CS 306 DATABASE SYSTEMS – PROJECT STEP 1
COUNTRIES, POLLUTIONS AND SOLUTIONS
GROUP 22

- Ahmet Alperen Güngör – 28847
- Anıl Arslan – 29468
- Batuhan İşildak - 29181
- Dediş Atakan Öz – 29414
- Eren Altın – 28986

GITHUB REPOSITORY: https://github.com/batuhanisildak-malwation/CS306_Group22

PROJECT DESCRIPTION

This database has been designed to manage and store data related to different aspects of countries and their environmental aspects, including population statistics, pollution, renewable energy sources, and causes of death. Main aim of this project is to promote awareness on some of the current global disasters by creating a database which has the capability of holding data about countries' death causes and pollution while classifying population into different generations. In addition, usage of renewable energy sources is also stored, which can give a deeper insight on how these energy sources can play a role in countries' mortality rates & causes.

ENTITIES IN THE ER DIAGRAM

- Country: There are 165 countries in this database, where each country as 4 attributes:

    1. id: Is the primary key - uniquely identifies countries.
    2. continentId: Foreign key that refers to continent table.
    3. name: Name of the country.
    4. isoCode: ISO code of the country.

- Continent: There are 6 continents: Africa, Asia, Europe, Oceania, North America and South America. Each continent has 3 attributes:

    1. id: Is the primary key – uniquely identifies continents.
    2. name: Name of the continent.
    3. isoCode: ISO code of the country.

- Pollution: Represents different types of pollutions that endangers our world. Each pollution information has 6 attributes:
    1. id: Is the primary key – uniquely identifies pollution information.
    2. countryId: Foreign key that refers to pollution table.
    3. type: Type of the pollution.
    4. amount: Amount of pollution (unit of this amount is given at another attribute).
    5. amountUnit: Unit of pollution.
    6. yearId: Foreign key that refers to year table.

- Death: Represents deaths, with causations and quantities. Has 5 attributes:
    1. id: Is the primary key – uniquely identifies death information.
    2. count: Number of deaths (cause of death is given as another entity).
    3. yearId: Foreign key that refers to year table.
    4. countryId: Foreign key that refers to country table.
    5. deathCauseId: Foreign key that refers to death cause table.

- Population: Represents number of people living at different countries. Has 4 attributes:
    1. id: Is the primary key – uniquely identifies population information.
    2. amount: Numerical quantity of population.
    3. countryId: Foreign key that refers to country table.
    4. yearId: Foreign key that refers to year table.

- Renewable Energy Source: Represents different types of renewable energy sources that are used to tackle environmental issues. Has 2 attributes:
    1. id: Is the primary key – uniquely identifies renewable energy sources.
    2. name: Names of the renewable energy sources.

- Year: Represents different years. Has 2 attributes:
    1. id: Is the primary key – uniquely identifies each year.
    2. value: Numerical value of the year.

- Death Cause: Causes of the different types of deaths. Has 2 attributes:
    1. id: Is the primary key – uniquely identifies death causes.
    2. name: Names of the death causes.

RELATIONSHIPS IN THE ER DIAGRAM

- uses: Relation between renewable country and energy sources, showing how much (percentage) a country uses a specific renewable energy source.
    - “id” is the primary key.
    - Many to many relationship.

- includes: Relation between country and continent, showing the continents where the countries are at.
    - Has key constraint, many to one relationship from country point of view.

- has1: Relation between country and population, showing the number of people living at a country.
    - Population is a weak entity.

    - Has key constraint, one to many relationship from county point of view.
    - Also has participation constraint for population.

- has2: Relation between country and pollution, showing the specifications about pollutions at a country.
    - Pollution is a weak entity.
    - Has key constraint, one to many relationship from country point of view.
    - Also has participation constraint for pollution.

- experiences: Relation between country and death, describing the number & type of deaths countries experience.
    - Death is a weak entity.
    - Has key constraint, one to many relationship from country point of view.
    - Also has participation constraint for death.

- leads: Relation between death and death cause, describing the causes that lead to death.
    - Death is a weak entity.
    - Has key constraint, one to one relationship.
    - Also has participation constraint for death.

- differsBy1: Relation between death and year, differentiating death data by year information.
    - Death is a weak entity.
    - Has key constraint, one to one relationship.
    - Also has participation constraint for death.

- differsBy2: Relation between pollution and year, differentiating pollution data by year information.
    - Pollution is a weak entity.
    - Has key constraint, one to one relationship.
    - Also has participation constraint for pollution.

- differsBy3: Relation between population and year, differentiating population data by year information.
    - Population is a weak entity.
    - Has key constraint, one to one relationship.
    - Also has participation constraint for population.

CSV FILES

We have 8 csv files: country.csv - continent.csv - population.csv - pollution.csv - year.csv – renewableEnergySources.csv - death.csv – deathCauses.csv. We cleaned the dataset by removing NA values and duplicate information. We transformed the specific information into respective id's that we use on our database system

ENTITY – RELATIONSHIP DIAGRAM