# Forecasting algorithm for CocaCola data

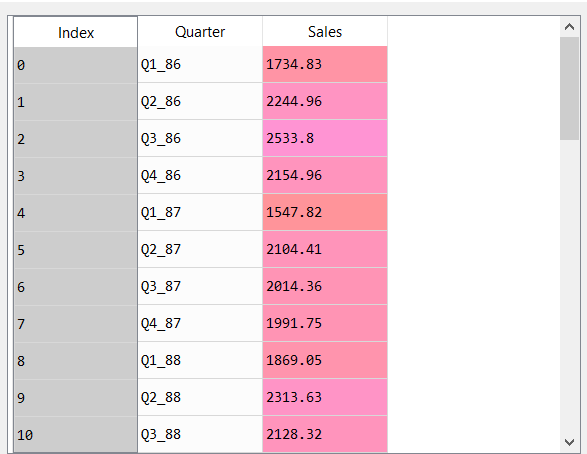
Problem Statement:

Forecast CocaCola Sales 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

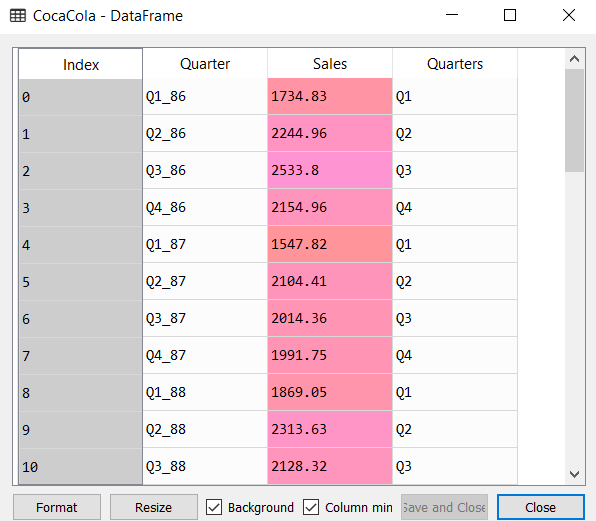
Forecasting.

Considering the CocaCola dataset as follows:



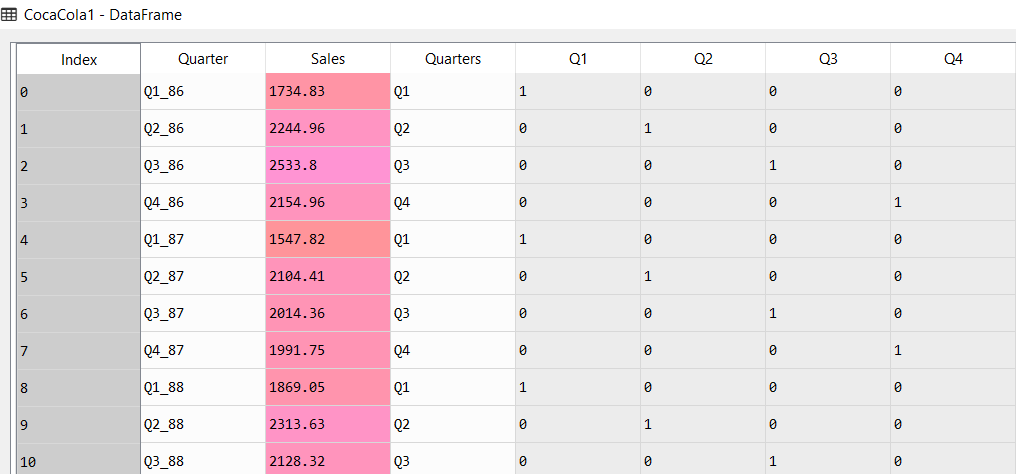
From the above dataset. The date containing the quarter is extracted

We get the following dataset:

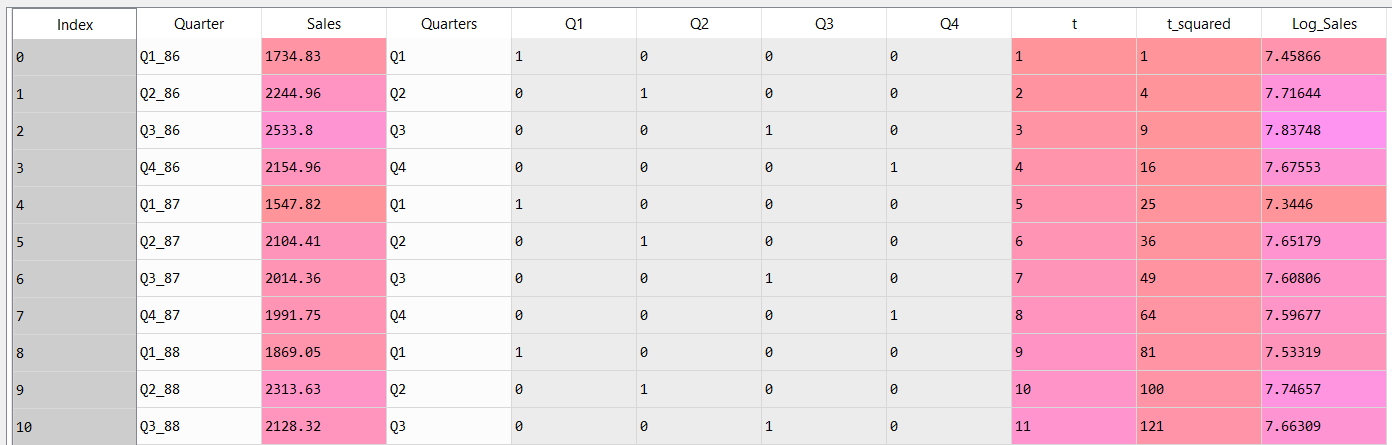


Once the quarters is extracted, we create dummy variables for all the months as shown below:

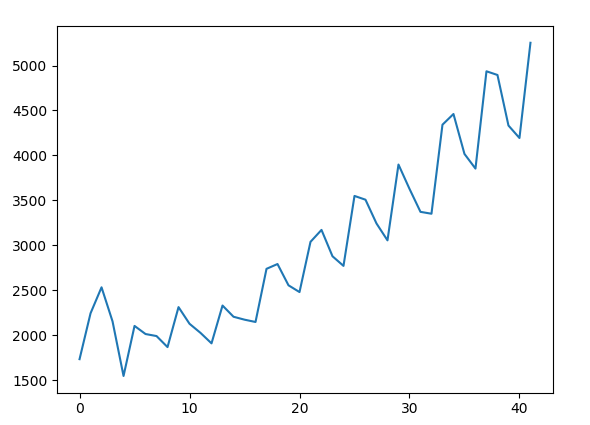
We then proceed to concatenate the above dummy variables with a new dataframe CocaCola1 as follows:



T, t-squared and the log value of passengers variable is calculated.

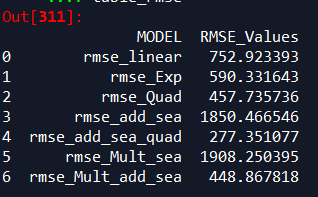


The following is the trend plot of Passengers variable



The above plot shows that the number of Sales are having and upward trend with additive seasonality.

After splitting the data into test and train we execute ordinary least square method. The results of the various models are as follow:



From the above results. The root mean square of additive seasonality quadratic model is having the least RMSE value. Hence we choose this model for prediction.