Martin Ngoh

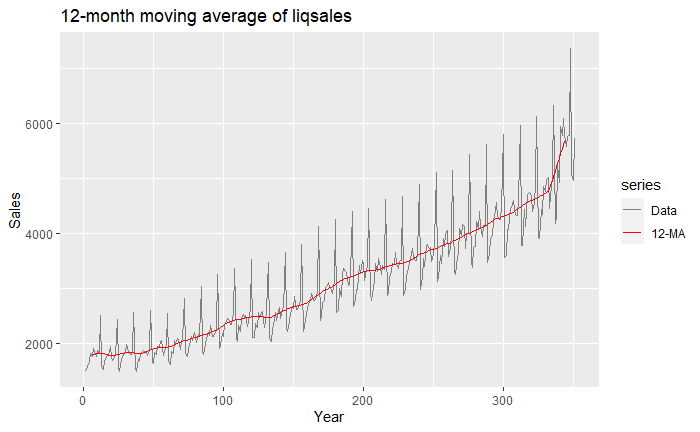
Marketing Project 1

Forecasting the Model

**Forecasting**

The first step in forecasting the model is to consider the Naïve Bayes forecasting method to set a baseline moving forward. We ran the model for the next 24 months and found that the p-value was significant showing that residuals are white noise and this is the incorrect model to use. It is important to consider this mode …

The next step in forecasting the model is too smooth using 12-month moving averages to uncover the trends in the data with minimal noise. **Plot xx** shows the moving average along with the trend in the data. We can see that the trend in liquor sales over time is increasing and has especially increased in the past 2 years. Smoothing the model this way allows for us to see the trends in the data without any seasonal effect.

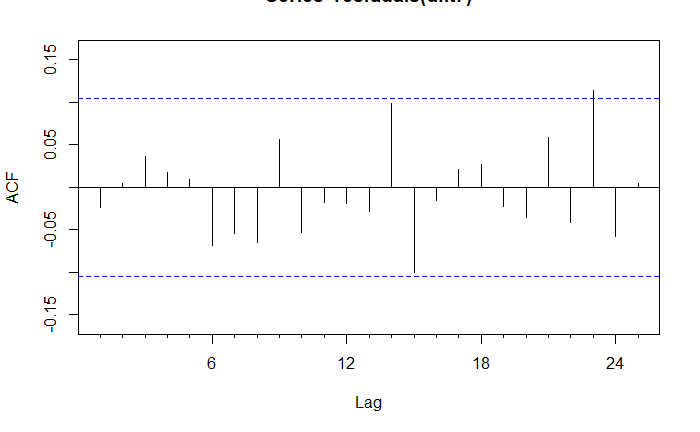
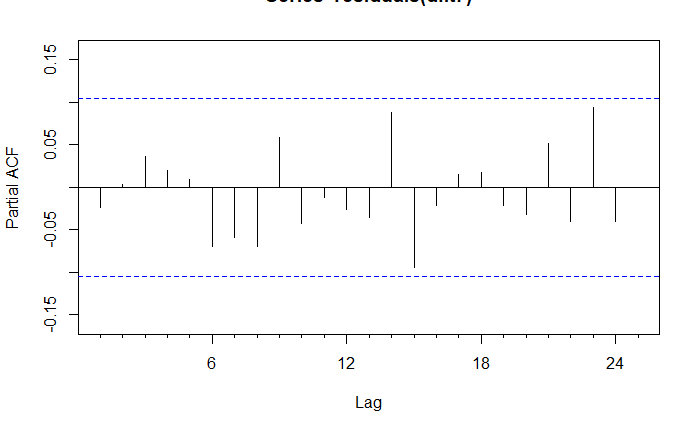


Next, we ran an Augmented Dickey-Fuller Test to verify whether or not our model was stationary. After running the model, we obtained a p-value of 0.04959 (.05 rounded by 3 decimal), suggesting our model is not stationary. Since the model is not stationary, we must use difference order starting with 1 differencing order. When using 1 order differencing, we saw values become negative in the partial ACF plot suggesting there is a moving average component to this model that should be added. To find the correct order differencing we ran 3 iterations of order differencing and found that are 3rd iteration yielded the lowest standard deviation and AIC suggesting this is the correct order differencing (d= 1 and D = 1). We then ran three iterations of the models with various combinations. For the first model we added an MA component to the non-seasonal only. The 2nd model we added model we added one-order MA component to the seasonal part only. The 3rd model combined a MA component to the non-seasonal and one to the seasonal side. **TABLE XX** shows the standard deviations and AIC of the 3 iterations of the models.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Standard Deviation | AIC | MAPE |
| 1 (One MA to non-seasonal) | 23394 | 4364.24 | 2.9614 |
| 2 (One MA to seasonal) | 28059 | 4430.99 | 3.7254 |
| 3 (One MA to both) | 18762 | 4296.96 | 2.9214 |

The table shows that we were able to reduce the standard deviation and AIC to the lowest values using model 3. To determine model 3 significance, we look at the MA significance for seasonal and non-seasonal which are confirmed to be statistically significant. **Plot x and Y below** show the ACF and Partial ACF plots for the residuals when running the 3rd iteration of the model. We then used the auto.arima function to compare our results with the auto generated results from the function. The results showed

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Standard Deviation | AIC | MAPE |
| Auto Arima | 18691 | 4294.36 | 2.8591 |
| 3 (One MA to both) | 18762 | 4296.96 | 2.9214 |

The auto generated model using the auto arima function in R provides a lower standard deviation, AIC, and MAPE value suggesting a better model. This version of the model utilized 1 seasonal MA and 2 non-seasonal MA. Moving forward, the coefficients of the auto arima model are all statistically significant and therefore we can use this model. **PLOT XX BELOW** shows the ACF and partial ACF of the residuals for the auto generated model.

The plots show that essentially all residuals fall within the non-significance level of the plot.

LJUNG BOX TEST OF RESIDUALS INSET HERE

FORECASTING + MAPE STATS HERE