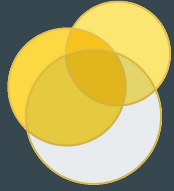# Fortune 500 Electoral Influence

## *1990-2012*

● ● ●

Paul Mack, Alden Golab

SUNLIGHT
FOUNDATION

## About the data

Sunlight Foundation State and
Federal Campaign Contributions
1990 - 2012
(Two Year Increments Corresponding to Cycle)

Federal law requires that campaigns and PACs disclose information about the identities of individuals, corporate or private, who donate over $200 to a in a given year or cycle.

Name
Residence
Employer/Organization
Recipient Information
Date

___

# 15.1 GB

Or 50.7 Million transactions for all
Federal and State level campaigns

# Top individual contributors tend to be lesser known

| | Contributions | Individual | | Contributions | Individual |
|---|---|---|---|---|---|
| 1 | $729,005,408.00 | Mr. Lee L Mercer Jr. | 5 | $28,215,750.00 | Thomas F. Steyer |
| 2 | $54,756,566.00 | Stephen L. Bing | 11 | $26,798,373.20 | James & Virginia Stower |
| 3 | $53,822,872.00 | Linda McMahon | 13 | $19,516,336.04 | James E. Pederson |
| 4 | $39,830,810.31 | Bob J. Perry | 15 | $17,889,650.00 | John R. Raese |

# Top organizations are well known heavy-weights

| Rank | Contributions | Organization |
|---|---|---|
| 1 | $1,348,150,439.38 | National Republican Congressional Committee |
| 2 | $991,885,410.28 | Republican National Committee |
| 3 | $777,393,014.00 | DNC Services Corps |
| 4 | $683,447,745.15 | Democratic Congressional Campaign Committee |
| 5 | $631,654,092.89 | Democratic Senatorial Campaign Committee |
| 11 | $137,855,333.64 | Service Employees International Union |
| 13 | $128,364,947.88 | California Teacher's Association |
| 15 | $97,952,177.43 | Philip Morris |
| 21 | $84,236,959.00 | National Rifle Association |
| 23 | $81,540,958.00 | National Association of Realtors |

There is substantial variation between the causes to whom Fortune 500 corporations donate and those to whom their employees donate.

# Resolve entities with substantial spelling variation and the occasional unicode character.

Challenge 1

# Sample of Goldman Sachs Variations

GOLDMAN SAKS      GOLDRAN SACHS

GOLDMAN, SACHS & CO      GOLDMAN SACHS & CO.

GOLDMAN- SACHS & CO.      GOLDMAN SACH & COMPANY

GOLMAN SACHS      GOLDMAN SACHS & COMPANY

GOLDMAN SACHS GROUP      GOLDMAN SACHS ASIA LLC

# Entity resolution algorithm

**Read in each entry and thoroughly clean all unicode**

**Loop through F500 authoritative list and match**

**Output JSON for later use**

```
GOLDMAN SAKS
```

```
Prudential
Financial

Humana, Inc.

The Goldman Sachs
Group, Inc.

Tesoro Corporation

Liberty Mutual
Holding Company
```

```
[{ 'GOLDMAN SAKS':

  'The Goldman Sachs
  Group, Inc.' }]
```

**Entity matching work**

```
1.  Remove stop words
2.  Strip misc. Chars
3.  Simple edit
    distance score
    using fuzzywuzzy
```

# Runtime.

# Runtime issues, even with parallelization

## Entity resolution algorithm

200,000 records:
    37 minutes 10 seconds on on local
512,729 records:
    5x the work per node
    AWS 20 nodes about 30 - 40 minutes

## Data aggregation algorithm

```
1.   Remove unicode encoding
2.   Strip characters
3.   Look-up organization namein
     JSON, held in memory on AWS
     using configure_options()
4.   Determine whether corporate
     or individual contribution
5.   Export to CSV
```

Achieved constant run time and could
process the full dataset in 30 minutes
on 40 nodes

# Entity reduction results from 1 percent of total data

**444** Uniques

**2000** Aliases

FMC  28 Aliases

AIG  26 Aliases

Coca-Cola  18 Aliases

Goldman Sachs  17 Aliases
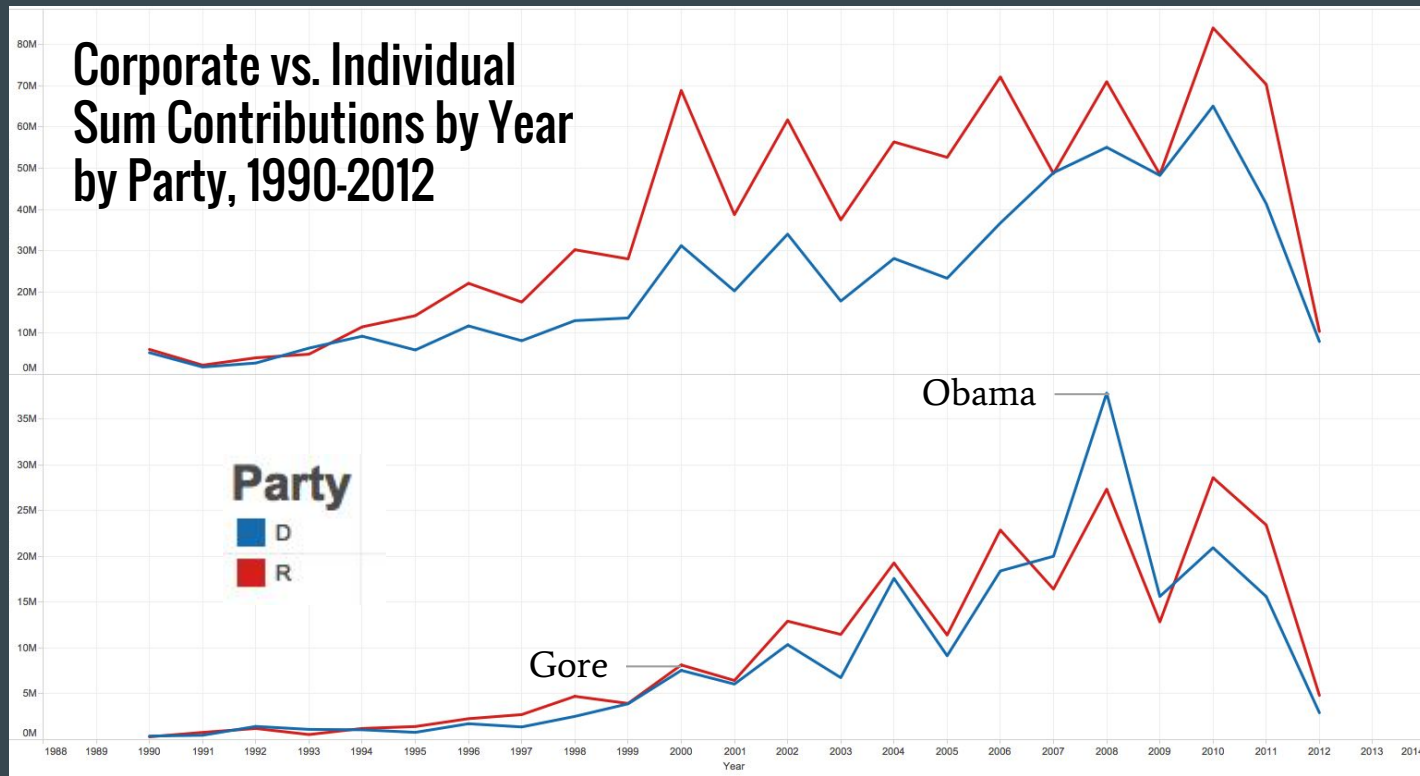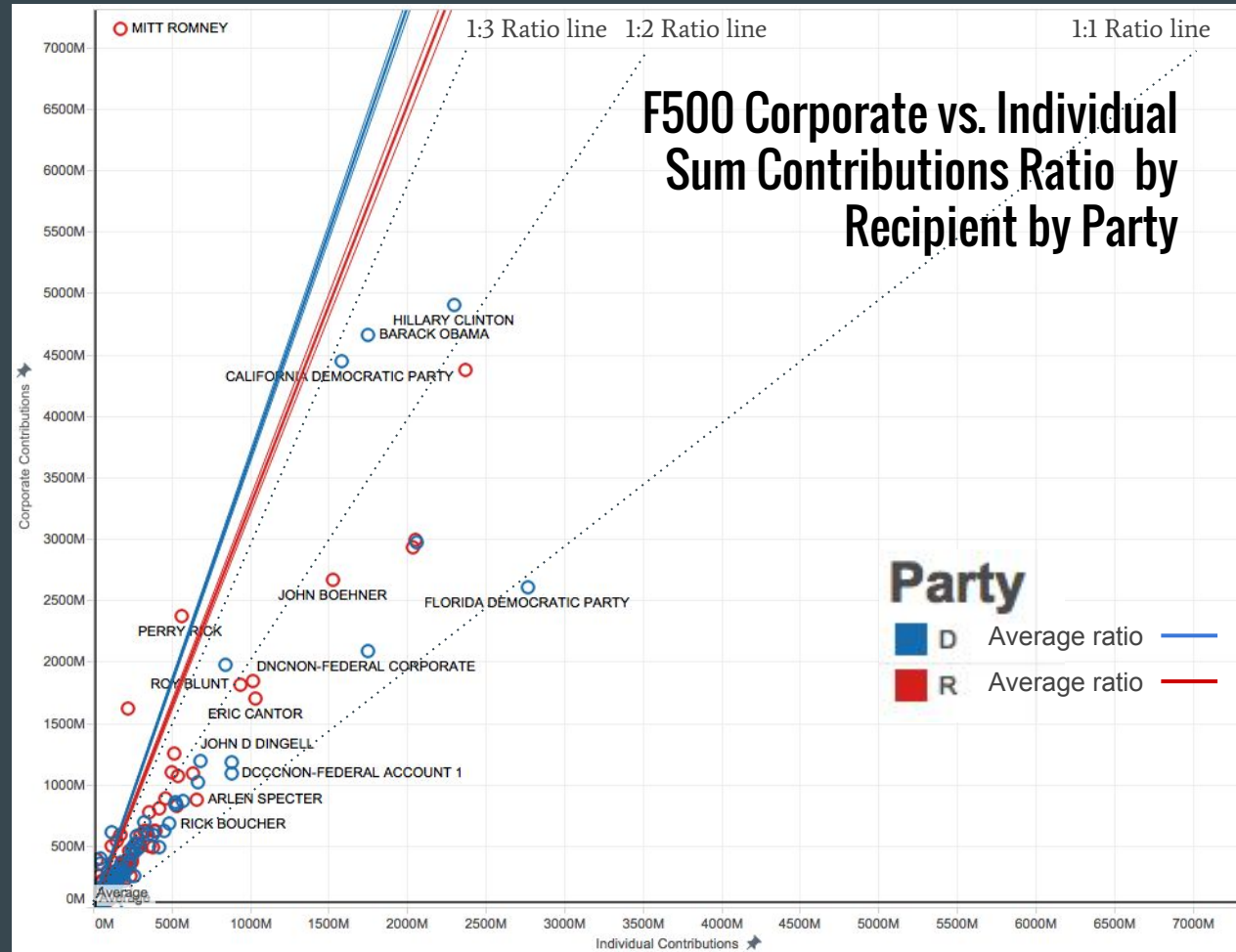
# Confirmed.

Hypothesis 1

# F500 Corporations don't always agree with employees

**Corporate Contributions**

**Individual Contributions**



Corporate vs. Individual
Sum Contributions by Year
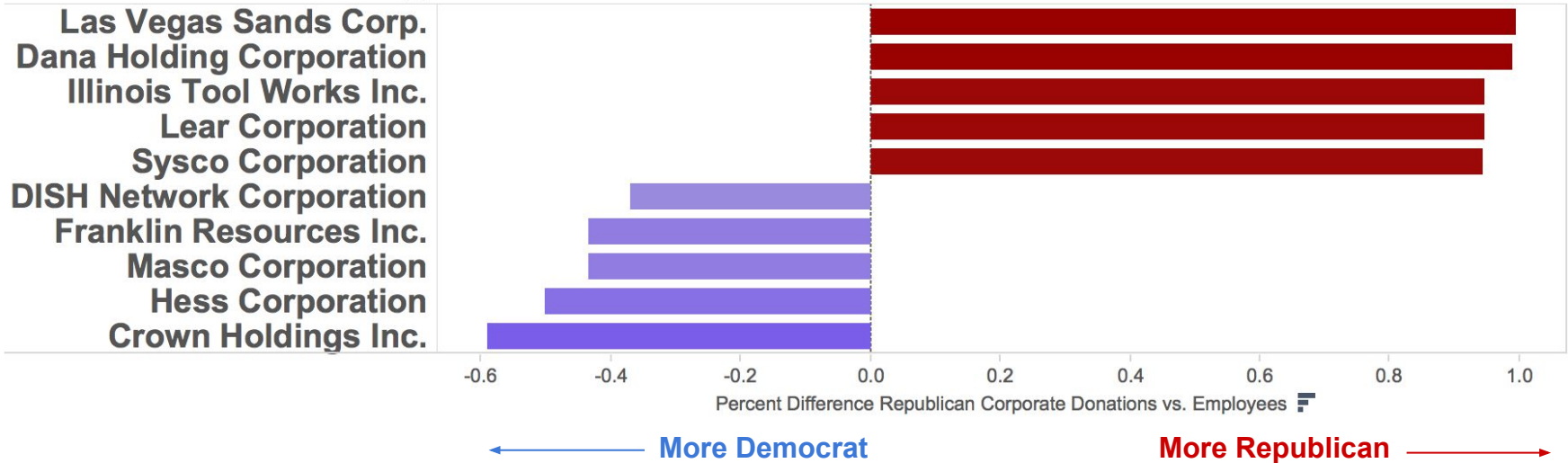by Party, 1990-2012

Party
D
R

Obama

Gore

While overall F500 corporations donate more to Republican candidates, when compared to their employees at a dollar-to-dollar ratio, corporations tend to slightly outweigh their employees on *Democratic* candidate and committee contributions



F500 Corporate vs. Individual Sum Contributions Ratio by Recipient by Party

1:3 Ratio line    1:2 Ratio line                    1:1 Ratio line

MITT ROMNEY

HILLARY CLINTON
BARACK OBAMA

CALIFORNIA DEMOCRATIC PARTY

JOHN BOEHNER

FLORIDA DEMOCRATIC PARTY

PERRY RICK

DNCNON-FEDERAL CORPORATE

ROY BLUNT

ERIC CANTOR

JOHN D DINGELL

DCCCNON-FEDERAL ACCOUNT 1

ARLEN SPECTER

RICK BOUCHER

Average

Corporate Contributions

Individual Contributions

**Party**

D    Average ratio
R    Average ratio

# Quantifying Disconnect



## Difference Percent Republican Contributions for Corporations vs. Individual, by Co.



Las Vegas Sands Corp.
Dana Holding Corporation
Illinois Tool Works Inc.
Lear Corporation
Sysco Corporation
DISH Network Corporation
Franklin Resources Inc.
Masco Corporation
Hess Corporation
Crown Holdings Inc.

-0.6   -0.4   -0.2   0.0   0.2   0.4   0.6   0.8   1.0

Percent Difference Republican Corporate Donations vs. Employees

← **More Democrat**          **More Republican** →

# Stuff We Want to Do

More specifically, we ran out of time to do this stuff because of AWS issues

## Analysis of Donor Networks

Extraction of all pairs and loading into a network analysis package, where edges represent money transfers, to study relationship of centrality to electoral success.

## Full Dataset Entity Resolution

Because of time constraints and issues with MRJob on AWS, we were constrained to Fortune 500 companies and their employees.

# Things We Learned

What we would tell ourselves

MRJob data passing protocols are difficult to understand, but useful

Reading in data from any source must deal with encoding issues by default

Entity reduction is computationally very expensive

AWS costs can balloon quickly

____

# Thanks!

● ● ●

Paul Mack, Alden Golab