

Project Documentation

Project Title: Predicting Plant Growth Stages with Environmental and Management Using Power BI.

Developed By: Dhanya K.

Tool Used: Microsoft Power BI.

1. Data Connection & Extraction

- **Step 1:** The raw dataset **plant_growth_data.csv** was identified as the source.
- **Step 2:** Opened Power BI Desktop -> Clicked **Get Data** ->Selected **Text/CSV**.
- **Step 3:** Loaded the file using **UTF-8 encoding** to ensure all characters were read correctly.

2. Data Transformation (ETL Process)

- **Step 1:** Opened **Power Query Editor** to clean the data.
- **Step 2: Created 'Growth_Milestone' Column:**
 - Used a Conditional Column to map the **Growth_Milestone** values.
 - Logic: If **Growth_Milestone** = 1, then "Success" else "Failure".
- **Step 3: Data Type Validation:**
 - Ensured **Temperature** and **Humidity** were set to Decimal Numbers.
 - Ensured **Soil_Type** and **Fertilizer_Type** were set to Text.
- **Step 4: Binning (Grouping):**
 - Created **Temp Bins** (Bin size 5) to categorize temperature ranges.
 - Created **Sunlight Bins** (Bin size 1) to smooth the line chart analysis.

3. Data Modeling & Calculations (DAX)

- **Step 1:** Created a specialized Measure Table to organize calculations.
- **Step 2:** Developed the Key Performance Indicator (KPI) measure:
 - **Measure Name:** Growth Success Rate %
 - **Formula:**
$$\text{DIVIDE}(\text{CALCULATE}(\text{COUNTROWS}(Data), Data[Growth_Milestone]=1), \text{COUNTROWS}(Data), 0)$$
- **Step 3:** Formatted the measure as a **Percentage (%)** with 1 decimal place.

4. Dashboard Visualization Development

The following visuals were created to address the problem statements:

- **KPI Card:** Displays the headline metric (49.7%) for immediate performance visibility.
- **Clustered Bar Chart:** Compares **Fertilizer_Type** against Success Rate. *Finding: Organic is #1.*
- **Scatter Plot:** Plots **Temperature** vs. **Humidity**. Reference lines were added at 20°C and 30°C to highlight the optimal growth window.
- **Matrix Visual:** Rows = **Soil_Type**, Columns = **Water_Frequency**. Conditional formatting was applied to highlight the "Sandy + Daily" high-performance cluster.
- **Slicers:** Added interactive filters for **Soil_Type** and **Fertilizer_Type** to allow end-users to drill down into the data.

5. Final Dashboard Layout

- **Theme:** Applied a "Green" colour theme to align with the agricultural context.
- **Colours Used:** Green, Red, Mustard Yellow, Blue.
- **Formatting:** Added a clear dashboard title, standardized font sizes, and enabled data labels on all charts for readability.
- **Insights Box:** Added a text summary of key findings directly on the dashboard page.

6. Conclusion

The project successfully transformed raw trial data into an interactive decision-support tool. The analysis proved that optimizing fertilizer inputs and watering schedules based on soil type can significantly improve crop yields.