In this Project I'll try to build a model that can help predict the price of used cars and help buyers know the real value of cars, so prices cannot be manipulated.

I used the dataset from Kaggle.com. It contains information about used Audi cars with many different features such as model, year, price, engine size, etc. This dataset contains more than 10669 records.

Before I start I did some data cleaning and remove the missing data and outliers, I also used Label Encoding Technique to make data analysis easier.

After preparing the data I did explore the data with different ways for example: heatmap-histogram-lineplot and boxplot. Then I split data into two different parts for model creation: 65% Train and 35% Test.

Finally, I build 3 different models: linear Regression- Random forest – CatBoost.

The model which has the higher accuracy was CatBoost with 96%. So, the prediction of this model is the nearest to actual price of cars.