**📝 Meeting Report**

**Date:** April 24, 2025  
**Subject:** Project Progress – Power Consumption Model  
**Attendees:** Project Team

**1. Data Adequacy Review**

* The current dataset sourced from Kaggle was evaluated for completeness and relevance.
* It was found to be **inadequate** in terms of granularity and volume, particularly for building a reliable machine learning model.
* As a result, it was concluded that **additional synthetic data** must be generated to move forward effectively.

**2. Project Objective**

* The primary goal of the project is to **develop a model that identifies households with higher energy consumption**.
* The analysis will be based on features such as:
  + Date and Time
  + House ID
  + Total Power Consumed (kV)
  + Meter Readings
  + Location
  + Breakdown of consumption across categories:
    - Entertainment
    - White Goods
    - Lighting
    - Air Conditioners
    - EV Items

**3. Plan for the Upcoming 2 Days**

* **Generate Data:** Create detailed, realistic power consumption data for **50 individual houses**.
* **Configure House Profiles:** Each house will be assigned unique characteristics (location, usage patterns, etc.) to simulate real-world diversity.
* **Prepare Dataset:** The synthetic dataset will be structured and stored in CSV format for model training and testing.

Date: April 28, 2025  
Topic: Data Generation Code Improvement

This is a brief summary of our meeting held on April 28, 2025 regarding improvements to the data generation process.

Key Update:  
We discussed and implemented changes to the entire codebase responsible for synthetic data generation. The primary improvement was in changing the data types and logic used in the script to enhance the quality and realism of the generated data.

**Meeting Report – April 29**

**Key Updates:**

1. **Enhancements to Dataset Structure:**
   * Added new columns including **date range** to support more robust model training and evaluation.
   * Ensured the dataset now better reflects real-world usage intervals and tracking needs.
2. **Data Quality Improvements:**
   * Addressed and **rectified issues** in the previously generated data (e.g., timestamp inconsistencies, logical errors in usage simulation).
   * Cleaned and validated appliance usage patterns for improved reliability.
3. **Dataset Optimization:**
   * Refined column organization and added **category-based summaries** (e.g., white\_goods, entertainment, lighting, etc.).
   * Enhanced compatibility with downstream **analytics and machine learning workflows**.

**Next Steps:**

* Begin integration with model training pipeline.
* Continue validating outputs with domain benchmarks.

**Meeting Report – May 1, 2025**

**Date:** May 1, 2025  
**Participants:** [Your Team Name or Attendees]  
**Duration:** [Optional – e.g., 30 mins]  
**Facilitator:** [Optional – e.g., Your Name]

**✅ Key Updates & Actions**

1. **GitHub Repository Created**
   * A new GitHub repository has been set up to centralize project collaboration and version control.
2. **Improved Data Quality**
   * Enhancements were made to increase the level of information in the dataset, leading to improved data quality and context.
3. **Organized Project Structure**
   * Separate folders were created for scripts, data, and configurations to streamline file management and development workflows.
4. **Data Generator Code Updated**
   * Modifications were made to the data generation script to support:
     + More accurate energy consumption logic
     + Proper handling of time-based usage
     + Structured CSV outputs with per-category and per-device detail