**Semester 1- Group 3- Sushma Gowri Deivanayagam Oluwatobiloba Oladimeji Sachin Suresh Dwij Shyamsunder Surya Amit Malwade**

**SQL GROUP PROJECT**

**Ontario community**

Table of Contents

[A. Introduction 1](#_Toc164015307)

[B. Data processing 1](#_Toc164015308)

[C. Analysis and Insights 4](#_Toc164015309)

[D. Visualizations 14](#_Toc164015310)

[E. Summary and Findings 16](#_Toc164015311)

# Introduction

For this project we have accessed Welcome Ontario dataset to get the Ontario Builds: Communities dataset and the links are given below:

[Welcome - Ontario Data Catalogue](https://data.ontario.ca/)

<https://data.ontario.ca/dataset/ontario-builds-key-infrastructure-projects/resource/36f92c5b-0c8b-4a4b-b4c5-d15a43894297>

This dataset contains valuable information related to key infrastructure projects in Ontario, which could be analyzed to extract insights and inform decision-making. Analysis of this dataset will help decision-making and project management within the government's infrastructure and community development initiatives.

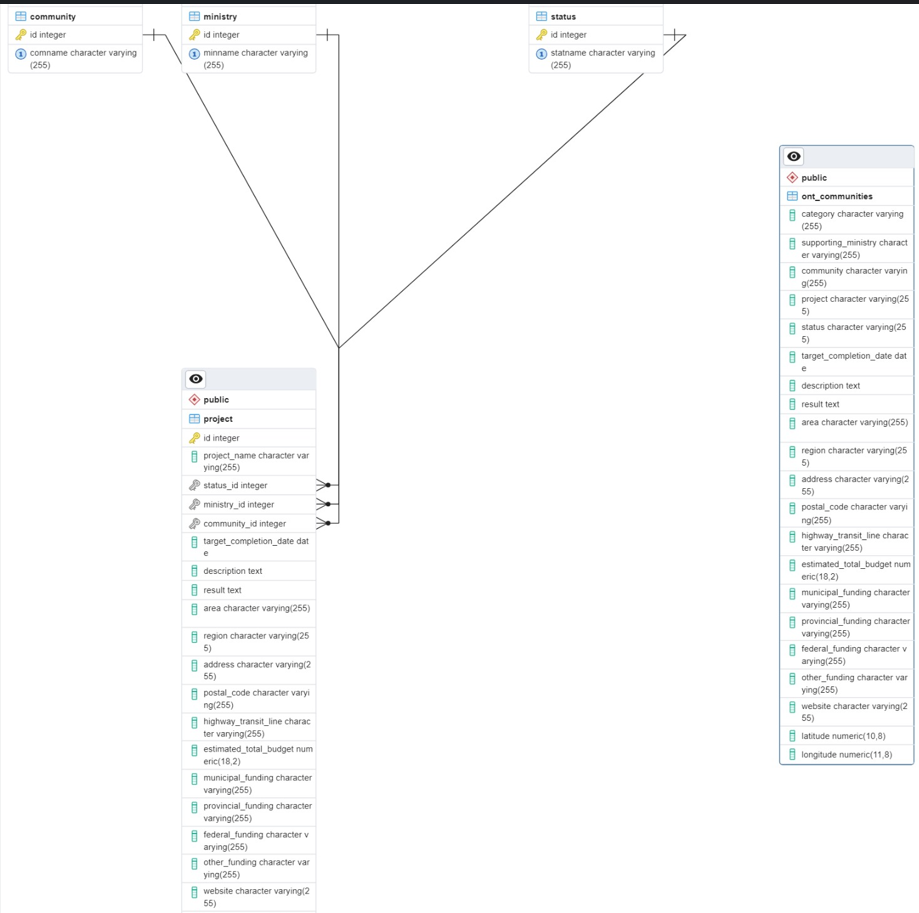
# Data processing

The dataset taken is of Ontario community Data Processing;

1. **Table Creation:**

The first step is to create a table called ont\_communities, and it has several columns that reflect different aspects of infrastructure projects. Each column has a list of the data types that are allowed, such as VARCHAR for text, DECIMAL for numbers, and DATE for dates.

1. **Setup relationships and attach the schema design (ERD) diagram as a picture.**



1. **Clean, Transform and Load into PostgreSQL/PGAdmin.**

Import the csv file for Data Cleaning Process

-- Update character/text columns with 'unknown' if empty

UPDATE ont\_communities –

If any character or text column is empty, the script updates it with 'unknown', and if any number column is empty, it updates it with 0. By doing this, consistency is guaranteed and processing mistakes are avoided.

Look for Row Duplication:  
  
To search the ont\_communities table for duplicate rows based on all columns, a SELECT query is utilized.  
Duplicate rows are identified using the HAVING clause with COUNT(\*) > 1.

In order to remove duplicate entries, a temporary table called temp\_unique\_rows that contains only unique rows from ont\_communities must first be created.

To eliminate all of the current data, the original ont\_communities table is shortened.  
The ont\_communities table receives the unique rows from the temporary table back.  
The temporary table is finally removed.

# Analysis and Insights

10 queries are created as below;

**Query 1:** **List of all projects along with their associated community names**

A screenshot of a computer

Description automatically generated

**Insight**: This query shows which projects are associated with which communities, providing a link between community development and project implementation. Comprehending which projects are associated with particular communities facilitates focused communication and interaction with locals, guaranteeing that their requirements and worries are taken care of during the project's duration.

**Query 2:** **Number of projects in each ministry**

A screenshot of a computer

Description automatically generated

**Insight**: This inquiry provides information about the emphasis and investment areas of the various government ministries and how their projects are distributed among them. The agency can evaluate resource allocation, track project progress, and guarantee alignment with strategic priorities established by various government departments by keeping track of the number of projects under each ministry.

**Query 3:** **Total budget allocated for projects in each region**

A screenshot of a computer

Description automatically generated

**Insight**: It is possible to determine whether regions receive more cash for community development or infrastructure projects by aggregating the budget according to regions. By breaking down budget allocation by region, it is possible to guarantee fair resource distribution and sufficient money for development projects in underserved areas.

**Query 4: Projects with their status**

A screenshot of a computer

Description automatically generated

**Insight**: This query helps stakeholders monitor progress by giving a summary of each project's current state. To guarantee on-time delivery and adherence to quality standards, monitoring project statuses facilitates efficient project management, the identification of bottlenecks, and the implementation of remedial measures.

**Query 5:** **Average budget of projects in each community**

A screenshot of a computer

Description automatically generated

**Insight**: This inquiry provides information on the average amount of funding allotted to various community projects, as well as the levels and priorities of investment. Analyzing the average budget for each community aids in determining the return on investment and in setting resource allocation priorities according to requirements and development objectives.

**Query 6:** **Projects with a completion date past the target completion date**

A screenshot of a computer

Description automatically generated

**Insight**: This search finds projects that are behind schedule, which may be a sign of delays or problems with the way the project is being carried out. Proactive involvement to address obstacles, reduce risks, and minimise disruptions to project timelines and overall programme delivery is made possible by identifying projects that are behind schedule.

**Query 7: Total funding received for projects in each region**

A screenshot of a computer

Description automatically generated

**Insight**: Decision-makers can evaluate the overall investment in infrastructure and community development across several geographic areas by adding up the total money for projects in each location. This understanding aids in resource allocation and strategic planning to resolve regional imbalances and encourage balanced growth.

**Query 8:** **Total budget allocated to projects for each ministry in a specific region**

A screenshot of a computer

Description automatically generated

**Insight**: This search offers a thorough breakdown of budget distribution by ministry for every region, giving valuable information on priorities and resource distribution. To maximize the impact of development projects, strategic planning, resource optimization, and coordination among various government agencies are aided by the evaluation of budget allocation by the ministry and region.

**Query 9:** **Projects with the highest estimated total budget**

A screenshot of a computer

Description automatically generated

**Insight**: Finding the projects in the dataset that have the largest costs can highlight important development or infrastructure initiatives. Setting priorities for large-scale projects necessitates thorough planning, risk analysis, and stakeholder involvement to guarantee successful execution and efficient use of available resources.

**Query 10: Number of projects in each status for a specific ministry**

A screenshot of a computer

Description automatically generated

**Insight:** By classifying projects according to their status, this query facilitates the tracking of project progress within each ministry. Ministry project status monitoring facilitates informed decision-making and ongoing project management practice improvement by offering insights into the performance, advancement, and difficulties encountered by each department.

# Visualizations

# Summary and Findings

Our study offers insightful information about community development and infrastructure projects based on an analysis of the Ontario Data Catalogue. The distribution of projects among ministries and regions, the average amount allotted to each community, and the state of projects about their intended completion dates are just a few of the significant themes that we discovered. Visual aids that improve comprehension of these insights include graphs that show project statuses and money allocations. Our findings underscore the significance of smart resource allocation, proactive project management to mitigate delays, and interagency coordination in ensuring efficient and equitable development. Our solution helps with well-informed decision-making and efficiently plans and executes development projects in Ontario by utilizing SQL queries and visualizations.