Predicting Resale Price of BMW Used Cars

Md Shahidullah Kawsar
Data Scientist, IDARE, Houston, TX, USA

Motivation for this project:

- Reduce bargaining: Make the buyer/seller's life easier
- Make a good price estimation
- Secured and no privacy interruption

How This Price Compares

Price range for similar **Toyota Camry**
Kelley Blue Book® Private Party Value

$6,395

Low   High

$5,521 – $7,810

About This Vehicle   ⓘ

🕐 Driven 182,400 miles

�️ Automatic transmission

🖌 Exterior: Silver · Interior: Grey

☆ 4/5 NHTSA safety rating ▾

⛽ Fuel type: Gasoline

🔋 22.0 MPG city · 32.0 MPG highway · 26.0 MPG combined

🪑 Seats 5 people

Goal

1. Find the important features to predict the resale price of BMW used cars
2. Design a machine learning model to predict the resale price with the lowest prediction error

BMW used car dataset

model	year	price	transmission	mileage	fuelType	tax	mpg	engineSize
5 Series	2014	11200	Automatic	67068	Diesel	125	57.6	2.0
6 Series	2018	27000	Automatic	14827	Petrol	145	42.8	2.0
5 Series	2016	16000	Automatic	62794	Diesel	160	51.4	3.0
1 Series	2017	12750	Automatic	26676	Diesel	145	72.4	1.5
7 Series	2014	14500	Automatic	39554	Diesel	160	50.4	3.0

- Number of observations (rows) = 10781
- Number of columns = 9 (8 feature variables and 1 target variable)
- Car Price is in $, Road tax is in $, engineSize is in Litre

Features and Target variable

Features	Data type
year	numeric
mileage	numeric
tax	numeric
mpg	numeric
engineSize	numeric

Features	Data type
model	category
transmission	category
fuelType	category

Target	Data type
Price	numeric

Numeric features
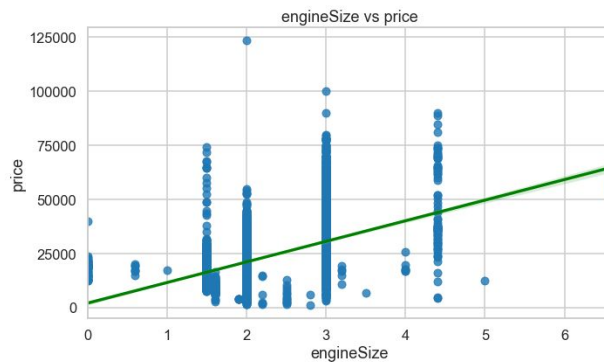
	number of observations	number of missing values	mean	standard deviation	minimum	median	maximum
year	10781	0	2017.08	2.35	1996.0	2017.0	2020.0
price	10781	0	22733.41	11415.53	1200.0	20462.0	123456.0
mileage	10781	0	25496.99	25143.19	1.0	18347.0	214000.0
tax	10781	0	131.7	61.51	0.0	145.0	580.0
mpg	10781	0	56.4	31.34	5.5	53.3	470.8
engineSize	10781	0	2.17	0.55	0.0	2.0	6.6

Categorical features

Percentage of cars in each model

Percentage of cars in each transmission type

Percentage of cars in each fuel type
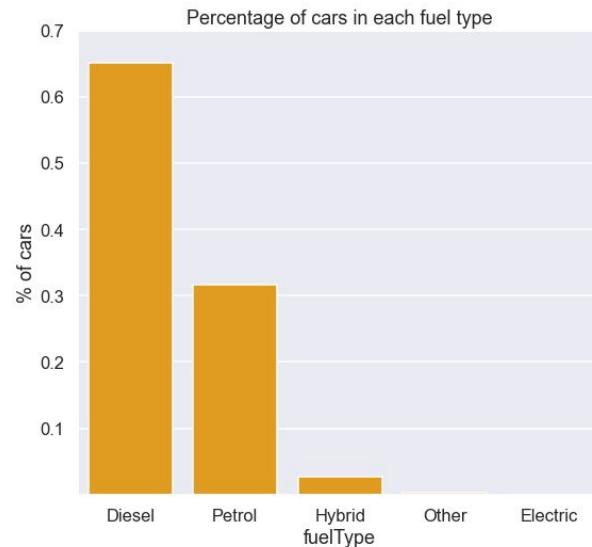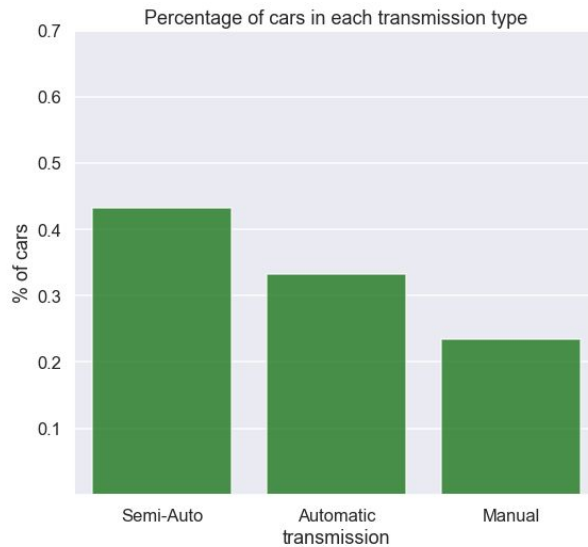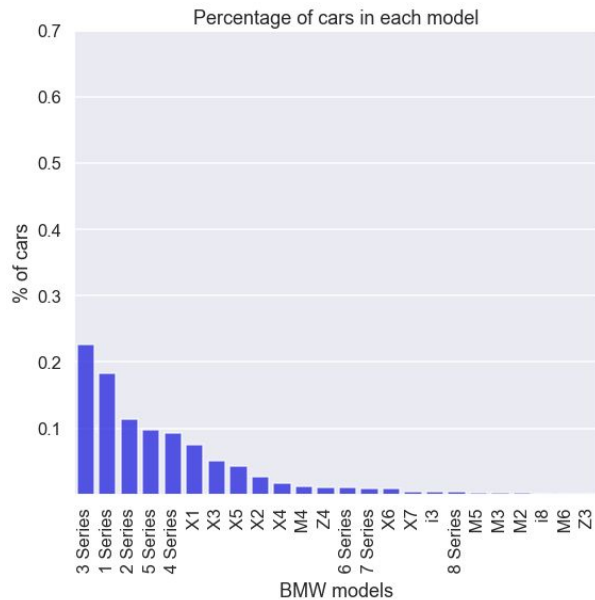
Correlation between the variables
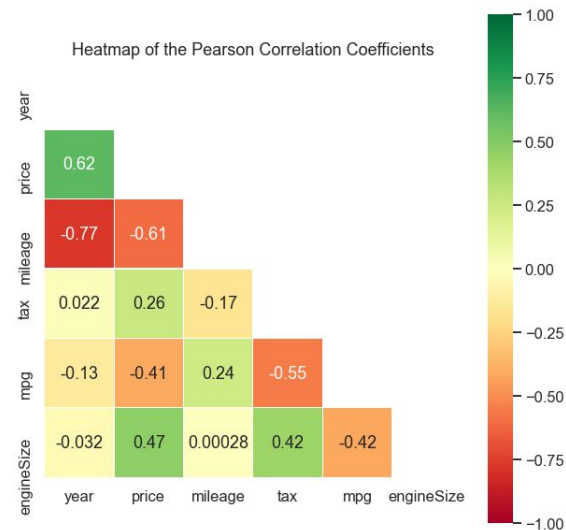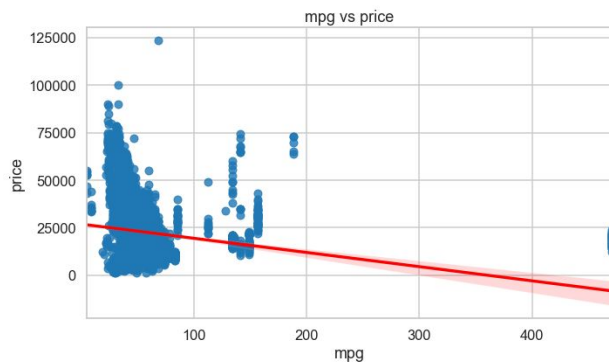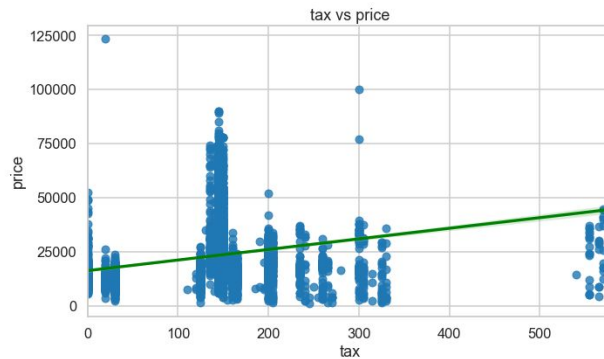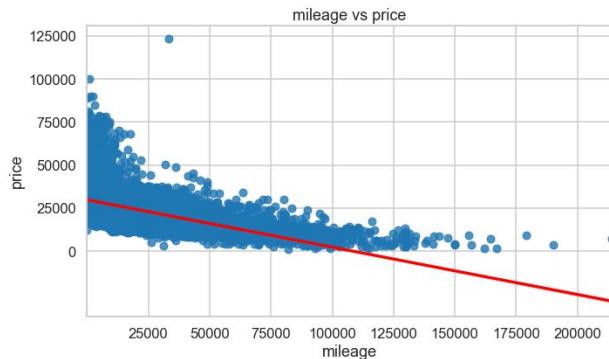
Result and Prediction: resale price of bmw used cars

model	year	transmission	mileage	fuelType	tax	mpg	engineSize	actual price ($)	predicted price ($)	absolute error ($)
1 Series	2013	Manual	89000	Diesel	0	74.3	1.6	6500	6546	46
X1	2013	Semi-Auto	48058	Diesel	160	52.3	2.0	10999	11039	40
1 Series	2019	Automatic	5771	Petrol	145	55.5	1.5	15989	15998	9
X1	2017	Manual	28205	Diesel	150	68.9	2.0	14500	14532	32
1 Series	2013	Manual	60000	Diesel	30	65.7	1.6	7990	8028	38

How good is the prediction?
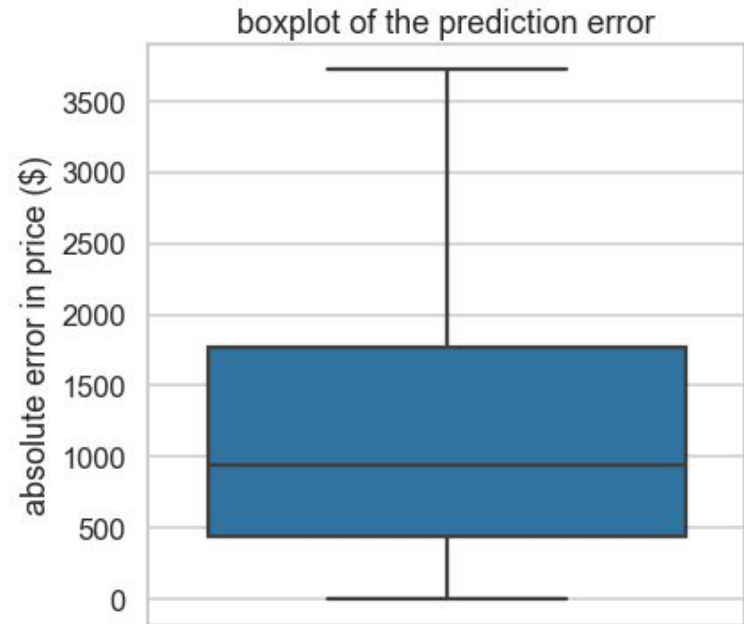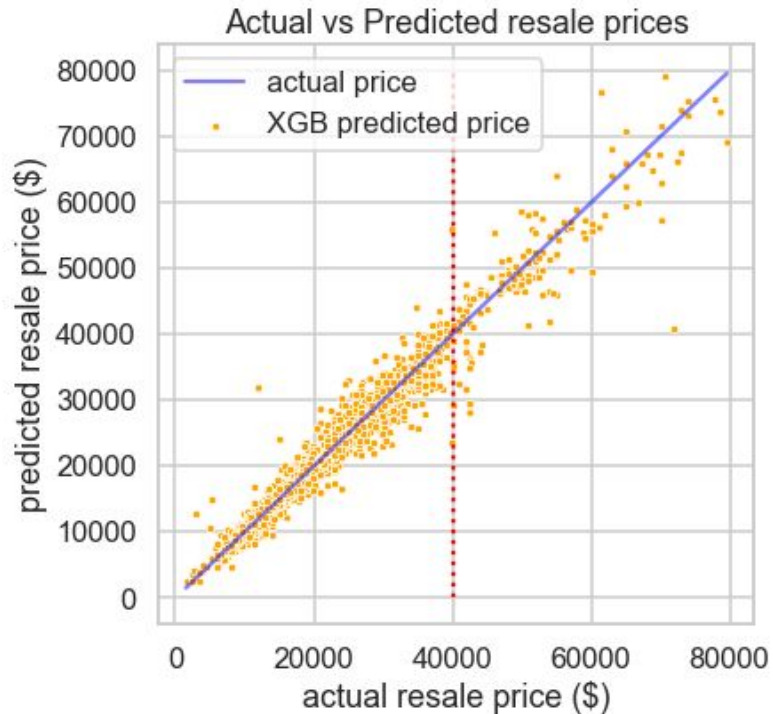
Conclusion and Future Work

- Most important features are engine size, year and transmission system
- Designed solution works really nice

- Improve the data collection
- Train the dataset for individual car model