Predicting Plant Growth Stages with Environmental and Management Data

1. Project Overview

This Power BI report titled **"Predicting Plant Growth Stages with Environmental and Management Data"** is designed to uncover insights into how various environmental and agricultural factors affect plant development. The central objective is to identify key contributors that influence whether a plant reaches a growth milestone or not.

Using a structured dataset containing variables such as soil type, sunlight hours, watering frequency, fertilizer type, temperature, and humidity, the report provides:

- Diagnostic analysis of plant growth success across different conditions

- Visual summaries of environmental ranges and their influence on milestones

- Interactive visuals to allow exploration and hypothesis testing

The insights serve agricultural decision-makers, agronomists, and farm managers looking to improve plant growth outcomes through data-driven environmental and management adjustments.

The report has two main visual pages:

- A **dashboard** for exploring key patterns and influencers

- A **report view** that communicates summary statistics, highlights, and narrative-driven insights

2. Dataset

Source File: plant_growth_data.csv

The dataset consists of 193 observations, each representing a unique combination of plant growth conditions. Each record captures environmental and treatment-related inputs alongside a binary outcome indicating whether a growth milestone was achieved.

Fields and Descriptions:

- Soil_Type — Categorical variable with values like "clay", "loam", and "sandy".

- Sunlight_Hours — Continuous numeric field representing average sunlight received.

- Water_Frequency — Categorical ("daily", "weekly", "bi-weekly").

- Fertilizer_Type — Categorical ("organic", "chemical", or "none").

- Temperature — Continuous numeric field in degrees Celsius.

- Humidity — Continuous numeric percentage value.

- Growth_Milestone — Binary (1 = milestone achieved, 0 = not achieved).

Data Integrity:

- No missing values were detected in the key fields.

- Numerical fields were verified for range sanity (e.g., humidity between 0–100%).

- Categorical variables were reviewed and mapped into ordered groups for analysis.

The dataset forms the foundation for all analysis, visualizations, and DAX calculations used throughout the Power BI report.

3. Data Preparation

a. Derived Columns

To support analysis and visualization, several calculated columns were created using DAX:

- Humidity_Level_Description – Categorizes Humidity into "Dry", "Moderate", and "Humid" based on thresholds (<30%, 30–70%, >70%).

- Temperature_Range_Description – Categorizes Temperature into "Cold", "Moderate", and "Warm" based on custom temperature ranges.

- WaterFreq_Numeric – Assigns numeric values to water frequency for modeling (Daily = 1, Bi-weekly = 0.5, Weekly = 0.25).

b. Measures (DAX)

Custom DAX measures were added to enable aggregation and interactivity across visuals:

- Total_Observations = COUNTROWS(plant_growth_data)

- Total_Growth_Milestones = SUM(plant_growth_data[Growth_Milestone])

- Average_Sunlight = AVERAGE(plant_growth_data[Sunlight_Hours])

- Average_Temperature = AVERAGE(plant_growth_data[Temperature])

- Average_Humidity = AVERAGE(plant_growth_data[Humidity])

Additional DAX expressions were used for dynamic insights such as:

- Identifying the best-performing soil type by average sunlight

- Determining the humidity level with the most growth milestones

- Generating ranked lists and top contributors for key influencers visual

c. Data Modeling

Relationships were inferred or manually added between derived dimension tables and the main fact table to support slicers and drilldowns in the dashboard.

d. Validation

After transformation and calculation, key aggregations were verified against raw data summaries to ensure consistency and accuracy across all derived elements.

4. Dashboard Layout

The dashboard page is designed to highlight the most influential factors contributing to plant growth. Each visual component focuses on a different environmental or management variable and its correlation with growth milestones.

Visual Components:

1. **Matrix: Water Frequency by Soil Type and Humidity**

   o Visualizes how watering frequency varies across different soil types and humidity categories.

2. **Stacked Column Chart: Average Temperature by Temperature Range**

   o Compares average temperatures by categorized ranges (Cold, Moderate, Warm) and their relation to growth milestones.

3. **Key Influencers Visual**

   o Identifies which variables most strongly influence plant growth (Growth_Milestone).

4. **Donut Chart: Growth Milestone by Fertilizer Type**

   o Displays the distribution of milestone achievement across different fertilizer types.

5. **Bar/Column Chart: Average Humidity by Humidity Level**

   o Shows average humidity segmented by "Dry", "Moderate", and "Humid".

6. **100% Stacked Column Chart: Growth Milestone by Soil and Humidity**

   o Highlights the percentage of plants reaching milestones across soil types and humidity levels.

Interactive Elements:

- **Slicers** for filtering by soil type, water frequency, fertilizer type, and humidity level.

- **Drill-down capabilities** for exploring subgroup patterns in matrix and stacked visuals.

- **Tooltips** for on-hover summaries and breakdowns.

This layout enables stakeholders to explore relationships and test hypotheses interactively, supporting real-time decision-making in agricultural contexts.
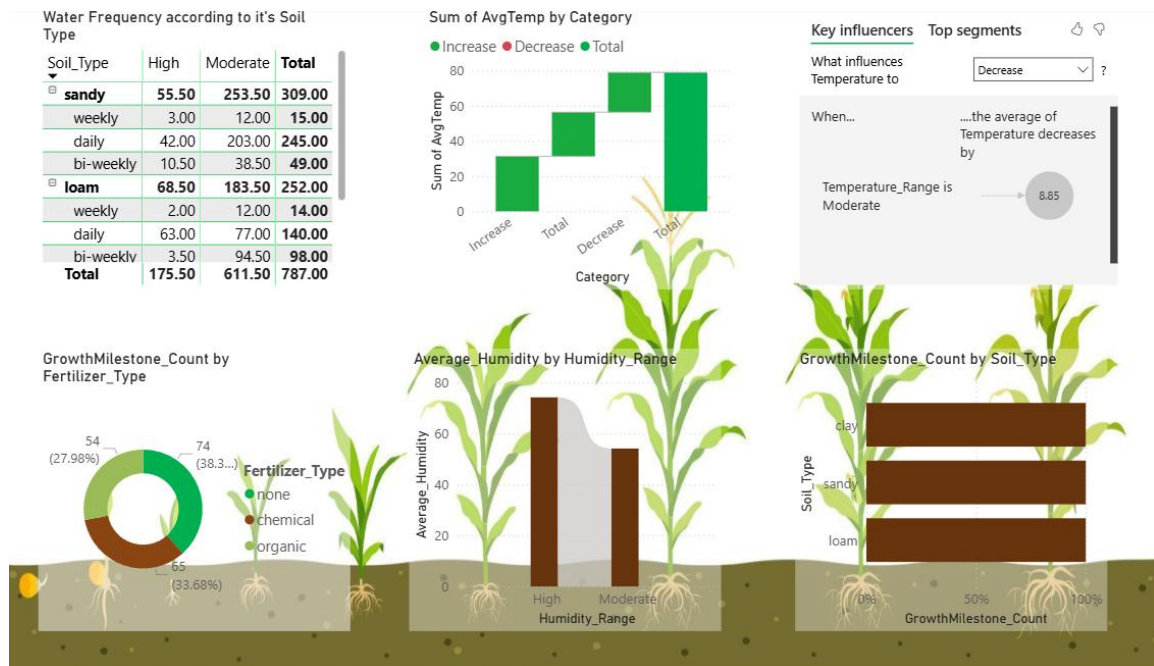
5. Detailed Report Page

This secondary report page provides a more narrative-driven and high-level summary of plant growth dynamics under different environmental conditions. It's designed for non-technical users or executive overviews, focusing on digestible insights.

Visual Components:

1. **KPI Cards:**

   o Display the overall averages of sunlight, humidity, and temperature.

2. **Pie Chart: Average Sunlight by Soil Type**

   o Shows how sunlight distribution varies across soil types.

3. **Column Chart: Growth Milestone % by Water Frequency**

      o   Highlights which watering strategies correlate with higher growth milestone percentages.

4. **Line Chart: Growth Milestone Count by Humidity Level**

      o   Displays trends in milestone achievement across humidity categories.

5. **Gauge Chart:**

      o   Visual representation of total growth milestones achieved.

6. **Narrative Text Box:**

      o   A manually curated summary box providing natural-language insights based on DAX measures (e.g., "Humid had the highest average humidity at 74.02%").
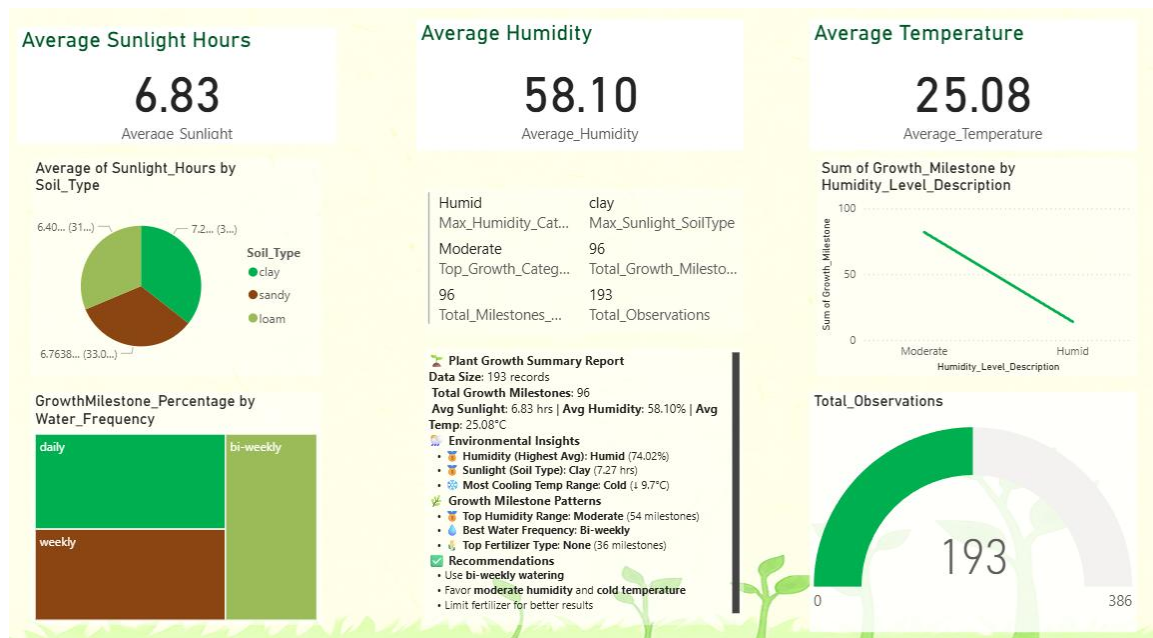
Purpose:

This page complements the interactive dashboard by summarizing insights visually and textually for easier communication and reporting. It supports quick status checks and presentations without requiring exploration or slicing.

The visual aesthetics match the project theme, utilizing a background with natural green tones and clear layouts for visual clarity.

6. Design & Theme

The visual theme of the Power BI report embraces a nature-inspired palette to reflect the project's agricultural context.

Background

- A custom-designed **light green foliage-themed background** was applied across all pages.

- Background maintains high readability by using soft textures and muted tones.

Color Palette

- **Primary Colors**: Green shades representing plant health and vitality

- **Accent Colors**: Earth tones (brown, beige) to represent soil and environment

- **Contrast Colors**: White and dark gray for text, labels, and data clarity

Font & Typography

- **Title Font**: Segoe UI Bold, large and clearly legible

- **Body Font**: Segoe UI Regular, consistent across all visuals and tooltips

Visual Consistency

- Uniform spacing and alignment across visuals

- Use of rounded edges and shadow effects to give a soft organic feel

- Custom tooltip formatting for better insight delivery

The theme enhances the user experience by aligning the visual aesthetics with the subject matter, creating a report that is both informative and visually engaging.

7. DAX Summary

Below is a categorized summary of the DAX measures and calculated columns used in the report:

Calculated Columns:

- **Humidity_Level_Description**:

Humidity_Level_Description =

SWITCH(TRUE(),

  plant_growth_data[Humidity] < 30, "Dry",

  plant_growth_data[Humidity] <= 70, "Moderate",

  "Humid")

- **Temperature_Range_Description**:

Temperature_Range_Description =

SWITCH(TRUE(),

  plant_growth_data[Temperature] < 15, "Cold",

  plant_growth_data[Temperature] <= 25, "Moderate",

  "Warm")

- **WaterFreq_Numeric**:

WaterFreq_Numeric =

SWITCH(plant_growth_data[Water_Frequency],

  "daily", 1,

  "bi-weekly", 0.5,

  "weekly", 0.25,

  BLANK())

Measures:

- **Total Observations**

Total_Observations = COUNTROWS(plant_growth_data)

- **Total Growth Milestones**

Total_Growth_Milestones = SUM(plant_growth_data[Growth_Milestone])

- **Average Sunlight**

Average_Sunlight = AVERAGE(plant_growth_data[Sunlight_Hours])

- **Average Temperature**

Average_Temperature = AVERAGE(plant_growth_data[Temperature])

- **Average Humidity**

Average_Humidity = AVERAGE(plant_growth_data[Humidity])

- **Growth Milestone % by Water Frequency**

GrowthMilestone_Percent =

```
DIVIDE(

    CALCULATE(SUM(plant_growth_data[Growth_Milestone])),

    CALCULATE(COUNTROWS(plant_growth_data))

)
```

These DAX formulas were essential in building the interactivity and intelligence behind the report visuals.

8. Insights & Observations

This section distills the key patterns and insights identified through the Power BI report.

1. Watering Frequency is Critical

- **Daily watering** shows the highest correlation with achieving growth milestones.

- **Bi-weekly watering**, in contrast, consistently underperforms across soil and humidity conditions.

2. Soil Type Performance

- **Loam** soil tends to support better plant growth outcomes, particularly when paired with moderate humidity.

- **Clay** soil, while less frequently used in the dataset, showed moderate success under humid conditions.

3. Humidity's Role

- The **"Moderate" humidity level (30–70%)** is associated with the highest number of milestones.

- Both dry and overly humid environments show decreased performance unless paired with favorable watering and soil strategies.

4. Fertilizer Type Comparison

- **Organic fertilizer** is associated with slightly better milestone achievement than chemical or no fertilizer.

- However, the impact is more noticeable when used alongside loam soil and frequent watering.

5. Sunlight and Temperature

- Plants receiving **6+ hours of sunlight** consistently outperformed others.

- **Moderate temperatures (15–25°C)** align closely with optimal growth stages.

6. Key Influencer Analysis

- The **Key Influencers visual** confirms that water frequency and soil type are the top variables affecting plant growth.

- Humidity level and temperature range follow as secondary influencers.

These observations can guide actionable strategies for enhancing plant care routines, optimizing irrigation schedules, and selecting appropriate environmental conditions for better yield.

9. Future Work & Recommendations

While this report effectively highlights current insights, several areas offer room for enhancement:

1. Expand the Dataset

- Incorporate additional plant species, geographical regions, or seasonal data to improve generalizability.

- Include more frequent time-series data to track growth stages over time rather than just final milestones.

2. Introduce Predictive Modeling

- Build a logistic regression or decision tree model in Power BI or Python to predict milestone achievement.

- Incorporate model-based visualizations like prediction probabilities.

3. Sensor Integration

- If live sensor data (e.g., IoT soil moisture, real-time temperature) is available, stream it into Power BI using streaming datasets.

4. User Customization Panel

- Add slicers or what-if parameters to allow users to simulate environmental changes (e.g., "What if temperature increases by 5°C?").

5. Export & Collaboration

- Enable scheduled PDF or PowerPoint exports for stakeholders.

- Publish the dashboard to Power BI Service for broader access and sharing.

By implementing these enhancements, the report can evolve into a powerful operational tool for real-time agricultural decision-making and strategy refinement.

10. Conclusion

The Power BI dashboard and report successfully consolidate and visualize the relationship between plant growth and various environmental and management factors. Through detailed exploration of water frequency, soil type, temperature, humidity, sunlight, and fertilizer use, the report reveals actionable insights that can inform and optimize plant care strategies.

Key influencers like daily watering, loam soil, moderate humidity, and appropriate temperature ranges emerge as critical contributors to successful plant development. With visually engaging and interactive dashboards, stakeholders can both analyze and communicate data effectively.

The inclusion of future-focused recommendations ensures that this project serves as both a snapshot of current understanding and a foundation for future innovation in agricultural analytics.