EDA Used Cars Dataset

Presented by: Rehab AL zaidi

Contents
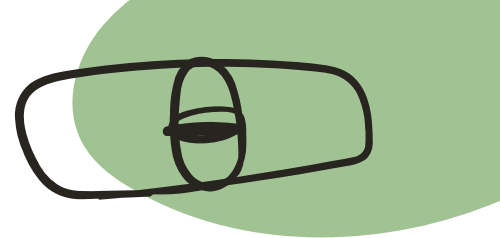
Introduction

When purchasing a car, there are many factors that go into making that decision. It's important that the car fits into the customer's budget as well as their lifestyle.

While purchasing a pre-owned vehicle is often a more economical option, the buyer may also consider the car's mileage, condition, manufacturer, etc. Using a dataset containing information about vehicles listed on Craigslist.

Dataset Description

Feature	Description
id	Vehicle Identification Number
region_url	Link to region page
region	Craigslist region in which this listing was posted
url	Link to listing
price	Price of vehicle
year	Year of vehicle
manufacturer	Manufacturer of vehicle
model	Model of vehicle

Dataset Description

Feature	Description
cylinders	Number of cylinders of vehicle
size	Size of vehicle
type	Vehicle type
paint_color	Color of vehicle
image_url	Link to image of vehicle
description	Listing description provided by owner
county	
state	State of Vehicle
lat	Latitude of vehicle (not precise but very close)

Dataset Description

| Feature	Description
drive	Drive of vehicle
condition	Condition of vehicle
fuel	taken by vehicle
VIN	Vehicle Identification Number
odometer	Vehicle type
paint_color	Color of vehicle
image_url	Link to image of vehicle
long	Longitude of vehicle (not precise but very close)
posting_date	Date posting of Vehicle

Preprocessing

01 Read Data

04 Visualization
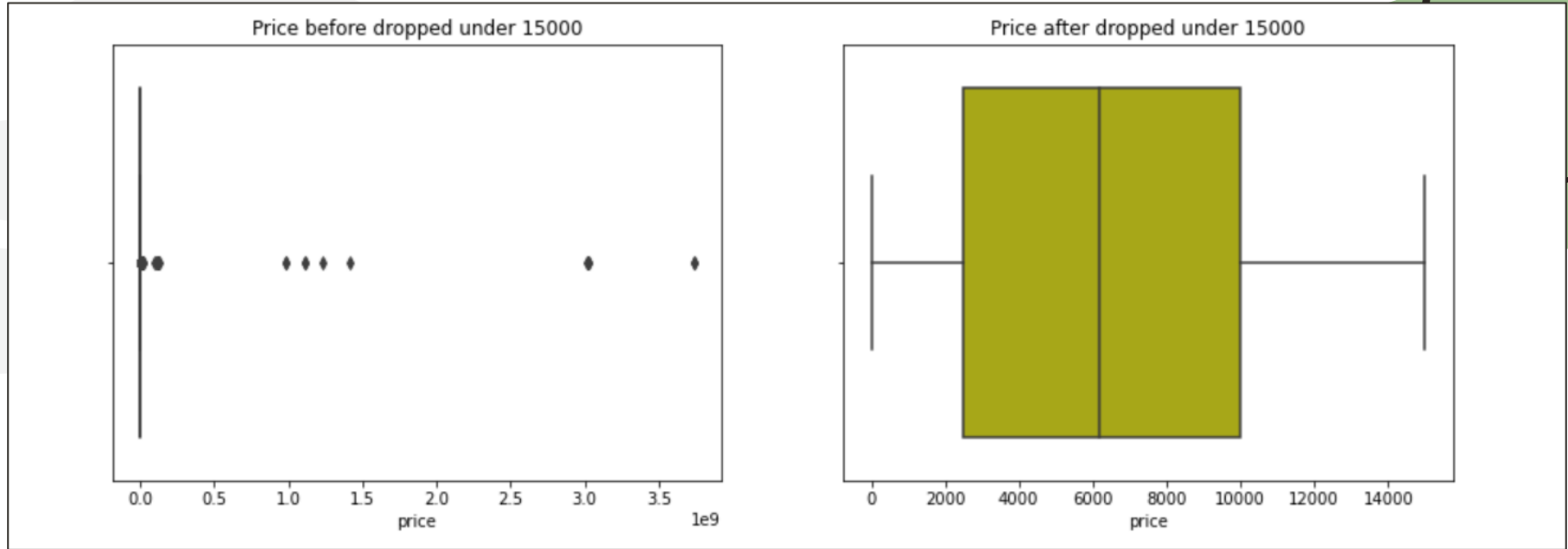
02 Check if there is missing values

03 Transformation

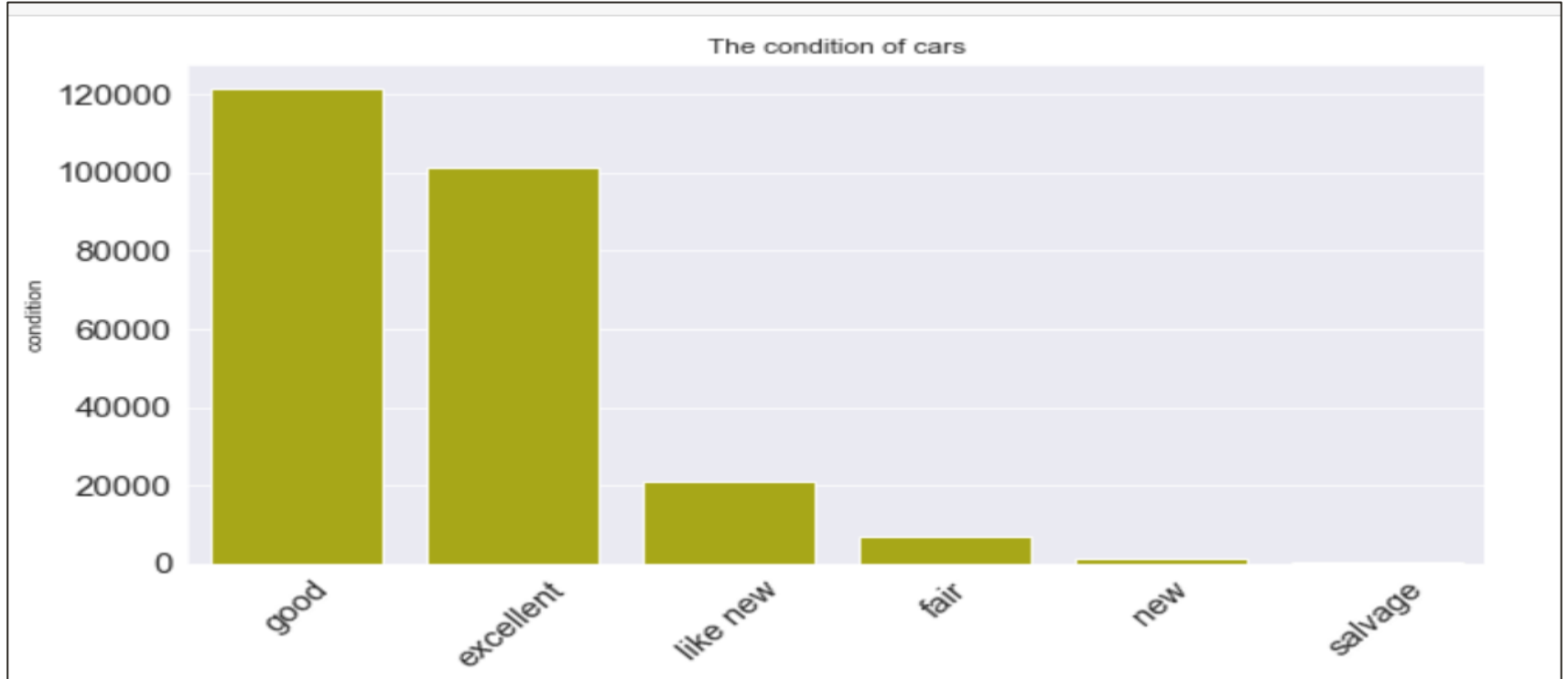outlier detection
in numerical
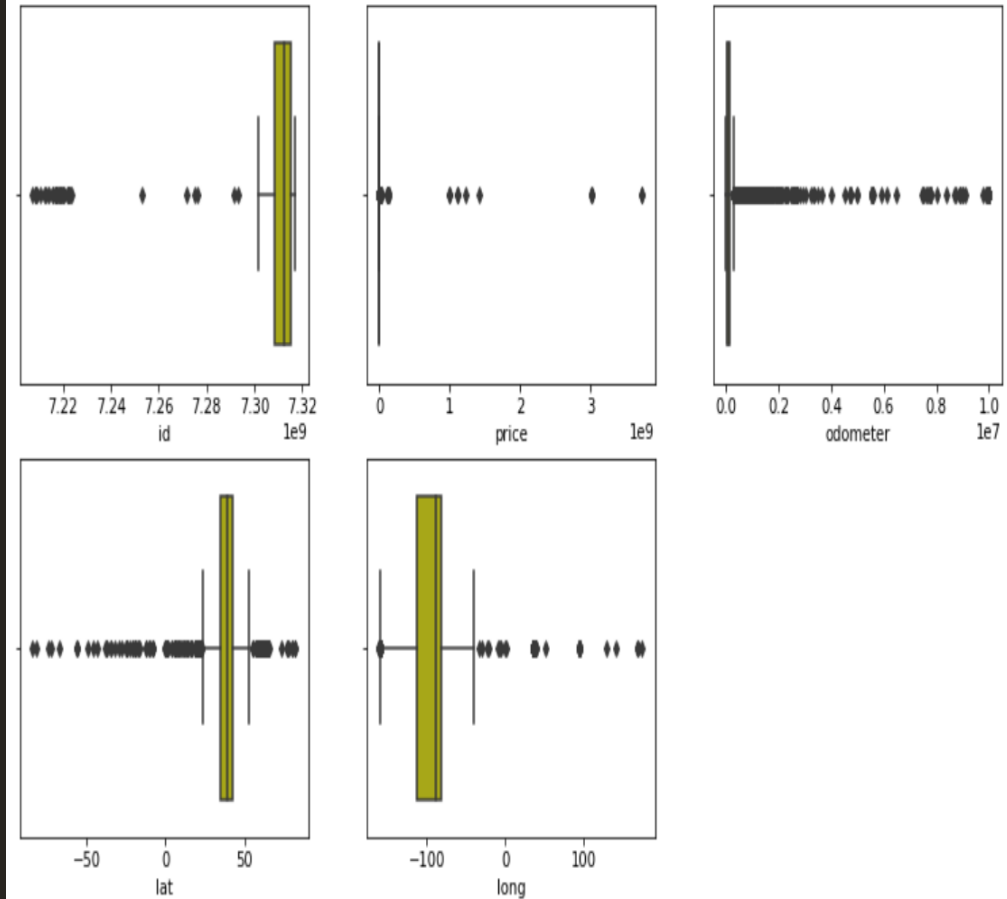
Based on boxplot visualization for numeric feature. We know that variable long, lat, and odomater has too many outliers, so that we can drop it from the table. Moreover we can see that price has so many outlier too, we can manipulate that variable.
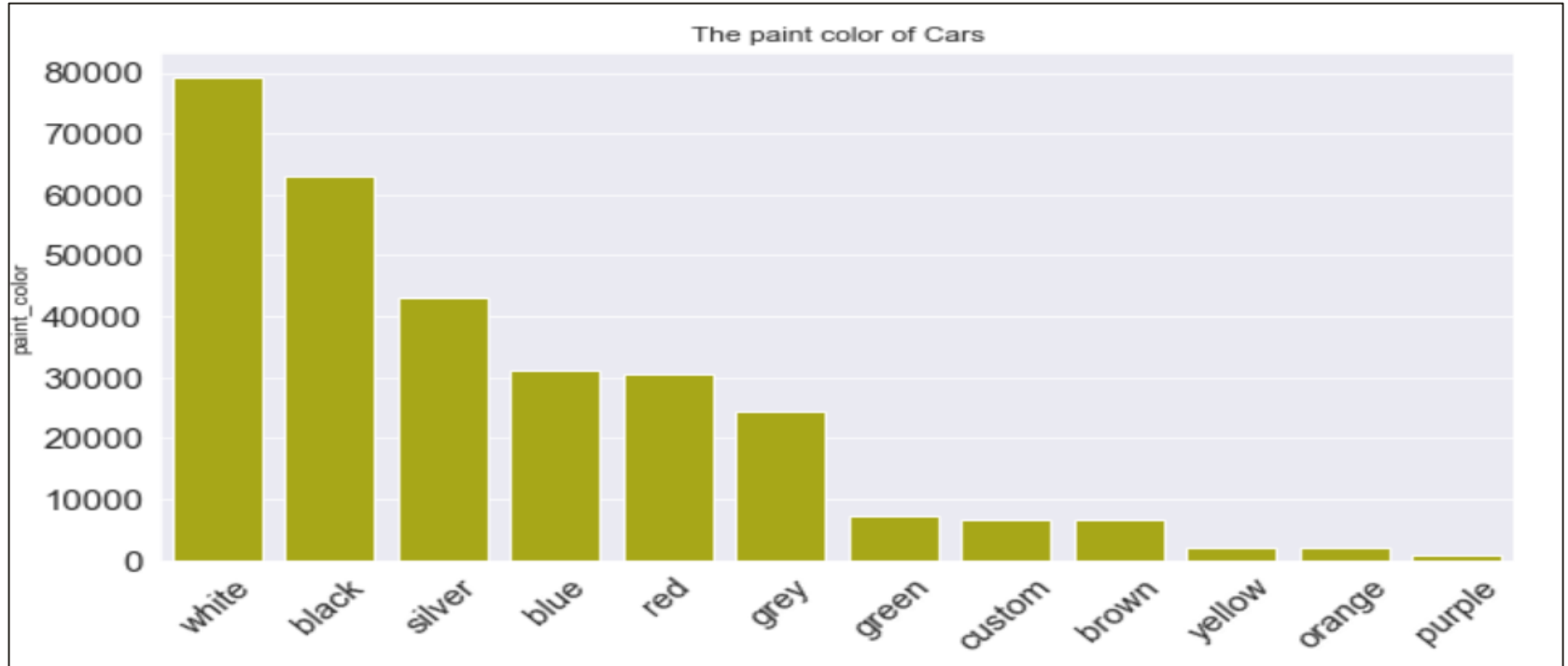
For Categorical:
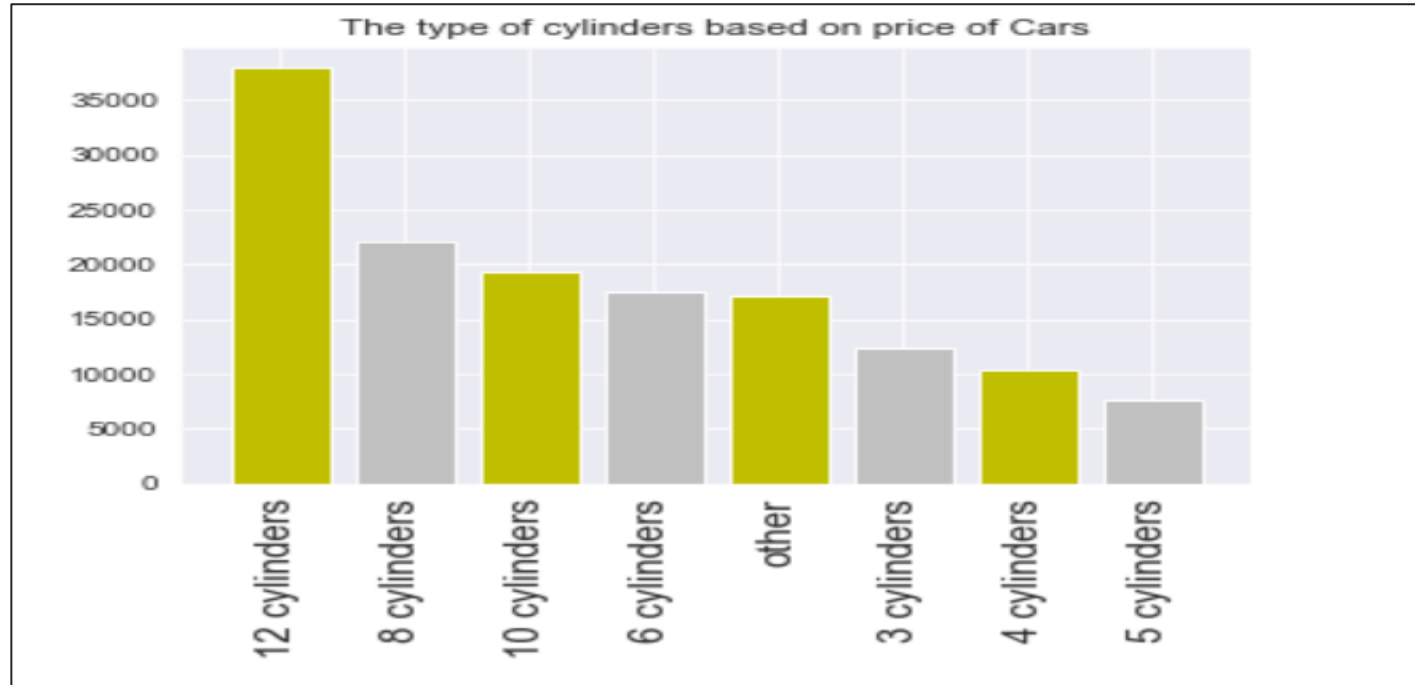
1- The condition of cars.
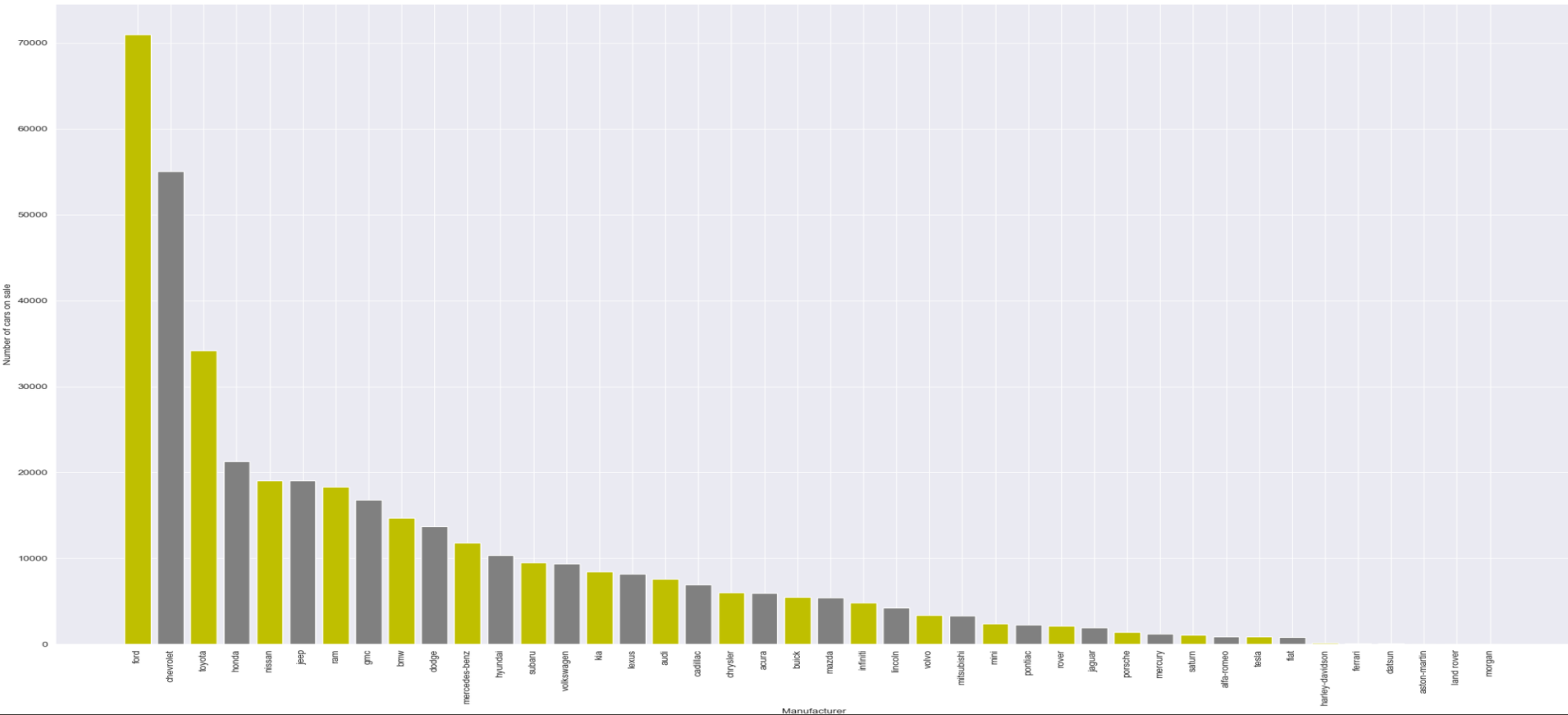
For Categorical:

2- The paint color of cars.

The paint color of Cars

For Categorical:
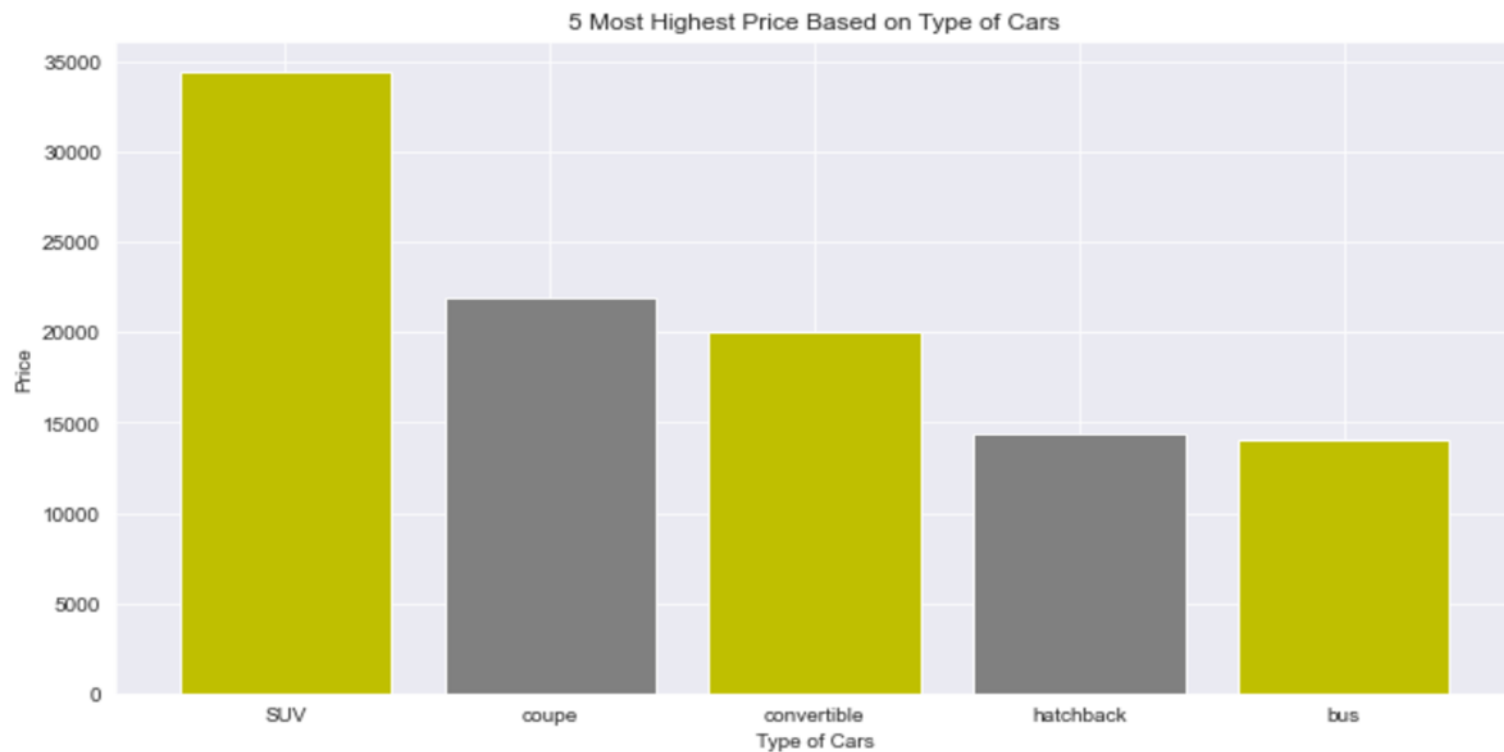
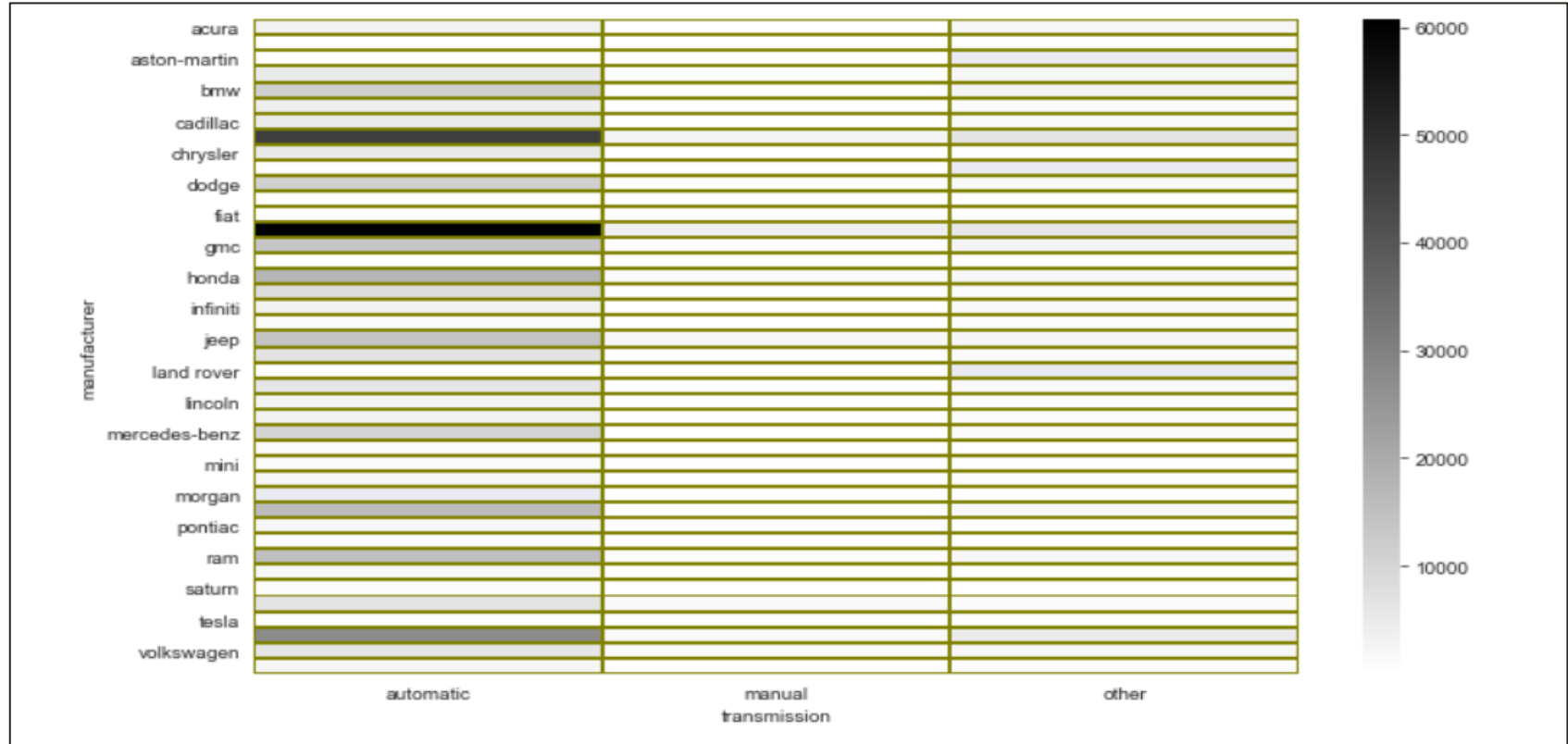3- The cylinders of cars.

The type of cylinders based on price of Cars

1- The most common manufacturers in dataset based on number of cars?

2-The Highest Price Based on Type Car?

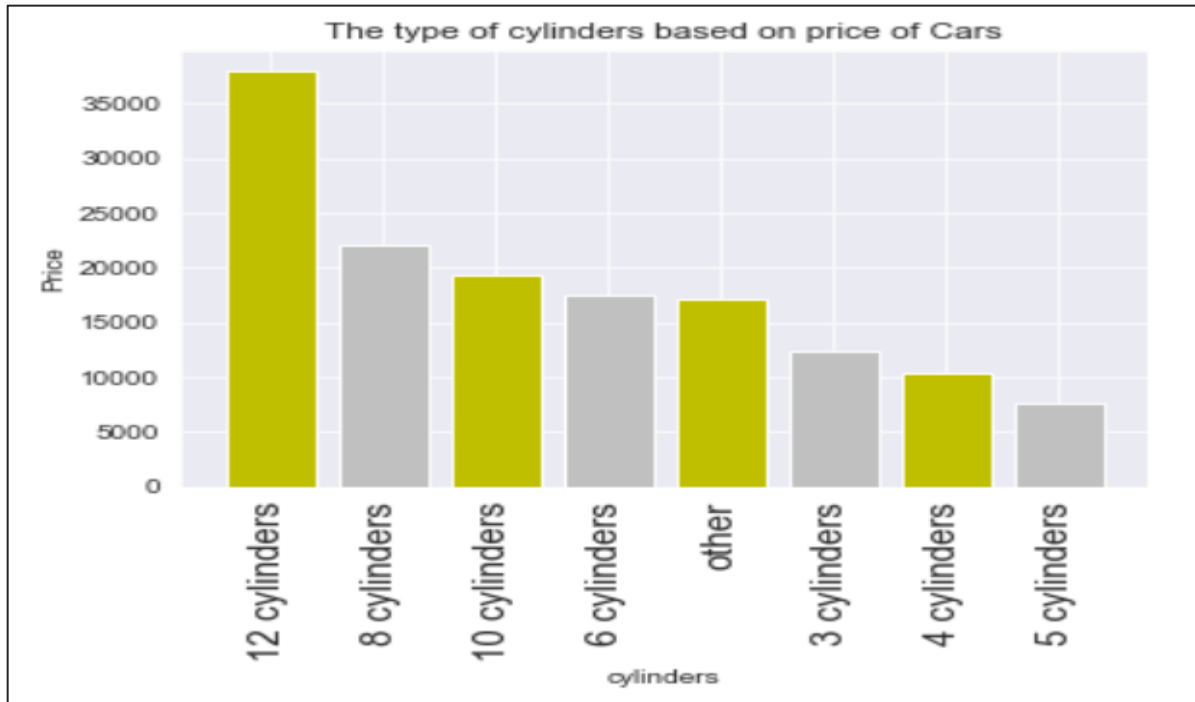5 Most Highest Price Based on Type of Cars

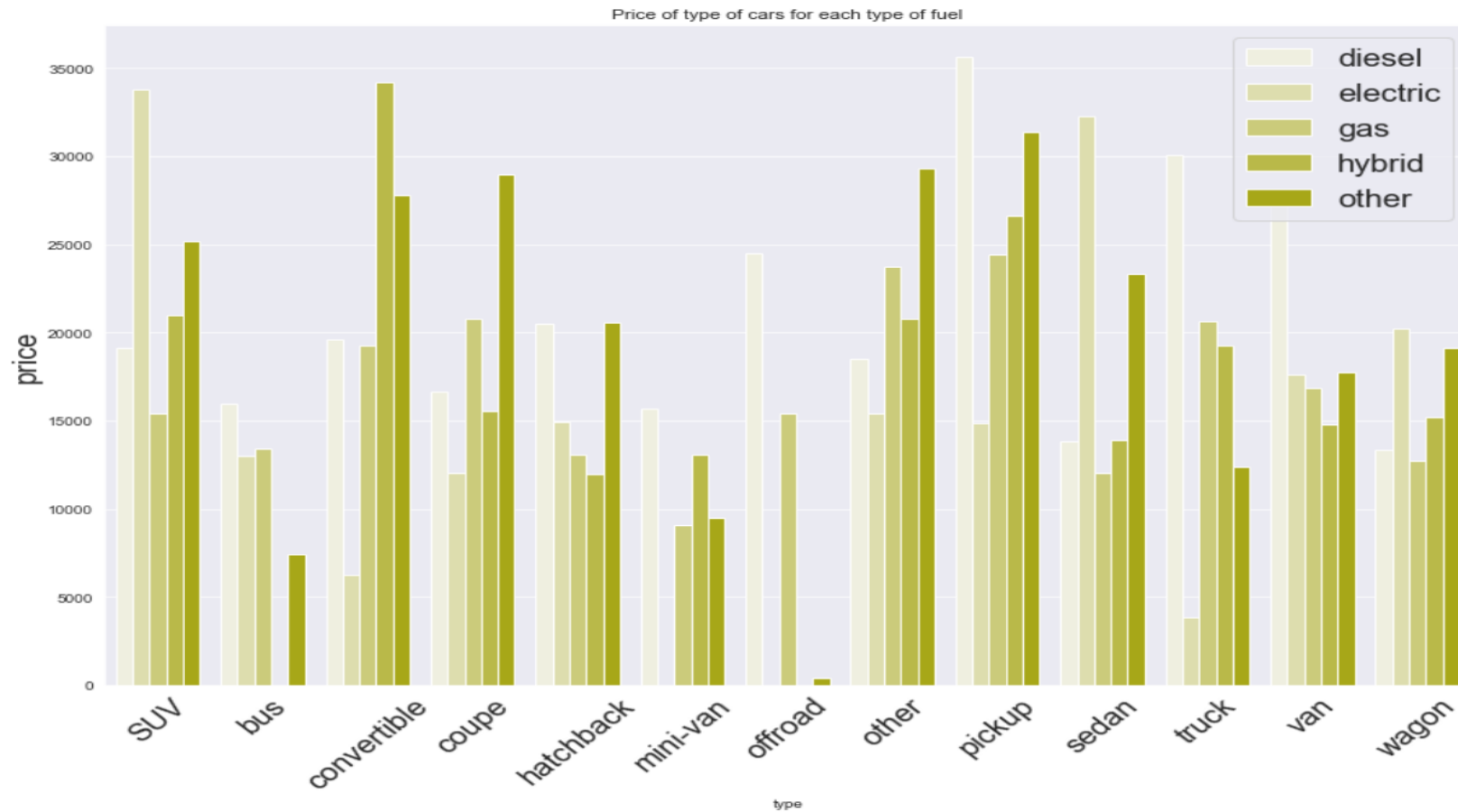# 3-The most transmission type used based on manufacturer?

It was no surprise that automatic used cars dominated the used cars market.

- The type of cylinders based on price of Cars.

- Price of type of cars for each type of fuel

Price of type of cars for each type of fuel

CONCLUSIONS

If you wanted to buy a certain type of used car, chart above could help you to decide what manufacturer to go for or type of cylinders .

So, those are some information we can obtain by doing EDA. There are much more insight from that one dataset if we want to dig deeper. Some code mistakes might be there, but you get the idea.

Thanks...