

# **Retail Giant Sales Forecasting**

**Submitted by-  
VAIBHAV BHARGAVA  
(DS-C42)**

# 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.

# Approach for the Assignment

1

- **Importing** “Global Superstore” Dataset

2

- **Understanding** basic structure of the data

3

- **Preparing the data** by changing the datatype of “Order Date”, concatenating “Market” and “Segment” Column

4

- Performing **Exploratory Data Analysis** on “Sales” and “Profit” columns

5

- **Pre-processing data** for Model Building which includes creating Pivot Table, Train-Test Split, calculation of COV and extracting Final Dataset

6

- **Time Series Decomposition(Additive and Multiplicative)** to check **Trend**, **Seasonality** and **Residual Noise** in the data

7

- Creating **Exponential Smoothing Models** which includes **Simple**, **Holt’s Method** and **Holt’s Winter Method**

8

- Checking **Stationarity** of Given Dataset using **Augmented Dickey-Fuller (ADF)** and **Kwiatkowski-Phillips-Schmidt-Shin (KPSS)** **Auto Regressive Test**

9

- Carrying out **Box-Cox Transformation** and **Differencing** to make Dataset Stationary and again doing **ADF** and **KPSS** test to confirm the same.

10

- Performing **Train-Test Split** and moving Forward for building **Auto-Regressive Models** like **AR,MA,ARMA,ARIMA,SARIMA** Models

11

- Evaluating all the models on the basis of **Mean Absolute Percentage Error (MAPE)**.

12

- Concluding the Best Model based on Lowest **MAPE** value

# Market Segments and their respective CoV Values

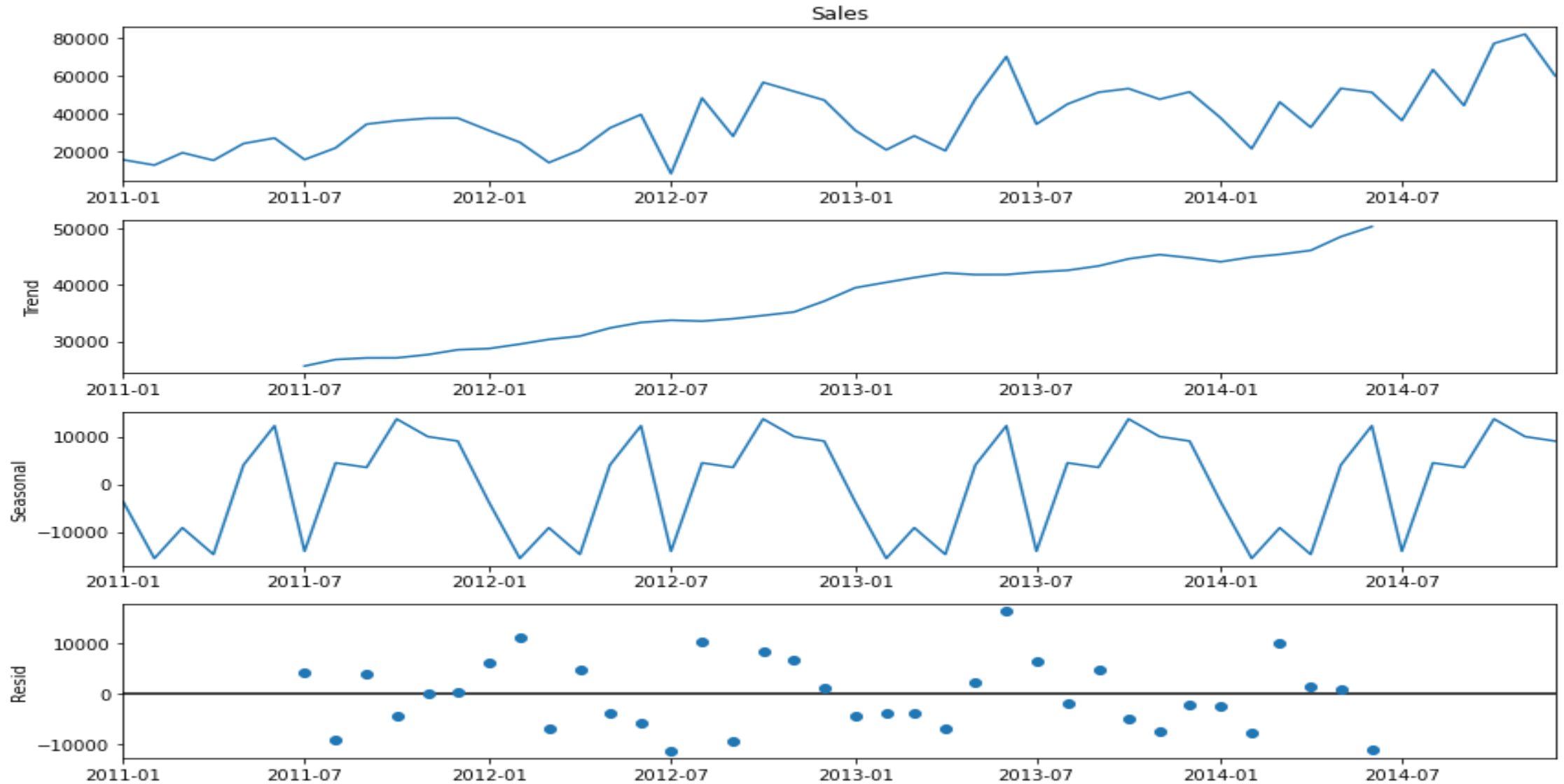
Market Segment	CoV
APAC-Consumer	0.596
EU-Consumer	0.647
LATAM-Consumer	0.681
EU-Corporate	0.689
APAC-Corporate	0.732
LATAM-Corporate	0.880
US-Corporate	1.027
APAC-Home Office	1.049
US-Consumer	1.095
EU-Home Office	1.114
Canada-Corporate	1.197
US-Home Office	1.217
LATAM-Home Office	1.344
Africa-Consumer	1.429

Market Segment	CoV
Canada-Consumer	1.476
Africa-Corporate	1.665
Africa-Home Office	1.990
Canada-Home Office	2.188
EMEA-Consumer	2.717
EMEA-Home Office	6.067
EMEA-Corporate	6.780

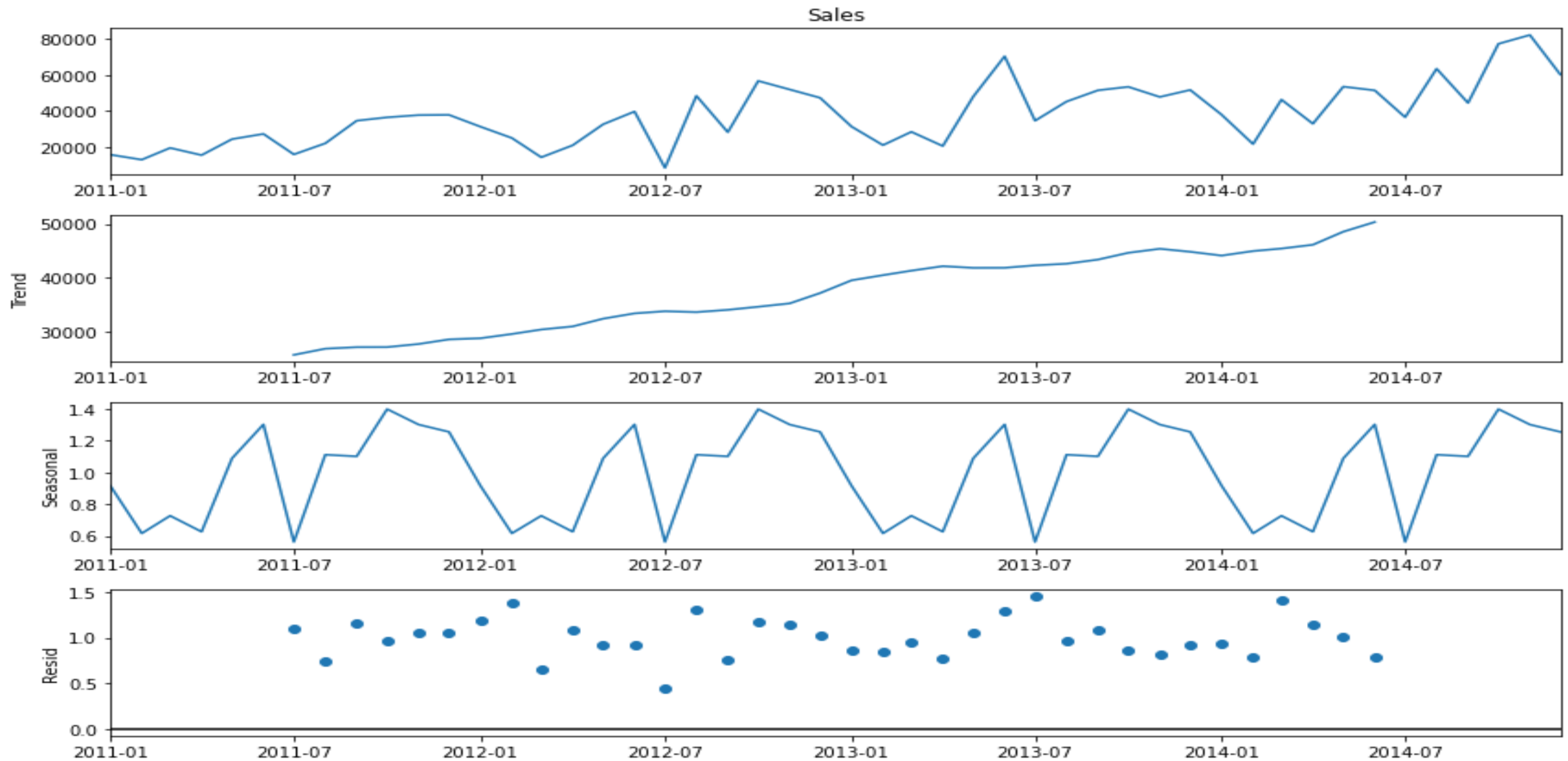
## Conclusion:-

Since we want very less variation in profit we will choose the Market Segment with least value of COV i.e. APAC-Customer for our further forecasting purposes.

# Time Series Decomposition-Additive

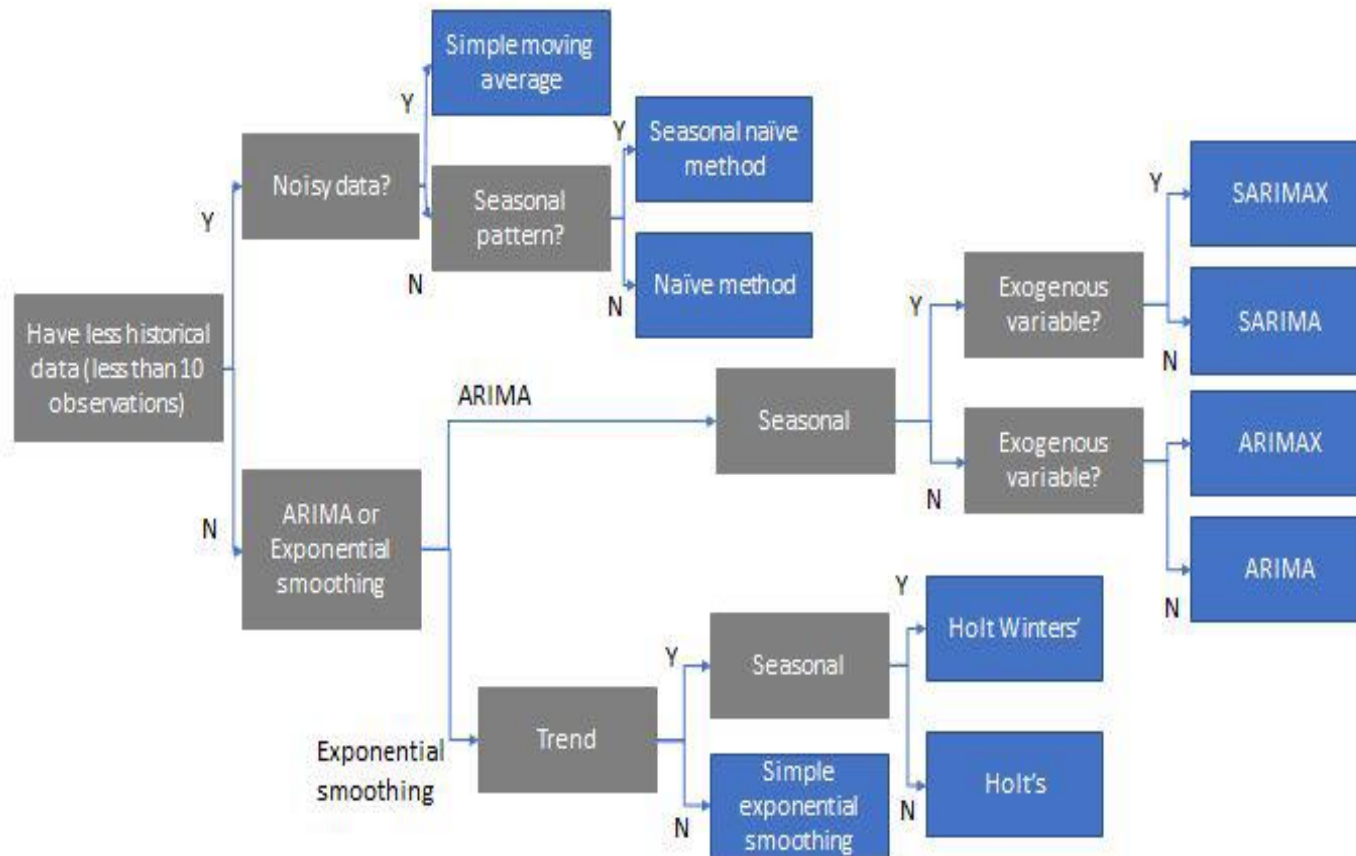


# Time Series Decomposition-Multiplicative



# Flowchart

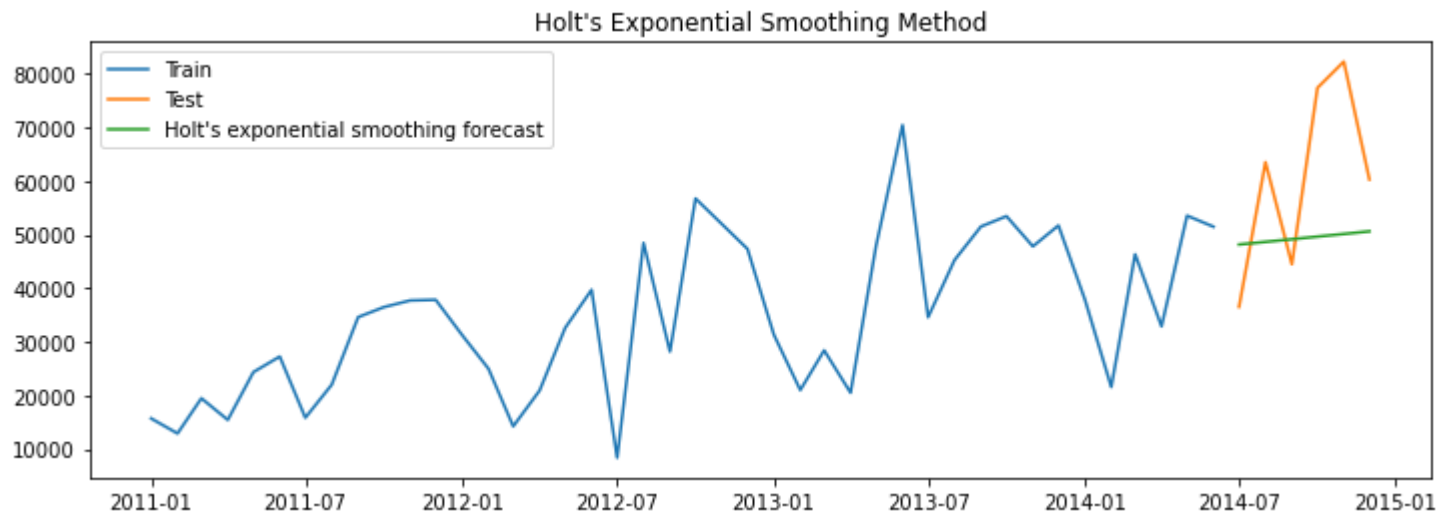
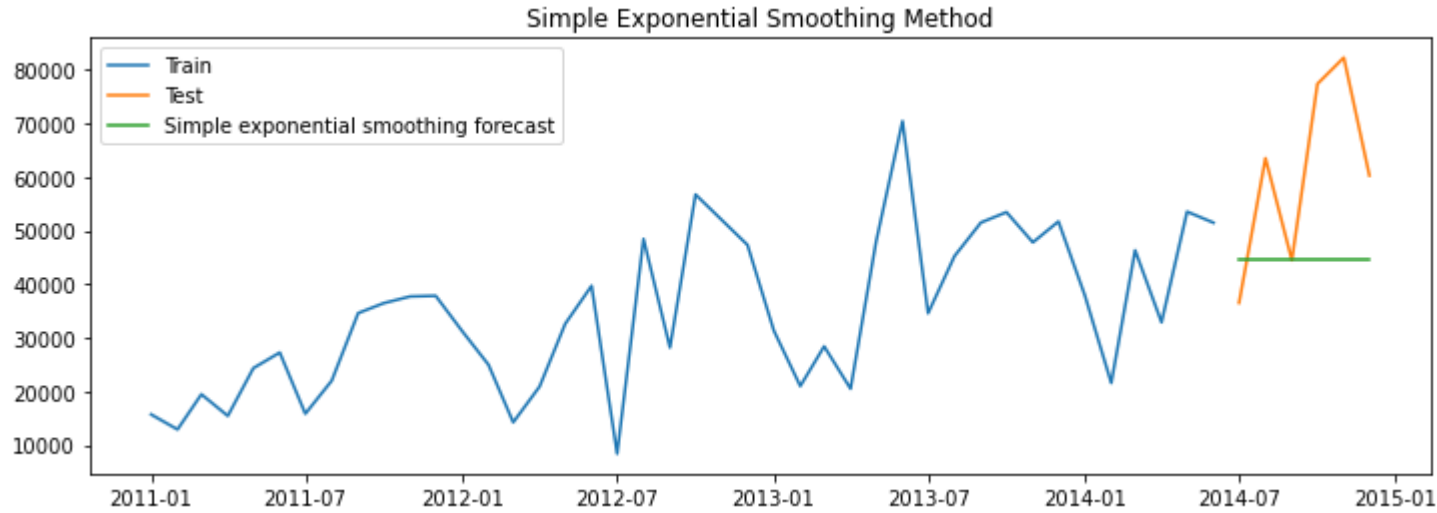
## Choosing the Right Time Series Method



### ❑ Conclusion:

Since the no of Data Points are more than 10 and Data Set is showing Trend and Seasonal pattern as shown in Previous Slides. Hence, we can conclude that Holt Winters from Exponential Smoothing and SARIMA Model from ARIMA Models will perform better as they take care of Seasonality and Trend better than other Models.

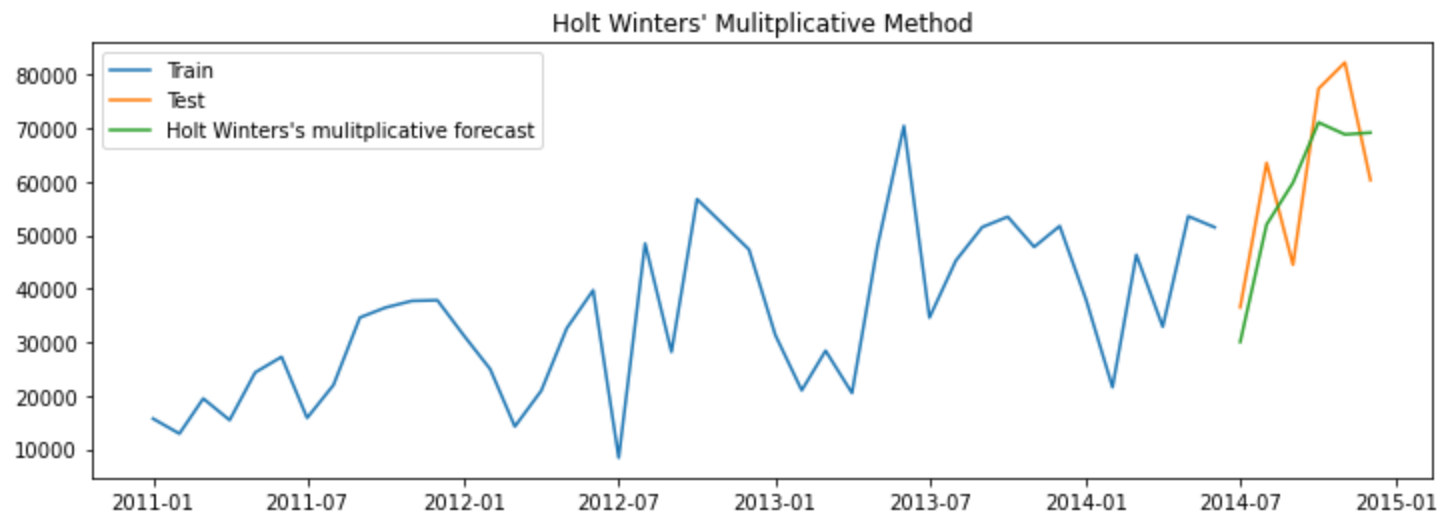
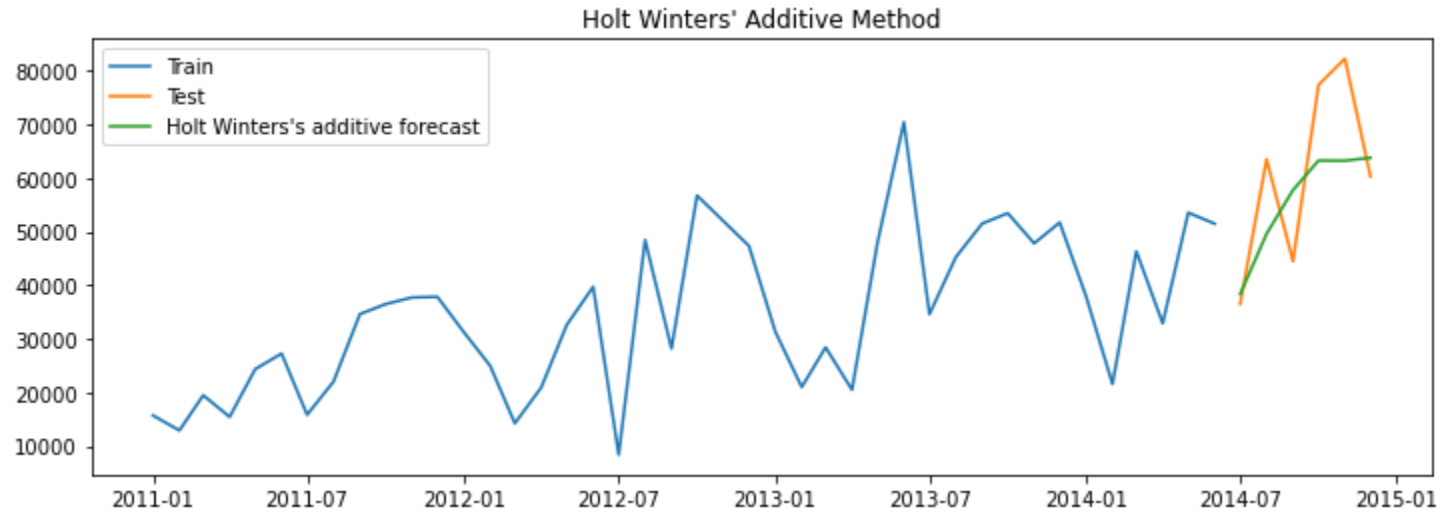
# Exponential Smoothing Models



Method	MAPE
Simple Exponential	27.70
Holt's Exponential	26.12

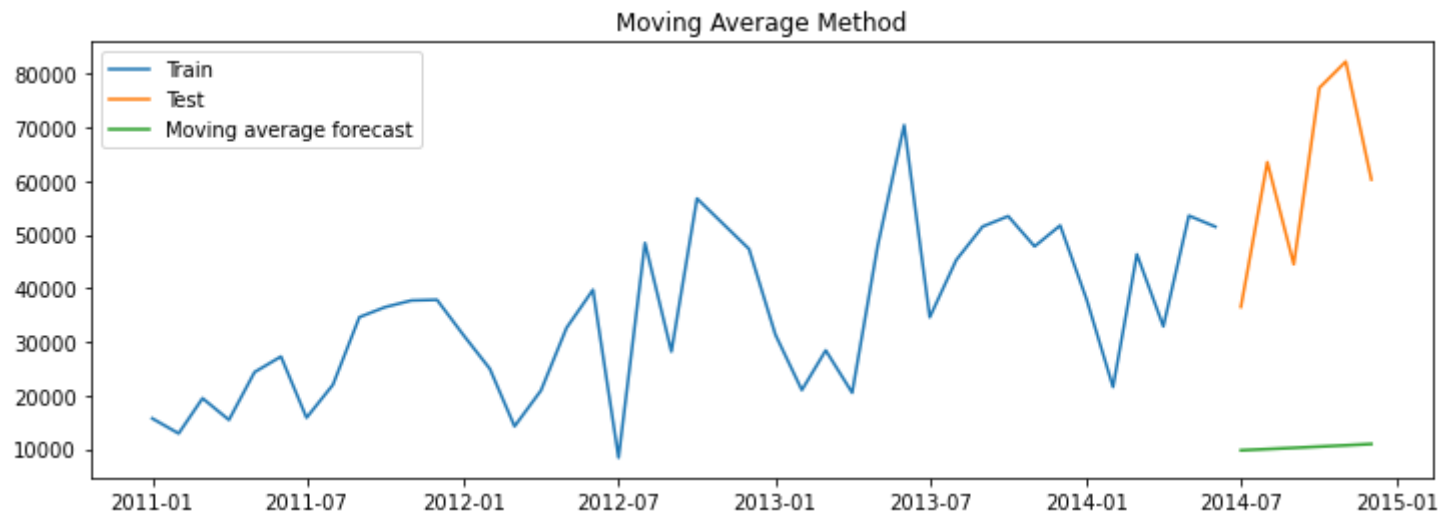
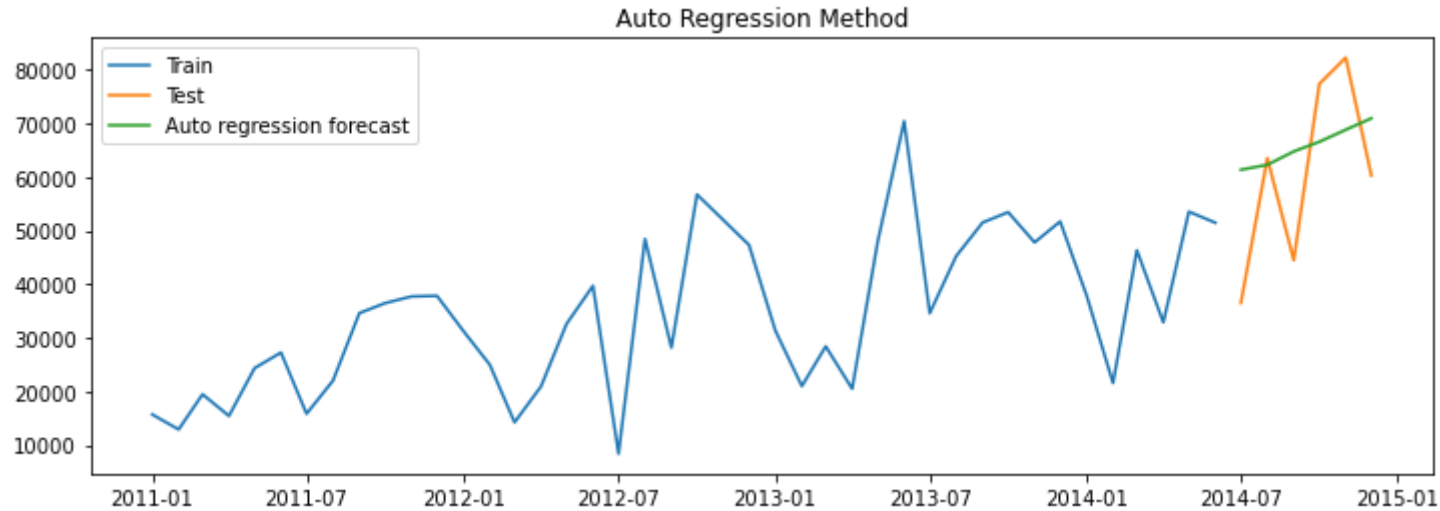


# Exponential Smoothing Models



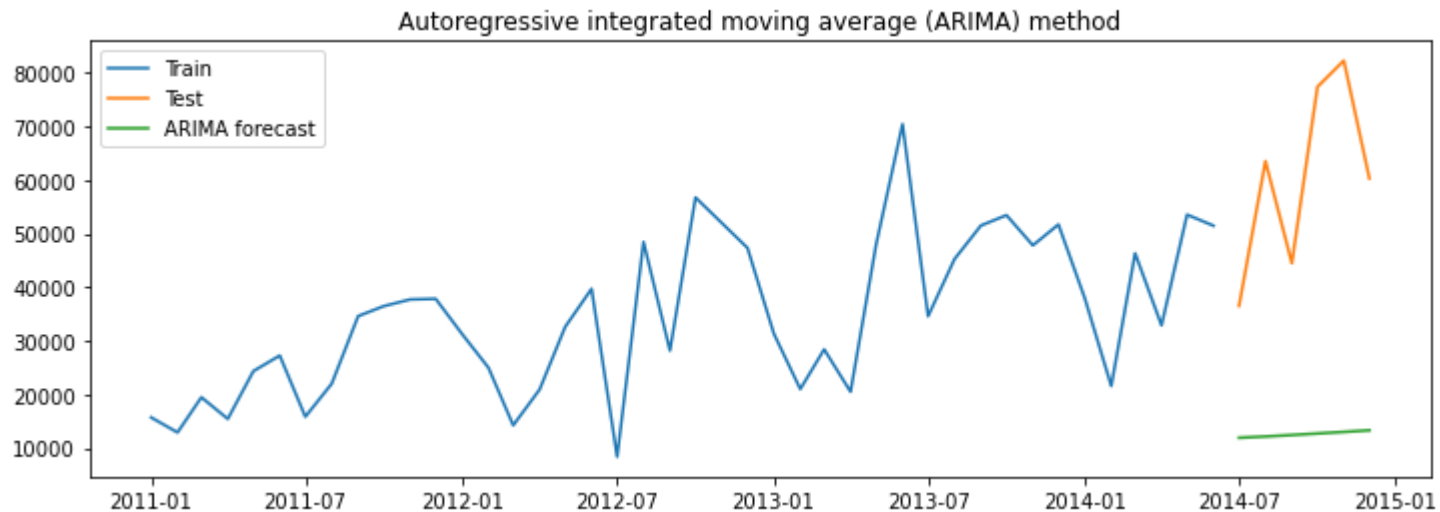
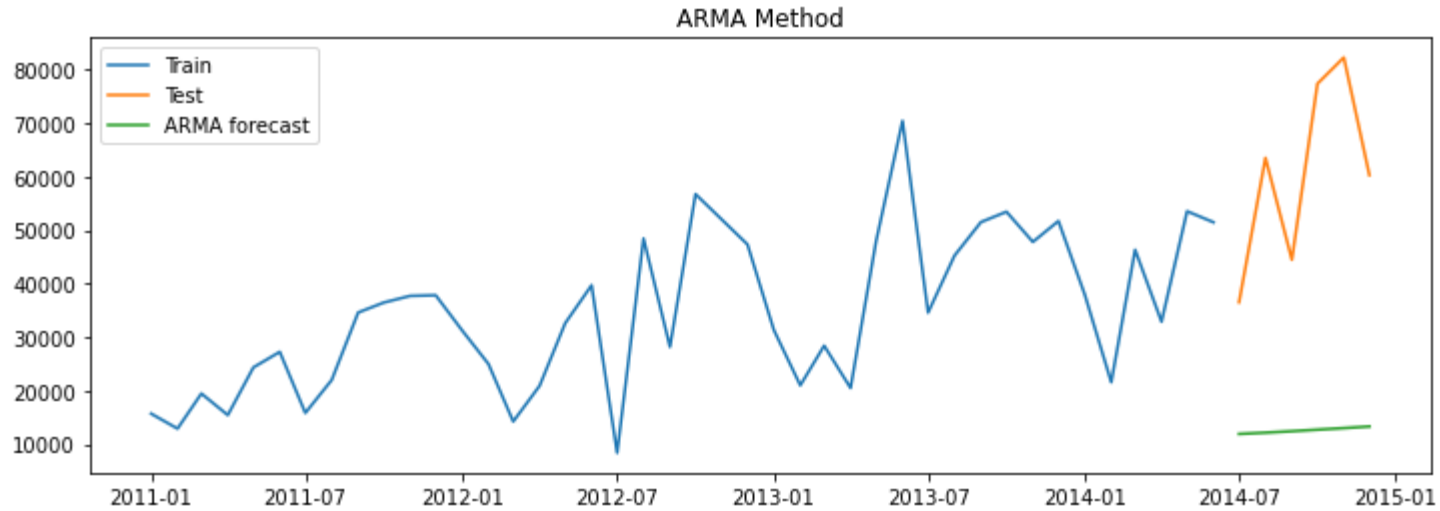
Method	MAPE
Simple Exponential	27.70
Holt's Exponential	26.12
Holt Winters Additive	17.32
Holt Winters Multiplicative	18.27

# Auto Regressive Models



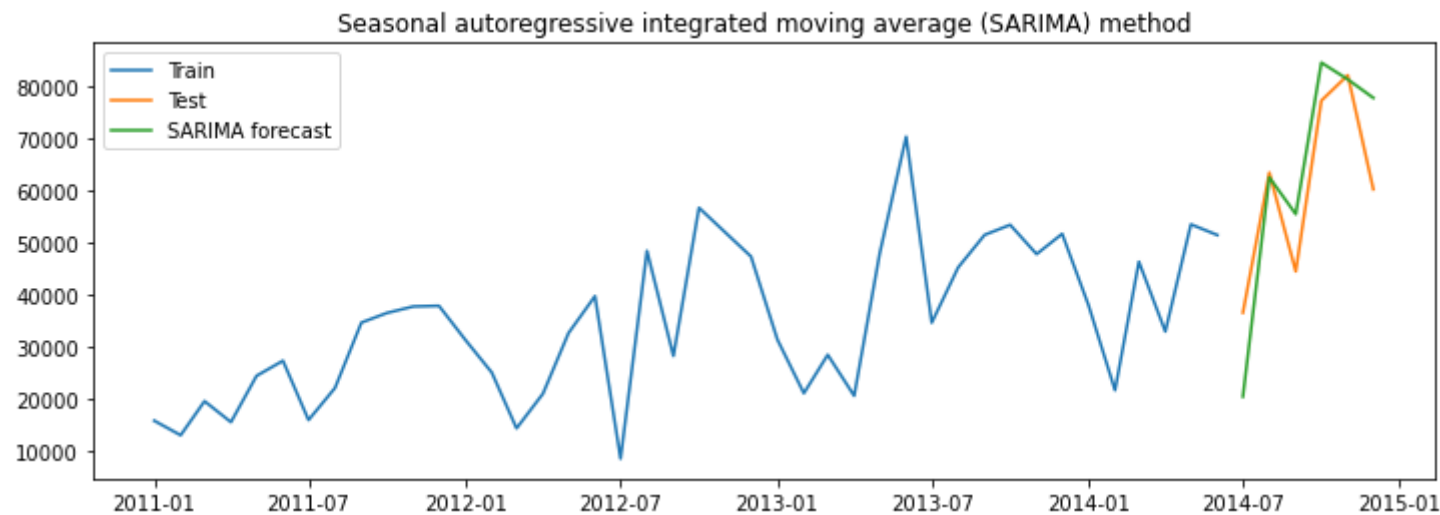
Method	MAPE
Simple Exponential	27.70
Holt's Exponential	26.12
Holt Winters Additive	17.32
Holt Winters Multiplicative	18.27
Auto Regressive(AR)	27.27
Moving Average(MA)	81.64

# Auto Regressive Models



Method	MAPE
Simple Exponential	27.70
Holt's Exponential	26.12
Holt Winters Additive	17.32
Holt Winters Multiplicative	18.27
Auto Regressive(AR)	27.27
Moving Average(MA)	81.64
ARMA	77.66
ARIMA	77.66

# Auto Regressive Models



Method	MAPE
Simple Exponential	27.70
Holt's Exponential	26.12
Holt Winters Additive	17.32
Holt Winters Multiplicative	18.27
Auto Regressive(AR)	27.27
Moving Average(MA)	81.64
ARMA	77.66
ARIMA	77.66
SARIMA	18.38

# Conclusion

- ❖ APAC-Consumer Segment proved to be most Profitable Segment based on COV Value.
- ❖ From the Simple Exponential Smoothing Methods Holt Winter's Additive Method performed best since it has lowest MAPE value among exponential Smoothing Methods
- ❖ From the Auto Regressive Models SARIMA method performed best since it has lowest MAPE value among auto regressive models
- ❖ Overall Holt Winter's Method performed best since it has lowest MAPE value among all.
- ❖ While Doing Time Series Decomposition we have observed that data has Trend and Seasonality hence Holt Winters or SARIMA model will perform better on this data since they capture Trend and Seasonality better than any other Models.
- ❖ While referring Forecast Plots of Holt Winter's and SARIMA Models we can see clearly that the model is capturing the Trend and Seasonality of Test Data quite well compared to other Models and hence has lower MAPE values.

***Thank you!***