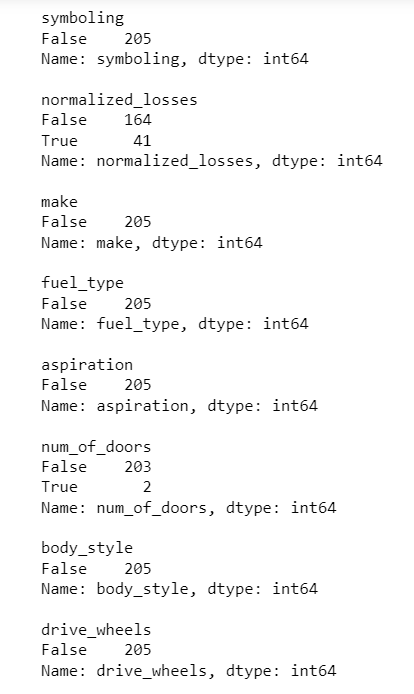
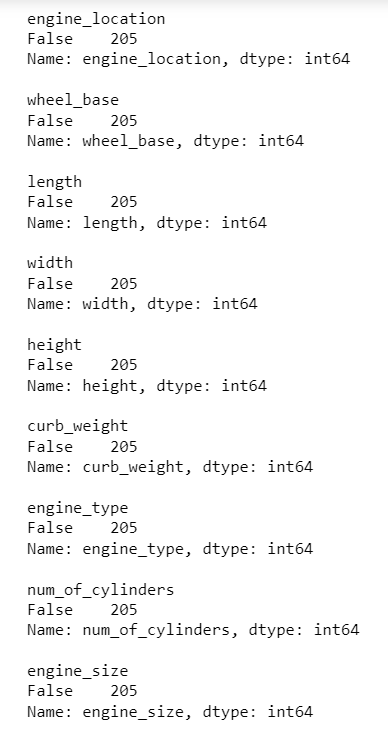
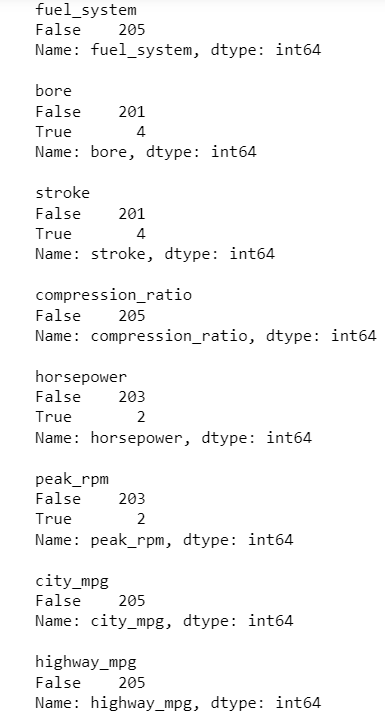
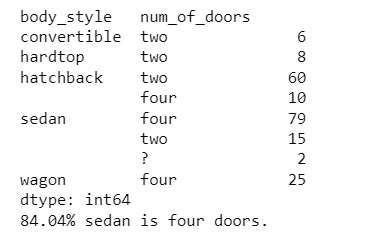
**Data Analysis**

During the study of the given data frame, it was decided to reject the zero values of the price, to replace all numerical values with the average, and to process the string values separately.

*84.04% Sedan is four doors.*

It was reasonable if missed data for the number of doors will be replace by four.



**Correlation**

After the correlation of the data, 21 features remained from the 26 presented, and the 3 most important features were also determined.

* Symboling has correlation 0.036, p-value is 0.6331842154850545
* Normalized losses has correlation 0.317, p-value is 1.4352759788069227e-05
* Make has correlation 0.005, p-value is 0.9465789197062073
* Num of doors has correlation 0.122, p-value is 0.10254970734462912
* Body style has correlation 0.085, p-value is 0.25421650817357294
* Drive wheels has correlation 0.565, p-value is 1.5008084787375338e-16
* Wheel base has correlation 0.612, p-value is 7.210669291835463e-20
* Length has correlation 0.716, p-value is 1.3237649365440924e-29
* Width has correlation 0.745, p-value is 4.483940735629761e-33
* Height has correlation 0.187, p-value is 0.012055937204577285
* Curb weight has correlation 0.838, p-value is 1.3532161301065677e-48
* Engine type has correlation 0.053, p-value is 0.4819564774757745
* Num of cylinders has correlation 0.155, p-value is 0.03756021162069747
* Engine size has correlation 0.725, p-value is 1.1540207363299491e-30
* Fuel system has correlation 0.671, p-value is 7.219421668533878e-25
* Compression ratio has correlation 0.022, p-value is 0.766261136862567
* Horsepower has correlation 0.769, p-value is 2.0073201202419802e-36
* Peak rpm has correlation 0.046, p-value is 0.5410315199111934
* City mpg has correlation 0.714, p-value is 2.591260715545443e-29
* Highway mpg has correlation 0.699, p-value is 1.0327686400384447e-27

*from the following features of the car, the most influential can be singled out:*

Curb weight has correlation 0.838, p-value is 1.3532161301065677e-4

Horsepower has correlation 0.769, p-value is 2.0073201202419802e-36

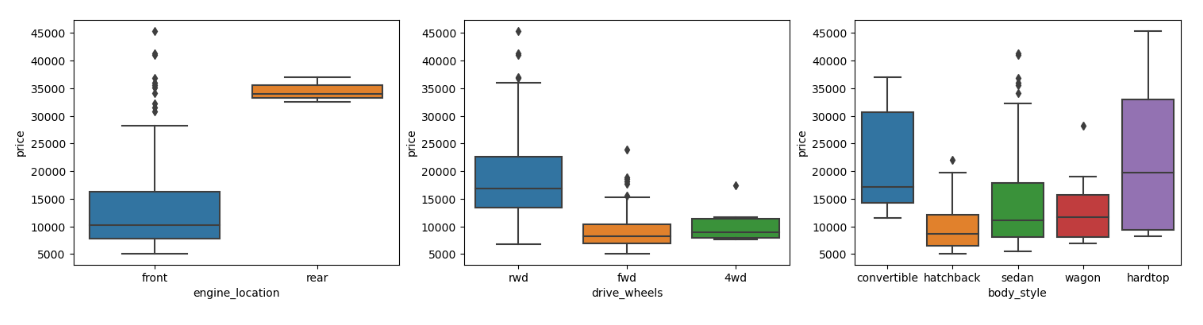
Engine size has correlation 0.725, p-value is 1.1540207363299491e-30

Highway mpg has correlation 0.699, p-value is 1.0327686400384447e-27

Fuel system has correlation 0.671, p-value is 7.219421668533878e-25

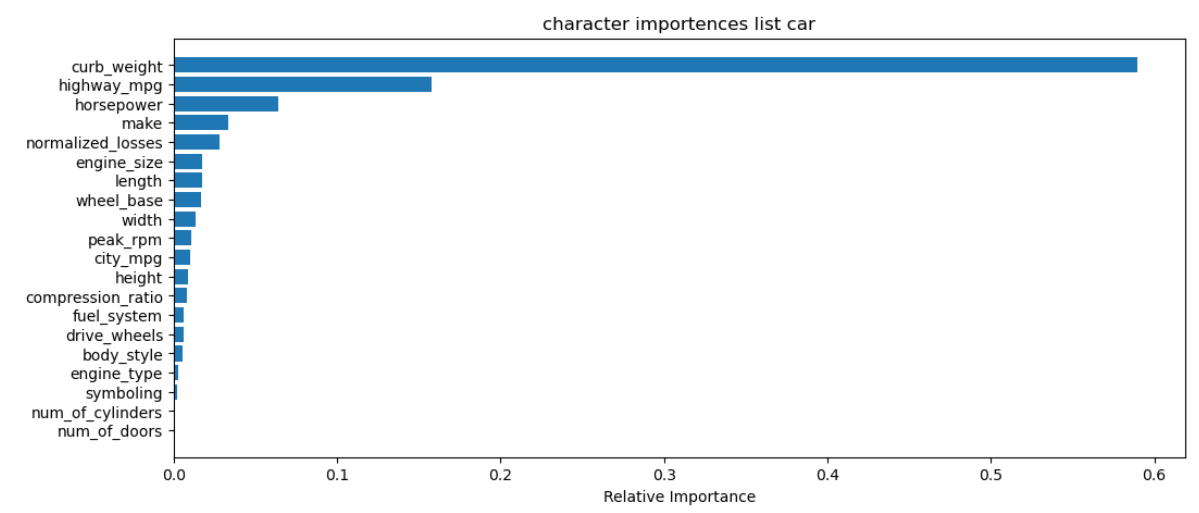
**Visualization**

The fuel consumption ratio in the city and outside the city increases, and the price decreases. The situation is reversed with the size of the engine, as the size increases, the price increases.

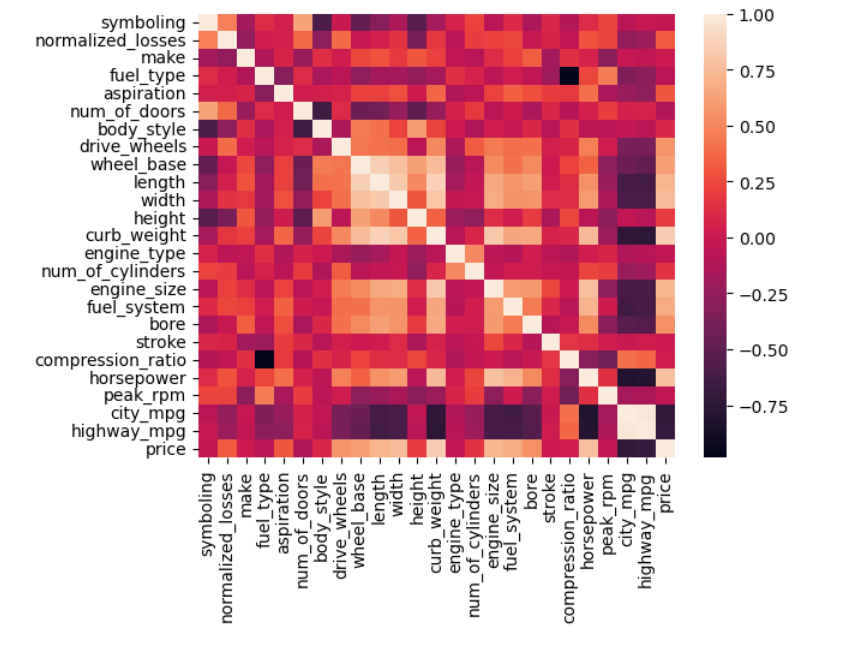


\* According to the chart above, rear-engine cars have an average price of 35,000, front-engined cars range in price from 8,000 to 30,000, and the median is around 10,000. But there are a small number that fall outside their chart . the price varies from 30 thousand to more than 45 thousand.

\* The driving wheels of expensive cars are of the rwd type, the median number is about 17,000 and the larger number is concentrated from 14,000 to about 23,000.



Using powerful tree-based ML models (random forest), calculated importences featyres



Correlation of the data that we marked as important was carried out