Retail Giant Sales Forecasting

SUBMISSION

Group Name:

1. Mohit Mamgain
2. Vishnu Poddar
3. Savita Upadhyay

Objective

- “Global Mart” is an online store super giant having worldwide operations. It takes orders and delivers across the globe and deals with all the major product categories - consumer, corporate & home office.
- Our aim is to forecast the sales and the demand for the next 6 months, that would help you manage the revenue and inventory accordingly.

Business Understanding

The store caters to 7 different market segments and in 3 major categories. We want to forecast at this granular level, so you subset your data into 21 (7*3) buckets before analyzing these data. But not all of these 21 market buckets are important from the store's point of view.

So you need to find out 2 most profitable (and consistent) segment from these 21 and forecast the sales and demand for these segments.

Data Cleaning and Preparation

- First segment the whole dataset into the 21 subsets based on the market and the customer segment level.
- Aggregate the 3 attributes - Sales, Quantity & Profit, over the Order Date to arrive at monthly values for these attributes.
- Calculate Coefficient of Variation(CV).
- Pick Top 2 Segments based on CV & Profit

Market	Segment	Profit_sum	Profit_cv
APAC	**Consumer**	**222817.560**	**63.21323**
EU	**Consumer**	**188687.707**	**62.43052**

- Aggregate the 3 attributes - Sales, Quantity & Profit, over the Market, Segment and month on the subset of Top 2 Segments.

Model building :Classical decomposition and auto ARIMA

1. Created time series for aggregated data of APAC,Consumer & EU,Cosumer subsets for first 48 months.
2. First 42 months will be used to train the model and the rest 6 months to test the model.
3. Smoothened time series using Moving Average method.
4. Built a model on the smoothed time series using classical decomposition
5. Model 1: Linear model : Created using lm() function from forecast package in R on train data.

i) Fitted a multiplicative model(as the amplitude has also increased with time) with trend and seasonality to the data. Modeled seasonality using a sinusoid function

ii)Modeled using one degree and two degree polynomial.

Model building (continued).

6. Model 2: Auto Arima Model: Created using auto.arima() function from forecast package in R train data

7.Checked if residual resembles white noise using below tests:

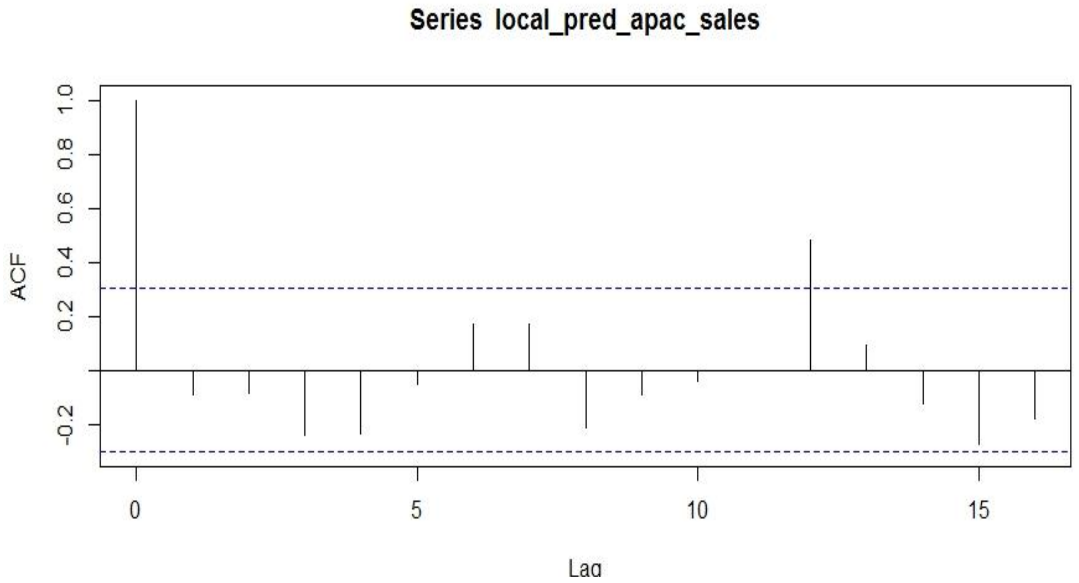
i) Augmented Dickey-Fuller Test

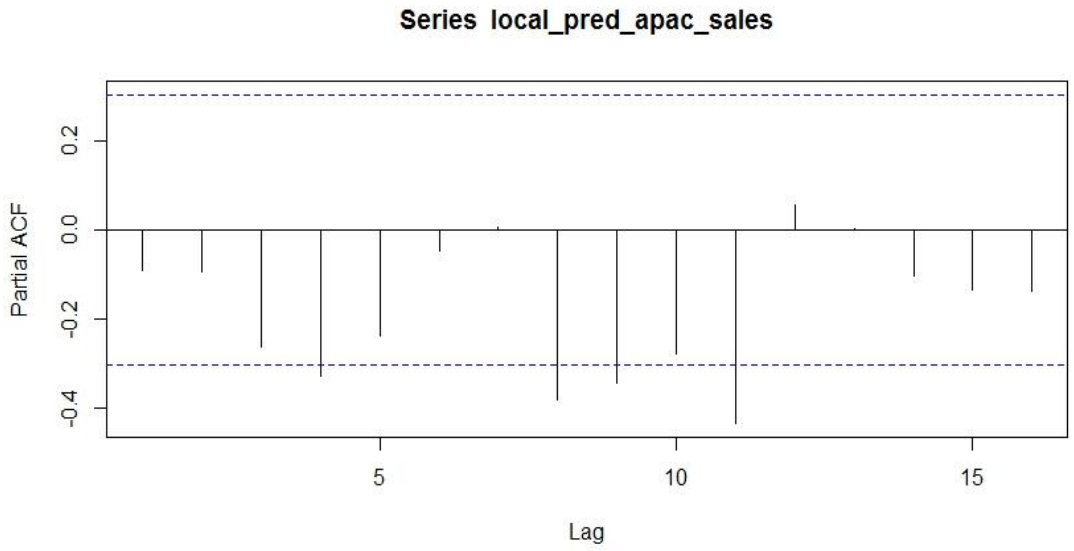
ii) KPSS Test

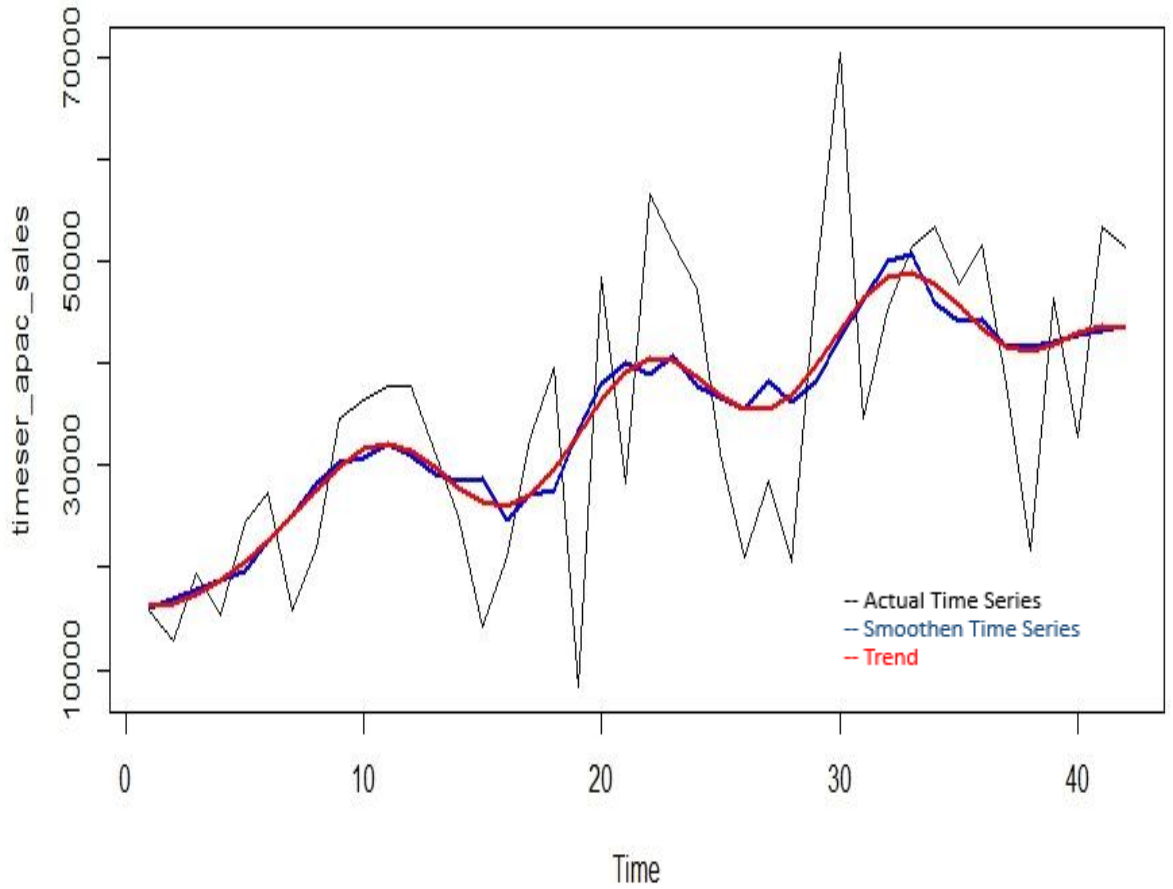
8. Both models were evaluated using Mean Absolute Percentage Error(MAPE).

Model Analysis & Results

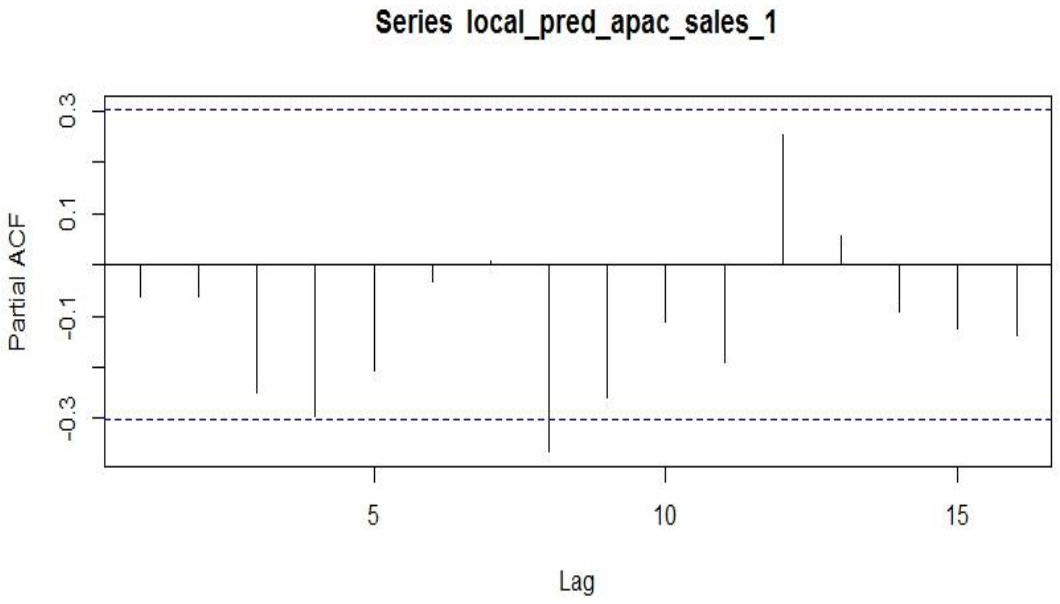
Time Series Analysis of Sales for CONSUMER Segment in APAC marketplace with Trendline Degree 3 – MAPE 37.38

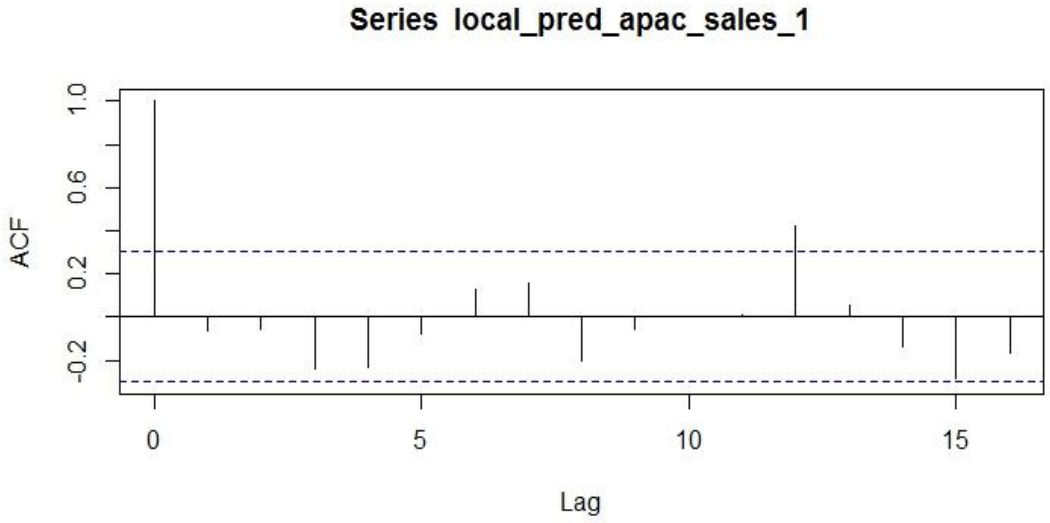


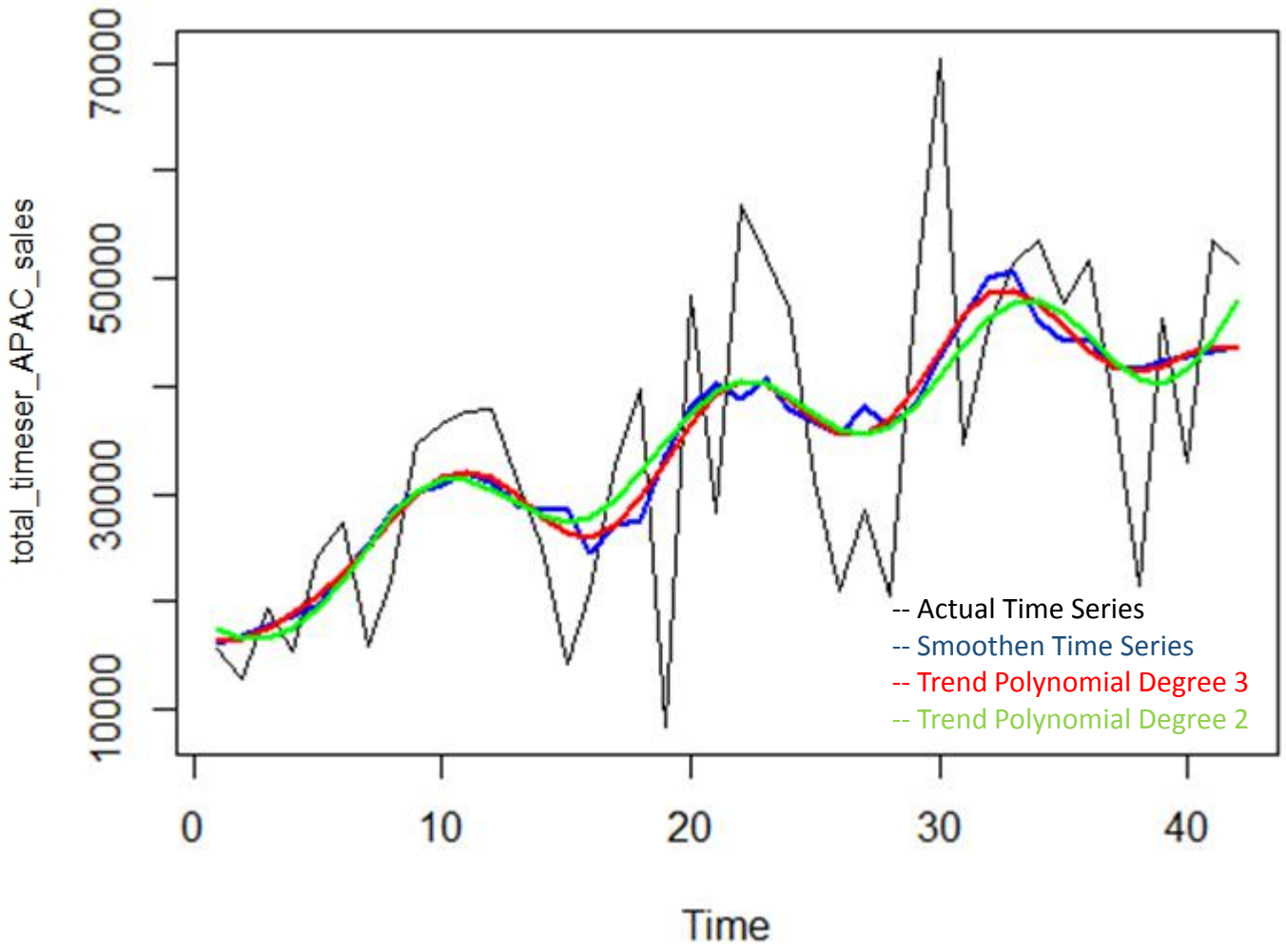




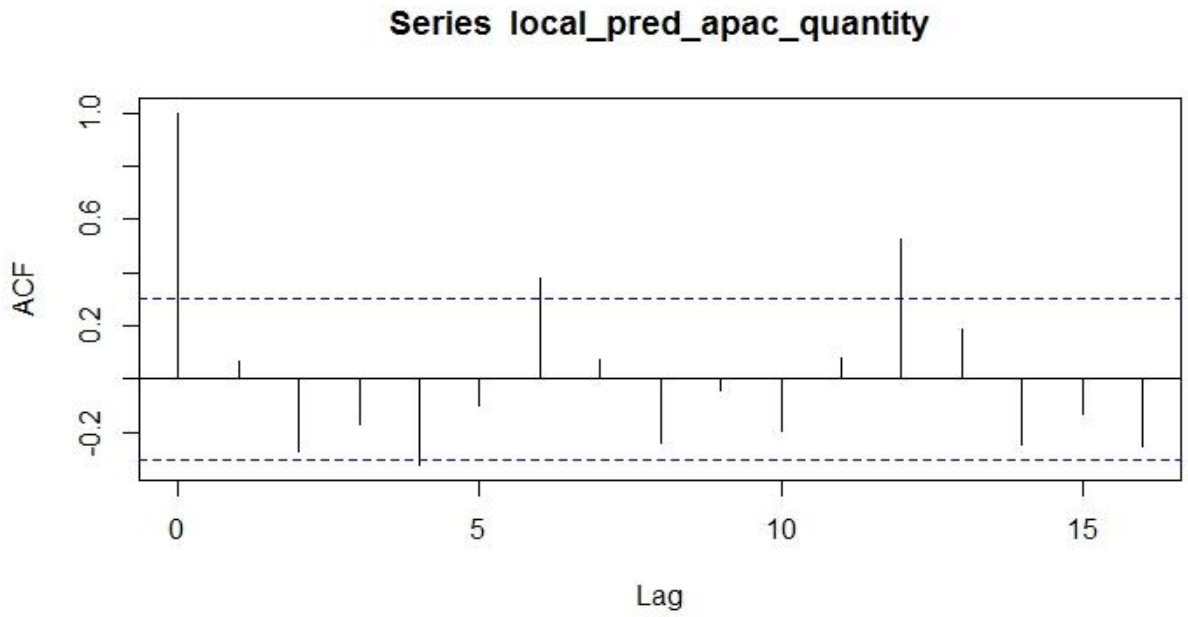
Time Series Analysis of Sales for CONSUMER Segment in APAC marketplace with Trendline Degree 2 – MAPE 27.7

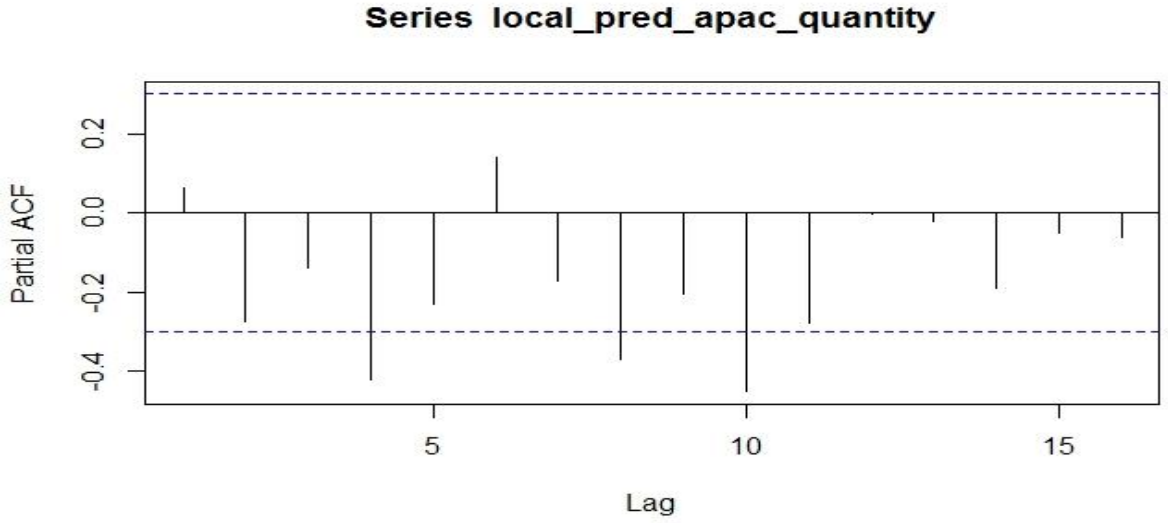


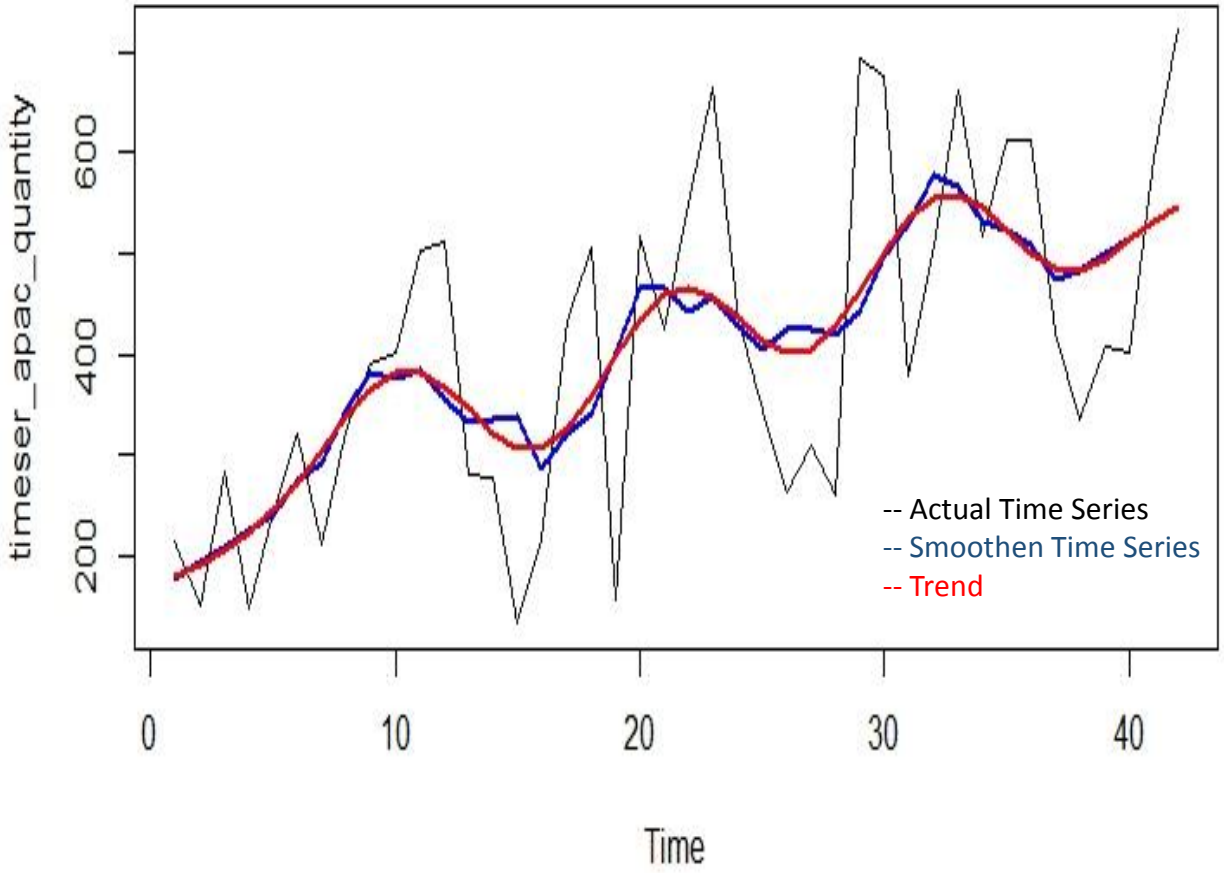




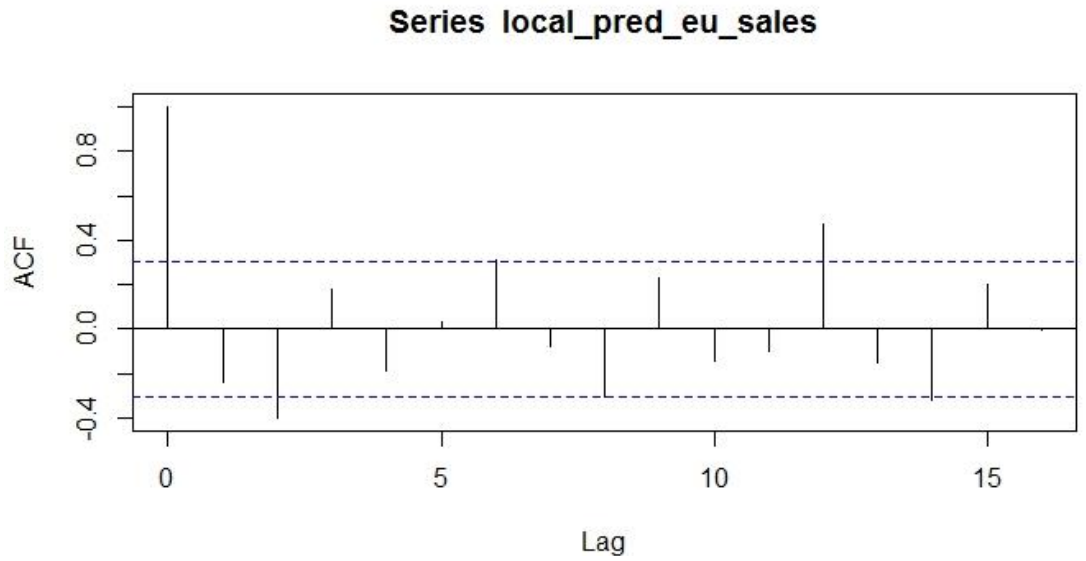
Time Series Analysis of Quantity for CONSUMER Segment in APAC marketplace with Trendline Degree 2 – 30.49

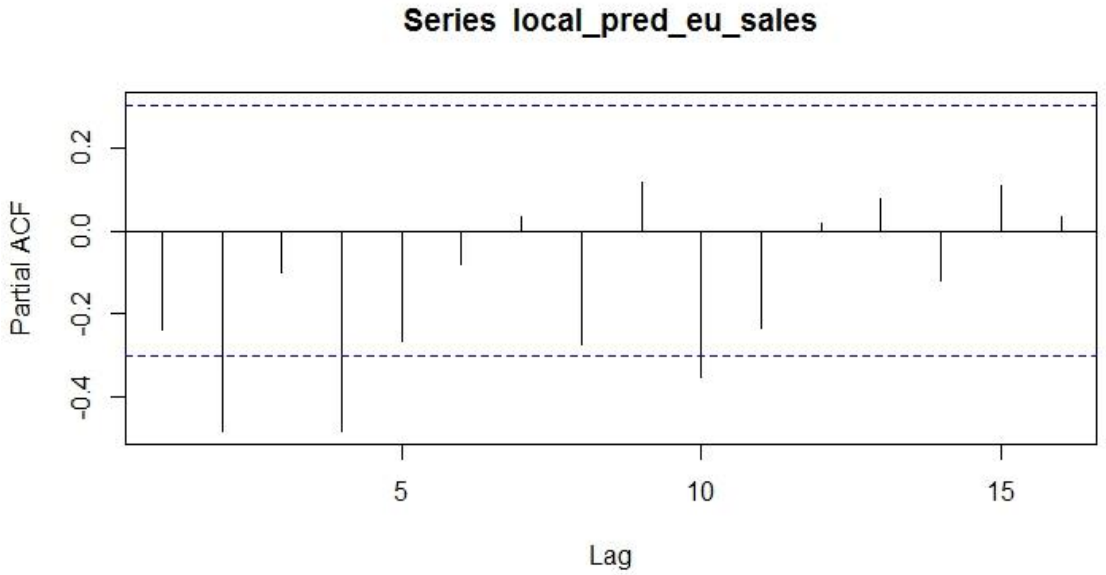


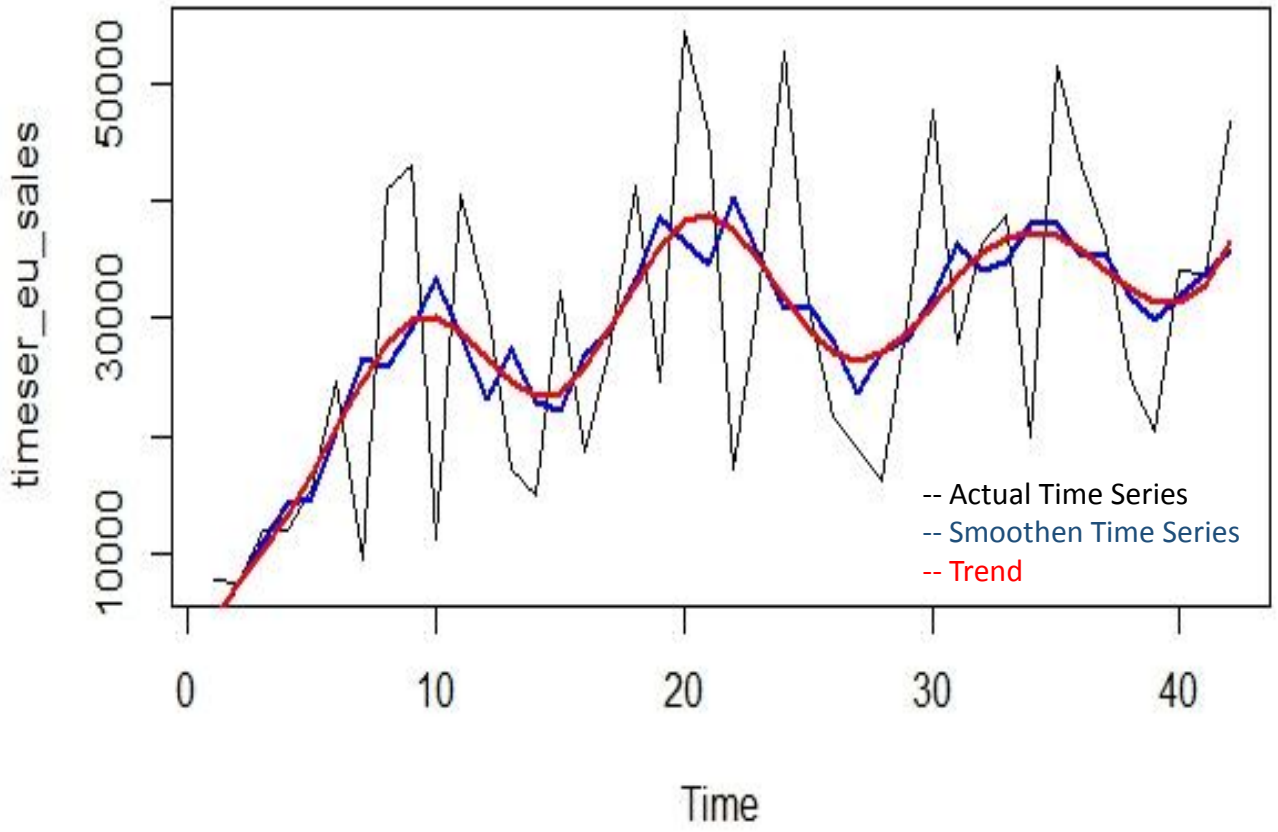




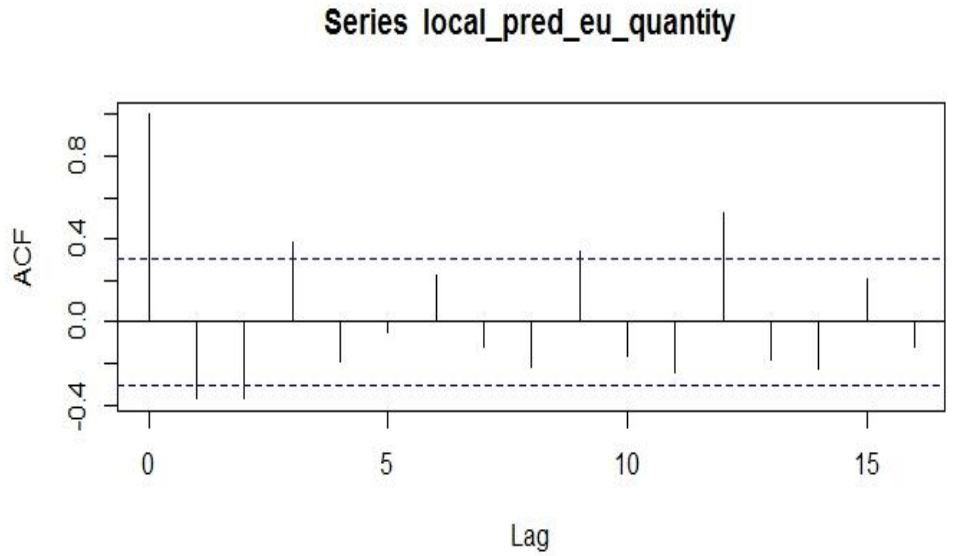
Time Series Analysis of Sales for CONSUMER Segment in EU marketplace with Trendline Degree 2 – MAPE 28.27

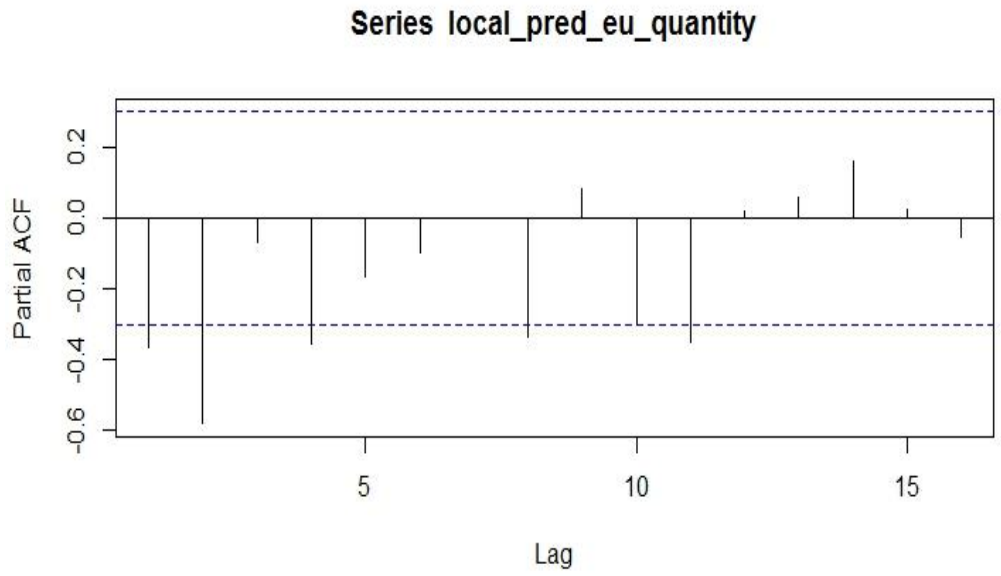


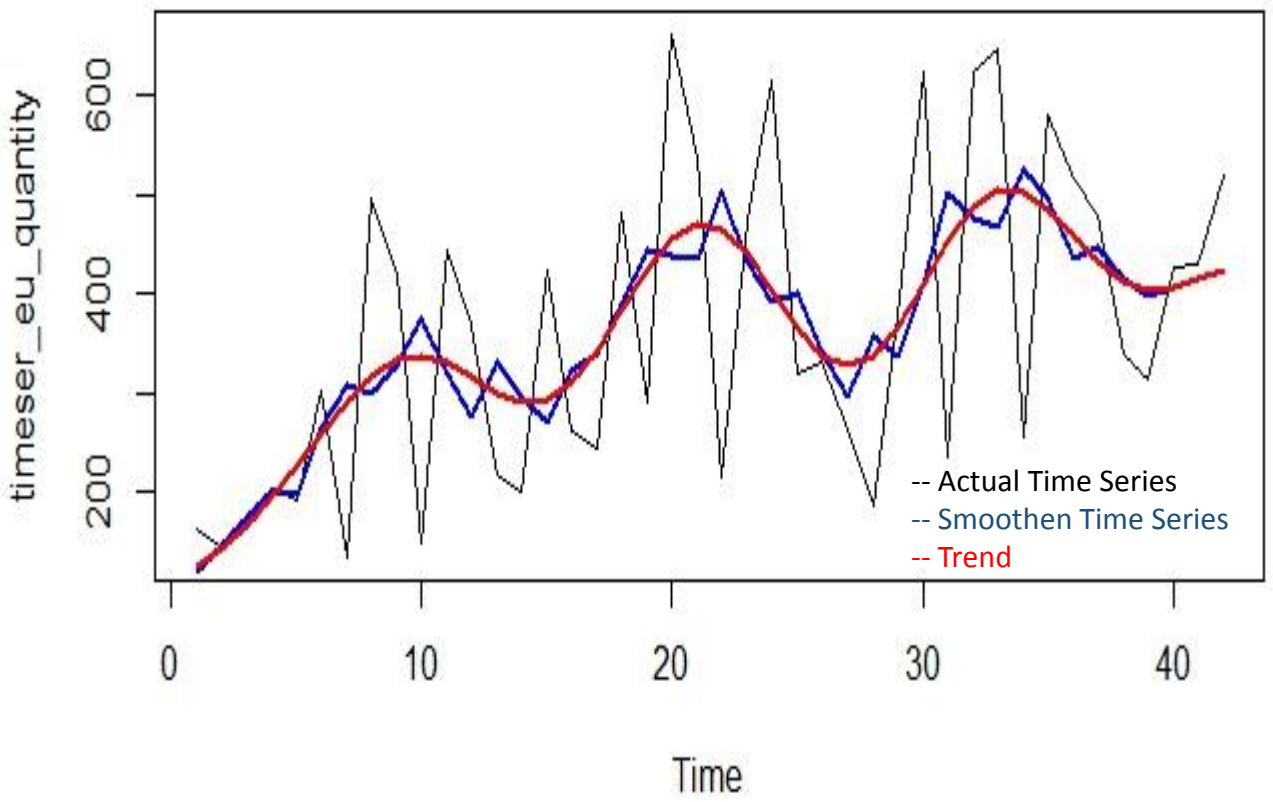




Time Series Analysis of Quantity for CONSUMER Segment in EU marketplace with Trendline Degree 2 – 30.49665







CONCLUSION

- APAC Consumer and EU Consumer are the top two profitable market segments .
- The company should focus on the APAC and Consumer segment as per the prediction of the time series model since the increase in sales is high with small fluctuation.
- As the demand is high in this region, highest number of resources should be deployed in this region. Deployment of resources can be proportional for the other regions too.

THANKS