Day 2 - Advanced Graphics in R

02 - Plotting Map Data

Karsten Maurer and Susan Vander Plas

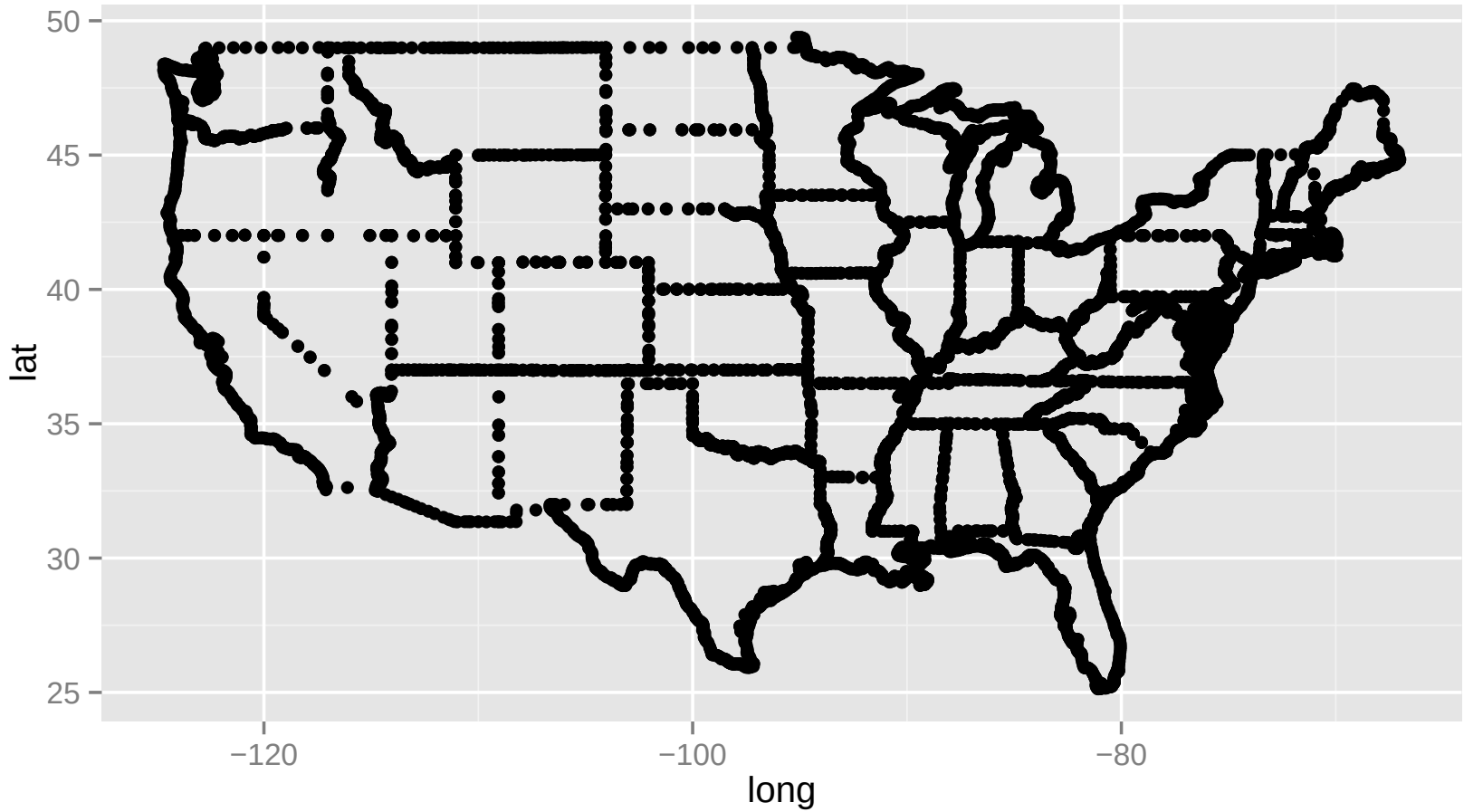
Iowa State University

March 12, 2014

What is a Map?

A bunch of latitude longitude points...

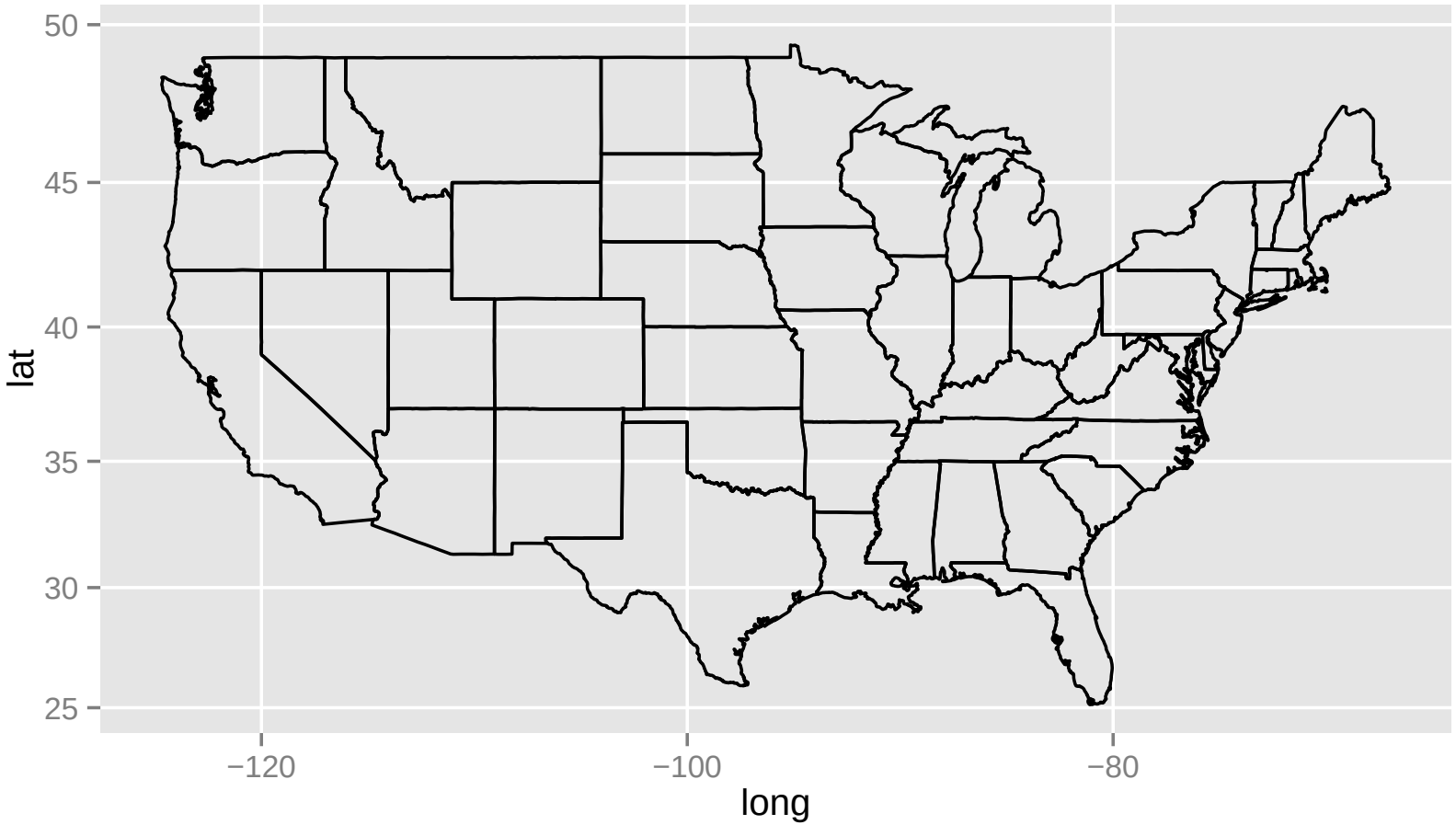
```
qplot(long, lat, geom = "point", data = states)
```



What is a Map?

... that are connected with lines in a very specific order.

```
qplot(long, lat, geom = "path", data = states, group = group) + coord_m
```



Basic Map Data

What needs to be in the data set in order to plot a basic map?

- Need latitude/longitude points for all map boundaries
- Need to know which boundary group all lat/long points belong
- Need to know the order to connect points within each group

Data for Building Basic State Map

Our states data has all necessary information

```
states <- map_data("state")
head(states)
```

```
##      long   lat group order  region subregion
## 1 -87.46 30.39     1     1 alabama      <NA>
## 2 -87.48 30.37     1     2 alabama      <NA>
## 3 -87.53 30.37     1     3 alabama      <NA>
## 4 -87.53 30.33     1     4 alabama      <NA>
## 5 -87.57 30.33     1     5 alabama      <NA>
## 6 -87.59 30.33     1     6 alabama      <NA>
```

Incorporating Information About States

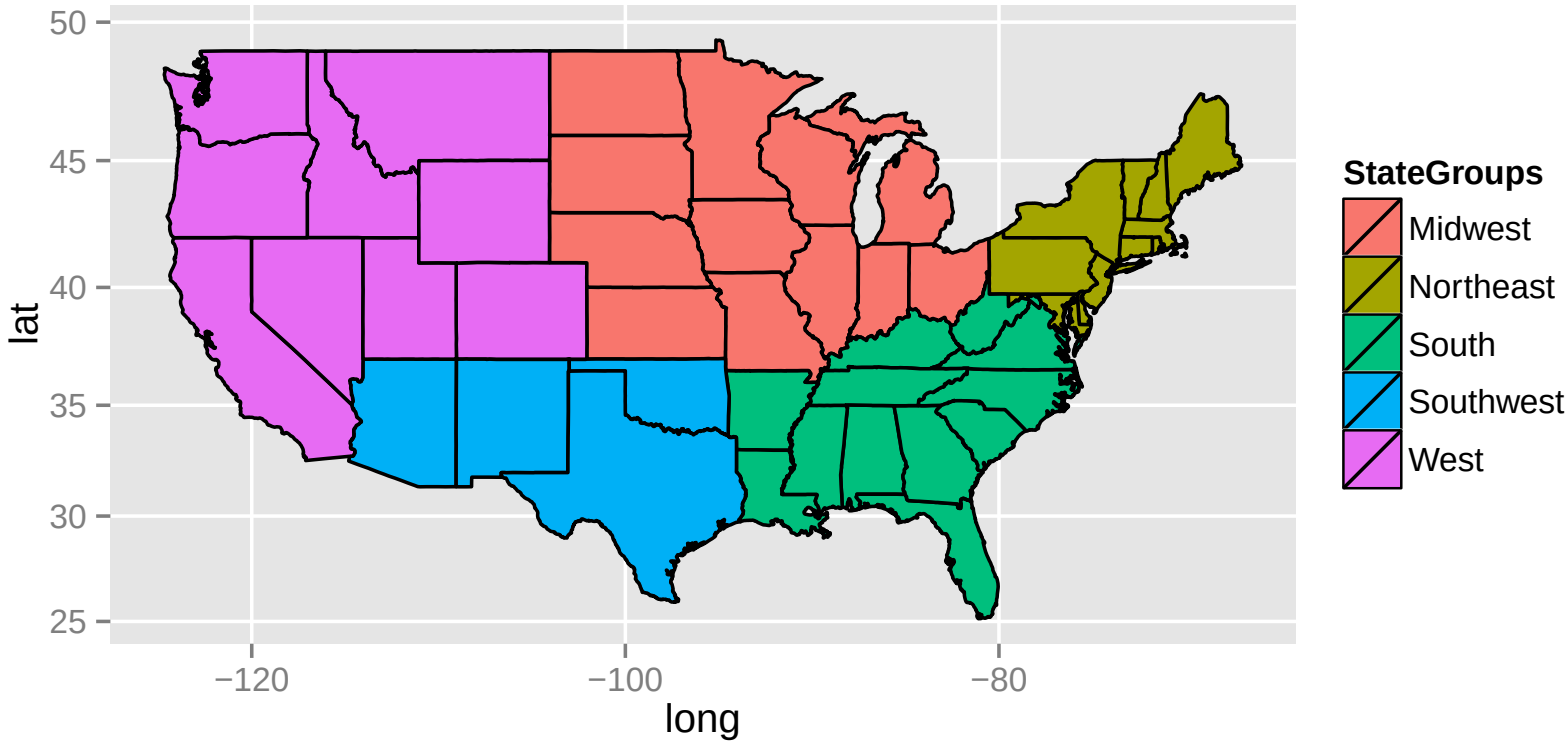
Want to incorporate additional information into the plot:

- Add other geographic information by adding geometric layers to the plot
- Add non-geopgraphic information by altering the fill color for each state
  - Use `geom="polygon"` to treat states as solid shapes to add color
  - Incorporate numeric information using color shade or intensity
  - Incorporate categorical informaion using color hue

Categorical Information Using Hue

If a categorical variable is assigned as the fill color then `qplot` will assign different hues for each category

```
qplot(long, lat, geom = "polygon", data = states.class.map, group = gro
    colour = I("black")) + coord_map()
```



Numerical Information Using Shade and Intensity

To show how was can add numerical information to map plots we will use the BRFSS data

- Behavioral Risk Factor Surveillance System
- 2008 telephone survey run by the Center for Disease Control (CDC)
- Ask a variety of questions related to health and wellness
- Cleaned data with state aggregated values posted on website

BRFSS Data Aggregated by State

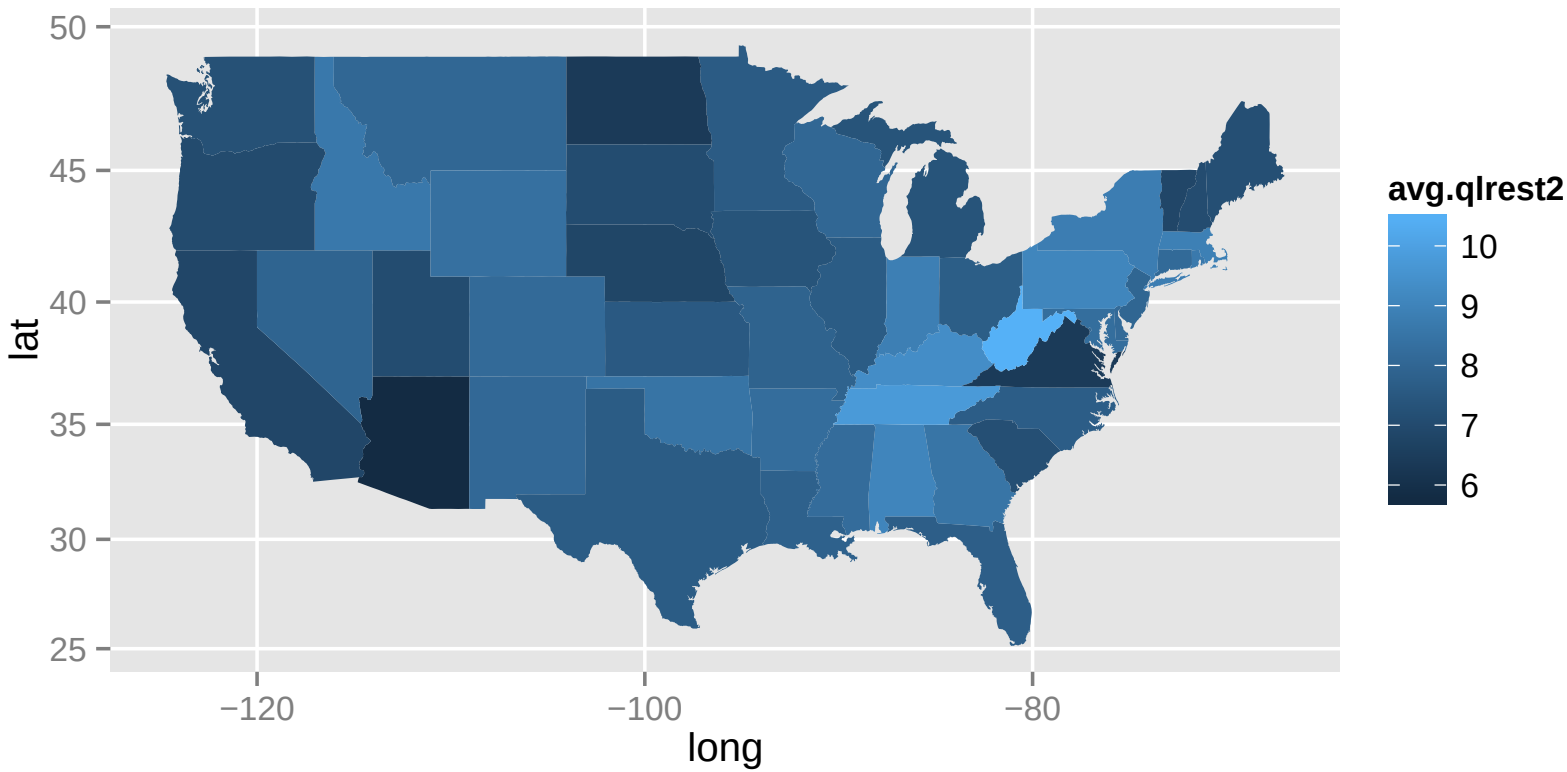
```
head(states.stats)
```

```
##   state.name avg.wt avg.qlrest2 avg.ht avg.bmi avg.drnk
## 1    alabama  180.7       9.051  168.0   29.00    2.333
## 2     alaska  189.3       8.381  172.1   28.91    2.324
## 3    arizona  169.7       5.770  168.3   27.05    2.407
## 4   arkansas  177.4       8.227  168.8   28.02    2.312
## 5 california  170.0       6.848  168.1   27.23    2.170
## 6   colorado  167.2       8.135  169.6   26.17    1.971
```

Numerical Information Using Shade and Intensity

Average number of days in the last 30 days of insufficient sleep by state

```
qplot(long, lat, geom = "polygon", data = states.map, group = group, fi
    coord_map()
```



BRFSS Data Aggregated by State and Gender

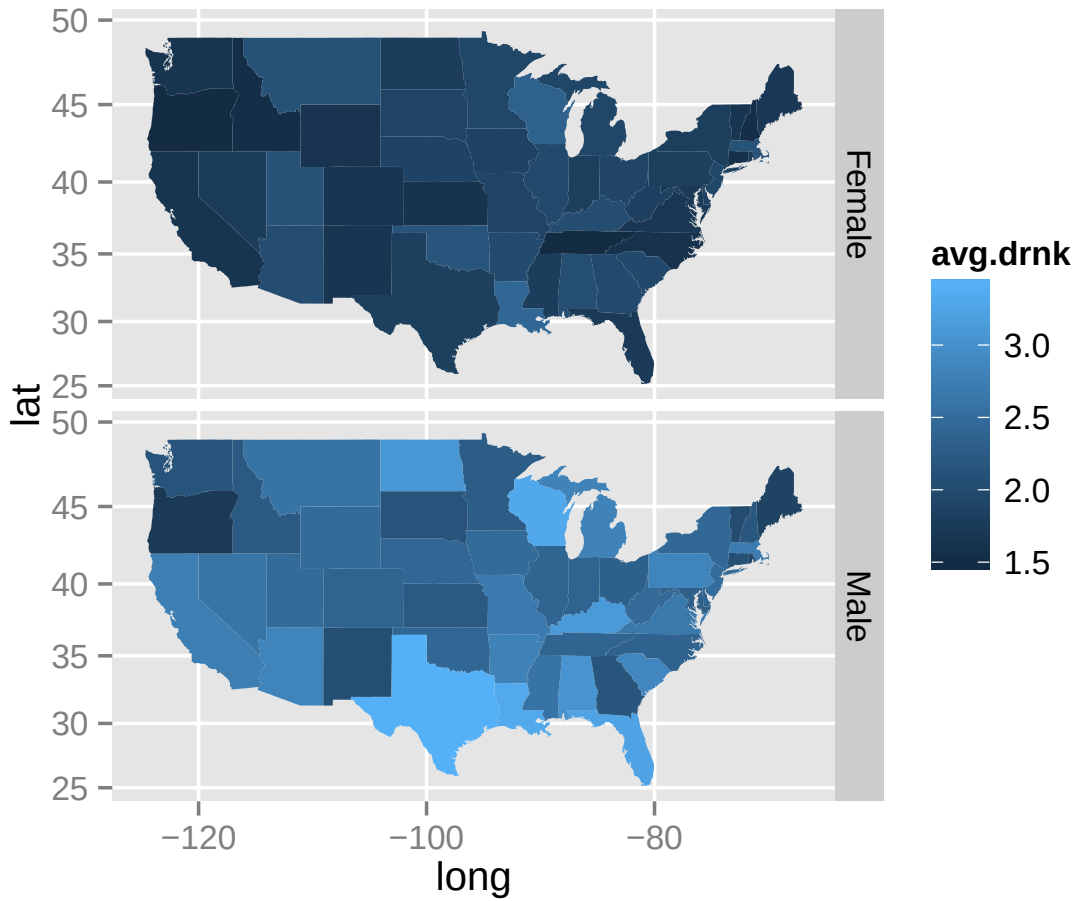
```
head(states.sex.stats)
```

```
##   state.name SEX avg.wt avg.qlrest2 avg.ht avg.bmi avg.drnk    sex
## 1    alabama   1  198.9       8.649  177.6   28.51    3.033   Male
## 2    alabama   2  173.0       9.225  164.0   29.21    2.042 Female
## 3     alaska   1  203.4       7.236  178.4   28.91    2.487   Male
## 4     alaska   2  169.6       9.907  163.1   28.89    2.103 Female
## 5    arizona   1  191.4       5.164  177.2   27.63    2.814   Male
## 6    arizona   2  156.2       6.143  162.7   26.68    2.027 Female
```

Adding Numerical Information

Average number of alcoholic drinks per day by state and gender

```
qplot(long, lat, geom = "polygon", data = states.sex.map, group = group
    coord_map() + facet_grid(sex ~ .)
```



Your Turn

Use `merge` to combine child healthcare data with maps information

Then use `qplot` to create a map of child healthcare undercoverage rate by state

Cleaning Up Your Maps

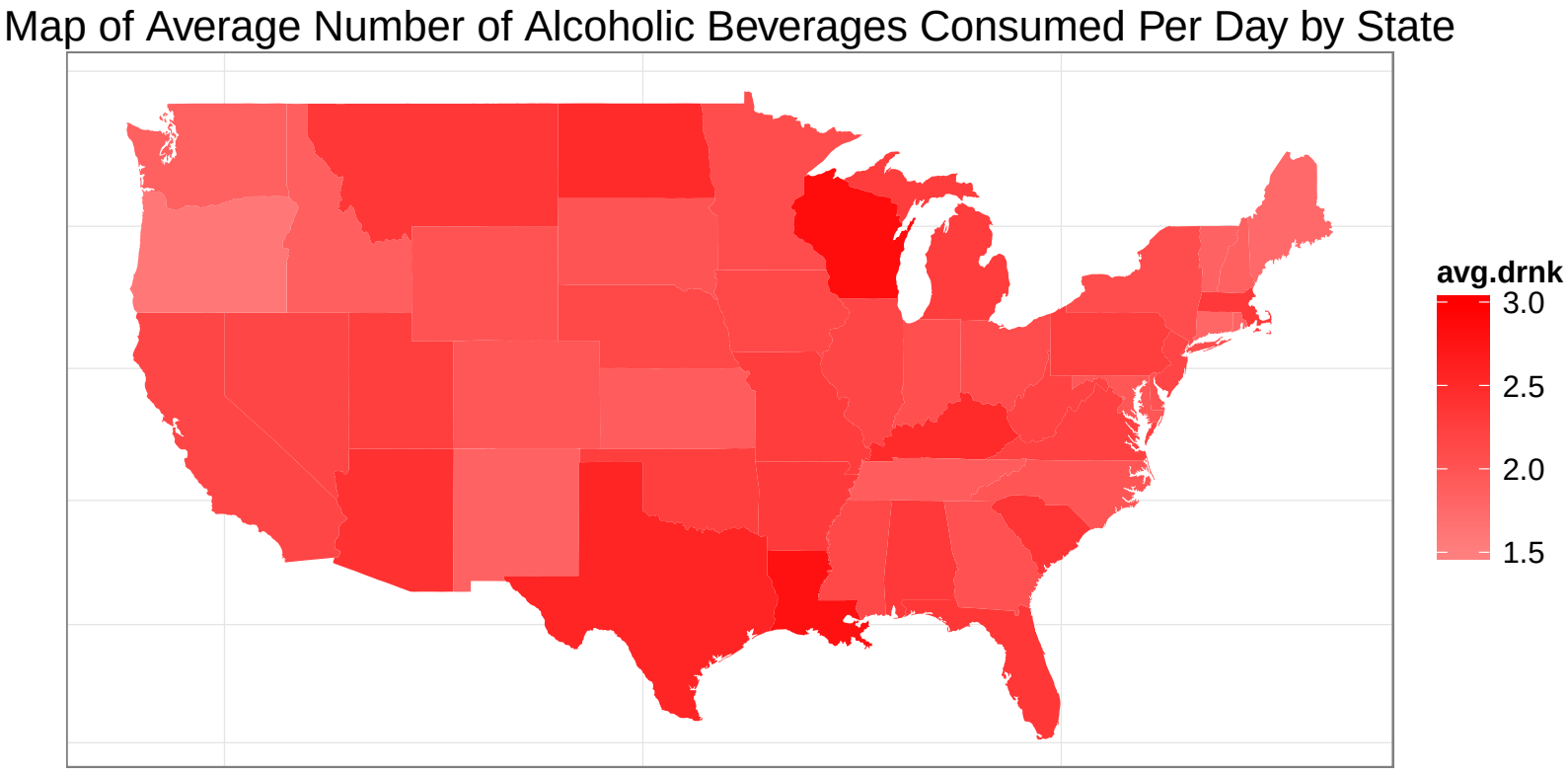
Use ggplot2 options to clean up your map!

- Adding Titles `+ ggtitle(...)`
- Might want a plain white background `+ theme_bw()`
- Extremely familiar geography may eliminate need for latitude and longitude axes `+ theme(...)`
- Want to customize color gradient `+ scale_fill_gradient2(...)`
- Keep aspect ratios correct `+ coord_map()`

Cleaning Up Your Maps

```
qplot(long, lat, geom="polygon", data=states.map, group=group, fill=avg
  coord_map() +  theme_bw() +
  scale_fill_gradient2(limits=c(1.5, 3),low="lightgray",high="red") +
  theme(axis.ticks = element_blank(),
        axis.text.x = element_blank(),
        axis.title.x=element_blank(),
        axis.text.y = element_blank(),
        axis.title.y=element_blank()) +
  ggtitle("Map of Average Number of Alcoholic Beverages Consumed Per Da
```

Cleaning Up Your Maps

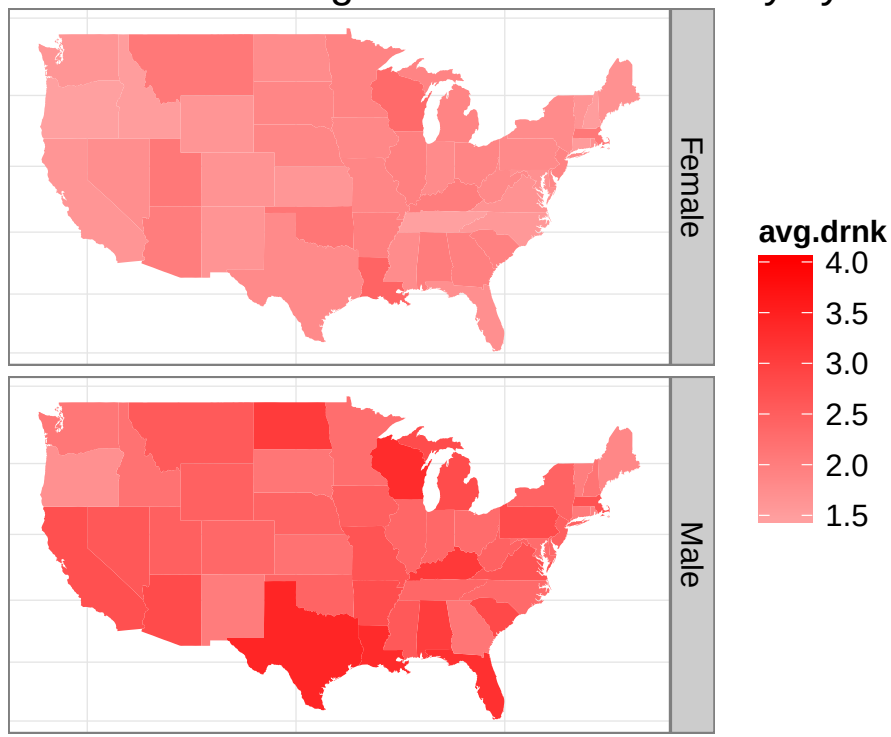


Cleaning Up Your Maps

```
qplot(long, lat, geom="polygon", data=states.sex.map, group=group, fill
  coord_map() + facet_grid(sex ~ .) +
  theme_bw() +
  scale_fill_gradient2(limits=c(1.5, 4),low="lightgray",high="red") +
  theme(axis.ticks = element_blank(),
        axis.text.x = element_blank(),
        axis.title.x=element_blank(),
        axis.text.y = element_blank(),
        axis.title.y=element_blank()) +
  ggtitle("Map of Average Number of Alcoholic Beverages Consumed Per Da
```

Cleaning Up Your Maps

- Average Number of Alcoholic Beverages Consumed Per Day by State and Gender



Your Turn

Use options to polish the look of your map of child healthcare undercoverage rate by state!