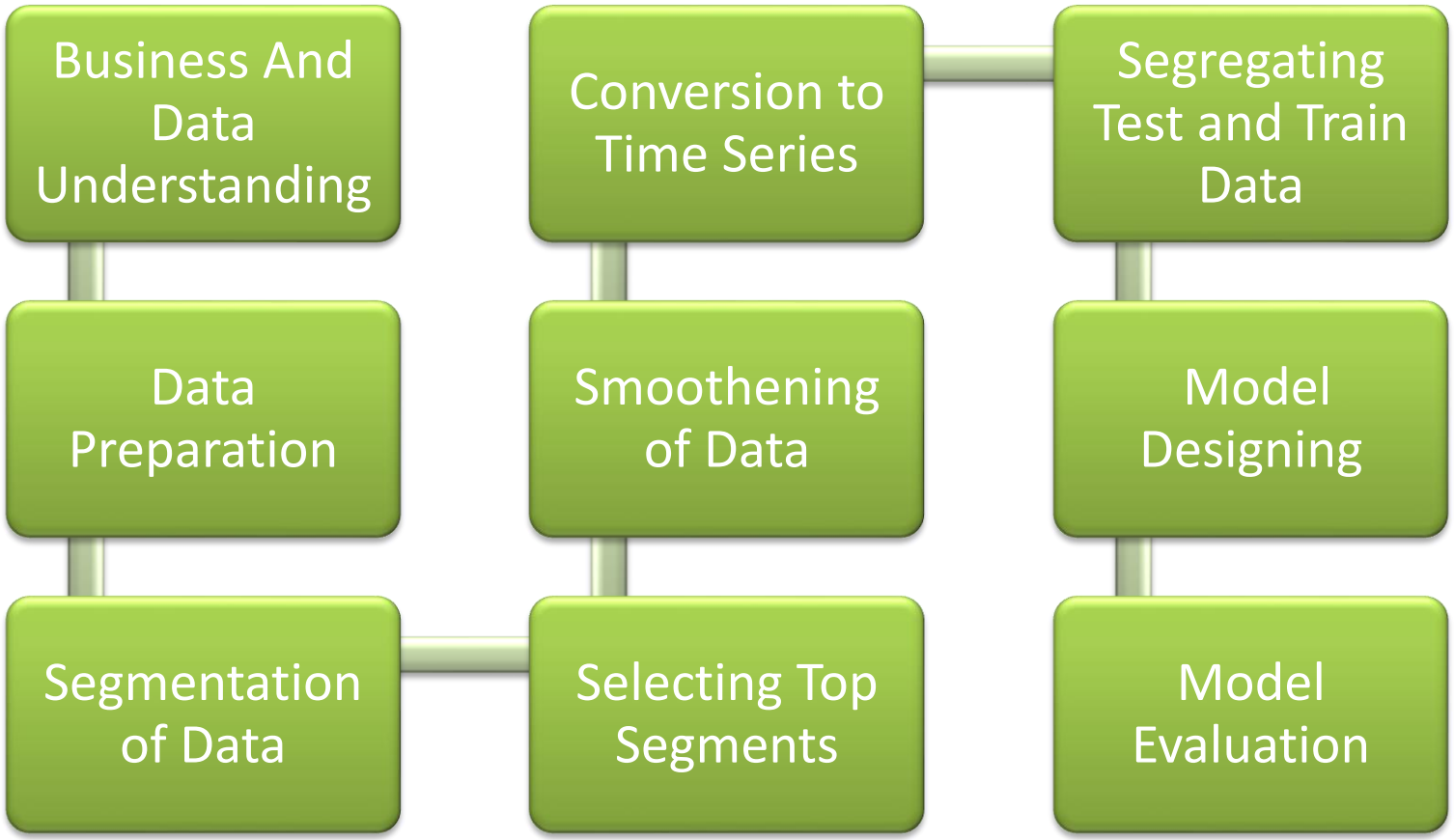
Retail-Giant Sales Forecasting Case Study

Shravani Karra

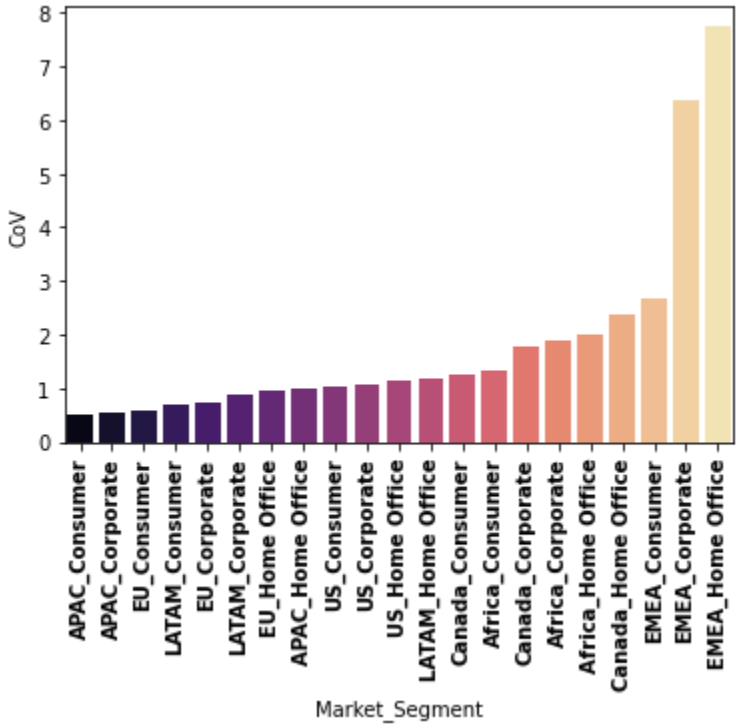


Problem Solving

Selection Of Top Two Buckets

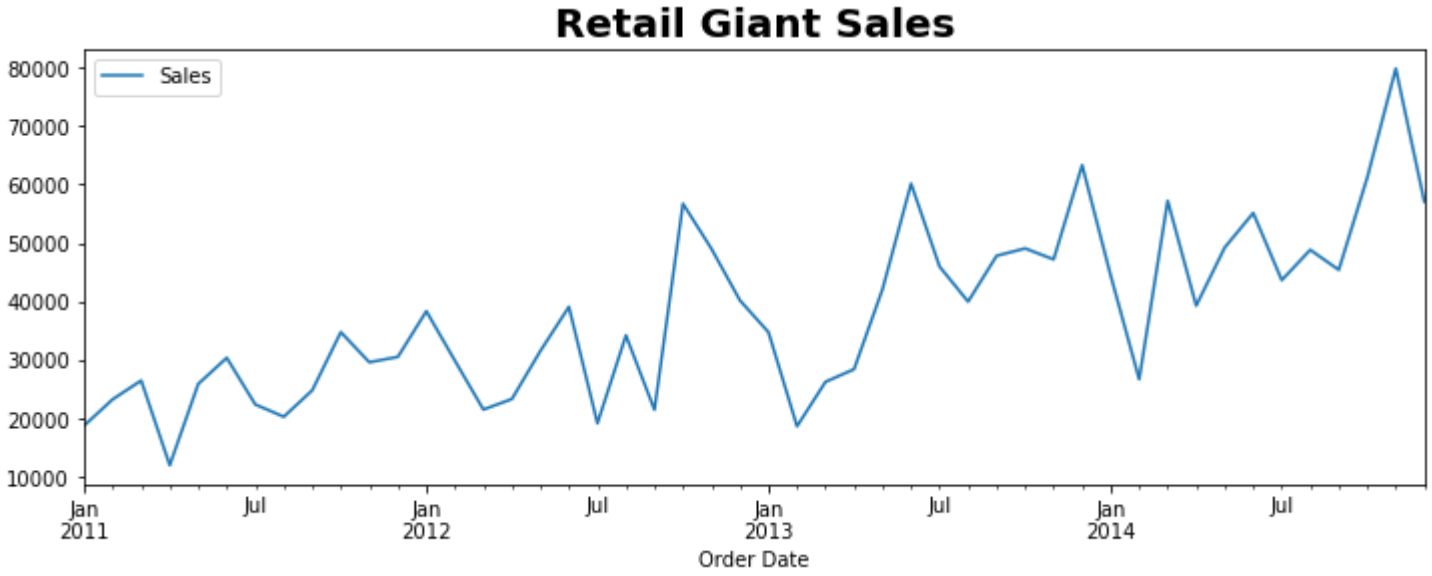
Calculate the CoV on the profit for each of the 21 market segments on the train data

•We compare the variance between the segments using the coefficient of variation which will normalise the standard deviation with the mean and give a comparative figure on the basis of which we can identify the most profitable market segment.
•We want to forecast the sales where the market segment is reliable or in other words, there is less variation in the profits.

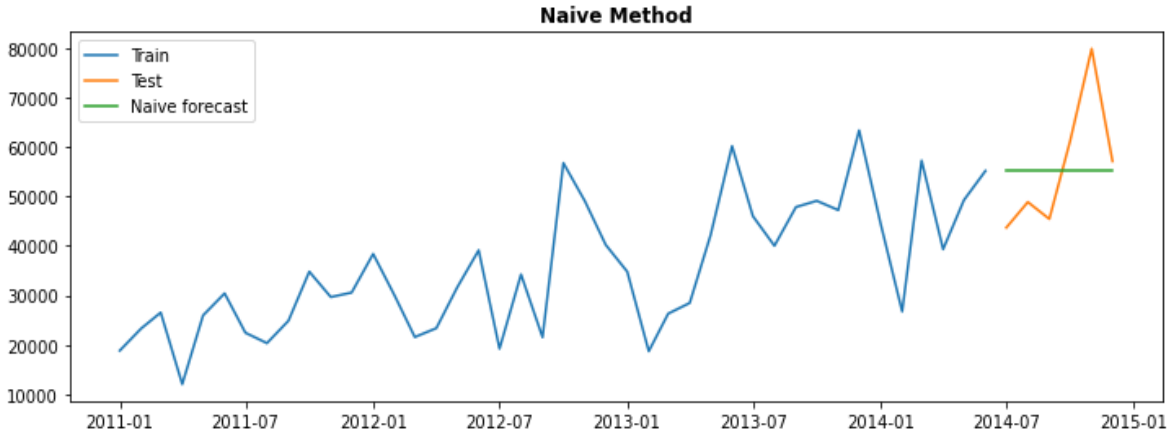


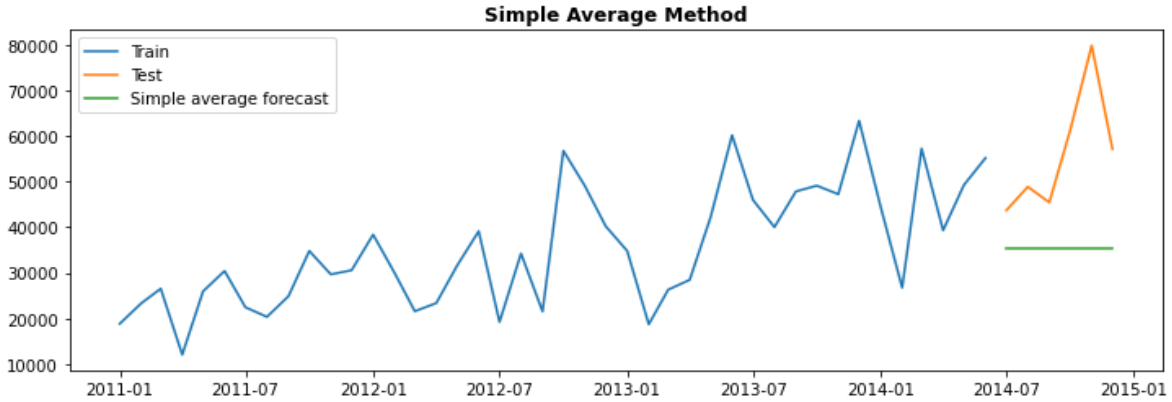
•We can see that Lowest CoV is 0.52272
•We can infer that the corresponding Marget Segment is "APAC_Consumer"
•The most profitable Market Segment is APAC_Consumer

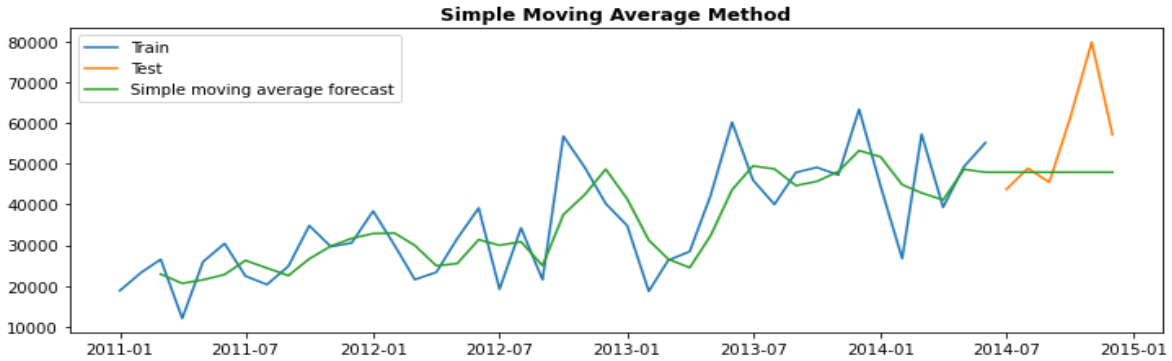
ANALYSIS FOR APAC CONSUMER



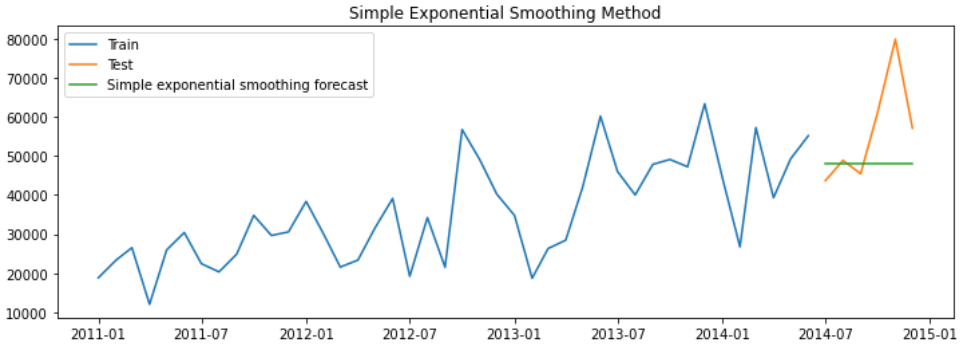
Simple time series methods



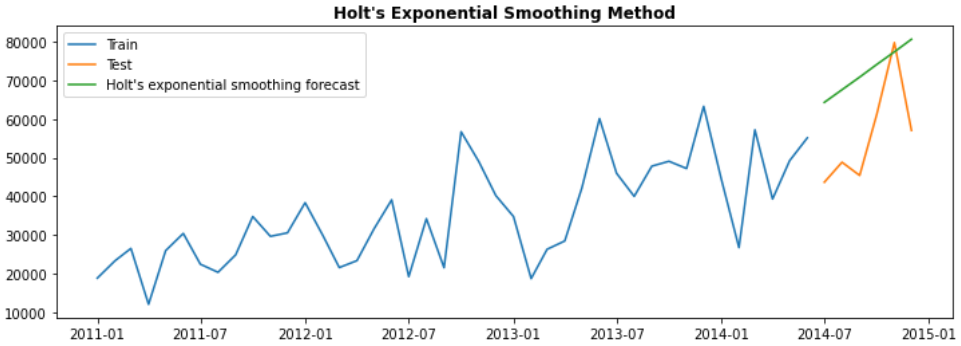


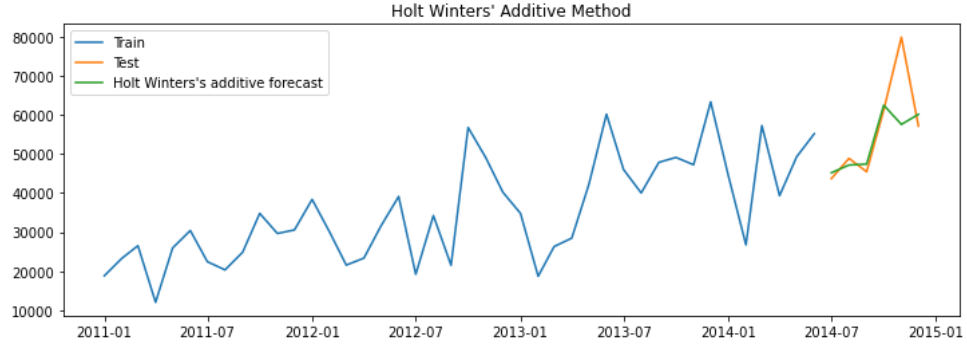


Simple exponential smoothing

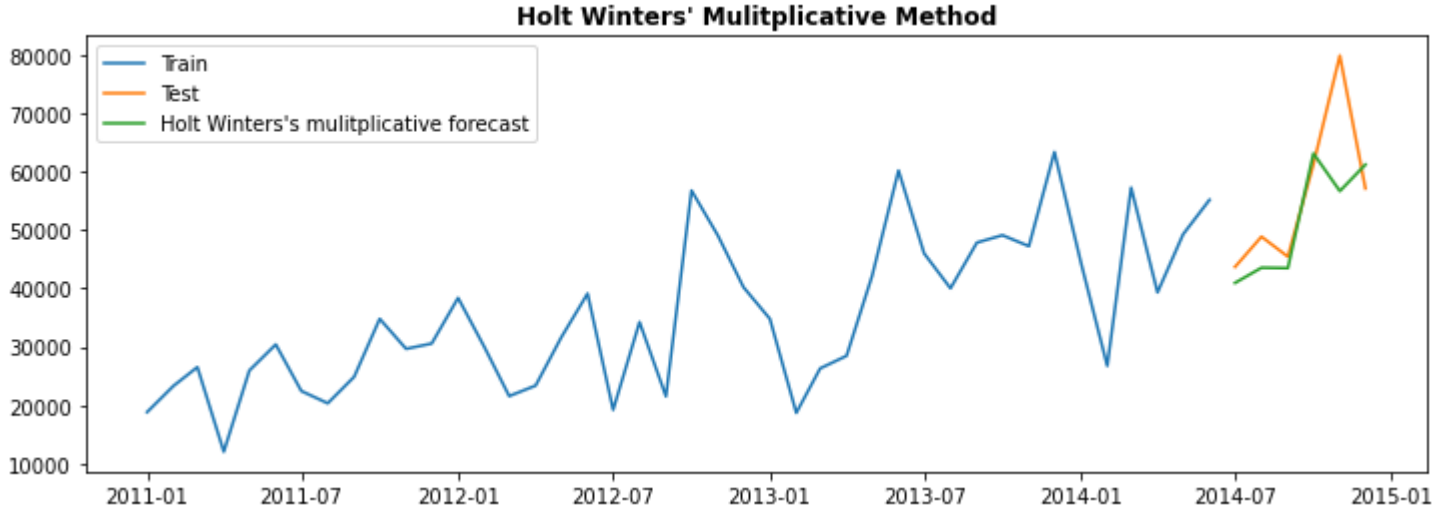


Holt's Exponential Smoothing

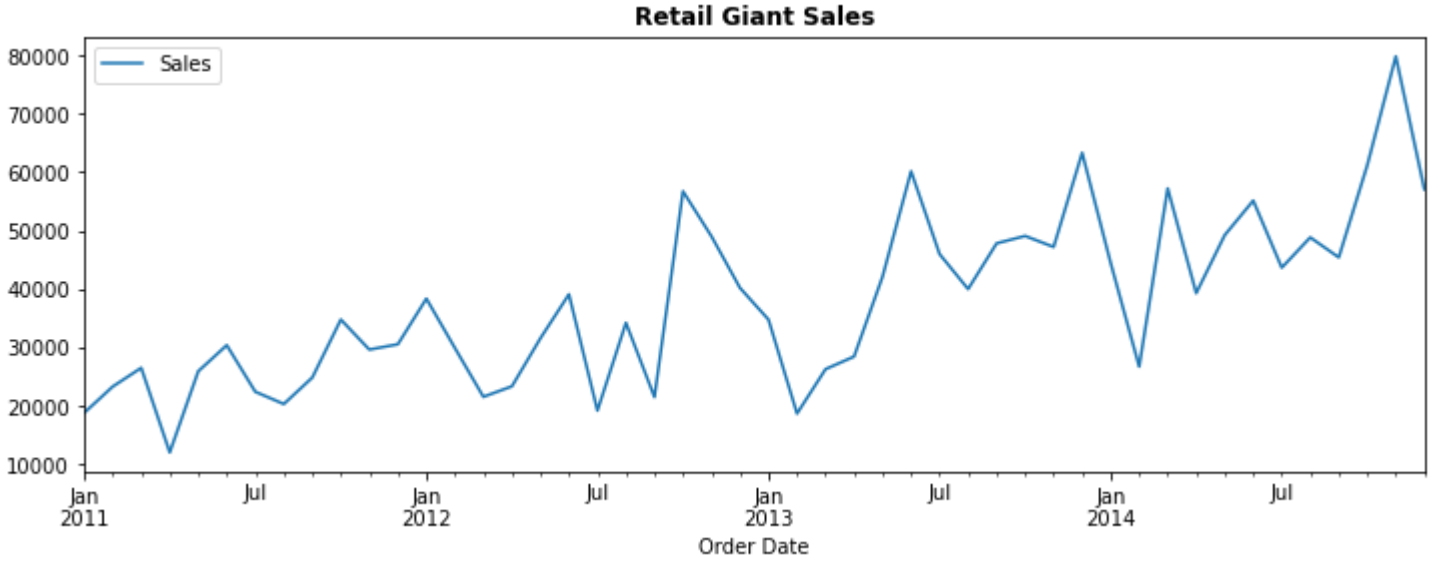




Holt Winter's multiplicative method with trend and seasonality



Stationarity vs non-stationary time series



Augmented Dickey-Fuller (ADF) test

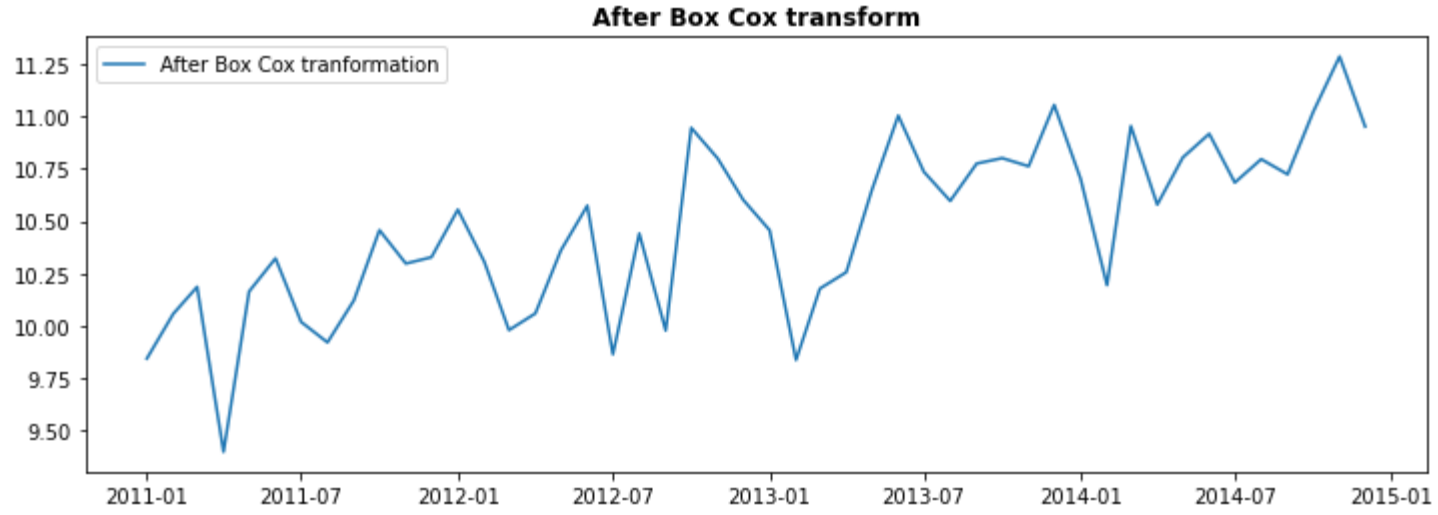
Null Hypothesis $(H0)$(H0): The series is not stationary $p-value$>0.05p−value>0.05
Alternate Hypothesis:$(H1)$(H1) The series is stationary $p-value$<=0.05

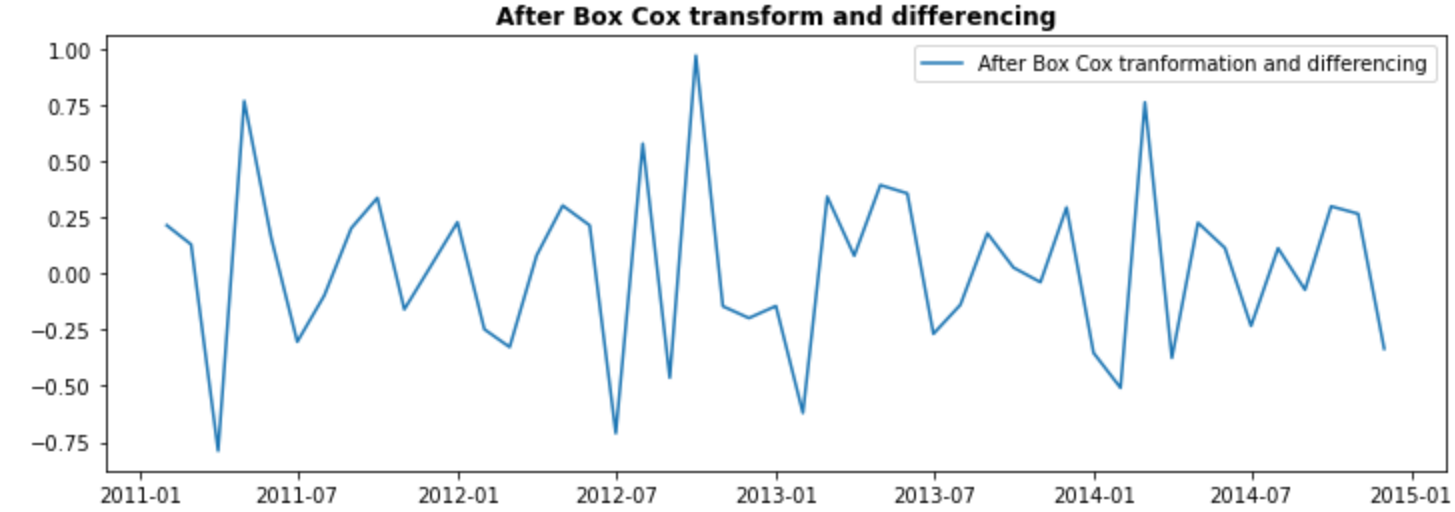
- **We can see that p-value is 0.011, which is less than 0.05**
- **So The series is stationary. And Reject the null hypothesis $(H0)$(H0)**

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test

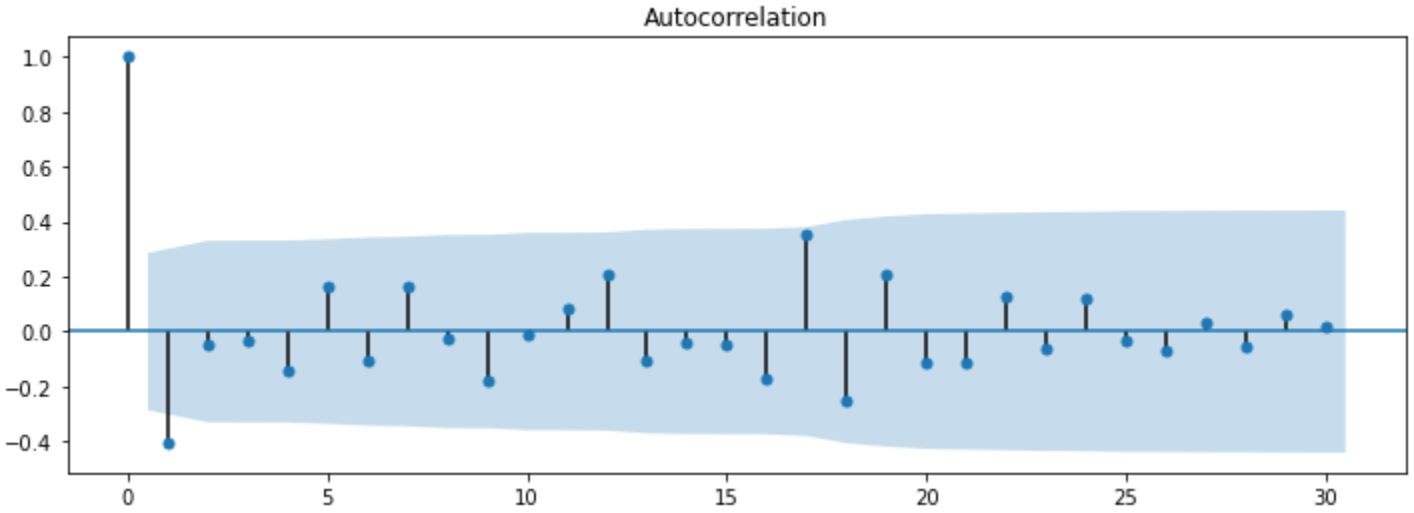
Null Hypothesis $(H0)$(H0): The series is stationary $p-value$>0.05p−value>0.05
Alternate Hypothesis:$(H1)$(H1) The series is not stationary $p-value$<=0.05p−value<=0.05

Box Cox transformation to make variance constant

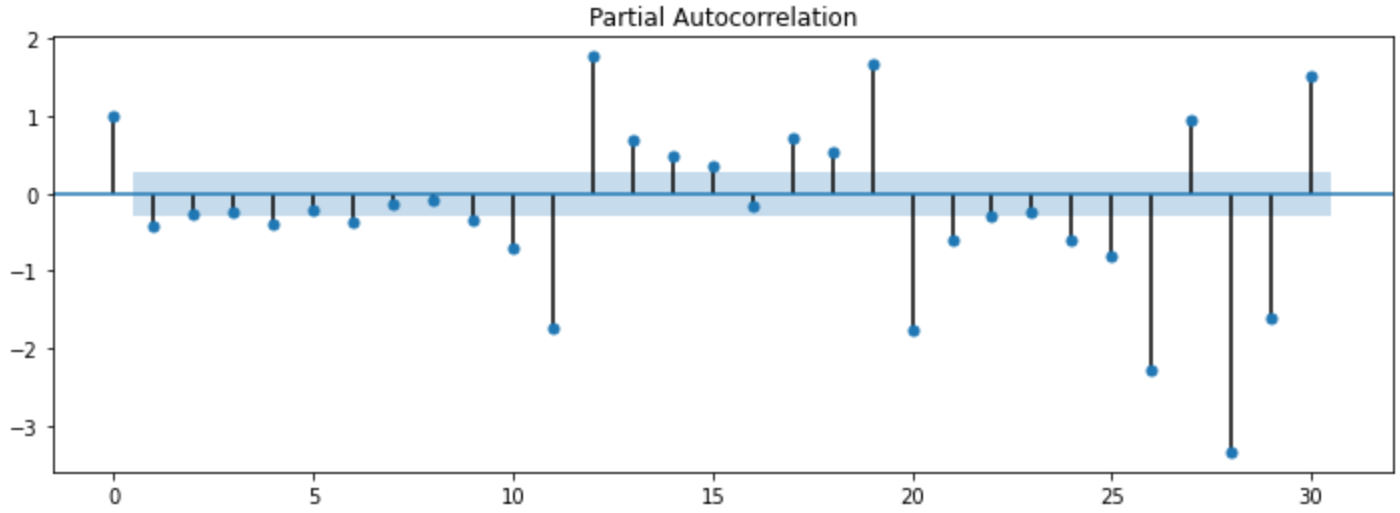


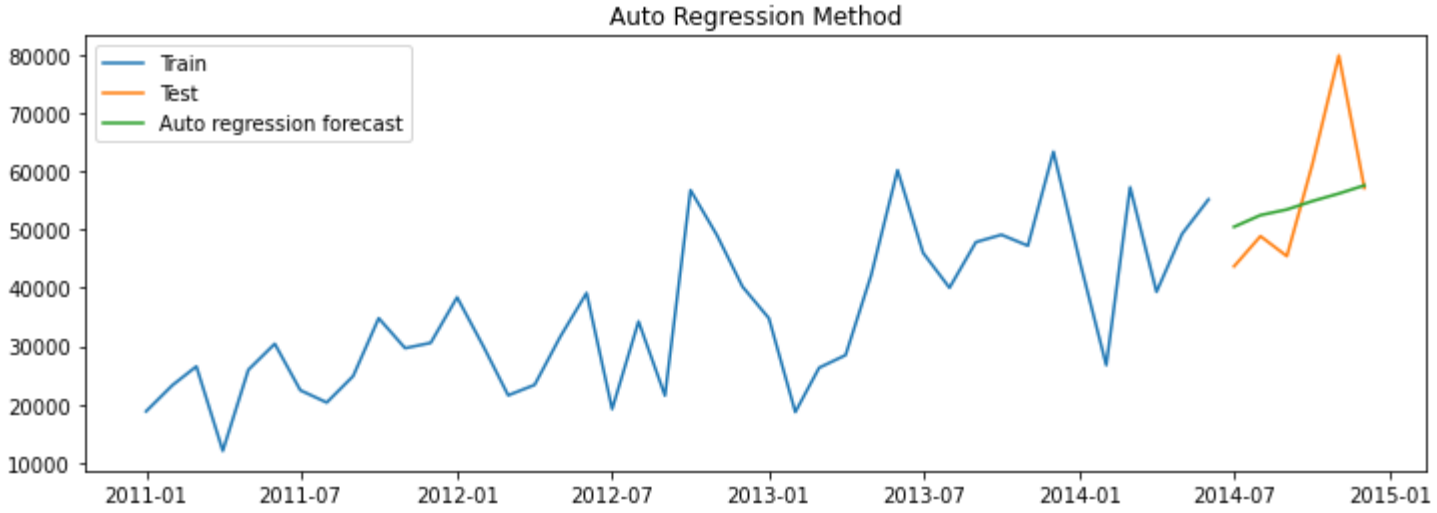


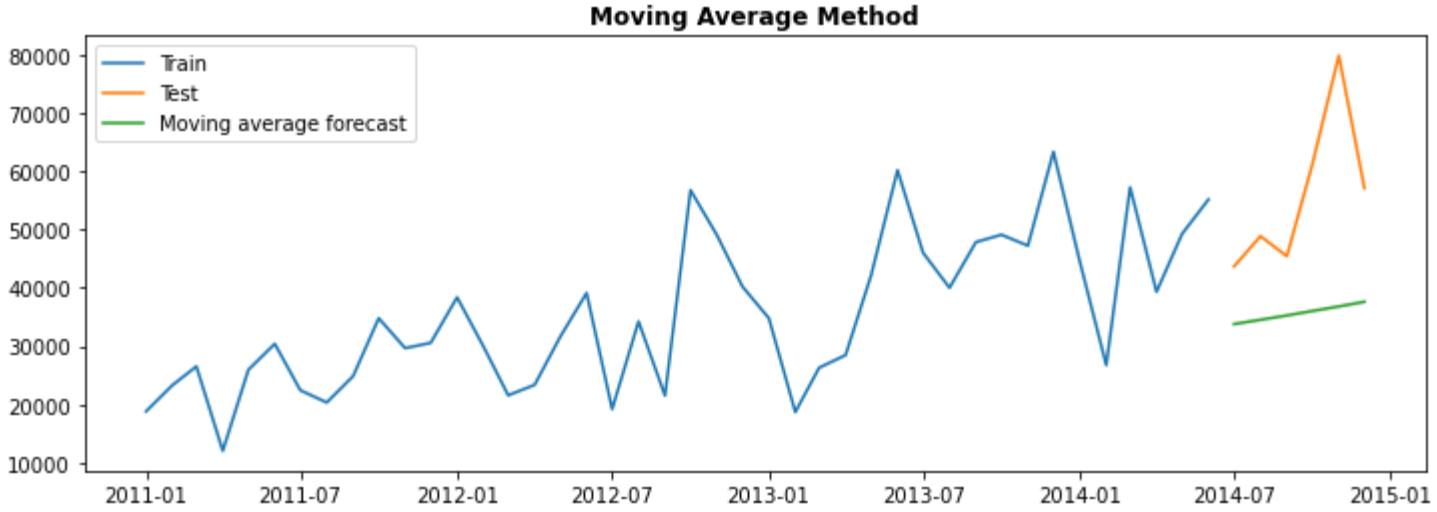
Autocorrelation function (ACF)

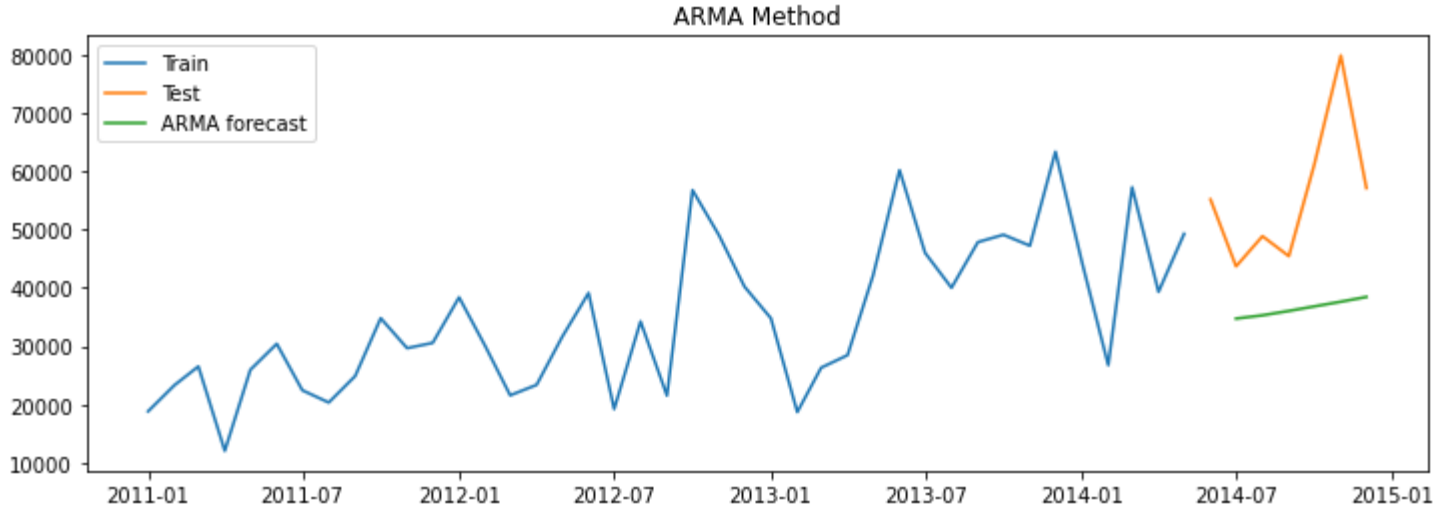


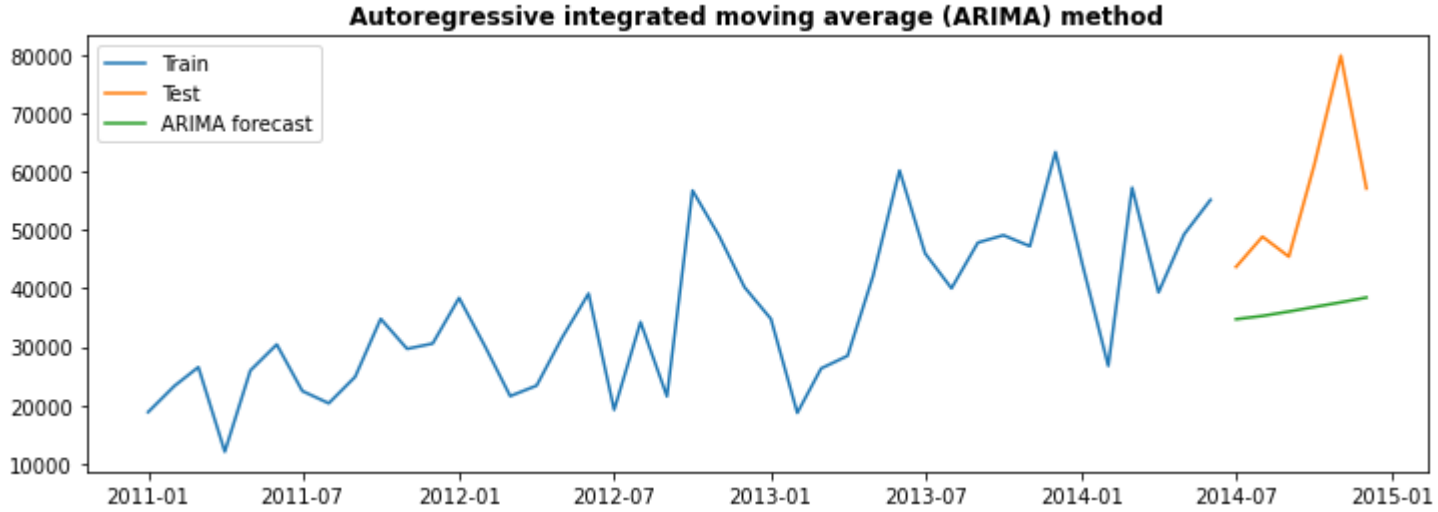
Partial autocorrelation function (PACF)




Auto Regression Method
Train
Test
Auto regression forecast
80000
70000
60000
50000
40000
30000
20000
10000
2011-01
2011-07
2012-01
2012-07
2013-01
2013-07
2014-01
2014-07
2015-01

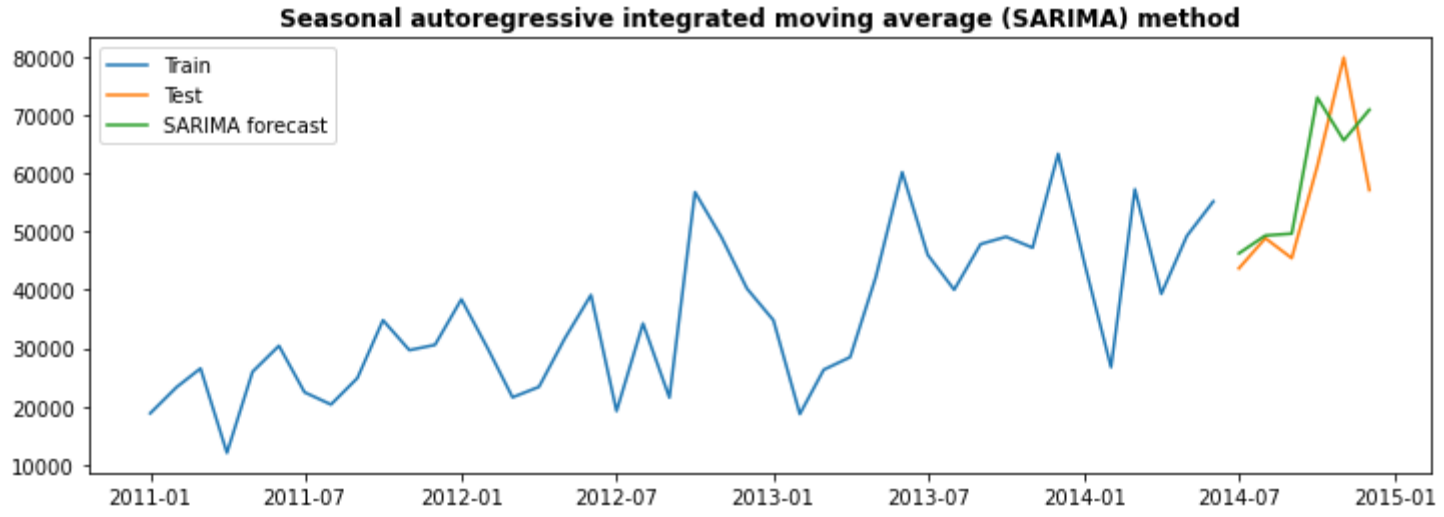

Moving Average Method
Train
Test
Moving average forecast
80000
70000
60000
50000
40000
30000
20000
10000
2011-01
2011-07
2012-01
2012-07
2013-01
2013-07
2014-01
2014-07
2015-01


ARMA Method
Train
Test
ARMA forecast
80000
70000
60000
50000
40000
30000
20000
10000
2011-01
2011-07
2012-01
2012-07
2013-01
2013-07
2014-01
2014-07
2015-01


Autoregressive integrated moving average (ARIMA) method
Train
Test
ARIMA forecast
80000
70000
60000
50000
40000
30000
20000
10000
2011-01
2011-07
2012-01
2012-07
2013-01
2013-07
2014-01
2014-07
2015-01

Seasonal auto regressive integrated moving average (SARIMA)

- SARIMA Model has both non seasonal elements and seasonal elements.
- SARIMA brings all the features of an ARIMA model with an extra feature - seasonality.
- SARIMA has six parameters along with seasonality.



Conclusion

- **Thus we can conclude that, Holt Winters additive method is the best forecasting method in the smoothing technique**

- **And SARIMA - Seasonal Autoregressive Integrated moving average is the best method in ARIMA set of techniques.**