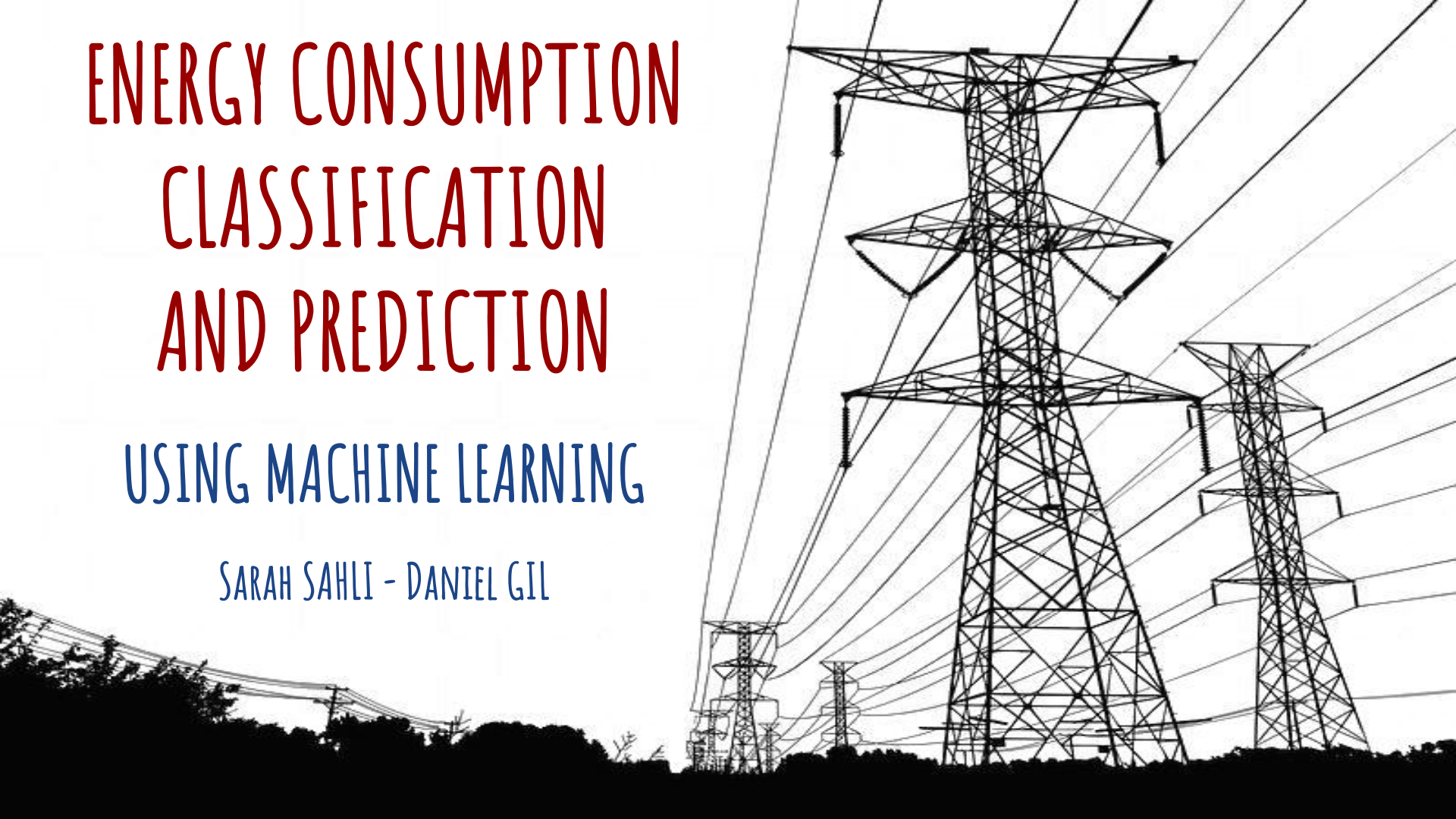
ENERGY CONSUMPTION CLASSIFICATION AND PREDICTION

USING MACHINE LEARNING

SARAH SAHLI - DANIEL GIL

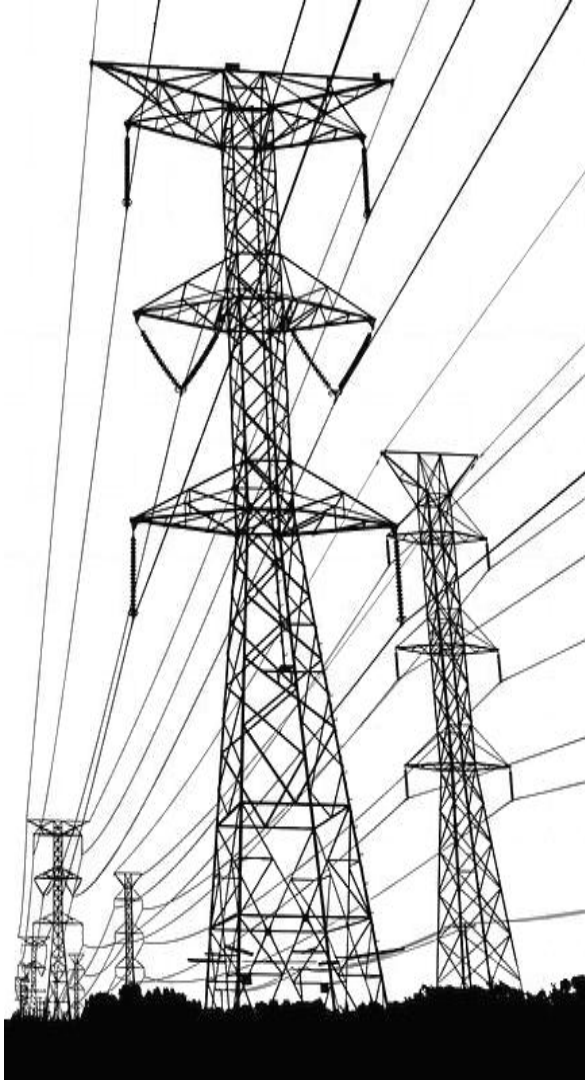


CONTEXT

NEW ENERGY COMMERCIALIZING COMPANY IN COLOMBIA WANTS TO DEFINE MORE DYNAMICS PRICING STRATEGIES

USING MACHINE LEARNING TECHNIQUES, THEY WANT TO BE ABLE TO:

- PROFILE THEIR CLIENTS ACCORDING TO THEIR ENERGY CONSUMPTION HABITS
- PREDICT THEIR ENERGY CONSUMPTION

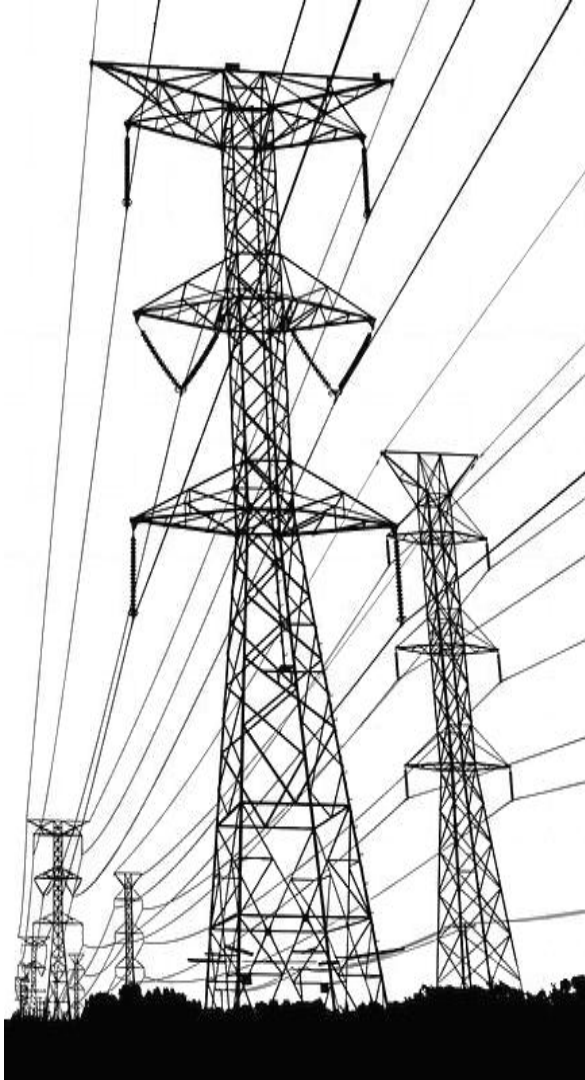


CASE STUDY DESCRIPTION

DATASET

- HOURLY INFORMATION OF ENERGY CONSUMPTION (WH)
- DURING A PERIOD OF 6 MONTHS
- FOR 233 HOUSEHOLDS

* NO INFORMATION DIRECTLY RELATED TO THE CLIENTS



CASE STUDY OBJECTIVES

1. CLASSIFY CLIENTS BASED ON THEIR ENERGY CONSUMPTION HABITS

1. ESTABLISH THE BASELINE OF ENERGY CONSUMPTION FOR EACH CLIENT AND PREDICT THEIR DAILY ENERGY CONSUMPTION

PRE-PROCESSING OF THE DATASET

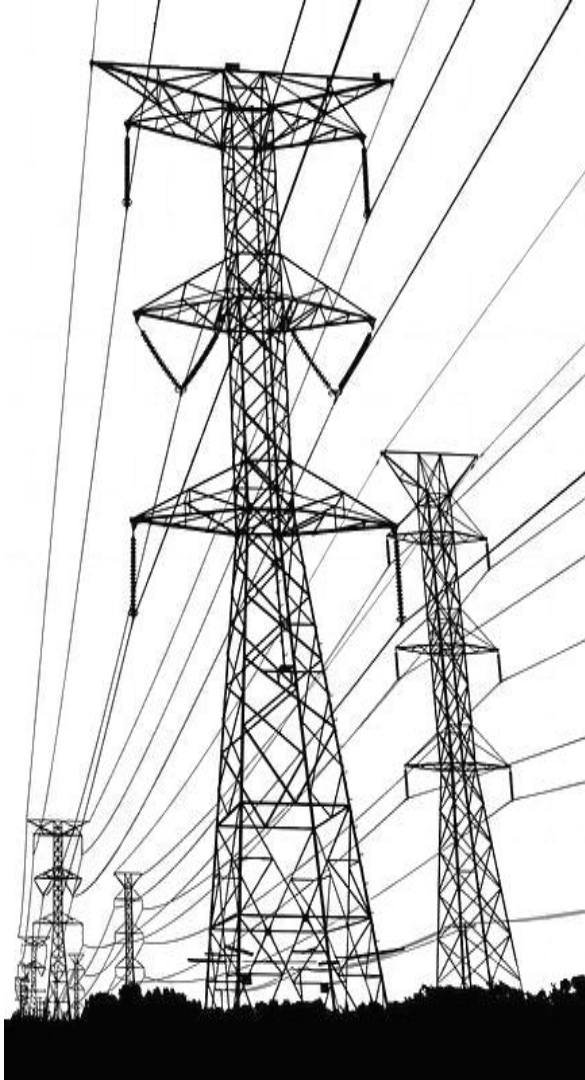
UPDATE DATA TYPES

- METER DEVICE SERIAL NUMBER ⟶ STRING
- DATETIME STAMP ⟶ UTC TO COLOMBIAN TIME ZONE

ESTABLISH ANALYSIS TIME FRAME

- UNIFY BEGINNING AND ENDING DATES
- 01/01/2023 - 26/06/2023

ERASE USELESS INFORMATION



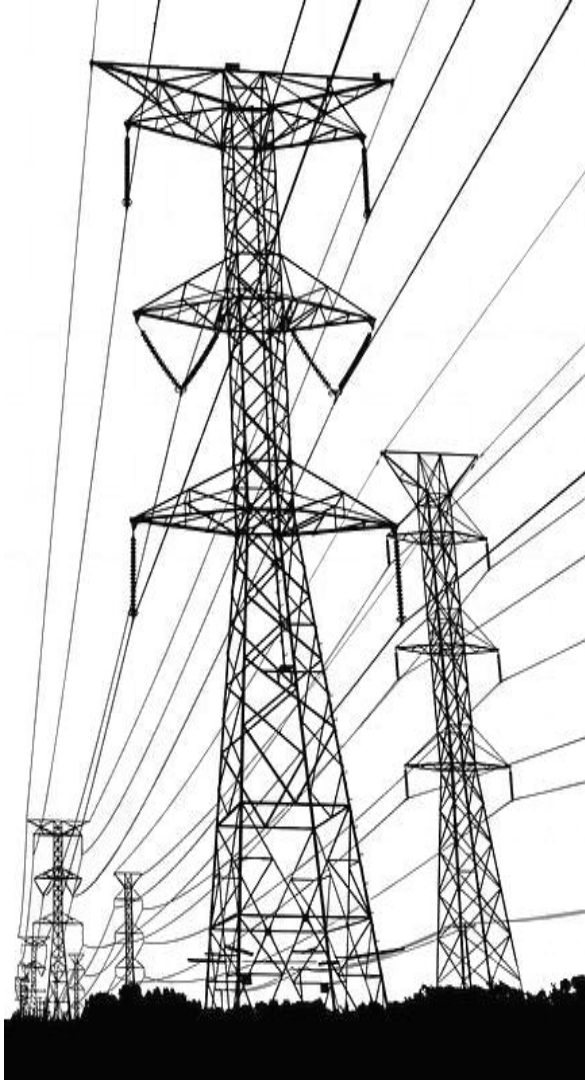
1. CLIENTS CLASSIFICATION

CLASSIFICATION ACCORDING TO THE ENERGY CONSUMPTION

VERY LOW → MINIMUM AVERAGE DAILY ENERGY CONSUMPTION. "NON-INHABITED HOUSEHOLD"

LOW →
MEDIUM →
HIGH →

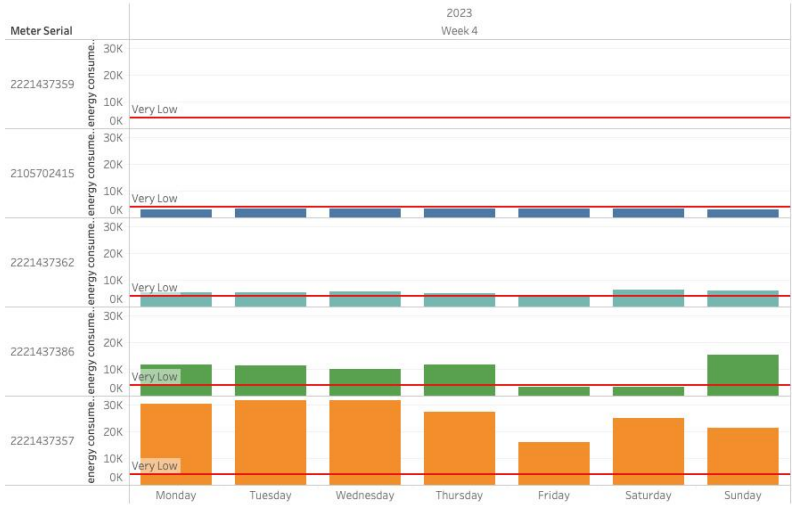
} UNSUPERVISED MACHINE LEARNING METHOD: K MEANS CLUSTERING



1. CLIENTS CLASSIFICATION

VERY LOW ENERGY CONSUMPTION (DAILY)

- FROM HISTORICAL DATA: 4000 WH / PERSON (COLOMBIA)



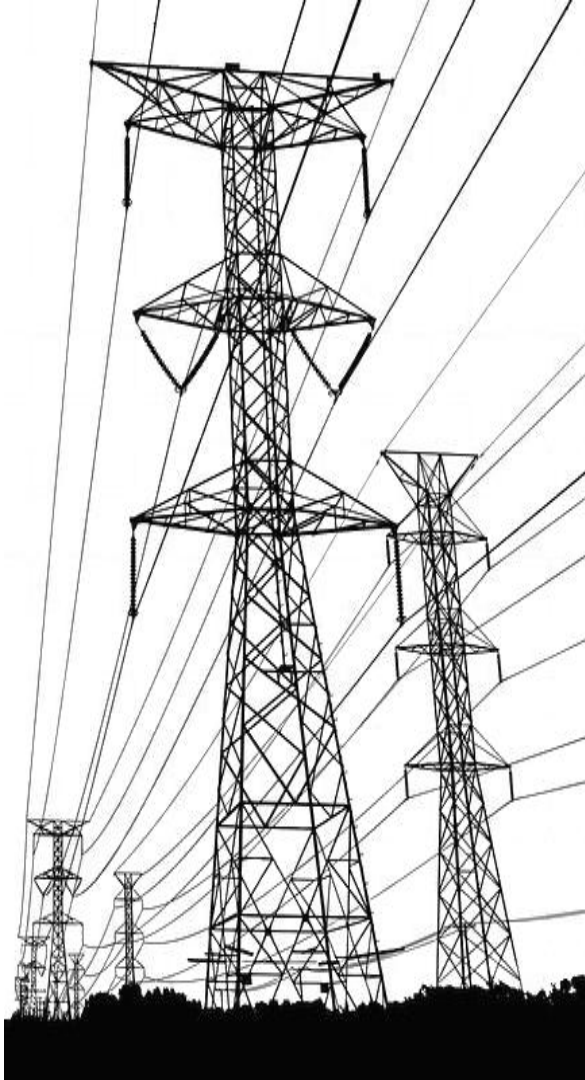
→ NO CONSUMPTION

→ VERY LOW CONSUMPTION

→ LOW CONSUMPTION

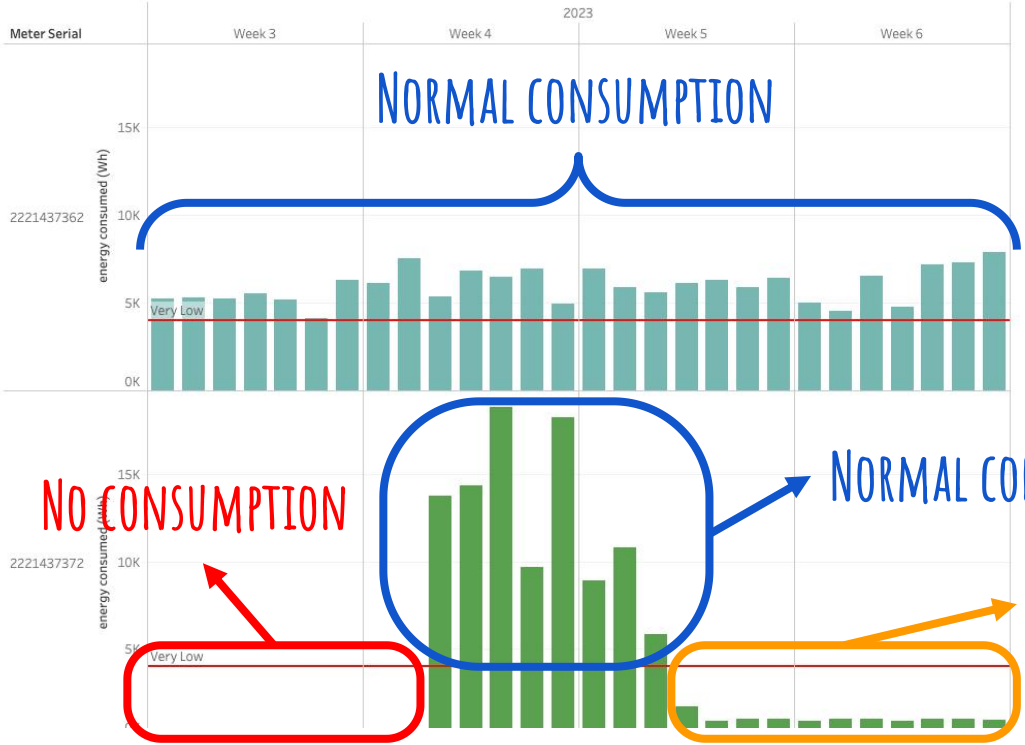
→ MIX VERY LOW - LOW

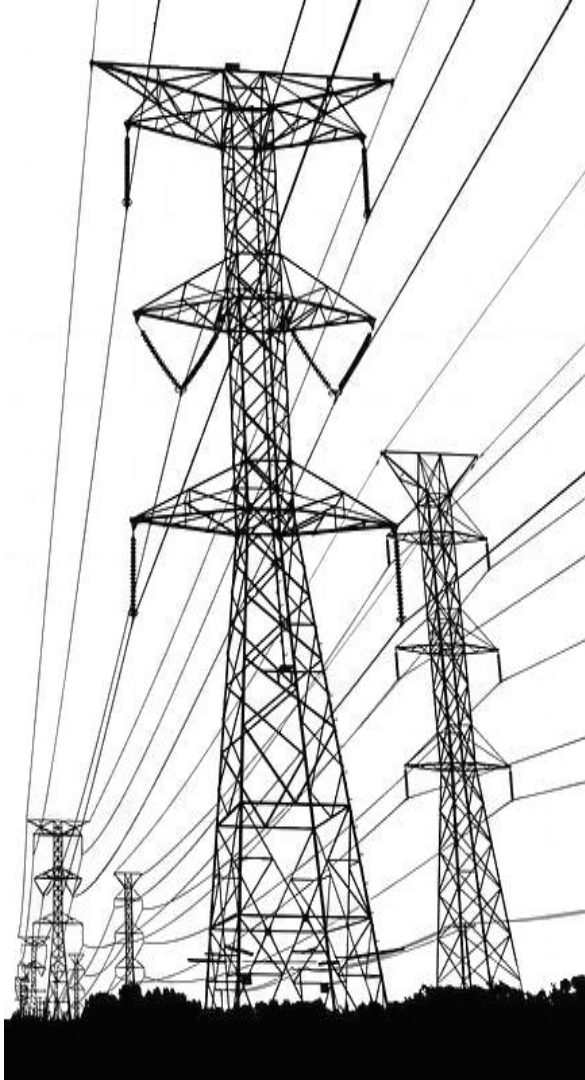
→ NORMAL CONSUMPTION



1. CLIENTS CLASSIFICATION

VERY LOW ENERGY CONSUMPTION (DAILY)





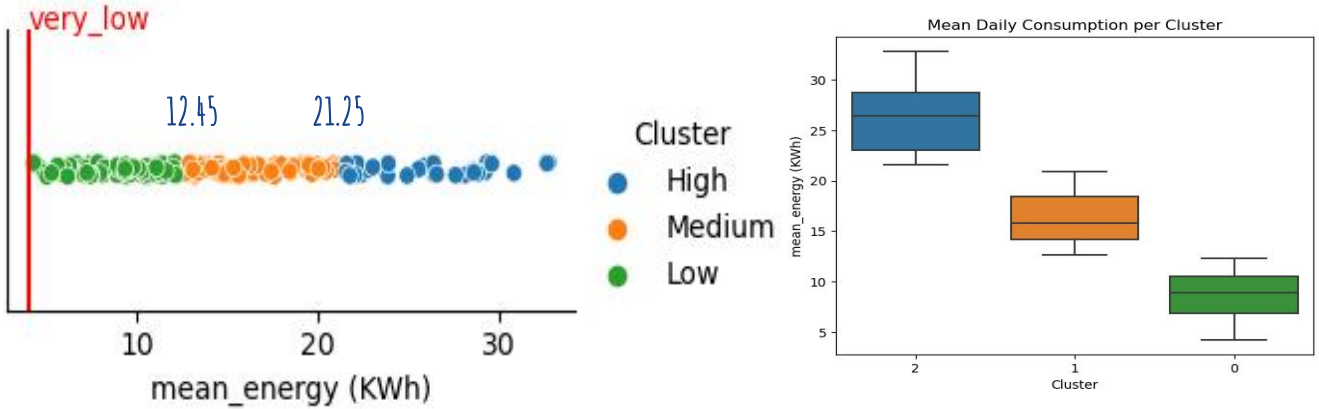
1. CLIENTS CLASSIFICATION

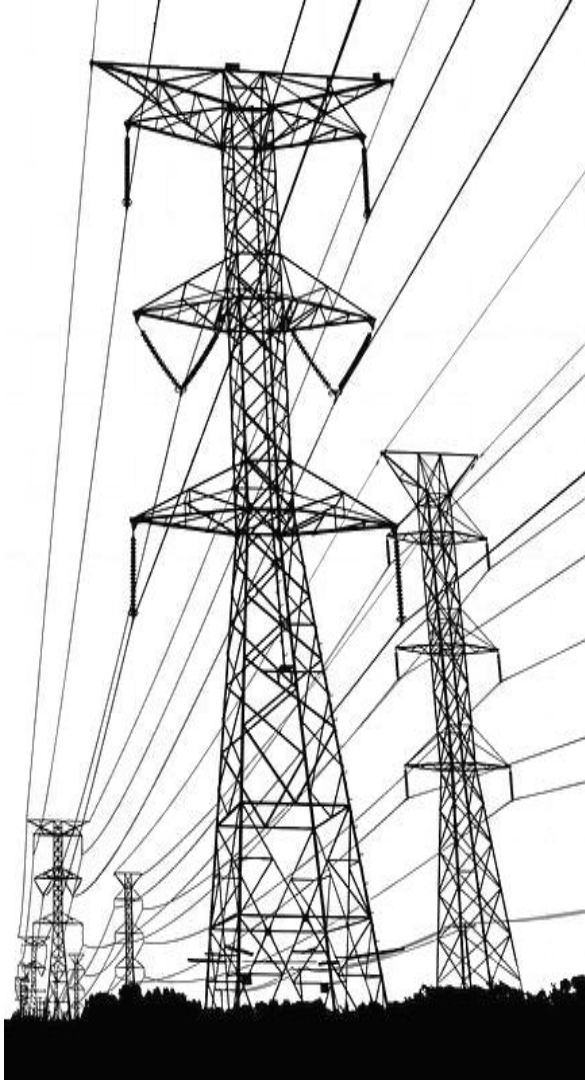
METHODOLOGY: K MEANS CLUSTERING

SCIKIT LEARN LIBRARY ⟶ SKLEARN.CLUSTER/KMEANS

KMEANS(N_CLUSTER = 3)

HIGH: 28 MEDIUM: 94 LOW: 104

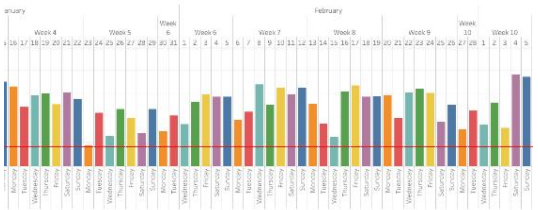


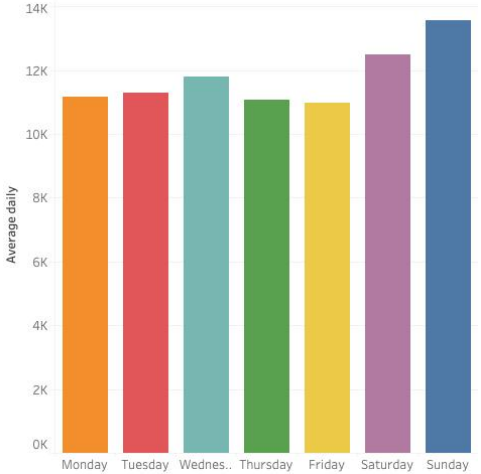


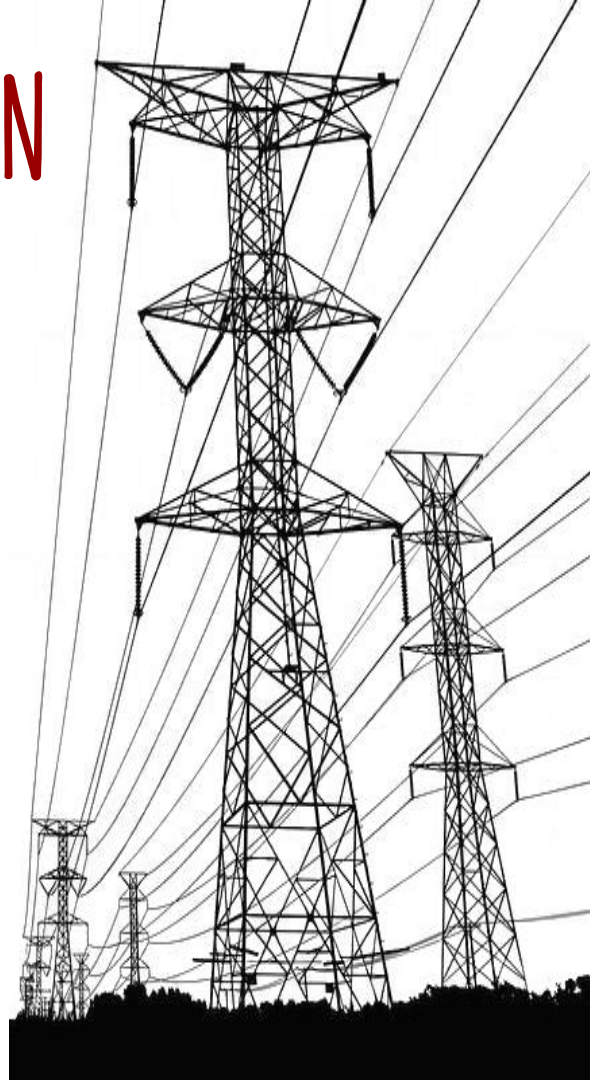
2. ENERGY CONSUMPTION PREDICTION

BASELINE OF ENERGY CONSUMPTION

AVERAGE DAILY ENERGY CONSUMPTION EXCLUDING DAYS WITH VERY LOW CONSUMPTION







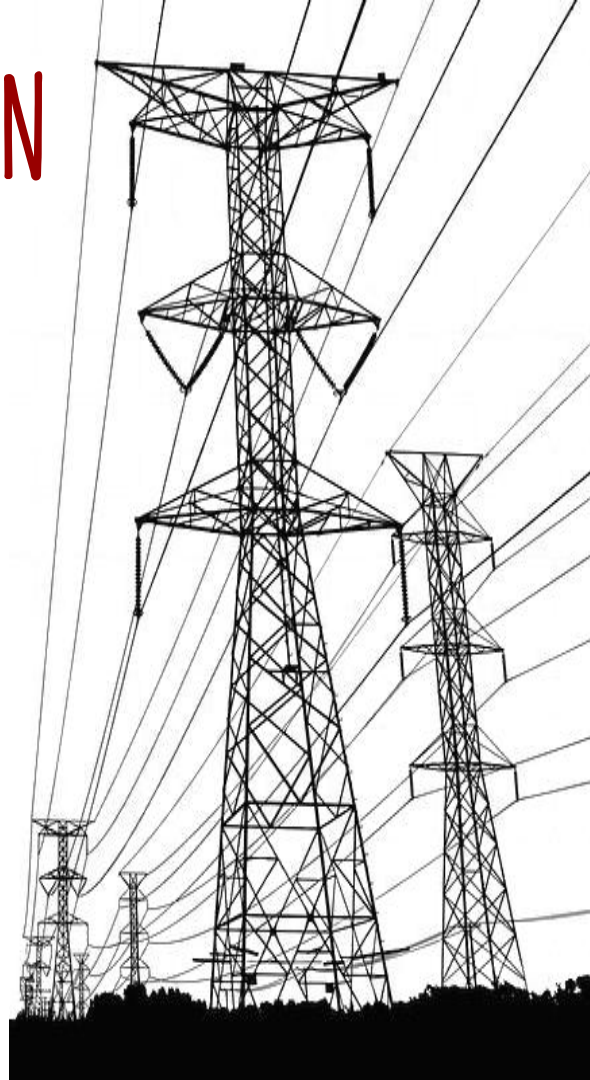
2. ENERGY CONSUMPTION PREDICTION

PREDICT ENERGY CONSUMPTION

- PREDICT DAILY ENERGY CONSUMPTION (HIGH PROFILE)
- LAST WEEK OF THE DATA SET

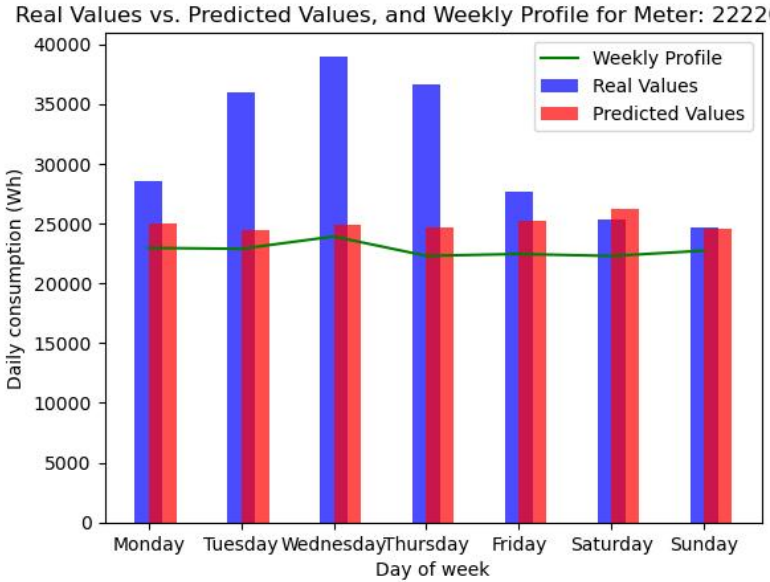
METHODOLOGY: PROPHET

- PROCEDURE FOR FORECASTING TIME SERIES

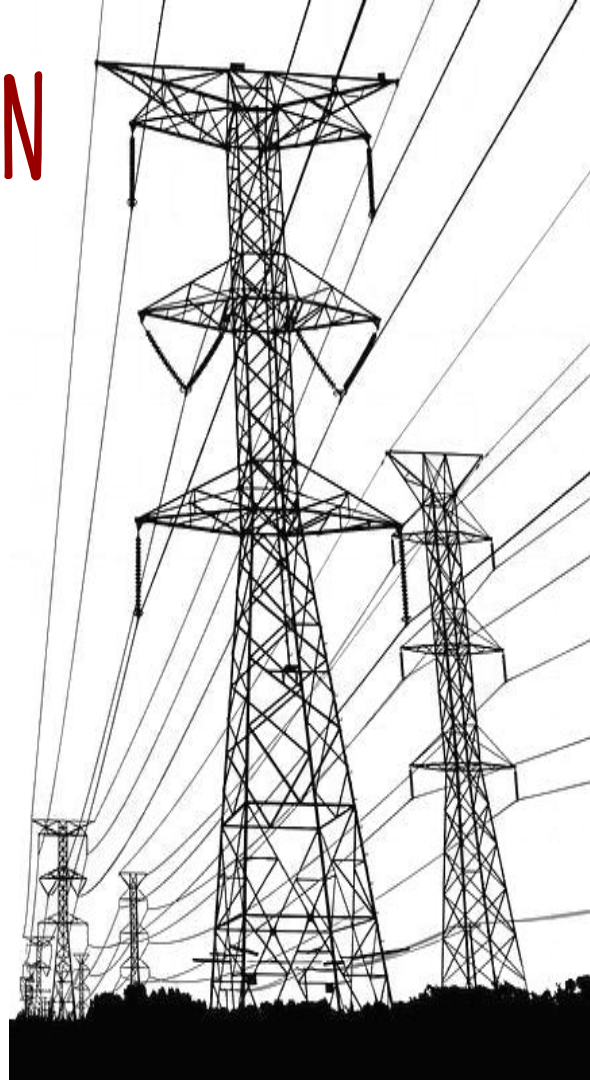


2. ENERGY CONSUMPTION PREDICTION

PREDICT ENERGY CONSUMPTION - RESULTS

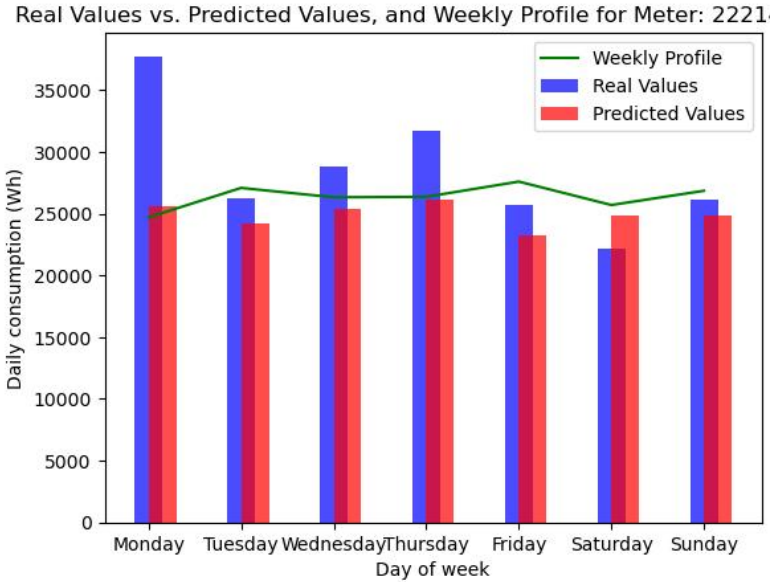


MEAN ABSOLUTE PERCENTAGE ERROR 17.9 %



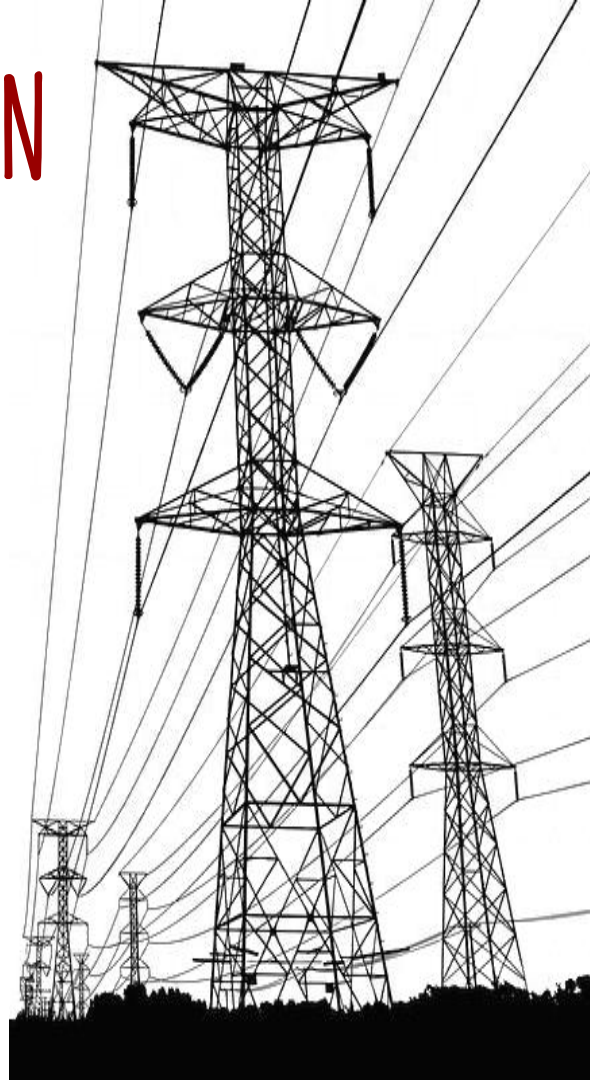
2. ENERGY CONSUMPTION PREDICTION

PREDICT ENERGY CONSUMPTION - RESULTS



MEAN ABSOLUTE PERCENTAGE ERROR
13.66 %

TOTAL M.A.P.E .7 METERS
18.66

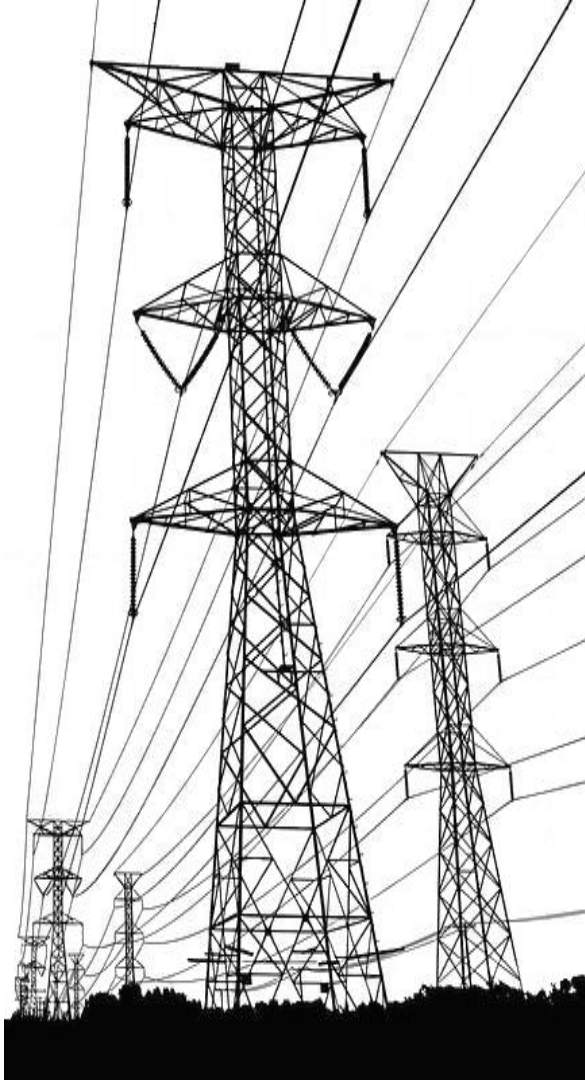


KEY TAKEAWAYS

CLIENTS CLASSIFICATION

- K MEANS METHOD OBTAINED 3 GROUPS WITH SIMILAR CONSUMPTION RANGE SIZES (LOW, MEDIUM, HIGH)

- MINORITY OF HOUSEHOLDS ON THE HIGH CONSUMPTION RANGE (13%)



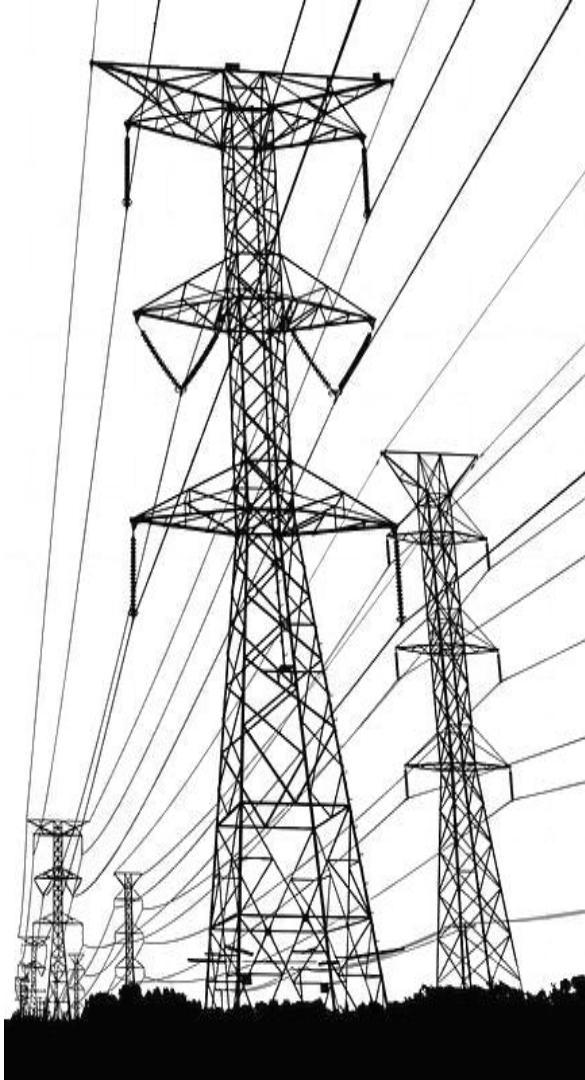
KEY TAKEAWAYS

PREDICT ENERGY CONSUMPTION

- PROPHET METHOD PREDICTED THE DAILY CONSUMPTION WITH 18.66% OF AVERAGE ERROR ON 7 METERS

- CONSIDERABLE ERROR DUE TO TESTING WEEKS FAR FROM NORMAL CONSUMPTION.

- DIFFICULTY TO PREDICT CONSUMPTION ON METERS WITH VERY LOW CONSUMPTION PERIODS.



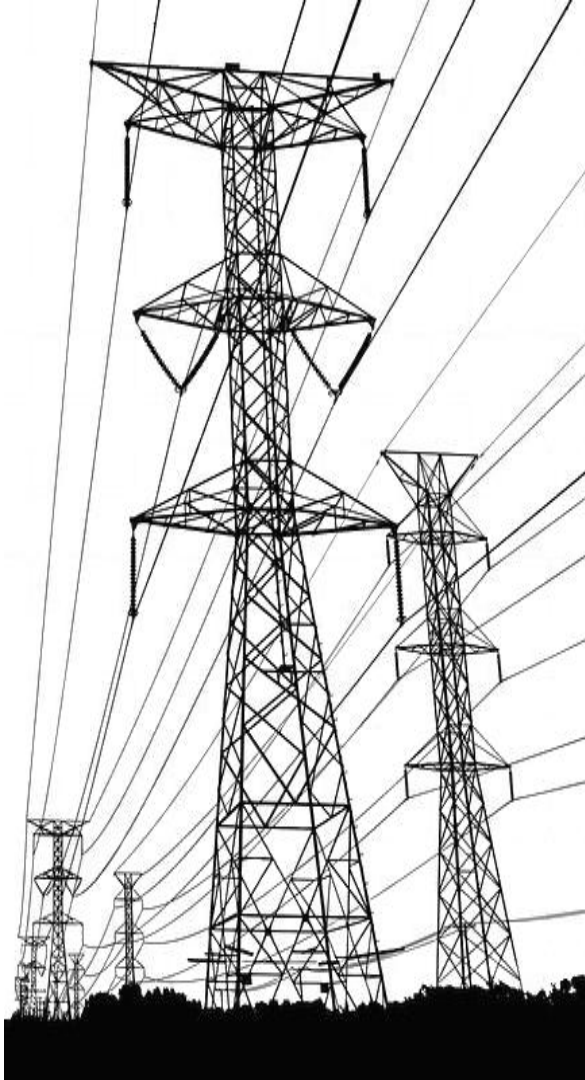
Next steps

Clients classification

- Optimize calculation of the "Very low" threshold
- Perform classification based on more than 1 parameter

Predict energy consumption

- Use a data set with a bigger time frame
- Improve the prediction including meter with mixed (Normal + very low) consumption



THANKS

ANY QUESTIONS?