**INFM 600 – ASSIGNMENT 2 – ELECTION DATASET ANALYSIS**

**Dataset** – **ELECTION DATASET**

**Source**

Sample Election Dataset provided in class for the analysis.

**Description of Dataset**

The data represent transactions between individuals and political committees and between political committees and candidates through the elections. This dataset presents relationships among individuals, the committees, and the candidates they support.

**Candidate Selection**

I assume I have been hired by the candidate Nunnelee, Patrick Alan with candidate ID – “H0MS01043” who stood for elections in year 2006. The candidate is from the city Tupelo and state MS.

**Areas of Campaign Selection**

The following areas would be of interest for analysis and improvements:

1. Identifying candidate donors.
2. Areas to allocate resources for campaigning.
3. **Identifying Candidate Donors**

Nunnelee, Patrick Alan is affiliated with committee “NUNNELEE FOR CONGRESS” with two committee ids “C00114132” and “C00030809”. We find this using the below SQL Command:

SELECT \* FROM election.committee2candidate

Where CAND\_ID = "H0MS01043";

The party with the above committee ids, is associated and receives funds with a list of individuals as mentioned in Table 1 with the SQL command used to generate the result:

SELECT CMTE\_ID, NAME, EMPLOYER, OCCUPATION, TRAN\_ID, TRANSACTION\_AMT, TRANSACTION\_DT FROM election.marylandcontributions

Where CMTE\_ID = ‘C00114132’ or CMTE\_ID = ‘C00030809’;

or

SELECT CMTE\_ID, NAME, EMPLOYER, OCCUPATION, TRAN\_ID, TRANSACTION\_AMT, TRANSACTION\_DT FROM election.marylandcontributions

WHERE CMTE\_ID in (SELECT CMTE\_ID FROM election.committee2candidate

WHERE CAND\_ID in ( SELECT CAND\_ID from election.candidate

WHERE CAND\_NAME = "Nunnelee, Patrick Alan"));

**Possible Improvements, Benefits & Information**

We observe from the output of the SQL query that there is a list of individuals from similar backgrounds and designations who provide funding support to the committees. These individuals, coming from the Pharmacy segment and being Veterinarian or owner/manager, have raised the funds in the year of 2015 for the election year of 2016 which is quite appropriate. The amount of funds raise by them can be correlated and analyzed for determination of this segment of people are capable of larger funding than others. Since the entity type “IND” and Transaction\_PGI is “P” for all individuals, considering these fields in the results would lead to no improvisation points, we do not bring them into consideration. A careful scrutiny of the result help us understand the below mentioned benefits and improvement points:

* There are only 2 individuals rendering Transaction amount more than $250 ($365 & $500) and the occupation of these individuals are owner/manager of their employee in Pharmacy Segment. Using the Query below, we see that there are a list of individuals capable of funding huge amounts starting from $30000 down to $1000 which belong to various multinational organizations with designations of Senior Vice President, Chairman, Businessman etc. The candidate and committee must do efforts and bring in donors and investors from such backgrounds which would help accumulate larger funds.

SELECT CMTE\_ID, NAME, EMPLOYER, OCCUPATION, TRAN\_ID, TRANSACTION\_AMT, TRANSACTION\_DT

FROM election.marylandcontributions

ORDER BY TRANSACTION\_AMT DESC;

* The other transaction amounts are either 100 or 250. These individuals should be motivated to raise party funds or certain schemes shall be in place which are long term beneficiary for the individuals to invest more in the hope of better returns in future.
* Since Pharmacy sector is prevalent in the list of donors for this committee, the candidate must have good relations in the pharmacy industry, and hence more contributors in the segment can be looked upon to match similar interests and convince them for raising funds towards this committee. This could be found using the below SQL command:

SELECT \* FROM election.marylandcontributions

where EMPLOYER LIKE '%pharmacy%';

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **CMTE\_ID** | **NAME** | **EMPLOYER** | **OCCUPATION** | **TRAN\_ID** | **TRANSACTION\_AMT** | **TRANSACTION\_DT** |
| **C00114132** | BROOKS, JOHN RUSSELL DR | SELF EMPLOYED | VETERINARIAN | 37855109 | 250 | 1152015 |
| **C00114132** | BRUNT, JANE DR | CAT HOSPITAL AT TOWSON | VETERINARIAN | 37855065 | 250 | 1162015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 20150406103818-208 | 100 | 3242015 |
| **C00030809** | LEIKACH, NEIL | CATONSVILLE PHARMACY LLC | OWNER/MANAGER | 20150406103818-209 | 100 | 3242015 |
| **C00030809** | FULTON, DAVID JR. | WHITESELLS PHARMACY | OWNER/MANAGER | 20150406103818-118 | 100 | 3242015 |
| **C00030809** | MINTZ, MARTIN B. | NORTHERN PCHY AND MED EQUIPMENT | OWNER/MANAGER | 20150406103818-245 | 100 | 3242015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 20150406103818-162 | 100 | 3242015 |
| **C00030809** | LEIKACH, NEIL | CATONSVILLE PHARMACY LLC | OWNER/MANAGER | 20150506145317-137 | 100 | 4222015 |
| **C00030809** | FULTON, DAVID JR. | WHITESELLS PHARMACY | OWNER/MANAGER | 20150506145317-76 | 100 | 4222015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 20150506145317-111 | 100 | 4222015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 20150506145317-136 | 100 | 4222015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 20150605132434-129 | 100 | 5202015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 20150605132434-158 | 100 | 5202015 |
| **C00030809** | LEIKACH, NEIL | CATONSVILLE PHARMACY LLC | OWNER/MANAGER | 20150605132434-159 | 100 | 5202015 |
| **C00030809** | FULTON, DAVID JR. | WHITESELLS PHARMACY | OWNER/MANAGER | 20150605132434-93 | 100 | 5202015 |
| **C00030809** | JOHNSON, JULIE S. | JOHNSON FAMILY PHARMACY | OWNER/MANAGER | 20150605132434-136 | 365 | 5282015 |
| **C00030809** | LEIKACH, NEIL | CATONSVILLE PHARMACY LLC | OWNER/MANAGER | 2015070885316-153 | 100 | 6242015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 2015070885316-152 | 100 | 6242015 |
| **C00030809** | MINTZ, MARTIN B. | NORTHERN PCHY AND MED EQUIPMENT | OWNER/MANAGER | 2015070885316-180 | 100 | 6242015 |
| **C00030809** | FULTON, DAVID JR. | WHITESELLS PHARMACY | OWNER/MANAGER | 2015070885316-85 | 100 | 6242015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 2015070885316-122 | 100 | 6242015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 20150805153945-135 | 100 | 7212015 |
| **C00030809** | LEIKACH, NEIL | CATONSVILLE PHARMACY LLC | OWNER/MANAGER | 20150805153945-136 | 100 | 7212015 |
| **C00030809** | MINTZ, MARTIN B. | NORTHERN PCHY AND MED EQUIPMENT | OWNER/MANAGER | 20150805153945-160 | 100 | 7212015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 20150805153945-109 | 100 | 7212015 |
| **C00030809** | BALCH, JOHN H. | PHARMACARE OF CUMBERLAND | OWNER/MANAGER | 20150805153945-14 | 500 | 7142015 |
| **C00030809** | MINTZ, MARTIN B. | NORTHERN PCHY AND MED EQUIPMENT | OWNER/MANAGER | 2015090885416-159 | 100 | 8202015 |
| **C00030809** | LEIKACH, NEIL | CATONSVILLE PHARMACY LLC | OWNER/MANAGER | 2015090885416-137 | 100 | 8202015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 2015090885416-136 | 100 | 8202015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 2015090885416-112 | 100 | 8202015 |
| **C00030809** | MINTZ, MARTIN B. | NORTHERN PCHY AND MED EQUIPMENT | OWNER/MANAGER | 20151005112319-148 | 100 | 9232015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 20151005112319-126 | 100 | 9232015 |
| **C00030809** | LEIKACH, NEIL | CATONSVILLE PHARMACY LLC | OWNER/MANAGER | 20151005112319-127 | 100 | 9232015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 20151005112319-103 | 100 | 9232015 |
| **C00114132** | BROOKS, JOHN RUSSELL DR |  | VETERINARIAN | 38421829 | 250 | 7272015 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG PHARMACY INC | OWNER/MANAGER | 20160104152321-117 | 100 | 12222015 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG PHARMACY | OWNER/MANAGER | 20160104152321-96 | 100 | 12222015 |

**Table 1: List of Candidate Donors for the candidate Nunnelee, Patrick Alan’s committee NUNNELEE FOR CONGRESS**

1. **Areas to allocate resources for campaigning**

Firstly, we determine the locations these committees for the candidate is based out of. This would help us understand the geography and potential of the committee, candidate is associated with. This can be obtained using the below SQL command and results in Table 2:

SELECT CMTE\_ID, NAME, CITY, STATE, ZIP\_CODE,TRANSACTION\_AMT

FROM election.committee2candidate

Where CAND\_ID IN (SELECT CAND\_ID FROM election.candidate

WHERE CAND\_NAME = "Nunnelee, Patrick Alan");

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CMTE\_ID** | **NAME** | **CITY** | **STATE** | **ZIP\_CODE** | **TRANSACTION\_AMT** |
| **C00114132** | NUNNELEE FOR CONGRESS | TUPELO | MS | 38802 | 2000 |
| **C00030809** | NUNNELEE FOR CONGRESS | TUPELO | MS | 38802 | -1500 |

**Table 2: Location of the Committee**

Next, we would determine the locations of the individuals supporting the committee through their funds to understand the dynamics of the individuals and determine potential locations to increase the fund raisers or do less campaigning in regions where the outcomes are relatively low. This can be determined by using the below SQL query and results in Table 3.

SELECT CMTE\_ID, NAME, CITY, STATE, ZIP\_CODE, TRANSACTION\_AMT

FROM election.marylandcontributions

WHERE CMTE\_ID in (SELECT CMTE\_ID FROM election.committee2candidate

WHERE CAND\_ID in ( SELECT CAND\_ID from election.candidate

WHERE CAND\_NAME = "Nunnelee, Patrick Alan"))

ORDER BY CITY ASC;

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CMTE\_ID** | **NAME** | **CITY** | **STATE** | **ZIP\_CODE** | **TRANSACTION\_AMT** |
| **C00030809** | JOHNSON, JULIE S. | ABERDEEN | MD | 21001 | 365 |
| **C00114132** | BRUNT, JANE DR | ANNAPOLIS | MD | 214032616 | 250 |
| **C00030809** | MINTZ, MARTIN B. | BALTIMORE | MD | 21234 | 100 |
| **C00030809** | LEIKACH, NEIL | BALTIMORE | MD | 212282305 | 100 |
| **C00030809** | MINTZ, MARTIN B. | BALTIMORE | MD | 21234 | 100 |
| **C00030809** | LEIKACH, NEIL | BALTIMORE | MD | 212282305 | 100 |
| **C00030809** | LEIKACH, NEIL | BALTIMORE | MD | 212282305 | 100 |
| **C00030809** | LEIKACH, NEIL | BALTIMORE | MD | 212282305 | 100 |
| **C00030809** | LEIKACH, NEIL | BALTIMORE | MD | 212282305 | 100 |
| **C00030809** | MINTZ, MARTIN B. | BALTIMORE | MD | 21234 | 100 |
| **C00030809** | LEIKACH, NEIL | BALTIMORE | MD | 212282305 | 100 |
| **C00030809** | MINTZ, MARTIN B. | BALTIMORE | MD | 21234 | 100 |
| **C00030809** | LEIKACH, NEIL | BALTIMORE | MD | 212282305 | 100 |
| **C00030809** | MINTZ, MARTIN B. | BALTIMORE | MD | 21234 | 100 |
| **C00030809** | BALCH, JOHN H. | CUMBERLAND | MD | 215022565 | 500 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | LEIKACH, DEANNA D. | FINKSBURG | MD | 210481633 | 100 |
| **C00030809** | FULTON, DAVID JR. | FREDERICK | MD | 21701 | 100 |
| **C00030809** | FULTON, DAVID JR. | FREDERICK | MD | 217015335 | 100 |
| **C00030809** | FULTON, DAVID JR. | FREDERICK | MD | 21701 | 100 |
| **C00030809** | FULTON, DAVID JR. | FREDERICK | MD | 21701 | 100 |
| **C00114132** | BROOKS, JOHN RUSSELL DR | KINGSVILLE | MD | 210870030 | 250 |
| **C00114132** | BROOKS, JOHN RUSSELL DR | KINGSVILLE | MD | 210870030 | 250 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |
| **C00030809** | HOSE, BRIAN M. | SHARPSBURG | MD | 217821626 | 100 |

**Table 3: List of Candidate Donors for the committee and their location details.**

**Possible Improvements, Benefits & Information**

Observations through the SQL queries are that the committee office is based out of Tupelo city in the state of MS, however, all the individuals supporting the committee are from the various cities from Maryland, per the data under study. This is found out using Order By function in the above query and Table 3. The individuals are from the cities Aberdeen, Annapolis, Baltimore, Cumberland, Finksburg, Frederick, Kingsville, Sharpsburg. There are other parameters which are there in the dataset but would not help in determining any help in location determination, hence excluded from the analysis report. Analysis of the output above helps us understand certain improvement points and benefits as listed below:

* The data can be analyzed to understand that aforementioned locations, the committee has been performing campaigning and hence has been able to receive funds for the election year. However, the funds are low for many of the cities, however the number of individuals re in good number like in Baltimore, Finksburg and Sharpsburg. These cities must stand utmost importance and more powerful campaigning shall be performed here to maintain and get new fund raisers from this place for the committee.
* Locations like Aberdeen, Annapolis, Cumberland and Kingsville do have a very few fund raisers, these locations can be made strong by making join more individuals to fund the committee by improving their campaign strategies in these locations.
* We determine the list of cities which are having individuals of high donation capacities and these locations can be considered as a potential locations to initiate campaigning to build in new individual fund raisers added to the committee funding list as also to pull in more funds in for the next elections. This can be obtained using the below SQL command and the results sample in table4 is below:

SELECT Name, CITY, STATE, ZIP\_CODE, TRANSACTION\_AMT

FROM election.marylandcontributions

ORDER BY TRANSACTION\_AMT DESC;

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **CITY** | **STATE** | **ZIP\_CODE** | **TRANSACTION\_AMT** |
| **EPSTEIN, MARK** | POTOMAC | MD | 20854 | 300000 |
| **NATIONAL NURSES UNITED** | SILVER SPRING | MD | 20910 | 200000 |
| **NATIONAL NURSES UNITED** | SILVER SPRING | MD | 20910 | 127000 |
| **RETAIL SYSTEMS AND SERVICE, INC.** | POTOMAC | MD | 20854 | 120000 |
| **ST. JOHN, EDWARD MR.** | BALTIMORE | MD | 212442677 | 100000 |
| **MCMAHON, JOHN** | BETHESDA | MD | 208143019 | 100000 |
| **MCMAHON, JOHN** | BETHESDA | MD | 20814 | 80000 |
| **ST. JOHN, EDWARD** | BALTIMORE | MD | 21244 | 80000 |
| **DELANEY, JOHN K.** | POTOMAC | MD | 208544613 | 67500 |
| **WALLACE, HENRY SCOTT** | BETHESDA | MD | 208173900 | 66800 |
| **PLASTER, MARK LEROY DR.** | ANNAPOLIS | MD | 21401 | 50000 |
| **ROSWELL, MARGORIE** | BALTIMORE | MD | 21218 | 50000 |
| **ROSWELL, MARGORIE** | BALTIMORE | MD | 21218 | 50000 |
| **MOSIER, FRANK JR.** | SAINT MICHAELS | MD | 21663 | 36575 |
| **MEYERHOFF, ROBERT E** | PHOENIX | MD | 21131 | 33400 |
| **BAINUM, SANDRA Y** | FULTON | MD | 207592531 | 33400 |
| **BAINUM, STEWART JR** | FULTON | MD | 207592531 | 33400 |
| **GOULD, MARY T MRS** | LAUREL | MD | 207235716 | 33400 |
| **GAINES, WILLIAM LEE MR. JR.** | BALTIMORE | MD | 212121053 | 33400 |
| **YAZDANI, BABAK** | POTOMAC | MD | 20859 | 33400 |
| **ISLAM, FRANK** | POTOMAC | MD | 208544117 | 33400 |
| **MOSIER, FRANK MR. JR.** | SAINT MICHAELS | MD | 216632950 | 33400 |
| **….** | …. | …. | …. | …. |

**Table 4: List of Candidate Donors for all locations grouped by Transaction Amount in Descending Order.**

* From Table 4, it is imperative to understand the locations which are having donors of high capacity donations, to list a few: Potomac, Silver Spring, Bethesda, Saint Michaels, Fulton etc. These are high potential locations the committee should plan strategy to campaign, since it might allow the committee to receive high amounts of funds from these locations.

**Extra Point Information: Potential Issue with the Data**

The data is quite extensive and descriptive to provide all relations and information for the candidates, committees and individuals with all details. However, there are certain points over the data which rise to potential issues in usage and understanding which can be observed through the dataset as listed below:

* There are a few attributes in all the Tables which are abbreviated, and we do not have a codebook for clear understanding, for eg, CAND\_PTY\_AFFILIATION, CAND\_ICI, CAND\_STATUS, CMTE\_DSGN CMTE\_FILING\_FREQ AMNDT\_IND, RPT\_TP. These fields might have some high importance for data analysis, however, since we do not have an understanding on their usage, we cannot take them under consideration for conclusions.
* The values for zip codes are incorrect for a few entries, eg: '220418795', which is not a correct zip code and if the search is based out of zip code, we may lose some potential data points.
* The names of the individuals or candidates are in the formal of “Last Name, Middle Name First Name” which should ideally be in a split of the three in different columns as the data might vary on different search patterns and we may not get proper results if the search is based out on Names.
* There are NULL values in the data which might cause inaccurate results for the analysis. For eg, there are no employer or occupation details for few individuals in the table “marylandcontributors”.
* There are varied formats of transaction ids which have been presented and it is difficult to understand if they are correct as per the underlying scheme or the values have lost format and have turned into hexadecimal or encoded format. Eg, '10144052', '95E082A9A60B2EDCA97', '23-389', 'SB23.5076' etc.

Word Count - 1210