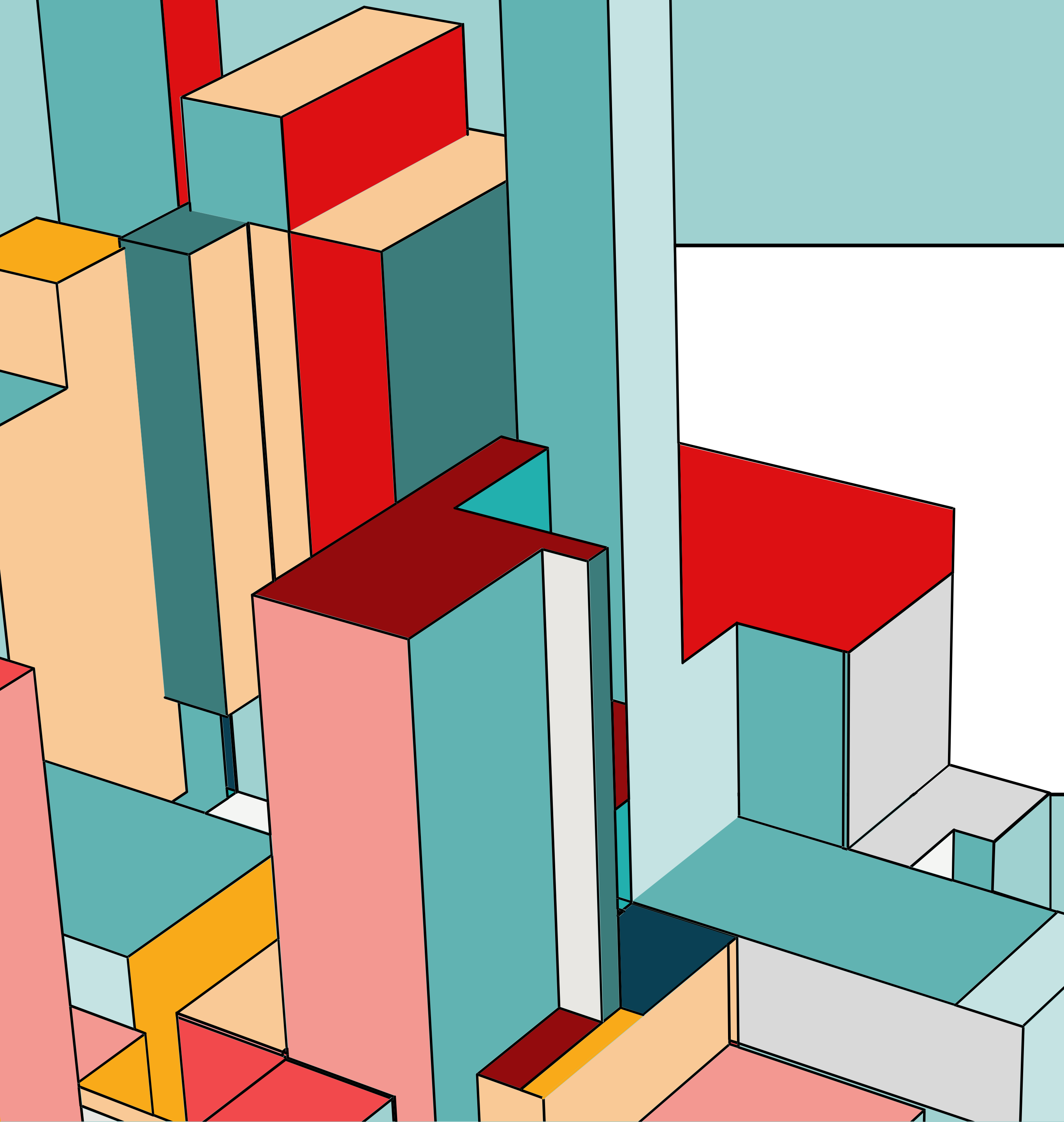


BUILDING ENERGY CONSUMPTION PREDICTOR (BECP)

Jose Correa

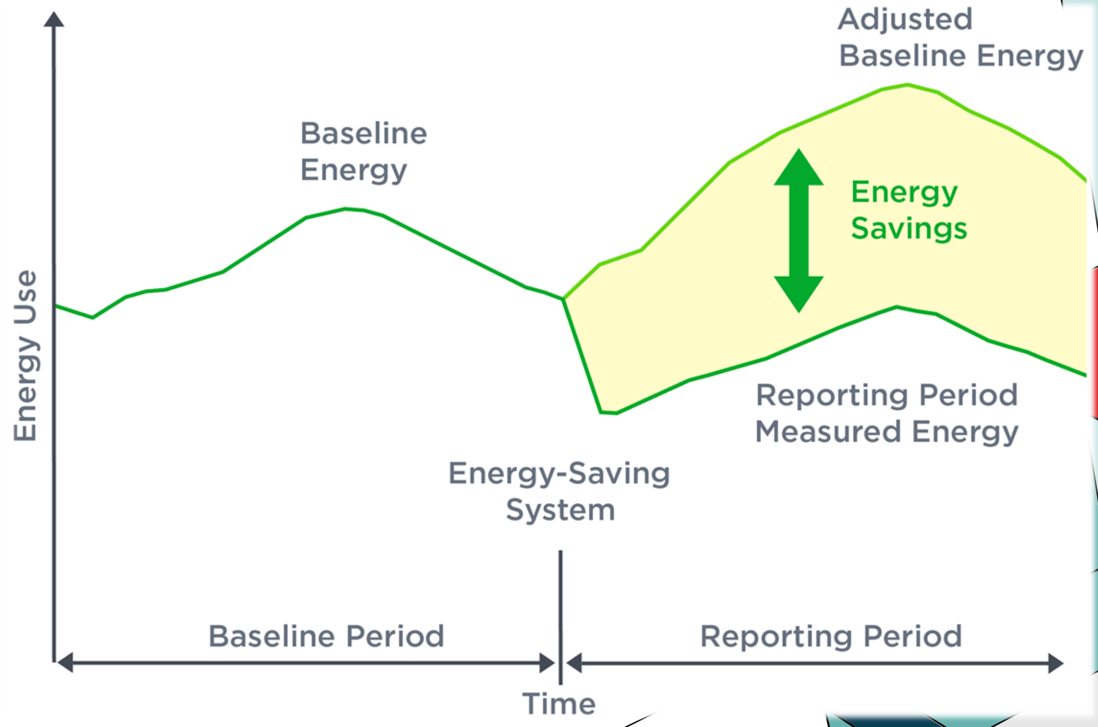
PROBLEM

PROBLEM STATEMENT / OPPORTUNITY YOU IDENTIFIED

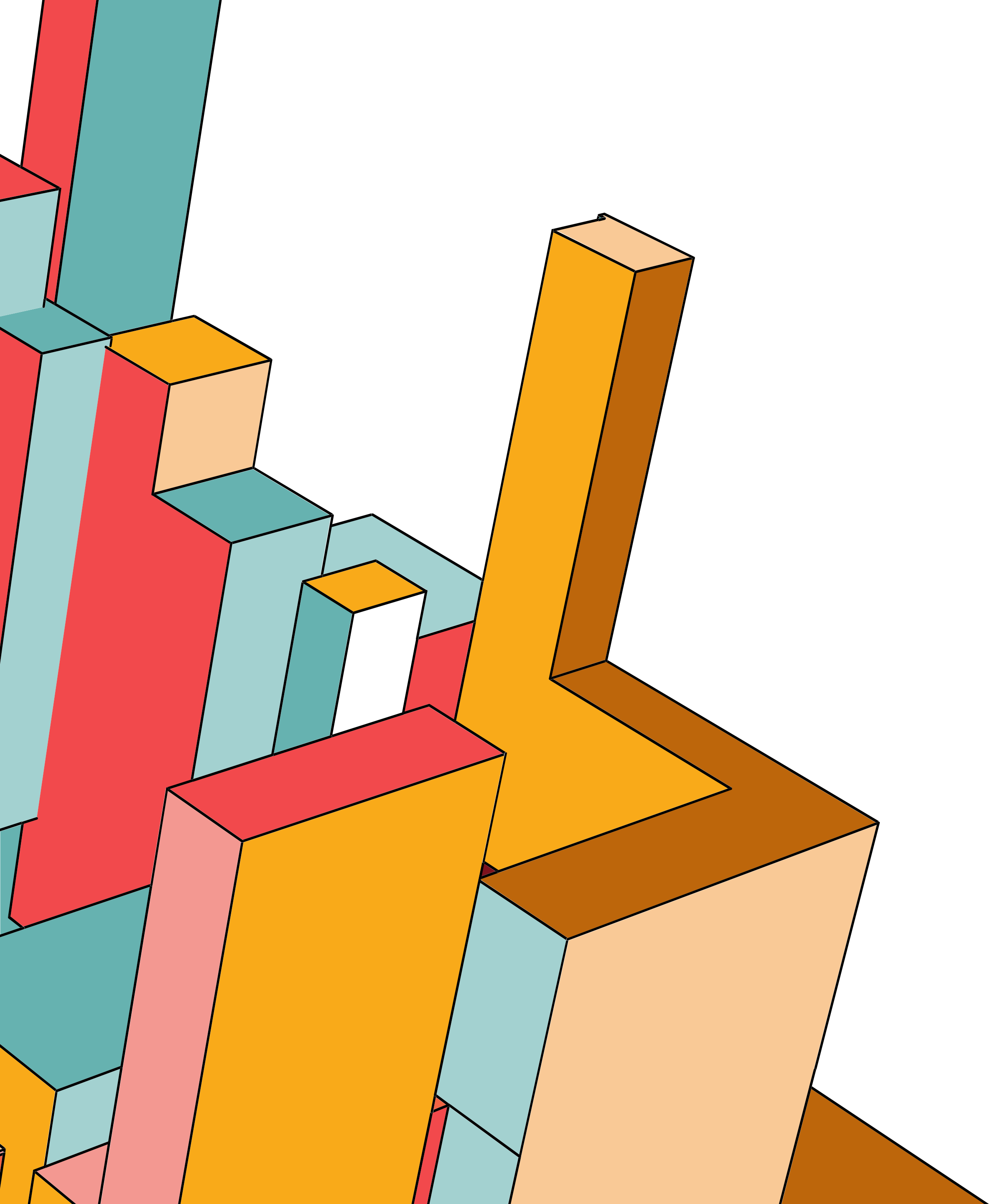
Inaccuracy evaluating the effectiveness of building retrofits and their impact on energy consumption.

SUBJECT AREA

Built environment
➡ Building energy efficiency retrofits.
➡ Incentives\Services related to 'pay-for-performance financing.'



Source: *https://resourcekraft.com/measurement-and-verification-epcs-ipmvp/*



SOLUTION

Implement the BECP Machine Learning model to evaluate the effectiveness of building retrofits by:

ACCURATE PREDICTIONS

Predicts building energy consumption for categories, such as: chilled water, electric, hot water, and steam meters.

PERFORMANCE ASSESSMENT

Gives a clear understanding of whether the retrofits are delivering the anticipated reduction in energy usage.

INVESTMENT DECISION SUPPORT

Offers stakeholders a reliable tool to make informed decisions about retrofit investments.

SOLUTION IMPACT

REAL STATE MARKETABILITY

- ✓ 9 out of 10 homebuyers prefer a home with energy-efficient features that lower energy costs[2]
- ✓ Home value increases by $20 for every $1 saved on energy bills[3]

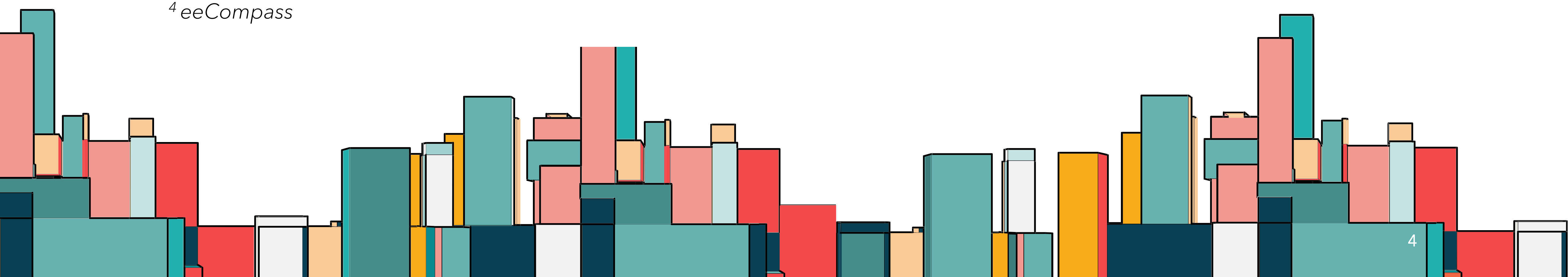
CLIMATE CHANGE

Buildings account for 30% of global energy consumption and 26% of global energy-related emissions[1]

ENERGY EFFICIENCY

Higher-efficiency equipment can reduce energy use by 50% for electric heating and cooling systems[4]

[1] International Energy Agency
[2] EnergyLogic
[3] National Renewable Energy Laboratory (NREL)
[4] eeCompass

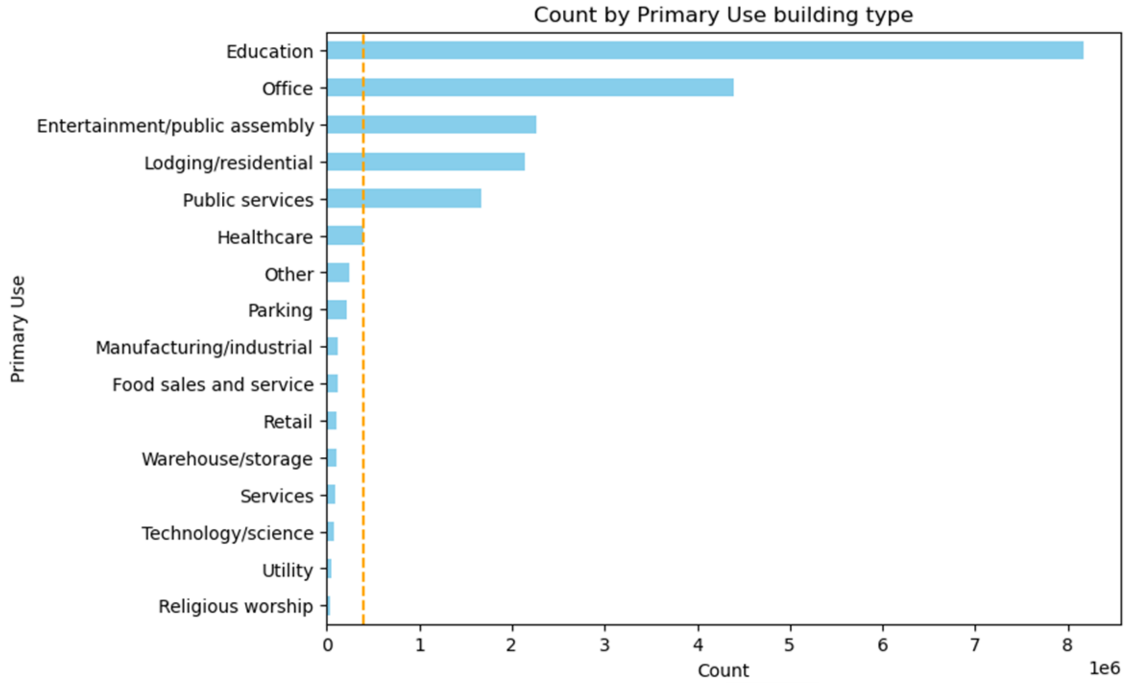


398527

BECP DATASET

MAIN FEATURES

- Data collected from more than 1,000 buildings over three years.
- Information stored in three CSV files
- Most columns contain numeric data, with only three categorical.
- The raw dataset comprises of 20 million rows, filtered down to a subset of 398,527 rows and 17 columns.

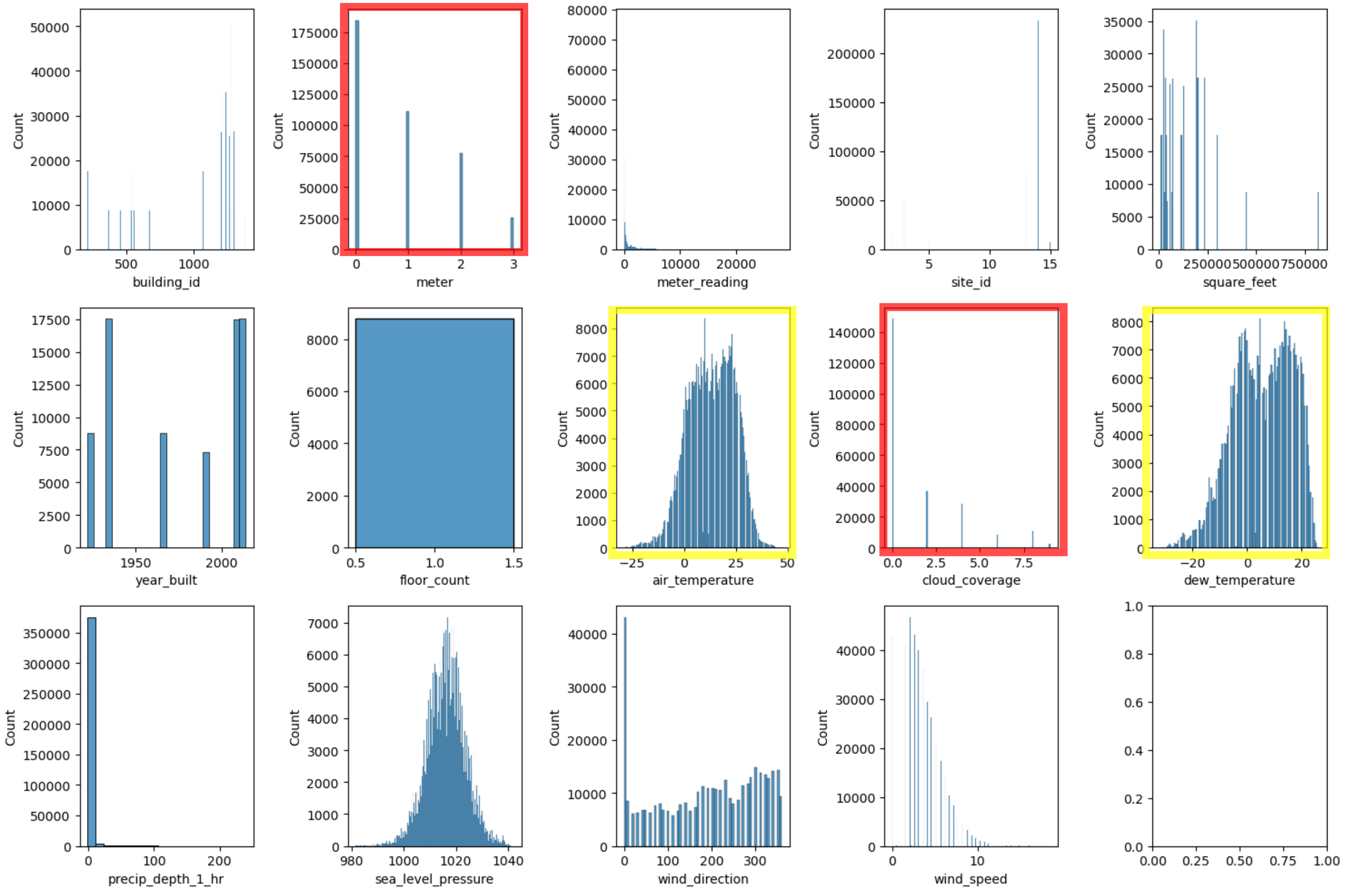


DATA QUALITY CONCERNS

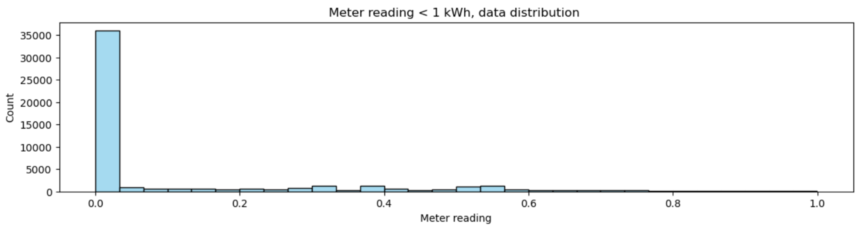
- 9% of energy data equal to zero.
- 80% missing values for ‘year_built’
- Energy measurements, heavily positive skewed .

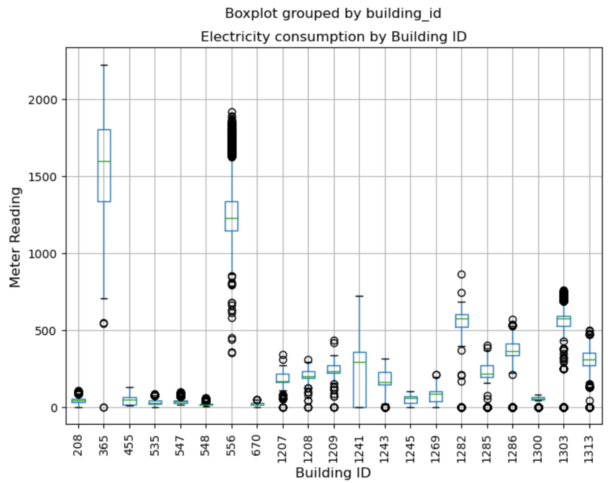
SUBSET EDA

Numeric data distribution



Energy meter reading distribution





Variables features:

- Bimodal distribution for air and dew temperature
- Categorical structure

NEXT STEPS

DATA PROCESSING

- Apply log transformation to energy consumption, heavily positive skew.
- Perform time-series analysis to uncover patterns, seasonality, or trends.

FEATURE ENGINEERGING

- Create derived variables: EUI (Wh/m2).
- Extract temporal features such as time of day, day of the week, seasonality, or holidays, which can impact energy usage.
- Group temperature ranges, humidity levels and other weather variables.
- Validate newly engineered features by assessing their correlation with the target variable.

BASELINE MODELING

- Handling missing values
- Train the time-series model
- Evaluate the model's performance
- Understand model's predictive capability and limitations

THANK YOU

Jose Correa