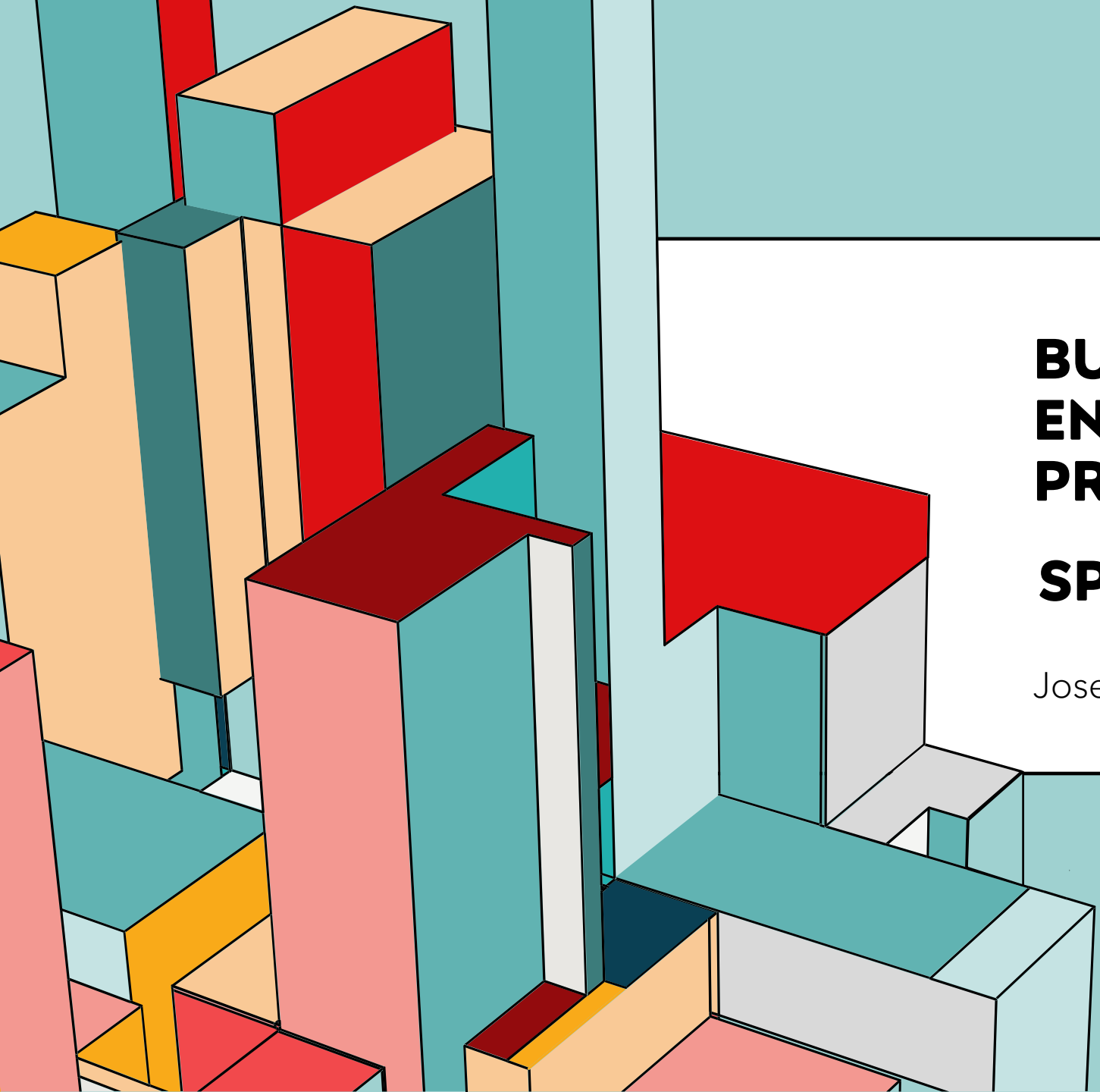


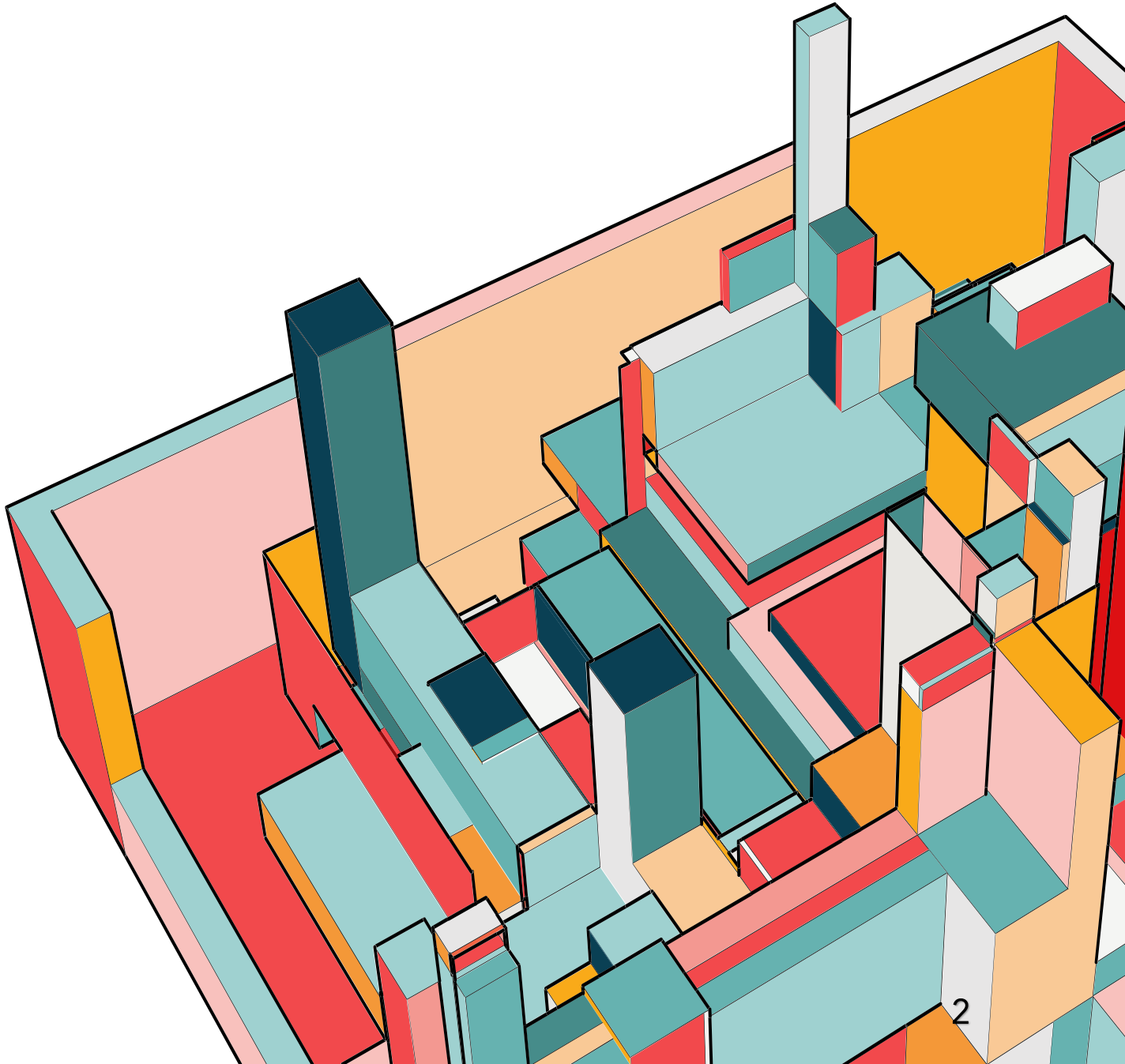
BUILDING ENERGY CONSUMPTION PREDICTOR (BECP)

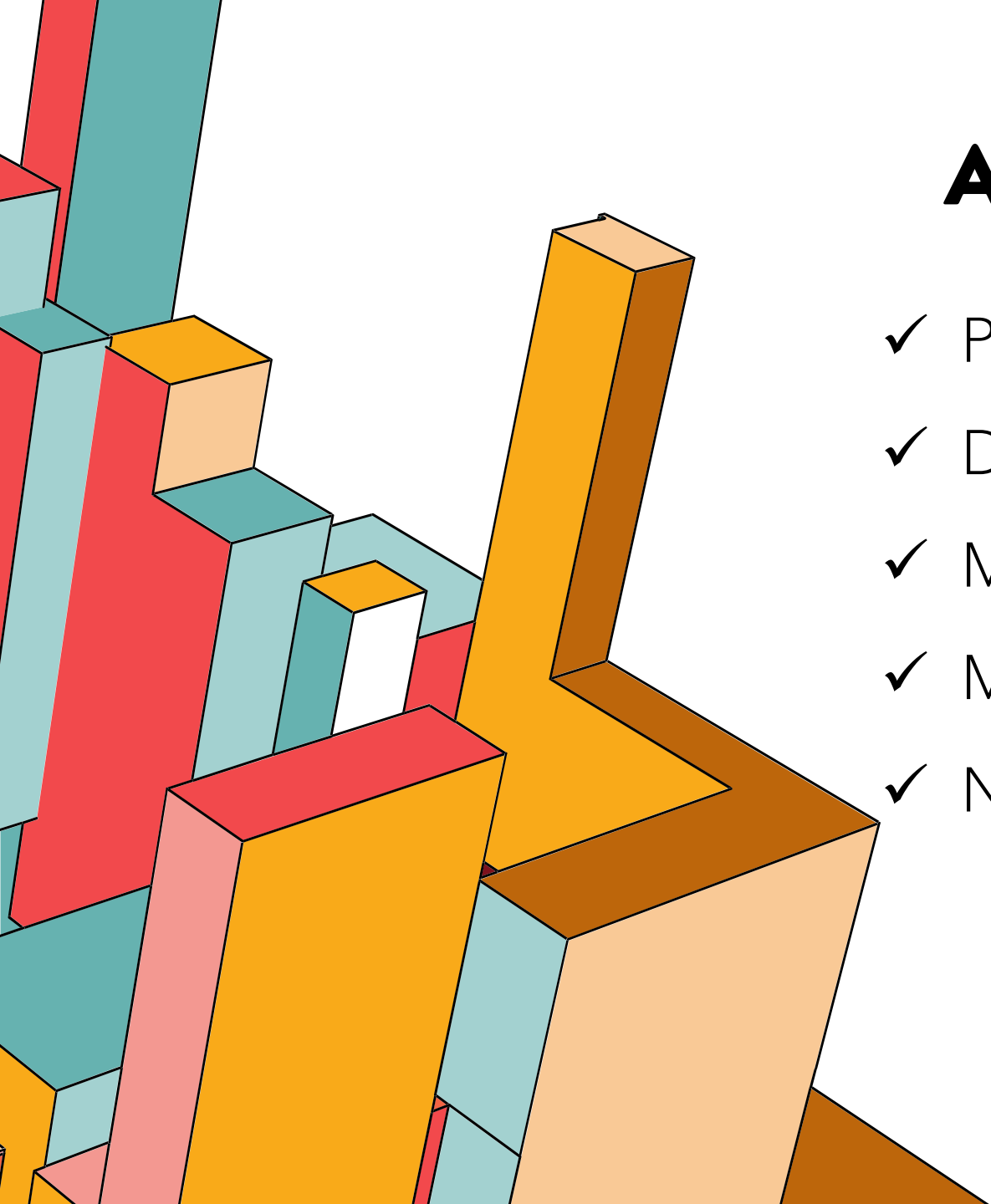
SPRINT 3

Jose Correa

OBJECTIVE

Present an advanced model for optimization, evaluation and interpretation.



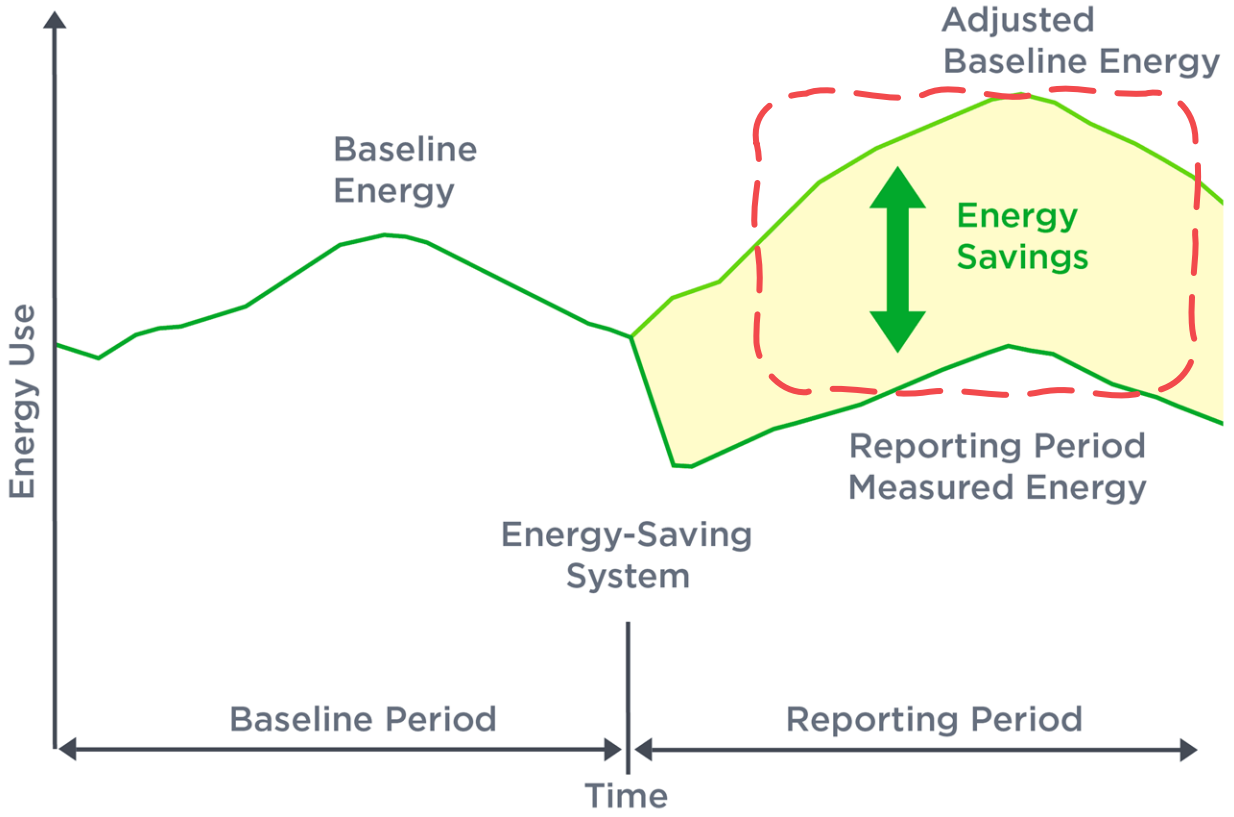


AGENDA

- ✓ Project overview
- ✓ Dataset description and Preprocessing
- ✓ Model comparison
- ✓ Model interpretation
- ✓ Next steps

PROBLEM STATEMENT

Inaccuracy when assessing how effectively building renovations impact energy consumption.



HOW BIG THE PROBLEM IS ?

The **health care** sector is one of the largest segments of the U.S. economy (**17% of U.S.GDP**)

The average hospital uses 2.5 times the amount of energy as other commercial buildings, adding up to 836 trillion Btu or **$5 billion annually**

Source: *ASHRAE and Department of Energy (DOE)*

PROPOSED SOLUTION

Promote **efficient building renovations** through the implementing of the BECP ML model.

This model **predicts energy consumption** by considering building features and weather conditions.

POTENTIAL IMPACT

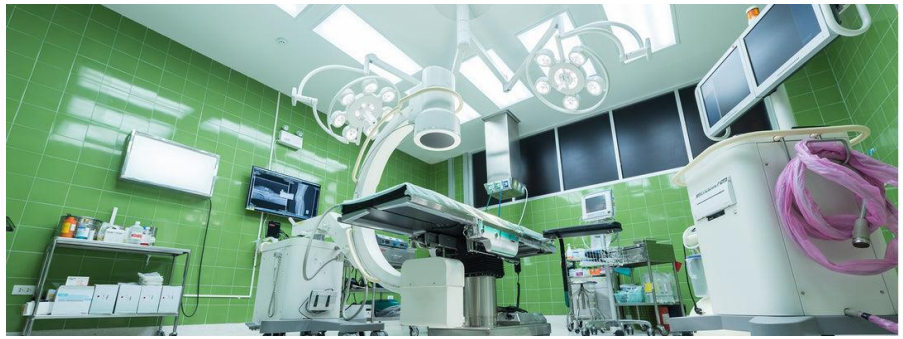
The average hospital energy cost annually

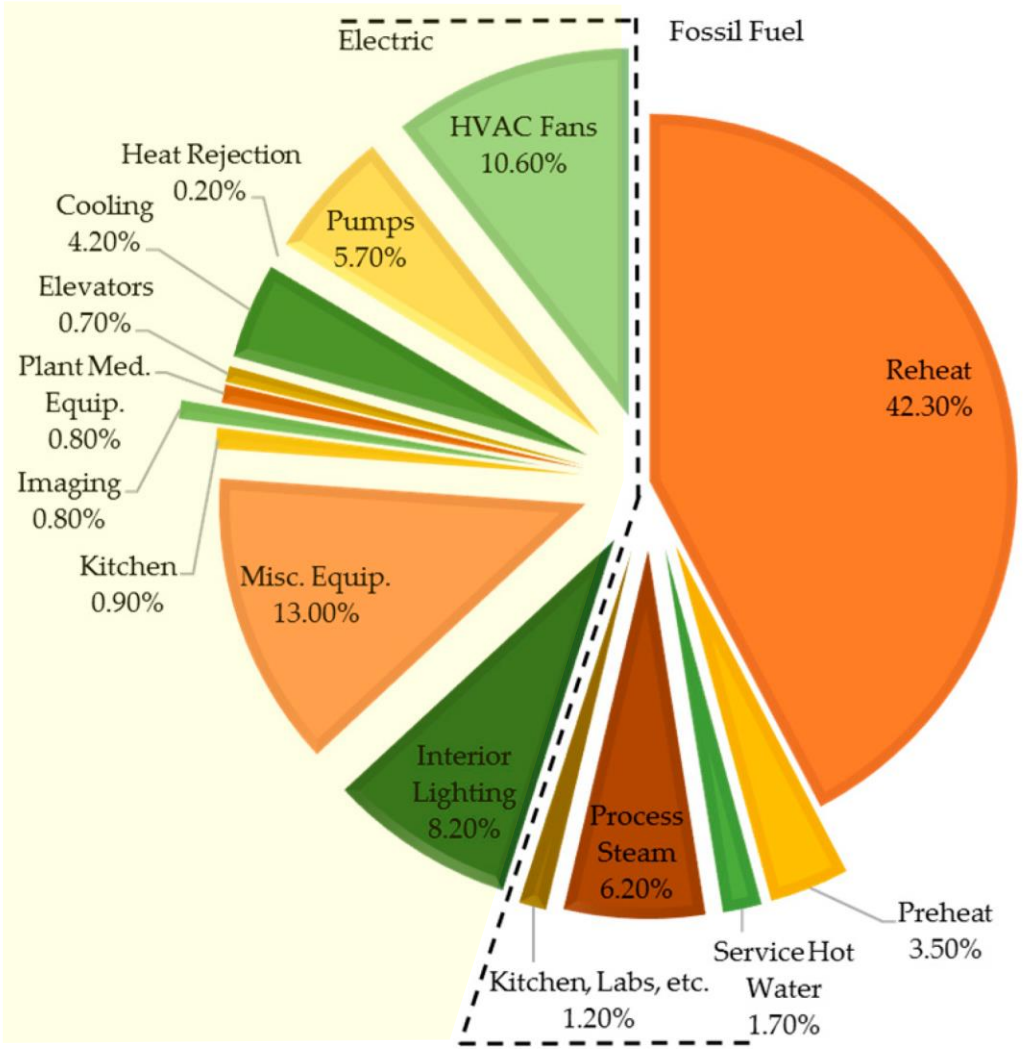
ENERGY EFFICIENCY

Higher-efficiency equipment can reduce energy use by 50% **($2.5 billion annual savings)**

***Source**: U.S Department of Energy*

ENERGY USAGE IN A HEALTH CARE BUILDING



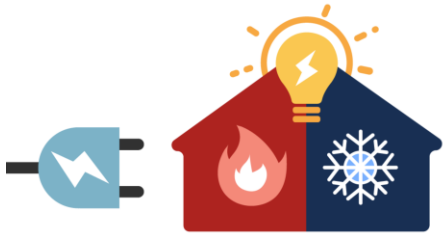


Source: *Energy Consumption Analysis and Characterization of Healthcare Facilities in the United States*

DATASET

https://www.kaggle.com/competitions/ashrae-energy-prediction/

3 x 



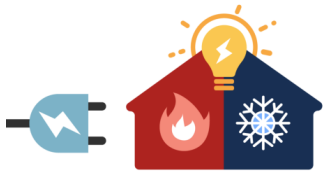
20M rows

400k rows

DATASET DETAILS

Target variable

Meter reading



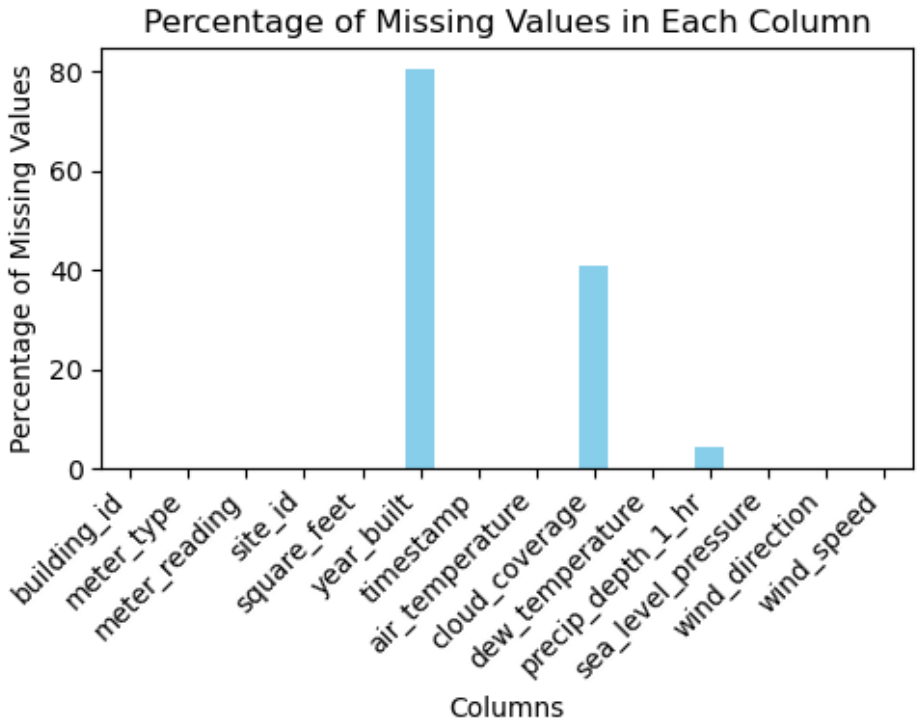
Features

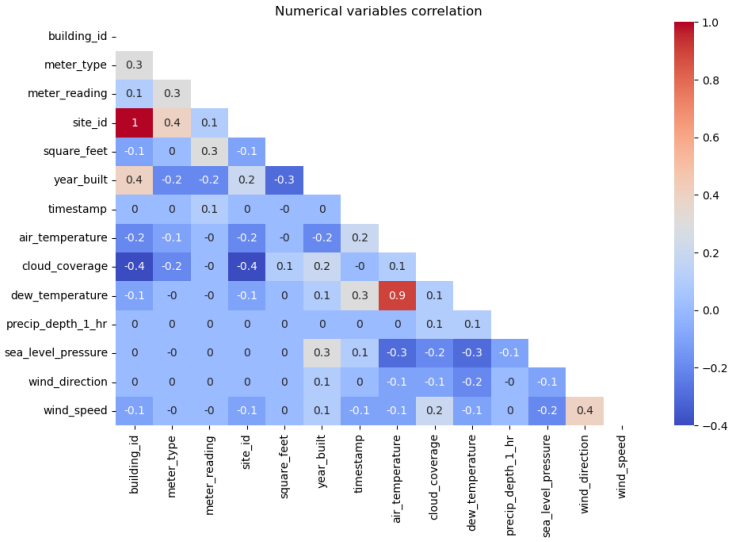
meter
meter timestamp
primary use
square feet
year built
floor count

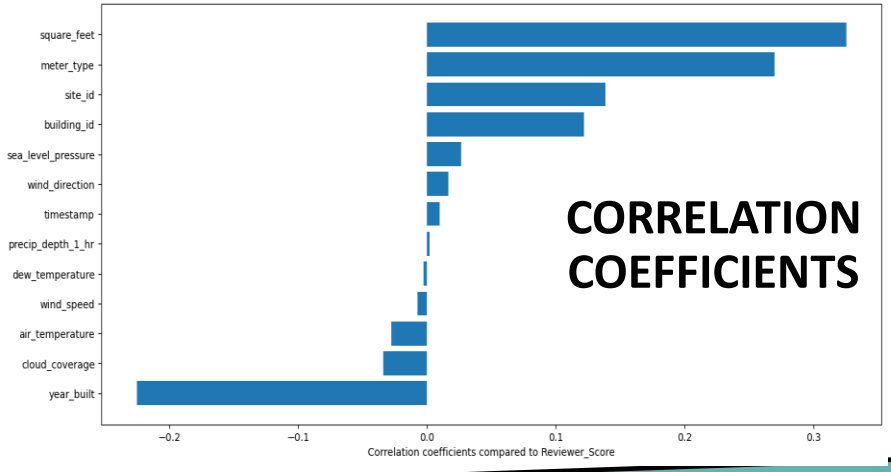
weather timestamp
air temperature
cloud coverage
dew temperature
precip depth 1 hr
sea level pressure
wind direction
wind speed

PREPROCESSING

NULL VALUES



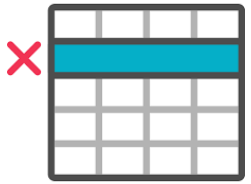




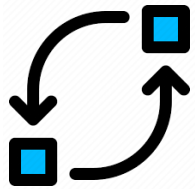
CORRELATION COEFFICIENTS

PREPROCESSING

year_built,
cloud_coverage,
precipitation,
primary_use,
floor_count,
meter_timestamp,



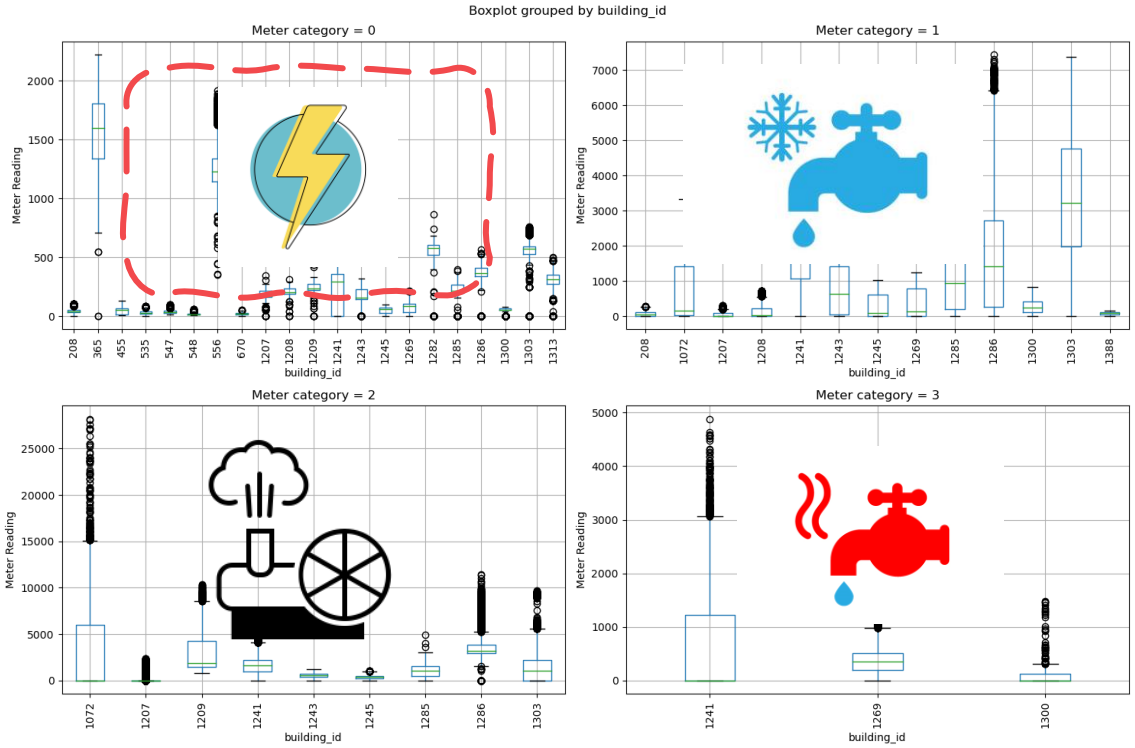
timestamp
0.13% (515
rows)



sea_level_pressure
wind_direction
wind_speed
air_temperature
dew_temperature

FINDINGS FROM EDA

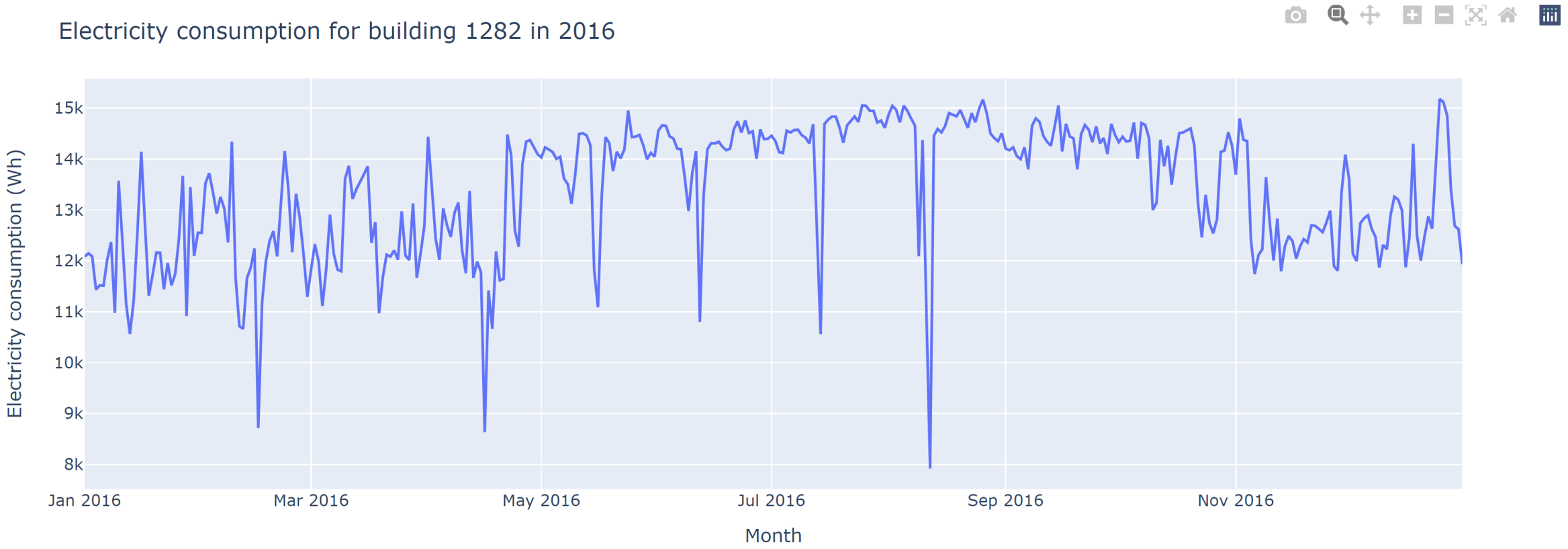
BUILDING AND METER TYPES



BUILDING ELECTRICITY

FINDINGS FROM EDA

Electricity for Building 1282



MODELLING

Baseline

Mean

AR

Auto-Regressive

ARIMA

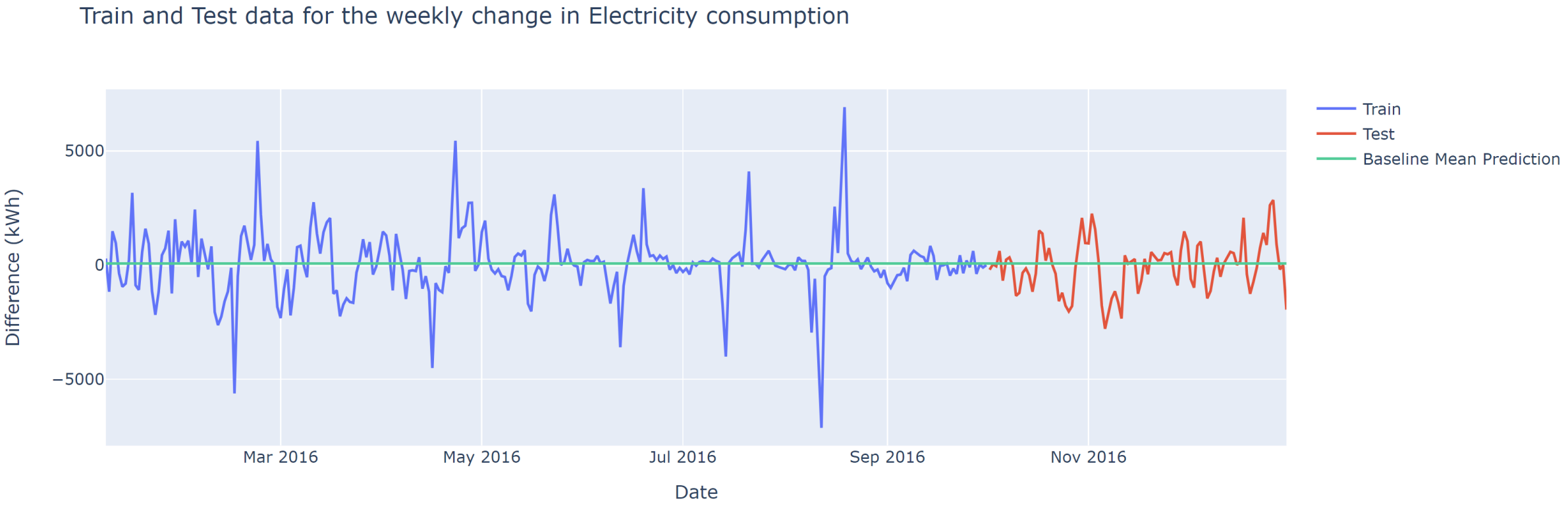
Auto-Regressive
Integrated
Moving Average

XGboost Regressor

Random Forest Regressor

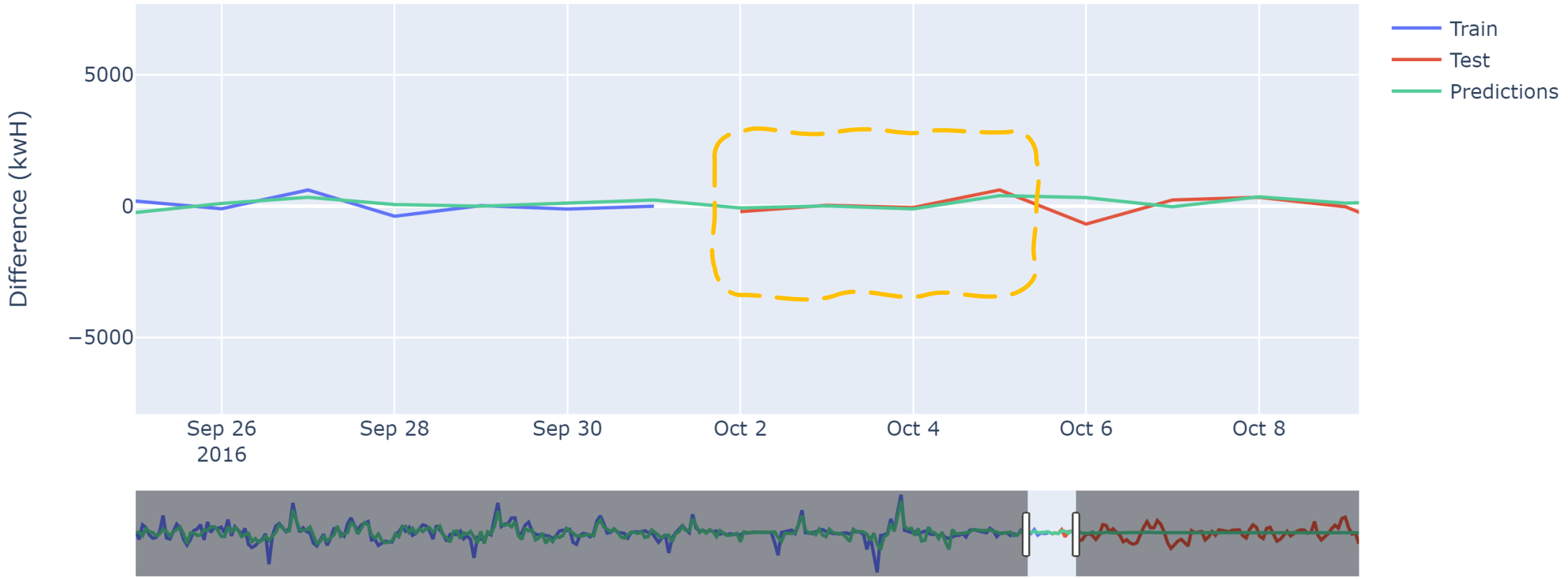
Multivariate time series analysis

BASELINE: FORECASTING



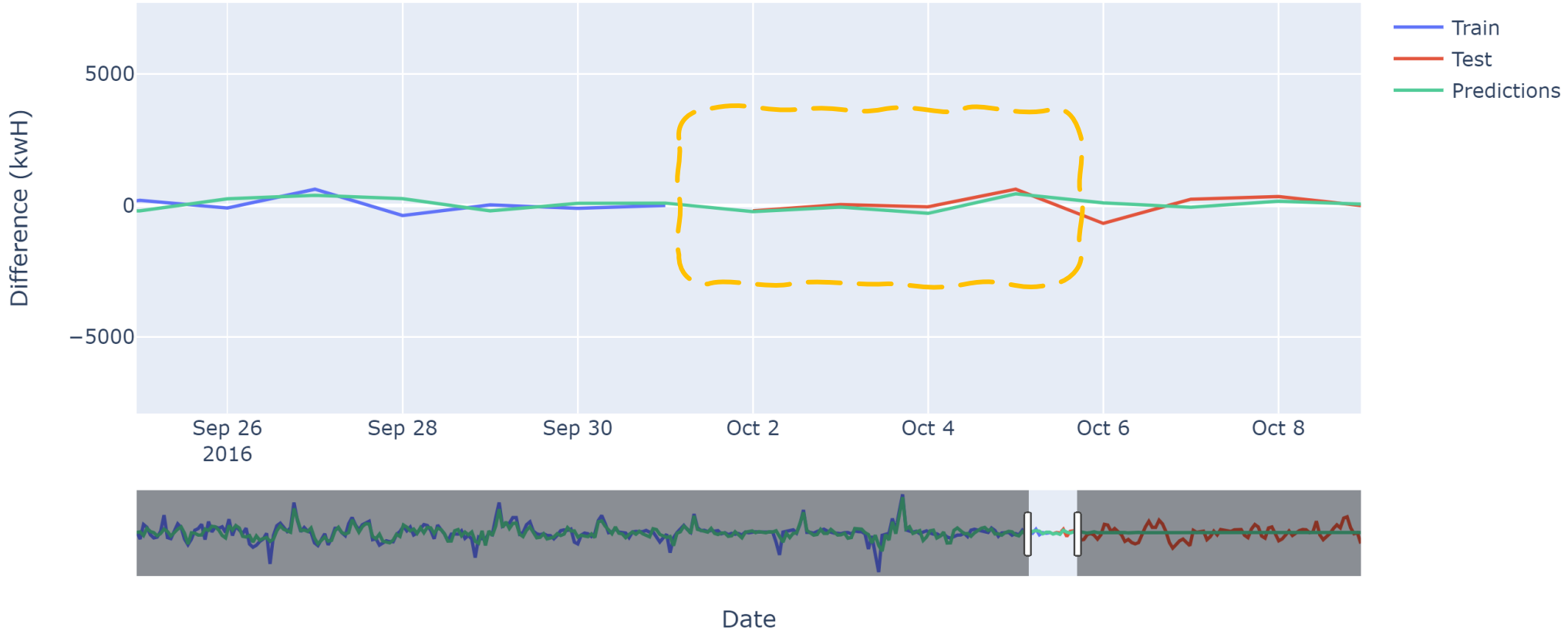
Train MAE = 908.7
Test MAE = 867.5

AUTO-REGRESSIVE (AR) MODEL



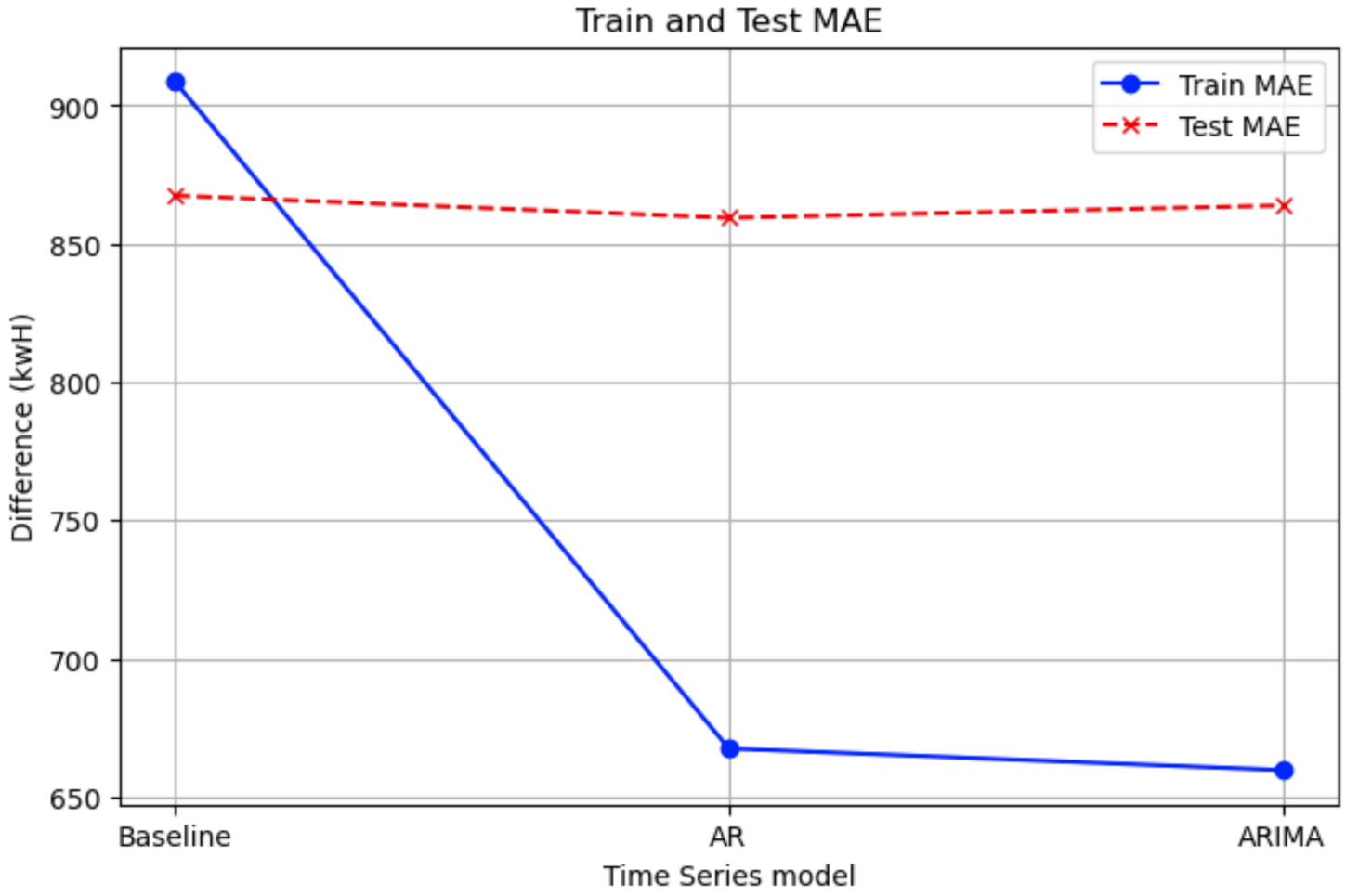
Train MAE = 667.6 (27% lower than baseline)
Test MAE = 859.4 (1% lower than baseline)

ARIMA: AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

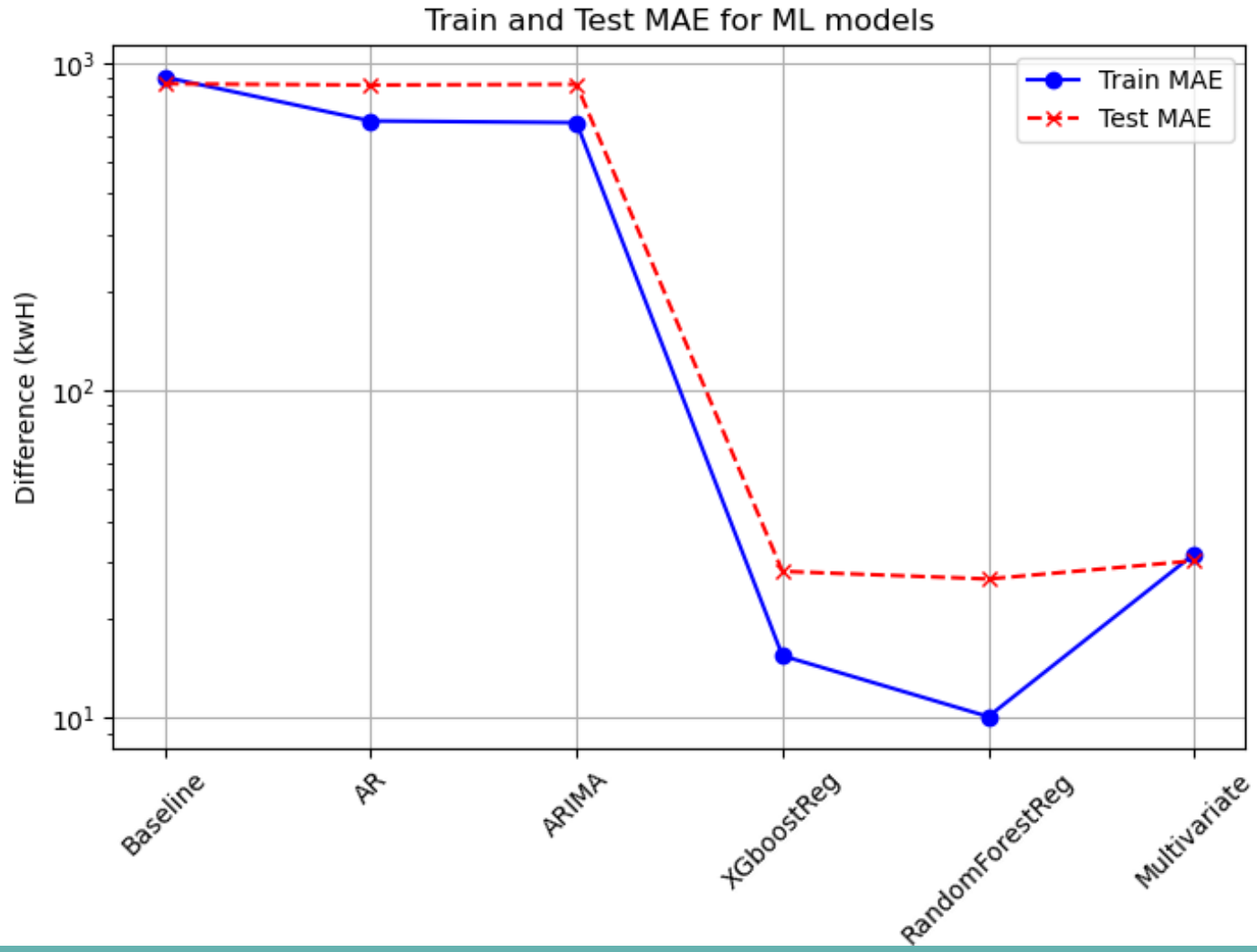


Train MAE = 659.8 (1% lower than AR)
Test MAE = 863.9 (1% higher than AR)

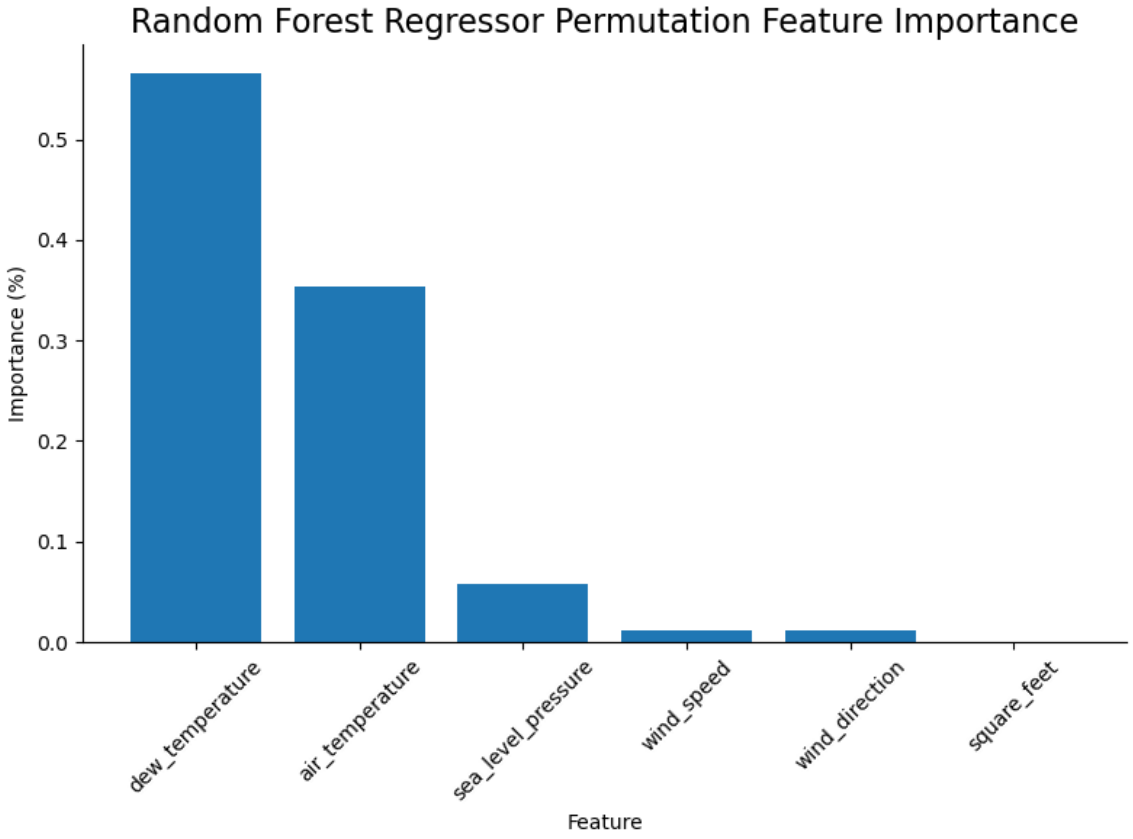
TIME SERIES MODEL COMPARISON

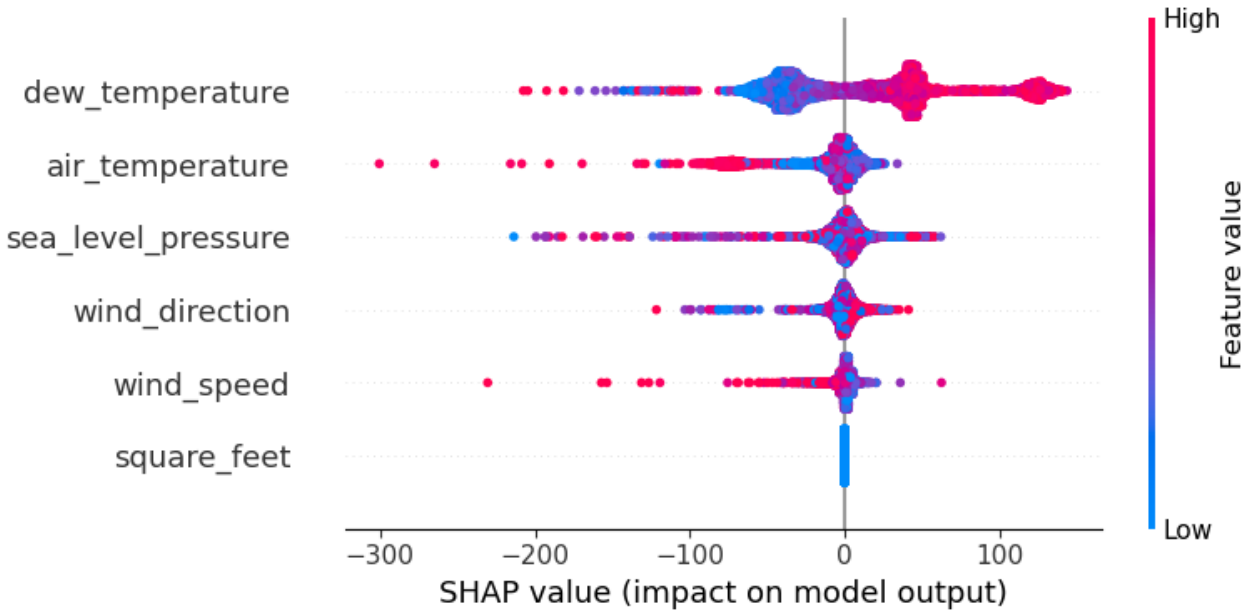


TIME SERIES + ENSEMBLE METHODS COMPARISON

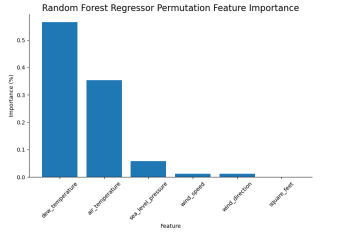


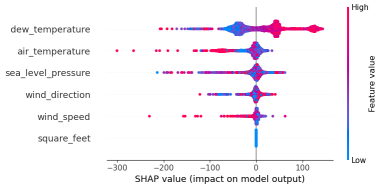
MODEL INTERPRETATION

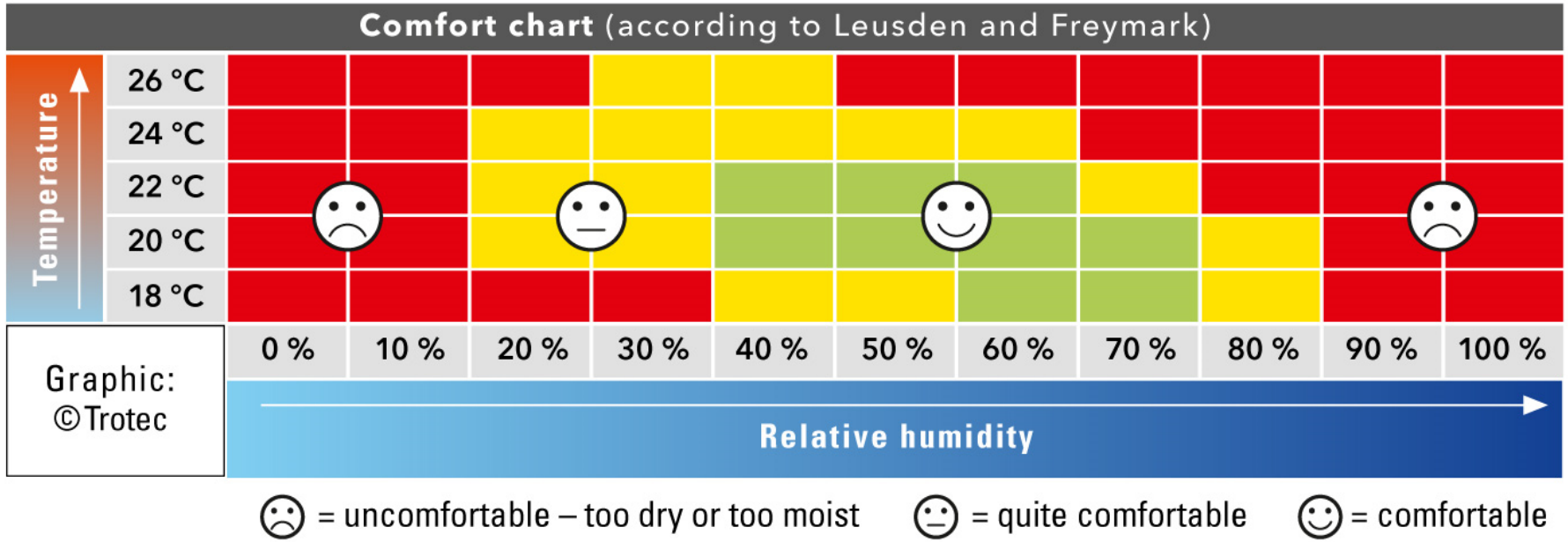




MODEL INTERPRETATION







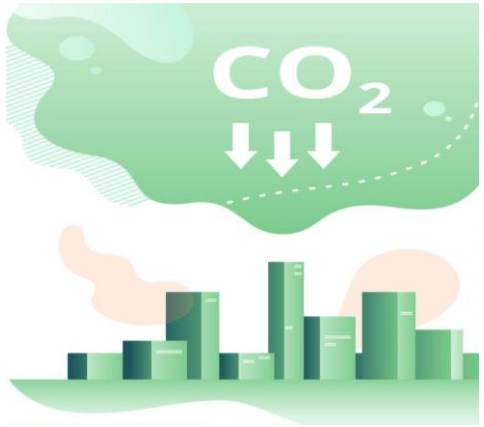
REAL WORLD USE OF THE MODEL

LIFE CHANGER FOR:

Patients are more comfortable, potentially aiding in faster recovery.

Healthcare staff improve productivity and well-being.

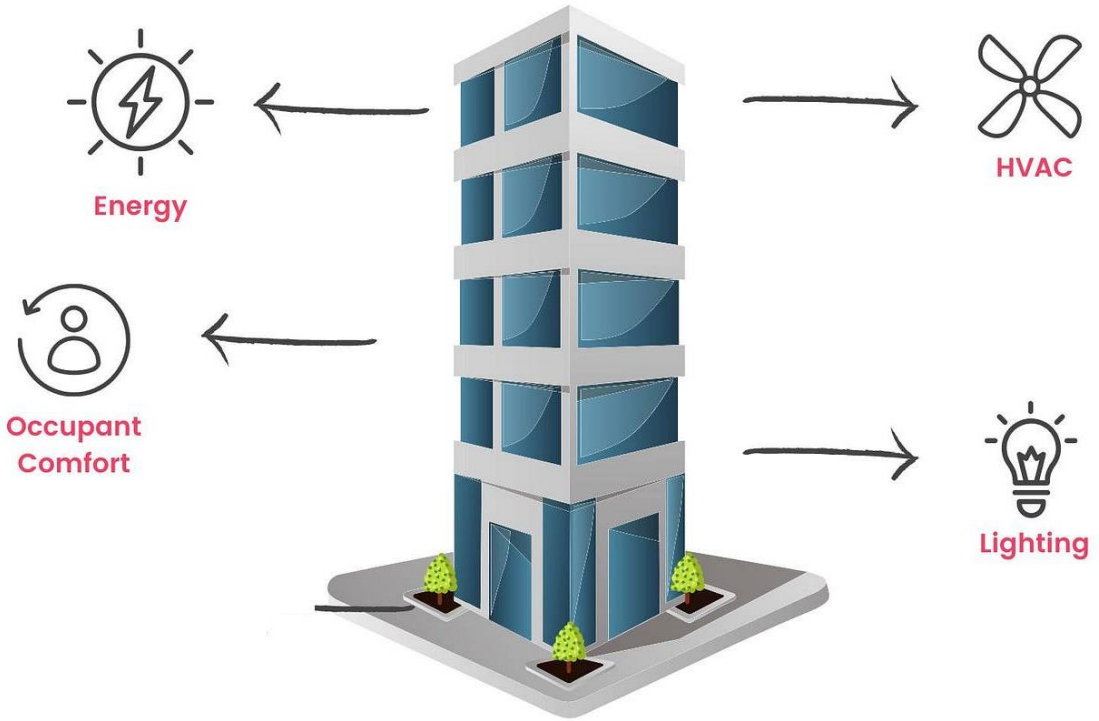
Environment with lower carbon emissions.



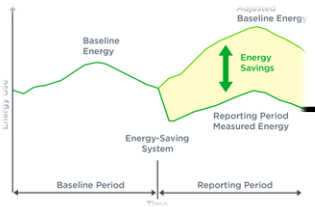
REAL WORLD USE OF THE MODEL

WHAT PRODUCTS OR SERVICES IS IT GOING TO CHANGE:

- Lighting systems.
- HVAC.
- Energy management systems.
- Renewable energy products.
- Improved maintenance.

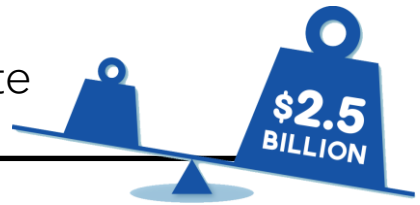


PROBLEM TIMELINE



Baseline	AR	ARIMA	Xgboost Reg	Random Forest Reg	Multivariate
867	859	864	28	27	30

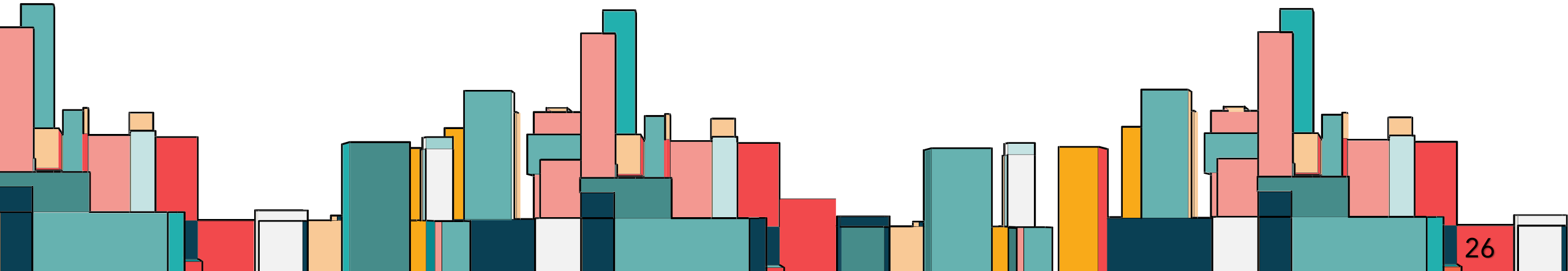
Test MAE



NEXT STEPS

RNN

Neural network



THANK YOU

Jose Correa