**Moving Average (MA)**

In the moving average section, we tried to take an average of the data within a time window. We did trial and error to split the time window to 4, 6, 12, and 18 weeks to select the best time window aka order, for the moving average. In the graph illustrated below, we can see that Moving Average 6 (MA6) has a better fit on the *WeeklySales* training set compared the rest of the three graphs. As result, we would go with MA6 for deeper analysis. With sufficient time, we also analyzed MA 4, MA12, and MA18 to ensure that our choice of MA is right.

Chart

Description automatically generated

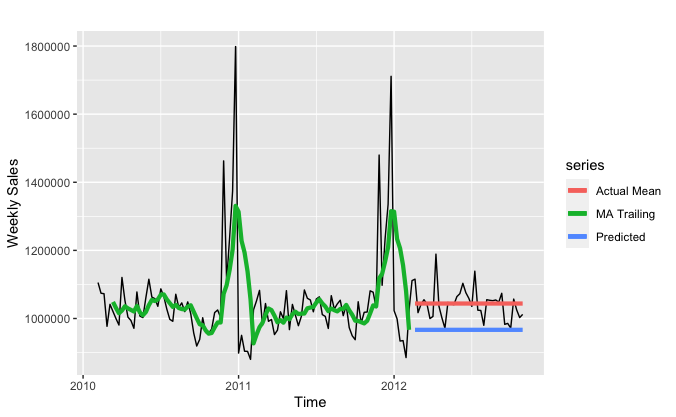
We further investigated MA 6 by applying **Trailing Moving Average**. After applied *rollmean* to the training data and set k=6, the graph is illustrated below:

Chart, histogram

Description automatically generated

**Trailing Average Smoother**

In the trailing average smoother section, we used *k=6* and the last moving average for prediction. Since we are using last moving average to fill in all the 38 prediction weeks (nValid), the prediction period on the graphs shows to be a straight line illustrated below. This is a naïve way to predict Walmart’s weekly sales for the next 38 weeks since it only uses the last moving average.



To better make prediction for the following 38 weeks of store sales, we improved the trailing model by employing a for loop to append last moving average of not only the last moving average, but the rolling average for better prediction. We tested different time window: *k = 4, k = 6, k=12, and k=18* to see which time window would have a better fit and accuracy with the testing set.

Graphical user interface, chart, application, histogram

Description automatically generated

As graph showed below, we can see that graph *Trailing 4* and *Trailing 6* has better fit for the rolling line and test line; however, when we look at the accuracy score, we can see that trailing moving average with 12 weeks window has the lowest root mean square deviation （RMSE）and mean absolute percentage error (MAPE). As show in the table below, trailing moving average has a RMSE of 50593.8 and MAPE of 3.359, which is the lowest among the four.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **ME** | **RMSE** | **MAE** | **MPE** | **MAPE** | **ACF1** | **Theil’s U** |
| **K = 4** | -3996.336 | 57580.03 | 42730.13 | -0.4816351 | 4.154352 | 0.2158959 | 2.842983 |
| **K = 6** | -5056.532 | 57088.12 | 42368.48 | -0.5701634 | 4.105455 | 0.23706 | 3.894264 |
| **K = 12** | 10326.97 | **50593.8** | 35462.27 | 0.9234055 | **3.359151** | 0.2333883 | 4.008283 |
| **K = 18** | 19367.91 | 56019.75 | 43931.72 | 1.754249 | 4.08718 | 0.2797184 | 5.85484 |

Conclusion: the best trailing moving average model is with time window of 12 weeks. It has RMSE of 50593.8 and MAPE of 3.359151.

**Simple Exponential Smoothing (SES)**

Simple Exponential Smoothing is a popular forecasting method. It is similar from moving average, but instead of taking a simple average over a time window, it weights average of all past values, so that the weights decrease exponentially into the past. We are using seasonality as a testing stage to see how the data would react. We are using the artificial neural network (ANN) model to forecast the predicted store sales. We used lag=1 to look for the data for previous week.

Chart

Description automatically generated

With the ANN model accuracy, the RMSE is 60904.42 and MAPE is 191.3546.

Holt’s Linear Trend Model

When we put the data into Holt’s Linear Trend Model (ANN) to forecast 38 weeks of store sale, we can see in the graph that the predicted value is a line. Compared to the actual test set, the predicted line is not so accurate. Since exponential smoothing does not take seasonality into consideration, it is not very accurate. It has a RMSE of 46881.7 and MAPE of 3.284567.

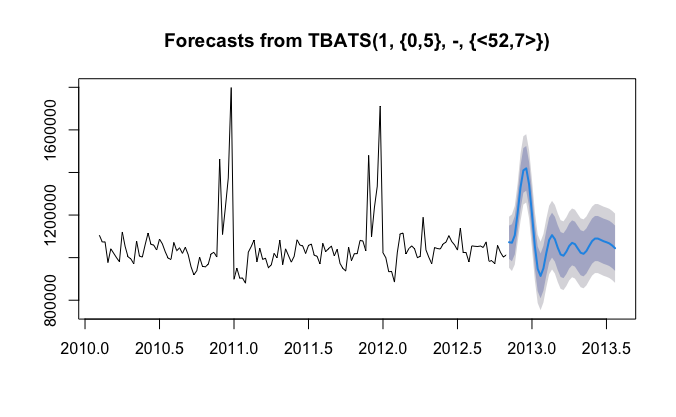
Chart, histogram

Description automatically generated

In conclusion for simple exponential smoothing, it is not a very good method to use to forecast or predict future values. Simple exponential smoothing does not take into consideration for trend and seasonality. In our graph, there are obvious seasonality at the end of the year of 2011 and 2012. So simple exponential smoothing is not a good predictor method for the Walmart dataset.

**TBATS Model**

One trivial model we used is TBATS model. TBATS model has the capability to deal with complex seasonality with no seasonality constraints. TBATS is an acronym for key features of the model. **T**: Trigonometric seasonality, **B**: Box-Cox transformation, **A**: ARIMA errors, **T**: Trend, **S**: seasonal components.



As illustrated in the graph above, the forecast seems very reasonable. TBATS model takes it roots in exponential smoothing methods. In the graph, 1 is the Box-Cox parameter, and {0,5} is the ARMA (0,5) model. <52,7> is the seasonality length and Fourier series.