Homework_R_Markdown

Eye nwp

2023-12-05

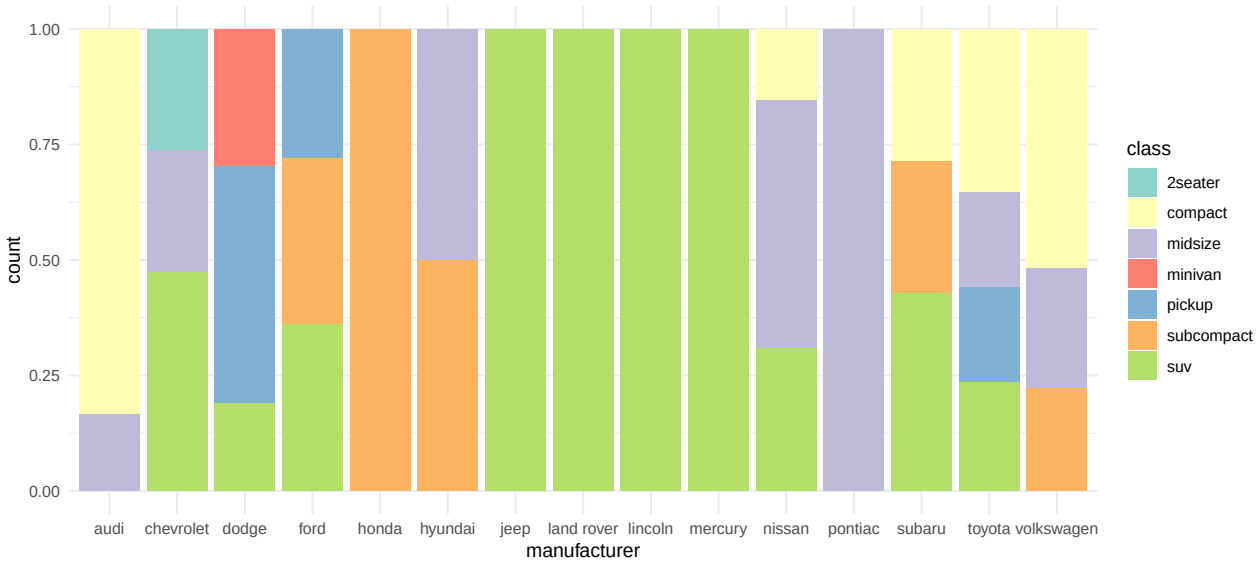
Explore data

```
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~
```

Question 1: How many car type each manufacturer has?

```
ggplot(mpg, aes(manufacturer, fill = class)) +
  geom_bar(position = "fill") +
  theme_minimal() +
  scale_fill_brewer(palette = "Set3")
```

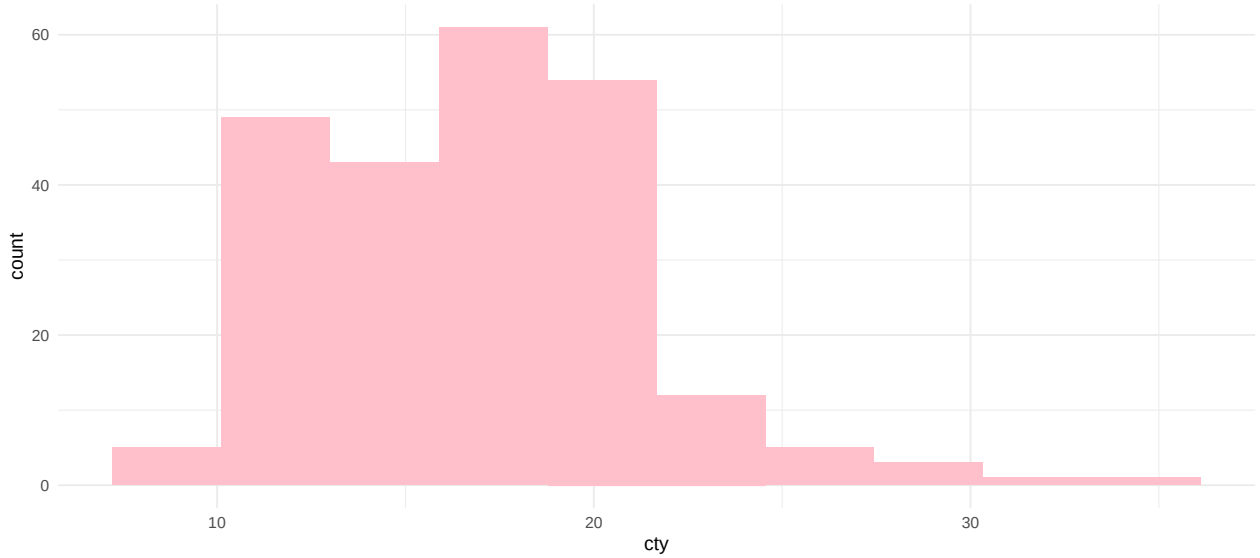


This bar chart shows the percentage of car type for each manufacturer. Jeep, Land Rover,

Linncoln, and Mercury only has suv whereas Toyota has the most various car type which are compact, midsize, pickup, and suv

Question 2: Distribution of cty

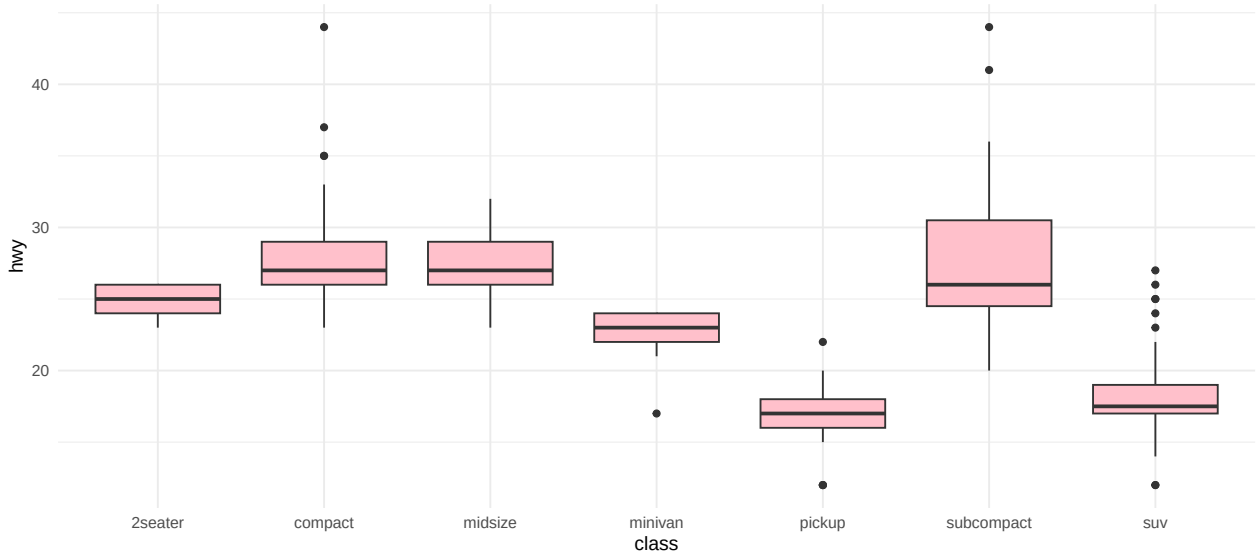
```
ggplot(mpg, aes(cty))+
  geom_histogram(bins = 10, fill = "pink") +
  theme_minimal()
```



Most of the car in this data set has city miles per gallon around 20.

Question 3: Distribution of hwy in each type of car

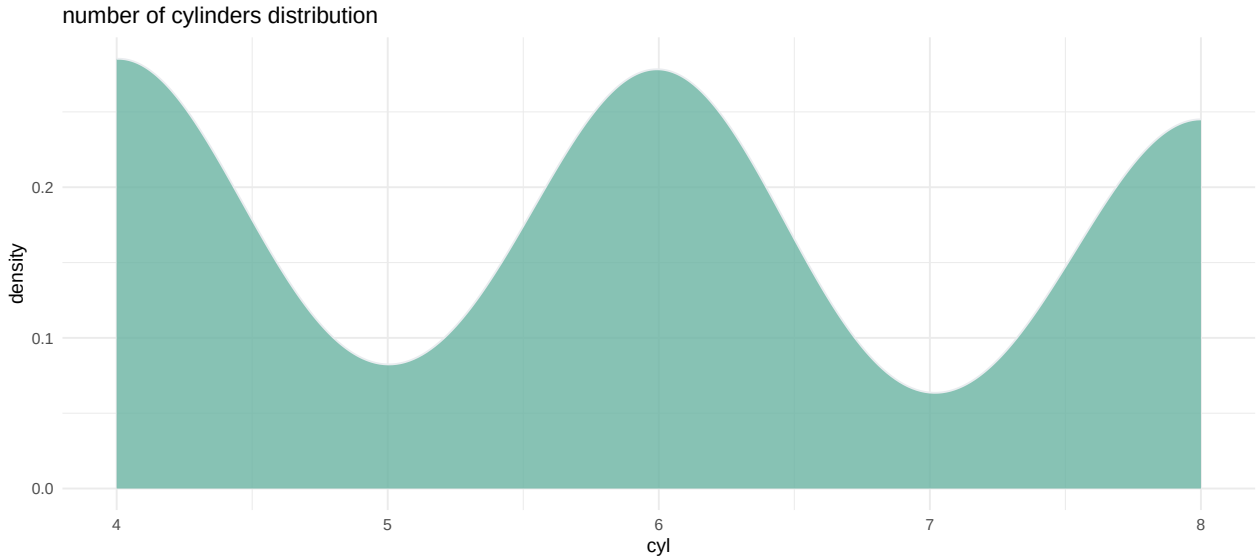
```
ggplot(mpg, aes(class,hwy))+
  geom_boxplot(fill = "pink") +
  theme_minimal()
```



This boxplot chart shows the distribution including median and outliner in highway miles per gallon by each type of car.

Question 4: Distribution of cty using density chart

```
ggplot(mpg, aes(cyl))+
  geom_density(fill="#69b3a2", color="#e9ecef", alpha=0.8) +
  ggtitle("number of cylinders distribution") +
  theme_minimal()
```

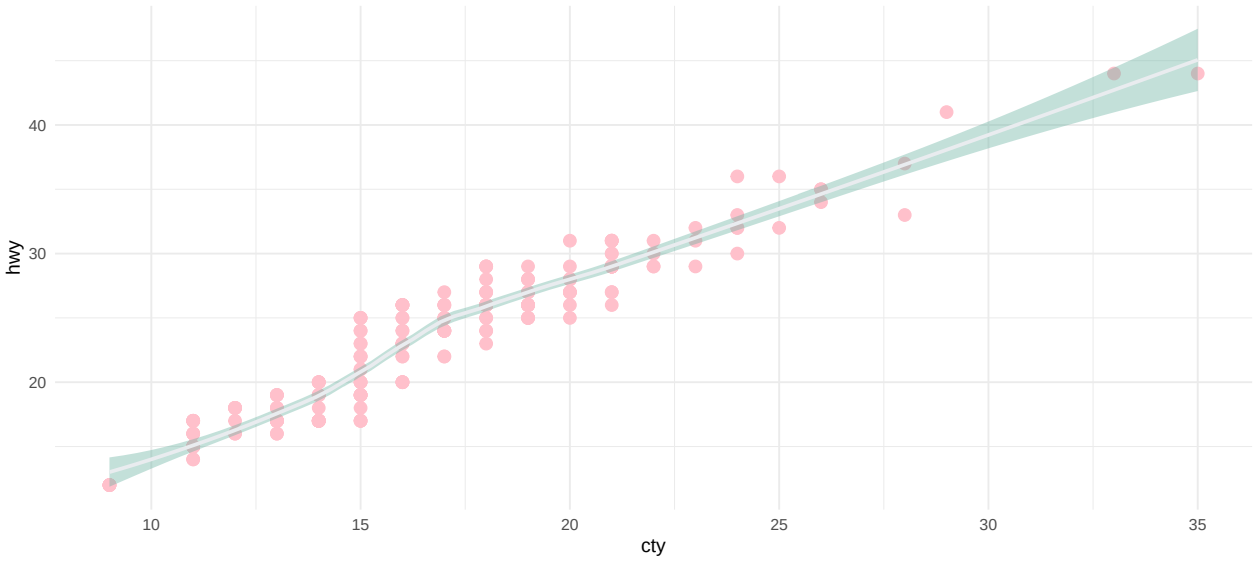


> This density chart shows the distribution of number of cylinder.

Question 5:Correlation between cty and hwy

```
ggplot(mpg, aes(cty, hwy)) +
  geom_point(color = "pink", size = 3) +
  geom_smooth(fill="#69b3a2", color="#e9ecef") +
  theme_minimal()
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



cty and hwy has correlation to each other.