Sriram Vasan

Business Report -Time series Forecast

Problem 1:

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

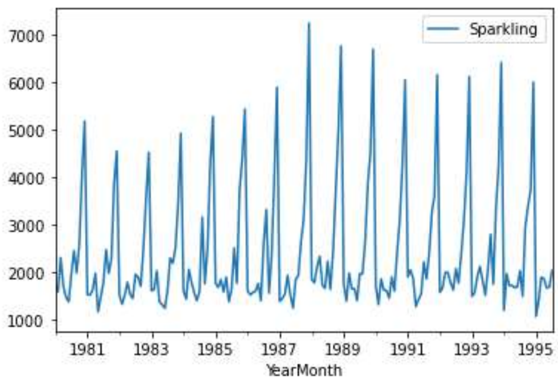
1.Read the data as an appropriate Time Series data and plot the data.

Head for sparkling and rose as appropriate time series data

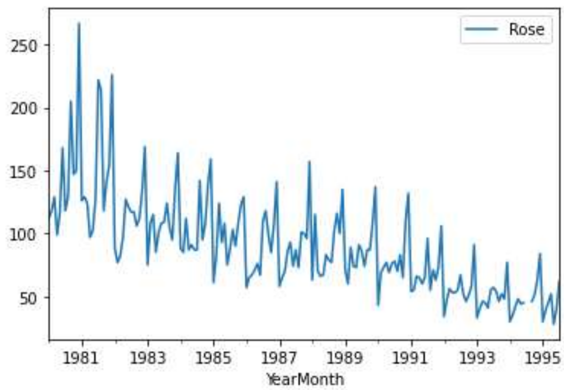
YearMonth	Sparkling
1980-01-01	1686
1980-02-01	1591
1980-03-01	2304
1980-04-01	1712
1980-05-01	1471

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

Sparkling



Rose



From the above graphs we can see that the data set rose has some null values

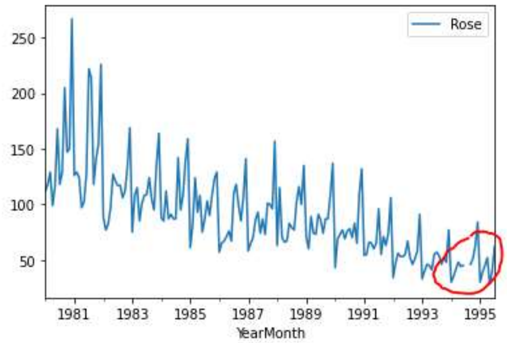
By performing interpolation for the specific months, we can impute the values.

YearMonth	Rose
1994-01-01	30.0
1994-02-01	35.0
1994-03-01	42.0
1994-04-01	48.0
1994-05-01	44.0
1994-06-01	45.0
1994-07-01	NaN
1994-08-01	NaN
1994-09-01	46.0
1994-10-01	51.0
1994-11-01	63.0
1994-12-01	84.0

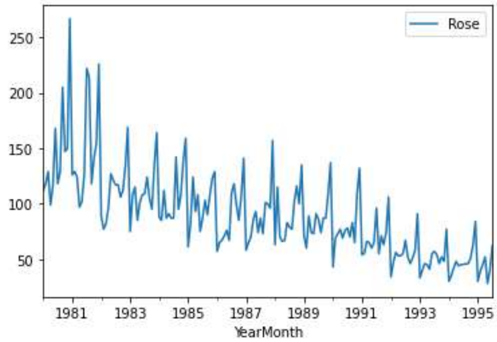
YearMonth	Rose
1994-01-01	30.000000
1994-02-01	35.000000
1994-03-01	42.000000
1994-04-01	48.000000
1994-05-01	44.000000
1994-06-01	45.000000
1994-07-01	45.333333
1994-08-01	45.666667
1994-09-01	46.000000
1994-10-01	51.000000
1994-11-01	63.000000
1994-12-01	84.000000

Graph of the interpolated view

Before interpolation

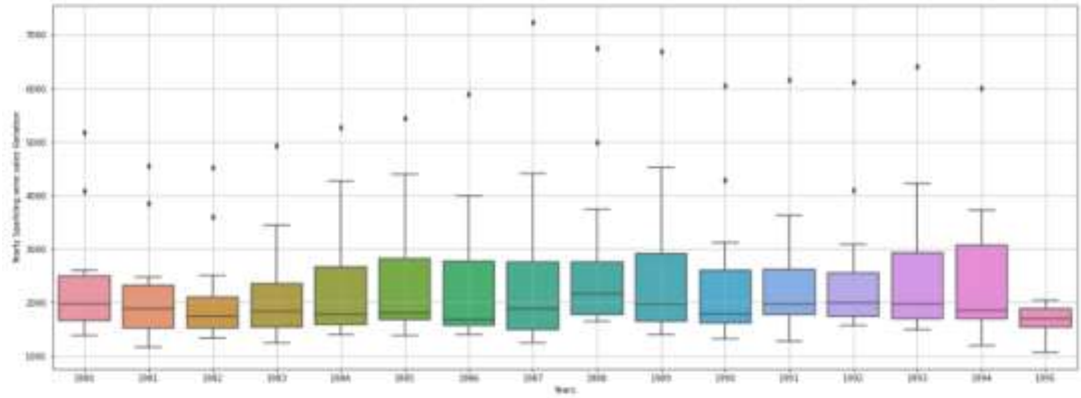


After interpolation

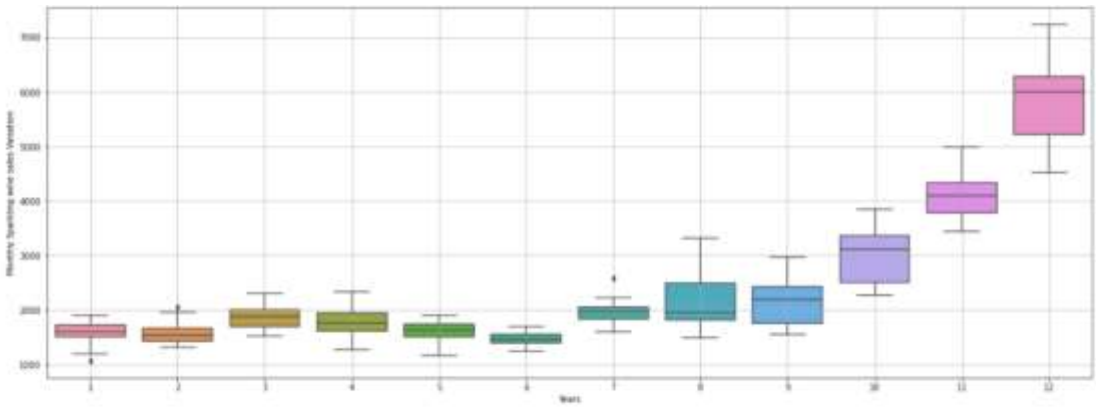


2.Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.

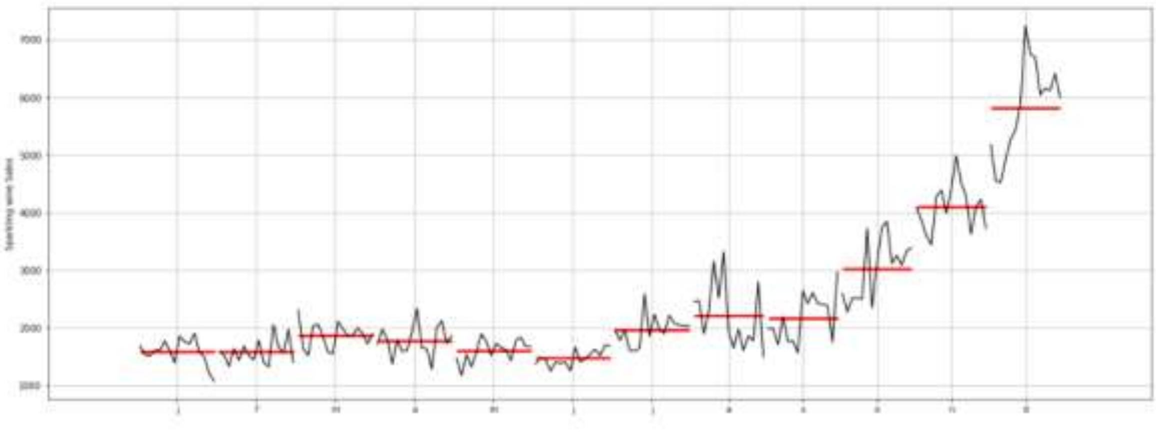
Sparkling yearly sales variation



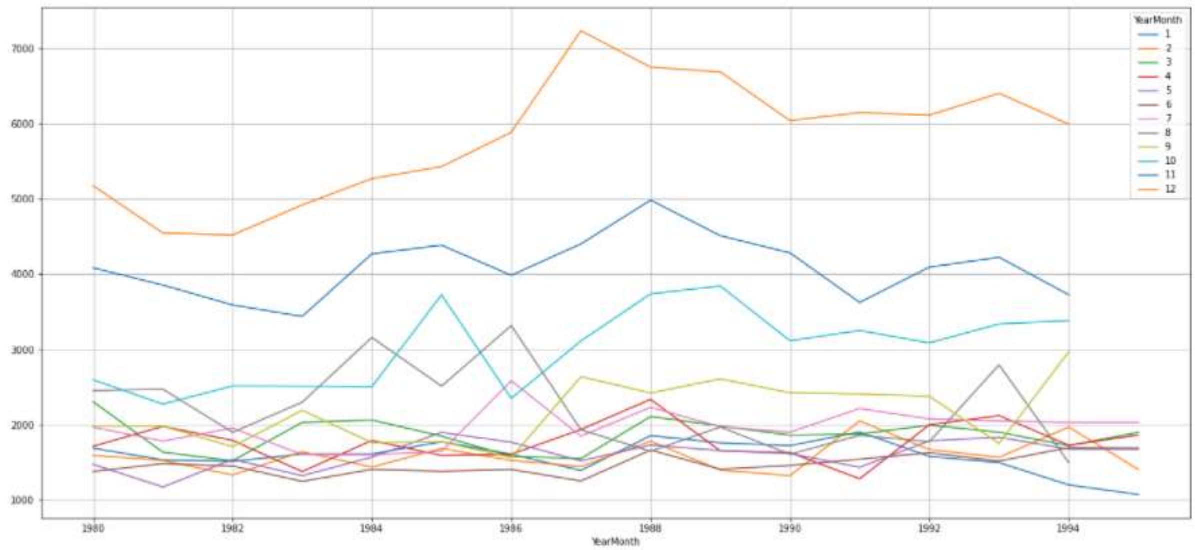
Sparkling monthly sales variation



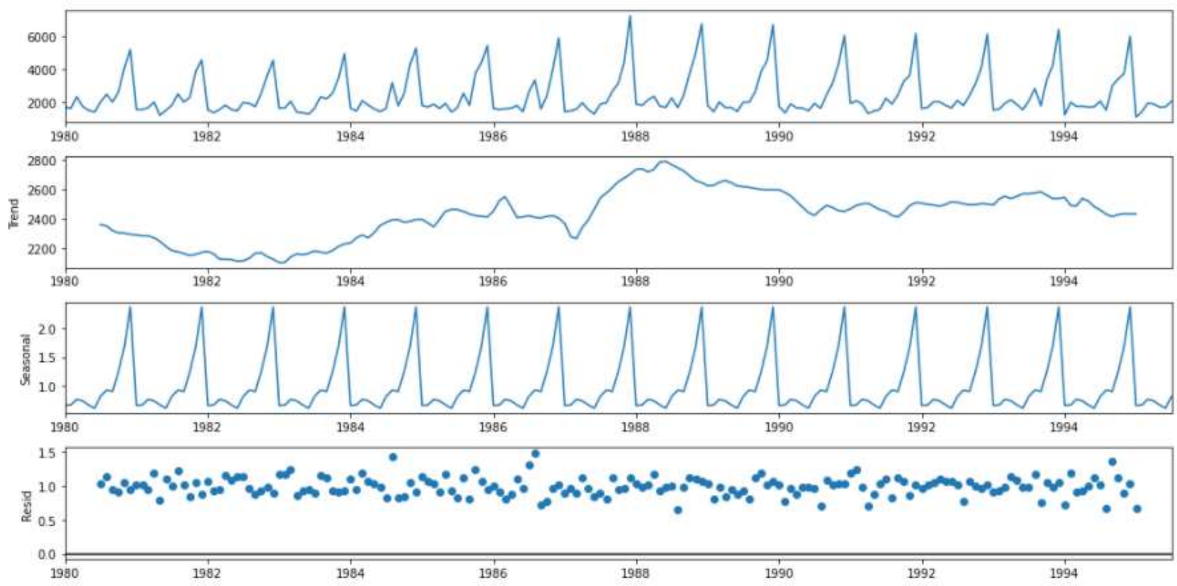
Sparkling wine monthly sales



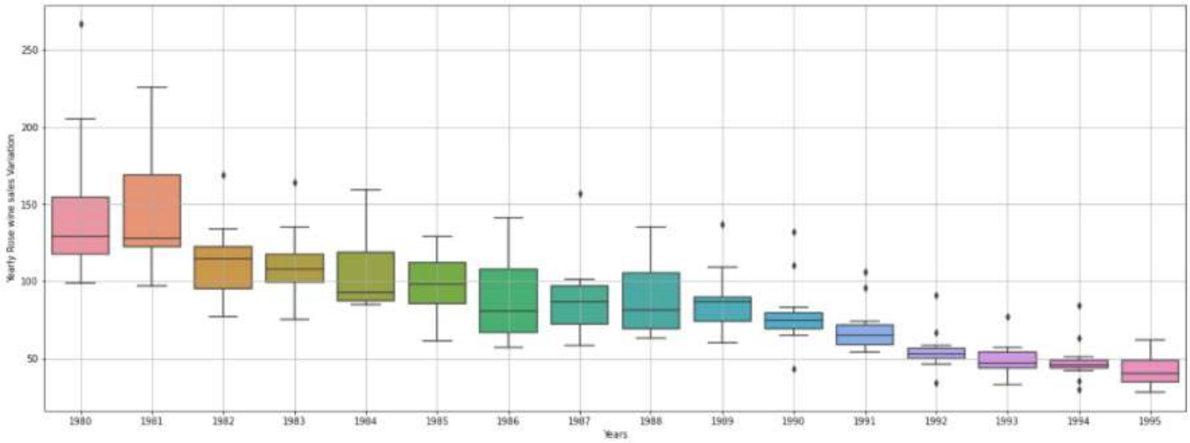
Sparkling wine monthly trend over years



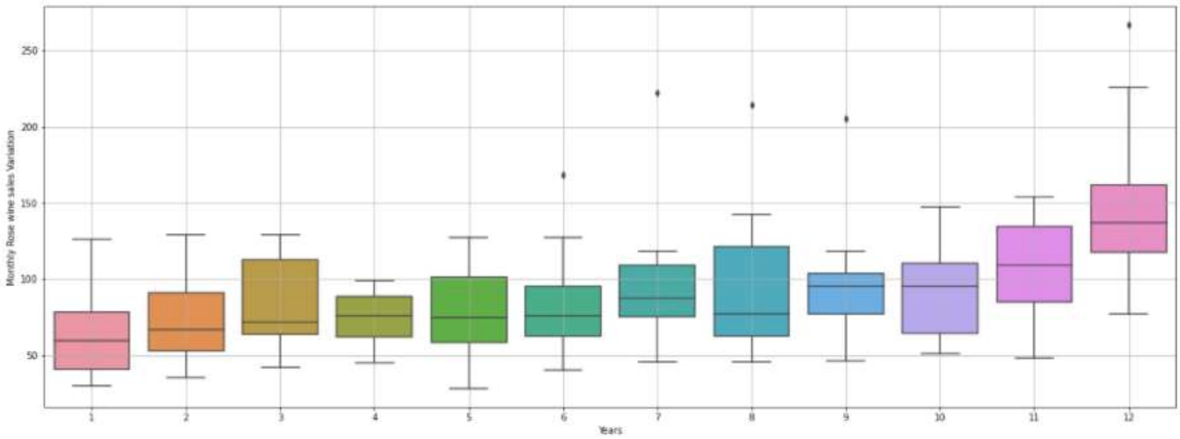
Sparkling wine seasonal decomposition



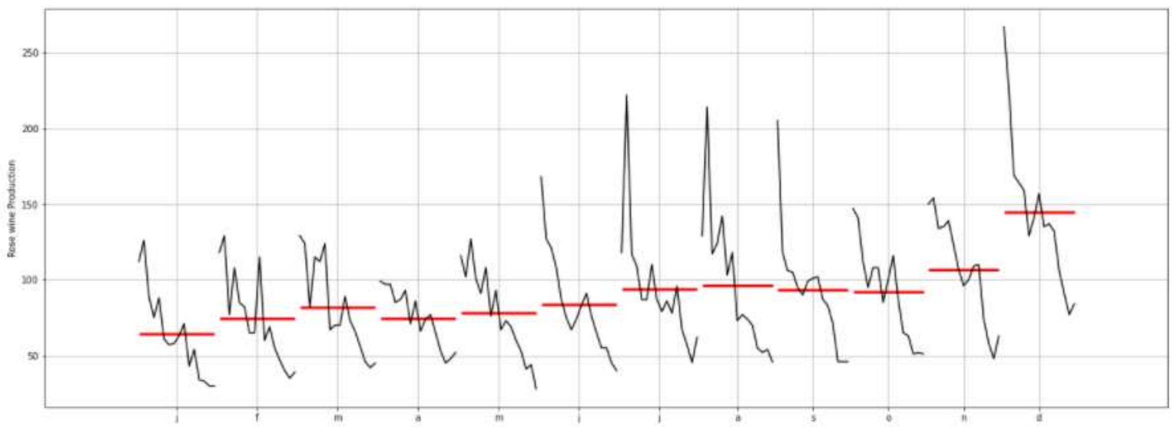
Rose yearly wine sales variation



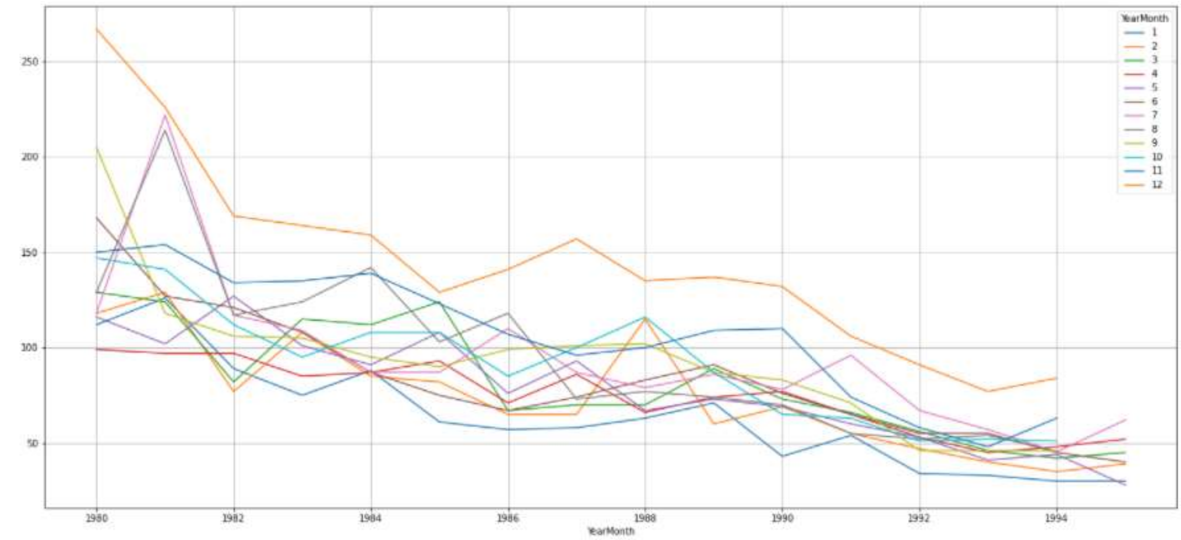
Rose monthly wine sales variation



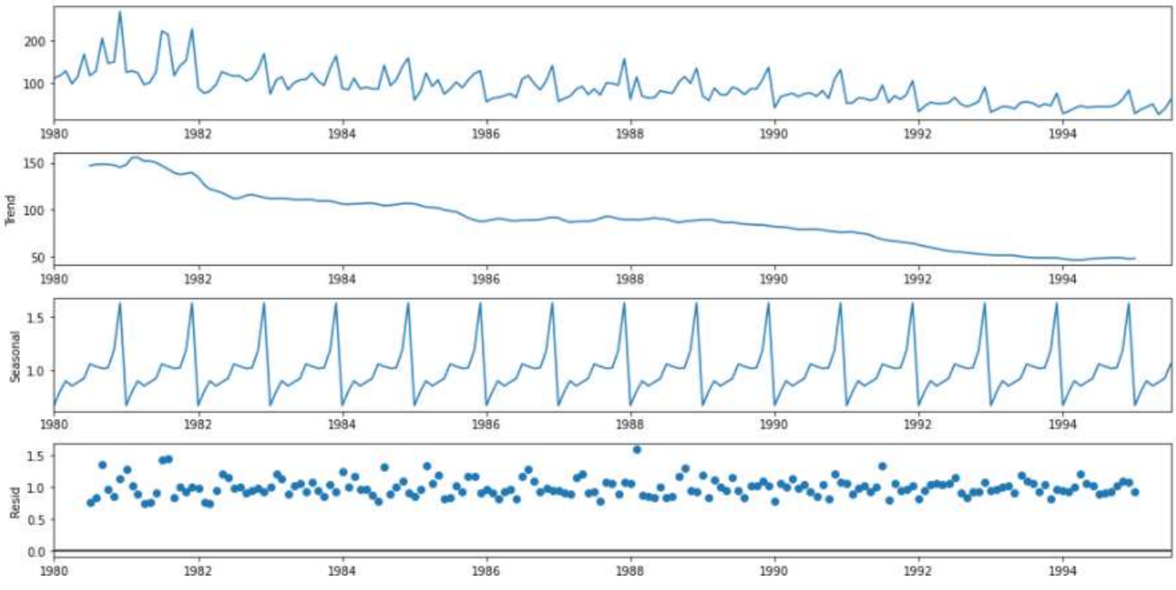
Rose wine sales monthly



Rose wine monthly trend over years



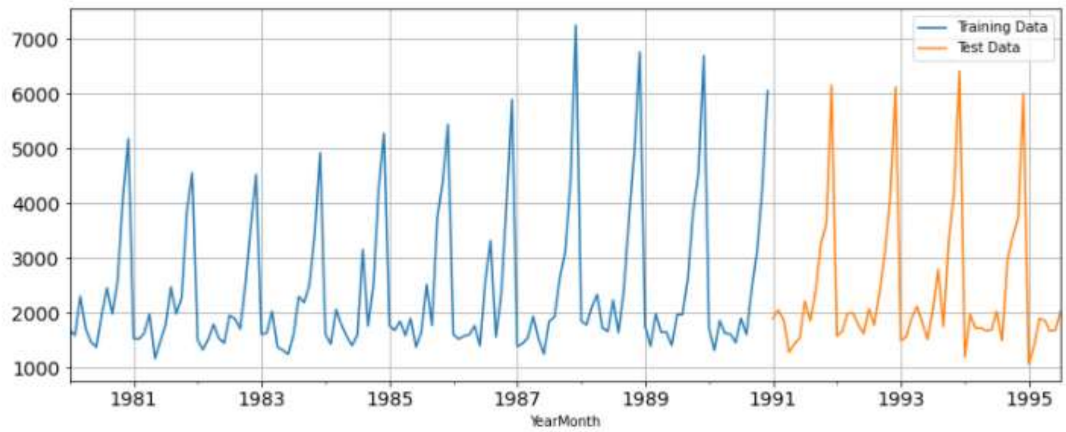
Rose wine seasonal decomposition



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

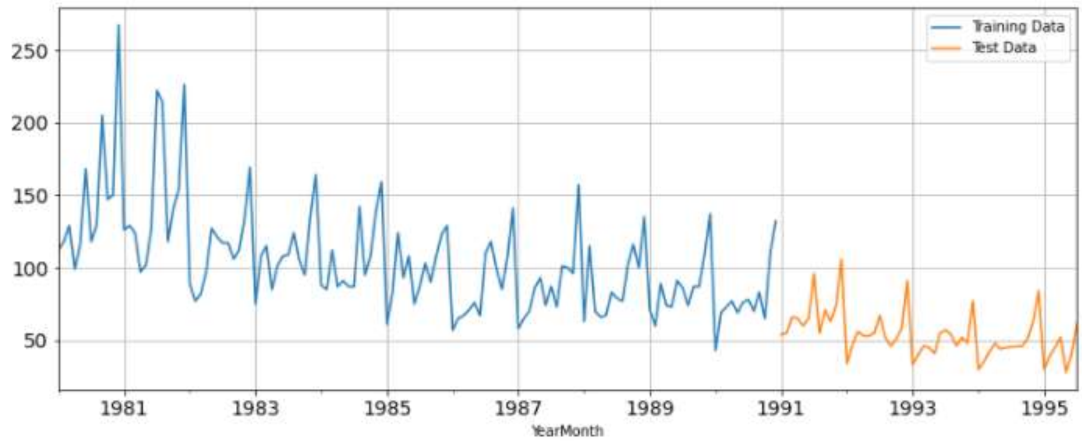
Sparkling

```
The train spark split is (132, 1)
The test spark split is (55, 1)
```



Rose

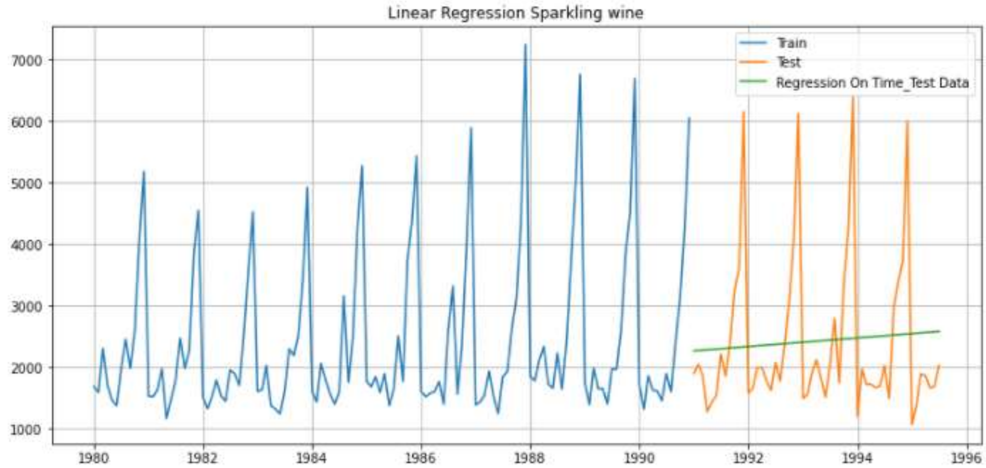
```
The train rose split is (132, 1)
The test rose split is (55, 1)
```



4. Build various 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, simple average models etc. should also be built on the training data and check the performance on the test data using RMSE.

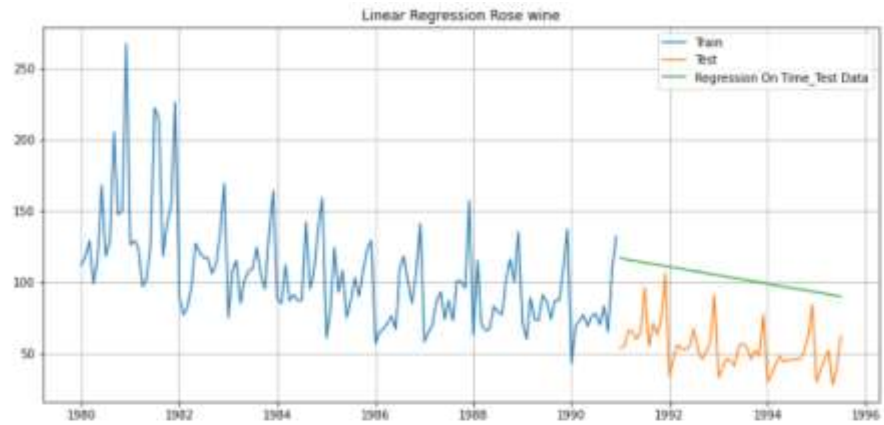
Model 1 : Linear Regression

Sparkling




```
For RegressionOnTime forecast on the Test Data for sparkling,  RMSE is
1275.867
```

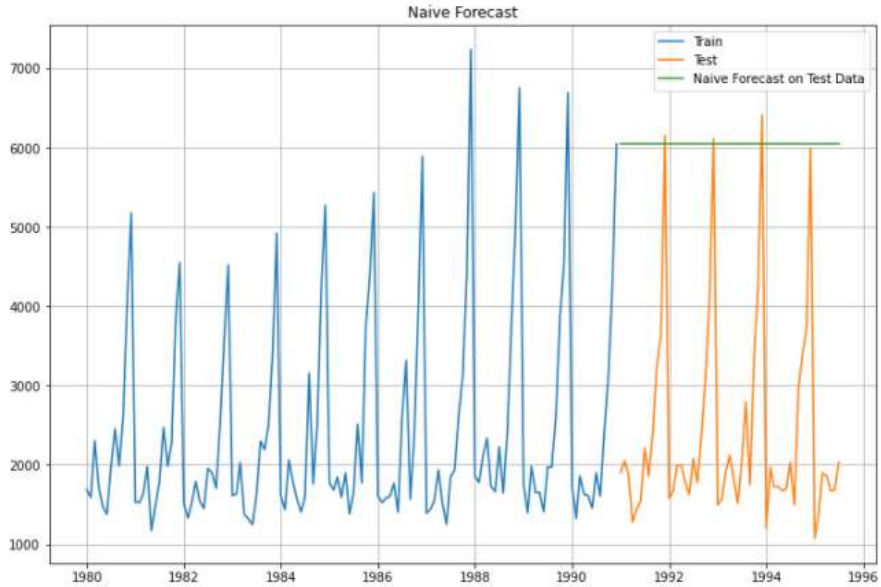
Rose




```
For RegressionOnTime forecast on the Test Data for rose wine,  RMSE is
51.433
```

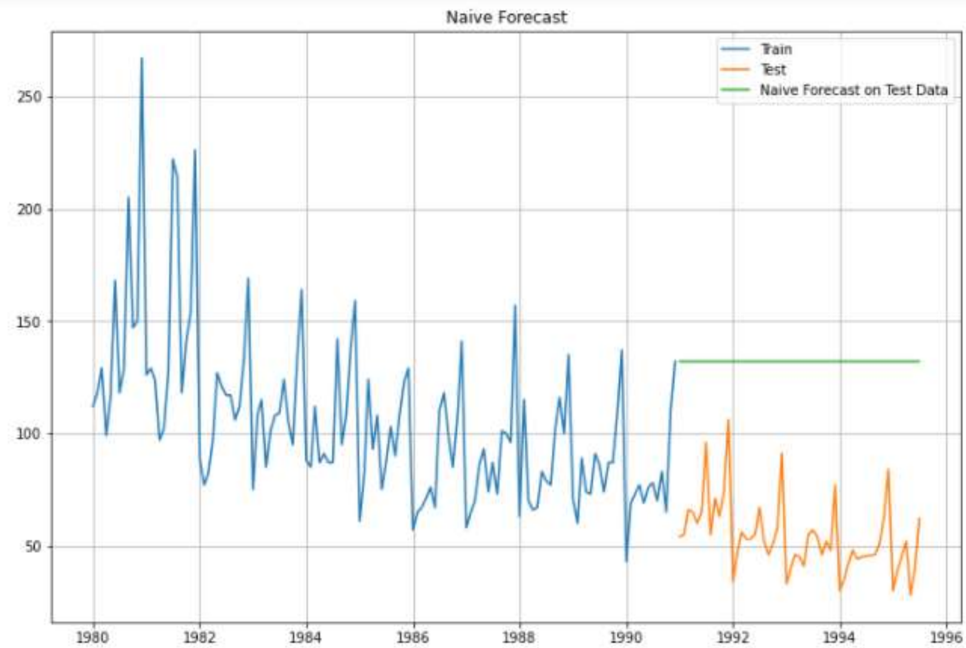
Model 2 : Naïve Approach

Sparkling



For RegressionOnTime forecast on the Test Data for sparkling wine, RMSE is 3864.279

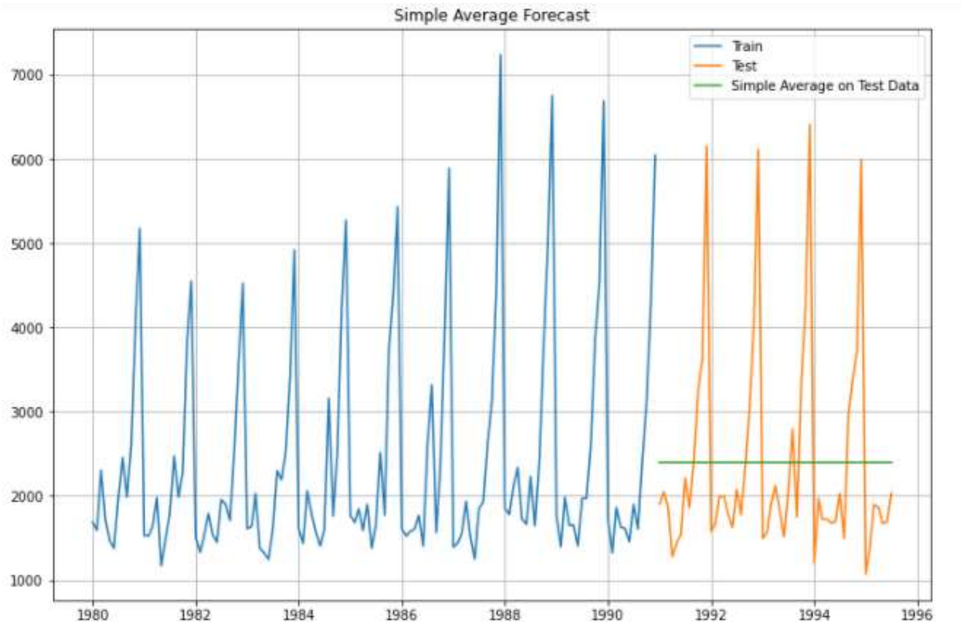
Rose



For RegressionOnTime forecast on the Test Data for rose wine, RMSE is 79.719

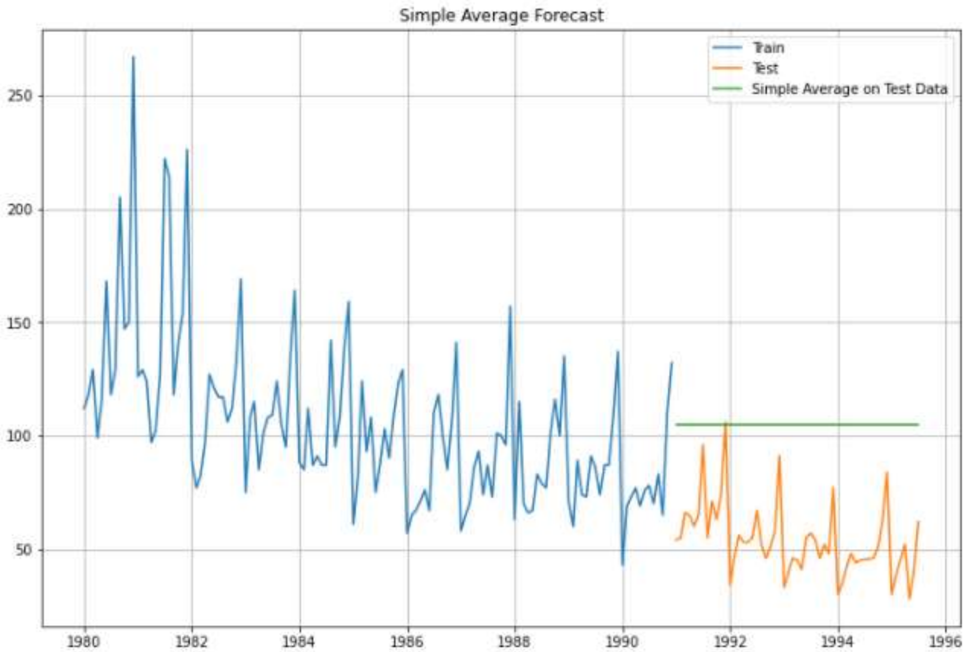
Model 3 : Simple Average

Sparkling Wine




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

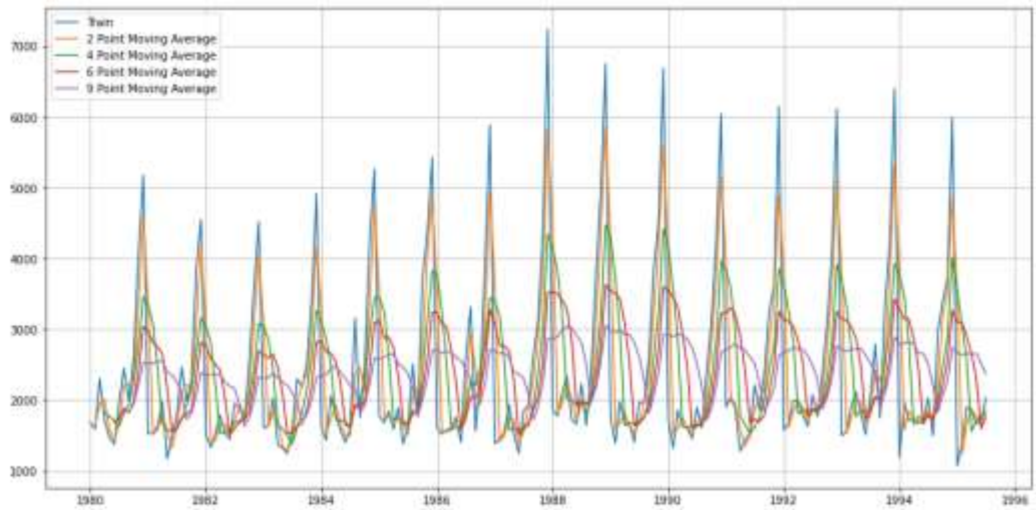
Rose Wine



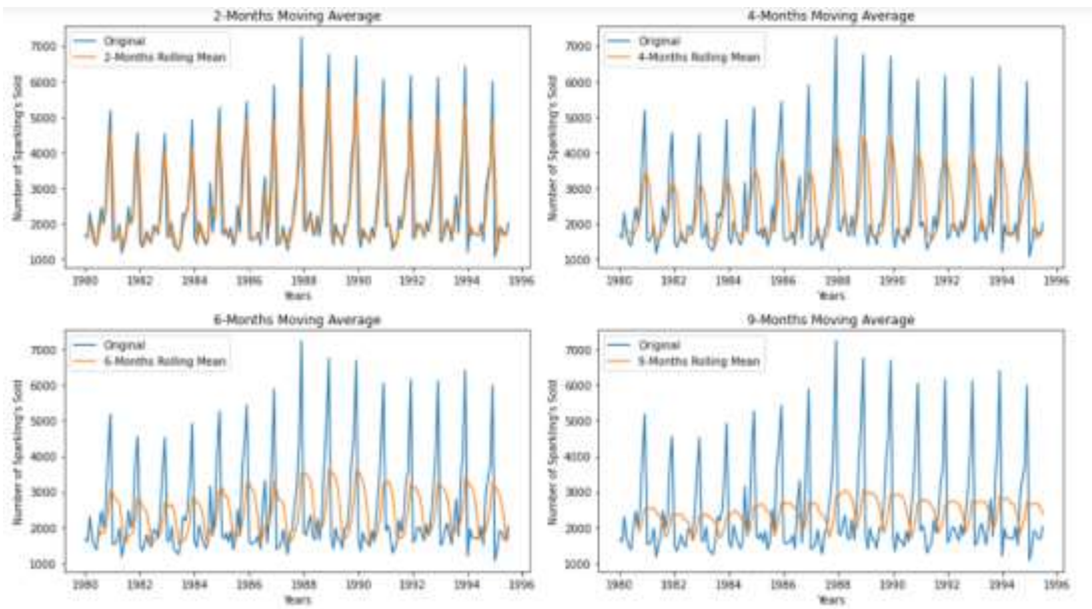

```
For Simple Average forecast on the Test Data,  RMSE is 53.461
```

Model 4 : Moving Average (MA)

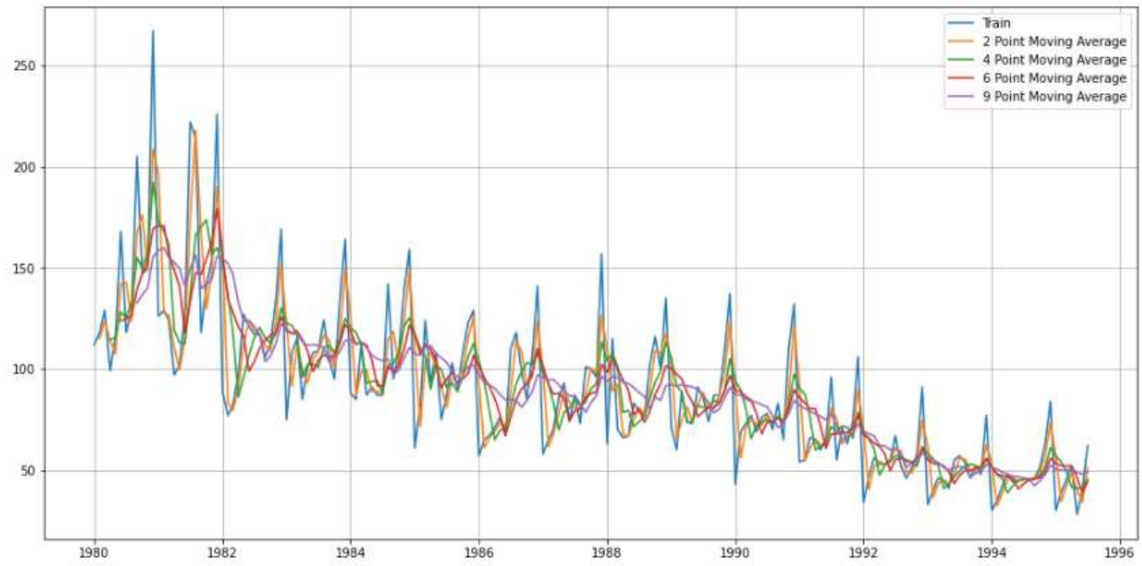
Sparkling Wine



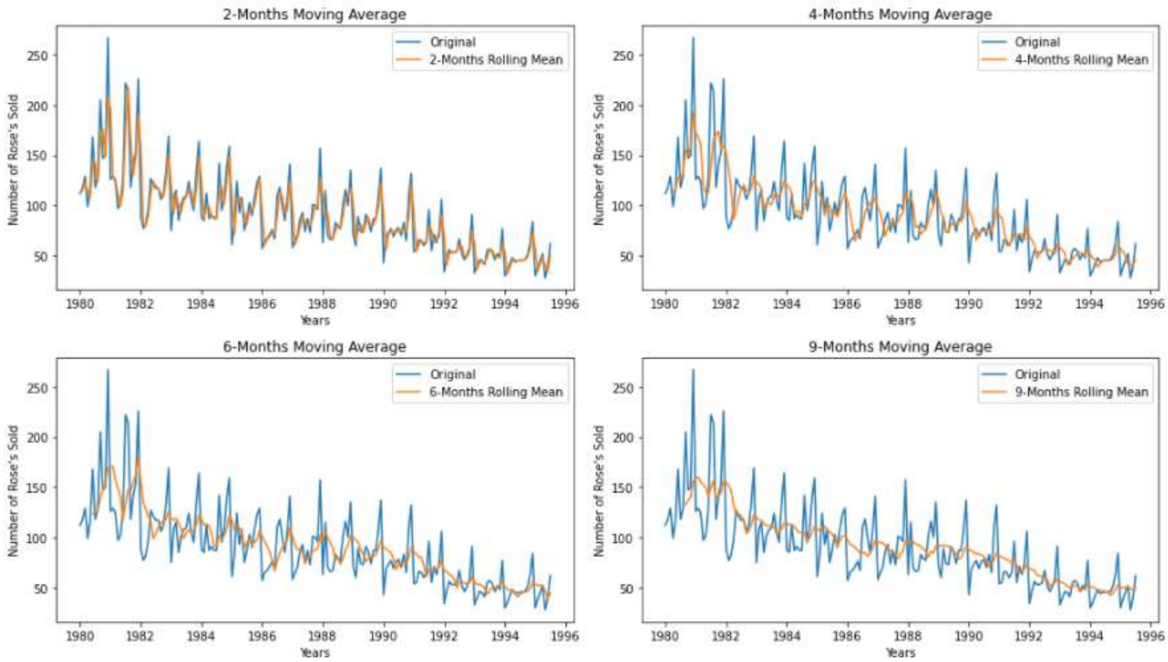

```
For 2 point Moving Average Model forecast on the Training Data,  RMSE i
s 813.401
For 4 point Moving Average Model forecast on the Training Data,  RMSE i
s 1156.590
For 6 point Moving Average Model forecast on the Training Data,  RMSE i
s 1283.927
For 9 point Moving Average Model forecast on the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



Rose Wine




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

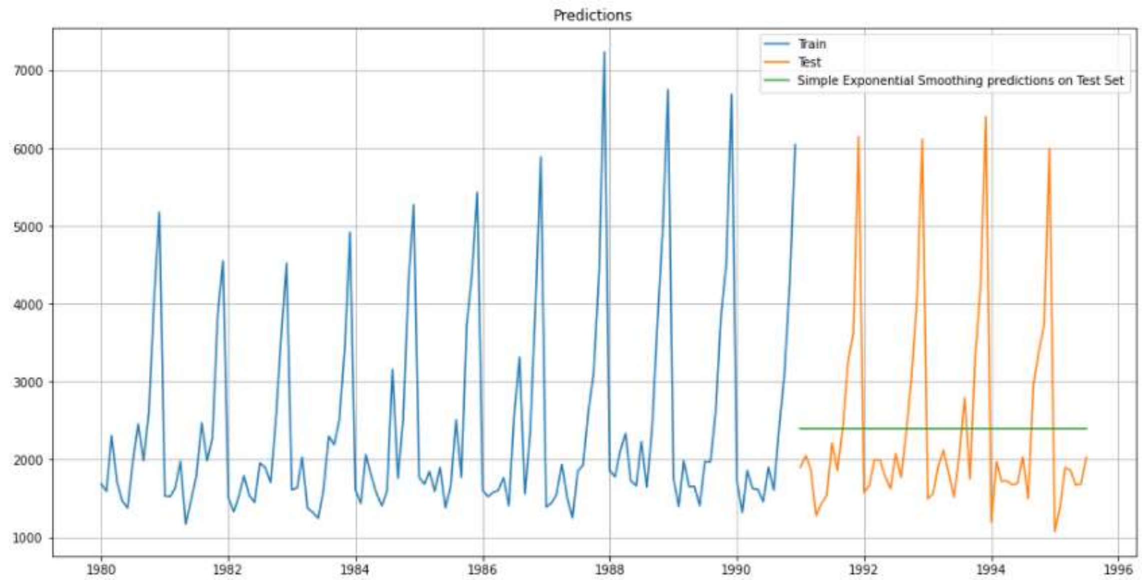


Model 5 : Simple Exponential Smoothing

Sparkling Wine

```
{'smoothing_level': 0.0,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 2403.7831046174856,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

YearMonth	Sparkling	predict
1991-01-01	1902	2403.783105
1991-02-01	2049	2403.783105
1991-03-01	1874	2403.783105
1991-04-01	1279	2403.783105
1991-05-01	1432	2403.783105

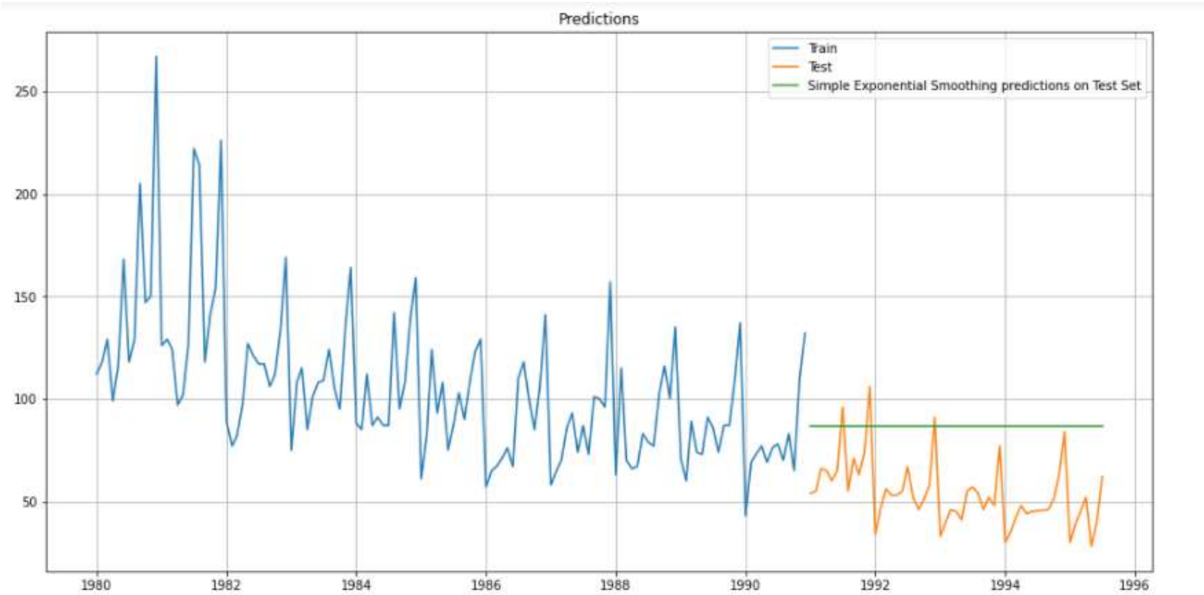


The RMSE for the sparkling wine with simple exponential smoothing is 1275.0818138832155

Rose Wine

```
{'smoothing_level': 0.09,
 'smoothing_trend': nan,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 134.54293365709344,
 'initial_trend': nan,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

YearMonth	Rose	predict
1991-01-01	54.0	86.89235
1991-02-01	55.0	86.89235
1991-03-01	66.0	86.89235
1991-04-01	65.0	86.89235
1991-05-01	60.0	86.89235



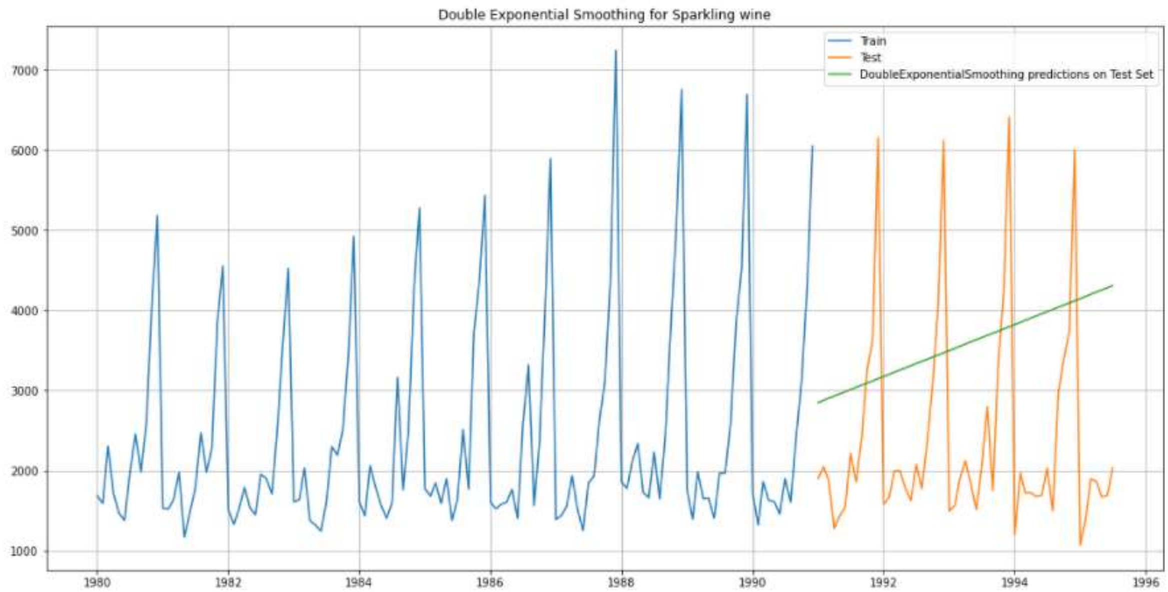
The RMSE for the rose wine with simple exponential smoothing is 36.6041980

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

Sparkling

```
{'smoothing_level': 0.1,
 'smoothing_trend': 0.1,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 2088.585790004531,
 'initial_trend': 12.744864084238706,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

	Sparkling	predict
YearMonth		
1991-01-01	1902	2848.795352
1991-02-01	2049	2875.806333
1991-03-01	1874	2902.817313
1991-04-01	1279	2929.828294
1991-05-01	1432	2956.839275

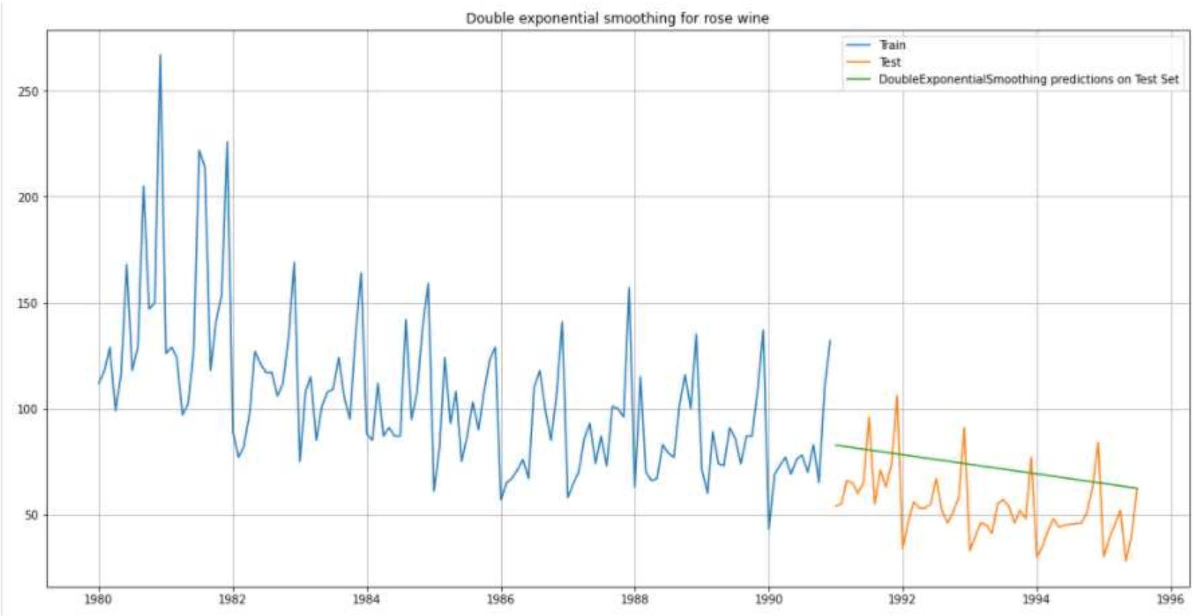



```
The RMSE for the Sparkling wine with simple exponential smoothing is  1779.
4248454808503
```

Rose

```
{'smoothing_level': 0.1,
 'smoothing_trend': 0.01,
 'smoothing_seasonal': nan,
 'damping_trend': nan,
 'initial_level': 139.63025186099742,
 'initial_trend': -0.45711420196168084,
 'initial_seasons': array([], dtype=float64),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

	Rose	predict
YearMonth		
1991-01-01	54.0	82.703413
1991-02-01	55.0	82.327171
1991-03-01	66.0	81.950929
1991-04-01	65.0	81.574686
1991-05-01	60.0	81.198444



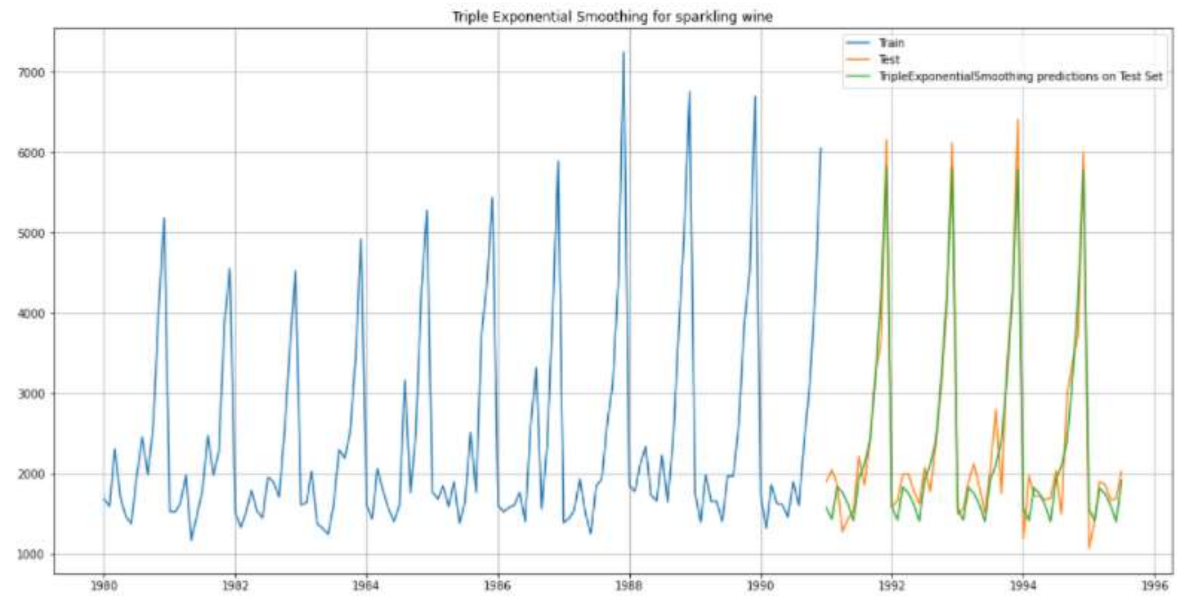
The RMSE for the rose wine with simple exponential smoothing is 23.5993312418369

Model 7 : Triple Exponential Smoothing (Holt – Winter's method)

Sparkling

```
{'smoothing_level': 0.4,
 'smoothing_trend': 0.1,
 'smoothing_seasonal': 0.2,
 'damping_trend': nan,
 'initial_level': 1575.1963858880351,
 'initial_trend': -0.8850775112912266,
 'initial_seasons': array([1.07752216, 1.0664483 , 1.3723685 , 1.23842597, 1.08452692,
        1.03695629, 1.40628503, 1.82484015, 1.56686977, 2.09347699,
        3.09790128, 4.05450111]),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

YearMonth	Sparkling	predict
1991-01-01	1902	1574.366143
1991-02-01	2049	1428.861569
1991-03-01	1874	1837.102381
1991-04-01	1279	1763.597882
1991-05-01	1432	1623.465570

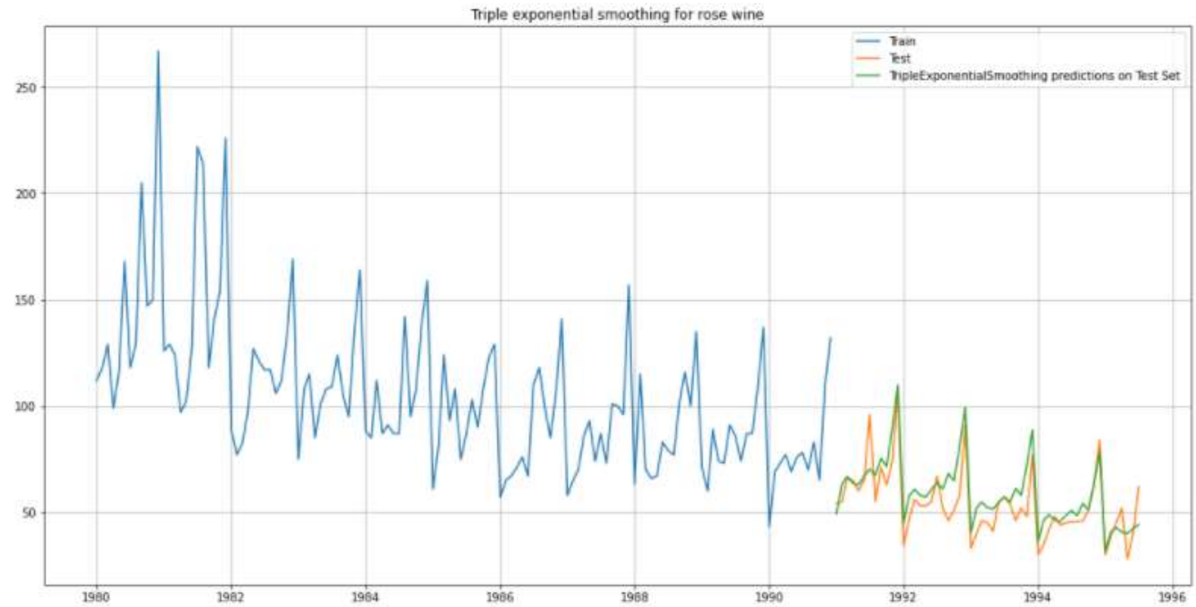


The RMSE for the sparkling wine with triple exponential smoothing is 311.518103875142

Rose

```
{'smoothing_level': 0.1,
 'smoothing_trend': 0.2,
 'smoothing_seasonal': 0.3,
 'damping_trend': nan,
 'initial_level': 62.203275614291485,
 'initial_trend': 0.7307412786209376,
 'initial_seasons': array([1.81649833, 1.89564733, 2.01859067, 1.71531351, 1.93799295,
        2.31651886, 2.56688374, 2.71535987, 2.42963831, 2.33757345,
        2.56325285, 3.82593526]),
 'use_boxcox': False,
 'lamda': None,
 'remove_bias': False}
```

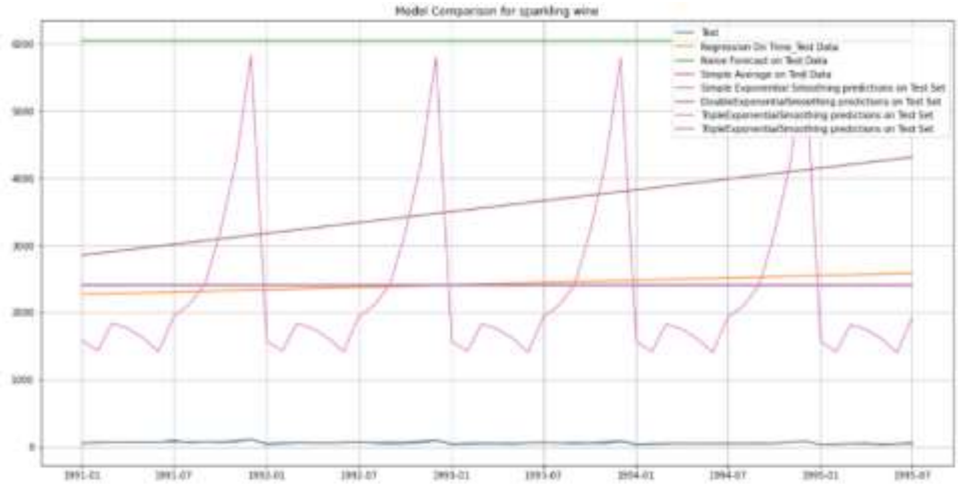
YearMonth	Rose	auto_predict
1991-01-01	54.0	49.241377
1991-02-01	55.0	63.343905
1991-03-01	66.0	66.770776
1991-04-01	65.0	63.625766
1991-05-01	60.0	62.955944



The RMSE for the rose wine with triple exponential smoothing is 9.896240973299607

Model Comparison

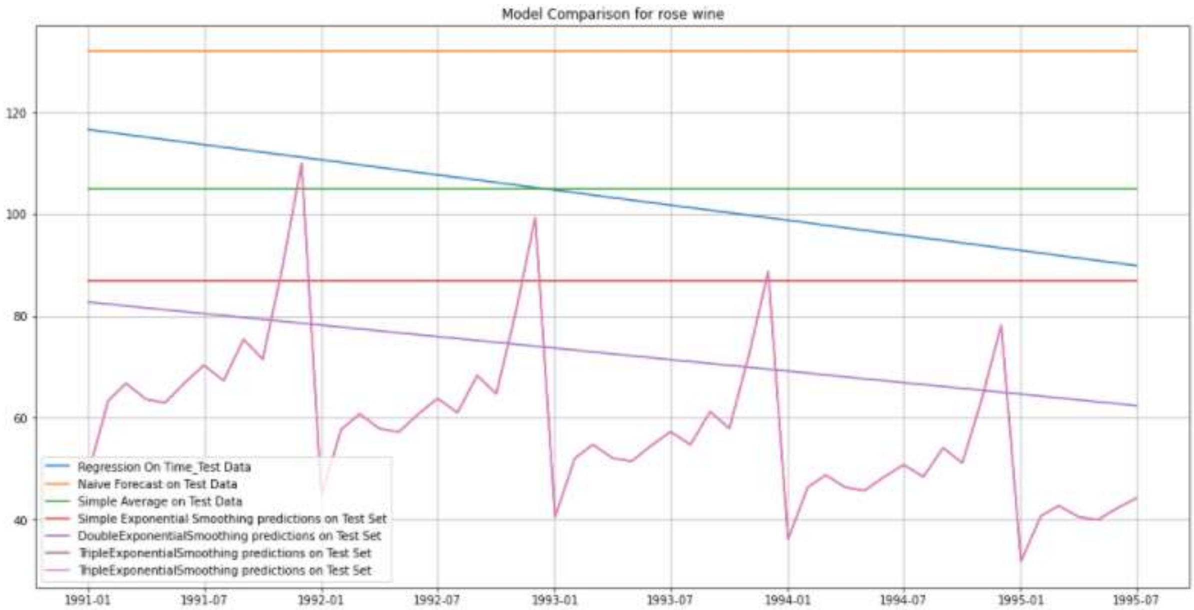
Sparkling



	Test RMSE
RegressionOnTimeSpark	1275.867052
NaiveModel	3864.279352
SimpleAverageModel	1275.081804
2 TMA	813.400684
4 TMA	1156.589694
6 TMA	1283.927428
9 TMA	1346.278315
SES alpha =0.07	1275.081814
DES alpha 0.1 beta 0.1	1779.424845
TES alpha 0.4 beta 0.1 gamma 0.2	311.518104

- From the above we can infer that , triple exponential smoothing gives the best performance since the dataset has all three , level , trend and seasonality when alpha is 0.4 , beta is 0.1 and gamma is 0.2
- The second-best model for the sparkling wine dataset is the 2 point trailing moving average , which fits well and shows a light lag.

Rose



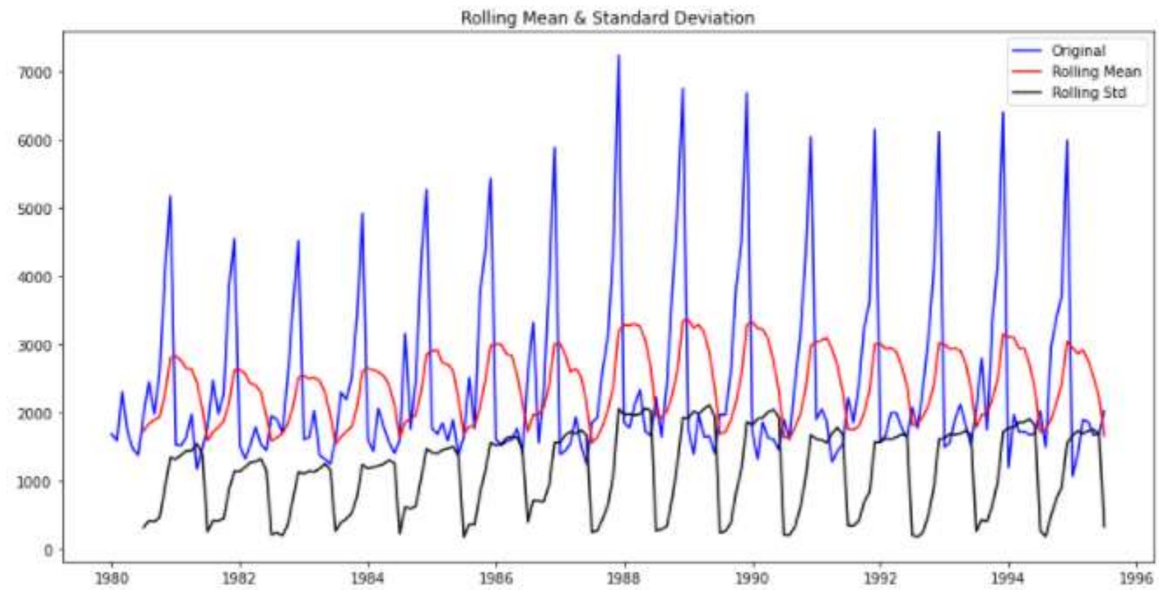
	Test RMSE
RegressionOnTime Rose	51.433312
NaiveModel	79.718773
SimpleAverageModel	53.460570
2 TMA	11.529278
4 TMA	14.451403
6 TMA	14.566327
9 TMA	14.727630
SES alpha 0.09	36.604198
DES alpha 0.1 beta 0.01	23.599331
TES alpha 0.1 beta 0.2 gamma 0.2	9.896241

- From the above we can infer that , triple exponential smoothing gives the best performance since the dataset has all three , level , trend and seasonality when alpha is 0.1 , beta is 0.2 and gamma is 0.2
- The second-best model following TES is 2 point trailing moving average.

5. Check for the stationarity of the data on which the model is being built on using appropriate statistical tests and also mention the hypothesis for the statistical test. If the data is found to be non-stationary, take appropriate steps to make it stationary. Check the new data for stationarity and comment. Note: Stationarity should be checked at alpha = 0.05.

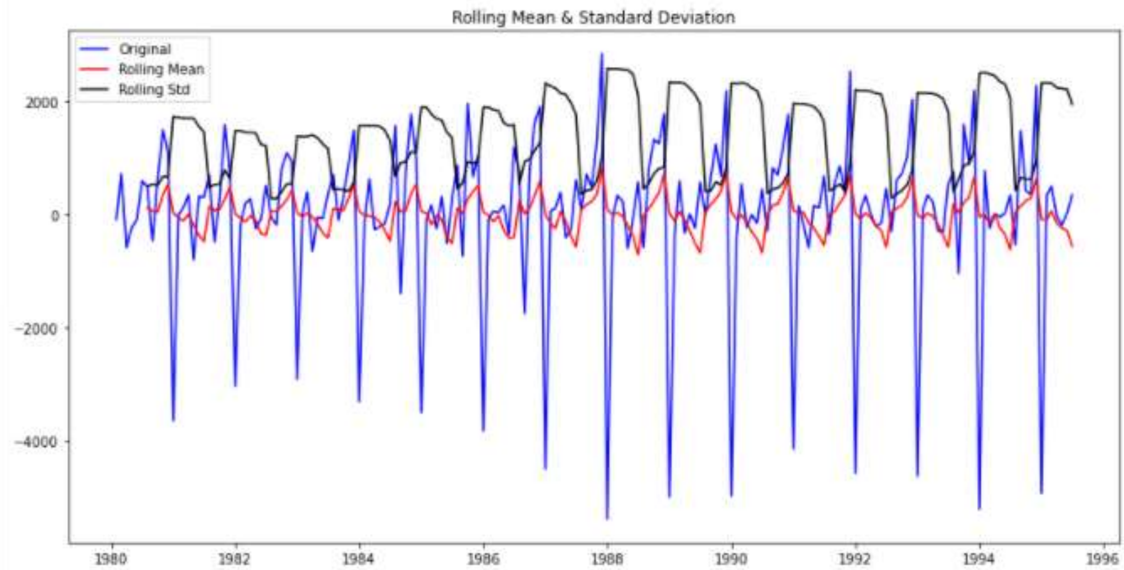
Sparkling

Checking Stationarity




```
Results of Dickey-Fuller Test:
Test Statistic                 -1.360497
p-value                         0.601061
#Lags Used                     11.000000
Number of Observations Used   175.000000
Critical Value (1%)            -3.468280
Critical Value (5%)            -2.878202
Critical Value (10%)           -2.575653
dtype: float64
```

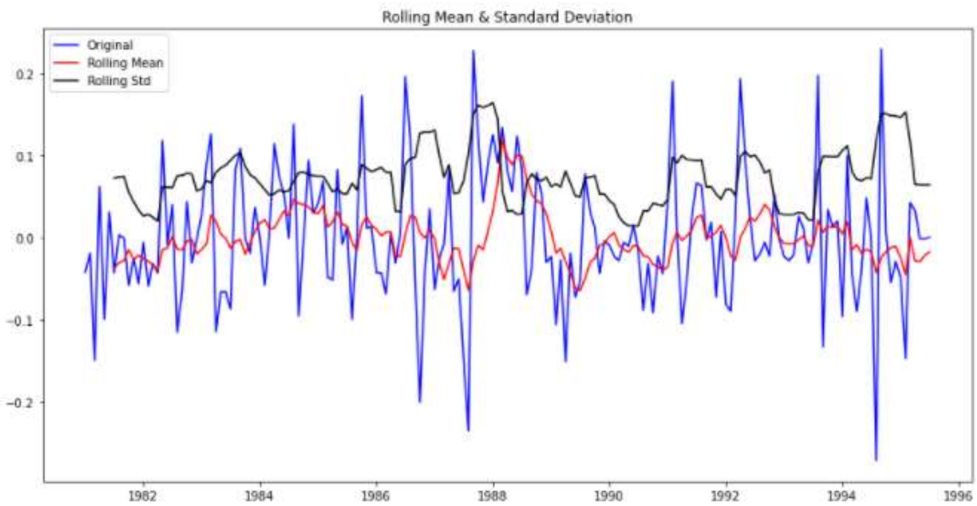
- We see that at 5% significant level the Time Series is non-stationary.
- Let us take a difference of order 1 and check whether the Time Series is stationary or not




```
Results of Dickey-Fuller Test:
Test Statistic                 -45.050301
p-value                          0.000000
#Lags Used                      10.000000
Number of Observations Used    175.000000
Critical Value (1%)             -3.468280
Critical Value (5%)             -2.878202
Critical Value (10%)            -2.575653
dtype: float64
```

We see that at α = 0.05 the Time Series is indeed stationary.

Difference of the log series

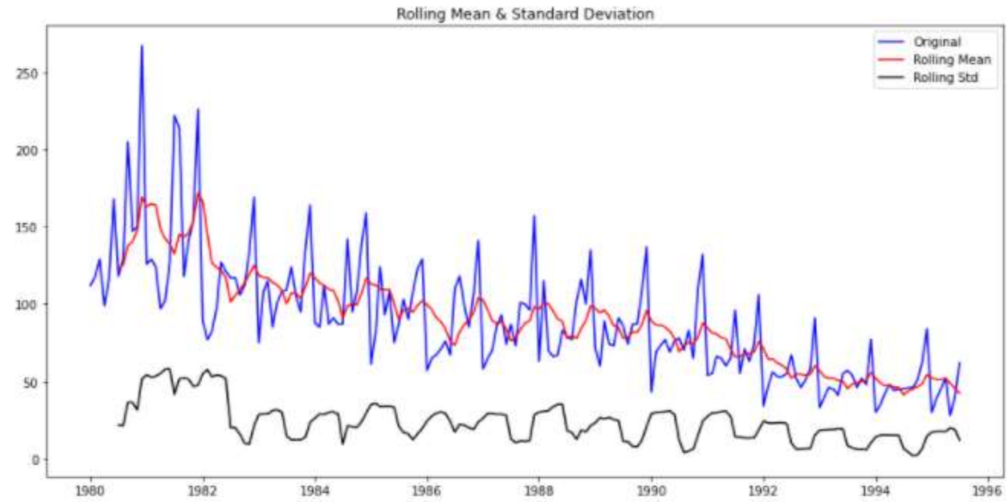



```
Results of Dickey-Fuller Test:
Test Statistic                  -5.183811
p-value                          0.000009
#Lags Used                      11.000000
Number of Observations Used    163.000000
Critical Value (1%)             -3.471119
Critical Value (5%)             -2.879441
Critical Value (10%)            -2.576314
dtype: float64
```

From the above the data is converted into stationary looking at the α value 0.05

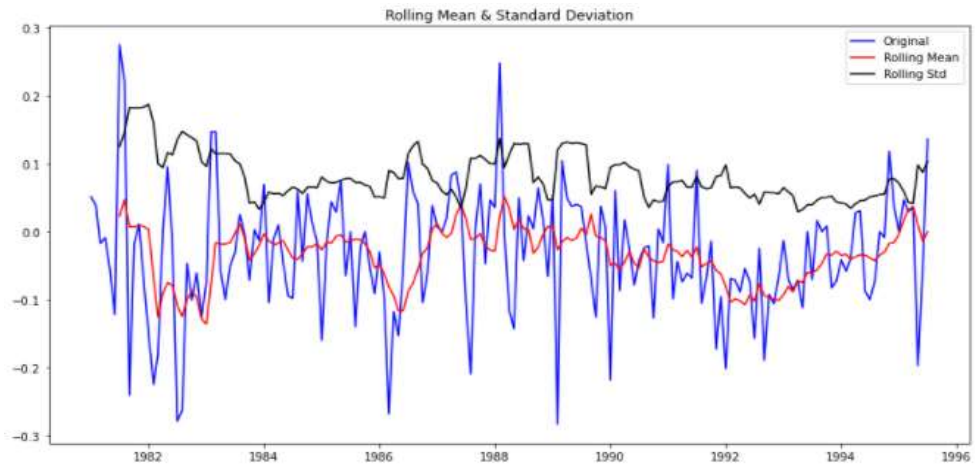
Rose

Checking Stationarity




```
Results of Dickey-Fuller Test:
Test Statistic                  -1.876699
p-value                          0.343101
#Lags Used                      13.000000
Number of Observations Used    173.000000
Critical Value (1%)             -3.468726
Critical Value (5%)             -2.878396
Critical Value (10%)            -2.575756
dtype: float64
```

We see that at 5% significant level the Time Series is non-stationary. Let us take a difference of order 1 and check whether the Time Series is stationary or not.




```
Results of Dickey-Fuller Test:
Test Statistic                  -3.932899
p-value                          0.001806
#Lags Used                      11.000000
Number of Observations Used    163.000000
Critical Value (1%)             -3.471119
Critical Value (5%)             -2.879441
Critical Value (10%)            -2.576314
dtype: float64
```

By applying log diff the data is converted into stationary 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.

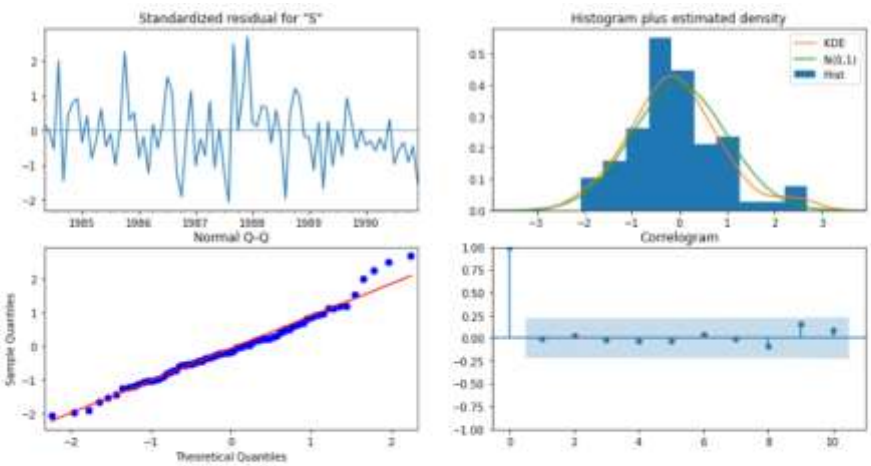
Model 8 : SARIMA

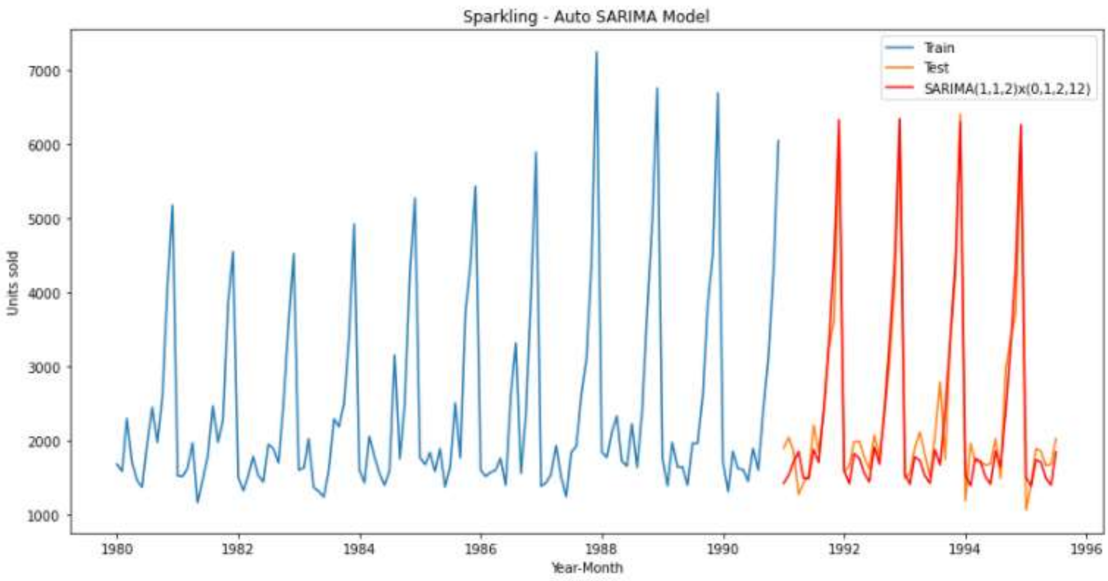
Sparkling Wine

AUTO SARIMA model

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                           Sparkling   No. Observations:                  132
Model:             SARIMAX(3, 1, 3)x(3, 1, [], 12)   Log Likelihood                -596.641
Date:                             Sat, 25 Sep 2021   AIC                           1213.283
Time:                                     15:04:53   BIC                           1237.103
Sample:                                 01-01-1980   HQIC                          1222.833
                                      - 12-01-1990
Covariance Type:                               opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -1.6130      0.176     -9.175      0.000      -1.958      -1.268
ar.L2         -0.6102      0.299     -2.039      0.041      -1.197      -0.024
ar.L3          0.0871      0.161      0.542      0.588      -0.228       0.402
ma.L1          0.9854      0.469      2.103      0.035       0.067       1.904
ma.L2         -0.8740      0.166     -5.259      0.000      -1.200      -0.548
ma.L3         -0.9465      0.486     -1.948      0.051      -1.899       0.006
ar.S.L12      -0.4525      0.142     -3.195      0.001      -0.730      -0.175
ar.S.L24      -0.2334      0.144     -1.617      0.106      -0.516       0.049
ar.S.L36      -0.1003      0.121     -0.826      0.409      -0.338       0.138
sigma2      1.839e+05   8.91e+04      2.063      0.039    9175.674    3.59e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):                 4.08
Prob(Q):                              0.93   Prob(JB):                         0.13
Heteroskedasticity (H):               0.73   Skew:                             0.48
Prob(H) (two-sided):                  0.41   Kurtosis:                         3.54
===================================================================================
```

The residual seems to follow the properties of white noise





The RMSE for the sparkling wine with SAMRIMA model is 331.638348031837

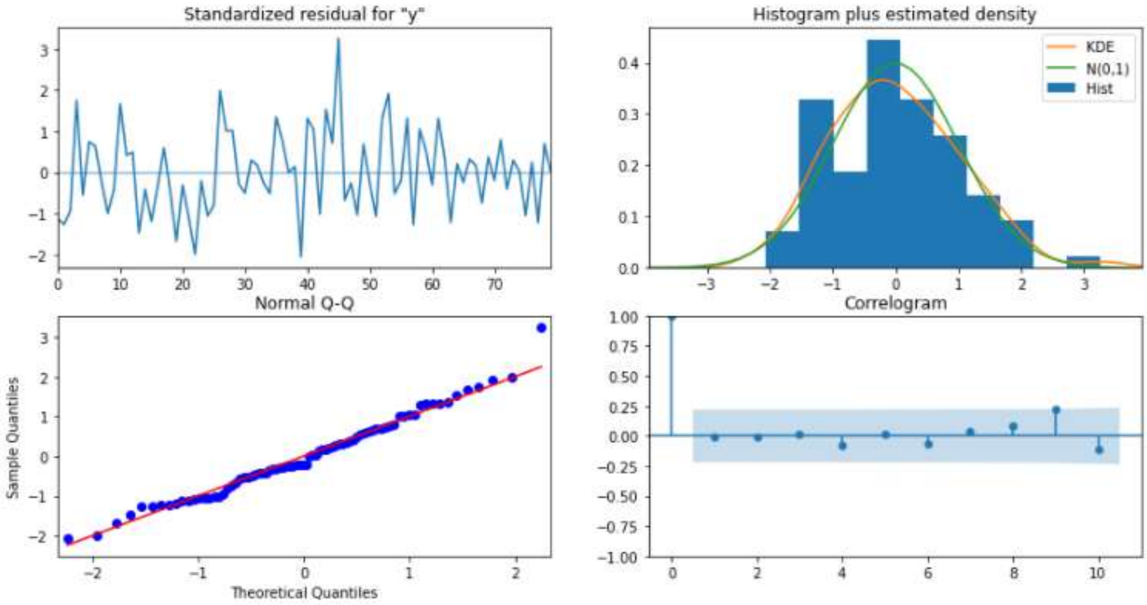
Rose Wine

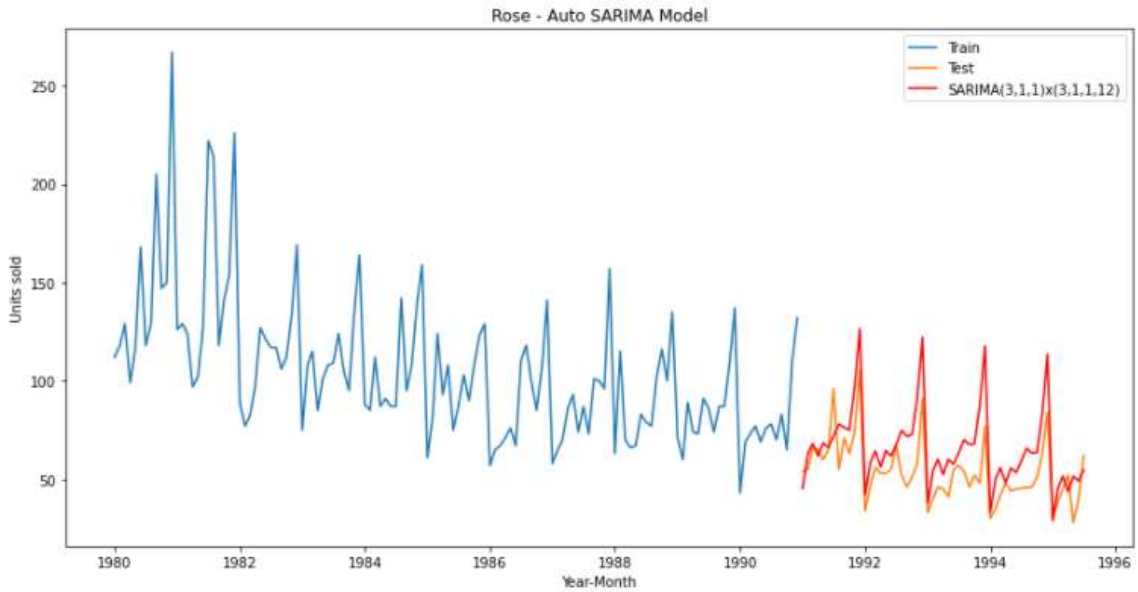
AUTO SARIMA model

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                                    y   No. Observations:                  132
Model:             SARIMAX(3, 1, 1)x(3, 1, 1, 12)   Log Likelihood                -331.681
Date:                             Sat, 25 Sep 2021   AIC                            681.363
Time:                                     15:05:03   BIC                            702.801
Sample:                                          0   HQIC                           689.958
                                             - 132
Covariance Type:                               opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.0173      0.151      0.114      0.909      -0.279       0.314
ar.L2         -0.0426      0.141     -0.302      0.763      -0.319       0.234
ar.L3         -0.0574      0.119     -0.484      0.628      -0.290       0.175
ma.L1         -0.9388      0.085    -11.105      0.000      -1.105      -0.773
ar.S.L12       0.0907      0.126      0.721      0.471      -0.156       0.337
ar.S.L24      -0.0437      0.108     -0.406      0.684      -0.255       0.167
ar.S.L36   -3.645e-05      0.053     -0.001      0.999      -0.103       0.103
ma.S.L12      -1.0000   2169.358     -0.000      1.000   -4252.864    4250.864
sigma2       185.3957   4.02e+05      0.000      1.000   -7.88e+05    7.88e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):                 2.56
Prob(Q):                              0.91   Prob(JB):                         0.28
Heteroskedasticity (H):               0.56   Skew:                             0.42
Prob(H) (two-sided):                  0.13   Kurtosis:                         3.22
===================================================================================

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

The residual seems to follow the properties of white noise

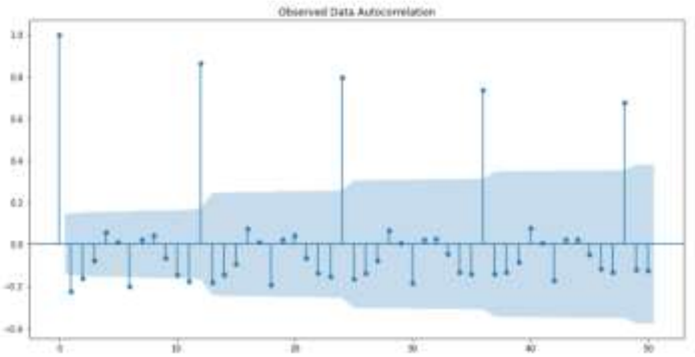


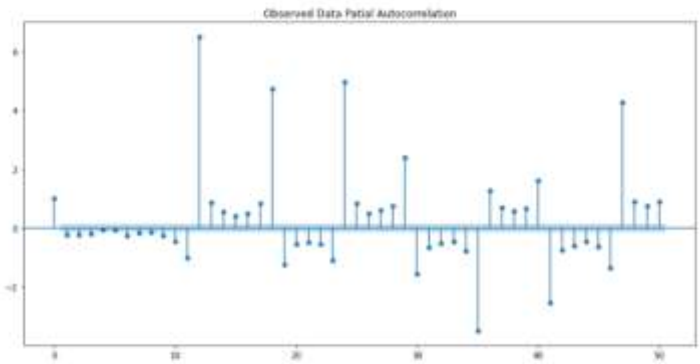



```
The RMSE for the rose wine with SAMRIMA model is  16.823783943392556
```

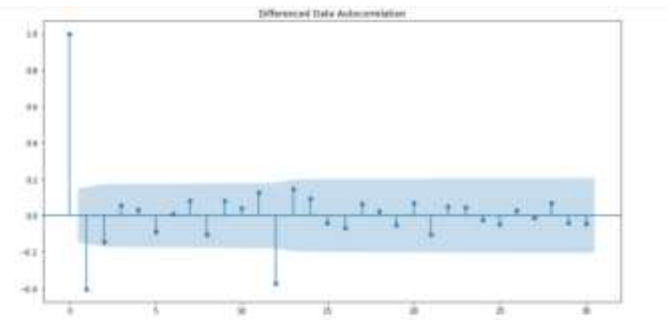
7. Build ARIMA/SARIMA models based on the cut-off points of ACF and PACF on the training data and evaluate this model on the test data using RMSE.

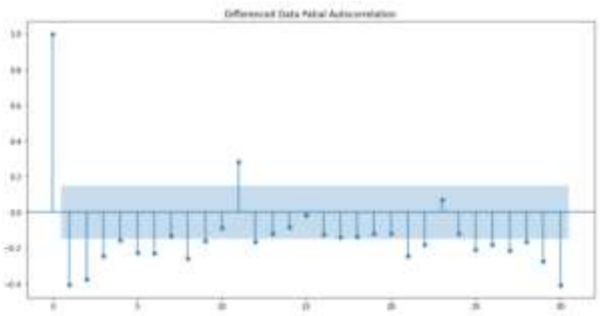
Sparkling Wine





We see that our ACF plot at the seasonal interval (12) does not taper off quickly. So, we go ahead and take a seasonal differencing of the original series. Before that let us look at the original series.

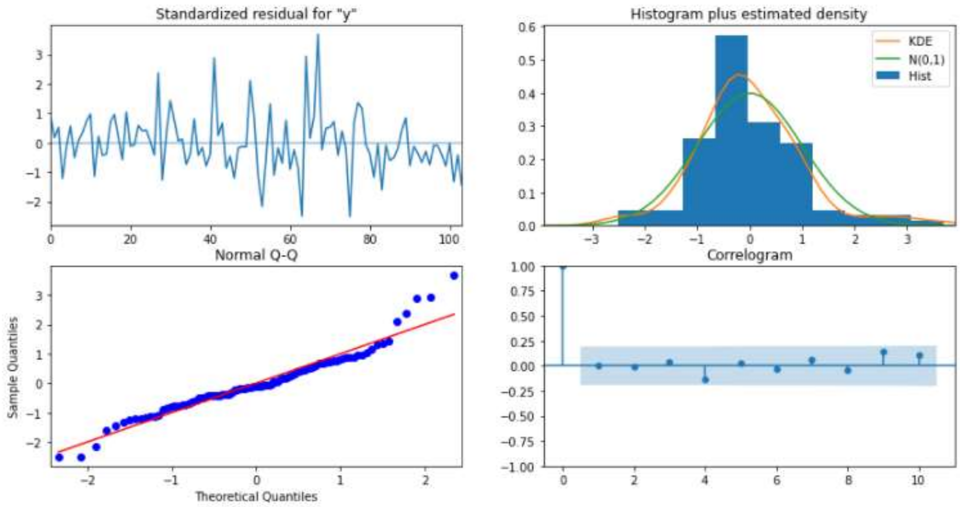




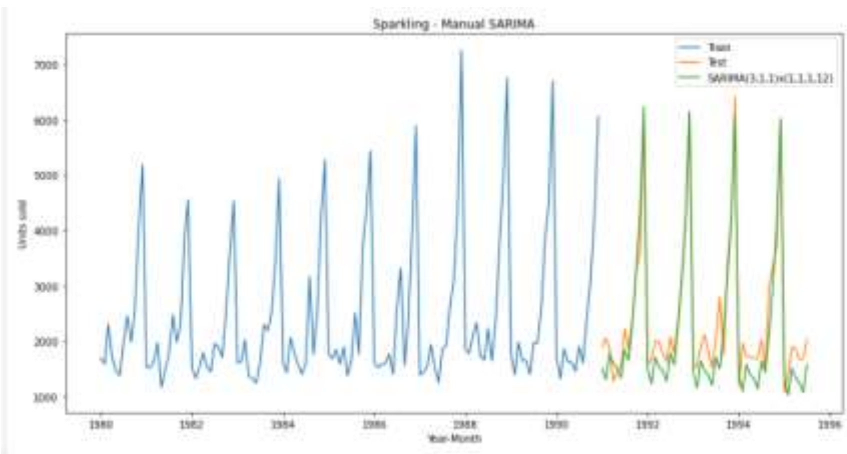
- From the above ACF graph , there are 3 significant values so p = 3, d = 1 and q = 1
- From PACF graph , there is some seasonality at 12, so P = 1 , and d = 1 and q =1

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                                y   No. Observations:                  132
Model:             SARIMAX(3, 1, 1)x(1, 1, 1, 12)   Log Likelihood                -772.847
Date:                          Sat, 25 Sep 2021   AIC                           1559.693
Time:                                  15:05:07   BIC                           1578.204
Sample:                                       0   HQIC                          1567.192
                                          - 132
Covariance Type:                            opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.1783      0.125      1.425      0.154      -0.067       0.424
ar.L2         -0.0904      0.127     -0.709      0.478      -0.340       0.159
ar.L3          0.0558      0.108      0.515      0.606      -0.156       0.268
ma.L1         -0.9293      0.063    -14.673      0.000      -1.053      -0.805
ar.S.L12      -0.0931      0.212     -0.439      0.660      -0.509       0.322
ma.S.L12      -0.3787      0.212     -1.789      0.074      -0.794       0.036
sigma2      1.658e+05     1.9e+04      8.713      0.000    1.29e+05    2.03e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):                28.49
Prob(Q):                              0.96   Prob(JB):                         0.00
Heteroskedasticity (H):               1.14   Skew:                             0.70
Prob(H) (two-sided):                  0.70   Kurtosis:                         5.14
===================================================================================
```

The residual seems to follow the properties of a white noise

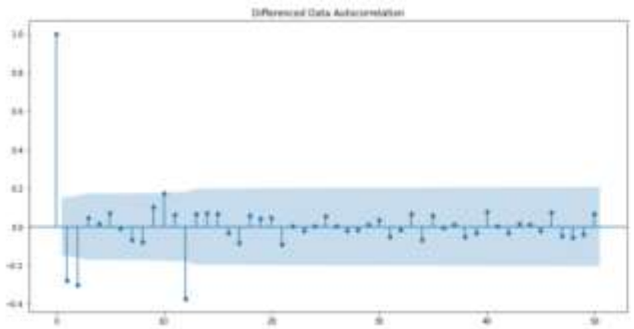


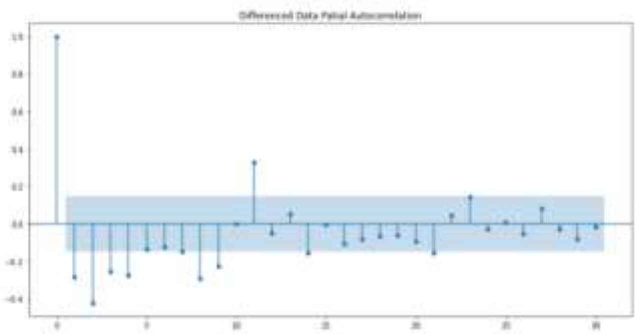
Prediction on test data



Rose Wine

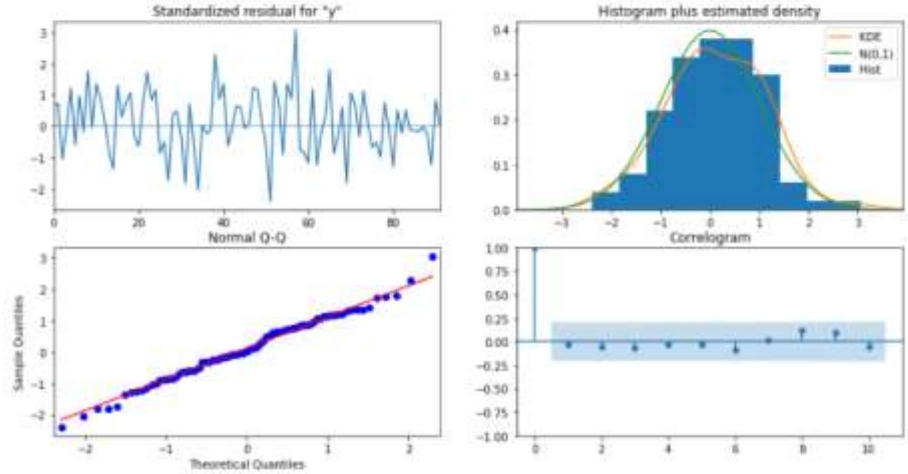
Taking the difference and take auto correlation and plotting ACF and PACF

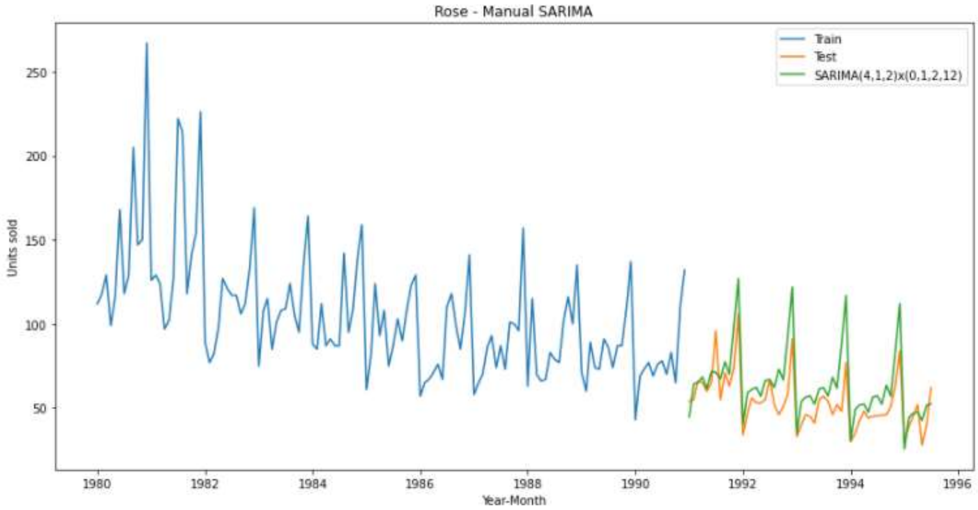




From above we can choose p = 4 , q = 2 and d = 1 , P = 0 , D= 1 , Q =2 , S = 12

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                                   y   No. Observations:                  132
Model:             SARIMAX(4, 1, 2)x(0, 1, 2, 12)   Log Likelihood                -384.369
Date:                             Sat, 25 Sep 2021   AIC                            786.737
Time:                                     15:05:12   BIC                            809.433
Sample:                                          0   HQIC                           795.898
                                             - 132
Covariance Type:                               opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.8967      0.132     -6.814      0.000      -1.155      -0.639
ar.L2          0.0165      0.171      0.097      0.923      -0.319       0.352
ar.L3         -0.1132      0.174     -0.650      0.515      -0.454       0.228
ar.L4         -0.1598      0.116     -1.380      0.168      -0.387       0.067
ma.L1          0.1508      0.174      0.866      0.387      -0.191       0.492
ma.L2         -0.8492      0.164     -5.166      0.000      -1.171      -0.527
ma.S.L12      -0.3907      0.102     -3.848      0.000      -0.590      -0.192
ma.S.L24      -0.0887      0.091     -0.977      0.329      -0.267       0.089
sigma2       238.9649      0.001   2.02e+05      0.000     238.963     238.967
===================================================================================
Ljung-Box (L1) (Q):                   0.06   Jarque-Bera (JB):                 0.01
Prob(Q):                              0.80   Prob(JB):                         0.99
Heteroskedasticity (H):               0.76   Skew:                            -0.01
Prob(H) (two-sided):                  0.46   Kurtosis:                         3.06
===================================================================================
```


Standardized residual for "y"
Histogram plus estimated density
KDE
N(0,1)
Hist
Normal Q-Q
Sample Quantiles
Theoretical Quantiles
Correlogram


Rose - Manual SARIMA
Train
Test
SARIMA(4,1,2)x(0,1,2,12)
Units sold
Year-Month

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.

	Test RMSE
TES alpha 0.4 beta 0.1 gamma 0.2	311.518104
AUTO SARIMA(3,1,3)x(3,1,0,12)	331.638348
SARIMA(3,1,1)x(1,1,1,12)	412.781138
2 TMA	813.400684
4 TMA	1156.589694
SimpleAverageModel	1275.081804
SES alpha =0.07	1275.081814
RegressionOnTimeSpark	1275.867052
6 TMA	1283.927428
9 TMA	1346.278315
DES alpha 0.1 beta 0.1	1779.424845
NaiveModel	3864.279352

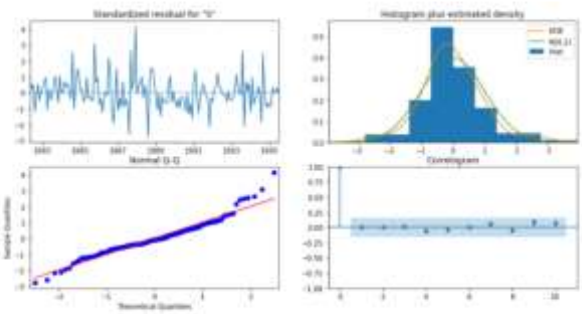
	Test RMSE
TES alpha 0.1 beta 0.2 gamma 0.2	9.896241
2 TMA	11.529278
4 TMA	14.451403
6 TMA	14.566327
9 TMA	14.727630
SARIMA(4,1,2)x(0,1,2,12)	15.377252
AUTO SARIMA(3,1,1)x(3,1,1,12)	16.823784
DES alpha 0.1 beta 0.01	23.599331
SES alpha 0.09	36.604198
RegressionOnTime Rose	51.433312
SimpleAverageModel	53.460570
NaiveModel	79.718773

From all of the models the naïve model seems to be doing worse for both the datasets

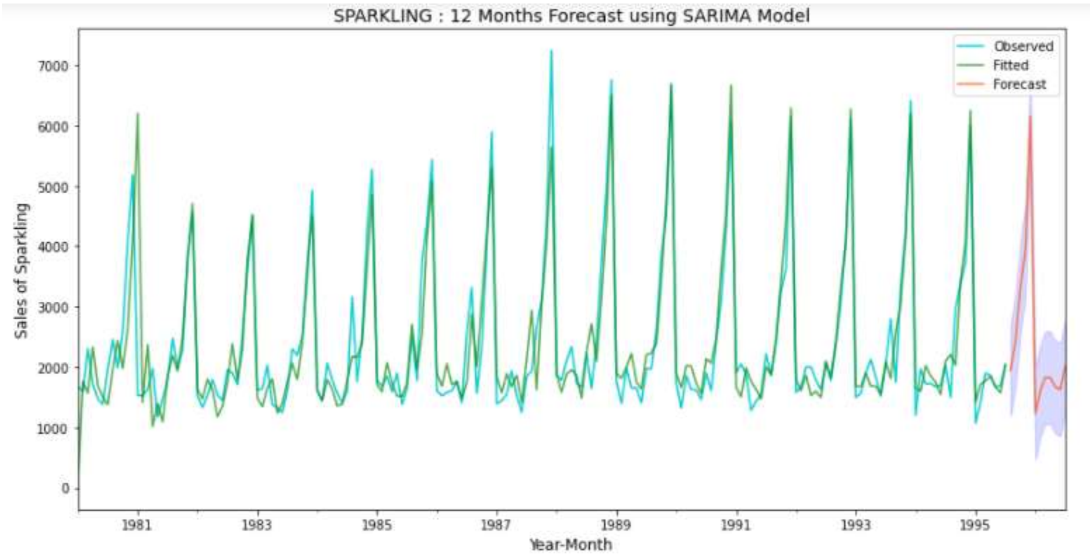
9. Based on the model-building exercise, build the most optimum model(s) on the complete data and predict 12 months into the future with appropriate confidence intervals/bands.

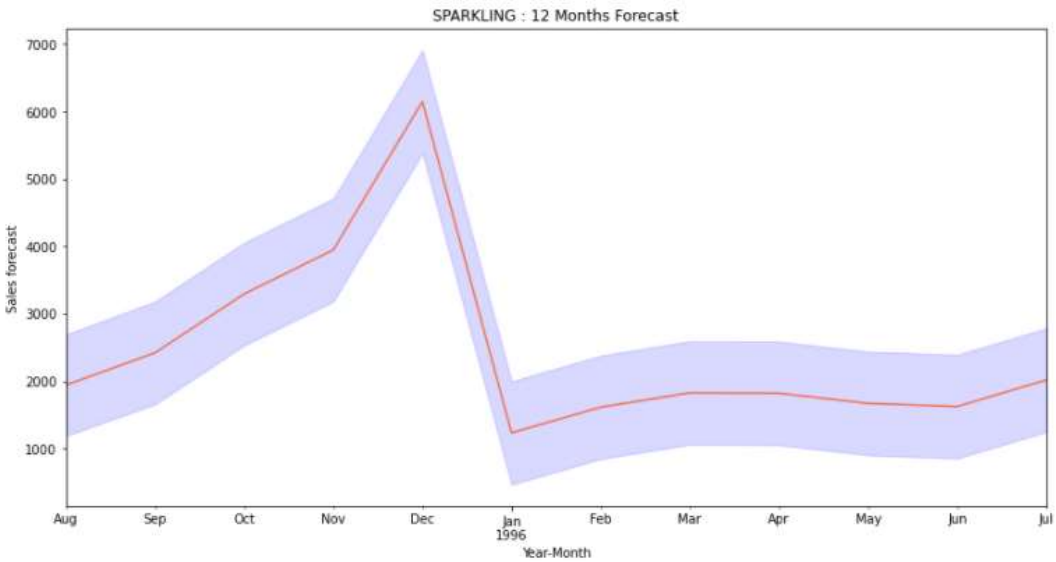
Taking the SARIMA model for the sparkling dataset

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                           Sparkling   No. Observations:                187
Model:             SARIMAX(3, 1, 1)x(1, 1, 1, 12)   Log Likelihood              -1172.978
Date:                             Sat, 25 Sep 2021   AIC                          2359.956
Time:                                     15:49:41   BIC                          2381.439
Sample:                                 01-01-1980   HQIC                         2368.680
                                      - 07-01-1995
Covariance Type:                               opg
==================================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
----------------------------------------------------------------------------------
ar.L1          0.1250      0.082      1.519      0.129      -0.036       0.286
ar.L2         -0.0654      0.099     -0.663      0.507      -0.259       0.128
ar.L3          0.0408      0.084      0.485      0.628      -0.124       0.206
ma.L1         -0.9640      0.033    -28.984      0.000      -1.029      -0.899
ar.S.L12      -0.0984      0.127     -0.772      0.440      -0.348       0.151
ma.S.L12      -0.4852      0.118     -4.110      0.000      -0.717      -0.254
sigma2      1.478e+05    1.3e+04     11.389      0.000    1.22e+05    1.73e+05
======================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):                 46.06
Prob(Q):                              0.93   Prob(JB):                          0.00
Heteroskedasticity (H):               0.89   Skew:                              0.66
Prob(H) (two-sided):                  0.69   Kurtosis:                          5.28
======================================================================================
```



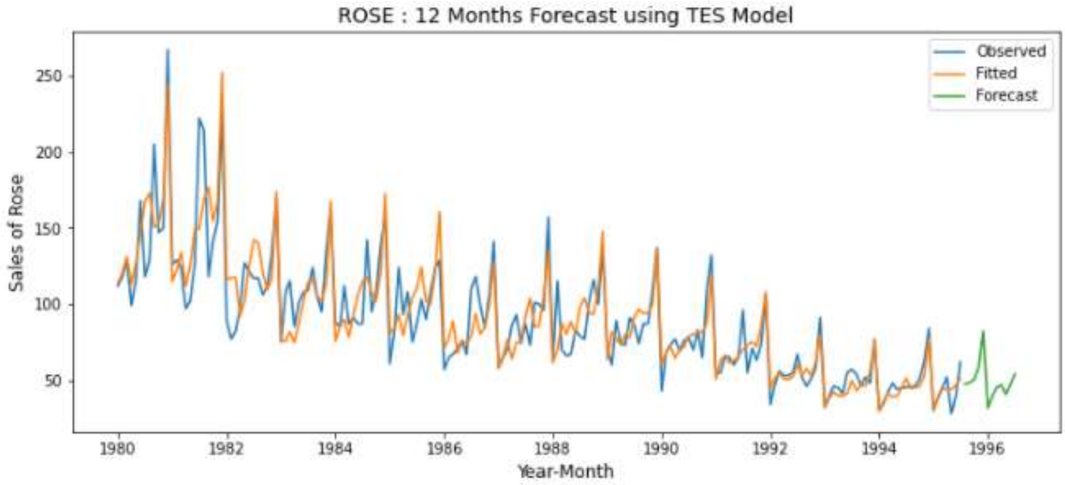
Forecast for the next 12 months

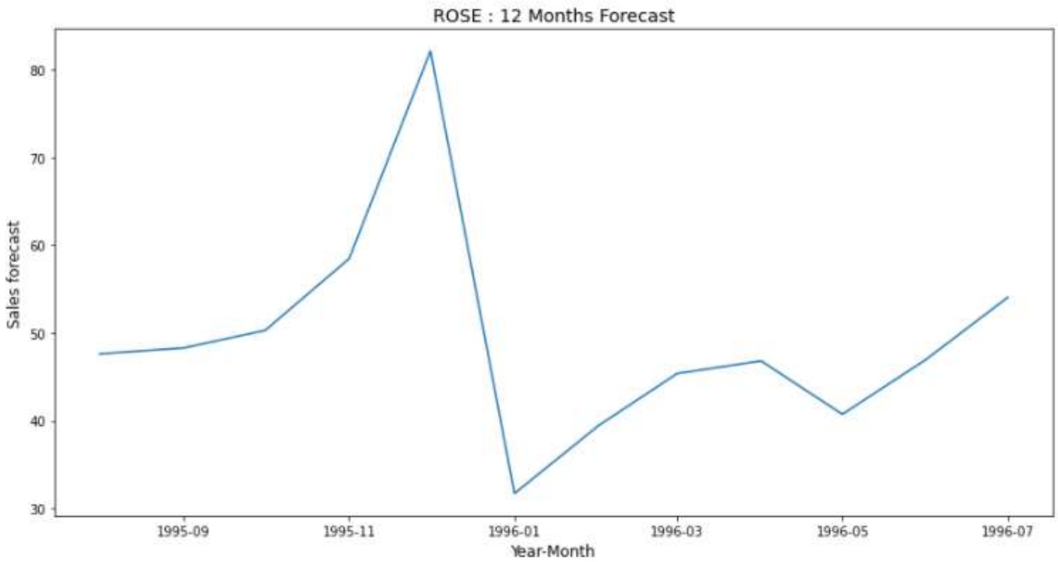




Using Triple Exponential Smoothing for forecasting rose

Forecast for the 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.

SPARKLING

```
Mean sales of sparkling wine 2461 bottles expected in the next 12
months
Max sales of sparkling wine of 6148  on December ,1995
Min sales of sparkling wine of 1230  on January ,1996
Sales of sparkling wine in the year 1995 after August would be 17744
Sales of sparkling wine in the year 1996 till July would be 11792
```

ROSE

```
Mean sales of Rose wine 49 bottles expected in the next 12 months
Max sales of Rose wine of 82 on December ,1995
Min sales of sparkling wine of 32  on January ,1996
Sales of Rose wine in the year 1995 after August would be 287
Sales of Rose wine in the year 1996 till July would be 11792
```

Inference from the models

- The model predicts an average sales of 2461 bottles of sparkling wine in the upcoming year
- The model predicts the max sales of 6148 bottles would be in the month December of 1995
- The model predicts the min sales of 1230 bottles would be in the month January of 1996
- The company can expect a sale of around 17744 bottles of sales in the year 1995 after August and around 11792 bottles till July
- The sales are usually high at the end of the year
- The wine company should look into why the sales of sparkling wine has no upward trend and make necessary promotions and marketing to get more profit and sales
- The model predicts on an average sale of 49 bottles in the upcoming 12 months
- The model predicts the max sales of 82 bottles would be in the month December of 1995
- The model predicts the minimum sales of 32 bottles would be in the month January of 1996
- From the insights rose wine sells very less and there is a negative trend on the product when compared to the sparkling wine.
- In case of the rose wine the company can only expect less sales , they can either make more promotions and offers or discontinue the product.