

## Report on Crop Yield and Data Analysis

### The average yields per hectare in Metric Tonnes

The average yield per hectare is **1.0434495204275063 Tonnes**. This is found by computing the mean of the Yield per Hectare column

### Explanation of the Erroneous Data

Some of the things showing erroneous or inconsistent data include the following

- Crazy box dimensions where we had outliers in the *'box1\_crop\_stands\_per\_square\_meter'* and the *'box2\_crop\_stands\_per\_square\_meter'* columns. This raises an alarm as to presence of inconsistencies in data entry
- Upon comparing the Enumerator comments about harvests against the actual comments documented during harvesting there is a disparity showing that there was erroneous data entered
- There are some instances where the dry weight was more than the wet weight which is true in practice, this suggests errors during data entry
- There are other situations where we do not have box dimensions but we end up with some yield data captured, this is erroneous since in order to capture yields we require box dimensions to be present
- In some cases, we have Zero wet weight yield but when recording the dry weight yield for the corresponding wet yields we find that the dry yield is **greater** than the wet yield. This is clearly an error
- I found that in some cases there are inconsistencies between the box weight recorded by the enumerators and the actual box weights (confirmed box weights)

## **Major Factors Affecting Crops Per District**

The major factors affecting the crops in different districts as shown by the analysis include the following:

- **Katsina,**
  - Poor Germination
  - Late planting
  - Pests and Diseases
  - Drought
  - Locust infestation
  
- **Zamfara,**
  - Drought
  - Locust infestation
  - Poor Germination
  
- **Niger**
  - Late planting
  - Pests and diseases
  - Drought
  - Poor germination
  - Weeds
  
- **Kebbi,**
  - Undocumented
  
- **Kaduna,**
  - Late planting
  - Poor germination
  - Drought
  - Animal/cattle encroachment

- **Sokoto,**

- Late planting
- Poor germination
- Weeds

- **Kano,**

- Poor Germination
- Pests and Diseases
- Drought
- Animal Encroachment
- Poor Germination
- Locust Infestation

- **Borno,**

- Late planting
- Drought
- Poor Germination

- **Bauchi,**

- poor germination
- other pests and diseases
- late planting
- weeds
- droughts

- **Yobe,**

- Late planting
- Animal/cattle encroachment
- Poor germination
- Weeds

- **Jigawa,**
  - Late planting
  - Poor germination
  
- **Adamawa,**
  - Late planting
  - Poor germination
  - drought
  - weeds
  - Locust infestation
  
- **Gombe,**
  - Undocumented
  
- **Taraba,**
  - Poor germination
  - Late planting
  - Pests and diseases
  - Drought
  - Locust infestation
  
- **Abia**
  - Undocumented

### **Dealing Enumerators with inconsistent or Wrong entries**

Based on the analysis, a few enumerators' data submissions raised suspicion to the level potentially warranting further investigation.

These enumerators don't necessarily need to be fired, instead, it would be a good idea to speak to understand areas where they might have some gaps in and provide them with the necessary training and resources to help them fill these gaps, they should be continuously monitored to ensure improvement. If there is no improvement after this then the actions can now be escalated

### **Assumption made**

For the missing entries, I made an assumption that for some they were undocumented and I therefore went ahead to impute them with 'Unknown'