

Political Contribution VS Election Results

Gerry Pratama
Hao Wu



The Relationship Between Total Political Contributions and Total Amount of People Vote



Our Datasets

Primary Dataset

Federal Election Commission:
Campaign Finance Data

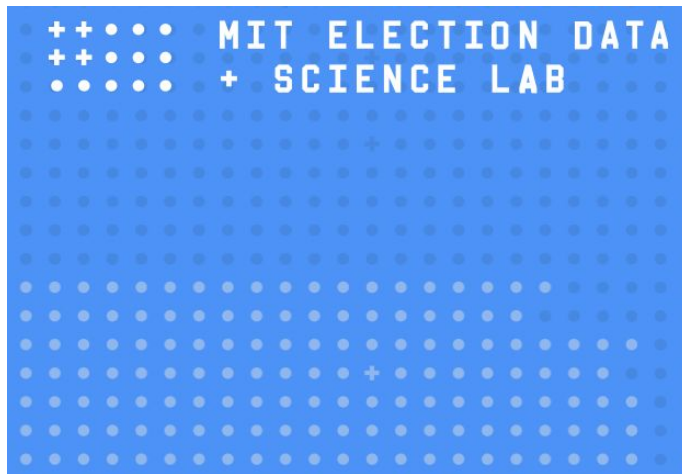
1. Candidate Master (7 files)
2. Committee Master (7 files)
3. Contributions from Committees to Candidates (7 files)



Secondary Dataset

MIT Election Lab:
US Election Data

1. Senate Elections
2. House Elections
3. President Elections





Transforming and Linking Datasets (SQL + Beam + DAG)

1. Create Staging and Modeled Dataset
2. Create Primary Key
3. Clean Column into Desired Type (Zip, Date, Year, Election Type)
4. **Transform Name Column in both DataSets and Link them**
 - a. Standardize Name Format: Last + First (UpperCase)
 - b. Get Rid of Middle Name
 - c. Get Rid of Suffix (Jr, Sr)
5. Link the Name with Candidate ID (PK, FK).

From DataSet One, FEC

Committee		
ID	STRING	
Name	STRING	
Treasurer	STRING	
Street1	STRING	
Street2	STRING	
City	STRING	
State	STRING	
ZIP	INTEGER	
Designation	STRING	
Type	STRING	
Party_Affiliation	STRING	
Filing_Frequency	STRING	
Category	STRING	
Year	INTEGER	
PK Label	STRING	

From DataSet One, FEC

Contributes		
Transaction_Type	STRING	
Date	DATE	
Amount	INTEGER	
FK Payee_ID	STRING	
PK Contribution_ID	STRING	
Year	INTEGER	
FK Committee_Label	STRING	
FK Candidate_Label	STRING	

From DataSet One, FEC

Supports		
PK,FK Committee_Label	STRING	
PK,FK Candidate_Label	STRING	

From DataSet One, FEC

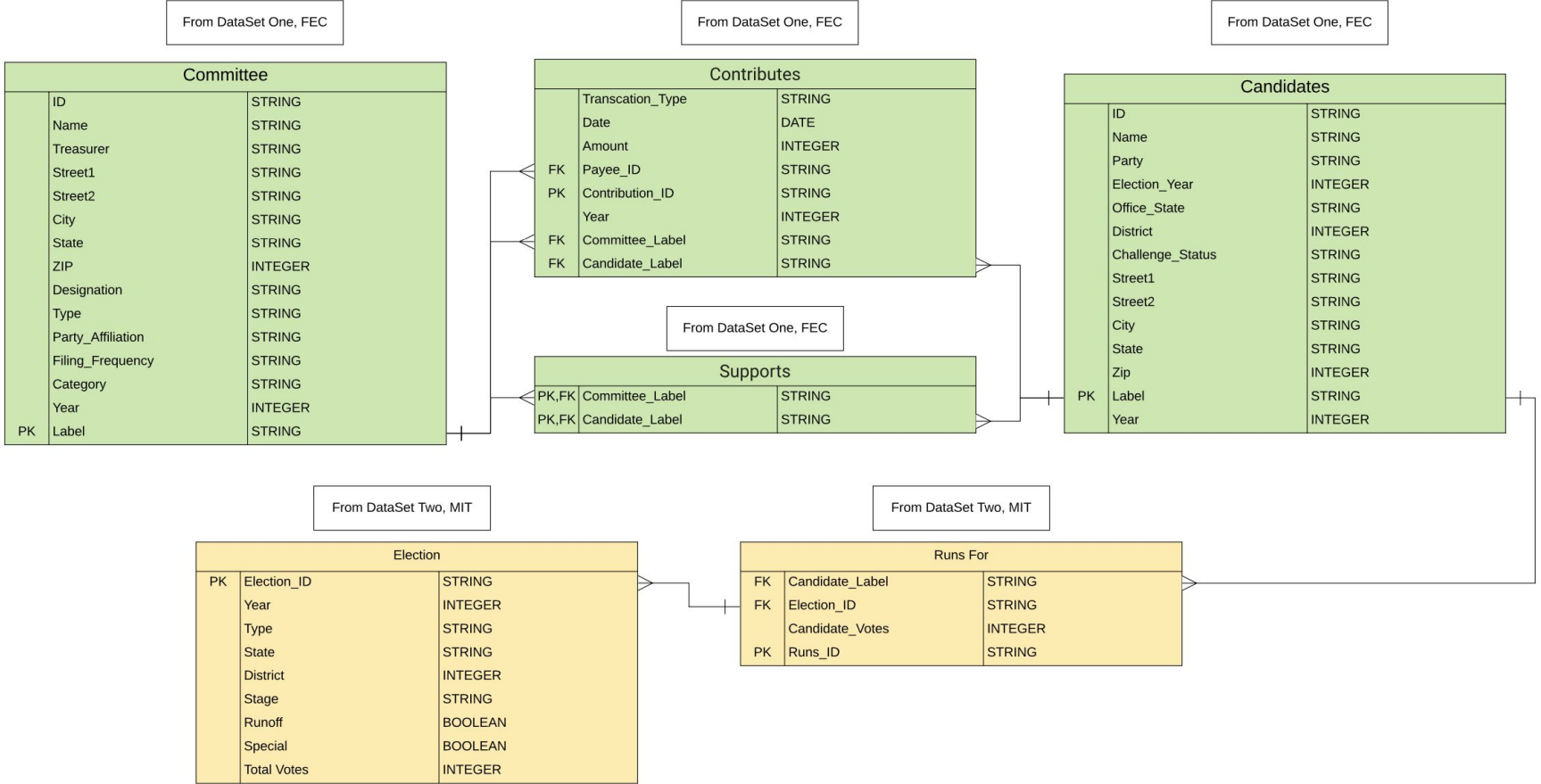
Candidates		
ID	STRING	
Name	STRING	
Party	STRING	
Election_Year	INTEGER	
Office_State	STRING	
District	INTEGER	
Challenge_Status	STRING	
Street1	STRING	
Street2	STRING	
City	STRING	
State	STRING	
Zip	INTEGER	
PK Label	STRING	
Year	INTEGER	

From DataSet Two, MIT

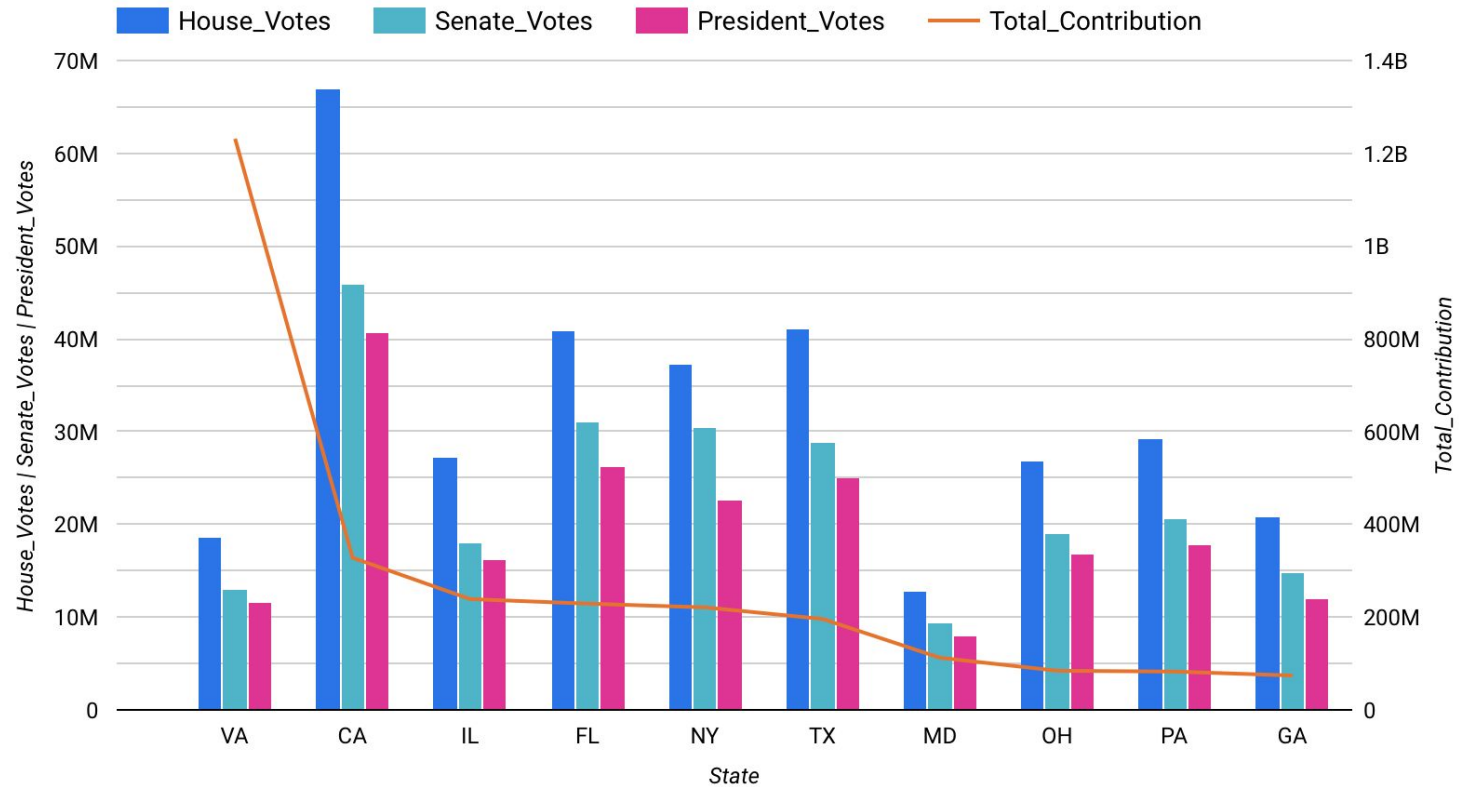
Election		
PK Election_ID	STRING	
Year	INTEGER	
Type	STRING	
State	STRING	
District	INTEGER	
Stage	STRING	
Runoff	BOOLEAN	
Special	BOOLEAN	
Total Votes	INTEGER	

From DataSet Two, MIT

Runs For		
FK Candidate_Label	STRING	
FK Election_ID	STRING	
Candidate_Votes	INTEGER	
PK Runs_ID	STRING	

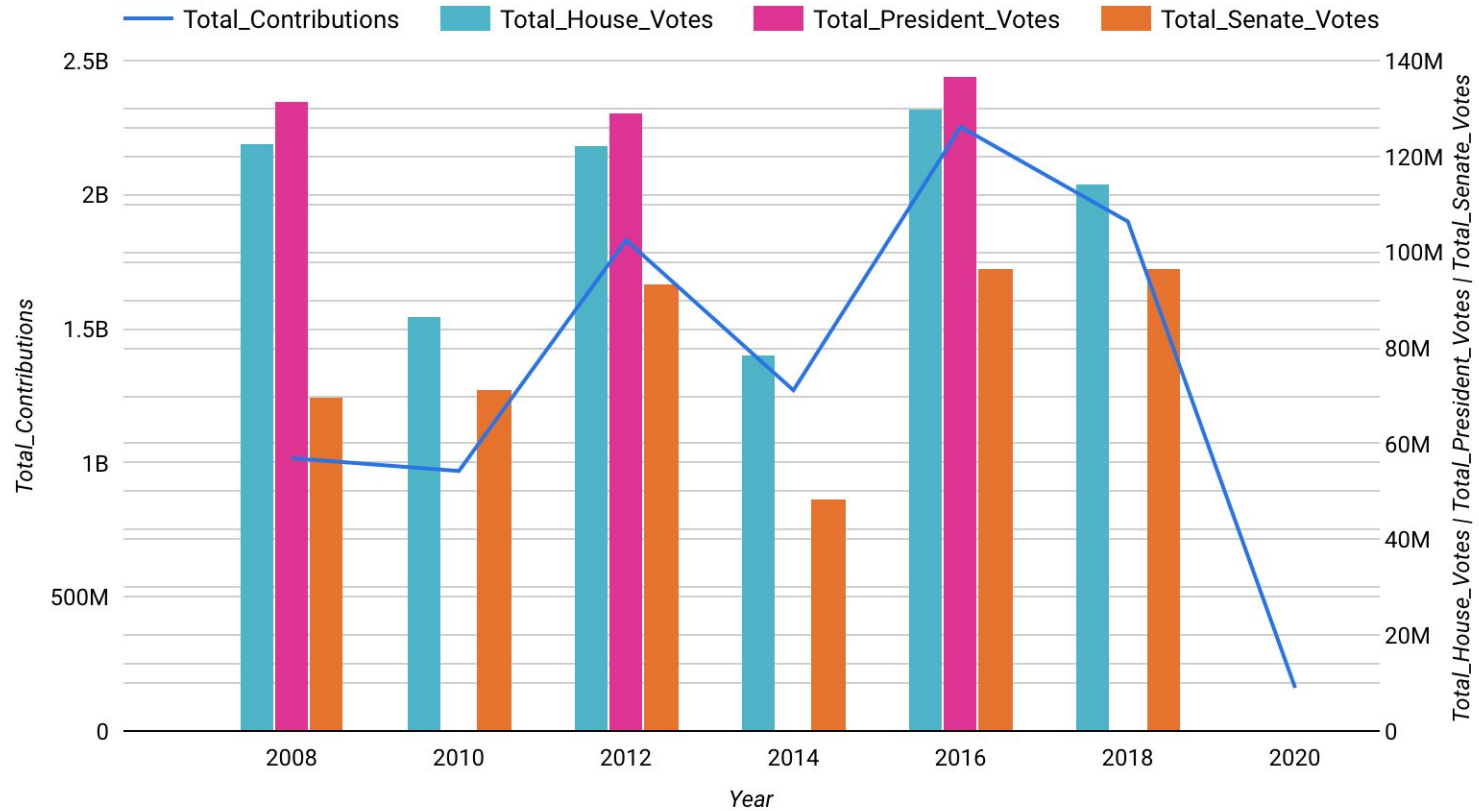


Total Contribution VS Total Amount of People Votes in Top 10 States



Total Amount of Contribution VS Total Amount of People Votes in House/Senate/Presidential Elections, **Group By State**

Total Contribution VS Total Amount of People Votes in Different Years



Total Amount of Contribution VS Total Amount of People Votes in House/Senate/Presidential Elections, **Group By Year**



Future Work

1. Include the total eligible voters/population in each state to find out the percentage of people vote in each state.
2. Include the overall economics data to explore its effects on the political contribution.
3. Investigate the effect of party affiliation to the political contribution amount.
4. Include individual Donors to have a comprehensive view of political campaign finance