ggplot2

Chia-Yu Wei

2023-11-28

ggplot2

The package which is useful and popular to be used on data visualization.

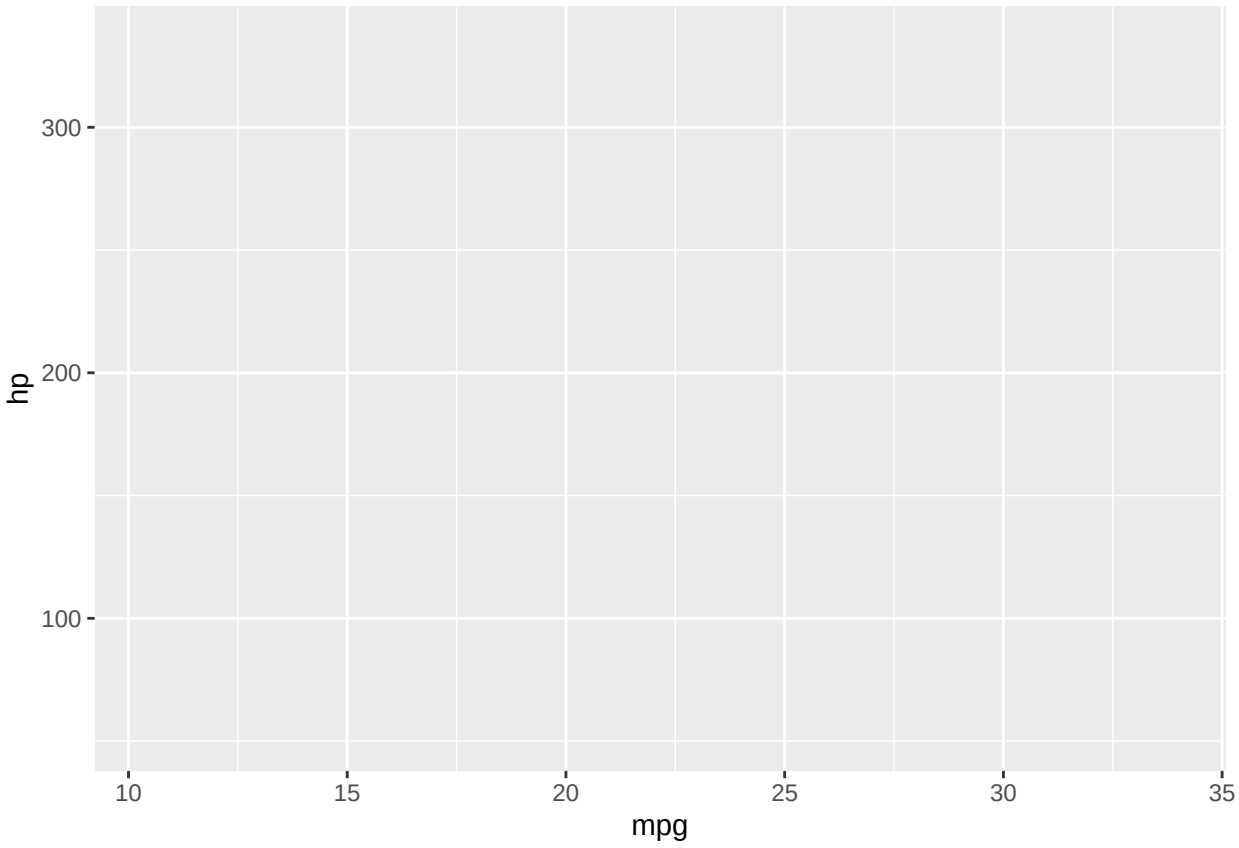
```
library(ggplot2)
```

Data (no plot shown yet)

```
ggplot(data = mtcars)
```

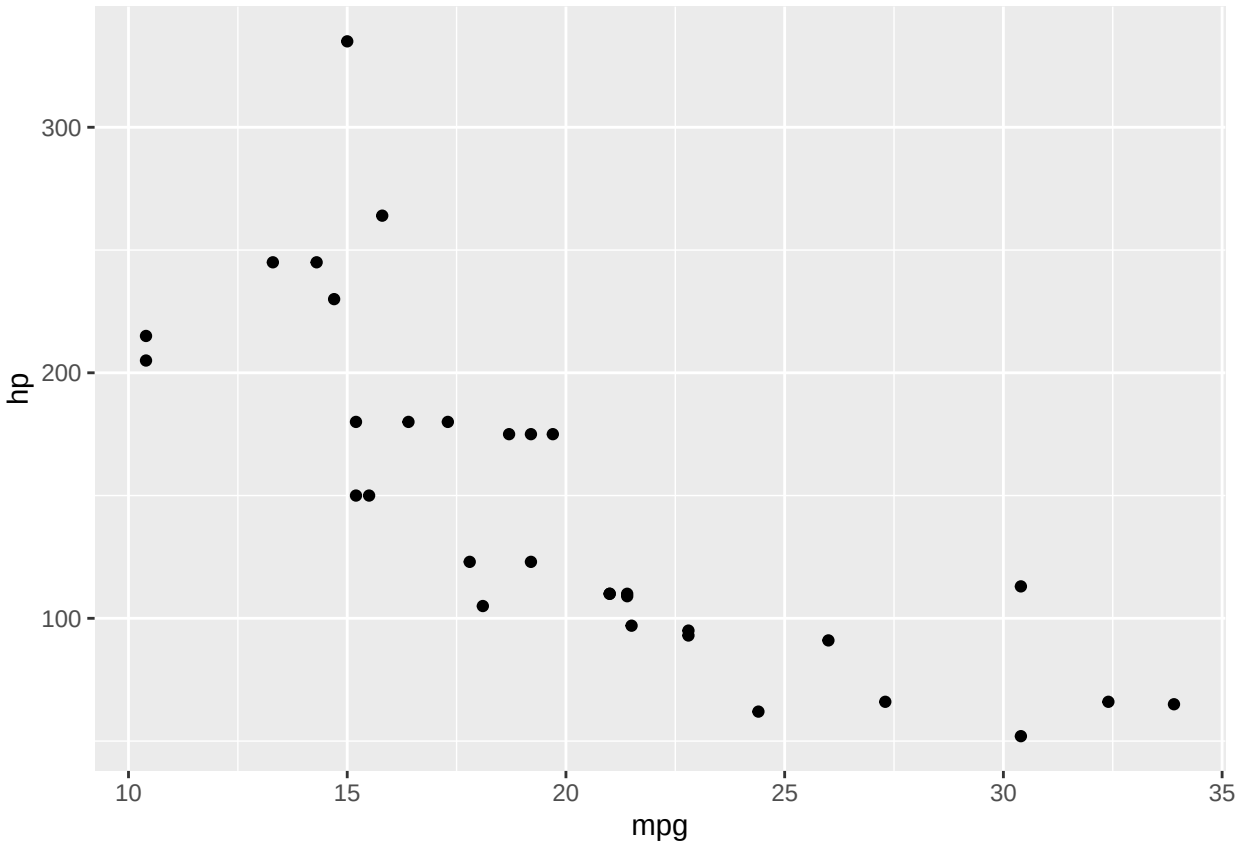

Data and Aesthetics

```
ggplot(data = mtcars, aes(x = mpg, y = hp))
```



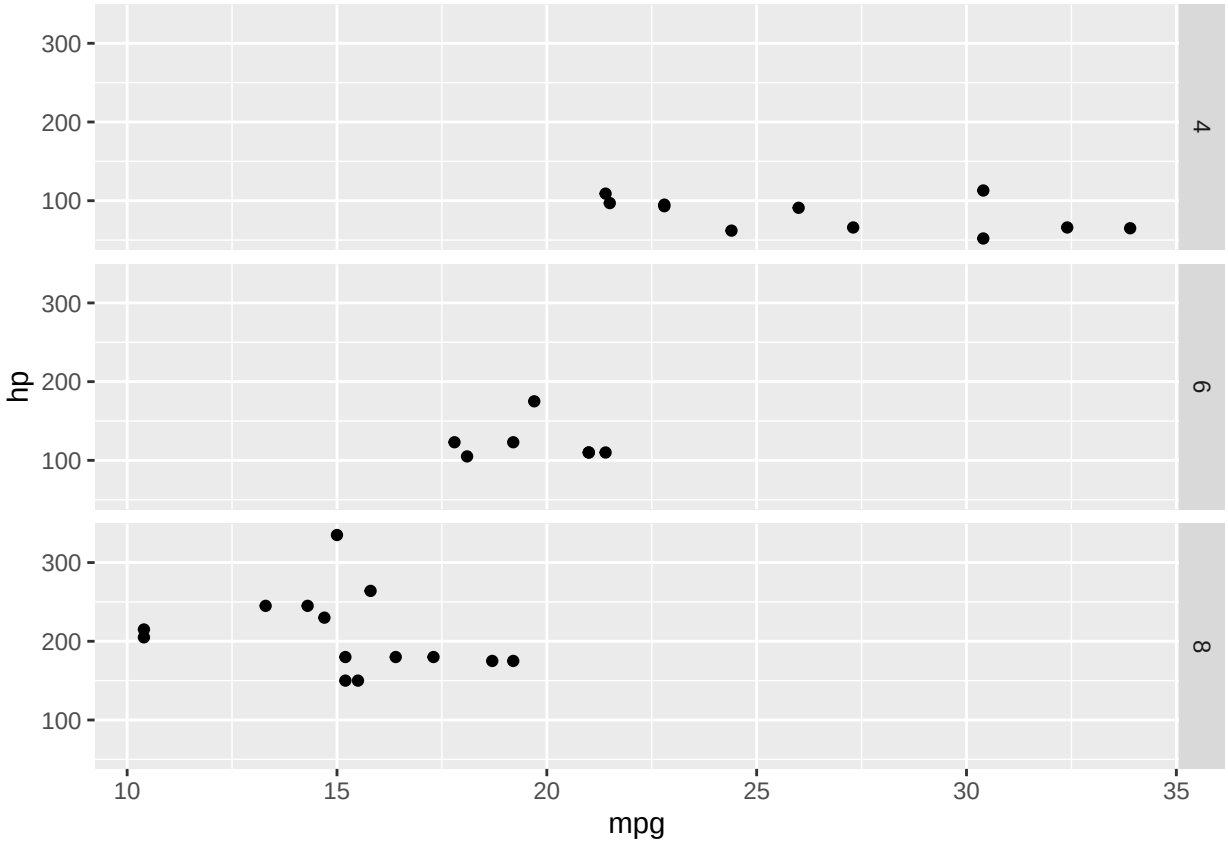
Data, Aesthetics, and Geometries

```
pl <- ggplot(data = mtcars, aes(x = mpg, y = hp))
pl + geom_point()
```



Data, Aesthetics, Geometries, and Facets

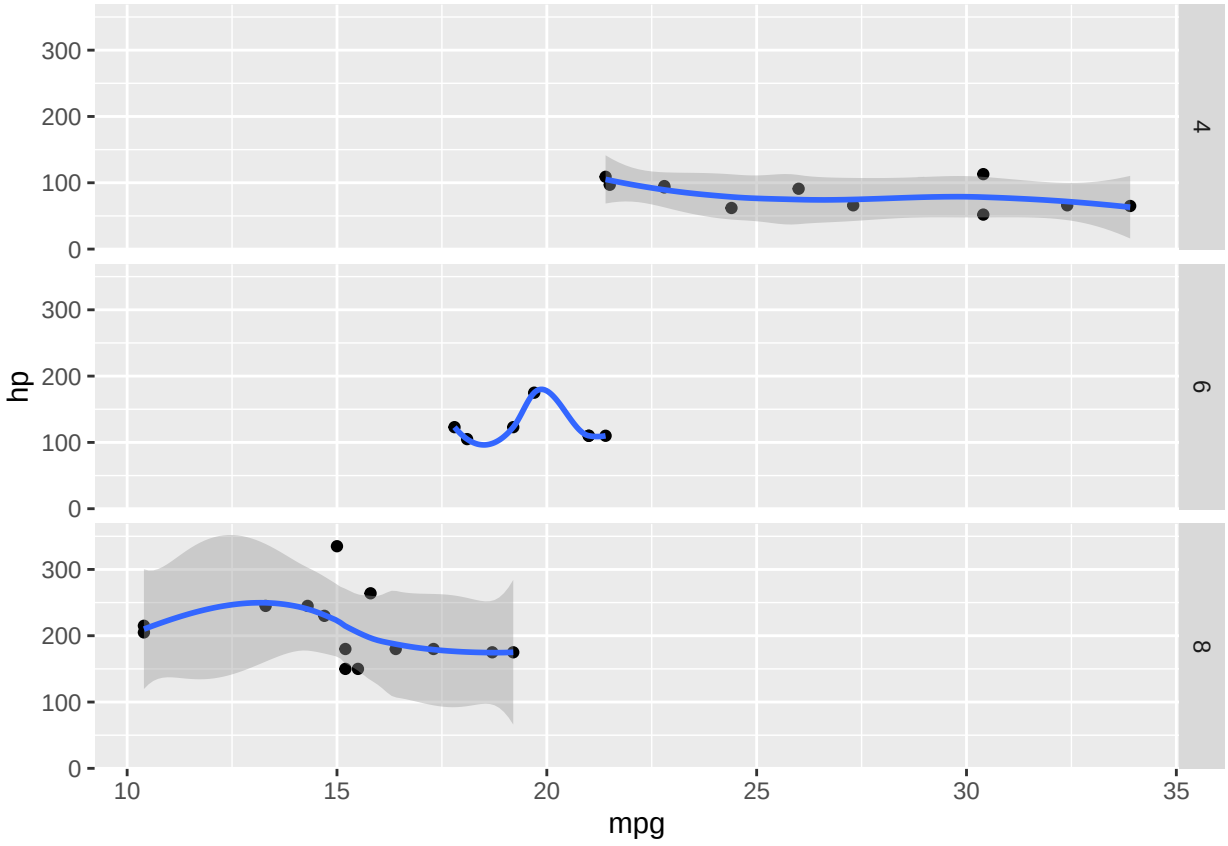
```
pl <- ggplot(data = mtcars, aes(x = mpg, y = hp)) +
  geom_point()
pl + facet_grid(cyl ~ .)
```



Data, Aesthetics, Geometries, Facets, and Statistics

```
pl <- ggplot(data = mtcars, aes(x = mpg, y = hp)) +
  geom_point()
pl + facet_grid(cyl ~ .) + stat_smooth()
```

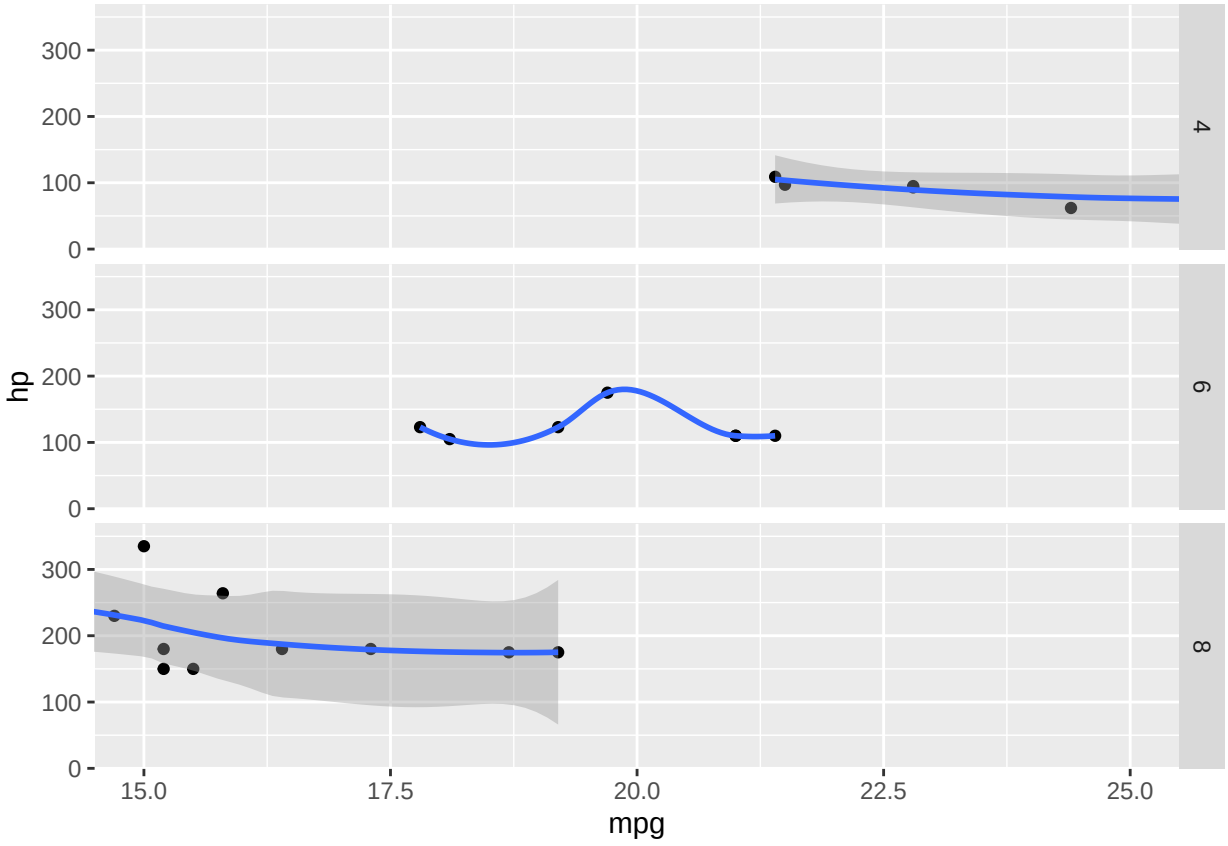
```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



Data, Aesthetics, Geometries, Facets, Statistics, and Coordinates

```
pl <- ggplot(data = mtcars, aes(x = mpg, y = hp)) +
  geom_point()
pl2 <- pl + facet_grid(cyl ~ .) + stat_smooth()
pl2 + coord_cartesian(xlim = c(15, 25))
```

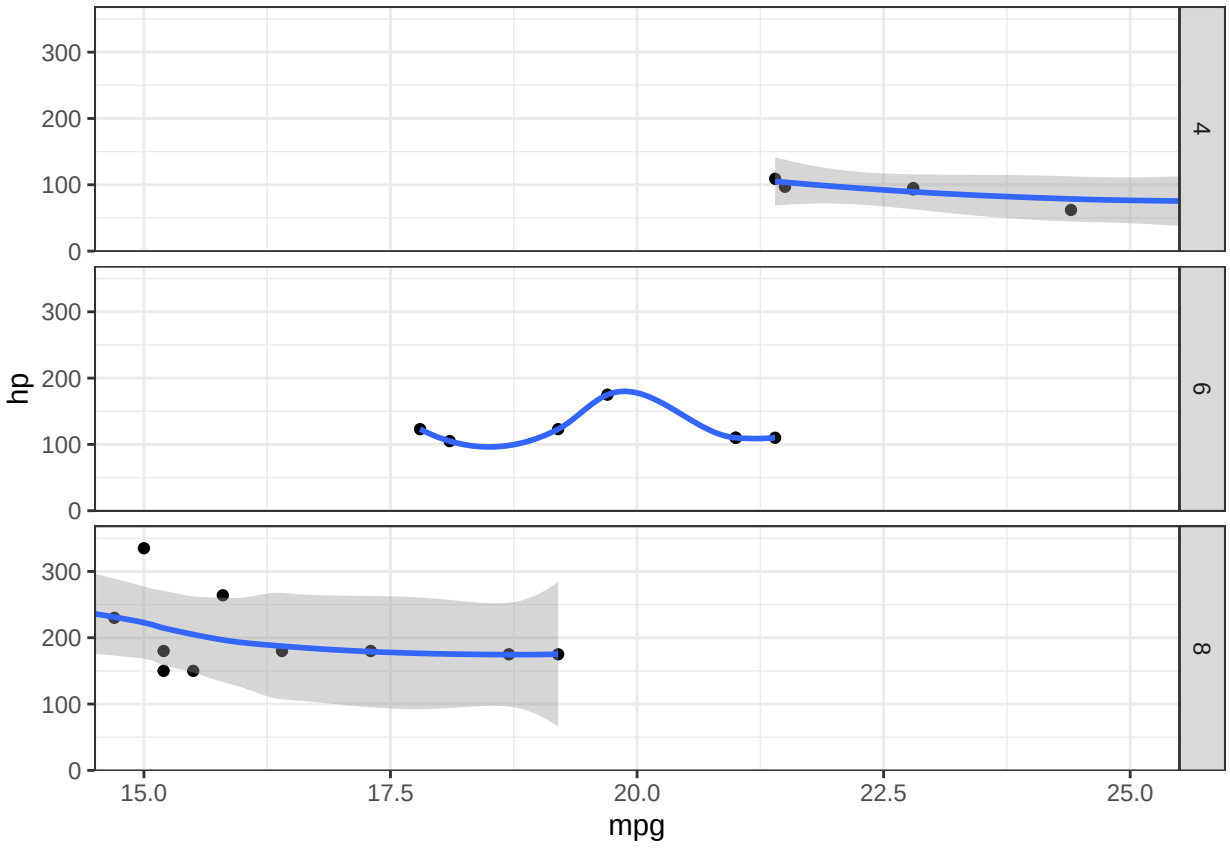
```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



Data, Aesthetics, Geometries, Facets, Statistics, Coordinates, and Theme

```
pl <- ggplot(data = mtcars, aes(x = mpg, y = hp)) +
  geom_point()
pl2 <- pl + facet_grid(cyl ~ .) + stat_smooth()
pl2 + coord_cartesian(xlim = c(15, 25)) + theme_bw()
```

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


hp
mpg
4
6
8
0
100
200
300
15.0
17.5
20.0
22.5
25.0

Histograms

Here is a Rstudio cheat sheet (https://statsandr.com/blog/files/ggplot2-cheatsheet.pdf) Use: + geom_histogram()

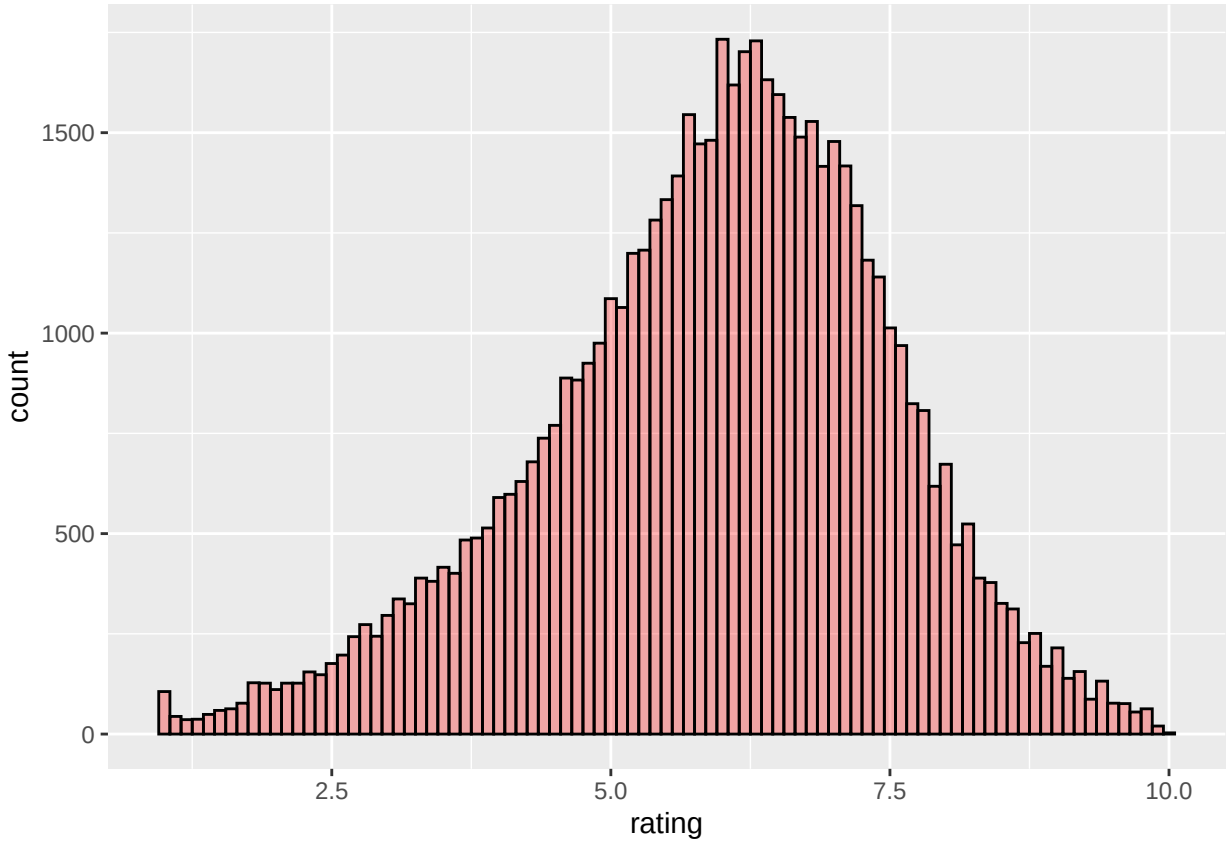
```
#install.packages('ggplot2movies')
library(ggplot2movies)
```

```
# Data + Aesthetics
colnames(movies)
```

```
##  [1] "title"       "year"        "length"      "budget"      "rating"
##  [6] "votes"       "r1"          "r2"          "r3"          "r4"
## [11] "r5"          "r6"          "r7"          "r8"          "r9"
## [16] "r10"         "mpaa"        "Action"      "Animation"   "Comedy"
## [21] "Drama"       "Documentary" "Romance"     "Short"
```

```
pl <- ggplot(movies, aes(x = rating))
```

```
# Data + Aesthetics + Geometries
pl2 <- pl + geom_histogram(binwidth = 0.1,
                           color = 'black',
                           fill = 'red',
                           alpha = 0.3)
print(pl2)
```

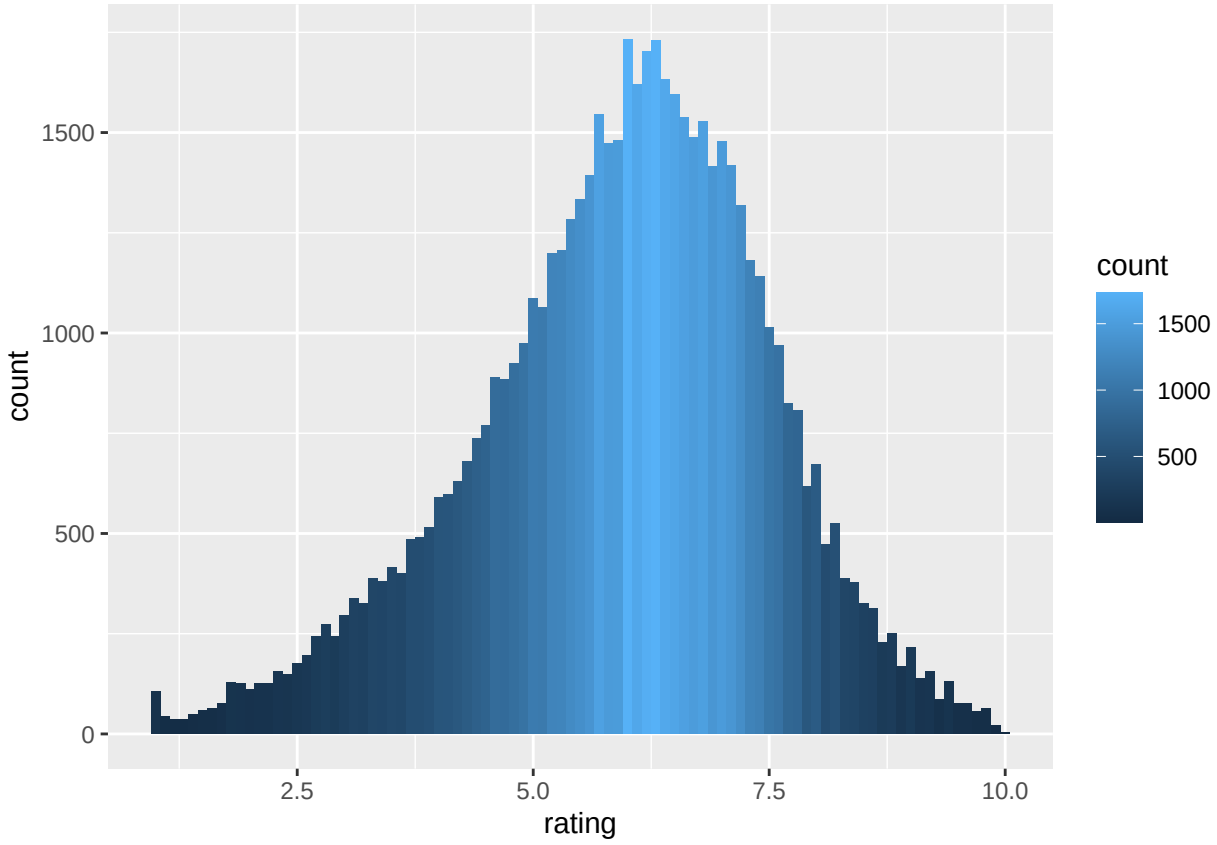


```
# Alpha: transparency
```

Adjusted Version (advanced)

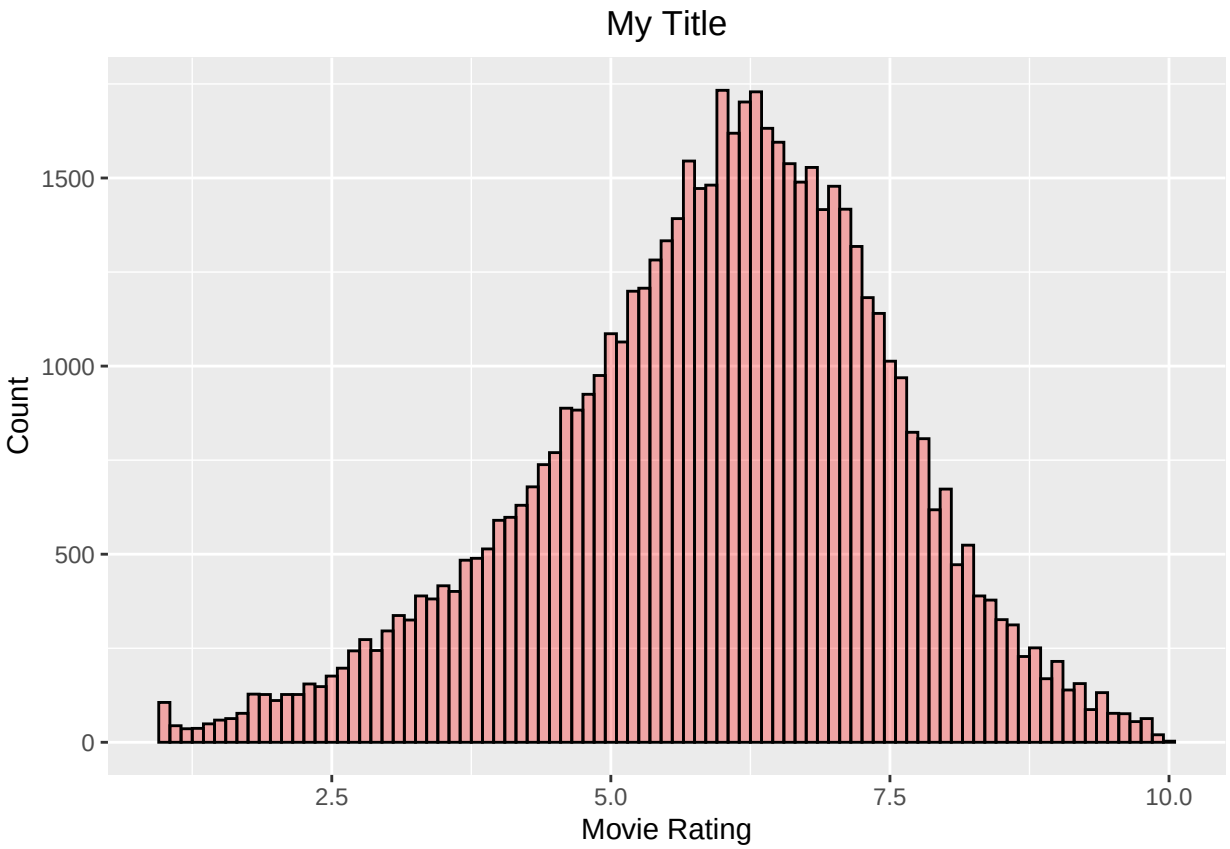
```
# Data + Aesthetics + Geometries
pl2_adj <- pl + geom_histogram(binwidth = 0.1,
                               aes(fill = ..count..))
print(pl2_adj)
```

```
## Warning: The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(count)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



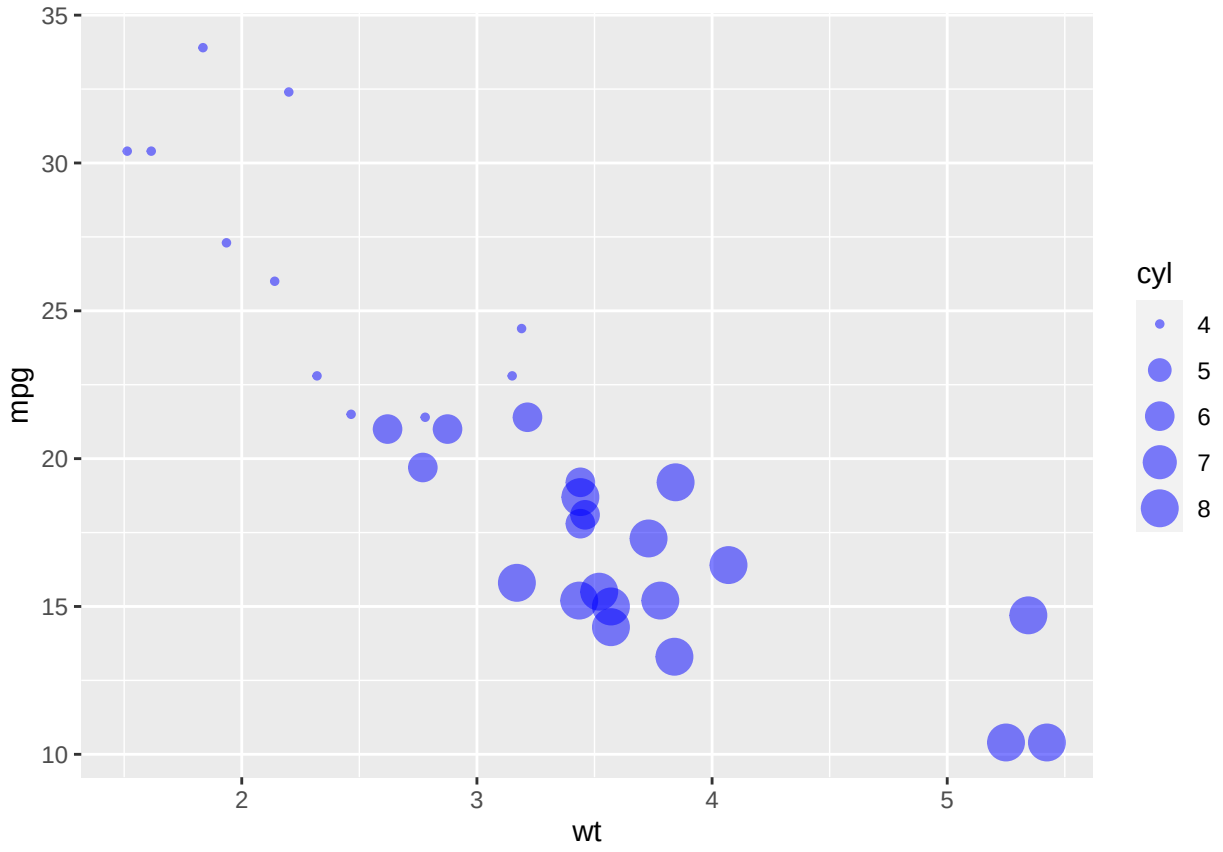

```
# Alpha: transparency
```

```
# Make it to center
(pl3 <- pl2 + xlab('Movie Rating') + ylab('Count') + ggtitle("My Title") + theme(plot.title = element_t
```


My Title
Count
1500
1000
500
0
2.5
5.0
7.5
10.0
Movie Rating

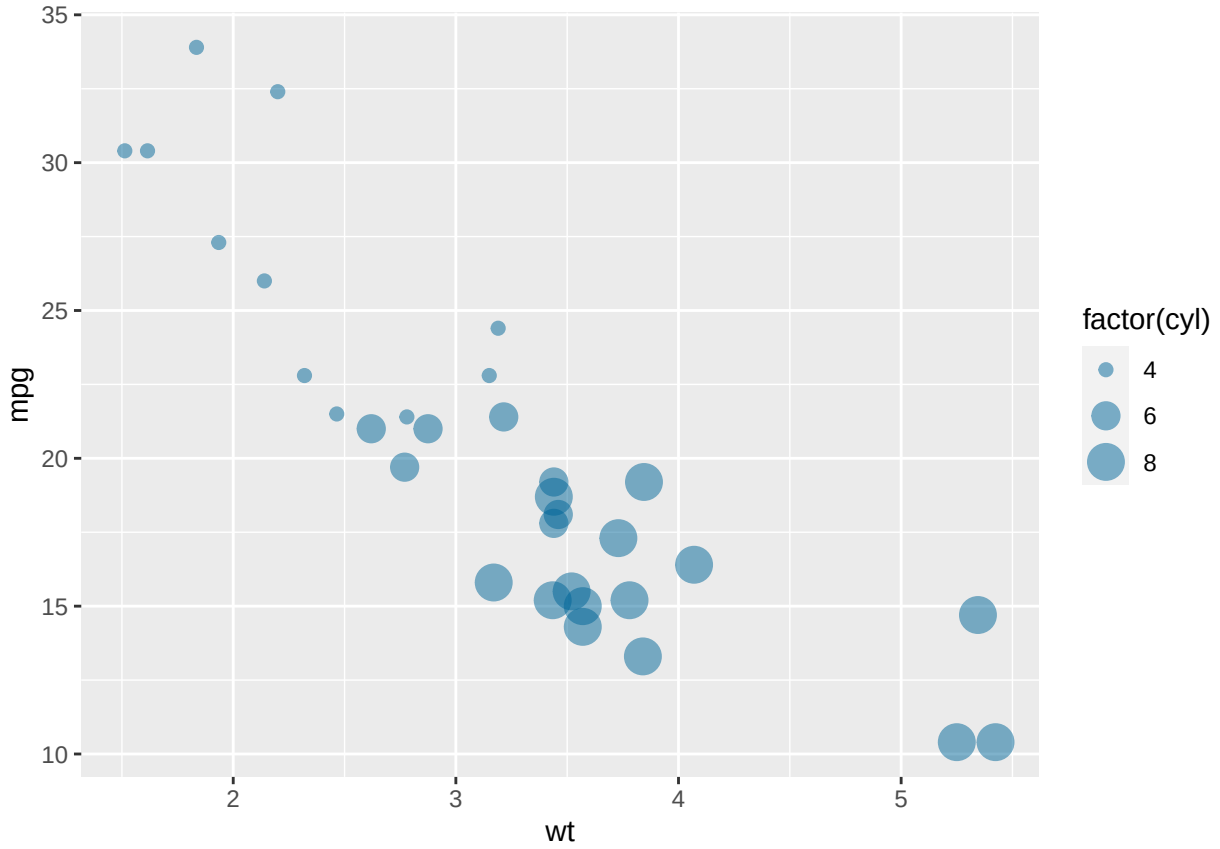
Plot

```
library(ggplot2)
df <- mtcars

# Data & Aesthetics
pl <- ggplot(df, aes(x = wt, y = mpg))

# Geometric
pl + geom_point(aes(size = cyl),  # The size changes with the 'cyl' data.
                col = 'blue',
                alpha = 0.5)
```




```
pl + geom_point(aes(size = factor(cyl)),  # The size changes with the 'cyl' data.
                col = '#006699',
                alpha = 0.5)
```

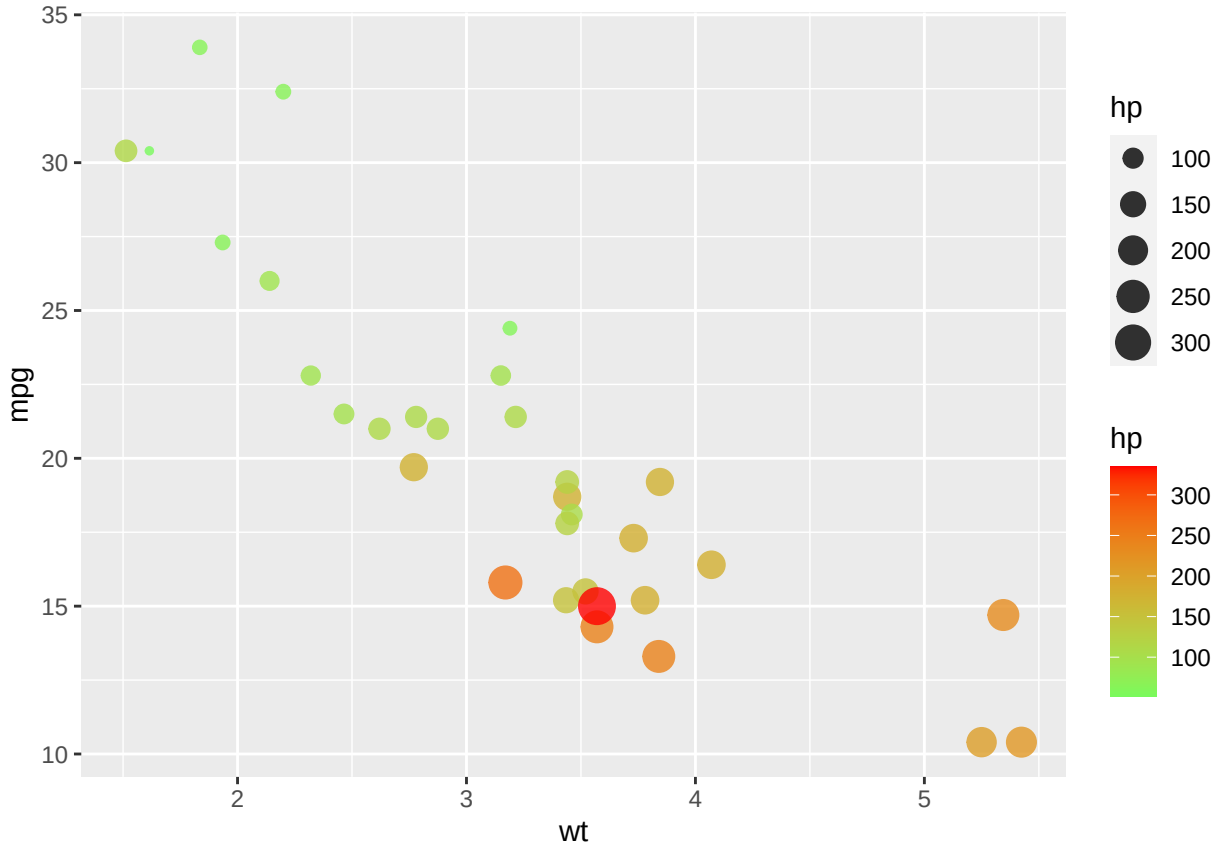
```
## Warning: Using size for a discrete variable is not advised.
```




```
# factor() remove the category that we don't use.
```

Shape

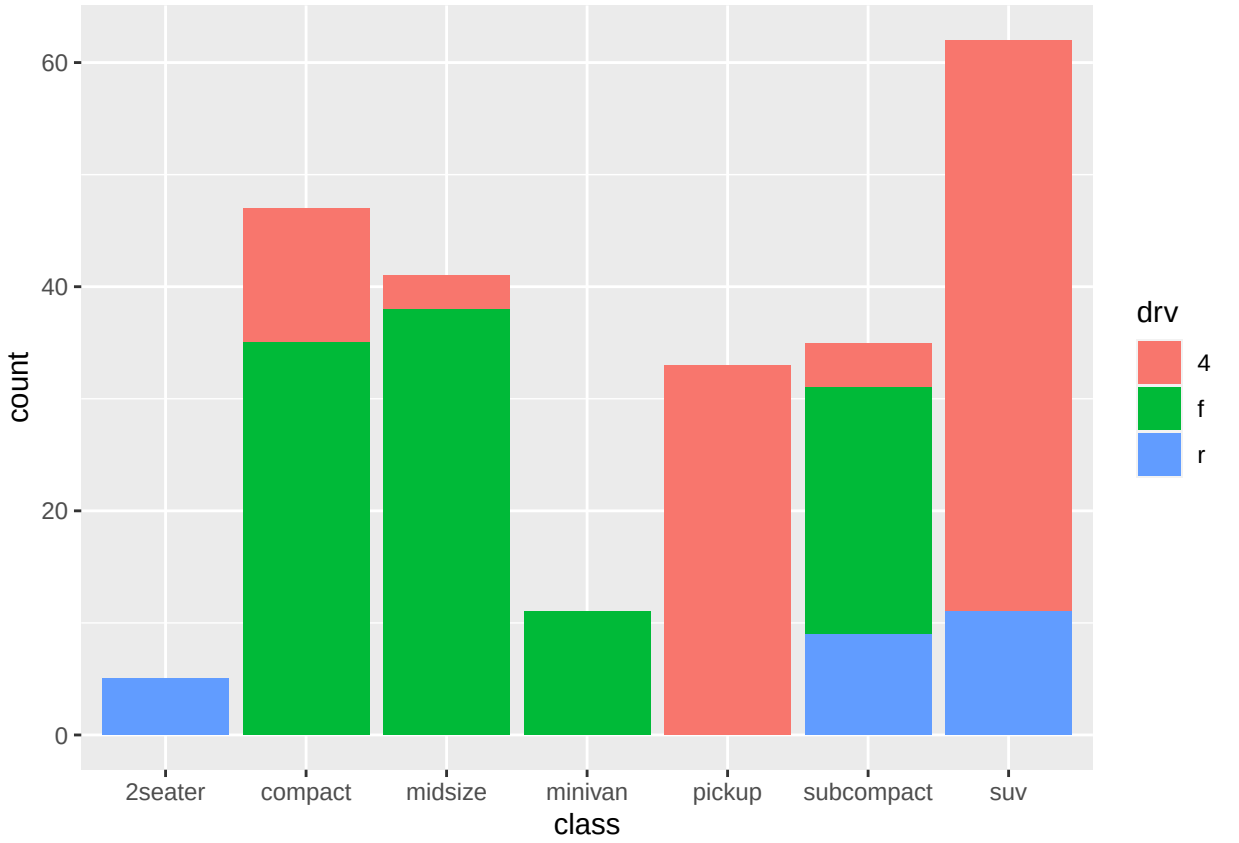
```
# The shape and color changes with the 'cyl' data.
pl2 <- pl + geom_point(aes(size = hp, color = hp),
                       alpha = 0.8)
pl2 + scale_color_gradient(low = '#78fb5c', high = 'red')
```



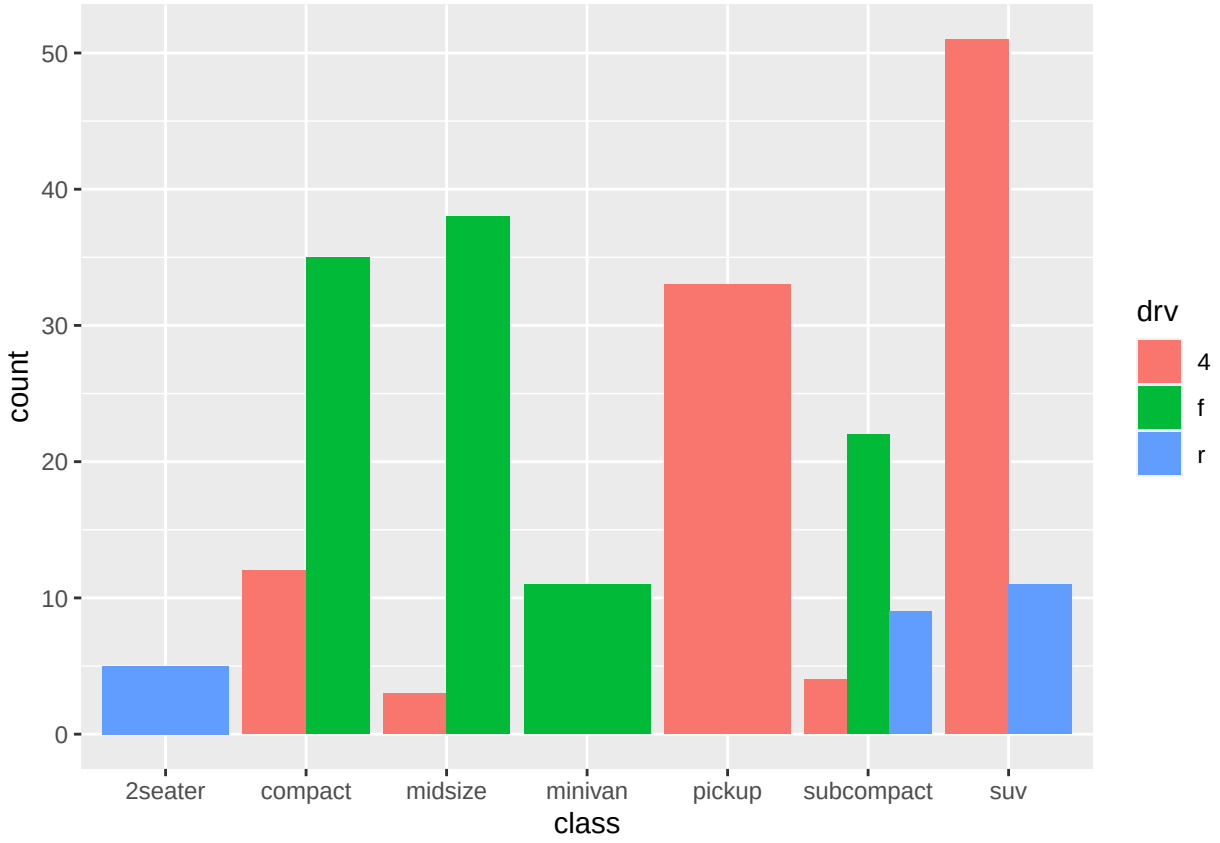
Here is the link to ggplot color code:(https://redketchup.io/color-picker)

Boxplot

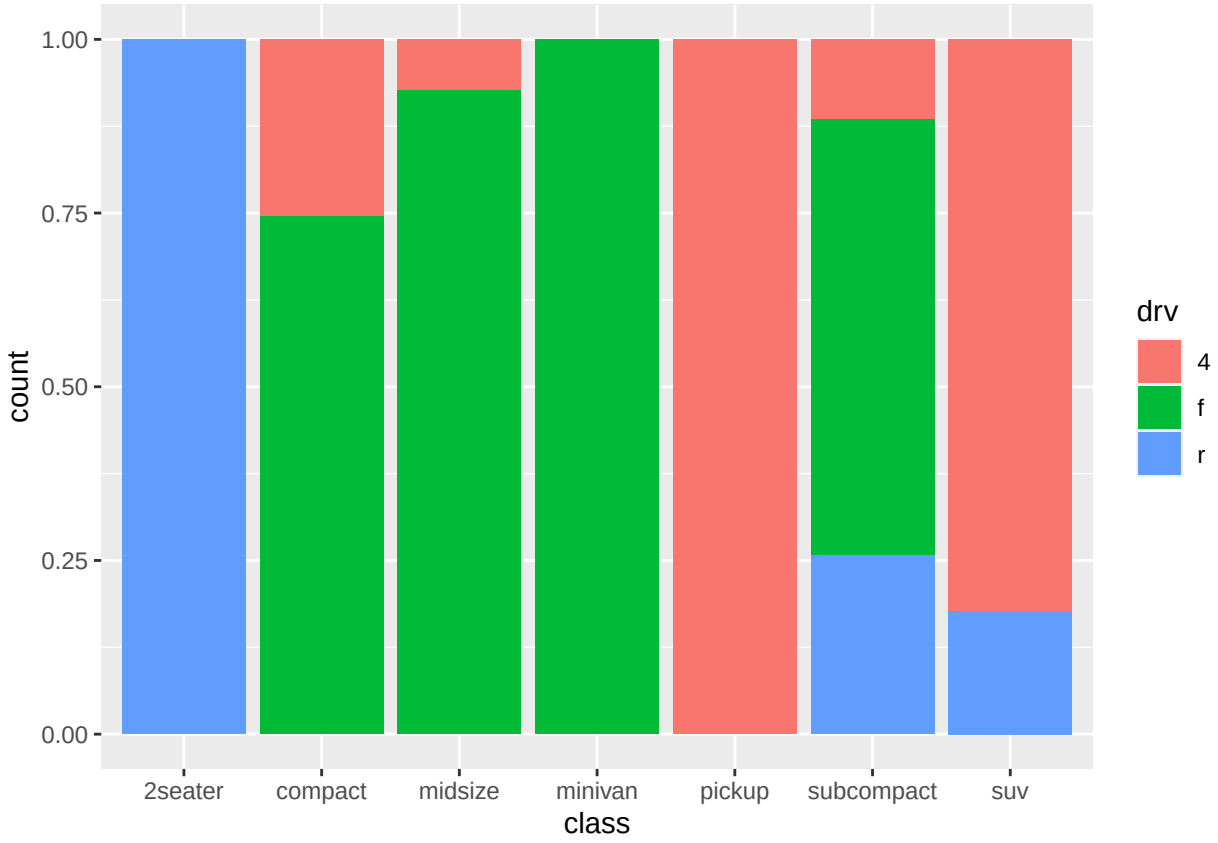
```
df <- mpg

pl <- ggplot(df, aes(x = class))
pl + geom_bar(aes(fill = drv))
```




```
pl + geom_bar(aes(fill = drv), position = "dodge")
```

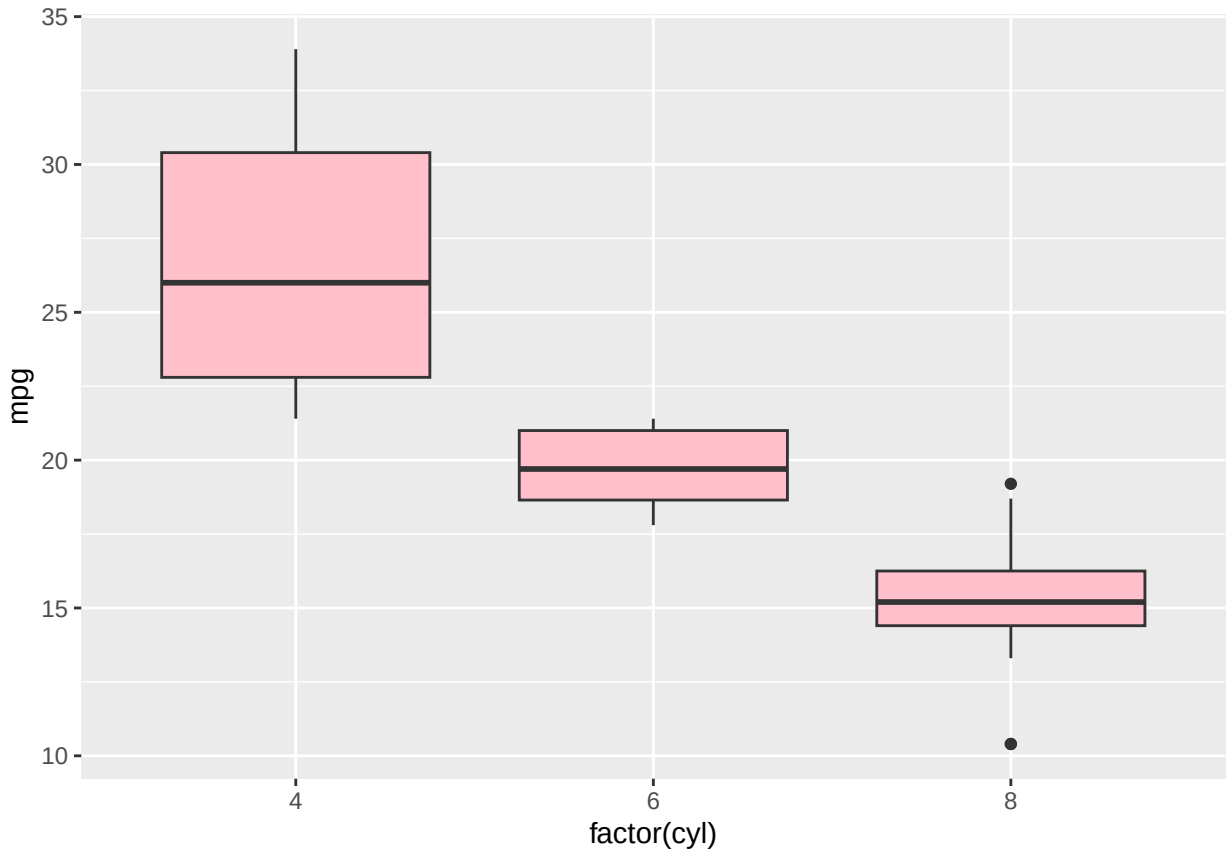



```
pl + geom_bar(aes(fill = drv), position = "fill")
```

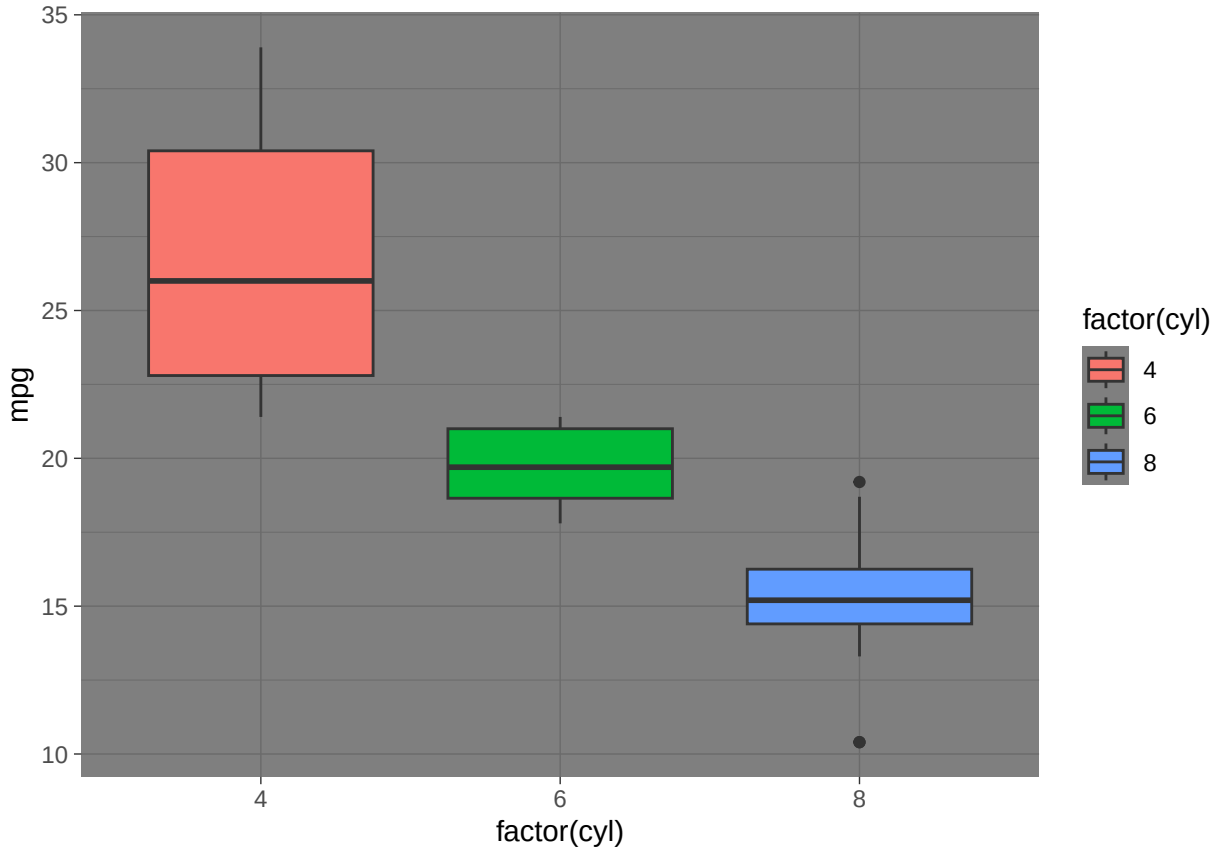

1.00
0.75
0.50
0.25
0.00
count
2seater
compact
midsize
minivan
pickup
subcompact
suv
class
drv
4
f
r

Boxplots

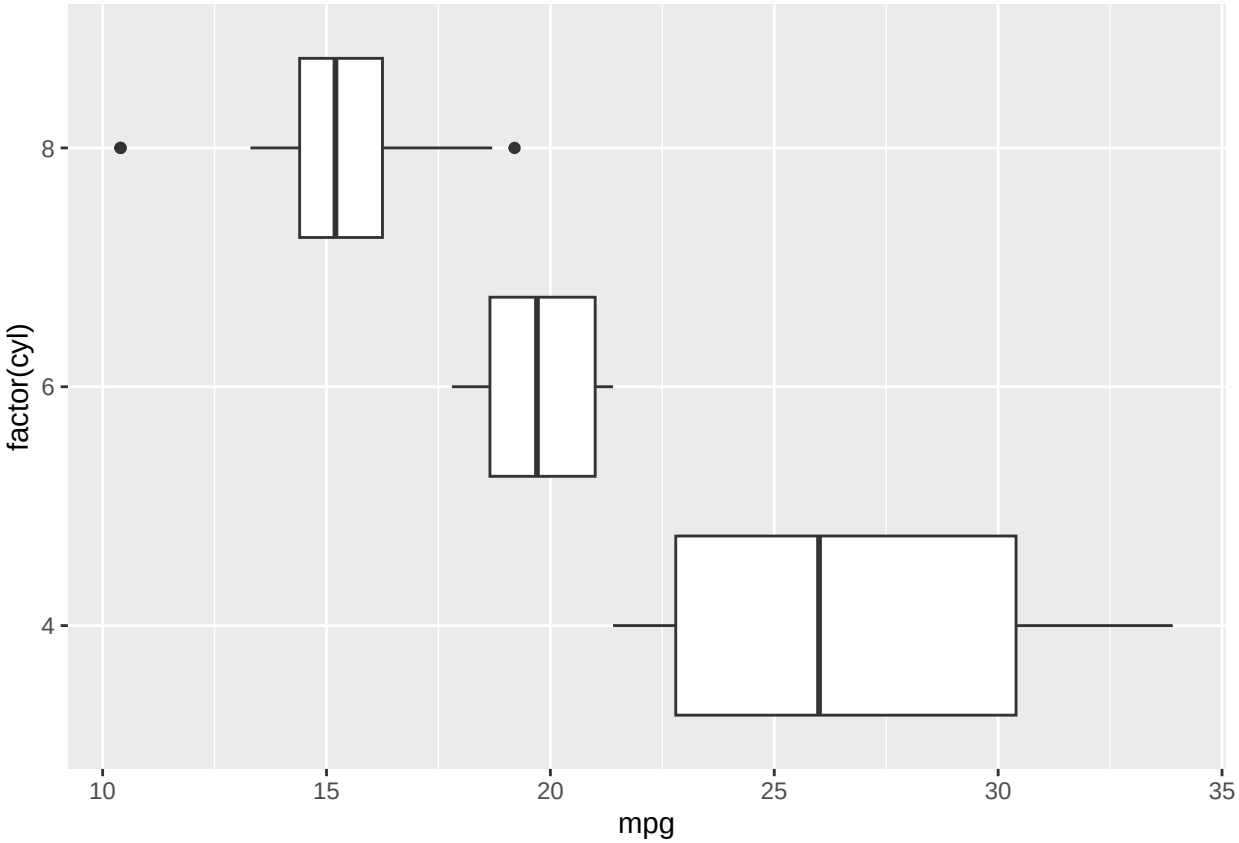
```
df <- mtcars
# X value should not be continuous vale, it should be category variable,
pl <- ggplot(df, aes(x = factor(cyl), y = mpg))
pl + geom_boxplot(fill = 'pink')
```




```
pl + geom_boxplot(aes(fill = factor(cyl))) + theme_dark()
```




```
# Flip the coordinate
pl + geom_boxplot() +coord_flip()
```


factor(cyl)
8
6
4
10
15
20
25
30
35
mpg