Homework_dataviz

lilseal

2023-12-14

Homework

Explore Data

```
## # 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~
```

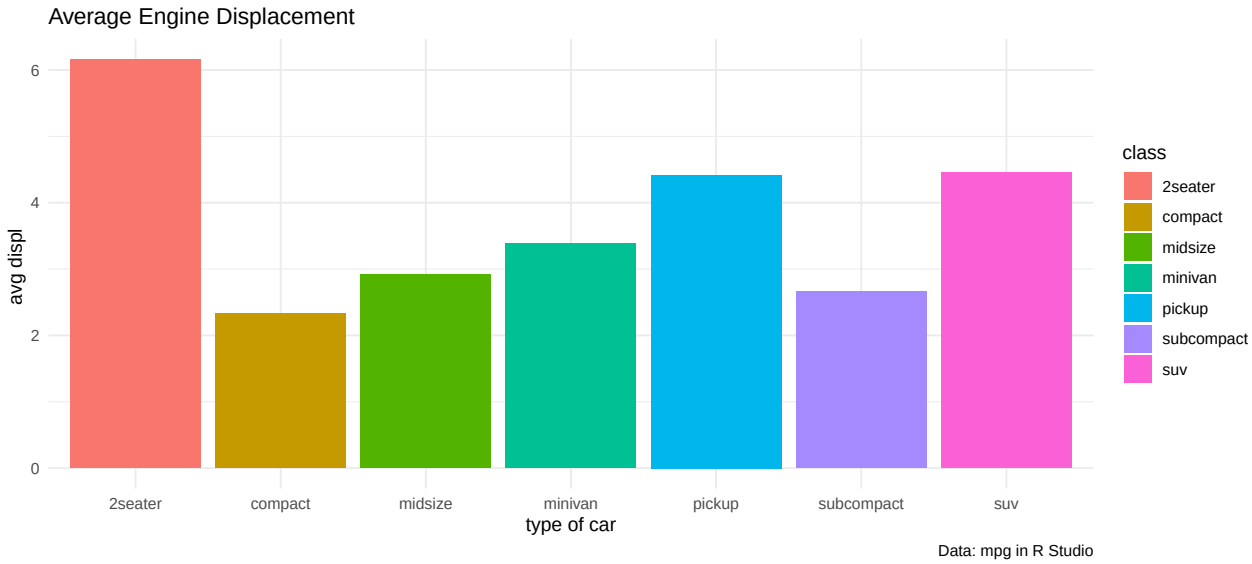
Q1 - Which type of car have the most average engine displacement

1. Aggregate mpg data

```
q1 <- mpg %>%
  group_by(class) %>%
  summarise(avg_displ = mean(displ))
```

2. Make a chart

```
ggplot(q1, aes(class, avg_displ, fill = class)) +
  geom_col() +
  theme_minimal() +
  labs(
    title = "Average Engine Displacement",
    caption = "Data: mpg in R Studio",
    x = "type of car",
    y = "avg displ ")
```



Summary The 2seater car have the most average engine displacement.

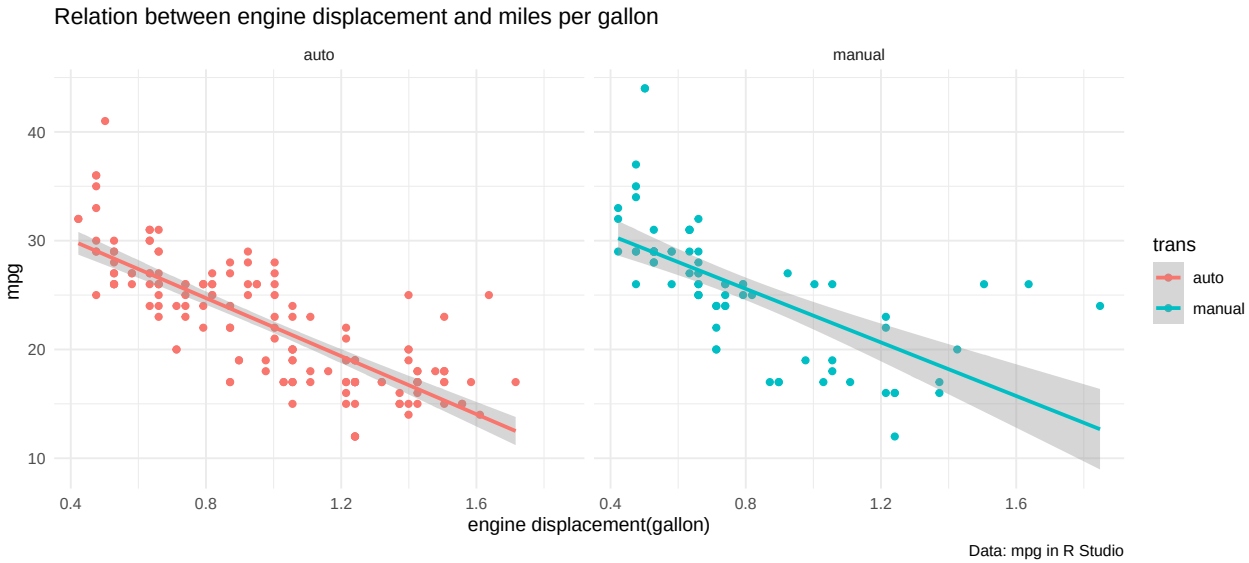
Q2 - Compare relationship between engine displacement and highway miles per gallon and divide in 2 chart by transmission (auto and manual)

1. Edit mpg data column trans and add new aggregate column(change litres to gallon)

```
q2 <- mpg %>%
  mutate(trans = if_else(grepl("^a", trans), "auto", "manual")) %>%
  mutate(dispg = displ *0.264)
```

2. Make a Chart

```
ggplot(q2, aes(dispg, hwy, col = trans)) +
  geom_point() +
  geom_smooth(method="lm") +
  theme_minimal() +
  facet_grid(~trans) +
  labs(
    title = "Relation between engine displacement and miles per gallon",
    caption = "Data: mpg in R Studio",
    x = "engine displacement(gallon)",
    y = "mpg ")
```



Summary Relationship between engine displacement and miles per gallon is inverse variation

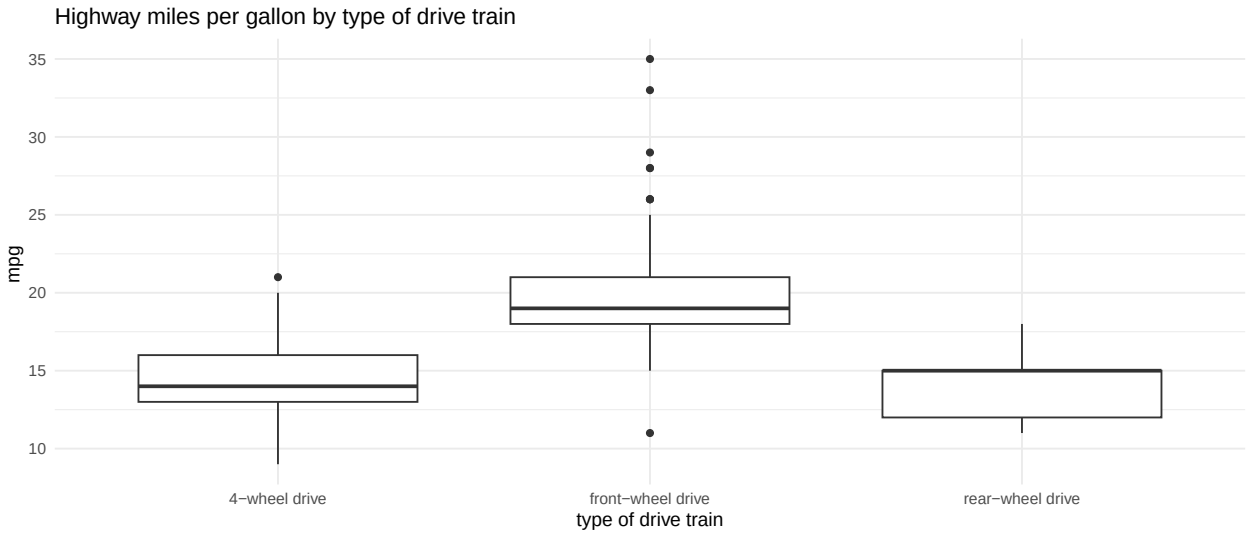
Q3 - Which type of drive train have the least city miles per gallon in Q1, Q2 and Q3

1. Edit mpg data column drv

```
q3 <- mpg %>%
  mutate(drv = if_else(grepl("f", drv), "front-wheel drive",
                       if_else(grepl("r", drv), "rear-wheel drive", "4-wheel drive")))
```

2. Make a Chart

```
ggplot(q3, aes(drv, cty)) +
  geom_boxplot() +
  theme_minimal() +
  labs(
    title = "Highway miles per gallon by type of drive train",
    caption = "Data: mpg in R Studio",
    x = "type of drive train",
    y = "mpg ")
```



Summary The most city miles per gallon in Q1,Q2,Q3 is *front-wheel drive* In Q1, the least city miles per gallon is *rear-wheel drive* In Q2, the least city miles per gallon is *4-wheel drive* In Q3, the least city miles per gallon is *rear-wheel drive*

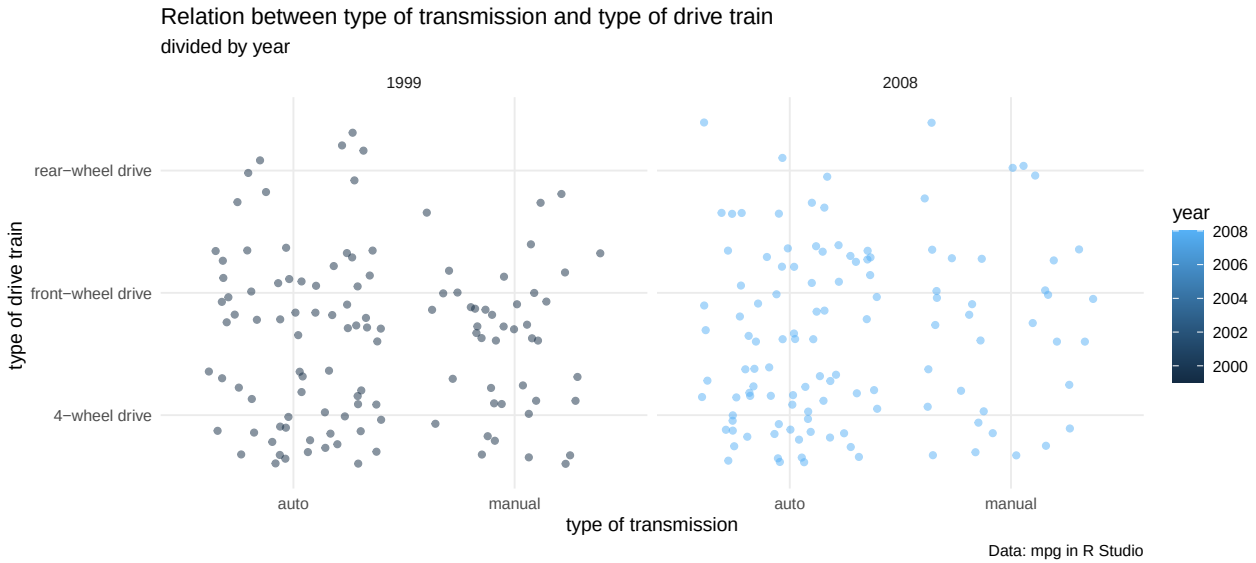
Q4 - Compare relation between type of transmission and type of drive train (divide in 2 chart by year)

1. Filter data and specify

```
q4 <- mpg %>%
  mutate(trans = if_else(grepl("^a", trans), "auto", "manual")) %>%
  mutate(drv = if_else(grepl("f", drv), "front-wheel drive",
                       if_else(grepl("r", drv), "rear-wheel drive", "4-wheel drive")))
```

2. Make a chart

```
ggplot(q4, aes(trans, drv, col = year)) +
  geom_jitter(alpha = 0.5) +
  facet_grid(~year) +
  theme_minimal() +
  labs(
    title = "Relation between type of transmission and type of drive train",
    subtitle = "divided by year",
    caption = "Data: mpg in R Studio",
    x = "type of transmission",
    y = "type of drive train ")
```



Summary The auto transmission is the most popular in both of 1999 and 2018 by the eay the rear-wheel drive is the least popular too

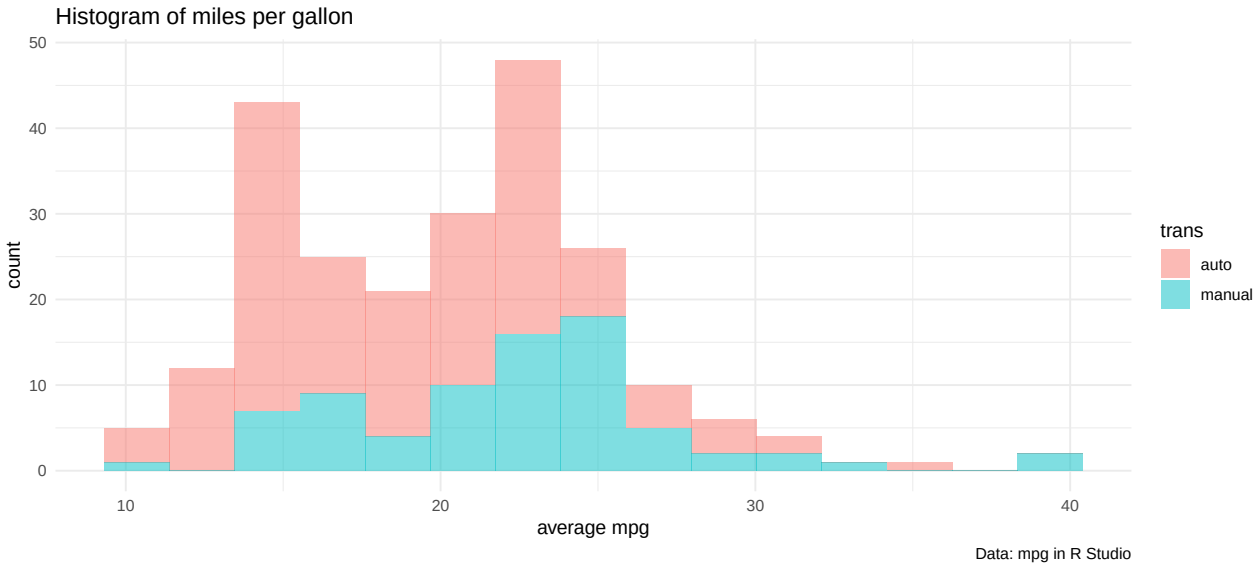
Q5 - Histogram of miles per gallon(average from cty and hwy) and fill color it by transmission

1. Aggregate data

```
q5 <- mpg %>%
  mutate(avg_mpg = (cty+hwy)/2) %>%
  mutate(trans = if_else(grepl("^a", trans), "auto", "manual"))
```

2. Make a chart

```
ggplot(q5, aes(avg_mpg, fill = trans)) +
  geom_histogram(bins = 15, alpha = 0.5) +
  theme_minimal() +
  labs(title = "Histogram of miles per gallon",
    caption = "Data: mpg in R Studio",
    x = "average mpg")
```



Summary The histogram of mpg in auto and manual transmission is *Positive skew*