Time Series Analysis– Case Study

Group Members:

1. Amitava Maity
2. Arpan Kumar Nandi
3. Ajay Kumar

Objective of this case study

In this case study we are required to forecast the sales and the demand for the next 6 months using time series analysis for online store (Global Mart). The results thus obtained would help Sales/Operational manager to manage revue and inventory accordingly.

The following questions are among the ones they would like answered

- What are top profitable Markets and segments?
- Whether their profits are consistent?
- Can we fit time series on their sales and quantity ?
- If so, how well can we predict?

The analysis needs to be done for answering below questions :

- Identify trends of Sales and forecast for next 6 months
- Identify trends of Quantity and forecast for next 6 months
- Measure the MAPE and decide that actually the model can be used for forecasting

Collect and Manage the Data

After load the datasets, we have found that we have following information for analysis. We have good dataset for 2012,2013 and 2014 and the data consist of -

1. Order Date and Shipment date information for Global Store
2. Demographic information (Market and Segment) for Global Store
3. Sales, profit and Demand information

Row ID	Customer Name	Market	Sales
Order ID	Segment	Region	Quantity
Order Date	City	Product ID	Discount
Ship Date	State	Category	Profit
Ship Mode	Country	Sub-Category	Shipping Cost
Customer ID	Postal Code	Product Name	Order Priority

Problem solving methodology

Step 1: **Data load cleanup**

Load data → Data cleanup → NA correction and Date formatting → Final dataset for analysis.

Step 2: **Selecting features to be analyzed and Prepare data**

Finalize columns to be analyzed → Order and Group the date based on Order Date, Market and Segment → Derive CV for Profit , sales and Quantity → Identify two segments for final Model Building

Step 3: **Model Building**

Build Time Series Model → Smoothing the time series → Identify Global Trend and subtract it from time series to generate local → Check whether local trend is stationary or non stationary → Build Autoarima model for same time series

Step 4: **Model Validation**

Predict the sales and Quantity using the Logical Decomposition and Autoarima model → Compare MAPE for both models and → Decide with model is the best for forecasting based on MAPE value → Plot the results

EDA

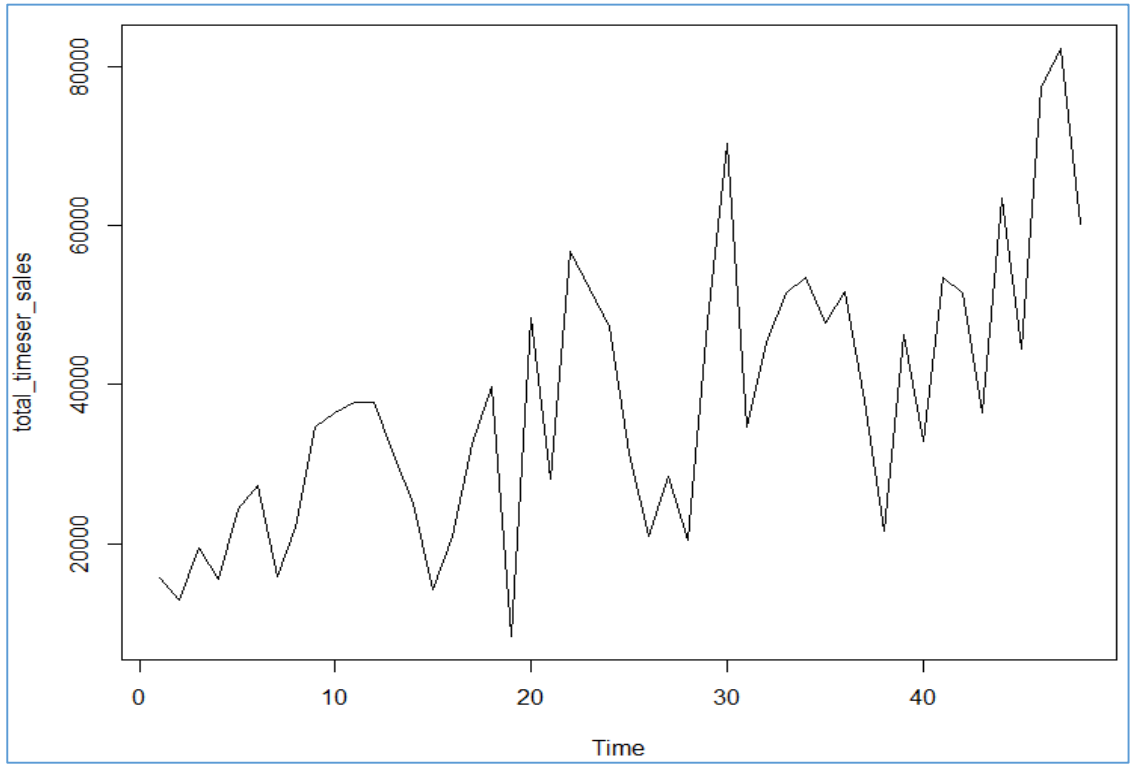
Summary of EDA

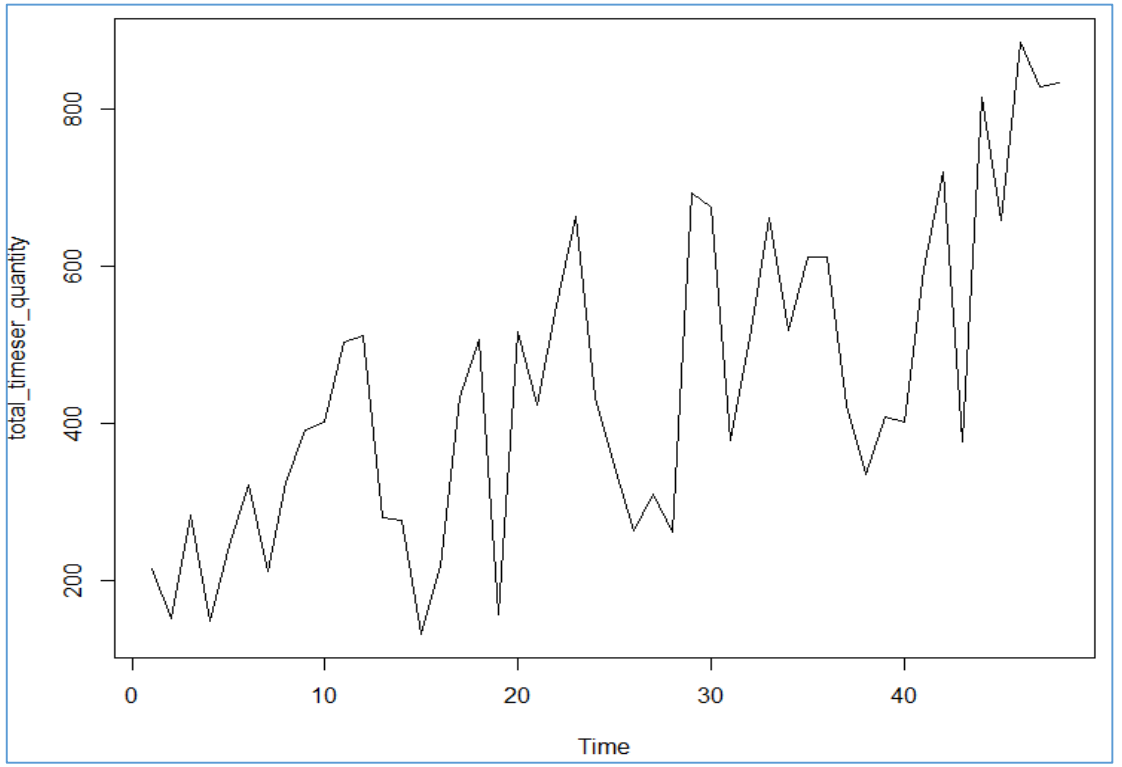
- Change the order date into monthly order date
- Group the data based on order date
- Group the data based on segments and Market
- Identify the top performing segments and Market –
    1. APAC – Consumer (Segment 1)
    2. EU – Consumer (Segment 2)

Model Building (Segment 1)

Generating time series for each segment using Classical decomposition (Segment 1 – APAC – Consumer)

Initial Timeseries plot for sales and quantity

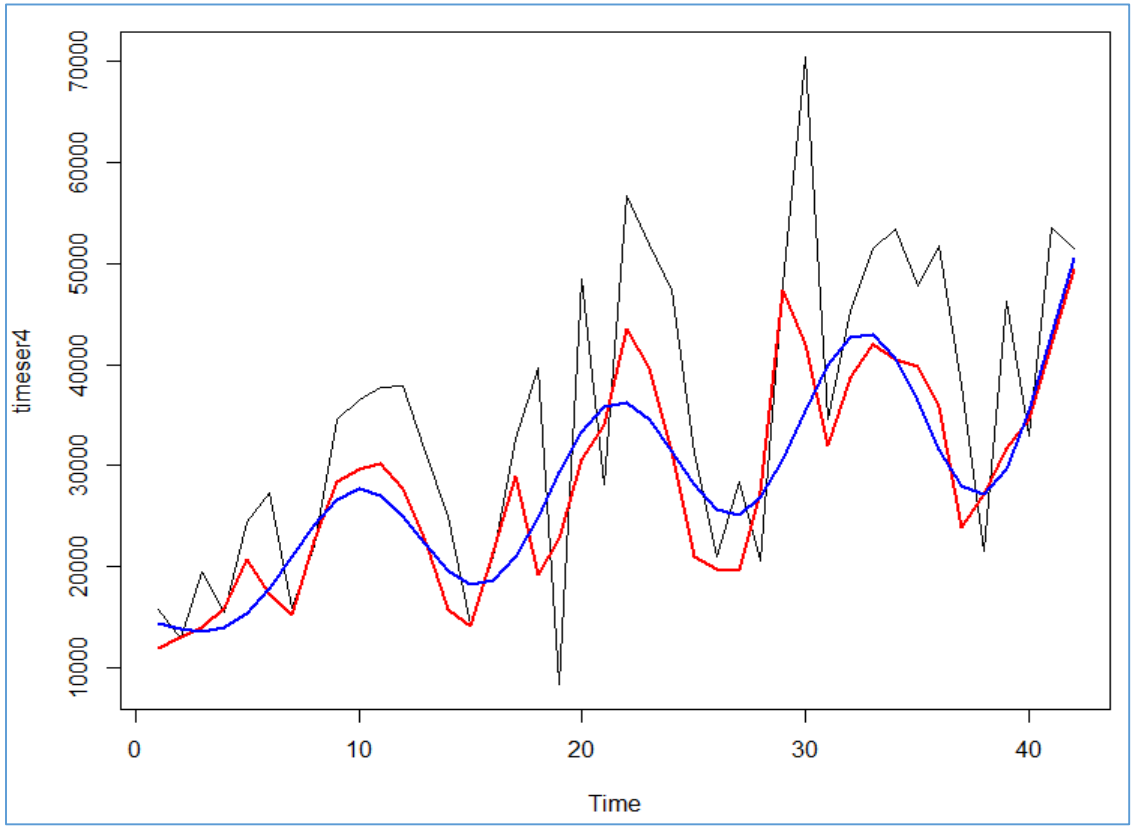


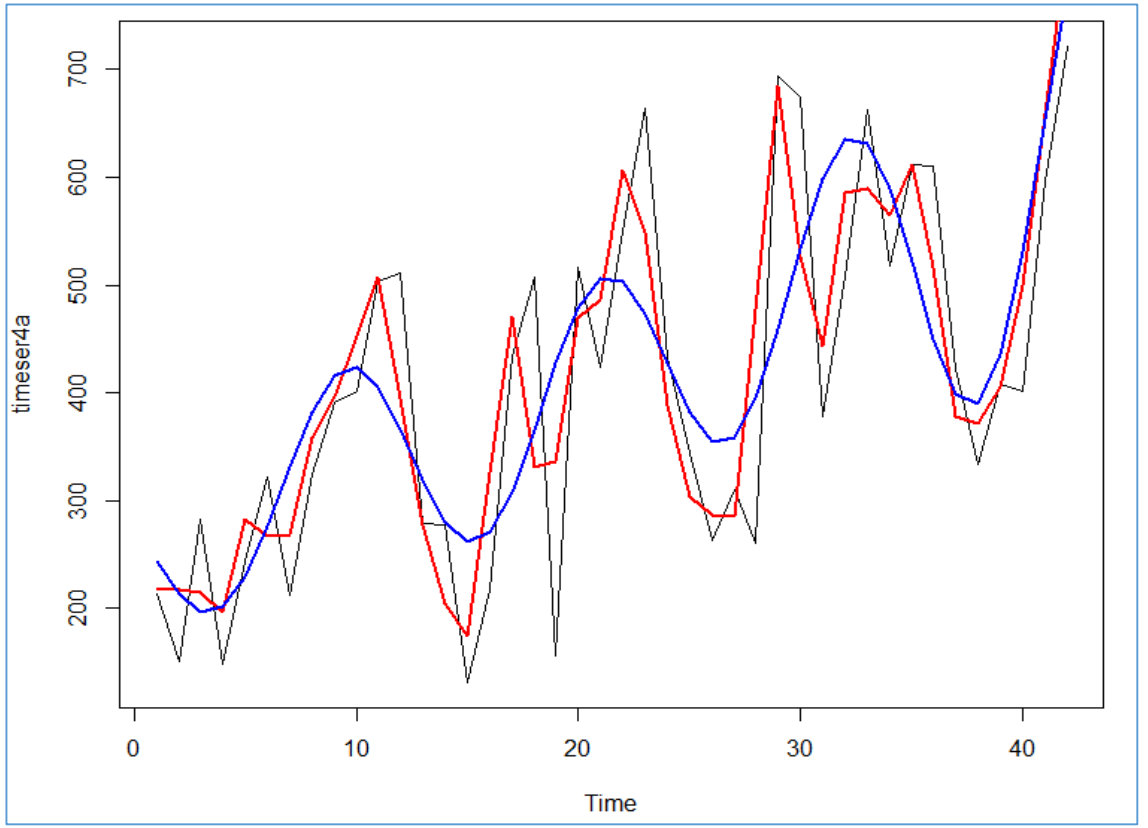


Model Building (Segment 1)

Generating time series for sales and Quantity

Generate Model for sales and Quantity





Identify stationary time series in local trend

ADF Test for Sales Data for 1st selected segment

P value is = 0.01 < 0.05 that means local series is a stationary time series

KPSS Test for Sales Data for 1st selected Segment

P value is = 0.1 > 0.05 that means local series is a stationary time series

ADF Test for Quantity Data for 1st selected segment

P value is = 0.01 < 0.05 that means local series is a stationary time series

KPSS Test for Quantity Data for 1st selected Segment

P value is = 0.1 > 0.05 that means local series is a stationary time series

Validate Model Accuracy for Sales (Segment 1)

MAPE for Sales Model (Classical decomposition)

43

Designed same model using Auto Arima and measure MAPE value

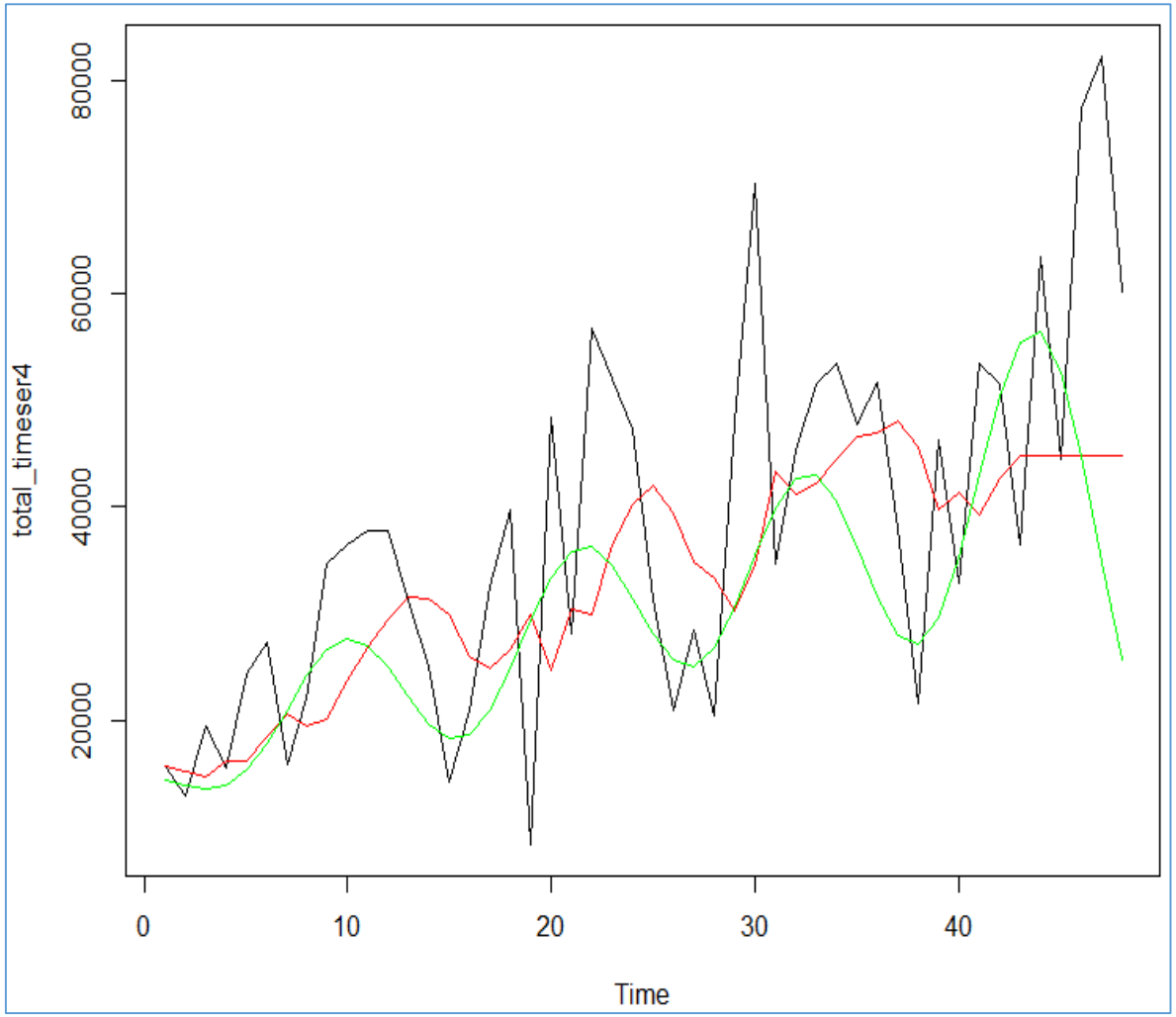
31

Clearly its visible that Classical decomposition is better model compare to Auto Arima model.

From plot also its visible that classical decomposition generates better prediction.

Green- Classical Decomposition

Red – Auto Arima



Validate Model Accuracy for Quantity(Segment 1)

MAPE for Sales Model (Classical decomposition)

43

Designed same model using Auto Arima and measure MAPE value

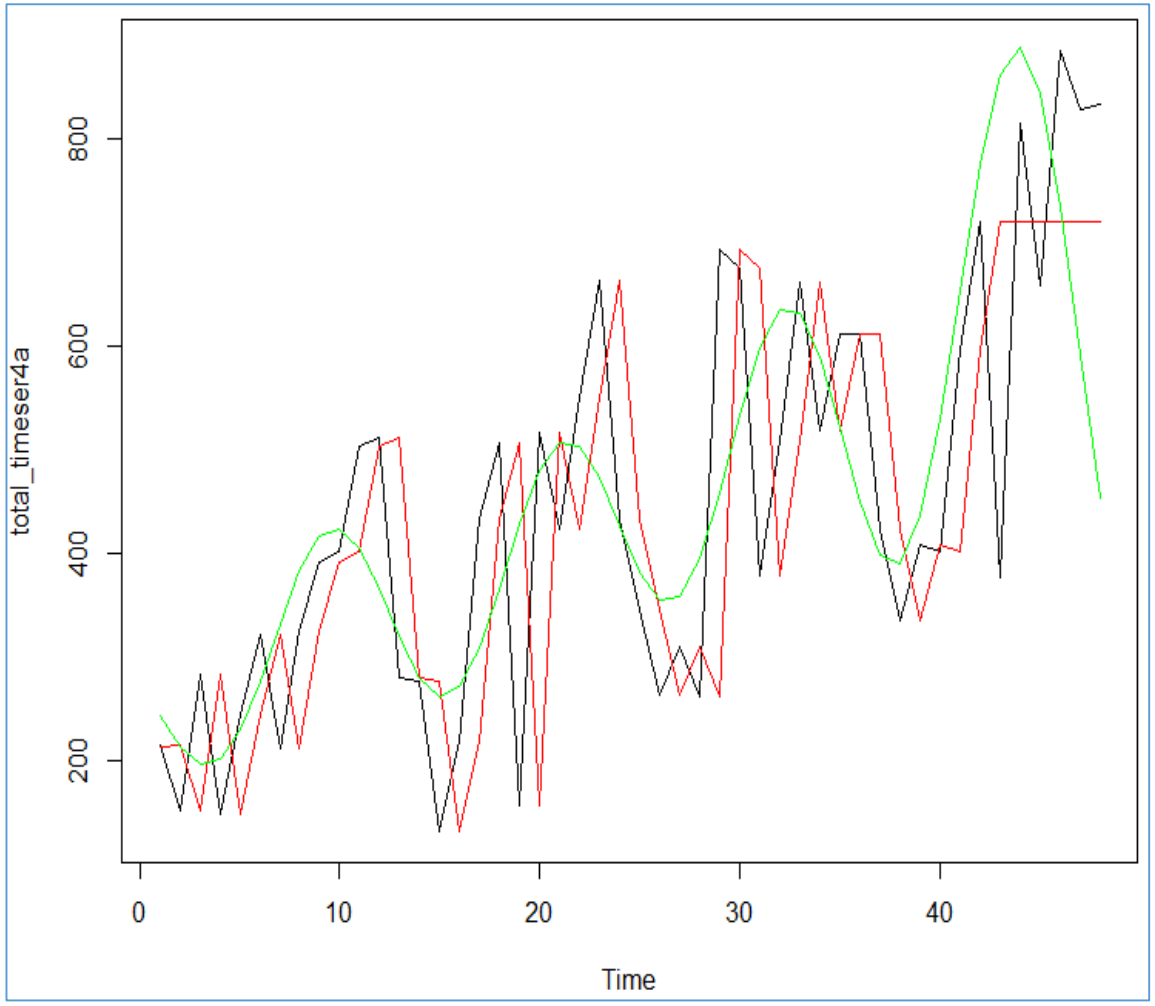
26.24

Clearly its visible that Classical decomposition is better model compare to Auto Arima model.

From plot also its visible that classical decomposition generates better prediction.

Green- Classical Decomposition

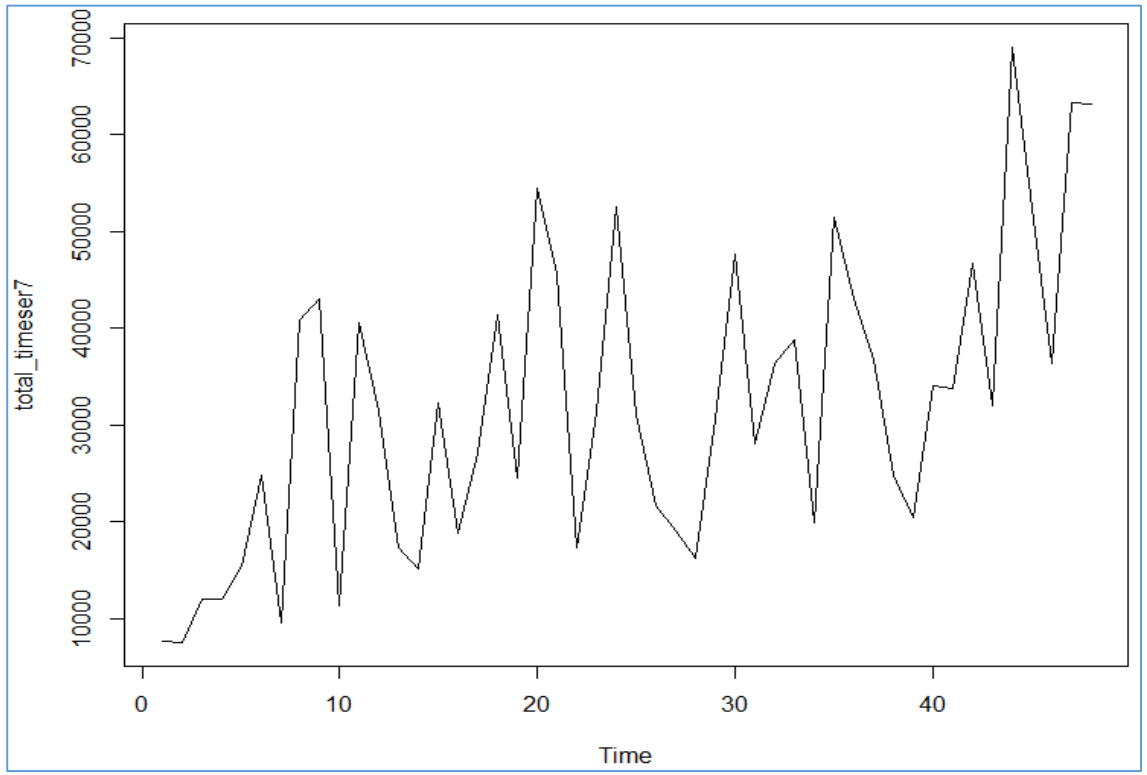
Red – Auto Arima

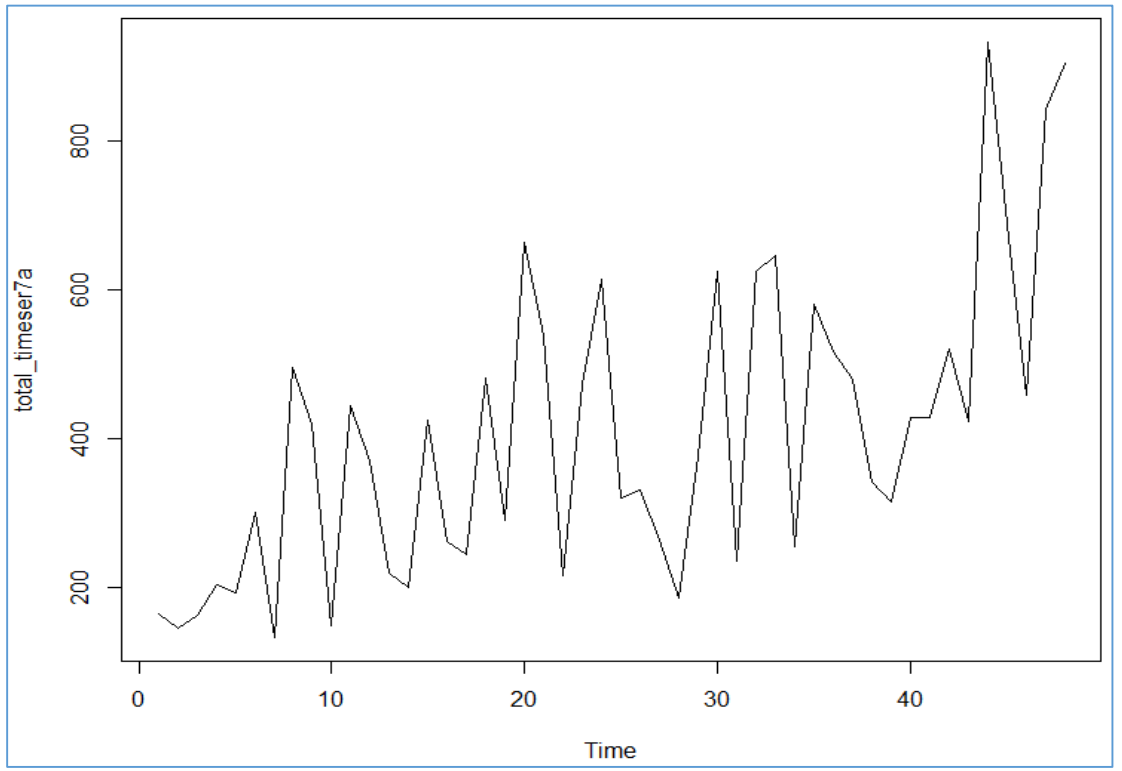


Model Building (Segment 2)

Generating time series for each segment using Classical decomposition (Segment 2 – EU – Consumer)

Initial Timeseries plot for sales and quantity

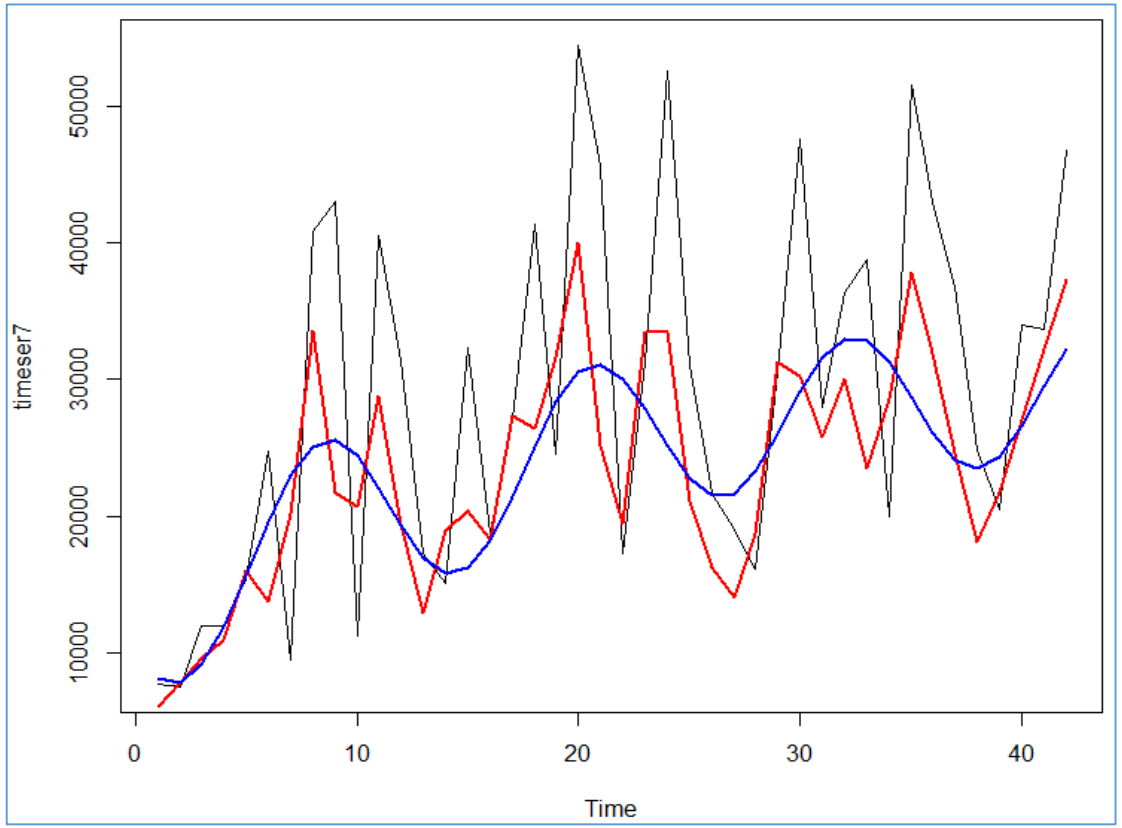


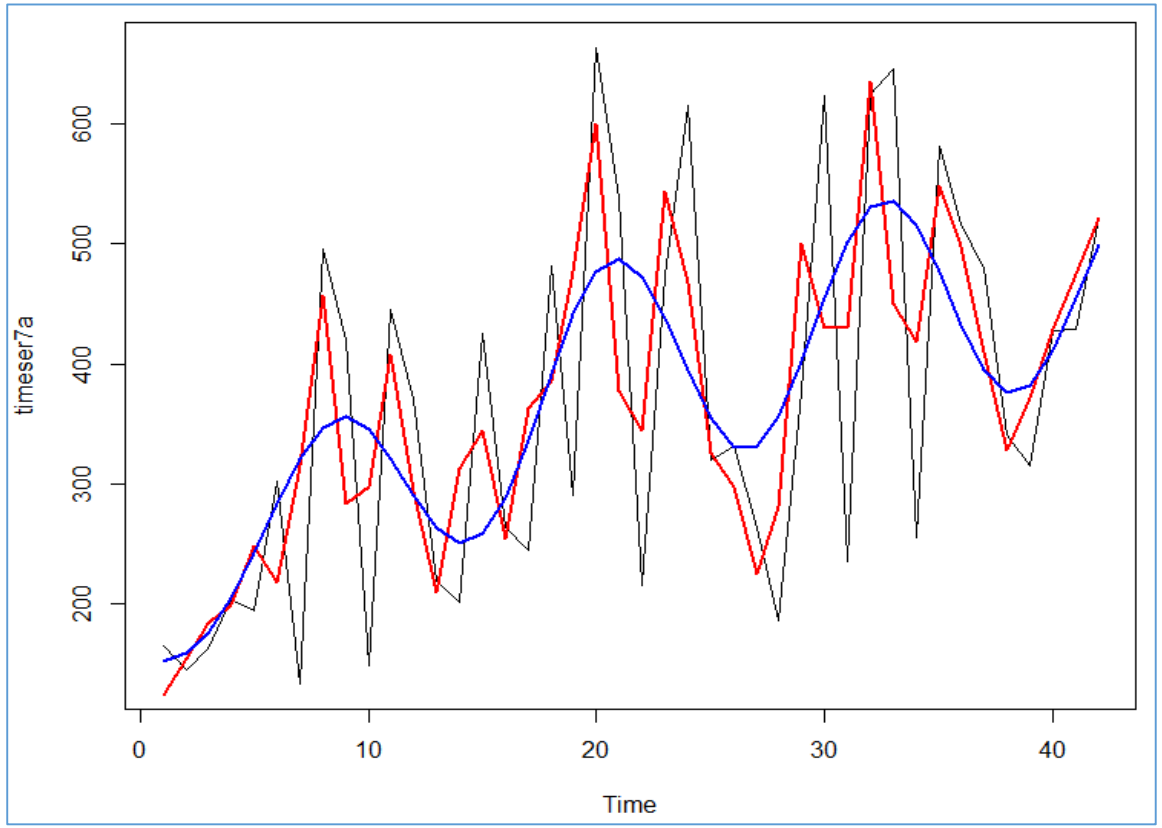


Model Building (Segment 2)

Generating time series for sales and Quantity

Generate Models for sales and Quantity





Identify stationary time series in local trend

ADF Test for Sales Data for 1st selected segment

P value is = 0.02463 < 0.05 that means local series is a stationary time series

KPSS Test for Sales Data for 1st selected Segment

P value is = 0.1 > 0.05 that means local series is a stationary time series

ADF Test for Quantity Data for 1st selected segment

P value is = 0.3513 > 0.05 that means local series is not a stationary time series

KPSS Test for Quantity Data for 1st selected Segment

P value is = 0.1 > 0.05 that means local series is a stationary time series

As we don't find any auto regressiveness property in ACF and PACF test, we can say the it's a stationary time series. And from auto Arima ADF and KPSS test it has been found as a stationary time series.

Validate Model Accuracy for Sales (Segment 2)

MAPE for Sales Model (Classical decomposition)

40

Designed same model using Auto Arima and measure MAPE value

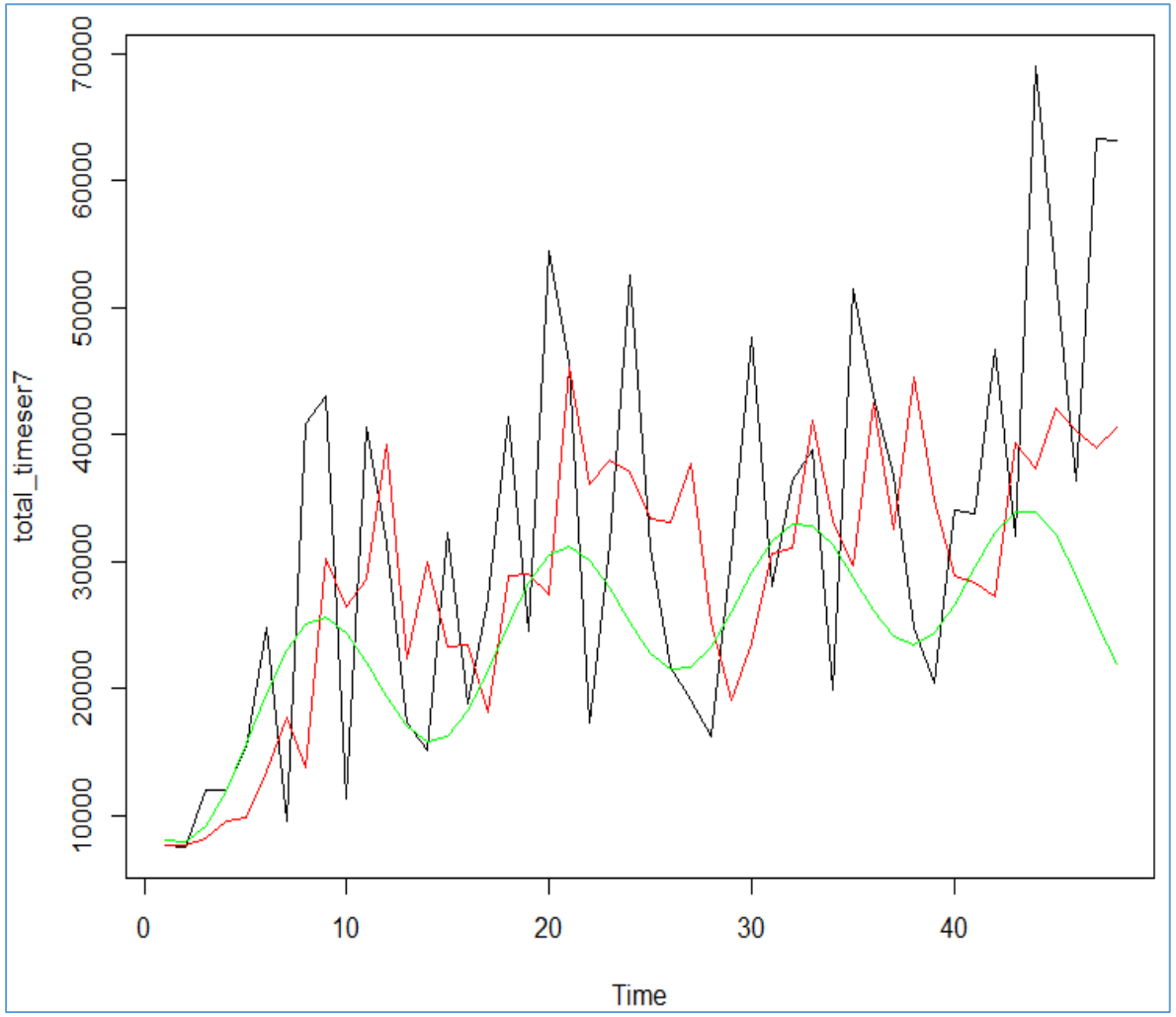
31

Clearly its visible that Classical decomposition is better model compare to Auto Arima model.

From plot also its visible that classical decomposition generates better prediction.

Green- Classical Decomposition

Red – Auto Arima



Validate Model Accuracy for Quantity(Segment 2)

MAPE for Sales Model (Classical decomposition)

33

Designed same model using Auto Arima and measure MAPE value

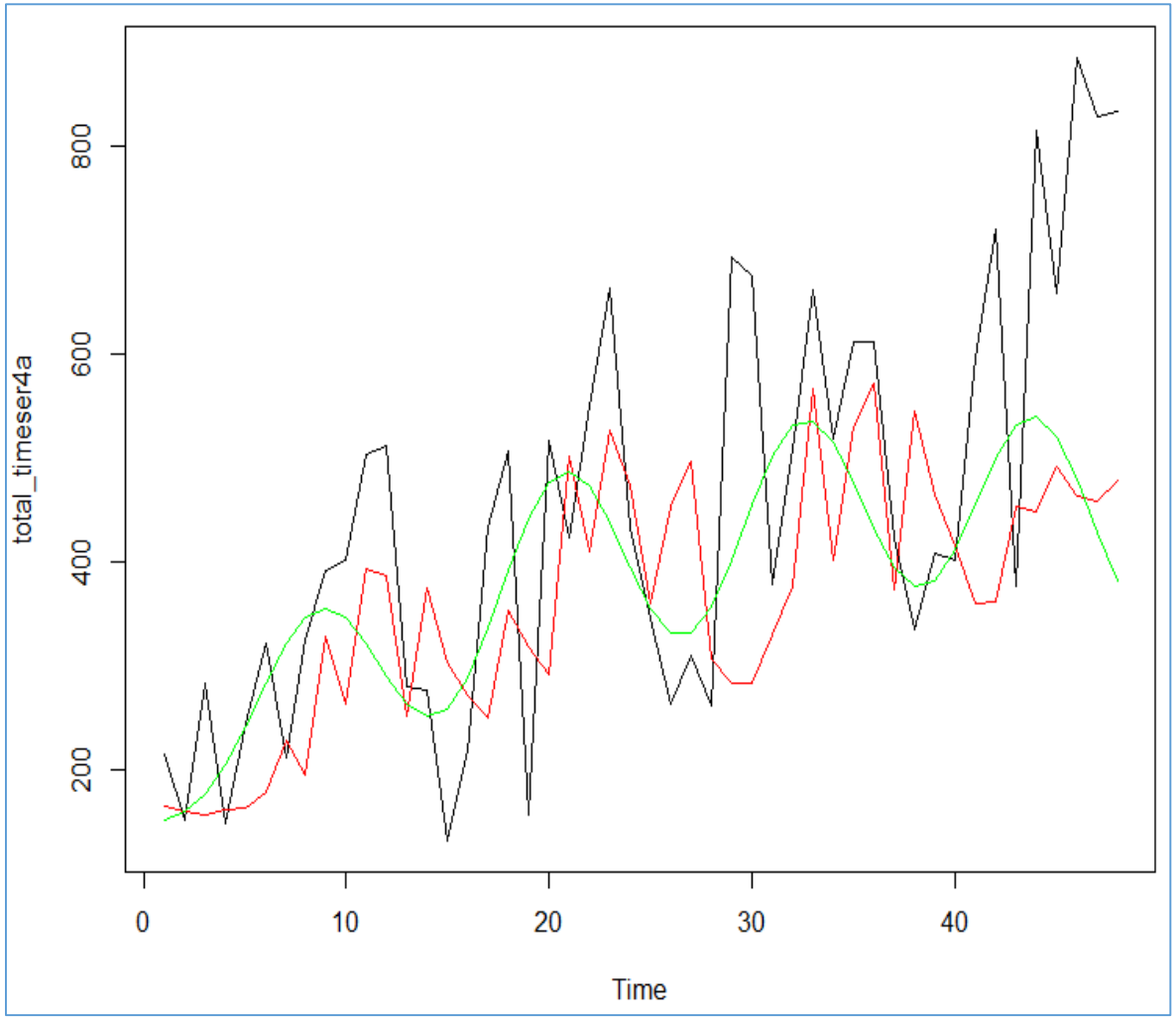
30.13

Clearly its visible that Classical decomposition is better model compare to Auto Arima model.

From plot also its visible that classical decomposition generates better prediction.

Green- Classical Decomposition

Red – Auto Arima



Summary

Throughout the analysis, we have learned several important things

- Classical decomposition gives the us bet fit model for above mentioned time series model.
- There is not auto regressiveness property in local trend
- MAPE value says that model is fairly a good model to predict sales and demand for next 6 months.