Data Collection Plan – Predicting Plant Growth Stages with Environmental and Management Data

➢ Purpose of Data Collection

▪ The goal of data collection in this project is to gather reliable and relevant environmental and management data that influence the stages of plant growth. This dataset serves as the foundation for building insightful Power BI visualizations to understand patterns and predict growth behavior effectively.

➢ Data Source

▪ The dataset used in this project was sourced from Kaggle:
🔗 Plant Growth Data Classification – Kaggle

➢ Dataset Description

▪ The dataset contains labeled data related to different plant growth stages along with multiple environmental and management-related features, including:

- **Soil Type**
- **Water Frequency**
- **Temperature**
- **Humidity**
- **Sunlight Hours**
- **Fertilizer Type**
- **Plant Growth Stage (Label)**

➢ Relevance to the Project

▪ These features directly affect plant development and are ideal for analyzing correlations and building predictive visuals. The classification label for plant growth stage helps measure how inputs influence outcomes.

➢ Data Selection Criteria

▪ Only relevant columns related to environmental and management factors were

retained.

▪ Non-essential or redundant columns (like plant ID, timestamps) were excluded.

▪ The dataset was checked for completeness, formatting consistency, and suitability for visualization.

➢ Data Accessibility

▪ The dataset is publicly available under open access for academic and research use.

▪ It was downloaded in CSV format for easy integration into Power BI and Excel.