Homework data viz

Sarayut Por

2023-12-06

```
library(tidyverse)
head(mpg)
```

```
## # A tibble: 6 x 11
##   manufacturer model displ  year   cyl trans      drv     cty   hwy fl    class
##   <chr>        <chr> <dbl> <int> <int> <chr>      <chr> <int> <int> <chr> <chr>
## 1 audi         a4      1.8  1999     4 auto(l5)   f        18    29 p     compa~
## 2 audi         a4      1.8  1999     4 manual(m5) f        21    29 p     compa~
## 3 audi         a4      2    2008     4 manual(m6) f        20    31 p     compa~
## 4 audi         a4      2    2008     4 auto(av)   f        21    30 p     compa~
## 5 audi         a4      2.8  1999     6 auto(l5)   f        16    26 p     compa~
## 6 audi         a4      2.8  1999     6 manual(m5) f        18    26 p     compa~
```

Chart 1: Displays the average fuel consumption level by vehicle type (cty)

- The chart shows how each type of car has the average fuel consumption in the city (cty).
- It can be observed that Some types of cars are more efficient. Other types of vehicles in the city, such as small cars, have higher average city fuel consumption than SUVs or off-road vehicles.

```
ggplot(data = mpg) +
  aes(x = class, y = cty, fill = class) +
  geom_bar(stat = "summary", fun = "mean") +
  labs(title = "Average City Mileage by Vehicle Class",
       x = "Vehicle Class",
       y = "Average City Mileage") +
  theme_minimal()
```

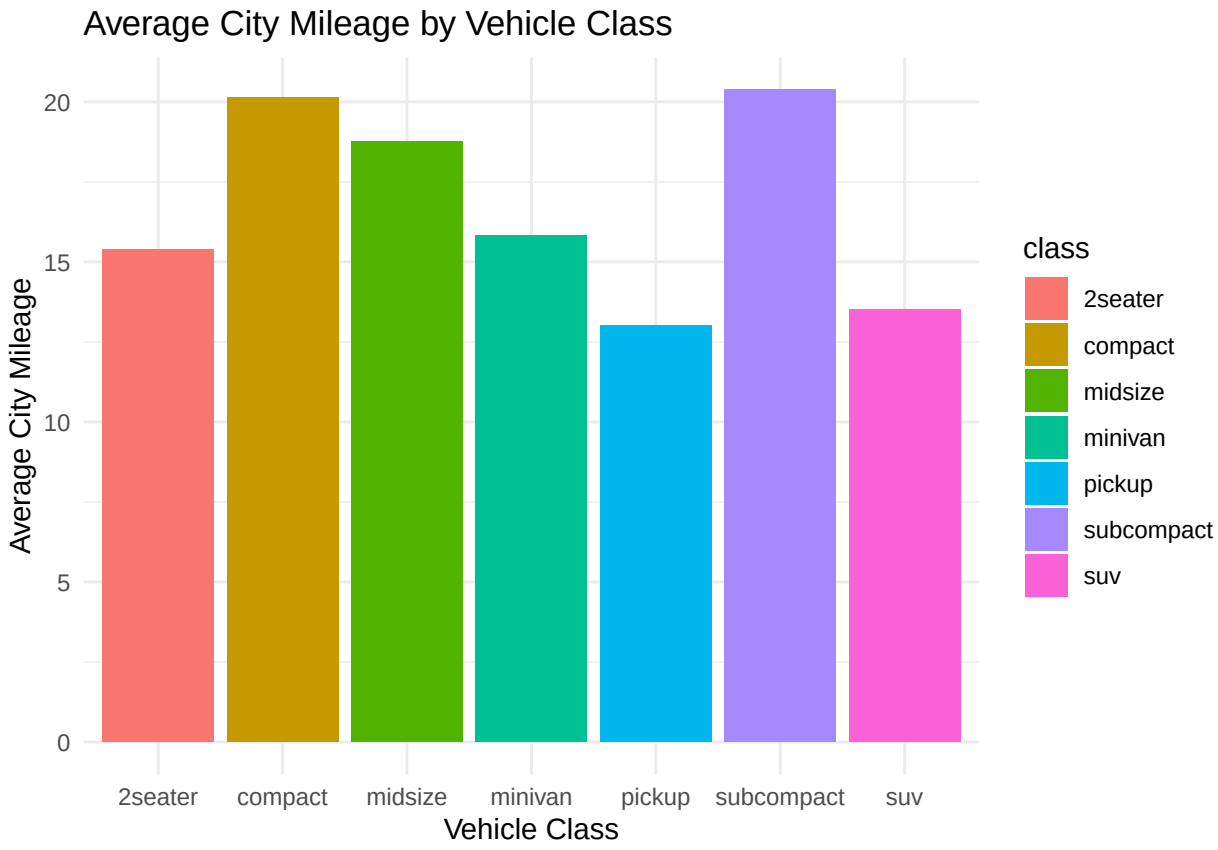


Chart 2: Shows the relationship between cty and hwy

- Show a positive relationship between city distance (cty) and highway (hwy).
- Cars that go far in the city tend to go further on the highway. Demonstrates consistency in the efficiency of oil use.

```
ggplot(data = mpg) +
  aes(x = cty, y = hwy) +
  geom_point(aes(color = class), alpha = 0.5) +
  geom_smooth(method = "lm", se = FALSE) +
  labs(title = "Relationship between City and Highway Mileage",
       x = "City Mileage",
       y = "Highway Mileage") +
  theme_minimal()
```

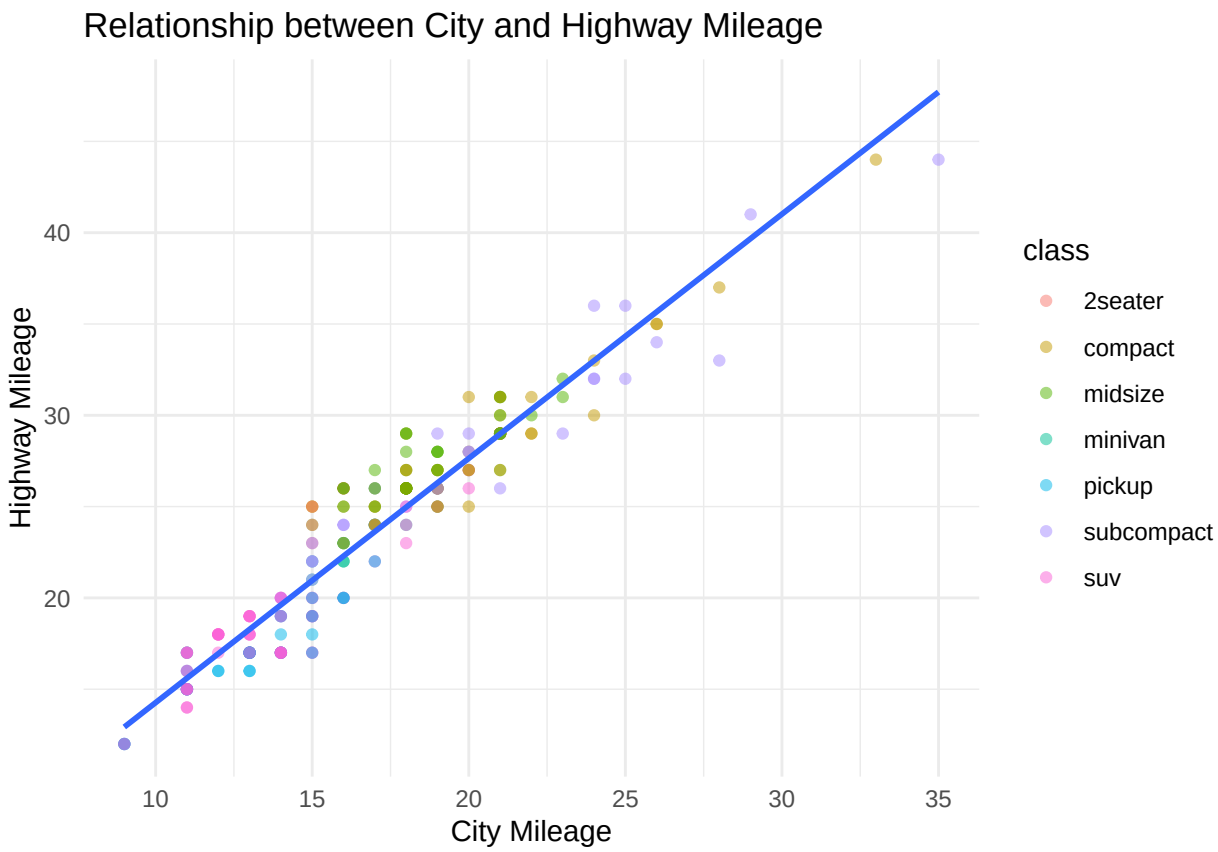


Chart 3: Shows the distribution of city distances for each vehicle type

- Boxplot helps highlight the distribution of city distance (cty) values for each vehicle type.
- It also helps highlight outliers. The vehicle categories with the lowest and highest values may be the focus for further analysis on their design or performance.

```
ggplot(data = mpg) +
  aes(x = class, y = cty) +
  geom_boxplot(aes(fill = class)) +
  labs(title = "City Mileage Distribution by Vehicle Class",
       x = "Vehicle Class",
       y = "City Mileage") +
  theme_minimal()
```

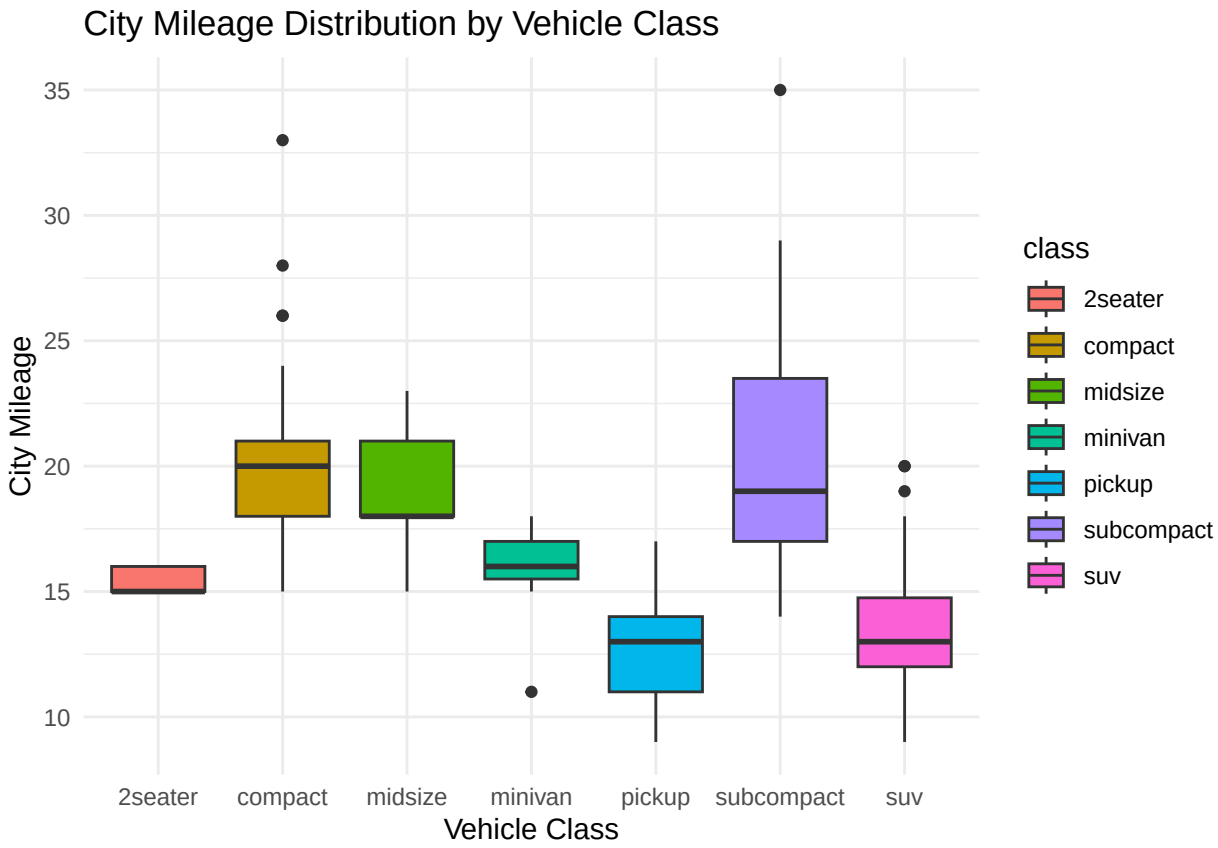


Chart 4: Shows the quantity of each type of vehicle according to the origin

- Shows the distribution of vehicle drive types in each vehicle category.
- It helps to see which types of cars are more or less by type of drive.

```
ggplot(data = mpg) +
  aes(x = class, fill = drv) +
  geom_bar(position = "stack") +
  labs(title = "Count of Vehicle Types by Drive Type",
       x = "Vehicle Class",
       y = "Count") +
  theme_minimal()
```

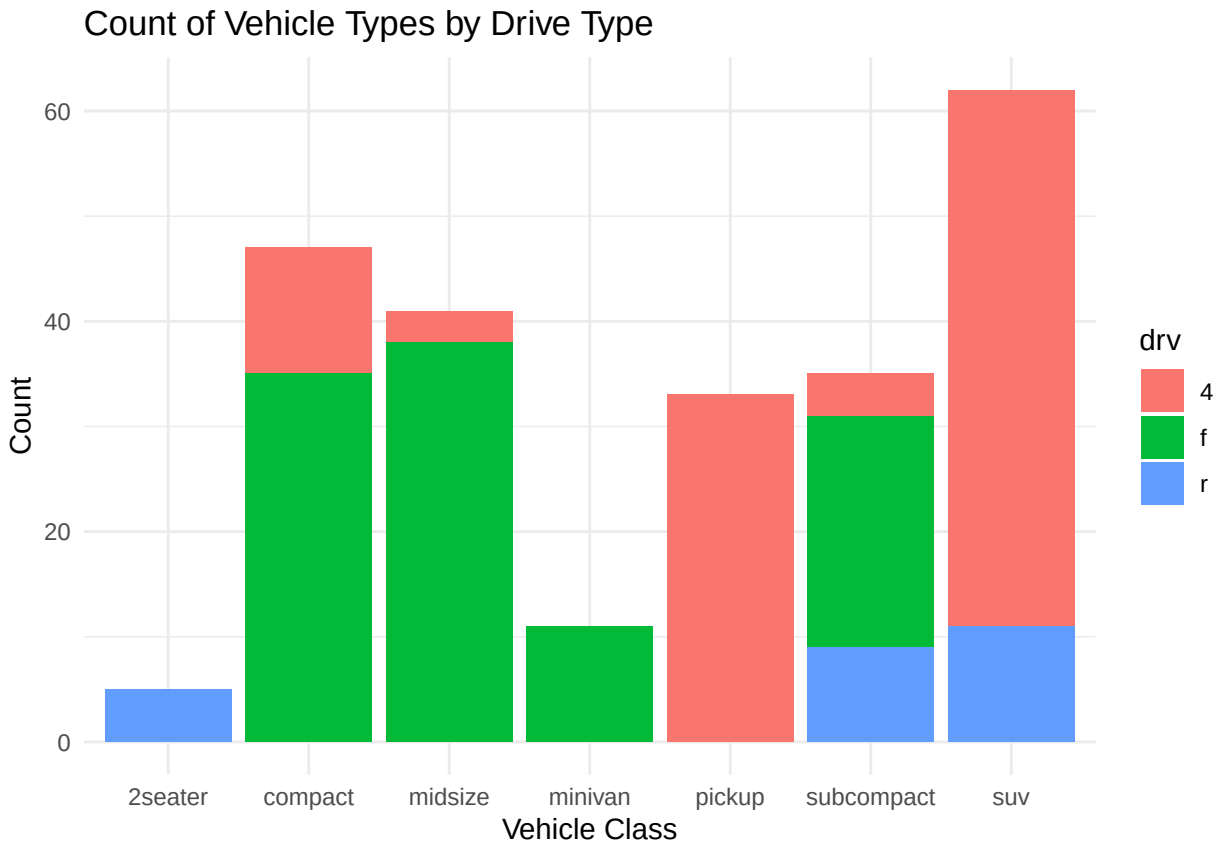
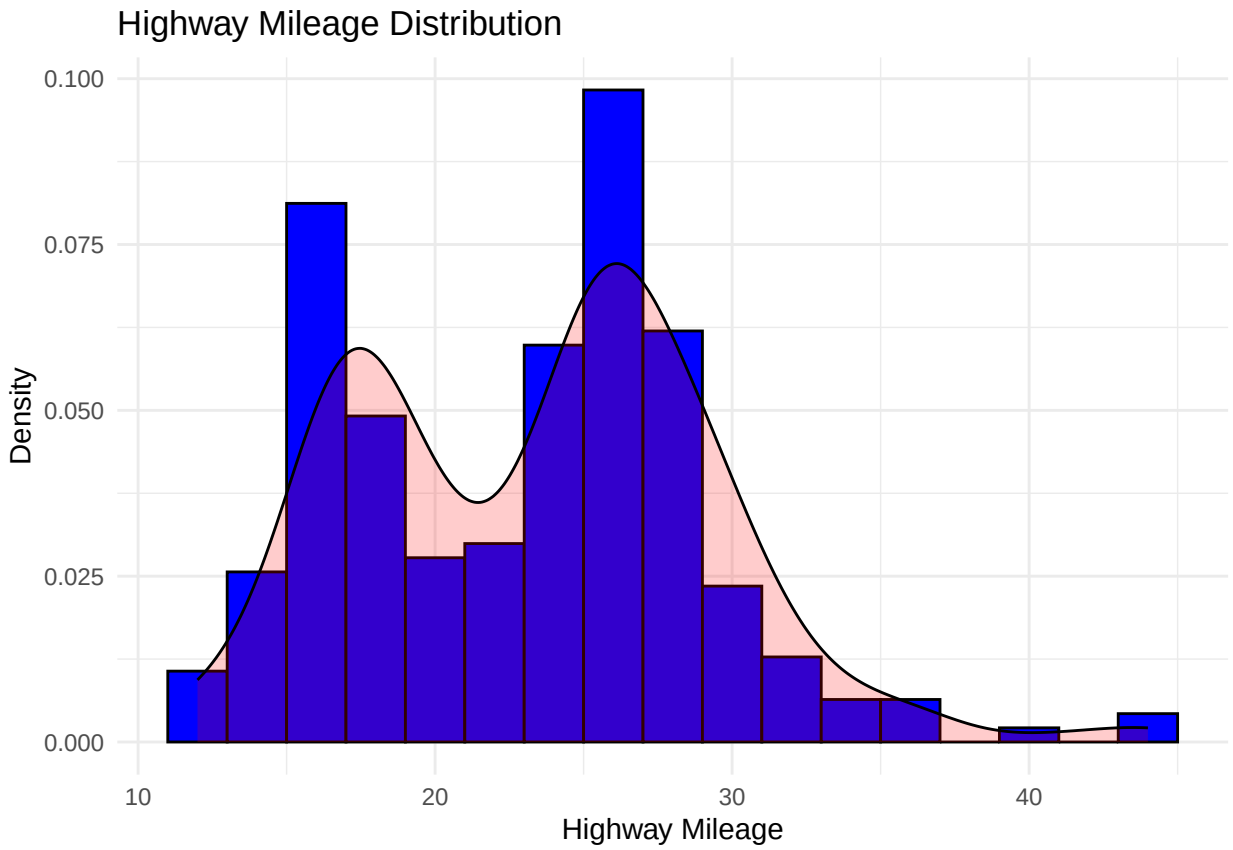


Chart 5: Shows hwy information and shows density usage

- Shows information about the distribution of distances on highways (hwy).
- It helps to see how the data is distributed. What kind of distribution is there (i.e. is it normal?) and helps highlight how often the data appear at different distances.

```
ggplot(data = mpg) +
  aes(x = hwy) +
  geom_histogram(aes(y = ..density..), binwidth = 2, fill = "blue", color = "black") +
  geom_density(alpha = 0.2, fill = "red") +
  labs(title = "Highway Mileage Distribution",
       x = "Highway Mileage",
       y = "Density") +
  theme_minimal()
```

```
## Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(density)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```


Highway Mileage Distribution
0.100
0.075
0.050
0.025
0.000
Density
10
20
30
40
Highway Mileage