Pushin Python Data Summary Report

Janaya Johnson

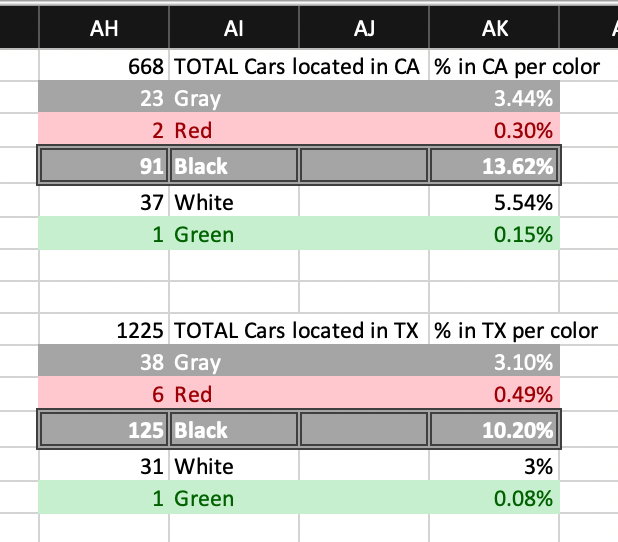
Yan Cai

Joshua Sehwani

The dataset we chose for our group project was car sales. So far, we have worked with python and SQL to explore the dataset. As we have learned new tools in the last week, we will be exploring the dataset more with Tableau. To explore the dataset more, we chose to work on exploring what is most popular within each state. For the purpose of this assignment, we chose to compare the purchase state with the maximum miles per gallon, engine type, fuel type, drive train, exterior car color, and the purchase price. The following is what we found:

We chose to analyze the states and what type of drivetrain was the most and least popular. We found that all-wheel drive was the most popular between all the states. Texas had the most cars bought with all-wheel drive. Of the 1225 cars that were bought in Texas, 547 of the cars chosen were all-wheel drive. However, the highest percentage of cars bought with all-wheel drive from one state was Wyoming, with 100% of the cars sold were bought with all-wheel drive. The smallest number of all-wheel drive cars purchased was from New Mexico, with 0% of the cars sold bought with rear-wheel drive. The least popular car by drivetrain was rear-wheel drive. Texas, again, has the highest number of buyers of rear-wheel drive. Montana had the highest number of cars with rear-wheel drive.

We chose to analyze the states and what type of fuel was the most and least popular. We found that gasoline was the most popular between all the states. Texas had the most cars with the gasoline fuel type. Delaware, Hawaii, Maine, Mississippi, North Dakota, New Mexico, and Vermont each had 100% of their car sales with gasoline powered cars. Flex fuel was the least popular car bought. One car was bought in each state: California, Illinois, Massachusetts, Pennsylvania, and Texas.

 Figure 1

We chose to analyze the most popular mpg indicator in this dataset. There are 9379 cars in this dataset. According to figure 2 we found the most popular mpg is 26, and there are 1078 of them in this data set. This is the largest number for a mpg indicator group. It takes 11.49% of total cars.

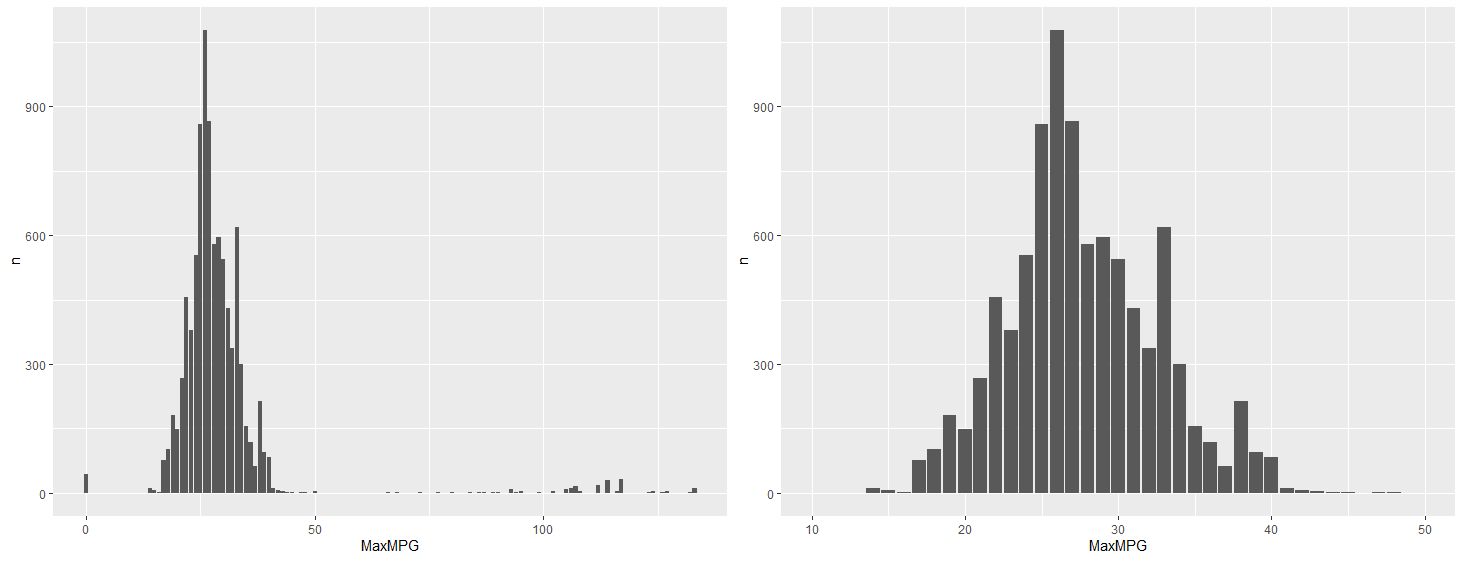


Figure 2

We chose to analyze the transmission preference of the fuel economy car (mpg at 30 or above). We realize most cars with 30 or more mpg use automatic transmission. That made us curious about the proportionate level of automatic transmission in the fuel economy car. After analysis, we found the result in figure 3, in which there are 3174 cars that have 30 mpg or more and 3080 of them use automatic transmission. It takes 97% of total cars. Based on this observation, we can conclude the automatic transmission is very popular for fuel economy cars.Figure 3

Within Figure 1, we are able to take different columns that separate the data and calculate different metrics based on the criteria within the columns. First we take the total number of cars that are sold within a particular state, in this case we chose California and Texas. The number of vehicles sold with a particular exterior color versus other exterior colors was the desired outcome with the colors being gray, red, black, white, and green. Simply we are able to find the percentage of each exterior color and how popular that color is among other colors within CA and TX. (note the percentages do not add up to 100% because there are numerous niche colors that were not being analyzed.)

Our findings show that a black exterior color dominated the market within the colors we analyzed with white and gray in the second and third spot respectively in both California and Texas. There were also similarly low percentages for red and green exterior color cars in both states as both colors struggled to gain even one percent market share.