



FALL SEMESTER 2025-26

DSCI-510 : Principles of Programming for Data Science

1. Name of the project and team members

How Do Heating Fuel Prices Compare Across California, and What Is Their Relationship to Energy Efficiency?

Name	USC ID	E-Mail
Abhey Sabesan Mageswaran Aryaan	7162-7286-71	sabesanm@usc.edu
Rylan Nathan Lewis	8358-1308-73	rylannat@usc.edu

2. What problem are you trying to solve?

This project aims to understand how California's heating-fuel and electricity operational metrics have changed over time, and what these changes reveal about the state's energy efficiency and consumption patterns. By analyzing trends in generation, fuel consumption, costs, stocks, heat content, and related operational indicators, the project seeks to answer whether California's energy usage has become more efficient between 2020 and 2025, how different fuel types contribute to total electricity production, and what operational behaviors correlate with cost or consumption changes. The overarching problem is to transform complex, raw energy-sector data into clear, data-driven insights about California's energy profile.

3. How will you collect data and from where?

The data will be collected directly from the **U.S. Energy Information Administration (EIA)** using the official **EIA v2 API**. Instead of downloading static files, the project retrieves data programmatically through dynamic API calls. A Python script sends sequential paginated requests to gather all monthly operational electricity data for California from **January 2020 to 2025**, including fields such as generation, total consumption, fuel-specific BTU values, stocks, sulfur content, cost metrics, and more. The script handles pagination, rate limiting, and JSON parsing to ensure the complete dataset, over 20,000 rows, is collected in a structured format. The data is then stored locally as a CSV for further processing.

4. What analysis will you do and what visualizations will you create?

After the dataset is cleaned and organized, the project will focus on exploring broad trends in California's electricity data. The analysis will look for meaningful patterns in areas such as consumption, generation, costs, and fuel-type usage over time. A range of high-level visualizations, such as trend lines, comparisons across fuel types, and simple relationship plots, will be used to highlight key insights.