

Causes of Death: STEP 2

Group AKA

Alaa Almouradi 00030123

Filzah Azeem 00030138

Fatma Khalil 00029999

M Safwan Yasin 00030037

Saleh Alshurafa 00030002

CS306 Spring 22-23

[GitHub Repository](#)

NOTE: All files and folders mentioned in this document can be found in our repository

Countries:

CountriesTable.csv

We completely changed the CountriesTable.csv. The older version of it contained 182 countries only and was missing a lot of countries that were in all of our collected data. Countries missing included: Qatar, Palestine, Syria, Somalia, Singapore, and more. Therefore, we replaced it with a new set that contains more countries and still follows the ISO3 standards. The new dataset contains 249 countries with their alpha-3 ISO codes.

SQL:

The "Countries" entity set was made into a stand alone table "countries" because each country can exist on its own.

***Note: The following CREATE statement exists in all the table creating ".sql" files**

```
CREATE TABLE countries (  
    countryName CHAR(100),  
    countryCode CHAR(5),  
    PRIMARY KEY (countryCode)  
);
```

Food Products:

Alaa Almouradi Table STEP2(Food).csv:

The original dataset was too big, therefore it was cut down so that it started from the same year the other datasets started. It was also cut down by changing the food types from specific types to more general types. A few countries were also removed because they did not exist in the new countries table. All this cleaning was done via filtering in excel.

Alaa Almouradi SQL STEP2(Food).sql:

The "Food Products" entity set and the "Manage" relationship set were made into one table ("foodManagementPerCountry") since the "Food Products" set has total participation constraint therefore it will always be in that relationship set. The "Food Products" set is a weak entity with the identifying relationship being "Manage" and the owner entities being from the "Countries" entity set, this is reflected in SQL by using "FOREIGN KEY" from "countries" table in the

"foodManagementPerCountry". The fact that a weak entity cannot exist without the owner entity is reflected in SQL by using "ON DELETE CASCADE" with the foreign key in the "foodManagementPerCountry" table.

```
CREATE TABLE foodManagementPerCountry (  
    countryCode CHAR(5),  
    typeOfFood CHAR(150),  
    year INT,  
    domestic_supply_quantity_1000_tonnes DOUBLE,  
    production_1000_tonnes DOUBLE,  
    FOREIGN KEY (countryCode)  
        REFERENCES countries (countryCode)  
        ON DELETE CASCADE,  
    PRIMARY KEY (countryCode, typeOfFood, year)  
);
```

Alaa Almouradi Logs STEP2(Food).log:

The logs show the successful creation of the tables and the imports of the data and the results of the select statements after each import show that the number of imported rows match the number of the csv file.

Mental Disorders:

M Safwan Yasin Table STEP2(Mental).csv

Since the entity named mental disorders is a weak entity, every record present in the entity is supposed to be linked to a record in the entity named country. However, this CSV file contained some records with country codes that were not present in the countries table since they were regions present inside countries, such as England, Scotland and Wales. Therefore, these records were removed from the mental disorders entity.

M Safwan Yasin SQL STEP2(Mental).sql

This file contains the create table statement for the entity set called "Mental Disorders" and the relationship set "Record". The table is named "mentalIllnessPerCountry" and it represents both the entity set and the relationship set since "Mental Disorders" is a weak entity that only exists when participating in the identifying relationship set (Record) with the owner entity i.e Countries. Using "countryCode" from the "countries" table as a foreign key in the "mentalIllnessPerCountry" table demonstrates the presence of an identifying relationship set.

Lastly the fact that a weak entity cannot exist without the presence of the owner entity being present is shown by using the “ON DELETE CASCADE” statement along with the foreign key in the “mentalIllnessPerCountry” table.

Name of entity set: Mental Disorders

Name of relationship set: Record

Name of table created: mentalIllnessPerCountry

SQL:

```
CREATE TABLE mentalIllnessPerCountry(  
    countryCode CHAR(5),  
    Year INT,  
    DALY DOUBLE,  
    prevalenceEating DOUBLE,  
    prevalenceDepressive DOUBLE,  
    FOREIGN KEY (countryCode)  
    REFERENCES countries(countryCode) ON DELETE CASCADE,  
    PRIMARY KEY (countryCode, Year)  
);
```

M Safwan Yasin log file STEP2(Mental).log

This is the log file that demonstrates the creation and import of data into the table called mentalIllnessPerCountry which is based upon the entity named Mental Disorders. The last entry in the log shows the presence of 6120 records in this table which matches with the number of records present in the CSV file.

Deaths:

Fatma Khalil Table STEP 2 (Death).csv

Some entities were not present in the SQL table due to the fact that they had no country code or were not present in the countries table, they were removed from the CSV files. That was necessary because the Death entity is a weak one, so the entities in it cannot exist if they are not present in the countries table.

Fatma Khalil MySQL Table STEP2 (Death).sql

I have converted the Death entity set and the Report relationship set into a table. The file contains the create table statements for the Countries table and the Deaths table. Since the entity is weak and cannot exist without a country code, we made the countryCode as a foreign key

constraint from the Countries table in the Deaths table, so when a country is dropped from the Countries table, it is also dropped from the Deaths table. The variable countryCode is a country's code, Year is the year where the deaths happen, cause is the cause of death, and numberOfDeaths is how many people died of a particular cause in a given year.

Name of entity set: Deaths

Name of relationship set: Report

Name of table created: Deaths

```
CREATE TABLE Deaths (  
    countryCode CHAR(5),  
    Year INT,  
    cause CHAR(255),  
    numberOfDeaths INT,  
    FOREIGN KEY (countryCode)  
        REFERENCES countries (countryCode)  
        ON DELETE CASCADE,  
    PRIMARY KEY (countryCode , Year , cause)  
);
```

Fatma Khalil Logs STEP2(Death).log

The log files contain not only the create statements, but also the imports from the CSV file into the table created in MySQL, at the end of the import, we have a table with 140760 rows. The multiple '*SELECT * cs306project.deaths;*' statements were for checking the import progress.

Substance Abuse:

Filzah Azeem Table STEP2(Substance).csv

Previously this table contained records with country codes that did not exist in CountriesTable.csv, so that data was removed as this is a weak entity set to ensure all data gets imported into the mySQL table.

Filzah Azeem SQL STEP2(Substance).sql

In this file I created the table 'substanceAbusePerCountry' by combining the entity set 'Substance Abuse' and the relationship set 'Calculate'. This table contains a foreign key 'countryCode' from the 'countries' table because 'Substance Abuse' is a weak entity set, and the foreign key constraint 'ON DELETE CASCADE' will delete anything in the 'substanceAbusePerCountry' table if it does not exist in the 'countries' table, because a weak entity cannot exist without a relationship with owner entity.

Name of entity set: Substance Abuse

Name of relationship set: Calculate

Name of table created: substanceAbusePerCountry

SQL:

```
CREATE TABLE substanceAbusePerCountry(  
    countryCode CHAR(5),  
    Year INTEGER,  
    prevalenceAlcoholUseDisorder_Percent DOUBLE,  
    prevalenceDrugUseDisorder_Percent DOUBLE,  
    FOREIGN KEY (countryCode)  
    REFERENCES countries(countryCode) ON DELETE CASCADE,  
    PRIMARY KEY (countryCode, Year)  
);
```

Filzah Azeem Logs STEP2(Substance).log

This is the log file that contains the create statements and shows the import of data from the csv files. At the end we see 6120 rows returned from the substanceAbusePerCountry Table, which is the same number of records in the Substance csv file.

Air Pollution:

Saleh Alshurafa SQL Files STEP2 (Air Pollution)

This folder contains 2 sql files: **Saleh Alshurafa SQL File(1) STEP2.sql** and **Saleh Alshurafa SQL File(2) STEP2.sql**.

In **File(1)**, you can find the SQL statements to create my tables. Two tables were created: Countries and AirPollution. The table Countries models the Countries entity. The AirPollution table models the weak entity Air Pollution and its identifying relationship, Monitor. The weak entity and its relationship are combined in one table because weak entities have total participation. Additionally, weak entities only exist with their owner entities (Countries entity) hence why countryCode is taken as a FOREIGN KEY from countries table. If the owner entity is deleted the weak entity is deleted with it and to demonstrate that ON DELETE CASCADE was used.

In **File(2)**, ALTER TABLE SQL statements were applied. These were used to change the names of the columns in the AirPollution table. Initially in **File(1)** when I first created the AirPollution table, I had used test names for the columns. Later, after importing all the data, I changed the names of the columns to shorthand names of the attribute names for better accuracy and understanding of the table.

Name of the entity set: Air Pollution (weak entity set)

Name of relationship set: Monitor

Name of the table created: AirPollution (combines Air Pollution and identifying relationship, Monitor)

```
CREATE TABLE AirPollution (  
    countryCode CHAR(5),  
    Year INT,  
    totalPollution DOUBLE,  
    meanAnnualExposure DOUBLE,  
    levelsExceedingWHO DOUBLE,  
    PRIMARY KEY (countryCode, Year),  
    FOREIGN KEY (countryCode) REFERENCES Countries(countryCode) ON DELETE  
    CASCADE  
);
```

Alter Table Statements

```
ALTER TABLE airpollution CHANGE totalPollution `Total including LUCF in tonnes`  
DOUBLE;  
ALTER TABLE airpollution CHANGE meanAnnualExposure `PM2.5, mean annual exposure  
(µg/m³)` DOUBLE;  
ALTER TABLE airpollution CHANGE levelsExceedingWHO `PM2.5, pop. exposed to levels >  
WHO guideline value (% TOT)` DOUBLE;
```

Saleh Alshurafa Log Files STEP2 (Air Pollution).log

This file shows the logs of my database. They demonstrate the success of the creation of the tables and the importing of data from the CSV files to the corresponding tables. The successful creation of the tables can be seen at the lines starting with [37, 17:57:08] and [38, 17:57:08]. The success of the import of the data from the CountriesTable.csv (from STEP1) to the Countries table can be seen at the line starting with [8, 17:59:13] which shows that the Countries table now has 249 entries, the same amount of records in CountriesTable.csv. The success of the import of the data from Saleh Alshurafa Table STEP1 (Air Pollution).csv to the AirPollution table can be seen at the line starting with [16, 18:02:40] which shows that the AirPollution table now has 2232 entries, the same amount of records in Saleh Alshurafa Table STEP1 (Air Pollution).csv. Finally, the success of the ALTER TABLE statements can be seen in the lines starting with [17, 18:05:23], [18, 18:05:23], and [19 18:05:23].