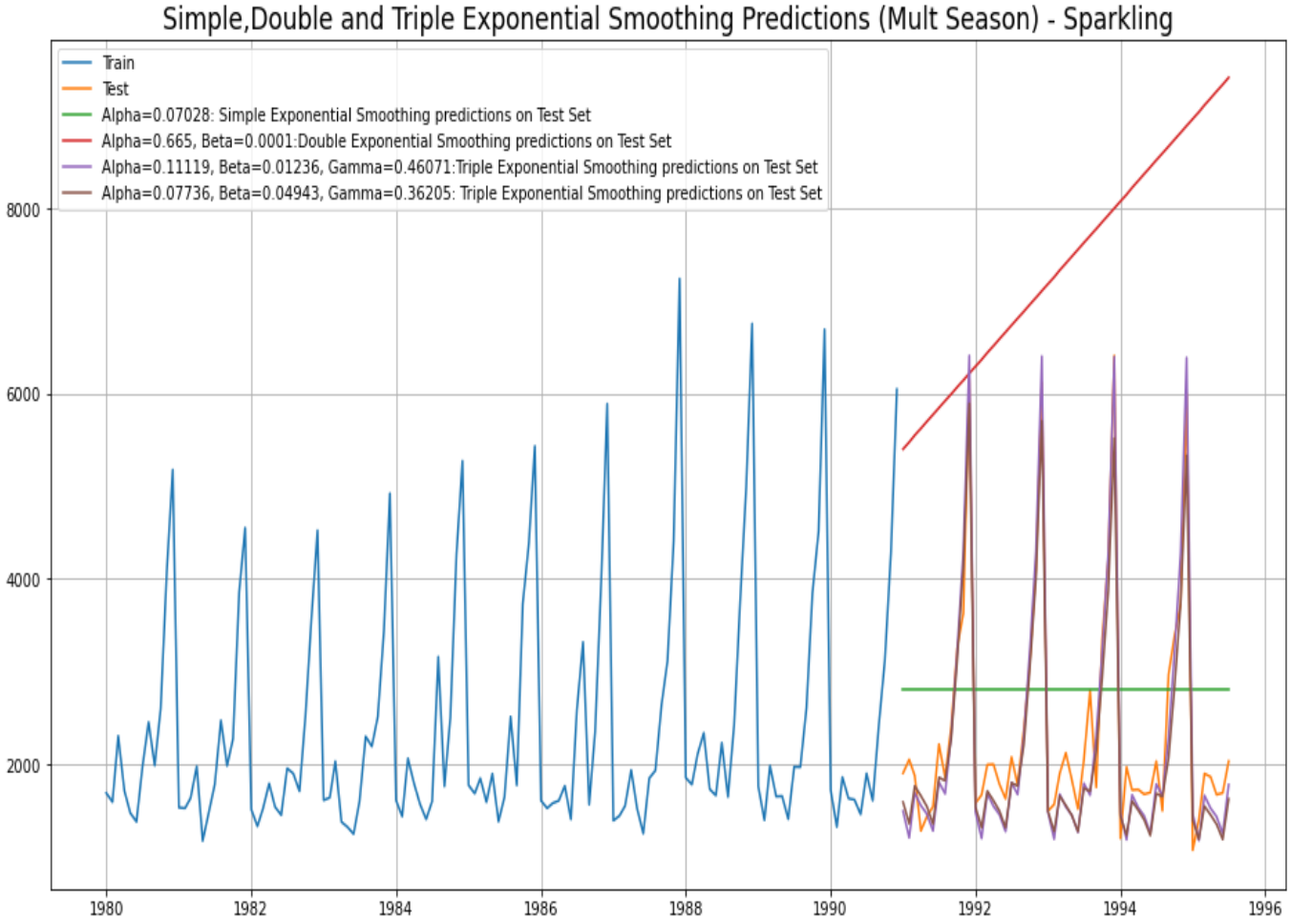
Observations:

Triple Exponential Smoothing has performed the best on the test as expected since the data had both trend and seasonality

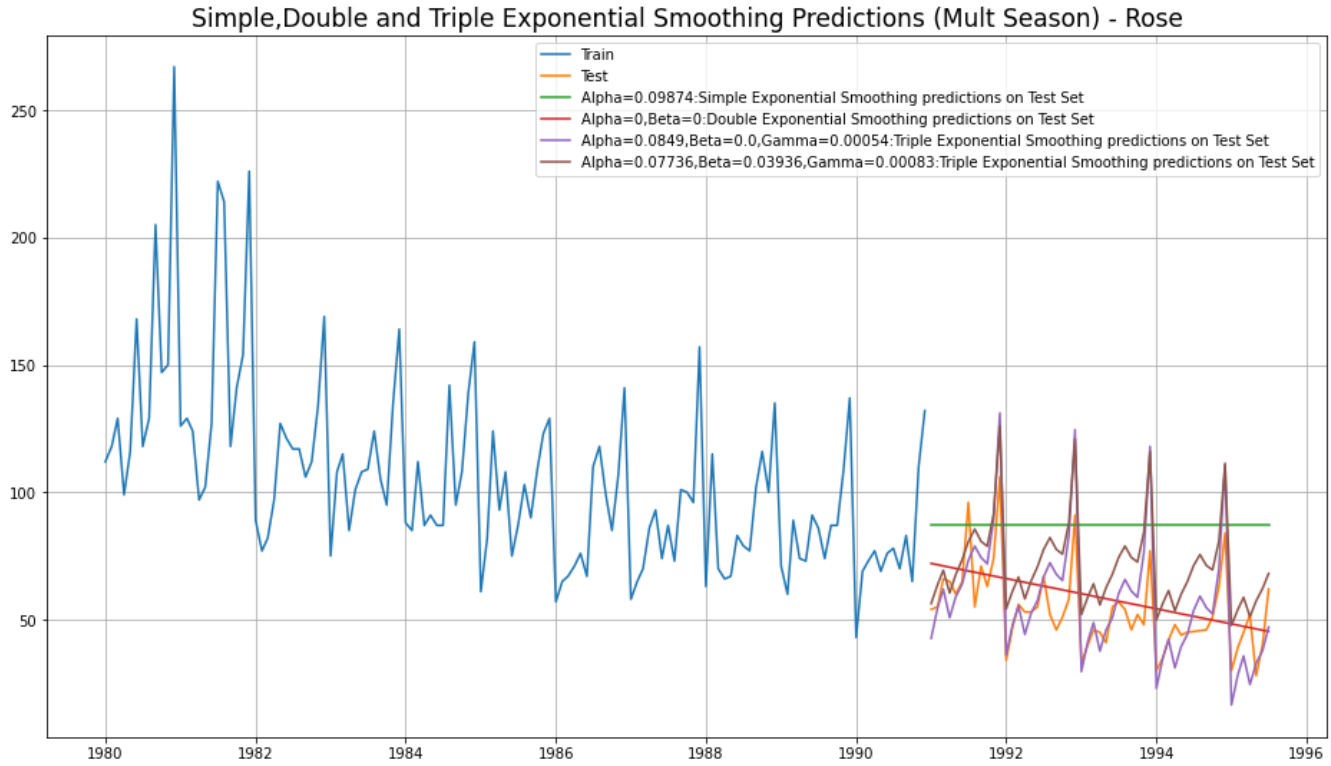
Multiplicative Season:

ETS(A, A, M) model - Taking MULTIPLICATIVE SEASONALITY

Triple Exponential Smoothing Sparkling:



Triple Exponential Smoothing Rose:



Model Evaluation:

	Test RMSE Rose	Test RMSE Sparkling
Triple Exponential Smoothing (Multiplicative Season)	19.11311	403.706228

Observations:

Best Model for Sparkling is Triple Exponential Smoothing (Additive Season)

Best Model for Rose till Now - 2 Pt Moving Average

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.

ADF test is a unit root test which determines whether there is a unit root and subsequently whether the series is non-stationary.

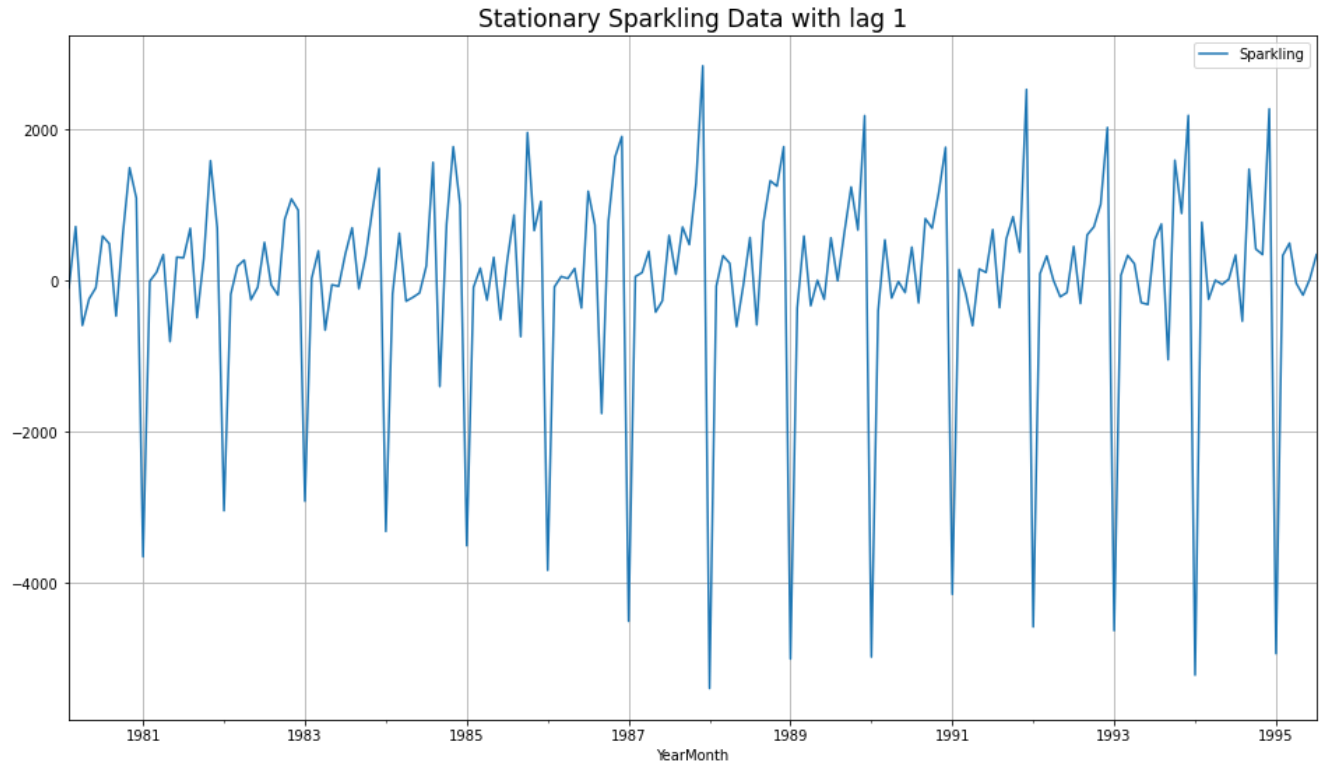
The hypothesis in a simple form for the ADF test is:

H0: The Time Series has a unit root and is thus non-stationary.

H1: The Time Series does not have a unit root and is thus stationary.

We would want the series to be stationary for building ARIMA models and thus we would want the p-valueof this test to be less than the α value. (0.05)

Sparkling:




```
DF test statistic is -44.912
DF test p-value is 0.0
Number of lags used 10
```

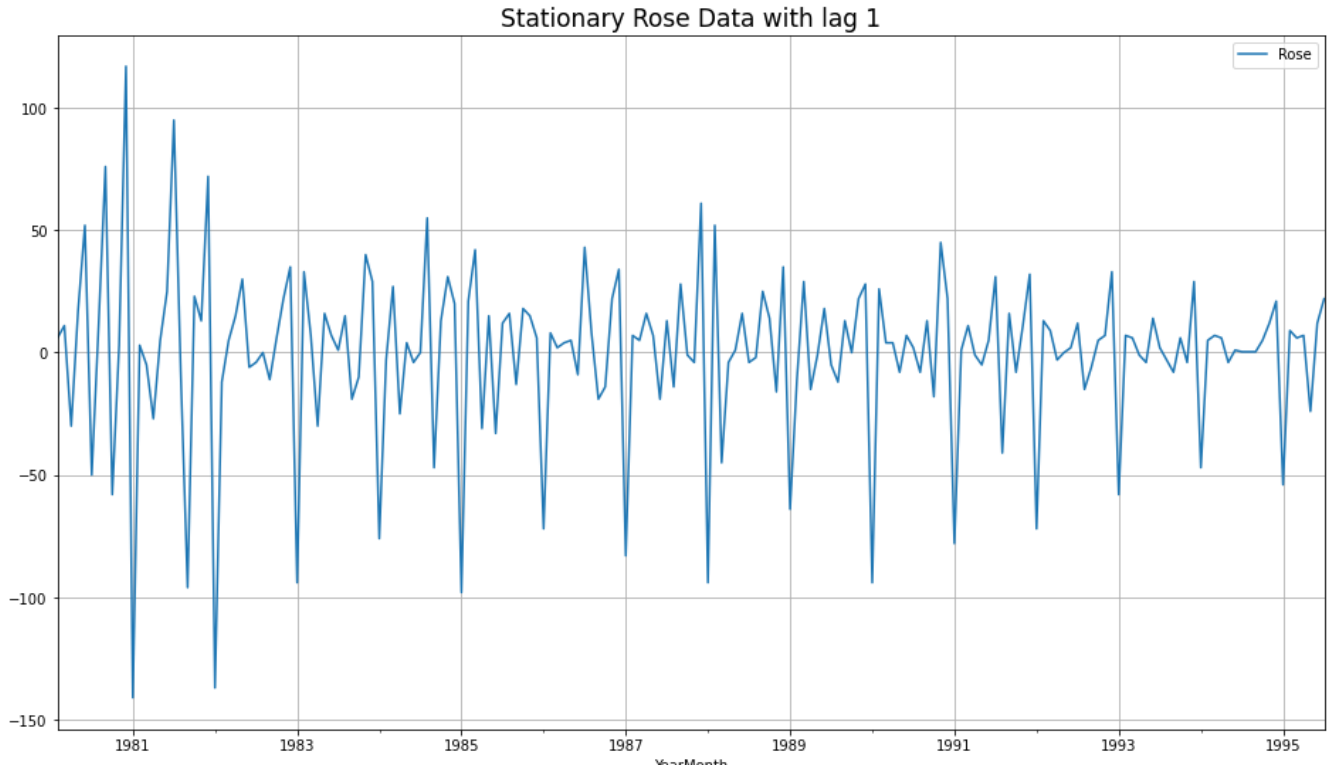
- We see that at 5% significant level the Time Series is non-stationary.

  Let us take one level of differencing to see whether the series becomes stationary.

```
DF test statistic is -44.912
DF test p-value is 0.0
Number of lags used 10
```

- We see that p-value < alpha=0.05 Hence, we reject the Null Hypothesis.
- We conclude that with a lag 1 - now the Sparkling data is Stationary

Rose:




```
DF test statistic is -2.240
DF test p-value is 0.4671371627793168
Number of lags used 13
```

- We see that at 5% significant level the Time Series is non-stationary.

  Let us take one level of differencing to see whether the series becomes stationary.

```
DF test statistic is -8.162
DF test p-value is 3.015976115827353e-11
Number of lags used 12
```

- We see that p-value < alpha=0.05 Hence, we reject the Null Hypothesis.
- We conclude that with a lag 1 - now the Sparkling data is Stationary

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.

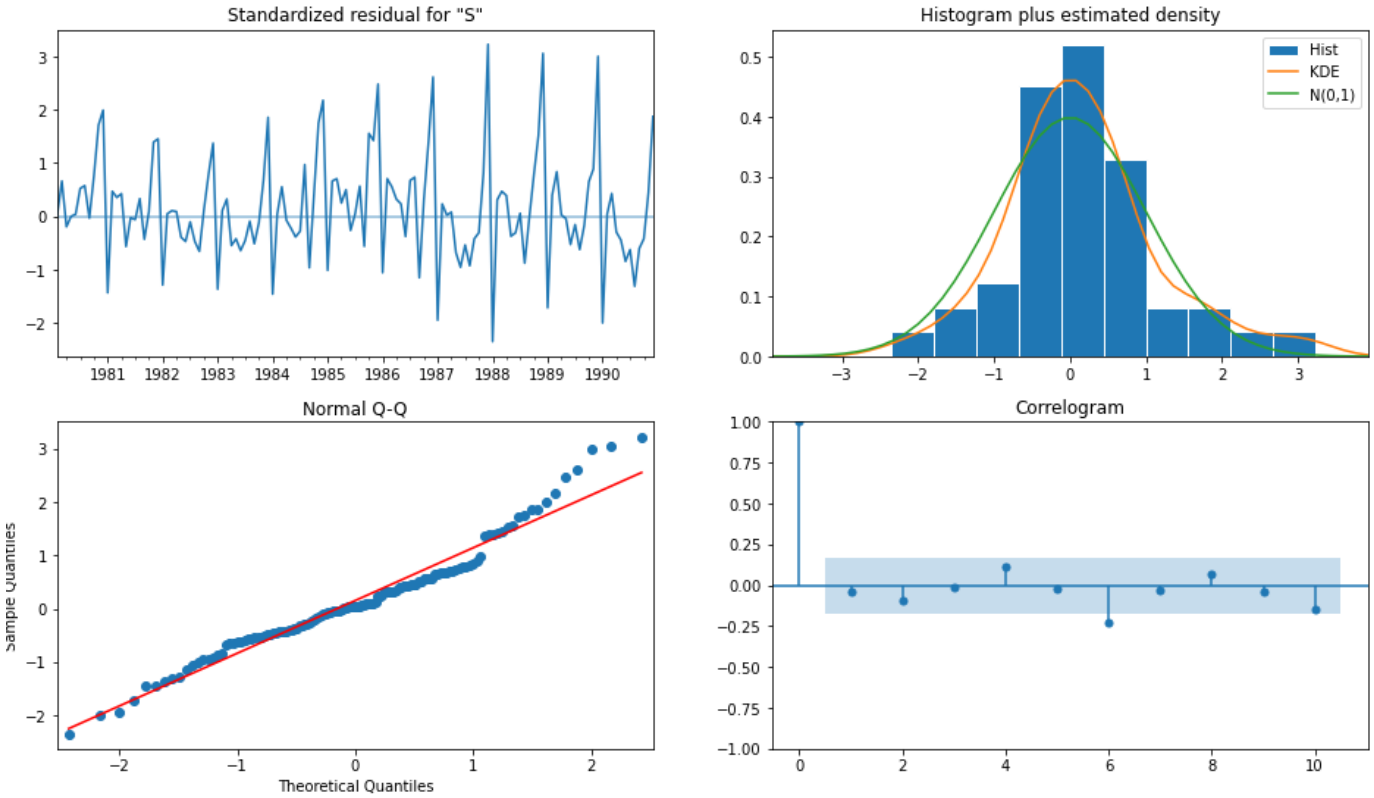
Sparkling:

We ran the automated ARIMA model for Sparkling Sales and sorted the AIC values output from lowest to highest.
We then proceeded to build the ARIMA model with the lowest Akaike Information Criteria and got the Test RMSE score 1299.97

The table showing the AIC values arranged in ascending order with various combinations of p, d and q

	param	AIC
10	(2, 1, 2)	2213.509212
15	(3, 1, 3)	2221.461689
14	(3, 1, 2)	2230.825009
11	(2, 1, 3)	2232.811211
9	(2, 1, 1)	2233.777626
3	(0, 1, 3)	2233.994858
2	(0, 1, 2)	2234.408323
6	(1, 1, 2)	2234.5272
13	(3, 1, 1)	2235.498899
7	(1, 1, 3)	2235.607815
5	(1, 1, 1)	2235.755095
12	(3, 1, 0)	2257.723379
8	(2, 1, 0)	2260.365744
1	(0, 1, 1)	2263.060016
4	(1, 1, 0)	2266.608539
0	(0, 1, 0)	2267.663036

Sparkling data Plot Diagnostics:



SARIMAX Results

Dep. Variable:	Sparkling	No. Observations:	132
Model:	ARIMA(2, 1, 2)	Log Likelihood	-1101.755
Date:	Sun, 09 Apr 2023	AIC	2213.509
Time:	18:41:15	BIC	2227.885
Sample:	01-01-1980 - 12-01-1990	HQIC	2219.351
Covariance Type:	opg		

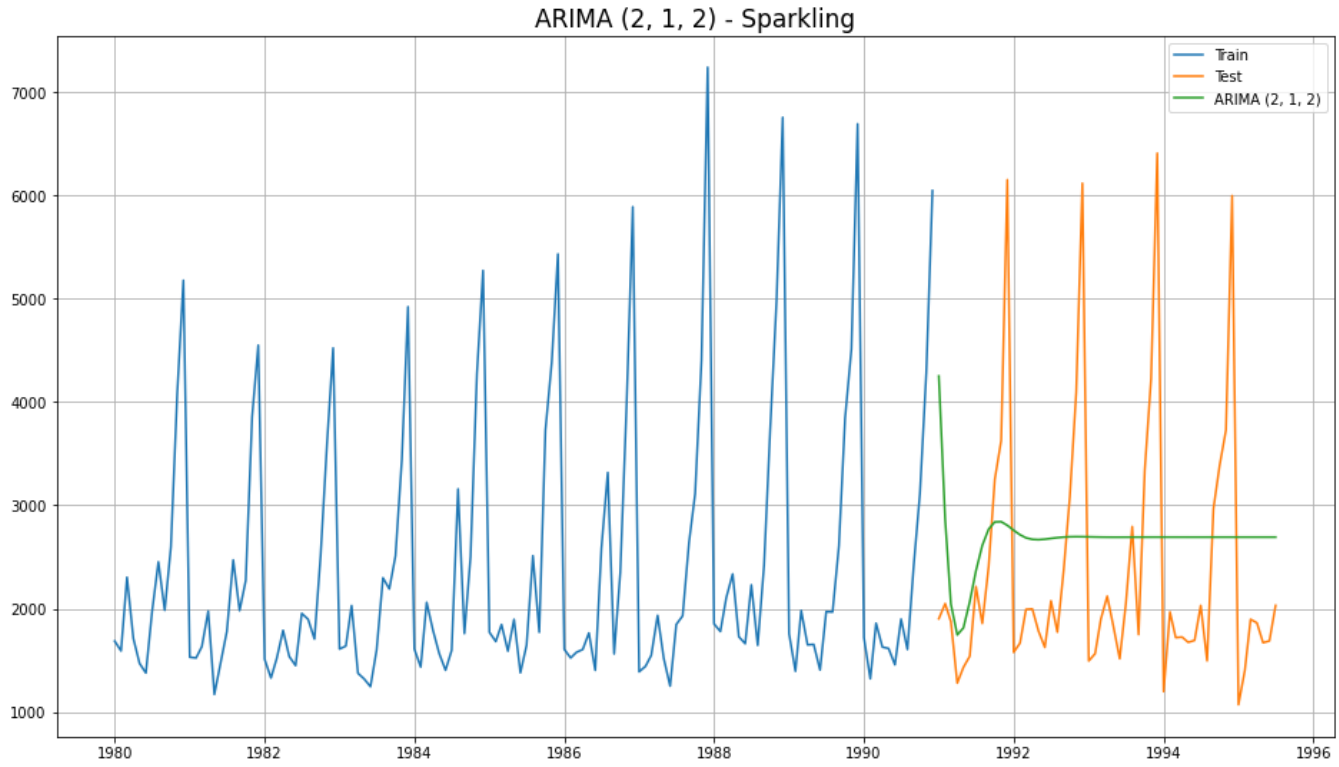
	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	1.3121	0.046	28.782	0.000	1.223	1.401
ar.L2	-0.5593	0.072	-7.740	0.000	-0.701	-0.418
ma.L1	-1.9917	0.109	-18.216	0.000	-2.206	-1.777
ma.L2	0.9999	0.110	9.109	0.000	0.785	1.215
sigma2	1.099e+06	1.99e-07	5.51e+12	0.000	1.1e+06	1.1e+06

Ljung-Box (L1) (Q):	0.19	Jarque-Bera (JB):	14.46
Prob(Q):	0.67	Prob(JB):	0.00
Heteroskedasticity (H):	2.43	Skew:	0.61
Prob(H) (two-sided):	0.00	Kurtosis:	4.08

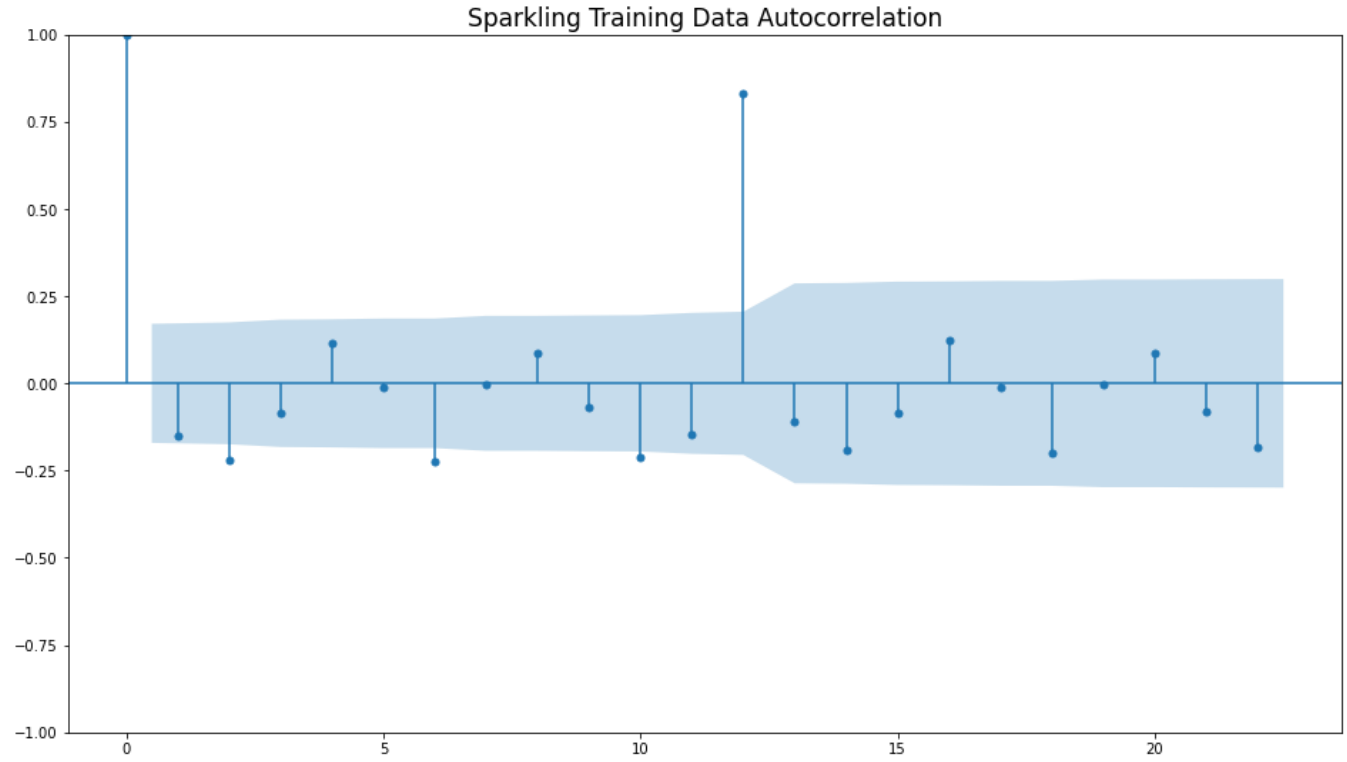
Model Valuation:

	RMSE
ARIMA(2,1,2)	1299.979821

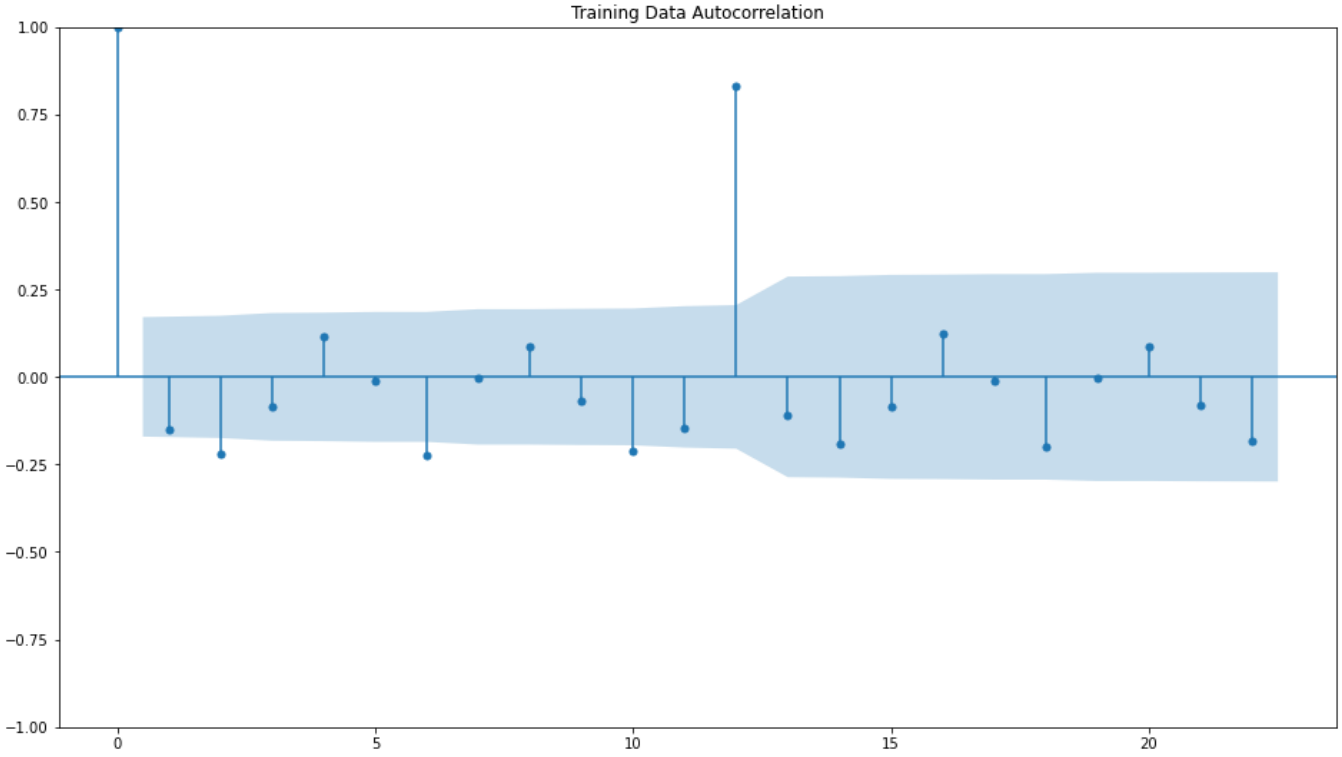
ARIMA (2, 1, 2) – Sparkling:



Sparkling Training Data Autocorrelation



SARIMA - SPARKLING DATA SET




```
Examples of the parameter combinations for the Model are
Model: (0, 1, 1)(0, 0, 1, 12)
Model: (0, 1, 2)(0, 0, 2, 12)
Model: (0, 1, 3)(0, 0, 3, 12)
Model: (1, 1, 0)(1, 0, 0, 12)
Model: (1, 1, 1)(1, 0, 1, 12)
Model: (1, 1, 2)(1, 0, 2, 12)
Model: (1, 1, 3)(1, 0, 3, 12)
Model: (2, 1, 0)(2, 0, 0, 12)
Model: (2, 1, 1)(2, 0, 1, 12)
Model: (2, 1, 2)(2, 0, 2, 12)
Model: (2, 1, 3)(2, 0, 3, 12)
Model: (3, 1, 0)(3, 0, 0, 12)
Model: (3, 1, 1)(3, 0, 1, 12)
Model: (3, 1, 2)(3, 0, 2, 12)
Model: (3, 1, 3)(3, 0, 3, 12)
```

SARIMAX Results

Dep. Variable:	Sparkling	No. Observations:	132
Model:	SARIMAX(3, 1, 1)x(3, 0, [], 12)	Log Likelihood	-685.894
Date:	Sun, 09 Apr 2023	AIC	1387.788
Time:	18:47:26	BIC	1407.963
Sample:	01-01-1980 - 12-01-1990	HQIC	1395.931
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.1615	0.150	1.075	0.282	-0.133	0.456
ar.L2	-0.0928	0.150	-0.618	0.537	-0.388	0.202
ar.L3	0.0916	0.136	0.676	0.499	-0.174	0.357
ma.L1	-0.9195	0.092	-10.033	0.000	-1.099	-0.740
ar.S.L12	0.5805	0.104	5.575	0.000	0.376	0.785
ar.S.L24	0.2559	0.119	2.159	0.031	0.024	0.488
ar.S.L36	0.2132	0.121	1.761	0.078	-0.024	0.451
sigma2	1.729e+05	2.18e+04	7.940	0.000	1.3e+05	2.16e+05

Ljung-Box (L1) (Q):	0.02	Jarque-Bera (JB):	18.78
Prob(Q):	0.88	Prob(JB):	0.00
Heteroskedasticity (H):	1.08	Skew:	0.47
Prob(H) (two-sided):	0.84	Kurtosis:	5.00

Predict on the Test Set using this model and evaluate the model.

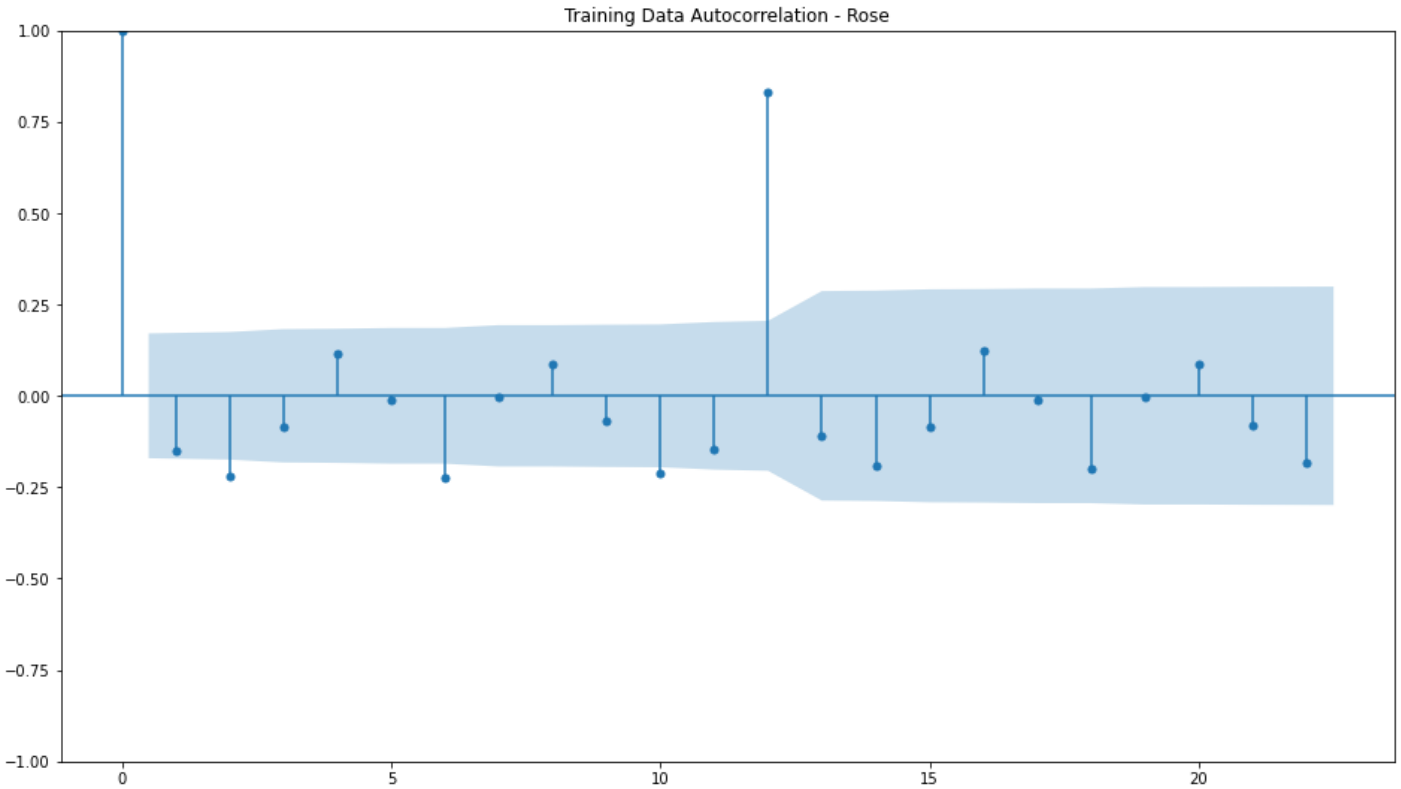
Sparkling	mean	mean_se	mean_ci_lower	mean_ci_upper
1991-01-01	1389.352969	415.861319	574.279761	2204.426178
1991-02-01	1224.672450	427.866006	386.070488	2063.274412
1991-03-01	1673.337773	428.010513	834.452583	2512.222963
1991-04-01	1533.304740	432.773829	685.083621	2381.525859
1991-05-01	1425.949551	435.887629	571.625497	2280.273605

	RMSE
ARIMA(2,1,2)	1299.979821
ARIMA(0,1,0)	3864.279352
SARIMA(3,1,1)(3,0,2,12)	601.243362

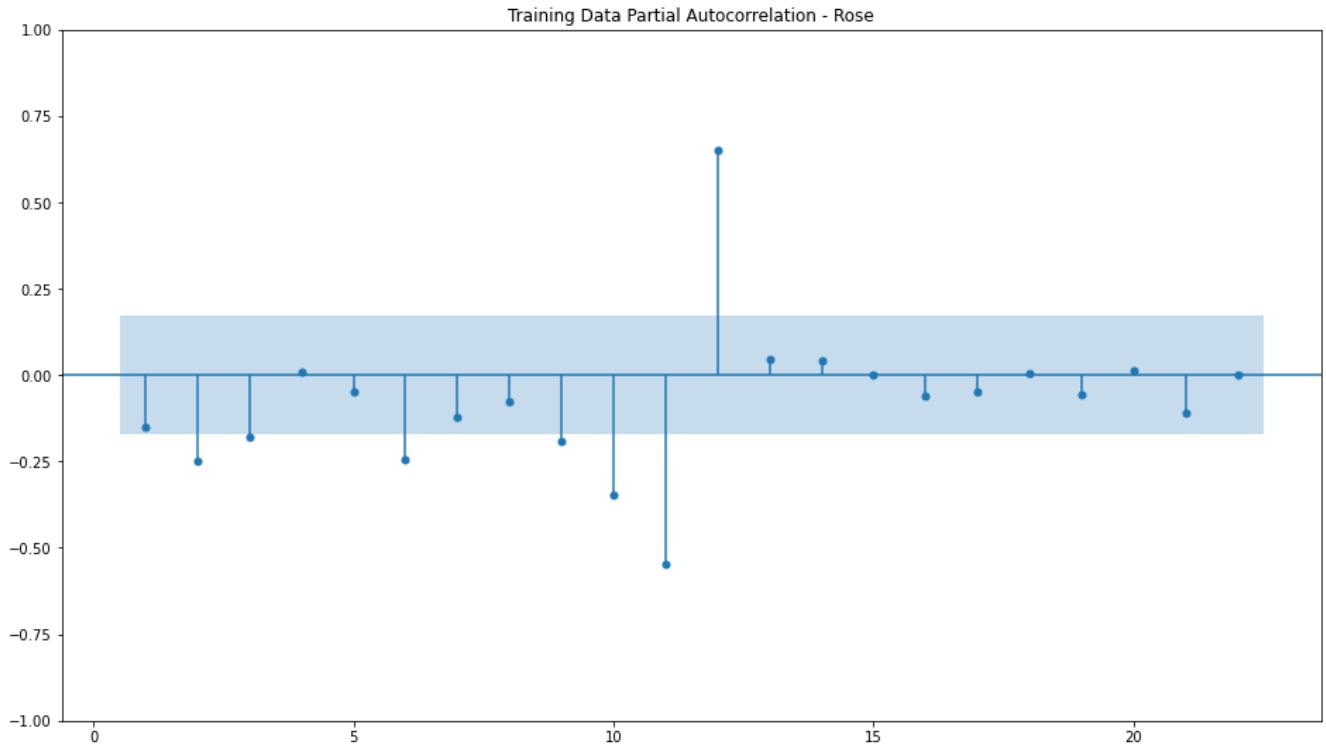
8. Based on the model-building exercise, build the most optimum model(s) on the complete data and predict 12 months into the future with appropriate confidence intervals/bands.

We observe the ACF plot for Rose Sales and observe seasonality at intervals 12, hence we run the Automated SARIMA models at seasonality 12

Auto Corelation:



Partial Auto Corelation



SARIMAX Results

Dep. Variable:	Sparkling	No. Observations:	132
Model:	SARIMAX(0, 1, 0)x(2, 1, [1, 2], 12)	Log Likelihood	-722.996
Date:	Sun, 09 Apr 2023	AIC	1455.991
Time:	18:47:27	BIC	1468.708
Sample:	01-01-1980 - 12-01-1990	HQIC	1461.128
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.S.L12	-0.2445	0.879	-0.278	0.781	-1.967	1.478
ar.S.L24	-0.2107	0.257	-0.820	0.412	-0.714	0.293
ma.S.L12	-0.1220	0.860	-0.142	0.887	-1.807	1.563
ma.S.L24	0.0444	0.502	0.088	0.930	-0.940	1.029
sigma2	2.806e+05	3.2e+04	8.764	0.000	2.18e+05	3.43e+05

Ljung-Box (L1) (Q):	12.20	Jarque-Bera (JB):	37.03
Prob(Q):	0.00	Prob(JB):	0.00
Heteroskedasticity (H):	0.76	Skew:	0.77
Prob(H) (two-sided):	0.44	Kurtosis:	5.66

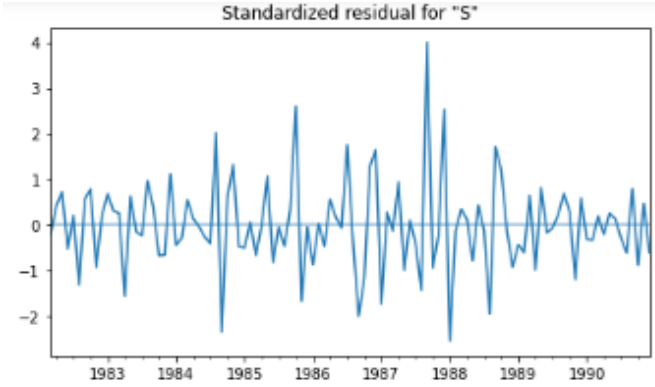
SARIMAX Results

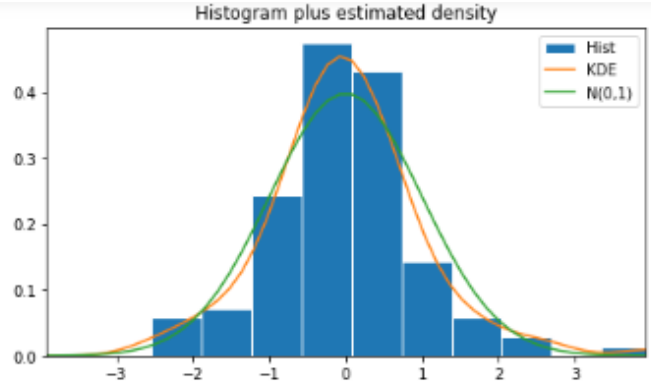
Dep. Variable:	Sparkling	No. Observations:	132
Model:	SARIMAX(0, 1, 0)x(3, 1, [1, 2], 12)	Log Likelihood	-638.304
Date:	Sun, 09 Apr 2023	AIC	1288.607
Time:	18:47:29	BIC	1303.120
Sample:	01-01-1980 - 12-01-1990	HQIC	1294.438
Covariance Type:	opg		

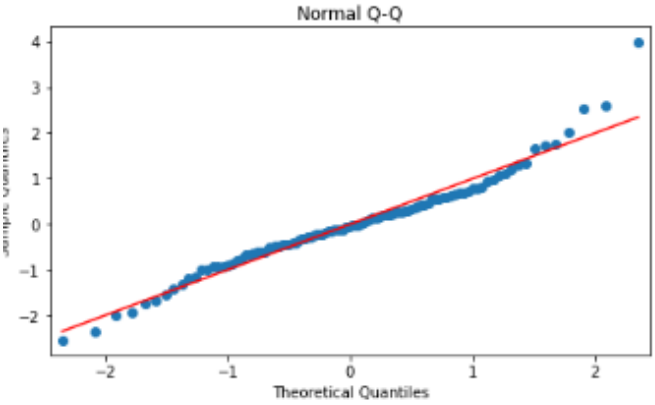
	coef	std err	z	P>|z|	[0.025	0.975]
ar.S.L12	-1.0545	0.201	-5.254	0.000	-1.448	-0.661
ar.S.L24	-0.9168	0.187	-4.913	0.000	-1.283	-0.551
ar.S.L36	-0.2828	0.128	-2.202	0.028	-0.534	-0.031
ma.S.L12	0.8582	0.339	2.533	0.011	0.194	1.522
ma.S.L24	0.8162	0.496	1.646	0.100	-0.156	1.788
sigma2	2.363e+05	9.04e+04	2.613	0.009	5.91e+04	4.14e+05

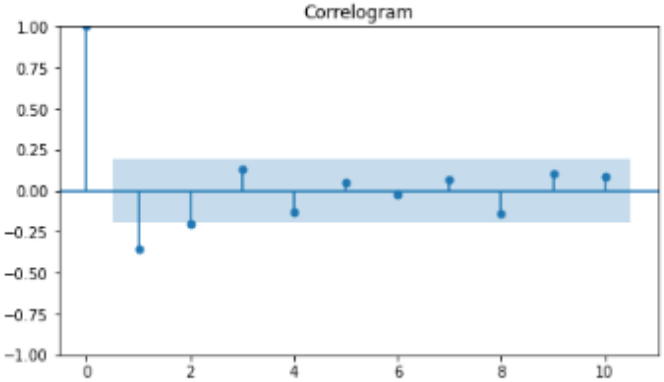
Ljung-Box (L1) (Q):	9.10	Jarque-Bera (JB):	48.00
Prob(Q):	0.00	Prob(JB):	0.00
Heteroskedasticity (H):	0.58	Skew:	1.01
Prob(H) (two-sided):	0.15	Kurtosis:	6.13

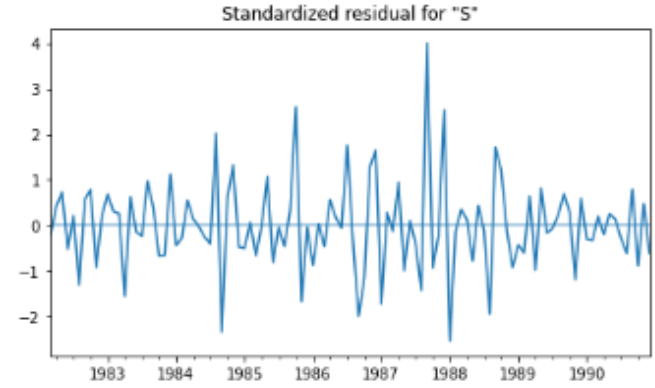
Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

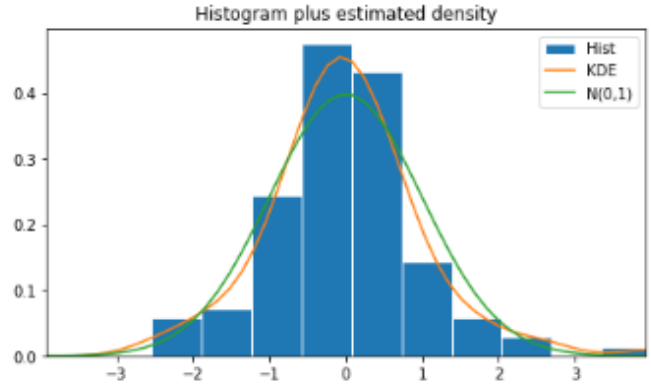

Standardized residual for "S"
1983
1984
1985
1986
1987
1988
1989
1990

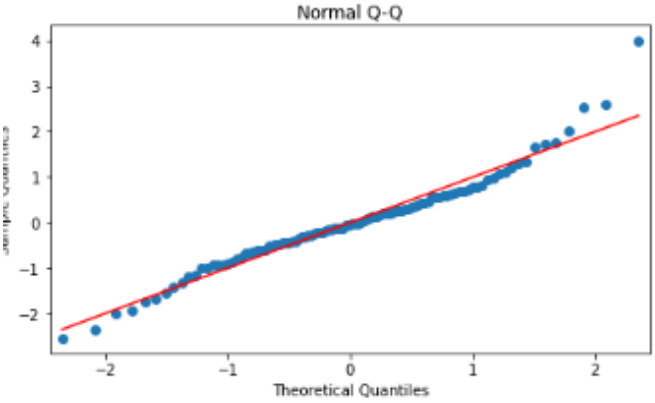

Histogram plus estimated density
Hist
KDE
N(0,1)

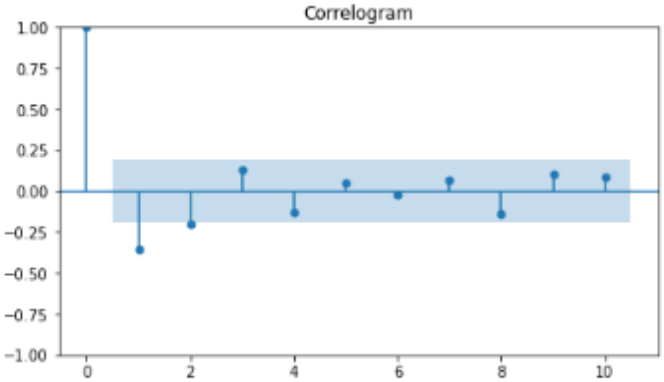

Normal Q-Q
Theoretical Quantiles


Correlogram


Standardized residual for "S"
1983
1984
1985
1986
1987
1988
1989
1990


Histogram plus estimated density
Hist
KDE
N(0,1)


Normal Q-Q
Theoretical Quantiles


Correlogram

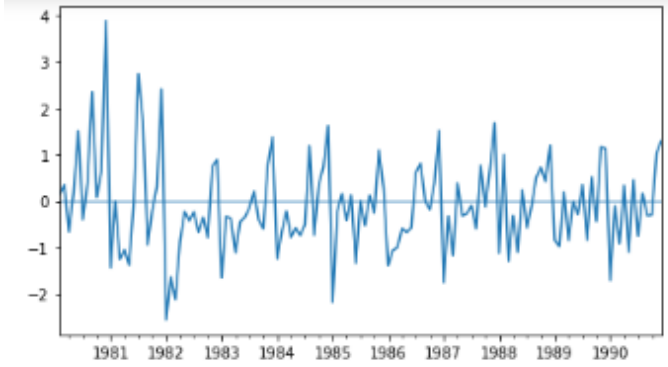
Observations:

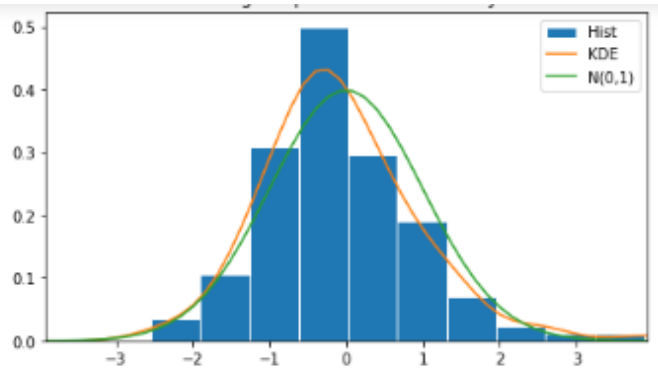
- Model diagnostics confirms that the model residuals are normally distributed.
- Standardized residual- Do not display any obvious seasonality
- Histogram plus estimated density - The KDE plot of the residuals is similar with the normal distribution, hence the model residuals are normally distributed Normal Q-Q plot
-
- There is an ordered distribution of residuals (blue dots) following the linear trend of the samples taken from a standard normal distribution with N(0, 1)Correlogram.
- The time series residuals have low correlation with lagged versions of itself.

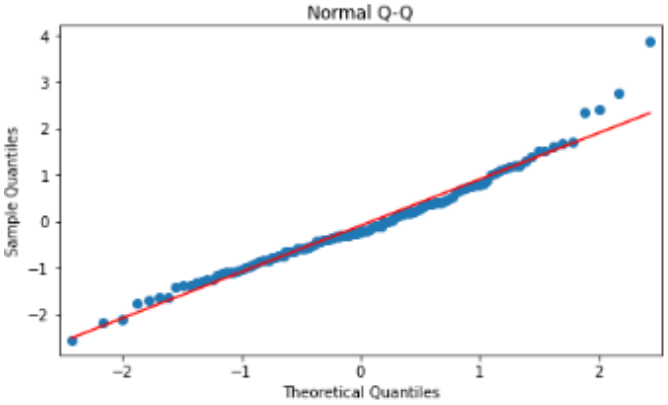
Rose:

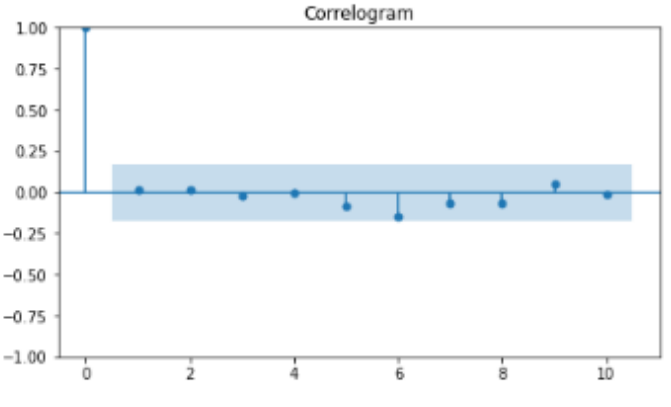
```
Examples of the parameter combinations for the Model
Model: (0, 1, 0)
Model: (0, 1, 1)
Model: (0, 1, 2)
Model: (0, 1, 3)
Model: (1, 1, 0)
Model: (1, 1, 1)
Model: (1, 1, 2)
Model: (1, 1, 3)
Model: (2, 1, 0)
Model: (2, 1, 1)
Model: (2, 1, 2)
Model: (2, 1, 3)
Model: (3, 1, 0)
Model: (3, 1, 1)
Model: (3, 1, 2)
Model: (3, 1, 3)
```

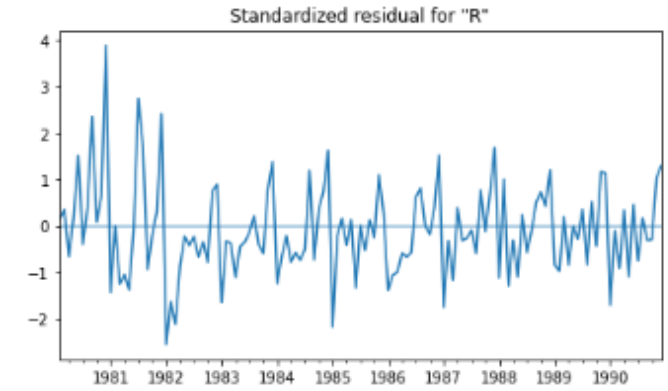
	param	AIC
11	(2, 1, 3)	1274.695356
15	(3, 1, 3)	1278.661965
2	(0, 1, 2)	1279.671529
6	(1, 1, 2)	1279.870723
3	(0, 1, 3)	1280.545376

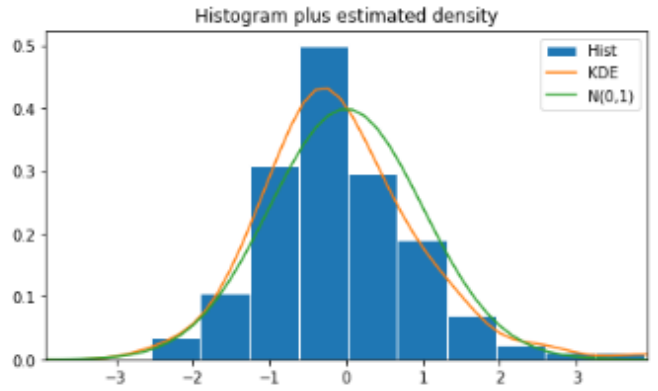

4
3
2
1
0
−1
−2
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990

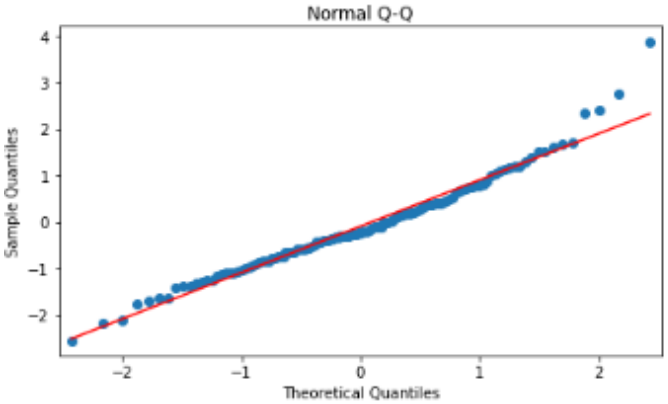

Hist
KDE
N(0,1)
0.5
0.4
0.3
0.2
0.1
0.0
−3
−2
−1
0
1
2
3

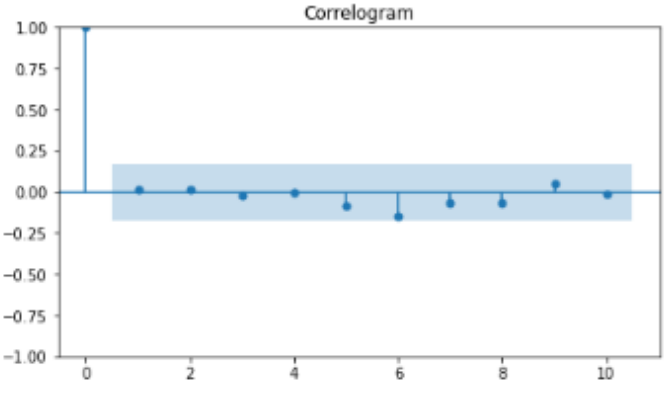

Normal Q-Q
Sample Quantiles
Theoretical Quantiles
4
3
2
1
0
−1
−2
−2
−1
0
1
2

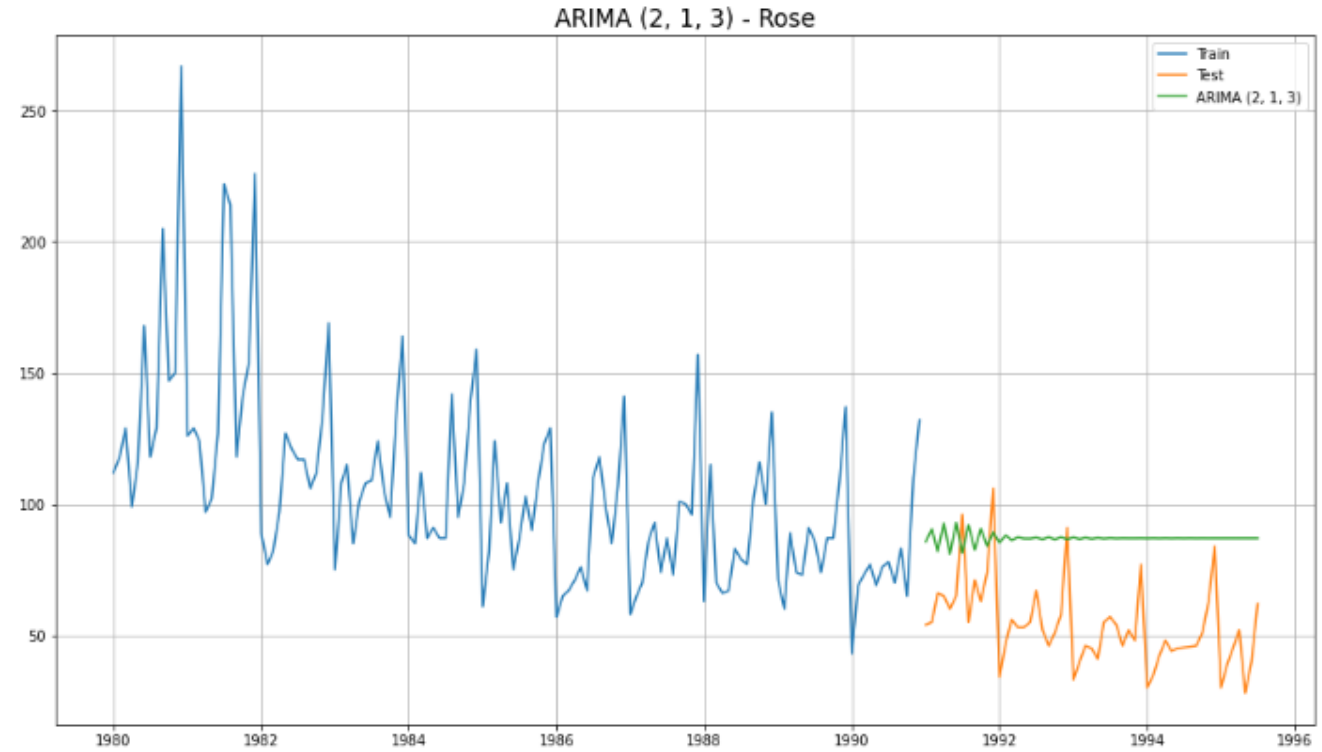

Correlogram
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
2
4
6
8
10

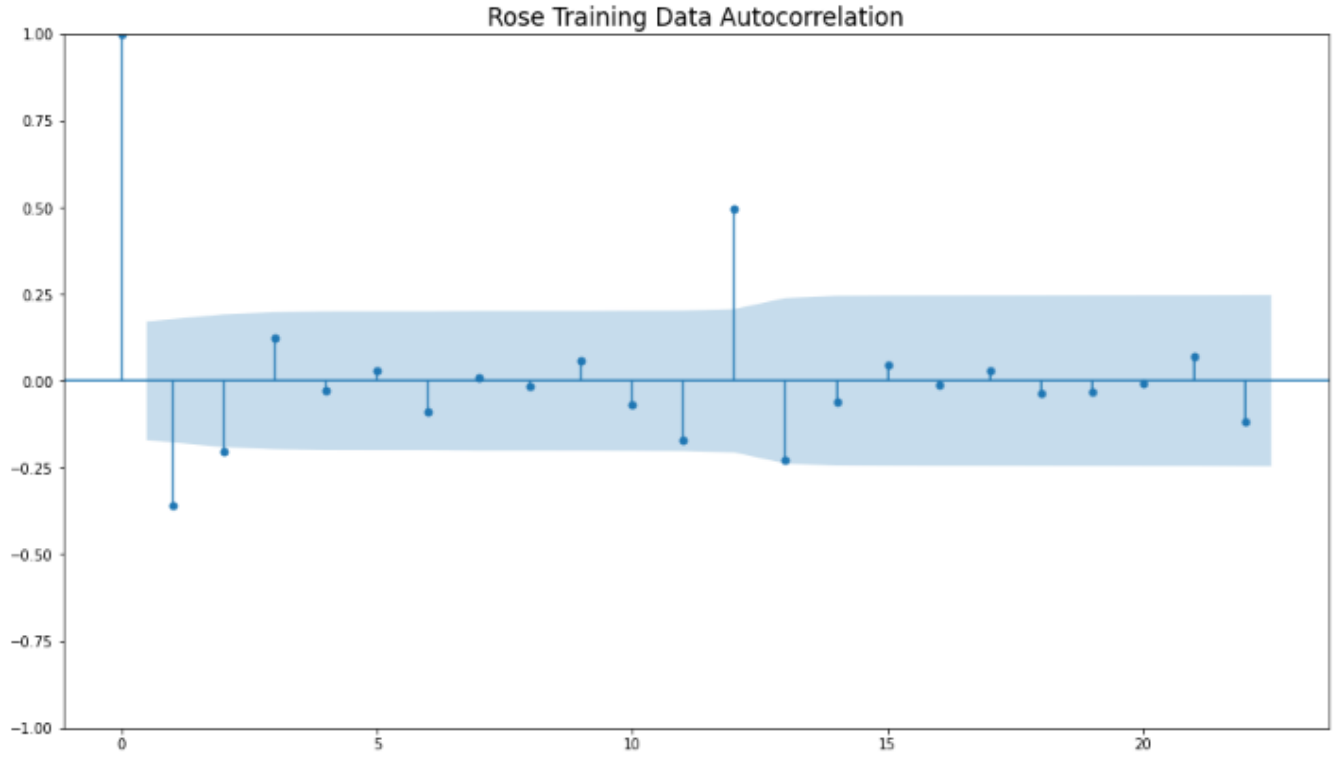

Standardized residual for "R"
4
3
2
1
0
−1
−2
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990

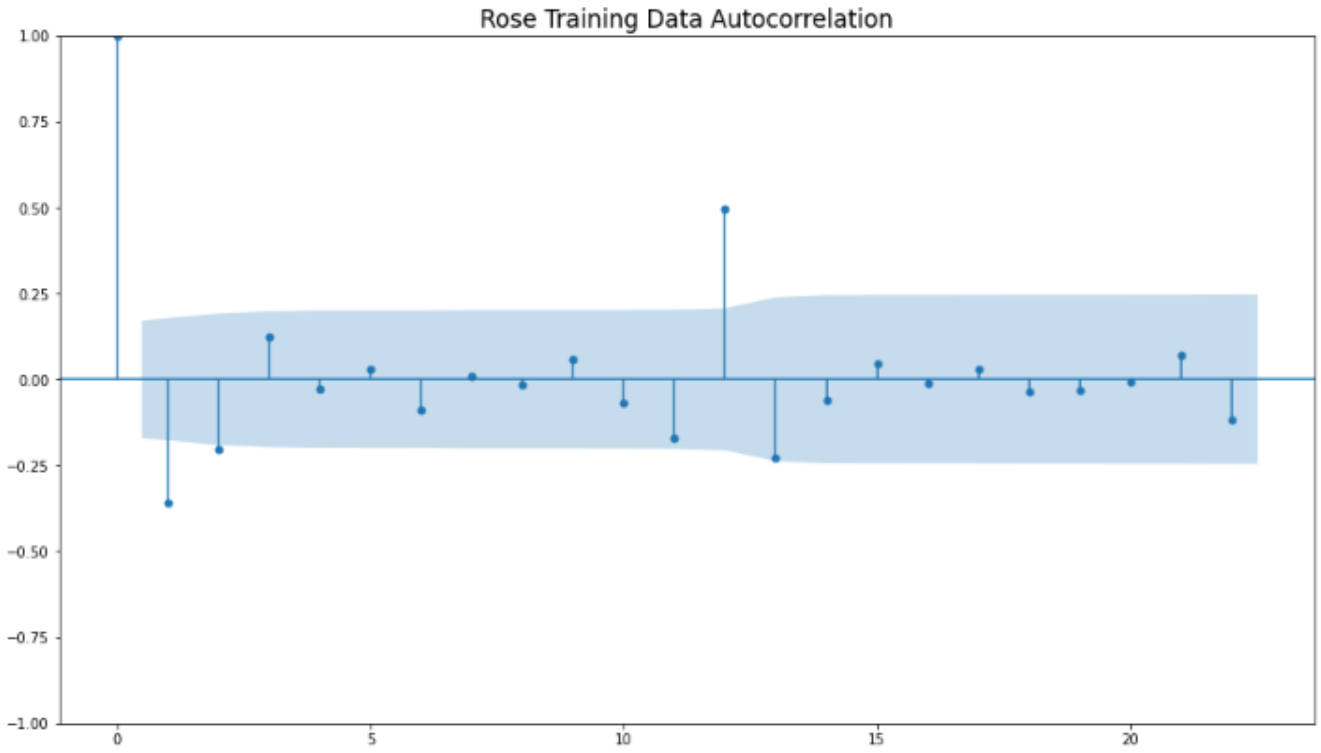

Histogram plus estimated density
Hist
KDE
N(0,1)
0.5
0.4
0.3
0.2
0.1
0.0
−3
−2
−1
0
1
2
3


Normal Q-Q
Sample Quantiles
Theoretical Quantiles
4
3
2
1
0
−1
−2
−2
−1
0
1
2


Correlogram
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
2
4
6
8
10

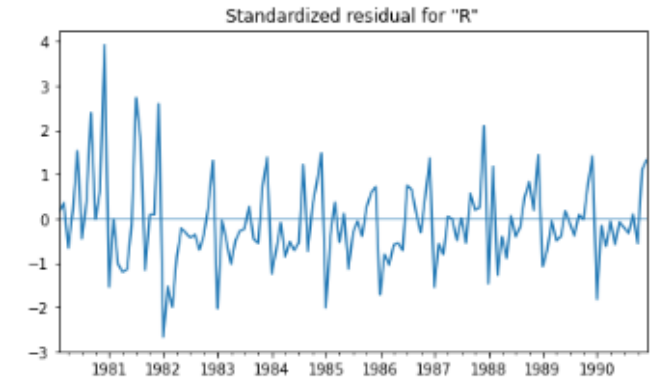

ARIMA (2, 1, 3) - Rose
Train
Test
ARIMA (2, 1, 3)
250
200
150
100
50
1980
1982
1984
1986
1988
1990
1992
1994
1996

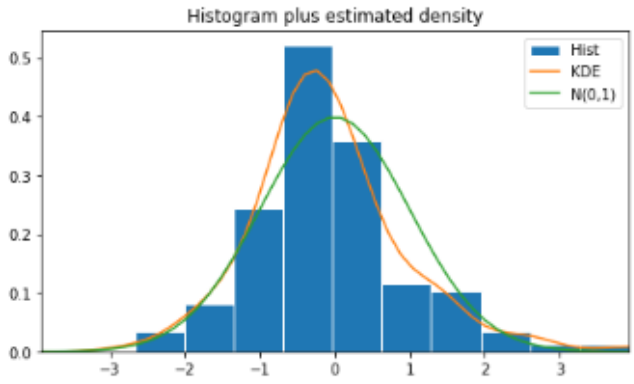

Rose Training Data Autocorrelation
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
0
5
10
15
20

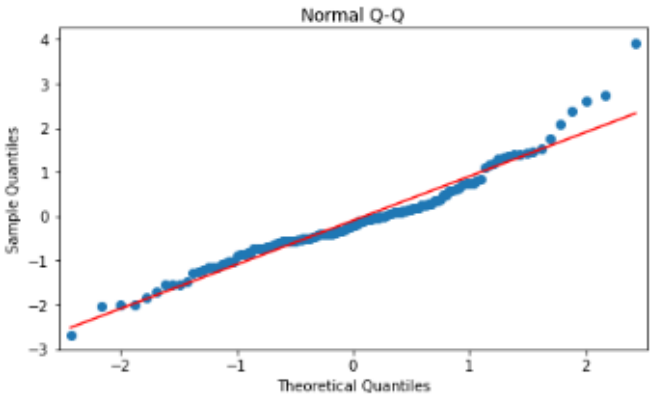


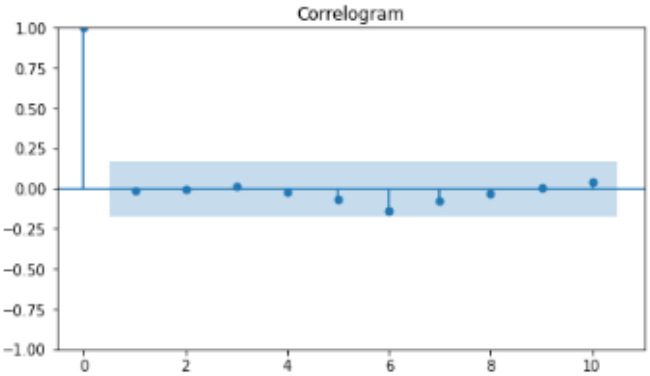
Here, we have taken alpha=0.05.

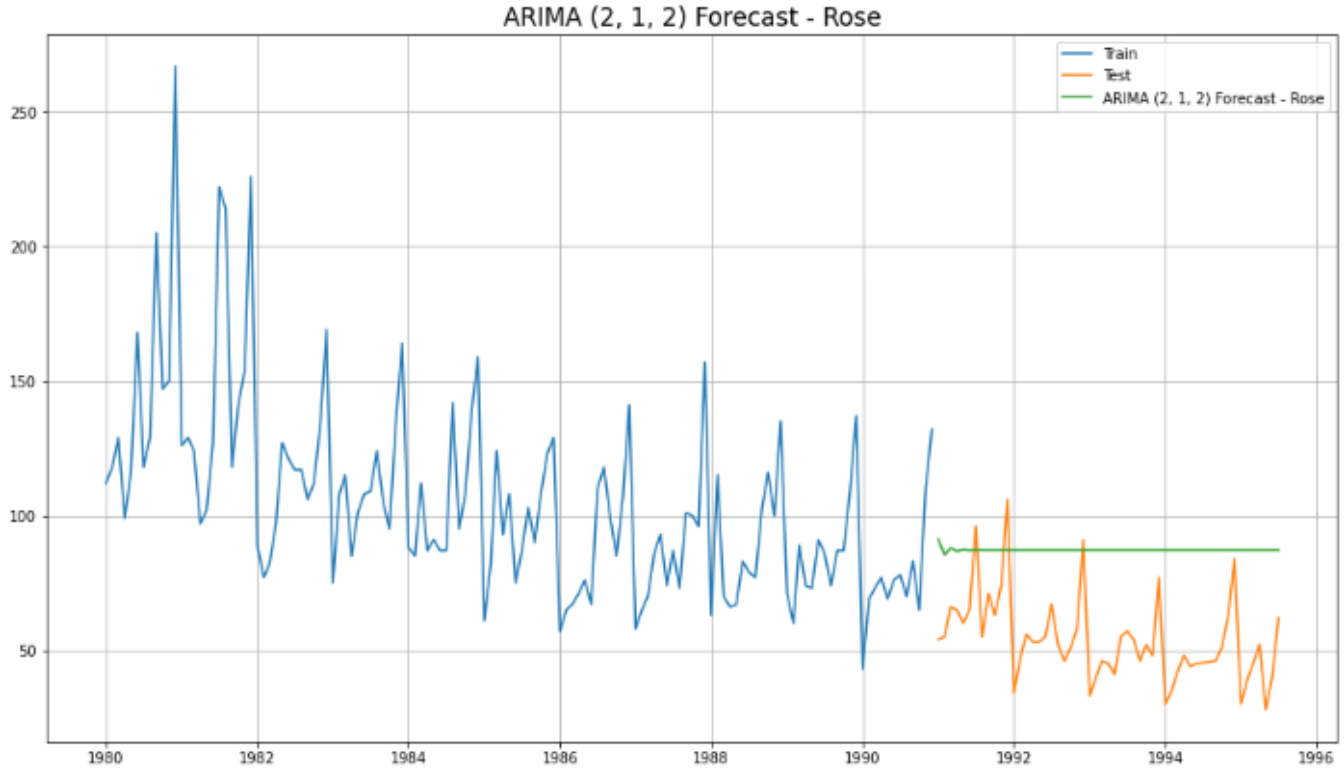
- The Auto-Regressive parameter in an ARIMA model is 'p' which comes from the significant lag before which the PACF plot cuts-off.
- The Moving-Average parameter in an ARIMA model is 'q' which comes from the significant lag before the ACF plot cuts-off.
- By looking at the above plots, we will take the value of p and q to be 2 and 2 respectively.

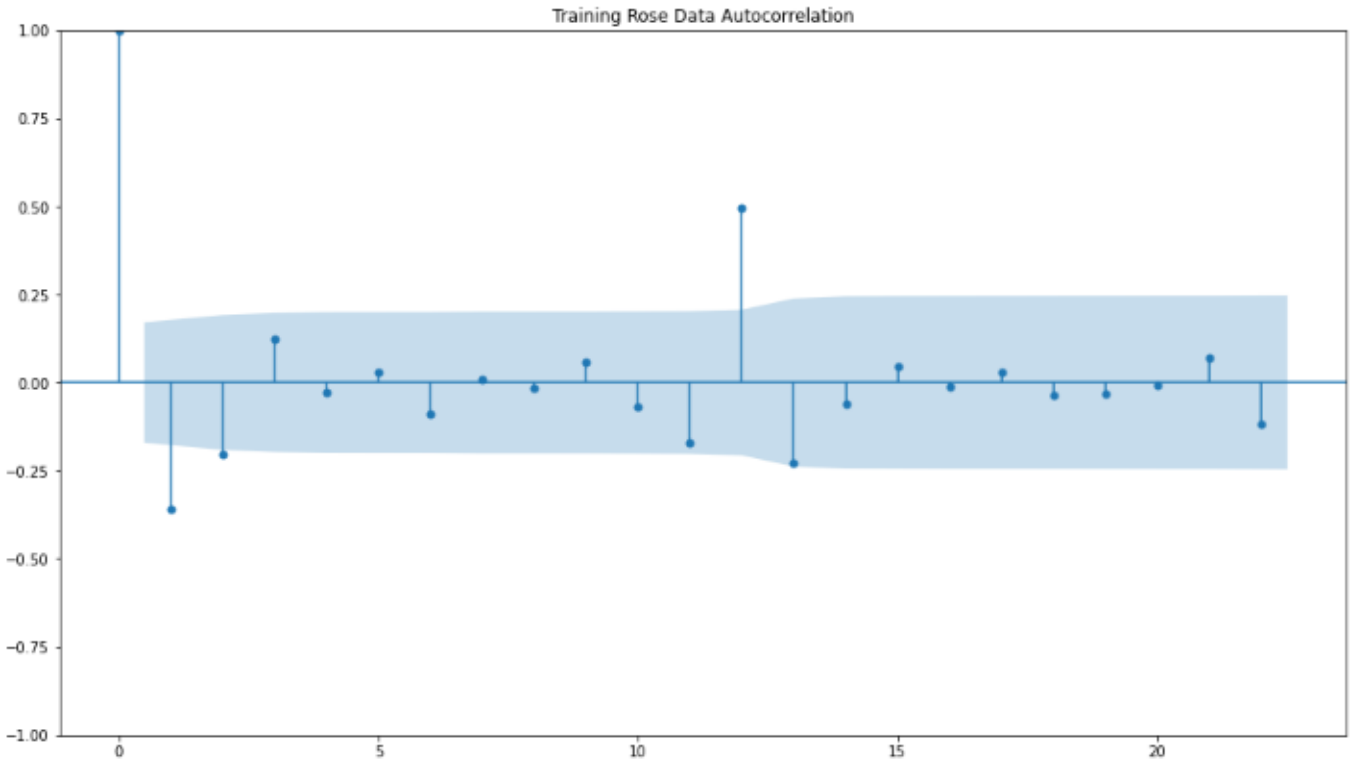

Standardized residual for "R"


Histogram plus estimated density
Hist
KDE
N(0,1)


Normal Q-Q
Sample Quantiles
Theoretical Quantiles


Correlogram


ARIMA (2, 1, 2) Forecast - Rose
Train
Test
ARIMA (2, 1, 2) Forecast - Rose

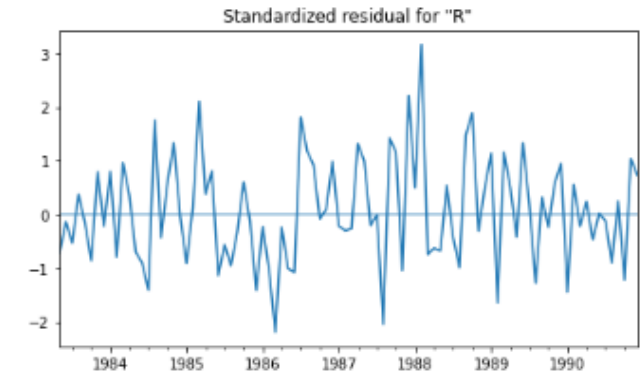


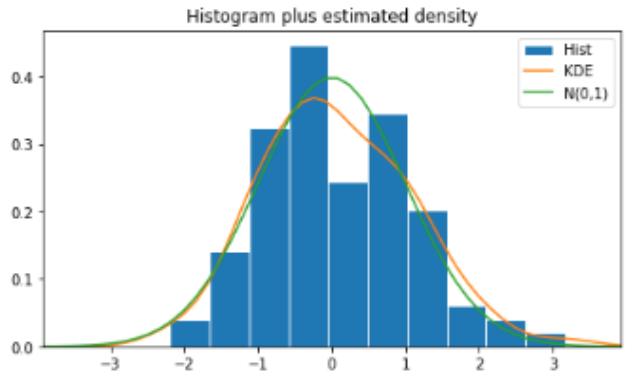
SARIMAX Results

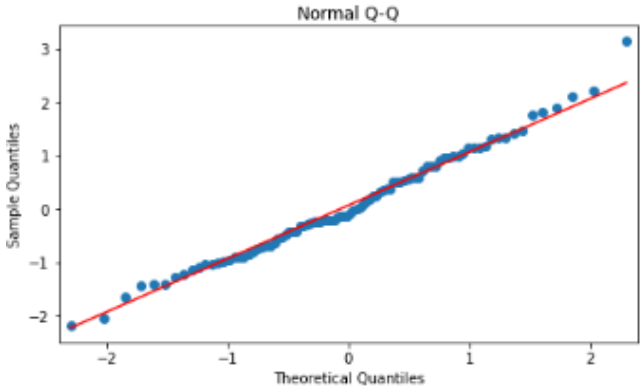
Dep. Variable:	Rose	No. Observations:	132
Model:	SARIMAX(3, 1, 1)x(3, 0, [1, 2], 12)	Log Likelihood	-377.200
Date:	Sun, 09 Apr 2023	AIC	774.400
Time:	18:52:31	BIC	799.618
Sample:	01-01-1980 - 12-01-1990	HQIC	784.578
Covariance Type:	opg		

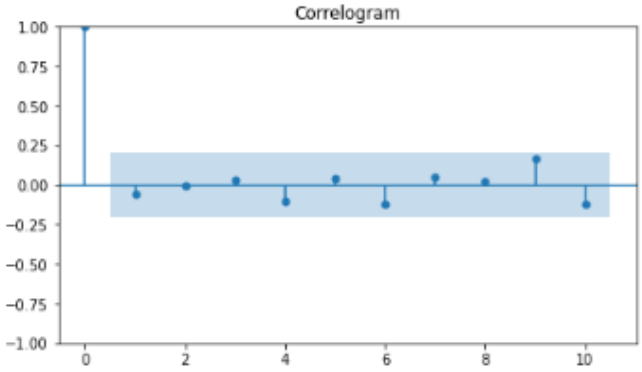
	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.0464	0.126	0.367	0.714	-0.202	0.294
ar.L2	-0.0060	0.120	-0.050	0.960	-0.241	0.229
ar.L3	-0.1808	0.098	-1.838	0.066	-0.374	0.012
ma.L1	-0.9370	0.067	-13.905	0.000	-1.069	-0.805
ar.S.L12	0.7639	0.165	4.640	0.000	0.441	1.087
ar.S.L24	0.0840	0.159	0.527	0.598	-0.229	0.397
ar.S.L36	0.0727	0.095	0.764	0.445	-0.114	0.259
ma.S.L12	-0.4969	0.250	-1.988	0.047	-0.987	-0.007
ma.S.L24	-0.2191	0.210	-1.044	0.296	-0.630	0.192
sigma2	192.1518	39.627	4.849	0.000	114.484	269.819

Ljung-Box (L1) (Q):	0.30	Jarque-Bera (JB):	1.64
Prob(Q):	0.58	Prob(JB):	0.44
Heteroskedasticity (H):	1.11	Skew:	0.33
Prob(H) (two-sided):	0.77	Kurtosis:	3.03

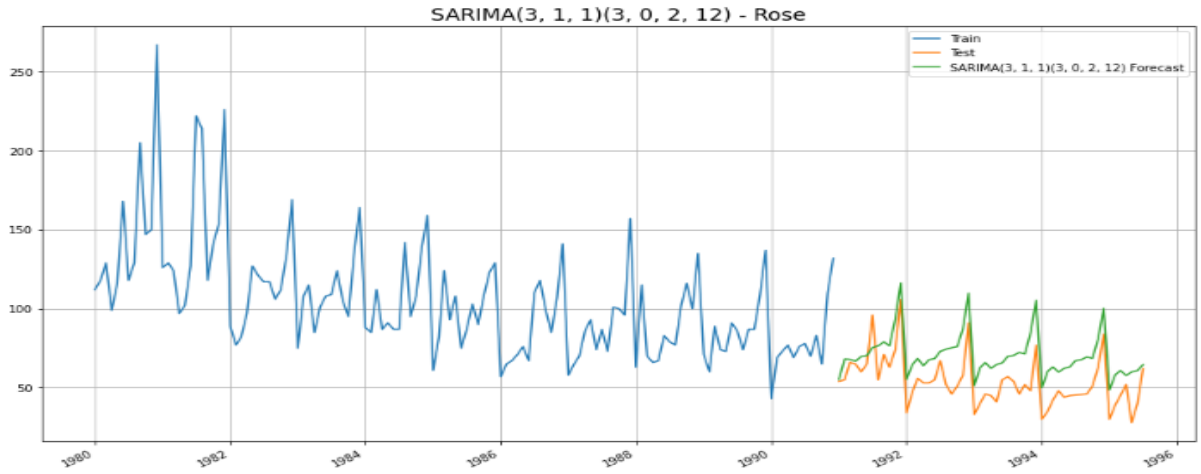








Predicted auto SARIMA:



Here, we have taken alpha=0.05.

- We are going to take the seasonal period as 12 We are taking the p value to be 2 and the q value also to be 2 as the parameters same as the ARIMA model.
- The Auto-Regressive parameter in an SARIMA model is 'P' which comes from the significant lag after which the PACF plot cuts-off to 0.
- The Moving-Average parameter in an SARIMA model is 'Q' which comes from the significant lag after which the ACF plot cuts-off to 3.

SARIMAX Results

Dep. Variable:	Rose	No. Observations:	132
Model:	SARIMAX(2, 1, 2)x(2, 1, 2, 12)	Log Likelihood	-379.498
Date:	Sun, 09 Apr 2023	AIC	776.996
Time:	18:54:38	BIC	799.692
Sample:	01-01-1980 - 12-01-1990	HQIC	786.156
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.8551	0.146	-5.837	0.000	-1.142	-0.568
ar.L2	-0.0022	0.125	-0.017	0.986	-0.247	0.242
ma.L1	0.0120	0.184	0.065	0.948	-0.348	0.372
ma.L2	-0.9434	0.150	-6.295	0.000	-1.237	-0.650
ar.S.L12	0.0349	0.185	0.188	0.851	-0.328	0.398
ar.S.L24	-0.0459	0.029	-1.599	0.110	-0.102	0.010
ma.S.L12	-0.7224	0.333	-2.172	0.030	-1.374	-0.071
ma.S.L24	-0.0771	0.212	-0.363	0.716	-0.493	0.339
sigma2	192.2014	39.474	4.869	0.000	114.834	269.569

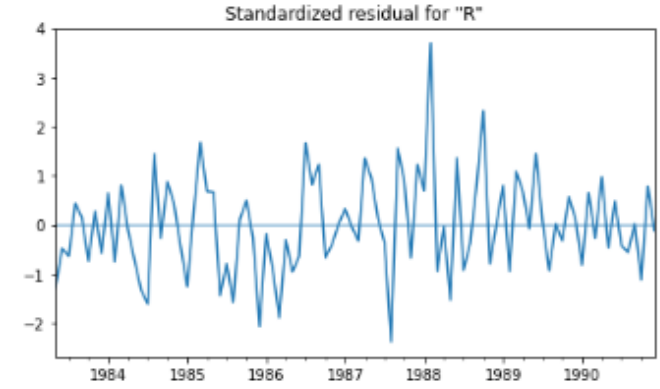
Ljung-Box (L1) (Q):	0.03	Jarque-Bera (JB):	7.06
Prob(Q):	0.86	Prob(JB):	0.03
Heteroskedasticity (H):	0.87	Skew:	0.45
Prob(H) (two-sided):	0.71	Kurtosis:	4.01

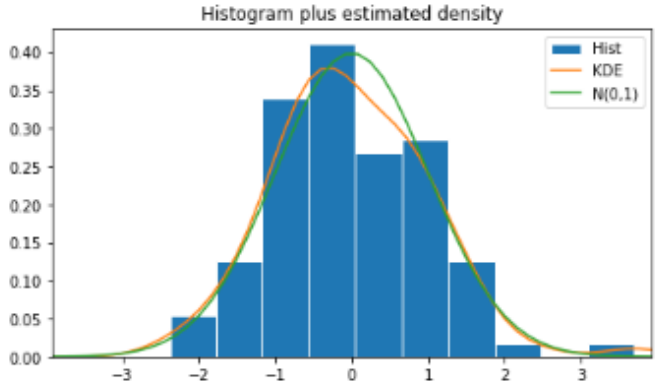
SARIMAX Results

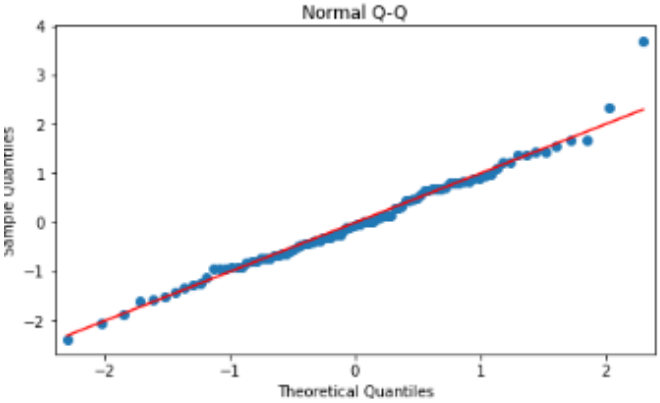
Dep. Variable:	Rose	No. Observations:	132
Model:	SARIMAX(2, 1, 2)x(3, 1, 2, 12)	Log Likelihood	-334.893
Date:	Sun, 09 Apr 2023	AIC	689.786
Time:	18:54:47	BIC	713.730
Sample:	01-01-1980 - 12-01-1990	HQIC	699.392
Covariance Type:	opg		

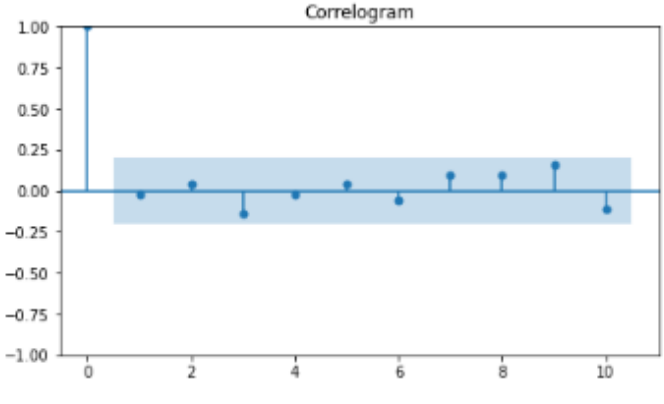
	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.7088	0.403	1.757	0.079	-0.082	1.500
ar.L2	-0.1501	0.176	-0.854	0.393	-0.495	0.194
ma.L1	-1.6097	0.422	-3.817	0.000	-2.436	-0.783
ma.L2	0.6494	0.397	1.638	0.101	-0.128	1.427
ar.S.L12	-0.0422	0.234	-0.180	0.857	-0.500	0.416
ar.S.L24	-0.0169	0.158	-0.107	0.915	-0.327	0.293
ar.S.L36	2.951e-06	0.067	4.42e-05	1.000	-0.131	0.131
ma.S.L12	-0.8403	87.318	-0.010	0.992	-171.981	170.300
ma.S.L24	-0.1604	13.772	-0.012	0.991	-27.152	26.831
sigma2	185.5467	1.62e+04	0.011	0.991	-3.16e+04	3.19e+04

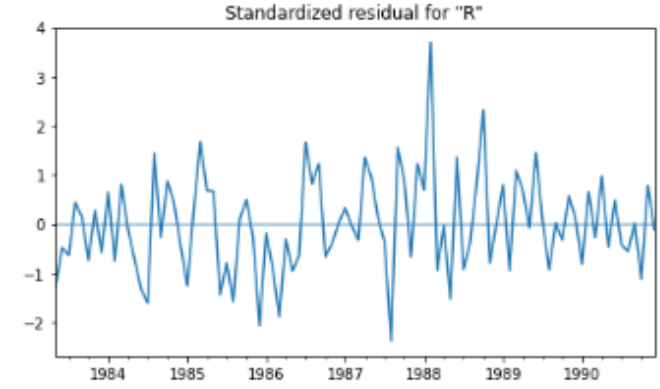
Ljung-Box (L1) (Q):	0.05	Jarque-Bera (JB):	4.60
Prob(Q):	0.82	Prob(JB):	0.10
Heteroskedasticity (H):	0.63	Skew:	0.48
Prob(H) (two-sided):	0.24	Kurtosis:	3.67

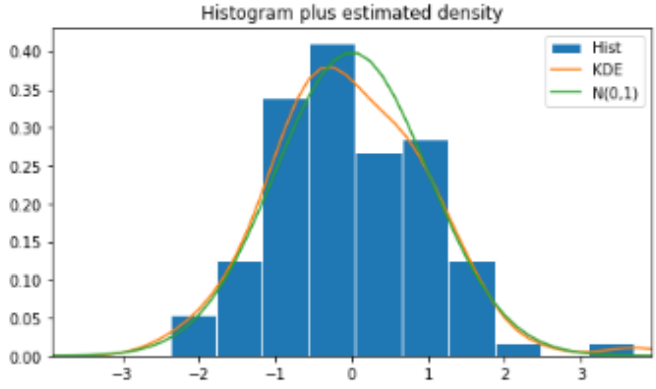

Standardized residual for "R"
1984
1985
1986
1987
1988
1989
1990

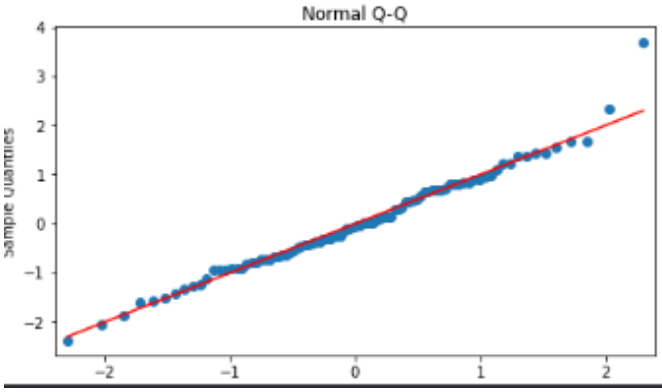

Histogram plus estimated density
Hist
KDE
N(0,1)

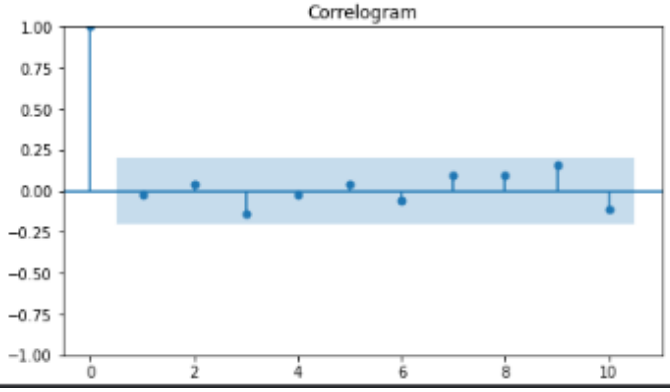

Normal Q-Q
Sample Quantiles
Theoretical Quantiles

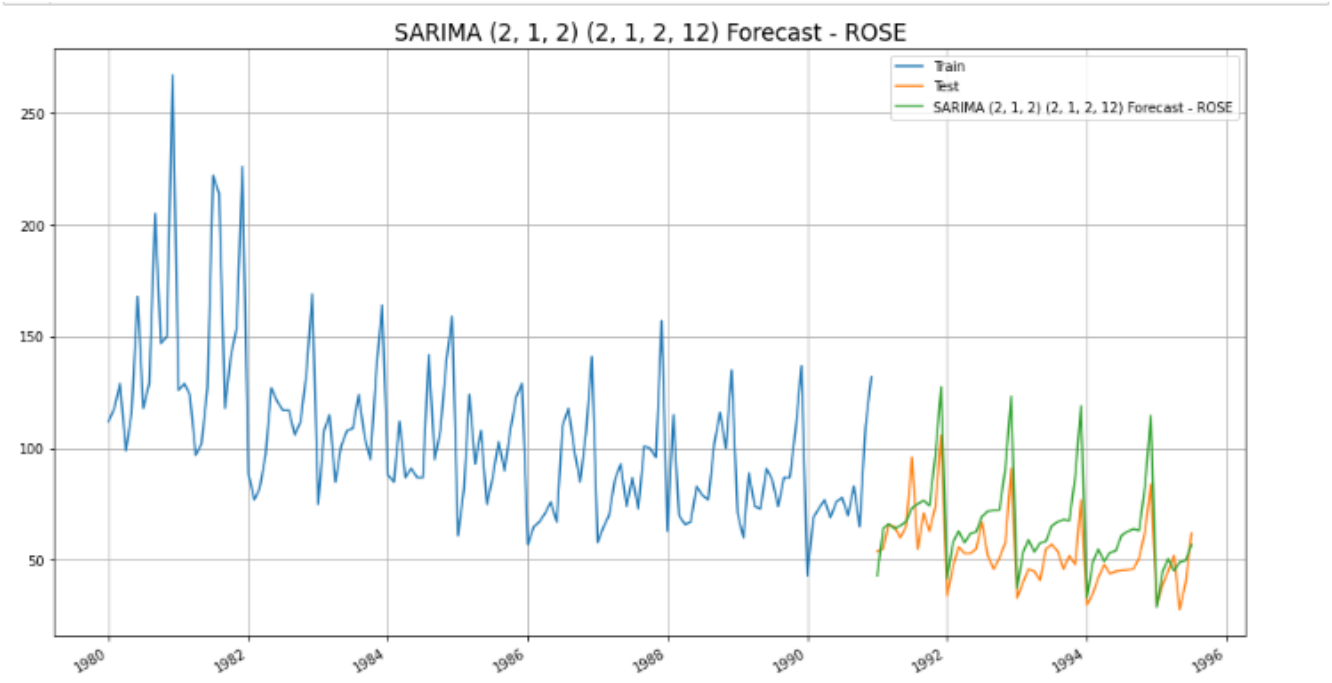

Correlogram


Standardized residual for "R"
1984
1985
1986
1987
1988
1989
1990


Histogram plus estimated density
Hist
KDE
N(0,1)


Normal Q-Q
Sample Quantiles


Correlogram



SARIMAX Results

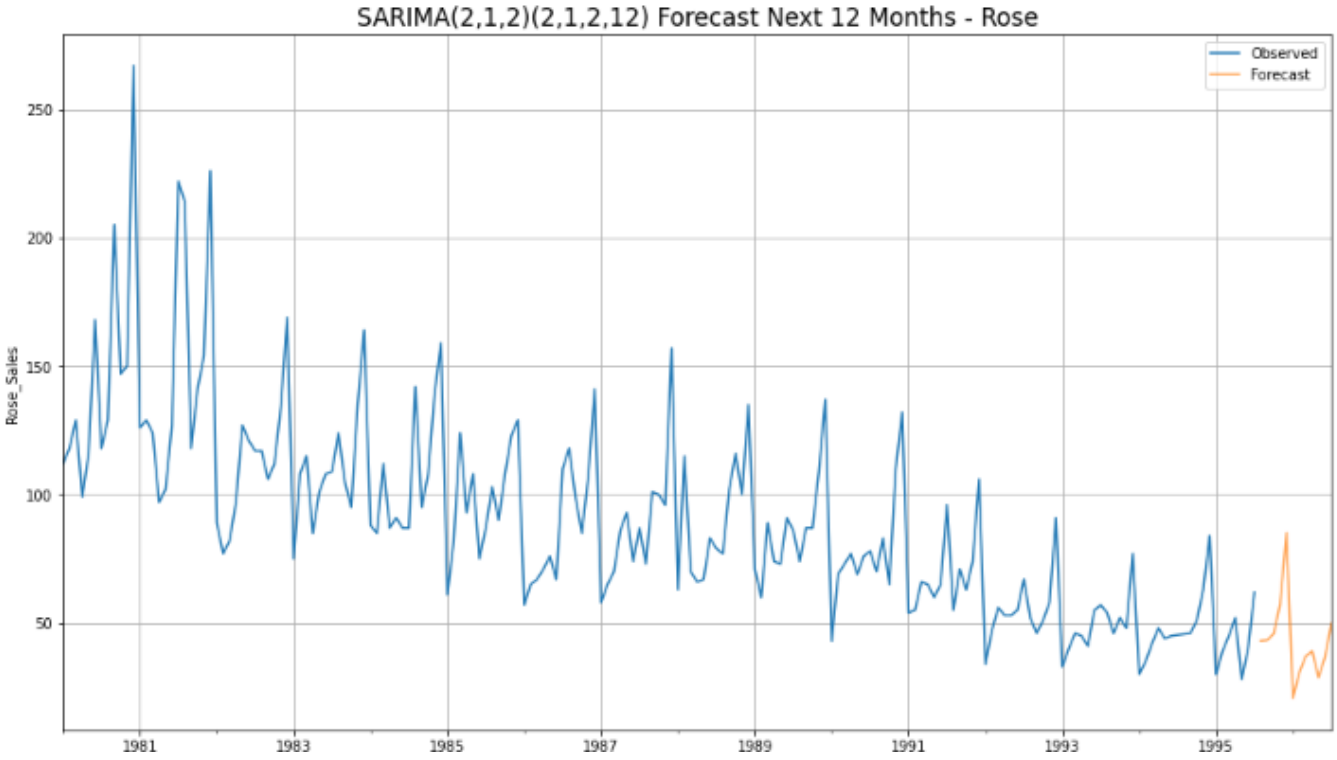
Dep. Variable:	Rose	No. Observations:	187
Model:	SARIMAX(2, 1, 2)x(2, 1, 2, 12)	Log Likelihood	-587.531
Date:	Sun, 09 Apr 2023	AIC	1193.062
Time:	18:56:35	BIC	1219.976
Sample:	01-01-1980 - 07-01-1995	HQIC	1203.997
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.8650	0.101	-8.557	0.000	-1.063	-0.667
ar.L2	0.0340	0.091	0.375	0.708	-0.144	0.211
ma.L1	0.0892	83.920	0.001	0.999	-164.392	164.570
ma.L2	-0.9108	76.438	-0.012	0.990	-150.726	148.904
ar.S.L12	0.0719	0.165	0.435	0.664	-0.252	0.396
ar.S.L24	-0.0357	0.017	-2.046	0.041	-0.070	-0.002
ma.S.L12	-0.6870	0.222	-3.095	0.002	-1.122	-0.252
ma.S.L24	-0.0549	0.150	-0.365	0.715	-0.350	0.240
sigma2	158.8958	1.33e+04	0.012	0.990	-2.6e+04	2.63e+04

Ljung-Box (L1) (Q):	0.05	Jarque-Bera (JB):	10.12
Prob(Q):	0.83	Prob(JB):	0.01
Heteroskedasticity (H):	0.53	Skew:	0.35
Prob(H) (two-sided):	0.03	Kurtosis:	4.08

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Rose	mean	mean_se	mean_ci_lower	mean_ci_upper
1995-08-01	43.083240	12.673822	18.243008	67.923474
1995-09-01	43.342804	12.959446	17.942757	68.742851
1995-10-01	45.827645	12.963459	20.419732	71.235557
1995-11-01	57.397751	13.124193	31.674804	83.120697
1995-12-01	85.075385	13.133043	59.335093	110.815677



Observations:

- Inference from Model diagnostics confirms that the model residuals are normally distributed.
- Standardized residual– Do not display any obvious seasonality
- Histogram plus estimated density - The KDE plot of the residuals is similar with the normal distribution; hence the model residuals are normally distributed
- Normal Q-Q plot – There is an ordered distribution of residuals (blue dots) following the linear trend of the samples taken from a standard normal distribution with N(0,
- Correlogram – The time series residuals have low correlation with lagged versions of itself

We see that the best model is the Triple Exponential Smoothing with multiplicative seasonality with the parameters α = 0.03, β = 0.03 and γ = 0.33.

9. 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 these insights to increase our sales further.
- We can introduce certain offers in November, December months to attract more customers.
- On Saturdays mean sales of the wine is highest. We can give certain offers to attract more customers.
- Year 1988 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 will be more or less same as that of previous year. Hence some important measures have to be taken to increase the trend. As the trend has been more or less constant throughout the years.
- The company should use the prediction results and capitalize on the high demand seasons and ensure to source and supply the high demand.
- The company should use the prediction results to plan the low demand seasons to stock as per the demand.
- The price of rose wine may be expensive than sparkling so seasonal discounts can help improve the sales of rose wine. Products that are discounted should be highlighted so consumers can see the savings prominently. Discounts can compel consumers to buy.
- As we know how the seasonality is in the prediction company cannot have the same stock through the year. You should create a dynamic consumer experience with fresh point-of-sale materials and well-stocked displays. Displays need to look fresh and interesting and tell a compelling story about

why the consumer should purchase the product.

- Seasonal memberships and discounts can be introduced. Consumers get very excited about savings and appreciate discounts being passed on. Many prominent retailers also have loyalty programs or club member cards that create excitement. A club-member price brings consumers back and improve sales.

- Events and tastings help draw consumers to your store and generate sales. Retailers with economies of scale successfully sample consumers on more profitable wines. Some even comparison-taste customer on national brands that are more expensive to demonstrate they are offering a less expensive but superior product. And bringing in celebrities, sommeliers or trade reps for tastings can help create excitement and drive traffic.