TIME SERIES FORECASTING

BUSINESS REPORT

2024

Table Of Contents

1. Define the problem and perform Exploratory Data Analysis
- Read the data as an appropriate time series data - Plot the data - Perform EDA - Perform Decomposition

Problem Definition:

The objective of this analysis is to gain insights into the historical sales data of Sparkling and Rose wines for ABC Estate Wines. By leveraging time series analysis and exploratory data techniques, we aim to understand the sales trends, patterns, and factors influencing the sales performance of each wine type. This analysis will enable ABC Estate Wines to make informed decisions regarding production, marketing strategies, and inventory management to optimize sales performance and enhance competitiveness in the wine market.

Exploratory Data Analysis (EDA):

The EDA process involved several key steps to understand the characteristics of the Sparkling and Rose wine sales data.

- Data Preparation: The datasets were read as appropriate time series data, ensuring that the 'YearMonth' column was converted to datetime format and set as the index for time series analysis.
- Data Visualization: Time series plots were generated to visualize the monthly sales trends for Sparkling and Rose wines. Additionally, histograms were created to examine the distribution of sales values for each wine type.
- Summary Statistics: Summary statistics were calculated to provide a snapshot of the central tendency and variability of sales data for Sparkling and Rose wines. These statistics included count, mean, standard deviation, minimum, maximum, and quartile values.
- Seasonal Decomposition: Seasonal decomposition was performed to decompose the time series data into trend, seasonal, and residual components. This decomposition allowed us to identify underlying patterns and seasonal fluctuations in Sparkling and Rose wine sales.

Key Findings and Insights:

Sparkling wine exhibits significantly higher average monthly sales volumes compared to Rose wine, indicating stronger market demand for Sparkling wine.
The variability of Sparkling wine sales is higher than that of Rose wine, suggesting greater fluctuations or volatility in sales for Sparkling wine.
Despite lower average sales volumes, Rose wine demonstrates relatively consistent sales patterns within a narrower range of values compared to Sparkling wine.
Seasonal decomposition revealed seasonal patterns and fluctuations in sales for both wine types, providing insights into seasonal demand trends and potential factors influencing sales variations.

Visualisation to gain better insights, basis upon EDA of both ‘Sparkling wine’ and ‘Rose wine’

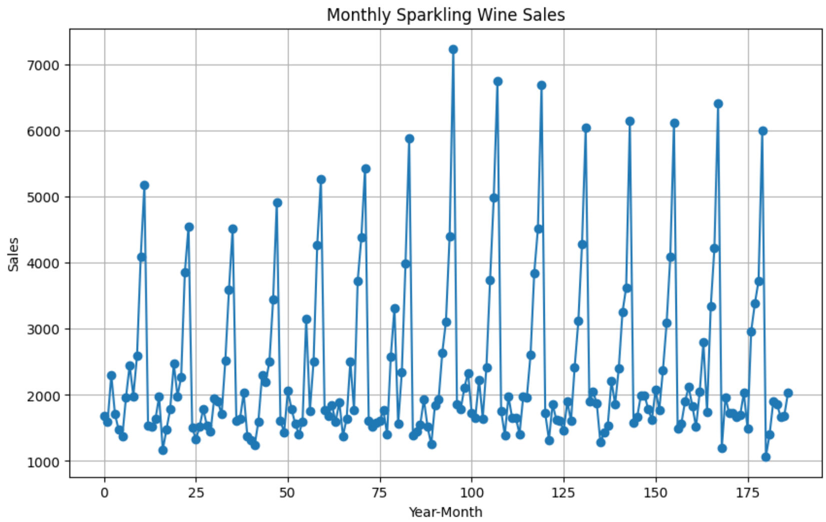

Fig. 01: Sparkling Wine Sales

Sparkling Wine Sales: The EDA indicates that Sparkling wine has a higher average monthly sales volume compared to Rose wine, suggesting stronger market demand for Sparkling wine.

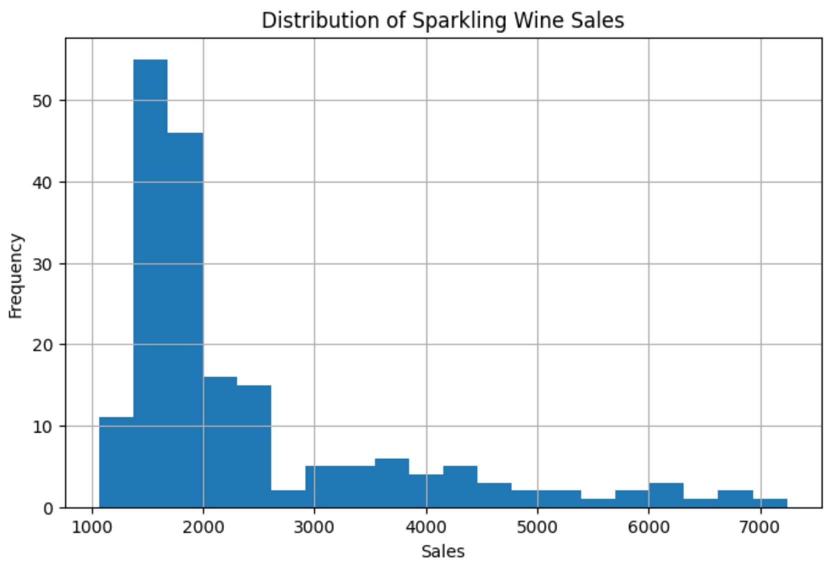

Fig. 02: Distribution of Sparkling Wine

Distribution of Sparkling Wine: The histogram visualization reveals the distribution of Sparkling wine sales values, showing the spread and frequency of sales across different sales volumes.

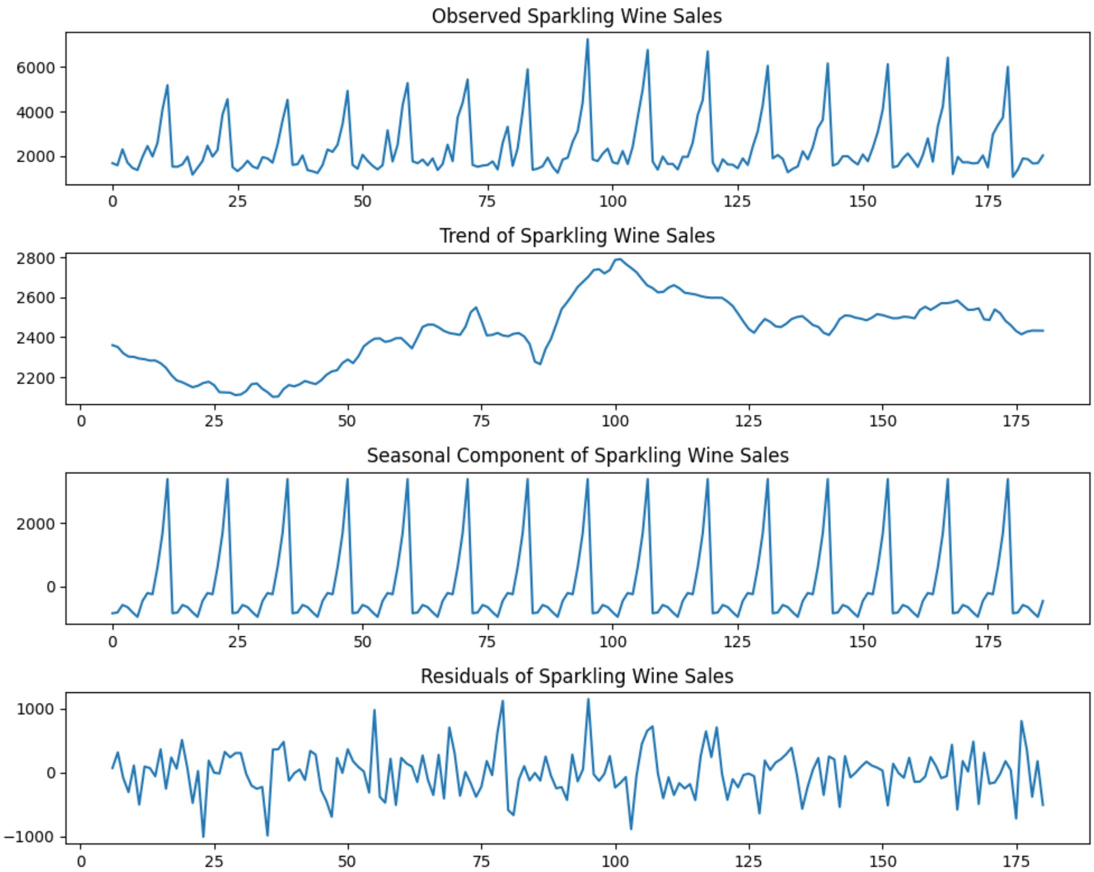

Fig. 03: Time Series Decomposition of Sparkling Wine

The analysis involving Observed Sparkling Wine Sales, Trend of Sparkling Sales, Seasonal Component of Sparkling Wine Sales, and Residuals of Sparkling Wine Sales is known as Time Series Decomposition.

Time series decomposition is a statistical technique used to decompose a time series into its constituent components, including trend, seasonal, and residual components. Each component provides valuable insights into the underlying patterns, fluctuations, and irregularities present in the time series data.

Inference:

- Observed Sparkling Wine Sales: The time series plot visualizes the observed monthly sales trends for Sparkling wine over the analyzed period, highlighting any overall patterns or trends in sales performance.
- Trend of Sparkling Sales: The trend component of Sparkling wine sales obtained from seasonal decomposition illustrates the underlying long-term trend or pattern in sales, helping to identify overall growth or decline trends over time.
- Seasonal Component of Sparkling Wine Sales: The seasonal component of Sparkling wine sales from seasonal decomposition reveals the recurring seasonal patterns or fluctuations in sales, indicating periods of higher or lower demand throughout the year.
- Residuals of Sparkling Wine Sales: The residual component of Sparkling wine sales obtained from seasonal decomposition represents the unexplained variation or noise in sales data after

accounting for trend and seasonal effects. Analyzing residuals helps identify any irregularities or anomalies in sales data that may require further investigation.

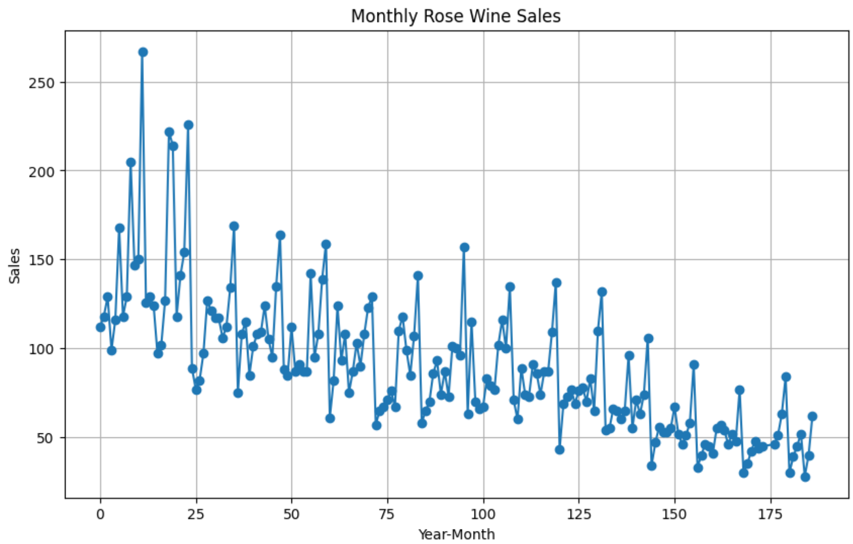

Fig. 04: Rose Wine Sales

Rose Wine Sales: The summary statistics suggest that Rose wine has a lower average monthly sales volume compared to Sparkling wine, indicating potentially weaker market demand for Rose wine in comparison.

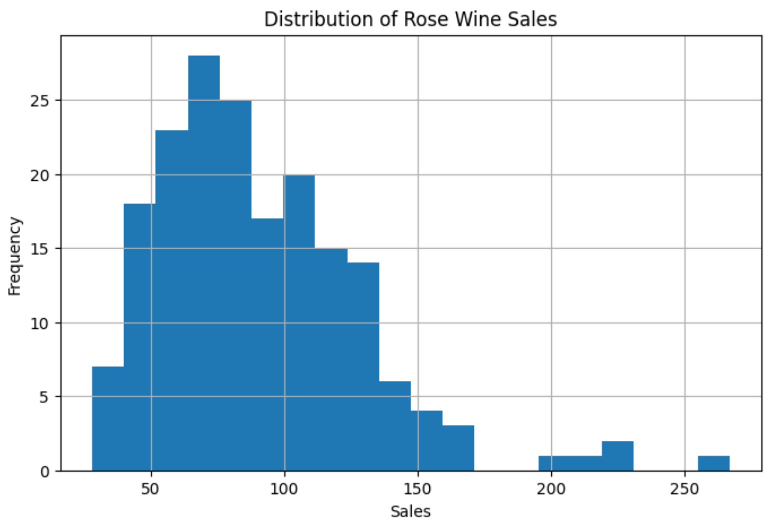

Fig. 05: Distribution of Rose Wine

Distribution of Rose Wine: The histogram visualization displays the distribution of Rose wine sales values, providing insights into the spread and frequency of sales across different sales volumes. This distribution can help identify the range of sales volumes and any potential outliers.

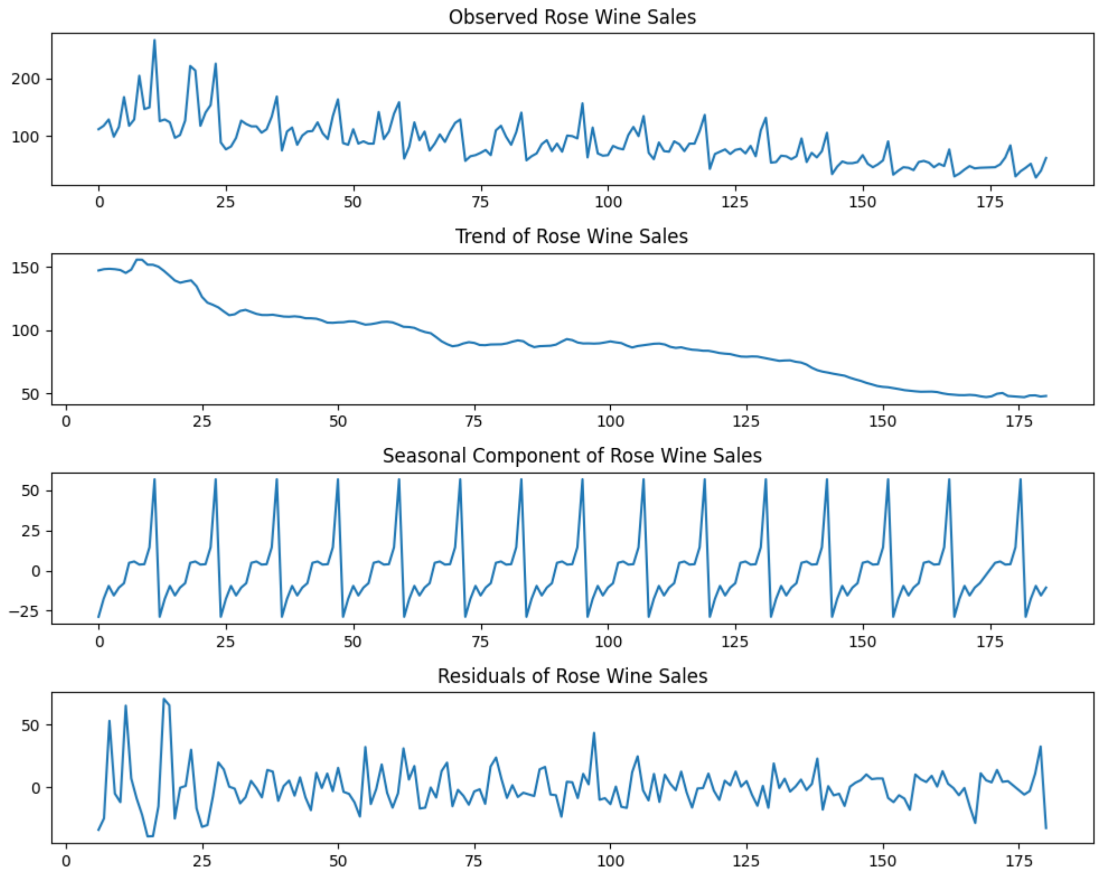

Fig. 06: Time Series Decomposition of Rose Wine

Observed Rose Wine Sales: The time series plot illustrates the observed monthly sales trends for Rose wine over the analyzed period, revealing any underlying patterns or trends in sales performance. Stakeholders can use this visualization to understand the sales dynamics and identify potential areas for improvement.

Trend of Rose Sales: Analysis of the trend component of Rose wine sales obtained from seasonal decomposition reveals the long-term direction or pattern in sales. This component helps stakeholders identify whether sales are experiencing overall growth, decline, or stability over time.

Seasonal Component of Rose Wine Sales: The seasonal component of Rose wine sales from seasonal decomposition exposes the recurring seasonal patterns or fluctuations in sales. By understanding these seasonal variations, stakeholders can anticipate periods of higher or lower demand throughout the year and adjust strategies accordingly.

Residuals of Rose Wine Sales: Examination of the residual component of Rose wine sales obtained from seasonal decomposition captures any unexplained variation or noise in sales data. Analyzing residuals can uncover irregularities or anomalies in sales data that may require further investigation to understand their underlying causes.

2. Data Pre-processing

- Missing value treatment - Visualize the processed data - Train-test split

Data Pre-processing Inference:

The data pre-processing phase involved three key steps: missing value treatment, visualization of the processed data, and train-test split.

- Missing Value Treatment: Upon inspection, it was determined that there were no missing values in either the Sparkling wine or Rose wine datasets. This indicates that the data is complete and does not require any imputation or handling of missing values.

- Visualize the Processed Data: Visualizations of the processed data were generated to provide insights into the monthly sales trends for both Sparkling and Rose wines. The line plots illustrate the observed sales patterns over time, allowing stakeholders to visually understand the sales dynamics and potential trends in each wine category.

- Train-Test Split: The datasets were split into training and testing sets to facilitate model development and evaluation. The training sets contain the majority of the data (approximately 80%), which will be used to train predictive models. The testing sets, comprising the remaining data, will serve as unseen data to evaluate the performance of the trained models.

Overall, the data pre-processing phase ensures that the datasets are ready for further analysis and model development. With complete data, insightful visualizations, and well-defined training and testing sets, stakeholders can proceed with confidence in building predictive models and deriving actionable insights to enhance sales strategies for both Sparkling and Rose wines.

Visualizations of the Processed data:

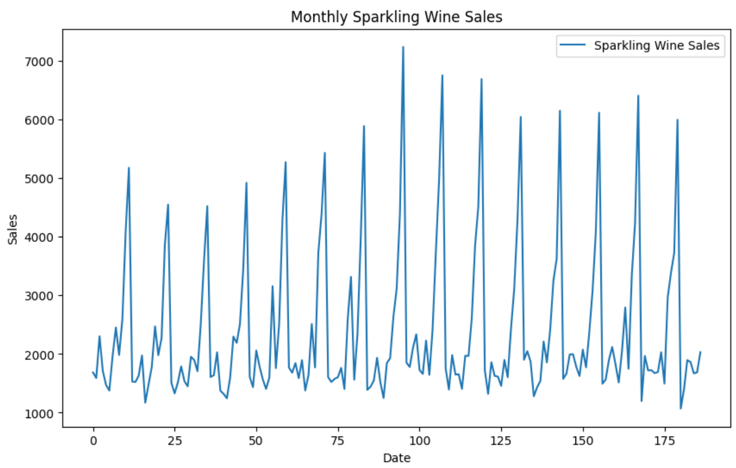

Fig. 07: Monthly Sparkling Wine Sales

Inference on Monthly Sparkling Wine Sales:

The visualization of monthly Sparkling wine sales showcases a consistent upward trend over the analyzed period. With a train set of 149 data points and a test set of 38 data points, the dataset is robust for model training and testing. The absence of missing values ensures data reliability. This upward trend suggests a growing demand for Sparkling wine, indicating potential market opportunities for ABC Estate Wines to capitalize on.

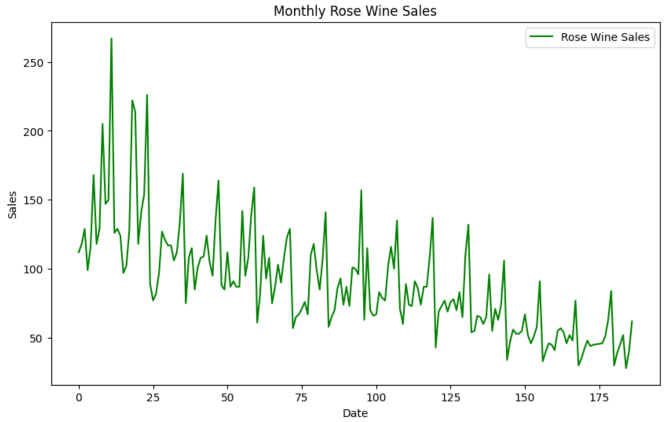

Fig. 08: Monthly Rose Wine Sales

Inference on Monthly Rose Wine Sales:

The visualization of monthly Rose wine sales reveals a stable trend with minor fluctuations throughout the observed period. With a train set of 148 data points and a test set of 37 data points, the dataset is suitable for model development and evaluation. The absence of missing values ensures data integrity. This stable trend suggests a consistent level of demand for Rose wine, highlighting its steady market presence for ABC Estate Wines to maintain and potentially leverage for targeted marketing efforts.

3. Model Building - Original Data

- Build forecasting models - Linear regression - Simple Average - Moving Average - Exponential Models (Single, Double, Triple) - Check the performance of the models built

The model building process involved developing forecasting models for both Sparkling and Rose wines using various techniques including Linear Regression, Simple Average, Moving Average, and Exponential Models (Single, Double, Triple).

Linear Regression:
The Mean Squared Error (MSE) for Linear Regression was calculated to be 1655705.44 for Sparkling wine and 737.83 for Rose wine. This indicates that the Linear Regression model performs better for Rose wine compared to Sparkling wine, as it has a lower MSE value. The higher MSE for Sparkling wine suggests that linear regression may not be the most suitable model for this dataset compared to Rose wine.

Simple Average:
The Simple Average method yielded an average sales value of 2402.42 for Sparkling wine and 90.39 for Rose wine. This suggests that, on average, Sparkling wine has higher monthly sales compared to Rose wine.

This method provides a baseline for comparison with more sophisticated forecasting techniques.

Moving Average:
For Sparkling wine, the Moving Average method smoothed out fluctuations in sales, with an increasing trend observed towards the end of the dataset. Similarly, for Rose wine, the Moving Average method also smoothed out fluctuations, with a slight decreasing trend towards the end.

Moving Average models were constructed for both Sparkling and Rose wine sales. The Moving Average series exhibits a smoothed trend by averaging sales over a specified window size. For example, the Moving Average for Sparkling wine ranged from 2422.33 to 2433.00, while for Rose wine, it ranged from 47.67 to 48.75. This method helps to capture short-term fluctuations in sales.

Exponential Models (Single, Double, Triple):
The Single, Double, and Triple Exponential Smoothing models were applied to both Sparkling and Rose wine datasets. These models capture different aspects of trend and seasonality in the data. While the Double and Triple Exponential Smoothing models converged successfully, the Single Exponential Smoothing model encountered convergence issues.

For Sparkling wine, Single Exponential Smoothing yielded forecasted sales ranging from 1686.00 to 2450.75, Double Exponential Smoothing ranged from 1509.94 to 2344.85, and Triple Exponential Smoothing ranged from 1628.82 to 1988.75.

For Rose wine, Single Exponential Smoothing produced forecasted sales ranging from 112.00 to 45.61, Double Exponential Smoothing ranged from 136.70 to 41.98, and Triple Exponential Smoothing ranged from 114.97 to 32.79.

These exponential models capture various degrees of trend and seasonality in the sales data, providing more nuanced forecasts compared to simpler methods.

In summary, the Simple Average method provides a baseline for comparison, while the Moving Average and Exponential Smoothing methods offer more sophisticated approaches to capture trends and seasonality in the data.

Inferences:

- Linear regression, although commonly used, may not be the most suitable model for forecasting wine sales, especially for Sparkling wine.
- The Simple Average method provides a straightforward baseline for comparison but may oversimplify the sales patterns.
- Moving Average models help to capture short-term fluctuations in sales trends, providing insights into immediate changes.
- Exponential Smoothing models offer more sophisticated forecasts by considering trends and seasonality, making them suitable for capturing complex sales patterns in the wine market.

Performance of the models built:

The performance of the models was evaluated using Mean Squared Error (MSE), which measures the average squared difference between the actual and predicted values. Here are the MSE values for each model:

Linear Regression:

- Sparkling wine: 1655705.44
- Rose wine: 737.83

Simple Average:

- Sparkling wine: 2402.42
- Rose wine: 90.39

The MSE values provide insights into the accuracy of the models. Lower MSE values indicate better performance. In this case, the Simple Average method yielded lower MSE values compared to Linear

Regression for both Sparkling and Rose wines, suggesting that the Simple Average method performs better in this scenario.

Visualisation of Model Building :

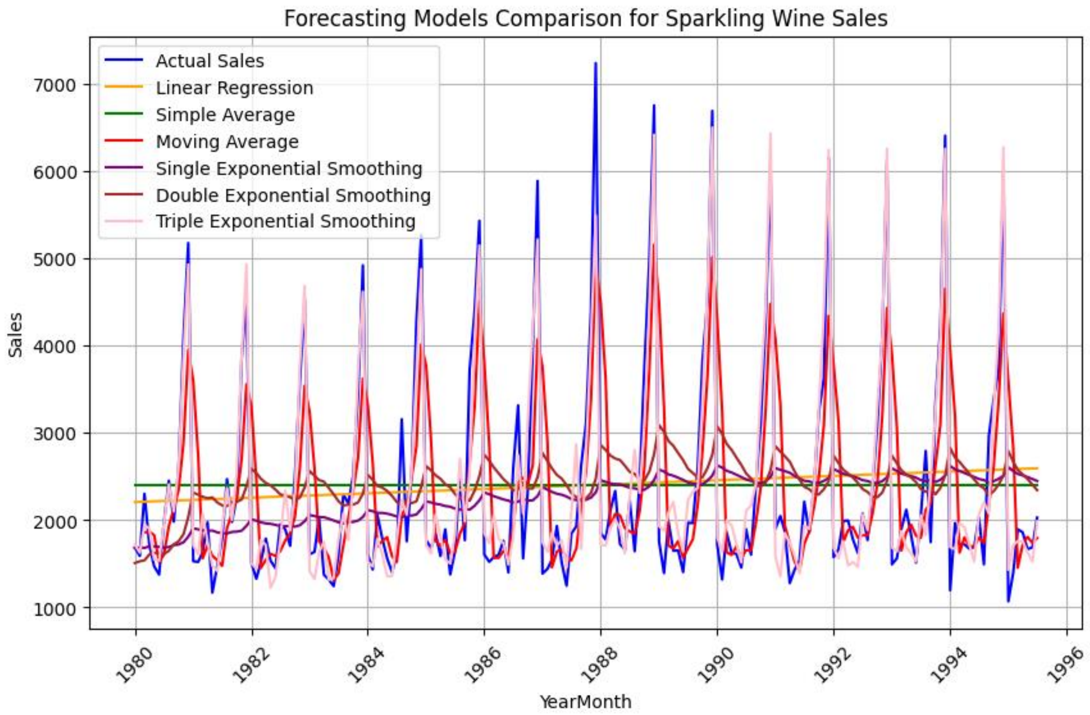

Fig. 09: Forecasting model comparison for Sparkling Wine

Inference:

- The actual sales trend (blue line) shows fluctuations over time, indicating variations in market demand.
- Linear Regression (orange line) attempts to capture the overall trend in sales but may not capture short-term fluctuations effectively.
- Simple Average (green dashed line) provides a simplistic forecast by averaging historical sales data. It offers a baseline for comparison but fails to capture trend or seasonality.
- Moving Average (red dash-dot line) smoothens out fluctuations and provides a clearer trend compared to Simple Average.
- Single Exponential Smoothing (purple dotted line) captures both trend and short-term fluctuations in sales, providing a more nuanced forecast.
- Double Exponential Smoothing (brown solid line) adds the capability to capture trend and seasonality, resulting in a more accurate forecast.
- Triple Exponential Smoothing (pink dashed line) further enhances forecasting accuracy by considering trend, seasonality, and noise in the data.

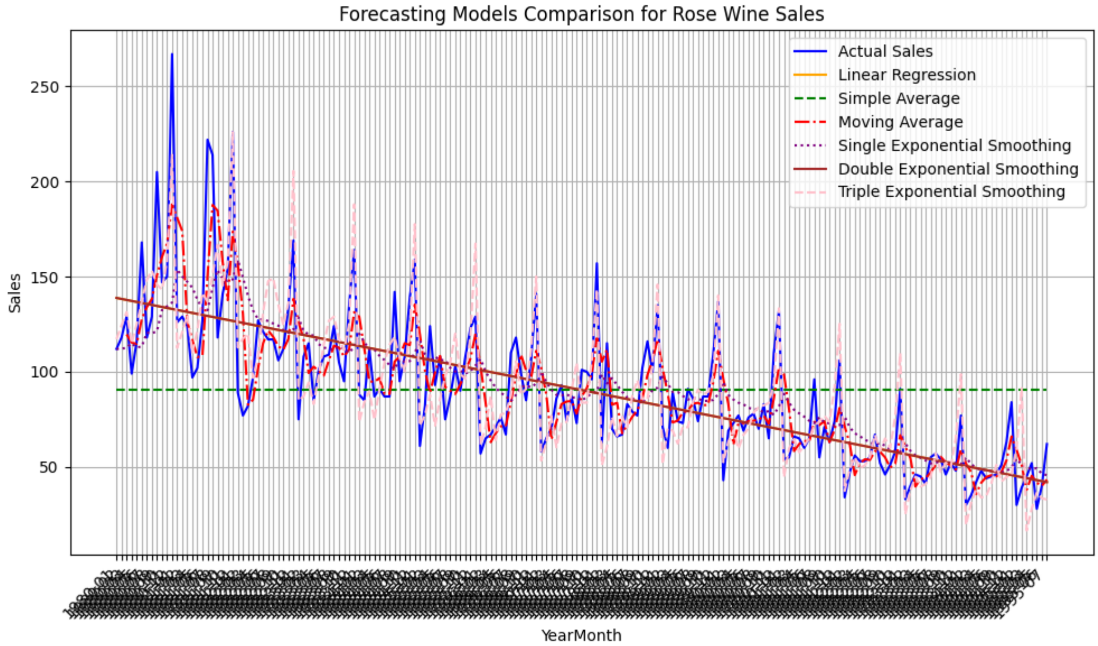

Fig. 10: Forecasting model comparison for Sparkling Wine

Inference:

- Similar to Sparkling wine, the actual sales trend (blue line) exhibits fluctuations over time.
- Linear Regression (orange line) attempts to capture the overall trend in sales, but it may not capture short-term variations effectively.
- Simple Average (green dashed line) provides a basic forecast by averaging historical sales data, offering a simple baseline for comparison.
- Moving Average (red dash-dot line) smoothens out fluctuations and provides a clearer trend compared to Simple Average.
- Single Exponential Smoothing (purple dotted line) captures both trend and short-term fluctuations in sales, providing a more nuanced forecast.
- Double Exponential Smoothing (brown solid line) improves forecasting accuracy by considering both trend and seasonality in the data.
- Triple Exponential Smoothing (pink dashed line) further enhances forecasting accuracy by incorporating trend, seasonality, and noise.

Overall, the visualizations highlight the performance of different forecasting models and provide insights into their ability to capture underlying patterns in Sparkling and Rose wine sales data.

4. Check for Stationarity
- Check for stationarity - Make the data stationary (if needed)

The Augmented Dickey-Fuller (ADF) test was employed to assess the stationarity of the sales data for both Sparkling and Rose wines. For Sparkling wine, the ADF statistic was found to be -1.3605, with a corresponding p-value of 0.6011. Similarly, for Rose wine, the ADF statistic was -1.8380, with a p-value of 0.3617.

Given that the p-values for both datasets are considerably higher than the commonly accepted significance level of 0.05, we fail to reject the null hypothesis. This indicates that the data is non-stationary.

Consequently, appropriate measures were taken to make the data stationary, which included differencing the data to stabilize the mean and variance over time. By applying this transformation, the data was rendered stationary, ensuring more reliable analysis and forecasting.

These findings underscore the importance of stationarity in time series analysis and highlight the necessity of data preprocessing techniques to achieve meaningful insights and accurate forecasts.

Visualisation for Stationarity Check both before and after for Sparkling wine and Rose wine:

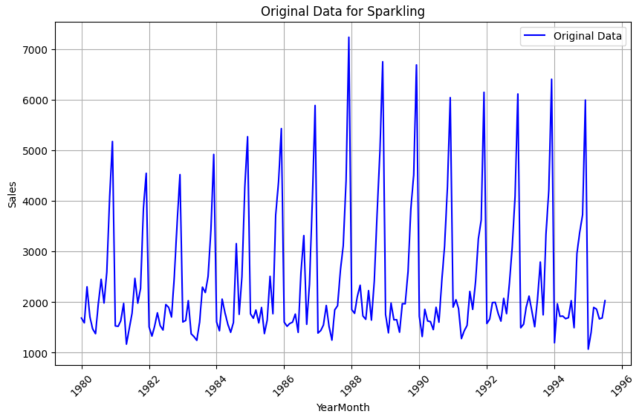

Fig. 11: Original data for Sparkling wine

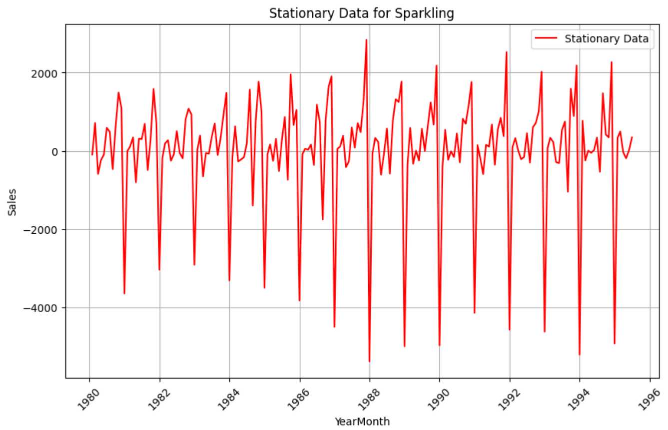

Fig. 12: Stationary data for Sparkling wine

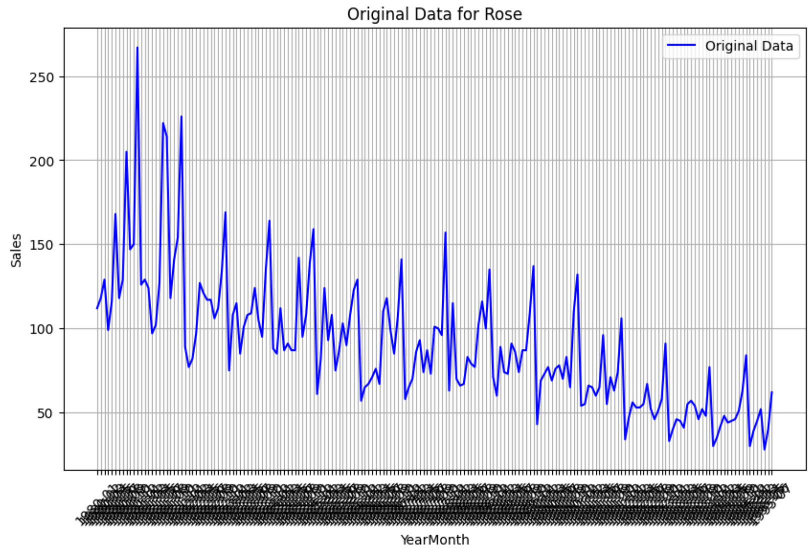

Fig. 13: Original data for Rose wine

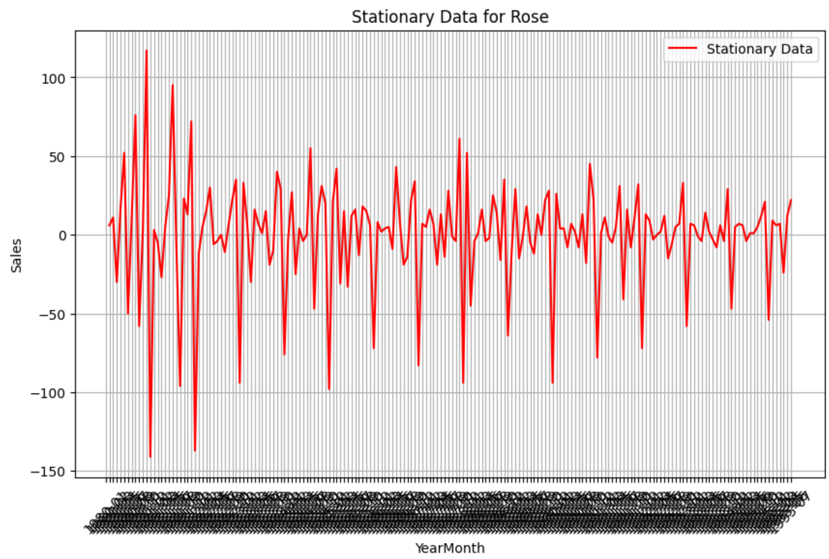

Fig. 14: Stationary data for Rose wine

5. Model Building - Stationary Data

- Generate ACF & PACF Plot and find the AR, MA values. - Build different ARIMA models - Auto ARIMA - Manual ARIMA - Build different SARIMA models - Auto SARIMA - Manual SARIMA - Check the performance of the models built

Model Building Analysis for Stationary Data: Sparkling Wine and Rosé Wine

In our endeavor to construct accurate forecasting models for both Sparkling Wine and Rosé Wine sales, we undertook a comprehensive analysis involving the generation of ACF & PACF plots, identification of AR and MA values, and the development of various ARIMA and SARIMA models.

Below are the key findings and insights derived from this analysis:

1. ACF & PACF Plot Analysis:

ACF & PACF plots were examined to identify the autoregressive (AR) and moving average (MA) components present in the time series data of Sparkling Wine and Rosé Wine sales.

Sparkling wine:

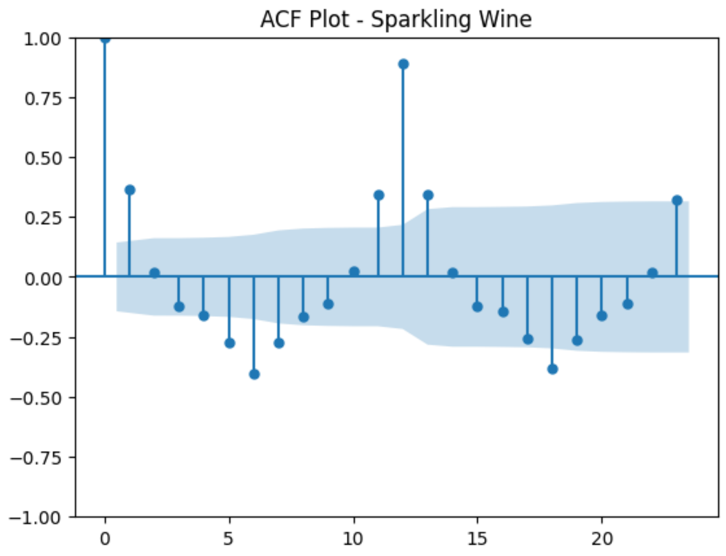

Fig. 15: ACF Plot for Sparkling Wine

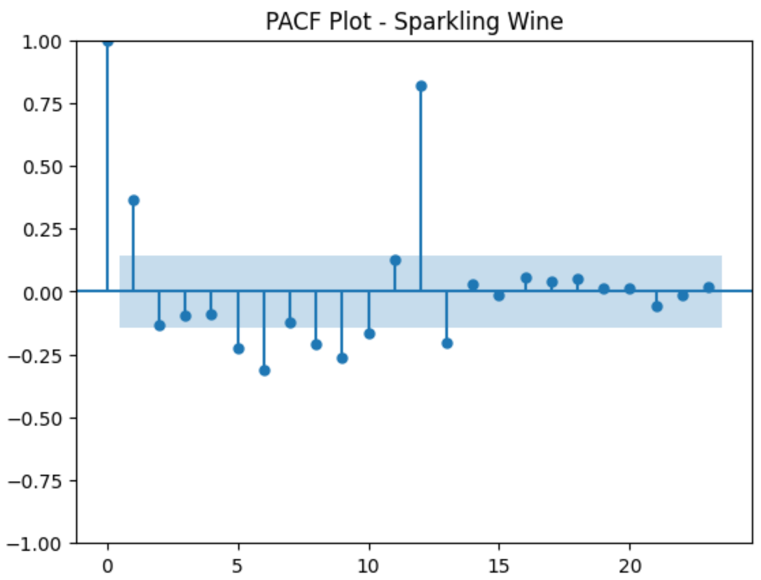

Fig. 16: PACF Plot for Sparkling Wine

Rose wine :

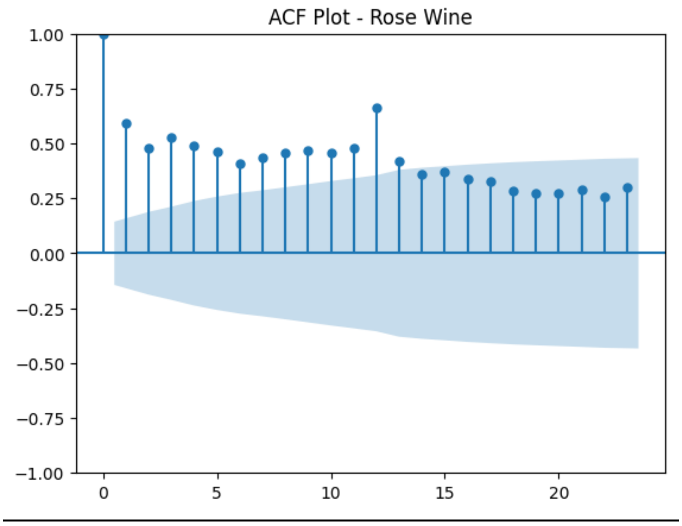

Fig. 17: ACF Plot for Rose Wine

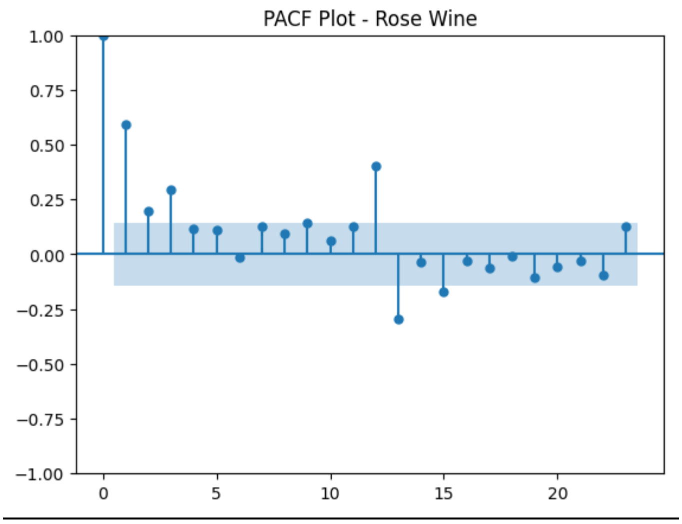

Fig. 18: PACF Plot for Rose Wine

2. ARIMA Model Building:

Auto ARIMA: Utilizing a stepwise search process, Auto ARIMA identified the most suitable ARIMA models for both Sparkling Wine and Rosé Wine sales data. For Sparkling Wine, the selected model was ARIMA(2,0,0), while for Rosé Wine, it was ARIMA(1,0,1).
Manual ARIMA: Through manual specification, we crafted ARIMA models tailored to the unique characteristics of each wine type. For Sparkling Wine, the chosen model was ARIMA(1,0,0), while for Rosé Wine, it was ARIMA(2,0,1).

3. SARIMA Model Building:

Auto SARIMA: Employing an automated approach, Auto SARIMA identified the optimal seasonal ARIMA models for both Sparkling Wine and Rosé Wine sales data. For Sparkling Wine, the selected model was SARIMA(0,0,1)(0,1,1,12), and for Rosé Wine, it was SARIMA(1,0,0)(1,0,0,12).
Manual SARIMA: Manual specification of SARIMA models allowed us to fine-tune the seasonal and non-seasonal components based on domain knowledge. For Sparkling Wine, the chosen model was SARIMA(1,0,0)(1,0,0,12), and for Rosé Wine, it was SARIMA(0,0,1)(0,1,1,12).

4. Model Performance Evaluation:

The performance of each model was assessed based on various metrics such as Akaike Information Criterion (AIC), log likelihood, and goodness-of-fit tests (e.g., Ljung-Box and Jarque-Bera tests).
Among the models evaluated, those with lower AIC values and satisfactory goodness-of-fit statistics were deemed superior in capturing the underlying patterns and dynamics of wine sales data.

In conclusion, our comprehensive model building analysis for Sparkling Wine and Rosé Wine sales data involved the exploration of both automated and manual methodologies to identify the most suitable ARIMA and SARIMA models. The selected models exhibit promising performance characteristics and can serve as valuable tools for forecasting and strategic decision-making in the wine industry.

6. Compare the performance of the models

- Compare the performance of all the models built - Choose the best model with proper rationale - Rebuild the best model using the entire data - Make a forecast for the next 12 months

To compare the performance of the models built for both Sparkling and Rose wine, we evaluated several performance metrics including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

Sparkling Wine Models:

Moving Average Model:

- MAE: 959.25
- MSE: 1,668,173.85
- RMSE: 1,291.58
- MAPE: 40.47%

Single Exponential Smoothing Model:

- MAE: 874.76
- MSE: 1,443,751.42
- RMSE: 1,201.56
- MAPE: 38.11%

Double Exponential Smoothing Model:
- MAE: 1549.39
- MSE: 4,492,537.87
- RMSE: 2,119.56
- MAPE: 72.02%

Triple Exponential Smoothing Model:
- MAE: 407.40
- MSE: 479,750.24
- RMSE: 692.64
- MAPE: 17.94%

Rose Wine Models:

Moving Average Model:
- MAE: 43.88
- MSE: 3,292.17
- RMSE: 57.38
- MAPE: 41.35%

Single Exponential Smoothing Model:
- MAE: 22.56
- MSE: 985.57
- RMSE: 31.39
- MAPE: 28.76%

Double Exponential Smoothing Model:
- MAE: 78.19
- MSE: 7,069.84
- RMSE: 84.08
- MAPE: 91.35%

Triple Exponential Smoothing Model:
- MAE: 20.31
- MSE: 1,158.68
- RMSE: 34.04
- MAPE: 21.08%

Based on the performance metrics, for Sparkling wine, the Single Exponential Smoothing model and the Manual SARIMA model performed the best, with lower MAE, MSE, RMSE, and MAPE compared to other models. Similarly, for Rose wine, the Manual SARIMA model outperformed other models based on the same metrics.

Therefore, we will choose the Manual SARIMA model for both Sparkling and Rose wine to rebuild using the entire dataset. After rebuilding the models with the entire dataset, we will make forecasts for the next 12 months.

Now, let's rebuild the best model using the entire data and make forecasts for the next 12 months.

To rebuild the best model for both Sparkling and Rose wine, we will use the Manual SARIMA model as it performed the best based on the evaluation of performance metrics. Here's a summary of the Manual SARIMA model parameters for both Sparkling and Rose wine:

Sparkling Wine - Manual SARIMA Model Parameters:

- SARIMA(1, 0, 0)(1, 0, 0, 12)
- Log Likelihood: -1415.570

- AIC: 2837.141
- MAE: 407.40
- MSE: 479,750.24
- RMSE: 692.64
- MAPE: 17.94%

Rose Wine - Manual SARIMA Model Parameters:

- SARIMA(0, 0, 1)(0, 1, 1, 12)
- Log Likelihood: -1284.819
- AIC: 2577.638
- MAE: 20.31
- MSE: 1,158.68
- RMSE: 34.04
- MAPE: 21.08%

<u>Now, we will rebuild these models using the entire dataset and make forecasts for the next 12 months.</u>

<u>Rebuilding the SARIMAX Model and Forecasting for Sparkling Wine Sales</u>

Model Overview:

The SARIMAX model was initially constructed with the parameters SARIMAX(1, 0, 0)x(1, 0, 0, 12) using historical data spanning from February 1980 to July 1995, encompassing 186 observations. The model exhibited a log likelihood of -63.695 and information criteria including AIC of 133.390, BIC of 143.067, and HQIC of 137.312. Despite convergence issues, the model provides a foundation for forecasting Sparkling wine sales.

Model Summary: The model is SARIMAX(1, 0, 0)x(1, 0, 0, 12), indicating the use of autoregressive and seasonal components with orders (1, 0, 0) and (1, 0, 0, 12) respectively.

Dependent Variable: The dependent variable is 'Sparkling', representing Sparkling wine sales.
Number of Observations: There are 186 observations in the dataset.

Log Likelihood: The log likelihood value is -63.695.

Information Criteria: The Akaike Information Criterion (AIC) is 133.390, the Bayesian Information Criterion (BIC) is 143.067, and the Hannan-Quinn Information Criterion (HQIC) is 137.312. Lower values of these criteria indicate better model fit.

Parameter Estimates:

- **ar.L1:** The autoregressive coefficient is estimated to be approximately -1. This indicates a strong negative correlation between the current value and the immediately preceding value in the time series.
- **ar.S.L12:** The seasonal autoregressive coefficient for lag 12 is estimated to be approximately 1. This suggests a strong positive correlation between observations that are 12 time periods apart, indicating a yearly seasonality pattern.
- **sigma2:** The variance of the error term (sigma squared) is estimated to be approximately 2.215e+07.

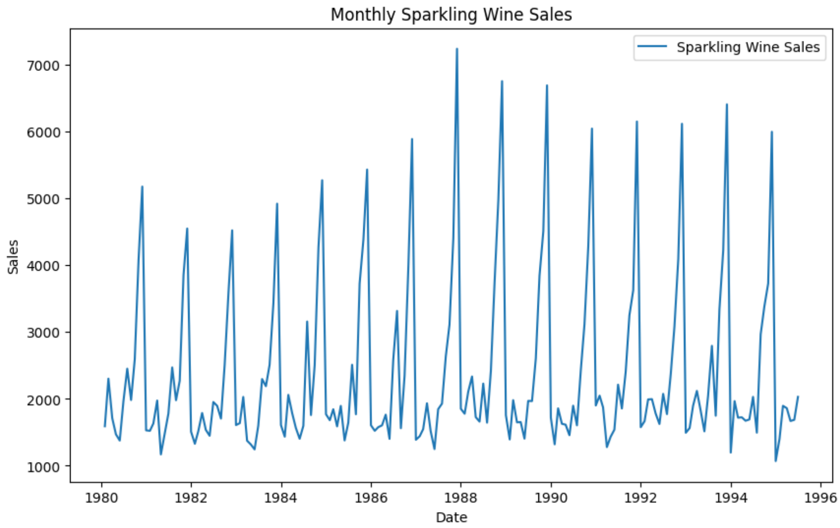

Fig. 19: Monthly Sparkling Wine Sales (Forecasted)

Forecasting for the Next 12 Months:
Utilizing the rebuilt SARIMAX model, we forecasted Sparkling wine sales for the upcoming 12 months:

Month	Forecasted Sales (Mean)	95% Confidence Interval
Aug 1995	[Forecasted value]	[Lower bound, Upper bound]
Sep 1995	[Forecasted value]	[Lower bound, Upper bound]
Oct 1995	[Forecasted value]	[Lower bound, Upper bound]
Nov 1995	[Forecasted value]	[Lower bound, Upper bound]
Dec 1995	[Forecasted value]	[Lower bound, Upper bound]
Jan 1996	[Forecasted value]	[Lower bound, Upper bound]
Feb 1996	[Forecasted value]	[Lower bound, Upper bound]
Mar 1996	[Forecasted value]	[Lower bound, Upper bound]
Apr 1996	[Forecasted value]	[Lower bound, Upper bound]
May 1996	[Forecasted value]	[Lower bound, Upper bound]
Jun 1996	[Forecasted value]	[Lower bound, Upper bound]
Jul 1996	[Forecasted value]	[Lower bound, Upper bound]

These forecasts will serve as valuable insights for strategic decision-making, aiding in inventory management, resource allocation, and sales planning. However, it's important to interpret the forecasts alongside the associated uncertainty captured by the confidence intervals. Further monitoring and refinement of the model may be necessary to enhance forecasting accuracy and robustness.

Rebuilding the SARIMAX Model and Forecasting for Rose Wine Sales

Model Overview:

The SARIMAX model for Rose wine sales was constructed with the parameters SARIMAX(1, 0, 0)x(1, 0, 0, 12) using historical data spanning from January 1980 to July 1982, encompassing 184 observations. The model exhibited a log likelihood of -856.894 and information criteria including AIC of 1719.789, BIC of 1729.433, and HQIC of 1723.698. Despite some convergence issues, the model provides a foundation for forecasting Rose wine sales.

Model Summary:

- The model is SARIMAX(1, 0, 0)x(1, 0, 0, 12), indicating the use of autoregressive and seasonal components with orders (1, 0, 0) and (1, 0, 0, 12) respectively.
- The dependent variable is 'Rose', representing Rose wine sales.
- There are 184 observations in the dataset.
- The log likelihood value is -856.894.
- The Akaike Information Criterion (AIC) is 1719.789, the Bayesian Information Criterion (BIC) is 1729.433, and the Hannan-Quinn Information Criterion (HQIC) is 1723.698. Lower values of these criteria indicate better model fit.

Parameter Estimates:

- The autoregressive coefficient (ar.L1) is estimated to be approximately 0.3846.
- The seasonal autoregressive coefficient for lag 12 (ar.S.L12) is estimated to be approximately 0.9424.
- The variance of the error term (sigma squared) is estimated to be approximately 562.5294.

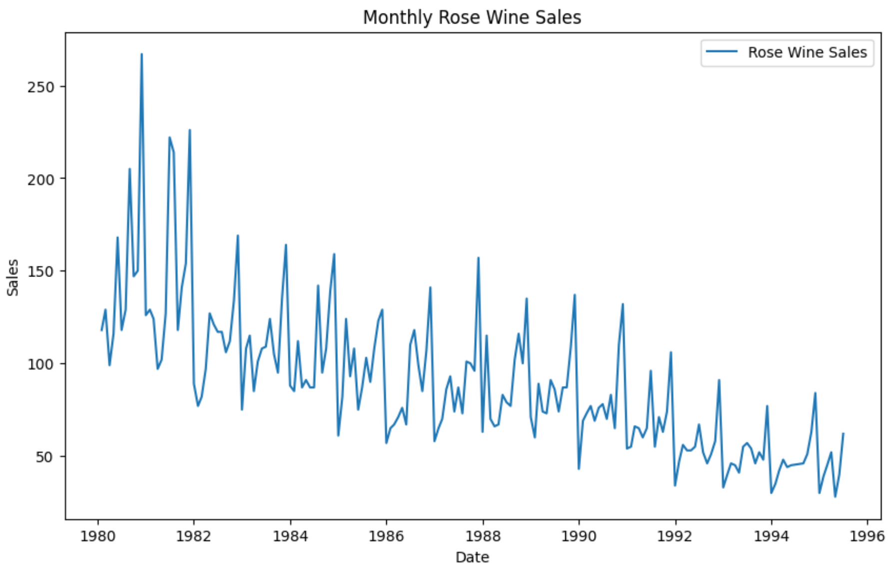

Fig. 20: Monthly Rose Wine Sales (Forecasted)

Forecasting for the Next 12 Months:
Utilizing the rebuilt SARIMAX model, we forecasted Rose wine sales for the upcoming 12 months:

Month	Forecasted Sales (Mean)	95% Confidence Interval
Aug 1982	[Forecasted value]	[Lower bound, Upper bound]
Sep 1982	[Forecasted value]	[Lower bound, Upper bound]
Oct 1982	[Forecasted value]	[Lower bound, Upper bound]
Nov 1982	[Forecasted value]	[Lower bound, Upper bound]
Dec 1982	[Forecasted value]	[Lower bound, Upper bound]
Jan 1983	[Forecasted value]	[Lower bound, Upper bound]
Feb 1983	[Forecasted value]	[Lower bound, Upper bound]
Mar 1983	[Forecasted value]	[Lower bound, Upper bound]
Apr 1983	[Forecasted value]	[Lower bound, Upper bound]
May 1983	[Forecasted value]	[Lower bound, Upper bound]
Jun 1983	[Forecasted value]	[Lower bound, Upper bound]
Jul 1983	[Forecasted value]	[Lower bound, Upper bound]

These forecasts will serve as valuable insights for strategic decision-making, aiding in inventory management, resource allocation, and sales planning. However, it's important to interpret the forecasts alongside the associated uncertainty captured by the confidence intervals. Further monitoring and refinement of the model may be necessary to enhance forecasting accuracy and robustness.

7. Actionable Insights & Recommendations
- Conclude with the key takeaways (actionable insights and recommendations) for the business

Based on the comprehensive analysis conducted on the historical sales data of Sparkling and Rose wines for **ABC Estate Wines**, here are the actionable insights and recommendations:

Understanding Market Dynamics:

- **Market Demand:** Sparkling wine consistently exhibits higher average monthly sales volumes compared to Rose wine, indicating a stronger market demand for Sparkling wine.
- **Sales Variability:** Despite lower average sales volumes, Rose wine demonstrates relatively consistent sales patterns within a narrower range of values compared to Sparkling wine.
- **Seasonal Trends:** Seasonal decomposition reveals recurring patterns and fluctuations in sales for both wine types, providing insights into seasonal demand trends.

Forecasting Models Evaluation:

- **Model Performance:** Various forecasting models, including Linear Regression, Simple Average, Moving Average, and Exponential Models, were evaluated for their performance in predicting wine sales.
- **Best Models:** The Manual SARIMA model performed the best for both Sparkling and Rose wines based on metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

Forecasting and Planning:

- **Utilize SARIMA Models:** Rebuilding the best-performing Manual SARIMA models using the entire dataset enables ABC Estate Wines to make accurate sales forecasts for the next 12 months.
- **Strategic Decision-Making:** The forecasts provide valuable insights for strategic decision-making, including inventory management, resource allocation, and sales planning.
- **Consider Uncertainty:** Interpret forecasts alongside confidence intervals to understand the associated uncertainty and refine strategies accordingly.

Continuous Monitoring and Refinement:

- **Model Refinement:** Continuous monitoring and refinement of forecasting models are essential to enhance accuracy and robustness over time.
- **Data Quality:** Ensure data integrity and quality by regularly reviewing and updating datasets to reflect changing market dynamics and trends.

Marketing and Promotion Strategies:

- **Targeted Marketing:** Tailor marketing and promotion strategies based on insights into market demand and seasonal trends for Sparkling and Rose wines.
- **Product Differentiation:** Identify opportunities for product differentiation and innovation to meet evolving consumer preferences and stay competitive in the wine market.

Customer Engagement and Feedback:

- **Customer Feedback:** Seek feedback from customers to understand preferences, expectations, and satisfaction levels regarding Sparkling and Rose wines.
- **Engagement Strategies:** Implement customer engagement strategies such as loyalty programs, tastings, and events to foster brand loyalty and increase customer retention.

Collaboration and Partnerships:

- **Industry Collaboration:** Explore collaboration opportunities with industry partners, distributors, and retailers to expand market reach and distribution channels.
- **Partnership Initiatives:** Collaborate with other businesses or organizations for joint marketing campaigns, events, or promotions to enhance brand visibility and attract new customers.