

1. Requirements

Types of Datasets Required for Project

Contributions by Individuals

- Properties include:
 - Required: ID Number
 - Allow Null: Date, String (Name, State), Number (Amount)
- Name, State, and ID fields will be used in the table on the Explore page. All other fields (recipient type, donor name, amount) will be used in the Analytics page for the ranking table, data bar, and calculating number of contributions made by relevant individuals.

Contributions from Committees to Candidates

- Properties include:
 - Required: ID Number
 - Allow Null: Date, String (Name, State, Candidate ID), Number (Amount)
- All fields will be used in the Analytics page for the ranking table, data bar, and calculating number of contributions made by relevant individuals.

Candidate Information

- Properties include:
 - Required: Candidate identification
 - Allow Null: Candidate name, Party affiliation, Year of election, Candidate state, Candidate office, Candidate district, Incumbent challenger status, Principal campaign committee, Mailing address
- Candidate name, affiliation, year of election, and state will be used in the Results Table. State will also be used to filter candidates. All other fields are used for the Analytics and Competitor Analytics page.

Committee Information

- Properties include:
 - Required: Committee identification
 - Allow Null: Committee name, Treasurer's name, Mailing address, Committee designation, Committee type, Committee party, Interest group category, Connected organization's category, Candidate identification
- All fields will be used in the committee ranking modal within the Analytics page to show generic info of the selected. ID is used by the explore table.

Linking Candidates with Committees

- Properties include:

- Required: Candidate identification, Candidate election year, FEC election year, Linkage ID
 - Allow Null: Committee identification, Committee type, Committee designation
- All fields will be used in the Analytics page to show the candidates a committee has donated to. Will also be used in the ranking of the amount donated to candidates from a selected committee.

2. Design

Preprocessing data

- Handling Null Fields
 - For All Donors
 - Donors are defined as individuals, committees, or others that have donated to a candidate. In order to be in the dataset, some amount must have been donated. We can therefore include all rows of the dataset to count total contributions and calculate the number of receipts of the recipient.
 - Any rows that do not include state will not be included in the heatmap but remain as part of the table and total contribution amount
 - Null strings will be set to “-” in the table
 - For Candidates
 - Any candidate without a location or office type will be excluded from the query. This is because the explore page is dependent on both fields to display results.

No post-processing needed

3. Evaluation

All data downloaded from <https://www.fec.gov> approx 1 hour to collect

1.1 Contributions by individuals

Size: 25.89 GB

Fields

- Recipient
 - Filer identification number
 - Entity type
 - CAN = Candidate
 - CCM = Candidate Committee
 - COM = Committee
 - IND = Individual (a person)
 - ORG = Organization (not a committee and not a person)

- PAC = Political Action Committee
 - PTY = Party Organization
- Donor
 - Contributor/Lender/Transfer Name
 - State
- Transaction
 - Transaction amount
 - Transaction Primary-general indicator
 - Other identification number

1.2 Contributions from committees to candidates & independent expenditures

Size: 87.5 MB

Fields

- Recipient
 - Filer identification number
 - Entity type
 - CAN = Candidate
 - CCM = Candidate Committee
 - COM = Committee
 - IND = Individual (a person)
 - ORG = Organization (not a committee and not a person)
 - PAC = Political Action Committee
 - PTY = Party Organization
- Donor
 - Contributor/Lender/Transfer Name
 - State
- Transaction
 - Transaction amount
 - Other identification number
 - Candidate ID

1.3 Candidate Master

Size: 794 KB

Fields

- Candidate identification
- Candidate name
- Party affiliation
- Year of election
- Candidate state
- Candidate office
 - H = House
 - P = President
 - S = Senate

- Candidate district
- Incumbent challenger status
- Principal campaign committee
- Mailing address
 - street
 - street2
 - city
 - state
 - ZIP code

1.4 Committee Master

Size: 2.23 MB

Fields

- Committee identification
- Committee name
- Treasurer's name
- Mailing address
 - street
 - street2
 - city
 - state
 - ZIP code
- Committee designation
- Committee type
- Committee party
- Interest group category
- Connected organization's category
- Candidate identification

1.5 Candidate-committee linkage

Size: 304 KB

Fields

- Candidate identification
- Candidate election year
- FEC election year
- Committee identification
- Committee type
- Committee designation
- Linkage ID

4. Contributions

UCR-SECURE wifi