CS 171/CSCI E-64: Visualization

Homework 2, Project Proposal

James Goodspeed – jgoodsp@fas.harvard.edu

Project Title

Campaign Donations for the 2012 Presidential election in the state of New Hampshire

Research questions and hypotheses

The primary question I'd like to answer is how are individual Presidential campaign donations broken down by town and by party in the state of New Hampshire. I'd like to correlate this to the per capita income level in each of the New Hampshire towns and see if there are any patterns such as whether 'rich' towns give predominantly to one party or another.

My hypothesis is that the southern portion of the state will have the majority of the donations across both parties. I believe that the Democratic donations will be centered mostly around the seacoast and around Dartmouth College in Hanover, traditionally liberal and wealthy areas of the state. I think the Republican donations could also be high around the seacoast, but also further north around Lake Winnipesauke. The northern portion of the state is generally more conservative, but, I think, less likely to contribute to Presidential campaigns.

Motivation

My main motivation is just interest. I live in New Hampshire and am interested in politics (my undergraduate degree is in Political Science) especially campaign finance reform. I came across the data and thought I could apply it to a region where I live to see if any interesting patterns emerged.

Data Source(s) and technical process:

The main data source for this project will be the New York Times Campaign Finance API. Specifically I will be pulling the Presidential State/Zip totals:
http://developer.nytimes.com/docs/read/campaign_finance_api#h3-pres-state-zip

To do this I will need to programmatically get a list of all the New Hampshire zip codes and pull the API data for each zip code in New Hampshire. The next step will be to aggregate the totals for all of the Republican candidates and all of the Democratic candidates for each town. This could be done with a shell script or using Google Refine.

From there I will need to map the town names to the zip codes. This could be done with a Python script or even a simple shell script. A data source for this exists here:
http://www.directorynh.com/NHReferenceDesk/ZipCodes.html

The data for the per capita income of each town will come from the US Census Bureau site:
http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk
The same mapping of town to zip will be done to incorporate this data.

CS 171/CSCI E-64: Visualization

Homework 2, Project Proposal

James Goodspeed – jgoodsp@fas.harvard.edu

An additional piece of data might be geographic locations (latitude and longitude) for the towns so that the data could be represented on a map of the state.

I will most likely store this data in a csv format. Merging of the data could be done with a combination of shell scripts, Google Refine and Google Fusion Tables.

Potential Visualizations

The most obvious visualization that springs to mind is a map of New Hampshire. Each town on the map could be shaded in varying intensities (depending on the campaign donation in that town) of either red for the Republican Party or Blue for the Democratic Party, or a combination of the two depending on the data. The darker the color the more money the people in that town donated to a political campaign.

I'm not sure of the best way (yet) to represent the per capita income levels for each town. A corresponding bar chart or some kind of indicator on the town that would show height based on income levels, but this might actually make the visualization more confusing. Another idea might be to allow the user to select whether they want to see per capita data or donation data on the map and let the user flip between the two. I'm looking forward to learning more about different ways of possibly presenting this data.

Yet another way to think about visualizing this is to use a scatter plot to show the correlation between donations and per capita income levels. On the X axis could be the donation levels per town, the Y axis could be the per capita level for that town and the dots could be color coded based on party.

A final option, suggested by Devin Shackle, would be a map like this showing the donation levels, but I'm not sure how I would correlate the per capita income levels in each town.

http://www-personal.umich.edu/~mejn/cartograms/hiv1024x512.png