

Data Collection and Preprocessing Phase

Date	June 2024
Team ID	Team: 739815
Project Title	Power Consumption Analysis For Households
Maximum Marks	2 Marks

Data Quality Report Template

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Source: Household Demographic Data

Data Source	Data Quality Issue	Severity	Resolution Plan
Household Demographic Data	Missing values in income level	Moderate	Develop a script to standardize timestamps to a single timezone and format. Validate adjustments with domain experts and stakeholders before implementation.

Data Source: Smart Meter Data

Data Source	Data Quality Issue	Severity	Resolution Plan
Smart Meter Data	Incomplete data for weekends	Moderate	Implement a weekend data collection strategy to ensure comprehensive coverage. Use interpolation techniques to estimate missing weekend data points.

Data Source: Household Survey Responses

Data Source	Data Quality Issue	Severity	Resolution Plan
Household Survey Responses	Inconsistent response formats	Low	Develop standardized response categories and mappings. Conduct data cleaning and transformation to unify formats across survey responses.