Project 4: Measure Energy Consumption

PHASE-5 PROJECT SUBMISSION

Team Member: Senoj S
Reg No: 31112120509

Objective:

The objective of this project is to create an automated system that measures energy consumption, analyzes the data, and provides visualizations for informed decision-making. This solution aims to enhance efficiency, accuracy, and ease of understanding in managing energy consumption across various sectors.

In this phase we concentrate on analyzing and visualizing the preprocessed data that we got from the previous phase.

Problem Statement:

Manual recording of energy consumption in houses may leads to minor errors or require time and human resource.Instead, a automated system like a software to monitor and measure day to day energy consumption in modern houses helps a lot.

Datasets Used:

All datasets are derived from kaggle link given in the project module

1.pjm_hourly_est.csv
2.PJM_Load_hourly.csv
3.PJME_hourly.csv
4.PJMW_hourly.csv
5.NI_hourly.csv
6.FE_hourly.csv
7.AEP_hourly.csv
8.COMED_hourly.csv
9.DEOK_hourly.csv
10.DOM_hourly.csv
11.DUQ_hourly.csv
12.EKPC_hourly.csv

Data preprocessing:

1.Data Inspection:

Begin by loading your dataset and taking a close look at its structure. Use functions like head(), info(), and describe() in Python's Pandas library to get a preliminary understanding of the data.

2.Handling Missing Data:

Identify the missing rows. Remove rows with missing values if they are relatively small in number and won't significantly impact the analysis. Impute missing values using methods such as mean, median, mode, or more sophisticated imputation techniques, like k-nearest neighbors or regression.

3.Handling Duplicates:

Check for and remove duplicate records, if any, as they can distort your analysis.

Visualization Techniques:

For visualizing the data we use basic level graphics plotting library Matplotlib,Seaborn for accessing various types of plotting and correlation heatmaps.