Data Collection and Preprocessing Phase

Date	31 June 2024
Team ID	739798
Project Title	Power Consumption Analysis for Households
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Report

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Report

Section	Description
Project Overview	The project aims to predict power consumed by a household based on the reading provided by them. Using a dataset with features such as Global active power, global reactive power, voltage, global intensity, sub metering values. The objective is to build a model that accurately predict the power consumption, which will be useful to manage the consumption.
Data Collection Plan	The dataset which contains a set of features through which power consumption can be calculated, is to be collected.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle, the popular platform for data science competitions and repositories. The provided datasets contain the collected

	information, variables such as global active power, global reactive power, global intensity, voltage, sub metering values for machine learning analysis.

Raw Data Sources Report

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle	The dataset contains the measurements of electric power consumption.	Household Electric Power Consumption (kaggle.com)	txt	126 MB	Public