ggplot2: Creating Effective Graphics in R

Patrick Gauding

University of Kansas - Department of Political Science

January 13th, 2018

Overview

Graphics in base R

- Base R has extensive graphic capabilities
  - Will give you what you want, if you're willing to work for it
  - Decent for quick and dirty exploratory plots, but leaves a lot to be desired for publication quality figures
  - Command structure (i.e. argument order) is decentralized, so must learn several differing structures
  - Not flexible with how data are formatted
- Essentially, using base R graphics is consigning yourself to dealing with the detritus of 30+ years of R development!

Graphics in ggplot2

- ggplot2 overcomes several of these difficulties
  - Utilizes the "Grammar of Graphics"
  - Standardized formatting
  - Highly extensible, endless options
  - But, bewildering at the extremes
- A pain at times in the beginning, but well worth the effort to sort out.

The Grammar of Graphics I

- A sequential ordering of essential elements
  - Data
  - Geometries (`geom_`)
  - Aesthetics (`aes()`)
  - Scales (`scale_`)
  - Statistics (`stat_`)
  - Coordinate systems (`coord_`)
  - Facets (`facet_`)
  - Visual Themes (`theme`)

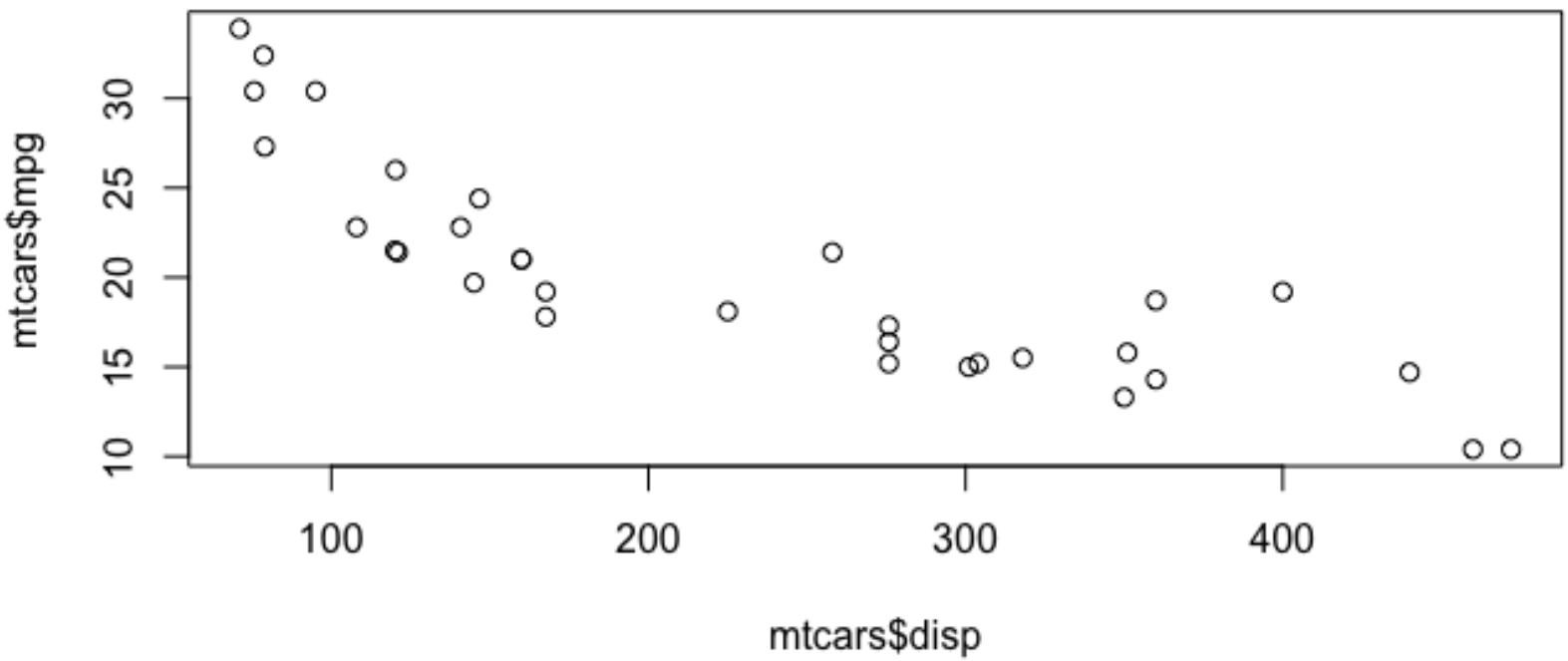
Syntax

```
53  p8 <- ggplot(mtcars, aes(x = disp, y = mpg, fill = factor(cyl))) +
54      geom_point(alpha = 0.6, shape = 20) +
55      geom_smooth(method = "lm", se = FALSE, linetype = 1) +
56      scale_x_continuous(limits = c(70, 500)) +
57      scale_color_discrete() +
58      coord_cartesian() +
59      ggtitle("MPG to Displacement") +
60      labs(x = "Displacement", y = "Miles per Gallon", fill = "Cylinders")
```

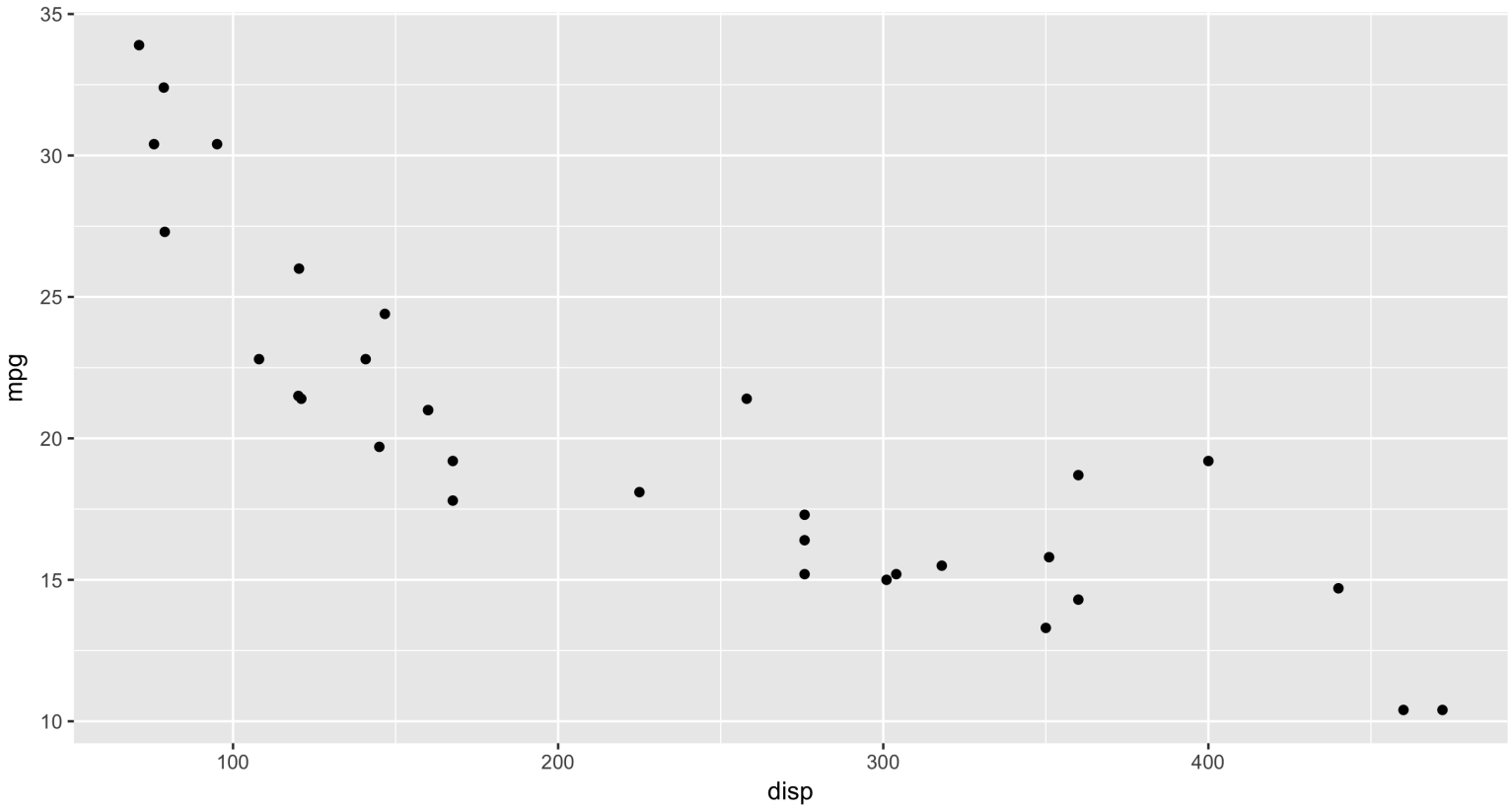
Scatter Plots I

```
17  plot(mtcars$disp, mtcars$mpg)
18  q <- plot(mtcars$disp, mtcars$mpg)
19  q
20
21  ggplot(mtcars, aes(x = disp, y = mpg)) +
22      geom_point()
23  ggsave("../output/plot2.png")
```

Scatter Plots II

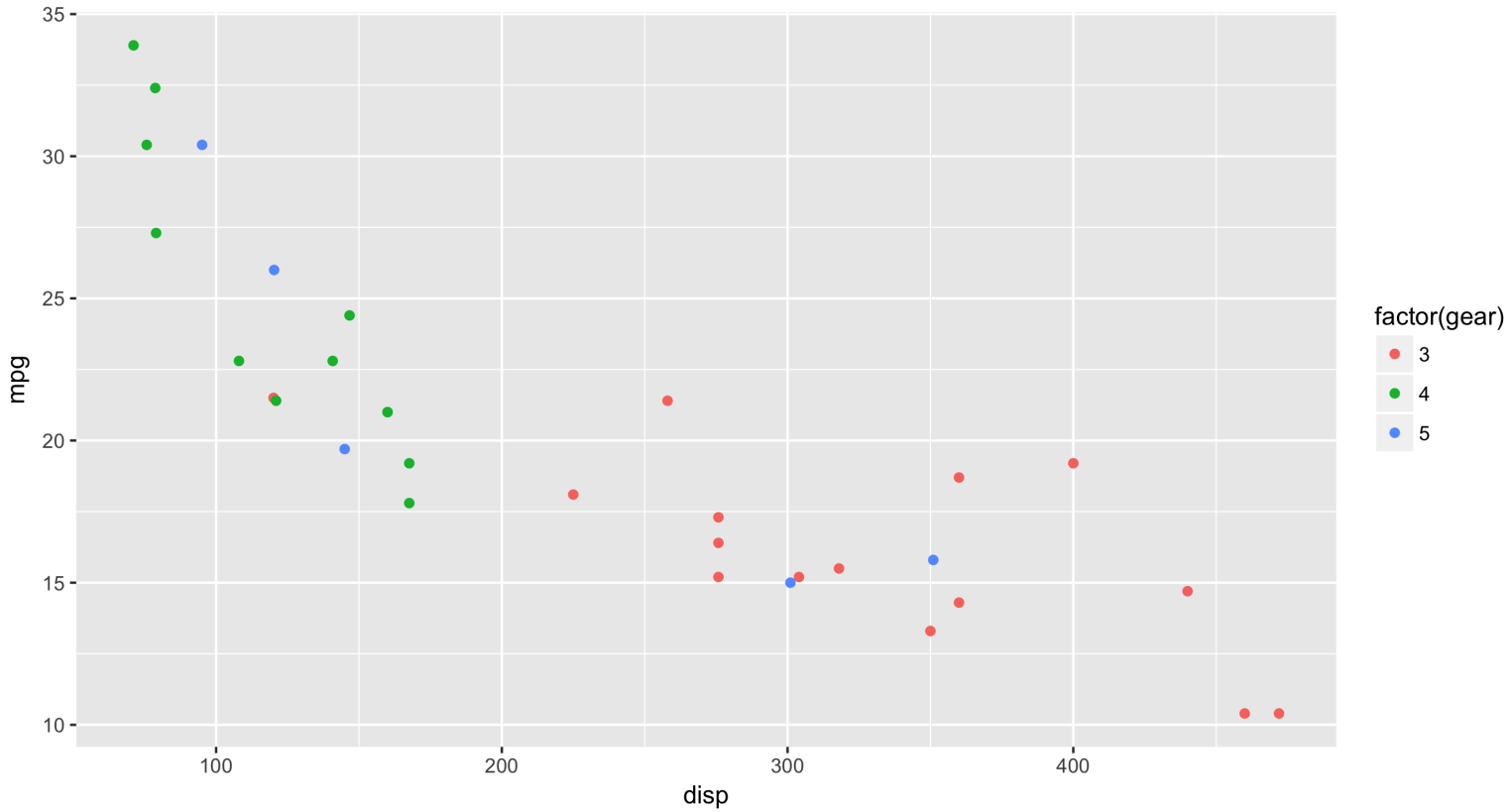


Scatter Plots III



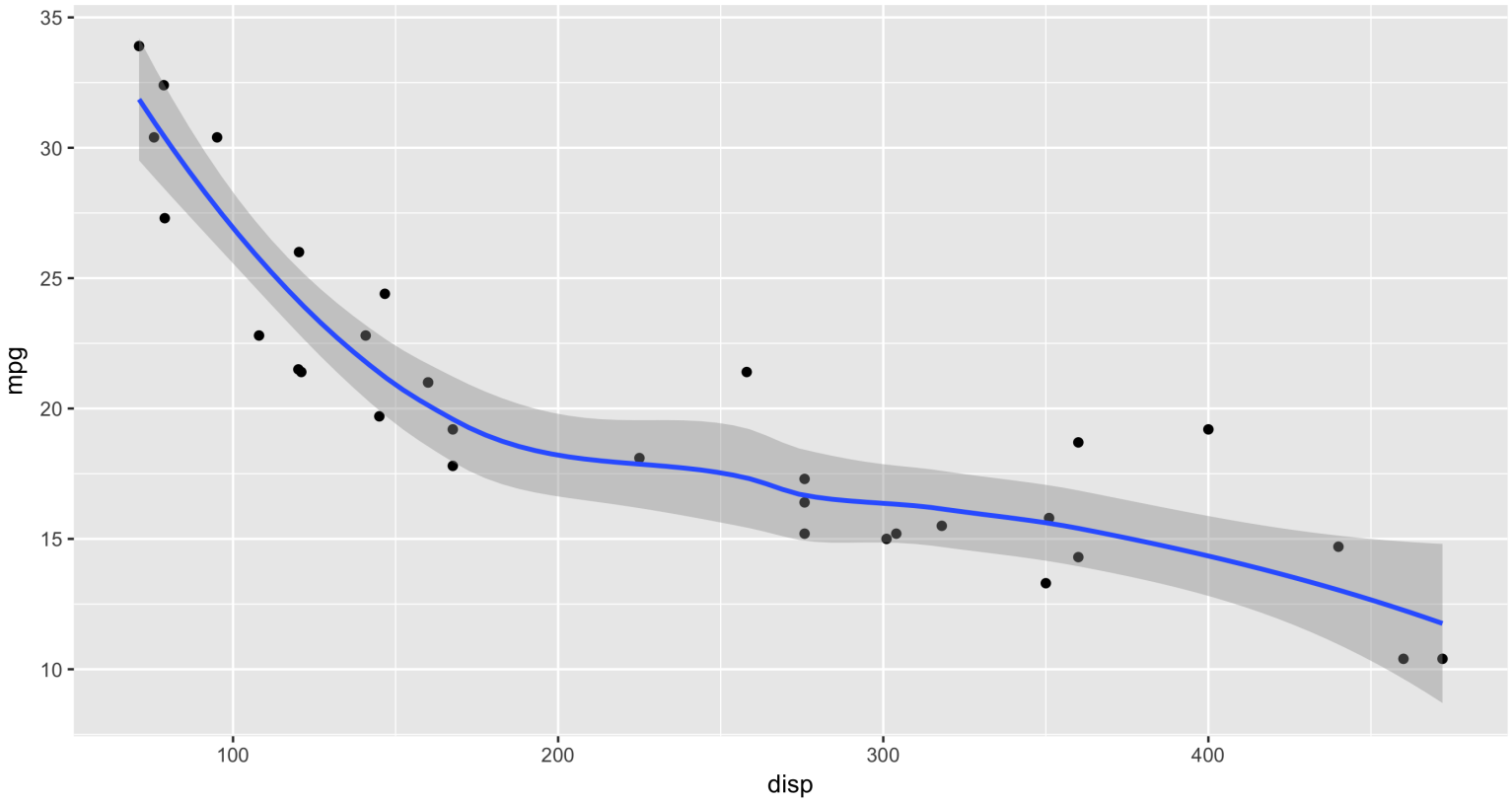
Scatter Plots cont. I

```
25  p <- ggplot(mtcars, aes(x = disp, y = mpg, col = factor(gear))) +
26      geom_point()
27  p
28  ggsave("../output/plot3.png")
```



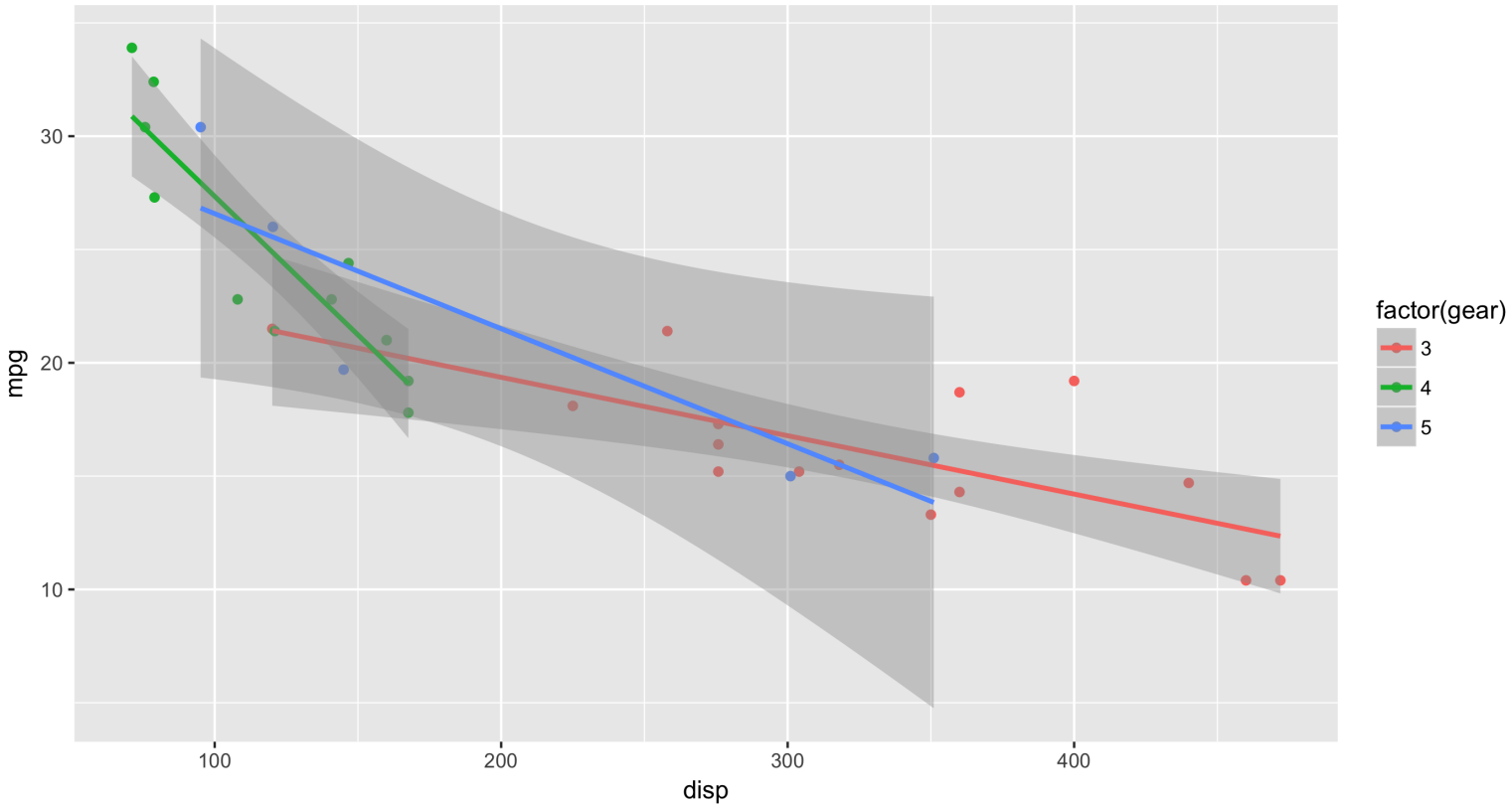
Scatter Plots cont. II

```
30 ggplot(mtcars, aes(x = disp, y = mpg)) +
31     geom_point() +
32     geom_smooth()
33 ggsave("../output/plot4.png")
34
```



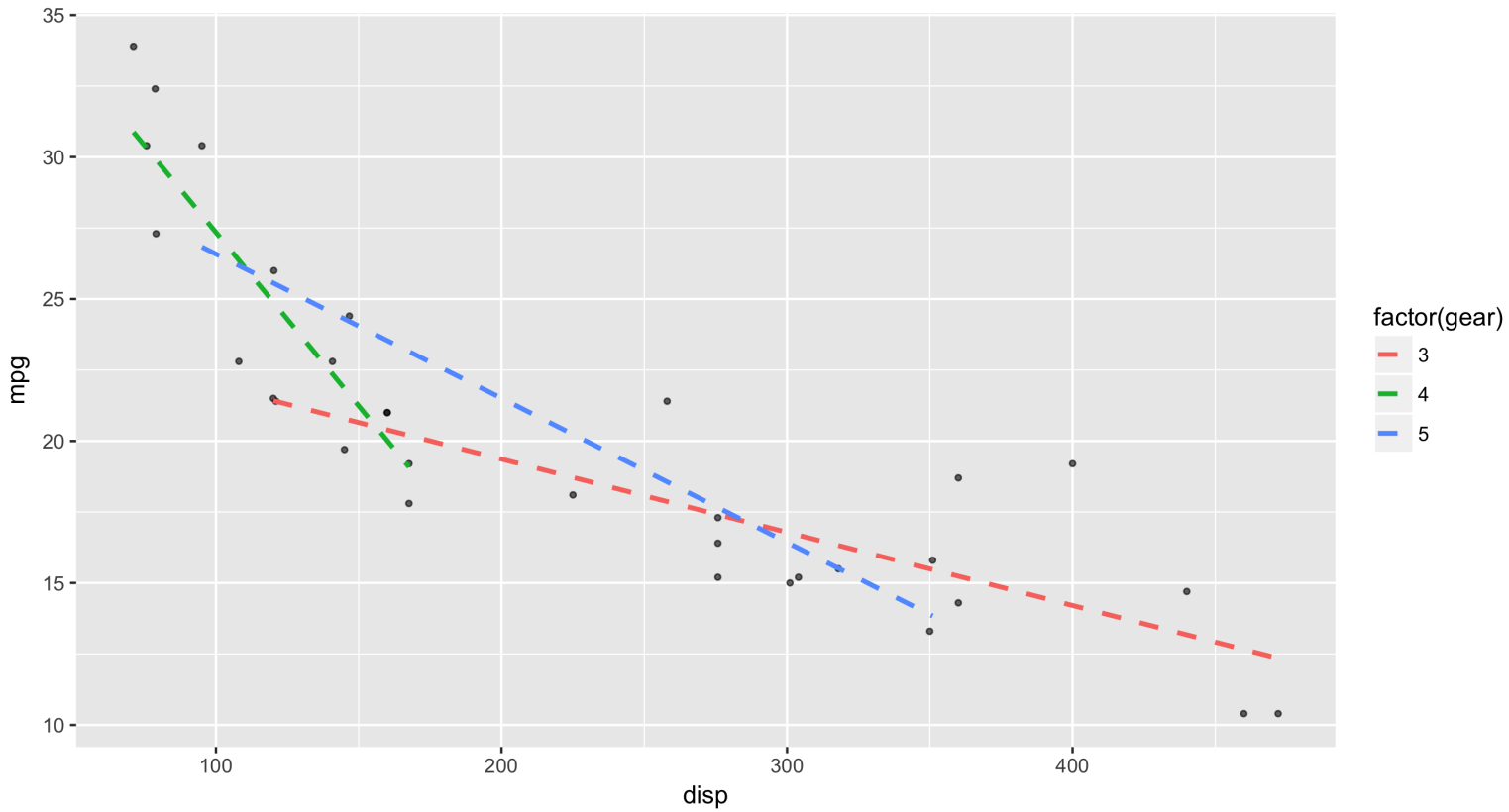
Linear Models I

```
35  ggplot(mtcars, aes(x = disp, y = mpg, col = factor(gear))) +
36      geom_point() +
37      geom_smooth(method = "lm")
38  p + geom_smooth(method = "lm")
39  ggsave("../output/plot5.png")
```



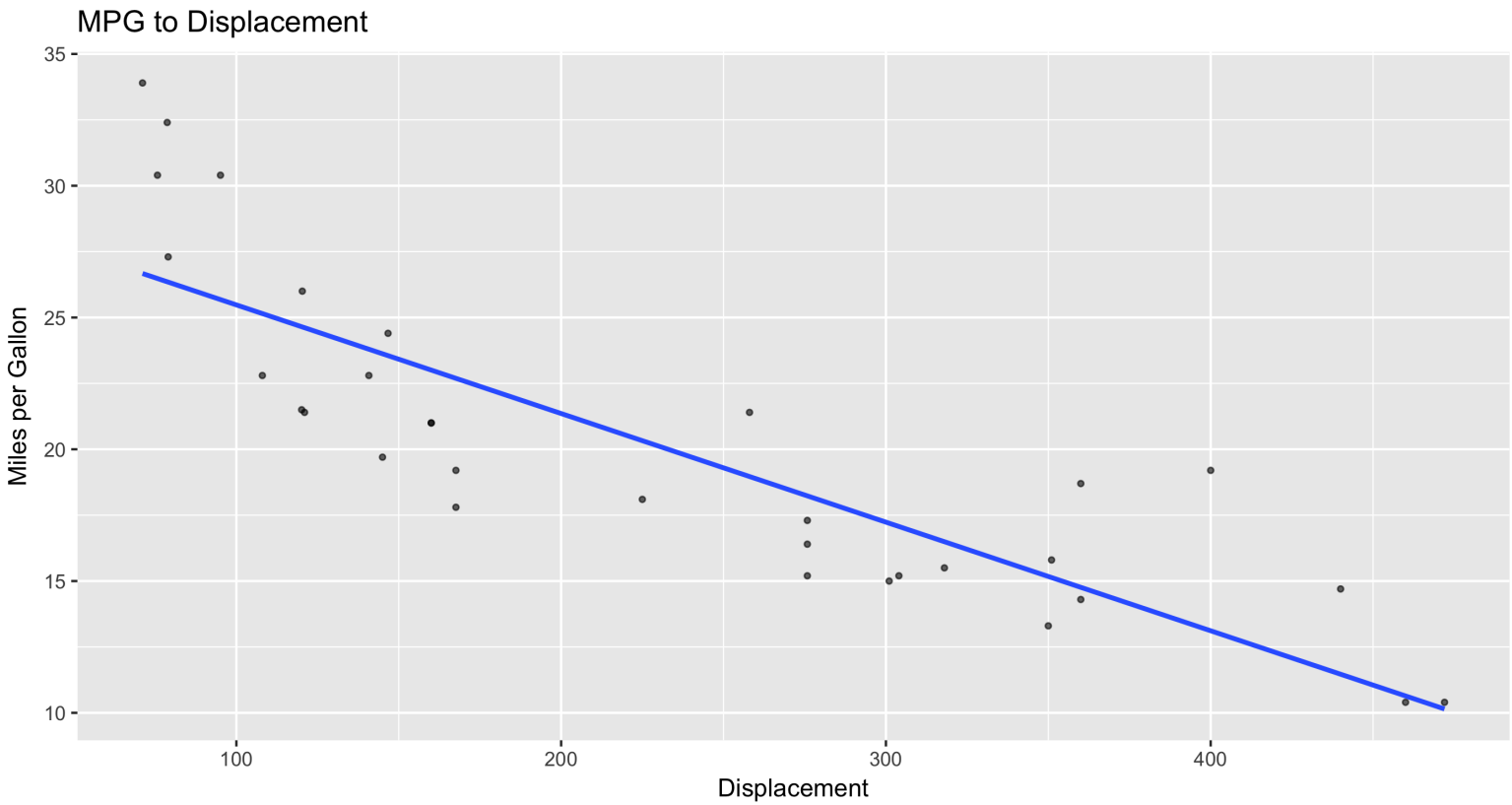
Linear Models II

```
41  ggplot(mtcars, aes(x = disp, y = mpg, col = factor(gear))) +
42      geom_point(col = "black", shape = 20, alpha = 0.6) +
43      geom_smooth(method = "lm", se = FALSE, linetype = 2, alpha = 0.6)
44  ggsave("../output/plot6.png")
45
```



Linear Models III

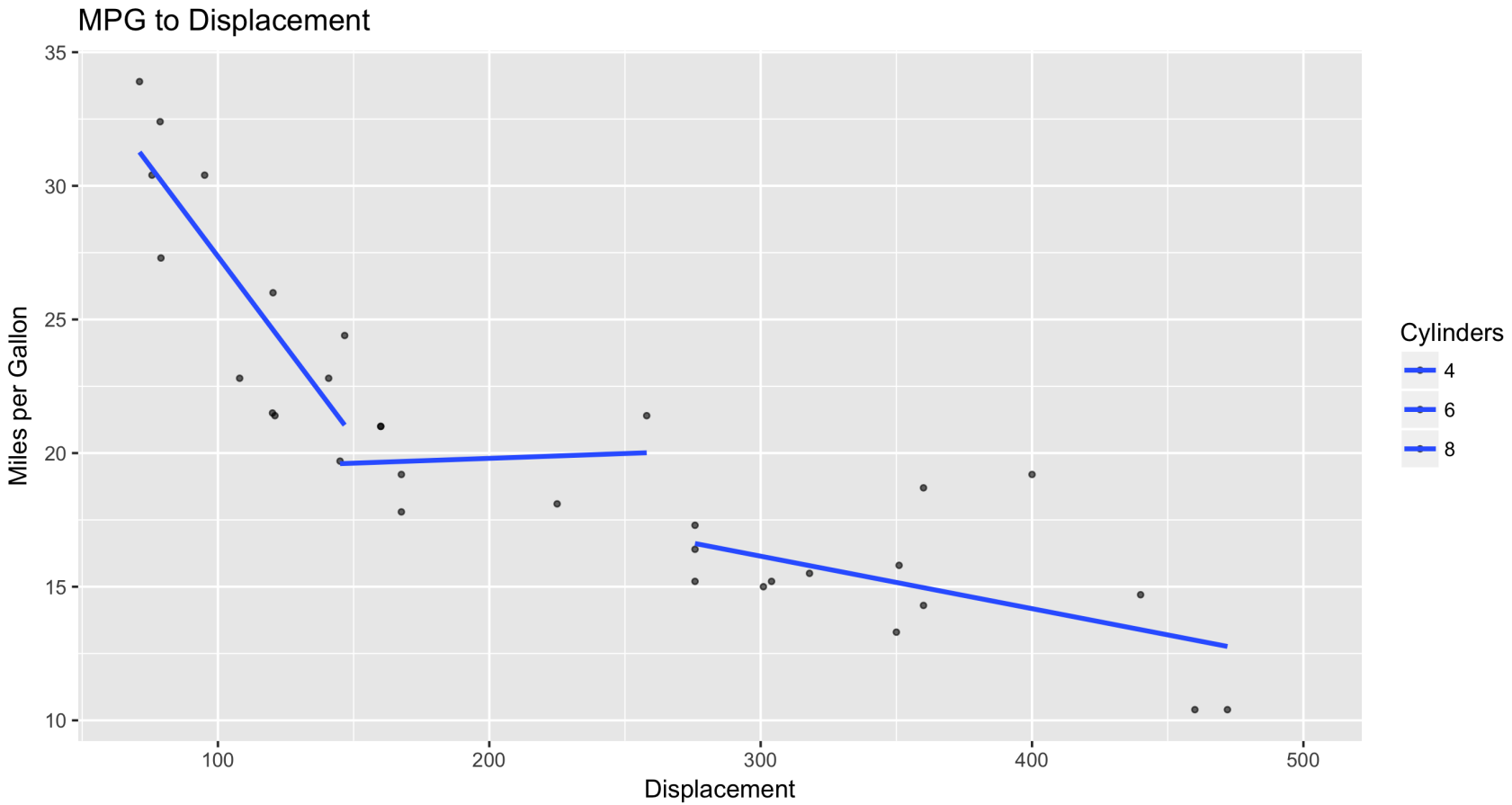
```
46  ggplot(mtcars, aes(x = disp, y = mpg)) +
47      geom_point(alpha = 0.6, shape = 20) +
48      geom_smooth(method = "lm", se = FALSE, linetype = 1) +
49      ggtitle("MPG to Displacement") +
50      labs(x = "Displacement", y = "Miles per Gallon", fill = "Cylinders")
51  ggsave("../output/plot7.png")
```



Linear Models IV

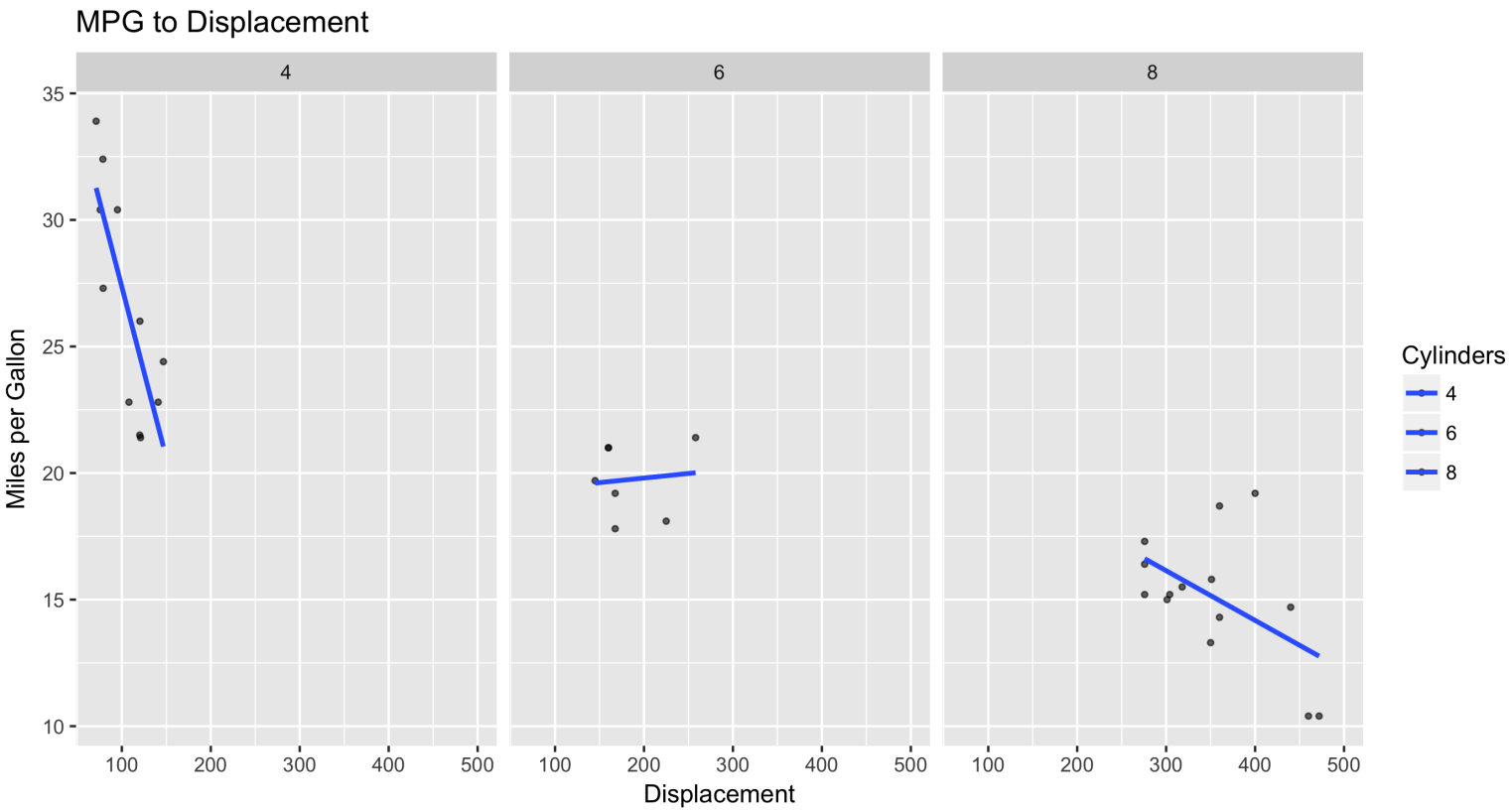
```
53  p8 <- ggplot(mtcars, aes(x = disp, y = mpg, fill = factor(cyl))) +
54      geom_point(alpha = 0.6, shape = 20) +
55      geom_smooth(method = "lm", se = FALSE, linetype = 1) +
56      scale_x_continuous(limits = c(70, 500)) +
57      scale_color_discrete() +
58      coord_cartesian() +
59      ggtitle("MPG to Displacement") +
60      labs(x = "Displacement", y = "Miles per Gallon", fill = "Cylinders")
61  p8
62  ggsave("../output/plot8.png")
```

Linear Models V



Facets

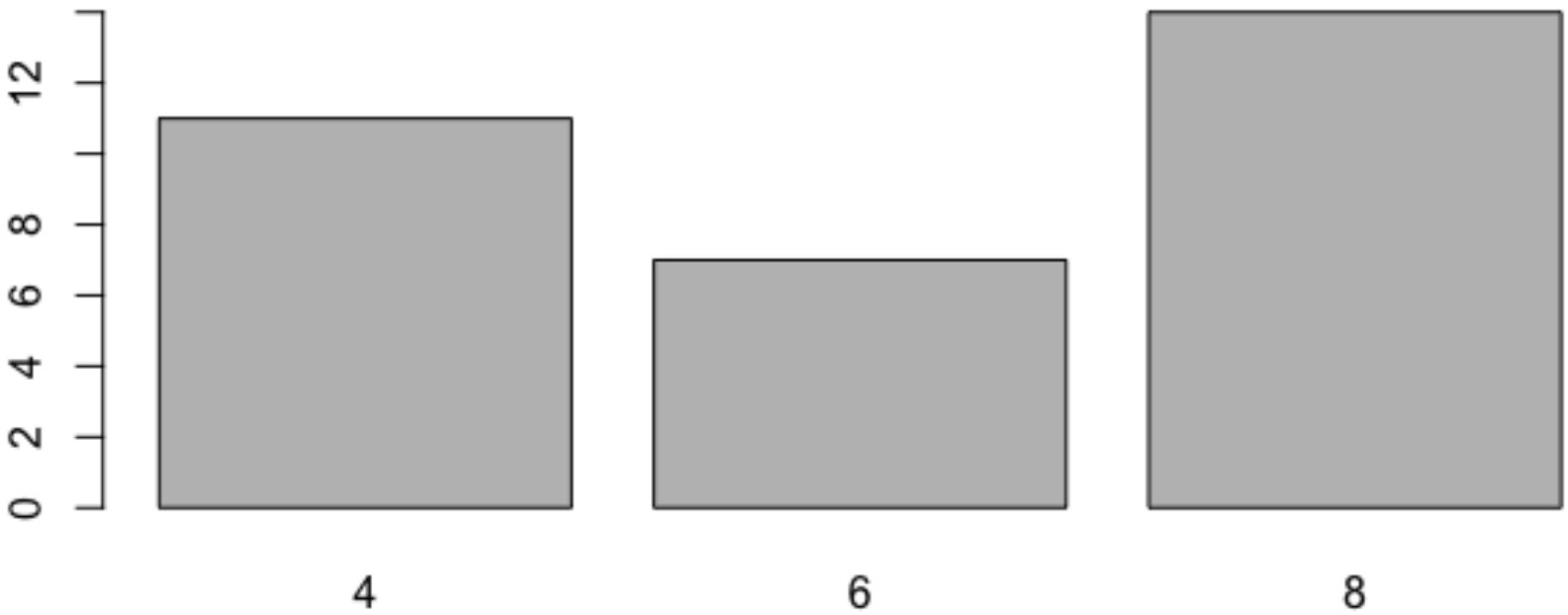
```
64  p8 + facet_grid(. ~ cyl)
65  ggsave('../output/plot8a.png')
```



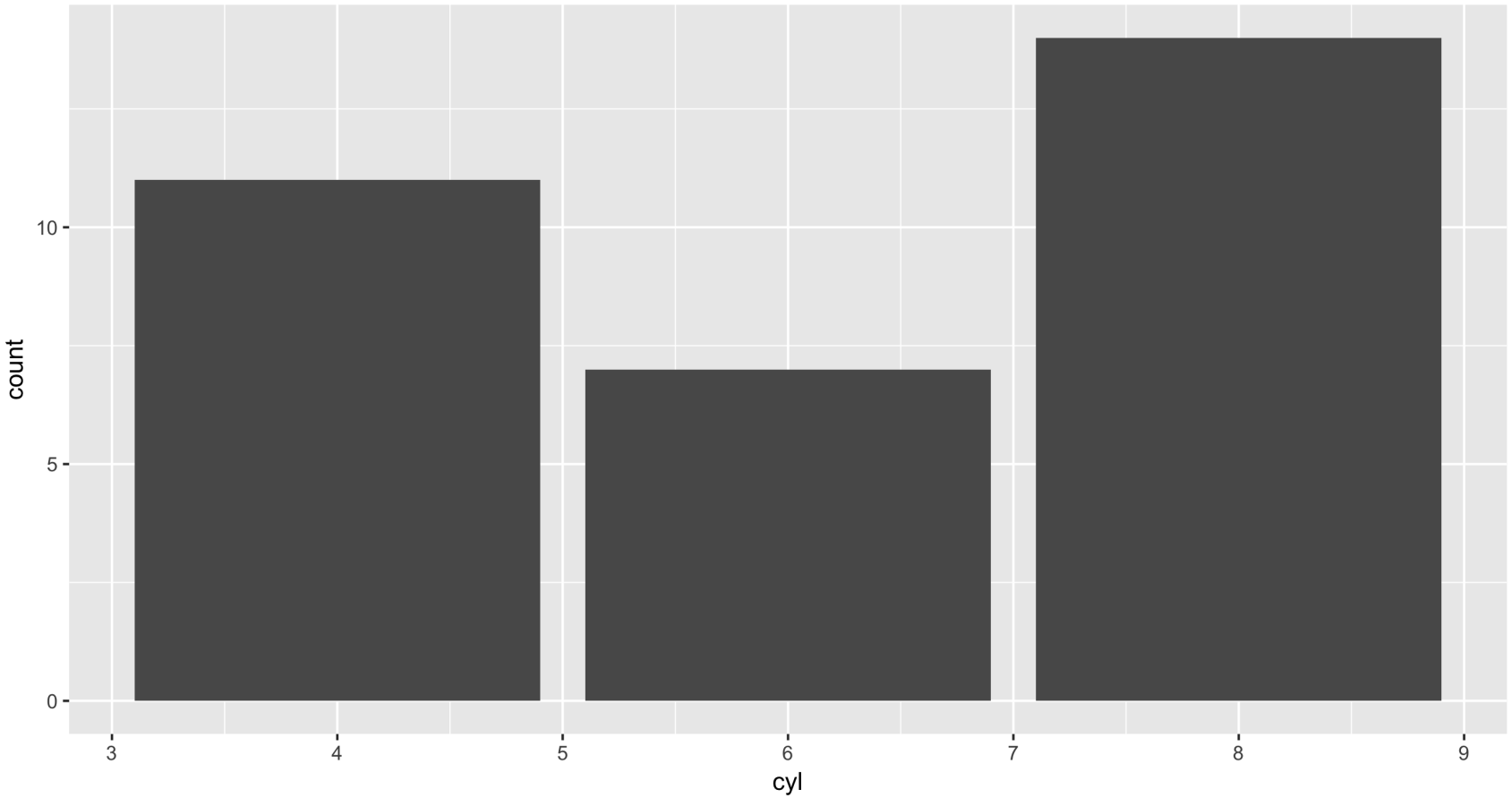
Bar Plots I

```
69  bartab <- table(mtcars$cyl)
70  barplot(bartab, beside = TRUE)
71
72  ggplot(mtcars, aes(x = cyl)) +
73    geom_bar()
74  ggsave('../output/plot10.png')
75
```

Bar Plots II



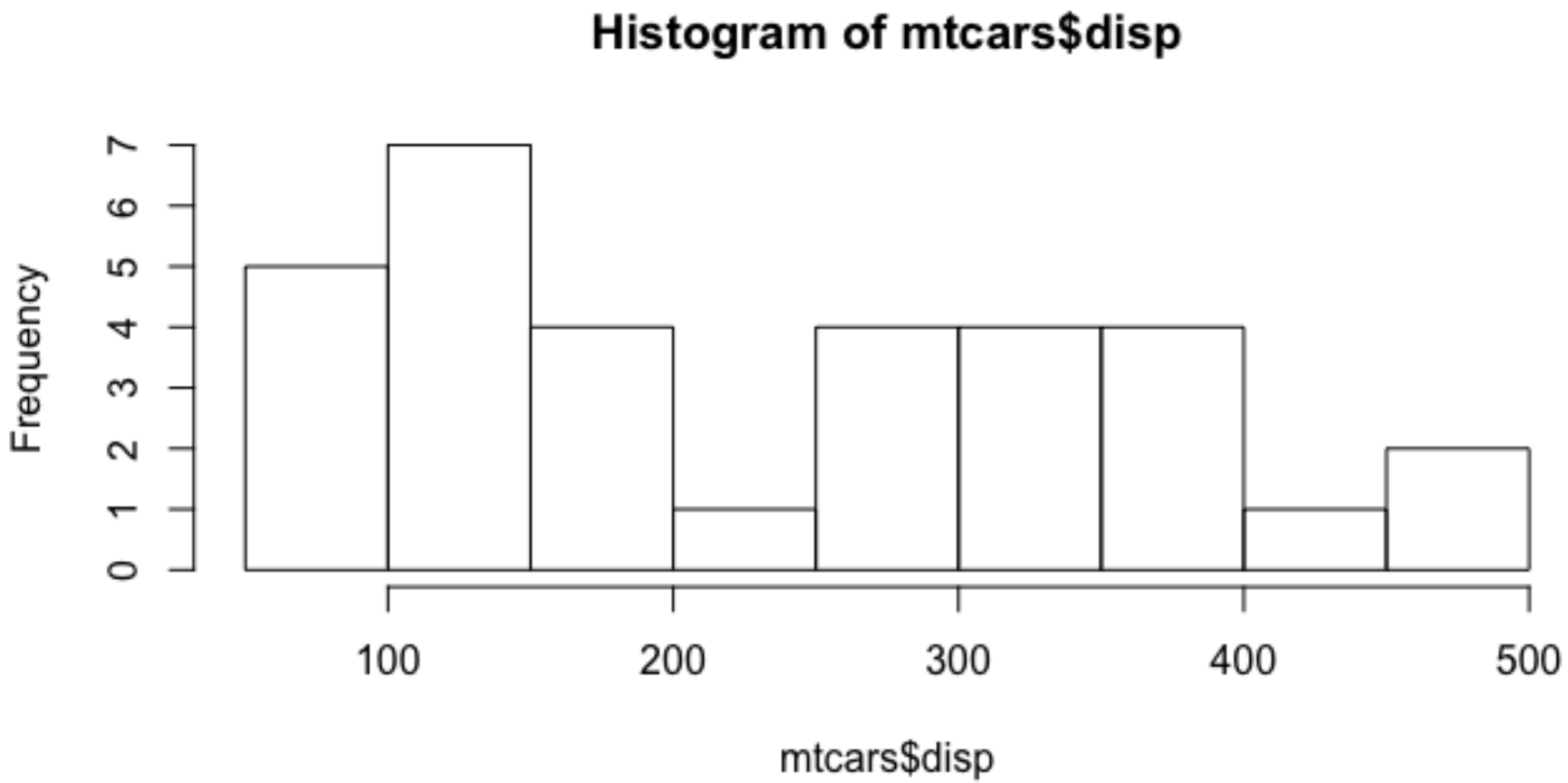
Bar Plots III



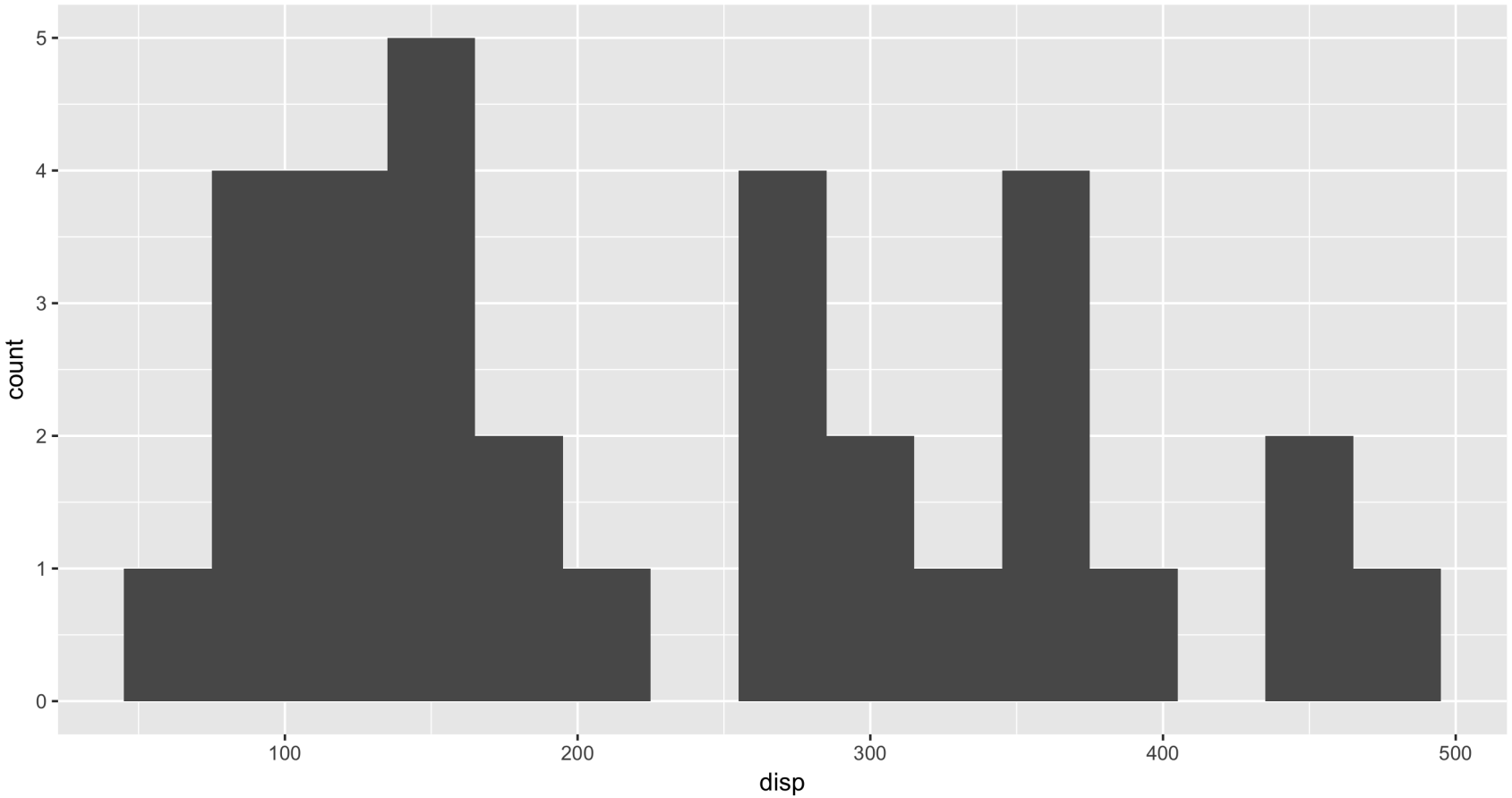
Histograms I

```
79  hist(mtcars$disp)
80
81  ggplot(mtcars, aes(x = disp)) +
82    geom_histogram(binwidth = 30)
83  ggsave('../output/plot12.png')
84
```

Histograms II



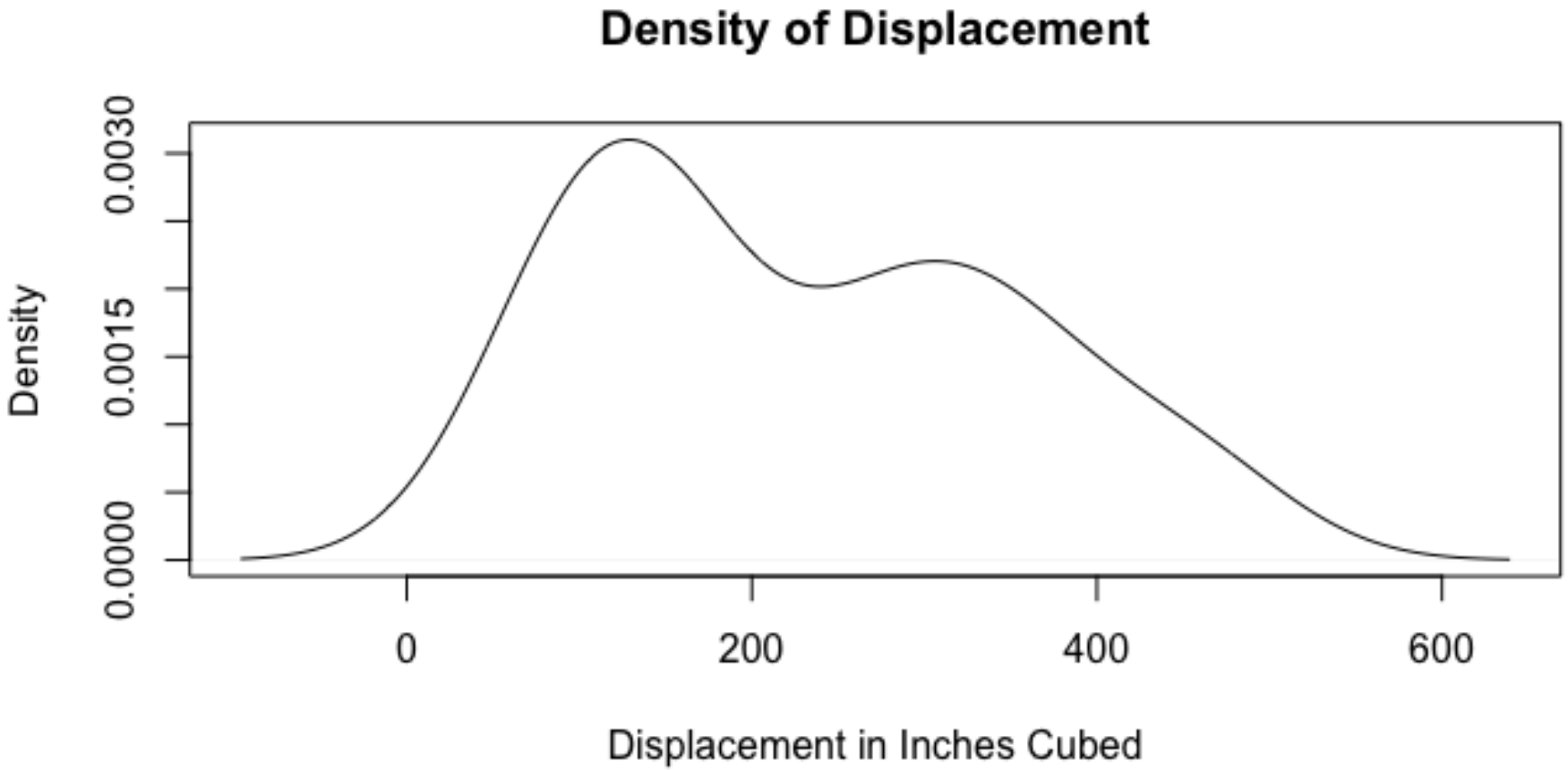
Histograms III



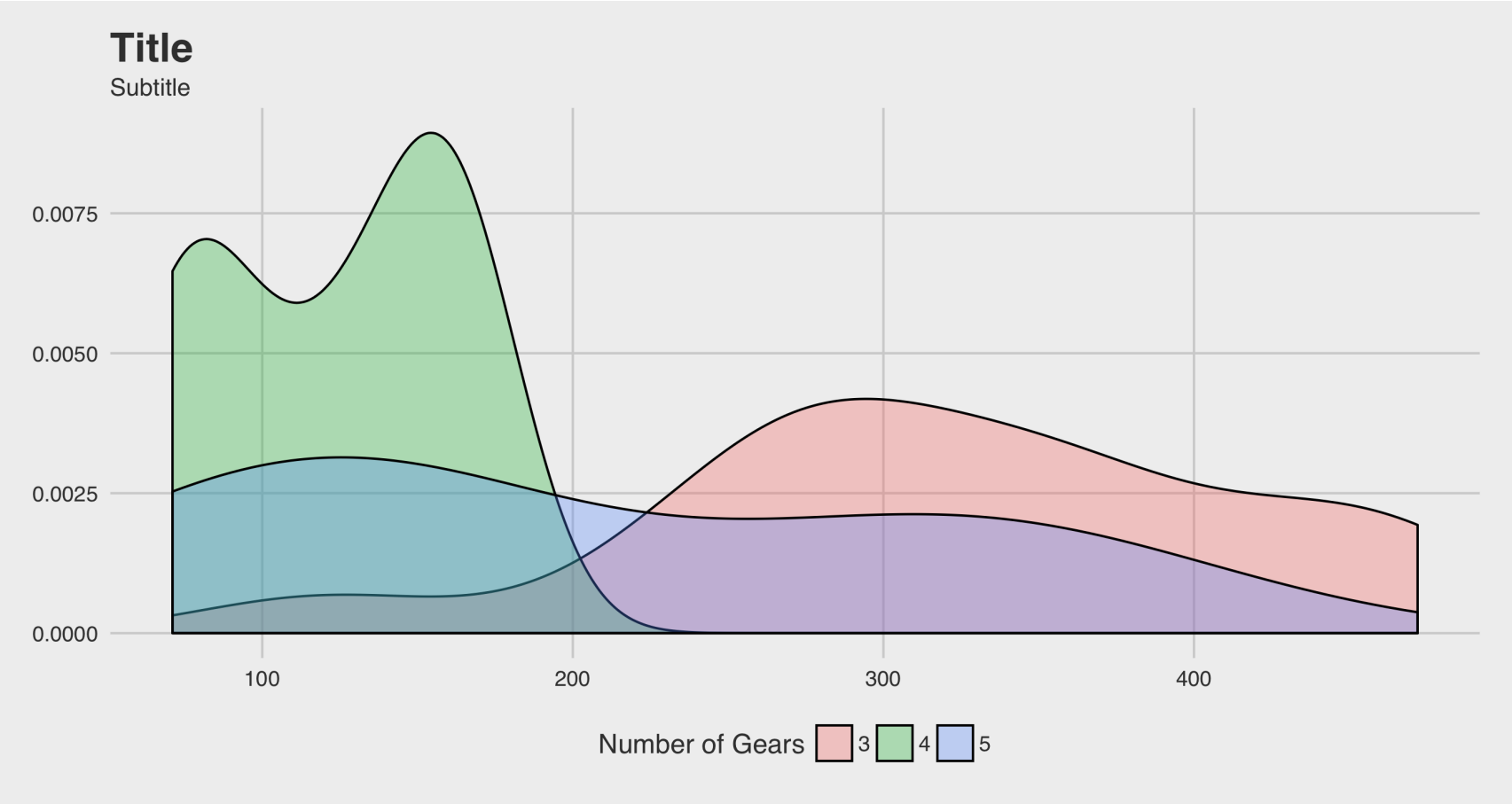
Density Plots I

```
87  plot(density(mtcars$disp),
88       main = "Density of Displacement",
89       xlab = "Displacement in Inches Cubed",
90       ylab = "Density")
91
92  ggplot(mtcars, aes(x = disp, fill = factor(gear))) +
93       geom_density(kernel = "gaussian", alpha = 0.3) +
94       ggtitle("Title", "Subtitle") +
95       labs(x = "Displacement", y = "Density", fill = "Number of Gears") +
96       theme_fivethirtyeight()
97  ggsave('../output/plot14.png')
98
```

Density Plots II



Density Plots III



Mapping I

```
119 m <- ggplot(fifty, aes(map_id = state)) +
120   geom_map(map = fifty_states) +
121   geom_map(fill="white", map= fifty_states, color="black") +
122   geom_map(data = listdf, map = fifty_states, color = "white") +
123   expand_limits(x = fifty_states$long, y = fifty_states$lat) +
124   coord_map() +
125   labs(x = "", y = "") +
126   ggtitle("States with At Least One LGBT Judge") +
127   theme_fivethirtyeight()
128 m
129 ggsave('../output/plot15.png')
```

Mapping II