

ELECTION OUTCOMES

Analyzing votes, money spent, and PAC data

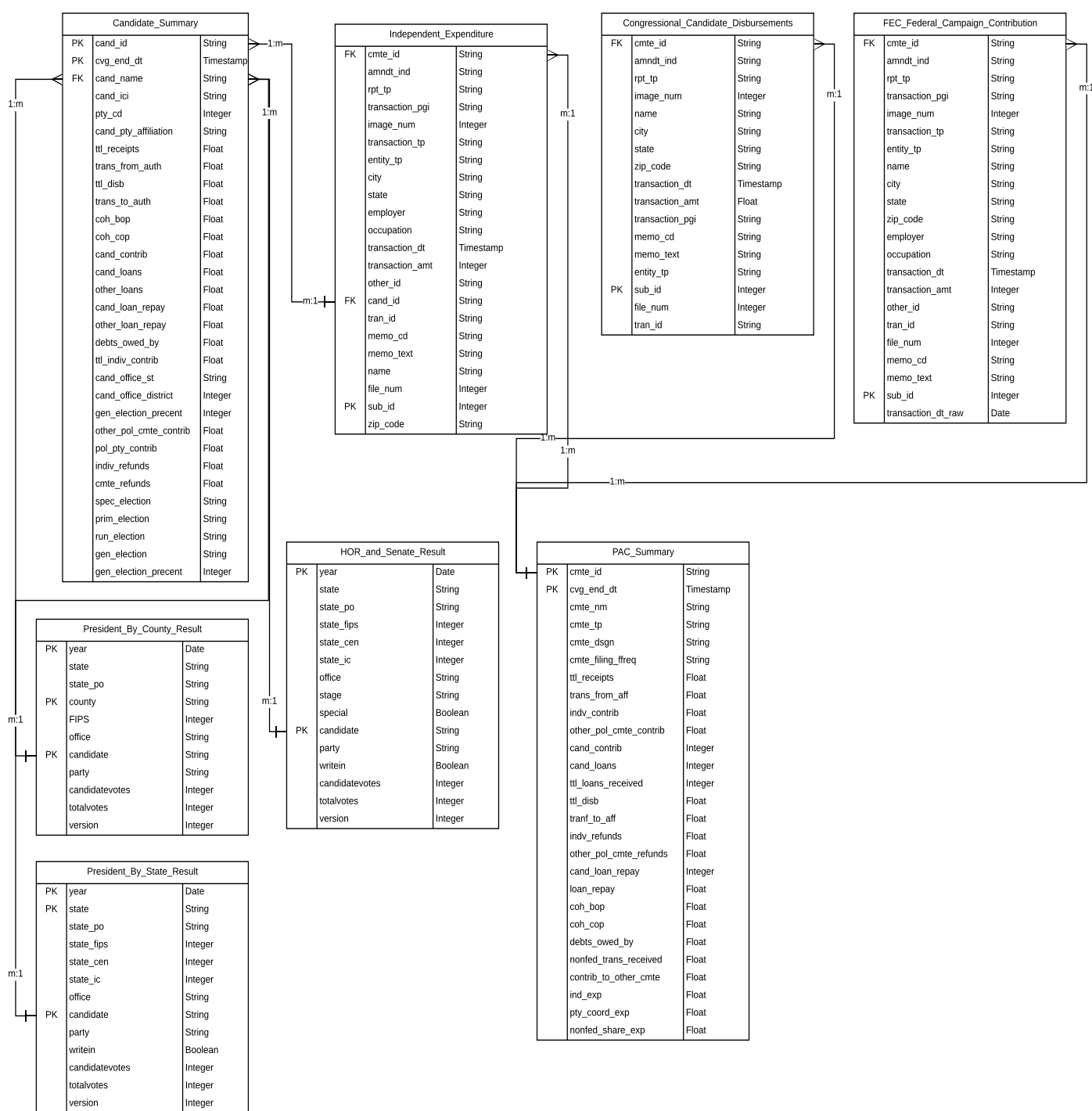
QUESTIONS

- Does money spent have an impact on votes received?
- Which party's PACs spend more money?
- Which campaigns spent the most money?

DATASETS

- FEC Information from Enigma Public
 - Main dataset
- Election Outcomes for President, Senate, and House of Representatives from Harvard Dataverse
 - Secondary dataset

ERD



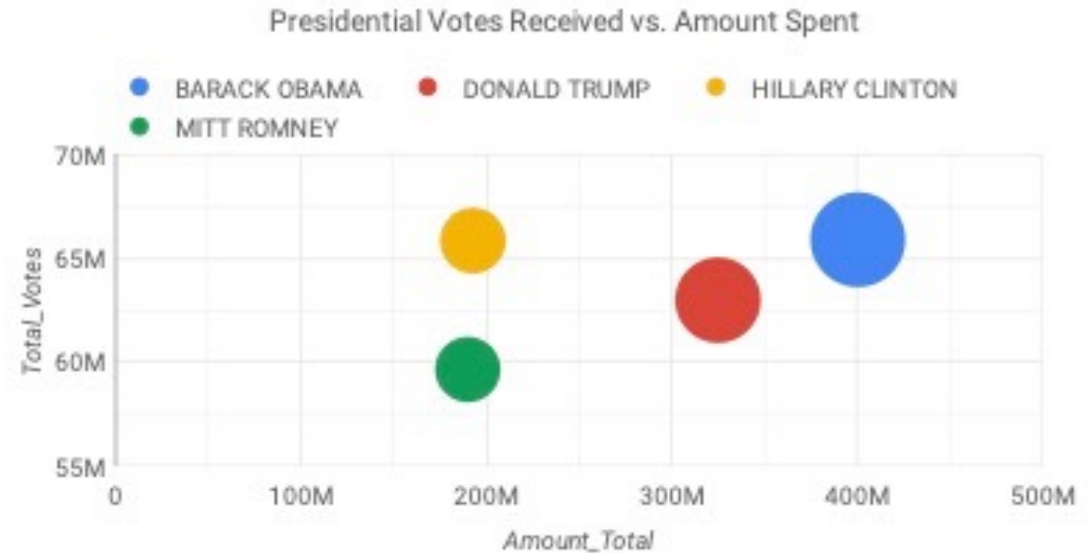
TRANSFORMS

- Primary dataset
 - Combined four years of tables from the original dataset to create FEC_Information dataset
 - Cleaned 'cand_name' to fit our scheme
 - Changed 'cvg_end_dt' from Integer to Date
 - Removed unnecessary columns
- Secondary dataset
 - Cleaned "candidate" column of all tables
 - Combined House_Of_Rep_Election_Result table with Senate_Election_Result table to form HOR_and_Senate_Result table
 - Changed datatype of 'year' column from an Integer to a Date for all tables
 - Dropped unnecessary column

AIRFLOW

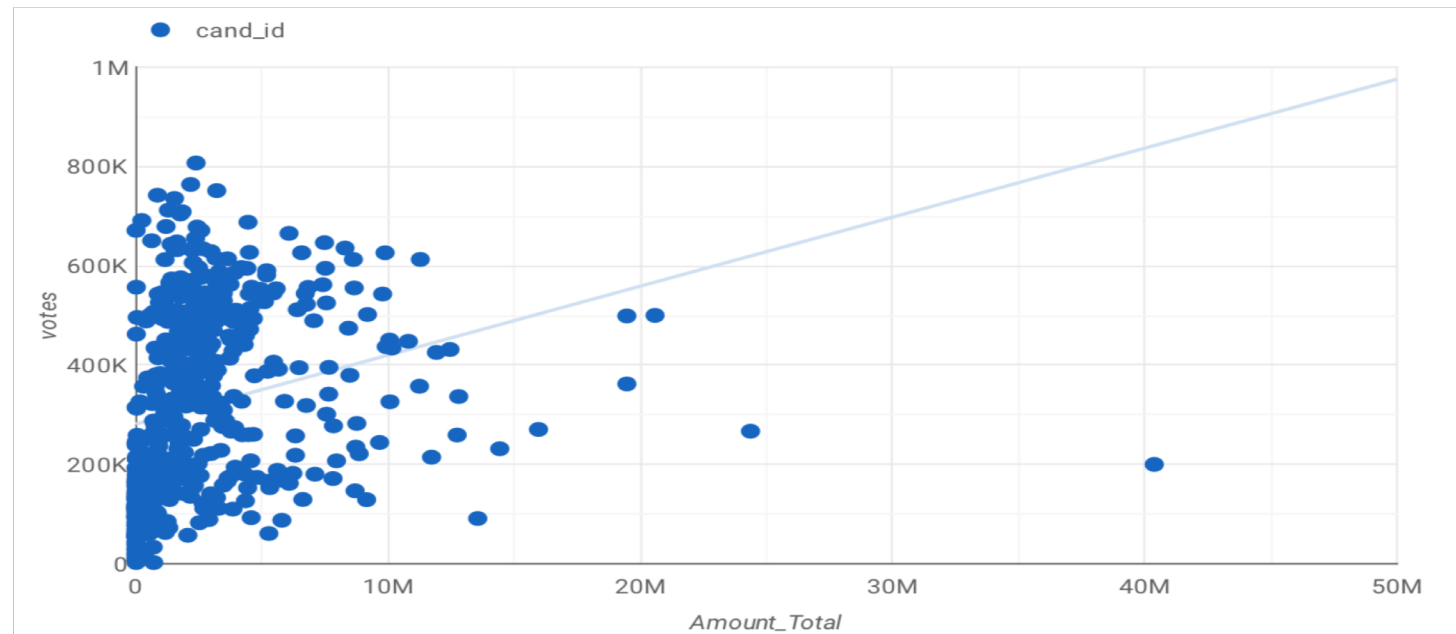
- <https://github.com/cs327e-spring2019/Not-Sure/edit/master/workflow.py>

PRESIDENTIAL VOTES RESULTS



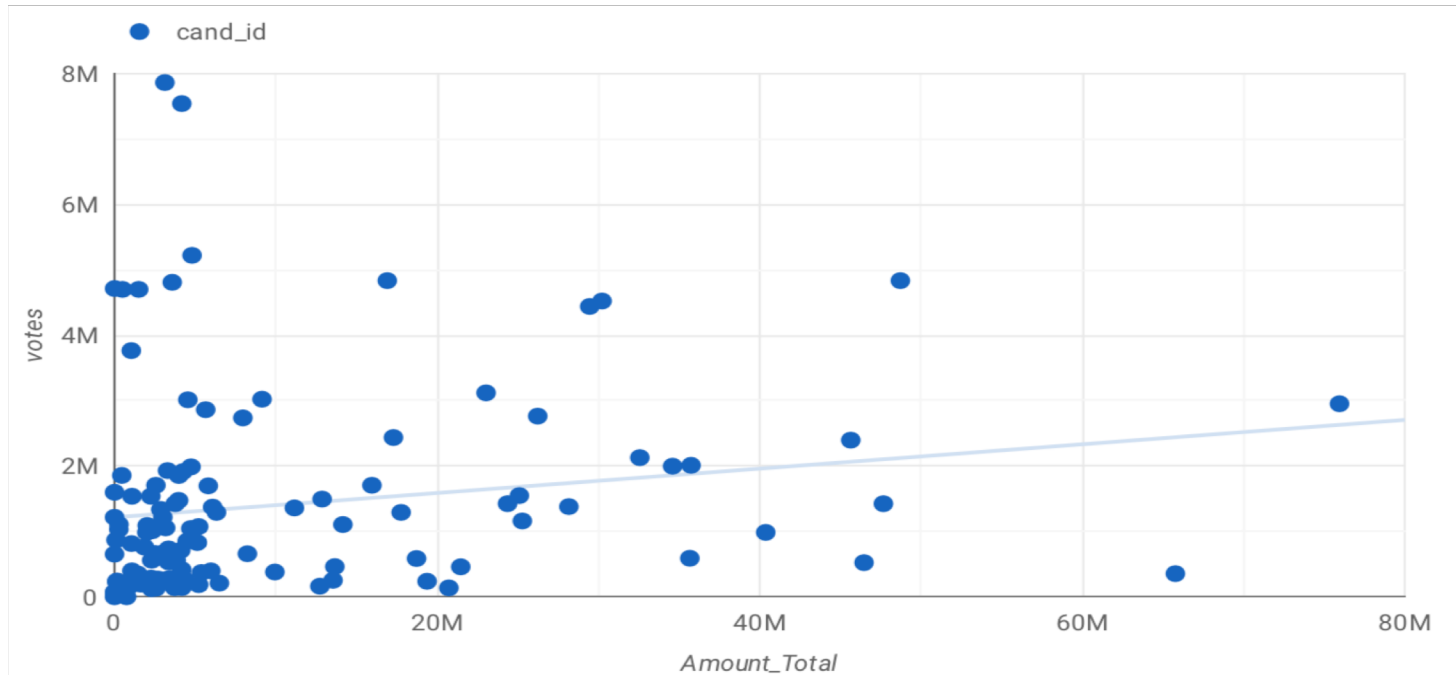
```
SELECT ie.cand_id, cs.cand_name, v.votes,
SUM(transaction_amt) AS Amount_Total
FROM (`not-sure-230422`.FEC_Information.Independent_Expenditure ie
JOIN `not-sure-230422`.FEC_Information.Candidate_Summary_Cleaned cs ON ie.cand_id =
cs.cand_id )
JOIN (Select SUM(r.candidatevotes) as votes, r.candidate as name0 from `not-sure-
230422`.Election_Outcome_Clean.President_By_State_Result`
r join (select distinct cand_name from `not-sure-
230422`.FEC_Information.Candidate_Summary_Cleaned` cs where date(cvg_end_dt) > '2018-
01-01') on r.candidate = cand_name where candidate = cand_name
group by r.candidate) v on cs.cand_name = name0
WHERE date(cvg_end_dt) > '2018-01-01'
GROUP BY ie.cand_id, cs.cand_name, v.votes
ORDER BY Amount_Total desc
limit 4
```

HOUSE OF REPRESENTATIVES VOTES RESULTS



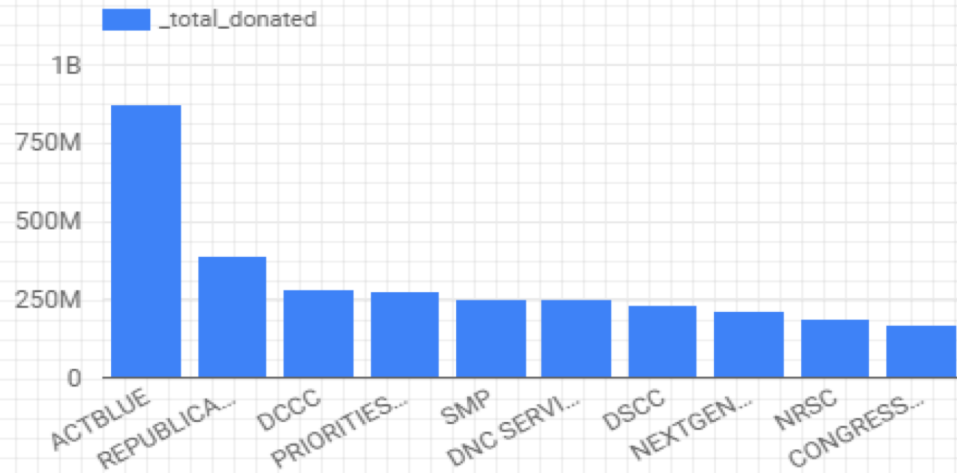
```
SELECT ie.cand_id, v.office, cs.cand_name, v.votes, SUM(transaction_amt) AS Amount_Total
FROM (`not-sure-230422`.FEC_Information.Independent_Expenditure ie
JOIN `not-sure-230422`.FEC_Information.Candidate_Summary_Cleaned cs ON ie.cand_id =
cs.cand_id )
JOIN (Select SUM(r.candidatevotes) as votes, r.candidate as name0, r.office FROM `not-sure-
230422`.`Election_Outcome_Clean.HOR_and_Senate_Result` r
JOIN (select distinct cand_name from `not-sure-
230422`.`FEC_Information.Candidate_Summary_Cleaned` cs WHERE date(cvg_end_dt) >
'2018-01-01') on r.candidate = cand_name WHERE office = "US House"
GROUP BY r.candidate, r.office) v on cs.cand_name = name0
WHERE date(cvg_end_dt) > '2018-01-01'
GROUP BY ie.cand_id, cs.cand_name, v.votes, v.office
ORDER BY Amount_Total desc
LIMIT 10
```


SENATE VOTES RESULTS

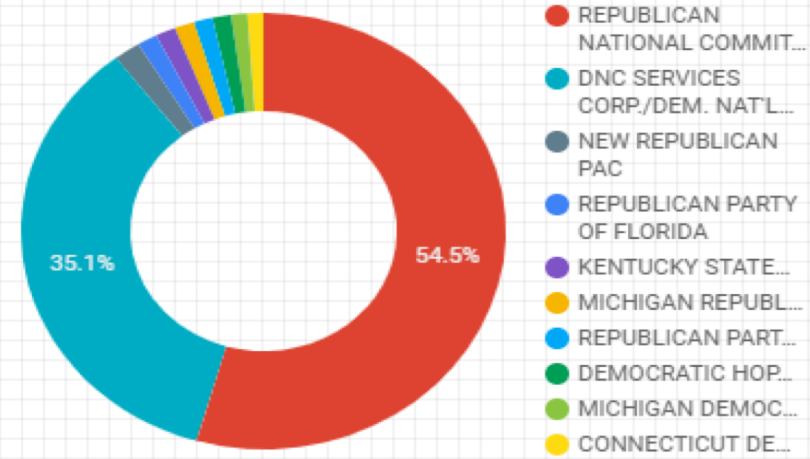


```
SELECT ie.cand_id, v.office, cs.cand_name, v.votes, SUM(transaction_amt) AS Amount_Total
FROM (`not-sure-230422`.FEC_Information.Independent_Expenditure ie
JOIN `not-sure-230422`.FEC_Information.Candidate_Summary_Cleaned cs ON ie.cand_id =
cs.cand_id )
JOIN (Select SUM(r.candidatevotes) as votes, r.candidate as name0, r.office FROM `not-sure-
230422`.`Election_Outcome_Clean.HOR_and_Senate_Result` r
JOIN (select distinct cand_name from `not-sure-
230422`.`FEC_Information.Candidate_Summary_Cleaned` cs WHERE date(cvg_end_dt) > '2018-
01-01') on r.candidate = cand_name WHERE office = "US Senate"
GROUP BY r.candidate, r.office) v on cs.cand_name = name0
WHERE date(cvg_end_dt) > '2018-01-01'
GROUP BY ie.cand_id, cs.cand_name, v.votes, v.office
ORDER BY Amount_Total desc
LIMIT 10
```

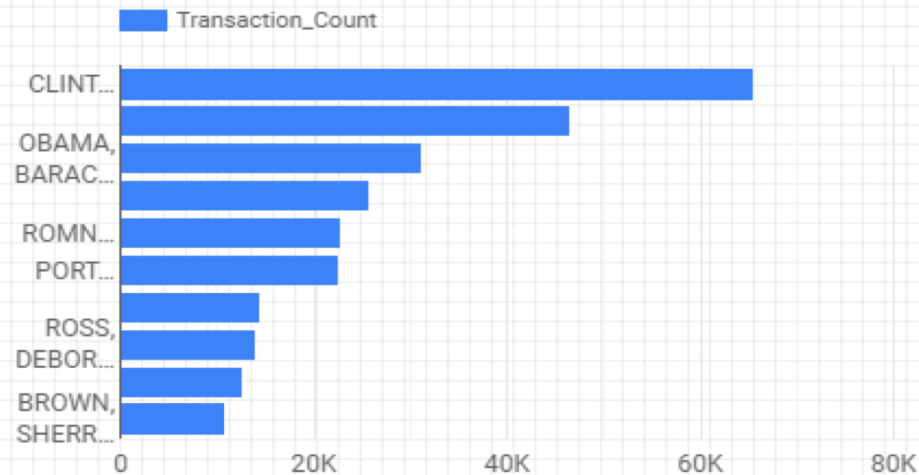
How much was donated to the top PAC's?



Distribution of money between PAC's associated with Dem or Rep



Top 10 Candidate/Campaign Transactions Count



Top 10 Candidate/Campaign Transaction Totals

