ANALYSIS OF TRENDS IN AGRICULTURE AND DEVELOPMENTS IN CONSUMPTION OF PLANT-DERIVED FOODS.

1. **Business Understanding & Problem Discovery**

With the ongoing pandemic and global warming conditions, more awareness is raising among us regarding vegetarian, vegan foods, and their benefits. People are moving more towards sustainable living and becoming more eco-friendly. This is changing the way we care for the land we live and the surrounding resources. The problem we now have is the management and regaining our land fertility, to increase farming. Farmers took a back step and moved to other industries because of no business in the agriculture and low harvest. Due to the increasing awareness of the vegetation, agriculture is slowly picking up. AI is being used to develop machines that would help in improving the farming and methods to increase the yield of the crop. With the consumption and harvest statistics, the trends can be determined and aid the farmers in producing certain crops in the respective consuming areas.

DATA REQUIREMENTS:

Dataset-1: <https://www.kaggle.com/datafiniti/vegetarian-vegan-restaurants>

This dataset has cuisines and restaurants that serve vegan or vegetarian foods. These places are sorted by the region and types of food most sold and the customer base.

Size: 24MB, CSV file

Dataset-2: <https://www.kaggle.com/unitednations/global-food-agriculture-statistics>

This dataset has the data of crops produced and harvested year wise and the yield statistics

Size: 7GB, 6 CSV files.

The test process would be, producing the data of foods and cuisines based on the majority sales of in an area. The training will be done using the dataset-1 and test using either user inputs or producing a data sheet.

1. **Data Acquisition & Understanding**

**EDA report:**

dataset 1



There are large number of missing values in some features. We can drop those features depending on their importance.

Below are the few of the important features required to categorize the desired output.

Graphical user interface, text, application

Description automatically generated

These features would help to categorizing the restaurant areas and respective foods for the output.

1. **Modelling**

After checking the dataset with possible classifiers, below were the scores of the classifiers:

Name Score

0 KNN 0.4928

1 DT 0.5628

2 RF 0.5656

Logistic regression 0.527

A screenshot of a computer

Description automatically generated

1. **Delivery and acceptance**

A simple input would be asked for the user.

Country?

Province?

Postal code?

Then we display cuisines and their addresses, food categories and Minimum and maximum prices.

Graphical user interface, text

Description automatically generated