GREENROOTS ALLIANCE

AGRICULTURAL DATASET ANALYSIS REPORT

Prepared by: Janada Kefas

Data Cleaning Summary

As part of my analysis for the Greenroots Alliance, I worked on a raw agricultural dataset covering farming activities across 35 states in Nigeria. The dataset included detailed records on costs (seeds, labor, fertilizer), yield outputs, and environmental factors affecting each farm.

To ensure accuracy and consistency, I carried out the following data cleaning steps:

- **Removed Duplicates:** I eliminated all duplicate entries to maintain data integrity.

- **Handled Missing Values:** Some rows lacked critical location data (State and LGA), so I removed them since they couldn't be matched to any region. For other columns, I standardized entries like "N/A," "NULL," and "MISSING VALUE" to **"Not Stated."**

- **Corrected Data Errors:**

    - I found a misnamed state labeled "BOLAJI." Since this row had a large amount of valid data, I replaced it with **FCT** instead of deleting it.

    - I fixed spelling errors and standardized inconsistent values such as recategorizing **"none"** under Fertilizer Used as **"Others."**

- **Standardized Date Formats:** All dates were reformatted to **DD-MM-YYYY** to ensure consistency across time-based analyses.

- **Cleaned Column Data Types:** I adjusted column formats to reflect correct data types (dates, numbers, text, etc.).

Additionally, I created three new columns to enhance insights:

- **Profit** – Calculated by subtracting total investment from revenue.

- **Experience Groups** – I grouped farmers based on years of experience to help assess performance by expertise.

- **Average Yield per Hectare** – Yield divided by farm size, offering a clearer measure of productivity.

- **Crop Duration** – Calculated by subtracting planting dates from harvest dates to understand growth cycles.

Key Findings and Observations

Top 5 High-Yield Crops: Sweet potato, sesame, rubber, cashew, and banana stood out with the highest average yields per hectare indicating areas of high productivity potential.

Top 5 States by Investment and Revenue: Bayelsa, Borno, FCT, Gombe, and Kebbi emerged as the top-performing states in terms of both investment and revenue, showing where agricultural growth can be further scaled.

Farmer Experience + Cooperatives = Higher Performance: I observed that farmers with **over 16 years of experience** who are also **members of cooperatives** tend to generate higher revenue and receive greater investments. This points to the value of supporting farmer education and group-based models.