

## **Progress Report**

# **Exploratory Analysis of EnergyPlus Simulation Data with Database Management and Web-Based Visualization**

Athul Jose P

Kunal Shankar

CptS 575 Data Science

Washington State University

## Number of Meetings Held Until Now

4

## Objective

This project analyzes data generated from numerous EnergyPlus simulations using Pacific Northwest National Laboratory (PNNL) Prototype Building Models and weather files. The goal is to perform an exploratory analysis to evaluate the richness of this data, uncover potential insights, and determine its usefulness for building performance modeling. We will also identify areas where the dataset could be improved for greater depth and diversity. Additionally, the project includes the development of a NoSQL database for efficient data storage and an upgrade to an existing web app for enhanced data analysis and visualization.

## Data Generation

The simulation data used in this project was generated using EnergyPlus, a state-of-the-art simulation tool widely used for modeling energy consumption in buildings. We chose two locations—Seattle, WA, and Florida—to account for completely different weather conditions and better understand how climatic variations affect building performance.

The dataset consists of simulations for two categories of buildings: large and small. For the purpose of our analysis, we decided to focus on the smaller building types, as they offer more manageable datasets and provide key insights into energy usage patterns without the complexity introduced by larger facilities. The buildings modeled include residential, manufactured, commercial, and small office buildings.

The data is stored in pickle files, organized by building type and location. Our planned analysis will include:

- **Time-Series Analysis:** To explore temporal patterns in energy consumption and operational schedules.
- **Visualization:** To identify trends and compare performance across different climate zones and building types.

We aim to extract meaningful insights from the data, such as seasonal energy usage patterns and the impact of operational schedules on overall consumption. These insights will be essential for improving building energy models and informing data-driven decisions in building performance optimization.

Additionally, we are learning from YouTube tutorials to implement a NoSQL database for storage when the time comes. The NoSQL database will allow efficient storage and retrieval of large simulation datasets, supporting future analysis and visualization needs.

## **Main Idea**

For now, we will be performing an exploratory analysis to understand patterns in the dataset, including energy consumption, climate zone behavior, and operational variations, and identify initial use cases for the data. After this step, we plan to enhance the dataset by adding more variables or improving data coverage based on insights gained from the exploratory analysis.

The main goal is to extend the functionality of the web app, as mentioned in the proposal, by adding improved data analysis and visualization features. Additionally, we plan to integrate a NoSQL database to facilitate efficient storage and retrieval of simulation data.