Retail Giant Sales Forecasting Assignment

Submitted by Akshay P

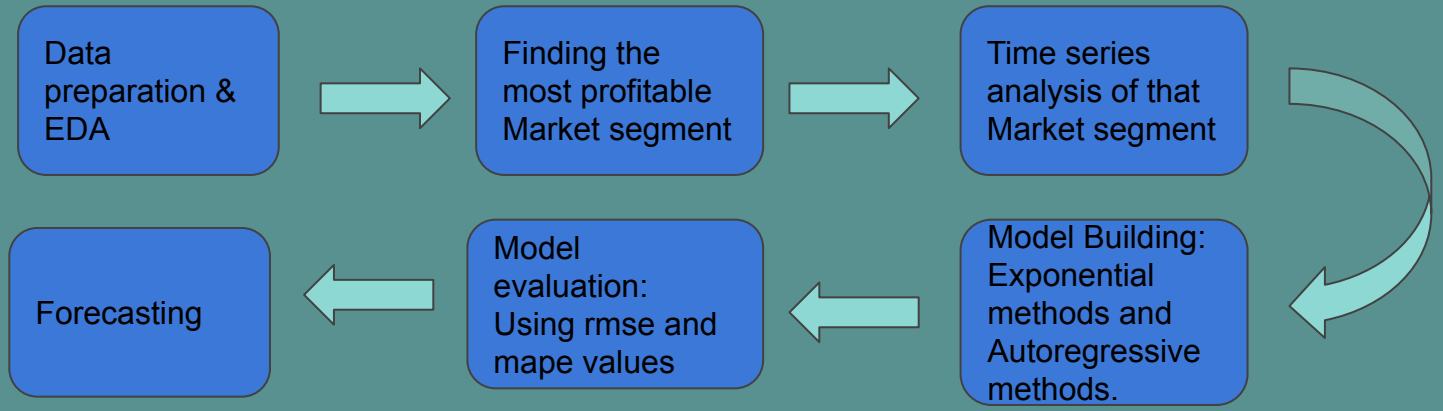
Problem statement

Global Mart is an online supergiant store that has worldwide operations. This store takes orders and delivers across the globe and deals with all the major product categories — consumer, corporate and home office.

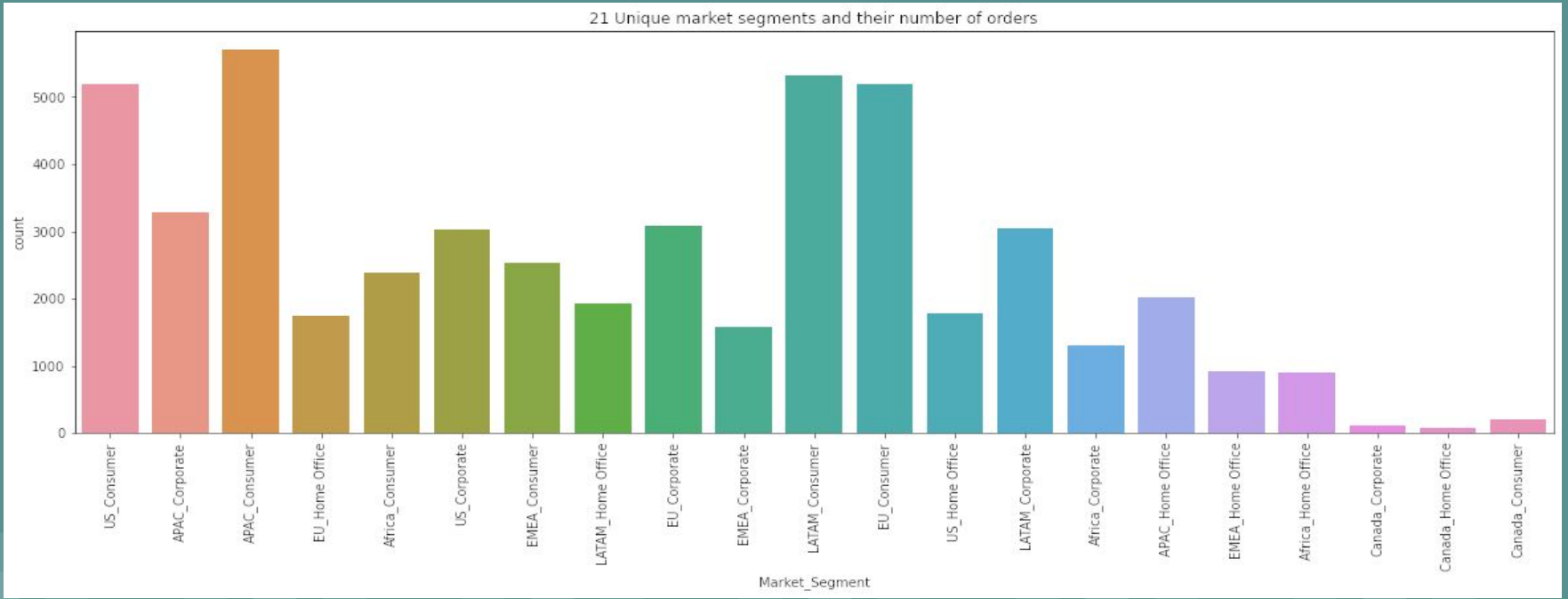
As a sales manager for this store, you have to forecast the sales of the products for the next 6 months, so that you have a proper estimate and can plan your inventory and business processes accordingly.

Objective : find out the most consistently profitable market-segment and forecast the sales of next 6 months for that single market-segment.

Analysis:steps



Number of orders placed in the 21 unique market segments



21 market Segments

US_Consumer	US_Corporate	US_Home Office
APAC_Consumer	APAC_Corporate	APAC_Home Office
Africa_Consumer	Africa_Corporate	Africa_Home Office
EU_Consumer	EU_Corporate	EU_Home Office
EMEA_Consumer	EMEA_Corporate'	EMEA_Home Office
LATAM_Consumer	LATAM_Corporate'	LATAM_Home Office
Canada_Consumer	Canada_Corporate	Canada_Home Office

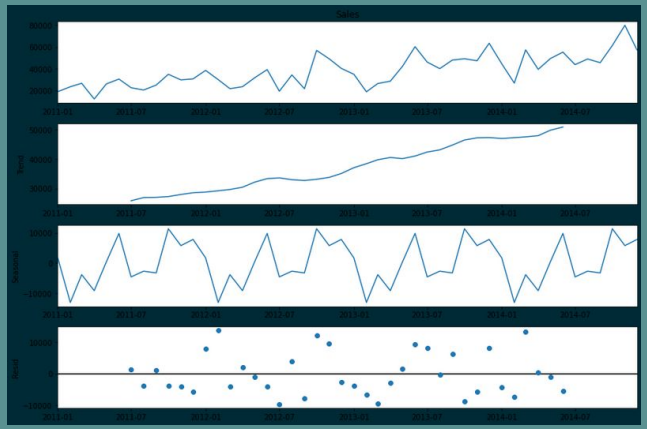
Identifying the most consistently profitable market segment

1. monthly aggregated profit data were created for each market segment
2. From the aggregated transaction data, coefficient of variation(Cov) was calculated for each market segment
3. The most profitable market segment should have the least value of CoV as it is a measure of variation (standard deviation normalised with mean). In other words it relates the standard deviation of the estimate to the value of estimate. Lesser CoV value means, lesser the variation within the profit data and more precise the estimate,so it is more reliable ->consistently profitable.
4. Here ‘APAC_Consumer’ has the least CoV, so it is the most consistently profitable market segment

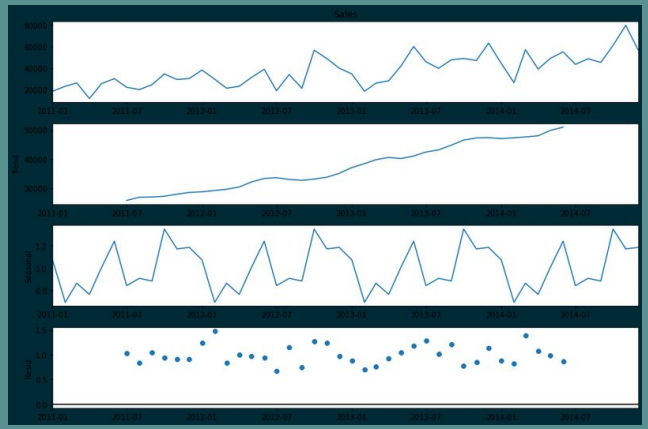
	Market_Segment	mean	std	CoV
0	APAC_Consumer	4400.894243	2300.457687	0.522725
1	APAC_Corporate	2574.919807	1364.837734	0.530051
12	EU_Consumer	3699.977143	2202.282289	0.595215
15	LATAM_Consumer	2295.555697	1569.632686	0.683770
13	EU_Corporate	2216.299429	1600.336696	0.722076
16	LATAM_Corporate	1122.633016	990.360880	0.882177
14	EU_Home Office	1224.456536	1148.627937	0.938072
2	APAC_Home Office	1511.088314	1523.508658	1.008219
18	US_Consumer	2686.740912	2715.031412	1.010530
19	US_Corporate	1754.199083	1880.200775	1.071829
20	US_Home Office	1132.065762	1272.476439	1.124030
17	LATAM_Home Office	818.398941	957.275713	1.169693
6	Canada_Consumer	225.987632	282.555788	1.250315
3	Africa_Consumer	957.707000	1254.932072	1.310351
7	Canada_Corporate	90.980294	162.493114	1.786025
4	Africa_Corporate	412.617571	780.566850	1.891744
5	Africa_Home Office	377.221071	759.322203	2.012937
8	Canada_Home Office	118.003750	279.632866	2.369695
9	EMEA_Consumer	423.960286	1124.552711	2.652495
10	EMEA_Corporate	182.642643	1160.698430	6.355024
11	EMEA_Home Office	84.231366	651.283095	7.732073

Decomposition of APAC_Consmer sales data

Additive decomposition

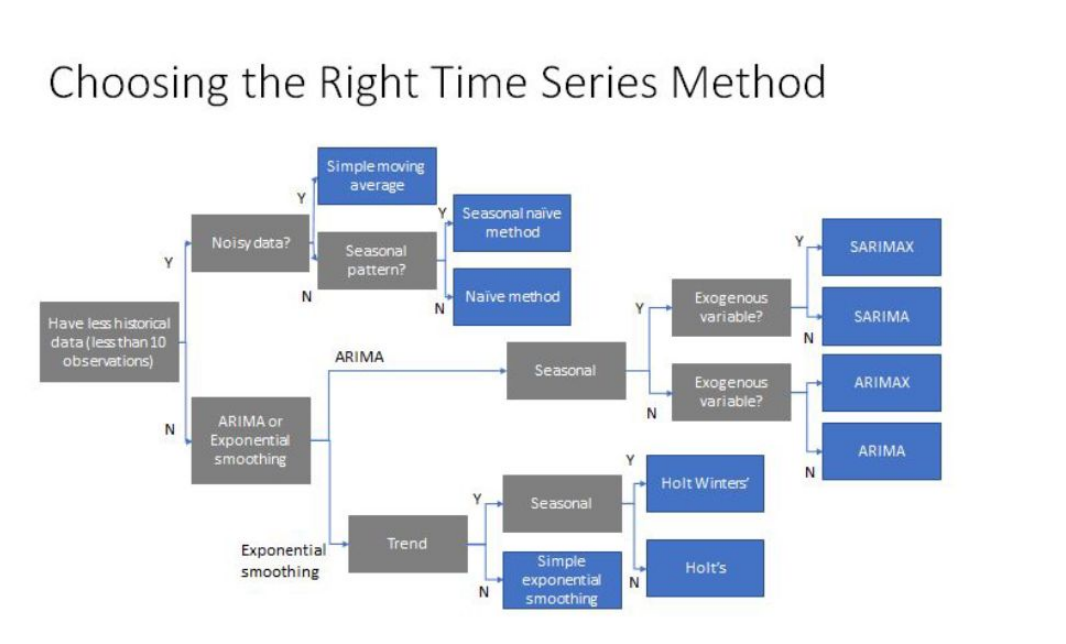


Multiplicative decomposition



In both the decomposition plots, there is a clear upward trend and a yearly seasonality.

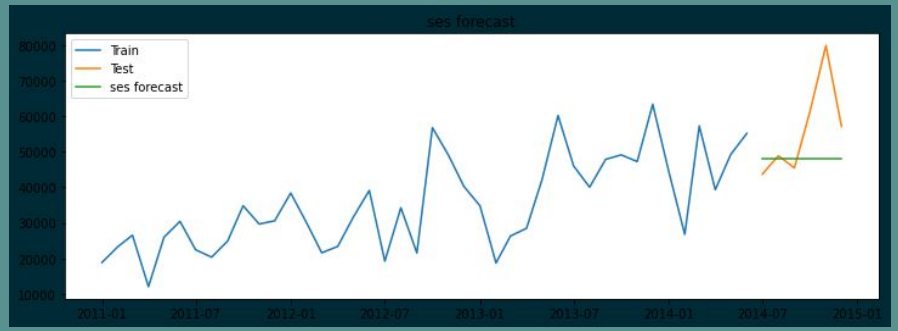
Optimum techniques for sales forecast



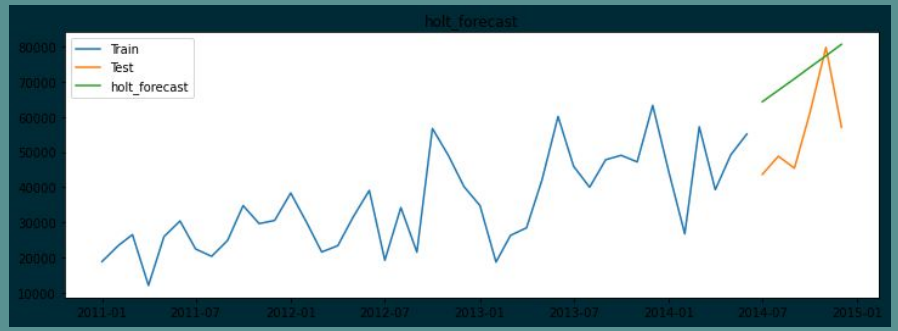
1. We have sales data for 42 months in train set, which is greater than 10 historical observations. So ARIMA set of methods or exponential smoothing methods will suit better in forecasting.
2. From the decomposition plots of sales data it was clear that it has both trend and seasonal components. So SARIMA would be able to forecast better than any other ARIMA methods and Holt Winters' method would be able to forecast better among all the exponential smoothing methods.

smoothing techniques : Simple exponential smoothing and Holt's exponential smoothing

Simple exponential smoothing



Holt's exponential smoothing

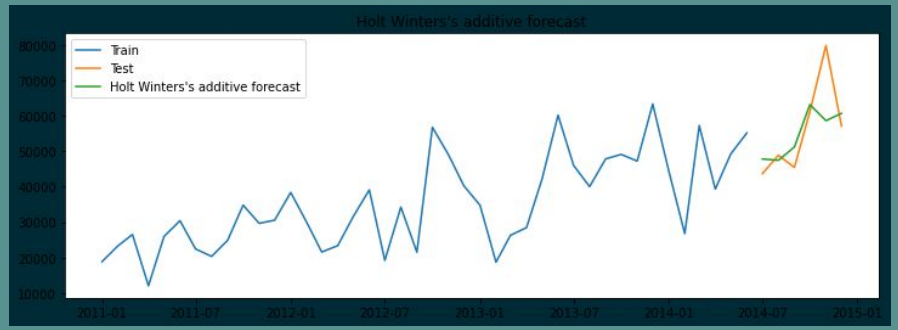


	Method	RMSE	MAPE
0	Simple exponential smoothing	14627.34	15.74
0	Holt's exponential smoothing	18976.37	34.57

Simple exponential model is able to capture only the level of the time series data while Holt's exponential smoothing is able to capture both the level and trend. However simple exponential smoothing has lower rmse and mape values compared to Holt's exponential smoothing, which means simple exponential smoothing is able to forecast better.

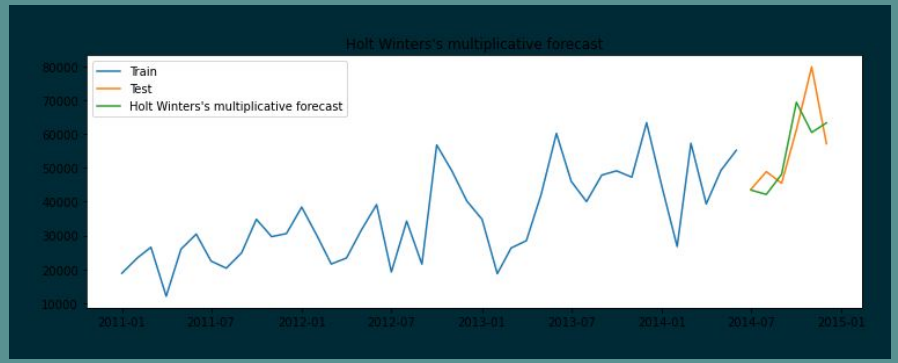
smoothing techniques :Holt-Winters' exponential smoothing - Additive & multiplicative

Holt-Winters' exponential smoothing - Additive



	Method	RMSE	MAPE
0	Simple exponential smoothing	14627.34	15.74
0	Holt's exponential smoothing	18976.37	34.57
0	Holt Winters's additive forecast	9306.82	10.17
0	Holt Winters' multiplicative method	9423.23	11.43

Holt-Winters' exponential smoothing - Multiplicative

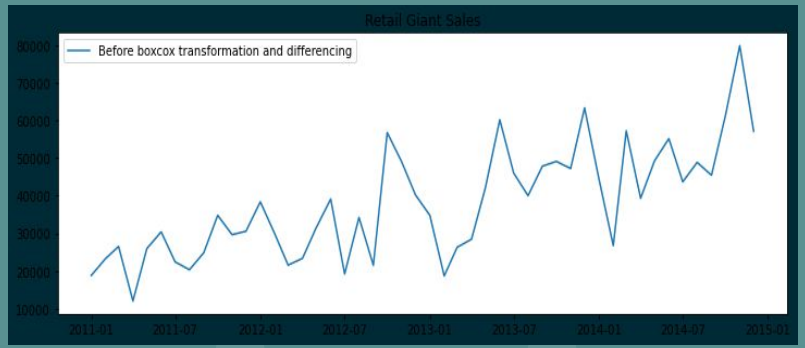


- Both of these methods are able to capture both trend and seasonality.
- Among the various smoothing techniques performed Holt Winters' additive forecast has the least mape and rmse values.

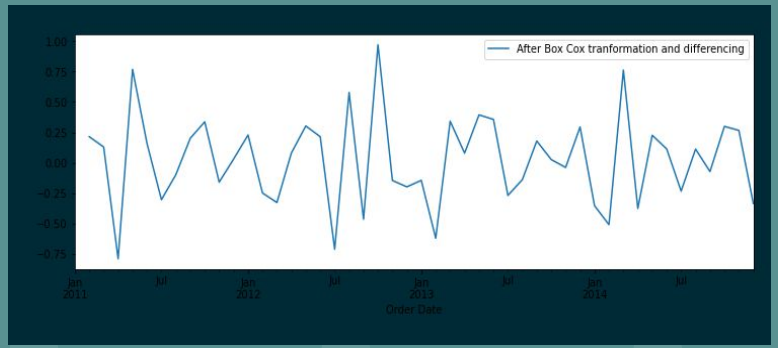
Box cox Transformation and Differencing

- In order to build ARIMA set of models it was needed to make sure that the data is stationary.
- After performing adf and kpss tests, it was clear that the data was not stationary.
- So boxcox transformation and differencing were done to make it stationary.

Sales data before boxcox transformation & differencing

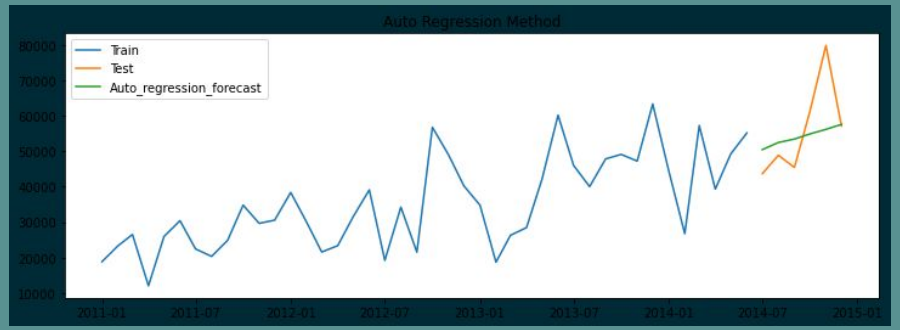


Sales data after boxcox transformation & differencing



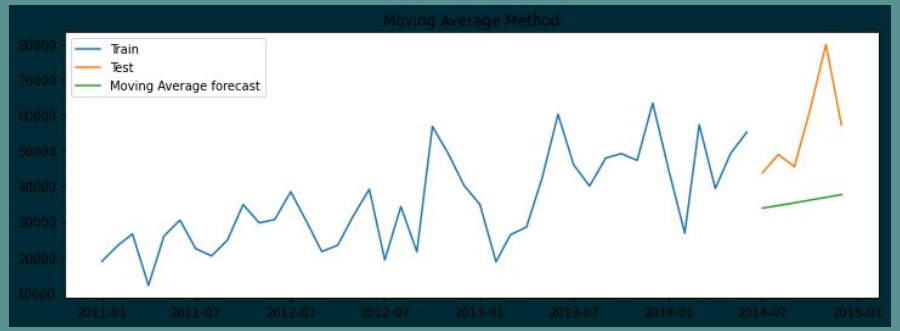
ARIMA set of techniques: Auto regression (AR) model & Moving average method (MA)

AR model - sales forecast



	Method	RMSE	MAPE
0	Simple exponential smoothing	14627.34	15.74
0	Holt's exponential smoothing	18976.37	34.57
0	Holt Winters's additive forecast	9306.82	10.17
0	Holt Winters' multiplicative method	9423.23	11.43
0	Autoregressive (AR) method	10985.28	13.56
0	Moving Average (MA) method	23360.02	33.93

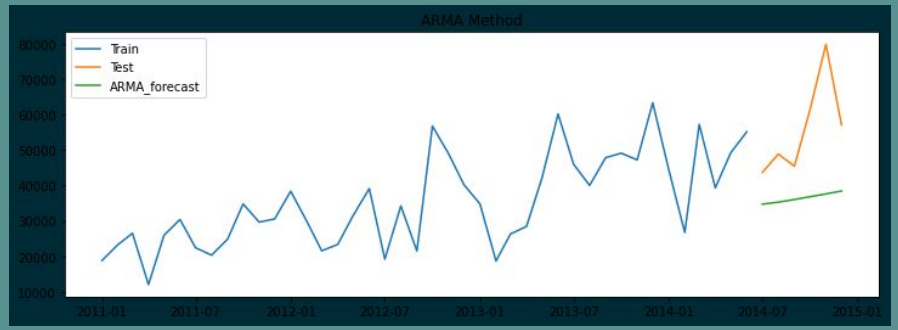
MA model - sales forecast



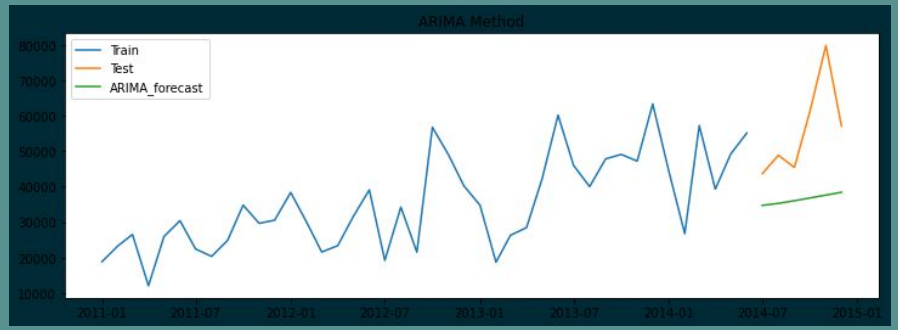
Both the models are able to capture the trend but fails to capture the seasonality.AR model has lesser mape and rmse values than that of Holt's exponential smoothing techniques.MA model forecasts better than Holt's exponential smoothing but worse than AR model.

ARIMA set of techniques: ARMA model and ARIMA model

ARMA model - sales forecast



ARIMA model - sales forecast

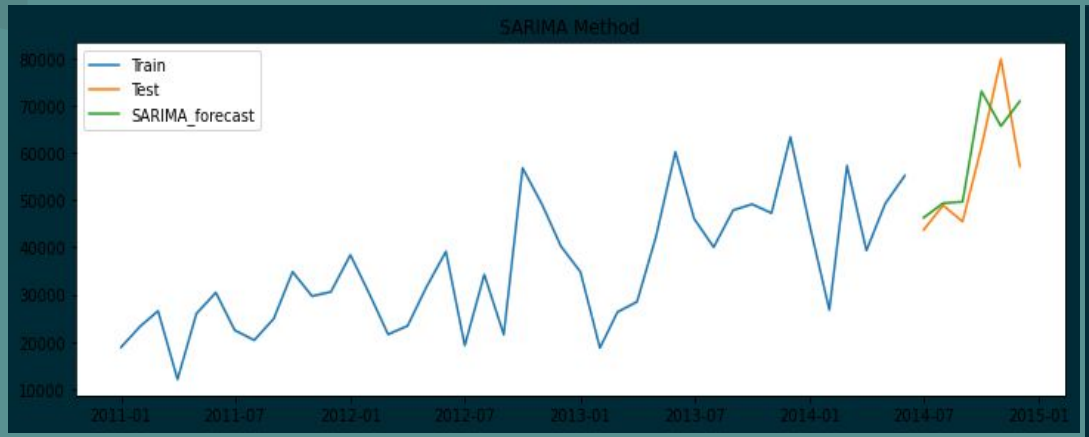


	Method	RMSE	MAPE
0	Simple exponential smoothing	14627.34	15.74
0	Holt's exponential smoothing	18976.37	34.57
0	Holt Winters's additive forecast	9306.82	10.17
0	Holt Winters' multiplicative method	9423.23	11.43
0	Autoregressive (AR) method	10985.28	13.56
0	Moving Average (MA) method	23360.02	33.93
0	Auto Regressive Moving Average (ARMA) method	22654.32	32.40
0	Auto Regressive Integrated Moving Average (ARI...	22654.32	32.40

Both the models perform better than MA model but worse than AR model.

ARIMA set of techniques: SARIMA model

SARIMA model - sales forecast



	Method	RMSE	MAPE
0	Simple exponential smoothing	14627.34	15.74
0	Holt's exponential smoothing	18976.37	34.57
0	Holt Winters's additive forecast	9306.82	10.17
0	Holt Winters' multiplicative method	9423.23	11.43
0	Autoregressive (AR) method	10985.28	13.56
0	Moving Average (MA) method	23360.02	33.93
0	Auto Regressive Moving Average (ARMA) method	22654.32	32.40
0	Auto Regressive Integrated Moving Average (ARI...	22654.32	32.40
0	Seasonal Auto Regressive Integrated Moving Ave...	9614.23	12.86

- SARIMA model is able to capture both trend and seasonality
- This model can forecast sales better than any other ARIMA set of techniques(very low rmse and maps values)
- Among all the techniques used so far (both exponential and ARIMA sets), only Holt Winters' additive method is able to forecast better than SARIMA model.

Conclusions

- 'APAC_Consumer' is the most consistently profitable market segment with least CoV value.
- Our initial time series data of monthly sales of 'APAC_Consumer' had,
  - more than 10 historical data points
  - There was an upward trend
  - There was a seasonal component

  So from the flow chart of time series models, it was expected that Holt Winters' smoothing method and SARIMA model can better forecast the monthly sales.

- After performing all the techniques , these predictions came out to be true.

Conclusions

- Holt Winters' additive forecast is the best smoothing technique for forecasting the sales of 'APAC_Consumer'.
- SARIMA model is the best autoregressive model for forecasting the monthly sales of 'APAC_Consumer'
- When sales were forecasted using these to models, the forecast was able to capture trend and seasonality very well than any other techniques.
- MAPE and RMSE values were very less for these two models, while other techniques resulted in higher mape and rmse values. Having a lower rmse and mape implies that the difference between the predicted and actual values were lesser in these model forecasts.