Global Mart – Time Series Analysis

SUBMISSION

Group members:

1. Vikas Jangra
2. Sudha Choudhary
3. Anjali Nair
4. Vernon Fernandes

Abstract

DATA UNDERSTANDING

- **Market Attributes** — 7 geographical market sector that the customer belongs to
- **Segment Attributes** — 3 major categories; consumer, corporate & home office
- **Data Period** — 4 years of time series data

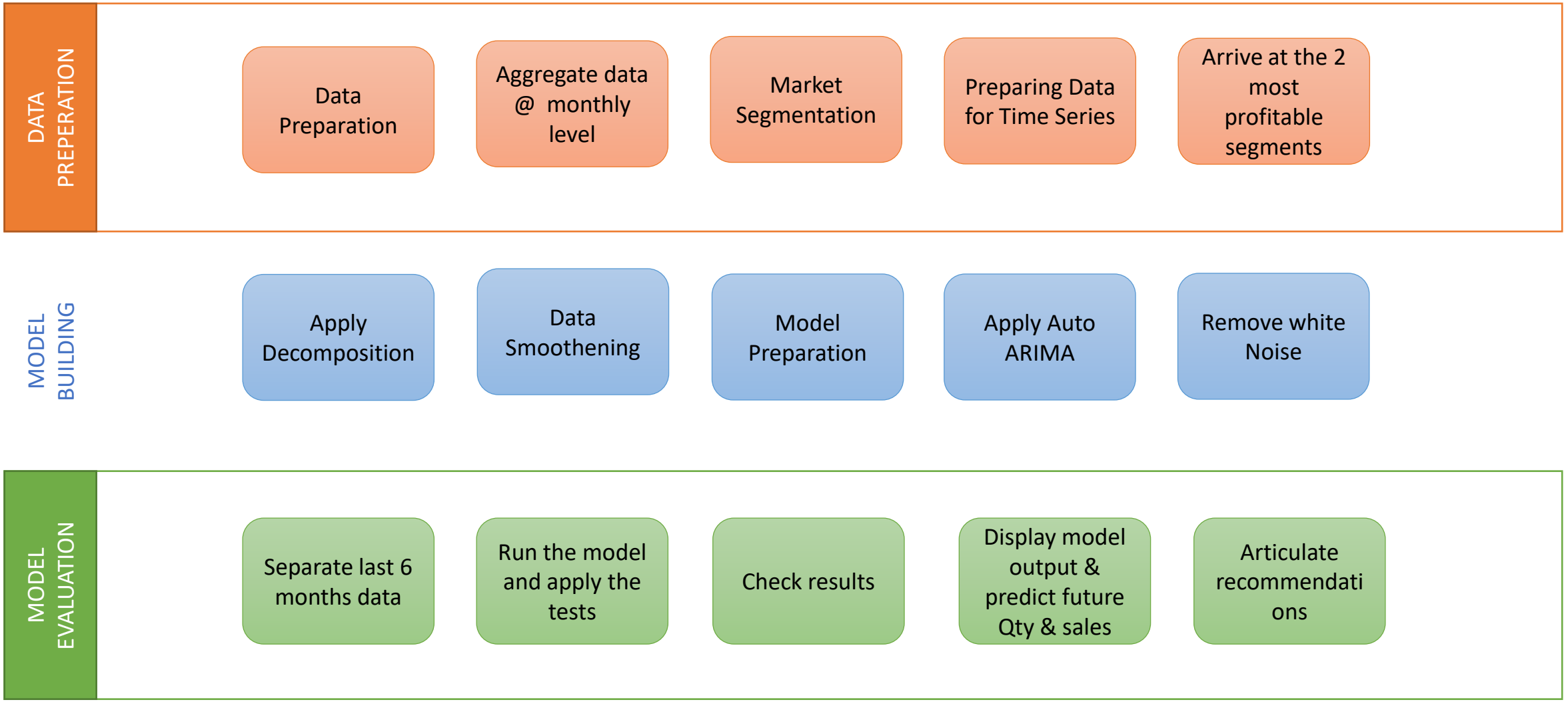
DATA PREPERATION

- **21 Subsets** — A combination of market attributes **X** segment attributes
- **Data Conversion** — convert the transaction-level data into a time series
- **Arrive at 2** — most consistent profitable segments

MODEL BUILDING & EVALUATION

- **Forecast** — sales and quantity for the next 6 months
- **Model Building** — using classical decomposition and auto ARIMA for forecasting
- **Model Evaluation** — using the out-of-sample figures

Steps


DATA PREPERATION
Data Preparation
Aggregate data @ monthly level
Market Segmentation
Preparing Data for Time Series
Arrive at the 2 most profitable segments
MODEL BUILDING
Apply Decomposition
Data Smoothening
Model Preparation
Apply Auto ARIMA
Remove white Noise
MODEL EVALUATION
Separate last 6 months data
Run the model and apply the tests
Check results
Display model output & predict future Qty & sales
Articulate recommendations

Top 2 Most Profitable Segments

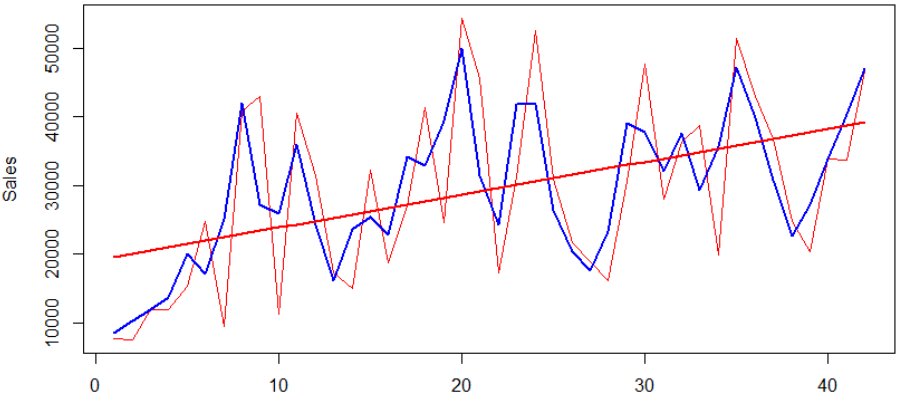
Rank	Group.1	sum_of_profit	mean_of_profit	sd_of_profit	CoV
1	APAC-Consumer	2,22,818	4,642	2,934	0.6321323
2	EU-Consumer	1,88,688	3,931	2,454	0.6243052
3	US-Consumer	1,34,119	2,794	2,829	1.01239
4	APAC-Corporate	1,29,737	2,703	1,887	0.6980869
5	EU-Corporate	1,23,394	2,571	1,964	0.7638072

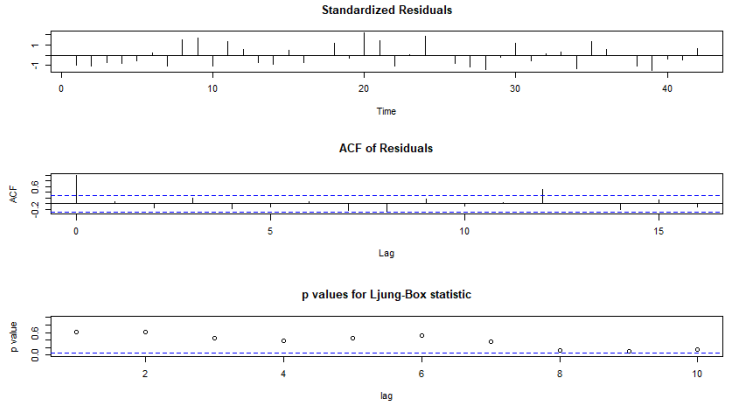
- Showcasing a list of top 5 by total sum of profit
- Group is a combination of Market & Segment

- APAC Consumer & EU Consumer are the 2 most consistently profitable segments based by CoV

EU Consumer - Sales

Classical Decomposition





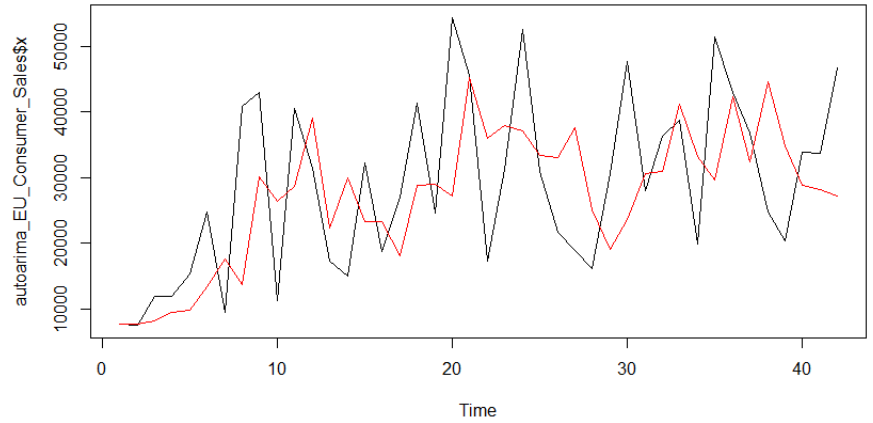

```
        Augmented Dickey-Fuller Test

data:  resi_EU_Consumer_Sales
Dickey-Fuller = -3.7581, Lag order = 3, p-value = 0.03254
alternative hypothesis: stationary
```

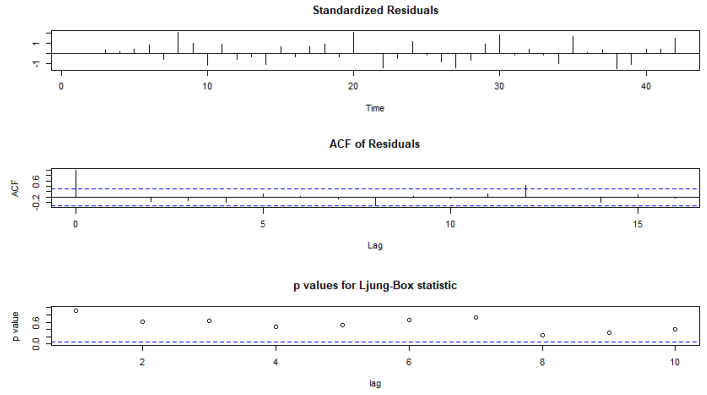
```
        KPSS Test for Level Stationarity

data:  resi_EU_Consumer_Sales
KPSS Level = 0.099616, Truncation lag parameter = 3, p-value = 0.1
```

- No Seasonality. Upward trend has been observed.| No Local trend thru ACF | Auto ARIMA confirms that there is no evidence of a local trend | Test confirms that residual is stationery

- Forecast with ARIMA (0,0,0) | MAPE: 57.26%

ARIMA






```
        Augmented Dickey-Fuller Test

data:  resi_auto_arima_EU_Consumer_Sales
Dickey-Fuller = -4.3522, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

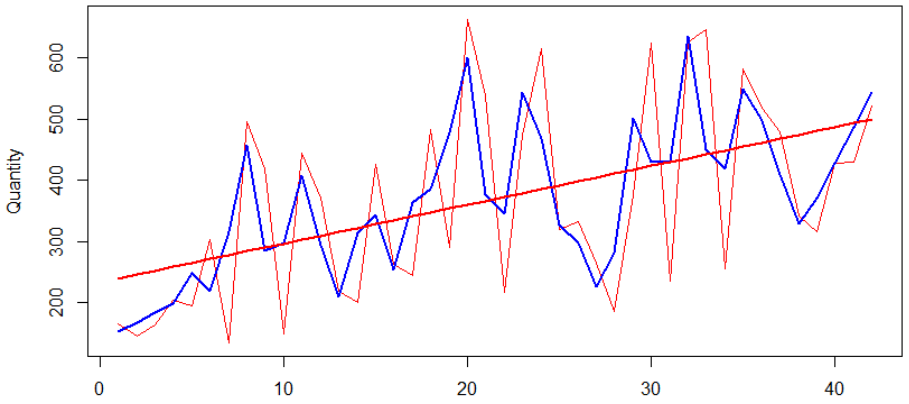
```
        KPSS Test for Level Stationarity

data:  resi_auto_arima_EU_Consumer_Sales
KPSS Level = 0.067962, Truncation lag parameter = 3, p-value = 0.1
```

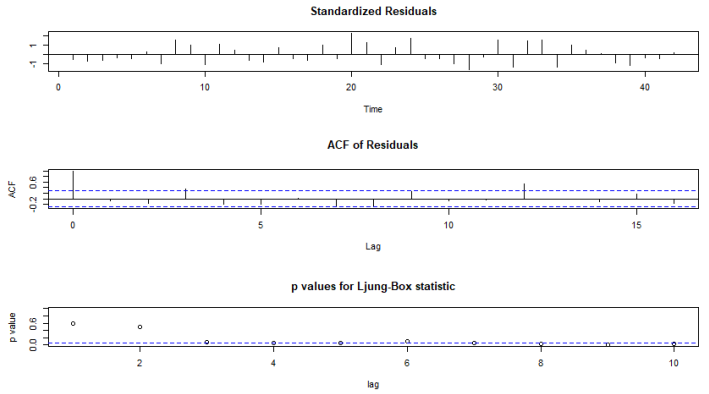
- Test confirms that residual is stationery | Forecast with ARIMA (2,1,0) | MAPE: 28.93%

AUTO ARIMA IS A VERY GOOD MODEL FOR US

EU Consumer - Quantity

Classical Decomposition



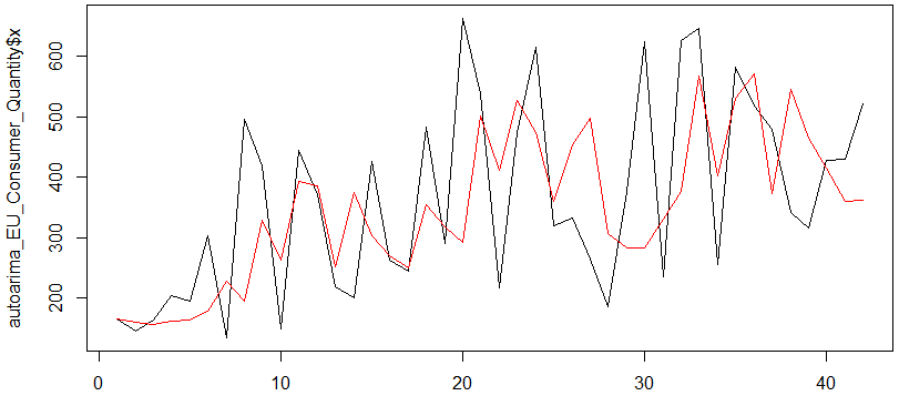


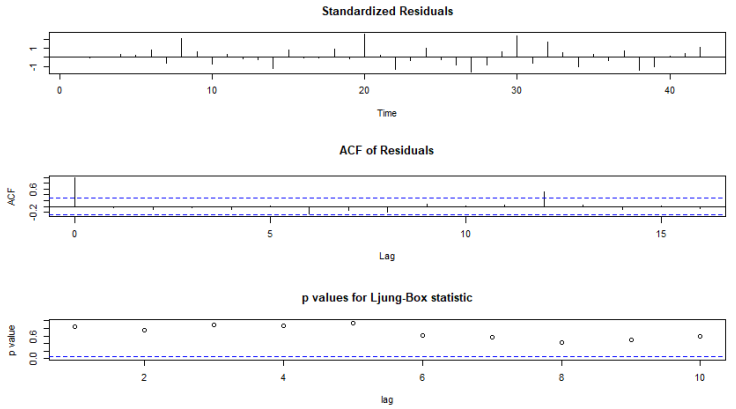

```
        Augmented Dickey-Fuller Test

data:  resi_EU_Consumer_Quantity
Dickey-Fuller = -3.1977, Lag order = 3, p-value = 0.1015
alternative hypothesis: stationary


        KPSS Test for Level Stationarity

data:  resi_EU_Consumer_Quantity
KPSS Level = 0.079663, Truncation lag parameter = 3, p-value = 0.1
```

- No Seasonality. Upward trend has been observed.| No Local trend thru ACF | Auto ARIMA confirms that there is no evidence of a local trend | Test confirms that residual is stationery

- Forecast with ARIMA (0,0,0) | MAPE: 60.45%

ARIMA



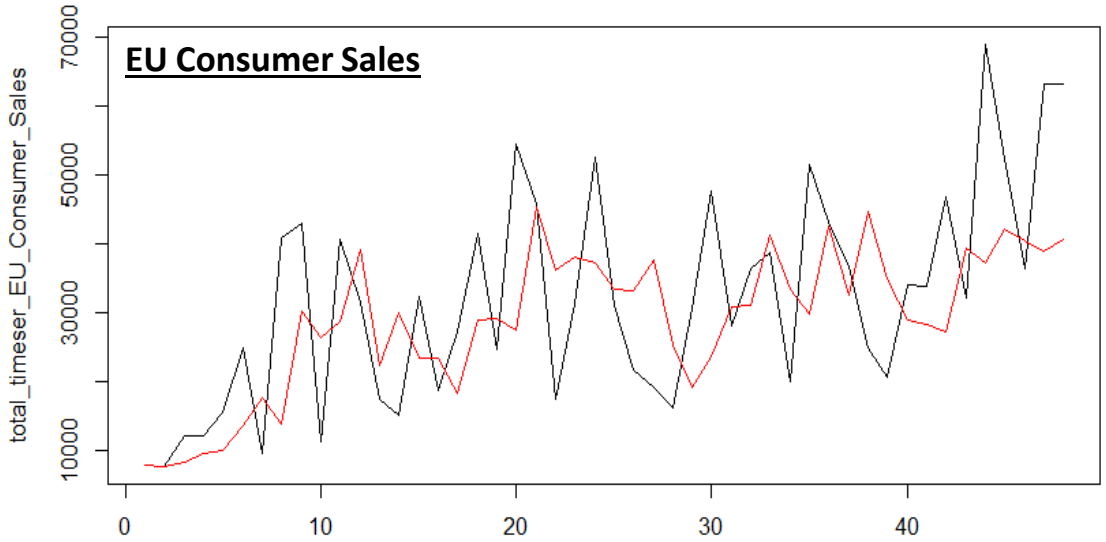



```
        Augmented Dickey-Fuller Test

data:  resi_auto_arima_EU_Consumer_Quantity
Dickey-Fuller = -3.5969, Lag order = 3, p-value = 0.04521
alternative hypothesis: stationary


        KPSS Test for Level Stationarity

data:  resi_auto_arima_EU_Consumer_Quantity
KPSS Level = 0.056232, Truncation lag parameter = 3, p-value = 0.1
```

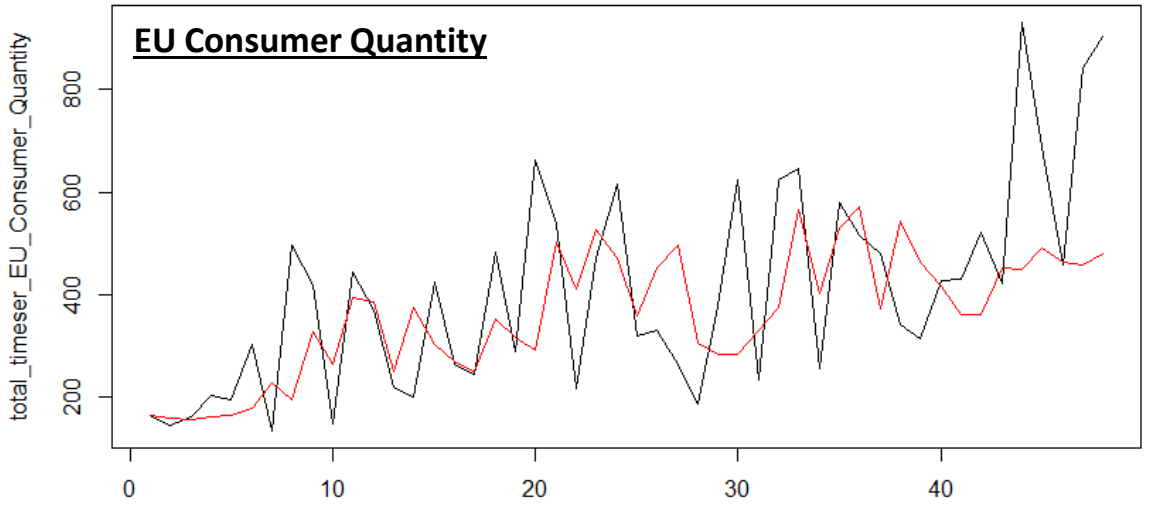
- Test confirms that residual is stationery | Forecast with ARIMA (2,1,0) | MAPE: 30.13%

AUTO ARIMA IS A VERY GOOD MODEL FOR US

Forecast EU Customer (Sales & Quantity)



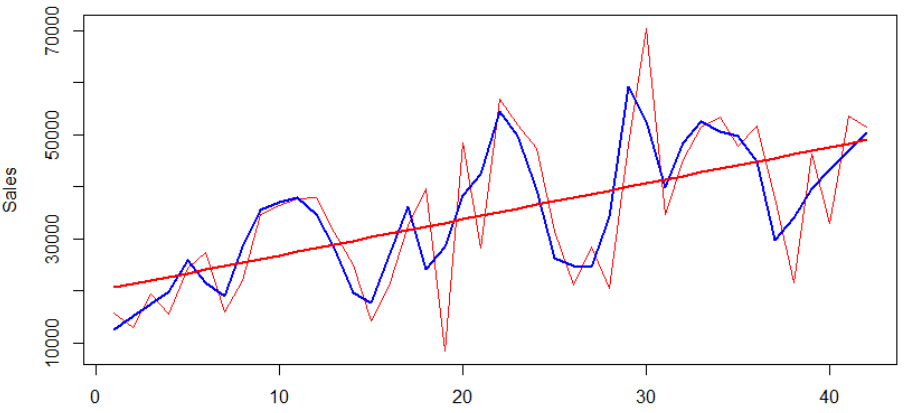
- EU Consumer Sales indicates a rise in the following 6 months.

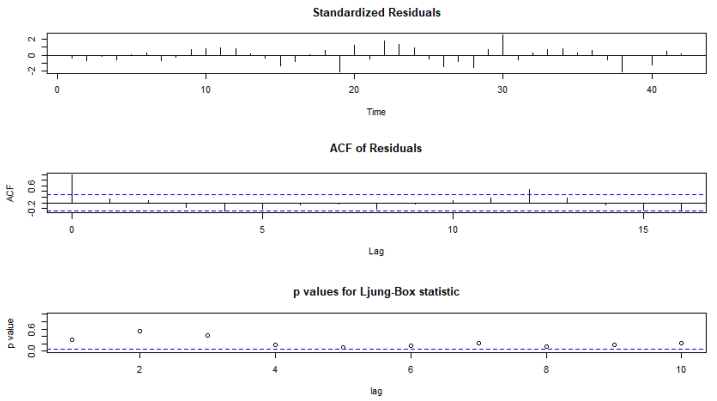


- EU Consumer Quantity is indicates a slow rise and the last 2 months showcases a plateau

APAC Consumer - Sales

Classical Decomposition



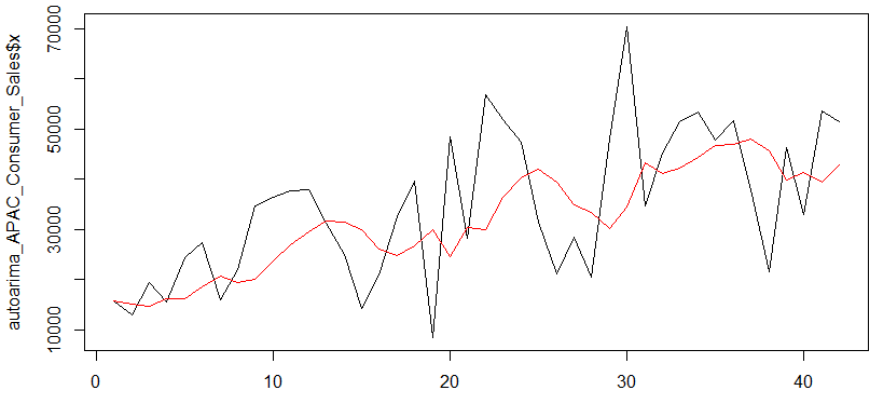


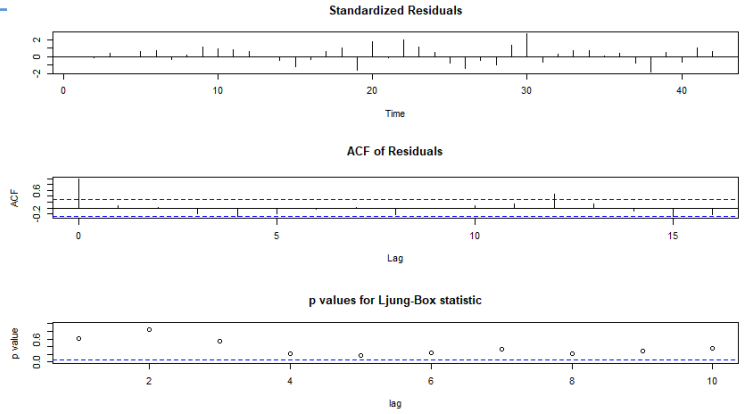

```
        Augmented Dickey-Fuller Test

data:  resi_APAC_Consumer_Sales
Dickey-Fuller = -4.0908, Lag order = 3, p-value = 0.01562
alternative hypothesis: stationary
```

```
        KPSS Test for Level Stationarity

data:  resi_APAC_Consumer_Sales
KPSS Level = 0.048624, Truncation lag parameter = 3, p-value = 0.1
```

- No Seasonality but the graph shows upward trend . No Local trend thru ACF | Auto ARIMA confirms that there is no evidence of a local trend | Test confirms that residual is stationery

ARIMA






```
        Augmented Dickey-Fuller Test

data:  resi_auto_arima_APAC_Consumer_Sales
Dickey-Fuller = -4.2563, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

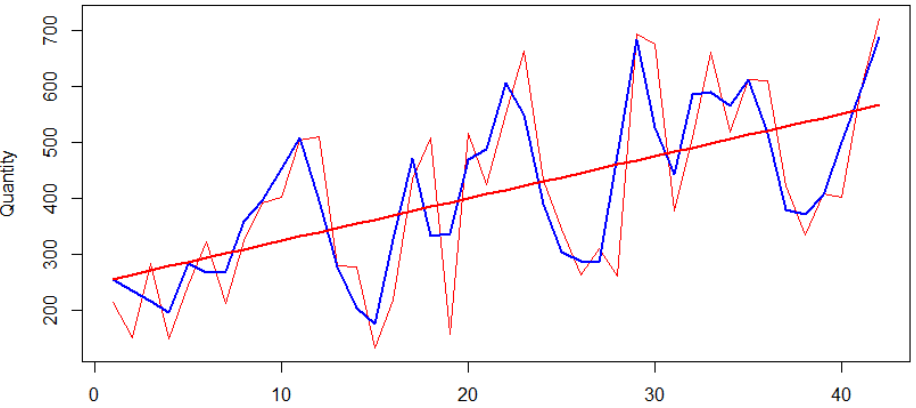
```
        KPSS Test for Level Stationarity

data:  resi_auto_arima_APAC_Consumer_Sales
KPSS Level = 0.043931, Truncation lag parameter = 3, p-value = 0.1
```

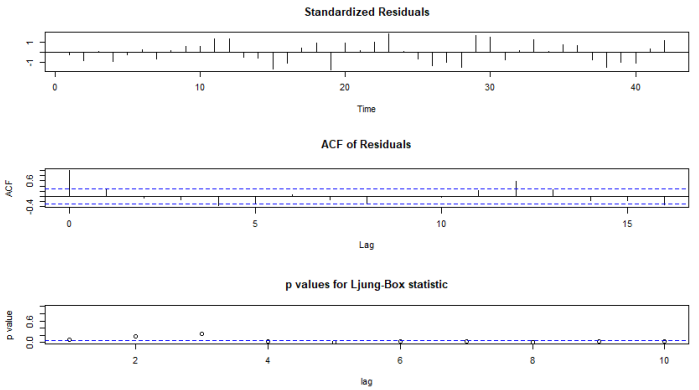
- Test confirms that residual is stationery | Forecast with ARIMA (0,1,1) | MAPE: 27.68%

AUTO ARIMA IS A VERY GOOD MODEL FOR US

APAC Consumer - Quantity

Classical Decomposition





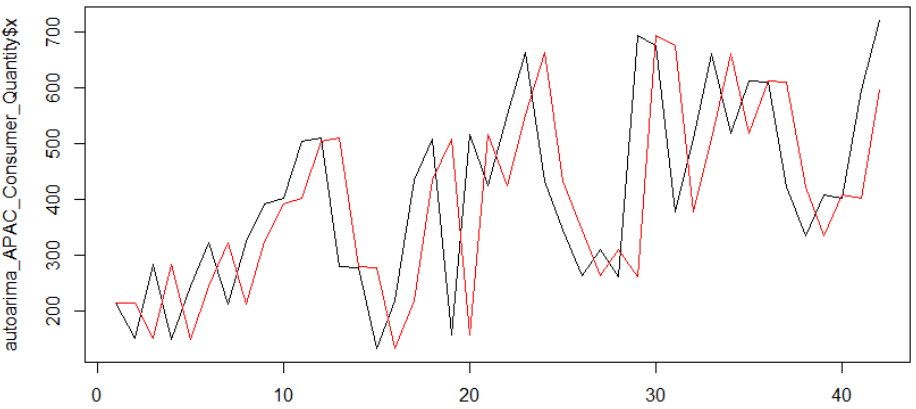

```
        Augmented Dickey-Fuller Test

data:  resi_APAC_Consumer_Quantity
Dickey-Fuller = -4.58, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

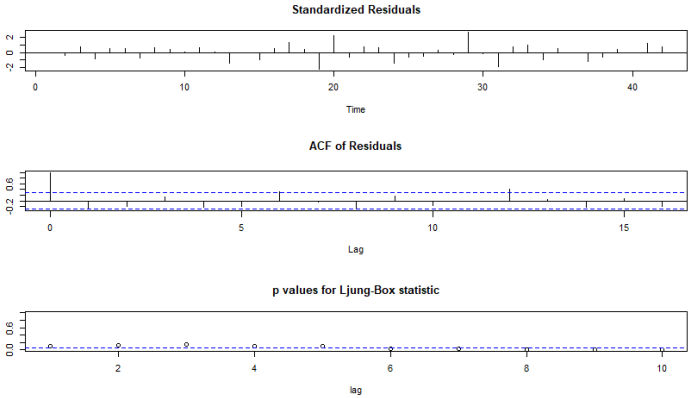
```
        KPSS Test for Level Stationarity

data:  resi_APAC_Consumer_Quantity
KPSS Level = 0.043212, Truncation lag parameter = 3, p-value = 0.1
```

- No Seasonality. Upward trend has been observed.| No Local trend thru ACF | Auto ARIMA confirms that there is no evidence of a local trend | Test confirms that residual is stationery

- Forecast with ARIMA (0,0,0) | MAPE: 59.60%

ARIMA





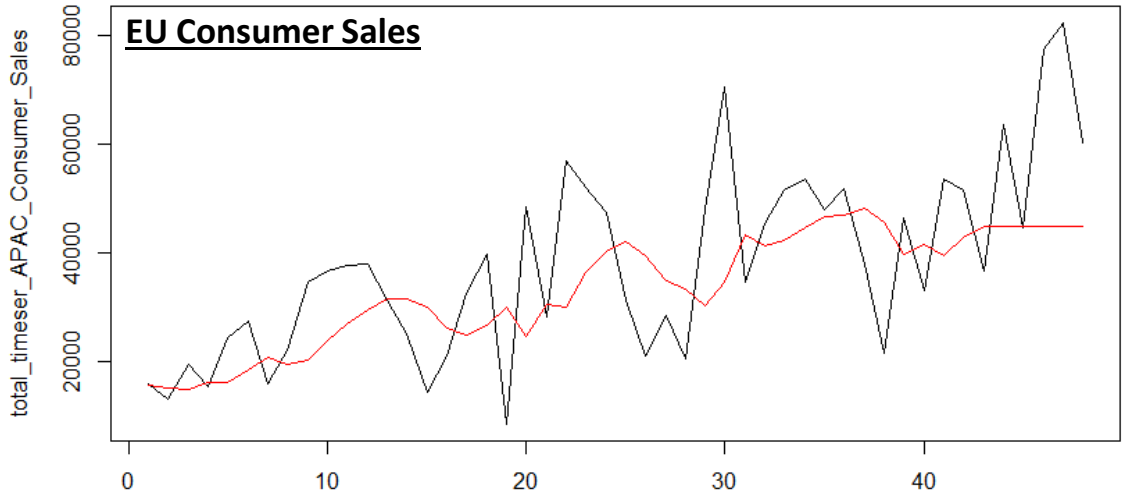

```
        Augmented Dickey-Fuller Test

data:  resi_auto_arima_APAC_Consumer_Quantity
Dickey-Fuller = -4.3326, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

```
        KPSS Test for Level Stationarity

data:  resi_auto_arima_APAC_Consumer_Quantity
KPSS Level = 0.04642, Truncation lag parameter = 3, p-value = 0.1
```

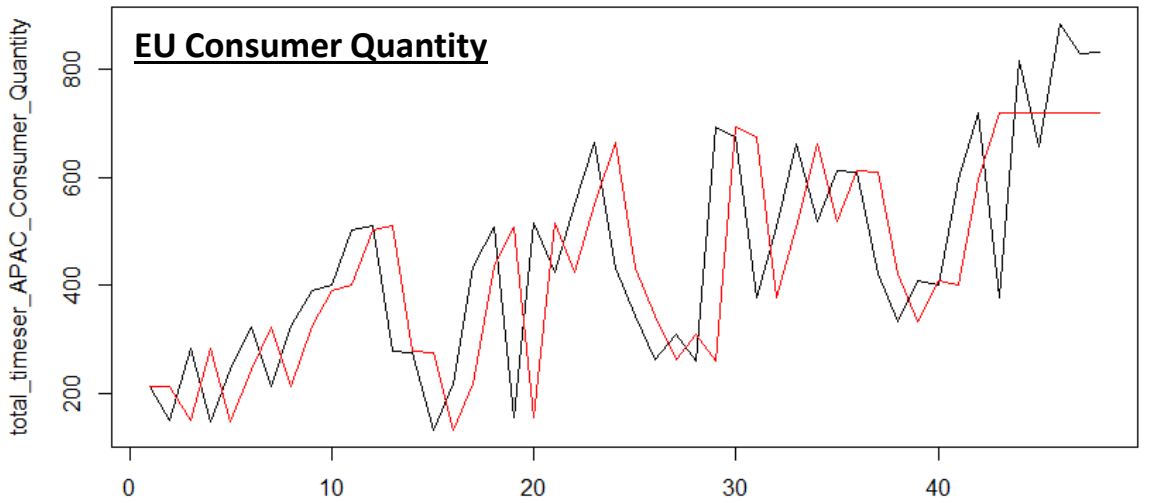
- Test confirms that residual is stationery | Forecast with ARIMA (0,1,0) | MAPE: 26.24%

AUTO ARIMA IS A VERY GOOD MODEL FOR US

Forecast APAC Customer (Sales & Quantity)



- APAC Consumer Sales is likely to plateau in the next 6 months



- APAC Consumer Quantity indicates a steep rise and then averaging number of customer below 700 in coming 6 months.

Summarizing Observations

- APAC Consumer & EU Consumer are the 2 most consistently profitable segments
- EU Consumer Sales indicates a rise in the following 6 months
- EU Consumer Quantity indicates a slow rise and the last 2 months showcases a plateau
- APAC Consumer Sales is likely to plateau in the next 6 months
- APAC Consumer Quantity indicates a steep rise and then averaging number of customer below 700 in coming 6 months.