Retail Giant Sales Forecast

Madhurima. D

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 & home office.

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

21 Market Segments

There are 21 unique "Market-Segments" for which the sales forecasts can be made.

- APAC-Consumer
- LATAM-Consumer
- US-Consumer
- EU-Consumer
- APAC-Corporate
- EU-Corporate
- LATAM-Corporate
- US-Corporate
- EMEA-Consumer
- Africa-Consumer
- APAC-Home Office
- LATAM-Home Office
- US-Home Office
- EU-Home Office
- EMEA-Corporate
- Africa-Corporate
- EMEA-Home Office
- Africa-Home Office
- Canada-Consumer
- Canada-Corporate
- Canada-Home Office

Understanding the market

Due to certain unpredictable circumstances in the market, as a company, we must prioritise only the best and most consistent market segment in terms of profitability. We must see which market segment is the most consistently profitable. And then, forecast the sales for that most consistently profitable market-segment only. This way we know that the market region our company is investing in will be beneficial for the company as the forecasts will be reliable.

To find the most consistently profitable market-segment you will be using a measure called "Coefficient of Variation (CoV)" which is nothing but the ratio of the standard deviation to mean for the data that it is being calculated for.

Coefficient of variation

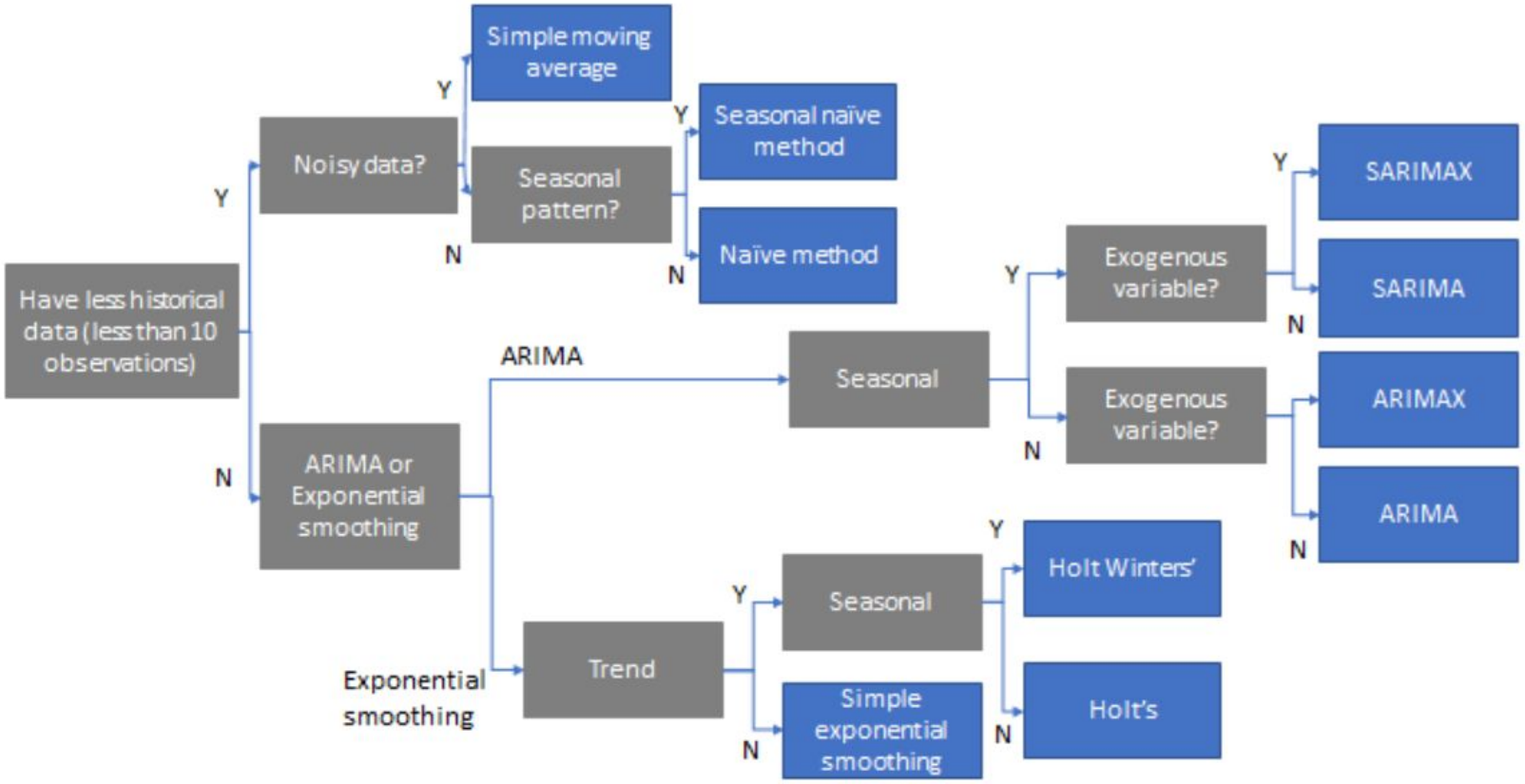
The coefficient of variation normalises the standard deviation with the mean and gives us a comparative figure on the basis of which we can identify the most profitable market segment.

As we can see in the table here, the CoV for the Market Segment 'APAC-Consumer' is the least here, hence we can say that this is the most consistently profitable Market Segment, as we want to forecast the sales for the market segment has less variation in the profits.

Market-Segment	
APAC-Consumer	0.522725
APAC-Corporate	0.530051
EU-Consumer	0.595215
LATAM-Consumer	0.683770
EU-Corporate	0.722076
LATAM-Corporate	0.882177
EU-Home Office	0.938072
APAC-Home Office	1.008219
US-Consumer	1.010530
US-Corporate	1.071829
US-Home Office	1.124030
LATAM-Home Office	1.169693
Canada-Consumer	1.250315
Africa-Consumer	1.310351
Canada-Corporate	1.786025
Africa-Corporate	1.891744
Africa-Home Office	2.012937
Canada-Home Office	2.369695
EMEA-Consumer	2.652495
EMEA-Corporate	6.355024
EMEA-Home Office	7.732073

Most Profitable Market Segment:
Asia-Pacific Consumer
(APAC-Consumer)

Choosing the Right Time Series Method



Optimum Technique Best for Sales Forecast

Smoothing Technique

Looking at the flowchart, we can say the best smoothing technique could be Holt Winter's Smoothing Technique as we the data given to us has more than 10 Observations and with a definite trend and an unsure seasonality.

Auto-Regressive Model

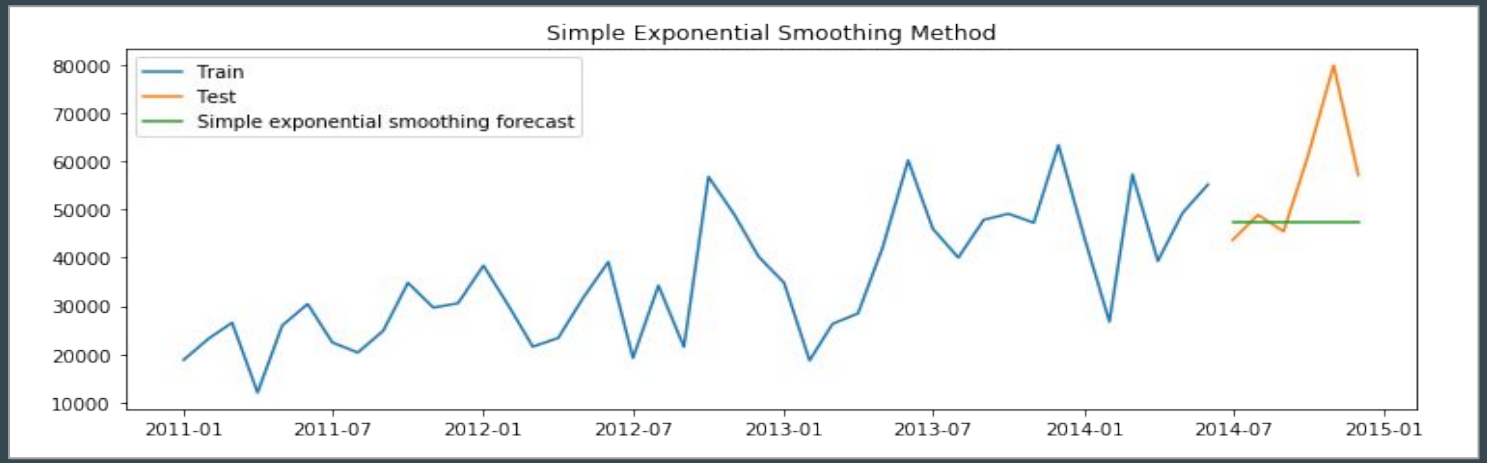
Looking at the flowchart, we can say that the best AR Model could be the SARIMA Model, Seasonal Autoregressive integrated moving average Model as the data has more than 10 Observations and with a definite trend and an unsure seasonality.

Time Series Analysis

We have applied all the below methods to see if the hypothesis we formed from the flowchart was correct or not based on the forecast plots and their MAPE values

- Simple exponential smoothing
- Holt's exponential smoothing
- Holt-Winters' exponential smoothing - Additive
- Holt-Winters' exponential smoothing - Multiplicative
- AR model
- 2. MA model
- 3. ARMA model
- 4. ARIMA model
- 5. SARIMA model

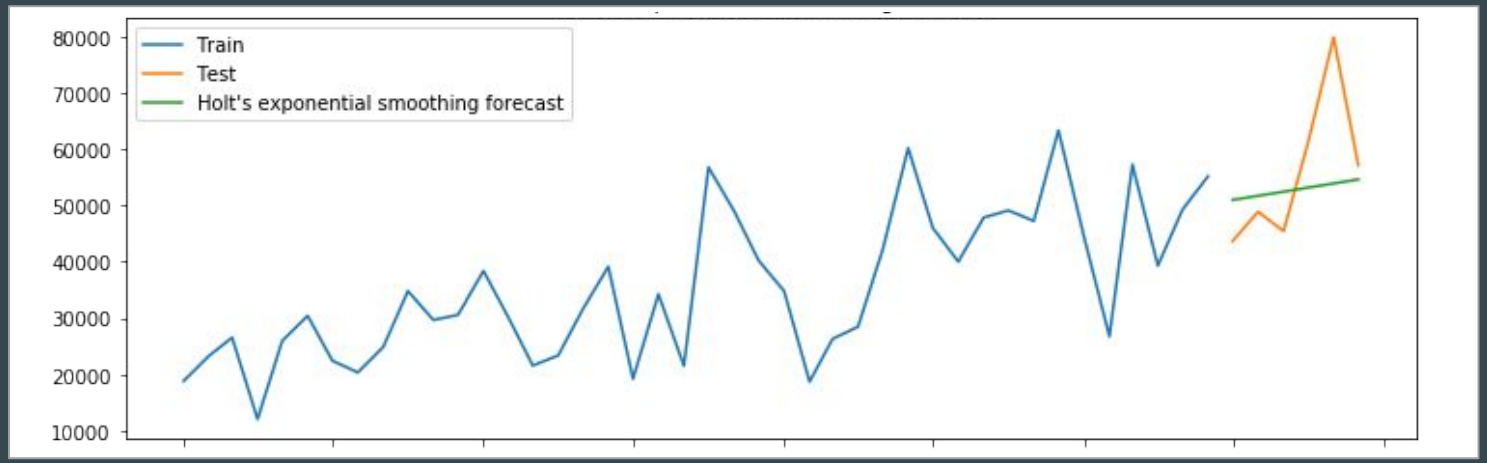
Simple Exponential Smoothing Technique



We see that the plot of the data has a trend which is moving upwards and also has a seasonality, but the forecast here is a straight line which does not show a Trend or does not capture the seasonality. But it captures the Level which is the baseline of the Time Series.

MAPE Value	15.99

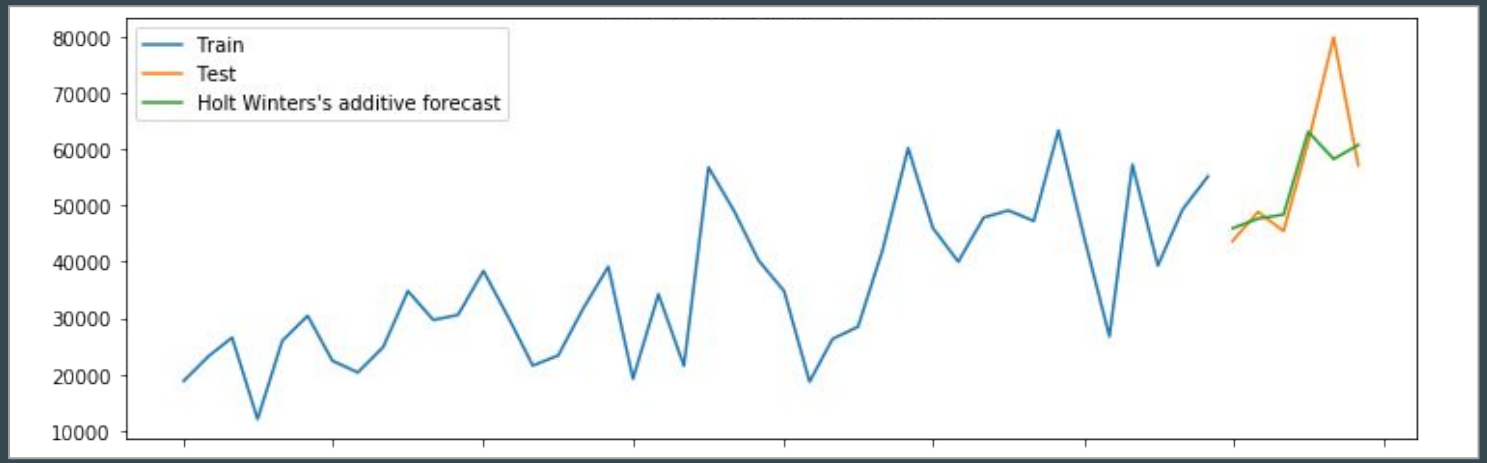
Holt's Exponential Smoothing Technique



We see that the forecast here is a line which captures a slightly upward Trend. But it does not capture the Seasonality of the Time Series yet.

MAPE Value	14.67

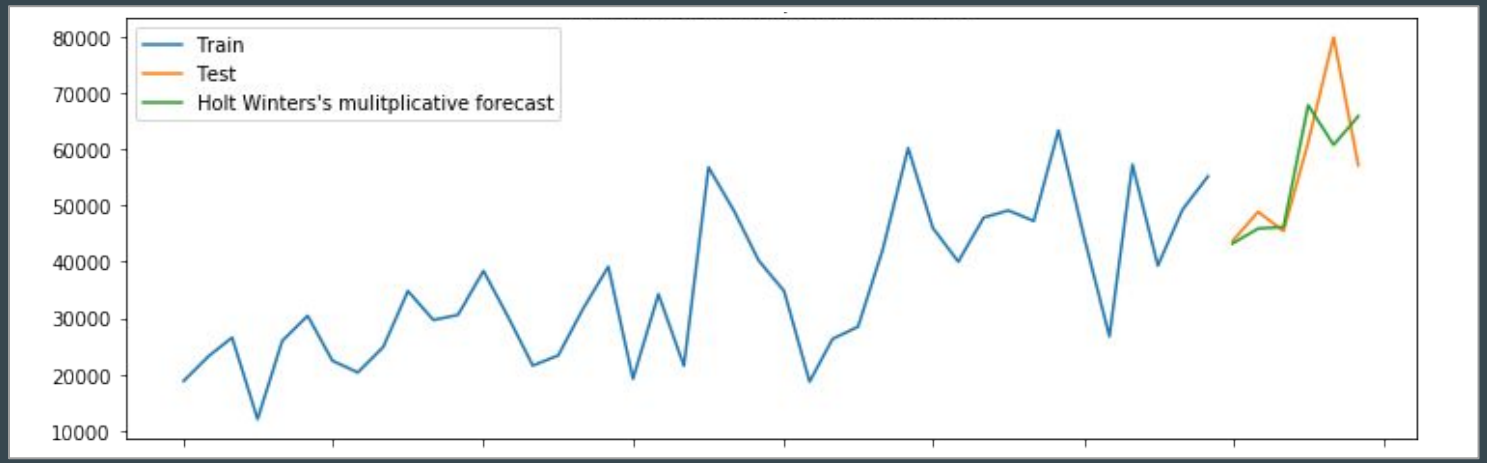
Holt-Winter's Additive Technique



We see that the forecast here looks good enough as it captures a the upward Trend and also the Seasonality of the Time Series, with a MAPE Value of 8.42 which is low as compared to the other methods.

MAPE Value	8.42

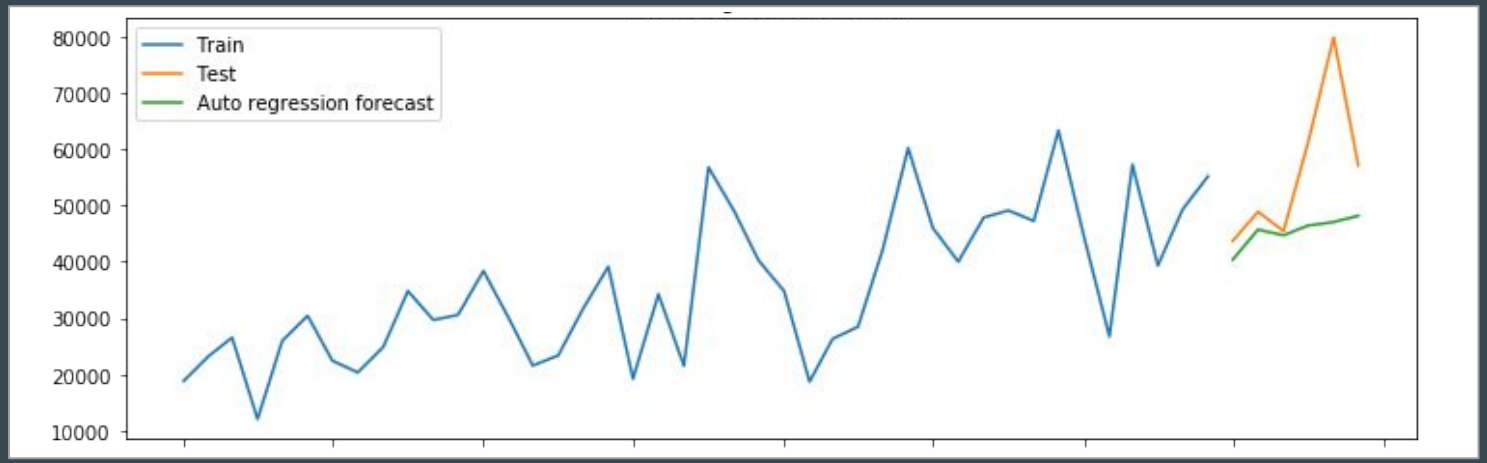
Holt-Winter's Multiplicative Technique



We see that the forecast here looks good enough as it captures an upward Trend and also the Seasonality of the Time Series. But the MAPE Value is 9.78 which is higher than the Holt Winter's Additive Forecast.

MAPE Value	9.78

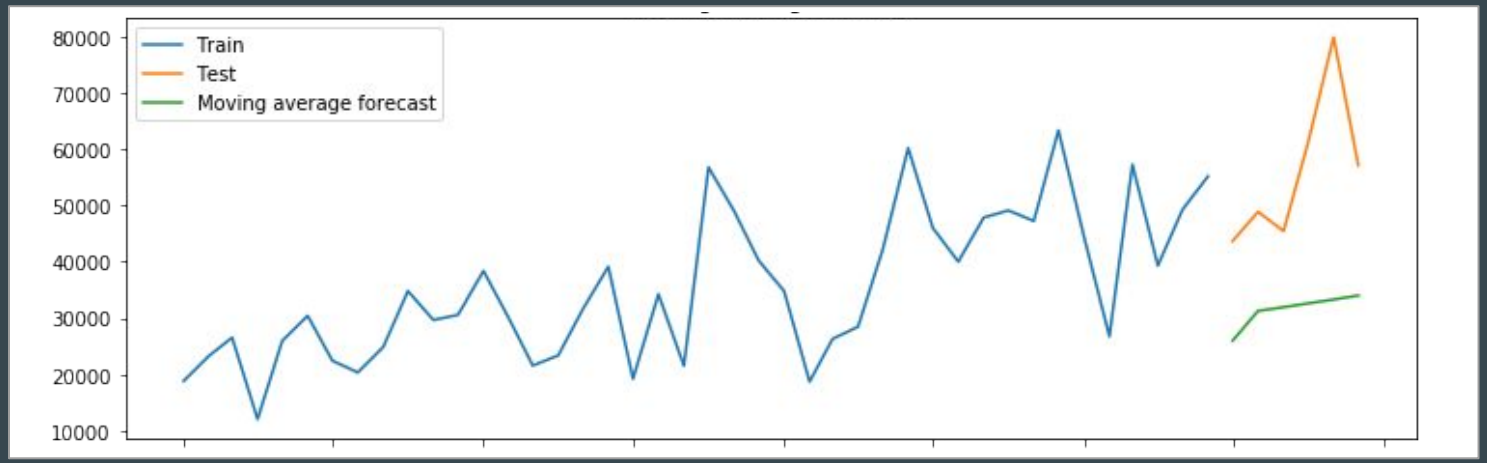
Simple Auto Regressive Model



We see that the forecast here has captured a slightly upward Trend and also a low Seasonality. But the MAPE Value is 16.11 which is quite high.

MAPE Value	16.11

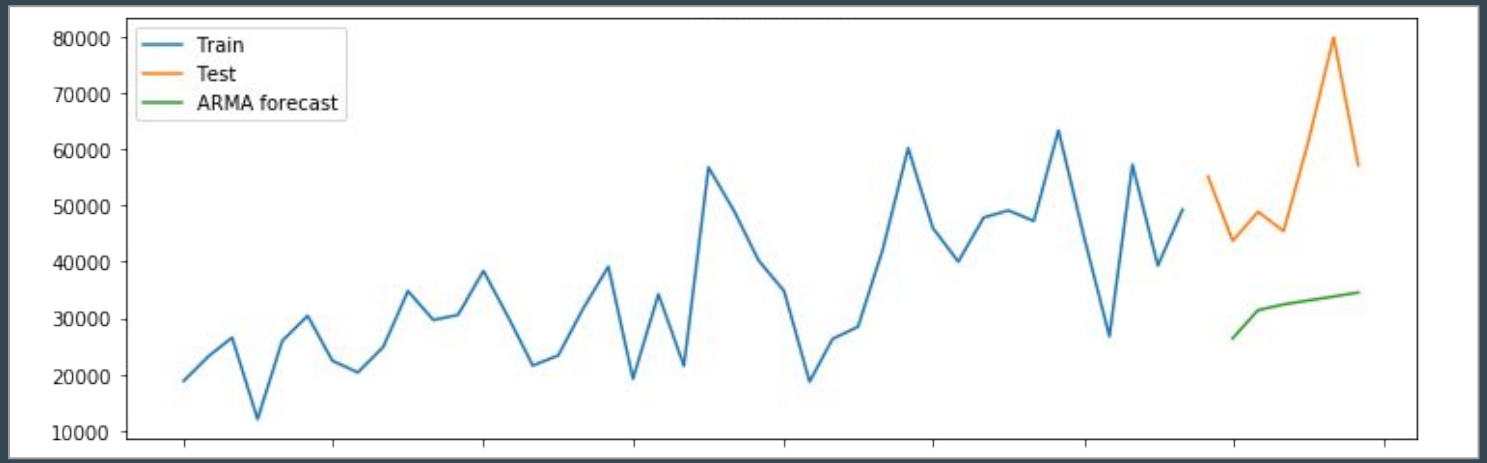
Moving Average Model



We see that the forecast here is nowhere close to the Actual values.

MAPE Value	41.97

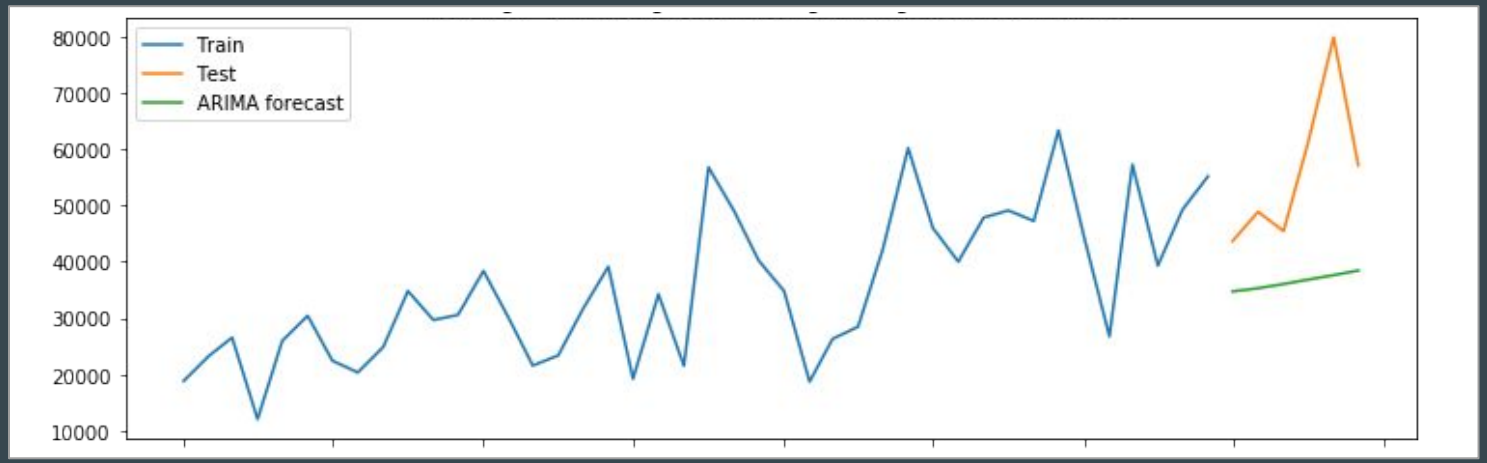
Auto-Regressive Moving Average Model



We see that the forecast here is nowhere close to the Actual values.

MAPE Value	41.17

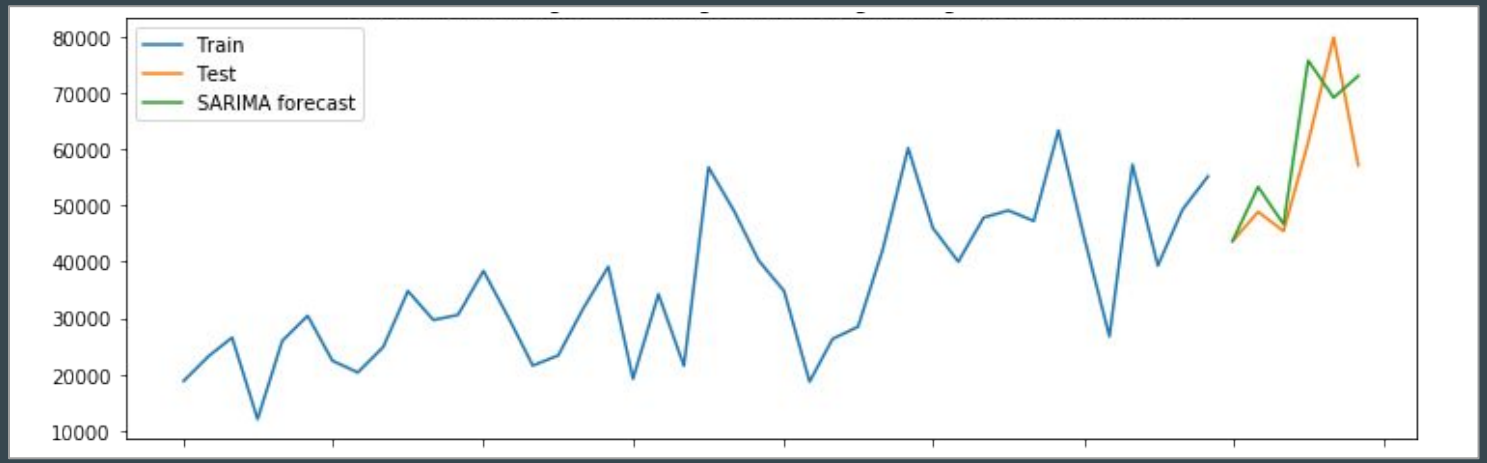
Auto-Regressive Integrated Moving Average Model



We see that the forecast here is nowhere close to the Actual values.

MAPE Value	16.11

Seasonal Auto-Regressive Integrated Moving Average Model



We see that the forecast here is quite close to the Actual values which captures both Trend and Seasonality.

MAPE Value	12.78

Conclusion

Best Smoothing Technique

The Holt Winter's Additive plot works best as it captures the Level, Trend and Seasonality. We can see that the forecast for Holt Winter's Additive Method fits better to the actual values with a least MAPE Value of 8.42 as compared to other models.

Best AR Model

The SARIMA Model is better to the actual values as compared with the other models which closely captures the seasonal component and the forecast fits better with a least MAPE Value of 12.78.