PROJECT REPORT

Date	**17 November, 2022**
Team ID	**PNT2022TMID45976**
Project Name	**Project - Estimate the crop yield during Data Analytics**

ESTIMATE THE CROP YIELD USING DATA ANALYTICS

1. INTRODUCTION

1.1 Project Overview:

Crop production in India is one of the most important sources of income and India is oneof the top countries to produce crops. Crop production may vary due to various factors such as season, temperature, place, year etc. As per this project we will be analyzing some important visualization, creating a dashboard and by going through these we will get most of the insights of Crop production in India. We are going to analyse the crop yield production using different attributes like year, area, season and so on. Using those attributes, we help to make a good crop yield production in future from the past data. It is quite helpful for the farmers and the people to get right crops to cultivate in a particularseason

1.2 Purpose:

Technology is used to make the people life better.It is to make and get a better crop yield production using data analytics in python from the past crop productions. It is toget insights from the past data with important parameters and help to make a good decisions for the crop yield production.

2. LITERATURE SURVEY:

2.1 Existing problem:

Crop Yield Prediction Using Machine Learning Algorithm [1] Machine learning (ML) is significant since it offers a decision-support tool for Crop Yield Prediction (CYP), which may help with decisions like which crops to cultivate and what to do during the crop's growing season. Crop yield estimation's major purpose is to boost agricultural crop production, and it does so using a variety of well-established models. To assist farmersin maximizing agricultural yield, timely instructions to forecast future crop output and analysis are required

Estimation of Crop Yield From Combined Optical and SAR Imagery Using Gaussian Kernel Regression [2] The synthetic aperture radar (SAR) interferometric coherence can complement optical data for the estimation of crop growth parameters, but it has not been yet investigated for predicting crop yield. Many studies have used machine-learningmethods, such as neural networks, random forest, and Gaussian process regression, to estimate crop yield from remotely sensed data.

The prediction accuracy was assessed using in situ measured yield data collected in 2019 and 2020 over Xinghua county in Jiangsu Province, China. In all cases, Gaussian kernel regression outperformed the probabilistic Gaussian regression and Bayesian linear inference.

Crop Yield Estimation in India Using Machine Learning [3] The main aim of this paper is to predict crop yield using area, yield, production, and area under irrigation. Four machinelearning techniques Decision Tree, Linear Regression, Lasso regression, and Ridge Regression have been applied to estimate the crop yield. Cross validations methods, for validation, mean absolute error, mean squared error, and root mean squared error, wereused to validate. The Decision tree outperforms other machine learning techniques.

2.2 References:

[1] **R. J, V. K. G. Kalaiselvi, A. Sheela, D. S. D and J. G, "Crop Yield Prediction Using Machine Learning Algorithm," 2021 4th International Conference on Computing andCommunications Technologies (ICCCT), 2021, pp. 611-616, doi: 10.1109/ICCCT53315.2021.9711853.**

[2] **Y. Alebele et al., "Estimation of Crop Yield From Combined Optical and SAR ImageryUsing Gaussian Kernel Regression," in IEEE Journal of Selected Topics in Applied EarthObservations and Remote Sensing, vol. 14, pp. 10520-10534, 2021, doi: 10.1109/JSTARS.2021.3118707.**

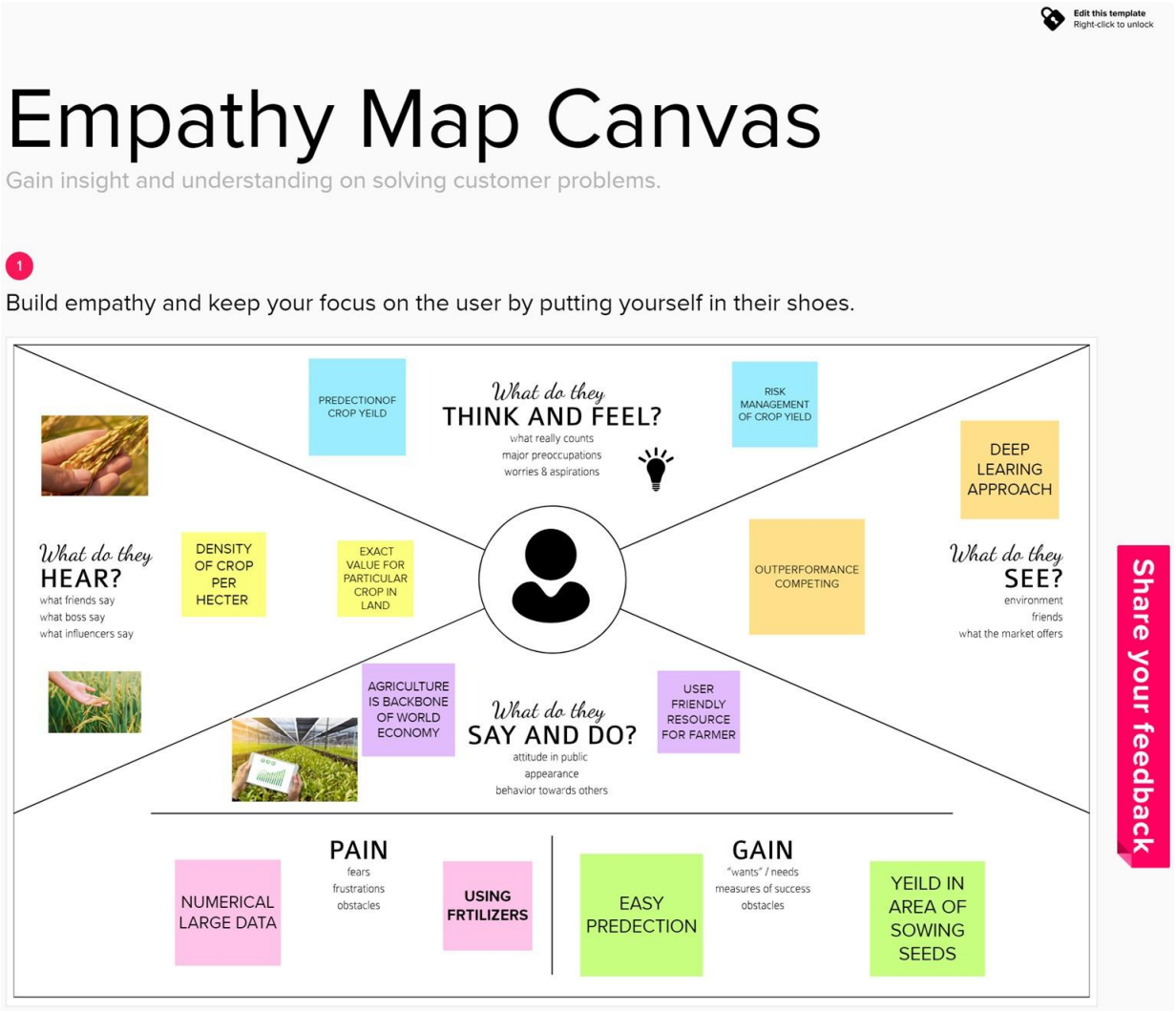
[3] **M. Kavita and P. Mathur, "Crop Yield Estimation in India Using Machine Learning," 2020 IEEE 5th International Conference on Computing Communication and Automation(ICCCA), 2020, pp. 220-224, doi: 10.1109/ICCCA49541.2020.9250915.**

2.3 Problem Statement Definition:

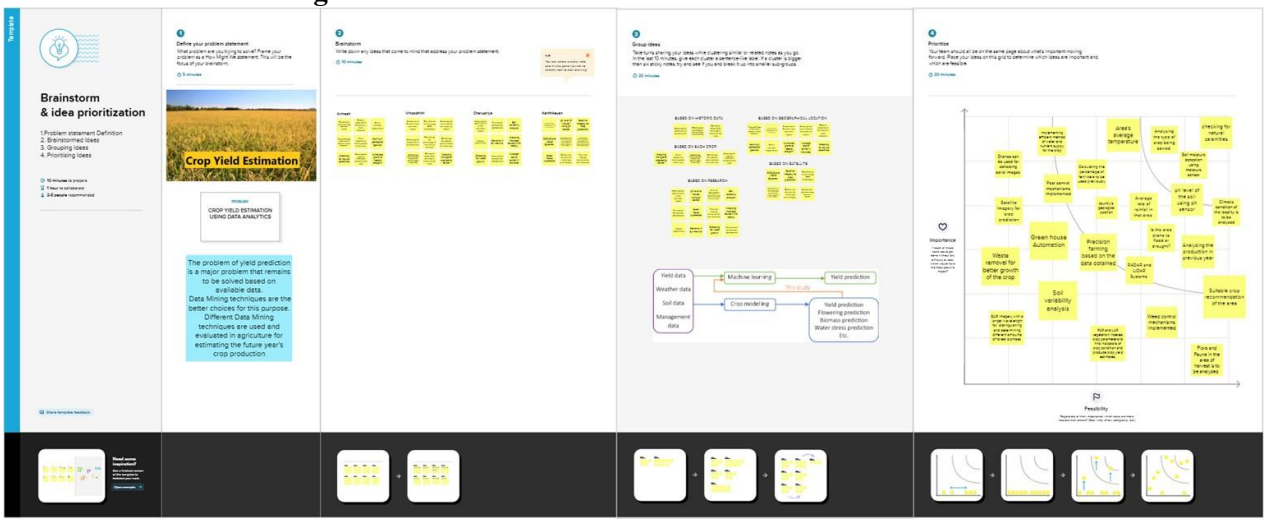
Crop production in India is one of the most important sources of income and India is oneof the top countries to produce crops. Using python, we will be analyzing some important visualization and getting interpretation and going through these we will get most of the insights of Crop production in India and it is used to help the people to make a better choice for cultivation depends on a season and location.

3. IDEATION & PROPOSED SOLUTION:

3.1 Empathy Map Canvas:



3.2 Ideation & Brainstorming:



3.3 Proposed Solution:

Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	CROP YIELD PREDICTION ANALYSIS: soil quality analysis to achieve high crop yield through out technology solution. To increase quality and yields, it is crucial to understand the current nutrient levels of the soil. Analyze pH to check for soil acidity and alkalinity, Conductivity to determine optimised fertilizer usage as well as Sodium, Potassium, Nitrate and Calcium levels. Fertilizer manufactures understand the importance of particle size. It directly affects certain aspects including release rates, fertilizer potency and also hazardous dust generation. To ensure quality and consistency, a minimum frequency of measurements must be made and our Particle Analyzers are ideal for this task.
2.	Idea/Solution description	ICP-OES and LaquaTwin ICP-OES is an analytical technique that is widely utilized throughout the agricultural industry and within research and development institutions. It is an ideal agriculture analysis technique to determine major and minor elements in soils and plants as well as detecting heavy metal contents. Our Ultima has been used in many crop science institutions to undertake such important content detection.

3.	Novelty/Uniqueness	How ICP-OES is used: • Drinking water quality and safety • Soil analysis • Environmental impact assessments • Food safety • Pharmaceutical analysis LquaTwin Detection systems used with ICP-OES.
4.	Social Impact/Customer satisfaction	The main objectives of this technique in prediction of crop-yield which can be extremely useful to farmers in planning for harvest and sale of grain harvest. Increasing crop yields is a high priority for growers.
5.	Business Model (Revenue Model)	The introduction of technology into the agricultural sector has led to a major rise in productivity. Technology improvements have made new concepts like precision agriculture and have observed and analysed the several crops that are grown, as well as their area and production rates in various states and districts.
6.	Scalability of the Solution	Production rate averaging seasons. Experence a rise in productivity. Yield average in seasons. Boots the crop productivity.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Variations in annual rainfall, average temperature, global increase of atmospheric CO2, and fluctuations in sea levels are some of the major manifestations of climate change, which negatively impact crop yields.
2.	Idea/Solution description	The regression analysis model between historical climatic data and yield data for food crops over the last 30 years in Nepal showed an increase in temperature of approximately 0.02–0.07°C per year in different seasons and a mixed trend in precipitation.
3.	Novelty/Uniqueness	No significant impact of climate variables on yields of all crops the regression analysis revealed negative relationships between maize yield and summer precipitation, between wheat yield and winter minimum temperature, and finally positive relationship was observed between millet yield and summer maximum temperature.

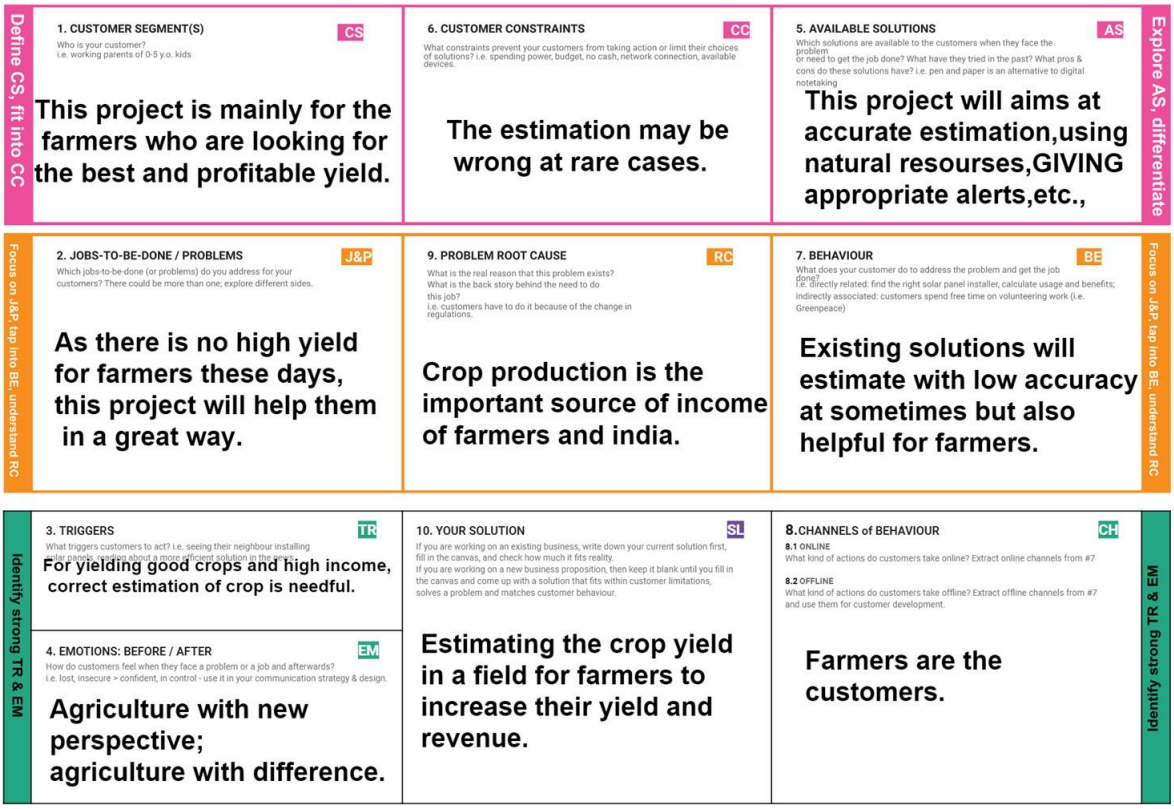
4.	Social Impact/Customer satisfaction	Though the climate changes and global warming, there will be an more yield in crops and grains. The nutrition also be increased by using the estimation of crop yield techniques and analysis.
5.	Business Model (Revenue Model)	Agriculture production and vuleruability. The accurate prediction of crop yield certainly benefits the farmers in choosing the right method to reduce the crop damage and it is used to produce the sustainability of the crops in any kind of environmental temperature. The aspiration of the planned method is to afford transparent, easily accessible, reproducible and for predicting the yield.
6.	Scalability of the Solution	• Increased productivity from warmer temperature. • Decreased moisture stress. • Possibility of growing new crops. • Accelerated. • Maturation rates.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Lack of mechanisation Some parts of country most of agricultural operations in lager parts are carried on human hands using simple and conventional tools.
2.	Idea/Solution description	The progress has been made for mechanizing agriculture. Strategies and programs in farming have been directed towards the replacement of traditional and inefficient implements by Improved ones, enabling the farmers to own tractors, power tillers, harvesters and other machines.
3.	Novelty/Uniqueness	The agriculture machines as also be developed for large industrial base. Strenuous efforts are made to encourage the farmers to adopt technically Advanced with agricultural equipment.
4.	Social Impact/Customer satisfaction	Time consumption in yield analysis. Cover crop and integrated pest management(IPM) have shown excellent result in increasing the profitability of farming in long-term.

		Yield production encourages the farmers. Economize the agricultural production process.
5.	Business Model (Revenue Model)	Without risk management Correct practises and plan for the long term gain can be achieved easily and reduces the capital investment.
6.	Scalability of the Solution	• Livestock management system. • Productivity of soil and water. • Complies with community norms and meet social expectations. • It will be a profitability of the sustainable farming

3.4 Proposed Solution Fit:

Project Title: Estimate the crop yield using data analytics **Project Design Phase-I - Solution Fit Template**



4. REQUIREMENT ANALYSIS:

4.1 Functional Requirement:

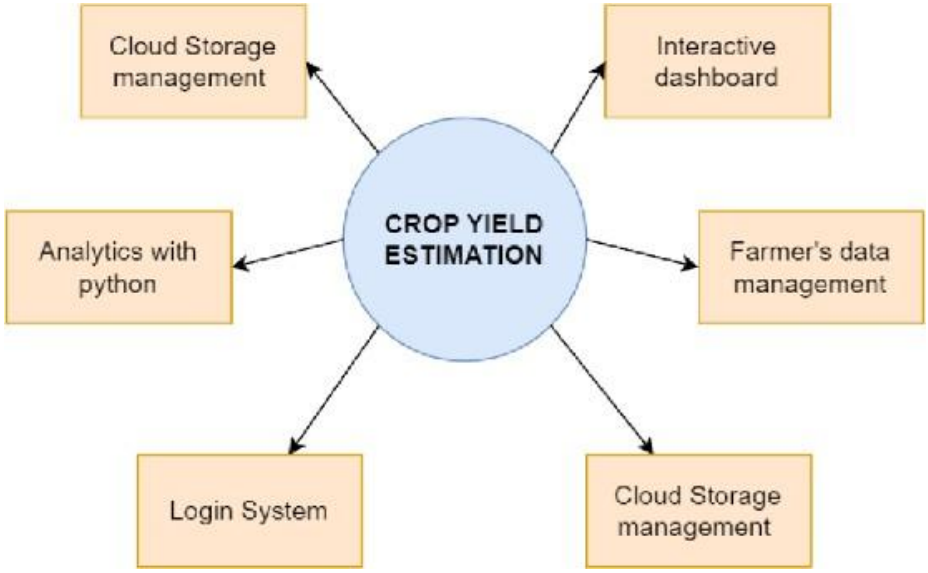
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Requirement	Knowledge of seeds ,crops ,mechanism ,soil ,climate & agriculture science. Right use of resources like soil and water. Time management .Market demand drive production.
FR-2	User Business rules	Three laws - the farmers produce trade and commerce, the farmers agreement of price assurance and farm services act and the essential commodities act
FR-3	User Factors	Crop prediction is highly sensitive to climate. It is affected by long-term trend in average rainfall and temperature, interannual climate variability, shocks during specific phenological stages and extreme weather events.
FR-4	Registration	Registering a new user through registration Crop yield estimates constitute a particular important productivity metric, both an aggregate level as well as in plot-level productivity analysis and impact evaluations of new technologies and policy interventions.
FR-5	User Objectives	Formulation and implementation of policies and programmes aimed at achieving rapid agricultural growth through optimum utilization of land, water, soil and plant resources of the state.
FR-6	User Improvement	It becomes necessary to increase the crop variety to produce disease-resistance off springs of the crops. It also helps in providing better and superior varieties based on the quality and quantity of the yield.
FR-7	Estimation and Analysis	Prediction of crop yield is made in estimation and the data is analysed to gain useful insights

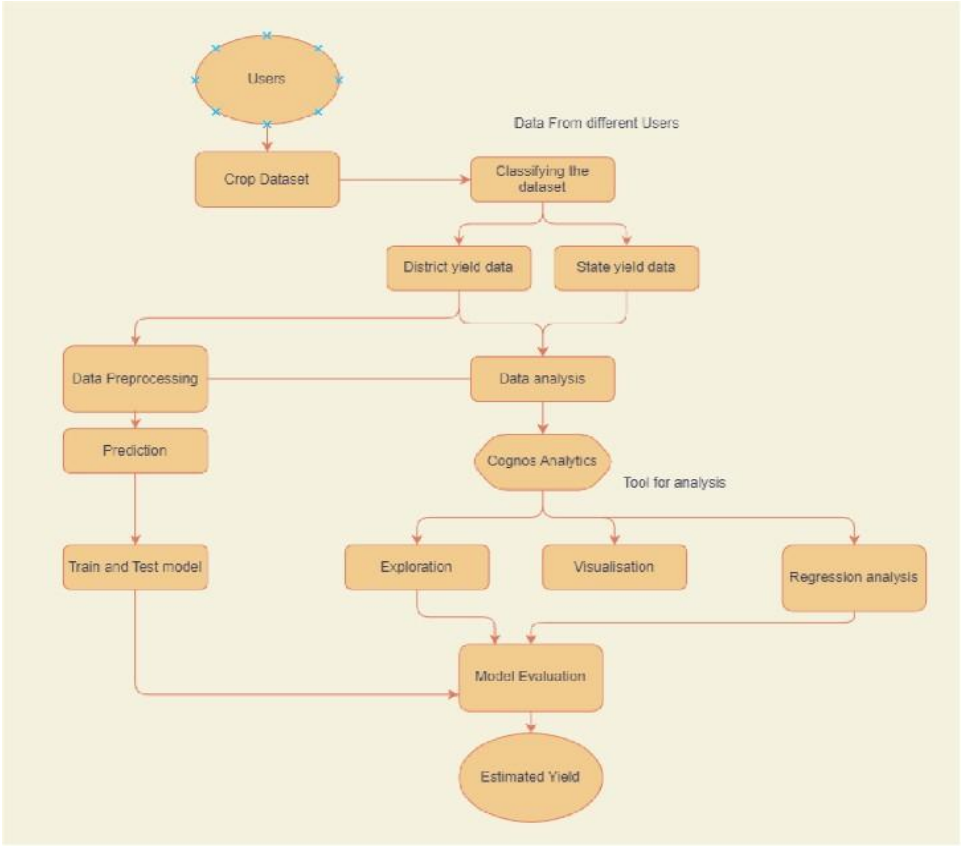
4.2 Non-Functional requirements:

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	To empower farmers and to increase the productivity there is need to provide the best dissemination tool for their farming activities.
NFR-2	Security	The developed ICT agriculture tools focus on very important agricultural services such as crop detection ,crop predictor will help farmers to make decision in future.
NFR-3	Reliability	This will remove multilingual issues and bridge the gap between farmers and technology.Effective tool that all farmers can use for management of all kind of crops
NFR-4	Performance	Multiple technologies and services that will improve the usability in agricultural activities.
NFR-5	Availability	Both website and mobile application interface and developed in local language and the content is available in localized language
NFR-6	Scalability	i)Increassed productivity from warm temperature ii)Decreased moisture stress iii)Possibility of growing new crops iv)Productivity of soil and water

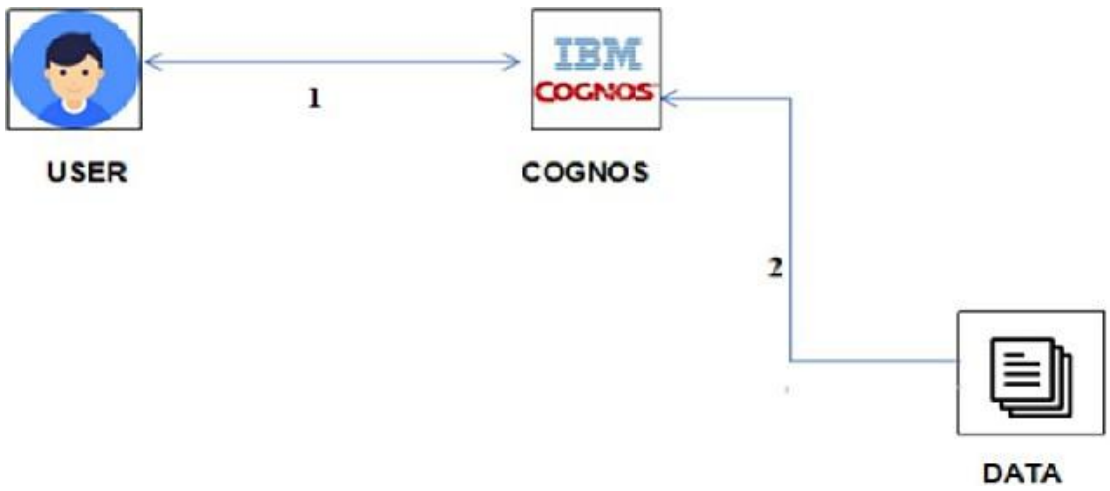
5. PROJECT DESIGN:

5.1 Data Flow Diagrams:





5.2 Solution & Technical Architecture:



5.3 User Stories:

User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user and Laptop users)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-6	Can use the methods provided in the dashboard		Medium	Sprint-1
	Invest	USN-7	With help of desired results obtained from application ,making profit or loss	Gain or Loss	High	Sprint-2
Administrator	Updating data		Collecting the data and storing it	Checking and updating dataset	High	Sprint-1
Customer (Web User)	Accessing the resources	USN -8	Using my own credentials for accessing the data	These resources cannot be accessed by others but only me	High	Sprint -1
	Satellite Visioning	USN -9	Having a view with geographic data		Medium	Sprint-2
Customer tools	Tools	USN -10	Analysis is performed by tools like cognos analytics	Ease of accessing the tools	High	Sprint 2

6. PROJECT PLANNING & SCHEDULING:

6.1 Sprint Planning & Estimation:

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	Avinash R
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	Charupriya S
Sprint-2		USN-3	As a user, I can register for the application through Facebook	2	High	Karthikeyan B
Sprint-1		USN-4	As a user, I can register for the application through Gmail	1	Medium	Vinoodhini D
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	Avinash R
Sprint-2	Dashboard	USN-6	Can use the methods provided in the dashboard	2	Medium	Karthikeyan B
Sprint-2		USN-7	With help of desired results obtained from application,making profit or loss and Collecting the data and storing it	1	High	Vinoodhini
Sprint 3		USN-8	Using my own credentials for accessing the data	1	High	Karthikeya
Sprint-3	Visualization	USN-9	Having a view with geographic data	2	High	Avinash R
Sprint 4		USN 10	Analysis is performed by tools like cognos analytics	1	High	Charupriya

6.2 Sprint Delivery Schedule:

Project Tracker, Velocity & Burndown Chart: (4 Marks)

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

6.3 Reports from JIRA

7.CODING & SOLUTIONING

6.4 Feature 1:

Exploratory Data Analysis:

```
df
```

	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production
0	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	1254.0	2000.0
1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2.0	1.0
2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	102.0	321.0
3	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	176.0	641.0
4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	720.0	165.0
...	...	...	...	...	...	...	...
246086	West Bengal	PURULIA	2014	Summer	Rice	306.0	801.0
246087	West Bengal	PURULIA	2014	Summer	Sesamum	627.0	463.0
246088	West Bengal	PURULIA	2014	Whole Year	Sugarcane	324.0	16250.0
246089	West Bengal	PURULIA	2014	Winter	Rice	279151.0	597899.0
246090	West Bengal	PURULIA	2014	Winter	Sesamum	175.0	88.0

246091 rows × 7 columns

```
[ ]  df.isnull().sum()
```

```
State_Name          0
District_Name       0
Crop_Year           0
Season              0
Crop                0
Area                0
Production       3730
dtype: int64
```

```
[ ]  df.dropna(inplace=True)
```

```
df.isnull().sum()
```

```
State_Name       0
District_Name    0
Crop_Year        0
Season           0
Crop             0
Area             0
Production       0
dtype: int64
```

```
df = df[df['Production'] !=0]
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 238838 entries, 0 to 246090
Data columns (total 7 columns):
 #   Column         Non-Null Count   Dtype
---  ------         --------------   -----
 0   State_Name     238838 non-null  object
 1   District_Name  238838 non-null  object
 2   Crop_Year      238838 non-null  int64
 3   Season         238838 non-null  object
 4   Crop           238838 non-null  object
 5   Area           238838 non-null  float64
 6   Production     238838 non-null  float64
dtypes: float64(2), int64(1), object(4)
memory usage: 14.6+ MB
```

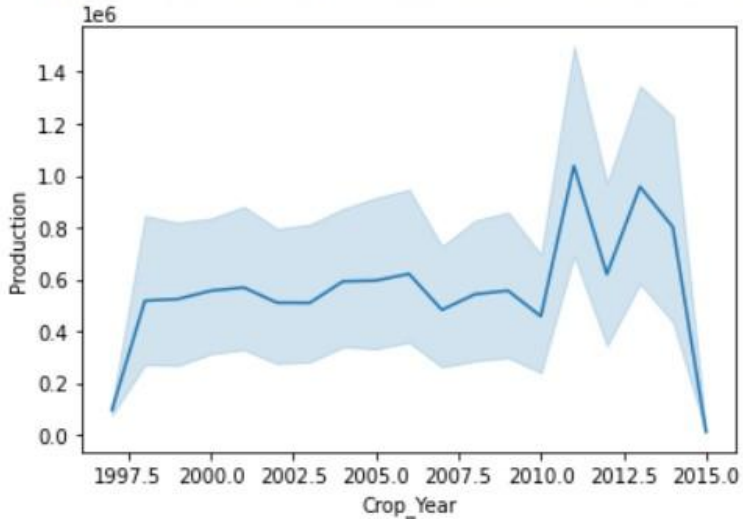
Data Visualisation

```
sns.lineplot(data["Crop_Year"],data["Production"])
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43:

Pass the following variables as keyword args: x, y. From version (

<matplotlib.axes._subplots.AxesSubplot at 0x7f43aeeeb950>
```

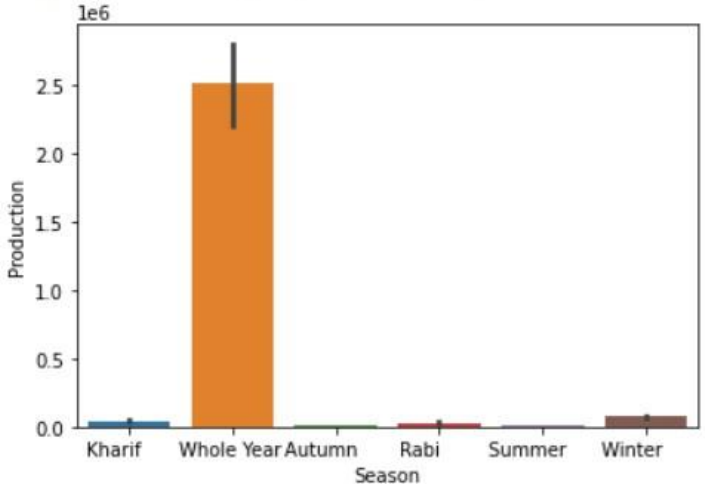


```
sns.barplot(data["Season"],data["Production"])
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:4

Pass the following variables as keyword args: x, y. From versio

<matplotlib.axes._subplots.AxesSubplot at 0x7f43ae7fe750>
```



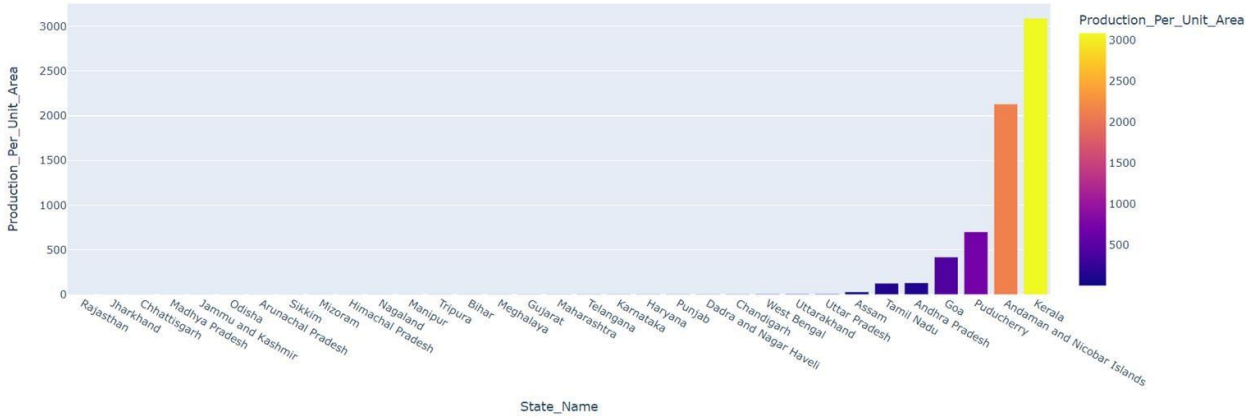

```
data.groupby("Season",axis=0).agg({"Production":np.sum})
```

Season	Production
Autumn	6.441377e+07
Kharif	4.029970e+09
Rabi	2.051688e+09
Summer	1.706579e+08
Whole Year	1.344248e+11
Winter	4.345498e+08

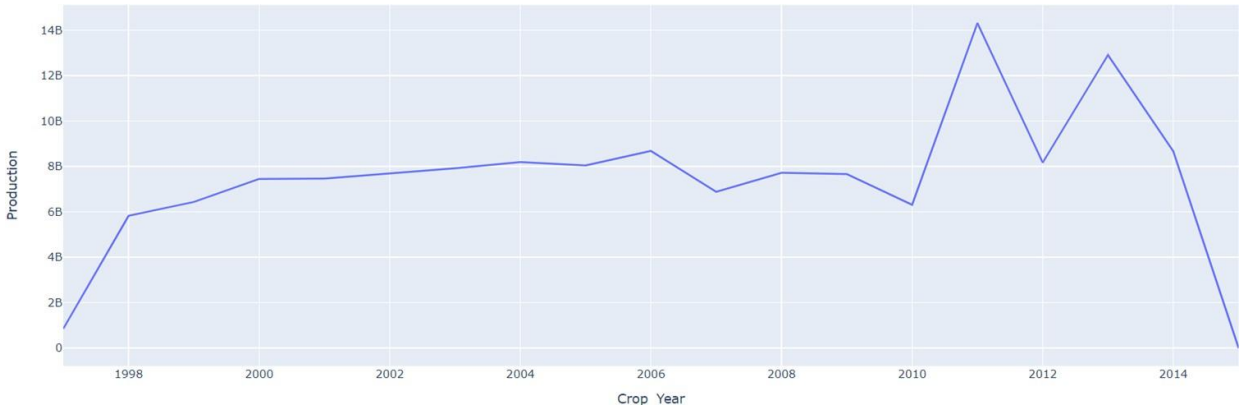
```
data["Crop"].value_counts()[:5]
```

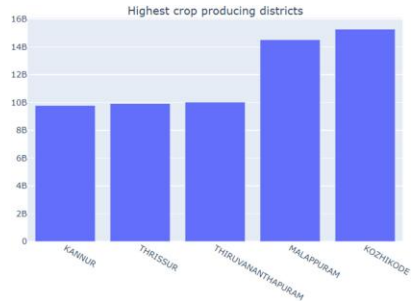
```
Rice                15078
Maize               13723
Moong(Green Gram)   10037
Urad                 9669
Groundnut            8732
Name: Crop, dtype: int64
```

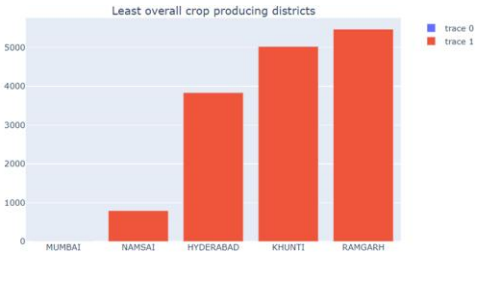
```
#PRODUCTIVITY OF DIFFERENT STATES
temp = data.groupby('State_Name')['Area', 'Production'].sum().reset_index()
temp['Production_Per_Unit_Area'] = temp['Production']/temp['Area']
temp = temp.sort_values(by='Production_Per_Unit_Area')
px.bar(temp, 'State_Name', 'Production_Per_Unit_Area', color='Production_Per_Unit_Area', )
```




```
#OVERALL PRODUCTION THROUGH YEARS
temp = data.groupby(by='Crop_Year')['Production'].sum().reset_index()
px.line(temp, 'Crop_Year', 'Production')
```







6.5 **Feature 2:**

Algorithm & Metric Evaluation:

SPLIT DATA INTO TRAINING AND TESTING SET

```
[ ]  X = f_data.drop("Production",axis=1)
     y = f_data[["Production"]]
```

```
[ ]  from sklearn.model_selection import train_test_split
     X_train,X_test,y_train,y_test = train_test_split(X,y,test_size=0.33, random_state=42)
```

```
[ ]  print("X_train :",X_train.shape)
     print("X_test :",X_test.shape)
     print("y_train :",y_train.shape)
     print("y_test :",y_test.shape)
```

```
X_train : (160021, 146)
X_test : (78817, 146)
y_train : (160021, 1)
y_test : (78817, 1)
```

MODEL

```
[ ]  from sklearn.linear_model import LinearRegression
     regression = LinearRegression()
     regression.fit(X_train, y_train)

     y_pred = regression.predict(X_test)
```

```
[ ]  from sklearn.metrics import mean_squared_error, r2_score
     mean_squared_error(y_test,y_pred)
     r2_score(y_test,y_pred)
```

```
1.0
```

MODEL PERFORMANCE METRICS

```
[ ] from sklearn.metrics import confusion_matrix, accuracy_score, precision_score, recall_score
```

```
[ ] cutoff = 0.7                              # decide on a cutoff limit
    y_pred_classes = np.zeros_like(y_pred)    # initialise a matrix full with zeros
    y_pred_classes[y_pred > cutoff] = 1
```

```
[ ] y_test_classes = np.zeros_like(y_pred)
    y_test_classes[y_test > cutoff] = 1
```

```
    # CONFUSION MATRIX
    conf_matrix=confusion_matrix(y_test_classes, y_pred_classes)
    conf_matrix
```

```
array([[  186,    10],
       [    0, 78621]])
```

```
[ ] # RECALL
```

```
1.0
```

```
TP = conf_matrix[1][1]
TN = conf_matrix[0][0]
FP = conf_matrix[0][1]
FN = conf_matrix[1][0]
print('True Positives:', TP)
print('True Negatives:', TN)
print('False Positives:', FP)
print('False Negatives:', FN)

# calculate accuracy
conf_accuracy = (float (TP+TN) / float(TP + TN + FP + FN))

# calculate the sensitivity
conf_sensitivity = (TP / float(TP + FN))

# calculate the specificity
conf_specificity = (TN / float(TN + FP))

# calculate precision
conf_precision = (TN / float(TN + FP))

# calculate recall
recall = recall_score(y_test_classes, y_pred_classes)

# calculate f_1 score
conf_f1 = 2 * ((conf_precision * conf_sensitivity) / (conf_precision + conf_sensitivity))

print('-'*50)
print(f'Accuracy: {round(conf_accuracy,2)}')
print(f'Sensitivity: {round(conf_sensitivity,2)}')
print(f'Specificity: {round(conf_specificity,2)}')
print(f'Precision: {round(conf_precision,2)}')
print('Recall: ',recall_score(y_test_classes, y_pred_classes))
print(f'f_1 Score: {round(conf_f1,2)}')
```

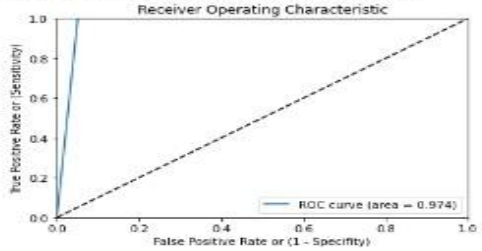
```
True Positives: 78621
True Negatives: 186
False Positives: 10
False Negatives: 0
--------------------------------------------------
Accuracy: 1.0
Sensitivity: 1.0
Specificity: 0.95
Precision: 0.95
Recall:  1.0
f_1 Score: 0.97
```

```
from sklearn.metrics import roc_curve
from sklearn.metrics import auc

fpr, tpr, thresholds = roc_curve(y_test_classes, y_pred_classes)
roc_auc = auc(fpr, tpr)


plt.plot(fpr, tpr, label='ROC curve (area = %0.3f)' % roc_auc)
plt.plot([0, 1], [0, 1], 'k--')  # random predictions curve
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.0])
plt.xlabel('False Positive Rate or (1 - Specifity)')
plt.ylabel('True Positive Rate or (Sensitivity)')
plt.title('Receiver Operating Characteristic')
plt.legend(loc="lower right")
```

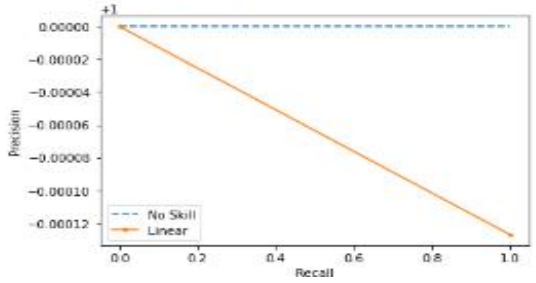
```
<matplotlib.legend.Legend at 0x7f43a921a590>
```



```
# precision-recall curve and f1
from sklearn.datasets import make_classification
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import precision_recall_curve
from sklearn.metrics import f1_score
from sklearn.metrics import auc
from matplotlib import pyplot

lr_precision, lr_recall, _ = precision_recall_curve(y_test_classes, y_pred_classes)
lr_f1, lr_auc = f1_score(y_test_classes, y_pred_classes), auc(lr_recall, lr_precision)
# summarize scores
print('Linear: f1=%.3f auc=%.3f' % (lr_f1, lr_auc))
# plot the precision-recall curves
no_skill = len(y_test[y_test==1]) / len(y_test)
pyplot.plot([0, 1], [no_skill, no_skill], linestyle='--', label='No Skill')
pyplot.plot(lr_recall, lr_precision, marker='.', label='Linear')
# axis labels
pyplot.xlabel('Recall')
pyplot.ylabel('Precision')
# show the legend
pyplot.legend()
# show the plot
pyplot.show()
```

```
Linear: f1=1.000 auc=1.000
```



7. TESTING

				Date	03-Nov-22							
				Team ID	PNT2022TMID21477							
				Project Name	Estimate the Crop Yield using							
				Maximum Marks	4 marks							
Test case ID	**Feature Type**	**Component**	**Test Scenario**	**Pre-Requisite**	**Steps To Execute**	**Test Data**	**Expected Result**	**Actual Result**	**Status**	**Comments**	**TC for Automation(Y/N**	**BUG ID**	**Executed By**
HomePage_TC_001	Functional	Home Page	Verify user is able to see the Login/Signup popup when user clicked on Login Button in the Homepage	Account to login or data to signup	1.Enter URL and click go 2.Click on Login Button 3.Verify login/Singup popup displayed or not	http://localhost:3000	Login/Signup popup should display	Working as expected	Pass		Y		Avinash.R
LoginPage_TC_002	Login Page	Home Page	Verify the UI elements in Login/Signup popup	Credentials to login	1.Enter URL and click go 2.Click on My Account dropdown button 3.Verify login/Singup popup with below UI elements: a.email text box b.password text box c.Login button d.New customer? Create account link	http://localhost:3000/login	Application should show below UI elements: a.email text box b.password text box c.Login button with orange colour d.New customer? Create account link e.Last password? Recovery password link	Working as expected	Pass		Y		Karthikeyan
LoginPage_TC_003	Functional	Login page	Verify user is able to log into application with Valid credentials	Account to login or data to signup	1.Enter URL(https://shopenzer.com/) and click go 2.Click on My Account dropdown button 3.Enter Valid username/email in Email text box 4.Enter valid password in password text box 5.Click on login button	Username: avinash@gmail.com password:Testing123	User should navigate to user account homepage	Working as expected	pass		Y		Vinoodhini D
				Maximum Marks	4 marks							
Test case ID	**Feature Type**	**Component**	**Test Scenario**	**Pre-Requisite**	**Steps To Execute**	**Test Data**	**Expected Result**	**Actual Result**	**Status**	**Comments**	**TC for Automation(Y/N**	**BUG ID**	**Executed By**
LoginPage_TC_004	Functional	Login page	Verify user is able to log into application with InValid credentials	Login credentials	1.Enter URL(https://shopenzer.com/) and click go 2.Click on My Account dropdown button 3.Enter Valid username/email in Email text box 4.Enter Invalid password in password text box 5.Click on login button	Username: avinash@gmail.com password: Testing123678686786876876	Application should show "Incorrect email or password " validation message.	Working as expected	Pass		Y		Avinash.R
LoginPage_TC_005	Functional	Login page	Verify user is able to log into application with InValid credentials	Account to login or data to signup	1.Enter URL(https://shopenzer.com/) and click go 2.Click on My Account dropdown button 3.Enter InValid username/email in Email text box 4.Enter Invalid password in password text box 5.Click on login button	Username: karthi@gmail.com password: Testing123678686786876876	Application should show "Incorrect email or password " validation message.	Working as expected	Pass		Y		Karthikeyan

7.1 Test Cases:

7.2 User Acceptance Testing:

2. Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	9	3	2	3	18
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	10	2	4	20	36
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	22	13	13	26	72

3. Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	7	0	0	7
Client Application	51	0	0	51
Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

8. RESULTS

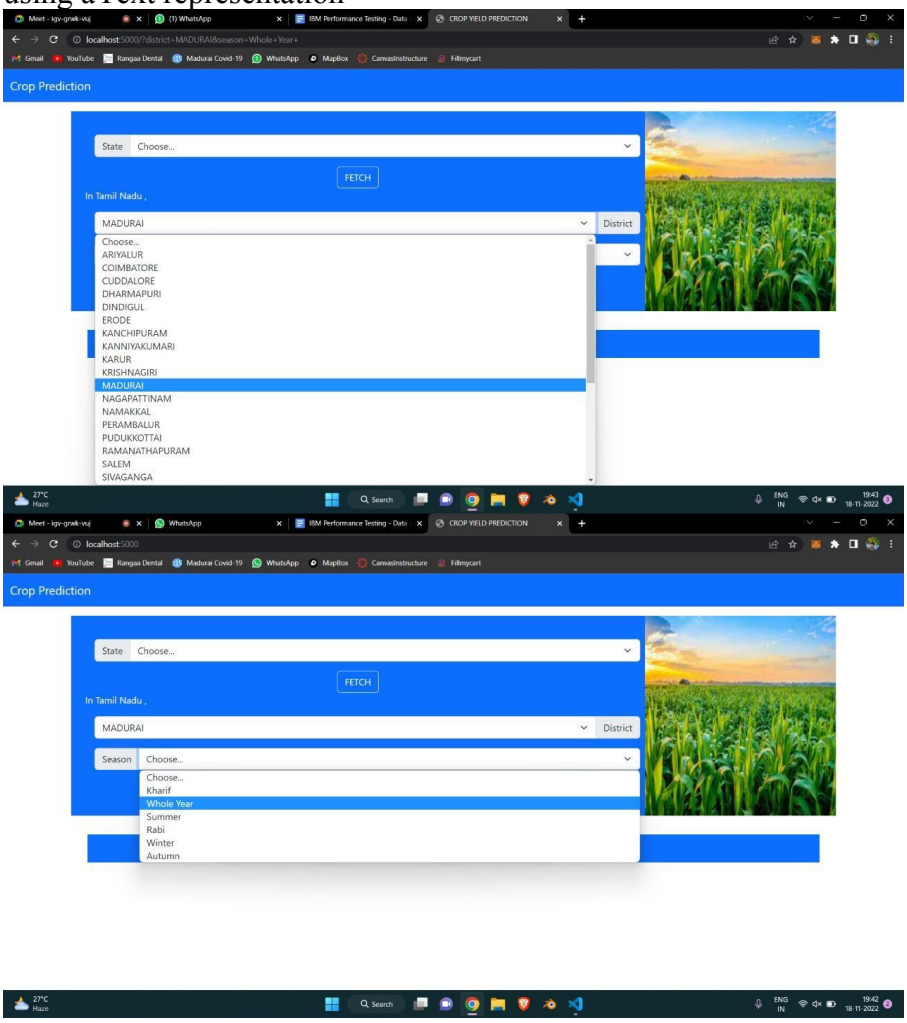
8.1 Performance Test:

Date	10 November 2022
Team ID	PNT2022TMID21477
Project Name	Estimate The Crop Yield Using Data Analytics
Maximum Marks	10 Marks

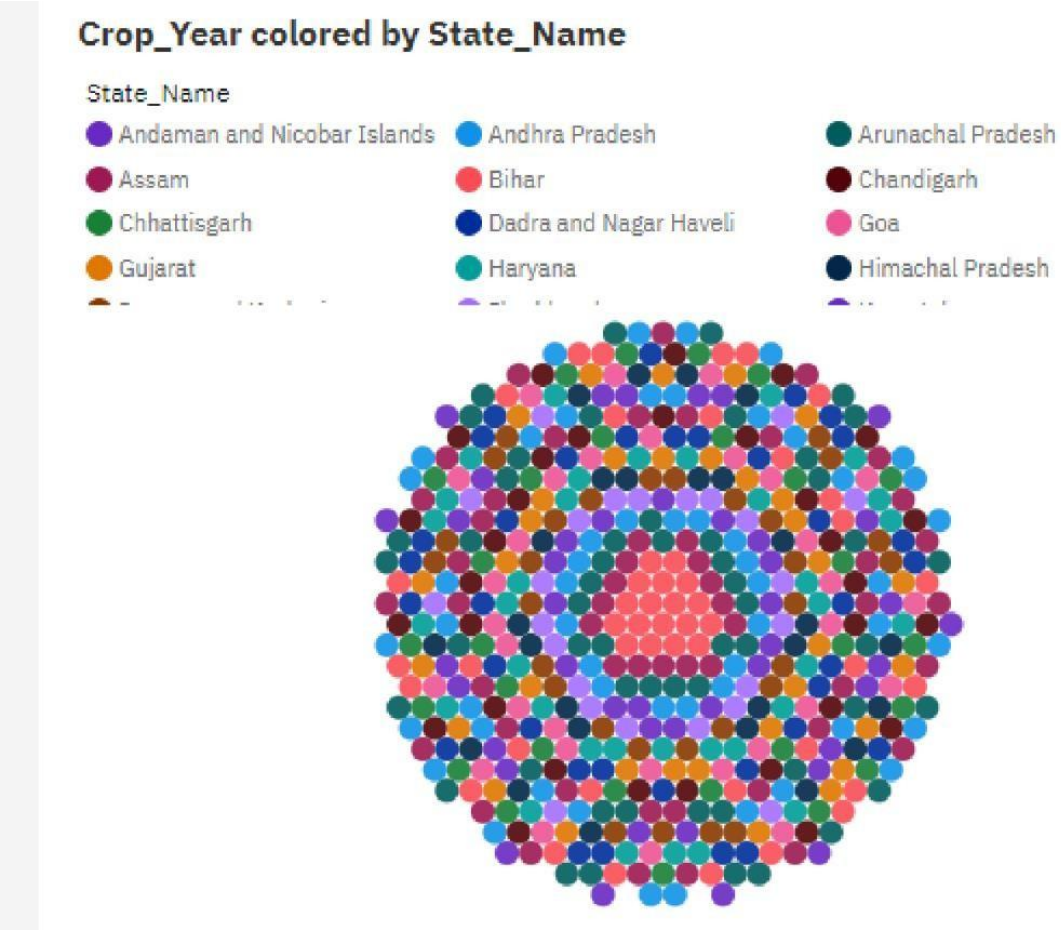
Model Performance Testing:

Project team shall fill the following information in the model performance testing template.

S.No.	Parameter	Screenshot / Values
1.	Dashboar ddesign	No of Visualizations / Graphs – 13
2.	Data Responsivene ss	Yes, the website is responsive completely,that is by resizing the browserwindow size as per the test scenario. **CROP PRODUCTION DATASET** The dataset contains 7 rows and 246091 record and dataset contains different state name, different district name, crop year ,crop, area, season and production
3.	Amount Datato Rendered (DB2 Metrics)	**To connect IBM Db2 database cloud with cognos analytics:** By using IBM Db2 to create Dashboard,Report,Story,Visualization andExploratory data analytics(EDA) , We have optimized the required data , and it fetches data based on the requirement.
4.	Utilization ofData Filters	Utilization of data filters - 12
5.	Effective UserStory	No of Scene Added – 8 To create the Registration page of the WebsiteTo create the Login page of the Website To create the Dashboard page of the Website To work on the given dataset,Understand the Dataset Load the dataset to Cloud platform then Build the required VisualizationsUsing the Crop production in Indian dataset,create various graphs and charts to highlight the insights and visualizations. Build a Visualizations to showcase Average Crop Production by Seasons Showcase the Yearly usage of Area in Crop Production

6.	Descriptiv eReports	No of Visualizations / Graphs – 5 Visualization1 - Average Crop Production by Seasons Visualization2 - Yearly usage of area in crop production Visualization3 - Top 10 States in Crop Yield Production by Area Visualization 4 - Crop Production by State Visualization5 - Represent the States with Seasonal Crop Production using aText representation 

VISUALISATIONS:



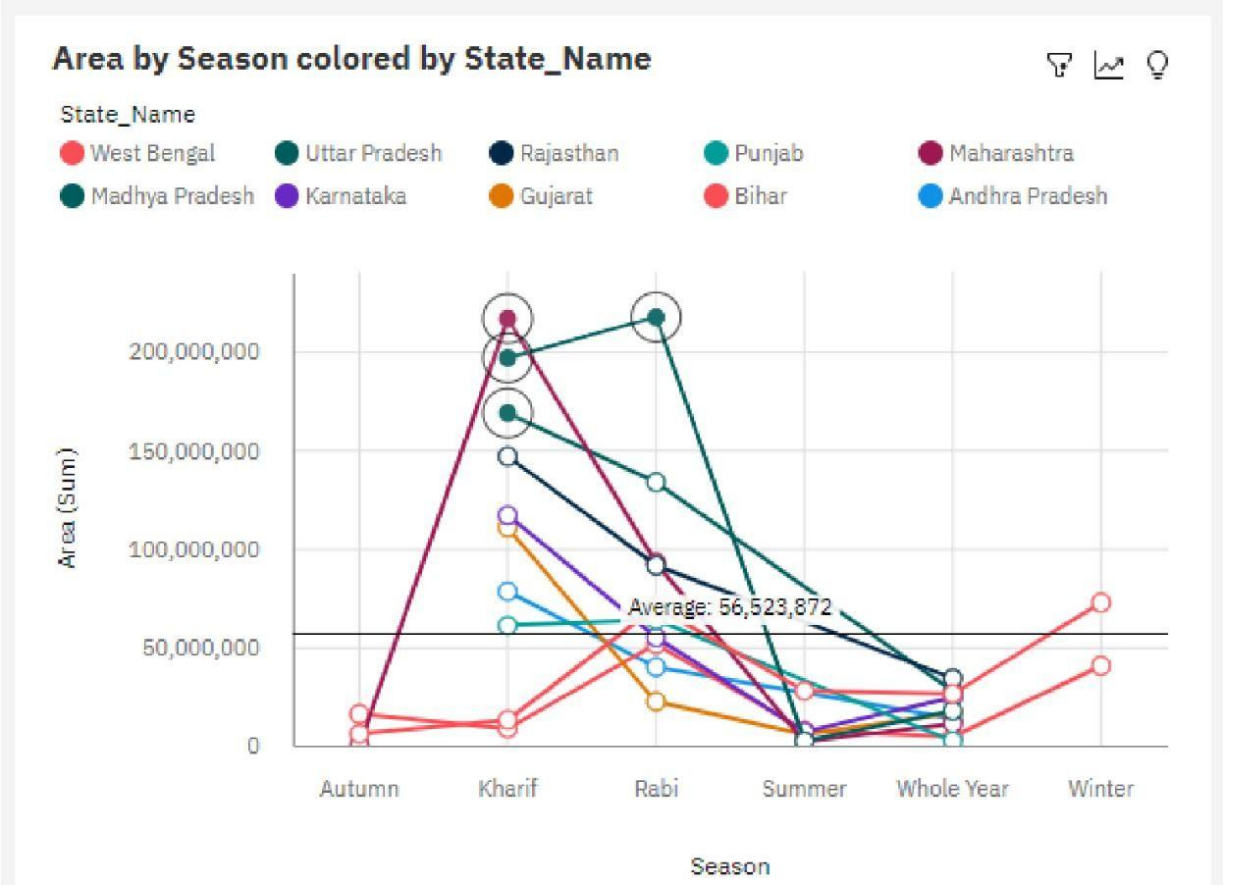
District_Name for State_Name and Crop_Year

District_Name	Andaman and Ni...	Andhra Pradesh	Arunachal Pradesh	Assam
1997	(no value)	13	13	
1998	(no value)	13	13	
1999	(no value)	13	13	
2000	2	13	13	
2001	2	13	13	
2002	2	13	15	
2003	2	13	15	
2004	2	13	16	
2005	2	13	16	
2006				

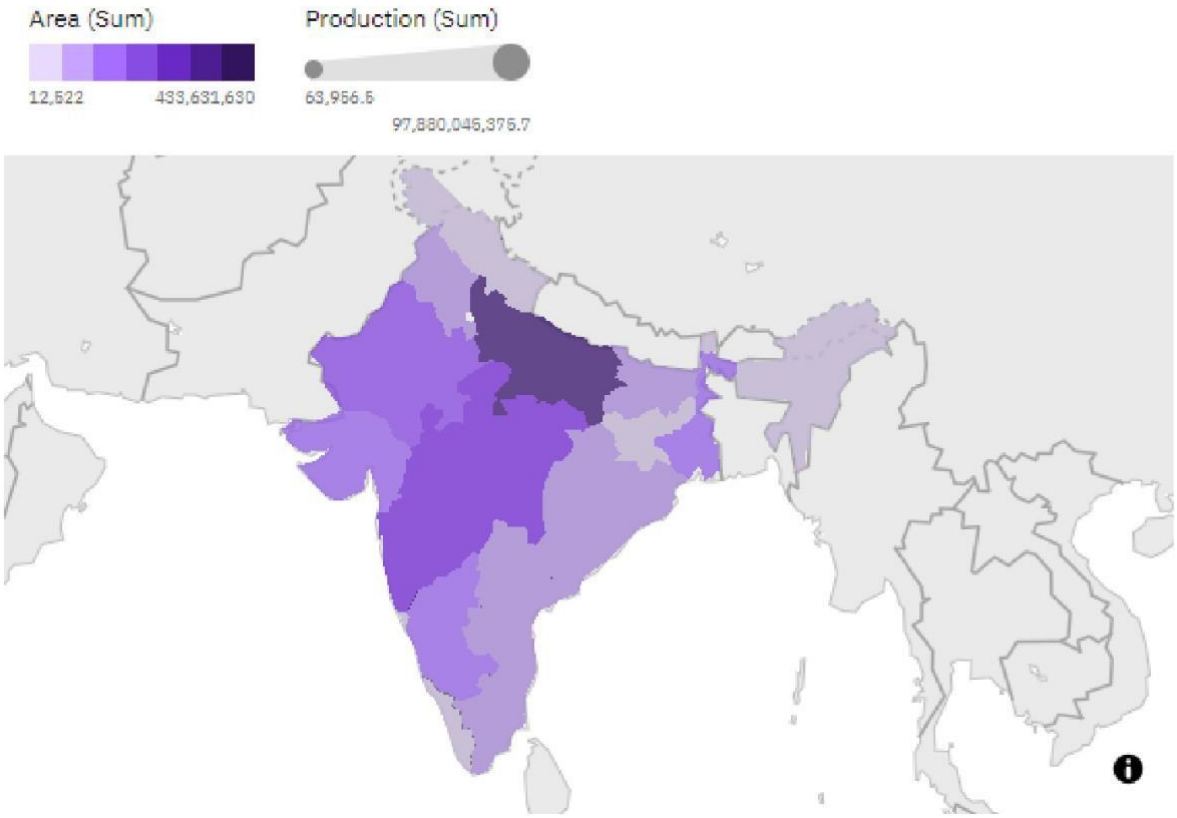

District_Name colored by Season
Season
Autumn
Kharif
Rabi
Summer
Whole Year
Winter

Season and Crop

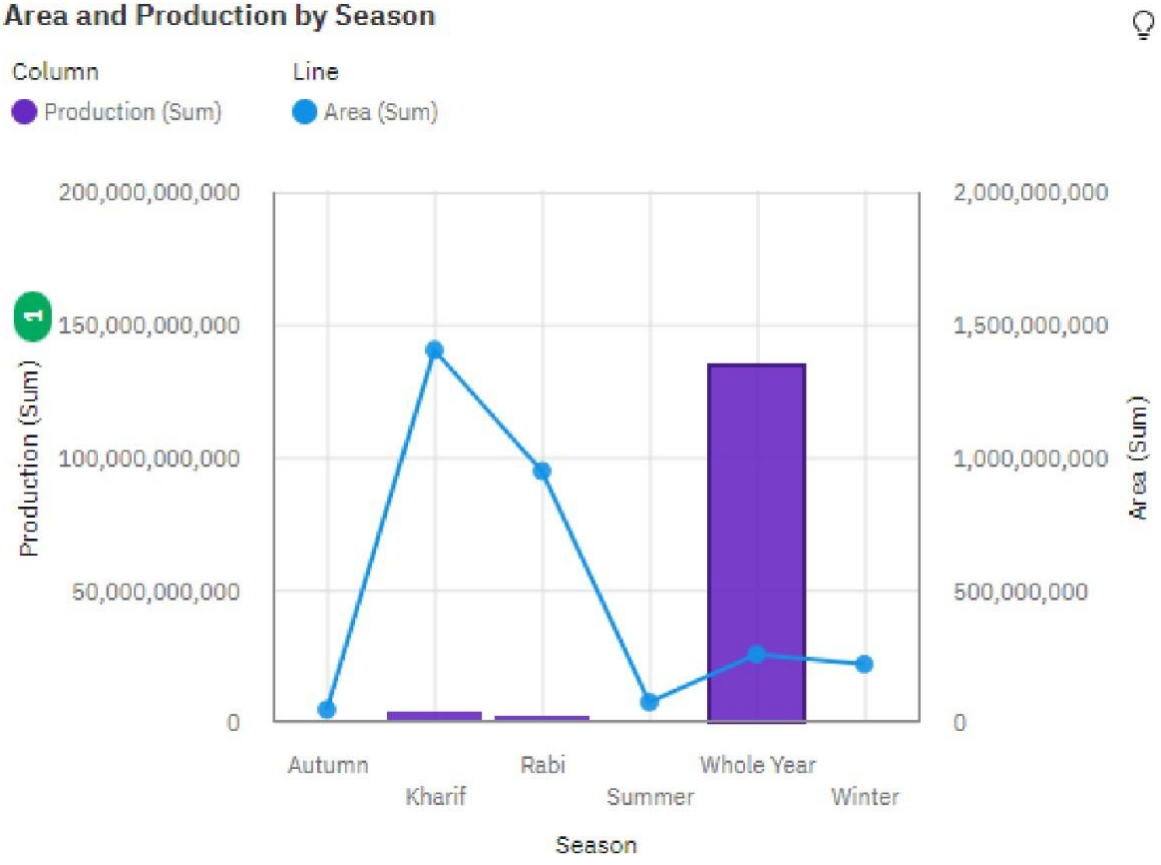
Crop	Season
Apple	Whole Year
Arcanut (Processed)	Whole Year
Arecanut	Kharif
	Rabi
	Whole Year
Arhar/Tur	Autumn
	Kharif
	Rabi
	Summer
	Whole Year
	Winter

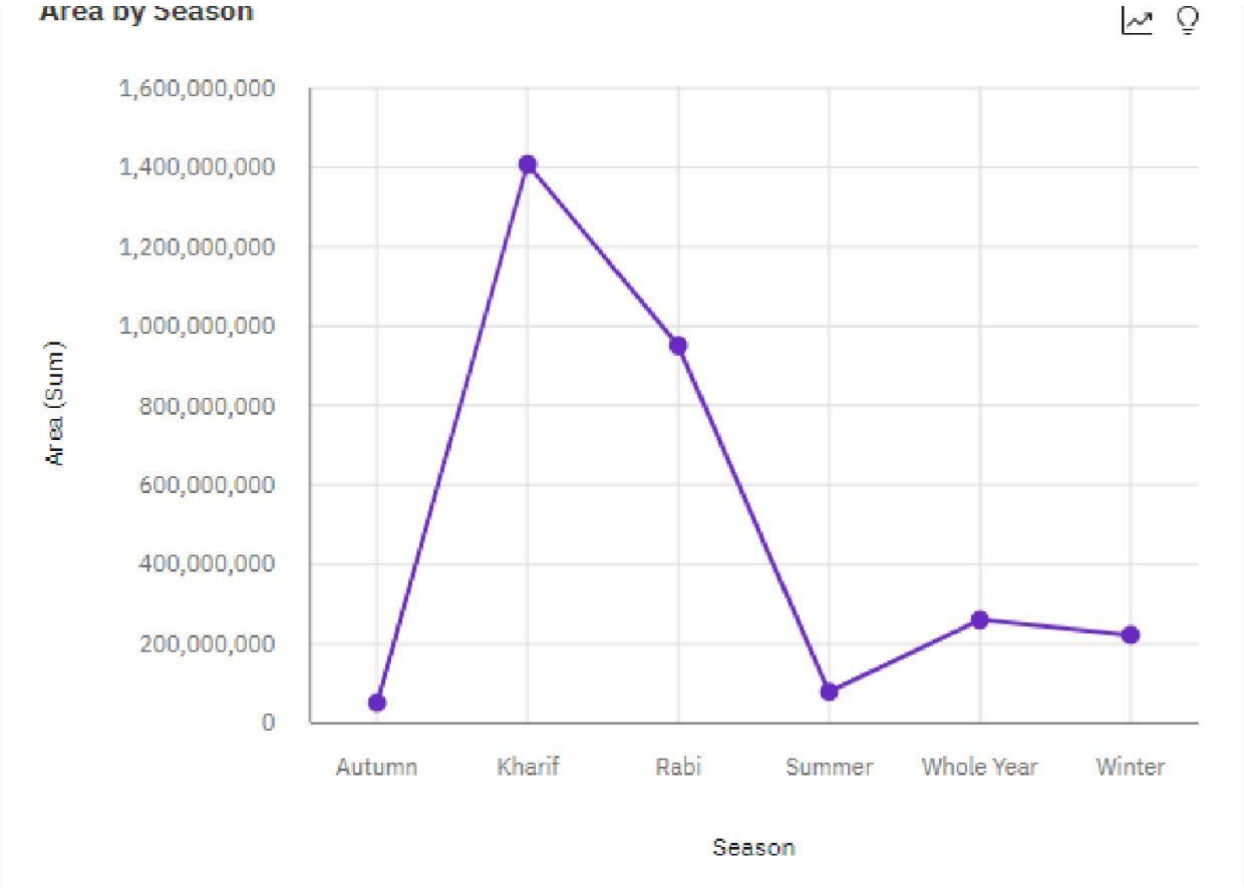

Area by Season colored by State_Name
State_Name
West Bengal
Uttar Pradesh
Rajasthan
Punjab
Maharashtra
Madhya Pradesh
Karnataka
Gujarat
Bihar
Andhra Pradesh
200,000,000
150,000,000
100,000,000
50,000,000
0
Area (Sum)
Average: 56,523,872
Autumn
Kharif
Rabi
Summer
Whole Year
Winter
Season

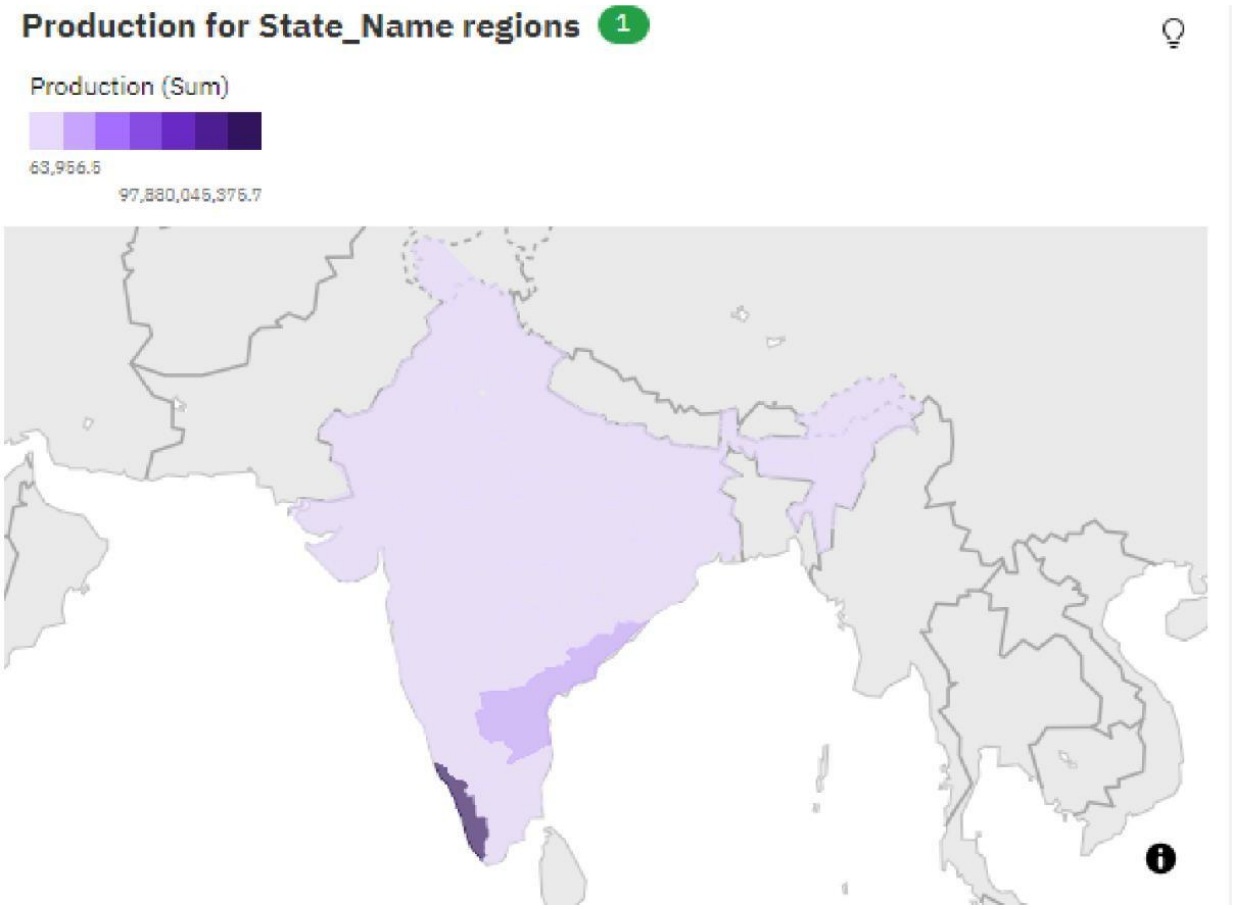
Area and Production for State_Name regions



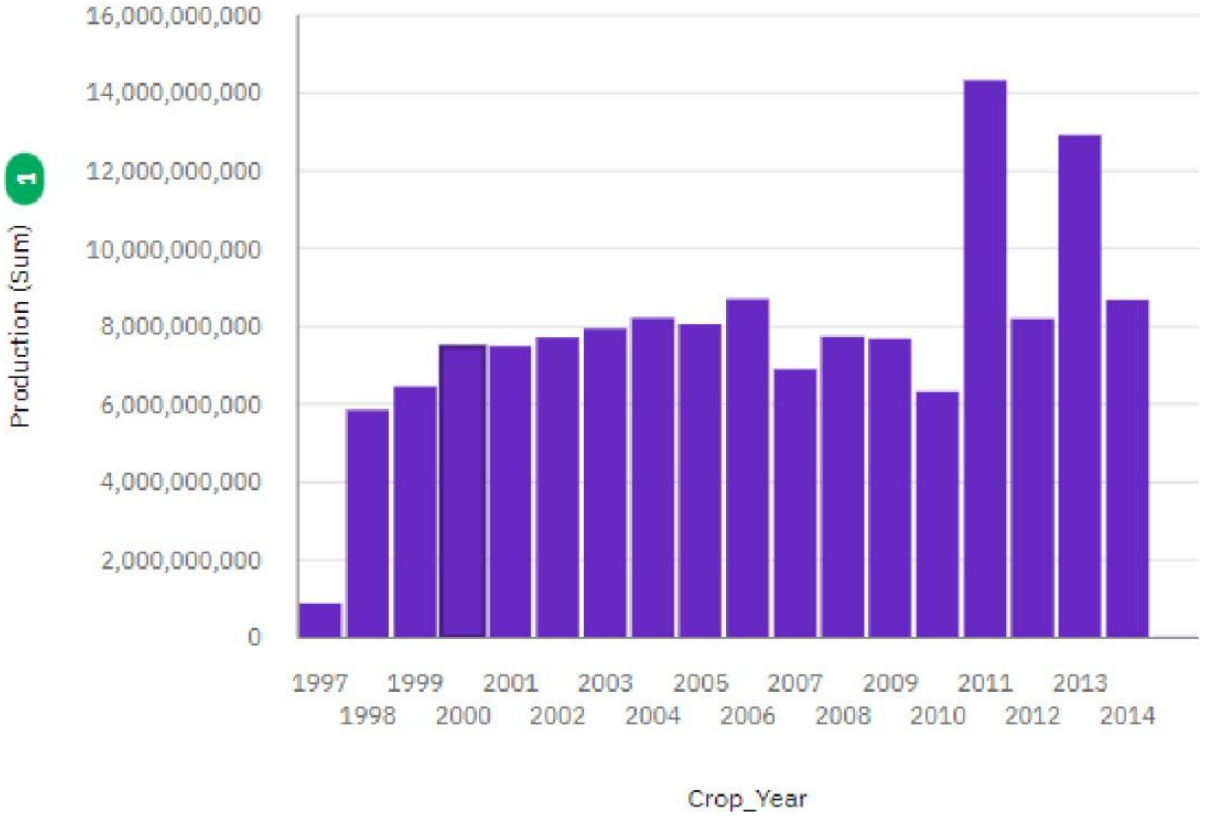
Area and Production by Season



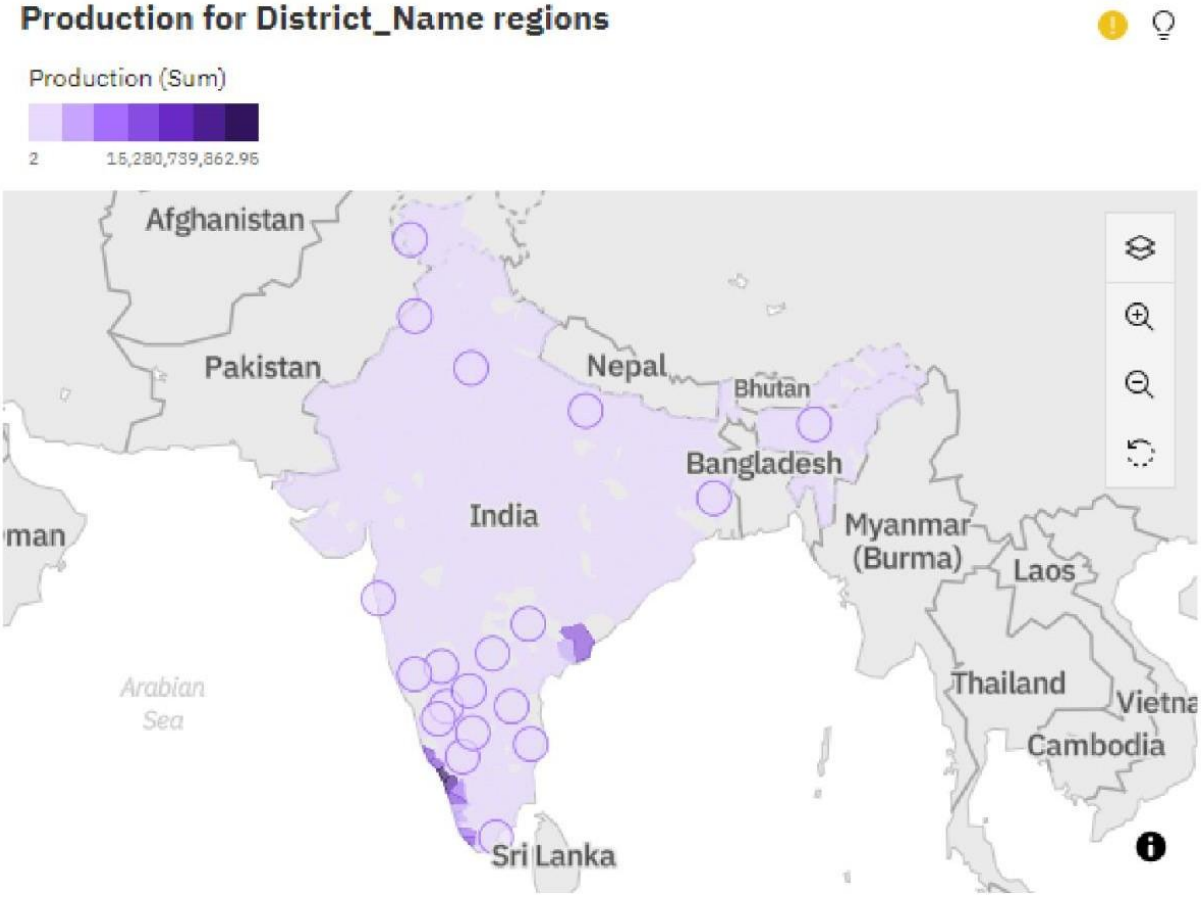

Area by Season
1,600,000,000
1,400,000,000
1,200,000,000
1,000,000,000
800,000,000
600,000,000
400,000,000
200,000,000
0
Area (Sum)
Autumn
Kharif
Rabi
Summer
Whole Year
Winter
Season


Production for State_Name regions
1
Production (Sum)
63,956.5
97,880,045,376.7

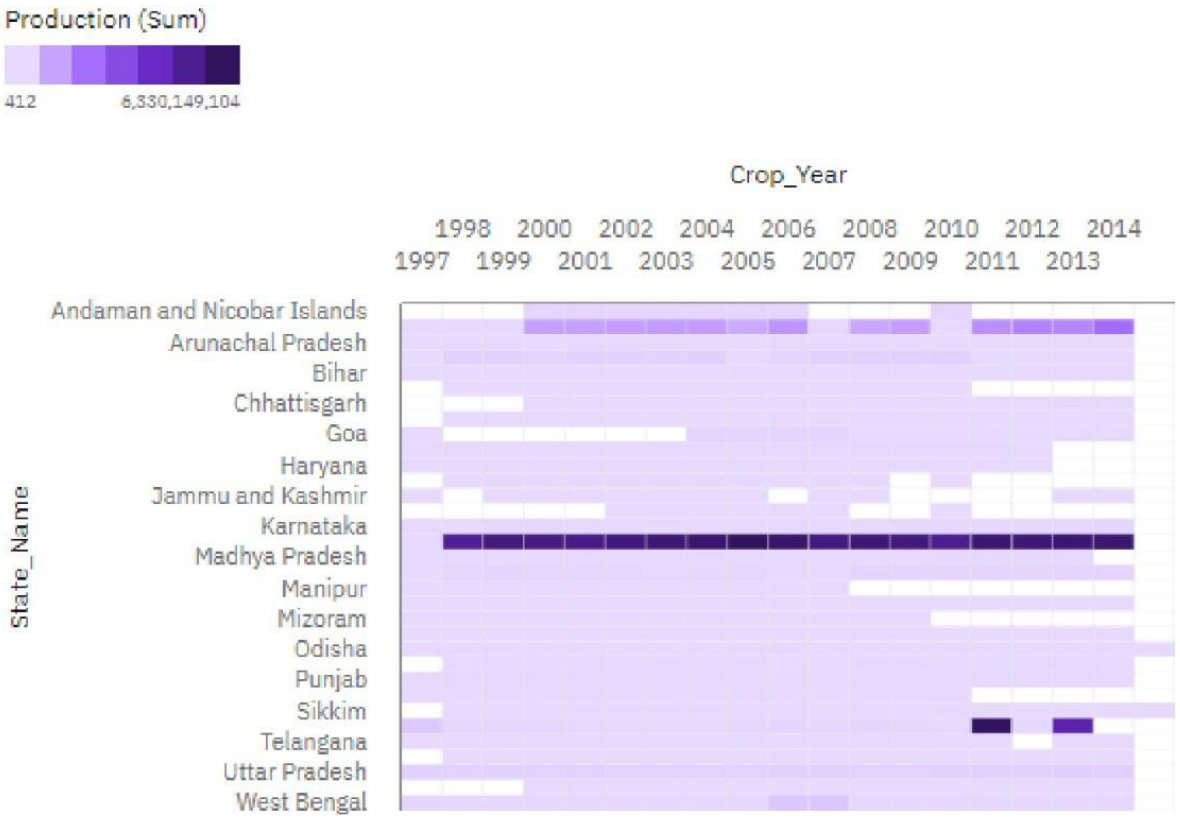
Production by Crop_Year

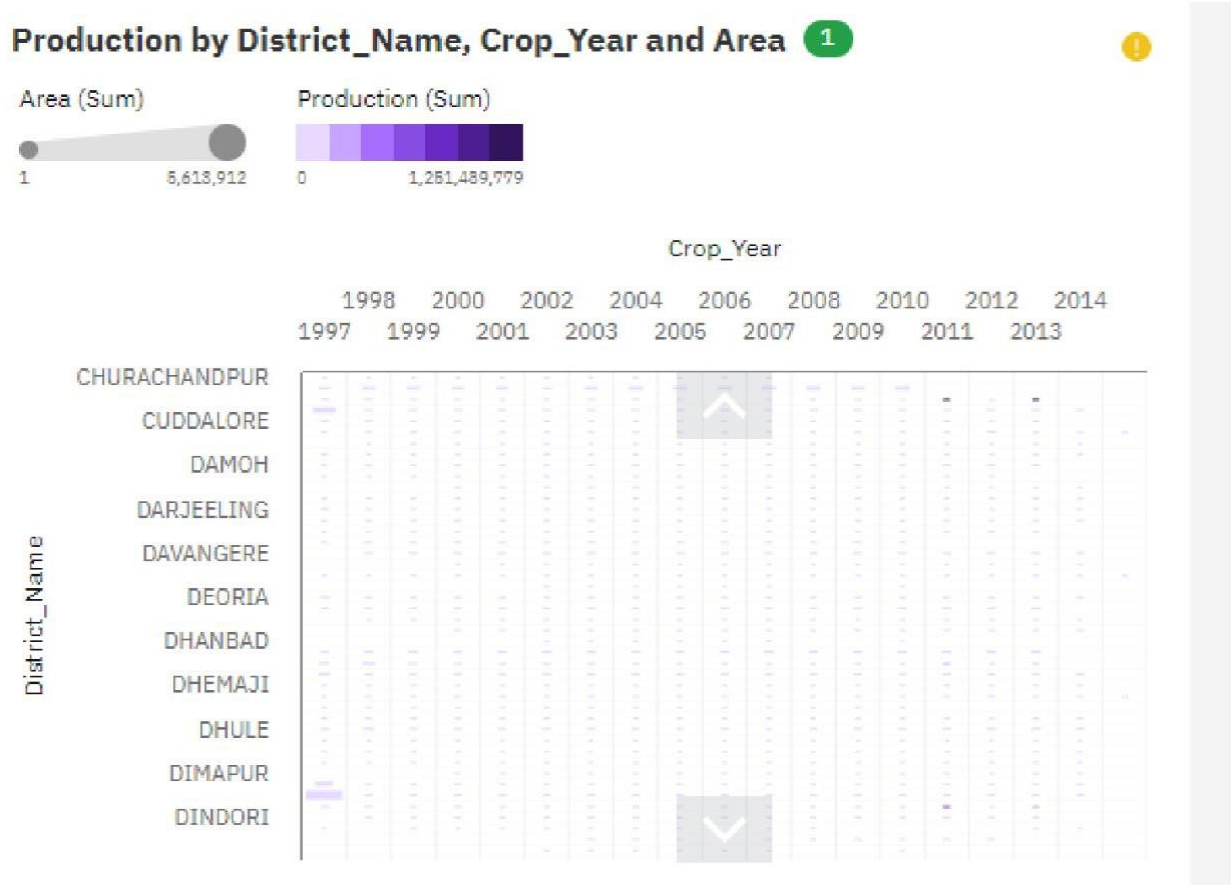


Production for District_Name regions

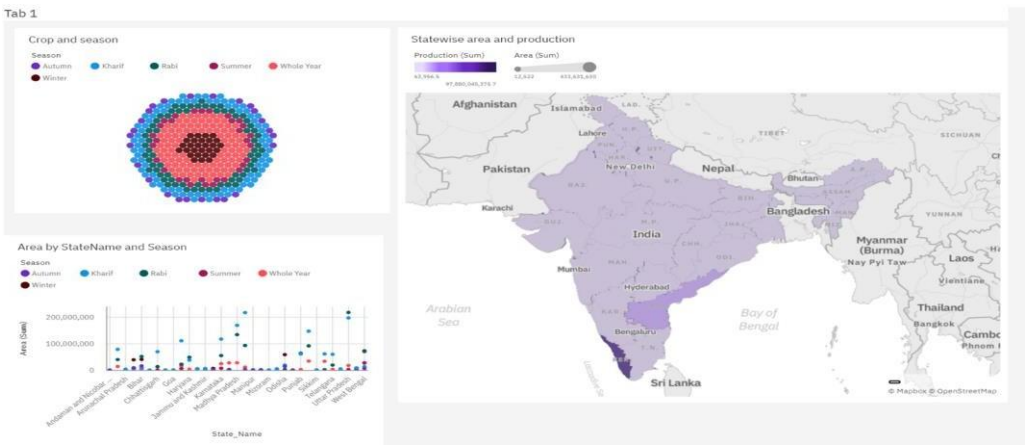


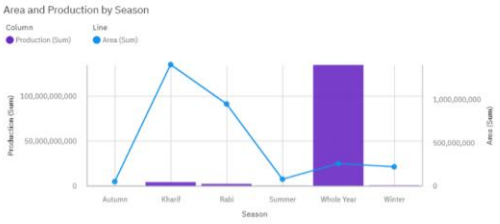
Production by State_Name and Crop_Year 1



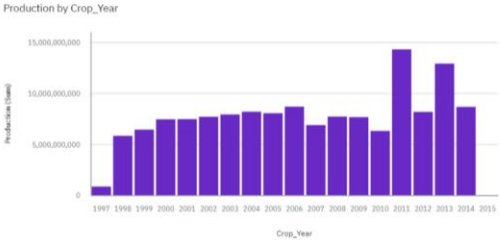

Production by District_Name, Crop_Year and Area
1
Area (Sum)
1
5,618,912
Production (Sum)
0
1,251,489,779
Crop_Year
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
District_Name
CHURACHANDPUR
CUDDALORE
DAMOH
DARJEELING
DAVANGERE
DEORIA
DHANBAD
DHEMAJI
DHULE
DIMAPUR
DINDORI

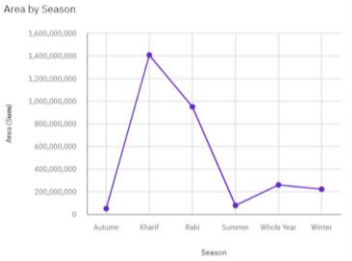
DASHBOARD:


Tab 1
Crop and season
Season
Autumn
Kharif
Rabi
Summer
Whole Year
Winter
Statewise area and production
Production (Sum)
Area (Sum)
Afghanistan
Islamabad
Lahore
Pakistan
New Delhi
Nepal
Bhutan
Karachi
Bangladesh
India
Myanmar (Burma)
Nay Pyi Taw
Laos
Vientiane
Mumbai
Hyderabad
Thailand
Bangkok
Arabian Sea
Bay of Bengal
Bengaluru
Sri Lanka
Area by StateName and Season
Area (Sum)
State_Name


Area and Production by Season
Column
Production (Sum)
Line
Area (Sum)
100,000,000,000
50,000,000,000
0
1,000,000,000
500,000,000
0
Autumn
Kharif
Rabi
Summer
Whole Year
Winter
Season

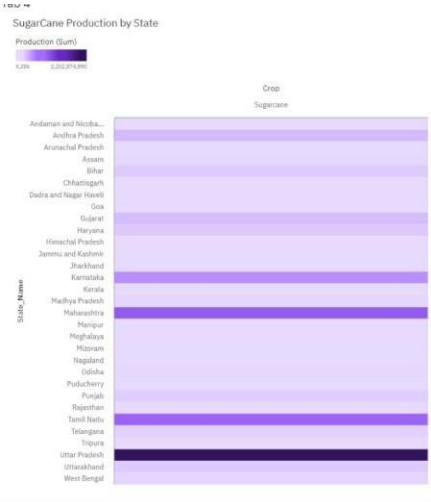

Production for State_Name regions
Production (Sum)
© Mapbox © OpenStreetMap

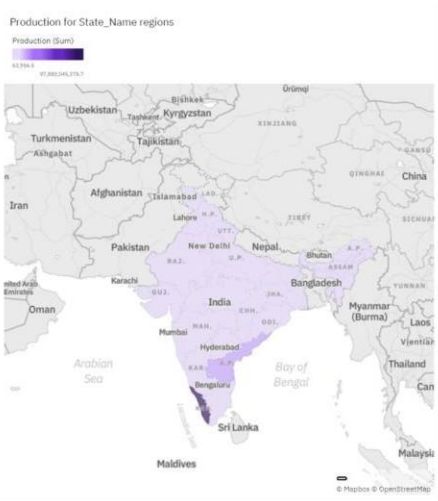

Production by Crop_Year
15,000,000,000
10,000,000,000
5,000,000,000
0
1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
Crop_Year

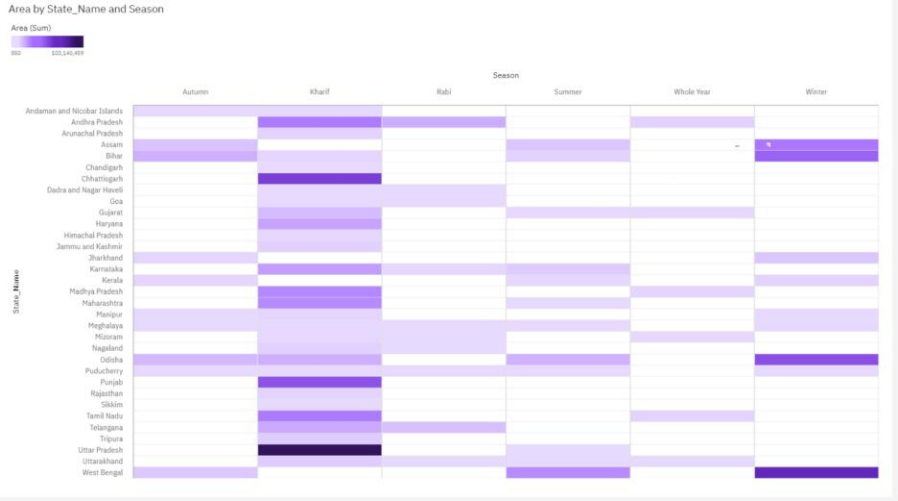

Area by Season
1,600,000,000
1,400,000,000
1,200,000,000
1,000,000,000
800,000,000
600,000,000
400,000,000
200,000,000
0
Autumn
Kharif
Rabi
Summer
Whole Year
Winter
Season


District wise production
Production (Sum)
Islamabad
Lahore
New Delhi
Nepal
Bhutan
Bangladesh
India
Mumbai
Hyderabad
Bengaluru
Sri Lanka
Bay of Bengal
© Mapbox © OpenStreetMap

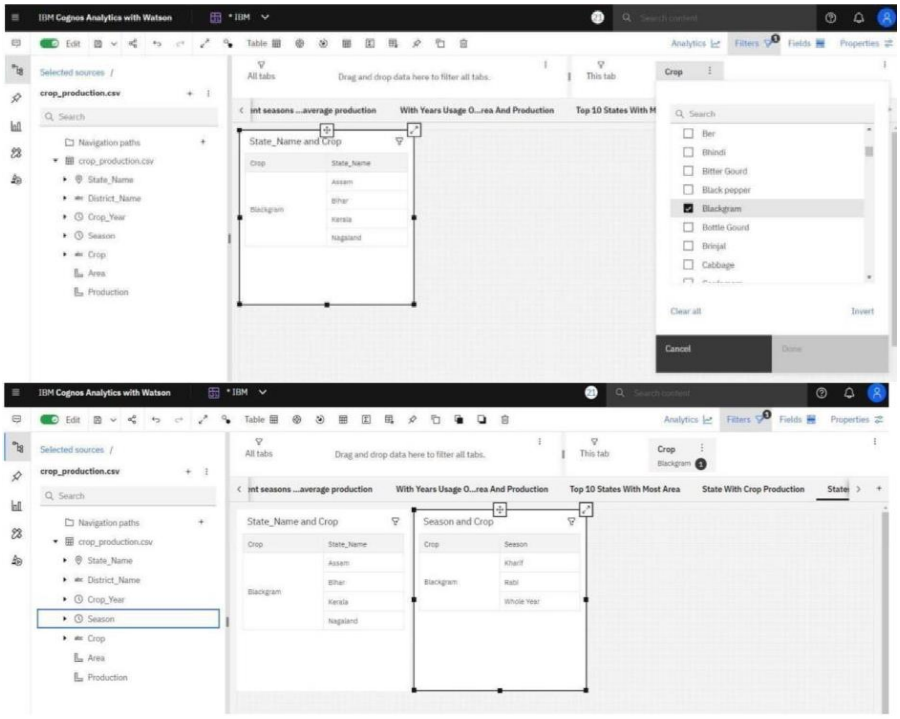

Production by State_Name and Crop_Year
Production (Sum)
Crop_Year
1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015
State_Name
Andaman and Nicoba...
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra and Nagar Haveli
Goa
Gujarat
Haryana
Himachal Pradesh
Jammu and Kashmir
Jharkhand
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Odisha
Puducherry
Punjab
Rajasthan
Sikkim
Tamil Nadu
Telangana
Tripura
Uttar Pradesh
Uttarakhand
West Bengal


SugarCane Production by State
Production (Sum)
Crop
Sugarcane
State_Name
Andaman and Nicoba...
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chhattisgarh
Dadra and Nagar Haveli
Goa
Gujarat
Haryana
Himachal Pradesh
Jammu and Kashmir
Jharkhand
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Odisha
Puducherry
Punjab
Rajasthan
Tamil Nadu
Telangana
Tripura
Uttar Pradesh
Uttarakhand
West Bengal


Production for State_Name regions
Production (Sum)
Ürümqi
Bishkek
Uzbekistan
Kyrgyzstan
Tashkent
Turkmenistan
Tajikistan
Ashgabat
Afghanistan
Islamabad
Lahore
Iran
China
Pakistan
New Delhi
Nepal
Bhutan
Karachi
Bangladesh
India
Myanmar (Burma)
Laos
Oman
Mumbai
Hyderabad
Bengaluru
Arabian Sea
Bay of Bengal
Thailand
Sri Lanka
Maldives
Malaysia
© Mapbox © OpenStreetMap



STATES WITH CROP PRODUCTION ALONG WITH THE SEASON:



9. ADVANTAGES & DISADVANTAGES

Advantages

- **People can easily identifies the crop choice for cultivation depends on the location,season and so on**
- **Crop production increases due to right choice of crop during cultivation**
- **It gives profit for the people who chose the seasonal crop in a particular locationDisadvantages**
- **Accuracy may be differ due to some unavoidable conditions**
- **Not at all the features that are impact the production be taken**

10.CONCLUSION

With the help of technology, people can can cultivate and harvest seasonal crops and get better yield with more accuracy. Hence the crop yield production can be increased and better understand of the nature of the soil for crop cultivation. The following steps can bedone for the estimation of crop yield production. Previous year data can be collected and get exploratory data analysis. After removing the null values and outliers, data visualisation is analysed and get insights for the future work

11.FUTURE SCOPE

In the future work, can implement with IOT devices as edge layer like temperature sensor, humidity sensor and so on. It is used to collect the real-time data and understand the realcondition of the soil and moisture level. So, people can easily find the better crop choicedepends on conditions and make better accuracy

12.APPENDIX

- **Sourcecode: https://github.com/IBM-EPBL/IBM-Project-47176-1660796949/tree/main/Final%20Deliverables**
- **Github link - https://github.com/IBM-EPBL/IBM-Project-47176-**
- **1660796949**

  Project demo link - https://github.com/IBM-EPBL/IBM-Project-47176-1660796949/tree/main/Final%20Deliverables/project%20demo