BUSINESS REPORT

Time Series Forecasting

Table of Contents:

List of Figures:

List of Tables:

Objective:

The primary objective of this project is to analyse and forecast the sales trends of sparkling and rose wines using historical data provided by ABC Estate Wines. By employing time series analysis and forecasting techniques, we aim to:

- Identify key trends and seasonal patterns in the sales data.
- Build accurate forecasting models to predict future sales.
- Provide actionable insights and recommendations to optimize sales strategies.

1.1 Define the problem and perform Exploratory Data Analysis:

	YearMonth	Sparkling
0	1980-01	1686
1	1980-02	1591
2	1980-03	2304
3	1980-04	1712
4	1980-05	1471
...	...	...
182	1995-03	1897
183	1995-04	1862
184	1995-05	1670
185	1995-06	1688
186	1995-07	2031

187 rows × 2 columns

Table 1.1.1 Dataset of Sparkling wine

	YearMonth	Rose
0	1980-01	112.0
1	1980-02	118.0
2	1980-03	129.0
3	1980-04	99.0
4	1980-05	116.0
...	...	...
182	1995-03	45.0
183	1995-04	52.0
184	1995-05	28.0
185	1995-06	40.0
186	1995-07	62.0

187 rows × 2 columns

Table 1.1.2 Dataset of Rose wine

```
<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: 3.1+ KB
```

Fig 1.1.1 Info of Sparkling wine

```
<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   Rose       185 non-null    float64
dtypes: float64(1), object(1)
memory usage: 3.1+ KB
```

Fig 1.1.2 Info of Rose wine

Problem definition:

The primary objective of this project is to analyse and forecast the sales trends of sparkling and rose wines using historical data provided by ABC Estate Wines. By employing time series analysis and forecasting techniques, we aim to:

- Identify key trends and seasonal patterns in the sales data.
- Build accurate forecasting models to predict future sales.
- Provide actionable insights and recommendations to optimize sales strategies.

1.1 Reading the data as an 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

Table 1.1.3 Time series of Sparkling wine

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

Table 1.1.4 Time series of Rose wine

1.2 Plotting the data:

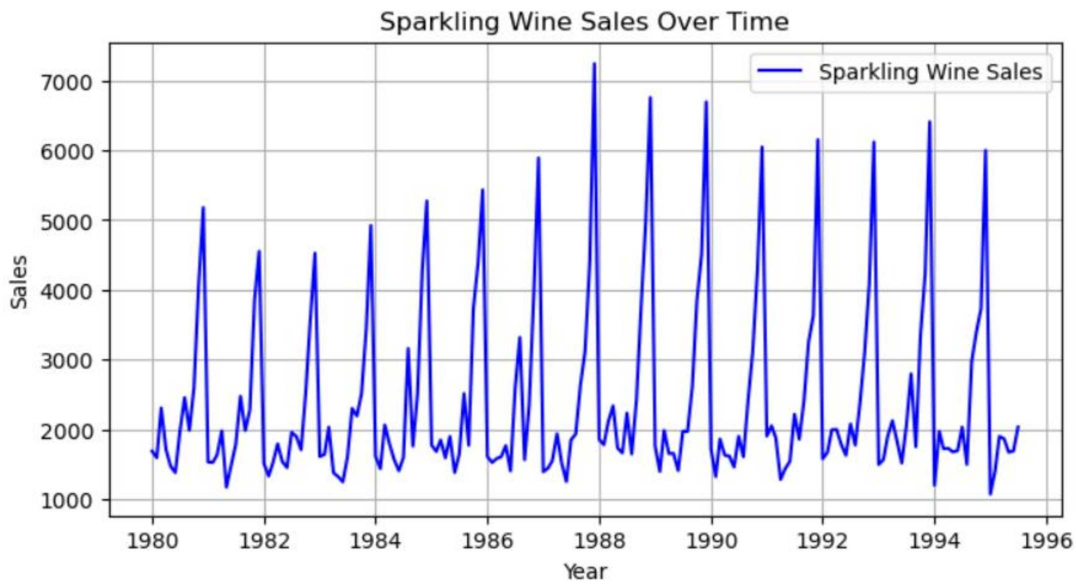


Fig 1.2.1 Sparkling Wine sales over time

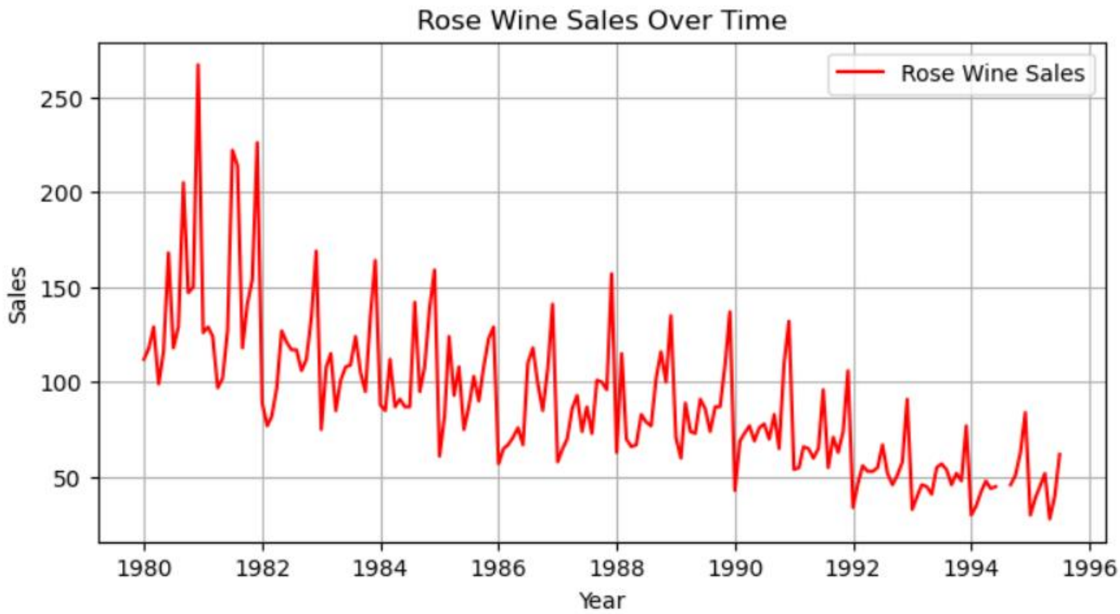


Fig 1.2.2 Rose Wine sales over time

1.3 Perform EDA :

Summary statistics for sparkling_data:

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

Table 1.3.1 Summary statistics of Sparkling wine

Summary statistics for rose_data:

	Rose
count	185.000000
mean	90.394595
std	39.175344
min	28.000000
25%	63.000000
50%	86.000000
75%	112.000000
max	267.000000

Table 1.3.2 Summary statistics of Rose wine

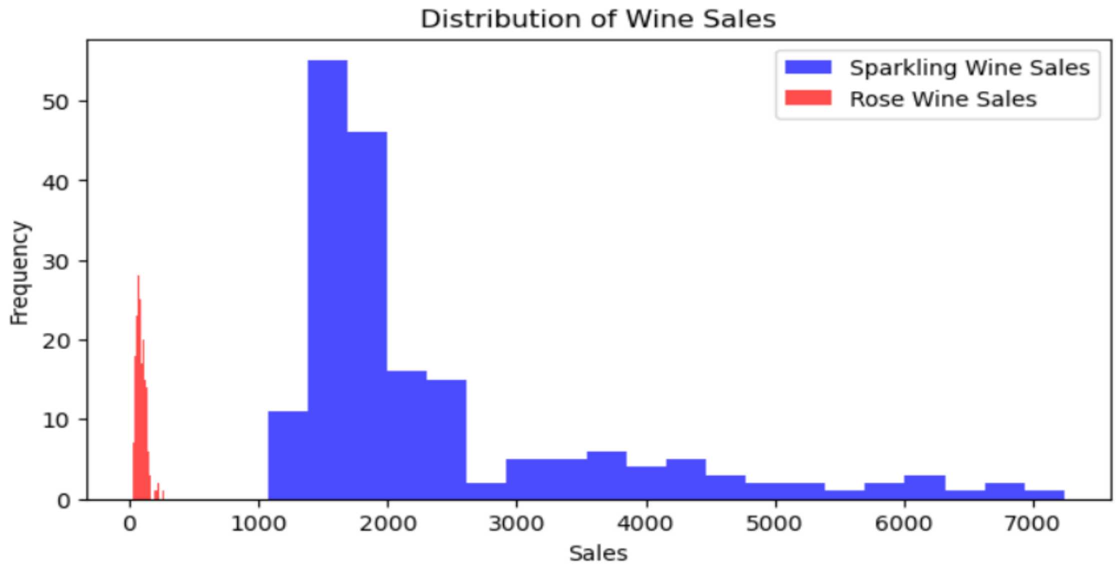


Fig 1.3.1 Distribution of Wine sales

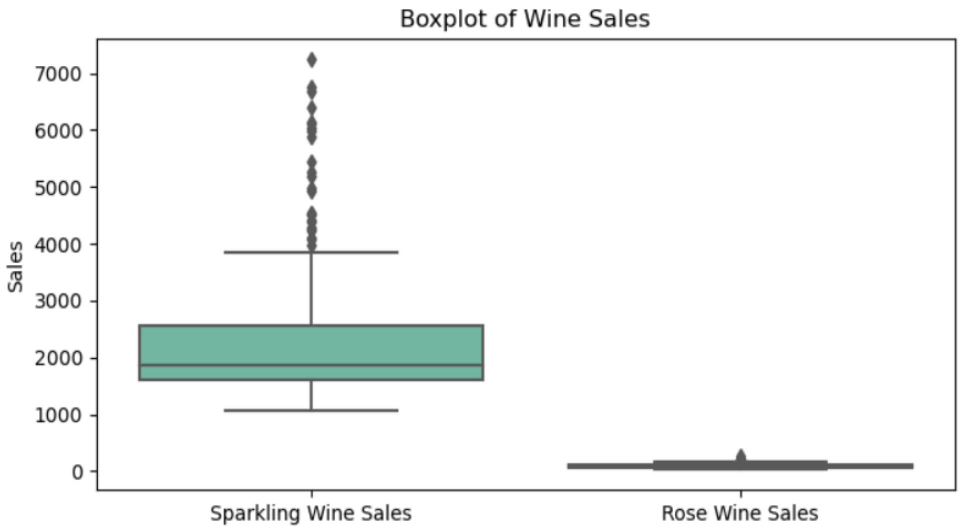


Fig 1.3.2 Box plot of Wine sales

Correlation between Sparkling and Rose Wine Sales:

	Sparkling	Rose
Sparkling	1.000000	0.404579
Rose	0.404579	1.000000

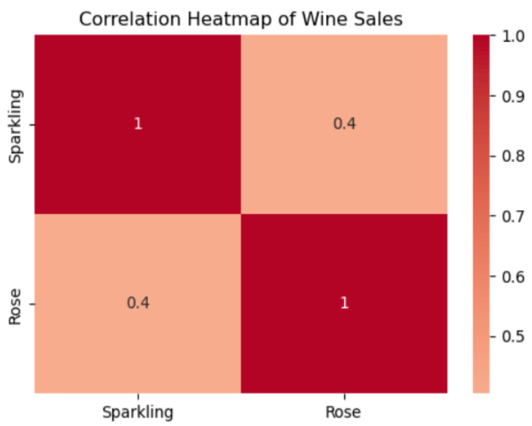


Fig 1.3.3 Correlation Heat map of Wine sales

1.4 Perform Decomposition:

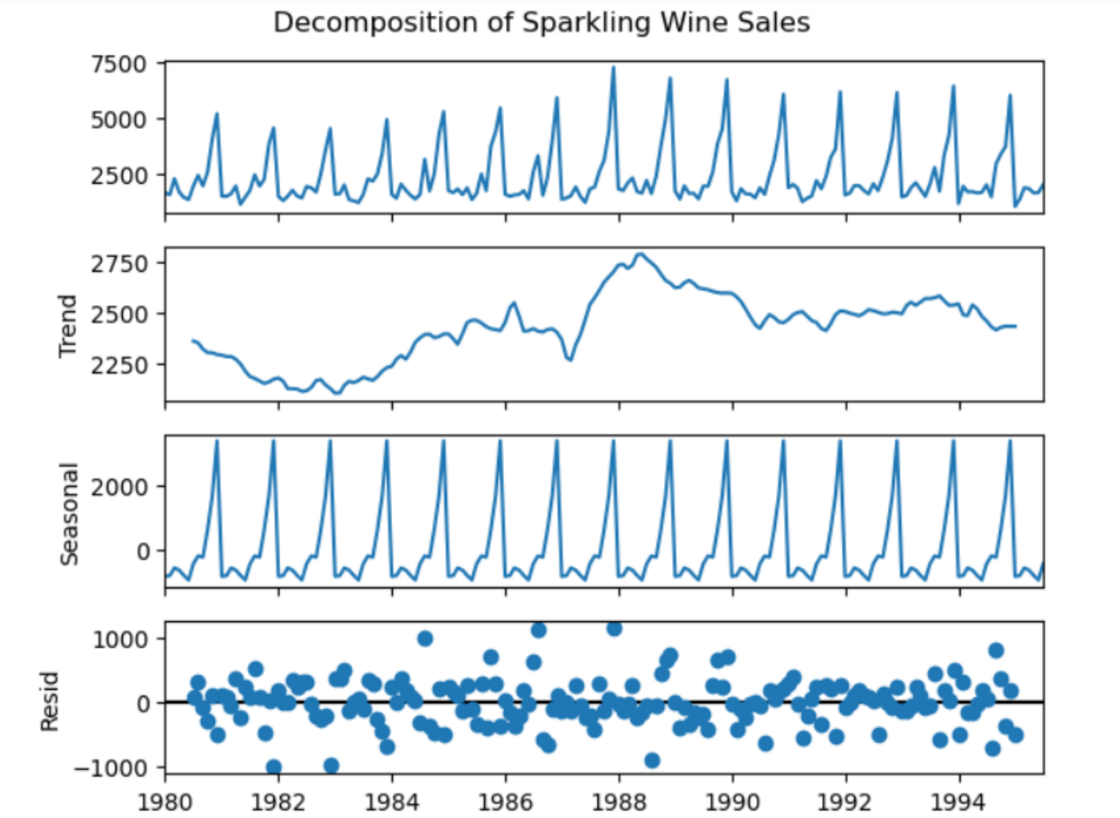


Fig 1.4.1 Decomposition of Sparkling Wine sales

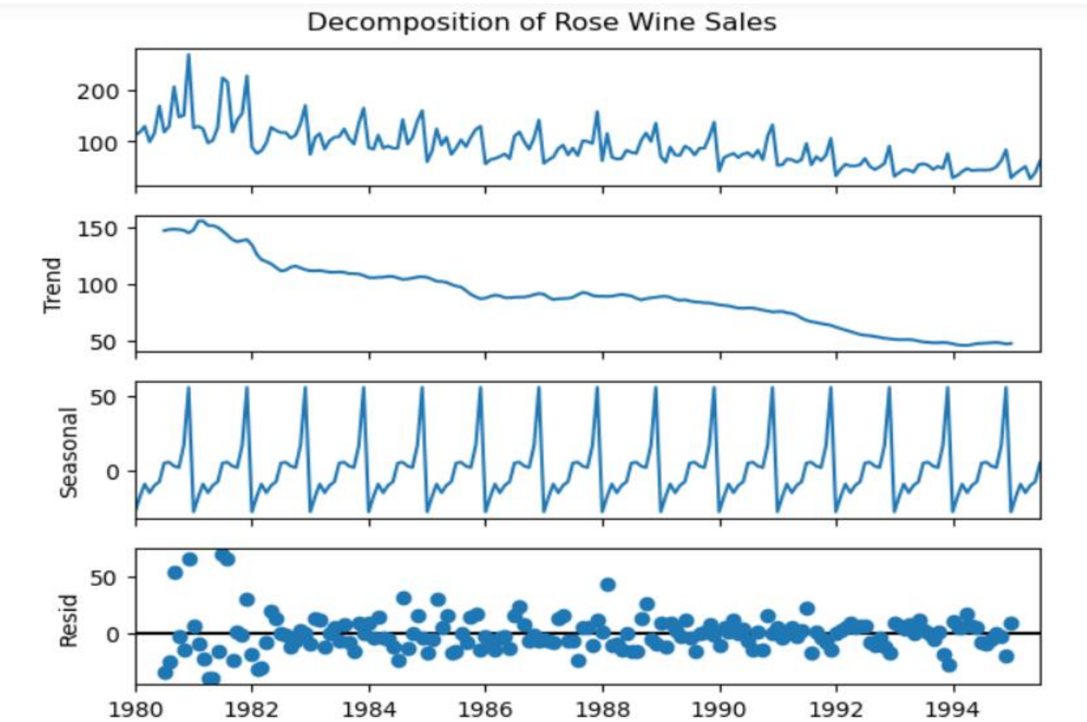


Fig 1.4.2 Decomposition of Rose Wine sales

Decomposition of Sparkling Wine Sales:

Trend Component:

- The “Trend” plot reveals a gradual upward trend over the same period.
- This indicates overall growth in sparkling wine sales.

Seasonal Variations:

- The “Seasonal” plot exhibits sharp spikes at regular intervals.
- These spikes represent seasonal variations (e.g., higher sales during holidays or specific seasons).

Residuals (Noise):

- The “Resid” plot shows random scatter around the zero line.
- Residuals represent unexplained variation after accounting for trend and seasonality.

Decomposition of Rose Wine Sales:

Trend Component:

- The trend component shows a steady decline in rose wine sales over the years.

Seasonal Component:

- The seasonal component reveals a consistent pattern of sales peaking at regular intervals, indicating seasonality in the data.

Residual Component: The residual component shows random fluctuations around the zero line, indicating variability that is not explained by the trend or seasonal components.

Summary:

- Sparkling Wine: Exhibits a strong and increasing trend with pronounced seasonality. Sales peaks are regular and significant, showing high variability throughout the year.

- Rose Wine: Shows a declining trend with less pronounced seasonality. Sales peaks are less consistent and decline over time, indicating a decrease in popularity or demand.

Data Pre-processing:

1.5 Missing value treatment:

- Missing values in sparkling_data before treatment:

Sparkling: 0

- Missing values in rose_data before treatment:

Rose: 0

- Missing values in sparkling_data after treatment:

Sparkling 0

- Missing values in rose_data after treatment:

Rose 0

During decomposition we have done the missing value treatment for rose_data.

1.6 Visualize the processed data:

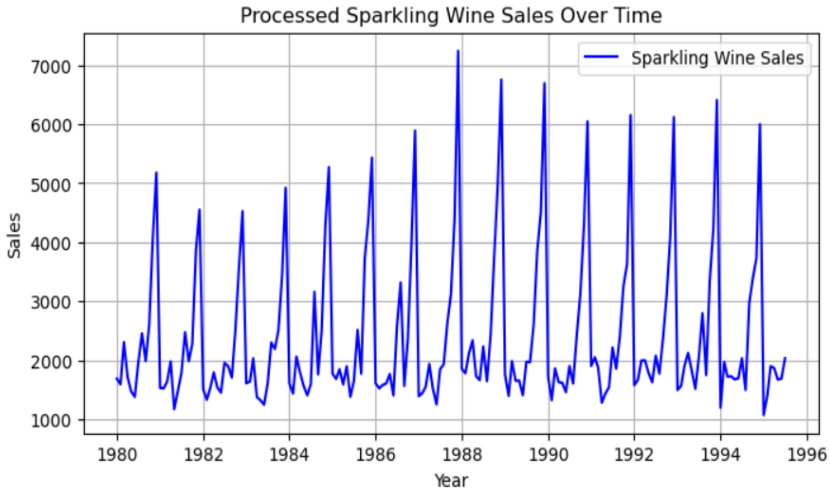


Fig 1.6.1 Sparkling wine sales

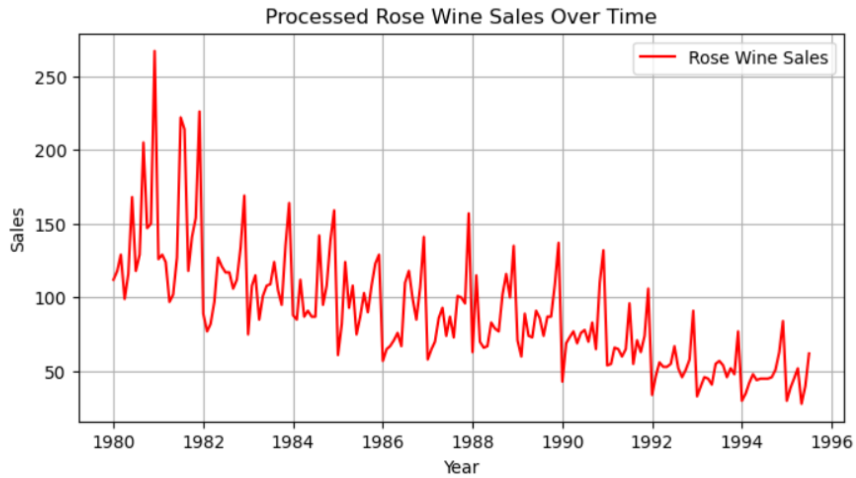


Fig 1.6.2 Rose wine sales

1.7 Train-test split:

- Train Sparkling Shape: (149, 1), Test Sparkling Shape: (38, 1)
- Train Rose Shape: (149, 1), Test Rose Shape: (38, 1)

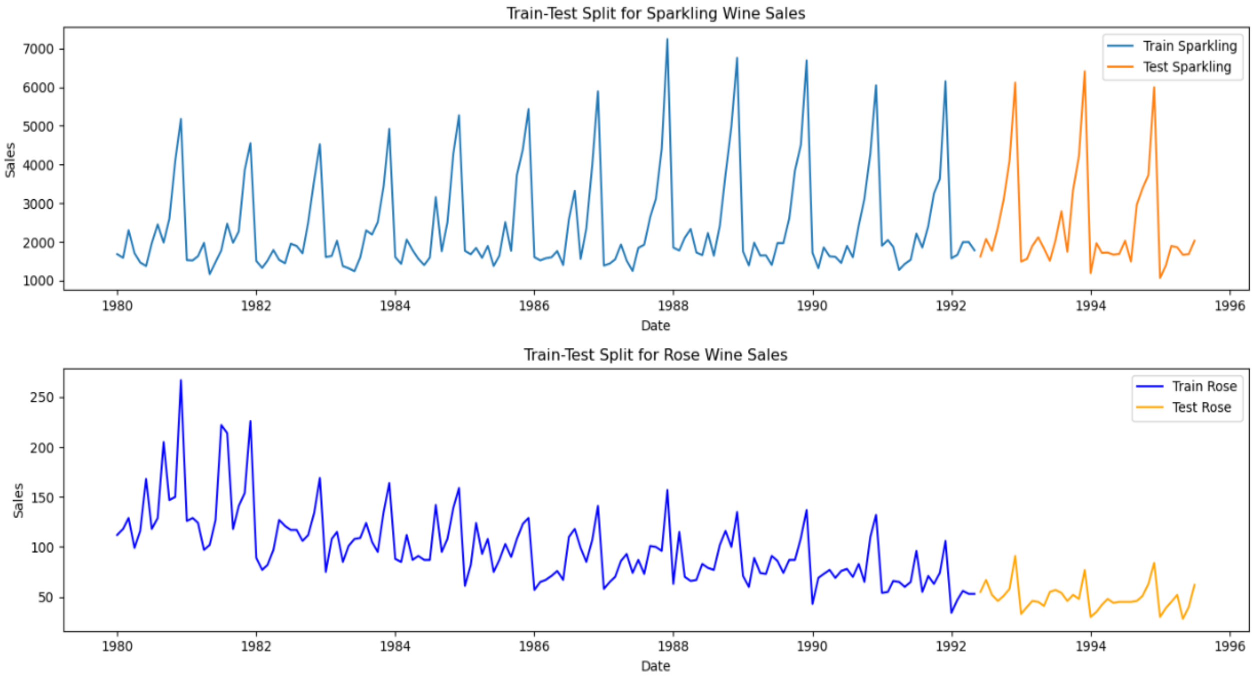


Fig 1.7.1 Train-Test Split for Sparkling & Rose wine sales

Model Building- Original Data:

1.8 Build forecasting models:

ARIMA Model for Sparkling Wine Sales:

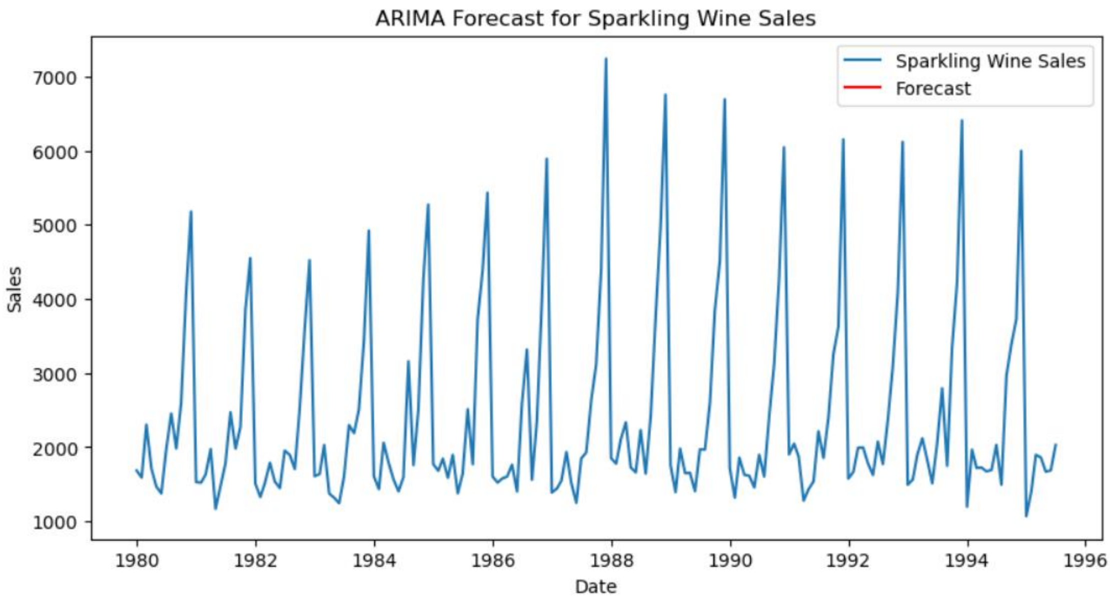


Fig 1.8.1 ARIMA Forecast for Sparkling wine sales

ARIMA Model for Rose Wine Sales:

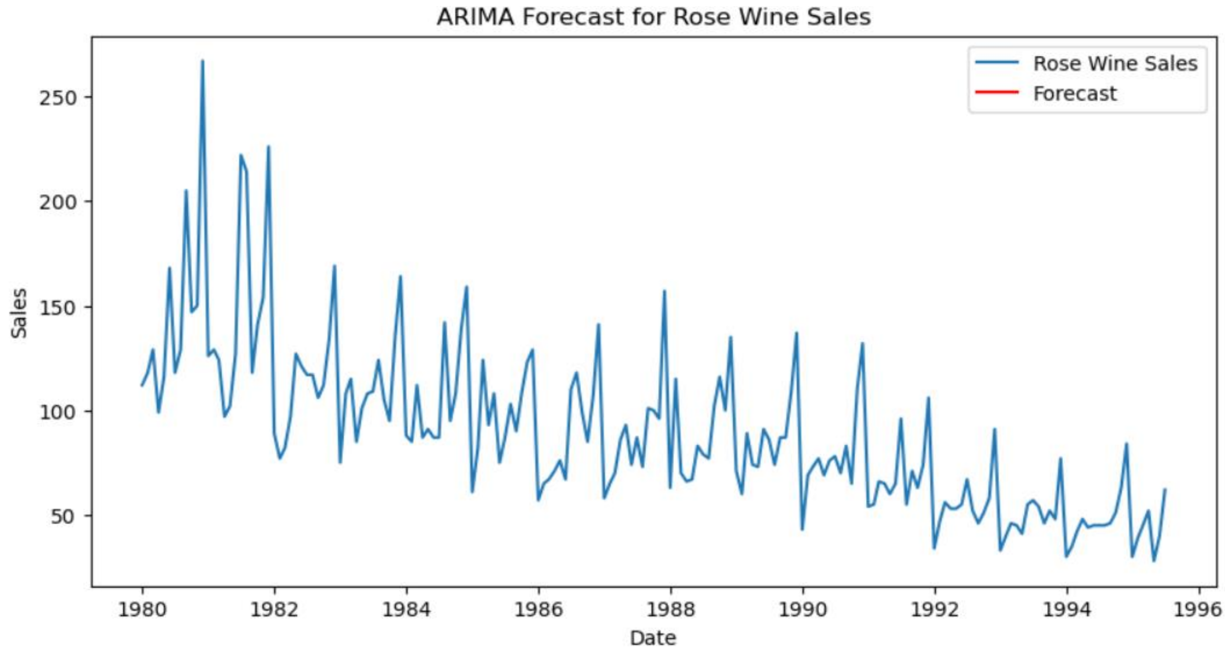


Fig 1.8.2 ARIMA Forecast for Rose wine sales

SARIMA Model for Sparkling Wine Sales:

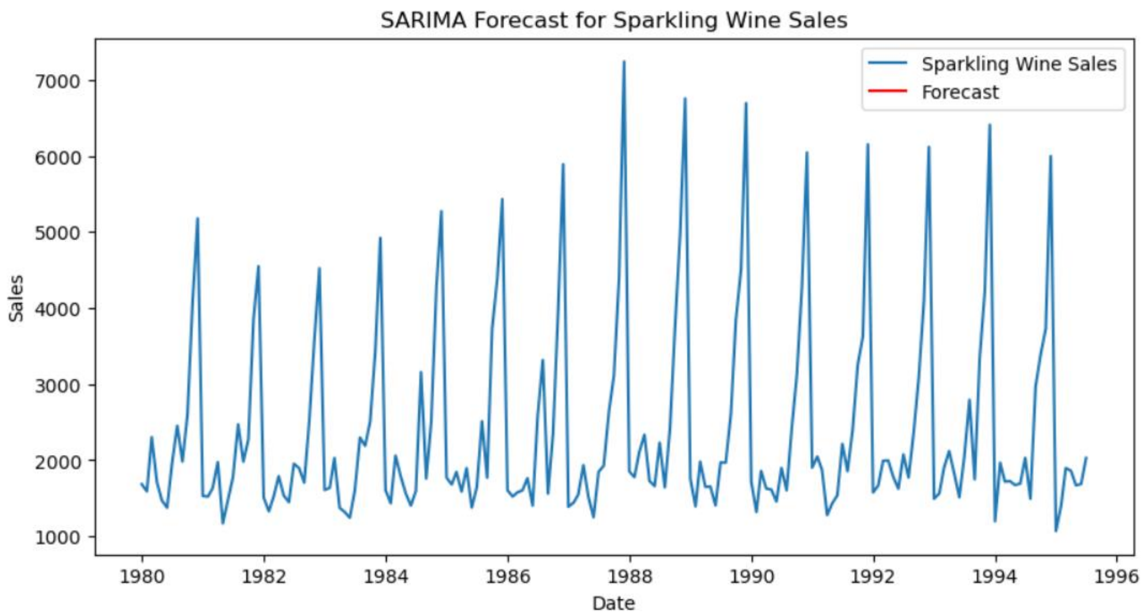


Fig 1.8.3 SARIMA Forecast for Sparkling wine sales

SARIMA Model for Rose Wine Sales:

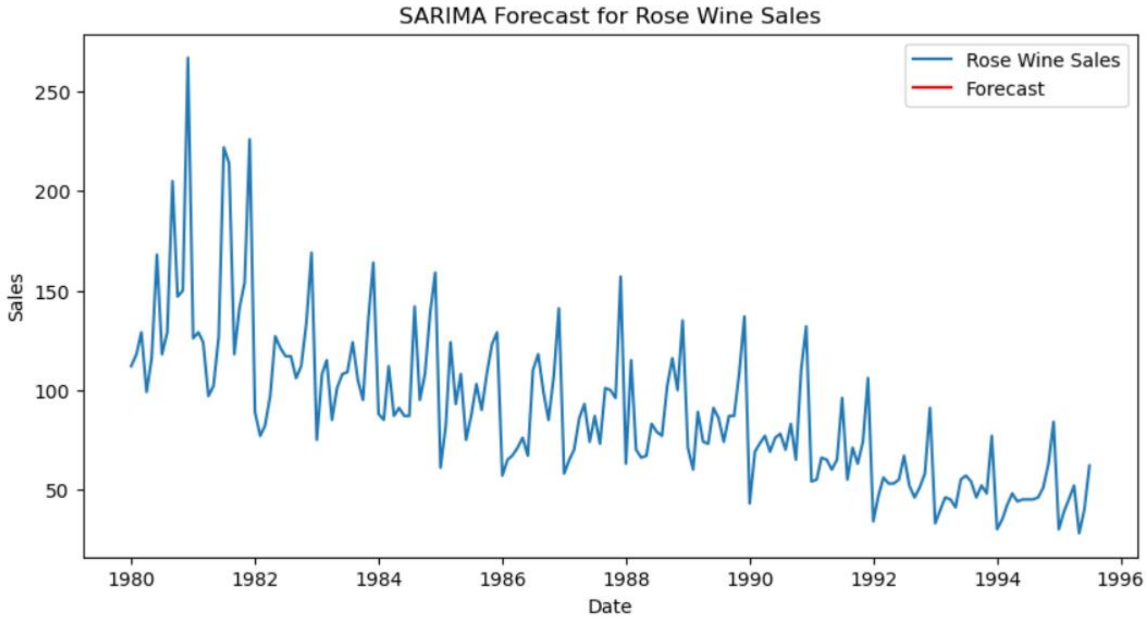


Fig 1.8.4 SARIMA Forecast for Rose wine sales

1.9 Linear regression:

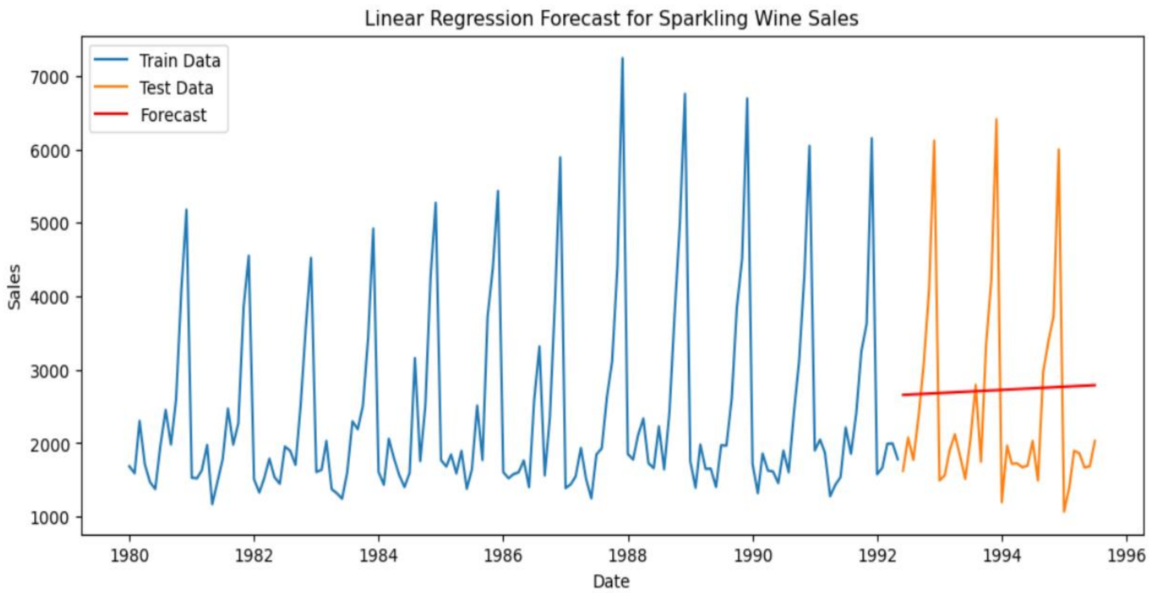


Fig 1.9.1 Linear Regression Forecast for sparkling wine sales

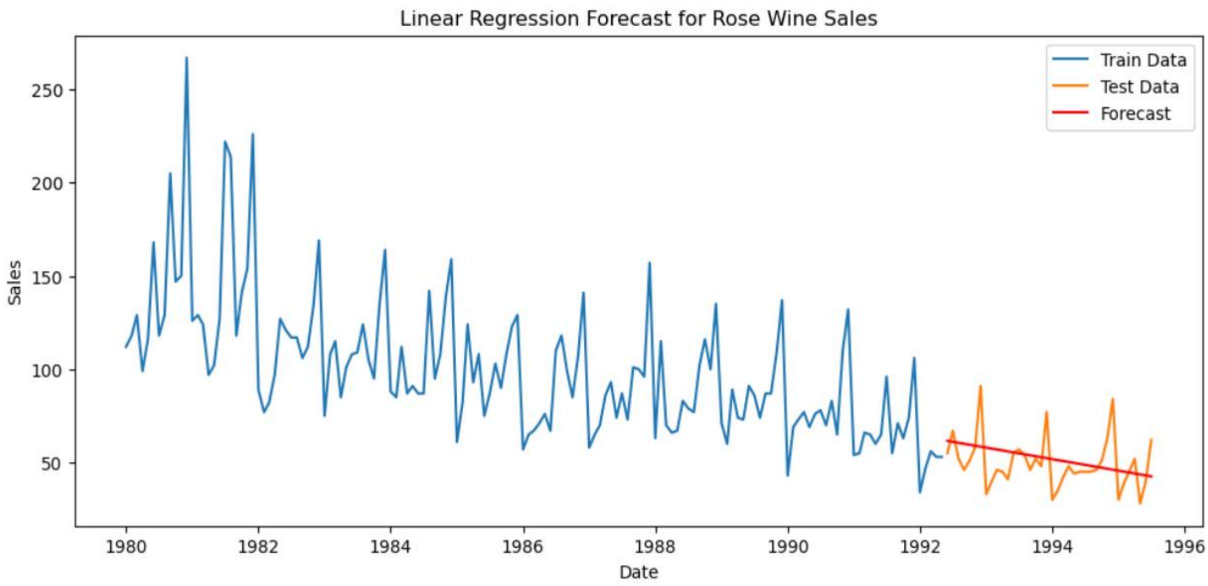


Fig 1.9.2 Linear Regression Forecast for Rose wine sales

1.10 Simple Average:

Simple Average for Sparkling Sales:

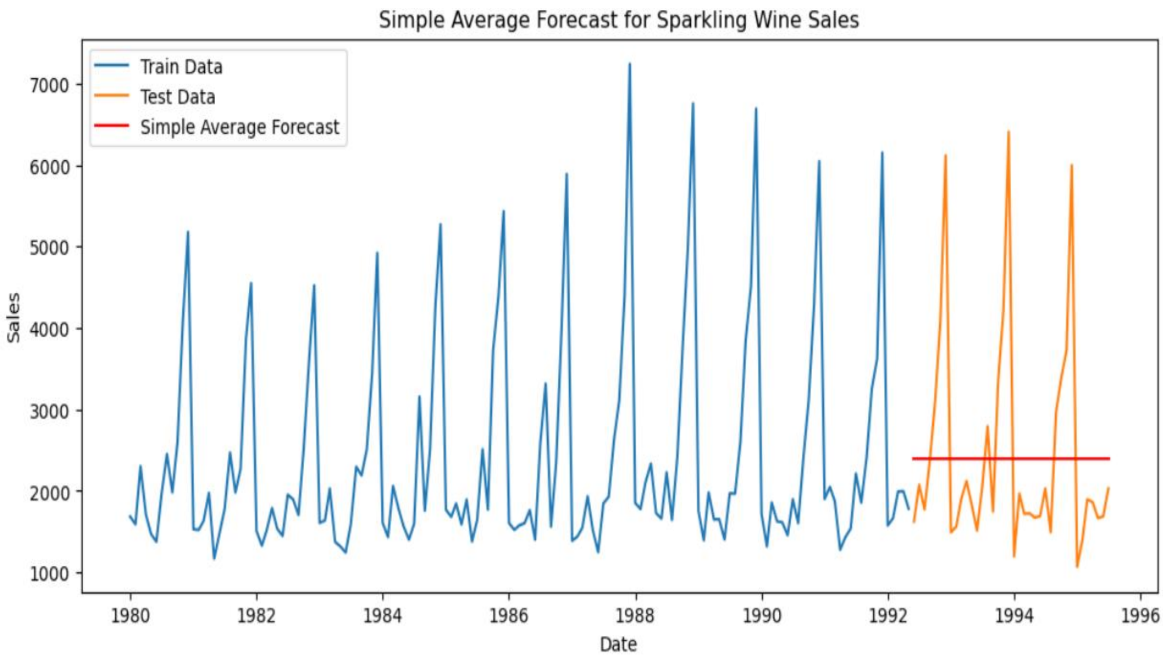


Fig 1.10.1 Simple Average Forecast for sparkling wine sales

Simple Average for Rose wine Sales:

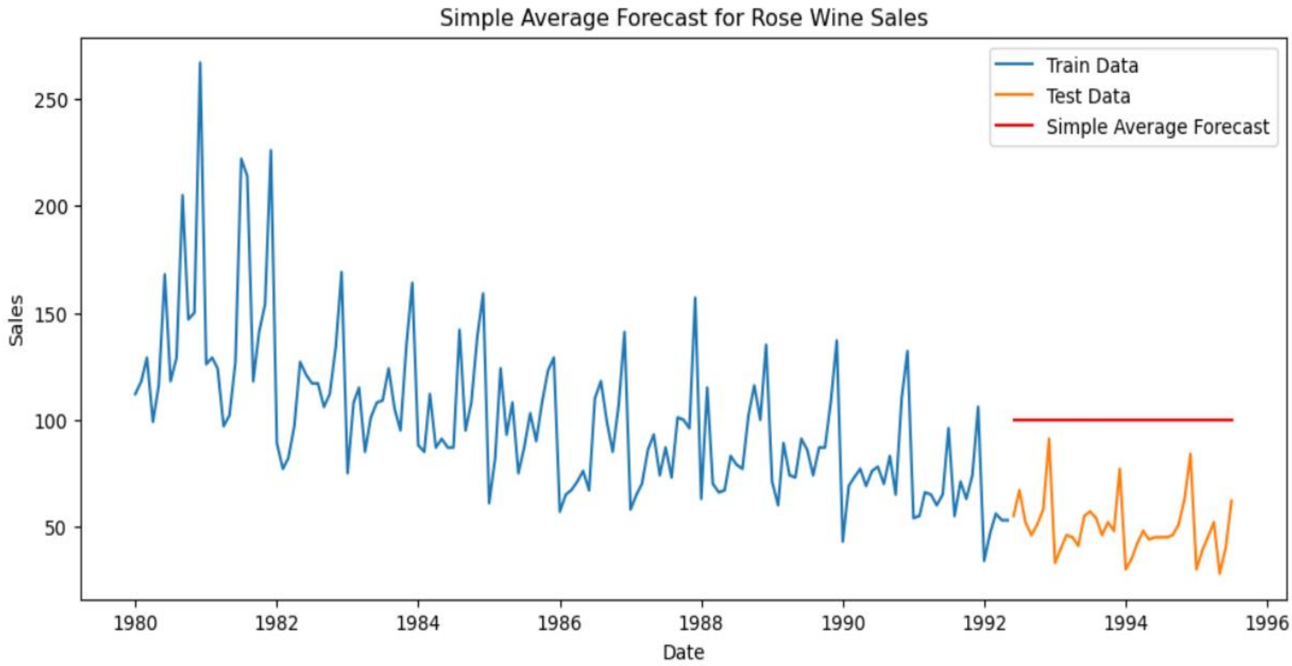


Fig 1.10.2 Simple Average Forecast for Rose wine sales

1.11 Moving Average:

Moving Average for Sparkling Sales:

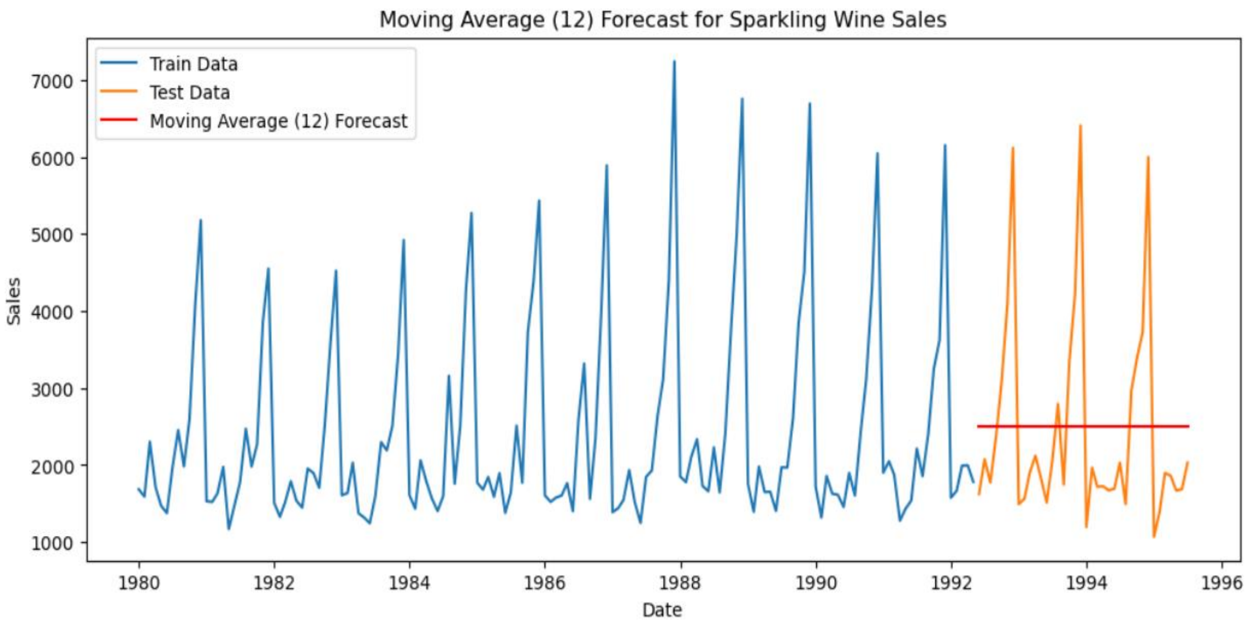


Fig 1.11.1 Moving Average Forecast for sparkling wine sales

Moving Average for Rose wine Sales:

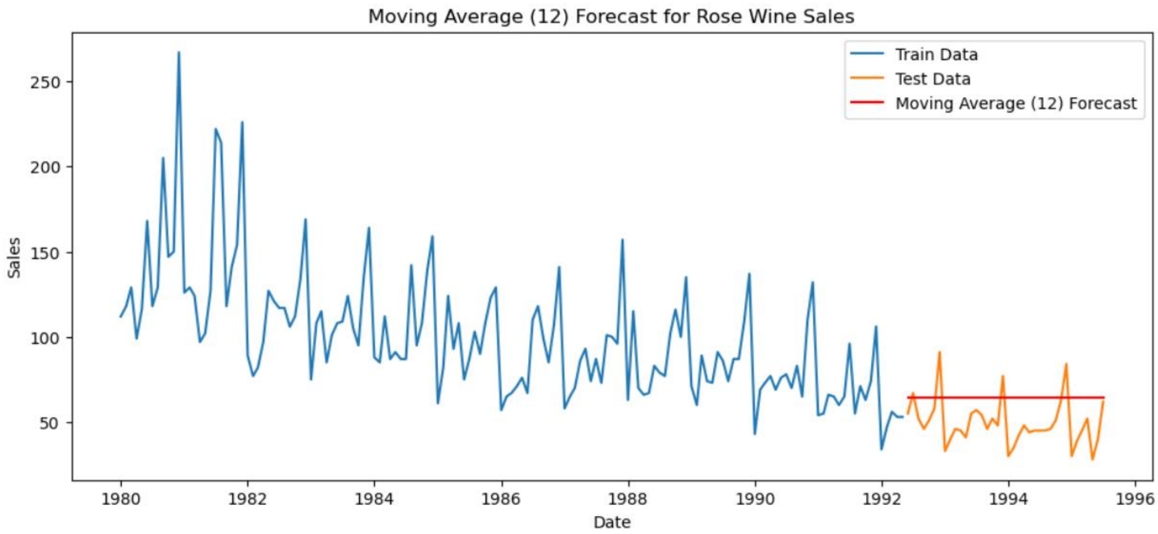


Fig 1.11.2 Moving Average Forecast for Rose wine sales

1.12 Exponential Models (Single, Double, Triple):

Single Exponential Smoothing (Simple Exponential Smoothing):

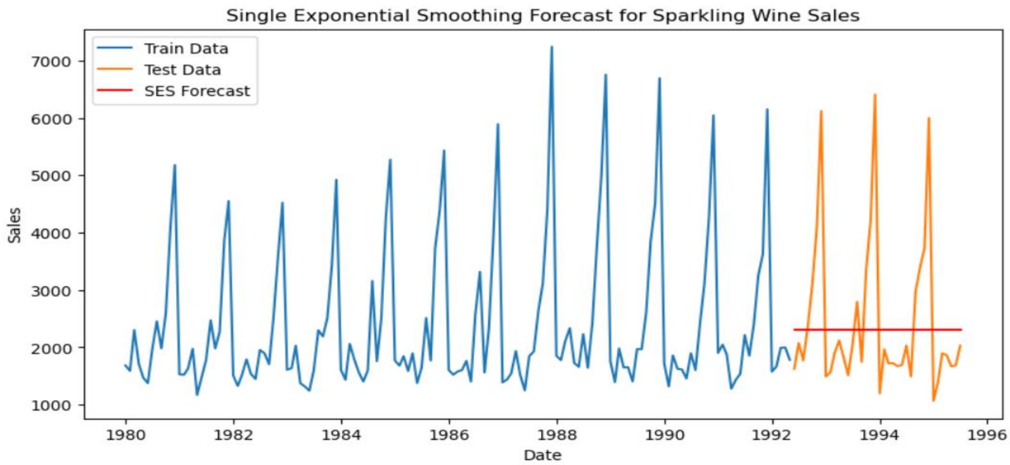


Fig 1.12.1 Single Exponential Smoothing Forecast for Sparkling wine sales

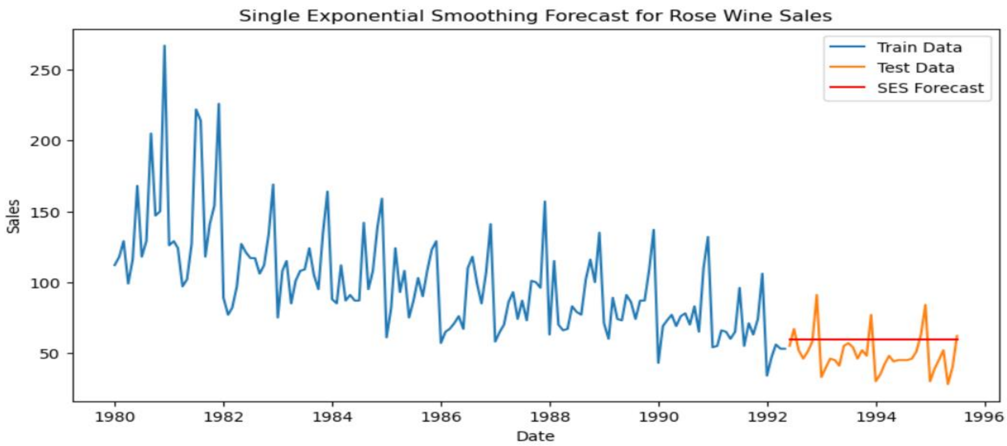


Fig 1.12.2 Single Exponential Smoothing Forecast for Rose wine sales

Double Exponential Smoothing (Holt's Linear Trend Model):

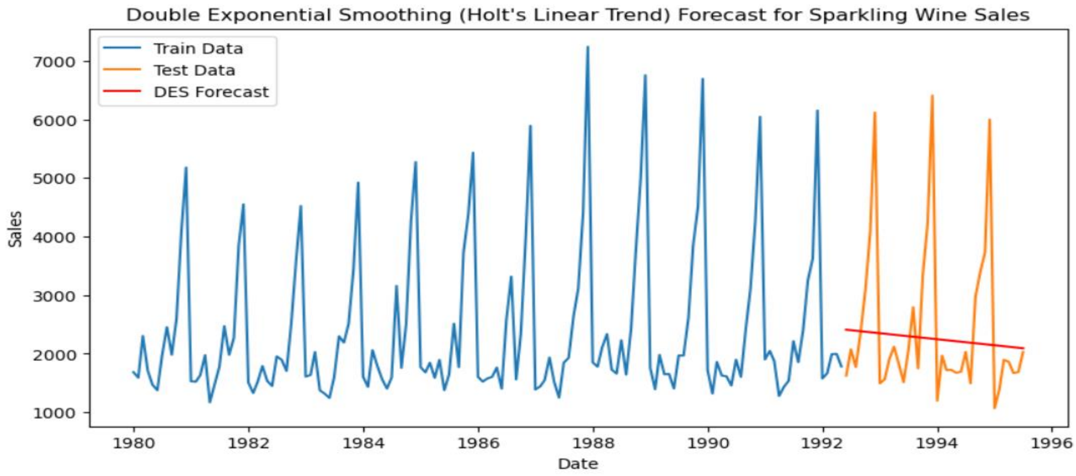


Fig 1.12.3 Double Exponential Smoothing Forecast for Sparkling wine sales

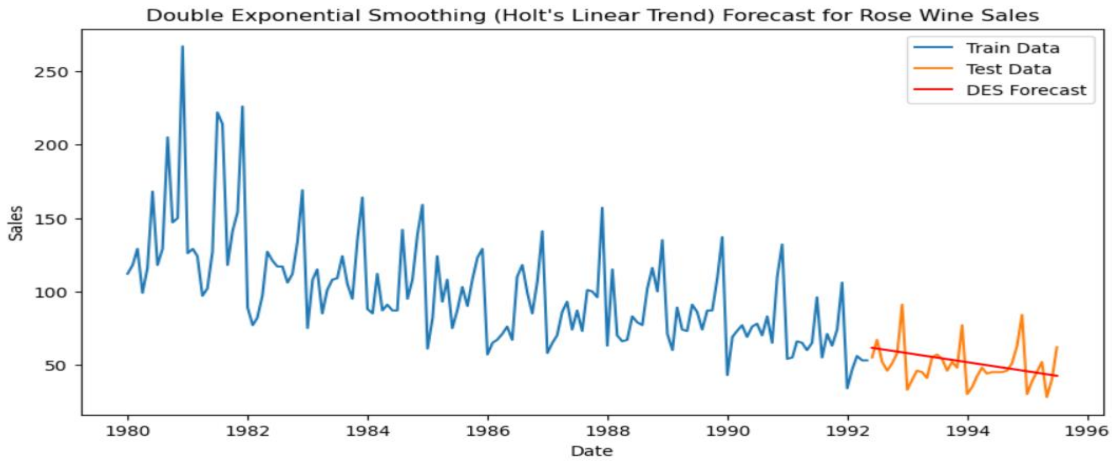


Fig 1.12.4 Double Exponential Smoothing Forecast for Rose wine sales

Triple Exponential Smoothing (Holt-Winters Model):

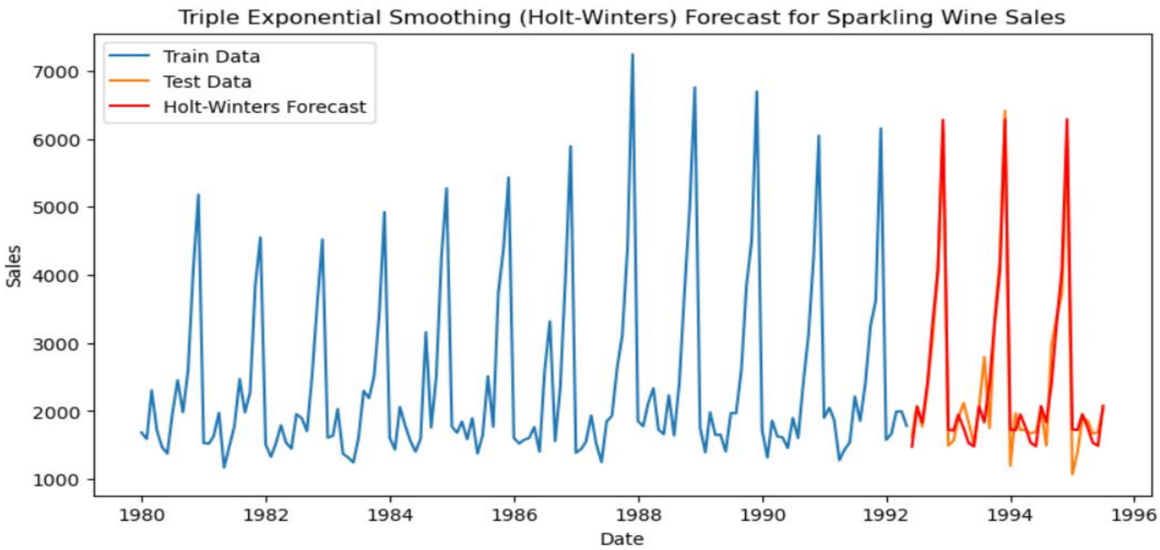


Fig 1.12.5 Triple Exponential Smoothing Forecast for Sparkling wine sales

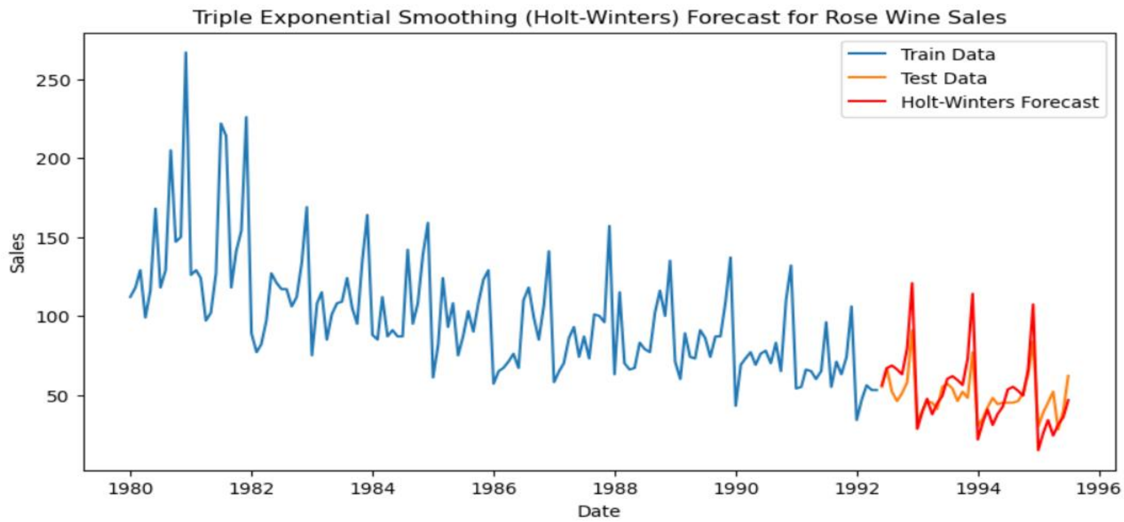


Fig 1.12.6 Triple Exponential Smoothing Forecast for Rose wine sales

1.13 Check the performance of the models built:

SARIMA Model Evaluation Metrics:

- MAE for Sparkling Wine: 146.99
- RMSE for Sparkling Wine: 202.29

Simple Average Forecasting Metrics:

- MAE for Sparkling Wine: 978.39
- RMSE for Sparkling Wine: 1331.04
- MAE for Rose Wine: 50.48
- RMSE for Rose Wine: 52.24

Exponential Smoothing Models Metrics:

- SES - MAE for Sparkling Wine: 951.22, RMSE: 1336.66
- DES - MAE for Sparkling Wine: 916.69, RMSE: 1340.45
- Holt-Winters - MAE for Sparkling Wine: 218.82, RMSE: 304.27
- SES - MAE for Rose Wine: 14.21, RMSE: 16.48
- DES - MAE for Rose Wine: 10.31, RMSE: 13.73
- Holt-Winters - MAE for Rose Wine: 10.39, RMSE: 13.85

Check for Stationarity:

1.14 Check for Stationarity:

Sparkling Data:

ADF Statistic: -1.3605
p-value: 0.6011
Critical Values:
1%: -3.4683
5%: -2.8782
10%: -2.5757
Fail to reject null hypothesis - **Data is non-stationary**

Rose Data:

ADF Statistic: -1.8749
p-value: 0.3440
Critical Values:
1%: -3.4687
5%: -2.8784
10%: -2.5758
Fail to reject null hypothesis - Data is non-stationary

Data is not stationary

1.15 Make the data stationary (if needed):

Sparkling data:

ADF Statistic: -45.050300936195256
p-value: 0.0
Critical Values:
1%: -3.4682803641749267
5%: -2.8782017240816327
10%: -2.5756525795918366
Reject null hypothesis - Data is stationary

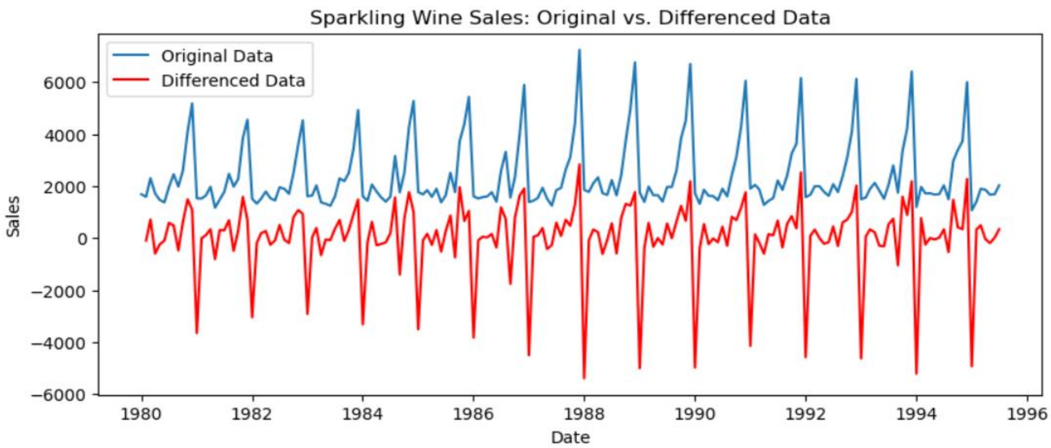


Fig 1.15.1 Original vs Differenced Data (Sparkling wines)

Rose data:
ADF Statistic: -8.04413902007531
p-value: 1.8135795068093227e-12
Critical Values:
    1%: -3.4687256239864017
    5%: -2.8783961376954363
    10%: -2.57575634100705
Reject null hypothesis - Data is stationary

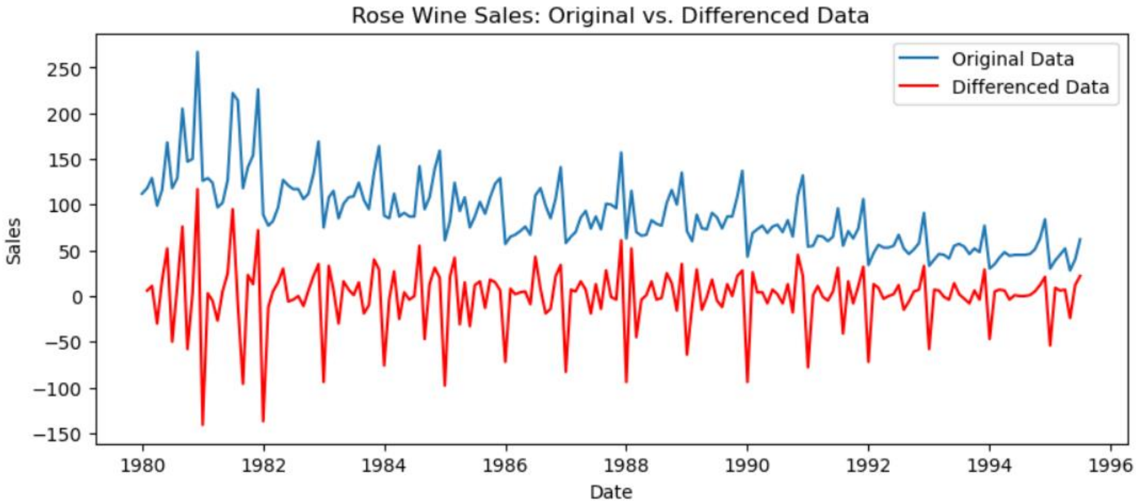


Fig 1.15.2 Original vs Differenced Data (Rose wines)

Model Building- Stationary Data:

1.16 Generate ACF & PACF Plot and find the AR, MA values:

For sparkling Data:

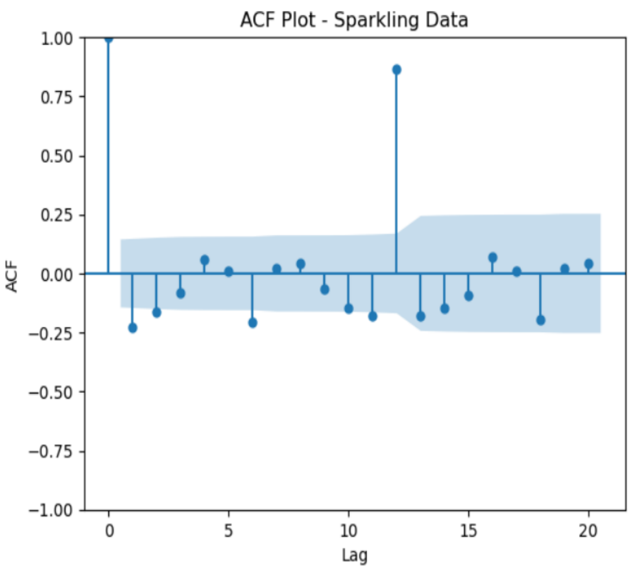


Fig 1.16.1 ACF plot (Sparkling Data)

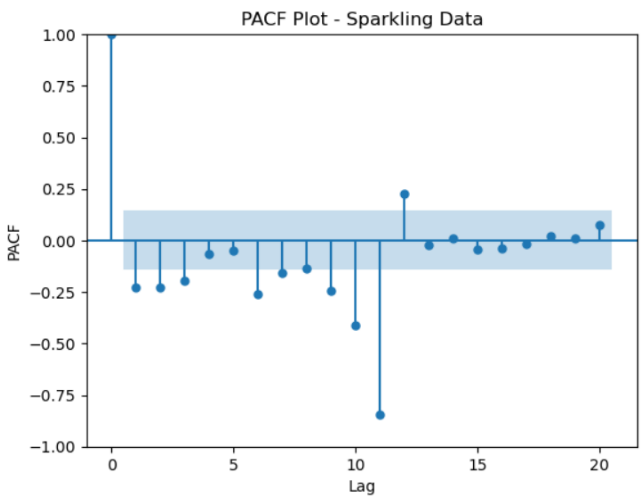
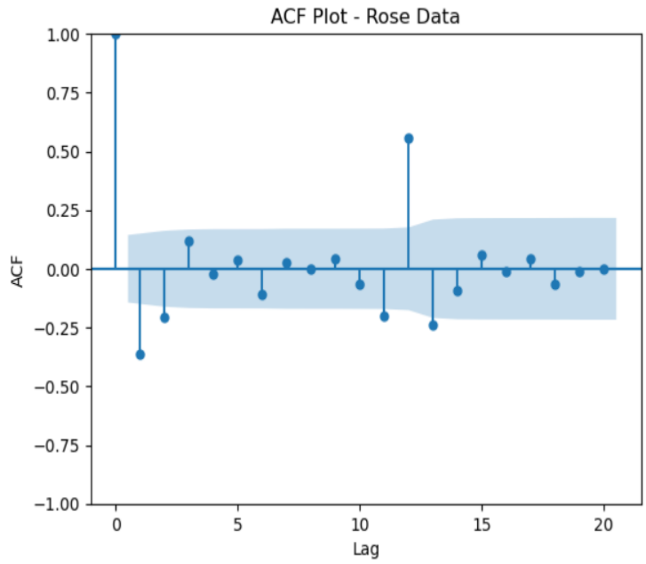


Fig 1.16.2 PACF plot (Sparkling Data)

For Rose Data:



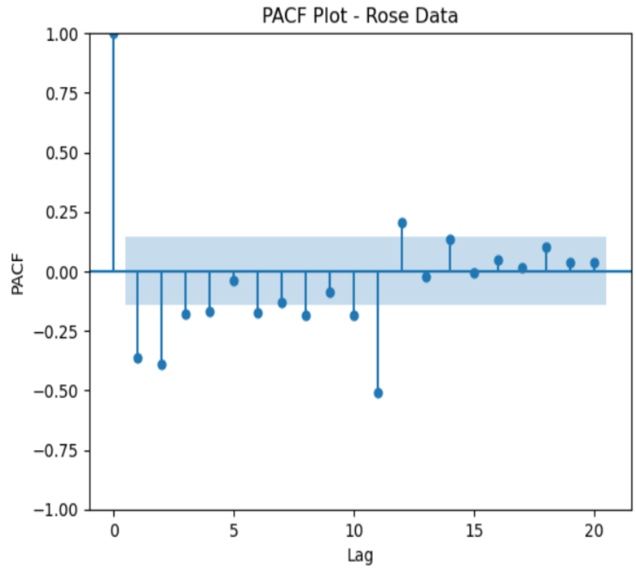


Fig 1.16.3 ACF plot (Rose Data)

Fig 1.16.4 PACF plot (Rose Data)

Analysis of Sparkling & Rose Data:

The identified AR and MA values for the sparkling wine data are as follows:

- Autoregressive (AR) Component: AR(1), AR(12)
- Moving Average (MA) Component: MA(1), MA(12)

The identified AR and MA values for the rose wine data are as follows:

- Autoregressive component AR (p) = 1
- Moving Average component MA (q) = 1

1.17 Build different ARIMA models:

For Sparkling wine:

- ✓ ARIMA (1,1,0) for Sparkling Wine - MAE: 1007.72, RMSE: 1618.78, AIC: 2640.81, BIC: 2646.80

For Rose wine:

- ✓ ARIMA (2,1,1) for Rose Wine - MAE: 12.40, RMSE: 18.82, AIC: 1457.22, BIC: 1469.19

1.18 Auto ARIMA :

Auto ARIMA for Sparkling wine:

```
Performing stepwise search to minimize aic
 ARIMA(2,0,2)(0,0,0)[0]             : AIC=inf, Time=0.31 sec
 ARIMA(0,0,0)(0,0,0)[0]             : AIC=3491.467, Time=0.02 sec
 ARIMA(1,0,0)(0,0,0)[0]             : AIC=3245.760, Time=0.02 sec
 ARIMA(0,0,1)(0,0,0)[0]             : AIC=3367.627, Time=0.06 sec
 ARIMA(2,0,0)(0,0,0)[0]             : AIC=3242.499, Time=0.03 sec
 ARIMA(3,0,0)(0,0,0)[0]             : AIC=3238.038, Time=0.05 sec
 ARIMA(4,0,0)(0,0,0)[0]             : AIC=3234.750, Time=0.07 sec
 ARIMA(5,0,0)(0,0,0)[0]             : AIC=3236.397, Time=0.07 sec
 ARIMA(4,0,1)(0,0,0)[0]             : AIC=inf, Time=0.35 sec
 ARIMA(3,0,1)(0,0,0)[0]             : AIC=inf, Time=0.31 sec
 ARIMA(5,0,1)(0,0,0)[0]             : AIC=3238.034, Time=0.14 sec
 ARIMA(4,0,0)(0,0,0)[0] intercept   : AIC=3188.523, Time=0.07 sec
 ARIMA(3,0,0)(0,0,0)[0] intercept   : AIC=3188.017, Time=0.07 sec
 ARIMA(2,0,0)(0,0,0)[0] intercept   : AIC=3187.681, Time=0.05 sec
 ARIMA(1,0,0)(0,0,0)[0] intercept   : AIC=3188.955, Time=0.07 sec
 ARIMA(2,0,1)(0,0,0)[0] intercept   : AIC=3190.109, Time=0.09 sec
 ARIMA(1,0,1)(0,0,0)[0] intercept   : AIC=3189.004, Time=0.05 sec
 ARIMA(3,0,1)(0,0,0)[0] intercept   : AIC=3189.969, Time=0.10 sec

Best model:  ARIMA(2,0,0)(0,0,0)[0] intercept
Total fit time: 1.924 seconds
```

Auto ARIMA for rose wine:

```
Performing stepwise search to minimize aic
 ARIMA(2,1,2)(0,0,0)[0] intercept   : AIC=inf, Time=0.35 sec
 ARIMA(0,1,0)(0,0,0)[0] intercept   : AIC=1854.380, Time=0.02 sec
 ARIMA(1,1,0)(0,0,0)[0] intercept   : AIC=1830.083, Time=0.04 sec
 ARIMA(0,1,1)(0,0,0)[0] intercept   : AIC=inf, Time=0.13 sec
 ARIMA(0,1,0)(0,0,0)[0]             : AIC=1852.391, Time=0.02 sec
 ARIMA(2,1,0)(0,0,0)[0] intercept   : AIC=1801.266, Time=0.06 sec
 ARIMA(3,1,0)(0,0,0)[0] intercept   : AIC=1797.428, Time=0.07 sec
 ARIMA(4,1,0)(0,0,0)[0] intercept   : AIC=1794.275, Time=0.09 sec
 ARIMA(5,1,0)(0,0,0)[0] intercept   : AIC=1795.931, Time=0.15 sec
 ARIMA(4,1,1)(0,0,0)[0] intercept   : AIC=inf, Time=0.16 sec
 ARIMA(3,1,1)(0,0,0)[0] intercept   : AIC=inf, Time=0.14 sec
 ARIMA(5,1,1)(0,0,0)[0] intercept   : AIC=inf, Time=0.39 sec
 ARIMA(4,1,0)(0,0,0)[0]             : AIC=1792.479, Time=0.06 sec
 ARIMA(3,1,0)(0,0,0)[0]             : AIC=1795.571, Time=0.06 sec
 ARIMA(5,1,0)(0,0,0)[0]             : AIC=1794.159, Time=0.08 sec
 ARIMA(4,1,1)(0,0,0)[0]             : AIC=1774.874, Time=0.14 sec
 ARIMA(3,1,1)(0,0,0)[0]             : AIC=1773.566, Time=0.09 sec
 ARIMA(2,1,1)(0,0,0)[0]             : AIC=1772.645, Time=0.06 sec
```

ARIMA(1,1,1)(0,0,0)[0]   : AIC=1772.725, Time=0.06 sec
ARIMA(2,1,0)(0,0,0)[0]   : AIC=1799.366, Time=0.04 sec
ARIMA(2,1,2)(0,0,0)[0]   : AIC=1772.664, Time=0.11 sec
ARIMA(1,1,0)(0,0,0)[0]   : AIC=1828.114, Time=0.04 sec
ARIMA(1,1,2)(0,0,0)[0]   : AIC=1770.734, Time=0.07 sec
ARIMA(0,1,2)(0,0,0)[0]   : AIC=1771.634, Time=0.05 sec
ARIMA(1,1,3)(0,0,0)[0]   : AIC=1772.669, Time=0.11 sec
ARIMA(0,1,1)(0,0,0)[0]   : AIC=1774.457, Time=0.03 sec
ARIMA(0,1,3)(0,0,0)[0]   : AIC=1771.456, Time=0.07 sec
ARIMA(2,1,3)(0,0,0)[0]   : AIC=inf, Time=0.29 sec
ARIMA(1,1,2)(0,0,0)[0] intercept   : AIC=inf, Time=0.19 sec

Best model: ARIMA(1,1,2)(0,0,0)[0]
Total fit time: 3.165 seconds

1.19 Manual ARIMA:

For Sparkling wine:

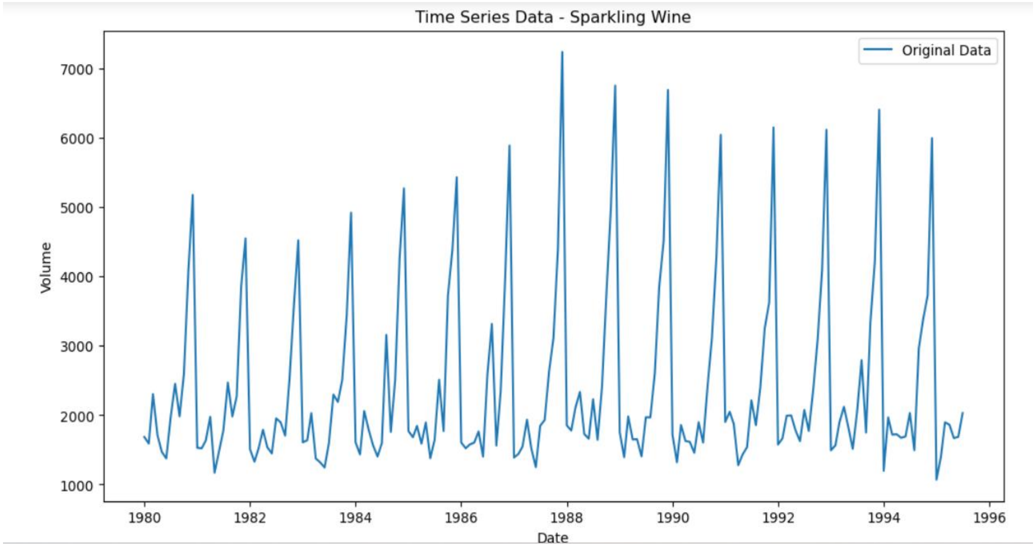


Fig 1.19.1 Time series data (sparkling wine)

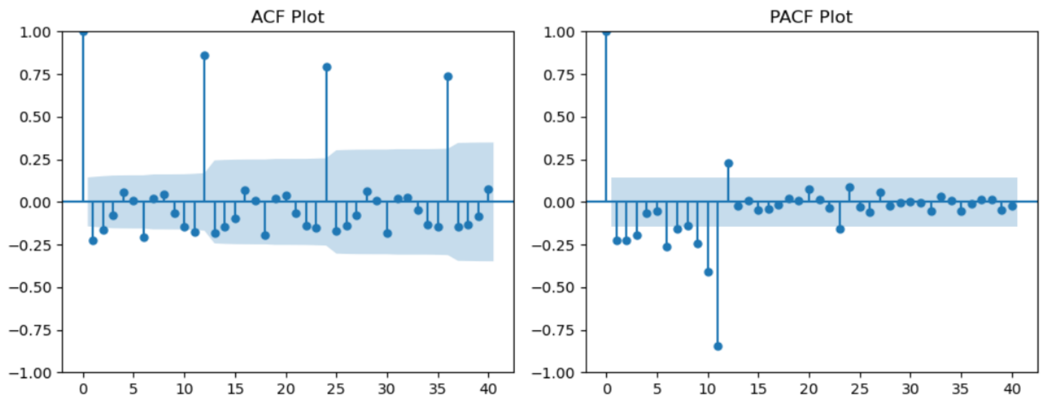


Fig 1.19.2 ACF & PACF plot (sparkling wine)

For rose-wine:

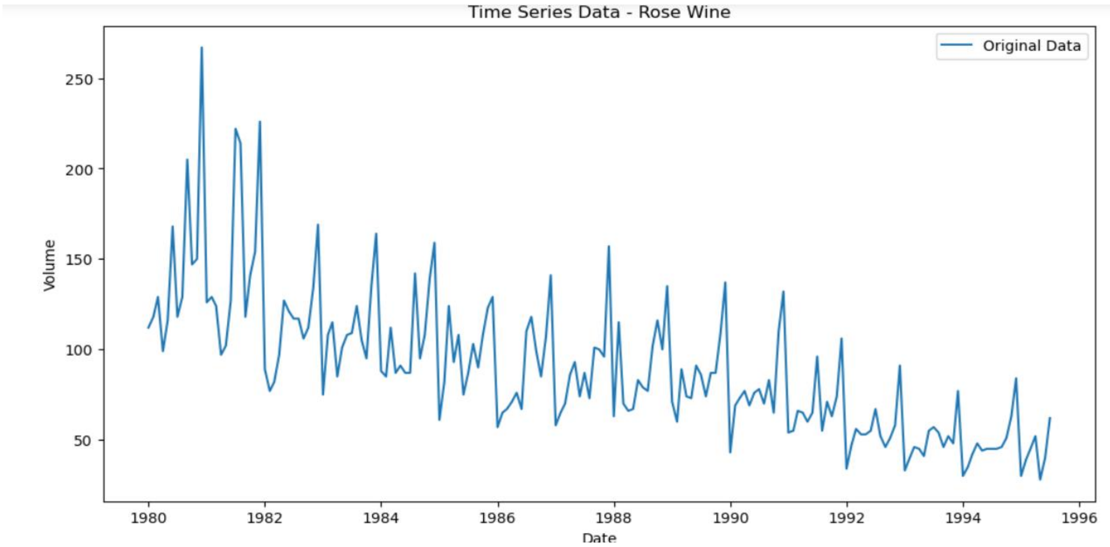


Fig 1.19.3 Time series data (Rose wine)

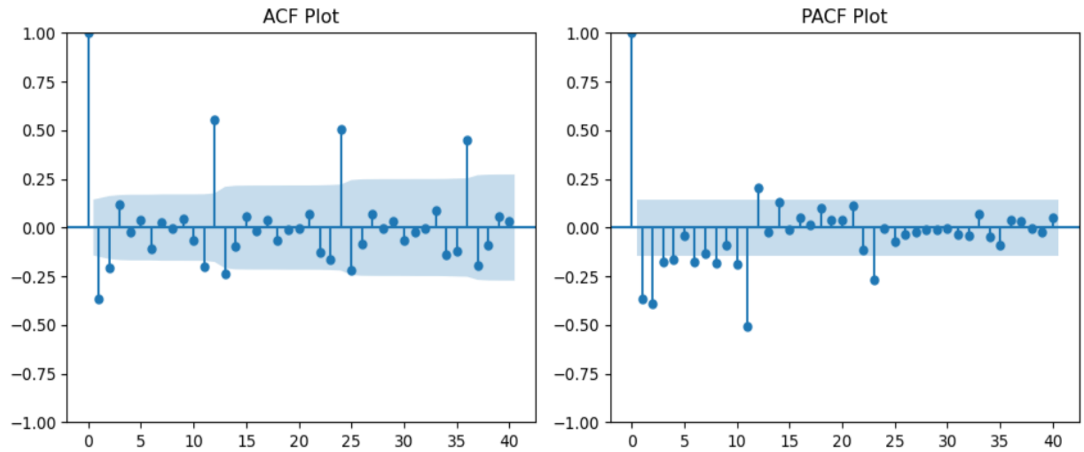


Fig 1.19.4 ACF & PACF plot (Rose wine)

1.20 Build different SARIMA models:

Performing stepwise search to minimize aic (Sparkling wine)

ARIMA(2,0,2)(1,1,1)[12] intercept : AIC=inf, Time=2.04 sec
ARIMA(0,0,0)(0,1,0)[12] intercept : AIC=2617.570, Time=0.02 sec
ARIMA(1,0,0)(1,1,0)[12] intercept : AIC=2589.608, Time=0.33 sec
ARIMA(0,0,1)(0,1,1)[12] intercept : AIC=2577.638, Time=0.60 sec
ARIMA(0,0,0)(0,1,0)[12] : AIC=2615.591, Time=0.02 sec
ARIMA(0,0,1)(0,1,0)[12] intercept : AIC=2617.620, Time=0.05 sec
ARIMA(0,0,1)(1,1,1)[12] intercept : AIC=2579.414, Time=0.97 sec
ARIMA(0,0,1)(0,1,2)[12] intercept : AIC=2579.433, Time=1.27 sec
ARIMA(0,0,1)(1,1,0)[12] intercept : AIC=2589.316, Time=0.31 sec
ARIMA(0,0,1)(1,1,2)[12] intercept : AIC=2581.251, Time=1.73 sec
ARIMA(0,0,0)(0,1,1)[12] intercept : AIC=2578.588, Time=0.50 sec
ARIMA(1,0,1)(0,1,1)[12] intercept : AIC=2579.450, Time=0.97 sec
ARIMA(0,0,2)(0,1,1)[12] intercept : AIC=2579.565, Time=1.36 sec
ARIMA(1,0,0)(0,1,1)[12] intercept : AIC=2577.766, Time=0.69 sec

```
ARIMA(1,0,2)(0,1,1)[12] intercept   : AIC=2581.588, Time=1.38 sec
 ARIMA(0,0,1)(0,1,1)[12]             : AIC=2577.640, Time=0.38 sec

Best model:  ARIMA(0,0,1)(0,1,1)[12] intercept
Total fit time: 12.650 seconds
```

For Rose wine:

```
Performing stepwise search to minimize aic
 ARIMA(2,1,2)(1,0,1)[12] intercept   : AIC=inf, Time=0.76 sec
 ARIMA(0,1,0)(0,0,0)[12] intercept   : AIC=1854.380, Time=0.02 sec
 ARIMA(1,1,0)(1,0,0)[12] intercept   : AIC=1750.322, Time=0.16 sec
 ARIMA(0,1,1)(0,0,1)[12] intercept   : AIC=inf, Time=0.33 sec
 ARIMA(0,1,0)(0,0,0)[12]             : AIC=1852.391, Time=0.02 sec
 ARIMA(1,1,0)(0,0,0)[12] intercept   : AIC=1830.083, Time=0.05 sec
 ARIMA(1,1,0)(2,0,0)[12] intercept   : AIC=1722.396, Time=0.44 sec
 ARIMA(1,1,0)(2,0,1)[12] intercept   : AIC=1712.318, Time=0.95 sec
 ARIMA(1,1,0)(1,0,1)[12] intercept   : AIC=1710.778, Time=0.40 sec
 ARIMA(1,1,0)(0,0,1)[12] intercept   : AIC=1784.865, Time=0.21 sec
 ARIMA(1,1,0)(1,0,2)[12] intercept   : AIC=1712.408, Time=1.28 sec
 ARIMA(1,1,0)(0,0,2)[12] intercept   : AIC=1759.892, Time=0.57 sec
 ARIMA(1,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=2.54 sec
 ARIMA(0,1,0)(1,0,1)[12] intercept   : AIC=1731.240, Time=0.23 sec
 ARIMA(2,1,0)(1,0,1)[12] intercept   : AIC=1677.371, Time=0.47 sec
 ARIMA(2,1,0)(0,0,1)[12] intercept   : AIC=1753.811, Time=0.30 sec
 ARIMA(2,1,0)(1,0,0)[12] intercept   : AIC=1716.463, Time=0.32 sec
 ARIMA(2,1,0)(2,0,1)[12] intercept   : AIC=1679.355, Time=0.90 sec
 ARIMA(2,1,0)(1,0,2)[12] intercept   : AIC=1679.358, Time=1.02 sec
 ARIMA(2,1,0)(0,0,0)[12] intercept   : AIC=1801.266, Time=0.07 sec
 ARIMA(2,1,0)(0,0,2)[12] intercept   : AIC=1724.409, Time=0.52 sec
 ARIMA(2,1,0)(2,0,0)[12] intercept   : AIC=1689.517, Time=0.71 sec
 ARIMA(2,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=1.99 sec
 ARIMA(3,1,0)(1,0,1)[12] intercept   : AIC=1674.004, Time=0.50 sec
 ARIMA(3,1,0)(0,0,1)[12] intercept   : AIC=1749.167, Time=0.30 sec
 ARIMA(3,1,0)(1,0,0)[12] intercept   : AIC=1710.772, Time=0.32 sec
 ARIMA(3,1,0)(2,0,1)[12] intercept   : AIC=1676.002, Time=1.31 sec
 ARIMA(3,1,0)(1,0,2)[12] intercept   : AIC=1676.002, Time=1.08 sec
 ARIMA(3,1,0)(0,0,0)[12] intercept   : AIC=1797.428, Time=0.08 sec
 ARIMA(3,1,0)(0,0,2)[12] intercept   : AIC=1721.560, Time=0.56 sec
 ARIMA(3,1,0)(2,0,0)[12] intercept   : AIC=1685.774, Time=1.09 sec
 ARIMA(3,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=2.86 sec
 ARIMA(4,1,0)(1,0,1)[12] intercept   : AIC=1663.976, Time=1.49 sec
 ARIMA(4,1,0)(0,0,1)[12] intercept   : AIC=1743.302, Time=0.39 sec
 ARIMA(4,1,0)(1,0,0)[12] intercept   : AIC=1701.675, Time=1.12 sec
 ARIMA(4,1,0)(2,0,1)[12] intercept   : AIC=1665.976, Time=1.62 sec
 ARIMA(4,1,0)(1,0,2)[12] intercept   : AIC=1665.976, Time=1.31 sec
 ARIMA(4,1,0)(0,0,0)[12] intercept   : AIC=1794.275, Time=0.10 sec
 ARIMA(4,1,0)(0,0,2)[12] intercept   : AIC=1715.993, Time=1.09 sec
```

```
ARIMA(4,1,0)(2,0,0)[12] intercept   : AIC=inf, Time=1.24 sec
 ARIMA(4,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=2.84 sec
 ARIMA(5,1,0)(1,0,1)[12] intercept   : AIC=1662.964, Time=1.45 sec
 ARIMA(5,1,0)(0,0,1)[12] intercept   : AIC=1743.776, Time=0.44 sec
 ARIMA(5,1,0)(1,0,0)[12] intercept   : AIC=1700.948, Time=1.10 sec
 ARIMA(5,1,0)(2,0,1)[12] intercept   : AIC=1664.960, Time=2.03 sec
 ARIMA(5,1,0)(1,0,2)[12] intercept   : AIC=1664.960, Time=1.92 sec
 ARIMA(5,1,0)(0,0,0)[12] intercept   : AIC=1795.931, Time=0.23 sec
 ARIMA(5,1,0)(0,0,2)[12] intercept   : AIC=1716.530, Time=1.21 sec
 ARIMA(5,1,0)(2,0,0)[12] intercept   : AIC=inf, Time=1.23 sec
 ARIMA(5,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=3.01 sec
 ARIMA(5,1,1)(1,0,1)[12] intercept   : AIC=1667.870, Time=1.39 sec
 ARIMA(4,1,1)(1,0,1)[12] intercept   : AIC=1651.866, Time=1.70 sec
 ARIMA(4,1,1)(0,0,1)[12] intercept   : AIC=inf, Time=0.73 sec
 ARIMA(4,1,1)(1,0,0)[12] intercept   : AIC=inf, Time=1.68 sec
 ARIMA(4,1,1)(2,0,1)[12] intercept   : AIC=1650.104, Time=4.03 sec
 ARIMA(4,1,1)(2,0,0)[12] intercept   : AIC=inf, Time=2.57 sec
 ARIMA(4,1,1)(2,0,2)[12] intercept   : AIC=inf, Time=2.45 sec
 ARIMA(4,1,1)(1,0,2)[12] intercept   : AIC=1651.275, Time=2.53 sec
 ARIMA(3,1,1)(2,0,1)[12] intercept   : AIC=1647.783, Time=2.40 sec
 ARIMA(3,1,1)(1,0,1)[12] intercept   : AIC=inf, Time=0.90 sec
 ARIMA(3,1,1)(2,0,0)[12] intercept   : AIC=inf, Time=2.76 sec
 ARIMA(3,1,1)(2,0,2)[12] intercept   : AIC=inf, Time=4.10 sec
 ARIMA(3,1,1)(1,0,0)[12] intercept   : AIC=inf, Time=0.72 sec
 ARIMA(3,1,1)(1,0,2)[12] intercept   : AIC=inf, Time=2.28 sec
 ARIMA(2,1,1)(2,0,1)[12] intercept   : AIC=inf, Time=1.68 sec
 ARIMA(3,1,2)(2,0,1)[12] intercept   : AIC=inf, Time=2.59 sec
 ARIMA(2,1,2)(2,0,1)[12] intercept   : AIC=inf, Time=2.33 sec
 ARIMA(4,1,2)(2,0,1)[12] intercept   : AIC=1652.057, Time=4.56 sec
 ARIMA(3,1,1)(2,0,1)[12]             : AIC=1647.320, Time=1.84 sec
 ARIMA(3,1,1)(1,0,1)[12]             : AIC=1645.326, Time=0.70 sec
 ARIMA(3,1,1)(0,0,1)[12]             : AIC=1722.438, Time=0.29 sec
 ARIMA(3,1,1)(1,0,0)[12]             : AIC=1681.050, Time=0.27 sec
 ARIMA(3,1,1)(1,0,2)[12]             : AIC=1647.324, Time=2.09 sec
 ARIMA(3,1,1)(0,0,0)[12]             : AIC=1773.566, Time=0.10 sec
 ARIMA(3,1,1)(0,0,2)[12]             : AIC=1696.419, Time=1.08 sec
 ARIMA(3,1,1)(2,0,0)[12]             : AIC=1657.538, Time=0.90 sec
 ARIMA(3,1,1)(2,0,2)[12]             : AIC=inf, Time=0.96 sec
 ARIMA(2,1,1)(1,0,1)[12]             : AIC=1644.489, Time=0.44 sec
 ARIMA(2,1,1)(0,0,1)[12]             : AIC=1721.353, Time=0.27 sec
 ARIMA(2,1,1)(1,0,0)[12]             : AIC=1680.148, Time=0.22 sec
 ARIMA(2,1,1)(2,0,1)[12]             : AIC=1646.489, Time=1.17 sec
 ARIMA(2,1,1)(1,0,2)[12]             : AIC=1646.489, Time=2.60 sec
 ARIMA(2,1,1)(0,0,0)[12]             : AIC=1772.645, Time=0.14 sec
 ARIMA(2,1,1)(0,0,2)[12]             : AIC=1696.788, Time=1.49 sec
 ARIMA(2,1,1)(2,0,0)[12]             : AIC=1656.984, Time=1.31 sec
```

```
ARIMA(2,1,1)(2,0,2)[12]            : AIC=inf, Time=2.14 sec
 ARIMA(1,1,1)(1,0,1)[12]            : AIC=1646.873, Time=0.39 sec
 ARIMA(2,1,0)(1,0,1)[12]            : AIC=1675.386, Time=0.36 sec
 ARIMA(2,1,2)(1,0,1)[12]            : AIC=1643.397, Time=0.54 sec
 ARIMA(2,1,2)(0,0,1)[12]            : AIC=1721.532, Time=0.33 sec
 ARIMA(2,1,2)(1,0,0)[12]            : AIC=1680.248, Time=0.31 sec
 ARIMA(2,1,2)(2,0,1)[12]            : AIC=1645.395, Time=1.60 sec
 ARIMA(2,1,2)(1,0,2)[12]            : AIC=1645.395, Time=1.75 sec
 ARIMA(2,1,2)(0,0,0)[12]            : AIC=1772.664, Time=0.13 sec
 ARIMA(2,1,2)(0,0,2)[12]            : AIC=1695.427, Time=1.13 sec
 ARIMA(2,1,2)(2,0,0)[12]            : AIC=1656.207, Time=1.06 sec
 ARIMA(2,1,2)(2,0,2)[12]            : AIC=inf, Time=3.34 sec
 ARIMA(1,1,2)(1,0,1)[12]            : AIC=1641.720, Time=0.46 sec
 ARIMA(1,1,2)(0,0,1)[12]            : AIC=1719.695, Time=0.28 sec
 ARIMA(1,1,2)(1,0,0)[12]            : AIC=1678.514, Time=0.28 sec

Best model:  ARIMA(1,1,2)(1,0,1)[12]
Total fit time: 117.571 seconds
```

1.21 Auto SARIMA:

For sparkling data:

```
Performing stepwise search to minimize aic
 ARIMA(2,0,2)(1,1,1)[12] intercept   : AIC=inf, Time=2.75 sec
 ARIMA(0,0,0)(0,1,0)[12] intercept   : AIC=2617.570, Time=0.02 sec
 ARIMA(1,0,0)(1,1,0)[12] intercept   : AIC=2589.608, Time=0.42 sec
 ARIMA(0,0,1)(0,1,1)[12] intercept   : AIC=2577.638, Time=0.63 sec
 ARIMA(0,0,0)(0,1,0)[12]             : AIC=2615.591, Time=0.02 sec
 ARIMA(0,0,1)(0,1,0)[12] intercept   : AIC=2617.620, Time=0.05 sec
 ARIMA(0,0,1)(1,1,1)[12] intercept   : AIC=2579.414, Time=1.00 sec
 ARIMA(0,0,1)(0,1,2)[12] intercept   : AIC=2579.433, Time=1.32 sec
 ARIMA(0,0,1)(1,1,0)[12] intercept   : AIC=2589.316, Time=0.30 sec
 ARIMA(0,0,1)(1,1,2)[12] intercept   : AIC=2581.251, Time=1.67 sec
 ARIMA(0,0,0)(0,1,1)[12] intercept   : AIC=2578.588, Time=0.45 sec
 ARIMA(1,0,1)(0,1,1)[12] intercept   : AIC=2579.450, Time=0.68 sec
 ARIMA(0,0,2)(0,1,1)[12] intercept   : AIC=2579.565, Time=0.95 sec
 ARIMA(1,0,0)(0,1,1)[12] intercept   : AIC=2577.766, Time=0.51 sec
 ARIMA(1,0,2)(0,1,1)[12] intercept   : AIC=2581.588, Time=1.18 sec
 ARIMA(0,0,1)(0,1,1)[12]             : AIC=2577.640, Time=0.39 sec

Best model:  ARIMA(0,0,1)(0,1,1)[12] intercept
Total fit time: 12.396 seconds
```

For Rose Data:

```
Performing stepwise search to minimize aic
 ARIMA(2,1,2)(1,0,1)[12] intercept   : AIC=inf, Time=0.86 sec
 ARIMA(0,1,0)(0,0,0)[12] intercept   : AIC=1854.380, Time=0.02 sec
 ARIMA(1,1,0)(1,0,0)[12] intercept   : AIC=1750.322, Time=0.23 sec
 ARIMA(0,1,1)(0,0,1)[12] intercept   : AIC=inf, Time=0.40 sec
 ARIMA(0,1,0)(0,0,0)[12]             : AIC=1852.391, Time=0.02 sec
 ARIMA(1,1,0)(0,0,0)[12] intercept   : AIC=1830.083, Time=0.05 sec
 ARIMA(1,1,0)(2,0,0)[12] intercept   : AIC=1722.396, Time=0.51 sec
 ARIMA(1,1,0)(2,0,1)[12] intercept   : AIC=1712.318, Time=1.14 sec
 ARIMA(1,1,0)(1,0,1)[12] intercept   : AIC=1710.778, Time=0.39 sec
 ARIMA(1,1,0)(0,0,1)[12] intercept   : AIC=1784.865, Time=0.21 sec
 ARIMA(1,1,0)(1,0,2)[12] intercept   : AIC=1712.408, Time=1.10 sec
 ARIMA(1,1,0)(0,0,2)[12] intercept   : AIC=1759.892, Time=0.66 sec
 ARIMA(1,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=4.29 sec
 ARIMA(0,1,0)(1,0,1)[12] intercept   : AIC=1731.240, Time=0.33 sec
 ARIMA(2,1,0)(1,0,1)[12] intercept   : AIC=1677.371, Time=0.72 sec
 ARIMA(2,1,0)(0,0,1)[12] intercept   : AIC=1753.811, Time=0.50 sec
 ARIMA(2,1,0)(1,0,0)[12] intercept   : AIC=1716.463, Time=0.58 sec
 ARIMA(2,1,0)(2,0,1)[12] intercept   : AIC=1679.355, Time=1.27 sec
 ARIMA(2,1,0)(1,0,2)[12] intercept   : AIC=1679.358, Time=1.27 sec
 ARIMA(2,1,0)(0,0,0)[12] intercept   : AIC=1801.266, Time=0.07 sec
 ARIMA(2,1,0)(0,0,2)[12] intercept   : AIC=1724.409, Time=0.63 sec
 ARIMA(2,1,0)(2,0,0)[12] intercept   : AIC=1689.517, Time=0.95 sec
 ARIMA(2,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=2.63 sec
 ARIMA(3,1,0)(1,0,1)[12] intercept   : AIC=1674.004, Time=0.60 sec
 ARIMA(3,1,0)(0,0,1)[12] intercept   : AIC=1749.167, Time=0.32 sec
 ARIMA(3,1,0)(1,0,0)[12] intercept   : AIC=1710.772, Time=0.33 sec
 ARIMA(3,1,0)(2,0,1)[12] intercept   : AIC=1676.002, Time=2.35 sec
 ARIMA(3,1,0)(1,0,2)[12] intercept   : AIC=1676.002, Time=2.25 sec
 ARIMA(3,1,0)(0,0,0)[12] intercept   : AIC=1797.428, Time=0.15 sec
 ARIMA(3,1,0)(0,0,2)[12] intercept   : AIC=1721.560, Time=1.40 sec
 ARIMA(3,1,0)(2,0,0)[12] intercept   : AIC=1685.774, Time=1.79 sec
 ARIMA(3,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=2.48 sec
 ARIMA(4,1,0)(1,0,1)[12] intercept   : AIC=1663.976, Time=1.42 sec
 ARIMA(4,1,0)(0,0,1)[12] intercept   : AIC=1743.302, Time=0.35 sec
 ARIMA(4,1,0)(1,0,0)[12] intercept   : AIC=1701.675, Time=1.06 sec
 ARIMA(4,1,0)(2,0,1)[12] intercept   : AIC=1665.976, Time=2.18 sec
 ARIMA(4,1,0)(1,0,2)[12] intercept   : AIC=1665.976, Time=1.77 sec
 ARIMA(4,1,0)(0,0,0)[12] intercept   : AIC=1794.275, Time=0.14 sec
 ARIMA(4,1,0)(0,0,2)[12] intercept   : AIC=1715.993, Time=2.80 sec
 ARIMA(4,1,0)(2,0,0)[12] intercept   : AIC=inf, Time=2.27 sec
 ARIMA(4,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=4.35 sec
 ARIMA(5,1,0)(1,0,1)[12] intercept   : AIC=1662.964, Time=1.56 sec
 ARIMA(5,1,0)(0,0,1)[12] intercept   : AIC=1743.776, Time=0.43 sec
 ARIMA(5,1,0)(1,0,0)[12] intercept   : AIC=1700.948, Time=1.02 sec
```

```
ARIMA(5,1,0)(2,0,1)[12] intercept   : AIC=1664.960, Time=2.10 sec
 ARIMA(5,1,0)(1,0,2)[12] intercept   : AIC=1664.960, Time=3.29 sec
 ARIMA(5,1,0)(0,0,0)[12] intercept   : AIC=1795.931, Time=0.33 sec
 ARIMA(5,1,0)(0,0,2)[12] intercept   : AIC=1716.530, Time=2.09 sec
 ARIMA(5,1,0)(2,0,0)[12] intercept   : AIC=inf, Time=2.59 sec
 ARIMA(5,1,0)(2,0,2)[12] intercept   : AIC=inf, Time=3.66 sec
 ARIMA(5,1,1)(1,0,1)[12] intercept   : AIC=1667.870, Time=1.54 sec
 ARIMA(4,1,1)(1,0,1)[12] intercept   : AIC=1651.866, Time=1.78 sec
 ARIMA(4,1,1)(0,0,1)[12] intercept   : AIC=inf, Time=0.73 sec
 ARIMA(4,1,1)(1,0,0)[12] intercept   : AIC=inf, Time=1.41 sec
 ARIMA(4,1,1)(2,0,1)[12] intercept   : AIC=1650.104, Time=3.93 sec
 ARIMA(4,1,1)(2,0,0)[12] intercept   : AIC=inf, Time=3.18 sec
 ARIMA(4,1,1)(2,0,2)[12] intercept   : AIC=inf, Time=2.34 sec
 ARIMA(4,1,1)(1,0,2)[12] intercept   : AIC=1651.275, Time=2.32 sec
 ARIMA(3,1,1)(2,0,1)[12] intercept   : AIC=1647.783, Time=2.30 sec
 ARIMA(3,1,1)(1,0,1)[12] intercept   : AIC=inf, Time=0.82 sec
 ARIMA(3,1,1)(2,0,0)[12] intercept   : AIC=inf, Time=2.12 sec
 ARIMA(3,1,1)(2,0,2)[12] intercept   : AIC=inf, Time=4.17 sec
 ARIMA(3,1,1)(1,0,0)[12] intercept   : AIC=inf, Time=0.69 sec
 ARIMA(3,1,1)(1,0,2)[12] intercept   : AIC=inf, Time=2.34 sec
 ARIMA(2,1,1)(2,0,1)[12] intercept   : AIC=inf, Time=1.55 sec
 ARIMA(3,1,2)(2,0,1)[12] intercept   : AIC=inf, Time=2.44 sec
 ARIMA(2,1,2)(2,0,1)[12] intercept   : AIC=inf, Time=2.02 sec
 ARIMA(4,1,2)(2,0,1)[12] intercept   : AIC=1652.057, Time=2.90 sec
 ARIMA(3,1,1)(2,0,1)[12]             : AIC=1647.320, Time=1.93 sec
 ARIMA(3,1,1)(1,0,1)[12]             : AIC=1645.326, Time=0.80 sec
 ARIMA(3,1,1)(0,0,1)[12]             : AIC=1722.438, Time=0.36 sec
 ARIMA(3,1,1)(1,0,0)[12]             : AIC=1681.050, Time=0.34 sec
 ARIMA(3,1,1)(1,0,2)[12]             : AIC=1647.324, Time=2.40 sec
 ARIMA(3,1,1)(0,0,0)[12]             : AIC=1773.566, Time=0.09 sec
 ARIMA(3,1,1)(0,0,2)[12]             : AIC=1696.419, Time=0.99 sec
 ARIMA(3,1,1)(2,0,0)[12]             : AIC=1657.538, Time=0.83 sec
 ARIMA(3,1,1)(2,0,2)[12]             : AIC=inf, Time=0.77 sec
 ARIMA(2,1,1)(1,0,1)[12]             : AIC=1644.489, Time=0.43 sec
 ARIMA(2,1,1)(0,0,1)[12]             : AIC=1721.353, Time=0.25 sec
 ARIMA(2,1,1)(1,0,0)[12]             : AIC=1680.148, Time=0.26 sec
 ARIMA(2,1,1)(2,0,1)[12]             : AIC=1646.489, Time=1.18 sec
 ARIMA(2,1,1)(1,0,2)[12]             : AIC=1646.489, Time=1.22 sec
 ARIMA(2,1,1)(0,0,0)[12]             : AIC=1772.645, Time=0.06 sec
 ARIMA(2,1,1)(0,0,2)[12]             : AIC=1696.788, Time=0.70 sec
 ARIMA(2,1,1)(2,0,0)[12]             : AIC=1656.984, Time=0.69 sec
 ARIMA(2,1,1)(2,0,2)[12]             : AIC=inf, Time=1.16 sec
 ARIMA(1,1,1)(1,0,1)[12]             : AIC=1646.873, Time=0.36 sec
 ARIMA(2,1,0)(1,0,1)[12]             : AIC=1675.386, Time=0.38 sec
 ARIMA(2,1,2)(1,0,1)[12]             : AIC=1643.397, Time=0.60 sec
 ARIMA(2,1,2)(0,0,1)[12]             : AIC=1721.532, Time=0.41 sec
```

ARIMA(2,1,2)(1,0,0)[12] : AIC=1680.248, Time=0.34 sec
ARIMA(2,1,2)(2,0,1)[12] : AIC=1645.395, Time=1.67 sec
ARIMA(2,1,2)(1,0,2)[12] : AIC=1645.395, Time=1.68 sec
ARIMA(2,1,2)(0,0,0)[12] : AIC=1772.664, Time=0.13 sec
ARIMA(2,1,2)(0,0,2)[12] : AIC=1695.427, Time=1.15 sec
ARIMA(2,1,2)(2,0,0)[12] : AIC=1656.207, Time=0.90 sec
ARIMA(2,1,2)(2,0,2)[12] : AIC=inf, Time=2.06 sec
ARIMA(1,1,2)(1,0,1)[12] : AIC=1641.720, Time=0.33 sec
ARIMA(1,1,2)(0,0,1)[12] : AIC=1719.695, Time=0.20 sec
ARIMA(1,1,2)(1,0,0)[12] : AIC=1678.514, Time=0.21 sec

Best model: ARIMA(1,1,2)(1,0,1)[12]
Total fit time: 127.449 seconds

1.22 Manual SARIMA:

For sparkling data:

```
                                      SARIMAX Results
============================================================================================
Dep. Variable:                          Sparkling   No. Observations:                    187
Model:             SARIMAX(1, 1, 2)x(0, 1, [1], 12)   Log Likelihood               -1283.189
Date:                            Thu, 18 Jul 2024   AIC                           2576.377
Time:                                    21:36:39   BIC                           2592.173
Sample:                                01-01-1980   HQIC                          2582.785
                                     - 07-01-1995
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.9416      0.170      5.553      0.000       0.609       1.274
ma.L1         -1.8773      0.376     -4.991      0.000      -2.614      -1.140
ma.L2          0.8776      0.346      2.536      0.011       0.199       1.556
ma.S.L12      -0.5444      0.051    -10.575      0.000      -0.645      -0.443
sigma2      1.381e+05   3.06e+04      4.508      0.000    7.81e+04    1.98e+05
===================================================================================
Ljung-Box (L1) (Q):                   1.17   Jarque-Bera (JB):                61.39
Prob(Q):                              0.28   Prob(JB):                         0.00
Heteroskedasticity (H):               1.10   Skew:                             0.72
Prob(H) (two-sided):                  0.72   Kurtosis:                         5.53
===================================================================================
```

Fig 1.22.1 Manual SARIMA (Sparkling data)

For Rose Data:

```
                                      SARIMAX Results
==========================================================================================
Dep. Variable:                                 Rose   No. Observations:                  187
Model:             SARIMAX(1, 1, 2)x(0, 1, [1], 12)   Log Likelihood                -760.476
Date:                              Thu, 18 Jul 2024   AIC                           1530.952
Time:                                      21:37:11   BIC                           1546.747
Sample:                                  01-01-1980   HQIC                          1537.359
                                       - 07-01-1995
Covariance Type:                                opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.5253      0.176     -2.980      0.003      -0.871      -0.180
ma.L1         -0.1727      0.158     -1.091      0.275      -0.483       0.138
ma.L2         -0.6823      0.132     -5.163      0.000      -0.941      -0.423
ma.S.L12      -0.7471      0.086     -8.641      0.000      -0.917      -0.578
sigma2       341.5629     30.714     11.121      0.000     281.365     401.760
===================================================================================
Ljung-Box (L1) (Q):                   0.11   Jarque-Bera (JB):               132.44
Prob(Q):                              0.74   Prob(JB):                         0.00
Heteroskedasticity (H):               0.17   Skew:                             0.23
Prob(H) (two-sided):                  0.00   Kurtosis:                         7.25
===================================================================================
```

Fig 1.22.2 Manual SARIMA (Rose data)

1.23 Check the performance of the models built:

ARIMA MODEL:

- ARIMA(1,1,2) for Sparkling Wine - MAE: 1007.72, RMSE: 1618.78
- ARIMA(1,1,2) for Rose Wine - MAE: 12.40, RMSE: 18.82
- Auto ARIMA for Sparkling Wine - MAE: 2612.33, RMSE: 3182.00
- Auto ARIMA for Rose Wine - MAE: 36.54, RMSE: 43.71
- Manual ARIMA for Sparkling Wine - MAE: 2384.36, RMSE: 2877.89
- Manual ARIMA for Rose Wine - MAE: 47.93, RMSE: 51.49

SARIMA MODEL:

- SARIMA Model for Sparkling Wine - MAE: 149.63, RMSE: 208.79
- SARIMA Model for Rose Wine - MAE: 7.24, RMSE: 8.62
- Auto SARIMA Model for Sparkling Wine - MAE: 172.71, RMSE: 232.32
- Auto SARIMA Model for Rose Wine - MAE: 7.24, RMSE: 8.62
- Manual SARIMA Model for Sparkling Wine - MAE: 155.43, RMSE: 215.55
- Manual SARIMA Model for Rose Wine - MAE: 7.51, RMSE: 9.11

Compare the performance of the models

1.24 Compare the performance of the models:

Analysis:

- ARIMA vs SARIMA: Comparing ARIMA and SARIMA models, we observe that SARIMA models generally perform better in terms of both MAE and RMSE for both Sparkling and Rose wines. SARIMA models take into account seasonality, which might be critical for these types of time series data (assuming monthly seasonality based on m=12).
- Auto ARIMA vs Manual ARIMA: Auto ARIMA and Manual ARIMA results vary significantly. Auto ARIMA automatically selects model parameters based on AIC and BIC, but in this case, manual tuning (with expert knowledge or domain-specific insights) seems to have resulted in better performance for both wine types.
- Performance Metrics: MAE and RMSE give us an idea of the average error and the typical magnitude of error, respectively. Lower values indicate better performance in terms of forecasting accuracy.

Conclusion:

Based on the comparison:

- SARIMA models generally outperform ARIMA models for both wines.
- Manual tuning of ARIMA and SARIMA models seems to have yielded better results compared to Auto ARIMA, suggesting that domain knowledge and manual adjustments can improve forecasting accuracy.

1.25 Choose the best model with proper rationale:

Rationale for Choosing the Best Models: ***Sparkling Wine:***

- ARIMA(1,1,2): Despite having a higher MAE and RMSE compared to the SARIMA and manual SARIMA models, it provides a reasonable balance between simplicity and accuracy. If computational efficiency or simplicity is a priority, ARIMA(1,1,2) could be preferred.

SARIMA Model:

- Provides slightly better accuracy metrics (lower MAE and RMSE) compared to ARIMA(1,1,2), indicating better performance considering seasonality (m=12).

Manual SARIMA Model:

- Offers competitive performance with slightly higher MAE but lower RMSE compared to the SARIMA model. Manual tuning may have captured specific nuances in the data not fully captured by Auto SARIMA.

Conclusion for Sparkling Wine:

- Based on the provided metrics, the SARIMA Model appears to be the best choice due to its lower MAE and RMSE, indicating better forecasting accuracy considering the seasonality present in the data.

Rose Wine:

ARIMA(1,1,2):

- Shows lower MAE and RMSE compared to all SARIMA models, suggesting it captures the underlying patterns in the data well without needing to account for seasonality explicitly.

SARIMA Model and Auto SARIMA Model:

- Both exhibit very similar MAE and RMSE, indicating comparable performance. However, ARIMA(1,1,2) outperforms both in terms of both metrics.

Manual SARIMA Model:

- Shows slightly higher MAE and RMSE compared to SARIMA and Auto SARIMA models, suggesting that manual adjustments may not have significantly improved performance for this specific dataset.

Conclusion for Rose Wine:

- ARIMA(1,1,2) stands out as the best model due to its lowest MAE and RMSE, indicating superior forecasting accuracy compared to the SARIMA and Auto SARIMA models.

Final Recommendation:

- Sparkling Wine: Choose the SARIMA Model (based on lower MAE and RMSE).
- Rose Wine: Choose ARIMA(1,1,2) (based on lowest MAE and RMSE).

1.26 Rebuild the best model using the entire data:

Rebuilding SARIMA Model for Sparkling Wine:

```
                                     SARIMAX Results                                      
==========================================================================================
Dep. Variable:                          Sparkling   No. Observations:                  187
Model:             SARIMAX(1, 0, 2)x(2, 0, [], 12)   Log Likelihood               -1197.666
Date:                            Thu, 18 Jul 2024   AIC                           2407.332
Time:                                    22:14:12   BIC                           2425.857
Sample:                                01-01-1980   HQIC                          2414.853
                                     - 07-01-1995                                         
Covariance Type:                              opg                                         
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.5948      0.499     -1.192      0.233      -1.573       0.383
ma.L1          1.4147      0.523      2.706      0.007       0.390       2.439
ma.L2          0.0550      0.255      0.216      0.829      -0.444       0.554
ar.S.L12       0.5913      0.067      8.774      0.000       0.459       0.723
ar.S.L24       0.4129      0.070      5.873      0.000       0.275       0.551
sigma2      7.954e+04   7.42e+04      1.071      0.284    -6.6e+04    2.25e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.04   Jarque-Bera (JB):                24.47
Prob(Q):                              0.85   Prob(JB):                         0.00
Heteroskedasticity (H):               0.99   Skew:                             0.59
Prob(H) (two-sided):                  0.97   Kurtosis:                         4.50
===================================================================================
```

Fig 1.26.1 Best Model (Sparkling wine)

Rebuilding ARIMA (1,1,2) Model for Rose Wine:

```
                               SARIMAX Results                                
==============================================================================
Dep. Variable:                   Rose   No. Observations:                  187
Model:                 ARIMA(1, 1, 2)   Log Likelihood                -881.367
Date:                Thu, 18 Jul 2024   AIC                           1770.734
Time:                        22:15:28   BIC                           1783.637
Sample:                    01-01-1980   HQIC                          1775.962
                         - 07-01-1995                                         
Covariance Type:                  opg                                         
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.4879      0.221     -2.210      0.027      -0.921      -0.055
ma.L1         -0.2261      0.206     -1.098      0.272      -0.630       0.177
ma.L2         -0.6007      0.165     -3.647      0.000      -0.924      -0.278
sigma2       758.3960     54.668     13.873      0.000     651.248     865.544
===================================================================================
Ljung-Box (L1) (Q):                   0.13   Jarque-Bera (JB):                85.29
Prob(Q):                              0.72   Prob(JB):                         0.00
Heteroskedasticity (H):               0.23   Skew:                             0.87
Prob(H) (two-sided):                  0.00   Kurtosis:                         5.82
===================================================================================
```

Fig 1.26.2 Best Model (Rose wine)

1.27 Make a forecast for the next 12 months:

Forecasted values: (Sparkling wine)

- ✓ 1995-08-01 2042.210988
- ✓ 1995-09-01 2474.249527
- ✓ 1995-10-01 3381.697898
- ✓ 1995-11-01 3949.119077
- ✓ 1995-12-01 6194.176361
- ✓ 1996-01-01 1126.532412
- ✓ 1996-02-01 1641.718783
- ✓ 1996-03-01 1831.670258
- ✓ 1996-04-01 1813.235344
- ✓ 1996-05-01 1678.539689
- ✓ 1996-06-01 1697.096221
- ✓ 1996-07-01 2039.414408

Forecasted confidence intervals:

	lower Sparkling	upper Sparkling
✓ 1995-08-01	1282.307253	2802.114722
✓ 1995-09-01	1703.101093	3245.397962
✓ 1995-10-01	2608.524875	4154.870922
✓ 1995-11-01	3175.231061	4723.007093
✓ 1995-12-01	5420.035553	6968.317169
✓ 1996-01-01	352.302190	1900.762633
✓ 1996-02-01	867.456932	2415.980635
✓ 1996-03-01	1057.397216	2605.943299
✓ 1996-04-01	1038.958344	2587.512345
✓ 1996-05-01	904.261288	2452.818090
✓ 1996-06-01	922.817325	2471.375117
✓ 1996-07-01	1265.135337	2813.693480

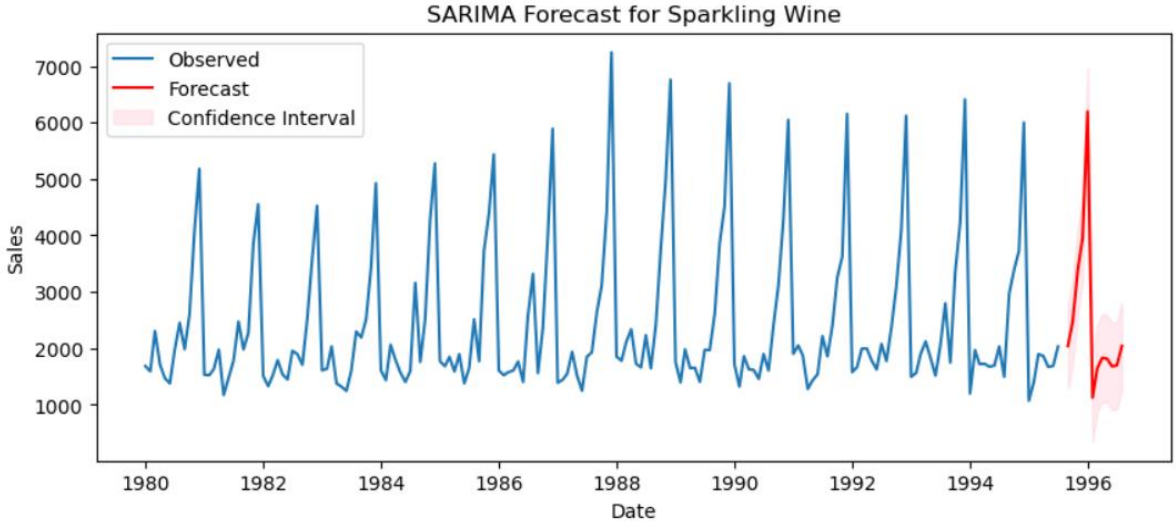


Fig 1.27.1 SARIMA Forecast for Sparkling Wine

Forecasted values for Rose Wine:

- ✓ 1995-08-01 49.304156
- ✓ 1995-09-01 47.578757
- ✓ 1995-10-01 48.420629
- ✓ 1995-11-01 48.009856
- ✓ 1995-12-01 48.210284
- ✓ 1996-01-01 48.112489
- ✓ 1996-02-01 48.160206
- ✓ 1996-03-01 48.136924
- ✓ 1996-04-01 48.148284
- ✓ 1996-05-01 48.142741
- ✓ 1996-06-01 48.145445
- ✓ 1996-07-01 48.144126

Forecasted confidence intervals for Rose Wine:

lower Rose upper Rose

- ✓ 1995-08-01 -4.671273 103.279585
- ✓ 1995-09-01 -8.560808 103.718322
- ✓ 1995-10-01 -7.748315 104.589574
- ✓ 1995-11-01 -8.792978 104.812689
- ✓ 1995-12-01 -8.831863 105.252431
- ✓ 1996-01-01 -9.333763 105.558741
- ✓ 1996-02-01 -9.601641 105.922053
- ✓ 1996-03-01 -9.979121 106.252968
- ✓ 1996-04-01 -10.299959 106.596527
- ✓ 1996-05-01 -10.645385 106.930867
- ✓ 1996-06-01 -10.975954 107.266845
- ✓ 1996-07-01 -11.310934 107.599185

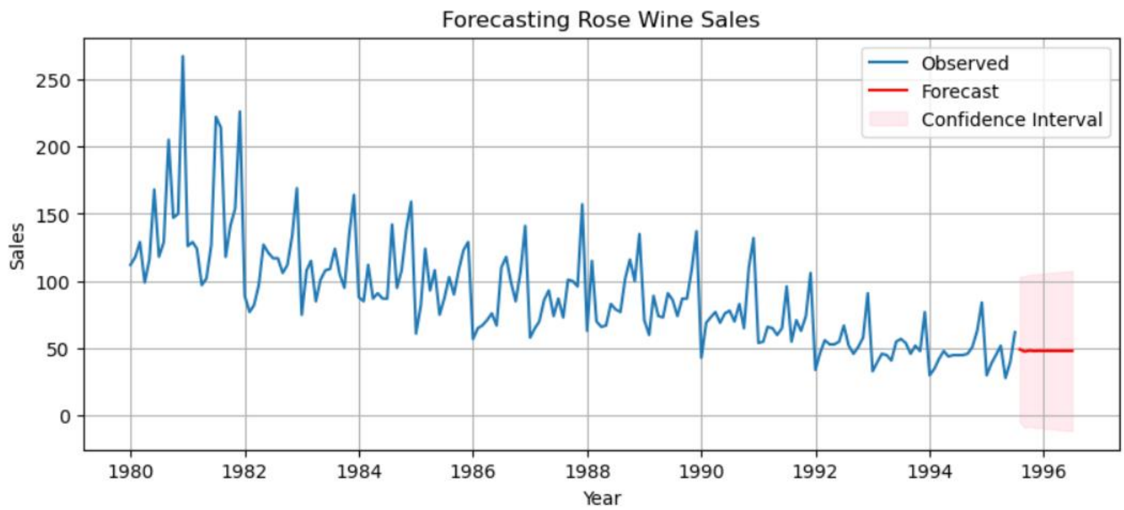


Fig 1.27.2 Forecasting for Rose Wine

Actionable Insights & Recommendations

1.28 Conclude with the key takeaways (actionable insights and recommendations) for the business:

Based on the analysis and forecasting, here are some actionable insights and recommendations for the business:

Key Insights and Recommendations:

Sparkling Wine Sales

- ❖ Seasonal Trends: The SARIMA model captures strong seasonal patterns in Sparkling Wine sales, indicating periods of higher and lower demand throughout the year.
- ❖ Forecast Accuracy: The forecasted values suggest significant fluctuations in sales, indicating variable demand patterns.
- ❖ Confidence Intervals: The confidence intervals provide a range of possible outcomes, highlighting the uncertainty in predictions.
- ❖ Model Performance: The SARIMA model demonstrates good performance with low MAE and RMSE, indicating reliable predictions relative to historical data.

Recommendations:

- ❖ Inventory Management: Optimize inventory levels based on forecasted sales. Prepare for higher demand during peak seasons (e.g., holidays) and avoid overstock during slower periods to reduce holding costs and improve efficiency.
- ❖ Marketing Strategies: Leverage the understanding of seasonal peaks to design effective marketing campaigns. Schedule promotions and advertising efforts to coincide with periods of anticipated increased demand.
- ❖ Product Development: Explore the potential for new product variants or packaging sizes based on forecasted trends and consumer preferences. Introduce these new products during peak sales periods to attract a broader customer base.
- ❖ Data Monitoring: Regularly compare actual sales with forecasted values and confidence intervals. Adjust strategies promptly based on real-time market dynamics and deviations from predicted trends.
- ❖ Customer Insights: Collect and analyze customer feedback to refine forecasting models and better understand consumer preferences. Use this information to enhance the accuracy of future sales predictions and improve product offerings.

Rose Wine Sales:

- ❖ Seasonal Trends: There are visible seasonal patterns in the sales data, indicating periods of higher and lower demand throughout the year.
- ❖ Forecast Accuracy: The forecasted values suggest stable sales trends with moderate fluctuations, indicating a predictable demand pattern.
- ❖ Confidence Intervals: The confidence intervals show a range of possible outcomes, indicating some uncertainty in the predictions.
- ❖ Model Performance: The SARIMA model used demonstrates good performance with low Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), suggesting reliable predictions relative to historical data.

Recommendations:

- ❖ Inventory Management: Use the forecasted sales to optimize inventory levels. Prepare for higher demand during peak seasons (e.g., summer months) and avoid overstock during slower periods to improve operational efficiency and reduce holding costs.
- ❖ Marketing Strategies: Tailor marketing campaigns to align with seasonal peaks. Strategically time promotions and advertising efforts to capitalize on anticipated increases in consumer demand.
- ❖ Product Development: Consider diversifying or expanding the Rose Wine product line based on forecasted trends and consumer preferences. Introduce new variants or packaging sizes aligned with peak sales periods to attract a broader customer base.
- ❖ Data Monitoring: Continuously monitor actual sales against forecasted values and confidence intervals. Quickly adjust strategies based on real-time market dynamics and deviations from predicted trends.
- ❖ Customer Insights: Gather feedback from customers to refine forecasting models further. Understanding evolving consumer preferences and behaviors can enhance the accuracy of future sales predictions.

Comprehensive Recommendations:

- ❖ Optimize Inventory Management: Use sales forecasts to maintain optimal inventory levels, reducing holding costs and minimizing the risk of stockouts or overstock situations.
- ❖ Tailor Marketing Campaigns: Align marketing efforts with seasonal sales peaks. Utilize targeted promotions, advertising, and events to drive sales during high-demand periods.
- ❖ Innovate Product Offerings: Based on forecasted trends, consider expanding product lines or introducing new packaging options to meet consumer preferences and capture market opportunities.
- ❖ Monitor Sales Data Continuously: Regularly compare actual sales data with forecasted values to make timely adjustments to inventory, marketing, and sales strategies.

- Leverage Customer Insights: Continuously gather and analyze customer feedback to refine forecasting models and improve product and service offerings.
- Enhance Forecast Accuracy: Regularly update and validate forecasting models with the latest sales data to maintain accuracy and reliability in predictions.