DATA VISUALIZATION USING TABLEAU

ASSIGNMENT 3

OCTOBER 9, 2022.

INDIANA UNIVERSITY
(BLOOMINGTON)

Authored by: VARSHA R
IU ID: 2000751388
Sem: Fall 2022

Introduction

Data visualization is achieved using Tableau, a software package focusing on business intelligence (BI). The word tableau refers to a graphic representation or description. As a novice with Tableau, I have experimented with the software and built a few visualizations. The purpose of this report is to discuss these visualizations. For each graph, the following information will be discussed:

- Plots built using the dataset.
- Plot analysis.
- Inferences drawn from the visualizations.

Dataset

> The UCI Machine Leaning Repository is a collection of databases, domain theories, and data generators that are used by the machine learning community for the empirical analysis of machine learning algorithms. - UCI

For the visualizations, we will use the following dataset from the UCI Machine Learning Repository:

Automobile data set: This data set consists of the specification of an auto in terms of various characteristics. Dataset includes various attributes that are listed below:

make: alfa-romero, audi, bmw, chevrolet, dodge, honda, isuzu, jaguar, mazda, mercedes-benz, mercury,, mitsubishi, nissan, peugot, plymouth, porsche, renault, saab, subaru, toyota, volkswagen, volvo
fuel-type: diesel, gas.
aspiration: std, turbo.
num-of-doors: four, two.
body-style: hardtop, wagon, sedan, hatchback, convertible.
engine-location: front, rear.
length: continuous from 141.1 to 208.1.
width: continuous from 60.3 to 72.3.
height: continuous from 47.8 to 59.8.
num-of-cylinders: eight, five, four, six, three, twelve, two.
horsepower: continuous from 48 to 288.
city-mpg: continuous from 13 to 49.
highway-mpg: continuous from 16 to 54.
price: continuous from 5118 to 45400.

Calculated fields:

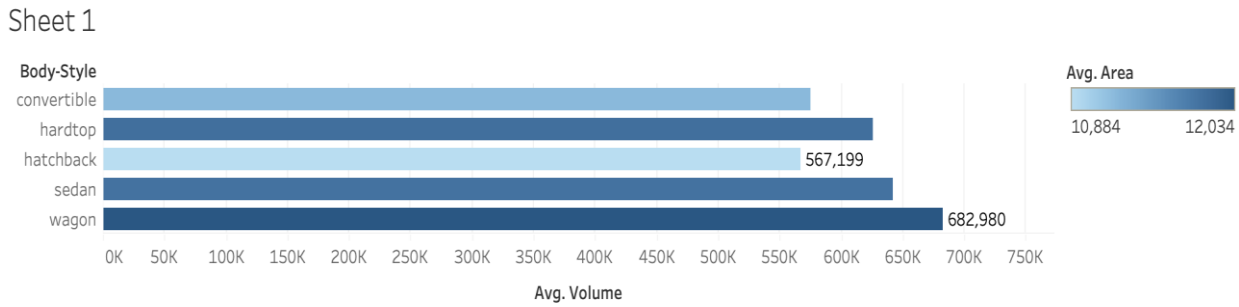
The following calculated fields are used for visualizations in this assignment. They are as follows:

- **Area** = The floor area that the automobile takes up. It is calculated using the dataset attributes as length X width.
- **Volume** = The total cubical space that the automobile takes up. It is calculated using the dataset attributes as length X width X height.
- **Highway vs city mpg** = The difference between mpgs of the automobile in highway and city. It is calculated using the dataset attributes as highway-mpg – city-mpg.

Visualizations

1)

○ Plot built using this dataset



Average of Volume for each Body-Style. Color shows average of Area.

○ Plot analysis

Attributes used for this plot: Body-style, volume, area
Columns: AVG(volume)
Rows: Body-style
Marks: color – AVG(Area)

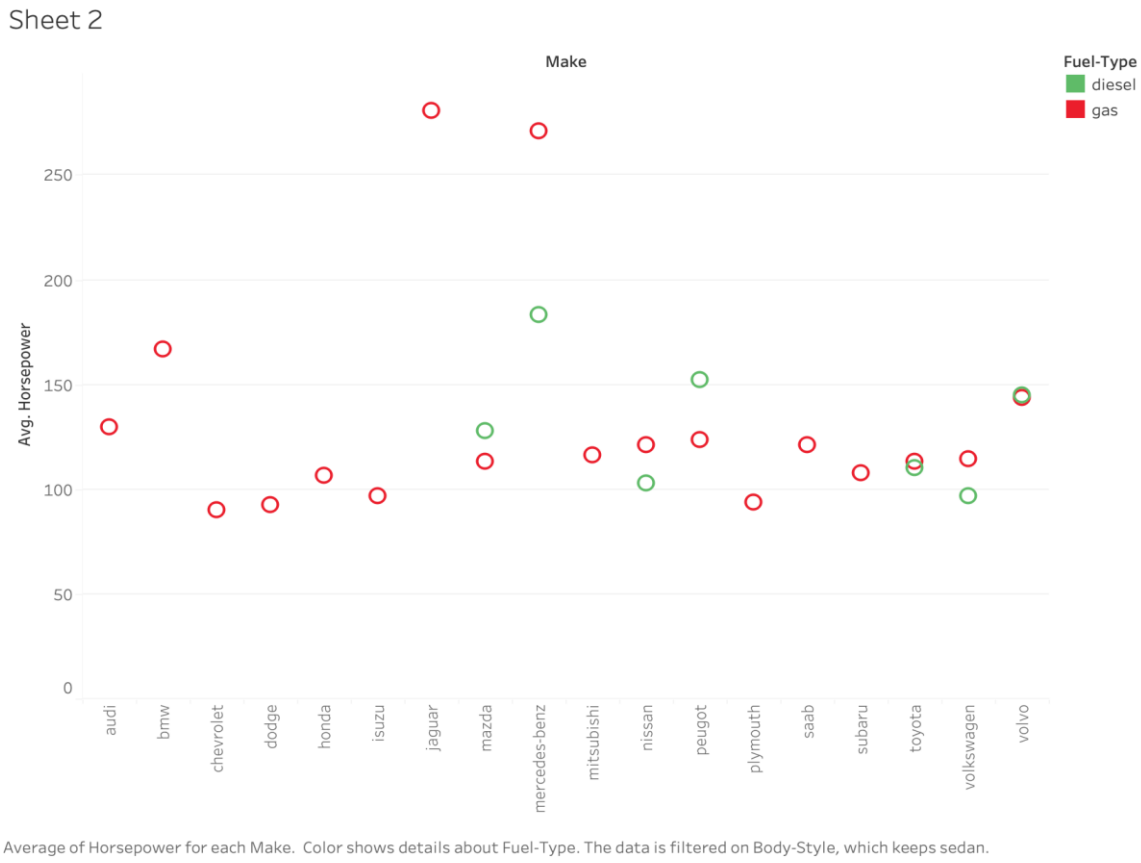
Here we are trying to compare and see which body-style occupies the most space.

○ Inferences drawn from the visualization

- On an average, wagons occupy the most area and has highest volumes which shows that wagons would require more parking space.
- On an average, hatchbacks occupy the least space.

2)

- **Plot built using this dataset**



Average of Horsepower for each Make. Color shows details about Fuel-Type. The data is filtered on Body-Style, which keeps sedan.

- **Plot analysis**

Attributes used for this plot: Make, Horsepower, Body-style, Fuel-type.
Columns: Make
Rows: AVG(Horsepower)
Marks: color – Fuel-type
Filters: Body-style - Sedan

Here we are trying to compare and see which automobile make that are specifically sedan body-style has the highest horsepower on overall that runs on gas.

- **Inferences drawn from the visualization**

- We can see that the sedans that run on gas has comparatively more horsepower than diesel sedan makes.
- A sedan that has the highest horsepower is made my jaguar that runs on gas and has a horsepower of 280.7.

3)

- ## Plot built using this dataset

Sheet 3

Make

Body-Style	alfa-rom..	audi	bmw	chevrolet	dodge	honda	isuzu	jaguar	mazda	mercede..	mercury	mitsubis..	nissan	peugot	plymouth	porsche	renault	saab	subaru	toyota	volkswa..	volvo
convertible	14,998									35,056						37,028				17,669	11,595	
hardtop										36,788			8,249			33,278				9,762		
hatchback	16,500			5,723	7,820	7,054	11,048		10,085		16,503	9,598	14,409		8,131	22,018	9,895	15,013	6,591	9,616	9,980	
sedan		17,647	26,119	6,575	7,620	9,945	6,785	34,600	11,464	33,074		8,434	8,605	15,759	7,151			15,433	9,071	9,542	9,674	18,727
wagon		18,920			8,921	7,295				28,248			9,916	15,018	8,921		9,295		9,342	9,836	12,290	16,293

Avg. Price: 5,723 – 37,028

Average of Price broken down by Make vs. Body-Style. Color shows average of Price. The marks are labeled by average of Price. The view is filtered on Exclusions (Body-Style,Make), which keeps 55 members.

Zoomed in image of the above visualization that contains the results:

Make

Body-Style	alfa-rom..	audi	bmw	chevrolet	dodge	honda	isuzu	jaguar	mazda	mercede..	mercury	mitsubis..	nissan	peugot	plymouth	porsche
convertible	14,998									35,056						37,028
hardtop										36,788			8,249			33,278
hatchback	16,500			5,723	7,820	7,054	11,048		10,085		16,503	9,598	14,409		8,131	22,018
sedan		17,647	26,119	6,575	7,620	9,945	6,785	34,600	11,464	33,074		8,434	8,605	15,759	7,151	
wagon		18,920			8,921	7,295				28,248			9,916	15,018	8,921	

- ## Plot analysis

Attributes used for this plot: Make, Horsepower, Body-style, Price
Columns: Make
Rows: Body-style
Marks: Color – AVG(Price), Label – AVG(Price)
Filters: Body-style that has no price value (null price value removed)

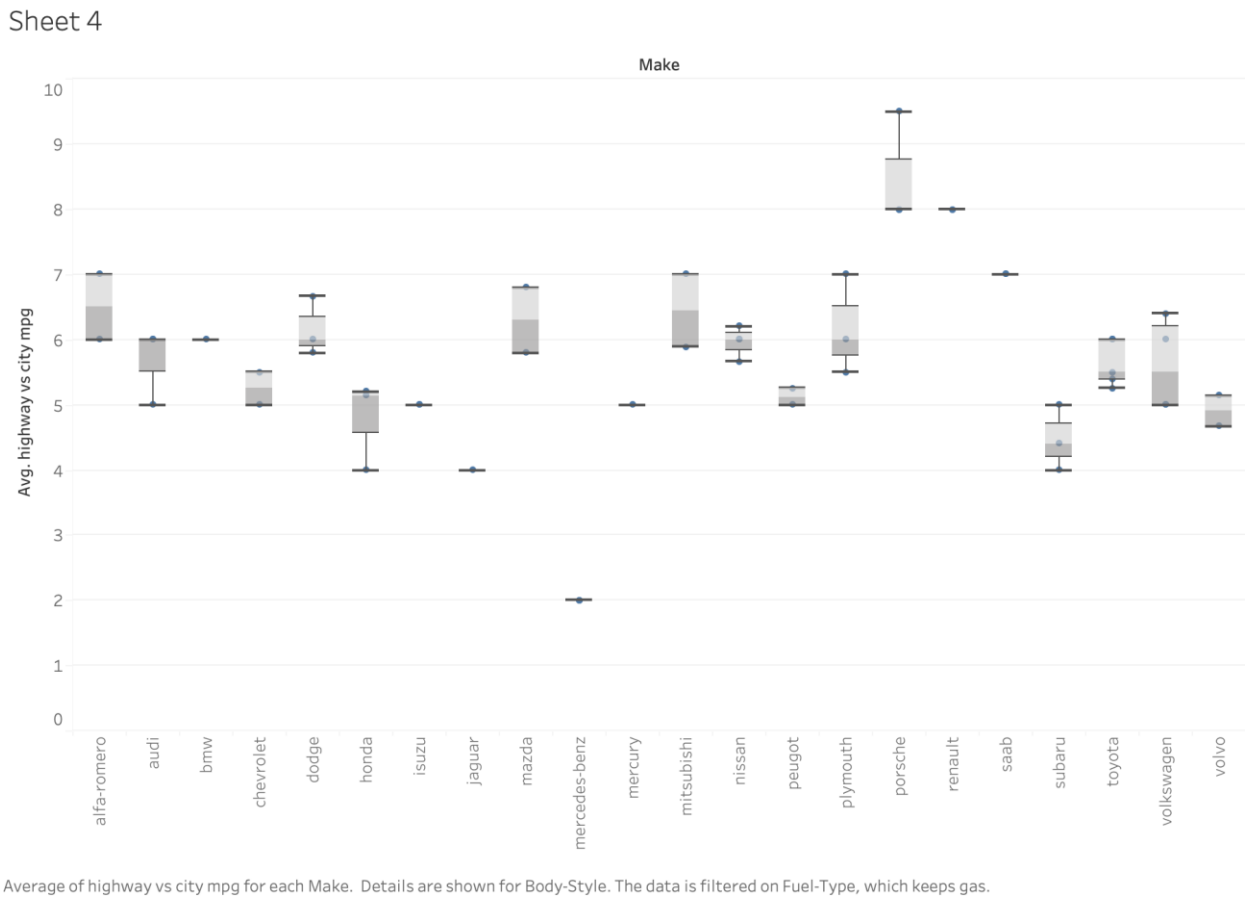
Here we are trying to compare and see which automobile make and body-style is the most expensive.

- ## Inferences drawn from the visualization

- The most expensive make is Porsche.
- The Most expensive body-shape of the Porsche is a convertible.
- The average price of this convertible Porsche is $37,028.

4)

- ## Plot built using this dataset



Average of highway vs city mpg for each Make. Details are shown for Body-Style. The data is filtered on Fuel-Type, which keeps gas.

- ## Plot analysis

Attributes used for this plot: Make, highway vs city mpg, Fuel-type, Body-style.
Columns: Make
Rows: AVG(highway vs city mpg)
Marks: Details – Body-style
Filters: Fuel-type

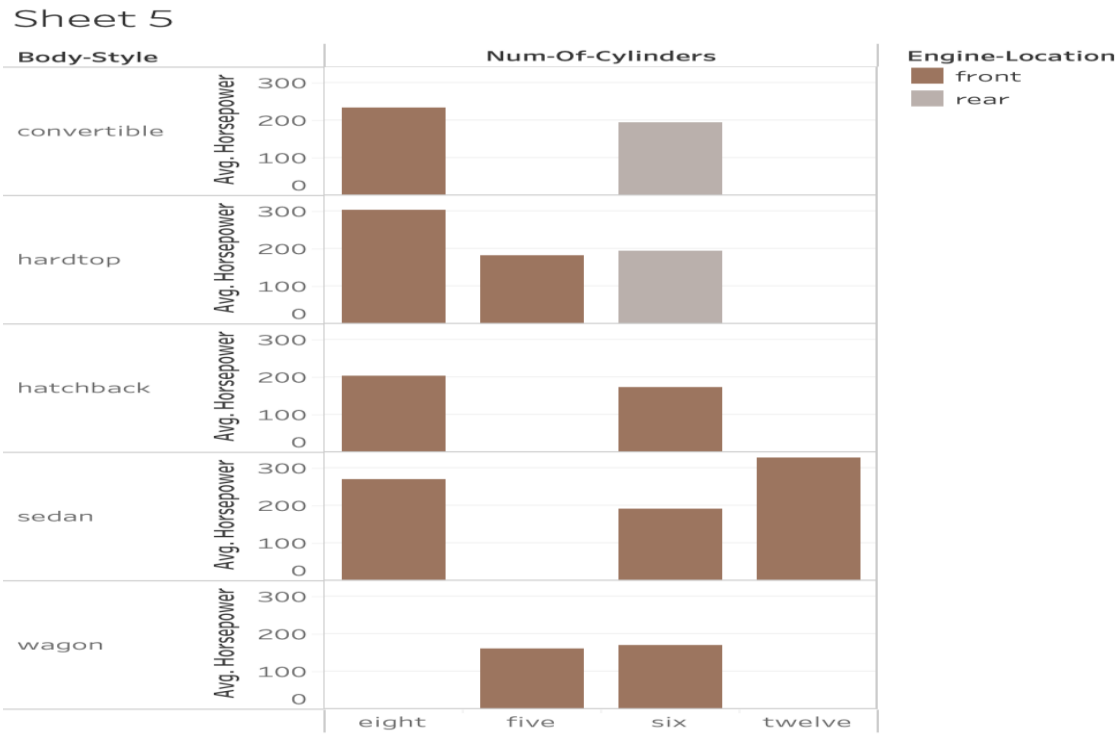
Here we are trying to compare and see which automobile make that runs on gas has the highest mpg difference between highways and cities.

- ## Inferences drawn from the visualization

- Porsche hatchback has the highest average difference of mpg in highways and cities.
- Most of the other makes' difference value ranges mostly from 4-7 mpg.

5)

- ## Plot built using this dataset



Average of Horsepower for each Num-Of-Cylinders broken down by Body-Style. Color shows details about Engine-Location. The view is filtered on average of Horsepower, which includes values greater than or equal to 150.0.

- ## Plot analysis

Attributes used for this plot: Body-style, Num-of-cylinders, Horsepower, Engine location
Columns: Num-of-cylinders
Rows: Body-style, AVG(Horsepower)
Marks: Color – Engine location
Filters: AVG(Horsepower)

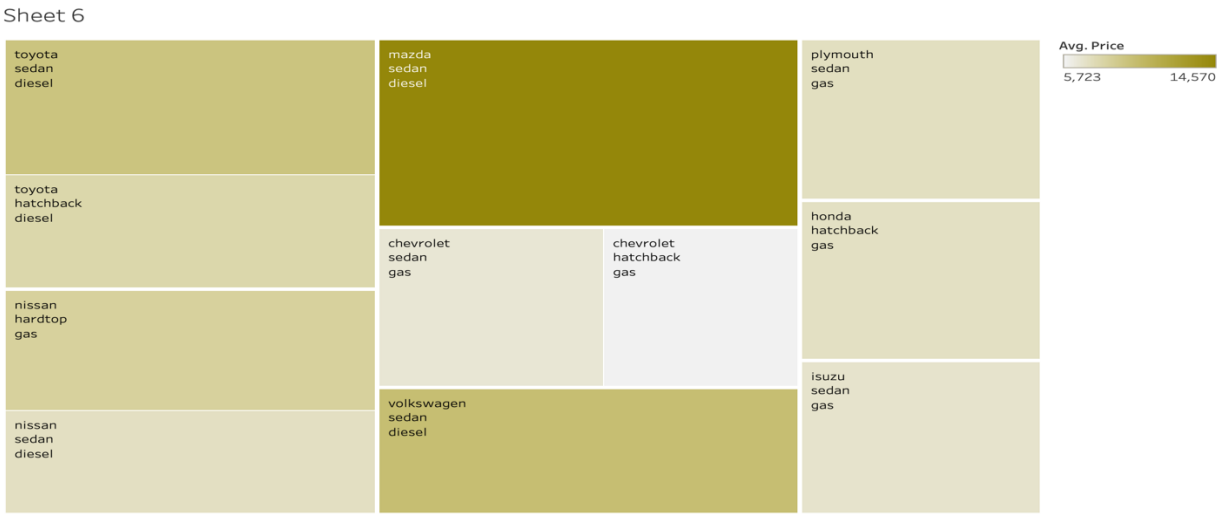
Here we are trying to compare and see what the average horsepower value that is above 150 and is for different body-styles with different number of cylinders.

- ## Inferences drawn from the visualization

- On an average, 12 cylinder sedans have that have the engines in the front have the most horsepower.
- Only hardtops and convertibles have automobiles that have the engines placed in the rear and has more than 150 horsepower. Coincidently, these also have six exact cylinders as another commonality.

6)

- **Plot built using this dataset**



Make, Body-Style and Fuel-Type. Color shows average of Price. Size shows average of Price. The marks are labeled by Make, Body-Style and Fuel-Type. The data is filtered on average of City-Mpg, which includes values greater than or equal to 30. The view is filtered on Exclusions (Body-Style,Fuel-Type,Make), which keeps 65 members.

- **Plot analysis**

Attributes used for this plot: Body-style, Make, Fuel-type, City-mpg, Price
Marks: Color – AVG(Price), Size – AVG(Price), Labels – Make, Body-style, Fuel-type
Filters: AVG(City-mpg), Null values removed

Here we are trying to compare and see which make and its body-style has the highest price that gives a city-mpg of at least 30. The null values are also filtered out.

- **Inferences drawn from the visualization**

  - On an average, sedans made by Mazda that run on diesel are the most expensive ones.
  - If we look closely, we can see that the top 3 expensive ones in this chart are all run by diesel.

References

http://archive.ics.uci.edu/ml/index.php