Team no. **61**, Team Name: **Beta Squad**

The raw datasets were taken from a public Github repository linked here. The

aim of the data preprocessing is to do the analysis and extract valuable

information for the further implementation of the AI model.

Columns:

- **Nitrogen:** ratio of content in soil
- **Phosphorous:** ratio of content in soil
- **Potassium:** ratio of content in soil
- **Temperature (Degree Celsius)**
- **Humidity:** relative humidity in percentage
- **Ph:** Ph value of soil
- **Rainfall (mm)**
- **label(Crop Names)**

Proposed Solution:

As we already discussed in the previous report, agricultural knowledge is supposed to be hierarchical, many novice farmers with improper knowledge and guidance are losing interest in farming and thus the agricultural quality is decreasing in India. We are planning to tackle this problem with the help of AI and Machine learning. The user just have to feed the application with geographical conditions and the model will suggest the best crop to be planted and the best course of action to be opted for better result. The UI will be very user friendly, and someone with even less tech knowledge will be able to enjoy the benefits.

Procedure followed for Data Processing:

After researching a lot on the internet, we got our hands on unprocessed and clean data, the first data set is about crop data and the second one is regarding the Fertilizer data. These datasets met all the requirements of the solution we intended to present.
We had two separate raw datasets, we implemented a Function for lowering the cases, and made some changes in the fertilizer dataset to match the names of

the crops. Using frame.appends function we made an array consisting of the intersecting crop names from the two datasets. With the help of the Random library in Python we deviated the N,P,K data keeping it close to the real life values. And finally merged the two datasets with the help of the array of crop names containing unique values.
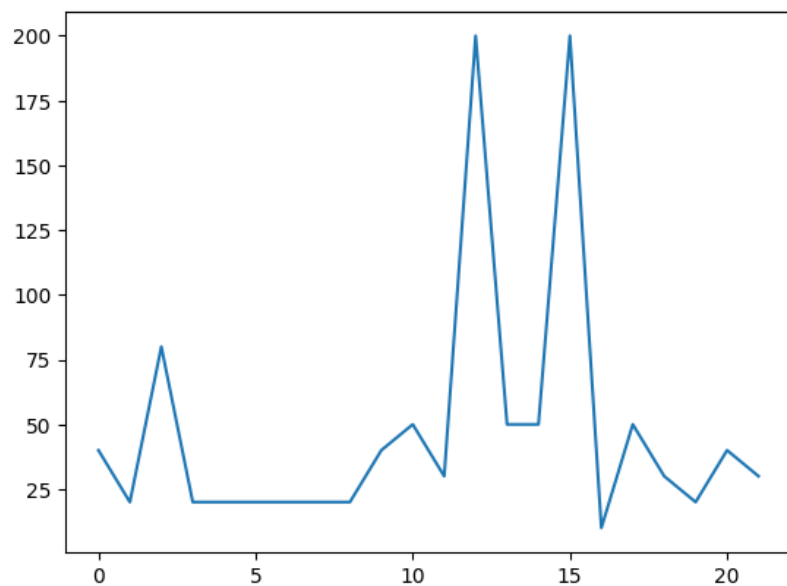
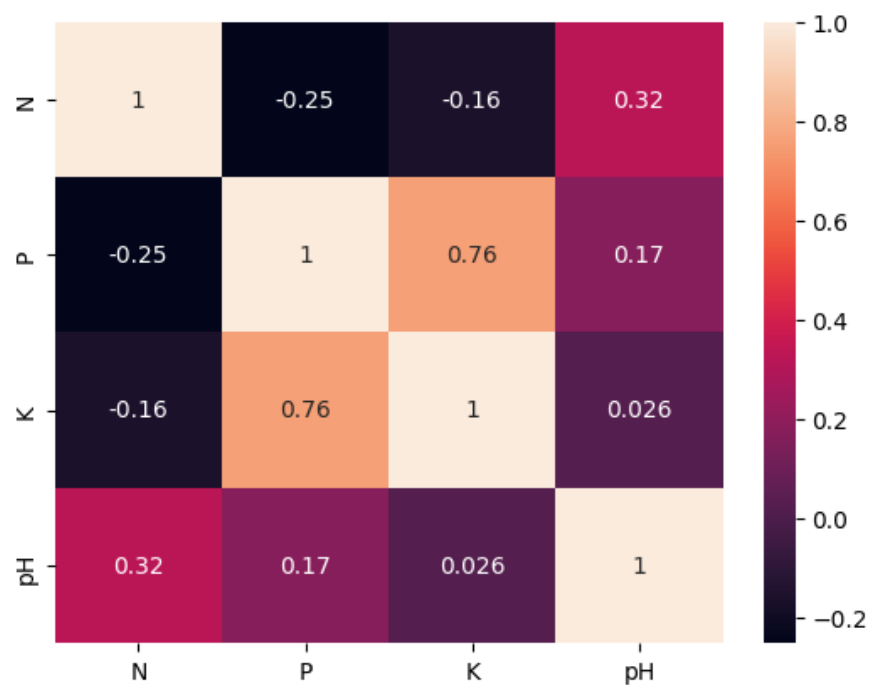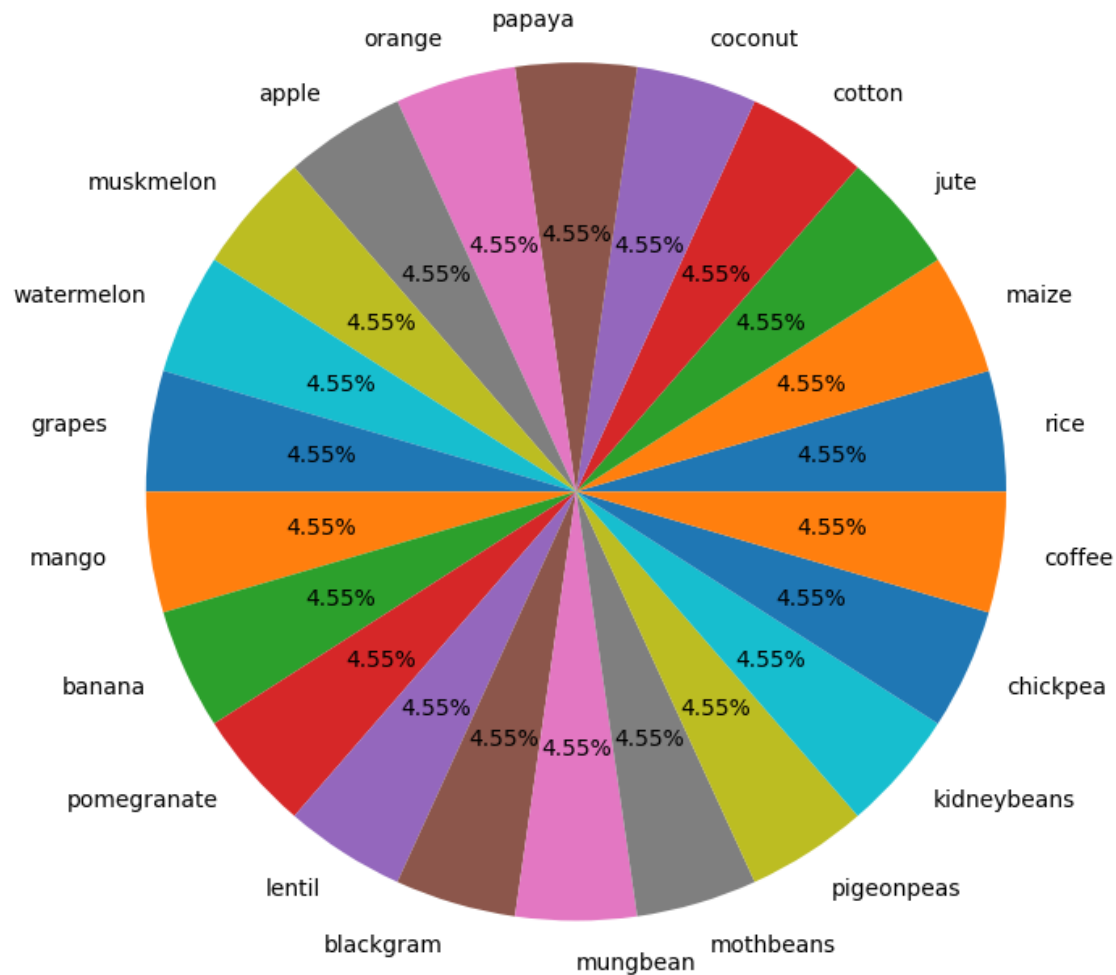## <u>Graphs from Data Visualization:</u>

Nitrogen:
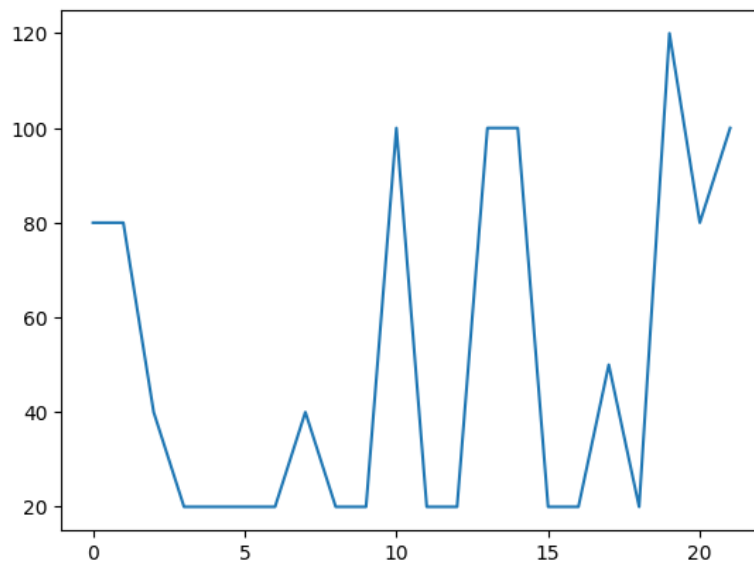
Phosphorous:
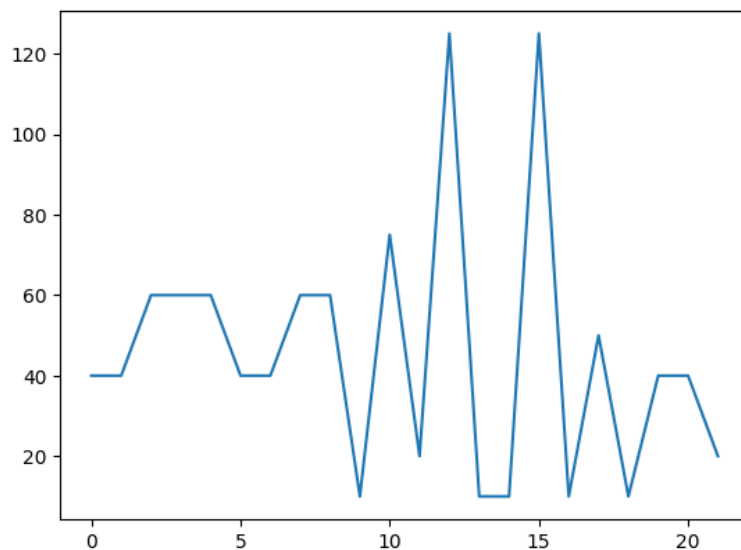
Potassium:

Heatmap of Merged Fertilizer dataset:

Pie chart showing equal distribution of crop data:

Github Repository link for Jupyter notebook:

https://github.com/isosceles45/61_BetaSquad_2/blob/main/61_DataProcessing.ipynb

Tech Stack :

Python Programming Language, Matplotlib library, Pandas library, Seaborn library, Jupyter Notebook, Microsoft Excel 2013.