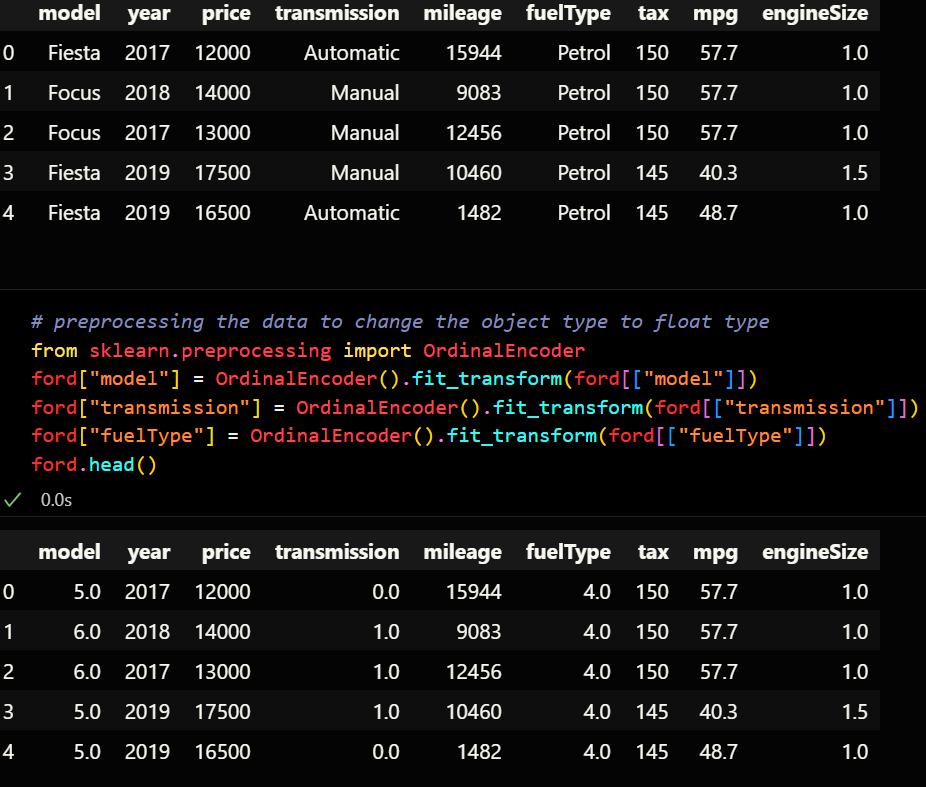
F20DL  
Group 13  
Lab 1

Dataset we explored:

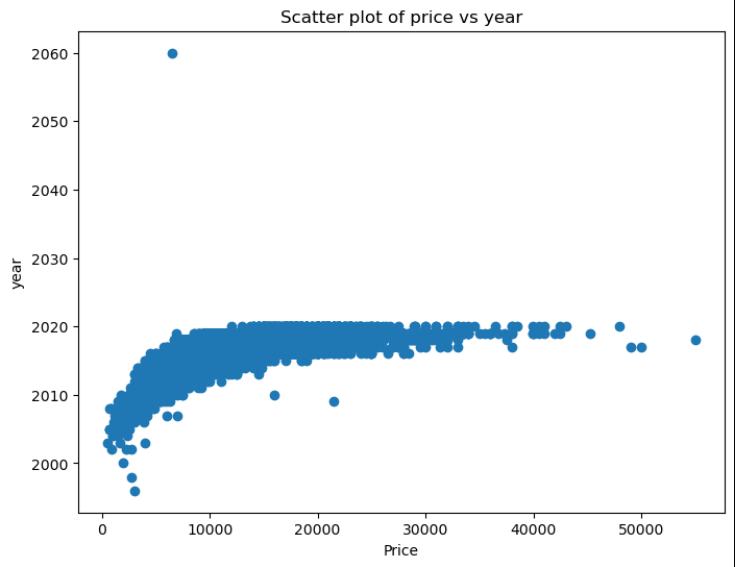
1. Ford Car Price Prediction - <https://www.kaggle.com/datasets/adhurimquku/ford-car-price-prediction>

In this dataset, we have a target attribute price of type float.   
in terms of preprocessing we have to change the attributes “transmission”,”fule type” and ”model”

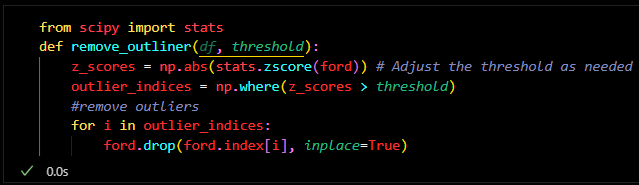


The correlation of our target is given as:  

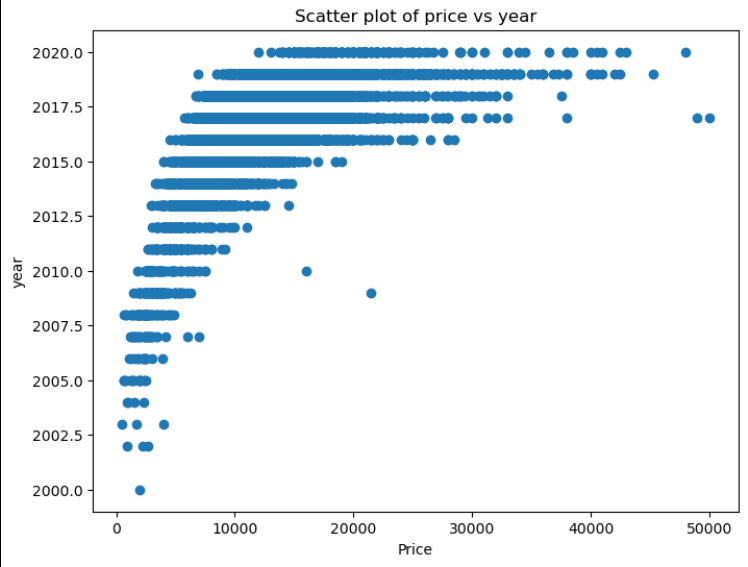

We saw that our dataset was ha outlining data points



So we ran a function to remove outlining data using scipy



The resulting graph is:



We ha to run preprocesssing so that our data only has numerical type data

Furthermore removing existing outliners increases our accuracy in the testing data

By just removing 9 the following data points:

