







## Course Administration

1. Sign up for consultations!
  - in lieu of class meeting April 13
  - sign up for slots April 7, 9 or 10
2. In-class workshop March 30: handout online (lecture 6)
3. If we have to move online: Blackboard Collaborate, WebEx, or Google Hangouts?
4. Anything else?





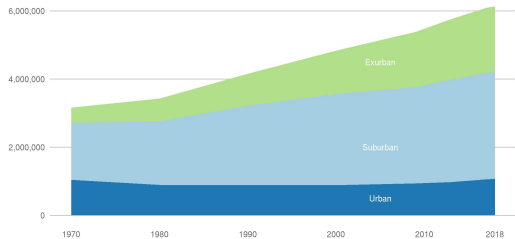
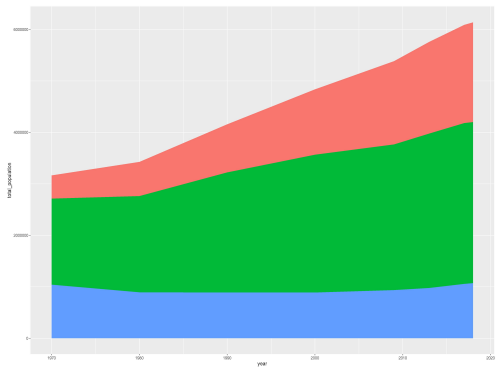








# A Before and After





## Helpful Standardizations

- Make a theme that you can use and modify for all of your charts
- Define a set of colors that you'll use throughout
- Define axis size labels you'll use throughout

For example,

```
axis_text_size <- 15
```

```
usual.colors <- c(‘‘#d4d4d4’’, ‘‘#39b3j2’’)
```

## Helpful Standardizations

- Make a theme that you can use and modify for all of your charts
- Define a set of colors that you'll use throughout
- Define axis size labels you'll use throughout

For example,

```
axis_text_size <- 15
```

```
usual.colors <- c(“#d4d4d4”, “#39b3j2”)
```

and then



## Next Week's Assignment

**Find a choropleth map.** Post link to google sheet by Wednesday noon. Don't wait till next week or you will forget!

Finder	Commenter
Betsy K.	Didem B.
Erik C.	Lydia G.
Josh F.	Neha M.

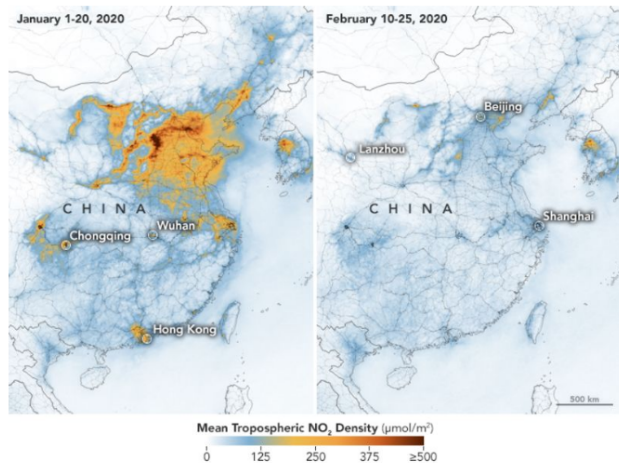








# Lindsay for Emily from *Bloomberg* “Air Pollution Vanishes Across China’s Industrial Heartland”



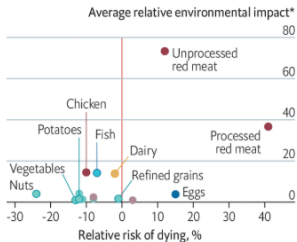
Nitrogen dioxide levels above China during Jan. and Feb., 2020. Source: NASA/ESA

# Neha for Josh from *The Economist*

Going vegan for two-thirds of meals could cut food-related carbon emissions by 60%

## High-steak diets

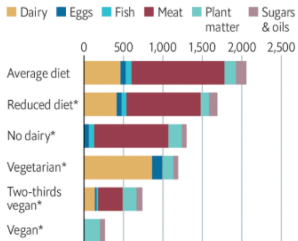
Health and environmental impact of one extra serving per day



Sources: \*Multiple health and environmental impacts of foods', by Clark et al., PNAS; 'Country-specific dietary shifts to mitigate climate and water crises', by Kim et al., Global Environmental Change

The Economist

United States, greenhouse-gas footprint  
kg of CO<sub>2</sub> equivalent per person per year



\*Vegetables=1 †Simulated diet, to reach 2,300 calories per day









# A Pathway to Answers

Start with the point

- What question are you trying to answer?
- What point are you trying to make?
- Which parts of the distribution are important?

And think about the data

- What question can your data answer?
- What level of aggregation is required for your point?



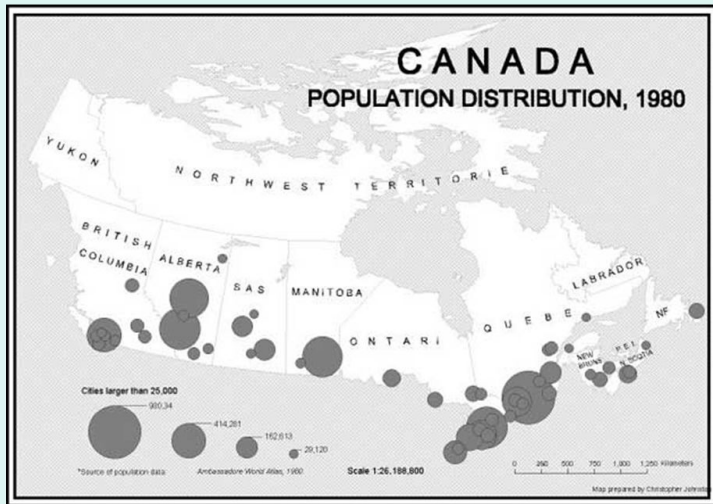
# B.1. Graduated Symbols

- Use symbol of graduated size to convey size or number
- Plot symbol at center of polygon
- Or at point location
- Used to convey absolute magnitudes – examples?

## B.1. Graduated Symbols

- Use symbol of graduated size to convey size or number
- Plot symbol at center of polygon
- Or at point location
- Used to convey absolute magnitudes – examples?
  - area
  - number of people
  - total home value

## Graduated Symbol Example



# Strengths and Weaknesses of Graduated Symbol Maps

What do you think?





# Best Practices for Graduated Symbol Maps

- Use them to convey magnitude
- Make symbols large enough to distinguish
- Be careful of overlap

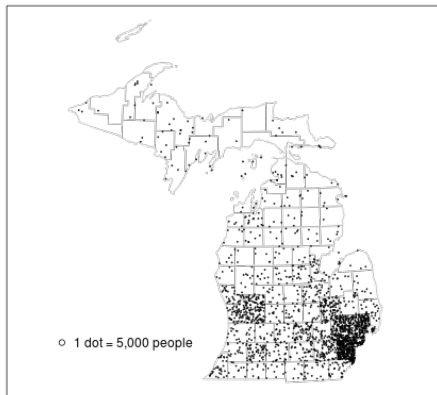
## B.2. Dot Density Maps

- Use dots within administrative unit polygons to represent magnitudes
- Similar to graduated symbol map, but can convey magnitude of more than one group
- Each dot can represent one unit, or can represent multiples, such as 10 people



# Dot Density Example

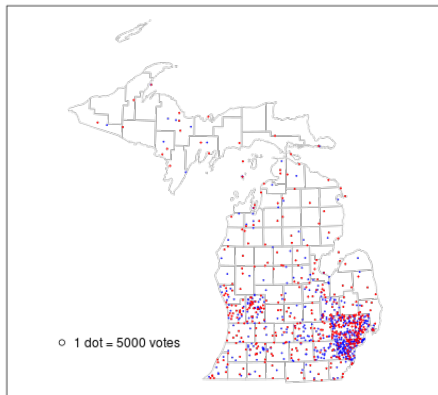
**Michigan 2010 Population  
Dot Density Map**



From [https://msu.edu/~ashton/classes/866/notes/lect20/dot\\_mapping.html](https://msu.edu/~ashton/classes/866/notes/lect20/dot_mapping.html)

## And With Two Variables

Michigan 2016 Election Dot Density Map



From [https://msu.edu/~ashton/classes/866/notes/lect20/dot\\_mapping.html](https://msu.edu/~ashton/classes/866/notes/lect20/dot_mapping.html)

# Strengths and Weaknesses

# Strengths and Weaknesses

- Strengths
  - In my opinion, frequently better at conveying magnitude than graduated symbols
- Weaknesses
  - Conveys a geographic specificity to data that do not exist
  - May generate confusion with specific points

## Dot Density Best Practices

- Use only when geographical granularity of data approximate granularity of depiction
- Use color as in our upcoming discussion of choropleth maps

## B.3. Choropleth Maps

- Used to show relative rates or intensities across space
- Examples?

## B.3. Choropleth Maps

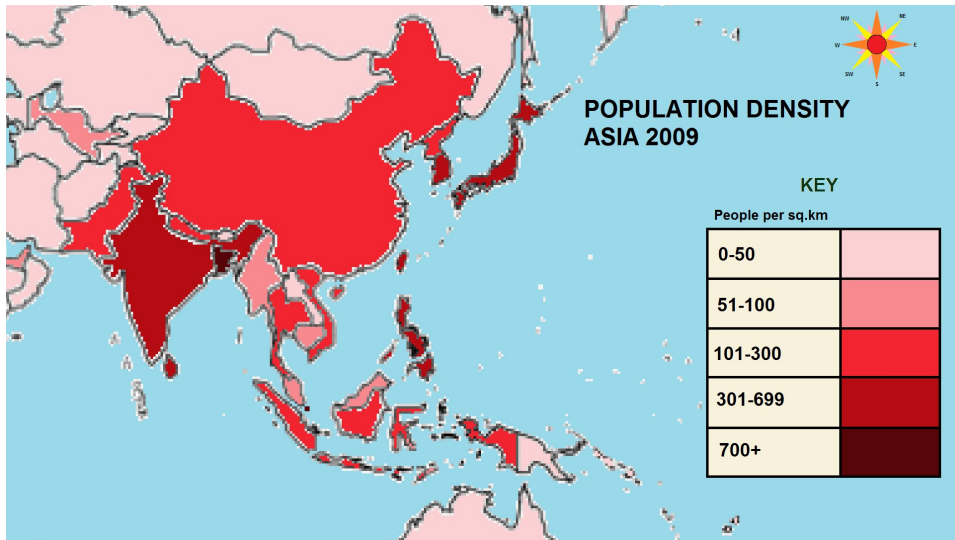
- Used to show relative rates or intensities across space
- Examples?
  - population density
  - share in poverty
  - share covered by health insurance
- these can be continuous: unclassed
- or broken up into categories: classed
- Also used to show categorical differences across space
- Examples?

## B.3. Choropleth Maps

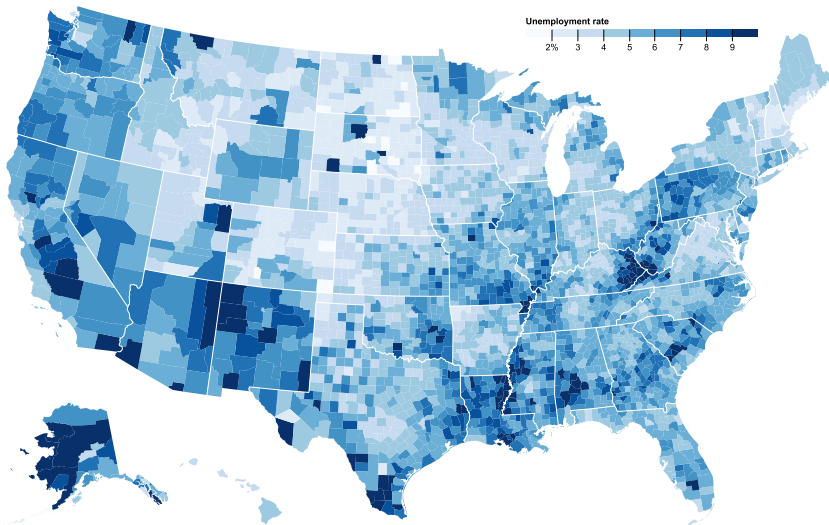
- Used to show relative rates or intensities across space
- Examples?
  - population density
  - share in poverty
  - share covered by health insurance
- these can be continuous: unclassed
- or broken up into categories: classed
- Also used to show categorical differences across space
- Examples?
  - ACA adoption or not
  - type of procurement legislation



# Choropleth with Intensity



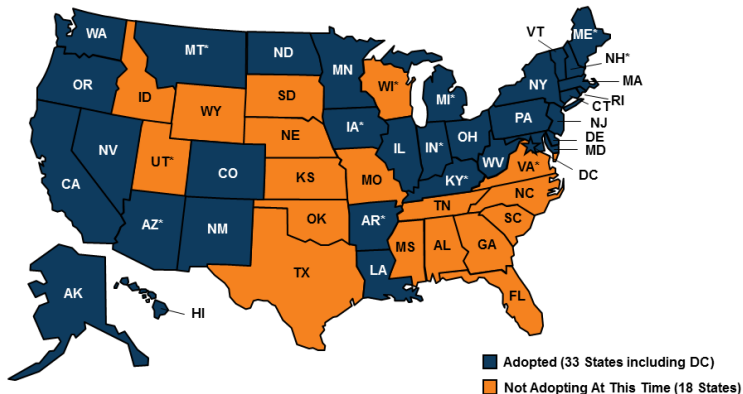
# Choropleth with Intensity





# Categorical Map

## Current Status of State Medicaid Expansion Decisions

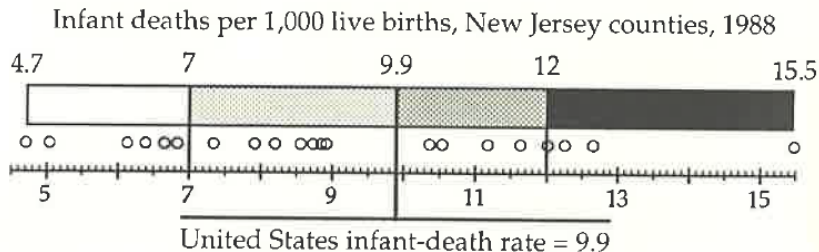


NOTES: Current status for each state is based on KFF tracking and analysis of state activity. \*AR, AZ, IA, IN, KY, MI, MT, and NH have approved Section 1115 expansion waivers. VA is considering adopting expansion in their FY 2019 state budget and UT passed a law directing the state to seek CMS approval to partially expand Medicaid to 100% FPL using the ACA enhanced match; see the link below for more detail. CMS approved the Kentucky HEALTH expansion waiver on January 12, 2018; implementation will begin in April 2018. ME adopted the Medicaid expansion through a ballot initiative in November 2017; the ballot measure requires submission of a state plan amendment (SPA) within 90 days and implementation of expansion within 180 days of the measure's effective date; however, the governor failed to meet the SPA submission deadline (April 3). WI covers adults up to 100% FPL in Medicaid, but did not adopt the ACA expansion.



## Better Yet, the Histogram Legend

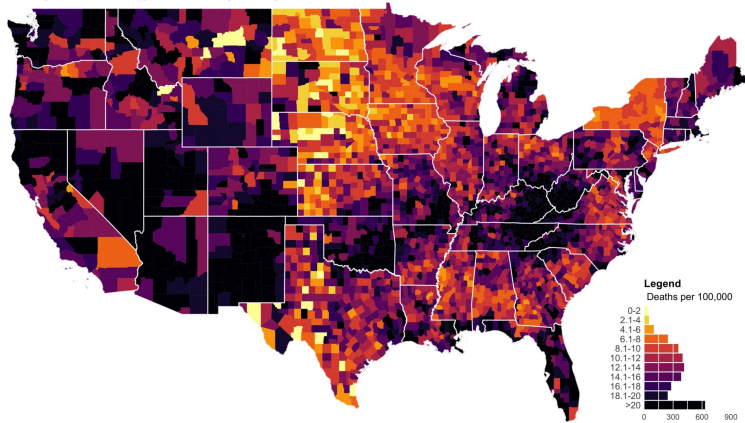
What does this add that the choropleth cannot convey?



Monomnier, Figure 6.10

# Another Histogram Legend

Drug poisoning deaths (2014)



Source: <https://blogs.cdc.gov/hchs-data-visualization/drug-poisoning-mortality/>

## Monmonier on Count vs. Intensity Data

- Monmonier says never use a [what kind of map] for count data
  - Why?



## Monmonier on Count vs. Intensity Data

- Monmonier says never use a [what kind of map] for count data
  - Why?
    - Because size should be the “principle visual variable” for such maps
- M. says use a choropleth for intensity

## Monmonier on Count vs. Intensity Data

- Monmonier says never use a [what kind of map] for count data
  - Why?
    - Because size should be the “principle visual variable” for such maps
- M. says use a choropleth for intensity
- Agree with overall sentiment, but not sure it holds in all cases

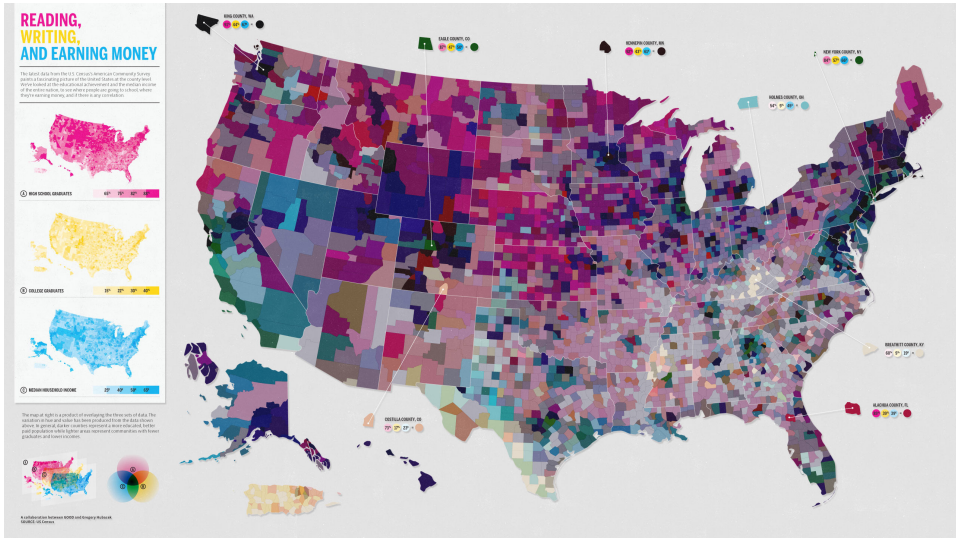
# Worst Practices

- Rainbow colors for classification
- Ones that are frequently bad ideas
  - Map total amounts
  - “geographic features that are continuous in nature”

# Worst Practices

- Rainbow colors for classification
- Ones that are frequently bad ideas
  - Map total amounts
  - “geographic features that are continuous in nature” “... because their distributions are not controlled by political or administrative subdivisions” (DTB, p. 104)

# Beautiful Confusing Map



# E. Best Practices

1. Categories
2. Colors
3. Histogram legend

## E.1. Categories

- 4 is great
- Don't use more than 5 or 6
- Use an intensity ramp only when you care very little about the exact values
- And comparison between values

## E.2 Colors

- Make the most intense color the largest value
- Avoid pattern fills if at all possible
- Make your legend a dot plot or histogram with the same colors
- Put anything else on map in a light color



# Use ColorBrewer

- Named after Cynthia Brewer
- <http://colorbrewer2.org/#type=sequential&scheme=BuGn&n=3>
- You say
  - number of classes
  - sequential or divergent or qualitative
  - multi-or single hue
  - your preferred color
  - color-blind friendly?
  - screen or printer?
  - and more...
- and it gives you a color scheme!

## F. Think about Goats







# Today in R Programming

- A. Choropleths
- B. Histogram legends
- C. Dot Density

## A. Choropleth Maps

```
p1 <- ggplot() +  
  geom_sf(data = polys,  
          mapping = aes(fill = fill.in.variable))
```

## A. Choropleth Maps

```
p1 <- ggplot() +  
  geom_sf(data = polys,  
          mapping = aes(fill = fill.in.variable))
```

Assigning colors with a vector

```
vectoro <- c("c1", "c2", "c3", "c4")  
scale_fill_manual(values = vectoro)
```



## A. Choropleth Maps

```
p1 <- ggplot() +  
  geom_sf(data = polys,  
          mapping = aes(fill = fill.in.variable))
```

Assigning colors with a vector

```
vectoro <- c("c1", "c2", "c3", "c4")  
scale_fill_manual(values = vectoro)
```

Use the same vector for your map and your histogram legend, below.

## B. Histogram Legends

- ▶ Can use a dot plot, as in Monmonier
  - ▶ `geom_dotplot()`
- ▶ Can use `geom_histogram()`
  - ▶ color bars with `fill`
- ▶ Can use `geom_bar()`

## B. Histogram Legends

- ▶ Can use a dot plot, as in Monmonier
  - ▶ `geom_dotplot()`
- ▶ Can use `geom_histogram()`
  - ▶ color bars with `fill`
- ▶ Can use `geom_bar()`
- ▶ **Colors must link with the plot – or the whole thing fails**

## Warning on the Histogram/Bar/Dot Legend

- ▶ need to trick R into making a bar chart
- ▶ assume we have four quartiles like this

```
ex.df[,c("quartile", "qmin", "qmax", "cato")]
```

```
##  quartile qmin qmax cato
## 1         1    0    1    1
## 2         2    1    2    1
## 3         3    2    4    1
## 4         4    4   23    1
```

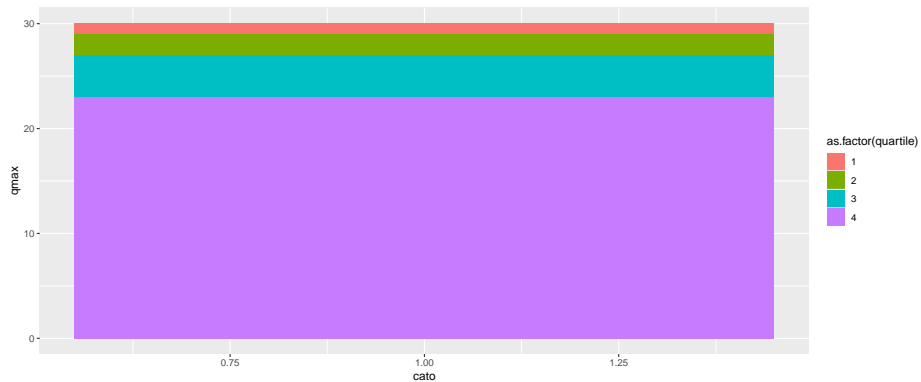
- ▶ plotting quartile vs max in a bar chart does not give what we want

## The Bad Graph

```
p3 <- ggplot() +  
  geom_col(data = ex.df,  
           mapping = aes(x = cato, y = qmax, fill = as.factor(quarter))
```

# The Bad Graph

p3



## We Need to Trick R

- ▶ add an “increment” variable
- ▶ and make the categories backwards

```
ex.df
```

```
##   quartile cato qmin qmax oppo.qs qrtl.incre
## 1         1     1    0    1        4          1
## 2         2     1    1    2        3          1
## 3         3     1    2    4        2          2
## 4         4     1    4   23        1         19
```

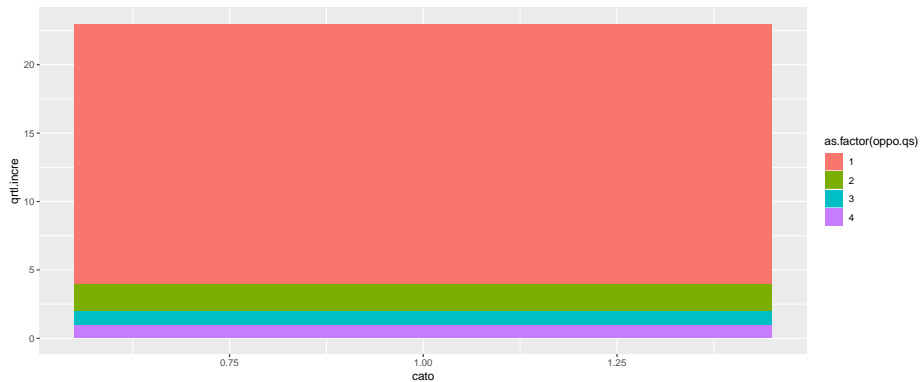
## The Better Graph

```
p4 <- ggplot() +  
  geom_col(data = ex.df,  
           mapping = aes(x = cato, y = qrtl.incre,  
                         fill = as.factor(oppo.qs)))
```



# The Better Graph

p4



## C. Dot Density Maps

- ▶ Make a simple feature with points: `st_sample()`
- ▶ Get those points into a matrix: `st_coordinates()`
- ▶ Make the matrix a dataframe: `as.data.frame()`
- ▶ Plot the points
- ▶ Make sure you plot your points on top of a map or they will look funny

## Dot Density Map code

```
h.map <-  
  ggplot() +  
    # this is the map you did before  
    geom_sf(data = bg2010,  
            fill = "transparent",  
            color = "white") +  
    # here are the points  
    geom_point(data = hdat$df,  
              mapping = aes(x=lon, y = lat),  
              shape = ".")
```

## Next Lecture

- Next week: spring break!
- Next next week: Line charts
- Read Few, parts of Chapter 10, pages 217-200, Chapter 13; Chang, Chapter 4