

Seattle Building Energy Analysis

Project Overview: Our project tackles climate pollution from Seattle's buildings, emphasizing the unequal impact on BIPOC communities through fossil fuel use. Utilizing the Racial and Social Equity Composite Index, we apply an equity lens to analyze energy usage and emissions across neighborhoods. Our analysis extends to developing forecasting models from the City of Seattle's Open Data, aiming to predict future emissions under current practices versus cleaner energy scenarios. This approach not only seeks to reduce the city's climate pollution but also ensures equitable benefits distribution, guiding policy and community actions towards a sustainable and inclusive urban environment.

Datasets Overview: We utilize various datasets to analyze energy usage and its implications on climate pollution in Seattle.

1. PSE Usage Data for OSE Climate Portal

Source: [Seattle City GIS Open Data](<https://data-seattlecitygis.opendata.arcgis.com/datasets/SeattleCityGIS::pse-usage-data-for-ose-climate-portal/about>)

Description: Aggregated gas consumption data by sector, quarter, and census tract.

Format: Shapefile

2. Racial and Social Equity Composite Index

Source: [Seattle City GIS Open Data](<https://data-seattlecitygis.opendata.arcgis.com/datasets/SeattleCityGIS::racial-and-social-equity-composite-index-current/about>)

Description: Data integrating race, ethnicity, socioeconomic status, and health disadvantages.

Format: Shapefile

3. Seattle's Building Energy Benchmarking (2015-2022)

Source:[City of Seattle's Official Open Data Portal](<https://www.seattle.gov/environment/climate-change/buildings-and-energy/energy-benchmarking/data-and-reports#individualbuildingdata>)

Description: Energy performance data for buildings over 20,000 square feet.

Format: CSV

4. Performance Ranges by Building Type (2015-2022)

Source: [City of Seattle's Official Open Data Portal](<https://www.seattle.gov/environment/climate-change/buildings-and-energy/energy-benchmarking/data-and-reports#summaryinformationbybuildingtype>)

Description: Summarized energy usage and characteristics, excluding known errors.

Format: CSV

Data Wrangling Techniques

Joining Data: Merging datasets by year (CSV data) and census tract (Shapefile data) to analyze emissions patterns.

Cleaning: Removing missing values, handling outliers, and standardizing data formats.

Transformation: Normalizing numerical data and encoding categorical variables for analysis.

Data Storage: Datasets are stored in CSV and Shapefile formats, facilitating access and manipulation for our analysis.

Documentation: Data Flow Diagram and Data Dictionary are in the `Documentation` directory, providing a visual overview of the data ingestion and analysis process, and offering essential information about each dataset's structure and content.