Time Series Forecasting

Data Science and Business Analytics (PGP-DSBA)
Online September_A 2021-22
Made By:
Anmol Tripathi

Table of Content:
Problem 1:

List of Tables:

List of Graph:

Problem 1:

Executive Summary:

Here in this problem, we are provided with the data for two firms Rose and Sparkling. Here, we will be using Time Series Forecasting in order to analyze the data and get the desired results.

Background of the problem:

For this particular assignment, the data of different types of wine sales in the 20th century is to be analyzed. 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 analyze and forecast Wine Sales in the 20th century.

Data Description:

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 187 entries, 1980-01-01 to 1995-07-01
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   Rose    185 non-null    float64
dtypes: float64(1)
memory usage: 2.9 KB
```

Table 1 Data Type for Rose

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 187 entries, 1980-01-01 to 1995-07-01
Data columns (total 1 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Sparkling  187 non-null    int64
dtypes: int64(1)
memory usage: 2.9 KB
```

Table 2 Data Type for Sparkling

Checking for Missing Values:

```
Sparkling    0
dtype: int64
```

```
Rose    2
dtype: int64
```

Table 3 Checking for Missing Values

Here we make use of interpolation method in order to treat the missing values for the case of Rose Data, after which the missing values are removed from the data set.

Sample Data Set:

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

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

Table 4 Sample Data set for Rose and Sparkling

The above table represents the actual representation of the data set.

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

Data Description:

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

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

Table 5: Sample Data set for Rose and Sparkling

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 187 entries, 1980-01-01 to 1995-07-01
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Rose    185 non-null    float64
dtypes: float64(1)
memory usage: 2.9 KB
```

Table 6: Data Type for Rose

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 187 entries, 1980-01-01 to 1995-07-01
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Rose    185 non-null    float64
dtypes: float64(1)
memory usage: 2.9 KB
```

Table 7: Data Type for Rose

Checking for Missing Values:

```
Sparkling    0
dtype: int64
```

```
Rose    2
dtype: int64
```

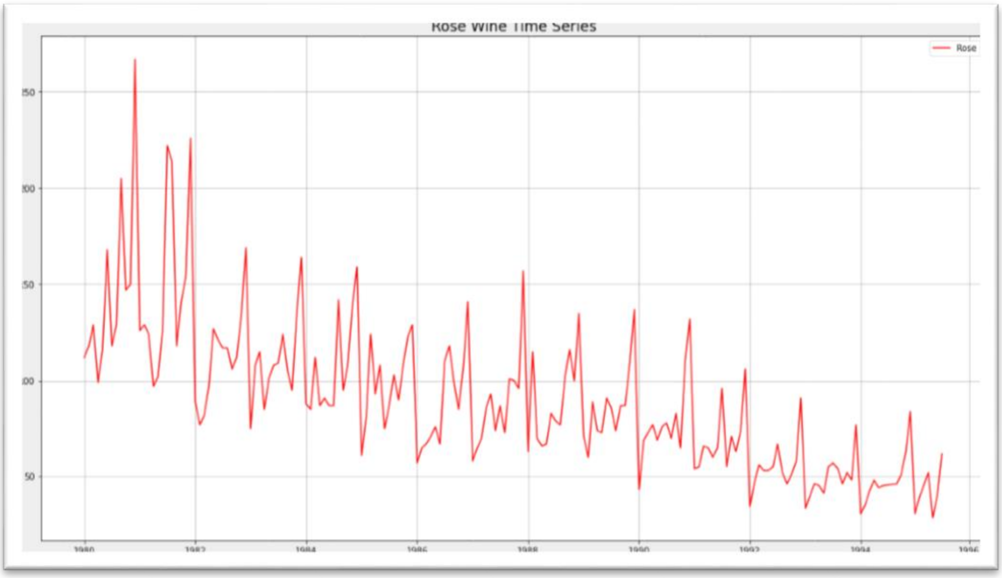
Table 8: Checking for missing Values

Fill Missing Values by Interpolation:

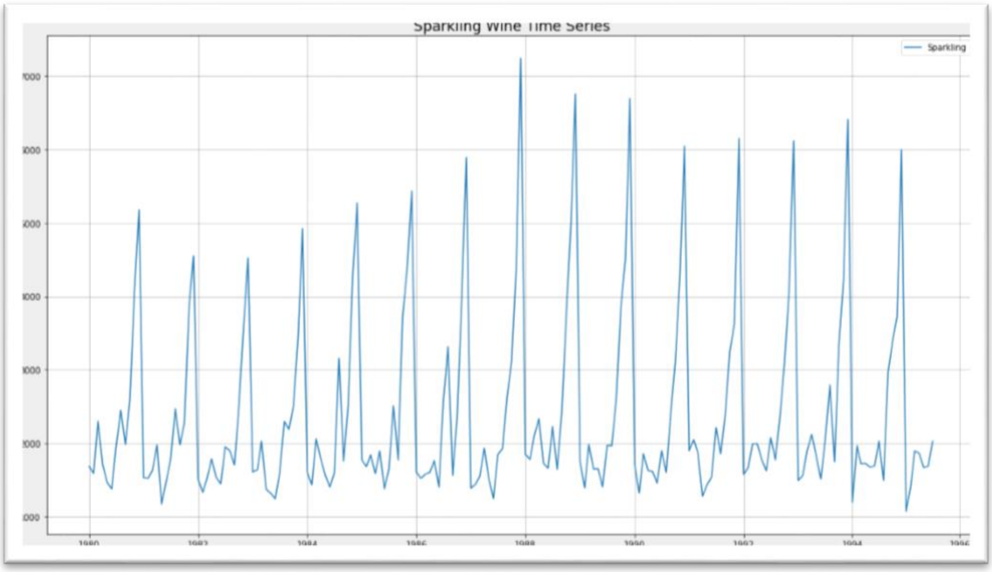
```
Rose     0
dtype: int64
```

Table 9: After filling missing values via interpolation

Plot the Time Series:



Graph 1: Rose Wine Time series Forecast



Graph 2: Sparkling Wine Time series Forecast

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

Check the basic measures of descriptive statistics of the Time Series:

	count	mean	std	min	25%	50%	75%	max
Rose	187.0	89.914	39.238	28.0	62.5	85.0	111.0	267.0

Table 10: Descriptive Analytics

	count	mean	std	min	25%	50%	75%	max
Sparkling	187.0	2402.417	1295.112	1070.0	1605.0	1874.0	2549.0	7242.0

Table 11: Descriptive Analytics

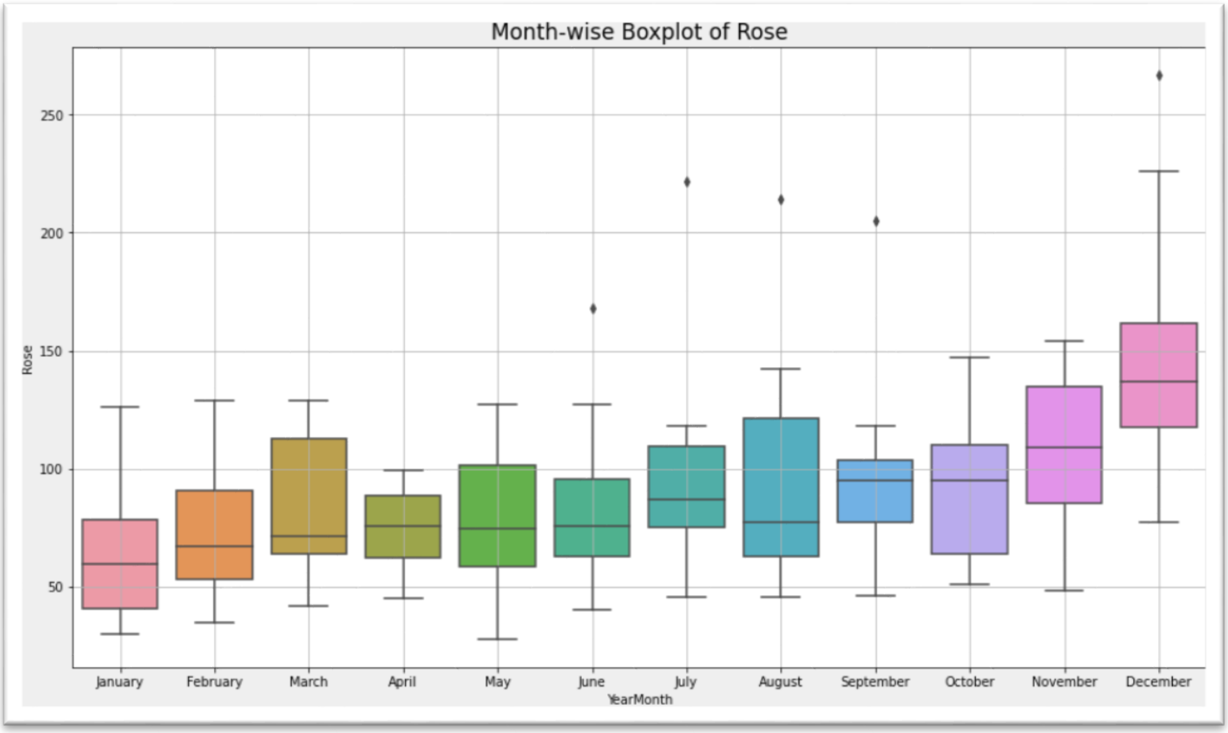
```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 187 entries, 1980-01-01 to 1995-07-01
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Rose    185 non-null    float64
dtypes: float64(1)
memory usage: 2.9 KB
```

Table 12: Data Type for Rose

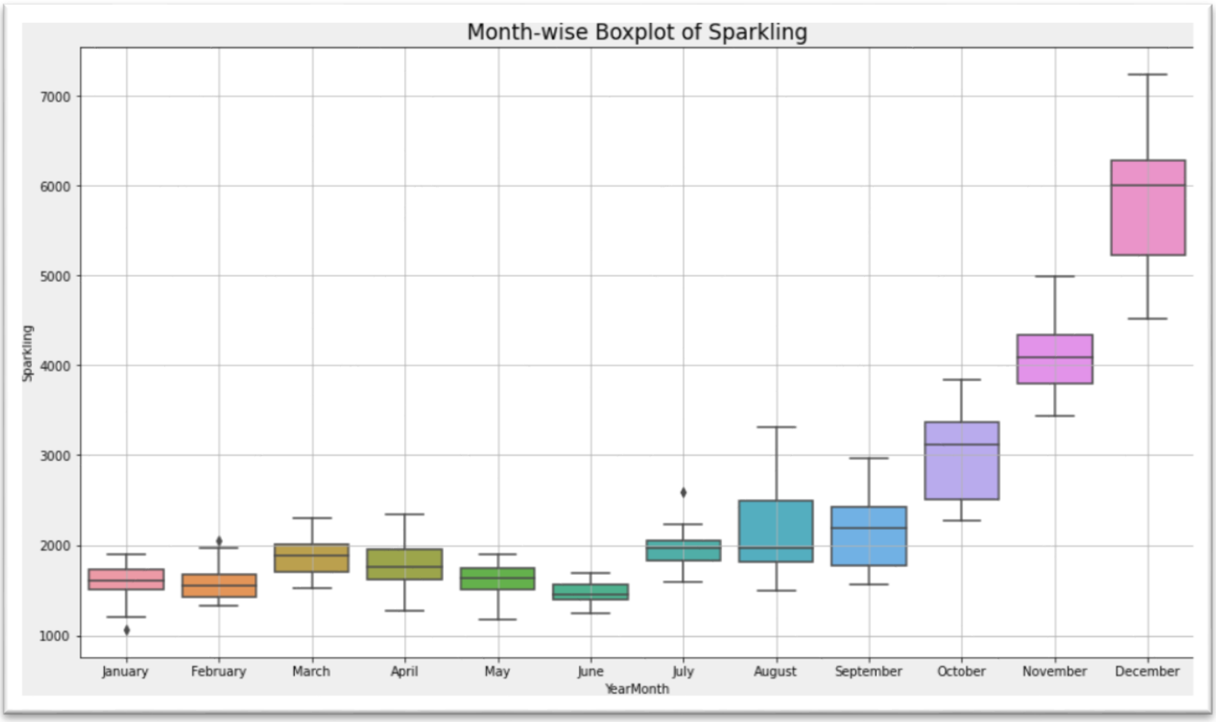
```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 187 entries, 1980-01-01 to 1995-07-01
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Rose    185 non-null    float64
dtypes: float64(1)
memory usage: 2.9 KB
```

Table 13: Data Type for Sparkling

Plot Box-plot Month wise:



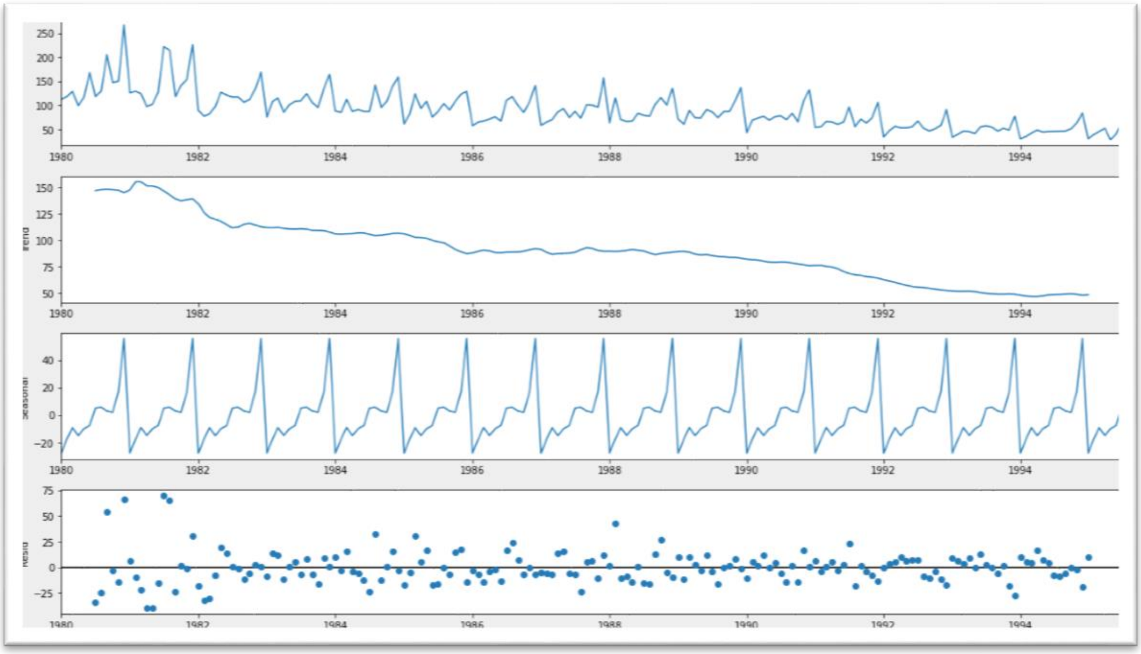
Graph 3: Box Plot



Graph 4: Box Plot

- Sales of both – Rose and Sparkling, show a spike in the last quarter of October to December.
- Spike is much more accentuated in sparkling sales.
- This spike may be due to the Holiday Season starting in October.

Additive Decomposition of Rose:



Graph 5: Decomposition Graph

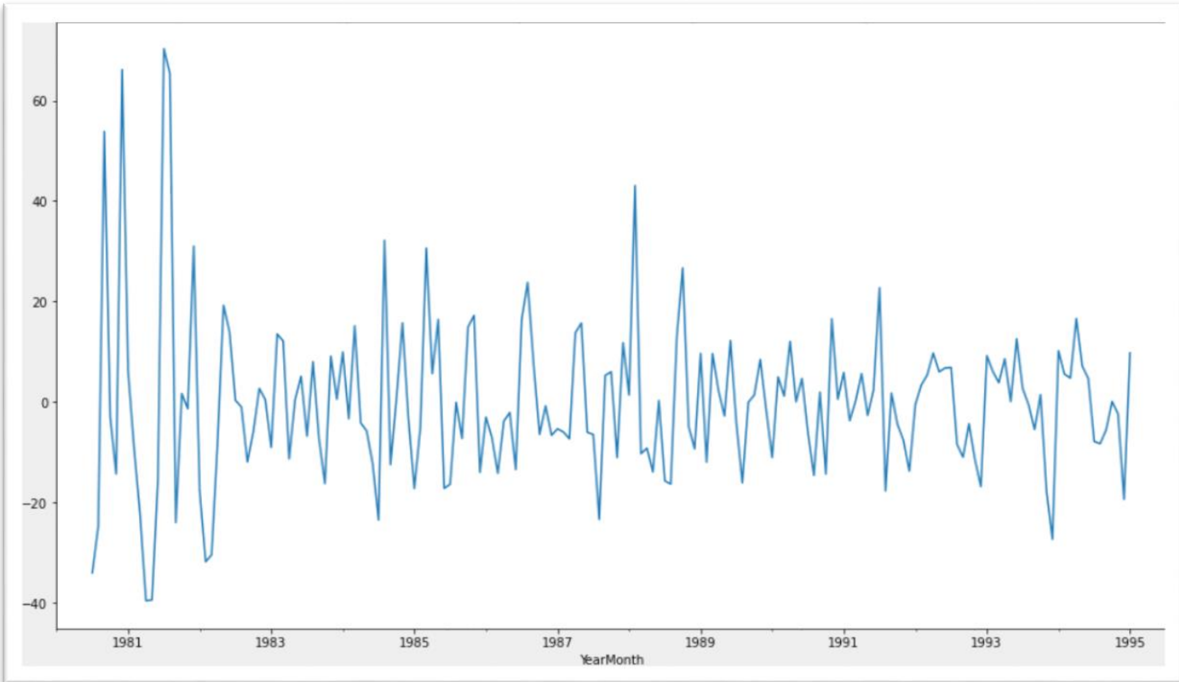
```
Trend
 YearMonth
1980-01-01           NaN
1980-02-01           NaN
1980-03-01           NaN
1980-04-01           NaN
1980-05-01           NaN
1980-06-01           NaN
1980-07-01    147.083333
1980-08-01    148.125000
1980-09-01    148.375000
1980-10-01    148.083333
1980-11-01    147.416667
1980-12-01    145.125000
Name: trend, dtype: float64
```

```
Seasonality
 YearMonth
1980-01-01   -27.908647
1980-02-01   -17.435632
1980-03-01    -9.285830
1980-04-01   -15.098330
1980-05-01   -10.196544
1980-06-01    -7.678687
1980-07-01     4.896908
1980-08-01     5.499686
1980-09-01     2.774686
1980-10-01     1.871908
1980-11-01    16.846908
1980-12-01    55.713575
Name: seasonal, dtype: float64
```

```
Residual
 YearMonth
1980-01-01          NaN
1980-02-01          NaN
1980-03-01          NaN
1980-04-01          NaN
1980-05-01          NaN
1980-06-01          NaN
1980-07-01   -33.980241
1980-08-01   -24.624686
1980-09-01    53.850314
1980-10-01    -2.955241
1980-11-01   -14.263575
1980-12-01    66.161425
Name: resid, dtype: float64
```

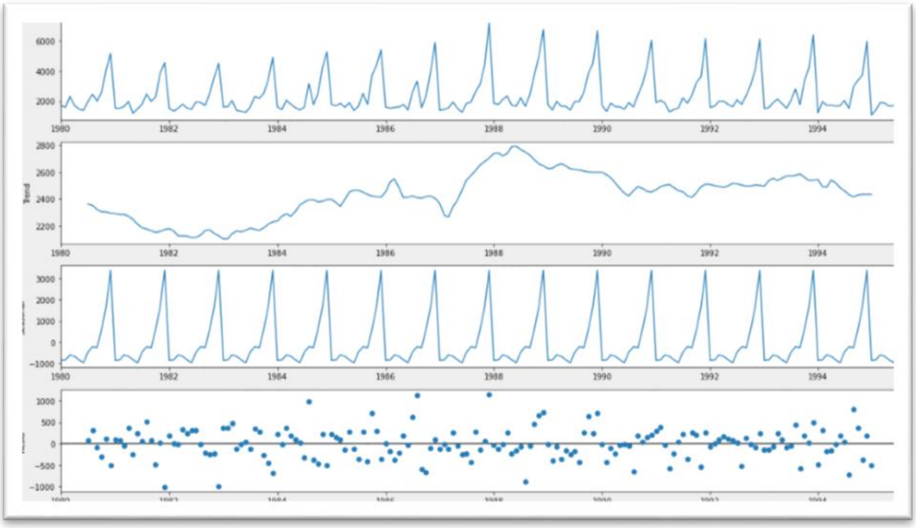
Table 14: Trend, Seasonal and Residual Values for Rose Data Set

Residual Plot:



Graph 6: Residual Plot

Additive Decomposition of Sparkling:



Graph 7: Decomposition Graph

Trend

YearMonth	
1980-01-01	NaN
1980-02-01	NaN
1980-03-01	NaN
1980-04-01	NaN
1980-05-01	NaN
1980-06-01	NaN
1980-07-01	2360.666667
1980-08-01	2351.333333
1980-09-01	2320.541667
1980-10-01	2303.583333
1980-11-01	2302.041667
1980-12-01	2293.791667

Name: trend, dtype: float64

Seasonality

YearMonth	
1980-01-01	-854.260599
1980-02-01	-830.350678
1980-03-01	-592.356630
1980-04-01	-658.490559
1980-05-01	-824.416154
1980-06-01	-967.434011
1980-07-01	-465.502265
1980-08-01	-214.332821
1980-09-01	-254.677265
1980-10-01	599.769957
1980-11-01	1675.067179
1980-12-01	3386.983846

Name: seasonal, dtype: float64

Residual

YearMonth	
1980-01-01	NaN
1980-02-01	NaN
1980-03-01	NaN
1980-04-01	NaN
1980-05-01	NaN
1980-06-01	NaN
1980-07-01	70.835599
1980-08-01	315.999487
1980-09-01	-81.864401
1980-10-01	-307.353290
1980-11-01	109.891154
1980-12-01	-501.775513

Name: resid, dtype: float64

Table 15: Trend, Seasonal and Residual Values for Sparkling Data Set

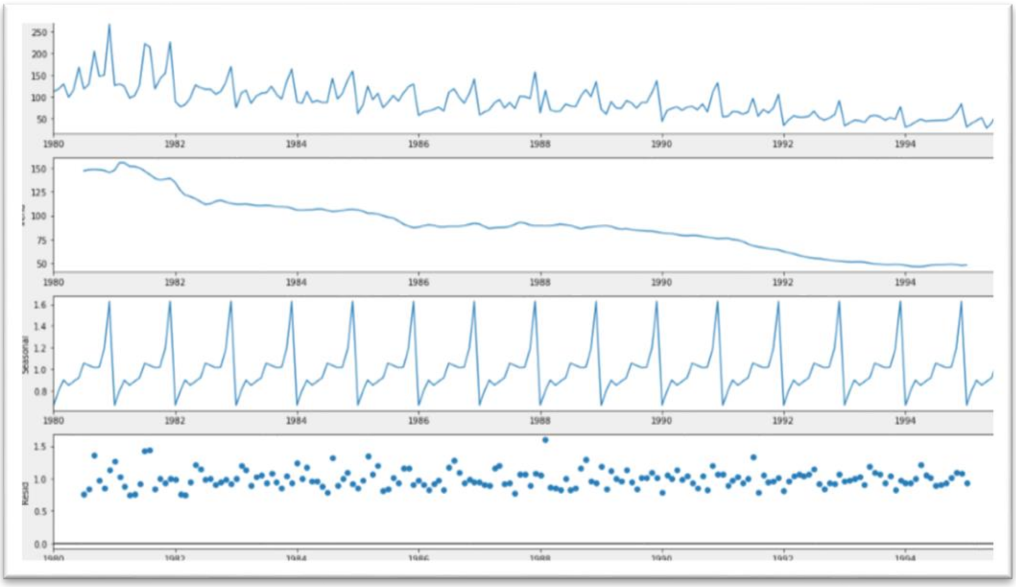
CVSA:

```
cvsa = residual.std()/residual.mean()
cvsa
```

```
-296.08335294142995
```

Table 16: CVSA Value

Multiplicative Decomposition of Rose:



Graph 8: Multiplicative Decomposition

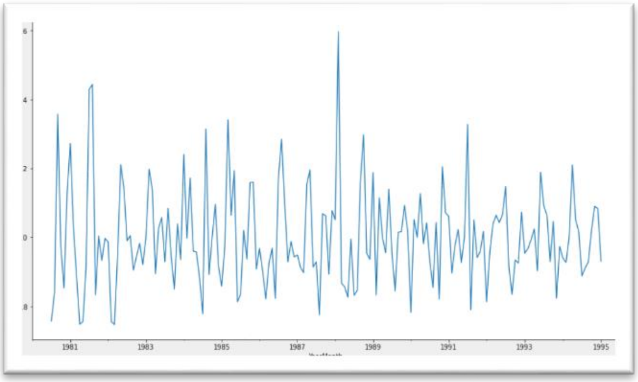
```
Trend
 YearMonth
1980-01-01           NaN
1980-02-01           NaN
1980-03-01           NaN
1980-04-01           NaN
1980-05-01           NaN
1980-06-01           NaN
1980-07-01    147.083333
1980-08-01    148.125000
1980-09-01    148.375000
1980-10-01    148.083333
1980-11-01    147.416667
1980-12-01    145.125000
Name: trend, dtype: float64
```

```
Seasonality
 YearMonth
1980-01-01    0.670111
1980-02-01    0.806163
1980-03-01    0.901164
1980-04-01    0.854024
1980-05-01    0.889415
1980-06-01    0.923985
1980-07-01    1.058038
1980-08-01    1.035881
1980-09-01    1.017648
1980-10-01    1.022573
1980-11-01    1.192349
1980-12-01    1.628646
Name: seasonal, dtype: float64
```

```
Residual
 YearMonth
1980-01-01         NaN
1980-02-01         NaN
1980-03-01         NaN
1980-04-01         NaN
1980-05-01         NaN
1980-06-01         NaN
1980-07-01    0.758258
1980-08-01    0.840720
1980-09-01    1.357674
1980-10-01    0.970771
1980-11-01    0.853378
1980-12-01    1.129646
Name: resid, dtype: float64
```

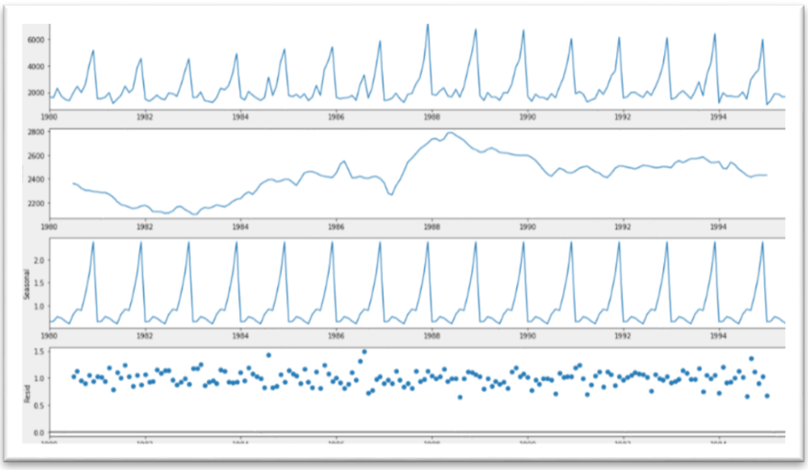
Table 17: Trend, Seasonal and Residual Values for Rose Data Set

Residual Plot:



Graph 9: Residual Plot:

Multiplicative Decomposition of Sparkling:



Graph 10: Multiplicative Decomposition

```
Trend
 YearMonth
1980-01-01            NaN
1980-02-01            NaN
1980-03-01            NaN
1980-04-01            NaN
1980-05-01            NaN
1980-06-01            NaN
1980-07-01    2360.666667
1980-08-01    2351.333333
1980-09-01    2320.541667
1980-10-01    2303.583333
1980-11-01    2302.041667
1980-12-01    2293.791667
Name: trend, dtype: float64
```

```
Seasonality
 YearMonth
1980-01-01    0.649843
1980-02-01    0.659214
1980-03-01    0.757440
1980-04-01    0.730351
1980-05-01    0.660609
1980-06-01    0.603468
1980-07-01    0.809164
1980-08-01    0.918822
1980-09-01    0.894367
1980-10-01    1.241789
1980-11-01    1.690158
1980-12-01    2.384776
Name: seasonal, dtype: float64
```

```
Residual
 YearMonth
1980-01-01         NaN
1980-02-01         NaN
1980-03-01         NaN
1980-04-01         NaN
1980-05-01         NaN
1980-06-01         NaN
1980-07-01    1.029230
1980-08-01    1.135407
1980-09-01    0.955954
1980-10-01    0.907513
1980-11-01    1.050423
1980-12-01    0.946770
Name: resid, dtype: float64
```

Table 18: Trend, Seasonal and Residual Values for Sparkling Data Set

CVSM:

```
cvsm = residual.std()/residual.mean()
cvsm

0.13886244212828963
```

Table 19: CVSM Value

Additive Models:

- The seasonality is relatively constant over time
- yt = Trend + Seasonalit y + Residual

Multiplicative Models:

- The seasonality increases or decreases over time. It is proportionate to the trend
- yt = Trend * Seasonalit y * Residual

Here by just observing the Residual patterns of Additive and Multiplicative models of Rose and Sparkling datasets. It seems that –

- Rose is Multiplicative
- Sparkling is Additive

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

- Both datasets of Rose and Sparkling are split at the year 1991
- Test datasets start at 1991

Shape for Rose Data:

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

Table 20: Shape for Rose Data

Shape for Sparkling Data:

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

Table 21: Shape for Sparkling Data

Rose Data for Train and Test:

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

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

First few rows of Rose Test Data

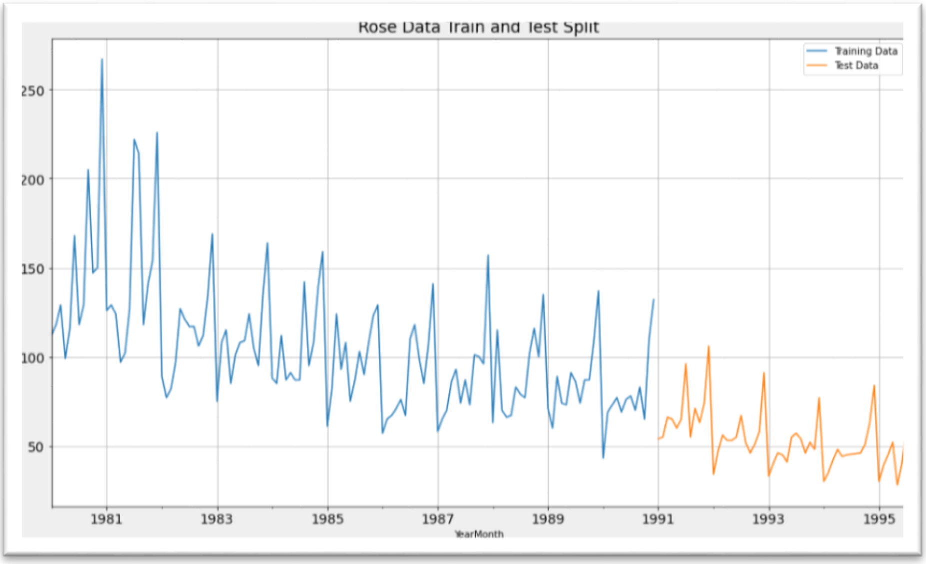
YearMonth	Rose
1991-01-01	54.0
1991-02-01	55.0
1991-03-01	66.0
1991-04-01	65.0
1991-05-01	60.0

Last few rows of Rose 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

Table 22: First few Data for Test and Training for Rose Data set

Graph for Rose Data Train Test Split Series:



Graph 11: Graph for Rose Data Train Test Split Series

Sparkling Data for Train and Test:

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

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

First few rows of Sparkling Test Data

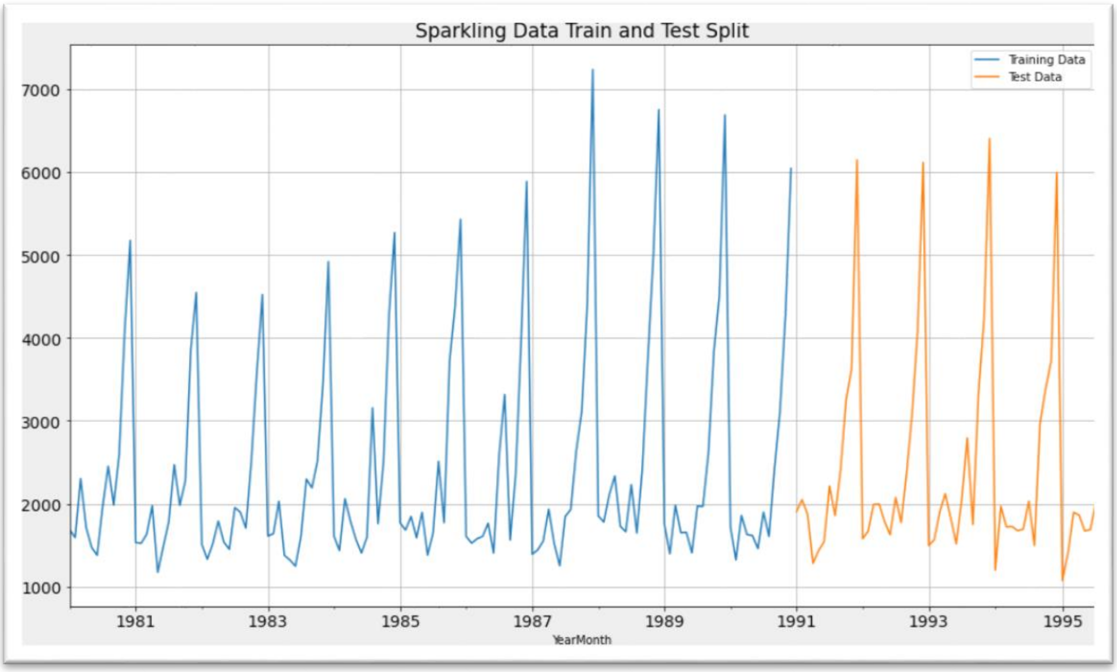
YearMonth	Sparkling
1991-01-01	1902
1991-02-01	2049
1991-03-01	1874
1991-04-01	1279
1991-05-01	1432

Last few rows of Sparkling Test Data

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

Table 23: First few Data for Test and Training for Sparkling Data set

Graph for Sparkling Data Train Test Split Series:



Graph 12: Graph for Sparkling Data Train Test Split Series

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:

a) For Rose Data:

Rose Split:

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

Table 24: Split data Set for Rose

Training and Testing Time Instance:

```
Training Time Instance :
 [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 3
4, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65,
66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97,
98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123,
124, 125, 126, 127, 128, 129, 130, 131, 132]
Test Time Instance :
 [133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157,
158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 18
3, 184, 185, 186, 187]
```

Table 25: Training and Testing Time Instance

Rose Data for Train and Test:

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

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

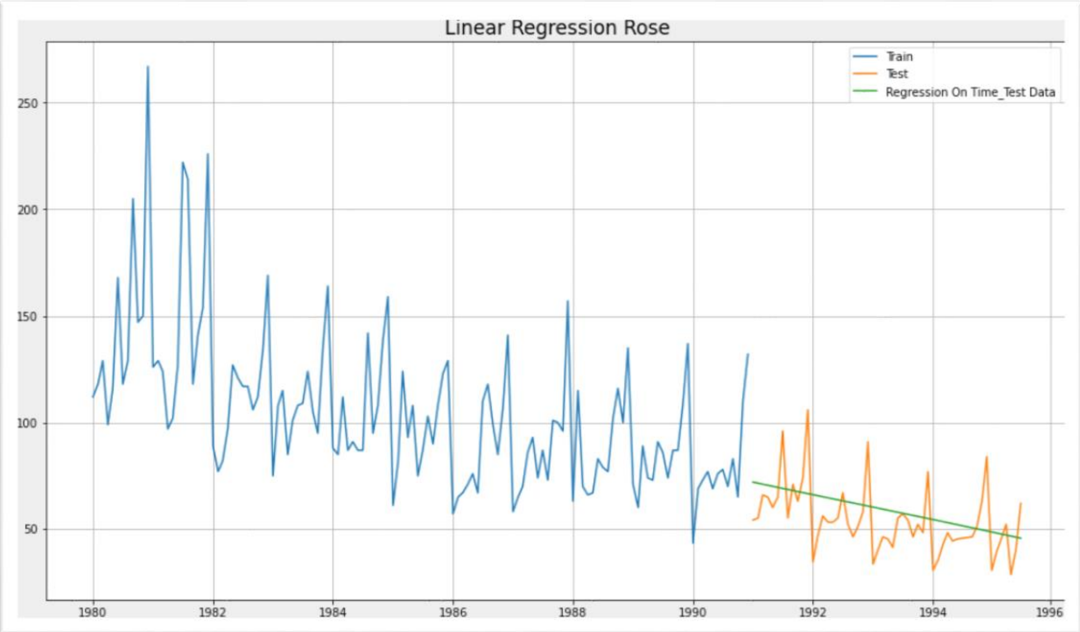
First few rows of Test Data

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 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

Table 26: First few Data for Test and Training for Rose Data set



Graph 13: Linear Regression

Model Evaluation:

```
For RegressionOnTime forecast on the Test Data,  RMSE is 15.269
```

Table 27: Model Evaluation

	Test RMSE Rose
RegressionOnTime	15.268955

Table 28: Model Evaluation

b) For Sparkling Data:

Sparkling Split:

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

Table 29: Sparkling Split

Training and Testing Time Instance:

```
Training Time instance
 [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 3
4, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65,
66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97,
98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123,
124, 125, 126, 127, 128, 129, 130, 131, 132]
Test Time instance
 [133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157,
158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 18
3, 184, 185, 186, 187]
```

Table 30: Training and Testing Time Instance

Sparkling Data for Train and Test:

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

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

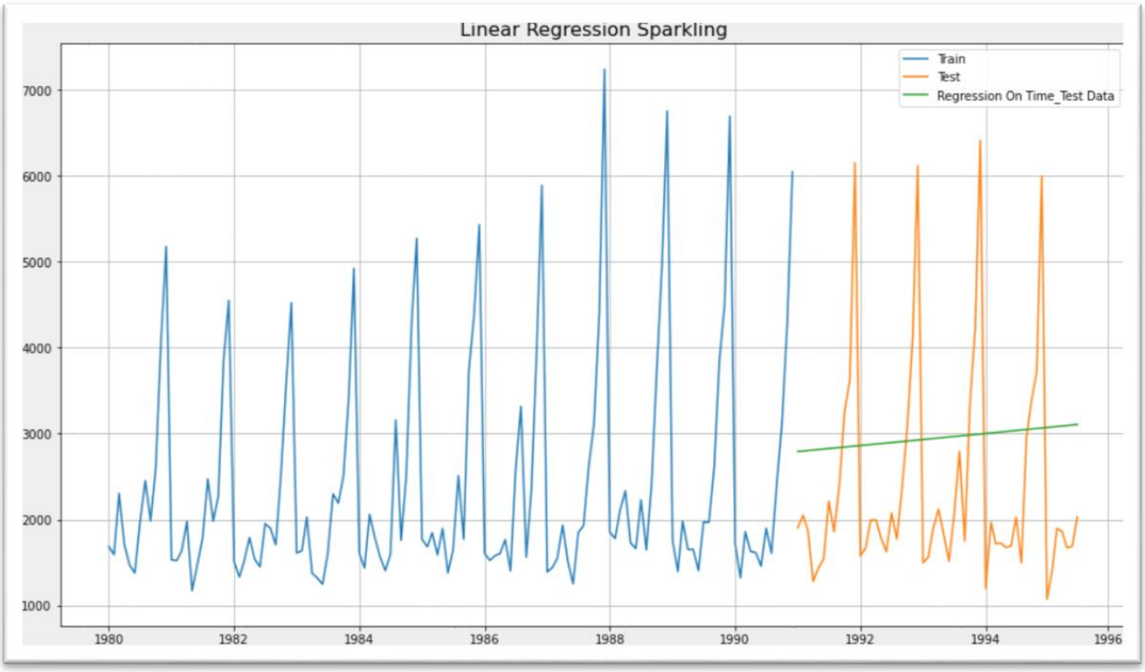
First few rows of Test Data

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 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

Table 31: First few Data for Test and Training for Rose Data set



Graph 14: Linear Regression

Model Evaluation:

```
For RegressionOnTime forecast on the Test Data,  RMSE is 1389.135
```

Table 32: Model Evaluation

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175

Table 33: Model Evaluation

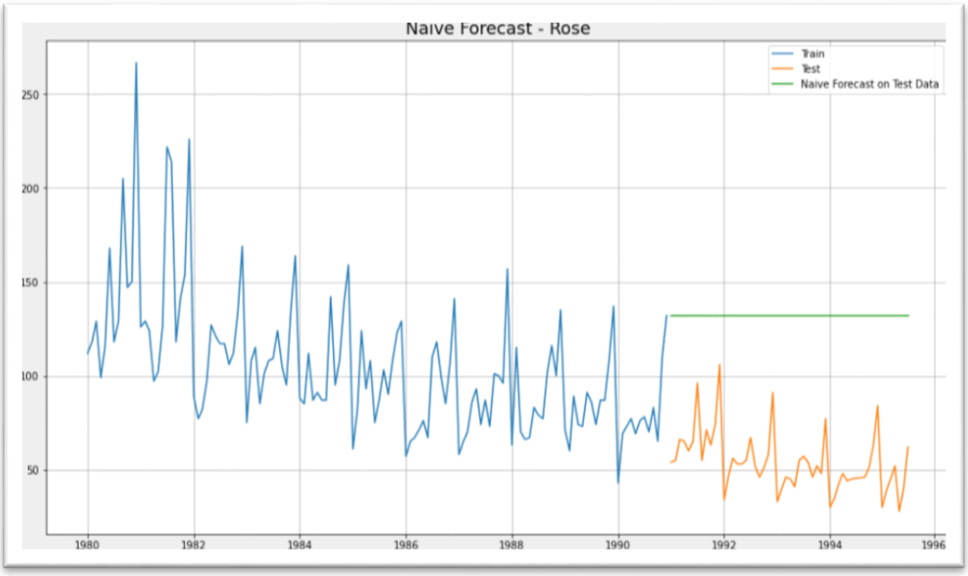
Model 2: Naive Approach:

a) For Rose 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

```
YearMonth
1991-01-01    132.0
1991-02-01    132.0
1991-03-01    132.0
1991-04-01    132.0
1991-05-01    132.0
Name: naive, dtype: float64
```

Table 34: Tail of Rose Data



Graph 15: Naïve Forecast

Model Evaluation:

```
For NaiveModel forecast on the Test Data,  RMSE is 79.719
```

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175

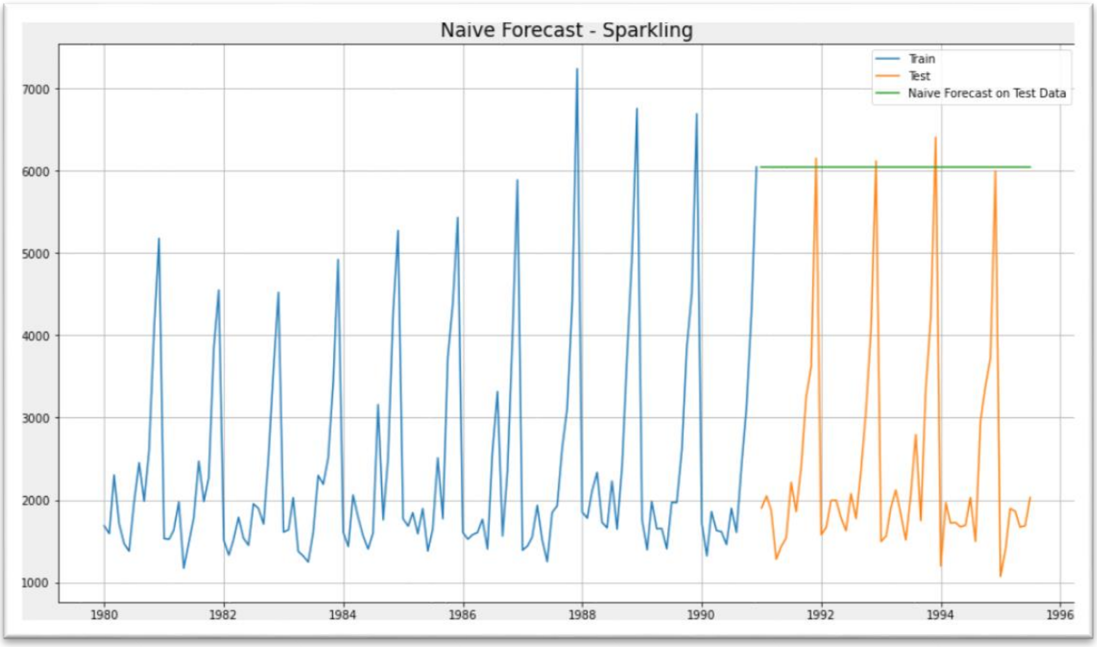
Table 35: Model Evaluation

b) For Sparkling Data:

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

```
YearMonth
1991-01-01    6047
1991-02-01    6047
1991-03-01    6047
1991-04-01    6047
1991-05-01    6047
Name: naive, dtype: int64
```

Table 36: Tail of Sparkling Data



Graph 16: Naïve Forecast

Model Evaluation:

```
For NaiveModel forecast on the Test Data,  RMSE is 3864.279
```

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175

	Test RMSE Rose
NaiveModel	79.718773

Table 37: Model Evaluation

	Test RMSE Rose	Test RMSE Sparkling
NaiveModel	79.718773	3864.279352

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352

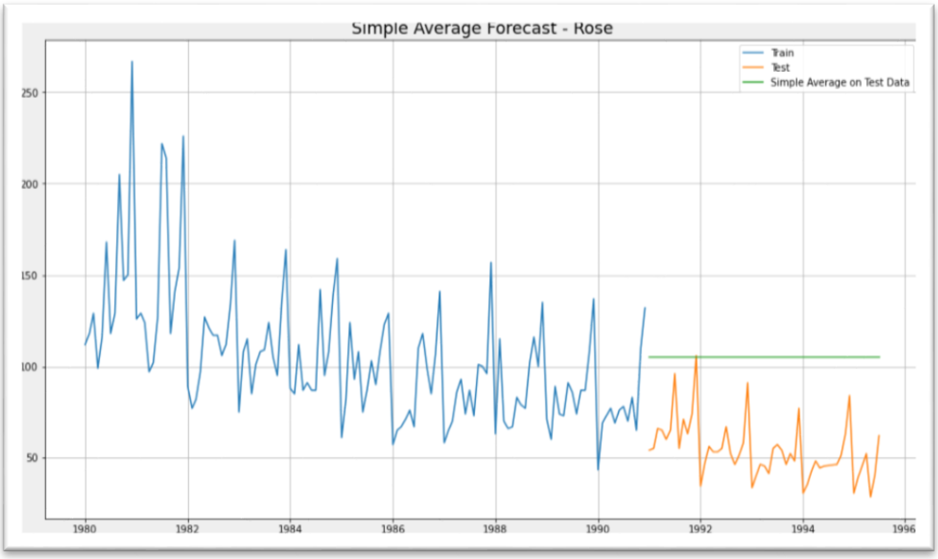
Table 38: Model Evaluation

Model 3: Simple Average:

a) For Rose Data:

YearMonth	Rose	mean_forecast
1991-01-01	54.0	104.939394
1991-02-01	55.0	104.939394
1991-03-01	66.0	104.939394
1991-04-01	65.0	104.939394
1991-05-01	60.0	104.939394

Table 39: Mean Forecast



Graph 17: Simple Average

Model Evaluation:

```
For Simple Average forecast on Rose Test Data,  RMSE is 53.461
```

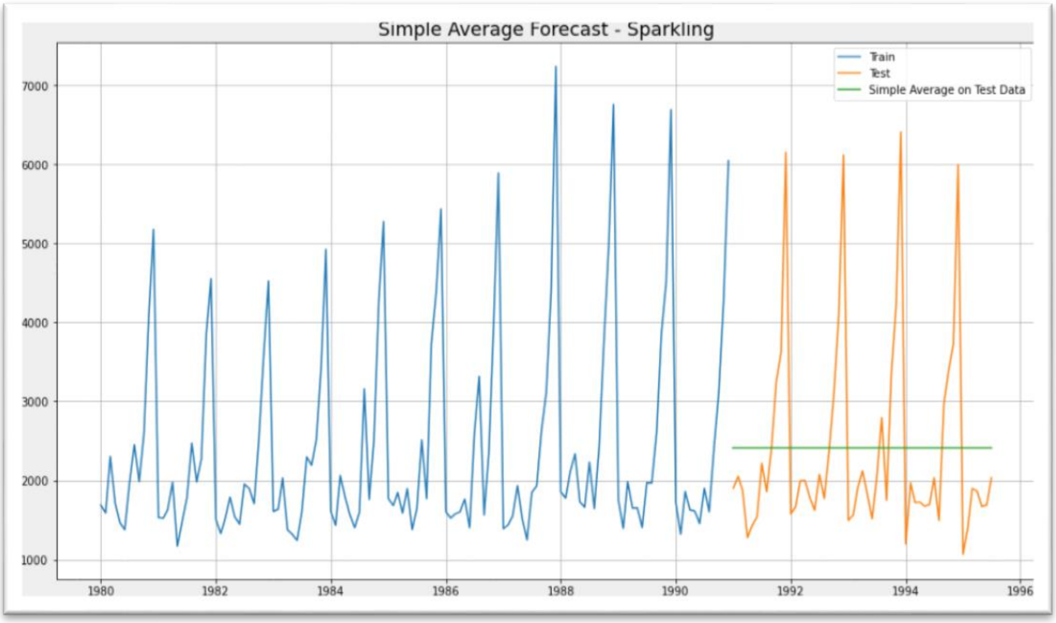
	Test RMSE Rose
SimpleAverageModel	53.46057

Table 40: Model Evaluation

b) For Sparkling Data:

YearMonth	Sparkling	mean_forecast
1991-01-01	1902	2403.780303
1991-02-01	2049	2403.780303
1991-03-01	1874	2403.780303
1991-04-01	1279	2403.780303
1991-05-01	1432	2403.780303

Table 41: Mean Forecast for Sparkling Data Set



Graph 18: Simple Average Forecast

Model Evaluation:

```
For Simple Average forecast on Sparkling Test Data,  RMSE is 1275.082
```

Table 42: Model Evaluation

	Test RMSE Sparkling
SimpleAverageModel	1275.081804

	Test RMSE Rose	Test RMSE Sparkling
SimpleAverageModel	53.46057	1275.081804

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352
SimpleAverageModel	53.460570	1275.081804

Table 43: Model Evaluation

Model 4: Moving Average (MA):

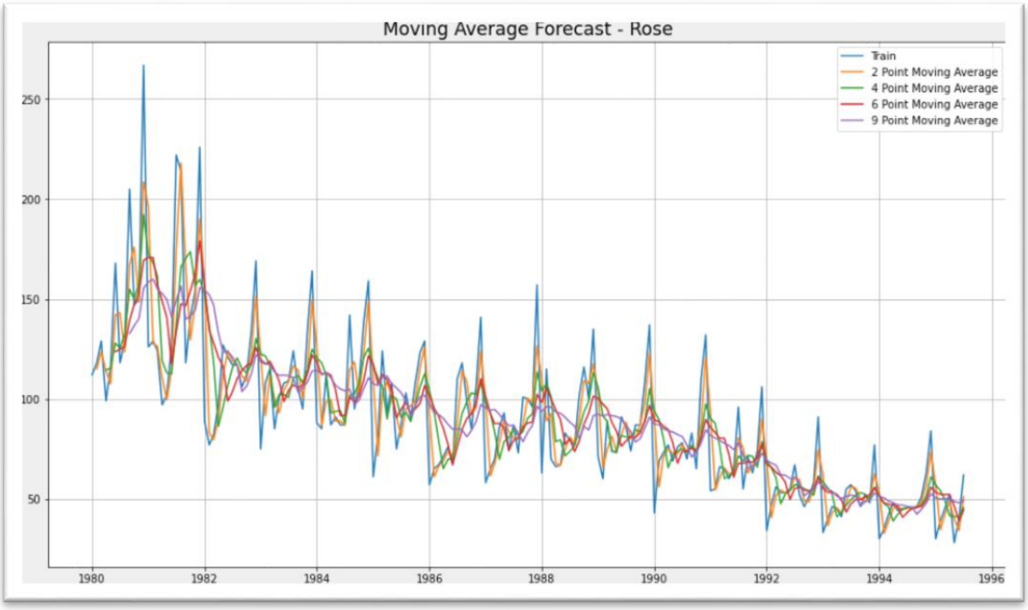
a) For Rose Data:

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

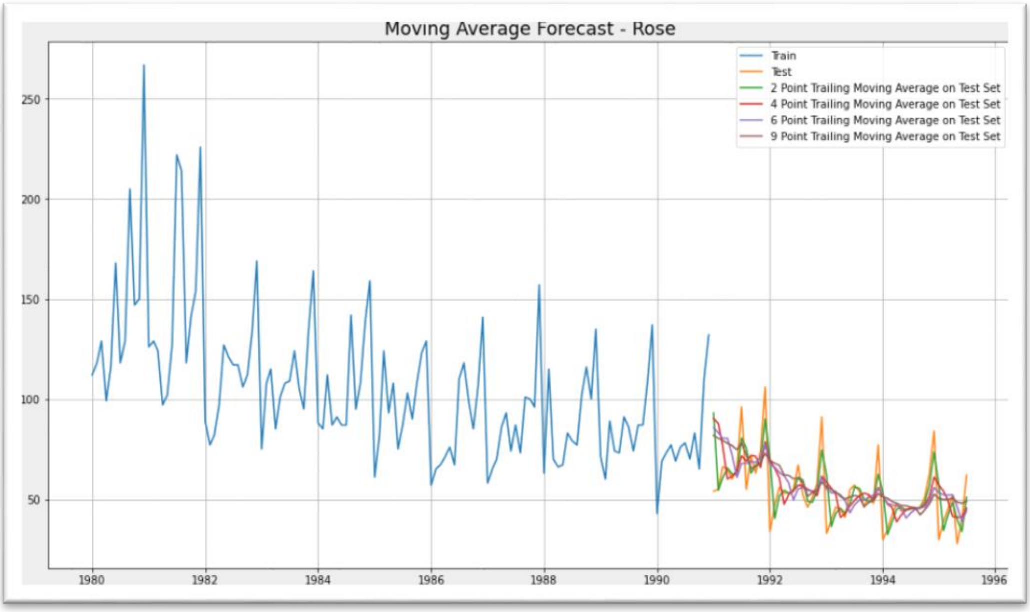
Table 44: Head of the Data set

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

Table 45: Trailing Moving Averages



Graph 19: Moving Average Forecast



Graph 20: Moving Average Forecast

Model Evaluation:

```
For 2 point Moving Average Model forecast on Rose Training Data,  RMSE is 11.529
For 4 point Moving Average Model forecast on Rose Training Data,  RMSE is 14.451
For 6 point Moving Average Model forecast on Rose Training Data,  RMSE is 14.566
For 9 point Moving Average Model forecast on Rose Training Data,  RMSE is 14.728
```

	Test RMSE Rose
2pointTrailingMovingAverage	11.529278
4pointTrailingMovingAverage	14.451403
6pointTrailingMovingAverage	14.566327
9pointTrailingMovingAverage	14.727630

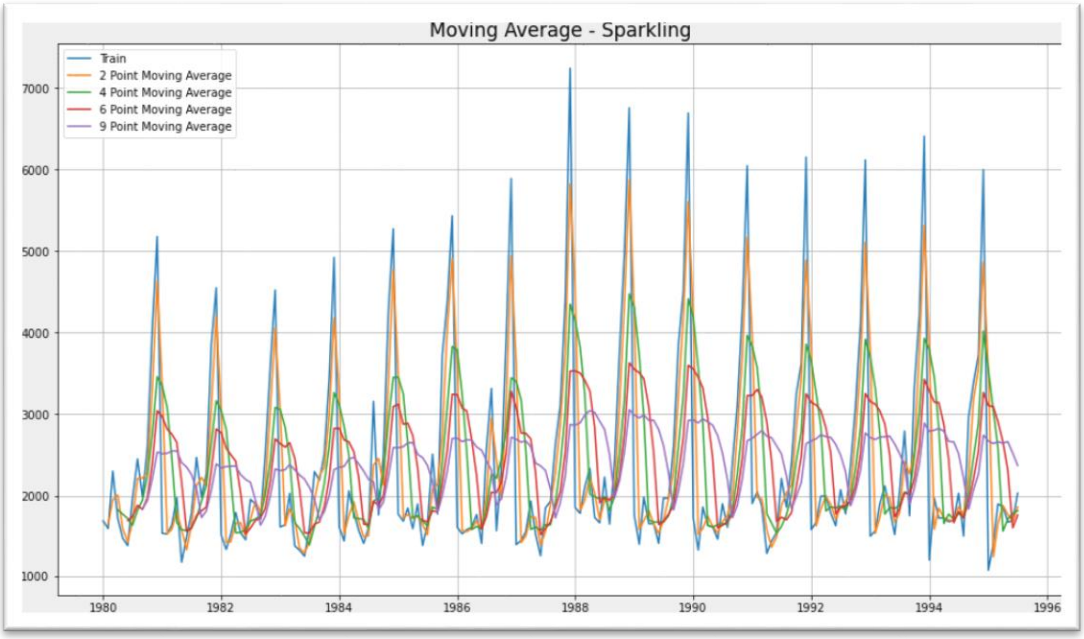
Table 46: RMSE Values

b) For Sparkling Data:

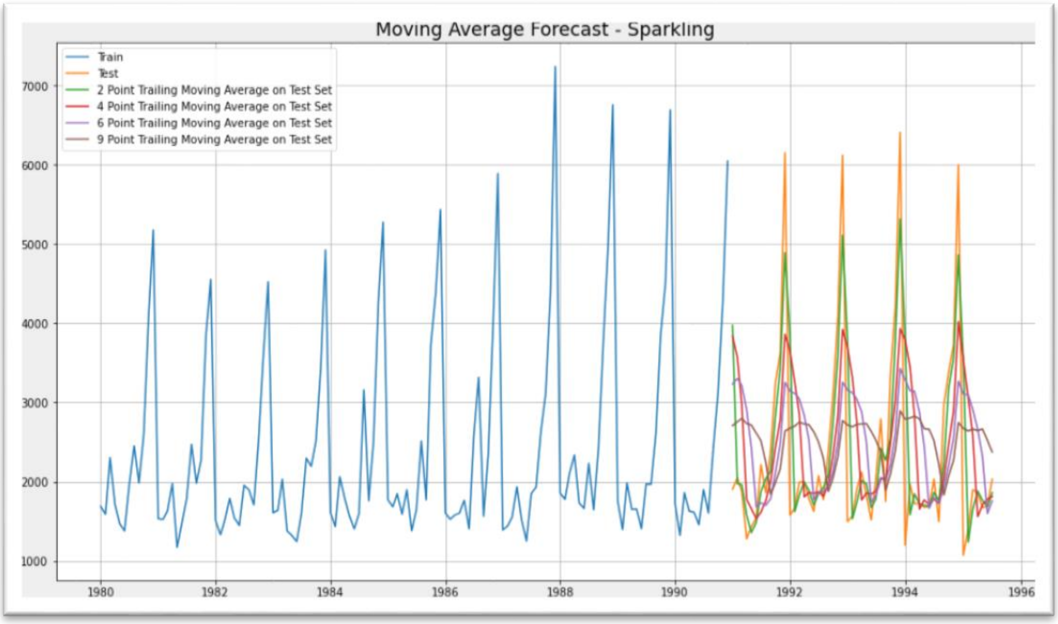
YearMonth	Sparkling
1980-01-01	1686
1980-02-01	1591
1980-03-01	2304
1980-04-01	1712
1980-05-01	1471

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

Table 47: Head of Sparkling Data along with Moving Averages Value



Graph 21: Moving Average Forecast



Graph 22: Moving Average Forecast

Model Evaluation:

```
For 2 point Moving Average Model forecast on Sparkling Training Data,  RMSE is 813.401
For 4 point Moving Average Model forecast on Sparkling Training Data,  RMSE is 1156.590
For 6 point Moving Average Model forecast on Sparkling Training Data,  RMSE is 1283.927
For 9 point Moving Average Model forecast on Sparkling Training Data,  RMSE is 1346.278
```

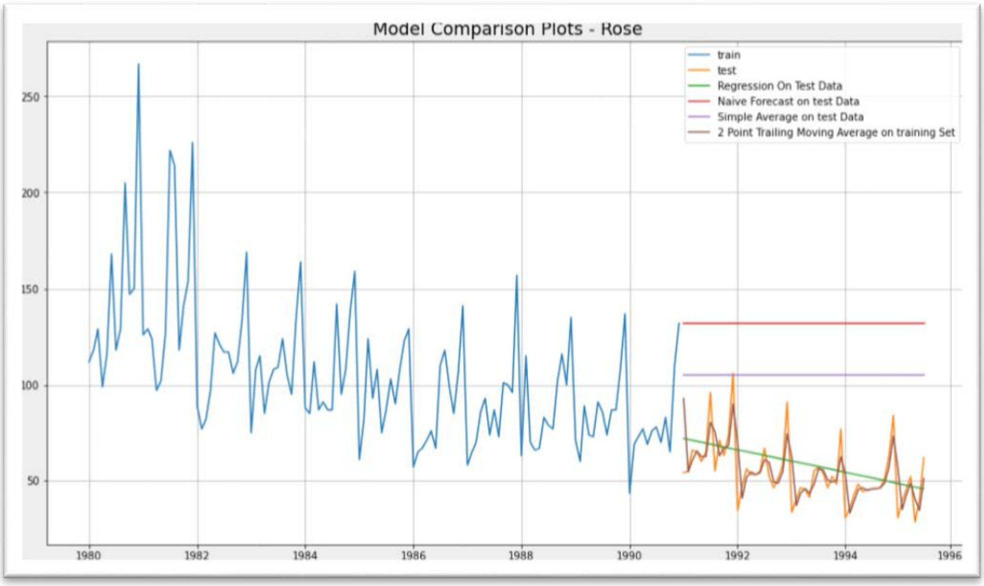
	Test RMSE Sparkling
2pointTrailingMovingAverage	813.400684
4pointTrailingMovingAverage	1156.589694
6pointTrailingMovingAverage	1283.927428
9pointTrailingMovingAverage	1346.278315

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

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

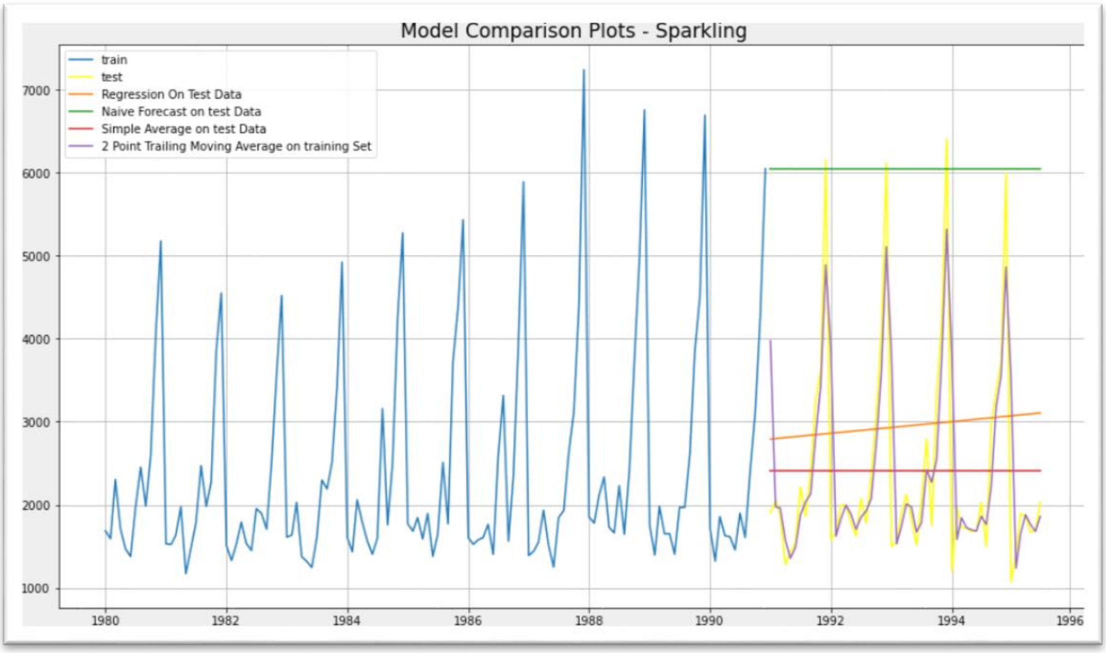
Table 48: Model Evaluation

Consolidated plots of all Models (For Rose Case):



Graph 23: Consolidated Plots of all Models

Consolidated plots of all Models (For Sparkling Case):



Graph 24: Consolidated Plots of all Models

- We have built 4 models till now for both Rose and Sparkling Wine datasets
- We fitted various models to the Train split and tested it on Test split. Accuracy metrics used is Root Mean Squared Error (RMSE) on Test data

Model 1 - Linear Regression:

- We regressed variables 'Rose' and 'Sparkling' against their individual time instances
- We modified the datasets and tagged individual sales to their time instances
- TEST RMSE ROSE = 15.27 | TEST RMSE SPARKLING = 1389.14

Model 2 - Naive Approach

- Naive approach says that prediction for tomorrow is same as today
- And, prediction for day-after is same as tomorrow
- So, effectively all future predictions are going to be same as today
- TEST RMSE ROSE = 79.72 | TEST RMSE SPARKLING = 3864.28

Model 3 - Simple Average

- All future predictions are the same as the simple average of all data till today
- TEST RMSE ROSE = 53.46 | TEST RMSE SPARKLING = 1275.08

Model 4 - Moving Average (MA)

- We calculate rolling means (Moving averages) over different intervals for the whole train data
- 2 Pt MA ====> means, we find average of 1st and 2nd to predict 3rd similarly, average of 2nd and 3rd to predict 4th and so on

- 4 Pt MA ====> means, we find average of 1st, 2nd, 3rd & 4th to predict 5th also, average of 2nd, 3rd, 4th & 5th to predict 6th and so on
  - 2 PT MA ====> TEST RMSE ROSE = 11.53 | TEST RMSE SPARKLING = 813.40
  - 4 PT MA ====> TEST RMSE ROSE = 14.45 | TEST RMSE SPARKLING = 1156.59
  - 6 PT MA ====> TEST RMSE ROSE = 14.57 | TEST RMSE SPARKLING = 1283.93

  - 9 PT MA ====> TEST RMSE ROSE = 14.73 | TEST RMSE SPARKLING = 1346.28

- Till now, Best Model which gives lowest RMSE score for both Rose and Sparkling is ——> 2 Pt Moving Average Model
- We'll continue to forecast using Exponential Smoothing Models for both datasets of Rose and Sparkling Wine Sales
- Exponential smoothing averages or exponentially weighted moving averages consist of forecast based on previous periods data with exponentially declining influence on the older observations
- Exponential smoothing methods consist of special case exponential moving with notation ETS (Error, Trend, Seasonality) where each can be None(N), Additive (N), Additive damped (Ad), Multiplicative (M) or Multiplicative damped (Md)
- One or more parameters control how fast the weights decay. The values of the parameters lie between 0 and 1

We'll build following Exponential Smoothing Models:

  - Single Exponential Smoothing with Additive Errors – ETS (A, N, N)
  - Double Exponential Smoothing with Additive Errors, Additive Trends – ETS (A, A, N)

- Triple Exponential Smoothing with Additive Errors, Additive Trends, Additive Seasonality – ETS (A, A, A)
- Triple Exponential Smoothing with Additive Errors, Additive Trends, Multiplicative Seasonality – ETS (A, A, M)
- Triple Exponential Smoothing with Additive Errors, Additive DAMPED Trends, Additive Seasonality – ETS (A, Ad, A)
- Triple Exponential Smoothing with Additive Errors, Additive DAMPED Trends, Multiplicative Seasonality – ETS (A, Ad, M)

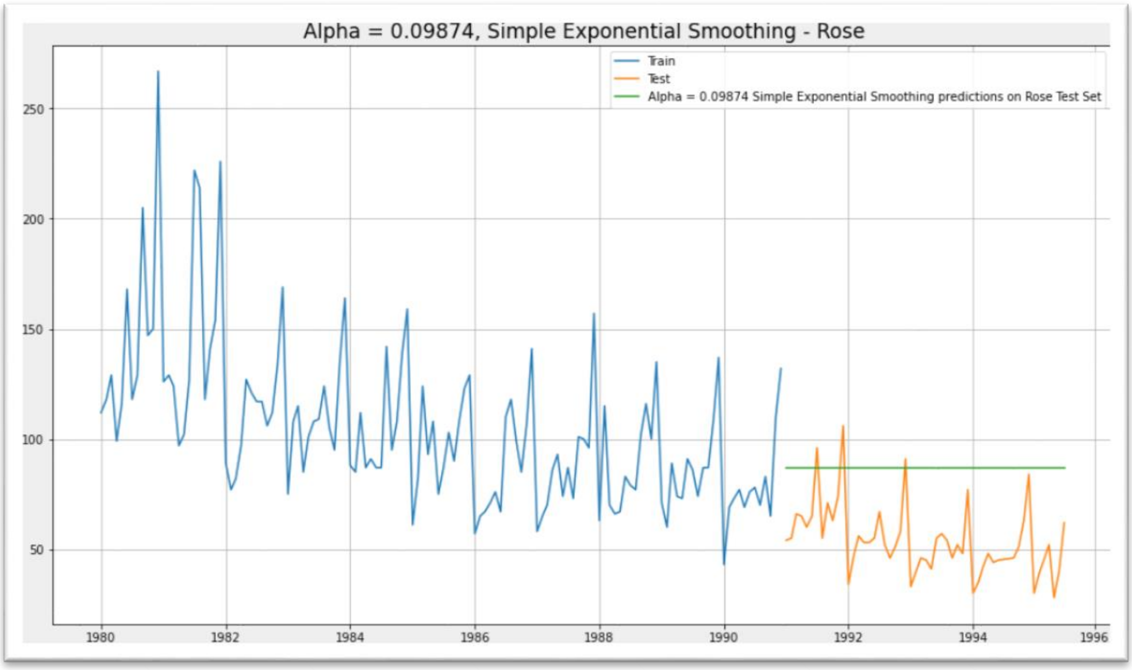
SES – ETS (A, N, N) - Simple Exponential Smoothing with additive errors - ROSE:

```
{'smoothing_level': 0.09874933517484011,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 134.38703609891138,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Table 49: Checking Parameters

```
1991-01-01    87.104984
1991-02-01    87.104984
1991-03-01    87.104984
1991-04-01    87.104984
1991-05-01    87.104984
1991-06-01    87.104984
1991-07-01    87.104984
1991-08-01    87.104984
1991-09-01    87.104984
1991-10-01    87.104984
1991-11-01    87.104984
1991-12-01    87.104984
1992-01-01    87.104984
1992-02-01    87.104984
1992-03-01    87.104984
1992-04-01    87.104984
1992-05-01    87.104984
1992-06-01    87.104984
1992-07-01    87.104984
1992-08-01    87.104984
1992-09-01    87.104984
1992-10-01    87.104984
1992-11-01    87.104984
1992-12-01    87.104984
1993-01-01    87.104984
1993-02-01    87.104984
1993-03-01    87.104984
1993-04-01    87.104984
1993-05-01    87.104984
1993-06-01    87.104984
1993-06-01    87.104984
1993-07-01    87.104984
1993-08-01    87.104984
1993-09-01    87.104984
1993-10-01    87.104984
1993-11-01    87.104984
1993-12-01    87.104984
1994-01-01    87.104984
1994-02-01    87.104984
1994-03-01    87.104984
1994-04-01    87.104984
1994-05-01    87.104984
1994-06-01    87.104984
1994-07-01    87.104984
1994-08-01    87.104984
1994-09-01    87.104984
1994-10-01    87.104984
1994-11-01    87.104984
1994-12-01    87.104984
1995-01-01    87.104984
1995-02-01    87.104984
1995-03-01    87.104984
1995-04-01    87.104984
1995-05-01    87.104984
1995-06-01    87.104984
1995-07-01    87.104984
Freq: MS, dtype: float64
```

Table 50: Training set for forecast



Graph 25: Simple Exponential Smoothing

```
SES RMSE: 36.79622798367751
SES RMSE (calculated using statsmodels): 36.796227983677504
```

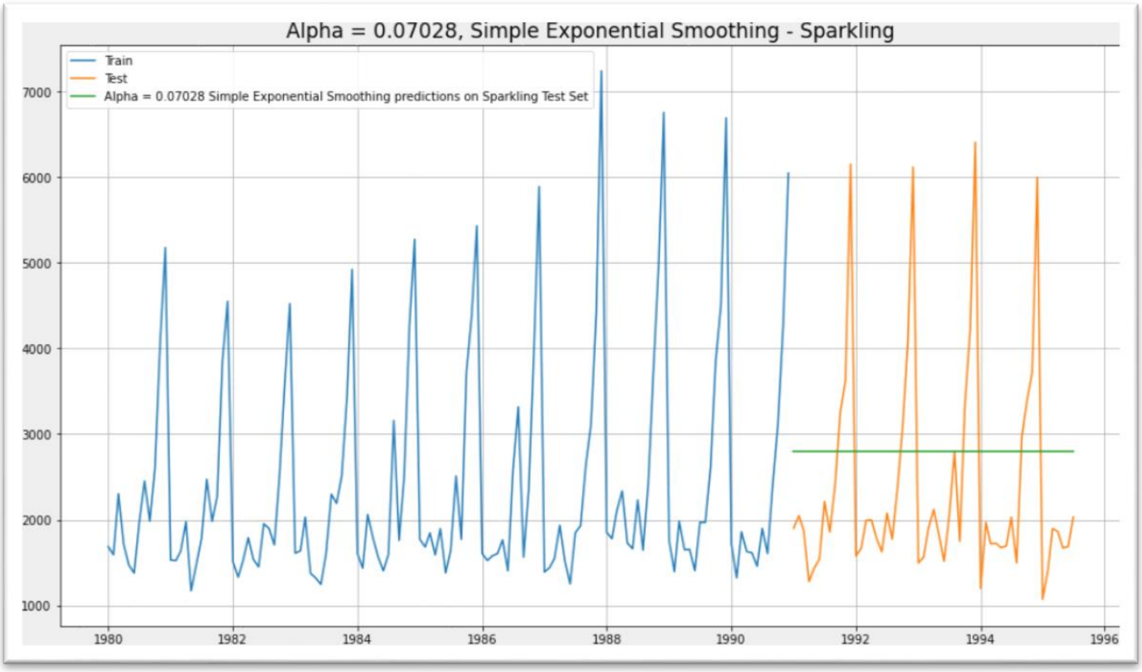
	Test RMSE Rose
Simple Exponential Smoothing	36.796228

Table 51: Model Evaluation

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

```
{'smoothing_level': 0.07029459943040381,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 1764.1004162520212,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

Table 52: Checking the parameters



Graph 26: Simple Exponential Smoothing

1991-01-01	2804.687529
1991-02-01	2804.687529
1991-03-01	2804.687529
1991-04-01	2804.687529
1991-05-01	2804.687529
1991-06-01	2804.687529
1991-07-01	2804.687529
1991-08-01	2804.687529
1991-09-01	2804.687529
1991-10-01	2804.687529
1991-11-01	2804.687529
1991-12-01	2804.687529
1992-01-01	2804.687529
1992-02-01	2804.687529
1992-03-01	2804.687529
1992-04-01	2804.687529
1992-05-01	2804.687529
1992-06-01	2804.687529
1992-07-01	2804.687529
1992-08-01	2804.687529
1992-09-01	2804.687529
1992-10-01	2804.687529
1992-11-01	2804.687529
1992-12-01	2804.687529
1993-01-01	2804.687529

1993-01-01	2804.687529
1993-02-01	2804.687529
1993-03-01	2804.687529
1993-04-01	2804.687529
1993-05-01	2804.687529
1993-06-01	2804.687529
1993-07-01	2804.687529
1993-08-01	2804.687529
1993-09-01	2804.687529
1993-10-01	2804.687529
1993-11-01	2804.687529
1993-12-01	2804.687529
1994-01-01	2804.687529
1994-02-01	2804.687529
1994-03-01	2804.687529
1994-04-01	2804.687529
1994-05-01	2804.687529
1994-06-01	2804.687529
1994-07-01	2804.687529
1994-08-01	2804.687529
1994-09-01	2804.687529
1994-10-01	2804.687529
1994-11-01	2804.687529
1994-12-01	2804.687529
1995-01-01	2804.687529

1994-08-01	2804.687529
1994-09-01	2804.687529
1994-10-01	2804.687529
1994-11-01	2804.687529
1994-12-01	2804.687529
1995-01-01	2804.687529
1995-02-01	2804.687529
1995-03-01	2804.687529
1995-04-01	2804.687529
1995-05-01	2804.687529
1995-06-01	2804.687529
1995-07-01	2804.687529
Freq: MS, dtype: float64	

Table 53: Training set for forecast

```
SES RMSE: 1338.0121443910186
SES RMSE (calculated using statsmodels): 2750.8781279287596
```

	Test RMSE Sparkling
Simple Exponential Smoothing	1338.012144

	Test RMSE Rose	Test RMSE Sparkling
Simple Exponential Smoothing	36.796228	1338.012144

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352
SimpleAverageModel	53.460570	1275.081804
2pointTrailingMovingAverage	11.529278	813.400684
4pointTrailingMovingAverage	14.451403	1156.589694
6pointTrailingMovingAverage	14.566327	1283.927428
9pointTrailingMovingAverage	14.727630	1346.278315
Simple Exponential Smoothing	36.796228	1338.012144

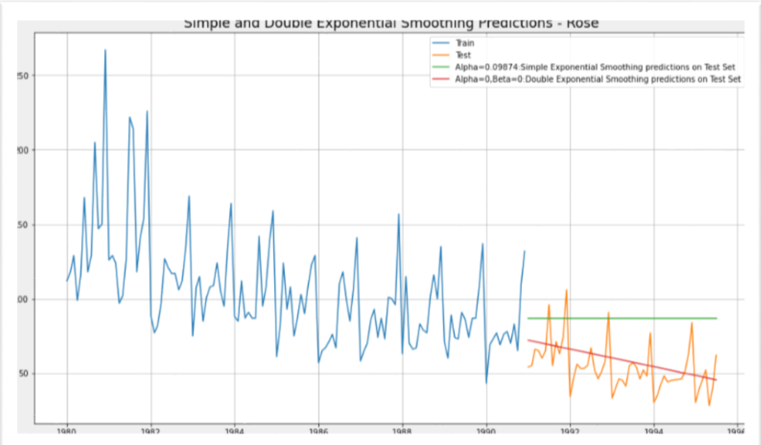
Table 54: Model Evaluation

- For Rose - Level Parameter, Alpha = 0.09874
- For Sparkling - Level Parameter, Alpha = 0.07028

Holt - ETS(A, A, N) - Holt's linear method with additive errors - Rose Double Exponential Smoothing – Rose

```
Holt model Exponential Smoothing Estimated Parameters :

{'smoothing_level': 1.9086427682180844e-08, 'smoothing_trend': 7.302464353829351e-09, 'smoothing_seasonal': nan, 'damping_tren
d': nan, 'initial_level': 137.81629861505857, 'initial_trend': -0.4943753249082896, 'initial_seasons': array([], dtype=float6
4), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 55: Checking for parameters



Graph 27: Double Exponential Smoothing

1991-01-01	72.064380
1991-02-01	71.570005
1991-03-01	71.075630
1991-04-01	70.581254
1991-05-01	70.086879
1991-06-01	69.592504
1991-07-01	69.098128
1991-08-01	68.603753
1991-09-01	68.109378
1991-10-01	67.615002
1991-11-01	67.120627
1991-12-01	66.626252
1992-01-01	66.131877
1992-02-01	65.637501
1992-03-01	65.143126
1992-04-01	64.648751
1992-05-01	64.154375
1992-06-01	63.660000
1992-07-01	63.165625
1992-08-01	62.671249
1992-09-01	62.176874
1992-10-01	61.682499
1992-11-01	61.188123
1992-12-01	60.693748
1993-01-01	60.199373
1993-02-01	59.704997
1992-03-01	65.143126
1992-04-01	64.648751
1992-05-01	64.154375
1992-06-01	63.660000
1992-07-01	63.165625
1992-08-01	62.671249
1992-09-01	62.176874
1992-10-01	61.682499
1992-11-01	61.188123
1992-12-01	60.693748
1993-01-01	60.199373
1993-02-01	59.704997
1993-03-01	59.210622
1993-04-01	58.716247
1993-05-01	58.221871
1993-06-01	57.727496
1993-07-01	57.233121
1993-08-01	56.738745
1993-09-01	56.244370
1993-10-01	55.749995
1993-11-01	55.255619
1993-12-01	54.761244
1994-01-01	54.266869
1994-02-01	53.772493
1994-03-01	53.278118
1994-04-01	52.783743
1994-05-01	52.289367
1994-06-01	51.794992
1994-07-01	51.300617
1994-08-01	50.806241
1994-09-01	50.311866
1994-10-01	49.817491
1994-11-01	49.323115
1994-12-01	48.828740
1995-01-01	48.334365
1995-02-01	47.839989
1995-03-01	47.345614
1995-04-01	46.851239
1995-05-01	46.356864
1995-06-01	45.862488
1995-07-01	45.368113

Freq: MS, dtype: float64

Table 56: Training set for forecast

Model Evaluation:

```
DES RMSE: 15.269327872865668
```

	Test RMSE Rose
Double Exponential Smoothing	15.269328

Table 57: Model Evaluation

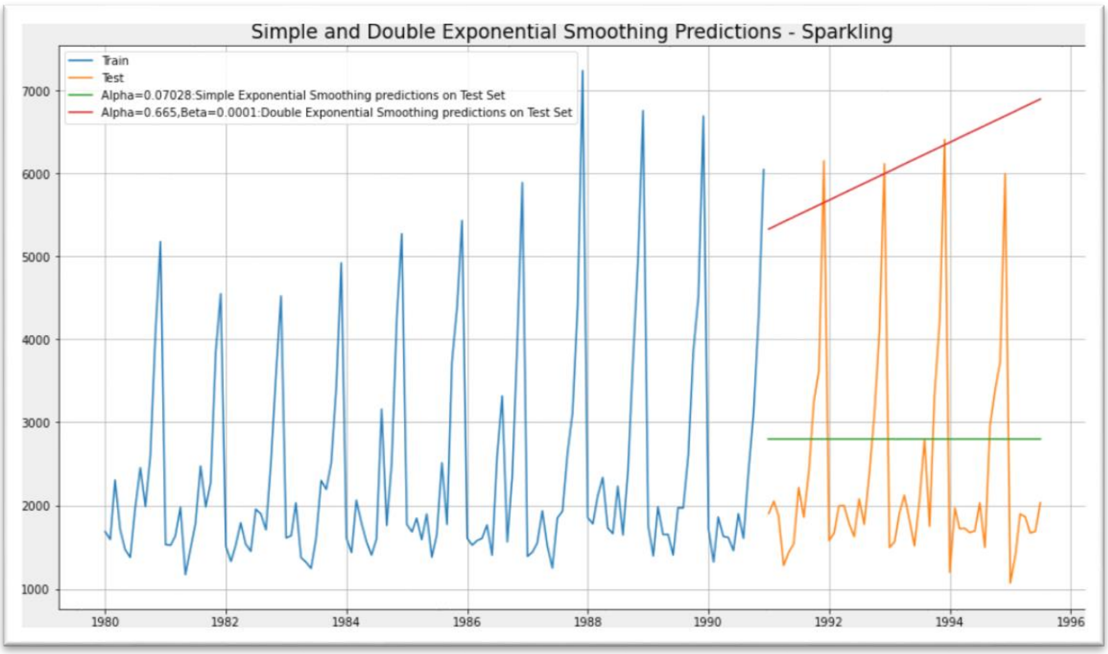
Holt - ETS(A, A, N) - Holt's linear method with additive errors - Sparkling Double Exponential Smoothing – Sparkling

```
Holt model Exponential Smoothing Estimated Parameters :

{'smoothing_level': 0.6638769092832238, 'smoothing_trend': 9.966251357628782e-05, 'smoothing_seasonal': nan, 'damping_trend': n
an, 'initial_level': 1502.5681711003654, 'initial_trend': 29.020225552837097, 'initial_seasons': array([], dtype=float64), 'use
_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 58: Checking for Parameters

1991-01-01	5330.501799	1992-01-01	5678.722657		
1991-02-01	5359.520204	1992-02-01	5707.741062		
1991-03-01	5388.538609	1992-03-01	5736.759467		
1991-04-01	5417.557013	1992-04-01	5765.777872		
1991-05-01	5446.575418	1992-05-01	5794.796277	1994-01-01	6375.164374
1991-06-01	5475.593823	1992-06-01	5823.814682	1994-02-01	6404.182779
1991-07-01	5504.612228	1992-07-01	5852.833086	1994-03-01	6433.201184
1991-08-01	5533.630633	1992-08-01	5881.851491	1994-04-01	6462.219588
1991-09-01	5562.649038	1992-09-01	5910.869896	1994-05-01	6491.237993
1991-10-01	5591.667443	1992-10-01	5939.888301	1994-06-01	6520.256398
1991-11-01	5620.685847	1992-11-01	5968.906706	1994-07-01	6549.274803
1991-12-01	5649.704252	1992-12-01	5997.925111	1994-08-01	6578.293208
1992-01-01	5678.722657	1993-01-01	6026.943516	1994-09-01	6607.311613
1992-02-01	5707.741062	1993-02-01	6055.961920	1994-10-01	6636.330018
1992-03-01	5736.759467	1993-03-01	6084.980325	1994-11-01	6665.348422
1992-04-01	5765.777872	1993-04-01	6113.998730	1994-12-01	6694.366827
1992-05-01	5794.796277	1993-05-01	6143.017135	1995-01-01	6723.385232
1992-06-01	5823.814682	1993-06-01	6172.035540	1995-02-01	6752.403637
1992-07-01	5852.833086	1993-07-01	6201.053945	1995-03-01	6781.422042
1992-08-01	5881.851491	1993-08-01	6230.072350	1995-04-01	6810.440447
1992-09-01	5910.869896	1993-09-01	6259.090754	1995-05-01	6839.458852
1992-10-01	5939.888301	1993-10-01	6288.109159	1995-06-01	6868.477256
1992-11-01	5968.906706	1993-11-01	6317.127564	1995-07-01	6897.495661
1992-12-01	5997.925111	1993-12-01	6346.145969	Freq: MS, dtype: float64	

Table 59: Training set for forecast



Graph 28: Double Exponential Smoothing

Model Evaluation:

```
DES RMSE: 3949.993290409098
```

	Test RMSE Sparkling
Double Exponential Smoothing	3949.99329

	Test RMSE Rose	Test RMSE Sparkling
Double Exponential Smoothing	15.269328	3949.99329

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352
SimpleAverageModel	53.460570	1275.081804
2pointTrailingMovingAverage	11.529278	813.400684
4pointTrailingMovingAverage	14.451403	1156.589694
6pointTrailingMovingAverage	14.566327	1283.927428
9pointTrailingMovingAverage	14.727630	1346.278315
Simple Exponential Smoothing	36.796228	1338.012144
Double Exponential Smoothing	15.269328	3949.993290

Table 60: Model Evaluation

- In Rose - DES has picked up the trend well. DES seems to perform better than SES here
- In Sparkling - DES shows a non-existent trend. DES does not perform well here

- Rose - Level parameter, Alpha = 0
  Trend parameter, Beta = 0

- Sparkling - Level parameter, Alpha = 0.665
  Trend parameter, Beta = 0.0001

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

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.08830330642635406, 'smoothing_trend': 6.730635331927582e-05, 'smoothing_seasonal': 0.004455138229351625,
'damping_trend': nan, 'initial_level': 146.88752868155674, 'initial_trend': -0.5492163940406024, 'initial_seasons': array([-31.
12207537, -18.81171138, -10.86052241, -21.52235816,
       -12.68359535,  -7.17529564,   2.7456236 ,   8.84900094,
         4.85724354,   2.9520333 ,  21.05004912,  63.29916317]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

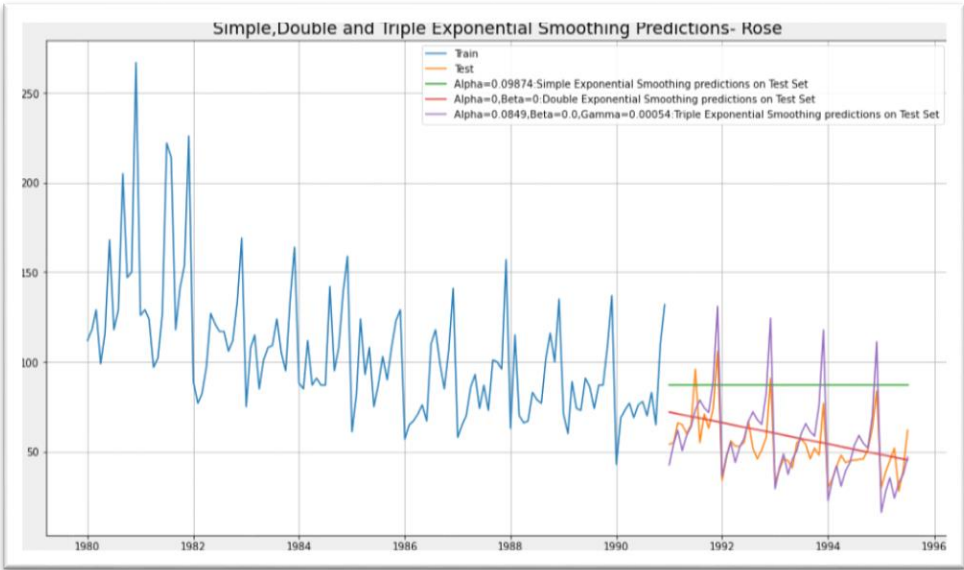
Table 61: Checking for Parameter

1991-01-01	42.672382	1993-01-01	29.491193
1991-02-01	54.439917	1993-02-01	41.258728
1991-03-01	61.841877	1993-03-01	48.660688
1991-04-01	50.636896	1993-04-01	37.455706
1991-05-01	58.918913	1993-05-01	45.737724
1991-06-01	63.870294	1993-06-01	50.689105
1991-07-01	73.240626	1993-07-01	60.059436
1991-08-01	78.790723	1993-08-01	65.609534
1991-09-01	74.257853	1993-09-01	61.076664
1991-10-01	71.805821	1993-10-01	58.624632
1991-11-01	89.354796	1993-11-01	76.173607
1991-12-01	131.072194	1993-12-01	117.891005
1992-01-01	36.081787	1994-01-01	22.900598
1992-02-01	47.849323	1994-02-01	34.668134
1992-03-01	55.251283	1994-03-01	42.070093
1992-04-01	44.046301	1994-04-01	30.865112
1992-05-01	52.328318	1994-05-01	39.147129
1992-06-01	57.279699	1994-06-01	44.098510
1992-07-01	66.650031	1994-07-01	53.468842
1992-08-01	72.200129	1994-08-01	59.018940
1992-09-01	67.667259	1994-09-01	54.486070
1992-10-01	65.215226	1994-10-01	52.034037
1992-11-01	82.764201	1994-11-01	69.583012
1992-12-01	124.481599	1994-12-01	111.300410

1995-01-01	16.310004
1995-02-01	28.077539
1995-03-01	35.479499
1995-04-01	24.274517
1995-05-01	32.556534
1995-06-01	37.507916
1995-07-01	46.878247

Freq: MS, dtype: float64

Table 62: Training set for forecast



Graph 29: Triple Exponential Smoothing

Model Evaluation:

TES RMSE: 14.265712941547475

	Test RMSE Rose
Triple Exponential Smoothing (Additive Season)	14.265713

Table 63: Model Evaluation

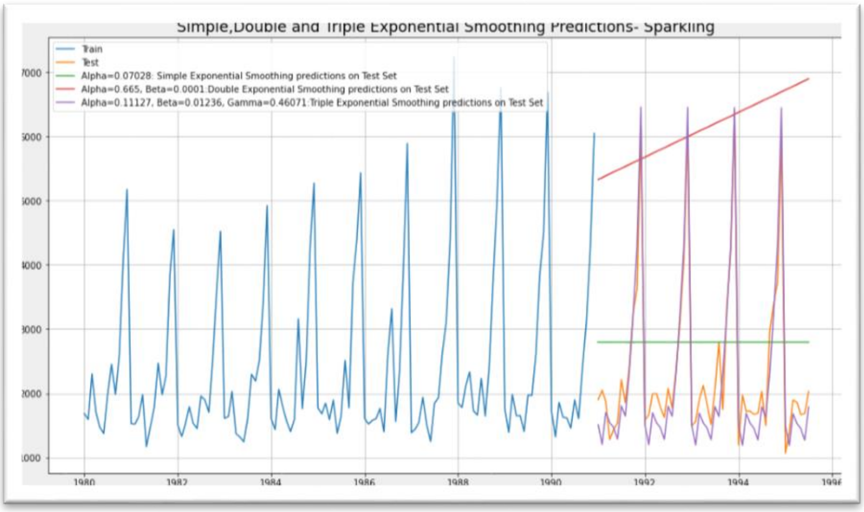
Holt-Winters - ETS(A, A, A) - Holt Winter's linear method with additive errors - SPARKLING:

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.10005373820823961, 'smoothing_trend': 0.010034490652580457, 'smoothing_seasonal': 0.5095957543425532, 'da
mping_trend': nan, 'initial_level': 2364.584774604334, 'initial_trend': -0.016752880078245408, 'initial_seasons': array([-653.8
2559323, -736.67734144, -368.25456128, -483.63906084,
       -826.15467946, -832.96819741, -386.3751117 ,   91.82676187,
       -261.32455153,  265.38968222, 1580.26233564, 2619.56221896]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 64: Checking for Parameters

1991-01-01	1509.969093	1993-01-01	1501.192388		
1991-02-01	1205.343244	1993-02-01	1196.566540		
1991-03-01	1702.386113	1993-03-01	1693.609408		
1991-04-01	1548.514691	1993-04-01	1539.737987		
1991-05-01	1467.824074	1993-05-01	1459.047370		
1991-06-01	1287.109239	1993-06-01	1278.332535		
1991-07-01	1804.027662	1993-07-01	1795.250957		
1991-08-01	1646.339830	1993-08-01	1637.563125		
1991-09-01	2326.596637	1993-09-01	2317.819933		
1991-10-01	3228.612283	1993-10-01	3219.835579		
1991-11-01	4303.269444	1993-11-01	4294.492739		
1991-12-01	6460.680111	1993-12-01	6451.903407		
1992-01-01	1505.580741	1994-01-01	1496.804036		
1992-02-01	1200.954892	1994-02-01	1192.178188		
1992-03-01	1697.997760	1994-03-01	1689.221056		
1992-04-01	1544.126339	1994-04-01	1535.349634		
1992-05-01	1463.435722	1994-05-01	1454.659018	1995-01-01	1492.415684
1992-06-01	1282.720887	1994-06-01	1273.944182	1995-02-01	1187.789835
1992-07-01	1799.639310	1994-07-01	1790.862605	1995-03-01	1684.832704
1992-08-01	1641.951478	1994-08-01	1633.174773	1995-04-01	1530.961282
1992-09-01	2322.208285	1994-09-01	2313.431581	1995-05-01	1450.270666
1992-10-01	3224.223931	1994-10-01	3215.447226	1995-06-01	1269.555830
1992-11-01	4298.881092	1994-11-01	4290.104387	1995-07-01	1786.474253
1992-12-01	6456.291759	1994-12-01	6447.515054	Freq: MS, dtype: float64	

Table 65: Training set for Forecast



Graph 30: Triple Exponential Smoothing

Model Evaluation:

```
TES RMSE: 379.6956857387101
```

	Test RMSE Sparkling
Triple Exponential Smoothing (Additive Season)	379.695686

	Test RMSE Rose	Test RMSE Sparkling
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352
SimpleAverageModel	53.460570	1275.081804
2pointTrailingMovingAverage	11.529278	813.400684
4pointTrailingMovingAverage	14.451403	1156.589694
6pointTrailingMovingAverage	14.566327	1283.927428
9pointTrailingMovingAverage	14.727630	1346.278315
Simple Exponential Smoothing	36.796228	1338.012144
Double Exponential Smoothing	15.269328	3949.993290
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686

Table 66: Model Evaluation

- In Rose & Sparkling - TES has picked up the trend and seasonality very well

- Rose:
  - Level parameter, Alpha = 0.0849
  - Trend parameter, Beta = 0.0
  - Seasonality parameter, Gamma = 0.00054

- Sparkling:
  - Level parameter, Alpha = 0.11127
  - Trend parameter, Beta = 0.01236
  - Seasonality parameter, Gamma = 0.46071

- Till now, Best Model for Rose ——> 2 Pt Moving Average Best Model for Sparkling ——> Holt-Winter - ETS (A, A, A)

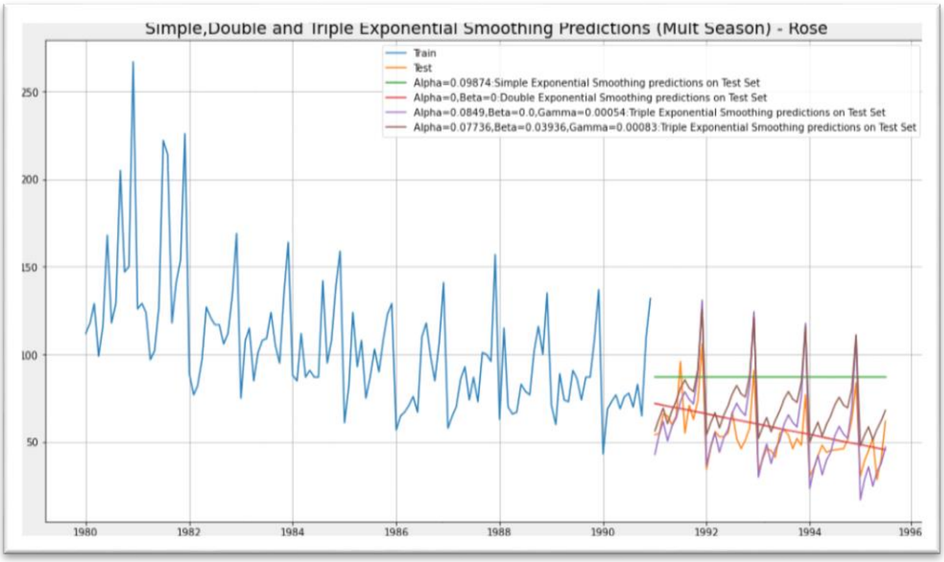
Holt-Winters - ETS(A, A, M) - Holt Winter's linear method - ROSE ETS(A, A, M) model - Taking MULTIPLICATIVE SEASONALITY

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.07132109562890512, 'smoothing_trend': 0.04553831096563722, 'smoothing_seasonal': 8.356711212063695e-07,
'damping_trend': nan, 'initial_level': 134.25655591779326, 'initial_trend': -0.8038265942903572, 'initial_seasons': array([0.83
746068, 0.94985307, 1.03812083, 0.90732186, 1.02043162,
       1.11131741, 1.22228039, 1.30104211, 1.23132915, 1.20610008,
       1.40577823, 1.93832412]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 67: Checking for Parameters

1991-01-01	56.334597	1993-01-01	52.038725		
1991-02-01	63.692059	1993-02-01	58.819652		
1991-03-01	69.388935	1993-03-01	64.063746		
1991-04-01	60.452304	1993-04-01	55.798066		
1991-05-01	67.770362	1993-05-01	62.535913		
1991-06-01	73.568837	1993-06-01	67.868180		
1991-07-01	80.653311	1993-07-01	74.383453		
1991-08-01	85.572391	1993-08-01	78.898513		
1991-09-01	80.724066	1993-09-01	74.407789		
1991-10-01	78.812306	1993-10-01	72.625445		
1991-11-01	91.559770	1993-11-01	84.348631		
1991-12-01	125.830690	1993-12-01	115.887785		
1992-01-01	54.186661	1994-01-01	49.890789		
1992-02-01	61.255856	1994-02-01	56.383449		
1992-03-01	66.726340	1994-03-01	61.401152		
1992-04-01	58.125185	1994-04-01	53.470947		
1992-05-01	65.153138	1994-05-01	59.918689	1995-01-01	47.742853
1992-06-01	70.718509	1994-06-01	65.017851	1995-02-01	53.947245
1992-07-01	77.518382	1994-07-01	71.248524	1995-03-01	58.738557
1992-08-01	82.235452	1994-08-01	75.561575	1995-04-01	51.143828
1992-09-01	77.565928	1994-09-01	71.249651	1995-05-01	57.301464
1992-10-01	75.718876	1994-10-01	69.532014	1995-06-01	62.167522
1992-11-01	87.954200	1994-11-01	80.743061	1995-07-01	68.113595
1992-12-01	120.859237	1994-12-01	110.916332	Freq: MS, dtype: float64	

Table 68: Training set for Forecast



Graph 31: Triple Smoothing with Multiplicative Seasonality

Report model accuracy:

```
TES_am RMSE_rose: 20.190997935033817
```

	Test RMSE Rose
Triple Exponential Smoothing (Multiplicative Season)	20.190998

Table 69: Model Evaluation

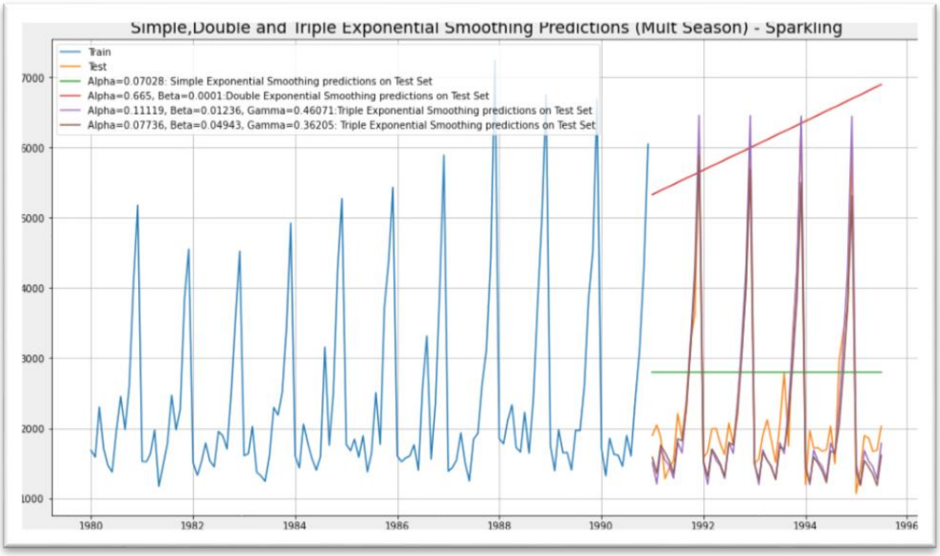
Holt-Winters - ETS(A, A, M) - Holt Winter's linear method - SPARKLING ETS(A, A, M) model - Taking MULTIPLICATIVE SEASONALITY

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.11194572287706502, 'smoothing_trend': 0.04979454913988668, 'smoothing_seasonal': 0.3616765678435302, 'dam
ping_trend': nan, 'initial_level': 2356.340229937152, 'initial_trend': -10.519480221963526, 'initial_seasons': array([0.7146511
8, 0.68302129, 0.90263858, 0.80589958, 0.65660325,
       0.65654363, 0.88525948, 1.132562  , 0.92225104, 1.21110112,
       1.8820382 , 2.38194187]), 'use boxcox': False, 'lamda': None, 'remove bias': False}
```

Table 70: Checking for Parameters

```
1991-01-01    1586.782642
1991-02-01    1355.896477
1991-03-01    1762.095344
1991-04-01    1655.471900
1991-05-01    1541.320914
1991-06-01    1354.477040
1991-07-01    1853.306666
1991-08-01    1820.010020
1991-09-01    2275.925030
1991-10-01    3120.245508
1991-11-01    4126.116727
1991-12-01    5885.056602
1992-01-01    1536.931357
1992-02-01    1313.187030
1992-03-01    1706.444942
1992-04-01    1603.050911
1992-05-01    1492.385418
1992-06-01    1311.359584
1992-07-01    1794.152902
1992-08-01    1761.764092
1992-09-01    2202.893649
1992-10-01    3019.852559
1992-11-01    3993.003270
1992-12-01    5694.685845
1993-01-01    1487.080071
1993-02-01    1270.477583
1993-03-01    1650.794539
1993-04-01    1550.629921
1993-05-01    1443.449922
1993-06-01    1268.242128
1993-07-01    1734.999138
1993-08-01    1703.518165
1993-09-01    2129.862267
1993-10-01    2919.459610
1993-11-01    3859.889814
1993-12-01    5504.315087
1994-01-01    1437.228785
1994-02-01    1227.768135
1994-03-01    1595.144137
1994-04-01    1498.208932
1994-05-01    1394.514427
1994-06-01    1225.124673
1994-07-01    1675.845374
1994-08-01    1645.272238
1994-09-01    2056.830886
1994-10-01    2819.066661
1994-11-01    3726.776358
1994-12-01    5313.944330
1995-01-01    1387.377499
1995-02-01    1185.058688
1995-03-01    1539.493734
1995-04-01    1445.787942
1995-05-01    1345.578931
1995-06-01    1182.007217
1995-07-01    1616.691610
Freq: MS, dtype: float64
```

Table 71: Training set for Forecast



Graph 32: Triple Smoothing with Multiplicative Seasonality

Report model accuracy:

```
TES_am RMSE_spark: 406.51016963157673
```

	Test RMSE Sparkling
Triple Exponential Smoothing (Multiplicative Season)	406.51017

	Test RMSE Rose	Test RMSE Sparkling
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.51017

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352
SimpleAverageModel	53.460570	1275.081804
2pointTrailingMovingAverage	11.529278	813.400684
4pointTrailingMovingAverage	14.451403	1156.589694
6pointTrailingMovingAverage	14.566327	1283.927428
9pointTrailingMovingAverage	14.727630	1346.278315
Simple Exponential Smoothing	36.796228	1338.012144
Double Exponential Smoothing	15.269328	3949.993290
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.510170

Table 72: Model Evaluation

- Rose:
  - Level parameter, Alpha = 0.07736
  - Trend parameter, Beta = 0.03936
  - Seasonality parameter, Gamma = 0.00083

- Sparkling:
  - Level parameter, Alpha = 0.07736
  - Trend parameter, Beta = 0.04943
  - Seasonality parameter, Gamma = 0.36205

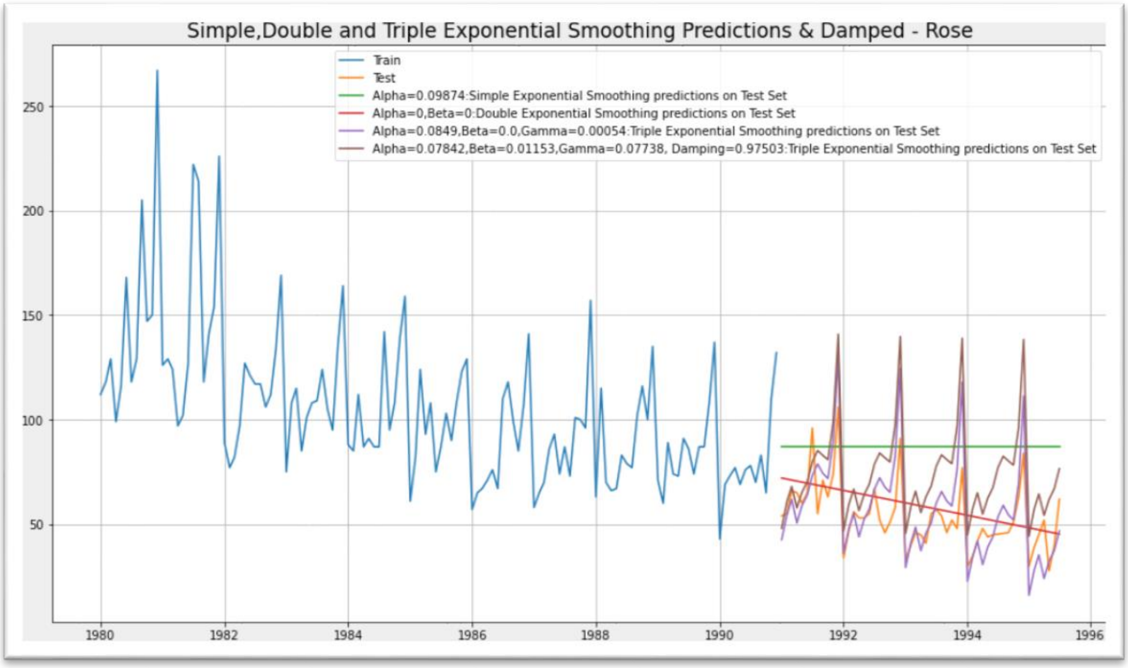
Holt-Winters - ETS(A, Ad, A) - Holt Winter's linear method with additive errors - ROSE USING DAMPED TREND

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.07848124317186382, 'smoothing_trend': 0.011340806743727252, 'smoothing_seasonal': 0.07684518378585875, 'd
amping_trend': 0.9774271738153909, 'initial_level': 153.42198975882621, 'initial_trend': -1.4893377479142316, 'initial_season
s': array([-30.39209525, -18.87017253, -10.88179619, -22.89643544,
       -13.58410591,  -6.64926869,   3.49196919,  10.51827686,
         6.1270493 ,   3.40480373,  21.30972085,  66.57108112]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 73: Checking for Parameters

```
1991-01-01     48.110296
1991-02-01     60.740191
1991-03-01     68.048590
1991-04-01     57.771836
1991-05-01     65.267144
1991-06-01     70.341710
1991-07-01     79.810801
1991-08-01     85.193659
1991-09-01     82.867091
1991-10-01     80.820114
1991-11-01     98.825637
1991-12-01    140.806084
1992-01-01     46.749347
1992-02-01     59.409963
1992-03-01     66.748388
1992-04-01     56.500984
1992-05-01     64.024978
1992-06-01     69.127584
1992-07-01     78.624080
1992-08-01     84.033726
1992-09-01     81.733342
1992-10-01     79.711957
1992-11-01     97.742493
1992-12-01    139.747390
1993-01-01     45.714550
1993-02-01     58.398525
1993-03-01     65.759781
1993-04-01     55.534693
1993-05-01     63.080499
1993-06-01     68.204424
1993-07-01     77.721759
1993-08-01     83.151772
1993-09-01     80.871296
1993-10-01     78.869370
1993-11-01     96.918926
1993-12-01    138.942414
1994-01-01     44.927744
1994-02-01     57.629479
1994-03-01     65.008095
1994-04-01     54.799974
1994-05-01     62.362365
1994-06-01     67.502500
1994-07-01     77.035680
1994-08-01     82.481180
1994-09-01     80.215841
1994-10-01     78.228710
1994-11-01     96.292728
1994-12-01    138.330350
1995-01-01     44.329497
1995-02-01     57.044736
1995-03-01     64.436551
1995-04-01     54.241332
1995-05-01     61.816333
1995-06-01     66.968794
1995-07-01     76.514020
Freq: MS, dtype: float64
```

Table 74: Training set for Forecast



Graph 33: Triple Smoothing using Damped Trend

Model Evaluation:

```
TES RMSE DAMPED ROSE: 25.660960416919345
```

	Test RMSE Rose
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.66096

Table 75: Model Evaluation

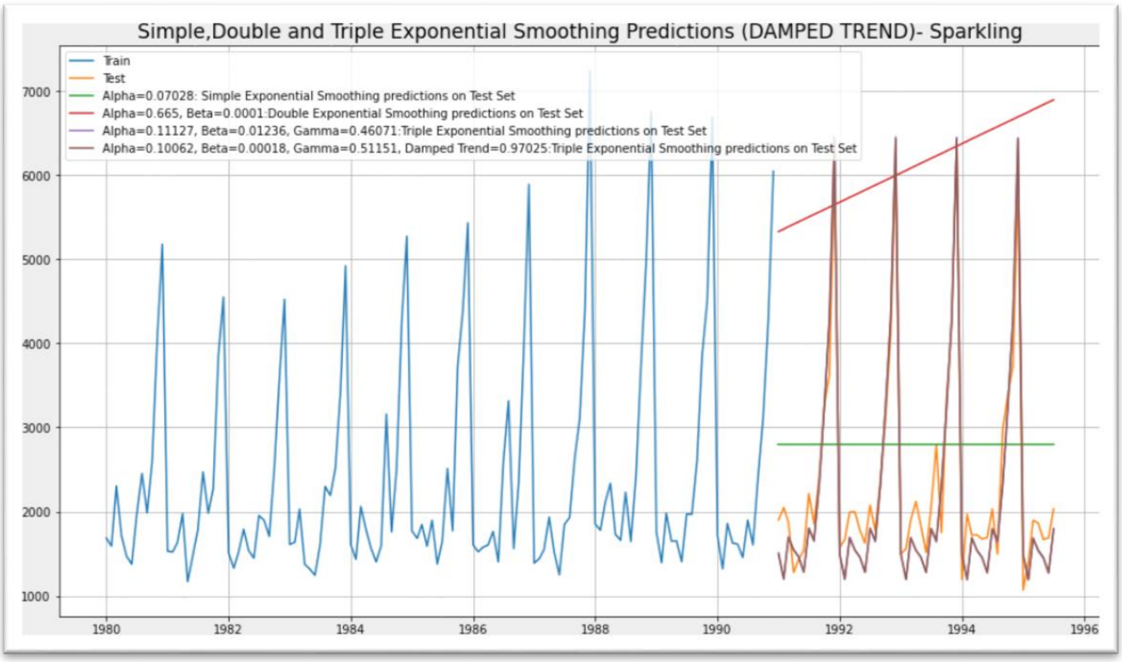
Holt-Winters - ETS(A, A, A) - Holt Winter's linear method with additive errors - SPARKLING USING DAMPED TREND

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.1057383191297317, 'smoothing_trend': 0.00014115807384965632, 'smoothing_seasonal': 0.48697109949814266,
'damping_trend': 0.9796782568408767, 'initial_level': 2361.578047768269, 'initial_trend': -1.9467315454588507, 'initial_season
s': array([-645.76716436, -730.40122176, -382.32172765, -478.92581014,
       -817.89058936, -824.7179298 , -385.27429991,    83.30682471,
       -250.13814372,  268.73180207, 1562.43394424, 2606.0934519 ]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 76: Checking for Parameters

1991-01-01	1493.356420	1993-01-01	1490.750162		
1991-02-01	1198.483188	1993-02-01	1195.929894		
1991-03-01	1690.910543	1993-03-01	1688.409137		
1991-04-01	1545.617609	1993-04-01	1543.167036		
1991-05-01	1461.436089	1993-05-01	1459.035315		
1991-06-01	1280.632383	1993-06-01	1278.280397		
1991-07-01	1802.636036	1993-07-01	1800.331847		
1991-08-01	1659.399156	1993-08-01	1657.141792		
1991-09-01	2321.017017	1993-09-01	2318.805527		
1991-10-01	3227.119547	1993-10-01	3224.952998		
1991-11-01	4301.654642	1993-11-01	4299.532121		
1991-12-01	6434.621194	1993-12-01	6432.541807		
1992-01-01	1491.893570	1994-01-01	1489.856440		
1992-02-01	1197.050066	1994-02-01	1195.054334		
1992-03-01	1689.506545	1994-03-01	1687.551369		
1992-04-01	1544.242143	1994-04-01	1542.326699		
1992-05-01	1460.088574	1994-05-01	1458.212056	1995-01-01	1489.157878
1992-06-01	1279.312252	1994-06-01	1277.473868	1995-02-01	1194.369968
1992-07-01	1801.342732	1994-07-01	1799.541707	1995-03-01	1686.880911
1992-08-01	1658.132135	1994-08-01	1656.367710	1995-04-01	1541.669866
1992-09-01	2319.775744	1994-09-01	2318.047175	1995-05-01	1457.568571
1992-10-01	3225.903498	1994-10-01	3224.210057	1995-06-01	1276.843460
1992-11-01	4300.463306	1994-11-01	4298.804278	1995-07-01	1798.924110
1992-12-01	6433.454068	1994-12-01	6431.828755	Freq: MS, dtype: float64	

Table 77: Training set for Forecast



Graph 34: Triple Smoothing using Damped Trend

Model Evaluation:

```
TES RMSE SPARK DAMPED: 377.36420514299726
```

	Test RMSE Sparkling
Triple Exponential Smoothing (Additive Season, Damped Trend)	379.695686

	Test RMSE Rose	Test RMSE Sparkling
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.66096	379.695686

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352
SimpleAverageModel	53.460570	1275.081804
2pointTrailingMovingAverage	11.529278	813.400684
4pointTrailingMovingAverage	14.451403	1156.589694
6pointTrailingMovingAverage	14.566327	1283.927428
9pointTrailingMovingAverage	14.727630	1346.278315
Simple Exponential Smoothing	36.796228	1338.012144
Double Exponential Smoothing	15.269328	3949.993290
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.510170
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.660960	379.695686

Table 78: Model Evaluation

- Rose:
  - Level parameter, Alpha = 0.07842
  - Trend parameter, Beta = 0.01153
  - Seasonality parameter, Gamma = 0.07738 Damping factor = 0.97503

- Sparkling:
  - Level parameter, Alpha = 0.10062
  - Trend parameter, Beta = 0.00018
  - Seasonality parameter, Gamma = 0.97025 Damping factor = 0.97025

Holt-Winters - ETS(A, A, M) - Holt Winter's linear method - ROSE USING DAMPED TREND

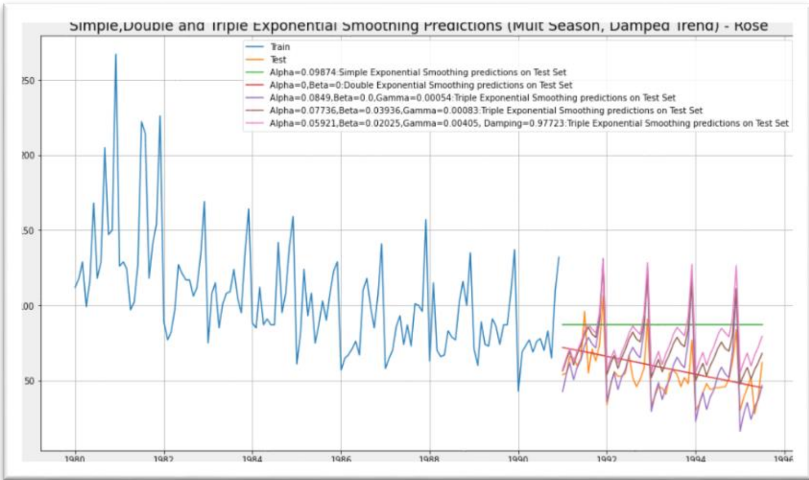
ETS(A, Ad, M) model - Taking MULTIPLICATIVE SEASONALITY

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.0002560473477695915, 'smoothing_trend': 3.3814318768645266e-07, 'smoothing_seasonal': 0.00016960815725286
792, 'damping_trend': 0.9790757197728251, 'initial_level': 165.96547270159138, 'initial_trend': -1.9265981623841522, 'initial_s
easons': array([0.70787477, 0.80549593, 0.87787457, 0.7704003 , 0.86151917,
       0.93514919, 1.0250353 , 1.08732193, 1.04699153, 1.01656203,
       1.18353599, 1.62323663]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 79: Checking for Parameters

1991-01-01	57.503594	1993-01-01	55.978300		
1991-02-01	65.347562	1993-02-01	63.648105		
1991-03-01	71.122016	1993-03-01	69.308608		
1991-04-01	62.334387	1993-04-01	60.776206		
1991-05-01	69.611882	1993-05-01	67.905953		
1991-06-01	75.457869	1993-06-01	73.645034		
1991-07-01	82.607574	1993-07-01	80.662040		
1991-08-01	87.517293	1993-08-01	85.496754		
1991-09-01	84.170879	1993-09-01	82.265949		
1991-10-01	81.630398	1993-10-01	79.819466		
1991-11-01	94.929799	1993-11-01	92.865484		
1991-12-01	130.040489	1993-12-01	127.268669		
1992-01-01	56.644700	1994-01-01	55.461252		
1992-02-01	64.390597	1994-02-01	63.072019		
1992-03-01	70.100885	1994-03-01	68.693896		
1992-04-01	61.456974	1994-04-01	60.248011		
1992-05-01	68.651272	1994-05-01	67.327673	1995-01-01	55.060085
1992-06-01	74.437061	1994-06-01	73.030515	1995-02-01	62.625045
1992-07-01	81.512042	1994-07-01	80.002538	1995-03-01	68.216951
1992-08-01	86.379526	1994-08-01	84.811828	1995-04-01	59.838193
1992-09-01	83.098212	1994-09-01	81.620212	1995-05-01	66.878997
1992-10-01	80.610661	1994-10-01	79.205592	1995-06-01	72.553722
1992-11-01	93.767382	1994-11-01	92.165718	1995-07-01	79.490844
1992-12-01	128.479675	1994-12-01	126.329071	Freq: MS, dtype: float64	

Table 80: Training set for Forecast



Graph 35: Triple Smoothing using Damped Trend with Multiplicative Seasonality

Report model accuracy:

```
TES_am RMSE_rose DAMPED: 26.29598142802485
```

	Test RMSE Rose
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	26.295981

Table 81: Model Evaluation

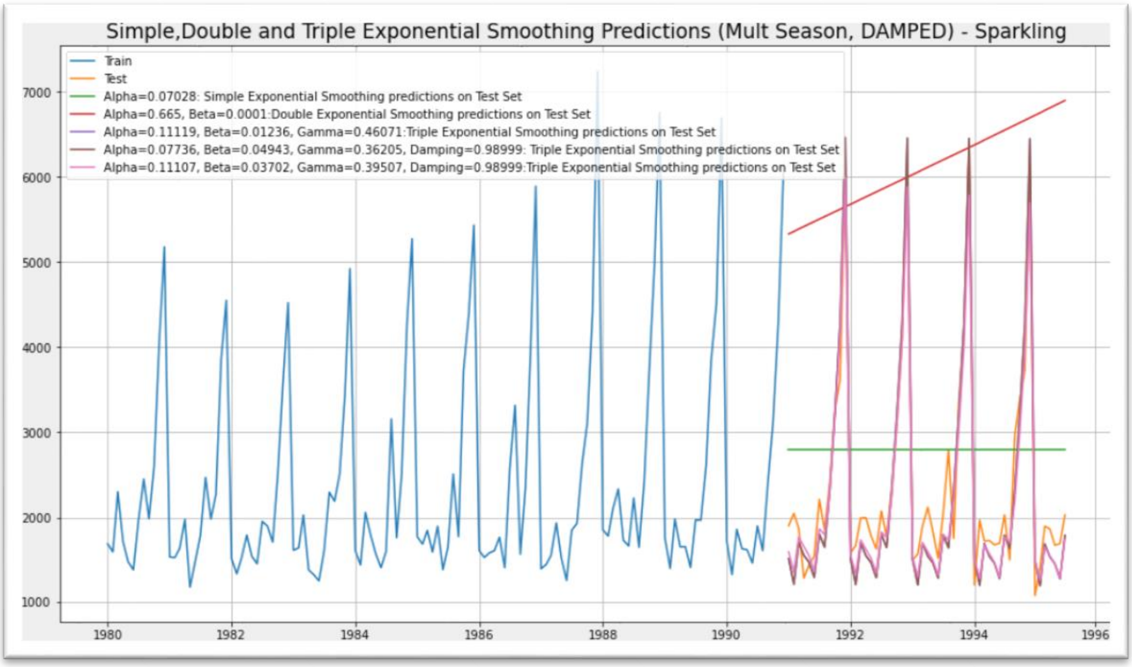
Holt-Winters – ETS (A, A, M) - Holt Winter's linear method - SPARKLING ETS (A, A, M) model - Taking MULTIPLICATIVE SEASONALITY USING DAMPED TREND

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.1110714622899222, 'smoothing_trend': 0.03702380844127912, 'smoothing_seasonal': 0.39507957727536136, 'dam
ping_trend': 0.9899999825826437, 'initial_level': 2356.5418308172734, 'initial_trend': -9.179892630347588, 'initial_seasons': a
rray([0.713876  , 0.68479146, 0.89985055, 0.80522628, 0.65413878,
       0.65498002, 0.88128754, 1.12310179, 0.91373324, 1.1919948 ,
       1.848147  , 2.33628145]), 'use boxcox': False, 'lamda': None, 'remove bias': False}
```

Table 82: Checking for Parameters

```
1991-01-01    1591.418267
1991-02-01    1348.088290
1991-03-01    1768.339414
1991-04-01    1654.050476
1991-05-01    1548.114090
1991-06-01    1363.572304
1991-07-01    1862.877257
1991-08-01    1797.372758
1991-09-01    2310.533752
1991-10-01    3163.683142
1991-11-01    4176.962819
1991-12-01    5994.207394
1992-01-01    1558.916323
1992-02-01    1320.782175
1992-03-01    1732.815783
1992-04-01    1621.096771
1992-05-01    1517.525754
1992-06-01    1336.853182
1992-07-01    1826.676879
1992-08-01    1762.735329
1992-09-01    2266.377494
1992-10-01    3103.726232
1992-11-01    4098.463283
1992-12-01    5882.497302
1993-01-01    1530.107096
1993-02-01    1296.578453
1993-03-01    1701.328181
1993-04-01    1591.887113
1993-05-01    1490.412721
1993-06-01    1313.169762
1993-07-01    1794.589419
1993-08-01    1732.033242
1993-09-01    2227.238063
1993-10-01    3050.581346
1993-11-01    4028.882497
1993-12-01    5783.479188
1994-01-01    1504.571040
1994-02-01    1275.124644
1994-03-01    1673.418052
1994-04-01    1565.996119
1994-05-01    1466.380144
1994-06-01    1292.177141
1994-07-01    1766.147586
1994-08-01    1704.819383
1994-09-01    2192.545471
1994-10-01    3003.474532
1994-11-01    3967.207154
1994-12-01    5695.711049
1995-01-01    1481.936270
1995-02-01    1256.108317
1995-03-01    1648.678942
1995-04-01    1543.046738
1995-05-01    1445.078035
1995-06-01    1273.569603
1995-07-01    1740.937180
Freq: MS, dtype: float64
```

Table 83: Training set for Forecast



Graph 36: Triple Smoothing using Damped Trend with Multiplicative Seasonality

Report model accuracy:

```
TES_am DAMPED RMSE_spark: 352.4433346626412
```

	Test RMSE Sparkling
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	352.443335

	Test RMSE Rose	Test RMSE Sparkling
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	26.295981	352.443335

	Test RMSE Rose	Test RMSE Sparkling
RegressionOnTime	15.268955	1389.135175
NaiveModel	79.718773	3864.279352
SimpleAverageModel	53.460570	1275.081804
2pointTrailingMovingAverage	11.529278	813.400684
4pointTrailingMovingAverage	14.451403	1156.589694
6pointTrailingMovingAverage	14.566327	1283.927428
9pointTrailingMovingAverage	14.727630	1346.278315
Simple Exponential Smoothing	36.796228	1338.012144
Double Exponential Smoothing	15.269328	3949.993290
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.510170
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.660960	379.695686
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	26.295981	352.443335

Table 84: Model Evaluation

- Rose:
  - Level parameter, Alpha = 0.05921
  - Trend parameter, Beta = 0.02025
  - Seasonality parameter, Gamma = 0.00405
  - Damping factor = 0.97723

- Sparkling:
  - Level parameter, Alpha = 0.11107
  - Trend parameter, Beta = 0.03702
  - Seasonality parameter, Gamma = 0.39507
  - Damping factor = 0.98999

Best Model for Rose till Now - 2 Pt Moving Average:

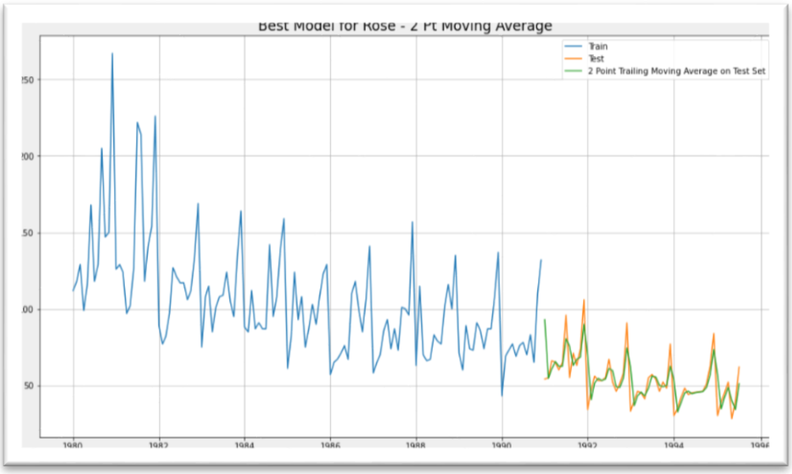
	Rose
YearMonth	
1980-01-01	112.0
1980-02-01	118.0
1980-03-01	129.0
1980-04-01	99.0
1980-05-01	116.0

Table 85: Data Set

Trailing moving:

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

Table 86: Trailing Moving Values



Graph 37: 2 Point Moving Average

- We conclude that models with least RMSE:
  Best Model for Rose ——> 2 Pt Moving Average
  Best Model for Sparkling ——> Holt-Winter Damped Trend ETS (A, Ad, M)

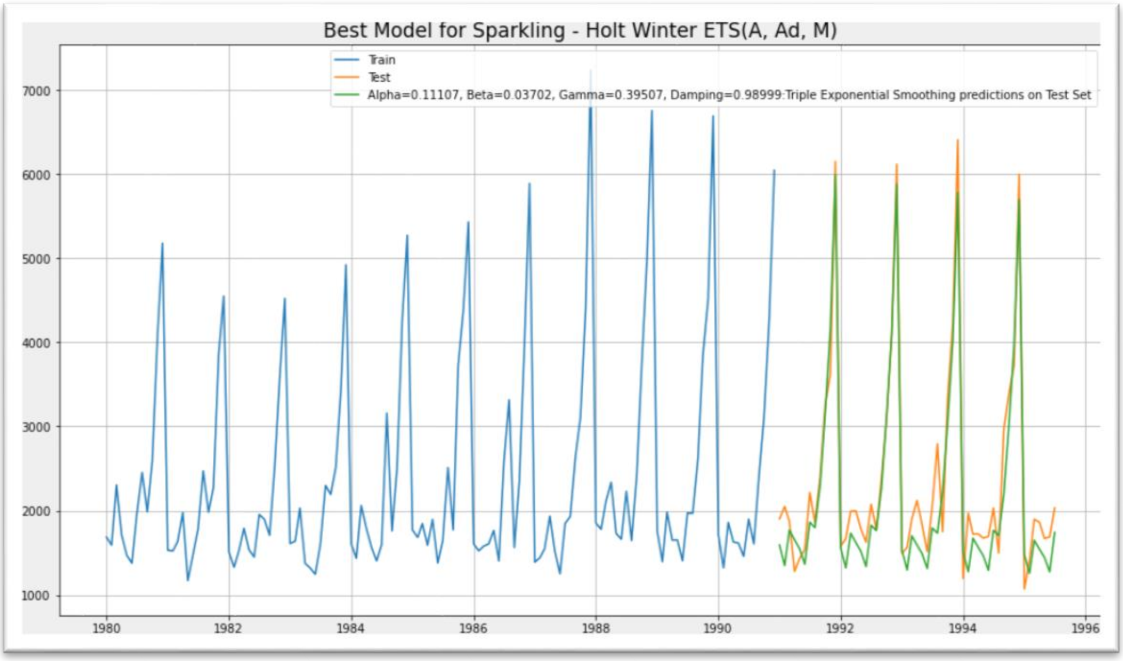
Holt-Winters - ETS(A, A, M) - Best Model for Sparkling Till Now ETS(A, Ad, M) model - Taking MULTIPLICATIVE SEASONALITY USING DAMPED TREND

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.1110714622899222, 'smoothing_trend': 0.03702380844127912, 'smoothing_seasonal': 0.395079577275361 36, 'dam
ping_trend': 0.9899999825826437, 'initial_level': 2356.5418308172734, 'initial_trend': -9.179892630347588, 'initial_seasons': a
rray([0.713876  , 0.68479146, 0.89985055, 0.80522628, 0.65413878,
       0.65498002, 0.88128754, 1.12310179, 0.91373324, 1.1919948 ,
       1.848147  , 2.33628145]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 87: Checking for Parameters

1991-01-01	1591.418267	1992-01-01	1558.916323		
1991-02-01	1348.088290	1992-02-01	1320.782175		
1991-03-01	1768.339414	1992-03-01	1732.815783		
1991-04-01	1654.050476	1992-04-01	1621.096771	1994-01-01	1504.571040
1991-05-01	1548.114090	1992-05-01	1517.525754	1994-02-01	1275.124644
1991-06-01	1363.572304	1992-06-01	1336.853182	1994-03-01	1673.418052
1991-07-01	1862.877257	1992-07-01	1826.676879	1994-04-01	1565.996119
1991-08-01	1797.372758	1992-08-01	1762.735329	1994-05-01	1466.380144
1991-09-01	2310.533752	1992-09-01	2266.377494	1994-06-01	1292.177141
1991-10-01	3163.683142	1992-10-01	3103.726232	1994-07-01	1766.147586
1991-11-01	4176.962819	1992-11-01	4098.463283	1994-08-01	1704.819383
1991-12-01	5994.207394	1992-12-01	5882.497302	1994-09-01	2192.545471
1992-01-01	1558.916323	1993-01-01	1530.107096	1994-10-01	3003.474532
1992-02-01	1320.782175	1993-02-01	1296.578453	1994-11-01	3967.207154
1992-03-01	1732.815783	1993-03-01	1701.328181	1994-12-01	5695.711049
1992-04-01	1621.096771	1993-04-01	1591.887113	1995-01-01	1481.936270
1992-05-01	1517.525754	1993-05-01	1490.412721	1995-02-01	1256.108317
1992-06-01	1336.853182	1993-06-01	1313.169762	1995-03-01	1648.678942
1992-07-01	1826.676879	1993-07-01	1794.589419	1995-04-01	1543.046738
1992-08-01	1762.735329	1993-08-01	1732.033242	1995-05-01	1445.078035
1992-09-01	2266.377494	1993-09-01	2227.238063	1995-06-01	1273.569603
1992-10-01	3103.726232	1993-10-01	3050.581346	1995-07-01	1740.937180
1992-11-01	4098.463283	1993-11-01	4028.882497	Freq: MS, dtype: float64	
1992-12-01	5882.497302	1993-12-01	5783.479188		

Table 88: Training set for Forecast



Graph 38: Best Model for Sparkling

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.

To Check Stationarity of Data:

- We use Augmented Dicky - Fuller (ADF) Test to check the Stationarity of Data
- Hypotheses of ADF Test:

  H(O): Time Series is not Stationary
  H(A): Time Series is Stationary

- So, for Industry standard (also given for this problem), the Confidence Interval is 95%.
- Hence, alpha = 0.05
- So, in ADF Test, if p-value < alpha ===> We reject the Null Hypothesis and hence conclude that given Time Series is Stationary
- So, in ADF Test, if p-value > alpha ===> We fail to reject the Null Hypothesis and hence conclude that given Time Series is Not Stationary
- If Time Series is not Stationary then we apply one level of differencing and check for Stationarity again.
- Again, if the Time Series is still not Stationary, we apply one more level of differencing and check for Stationarity again
- Generally, with max 2 levels of differencing, Time Series becomes Stationary
- Once the Time Series is Stationary then we are ready to apply ARIMA / SARIMA models

Stationarity of Rose Wine Dataset:

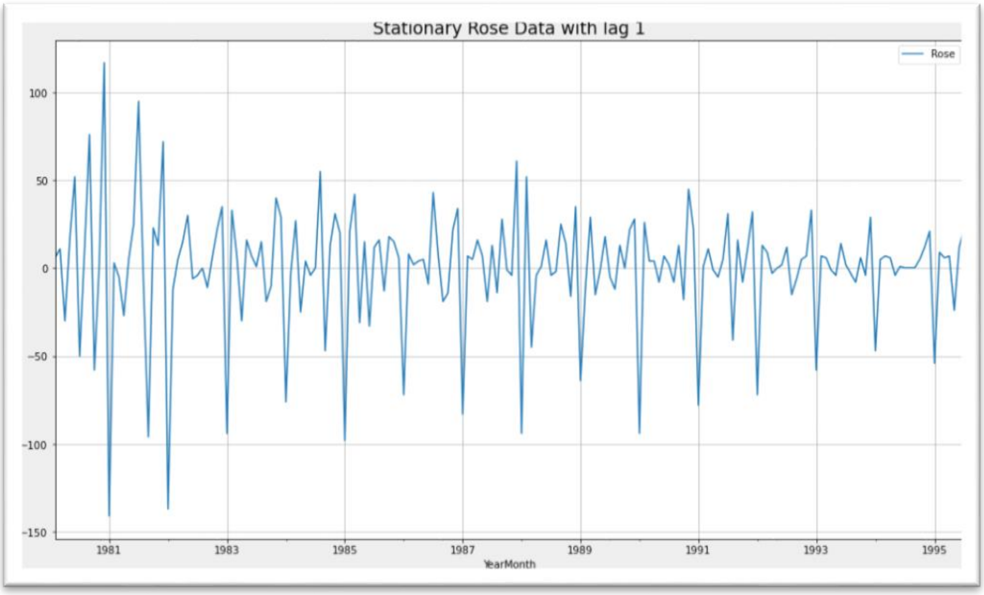
- Augmented Dicky- Fuller Test was applied to the whole Rose dataset
- We found, p-value = 0.4671
- Here, p-value > alpha=0.05
- We fail to reject the Null Hypothesis and hence conclude that Rose Wine Time Series is Not Stationary
- We take 1 level of differencing and check again for Stationarity
- Now, p-value = 3.0159e-11 0.00
- Now, p-value < alpha=0.05
- Now, we reject the Null Hypothesis and conclude that Rose Time Series is Stationary with a lag of 1

Stationarity for Rose Data:

```
DF test statistic is -2.240
DF test p-value is 0.4671371627793208
Number of lags used 13
```

```
DF test statistic is -8.162
DF test p-value is 3.015976115826596e-11
Number of lags used 12
```

Table 89: Stats for Checking Stationarity



Graph 39: Stationarity of Rose Data

Stationarity of Sparkling Wine Dataset:

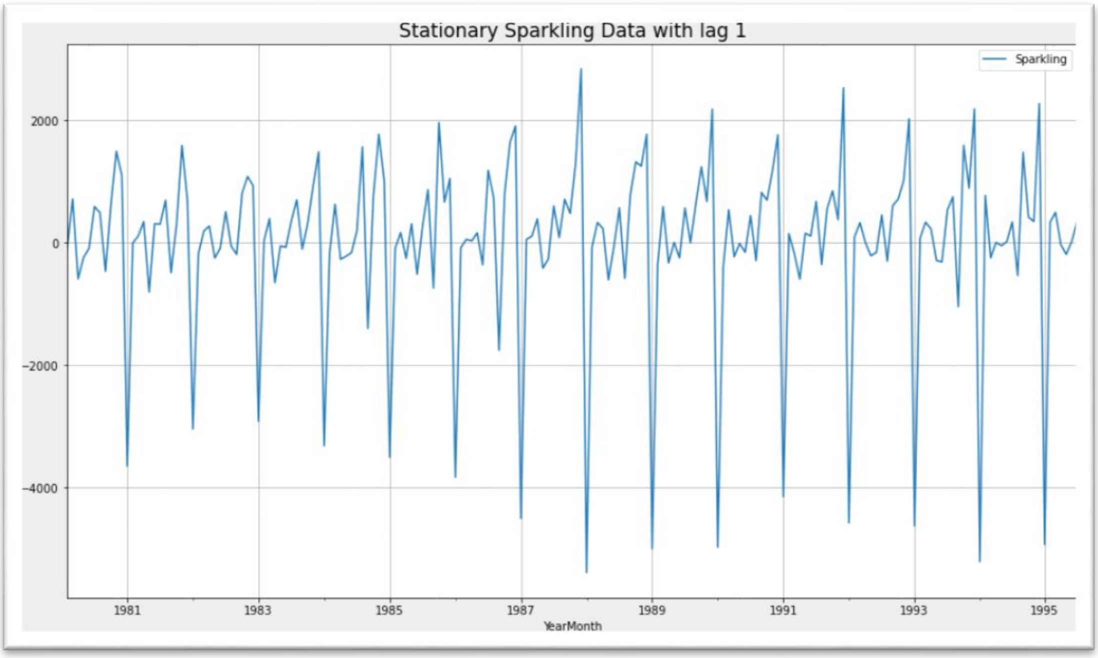
- Augmented Dicky-Fuller Test was applied to the whole Sparkling dataset • We found, p-value = 0.70559
- Here, p-value > alpha=0.05
- We fail to reject the Null Hypothesis and hence conclude that Sparkling Wine Time Series is Not Stationary
- We take 1 level of differencing and check again for Stationarity
- Now, p-value = 0.00
- Now, p-value < alpha=0.05
- Now, we reject the Null Hypothesis and conclude that Sparkling Time Series is Stationary with a lag of 1

Stationarity for Sparkling Data:

```
DF test statistic is -1.798
DF test p-value is 0.7055958459932035
Number of lags used 12
```

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

Table 90: Stats for Checking Stationarity



Graph 40: Stationarity of Sparkling Data

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.

ARIMA / SARIMA Models:

- ARIMA is an acronym for Auto-Regressive Integrated Moving Average
- SARIMA stands for Seasonal ARIMA, when the TS has seasonality
- ARIMA / SARIMA are forecasting models on Stationary Time Series
- ARIMA / SARIMA Modelling on Train Rose & Sparkling Data:
- We check for stationarity of Train Rose & Sparkling data by using Augmented Dicky Fuller Test
- We take a difference of 1 and make both these datasets Stationary
- We apply the following iterations to both these datasets:

1. ARIMA Automated
2. SARIMA Automated

1) ARIMA Automated:

- We create a grid of all possible combinations of (p, d, q)
- Range of p = Range of q = 0 to 3, Constant d = 1
- Few Examples of the grid –

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)

- We fit ARIMA models to each of these combinations for both datasets
- We choose the combination with the least Akaike Information Criteria (AIC)
- We fit ARIMA to this combination of (p, d, q) to the Train set and forecast on the Test set
- Finally, we check the accuracy of this model by checking RMSE of Test set
- For Rose, Best Combination with Least AIC is - (p, d, q) —-> (2, 1, 3)
- For Sparkling, Best Combination with Least AIC is - (p, d, q) —-> (2, 1, 2)

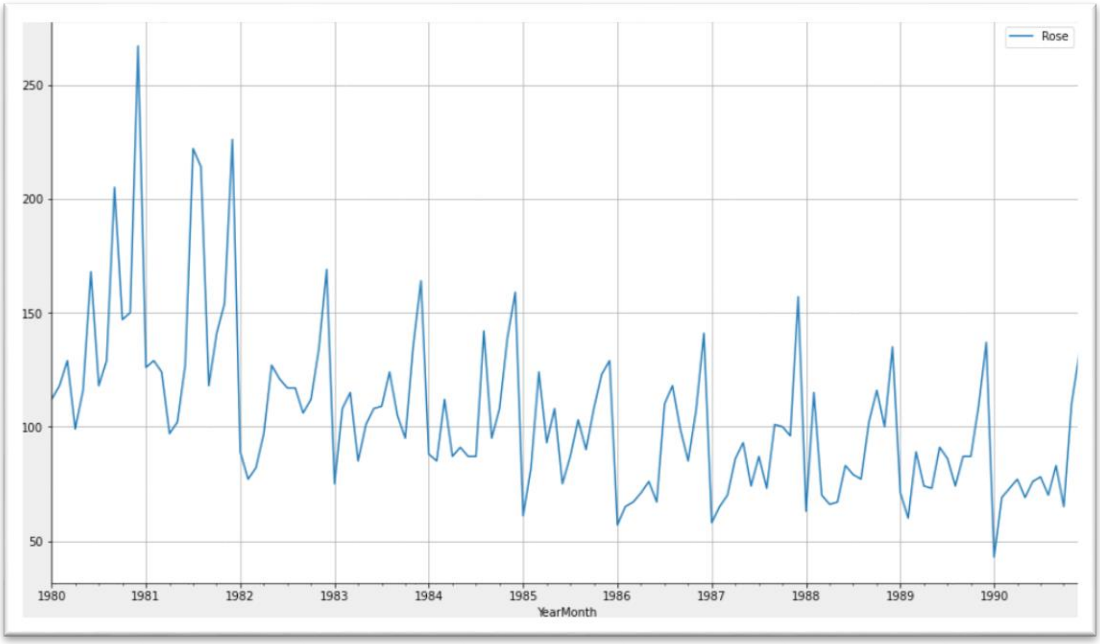
2) SARIMA Automated:

- We create a grid of all possible combinations of (p, d, q) along with Seasonal (P, D, Q) & Seasonality of 12 (for both datasets)
- Range of p = Range of q = 0 to 3, Constant d = 1
- Range of Seasonal P = Range of Seasonal Q = 0 to 3, Constant D = 1, Seasonality m = 12
- Few Examples of the grid (p, d, q) (P, D, Q, m):

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)

- We fit SARIMA models to each of these combinations and select with least AIC
- We fit SARIMA to this best combination of (p, d, q) (P, D, Q, m) to the Train set and forecast on the Test set. Then, we check accuracy using RMSE on Test set
- For Rose, Best Combination with Least AIC is - (3, 1, 1) (3, 0, 2, 12)
- For Sparkling, Best Combination with low AIC and low Test RMSE is - (3, 1, 1) (3, 0, 0, 12)

ARIMA / SARIMA of Rose Train Data:

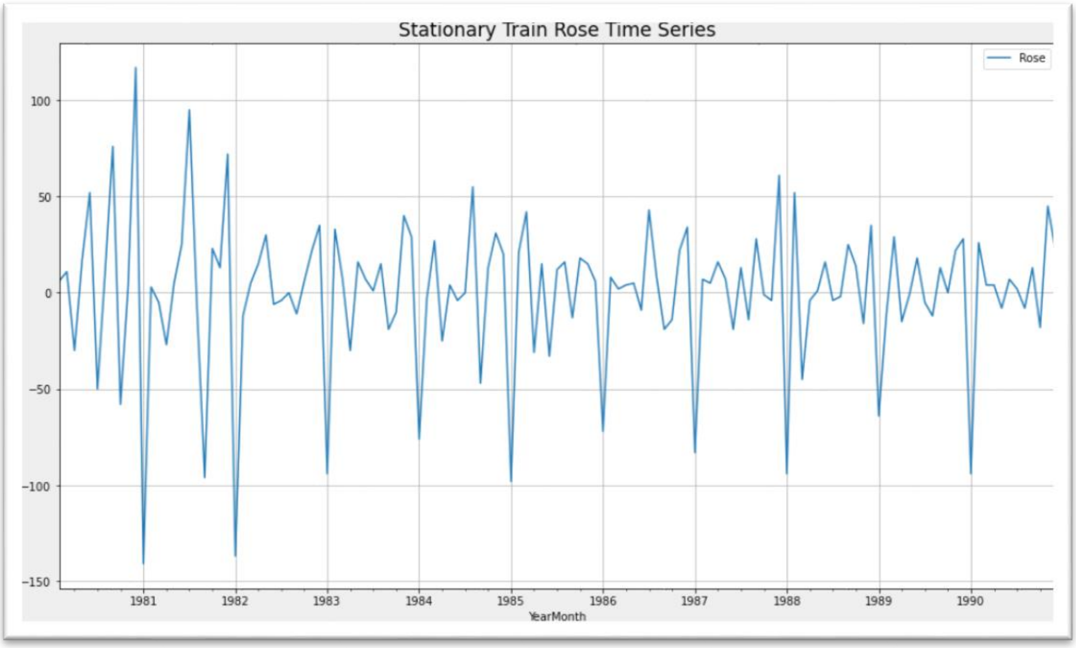


Graph 41: ARIMA / SARIMA of Rose data

```
DF test statistic is -1.686
DF test p-value is 0.7569093051047057
Number of lags used 13
```

```
DF test statistic is -6.804
DF test p-value is 3.894831356782385e-08
Number of lags used 12
```

Table 91: Stats for Checking Stationarity



Graph 42: Stationarity of Rose data

Build an Automated version of an ARIMA model for which the best parameters are selected in accordance with the lowest Akaike Information Criteria (AIC):

```
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)
```

Table 92: Parameter Combination of Model

```
ARIMA(0, 1, 0) - AIC:1333.1546729124348
ARIMA(0, 1, 1) - AIC:1282.3098319748315
ARIMA(0, 1, 2) - AIC:1279.6715288535752
ARIMA(0, 1, 3) - AIC:1280.5453761734652
ARIMA(1, 1, 0) - AIC:1317.3503105381526
ARIMA(1, 1, 1) - AIC:1280.5742295380076
ARIMA(1, 1, 2) - AIC:1279.8707234231915
ARIMA(1, 1, 3) - AIC:1281.870722330997
ARIMA(2, 1, 0) - AIC:1298.6110341604908
ARIMA(2, 1, 1) - AIC:1281.507862186858
ARIMA(2, 1, 2) - AIC:1281.8707222264168
```

Table 93: AIC Values

```
ARIMA(2, 1, 3) - AIC:1274.6953190416875
ARIMA(3, 1, 0) - AIC:1297.4810917271702
ARIMA(3, 1, 1) - AIC:1282.419277627203
ARIMA(3, 1, 2) - AIC:1283.720740597714
ARIMA(3, 1, 3) - AIC:1278.6543993387522
```

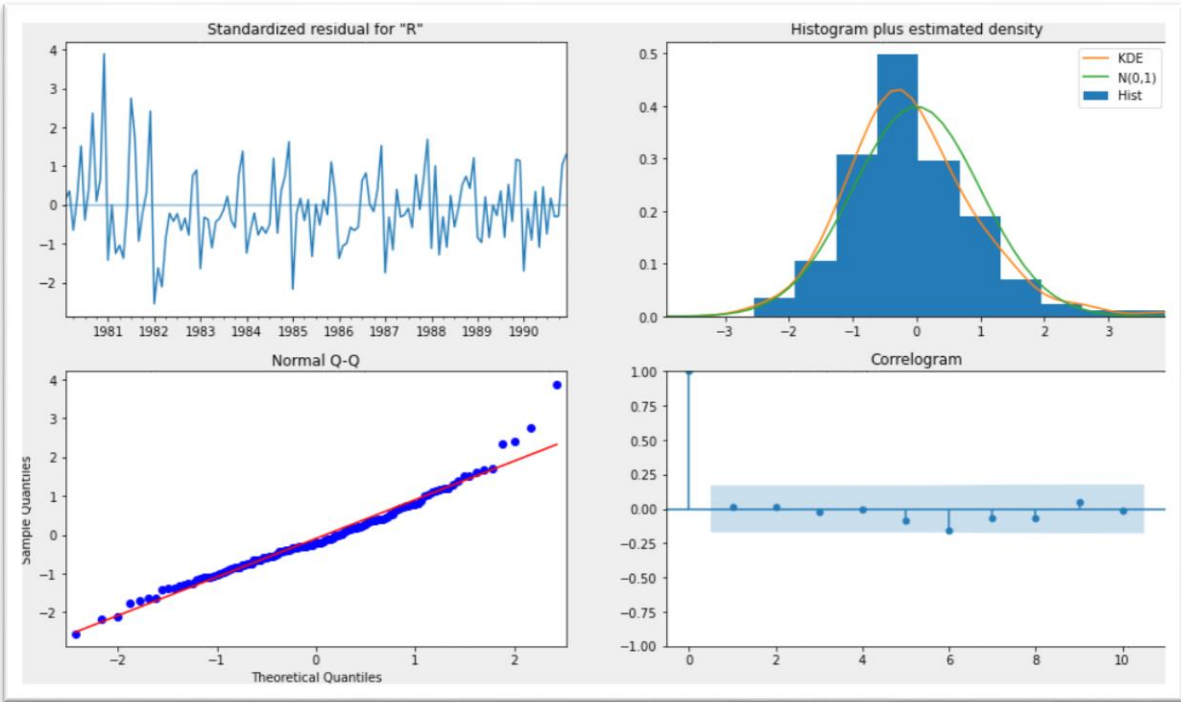
	param	AIC
11	(2, 1, 3)	1274.695319
15	(3, 1, 3)	1278.654399
2	(0, 1, 2)	1279.671529
6	(1, 1, 2)	1279.870723
3	(0, 1, 3)	1280.545376

Table 94: Parameter for Minimum AIC in a sorted way

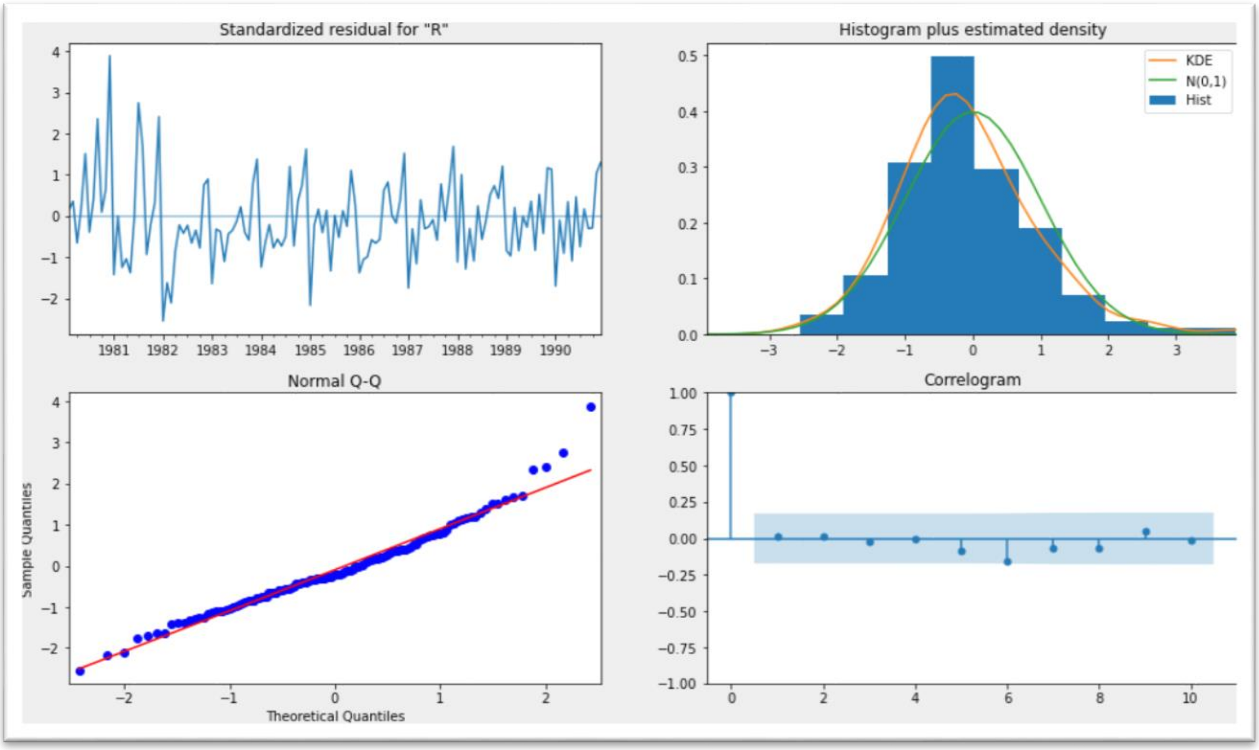
```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   Rose   No. Observations:                  132
Model:                 ARIMA(2, 1, 3)   Log Likelihood                -631.348
Date:                Fri, 25 Mar 2022   AIC                           1274.695
Time:                        12:24:18   BIC                           1291.947
Sample:                    01-01-1980   HQIC                          1281.705
                         - 12-01-1990
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -1.6780      0.084    -20.029      0.000      -1.842      -1.514
ar.L2         -0.7287      0.084     -8.697      0.000      -0.893      -0.565
ma.L1          1.0447      0.616      1.695      0.090      -0.163       2.253
ma.L2         -0.7716      0.132     -5.856      0.000      -1.030      -0.513
ma.L3         -0.9044      0.558     -1.620      0.105      -1.999       0.190
sigma2       858.9120    517.873      1.659      0.097    -156.100    1873.924
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):                24.43
Prob(Q):                              0.88   Prob(JB):                         0.00
Heteroskedasticity (H):               0.40   Skew:                             0.71
Prob(H) (two-sided):                  0.00   Kurtosis:                         4.57
===================================================================================

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

Table 92: Model Summary

Rose Train Diagnostics plot:



Graph 43: Diagnostic Plot



Graph 44: Diagnostic Plot

Predict on the Rose Test Set using this model and evaluate the model:

1991-01-01	85.592932
1991-02-01	90.547211
1991-03-01	81.963691
1991-04-01	92.756336
1991-05-01	80.901574
1991-06-01	92.928681
1991-07-01	81.386383
1991-08-01	91.989584
1991-09-01	82.608874
1991-10-01	90.622618
1991-11-01	84.011748
1991-12-01	89.264777
1992-01-01	85.267859
1992-02-01	88.146550
1992-03-01	86.228852
1992-04-01	87.348913
1992-05-01	86.866964
1992-06-01	86.859438
1992-07-01	87.223281
1992-08-01	86.618243
1992-09-01	87.368341
1992-10-01	86.550602
1992-11-01	87.376131
1992-12-01	86.586822
1993-01-01	87.309677
1993-02-01	86.671936
1993-03-01	87.215284

1993-04-01	86.768300
1993-05-01	87.122375
1993-06-01	86.853976
1993-07-01	87.046318
1993-08-01	86.919164
1993-09-01	86.992360
1993-10-01	86.962199
1993-11-01	86.959467
1993-12-01	86.986031
1994-01-01	86.943449
1994-02-01	86.995543
1994-03-01	86.939161
1994-04-01	86.995806
1994-05-01	86.941844
1994-06-01	86.991112
1994-07-01	86.947765
1994-08-01	86.984597
1994-09-01	86.954382
1994-10-01	86.978242
1994-11-01	86.960224
1994-12-01	86.973070
1995-01-01	86.964645
1995-02-01	86.969421
1995-03-01	86.967547
1995-04-01	86.967211
1995-05-01	86.969140
1995-06-01	86.966148
1995-07-01	86.969763

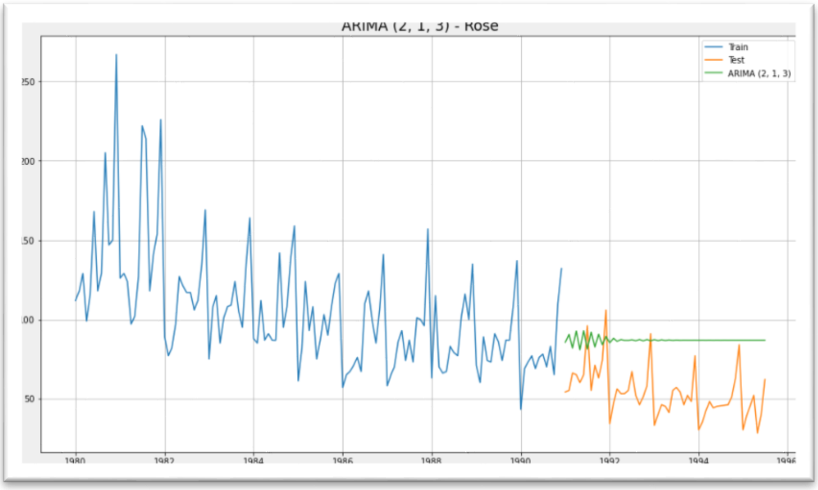
Table 96: Prediction on Rose Test Set

RSME and MAPE Values:

RMSE: 36.81634749426474
MAPE: 75.84583635258521

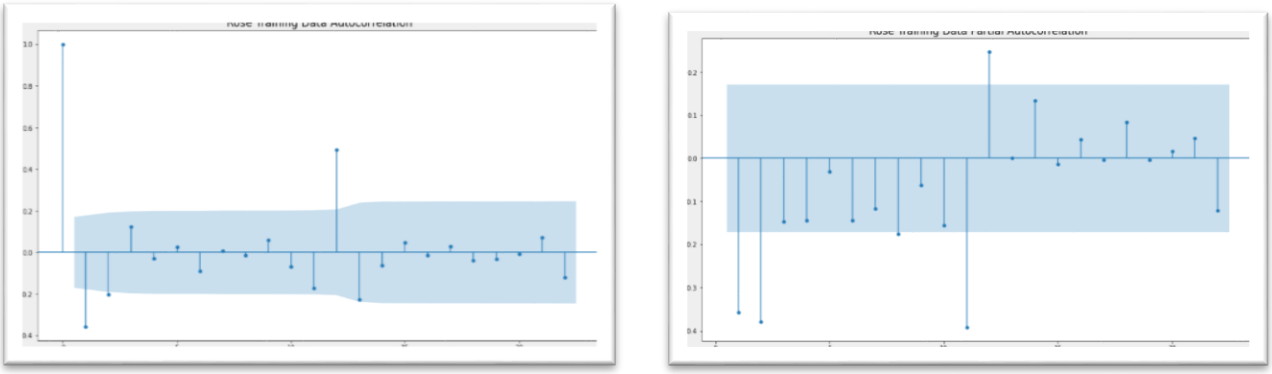
	Test RMSE Rose	Test MAPE Rose
ARIMA(2,1,3)	36.816347	75.845836

Table 94: RSME and MAPE Values



Graph 45: ARIMA Plot

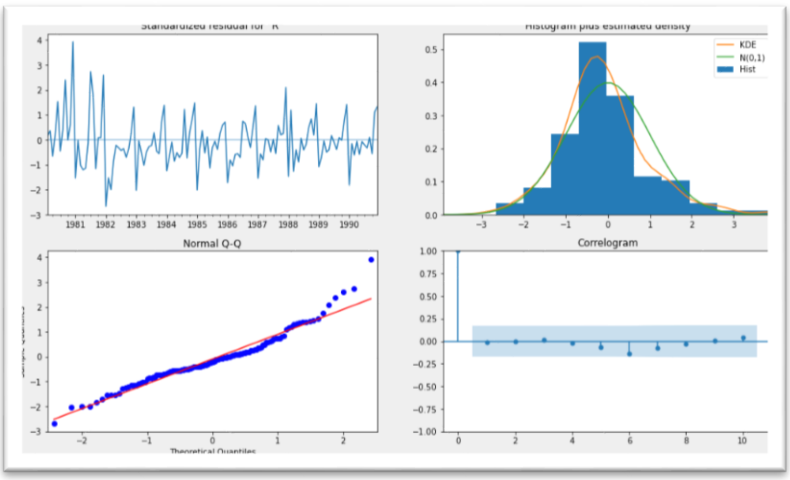
Build a version of the ARIMA model for which the best parameters are selected by looking at the ACF and the PACF plots on ROSE dataset:



Graph 46: Data Autocorrelation

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   Rose   No. Observations:                  132
Model:                 ARIMA(2, 1, 2)   Log Likelihood                -635.935
Date:                Fri, 25 Mar 2022   AIC                           1281.871
Time:                        12:24:23   BIC                           1296.247
Sample:                    01-01-1980   HQIC                          1287.712
                         - 12-01-1990
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.4540      0.469     -0.969      0.333      -1.372       0.464
ar.L2          0.0001      0.170      0.001      0.999      -0.334       0.334
ma.L1         -0.2541      0.459     -0.554      0.580      -1.154       0.646
ma.L2         -0.5984      0.430     -1.390      0.164      -1.442       0.245
sigma2       952.1601     91.424     10.415      0.000     772.973    1131.347
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):                34.16
Prob(Q):                              0.88   Prob(JB):                         0.00
Heteroskedasticity (H):               0.37   Skew:                             0.79
Prob(H) (two-sided):                  0.00   Kurtosis:                         4.94
===================================================================================

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

Table 95: Model Summary



Graph 47: Diagnostic Plot

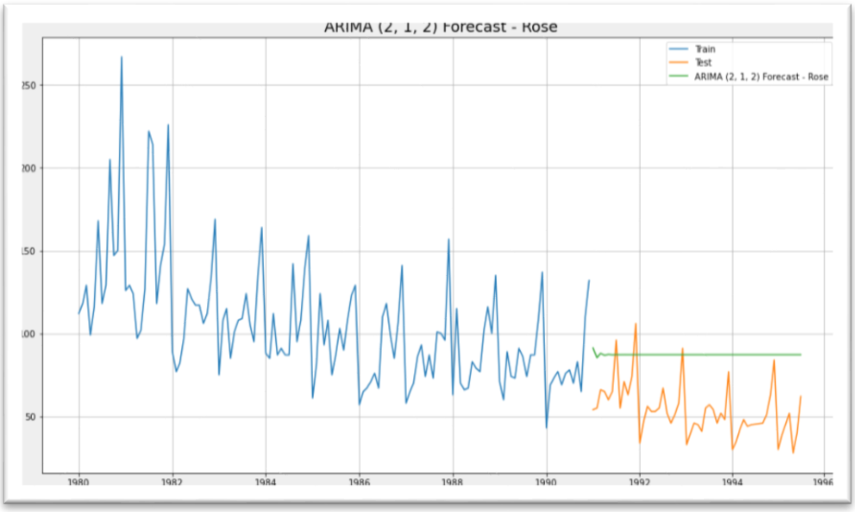
Predict on the Test Set using this model and evaluate the model:

```
RMSE: 36.8711966230642
MAPE: 76.05621272520435
```

	Test RMSE Rose	Test MAPE Rose
ARIMA(2,1,3)	36.816347	75.845836
ARIMA(2,1,2)	36.871197	76.056213

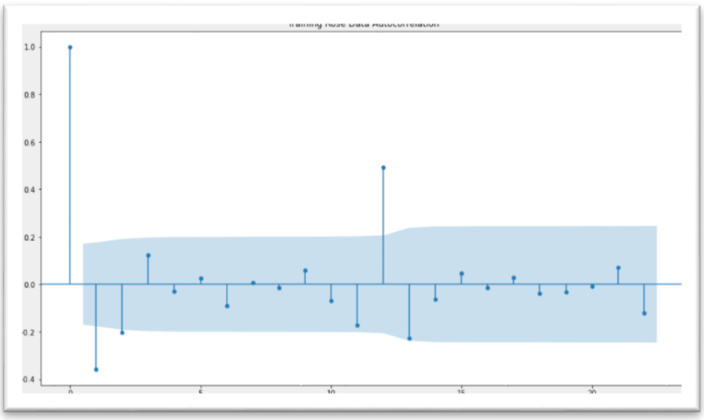
	Test RMSE Rose	Test MAPE Rose
ARIMA(2,1,3)	36.816347	75.845836
ARIMA(2,1,2)	36.871197	76.056213

Table 99: Model Evaluation



Graph 48: ARIMA Forecast

Build an Automated version of a SARIMA model for which the best parameters are selected in accordance with the lowest Akaike Information Criteria (AIC) - ROSE DATA:



Graph 49: Data Autocorrelation

```
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)
```

Table 100: Parameter Combination

```
SARIMA(0, 1, 0)x(0, 0, 0, 12) - AIC:1323.9657875279158
SARIMA(0, 1, 0)x(0, 0, 1, 12) - AIC:1145.4230827207525
SARIMA(0, 1, 0)x(0, 0, 2, 12) - AIC:976.4375296380895
SARIMA(0, 1, 0)x(0, 0, 3, 12) - AIC:3089.9353646004197
SARIMA(0, 1, 0)x(1, 0, 0, 12) - AIC:1139.921738995602
SARIMA(0, 1, 0)x(1, 0, 1, 12) - AIC:1116.0207869386875
SARIMA(0, 1, 0)x(1, 0, 2, 12) - AIC:969.6913635752474
SARIMA(0, 1, 0)x(1, 0, 3, 12) - AIC:3743.561624129019
SARIMA(0, 1, 0)x(2, 0, 0, 12) - AIC:960.8812220353041
SARIMA(0, 1, 0)x(2, 0, 1, 12) - AIC:962.8794540697533
SARIMA(0, 1, 0)x(2, 0, 2, 12) - AIC:955.5735408945699
SARIMA(0, 1, 0)x(2, 0, 3, 12) - AIC:5047.932298726474
SARIMA(0, 1, 0)x(3, 0, 0, 12) - AIC:850.7535403931095
SARIMA(0, 1, 0)x(3, 0, 1, 12) - AIC:851.7482702789318
SARIMA(0, 1, 0)x(3, 0, 2, 12) - AIC:850.5304136127079
SARIMA(0, 1, 0)x(3, 0, 3, 12) - AIC:3816.5968319327458
SARIMA(0, 1, 1)x(0, 0, 0, 12) - AIC:1263.5369097383968
SARIMA(0, 1, 1)x(0, 0, 1, 12) - AIC:1098.5554825918334
SARIMA(0, 1, 1)x(0, 0, 2, 12) - AIC:923.6314049383752
SARIMA(0, 1, 1)x(0, 0, 3, 12) - AIC:3862.5806548645746
SARIMA(0, 1, 1)x(1, 0, 0, 12) - AIC:1095.793632491829
SARIMA(0, 1, 1)x(1, 0, 1, 12) - AIC:1054.7434330947895
SARIMA(0, 1, 1)x(1, 0, 2, 12) - AIC:918.8573483315454
SARIMA(0, 1, 1)x(1, 0, 3, 12) - AIC:3700.5078389923465
SARIMA(0, 1, 1)x(2, 0, 0, 12) - AIC:914.5982866535874
SARIMA(0, 1, 1)x(2, 0, 1, 12) - AIC:915.3332430461671
SARIMA(0, 1, 1)x(2, 0, 2, 12) - AIC:901.1988261883185
SARIMA(0, 1, 1)x(2, 0, 3, 12) - AIC:5356.640813592264
SARIMA(0, 1, 1)x(3, 0, 0, 12) - AIC:798.5889764821379
SARIMA(0, 1, 1)x(3, 0, 1, 12) - AIC:800.4844931681861
SARIMA(0, 1, 1)x(3, 0, 2, 12) - AIC:801.0595269476901
```

```
SARIMA(0, 1, 1)x(3, 0, 3, 12) - AIC:3674.5259016940986
SARIMA(0, 1, 2)x(0, 0, 0, 12) - AIC:1251.667543054105
SARIMA(0, 1, 2)x(0, 0, 1, 12) - AIC:1083.4866975264968
SARIMA(0, 1, 2)x(0, 0, 2, 12) - AIC:913.4938486617705
SARIMA(0, 1, 2)x(0, 0, 3, 12) - AIC:2819.5396142293735
SARIMA(0, 1, 2)x(1, 0, 0, 12) - AIC:1088.8332843413646
SARIMA(0, 1, 2)x(1, 0, 1, 12) - AIC:1045.540093365143
SARIMA(0, 1, 2)x(1, 0, 2, 12) - AIC:904.8310913484157
SARIMA(0, 1, 2)x(1, 0, 3, 12) - AIC:3107.35193983523
SARIMA(0, 1, 2)x(2, 0, 0, 12) - AIC:913.0105912257973
SARIMA(0, 1, 2)x(2, 0, 1, 12) - AIC:914.1707545029767
SARIMA(0, 1, 2)x(2, 0, 2, 12) - AIC:887.937508568104
SARIMA(0, 1, 2)x(2, 0, 3, 12) - AIC:5214.380245241165
SARIMA(0, 1, 2)x(3, 0, 0, 12) - AIC:800.1020189374518
SARIMA(0, 1, 2)x(3, 0, 1, 12) - AIC:801.9941080532727
SARIMA(0, 1, 2)x(3, 0, 2, 12) - AIC:802.5206528372927
SARIMA(0, 1, 2)x(3, 0, 3, 12) - AIC:2235.02251887393
SARIMA(0, 1, 3)x(0, 0, 0, 12) - AIC:1243.950121673909
SARIMA(0, 1, 3)x(0, 0, 1, 12) - AIC:1076.7632646089203
SARIMA(0, 1, 3)x(0, 0, 2, 12) - AIC:905.692653338092
SARIMA(0, 1, 3)x(0, 0, 3, 12) - AIC:3668.9702016428
SARIMA(0, 1, 3)x(1, 0, 0, 12) - AIC:1089.2051361212261
SARIMA(0, 1, 3)x(1, 0, 1, 12) - AIC:1034.8411212936885
SARIMA(0, 1, 3)x(1, 0, 2, 12) - AIC:896.8362523341466
SARIMA(0, 1, 3)x(1, 0, 3, 12) - AIC:3739.6836930694526
SARIMA(0, 1, 3)x(2, 0, 0, 12) - AIC:914.9477108987188
SARIMA(0, 1, 3)x(2, 0, 1, 12) - AIC:916.0630996888726
SARIMA(0, 1, 3)x(2, 0, 2, 12) - AIC:880.5509788251176
SARIMA(0, 1, 3)x(2, 0, 3, 12) - AIC:4920.088773452565
SARIMA(0, 1, 3)x(3, 0, 0, 12) - AIC:802.0987124746902
SARIMA(0, 1, 3)x(3, 0, 1, 12) - AIC:803.9928772340271
SARIMA(0, 1, 3)x(3, 0, 2, 12) - AIC:804.5150720444066
SARIMA(0, 1, 3)x(3, 0, 3, 12) - AIC:3664.6447598339
SARIMA(1, 1, 0)x(0, 0, 0, 12) - AIC:1308.161871082466
SARIMA(1, 1, 0)x(0, 0, 1, 12) - AIC:1135.2955447585712
SARIMA(1, 1, 0)x(0, 0, 2, 12) - AIC:963.9405391257687
SARIMA(1, 1, 0)x(0, 0, 3, 12) - AIC:3700.2406067335487
SARIMA(1, 1, 0)x(1, 0, 0, 12) - AIC:1124.8860786804528
SARIMA(1, 1, 0)x(1, 0, 1, 12) - AIC:1105.4080055028642
SARIMA(1, 1, 0)x(1, 0, 2, 12) - AIC:958.5001972948144
SARIMA(1, 1, 0)x(1, 0, 3, 12) - AIC:3230.3602345110353
SARIMA(1, 1, 0)x(2, 0, 0, 12) - AIC:939.098477866423
SARIMA(1, 1, 0)x(2, 0, 1, 12) - AIC:940.9087133661069
SARIMA(1, 1, 0)x(2, 0, 2, 12) - AIC:942.2973103071042
SARIMA(1, 1, 0)x(2, 0, 3, 12) - AIC:5566.424430471874
SARIMA(1, 1, 0)x(3, 0, 0, 12) - AIC:819.3931032272714
SARIMA(1, 1, 0)x(3, 0, 1, 12) - AIC:821.1501371573529
```

```
SARIMA(1, 1, 0)x(3, 0, 2, 12) - AIC:819.1305086853257
SARIMA(1, 1, 0)x(3, 0, 3, 12) - AIC:3496.3941152286634
SARIMA(1, 1, 1)x(0, 0, 0, 12) - AIC:1262.1840064255503
SARIMA(1, 1, 1)x(0, 0, 1, 12) - AIC:1094.3172708640618
SARIMA(1, 1, 1)x(0, 0, 2, 12) - AIC:923.0862224063877
SARIMA(1, 1, 1)x(0, 0, 3, 12) - AIC:3968.151036415084
SARIMA(1, 1, 1)x(1, 0, 0, 12) - AIC:1083.3937965031319
SARIMA(1, 1, 1)x(1, 0, 1, 12) - AIC:1054.718054713631
SARIMA(1, 1, 1)x(1, 0, 2, 12) - AIC:916.3549428507811
SARIMA(1, 1, 1)x(1, 0, 3, 12) - AIC:3749.0936590196657
SARIMA(1, 1, 1)x(2, 0, 0, 12) - AIC:905.924906083923
SARIMA(1, 1, 1)x(2, 0, 1, 12) - AIC:907.2972867470744
SARIMA(1, 1, 1)x(2, 0, 2, 12) - AIC:900.6725795936848
SARIMA(1, 1, 1)x(2, 0, 3, 12) - AIC:5152.38589464608
SARIMA(1, 1, 1)x(3, 0, 0, 12) - AIC:789.2360439386595
SARIMA(1, 1, 1)x(3, 0, 1, 12) - AIC:790.9603447839742
SARIMA(1, 1, 1)x(3, 0, 2, 12) - AIC:790.8113850256015
SARIMA(1, 1, 1)x(3, 0, 3, 12) - AIC:3780.1436741469934
SARIMA(1, 1, 2)x(0, 0, 0, 12) - AIC:1251.9495040706115
SARIMA(1, 1, 2)x(0, 0, 1, 12) - AIC:1085.486192810099
SARIMA(1, 1, 2)x(0, 0, 2, 12) - AIC:915.4938402461073
SARIMA(1, 1, 2)x(0, 0, 3, 12) - AIC:3854.787036180154
SARIMA(1, 1, 2)x(1, 0, 0, 12) - AIC:1090.7760928072626
SARIMA(1, 1, 2)x(1, 0, 1, 12) - AIC:1042.6183210782174
SARIMA(1, 1, 2)x(1, 0, 2, 12) - AIC:906.7318533852929
SARIMA(1, 1, 2)x(1, 0, 3, 12) - AIC:3197.767324309791
SARIMA(1, 1, 2)x(2, 0, 0, 12) - AIC:906.1690197368665
SARIMA(1, 1, 2)x(2, 0, 1, 12) - AIC:907.4597832265824
SARIMA(1, 1, 2)x(2, 0, 2, 12) - AIC:896.6869638209376
SARIMA(1, 1, 2)x(2, 0, 3, 12) - AIC:5278.451258125427
SARIMA(1, 1, 2)x(3, 0, 0, 12) - AIC:791.0073801542183
SARIMA(1, 1, 2)x(3, 0, 1, 12) - AIC:792.6312809606984
SARIMA(1, 1, 2)x(3, 0, 2, 12) - AIC:792.1390325086993
SARIMA(1, 1, 2)x(3, 0, 3, 12) - AIC:3667.7075335394716
SARIMA(1, 1, 3)x(0, 0, 0, 12) - AIC:1245.5463125316312
SARIMA(1, 1, 3)x(0, 0, 1, 12) - AIC:1072.9849747442865
SARIMA(1, 1, 3)x(0, 0, 2, 12) - AIC:907.4130105119083
SARIMA(1, 1, 3)x(0, 0, 3, 12) - AIC:3768.8544065088495
SARIMA(1, 1, 3)x(1, 0, 0, 12) - AIC:1082.3281487675906
SARIMA(1, 1, 3)x(1, 0, 1, 12) - AIC:1036.195093007309
SARIMA(1, 1, 3)x(1, 0, 2, 12) - AIC:897.5232291576681
SARIMA(1, 1, 3)x(1, 0, 3, 12) - AIC:3306.7936461131985
SARIMA(1, 1, 3)x(2, 0, 0, 12) - AIC:908.26876118975
SARIMA(1, 1, 3)x(2, 0, 1, 12) - AIC:909.7955658923438
SARIMA(1, 1, 3)x(2, 0, 2, 12) - AIC:880.2886588076616
SARIMA(1, 1, 3)x(2, 0, 3, 12) - AIC:5061.734322510496
SARIMA(1, 1, 3)x(3, 0, 0, 12) - AIC:792.9903134043286
```

```
SARIMA(1, 1, 3)x(3, 0, 1, 12) - AIC:794.6310138939964
SARIMA(1, 1, 3)x(3, 0, 2, 12) - AIC:794.0467929600394
SARIMA(1, 1, 3)x(3, 0, 3, 12) - AIC:3581.3570292711483
SARIMA(2, 1, 0)x(0, 0, 0, 12) - AIC:1280.2537561535767
SARIMA(2, 1, 0)x(0, 0, 1, 12) - AIC:1128.7773704711312
SARIMA(2, 1, 0)x(0, 0, 2, 12) - AIC:958.0793208829957
SARIMA(2, 1, 0)x(0, 0, 3, 12) - AIC:4010.6902948713973
SARIMA(2, 1, 0)x(1, 0, 0, 12) - AIC:1099.5086021575923
SARIMA(2, 1, 0)x(1, 0, 1, 12) - AIC:1076.7863198641219
SARIMA(2, 1, 0)x(1, 0, 2, 12) - AIC:951.1988165559179
SARIMA(2, 1, 0)x(1, 0, 3, 12) - AIC:4048.3739222801732
SARIMA(2, 1, 0)x(2, 0, 0, 12) - AIC:924.6004792645443
SARIMA(2, 1, 0)x(2, 0, 1, 12) - AIC:925.9757801384217
SARIMA(2, 1, 0)x(2, 0, 2, 12) - AIC:927.8380693280769
SARIMA(2, 1, 0)x(2, 0, 3, 12) - AIC:5430.619788752883
SARIMA(2, 1, 0)x(3, 0, 0, 12) - AIC:806.0536407409069
SARIMA(2, 1, 0)x(3, 0, 1, 12) - AIC:808.0472637903475
SARIMA(2, 1, 0)x(3, 0, 2, 12) - AIC:806.4988191370863
SARIMA(2, 1, 0)x(3, 0, 3, 12) - AIC:4044.7448436338773
SARIMA(2, 1, 1)x(0, 0, 0, 12) - AIC:1263.2315231799707
SARIMA(2, 1, 1)x(0, 0, 1, 12) - AIC:1094.209349194945
SARIMA(2, 1, 1)x(0, 0, 2, 12) - AIC:922.9408472071682
SARIMA(2, 1, 1)x(0, 0, 3, 12) - AIC:3952.564509209372
SARIMA(2, 1, 1)x(1, 0, 0, 12) - AIC:1071.4249601101308
SARIMA(2, 1, 1)x(1, 0, 1, 12) - AIC:1052.9244471207162
SARIMA(2, 1, 1)x(1, 0, 2, 12) - AIC:916.2424912822544
SARIMA(2, 1, 1)x(1, 0, 3, 12) - AIC:3573.9574530070927
SARIMA(2, 1, 1)x(2, 0, 0, 12) - AIC:896.5181608099839
SARIMA(2, 1, 1)x(2, 0, 1, 12) - AIC:897.6399565369456
SARIMA(2, 1, 1)x(2, 0, 2, 12) - AIC:899.4835866291438
SARIMA(2, 1, 1)x(2, 0, 3, 12) - AIC:5359.986286633573
SARIMA(2, 1, 1)x(3, 0, 0, 12) - AIC:785.3932600272979
SARIMA(2, 1, 1)x(3, 0, 1, 12) - AIC:787.3639259079526
SARIMA(2, 1, 1)x(3, 0, 2, 12) - AIC:787.0598889517852
SARIMA(2, 1, 1)x(3, 0, 3, 12) - AIC:3574.7164058641156
SARIMA(2, 1, 2)x(0, 0, 0, 12) - AIC:1253.9102116146698
SARIMA(2, 1, 2)x(0, 0, 1, 12) - AIC:1085.9643552598225
SARIMA(2, 1, 2)x(0, 0, 2, 12) - AIC:916.3258311106489
SARIMA(2, 1, 2)x(0, 0, 3, 12) - AIC:3652.5374890303156
SARIMA(2, 1, 2)x(1, 0, 0, 12) - AIC:1073.2912713714093
SARIMA(2, 1, 2)x(1, 0, 1, 12) - AIC:1044.1909352203709
SARIMA(2, 1, 2)x(1, 0, 2, 12) - AIC:907.6661488809933
SARIMA(2, 1, 2)x(1, 0, 3, 12) - AIC:3859.5457090138484
SARIMA(2, 1, 2)x(2, 0, 0, 12) - AIC:897.3464441325917
SARIMA(2, 1, 2)x(2, 0, 1, 12) - AIC:898.3781891586497
SARIMA(2, 1, 2)x(2, 0, 2, 12) - AIC:890.6687982160834
SARIMA(2, 1, 2)x(2, 0, 3, 12) - AIC:4927.208569638218
```

```
SARIMA(2, 1, 2)x(3, 0, 0, 12) - AIC:785.1199713052301
SARIMA(2, 1, 2)x(3, 0, 1, 12) - AIC:786.9410035473928
SARIMA(2, 1, 2)x(3, 0, 2, 12) - AIC:786.9640483730168
SARIMA(2, 1, 2)x(3, 0, 3, 12) - AIC:3391.0913820554165
SARIMA(2, 1, 3)x(0, 0, 0, 12) - AIC:1237.8701696494174
SARIMA(2, 1, 3)x(0, 0, 1, 12) - AIC:1067.8073333521259
SARIMA(2, 1, 3)x(0, 0, 2, 12) - AIC:908.3349290993127
SARIMA(2, 1, 3)x(0, 0, 3, 12) - AIC:3873.7153345713123
SARIMA(2, 1, 3)x(1, 0, 0, 12) - AIC:1073.2563230708904
SARIMA(2, 1, 3)x(1, 0, 1, 12) - AIC:1024.046196953152
SARIMA(2, 1, 3)x(1, 0, 2, 12) - AIC:890.9483345962153
SARIMA(2, 1, 3)x(1, 0, 3, 12) - AIC:3241.04865670104
SARIMA(2, 1, 3)x(2, 0, 0, 12) - AIC:893.3579301943331
SARIMA(2, 1, 3)x(2, 0, 1, 12) - AIC:903.2328085811615
SARIMA(2, 1, 3)x(2, 0, 2, 12) - AIC:879.2191804131846
SARIMA(2, 1, 3)x(2, 0, 3, 12) - AIC:4993.340661444756
SARIMA(2, 1, 3)x(3, 0, 0, 12) - AIC:782.4074207906145
SARIMA(2, 1, 3)x(3, 0, 1, 12) - AIC:783.4101466493466
SARIMA(2, 1, 3)x(3, 0, 2, 12) - AIC:781.0452516000171
SARIMA(2, 1, 3)x(3, 0, 3, 12) - AIC:3687.6284583854635
SARIMA(3, 1, 0)x(0, 0, 0, 12) - AIC:1269.8130062641508
SARIMA(3, 1, 0)x(0, 0, 1, 12) - AIC:1119.1703158042228
SARIMA(3, 1, 0)x(0, 0, 2, 12) - AIC:953.3440589118012
SARIMA(3, 1, 0)x(0, 0, 3, 12) - AIC:4012.690294870912
SARIMA(3, 1, 0)x(1, 0, 0, 12) - AIC:1080.546845672649
SARIMA(3, 1, 0)x(1, 0, 1, 12) - AIC:1065.5569726784586
SARIMA(3, 1, 0)x(1, 0, 2, 12) - AIC:943.6865996304871
SARIMA(3, 1, 0)x(1, 0, 3, 12) - AIC:3730.662828037154
SARIMA(3, 1, 0)x(2, 0, 0, 12) - AIC:909.72031980122
SARIMA(3, 1, 0)x(2, 0, 1, 12) - AIC:911.0620713976488
SARIMA(3, 1, 0)x(2, 0, 2, 12) - AIC:913.0438816472875
SARIMA(3, 1, 0)x(2, 0, 3, 12) - AIC:5274.645502607426
SARIMA(3, 1, 0)x(3, 0, 0, 12) - AIC:796.0564701527965
SARIMA(3, 1, 0)x(3, 0, 1, 12) - AIC:795.5410167661927
SARIMA(3, 1, 0)x(3, 0, 2, 12) - AIC:794.6677017193509
SARIMA(3, 1, 0)x(3, 0, 3, 12) - AIC:3666.9857801847747
SARIMA(3, 1, 1)x(0, 0, 0, 12) - AIC:1255.0658180709434
SARIMA(3, 1, 1)x(0, 0, 1, 12) - AIC:1095.6894641228312
SARIMA(3, 1, 1)x(0, 0, 2, 12) - AIC:923.8872427547557
SARIMA(3, 1, 1)x(0, 0, 3, 12) - AIC:3954.564509209606
SARIMA(3, 1, 1)x(1, 0, 0, 12) - AIC:1064.614913651313
SARIMA(3, 1, 1)x(1, 0, 1, 12) - AIC:1046.0837554245256
SARIMA(3, 1, 1)x(1, 0, 2, 12) - AIC:917.0308009384311
SARIMA(3, 1, 1)x(1, 0, 3, 12) - AIC:3753.0919510604786
SARIMA(3, 1, 1)x(2, 0, 0, 12) - AIC:887.3204459899118
SARIMA(3, 1, 1)x(2, 0, 1, 12) - AIC:888.3279208537639
SARIMA(3, 1, 1)x(2, 0, 2, 12) - AIC:890.1571522289563
```

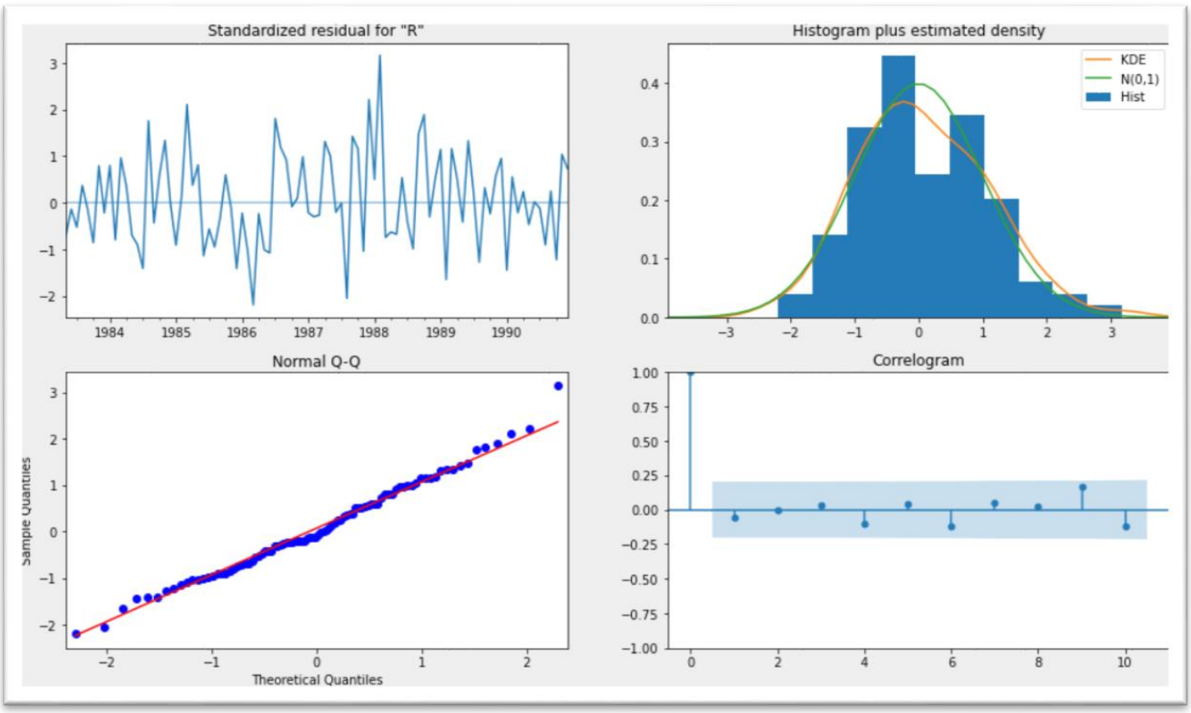
```
SARIMA(3, 1, 1)x(2, 0, 3, 12) - AIC:5001.643847802581
SARIMA(3, 1, 1)x(3, 0, 0, 12) - AIC:775.4266990255678
SARIMA(3, 1, 1)x(3, 0, 1, 12) - AIC:775.4953300800184
SARIMA(3, 1, 1)x(3, 0, 2, 12) - AIC:774.4002868175202
SARIMA(3, 1, 1)x(3, 0, 3, 12) - AIC:3894.7333312767314
SARIMA(3, 1, 2)x(0, 0, 0, 12) - AIC:1255.9835652382103
SARIMA(3, 1, 2)x(0, 0, 1, 12) - AIC:1086.7092402538601
SARIMA(3, 1, 2)x(0, 0, 2, 12) - AIC:917.798011027974
SARIMA(3, 1, 2)x(0, 0, 3, 12) - AIC:3932.6535255037943
SARIMA(3, 1, 2)x(1, 0, 0, 12) - AIC:1066.2776211376683
SARIMA(3, 1, 2)x(1, 0, 1, 12) - AIC:1046.126051830165
SARIMA(3, 1, 2)x(1, 0, 2, 12) - AIC:909.2875089213414
SARIMA(3, 1, 2)x(1, 0, 3, 12) - AIC:3808.325806015446
SARIMA(3, 1, 2)x(2, 0, 0, 12) - AIC:889.3191140331114
SARIMA(3, 1, 2)x(2, 0, 1, 12) - AIC:890.3277441304256
SARIMA(3, 1, 2)x(2, 0, 2, 12) - AIC:896.1432066357379
SARIMA(3, 1, 2)x(2, 0, 3, 12) - AIC:3462.0574428122018
SARIMA(3, 1, 2)x(3, 0, 0, 12) - AIC:777.2505642875062
SARIMA(3, 1, 2)x(3, 0, 1, 12) - AIC:776.3462148381889
SARIMA(3, 1, 2)x(3, 0, 2, 12) - AIC:774.8809402229421
SARIMA(3, 1, 2)x(3, 0, 3, 12) - AIC:3826.8453340343267
SARIMA(3, 1, 3)x(0, 0, 0, 12) - AIC:1243.7385974840606
SARIMA(3, 1, 3)x(0, 0, 1, 12) - AIC:1060.539766007896
SARIMA(3, 1, 3)x(0, 0, 2, 12) - AIC:910.3340156219601
SARIMA(3, 1, 3)x(0, 0, 3, 12) - AIC:3318.9727030909607
SARIMA(3, 1, 3)x(1, 0, 0, 12) - AIC:1066.861879262267
SARIMA(3, 1, 3)x(1, 0, 1, 12) - AIC:1039.766238025352
SARIMA(3, 1, 3)x(1, 0, 2, 12) - AIC:900.048370031827
SARIMA(3, 1, 3)x(1, 0, 3, 12) - AIC:3617.0907895217465
SARIMA(3, 1, 3)x(2, 0, 0, 12) - AIC:884.0463711780098
SARIMA(3, 1, 3)x(2, 0, 1, 12) - AIC:885.9217105650052
SARIMA(3, 1, 3)x(2, 0, 2, 12) - AIC:880.5595715451975
SARIMA(3, 1, 3)x(2, 0, 3, 12) - AIC:5062.831394617674
SARIMA(3, 1, 3)x(3, 0, 0, 12) - AIC:775.561018503348
SARIMA(3, 1, 3)x(3, 0, 1, 12) - AIC:776.369040262433
SARIMA(3, 1, 3)x(3, 0, 2, 12) - AIC:778.231623703088
SARIMA(3, 1, 3)x(3, 0, 3, 12) - AIC:3664.146571231692
```

	param	seasonal	AIC
222	(3, 1, 1)	(3, 0, 2, 12)	774.400287
238	(3, 1, 2)	(3, 0, 2, 12)	774.880940
220	(3, 1, 1)	(3, 0, 0, 12)	775.426699
221	(3, 1, 1)	(3, 0, 1, 12)	775.495330
252	(3, 1, 3)	(3, 0, 0, 12)	775.561019

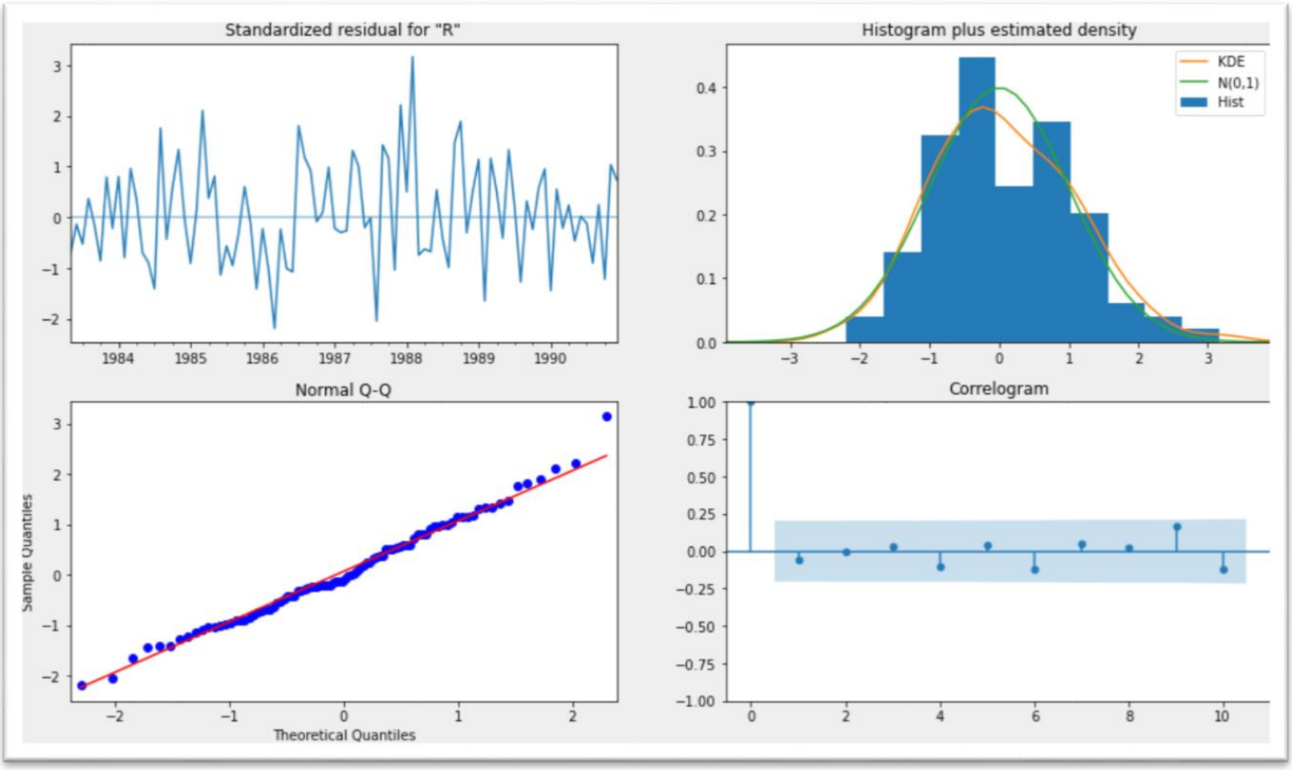
Table 101: AIC Values

```
                                      SARIMAX Results
============================================================================================
Dep. Variable:                                 Rose   No. Observations:                  132
Model:             SARIMAX(3, 1, 1)x(3, 0, [1, 2], 12)   Log Likelihood              -377.200
Date:                              Fri, 25 Mar 2022   AIC                            774.400
Time:                                      12:27:02   BIC                            799.618
Sample:                                  01-01-1980   HQIC                           784.578
                                       - 12-01-1990
Covariance Type:                                opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.0464      0.126      0.367      0.714      -0.201       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    -13.907      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       192.1336     39.623      4.849      0.000     114.474     269.793
===================================================================================
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
===================================================================================

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

Table 102: Model Summary



Graph 50: Diagnostic Plot



Graph 51: Diagnostic Plot

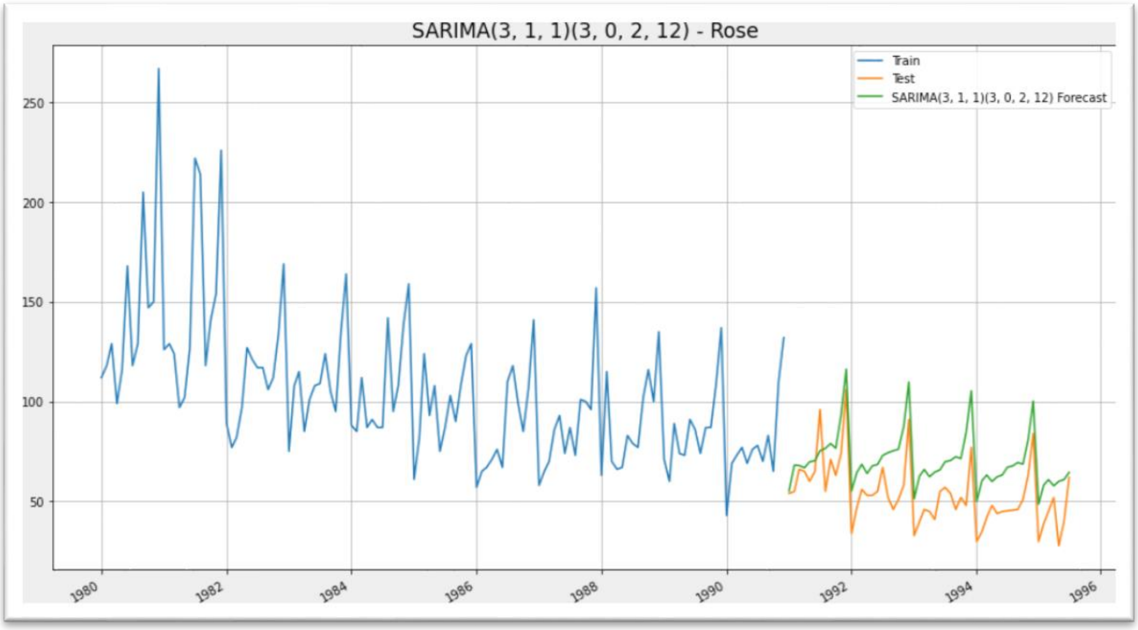
Predict on the Test Set using this model and evaluate the model:

Rose	mean	mean_se	mean_ci_lower	mean_ci_upper
1991-01-01	55.234567	13.906837	27.977668	82.491467
1991-02-01	68.122392	13.990245	40.702016	95.542768
1991-03-01	67.908360	14.011274	40.446767	95.369952
1991-04-01	66.785587	14.098600	39.152838	94.418336
1991-05-01	69.760565	14.107959	42.109474	97.411656

```
RMSE: 18.881428212560966
MAPE: 36.37439460579313
```

	Test RMSE Rose	Test MAPE Rose
ARIMA(2,1,3)	36.816347	75.845836
ARIMA(2,1,2)	36.871197	76.056213
SARIMA(3, 1, 1)(3, 0, 2, 12)	18.881428	36.374395

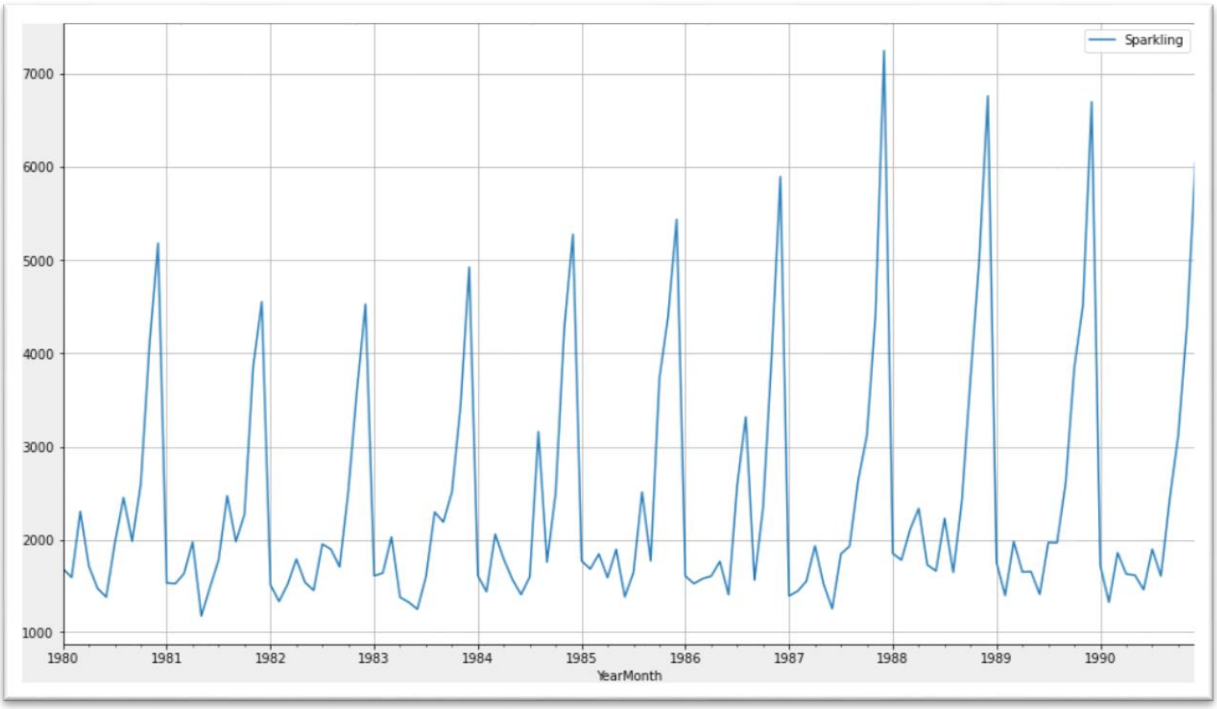
Table 103: Model Evaluation Values



Graph 52: SARIMA Forecast

ARIMA / SARIMA Modelling on SPARKLING dataset:

Check for stationarity of the Training Data - Sparkling:

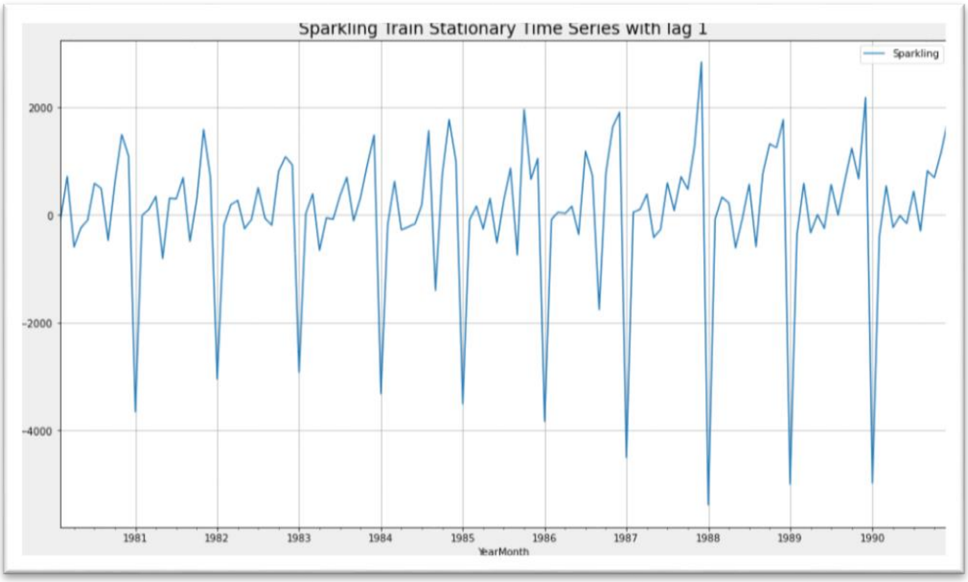


Graph 53: Checking for Stationarity of the Training Data

```
DF test statistic is -2.062
DF test p-value is 0.5674110388593696
Number of lags used 12
```

```
DF test statistic is -7.968
DF test p-value is 8.479210655516242e-11
Number of lags used 11
```

Table 104: Checking for Stationarity Values



Graph 54: Stationarity with Lag 1

Build an Automated version of an ARIMA model for which the best parameters are selected in accordance with the lowest Akaike Information Criteria (AIC) - SPARKLING:

```
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)
```

Table 105: Parameter Combination

```
ARIMA(0, 1, 0) - AIC:2267.6630357855465
ARIMA(0, 1, 1) - AIC:2263.060015591336
ARIMA(0, 1, 2) - AIC:2234.408323130674
ARIMA(0, 1, 3) - AIC:2233.9948577476116
ARIMA(1, 1, 0) - AIC:2266.6085393190087
ARIMA(1, 1, 1) - AIC:2235.7550946742404
ARIMA(1, 1, 2) - AIC:2234.5272004519366
ARIMA(1, 1, 3) - AIC:2235.6078101124103
ARIMA(2, 1, 0) - AIC:2260.365743968097
ARIMA(2, 1, 1) - AIC:2233.7776262581274
ARIMA(2, 1, 2) - AIC:2213.50921703971
ARIMA(2, 1, 3) - AIC:2232.983057575394
ARIMA(3, 1, 0) - AIC:2257.72337899794
ARIMA(3, 1, 1) - AIC:2235.4989865071907
ARIMA(3, 1, 2) - AIC:2230.7572943437854
ARIMA(3, 1, 3) - AIC:2221.4519770502657
```

Table 106: Prediction on Sparkling Test Set

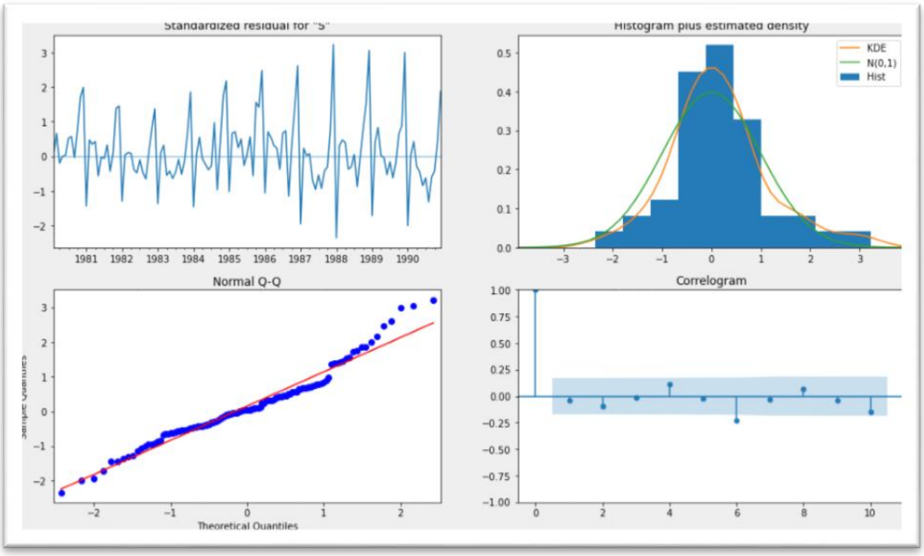
	param	AIC
10	(2, 1, 2)	2213.509217
15	(3, 1, 3)	2221.451977
14	(3, 1, 2)	2230.757294
11	(2, 1, 3)	2232.983058
9	(2, 1, 1)	2233.777626

Table 107: AIC Values

```
                               SARIMAX Results
==============================================================================
Dep. Variable:              Sparkling   No. Observations:                  132
Model:                 ARIMA(2, 1, 2)   Log Likelihood               -1101.755
Date:                Fri, 25 Mar 2022   AIC                           2213.509
Time:                        12:27:33   BIC                           2227.885
Sample:                    01-01-1980   HQIC                          2219.351
                         - 12-01-1990
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          1.3121      0.046     28.786      0.000       1.223       1.401
ar.L2         -0.5593      0.072     -7.731      0.000      -0.701      -0.417
ma.L1         -1.9916      0.110    -18.184      0.000      -2.206      -1.777
ma.L2          0.9999      0.110      9.093      0.000       0.784       1.215
sigma2      1.099e+06      2e-07   5.49e+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
===================================================================================
```

Table 108: Model Summary

Diagnostics plot - Sparkling:



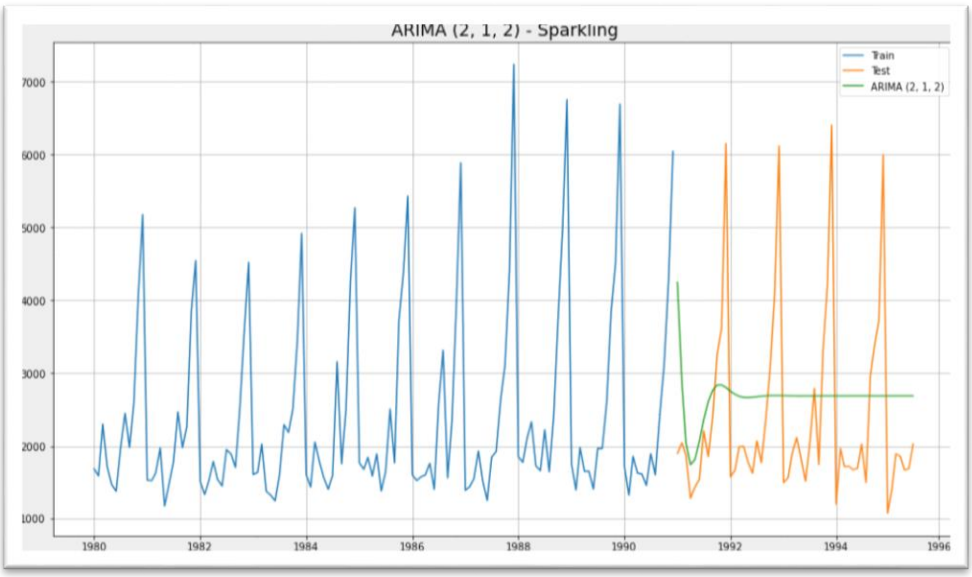
Graph 55: Diagnostic Plot

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

```
RMSE: 1299.980372953183
MAPE: 47.10001658999489
```

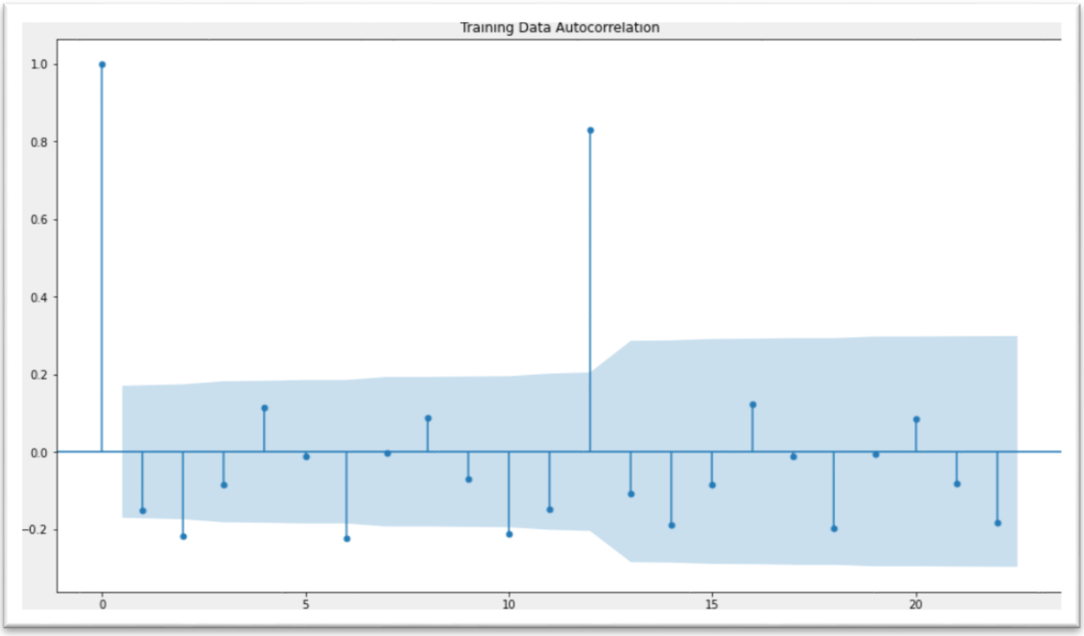
	RMSE	MAPE
ARIMA(2,1,2)	1299.980373	47.100017

Table 109: RMSE and MAPE Values



Graph 56: ARIMA Plot

Build an Automated version of a SARIMA model for which the best parameters are selected in accordance with the lowest Akaike Information Criteria (AIC) - SPARKLING:



Graph 57: Data Autocorrelation

```
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)
```

Table 110: Parameter Combination

```
SARIMA(0, 1, 0)x(0, 0, 0, 12) - AIC:2251.3597196862966
SARIMA(0, 1, 0)x(0, 0, 1, 12) - AIC:1956.261461684553
SARIMA(0, 1, 0)x(0, 0, 2, 12) - AIC:1723.1533640237064
SARIMA(0, 1, 0)x(0, 0, 3, 12) - AIC:4164.182719717596
SARIMA(0, 1, 0)x(1, 0, 0, 12) - AIC:1837.4366022456675
SARIMA(0, 1, 0)x(1, 0, 1, 12) - AIC:1806.9905301389308
SARIMA(0, 1, 0)x(1, 0, 2, 12) - AIC:1633.2108735792024
SARIMA(0, 1, 0)x(1, 0, 3, 12) - AIC:3491.2915843395344
```

```
SARIMA(0, 1, 0)x(2, 0, 0, 12) - AIC:1648.3776153470856
SARIMA(0, 1, 0)x(2, 0, 1, 12) - AIC:1647.2054158613328
SARIMA(0, 1, 0)x(2, 0, 2, 12) - AIC:1630.9898053920729
SARIMA(0, 1, 0)x(2, 0, 3, 12) - AIC:2924.5823460655906
SARIMA(0, 1, 0)x(3, 0, 0, 12) - AIC:1467.4574095308408
SARIMA(0, 1, 0)x(3, 0, 1, 12) - AIC:1469.187105262566
SARIMA(0, 1, 0)x(3, 0, 2, 12) - AIC:1471.059453006515
SARIMA(0, 1, 0)x(3, 0, 3, 12) - AIC:2963.6771092363338
SARIMA(0, 1, 1)x(0, 0, 0, 12) - AIC:2230.1629078505825
SARIMA(0, 1, 1)x(0, 0, 1, 12) - AIC:1923.7688649566726
SARIMA(0, 1, 1)x(0, 0, 2, 12) - AIC:1692.7089572748998
SARIMA(0, 1, 1)x(0, 0, 3, 12) - AIC:3555.7996335239195
SARIMA(0, 1, 1)x(1, 0, 0, 12) - AIC:1797.179588183803
SARIMA(0, 1, 1)x(1, 0, 1, 12) - AIC:1738.0903193766617
SARIMA(0, 1, 1)x(1, 0, 2, 12) - AIC:1570.1509144566762
SARIMA(0, 1, 1)x(1, 0, 3, 12) - AIC:3289.4091887195327
SARIMA(0, 1, 1)x(2, 0, 0, 12) - AIC:1605.6751954175065
SARIMA(0, 1, 1)x(2, 0, 1, 12) - AIC:1599.224508535863
SARIMA(0, 1, 1)x(2, 0, 2, 12) - AIC:1570.3683738379118
SARIMA(0, 1, 1)x(2, 0, 3, 12) - AIC:2781.528924313392
SARIMA(0, 1, 1)x(3, 0, 0, 12) - AIC:1428.4607679626317
SARIMA(0, 1, 1)x(3, 0, 1, 12) - AIC:1428.872799048972
SARIMA(0, 1, 1)x(3, 0, 2, 12) - AIC:1429.589188208307
SARIMA(0, 1, 1)x(3, 0, 3, 12) - AIC:2582.922206064308
SARIMA(0, 1, 2)x(0, 0, 0, 12) - AIC:2187.441010168703
SARIMA(0, 1, 2)x(0, 0, 1, 12) - AIC:1887.9128007200654
SARIMA(0, 1, 2)x(0, 0, 2, 12) - AIC:1659.878989129546
SARIMA(0, 1, 2)x(0, 0, 3, 12) - AIC:4162.635770743183
SARIMA(0, 1, 2)x(1, 0, 0, 12) - AIC:1790.032633217488
SARIMA(0, 1, 2)x(1, 0, 1, 12) - AIC:1724.1675071097636
SARIMA(0, 1, 2)x(1, 0, 2, 12) - AIC:1557.160506866491
SARIMA(0, 1, 2)x(1, 0, 3, 12) - AIC:3901.9109450327037
SARIMA(0, 1, 2)x(2, 0, 0, 12) - AIC:1603.9654774511769
SARIMA(0, 1, 2)x(2, 0, 1, 12) - AIC:1600.543515674843
SARIMA(0, 1, 2)x(2, 0, 2, 12) - AIC:1557.1215627001934
SARIMA(0, 1, 2)x(2, 0, 3, 12) - AIC:3387.888170272835
SARIMA(0, 1, 2)x(3, 0, 0, 12) - AIC:1428.5993407934304
SARIMA(0, 1, 2)x(3, 0, 1, 12) - AIC:1429.7448450098339
SARIMA(0, 1, 2)x(3, 0, 2, 12) - AIC:1430.6883845666819
SARIMA(0, 1, 2)x(3, 0, 3, 12) - AIC:3219.4767880264126
SARIMA(0, 1, 3)x(0, 0, 0, 12) - AIC:2168.092540843793
SARIMA(0, 1, 3)x(0, 0, 1, 12) - AIC:1873.9770701767752
SARIMA(0, 1, 3)x(0, 0, 2, 12) - AIC:1643.297366121778
SARIMA(0, 1, 3)x(0, 0, 3, 12) - AIC:3755.8482960561887
SARIMA(0, 1, 3)x(1, 0, 0, 12) - AIC:1792.031817660381
SARIMA(0, 1, 3)x(1, 0, 1, 12) - AIC:1710.6404077706738
```

```
SARIMA(0, 1, 3)x(1, 0, 2, 12) - AIC:1542.741903297107
SARIMA(0, 1, 3)x(1, 0, 3, 12) - AIC:3505.589318185014
SARIMA(0, 1, 3)x(2, 0, 0, 12) - AIC:1604.5715806973774
SARIMA(0, 1, 3)x(2, 0, 1, 12) - AIC:1601.0674829618224
SARIMA(0, 1, 3)x(2, 0, 2, 12) - AIC:1543.0489044195683
SARIMA(0, 1, 3)x(2, 0, 3, 12) - AIC:2431.798319123823
SARIMA(0, 1, 3)x(3, 0, 0, 12) - AIC:1429.6529456421736
SARIMA(0, 1, 3)x(3, 0, 1, 12) - AIC:1441.14351753829
SARIMA(0, 1, 3)x(3, 0, 2, 12) - AIC:1431.154479214692
SARIMA(0, 1, 3)x(3, 0, 3, 12) - AIC:2105.78514182967
SARIMA(1, 1, 0)x(0, 0, 0, 12) - AIC:2250.3181267386713
SARIMA(1, 1, 0)x(0, 0, 1, 12) - AIC:1954.3938339904792
SARIMA(1, 1, 0)x(0, 0, 2, 12) - AIC:1721.268847635562
SARIMA(1, 1, 0)x(0, 0, 3, 12) - AIC:3762.6051065523843
SARIMA(1, 1, 0)x(1, 0, 0, 12) - AIC:1811.244027933166
SARIMA(1, 1, 0)x(1, 0, 1, 12) - AIC:1788.5343592674053
SARIMA(1, 1, 0)x(1, 0, 2, 12) - AIC:1616.489440258681
SARIMA(1, 1, 0)x(1, 0, 3, 12) - AIC:3516.3061855701303
SARIMA(1, 1, 0)x(2, 0, 0, 12) - AIC:1621.6355080128965
SARIMA(1, 1, 0)x(2, 0, 1, 12) - AIC:1617.1356134186935
SARIMA(1, 1, 0)x(2, 0, 2, 12) - AIC:1616.541206747189
SARIMA(1, 1, 0)x(2, 0, 3, 12) - AIC:2983.227603067336
SARIMA(1, 1, 0)x(3, 0, 0, 12) - AIC:1440.5134640370259
SARIMA(1, 1, 0)x(3, 0, 1, 12) - AIC:1442.1719780402789
SARIMA(1, 1, 0)x(3, 0, 2, 12) - AIC:1443.7660392295427
SARIMA(1, 1, 0)x(3, 0, 3, 12) - AIC:2818.1969279886844
SARIMA(1, 1, 1)x(0, 0, 0, 12) - AIC:2204.9340491545727
SARIMA(1, 1, 1)x(0, 0, 1, 12) - AIC:1907.3558974131313
SARIMA(1, 1, 1)x(0, 0, 2, 12) - AIC:1678.098135264311
SARIMA(1, 1, 1)x(0, 0, 3, 12) - AIC:6588.100573704862
SARIMA(1, 1, 1)x(1, 0, 0, 12) - AIC:1775.1424469157992
SARIMA(1, 1, 1)x(1, 0, 1, 12) - AIC:1739.7167467427319
SARIMA(1, 1, 1)x(1, 0, 2, 12) - AIC:1571.3248863624835
SARIMA(1, 1, 1)x(1, 0, 3, 12) - AIC:3345.408537936073
SARIMA(1, 1, 1)x(2, 0, 0, 12) - AIC:1590.616160688121
SARIMA(1, 1, 1)x(2, 0, 1, 12) - AIC:1586.3140615064303
SARIMA(1, 1, 1)x(2, 0, 2, 12) - AIC:1571.8069967824815
SARIMA(1, 1, 1)x(2, 0, 3, 12) - AIC:2890.330251684579
SARIMA(1, 1, 1)x(3, 0, 0, 12) - AIC:1414.9112074837815
SARIMA(1, 1, 1)x(3, 0, 1, 12) - AIC:1415.9502497589062
SARIMA(1, 1, 1)x(3, 0, 2, 12) - AIC:1417.0903539882893
SARIMA(1, 1, 1)x(3, 0, 3, 12) - AIC:2723.834547003055
SARIMA(1, 1, 2)x(0, 0, 0, 12) - AIC:2188.463345050487
SARIMA(1, 1, 2)x(0, 0, 1, 12) - AIC:1889.7708307499424
SARIMA(1, 1, 2)x(0, 0, 2, 12) - AIC:1659.629142155597
SARIMA(1, 1, 2)x(0, 0, 3, 12) - AIC:4409.2972719184
SARIMA(1, 1, 2)x(1, 0, 0, 12) - AIC:1771.8259799131204
```

```
SARIMA(1, 1, 2)x(1, 0, 1, 12) - AIC:1723.9952183519094
SARIMA(1, 1, 2)x(1, 0, 2, 12) - AIC:1555.5842471707501
SARIMA(1, 1, 2)x(1, 0, 3, 12) - AIC:4149.930431367272
SARIMA(1, 1, 2)x(2, 0, 0, 12) - AIC:1588.4216932678253
SARIMA(1, 1, 2)x(2, 0, 1, 12) - AIC:1585.525085145794
SARIMA(1, 1, 2)x(2, 0, 2, 12) - AIC:1555.934562709294
SARIMA(1, 1, 2)x(2, 0, 3, 12) - AIC:3636.498380912535
SARIMA(1, 1, 2)x(3, 0, 0, 12) - AIC:1413.8102394942216
SARIMA(1, 1, 2)x(3, 0, 1, 12) - AIC:1414.9146975917229
SARIMA(1, 1, 2)x(3, 0, 2, 12) - AIC:1415.9690333276847
SARIMA(1, 1, 2)x(3, 0, 3, 12) - AIC:3464.4980853091424
SARIMA(1, 1, 3)x(0, 0, 0, 12) - AIC:2171.026403977742
SARIMA(1, 1, 3)x(0, 0, 1, 12) - AIC:1869.938145529668
SARIMA(1, 1, 3)x(0, 0, 2, 12) - AIC:1646.5176726123618
SARIMA(1, 1, 3)x(0, 0, 3, 12) - AIC:3441.624720132434
SARIMA(1, 1, 3)x(1, 0, 0, 12) - AIC:1772.3921596103721
SARIMA(1, 1, 3)x(1, 0, 1, 12) - AIC:1711.9554897979485
SARIMA(1, 1, 3)x(1, 0, 2, 12) - AIC:1542.5744605345103
SARIMA(1, 1, 3)x(1, 0, 3, 12) - AIC:3191.975309318338
SARIMA(1, 1, 3)x(2, 0, 0, 12) - AIC:1590.4160307386474
SARIMA(1, 1, 3)x(2, 0, 1, 12) - AIC:1587.3669049138343
SARIMA(1, 1, 3)x(2, 0, 2, 12) - AIC:1543.1292697816662
SARIMA(1, 1, 3)x(2, 0, 3, 12) - AIC:2717.6533581564445
SARIMA(1, 1, 3)x(3, 0, 0, 12) - AIC:1415.7723354506957
SARIMA(1, 1, 3)x(3, 0, 1, 12) - AIC:1416.7509722069967
SARIMA(1, 1, 3)x(3, 0, 2, 12) - AIC:1417.7649186698627
SARIMA(1, 1, 3)x(3, 0, 3, 12) - AIC:2575.3883409615573
SARIMA(2, 1, 0)x(0, 0, 0, 12) - AIC:2227.302761872421
SARIMA(2, 1, 0)x(0, 0, 1, 12) - AIC:1946.438343540748
SARIMA(2, 1, 0)x(0, 0, 2, 12) - AIC:1711.4123039833196
SARIMA(2, 1, 0)x(0, 0, 3, 12) - AIC:3327.6311027253087
SARIMA(2, 1, 0)x(1, 0, 0, 12) - AIC:1780.764606605774
SARIMA(2, 1, 0)x(1, 0, 1, 12) - AIC:1757.150660787284
SARIMA(2, 1, 0)x(1, 0, 2, 12) - AIC:1600.9702204970727
SARIMA(2, 1, 0)x(1, 0, 3, 12) - AIC:3696.9640106244
SARIMA(2, 1, 0)x(2, 0, 0, 12) - AIC:1592.240346481058
SARIMA(2, 1, 0)x(2, 0, 1, 12) - AIC:1587.634498977404
SARIMA(2, 1, 0)x(2, 0, 2, 12) - AIC:1585.9191732213492
SARIMA(2, 1, 0)x(2, 0, 3, 12) - AIC:3172.31303765675
SARIMA(2, 1, 0)x(3, 0, 0, 12) - AIC:1411.9449728031302
SARIMA(2, 1, 0)x(3, 0, 1, 12) - AIC:1413.1017619794814
SARIMA(2, 1, 0)x(3, 0, 2, 12) - AIC:1414.1975701530591
SARIMA(2, 1, 0)x(3, 0, 3, 12) - AIC:2263.4555896136367
SARIMA(2, 1, 1)x(0, 0, 0, 12) - AIC:2199.8586131454495
SARIMA(2, 1, 1)x(0, 0, 1, 12) - AIC:1905.0209495018182
SARIMA(2, 1, 1)x(0, 0, 2, 12) - AIC:1675.4234080336553
SARIMA(2, 1, 1)x(0, 0, 3, 12) - AIC:3134.4114078345547
```

SARIMA(2, 1, 1)x(1, 0, 0, 12) - AIC:1792.823429063149
SARIMA(2, 1, 1)x(1, 0, 1, 12) - AIC:1740.0911247188058
SARIMA(2, 1, 1)x(1, 0, 2, 12) - AIC:1571.98882783516
SARIMA(2, 1, 1)x(1, 0, 3, 12) - AIC:3376.367685084636
SARIMA(2, 1, 1)x(2, 0, 0, 12) - AIC:1577.1235060896936
SARIMA(2, 1, 1)x(2, 0, 1, 12) - AIC:1573.159584929901
SARIMA(2, 1, 1)x(2, 0, 2, 12) - AIC:1572.342865290834
SARIMA(2, 1, 1)x(2, 0, 3, 12) - AIC:3140.735319652319
SARIMA(2, 1, 1)x(3, 0, 0, 12) - AIC:1402.4673337237525
SARIMA(2, 1, 1)x(3, 0, 1, 12) - AIC:1403.1307561900896
SARIMA(2, 1, 1)x(3, 0, 2, 12) - AIC:1403.8241327906997
SARIMA(2, 1, 1)x(3, 0, 3, 12) - AIC:2598.4050987458068
SARIMA(2, 1, 2)x(0, 0, 0, 12) - AIC:2176.8681144920774
SARIMA(2, 1, 2)x(0, 0, 1, 12) - AIC:1892.2372645539306
SARIMA(2, 1, 2)x(0, 0, 2, 12) - AIC:1661.5523433574174
SARIMA(2, 1, 2)x(0, 0, 3, 12) - AIC:3582.9930369657127
SARIMA(2, 1, 2)x(1, 0, 0, 12) - AIC:1757.2140931725573
SARIMA(2, 1, 2)x(1, 0, 1, 12) - AIC:1725.3955655750103
SARIMA(2, 1, 2)x(1, 0, 2, 12) - AIC:1557.340401979642
SARIMA(2, 1, 2)x(1, 0, 3, 12) - AIC:2584.5884451589836
SARIMA(2, 1, 2)x(2, 0, 0, 12) - AIC:1576.0457450685449
SARIMA(2, 1, 2)x(2, 0, 1, 12) - AIC:1573.5476006755023
SARIMA(2, 1, 2)x(2, 0, 2, 12) - AIC:1557.8401269085794
SARIMA(2, 1, 2)x(2, 0, 3, 12) - AIC:2798.1579802118194
SARIMA(2, 1, 2)x(3, 0, 0, 12) - AIC:1401.4376190410158
SARIMA(2, 1, 2)x(3, 0, 1, 12) - AIC:1402.4901158302162
SARIMA(2, 1, 2)x(3, 0, 2, 12) - AIC:1403.435899463235
SARIMA(2, 1, 2)x(3, 0, 3, 12) - AIC:2807.6805776627652
SARIMA(2, 1, 3)x(0, 0, 0, 12) - AIC:2171.0395889190677
SARIMA(2, 1, 3)x(0, 0, 1, 12) - AIC:1866.7938191126489
SARIMA(2, 1, 3)x(0, 0, 2, 12) - AIC:1646.707278415763
SARIMA(2, 1, 3)x(0, 0, 3, 12) - AIC:3777.864780828321
SARIMA(2, 1, 3)x(1, 0, 0, 12) - AIC:1757.8357860065175
SARIMA(2, 1, 3)x(1, 0, 1, 12) - AIC:1710.8242028789937
SARIMA(2, 1, 3)x(1, 0, 2, 12) - AIC:1542.52884718371
SARIMA(2, 1, 3)x(1, 0, 3, 12) - AIC:3585.4226060978576
SARIMA(2, 1, 3)x(2, 0, 0, 12) - AIC:1582.4492082583247
SARIMA(2, 1, 3)x(2, 0, 1, 12) - AIC:1573.7195860232612
SARIMA(2, 1, 3)x(2, 0, 2, 12) - AIC:1543.9976754377026
SARIMA(2, 1, 3)x(2, 0, 3, 12) - AIC:3077.0824284458863
SARIMA(2, 1, 3)x(3, 0, 0, 12) - AIC:1399.7898668536177
SARIMA(2, 1, 3)x(3, 0, 1, 12) - AIC:1400.129153193059
SARIMA(2, 1, 3)x(3, 0, 2, 12) - AIC:1408.788533040485
SARIMA(2, 1, 3)x(3, 0, 3, 12) - AIC:2907.4813149983206
SARIMA(3, 1, 0)x(0, 0, 0, 12) - AIC:2208.40250139061
SARIMA(3, 1, 0)x(0, 0, 1, 12) - AIC:1946.368629141873
SARIMA(3, 1, 0)x(0, 0, 2, 12) - AIC:1709.5089923167952

```
SARIMA(3, 1, 0)x(0, 0, 3, 12) - AIC:3723.7917300744066
SARIMA(3, 1, 0)x(1, 0, 0, 12) - AIC:1762.7961104702983
SARIMA(3, 1, 0)x(1, 0, 1, 12) - AIC:1740.3783834119197
SARIMA(3, 1, 0)x(1, 0, 2, 12) - AIC:1600.0472030096348
SARIMA(3, 1, 0)x(1, 0, 3, 12) - AIC:3578.386528100591
SARIMA(3, 1, 0)x(2, 0, 0, 12) - AIC:1578.2748025558474
SARIMA(3, 1, 0)x(2, 0, 1, 12) - AIC:1572.9330814124191
SARIMA(3, 1, 0)x(2, 0, 2, 12) - AIC:1571.6808190273023
SARIMA(3, 1, 0)x(2, 0, 3, 12) - AIC:3104.260483386371
SARIMA(3, 1, 0)x(3, 0, 0, 12) - AIC:1396.8678620477535
SARIMA(3, 1, 0)x(3, 0, 1, 12) - AIC:1397.6774721256184
SARIMA(3, 1, 0)x(3, 0, 2, 12) - AIC:1398.3074722328342
SARIMA(3, 1, 0)x(3, 0, 3, 12) - AIC:2539.69508581453
SARIMA(3, 1, 1)x(0, 0, 0, 12) - AIC:2188.222098341979
SARIMA(3, 1, 1)x(0, 0, 1, 12) - AIC:1906.9690713100724
SARIMA(3, 1, 1)x(0, 0, 2, 12) - AIC:1677.4234050669895
SARIMA(3, 1, 1)x(0, 0, 3, 12) - AIC:3382.2984417261123
SARIMA(3, 1, 1)x(1, 0, 0, 12) - AIC:1746.2222518906797
SARIMA(3, 1, 1)x(1, 0, 1, 12) - AIC:1727.3085042691016
SARIMA(3, 1, 1)x(1, 0, 2, 12) - AIC:1573.684497306616
SARIMA(3, 1, 1)x(1, 0, 3, 12) - AIC:3375.189203748094
SARIMA(3, 1, 1)x(2, 0, 0, 12) - AIC:1564.9377505377156
SARIMA(3, 1, 1)x(2, 0, 1, 12) - AIC:1576.2959850977204
SARIMA(3, 1, 1)x(2, 0, 2, 12) - AIC:1576.3498127922426
SARIMA(3, 1, 1)x(2, 0, 3, 12) - AIC:3142.735237055727
SARIMA(3, 1, 1)x(3, 0, 0, 12) - AIC:1387.7883313226519
SARIMA(3, 1, 1)x(3, 0, 1, 12) - AIC:1388.6814847337996
SARIMA(3, 1, 1)x(3, 0, 2, 12) - AIC:1389.1958976359617
SARIMA(3, 1, 1)x(3, 0, 3, 12) - AIC:2937.167808877823
SARIMA(3, 1, 2)x(0, 0, 0, 12) - AIC:2187.3147271465705
SARIMA(3, 1, 2)x(0, 0, 1, 12) - AIC:1887.780425765844
SARIMA(3, 1, 2)x(0, 0, 2, 12) - AIC:1659.6255032978179
SARIMA(3, 1, 2)x(0, 0, 3, 12) - AIC:3861.6803350786818
SARIMA(3, 1, 2)x(1, 0, 0, 12) - AIC:1744.9635865774947
SARIMA(3, 1, 2)x(1, 0, 1, 12) - AIC:1727.571616541588
SARIMA(3, 1, 2)x(1, 0, 2, 12) - AIC:1559.4364186624923
SARIMA(3, 1, 2)x(1, 0, 3, 12) - AIC:3622.5257755369016
SARIMA(3, 1, 2)x(2, 0, 0, 12) - AIC:1567.9873069513087
SARIMA(3, 1, 2)x(2, 0, 1, 12) - AIC:1560.7894243367928
SARIMA(3, 1, 2)x(2, 0, 2, 12) - AIC:1559.8091487949541
SARIMA(3, 1, 2)x(2, 0, 3, 12) - AIC:3099.981096961768
SARIMA(3, 1, 2)x(3, 0, 0, 12) - AIC:1398.4779656333324
SARIMA(3, 1, 2)x(3, 0, 1, 12) - AIC:1388.6026133330377
SARIMA(3, 1, 2)x(3, 0, 2, 12) - AIC:1389.701991903919
SARIMA(3, 1, 2)x(3, 0, 3, 12) - AIC:2852.2599213812928
SARIMA(3, 1, 3)x(0, 0, 0, 12) - AIC:2155.7749542046586
SARIMA(3, 1, 3)x(0, 0, 1, 12) - AIC:1874.7818327814907
```

```
SARIMA(3, 1, 3)x(0, 0, 2, 12) - AIC:1643.7830065692258
SARIMA(3, 1, 3)x(0, 0, 3, 12) - AIC:20.0
SARIMA(3, 1, 3)x(1, 0, 0, 12) - AIC:1755.7409585506543
SARIMA(3, 1, 3)x(1, 0, 1, 12) - AIC:1709.2985080868928
SARIMA(3, 1, 3)x(1, 0, 2, 12) - AIC:1546.426687495325
SARIMA(3, 1, 3)x(1, 0, 3, 12) - AIC:3416.560076039803
SARIMA(3, 1, 3)x(2, 0, 0, 12) - AIC:1567.1767106208517
SARIMA(3, 1, 3)x(2, 0, 1, 12) - AIC:1561.9915529629466
SARIMA(3, 1, 3)x(2, 0, 2, 12) - AIC:1545.7744403964616
SARIMA(3, 1, 3)x(2, 0, 3, 12) - AIC:2211.780676342169
SARIMA(3, 1, 3)x(3, 0, 0, 12) - AIC:1389.1420909426768
SARIMA(3, 1, 3)x(3, 0, 1, 12) - AIC:1393.1698775529776
SARIMA(3, 1, 3)x(3, 0, 2, 12) - AIC:1391.6961897760914
SARIMA(3, 1, 3)x(3, 0, 3, 12) - AIC:2902.6859025010867
```

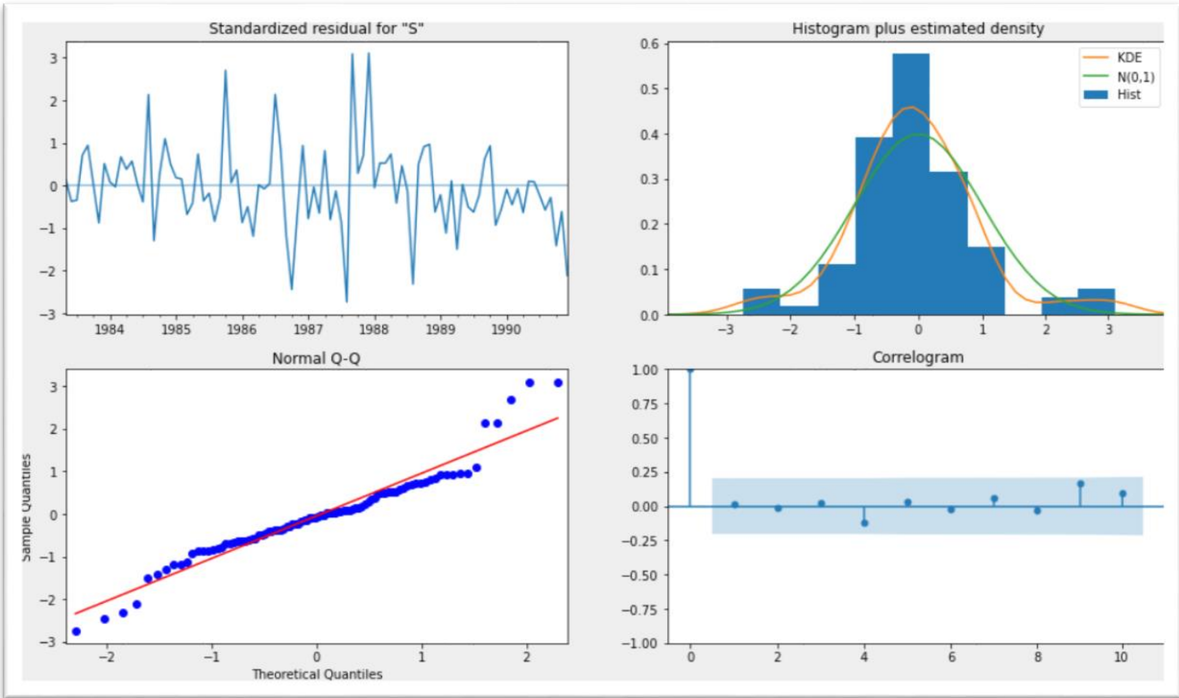
Table 111: Prediction on Sparkling Test Set

	param	seasonal	AIC
243	(3, 1, 3)	(0, 0, 3, 12)	20.000000
220	(3, 1, 1)	(3, 0, 0, 12)	1387.788331
237	(3, 1, 2)	(3, 0, 1, 12)	1388.602613
221	(3, 1, 1)	(3, 0, 1, 12)	1388.681485
252	(3, 1, 3)	(3, 0, 0, 12)	1389.142091
222	(3, 1, 1)	(3, 0, 2, 12)	1389.195898
238	(3, 1, 2)	(3, 0, 2, 12)	1389.701992
254	(3, 1, 3)	(3, 0, 2, 12)	1391.696190
253	(3, 1, 3)	(3, 0, 1, 12)	1393.169878
204	(3, 1, 0)	(3, 0, 0, 12)	1396.867862
205	(3, 1, 0)	(3, 0, 1, 12)	1397.677472
206	(3, 1, 0)	(3, 0, 2, 12)	1398.307472
236	(3, 1, 2)	(3, 0, 0, 12)	1398.477966
188	(2, 1, 3)	(3, 0, 0, 12)	1399.789867
189	(2, 1, 3)	(3, 0, 1, 12)	1400.129153
172	(2, 1, 2)	(3, 0, 0, 12)	1401.437619
156	(2, 1, 1)	(3, 0, 0, 12)	1402.467334
173	(2, 1, 2)	(3, 0, 1, 12)	1402.490116
157	(2, 1, 1)	(3, 0, 1, 12)	1403.130756
174	(2, 1, 2)	(3, 0, 2, 12)	1403.435899
158	(2, 1, 1)	(3, 0, 2, 12)	1403.824133
190	(2, 1, 3)	(3, 0, 2, 12)	1408.788533
140	(2, 1, 0)	(3, 0, 0, 12)	1411.944973
141	(2, 1, 0)	(3, 0, 1, 12)	1413.101762
108	(1, 1, 2)	(3, 0, 0, 12)	1413.810239
142	(2, 1, 0)	(3, 0, 2, 12)	1414.197570
92	(1, 1, 1)	(3, 0, 0, 12)	1414.911207
109	(1, 1, 2)	(3, 0, 1, 12)	1414.914698
124	(1, 1, 3)	(3, 0, 0, 12)	1415.772335
93	(1, 1, 1)	(3, 0, 1, 12)	1415.950250
110	(1, 1, 2)	(3, 0, 2, 12)	1415.969033
125	(1, 1, 3)	(3, 0, 1, 12)	1416.750972
94	(1, 1, 1)	(3, 0, 2, 12)	1417.090354
126	(1, 1, 3)	(3, 0, 2, 12)	1417.764919
28	(0, 1, 1)	(3, 0, 0, 12)	1428.460768
44	(0, 1, 2)	(3, 0, 0, 12)	1428.599341
29	(0, 1, 1)	(3, 0, 1, 12)	1428.872799
30	(0, 1, 1)	(3, 0, 2, 12)	1429.589188
60	(0, 1, 3)	(3, 0, 0, 12)	1429.652946
45	(0, 1, 2)	(3, 0, 1, 12)	1429.744845

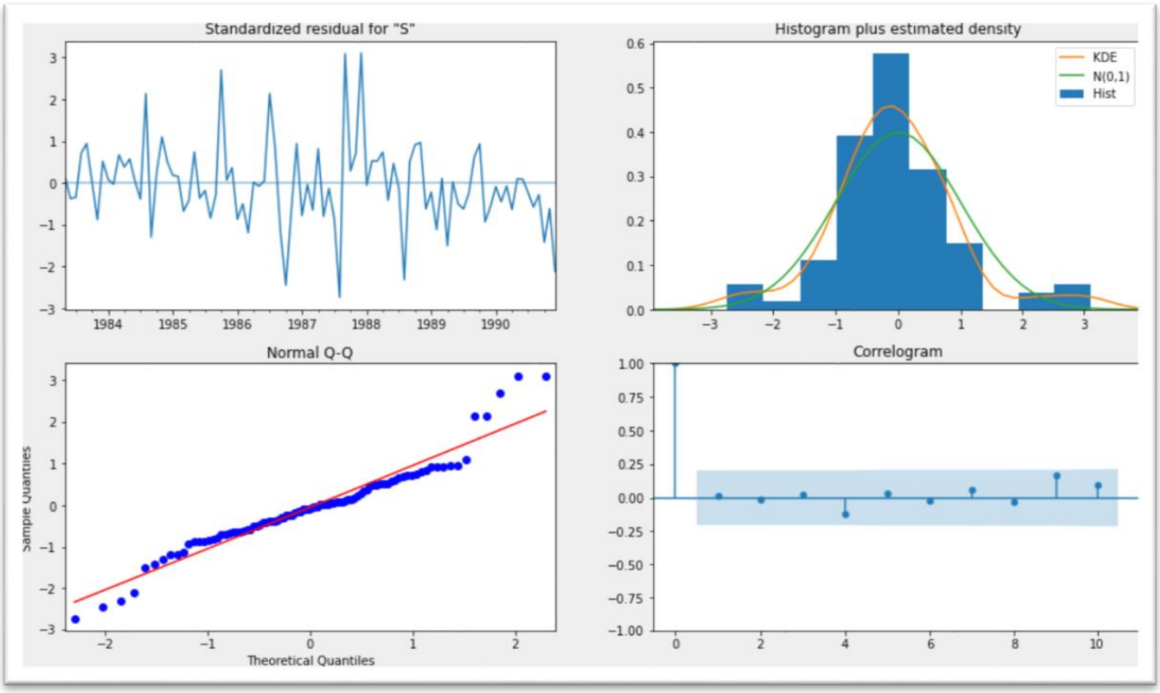
Table 112: AIC Values

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                            Sparkling   No. Observations:                132
Model:             SARIMAX(3, 1, 1)x(3, 0, [], 12)   Log Likelihood                -685.894
Date:                              Fri, 25 Mar 2022   AIC                           1387.788
Time:                                      12:31:50   BIC                           1407.963
Sample:                                  01-01-1980   HQIC                          1395.931
                                       - 12-01-1990
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     -0.617      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       0.5804      0.104      5.574      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
===================================================================================

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

Table 113: Model Summary



Graph 58: Diagnostic Plot



Graph 59: Diagnostic Plot

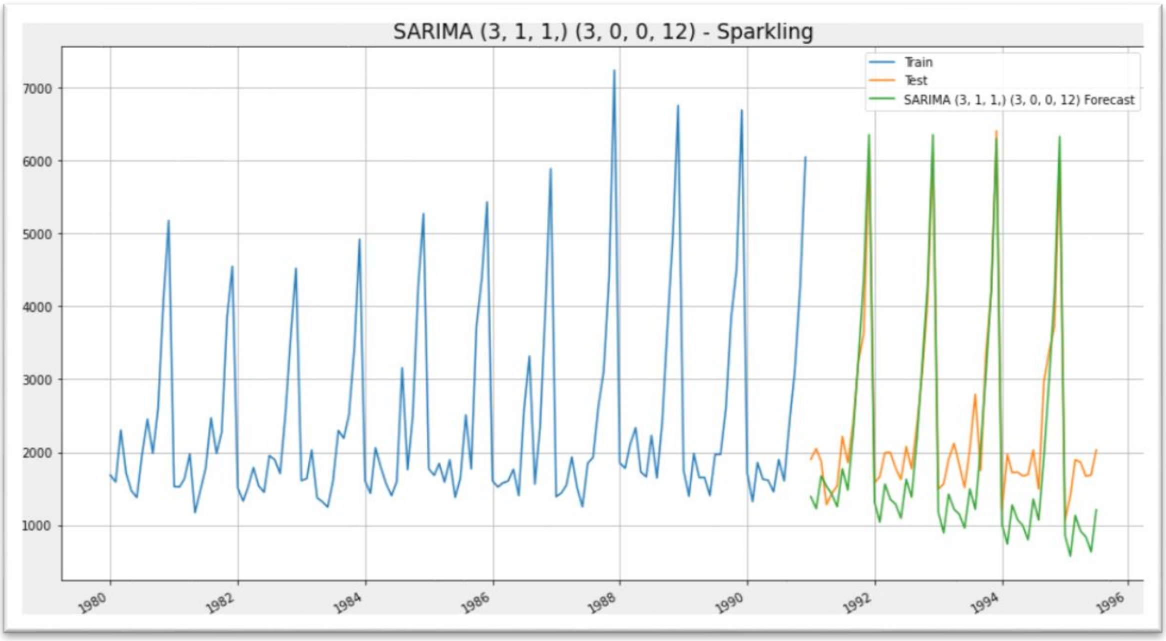
Predict on the Test Set using this model and evaluate the model:

Sparkling	mean	mean_se	mean_ci_lower	mean_ci_upper
1991-01-01	1389.352726	415.861210	574.279731	2204.425720
1991-02-01	1224.663406	427.865187	386.063050	2063.263762
1991-03-01	1673.330013	428.009765	834.446288	2512.213737
1991-04-01	1533.303629	432.773658	685.082846	2381.524411
1991-05-01	1425.948866	435.887464	571.625136	2280.272597

```
RMSE: 601.2443960941916
MAPE: 25.870720702982446
```

	RMSE	MAPE
ARIMA(2,1,2)	1299.980373	47.100017
ARIMA(0,1,0)	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	601.244396	25.870721

Table 114: Model Evaluation



Graph 60: SARIMA PLOT

Till Now, Best Model for Rose with Least RMSE —-> SARIMA (3, 1, 1) (3, 0, 2, 12)
Till Now, Best Model for Sparkling with Least RMSE ->SARIMA (3, 1, 1) (3, 0, 0, 12)

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.

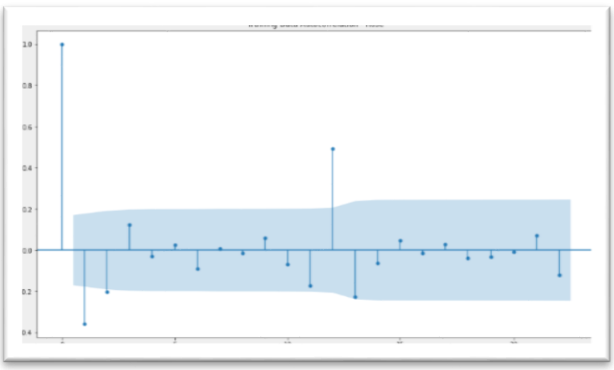
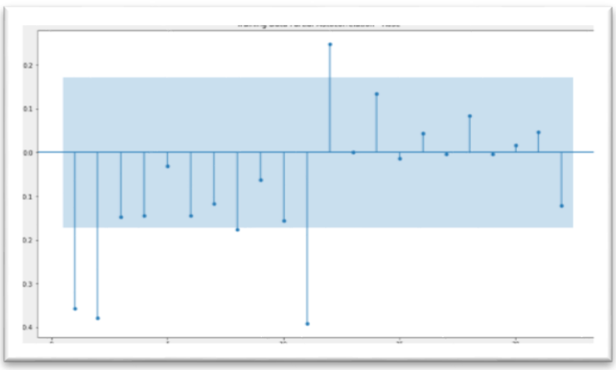
Auto-Correlation Function (ACF):

- Autocorrelation refers to how correlated a time series is with its past values.
- 'Auto' part of Autocorrelation refers to Correlation of any time instance with its previous time instance in the SAME Time Series
- ACF is the plot used to see the correlation between the points, up to and including the lag unit
- ACF indicates the value of 'q' - which is the Moving Average parameter in ARIMA / SARIMA models

Partial Auto-Correlation Function (PACF):

- Partial Autocorrelation refers to how correlated a time series is with its past lag values.
- PACF is the plot used to see the correlation between the lag points
- PACF indicates the value of 'p' - which is the Auto-Regressive parameter in ARIMA / SARIMA models

Build a version of the SARIMA model for which the best parameters are selected by looking at the ACF and the PACF plots. - Seasonality at 12:

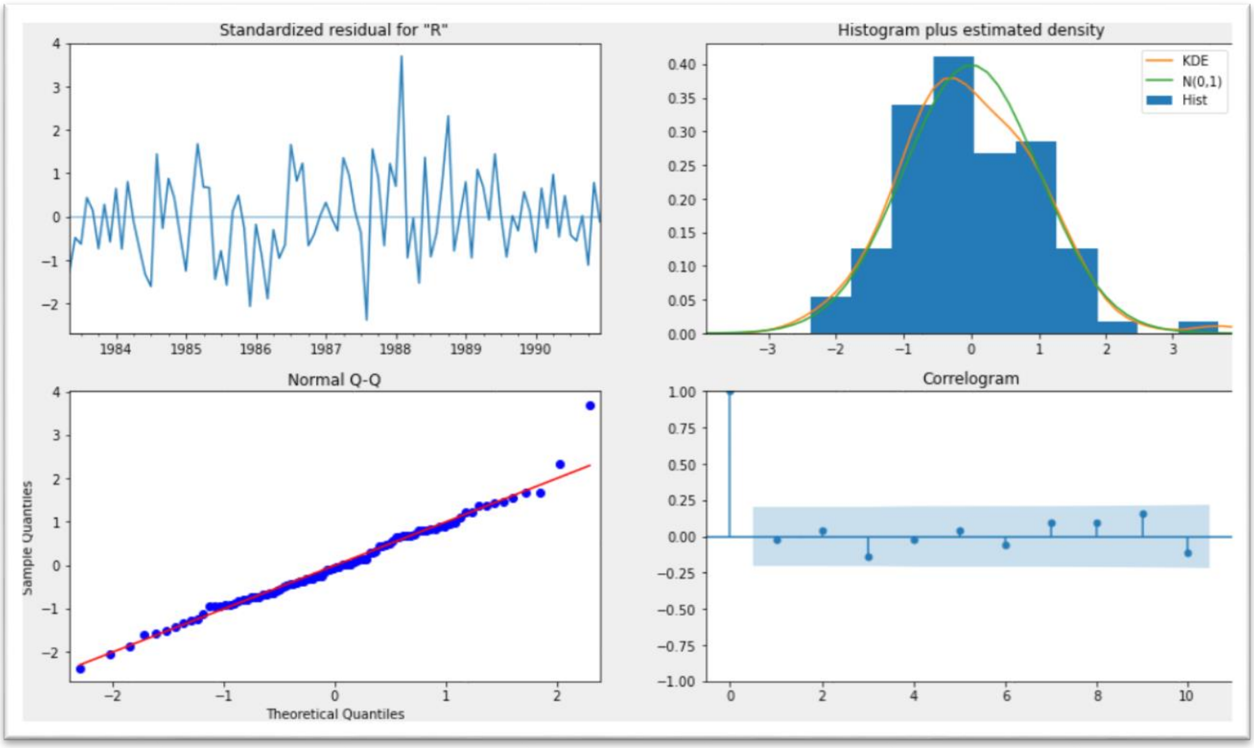
Graph 61: Data Autocorrelation

```
                                    SARIMAX Results
==========================================================================================
Dep. Variable:                                Rose   No. Observations:                  132
Model:             SARIMAX(2, 1, 2)x(2, 1, 2, 12)   Log Likelihood                -379.498
Date:                             Fri, 25 Mar 2022   AIC                            776.996
Time:                                     12:27:18   BIC                            799.692
Sample:                                 01-01-1980   HQIC                           786.156
                                      - 12-01-1990
Covariance Type:                               opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.8551      0.146     -5.838      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      0.066      0.948      -0.348       0.372
ma.L2         -0.9435      0.150     -6.291      0.000      -1.237      -0.650
ar.S.L12       0.0347      0.185      0.187      0.851      -0.328       0.397
ar.S.L24      -0.0459      0.029     -1.598      0.110      -0.102       0.010
ma.S.L12      -0.7223      0.333     -2.172      0.030      -1.374      -0.071
ma.S.L24      -0.0772      0.212     -0.364      0.716      -0.493       0.339
sigma2       192.1831     39.484      4.867      0.000     114.795     269.571
===================================================================================
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
===================================================================================

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

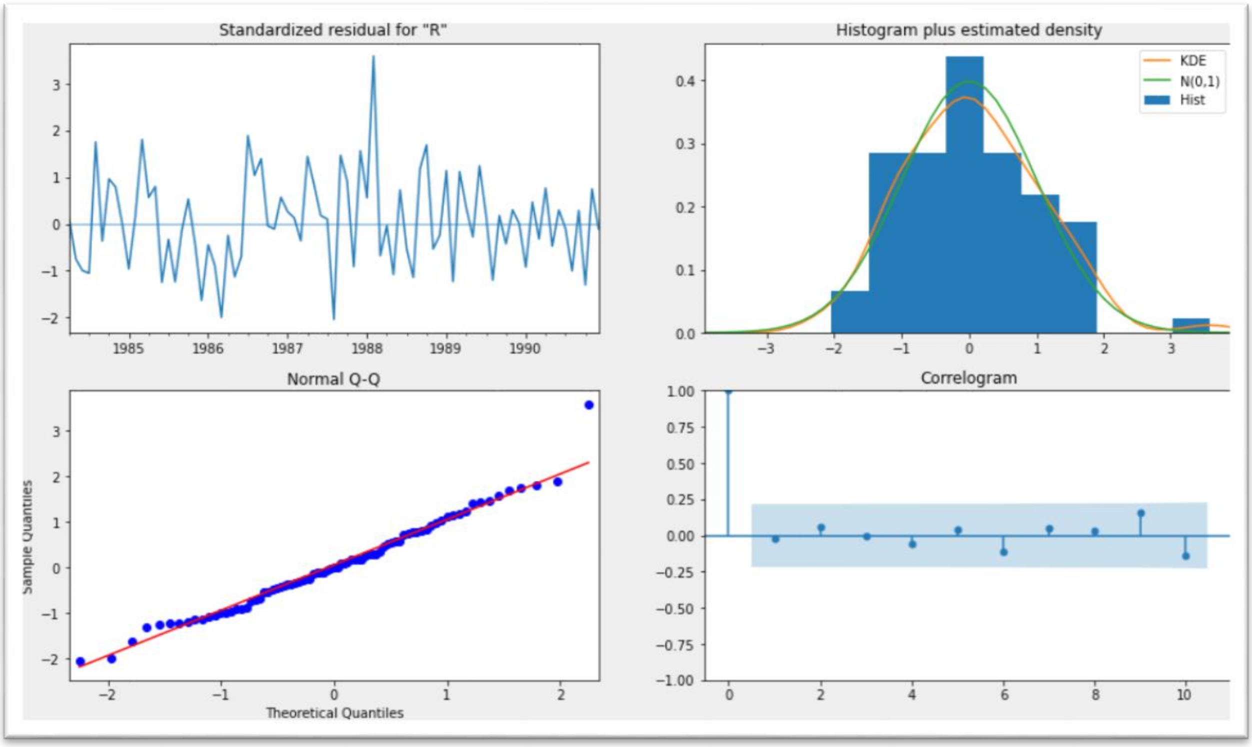
Table 115: Model Summary

```
                                    SARIMAX Results
==========================================================================================
Dep. Variable:                                Rose   No. Observations:                  132
Model:             SARIMAX(2, 1, 2)x(3, 1, 2, 12)   Log Likelihood                -334.893
Date:                             Fri, 25 Mar 2022   AIC                            689.786
Time:                                     12:27:24   BIC                            713.730
Sample:                                 01-01-1980   HQIC                           699.392
                                      - 12-01-1990
Covariance Type:                               opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.7095      0.403      1.760      0.078      -0.081       1.500
ar.L2         -0.1500      0.176     -0.854      0.393      -0.494       0.194
ma.L1         -1.6104      0.421     -3.822      0.000      -2.436      -0.785
ma.L2          0.6500      0.396      1.641      0.101      -0.127       1.427
ar.S.L12      -0.0423      0.234     -0.181      0.856      -0.500       0.416
ar.S.L24      -0.0168      0.158     -0.107      0.915      -0.327       0.293
ar.S.L36     3.03e-06      0.067   4.53e-05      1.000      -0.131       0.131
ma.S.L12      -0.8391    113.764     -0.007      0.994    -223.813     222.135
ma.S.L24      -0.1604     18.489     -0.009      0.993     -36.397      36.076
sigma2       185.7704   2.11e+04      0.009      0.993   -4.13e+04    4.16e+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
===================================================================================

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

Table 116: Model Summary



Graph 62: Diagnostic Plot



Graph 63: Diagnostic Plot

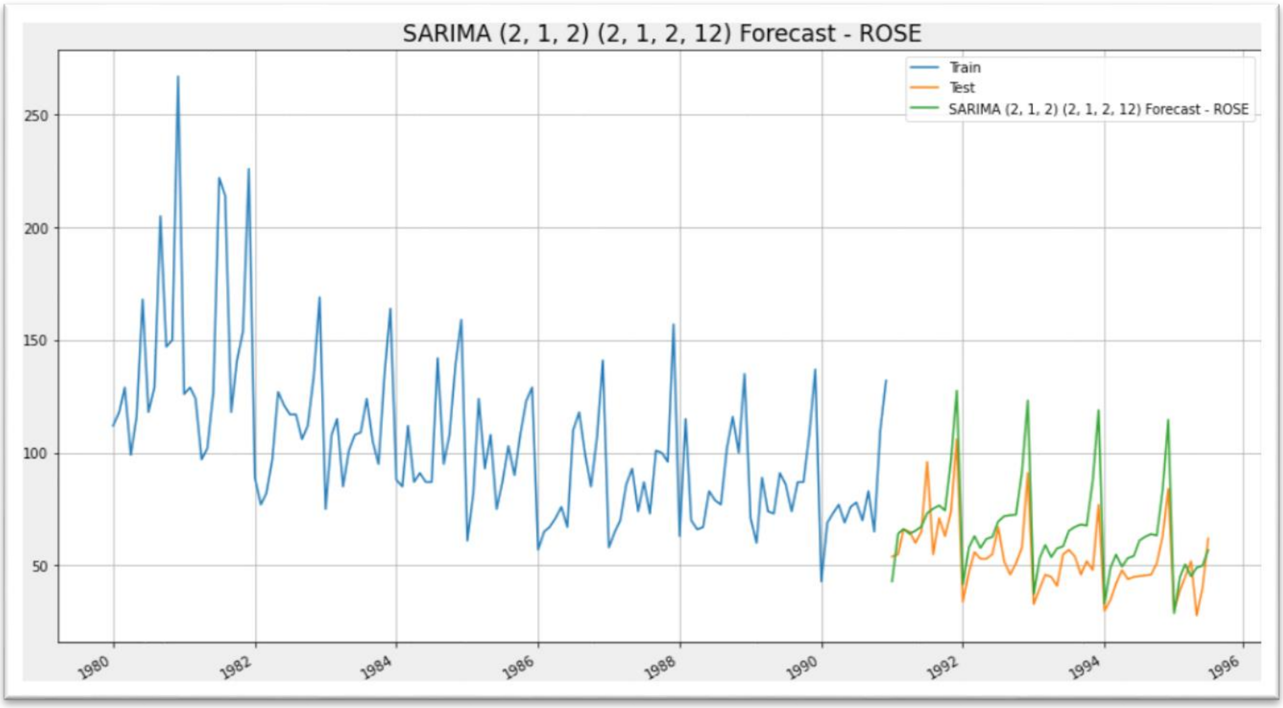
Predict on the Test Set using this model and evaluate the model:

```
RMSE: 16.550522730402722
MAPE: 25.47665331560397
```

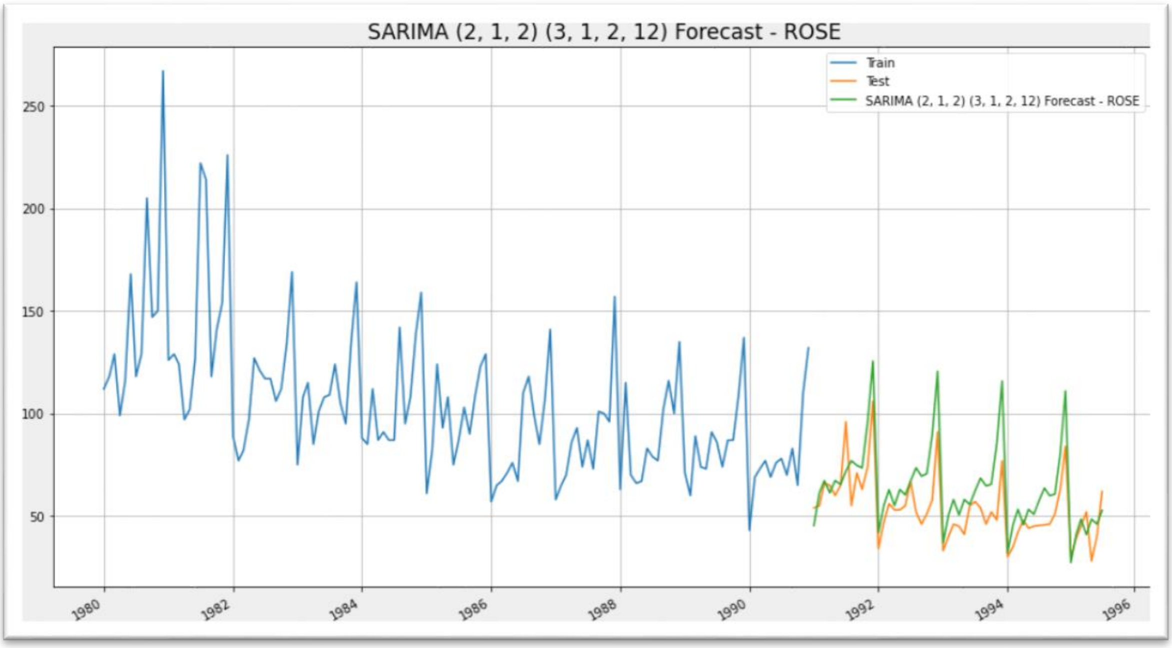
```
RMSE: 15.358805058999812
MAPE: 22.959850074445416
```

	Test RMSE Rose	Test MAPE Rose
ARIMA(2,1,3)	36.816347	75.845836
ARIMA(2,1,2)	36.871197	76.056213
SARIMA(3, 1, 1)(3, 0, 2, 12)	18.881428	36.374395
SARIMA(2,1,2)(3,1,2,12)	15.358805	22.959850

Table 117: Model Accuracy



Graph 64: SARIMA PLOT



Graph 65: SARIMA PLOT

Building the most optimum model on the Full Data:

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                                  Rose   No. Observations:               187
Model:             SARIMAX(2, 1, 2)x(2, 1, 2, 12)   Log Likelihood                -587.531
Date:                            Fri, 25 Mar 2022   AIC                           1193.062
Time:                                    12:27:29   BIC                           1219.976
Sample:                                01-01-1980   HQIC                          1203.997
                                     - 07-01-1995
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.8649      0.101     -8.572      0.000      -1.063      -0.667
ar.L2          0.0340      0.090      0.376      0.707      -0.143       0.211
ma.L1          0.0892   1649.995    5.4e-05      1.000   -3233.842    3234.020
ma.L2         -0.9108   1502.873     -0.001      1.000   -2946.488    2944.666
ar.S.L12       0.0719      0.166      0.434      0.664      -0.253       0.396
ar.S.L24      -0.0357      0.017     -2.045      0.041      -0.070      -0.001
ma.S.L12      -0.6869      0.222     -3.090      0.002      -1.123      -0.251
ma.S.L24      -0.0549      0.151     -0.365      0.715      -0.350       0.240
sigma2       158.9029   2.62e+05      0.001      1.000   -5.14e+05    5.14e+05
===================================================================================
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).
```

Table 118 Model Summary

Evaluate the model on the whole data and predict 12 months into the future:

Rose	mean	mean_se	mean_ci_lower	mean_ci_upper
1995-08-01	43.082945	12.674005	18.242352	67.923539
1995-09-01	43.341394	12.959728	17.940794	68.741993
1995-10-01	45.826993	12.963722	20.418566	71.235421
1995-11-01	57.396972	13.124452	31.673519	83.120425
1995-12-01	85.074678	13.133279	59.333924	110.815431

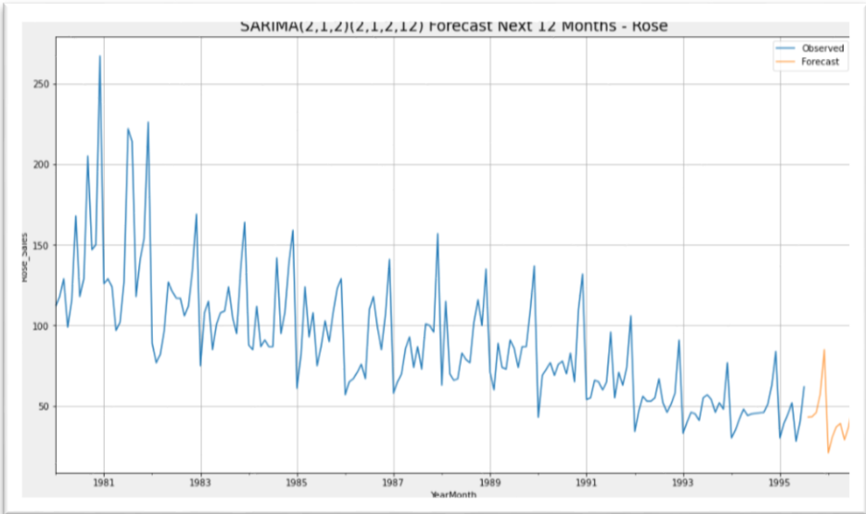
```
RMSE of the Full Model 35.910838112663996
```

	Test RMSE Rose	Test MAPE Rose
ARIMA(2,1,3)	36.816347	75.845836
ARIMA(2,1,2)	36.871197	76.056213
SARIMA(3, 1, 1)(3, 0, 2, 12)	18.881428	36.374395
SARIMA(2,1,2)(3,1,2,12)	15.358805	22.959850

	Test RMSE Rose	Test MAPE Rose
SARIMA(2,1,2)(3,1,2,12)	15.358805	22.959850
SARIMA(3, 1, 1)(3, 0, 2, 12)	18.881428	36.374395
ARIMA(2,1,3)	36.816347	75.845836
ARIMA(2,1,2)	36.871197	76.056213

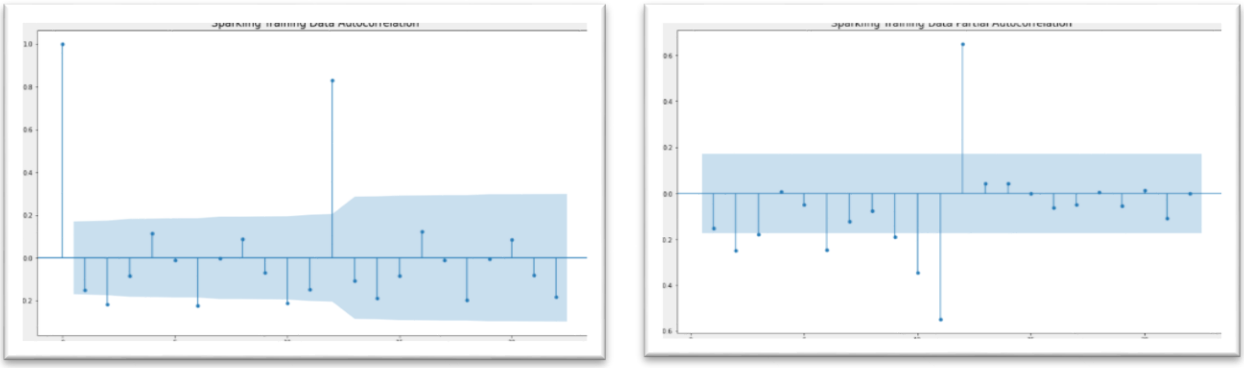
	Test RMSE Rose	Test RMSE Sparkling	Test MAPE Rose
RegressionOnTime	15.268955	1389.135175	NaN
NaiveModel	79.718773	3864.279352	NaN
SimpleAverageModel	53.460570	1275.081804	NaN
2pointTrailingMovingAverage	11.529278	813.400684	NaN
4pointTrailingMovingAverage	14.451403	1156.589694	NaN
6pointTrailingMovingAverage	14.566327	1283.927428	NaN
9pointTrailingMovingAverage	14.727630	1346.278315	NaN
Simple Exponential Smoothing	36.796228	1338.012144	NaN
Double Exponential Smoothing	15.269328	3949.993290	NaN
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686	NaN
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.510170	NaN
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.660960	379.695686	NaN
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	26.295981	352.443335	NaN
ARIMA(2,1,3)	36.816347	NaN	75.845836
ARIMA(2,1,2)	36.871197	NaN	76.056213
SARIMA(3, 1, 1)(3, 0, 2, 12)	18.881428	NaN	36.374395
SARIMA(2,1,2)(3,1,2,12)	15.358805	NaN	22.959850

Table 119: Model Accuracy



Graph 66: SARIMA PLOT

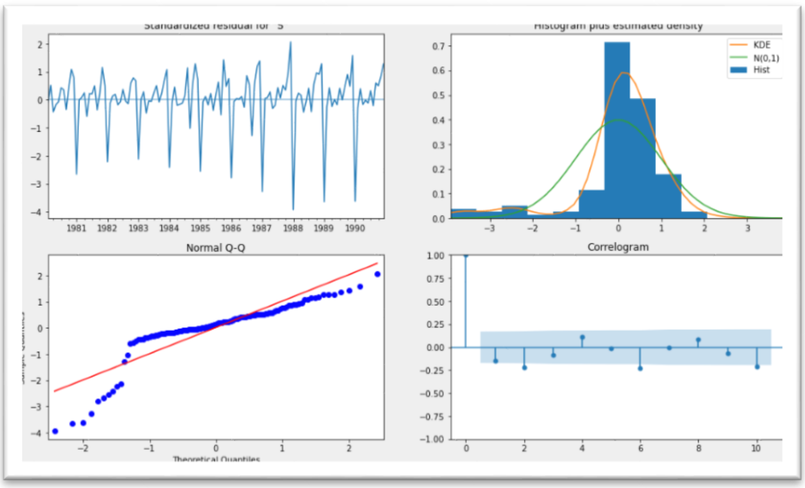
Build a version of the ARIMA model for which the best parameters are selected by looking at the ACF and the PACF plots - SPARKLING:



Graph 67: Data Autocorrelation

```
                               SARIMAX Results
==============================================================================
Dep. Variable:              Sparkling   No. Observations:                  132
Model:                 ARIMA(0, 1, 0)   Log Likelihood               -1132.832
Date:                Fri, 25 Mar 2022   AIC                           2267.663
Time:                        12:27:35   BIC                           2270.538
Sample:                    01-01-1980   HQIC                          2268.831
                         - 12-01-1990
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
sigma2      1.885e+06   1.29e+05     14.658      0.000    1.63e+06    2.14e+06
===================================================================================
Ljung-Box (L1) (Q):                   3.07   Jarque-Bera (JB):               198.83
Prob(Q):                              0.08   Prob(JB):                         0.00
Heteroskedasticity (H):               2.46   Skew:                            -1.92
Prob(H) (two-sided):                  0.00   Kurtosis:                         7.65
===================================================================================
```

Table 120: Model Summary



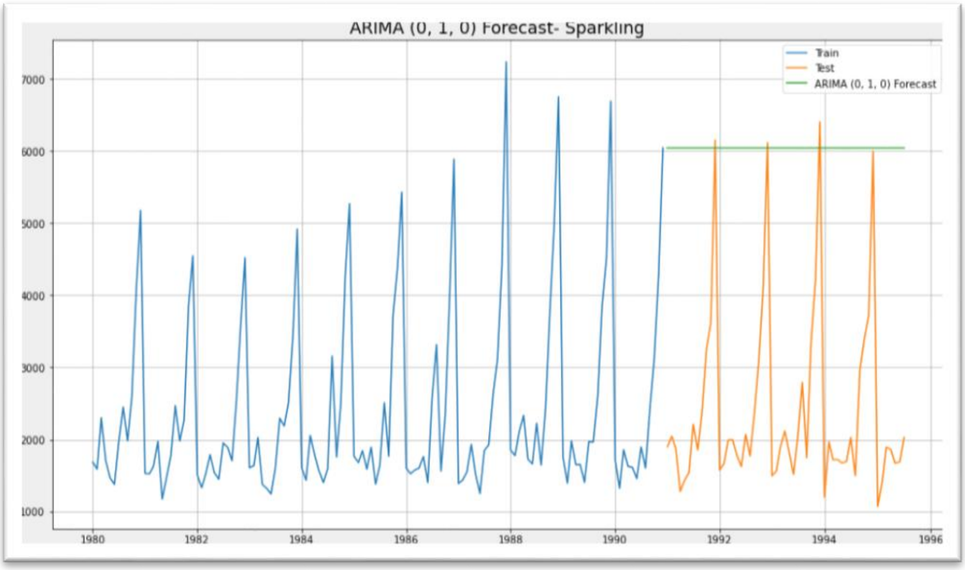
Graph 68: Diagnostic Plot

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

```
RMSE: 3864.2793518443914
MAPE: 201.32764950352743
```

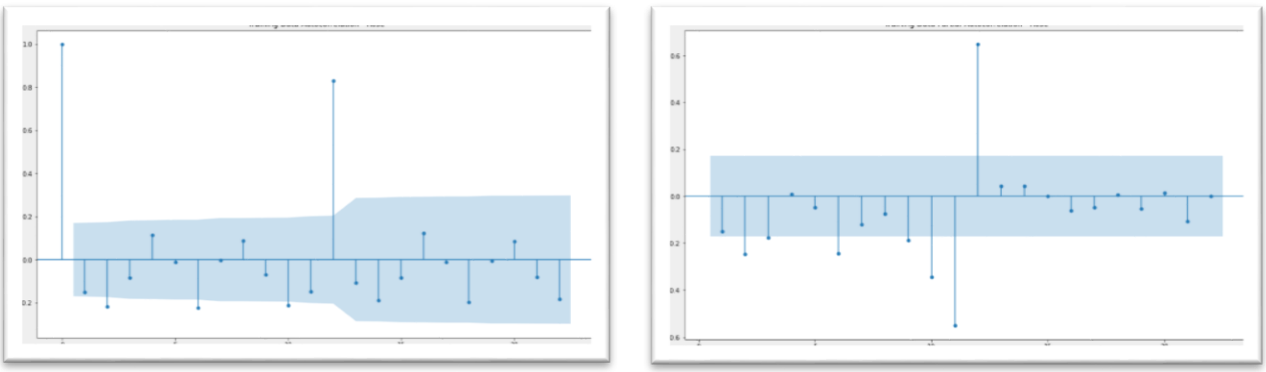
	RMSE	MAPE
ARIMA(2,1,2)	1299.980373	47.100017
ARIMA(0,1,0)	3864.279352	201.327650

Table 121: Model Accuracy



Graph 69: ARIMA Plot

Build a version of the SARIMA model for which the best parameters are selected by looking at the ACF and the PACF plots. - Seasonality at 12 - SPARKLING:



Graph 70: Data Autocorrelation

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                              Sparkling   No. Observations:                  132
Model:             SARIMAX(0, 1, 0)x(1, 1, [1], 12)   Log Likelihood                -811.162
Date:                                Fri, 25 Mar 2022   AIC                           1628.324
Time:                                        12:31:52   BIC                           1636.315
Sample:                                    01-01-1980   HQIC                          1631.563
                                         - 12-01-1990
Covariance Type:                                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.S.L12       0.1482      0.223      0.664      0.507      -0.289       0.586
ma.S.L12      -0.5732      0.217     -2.640      0.008      -0.999      -0.148
sigma2      2.577e+05   2.63e+04      9.806      0.000    2.06e+05    3.09e+05
===================================================================================
Ljung-Box (L1) (Q):                  13.54   Jarque-Bera (JB):                27.17
Prob(Q):                              0.00   Prob(JB):                         0.00
Heteroskedasticity (H):               0.73   Skew:                             0.59
Prob(H) (two-sided):                  0.36   Kurtosis:                         5.19
===================================================================================

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

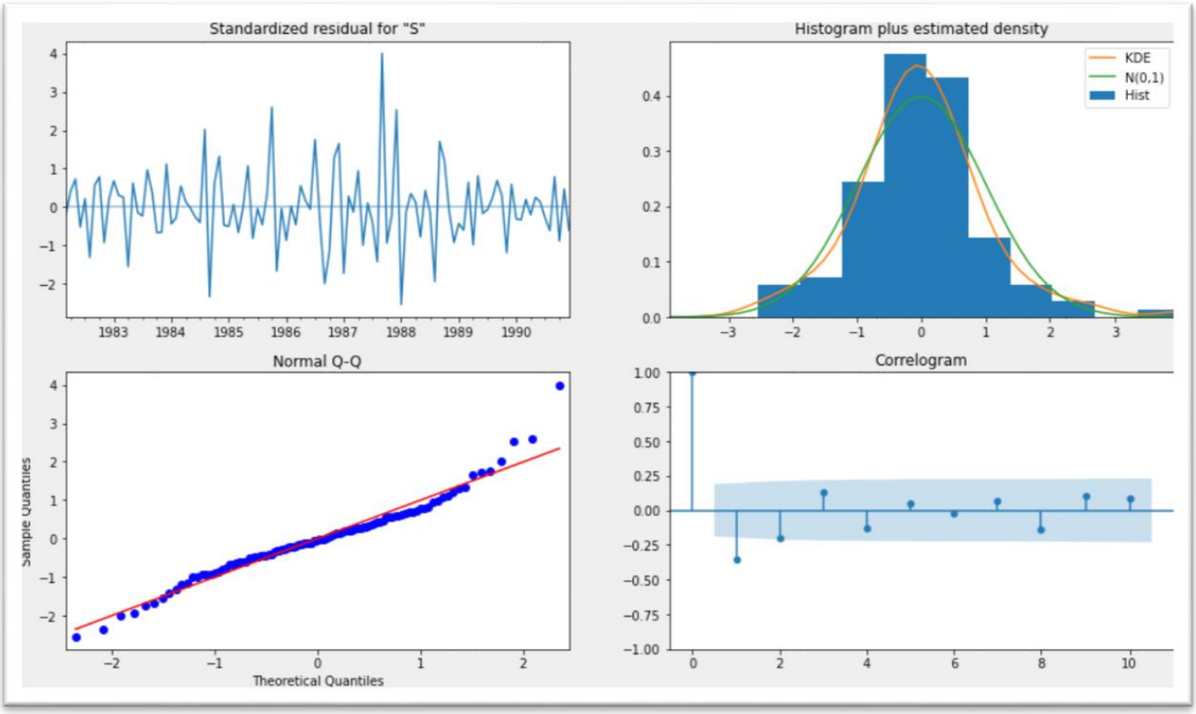
Table 122: Model Summary

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                              Sparkling   No. Observations:                  132
Model:          SARIMAX(0, 1, 0)x(2, 1, [1, 2], 12)   Log Likelihood                -722.996
Date:                                Fri, 25 Mar 2022   AIC                           1455.991
Time:                                        12:31:52   BIC                           1468.708
Sample:                                    01-01-1980   HQIC                          1461.128
                                         - 12-01-1990
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
===================================================================================

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

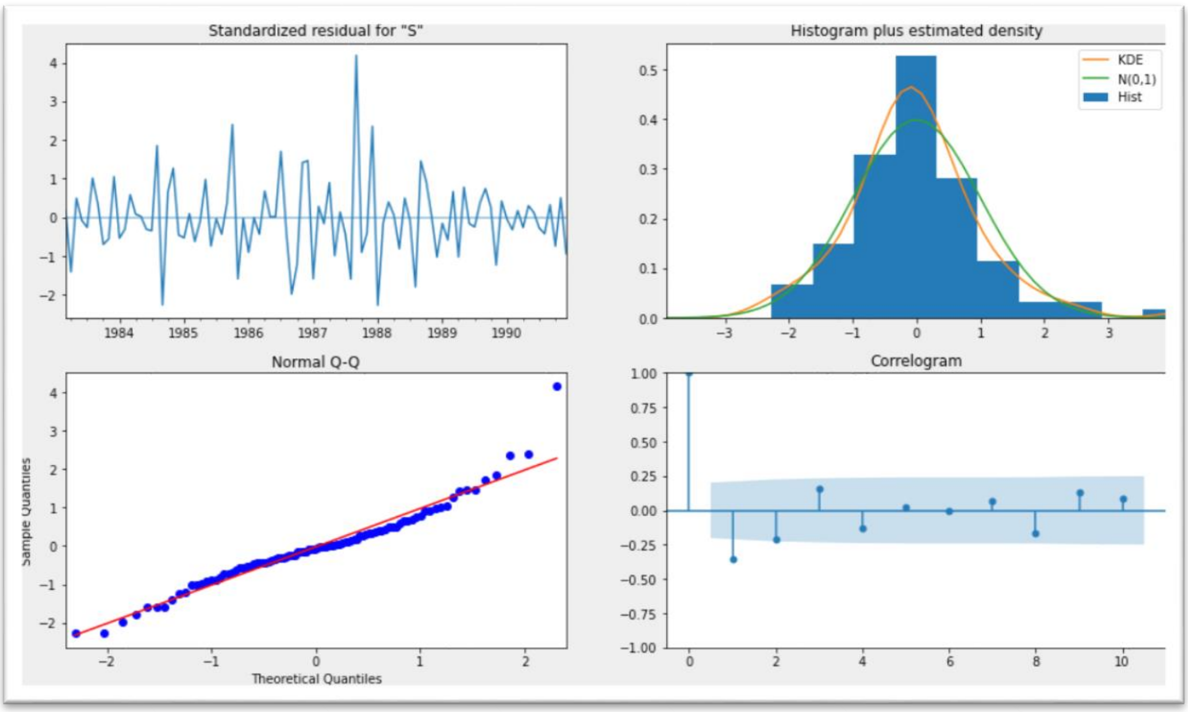
Table 123: Model Summary

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                              Sparkling   No. Observations:                  132
Model:          SARIMAX(0, 1, 0)x(3, 1, [1, 2], 12)   Log Likelihood                -638.304
Date:                                Fri, 25 Mar 2022   AIC                           1288.607
Time:                                        12:31:54   BIC                           1303.120
Sample:                                    01-01-1980   HQIC                          1294.438
                                         - 12-01-1990
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      -0.535      -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).
```

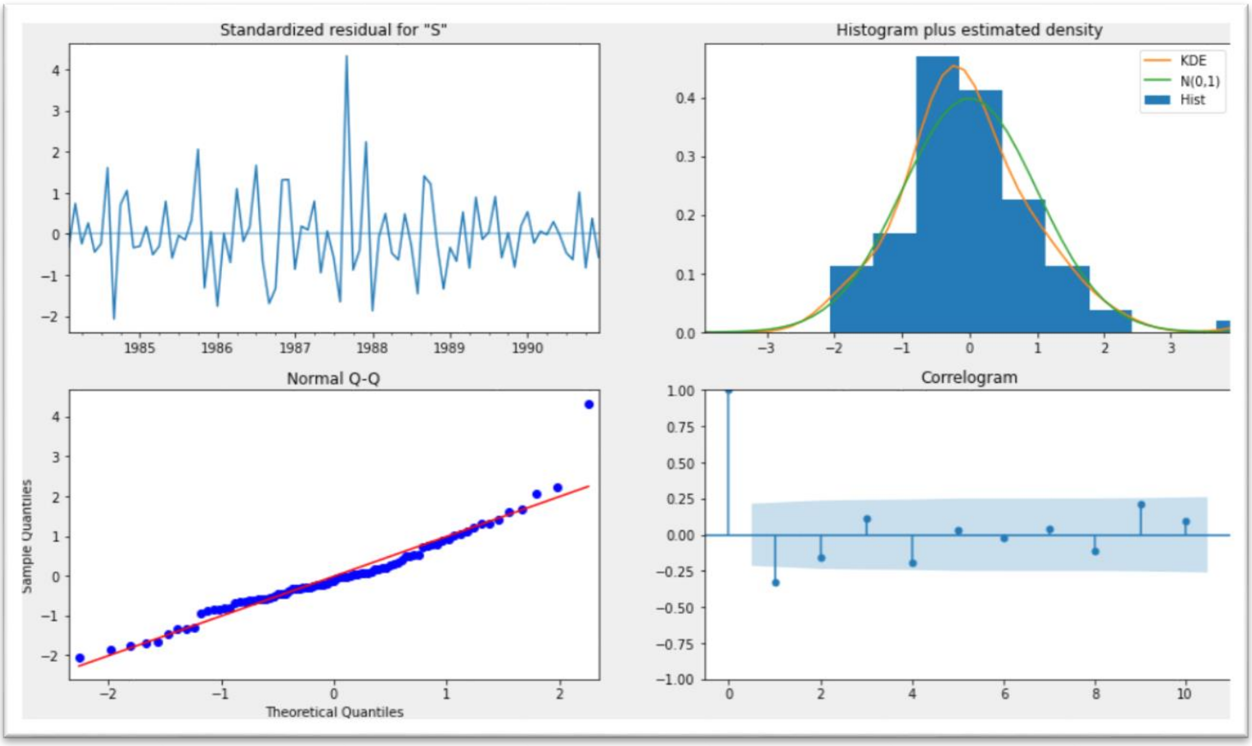
Table 124: Model Summary



Graph 71: Diagnostic Plot



Graph 72: Diagnostic Plot



Graph 73: Diagnostic Plot

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

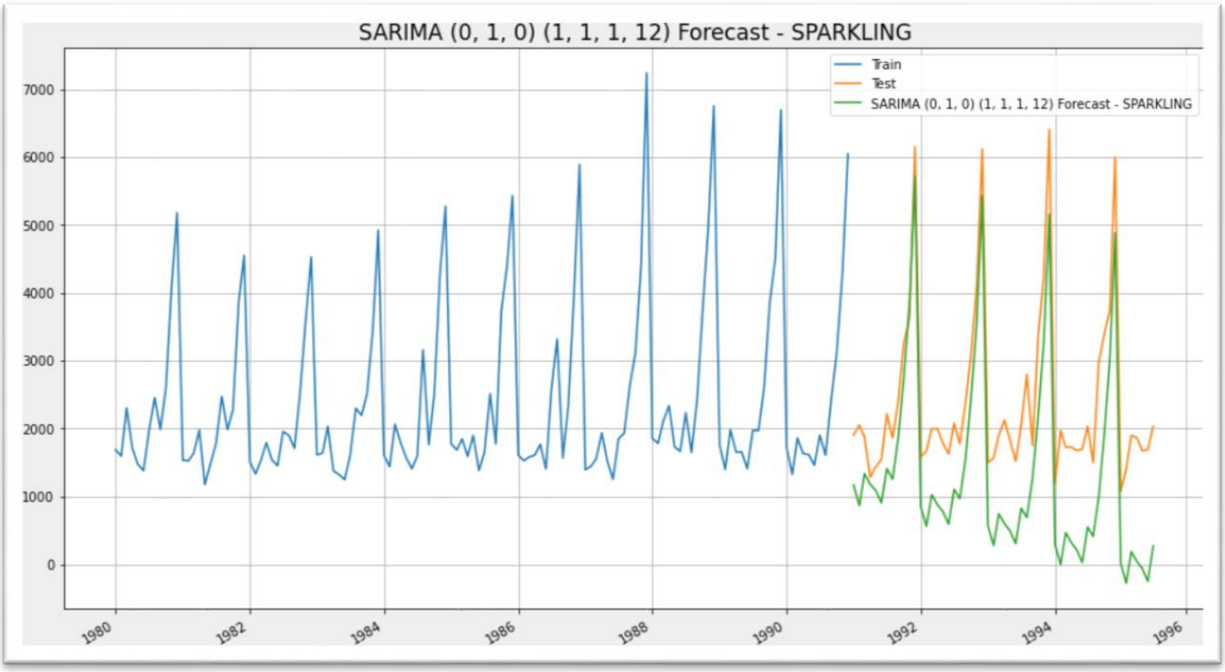
```
RMSE: 1189.8357829699708
MAPE: 54.87253569306072
```

```
RMSE: 1757.7268671724094
MAPE: 81.78523424489413
```

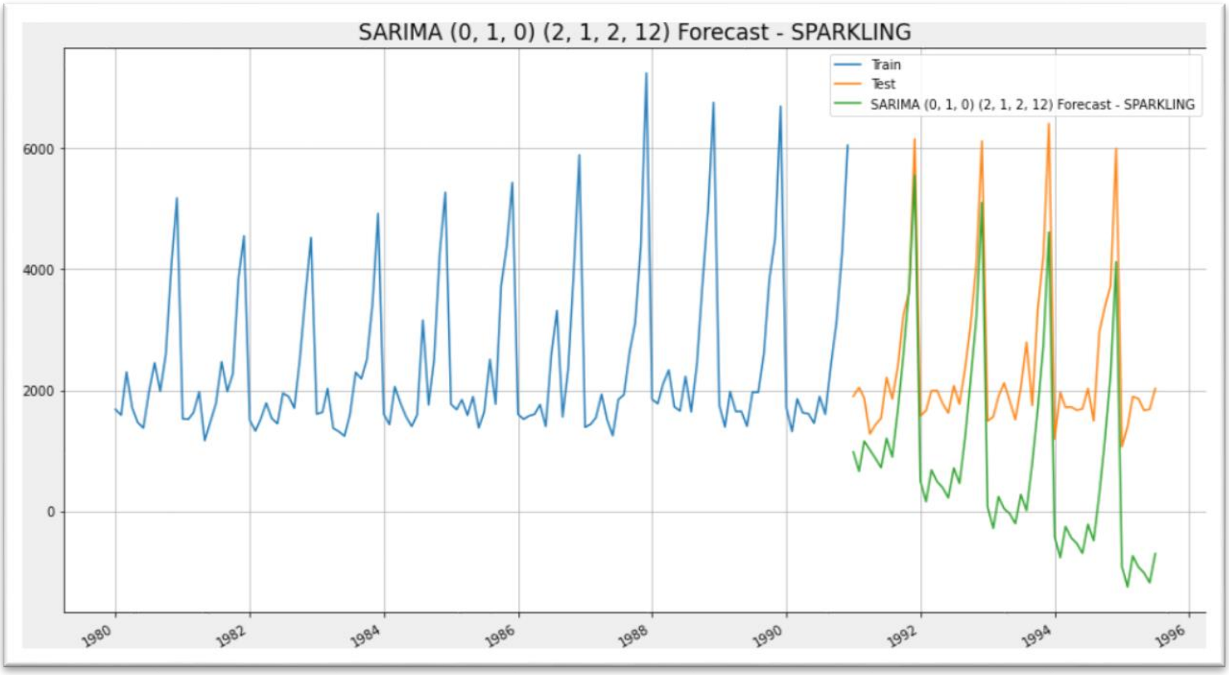
```
RMSE: 1551.6462725948265
MAPE: 71.56619627019045
```

	RMSE	MAPE
ARIMA(2,1,2)	1299.980373	47.100017
ARIMA(0,1,0)	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	601.244396	25.870721
SARIMA(0,1,0)(3,1,2,12)	1189.835783	54.872536
SARIMA(0,1,0)(2,1,2,12)	1757.726867	81.785234
SARIMA(0,1,0)(3,1,2,12)	1551.646273	71.566196

Table 125: Model Accuracy



Graph 74: SARIMA Plot



Graph 75: SARIMA Plot

Building the most optimum model on the Full Data:

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                                 Rose   No. Observations:                187
Model:             SARIMAX(2, 1, 2)x(2, 1, 2, 12)   Log Likelihood                -587.531
Date:                            Fri, 25 Mar 2022   AIC                           1193.062
Time:                                    12:32:00   BIC                           1219.976
Sample:                                01-01-1980   HQIC                          1203.997
                                     - 07-01-1995
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.8649      0.101     -8.572      0.000      -1.063      -0.667
ar.L2          0.0340      0.090      0.376      0.707      -0.143       0.211
ma.L1          0.0892   1649.995    5.4e-05      1.000   -3233.842    3234.020
ma.L2         -0.9108   1502.873     -0.001      1.000   -2946.488    2944.666
ar.S.L12       0.0719      0.166      0.434      0.664      -0.253       0.396
ar.S.L24      -0.0357      0.017     -2.045      0.041      -0.070      -0.001
ma.S.L12      -0.6869      0.222     -3.090      0.002      -1.123      -0.251
ma.S.L24      -0.0549      0.151     -0.365      0.715      -0.350       0.240
sigma2       158.9029   2.62e+05      0.001      1.000   -5.14e+05    5.14e+05
===================================================================================
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).
```

Table 126: Model Summary

Evaluate the model on the whole data and predict 12 months into the future:

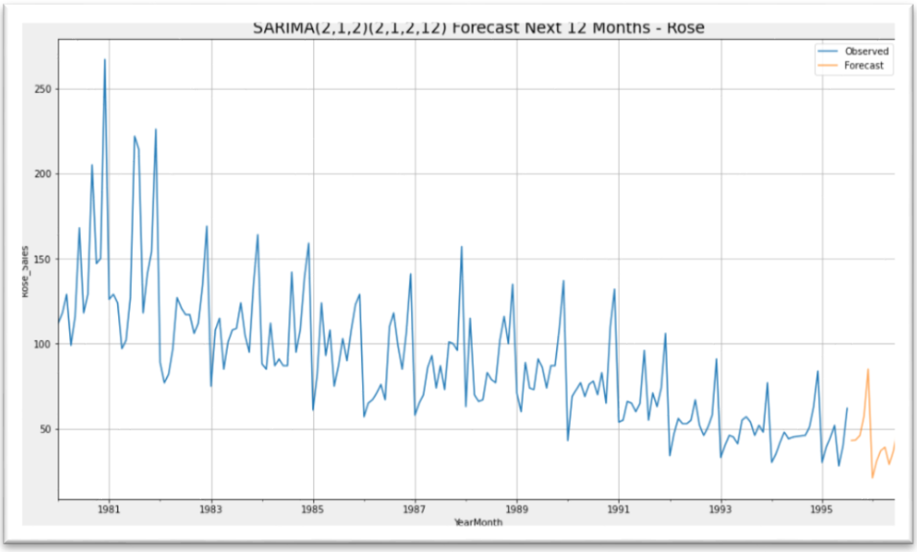
Rose	mean	mean_se	mean_ci_lower	mean_ci_upper
1995-08-01	43.082945	12.674005	18.242352	67.923539
1995-09-01	43.341394	12.959728	17.940794	68.741993
1995-10-01	45.826993	12.963722	20.418566	71.235421
1995-11-01	57.396972	13.124452	31.673519	83.120425
1995-12-01	85.074678	13.133279	59.333924	110.815431

	RMSE	MAPE
ARIMA(2,1,2)	1299.980373	47.100017
ARIMA(0,1,0)	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	601.244396	25.870721
SARIMA(0,1,0)(3,1,2,12)	1189.835783	54.872536
SARIMA(0,1,0)(2,1,2,12)	1757.726867	81.785234
SARIMA(0,1,0)(3,1,2,12)	1551.646273	71.566196

```
RMSE of the Full Model 35.910838112663996
```

	Test RMSE Rose	Test RMSE Sparkling	RMSE	MAPE
RegressionOnTime	15.268955	1389.135175	NaN	NaN
NaiveModel	79.718773	3864.279352	NaN	NaN
SimpleAverageModel	53.460570	1275.081804	NaN	NaN
2pointTrailingMovingAverage	11.529278	813.400684	NaN	NaN
4pointTrailingMovingAverage	14.451403	1156.589694	NaN	NaN
6pointTrailingMovingAverage	14.566327	1283.927428	NaN	NaN
9pointTrailingMovingAverage	14.727630	1346.278315	NaN	NaN
Simple Exponential Smoothing	36.796228	1338.012144	NaN	NaN
Double Exponential Smoothing	15.269328	3949.993290	NaN	NaN
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686	NaN	NaN
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.510170	NaN	NaN
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.660960	379.695686	NaN	NaN
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	26.295981	352.443335	NaN	NaN
ARIMA(2,1,2)	NaN	NaN	1299.980373	47.100017
ARIMA(0,1,0)	NaN	NaN	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	NaN	NaN	601.244396	25.870721
SARIMA(0,1,0)(3,1,2,12)	NaN	NaN	1189.835783	54.872536
SARIMA(0,1,0)(2,1,2,12)	NaN	NaN	1757.726867	81.785234
SARIMA(0,1,0)(3,1,2,12)	NaN	NaN	1551.646273	71.566196

Table 127: Model Accuracy



Graph 76: SARIMA Plot

- In all Manual methods, Best Model for Rose with Least RMSE

—-> SARIMA (2, 1, 2) (3, 1, 2, 12)

- In all Manual methods, Best Model for Sparkling with Least RMSE

—-> SARIMA (0, 1, 0) (1, 1, 1, 12)

Seasonal P and Q - it was difficult to gauge the correct values here as the data was not enough and cutoffs were not visible.

Hence, we tried multiple combinations of Seasonal P and Q as given above.

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.

Evaluation of Models on whole Data:

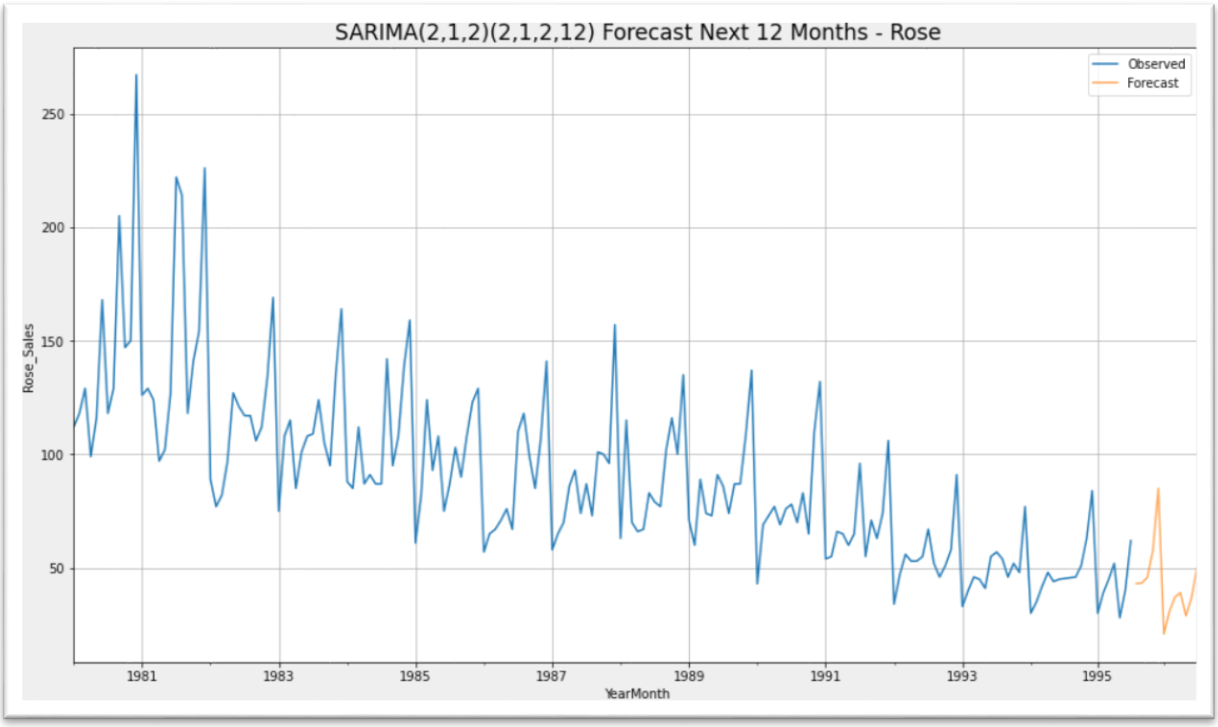
Rose	mean	mean_se	mean_ci_lower	mean_ci_upper
1995-08-01	43.082945	12.674005	18.242352	67.923539
1995-09-01	43.341394	12.959728	17.940794	68.741993
1995-10-01	45.826993	12.963722	20.418566	71.235421
1995-11-01	57.396972	13.124452	31.673519	83.120425
1995-12-01	85.074678	13.133279	59.333924	110.815431

	RMSE	MAPE
ARIMA(2,1,2)	1299.980373	47.100017
ARIMA(0,1,0)	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	601.244396	25.870721
SARIMA(0,1,0)(3,1,2,12)	1189.835783	54.872536
SARIMA(0,1,0)(2,1,2,12)	1757.726867	81.785234
SARIMA(0,1,0)(3,1,2,12)	1551.646273	71.566196

```
RMSE of the Full Model 35.910838112663996
```

	Test RMSE Rose	Test RMSE Sparkling	RMSE	MAPE
RegressionOnTime	15.268955	1389.135175	NaN	NaN
NaiveModel	79.718773	3864.279352	NaN	NaN
SimpleAverageModel	53.460570	1275.081804	NaN	NaN
2pointTrailingMovingAverage	11.529278	813.400684	NaN	NaN
4pointTrailingMovingAverage	14.451403	1156.589694	NaN	NaN
6pointTrailingMovingAverage	14.566327	1283.927428	NaN	NaN
9pointTrailingMovingAverage	14.727630	1346.278315	NaN	NaN
Simple Exponential Smoothing	36.796228	1338.012144	NaN	NaN
Double Exponential Smoothing	15.269328	3949.993290	NaN	NaN
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686	NaN	NaN
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.510170	NaN	NaN
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.660960	379.695686	NaN	NaN
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	26.295981	352.443335	NaN	NaN
ARIMA(2,1,2)	NaN	NaN	1299.980373	47.100017
ARIMA(0,1,0)	NaN	NaN	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	NaN	NaN	601.244396	25.870721
SARIMA(0,1,0)(3,1,2,12)	NaN	NaN	1189.835783	54.872536
SARIMA(0,1,0)(2,1,2,12)	NaN	NaN	1757.726867	81.785234
SARIMA(0,1,0)(3,1,2,12)	NaN	NaN	1551.646273	71.566196

Table 128: Model Accuracy



Graph 77: SARIMA Plot

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.

Best Models as per the Least RMSE on ROSE Test set ——>
- 2 Pt Trailing Moving Average
- Triple Exponential Smoothing (Additive Seasonality)

Best Model as per the Least RMSE on SPARKLING Test set ——>
- Triple Exponential Smoothing (Multiplicative Season, Damped Trend) ,
- $\alpha = 0.11107$ $\beta = 0.03702$ $\gamma = 0.3950$

Building the most optimum model on the Full Data:

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                               Rose   No. Observations:                  187
Model:             SARIMAX(2, 1, 2)x(2, 1, 2, 12)   Log Likelihood                -587.531
Date:                            Fri, 25 Mar 2022   AIC                           1193.062
Time:                                    12:32:00   BIC                           1219.976
Sample:                                01-01-1980   HQIC                          1203.997
                                     - 07-01-1995
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.8649      0.101     -8.572      0.000      -1.063      -0.667
ar.L2          0.0340      0.090      0.376      0.707      -0.143       0.211
ma.L1          0.0892   1649.995   5.4e-05      1.000   -3233.842    3234.020
ma.L2         -0.9108   1502.873     -0.001      1.000   -2946.488    2944.666
ar.S.L12       0.0719      0.166      0.434      0.664      -0.253       0.396
ar.S.L24      -0.0357      0.017     -2.045      0.041      -0.070      -0.001
ma.S.L12      -0.6869      0.222     -3.090      0.002      -1.123      -0.251
ma.S.L24      -0.0549      0.151     -0.365      0.715      -0.350       0.240
sigma2       158.9029   2.62e+05      0.001      1.000   -5.14e+05    5.14e+05
===================================================================================
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).
```

Table 129: Model Summary

Evaluate the model on the whole data and predict 12 months into the future:

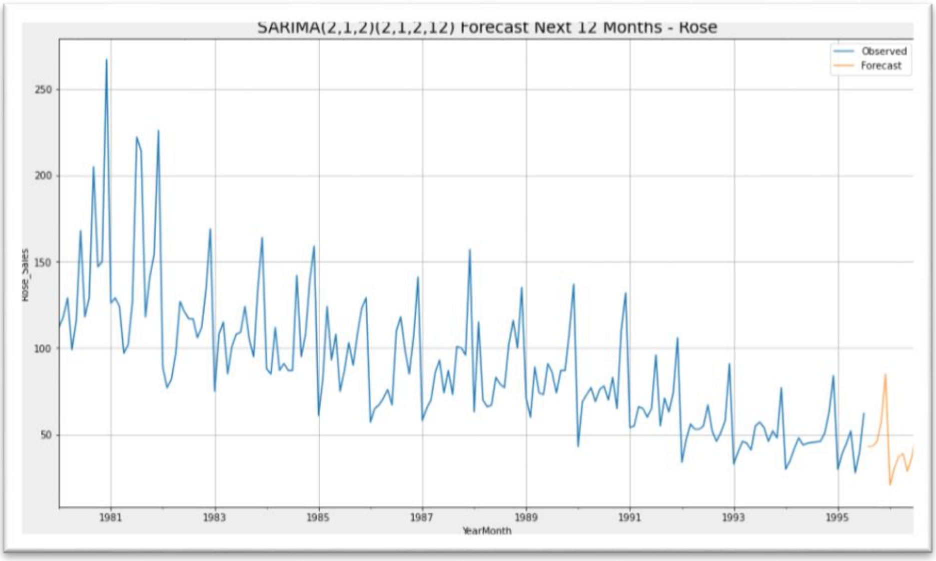
Rose	mean	mean_se	mean_ci_lower	mean_ci_upper
1995-08-01	43.082945	12.674005	18.242352	67.923539
1995-09-01	43.341394	12.959728	17.940794	68.741993
1995-10-01	45.826993	12.963722	20.418566	71.235421
1995-11-01	57.396972	13.124452	31.673519	83.120425
1995-12-01	85.074678	13.133279	59.333924	110.815431

```
RMSE of the Full Model 35.910838112663996
```

	Test RMSE Rose	Test RMSE Sparkling	RMSE	MAPE
RegressionOnTime	15.268955	1389.135175	NaN	NaN
NaiveModel	79.718773	3864.279352	NaN	NaN
SimpleAverageModel	53.460570	1275.081804	NaN	NaN
2pointTrailingMovingAverage	11.529278	813.400684	NaN	NaN
4pointTrailingMovingAverage	14.451403	1156.589694	NaN	NaN
6pointTrailingMovingAverage	14.566327	1283.927428	NaN	NaN
9pointTrailingMovingAverage	14.727630	1346.278315	NaN	NaN
Simple Exponential Smoothing	36.796228	1338.012144	NaN	NaN
Double Exponential Smoothing	15.269328	3949.993290	NaN	NaN
Triple Exponential Smoothing (Additive Season)	14.265713	379.695686	NaN	NaN
Triple Exponential Smoothing (Multiplicative Season)	20.190998	406.510170	NaN	NaN
Triple Exponential Smoothing (Additive Season, Damped Trend)	25.660960	379.695686	NaN	NaN
Triple Exponential Smoothing (Multiplicative Season, Damped Trend)	26.295981	352.443335	NaN	NaN
ARIMA(2,1,2)	NaN	NaN	1299.980373	47.100017
ARIMA(0,1,0)	NaN	NaN	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	NaN	NaN	601.244396	25.870721
SARIMA(0,1,0)(3,1,2,12)	NaN	NaN	1189.835783	54.872536
SARIMA(0,1,0)(2,1,2,12)	NaN	NaN	1757.726867	81.785234
SARIMA(0,1,0)(3,1,2,12)	NaN	NaN	1551.646273	71.566196

	RMSE	MAPE
ARIMA(2,1,2)	1299.980373	47.100017
ARIMA(0,1,0)	3864.279352	201.327650
SARIMA(3,1,1)(3,0,2,12)	601.244396	25.870721
SARIMA(0,1,0)(3,1,2,12)	1189.835783	54.872536
SARIMA(0,1,0)(2,1,2,12)	1757.726867	81.785234
SARIMA(0,1,0)(3,1,2,12)	1551.646273	71.566196

Table 130: Model Accuracy



Graph 78: SARIMA Plot

Best Model for ROSE with Least RMSE - 2 Pt Moving Average:

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

Table 131: Rose data set

Trailing moving averages:

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

YearMonth	Rose	Trailing_2
1995-03-01	45.0	42.0
1995-04-01	52.0	48.5
1995-05-01	28.0	40.0
1995-06-01	40.0	34.0
1995-07-01	62.0	51.0

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

	Rose	Trailing_2	forecast_12
1996-03-01	NaN	NaN	NaN
1996-04-01	NaN	NaN	NaN
1996-05-01	NaN	NaN	NaN
1996-06-01	NaN	NaN	NaN
1996-07-01	NaN	NaN	NaN

```
(199, 3)
```

	Rose	Trailing_2	forecast_12
1980-01-01	112.0	0.0	0.0
1980-02-01	118.0	115.0	115.0
1980-03-01	129.0	123.5	123.5
1980-04-01	99.0	114.0	114.0
1980-05-01	116.0	107.5	107.5
...	...	...	...
1996-03-01	0.0	0.0	0.0
1996-04-01	0.0	0.0	0.0
1996-05-01	0.0	0.0	0.0
1996-06-01	0.0	0.0	0.0
1996-07-01	0.0	0.0	0.0

199 rows × 3 columns

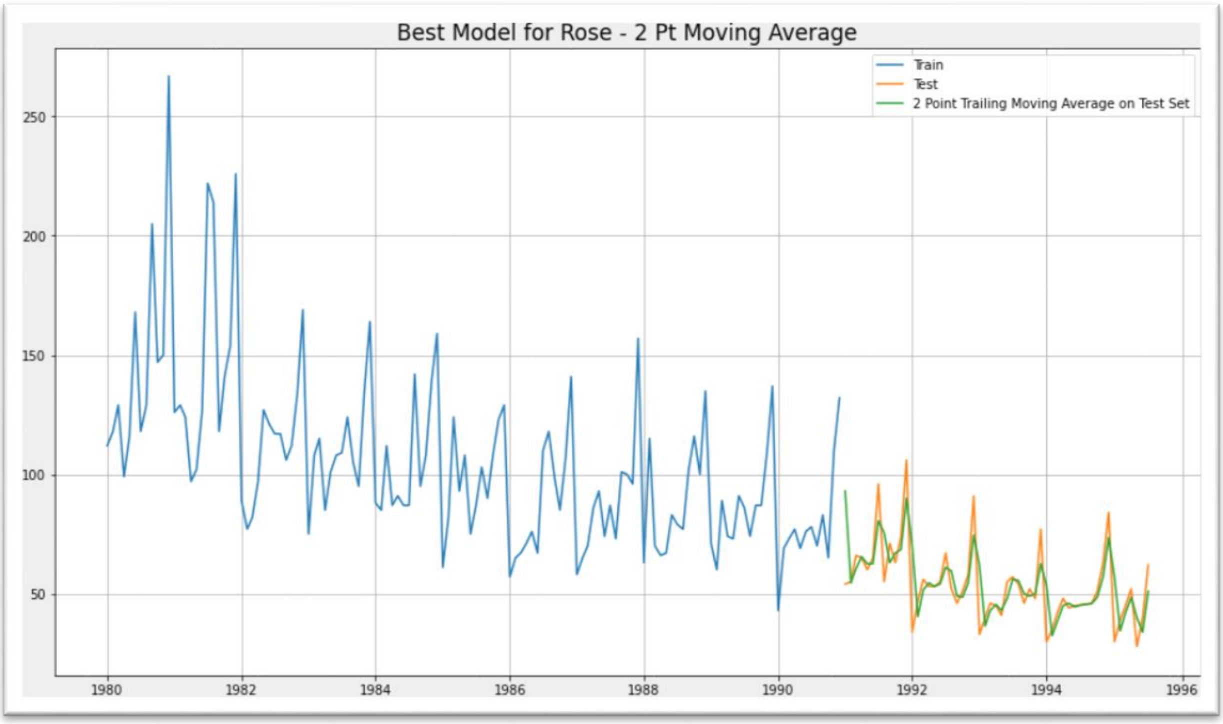
```
1980-01-01    112.0
1980-02-01    118.0
Name: Rose, dtype: float64
```

Table 132: Trailing Moving Averages

[115.0,	166.0,	100.0,	123.5,	114.0,	92.5,	87.0,	65.5,	54.5,
123.5,	129.5,	93.0,	149.0,	108.5,	68.0,	98.0,	62.5,	74.5,
114.0,	147.5,	104.5,	110.0,	92.0,	66.5,	123.0,	62.5,	62.0,
107.5,	190.0,	108.5,	71.5,	96.0,	75.0,	90.0,	80.5,	36.5,
142.0,	157.5,	116.5,	103.0,	124.0,	81.0,	56.0,	75.5,	43.0,
143.0,	83.0,	114.5,	108.5,	99.5,	78.0,	71.0,	63.0,	45.5,
123.5,	79.5,	100.0,	100.5,	61.5,	89.5,	75.0,	67.0,	43.0,
167.0,	89.5,	115.0,	91.5,	67.5,	109.0,	73.0,	68.5,	48.0,
176.0,	112.0,	149.5,	81.0,	78.0,	108.0,	72.5,	90.0,	56.0,
148.5,	124.0,	126.0,	95.0,	89.5,	117.5,	77.0,	70.0,	55.5,
208.5,	119.0,	86.5,	96.5,	83.5,	103.0,	74.0,	40.5,	50.0,
196.5,	117.0,	98.5,	99.0,	80.5,	65.5,	76.5,	51.5,	49.0,
127.5,	111.5,	99.5,	115.5,	80.0,	74.5,	74.0,	54.5,	50.0,
126.5,	109.0,	89.0,	126.0,	87.0,	81.5,	87.5,	53.0,	62.5,
110.5,	123.0,	89.0,	93.0,	100.5,	73.5,	121.0,	54.0,	53.5,
99.5,	151.5,	87.0,	61.0,	98.0,	82.0,	93.0,	61.0,	32.5,
114.5,	122.0,	114.5,	66.0,	126.5,	88.5,	54.5,	59.5,	38.5,
174.5,	91.5,	118.5,	69.0,	110.0,	80.0,	60.5,	49.0,	45.0,
218.0,	111.5,	101.5,	73.5,	89.0,	80.5,	65.5,	48.5,	46.0,

```
44.5,
45.166666666666667,
45.5,
45.833333333333333,
48.5,
57.0,
73.5,
57.0,
34.5,
42.0,
48.5,
40.0,
34.0,
51.0,
nan,
nan,
nan,
nan,
nan,
```

```
nan,
nan,
nan,
nan,
nan,
nan,
nan,
nan,
nan,
nan,
nan,
nan,
nan]
```



Graph 79: Best Model for Rose

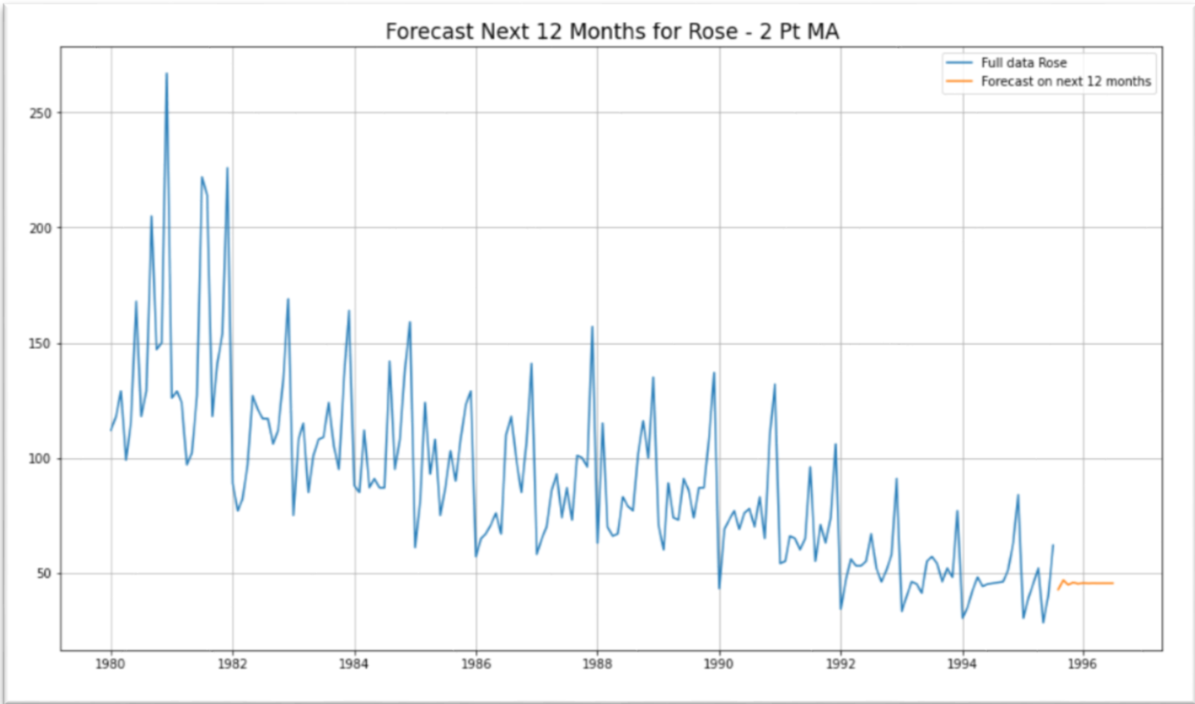
FORECAST ON NEXT 12 MONTHS - ROSE:

(Using 2 Pt Moving Average Model):

YearMonth	Rose	forecast
1980-01-01	112.0	NaN
1980-02-01	118.0	115.0
1980-03-01	129.0	123.5
1980-04-01	99.0	114.0
1980-05-01	116.0	107.5

YearMonth	Rose	forecast
1996-03-01	NaN	45.355469
1996-04-01	NaN	45.322266
1996-05-01	NaN	45.338867
1996-06-01	NaN	45.330566
1996-07-01	NaN	45.334717

Table 133: 2 Pt Moving Average Model



Graph 80: Forecast for next 12 months

Holt-Winters – ETS (A, A, M) - Best Model for Sparkling till Now ETS (A, Ad, M) model - Taking MULTIPLICATIVE SEASONALITY USING DAMPED TREND

```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.07571766443273568, 'smoothing_trend': 0.06490063802484945, 'smoothing_seasonal': 0.27386369588276294, 'da
mping_trend': 0.9899999423037755, 'initial_level': 2356.5390698560527, 'initial_trend': -9.255718981252874, 'initial_seasons':
array([0.7305313 , 0.70065313, 0.89858585, 0.8161961 , 0.6796096 ,
       0.66533713, 0.88873825, 1.1467924 , 0.93452548, 1.2605895 ,
       1.89257504, 2.4489025 ]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

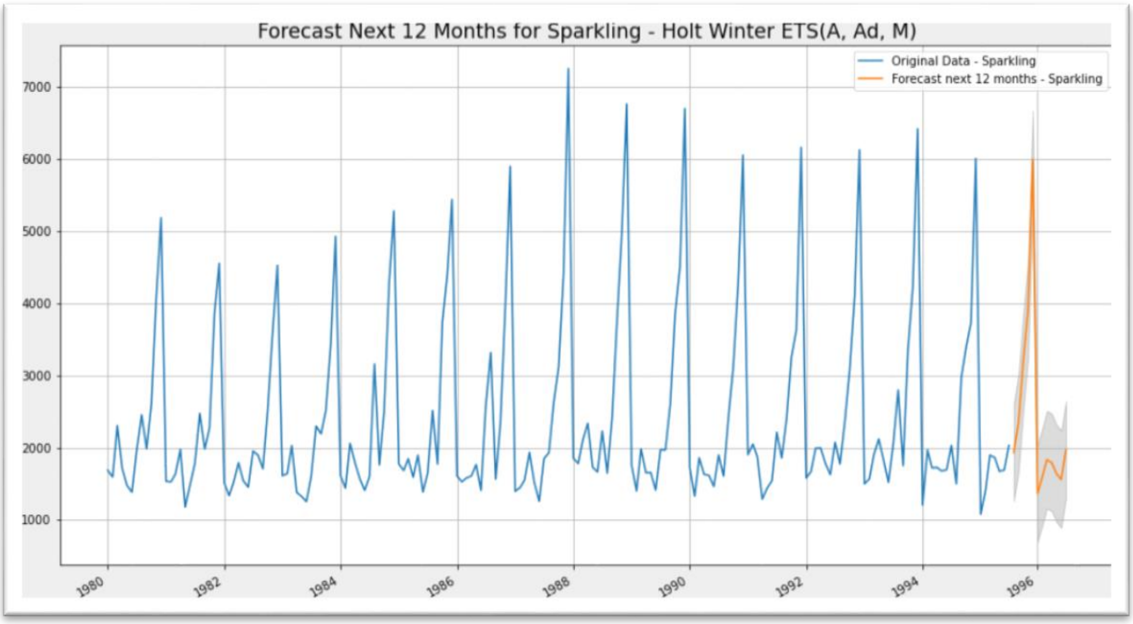
Table 134: Checking for Parameters

```
1995-08-01    1931.624740
1995-09-01    2352.139809
1995-10-01    3179.864641
1995-11-01    3918.101027
1995-12-01    5985.869206
1996-01-01    1357.348823
1996-02-01    1599.120616
1996-03-01    1830.277780
1996-04-01    1790.979274
1996-05-01    1641.918645
1996-06-01    1556.396563
1996-07-01    1965.855939
Freq: MS, dtype: float64
```

Table 135: AIC Values

	lower_CI	prediction	upper_ci
1995-08-01	1251.886902	1931.624740	2611.362577
1995-09-01	1672.401972	2352.139809	3031.877647
1995-10-01	2500.126803	3179.864641	3859.602478
1995-11-01	3238.363189	3918.101027	4597.838864
1995-12-01	5306.131368	5985.869206	6665.607043

Table 136: Prediction Values



Graph 81: Forecast for next 12 Months

Building the second most optimum model on ROSE - TES ETS (A, A, A):

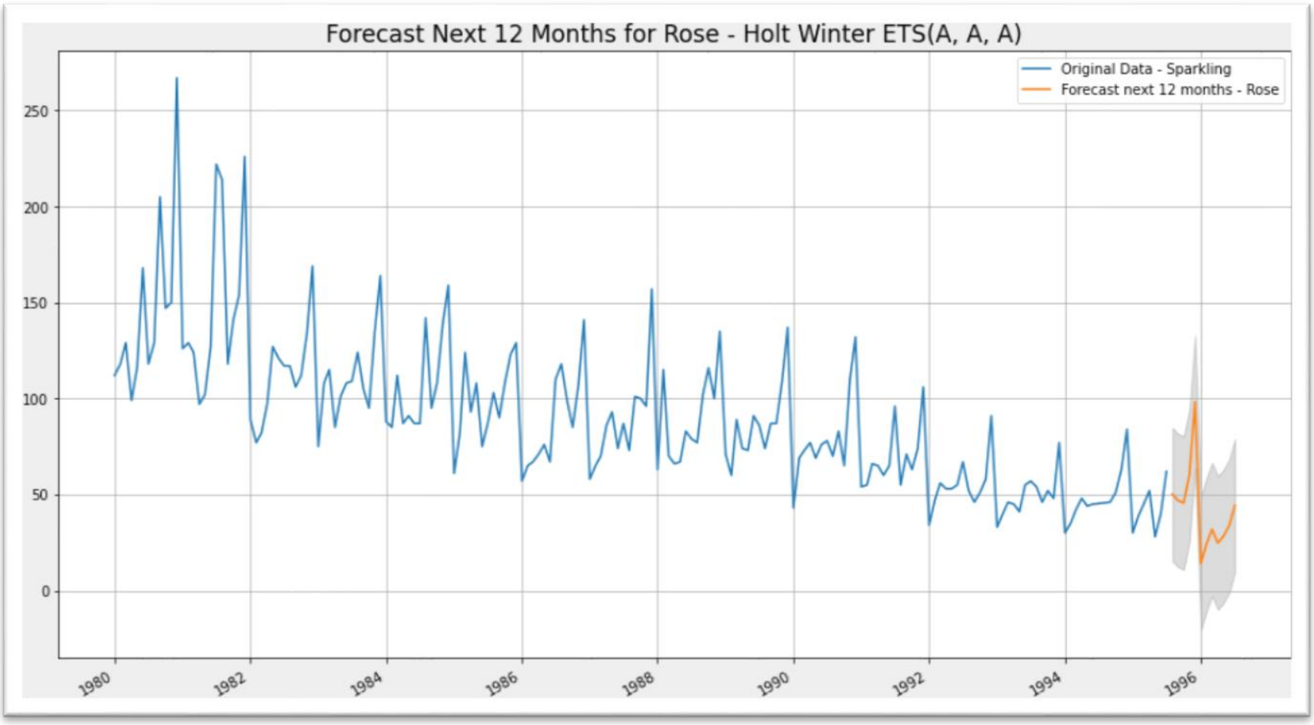
```
==Holt Winters model Exponential Smoothing Estimated Parameters ==

{'smoothing_level': 0.09936159631278565, 'smoothing_trend': 0.0002309325981159989, 'smoothing_seasonal': 0.0023514362014439823,
'damping_trend': nan, 'initial_level': 145.26022085722158, 'initial_trend': -0.5364025702831309, 'initial_seasons': array([-28.
09736871, -17.27252389,  -9.16094678, -15.76250213,
       -11.75395333,  -5.66922986,   5.40671588,   5.33567347,
         2.61441889,   1.84378355,  16.95414129,  55.57840214]), 'use_boxcox': False, 'lamda': None, 'remove_bias': False}
```

Table 137: Checking for Parameter

```
1995-08-01    50.159650
1995-09-01    46.908166
1995-10-01    45.603978
1995-11-01    60.177419
1995-12-01    98.266648
1996-01-01    14.054484
1996-02-01    24.343651
1996-03-01    31.917314
1996-04-01    24.780731
1996-05-01    28.244507
1996-06-01    33.790540
1996-07-01    44.305460
Freq: MS, dtype: float64
```

	lower_CI	prediction	upper_ci
1995-08-01	15.410335	50.159650	84.908965
1995-09-01	12.158851	46.908166	81.657481
1995-10-01	10.854663	45.603978	80.353293
1995-11-01	25.428103	60.177419	94.926734
1995-12-01	63.517333	98.266648	133.015964

Table 138: AIC and Prediction Values



Graph 82: Forecast for next 12 Months

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

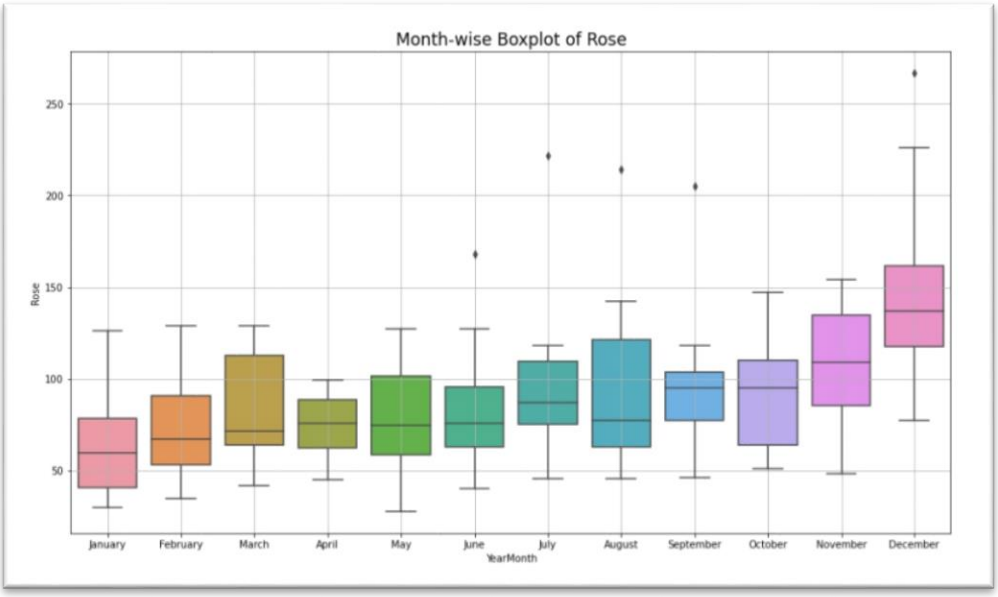
Rose Wine Sales - Comments:

- Rose wine shows a clear trend of declining sales since 1980
  - This shows decline in popularity of this variant of wine

- Also, there is a clear spike in sales seen in the last quarter of every year from Oct to Dec
  - This might be due to the Holiday season in this period
  - Highest peak in sales is seen in Dec every year

- There is also an instant crashing slump in sales in the first quarter of every year from Jan
  - This might be due to the after effect or hangover of Holidays

- Sales slowly pick up only after May-June



Graph 83: Box Plot

Rose Wine Sales - Forecast Models:

- Top 2 best models as per lowest Test RMSE were found to be:
  -2 Pt Moving Average and Holt-Winters - Additive Seasonality & Trend

- 2 Pt Moving Average model, when used for forecasting do not seem to give good predictions. Forecast values level out after a few iterations

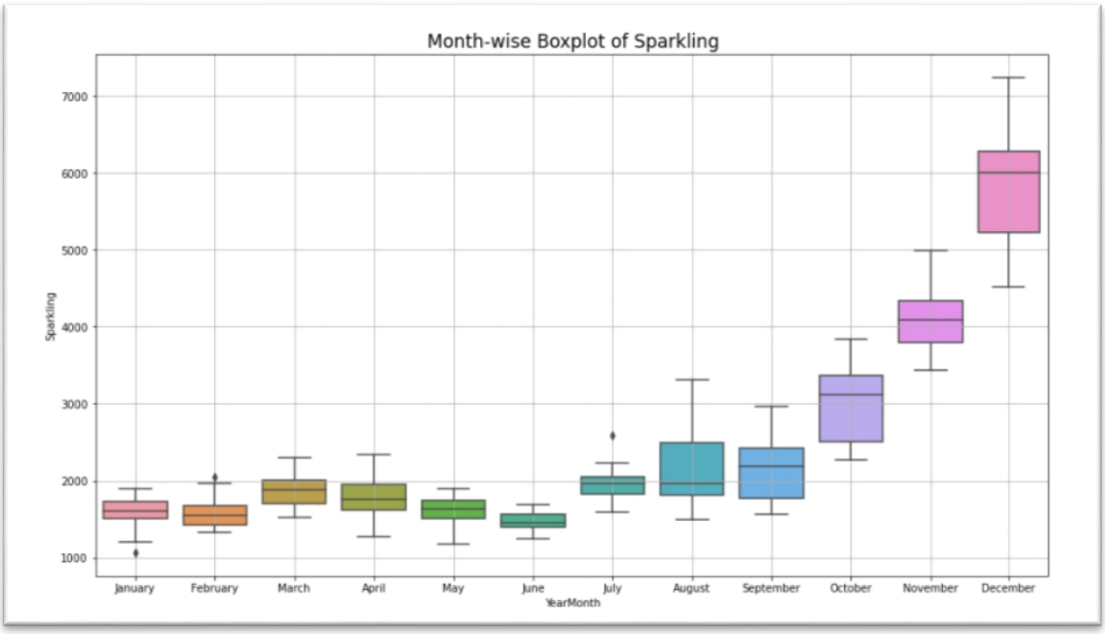
- Holt-Winters seems to give a consistent forecast with respect to the data
- Hence, for final forecast of Rose Wine Sales - we choose Holt-Winters

Rose Wine Sales - Suggestions:

- Firstly, Holiday season is around the corner and forecast shows increasing sales and sharp peak in Dec. Hence, Company should stock up
- But Declining sales of Rose Wine over the long period should be investigated with more data crunching
- Company can rebrand its Rose variant along-with a new Wine-master
- Company should take advantage of the oncoming spike from Aug-Oct by introducing aggressive offers and Ad campaigns.
- This will entice first time Wine drinkers and fence sitters (who don't have specific loyalties to any particular brand)
- Still if there is no significant upward trend in sales by this Dec, then Company has 2 options - invest in R&D or think of discontinuing this variant and come up with something completely new

Sparkling Wine Sales - Comments:

- Sparkling wine sales don't show any upward or downward trend
  - This shows flat sales over long term range



Graph 84: Box Plot

- Also, there is very high spike in sales seen in the last quarter of every year from Oct to Dec
  - This might be due to the Holiday season in this period
  - Highest peak in sales is seen in Dec every year
  - Dec sales are almost 3 times of Sep sales

- Similar to Rose Wine, even in Sparkling sales, an instant crashing slump is seen in the first quarter of every year from Jan
  - This might be due to the after effect or hangover of Holidays

- Sales slowly pick up only from Jul-Aug

Sparkling Wine Sales - Forecast Models:

- Triple Exponential Smoothing - Holt-Winters Models perform the best on Sparkling datasets, considering the least RMSE on Test data

- There has been incremental improvements in Test RMSE with each tuning of parameters

- Finally, for forecast of Sparkling Wine Sales - we choose Holt-Winters with Multiplicative Seasonality and Additive Damped Trend

Sparkling Wine Sales - Suggestions:

- Even for Sparkling, Holiday season is around the corner and forecast shows increasing sales and sharp peak in Dec. Hence, Company should stock up

- Sparkling wine has great holiday sales, so this shows popularity.

- So, no need to introduce any offers here but hammering Ads are suggested in these times of Oct-Dec. This will drive sales even further.

- Sparkling wines are generally associated with celebrations and mainly to burst open.

- A special designer bottle can be introduced at a cheaper price just for bursting. This will maximize profits

- Year on Year sales do not show any significant increase or decrease

- Though, Holiday spikes are extreme, but general Year on Year sales need to be investigated more. Early period from Jan should be used to do this deep dive