



Car Sales Data Report

12/12/2022

DS311 Professor Norman Lo

Angie Llamas, Barbara Wallen, Chris Albert & Kris Guan
Passiflora Team

Overview

The Used Car Data set are indexed rows of car listings from various sellers across the United States. The columns have various attributes of the listings such as the naming of the current mileage of the car, the make, the drive-type and any other relevant information regarding a car listing. Our main objective of this data analysis was to find the answers to high level questions that prospective buyers or sellers would need to know before using the marketplace that the dataset is from.

Goals in Data Deep Dive:

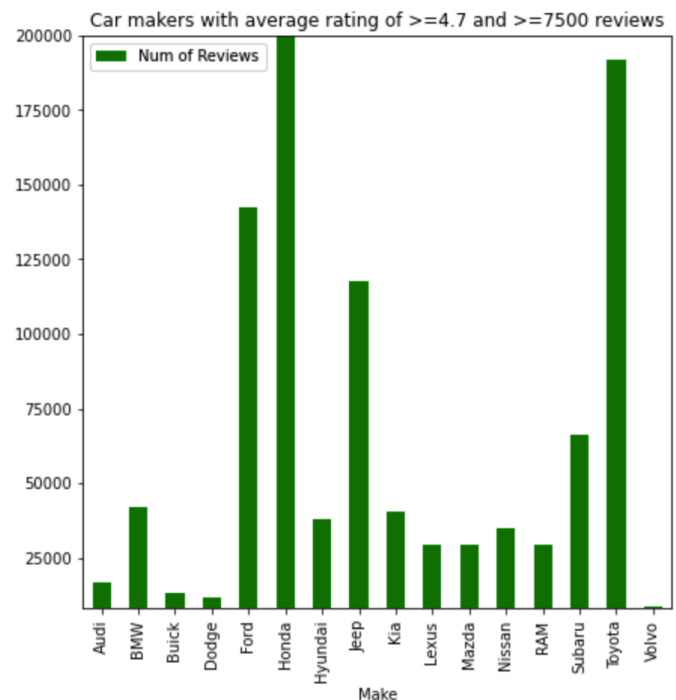
1. Find the most popular cars in the dataset.
2. Find the most economic car, based on number of reviews, mileage, max MPG.
3. Where are the most used/new cars?
4. Is it better to buy from a private seller or dealer?
5. Does the price vary by region?

Question 1: What is the most popular car?

To first find the most popular car we need to define what makes a car popular? Inferring that popularity is directly connected to units sold, we based our popularity on the Rating column and The Reviews column. The Reviews column is the sample size from which the Rating is based on. The more reviews means the more popular and less erroneous the Rating is. There are other rating columns in the dataset that could be a driver to why people buy a car, but it does not necessarily make it popular.

After setting a benchmark of rating that shows a benchmark that visually separates the car makers, we began to see clearly what the three most popular car manufacturers in the dataset were. They were Honda, Toyota, and Ford. Then applying the same method of data filtering grouped by each of the most popular manufacturers, we were able to find the most popular car model from each of the top three makes. The top three makes were the

- 2018 Ford 150 XL
- 2017 Toyota Camry XLE
- 2018 Honda Civic EX



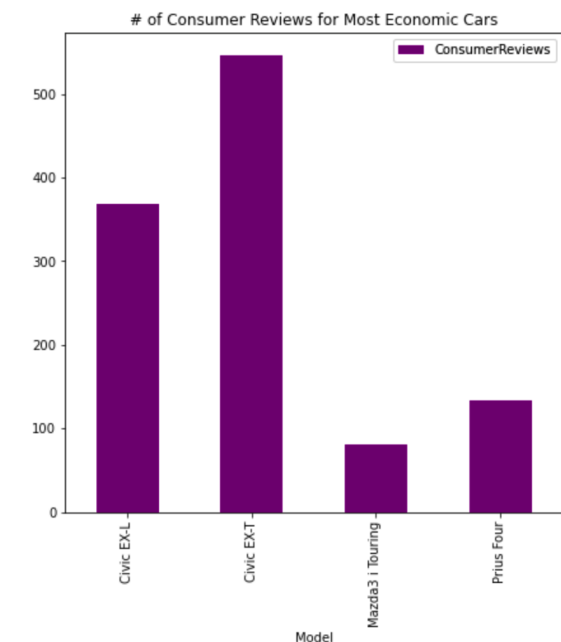
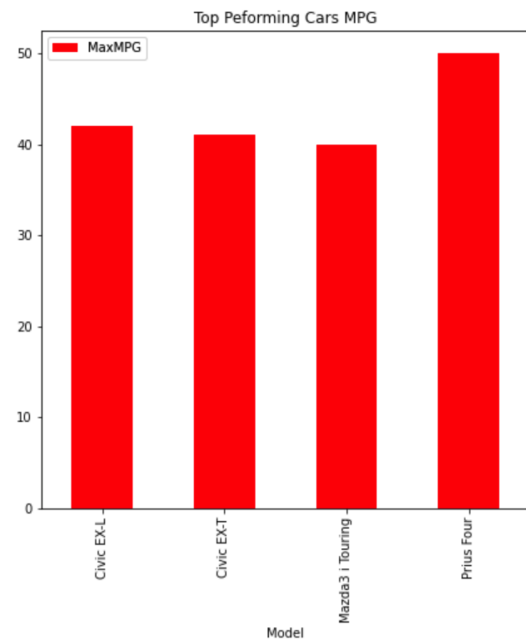
Finding the years for the most popular models is the same method as the prior two filters. Group by the most popular model this time and apply a filter for ratings and reviews, and then select which year has the highest reviews in tandem with the highest rating.

Question 2: What is the most Economic Car?

To find the most economical car, we decided to base what economic means on the Value for Money Rating, the Reliability Rating, the Reviews Column, and the Max MPG Column. We felt that these columns are what people would look for in the marketplace when deciding to buy an everyday car that will not break down often, get good gas mileage, and does not cost a fortune.

After filtering Cars with a certain threshold for number of reviews and at least a 4.7 Rating for Value for Money, Reliability, and at least 40 MPG. We were able to see which models were the most economic out of the dataset.

We then sorted the cars by the number of reviews they each had. The car with the highest number of reviews would then be the most economic car because the larger the number of reviews means that the sample size is therefore larger. The larger the sample size means there are less standard deviations of the average reviews. Meaning the Value, Reliability, MPG are all thus more concrete. After the analysis, we were able to find that Honda Civic EX-T was the most economic car in the dataset.



Question 3: Where are the most used/new cars?

For the third question, we found that the dataset has 9379 observations and 32 variables. The geographical level used in this analysis is state and will be used to determine the most used car or new car in this platform.

- The results show that there are more used cars (7900) as compared to new ones (1479)

Used	7900
BMW Certified	247
Mercedes-Benz Certified	223
Honda Certified	208
Toyota Certified	153
Cadillac Certified	105
Ford Certified	71
Subaru Certified	58
Jeep Certified	55
Nissan Certified	50
Acura Certified	47
Chevrolet Certified	40
INFINITI Certified	34
Kia Certified	34
Volvo Certified	33
Porsche Certified	25
RAM Certified	24
Buick Certified	20
Volkswagen Certified	17
GMC Certified	13
Dodge Certified	10
Alfa Romeo Certified	7
Bentley Certified	2
MINI Certified	1
Ferrari Certified	1
Maserati Certified	1

- The state with the highest count of used cars is TX with 1069 used cars.
- The state with the newest cars is FL with a count of 157 cars.

TX	Acura Certified	3	FL	Acura Certified	5
	Alfa Romeo Certified	3		BMW Certified	13
	BMW Certified	24		Buick Certified	1
	Buick Certified	2		Cadillac Certified	23
	Cadillac Certified	18		Chevrolet Certified	2
	Chevrolet Certified	4		Ford Certified	6
	Dodge Certified	1		GMC Certified	1
	Ford Certified	11		Honda Certified	18
	GMC Certified	1		INFINITI Certified	3
	Honda Certified	12		Jeep Certified	4
	INFINITI Certified	17		Kia Certified	14
	Jeep Certified	1		MINI Certified	1
	Mercedes-Benz Certified	19		Mercedes-Benz Certified	32
	Nissan Certified	9		Nissan Certified	6
	Porsche Certified	3		Porsche Certified	2
	RAM Certified	2		RAM Certified	2
	Subaru Certified	4		Subaru Certified	2
	Toyota Certified	12		Toyota Certified	16
	Used	1069		Used	731
	Volvo Certified	10		Volkswagen Certified	2
				Volvo Certified	4

Also, from this dataset, we were able to know which brand and where used cars came from. In Texas, the most used car is Toyota with 128 cars, Mercedes-Benz 89, and 81 from Ford. And the newest car is BMW certified with 24 cars.

In Florida, the most used cars are Mercedes and Toyota with 81 cars and the new cars are also Mercedes-Benz with 32 cars. Interestingly, results from both states where Mercedes-Benz seems to be the top favorite car.

There's an article from [CNBC news](#), that states why Mercedes-Benz became a prestigious car, meaning that the reason why people would buy used or new, is because these cars are well designed, well made, and high quality and performance.

Question 4: Is it better to buy from a dealer or private seller?

For the fourth question, we were asked to find out if it's better to buy from a dealer or a private seller. We found that, are 9339 cars bought from dealers and 40 from private sellers.

- We wanted to focus on Texas and Florida, to compare with question 3.

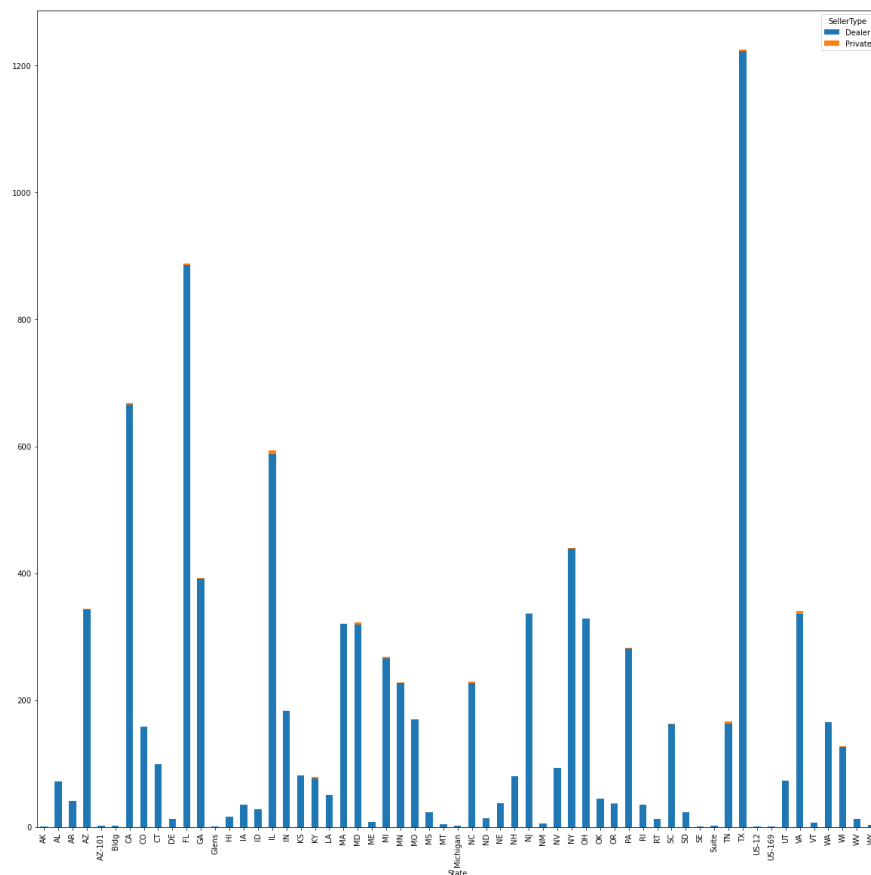
TX	Dealer	1223
	Private	2

FL	Dealer	886
	Private	2

- The state that has the most private seller is Chicago

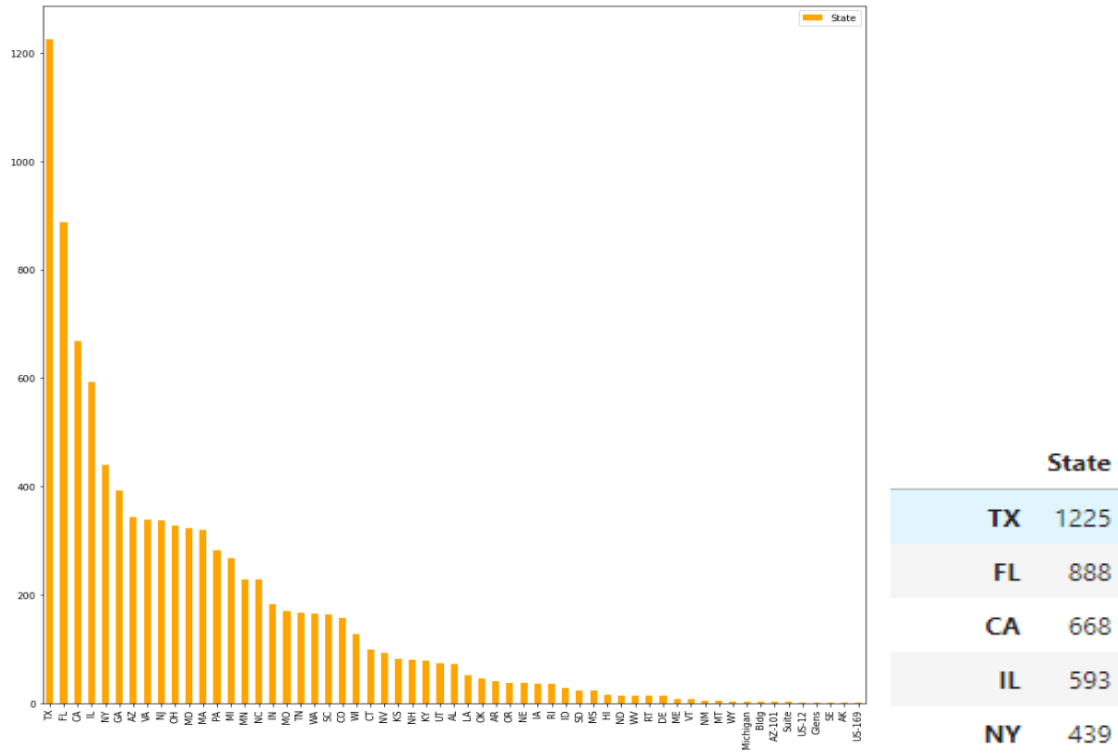
IL	Dealer	588
	Private	5

Well, this dataset doesn't give us the reason behind it but after some research, this can be explained by the fact that buying cars from dealers may be the best option because it offers more protection, handles paperwork and registration, and financing options. One of the biggest reasons why someone would buy a car from a private is for the price and if there is a specific car that you need/want.



Question 5: Does the price vary by region?

For the fifth question, we found that this dataset contains the car sale listing with over 9000 different listings across the country. After exploring this data to find insight that prices of car sales vary from region and some states are best for buying or selling.



The top five states with the most listings: Texas with 1225, Florida with 888, California with 668, Illinois with 593 and New York with 439.

Zip Code 76051, 75093 and 75034 are located in the state of Texas.

Zipcode	
76051	69
75093	65
30096	63
29607	48
75034	46

Texas has the most car sales listing, it means that compared with other states, Texas has more cars with different models or prices in the market across the country. Texas is best for car buying or selling.



