Hyperiondev

TASK

Exploratory Data Analysis on the Automobile Data Set

Visit our website

Introduction

This dataset was full of information about cars.

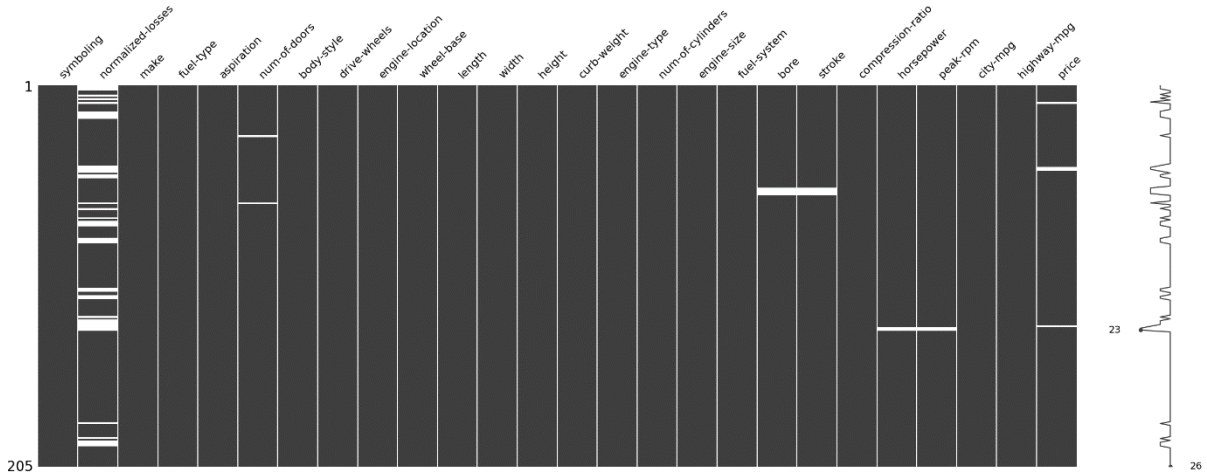
DATA CLEANING

First, I replaced all the '?' with NaN so that missing values were recognised.

I next made sure that all attributes were the correct data type and corrected the ones that were not.

MISSING DATA

I then checked for any missing values using missingno:



It can be seen that most of the missing values are in normalized-losses so I decided not the use this attribute in my analysis.

I then decided to drop the remaining rows the missing values.

DATA STORIES AND VISUALISATIONS

I then tackled 5 questions:

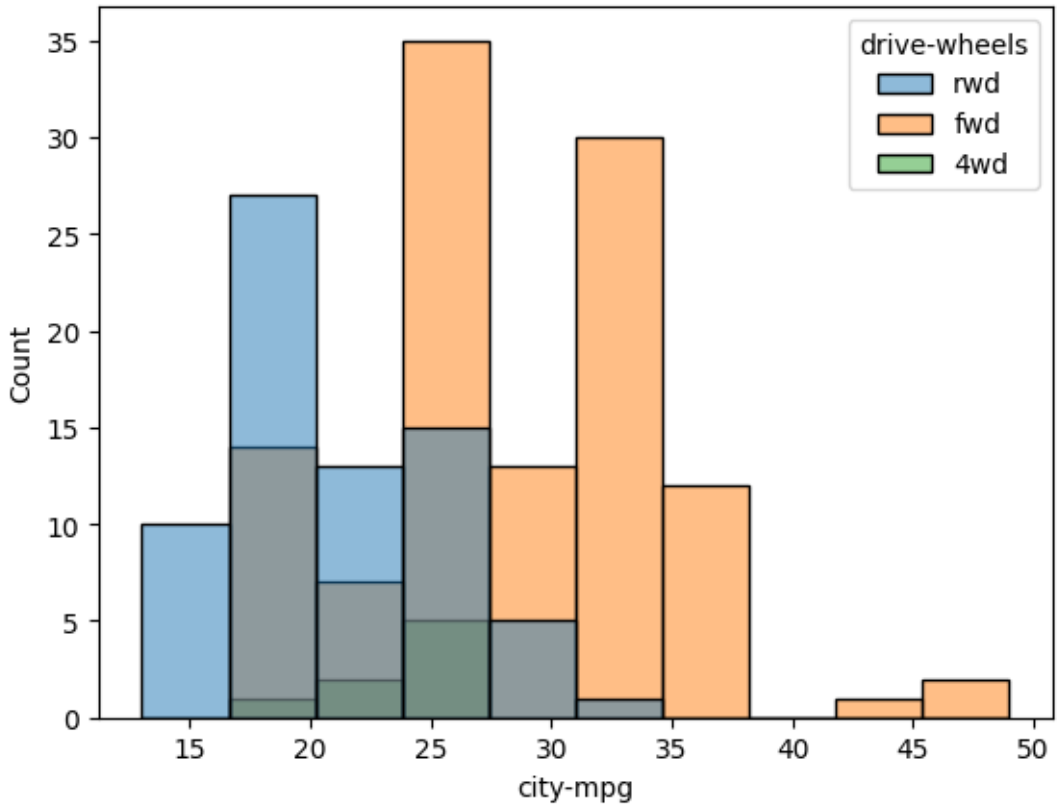
What are the top 5 most expensive cars?

I sorted the dataset by price and isolated the top 5 cars to show which were the most expensive.

	symboling	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheels	engine-location	wheel-base	length	...	engine-size	fuel-system	bore	stroke	compression-ratio	horsepower
74	1	mercedes-benz	gas	std	two	hardtop	rwd	front	112.0	199.2	...	304	mpfi	3.80	3.35	8.0	184.0
16	0	bmw	gas	std	two	sedan	rwd	front	103.5	193.8	...	209	mpfi	3.62	3.39	8.0	182.0
73	0	mercedes-benz	gas	std	four	sedan	rwd	front	120.9	208.1	...	308	mpfi	3.80	3.35	8.0	184.0
128	3	porsche	gas	std	two	convertible	rwd	rear	89.5	168.9	...	194	mpfi	3.74	2.90	9.5	207.0
17	0	bmw	gas	std	four	sedan	rwd	front	110.0	197.0	...	209	mpfi	3.62	3.39	8.0	182.0

What type of drive wheels are the most fuel efficient in the city?

I created a histogram with all 3 drive wheel types on it.



It can be seen that forward drive is, on average, the most fuel efficient in the city.

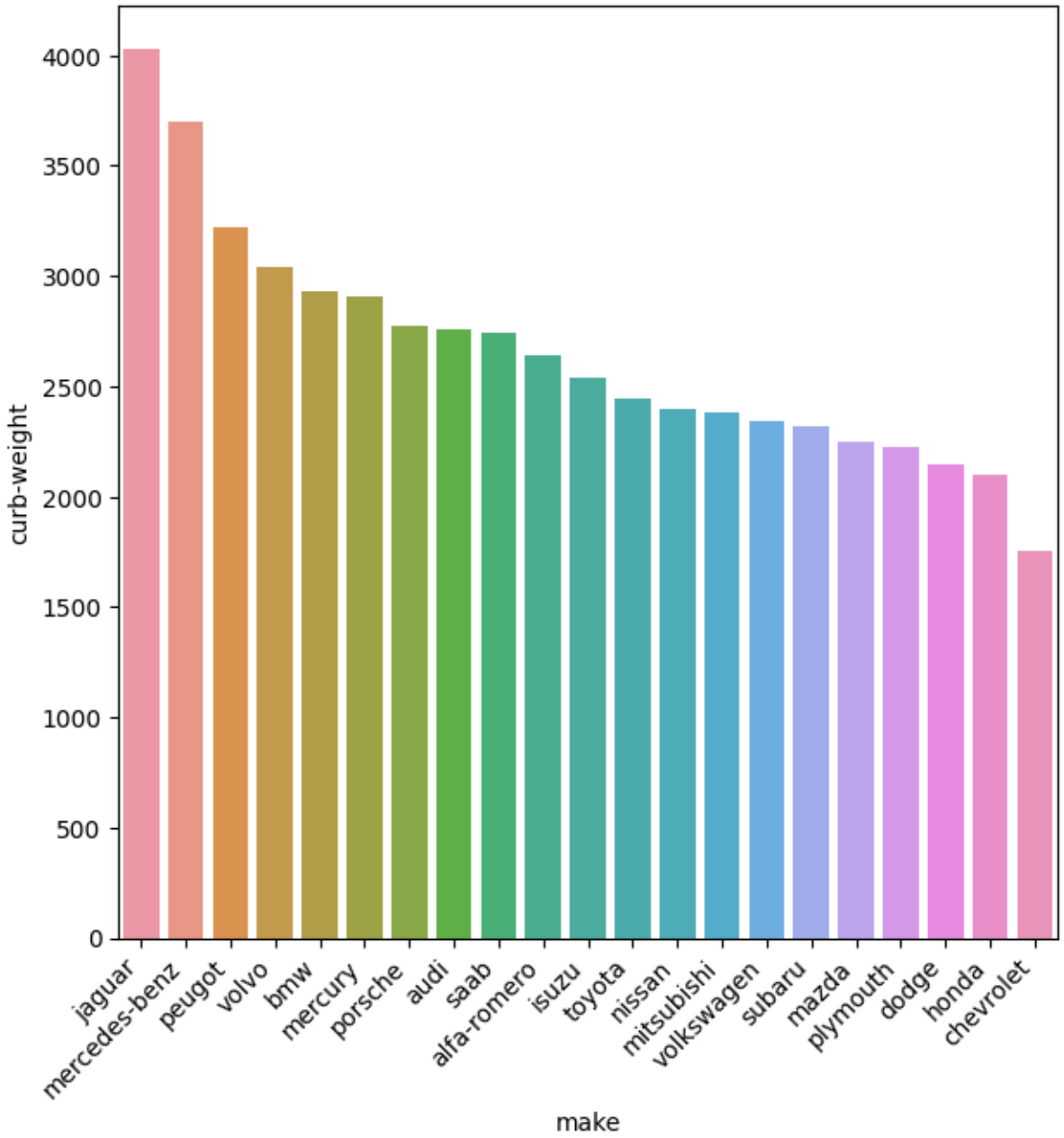
What are the top 5 heaviest cars?

I sorted the dataset by curb weight and isolated the top 5 cars to show which were the heaviest.

	symboling	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheels	engine-location	wheel-base	length	...	engine-size	fuel-system	bore	stroke	compression-ratio	horsepower
48	0	jaguar	gas	std	four	sedan	rwd	front	113.0	199.6	...	258	mpfi	3.63	4.17	8.1	176.0
47	0	jaguar	gas	std	four	sedan	rwd	front	113.0	199.6	...	258	mpfi	3.63	4.17	8.1	176.0
49	0	jaguar	gas	std	two	sedan	rwd	front	102.0	191.7	...	326	mpfi	3.54	2.76	11.5	262.0
73	0	mercedes-benz	gas	std	four	sedan	rwd	front	120.9	208.1	...	308	mpfi	3.80	3.35	8.0	184.0
70	-1	mercedes-benz	diesel	turbo	four	sedan	rwd	front	115.6	202.6	...	183	idi	3.58	3.64	21.5	123.0

What's the heaviest make on average?

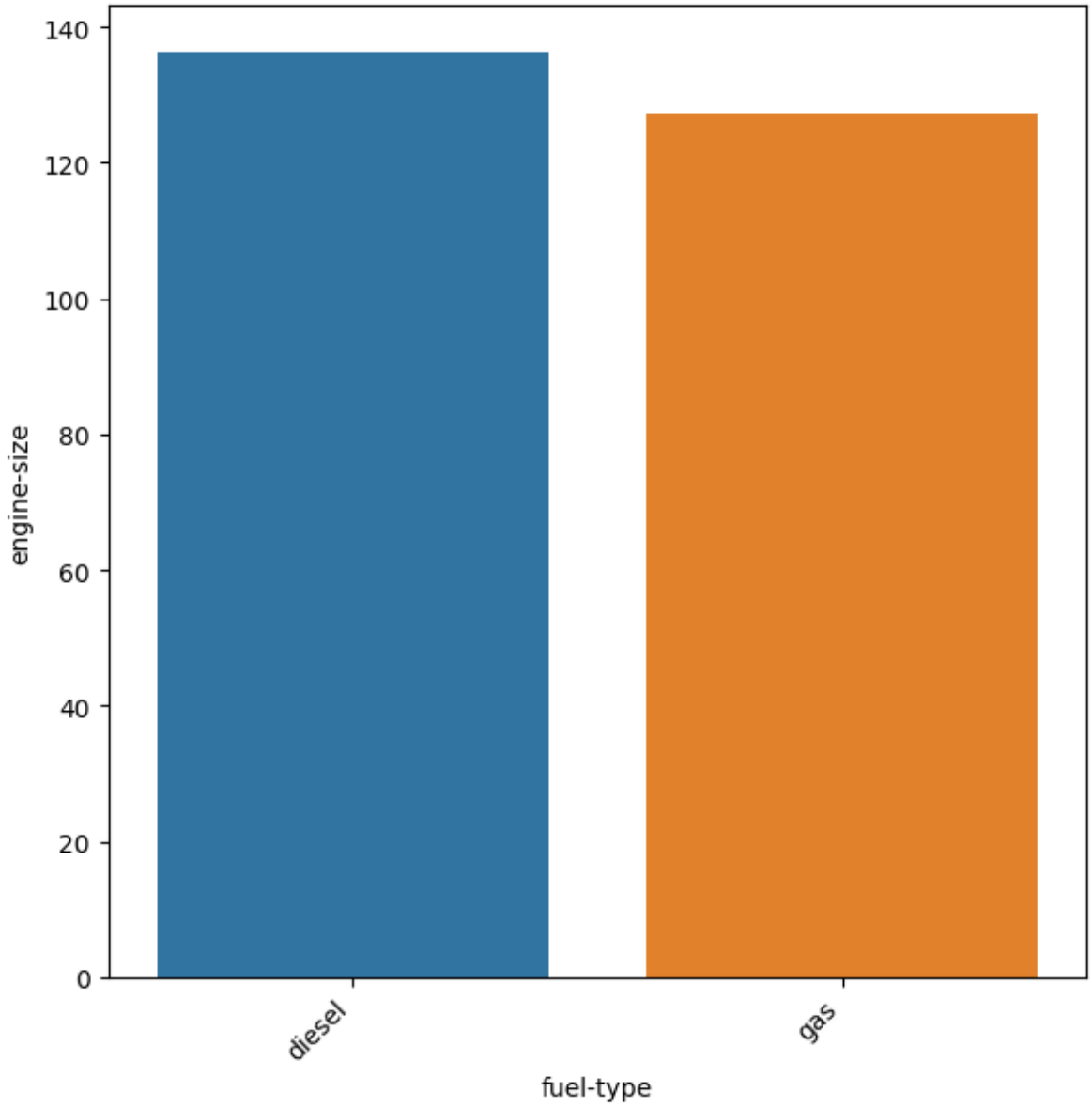
I grouped the dataset by the make of the cars and then found the mean weight of each group. I then sorted the dataset by curb weight and created a bar chart to show the information.



It can be seen that Jaguar's, on average, are the heaviest make.

Which fuel type has the bigger engine size?

I grouped the dataset by the fuel type of the cars and then found the mean engine size of each group. I then sorted the dataset by engine size and created a bar chart to show the information.



It can be seen that diesel engine's, on average, are bigger.

THIS REPORT WAS WRITTEN BY : Robert Deed