Week 2 Assignments

1. First assignment (40 Points)
   1. Open Missouri Beer Production.csv file.
   2. Plot the dataset. Do you see any trend and seasonality? Very briefly Explain.

A graph of beer production

Description automatically generated

* Yes, there is seasonality for every year in the month December there is increase in sales.
* There is a trend showing the increase in sales every year.
  1. Create a timeseries data using R ts() function with seasonal frequency 12 (months in a year) and another timeseries data with seasonal frequency of 4 (quarters in a year) . Compare these graph and write down, your understanding.

A graph of different types of beer production

Description automatically generated

* 1. Choose the correct timeseries and partition the timeseries into training and validation. Consider 75% and 25%.
  2. Forecast validation with naïve, snaive, and trailing moving average. Use W = 5 for MA.
  3. Which one is a better model.

A computer screen shot of numbers and letters

Description automatically generated

* 1. Forecast January 24 with naïve, snaive, and moving average.

A white background with black text

Description automatically generated

* 1. Put training data, validation, and all forecasted data in one plot.

A graph with different colored lines

Description automatically generated

* 1. Get the RMSE, MAE, and MAPE. Which model gives a better performance.

| Model | RMSE | MAE | MAPE |
| --- | --- | --- | --- |
| Naive Model | 1431.32 | 1224.60 | 28.28 |
| Seasonal Naive | 633.91 | 510.13 | 12.34 |
| Trailing Moving Average | 1060.75 | 797.64 | 18.66 |

* Lower values of RMSE, MAE, and MAPE indicate better performance. Based on these metrics, the Seasonal Naive model appears to perform better than the other models.

1. Second assignment (Use xl to do the following work. R work won’t be accepted) (30 Points)
   1. Open DepartmentSales.xlsx file
   2. Plot the dataset. Do you see any trend and seasonality? Very briefly Explain.

Upon examining the sales data provided, it's evident that there is both a trend and seasonality present.

***Trend:***

* There is a clear upward trend in sales over the years. The sales figures consistently increase from Q1-2019 to Q4-2023.
* This suggests that overall, the business is experiencing growth in sales over time.

***Seasonality:***

* There is a repeating pattern of higher sales in the fourth quarter (Q4) of each year.
* This indicates a seasonal trend where sales peak towards the end of each calendar year.
* The sales figures for Q4 consistently surpass those of the other quarters every year.

The data suggests a positive trend of increasing sales over time, coupled with a seasonal pattern of higher sales in the fourth quarter of each year. This combination of trend and seasonality indicates a healthy and cyclical pattern in the sales data.

* 1. Use Trailing and Center moving average and build for up to “Q4-2022”. Use W = 3 for MA.
  2. Plot dataset, Center MA and Trailing MA together. Very briefly explain the results.

***Interpretation:***

* The Trailing Moving Average (TMA) captures short-term trends, showing an increase in sales every fourth quarter. Centered Moving Average (CMA) smooths fluctuations, revealing the overall sales trend and highlighting quarterly decreases. TMA excels in short-term forecasting, while CMA emphasizes long-term trend analysis, aiding in decision-making.
* Trailing Moving Average (TMA) and Centered Moving Average (CMA) values are lower compared to original sales figures as they represent averages, smoothing out fluctuations. This approach provides a clearer picture of long-term trends but may underestimate peak sales periods.
  1. Forecast for 4 quarters (Q1-2023:Q4-2023) with the training MA, naïve, and snaive.

***Forecasted Results:***

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Sales with TMA | Sales with Naïve | Sales with Snaive |
| Q1-2023 | 68325 | 85175 | 51640 |
| Q2-2023 | 68325 | 85175 | 54119 |
| Q3-2023 | 68325 | 85175 | 65681 |
| Q4-2023 | 68325 | 85175 | 85175 |

* 1. Plot dataset from Q1-2019 to Q4-2023 with MA, naïve, snaive. Explain the results.
     + Naïve sales remain constant as it relies on the most recent observed value.
     + Sales with TMA reflect an averaged representation of the data.
     + Seasonal Naïve sales reflect recent seasonal patterns by considering the most recent season's value.
  2. Use the three models and forecast for last 2 quarters of 2024.

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Sales with TMA | Sales with Naïve | Sales with Snaive |
| Q3-2024 | 74566.67 | 92183 | 71486 |
| Q4-2024 | 74566.67 | 92183 | 92183 |