Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	xxxxxx
Project Title	Predicting The Energy Output Of Wind Turbine Based On Weather Condition
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

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 Template

Section	Description
Project Overview	Developing model for predicting wind turbine power output using meteorological data was developed to increase energy efficiency and optimize turbine performance by predicting energy production based on current and historical weather patterns.
Data Collection Plan	Collecting real-time and historical weather data, including wind speed, temperature, and humidity, from **Kaggle datasets.** Gathering corresponding wind turbine energy output data to train and validate the predictive model.
Raw Data Sources Identified	1. Kaggle datasets 2. OpenWeather API

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	The dataset comprises comprehensive weather parameters alongside corresponding wind turbine energy outputs, sourced from meteorological stations, Kaggle, and turbine operators.	https://www.kaggle.com/datasets/berkerisen/wind-turbine-scada-dataset	CSV	3.5 MB	Public