Data Appendix

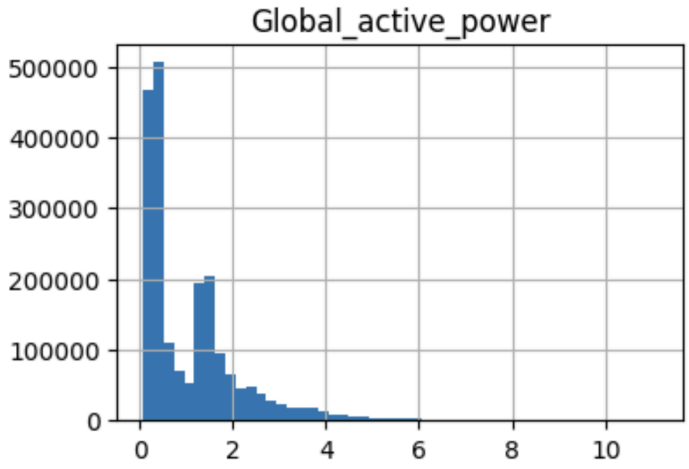
Analysis Data File: Power Consumption Analysis

Unit of Observation:
Each row in this dataset represents a single observation of power consumption for a household, recorded at one-minute intervals over 47 months. The dataset includes various features related to power consumption, voltage, and energy sub-metering, with some instances containing missing values.

Overview of Analysis Data File:
The cleaned dataset, `power_data_cleaned`, is derived from the original file by removing rows with missing values and resetting the index. This cleaned dataset serves as the basis for subsequent analysis.

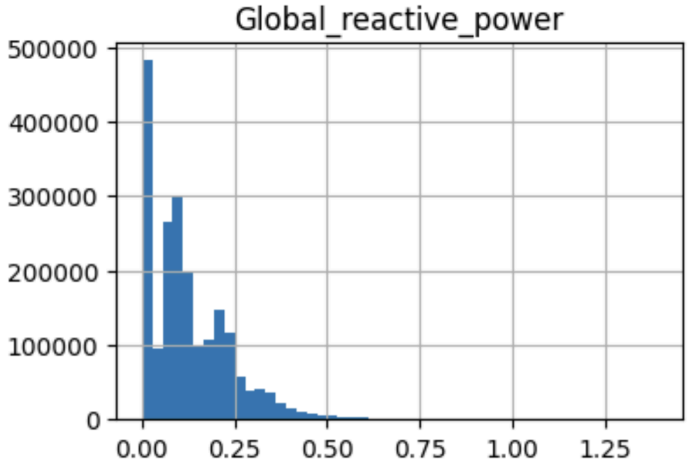
Variables in the Analysis Data File:

1. **Global Active Power**
    - **Definition:** The total active power consumed by the household in kilowatts.
    - **Type:** Quantitative
    - **Summary Statistics:**
        - **Mean:** 1.0916 kW
        - **Standard Deviation:** 1.0573
    - **Visualization:** Histogram illustrating the distribution of global active power usage.

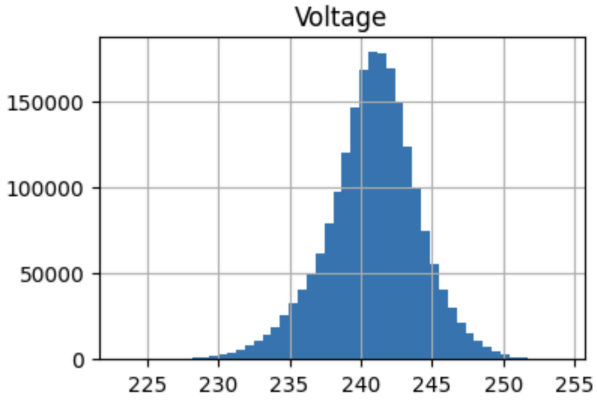


2. **Global Reactive Power**
    - **Definition:** The reactive power consumed in the household, representing non-productive power consumption.

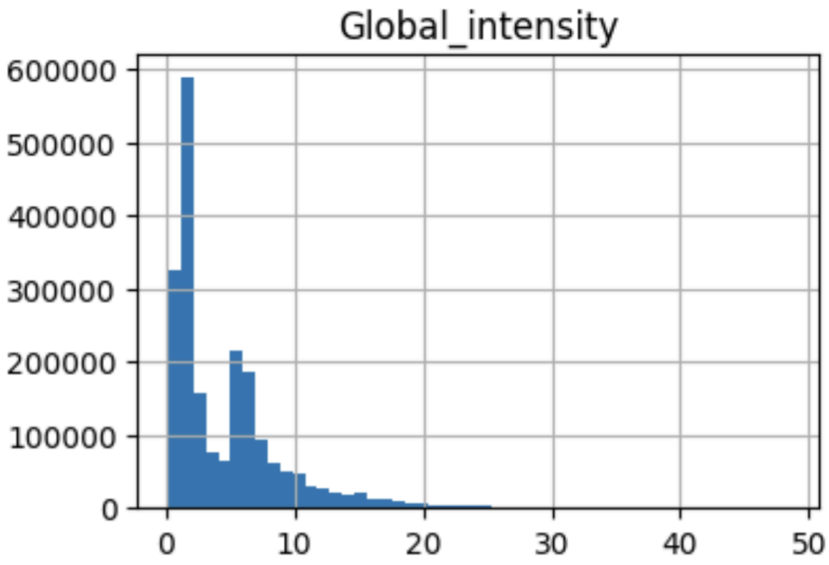
- **Type:** Quantitative
- **Summary Statistics:**
  - **Mean:** 0.1237 kW
  - **Standard Deviation:** 0.1127
- **Visualization:** Histogram showing the distribution of global reactive power usage.



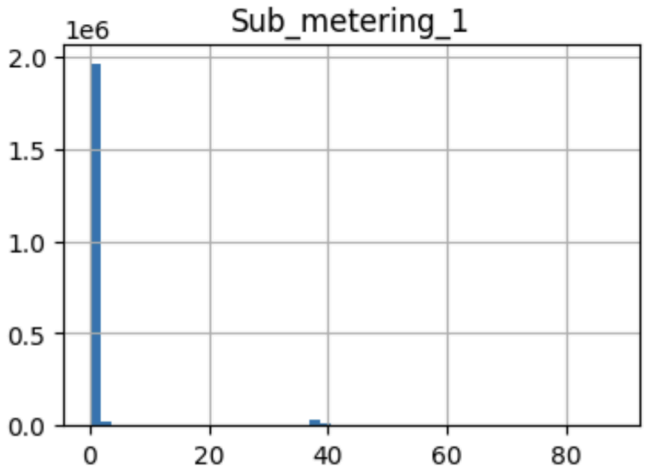
3. **Voltage**
   - **Definition:** Voltage measured at the household's distribution panel.
   - **Type:** Quantitative
   - **Summary Statistics:**
     - **Mean:** 240.8399 V
     - **Standard Deviation:** 3.2399
   - **Visualization:** Line plot showing voltage fluctuations over time.



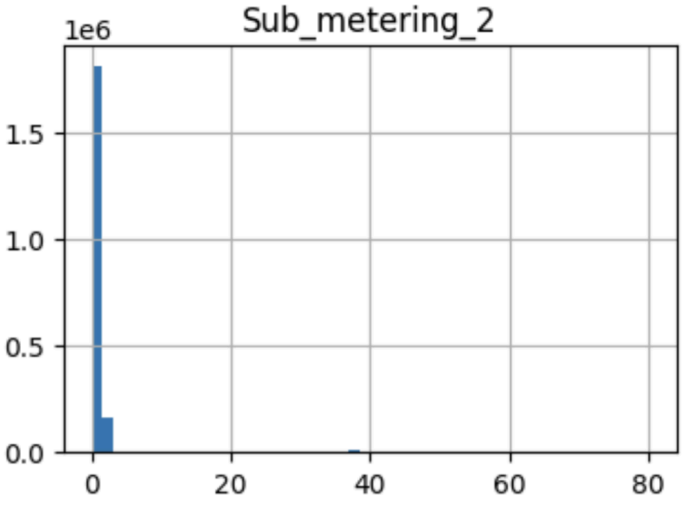
4. **Global Intensity**
    - **Definition:** Total current intensity of the household in amperes.
    - **Type:** Quantitative
    - **Summary Statistics:**
        - **Mean:** 4.6278 A
        - **Standard Deviation:** 4.4444
    - **Visualization:** Histogram illustrating the distribution of global intensity.



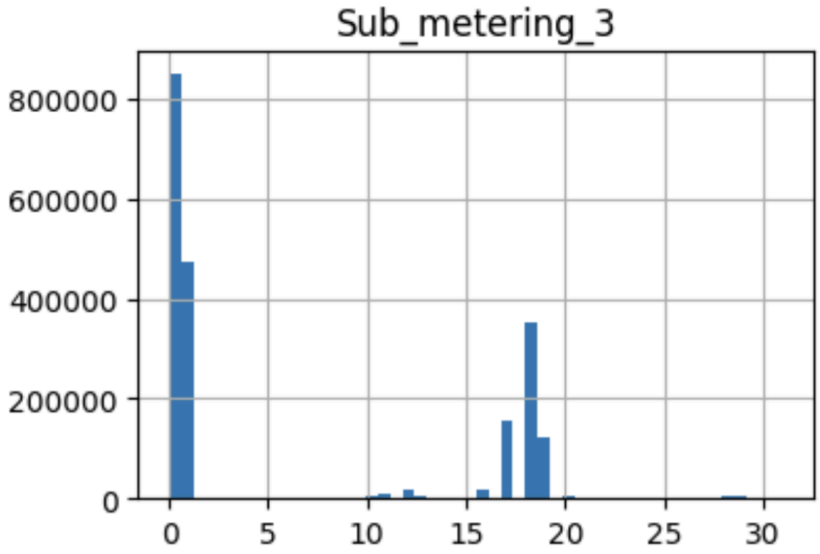
5. **Sub-Metering 1**
    - **Definition:** Energy sub-metering for kitchen appliances.
    - **Type:** Quantitative
    - **Summary Statistics:**
        - **Mean:** 1.1219 W
        - **Standard Deviation:** 6.1530
    - **Visualization:** Histogram showing the distribution of sub-metering 1 values.



6. **Sub-Metering 2**
    - **Definition:** Energy sub-metering for laundry appliances.
    - **Type:** Quantitative
    - **Summary Statistics:**
        - **Mean:** 1.2985 W
        - **Standard Deviation:** 5.8220
    - **Visualization:** Histogram showing the distribution of sub-metering 2 values.



7. **Sub-Metering 3**
    - **Definition:** Energy sub-metering for climate control systems.
    - **Type:** Quantitative
    - **Summary Statistics:**
        - **Mean:** 6.4584 W
        - **Standard Deviation:** 8.4372
    - **Visualization:** Histogram showing the distribution of sub-metering 3 values.



Data Cleaning and Preprocessing:
Rows with missing values were removed to ensure consistency. Relevant data transformations, including resampling for time-series analysis and standardization, were performed to prepare the data for statistical analysis.

Visualization and Summary:
Various plots, including histograms, line plots, and time-series graphs, were generated to understand the distribution and trends across key variables like power consumption, voltage, and sub-metering categories. These visualizations are available in the Data Appendix for further reference.