Hypothesis: “Sports cars are the least efficient type of car” based on the dataset *mpg*.

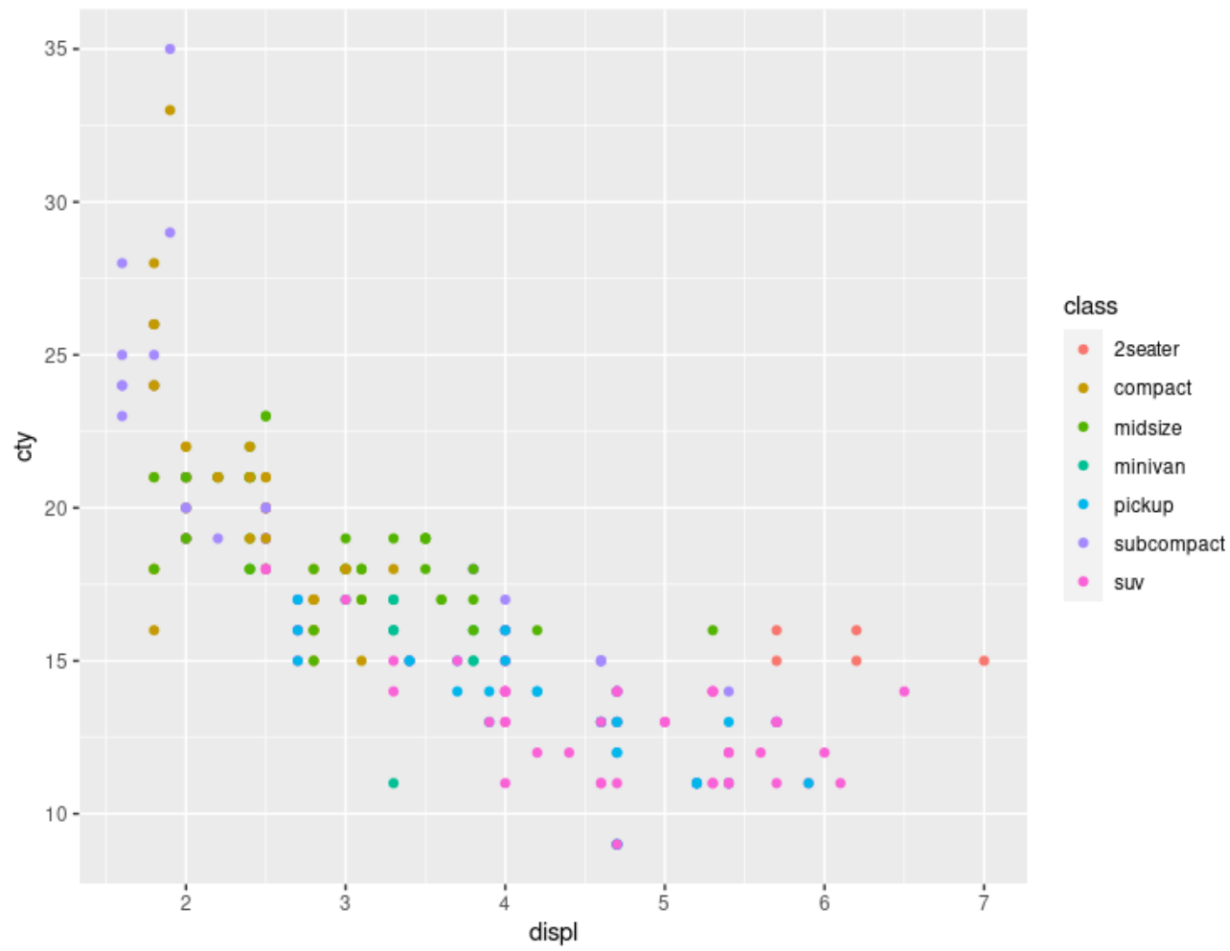
Preliminaries
Efficiency was considered in terms of *hwy* or *cty* measurements.

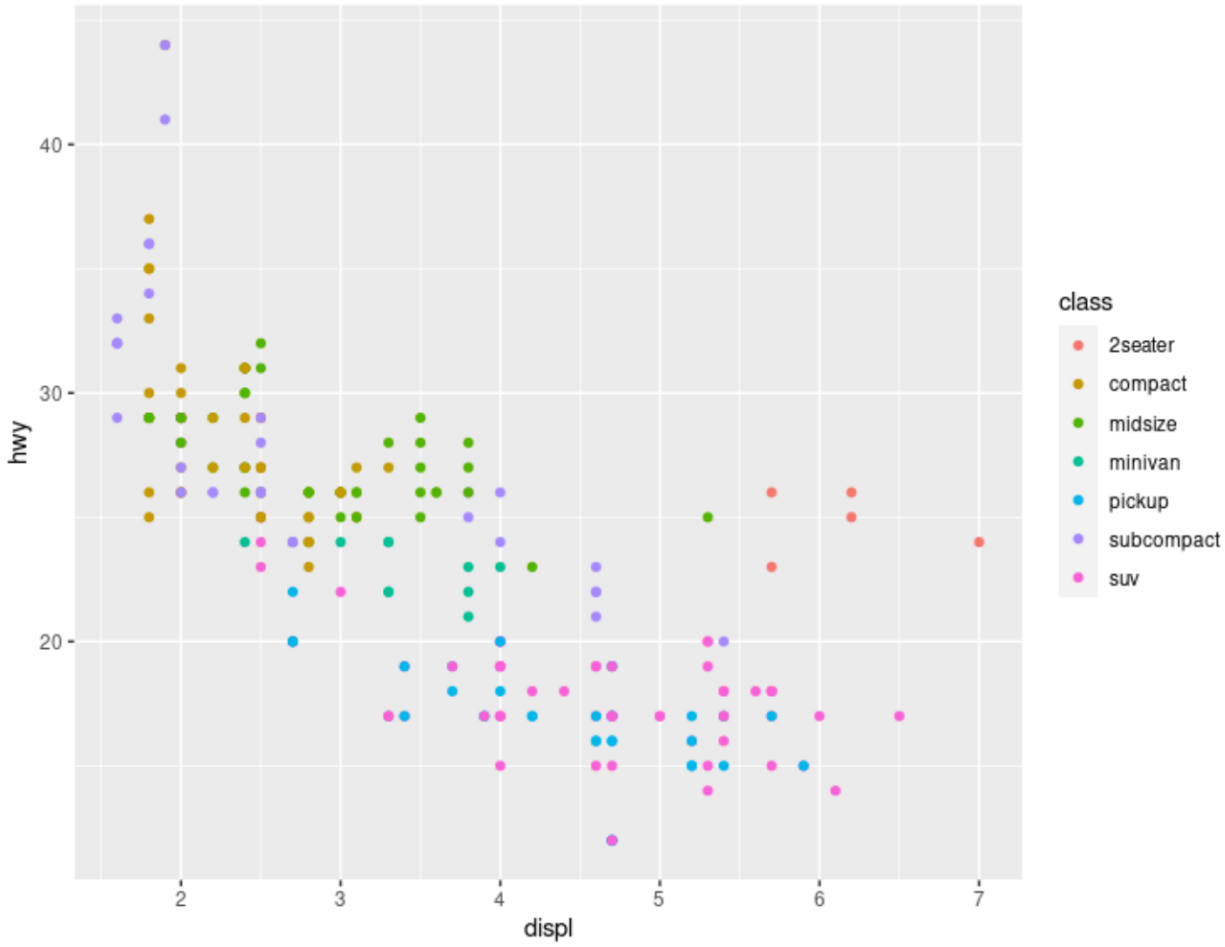
Sports cars were considered as the “2seater” class in the dataset.

Overview
In terms of analysis, we first consider all vehicle classes, plotting efficiency ratings versus displacement (the first two plots). Based on these visualizations, it does not seem that the sports cars are the least efficient, with certain other classes (pickups or SUVs) achieving lower efficiency ratings.

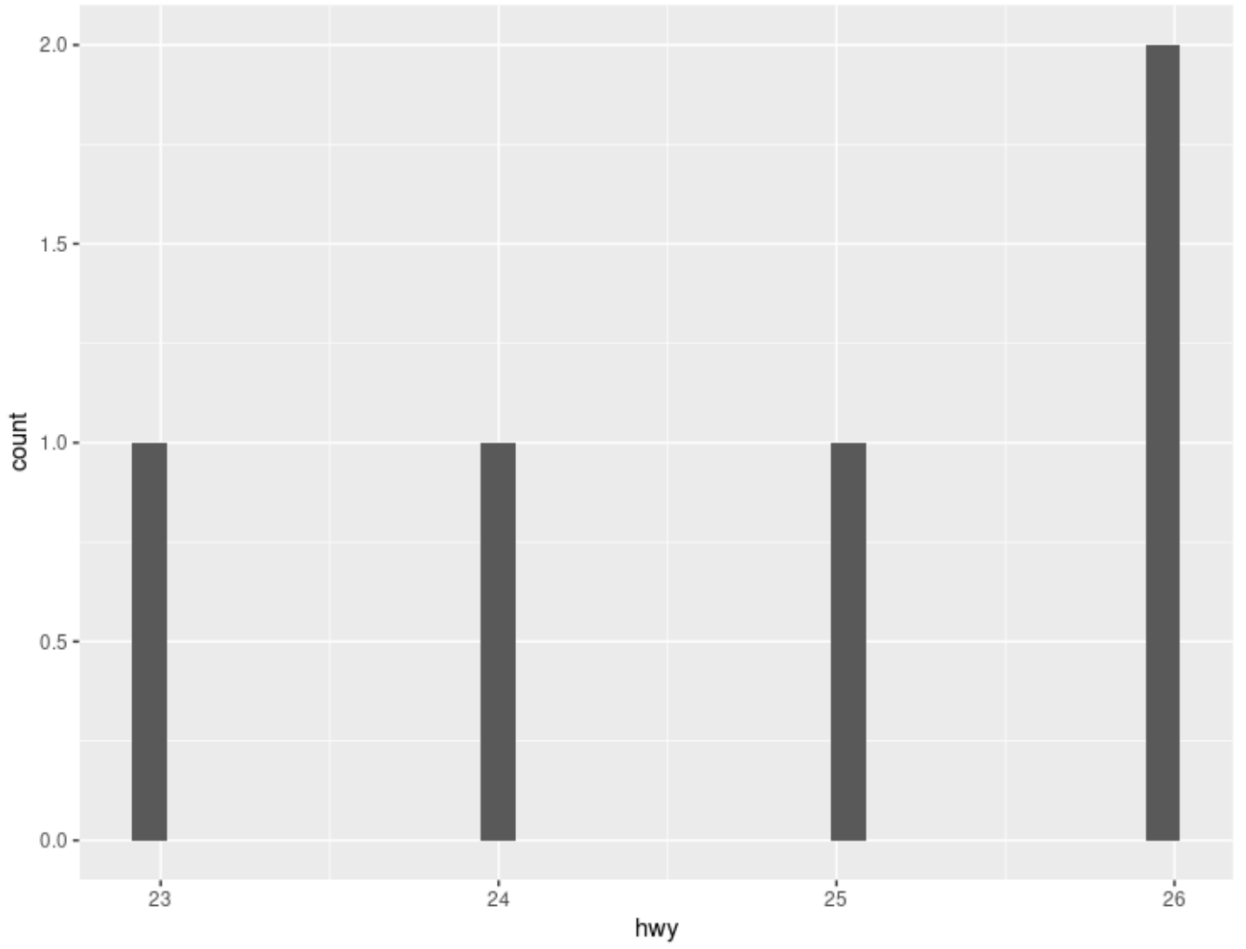
To compare these classes, we did subsequent analysis, showing spread of fuel efficiencies within each class (in terms of histograms). We also determined mean values, medians and quartiles for each class in terms of highway fuel efficiency.

Conclusion
From these analyses, it seems that the hypothesis is false. Pickups and SUVs are less efficient than sports cars. Some exceptions, however, exist. For instance, in terms of *cty* efficiency, one pickup exists which outperforms all the sports cars for *cty* efficiency in the dataset. However, as shown by the mean values of each respective class, sports cars are not the least efficient class of car.

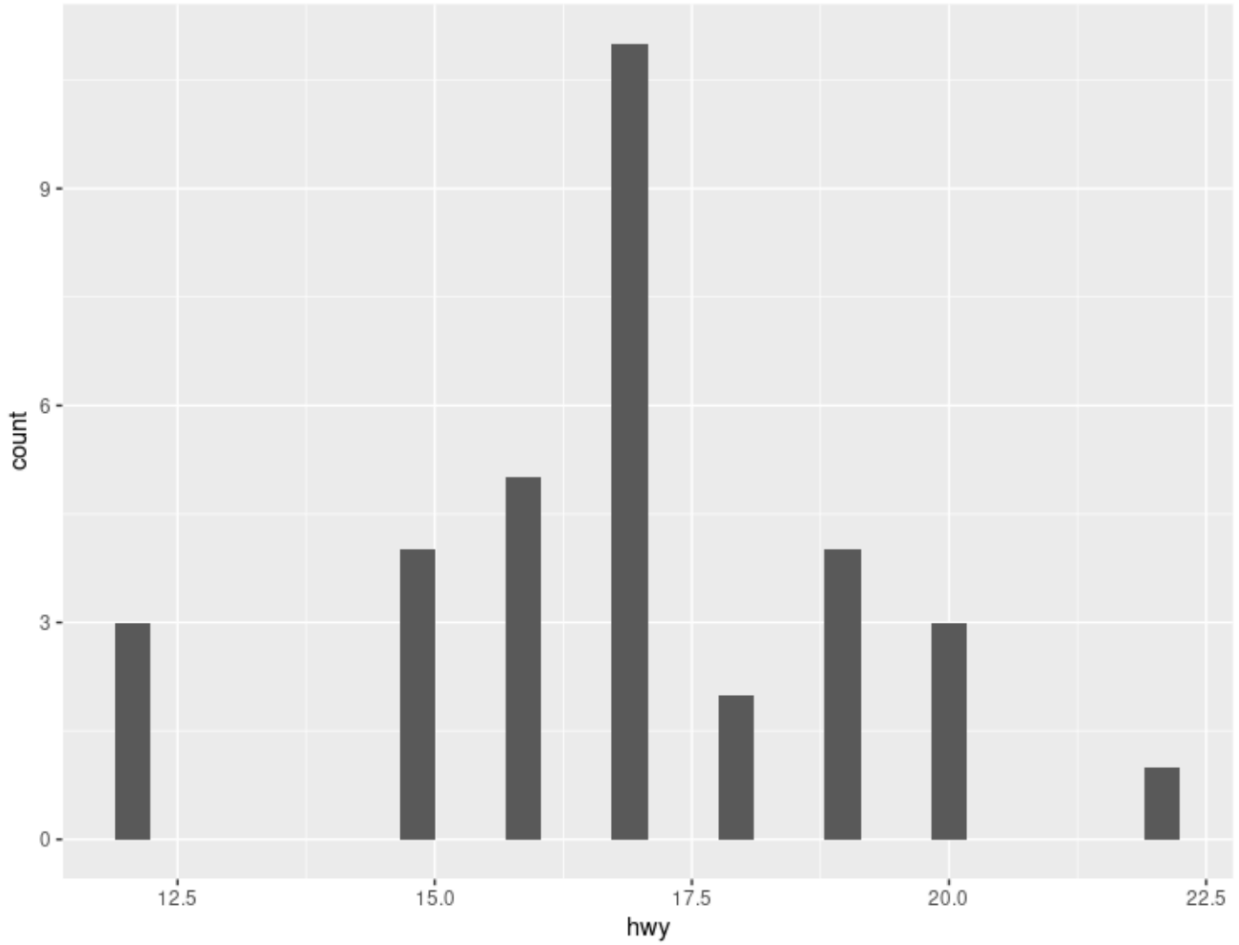

35
30
25
20
15
10
cty
2
3
4
5
6
7
displ
class
2seater
compact
midsize
minivan
pickup
subcompact
suv


hwy
40
30
20
2
3
4
5
6
7
displ
class
2seater
compact
midsize
minivan
pickup
subcompact
suv

Sports car histogram (hwy)



Pickup histogram (hwy)



R Code:

```
# load the mpg library and other useful functions
library(tidyverse)

# definitions
  # efficiency: cty or hwy
  # sports cars: 2 seaters with rwd
```

```
# visualizations
  # show cty versus displacement by class
  ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y=cty, color=class))

  # show hwy versus displacement by class
  ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y=hwy, color=class))

  # extraction : from visualizations, we can look at
    # vehicles with less than 20 mpg (hwy) OR less than 12 mpg (cty)

    # get just the sports cars from mpg
    mpg_sport <- dplyr::filter(mpg, class=="2seater" & ((hwy < 30) | (cty < 20)))
    ggplot(data = mpg_sport) + geom_histogram(mapping = aes(hwy))
    summary(mpg_sport$cty)
    summary(mpg_sport$hwy)

    # get just the suvs from from mpg
    mpg_suv <- dplyr::filter(mpg, class=="suv" & ((hwy < 30) | (cty < 20)))
    ggplot(data = mpg_suv) + geom_histogram(mapping = aes(hwy))

    # get just the minivans from mpg
    mpg_minivan <- dplyr::filter(mpg, class=="minivan" & ((hwy < 30) | (cty < 20)))
    ggplot(data = mpg_minivan) + geom_histogram(mapping = aes(hwy))

    # get just the pickups from mpg
    mpg_pickup <- dplyr::filter(mpg, class=="pickup" & ((hwy < 30) | (cty < 20)))
    ggplot(data = mpg_pickup) + geom_histogram(mapping = aes(hwy))
    summary(mpg_pickup$cty)
    summary(mpg_pickup$hwy)

    # as a comparison, midsize
    mpg_midsize <- dplyr::filter(mpg, class=="midsize" & ((hwy < 30) | (cty < 20)))
    ggplot(data = mpg_midsize) + geom_histogram(mapping = aes(hwy))
```