Practice - R for Data Science

Ben Nickley

December 7, 2017

```
library(ggplot2)
```

```
library(knitr)
opts_chunk$set(tidy.opts=list(width.cutoff=60),tidy=TRUE)
```

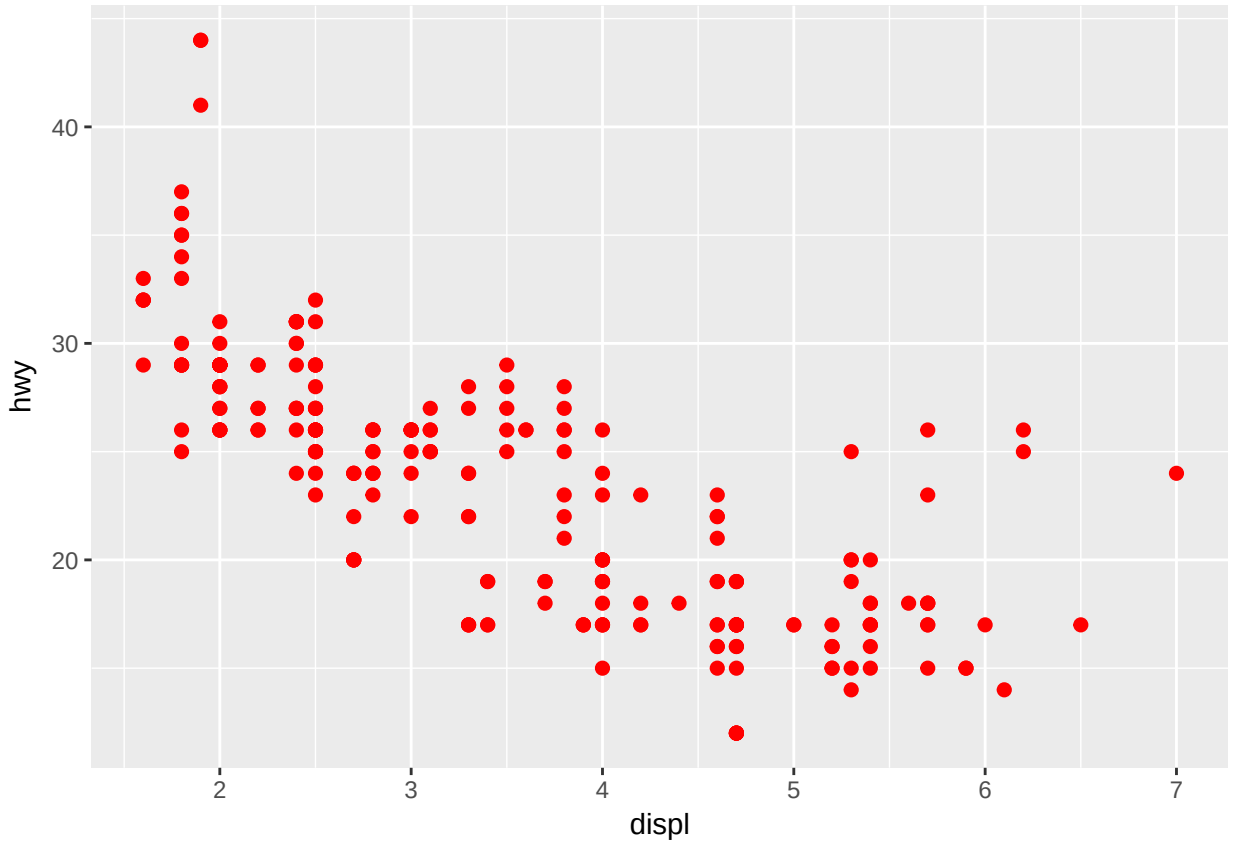
Ch 3 - Data Visualization

R for Data Science - Ch 3

ggplot creates the coordinate system w/ the first arguemnt, which tells ggplot what data you are using

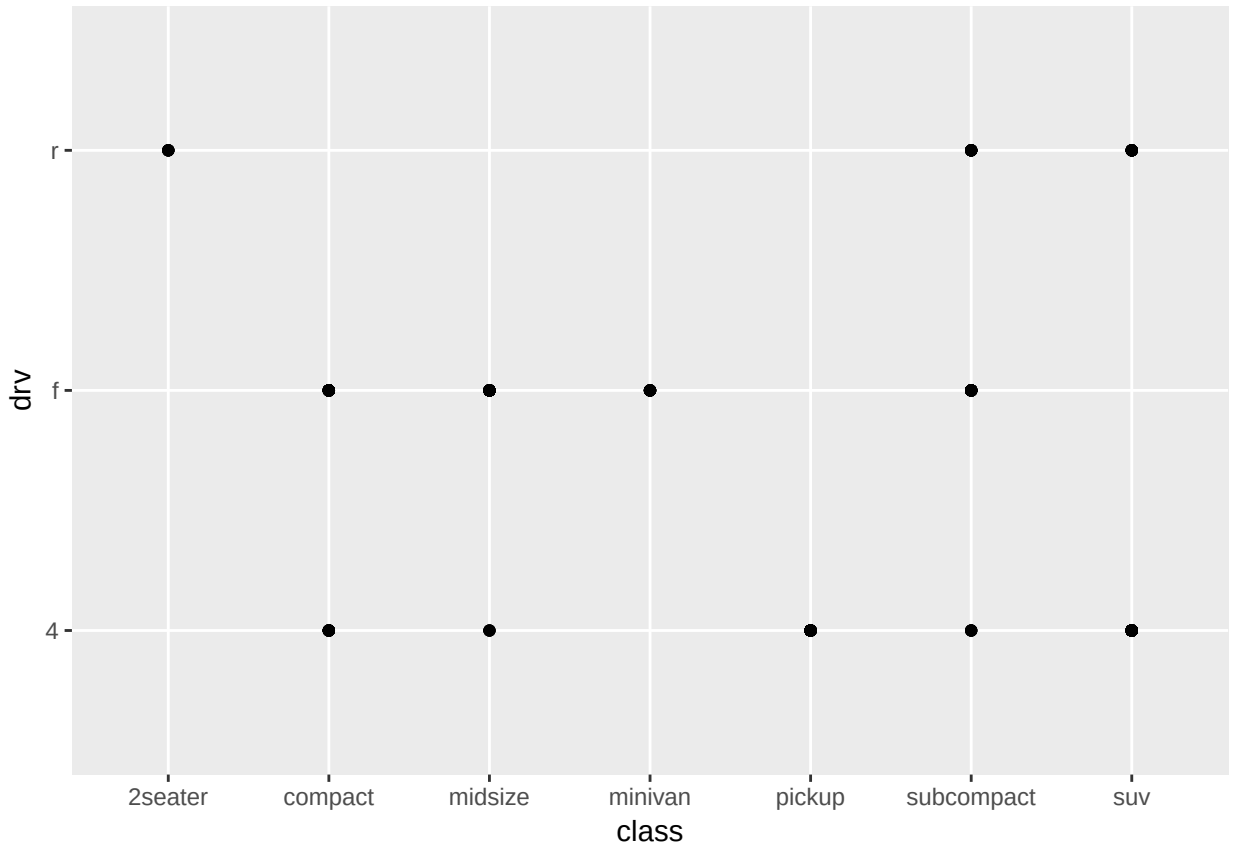
- add layers with subsequent arguemnts (i.e. geom_point)

```
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy),
    color = "red", size = 2)
```

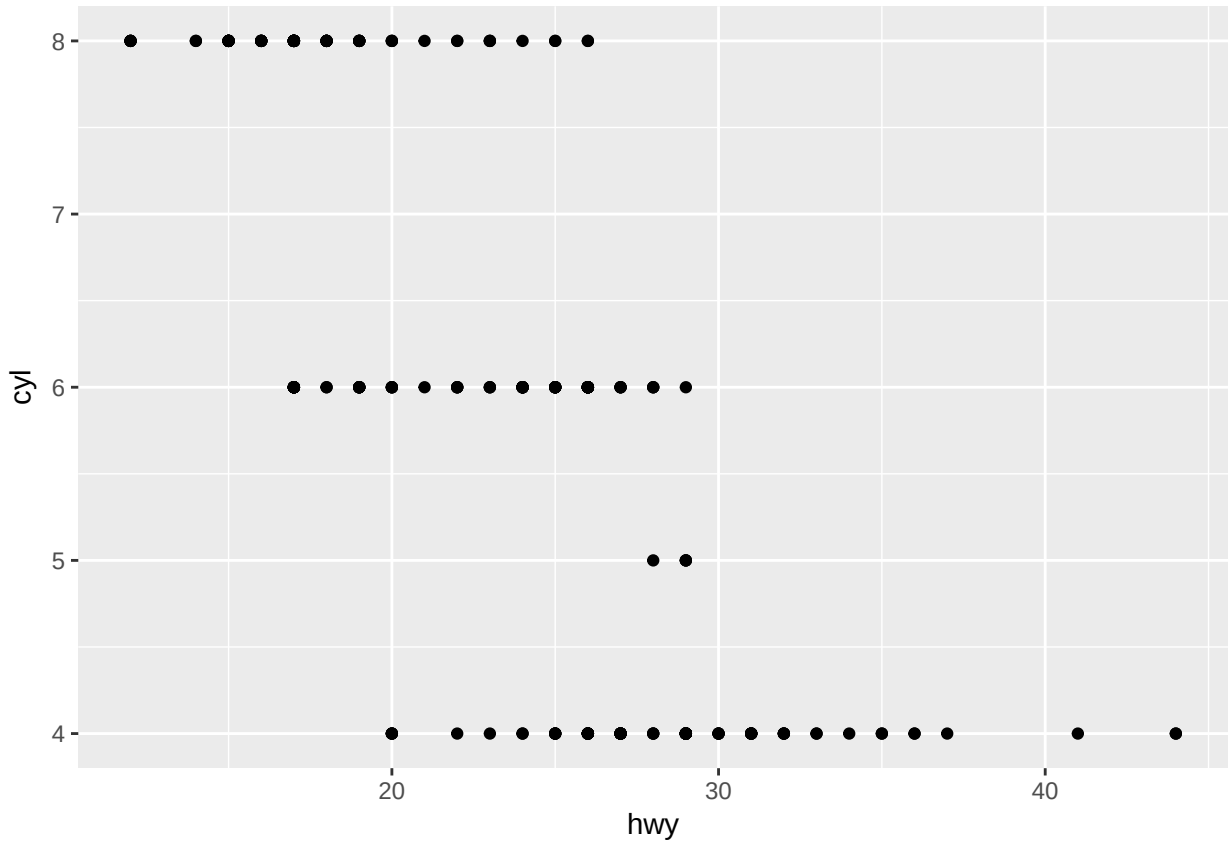


This is the ggplot skeleton: ggplot(data =) + (mapping = aes())

```
ggplot(data = mpg) + geom_point(mapping = aes(x = class, y = drv))
```

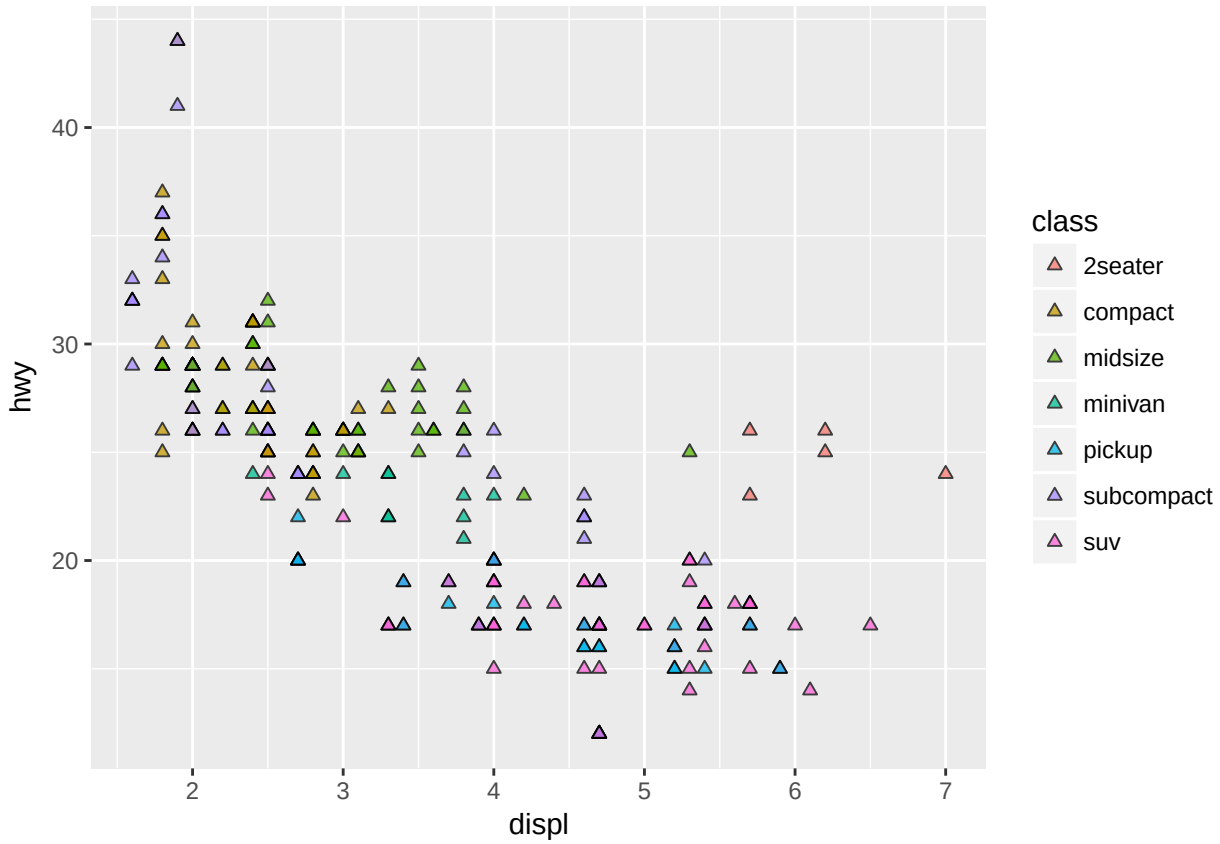



```
ggplot(data = mpg) + geom_point(mapping = aes(x = hwy, y = cyl))
```

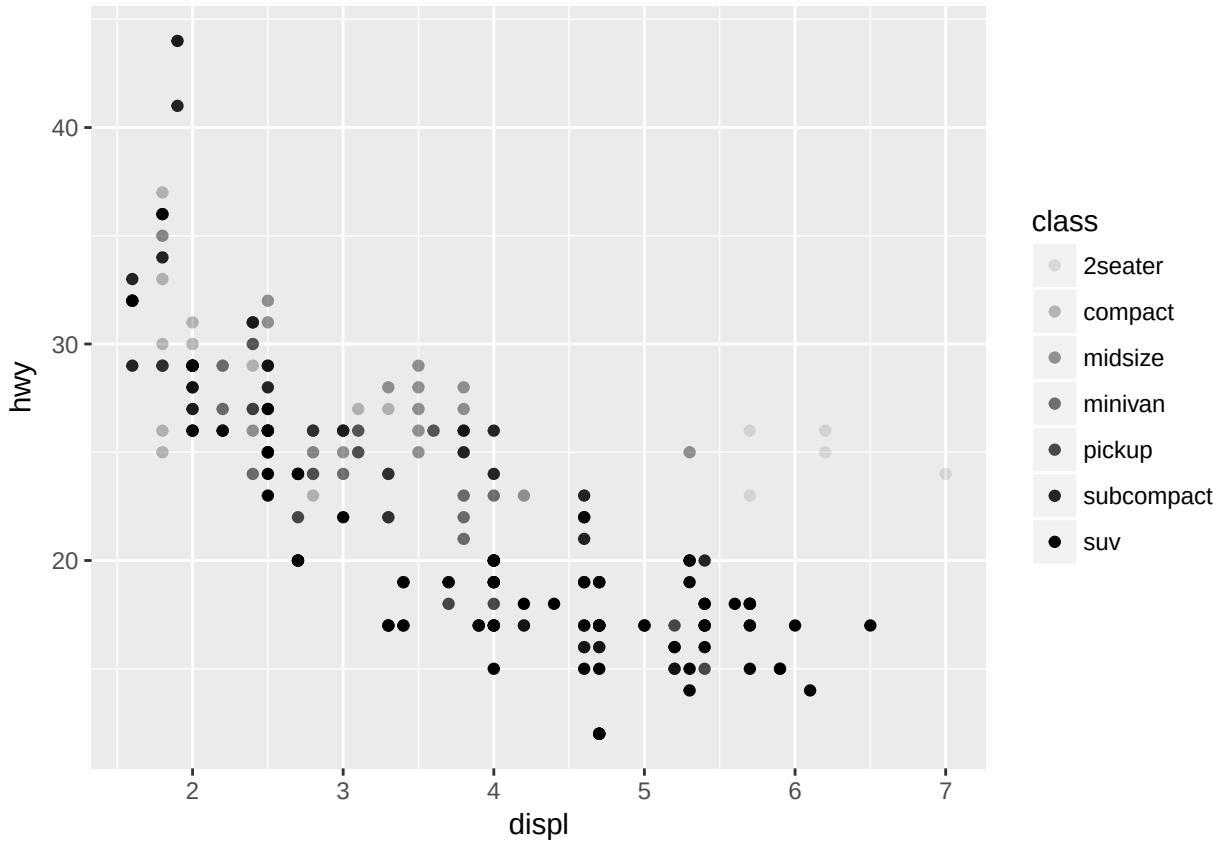


Map 'class' of car to different aesthetics (aes)

```
# Left
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy,
    fill = class), shape = 24, color = "black", alpha = 0.75)
```



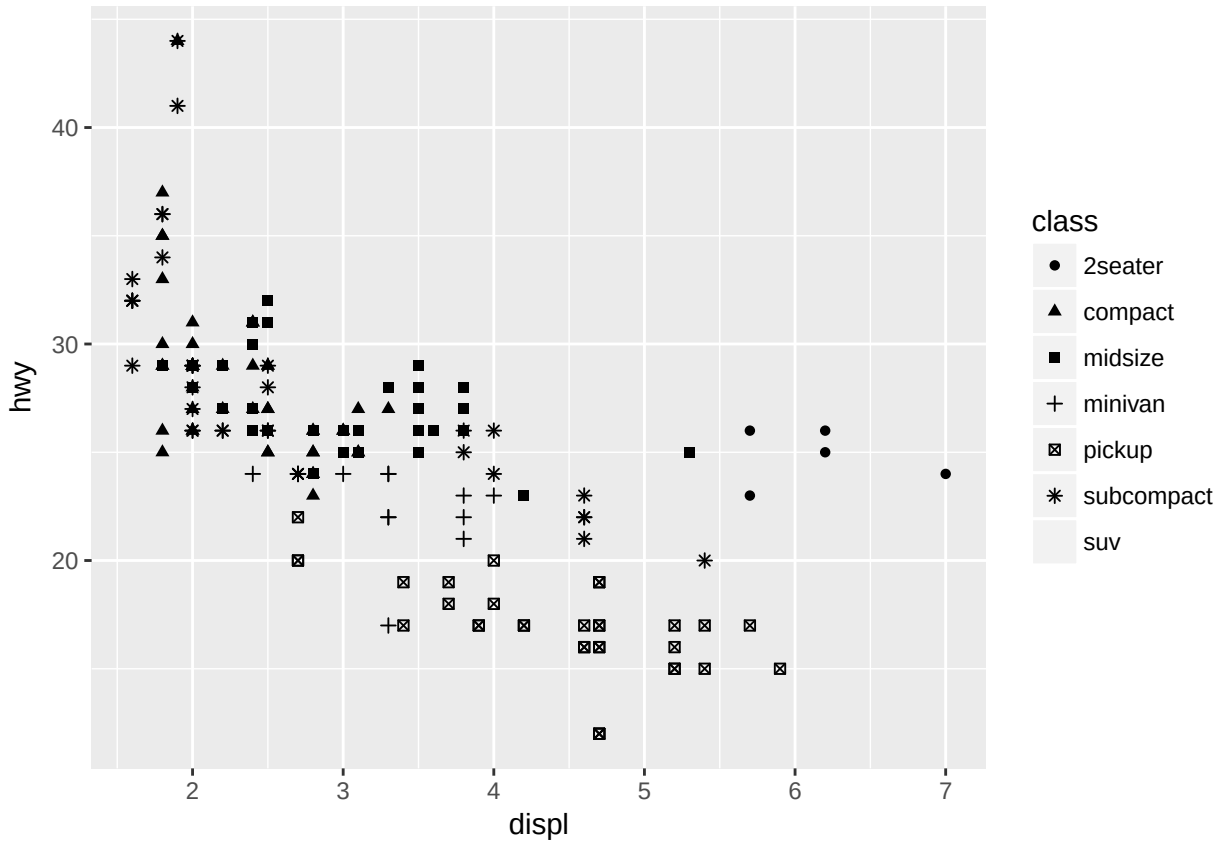

```
# Middle
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy,
    alpha = class))
```




```
# Right
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy,
    shape = class))
```

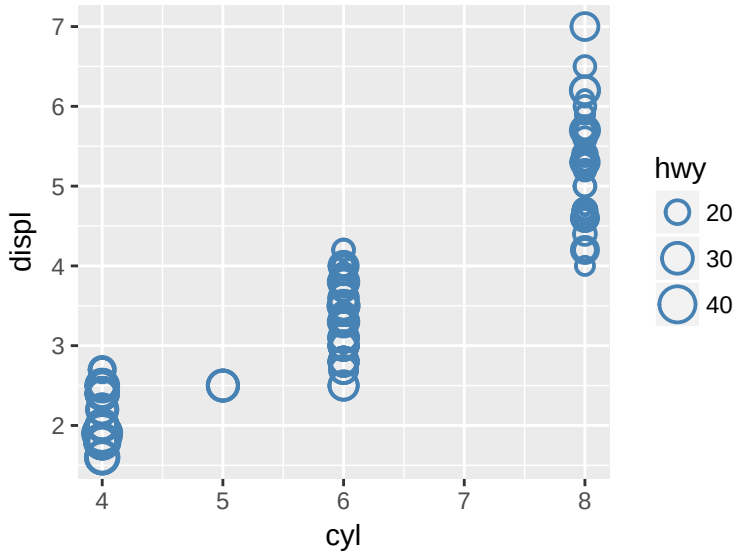
```
## Warning: The shape palette can deal with a maximum of 6 discrete values
## because more than 6 becomes difficult to discriminate; you have 7.
## Consider specifying shapes manually if you must have them.
```

```
## Warning: Removed 62 rows containing missing values (geom_point).
```

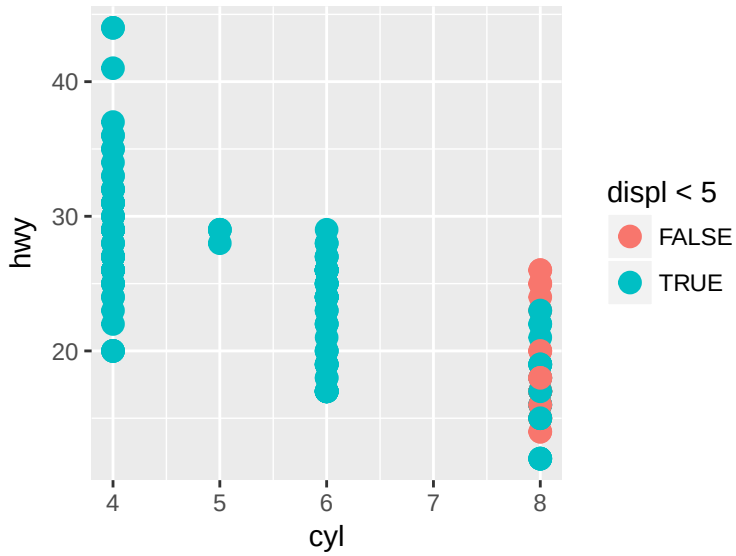


3. Map a continuous variable to color, size, and shape. How do these aesthetics behave differently for categorical vs. continuous variables?

```
ggplot(data = mpg) + geom_point(mapping = aes(x = cyl, y = displ,
    size = hwy), stroke = 1, shape = 1, color = "steelblue")
```




```
# What happens if you map an aesthetic to something other
# than a variable name, like aes(colour = displ < 5)?
ggplot(data = mpg) + geom_point(mapping = aes(x = cyl, y = hwy,
    colour = displ < 5), stroke = 2)
```




```
ggplot(data = mpg) + geom_point(mapping = aes(x = cyl, y = hwy,
    colour = fl), stroke = 2, shape = 20)
```