

# **Advanced Metering Infrastructure (AMI) Data Analysis**

Presenter

Abu Junior Vandi & Joseph Prince Conteh



# Project Overview

Advanced Metering Infrastructure (AMI) Data Analysis

Introduction & Objectives

Dataset Overview

Tools &  
Technologies  
Used

Data  
Preprocessing

Key Findings &  
Visualizations

# Introduction & Objectives

- Simulate an AMI environment for data analysis
- Analyze energy consumption patterns & Anomalies
- Forecast energy demand for optimization
- Provide actionable insights for utilities

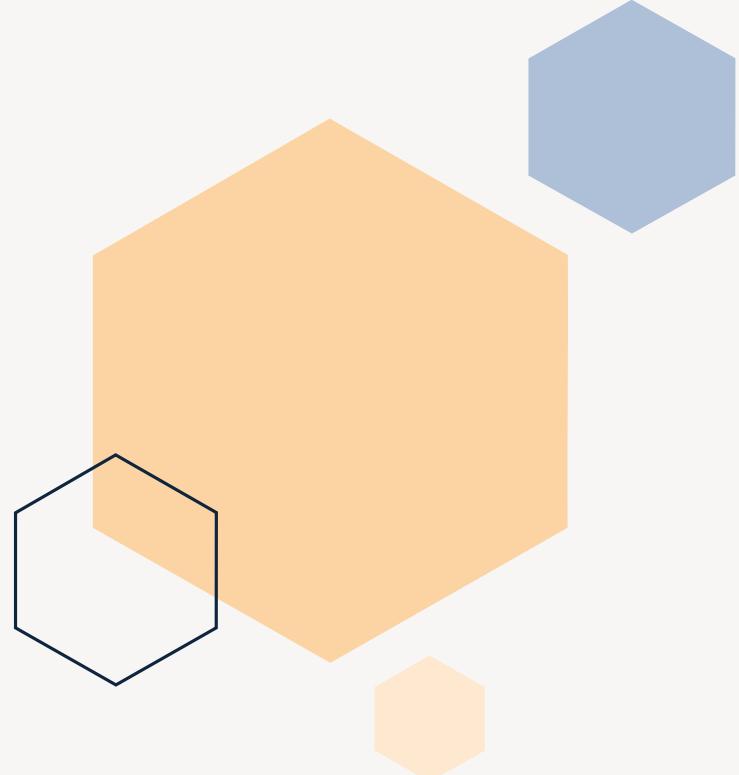
Advanced Metering Infrastructure (AMI) Data Analysis





## ❑ Primary goals

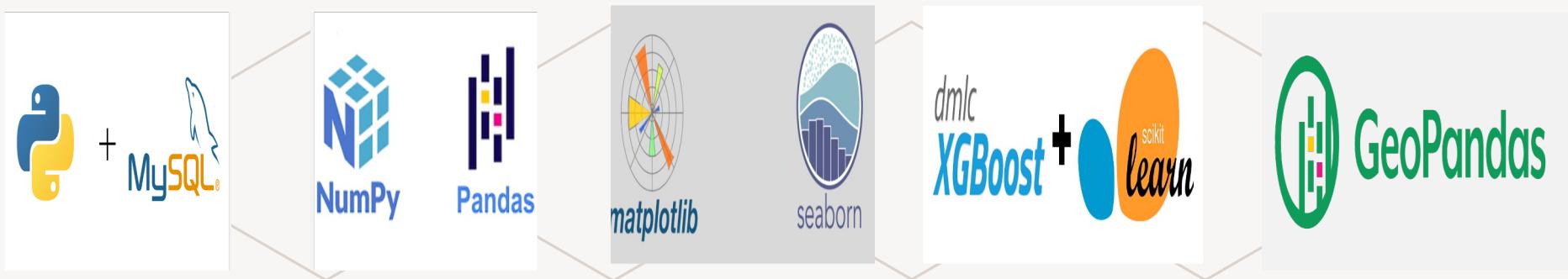
- ❑ Hourly consumption readings
- ❑ Includes billing, downtime, and location data
- ❑ Categorized by meter type & seasonality
- ❑ 1,000 meters over 1 year (2023)



## ❑ Tools & Technologies Used

- ❑ Python, MySQL, Pandas, NumPy
- ❑ Matplotlib, Seaborn for visualization
- ❑ Scikit-learn & XGBoost for ML modeling
- ❑ GeoPandas & Folium for GIS analysis

# Tools & Technologies Details



## Python & MySQL

Python, used for data processing and machine learning, was integrated with MySQL to efficiently store and manage AMI data.

## Pandas, NumPy

Pandas – Used for data manipulation and cleaning;  
NumPy – Handles numerical computations and array operations.

## Matplotlib, Seaborn

Visualization libraries used to create graphs, charts, and plots for data analysis and insights.

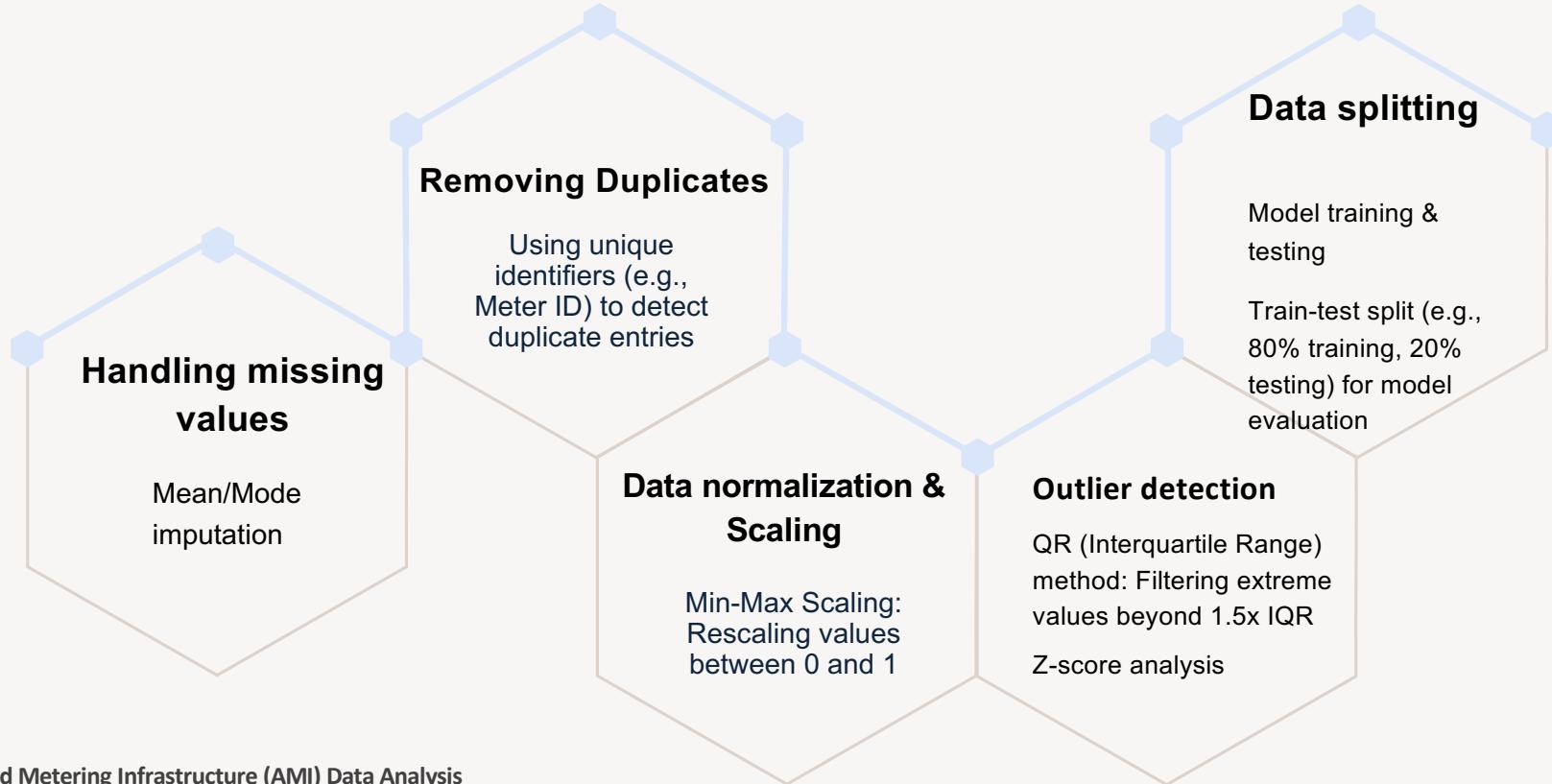
## Scikit-learn & XGBoost

Scikit-learn offers ML for clustering, anomaly detection, and predictive modeling, while XGBoost, for forecasting and prediction accuracy.

## GeoPandas & Folium

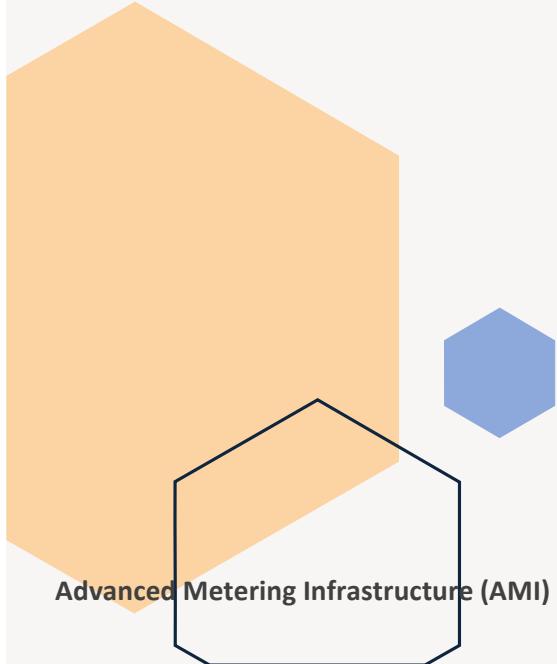
GIS tools used for mapping energy consumption patterns and performing spatial analysis.

# Data Preprocessing

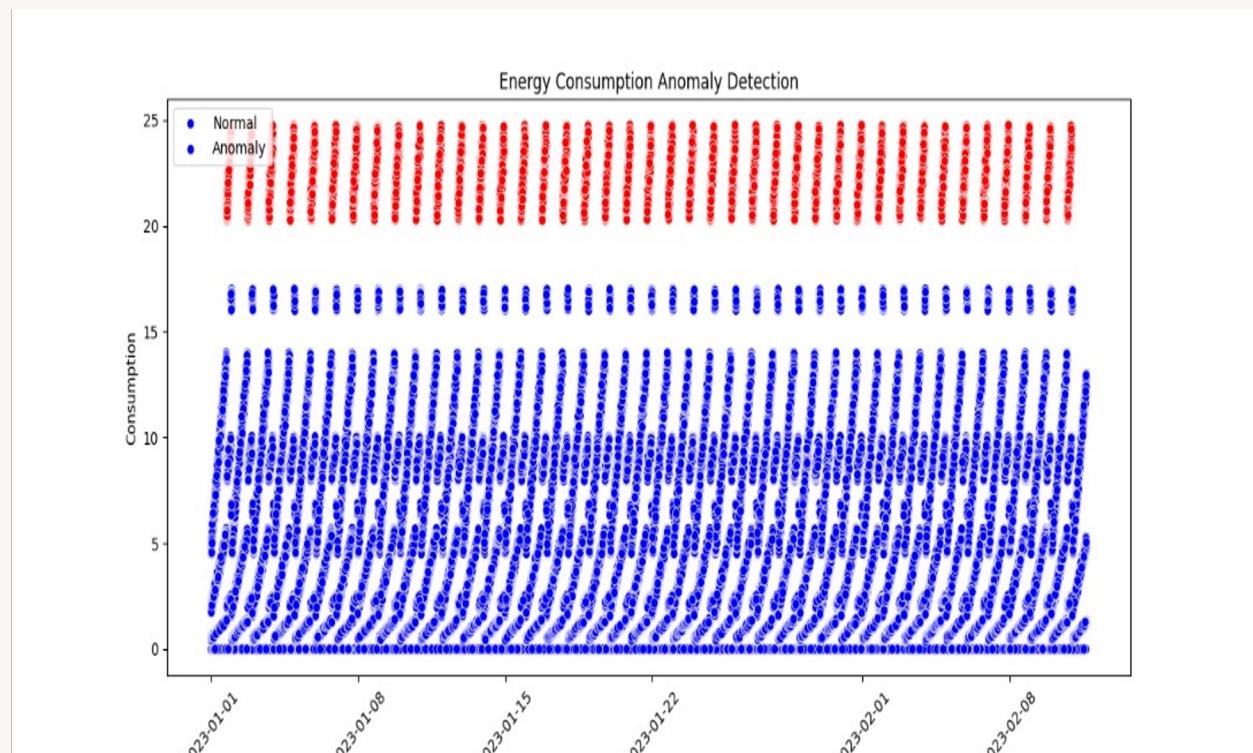
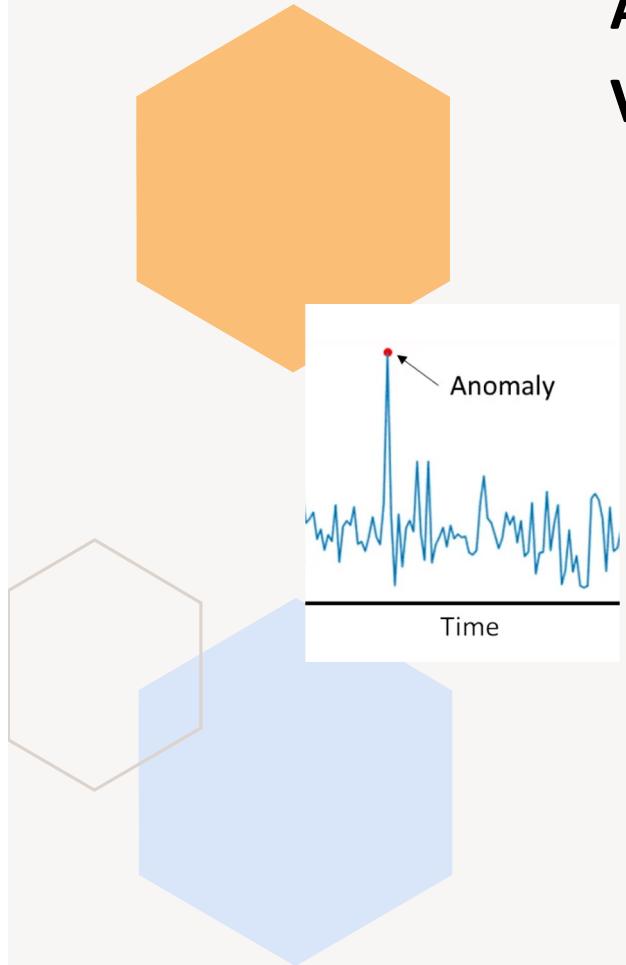


# Exploratory Data Analysis (EDA)

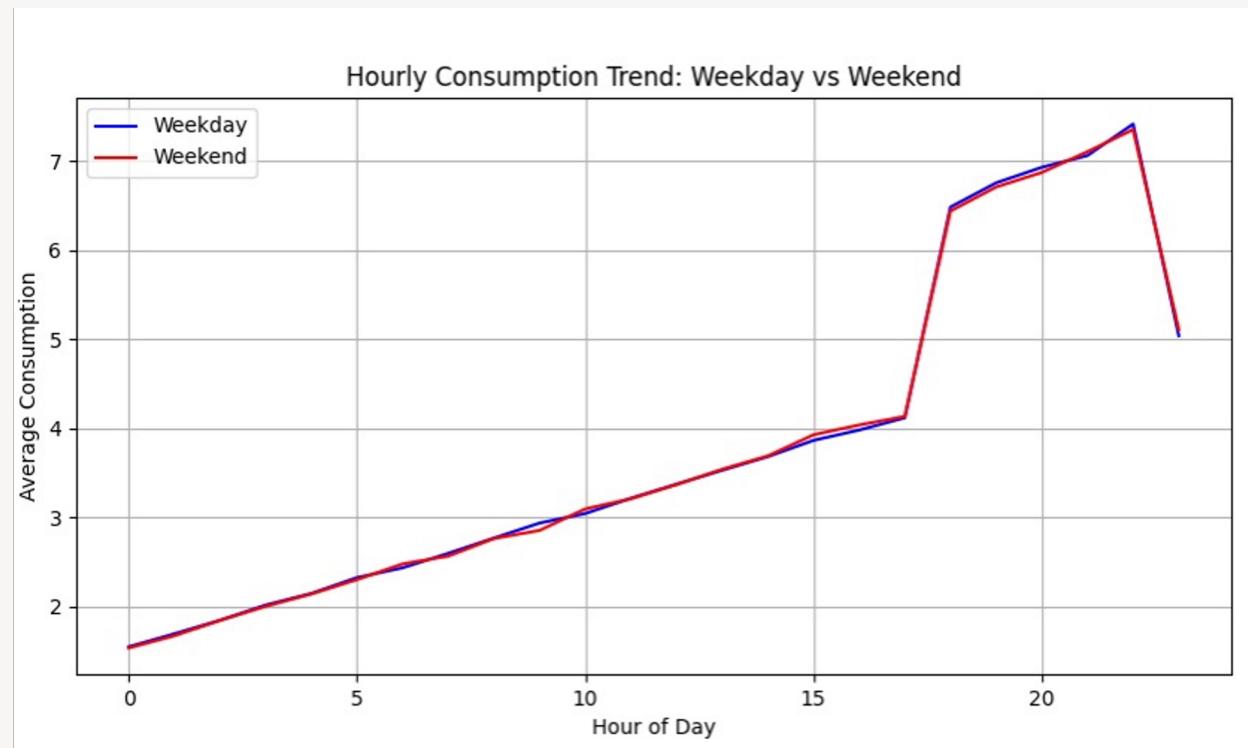
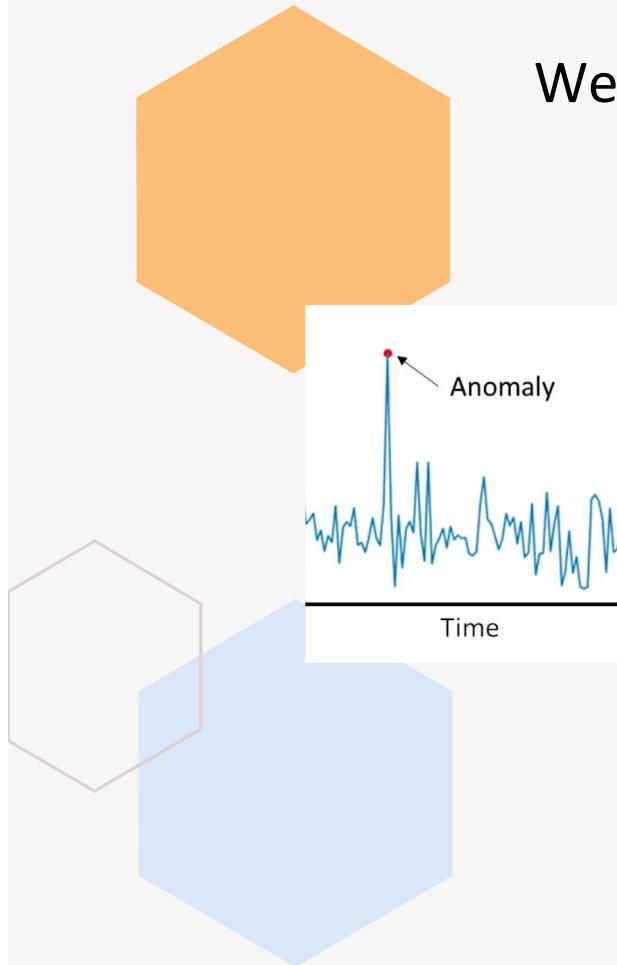
- Anomaly Detection using Z-score
- Weekday and Weekend Analysis
- Consumption Clustering via K-Means
- Load Forecasting using ML models
- Billing Analysis for discrepancy detection



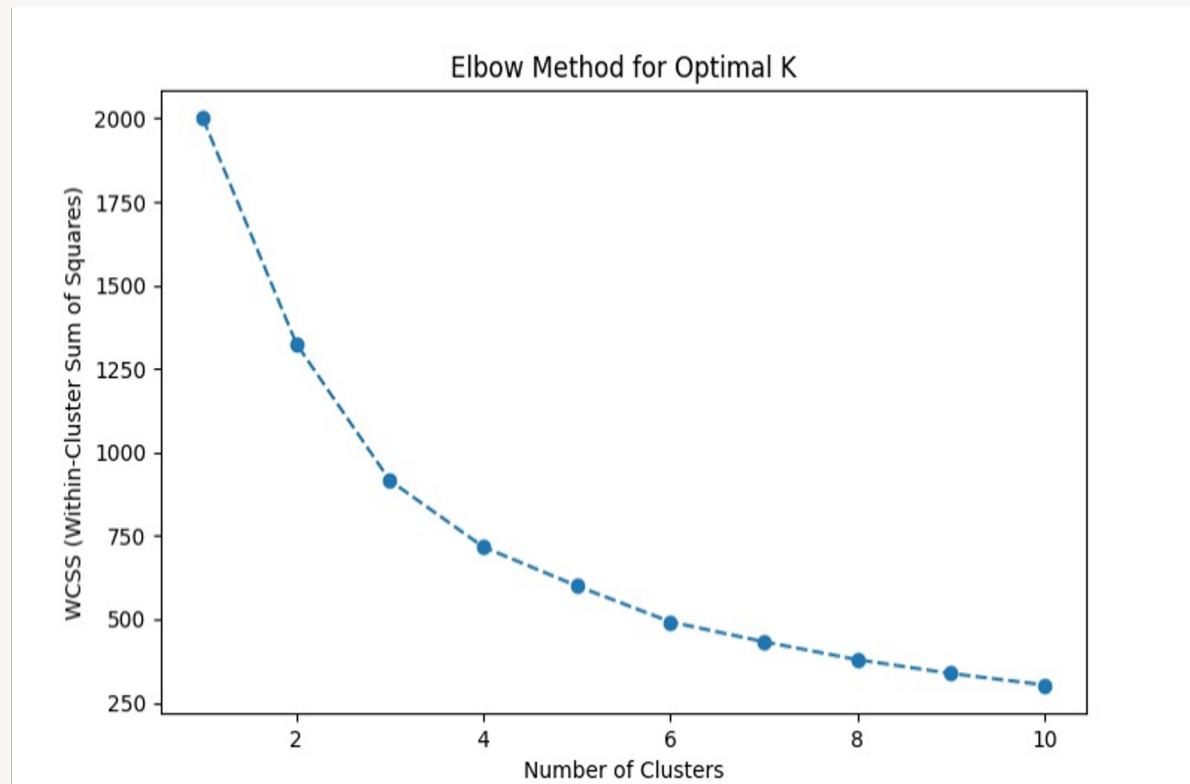
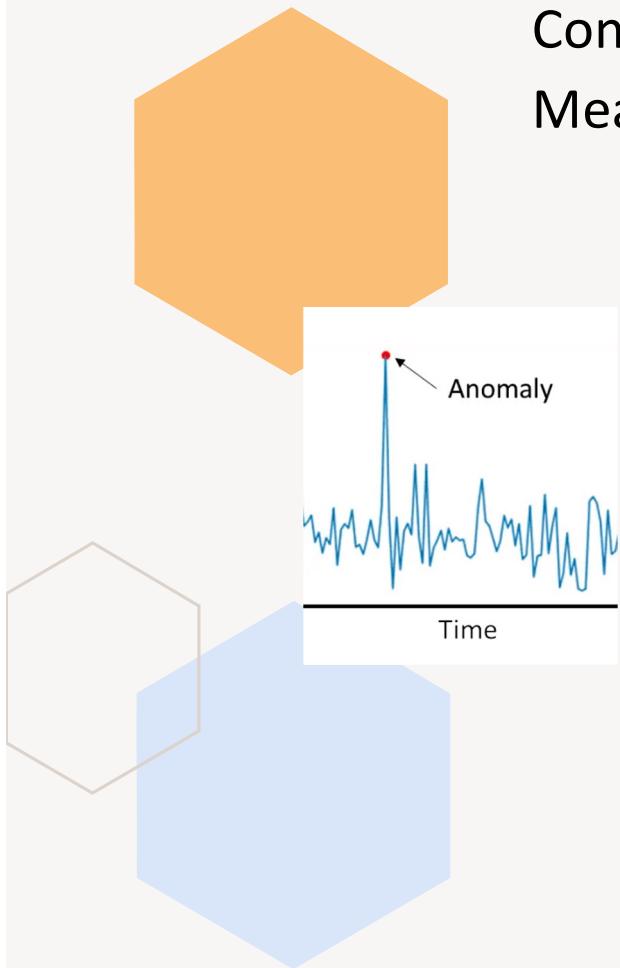
# Anomaly Detection Visualization



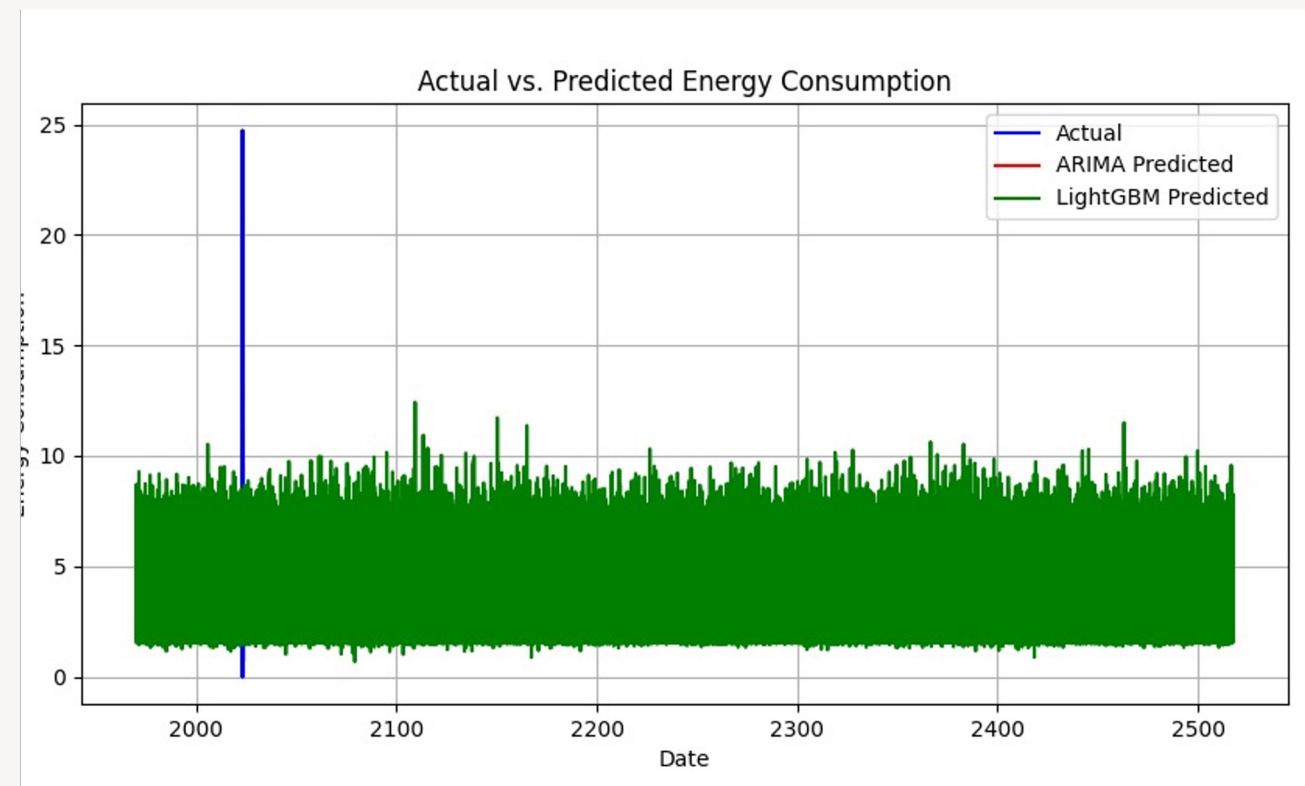
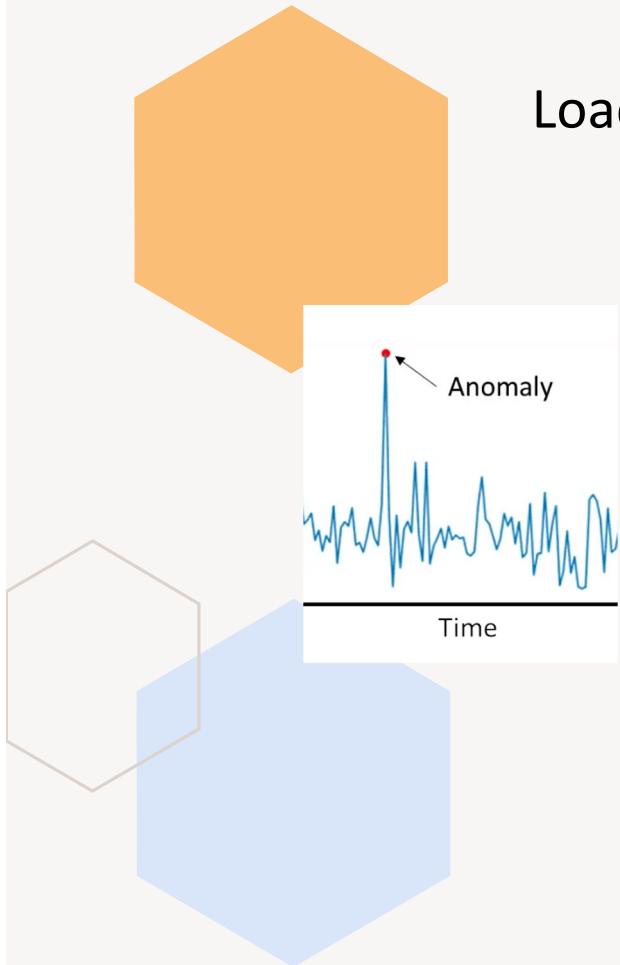
## Weekday and Weekend Analysis



## Consumption Clustering via K-Means



# Load Forecasting



# Challenges & Solutions

## 📌 Data Volume & Processing Efficiency

**Challenge:** 9M+ hourly readings in MySQL caused delays due to high computational demands.

**Solution:** Chunk-wise processing & optimized SQL queries improved efficiency and reduced response time.

## 📌 Handling Missing & Anomalous Data

**Challenge:** Missing values and outliers led to biased results.

**Solution:** Used statistical imputation (mean/median) and Z-score analysis to detect and remove anomalies.

## 📌 GIS Data Integration

**Challenge:** High-granularity GIS data was difficult to process and visualize in the terminal.

**Solution:**

- Used **Folium & GeoPandas** for interactive HTML-based maps.



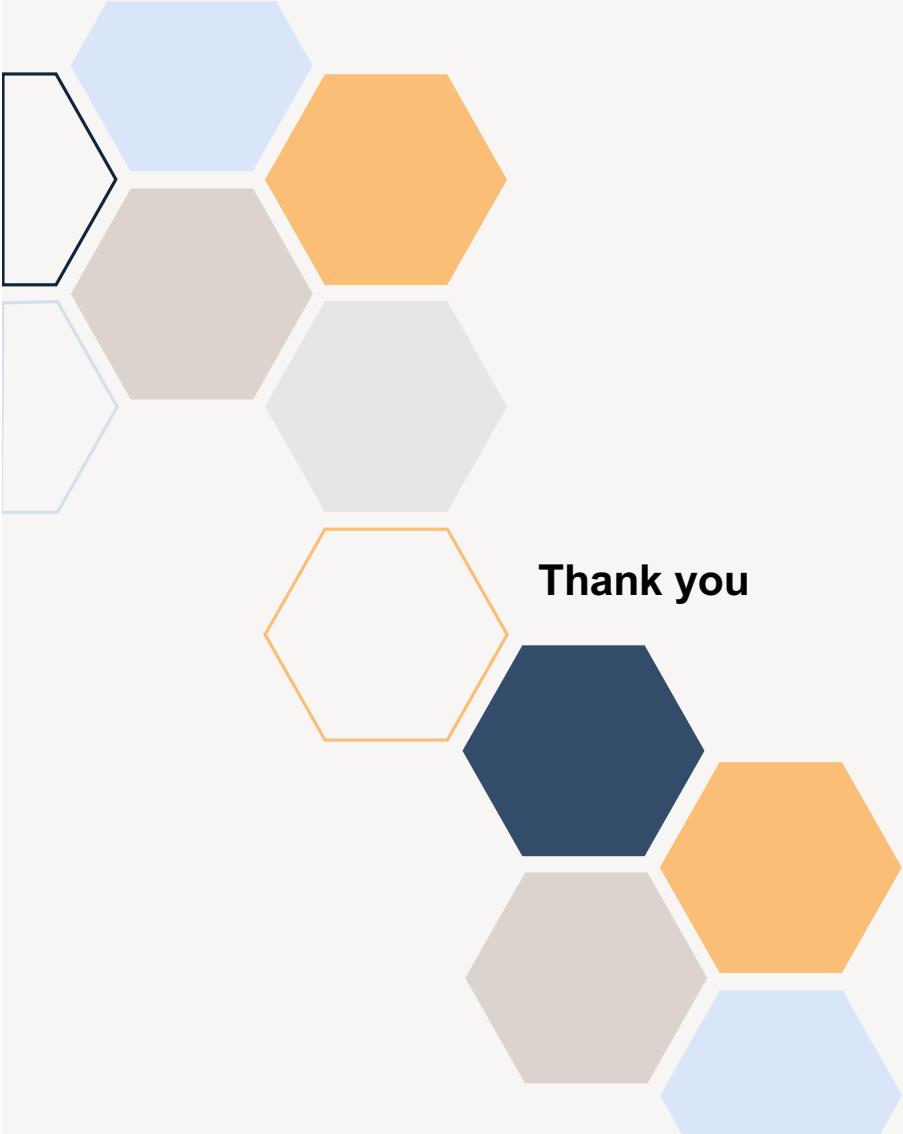
# Meet My Team



**Abu Junior Vandi**



**Joseph Prince Conteh**



# QUESTION



Abu Jr Vandi

[abujuniorv@gmail.com](mailto:abujuniorv@gmail.com) - [josephprinceconteh@gmail.com](mailto:josephprinceconteh@gmail.com)

+232 73914398

Joseph Prince Conteh

+23276366095

