ggplot2

The grammar of graphics

Comprised of **building blocks** of plots that we can combine to create just about any plot we would like

Building blocks

- data
- geometric object (the marks we actually draw)
- aesthetic mappings (how we draw the marks)
- statistical transformations (how we transform data before plotting)
- scales (ranges of values, colors, shapes, sizes, etc.)
- faceting (small multiples)

Geometric objects

- In `ggplot2`, the type of marks we draw are set by `geoms`
- Examples:
  - `geom_point`
  - `geom_line`
  - `geom_bar`
  - `geom_boxplot`
- Data vis cheatsheet contains a more complete list

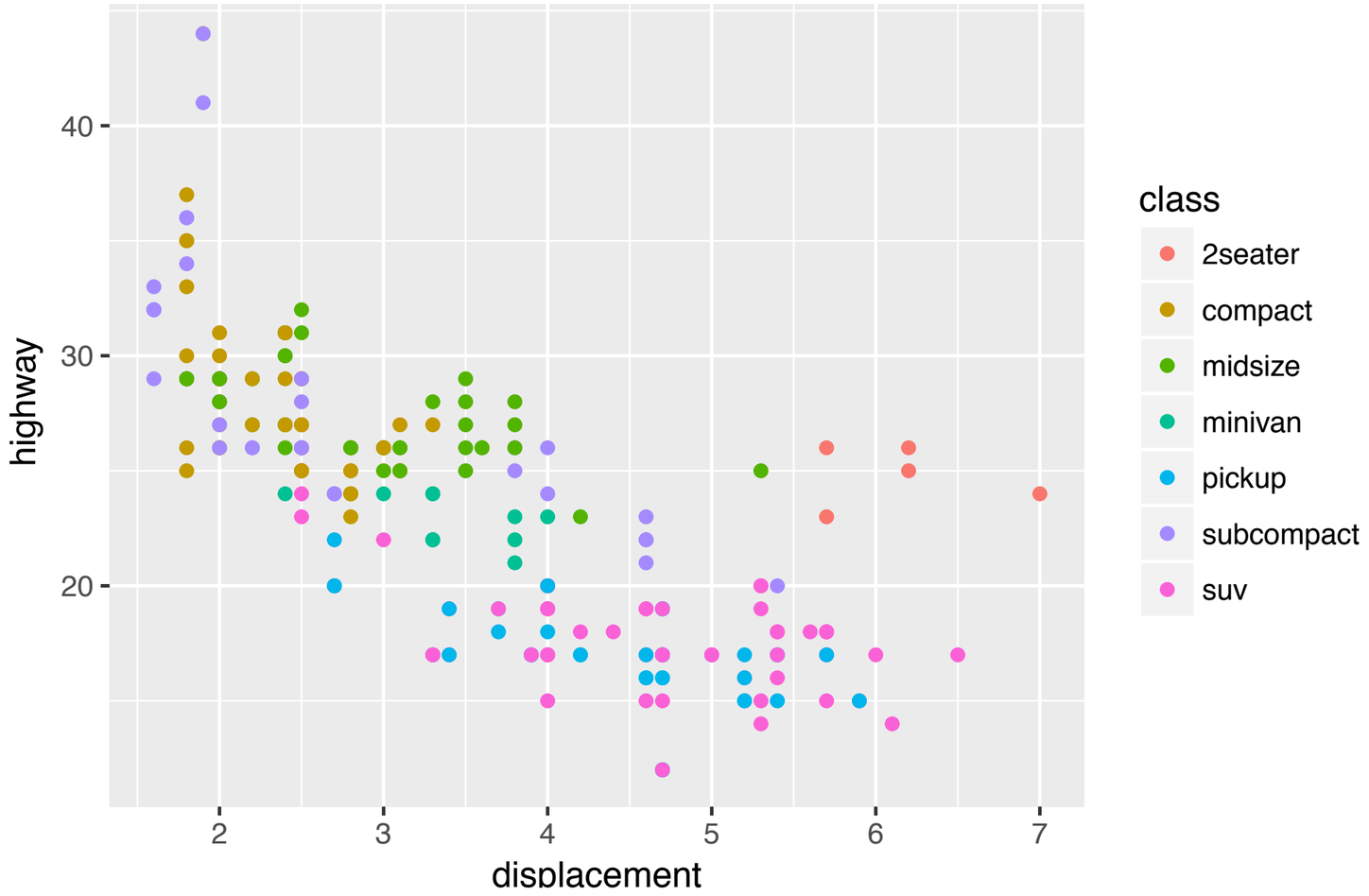
Aesthetic mapping

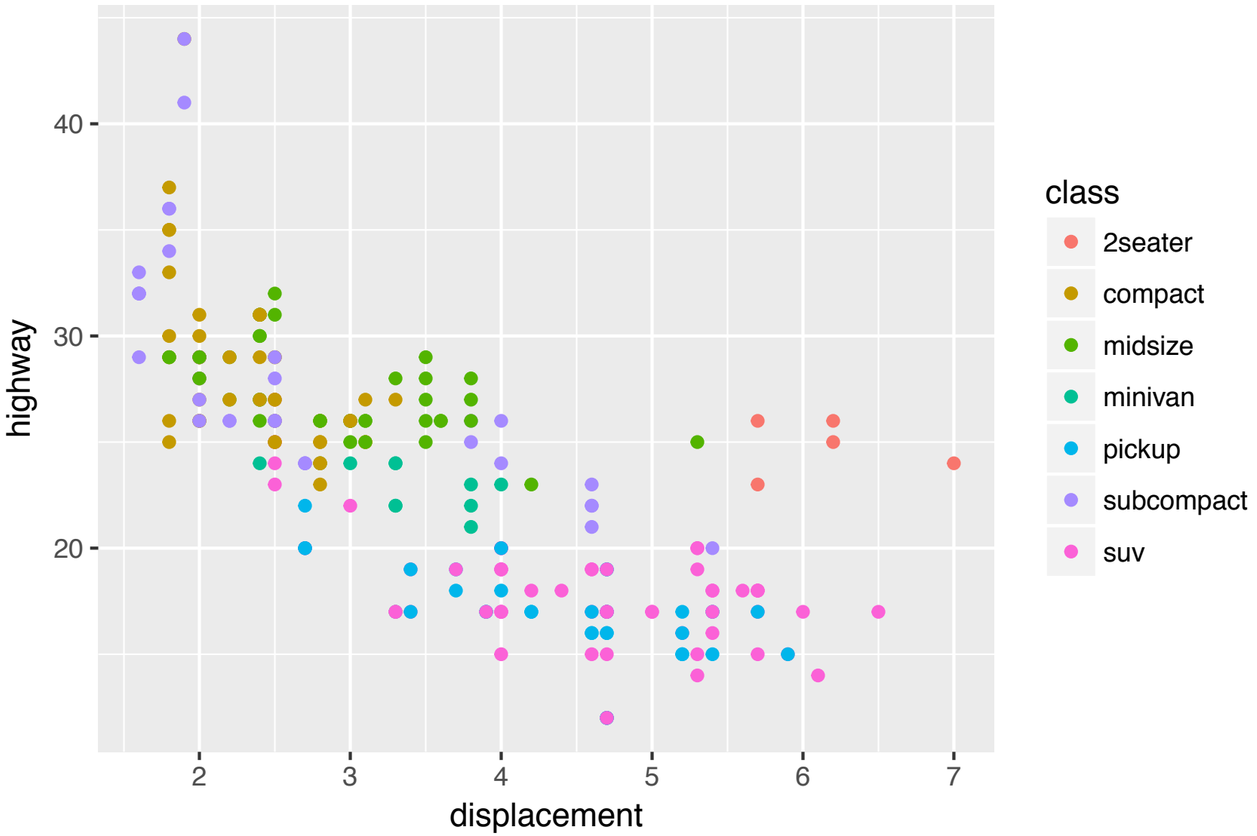
- In `ggplot2`, set with the `aes()` function
- `aes()` **maps variables** to aesthetics
- Different geoms accept different aesthetics
  - `position` (on the x and y axes)
  - `color` ("outside" color)
  - `fill` ("insde" color)
  - `shape`

Key questions

1. What do we want R to do? (What is the goal?)
2. What does R need to know?

How do we make this plot?





How do we make this plot?

1. Goal: scatterplot = plot with points
   - `ggplot() + geom_point()`
2. What does R need to know?
   - data source: `data = mpg`
   - aesthetics:

     ```
     aes(x = displacement,
     y = highway,
     color = class)
     ```

```
ggplot(data = mpg,
  aes(x=displacement, y=highway, color=class)) +
  geom_point()
```