Forecast the CocaCola prices and Airlines Passengers data set. Prepare a document for each model explaining how many dummy variables you have created and RMSE value for each model. Finally which model you will use for Forecasting.

Airline Passengers:

I have created 18 dummy variables, 12 are month, 6 are 't', 't_squared', 'log_passengers', 'Date', 'month', and 'year'

I created model using Model Based Approach using following alogirthms

Linear Model

RMSE = 57.00014788256584

Exponential Model

RMSE = 314.4271317504474

Quadratic

RMSE = 58.49427463947253

Additive Seasonality

RMSE = 132.25413439949477

Additive Quadratic Seasonality

RMSE = 39.75976633806374

Multiplicative Seasonality

RMSE = 314.92426103796834

Multiplicative Additive Seasonality

RMSE = 314.92426103796834

Since, Additive Quadratic Seasonality has lesser RMSE value than others so I will be using Additive Quadratic Seasonality.

The visualizations can be seen in ipynb file.

CocaCola Sales

I have created one extra column i.e. Year which I extracted from Quarter column and I changed the format of Quarter column to Timestamp format, which can be seen in the .ipynb file.

Here, I created model using Data-Driven Models which are given below

Simple Exponential Method

MAPE = 25.002287586785418

Holt Method

MAPE = 22.804732952040638

Holts winter exponential smoothing with additive seasonality and additive trend

MAPE = 10.960443432621751

Holts winter exponential smoothing with multiplicative seasonality and additive trend

MAPE = 7.606623972928795

Since, Holts winter exponential smoothing with multiplicative seasonality and additive trend lesser MAPE value than other models so I used Holts winter exponential smoothing with multiplicative seasonality and additive trend to build final model.

The visualizations can be seen in .ipynb file.

ARIMA model

I have built ARIMA model for the ColaCola Sales data.

which has RMSE value as 1980.88 whose visualizations can be seen in the .ipynb file.