Data Collection and Preprocessing Phase

Date	15 dec 2025
Team ID	SWUID20250236463
Project Title	Predicting Plant Growth Stages with Environmental and Management Data Using Power BI
Maximum Marks	2 Marks

Data Collection Plan

Section	Description
Project Overview	**This project focuses on analyzing plant growth patterns based on environmental and agricultural factors such as soil type, water frequency, temperature, humidity, fertilizer usage, and sunlight hours. The objective is to understand how these factors influence plant growth milestones and overall plant health using data visualization and analytical techniques in Power BI.**
Data Collection Plan	**The data is collected from an open-source agricultural dataset available online. The dataset includes structured tabular data related to plant growth conditions and outcomes. Data is downloaded in CSV format and imported directly into Power BI for further transformation and analysis using Power Query.**
Raw Data Sources Identified	**The raw dataset contains information such as plant identifiers, soil type, water frequency, fertilizer type, temperature, humidity, sunlight hours, and growth milestone indicators. This dataset serves as the primary source for all analysis and visualizations in the project.**

Raw Data Sources

Source Name	Description	Location/URL	Format	Size	Access Permissions
Plant Growth Dataset	Contains detailed records of plant growth conditions including soil type, water frequency, fertilizer type, temperature, humidity, sunlight hours, and growth milestones.	Public open-source dataset repository	CSV	~5 MB	Public
Environmental Factors Reference Data	Supplementary reference data used to categorize temperature and humidity into descriptive ranges for analysis.	Local system (derived during preprocessing)	Excel	~1 MB	Private (project use only)