Used Car Price Predictions
on eBay

Team 6

Our Team

Chris Chang
U98764222
cchang72@bu.edu

XX DRIVER LICENSE XX

Ting-Hsuan Chen
U51060396
aliciac@bu.edu

XX DRIVER LICENSE XX

Haiyuan Zhang
U68804194
haz155@bu.edu

XX DRIVER LICENSE XX

Mekhal Raj R
U31160630
mekhal@bu.edu

XX DRIVER LICENSE XX

Yifang (Harry) He
U94086175
yifanghe@bu.edu

XX DRIVER LICENSE XX

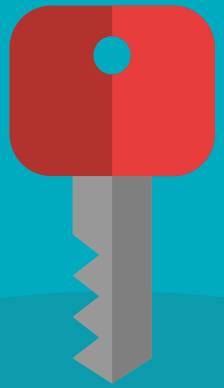
Mohammed Al Qenae
U04709532
alqenae@bu.edu

XX DRIVER LICENSE XX

Table of Contents

01

Intro to Problem



Objective / Bigger Picture

- To predict listing price of used cars on eBay based on predictors such as vehicle type, gearbox, fuel type, and kilometers on car.

- To provide insight/guidance for future buyers and sellers on the prices they should list or bid on eBay for used cars.

Our Dataset

Data Source

https://towardsdatascience.com/predicting-the-price-of-used-cars-891d13faf3fc

Columns

19

Rows

50000

Variable Types

Character, Integer, Other

The features of our dataset

Price	Vehicle Type	Year of Registration	Gearbox	Power PS
Model	Kilometer	Fuel Type	Brand	Not Repaired Damage

Data Cleaning

1st

Dropped
Date Crawled, Seller, Offer Type, abtest, Month of Registration, Date Created, Postal Code, Last Seen, Name

2nd

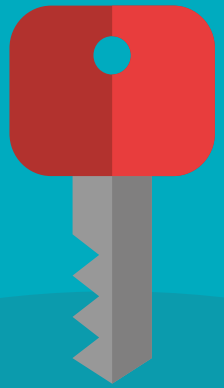
Removed NAs, zeroes, and blank spaces

3rd

Filtered year of registration and price outliers

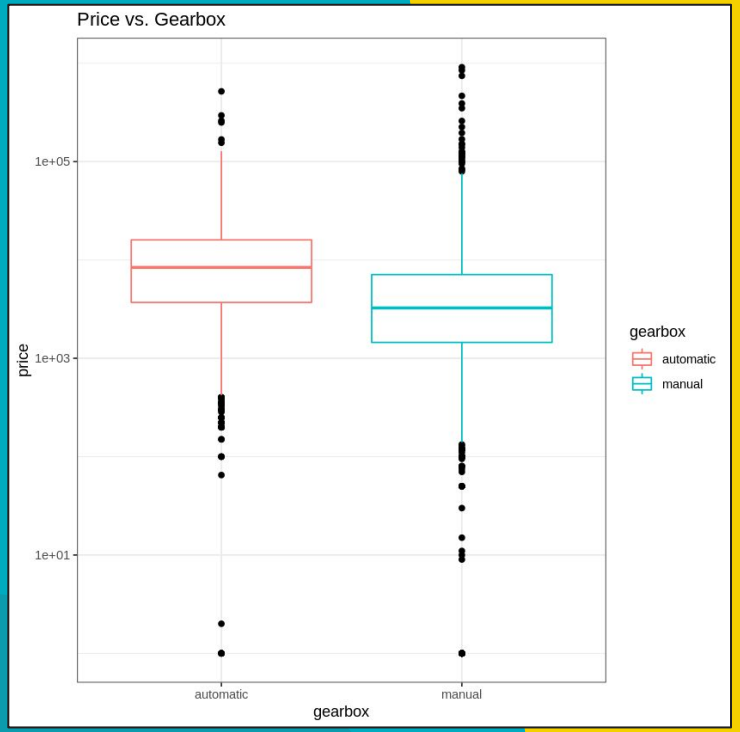
02

Exploratory Data Analysis

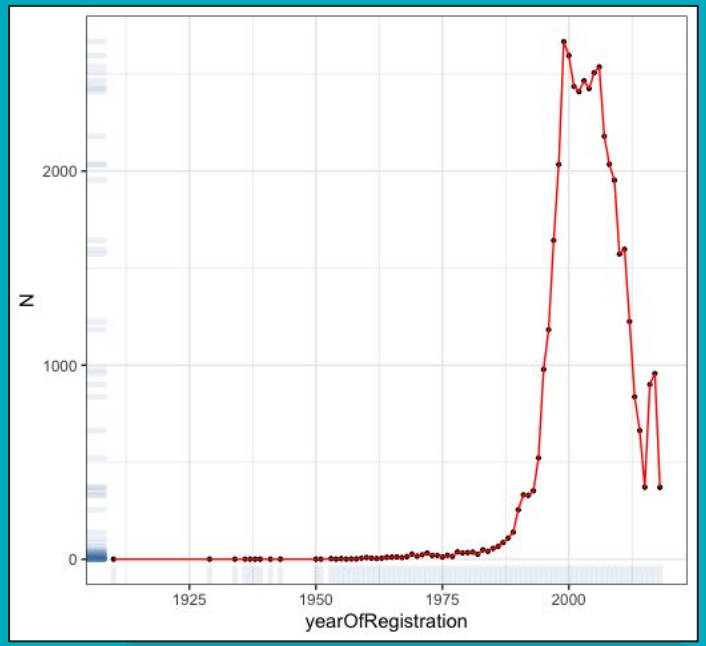


Gearbox

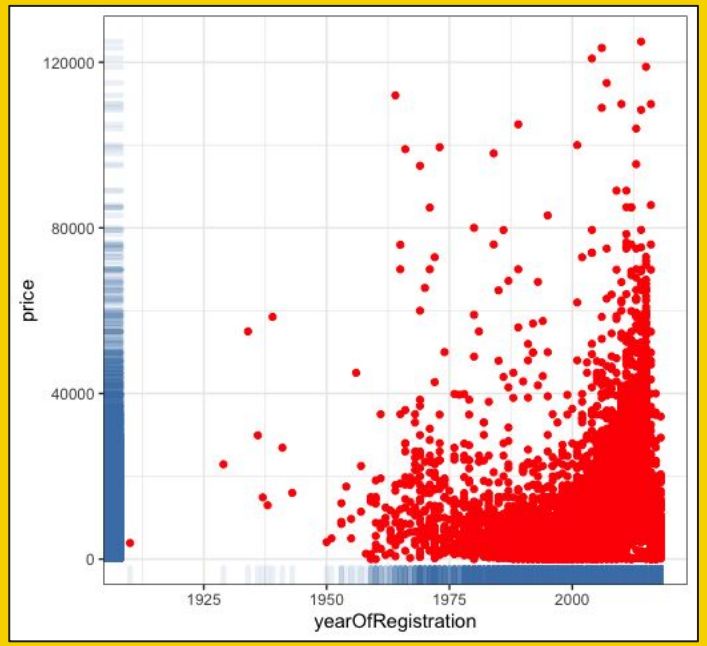
Price vs. Gearbox



Year of Registration



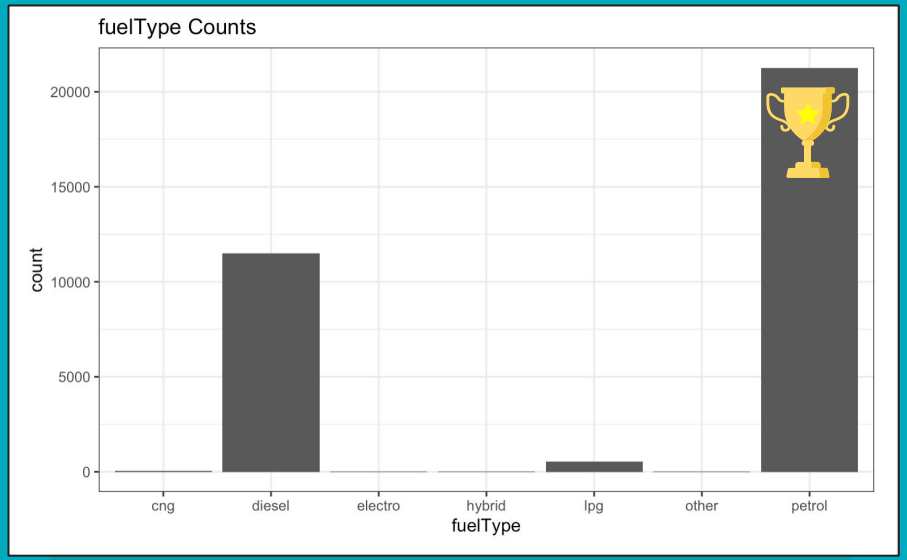
Counts of cars registered for each year



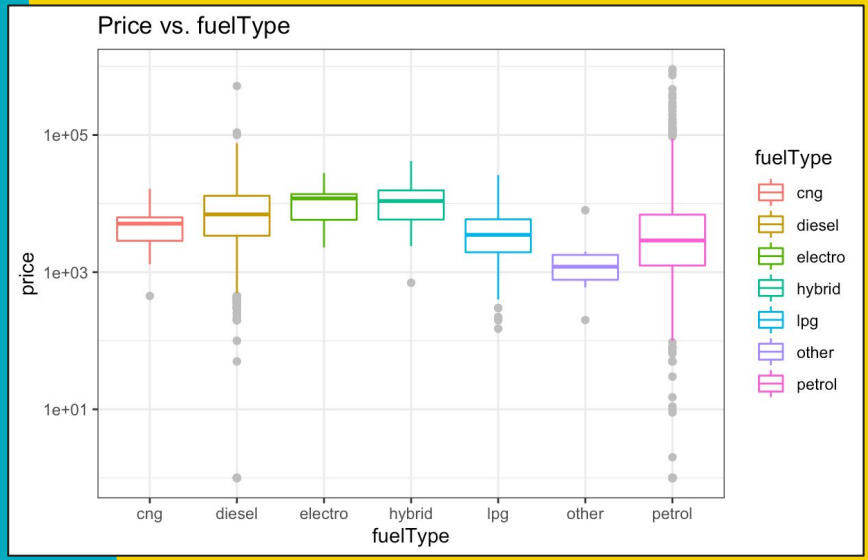
Relationship between price and registration year

Fuel Type

Count of Fuel Type

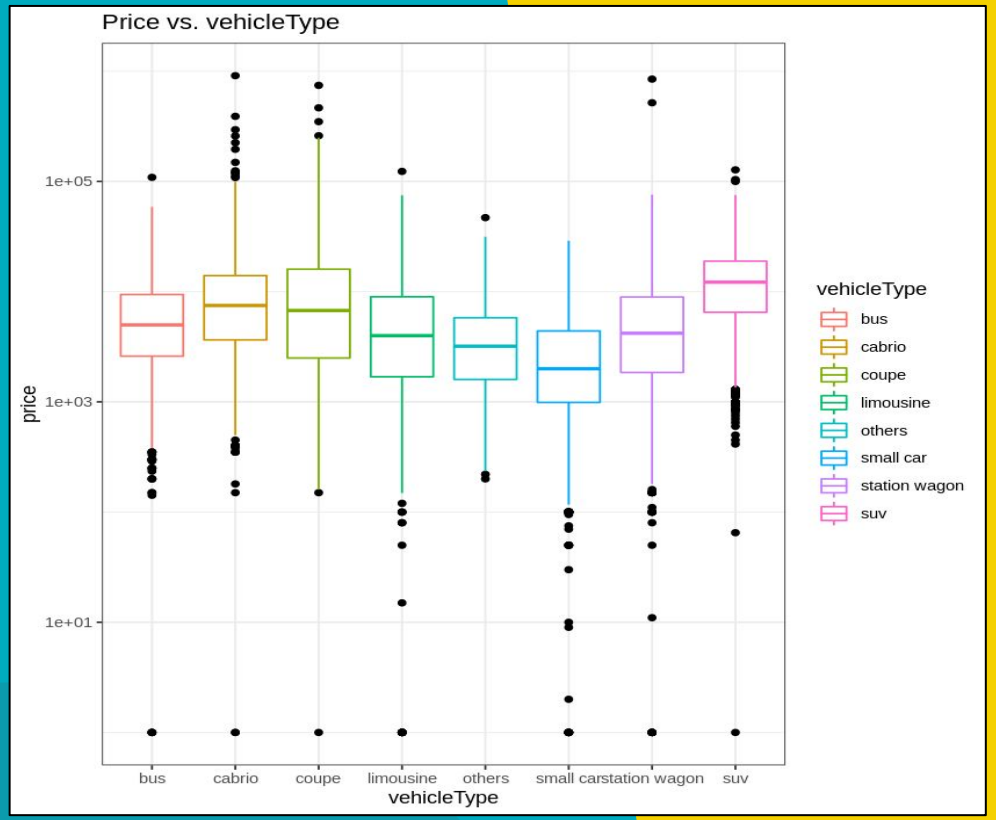


Price vs. Fuel Type



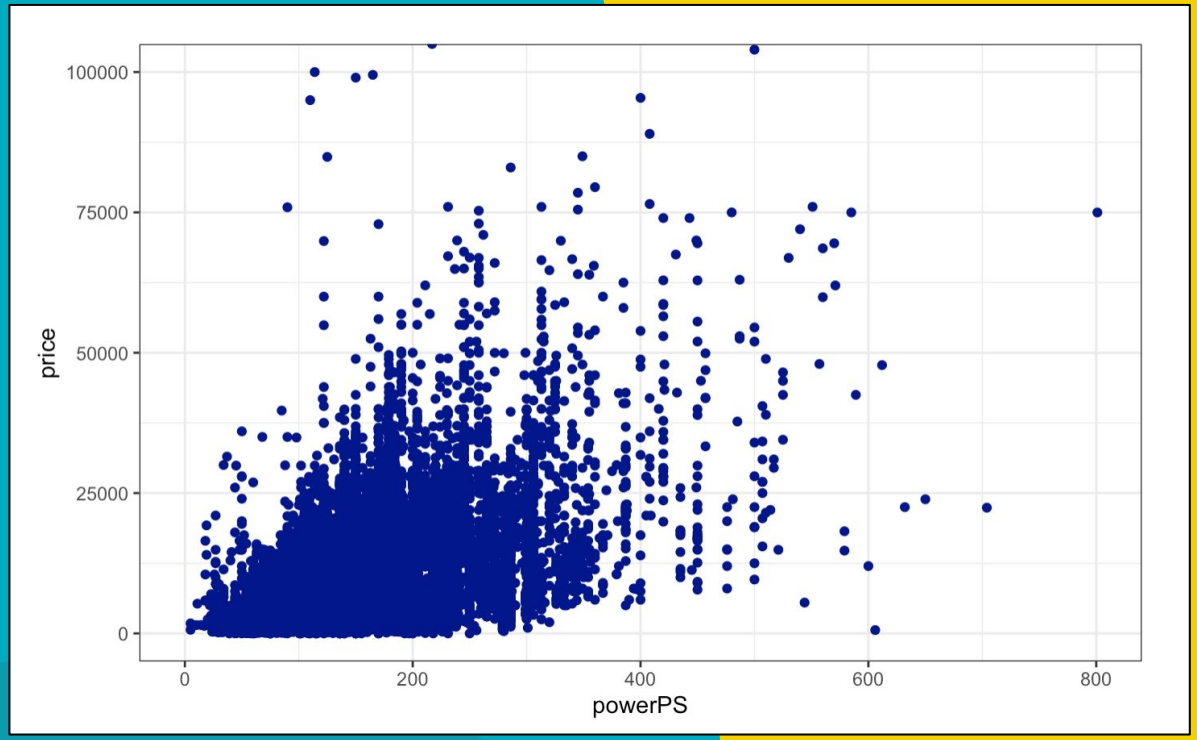
Vehicle Type

Vehicle Type vs. Price



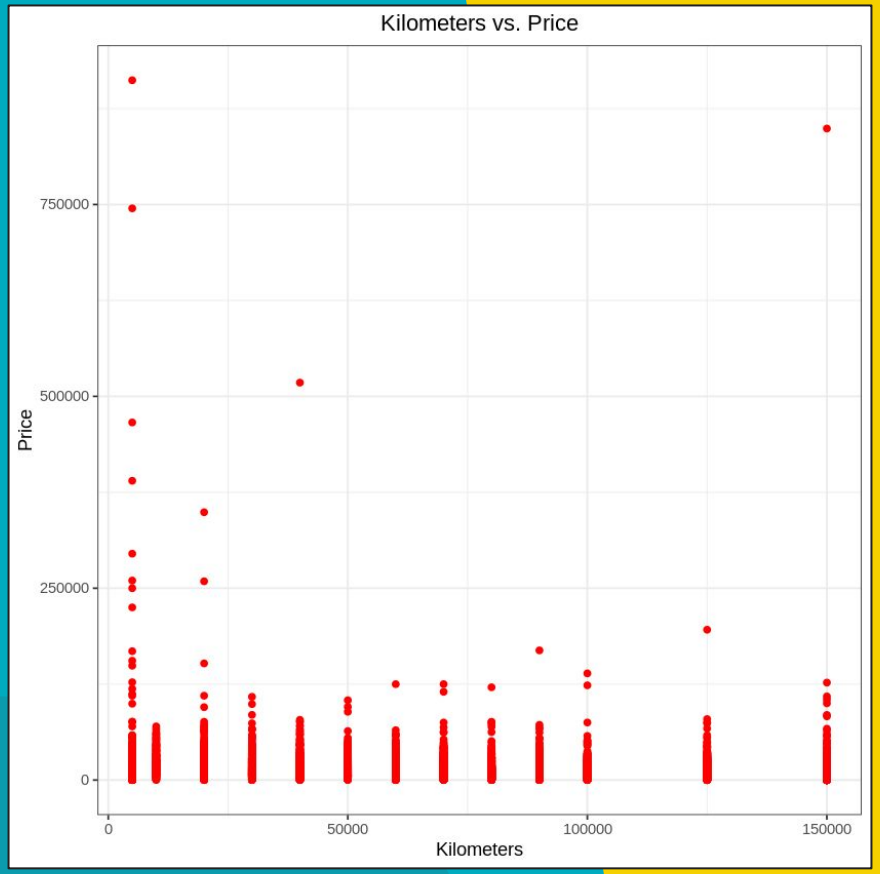
PowerPS

PowerPS vs. Price



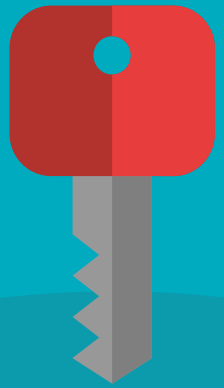
Kilometers

KM vs. Price



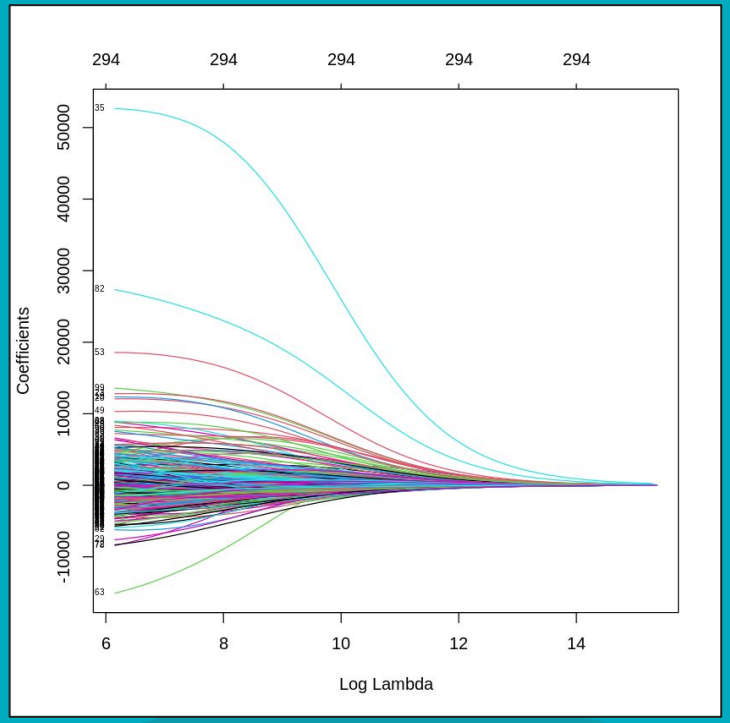
03

Machine Learning Models

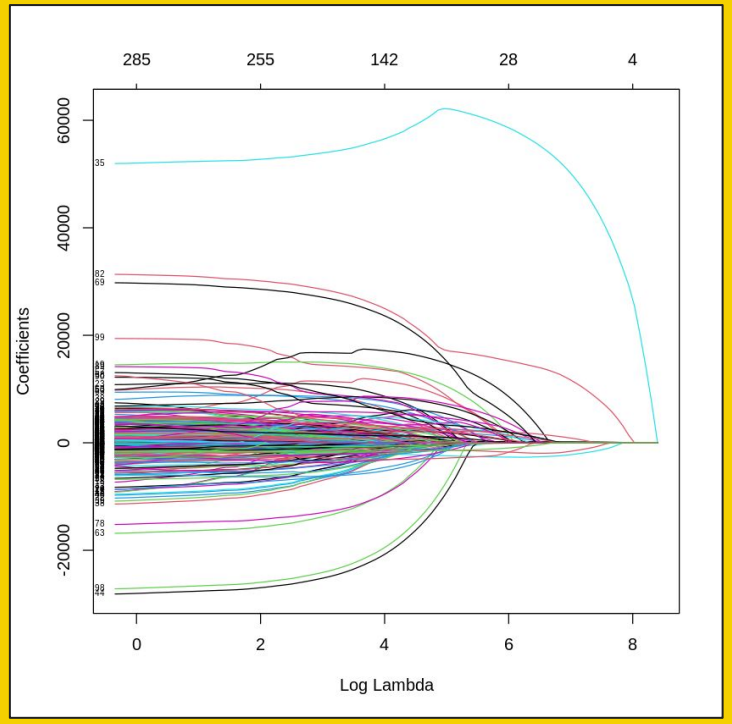


Ridge and Lasso

Test RMSE for OLS : 11938.486

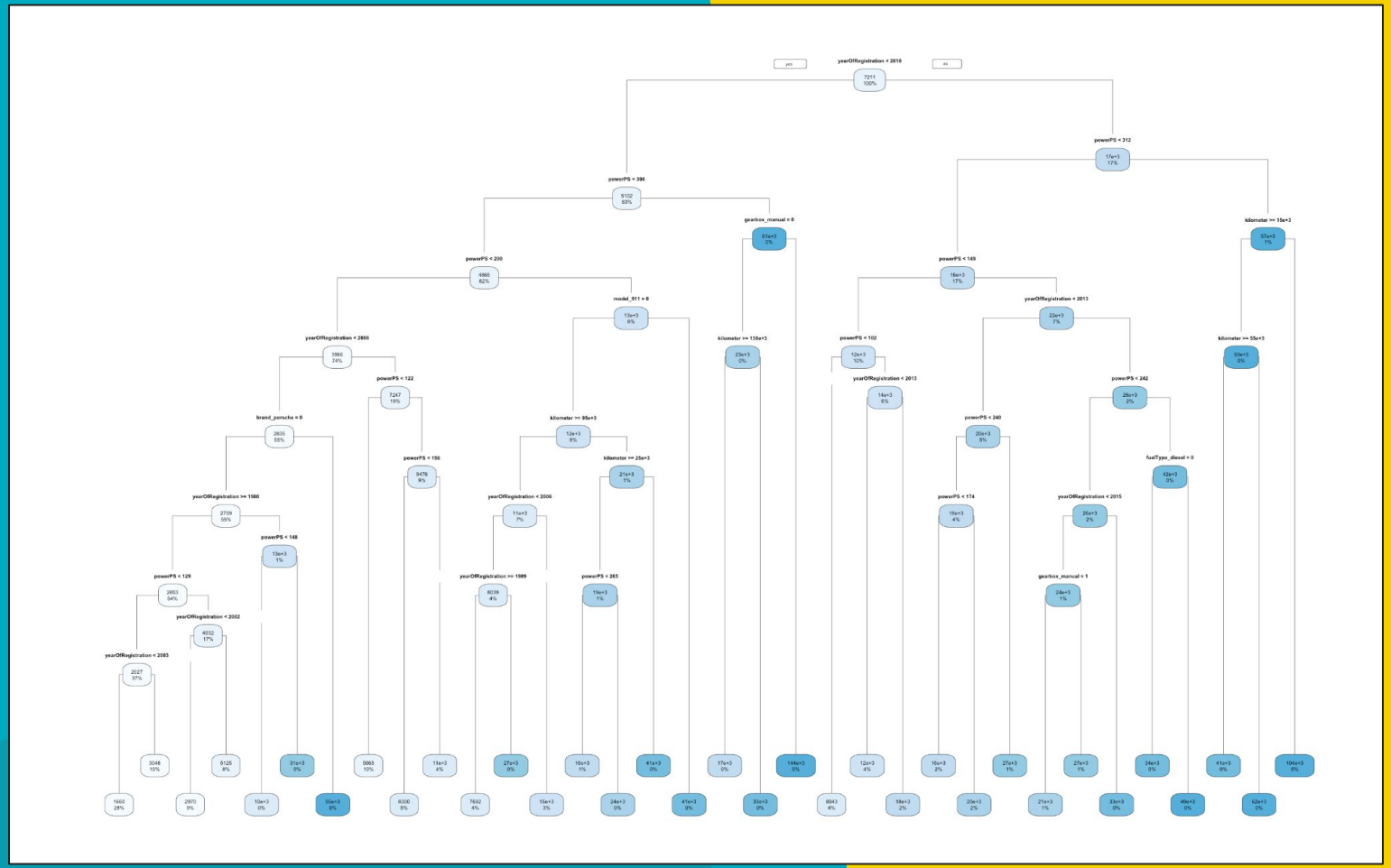


Test RMSE for Ridge: 7155.59

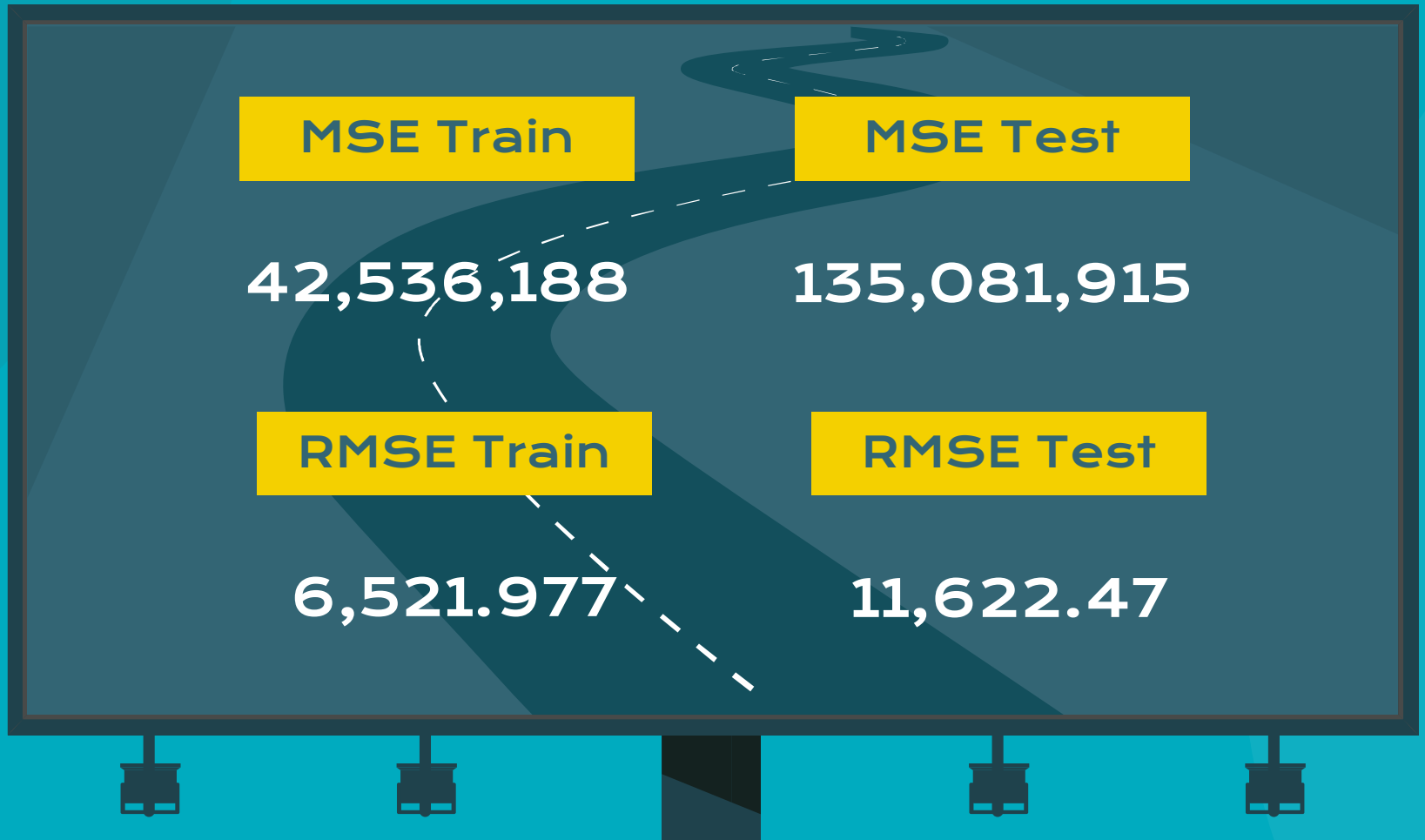


Test RMSE for Lasso: 7195.74

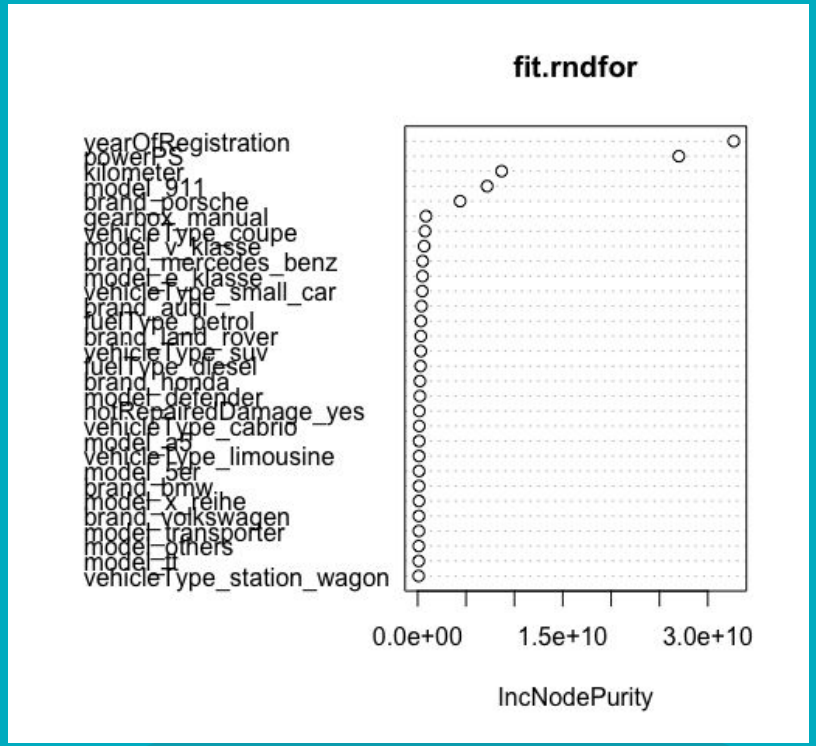
Regression Tree



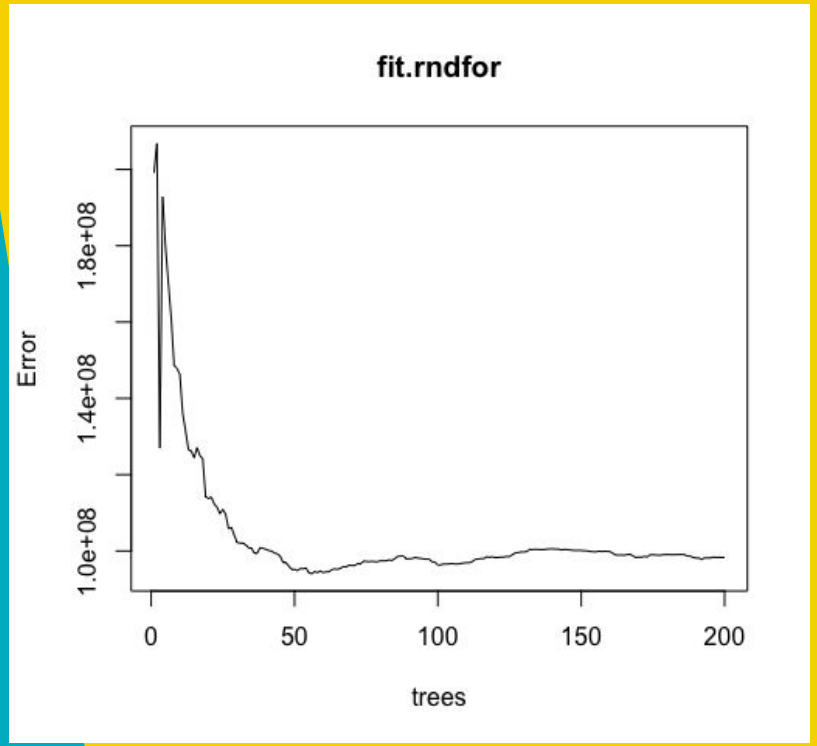
Regression Tree



Random Forest



Predictor Importance



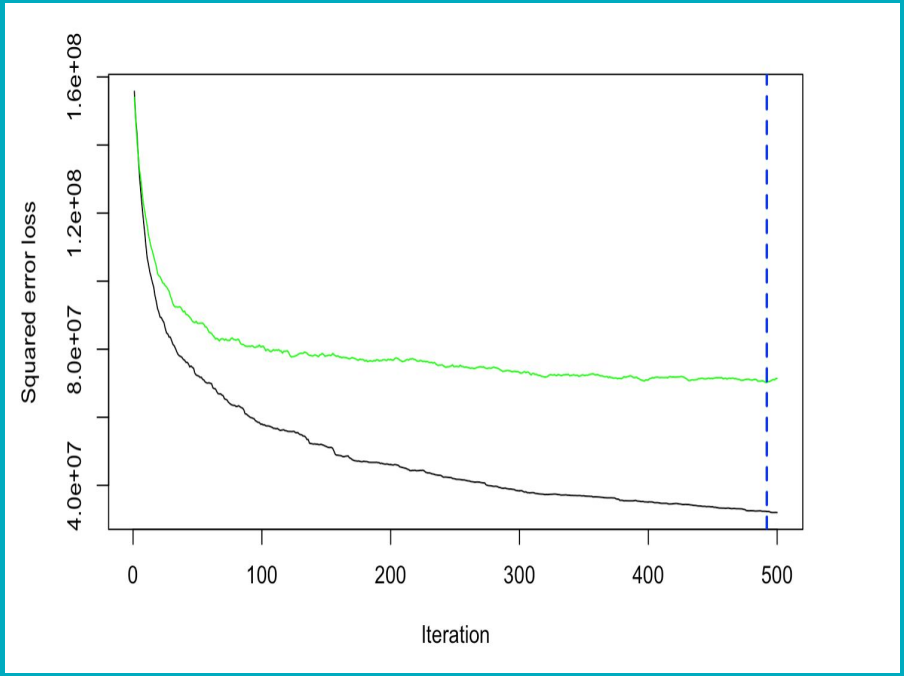
OOB Error

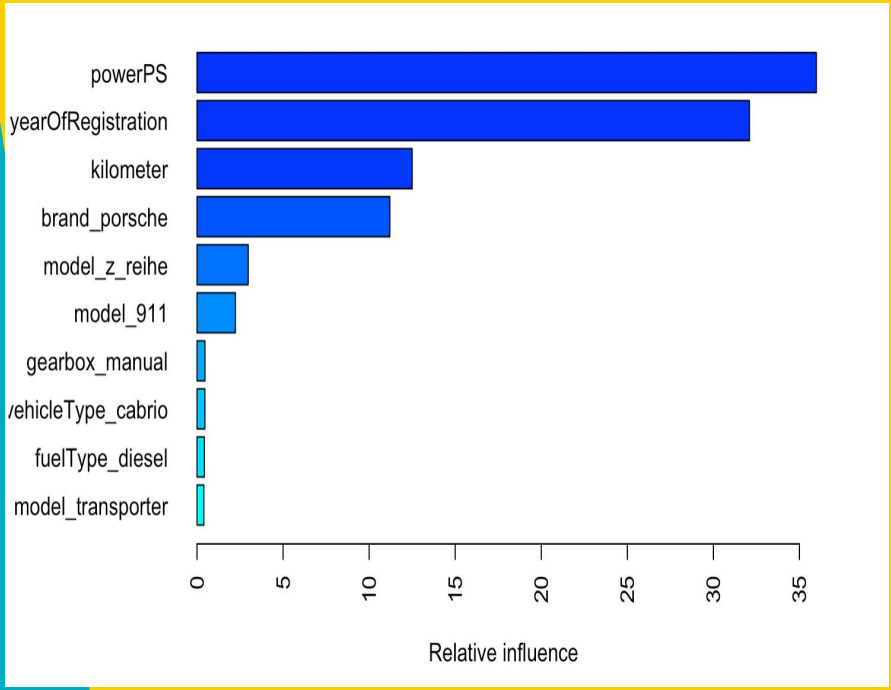
Random Forest

Train RMSE: 2091.875

Test RMSE: 4913.075

Boosting





Min Test RMSE:
8379.893

MSE Comparisons

OLS

RMSE Train : 9609.137

RMSE Test : 11938.49

Ridge

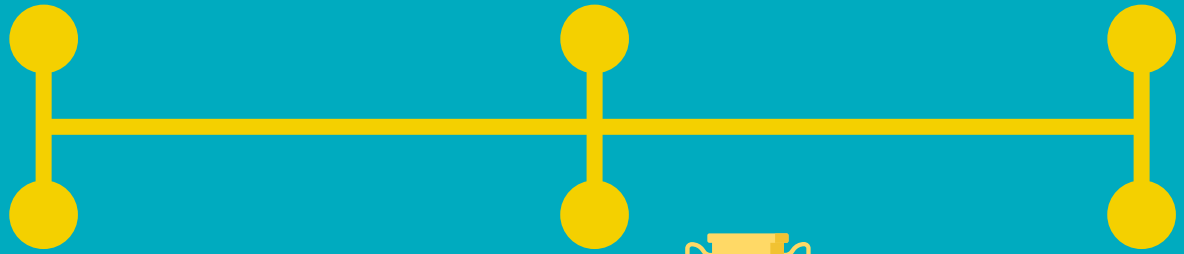
RMSE Train : 11442.35

RMSE Test : 7155.59

Lasso

RMSE Train : 11426.85

RMSE Test : 7195.74



Regression Tree

RMSE Train : 6521.98

RMSE Test : 11622.47

Random Forests

RMSE Train : 2091.875

RMSE Test : 4913.075

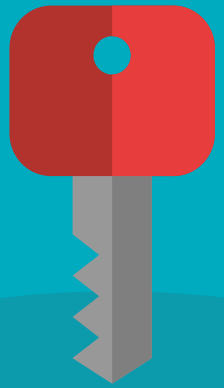
GBM Boosting

RMSE Train : 12395.44

RMSE Test : 8379.893

04

Conclusion



Our biggest challenge

THANKS

Do you have any questions?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik** and illustrations by **Stories**