

RETAIL-GIANT SALES FORECASTING

TIME SERIES ANALYSIS

Team Members

Praval Rastogi

Deepsi Shrivastava

Prakhar Shukla

Lakshmipathi Kakarla

Analysis Overview

- Business Objective
- Data Exploration
- Data Cleaning and Manipulation
- Data Analysis
- Time Series Modelling
- Conclusion

Business Objective

Business Objective

- Forecast the sales and the demand for the next 6 months

Data Exploration

Data Exploration

- Sales transaction data provided for four years (from Jan 2011 to Dec 2014)
- Contains 51290 sales transaction records with 24 attributes
- Significant Attributes:
 - Market
 - Segment
 - Sales - Total Value of the Sale
 - Quantity - Sold Quantity
 - Profit - Profit On Total Sale Value

Data Preparation and Processing

Data Preparation and Processing

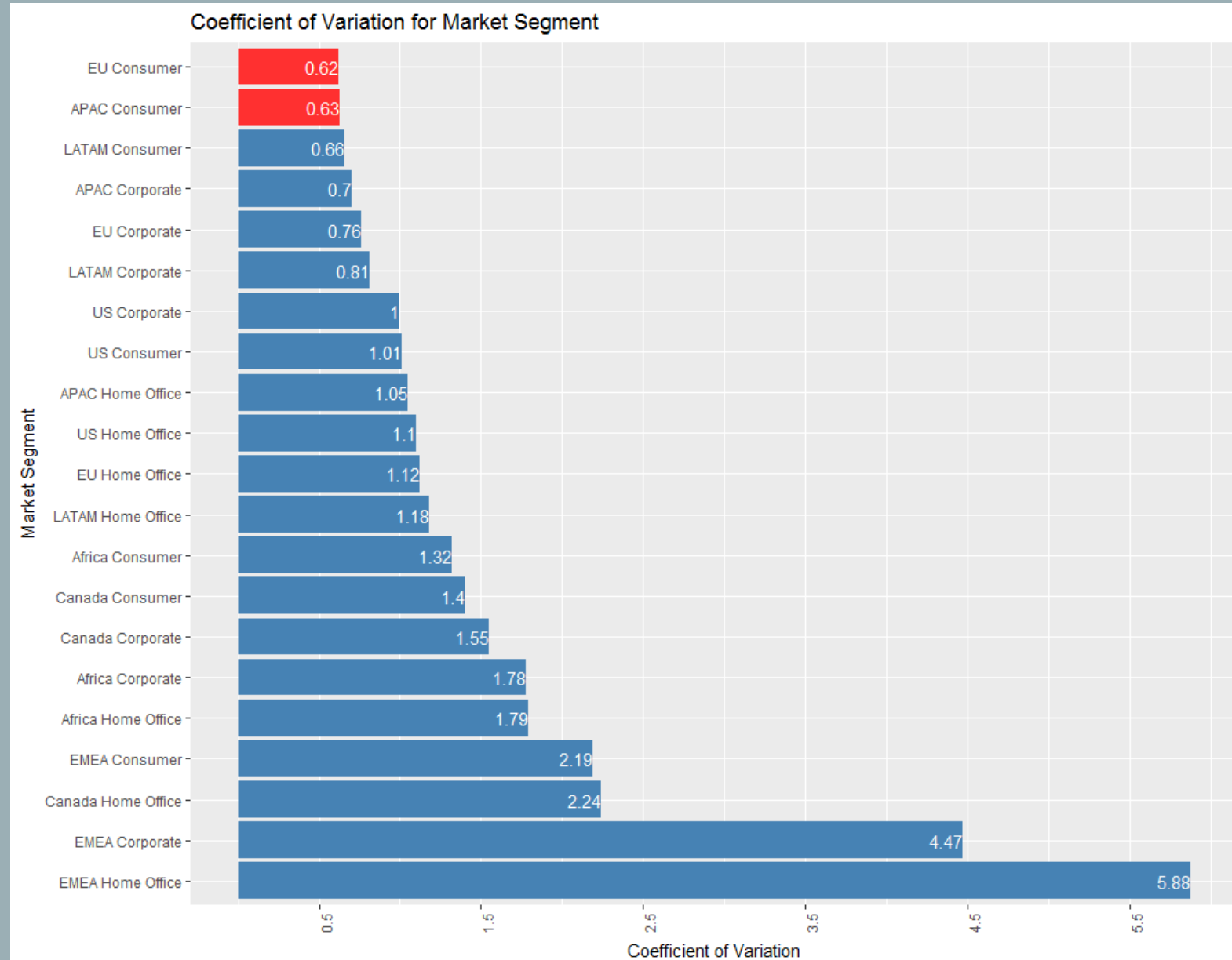
- Formatted “Order Date” and “Ship Date” attribute to proper date format
- Derived new attribute based on “Market” and “Segment” attributes to identify most profitable market segment
- Derived Year and Month from formatted “Order Date” attribute for time series analysis

Assumptions

- Missing Values:
 - Postal Code attribute having NA values. As per our analysis, Postal Code not present for other than US Market. Missing values not replaced as it's an insignificant attribute for time series analysis.

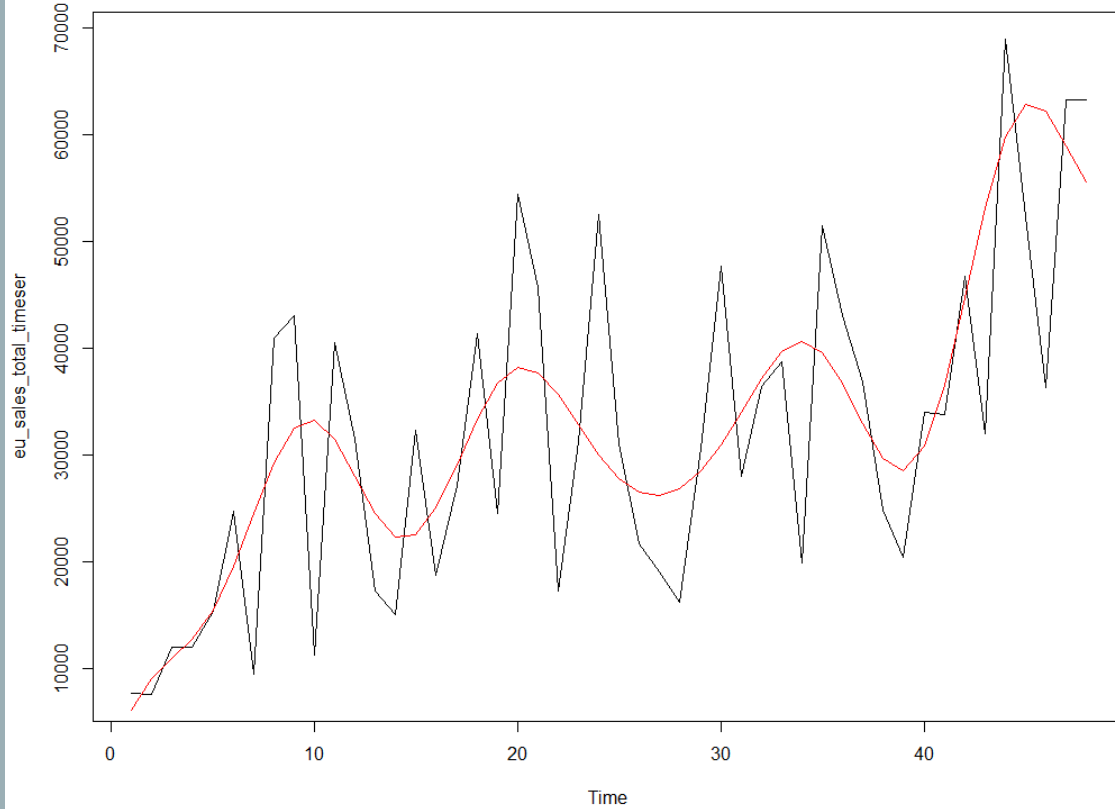
Data Analysis - Top 2 Profitable Market Segments

- Based on the least coefficient of variation of the Profit for all 21 market segments, following 2 market segments identified as top 2 most profitable market segments.
 - EU Consumer
 - APAC Consumer



Model Building :: EU Consumer - Sales Forecasting

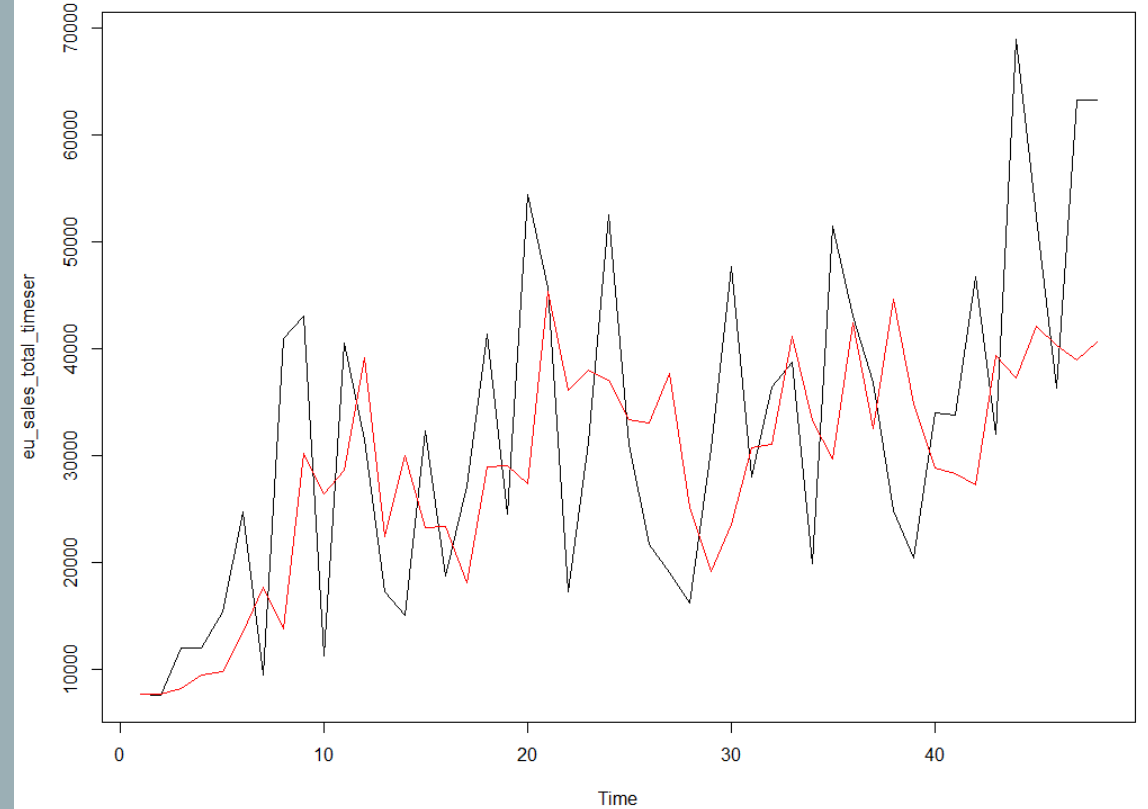
EU Consumer - Sales Forecasting - Classical Decomposition



Classical Decomposition

- ARIMA(0,0,0) with zero mean
- ADF Test: p-value = 0.01
- KPSS Test: p-value = 0.1
- AIC = 892.44
- MAPE : 31.58

EU Consumer - Sales Forecasting - Auto Arima

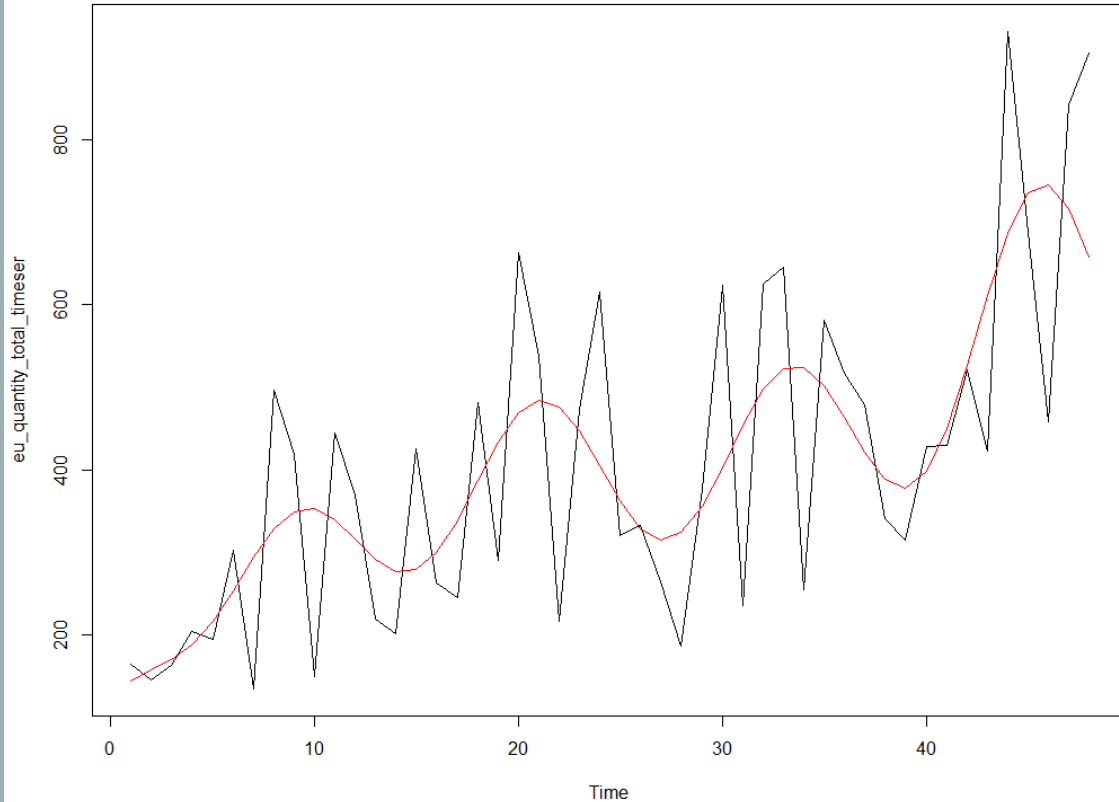


Auto ARIMA

- ARIMA(2,1,0)
- ADF Test: p-value = 0.01
- KPSS Test: p-value = 0.1
- AIC = 897.67
- MAPE : 28.92

Model Building :: EU Consumer - Quantity Forecasting

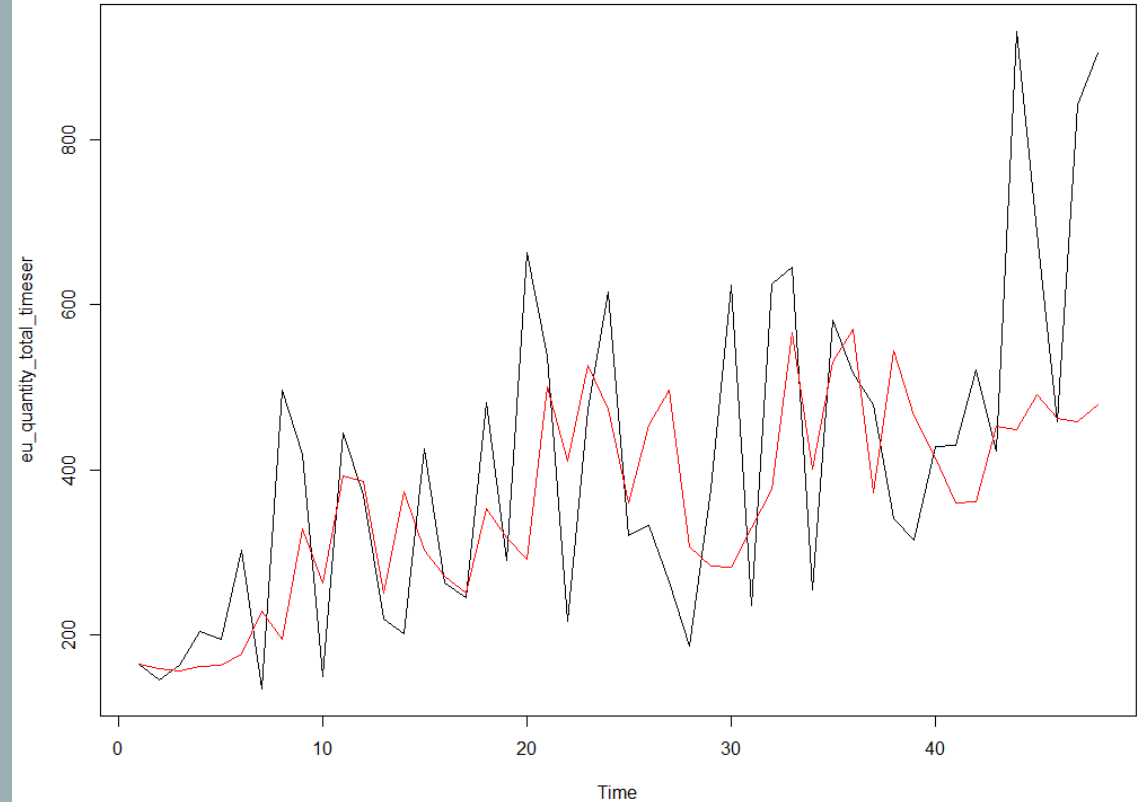
EU Consumer - Quantity Forecasting - Classical Decomposition



Classical Decomposition

- ARIMA(2,0,0) with zero mean
- ADF Test: p-value = 0.01
- KPSS Test: p-value = 0.1
- AIC = 497.79
- MAPE : 30.39

EU Consumer - Quantity Forecasting - Auto Arima

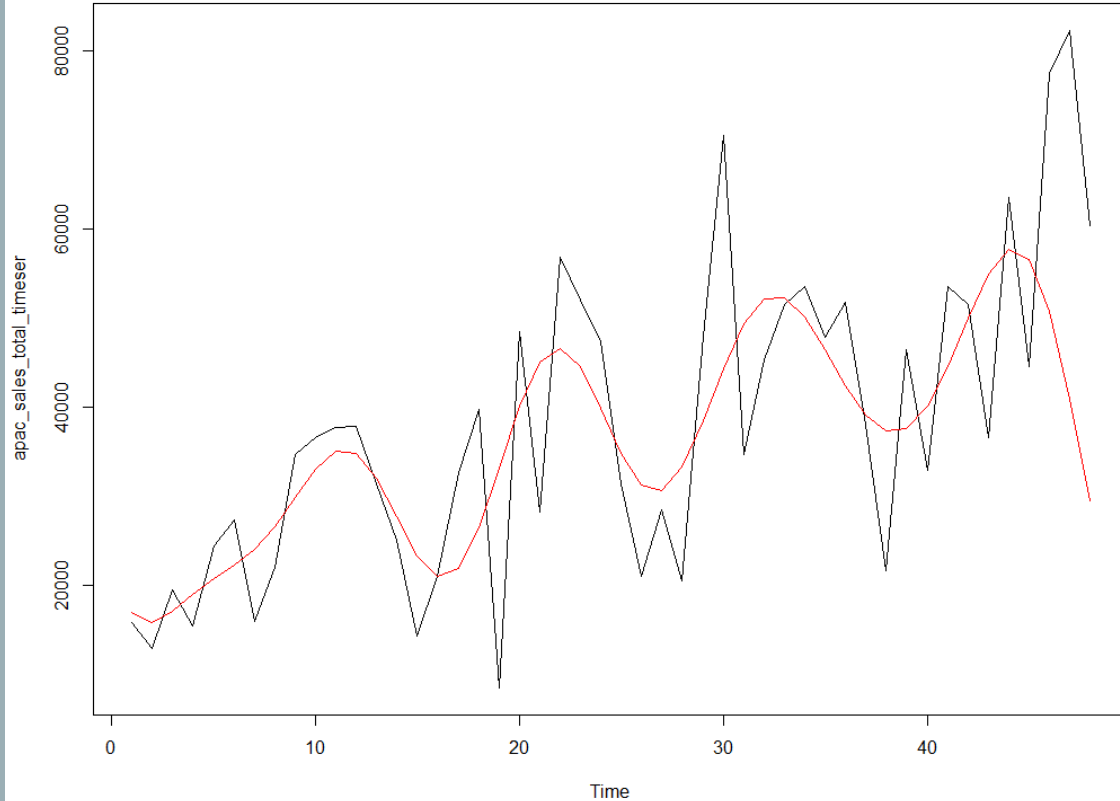


Auto ARIMA

- ARIMA(2,1,0)
- ADF Test: p-value = 0.04
- KPSS Test: p-value = 0.1
- AIC = 529.8
- MAPE : 30.13

Model Building :: APAC Consumer - Sales Forecasting

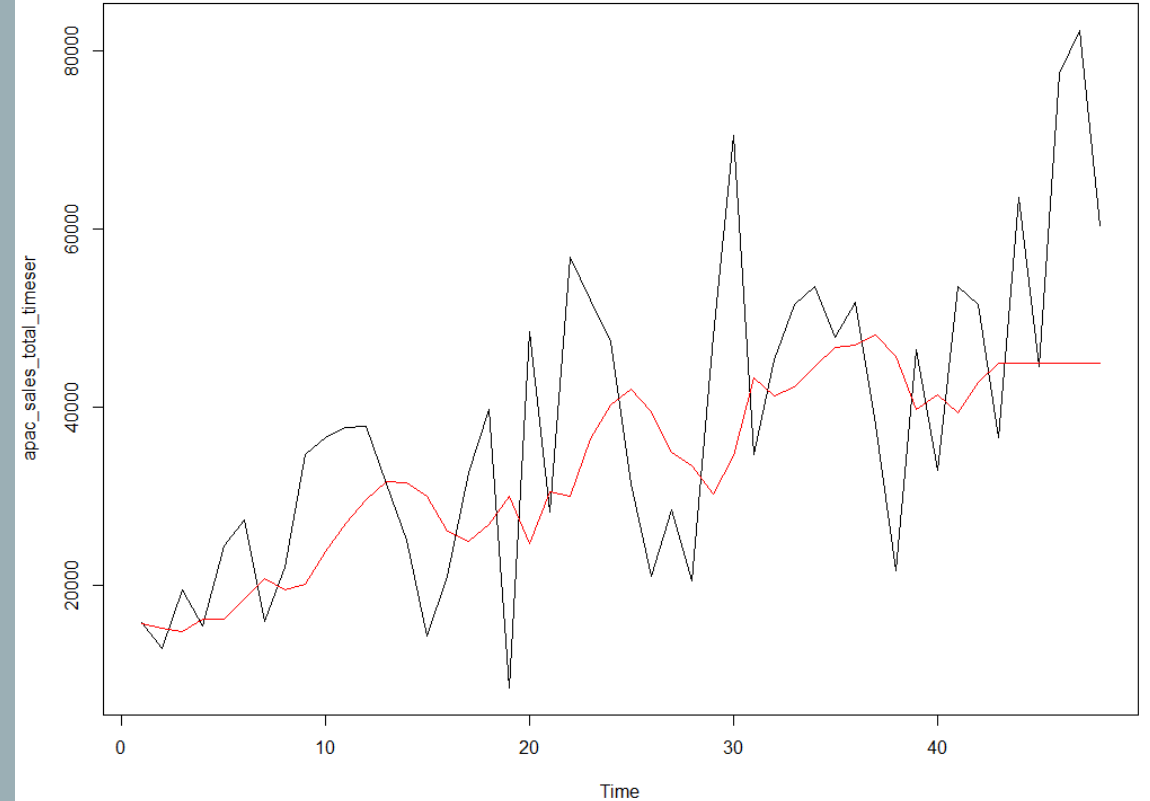
APAC Consumer - Sales Forecasting - Classical Decomposition



Classical Decomposition

- ARIMA(0,0,0) with zero mean
- ADF Test: p-value = 0.01
- KPSS Test: p-value = 0.1
- AIC = 888.99
- MAPE : 37.10

APAC Consumer - Sales Forecasting - Auto Arima

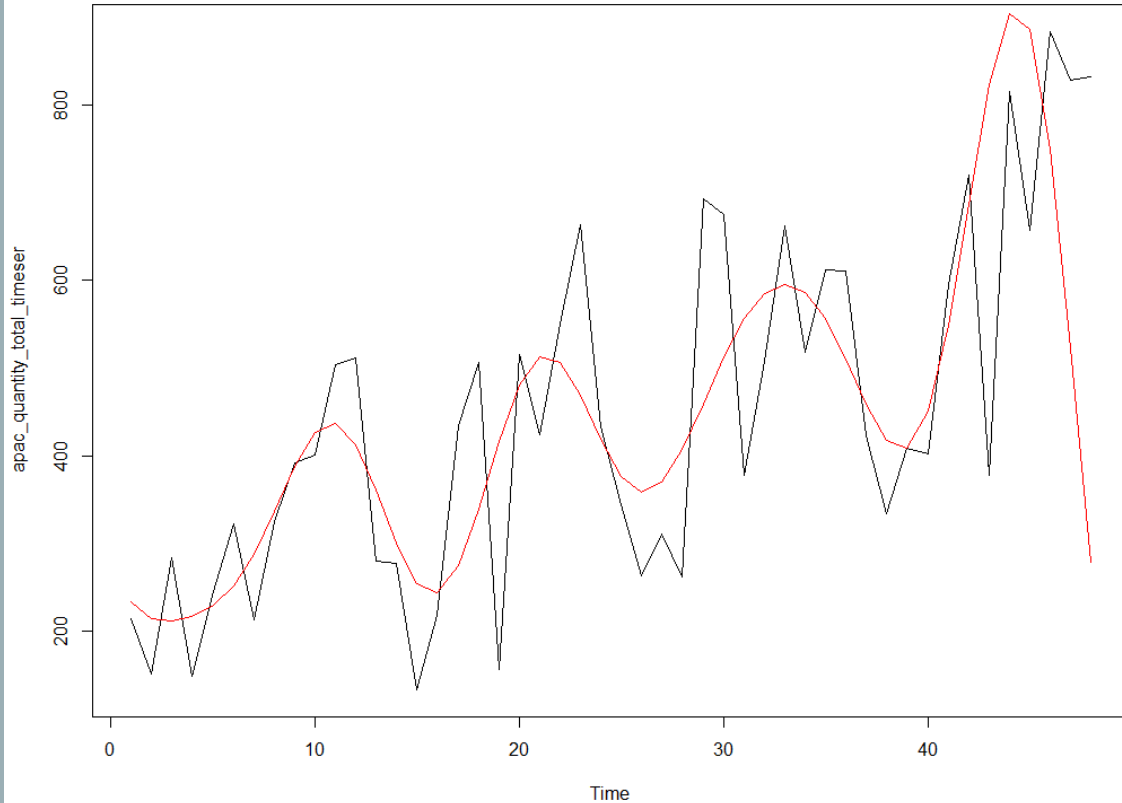


Auto ARIMA

- ARIMA(0,1,1)
- ADF Test: p-value = 0.01
- KPSS Test: p-value = 0.1
- AIC = 898.23
- MAPE : 27.68

Model Building :: APAC Consumer - Quantity Forecasting

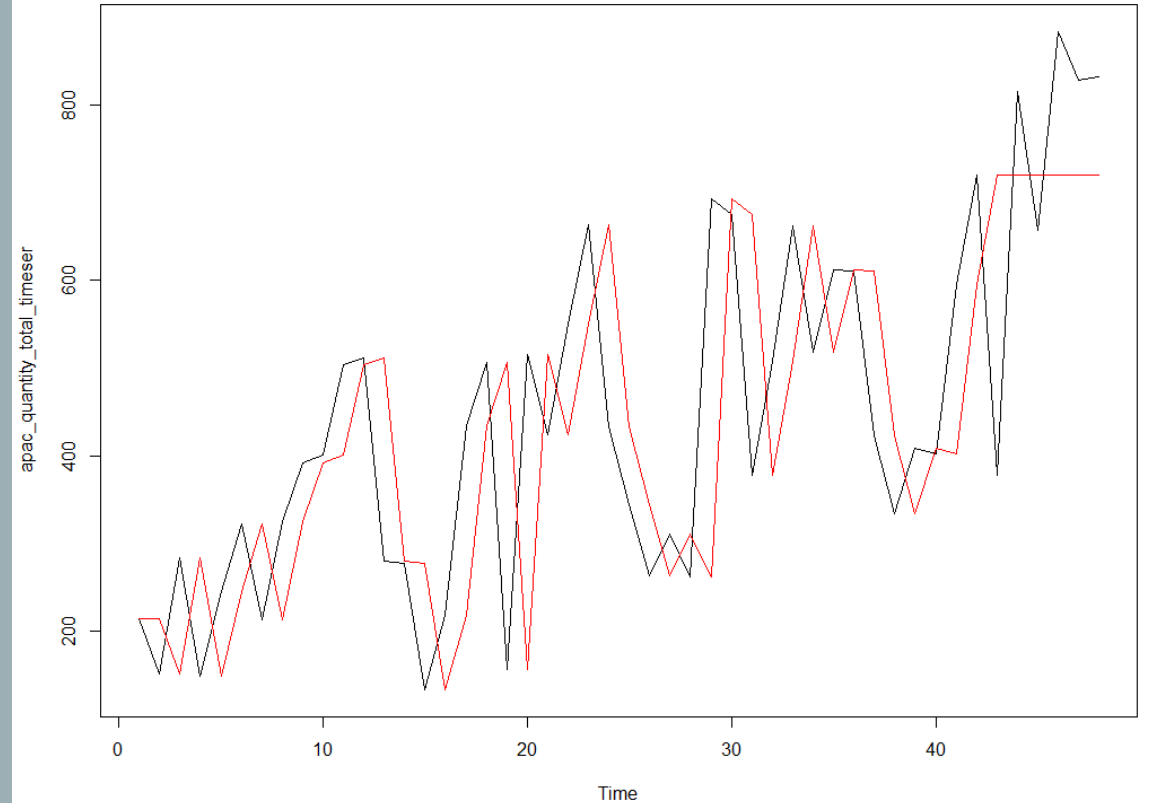
APAC Consumer - Quantity Forecasting - Classical Decomposition



Classical Decomposition

- ARIMA(0,0,0) with zero mean
- ADF Test: p-value = 0.01
- KPSS Test: p-value = 0.1
- AIC = 508.46
- MAPE : 47.11

APAC Consumer - Quantity Forecasting - Auto Arima



Auto ARIMA

- ARIMA(0,1,0)
- ADF Test: p-value = 0.01
- KPSS Test: p-value = 0.1
- AIC = 534.14
- MAPE : 26.244

Conclusion

- Sales across all analyzed market segments have shown global trend and periodic seasonal behavior
- Model Evaluation
 - EU Consumer - Sales Forecasting
 - Based on MAPE value, Auto ARIMA Model identified as better model than Classical Decomposition
 - EU Consumer - Quantity Forecasting
 - Based on MAPE value, Auto ARIMA Model identified as better model than Classical Decomposition
 - APAC Consumer - Sales Forecasting
 - Based on MAPE value, Auto ARIMA Model identified as better model than Classical Decomposition
 - APAC Consumer - Quantity Forecasting
 - Based on MAPE value, Auto ARIMA Model identified as better model than Classical Decomposition