# Energy Efficiency Prediction

Predicting EUI and Energy Efficiency Classification

Anurag Tadapaneni

# Objective

The goal of this project is twofold. The primary objective is to predict the EUI (Energy Use Intensity) of buildings based on their basic characteristics. The secondary objective is to classify these buildings as efficient or not.

# Dataset Overview

The dataset contains 21,132 buildings in NYC, with 53 columns per building containing features like size, year built, fuel type, etc . . .

Key Features
- GFA (Gross Floor Area)
- Building Type
- Number of floors
- Year Built
- Energy Distribution

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | gfa | eui | value | eui_diff | ghg_diff | electricity | natural_ga |
| 2 | 1.353092 | 0.112906 | 2.851183 | -0.11228 | -0.12316 | 1.957964 | 0.083494 |
| 3 | -0.56772 | 1.851008 | -0.01665 | -0.01936 | -0.03948 | 0.214964 | -0.54982 |
| 4 | -0.75113 | 0.078951 | -0.26586 | -0.06569 | -0.09072 | -0.5272 | -0.46222 |
| 5 | -0.6459 | -0.47071 | -0.44311 | -0.11482 | -0.13919 | -0.37364 | -0.66626 |
| 6 | -0.62971 | -1.42783 | -0.33141 | -0.08693 | -0.11945 | -0.57765 | -0.66077 |
| 7 | -0.62666 | -0.54286 | -0.57895 | -0.07291 | -0.09121 | -0.5052 | -0.46216 |
| 8 | 0.198841 | -0.17996 | 0.137108 | -0.01567 | 0.16178 | -0.01527 | 0 |
| 9 | -0.54972 | -1.1095 | -0.24637 | -0.08079 | -0.08171 | -0.42714 | -0.68869 |
| 10 | -0.71445 | -0.57469 | -0.42333 | -0.13617 | -0.16335 | -0.49338 | -0.616 |
| 11 | -0.76791 | -0.85271 | -0.50967 | -0.22177 | -0.24126 | -0.60735 | -0.59739 |
| 12 | -0.76014 | -0.77206 | -0.33249 | 0.148809 | 0.400634 | -0.60135 | -0.80263 |
| 13 | -0.57456 | 0.025895 | -0.5087 | -0.03369 | -0.05008 | -0.59098 | -0.77785 |
| 14 | -0.7648 | 0.095928 | -0.33807 | -0.15345 | -0.18114 | -0.36555 | -0.72985 |

# Preprocessing

A lot of the dataset had missing data across multiple columns

- Imputed energy distribution with the building type median
- Dropped buildings with missing data like GFA or year built, as well as outliers

Predicted vs Actual EUI
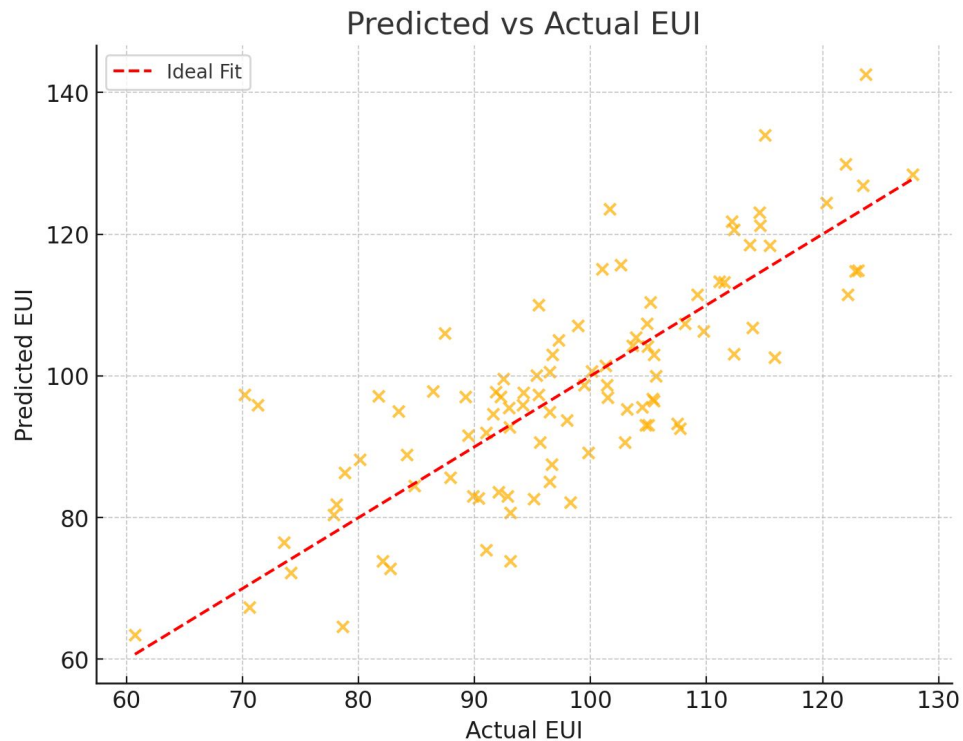
MSE: 103.76

MAE: 9.44

Acc: 68.42%

# Regression Results

# Classifying Results

- A building was deemed energy efficient if its predicted EUI was below the median EUI in the US
- While there are no hard statistic to reference this with, the data does correlate to NYC's general energy efficiency report distribution, with almost 50% of the city getting a D or F



Efficient vs Inefficient Buildings
Efficient 12.02%
Inefficient 87.98%
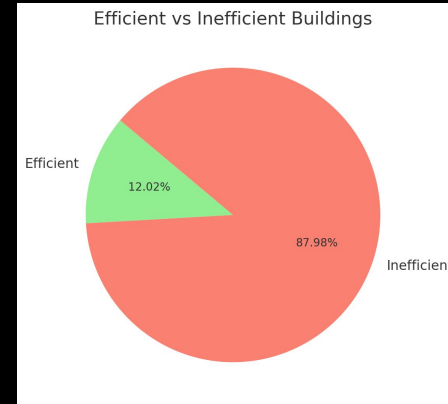
# What's Next?

The goal is to expand on this project to not only predict the buildings current EUI, but future EUI in the event of a buildings retrofit. The goal is to identify buildings in NYC that will benefit the most from a potential retrofit in order to improve its energy efficiency.