CS306 Group Project – Step 2

Group Number 16

Berke Ceylan	27895
İsmail Çakmak	29496
Nisa Erdal	28943
Salih Kaya	27890
Yunus Topçu	28880

Can be accessed at https://github.com/Salihmaya/CS306_Group16

Tobacco Consumption: Causes and Effects

Introduction to Step 2

In this step of the project, the final form of the ER Diagram was converted into a relational model and tables were created for each of the prepared ".csv" files. Tables were created in SQL by each of the group members for each of the five tables, combined with the "countries.csv" dataset. For the purposes of familiarizing with MySQL databases in a user friendly way, MySQL Workbench environment was utilized to write SQL statements, to import respective ".csv" files and to export "log" files for each process.

Revisions for the "countries.csv" File

During the import of the "cigarette_advertisements.csv", a mismatch between the iso_code attributes of the said file and "countries.csv" has been spotted. "countries.csv" had 113 different iso_code entries while the "cigarette_advertisements.csv" had 197 different iso_codes entries this caused a violation of the foreign key constraints of the table.

To fix this issue, "countries.csv" was updated to include all 197 different iso_codes and the problem was resolved.

Workflow

- For each of the following datasets:
    - Lung cancer death rates
    - Share of deaths that are attributed to smoking
    - Government health expenditure as a share of GDP
    - Support to help quit smoking
    - Enforcement of bans on tobacco advertising

    - ".sql" files were written to create two tables -for the dataset and the "countries.csv" dataset- deriving from the ER Diagram.

    - A Schema was created in MySQL Workbench and both tables were created by executing the SQL statements.

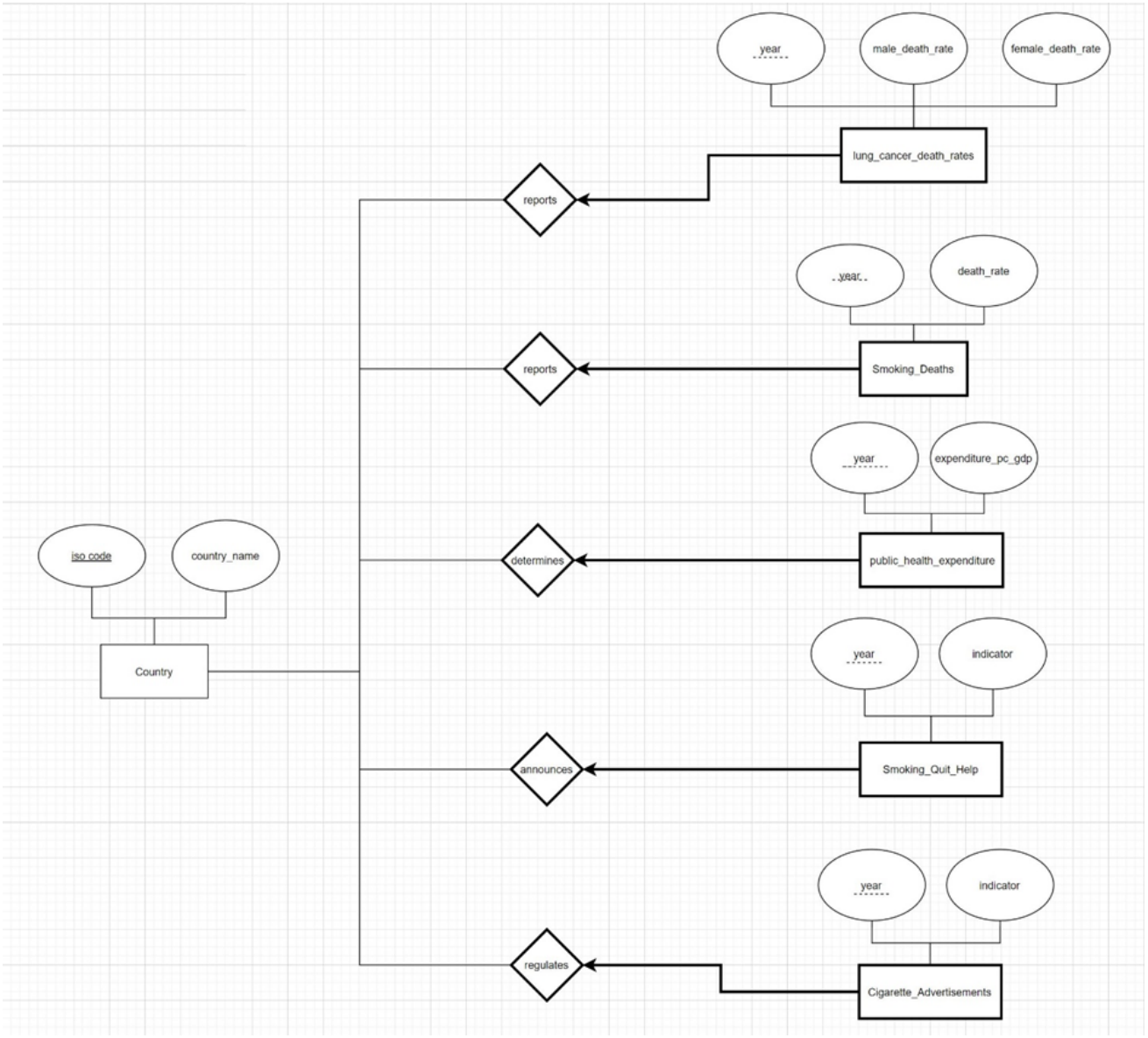
    - Through the "import .csv" function of MySQL Workbench, both datasets were loaded onto the tables.

    - In order to check for any errors during the table creation and / or the ".csv" imports, a ".log" file was exported from MySQL Workbench.

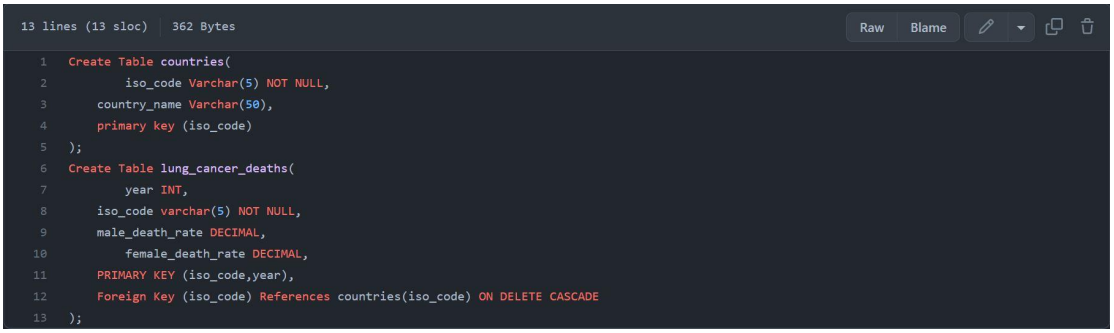
    - Both the ".sql" and ".log" files were committed to the project GitHub folder, under the folder name of "SQLFiles".

SQL Statements for Each Dataset

For reference purposes, here is the ER Diagram from Step 1.



lung_cancer_deaths.sql



```
13 lines (13 sloc)  362 Bytes                                                    Raw  Blame

1  Create Table countries(
2          iso_code Varchar(5) NOT NULL,
3      country_name Varchar(50),
4      primary key (iso_code)
5  );
6  Create Table lung_cancer_deaths(
7          year INT,
8      iso_code varchar(5) NOT NULL,
9      male_death_rate DECIMAL,
10         female_death_rate DECIMAL,
11     PRIMARY KEY (iso_code,year),
12     Foreign Key (iso_code) References countries(iso_code) ON DELETE CASCADE
13 );
```

The first SQL statement creates a "countries" table with columns for the country's ISO code and name, where the ISO code is defined as the primary key.

The second SQL statement creates a "lung_cancer_deaths" table to represent the relation set between countries and lung cancer deaths, with columns for the year, ISO code, and male/female death rates. The table's primary key is a combination of the ISO code and year columns.

A foreign key constraint is added to the "lung_cancer_deaths" table, referencing the "countries" table's ISO code column. This ensures that records can only be added to the "lung_cancer_deaths" table for countries that already exist in the "countries" table, and that if a country is deleted from the "countries" table, any associated lung cancer death records will also be deleted.

death-rate-smoking.sql

```
13 lines (13 sloc)  514 Bytes                                                    Raw  Blame

1  USE CS306_Project;
2  CREATE TABLE countries(iso_code VarChar(5) NOT NULL,
3                         country_name VarChar(50),
4                         primary key (iso_code)
5                         );
6  CREATE TABLE death_rate_smoking(iso_code VarChar(50),
7                                  year int,
8                                  death_rate double,
9                                  foreign key(iso_code) references countries(iso_code) ON DELETE CASCADE,
10                                 primary key(year)
11                                 );
12 SELECT * FROM death_rate_smoking;
13 SELECT * FROM countries;
```

The first SQL statement selects the "CS306_Project" database for use, then creates a table called "countries" with columns for the country's ISO code and name, where the ISO code is defined as the primary key.

Next, a table called "death_rate_smoking" is created to represent the relation set between countries and their smoking-related death rates. This table includes columns for the ISO code, year, and death rate, and the primary key is set to the year column.

A foreign key constraint is added to the "death_rate_smoking" table, referencing the "countries" table's ISO code column. This ensures that records can only be added to the "death_rate_smoking" table for countries that already exist in the "countries" table, and that if a country is deleted from the "countries" table, any associated smoking-related death rate records will also be deleted.

public_health_expenditure.sql

12 lines (12 sloc) | 340 Bytes — Raw | Blame

```sql
Create Table countries(
        iso_code Varchar(5) NOT NULL,
    country_name Varchar(50),
    primary key (iso_code)
);
Create Table public_health_expenditure(
        iso_code varchar(5) NOT NULL,
        year INT,
    expenditure_pc_gdp double,
    primary key (iso_code, year),
    foreign key (iso_code) references countries(iso_code) ON DELETE CASCADE
);
```

The first SQL statement creates a "countries" table with columns for the country's ISO code and name, where the ISO code is defined as the primary key.

The second SQL statement creates a "public_health_expenditure" table to represent the relation set between countries and their public health expenditure, with columns for the ISO code, year, and expenditure as a percentage of GDP. The primary key is defined as a combination of the ISO code and year columns.

A foreign key constraint is added to the "public_health_expenditure" table, referencing the "countries" table's ISO code column. This ensures that records can only be added to the "public_health_expenditure" table for countries that already exist in the "countries" table, and that if a country is deleted from the "countries" table, any associated public health expenditure records will also be deleted.

smoking_quit_help.sql

15 lines (12 sloc) | 346 Bytes — Raw | Blame

```sql
create Table countries(
        iso_code Varchar(5) NOT NULL,
    country_name Varchar(50),
    primary key (iso_code)
);


create Table smoking_quit_help(

        iso_code varchar(5) NOT NULL,
    year INT,
    indicator double,
    primary key(iso_code, year),
    foreign key (iso_code) references countries(iso_code) ON DELETE CASCADE
    );
```

The first SQL statement creates a "countries" table with columns for the country's ISO code and name, where the ISO code is defined as the primary key.

The second SQL statement creates a "smoking_quit_help" table to represent the relation set between countries and their smoking cessation support indicators, with columns for the ISO code, year, and indicator value. The primary key is defined as a combination of the ISO code and year columns.

A foreign key constraint is added to the "smoking_quit_help" table, referencing the "countries" table's ISO code column. This ensures that records can only be added to the "smoking_quit_help" table for countries that already exist in the "countries" table, and that if a country is deleted from the "countries" table, any associated smoking cessation support indicator records will also be deleted.

cigarette-advertisements.sql

19 lines (14 sloc) | 395 Bytes — Raw | Blame

```sql
CREATE DATABASE cs306_project;
USE cs306_project;

Create TABLE Countries(
        iso_code Varchar(5) NOT NULL,
    country_name Varchar(50),
    primary key (iso_code)
);

CREATE TABLE Cigarette_advertisements(
        iso_code VARCHAR(3) NOT NULL,
    year INT,
    ban_indicator INT,

    PRIMARY KEY (iso_code, year),
    FOREIGN KEY (iso_code) REFERENCES Countries(iso_code)  ON DELETE CASCADE
);


```

The first SQL statement creates a new database named "cs306_project" and switches the current database context to it.

The next SQL statement creates a "Countries" table with columns for the country's ISO code and name, where the ISO code is defined as the primary key.

Then, a "Cigarette_advertisements" table is created to represent the relation set between countries and their cigarette advertisement ban indicator values, with columns for the ISO code, year, and ban indicator value. The primary key is defined as a combination of the ISO code and year columns.

A foreign key constraint is added to the "Cigarette_advertisements" table, referencing the "Countries" table's ISO code column. This ensures that records can only be added to the "Cigarette_advertisements" table for countries that already exist in the "Countries" table, and that if a country is deleted from the "Countries" table, any associated cigarette advertisement ban indicator records will also be deleted.

References

Our World in Data. (n.d.). *Enforcement of bans on tobacco advertising*. Retrieved March 25, 2023, from https://ourworldindata.org/grapher/enforcement-of-bans-on-tobacco-advertising

Our World in Data. (n.d.). *Share of deaths from smoking*. Retrieved March 25, 2023, from https://ourworldindata.org/grapher/share-deaths-smoking

Our World in Data. (n.d.). *Lung cancer deaths per 100,000 by sex, 1950-2002*. Retrieved March 25, 2023, from https://ourworldindata.org/grapher/lung-cancer-deaths-per-100000-by-sex-1950-2002?country=ESP-1~ESP~ESP~USA-1~USA~USA

Our World in Data. (n.d.). *Support to help quit tobacco use*. Retrieved March 25, 2023, from https://ourworldindata.org/grapher/support-to-help-to-quit-tobacco-use

Our World in Data. (n.d.). *Public health expenditure (% of GDP)*. Retrieved March 25, 2023, from https://ourworldindata.org/grapher/public-health-expenditure-share-gdp-owid

MySQL Workbench. (Version 8.0.32). [Computer software]. Oracle Corporation. Retrieved from https://www.mysql.com/products/workbench/ on March 25, 2023.