MAJOR PROJECT

AGROINSIGHT

CROP PRODUCTION DYNAMICS IN ANDHRA PRADESH

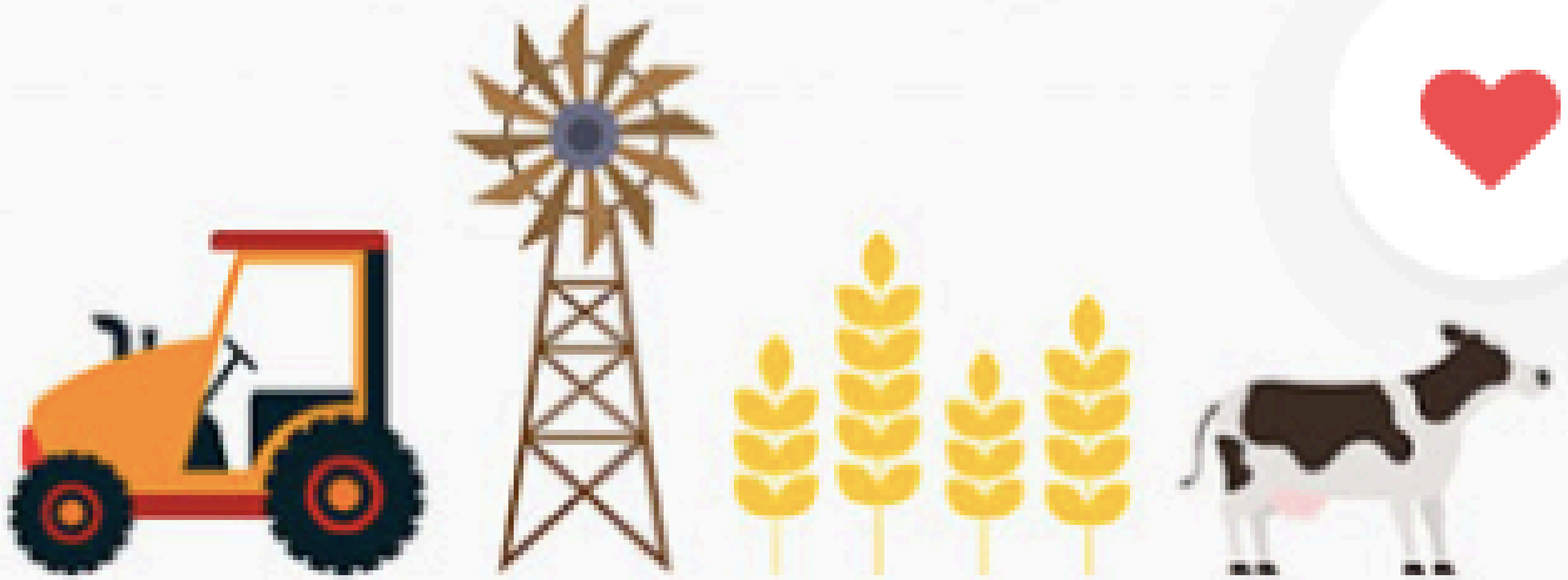
Team:
Bhavya Gupta
Nishtha Kohli
Vanshika Kathuria
Yash Thakran

Mentor
Ms Meenakshi Sihag

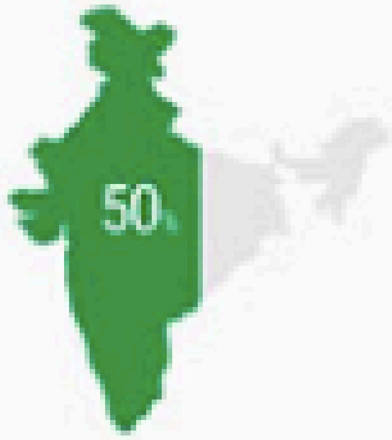
Agenda

INDIA

AS AGRICULTURAL ECONOMY

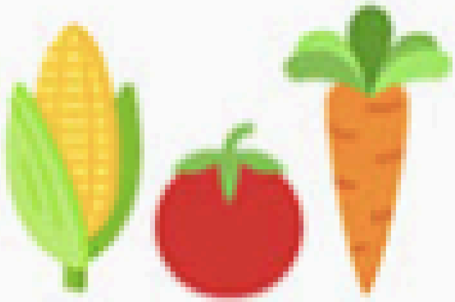


50% India is primarily an agriculture-based economy, employing **nearly** of the country's workforce



17% Agriculture contributes **nearly** to India's gross domestic product (GDP)

7% India contributes **at least** of the world's total agricultural output



86% of the country's farmlands are cultivated by small and marginal farmers

Why we chose Andhra Pradesh

Choosing Andhra Pradesh as the focus for our crop recommendation and yield prediction project was influenced by several factors. Here are some reasons why we have chosen Andhra Pradesh:

Agricultural Significance

Andhra Pradesh is one of the significant agricultural states in India, known for its diverse range of crops and farming practices. The state's economy is heavily dependent on agriculture, making it an ideal location for a project that aims to enhance agricultural practices.

Variability in Climate and Geography

Andhra Pradesh exhibits diverse climatic conditions and geographical features, including coastal regions, plains, and hilly areas. This diversity can impact crop growth and yield, making it an interesting region for study.

Back to Agenda Page

Data Availability

Access to historical and relevant agricultural data for Andhra Pradesh may have played a role in the decision. The availability of comprehensive datasets is crucial for training and validating machine learning models.

Government Initiatives

The state's government might have shown interest or initiated programs related to agriculture, creating an opportunity for collaboration or support for your project.

Societal Impact

Addressing agricultural challenges in Andhra Pradesh could have a significant impact on the livelihoods of farmers and contribute to overall food security in the region.

Back to Agenda Page

Functionality

Data Collection	Gathering data from various sources, including weather patterns, soil quality, historical crop yields.
Data Preprocessing	Cleaning and organizing the collected data to ensure accuracy and consistency, including handling missing values and outliers.
Predictive Modeling	Employing machine learning algorithms to create predictive models that can forecast crop yields based on historical data and current conditions.

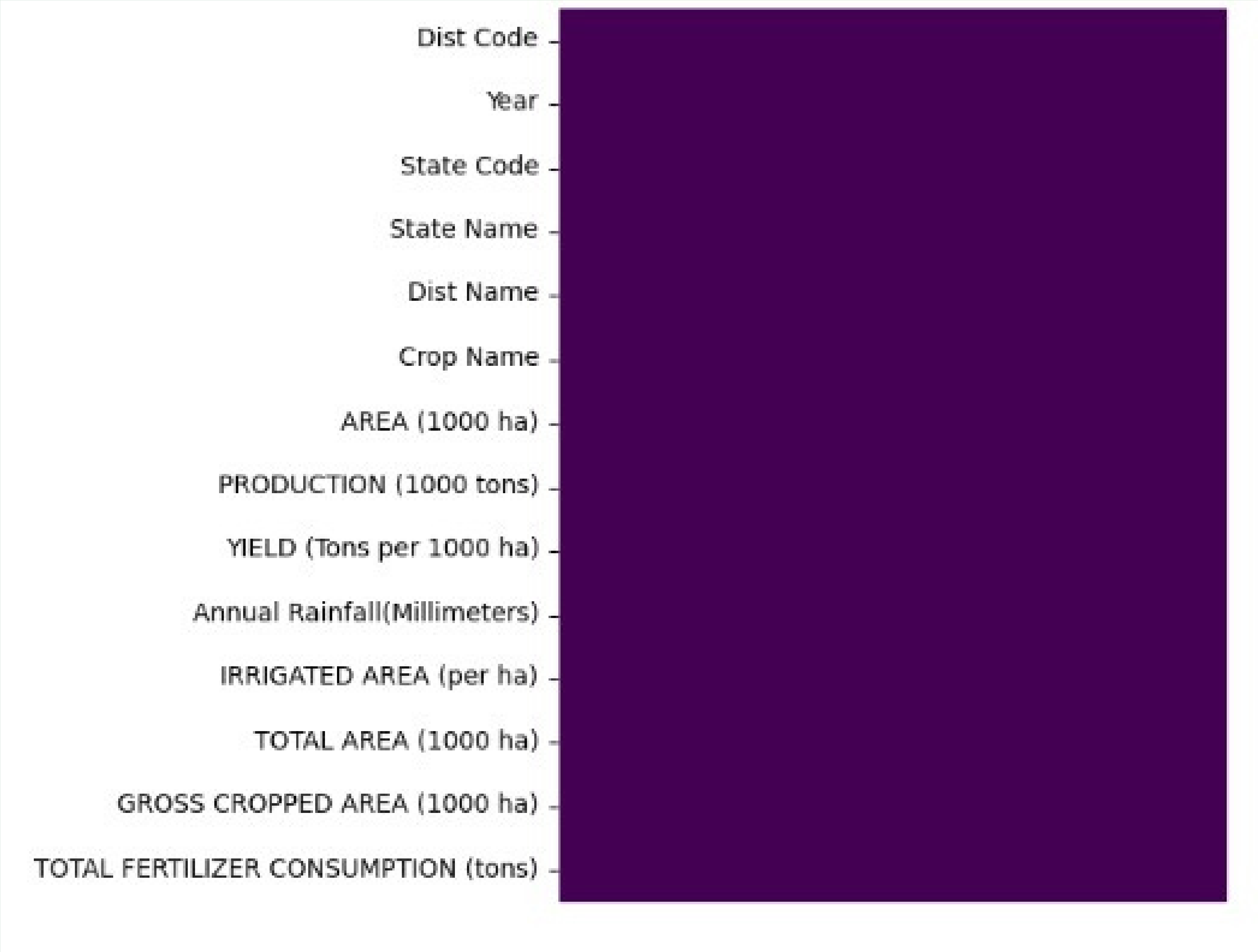
Back to Agenda Page

Yield Forecasting	Utilizing the predictive models to estimate potential crop yields for different crops and regions, taking into account factors such as weather conditions and agricultural practices.
Resource Optimization	Analyzing resource utilization, such as water, fertilizers, and pesticides, to identify areas where efficiency can be improved and resources can be used more sustainably.
Fuzzy Implementation	In agricultural data prediction, dealing with uncertainty is particularly important due to the inherent variability in factors such as weather, soil conditions, and crop growth patterns.
Data Visualization	Presenting the analysis results in a visually accessible format, such as charts or maps, to facilitate better understanding and decision-making for farmers and policymakers.

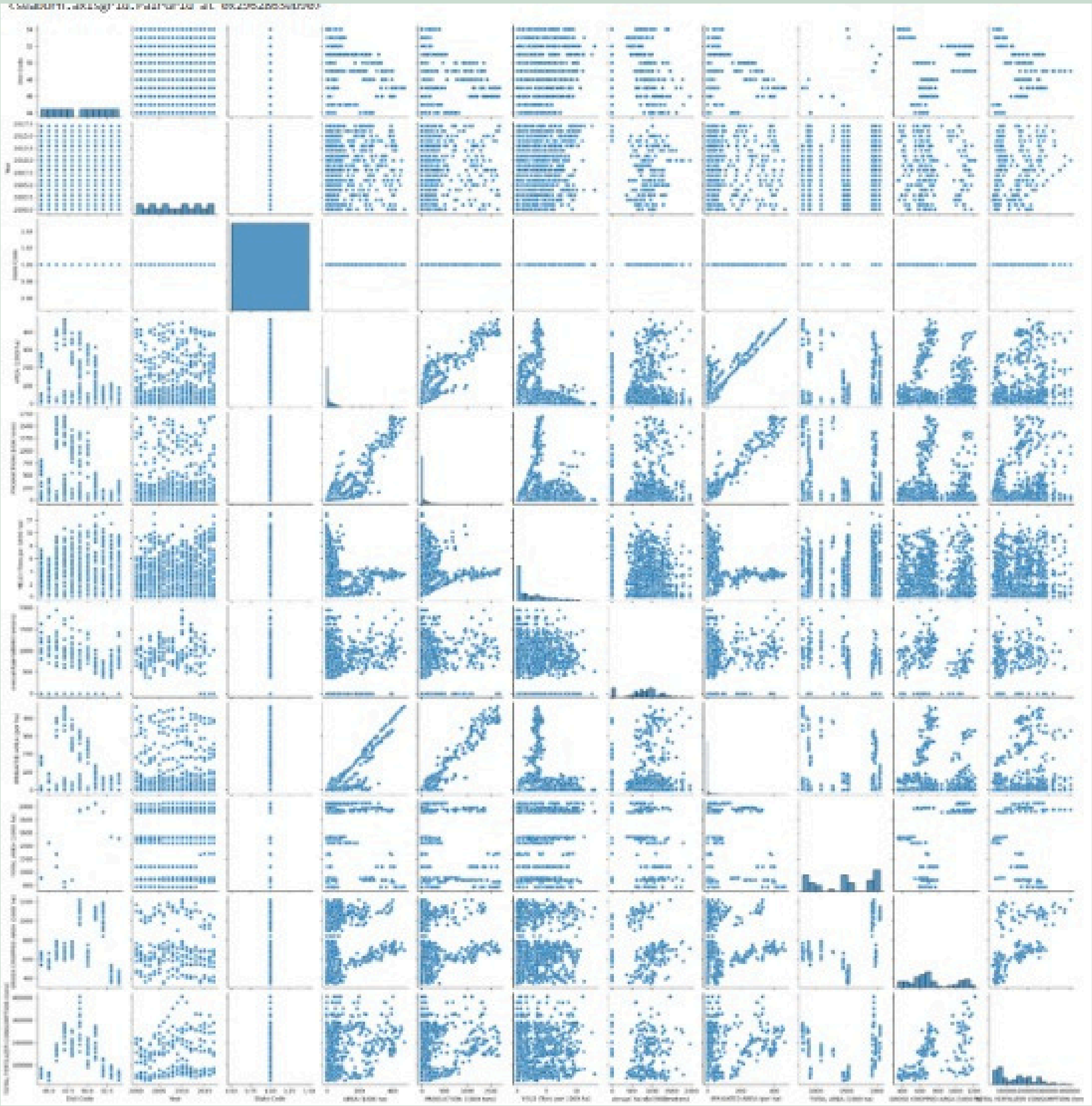
Dataset Overview

	Dist Code	Year	State Code	State Name	Dist Name	Crop Name	AREA (1000 ha)	PRODUCTION (1000 tons)	YIELD (Tons per 1000 ha)	Annual Rainfall(Millimeters)	IRRIGATED AREA (per ha)	TOTAL AREA (1000 ha)	GROSS CROPPED AREA (1000 ha)	TOTAL FERTILIZER CONSUMPTION (tons)
0	44	2000	1	Andhra Pradesh	Srikakulam	Rice	274.60	511.34	1.862127	982.0	177.21	899.31	638.03	56733
1	44	2001	1	Andhra Pradesh	Srikakulam	Rice	235.44	502.11	2.132645	1053.0	153.83	899.00	614.00	52182
2	44	2002	1	Andhra Pradesh	Srikakulam	Rice	201.50	328.50	1.630273	799.0	138.48	899.00	539.50	44893
3	44	2003	1	Andhra Pradesh	Srikakulam	Rice	246.96	545.82	2.210155	1306.0	163.83	899.31	620.14	55358
4	44	2004	1	Andhra Pradesh	Srikakulam	Rice	256.78	601.96	2.344264	970.0	167.72	899.31	631.74	57747
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1381	54	2013	1	Andhra Pradesh	Chittoor	Cotton	0.35	0.16	0.457143	1418.9	0.35	1515.10	417.07	47808
1382	54	2014	1	Andhra Pradesh	Chittoor	Cotton	0.85	0.36	0.423529	NaN	0.81	1515.10	378.84	50883
1383	54	2015	1	Andhra Pradesh	Chittoor	Cotton	1.14	0.36	0.315789	NaN	1.14	1515.10	430.20	67508
1384	54	2016	1	Andhra Pradesh	Chittoor	Cotton	0.78	0.31	0.397436	NaN	0.76	1515.10	404.22	60481
1385	54	2017	1	Andhra Pradesh	Chittoor	Cotton	0.50	0.24	0.480000	NaN	0.50	1515.10	420.58	68242

HeatMap: Checking for Null values

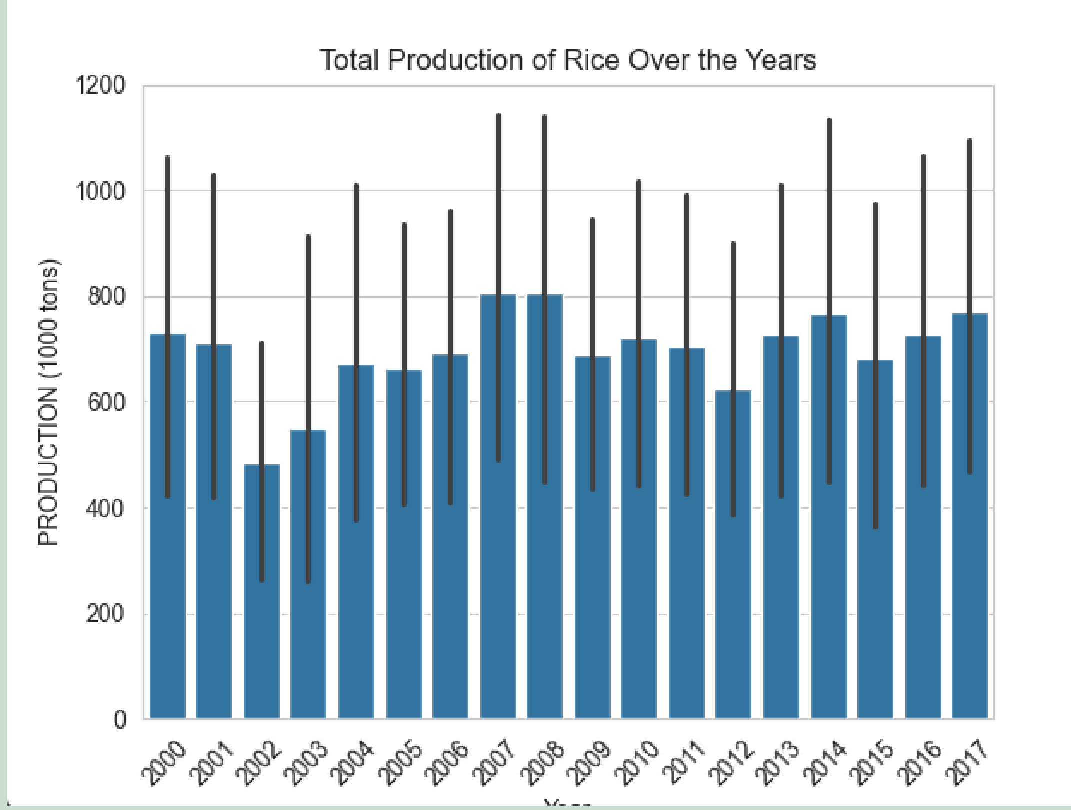


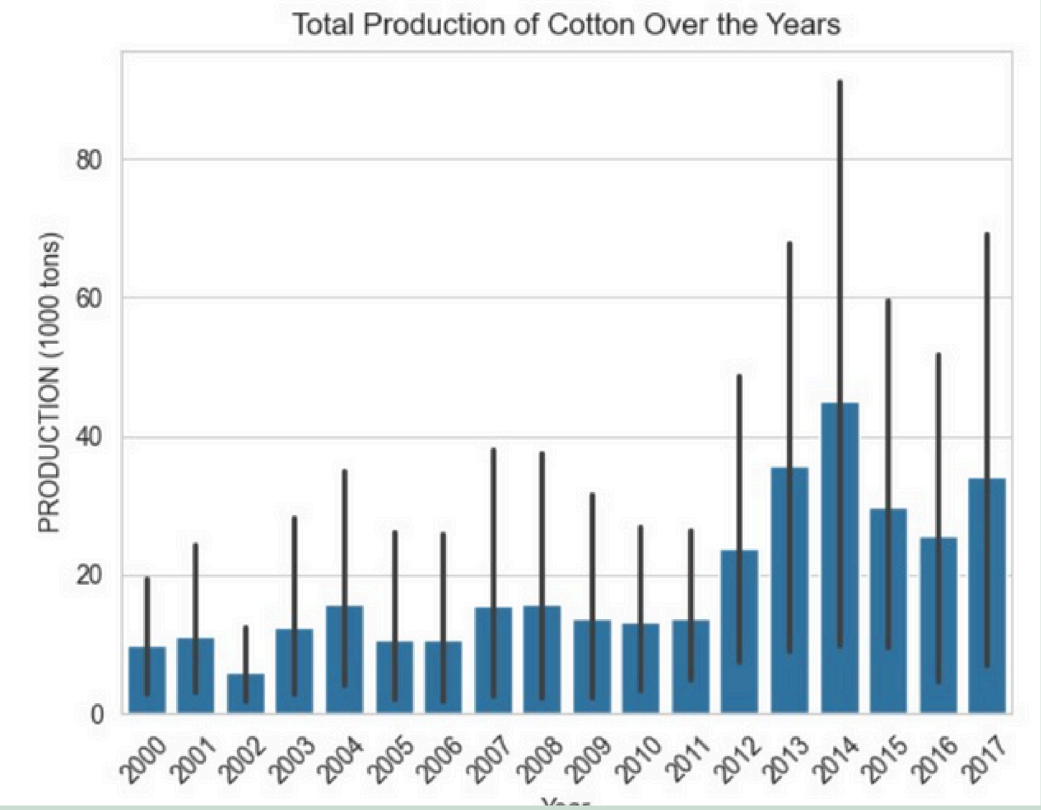
PairPlots for different parameters

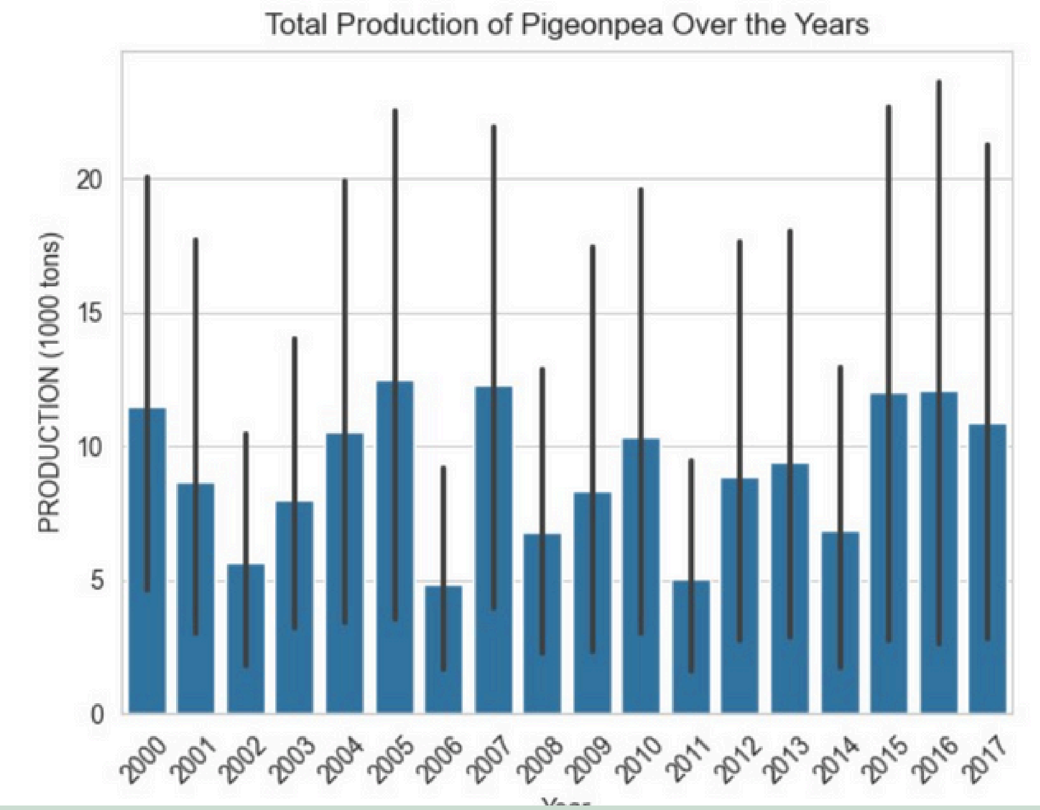


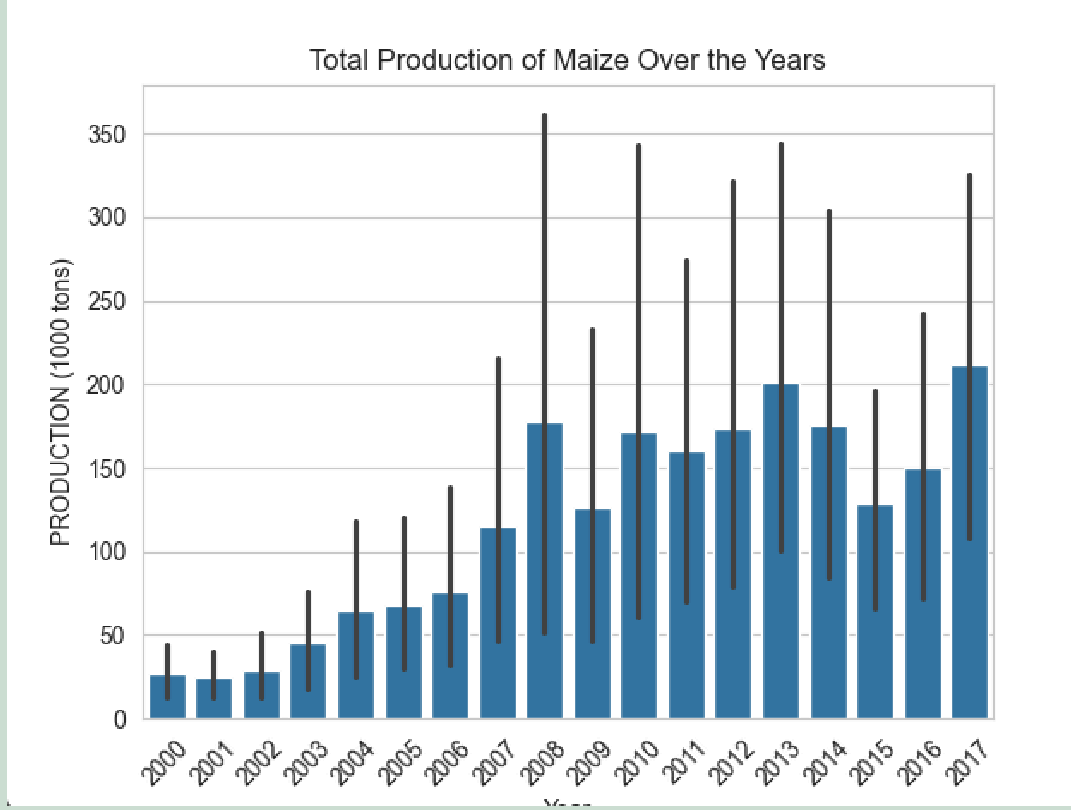
Total Production Over the years

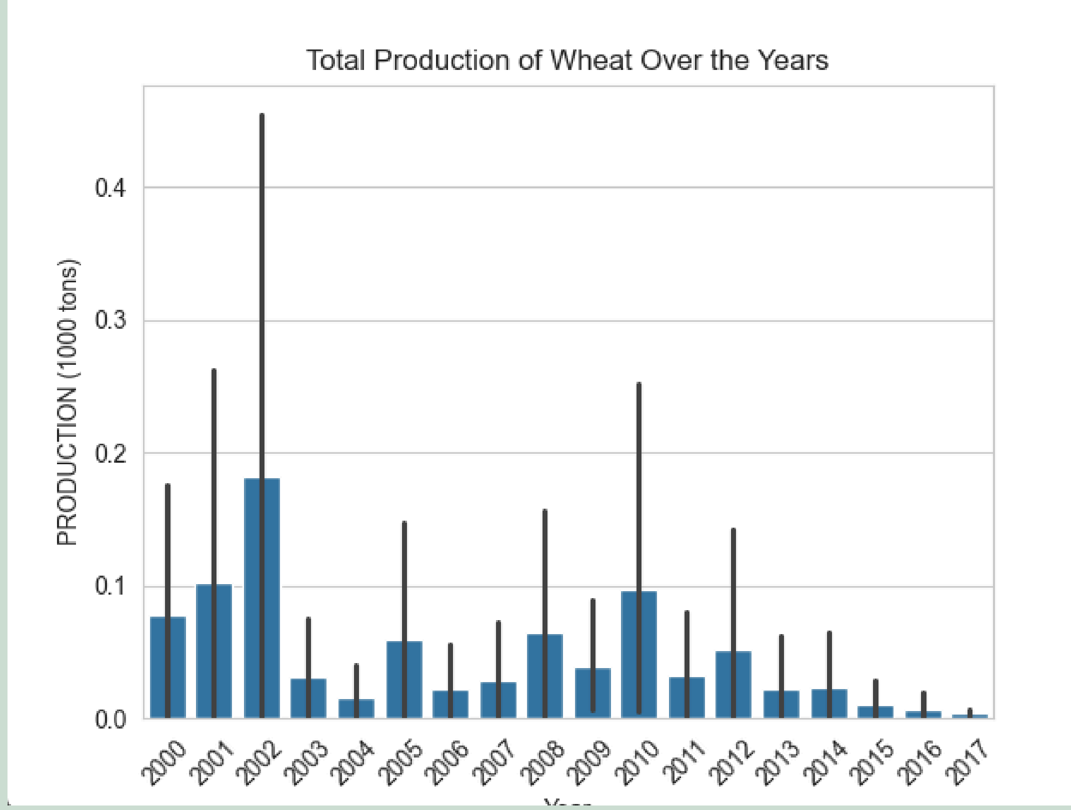
of major crops

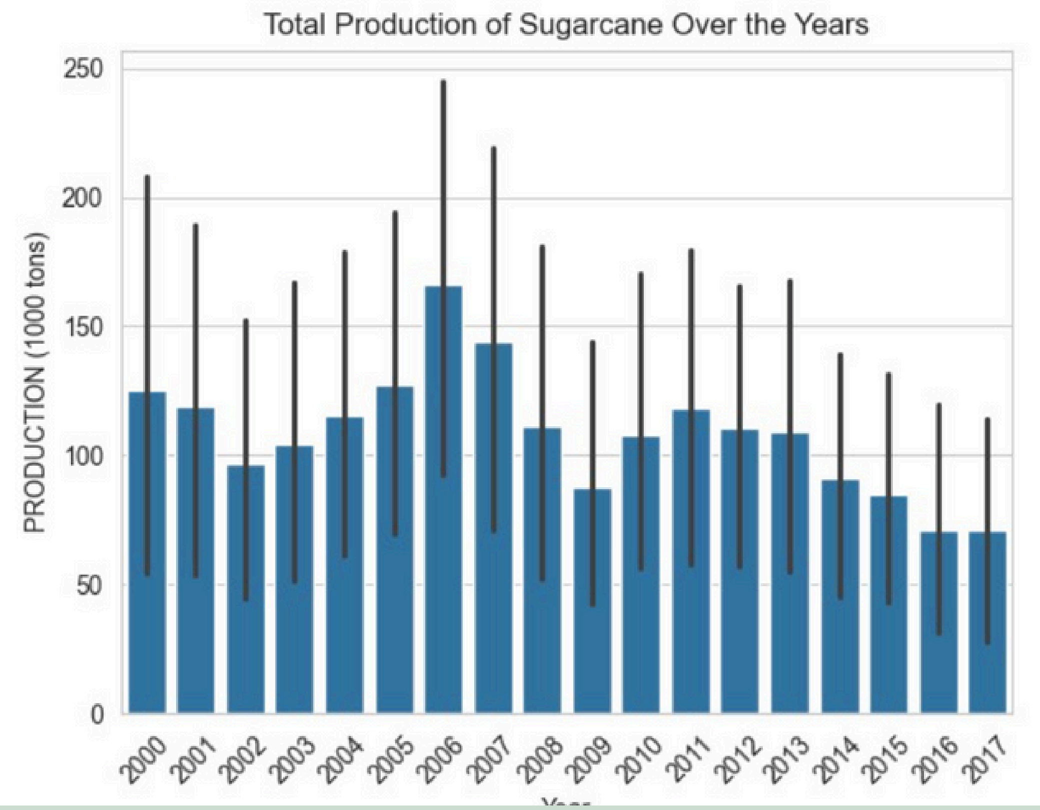




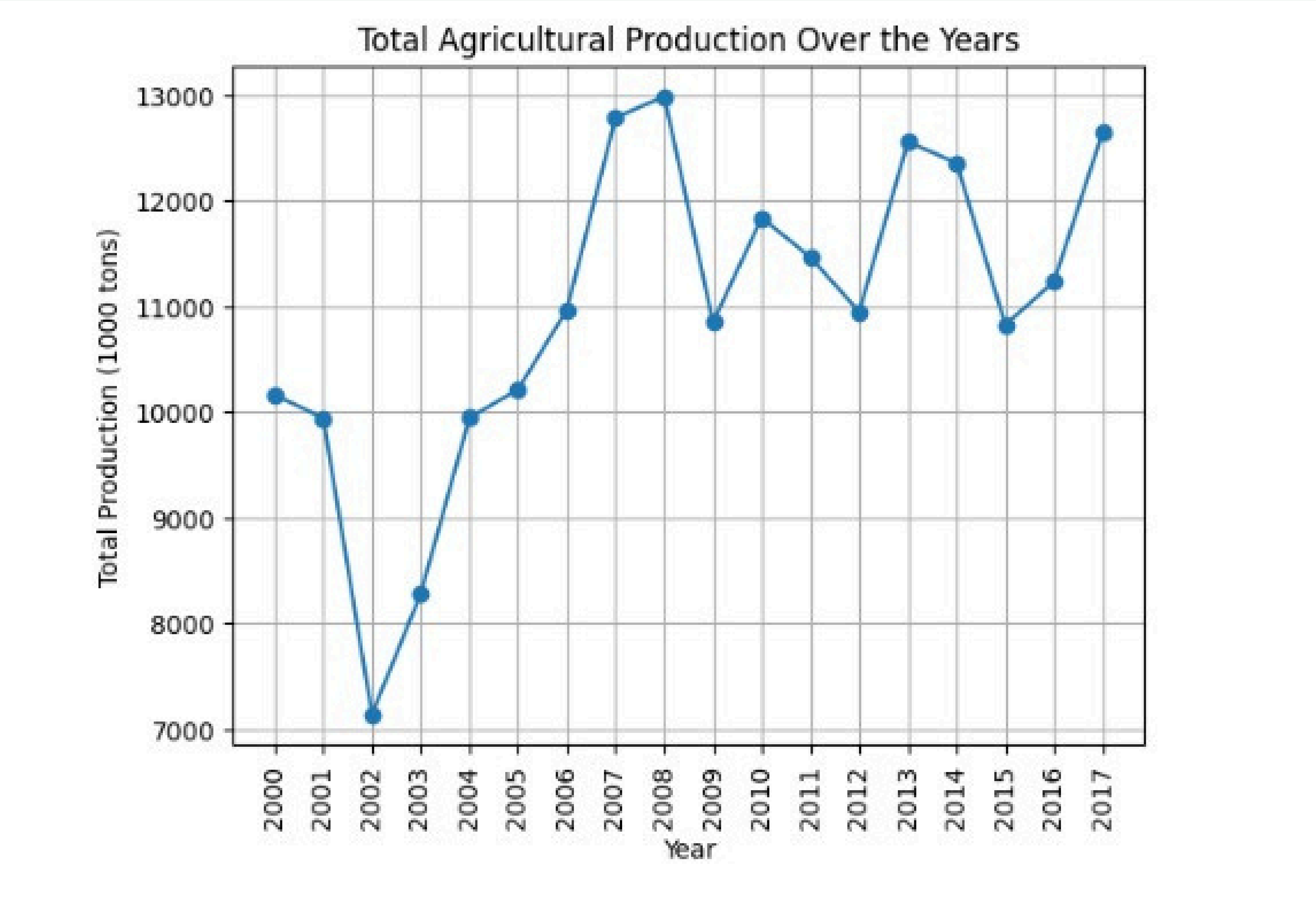






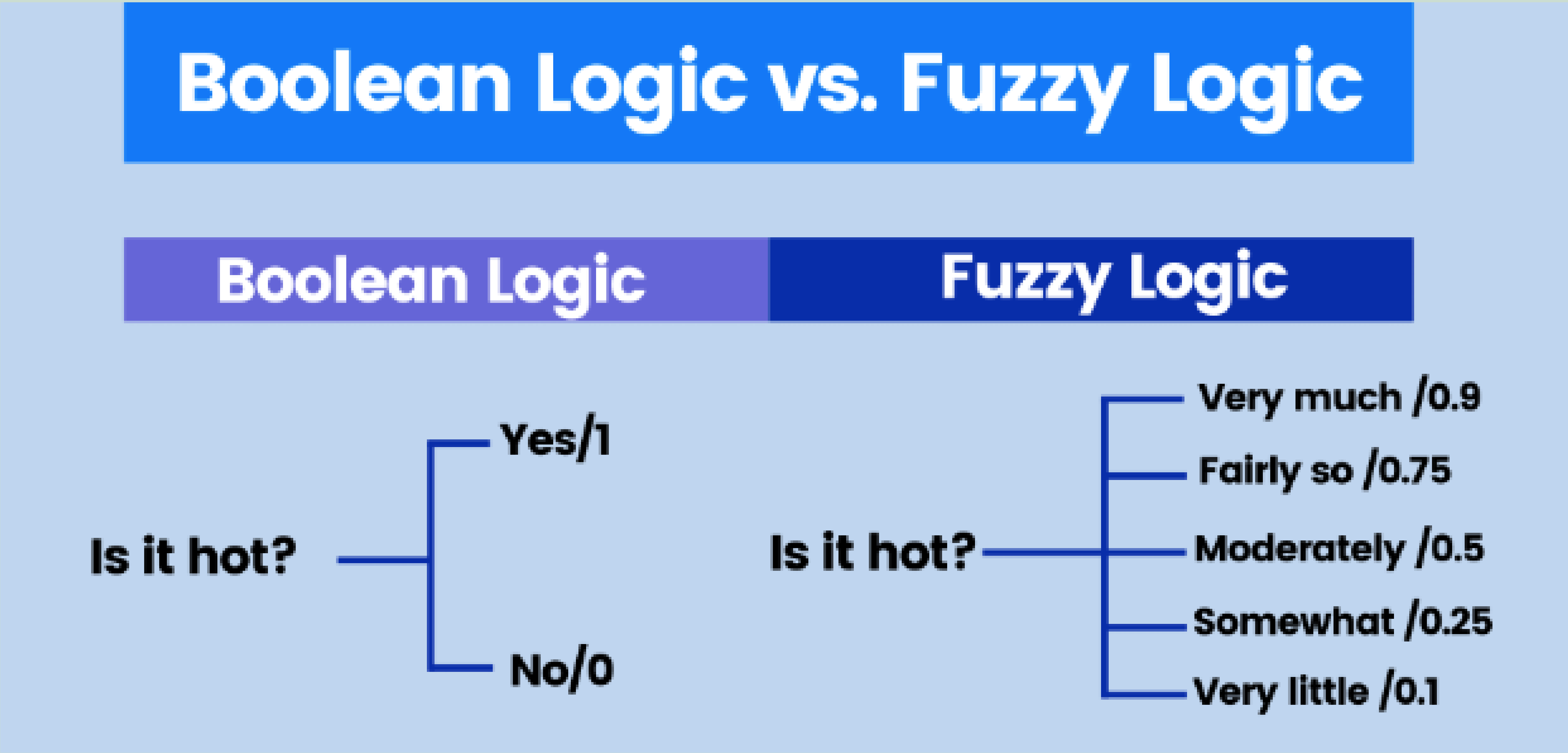


Total agricultural production

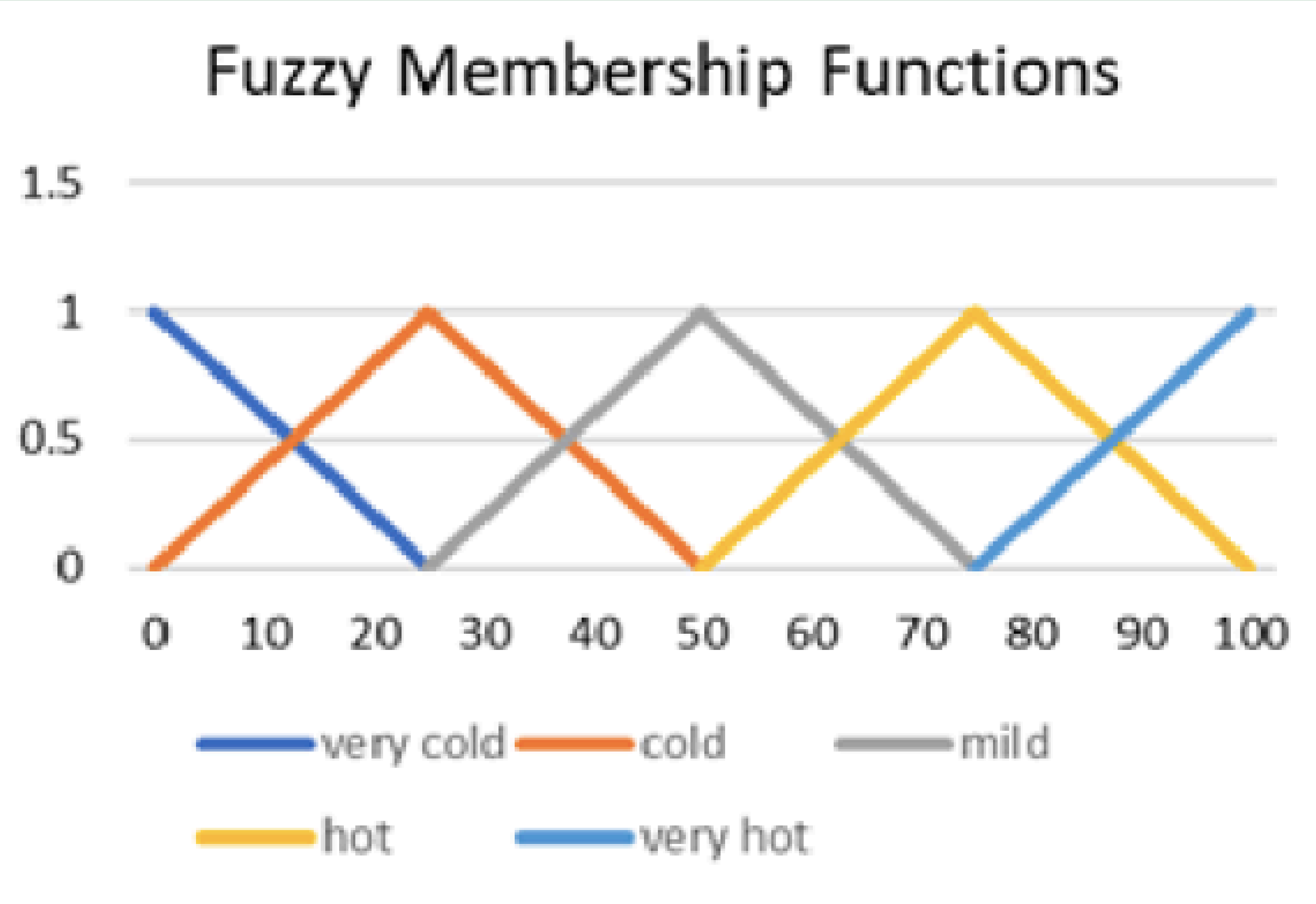


Fuzzy Logic for Handling Uncertainty

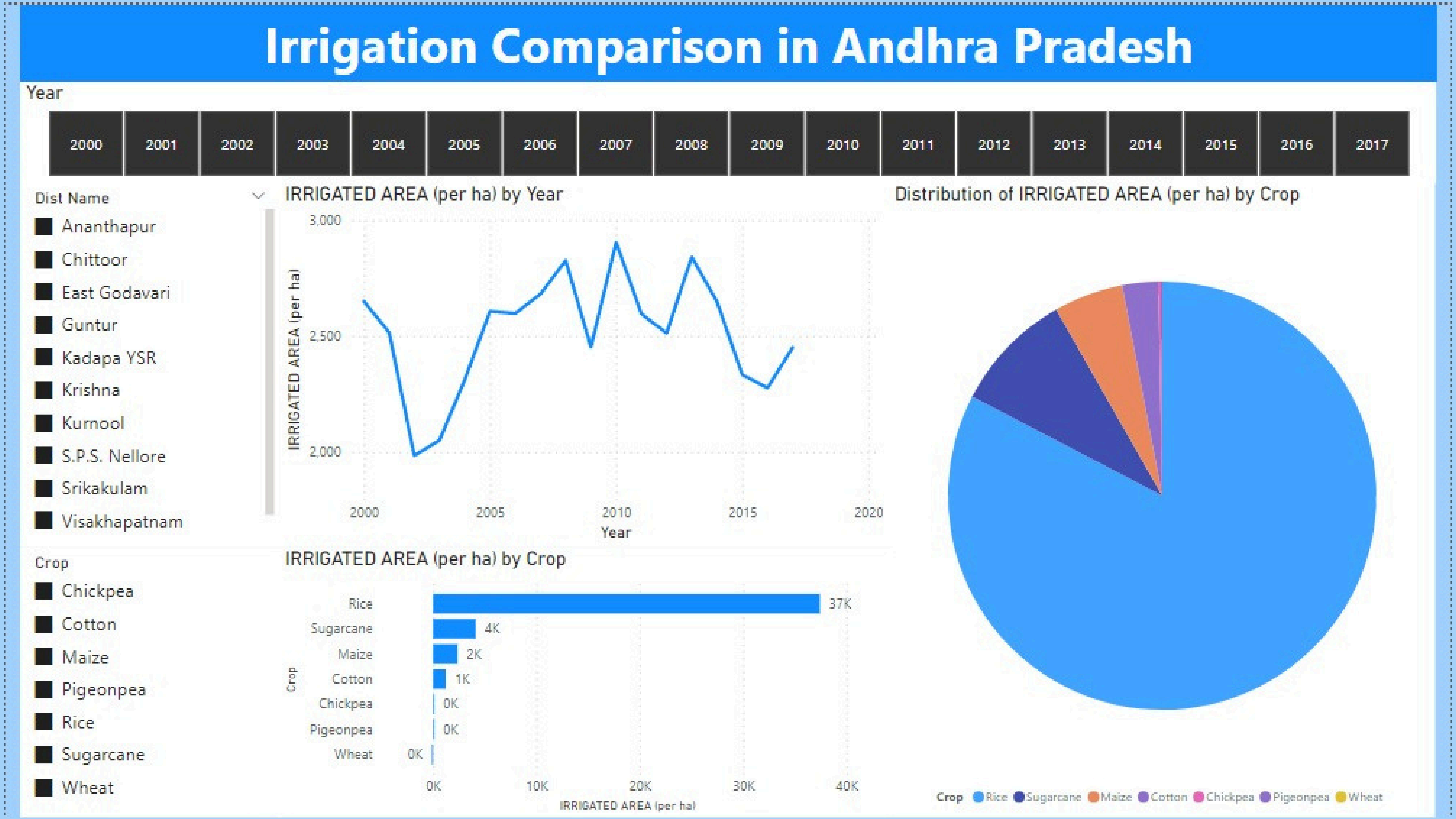
What does fuzzy mean and why is it important?



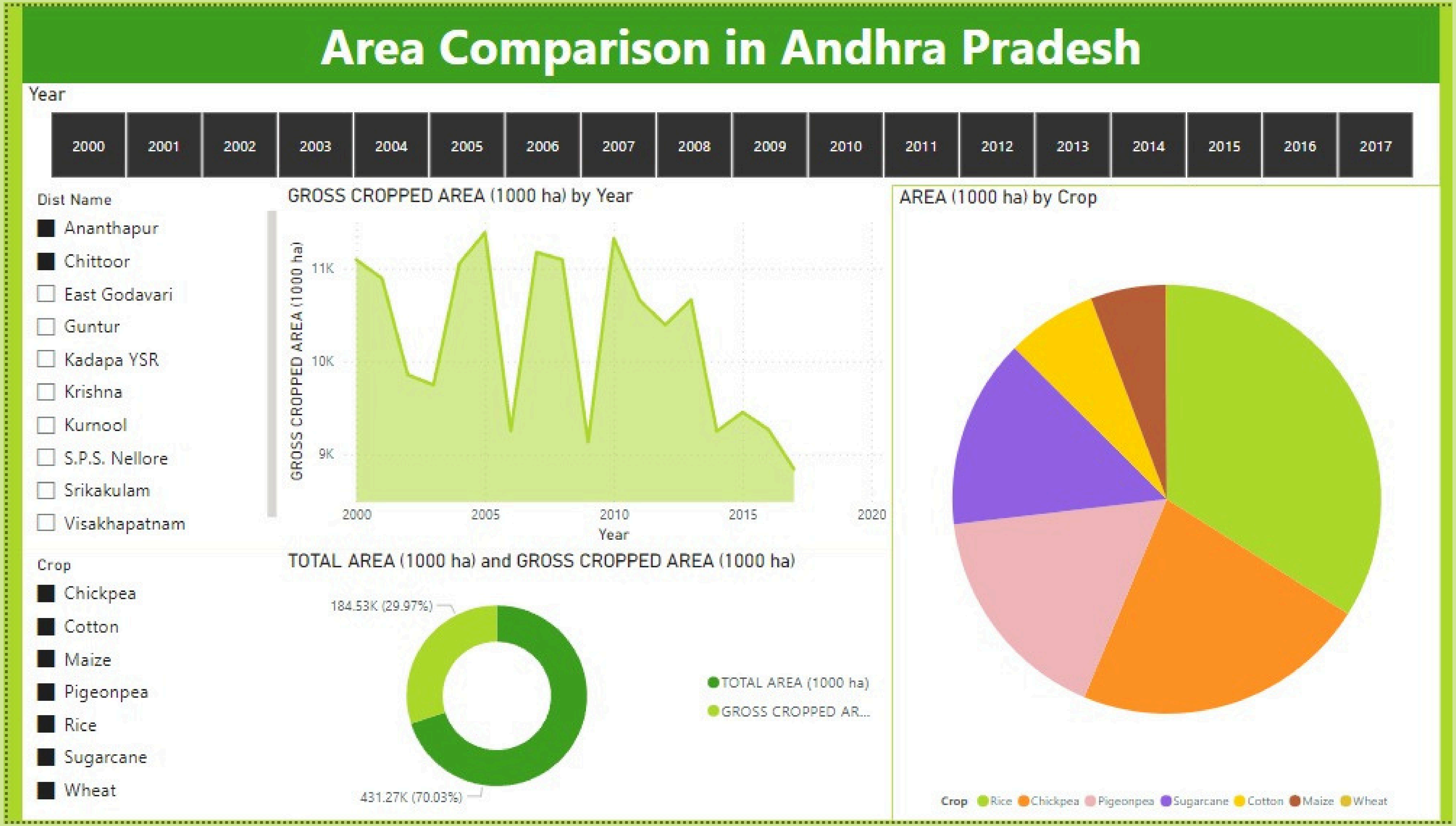
Fuzzy Membership Function: Triangular



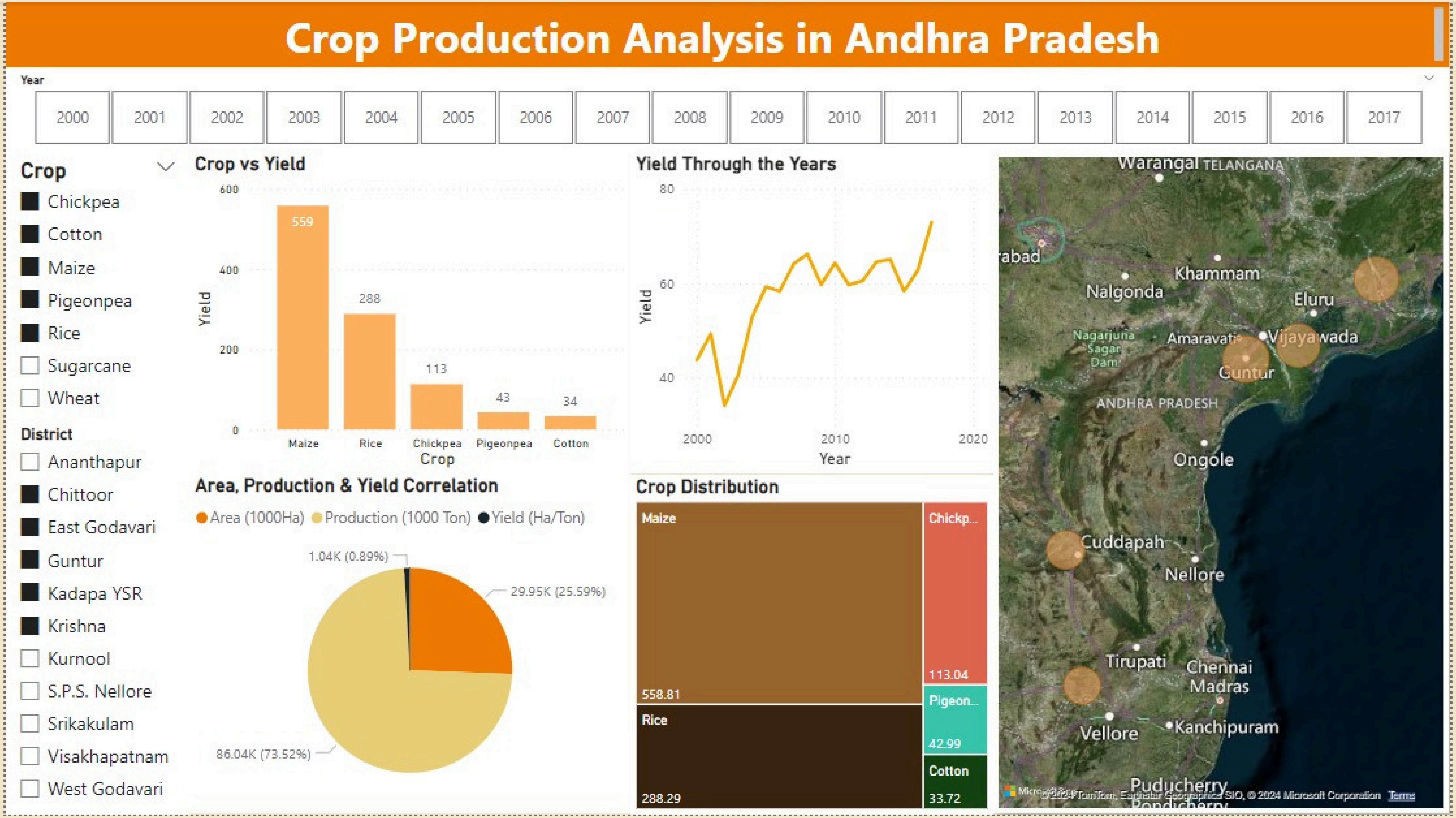
Trend Report: Irrigation Comparison



Trend Report: Area Comparision



Trend Report: Production Analysis



Advantages of using ML Algorithms

Enhanced Effectiveness

This approach proves more effective and precise in identifying patterns, saving farmers considerable time and resources. Machine learning can swiftly evaluate a larger volume of data, surpassing the efficiency of traditional methods.

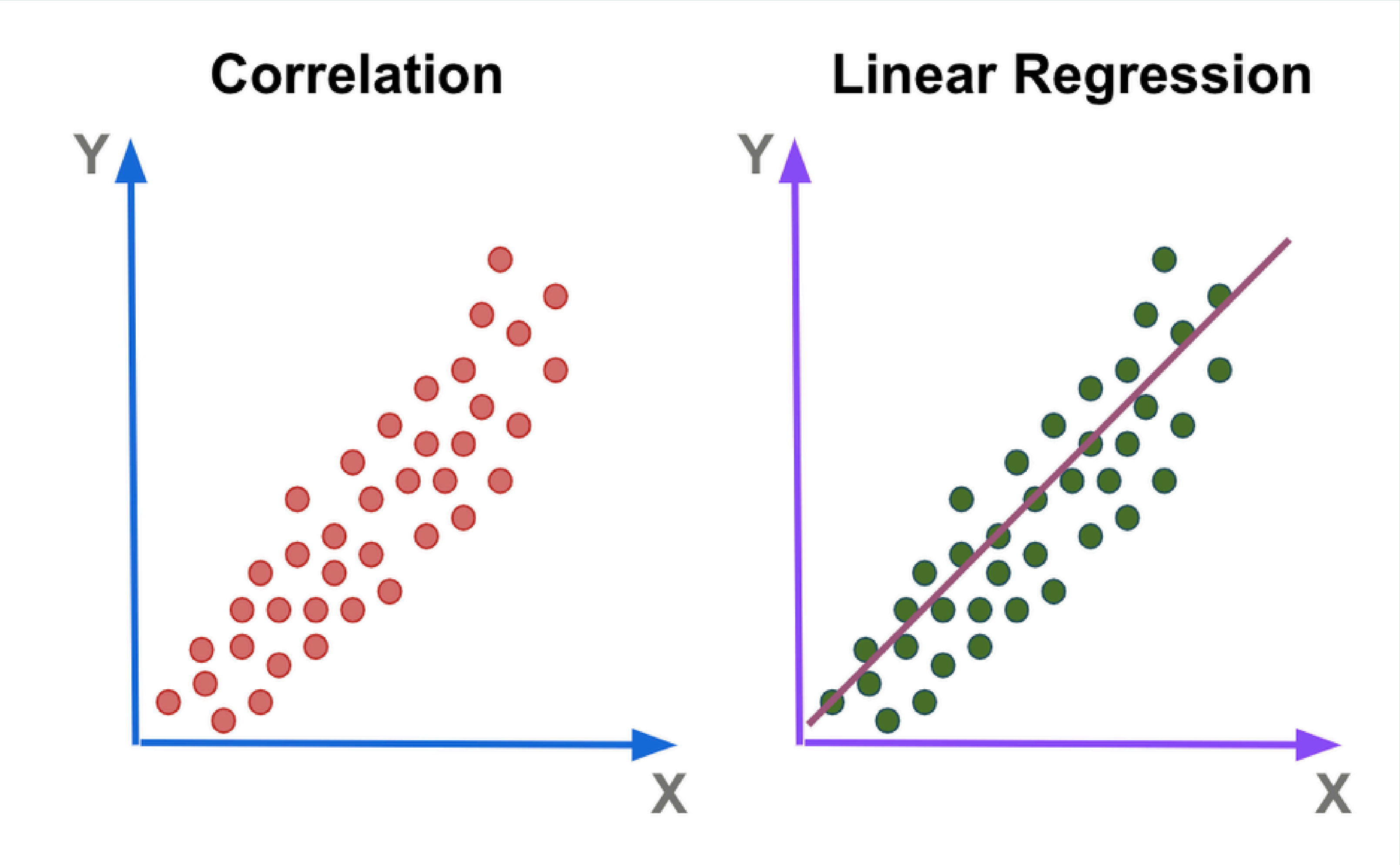
Increased Crop Yield

Leveraging diverse data sources, including weather patterns, soil quality, and historical machine learning algorithms, empowers farmers to make well-informed decisions that contribute to heightened crop yields. The comprehensive analysis facilitates strategic planning and resource allocation.

Cost Reduction

Machine learning aids farmers in optimizing resource utilization, such as water, fertilizer, and pesticides, by providing insights into crop development and health. This not only saves costs but also diminishes the environmental impact of agriculture, aligning with sustainable practices.

Linear Regression



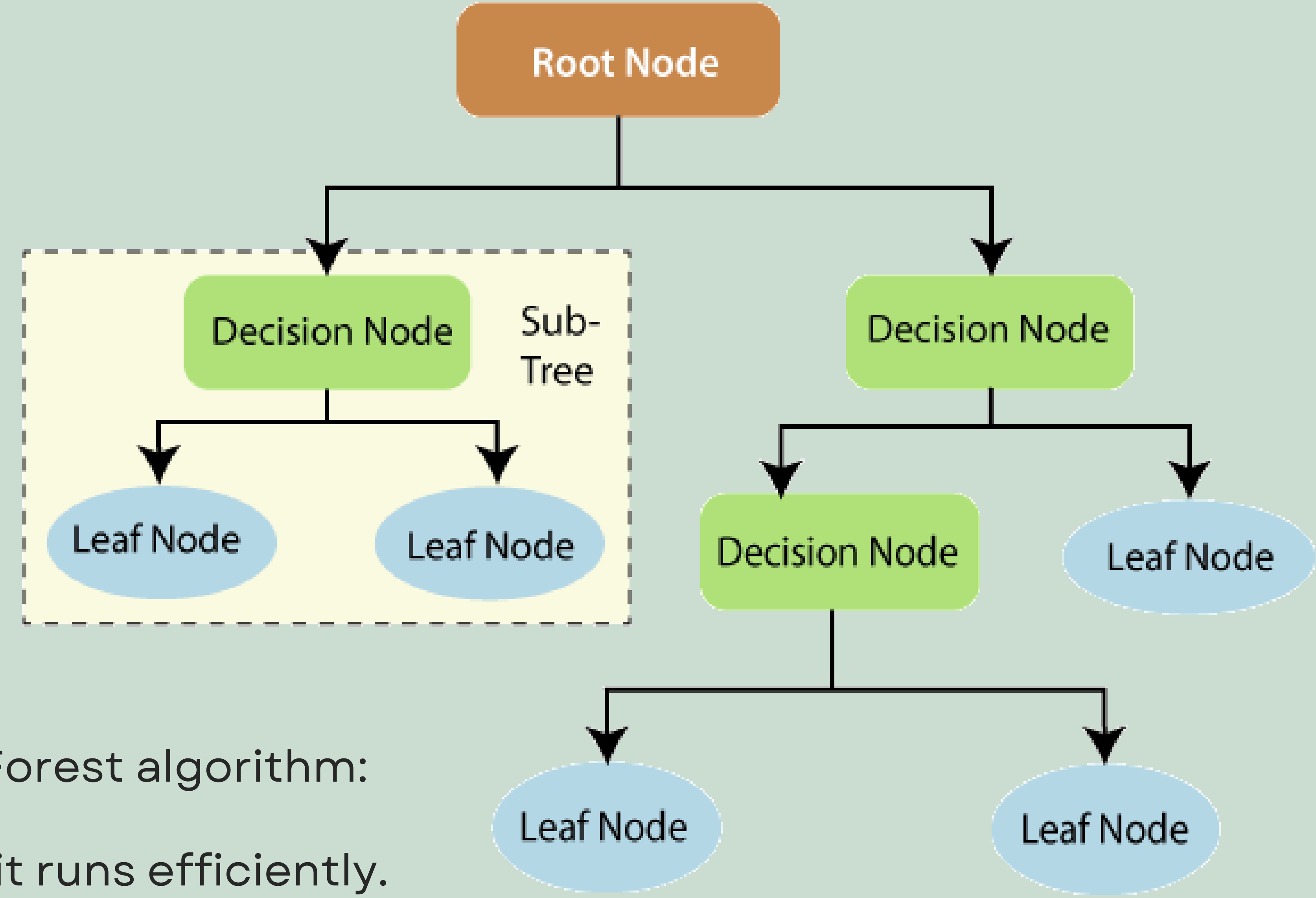
Linear regression can be employed to model the relationship between independent variables (such as location, season, and crop type) and the dependent variable (crop yield). By training the model on historical data, the algorithm learns the coefficients of the linear equation that best fits the observed yield data.

Linear regression allows you to assess the significance of different features in predicting crop yield. This can help identify which factors, such as specific locations, seasons, or crop types, have a more pronounced impact on the overall yield. Linear regression models are relatively interpretable, making it easier to understand the influence of each input variable on the predicted outcome. This interpretability is valuable for communicating findings to stakeholders, including farmers and agricultural policymakers.

Linear regression is suitable for predicting continuous outcomes, making it well-suited for tasks like crop yield prediction where the output is a numerical value.

Random Forest Algorithm

Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

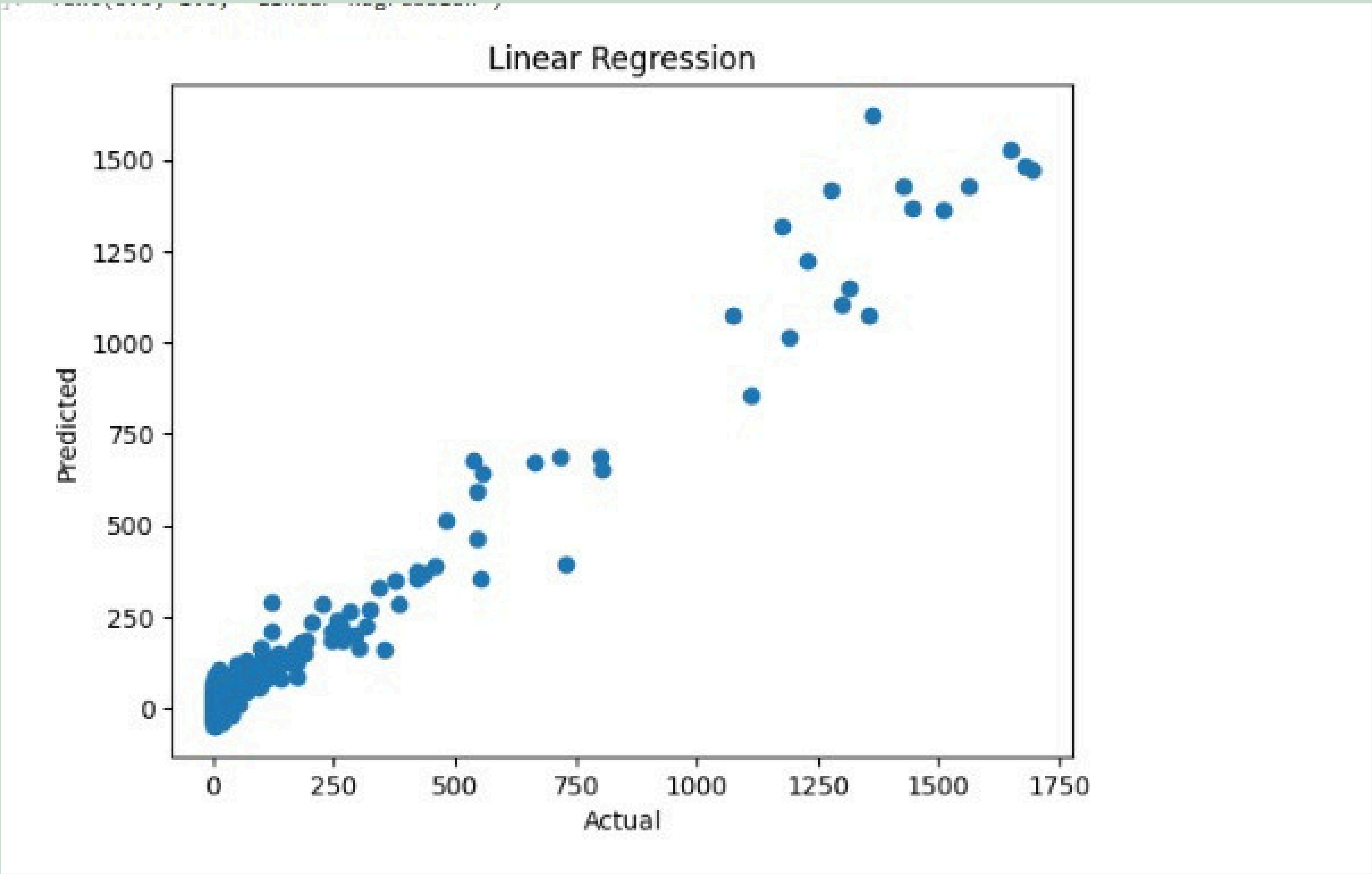


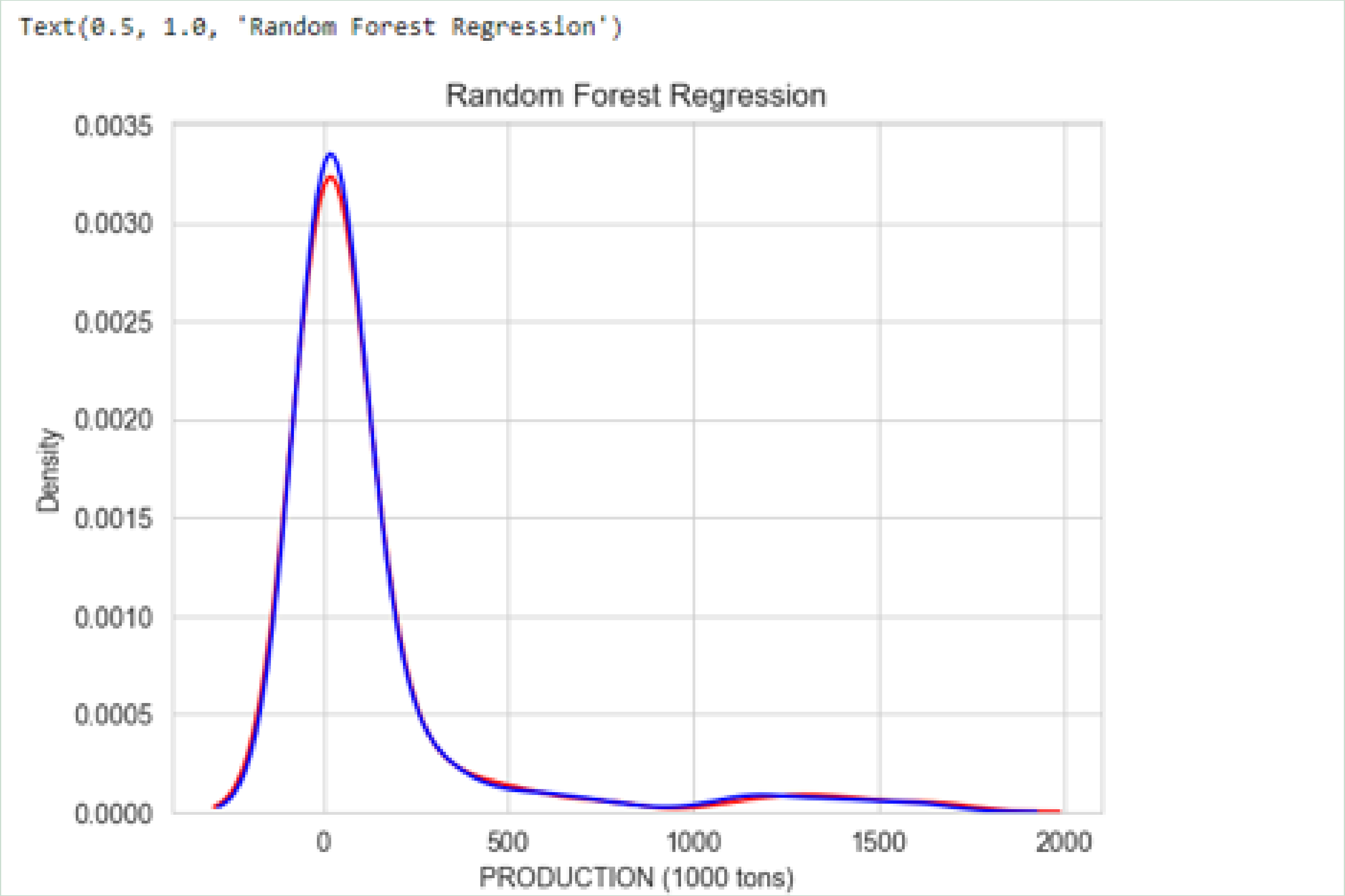
Why use Random Forest?
Below are some points that explain why we should use the Random Forest algorithm:

- It takes less training time as compared to other algorithms.
- It predicts output with high accuracy, even for the large dataset it runs efficiently.
- It can also maintain accuracy when a large proportion of data is missing.

Machine Learning model outputs

Linear regression & Random Forest Regression





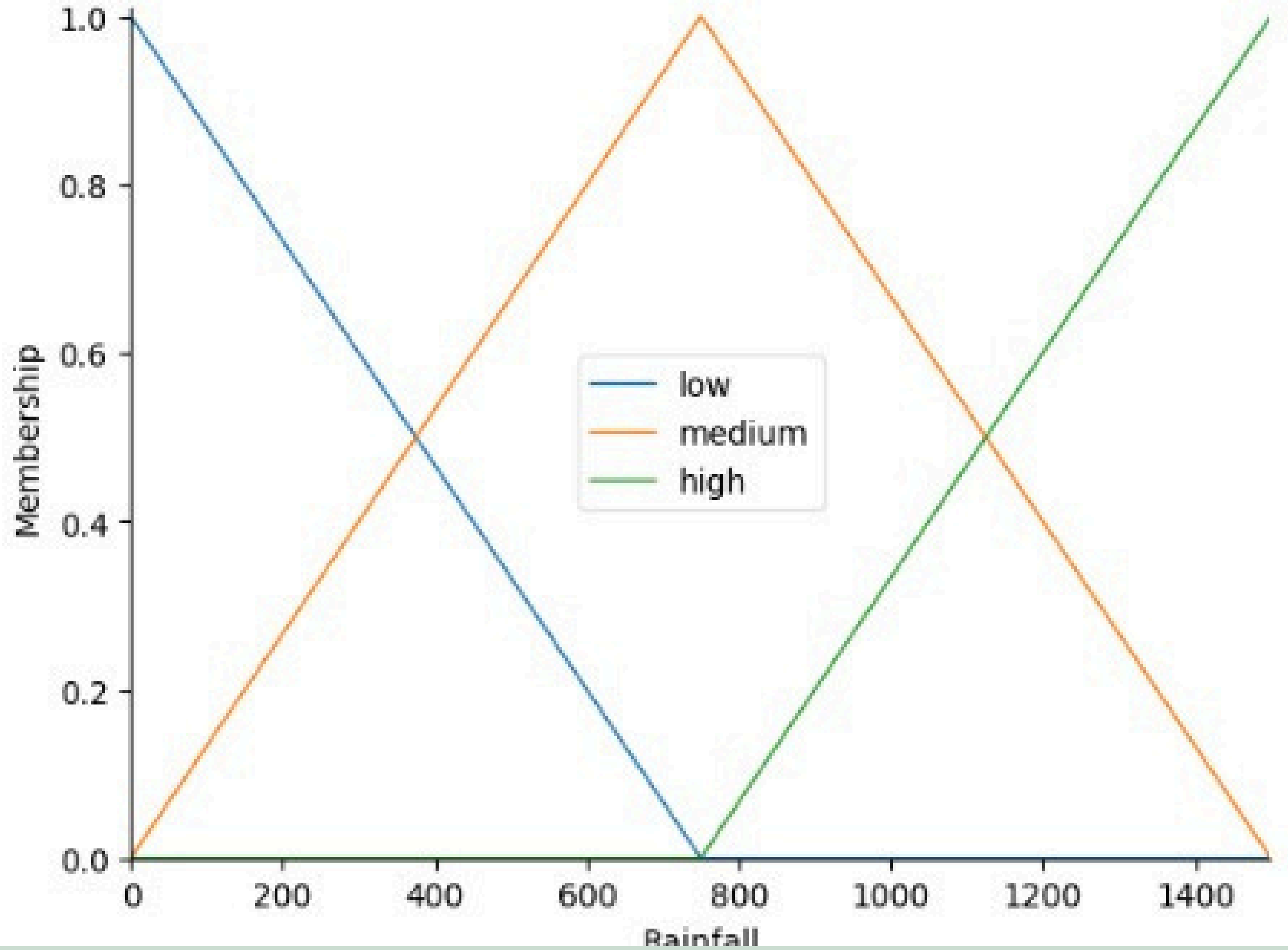
Implementing Fuzzy Logic

[77]:
```
#Defining fuzzy variable for rainfall
rainfall = ctrl.Antecedent(np.arange(0, 1500, 1), 'Rainfall')

#Defining membership function for rainfall
rainfall['low'] = fuzz.trimf(rainfall.universe, [0, 0, 750])
rainfall['medium'] = fuzz.trimf(rainfall.universe, [0, 750, 1500])
rainfall['high'] = fuzz.trimf(rainfall.universe, [750, 1500, 1500])

#Displaying membership function
rainfall.view()
```

```
C:\Users\dell\AppData\Roaming\Python\Python310\site-packages\skfuzzy\control\fuzzyvariable.py:122: UserWarning: Matplotlib is currently using module://matplotlib_inline.backend_inline, which is a non-GUI backend, so cannot show the figure.
  fig.show()
```



Fuzzy Inference & Rules Table

S. No	Rule
1.	IF Rainfall is Low AND Irrigated Area is Medium THEN Production is Medium
2.	IF Rainfall is Low AND Irrigated Area is Small THEN Production is Low
3.	IF Rainfall is Low AND Irrigated Area is Large THEN Production is Medium
4.	IF Rainfall is Medium AND Irrigated Area is Large THEN Production is High
5.	IF Rainfall is Medium AND Irrigated Area is Medium THEN Production is High
6.	IF Rainfall is Medium AND Irrigated Area is Small THEN Production is Medium
7.	IF Rainfall is High AND Irrigated Area is Small THEN Production is Medium
8.	IF Rainfall is High AND Irrigated Area is Medium THEN Production is High
9.	IF Rainfall is High AND Irrigated Area is Large THEN Production is High

Rule	Rainfall	Irrigated Area	Production
Rule 1	Low	Small	Low
Rule 2	Low	Medium	Medium
Rule 3	Low	Large	Medium
Rule 4	Medium	Small	Medium
Rule 5	Medium	Medium	High
Rule 6	Medium	Large	High
Rule 7	High	Small	Medium
Rule 8	High	Medium	High
Rule 9	High	Large	High

```
      Rainfall  Irrigated_area      Impact
0        982.0          177.21  938.218179
1       1053.0          153.83  884.215250
2        799.0          138.48  860.258653
3       1306.0          163.83  909.622858
4        970.0          167.72  918.417090
...        ...             ...         ...
1381    1418.9            0.35  749.999360
1382       0.0            0.81  255.626346
1383       0.0            1.14  257.895739
1384       0.0            0.76  255.281276
1385       0.0            0.50  253.482089

[1386 rows x 3 columns]
```

Here Impact is the Output Production values.

RESULTS & CONCLUSIONS

The random forest algorithm outperformed linear regression in handling the complexity and non-linearity of agricultural data, resulting in more accurate production predictions. Integrating fuzzy logic techniques further enhanced the models' ability to account for uncertainties, providing more reliable recommendations and predictions.

The error analysis revealed that the random forest model achieved an MSE value of 1603.84 and MAE value of 15.0, indicating a strong correlation between the predicted and actual crop yields. The visualization tools offered a clear understanding of the factors influencing crop selection and yield, empowering farmers and policymakers with actionable insights in the face of uncertainty.

1

Market Price Analysis

2

Correlation of datasets

3

Improving accuracy

4

Scalable in terms of states

5

Parameters expansion

FUTURE SCOPE

features to upgrade and improve

REFERENCES

1. http://data.icrisat.org/dld/src/crops.html
2. https://agmarknet.gov.in/PriceTrends/
3. https://www.indiastat.com/
4. https://towardsdatascience.com/fuzzy-inference-system-implementation-in-python-8af88d1f0a6e
5. https://pib.gov.in/PressReleaseIframePage.aspx?PRID=2002012
6. https://cacp.dacnet.nic.in/ViewReports.aspx?Input=2&PageId=39&KeyId=669
7. https://data.gov.in/catalog/district-wise-season-wise-crop-production-statistics-0
8. https://ieeexplore.ieee.org/document/10029009
9. https://www.isroset.org/pub_paper/IJSRMSS/9-IJSRMSS-01819.pdf
10. https://notebook.community/castelao/CoTeDe/docs/notebooks/fuzzy_logic