Ebanezare Tadele

INFM 600

Assignment: SQL Queries

October 19, 2018

Opportunities for support:

1. **Identifying candidate donors**
2. **Determine influential people in the political world**
3. **Identifying candidate donors:**

Phase One:

* Identify Areas of Improvement
* Benefit Provided from Improvement
* Key Areas of Information from Data

I want to identify all the donors who donated the highest amount in a single donation. By identifying presumable wealthy donors, we can tap into their network and increase donations from those in their network who has not donated by incentivizing already top donors to network on our behalf. Grass-root donors take much more man power, time and resources to require donations from without the return on investment. By targeting big donors, we can strategize on how to cater to their vote and donations. Doing so we can raise much more with less donors. I will be looking to filter Maryland Contributors by donation amount. Once I identify how much is the highest amount of donations, I will filter donors by donation amount from the state of Maryland. Now that I have identified Maryland donors who have donate the most, I can focus my attention on them and their networks.

Phase Two:

* Queries used to accomplish goals.
* Expectation from queries searched.
* Data found using the queries.
* Potential issue with the data.

First Query:

**SELECT**

CMTE\_ID,

NAME,

STATE,

EMPLOYER,

OCCUPATION,

TRANSACTION\_DT,

TRANSACTION\_AMT,

TRAN\_ID

**FROM**

marylandcontributions

**WHERE**

STATE = 'MD'

**ORDER** **BY** TRANSACTION\_AMT **DESC**;

Second Query:

**SELECT**

COUNT(CMTE\_ID)

**FROM**

marylandcontributions

**WHERE**

STATE = 'MD'

**AND** TRANSACTION\_AMT **>=** '5000';

Third Query:

**SELECT**

**AVG**(TRANSACTION\_AMT)

**FROM**

marylandcontributions

**WHERE**

STATE = 'MD'

**AND** TRANSACTION\_AMT **>=** '5000';

Fourth Query:

**SELECT**

CMTE\_ID,

NAME,

STATE,

EMPLOYER,

OCCUPATION,

TRANSACTION\_DT,

TRANSACTION\_AMT,

TRAN\_ID

**FROM**

marylandcontributions

**WHERE**

STATE = 'MD'

**AND** TRANSACTION\_AMT **>=** '5000'

**ORDER** **BY** TRANSACTION\_AMT **DESC**;

What I did not expect was the high “TRANSACTION\_AMT” from donors. But once I read through the type of professions and the companies they own or work for, I was less surprised since obviously those individuals must be making a lot of money already. I did expect high level professionals since my query was for donations more than $5,000.

I found that there are 438 Maryland donors who donated more $5,000. The average donation was around $11,000 and the highest donation being $300,000.

A potential issue with the data that I found is that not all donors have their employer or occupation listed. That will make it difficult to know what time of professional influence those donors have without that information. Another issue is that there are some rows that show donors with a negative “TRANSACTION\_AMT” which is unclear why or how that could be the case.

1. **Determine influential people in the political world.**

Phase One:

* Identify Areas of Improvement
* Benefit Provided from Improvement
* Key Areas of Information from Data

I want to identify influential people in the political world by the donation amounts from different committees in Maryland. By identifying committees with high donor transactions from Maryland, we can partner and strategize with them on ways to leverage their networks and learn from their best practices. Also, we can identify the top potential candidates who are best resourced which would help with strategic planning. The benefit provided is that we don’t have to waste time recreating the wheel, maximizes our efforts in impact, as well as create strategic planning around high funded candidates. Once I identify the range of donations by committees, I will filter committees by donation amount from the state of Maryland. Now that I have identified Maryland committees who have donated the most, I can focus my attention on them and their networks.

Phase Two:

* Show the queries that you will use to accomplish these goals.
* Explain what you expect your queries to provide,
* Show the data (or at the very least, a sample of the data) that you found using your queries.
* For 1 extra point: find and describe a potential issue with the data.

First Query:

**SELECT**

CMTE\_ID,

TRANSACTION\_TP,

NAME,

STATE,

EMPLOYER,

OCCUPATION,

TRANSACTION\_AMT,

CAND\_ID

**FROM**

committee2candidate

**WHERE**

State = 'MD'

**ORDER** **BY** TRANSACTION\_AMT **DESC**;

Second Query:

**SELECT**

**COUNT**(CMTE\_ID)

**FROM**

committee2candidate

**WHERE**

STATE = 'MD'

**AND** TRANSACTION\_AMT **>=** '5000';

Third Query:

**SELECT**

CMTE\_ID,

TRANSACTION\_TP,

NAME,

STATE,

EMPLOYER,

OCCUPATION,

TRANSACTION\_AMT,

CAND\_ID

**FROM**

committee2candidate

**WHERE**

State = 'MD'

**AND** TRANSACTION\_AMT **>=** '5000'

**ORDER** **BY** TRANSACTION\_AMT **DESC**;

I had expected a lot more committees who’s “TRANSACTION\_AMT” would have been more than the $5,000 threshold I set in my searches as well as the number of committees who contributed at least $5,000. Since people donate to committees who donate to candidates, the assumption is that those committees would donate much more than any one contributor could and or would.

The total committees who did contributed $5,000 was only 8. One committee was for candidates running for Senate, five for Congress and the rest did not specify. The top committees are: 'ANDY HARRIS FOR CONGRESS', 'CUMMINGS FOR CONGRESS CAMPAIGN COMMITTEE', 'DONNA EDWARDS FOR SENATE', 'DUTCH RUPPERSBERGER FOR CONGRESS', 'FRIENDS OF JOHN DELANEY’, ‘HOYER FOR CONGRESS’, ‘MIKULSKI FOR SENATE COMMITTEE', and 'VAN HOLLEN FOR CONGRESS'.

An issue with the data that I had found was being able to relate the committee with the candidate to filter out their information by political affiliation. That made it not possible to identify which committees represent candidates from certain affiliations. Again, there was an issue that some rows that show committees with a negative “TRANSACTION\_AMT” which is unclear why or how that could be the case.