RETAIL-GIANT SALES FORECASTING CASE STUDY

SUBMISSION

Group Name:

1. Nirmal Maheshwari
2. Vijaya Raman
3. Shwetank Pandey
4. Vijay Mane

Business 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.

Now as a sales/operations manager, you want to finalize the plan for the next 6 months. So, you want to forecast the sales and the demand for the next 6 months, that would help you manage the revenue and inventory accordingly.

The store caters to 7 different market segments and in 3 major categories. You want to forecast at this granular level, so you subset your data into 21 (7*3) buckets before analyzing the 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

Problem Solving Methodology- CRISP DM Framework

Business Objective : Explained Above Slide

Data Understanding The data currently has the transaction level data, where each row represents a particular order made on the online store. There are 24 attributes related to each such transaction. The “Market” attribute has 7-factor levels representing the geographical market sector that the customer belongs to. The “Segment” attribute tells which of the 3 segments that customer belongs to

Data Preparation : Cleaning of the data available by removing the columns having very large NA values and doing necessary EDA . After dividing the data into 21 subsets according to the market and the segment, need to find the two most profitable and consistent segment by using the coefficient of variation as the matrix.

Modelling: Applying the time series model for the top two segments using the classical decomposition and auto-arima method and forecasting the last 6 months sales and demand value for each of the segment

Evaluation: Evaluating the model by using the MAPE value.

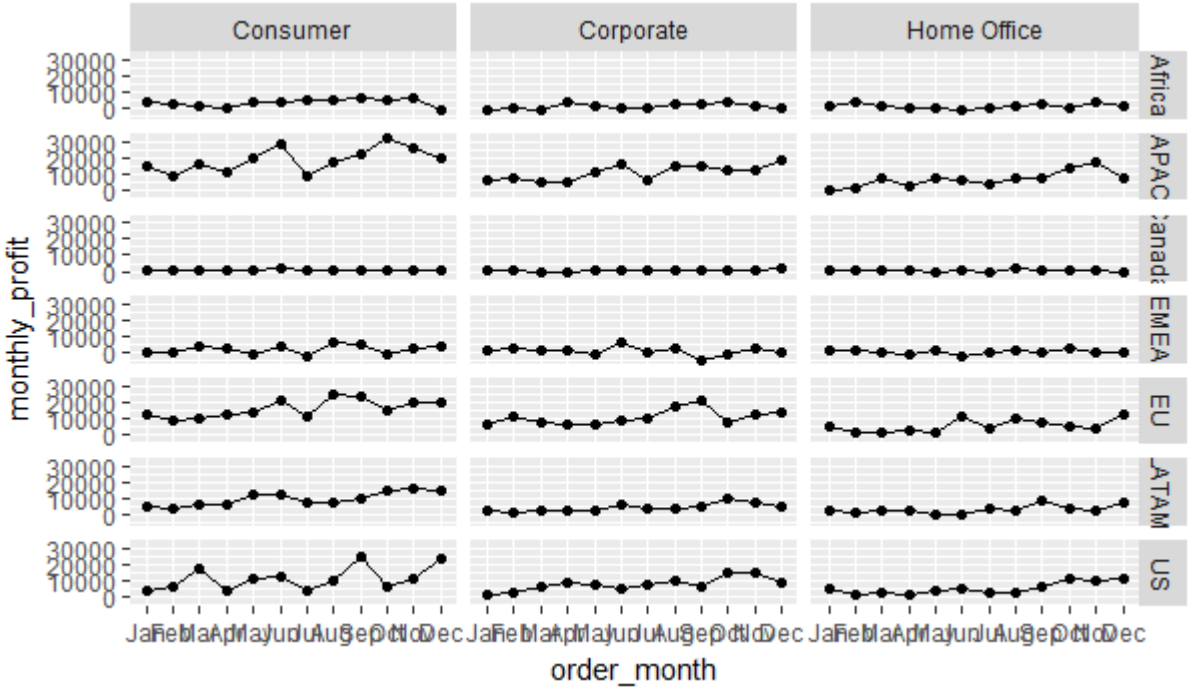
Tools Used : R, Excel

Data Understanding And Preparation

Strategy And Steps Followed

1. Reading CSV Data available in R.
2. Removing the columns which has large percentage of NA values.
3. Finding the unique Market, Segment and the 21 buckets in which the data is divided.
4. Converting the order date in correct format and extracting the month from the date.
5. Aggregating profit on the basis of 21 buckets
6. Finding the top two segments by finding out the coefficient of variation for all the 21 buckets.

Profit Graph for Each Bucket



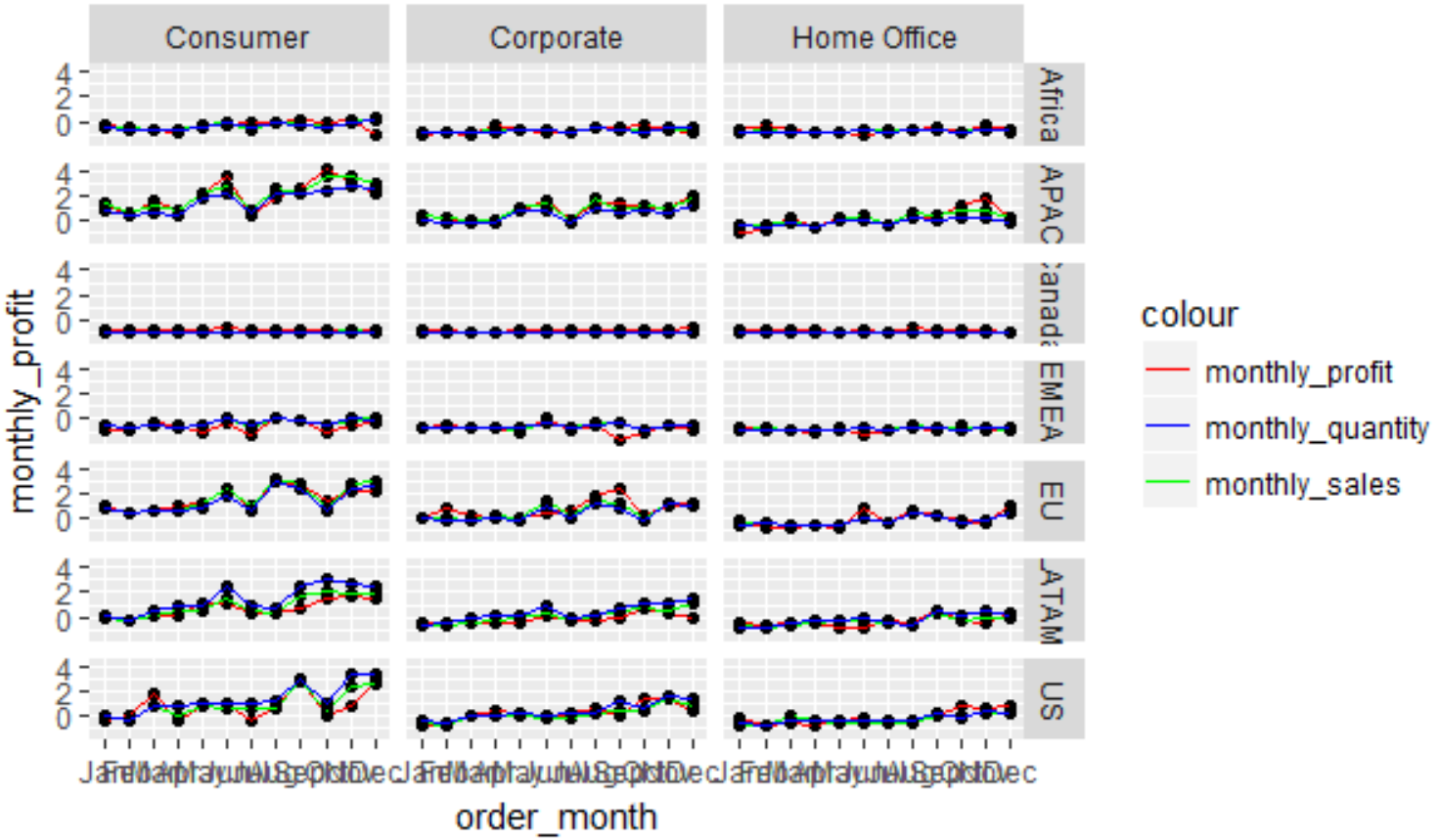
From the graph it is visible that the Canada Market is a straight line, it is not changing with the month, most profitable seems to be the consumer segment for which the graph is showing some trend.

Two Segments

- After finding the coefficient of variation for each of the 21 buckets, the two most profitable and the consistent segment are below.
- Formulae used for calculating coefficient of variation
- Coefficient of variation = standard deviation/ mean

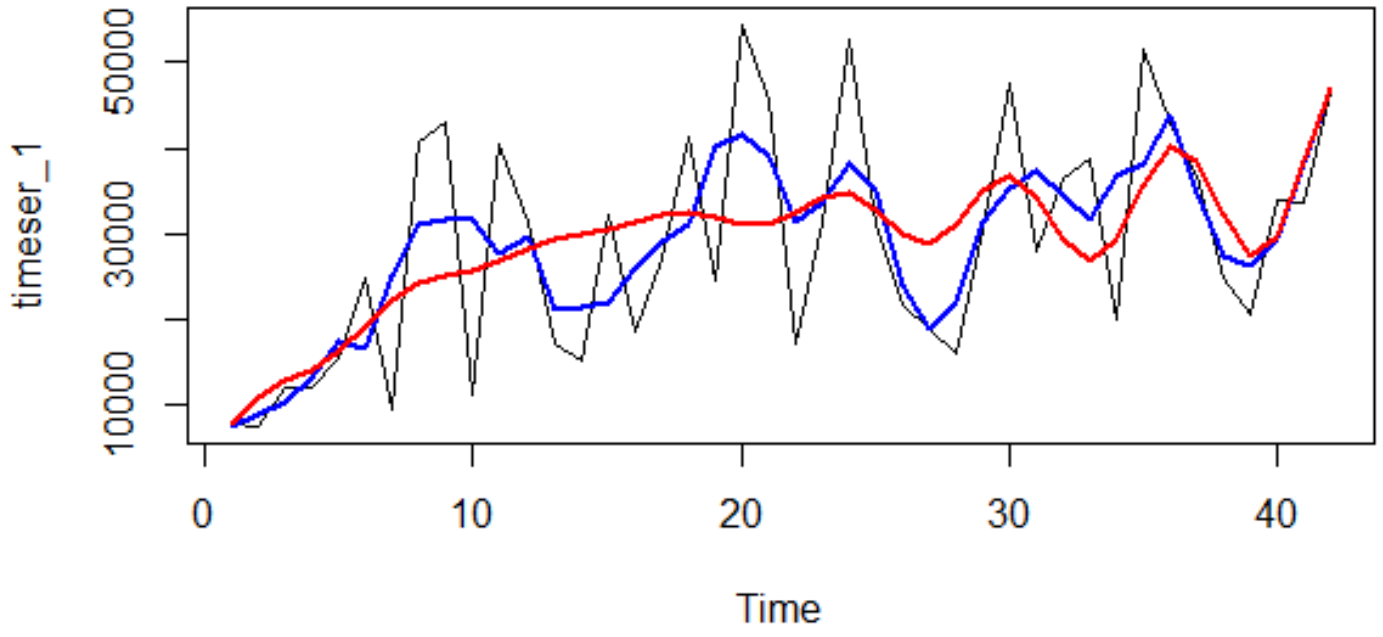
	Market	Segment
1	EU	Consumer
2	APAC	Consumer

Complete Plot for Sales, Profit and Demand

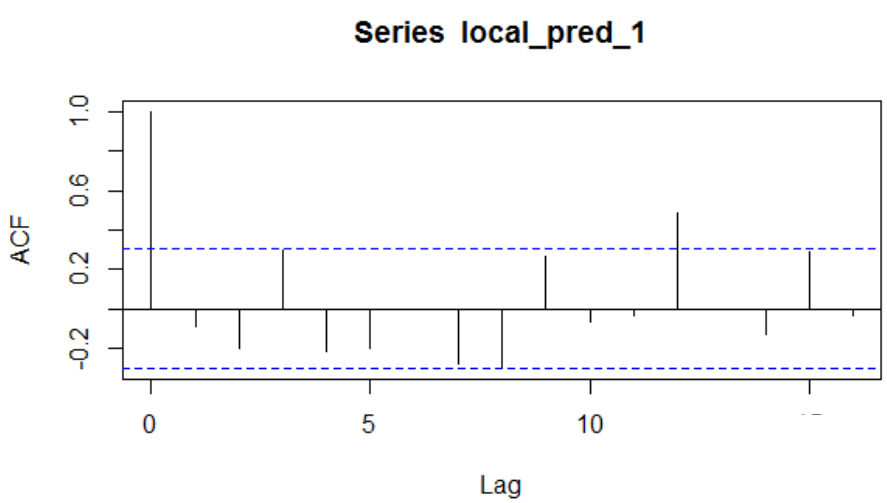


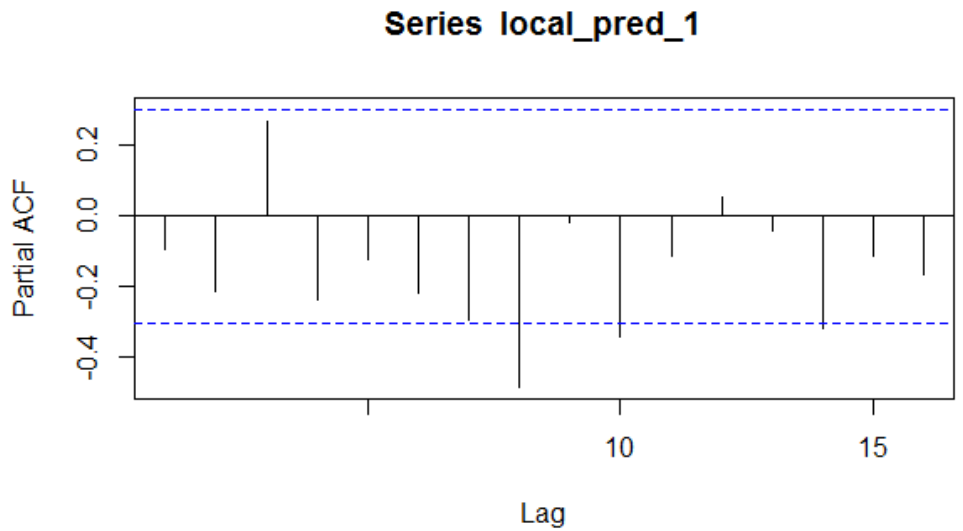
EU Consumer Segment Sales Forecasting

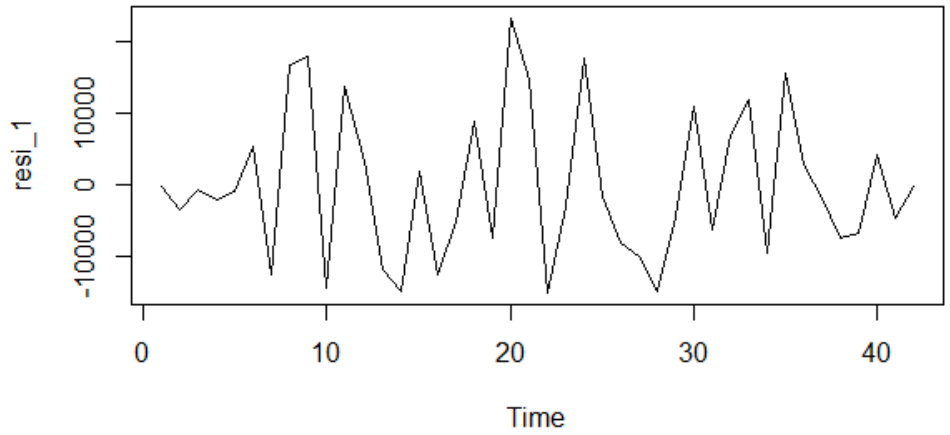
- Time series graph along with the smoothened graph on it shown in blue and the globally predicted part shown in red.



ACF, PACF and Residual Plot



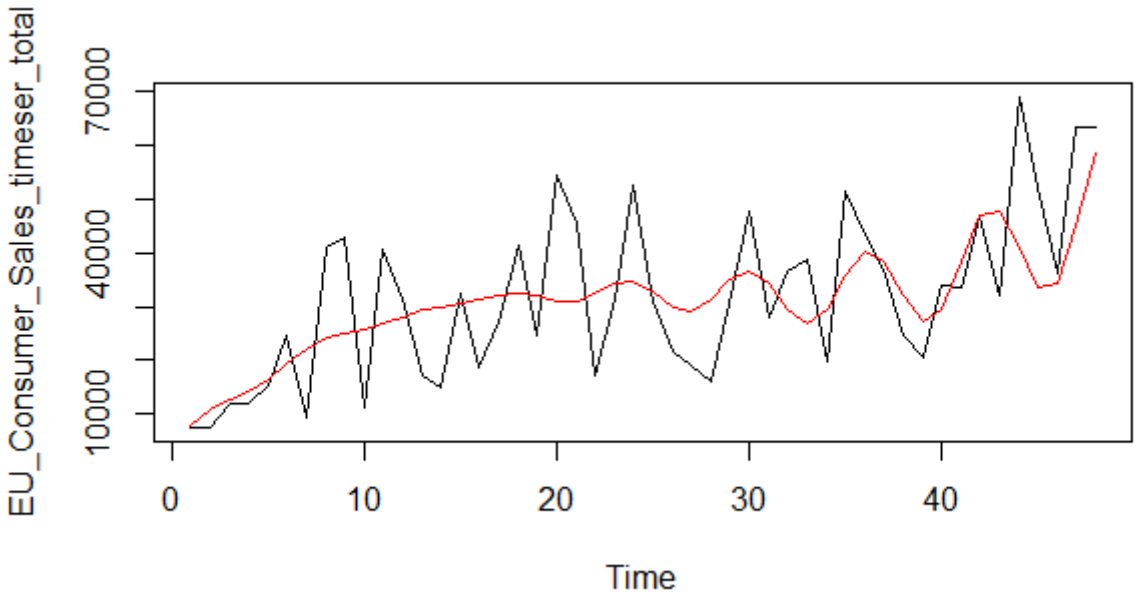




MAPE Value and Complete Forecasting Graph

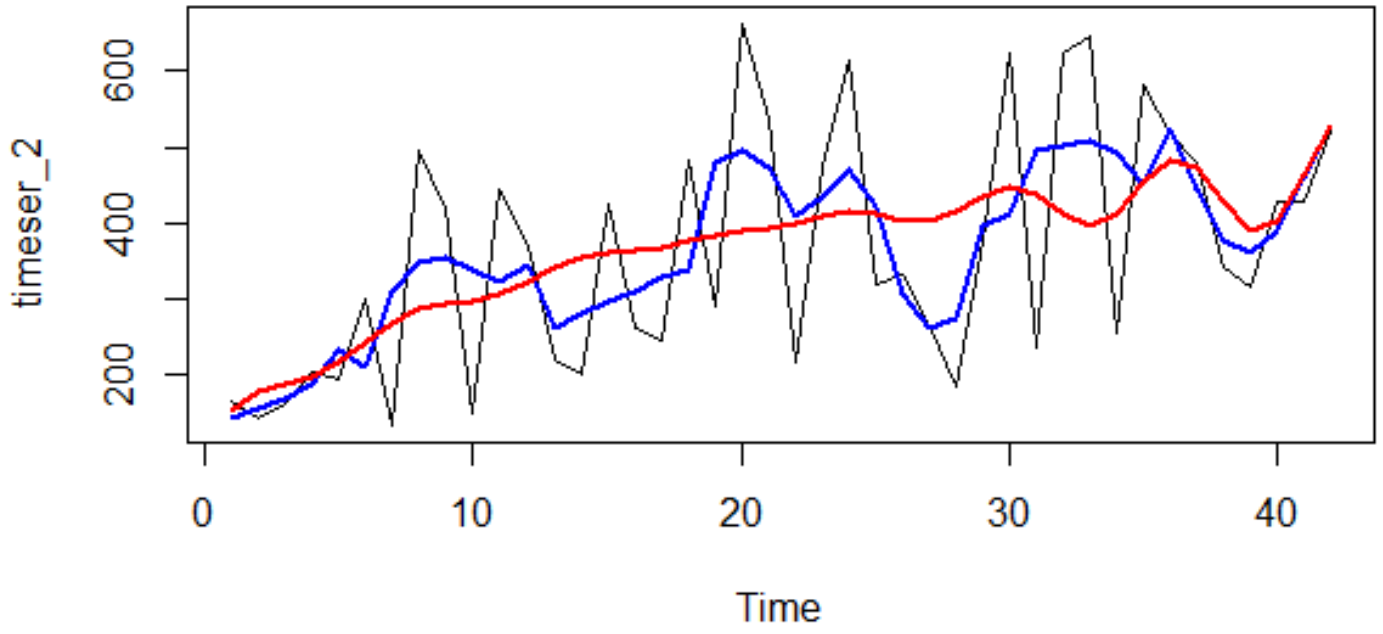
The MAPE value for accuracy of EU Consumer Sales prediction for the first 42 months comes out to be **28.04386** which is quite low.

Below is the complete graph with the forecasting line.

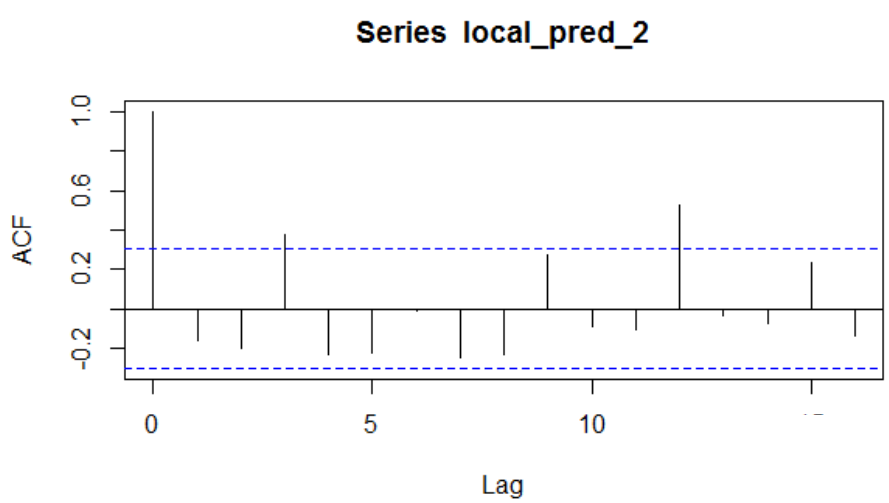


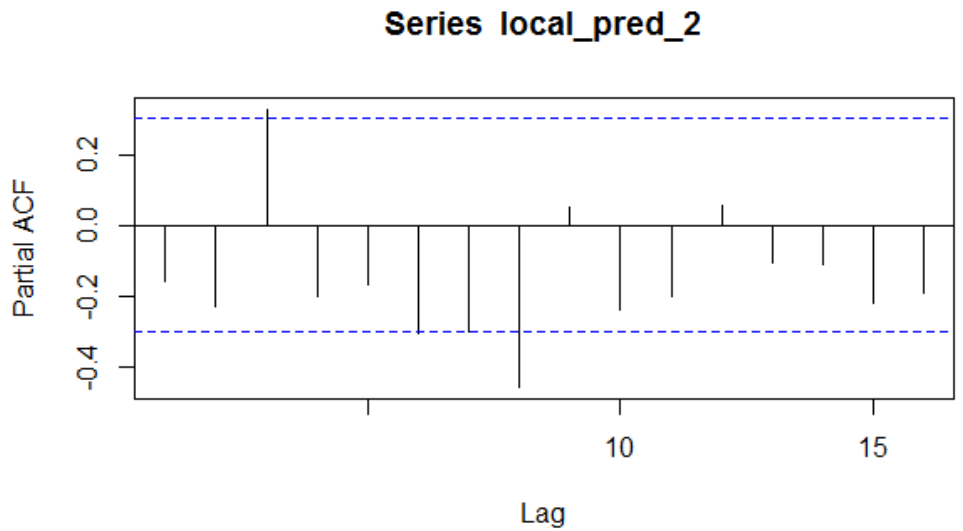
EU Consumer Segment Demand Forecasting

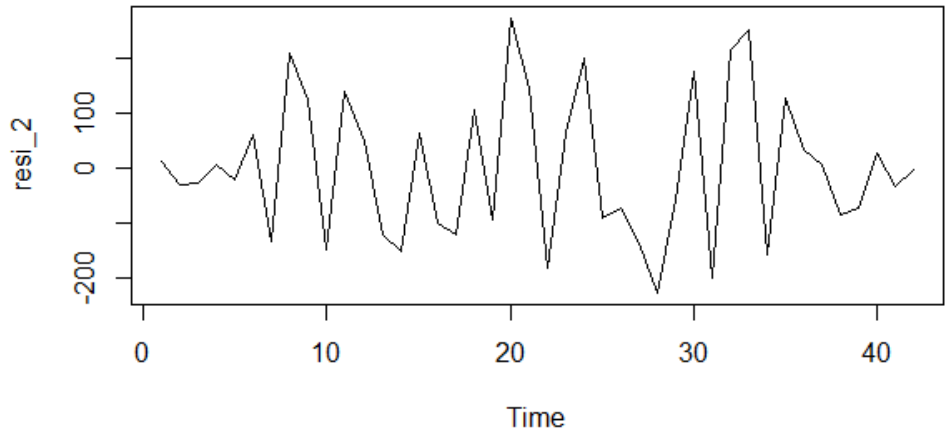
- Time series graph along with the smoothened graph on it shown in blue and the globally predicted part shown in red.



ACF, PACF and Residual Plot



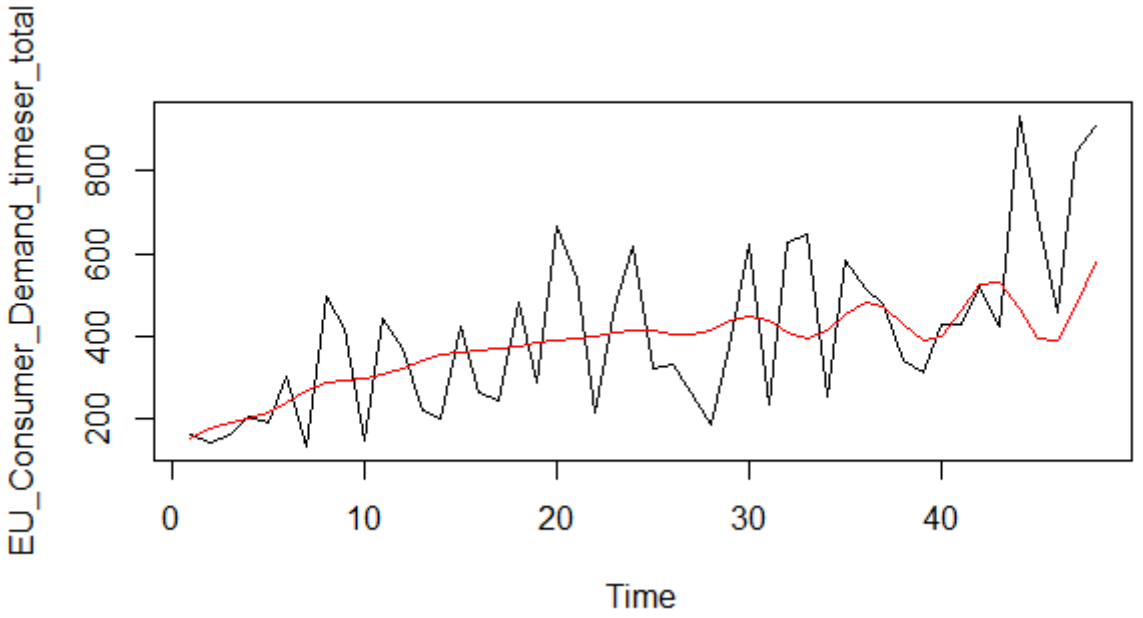




MAPE Value and Complete Forecasting Graph

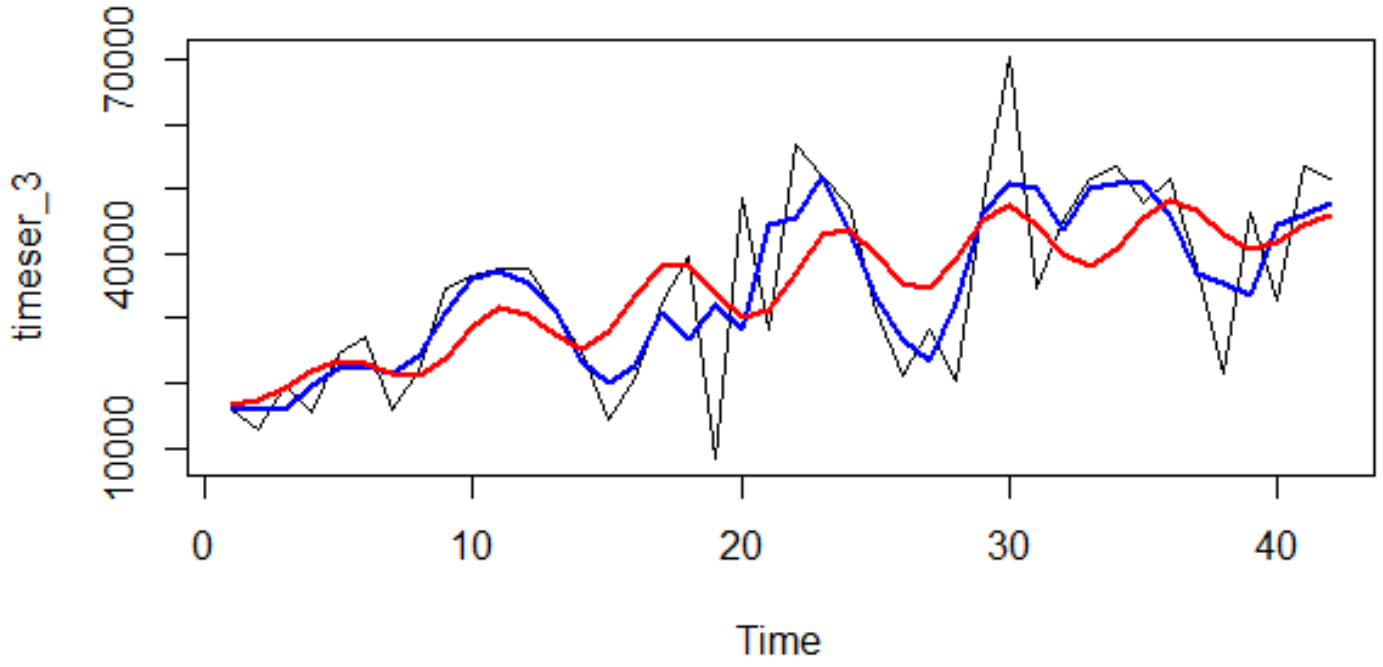
The MAPE value for accuracy of EU Consumer Demand prediction for the first 42 months comes out to be **35.31684** which is quite low.

Below is the complete graph with the forecasting line.

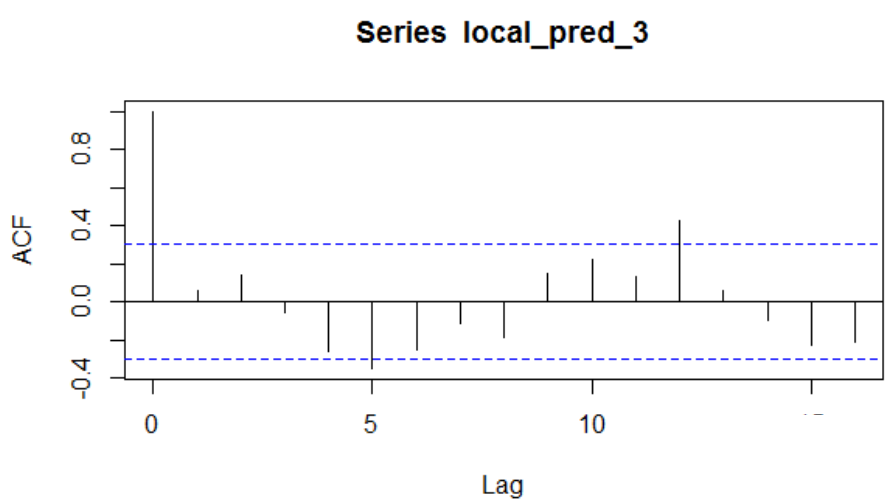


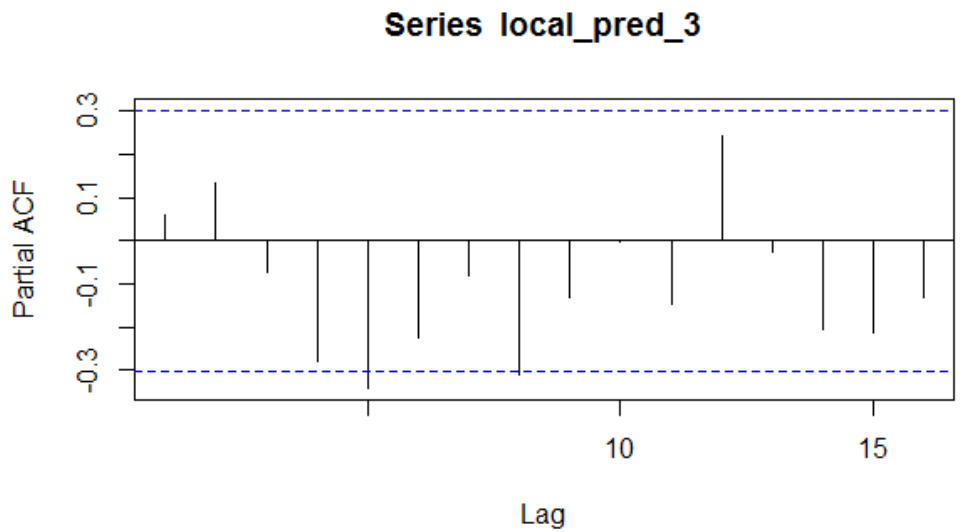
APAC Consumer Segment Sales Forecasting

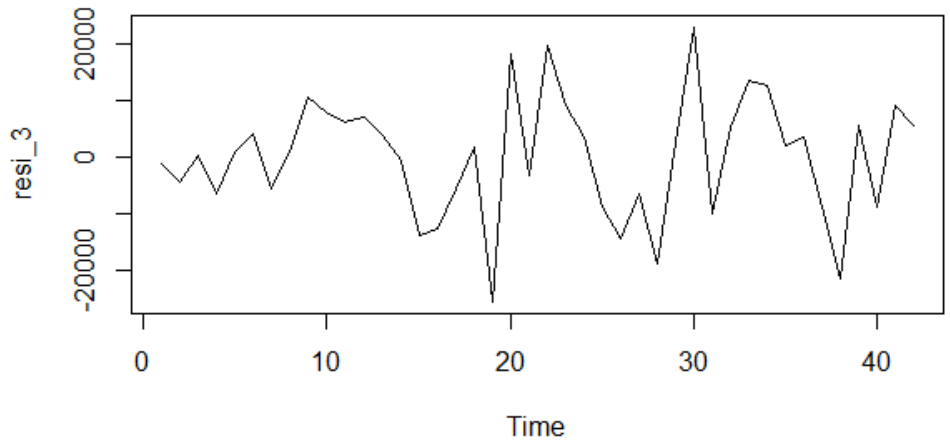
- Time series graph along with the smoothened graph on it shown in blue and the globally predicted part shown in red.



ACF, PACF and Residual Plot



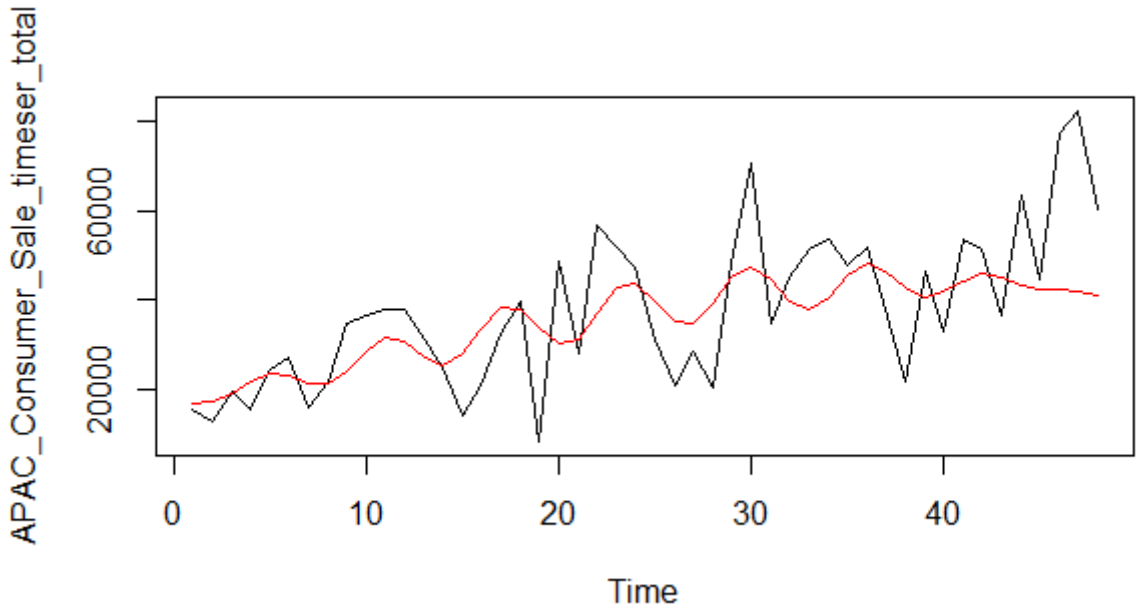




MAPE Value and Complete Forecasting Graph

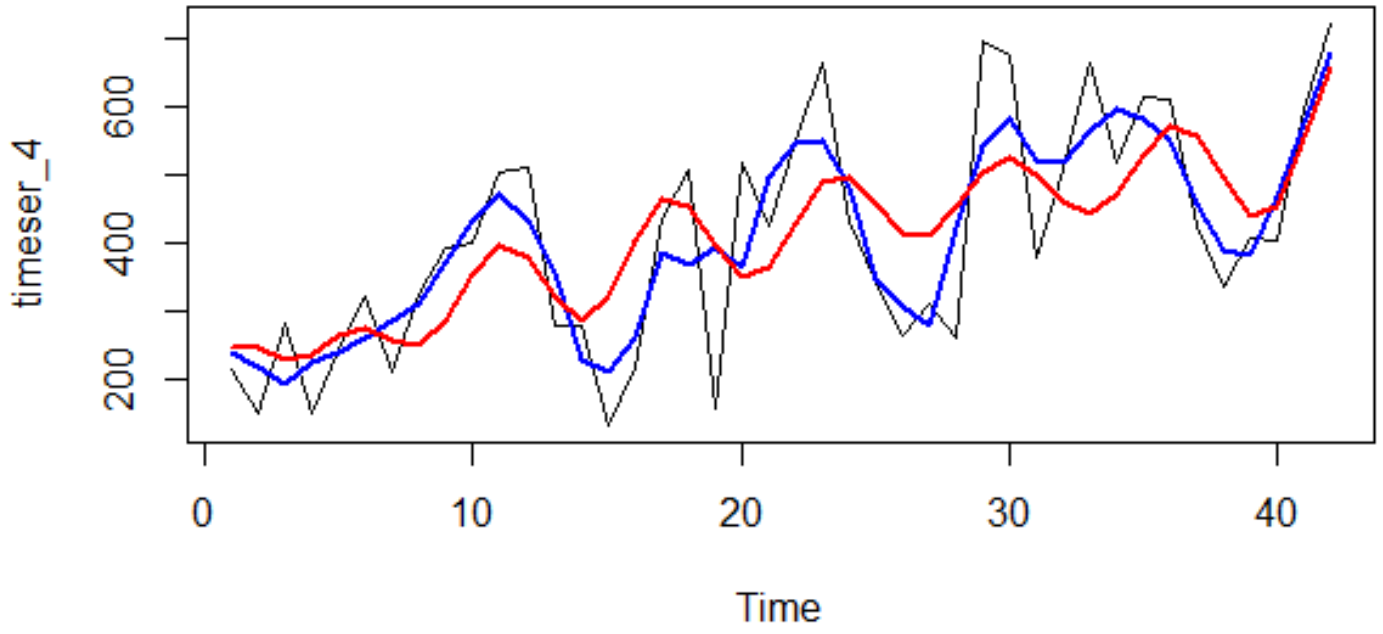
The MAPE value for accuracy of EU Consumer Demand prediction for the first 42 months comes out to be **31.15703** which is quite low.

Below is the complete graph with the forecasting line.

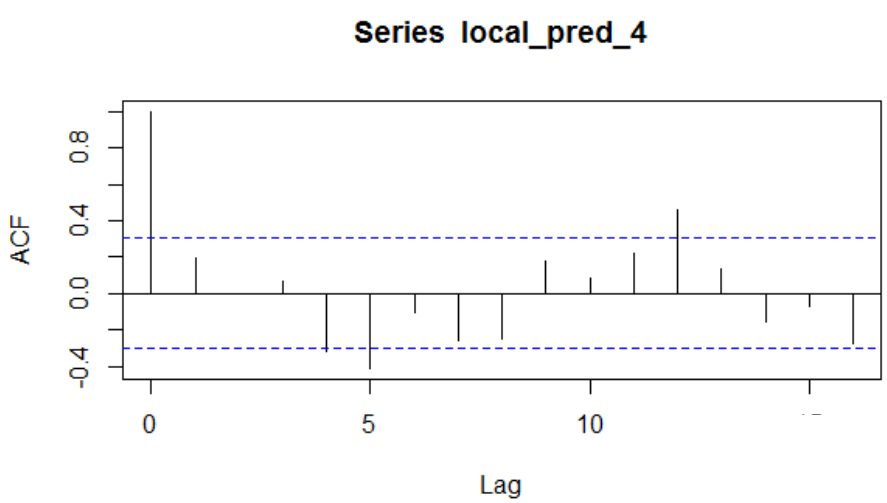


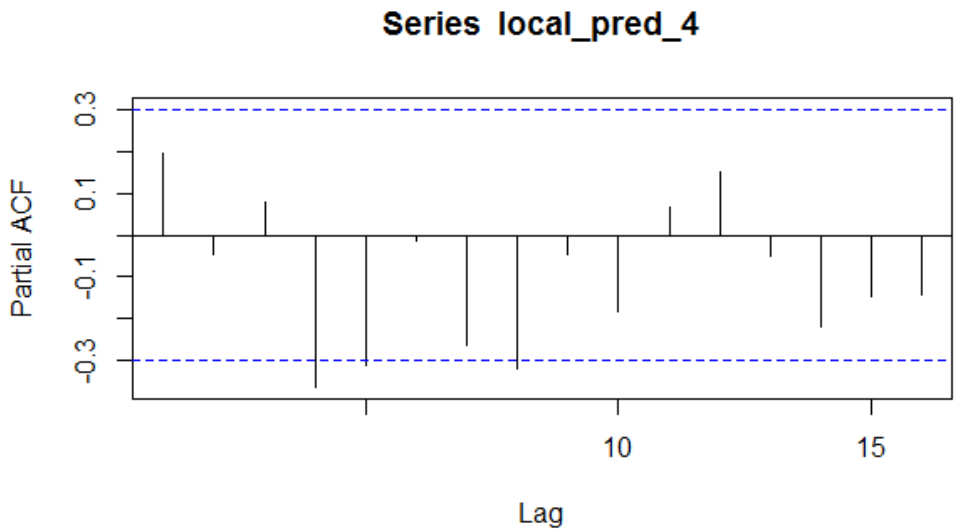
APAC Consumer Segment Demand Forecasting

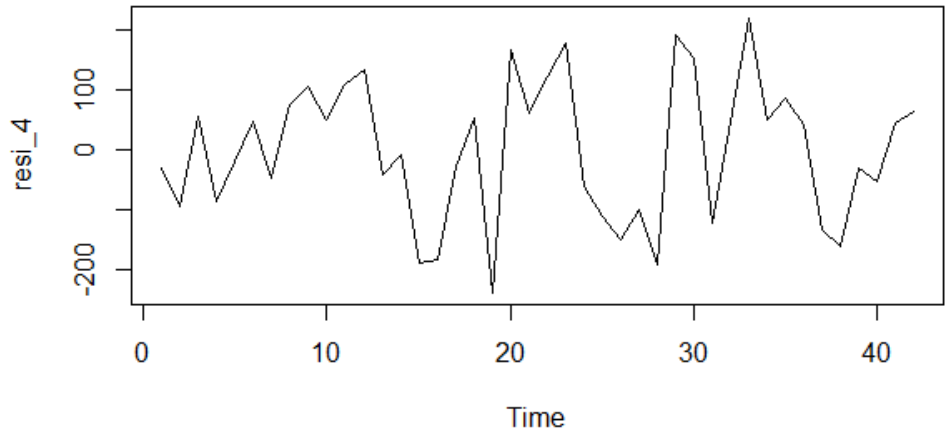
- Time series graph along with the smoothened graph on it shown in blue and the globally predicted part shown in red.



ACF, PACF and Residual Plot



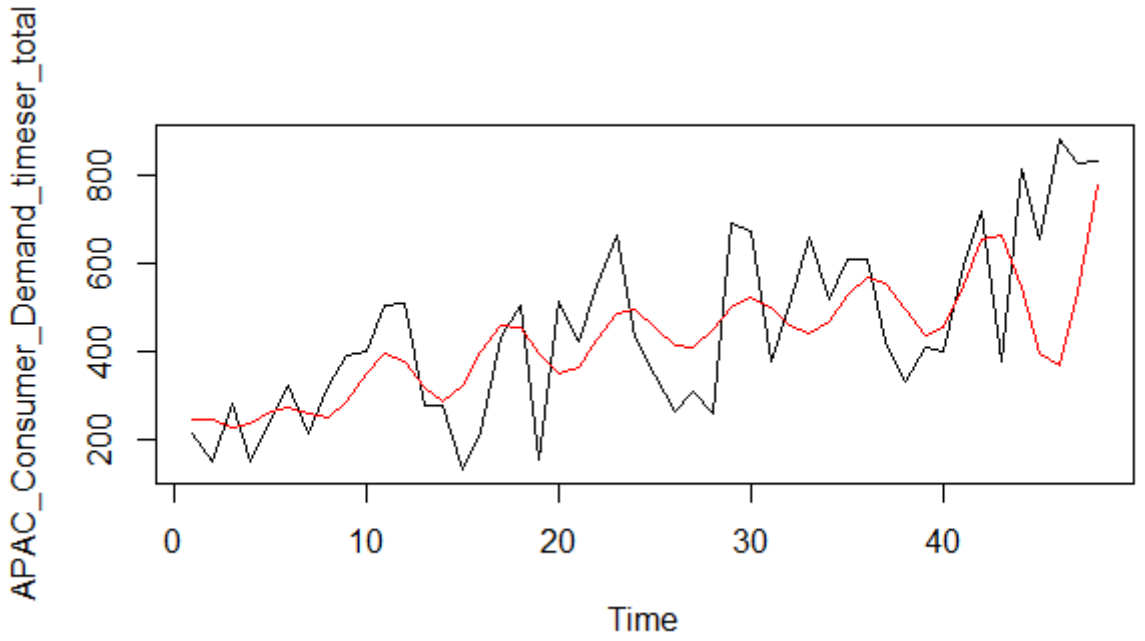




MAPE Value and Complete Forecasting Graph

The MAPE value for accuracy of EU Consumer Demand prediction for the first 42 months comes out to be **41.49113** which is quite low.

Below is the complete graph with the forecasting line.



Conclusion

- The two most profitable buckets out of 21 are EU, Consumer and APAC, Consumer segment.
- The MAPE values of all the predictions are low.
- The multiplicative model has been followed for modeling out the globally predictable part for both the segments
- The ACF,PACF plots for all the segments shows that the locally predictable part is itself weakly stationary and residuals for all the segments came out to be pure white noise.