

ECE 9014 Final Project

Deliverable 1

**US Electricity database**

– Student Number

Yuanpeng Wang – 251343797

Table of Contents

[Dataset 1](#_Toc149172066)

[Description 1](#_Toc149172067)

[Business Rules 5](#_Toc149172068)

[Entity Relationship Diagrams 5](#_Toc149172069)

[SQL Implementation 6](#_Toc149172070)

[Relational Model 6](#_Toc149172071)

[Screenshots 6](#_Toc149172072)

# Dataset

### Description

**What do the tables contain? What do the columns mean?**

The 1st table contains the fundamental information of the regulatory of each state in US including regulatory ID, regulatory name, and office location, telephone number, postcode.

In columns, regulatory ID is the abbreviation of each state and its regulatory; Name is the name of the regulatory of each state. The address, contact, and postcode are the location, telephone number, and postcode of the regulatory office, respectively.

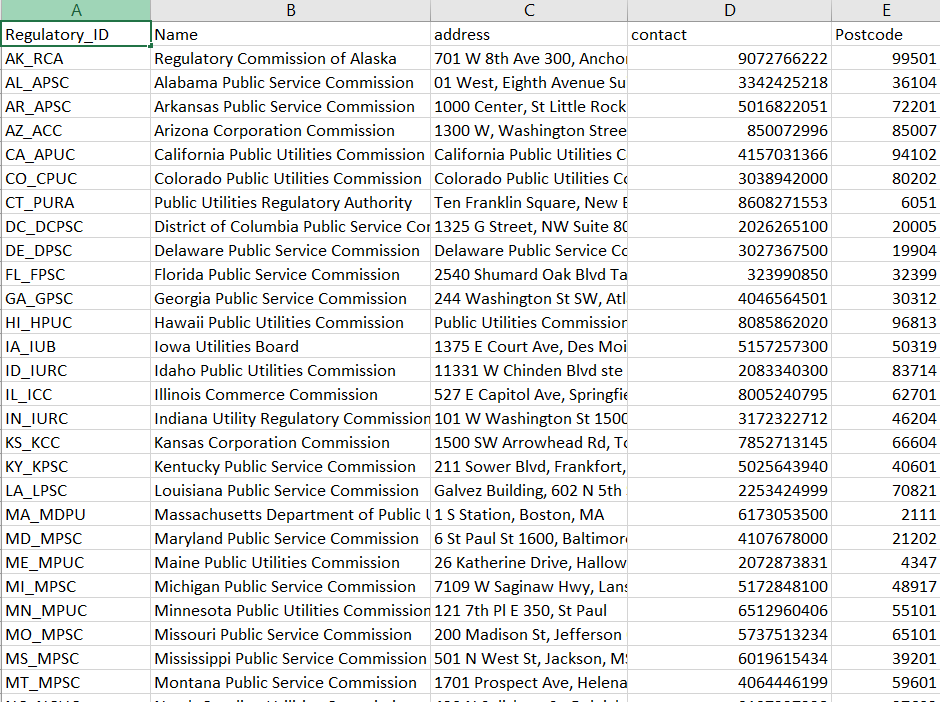


Fig. 1 regulatory table

The 2nd table contains information regarding electricity usage in four different fields, residential, industry, commercial, and transportation, in each state from 2011 to 2021, including group ID, regulatory ID, year, electricity usage, and unit price.

In columns, the group ID is the abbreviation of each state and its field. RS stands for residential. IS stands for industry. CM stands for commercial. TS stands for transportation. In\_year means the year of the data for each state. Electricity usage means the total usage of each state. The unit price is the price for using one cent of electricity.

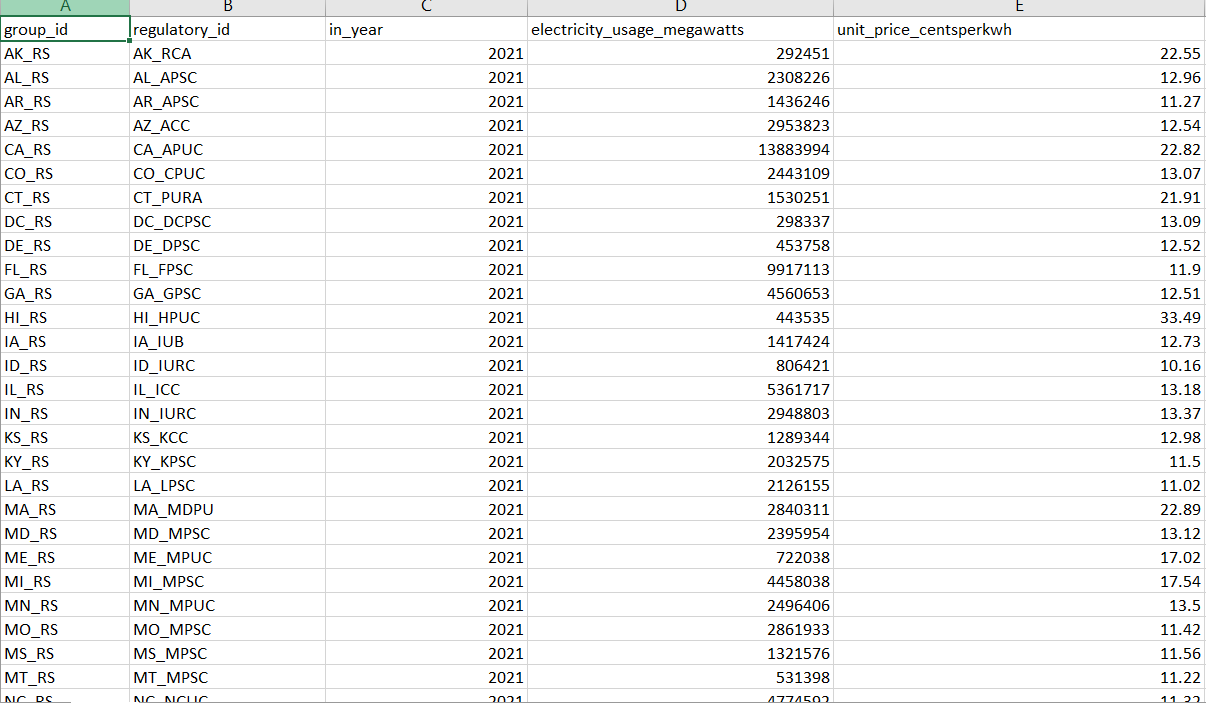


Fig. 2 Electricity usage table

The 3rd table contains information about the user numbers in each state.

In columns, the user groups are the abbreviation of each state and its user number.

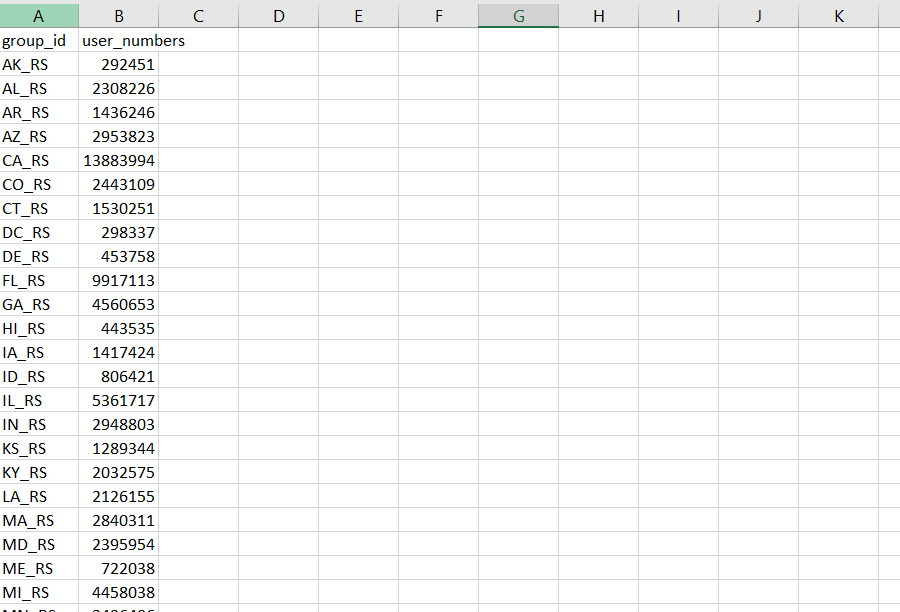


Fig. 3 User numbers table

The 4th table contains information regarding fossil fuel consumption statistics including fuel type, state, year, regulatory ID, and fossil consumption.

In columns, fossil fuel includes coal, petroleum, and natural gas. Us\_state is the abbreviation of each state. Year in terms of the year of the data. Fossil consumption is the amount of fuel units in tons.

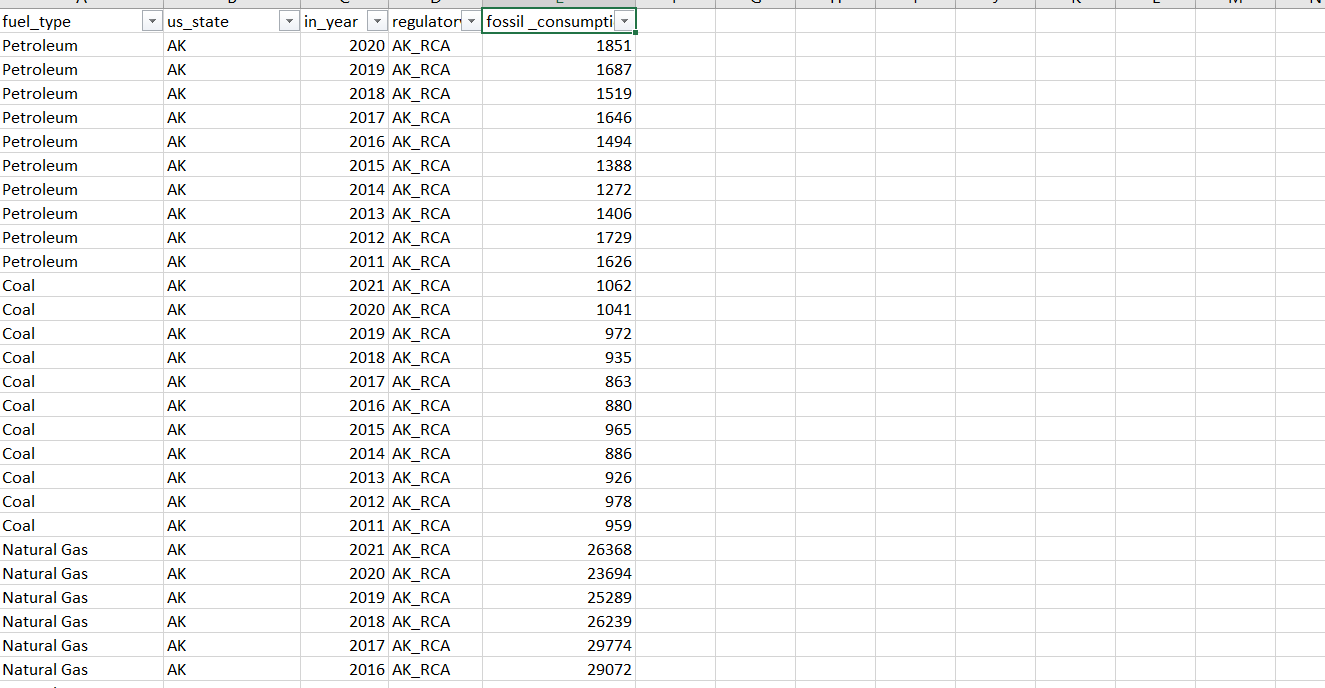


Fig. 4 fossil consumption table

The 5th table contains information about the power plants in US. Including fuel type, state, year, generator, facilities, named capacity, and summer capacity.

In columns, the fuel type includes Coal, Petroleum Natural Gas, and Hydroelectric Conventional. The fuel type is up to 13 different kinds of fuel. Generator is the number of generators run in the state. Named capacity means the average capacity of each state.

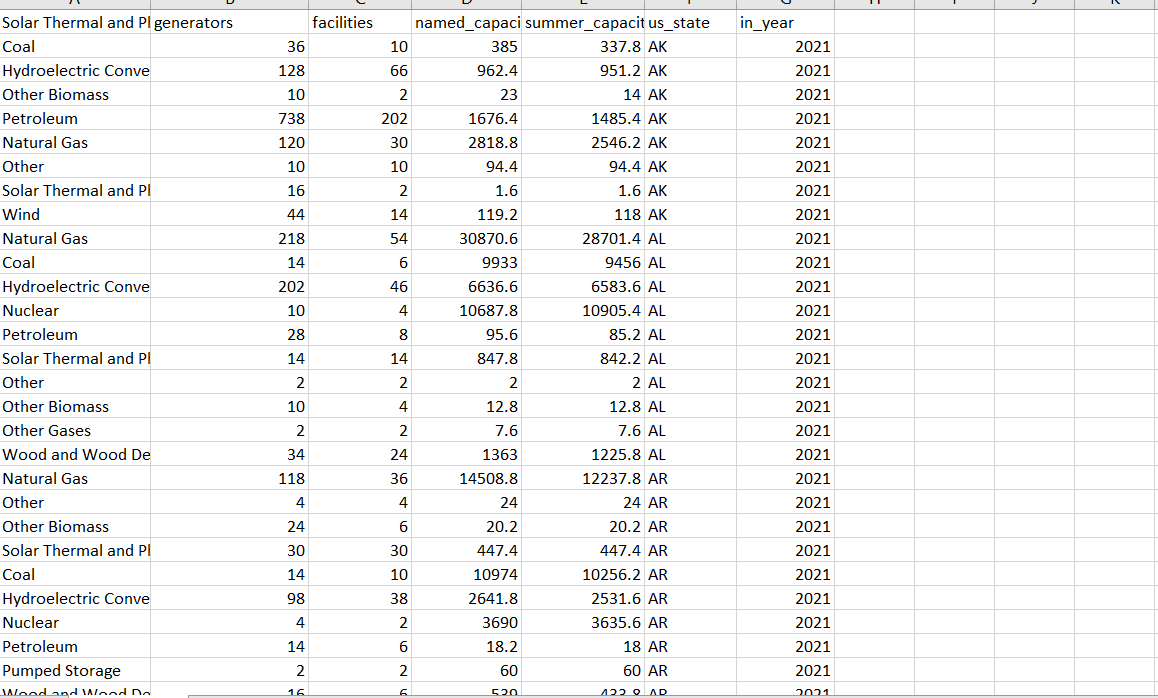


Fig. 5 power plants table

The 6th table shows the electricity production of each state. Including fuel type, regulatory ID, state, year, and the produced amount.

In columns, the types of fuel are the same as the types of fuel in power plants. The produced amount means the average amount of fuel produced by the state in the year.

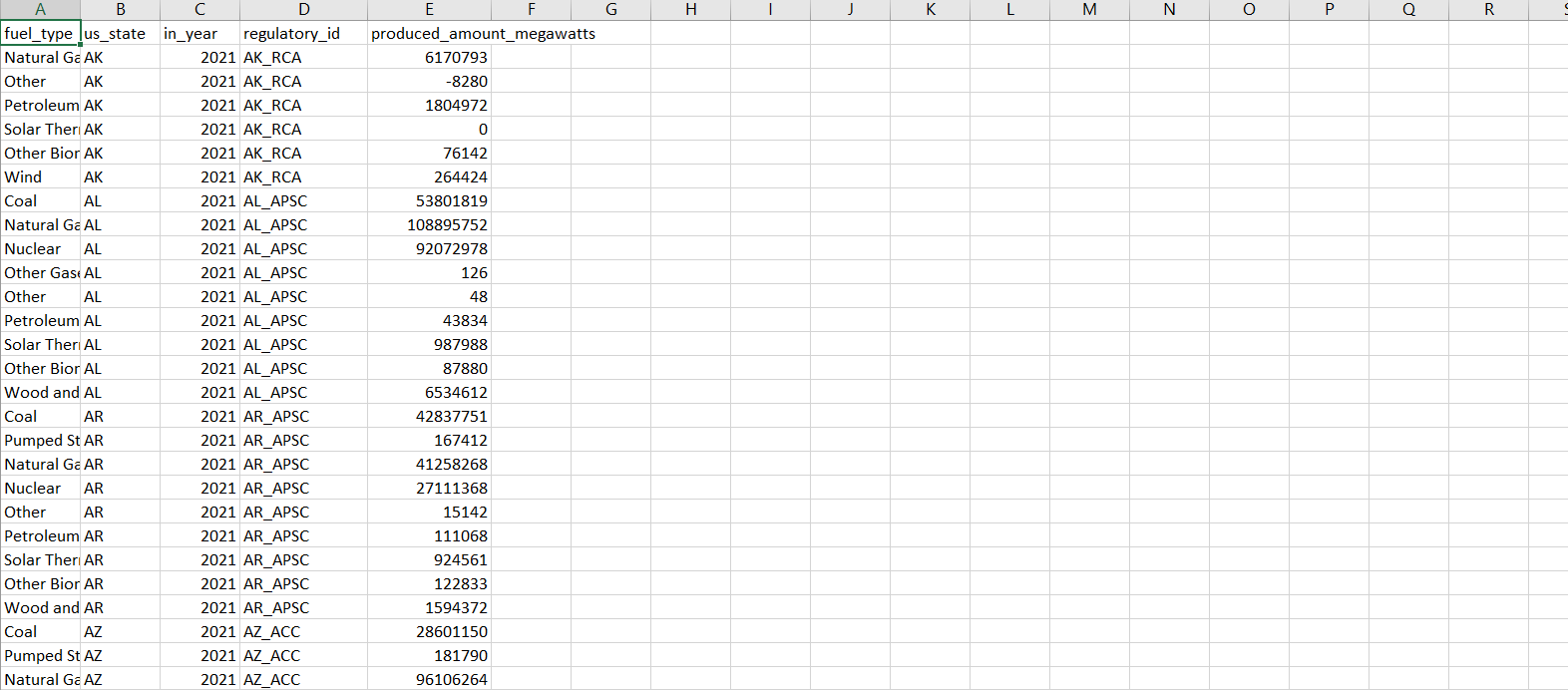
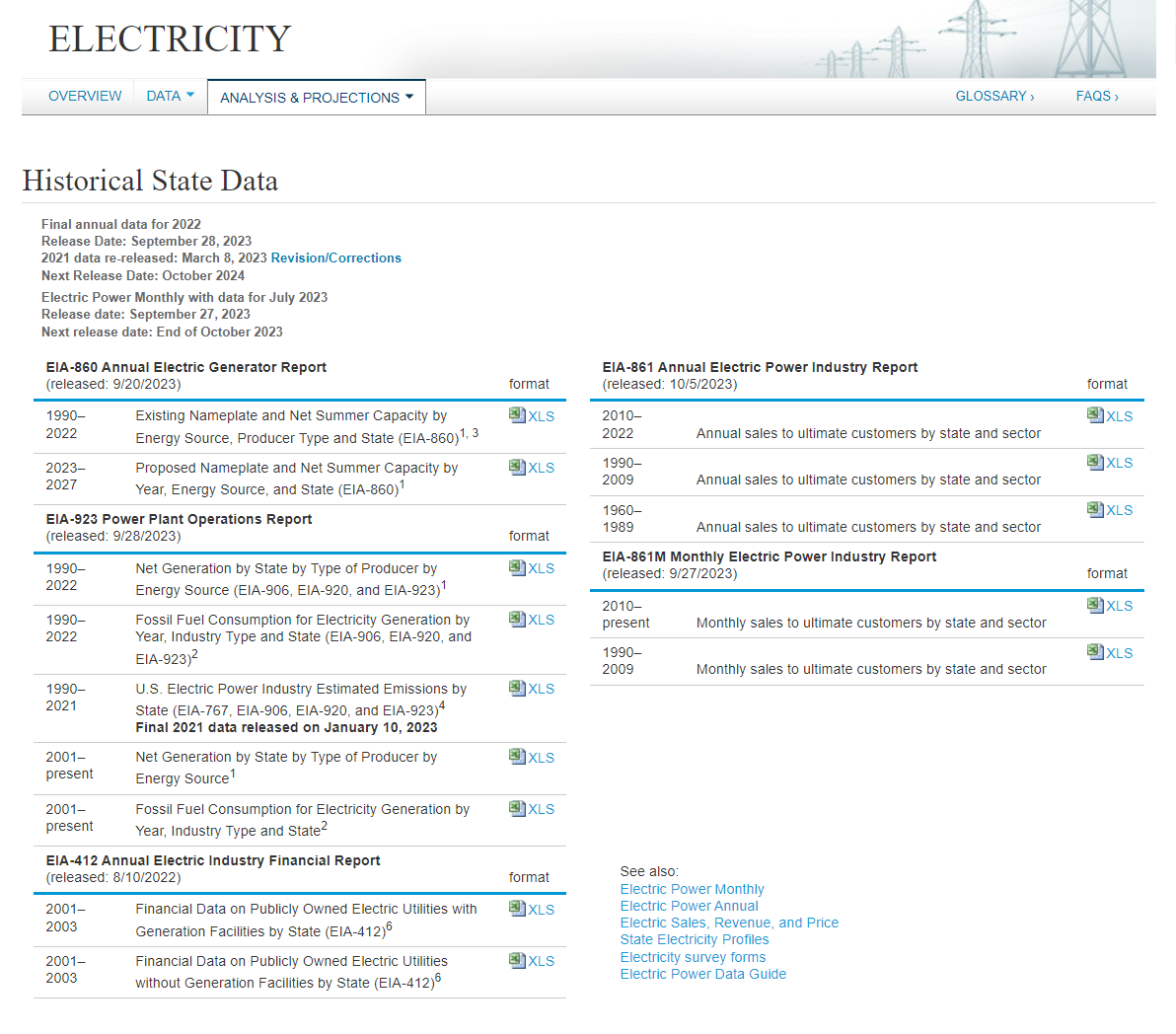


Fig. 6 electricity production table

**Where did the data come from?**

Source from EIA to build US Electricity information database in this project

* The electricity dataset selected for this project was from US Energy Information Administration
* The data in those 5 Excels are massive, tailored data are used for this project
* The dataset used in this project is to design a database to help authorities monitor analyze, manage, and regulate the electricity industry efficiently.



Detail electricity capacity is included in the excel by energy source

Consumer usage information is included

Fig. 7 the source of data

**As a high-level overview, what are the relationships between the tables?**

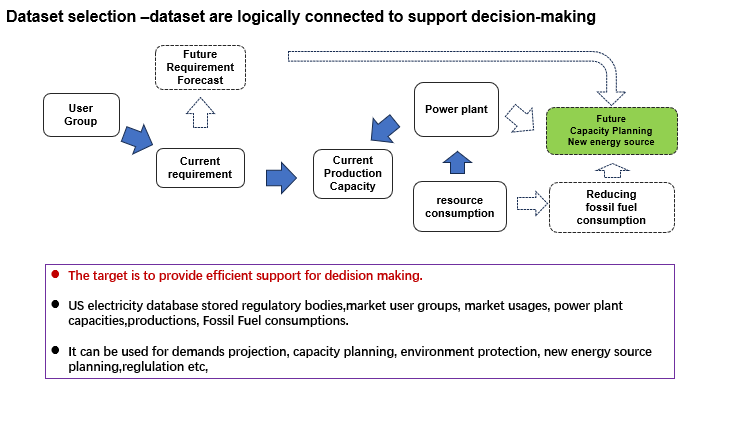


Fig. 8 The entity relationships and their functionality

### Business Rules

* Electricity regulatory bodies monitor, manage, and regulate the whole industry efficiently required to build an electricity industry database to support efficient analysis before decision-making
* Each authority in the states has its name, address, contact, postcode
* The authority in each state monitor electricity market information, like electricity usage, unit price, and market volume， The user group reported that information by different application like residential, industry, commercial, transportation
* The power plant has its existing power capacity and produce electricity based on market requirement, and report to authorities. The power plant also has its planning capacity
* The power plant consumes fossil energy which includes coal, Petaluma, and gas, to produce electricity and will be monitored by authorities
* The power plants produce electricity from 13 energy types, Coal, Petroleum Natural Gas, Hydroelectric Conventional, Wind, Other Biomass, Other, Other Gases, Nuclear, Wood and Wood Derived Fuels, Pumped Storage, Solar Thermal and Solar Thermal, and Geothermal, those will be stored in a database.

### Entity Relationship Diagrams

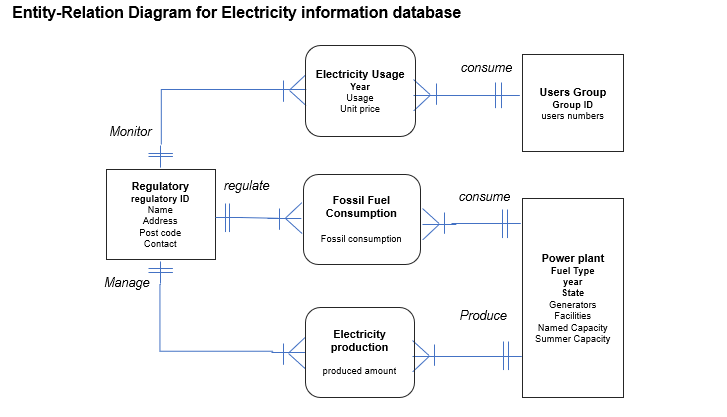


Fig. 9 entity relationship diagrams

# SQL Implementation

### Relational Model

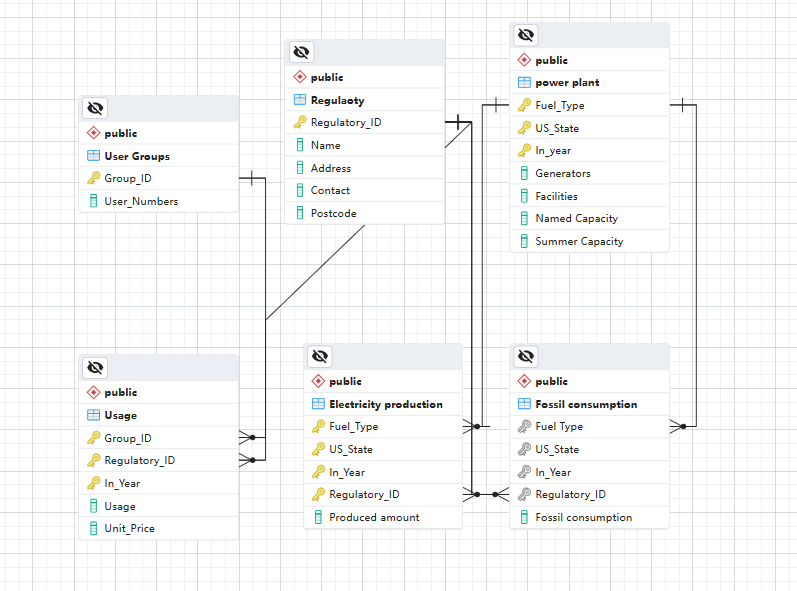


Fig. 10 Relational model

### Screenshots

**Used shell script to implement data from CSV files:**

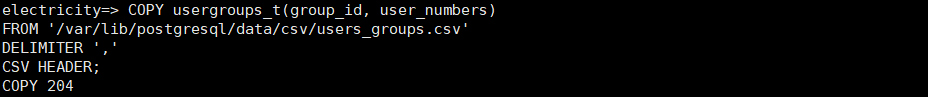


Fig. 11 Implement the user groups table

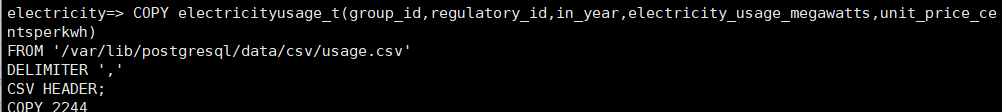


Fig. 12 Implement the electricity usage table

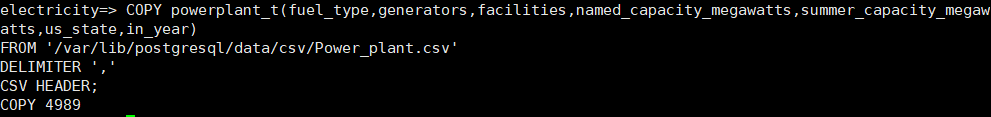


Fig. 13 Implement the power plant

**Populated database:**

****

Fig. 14 The user groups table

****

Fig. 15 The regulatory table

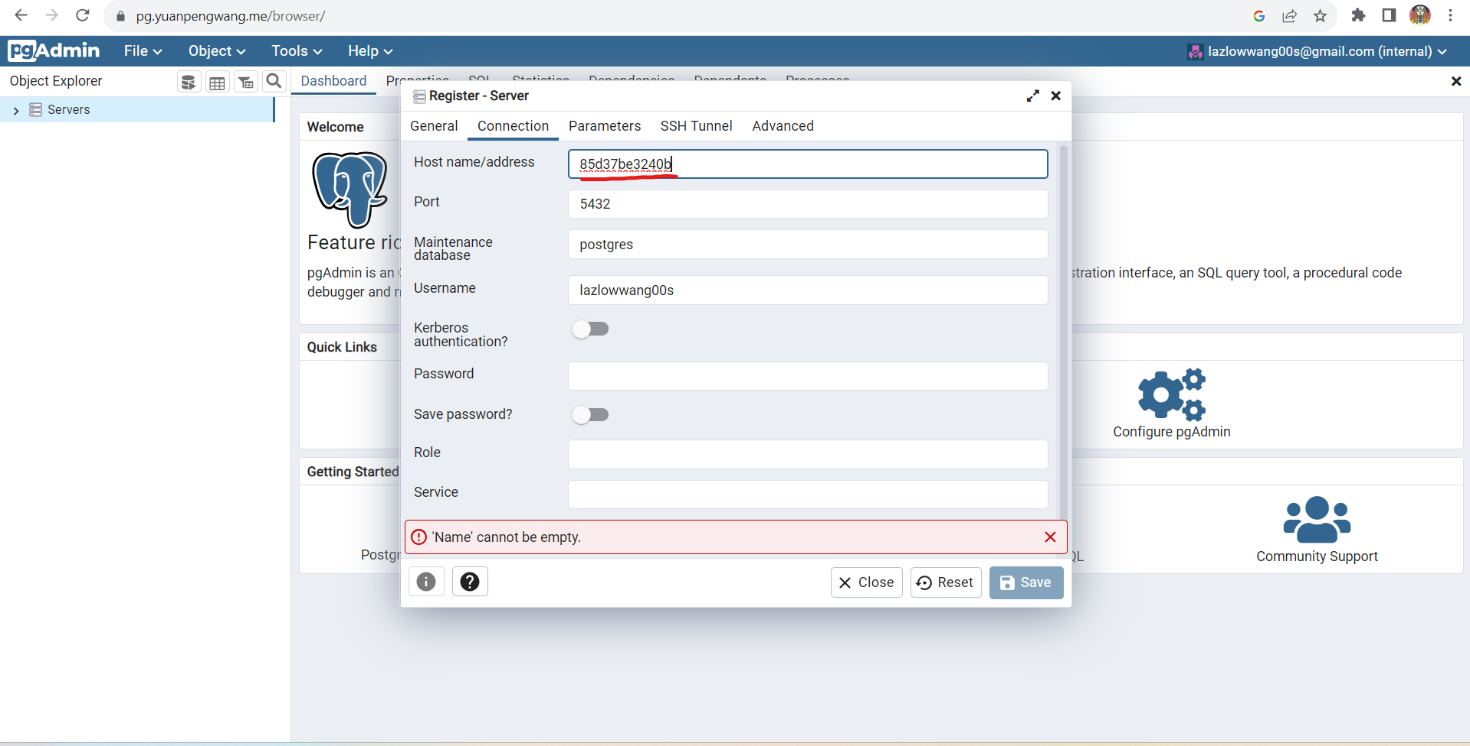


Fig. 16 Pgadmin login and connect to the container

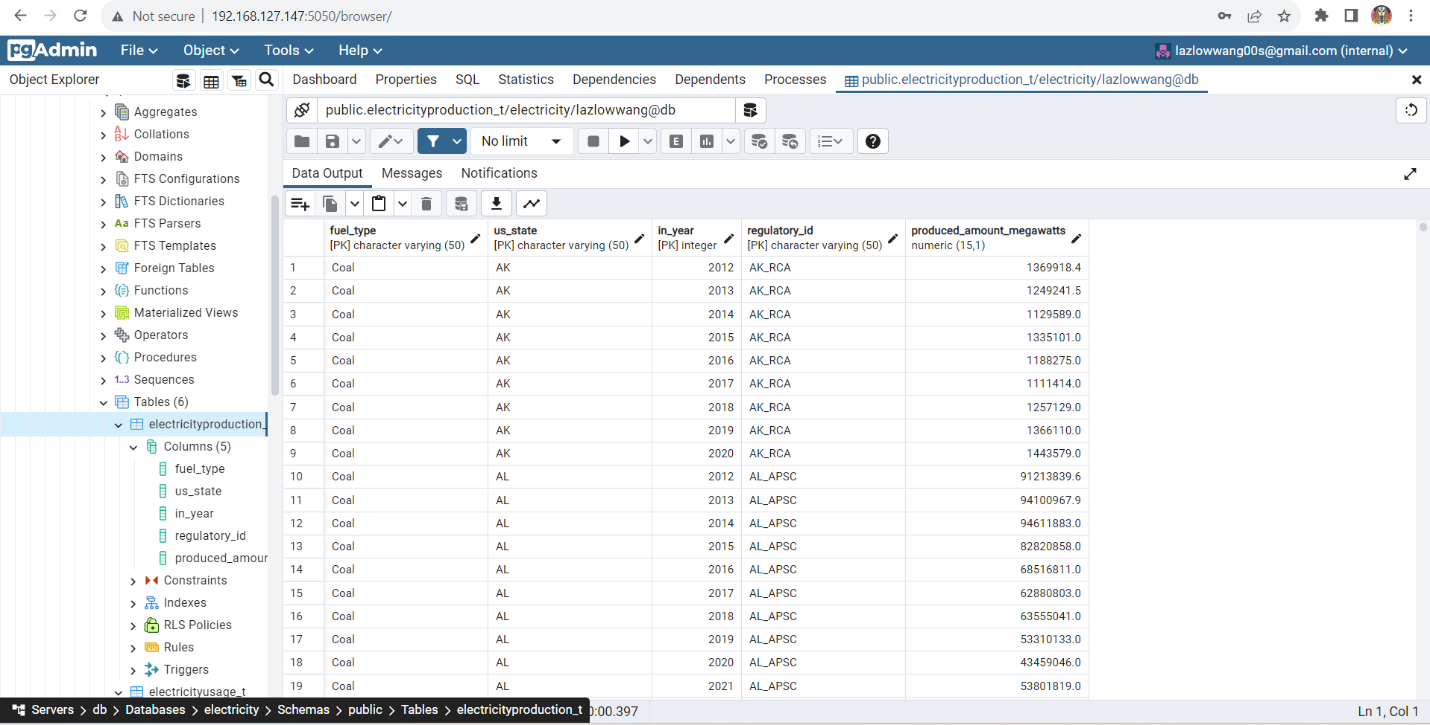


Fig. 17 The Populated table

You can use this link to access my Pgadmin and check my group project.  <https://pg.yuanpengwang.me/>  The email is [lazlowwang00s@gmail.com](mailto:lazlowwang00s@gmail.com) and the password is 123456   Name: db   Hostname: 85d37be3240b Port: 5432

Username: lazlowwang  Password: wwyypp123