INFO 430: RELATIONAL DATABASE MANAGEMENT SYSTEMS

DATABASE PROJECT 2: Developing a Data Warehouse for Greenhouse Gas Emissions

Project 2 Deliverable 1: Project Proposal and Conceptual Data Modeling

Note: For project 2, ALL groups are working on the same topic provided by me.

Background

Climate change is one of the hotly debated topics around the world today. One the main culprits cited as the cause of climate change (global warming in particular) is greenhouse gas emissions.

For project 2, you are going to use the annual country-level emissions by sector dataset for 2015-2021 to create a data warehouse. This dataset, provided by Climate TRACE is available from Kaggle.com (https://www.kaggle.com/datasets/michaelbryantds/greenhouse-gas-emissions-dataset).

This dataset spans 70,000+ individual sites that represent the top known sources of emissions in the across the following sectors: power sector, oil and gas production and refining, shipping, aviation, mining, waste, agriculture, road transportation, and the production of steel, cement, and aluminum. The data is stored in 54 CSV files across 706 columns.

What You Should Do

For project 2 deliverable, your mail goal is to develop a dimensional model to represent the greenhouse gas emissions data warehouse. Please spend time to review and understand the data to enable you to develop a good dimensional model to represent the data.

As you review the data, think of the type of dimensional model that would be suitable for this dataset (i.e., either star schema, snowflake schema, or fact constellations).

Thereafter, draw your dimensional model using app-diagrams.net tool. Export your dimensional model as a PDF file.

What to Submit:

- Original dimensional model as an app-diagrame.net (i.e., ".drawio" extension) file
- A PDF version of the dimensional model

NOTE: Make sure that you have included the names of group members for all submitted work.