

Basic Info

Title:

Can I Afford To Teach? Public School Teacher Salaries by State/County as compared to the Cost of Living Index.

Group Members:

Estee Rebibo – erebibo@wustl.edu – 431757

Dan Finn – dmfinn@wustl.edu – 434967

Fiona Hon – fhon@wustl.edu – 434944

Charlotte Maguire – cmaguire@wustl.edu – 435067

Project Repository:

<https://github.com/finnd/557finalproject>

Background Information and Motivation

Public school teacher salary and compensation has been a widely discussed topic since the dawn of the Internet age. However, recently, we have seen a growing number of strikes, protests, and civic involvement from public teachers who are unsatisfied with their compensation and, despite providing a necessary and critical public service, often are living at or below the family poverty line.

Recent teacher strikes have occurred in West Virginia, Oklahoma, Kentucky, and

Ohio. More information about the specifics of these strikes can be found here:

goo.gl/ei3sW4.

Thus, the motivation of our project is the following: **Create a tool which current or prospective teachers can visit and use said tool to make an informed decision about whether being a public school teacher in a given area is financially viable for their lifestyle.**

Project Objectives

Questions:

Where in the United States has the highest average or median public school teacher salary?

Where in the United States has the lowest average or median public school teacher salary?

Where in the United States, when cross-referenced with Cost of Living information, is the most affordable for a public school teacher to live and work?

Where in the United States, when cross-referenced with Cost of Living information, is the least affordable for a public school teacher to live and work?

I am a current or prospective teacher working or trying to work in XXXXX, is this a fiscally responsible decision given my lifestyle and demographics?

Benefits:

- 1) A useful, effective tool that can be used by real world teachers to analyze their compensation and compare it to other areas and states.
- 2) An increased understanding of the fiscal landscape that public school teachers face, and in which teachers areas are in need of the most assistance.
- 3) Insight into which states most effectively base teacher salary on COI indices and create a fiscally welcoming environment for teachers.

Data and Data Processing

Data:

We will gather our data from a variety of sources. State by state teacher salary data is available from a number of sources. The most in-depth source of state by state data we found was provided by the National Education Association, and is available here: <https://goo.gl/Rv2Wbh>

County by county data is more difficult to obtain but is present online. Most, if not all, states publish their fiscal information online every year. We have found county by county salary data for many states, and are continuing to find more. We will not link all 35+ datasets we have found so far here, but examples are below:

Virginia: <https://goo.gl/2ssJa9>

Missouri: <https://goo.gl/vNh2VR>

For the county data, we will acquire it from a mix of manual acquisition, webscraping with beautifulsoup, and available datasets.

Cost of living index data is available in many places. The Missouri Economic Research and Information Center provides a nice dataset here:

<https://goo.gl/VfDDBL>

Data Processing:

We do not expect substantial data cleaning to be necessary. The state by state is already cleaned and easily accessible in a programmatic manner. The county data will be slightly more difficult to conglomerate given that it is coming from a variety of sources, but the output from each source regardless of acquisition method (manual, webscaping, or dataset) will be average salary, county, and zipcode. We will generate a csv file containing COI information from the link provided above.

Data processing, cleaning, and agglomeration will be done using **Python**. CSV files will be manipulated using **PANDAS**. Webscraping will be done using **BeautifulSoup**. Any other required data manipulation will be done using standard python libraries unless a more efficient solution is found.

Visualization Design

We decided that it would be easiest to display our data using an outline of the U.S. map in order to display the overall country's data and correlation at a glance.

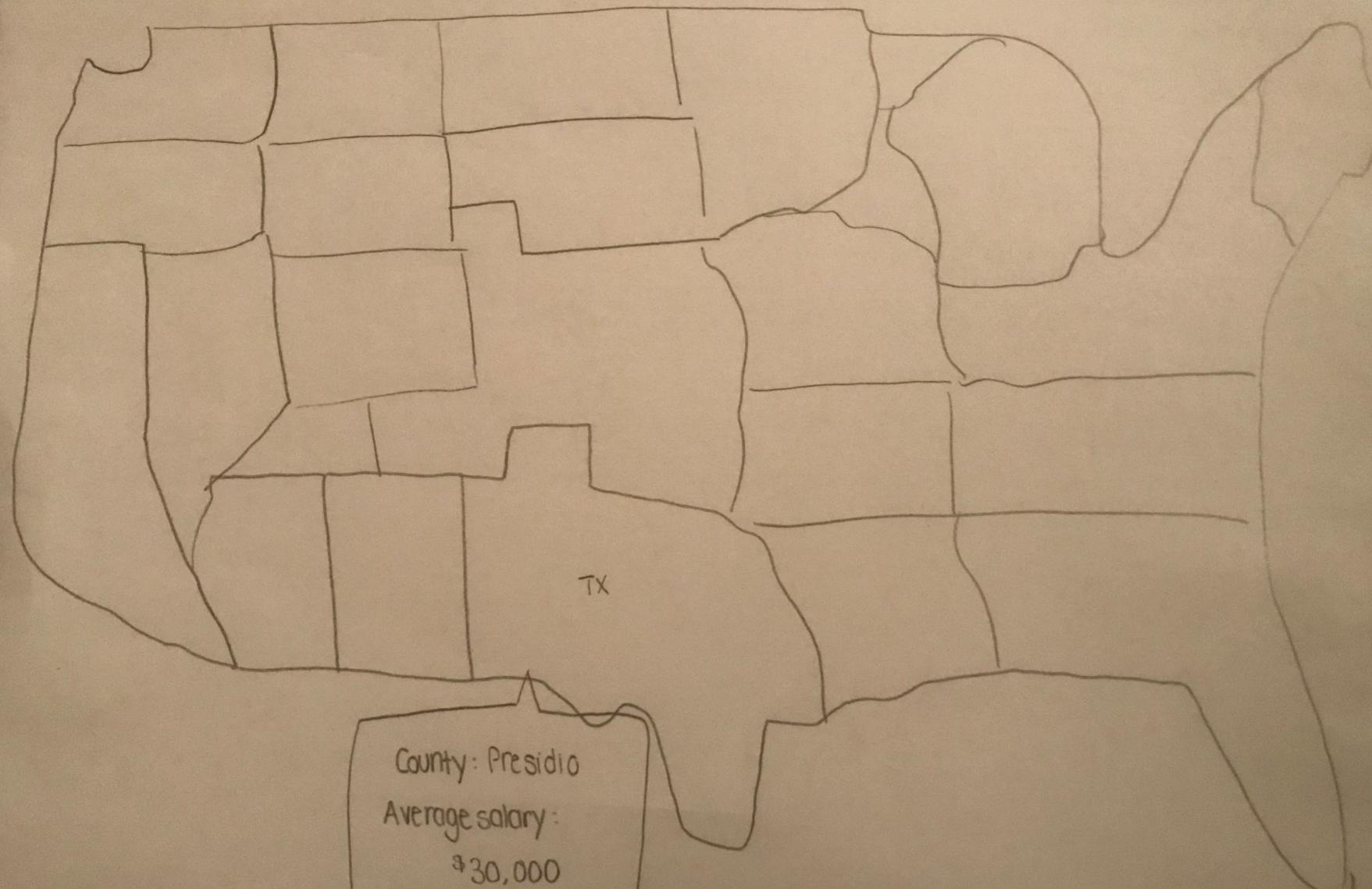
Our first idea was to create a zoom-able map of the U.S. and implement tooltip hovering over each state or county that would display the average salary and the cost of living.

Our second idea was to divide the web page into three sections. The left section would incorporate a map of the U.S. on the left, and when a user selects a state, it would highlight, and a zoomed in version would appear in the section on the top right with all of the counties. When a user selects a county, the bottom right section would display the information over the past five years in a line graph.

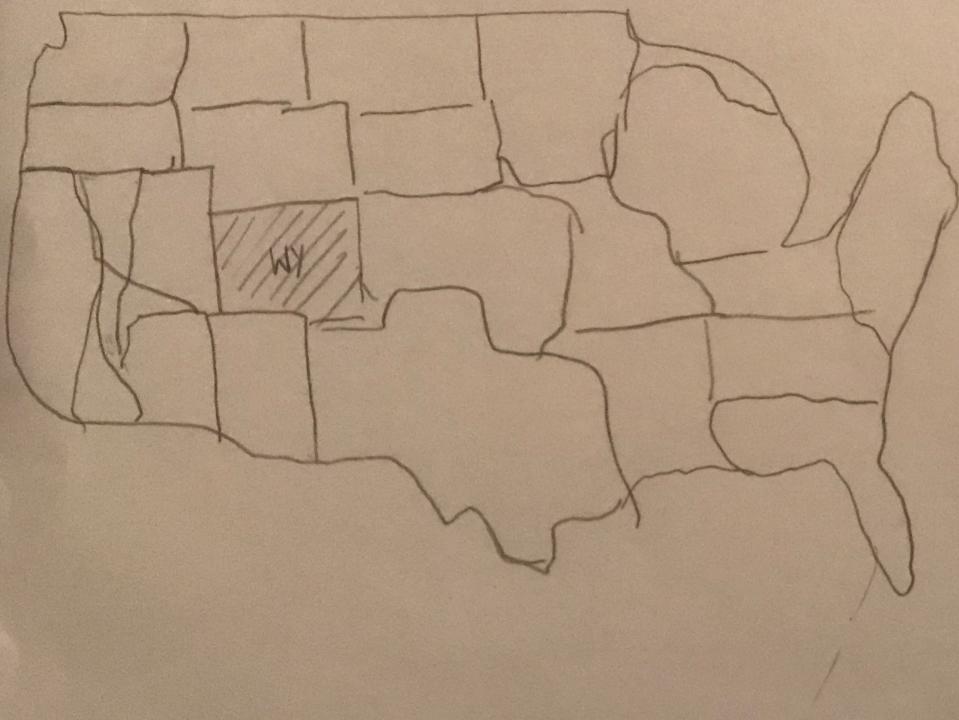
Our third idea was to create a map that would outline all each of the U.S. counties and generate a heat map; this would also implement tooltip hovering displaying the average salary and cost of living of each county.

Our final idea was a combination of all three: it would include those three sections and have a heat map for each state and each county, but instead of a line graph in the bottom right, it would simply list the necessary information.

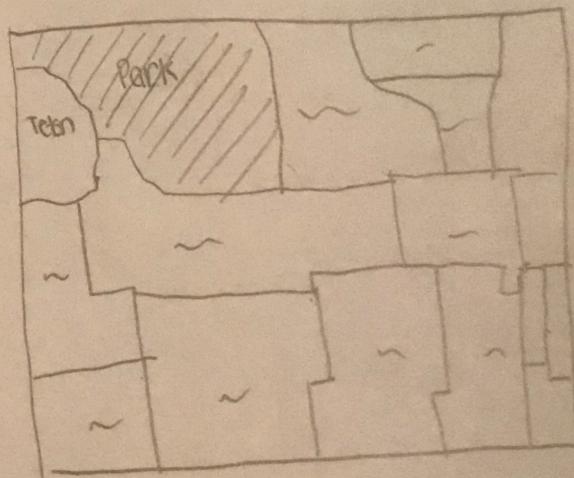
Our sketches follow on the next 4 pages.



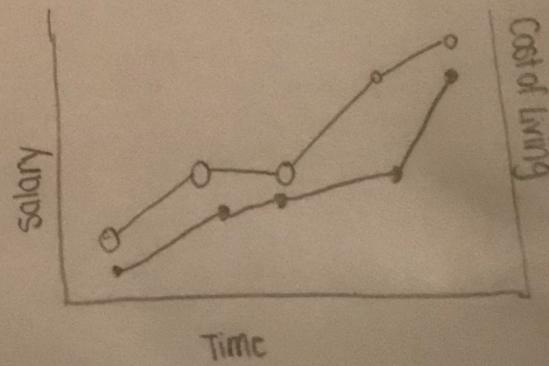
* zoomable map and have a tooltip
hover displaying info



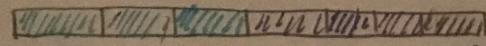
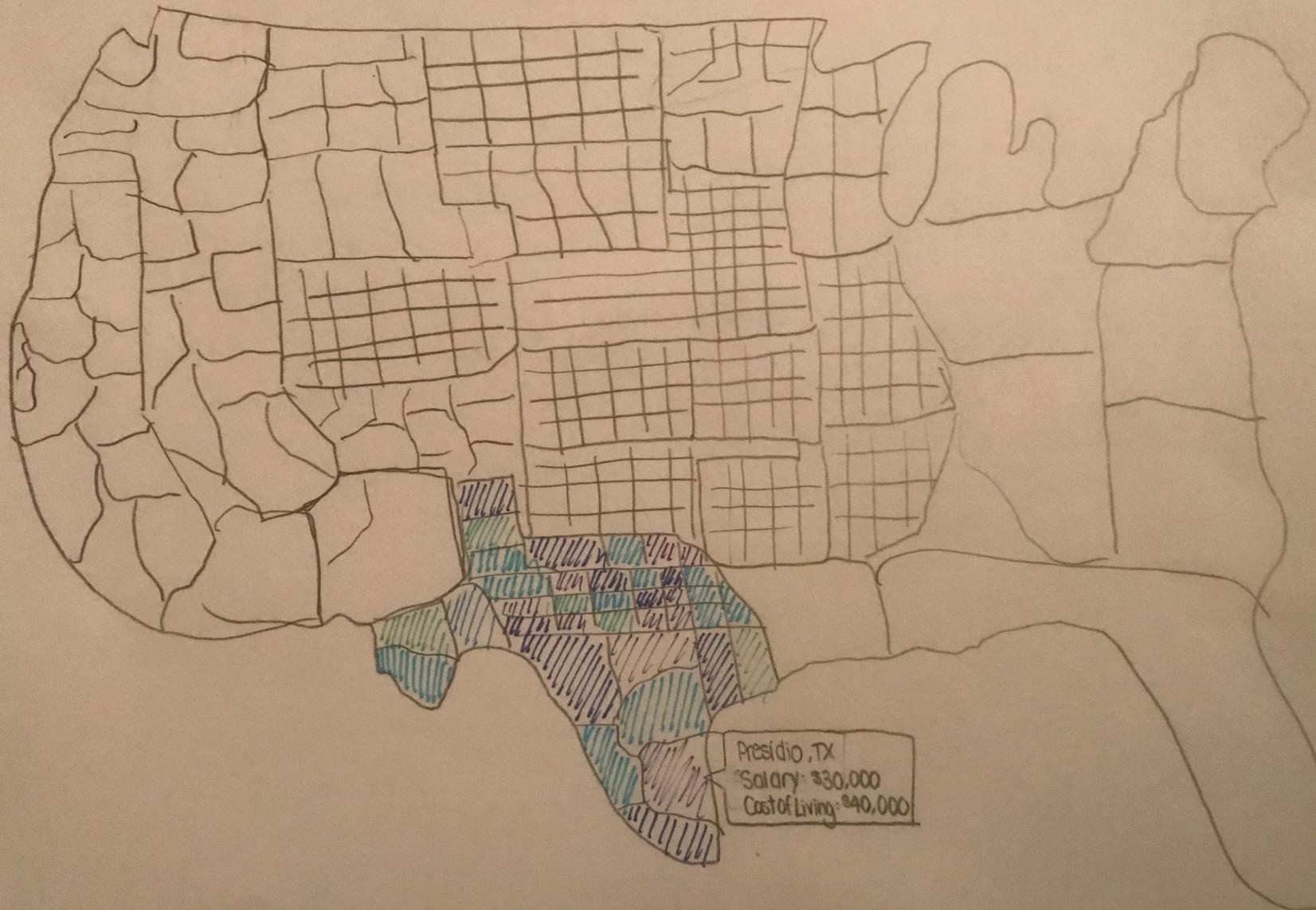
WYOMING



County: Park



* Would show all counties on
the entire map *

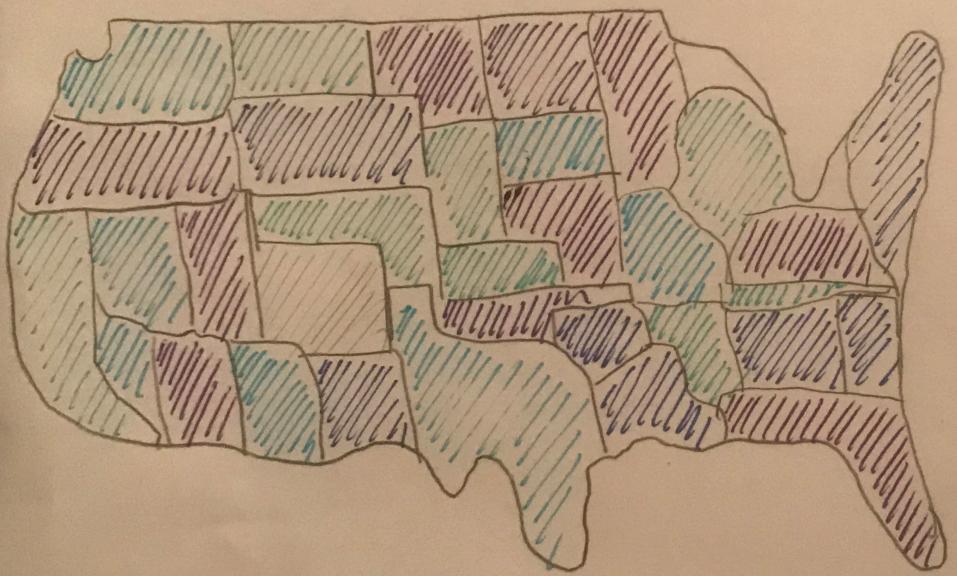


20,000

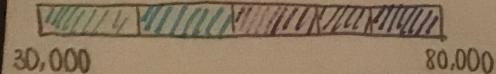
80,000

* the rest of the map would be
colored as a heat map like Texas *

* tooltip hovering
would be implemented *

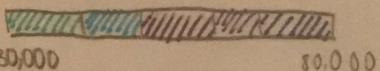
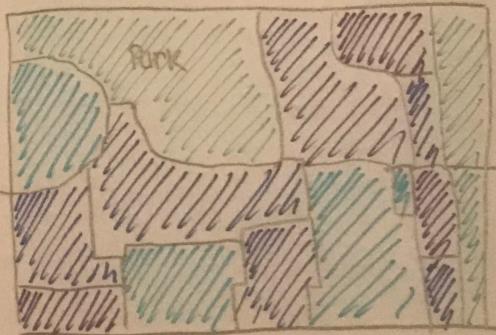


WYOMING



* user selects a state; will highlight
* includes an overall heat map
by state *

* user selects a state, and a
zoomed in image w/ counties
appears*
counties will be heat mapped



Park, WY

Average Salary in 2017: \$50,000

Cost of Living in 2017: \$49,000

Salary : Cost of Living in Park : $50/49$

Salary : Cost of Living in Wyoming : $50/47$

* user selects a
county; will highlight

Must Have Features

- 1) Ability to view and compare state by state average public school teacher salaries for the past 3 years at minimum.
- 2) Ability to view and compare state by state cost of living indices for the past 3 years at minimum.
- 3) Ability to view and compare county by county average public school teacher salaries*
- 4) Ability to view and compare county by county cost of living indices*.
- 5) Intuitive and easy to use visualization that incorporates at least 1 map projection.

* Some data may not be available for all counties.

Optional Features

- 1) Ability to view private school teacher salaries by state and county
- 2) Additional historical data (perhaps the past 5 years)
- 3) Ability to view information regarding public school funding in a given state.

Project Schedule

Week 1: Monday, April 2-Sunday, Apr 8

- Project proposal due Tuesday, April 3 @midnight
 - Identify data and project goals.
 - Create a visualization idea that works well with the data we want to represent.
 - Make sketches for project proposal.
- Gather data
 - Write scripts to parse and pull data from public state sites.
 - Download any available csv/tsv files with relevant datasets.
 - Manually acquire data that is not available programmatically

Week 2: Monday, April 9-Sunday, Apr 15

- Scrape and clean data.
- Create a sample visualization for one State with the counties.
- Ask questions to classmates on Thurs on any questions we've run into based on cleaned data and sample visualization.

Week 3: Monday, April 16 –Sunday, Apr 22

- Using sample visualization as model, use the cleaned data to make the visualization for all states/counties.
- Ask questions to classmates on Thurs on any questions we've run into based on cleaned data and sample visualization

Week 4: Monday, April 23-Sunday, Apr 29

- Incorporate interactivity to the visualization, using one state as a model.
 - Create a tooltip that gives relevant information for each county in a given state.
 - When user selects state, create zoomed in heat map of one state and its counties in separate part of screen with more detailed information.
- Ask questions to classmates on Thurs on any questions we've run into based on cleaned data and sample visualization.

Week 5: Monday, April 30-Sunday, May 6

- Incorporate interactivity to the entire visualization, using the sample state from week 4 as a model.
 - Create a tooltip that gives relevant information for each county in a given state.
 - When user selects a state, create zoomed in heat map of that state and its counties in separate part of screen with more detailed information.
- Ask questions to classmates on Thurs on any questions we've run into based on cleaned data and sample visualization.

Week 6: Monday, May 7-Wednesday, May 9

- Refine visualization and debug any issues that have come up.
- Make sure entire process book is good to go, and that visualization is posted to website.
- Review and practice presentation.
- Present visualization on May 9.