# Data Collection and Preprocessing Phase

| | |
|---|---|
| Date | 25 july 2025 |
| Team ID | Satya Prakash Sharma |
| Project Title | **Predicting Plant Growth Stages With Environmental And Management** |
| 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 | This project aims to analyze environmental and management data to predict plant growth stages using Power BI. The key objective is to optimize plant health and productivity through data-driven insights. |
| Data Collection Plan | The dataset was sourced from Kaggle, a trusted platform for public datasets. It provides real-world agricultural data suitable for classification and analysis tasks. |
| Raw Data Sources Identified | **Kaggle Dataset**: Plant Growth Data - Classification — contains features like soil type, sunlight, temperature, humidity, fertilizer, and watering frequency. |

**Raw Data Sources Template**

| Source Name | Description | Location/URL | Format | Size | Access Permissions |
|---|---|---|---|---|---|
| Dataset 1 | Contains data related to plant growth stages based on environmental and management variables like sunlight, soil type, fertilizer, temperature, and humidity. | https://www.kaggle.com/datasets/gorororororo23/plant-growth-data-classification | excel | ~0.05 GB | Public |
| Dataset 2 | Includes a dataset on different plant types with features like water frequency, fertilizer | https://www.kaggle.com/datasets/manutrex78/plant-healt- | Excel | ~0.02 GB | public |

| | | | | | |
|---|---|---|---|---|---|
| | type, growth stage, and other environmental inputs. | dataset | | | |
| … | … | … | … | … | … |