Time Series Project – Retail-Giant Sales Forecasting Case

Background:

"Global Mart" is an online store super giant having worldwide operations. The store caters to 7 different market segments and 3 major categories.

Aim:

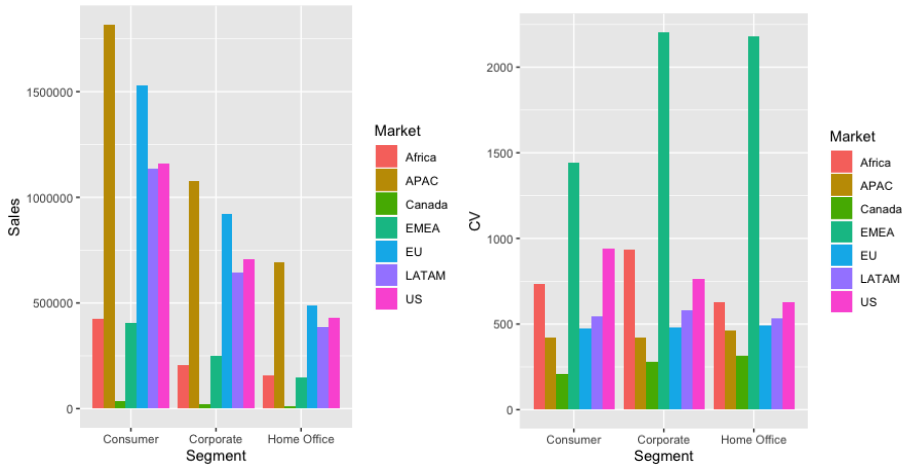
1. Analyse given data to identify 2 most profitable segments.
2. Forecast the sales and the demand for the next 6 months.

Note: CV (Coefficient of Variation)
= sd (Profit) * 100/mean (Profit)

First, we read the data and find that there are lots of missing value in Postal. Code, then we delete it for further calculating since it is not important for forecasting

Detail Steps:

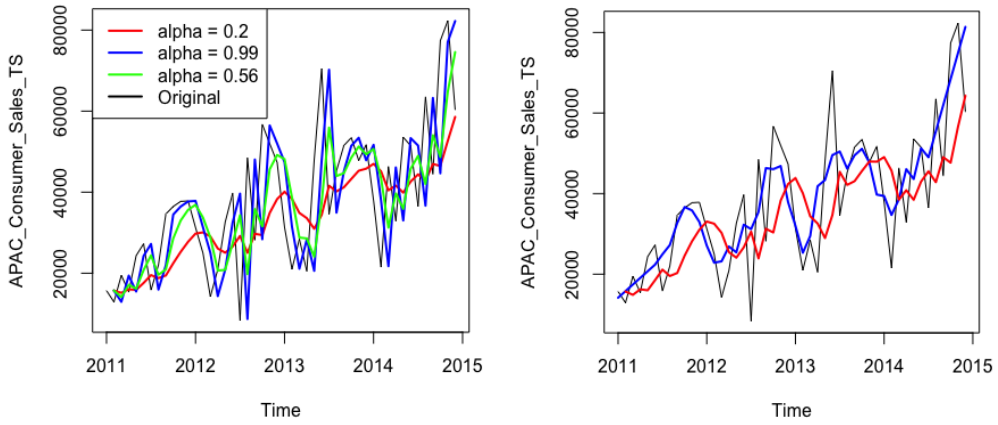
1. Data Preparation
   1) Read data → there are 3 levels of segments and 7 levels of markets
      So, we created 21 data subset buckets based on Market & Segment they belong
   2) Identify 2 most profitable segments:
      Data Aggregation → Aggregated data in each bucket by sales, quantity, profit & CV
      Then using Profit & CV to find the top2 most profitable segments
      → APAC_Consumer with sum(profit) = 222818
        EU_Consumer with sum(profit) = 188688
   3) Data Visualization → Plot Data



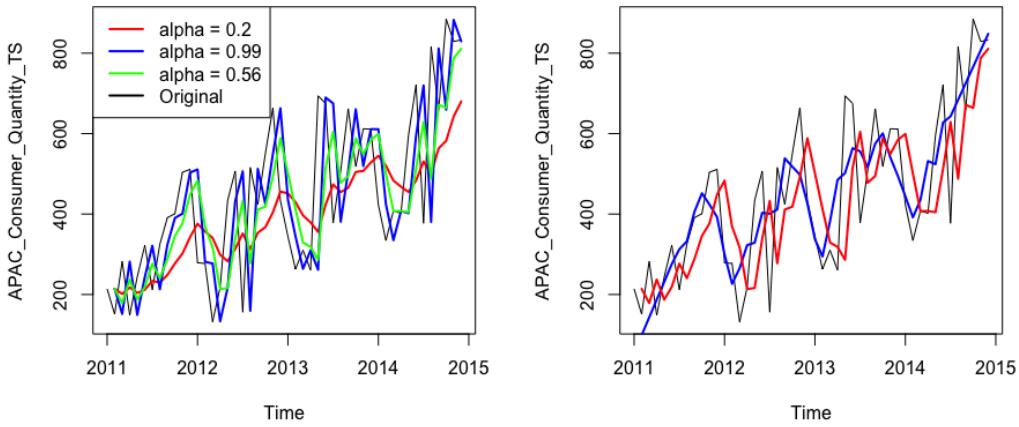
2. Modeling
    1) Subset top2 profitable market_segments
    2) Create Time Series of the top2 for first 48 months
    3) Smoothened time series using MA and Holt Winters smoothing
       → smoothen time series to identify trend & seasonality
    4) Time series data was divided into train (1-42 month), validation (43-48 month) & test sets (49-54 month)
    5) Model1: Linear model; Model2: Auto ARIMA Model

Plots:

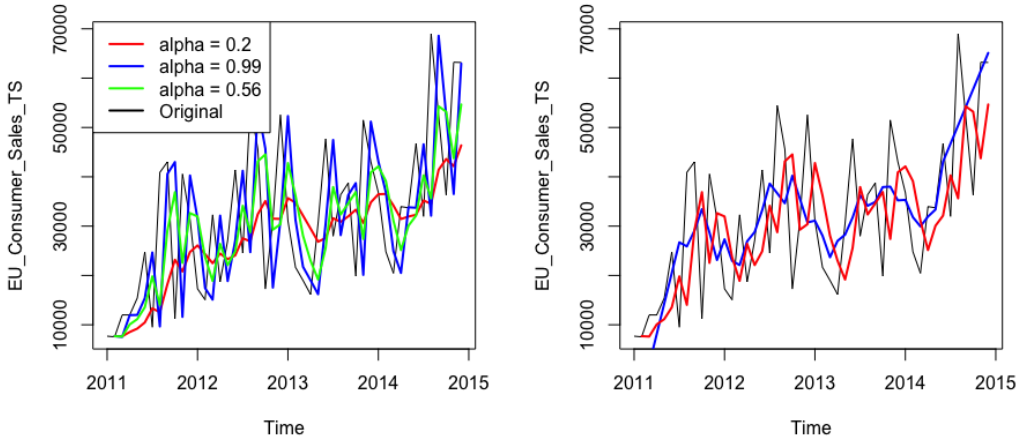
I. APAC Consumer Sales



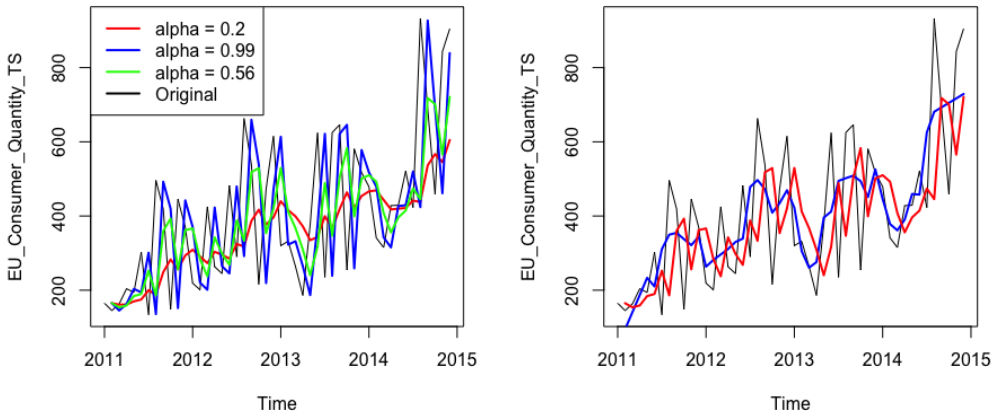
II. APAC Consumer Quantity



III. EU Consumer Sales



IV. EU Consumer Quantity



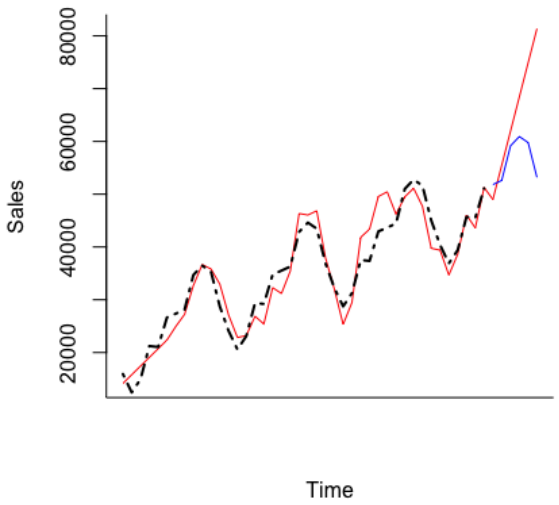
3. Model Evaluation
   1) Both models are evaluated using Mean Absolute Percentage Error (MAPE)
   2) ACF plots of residuals are used to check that it resembles white noise

4. Forecasting
   1) Use best model to forecast future 6 months Sales
   2) Repeat same for both Market-Segments Sales & Quantity
   3) Prepare 4 forecasts for top2 segments by Sales & Quantity

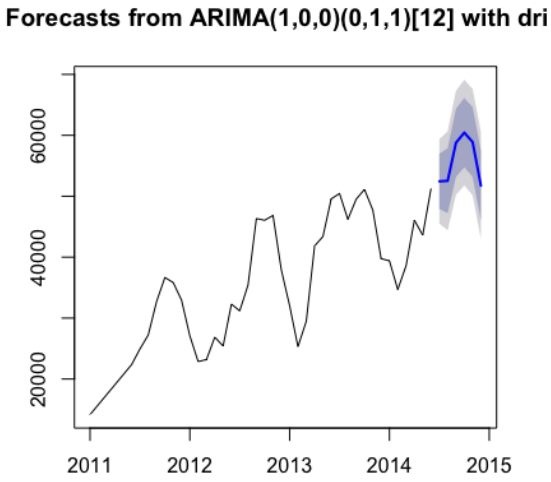
Results:

1a. APAC Consumer Sales Forecast on validation set

Linear Model Forecast

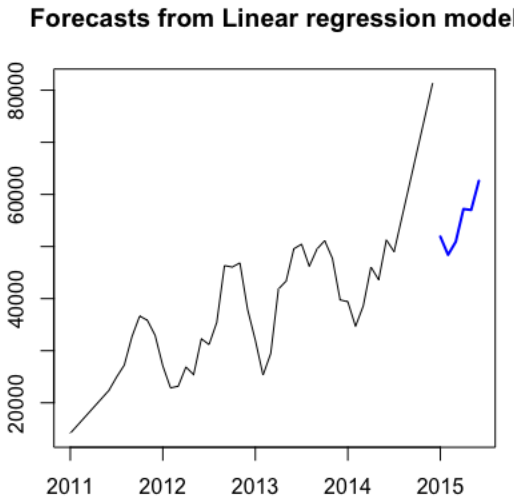


Auto ARIMA Model Forcast

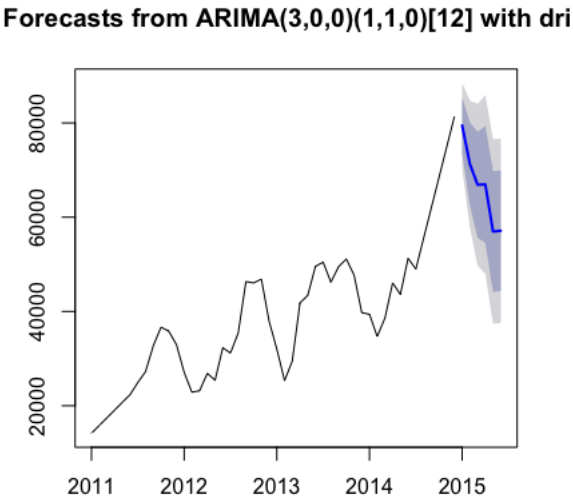


1b. APAC Consumer Sales Forecast on test set

Linear Model Forecast

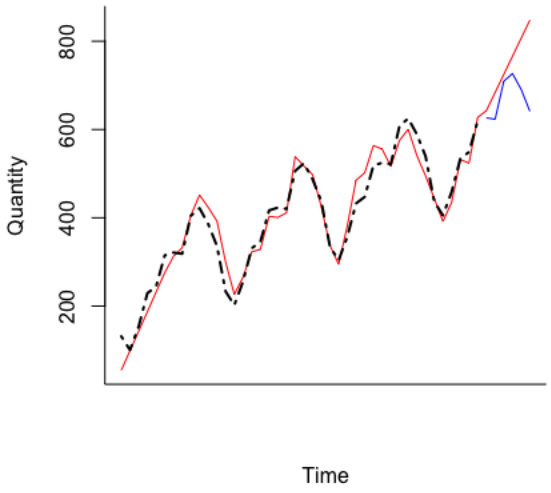


Auto ARIMA Model Forcast

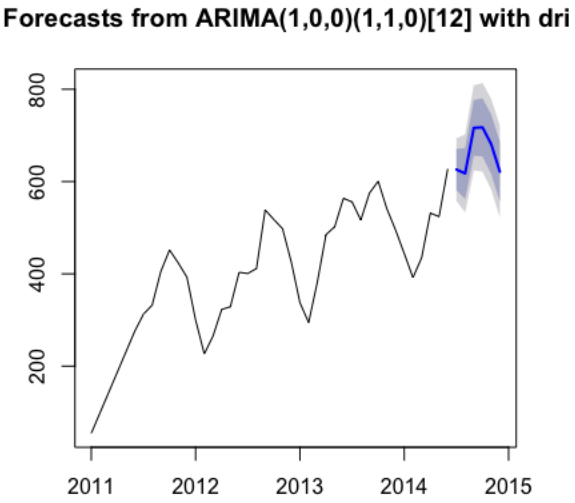


2a. APAC Consumer Quantity Forecast on validation set

Linear Model Forecast

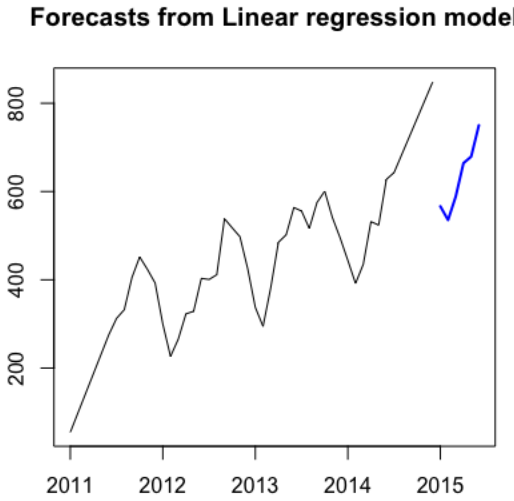


Auto ARIMA Model Forcast

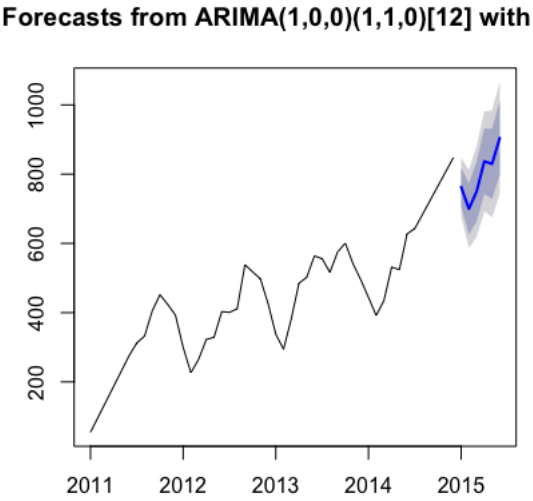


2b. APAC Consumer Sales Forecast on test set

Linear Model Forecast

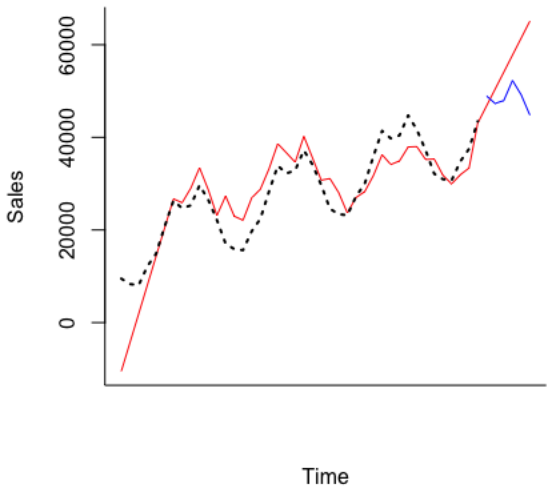


Auto ARIMA Model Forcast

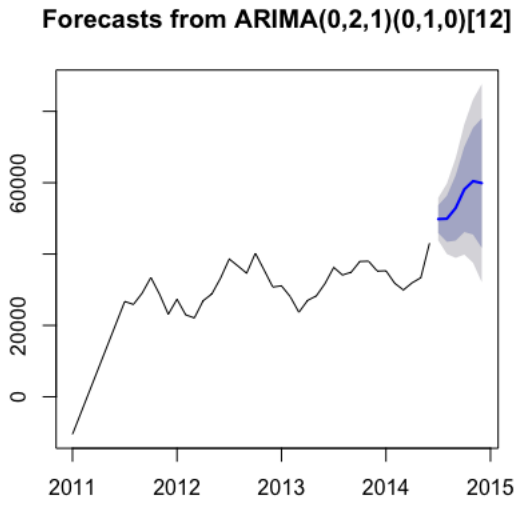


3a. EU Consumer Sales Forecast on validation set

Linear Model Forecast

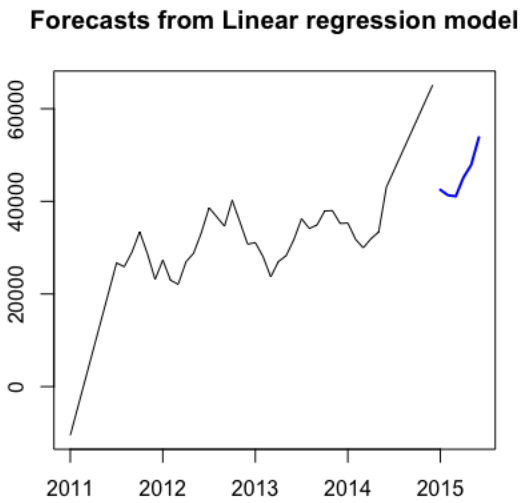


Auto ARIMA Model Forcast

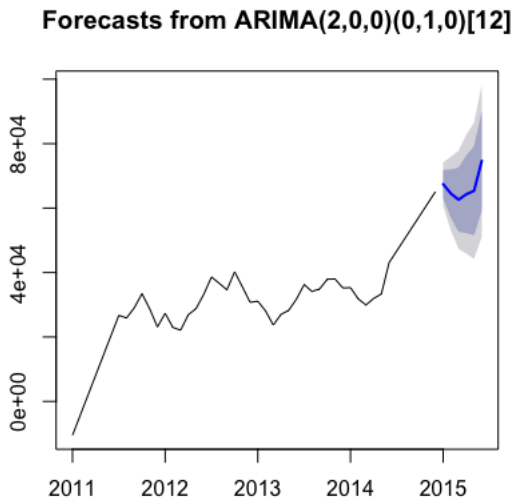


3b. EU Consumer Sales Forecast on test set

Linear Model Forecast

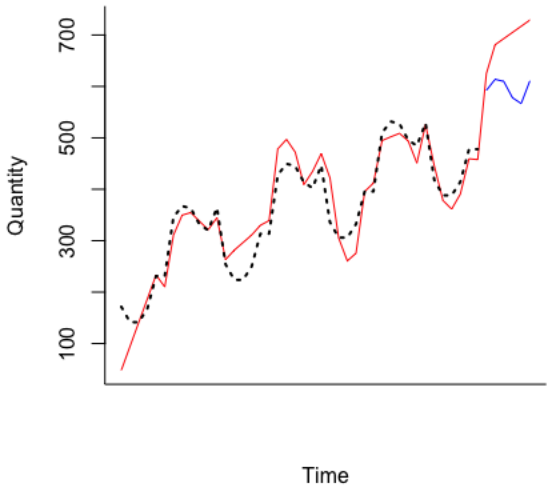


Auto ARIMA Model Forcast

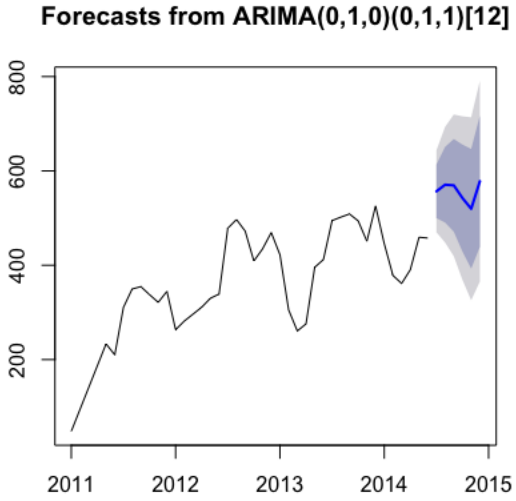


4a. EU Consumer Quantity Forecast on validation set

Linear Model Forecast

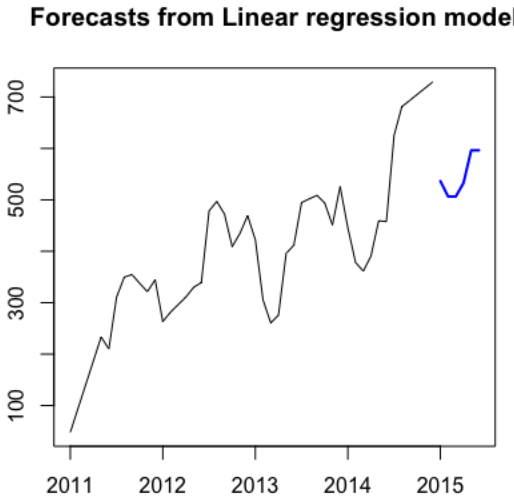


Auto ARIMA Model Forcast

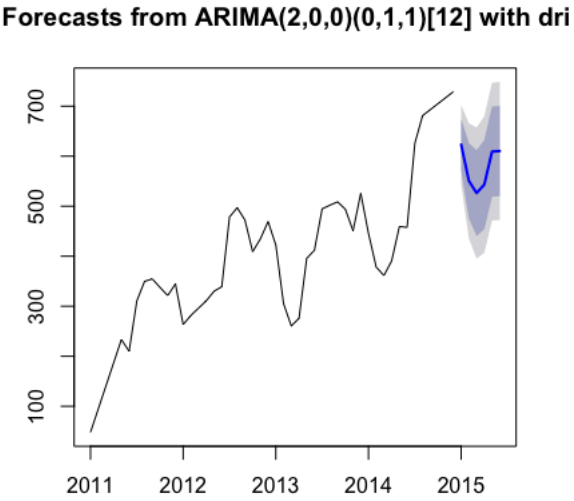


4b. EU Consumer Quantity Forecast on test set

Linear Model Forecast



Auto ARIMA Model Forcast



Conclusion:

1. The most 2 profitable market segments are APAC Consumers and EU Consumers
2. We create total 8 forecasting models for top2 segments out of which 4 best were selected for forecasting future 6 months sales & quantity for months Jan.2015 to Jun.2015
3. Summary of 4 key forecasts on test data:
    a) APAC Consumer Sales is likely to decrease next 6 months with large fluctuations
    b) APAC Consumer is likely to rise steeply in coming 6 months
    c) EU Consumer Sales may show slow rise in coming 6 months
    d) EU Consumer Quantity is like to drop during initial 1 or 2 months and then rise rapidly in next 3 months, eventually reaching a plateau.