

Used Car Price Prediction

Ford vs Toyota

- Joseph Onwukeme
- Matthew Witschorke
- Sam Slomowitz
- Yarely Vargas



Data Story



Meet Jim



Jim wants to buy a car



Jim is a data scientist so he...

kaggle



Builds a model to predict price

AGENDA

- 1 Data
- 2 Analysis
- 3 Hypothesis
- 4 Predictions
- 5 Conclusions
- 6 Q&A



About our Data

PURPOSE:

To help others determine how much to pay for a used car (based on Market price)

HOW:

Contains web-scraped data on 100,000 used cars sold in 2020 divided into CSV based on brands. (In this project we focused on Toyota & Ford).

GETS BETTER:

The data set contains mostly clean data with variables such as price, mileage, mpg, year of model, and engine size.

kaggle



Data Cleaning

Price

Only kept the mean price + 2 standard deviations

Mileage

Only kept the mean mileage + 2 standard deviations

Years

Only included 2013-2020

Models

Only kept Aygo, Yaris, Fiesta, Focus

Engine

Size not equal to 0



Models in Data

Aygo



Toyota

Yaris



Focus



Ford

Fiesta





Snapshot of Our Data

Model	Year	Price	Transmission	Mileage	Fuel Type	Tax	mpg	Engine Size
Aygo	2017	15,000	Manual	7,995	Petrol	145	68.9	1.0
Aygo	2018	13,881	Manual	12,447	Petrol	150	56.5	1.0
Yaris	2020	19,276	Automatic	3,113	Petrol	145	47.9	1.5
Yaris	2020	18,908	Automatic	1,133	Petrol	150	47.9	1.5
Fiesta	2019	19,999	Manual	9,000	Petrol	145	40.3	1.5
Fiesta	2019	19,998	Manual	3,361	Petrol	145	40.3	1.5
Focus	2019	26,995	Manual	3,500	Diesel	145	50.4	2.0
Focus	2019	26,950	Manual	7,042	Diesel	145	50.4	2.0

Model	Price mean	Price std	Price var	Price count
Aygo	8035.19	1534.56	2.35E+06	1882
Fiesta	10439.25	2625.35	6.893+6	6168
Focus	13721.95	3845.55	1.48E+07	4174
Yaris	10832.58	2217.99	4.92E+06	2009



Hypotheses



1. There is a difference between sell price for models within brands.

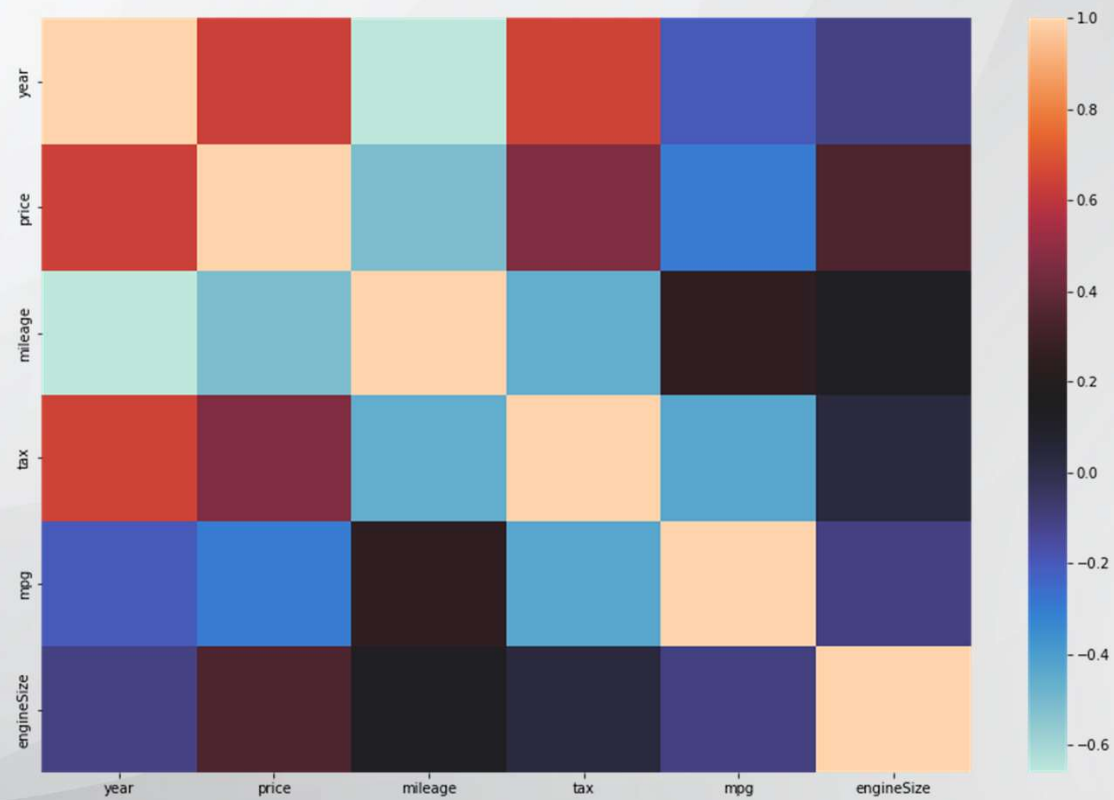


2. The engine size and year of a used car impacts the market price of the car.

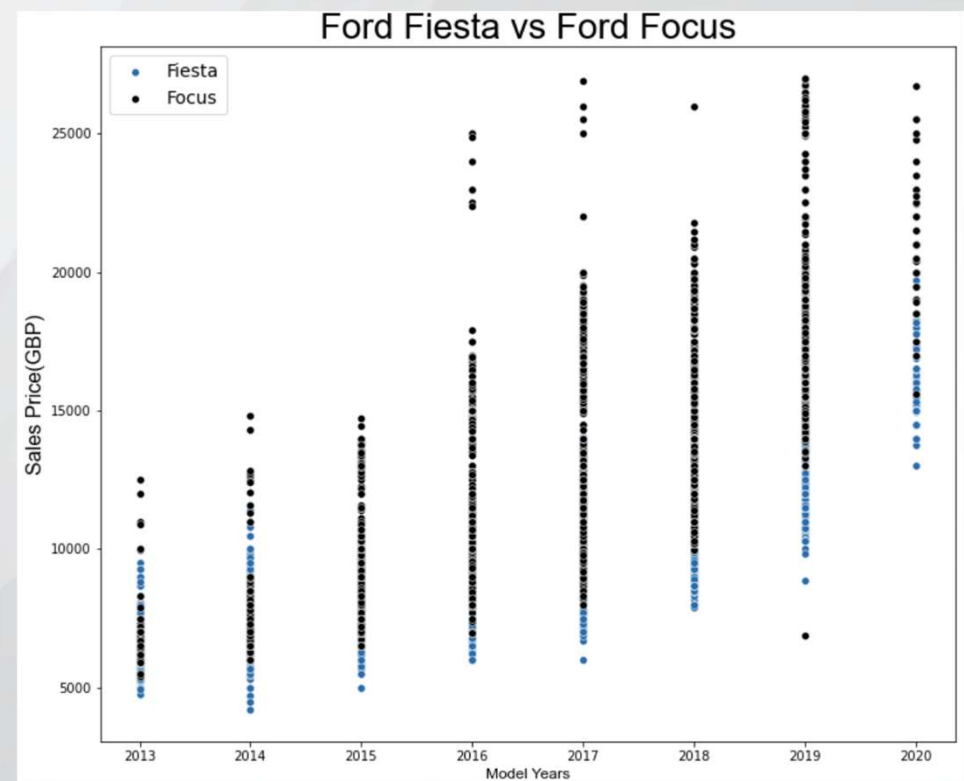
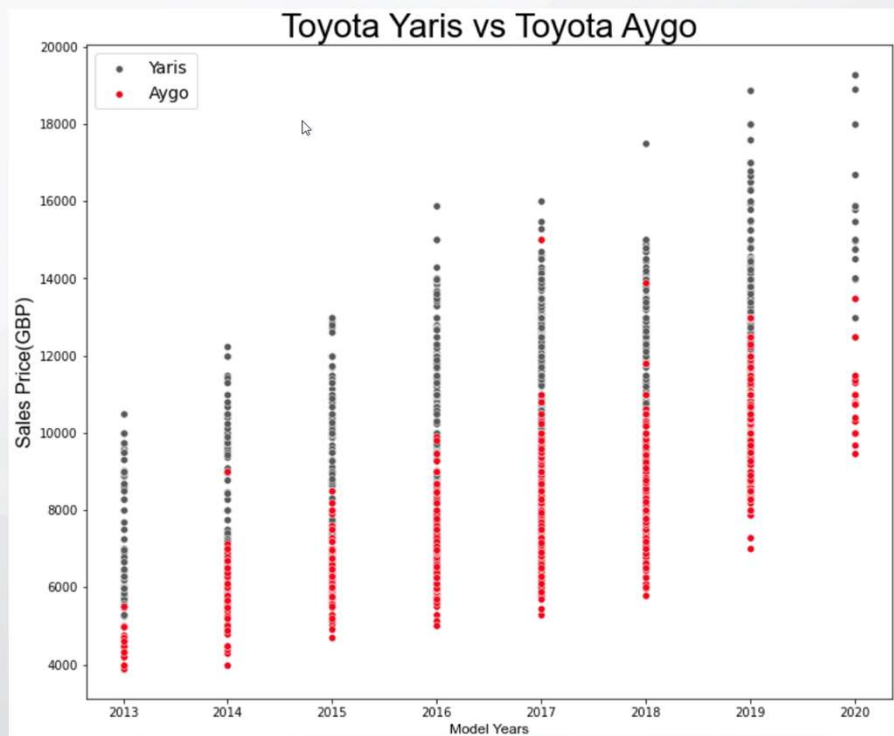


3. There is a difference in Market price between Ford and Toyota.

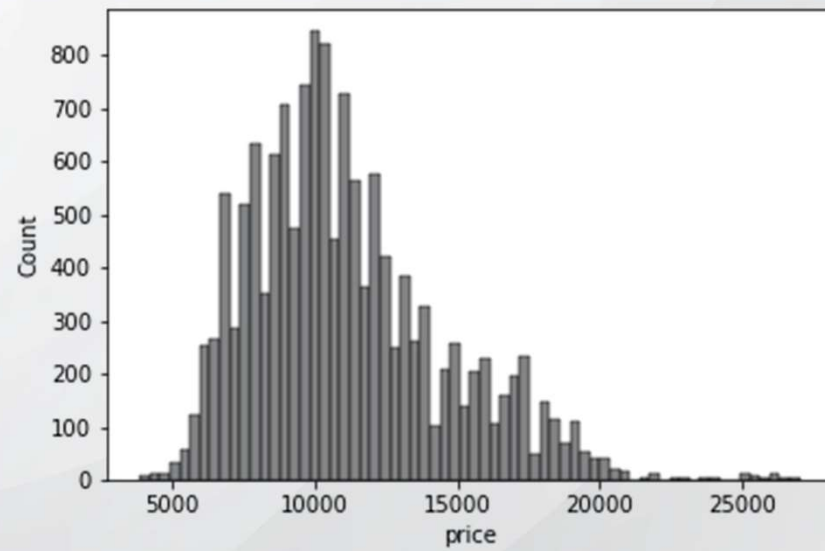
Correlation Heat Map



Price vs Model



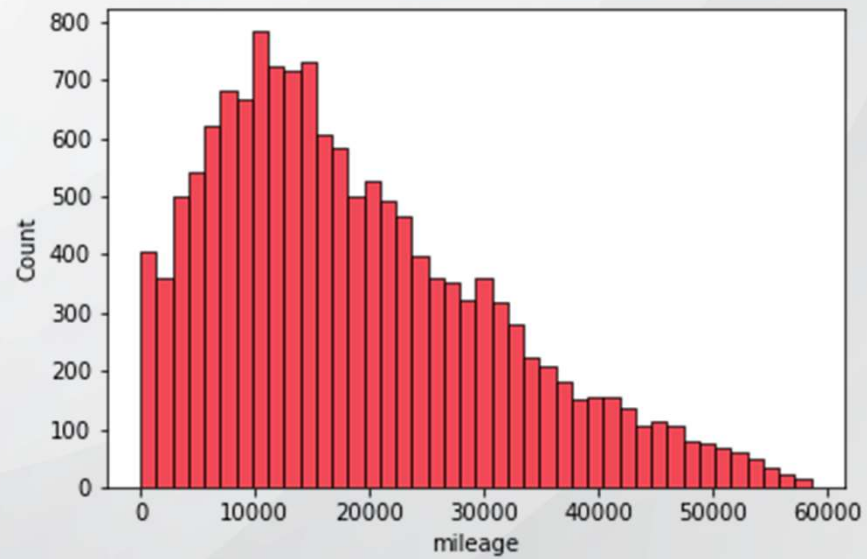
Histogram: Price



price



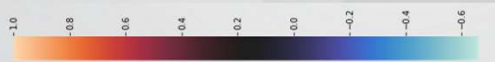
Histogram: Mileage



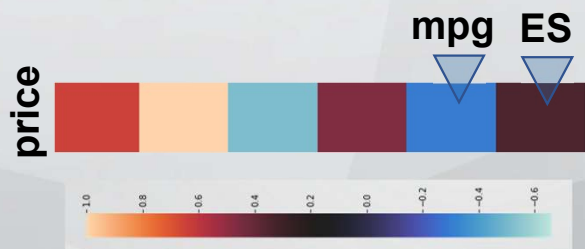
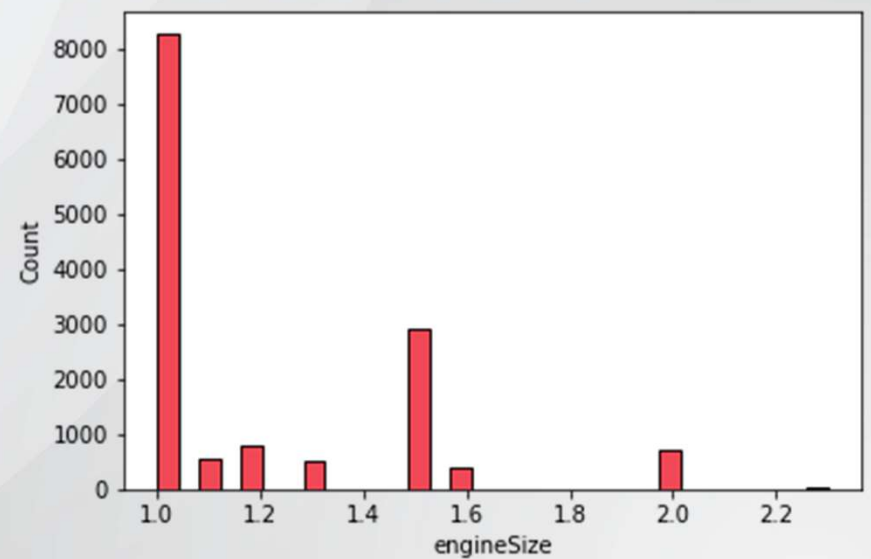
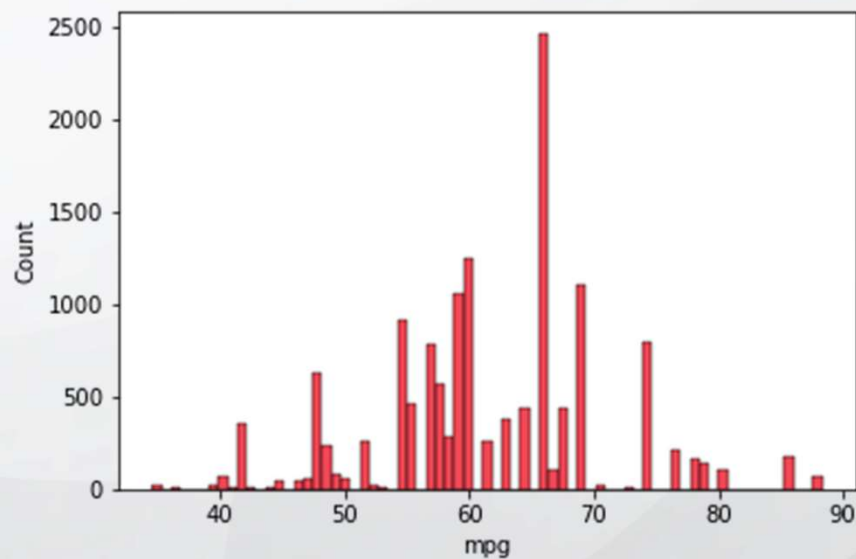
mileage



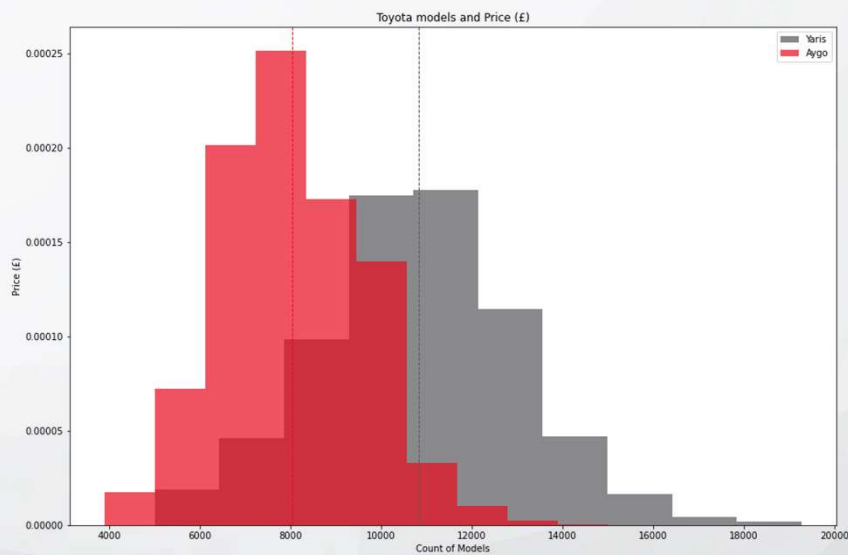
price



Histogram: mpg and Engine Size

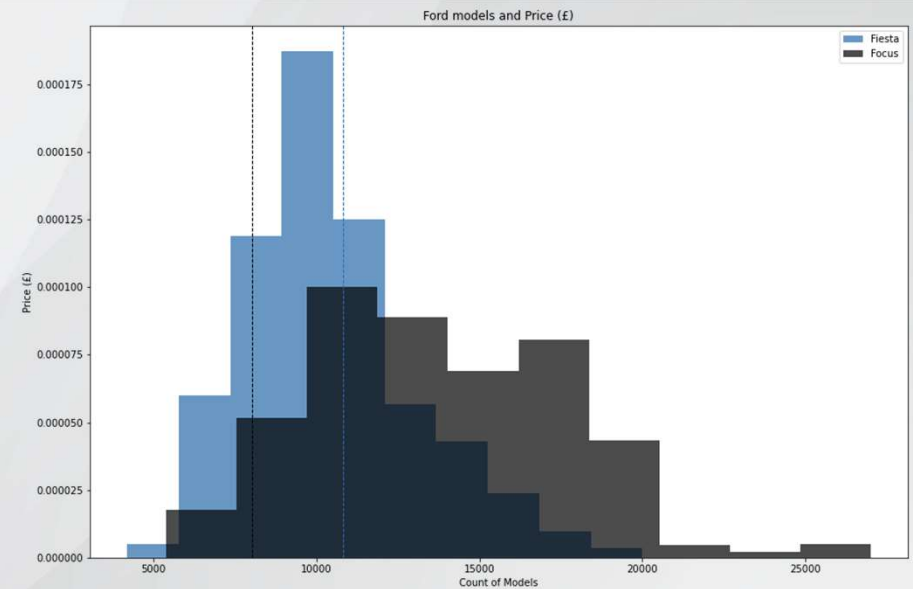


1. There is a difference between sell price for models within brands



$p < 0.0001$

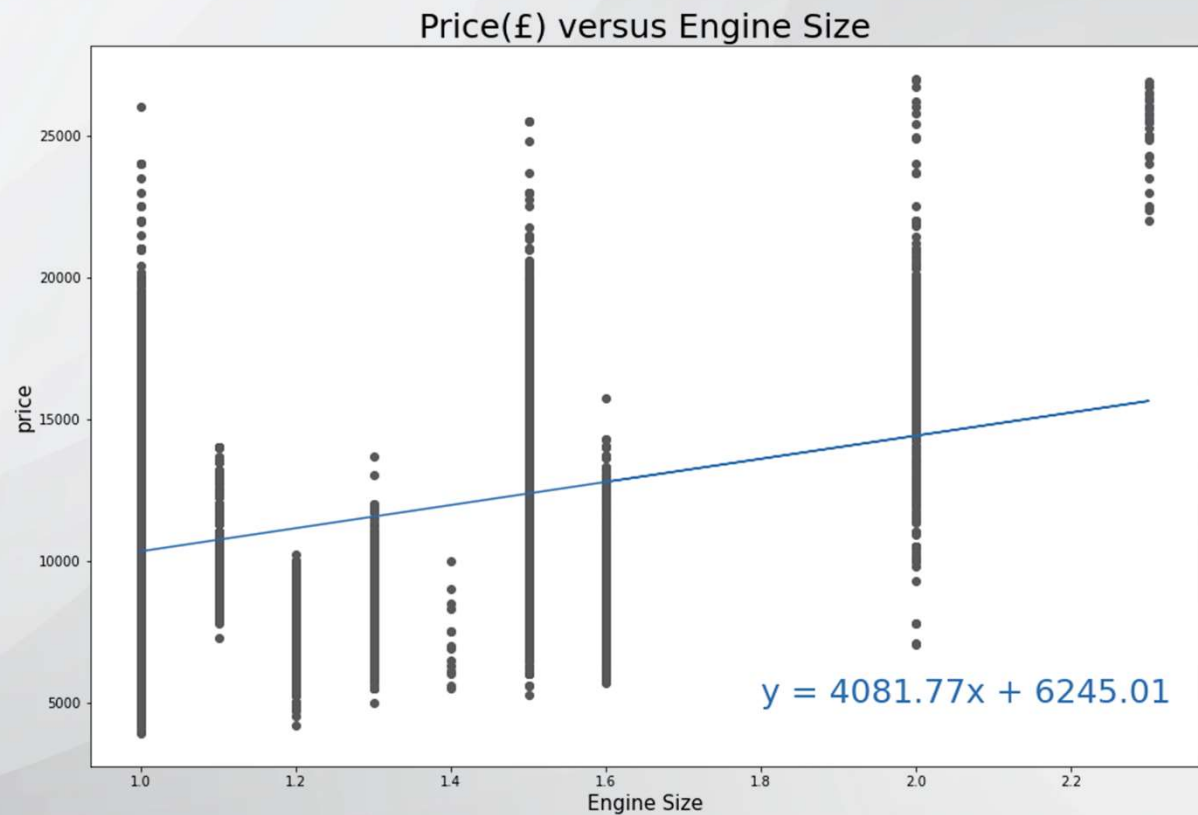
Independent t-test
on both brands



$p < 0.0001$

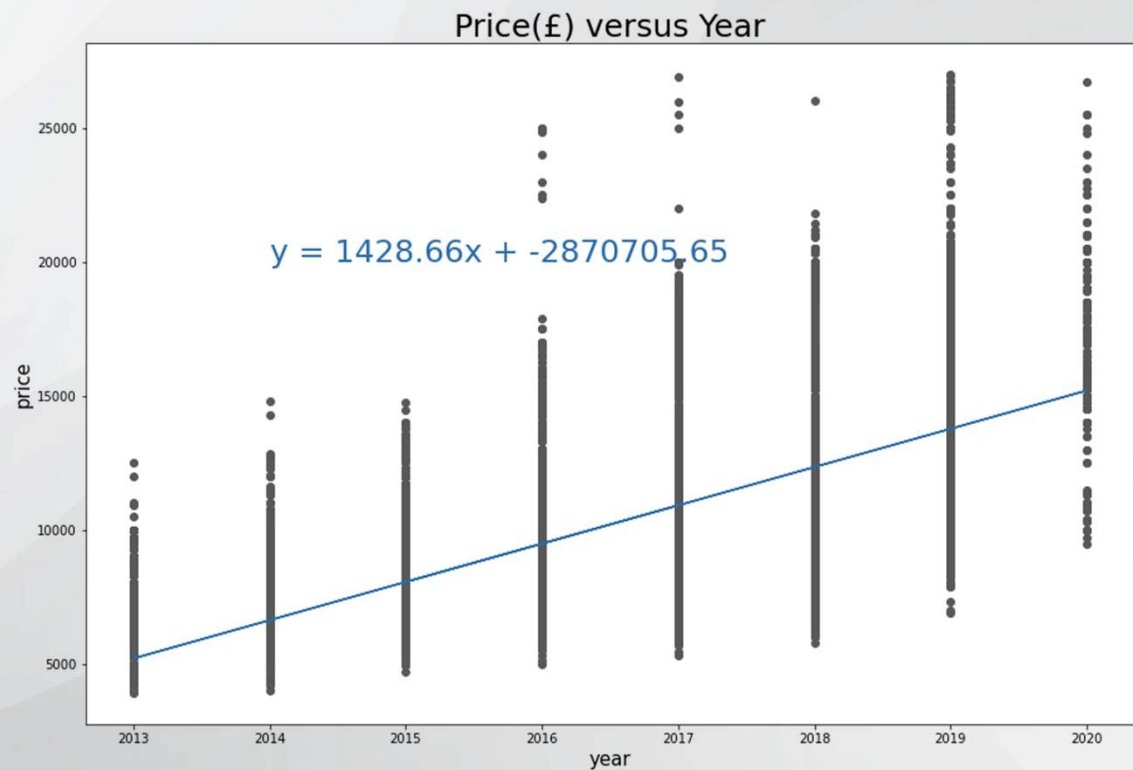
2. The engine size of a used car impacts the market price of the car.

The r-squared is:
0.1164



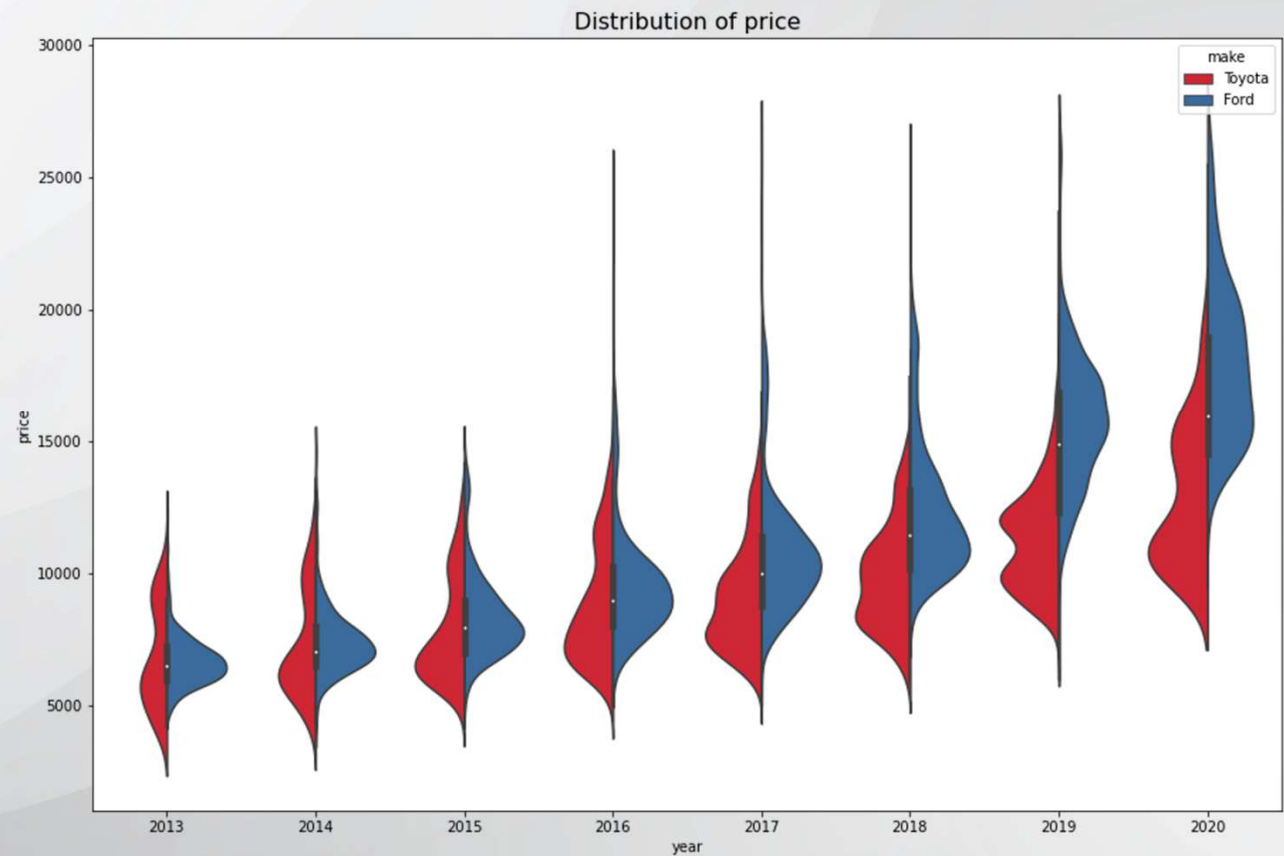
2. The year of a used car impacts the market price of the car.

The r-squared is: 0.41



3. There is a difference in Market price between Ford and Toyota

ANOVA had a p-value less than 0.0001 ($p < 0.0001$)





Hypotheses Review



There is a difference between sell price for models within brands



The engine size and year of a used car impacts the market price of the car.



There is a difference in Market price between Ford and Toyota

Live! Price Prediction

Models:

Fiesta • Focus • Aygo • Yaris



Mileage:

9,000 to 30,000



MPG:

55 to 65
(integers only)



Engine Size:

1 to 1.5
(in increments of 0.1)

Disclaimer: Price prediction does not consider, year of used car, damage of used car, fuel type, transmission type or emotional attachment to used car. Model should be used purely as estimate and not fact.



Limitations



Low r -squared value for three major simple linear regressions



Our model is not generalizable to other countries or regions of the world.



There are other variable to used car sales than car metrics such as state of the economy, oil prices, etc.

Future Work



Explore year 2020 for unique impact of oil price on transmission type.



Web-scrap data from 2022 and compare it with data of 2020



Use APIs to generate new data set on new cars



Apply code to U.S. car data set and compare with the U.K.



Apply our code to the other CSV files





Call to Action



With a budget of £10,000 to £18,000 and to retain value over time, buy

- a 2017 for £11,000,
- a 2018 for £12,000,
- a 2019 for £13,000
- or a 2020 for £14,000



Buy a car with an engine size between 1.2 and 1.6



Summary: 2017, £11,000, 1.3 engine size

Thank You!

Q & A

