**Forecasting – Time Series**

**Airline Data**

**Package used in R:**

Readxl : Used to Load , open Excel Data

Forecast: Used to Forecast using ARIMA

**Package used in Python** :

Pandas : Used for data manipulation

Numpy: Used for Scientific calculation

statsmodels.formula.api : Used to Forecast model Based Time Calculation

**Loading the data**

Loading the Airline dataset in R and Python.

**EDA**

No NaN Data Found in the dataset.

Month variable was created and dummy variable was created for every month.

t and t square, Log value for target variable was created for further analysis.

**Data Partitioning**

Airline Data has 8 years of month wise data,7 years data has been taken for training and 1 year data has been taken for testing purpose

**Modeling**

|  |  |
| --- | --- |
| Model | RMSE |
| Linear | 53.199 |
| Exponential | 46.057 |
| Quadratic | 48.051 |
| Additive Seasonality | 132.819 |
| Additive Seasonality with Linear | 35.348 |
| Additive Seasonality with Quadratic | 26.36 |
| Multiplicative Seasonality | 140.063 |
| Multiplicative Seasonality Linear | 10.519 |

Multiplicative Seasonality Linear has the least RMSE value. Hence we can use this method to forecast.

**Coca Cola Data**

**Package used in R:**

Forecast : Used for Forecasting

Fpp : Used for Exponential Smoothing

Smooth : Used for Smoothing and MAPE

Tseries : Used for converting the data to Time Series data

Readxl : Used for Read,Loading the Excel data

**Package used in Python** :

Pandas : Used for Data Manipulation

Numpy : Used For Scientific Calculation

matplotlib.pyplot : Used for Data Visualization

Seaborn : Used for Data Visualization

statsmodels.tsa.holtwinters

SimpleExpSmoothing : Used for Forecast Simple Exponential Smoothing Method

Holt : Used for Forecast Holt Method

ExponentialSmoothing : Used for Forecast with Exponential Smoothing Method

statsmodels.graphics.tsaplots

tsa\_plots : Used to Plot ACF plot

**Loading the data**

Loading the Coca Cola dataset in R and Python.

**EDA**

No NaN Data Found in the dataset.

Converted the sales to Time Series Object

**Data Partitioning**

Coca Cola Data has 42 Quarter, First 38 data has been taken for training and Last 4 Quarter data has been taken for testing purpose.

**Modeling**

|  |  |
| --- | --- |
| Model | Value |
| Simple Exponential Method | 9.765 |
| Holts Method | 9.776 |
| Holts Winter seasonal="add",trend="add" | 3.108 |
| Holts Winter seasonal="add",trend="mul" | 3.161 |
| Holts Winter seasonal="mul",trend="add" | 2.359 |
| Holts Winter seasonal="mul",trend="mul" | 2.183 |

Holts Winter with Multiplicative Seasonality and Multiplicative Trend is the best method to proceed further.