GLOBAL SUPER STORE CASE STUDY

SUBMISSION

Group Members

1. Abhishek Ranjan
2. Maxim Rohit
3. Viveka Manikonda
4. Yogesh Kulkarni

Business Understanding and Objectives

Background

"Global Mart" is an online store super giant with operations across world and deals with major product categories including consumers, corporate & home office

Data availability & key detail

Data available is from Jan 2011 to Dec 2014 with 24 attributes across 51,290 transactions. Among the 24 attributes, data is segments using 7 different market segments and 3 major categories

Key Objectives

To better manage revenue and inventory we would like to forecast sales and the demand for the next 6 months, for two most profitable segments (out of 21 segments ie., 7 different markets and 3 major categories)

Analysis Approach

Data Sourcing & Preparation

Data set contain 51,290 transactions with 24 attributes each

A quick check on data indicate no duplicate entries or missing detail across attributes except for postal code which is missing for 41,296 transactions. Postal code data is not used for building model hence it is ignored.

Date columns are observed to be in character format and converted to appropriate date formats for analysis purpose

New columns were created on 'Order Date' column to observe seasonality in monthly sales, order quantity and profits

Whole data set is grouped by 'Market' category and 'Segment' categories into 21 subset categories

Coefficient of variation is used to identify two most consistently profitable segments

Exploratory Data Analysis

Exploratory Data Analysis indicate sales has been consistently growing the time period albeit with some seasonality – (i) sales are higher in second half on average particularly towards end of year reflecting holiday demand (ii) June is most active period in first half and December in second half (iii) lately August and November month sales are increasingly may due to promotional offers

Consumer segment represent largest share in Global Mark Sales and has been consistently growing over the time period

Sales by APAC and EU markets have grown significantly over the time period, followed by LATAM market

All categories experienced growth with sharp rise in demand for technology and office supplies

Coefficient of Variation by Market segments (21)indicate "Consumer APAC" and "Consumer EU" divisions have consistent profitability

Model Building & Evaluation

For both Consumer APAC and Consumer EU segments, data is grouped by monthly by year and identified to be spread over 48 month period

Of which, train set is recognised as first 42 month data and remaining 6 months are reserved as test set

Auto arima model is run on train set and removed from train set time series(ts) to arrive at residual ts

Both ADF & KPSS tests are run to confirm residual ts is white noise and then we used Auto arima to forecast values and MAPE is calculated

Time series is smoothed by choosing a particular width through incremental value choice of width

Both Additive and Multiplicative regression lines are run and a particular model is chosen for its low complexity, while ensuring error percentage is within acceptable limits

Once a particular type of regression is chosen, it is then used to predict train set historical values, which is termed as "global predict"

Then "local predict" is calculated by removing "global predict" from original train time series

We use ACF and PACF tests to determine whether "local predict" time series is stationary or not

Further, we run a auto arima model on "local predict" time series and remove the auto arima fit model from "local predict" to get residual time series

ADF and KPSS tests are repeated on residual time series to check on stationary

We use the regression based global predict and auto arima local predict model and forecast for next 6 month period

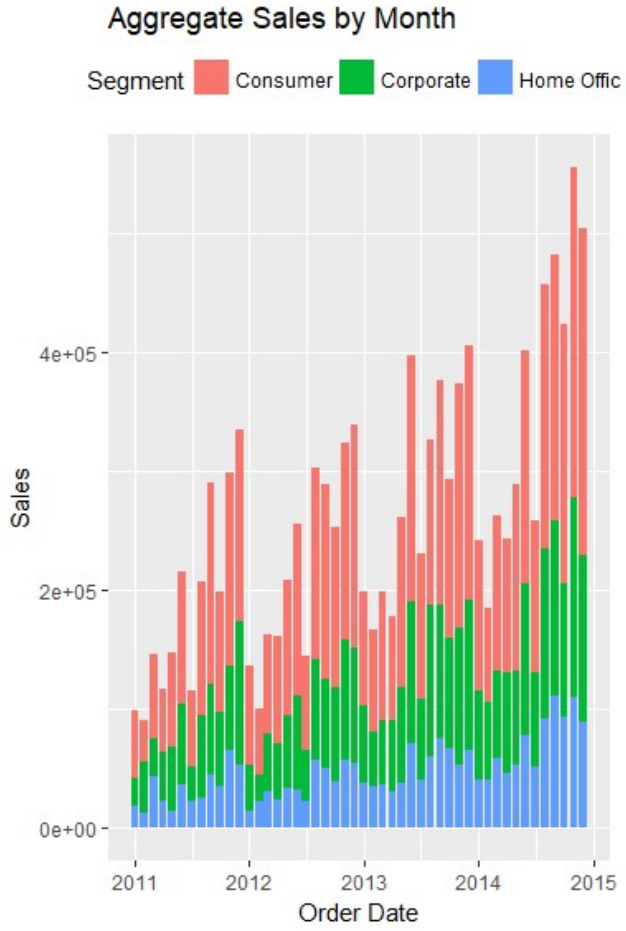
We use original test values and above calculated values to determine MAPE for comparisons among models

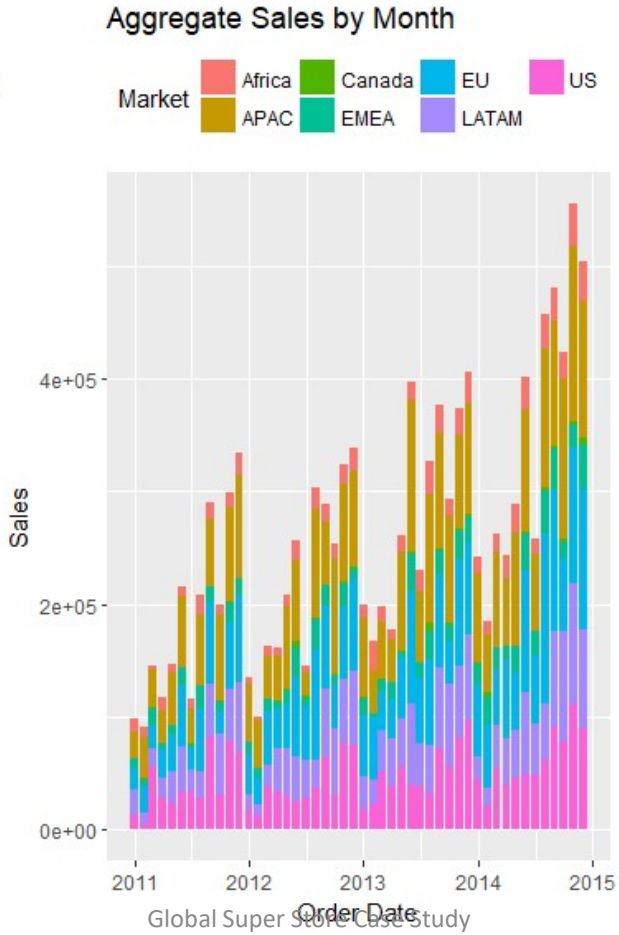
Monthly Trends

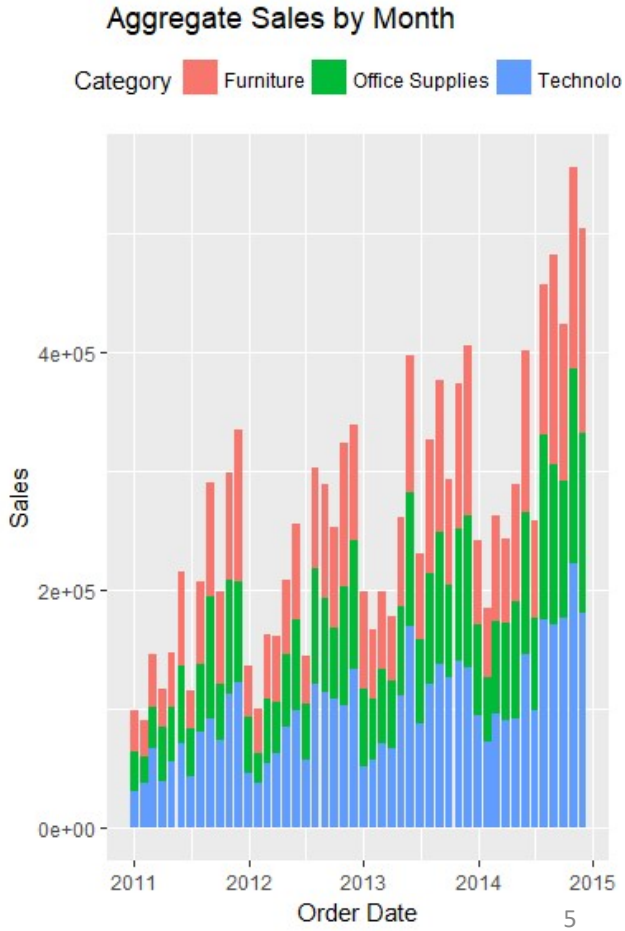
- Data contains seasonality (increase towards end of year – Nov, Dec) and overall trend is increasing
- Average Sales, Quantity and Profit low in 1st Half vs 2nd Half – may due to promotional offers and general increase in demand due to holiday season
- Although Sales Volume and Value increased in 2H of 2014, profitability reduced which may indicate price competitive measures
- In the first half, June month recorded high sales volume and value. Also, at the end of each quarter, there is a general uptick in sales volume and value – if this is corrected through sales policy change it would be easy to streamline inventory management through the quarters and avoid shortages

Monthly Trends

- Increase in demand during end of year is largely driven by Consumer Segment and Home Office segment, while demand from Corporate segment remained on average constant through second half of the year
- EU, US & LATAM clearly show an increase in demand through later months in a year – most likely due to holiday seasons
- While Technology remain most popular product category during second half of year, there is an increase in 'Furniture' products especially in 2013 and 2014

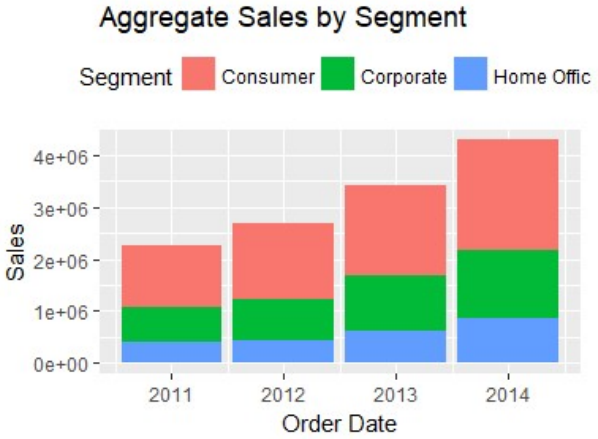


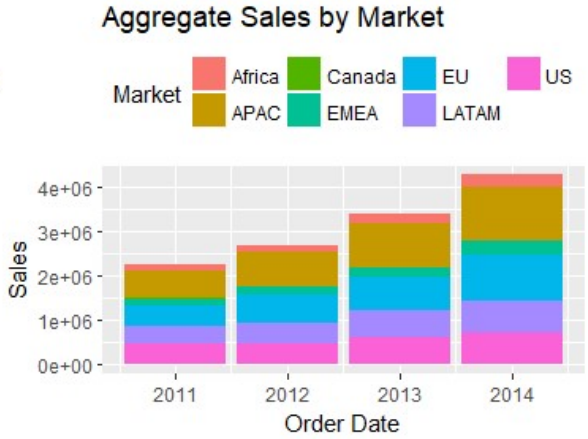


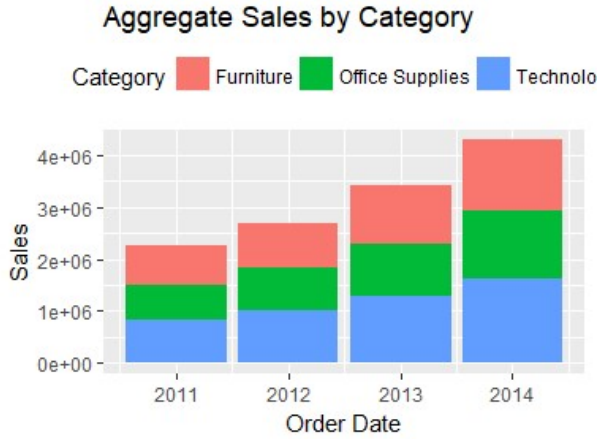


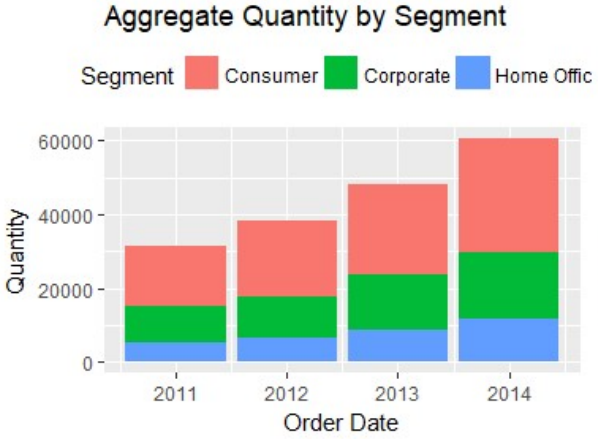
Yearly Trends

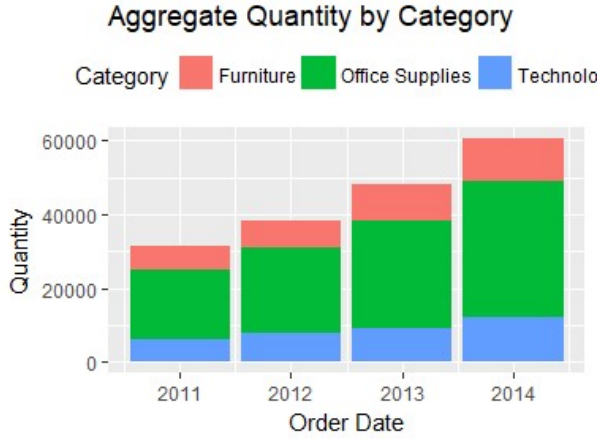
- Growth rates in consumer segment outperform other segments (Corporate and Home Office)
- APAC which is nearly same as that of EU market in 2011 outperformed in terms of volume and value, as the economies improve rising income levels at consumer levels
- Sales contribution from Office Supplies grew significantly over the time period along with volume growth. Technology product category remains low volume high value product business









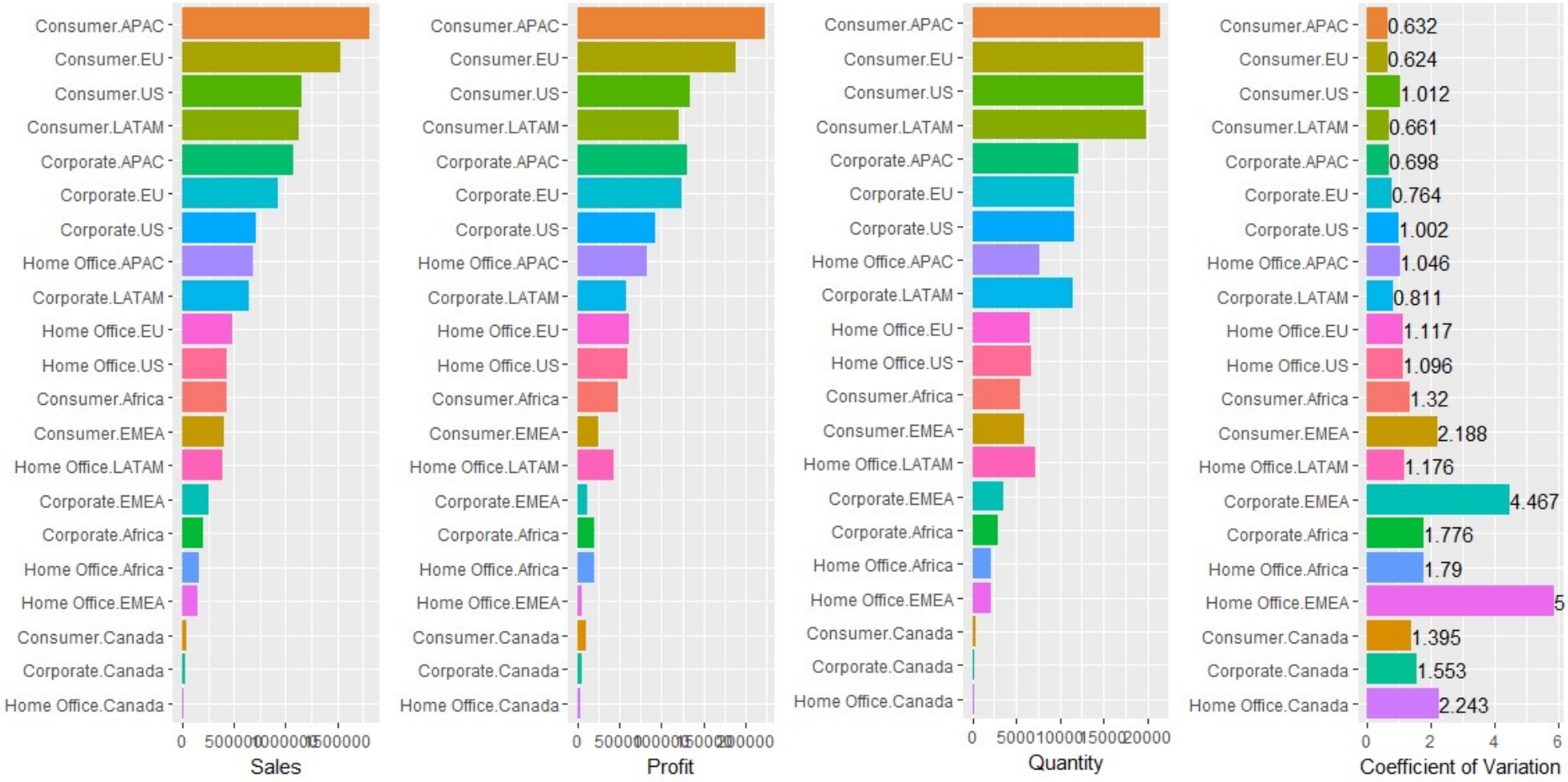


Categorical Analysis

- Consumer is the largest segment for Global Mart contributing over 50% of revenue. Profitability of segments indicate all three Consumer(11.5%), Corporate (11.5%) and Home Office (12.0%) have nearly same profit margins
- APAC and EU remain the large two sectors with high Sales and Profit contribution
- Quantity supplied to EU, LATAM and US markets remain near same but sales and profit remain small for LATAM and US, which indicate avg. selling prices of goods in these markets are low
- In terms of profitability, Technology remains top at 14.0%, closely followed by Office Supplies at (13.7%) and Furniture products have low profit margin of (6.9%)

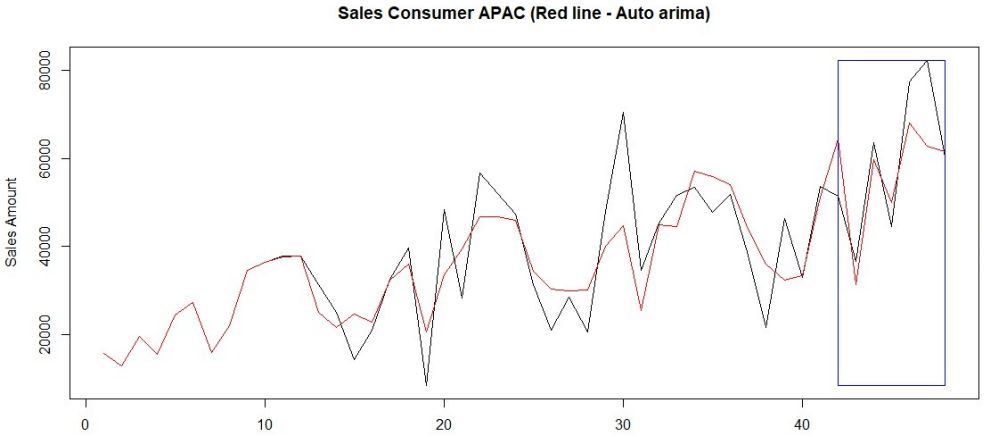
21 Segments overview

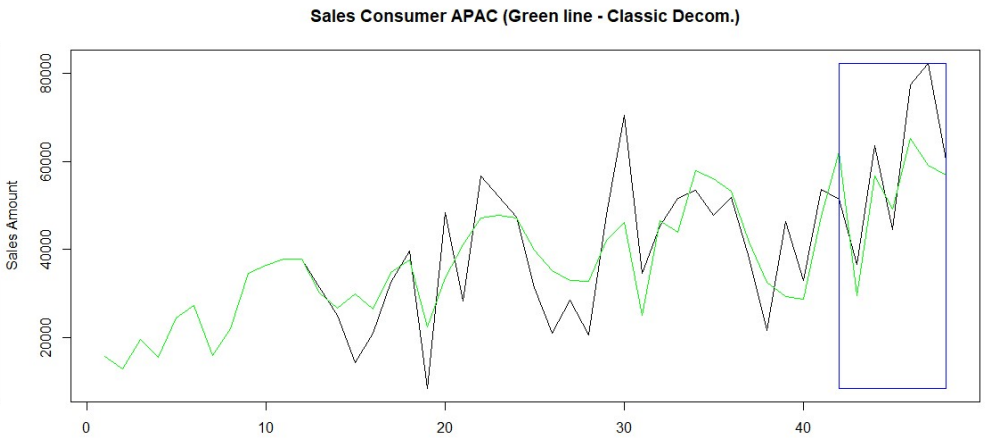
- Consumer APAC and Consumer EU lead in terms of overall Sales, Quantity and Profit
- Coefficient of variation in Profit remains low for Consumer EU and Consumer APAC
- Thus, these two can be considered as the most profitable segments for forecast analysis

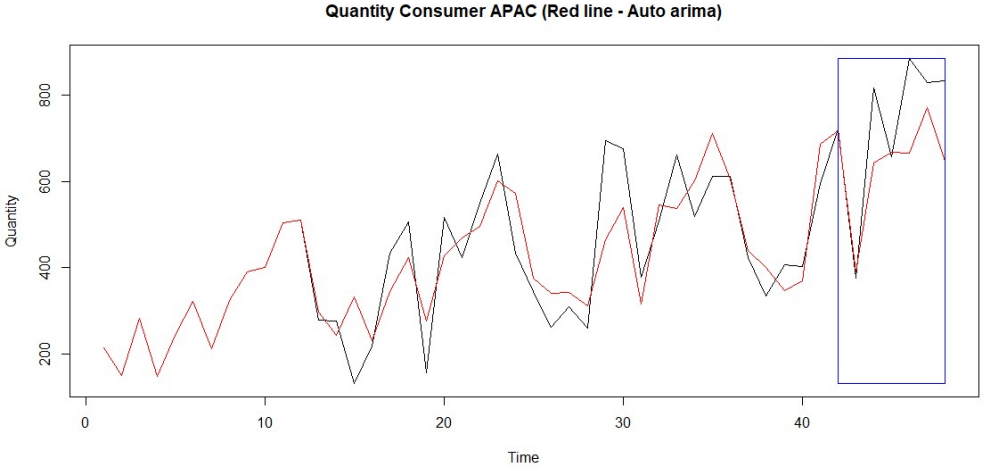


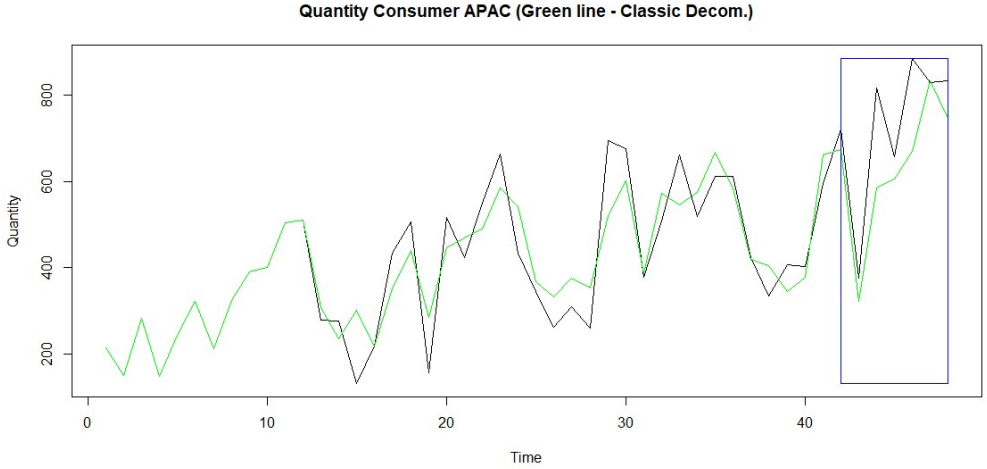
Consumer APAC Division – Sales & Quantity forecasting

- ADF Test on auto arima model indicates some residual component, hence we isolate it using classic decomposition models
- MAPE values of Classic decomposition remain within acceptable levels(14-15). Forecast data is highlighted in Blue box



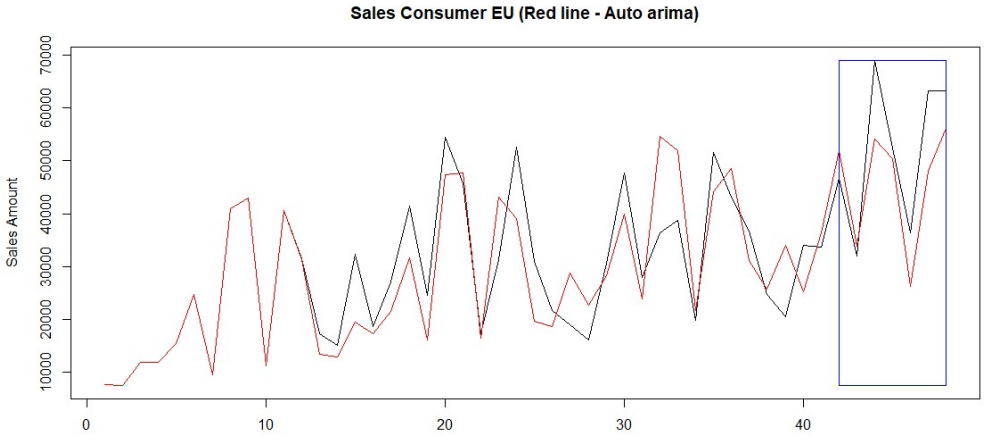


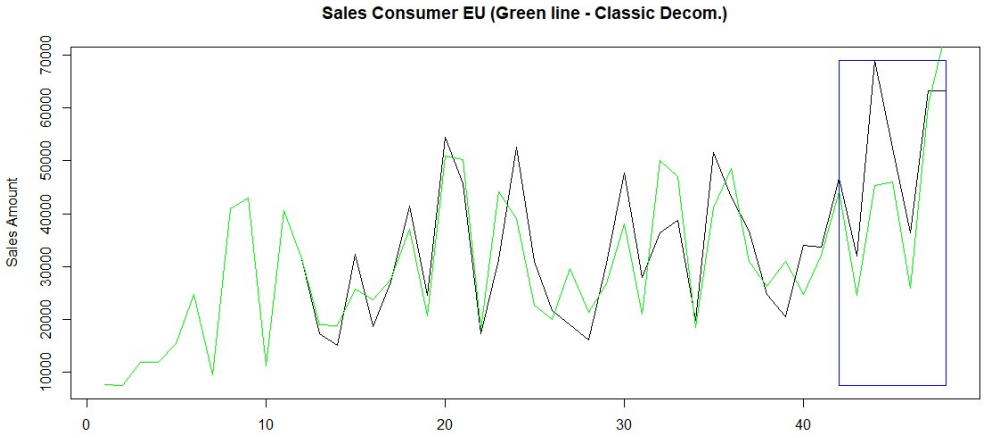


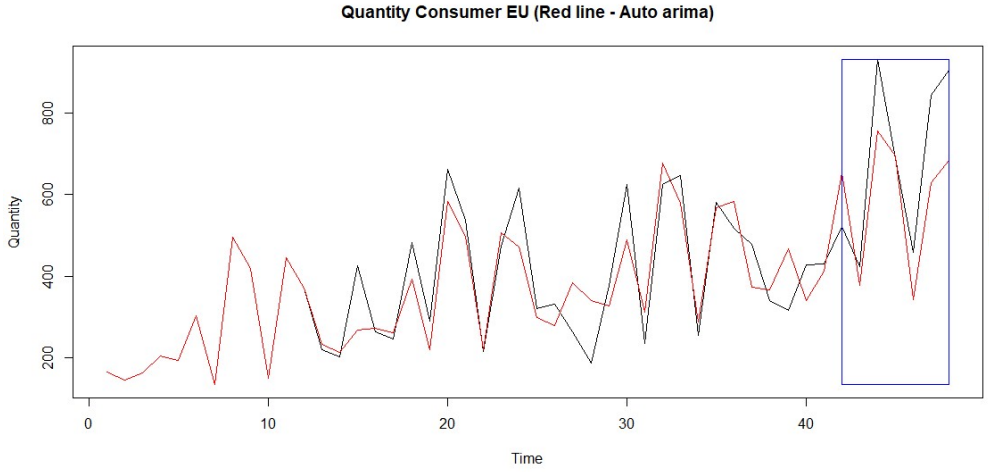


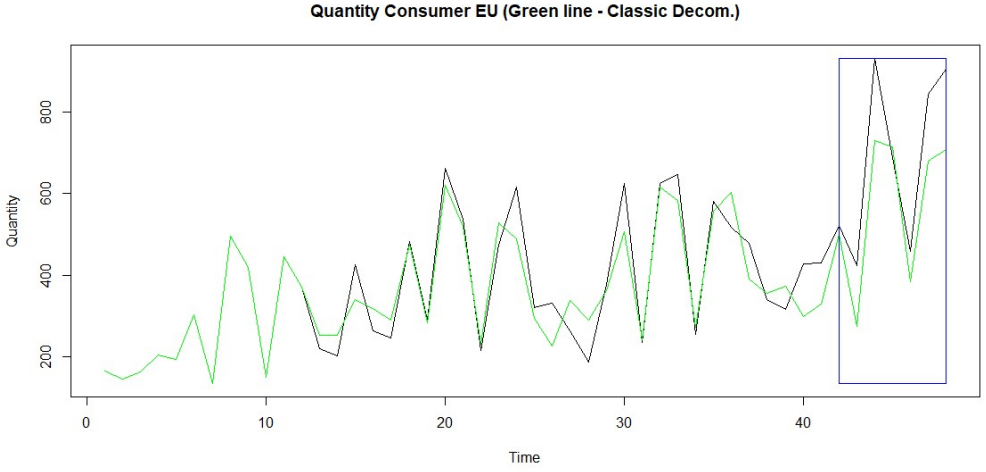
Consumer EU Division – Sales & Quantity forecasting

- ADF Test on auto arima model indicates some residual component, hence we isolate it using classic decomposition models
- MAPE values of Classic decomposition remain within acceptable levels(19-20). Forecast data is highlighted in Blue box









Conclusion

- **Based on data provided, Consumer APAC & Consumer EU segments have demonstrated most & consistent profitability**
- **Timeseries of this data shows strong growing trend for both segments in sales as well as demand along with seasonality. The model built can provide insight to how sales and demand will change in future**
- Model specifics:
- ADF tests on Auto Arima models on Sales & Quantity for both segments indicate that there may be some local regressive behaviour.
- Consumer APAC timeseries required smoothing over a longer term of 8 periods, as compared to the simpler Consumer EU time series which we smoothed with 6 periods
- With Classic decomposition models on smoothened series, ADF & KPSS test results indicate that we are better able to separate white noise, thus prove to be better choice for forecast modelling

	Consumer APAC		Consumer EU	
	Sales	Demand/Quantity	Sales	Demand/Quantity
Auto Arima-MAPE	11.77618	13.36779	15.75223	17.61852
Classic decomposition-MAPE	14.97376	14.38483	20.13542	19.71775
Classic decomposition Polynomial & coeffs	Multiplicative 2nd degree polynomial with x=0.5	Multiplicative 3rd degree polynomial with x=0.5	Multiplicative 3rd degree polynomial with x=0.4	Multiplicative 2nd degree polynomial with x=0.4