Final Report

README

Description

This database application was created in order for scientists and students to look into fossil fuel emissions of different countries. It allows the user to insert, update, delete, view, and compare data surrounding the US, Canada and Mexico's climate change impacts. The tables in the database are related to both emissions and the possible repercussions of climate change on each of the countries. The application utilizes Python as the programming language which is connected to a MySQL workbench database.

Features

- Insert new data into any of the tables in the database
- Delete tuples from any of the tables in the database
- Update data in the database
- View the tuples in a table in its entirety or based on a condition
- Compare pollution of bodies of water
- Compare GDPs of countries
- View visualizations of the data in the database
- View the total emissions for each country
- …

Requirements

- Python 3.3 or above
- MySQL workbench
- Python libraries: pymysql (for database connection), matploylib.pyplot (for displaying tables)

Setup

1. Open an environment in which you can run Python code.
2. If needed, install the necessary libraries: pip install pymysql/pip install matplotlib
3. Import the required libraries: import pymysql/ import matplotlib.pyplot as plt
4. Import the database schema into MySQL workbench using the accompanying SQL file: emissionDB.sql

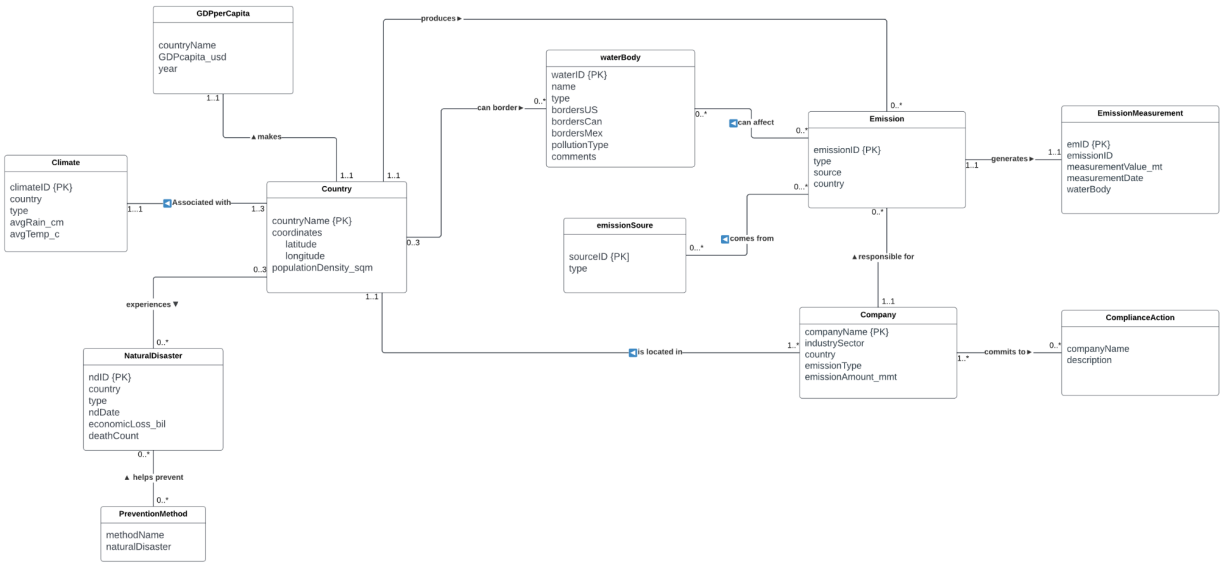
Usage

1. Run the database application: emissiondbApplication.py
2. Use your username and password to connect to the appropriate connection that the MySQL database is located in.
3. Choose if you want to insert, delete, view, or update something in the database. You also have the option to exit the program if you don't wish to accomplish any of these tasks.
4. Based on which of the following you chose, you will be prompted in order to complete the task effectively.
5. If you exit the program, you will be asked if you would like to compare or explore more using various procedures which you will also be appropriately prompted in order to accomplish.
6. If you answer no to performing the various procedures, you will be asked if you want to peruse any of the four visualizations created from the database
7. If you choose to not see any database visuals, the connection and the application will be ended.

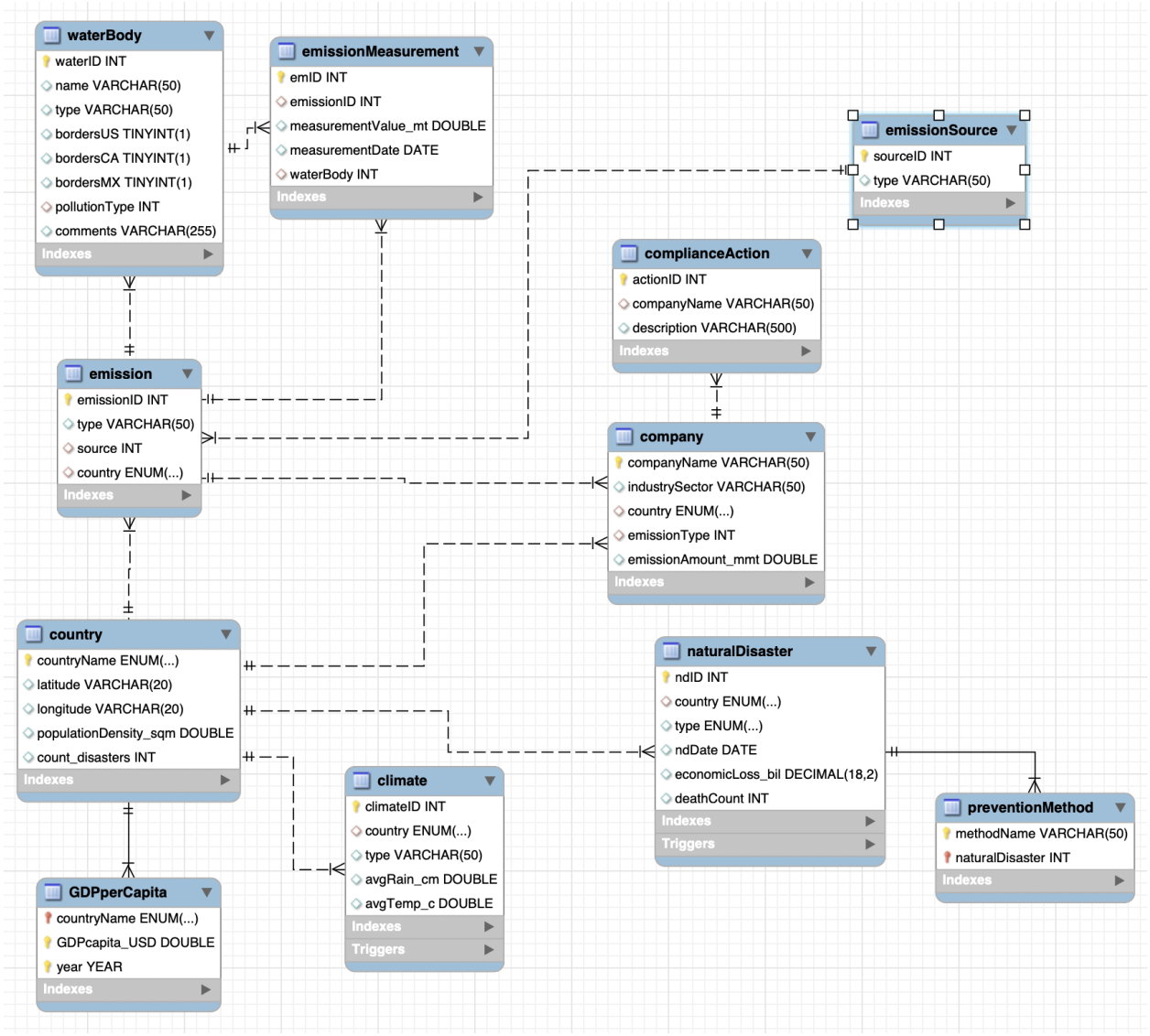
Technical Specifications

For our project, we used MySQL and MySQL workbench. This allowed us to effectively work with and store within our database. As all team members have experience using Python and its libraries, that is the language we used to increase efficiency and understanding while interacting with the database to provide a common-line user flow. The libraries we used within the program were pymysql in order to connect the database and matplotlib to create visualizations.

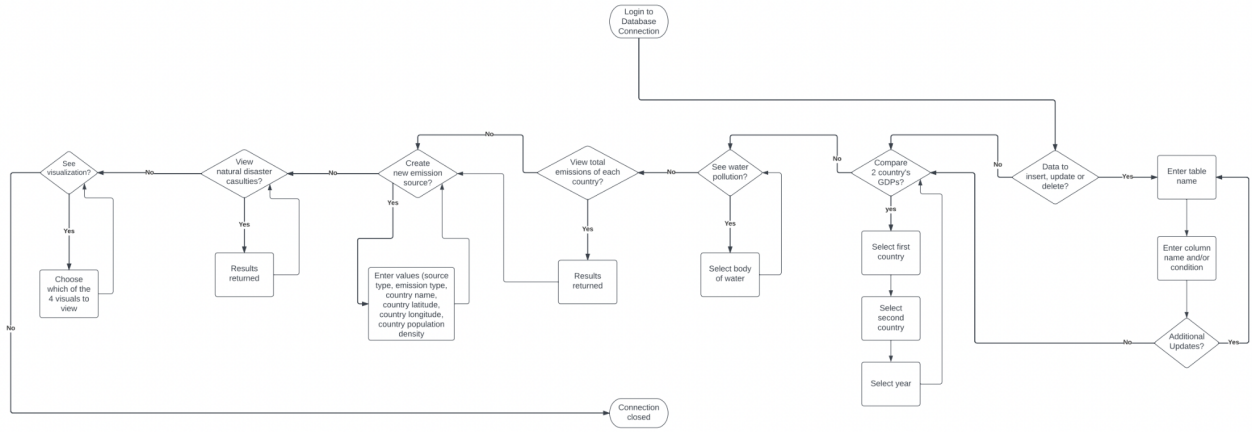
UML Conceptual Design



Logical Design



Final User Flow



Lessons Learned

Technical Expertise

Our group learned a variety of new technical skills throughout the project. In general, all of our SQL knowledge improved significantly from being able to work from start to finish on creating our schema. We all gained a deeper understanding and appreciation for the work and steps that go into making a database schema as effective and efficient as possible. We also improved our knowledge of UML notation and the steps towards creating an effective schema. While creating our functions, procedures, and triggers, we gained more insight on how to create useful database programming objects that were applicable to our schema and provided important information to those who would utilize them. Creating an application applicable in a real world setting, for both students and scientists, was the most important lesson learned. By creating a project with a real purpose, we had the opportunity to create something real world use, helping us look at each step that must be taken to do so.

Insights, time management insights, data domain insights

Ensuring the connectivity of various tables in order to easily perform necessary procedures was one of our greatest insights. Ensuring third normal form and limiting redundancies came into practice when improving our schema. Initially, there were quite a few redundancies; Once we cleaned this up, it caused the system to run much more efficiently. We also wanted to make sure that the data that we were inserting was accurate. We also based our data off of accurate US data, and created values for Canada and Mexico based on a scale of emissions comparative to the US in order to depict an accurate scale of how the US compares to the two other countries. It took extra time to go through the schema and find data from the internet that was based on accurate ratios and showed relevance. It was worth the extra time spent to show how countries generally compare. Finally, we were able to gain insight in troubleshooting in MySQL and in our database application that went beyond figuring out how to produce a homework answer correctly.

Realized or contemplated alternative designs/approaches to the project

We initially went too in depth in our UML notation tables and had too many dependencies and weak entities. For the final version of the project, we took a step back from our original plan and made the design much more broad than it was previously. Instead of using niche, specific tables we chose tables that broadly related to climate change. As a group this taught us the importance of the design process as it would’ve been incredibly difficult to implement the schema if we had stuck with our original UML model. Lastly, we were able to discover and learn about validating user input into the database application. Performing this type of validation and giving guidance to a user helped us create an easier and more reliable application. We grew confidence in our database designing through growing insight and experience from working on an effective database application.

Document any code not working in this section

From our knowledge and testing, there doesn't seem to be any code that isn't working in any of our sections.

Future Work

This database could be used by students and scientists alike to see the climate data affecting Canada, Mexico and the United States. As climate change becomes increasingly apparent, it becomes even more important to represent accurate and recent data contributing to our changing planet. It is one of the best ways to track how we are impacting the world and where we can implement improvements. For those who are curious, it also has the ability to compare countries which could allow the user to see which strategies for addressing climate change are working versus the practices that are not as effective.

Ideally this database would contain data from every country around the world. However, creating this big of a connection among countries takes a greater amount of work to insert much more data. Data on a few North American countries serves as a basis for smaller, but equally as important comparisons on countries the United States is physically closest to. There is possibility for other tables that are more specific and in depth pertaining to changes in our climate additional causes, since climate change causes are not mutually exclusive. Due to the nature of this project, we decided to keep our tables more broad in order to not overcomplicate. There are a variety of different ways that climate change is impacting Earth so we had to be selective to choose the tables that were most relatable to one another.