GLOBAL MART FORECAST CASE STUDY

SUBMISSION

Group Name:

1. Vikash Sharma
2. Vivek Sachdeva
3. Nilaksh Bajpai

Business Understanding

Problem Statement:

- Global Mart, an online store super giant, takes orders and delivers across the globe. It caters to 7 market segments (Africa, APAC, Canada, EMEA, EU, LATAM & US) and 3 categories (Consumer, Corporate & Home Office). At a granular level, not all of these 21 (7*3=21) buckets are important from the store's point of view.
- We need to find out 2 most profitable (and consistent) segment from these 21 and forecast the sales and demand for these segments.
- To manage revenue and inventory, Sales/Operations Manager wants to forecast:
  - Sales for next 6 months
  - Demand for next 6 months

Analysis Objectives:

- Find the 2 most profitable and consistently profitable segments based on coefficient of variation in each bucket
- Forecast sales and quantity for next 6 months using classical decomposition and auto ARIMA.
- Evaluate Forecast Accuracy based on MAPE (Mean Absolute Percentage Error)

Analysis Approach

Data Understanding & Preparation

- Convert Transaction level data into Time Series
- Aggregate over Order Date to arrive at monthly values for Sales, Quantity and Profit

Calculate Coefficient of Variation (COV) for each Market and Segment bucket

Choose the 2 buckets with least COV as the most consistently profitable ones for forecasting sales and demand

Model Creation

Split Data into first 42 months as Training set and last 6 months as Testing set.

Smoothen training data and Forecast using Classical Decomposition

Forecast using Auto ARIMA method

Model Evaluation & Deployment

Run Classical Decomposition and Auto ARIMA models on 6 months Testing data

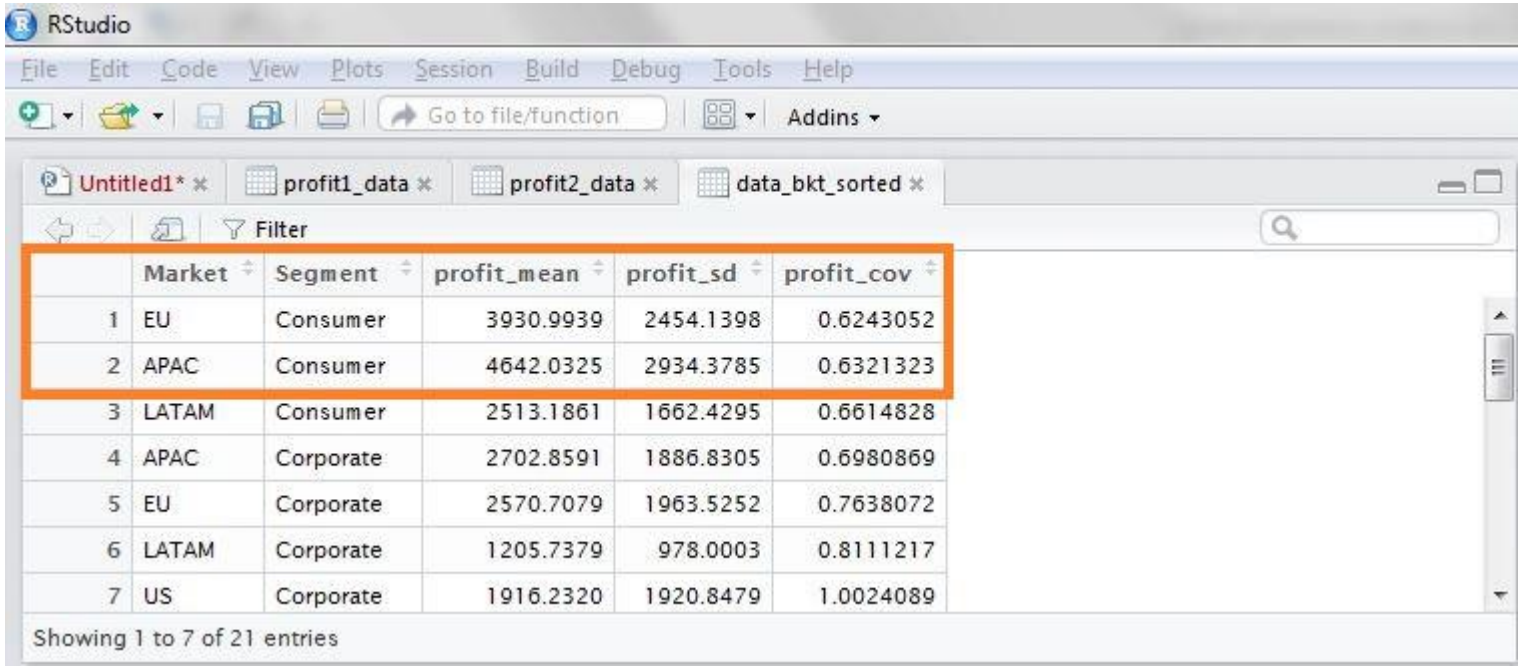
Calculate MAPE for both Classical Decomposition and Auto ARIMA Methods to measure forecast accuracy

Choose method with lower MAPE as the recommended forecast

Data Understanding & Preparation

- **Continuous Variables to be Forecasted:** Sales, Quantity and Profit
- **Time Series Indicator Variable:** Order Date.
- **Categorical Variables:**
  - Market – 7 levels: Africa, APAC, Canada, EMEA, EU, LATAM, US
  - Segment – 3 levels: Consumer, Corporate, Home Office
- **Redundant Variables:** All other variables have not been considered for our analysis.
- **Derived Metrics/Variables:**
  - Months (Order Month since Jan 2011) – Number of months from Jan 2011 (first month of orders) to the month in which order has been placed
  - Profit_Mean – Mean of monthly level profits calculated for each market and segment.
  - Profit_SD – Standard deviation of monthly level profits calculated for each market and segment.
  - Profit_COV (Coefficient of Variation) – This is a measure of consistency in profits. It is equal to the standard deviation divided by the mean. Here, coefficient of variation (COV) is calculated for each market and segment. The market and segment with the least COV is the most consistently profitable one.

Most Consistently Profitable Segments



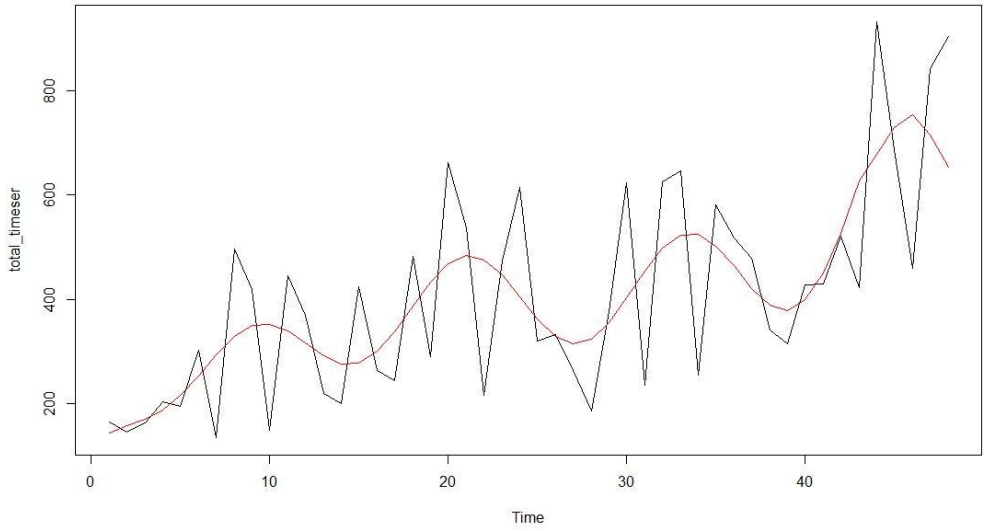
	Market	Segment	profit_mean	profit_sd	profit_cov
1	EU	Consumer	3930.9939	2454.1398	0.6243052
2	APAC	Consumer	4642.0325	2934.3785	0.6321323
3	LATAM	Consumer	2513.1861	1662.4295	0.6614828
4	APAC	Corporate	2702.8591	1886.8305	0.6980869
5	EU	Corporate	2570.7079	1963.5252	0.7638072
6	LATAM	Corporate	1205.7379	978.0003	0.8111217
7	US	Corporate	1916.2320	1920.8479	1.0024089

Showing 1 to 7 of 21 entries

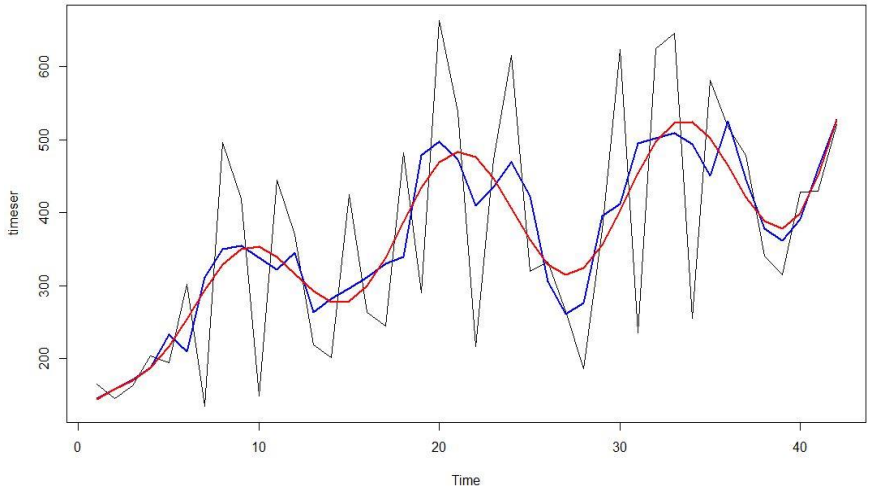
- Most consistently profitable buckets have been calculated based on the coefficient of variation (COV).
- Coefficient of Variation = Standard deviation/Mean. So, the lesser the coefficient of variation, the greater is the consistency.
- As shown in picture above, the two most consistently profitable market-segment buckets (having least COV) are:
  - ✓ Most consistently profitable customer is EU Consumer
  - ✓ 2nd most consistently profitable customer is APAC Consumer

EU Consumer Demand – Classical Decomposition

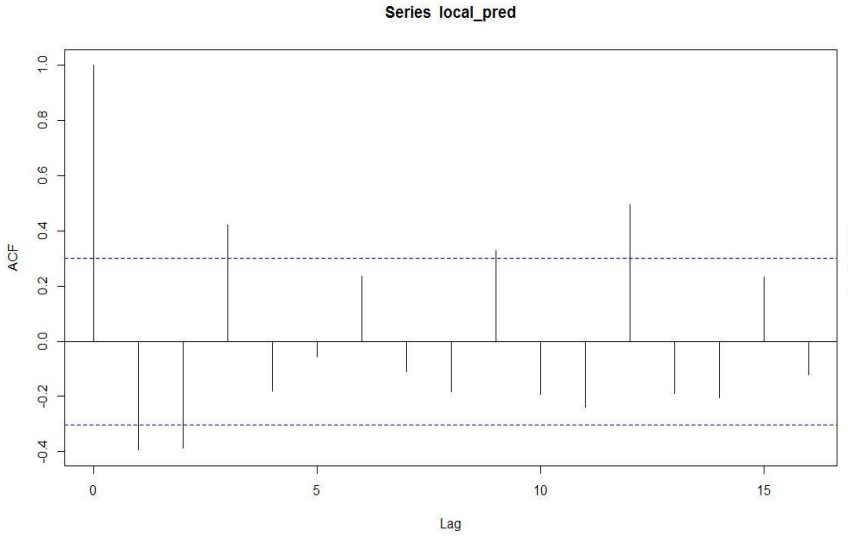
Model with Forecast as Red Line



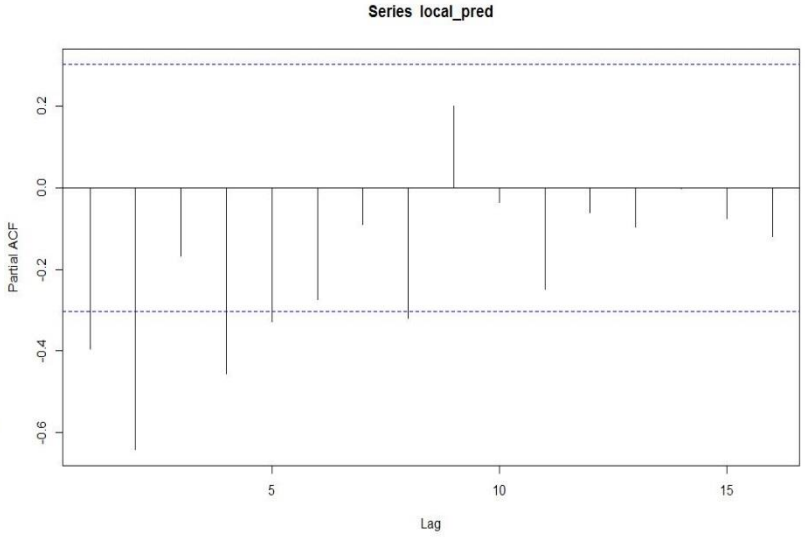
- ✓ Global Model: Multiplicative Trend with Sinusoidal Seasonality
- ✓ Local Model: ARIMA(2,0,0) with zero mean
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 31.45475



Global Model – Blue for Smoothing, Red for Forecast



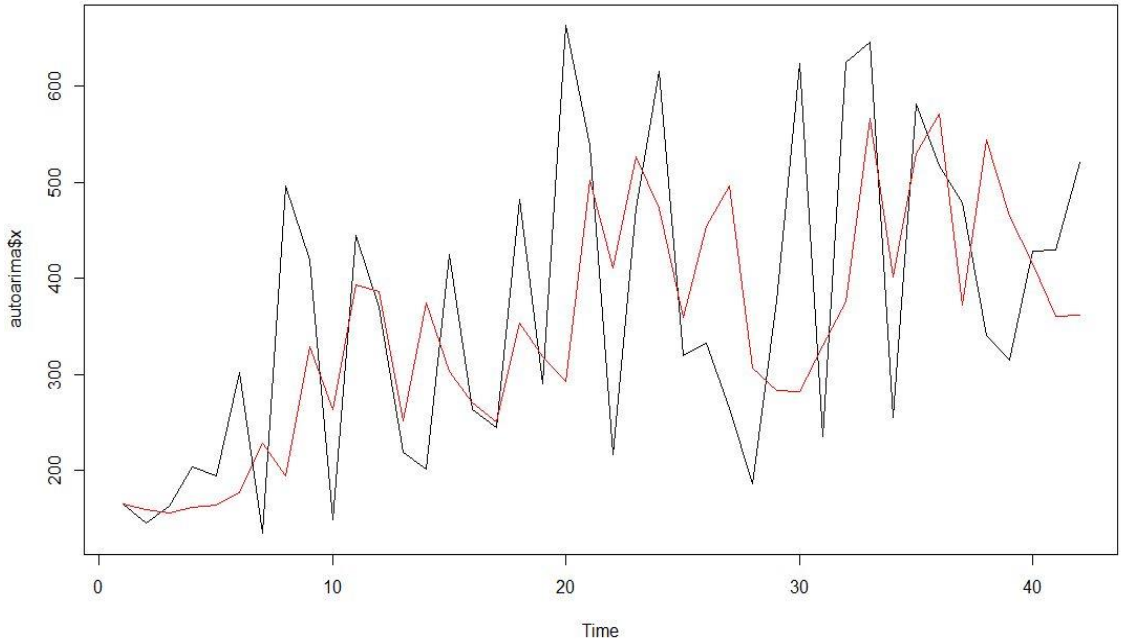
Local Model ARIMA – ACF



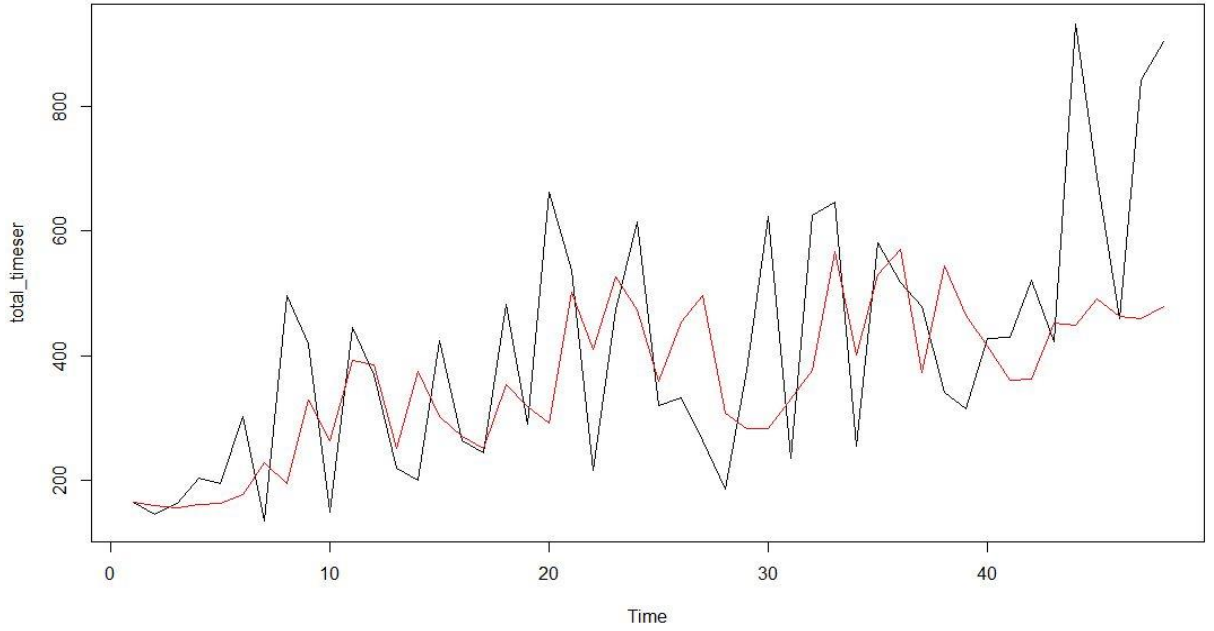
Local Model ARIMA – PACF

EU Consumer Demand – Auto ARIMA

Auto ARIMA Model



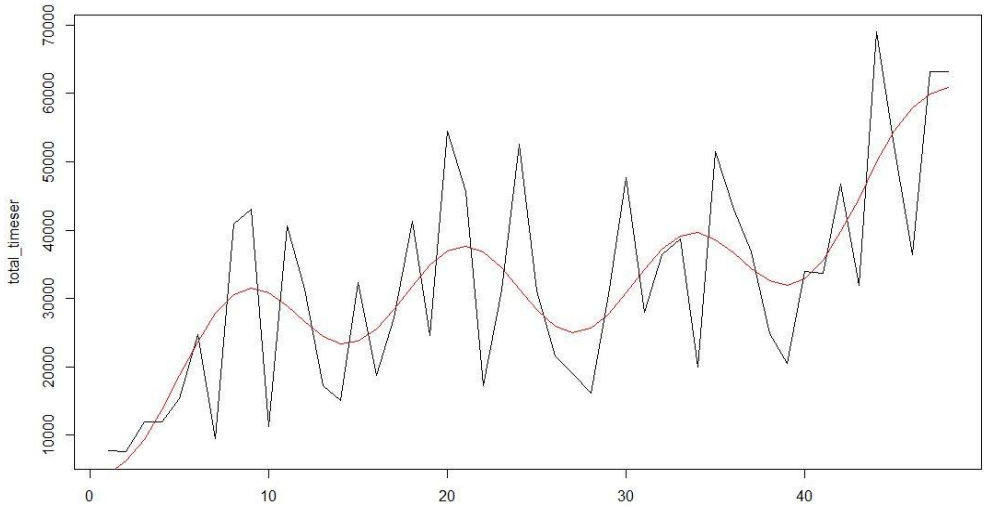
Model with Forecast as Red Line



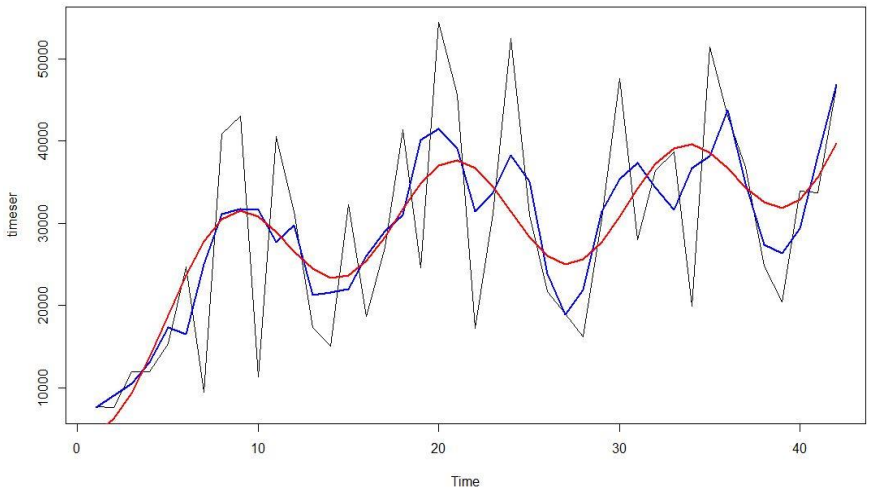
- ✓ Auto Model: ARIMA(2,1,0)
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 30.13319

EU Consumer Sales – Classical Decomposition

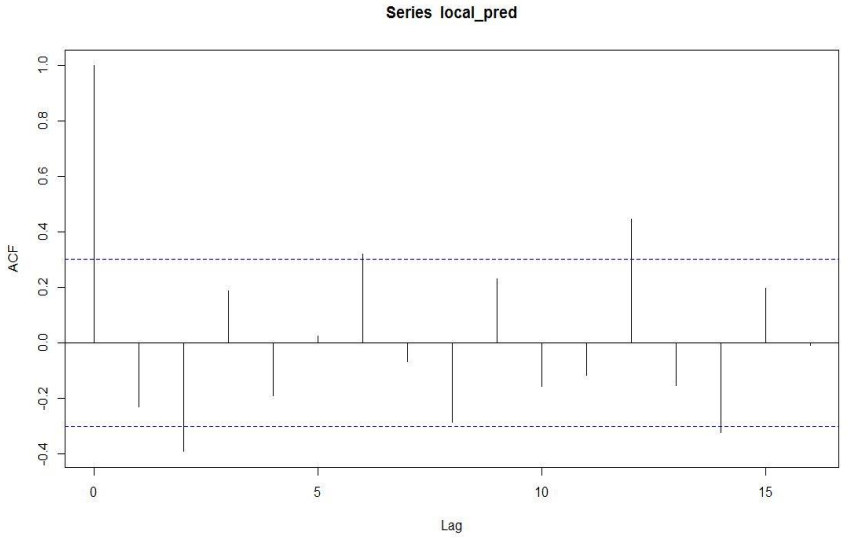
Model with Forecast as Red Line



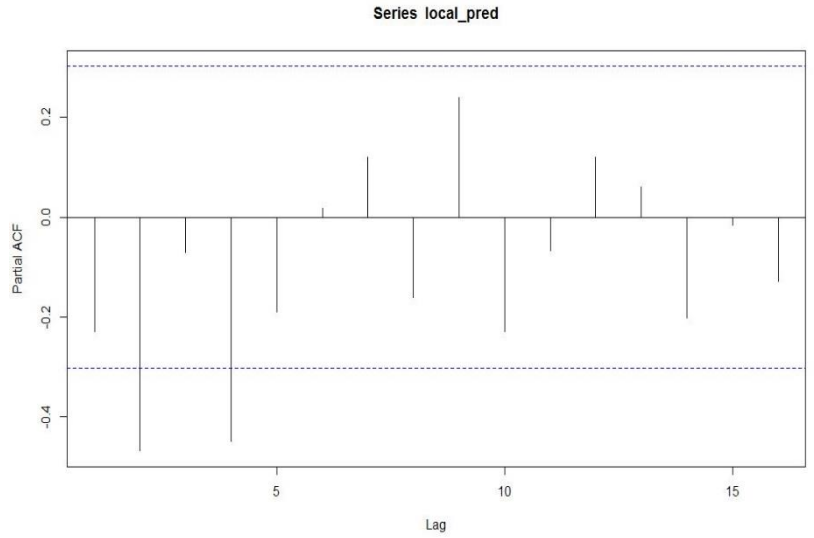
- ✓ Global Model: Additive Trend with Sinusoidal Seasonality
- ✓ Local Model: ARIMA(0,0,1) with zero mean
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 23.18592



Global Model – Blue for Smoothing, Red for Forecast



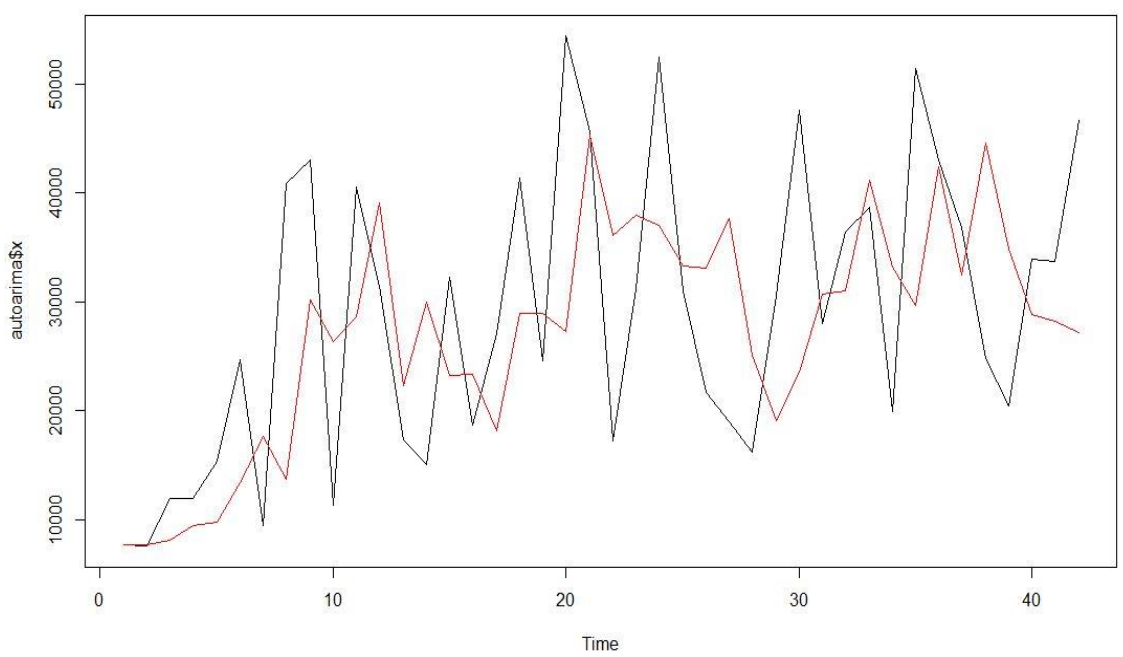
Local Model ARIMA – ACF



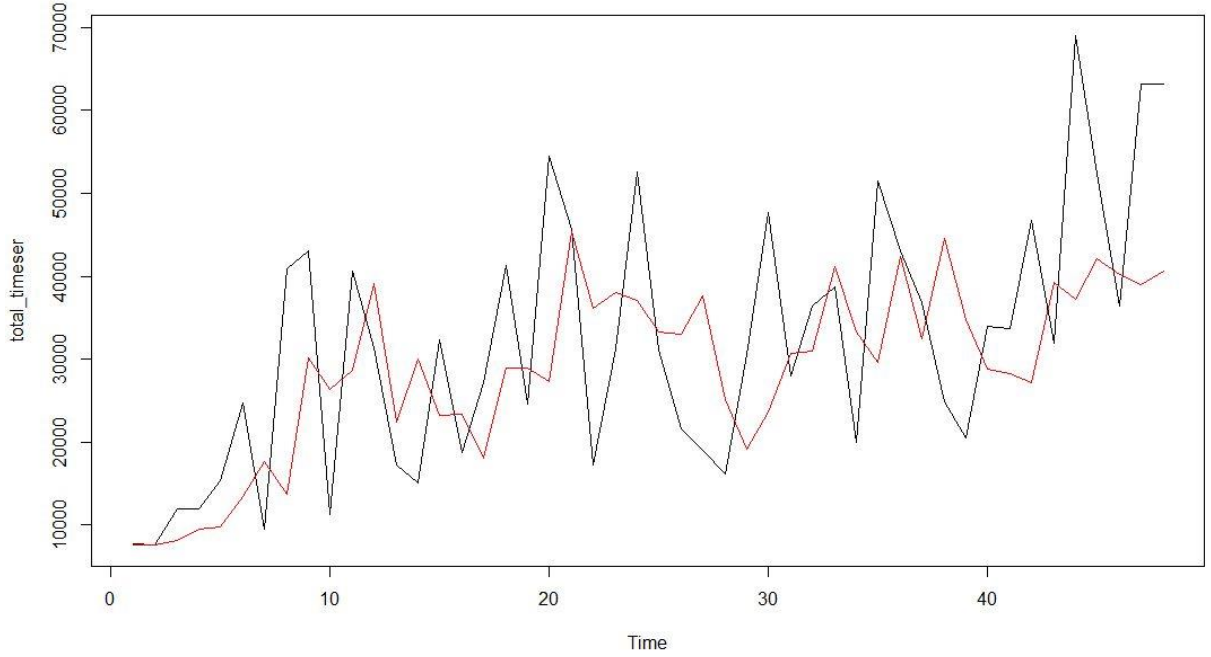
Local Model ARIMA – PACF

EU Consumer Sales – Auto ARIMA

Auto ARIMA Model



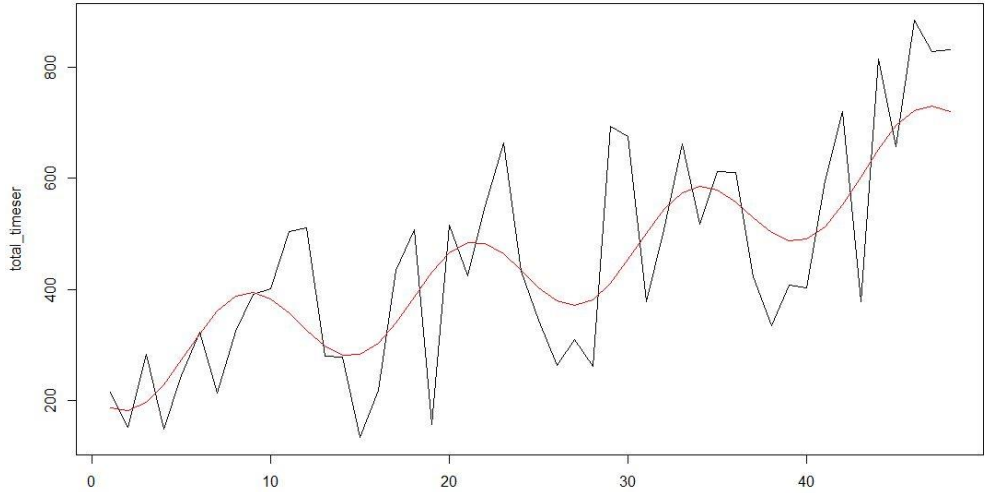
Model with Forecast as Red Line



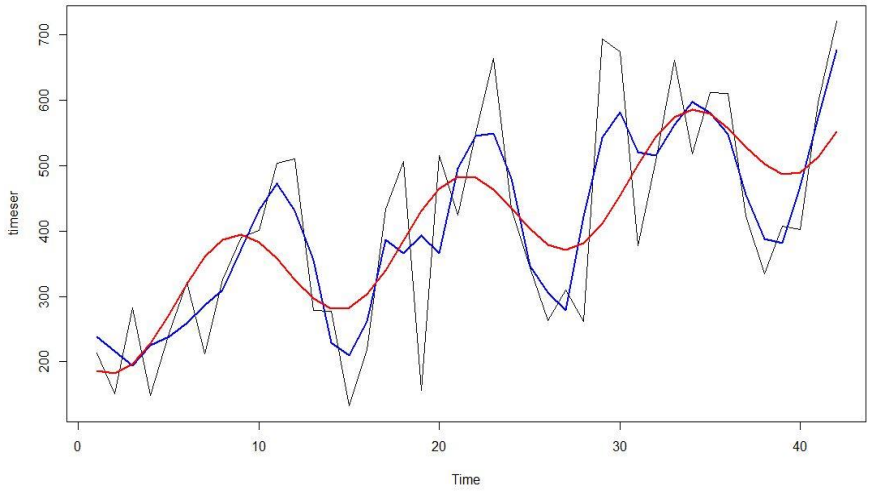
- ✓ Auto Model: ARIMA(2,1,0)
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 28.9226

APAC Consumer Demand – Classical Decomposition

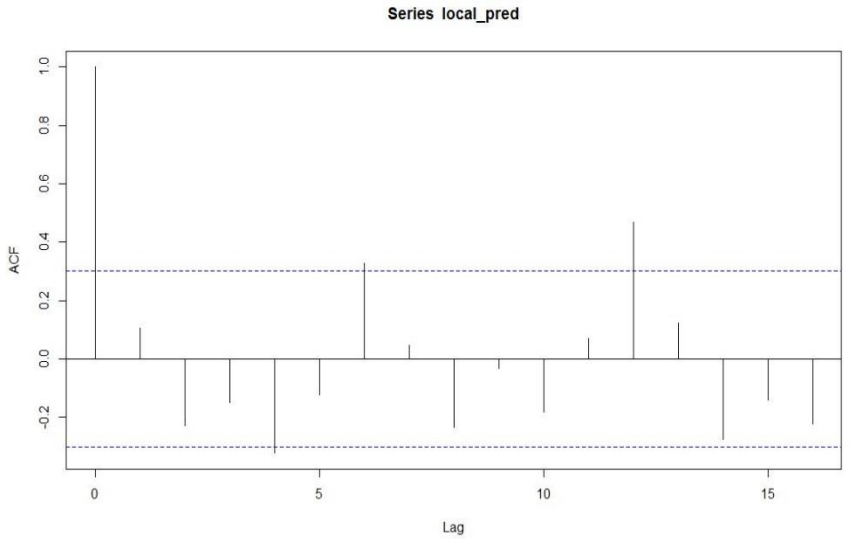
Model with Forecast as Red Line



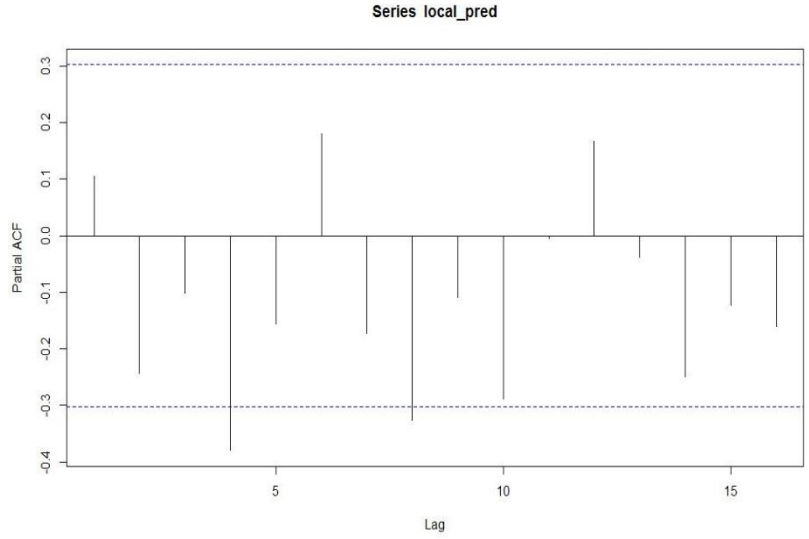
- ✓ Global Model: Additive Trend with Sinusoidal Seasonality
- ✓ Local Model: ARIMA(0,0,0) with zero mean (No local component)
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 21.53303



Global Model – Blue for Smoothing, Red for Forecast



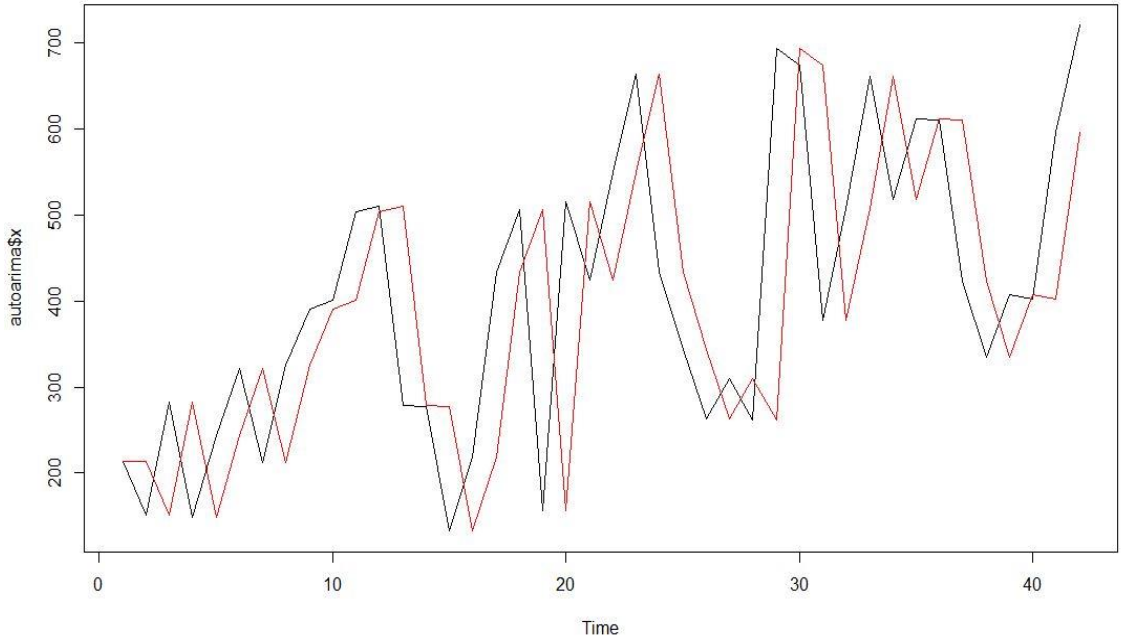
Local Model ARIMA – ACF



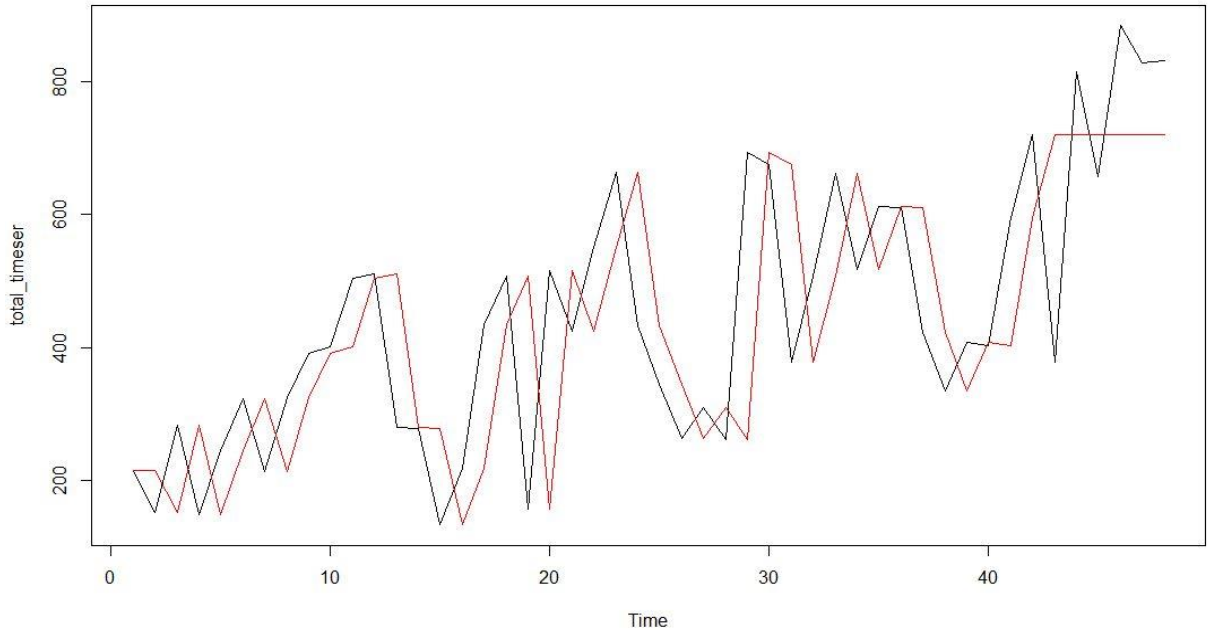
Local Model ARIMA – PACF

APAC Consumer Demand – Auto ARIMA

Auto ARIMA Model



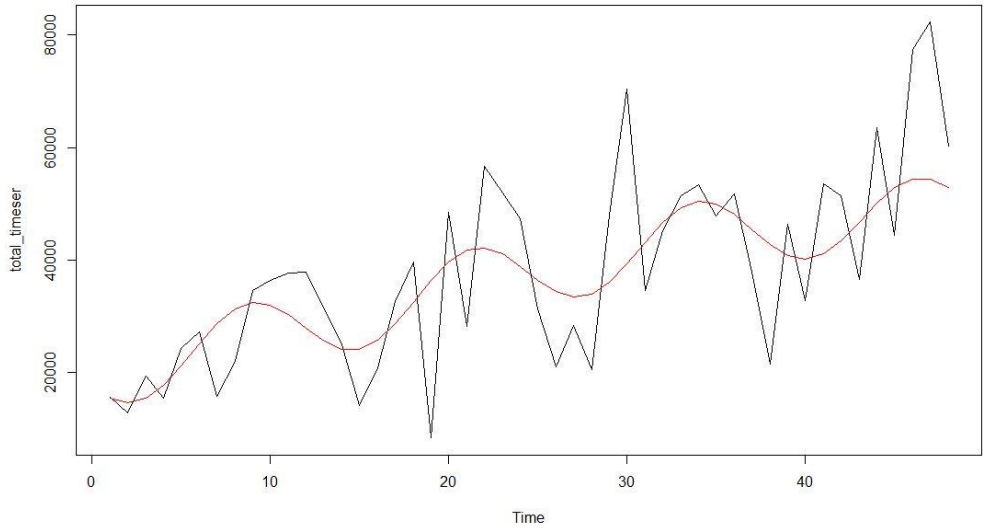
Model with Forecast as Red Line



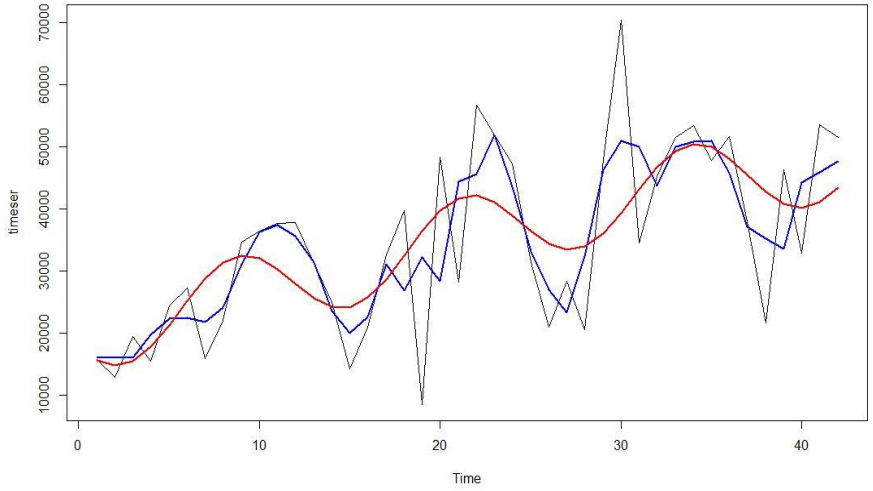
- ✓ Auto Model: ARIMA(0,1,0)
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 26.24458

APAC Consumer Sales – Classical Decomposition

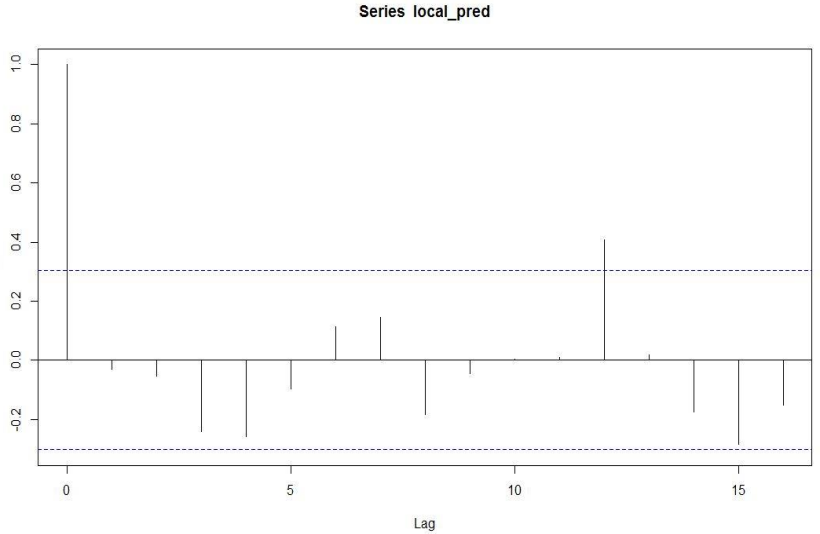
Model with Forecast as Red Line



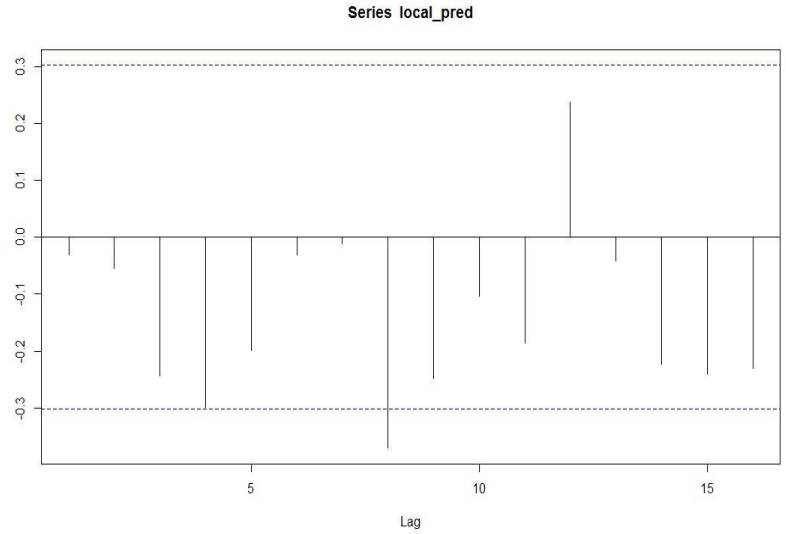
- ✓ Global Model: Additive Trend with Sinusoidal Seasonality
- ✓ Local Model: ARIMA(0,0,0) with zero mean (No local component)
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 23.93068



Global Model – Blue for Smoothing, Red for Forecast



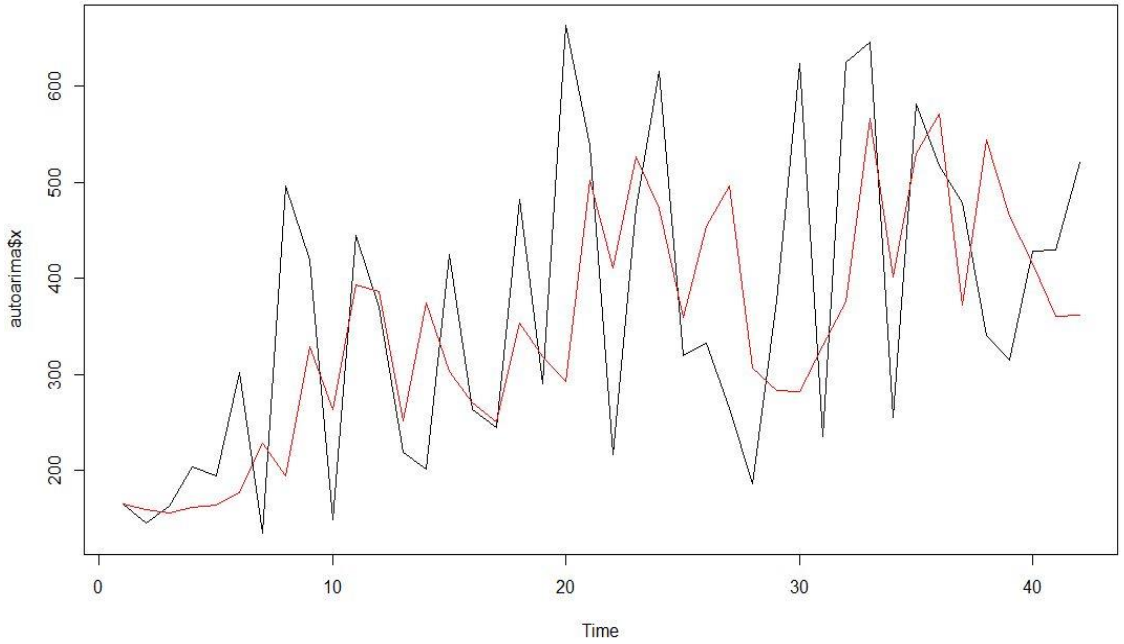
Local Model ARIMA – ACF



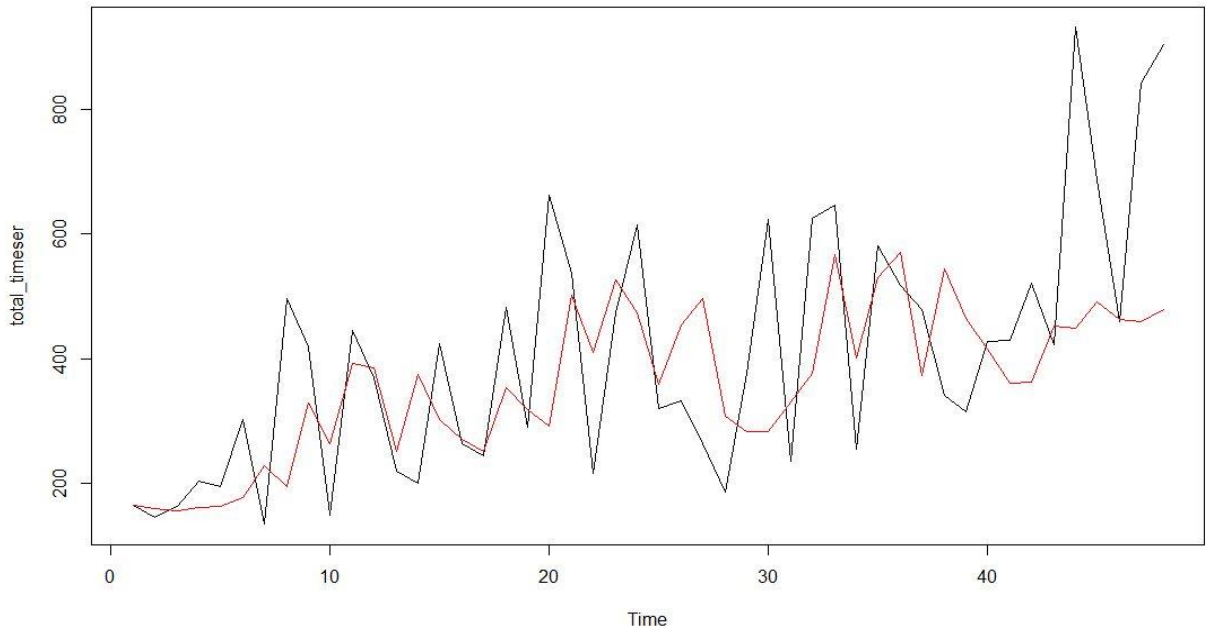
Local Model ARIMA – PACF

APAC Consumer Sales – Auto ARIMA

Auto ARIMA Model



Model with Forecast as Red Line



- ✓ Auto Model: ARIMA(0,1,1)
- ✓ Residuals are stationary based on both ADF and KPSS Tests
- ✓ MAPE (Mean Absolute Percentage Error) = 27.68952

Model Evaluation

MAPE (Mean Absolute Percentage Error) Comparison

Bucket	Classical Decomposition	Auto ARIMA	Recommended Method
EU Consumer Demand	31.45475	30.13319	Auto ARIMA
EU Consumer Sales	23.18592	28.9226	Classical Decomposition
APAC Consumer Demand	21.53303	26.24458	Classical Decomposition
APAC Consumer Sales	23.93068	27.68952	Classical Decomposition

Model Observations:

- Additive Trends show lower MAPE than Multiplicative Trends in Classical Decomposition.
- For EU Consumer Demand, we have used Multiplicative Trend. This is because when we used the Additive Trend, after removing the local and global components, the residuals were not stationary as per ADF test.

Model Deployment

Recommendations to Management:

- We recommend the following methods for forecasting:
  - Auto ARIMA for EU Consumer Demand
  - Classical Decomposition for EU Consumer Sales, APAC Consumer Demand and APAC Consumer Sales
- These recommendations are based on our analysis of transactions in the period from 2011 to 2014. So, they should be valid as long as the market conditions are similar to the transaction patterns seen during the aforementioned period.
- As an industry standard and given that our MAPE also lies between 25% and 30%, we can assume around 70% to 75% accuracy for our forecasts. Planning for safety stock and/or reserve inventory should be done accordingly.
- This forecast may not be so accurate if we see shifts in purchasing patterns, unprecedented trade discounts, promotional activities arising due to local dynamics like festivals, special events, incidents, etc which are not a part of the transactions considered in our analysis.

THANK YOU