PROJECT

TIME SERIES FORECASTING

Table of contents

Content	Page No.
Q1. Read the data as an appropriate Time Series data and plot the data.	8
Q2. Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.	10
Q3. Split the data into training and test. The test data should start in 1991.	22
Q4. 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.	24
Q5. Check for the stationarity of the data on which the model is being built 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.	47
Q6. 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.	52
Q7. 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.	60
Q8. Build a table with all the models built along with their corresponding parameters and the respective RMSE values on the test data.	68
Q9. 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.	69
Q10. Comment on the model thus built and report your findings and suggest the measures that the company should be taking for future sales.	70

List of Figures

Figure	Page No.
Figure 1. Time Series Plot.	10
Figure 2. Box Plot of Yearly Sales.	10
Figure 3. Yearly Mean Sales.	11
Figure 4. Yearly Total Wine Sales.	11
Figure 5. Quarterly Resampled Time Series Plot.	12
Figure 6. Box Plot of Quarterly Sales.	12
Figure 7. Quarterly Mean Sales.	14
Figure 8. Quarterly Total Wine Sales.	14
Figure 9. Box Plot of Monthly Sales.	15
Figure 10. Monthly Mean Sales.	16
Figure 11. Monthly Total Sales.	16
Figure 12. Monthly Time Series.	17
Figure 13. Yearly Sales across Months.	18
Figure 14. Empirical Cumulative Distribution of Sales.	19
Figure 15. Trend, Seasonality, and Residual Plots after Additive Decomposition.	20
Figure 16. Trend, Seasonality, and Residual Plots after Multiplicative Decomposition.	21
Figure 17. Train and Test Datasets Plot.	23
Figure 18. Plot of Forecasted Sales in Linear Regression Model.	25
Figure 19. Plot of Forecasted Sales in Naive Model.	26
Figure 20. Plot of Forecasted Sales in Simple Average Model.	27
Figure 21. Plot of Forecasted Sales in Simple Exponential Smoothing Optimized Model.	29
Figure 22. Plot of Forecasted Sales in Simple Exponential Smoothing Iteration Model.	31
Figure 23. Plot of Forecasted Sales in Double Exponential Smoothing Optimized Model.	32
Figure 24. Plot of Forecasted Sales in Double Exponential Smoothing Iteration Model.	34

Figure 25. Plot of Forecasted Sales in TES with the additive trend and additive seasonality optimized model.	35
Figure 26. Plot of Forecasted Sales in TES with the additive trend and additive seasonality iteration model.	37
Figure 27. Plot of Forecasted Sales in TES with the additive trend and multiplicative seasonality optimized model.	38
Figure 28. Plot of Forecasted Sales in TES with the additive trend and multiplicative seasonality iteration model.	40
Figure 29. Plot of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality optimized model.	41
Figure 30. Plot of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality iteration model.	43
Figure 31. Plot of Forecasted Sales in TES with the multiplicative trend and additive seasonality optimized model.	44
Figure 32. Plot of Forecasted Sales in TES with the multiplicative trend and additive seasonality iteration model.	46
Figure 33. Plot of Differenced Time Series	48
Figure 34. Autocorrelation Plot of Whole Time Series	49
Figure 35. Autocorrelation Plot of Differenced Whole Time Series	49
Figure 36. Partial Autocorrelation Plot of Whole Time Series	50
Figure 37. Partial Autocorrelation Plot of Differenced Whole Time Series	50
Figure 38. Plot of Differenced Training Dataset	52
Figure 39. Diagnostics Plots of an automated ARIMA model.	54
Figure 40. Plot of Forecasted Sales in an automated ARIMA model.	55
Figure 41. Diagnostics Plots of an automated SARIMA model.	58
Figure 42. Plot of Forecasted Sales in an automated SARIMA model.	59
Figure 43. Autocorrelation Plot of Differenced Training Dataset	60
Figure 44. Partial Autocorrelation Plot of Differenced Training Dataset	60
Figure 45. Diagnostics Plots of a Manual ARIMA model.	62
Figure 46. Plot of Forecasted Sales in a Manual ARIMA model.	62
Figure 47. Plot of Seasonal Differenced Training Dataset	64
Figure 48. Autocorrelation Plot of Seasonal Differenced Training Dataset	64

Figure 49. Partial Autocorrelation Plot of Seasonal Differenced Training Dataset	64
Figure 50. Diagnostics Plots of a Manual SARIMA model.	66
Figure 51. Plot of Forecasted Sales in a Manual SARIMA model.	67
Figure 52. Plot of Forecasted Sales with 95% Confidence Interval.	69

List of Tables

Table	Page No.
Table 1. Sample of the Time Series.	8
Table 2. Data Types of All Features in the Dataset.	8
Table 3. Sample of the Time Series after Creating Timestamp.	9
Table 4. Five Number Summary.	9
Table 5. Yearly Mean and Total Wine Sales.	12
Table 6. Quarterly Resampled Time Series.	12
Table 7. Quarterly Mean and Total Sales.	14
Table 8. Monthly Mean and Total Sales.	17
Table 9. Yearly Sales across Months.	18
Table 10. Components of both Additive and Multiplicative Decomposition.	22
Table 11. Sample Train and Test Datasets.	23
Table 12. Sample Train and Test Datasets for Linear Regression Model.	24
Table 13. Sample of Forecasted Sales in Linear Regression Model.	25
Table 14. Sample of Forecasted Sales in Naive Model.	26
Table 15. Sample of Forecasted Sales in Simple Average Model.	27
Table 16. Smoothing Parameters in Simple Exponential Smoothing Optimized Model.	28
Table 17. Sample of Forecasted Sales in Simple Exponential Smoothing Optimized Model.	29
Table 18. SES models with low test RMSE values.	30
Table 19. Sample of Forecasted Sales in Simple Exponential Smoothing Iteration Model.	30
Table 20. Smoothing Parameters in Double Exponential Smoothing Optimized Model.	32

Table 21. Sample of Forecasted Sales in Double Exponential Smoothing Optimized Model.	32
Table 22. DES models with low test RMSE values.	33
Table 23. Sample of Forecasted Sales in Double Exponential Smoothing Iteration Model.	33
Table 24. Smoothing Parameters in TES with the additive trend and additive seasonality optimized model.	34
Table 25. Sample of Forecasted Sales in TES with the additive trend and additive seasonality optimized model.	35
Table 26. TES with the additive trend and additive seasonality models with low test RMSE values.	36
Table 27. Sample of Forecasted Sales in TES with the additive trend and additive seasonality Iteration Model.	36
Table 28. Smoothing Parameters in TES with the additive trend and multiplicative seasonality optimized model.	37
Table 29. Sample of Forecasted Sales in TES with the additive trend and multiplicative seasonality optimized model.	38
Table 30. TES with the additive trend and multiplicative seasonality models with low test RMSE values.	39
Table 31. Sample of Forecasted Sales in TES with the additive trend and multiplicative seasonality Iteration Model.	39
Table 32. Smoothing Parameters in TES with the multiplicative trend and multiplicative seasonality optimized model.	40
Table 33. Sample of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality optimized model.	41
Table 34. TES with the multiplicative trend and multiplicative seasonality models with low test RMSE values.	42
Table 35. Sample of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality Iteration Model.	42
Table 36. Smoothing Parameters in TES with the multiplicative trend and additive seasonality optimized model.	43
Table 37. Sample of Forecasted Sales in TES with the multiplicative trend and additive seasonality optimized model.	44

Table 38. TES with the multiplicative trend and additive seasonality models with low test RMSE values.	45
Table 39. Sample of Forecasted Sales in TES with the multiplicative trend and additive seasonality Iteration Model.	45
Table 40. Top five ARIMA models with low AIC values.	53
Table 41. Summary of an automated ARIMA model.	53
Table 42. Sample of Forecasted Sales in an automated ARIMA Model.	55
Table 43. Top five SARIMA models with low AIC values.	56
Table 44. Summary of an automated SARIMA model.	57
Table 45. Sample of Forecasted Sales in an automated SARIMA Model.	58
Table 46. Summary of a Manual ARIMA model.	61
Table 47. Sample of Forecasted Sales in a Manual ARIMA Model.	62
Table 48. Summary of a Manual SARIMA model.	65
Table 49. Sample of Forecasted Sales in a Manual SARIMA Model.	66
Table 50. All Forecast Models with Respective Parameters and RMSE’s	68
Table 51. Forecasted Sales with 95% Confidence Interval.	69

PROBLEM 1

Problem Statement

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.

Q1. Read the data as an appropriate Time Series data and plot the data.

Sample of the Time Series

	YearMonth	Sparkling
0	1980-01	1686
1	1980-02	1591
2	1980-03	2304
3	1980-04	1712
4	1980-05	1471
5	1980-06	1377
6	1980-07	1966
7	1980-08	2453
8	1980-09	1984
9	1980-10	2596

Table 1. Sample of the Time Series.

Basic Information of the Dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 187 entries, 0 to 186
Data columns (total 2 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   YearMonth  187 non-null    object
 1   Sparkling  187 non-null    int64
dtypes: int64(1), object(1)
memory usage: 3.0+ KB
```

Data Types of Variables

	Data Type
YearMonth	object
Sparkling	int64

	Data Type
YearMonth	datetime64[ns]
Sparkling	int64

Table 2. Data Types of All Features in the Dataset.

In original data, Year-Month columns are of object datatype. It is converted into a time series stamp and the same is shown in the above table.

YearMonth	Sparkling
1980-01-01	1686
1980-02-01	1591
1980-03-01	2304
1980-04-01	1712
1980-05-01	1471
1980-06-01	1377
1980-07-01	1966
1980-08-01	2453
1980-09-01	1984
1980-10-01	2596

Table 3. Sample of the Time Series after Creating Timestamp.

Insights

1. There are 2 features (columns) with 187 observations (rows) in the dataset.
2. The dataset has one numerical feature i.e., It indicates the number of sparkling wine sales. This is the forecast variable.
3. Another feature is the year-month timestamp and it is made as an index for this time series.
4. There **are no null values** in the dataset.

Description of the Dataset

Five number summaries

	count	mean	std	min	25%	50%	75%	max
Sparkling	187.0	2402.417112	1295.11154	1070.0	1605.0	1874.0	2549.0	7242.0

Table 4. Five Number Summary.

Insights:

1. The Minimum number of sales in a month is 1070.
2. The Maximum number of sales in a month is 7242.
3. Mean sales per month is 2402.

Plot the Time Series to understand the behavior of the data

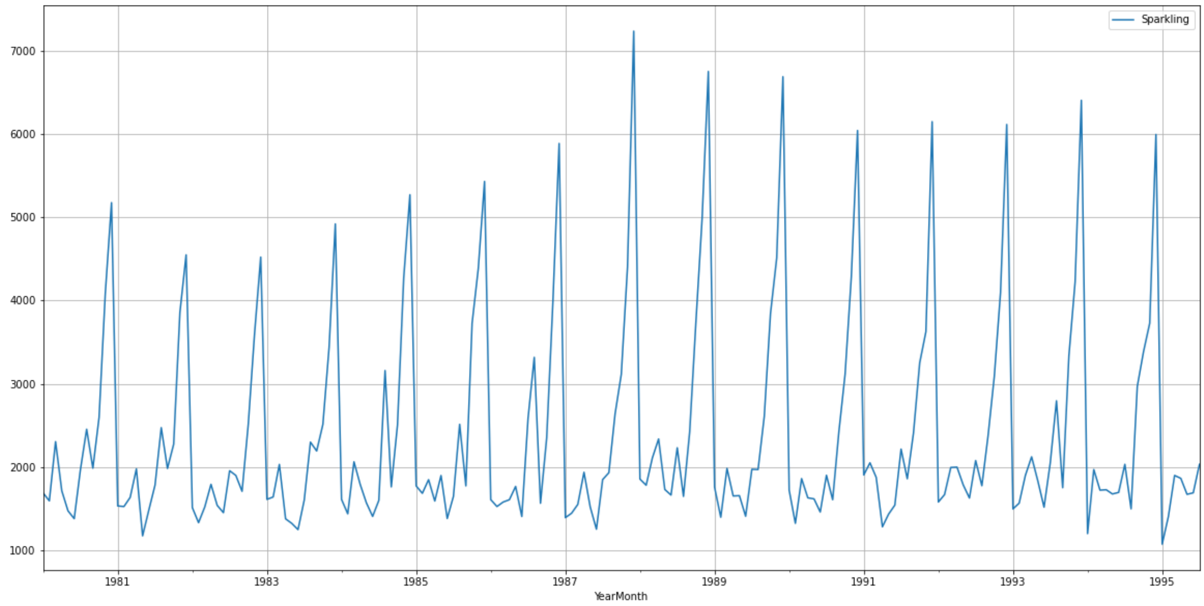


Figure 1. Time Series Plot.

- From the above plot, it can be noticed that wine sales **data has seasonality**. Every December month of the year sales is reaching the maximum value.

Q2. Perform appropriate Exploratory Data Analysis to understand the data and also perform decomposition.

EXPLORATORY DATA ANALYSIS

- There **are no null values** in the dataset.

Yearly Sales Analysis - Box Plot

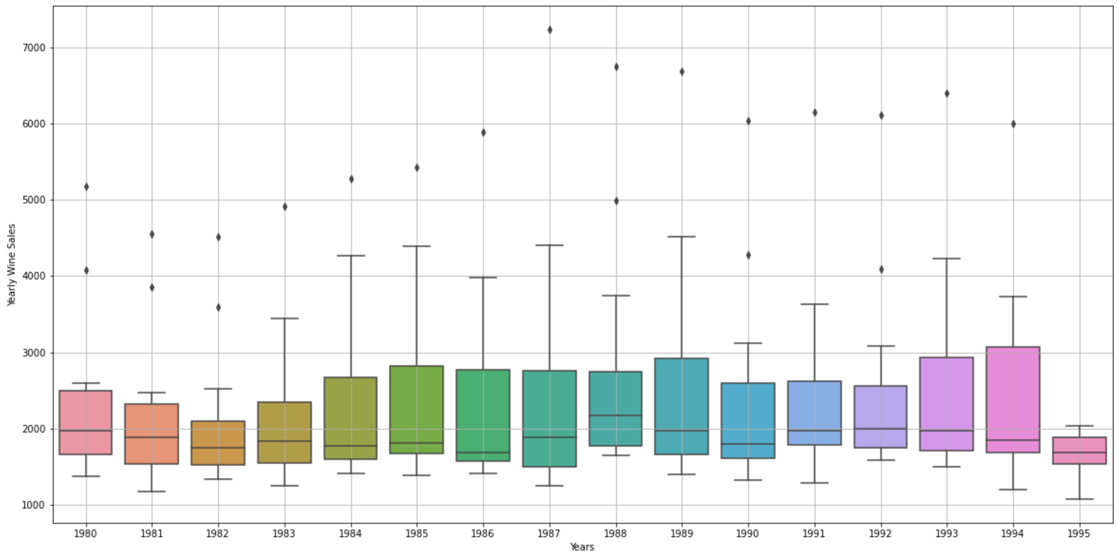


Figure 2. Box Plot of Yearly Sales.

- As we got to know from the Time Series plot, the boxplots over here also indicate a measure of the trend is present. Also, we see that the sales of wine have some outliers for certain years.

Yearly Mean Sales

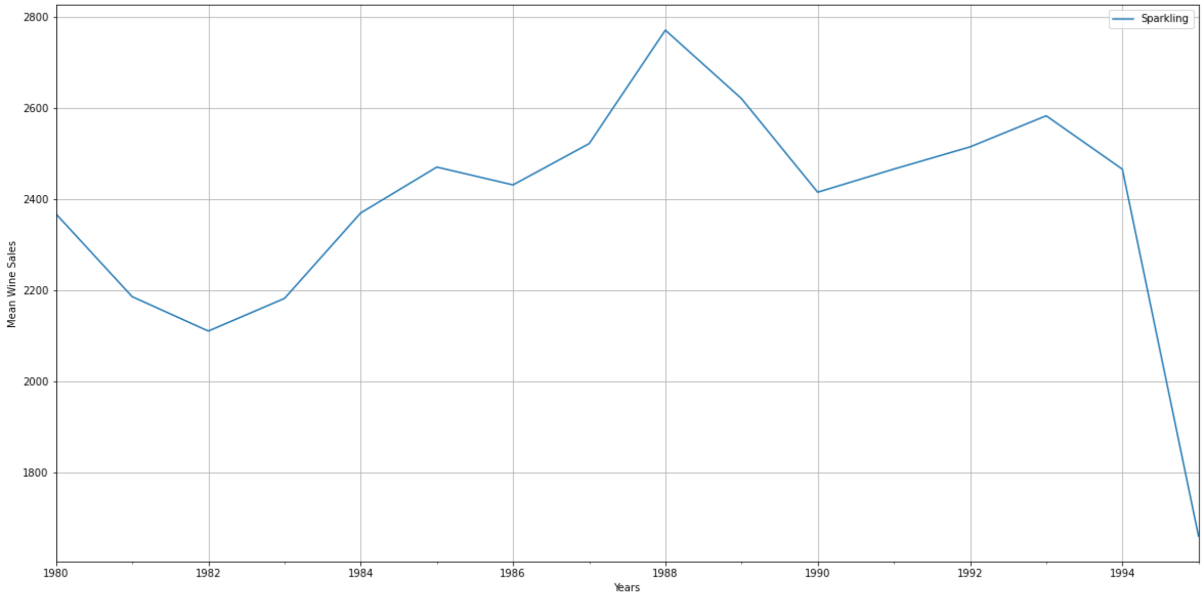


Figure 3. Yearly Mean Sales.

Insights

- Mean wine sales in a year is maximum in 1988.
- Mean wine sales in a year is minimum in 1995.
- Initially mean wine sales are increasing up to 1988 and then start decreasing.

Yearly Total Sales

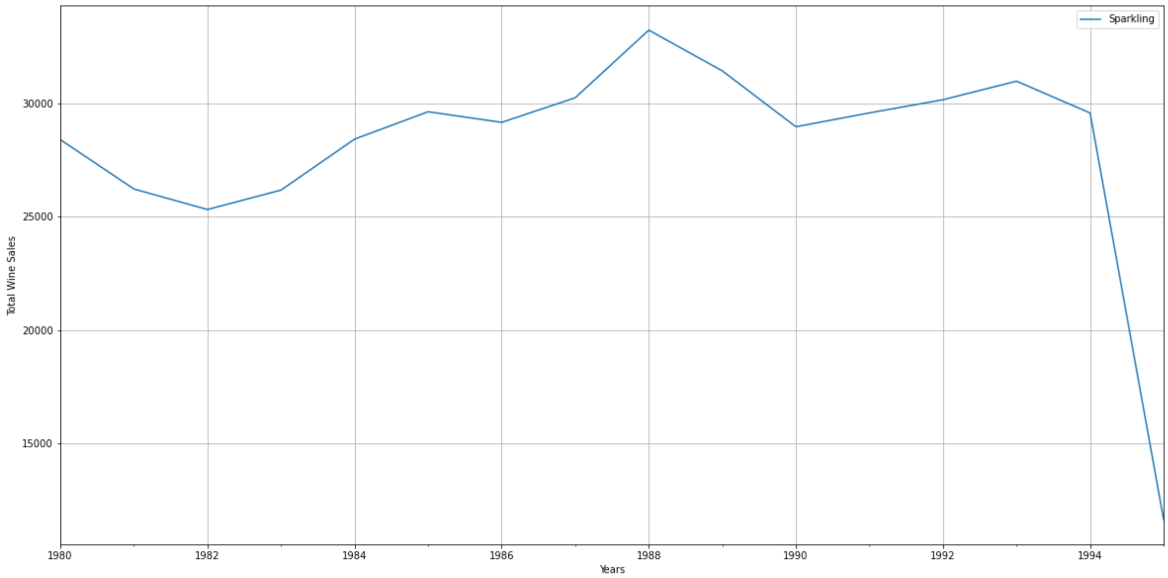


Figure 4. Yearly Total Wine Sales.

Year	Mean Sales	Total Sales
1980	2367.2	28406
1981	2185.6	26227
1982	2110.1	25321
1983	2181.7	26180
1984	2369.2	28431
1985	2470.0	29640
1986	2430.8	29170
1987	2521.5	30258
1988	2770.5	33246
1989	2620.2	31443
1990	2414.8	28977
1991	2465.6	29587
1992	2514.2	30171
1993	2582.6	30991
1994	2465.3	29584
1995	1660.0	11620

Table 5. Yearly Mean and Total Wine Sales.

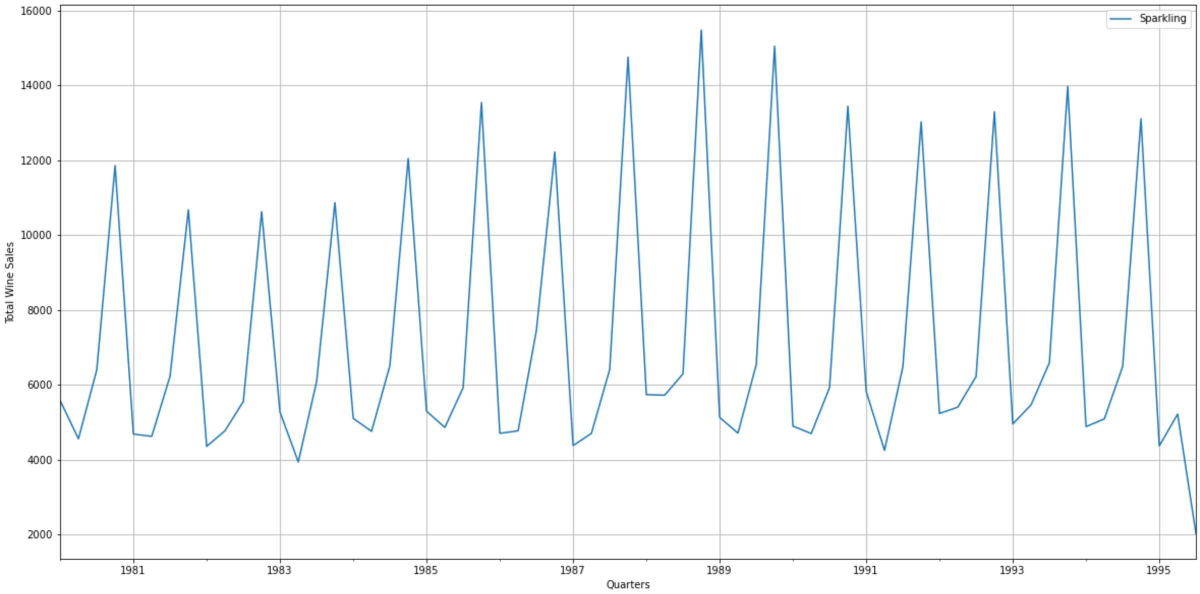
Insights

- Total wine sales in a year are following almost the same pattern as that of the mean wine sales in a year.
- Total wine sales in a year are maximum in 1988.
- Total wine sales in a year are minimum in 1995.
- Initially total wine sales are increasing up to 1988 and then start decreasing.

Quarterly Sales Analysis

YearMonth	Sparkling
1980-03-31	5581
1980-06-30	4560
1980-09-30	6403
1980-12-31	11862
1981-03-31	4686

Table 6. Quarterly Resampled Time Series.



. Figure 5. Quarterly Resampled Time Series Plot.

Box Plot of Quarterly Sales

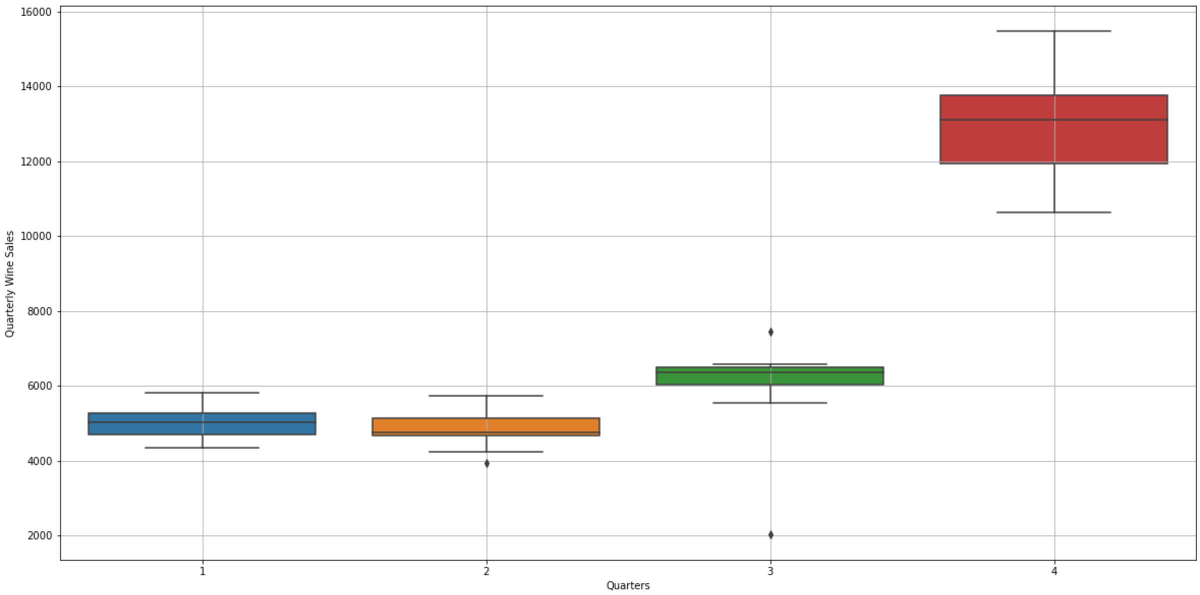


Figure 6. Box Plot of Quarterly Sales.

Insights

- Sales are increasing gradually from quarter1 to quarter 3 and then increase exponentially from quarter 3 to quarter 4.
- Sales of wine have some outliers in the second and third quarters.

Quarterly Mean Sales

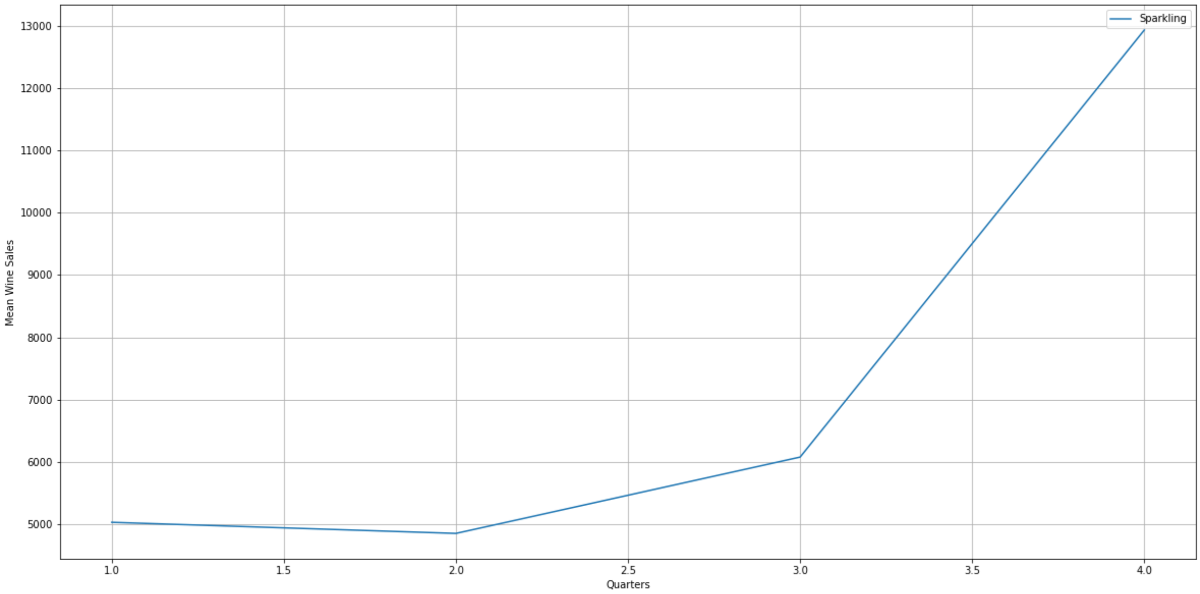


Figure 7. Quarterly Mean Sales.

Quarterly Total Sales

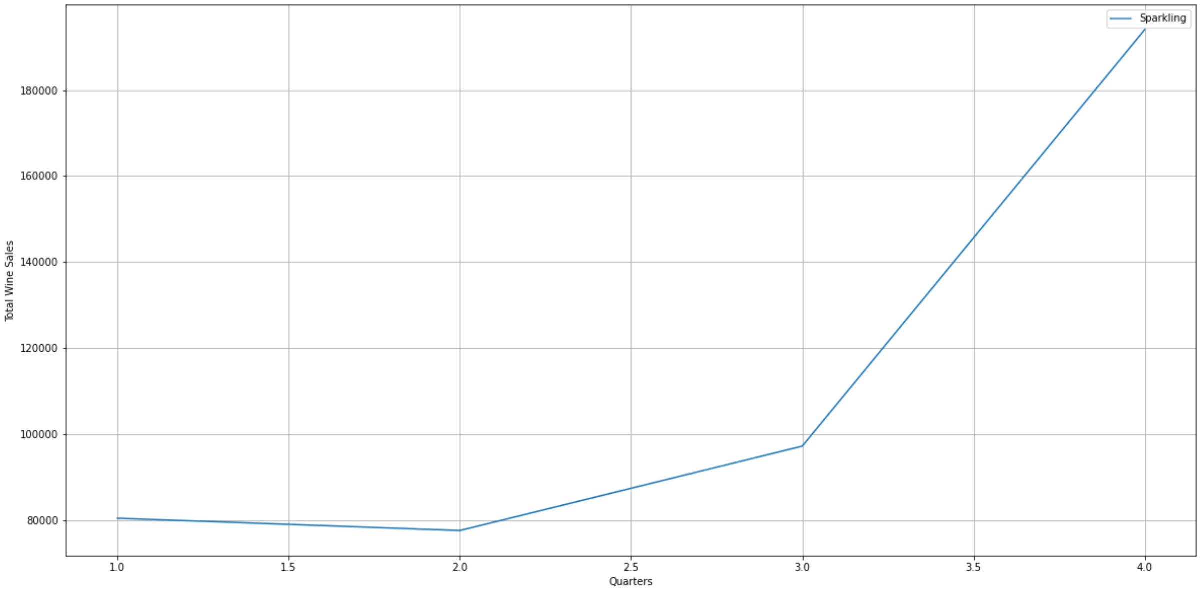


Figure 8. Quarterly Total Wine Sales.

Quarter	Mean Sales	Total Sales
1	5027.2	80435
2	4848.2	77572
3	6074.9	97198
4	12936.5	194047

Table 7. Quarterly Mean and Total Sales.

Insights

- Mean sales are increasing gradually from quarter1 to quarter 3 and then increase exponentially from quarter 3 to quarter 4.
- Total wine sales in a quarter are following almost the same pattern as that of the mean wine sales in a quarter.
- Total sales are increasing gradually from quarter1 to quarter 3 and then increase exponentially from quarter 3 to quarter 4.

Monthly Sales Analysis

Box Plot of Monthly Sales

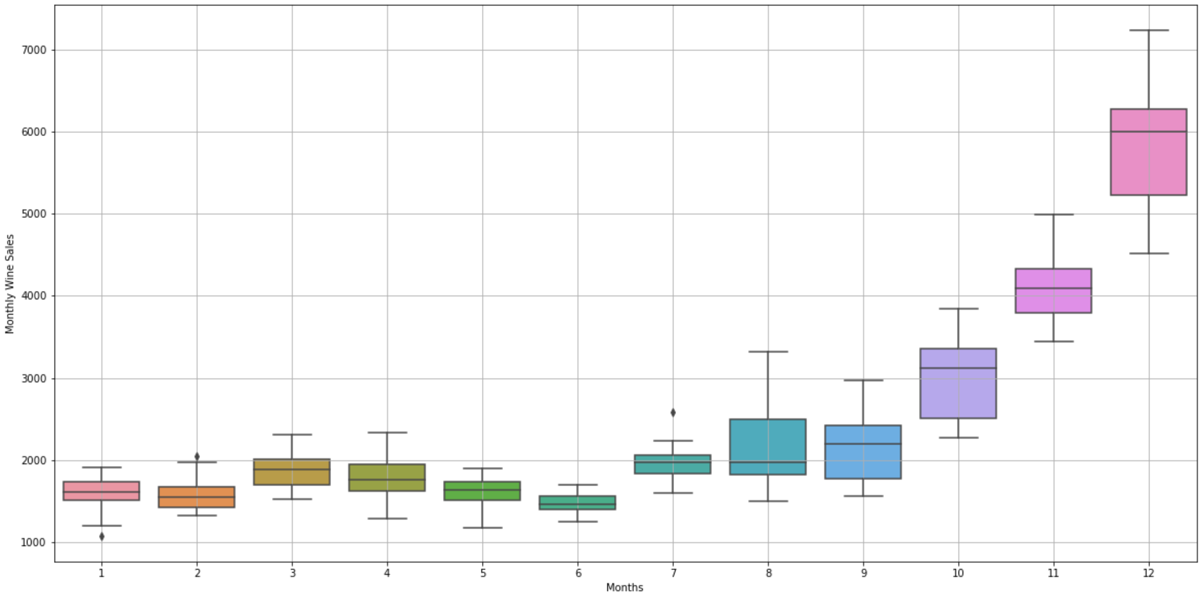


Figure 9. Box Plot of Monthly Sales.

Insights

- Sales are not following any significant pattern in the first half of the year and the second half increasing exponentially.
- Sales of wine have some outliers in a few months.

Monthly Mean Sales

- Mean sales in a month are not following any significant pattern up to September month and then increase exponentially.

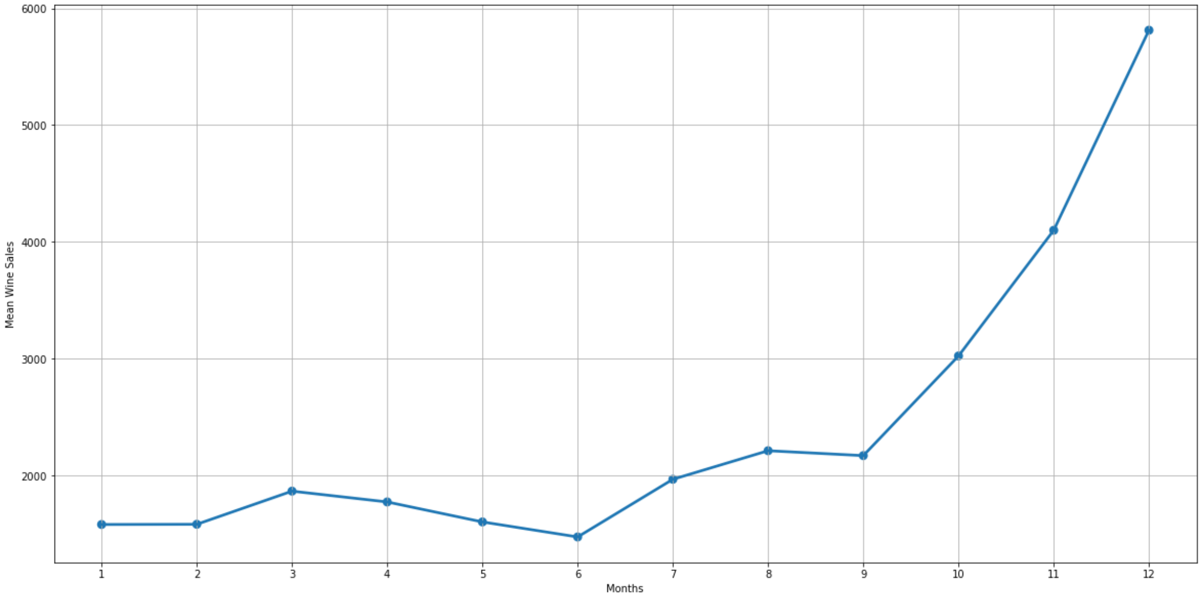


Figure 10. Monthly Mean Sales.

Monthly Total Sales

- Total sales in a month are following almost the same pattern as that of the mean sales in a month.
- Total sales in a month are not following any significant pattern up to September month and then increase exponentially.

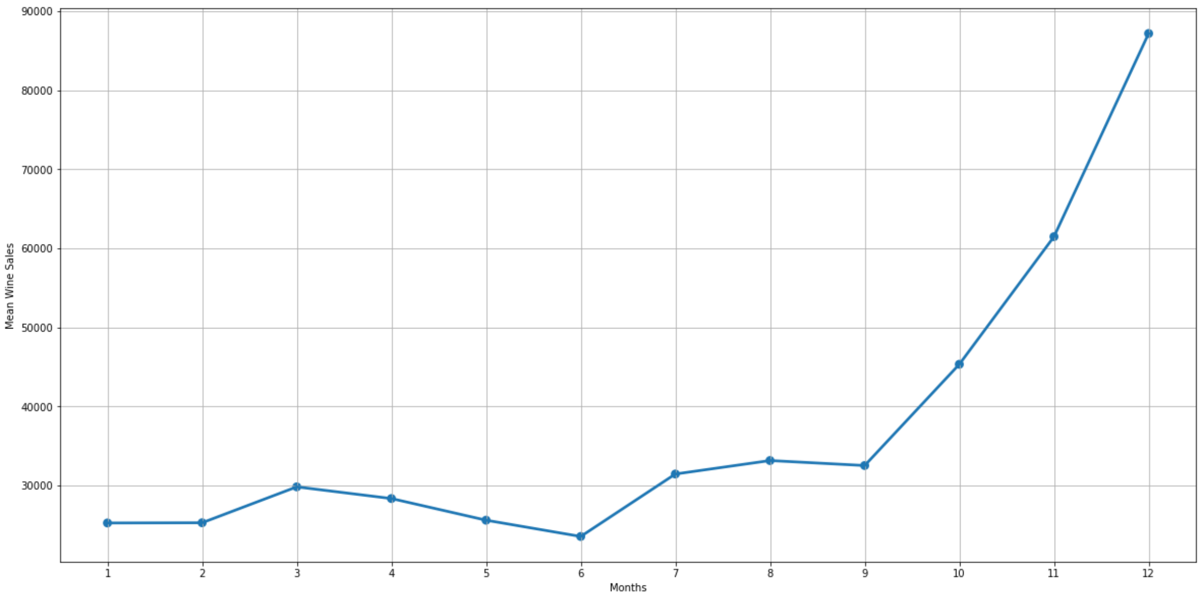


Figure 11. Monthly Total Sales.

Month	Mean Sales	Total Sales
1	1580.0	25280
2	1581.7	25307
3	1865.5	29848
4	1773.1	28369
5	1601.9	25631
6	1473.2	23572
7	1967.1	31474
8	2211.8	33177
9	2169.8	32547
10	3023.5	45353
11	4099.8	61497
12	5813.1	87197

Table 8. Monthly Mean and Total Sales.

Monthly Time Series

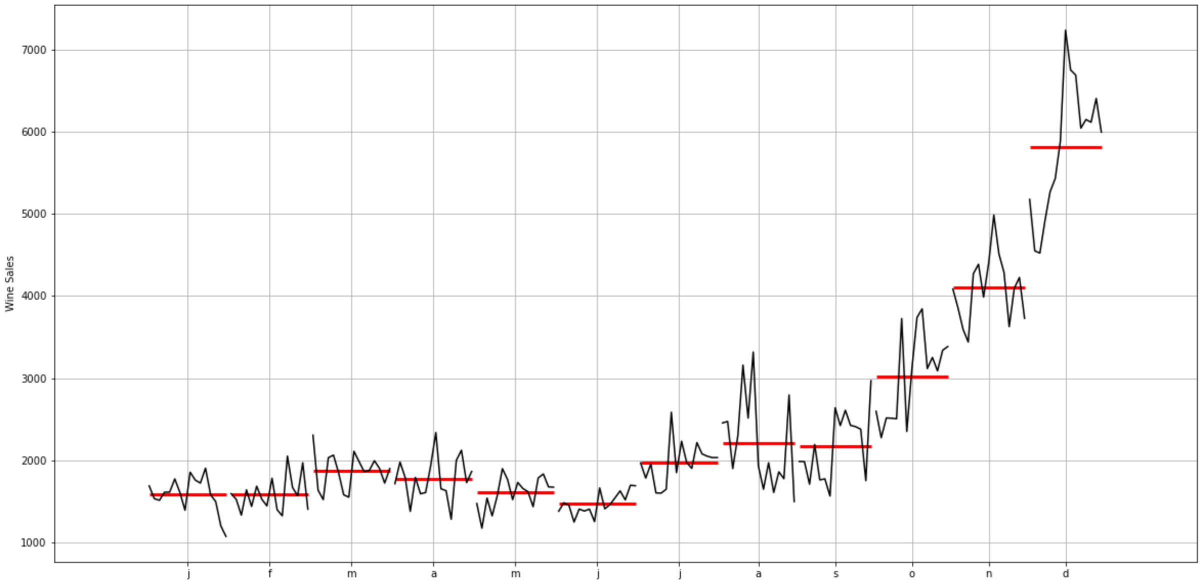


Figure 12. Monthly Time Series.

- Sales are not following any significant pattern up to September month and then increase exponentially.

Yearly Sales across Months

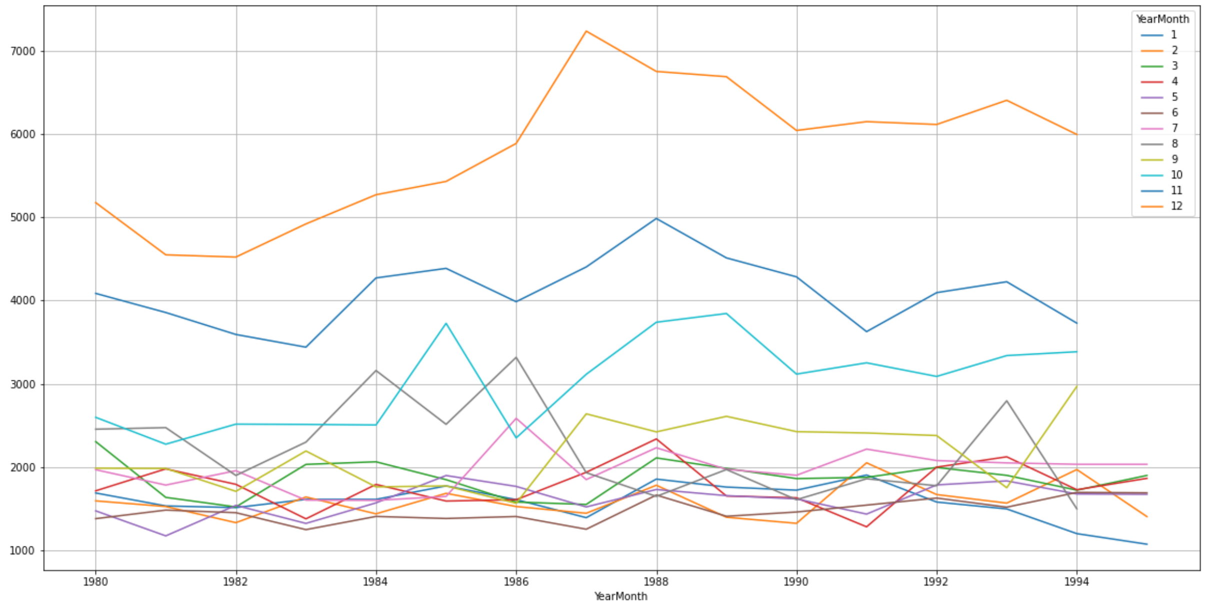


Figure 13. Yearly Sales across Months.

YearMonth	1	2	3	4	5	6	7	8	9	10	11	12
YearMonth												
1980	1686.0	1591.0	2304.0	1712.0	1471.0	1377.0	1966.0	2453.0	1984.0	2596.0	4087.0	5179.0
1981	1530.0	1523.0	1633.0	1976.0	1170.0	1480.0	1781.0	2472.0	1981.0	2273.0	3857.0	4551.0
1982	1510.0	1329.0	1518.0	1790.0	1537.0	1449.0	1954.0	1897.0	1706.0	2514.0	3593.0	4524.0
1983	1609.0	1638.0	2030.0	1375.0	1320.0	1245.0	1600.0	2298.0	2191.0	2511.0	3440.0	4923.0
1984	1609.0	1435.0	2061.0	1789.0	1567.0	1404.0	1597.0	3159.0	1759.0	2504.0	4273.0	5274.0
1985	1771.0	1682.0	1846.0	1589.0	1896.0	1379.0	1645.0	2512.0	1771.0	3727.0	4388.0	5434.0
1986	1606.0	1523.0	1577.0	1605.0	1765.0	1403.0	2584.0	3318.0	1562.0	2349.0	3987.0	5891.0
1987	1389.0	1442.0	1548.0	1935.0	1518.0	1250.0	1847.0	1930.0	2638.0	3114.0	4405.0	7242.0
1988	1853.0	1779.0	2108.0	2336.0	1728.0	1661.0	2230.0	1645.0	2421.0	3740.0	4988.0	6757.0
1989	1757.0	1394.0	1982.0	1650.0	1654.0	1406.0	1971.0	1968.0	2608.0	3845.0	4514.0	6694.0
1990	1720.0	1321.0	1859.0	1628.0	1615.0	1457.0	1899.0	1605.0	2424.0	3116.0	4286.0	6047.0
1991	1902.0	2049.0	1874.0	1279.0	1432.0	1540.0	2214.0	1857.0	2408.0	3252.0	3627.0	6153.0
1992	1577.0	1667.0	1993.0	1997.0	1783.0	1625.0	2076.0	1773.0	2377.0	3088.0	4096.0	6119.0
1993	1494.0	1564.0	1898.0	2121.0	1831.0	1515.0	2048.0	2795.0	1749.0	3339.0	4227.0	6410.0
1994	1197.0	1968.0	1720.0	1725.0	1674.0	1693.0	2031.0	1495.0	2968.0	3385.0	3729.0	5999.0
1995	1070.0	1402.0	1897.0	1862.0	1670.0	1688.0	2031.0	NaN	NaN	NaN	NaN	NaN

Table 9. Yearly Sales across Months.

- From the above plot and pivot table, it can be noticed that from September month onwards sales increase drastically and reaches the maximum in December month.

Empirical Cumulative Distribution of Sales

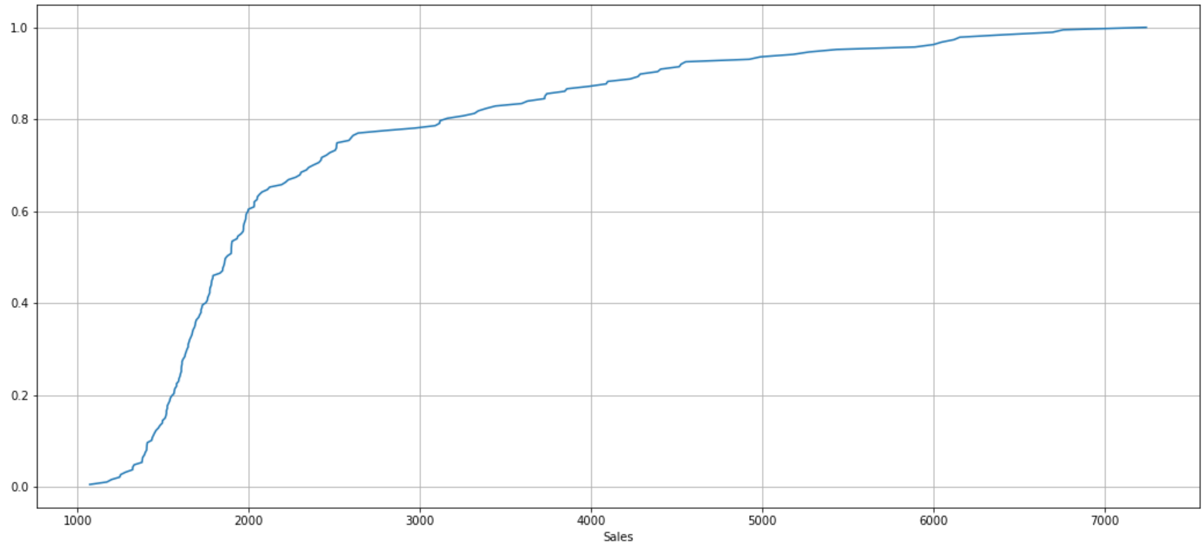


Figure 14. Empirical Cumulative Distribution of Sales.

- From the above plot, the probability to have sales more/less than a certain value can be found. For example, the probability to have sales of more than 2000 is 0.4.

DECOMPOSITION

Additive Decomposition

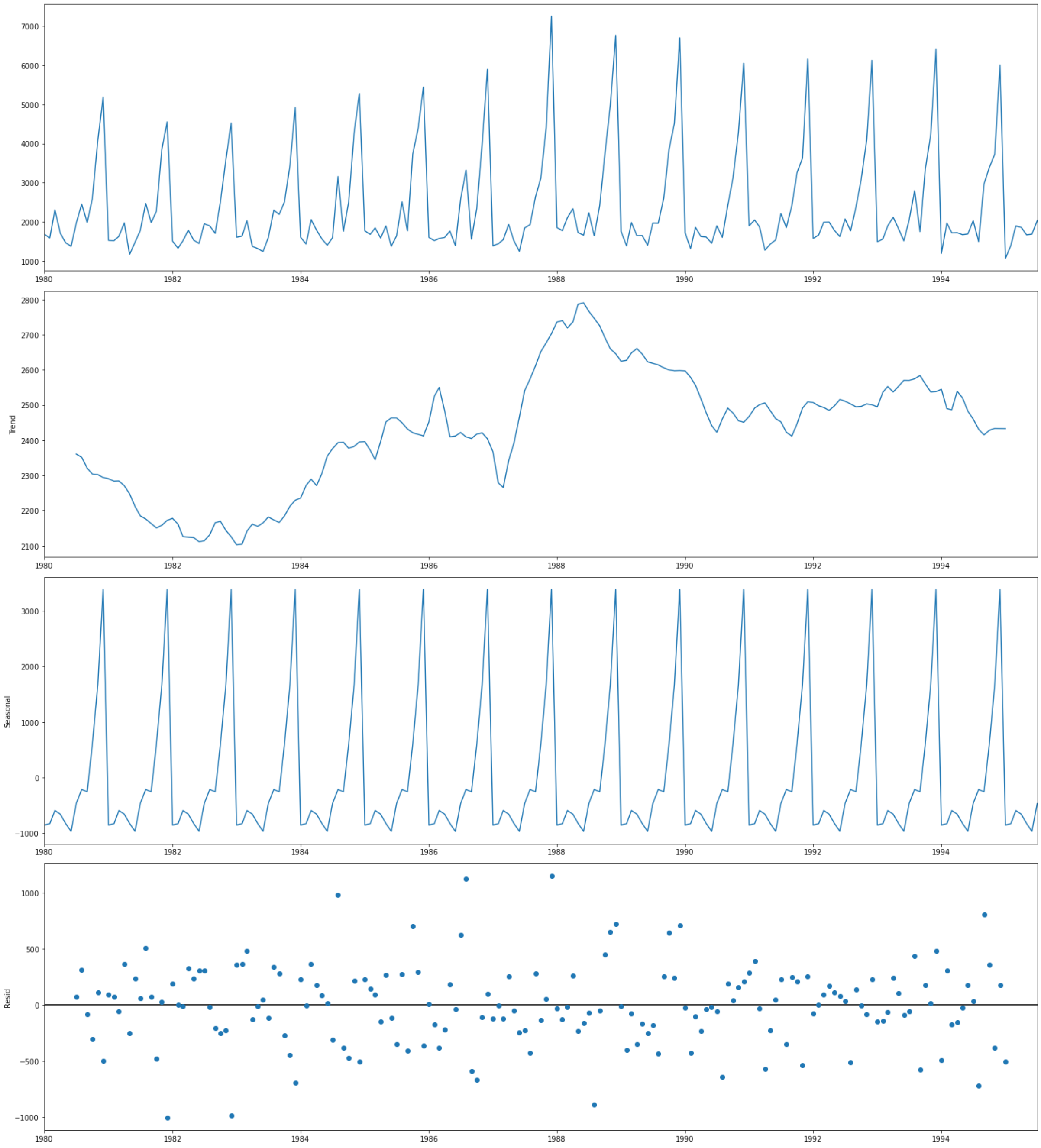


Figure 15. Trend, Seasonality, and Residual Plots after Additive Decomposition.

Insights

- Increasing trend is observed up to the year 1988 atter that decreasing trend is observed.
- Yearly seasonality is present in the time series. Sales reach to maximum in December every year.
- Residuals are not following any pattern.

Multiplicative Decomposition

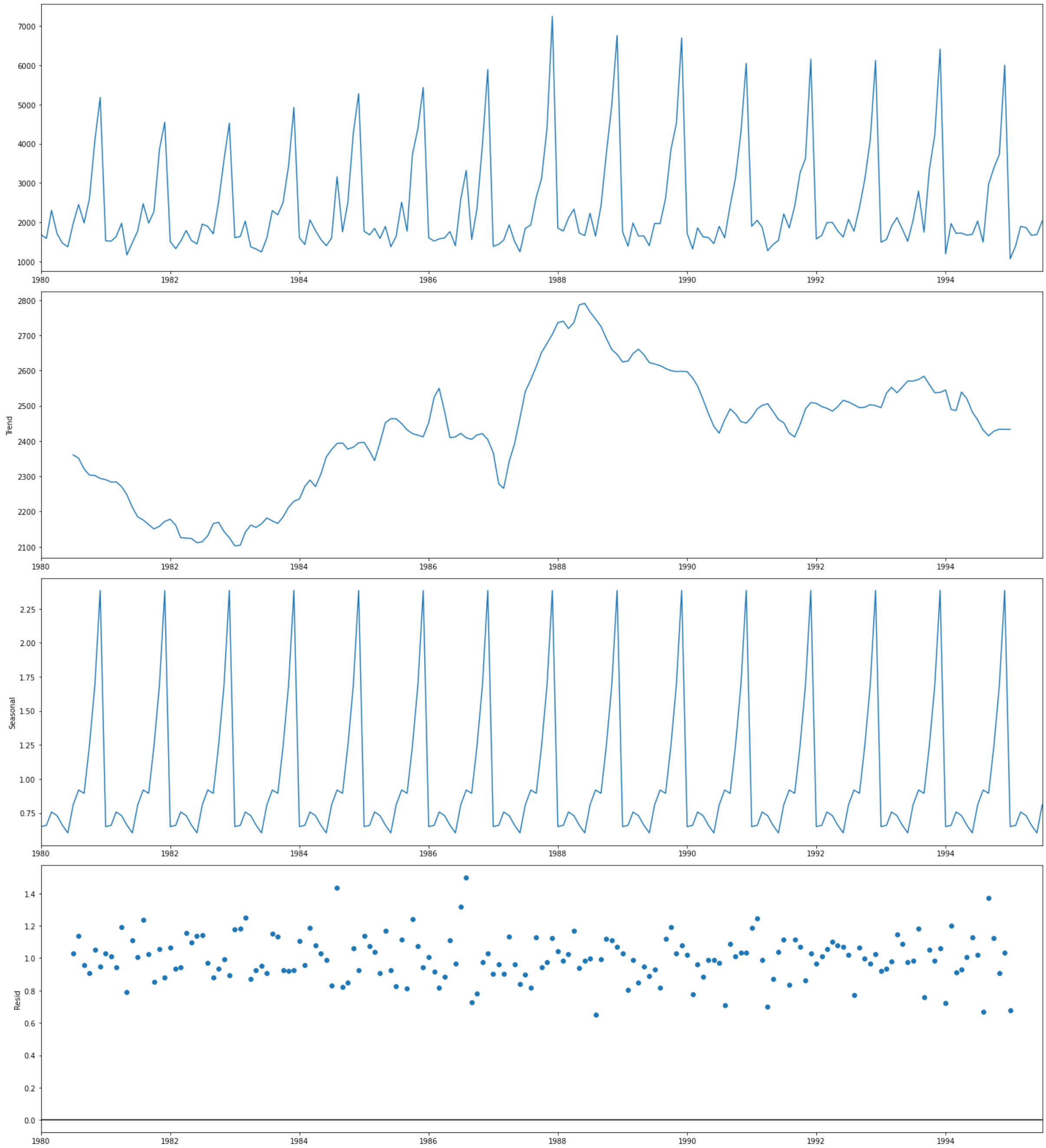


Figure 16. Trend, Seasonality, and Residual Plots after Multiplicative Decomposition.

Insights

- Increasing trend is observed up to the year 1988 atter that decreasing trend is observed.
- Yearly seasonality is present in the time series. Sales reach to maximum in December every year.
- Residuals are not following any pattern.
- As residuals are randomly distributed in both additive and multiplicative decomposition. It is recommended to try both of them in forecasting to find out the best model.

YearMonth	trend	seasonal	resid
1980-01-01	NaN	-854.3	NaN
1980-02-01	NaN	-830.4	NaN
1980-03-01	NaN	-592.4	NaN
1980-04-01	NaN	-658.5	NaN
1980-05-01	NaN	-824.4	NaN
1980-06-01	NaN	-967.4	NaN
1980-07-01	2360.7	-465.5	70.8
1980-08-01	2351.3	-214.3	316.0
1980-09-01	2320.5	-254.7	-81.9
1980-10-01	2303.6	599.8	-307.4
1980-11-01	2302.0	1675.1	109.9
1980-12-01	2293.8	3387.0	-501.8
1981-01-01	2290.4	-854.3	93.9
1981-02-01	2283.5	-830.4	69.9
1981-03-01	2284.1	-592.4	-58.8
1981-04-01	2270.5	-658.5	363.9
1981-05-01	2247.5	-824.4	-253.1
1981-06-01	2211.8	-967.4	235.7
1981-07-01	2184.7	-465.5	61.8
1981-08-01	2175.8	-214.3	510.5
1981-09-01	2163.0	-254.7	72.7
1981-10-01	2150.4	599.8	-477.2
1981-11-01	2158.0	1675.1	24.0

YearMonth	trend	seasonal	resid
1980-01-01	NaN	0.6	NaN
1980-02-01	NaN	0.7	NaN
1980-03-01	NaN	0.8	NaN
1980-04-01	NaN	0.7	NaN
1980-05-01	NaN	0.7	NaN
1980-06-01	NaN	0.6	NaN
1980-07-01	2360.7	0.8	1.0
1980-08-01	2351.3	0.9	1.1
1980-09-01	2320.5	0.9	1.0
1980-10-01	2303.6	1.2	0.9
1980-11-01	2302.0	1.7	1.1
1980-12-01	2293.8	2.4	0.9
1981-01-01	2290.4	0.6	1.0
1981-02-01	2283.5	0.7	1.0
1981-03-01	2284.1	0.8	0.9
1981-04-01	2270.5	0.7	1.2
1981-05-01	2247.5	0.7	0.8
1981-06-01	2211.8	0.6	1.1
1981-07-01	2184.7	0.8	1.0
1981-08-01	2175.8	0.9	1.2
1981-09-01	2163.0	0.9	1.0
1981-10-01	2150.4	1.2	0.9
1981-11-01	2158.0	1.7	1.1

Table 10. Components of both Additive and Multiplicative Decomposition.

Q3. Split the data into training and test. The test data should start in 1991.

- Data is divided into train and test sets. The sales before 1991 are included in the train set. The sales from 1991 are included in the test set.
- Test Dataset has 4 years 7 months data. It has four full cycles.
- Train set has 132 records and test set has 55 records.

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

First Few Rows of Train Data

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

First Few Rows of Test Data

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

Last Few Rows of Train Data

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

Last Few Rows of Test Data

Table 11. Sample Train and Test Datasets.

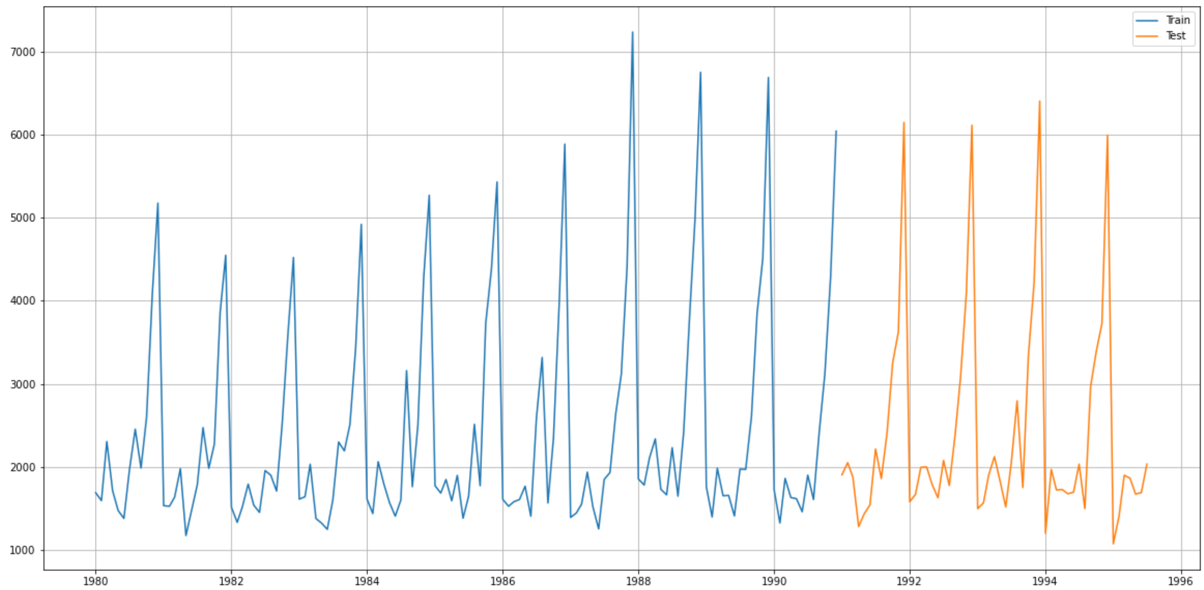


Figure 17. Train and Test Datasets Plot.

Q4. 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, 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

- In the linear regression model, we are going **to regress the 'Sparkling' variable against the order of the occurrence**. For this, we need to add time instance to our training data before fitting it into a linear regression.
- We have successfully generated the numerical time instance order for both the training and test set.

```
Training Time
 [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
 [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]
```

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

First Few Rows of Train Data

	Sparkling	time
YearMonth		
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

First Few Rows of Test Data

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

Last Few Rows of Train Data

	Sparkling	time
YearMonth		
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

Last Few Rows of Test Data

Table 12. Sample Train and Test Datasets for Linear Regression Model.

YearMonth	Sparkling	forecast_lr
1991-01-01	1902	2791.7
1991-02-01	2049	2797.5
1991-03-01	1874	2803.3
1991-04-01	1279	2809.2
1991-05-01	1432	2815.0
1991-06-01	1540	2820.8
1991-07-01	2214	2826.6
1991-08-01	1857	2832.5
1991-09-01	2408	2838.3
1991-10-01	3252	2844.1

Table 13. Sample of Forecasted Sales in Linear Regression Model.

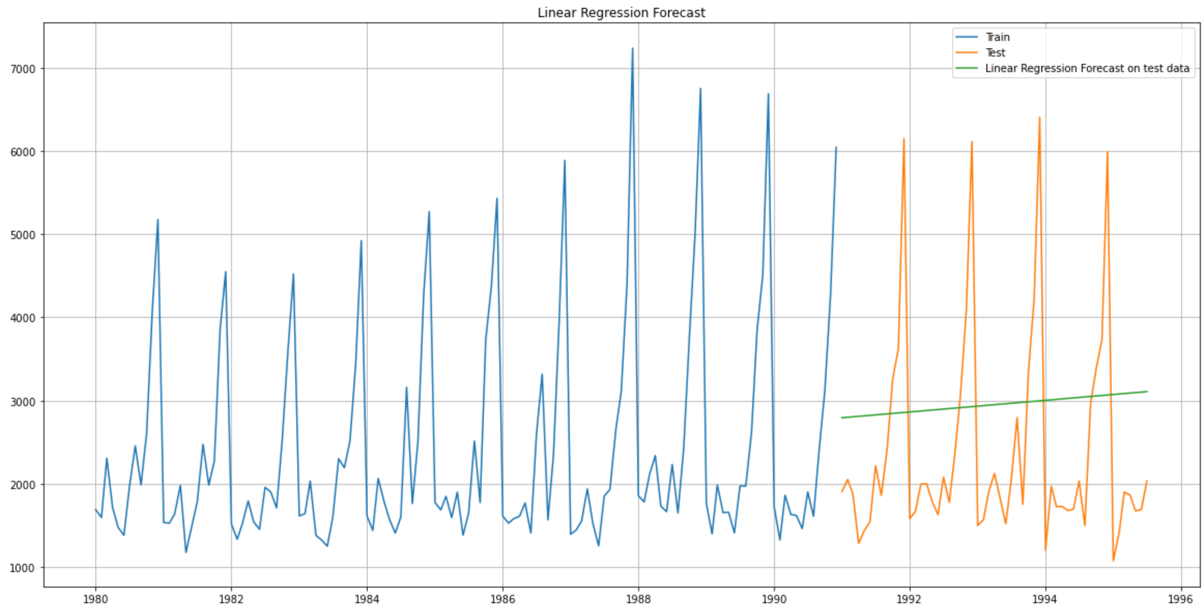


Figure 18. Plot of Forecasted Sales in Linear Regression Model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Linear Regression may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE’s of different forecast models in subsequent sections.
- RMSE in Linear Regression Model is 1389.1

Model 2: Naïve Approach

- In this model, the forecast for tomorrow is the same as today and the forecast for the day after tomorrow is tomorrow therefore the forecast for the day after tomorrow is also today.

YearMonth	Sparkling	forecast_naive
1991-01-01	1902	6047
1991-02-01	2049	6047
1991-03-01	1874	6047
1991-04-01	1279	6047
1991-05-01	1432	6047
1991-06-01	1540	6047
1991-07-01	2214	6047
1991-08-01	1857	6047
1991-09-01	2408	6047
1991-10-01	3252	6047

Table 14. Sample of Forecasted Sales in Naive Model.

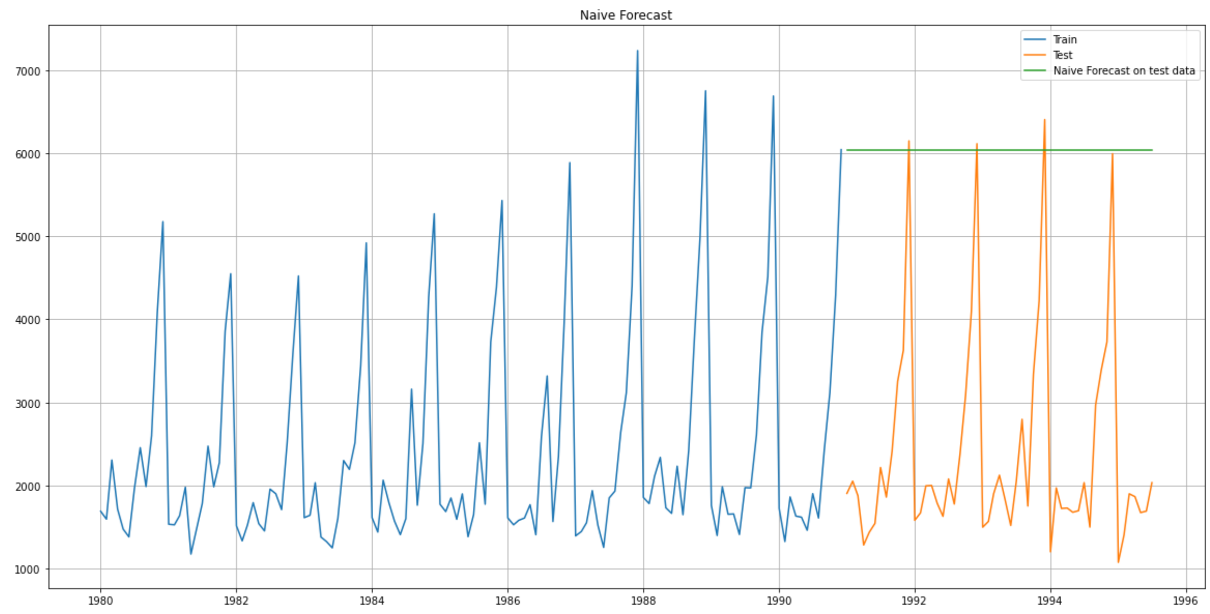


Figure 19. Plot of Forecasted Sales in Naive Model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Naïve model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.

- RMSE in Naive Model is 3864.3

Model 3: Simple Average Approach

- In this model, the forecast for the future is the **simple average of training values**.

YearMonth	Sparkling	forecast_sa
1991-01-01	1902	2403.8
1991-02-01	2049	2403.8
1991-03-01	1874	2403.8
1991-04-01	1279	2403.8
1991-05-01	1432	2403.8
1991-06-01	1540	2403.8
1991-07-01	2214	2403.8
1991-08-01	1857	2403.8
1991-09-01	2408	2403.8
1991-10-01	3252	2403.8

Table 15. Sample of Forecasted Sales in Simple Average Model.

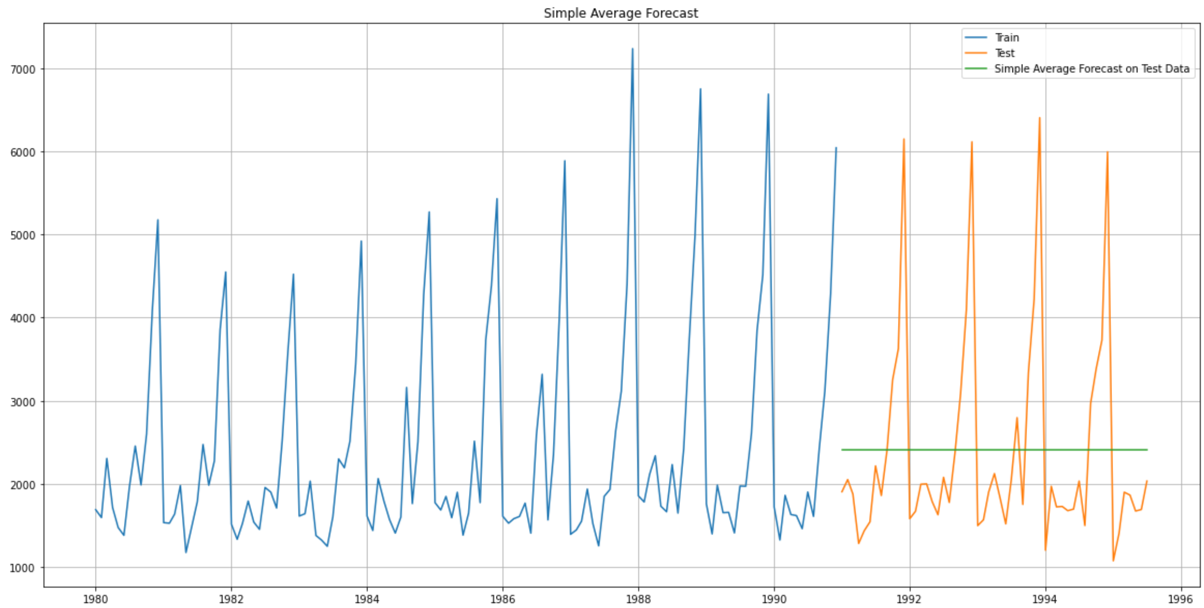


Figure 20. Plot of Forecasted Sales in Simple Average Model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Simple Average model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in Simple Average Model is 1275.1

Exponential Smoothing

- Exponential smoothing methods consist of flattening time series data.
- Exponential smoothing averages or exponentially weighted moving averages consist of forecasts 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.
- These parameters have values between 0 and 1

Model 4: Simple Exponential Smoothing

- The simplest of the exponentially smoothing methods is called simple exponential smoothing (SES).
- This method is suitable for forecasting data when a **time series does not have trend and seasonality**.
- In simple exponential smoothing, the forecast at the time (t + 1) is given by,

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$$

- Parameter **α** is called the smoothing constant and its value lies between 0 and 1. Since the model uses only one smoothing constant, it is called Single Exponential Smoothing.

I. Optimized Model

A simple exponential smoothing model is built and fitted with optimum parameters.

Parameters

	name	param	optimized
smoothing_level	alpha	0.070291	True
initial_level	l.0	1764.013706	True

Table 16. Smoothing Parameters in Simple Exponential Smoothing Optimized Model.

YearMonth	Sparkling	forecast_ses_optimized
1991-01-01	1902	2804.7
1991-02-01	2049	2804.7
1991-03-01	1874	2804.7
1991-04-01	1279	2804.7
1991-05-01	1432	2804.7
1991-06-01	1540	2804.7
1991-07-01	2214	2804.7
1991-08-01	1857	2804.7
1991-09-01	2408	2804.7
1991-10-01	3252	2804.7

Table 17. Sample of Forecasted Sales in Simple Exponential Smoothing Optimized Model.

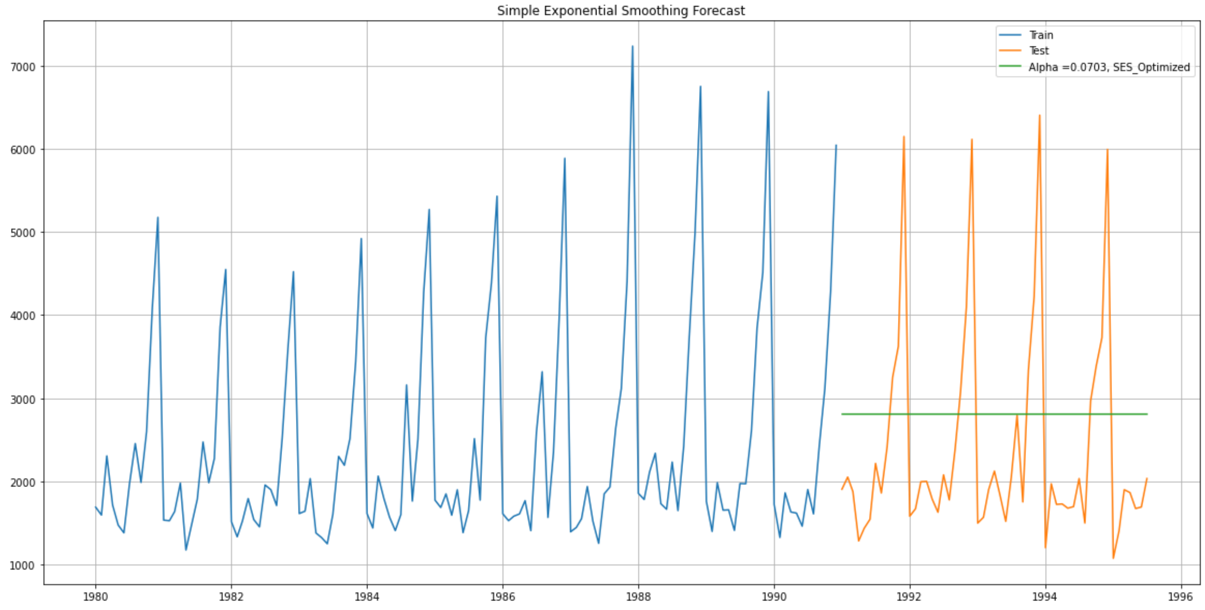


Figure 21. Plot of Forecasted Sales in Simple Exponential Smoothing Optimized Model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Simple Exponential Smoothing Optimized Model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in Simple Exponential Smoothing Optimized Model is 1338.

II. Iteration Model – Finding best **α** to minimize RMSE on the test dataset.

- The higher the alpha value more weightage is given to the more recent observation. That means, what happened recently will happen again. We will run a loop with different alpha values to understand which particular value works best for alpha on the test set.
- Different SES models are built and fitted with different α values (0.1 to 1) and RMSE values are calculated on the test dataset. Below are the top five models with low RMSE values on the test dataset.

	Alpha_Values	RMSE_Train	RMSE_Test
0	0.1	1336.4	1375.4
1	0.2	1357.0	1595.2
2	0.3	1360.0	1935.5
3	0.4	1352.9	2311.9
4	0.5	1344.2	2666.4

Table 18. SES models with low test RMSE values.

From the above table, it can be noticed that the SES model with smoothing constant 0.1 can be considered as the best and final simple exponential smoothing model.

	Sparkling	forecast_ses_optimized	forecast_ses
YearMonth			
1991-01-01	1902	2804.7	2914.8
1991-02-01	2049	2804.7	2914.8
1991-03-01	1874	2804.7	2914.8
1991-04-01	1279	2804.7	2914.8
1991-05-01	1432	2804.7	2914.8
1991-06-01	1540	2804.7	2914.8
1991-07-01	2214	2804.7	2914.8
1991-08-01	1857	2804.7	2914.8
1991-09-01	2408	2804.7	2914.8
1991-10-01	3252	2804.7	2914.8

Table 19. Sample of Forecasted Sales in Simple Exponential Smoothing Iteration Model.

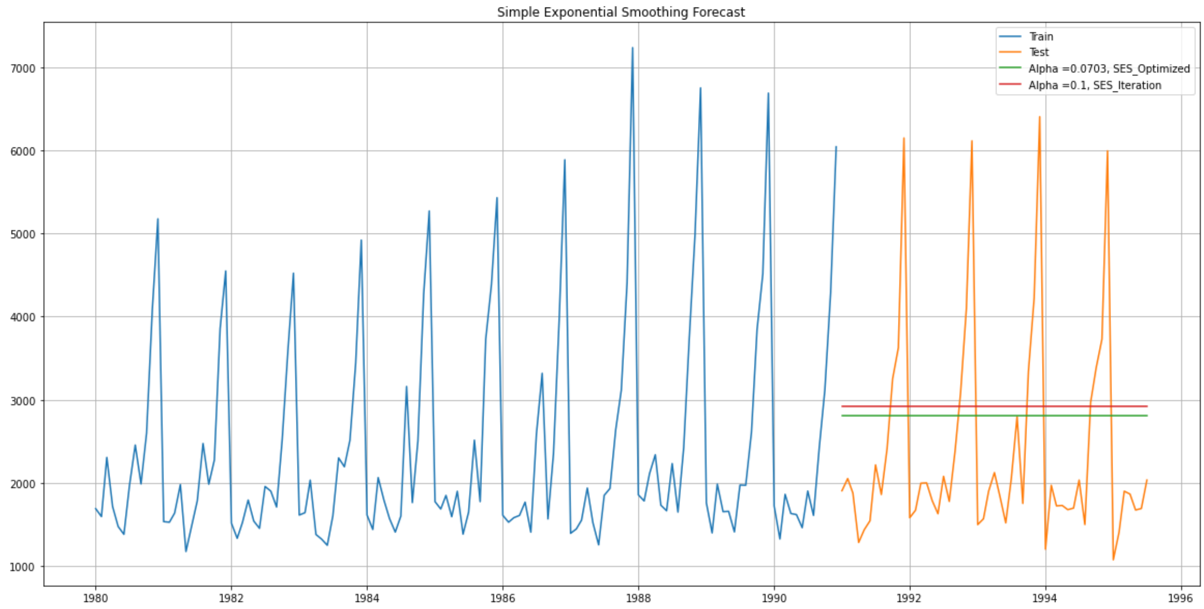


Figure 22. Plot of Forecasted Sales in Simple Exponential Smoothing Iteration Model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Simple Exponential Smoothing Iteration Model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in Simple Exponential Smoothing Iteration Model is 1375.4

Model 5: Double Exponential Smoothing

- This model is applicable when the dataset has a trend but no seasonality.
- Two separate components are considered: Level and Trend.
- Level is the local mean.
- One smoothing parameter α corresponds to the level series which has values between 0 and 1.
- A second smoothing parameter β corresponds to the trend series which has values between 0 and 1.

I. Optimized Model

A double exponential smoothing model is built and fitted with optimum parameters.

	name	param	optimized
smoothing_level	alpha	0.665000	True
smoothing_trend	beta	0.000100	True

Table 20. Smoothing Parameters in Double Exponential Smoothing Optimized Model.

YearMonth	Sparkling	forecast_des_optimized
1991-01-01	1902	5401.7
1991-02-01	2049	5476.0
1991-03-01	1874	5550.3
1991-04-01	1279	5624.5
1991-05-01	1432	5698.8
1991-06-01	1540	5773.1
1991-07-01	2214	5847.4
1991-08-01	1857	5921.6
1991-09-01	2408	5995.9
1991-10-01	3252	6070.2

Table 21. Sample of Forecasted Sales in Double Exponential Smoothing Optimized Model.

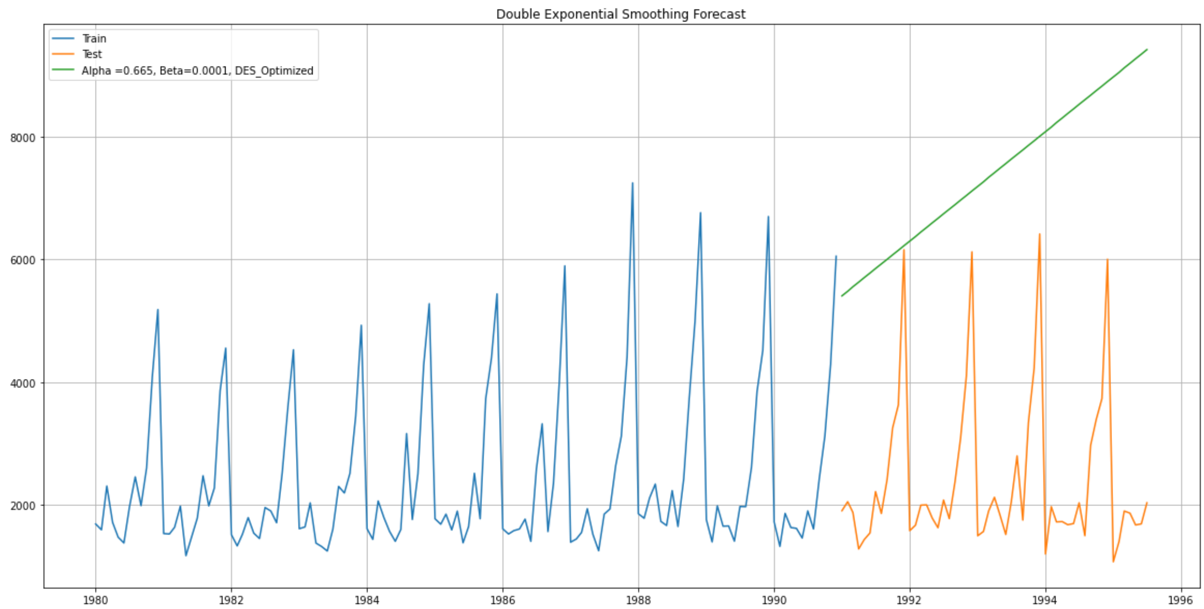


Figure 23. Plot of Forecasted Sales in Double Exponential Smoothing Optimized Model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Double Exponential Smoothing Optimized Model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in Double Exponential Smoothing Optimized Model is 5291.9

II. Iteration Model – Finding best **α** and **β** to minimize RMSE on the test dataset.

- We will run a loop with different **α and β** values to understand which particular values work best for **α and β** on the test set.
- Different DES models are built and fitted with different **α and β** values (0.1 to 1) and RMSE values are calculated on the test dataset. Below are the top five models with low RMSE values on the test dataset.

	Alpha_Values	Beta Values	RMSE_Train	RMSE_Test
0	0.1	0.1	1373.3	1777.7
1	0.1	0.2	1403.5	2599.3
10	0.2	0.1	1413.4	3611.8
2	0.1	0.3	1434.8	4287.5
20	0.3	0.1	1428.6	5908.2

Table 22. DES models with low test RMSE values.

From the above table, it can be noticed that the DES model with smoothing constants α=0.1 & β=0.1 can be considered as the best and final double exponential smoothing model.

	Sparkling	forecast_des_optimized	forecast_des
YearMonth			
1991-01-01	1902	5401.7	2847.7
1991-02-01	2049	5476.0	2874.6
1991-03-01	1874	5550.3	2901.6
1991-04-01	1279	5624.5	2928.6
1991-05-01	1432	5698.8	2955.5
1991-06-01	1540	5773.1	2982.5
1991-07-01	2214	5847.4	3009.5
1991-08-01	1857	5921.6	3036.4
1991-09-01	2408	5995.9	3063.4
1991-10-01	3252	6070.2	3090.4

Table 23. Sample of Forecasted Sales in Double Exponential Smoothing Iteration Model.

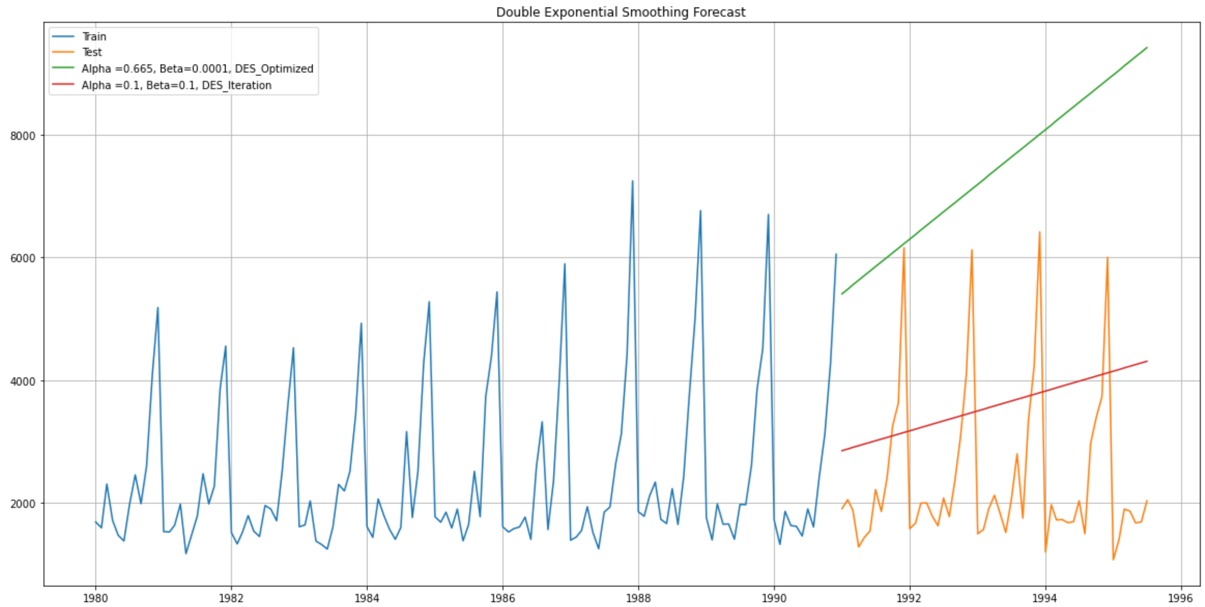


Figure 24. Plot of Forecasted Sales in Double Exponential Smoothing Iteration Model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Double Exponential Smoothing Iteration Model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in Double Exponential Smoothing Iteration Model is 1777.7

Model 6: Triple Exponential Smoothing with Additive trend & Additive seasonality

I. Optimized Model

A TES with the additive trend and additive seasonality model is built and fitted with optimum parameters.

Parameters

	name	param	optimized
smoothing_level	alpha	0.111272	True
smoothing_trend	beta	0.012361	True
smoothing_seasonal	gamma	0.460718	True

Table 24. Smoothing Parameters in TES with the additive trend and additive seasonality optimized model.

	Sparkling	forecast_tes_add_add_optimized
YearMonth		
1991-01-01	1902	1490.4
1991-02-01	2049	1204.5
1991-03-01	1874	1688.7
1991-04-01	1279	1551.2
1991-05-01	1432	1461.2
1991-06-01	1540	1278.6
1991-07-01	2214	1804.9
1991-08-01	1857	1679.0
1991-09-01	2408	2315.4
1991-10-01	3252	3225.0

Table 25. Sample of Forecasted Sales in TES with the additive trend and additive seasonality optimized model.

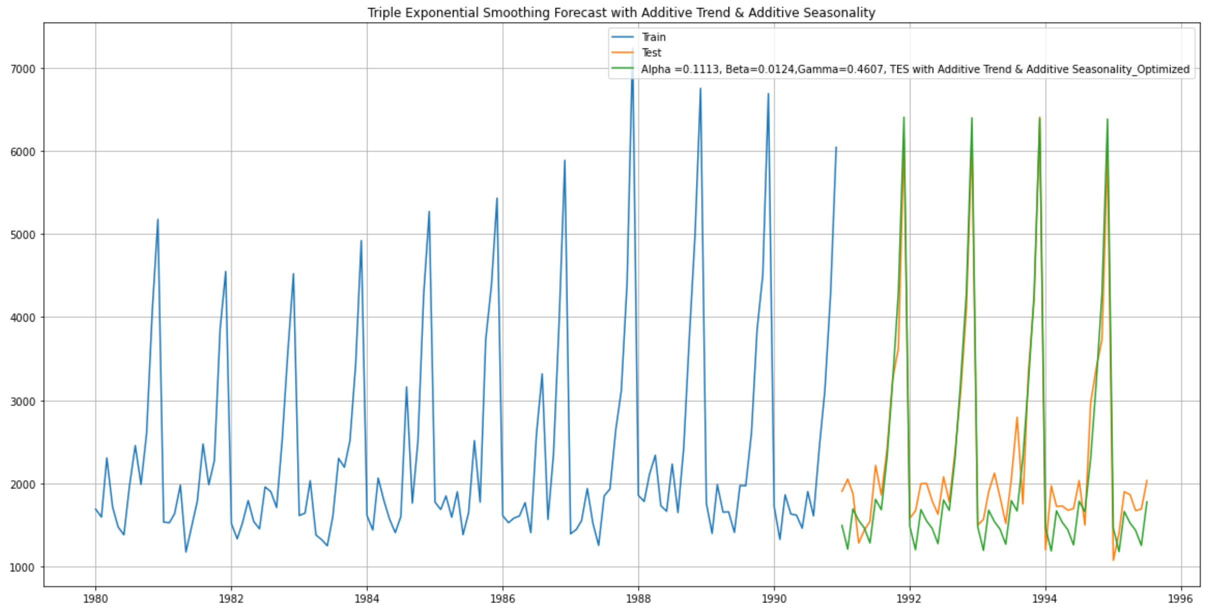


Figure 25. Plot of Forecasted Sales in TES with the additive trend and additive seasonality optimized model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the TES with the additive trend and additive seasonality optimized model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.

- RMSE in TES with the additive trend and additive seasonality optimized model is 378.9

II. Iteration Model – Finding best **α**, **β,** and **γ** to minimize RMSE on the test dataset.

- We will run a loop with different **α, β**, and **γ** values to understand which particular values work best for **α, β**, and **γ** on the test set.
- Different TES models are built and fitted with different **α, β**, and **γ** values (0.1 to 1), and RMSE values are calculated on the test dataset. Below are the top five models with low RMSE values on the test dataset.

	Alpha_Values	Beta Values	Gamma Values	RMSE_Train	RMSE_Test
30	0.1	0.4	0.1	451.6	342.9
110	0.2	0.2	0.1	455.0	343.1
156	0.2	0.6	0.7	492.6	348.8
200	0.3	0.1	0.1	454.3	390.8
20	0.1	0.3	0.1	443.2	391.3

Table 26. TES with the additive trend and additive seasonality models with low test RMSE values.

From the above table, it can be noticed that the TES model with smoothing constants α=0.1, β=0.4, and γ=0.1 can be considered as the best and final triple exponential smoothing model with additive trend and additive seasonality.

YearMonth	Sparkling	forecast_tes_add_add_optimized	forecast_tes_add_add
1991-01-01	1902	1490.4	1613.4
1991-02-01	2049	1204.5	1478.8
1991-03-01	1874	1688.7	1833.2
1991-04-01	1279	1551.2	1760.8
1991-05-01	1432	1461.2	1555.7
1991-06-01	1540	1278.6	1438.9
1991-07-01	2214	1804.9	1955.5
1991-08-01	1857	1679.0	2303.7
1991-09-01	2408	2315.4	2219.4
1991-10-01	3252	3225.0	3003.4

Table 27. Sample of Forecasted Sales in TES with the additive trend and additive seasonality Iteration Model.

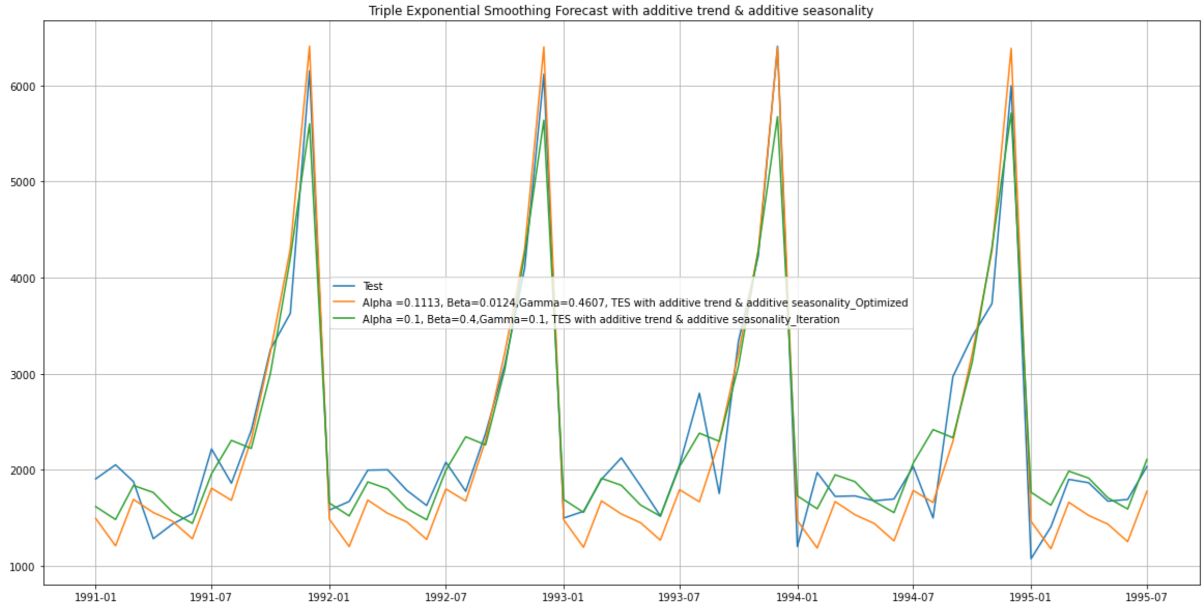


Figure 26. Plot of Forecasted Sales in TES with the additive trend and additive seasonality iteration model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the triple exponential smoothing with the additive trend and additive seasonality iteration model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in triple exponential smoothing with the additive trend and additive seasonality iteration model is 342.9

Model 7: Triple Exponential Smoothing with Additive Trend & Multiplicative Seasonality.

I. Optimized Model

A TES with the additive trend and multiplicative seasonality model is built and fitted with optimum parameters.

	name	param	optimized
smoothing_level	alpha	0.111338	True
smoothing_trend	beta	0.049505	True
smoothing_seasonal	gamma	0.362080	True

Table 28. Smoothing Parameters in TES with the additive trend and multiplicative seasonality optimized model.

YearMonth	Sparkling	forecast_tes_add_mult_optimized
1991-01-01	1902	1587.5
1991-02-01	2049	1356.4
1991-03-01	1874	1762.9
1991-04-01	1279	1656.2
1991-05-01	1432	1542.0
1991-06-01	1540	1355.1
1991-07-01	2214	1854.2
1991-08-01	1857	1820.5
1991-09-01	2408	2277.0
1991-10-01	3252	3122.0

Table 29. Sample of Forecasted Sales in TES with the additive trend and multiplicative seasonality optimized model.

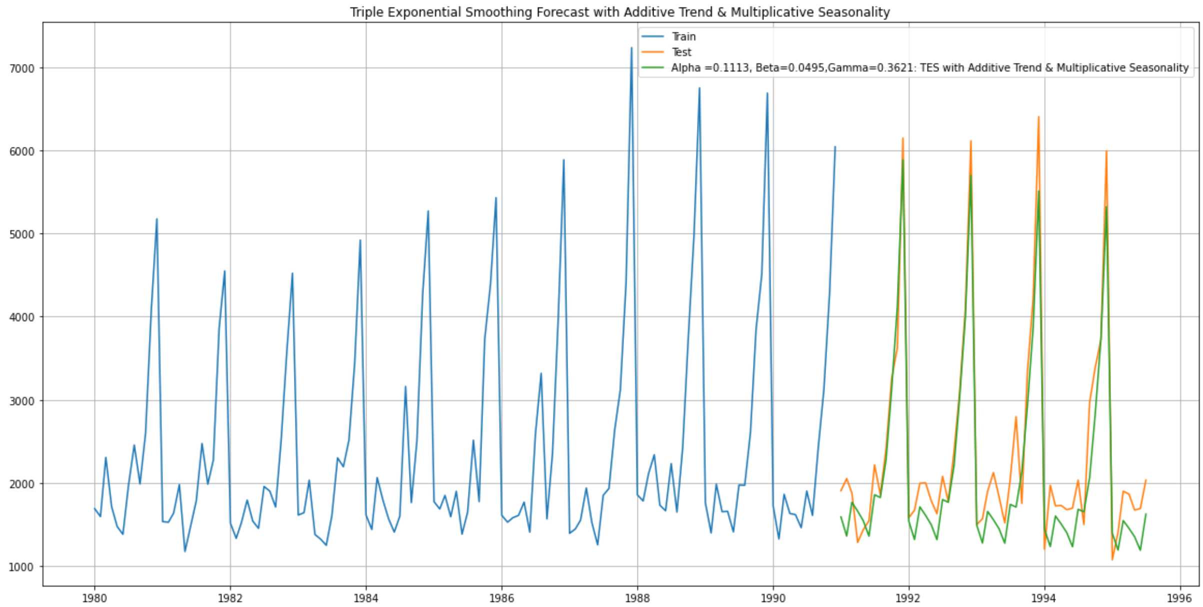


Figure 27. Plot of Forecasted Sales in TES with the additive trend and multiplicative seasonality optimized model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the TES with the additive trend and multiplicative seasonality optimized model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.

- RMSE in TES with the additive trend and multiplicative seasonality optimized model is 404.3

II. Iteration Model – Finding best **α**, **β,** and **γ** to minimize RMSE on the test dataset.

- We will run a loop with different **α, β**, and **γ** values to understand which particular values work best for **α, β**, and **γ** on the test set.
- Different TES models are built and fitted with different **α, β**, and **γ** values (0.1 to 1), and RMSE values are calculated on the test dataset. Below are the top five models with low RMSE values on the test dataset.

	Alpha_Values	Beta Values	Gamma Values	RMSE_Train	RMSE_Test
301	0.4	0.1	0.2	384.5	317.4
211	0.3	0.2	0.2	388.5	329.0
200	0.3	0.1	0.1	388.2	337.1
110	0.2	0.2	0.1	398.5	340.2
402	0.5	0.1	0.3	396.6	345.9

Table 30. TES with the additive trend and multiplicative seasonality models with low test RMSE values.

From the above table, it can be noticed that the TES model with smoothing constants α=0.4, β=0.1, and γ=0.2 can be considered as the best and final triple exponential smoothing model with additive trend and multiplicative seasonality.

	Sparkling	forecast_tes_add_mult_optimized	forecast_tes_add_mult
YearMonth			
1991-01-01	1902	1587.5	1558.1
1991-02-01	2049	1356.4	1425.2
1991-03-01	1874	1762.9	1844.1
1991-04-01	1279	1656.2	1791.8
1991-05-01	1432	1542.0	1658.9
1991-06-01	1540	1355.1	1457.6
1991-07-01	2214	1854.2	2008.6
1991-08-01	1857	1820.5	2194.6
1991-09-01	2408	2277.0	2505.0
1991-10-01	3252	3122.0	3279.3

Table 31. Sample of Forecasted Sales in TES with the additive trend and multiplicative seasonality Iteration Model.

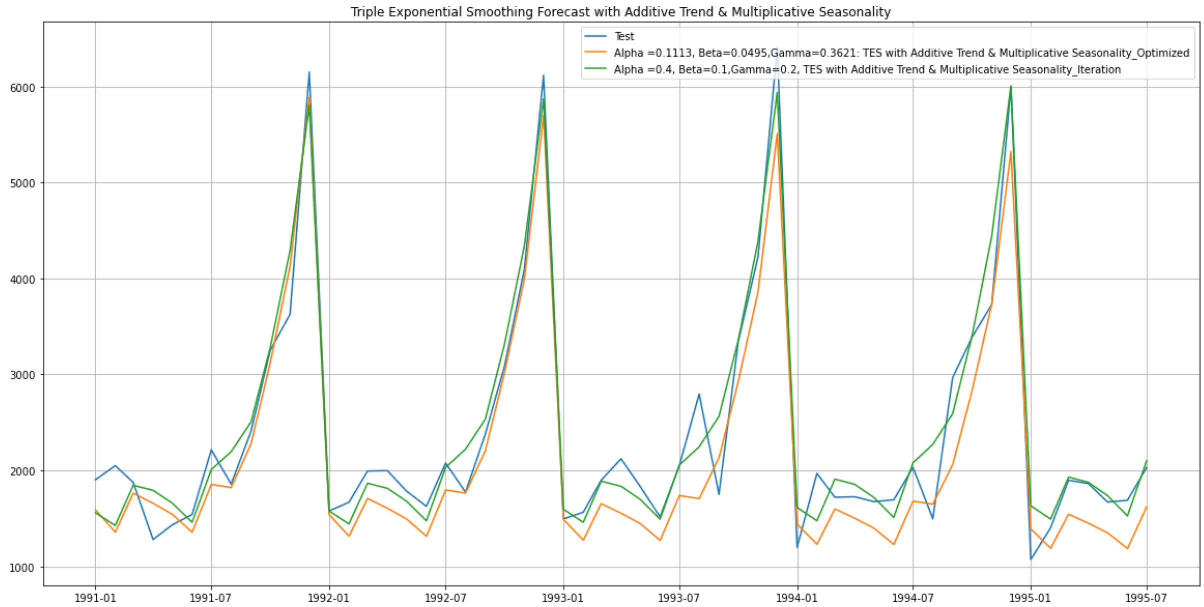


Figure 28. Plot of Forecasted Sales in TES with the additive trend and multiplicative seasonality iteration model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the triple exponential smoothing with the additive trend and multiplicative seasonality iteration model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in triple exponential smoothing with the additive trend and multiplicative seasonality iteration model is 317.4

Model 8: Triple Exponential Smoothing with Multiplicative Trend & Multiplicative Seasonality.

I. Optimized Model

A TES with the multiplicative trend and multiplicative seasonality model is built and fitted with optimum parameters.

	name	param	optimized
smoothing_level	alpha	0.111067	True
smoothing_trend	beta	0.049361	True
smoothing_seasonal	gamma	0.362182	True

Table 32. Smoothing Parameters in TES with the multiplicative trend and multiplicative seasonality optimized model.

YearMonth	Sparkling	forecast_tes_mult_mult_optimized
1991-01-01	1902	1591.3
1991-02-01	2049	1360.4
1991-03-01	1874	1767.9
1991-04-01	1279	1661.6
1991-05-01	1432	1547.4
1991-06-01	1540	1360.5
1991-07-01	2214	1862.1
1991-08-01	1857	1829.5
1991-09-01	2408	2287.1
1991-10-01	3252	3137.0

Table 33. Sample of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality optimized model.

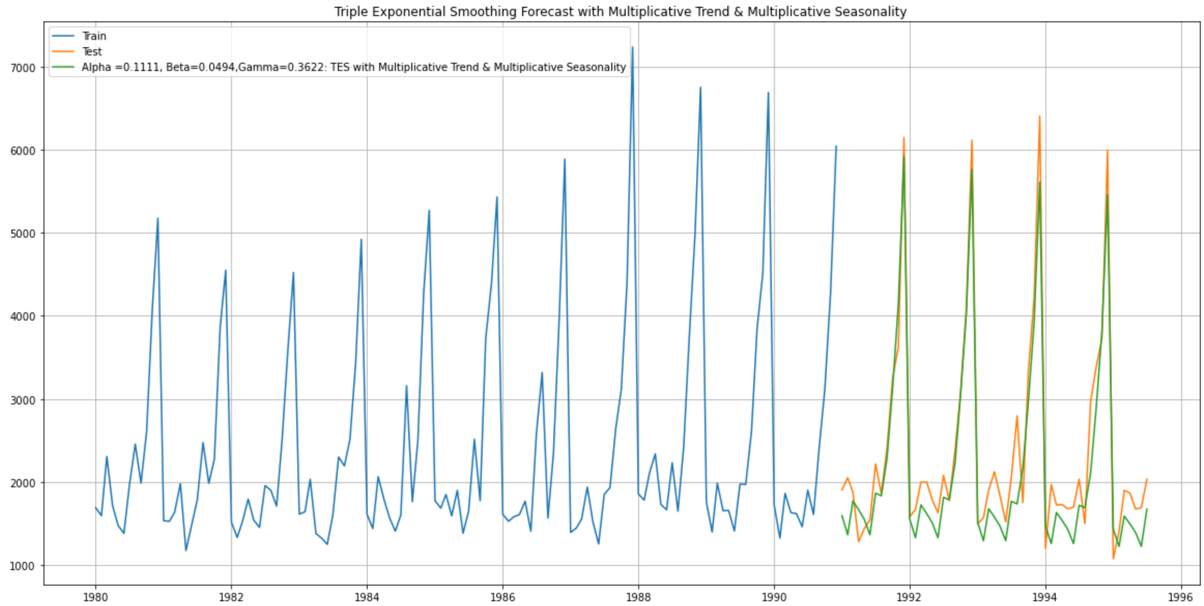


Figure 29. Plot of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality optimized model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the TES with the multiplicative trend and multiplicative seasonality optimized model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.

- RMSE in TES with the multiplicative trend and multiplicative seasonality optimized model is 380.4

II. Iteration Model – Finding best **α**, **β,** and **γ** to minimize RMSE on the test dataset.

- We will run a loop with different **α, β**, and **γ** values to understand which particular values work best for **α, β**, and **γ** on the test set.
- Different TES models are built and fitted with different **α, β**, and **γ** values (0.1 to 1), and RMSE values are calculated on the test dataset. Below are the top five models with low RMSE values on the test dataset.

	Alpha_Values	Beta Values	Gamma Values	RMSE_Train	RMSE_Test
245	0.4	0.1	0.3	381.1	326.6
163	0.3	0.1	0.2	376.0	342.5
90	0.2	0.2	0.1	396.0	345.9
109	0.2	0.4	0.2	401.7	349.4
182	0.3	0.3	0.3	396.7	353.6

Table 34. TES with the multiplicative trend and multiplicative seasonality models with low test RMSE values.

From the above table, it can be noticed that the TES model with smoothing constants α=0.4, β=0.1, and γ=0.3 can be considered as the best and final triple exponential smoothing model with multiplicative trend and multiplicative seasonality.

	Sparkling	forecast_tes_mult_mult_optimized	forecast_tes_mult_mult
YearMonth			
1991-01-01	1902	1591.3	1502.6
1991-02-01	2049	1360.4	1324.1
1991-03-01	1874	1767.9	1761.9
1991-04-01	1279	1661.6	1689.6
1991-05-01	1432	1547.4	1605.2
1991-06-01	1540	1360.5	1410.8
1991-07-01	2214	1862.1	1940.2
1991-08-01	1857	1829.5	1981.5
1991-09-01	2408	2287.1	2533.8
1991-10-01	3252	3137.0	3322.3

Table 35. Sample of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality Iteration Model.

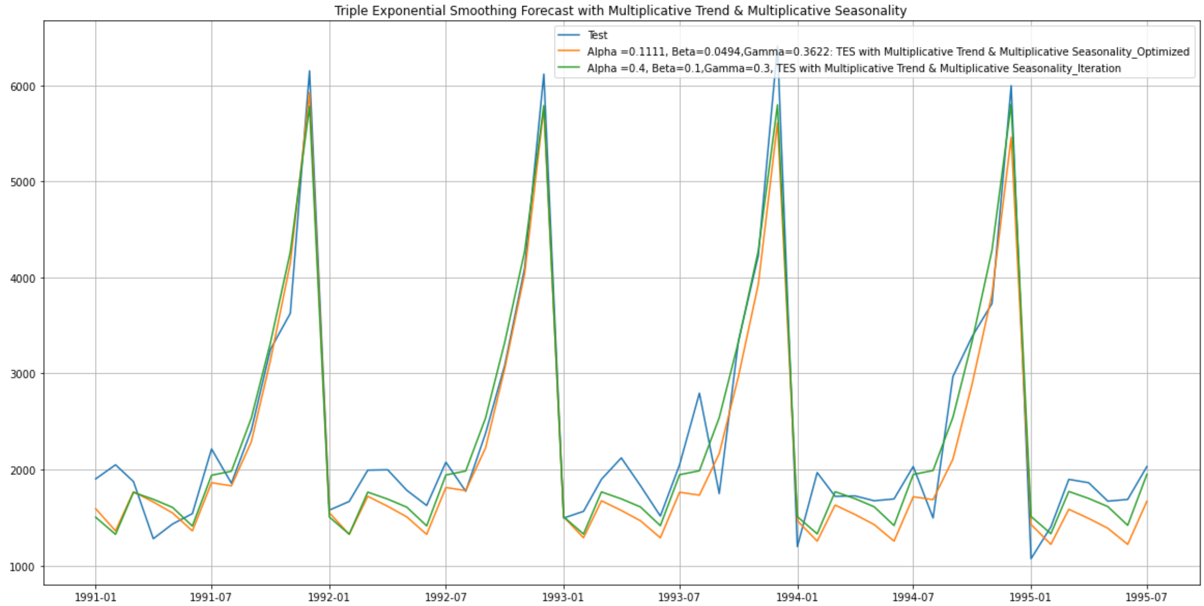


Figure 30. Plot of Forecasted Sales in TES with the multiplicative trend and multiplicative seasonality iteration model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the triple exponential smoothing with the multiplicative trend and multiplicative seasonality iteration model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in triple exponential smoothing with the multiplicative trend and multiplicative seasonality iteration model is 326.6

Model 9: Triple Exponential Smoothing with Multiplicative Trend & Additive Seasonality.

I. Optimized Model

A TES with the multiplicative trend and additive seasonality model is built and fitted with optimum parameters.

	name	param	optimized
smoothing_level	alpha	0.115477	True
smoothing_trend	beta	0.013339	True
smoothing_seasonal	gamma	0.456513	True

Table 36. Smoothing Parameters in TES with the multiplicative trend and additive seasonality optimized model.

YearMonth	Sparkling	forecast_tes_mult_add_optimized
1991-01-01	1902	1483.3
1991-02-01	2049	1199.0
1991-03-01	1874	1682.8
1991-04-01	1279	1546.7
1991-05-01	1432	1456.4
1991-06-01	1540	1273.8
1991-07-01	2214	1800.9
1991-08-01	1857	1678.5
1991-09-01	2408	2312.0
1991-10-01	3252	3221.6

Table 37. Sample of Forecasted Sales in TES with the multiplicative trend and additive seasonality optimized model.

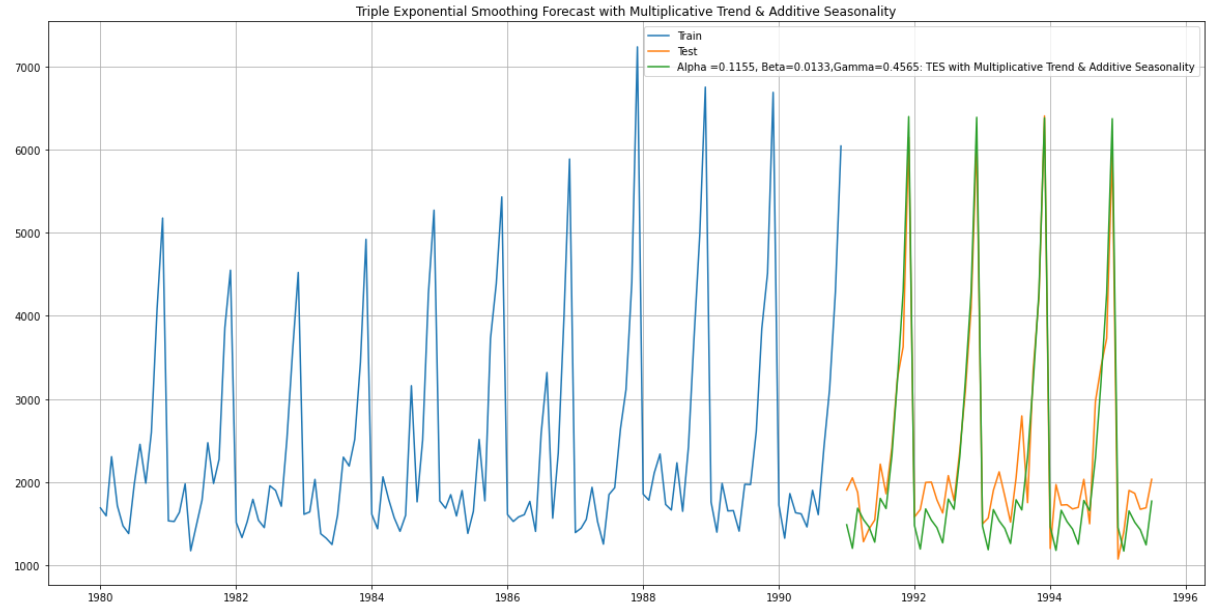


Figure 31. Plot of Forecasted Sales in TES with the multiplicative trend and additive seasonality optimized model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the TES with the multiplicative trend and additive seasonality optimized model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.

- RMSE in TES with the multiplicative trend and additive seasonality optimized model is 381.2

II. Iteration Model – Finding best **α**, **β,** and **γ** to minimize RMSE on the test dataset.

- We will run a loop with different **α, β**, and **γ** values to understand which particular values work best for **α, β**, and **γ** on the test set.
- Different TES models are built and fitted with different **α, β**, and **γ** values (0.1 to 1), and RMSE values are calculated on the test dataset. Below are the top five models with low RMSE values on the test dataset.

	Alpha_Values	Beta Values	Gamma Values	RMSE_Train	RMSE_Test
244	0.4	0.1	0.2	440.0	341.7
27	0.1	0.4	0.1	456.9	341.8
81	0.2	0.1	0.1	443.2	356.6
90	0.2	0.2	0.1	458.0	359.6
18	0.1	0.3	0.1	446.6	375.2

Table 38. TES with the multiplicative trend and additive seasonality models with low test RMSE values.

From the above table, it can be noticed that the TES model with smoothing constants α=0.4, β=0.1, and γ=0.2 can be considered as the best and final triple exponential smoothing model with multiplicative trend and additive seasonality.

YearMonth	Sparkling	forecast_tes_mult_add_optimized	forecast_tes_mult_add
1991-01-01	1902	1483.3	1265.1
1991-02-01	2049	1199.0	1187.8
1991-03-01	1874	1682.8	1695.7
1991-04-01	1279	1546.7	1693.7
1991-05-01	1432	1456.4	1608.5
1991-06-01	1540	1273.8	1461.4
1991-07-01	2214	1800.9	2034.5
1991-08-01	1857	1678.5	2227.7
1991-09-01	2408	2312.0	2491.3
1991-10-01	3252	3221.6	3281.0

Table 39. Sample of Forecasted Sales in TES with the multiplicative trend and additive seasonality Iteration Model.

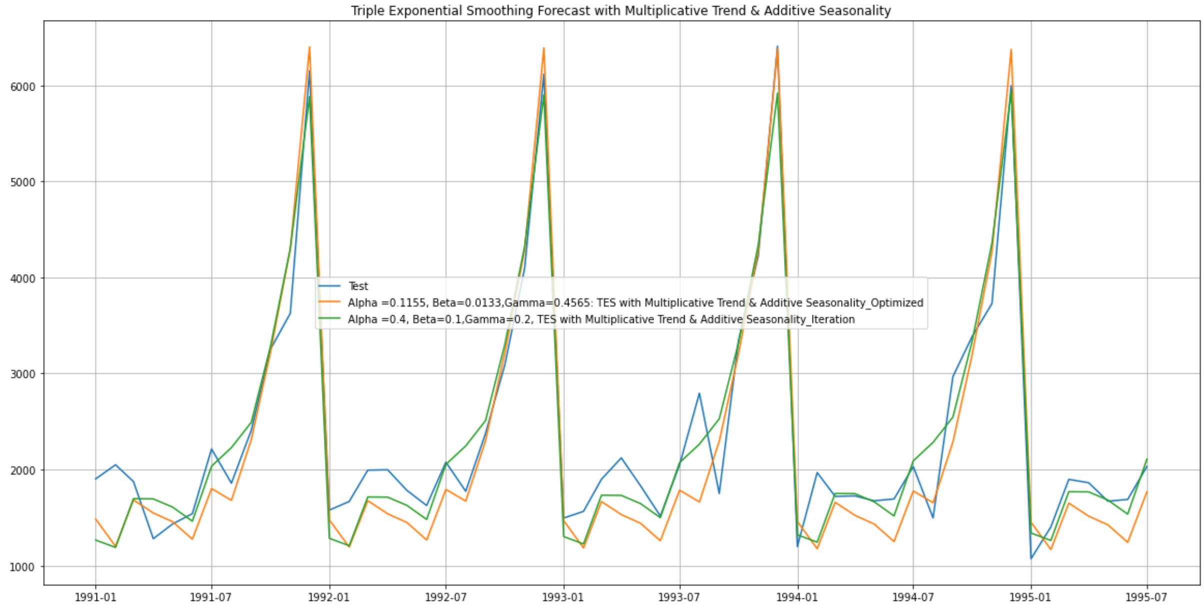


Figure 32. Plot of Forecasted Sales in TES with the multiplicative trend and additive seasonality iteration model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is approximately matching with the actual plot of the test set**. Hence, the triple exponential smoothing with the multiplicative trend and additive seasonality iteration model may be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in triple exponential smoothing with the multiplicative trend and additive seasonality iteration model is 341.6

Q5. Check for the stationarity of the data on which the model is being built 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.

Check for Stationarity of Whole Time Series

The Augmented Dickey-Fuller test is a unit root test that determines whether there is a unit root and subsequently whether the series is non-stationary.

The hypothesis for the ADF test is

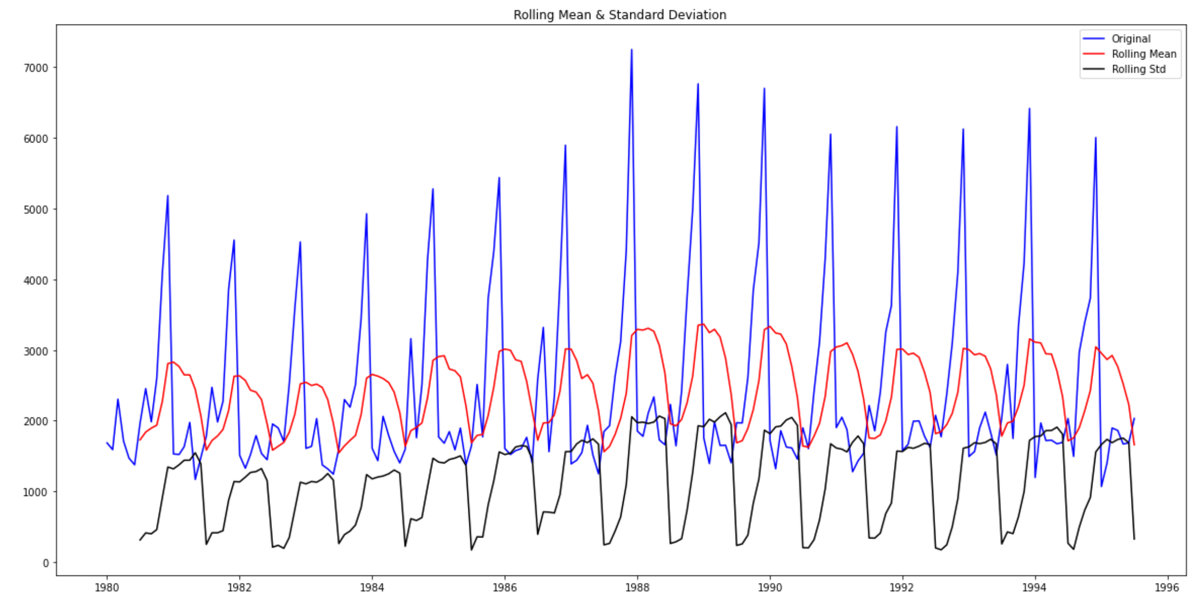
H0: The Time Series has a unit root and is thus non-stationary

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

Significance level α = 0.05

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

Results of Dickey-Fuller Test on Whole Time Series

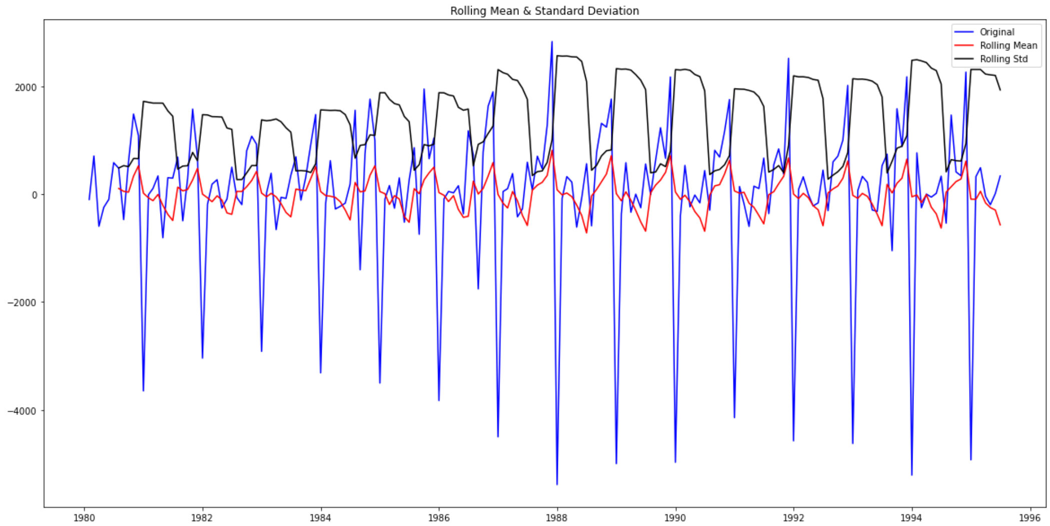



```
Results of Dickey-Fuller Test:
Test Statistic                  -1.360497
p-value                          0.601061
#Lags Used                      11.000000
Number of Observations Used    175.000000
Critical Value (1%)             -3.468280
Critical Value (5%)             -2.878202
Critical Value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

As the P-value in Dicky-Fuller test is 0.6 which is more than the significance level (0.05), we failed to reject the null hypothesis. Hence, the given **time series is non-stationary**.

To make time-series stationary, let us take first differencing on whole time series and perform Dicky-Fuller test one again.

Results of Dickey-Fuller Test on Differenced Whole Time Series




```
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
```

As the P-value in Dicky-Fuller test is 0 which is less than the significance level (0.05), we can reject the null hypothesis. Hence, **one level differenced time series is stationary**.

Plotting Differenced Time Series

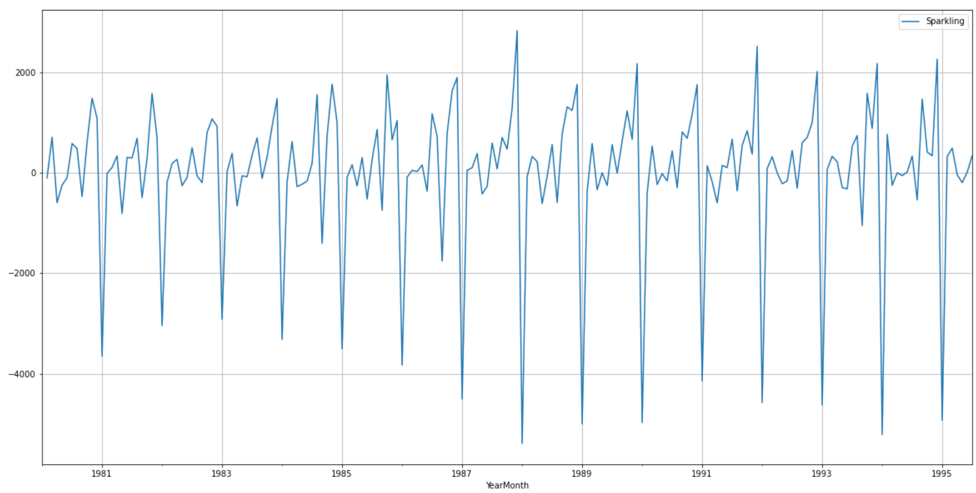


Figure 33. Plot of Differenced Time Series

Plot the Autocorrelation and the Partial Autocorrelation function plots on the whole data

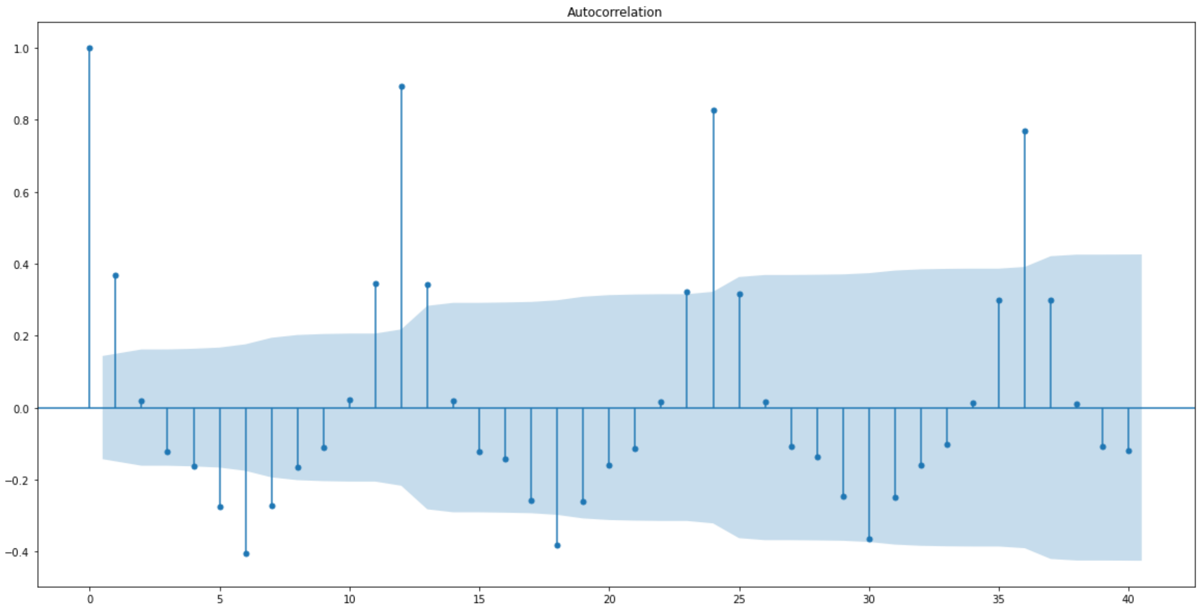


Figure 34. Autocorrelation Plot of Whole Time Series

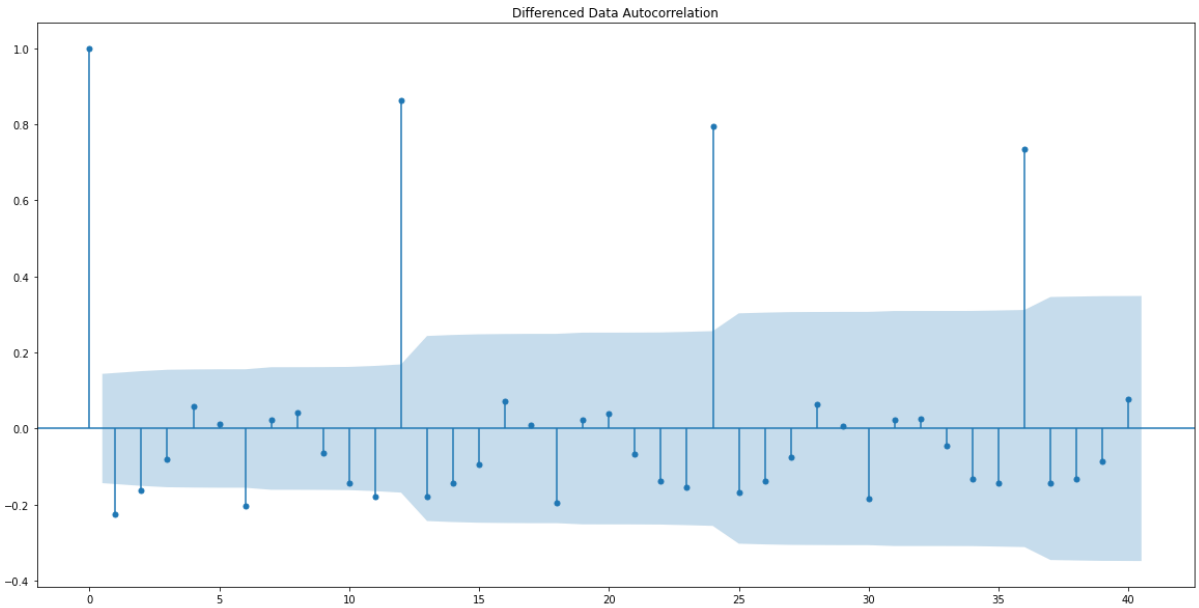


Figure 35. Autocorrelation Plot of Differenced Whole Time Series

From the above plot, it can be noticed that there is a seasonality in the time series.

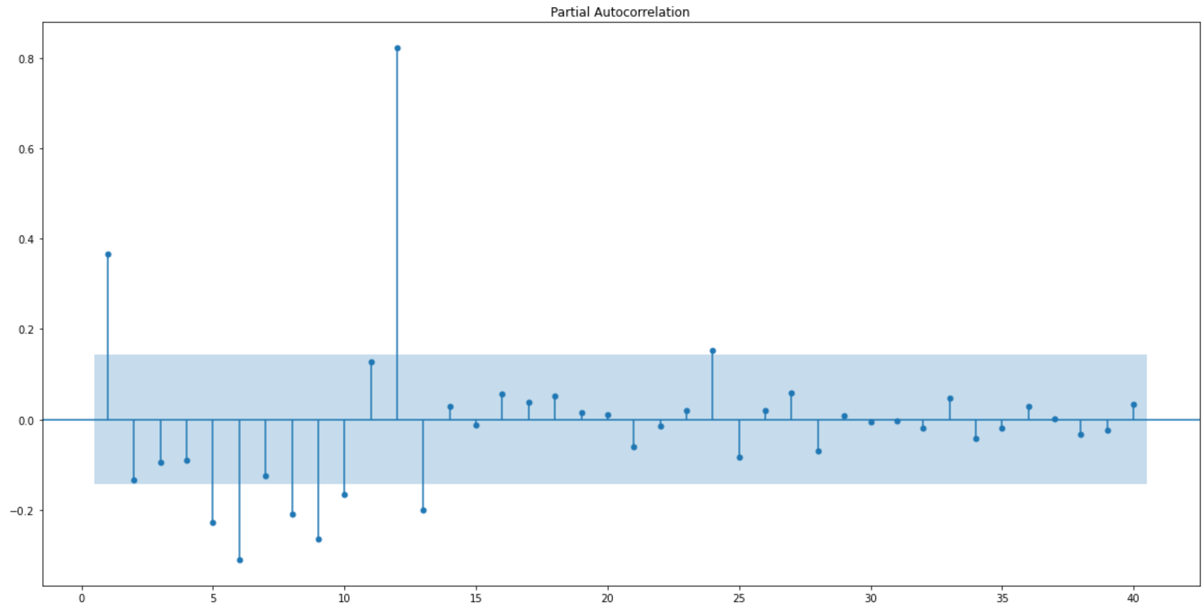


Figure 36. Partial Autocorrelation Plot of Whole Time Series

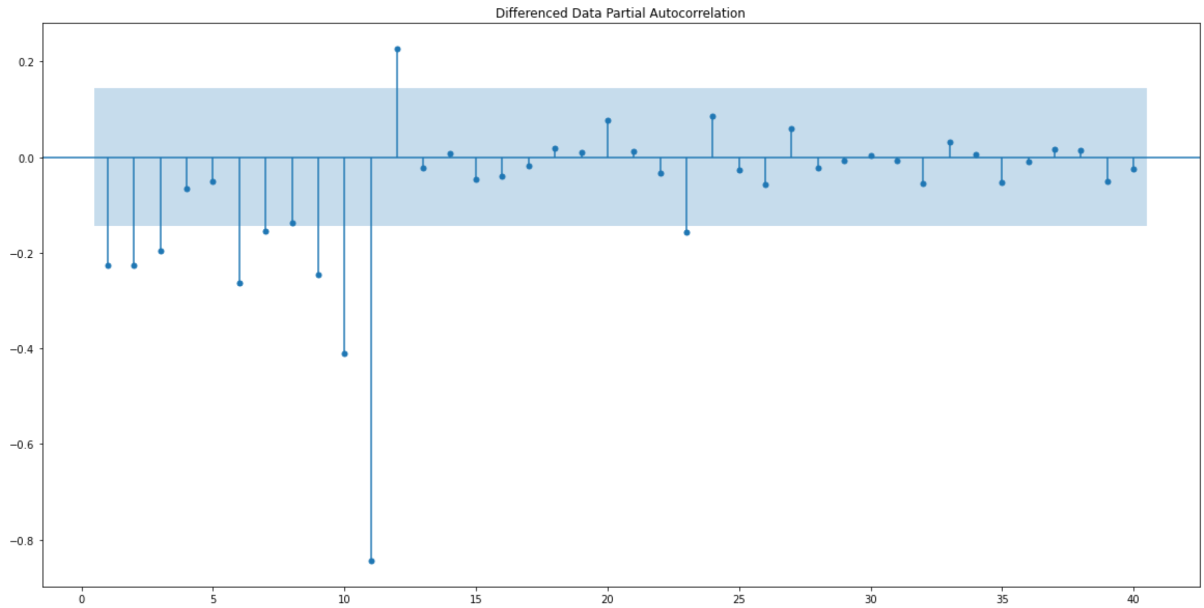


Figure 37. Partial Autocorrelation Plot of Differenced Whole Time Series

Check for Stationarity of Train Dataset

Results of Dickey-Fuller Test on Train Dataset

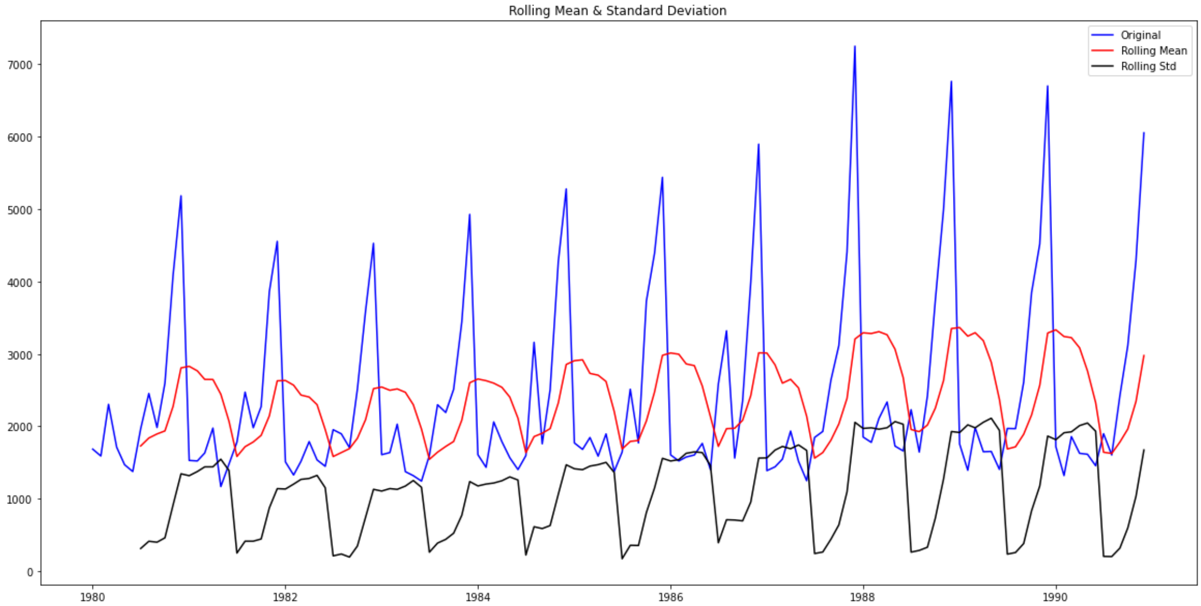
The Augmented Dickey-Fuller test is a unit root test that determines whether there is a unit root and subsequently whether the series is non-stationary.

The hypothesis for the ADF test is

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

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

Significance level, α = 0.05

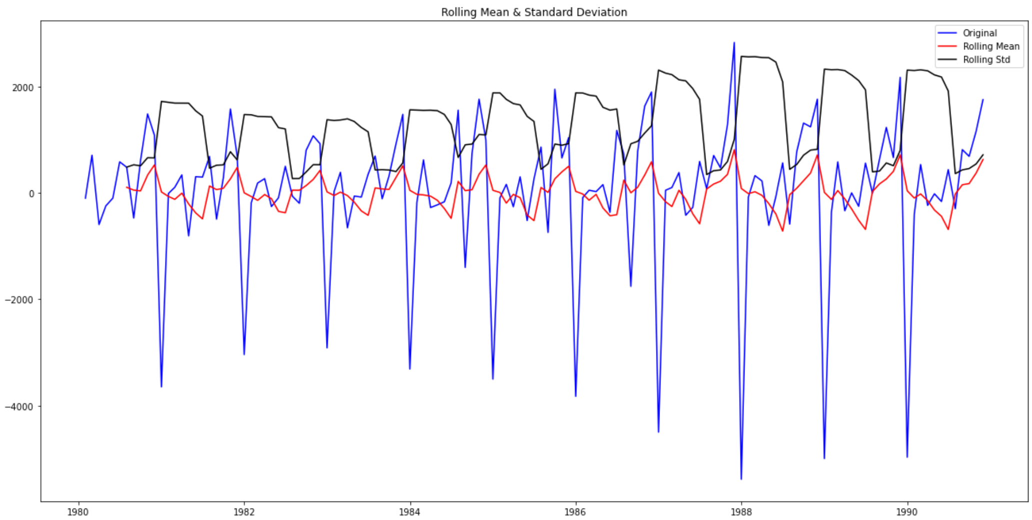



```
Results of Dickey-Fuller Test:
Test Statistic                  -1.208926
p-value                          0.669744
#Lags Used                      12.000000
Number of Observations Used    119.000000
Critical Value (1%)             -3.486535
Critical Value (5%)             -2.886151
Critical Value (10%)            -2.579896
dtype: float64
```

As the P-value in Dicky-Fuller test is 0.67 which is more than the significance level (0.05), we failed to reject the null hypothesis. Hence, the training dataset **is non-stationary**.

To make time-series stationary, let us take first differencing on the training dataset and perform Dicky-Fuller test once again.

Results of Dickey-Fuller Test on Differenced Training Dataset



```
Results of Dickey-Fuller Test:
Test Statistic                 -8.005007e+00
p-value                         2.280104e-12
#Lags Used                      1.100000e+01
Number of Observations Used     1.190000e+02
Critical Value (1%)            -3.486535e+00
Critical Value (5%)            -2.886151e+00
Critical Value (10%)           -2.579896e+00
dtype: float64
```

As the P-value in Dicky-Fuller test is 0 which is less than the significance level (0.05), we can reject the null hypothesis. Hence, **one level differenced training dataset is stationary**.

Plotting Differenced Time Series

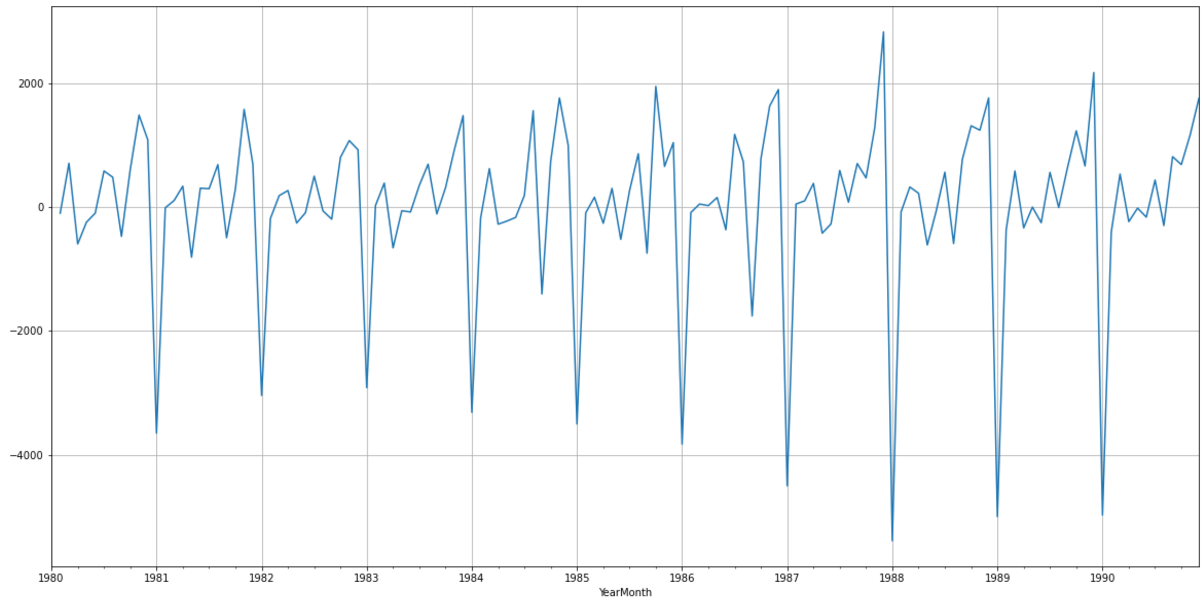


Figure 38. Plot of Differenced Training Dataset

Q6. 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.

Automated Version of ARIMA Model

The following values are considered for the parameters p, d, and q to build to various ARIMA models.

Autoregressive Component, p = 0, 1, 2, 3

Moving average component, q = 0, 1, 2, 3

Differencing Component, d = 1

order = [(0, 1, 0), (0, 1, 1), (0, 1, 2), (0, 1, 3), (1, 1, 0), (1, 1, 1), (1, 1, 2), (1, 1, 3), (2, 1, 0), (2, 1, 1), (2, 1, 2), (2, 1, 3), (3, 1, 0), (3, 1, 1), (3, 1, 2), (3, 1, 3)]

Various ARIMA models are built and fitted by using above order values and for every model AIC value is calculated. The ARIMA model with lowest AIC value is considered as the best model. Again, the best ARIMA model is built by using corresponding order parameters for forecasting purpose.

	param	AIC
10	(2, 1, 2)	2213.5
15	(3, 1, 3)	2221.5
14	(3, 1, 2)	2231.0
11	(2, 1, 3)	2232.9
9	(2, 1, 1)	2233.8

Table 40. Top five ARIMA models with low AIC values.

ARIMA model with order (2, 1, 2)

Summary

Dep. Variable:	Sparkling	**No. Observations:**	132
Model:	ARIMA(2, 1, 2)	**Log Likelihood**	-1101.755
Date:	Sun, 16 Jan 2022	**AIC**	2213.509
Time:	15:49:32	**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	28.781	0.000	1.223	1.401
ar.L2	-0.5593	0.072	-7.741	0.000	-0.701	-0.418
ma.L1	-1.9917	0.109	-18.218	0.000	-2.206	-1.777
ma.L2	0.9999	0.110	9.109	0.000	0.785	1.215
sigma2	1.099e+06	1.99e-07	5.51e+12	0.000	1.1e+06	1.1e+06

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

Table 41. Summary of an automated ARIMA model.

Diagnostics Plots

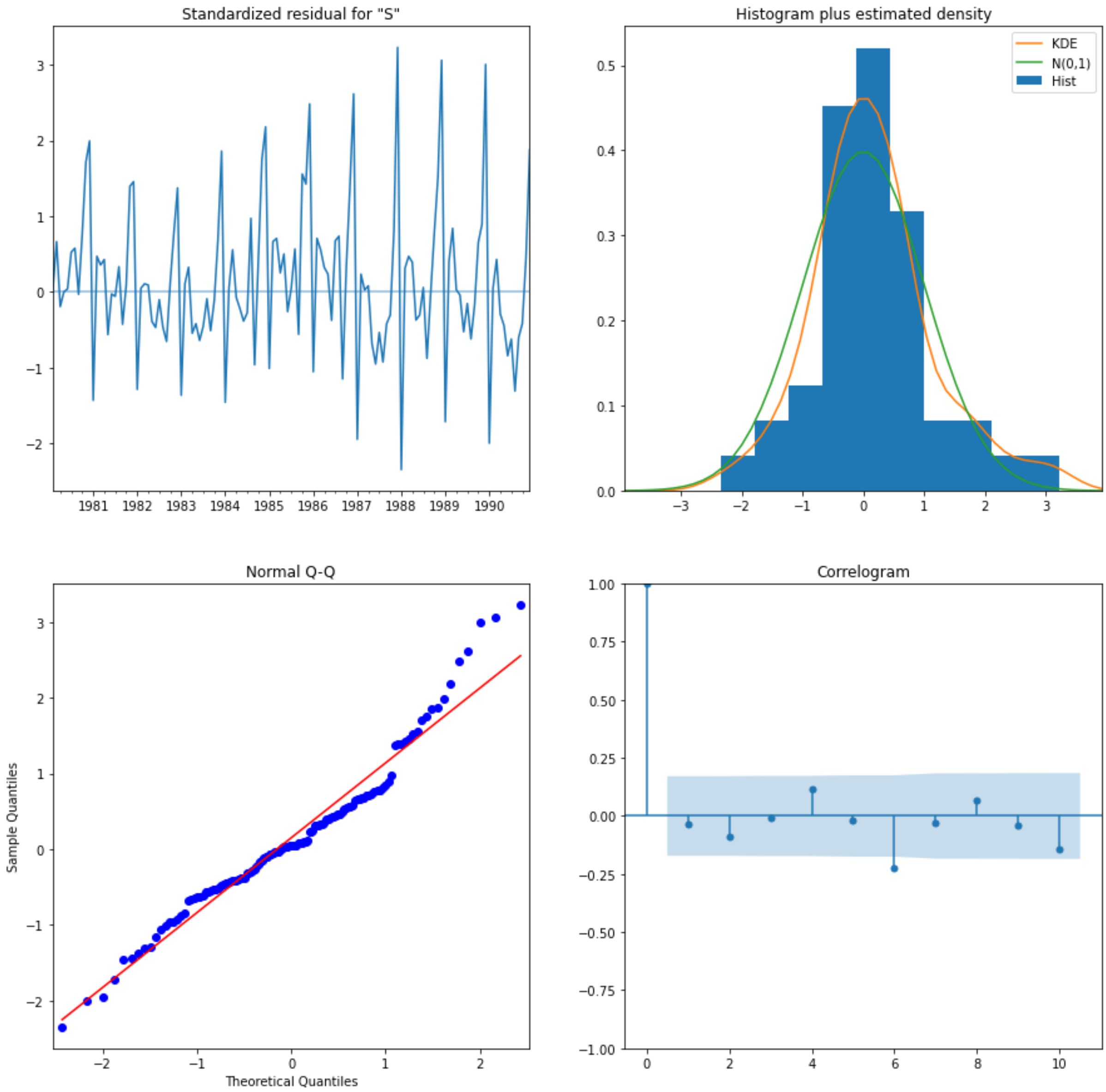


Figure 39. Diagnostics Plots of an automated ARIMA model.

Insights

- Autoregressive model - Lag 1 series has Z value (28.781). It means that the forecast for this month is largely influenced by last month value.
- Moving Average model - Lag 1 series has Z value (-18.218). It means that the forecast for this month is largely influenced by last month’s error term.

YearMonth	Sparkling	forecast_ARIMA_auto
1991-01-01	1902	4252.3
1991-02-01	2049	2863.1
1991-03-01	1874	2044.0
1991-04-01	1279	1746.2
1991-05-01	1432	1813.6
1991-06-01	1540	2068.6
1991-07-01	2214	2365.5
1991-08-01	1857	2612.5
1991-09-01	2408	2770.4
1991-10-01	3252	2839.5

Table 42. Sample of Forecasted Sales in an automated ARIMA Model.

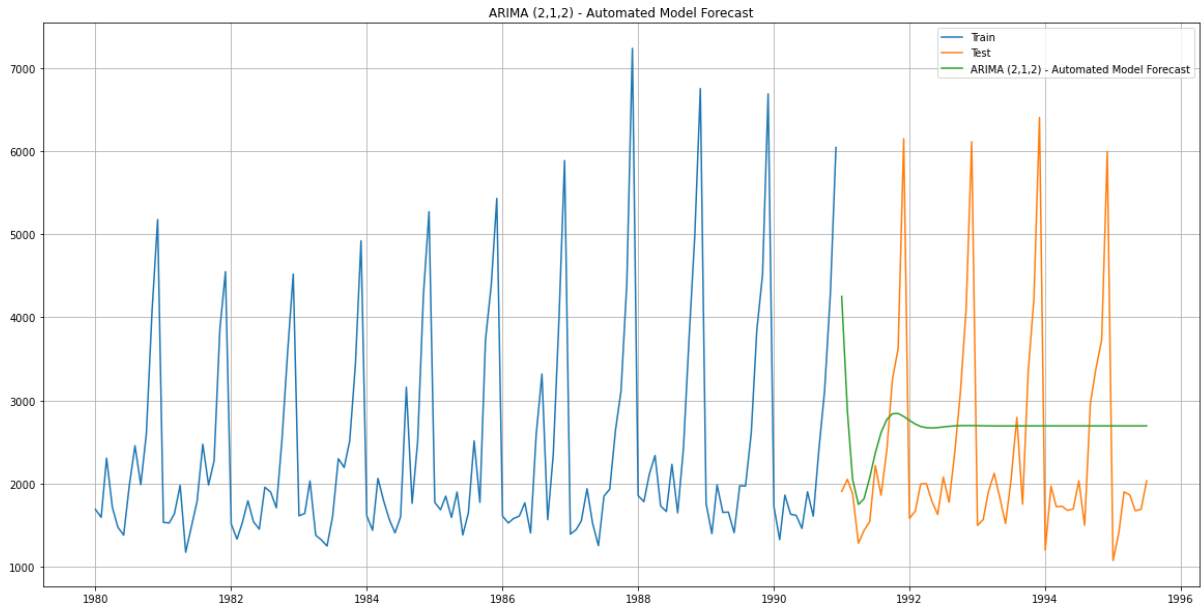


Figure 40. Plot of Forecasted Sales in an automated ARIMA model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the automated ARIMA model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in an automated ARIMA model is 1299.9

Automated Version of SARIMA Model

The following values are considered for the parameters (p, d, q), (P, D, Q, s) to build to various ARIMA models.

Autoregressive Component, p = 0, 1, 2, 3

Moving average component, q = 0, 1, 2, 3

Differencing Component, d = 1

Seasonal Autoregressive Component, P = 0, 1, 2, 3

Seasonal Moving average component, Q = 0, 1, 2, 3

Seasonal Differencing Component, D = 0, 1

Periodicity, s = 6

order = [(0, 1, 0), (0, 1, 1), (0, 1, 2), (0, 1, 3), (1, 1, 0), (1, 1, 1), (1, 1, 2), (1, 1, 3), (2, 1, 0), (2, 1, 1), (2, 1, 2), (2, 1, 3), (3, 1, 0), (3, 1, 1), (3, 1, 2), (3, 1, 3)]

Seasonal order = [(0, 0, 0, 6), (0, 0, 1, 6), (0, 0, 2, 6), (0, 0, 3, 6), (0, 1, 0, 6), (0, 1, 1, 6), (0, 1, 2, 6), (0, 1, 3, 6), (1, 0, 0, 6), (1, 0, 1, 6), (1, 0, 2, 6), (1, 0, 3, 6), (1, 1, 0, 6), (1, 1, 1, 6), (1, 1, 2, 6), (1, 1, 3, 6), (2, 0, 0, 6), (2, 0, 1, 6), (2, 0, 2, 6), (2, 0, 3, 6), (2, 1, 0, 6), (2, 1, 1, 6), (2, 1, 2, 6), (2, 1, 3, 6), (3, 0, 0, 6), (3, 0, 1, 6), (3, 0, 2, 6), (3, 0, 3, 6), (3, 1, 0, 6), (3, 1, 1, 6), (3, 1, 2, 6), (3, 1, 3, 6)]

Various SARIMA models are built and fitted by using the above order values and for every model, the AIC value is calculated. The SARIMA model with the lowest AIC value is considered the best model. Again, the best SARIMA model is built by using corresponding order parameters for forecasting purposes.

	param	param_seasonal	AIC
375	(2, 1, 3)	(2, 1, 3, 6)	1540.9
495	(3, 1, 3)	(1, 1, 3, 6)	1541.0
111	(0, 1, 3)	(1, 1, 3, 6)	1544.0
503	(3, 1, 3)	(2, 1, 3, 6)	1544.7
239	(1, 1, 3)	(1, 1, 3, 6)	1544.9

Table 43. Top five SARIMA models with low AIC values.

SARIMA model with order (2, 1, 3) (2, 1, 3, 6)

Summary

Dep. Variable:	Sparkling	**No. Observations:**	132
Model:	SARIMAX(2, 1, 3)x(2, 1, 3, 6)	**Log Likelihood**	-759.447
Date:	Sun, 16 Jan 2022	**AIC**	1540.893
Time:	17:14:28	**BIC**	1569.875
Sample:	01-01-1980	**HQIC**	1552.632
	- 12-01-1990		
Covariance Type:	opg		

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
ar.L1	-1.7472	0.066	-26.431	0.000	-1.877	-1.618
ar.L2	-0.7890	0.072	-10.911	0.000	-0.931	-0.647
ma.L1	1.0811	0.300	3.601	0.000	0.493	1.669
ma.L2	-0.7653	0.127	-6.018	0.000	-1.015	-0.516
ma.L3	-0.8964	0.272	-3.301	0.001	-1.429	-0.364
ar.S.L6	-1.0887	0.673	-1.617	0.106	-2.408	0.231
ar.S.L12	-0.0656	0.693	-0.095	0.925	-1.423	1.292
ma.S.L6	0.4196	0.650	0.645	0.519	-0.855	1.694
ma.S.L12	-0.6923	0.335	-2.068	0.039	-1.348	-0.036
ma.S.L18	0.0774	0.522	0.148	0.882	-0.946	1.101
sigma2	1.13e+05	4.52e-06	2.5e+10	0.000	1.13e+05	1.13e+05

Ljung-Box (L1) (Q):	0.00	**Jarque-Bera (JB):**	12.52
Prob(Q):	0.98	**Prob(JB):**	0.00
Heteroskedasticity (H):	1.37	**Skew:**	0.41
Prob(H) (two-sided):	0.37	**Kurtosis:**	4.49

Table 44. Summary of an automated SARIMA model.

Insights

- Autoregressive model - Lag 1 series has the highest Z value (-26.431). It means that forecast for this month is largely influenced by last month value.
- Autoregressive model - Lag 2 series has the second-highest Z value (-10.911). It means that forecast for this month is also influenced by two months before value.

Diagnostics Plots

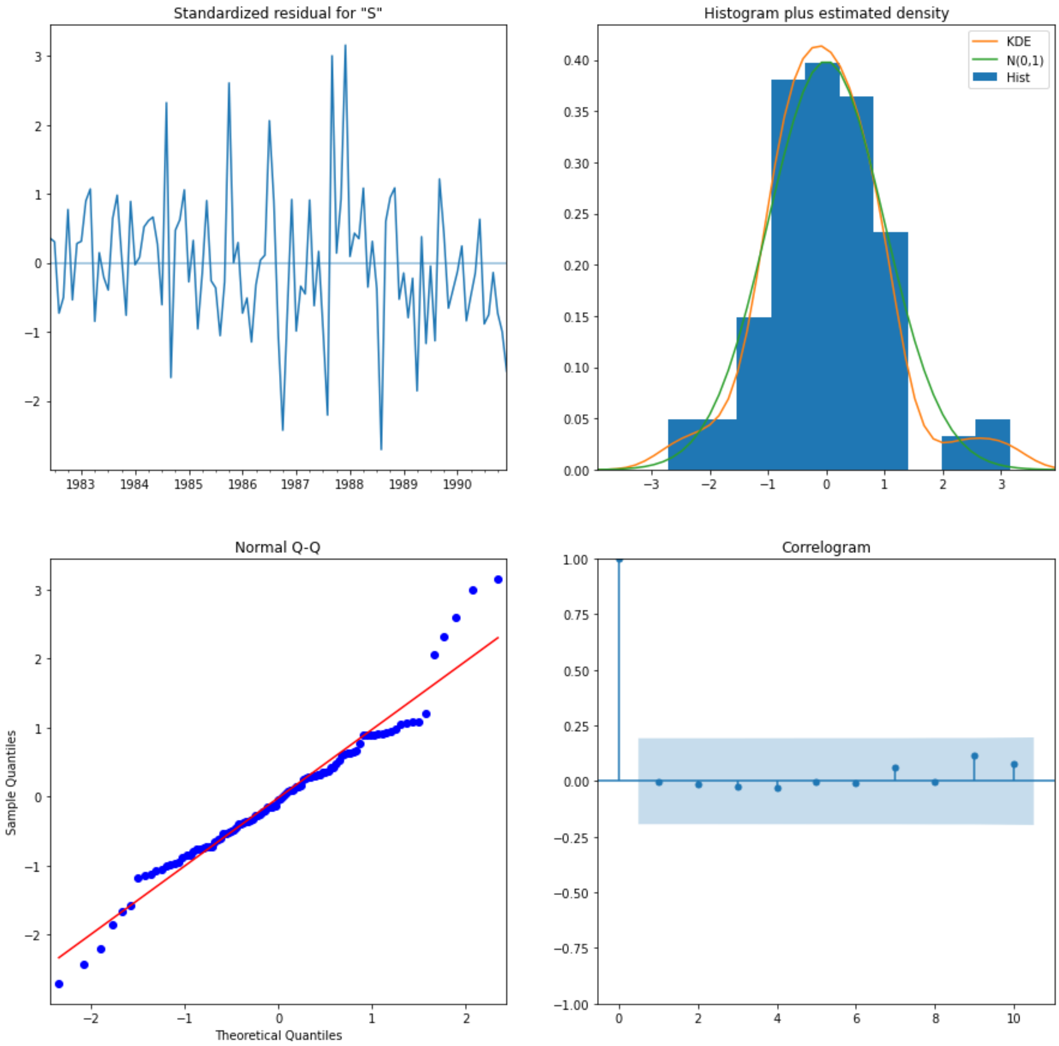


Figure 41. Diagnostics Plots of an automated SARIMA model.

YearMonth	Sparkling	forecast_ARIMA_auto	forecast_SARIMA_auto
1991-01-01	1902	4252.3	1402.7
1991-02-01	2049	2863.1	1010.0
1991-03-01	1874	2044.0	1703.2
1991-04-01	1279	1746.2	1590.4
1991-05-01	1432	1813.6	1162.4
1991-06-01	1540	2068.6	983.7
1991-07-01	2214	2365.5	1710.8
1991-08-01	1857	2612.5	1704.2
1991-09-01	2408	2770.4	2143.5
1991-10-01	3252	2839.5	3138.9

Table 45. Sample of Forecasted Sales in an automated SARIMA Model.

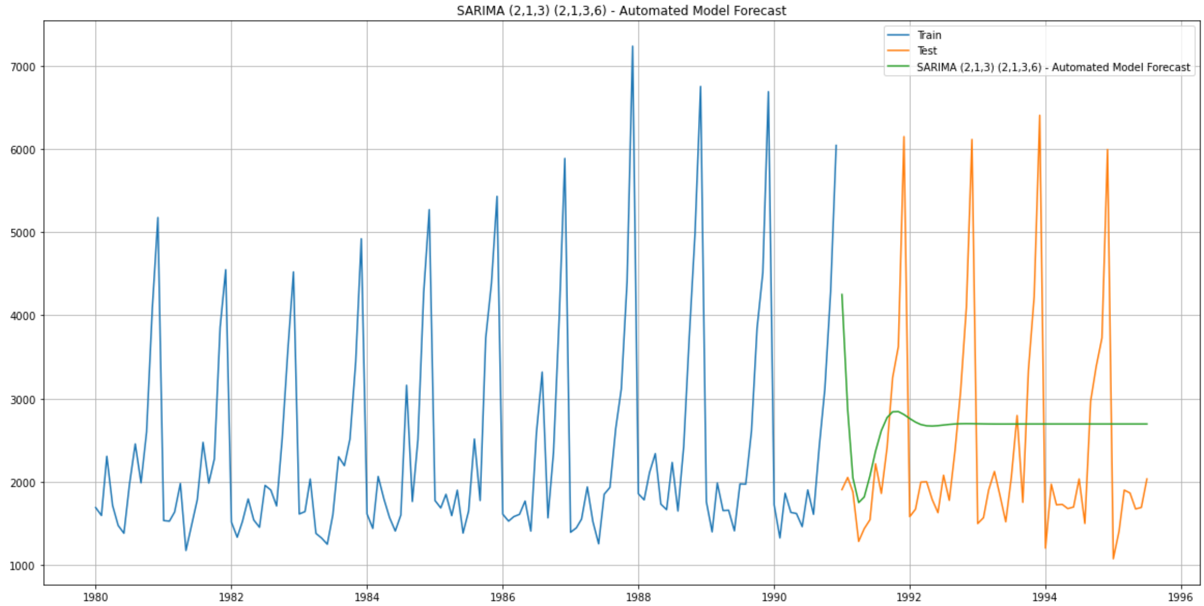


Figure 42. Plot of Forecasted Sales in an automated SARIMA model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the automated SARIMA model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in an automated SARIMA model is 784.1

Q7. 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.

Manual Version of ARIMA Model

The values for the parameters p and q are found by reading PACF and ACF plots of training data respectively.

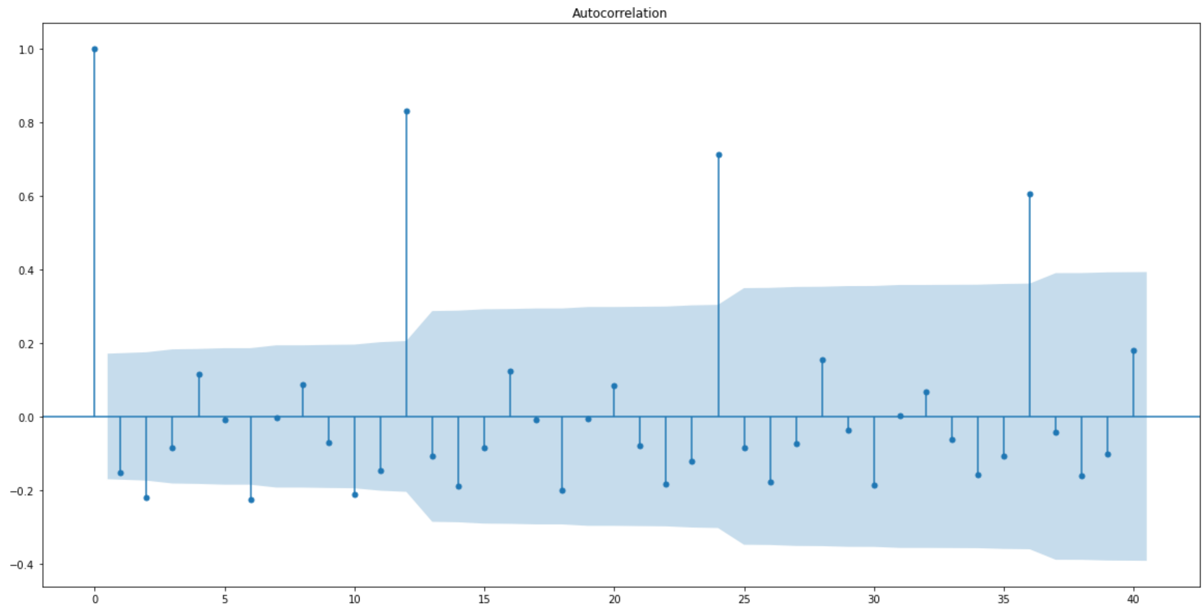


Figure 43. Autocorrelation Plot of Differenced Training Dataset

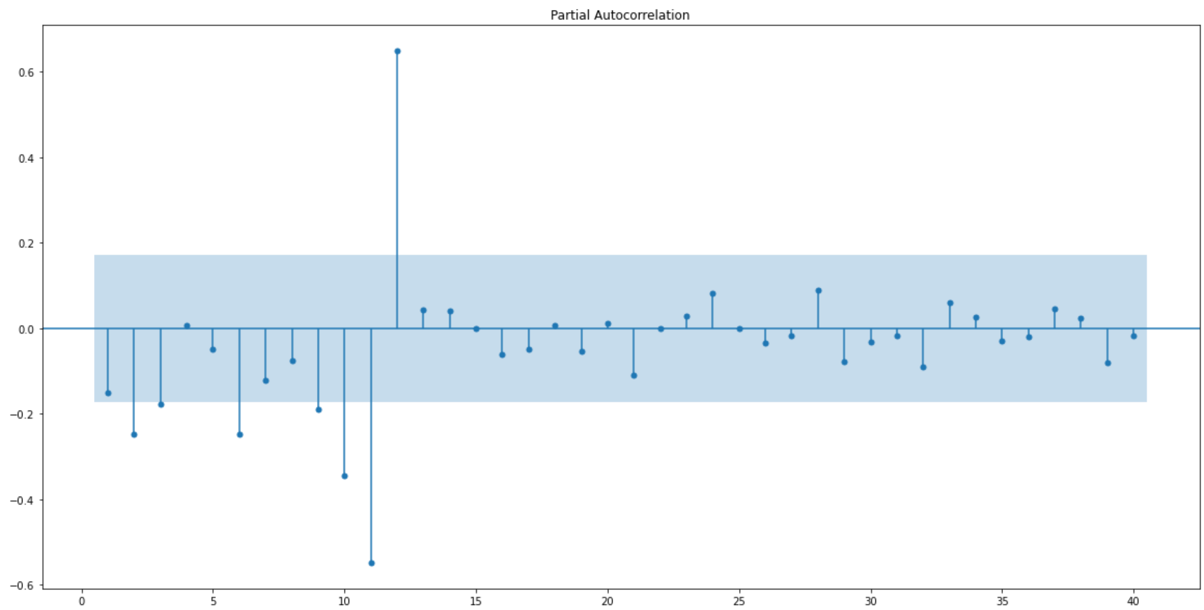


Figure 44. Partial Autocorrelation Plot of Differenced Training Dataset

Insights

- The Auto-Regressive parameter in an ARIMA model is 'p' which comes from the significant lag before which the PACF plot cuts off to 0.

- The Moving-Average parameter in an ARIMA model is 'q' which comes from the significant lag before the ACF plot cuts off to 0.
- Differencing Component, d is taken as 1 to make the time series stationarity.
- Final order to be considered is (0, 1, 0) to build the manual ARIMA model

Manual ARIMA model with order (0, 1, 0)

Summary

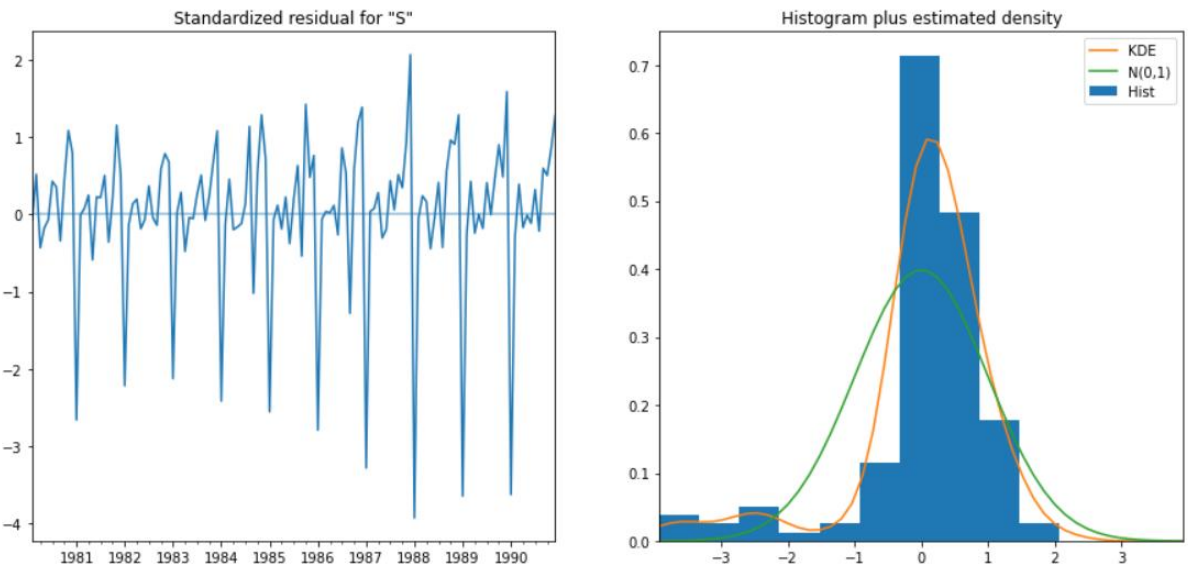
Dep. Variable:	Sparkling	**No. Observations:**	132
Model:	ARIMA(0, 1, 0)	**Log Likelihood**	-1132.832
Date:	Sun, 16 Jan 2022	**AIC**	2267.663
Time:	18:20:01	**BIC**	2270.538
Sample:	01-01-1980	**HQIC**	2268.831
	- 12-01-1990		
Covariance Type:	opg		

	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 46. Summary of a Manual ARIMA model.

Diagnostics Plots



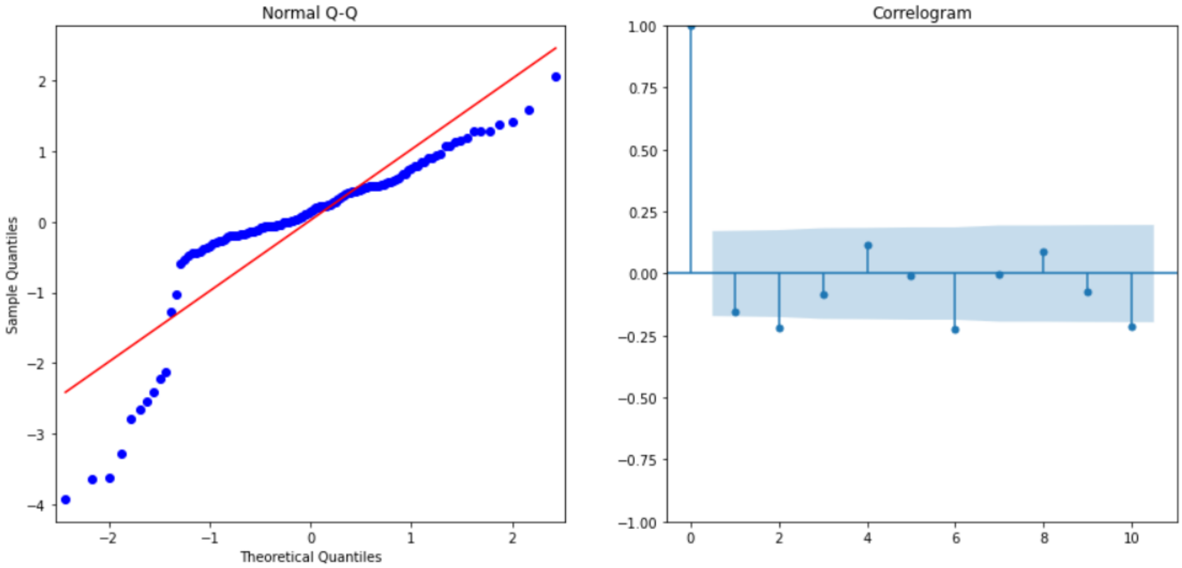


Figure 45. Diagnostics Plots of a Manual ARIMA model.

YearMonth	Sparkling	forecast_ARIMA_manual
1991-01-01	1902	6047.0
1991-02-01	2049	6047.0
1991-03-01	1874	6047.0
1991-04-01	1279	6047.0
1991-05-01	1432	6047.0
1991-06-01	1540	6047.0
1991-07-01	2214	6047.0
1991-08-01	1857	6047.0
1991-09-01	2408	6047.0
1991-10-01	3252	6047.0

Table 47. Sample of Forecasted Sales in a Manual ARIMA Model.

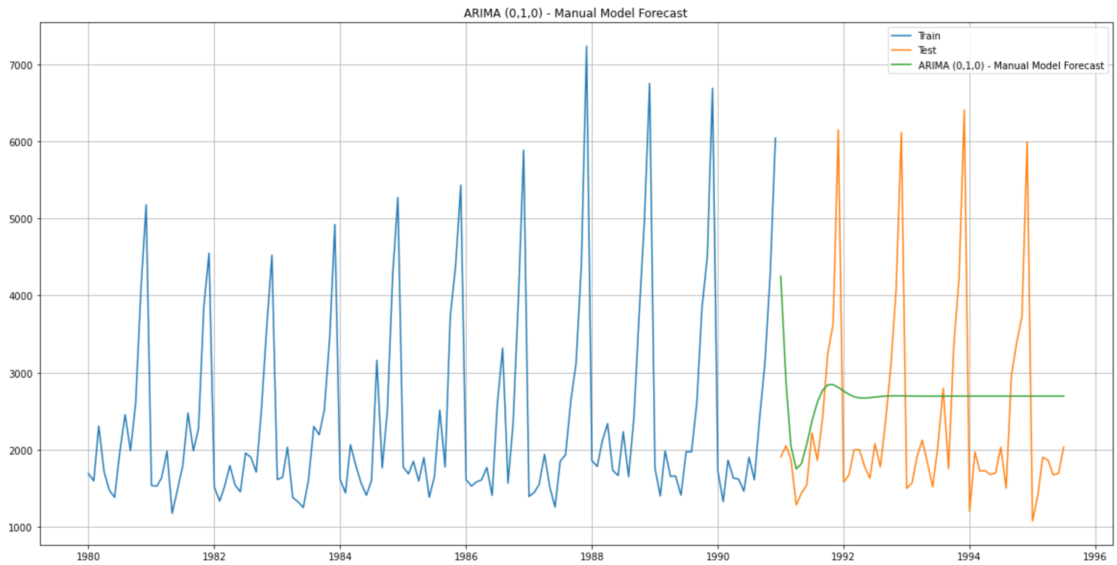


Figure 46. Plot of Forecasted Sales in a Manual ARIMA model.

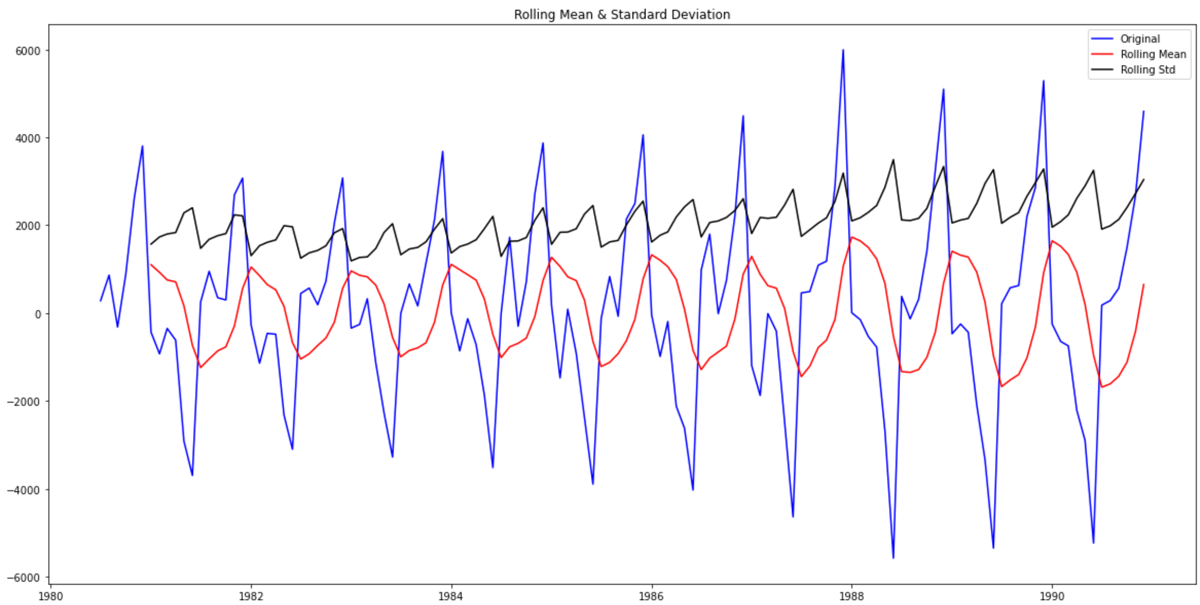
Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Manual ARIMA model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in a Manual ARIMA model is 3864.3

Manual Version of SARIMA Model

- The values for the parameters p and q are found by reading PACF and ACF plots of training data respectively.
- The values for the parameters P and Q are found by reading PACF and ACF plots of seasonally differenced training data respectively.

Results of Dickey-Fuller Test on Seasonal Differenced Training Dataset




```
Results of Dickey-Fuller Test:
Test Statistic                  -8.181919e+00
p-value                          8.088278e-13
#Lags Used                       6.000000e+00
Number of Observations Used      1.190000e+02
Critical Value (1%)             -3.486535e+00
Critical Value (5%)             -2.886151e+00
Critical Value (10%)            -2.579896e+00
dtype: float64
```

As the P-value in Dicky-Fuller test is 0 which is less than the significance level (0.05), we can reject the null hypothesis. Hence, the **seasonal differenced training dataset is stationary**.

Plotting Seasonal Differenced Time Series

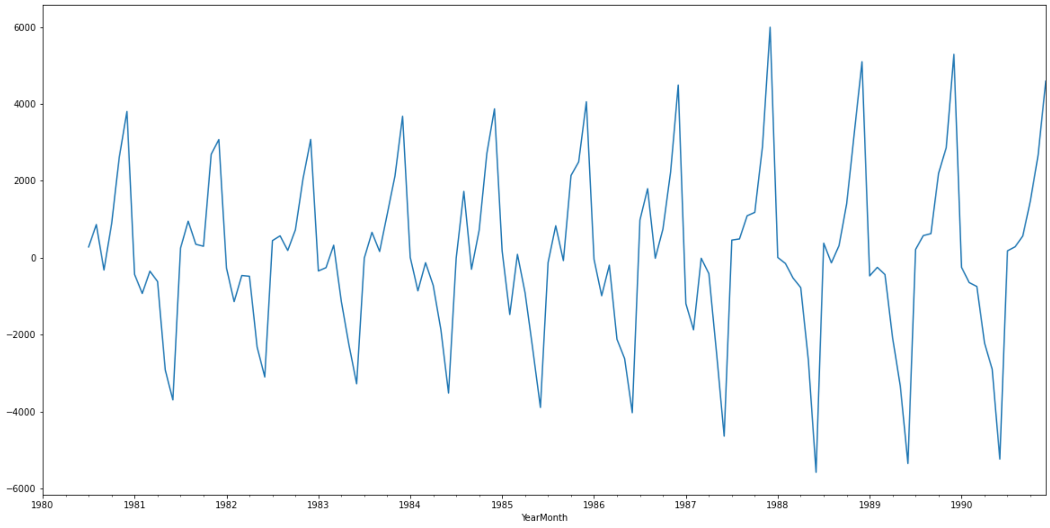


Figure 47. Plot of Seasonal Differenced Training Dataset

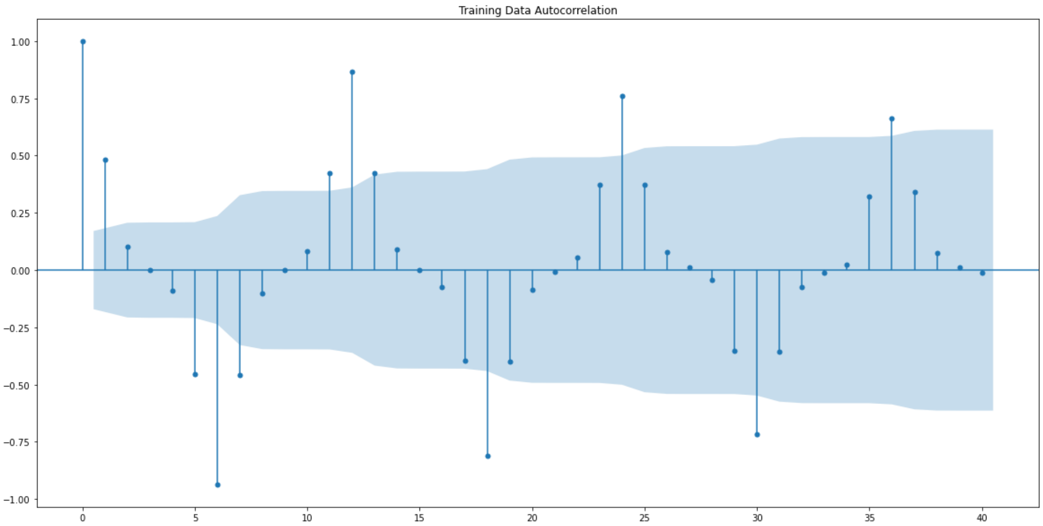


Figure 48. Autocorrelation Plot of Seasonal Differenced Training Dataset

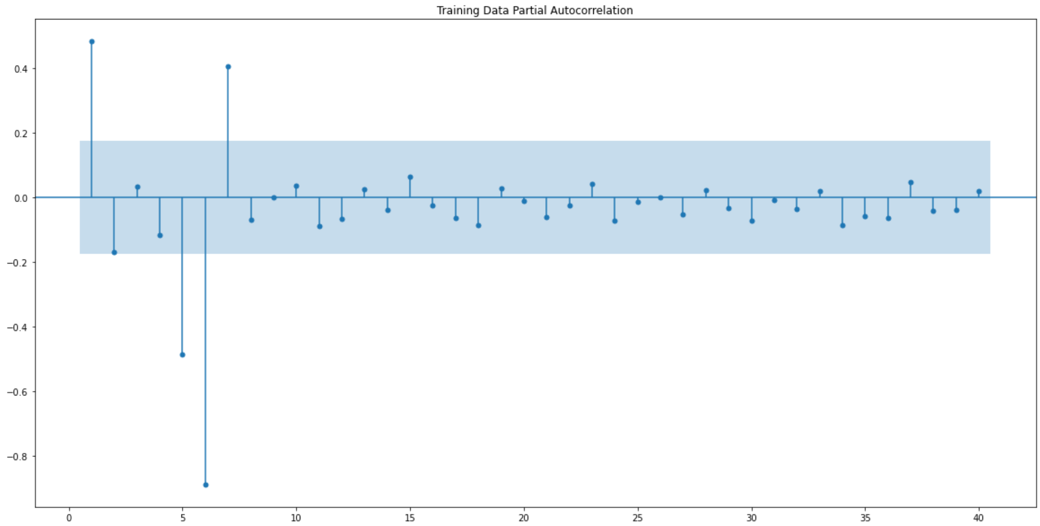


Figure 49. Partial Autocorrelation Plot of Seasonal Differenced Training Dataset

Insights

- We are taking the p-value to be 0 and the q value also to be 0 as the parameters same as the ARIMA model.
- Differencing Component, d is taken as 1 to make the time series stationarity.
- The Auto-Regressive parameter in a SARIMA model is 'P' which comes from the significant lag after which the PACF plot cuts off to 2.
- The Moving-Average parameter in a SARIMA model is 'Q' which comes from the significant lag after which the ACF plot cuts off to 1.
- Periodicity, s is taken as 6.
- Final order to be considered is (0, 1, 0) (2, 0, 1, 6) to build the manual SARIMA model

Manual SARIMA model with order (0, 1, 0) (2, 0, 1, 6)

Summary

Dep. Variable:	Sparkling	**No. Observations:**	132
Model:	SARIMAX(0, 1, 0)x(2, 0, [1], 6)	**Log Likelihood**	-916.600
Date:	Sun, 16 Jan 2022	**AIC**	1841.199
Time:	18:20:03	**BIC**	1852.316
Sample:	01-01-1980	**HQIC**	1845.713
	- 12-01-1990		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.S.L6	0.0067	0.063	0.106	0.916	-0.117	0.131
ar.S.L12	0.9490	0.033	28.908	0.000	0.885	1.013
ma.S.L6	-0.0740	0.141	-0.525	0.599	-0.350	0.202
sigma2	2.895e+05	2.67e+04	10.845	0.000	2.37e+05	3.42e+05

Ljung-Box (L1) (Q):	13.33	**Jarque-Bera (JB):**	37.71
Prob(Q):	0.00	**Prob(JB):**	0.00
Heteroskedasticity (H):	2.02	**Skew:**	0.67
Prob(H) (two-sided):	0.03	**Kurtosis:**	5.41

Table 48. Summary of a Manual SARIMA model.

Insights

- Seasonal Autoregressive model - Lag 12 series has the highest Z value (28.908). It means that forecast for this month is largely influenced by the 12 months before value.

Diagnostics Plots

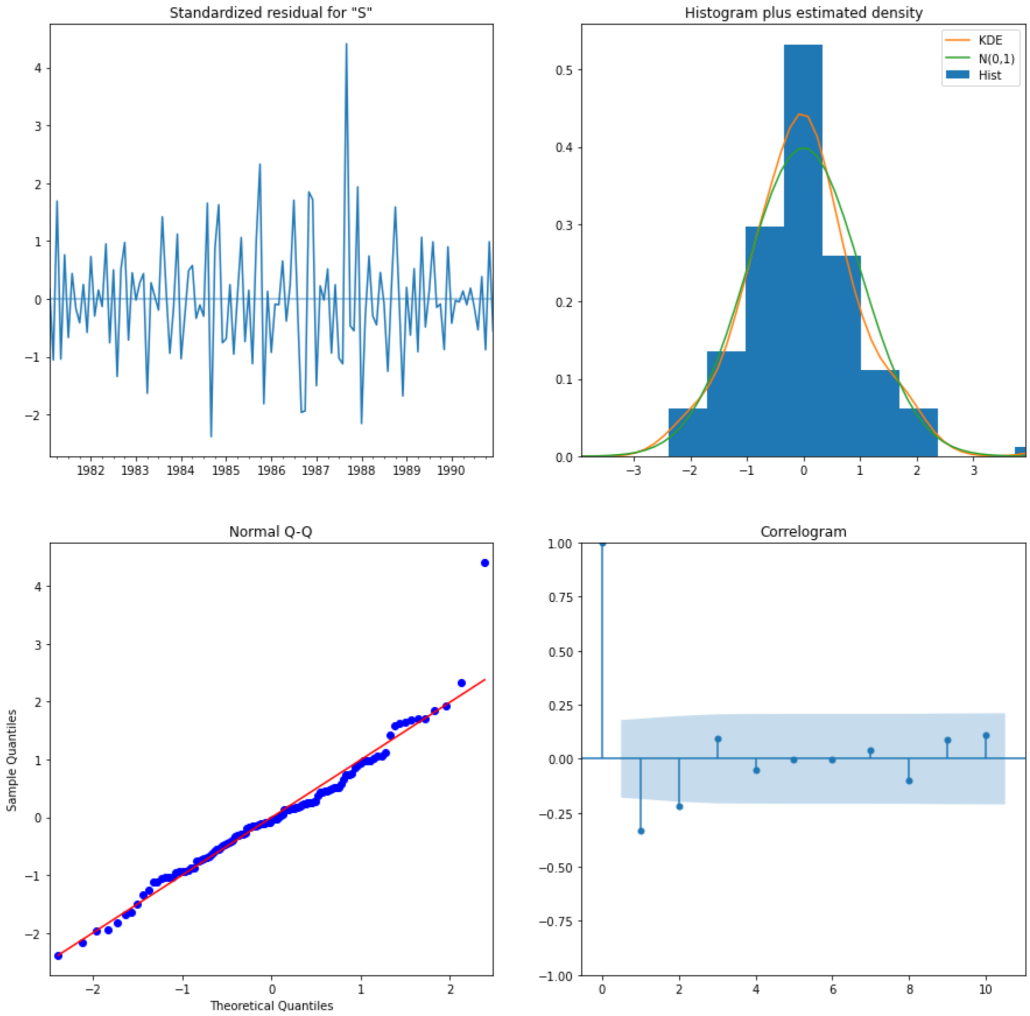


Figure 50. Diagnostics Plots of a Manual SARIMA model.

YearMonth	Sparkling	forecast_ARIMA_manual	forecast_SARIMA_manual
1991-01-01	1902	6047.0	1335.6
1991-02-01	2049	6047.0	976.4
1991-03-01	1874	6047.0	1477.2
1991-04-01	1279	6047.0	1297.8
1991-05-01	1432	6047.0	1254.0
1991-06-01	1540	6047.0	1138.0
1991-07-01	2214	6047.0	1525.8
1991-08-01	1857	6047.0	1244.4
1991-09-01	2408	6047.0	2025.0
1991-10-01	3252	6047.0	2680.5

Table 49. Sample of Forecasted Sales in a Manual SARIMA Model.

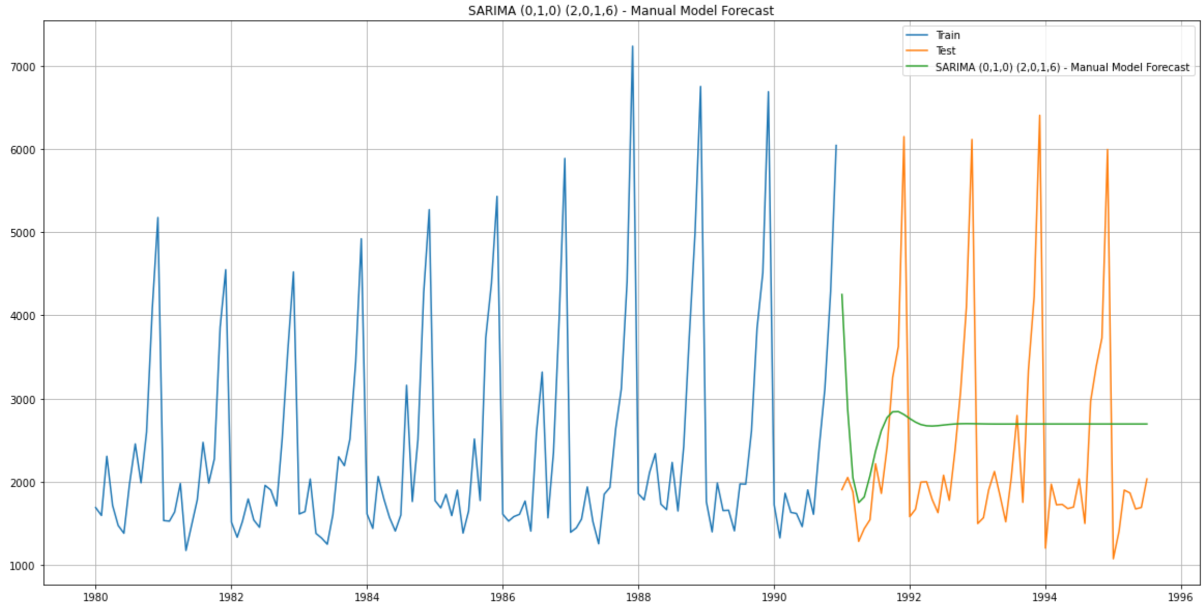


Figure 51. Plot of Forecasted Sales in a Manual SARIMA model.

Insights

- From the above plot, it can be noticed that **the plot of forecast sales is not matching with the actual plot of the test set**. Hence, the Manual SARIMA model may not be an appropriate model to forecast sales in this project. It will be concluded after comparing RMSE's of different forecast models in subsequent sections.
- RMSE in a Manual SARIMA model is 1239.3

Q8. 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.

Model	RMSE
Linear_Regression	1389.1
Naive_Forecast	3864.3
Simple_Average	1275.1
Simple_Exponential_Smoothing_optimized, Alpha =0.0703,	1338.0
Simple_Exponential_Smoothing_iteration, Alpha =0.1	1375.4
Double_Exponential_Smoothing_optimized, Alpha=0.665, Beta=0.0001	5291.9
Double_Exponential_Smoothing_iteration, Alpha=0.1, Beta=0.1	1777.7
Triple Exponential_Smoothing_with additive trend & additive seasonality_optimized, Alpha=0.1113, Beta=0.0124, Gamma=0.4607	379.0
Triple Exponential_Smoothing_with additive trend & additive seasonality_iteration, Alpha=0.1, Beta=0.4, Gamma=0.1	342.9
Triple Exponential_Smoothing_with additive trend & multiplicative seasonality_optimized, Alpha=0.1113, Beta=0.0495, Gamma=0.3621	404.3
Triple Exponential_Smoothing_with additive trend & multiplicative seasonality_iteration, Alpha=0.4, Beta=0.1, Gamma=0.2	317.4
Triple Exponential_Smoothing_with multiplicative trend & multiplicative seasonality_optimized, Alpha=0.1111, Beta=0.0494, Gamma=0.3622	380.4
Triple Exponential_Smoothing_with multiplicative trend & multiplicative seasonality_iteration, Alpha =0.4, Beta=0.1,Gamma=0.3	326.6
Triple Exponential_Smoothing_with multiplicative trend & additive seasonality_optimized, Alpha =0.1155, Beta=0.0133,Gamma=0.4565	381.2
Triple Exponential_Smoothing_with multiplicative trend & additive seasonality_iteration, Alpha =0.4, Beta=0.1,Gamma=0.2	341.7
ARIMA_Automated (2,1,2)	1300.0
ARIMA_ACF Plot (0,1,0)	3864.3
SARIMA_Automated (2,1,3)(2,1,3,6)	784.1
SARIMA_ACF Plot (0,1,0)(2,0,1,6)	1239.3

Let us sort the data frame by RMSE values to find the best model

Model	RMSE
Triple Exponential_Smoothing_with additive trend & multiplicative seasonality_iteration, Alpha=0.4, Beta=0.1, Gamma=0.2	317.4
Triple Exponential_Smoothing_with multiplicative trend & multiplicative seasonality_iteration, Alpha =0.4, Beta=0.1,Gamma=0.3	326.6
Triple Exponential_Smoothing_with multiplicative trend & additive seasonality_iteration, Alpha =0.4, Beta=0.1,Gamma=0.2	341.7
Triple Exponential_Smoothing_with additive trend & additive seasonality_iteration, Alpha=0.1, Beta=0.4, Gamma=0.1	342.9
Triple Exponential_Smoothing_with additive trend & additive seasonality_optimized, Alpha=0.1113, Beta=0.0124, Gamma=0.4607	379.0
Triple Exponential_Smoothing_with multiplicative trend & multiplicative seasonality_optimized, Alpha=0.1111, Beta=0.0494, Gamma=0.3622	380.4
Triple Exponential_Smoothing_with multiplicative trend & additive seasonality_optimized, Alpha =0.1155, Beta=0.0133,Gamma=0.4565	381.2
Triple Exponential_Smoothing_with additive trend & multiplicative seasonality_optimized, Alpha=0.1113, Beta=0.0495, Gamma=0.3621	404.3
SARIMA_Automated (2,1,3)(2,1,3,6)	784.1
SARIMA_ACF Plot (0,1,0)(2,0,1,6)	1239.3
Simple_Average	1275.1
ARIMA_Automated (2,1,2)	1300.0
Simple_Exponential_Smoothing_optimized, Alpha =0.0703,	1338.0
Simple_Exponential_Smoothing_iteration, Alpha =0.1	1375.4
Linear_Regression	1389.1
Double_Exponential_Smoothing_iteration, Alpha=0.1, Beta=0.1	1777.7
Naive_Forecast	3864.3
ARIMA_ACF Plot (0,1,0)	3864.3
Double_Exponential_Smoothing_optimized, Alpha=0.665, Beta=0.0001	5291.9

Table 50. All Forecast Models with Respective Parameters and RMSE's

From the above table, it is evident that the best or most optimum model is the Triple Exponential Smoothing with Additive Trend & Multiplicative seasonality with the Parameters $\alpha = 0.4$, $\beta = 0.1$ and $\gamma = 0.2$.

Q9. 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.

- Let us build Triple Exponential Smoothing with Additive Trend & Multiplicative seasonality with the Parameters $\alpha = 0.4$, $\beta = 0.1$ and $\gamma = 0.2$ (most optimum model) on full data.
- RMSE obtained in full model is 377.3

	lower_CI	Forecast	upper_ci
1995-08-01	1321.9	2063.4	2804.8
1995-09-01	1838.3	2579.8	3321.3
1995-10-01	2676.6	3418.1	4159.6
1995-11-01	3567.1	4308.6	5050.1
1995-12-01	5874.3	6615.8	7357.3
1996-01-01	825.5	1566.9	2308.4
1996-02-01	1111.0	1852.5	2593.9
1996-03-01	1359.6	2101.0	2842.5
1996-04-01	1282.7	2024.2	2765.7
1996-05-01	1094.1	1835.5	2577.0
1996-06-01	971.7	1713.2	2454.7
1996-07-01	1435.7	2177.2	2918.6

Table 51. Forecasted Sales with 95% Confidence Interval.

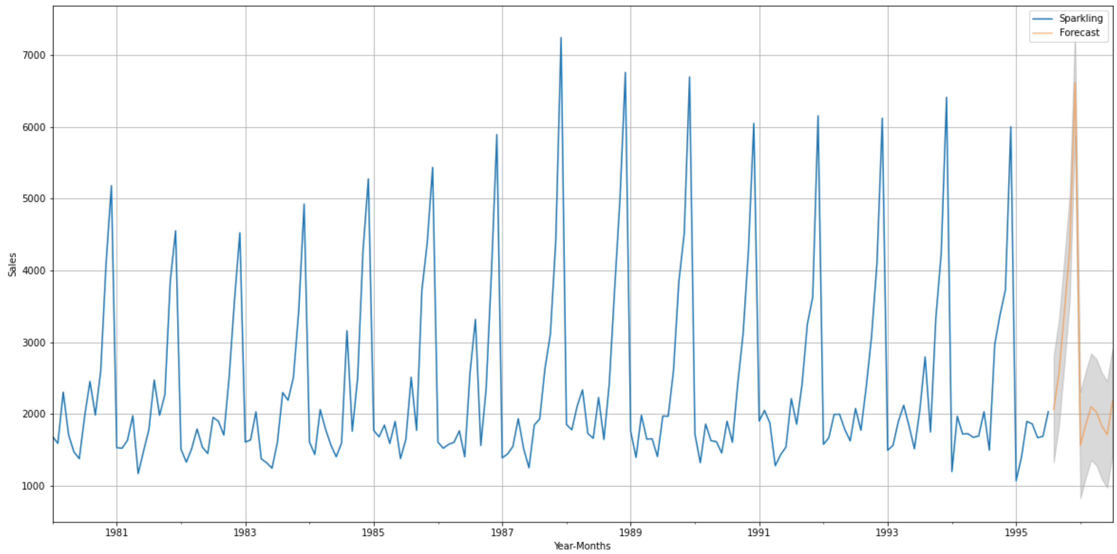


Figure 52. Plot of Forecasted Sales with 95% Confidence Interval.

Q10. Comment on the model thus built and report your findings and suggest the measures that the company should be taking for future sales.

- The best **or most optimum model is the Triple Exponential Smoothing with Additive Trend & Multiplicative seasonality** with the Parameters $\alpha = 0.4$, $\beta = 0.1$ and $\gamma = 0.2$.
- RMSE obtained in the most optimum model on the test data set is 317.4
- RMSE obtained when this model is fitted on the full dataset is 377.3.
- The increase in RMSE is marginal. Hence, the model is stable and This model can be used for forecasting future sales.
- The plot of 12 months forecasted sales is following the approximately same pattern as that of the original time series.
- It is expected that every December wines sales reach to maximum. Hence, the company should be ready with production to meet the demand at end of the year.
- The sales of wine have some outliers for certain years
- Initially mean wine sales in a year are increasing up to 1988 and then start decreasing.
- Total wine sales in a year are following almost the same pattern as that of the mean wine sales in a year.
- Sales are increasing gradually from quarter1 to quarter 3 and then increase exponentially from quarter 3 to quarter 4.
- Mean sales in a month are not following any significant pattern up to September month and then increase exponentially.
- Total sales in a month are following almost the same pattern as that of the mean sales in a month.