0. Name of the group members and CNetID

Linh Dinh: ldinh
Eric Son: yson
Emily Zhang: zhange

1. A brief overview of the final project (200 words maximum)

Our project provides utilizable insights about the energy consumption trends in the U.S. at a state level, and thus promotes more conscious utilization of energy and an awareness of the states' transition to renewable energy. Users can use the search tool with data visualization option to look at their state or choose to compare with another state its energy consumption patterns and prices in 2017 (the latest year available), as well as historical time series of energy consumption by source. We retrieved state-level data for our project from the U.S. Energy Information Administration (EIA) via web scraping and API.

2. The overall structure of the software (1-page maximum)

How to run our software?

requirement.txt contains the required packages to run our program
To install the required packages in a new environment:

- $ chmod +x install.sh
- $ ./install.sh

Run shell script: $ ./run_software.sh
Available options:

1. -d will re-pull the data from web scraping and API
2. -g will recreate graphs from the database

Structure of our software:

There are 3 main components:

1. Data collecting and schema building scripts:
   - webscraper.py
     - Scrapes 2017 data from EIA website. The script crawls the index page and retrieves the datasets of interest for year 2017. (https://www.eia.gov/state/seds/seds-data-complete.php?sid=US) The relevant datasets are then saved in the "data" folder.
   - pull_api.py
     - Leverages EIA's API to retrieve trend data for a longer time period (1960 - 2017), which is not available in a web page format. The relevant datasets are saved in the "data" folder.
   - db.sql
     - sql file that converts data retrieved from web scraping and API into a sql3.lite database.

2. Data visualization/display scripts:
   - make_2017_graph.py

    - Python script that creates graphs showing the total energy consumption pattern by state in 2017. All the graphs are created and saved in the "static/graphs" folder. When a user selects a state, we'll find the correct path of the state's graph to display.
- make_1960_2017_graph.py
    - Python script to create graphs showing the total energy consumption pattern by state for 1960-2017. All the graphs are created and saved in the "static/graphs" folder. When a user selects a state, we'll find the correct path of the state's graph to display.
- make_2017_table.py
    - Python script that queries data on the total energy consumption pattern by state in 2017 from the database we created above. Convert the data into a formatted table to be displayed when a user selects a state.

3. Django scripts:

    Following the standard structure of a Django application, there are 3 main folders:
    - "search": folder contains html template and codes to display the correct visualizations selected by the users.
    - "static": folder contains css style and graphs generated from our data visualization scripts (to be displayed as static components in css structure).
    - "ui": Django default setup and scripts.

3. What the project tried to accomplish and what it actually accomplished (200 words)

Our project provides a visual representation of each states' energy consumption patterns in 2017 and in the last six decades. It also draws visual distinction between the states with varying consumption trends. We initially discussed using several data sources such as renewable energy potential, locations of power plants, retail energy prices, and renewable energy incentives. Further possible improvements to the interface include analyzing the correlation between retail electricity prices and state's composition of energy sources to see how much "cheaper" (if any) traditional energy sources (coal, oil, gas) are compared to nuclear power and renewable sources. We believe the current data and visualization of our project is sufficient to spark interests and questions in the users of our software.