Homework 1

Sukii

2023-12-11

Get library

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become error
```

Explore Data = mpg

```r
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~
```

Details ;

manufacturer : manufacturer mane

model : model name

displ : engine displacement, in litres

year : year of munufacture

cyl : number of cylinders

trans : type of transmission

drv : the type of drive train, where f = front-wheel drive, r = rear wheel drive, 4 = 4wd cty : city miles per gallon

hwy : highway miles per gallon

fl : fuel type

class : "type" of car

Create Charts

1. Market Share

```r
ggplot(mpg, aes(class, fill=class)) +
  geom_bar(alpha=.8) +
  labs(
    title="The most market share",
    subtitle = "By type",
    caption = "Data : mpg data",
    x="Class type",
    y="Number of class"
  ) +
  theme_minimal() +
  scale_fill_brewer(palette="Paired")
```
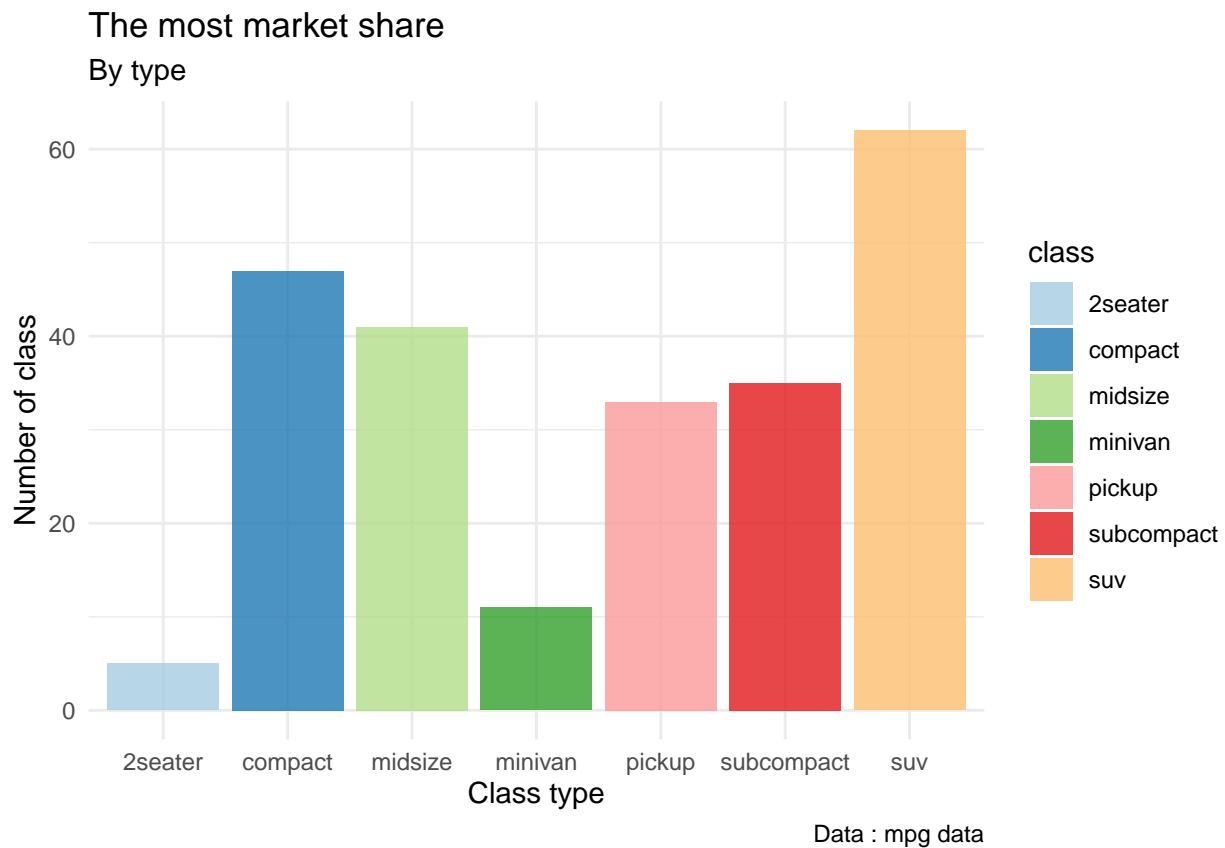

Fig. 1 shown class type of cars in market. This chart show suv type is the most market share

2. Scatter plot of cyl vs. hwy, cyl vs. cty

```r
d_hwy <- ggplot(mpg, aes(cyl, hwy)) +
  geom_point() +
  labs(x = "Miles per gallon",
```
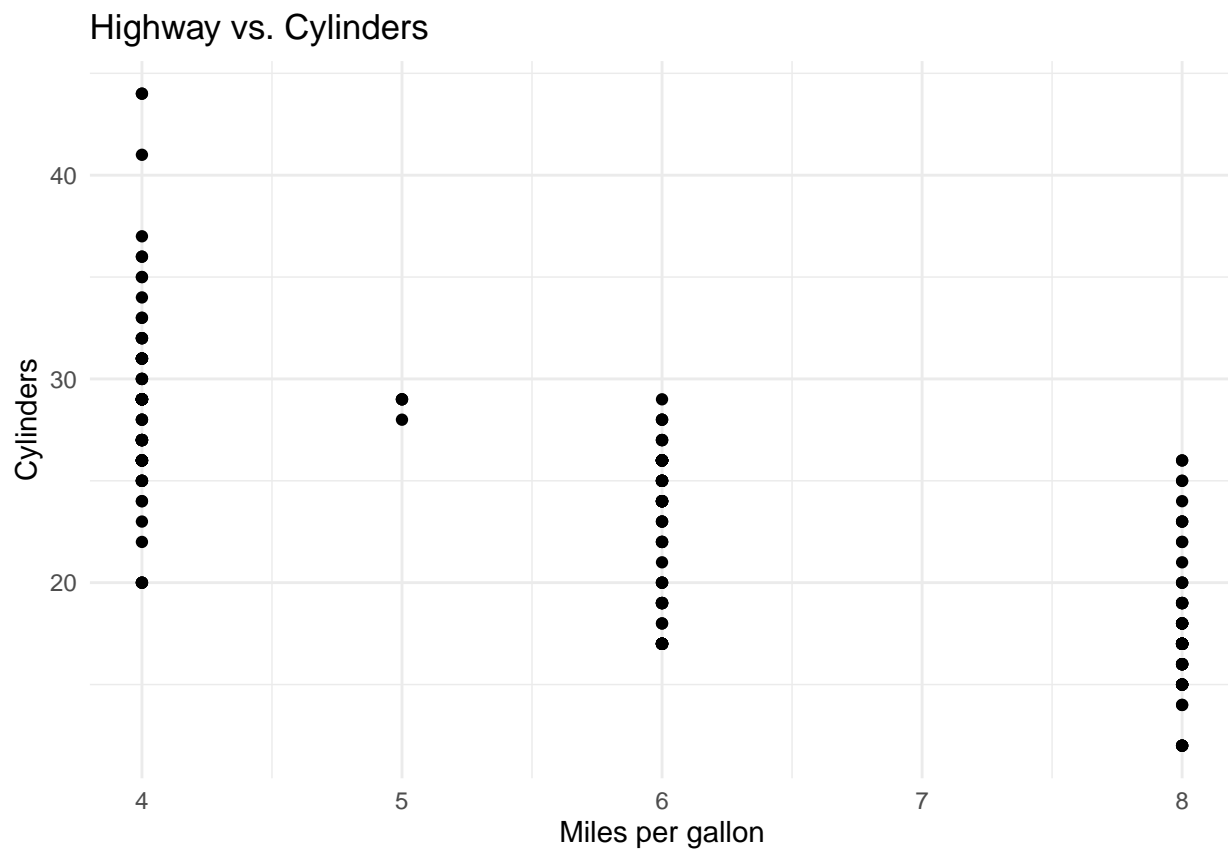
```
      y = "Cylinders",
      title = "Highway vs. Cylinders") +
  theme_minimal()

d_cty <- ggplot(mpg, aes(cyl, cty)) +
  geom_point() +
  labs(x = "Miles per gallon",
      y = "Cylinders",
      title = "City vs. Cylinders") +
  theme_minimal()

d_hwy
```

Highway vs. Cylinders

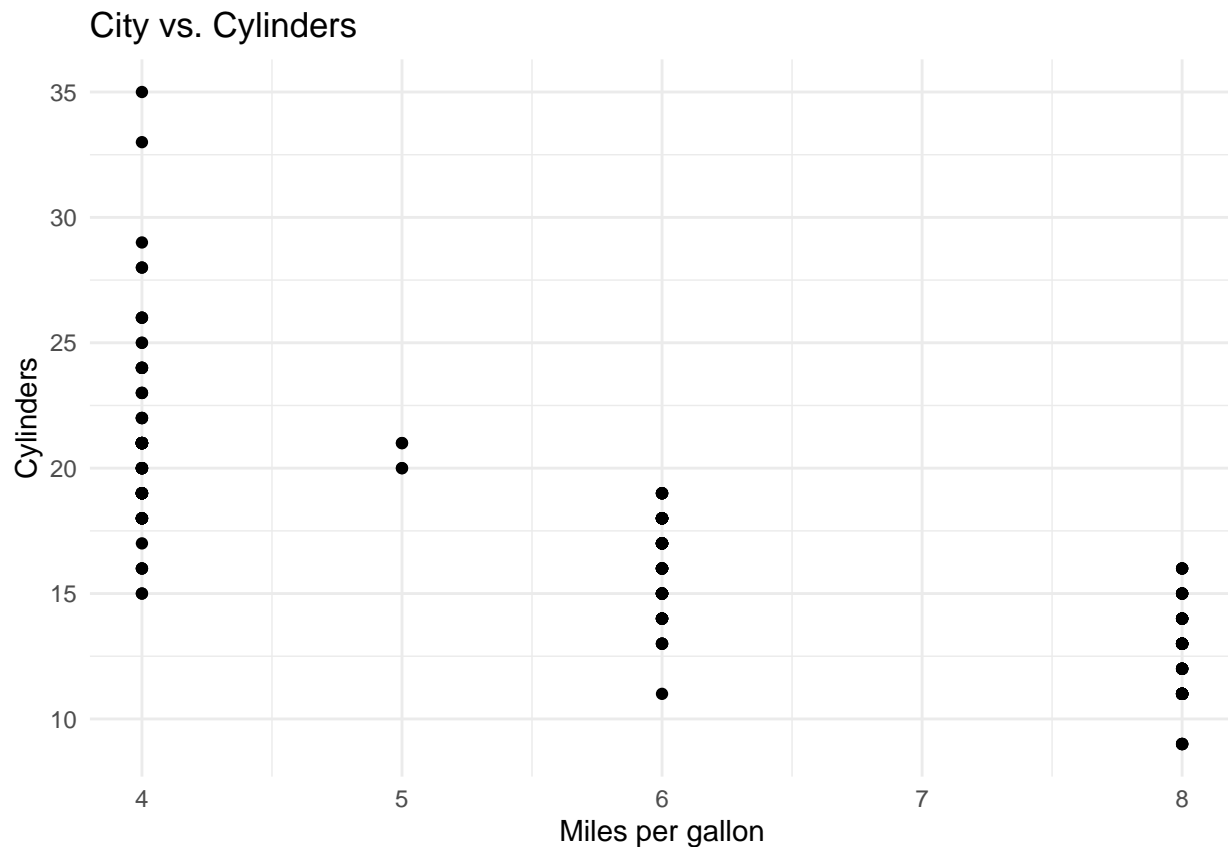
```
d_cty
```

City vs. Cylinders

Fig. 2,3 shown relationship between cylinders and distance in miles per gallon, from this example show cars which have 4 cylinders be appropriate than 5, 6 or 8 cylinders no matter where city or highway.

3. Histogram of the distribution of cylinders

```
ggplot(mpg, aes(cyl)) +
  geom_histogram() +
  labs(x = "Miles per gallon",
       y = "Cylinders",
       title = "City vs. Cylinders") +
  theme_minimal()
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
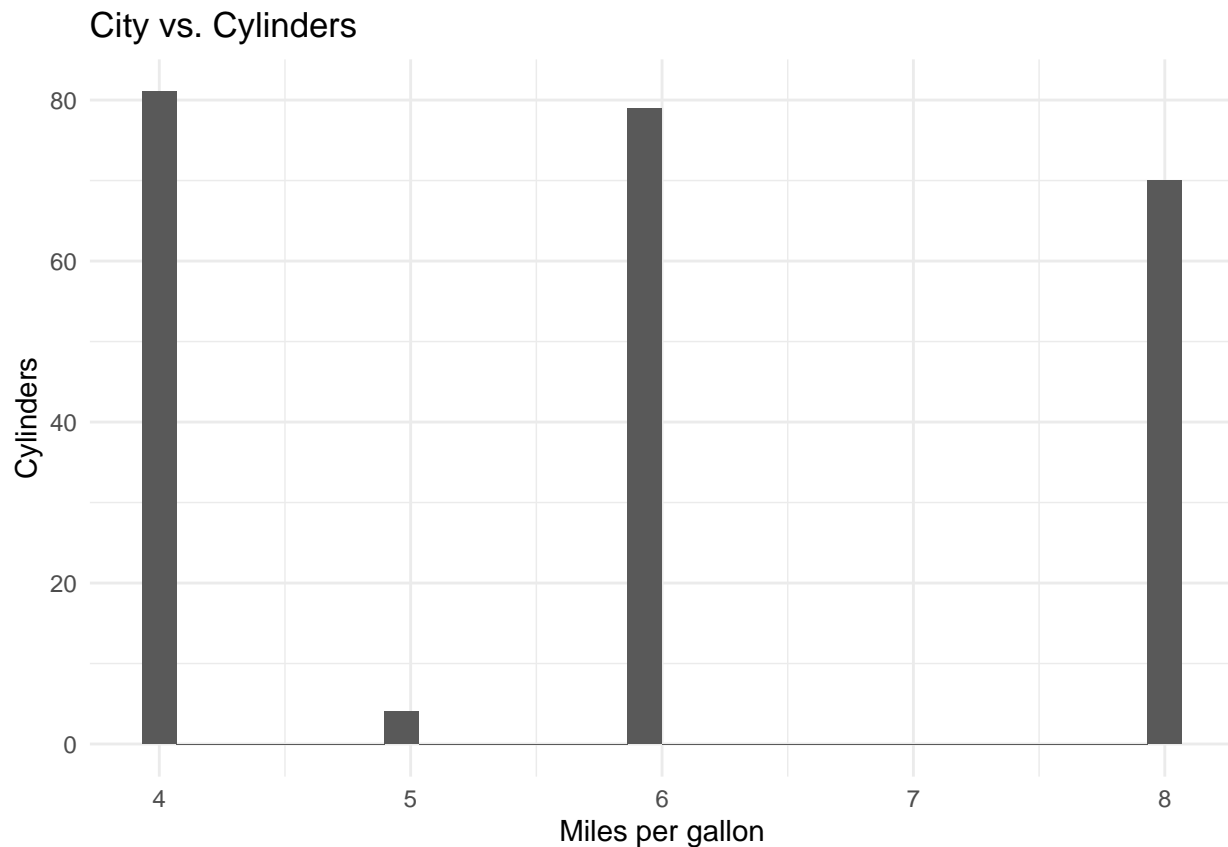
City vs. Cylinders

Fig. 4 show which 4 and 6 cylinders are adopted more than 8. reason is 4, 6 cylinders have good performance than 8; distance in city, highway per gallon

4. Compare hit car before and after year 2000 by manufacturer

```
d_year1 <- mpg %>%
  filter(year >= 2000)

d_year2 <- mpg %>%
  filter(year < 2000)


ggplot(d_year1, aes(manufacturer, fill=manufacturer)) +
  geom_bar() +
  labs(x = "Manufacturer",
       title = "Car after year 2000") +
  theme_minimal()
```
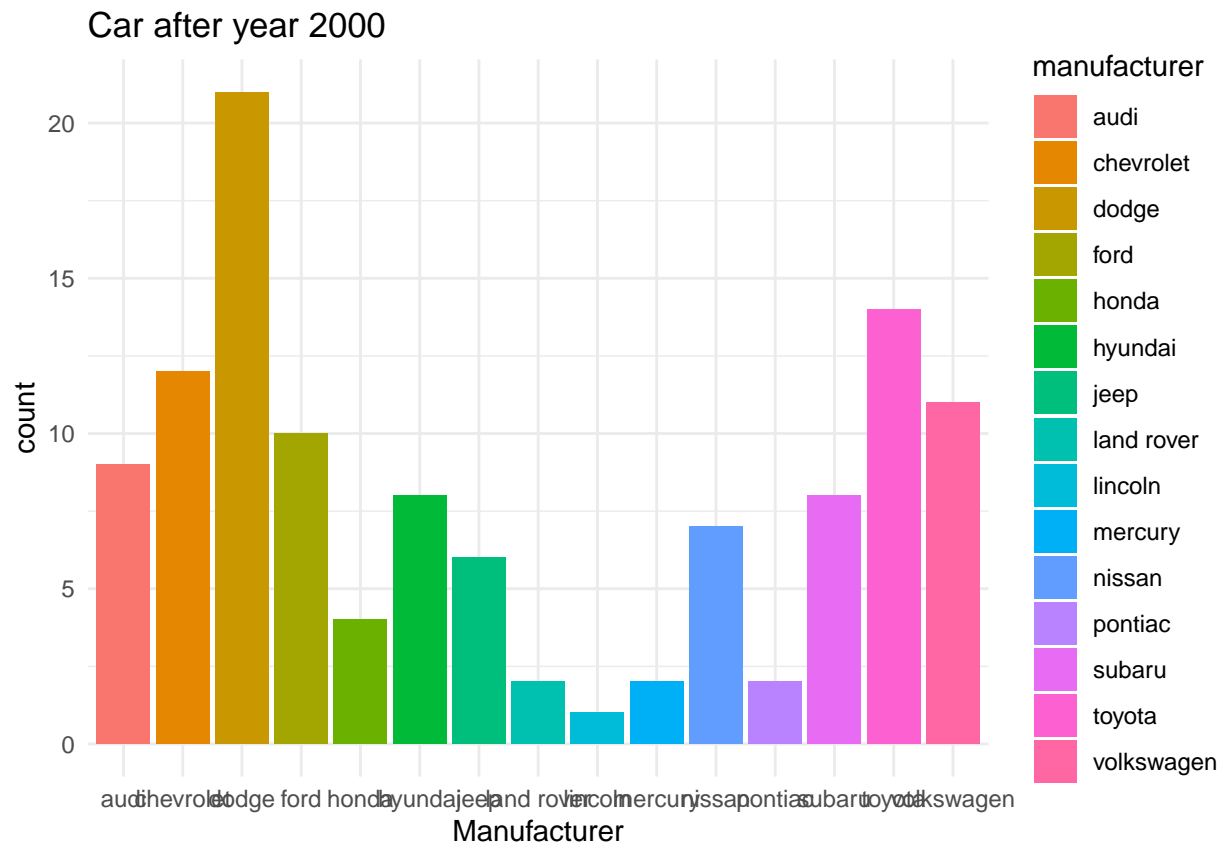
Car after year 2000


```
ggplot(d_year2, aes(manufacturer, fill=manufacturer)) +
  geom_bar() +
  labs(x = "Manufacturer",
       title = "Car before year 2000") +
  theme_minimal()
```
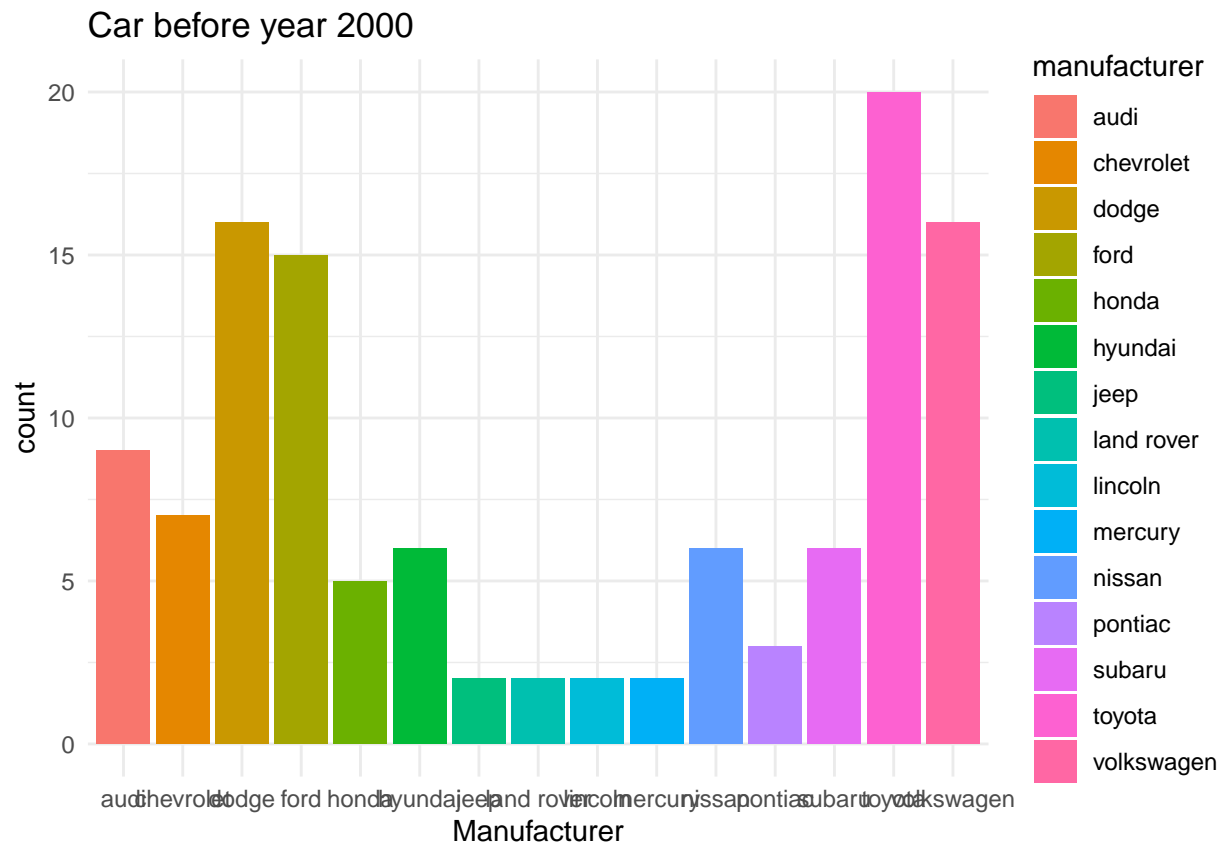
Car before year 2000

Fig. 5,6 shown before year 2000, Toyota is the most popular. But after year 2000, is Dodge

5. Cty and Hwy

```
ggplot(mpg, aes(hwy, cty)) +
  geom_point(color="red") +
  geom_smooth(method="loess", se=TRUE,
              fill="green",
              color="blue") +
  labs(x = "Highway",
       y = "City",
       title = "City Highway") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```
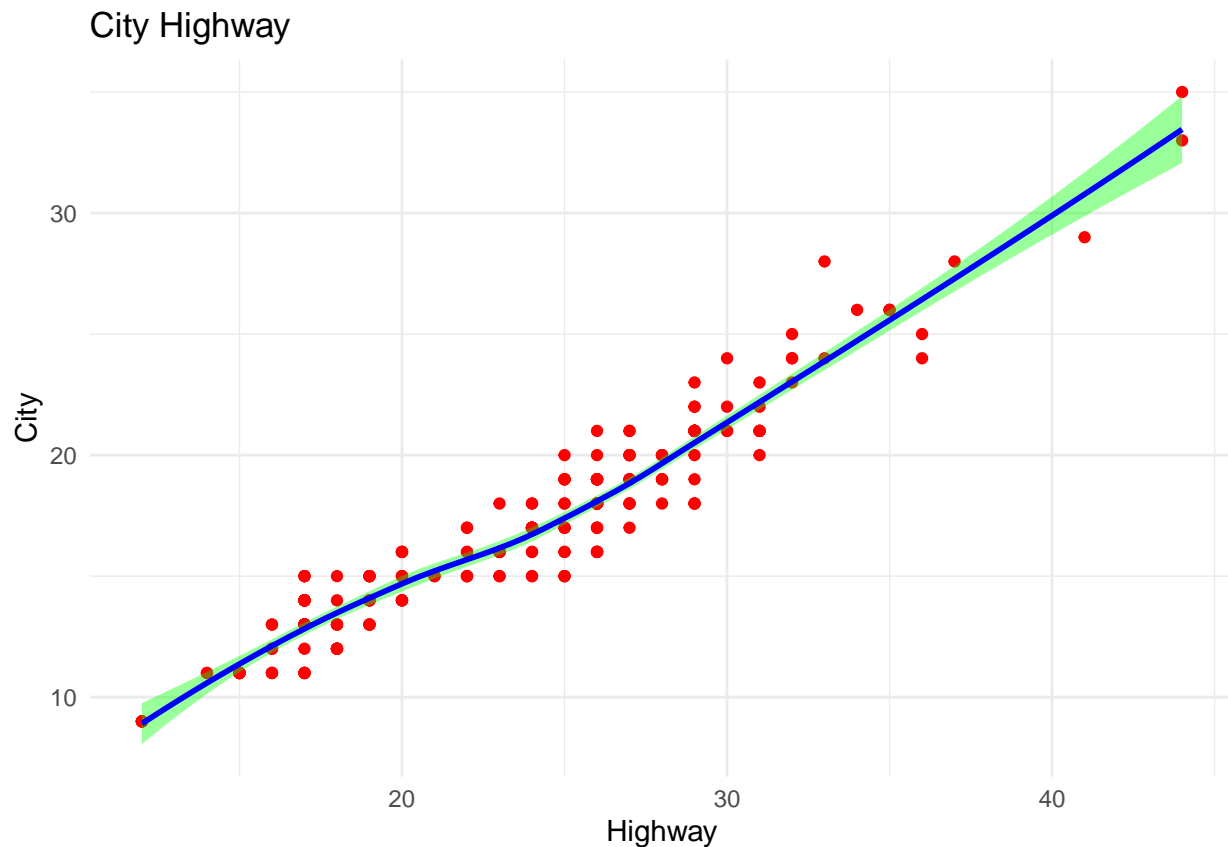
City Highway

Fig. 7 shown distance between city and highway.

Suggest car if you want both city and highway

```
sug_car <- mpg %>%
  filter((hwy > 25 & hwy < 30) & (cty > 15 & cty < 20))

sug_car
```

```
## # A tibble: 46 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~ f        18    29 p     comp~
##  2 audi         a4           2.8  1999     6 auto~ f        16    26 p     comp~
##  3 audi         a4           2.8  1999     6 manu~ f        18    26 p     comp~
##  4 audi         a4           3.1  2008     6 auto~ f        18    27 p     comp~
##  5 audi         a4 quattro   1.8  1999     4 manu~ 4        18    26 p     comp~
##  6 audi         a4 quattro   2    2008     4 auto~ 4        19    27 p     comp~
##  7 chevrolet    corvette     5.7  1999     8 manu~ r        16    26 p     2sea~
##  8 chevrolet    corvette     6.2  2008     8 manu~ r        16    26 p     2sea~
##  9 chevrolet    malibu       2.4  1999     4 auto~ f        19    27 r     mids~
## 10 chevrolet    malibu       3.1  1999     6 auto~ f        18    26 r     mids~
## # i 36 more rows
```