A BETTER APPROACH TO CAR PRICE PREDICTION

Juan Herrera

The decrease of car production led used car prices to go up

Project Goal

Create a machine learning model that accurately predicts the price of a car

Data

Used

Save

2020 Subaru Forester Sport

25,531 mi.

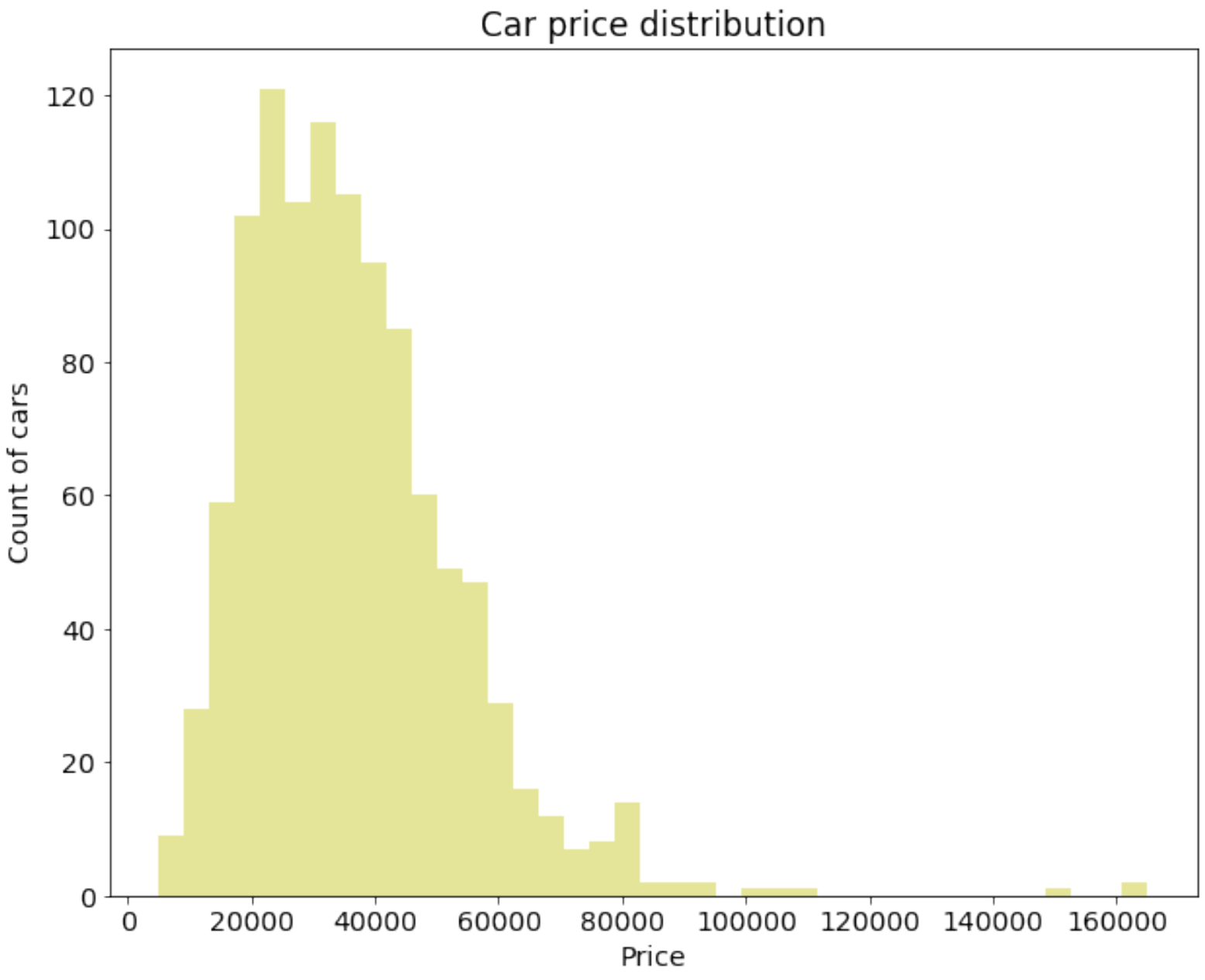
$32,557 $1,400 price drop

Great Deal | $1,308 under | Home Delivery | Virtual Appointments

Basics

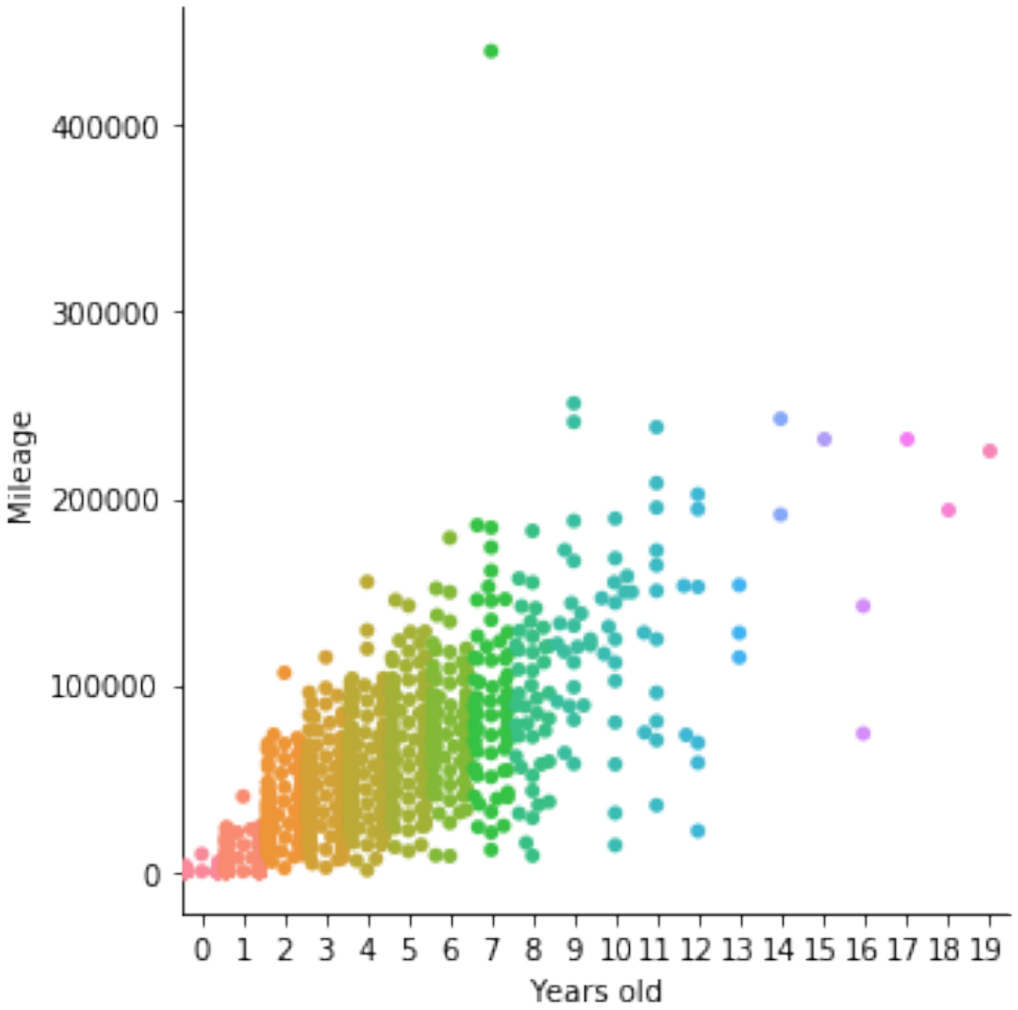
Exterior color	Magnetite Gray Metallic
Interior color	Gray
Drivetrain	All-wheel Drive
MPG	26–33
Fuel type	Gasoline
Transmission	Automatic CVT
Engine	2.5L H4 16V GDI DOHC
VIN	JF2SKARC8LH487218
Stock #	A015174A
Mileage	25,531 mi.
Vehicle history	Free CARFAX 1-Owner Report ↗

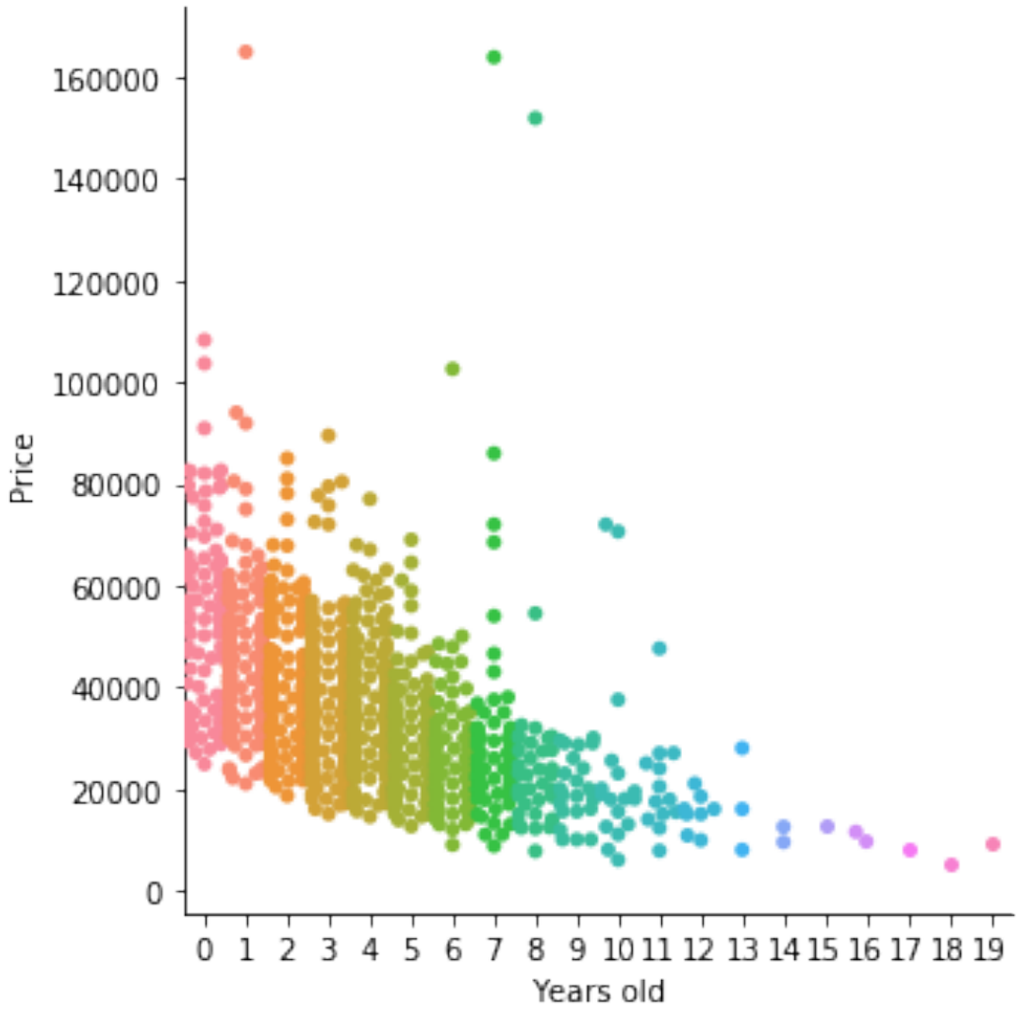
Distribution of Target Variable



- 1183 rows of data
- 10 columns

EDA





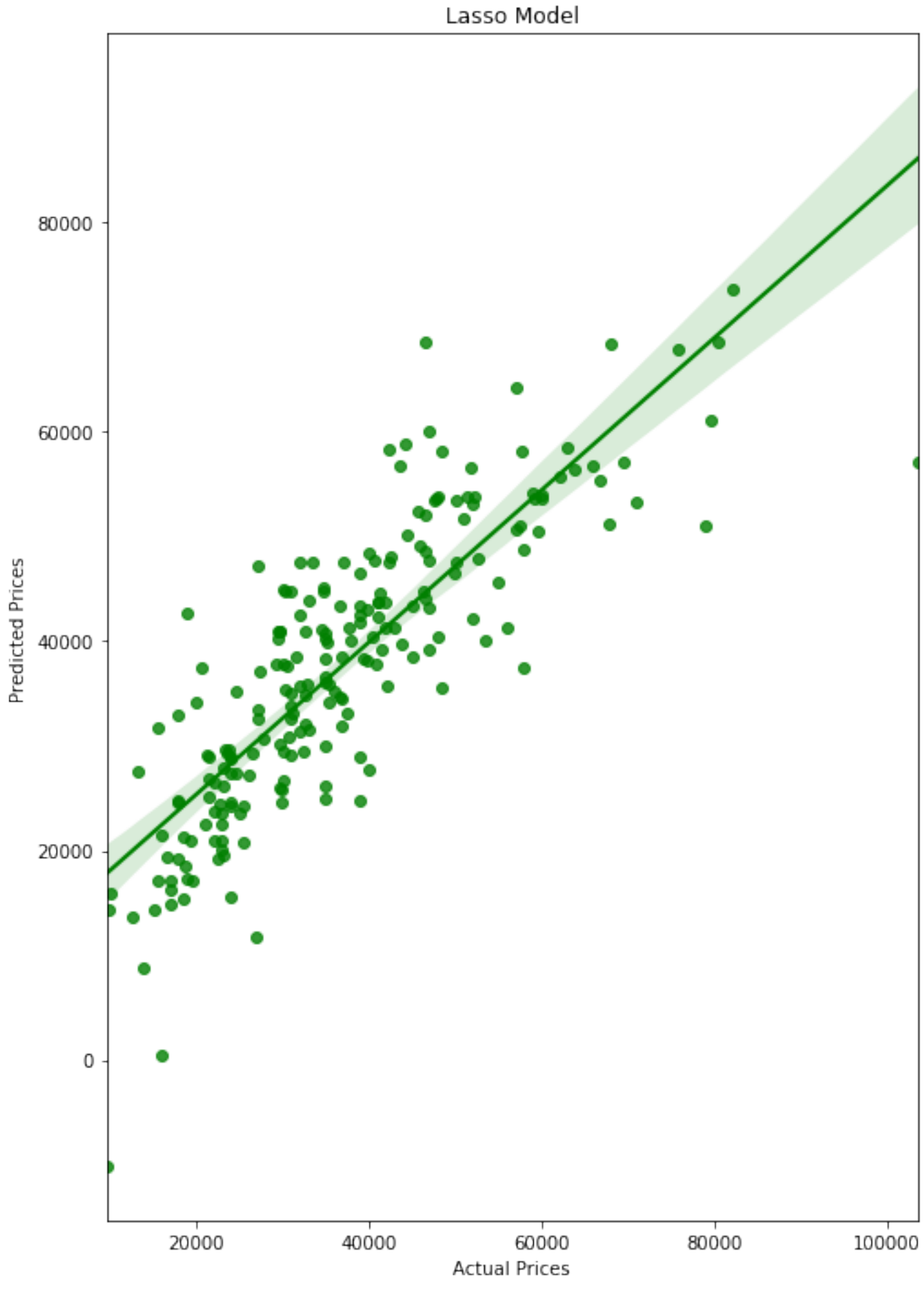
Older cars have more miles and are less expensive

Final Model Performance:

Lasso Regression

R-squared = 69.05

RMSE = $93.10



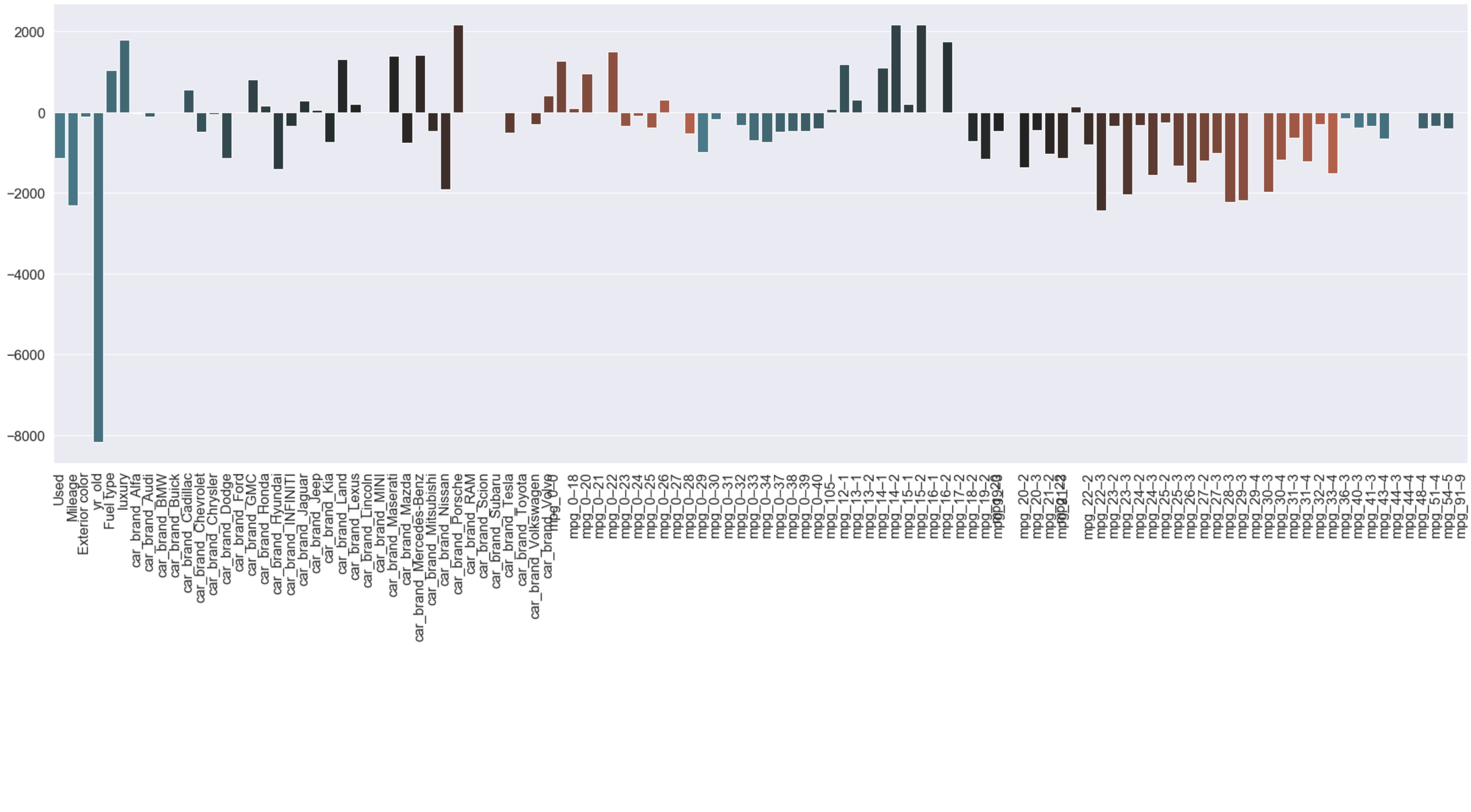
Prediction

Used Land Rover 2015 with 83603 miles

Actual Price: 46557

Predicted Price: 48551.13

Model Coefficients

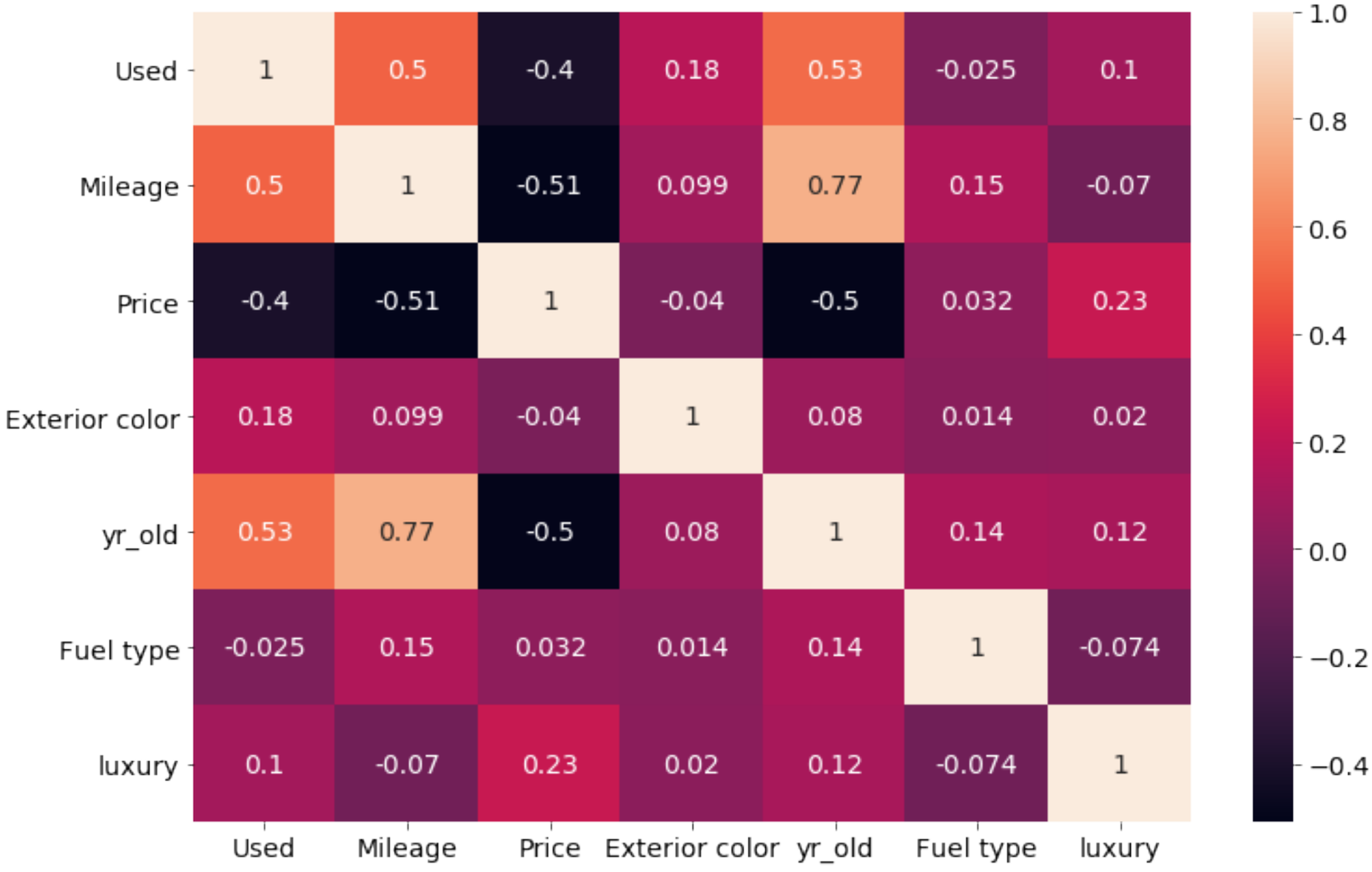


Future Work

1. **Sentiment analysis to understand customer's reviews**
2. **Use seller's info to improve model performance**

Other Charts

Correlation Heatmap



Residuals


60000
40000
20000
0
−20000
−10000
0
10000
20000
30000
40000
Price

Other Models

	Model	Training_score	Testing_score
0	OLS	72.8	72.8
1	Validation	79.5	69.3
2	Ridge	79.1	69.1
3	Lasso	79.1	69.0
4	Random Forest	64.8	64.8