INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE

Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

Final Project Progress Report

Introduction:

  Our team created a database containing information about campaign finance expenditures. This refers to all money spent during a political campaign and those who donated to said campaign. With the creation of this database, we hope people will use it to understand the complexities of campaign finance expenditures and see how much money can influence a campaign. For example, you can find different trends within the dataset, correlating with campaign success or failure. This data also shows transparency and accountability, showing the public where funding is used. This can prevent misuse of funds, reducing the risk of corruption.
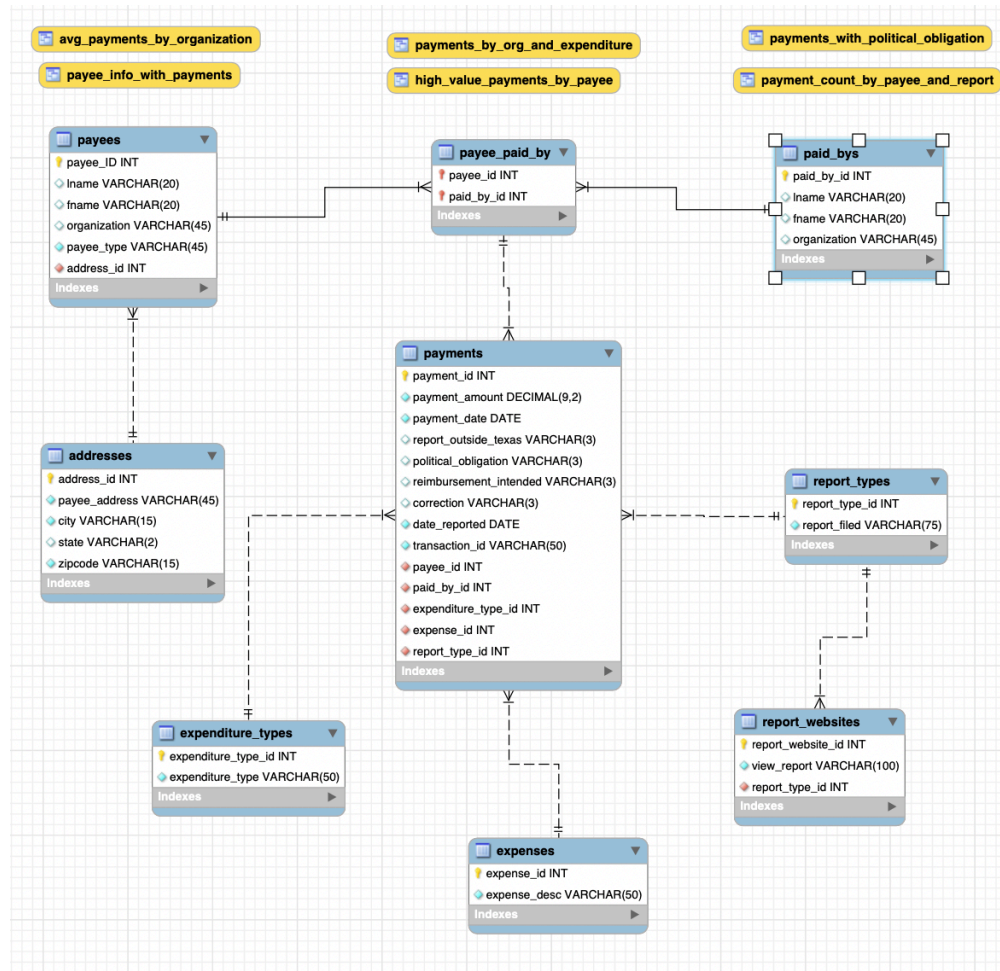
  We built our database based on the data from the official city of Austin's open data portal, with the latest year being 2022. We wanted to emphasize all of the campaigns' expenses and the information about the payees who helped fund them. Originally, the data we chose to use lacked scope and needed more realism. Still, as we moved toward finalizing the database we found how different attributes in our dataset related to each other and how some parts could be simplified through normalization. With this, our database can be easily navigated especially without all the data clutter.

Database Description:

  Our database is constructed around Campaign Finance Expenditures reported by the government of Austin, Texas. With this database, people will be able to analyze and observe how people fund and allocate funds for their campaigns. Our database consists of 9 tables, with our payments table being the main one because it has multiple one-to-many relationships.

Logical Design

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

Our goal was to create a database that was very clear and easy to understand, which would allow the users to understand the payees, paid bys, payments, expenditures, and reports. We also wanted to create database that could demonstrate relationships between each table and what they portray. To create this we went through many trials and tribulations and after many renditions, we centered our design around payments.

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

The payments table is connected to almost every single table, excluding the addresses and report_websites table. The payments table is connected to expenditure_type, expense, report_type, and the payee_paid_by table, which is a linking tables between payees and paid_bys.

We wanted the user to understand the connections between each table, so within our design, for each relationship there is a foreign key. Since majority of the tables in our design is connected to the payments table as a one-to-many relationship, these tables are exhibited as a foreign key within the payments table. This allows for users to easily understand where and what is being connected to the payments table. This structure can allow for other users to plug in different campaign finance datasets based in different cities and states, and would allow for them to break down this dataset in similar fashion.

Physical Database

Our database allows for people to understand where money comes from in campaigns, who they go to, how much money spent, date of expenses, reason for the expenses, when and the type of report of the expense, and where you can find the report. This database is seeking to inform people on how money is being used during a campaign cycle.

From this we created our main table, payments, which will store all the information about the payments from the campaigns. Connected to this table there is; expediture_types, storing the type of expenditures, expenses, which stores the different types of expenses, report_types, storing the information regarding the report types, and payee_paid_by, which stores information about the id of payees and paid bys. All four of these tables have a one-to-many relationship with payments. The table report_types has a one-to-many relationship with report_websites, which stored the websites that the reports can be stored in. The payee_paid_by table has a many-to-one relationship with both payees, which stored who is getting paid, and paid_bys, who is paying.

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

The payee_paid_by serves as a linking table between payees and paid_bys because payees and paid_bys have a many-to-many relationship. The payees table also has a one-to-many relationship with the addresses table, which stores the addresses of the payees.

Sample Data

The data that we used for our database was provided by the official City of Austin Data Portal, which was provided to us through the ELMS website. This data has been useful throughout our whole project, as it has helped us shape our database. Initially, our sample data lacked scope and needed more realism, so we realized we needed to base our sample data on actual contribution records with a focused scope. This scope includes; payee, payment amount, payment date, payee type, address, expenditure type, expense description, report date, report type, report website, and transaction id. Initially, this data was compacted into a single table, that was very hard to understand, so it was necessary to clean the data.

To reduce the amount of unnecessary data, we used thenormalization and went over our data tables to ensure logical and efficient design. This split the original one table, into nine separate tables. One example can be found in our "payees" table;

payee_ID	lname	fname	organization	payee_type	address_id
1	NULL	NULL	ActBlue	ENTITY	1
2	Aggarwal	Akshay	NULL	INDIVIDUAL	2
3	Alemao	Andrew	NULL	INDIVIDUAL	3
4	NULL	NULL	American Bank	ENTITY	4
5	NULL	NULL	Anedot	ENTITY	5
6	Aslam	Muneeb	NULL	INDIVIDUAL	6
7	NULL	NULL	Austin Chronicle	ENTITY	7
8	NULL	NULL	Austin Web Design	ENTITY	8

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

Views/ Queries:

Queries	Req 1 (Join)	Req 2 (Filter)	Req 3 (Aggregate)	Req 4 (Linking)	Req 5 (Sub-Query)
Query 1	X	X		X	
Query 2	X	X	X		X
Query 3	X	X	X		
Query 4	X	X	X		
Query 5	X	X	X		
Query 6	X	X	X		X

Description what each query:

Query 1: Creates a view that shows the payments made by an individual (paid by) to an organization (payee), and if the payment because of political obligation

Query 2: Creates a view that shows the amount of payments per report type for an individual payee.

Query 3: Creates a view that shows the total amount of payments for an individual payee, and the address information associated with the payees.

Query 4: Creates a view that shows the number of payments for a specific expense, for a organization (paid by) and the expenditure type.

Query 5: Creates a view that shows the average payment amount for an organization (paid by) within the first four months of the year.

Query 6: Creates a view that identifies payments that are above the average payment amount for each payee, specifically identifying individual payees.

Changes from the Original Design:

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

During our work on the project, we made several integral changes to improve our database design and structure. We had many redundant tables and columns that made the structure way more complicated than it needed to be, we removed these redundancies and normalized the database, simplifying our tables and relationships. We also kept expanding the project slightly by adding more refined and detailed information to more of the sections, which required us to re-adjust our ERD, queries, views, etc. These changes helped us create a more efficient, transparent, and user-friendly database.

Database Ethics Considerations:

When creating a database about campaign finance expenditures, we had multiple ethical considerations to ensure this data isn't used to take advantage of anyone. The first thing we wanted to emphasize was transparency. By using an official government source and not altering the data we obtained, we ensure the public that we are publicly sourced. This also helps keep our data accountable and accurate since we are using a public database. Another concern we realized we would face would be data privacy, especially since this data also includes personal information about individuals and groups. To ensure this privacy isn't breached, we wanted to make sure any sensitive information shared follows privacy compliance laws and that anything too sensitive isn't included. We also realized that certain groups or individuals might be represented more in the database, and we wanted to make sure that there was no bias at all in the end product of our database. We want to make sure every group or individual that donated or had expenditures is included in the database, with no bias towards anyone. By doing this, it makes our database representative and accurate, preventing any group from being excluded. The last consideration we wanted to make clear is that this dataset will be labeled under the term free use. We want to make it clear that this database will have equitable access, meaning that it will be

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

free to use by anyone. This has been at the core of the creation of our database since the first report, making sure this database is representative and equitable.

Lessons Learned:

Our biggest challenge in this project was the formation of our ERD and normalizing our data. Initially, we found it hard to organize our tables and rows into logical groupings and also found that we had redundant attributes. For example, we had attributes such as payees_paid_by and other repeating values that were not needed. We also had some other cases of this, but after going to the professor and talking amongst ourselves, we were able to get rid of those attributes and reduce the size of our data to be more efficient. Removing such columns and attributes also forced us to change parts of our ERD as well, but it was worth it in the end to create an optimized database. Another problem we had for most of our project was communication, where for certain parts of the work everyone wasn't sure on what to do specifically. This was especially a problem when deadlines were coming up, and we had only done the minimum work a day or two before. This was something we had to change and were able to by starting to talk about assignments at least 5 days in advance, and how we would approach them. This saved us a lot of time, and prevented us from stressing about assignments the day of anymore.

Potential Future Work:

In the future, our project could expand by making the data we collect broader and also have real-time updates. This project focused on the data we collected from the official city of Austin's data portal, but in the future, we could also include data collected from other official government sites. We could also merge our database with other datasets, which could show off more of the political landscape. This would include things such as voter demographics, election results, specific activities, and more to emphasize the effectiveness of using campaign finance.

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

This could all later on be implemented into some sort of public-use government website, where the public and officials can analyze the trends themselves from the database which could help with campaigns and deciding how to allocate resources.

INST 327
Section: WB21
Project Final Report
8/13/2024
Team RRRE
Ryan McNeill,
Tariq Witherspoon,
Rheymar Devera,
Ethan Schwartzberg

References

"Campaign Finance - Contributions." *Data.AustinTexas.Gov - The Official City of Austin Open Data Portal*, City of Austin, Texas, 22 July 2024, data.austintexas.gov/City-Government/Campaign-Finance-Contributions/3kfv-biw6/about_data.

"Campaign Finance - Expenditures." *Data.AustinTexas.Gov - The Official City of Austin Open Data Portal*, City of Austin, Texas, 22 July 2024, data.austintexas.gov/City-Government/Campaign-Finance-Expenditures/gd3e-xut2/about_data.