

Individual Household Electric Power Consumption

The main objective of this Exploratory Data Analysis (EDA) project is to examine household electricity usage patterns to derive meaningful insights, detect inefficiencies, and enhance energy management

Problem Statement:

Electricity consumption in households fluctuates throughout the day and across various appliances. Gaining a clear understanding of these patterns is essential for optimizing energy efficiency, lowering expenses, and promoting sustainable power usage. However, many households lack detailed information on their electricity consumption, the appliances that contribute most to their energy use, and the effects of voltage fluctuations on overall power consumption.

EDA QUESTIONS FOR THE DATASET

Basic Data Understanding

1. What is the total number of records in the dataset?
2. Are there any missing values in the dataset? If so, in which columns?
3. What are the data types of each column?
4. What is the range of dates present in the dataset?

Descriptive Statistics

5. First you have to make a `Total_Submetering` column by adding the all `submeterings`.
6. What are the summary statistics (mean, median, min, max, standard deviation) for `global_active_power`, `global_reactive_power`, `voltage`, and `global_intensity`?
7. What is the mean and standard deviation of the `Total_Submetering` column?
8. What is the distribution of `voltage` in the dataset? (e.g., histogram or KDE plot)
9. What is the correlation between `global_active_power` and `global_intensity`?
10. How does `global_active_power` vary with `voltage`?
11. What are the minimum and maximum values of `global_active_power`, and at what times did they occur?

Time-Based Analysis

11. How does `global_active_power` change over different hours of the day? (Hourly trend)
12. Are there any noticeable trends in power consumption over different days of the week?
13. How does `Total_Submetering` change over different hours of the day?
14. How does `voltage` fluctuate throughout a single day?
15. What is the average `global_active_power` usage for each month?

Sub-Metering Analysis

16. Which sub-metering category (sub_metering_1, sub_metering_2, or sub_metering_3) consumes the most power on average?
17. How does Total_Submetering compare with global_active_power? Is there a strong correlation?
18. How much energy is consumed outside of the three sub-metering categories (global_active_power - Total_Submetering)?
19. Are there any specific time periods where Total_Submetering is significantly higher or lower than usual?
20. Which hour of the day has the highest Total_Submetering on average?