Data Visualization with R

Agenda

- understand grammar of graphics
- understand geoms
- explore aesthetics to modify geoms
- use facets for sub plots
- change coordinate system

Libraries

```
library(ggplot2)
library(readr)
library(descriptr)
```

Coordinate System

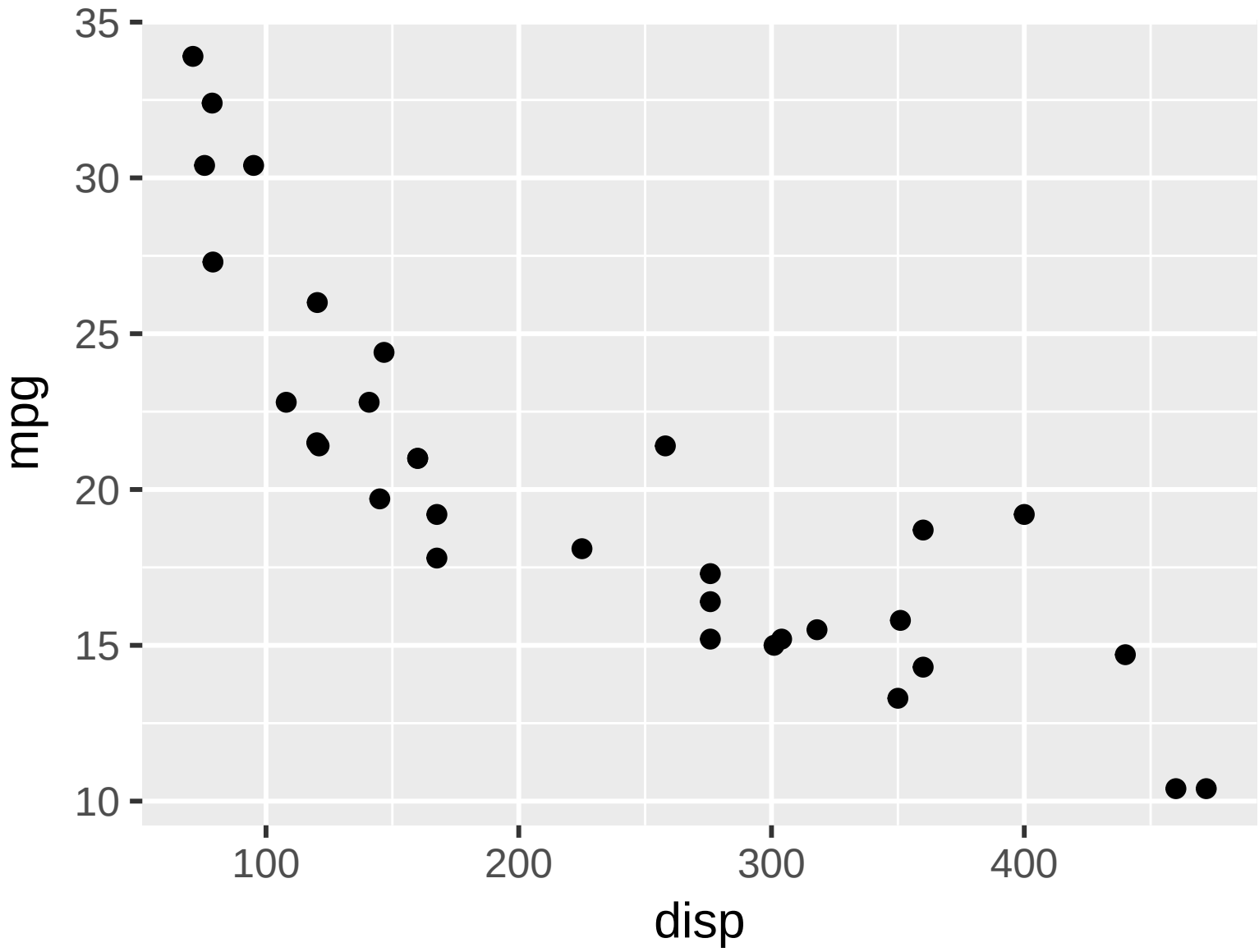
```
ggplot()
```

Data Set

```
ggplot(data = mtcarz)
```

Geoms

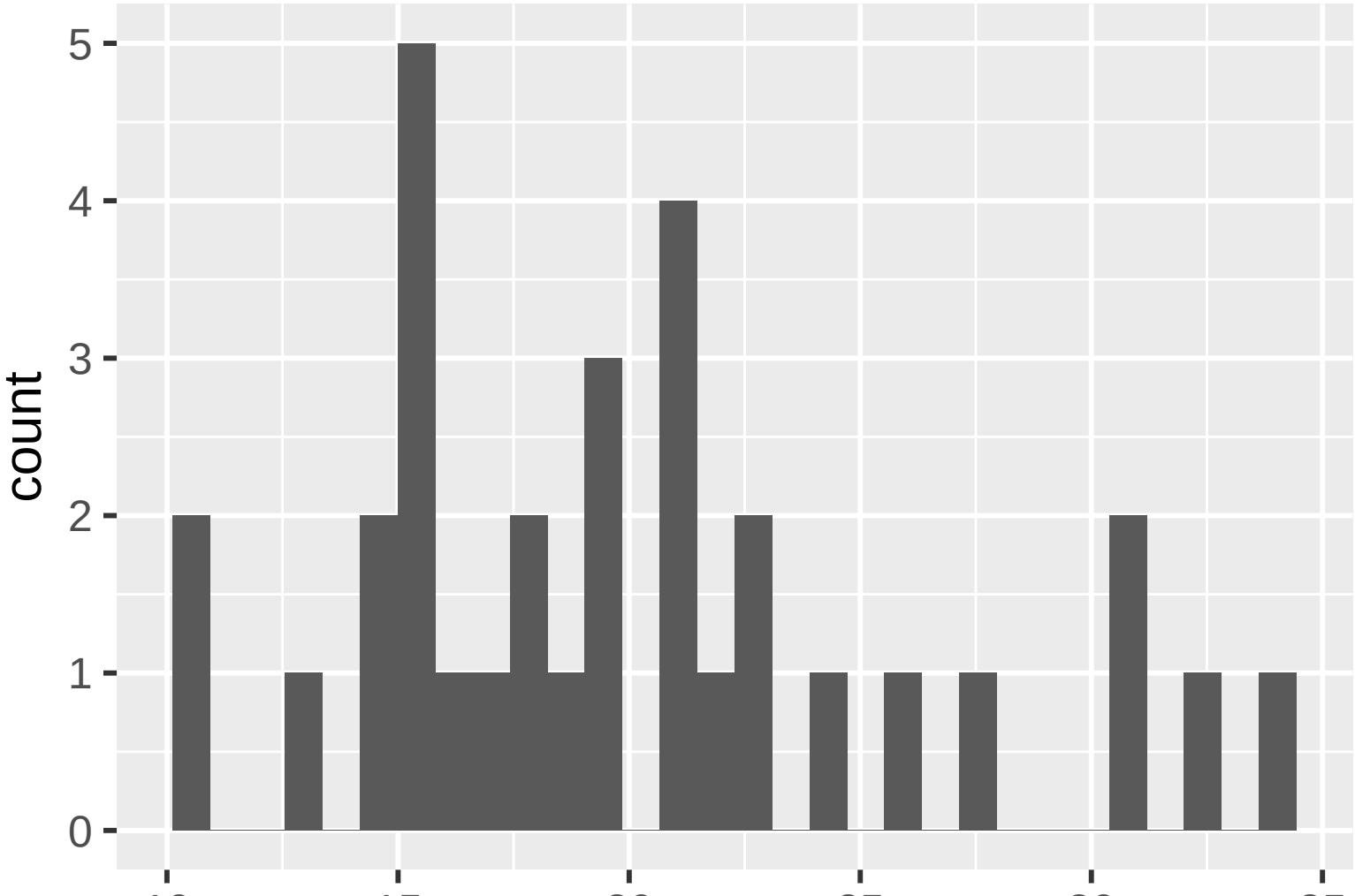
```
ggplot(data = mtcarz) +
  geom_point(mapping = aes(x = disp, y = mpg))
```



Geoms

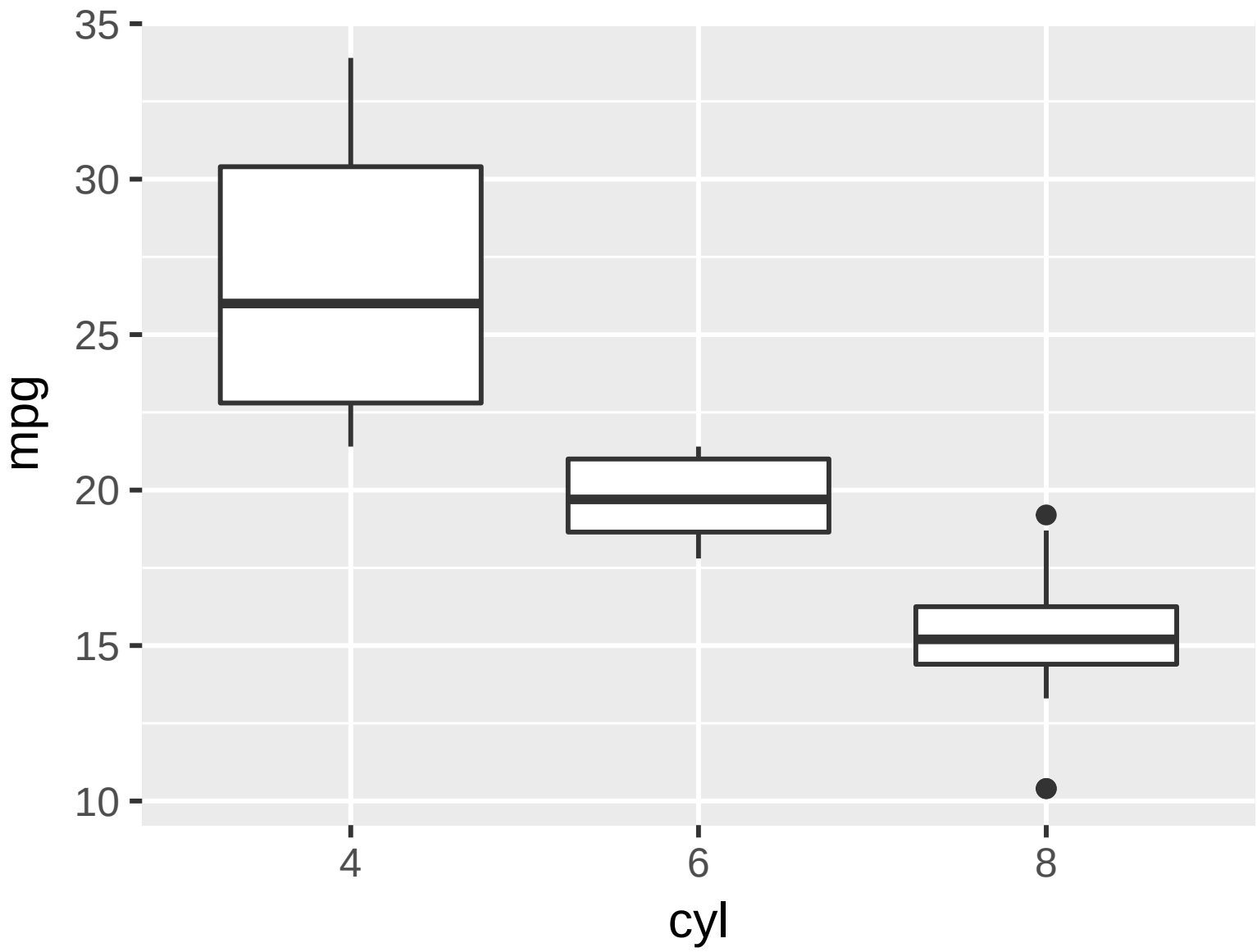
```
ggplot(data = mtcarz) +
  geom_histogram(mapping = aes(x = mpg))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `
```



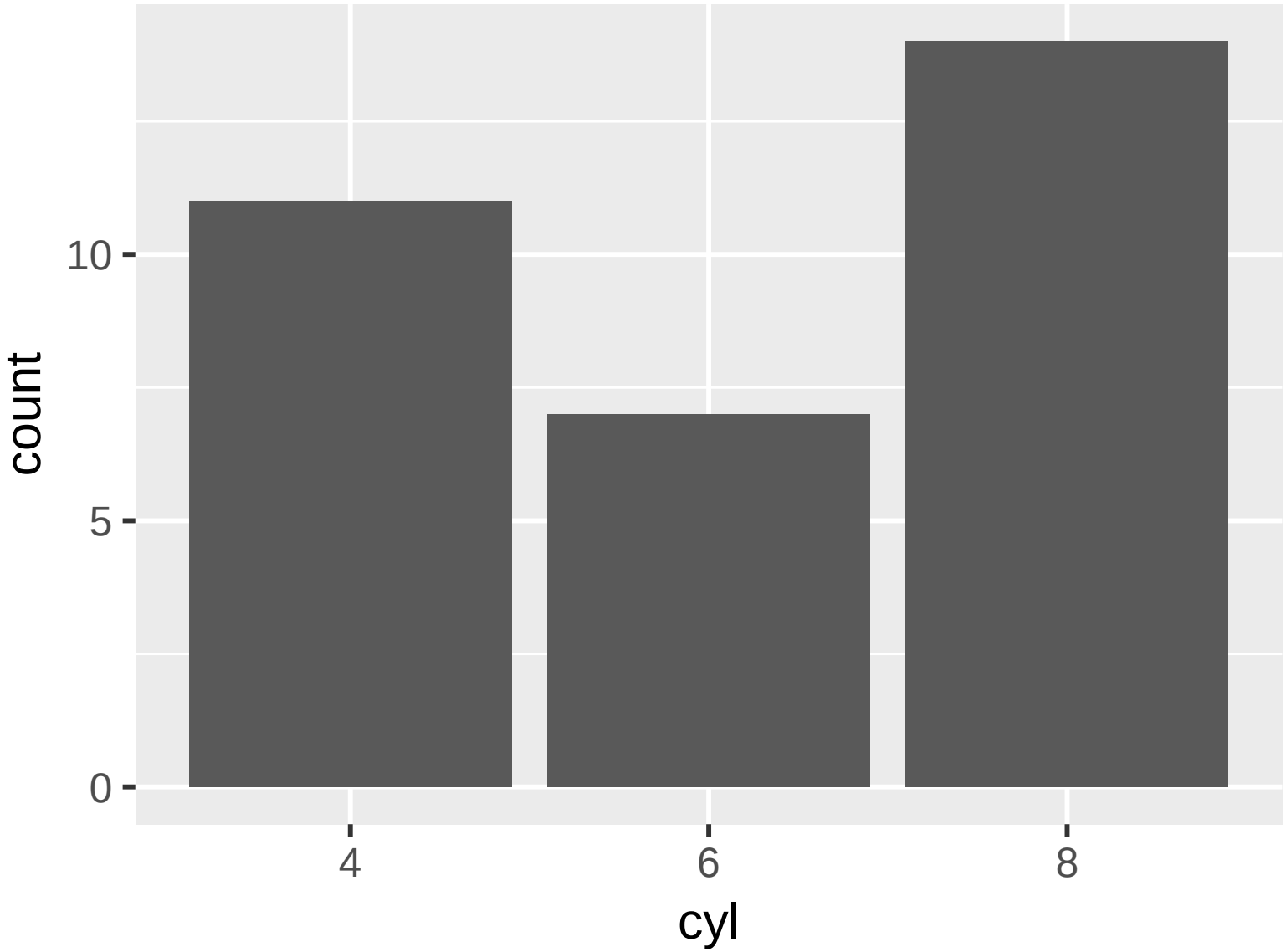
Geoms

```
ggplot(data = mtcarz) +
  geom_boxplot(mapping = aes(x = cyl, y = mpg))
```



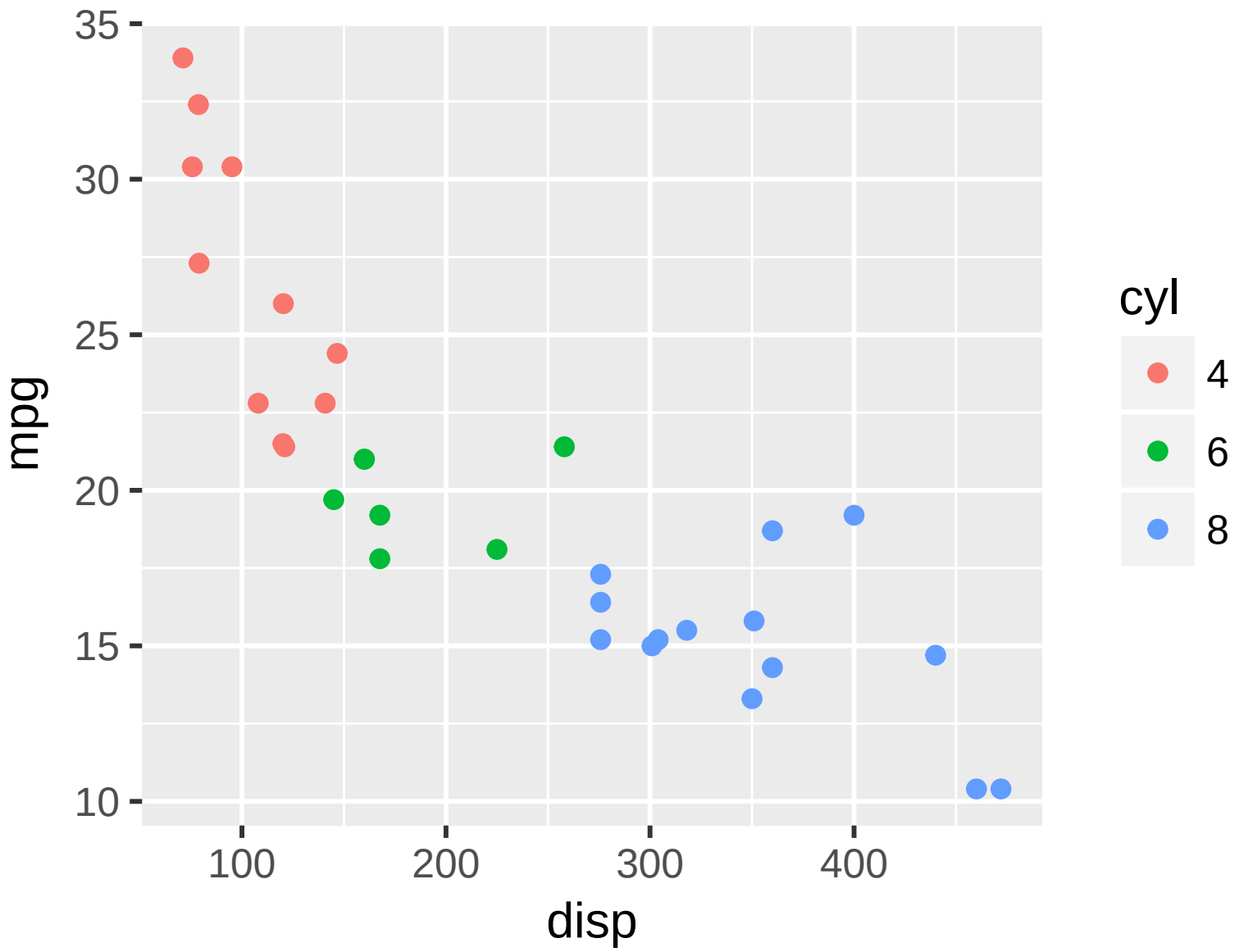
Geoms

```
ggplot(data = mtcarz) +
  geom_bar(mapping = aes(x = cyl))
```



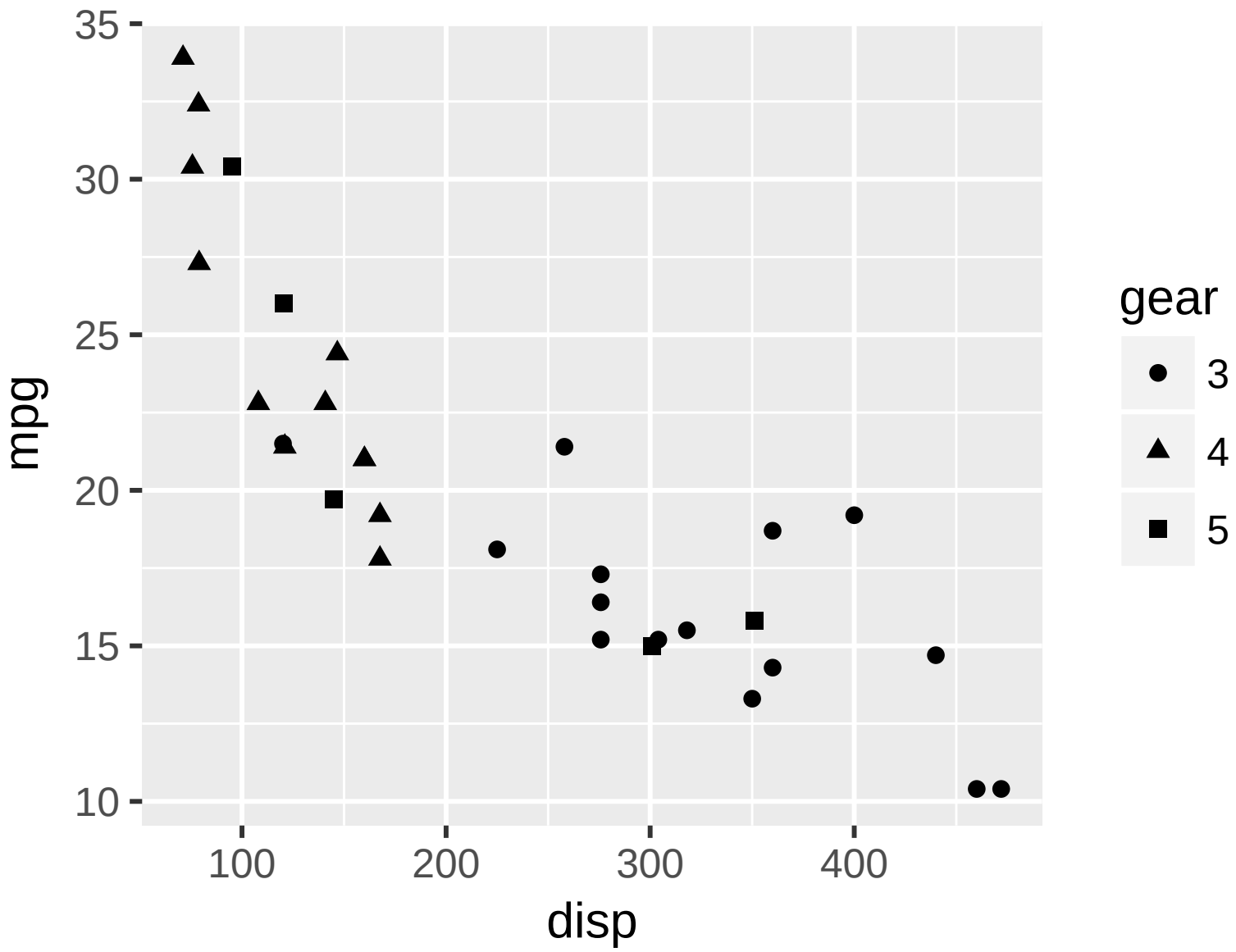
Map Color

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg, color = cyl))
```



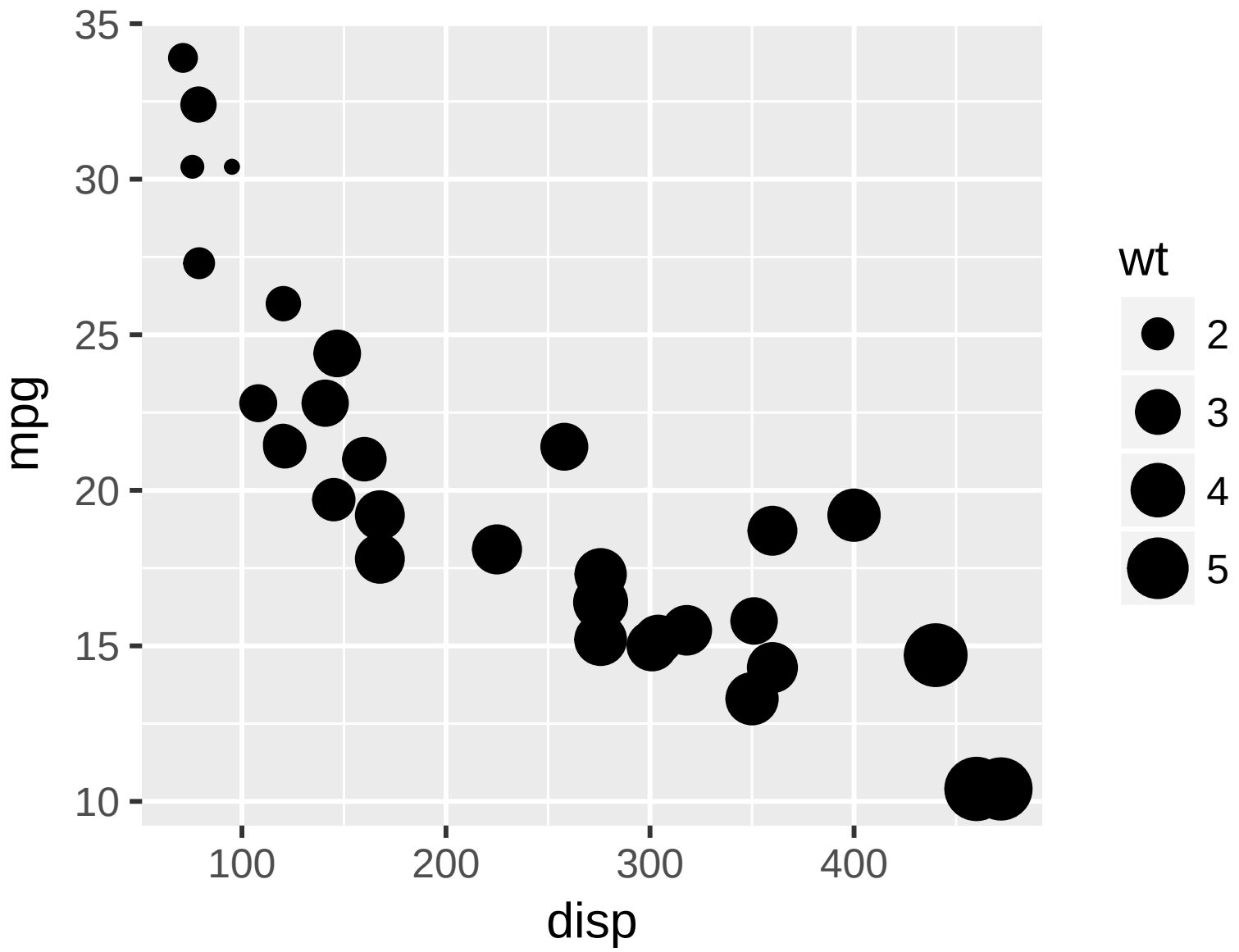
Map Shape

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg, shape = gear))
```



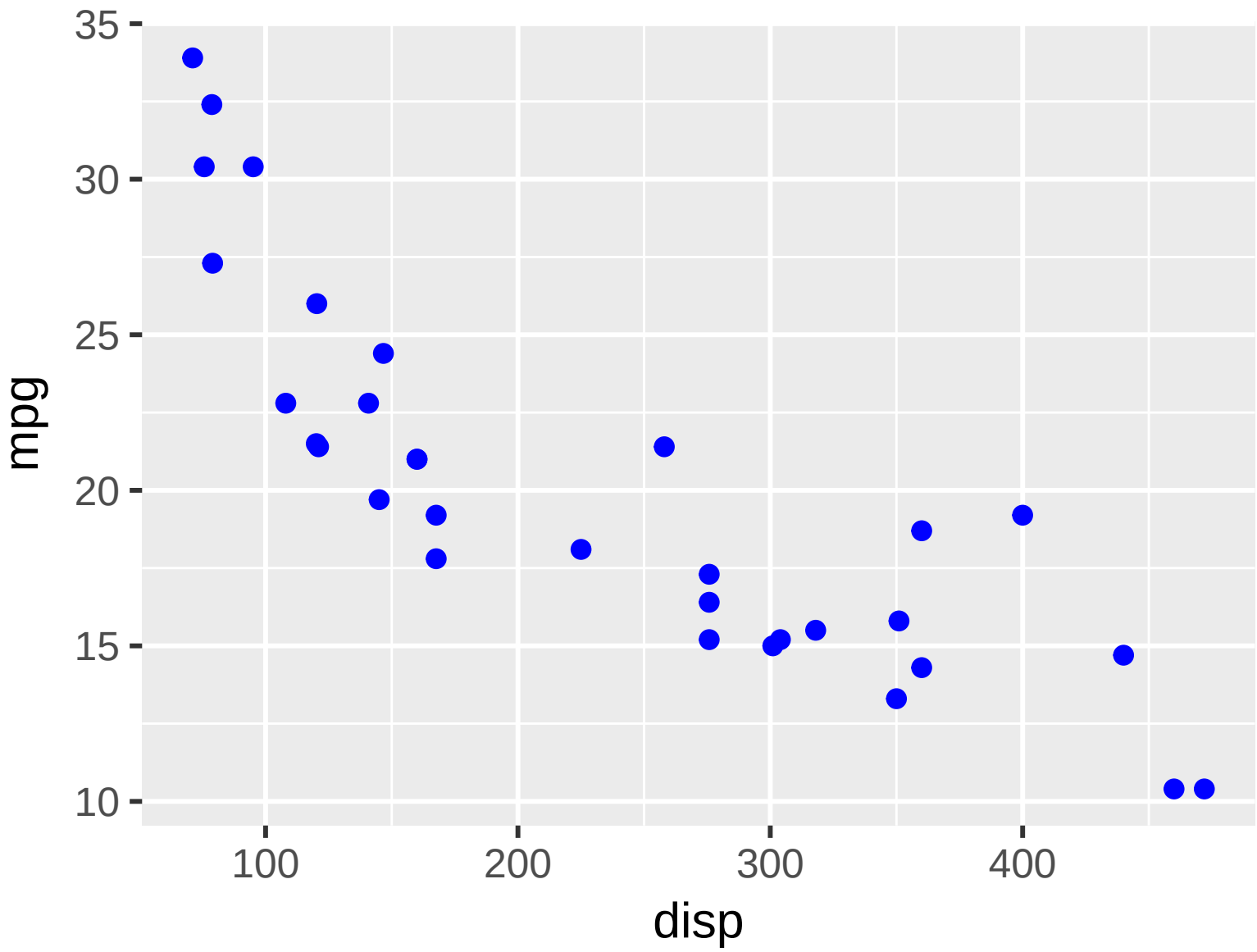
Map Size

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg, size = wt))
```



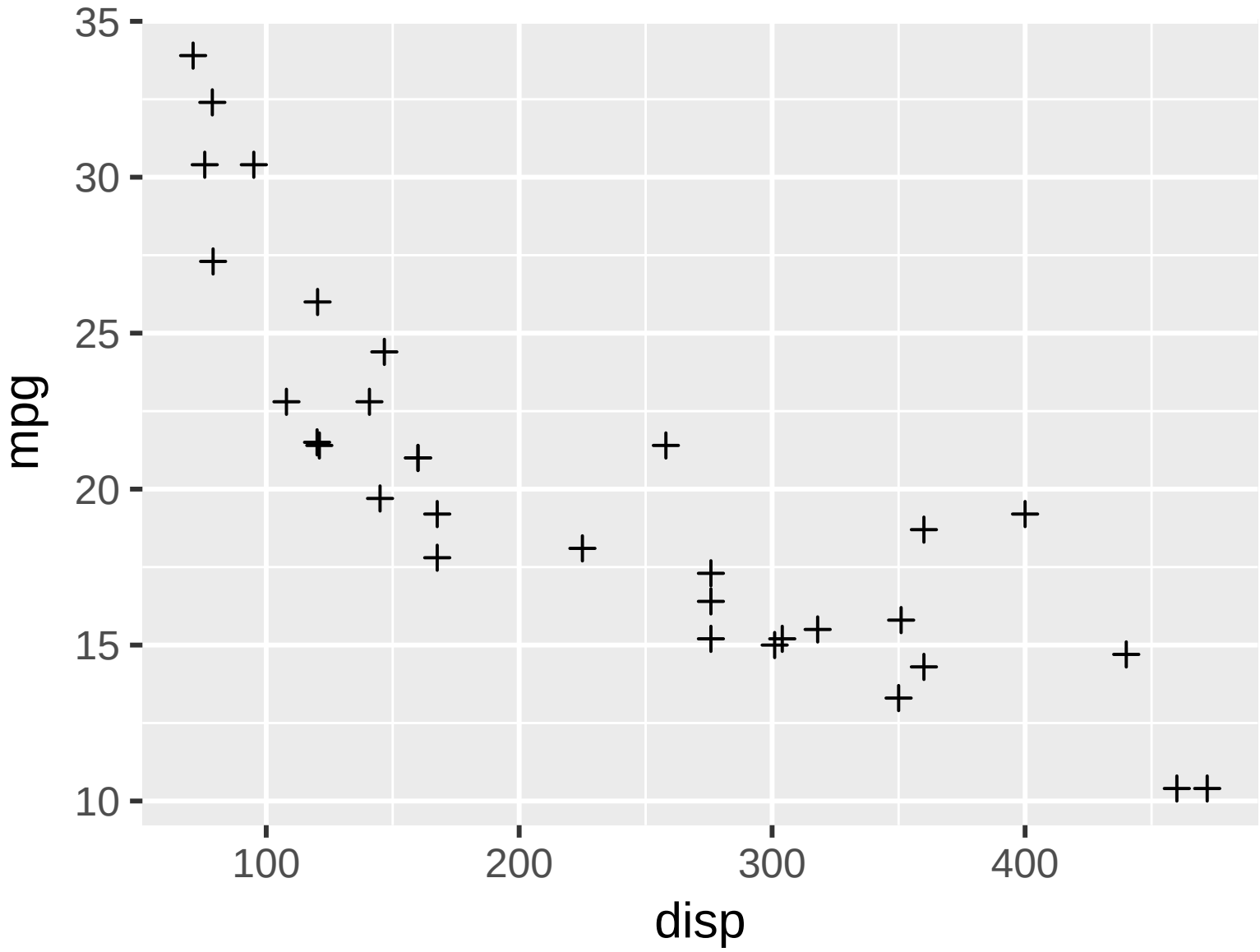
Color = blue

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), color = "blue")
```



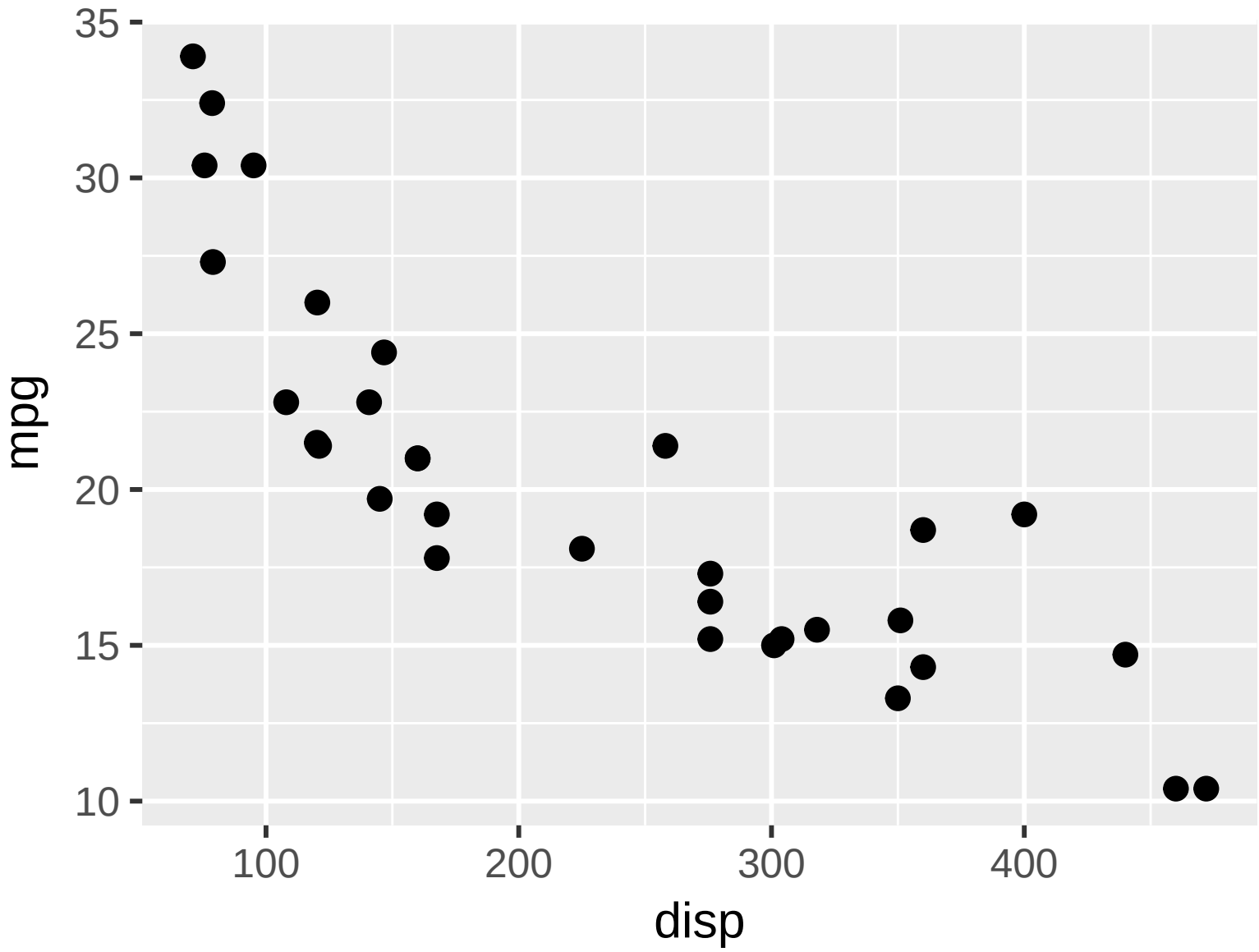
Shape = 3

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), shape = 3)
```



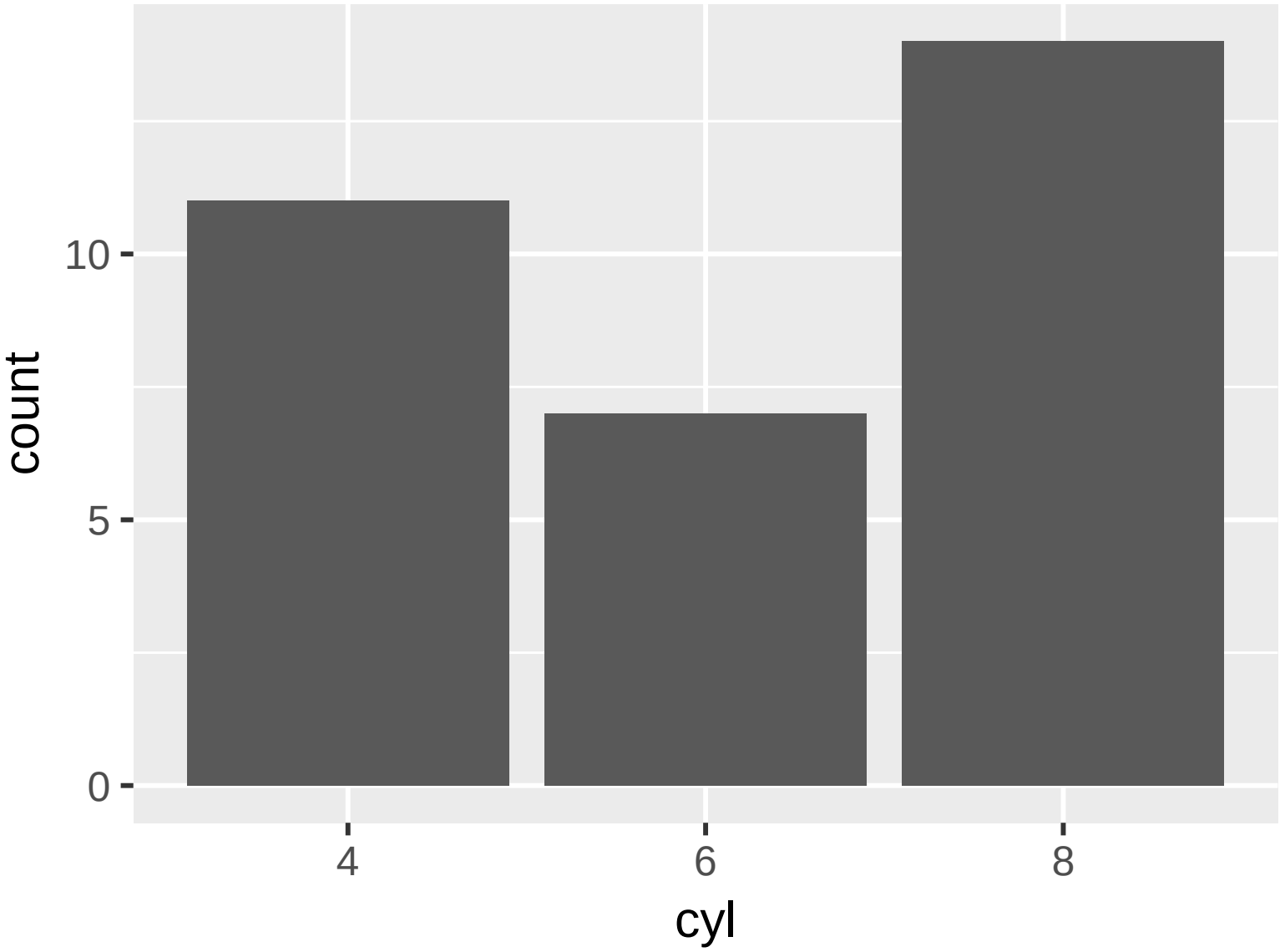
Size = 2

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), size = 2)
```



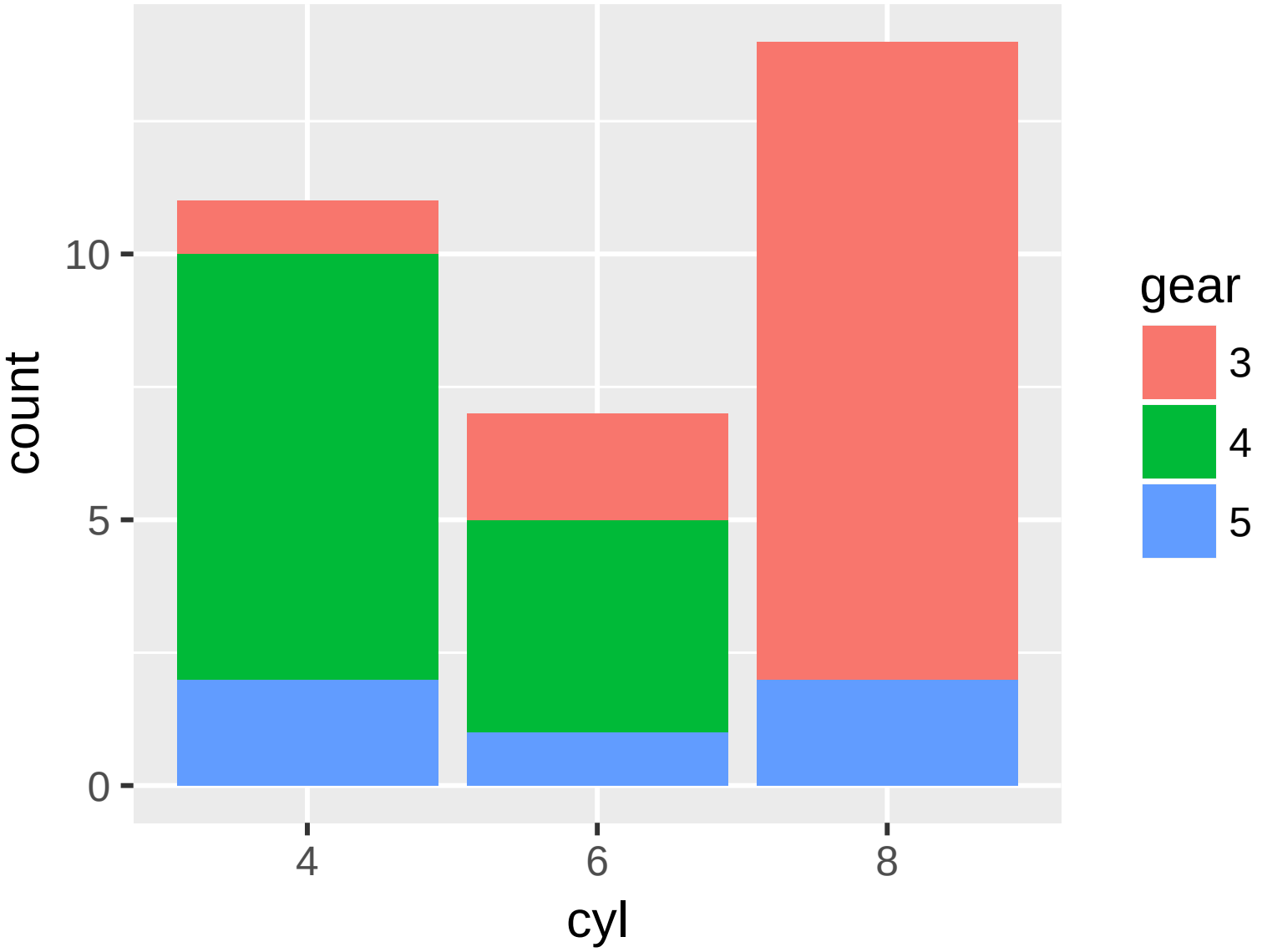
Bar chart

```
ggplot(data = mtcarz) +
  stat_count(mapping = aes(x = cyl))
```



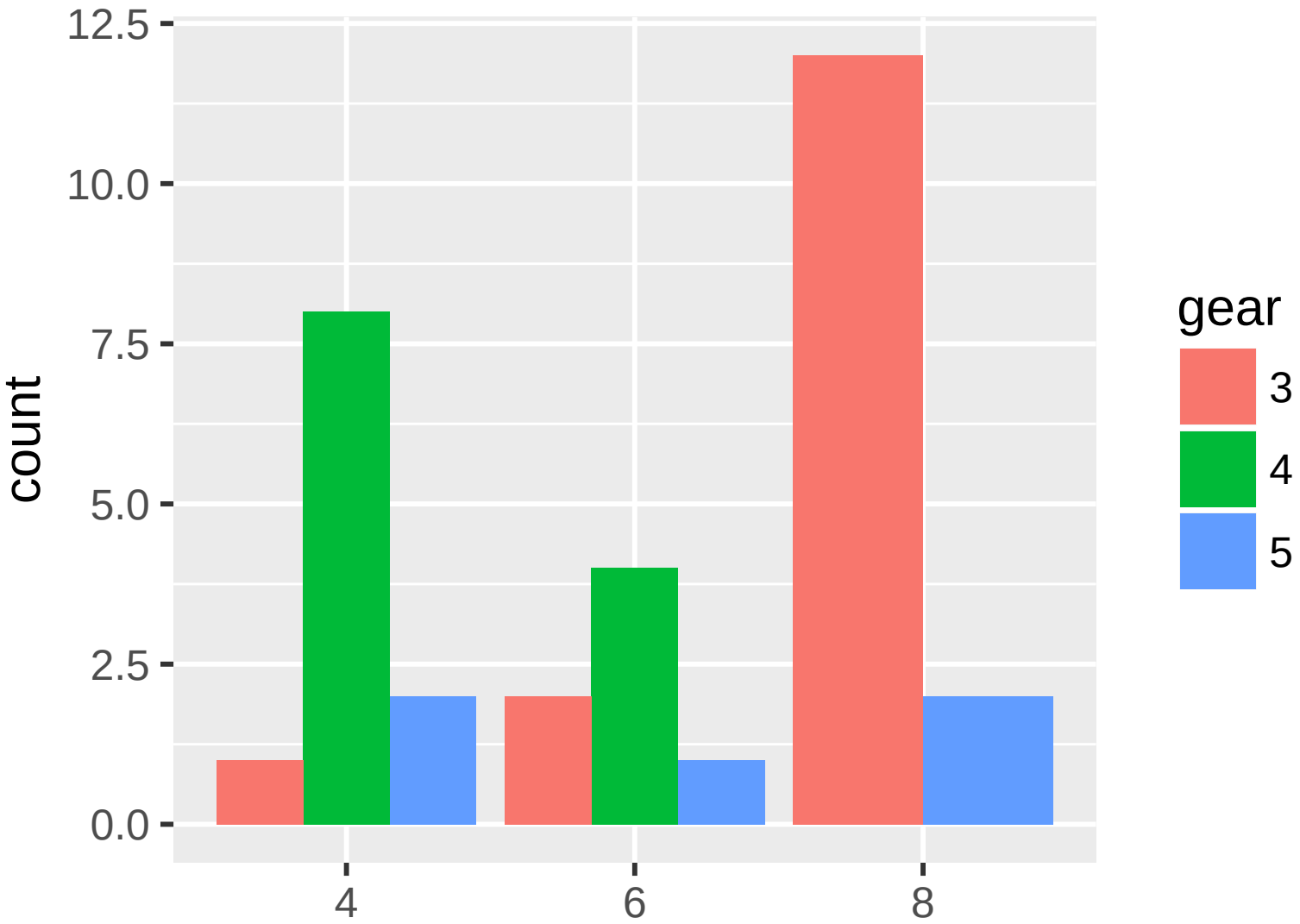
Position

```
ggplot(data = mtcarz) +
  geom_bar(mapping = aes(x = cyl, fill = gear))
```



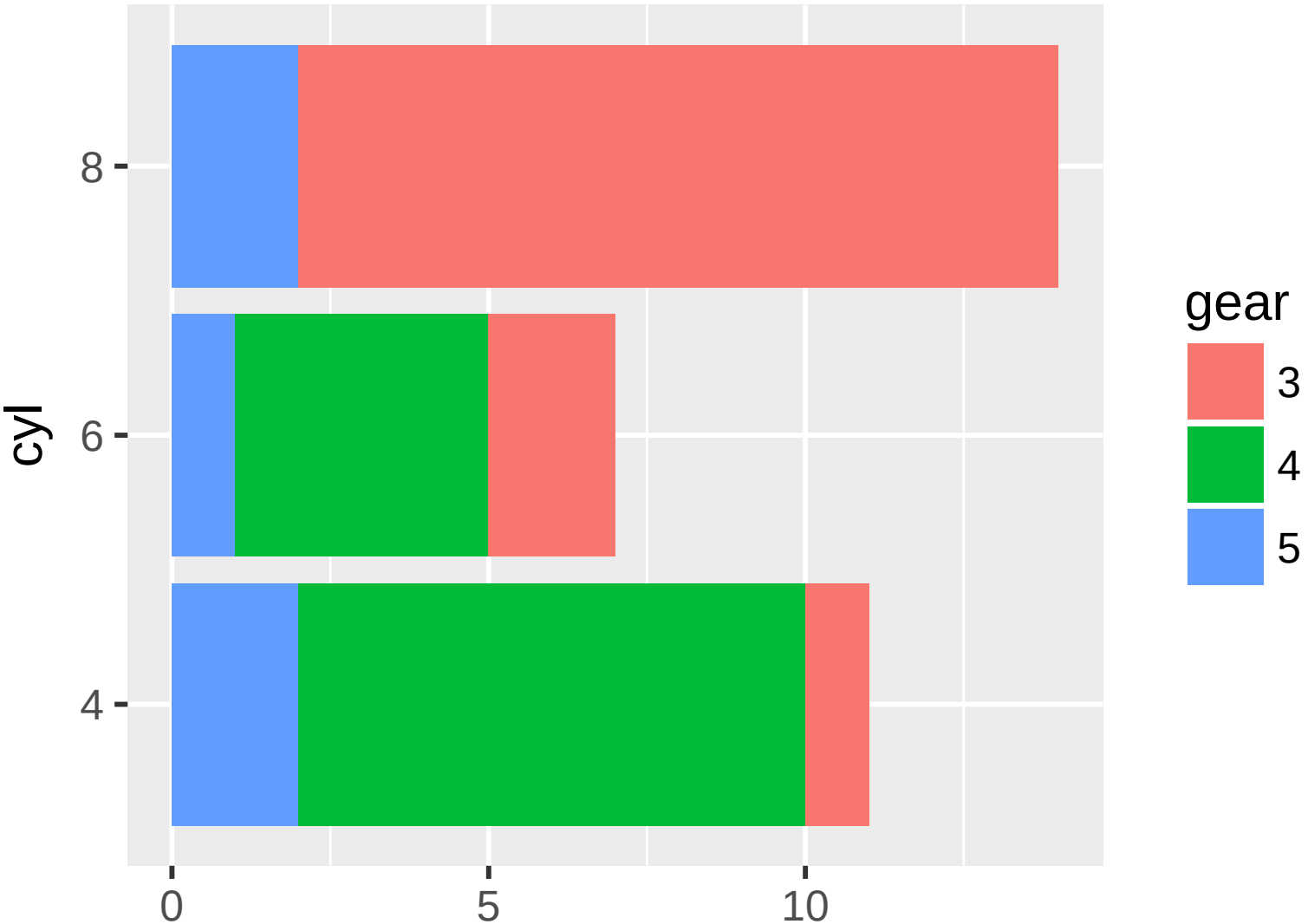
Position

```
ggplot(data = mtcarz) +
  geom_bar(mapping = aes(x = cyl, fill = gear),
    position = "dodge")
```



Flip Coordinates

```
ggplot(data = mtcarz) +
  geom_bar(mapping = aes(x = cyl, fill = gear)) +
  coord_flip()
```

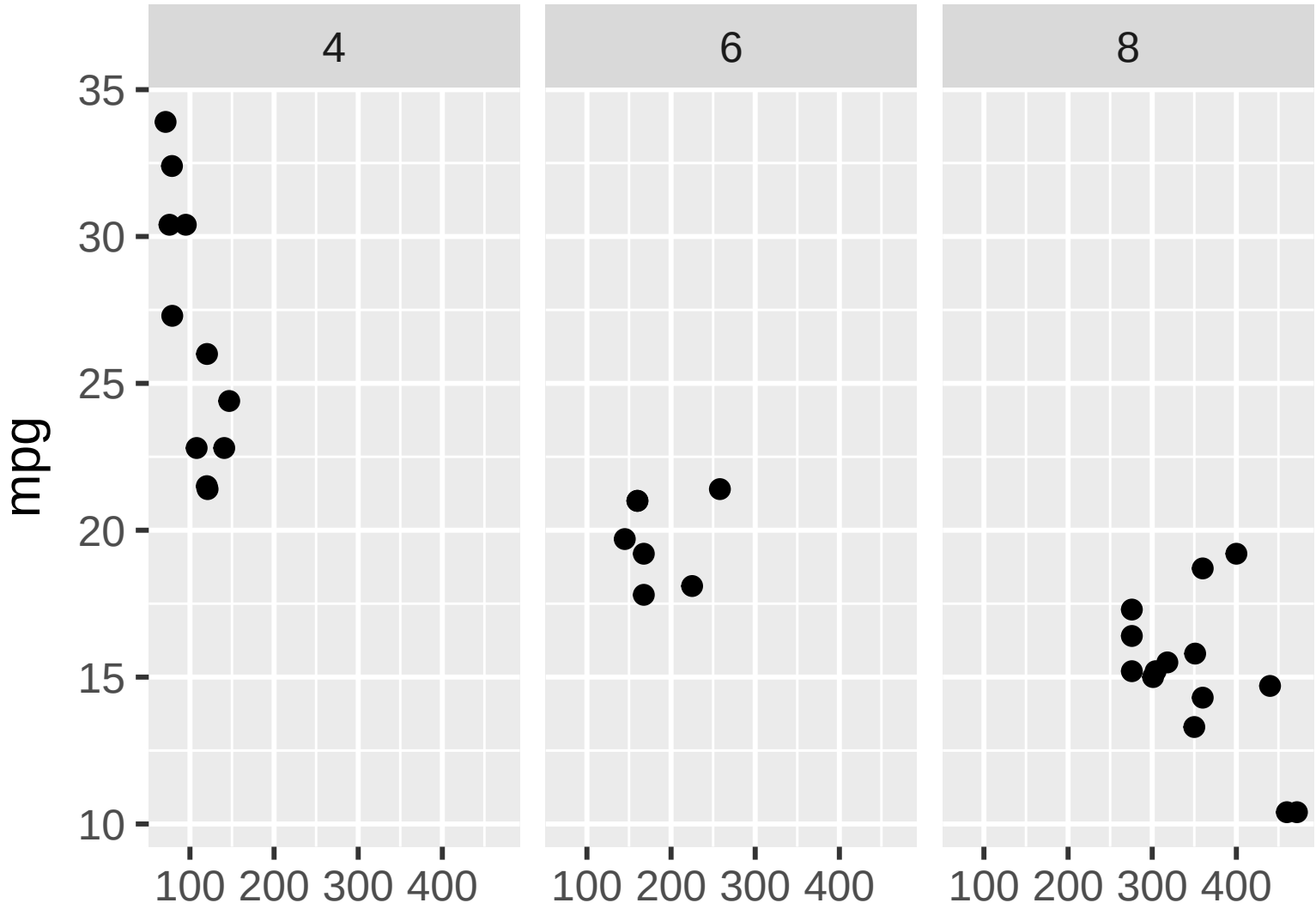


Facets

- split the plot into sub plots each of which
- displays a subset of the data
- examine the relationship between disp and mpg for cyl

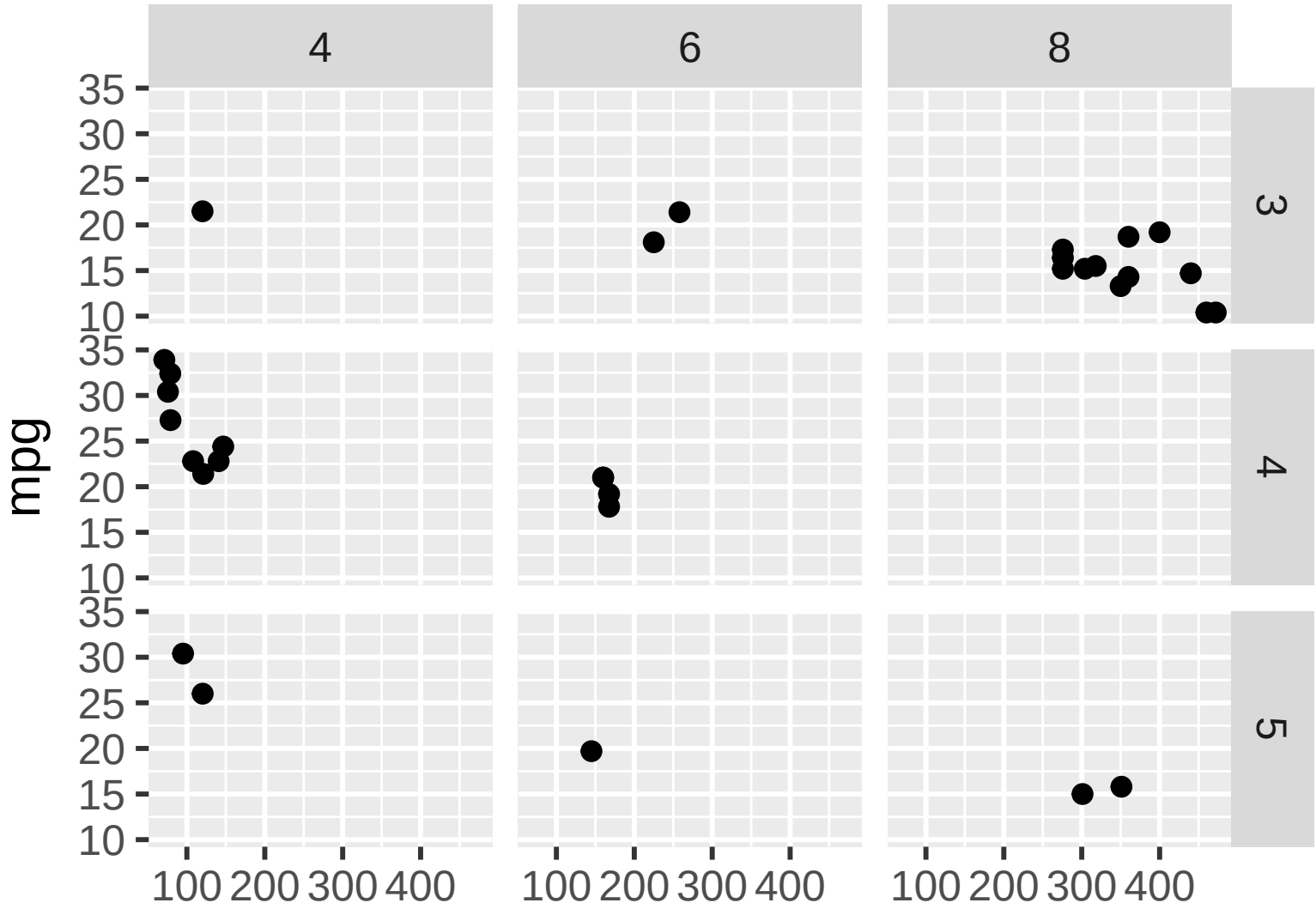
Facets

```
ggplot(data = mtcarz) +
  geom_point(mapping = aes(x = disp, y = mpg)) +
  facet_wrap(~ cyl)
```



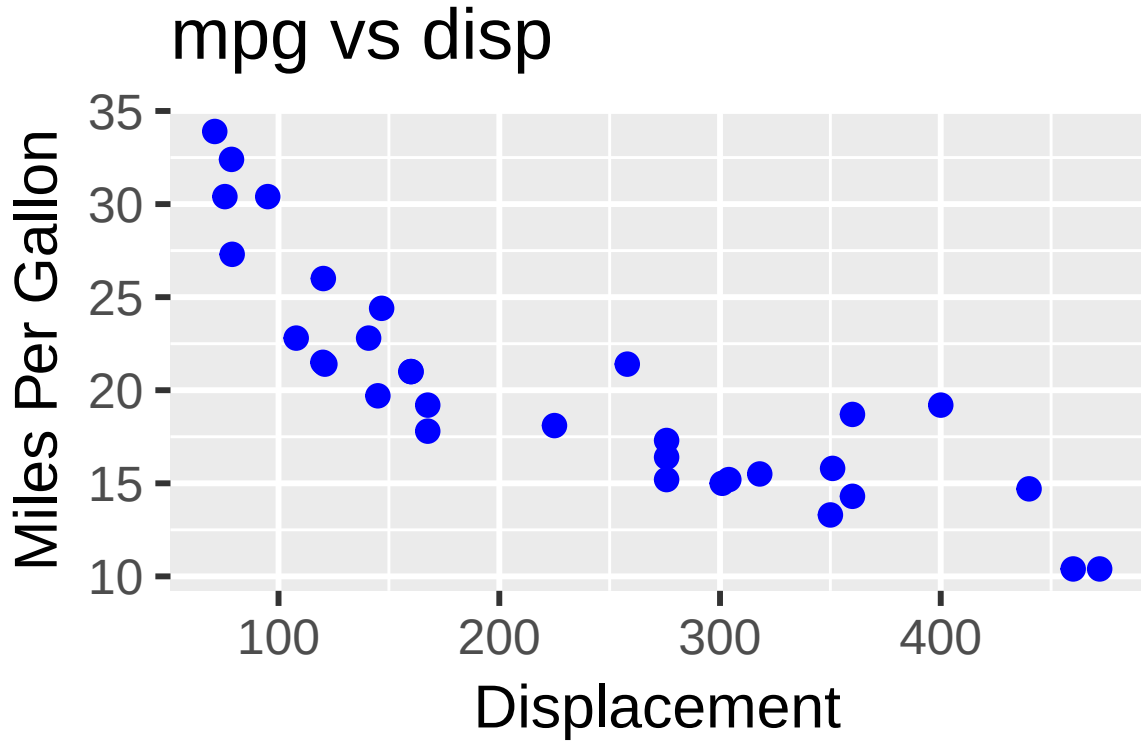
Facets

```
ggplot(data = mtcarz) +
  geom_point(mapping = aes(x = disp, y = mpg)) +
  facet_grid(gear ~ cyl)
```



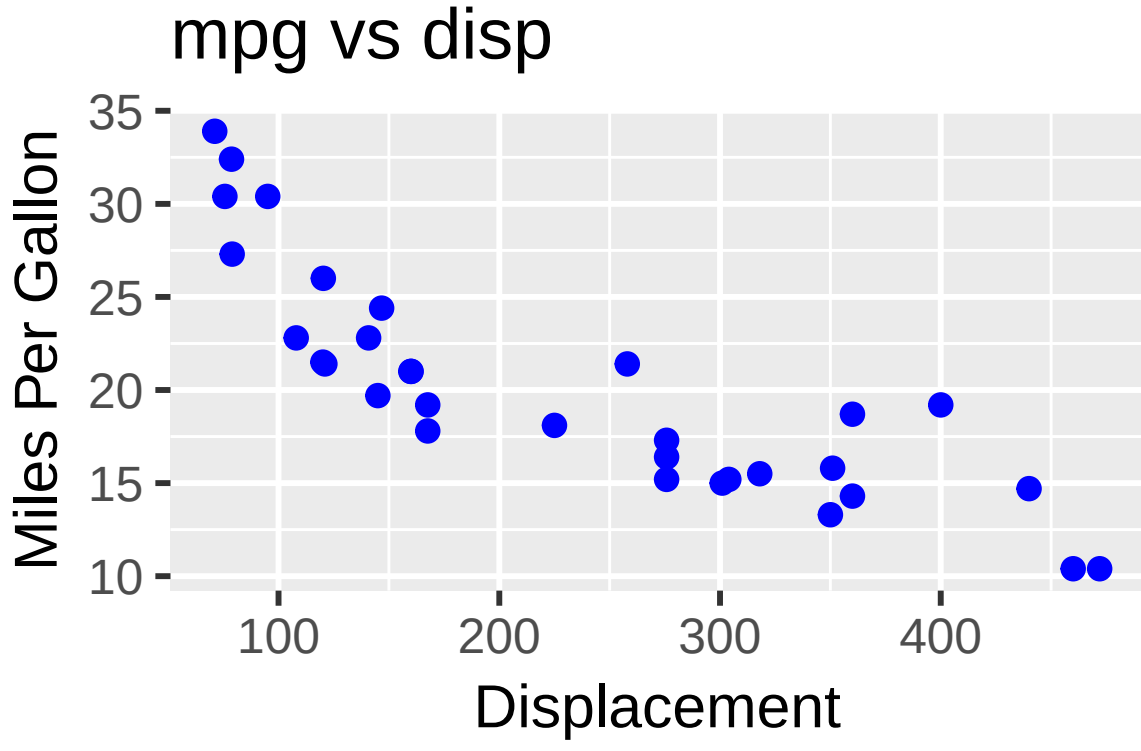
Title, Axis Labels & Limits

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), color = "blue") +
  labs(title = "mpg vs disp", x = "Displacement",
    y = "Miles Per Gallon")
```



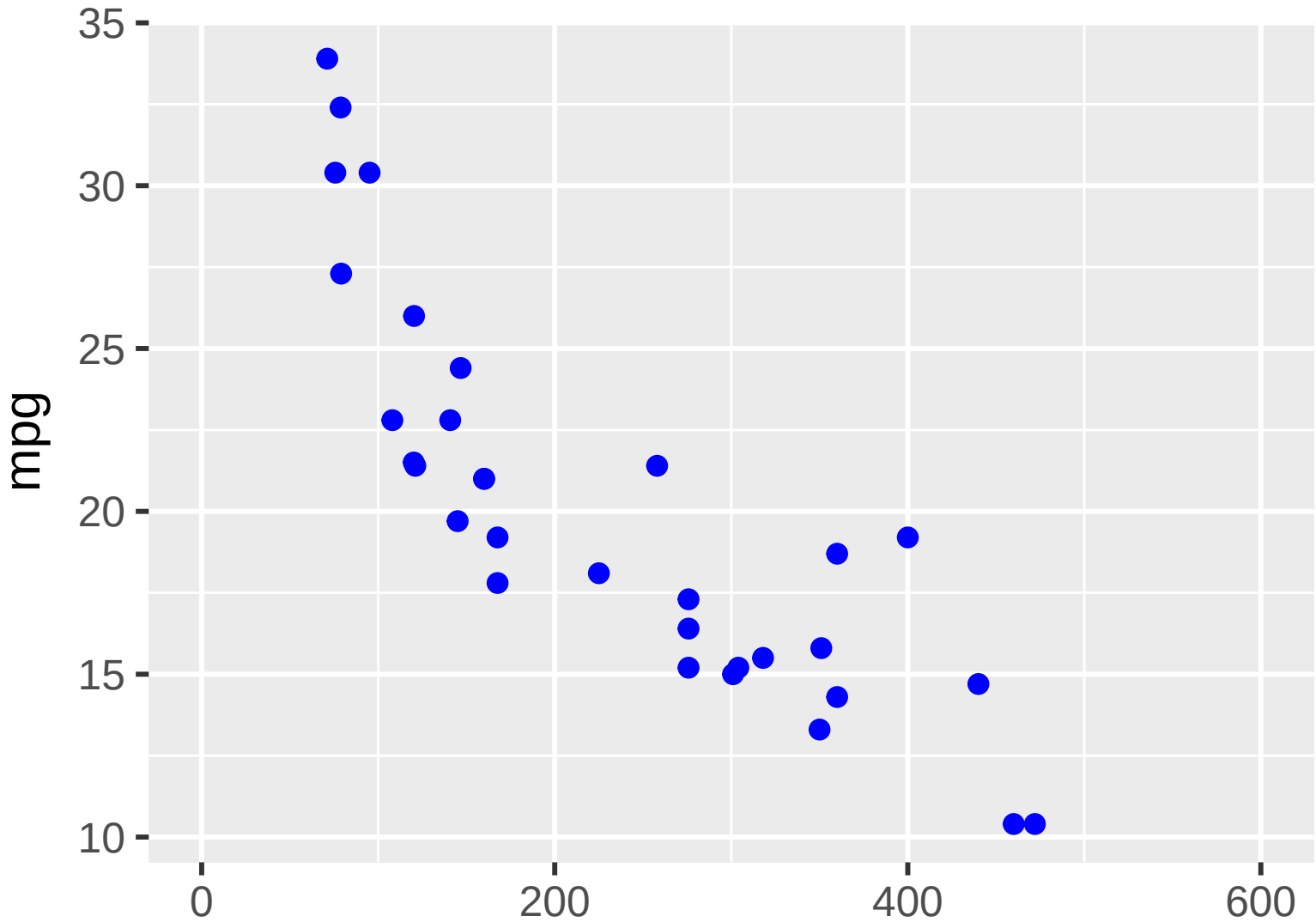
Title

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), color = "blue") +
  ggtitle("mpg vs disp") + xlab("Displacement") +
  ylab("Miles Per Gallon")
```



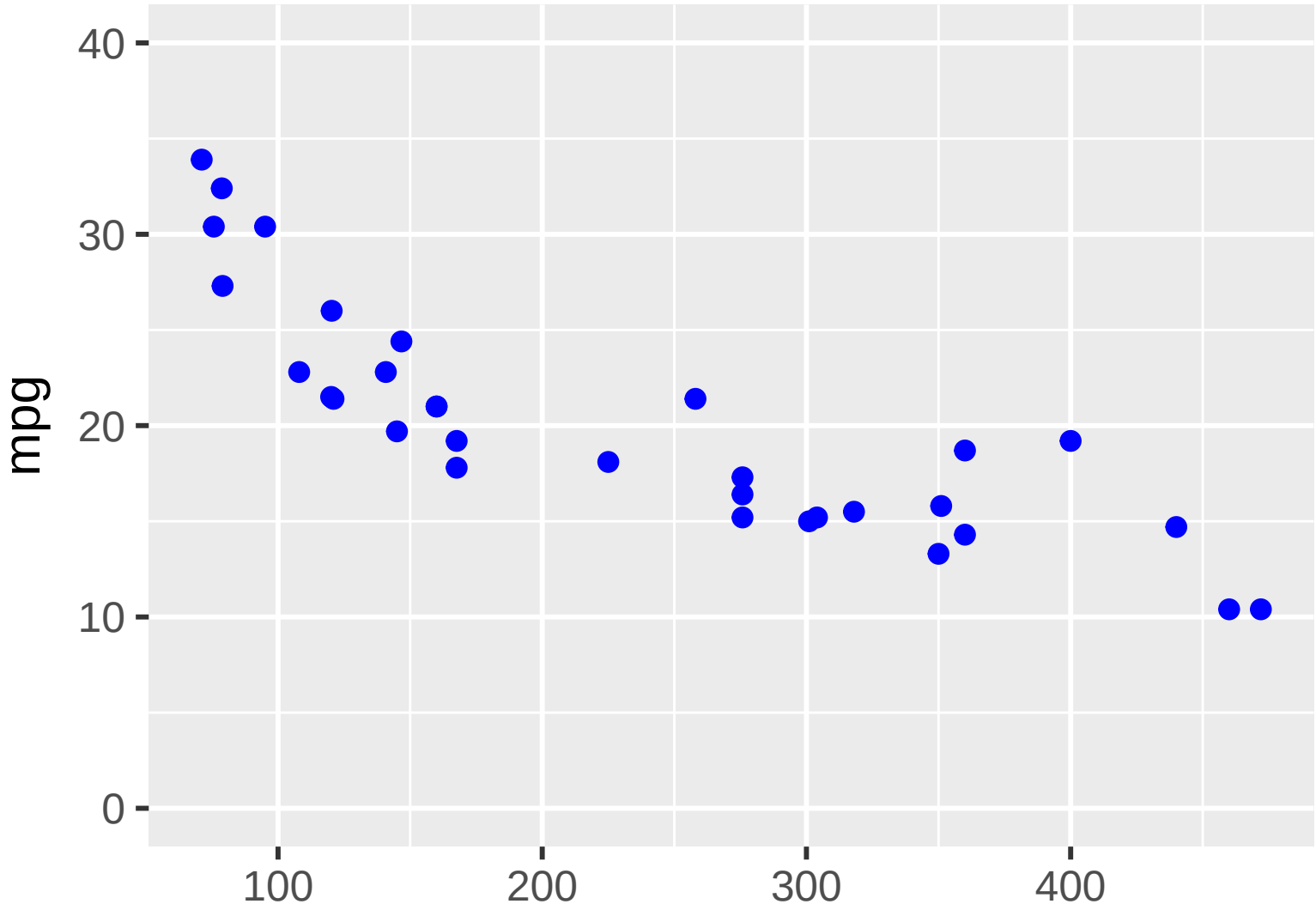
Axis Limits

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), color = "blue") +
  xlim(c(0, 600))
```



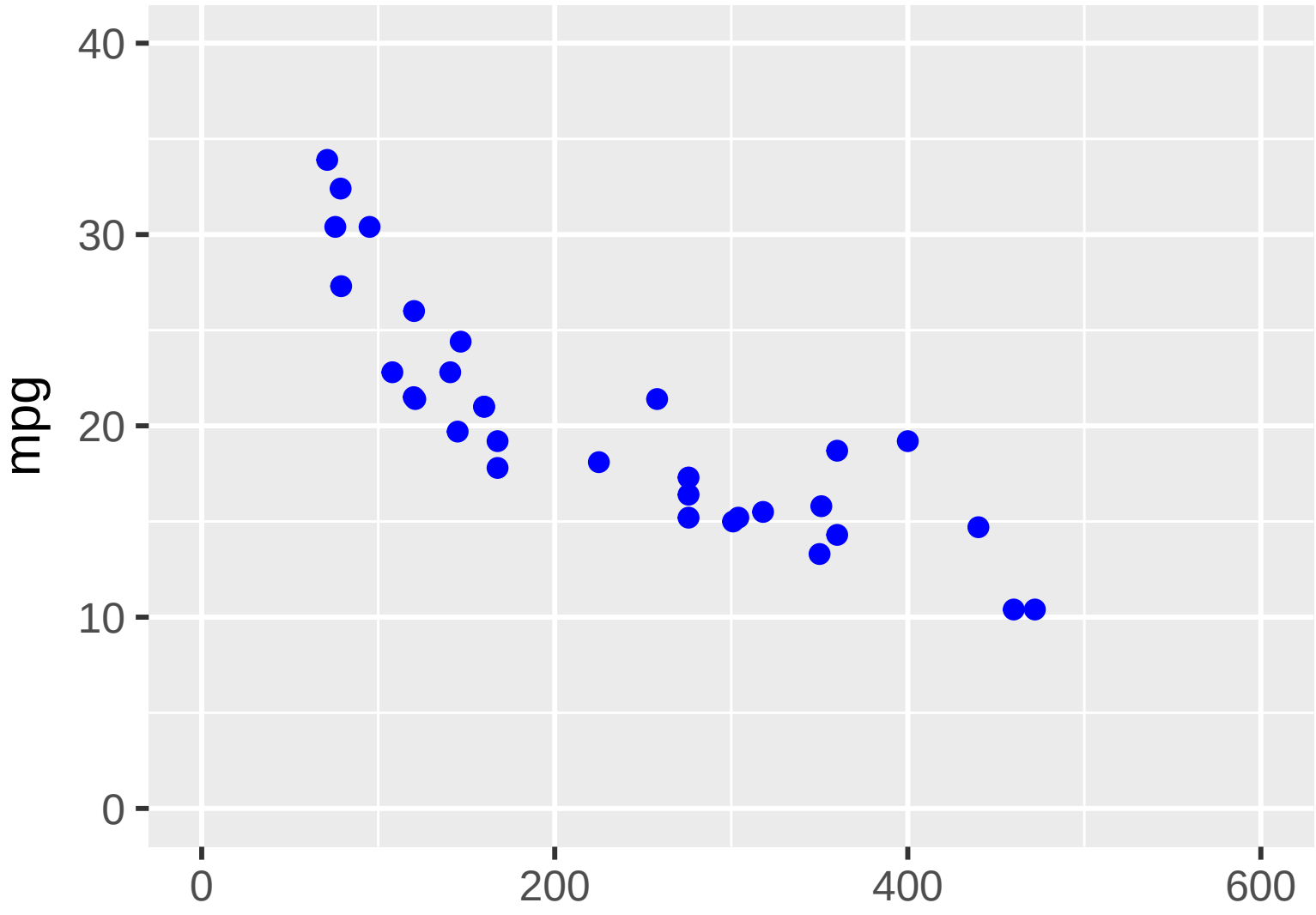
Axis Limits

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), color = "blue") +
  ylim(c(0, 40))
```



Axis Limits

```
ggplot(data = mtcarz) +
  geom_point(aes(x = disp, y = mpg), color = "blue") +
  expand_limits(x = c(0, 600), y = c(0, 40))
```

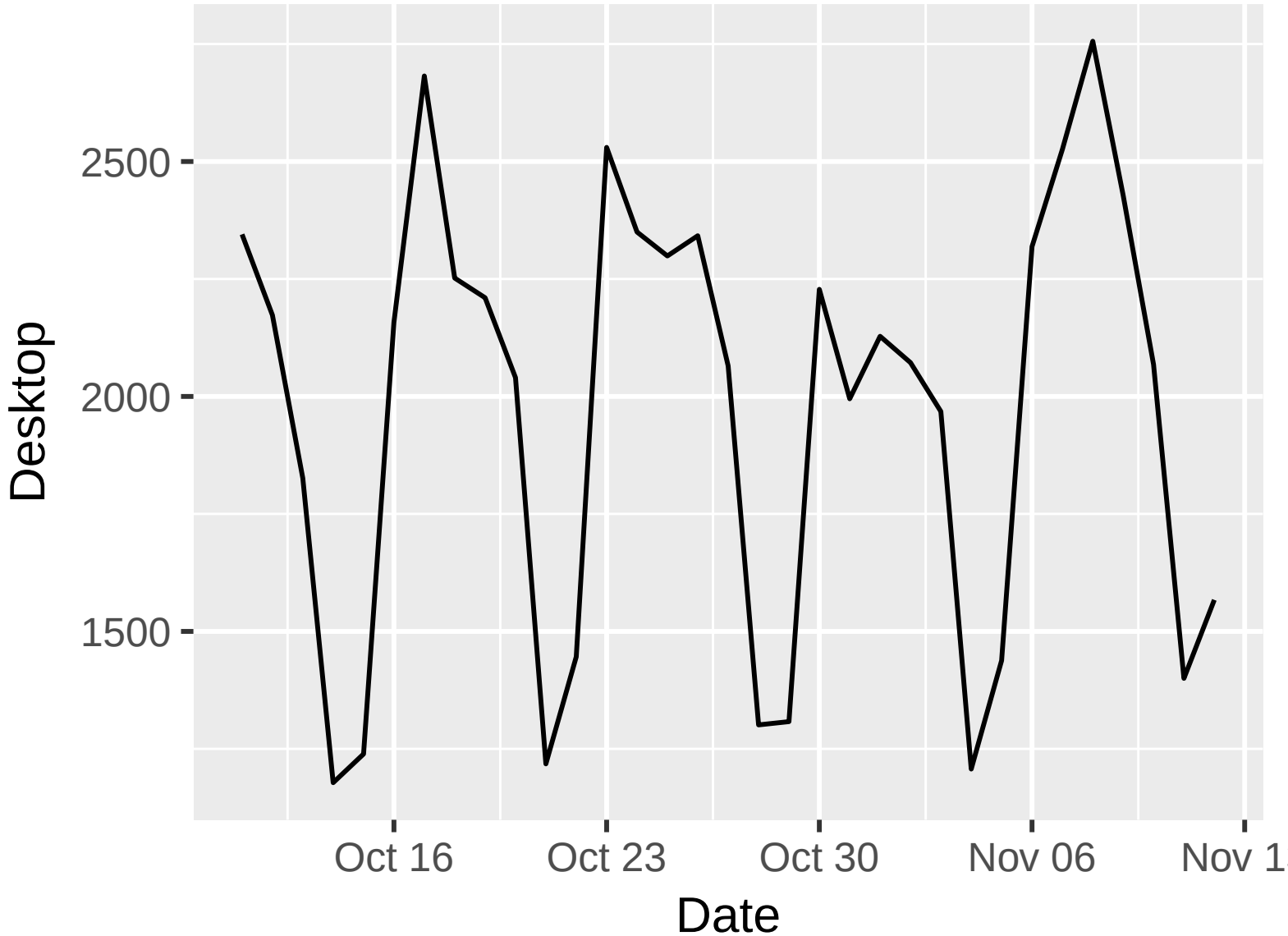


Line Chart: Data

```
device <- read_csv("data/users_device.csv",
           col_types = list(col_date(format = "%m/%d/%y"),
           col_integer(), col_integer(), col_integer()))
```

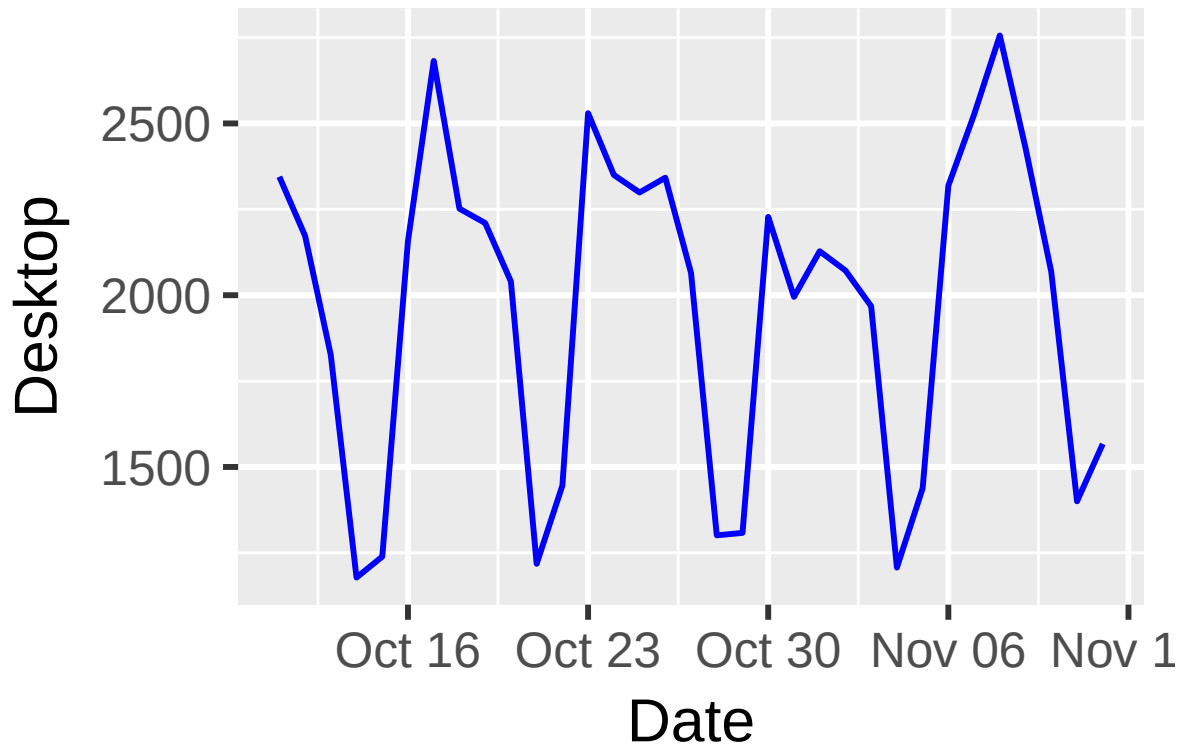
Line Chart

```
ggplot(data = device) +
  geom_line(mapping = aes(x = Date, y = Desktop))
```



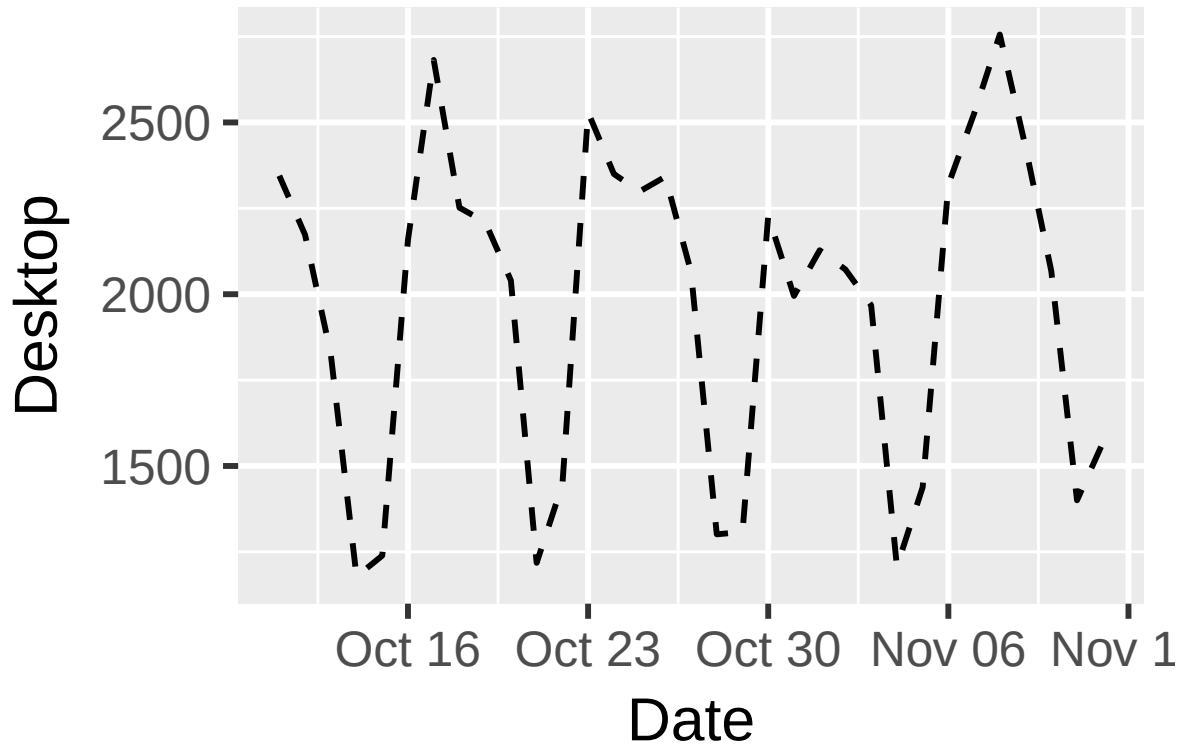
Line Color

```
ggplot(data = device) +
  geom_line(mapping = aes(x = Date, y = Desktop),
    color = "blue")
```



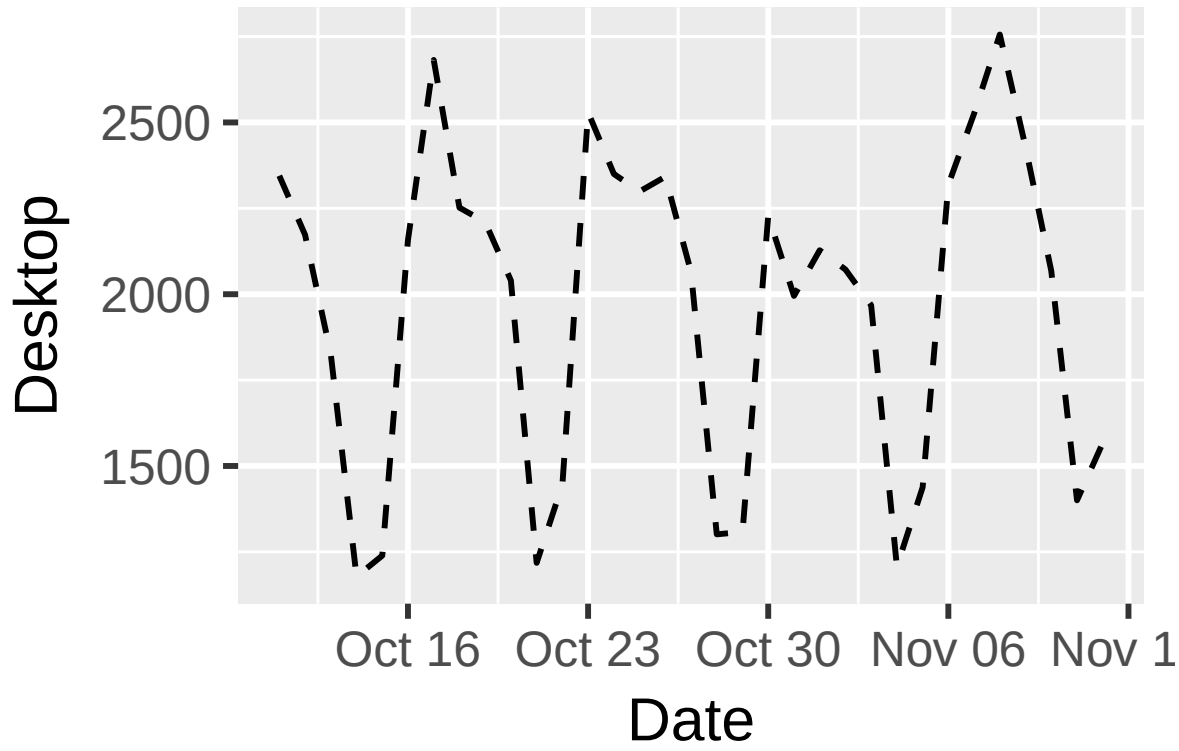
Line Type

```
ggplot(data = device) +
  geom_line(mapping = aes(x = Date, y = Desktop),
    linetype = 2)
```



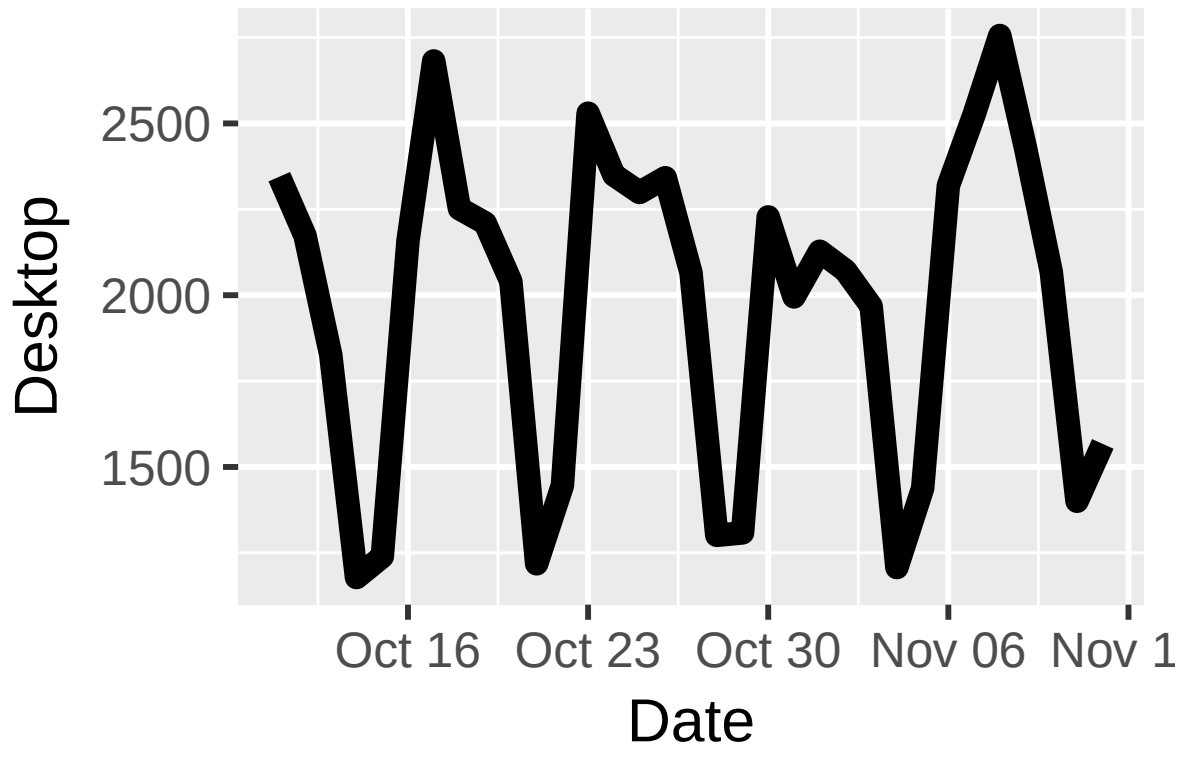
Line Type

```
ggplot(data = device) +
  geom_line(mapping = aes(x = Date, y = Desktop),
    linetype = "dashed")
```



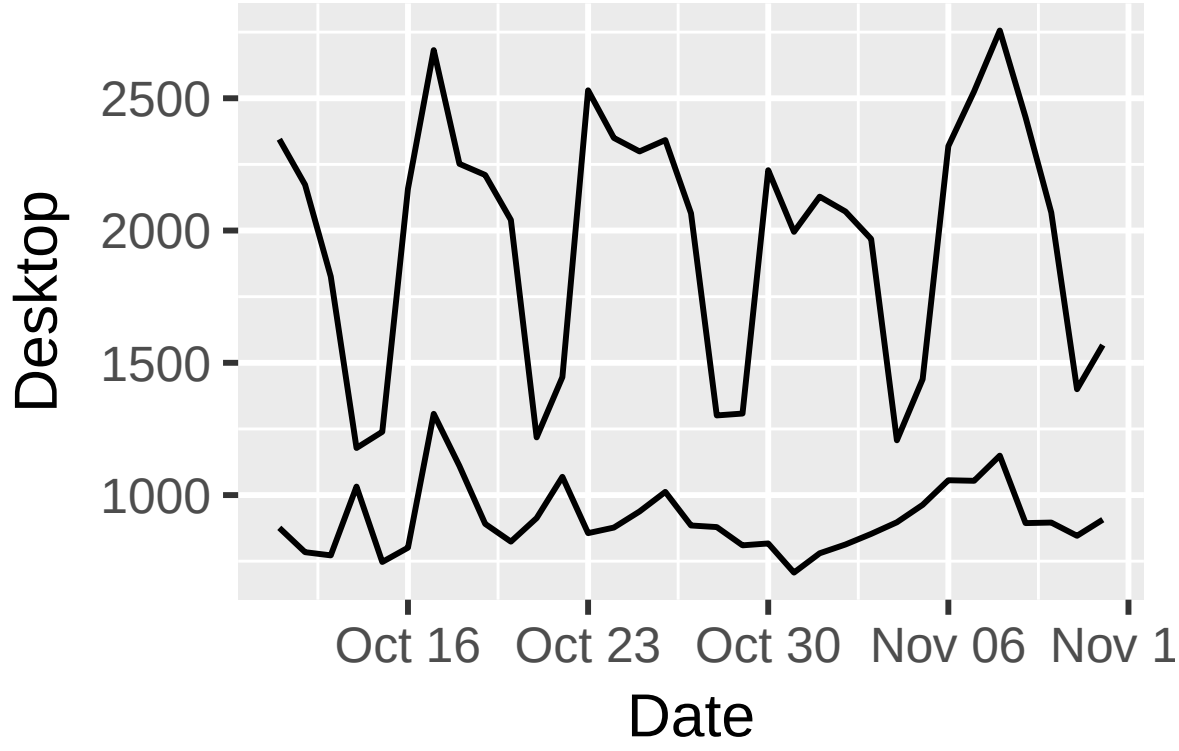
Line Width

```
ggplot(data = device) +
  geom_line(mapping = aes(x = Date, y = Desktop),
    size = 2)
```



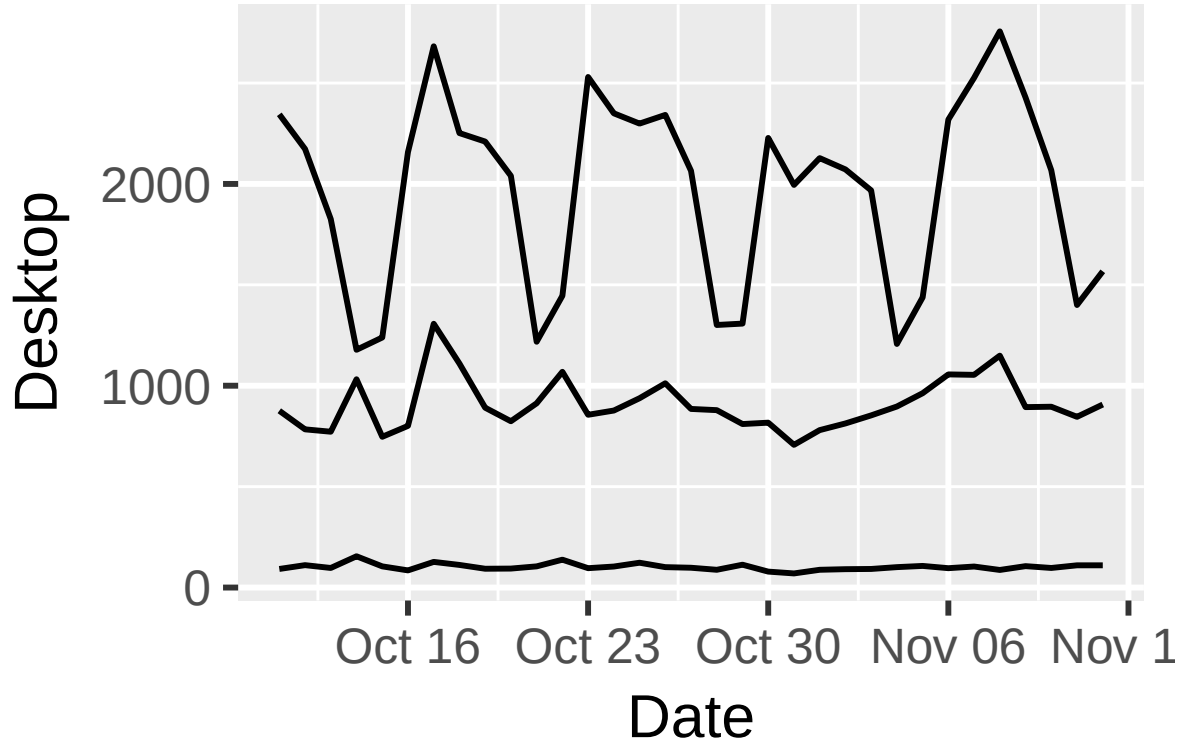
Multiple Lines

```
ggplot(data = device) +
  geom_line(mapping = aes(x = Date, y = Desktop)) +
  geom_line(mapping = aes(x = Date, y = Mobile))
```



Multiple Lines

```
ggplot(data = device) +
  geom_line(mapping = aes(x = Date, y = Desktop)) +
  geom_line(mapping = aes(x = Date, y = Mobile)) +
  geom_line(mapping = aes(x = Date, y = Tablet))
```

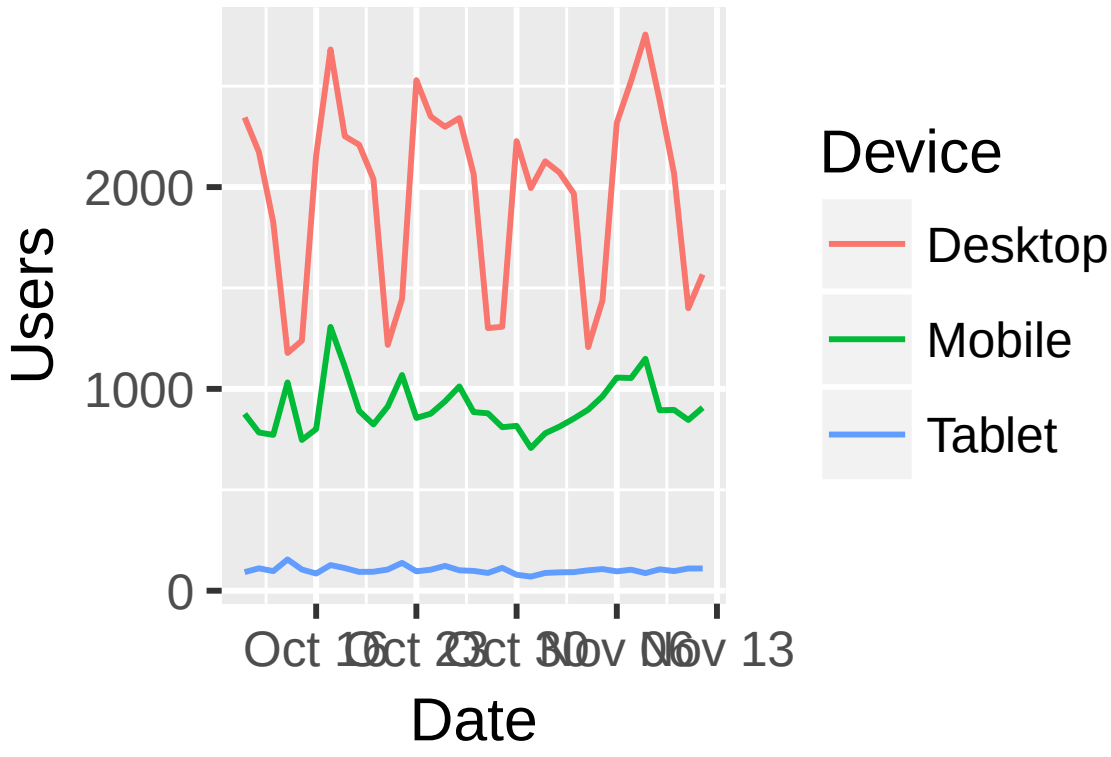


Line Chart: Data

```
tidy_device <- read_csv("data/tidy_users.csv",
                        col_types = list(col_date(format = "%m/%d/%y
                        col_character(), col_integer()))
```

Multiple Lines

```
ggplot(data = tidy_device) +
  geom_line(mapping = aes(x = Date, y = Users,
    group = Device, color = Device))
```



Thank You

For more information please visit our website
www.rsquaredacademy.com