HUMBER INSTITUTE OF TECHNOLOGY

AND ADVANCED LEARNING

BIA-5401-0GA: Business Intelligence

GROUP CASE STUDY 02

Submitted by: Group 4 Grade/Comments:

First Name	Last Name	Student Number
Amarachi	Ezeji	N01510367
Aparajita	Roy	N01511087
Ayesha	Karadia	N01514568
Gbenga	Adebowale	N01513109
Ivin	Alexander	N01474172
Navdeep	Chauhan	N01514720

Submitted to: Professor Haytham Qushtom

Submission Date: 07/16/2023

TABLE OF CONTENTS

INTRODUCTION

This study attempts to describe the procedure and significance of handling donation data for an annual fundraising campaign for a pet rescue organisation. The nonprofit, which is dedicated to saving and caring for animals in need, organises a citywide campaign to gather money for their admirable cause. The donation drive entails segmenting the city into donation areas, allocating committed teams of volunteers to each area, and requesting money from locals by going door to door.

Each volunteer received six postal codes for a specific region. These dedicated volunteers are tasked with gathering donations made by locals in the form of cash, cheques, or credit card payments. To ensure data accuracy and reliability, each group leader compiles a weekly donation record that is then sent to the charity's head office.

The various elements of donation data management and analysis will be covered in more detail in the sections that follow. It will go over the procedures for gathering, validating, and loading data as well as the many analytical methods that can be used to glean useful information. The paper will also examine the possible advantages of effective data management, such as better decision-making, focused fundraising, and better resource allocation.

The success of a pet rescue charity's yearly fundraising campaign depends heavily on the efficient management and analysis of donation data. The organisation can improve its fundraising efforts, gain a better understanding of donor behaviour, and have a bigger impact on their mission to rescue and care for animals in need by adopting a systematic approach to data collection, validation, and analysis.

DATA PREPARATION

The volunteers were tasked with collecting donations in Sudbury, Brampton, Oshawa, Scarborough, Toronto, and York. Each volunteer kept track of the information about the contributors on paper. We structured the data after merging them all together after data collection. The information included the contributors' name, donation amount, date, and method of payment, as well as the addresses of the donors. We chose to change the data format after deliberating as a team. We created two columns with the donors' first and last names for their names. Additionally, we separated the donors' addresses into four columns: address (unit or block number), city, province, and postal code. We accomplished this so that working with data in SQL will be simpler for us in the future. For each volunteer, we added a second column for the group leader.

As time went on, we realised that these variables also require being given distinctive keys. We were aware that we would need to create three tables from this information: Address, Donation, and Volunteer. Therefore, we gave each address a unique ID and each donor a unique ID. As follows, our dataset includes many types of attributes:

Address_ID: A numerical unique key given to address.

Street_Address: Address of the Donor’s which includes unit number and building name.

City: The name of the city where the donor resides.

Province: The name of the province where the donors reside.

Postal_code: The postal code where the donor resides.

Donation_ID: A numerical unique key given to each donor.

Donor_FirstName: The first name of the donor.

Donor_LastName: The last name of the donor.

Donation_Date: The date at which donation was collected from the donor.

Donation_Amount: The amount of money donated by the donor.

Payment_Method: The method of payment used by donor to donate money.

Volunteer_id: A numerical unique key given to each volunteer.

Volunteer_FirstName: The first name of the Volunteer.

Volunteer_LastName: The last name of the Volunteer.

Group_Leader: The name of the group leader which were assigned to each volunteer.

There are 100 rows and 15 columns of text and numeric data in the dataset. Five of the fifteen columns are numerical, and nine are character columns.

DIMENSIONAL MODELLING

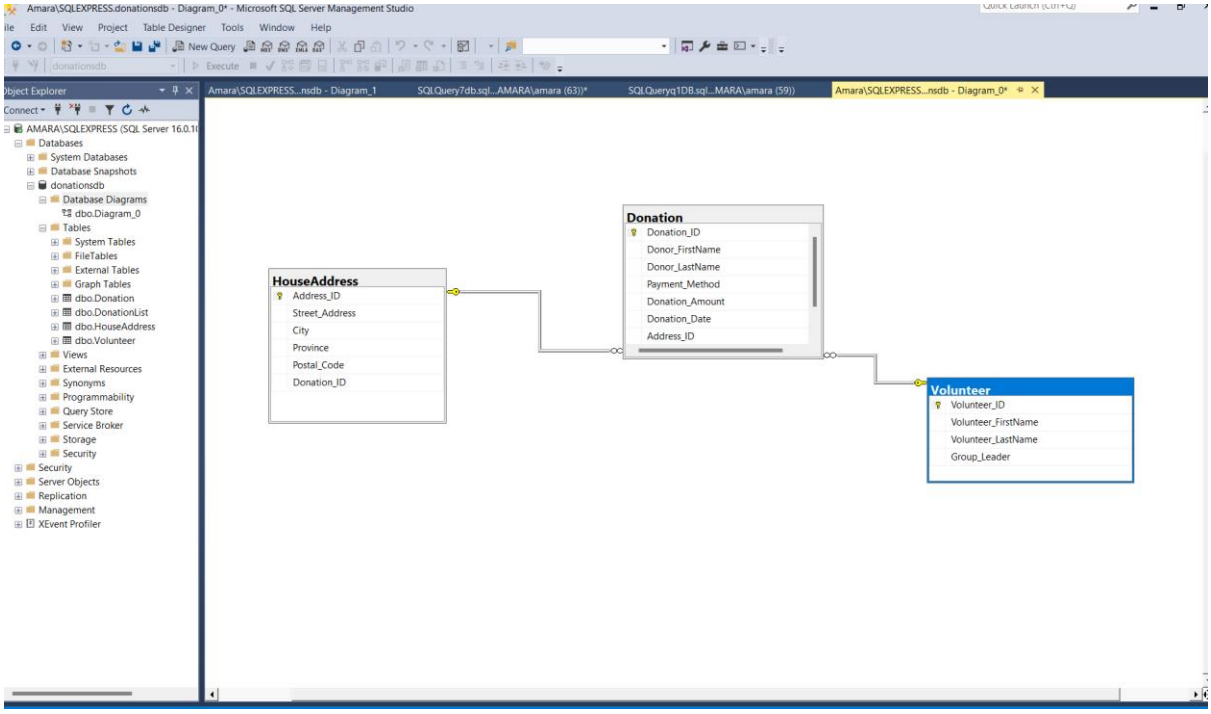

Image 1 – Final Star Schema

The data model for the four tables in the above *Image 1* mentioned can be represented using entity-relationship (ER) modeling. Below is the explanation of each table and its relationships:

Donation List:

This table represents a list of donations made, including information about the donor, the donation itself, and the volunteer involved.

It includes foreign key references to the House Address table (Address_ID) and Volunteer table (Volunteer_ID).

House Address:

This table represents the addresses of the houses associated with the donations.

It includes a foreign key reference to the Donation table (Donation_ID).

Donation:

It includes foreign key references to the Volunteer table (Volunteer_ID) and House Address table (Address_ID).

Volunteer:

This table represents information about the volunteers involved in the donation process.

It includes a foreign key reference to the Donation List table (Volunteer_ID).

The relationships among these tables can be described as follows:

The Donation List table has a one-to-many relationship with the House Address table based on the Donation_ID. This means that each donation in the Donation List can have multiple associated addresses in the House Address table.

The Donation List table has a one-to-many relationship with the Donation table based on the Donation_ID. This means that each donation in the Donation List can have multiple associated records in the Donation table.

The Donation table has a many-to-one relationship with the Volunteer table based on the Volunteer_ID. This means that multiple donations can be associated with a single volunteer.

The Donation table has a many-to-one relationship with the House Address table based on the Address_ID. This means that multiple donations can be associated with a single address.

The Volunteer table has a one-to-many relationship with the Donation List table based on the Volunteer_ID. This means that each volunteer can be associated with multiple donations in the Donation List table.

Donations, donor information, volunteer information, and dwelling addresses may all be tracked using this data model. It enables users to link donations to specific addresses and volunteers, giving you a complete picture of the donation process.

ANALYTIC QUESTIONS AND INSIGHTS

The analysis was done to answer the following questions.

- The average and sum of the donation by day, month, and year.

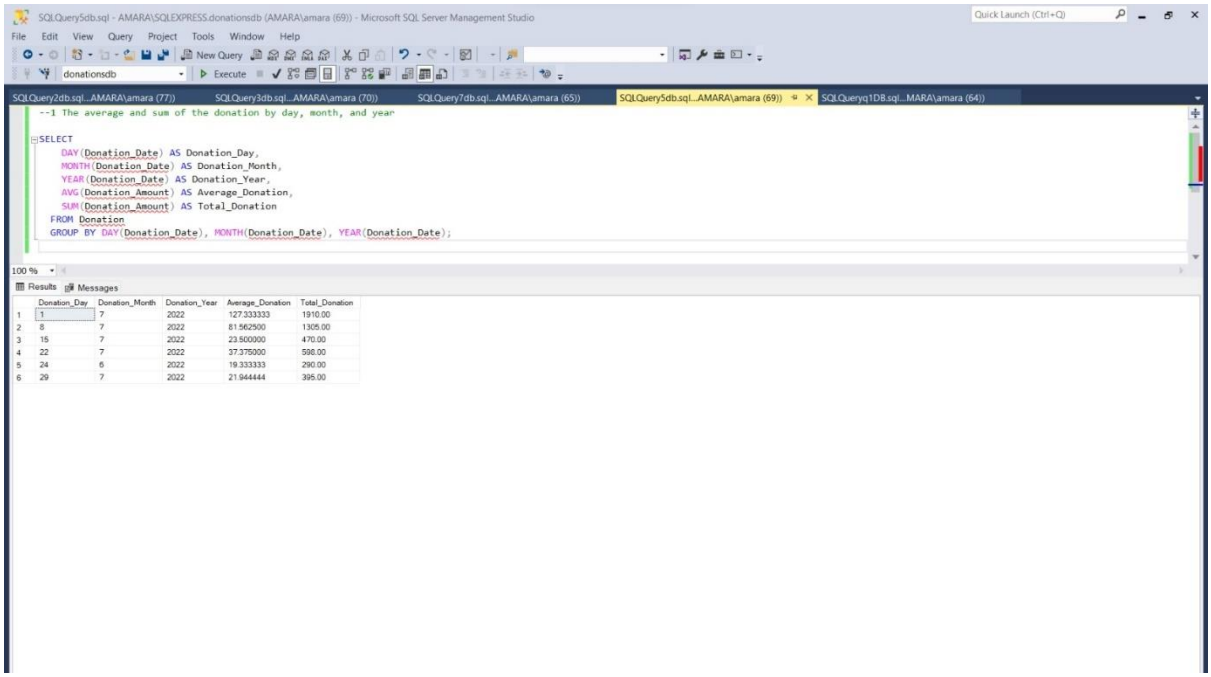

Image 2. Results for the average and sum of the donation by day, month and year

The above code in *Image 2* analyses the Donation table to determine the average and amount of donations made on a daily, monthly, and annual basis. The code extracts the day, month, and year from the Donation_Date column using the DAY (), MONTH (), and YEAR () methods, accordingly. This enables donation groups to be formed based on these time components.

The SELECT query returns the columns Donation_Day, Donation_Month, and Donation_Year, which correspond to the temporal components of the Donation_Date. It also uses the AVG () method to compute the average donation amount and the SUM () function to obtain the overall donation amount.

The GROUP BY clause is used to group the organize donations based on the Donation_Date column's day, month, and year values. This guarantees that the average and total computations are run for each unique day, month, and year combination.

We gain significant insights about donation trends over time by running this code. The average donation amount gives information on the normal contribution made on a given day, month, or year, allowing you to better understand donor behaviour and preferences. The total of donations represents the entire financial effect of contributions throughout certain time periods, allowing for comparisons and trend analysis. This information can be utilized to make strategic decisions, identify peak donation periods, evaluate campaign efficiency, and analyse overall fundraising performance.

- 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.

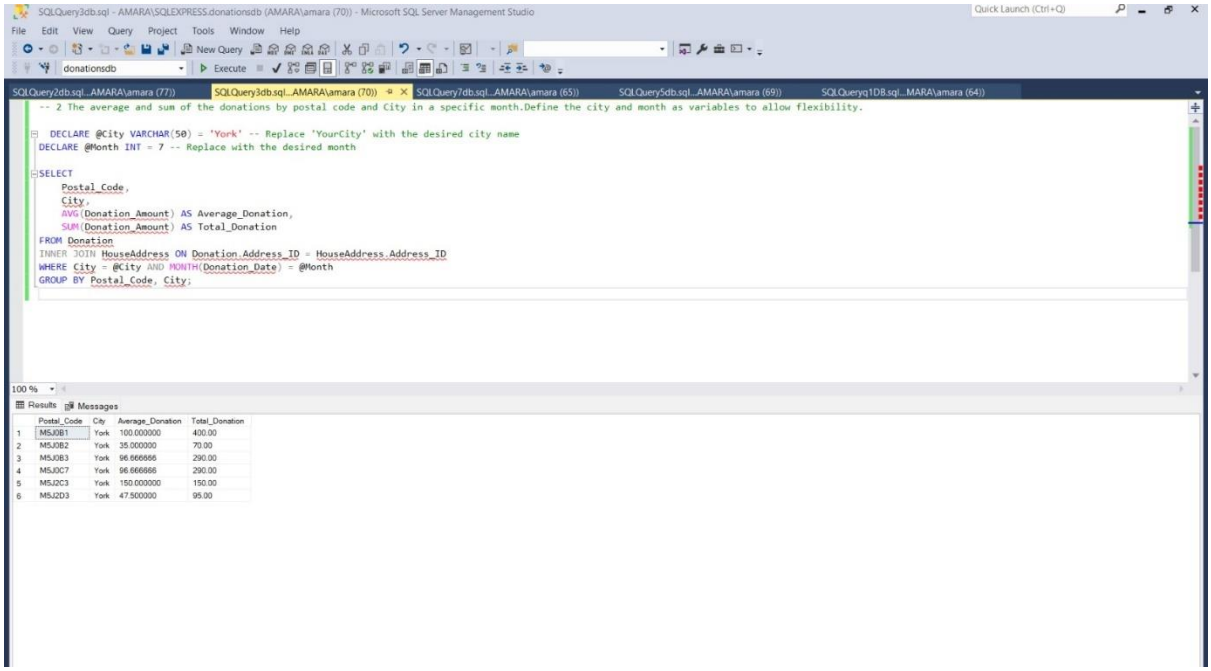

Image 3. Results showing the average and sum of the donations by postal code in a specific month

This question seeks to understand the breakdown of donation amounts based on different payment methods such as cash, checks, and credit card payments. By analysing the distribution, the charity can gain insights into donor preferences and identify which payment methods are more popular or effective in generating higher donation amounts. This information can help them tailor their fundraising strategies and payment options accordingly.

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

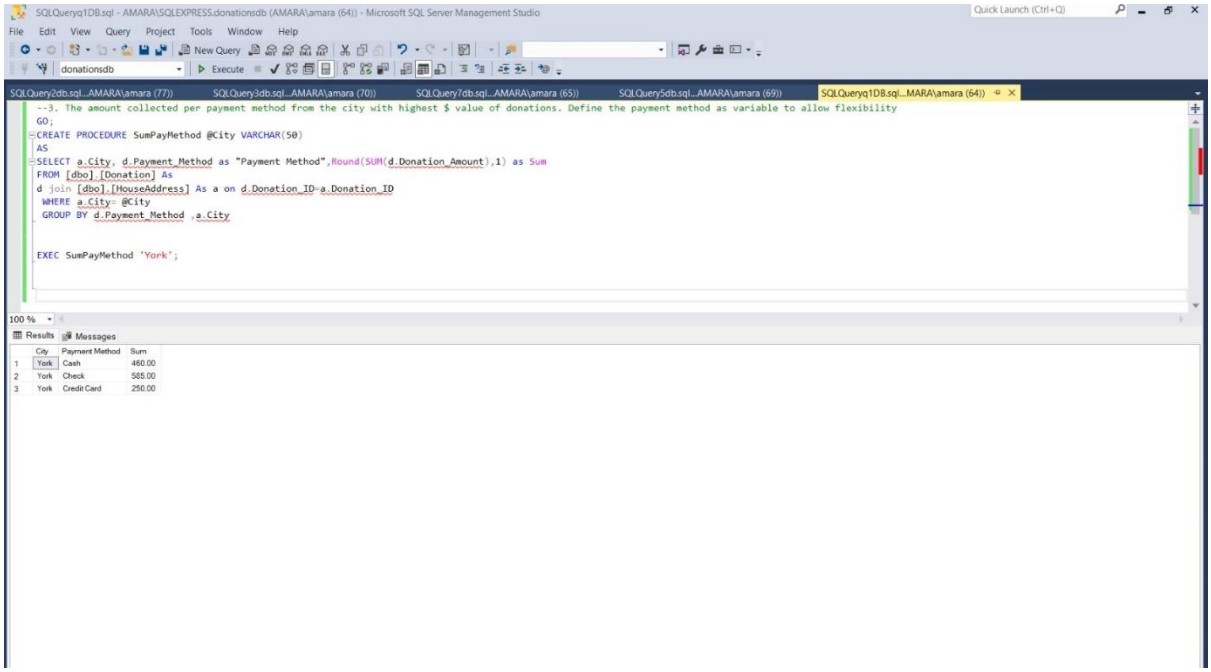

Image 4. Results for the amount collected per payment method from the city with highest value of donations

Define the variables for city and month to allow flexibility in the analysis. This inquiry aims to explore donation patterns within different geographic areas and time limits, helping the charity identify potential trends and prioritize efforts for fundraising initiatives.

These business questions can be answered using stored procedures in the SQL database, allowing the charity to retrieve the desired insights from the donation data and make informed decisions to enhance their fundraising campaign.

CONCLUSION

As we analyzed our data, we discovered that the method of payment had an impact on the number of donors who were willing to provide. People typically prefer to pay with checks or cash, and if they choose checks, there is a good risk that they may incur significant financial obligations. We also knew that addresses with the postal code "M5J0B1" had the most donations overall and the highest average gift in the York region.