GLOBAL MART – RETAIL GIANT SALES FORECASTING

COURSE 4 | PREDICTIVE ANALYTICS

Submitted by

Madhavi Ayyagari

ANALYSIS OBJECTIVE

1. To identify the most profitable market segments based on attributes Sales, Quantity & Profit
2. To forecast the sales/demand for the next 6 months using the time series (Classical decomposition and auto ARIMA)

DATA INPUTS

Global Store Sales data was provided which had Order details, Shipping details, Customer details, Market segment and Product category details.

As per the business understanding we used these details from the data given

1. Order date
2. Product Category

   Consumer, Corporate & Home Office
3. Market segment

   Africa, APAC, Canada, EMEA, EU, LATAM, US
4. Sales Amount
5. Quantity
6. Profit

TIME SERIES ANALYSIS – METHODOLOGY

1. DATA PREPERATION

1. Handling the NA values for Postal code
2. Changing char type for order date
3. Creating subsets based on the 7 markets and 3 segments into 21 buckets.
4. Deriving the summary sales data and Monthly summary data with Total Profit, Profit percentage, Quantity of Sales.

2. ANALYSING THE DATA BASED ON PLOTS

1. We have calculated the standard deviation, mean and found out the CV and thus calculated.
2. Using the resulted values we plotter bar graphs and found the Top 2 profitable market segments.

3. DECOMPOSING

1. Creating the timer series and further decompose to analyse
   Trend,

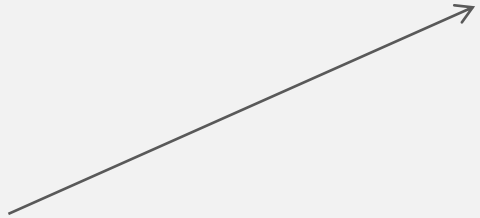
4. SMOOTHENING

1. Testing the model using the test dataset for determining accuracy, specificity and sensitivity – Confusion matrix

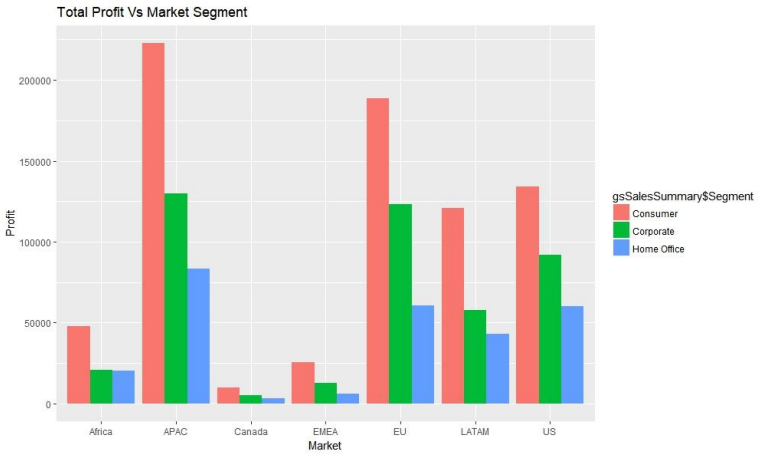
5. ARIMA AND MODEL TESTING

1. Checking the results and remodelling based on accuracy measured by MAPE method and matching the forecast values.
2. If the results are acceptable the final model is declared

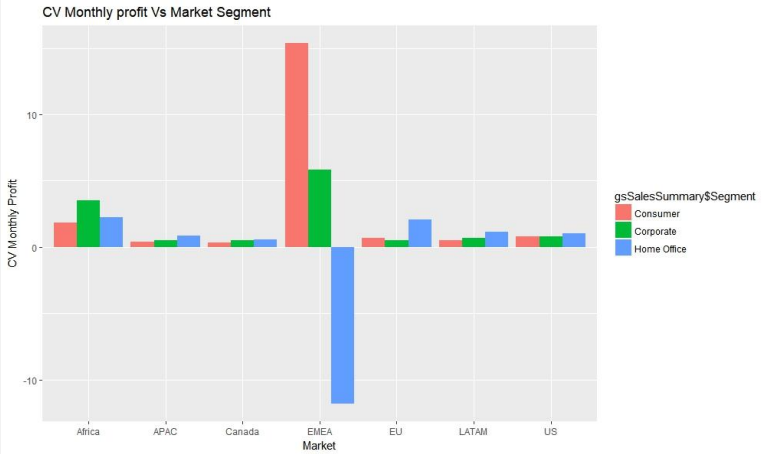
6. FORECASTING AND FINAL MODEL



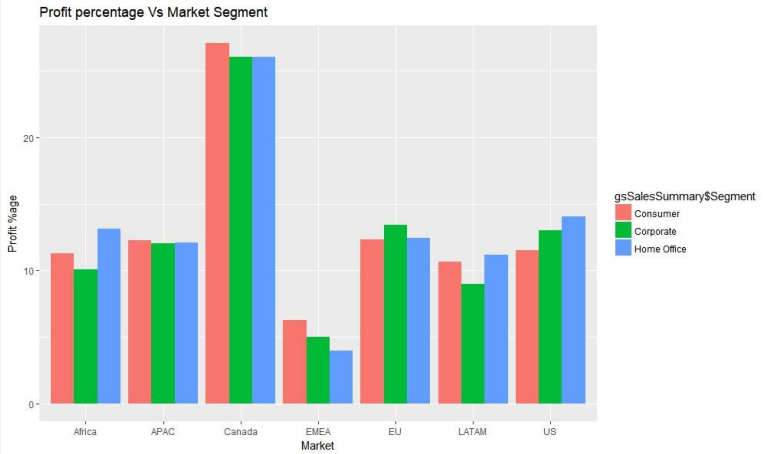
ANALYSING THE DATA BASED ON PLOTS



Top 10 Market segments based on the total maximum profits



Among the above the top 10 Market segments further selected based on maximum profit percentage



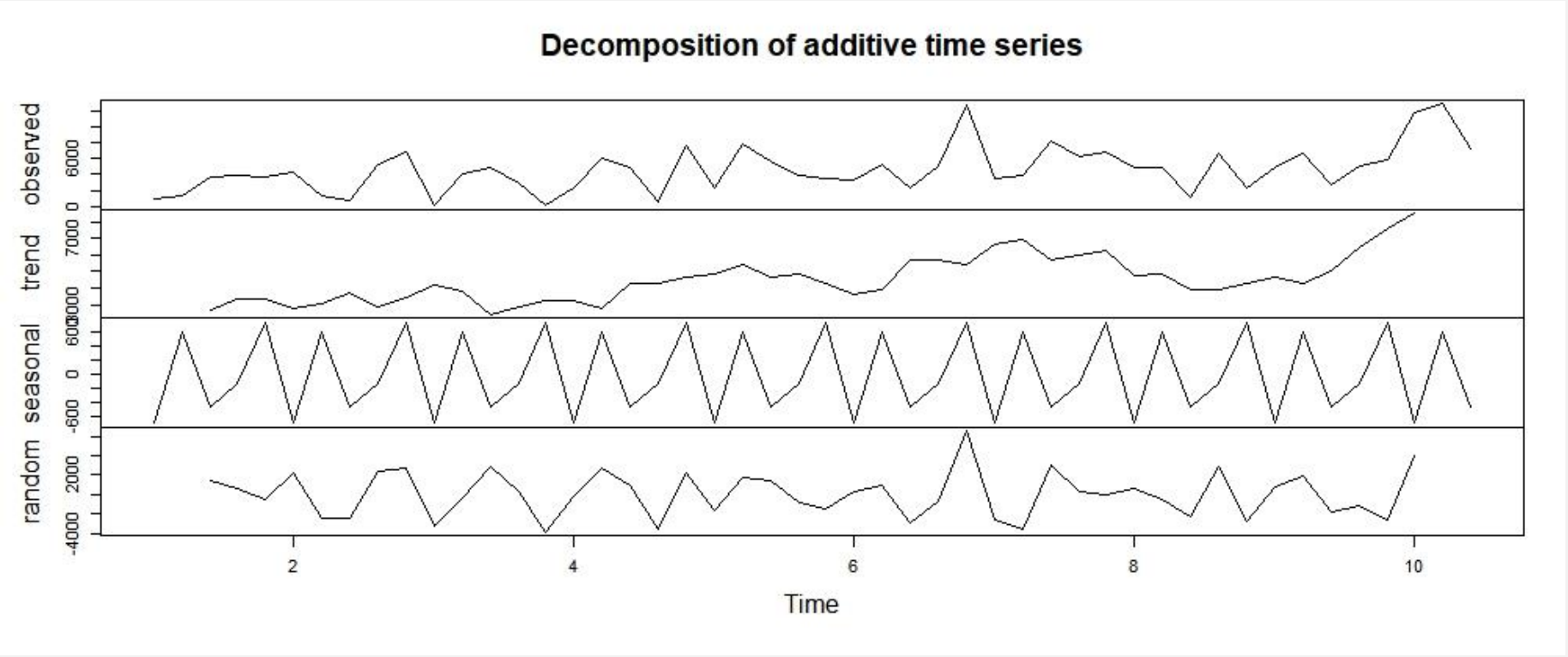
Among the above the top 10 Market segments further selected based on the low CV

Inference

Therefore the 2 most profitable segments which have to be considered for Model building are:

1. APAC Consumer
2. EU Consumer

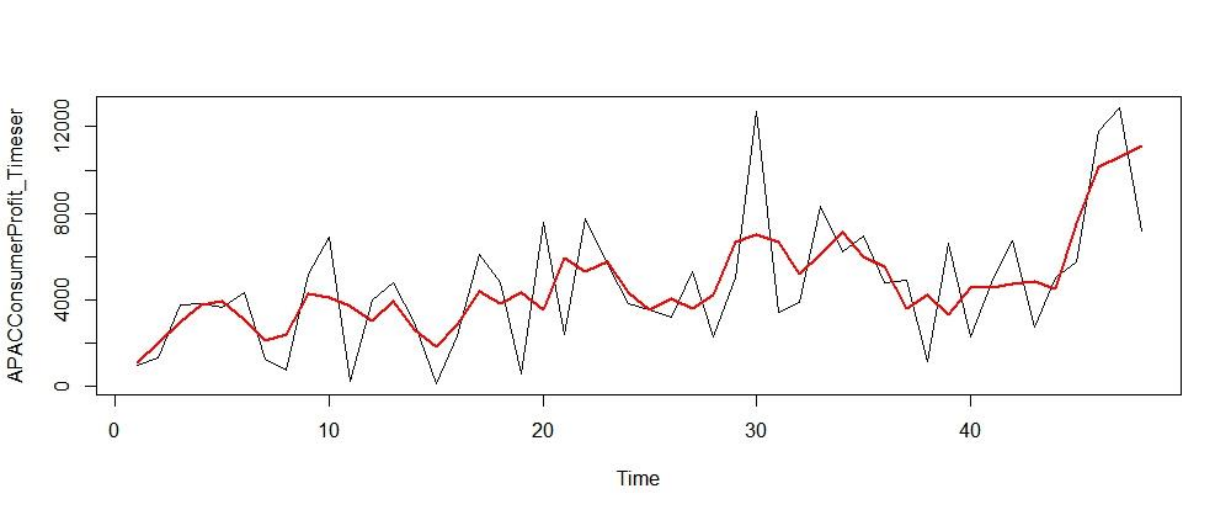
DECOMPOSE THE DATA FOR FURTHER ANALYSIS

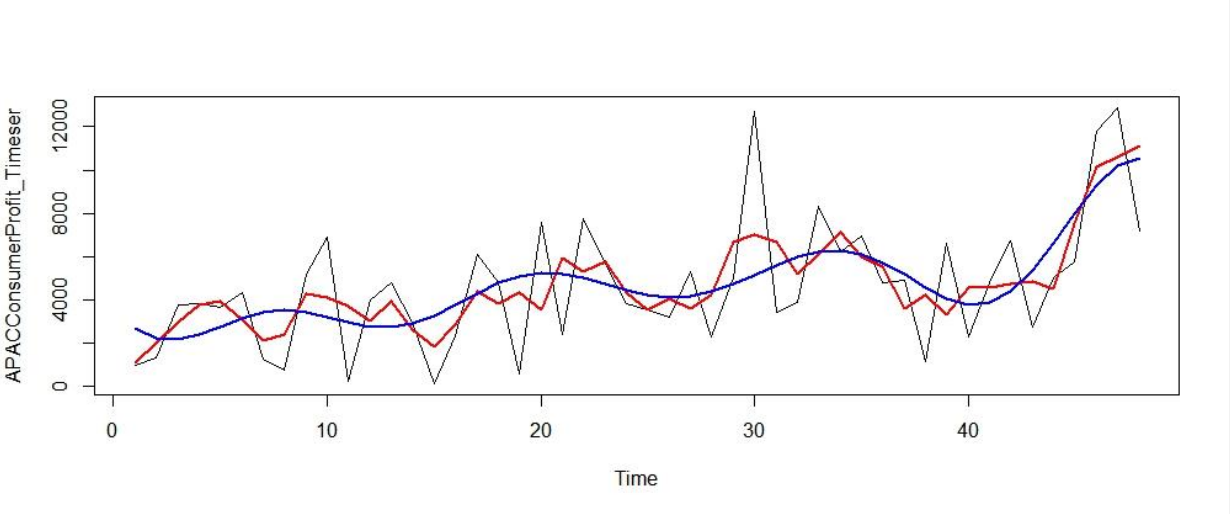


Inference

1. There is an upward trend with a high wavelength sine curve
2. Seasonality seems like a sine curve

SMOOTHENING CURVE USING BOTH NORMAL AND EXPONENTIAL METHODS





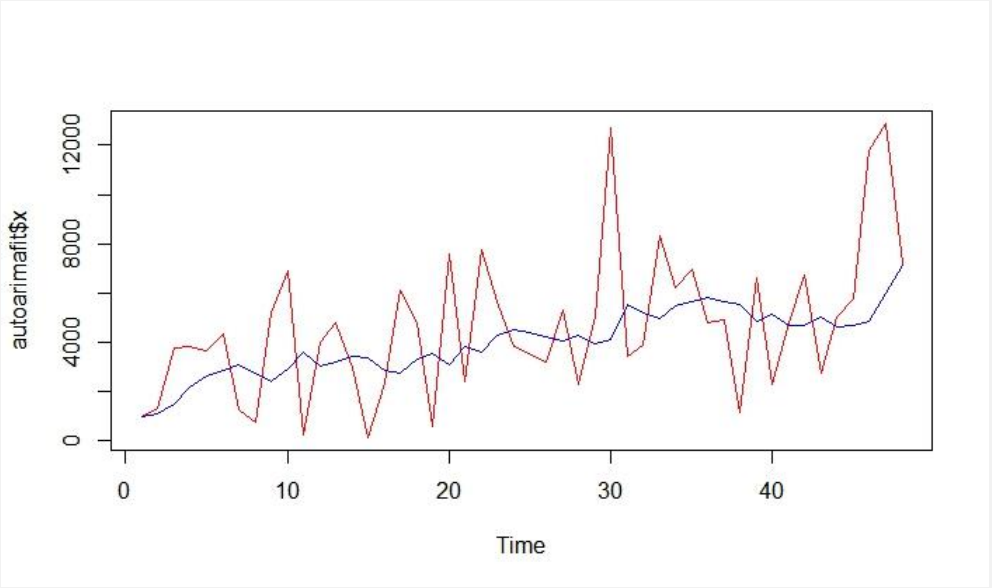
Inference

1. Exponential smoothening reduces the sloes further thus we are using the normal smoothening timeseries for analysis to obtain accuracy

ARIMA AND MODEL TESTING USING MAPE

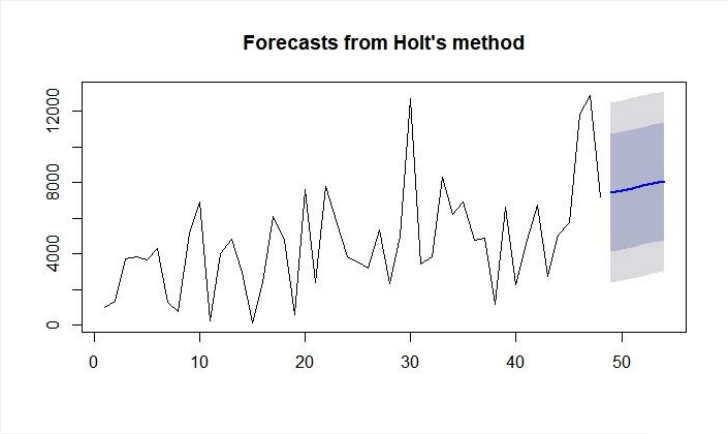
AURO ARIMA:

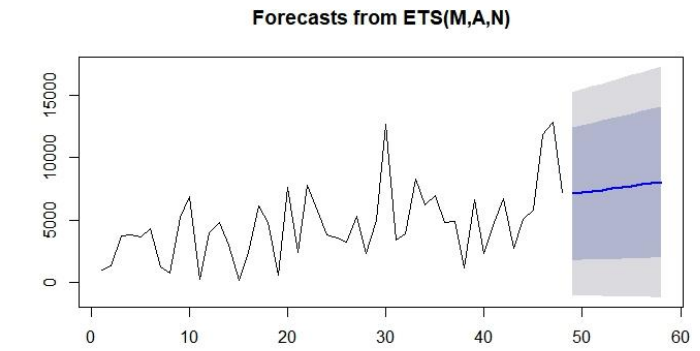
The autoarima series was similar to the series obtained after smoothening



MODEL EVALUATION AND MAPE

Using MAPE the model is tested and has the forecasts matched.





FINAL MODEL FOR FORECASTING

FINAL MODEL WITH AN ACCURACY OF 57% WAS OBTAINED AND FINALISED

Based on the model obtained the process was repeated for calculating the forecasted sales for APAC consumer and EU consumer.