
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 Data Collected

Data we've collected is as follows:

1. OCPF data:

The data indicates who contributed how much money to whom.

The data is in Microsoft Access Database, We only list useful attributes, and the structure of the data is:

Table	vupload_master	
Info	information of the receiver	
Attributes	report_id	the identifier for distinct report
	cpf_id	the identifier for distinct donation receiver
	full_name	the name of the receiver

Table	vupload_tcurrent_receipt	
Info	information of the contributor	
Attributes	report_id	which report this entry is belonging to
	date	when the contribution happened
	first_name	the first name of the contributor
	last_name	the last name of the contributor
	city	the city of the contributor
	state	the state of the contributor
	occupation	the occupation of the contributor
	employer	the company that the contributor works in
	amount	how many money does the contributor contribute

2. List of keywords of industry name:

Some list like this, which we can use the keywords to determine which industry contributed how much money. The list is shown as Table 1

Industry/ Sector	Occupation keywords	Company/ Em- ployer keywords	Names of PACS
Real Estate	Architect (am- biguous)	Property Manage- ment	Greater Boston Real Estate Board PAC
	Developer (ambiguous)	Construction	NAIOP MA Commercial Real Estate
	Builder	Building	Development Political Ac- tion Committee
	Carpenter	Property	Real Estate Bar Association for MA PAC
	Painter	Surveyor	
	Realtor	Real Estate	
	Contractor	Civil Engineering	
	Planner		
	Environmental Reviewer		
	Surveyor		

Table 1: Keywords for Employers and Occupations

3. Registered-candidates

Name, cpf_id, office type of all registered candidates in house or senate.

4. Registered-PACs

Name, cpf_id of all the registered political action committees.

5. The Information of City Council Committees is shown as Table 2

6. Business information related to Real Estate

We scrape the website (<https://www.cityofboston.gov/cityclerk/dbasearch/Default.aspx>) to get all business information related to Real Estate based on the keywords.

The scrape is divided into two parts.

- i) Scrape all businesses that have keywords in its' name. The keywords of company and employer are listed in Table 1 . The results are shown as Figure 1 . No results are found of the keyword Civil Engineering.
- ii) Scrape all businesses with the type of Real Estate. The results are shown as Figure 2 .

file number	name	type
cc311344	charleagle property management	property management
cc371471	j & a property management	property management
cc392428	flynn & sons property management	property management
cc303728	j a l property management	property management services
cc364756	light hill property management	real estate property management
cc341661	rosamond property management	property management
cc386790	community real estate & property management	real estate and property management
cc318802	extra mile property management	property management
cc357050	modernista property management	property management
cc338025	southampton street property management	property management
cc342261	sgh property management	property management
cc371471	j & a property management	property management
cc418802	extra mile property management	property management
cc351079	real property management boston	property management
cc391830	neighborhood property management	property management
cc364756	light hill property management	real estate property management
cc386790	community real estate & property management	real estate and property management
cc341316	muy bueno cleaning & property management	cleaning & property management
cc358019	a-lll property management	property management
cc357050	modernista property management	property management

(a) Property Management

file number	name	type
cc175919	santiago & sons construction co.	construction
cc177281	dejoy construction	construction
cc199937	sdh construction and design	construction
cc304266	jf construction	construction
cc333477	hurtado construction	construction
cc177167	fort hill construction	general contractor
cc199910	i & e construction	construction business
cc302183	mario's construction	construction
cc310101	wood and stone construction	construction
cc337823	j s d construction	construction
cc364309	alpa construction	general contractor
cc366331	y framing and construction	home improvement
cc377810	newton construction	home improvement
cc381105	rizzo construction group	construction and contracting
cc387620	w.d.s. commercial construction	commercial construction
cc364357	mass properties maintenance & construction	maintenance/construction
cc371663	coastal construction & management co.	general construction
cc377998	r & n construction	construction
cc384648	new world construction	construction
cc405932	azabache construction	construction/maintenance

(b) Construction

file number	name	type
cc175947	chenub property services	property services
cc318787	boston property advisors	real estate appraisal
cc366842	commercial property services	property management
cc341805	mack property management	multi-family property management
cc346638	charles pessia professional property services	contractor
cc309221	boston property rentals	property management
cc338025	southampton street property management	property management
cc406140	king craft property management	property management
cc414806	next property management	multi-family property management
cc304091	cmd property services	construction
cc383284	vision's property cleaning and lawn care services	cleaning & landscaping
cc406140	king craft property management	property management
cc317553	property management solutions	property maintenance
cc351079	real property management boston	property management
cc367874	j & b complete property maintenance	janitorial cleaning
cc191830	neighborhood property management	property management
cc313413	bolado property services	web design/sales/serv
cc341316	muy bueno cleaning & property management	cleaning & property management
cc363562	dpy property maintenance	property maintenance
cc358019	a-lll property management	property management

(c) Property

file number	name	type
cc339056	new england building maintenance / carpet cleaning	maintenance / cleaning
cc362996	academic shipbuilding center trump	online education
cc362999	architecture shipbuilding center trump	online education
cc387240	the rock building service	cleaning services
cc413238	re building maintenance & services	construction/maintenance
cc362799	creative building associates	construction/remodeling
cc362997	information shipbuilding center trump	online education
cc363001	shipbuilding center for social systems trump	online education
cc392432	the regency building company	residential apartments
cc416071	woodchuck's building & home center	bldg. material sales
cc339056	new england building maintenance / carpet cleaning	maintenance / cleaning
cc362996	academic shipbuilding center trump	online education
cc362999	architecture shipbuilding center trump	online education
cc387240	the rock building service	cleaning services
cc413238	re building maintenance & services	construction/maintenance
cc362799	creative building associates	construction/remodeling
cc362997	information shipbuilding center trump	online education
cc363001	shipbuilding center for social systems trump	online education
cc392432	the regency building company	residential apartments
cc416071	woodchuck's building & home center	bldg. material sales

(d) Building

file number	name	type
cc199075	chestnut homes real estate	real estate
cc199319	hammond residential real estate	residential real estate brokerage
cc199326	hammond residential real estate	residential real estate brokerage
cc199332	hammond residential real estate	
cc199336	hammond real estate	real estate
cc199302	caldwell real estate	real estate
cc199323	hammond residential real estate	residential real estate brokerage
cc199327	hammond residential real estate	residential real estate brokerage
cc199335	hammond residential real estate	
cc199339	hammond residential real estate	
cc333538	hart real estate	real estate
cc363822	the french real estate company	real estate brokerage
cc366541	ed maher real estate	real estate
cc401714	t - evolution real estate services	real estate
cc423154	dhci real estate group	real estate office
cc341690	your neighborhood real estate management group	real estate management
cc372116	technology real estate	real estate
cc363495	neponset river real estate	real estate
cc411026	crpboston real estate	real estate
cc349760	grandbridge real estate capital	commercial mortgage broker

(e) Real Estate

file number	name	type
cc357758	feldman land surveyors	surveying

(f) Surveyor

Figure 1: All businesses have keywords in its' name

file number	name	type
cc175061	brighton realty	real estate
cc175204	southview realty group	real estate
cc177200	a+ realty	real estate
cc199330	hammond residential	residential real estate brokerage
cc200188	s o industries	real estate

Figure 2: All businesses with the type of Real Estate

Occupation	City Council Committees' Name	Name in OCPF Data	CPF_ID
Real estate	KENZIE BOK	Priscilla MacKenzie Bok	17164
	ANNISSA ESSAIBI GEORGE	Annissa Essaibi George	15618
	FRANK BAKER	Frank Baker	15333
	ANDREA CAMPBELL	Andrea Joy Campbell	15931
	MICHAEL FLAHERTY	Michael F. Flaherty Jr.	12892
	ED FLYNN	Edward Michael Flynn	14391
	JULIA MEJIA	Julia M. Mejia	17092
	LYDIA EDWARDS	Lydia Edwards	16314
	LIZ BREADON	Elizabeth A. Breadon	17111
	MICHELLE WU	Michelle Wu	15563
	RICARDO ARROYO	Ricardo N. Arroyo	17105
Law Enforcement	JULIA MEJIA	Julia M. Mejia	17092
	ED FLYNN	Edward Michael Flynn	14391
	RICARDO ARROYO	Ricardo N. Arroyo	17105
	ANNISSA ESSAIBI GEORGE	Annissa Essaibi George	15618
	MATT O'MALLEY	Matthew J. O'Malley	14092
	MICHELLE WU	Michelle Wu	15563
	ANDREA CAMPBELL	Andrea Joy Campbell	15931
	MICHAEL FLAHERTY	Michael F. Flaherty Jr.	12892
Energy	MATT O'MALLEY	Matthew J. O'Malley	14092
	MICHELLE WU	Michelle Wu	15563
	KENZIE BOK	Priscilla MacKenzie Bok	17164
	LIZ BREADON	Elizabeth A. Breadon	17111
	LYDIA EDWARDS	Lydia Edwards	16314
Healthcare	LIZ BREADON	Elizabeth A. Breadon	17111
	MATT O'MALLEY	Matthew J. O'Malley	14092
	ANNISSA ESSAIBI GEORGE	Annissa Essaibi George	15618
	JULIA MEJIA	Julia M. Mejia	17092
	MICHELLE WU	Michelle Wu	15563
Higher Education	ANNISSA ESSAIBI GEORGE	Annissa Essaibi George	15618
	ANDREA CAMPBELL	Andrea Joy Campbell	15931
	RICARDO ARROYO	Ricardo N. Arroyo	17105
	LYDIA EDWARDS	Lydia Edwards	16314
	MICHAEL FLAHERTY	Michael F. Flaherty Jr.	12892
	KIM JANEY	Kim Janey	16537
	MATT O'MALLEY	Matthew J. O'Malley	14092

Table 2: City Council Committees Information

2 Data to be Collected

We are trying to collect the overall description of a company.

We tried to scrape the information from the <http://corp.sec.state.ma.us/CorpWeb/CorpSearch/CorpSearch.aspx>. However, the website is protected by anti-scrape methods. It's really hard to scrape the data. We are discussing this issue with our PMs.

3 Analyze Real Estate industry Based on Keywords

1. Overview.

This is the analyze of OCPF data with keywords from Real Estate.

The number of transactions is **17286**, the total amount is **\$6889810.82**, and the number of receivers is **1012**.

The receivers who receive the most amount is **Martin J. Walsh** whose CPF_ID is **13171** and the total amount is **\$781707.1**

The receivers who receive the least amount is **1199 SEIU MA PAC** whose CPF_ID is **80769** and the total amount is **\$5.0**

2. Compute by Season

The result is shown as Table 3 and Figure 3

Season	Amount
1	\$1134122.81
2	\$1901463.20
3	\$1347196.58
4	\$2507028.23

Table 3: Compute by Season

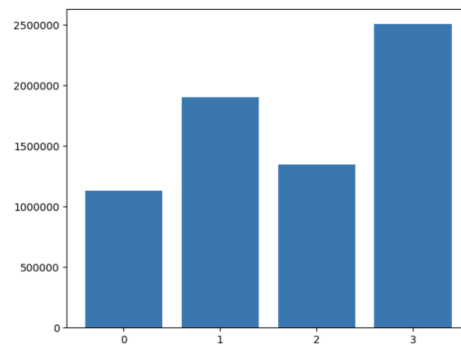


Figure 3: Compute by Season

3. Compute by Year

The result is shown as Table 4 and Figure 4

Year	Amount
2016	\$1942829.81
2017	\$1736838.23
2018	\$1739569.30
2019	\$1470573.48

Table 4: Compute by Year

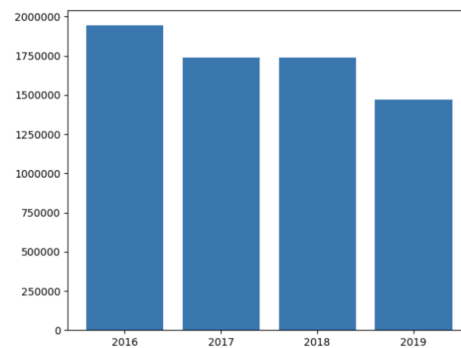


Figure 4: Compute by Year

4. Compute by Size

The result is shown as Table 5 and Figure 5

Size	Number	Amount
< \$25	422	\$4304.76
\$25 – \$99	1884	\$86518.04
\$100 – \$249	4903	\$669819.92
\$250 – \$499	3714	\$976471.49
\$500 – \$1000	3833	\$1942196.61
> \$1000	2530	\$3210500.00

Table 5: Compute by Size

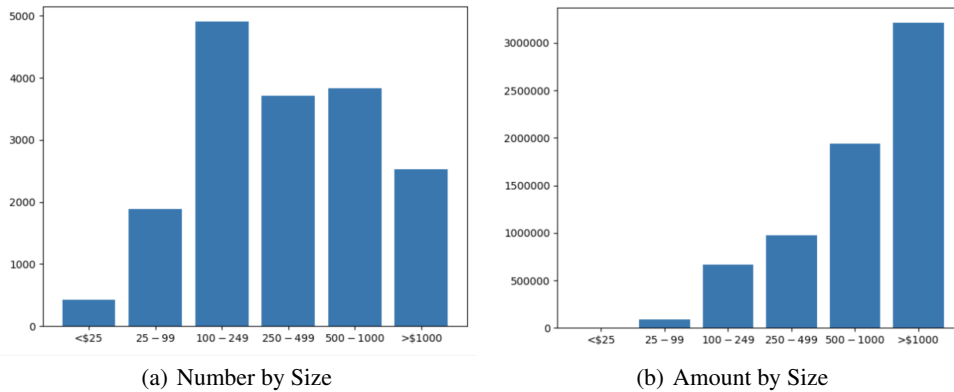


Figure 5: Compute by Size

5. Compute by City

The number of cities is **770**.

The city that donates the most is **Boston** with the amount of **\$953902.84**, and the cities that donate the least are **Buffalo, Berkeley, San Pablo, Sunnyvale, Los Gatos, South Charleston, Locust Grove, Lemon Grove** with the amount of **\$5.0**

6. Compute by State

The number of states is **42**.

The state that donates the most is **MA** with the amount of **\$6130812.41**, and the state that donates the least is **AL** with the amount of **\$5.0**

The amount in state(MA) is **\$6130812.41**, and the amount out of state is **\$758998.41**

7. Compute by Contributor Type

The result is shown as Table 6

Contributor Type	Amount
Individual	\$6886760.82
Other	\$3050.0

Table 6: Compute by Contributor Type

8. Compute by City Council Committees

The result of Real Estate is shown as Table 7

City Council Committees' Name	Name in OCPF Data	CPF_ID	Amount
KENZIE BOK	Priscilla MacKenzie Bok	17164	\$ 7500.0
ANNISSA ESSAIBI GEORGE	Annisssa Essaibi George	15618	\$ 35895.0
FRANK BAKER	Frank Baker	15333	\$ 70750.0
ANDREA CAMPBELL	Andrea Joy Campbell	15931	\$ 19445.0
MICHAEL FLAHERTY	Michael F. Flaherty Jr.	12892	\$ 122300.0
ED FLYNN	Edward Michael Flynn	14391	\$ 56350.0
JULIA MEJIA	Julia M. Mejia	17092	\$ 4530.0
LYDIA EDWARDS	Lydia Edwards	16314	\$ 22151.56
MICHELLE WU	Micheele Wu	15563	\$ 57630.0
RICARDO ARROYO	Ricardo N. Arroyo	17105	\$ 13850.0

Table 7: Compute by City Council Committees