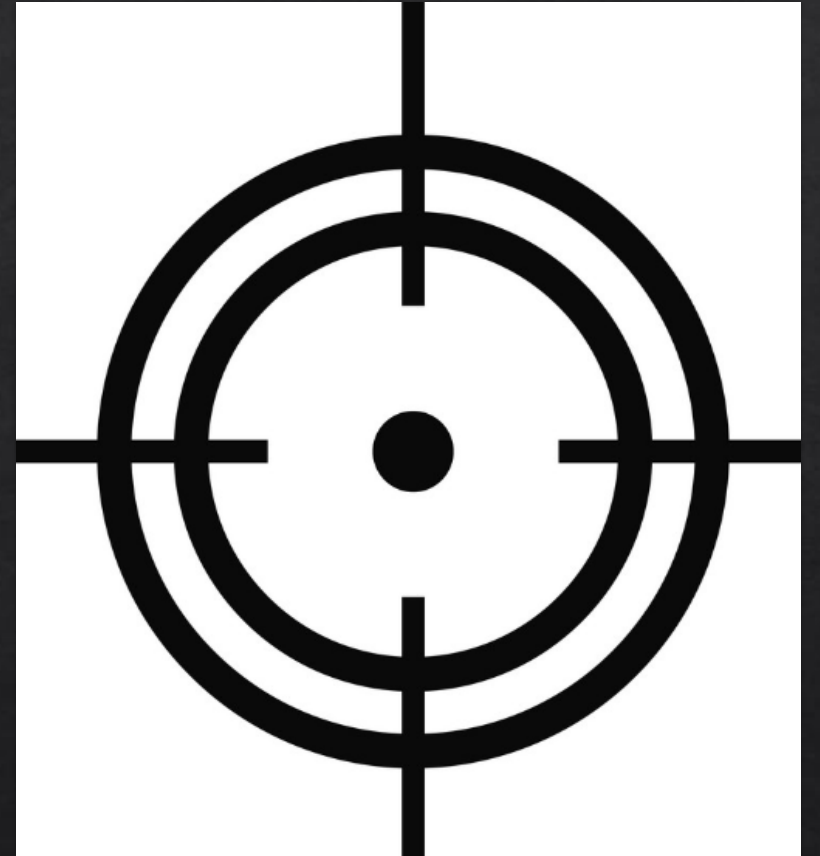


Car Price Prediction

By Abdul Mawar

Aim:

- ◆ Create a statistical model
- ◆ Understand what features influence the price



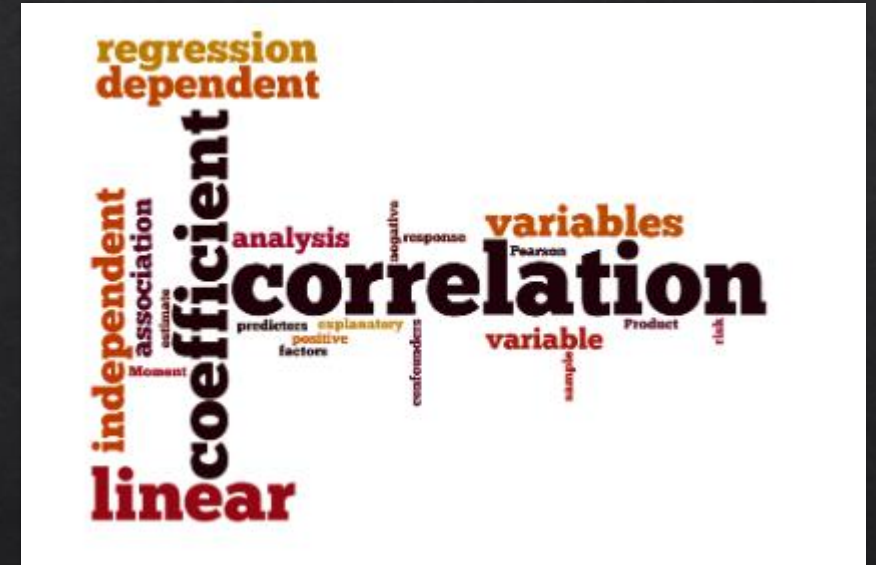
Value:

- ◆ Confidence in the price tag
- ◆ Competitive price advantage



Methodology:

- ◇ Data Science Process:
 - ◇ OSEMN
- ◇ Create a Multiple Linear Regression
 - ◇ OLS (Ordinary Least Squares)



The Data:

- ◆ Dependent Variable:

- ◆ Price

- ◆ Independent Variables

- ◆ Brand

- ◆ Body

- ◆ Mileage

- ◆ Engine Volume

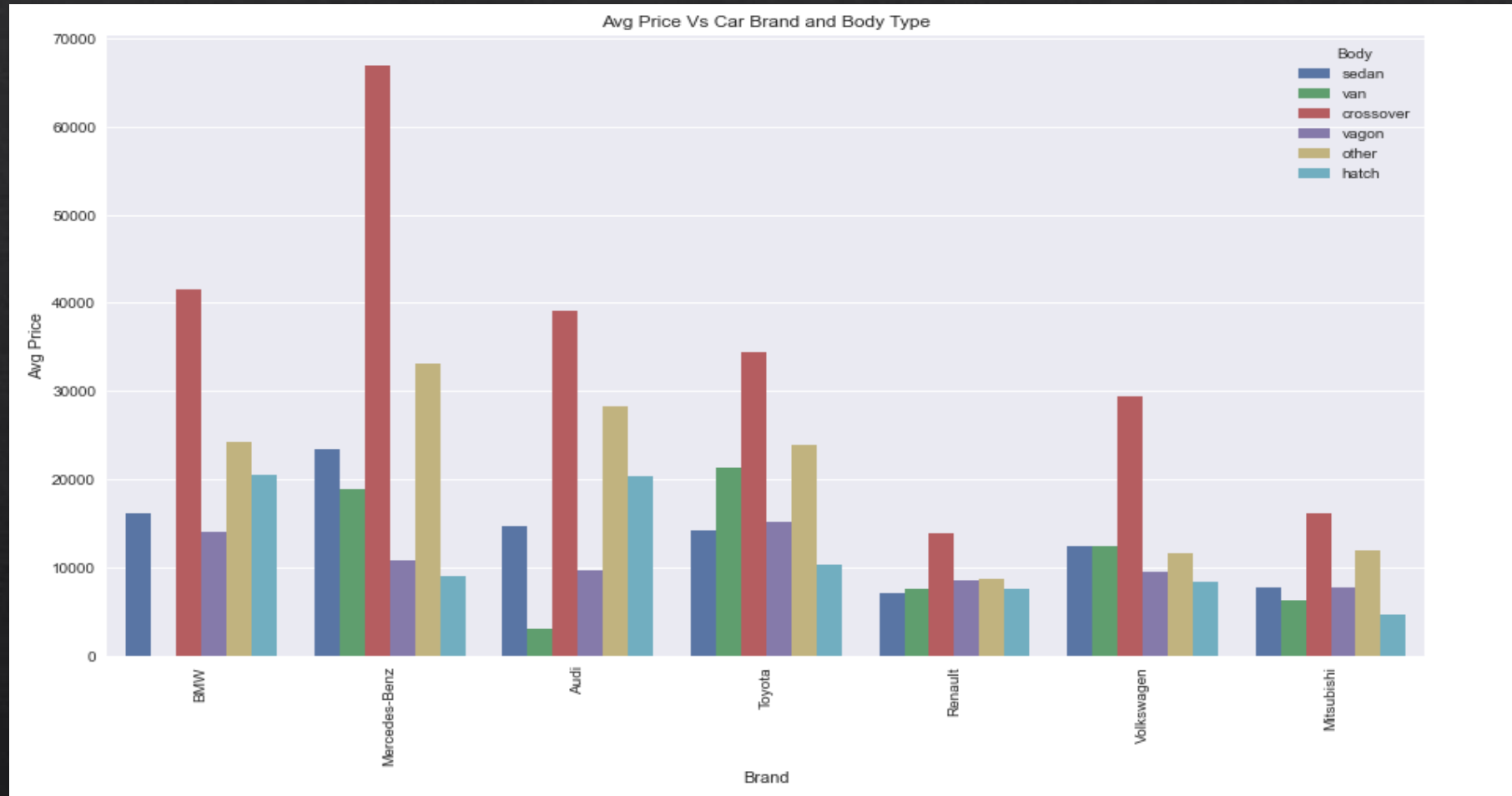
- ◆ Engine Type

- ◆ Registration

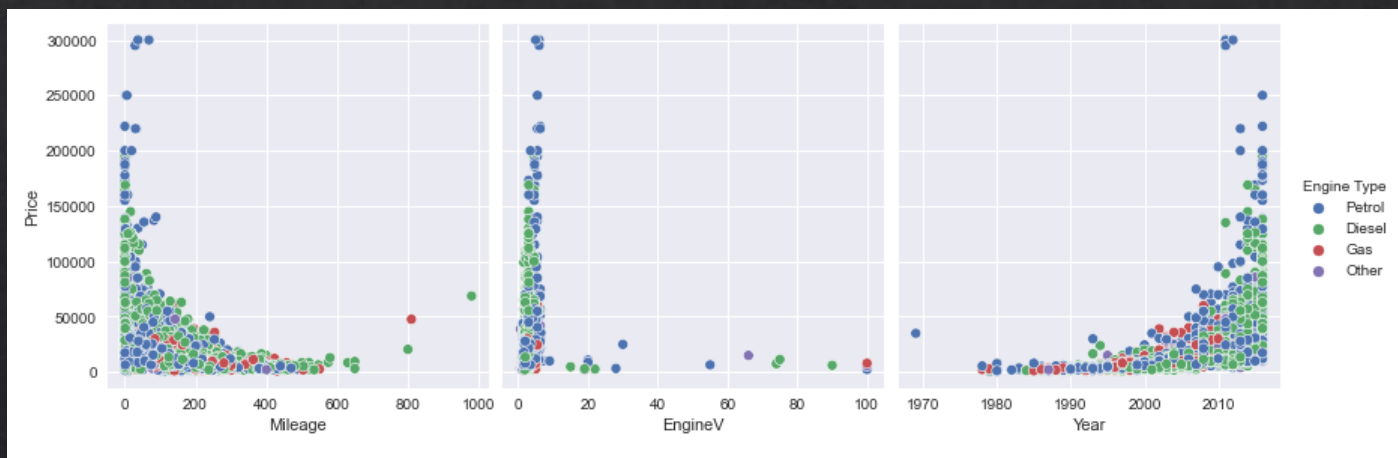
- ◆ Year

	Brand	Price	Body	Mileage	EngineV	Engine Type	Registration	Year
0	BMW	4200.0	sedan	277	2.0	Petrol	yes	1991
1	Mercedes-Benz	7900.0	van	427	2.9	Diesel	yes	1999
2	Mercedes-Benz	13300.0	sedan	358	5.0	Gas	yes	2003
3	Audi	23000.0	crossover	240	4.2	Petrol	yes	2007
4	Toyota	18300.0	crossover	120	2.0	Petrol	yes	2011

Data Exploration:

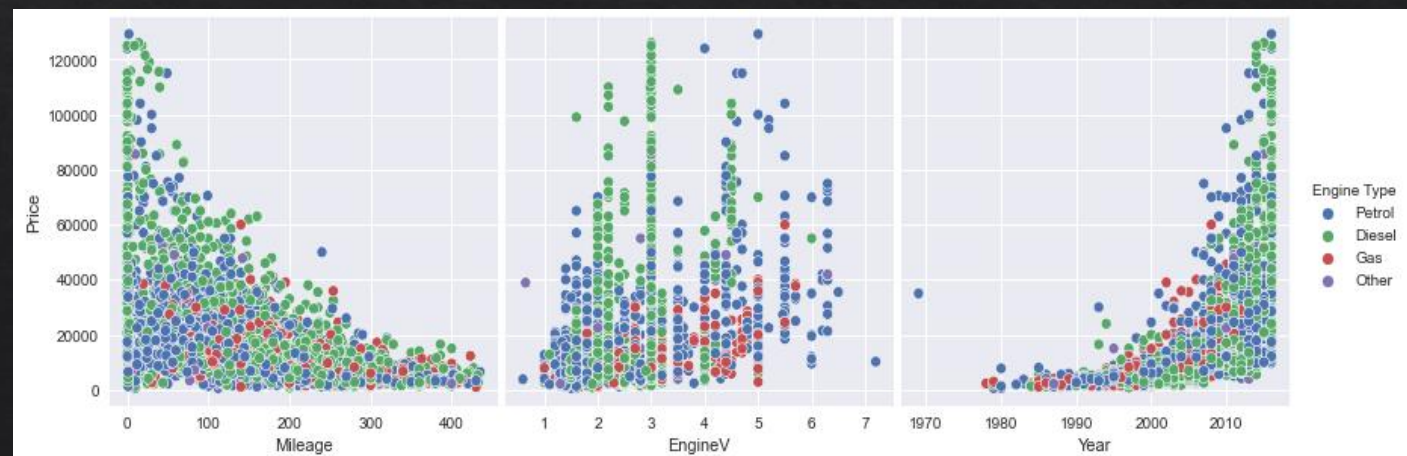


Data Exploration:



Post

Scrub



Outcome:



	coef

Intercept	-148.9026
Mileage	-0.5915
EngineV	0.2364
Year	0.0788
Brd_BMW	0.0744
Brd_Mercedes_Benz	0.0334
Brd_Mitsubishi	-0.4759
Brd_Renault	-0.5675
Brd_Toyota	-0.2169
Brd_Volkswagen	-0.2148
Bdy_hatch	-0.3004
Bdy_other	-0.2413
Bdy_sedan	-0.2752
Bdy_vagon	-0.2997
Bdy_van	-0.4086
EngT_Gas	-0.1148
EngT_Other	-0.0761
EngT_Petrol	-0.1005
Rego_yes	0.8989

Recommendation:

- ◆ Further analysis on the correlation of the independent variables
 - ◆ Brand and Year – Are the brands of cars evenly distributed by their age
 - ◆ Brand and Body – Is there a significant correlation, should Body be removed in the model

Thank you

Abdul Mawar

abdulmawar@gmail.com

<https://github.com/amawar/Capstone-Project.git>