American Electoral Politics:

The Past 100 Years

RYAN KERR, EVAN SANDHOEFNER, MENGTING ZHANG
CS171 SPRING 2015
TF: KEVIN SUN

TABLE OF CONTENTS

INTRODUCTION

Overview and Motivation

For our project, we initially brainstormed several areas of interest, then attempted to find data for those target areas. Politics, especially those involved in the United States presidential election, was one of those areas where our interest intersected with public data that was publicly accessible. The history of the presidential election is not a familiar topic for some, and with this project, we hoped to educate the general public about how the process and the political parties involved have changed over time.

Related Work

We began brainstorming the overall structure of our data visualization by looking through some of the previous projects listed in the Hall of Fame tab of the course website. We were inspired by the project *Century of Corn* and were really interested in creating a choropleth map, especially since none of the course problem sets this year had tackled the implementation of maps. This was relevant to our interest in U.S. electoral votes for presidential elections, and so we decided to visualize our data geographically. In our search for data, we found many maps that visualized election results, but found that most looked like this:

This visualization shows a static image for the results of the Presidential election. It links to results for the Senate seats and elections, but each view is itself static.

Our visualization aims to be more informative by turning that static visualization into one that is selectively interactive. We also aim to encode margin of victory for each state, meaning that states that voted very strongly in one direction would have more intense color. We also know that there are a multitude of ways to analyze election results, whether through a state-by-state comparison, a closer look at how certain demographics voted, or general political party trends through the years. We also want to visualize that data to reveal trends through interactive views below the main visualization.

Questions

The primary goal of our visualization is to display voting trends in the United States, as well as the changes in political party support through the history of the country. In addition to currently existing parties, the data we visualize will include parties that have faded out of existence. A benefit of this visualization is that we will also be able to visualize the history of the United States and view the growth of not only the states as they exist currently, but also how parties and the election process have evolved over time. Through the creation of this visualization, we wanted to also be able to learn about how certain demographic factors relate to both voting and voter turnout. Not only will we be able to view how each state changes through elections, but also how states differ with each other.

Data

We have combined data from a number of sources to obtain a comprehensive understanding of changes in American electoral politics over the centuries, and of differences between states for a given election year. Our data files will be a comprehensive list of every presidential election's voting breakdown by political party. We also have collected data for every state that includes various demographic statistics. A list of sources and the specific data we located on each site will be detailed further in this process book.

Tasks

Our visualization will require a map to visualize our data, with the intensity of the color of the state representing the winning party and by how much they won. We will implement a slider that will allow the user to select a particular election year and update the views accordingly. This feature will also include political parties that had previously been represented in elections. The user will also be able to click on a state to show more detailed statistics for that state during that particular election, which will be represented visually below the United States map.

We would also like to implement a timeline that will allow the user to enter a story mode, which will display a certain number of significant events in the history of the United States and select the corresponding elections. This will ideally communicate how voting was potentially influenced in that period of time. Our visualization will also include greater interactivity through the ability to filter results, like showing only one party (color) on the map.

Users

The target audience of our visualization is the general population of the United States, especially those interested in political trends. We want to educate those who are most directly affected by the structure of our government by providing them with a visualization of a comprehensive history of the presidential elections.

Data Collection (Source, scraping method, cleanup, etc)

http://www.presidency.ucsb.edu/showelection.php?year=2012
> The above link has data for each presidential election from 1789 to 2012. The data include, for each state: total votes and votes for each major candidate.

https://docs.google.com/spreadsheets/d/1bH38j6_e8yA9xq8OMlyLOL6h_iTS7ABQMKNxzFgKBDo/edit#gid=435419492
> The above link has national turnout rate for 1789-2012.

https://docs.google.com/spreadsheets/d/1or-N33CpOZYQ1UfZo0h8yGPSyz0Db-xjmZOXg3VJi-Q/edit#gid=1670431880
> The above link has the following variables for 1980-2014, by state: voting-eligible population, votes counted, voting-age population, % non-citizen, and population of prison/probation/parole/ineligible felon. Unfortunately, there are some missing values in the spreadsheet. We will attempt to fill in those values from another data set, or find a better data set and not use this one at all.

http://www.presidency.ucsb.edu/data.php
> The above link contains a ton of other data that we'll think of as best-case features for now. This includes not only more detailed election data, but also data on the presidencies themselves, which could be an entirely new component of the overall visualization. This includes things like number of vetoes, number of executive orders, and approval ratings. Approval ratings in particular have a direct link to election data - we may be curious to see, for example, if there are presidents who were elected by a landslide but who ended up with low approval ratings. This whole section is a "maybe" feature, because it would be an entirely separate and non-essential wing of the overall visualization.

https://cse.google.com/cse/publicurl?cx=002720237717066476899:v2wv26idk7m
> The above link leads to Google's Datasets Search Engine, which we have been using to find a lot of this data. We will continue to use it, in addition to general Google and sites that specifically compile political/electoral/demographic data from various sources, to continue bolstering the above data and introducing the below data.

http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

> The above link yields race and ethnicity data by state and year from American FactFinder, powered by the U.S. Census Bureau.

https://www.census.gov/popest/data/state/asrh/1980s/tables/estage80.txt

> The above link yields age data by state and year from census.gov.

http://en.wikipedia.org/wiki/List_of_U.S._states_by_GDP

> The above link yields gross state product data by year from Wikipedia.

http://www.census.gov/hhes/www/income/data/statemedian/

> The above link yields median household income from 1984-2013 from census.gov.

The following is a run-down of what each file in our repo's Data folder contains.

- us_states.json and us_states_small.json: not created by us; downloaded for the purpose of generating our central map
- forDataCleaning: a working file (nothing to see here) for parsing and cleaning and formatting tables from the Web; basically holds scraps of various data sets
- masterSpread: a single (cleaned and parsed and formatted) spreadsheet compiling many data sources
- jsonMain: simply masterSpread converted to .json; the only data file we created to load into D3
- medInc, medInc_2013_1984, 2013_inc: files that hold data for median household income from the years 1984 to 2013; the main one used is medInc.json
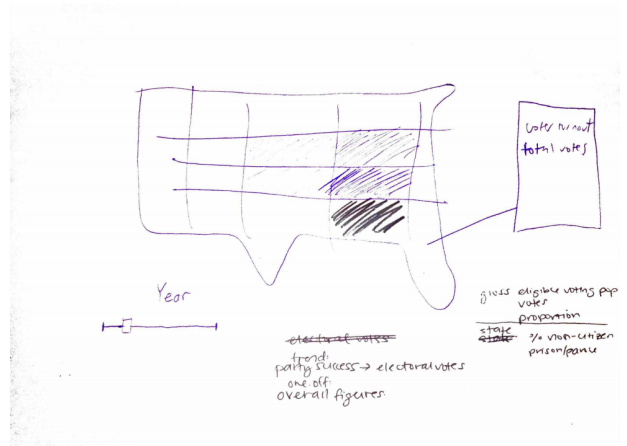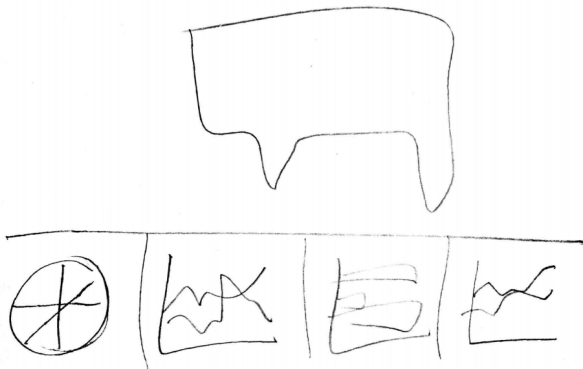
Cleaning, parsing, and formatting the data has been reasonably straightforward. Most of our data come from tables. Some are HTML, others tab-delimited, others comma-delimited, others fixed-width .txt. The big-picture process is: copy and paste table into Excel (perhaps using forDataCleaning as an intermediary, or going straight to masterSpread). Cut out unwanted data. Layout table and format values to match desired layout and formatting in masterSpread. Formatting values involves removing commas from numbers over 1,000, removing % signs from percent values, removing asterisks from a few cells (flagged for notes by our sources), etc. Finally, convert csv files (masterSpread) to json files (jsonMain) using:

http://www.convertcsv.com/csv-to-json.htm
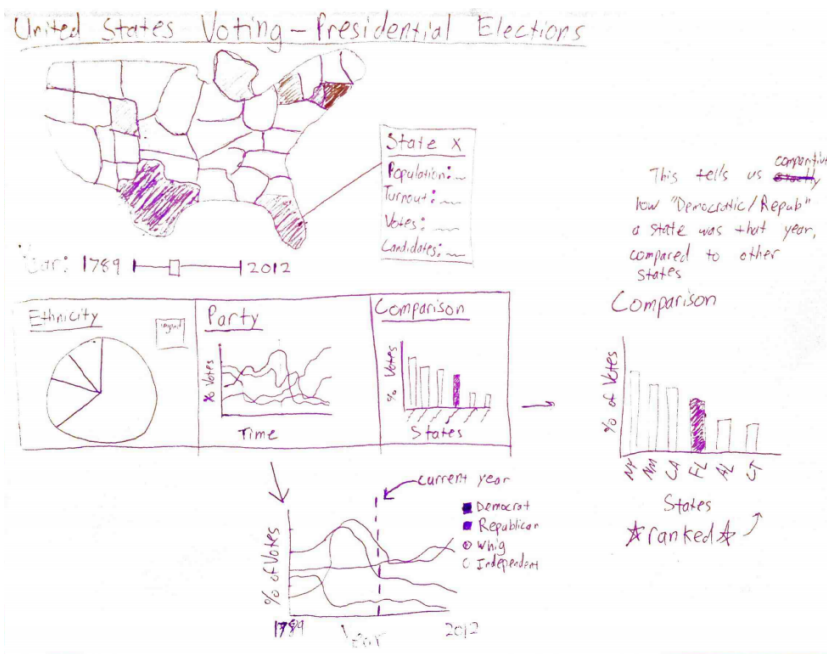
PROCESS & DESIGN EVOLUTION

First Concepts

In all of our prototypes, the main visualization is the United States. We focused on developing the main map to include interactivity through hover-over pop-ups that would give more information about the particular state. The main map would show the national level of politics, while the smaller views at the bottom would show the more local statistics at the state level.
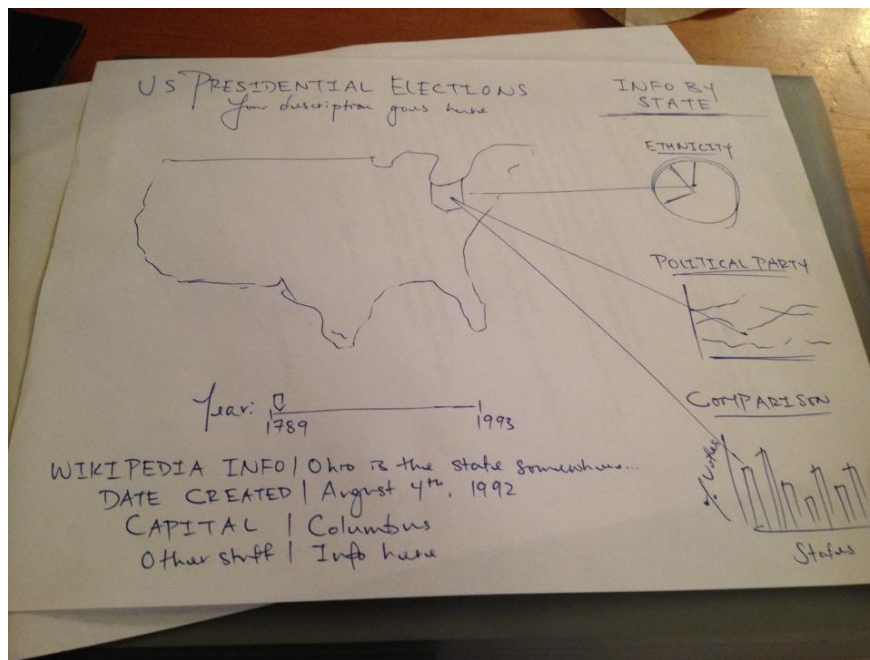

Design Studio
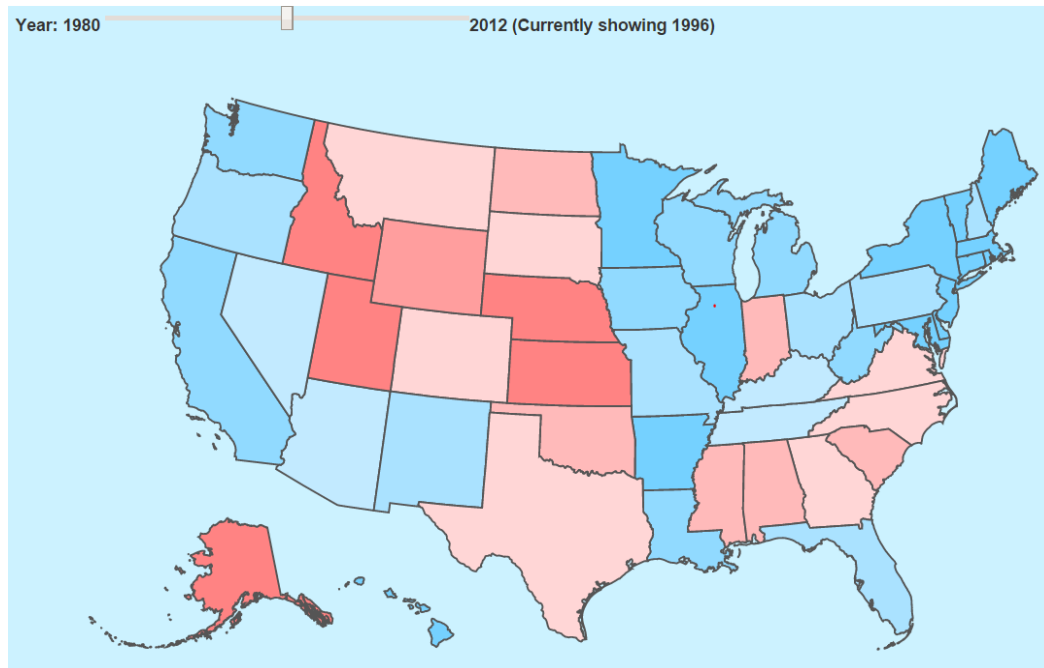
partner group: Angela Jiang, Jason Shen, Kit Wu

At our design studio, we presented the above prototype. From our first sketches, we turned our vague concepts of the smaller views into more developed views with specific datasets. Feedback that we received included concerns about including political parties that once existed, as well as the change in US states over time. These concerns we plan on addressing by 1) including different colours for past parties (i.e. Whigs will be encoded with shades of green) and 2) to simplify the task at hand, we will keep the current map of the 50 United States, but keep states greyed out until they come into existence. We also received the suggestion to include exogenous information per state (such as gross state product and median age), possibly from Wikipedia. Below is a new alternative mockup for the overall view, produced during the studio.

Milestone 1 Design

Map

Our main visualization, as it existed in its current, preliminary form.

The map shows the election results for each state for each election. The states are colored red or blue if they voted republican or democrat, respectively. There are also bins for the intensity of the color, where lighter reds/blues represent a narrower margin of victory. The slider at the top of the map can move from 1980 to 2012, and as it does, updates the graph to the relevant election year data.

We showed this preliminary visualization to a group of non-visualization students and a few questions were brought up.

1) How do we know what year the current map is showing?
To address this we added the "Currently showing YEAR" feature next to the slider.
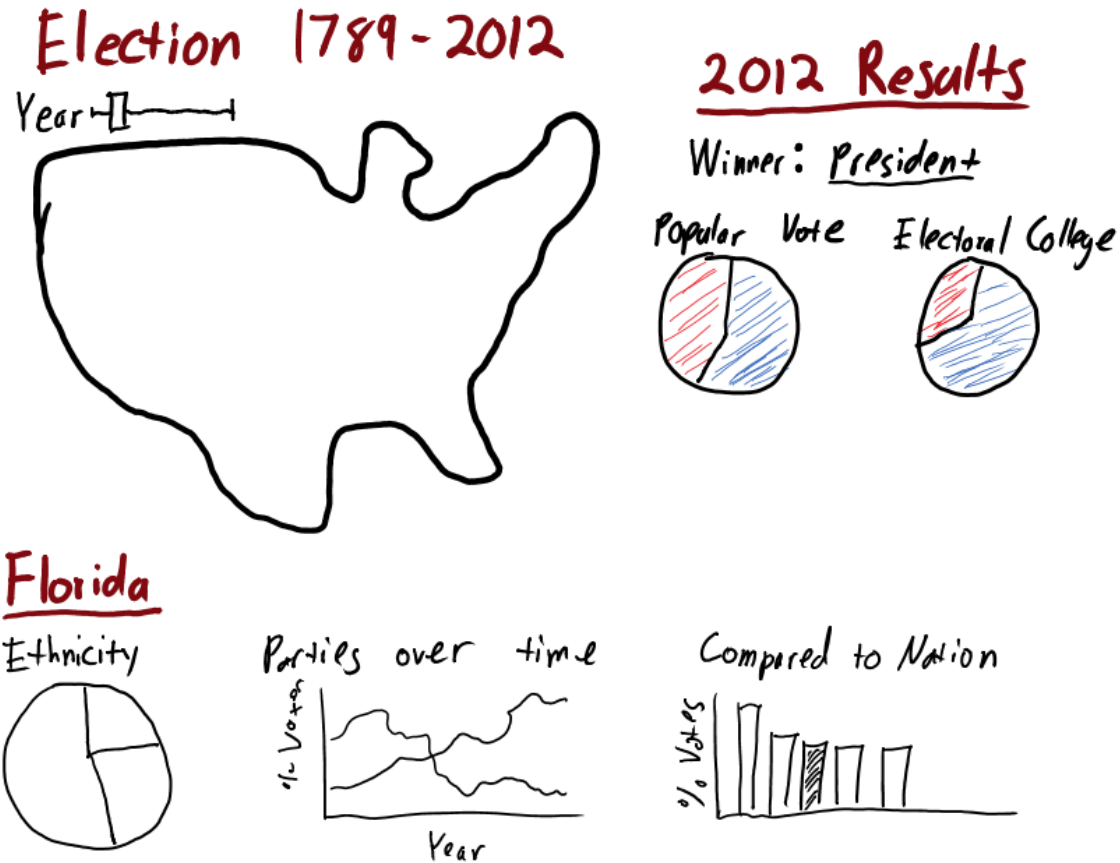
2) What does pink represent?
There was some confusion regarding what the lighter colors meant. We now plan to add some sort of legend or key to let users know what each color represents.

3) How do we know who won the election?
We realized that one of the most integral pieces of information is who won the election, so we revised our design to have the right panel to be overall election data (detailed in additional views)

Additional Views

Below was the most updated sketch of our additional views:
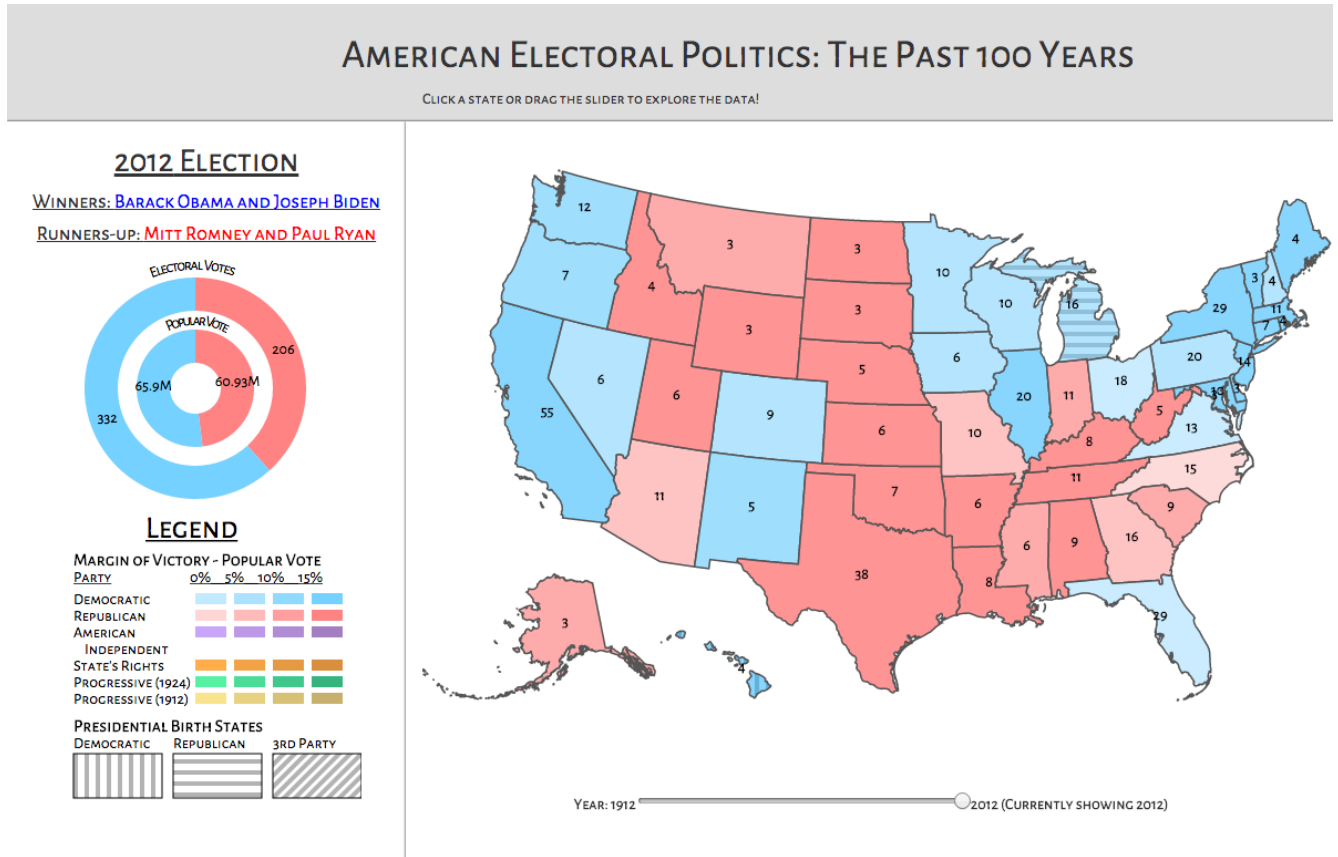

On the right panel we will have a view that states the president who won and the president he was running against. We will display the voting data for that election by both raw votes and electoral votes.

On the bottom we will have our state-specific visualizations. These will appear when the user clicks on a state. These visualizations will then be updated each time the slider moves, so you can see state-specific trends over time. The views will be demographic breakdown pie chart, party votership over time line graph, and a comparison to the nation for that particular election in bar graph form.
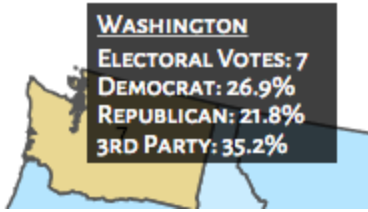
FINAL IMPLEMENTATION

Main Visualization

Our main visualization is pictured above. The center is the choropleth of the United States, which colours each state with its margin of victory for the political party that won its electoral votes. Below the graph of the United States, there is a slider that enables the user view how the United States voted starting in 1912 all the way to the most recent election in 2012.

To the left of the choropleth, we have a legend that details what margin of victory corresponds with each shade of colour. Above this legend is another key feature, the two donut charts. The outer ring visually depicts the amount of electoral votes that each candidate won, with the actual numbers laid over the colours. The inner ring does the same, but for the popular vote. These metrics are useful to display in the main feature so that the user can gain a better sense of how the votes in the states, displayed in the main choropleth, tally up to result in the winner. As the slider changes, so do the donut charts and the label of the election winner and runners-up.
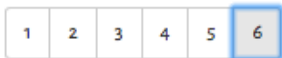
Furthermore, the user can mouse-over the choropleth to display further information about each state. Inside the box, the user is told what state they are learning more about. They are also told the number of electoral votes assigned to each state (helpful for smaller states), as well as the percentage of voters who voted for each major party in that state.

Each state can also be clicked, which affects the views below. We will detail these features in the next section about additional features.

Additional Features

Story Timeline

IN 2012, OBAMA'S SUCCESSES IN OBTAINING A MAJOR PARTY'S NOMINATION AND WINNING THE GENERAL ELECTION WERE BOTH FIRSTS FOR AN AFRICAN AMERICAN. WITH 365 ELECTORAL VOTES TO MCCAIN'S 173, OBAMA RECEIVED THE LARGEST PERCENTAGE OF THE POPULAR VOTE FOR A DEMOCRAT SINCE LYNDON JOHNSON IN 1964.

THIS WAS ALSO THE FIRST ELECTION IN WHICH NEITHER CANDIDATE WAS BORN IN THE CONTIGUOUS U.S. OBAMA WAS BORN IN HAWAII AND MCCAIN WAS BORN IN A MILITARY BASE IN PANAMA.

To the right of the choropleth is another feature of our visualization: the story feature. On page load, we see several options, along with text telling the user to pick one of the above options. Upon choosing one, we get this:

This fun feature gives a short description of the given year's election, as well as the headshot of the winning presidential candidate. Our visualization provides only six options for this timeline; further time would have allowed us to look closer at the history of each election in order to put together a longer story. However, we felt that these six elections provided a story that had enough facts for viewers to learn about important elections in history.

WINNER BARACK OBAMA!
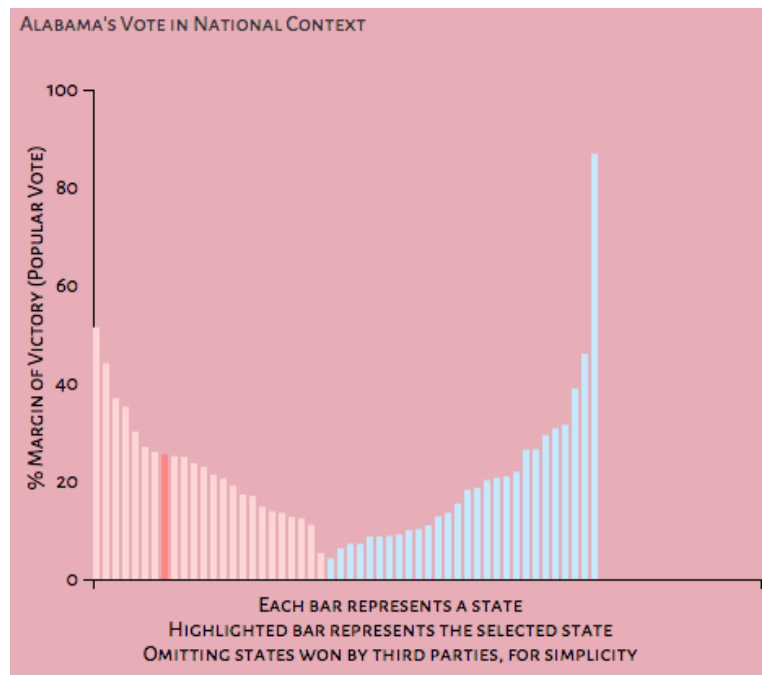
Per State Views

Our main choropleth also allows the user to click on each state. Thus, below the main choropleth, we have additional views that give data on the chosen state. The title informs the user of which state they have selected, as well the year for which data is displayed. For the convenience of the user, we have also included a drop-down list that will allow the user to select a state from the list of all states and the District of Columbia. The following views change accordingly per state. By default, the selected state on page load is Alabama.

The first row of views, whose background colors are shades of pink, we have implemented views that change per year. The row of features whose background is tinted blue-green gives data about each state that spans all of the elections years from 1912 to 2012.

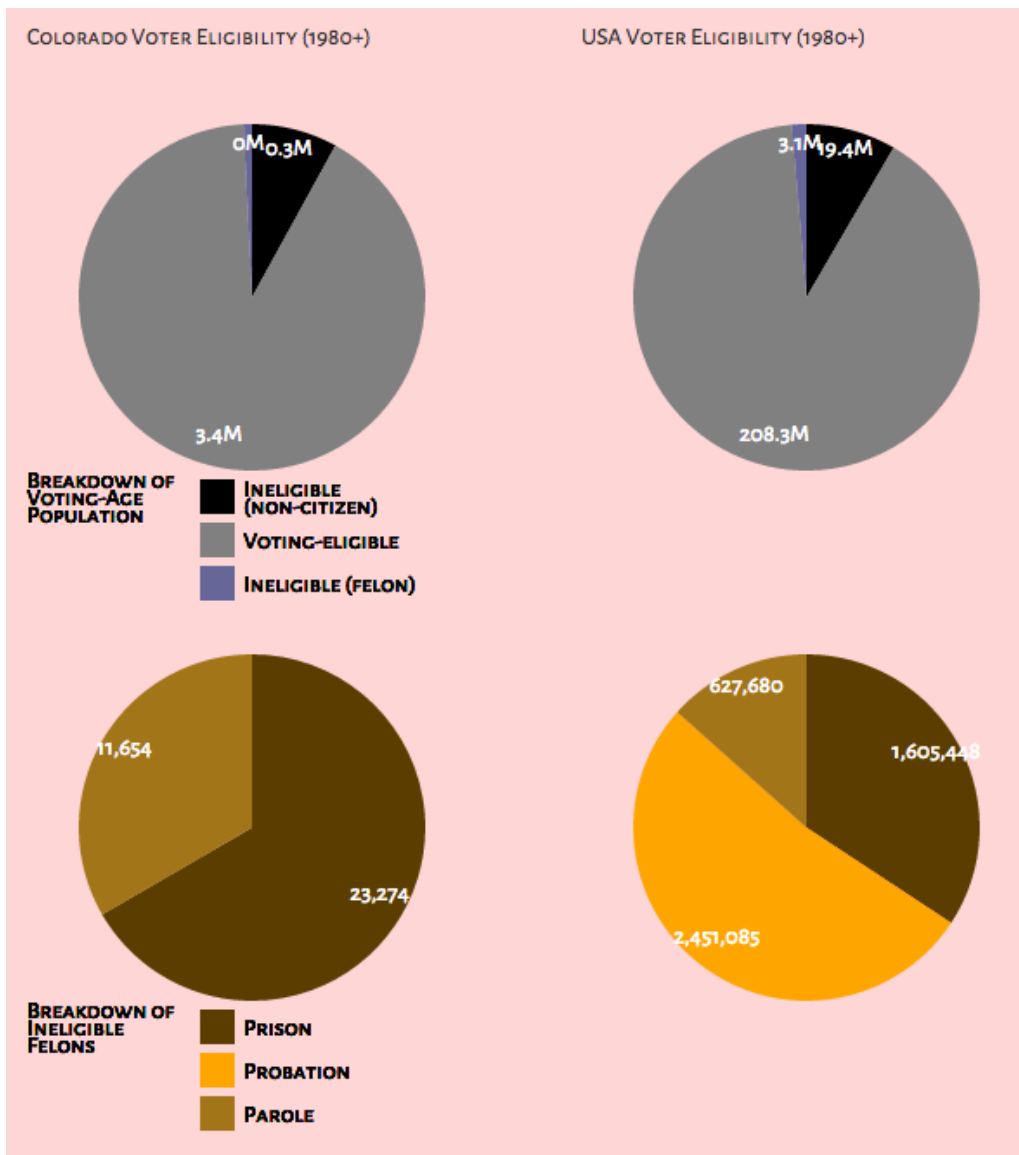**State to Nation Comparison**

For this view, we have implemented a dynamic bar chart that changes as you move the slider. The bar that has been highlighted (is darker in comparison) will represent the state that you have selected. As the year is changed, you will be able to see how that state changes in the context of the entire nation's voting margins.

To the right of the previous view, we have created pie charts to give the user some information about demographics residing within each state.
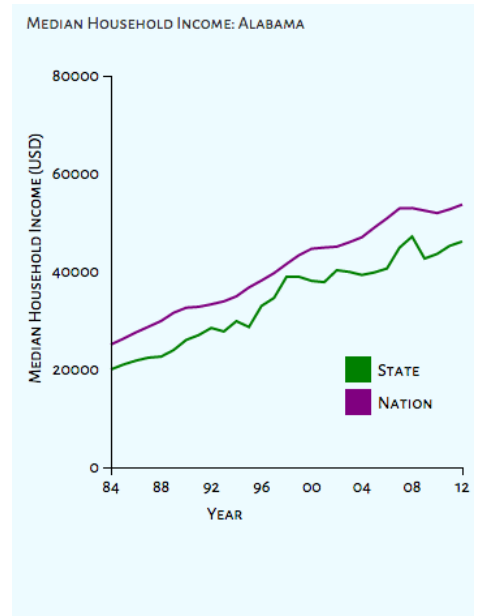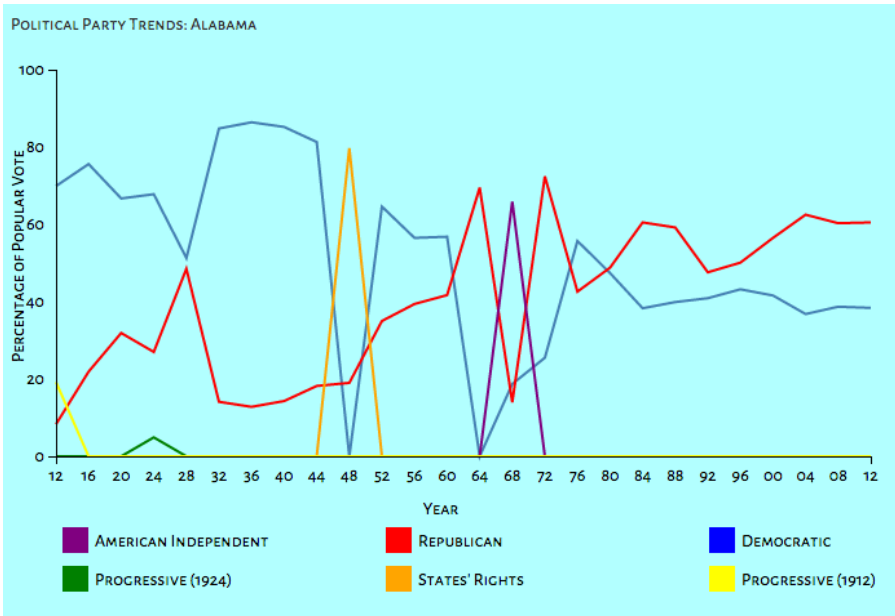
The upper two pie charts detail the breakdown of the voting-age population, including those ineligible due to felon status. The lower two pie charts then relate to the above pies in that they give the breakdown of ineligible felons in each state. Each slice in the pie charts also gives those number breakdowns, in order to impart more information to the user.

As you change the year, the user can compare each state's breakdown to the nation's breakdown. Due to limited data, we display pie charts for each state beginning in 1980 and continuing onwards to 2012.
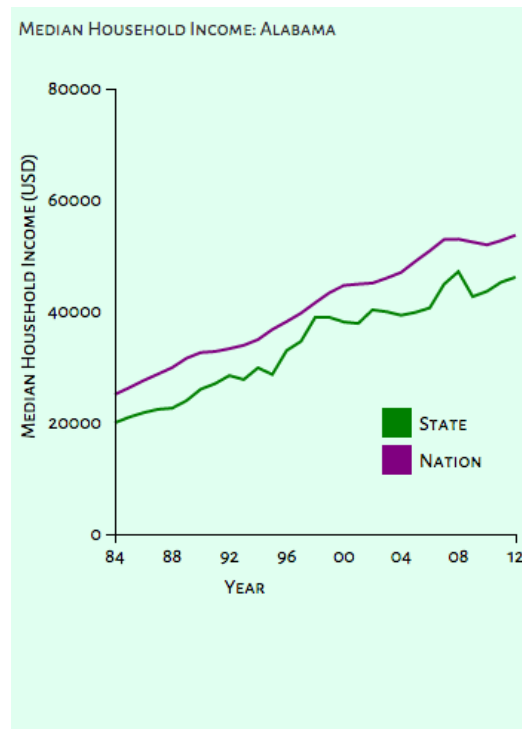
State Trends

The last row of views in our visualization, all tinted a blue-green, gives data about each state that spans all of the elections years from 1912 to 2012. This overview aims to give our audience a comprehensive understanding of how trends have changed in the past 100 years of elections, for the variables of political party and voter turnout. Legends below each picture provide information about each coloured line.

We have also created an additional graph (pictured below) for each state's median household income; however, this was limited in our data to the years of 1984 to 2012.

EVALUATION
What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?

Milestone 1
We haven't finished inputting all the data into our existing structure, so we aren't ready to draw sweeping conclusions about electoral or state trends over a long time period. At the moment we have a relatively small sample size (1980 to present) - less for some variables. We also have only implemented the bare essential component (choropleth and slider) of the overall visualization suite.

Looking at what we already have, however, does already display some interesting trends. We can see both individual election trends, like Ronald Reagan's strong victories. we can also see the general trends in the past few elections where very few states have changed which party they vote for. There are countries that are almost always more partisan than others, and we can clearly see the "swing" states like Ohio by their pale color.

Also, the sheer process of cleaning the data unearths a lot of fascinating outliers and noteworthy data points, such as faithless electors, states with unusual electoral law, etc. These outliers will be noted in the end story of our overall visualization!

In terms of how well our visualization works right now, we obviously have a long way to go toward the final implementation. Further improving our implementation will take the form of finishing up the ideas mentioned throughout this book, as well as potentially drawing in more and new data fields for state-level analysis.

Final Evaluation

As indicated elsewhere in this process book, much can be learned from this visualization, particularly by looking at trends over time for individual states and the nation as a whole. A mere glance at our visualization can leave the user with a sizable amount of information regarding the nation's electoral history and how particular regions of the nation might vote.

One of the most interesting conclusions about our data is the amount of variation between states. Depending on which states are chosen by the user, a variety of trends can be seen, which may inspire further research in to each state's history. This was one of the main goals of our inspiration, and seeing our work realised in this visualization is incredibly satisfying.

Over the course of this project, we have learned much more about the process of data visualization. There is always a greater number of features and interactivity that we could have implemented without the time constraint, but we have learned a great amount through the making of this project. Had we had an even longer period to work on this, our visualization would likely draw in more thorough datasets to provide users with a fuller story of what we currently have.

SOURCES

CODING
http://bl.ocks.org/mbostock/4060606:
  Choropleth (Bostock, Mike; 2012): Basic choropleth example.

http://d3js.org/:
  d3.js (Bostock, Mike; 2015): D3 library.

DATA
http://www.presidency.ucsb.edu/showelection.php?year=2012
  presidential elections 1789 to 2012, per state: total votes & per candidate

https://docs.google.com/spreadsheets/d/1bH38j6_e8yA9xq8OMIyLOL6h_iTS7ABQ
MKNxzFgKBDo/edit#gid=435419492
  The above link has national turnout rate for 1789-2012.

https://docs.google.com/spreadsheets/d/1or-N33CpOZYQ1UfZo0h8yGPSyz0Db-xj
mZOXg3VJi-Q/edit#gid=1670431880
  The above link has the following variables for 1980-2014, by state:
  voting-eligible population, votes counted, voting-age population, %
  non-citizen, and population of prison/probation/parole/ineligible felon.

http://www.presidency.ucsb.edu/data.php
  more detailed election data + data on the presidencies themselves

https://cse.google.com/cse/publicurl?cx=002720237717066476899:v2wv26idk7m
  Google's Datasets Search Engine

http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src
=bkmk
  race and ethnicity data by state and year

https://www.census.gov/popest/data/state/asrh/1980s/tables/estage80.txt
  age data by state and year

http://en.wikipedia.org/wiki/List_of_U.S._states_by_GDP
  gross state product data by year

http://www.census.gov/hhes/www/income/data/statemedian/
  The above link yields median household income from 1984-2013 from
  census.gov.

http://en.wikipedia.org/wiki/United_States_presidential_election
  For each story, we pulled snippets + pictures on each election year page.