



US Cars Dataset

08.Oct.2022

Brian Yoo

Executive Summary

This dataset shows the sales information of used cars uploaded on a website. The website lets users explore key information that customers should consider when buying a car at the used car market. This dataset has information on car's detailed information, such as price, color, brand, model, location, condition, mileage, year, vin, and lot. This analysis will give sellers and buyers more detailed insights of the used car market and future trends.

Data Source

Kaggle link

Limitations and Ethics

There is none of the columns that shows customers' activity

There is no way I can check if the cars with 0 prices are the real existing price or just missing data.

Data Cleaning and Data Consistency Check

Check for mixed data type

Check for duplicated data

Check for missing data

Replace column

Drop irrelevant column

Data Profile

Raw Dataset: 2499 x 13

Cleaned Dataset: 2456 x 9

Questions

Cars each state have

The distribution of mileage

The relationship of price and mileage

Car price by the year

The highest price of the car by each brand

Column Details

| Column | Column Description | Data Type | Time Variant |
|--------------|--|--------------------------|--------------|
| Unnamed | | | No |
| Price | The sale price of the vehicle in the ad | Quantitative, Continuous | No |
| Brand | The brand of car | Qualitative, Ordinal | No |
| Model | model of the vehicle | Qualitative, Nominal | No |
| Year | The vehicle registration year | Quantitative, Discrete | No |
| Title Status | This feature included binary classification, which are clean title vehicles and salvage insurance | Qualitative, Ordinal | No |
| Mileage | miles traveled by vehicle | Quantitative, Continuous | No |
| Color | Color of the vehicle | Qualitative, Nominal | No |
| Vin | The vehicle identification number is a collection of 17 characters (digits and capital letters) | Qualitative, Nominal | No |
| Lot | A lot number is an identification number assigned to a particular quantity or lot of material from a single manufacturer. For cars, a lot number is combined with a serial number to form the Vehicle Identification Number. | Qualitative, Nominal | No |
| State | The location in which the car is being available for purchase | Qualitative, Nominal | No |
| Country | The location in which the car is being available for | Qualitative, Nominal | No |

| | | | |
|-----------|----------------|-----------------------------|-----|
| | purchase | | |
| Condition | Remaining time | Quantitative, Continuous | Yes |