**AgriSmart AI: Climate Data Cleaning and Analysis Report**

**Student Name:** WAKINYA ESAU JOHN  
**Registration Number:** 2300718448  
**Course:** BSSE2301 - Software Engineering

### 1. Objective

To clean and analyze the dataset climate\_action\_data.csv containing 913 records and 10 variables related to soil and environmental data. The goal is to prepare the data for AI-driven agricultural decisions.

### Data Cleaning Summary

### ****Initial Issues Identified:****

Missing values in multiple columns.

Erroneous entries marked as 'error'.

Incorrect data types (string instead of numeric).

Duplicate records.

**Cleaning Actions Taken:**

Replaced 'error' values with NaN.

Converted all columns to numeric where applicable.

Dropped duplicate records.

Dropped rows with more than 3 missing values.

Filled remaining missing numeric values with column medians.

**Final Dataset Shape:** 899 records, 10 columns.

**Variables influencing Fertilizer recommendations**

**Strongest influencers**

Fertilizer\_Recommended(kg/ha) 1.000000

Soil\_pH 0.083676

Humidity(%) 0.015300

Soil\_Moisture(%) -0.002757

Temperature(C) -0.007206

***Crop with the highest average soil moisture*** is **Wheat 47.195731**

**Irrigation Adjustments for Crops >30°C:**

Beans Maintain current

Lettuce Maintain current

Maize Maintain current

Tomatoes Maintain current

Wheat Maintain current

### 6. Conclusion

The cleaned dataset is ready for AI modeling and provides accurate inputs for AgriSmart AI's decision-making systems. Key relationships between environmental factors and fertilizer recommendations were identified, and actionable advice for irrigation optimization under high temperature conditions was proposed.