
Campaign Finance City Council and State Representative Scorecard

Yuang Liu (U99473611), Jun Xiao(U85900288), Qi Yin (U31787103), Yiming Miao(U76158067)

Department of Computer Science
Boston University
yaliu@bu.edu

1 Motivation and background

Solomon Kahn has built Explore Campaign Finance to help citizens and journalists explore how much money is given to political candidates and by whom. An overwhelming percentage of Americans believe there is too much money in politics, but even educated people have a lack of basic understanding of the facts of how much money is in politics. This project will arm them with information about any politician they are interested in.

However, the website hasn't been updated since 2016 though. Our client wants to continue this project to have an insight into the data from 2016. Last year the students examined publicly available campaign contributions data to city council members and categorized them by industry. They then created a scorecard for each city council member and classified donations by industry, donation size, and employment type.

2 Goals, hypothesis, outcomes

1) Goals:

- a. Our main goal is to analyze the average money donated from specific industries by scraping the donation data and industry information for classification.
- b. And we will also draw the outlier of our donors, who they are, what occupation they are, etc.
- c. Besides, we will analyze the influence directed toward the State Legislature and City Council Committee Members by identifying and classifying donations to city council members and state representatives in the key committees given by our client.

2) Hypothesis:

- a. All of the donation data are objective and correct.
- b. All donor and campaign information is filled in correctly and all data is based on actual conditions.
- c. Most of the employers' information can correspond to related industries.

Under these conditions, we can perform better processing and analysis based on the data, so as to obtain more objective and practical statistical results.

3) Outcomes:

Hopefully, our project will help our client, Progressive Massachusetts, to have an insight into how much money specific industries donate and how much money is one specific committee received.

3 Non-goals

- 1) We will not judge the accuracy of the original data.
- 2) We will not manually correct the collected data to prevent it from affecting the authenticity of statistical results.

4 End product

The definitions of done for this project

- 1) Get enough information to determine which industry a company belongs to.
- 2) Fully classify the data that will provide enough information to answer the open question.

By the end of the project, we will deliver these products:

- 1) Script that will scrapy the Secretary of State Website.
- 2) Script that will run over all the data and classify the data by industry.
- 3) Script that can produce image illustration over the classification results such as Figure 1.

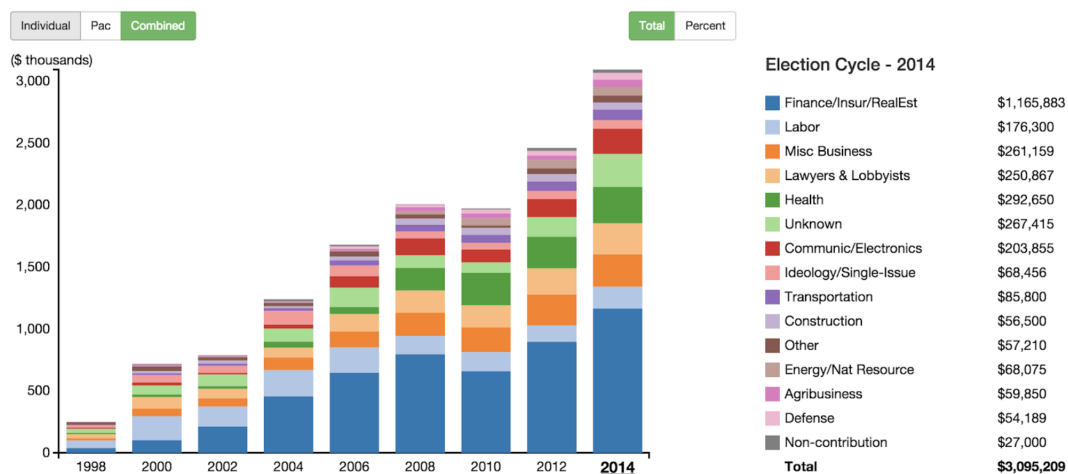


Figure 1: The amount that politicians raise from different industries, in total and on a percentage basis. The companies and pacs that donate the most, and the people within those companies who made those donations[1].

5 Open questions

- 1) Is there a pattern or insights of these individuals such as what time of year do they donate or are there some people donating to multiple candidates from the same companies?
- 2) Do members of the committees most relevant to a given industry receive a significantly larger volume of donations?
- 3) How much are PACs and Unions giving?
- 4) Who are the largest industry donors including PACs, Unions, Companies, Individuals?
- 5) What percent of Reps overall receive donations from the same firm by industry.

6 General approach

- 1) Review the Fall Semester team's work.
- 2) Collect campaign donations data from OCPF, which we can know donors donates how much money to whom.
- 3) Collect industry information data from the Secretary of State Website, which can help us know which category a company belongs to.
- 4) Collect industry information from other sources such as registered lobbyists, PACS and Unions/Associations to get more information indicates specific industry of a business.
- 5) Clean and classification, where we first clean the data and then classify all companies by industry, classify OCPF data by occupation and employer.
- 6) Analyze the classified data and get the average donations from specific industries.

References

[1] <http://explorecampaignfinance.com/>