SDG-2-Zero-Hunger

Team - Hackatoons

Agriculture has been the backbone of Telangana's economy for a long time. There are multiple ways to increase and improve the crop yield and the quality of the crops. In this case study, we have directed to predict the price of the crop, for which we used real time data of crops Groundnut, Maize, and Bengal Gram. We have analyzed the type of crop, current price, and its location of harvest in Telangana. By analyzing these features we could predict the price of the crop.

METHODOLOGY:

The following workflow has been followed for developing our model:

1.Data Preprocessing: Data Preprocessing is a data mining technique used to transform the raw data into useful and efficient format. The data here goes through 2 stages:Data Cleaning and Data Transformation.

- Data Cleaning: It is very important for data to be error free and free of unwanted data. **- Data Transformation:** Data Transformation is transformation of the datasets mathematically.

2.Exploratory data analysis: is an approach to understand the datasets more keenly by the means of visual elements like scatter plots, bar plots, etc. This allows us to identify the trends in the data more accurately and to perform analysis accordingly.

3. Data Visualization: This has been implemented using Tableau. The adjacent chart is an example of how we used Data Visualization concept. The Chart is called Packed Bubble Chart. Here size and optionally colour of bubbles are used to visualize the data for which we need 1 or 2 dimensions against 1 or 2 measures to build this chart. **4.XGBoost Model:** XGBoost is a decision-tree-based ensemble Machine Learning algorithm that uses a gradient boosting framework. In prediction problems involving unstructured data (images, text, etc.) artificial neural networks tend to outperform all other algorithms or frameworks.

5. K Fold Cross Validation: k-fold cross validation is a procedure used to estimate the skill of the model on new data.
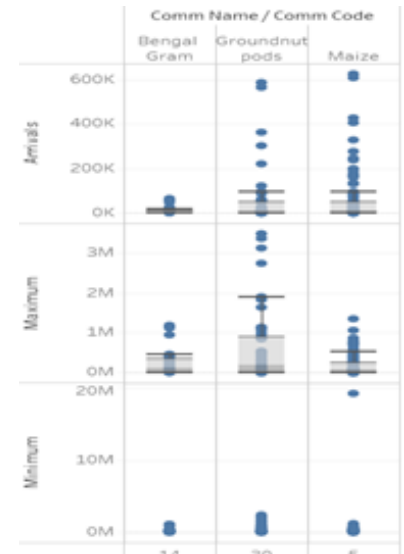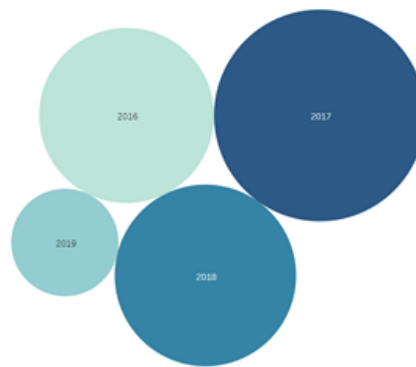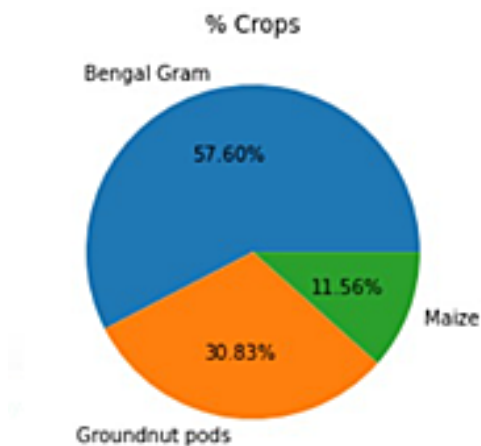
RESULT:

In order to predict the price, we have studied the type of crop, variety of crop and the current price of the crop.

SDG-2-Zero-Hunger

Team - Hackatoons

The statistics are as below:

Conclusion:

In this study, certain Data Analytics techniques and Machine Learning model were adopted to estimate crop price analysis with real-time data. Linear Regression is employed for discovering important information from the agricultural datasets. Using this real-time data and Machine Learning model we could predict the price of the crop at different regions of Telangana for different volumes. XG Boost Model is giving the highest accuracy of 99.9960 %, hence we will use XGBoost Model for prediction and is reliable.