**"Group Case Study #2"**

**Business Intelligence for BIA**

**(BIA-5401)**

**Submitted to:**

**Dr. Haytham Qushtom**

**Professor,**

**Humber College**

**Submitted by:**

**Prappan Batra (N01579150)**

**Dhairya Dangi (N01580705)**

**Param Panchal (N01579822)**

**Vrajkumar Patel (N01581006)**

**Pravina Prajapati (N01579926)**

**Harsh Agrawal (N 01580593)**

**Submitted On:**

**June 10, 2024**

Table of Contents

[**1.** **Introduction** 3](#_Toc168933393)

[**2.** **Data Description** 3](#_Toc168933394)

[**3.** **Schema Description** 4](#_Toc168933395)

[**4.** **Analytical Question:** 9](#_Toc168933396)

1. **Introduction**

Through improved management and analysis of the funds gathered during their yearly fundraising campaign, a pet rescue charity hopes to benefit from this report. We hope to provide helpful insights that will enable the organization to enhance its donation collecting and make informed decisions by utilizing specialized methodologies for data analysis and a well-structured database.

The pet rescue charity holds a yearly donation drive where volunteers go door-to-door in different areas of the city to collect donations. Each area includes about six postal codes and is assigned a group of volunteers. These volunteers collect donations in cash, checks, or credit card payments. At the end of each week, the group leaders create a list of collected donations and send it to the charity’s main office. The staff at the main office then enter this data into a central database, checking for any errors and removing invalid entries. This collected data is then used to analyze how well the donation drive is going and find ways to improve it.

This assignment will involve building a database to house all of the donation data, organizing it in a way that makes it easy to analyze, and employing unique procedures to provide answers to a few key questions. The organization will be better able to make decisions to enhance their fundraising efforts by using this information to better understand their donation trends.

1. **Data Description**

**Source:**

The data used in this project comes from two main sources:

1. **Central Donations Repository:** This is a database where the main office of the charity stores all the donation data collected during the annual drive. This repository includes several tables to store different types of information.
2. **DonorsList.csv Files:** These are comma-separated files sent by the volunteer group leaders at the end of each week. Each file contains a list of donations collected in a specific area by a group of volunteers. Each group member in our project has provided a separate list for their area with at least 15 entries.
3. **Schema Description**

**Importing CSV Files:** We created the Donor\_Data table to store the donor information collected from various CSV files. These files represent different areas where donations were collected. We used a stored procedure to import these files into the Donor\_Data table. The CSV files include East York.csv, Etobicoke.csv, North York.csv, Scarborough.csv, Toronto.csv, and York.csv

**Stored Procedure for Importing Data:** We created a stored procedure named ImportDonorsCSVFiles to handle the bulk import of data from the CSV files into the Donor\_Data table. This procedure disables triggers, performs the bulk inserts for each file, and then re-enables the triggers.

**Executing the Stored Procedure:** We executed the ImportDonorsCSVFiles stored procedure to populate the Donor\_Data table with data from the six CSV files. This step ensured that all donor information was consolidated into a single table for further processing and analysis.

The Central Donations Repository is structured into several tables, each designed to store specific information. Here's a detailed description of each table's schema:

**Donor\_Data description:**

* **Address (VARCHAR(255)):** This column stores the full address of the donor. The address typically includes the unit number, street number, street name, street type, street direction, postal code, city, and province. This data is later parsed and split into individual components for more detailed analysis.
* **DonorName (VARCHAR(255)):** This column stores the full name of the donor. It is a combination of the donor's first and last names.
* **Date\_of\_Donation (VARCHAR(50)):** This column stores the date when the donation was made. The date is stored as a string and may need to be converted to a date type for certain analyses.
* **Time (VARCHAR(255)):** This column stores the time when the donation was made. This data can be used to analyze donation patterns throughout the day.
* **Amount (VARCHAR(255)):** This column stores the amount of the donation. The amount is stored as a string and may need to be converted to a numeric type for financial calculations.
* **Type (VARCHAR(50)):** This column indicates the payment method used for the donation (e.g., Cash, Check, Credit Card). This information is essential for understanding the distribution of payment methods.
* **VolunteerID (VARCHAR(255)):** This column stores the identifier of the volunteer who collected the donation. It helps in tracking which volunteer collected which donation and can be used to analyze the performance of individual volunteers or groups.
* **ID (INT IDENTITY(1,1) PRIMARY KEY):** This column is an auto-incrementing primary key that uniquely identifies each donation record. It ensures that each entry in the Donor\_Data table can be uniquely referenced.

**Address Table Description:**

The Address table is created to store detailed information about the addresses of the donors. This table was populated by parsing the address information from the Donor\_Data table. Below is the schema of the Address table:

* **addressID (INT Primary Key):** This is a unique identifier for each address. It is an auto-incrementing primary key.
* **unit\_num (VARCHAR(50)):** This column stores the unit number of the donor's address, if applicable (e.g., apartment or suite number).
* **street\_number (VARCHAR(50)):** This column stores the number of the building on the street where the donor lives.
* **street\_name (VARCHAR(100)):** This column stores the name of the street where the donor lives.
* **street\_type (VARCHAR(50)):** This column stores the type of the street (e.g., Avenue, Street, Boulevard).
* **street\_direction (VARCHAR(50)):** This column stores the direction of the street, if applicable (e.g., North, South).
* **postal\_code (VARCHAR(10)):** This column stores the postal code of the donor's address.
* **city (VARCHAR(100)):** This column stores the city where the donor lives.
* **province (VARCHAR(50)):** This column stores the province where the donor lives.

**Volunteer Table:**

The Volunteer table is created to store information about the volunteers who help in the donation drive. This table includes details about each volunteer, including their name and the group leader they report to. Below is the schema of the Volunteer table:

* **volunteerID (INT Primary Key identity(1,1)):** This is a unique identifier for each volunteer. It is an auto-incrementing primary key.
* **first\_name (VARCHAR(50)):** This column stores the first name of the volunteer.
* **last\_name (VARCHAR(50)):** This column stores the last name of the volunteer.
* **Group\_Leader\_ID (INT):** This column stores the identifier of the group leader to whom the volunteer reports. This helps in organizing the volunteers into groups for better management.

**Payment Method Table:**

The Payment\_Method table is created to store information about the various payment methods used by donors to make contributions. This table helps in categorizing and analyzing the donations based on the payment methods. Below is the schema of the Payment\_Method table:

* **Payment\_Method\_ID (INT Primary Key Identity(1,1)):** This is a unique identifier for each payment method. It is an auto-incrementing primary key.
* **Payment\_Type (VARCHAR(50)):** This column stores the type of payment method used by the donor (e.g., Cash, Credit, Check).

**Donation Table:**

The Donation table stores information about individual donations made by donors during the charity's annual donation drive. This table contains details such as the donor's name, donation date, amount, and payment method. Below is the schema of the Donation table:

* **Donation\_ID (INT Primary Key):** This is a unique identifier for each donation. It serves as the primary key for the table.
* **donor\_first\_name (VARCHAR(50)):** This column stores the first name of the donor who made the donation.
* **donor\_last\_name (VARCHAR(50)):** This column stores the last name of the donor who made the donation.
* **donation\_date (DATE):** This column stores the date when the donation was made.
* **donation\_amount (INT):** This column stores the amount of the donation in monetary units.
* **payment\_method\_ID (INT):** This column stores the identifier of the payment method used for the donation. It is a foreign key referencing the Payment\_Method table.

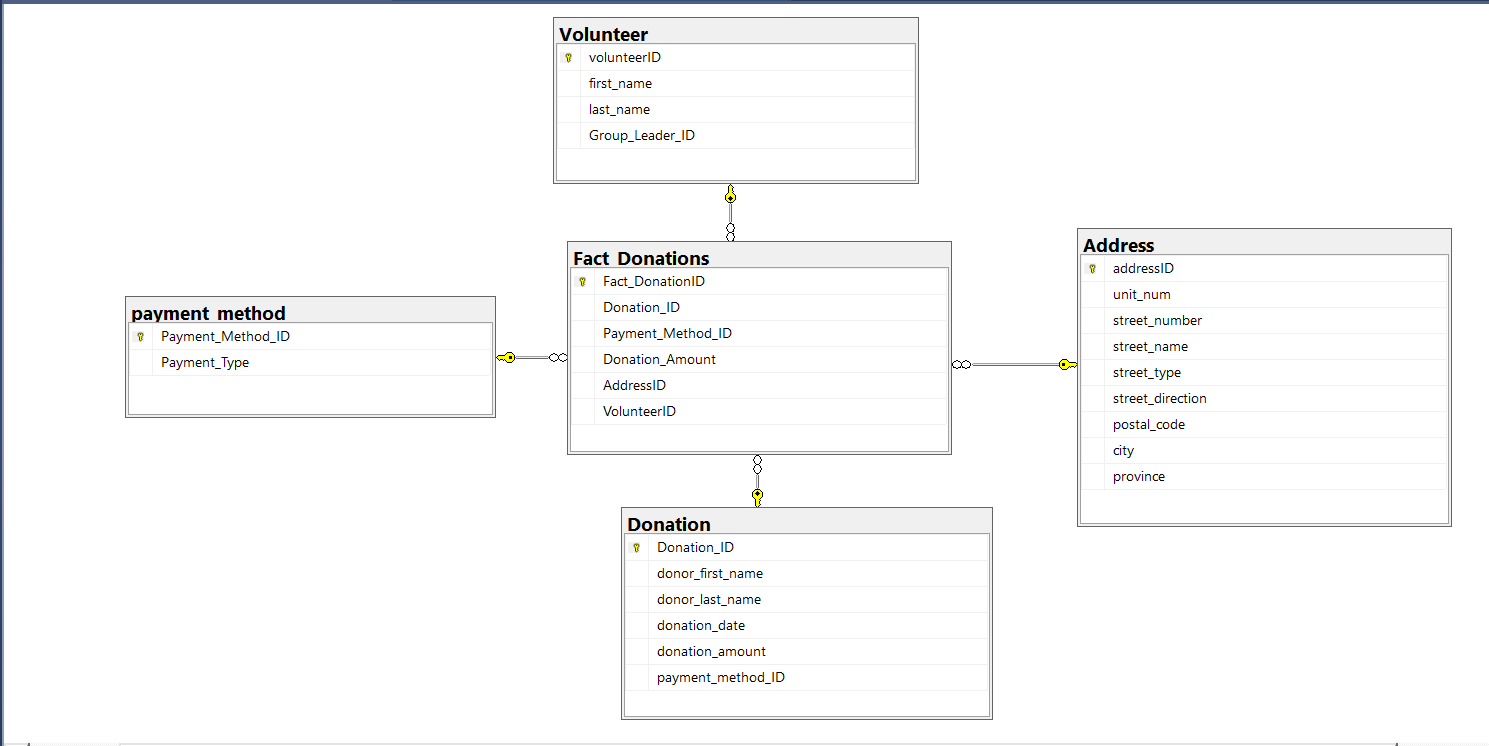
**Fact\_Donations Table:**

The Fact\_Donations table is a key component in a data warehousing architecture, particularly in the context of dimensional modeling. It serves as a central repository for storing transactional data related to donations, providing a foundation for analytical queries and reporting. Below is the explanation of the Fact\_Donations table:

* **Fact\_DonationID (INT IDENTITY(1,1) PRIMARY KEY): In the fact table, this column acts as a unique identifier for every donation transaction.**
* **Donation\_ID (INT): This column connects each donation transaction to its matching record in the donation dimension by referencing the main key of the Donation table. It creates a link of one to many between donations and the information they are associated with.**
* **Payment\_Method\_ID (INT):** This column references the primary key of the Payment\_Method table, indicating the payment method used for each donation. It establishes a one-to-many relationship between payment methods and donation transactions.
* **Donation\_Amount (INT):** This column stores the monetary value of each donation transaction, representing the amount contributed by the donor.
* **AddressID (INT):** This column references the primary key of the Address table, linking each donation transaction to the donor's address information. It establishes a one-to-many relationship between donation transactions and donor addresses.
* **VolunteerID (INT):** This column references the primary key of the Volunteer table, indicating the volunteer responsible for collecting each donation. It establishes a one-to-many relationship between volunteers and donation transactions.
* **Foreign Key Constraints:** The Donation\_ID, Payment\_Method\_ID, AddressID, and VolunteerID columns are all foreign keys that enforce referential integrity with their respective dimension tables. These constraints ensure that only valid and existing values are stored in the fact table, maintaining data consistency.

**Dimensional Modeling (Star Schema):**

After creating Fact\_donations table, we created a star schema to analyze the data.



**Fact Table (Fact\_Donations):**

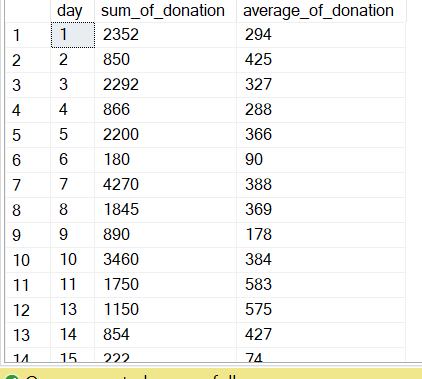
* The Fact\_Donations table is the central table that stores transactional data related to donations made during the charity's annual fundraising drive.
* It contains quantitative data, or "facts", such as donation amounts, and serves as the foundation for analytical queries and reporting.
* Each row in the Fact\_Donations table represents a specific donation transaction made by a donor.

**Dimension Tables:**

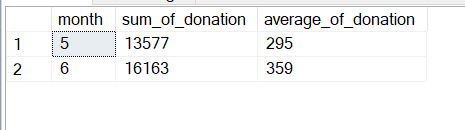
* Dimension tables provide descriptive context to the data stored in the fact table.
* In this case, several dimension tables are mentioned:
  + **Volunteer Dimension Table:** This table stores information about the volunteers who collect donations, such as their names and id.
  + **Donations Dimension Table:** This table could contain additional details about the donations themselves, such as donation types, purposes, etc.
  + **Payment\_Method Dimension Table:** This table holds information about the various payment methods used by donors to make contributions, such as cash, credit card, or check.
  + **Address Dimension Table:** This table stores detailed information about the addresses of the donors, including unit numbers, street names, postal codes, etc.

1. **Analytical Question:**
2. **The average and sum of the donation by day, month, and year**

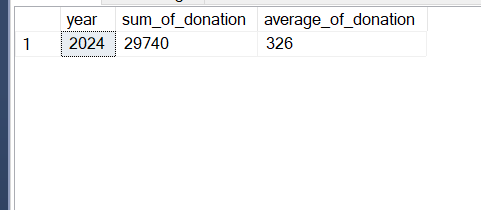
* By Day:



* By Month:



* By Year:



1. **The average and sum of the donations by postal code and City in a specific month. define the city and month as variables to allow flexibility.**

* Here, we set variables for city = Etobicoke and Month = 5

A screenshot of a computer

Description automatically generated

1. **The amount collected per payment method from the city with highest $ value of donations. Define the payment method as variable to allow flexibility.**

* Here, we have set variable payment\_method = “Check”

