CARS RENTAL COMPANY

By Annal Ali Al-Beeshi

OUTLINE

INTRODUCTION 1

Motivation 2

Data Description 3

4 **Methodology**

5 CONCLUSIONS

INTRODUCTION

Motivation

➢ One of the most important motives that we seek to achieve is to analyze the data by answering some questions:

1. **What is the brand with the highest price and in which state it is?**
2. **Does the model year affect the price of the car?**
3. **What states have clean cars, and what are the most popular brands in each state?**
4. **What are the most common car colors in each state?**

Data Description

Data Description

➢ This data file includes Information about 28 brands of vehicles for hire in US. Twelve features were assembled for each car in the dataset.

➢ 13 columns, 12 features for each car, and 2499 rows, a dataset has three types of data, as shown in Table 1, which includes lists the features, data type, and description for each column.

Table 1: Describe columns by clarifying features and data type.

Features	Data Type	Description
Price	Integer	The sale price of the vehicle in the ad.
Years	Integer	The vehicle registration year.
Brand	String	The brand of car.
Model	String	model of the vehicle.
Color	String	Color of the vehicle.
State/City	String	The location in which the car is being available for purchase.
Mileage	Float	miles traveled by vehicle.
Vin	String	The vehicle identification number is a collection of 17 characters (digits and capital letters).
Title Status	String	This feature included binary classification, which are clean title vehicles and salvage insurance.
Lot	Integer	A lot of numbers are identification the number assigned for a particular quantity or a lot of materials from a single manufacturer. For cars, a lot numbers are combined with a serial number to form the Vehicle Identification Number.
Condition	String	Time

Methodology

Methodology

Data Exploring
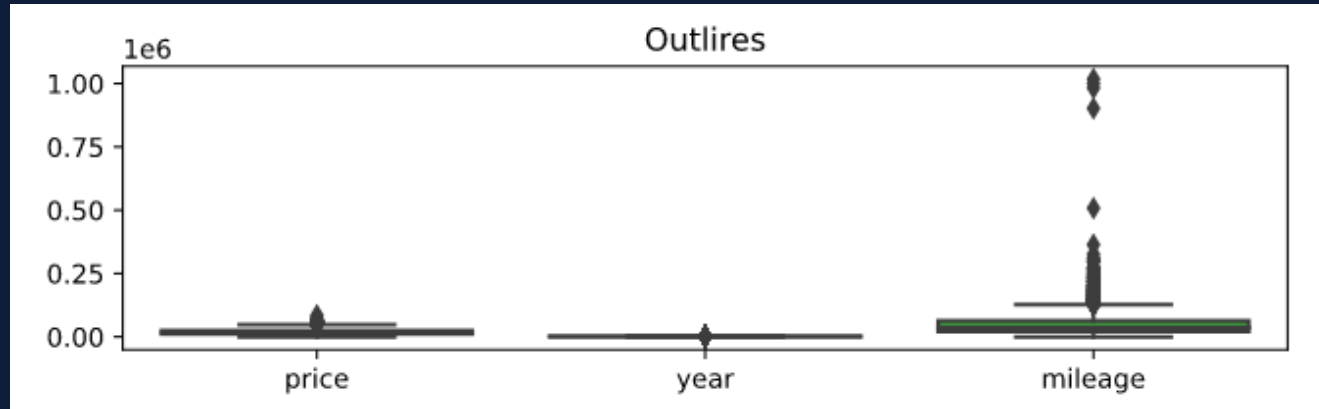
Data Cleaning & Filtering

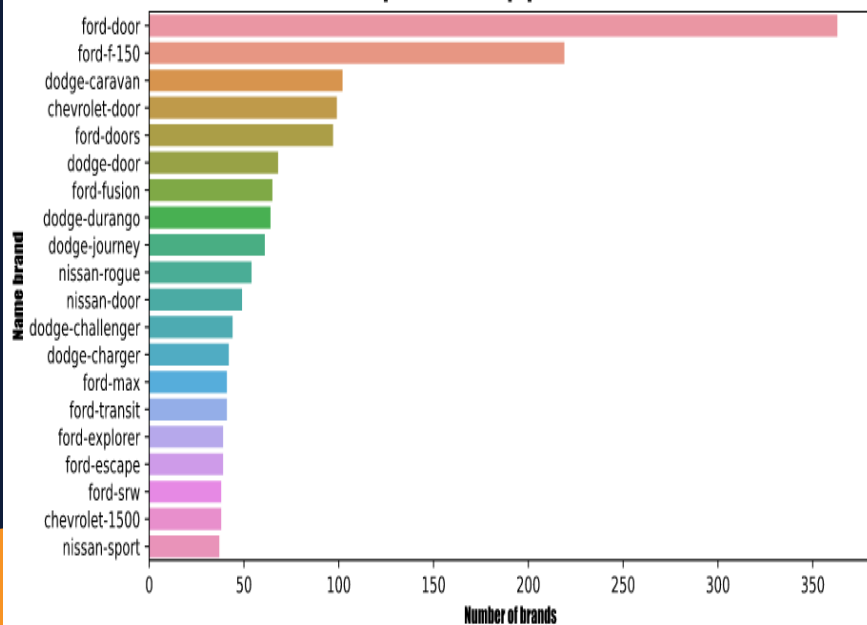**Features Selection**

Analysis & Visualization

Methodology…

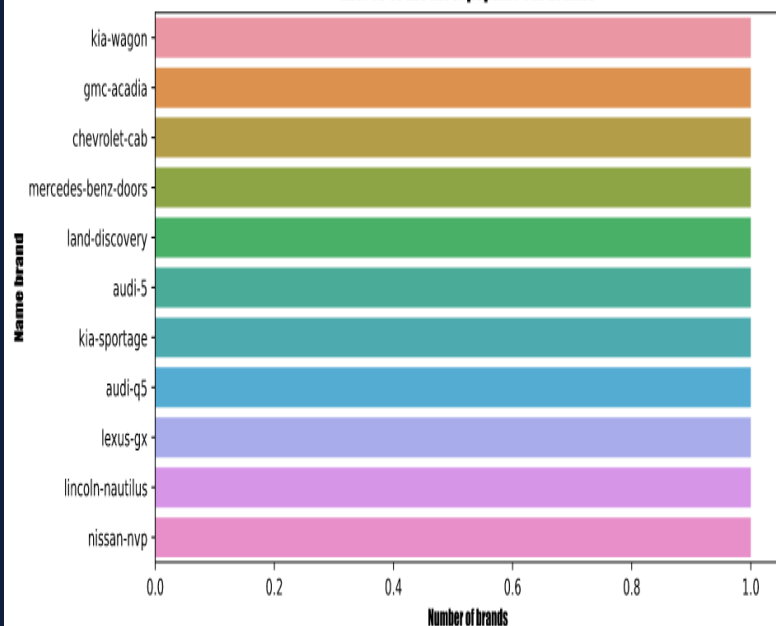**visualize numeric data for check from Outlires value.**
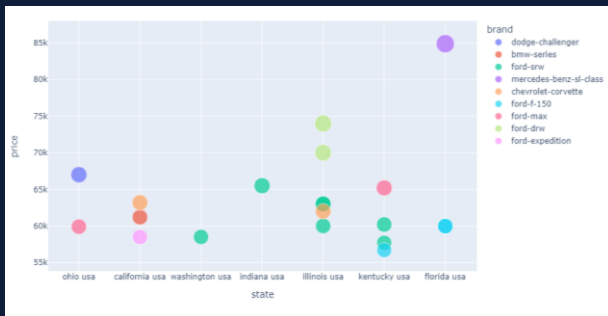
Analyze each column in the dataset
Brand

Questions

Questions

1- What is the brand with the highest price and in which cities it is?

➢ **The first step** is to display prices from the highest to the lowest.

➢ **The second step** is to determine the high prices only

	price	brand	year	title_status	mileage	color	state	condition	price_binncd	price_cat
502	84900.0	mercedes-benz-sl-class	2017	clean vehicle	25302	silver	florida usa	2 days left	High	High(19091-84900)
1340	74000.0	ford-drw	2019	clean vehicle	10536	no_color	illinois usa	2 days left	High	High(19091-84900)
1336	70000.0	ford-drw	2019	clean vehicle	9643	no_color	illinois usa	2 days left	High	High(19091-84900)
277	67000.0	dodge-challenger	2019	clean vehicle	10944	blue	ohio usa	21 hours left	High	High(19091-84900)
1215	65500.0	ford-srw	2019	clean vehicle	6500	black	indiana usa	21 hours left	High	High(19091-84900)
...	...	...	...	...	...	...	...	...	...	...
386	25.0	gmc-door	2003	salvage insurance	138441	white	michigan usa	2 days left	Low	LOW(0-25)
383	25.0	chevrolet-door	2007	salvage insurance	194903	gray	virginia usa	2 days left	Low	LOW(0-25)
362	25.0	ford-pickup	1994	salvage insurance	206162	white	georgia usa	2 days left	Low	LOW(0-25)
359	25.0	chevrolet-door	2008	salvage insurance	130623	white	california usa	8 days left	Low	LOW(0-25)
470	25.0	ford-pk	2010	salvage insurance	155242	gray	oklahoma usa	17 hours left	Low	LOW(0-25)

Questions

2- Does the model year affect the price of the car? Yes

The effect of the model year on prices

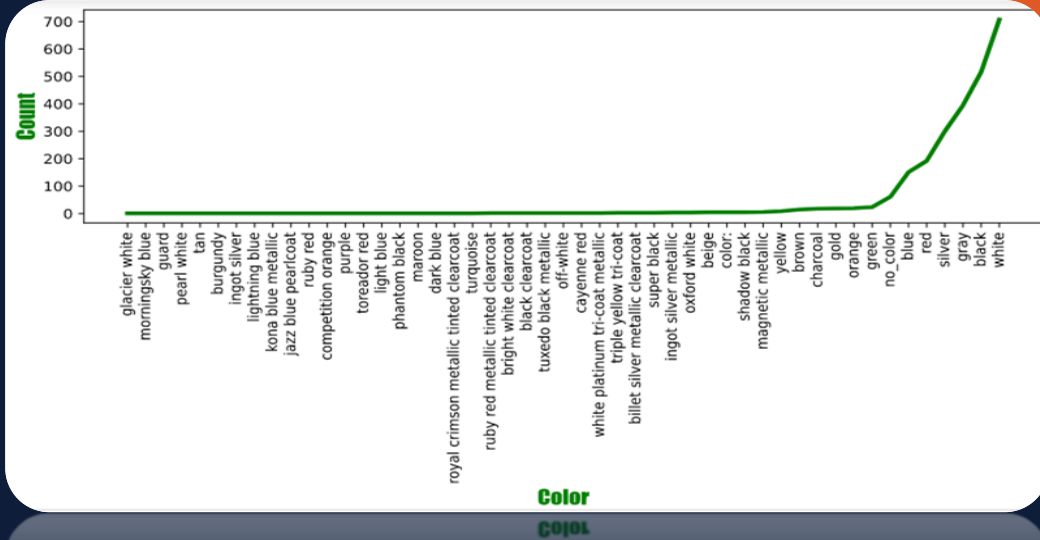We observe not strong relationship of price and year.

Questions

3- What states have clean cars, and what are the most popular brands in each state?

Questions

4- Since white is the most common color in cars, we want to know what cities white colors are in, and how many white cars are in each brand?

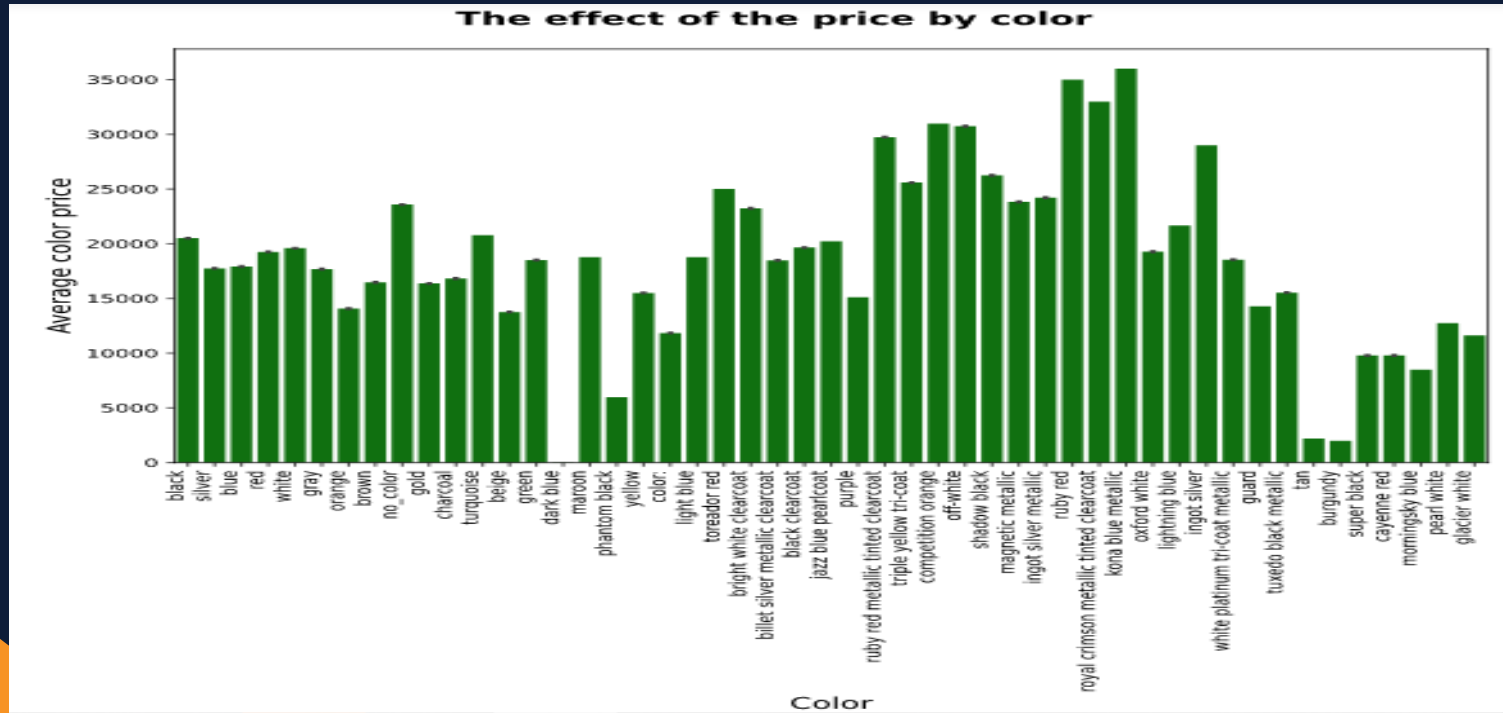| state	brand	
texas usa	ford-door	39
pennsylvania usa	ford-f-150	20
michigan usa	ford-f-150	19
oklahoma usa	ford-f-150	10
texas usa	ford-mpv	9
		..
missouri usa	ford-expedition	1
	ford-f-750	1
	ford-taurus	1
	ford-transit	1
wisconsin usa	nissan-pathfinder	1

We noticed that the number of white color in cars was the highest, but in the low price rate.

	price	brand	year	title_status	mileage	color	state	condition	price_binncd	price_cat
5	5700.0	dodge-mpv	2018	clean vehicle	45561	white	texas usa	2 days left	Low	Medium(26-19090)
11	12920.0	gmc-mpv	2017	clean vehicle	39650	white	california usa	20 hours left	Low	Medium(26-19090)
15	12710.0	gmc-door	2017	clean vehicle	25747	white	california usa	20 hours left	Low	Medium(26-19090)
18	5210.0	ford-mpv	2017	clean vehicle	35714	white	texas usa	2 days left	Low	Medium(26-19090)
20	12520.0	gmc-door	2017	clean vehicle	30114	white	california usa	20 hours left	Low	Medium(26-19090)

Questions

5- We want to know the effect of the price by color

The effect of the price by color

CONCLUSIONS