BLG 317E - DATABASE SYSTEMS

FALL 2024 (CRN: 13591)

PROJECT PROPOSAL

*Note that your selected dataset should be extensive enough to allow at least 5 main tables, and each table must have at least 3 non-key columns. Each team member must be responsible for at least one main table. Do your best in identifying table requirements. You will be asked to change your selected dataset if it does not seem to meet the minimum requirements. Please see the project evaluation and project self-assessment forms before finalizing your dataset to get an idea about the project evaluation criteria and process.*

|  |  |
| --- | --- |
| Project Group Name | Real Data |
| Dataset title | Sustainability Development Goals |
| Dataset URL | https://www.kaggle.com/datasets/hn4ever/world-development-indicators-by-countries? |
| Data description | This project utilizes the “World Development Indicators by Countries” dataset from Kaggle, which provides a wide range of sustainability and development metrics. Our team has organized this data into five key domains, with each group member responsible for one main data table: Freshwater Usage, Healthcare Systems, Greenhouse Gas Emissions, Sustainable Energy, and General Sustainability.  For this project, the data is modeled using a relational database approach, transforming the original "wide" data into a normalized "long" format. The schema is built around a single, central **countries** table (managed by Gülbahar Karabaş), which acts as a shared dimension table. To provide common details (e.g., region, country\_code) not present in the core dataset files, this central table will be manually created and populated by Gülbahar Karabaş.  Each main data table (e.g., **health\_system**, **energy\_data**) is linked to the **countries** table via a foreign key (**country\_id**). In addition, each group member has created their own **indicator\_details** table to provide specific metadata related to their topic. The **indicator\_id** values generated in these tables are used as foreign keys within their respective main data tables.  As a result, each main table contains one **country\_id** foreign key from the central **countries** table and **indicator\_id** foreign key from the related **indicator\_details** table — forming a total of **six different foreign keys** across the schema. This normalized structure ensures data integrity and supports efficient querying for both topic-specific and cross-domain analysis. |
| Member 1  Signature | |  |  | | --- | --- | | Full name | Muhammet Tuncer | | Student no | 820230314 | | Responsibility: | CREATE TABLE freshwater\_indicator\_details (  freshwater\_indicator\_id INT PRIMARY KEY,  indicator\_name VARCHAR(255) NOT NULL,  description TEXT,  unit\_of\_measure VARCHAR(100)  );  CREATE TABLE freshwater\_data (  data\_id INT PRIMARY KEY AUTO\_INCREMENT,  country\_id INT,  freshwater\_indicator\_id INT,  year INT,  indicator\_value FLOAT,  source\_notes VARCHAR(255),  FOREIGN KEY (country\_id) REFERENCES countries(country\_id),  FOREIGN KEY (­freshwater \_indicator\_id) REFERENCES freshwater\_indicator\_details(freshwater\_indicator\_id)  ); | |
| Member 2  Signature | |  |  | | --- | --- | | Full name | Gülbahar Karabaş | | Student no | 150210085 | | Responsibility: | CREATE TABLE countries (  country\_id INT PRIMARY KEY,  country\_name VARCHAR(100) NOT NULL UNIQUE,  country\_code VARCHAR(3) UNIQUE,  region VARCHAR(100)  );  CREATE TABLE health\_system (  row\_id INT PRIMARY KEY,  country\_id INT,  health\_indicator\_id INT,  indicator\_value FLOAT,  year INT,  source\_notes VARCHAR(255),  FOREIGN KEY (country\_id) REFERENCES countries(country\_id),  FOREIGN KEY (health\_indicator\_id) REFERENCES health\_indicator\_details(health\_indicator\_id)  );  CREATE TABLE health\_indicator\_details (  health\_indicator\_id INT PRIMARY KEY,  indicator\_name varchar(100),  indicator\_description TEXT,  unit\_symbol VARCHAR(20)  ); | |
| Member 3  Signature | |  |  | | --- | --- | | Full name | Fatih Serdar Çakmak | | Student no | 820230326 | | Responsibility: | CREATE TABLE ghg\_indicator\_details (  ghg\_indicator\_id INT PRIMARY KEY,  indicator\_name VARCHAR(100),  indicator\_description VARCHAR(200),  unit\_symbol VARCHAR(20)  );  CREATE TABLE greenhouse\_emissions (  row\_id INT PRIMARY KEY,  country\_id INT,  ghg\_indicator\_id INT,  indicator\_value INT,  share\_of\_total\_pct INT,  uncertainty\_pct INT,  FOREIGN KEY (country\_id) REFERENCES countries(country\_id),  FOREIGN KEY (ghg\_indicator\_id) REFERENCES ghg\_indicator\_details(ghg\_indicator\_id)  ); | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Member 4  Signature | |  |  | | --- | --- | | Full name | Atahan Evintan | | Student no | 820230334 | | Responsibility: | CREATE TABLE energy\_indicator\_details (  energy\_indicator\_id INT PRIMARY KEY,  indicator\_code VARCHAR(50),  indicator\_name VARCHAR(255),  indicator\_description TEXT,  measurement\_unit VARCHAR(50)  );  CREATE TABLE energy\_data (  data\_id INT PRIMARY KEY,  country\_id INT,  energy\_indicator\_id INT,  year INT,  indicator\_value FLOAT,  data\_source VARCHAR(255),  FOREIGN KEY (country\_id) REFERENCES countries(country\_id),  FOREIGN KEY (energy\_indicator\_id) REFERENCES energy\_indicator\_details(energy\_indicator\_id)  ); | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Member 5  Signature | |  |  | | --- | --- | | Full name | Salih Sefer | | Student no | 820230313 | | Responsibility: | CREATE TABLE sustainability\_indicator\_details (  sus\_indicator\_id INT PRIMARY KEY,  indicator\_code VARCHAR(50),  indicator\_name VARCHAR(255),  indicator\_description TEXT  );  CREATE TABLE sustainability\_data (  data\_id INT PRIMARY KEY,  country\_id INT,  sus\_indicator\_id INT  year INT,  indicator\_value FLOAT,  source\_note VARCHAR(255),  FOREIGN KEY (country\_id) REFERENCES countries(country\_id),  FOREIGN KEY (sus\_indicator\_id) REFERENCES sustainability\_indicator\_details(sus\_indicator\_id)  ); | |