

Homework Data Viz

PiyaY

2024-02-24

Homework

Library

```
library(tidyverse)
library(ggthemes)
library(patchwork)
```

mpg dataset

```
mpg |> head()
```

```
## # 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~
```

Transform Data set

```
## change year int -> char
mpg$year <- as.character(mpg$year)

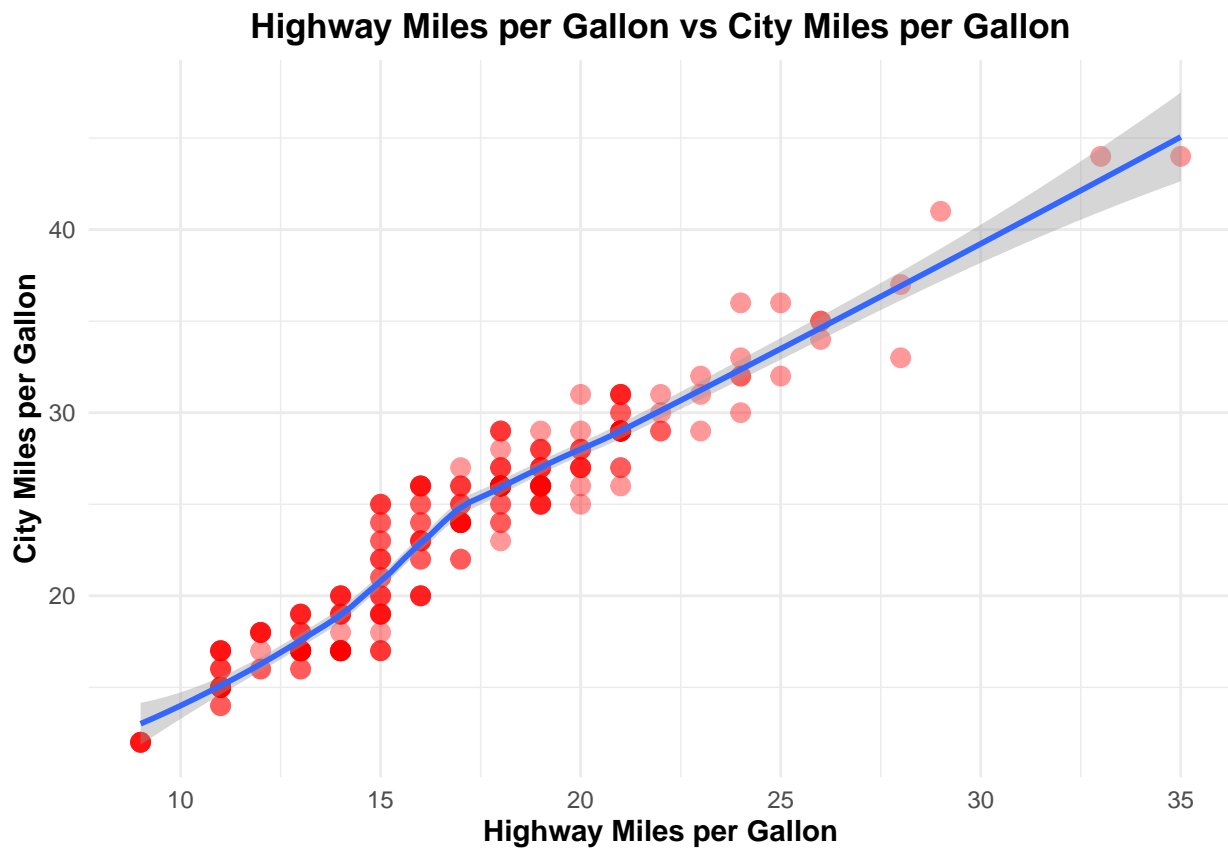
## make car_gear for easy understand
mpg <- mpg |>
  mutate(car_gear = if_else(grepl("^auto", trans), "Auto", "Manual"))
```

Picture setting

```
## Title and axis name setting
pic_setting <- theme(axis.title = element_text(face="bold"),
  plot.title = element_text(hjust = 0.5, face = "bold"))
```

city miles per gallon vs highway miles per gallon

```
ggplot(mpg, aes(cty, hwy)) +
  geom_point(alpha = 0.4, size = 3, col = "red") +
  geom_smooth() +
  theme_minimal() +
  labs(title = "Highway Miles per Gallon vs City Miles per Gallon",
       y = "City Miles per Gallon",
       x = "Highway Miles per Gallon") +
  pic_setting
```



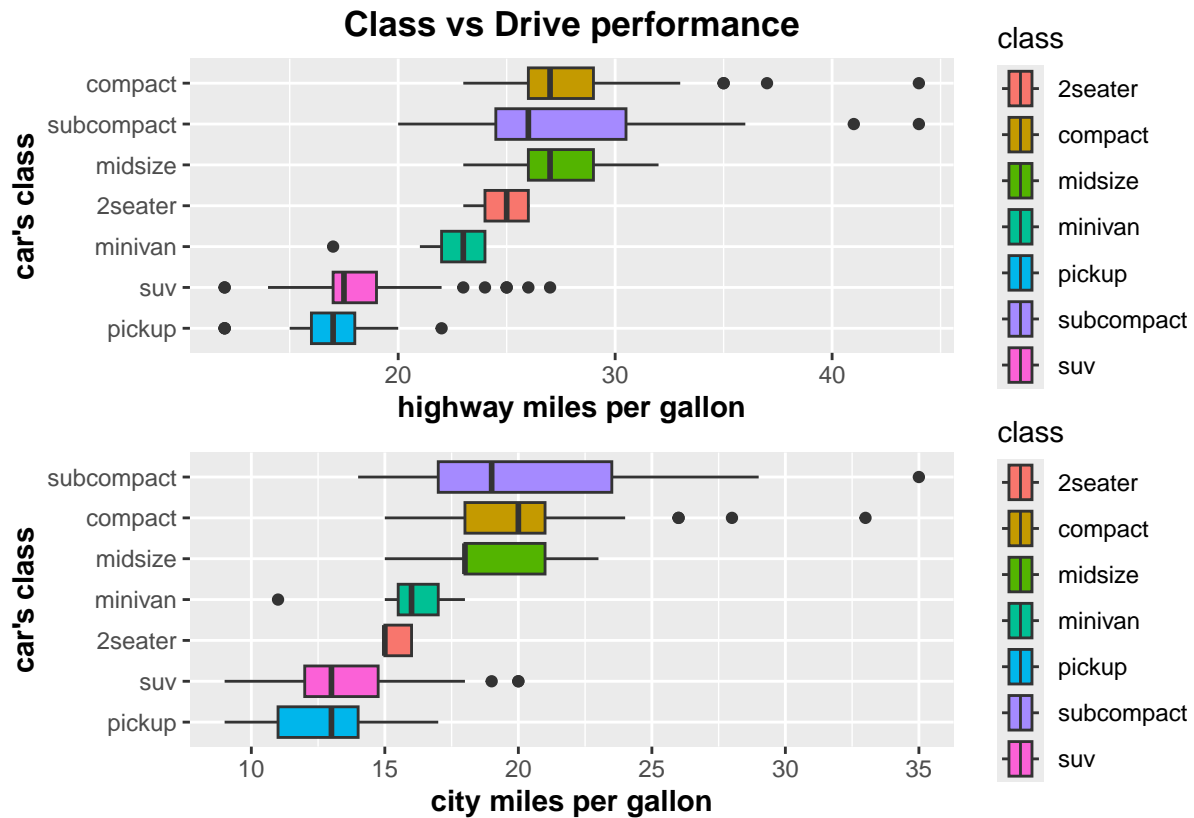
They have correlation.

class vs performance

```
p1 <- ggplot(mpg, aes(y = reorder(class, hwy), x = hwy, fill = class)) +
  geom_boxplot() +
  labs(y = "car's class", x = "highway miles per gallon",
       title = "Class vs Drive performance") +
  pic_setting

p2 <- ggplot(mpg, aes(y = reorder(class, cty), x = cty, fill = class)) +
  geom_boxplot() +
  labs(y = "car's class", x = "city miles per gallon") +
  theme(axis.title = element_text(face="bold"))

p1 / p2
```

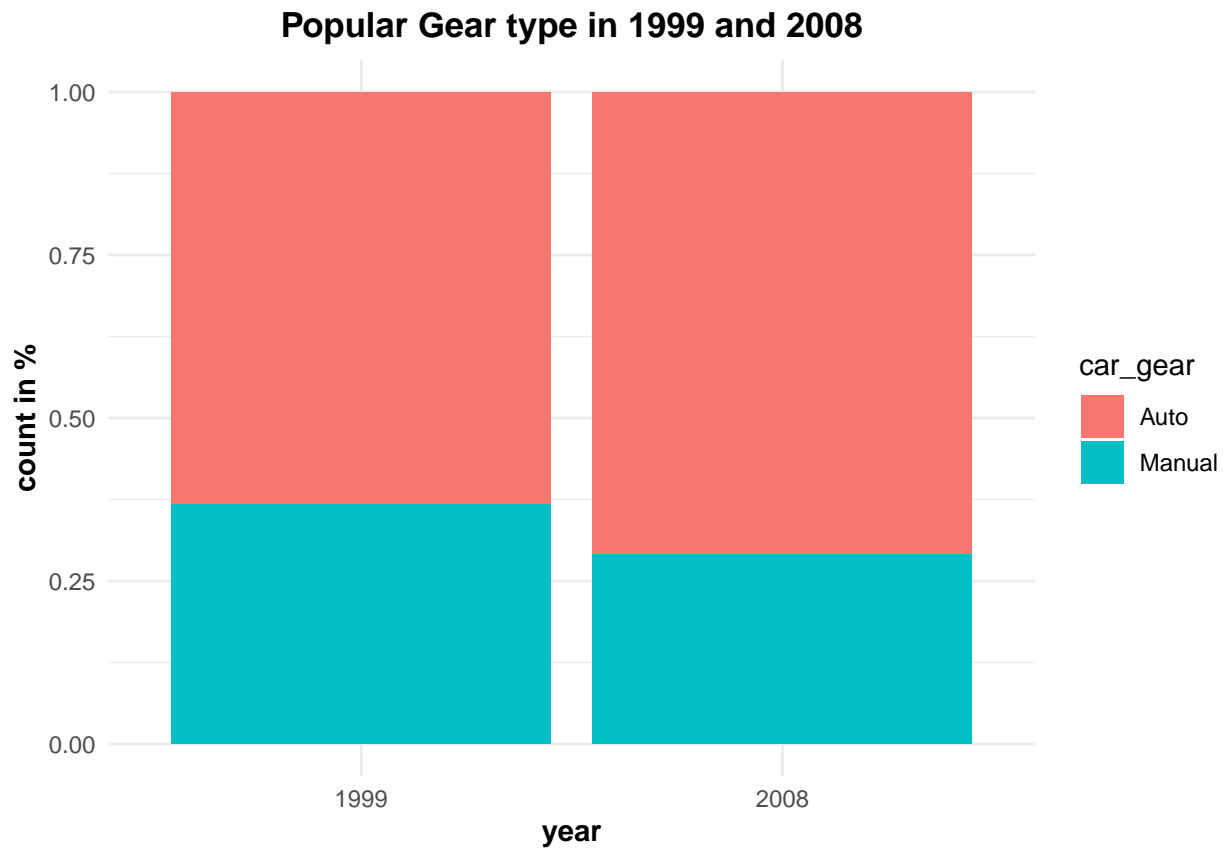


- compact and midsize is best for drive in highway.
- compact is best for city.

summary : Buying a compact car is probably the best value.

popular car gear in 1999 and 2008

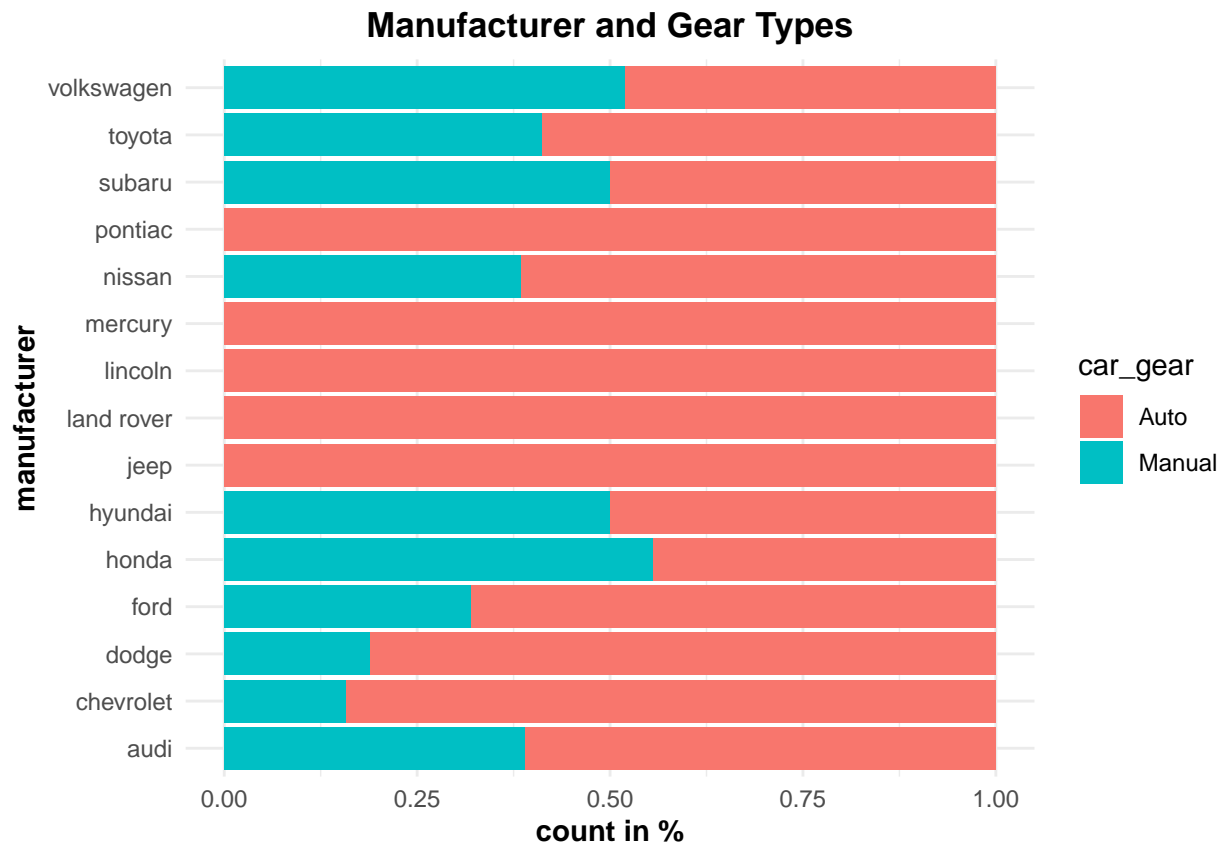
```
ggplot(mpg, aes(year, fill = car_gear)) +
  geom_bar(position = "fill") +
  theme_minimal() +
  labs( title = "Popular Gear type in 1999 and 2008",
        y = "count in %") +
  pic_setting
```



The popularity of the production of automatic transmissions has increased.

manufacturer and gear type trend

```
ggplot(mpg, aes(y = manufacturer, fill = car_gear)) +  
  geom_bar(position = "fill") +  
  theme_minimal() +  
  labs(title = "Manufacturer and Gear Types",  
        x = "count in %") +  
  pic_setting
```

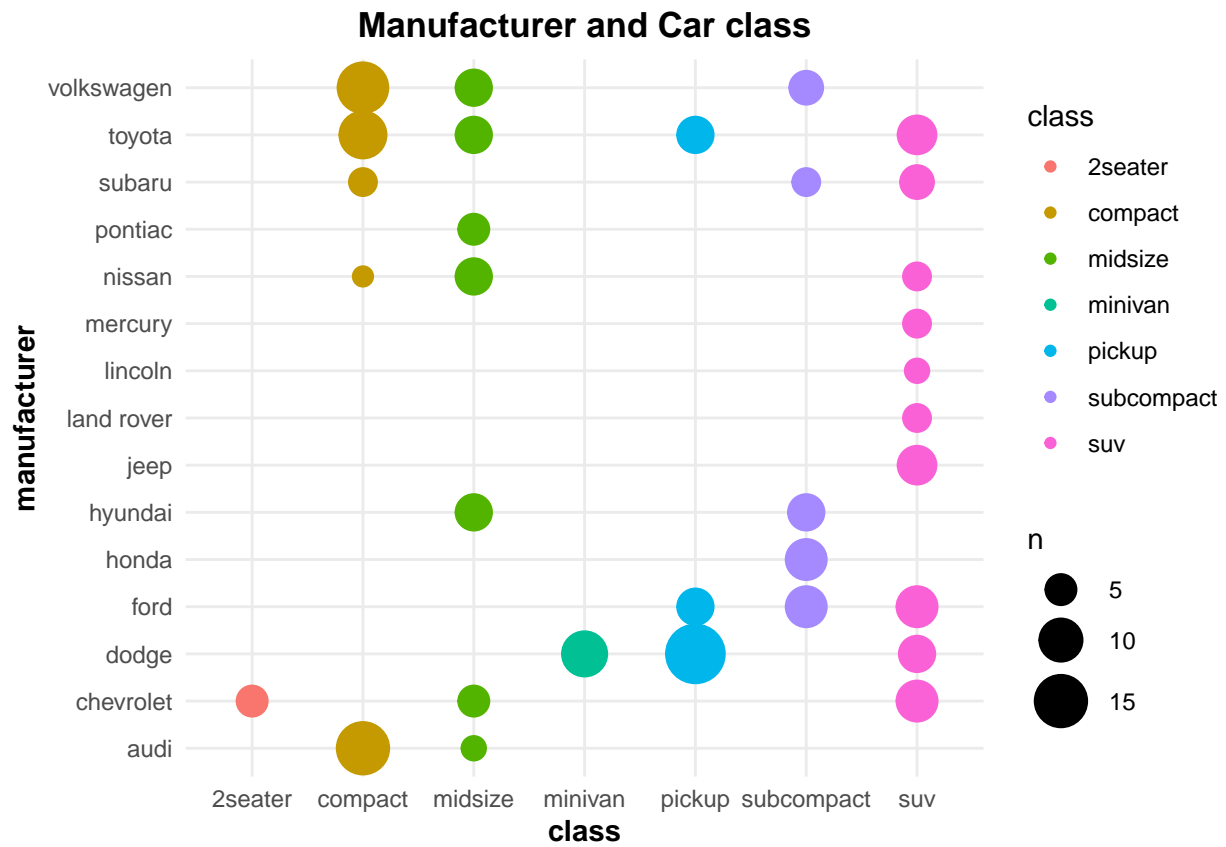


The popularity of production is interesting.

1. Mercury, lincoln, land rover, jeep only produce automatic transmissions.
2. Honda, Hyundai, Subaru, and Volkswagen, half of them still produce automatic transmissions. It can be noted that 3 of the 4 are Japanese companies.

manufacturer and class

```
ggplot(mpg, aes(y = manufacturer, x = class, col = class)) +
  geom_count() +
  scale_size_area(max_size = 10) +
  theme_minimal() +
  labs(title = "Manufacturer and Car class") +
  pic_setting
```



Mercury , Lincoln , Land Rover, Jeep only produce SUV.