**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 6 August 2025 |
| Team ID | xxxxxx |
| Project Title | Predicting Plant Growth Stages Using Environmental & Management Data in Power BI |
| Maximum Marks | 10 Marks |

**Data Exploration and Preprocessing Template**

Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | Dataset contains ~10,000 records with environmental parameters (temperature, humidity, soil moisture) and management factors (irrigation, fertilizer). |
| Data Cleaning | Handle missing values, remove duplicates, and correct inconsistent entries for accurate analysis. |
| Data Transformation | Use Power Query for filtering, sorting, pivoting, and creating calculated columns such as average growth rate. |
| Data Type Conversion | Rectify incorrect data types (e.g., converting text-based numbers to numeric formats, dates to date format). |
| Column Splitting and Merging | Split date/time into separate fields; merge related attributes for analysis. |
| Data Modeling | Define relationships between environmental and management data tables, create measures in Power BI for trend analysis. |
| Save Processed Data | Store the cleaned and processed dataset in .xlsx and .csv format for use in prediction dashboards. |