ESTIMATION OF CROP YIELD USING DATA ANALYTICS

Domain: Data Analytics
Team Id: PNT2022TMID34790

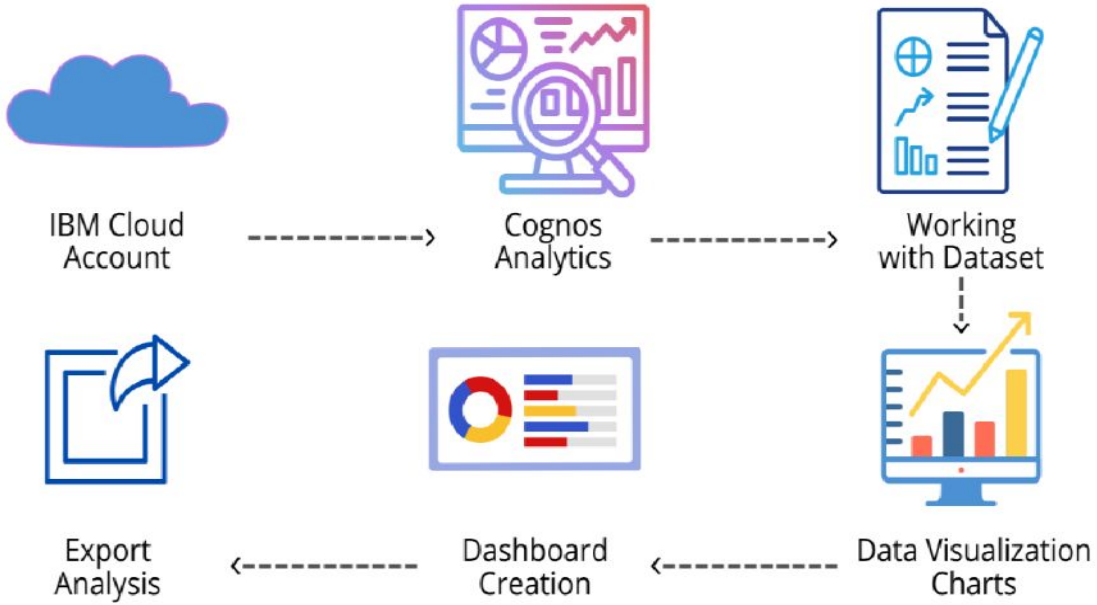
INTRODUCTION

- In agriculture, the yield is a measurement of the amount of a crop grown, or product such as wool, meat or milk produced, per unit area of land.
- Weather plays an important role in Agriculture Production.
- A detailed knowledge of rainfall regime at a place is an important prerequisite.
- Soil fertility refers to the inherent capacity of soil to supply nutrients in adequate amount and in suitable proportion for crop growth and crop yield.

OBJECTIVES

- ➔ To prepare the fields for sowing of crops with adequate availability of seed zone.
- ➔ To contribute optimal crop growth,development and yield.
- ➔ To predict appropriate crop yield using data analytics and machine learning.

PROJECT FLOW CHART



PROPOSED METHODOLOGY

- Increase in knowledge of agriculture in terms of crop production from sowing the seeds till harvesting. It gives an overview on the total crop yield before even sowing the seed from the previous datasets.

- Comparison of the subsequent crop yield predictions to be made with the entire set of existing available data and can be dedicated to suitable approaches for improving the efficiency of the proposed technique.

FUNCTIONAL REQUIREMENTS

User Registration

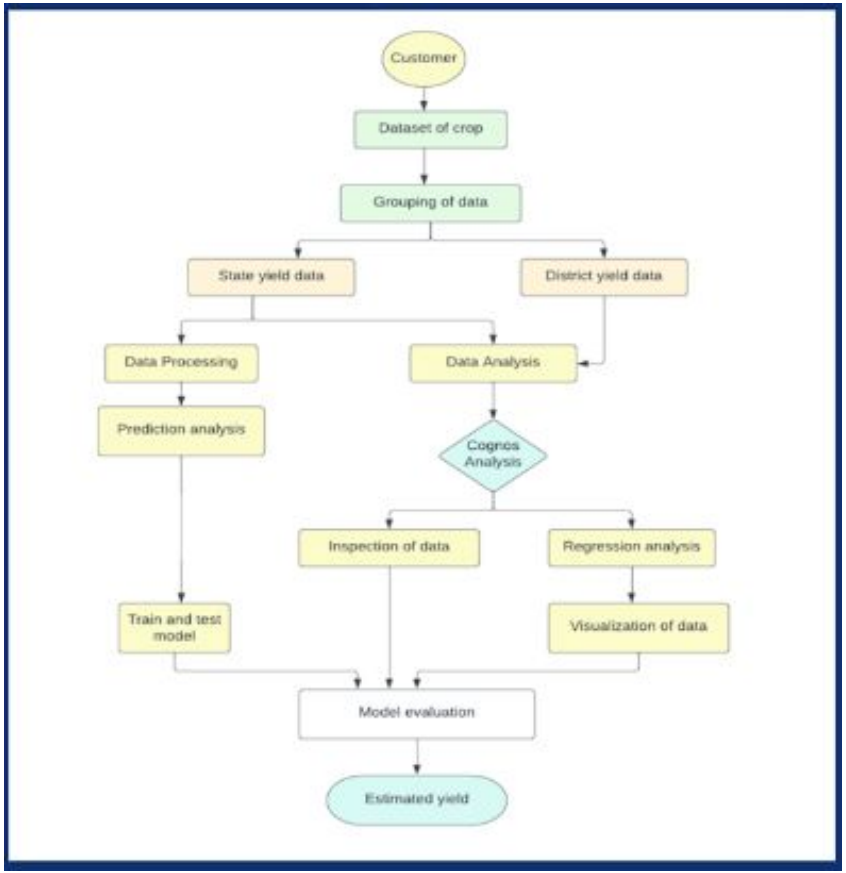
User Confirmation

Data collection

Data processing

Data visualization

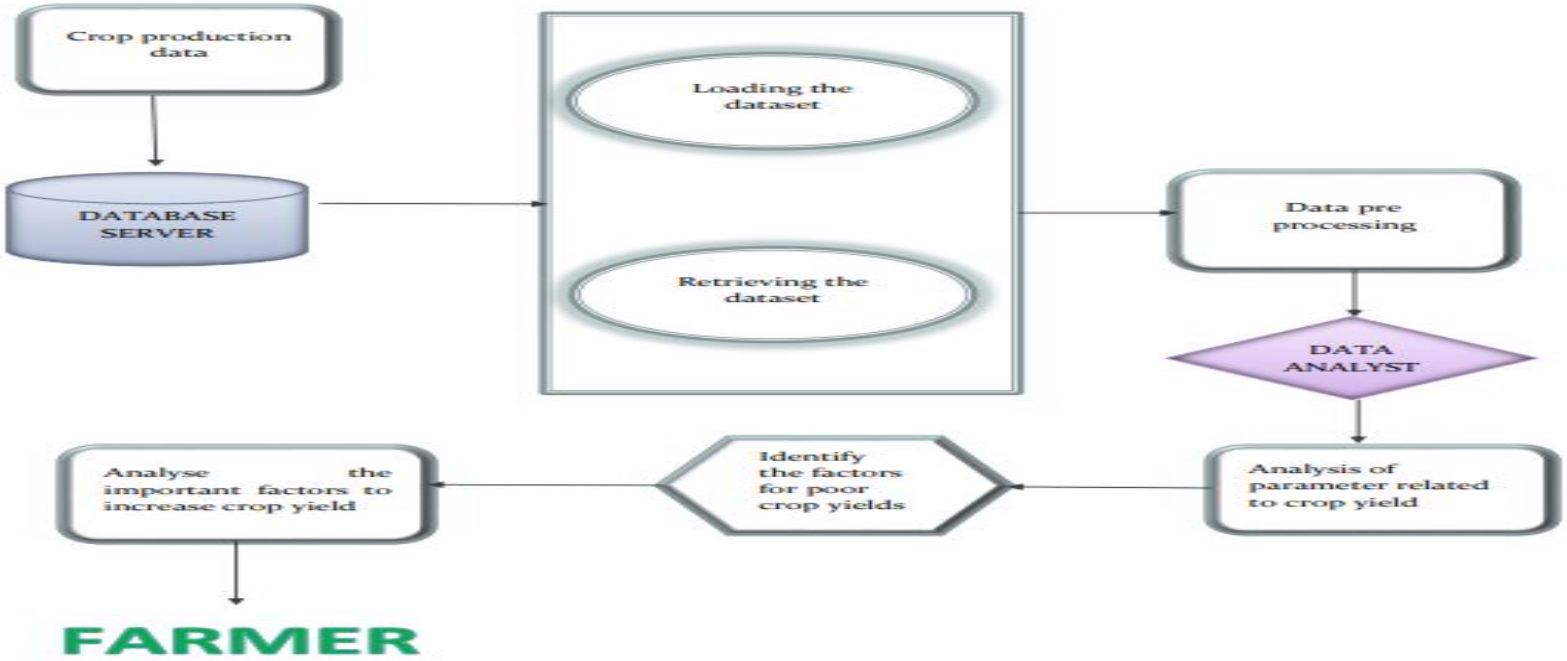
FLOWCHART


Customer
Dataset of crop
Grouping of data
State yield data
District yield data
Data Processing
Data Analysis
Prediction analysis
Cognos Analysis
Inspection of data
Regression analysis
Train and test model
Visualization of data
Model evaluation
Estimated yield

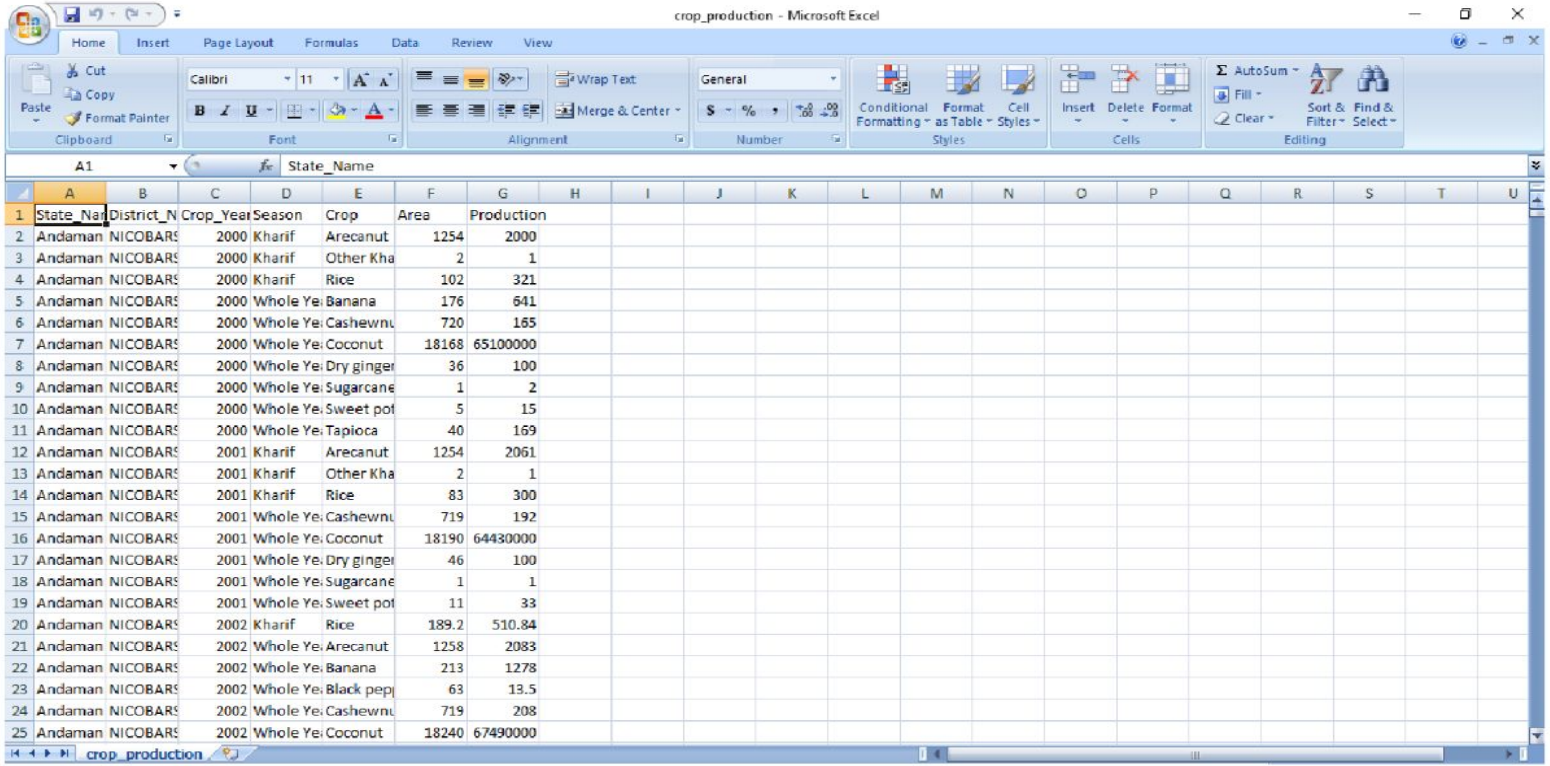
PROJECT DESIGN

- It is completely user friendly. As it is designed to provide support for agriculturalists and farmers, the user experience is simple and efficient.
- Data analytics help the farmers to predict the market conditions, climatic changes, factor-in inflation that will help them plan the entire process even before sowing the seeds.
- Can be made available in all platforms with minimum constraints and a stable speed network connection.

SOLUTION ARCHITECTURE

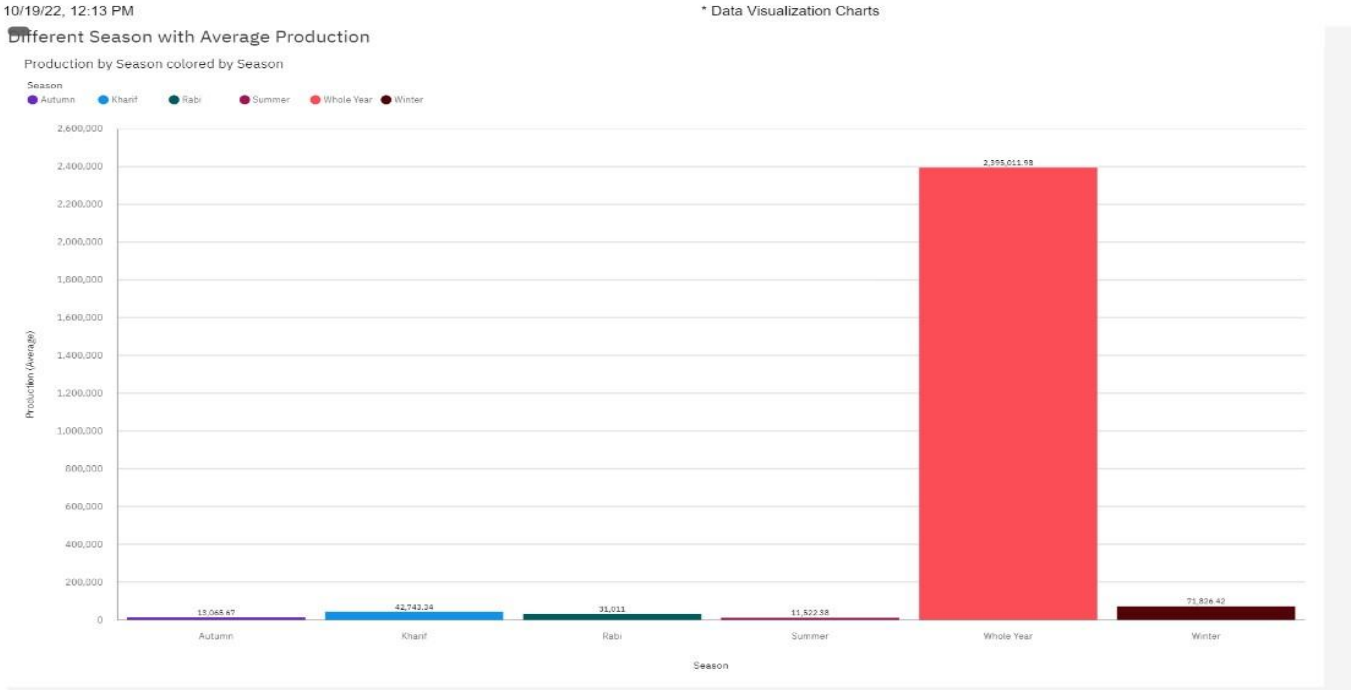


DATASET OF CROPS

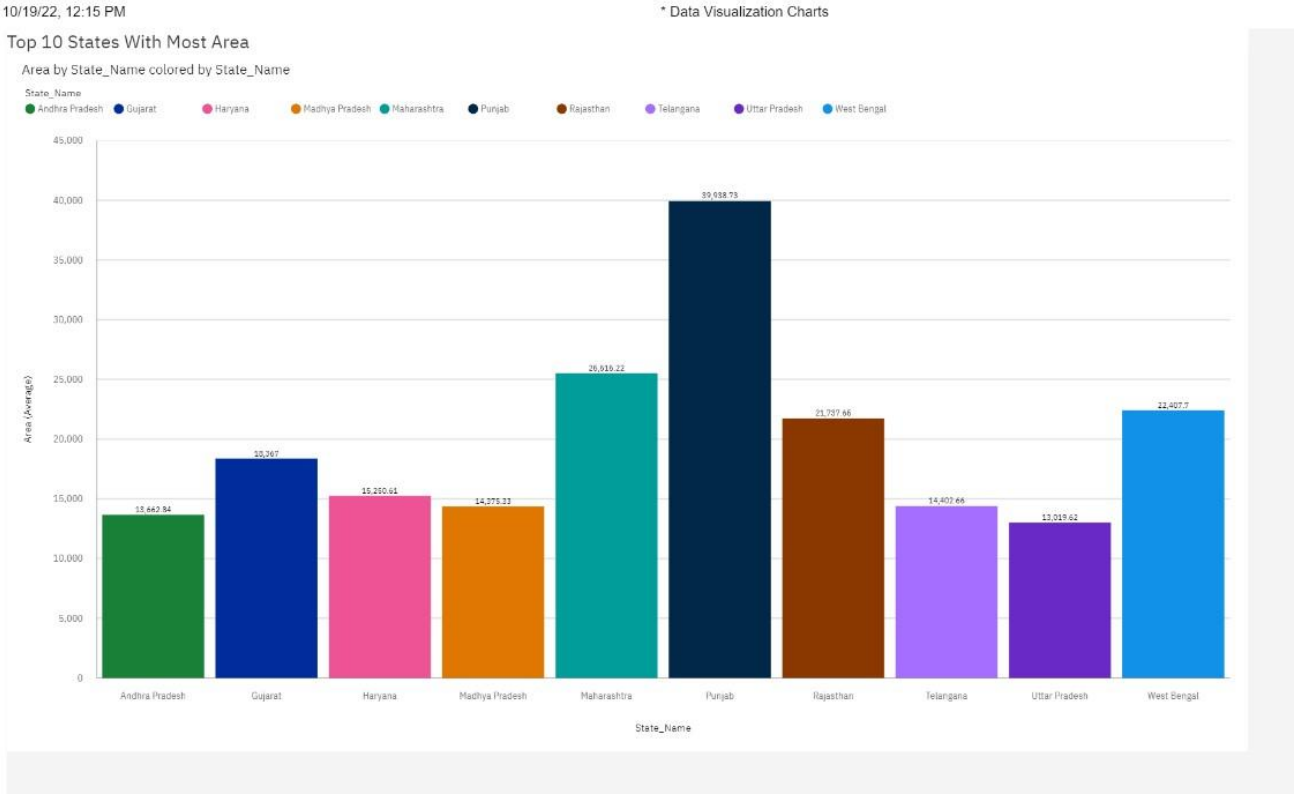


	A	B	C	D	E	F	G
1	State_Nat	District_N	Crop_Year	Season	Crop	Area	Production
2	Andaman	NICOBARS	2000	Kharif	Arecanut	1254	2000
3	Andaman	NICOBARS	2000	Kharif	Other Kha	2	1
4	Andaman	NICOBARS	2000	Kharif	Rice	102	321
5	Andaman	NICOBARS	2000	Whole Ye	Banana	176	641
6	Andaman	NICOBARS	2000	Whole Ye	Cashewnu	720	165
7	Andaman	NICOBARS	2000	Whole Ye	Coconut	18168	65100000
8	Andaman	NICOBARS	2000	Whole Ye	Dry ginger	36	100
9	Andaman	NICOBARS	2000	Whole Ye	Sugarcane	1	2
10	Andaman	NICOBARS	2000	Whole Ye	Sweet pot	5	15
11	Andaman	NICOBARS	2000	Whole Ye	Tapioca	40	169
12	Andaman	NICOBARS	2001	Kharif	Arecanut	1254	2061
13	Andaman	NICOBARS	2001	Kharif	Other Kha	2	1
14	Andaman	NICOBARS	2001	Kharif	Rice	83	300
15	Andaman	NICOBARS	2001	Whole Ye	Cashewnu	719	192
16	Andaman	NICOBARS	2001	Whole Ye	Coconut	18190	64430000
17	Andaman	NICOBARS	2001	Whole Ye	Dry ginger	46	100
18	Andaman	NICOBARS	2001	Whole Ye	Sugarcane	1	1
19	Andaman	NICOBARS	2001	Whole Ye	Sweet pot	11	33
20	Andaman	NICOBARS	2002	Kharif	Rice	189.2	510.84
21	Andaman	NICOBARS	2002	Whole Ye	Arecanut	1258	2083
22	Andaman	NICOBARS	2002	Whole Ye	Banana	213	1278
23	Andaman	NICOBARS	2002	Whole Ye	Black pep	63	13.5
24	Andaman	NICOBARS	2002	Whole Ye	Cashewnu	719	208
25	Andaman	NICOBARS	2002	Whole Ye	Coconut	18240	67490000

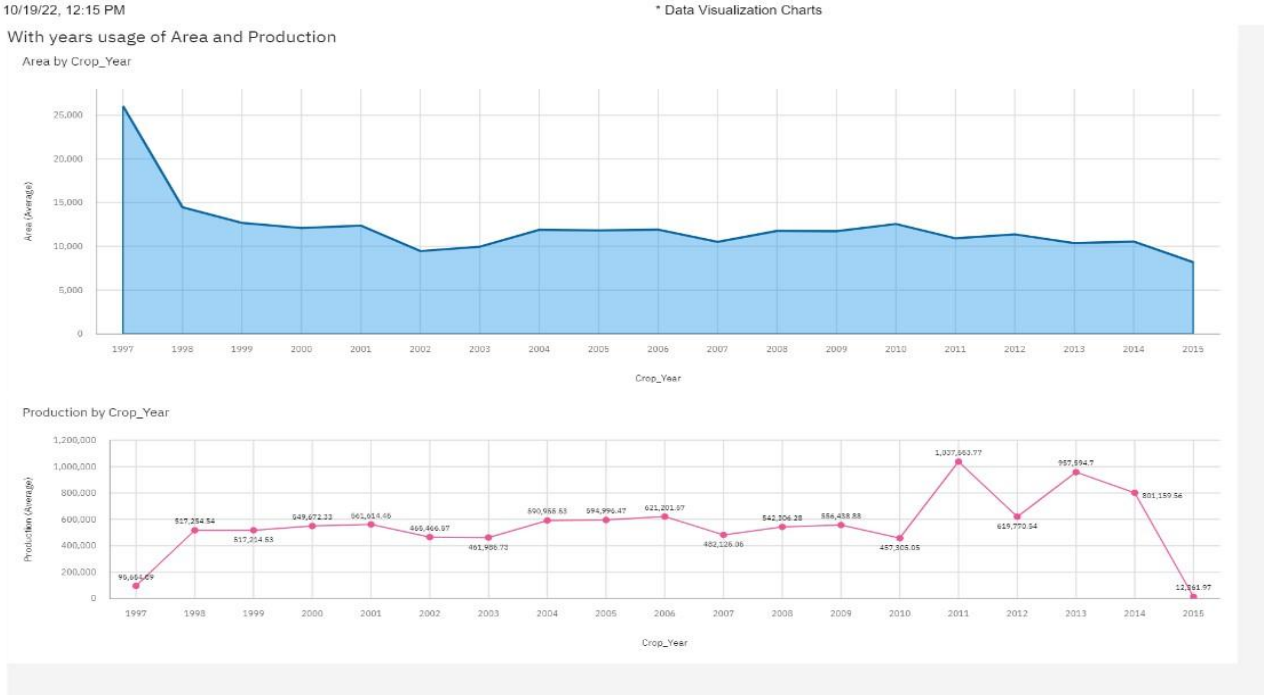
Visualization on Season With average Production



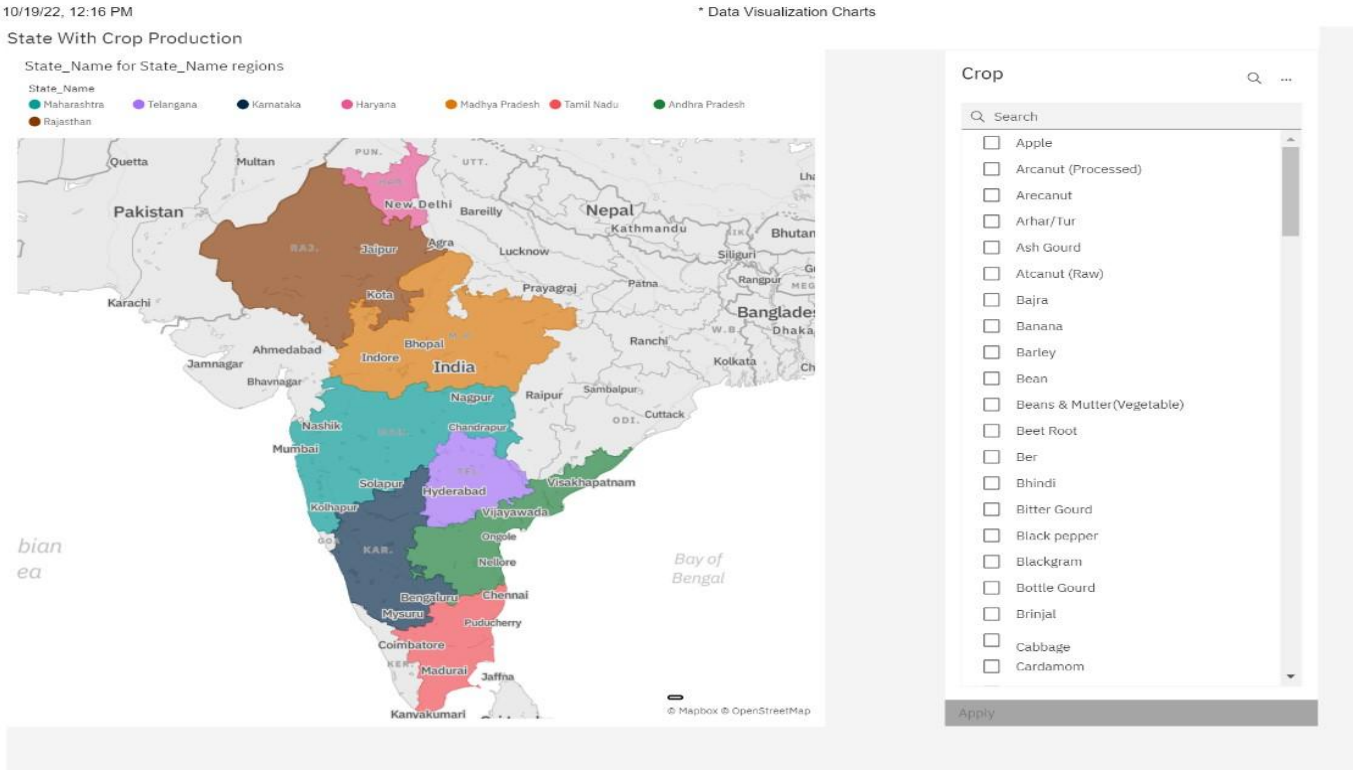
Visualization on top 10 States with most area



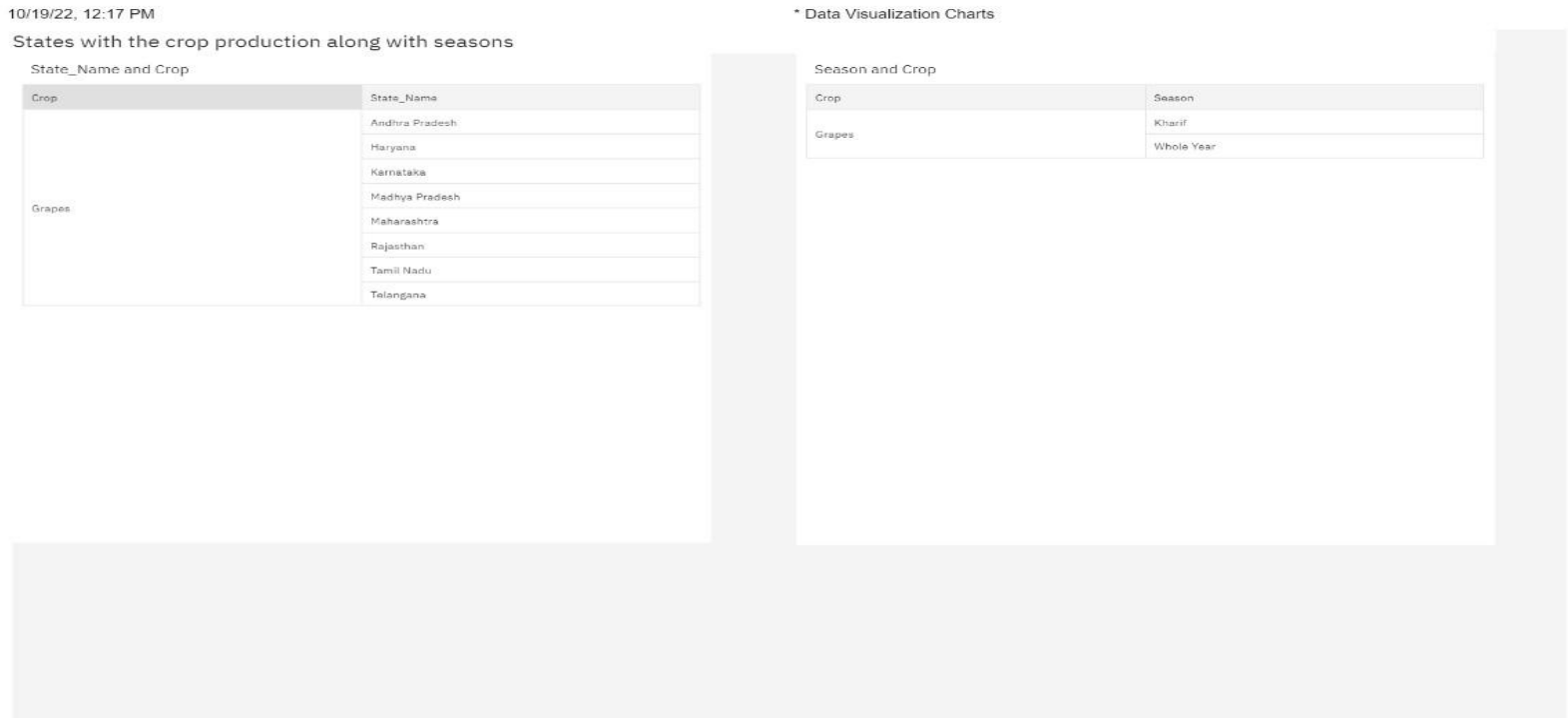
Visualization with years usage of Area and Production

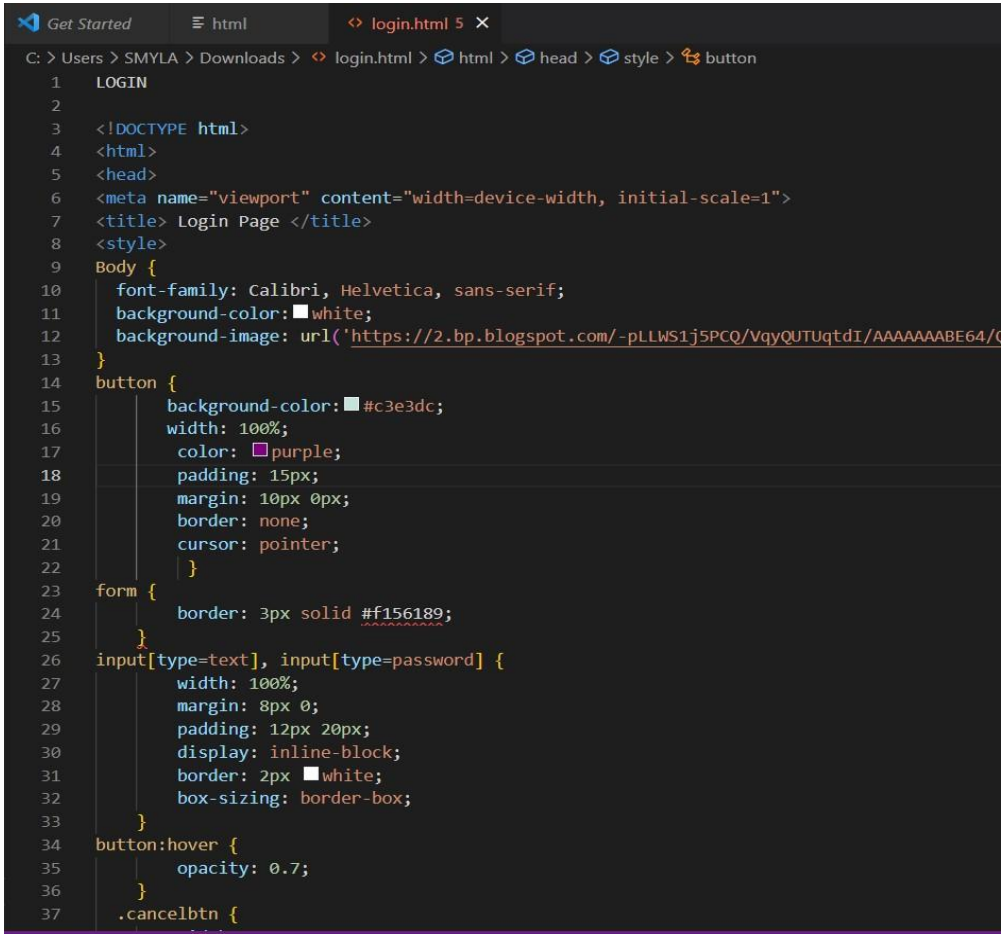


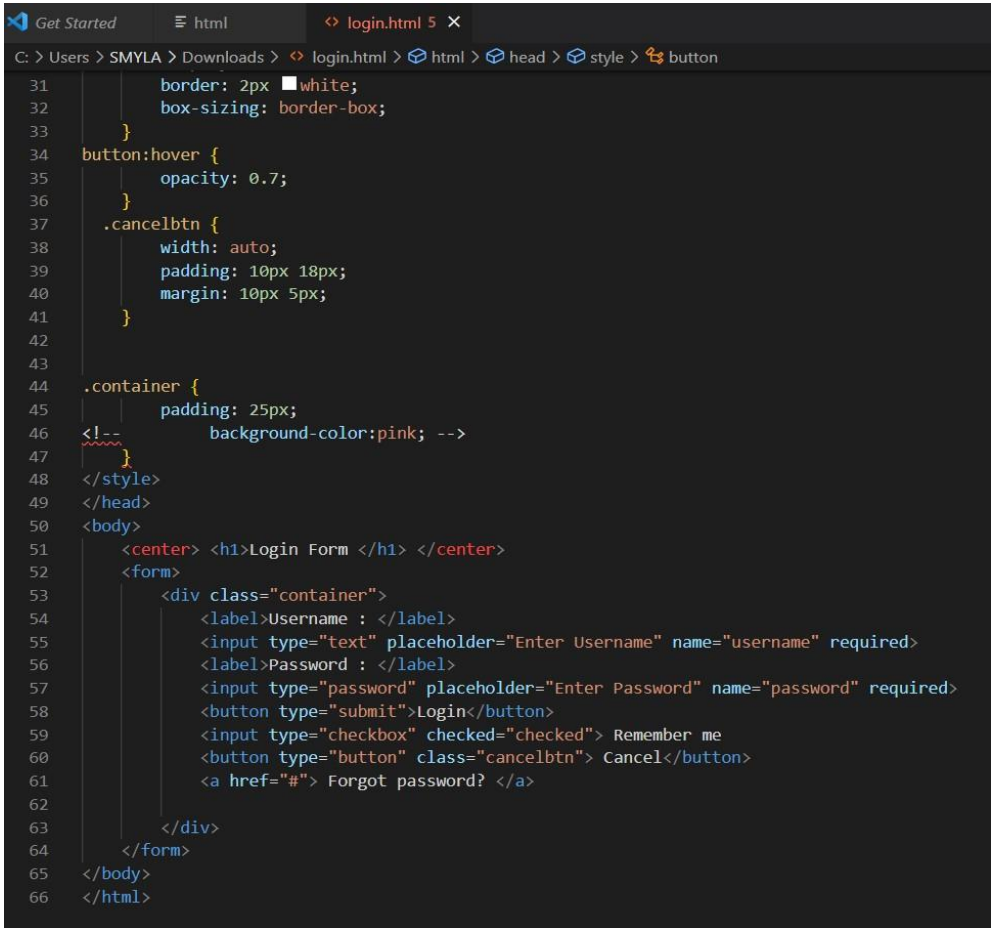
Visualization on State with Corp Production



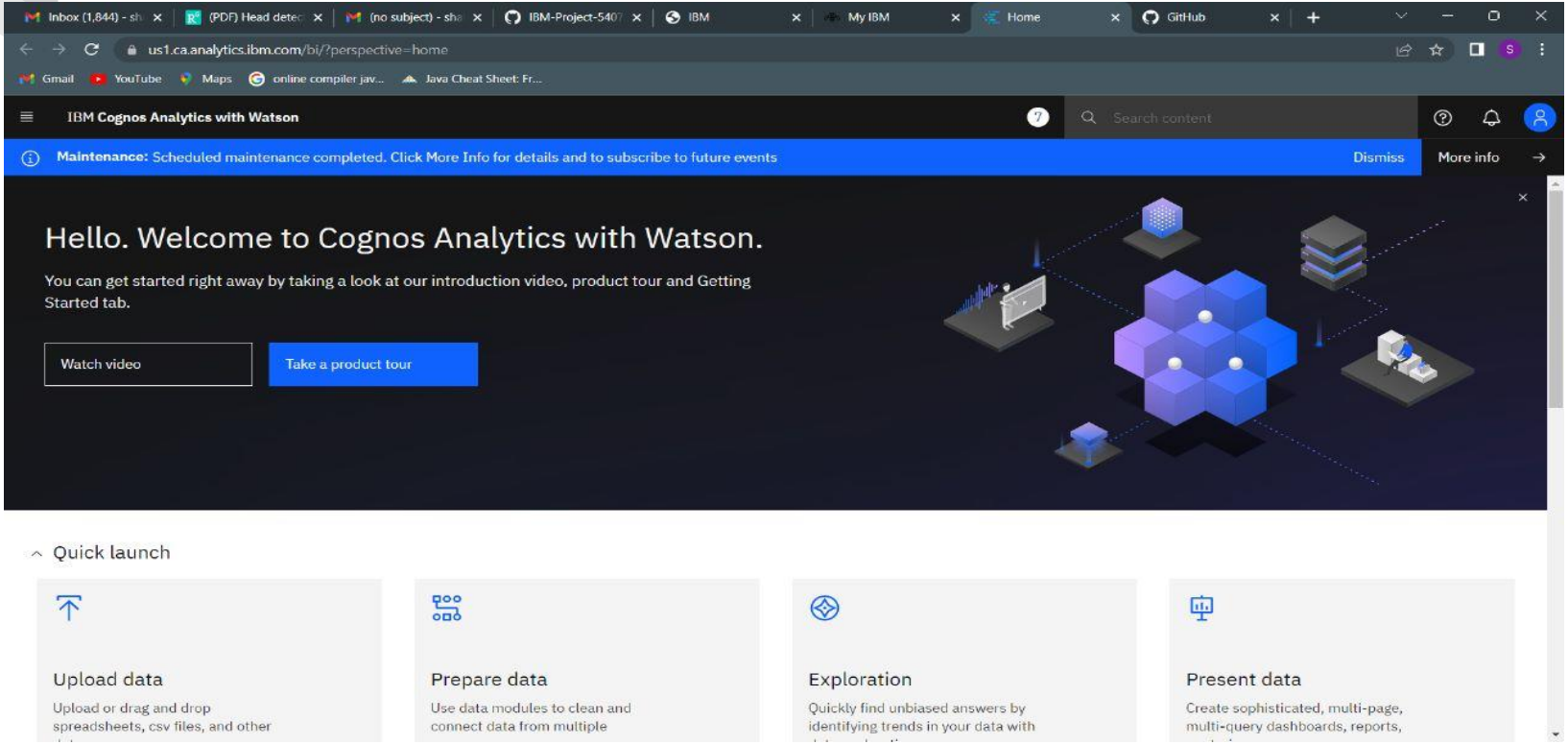
Visualization on States with the Crop Production along with Season created using the COGNOS platform

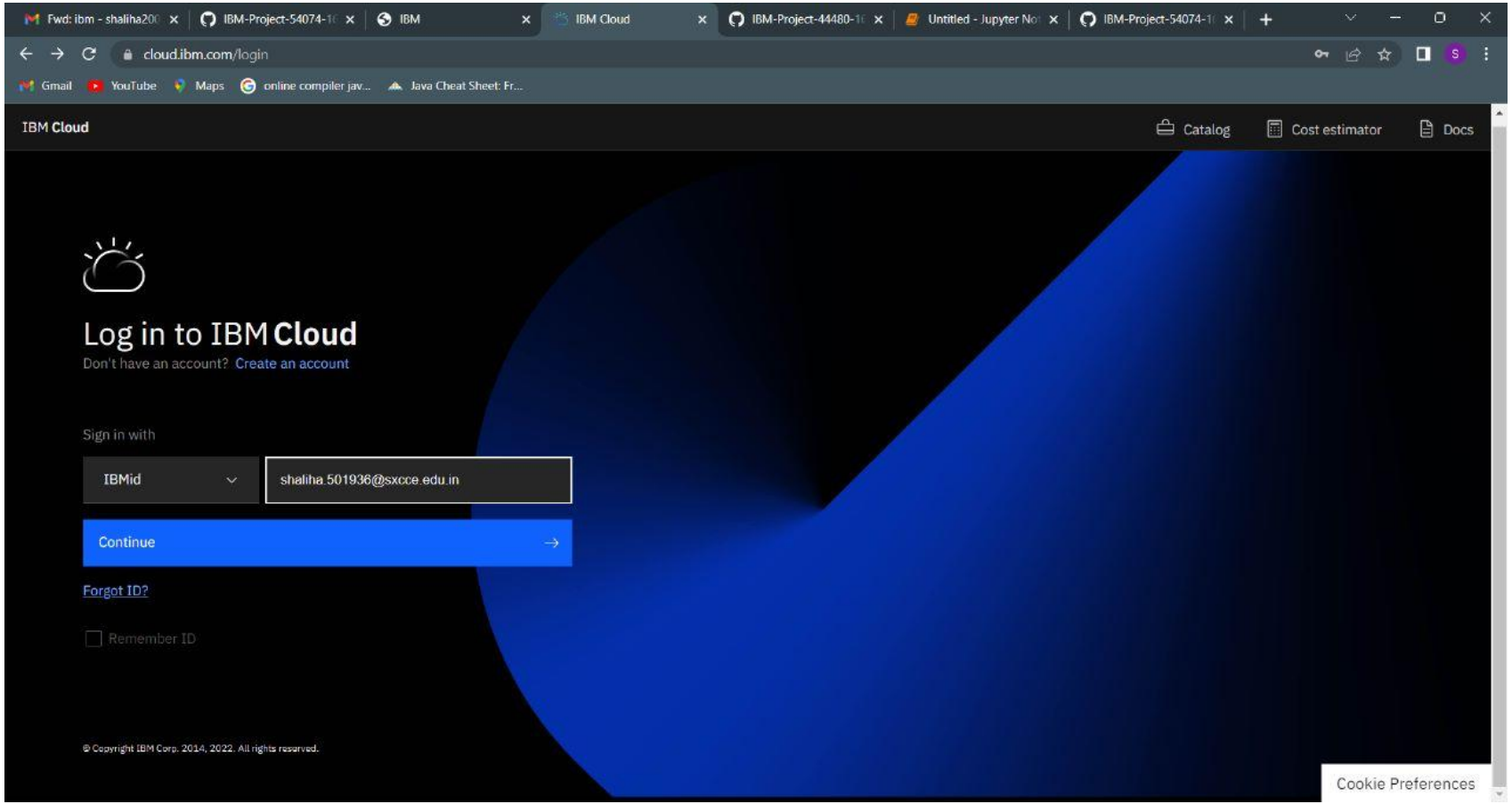


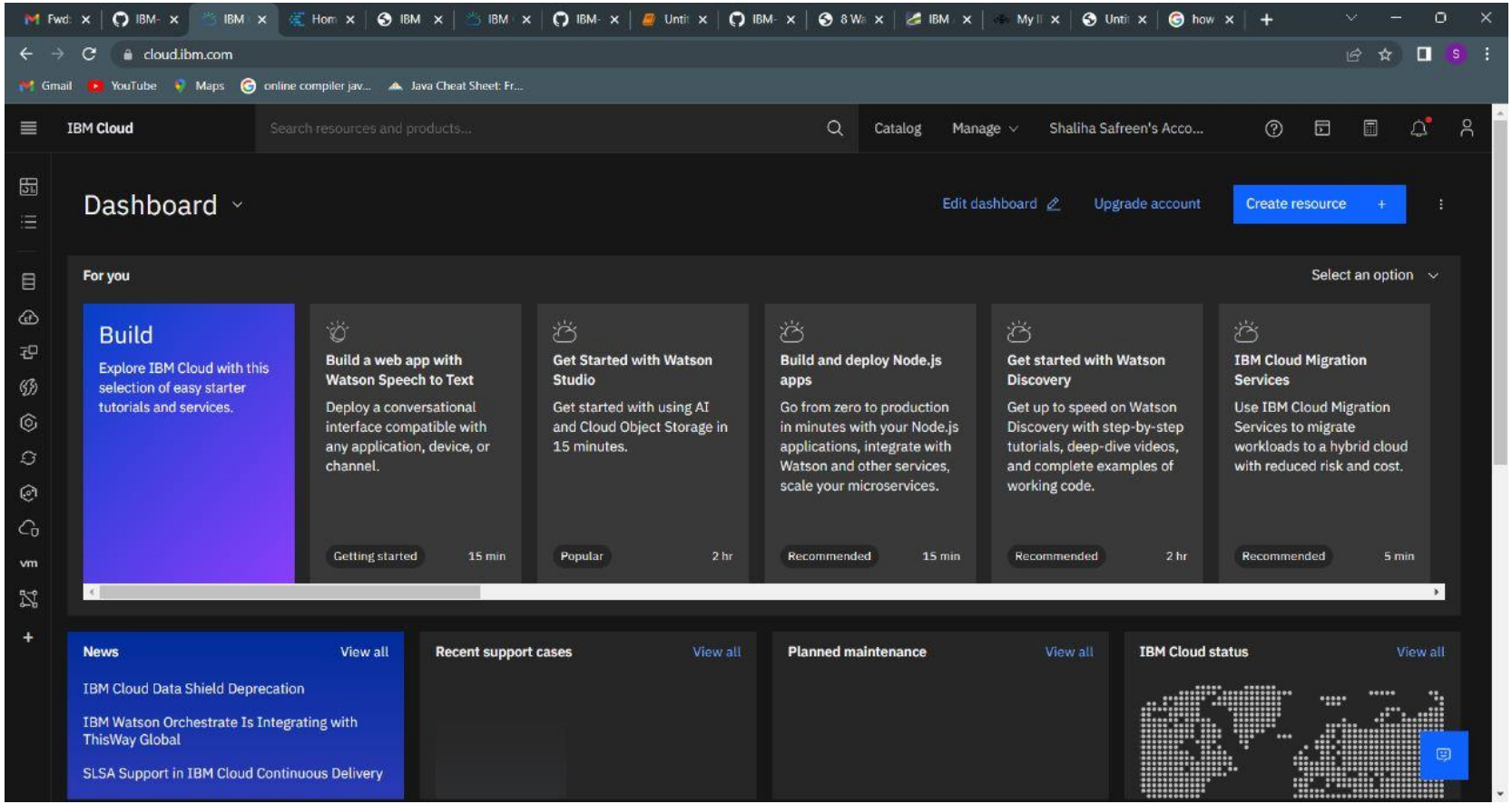

Get Started
html
login.html 5
C: > Users > SMYLA > Downloads > login.html > html > head > style > button
LOGIN
<!DOCTYPE html>
<html>
<head>
<meta name="viewport" content="width=device-width, initial-scale=1">
<title> Login Page </title>
<style>
Body {
font-family: Calibri, Helvetica, sans-serif;
background-color: white;
background-image: url('https://2.bp.blogspot.com/-pLLWS1j5PCQ/VqyQUTUqtdI/AAAAAAABE64/
}
button {
background-color: #c3e3dc;
width: 100%;
color: purple;
padding: 15px;
margin: 10px 0px;
border: none;
cursor: pointer;
}
form {
border: 3px solid #f156189;
}
input[type=text], input[type=password] {
width: 100%;
margin: 8px 0;
padding: 12px 20px;
display: inline-block;
border: 2px white;
box-sizing: border-box;
}
button:hover {
opacity: 0.7;
}
.cancelbtn {


Get Started
html
login.html 5
C: > Users > SMYLA > Downloads > login.html > html > head > style > button
31 border: 2px white;
32 box-sizing: border-box;
33 }
34 button:hover {
35 opacity: 0.7;
36 }
37 .cancelbtn {
38 width: auto;
39 padding: 10px 18px;
40 margin: 10px 5px;
41 }
42
43
44 .container {
45 padding: 25px;
46 <!-- background-color:pink; -->
47 }
48 </style>
49 </head>
50 <body>
51 <h1>Login Form </h1>
52 <form>
53
54 <label>Username : </label>
55 <input type="text" placeholder="Enter Username" name="username" required>
56 <label>Password : </label>
57 <input type="password" placeholder="Enter Password" name="password" required>
58 <button type="submit">Login</button>
59 <input type="checkbox" checked="checked"> Remember me
60 <button type="button" class="cancelbtn"> Cancel</button>
61 <a href="#"> Forgot password? </a>
62
63
64 </form>
65 </body>
66 </html>

Creating the dashboard in COGNOS




Fwd: ibm - shaliha200
IBM-Project-54074-1
IBM
IBM Cloud
IBM-Project-44480-1
Untitled - Jupyter No
IBM-Project-54074-1
cloud.ibm.com/login
Gmail
YouTube
Maps
online compiler jav...
Java Cheat Sheet: Fr...
IBM Cloud
Catalog
Cost estimator
Docs
Log in to IBM Cloud
Don't have an account? Create an account
Sign in with
IBMid
shaliha.501936@sxcce.edu.in
Continue
Forgot ID?
Remember ID
© Copyright IBM Corp. 2014, 2022. All rights reserved.
Cookie Preferences


cloud.ibm.com
Gmail
YouTube
Maps
online compiler jav...
Java Cheat Sheet: Fr...
IBM Cloud
Search resources and products...
Catalog
Manage
Shaliha Safreen's Acco...
Dashboard
Edit dashboard
Upgrade account
Create resource
For you
Select an option
Build
Explore IBM Cloud with this selection of easy starter tutorials and services.
Build a web app with Watson Speech to Text
Deploy a conversational interface compatible with any application, device, or channel.
Getting started
15 min
Get Started with Watson Studio
Get started with using AI and Cloud Object Storage in 15 minutes.
Popular
2 hr
Build and deploy Node.js apps
Go from zero to production in minutes with your Node.js applications, integrate with Watson and other services, scale your microservices.
Recommended
15 min
Get started with Watson Discovery
Get up to speed on Watson Discovery with step-by-step tutorials, deep-dive videos, and complete examples of working code.
Recommended
2 hr
IBM Cloud Migration Services
Use IBM Cloud Migration Services to migrate workloads to a hybrid cloud with reduced risk and cost.
Recommended
5 min
News
View all
IBM Cloud Data Shield Deprecation
IBM Watson Orchestrate Is Integrating with ThisWay Global
SLSA Support in IBM Cloud Continuous Delivery
Recent support cases
View all
Planned maintenance
View all
IBM Cloud status
View all

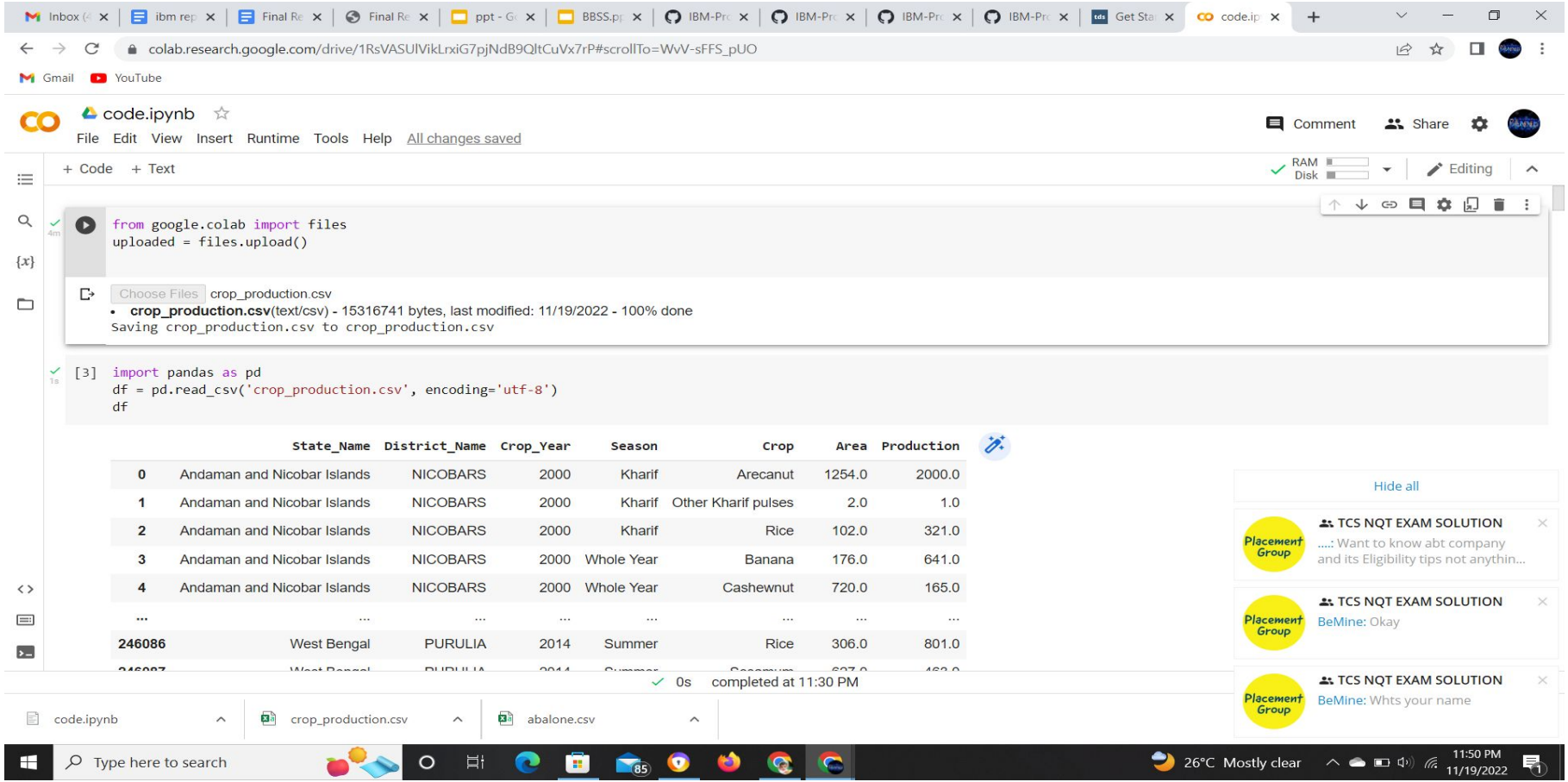
MERITS

- Weather aberration can cause physical damage to crops
- Help in cut costs.
- Product higher crop yield.
- Prevent over or under watering.

DEMERITS

- Sudden change in weather can cause crop damage.

Output:


colab.research.google.com/drive/1RsVASUlVikLrxiG7pjNdB9QltCuVx7rP#scrollTo=WvV-sFFS_pUO
code.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
from google.colab import files
uploaded = files.upload()
Choose Files crop_production.csv
crop_production.csv(text/csv) - 15316741 bytes, last modified: 11/19/2022 - 100% done
Saving crop_production.csv to crop_production.csv
[3] import pandas as pd
df = pd.read_csv('crop_production.csv', encoding='utf-8')
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
0s completed at 11:30 PM
code.ipynb
crop_production.csv
abalone.csv
Hide all
TCS NQT EXAM SOLUTION
....: Want to know abt company and its Eligibility tips not anythin...
TCS NQT EXAM SOLUTION
BeMine: Okay
TCS NQT EXAM SOLUTION
BeMine: Whts your name
Type here to search
26°C Mostly clear
11:50 PM
11/19/2022

Inbox (4 | ibm rep | Final Re | Final Re | ppt - G | BBSS.pp | IBM-Pro | IBM-Pro | IBM-Pro | IBM-Pro | Get Sta | code.ip
colab.research.google.com/drive/1RsVASUlVikLrxiG7pjNdB9QltCuVx7rP#scrollTo=WvV-sFFS_pUO
Gmail YouTube

code.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk Editing

[3]

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

[4]
```
df = df[df['State_Name'] == "Andhra Pradesh"]
```

[5]
```
df['Yield'] = df['Production']/df['Area']
df
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  """Entry point for launching an IPython kernel.
```

	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production	Yield
203	Andhra Pradesh	ANANTAPUR	1997	Kharif	Arhar/Tur	21400.0	2600.0	0.121495
204	Andhra Pradesh	ANANTAPUR	1997	Kharif	Bajra	1400.0	500.0	0.357143
205	Andhra Pradesh	ANANTAPUR	1997	Kharif	Castor seed	1000.0	100.0	0.100000

0s completed at 11:30 PM

code.ipynb crop_production.csv abalone.csv Show all

Type here to search 26°C Mostly clear 11:51 PM 11/19/2022

Inbox | ibm re | Final F | Final F | ppt - (| BBSS.p | IBM-P | IBM-P | IBM-P | IBM-P | Get St | code. | screer
colab.research.google.com/drive/1RsVASUlVikLrxiG7pjNdB9QltCuVx7rP#scrollTo=WvV-sFFS_pUO
Gmail YouTube

code.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk Editing

[5]

206	Andhra Pradesh	ANANTAPUR	1997	Kharif	Cotton(lint)	7300.0	9400.0	1.287671
207	Andhra Pradesh	ANANTAPUR	1997	Kharif	Dry chillies	3700.0	7100.0	1.918919
...	...	...	...	...	...	...	...	...
9826	Andhra Pradesh	WEST GODAVARI	2014	Rabi	Tobacco	28046.0	89107.0	3.177173
9827	Andhra Pradesh	WEST GODAVARI	2014	Rabi	Tomato	224.0	3035.0	13.549107
9828	Andhra Pradesh	WEST GODAVARI	2014	Rabi	Urad	5418.0	4833.0	0.892027
9829	Andhra Pradesh	WEST GODAVARI	2014	Whole Year	Coconut	21729.0	718991000.0	33089.005477
9830	Andhra Pradesh	WEST GODAVARI	2014	Whole Year	Rice	409286.0	1624324.0	3.968677

9628 rows × 8 columns

[6]

```
import matplotlib.pyplot as plt
import seaborn as sb

C_mat = df.corr()
fig = plt.figure(figsize = (15,15))

sb.heatmap(C_mat, vmax = .8, square = True)
plt.show()
```

0s completed at 11:30 PM

code.ipynb crop_production.csv abalone.csv Show all

Type here to search 26°C Mostly clear 11:53 PM 11/19/2022

Inbox | ibm re | Final F | Final F | ppt - | BBSS.j | IBM-P | IBM-P | IBM-P | IBM-P | Get St | code.j | scree | +
colab.research.google.com/drive/1RsVASUlVikLrxiG7pjNdB9QltCuVx7rP#scrollTo=WvV-sFFS_pUO
Gmail YouTube

code.ipynb

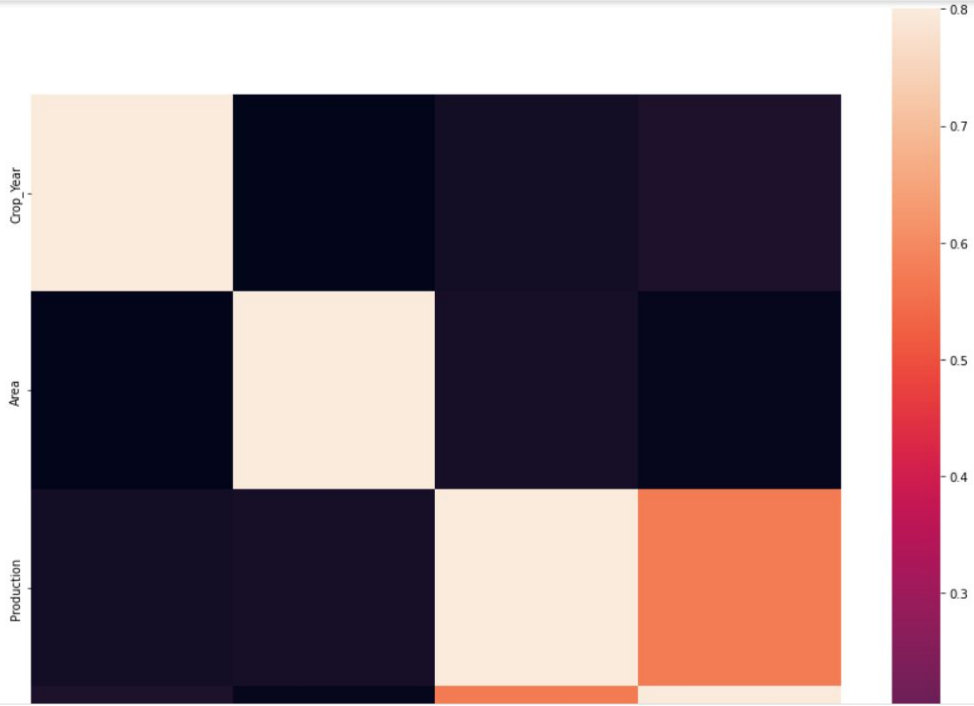
File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk Editing

[6]



0s completed at 11:30 PM

Type here to search
26°C Mostly clear
11:55 PM
11/19/2022

Inbox | ibm re | Final F | Final F | ppt - | BBSS.j | IBM-P | IBM-P | IBM-P | IBM-P | Get St | code.i | screer

colab.research.google.com/drive/1RsVASUlVikLrxiG7pjNdB9QltCuVx7rP#scrollTo=WvV-sFFS_pUO

Gmail YouTube

code.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk Editing

[12] df = df.fillna(df.mean())

[13]
```
from sklearn.model_selection import train_test_split
b = df['Yield']
a = df.drop('Yield', axis = 1)

a_train, a_test, b_train, b_test = train_test_split(a, b, test_size = 0.3, random_state = 42)

print(a_train)
print(a_test)
print(b_train)
print(b_test)
```

```
          Area  ANANTAPUR  CHITTOOR  EAST GODAVARI  GUNTUR  KADAPA  KRISHNA  \
1711  0.000246          0         1              0       0       0        0
5452  0.031408          0         0              0       0       0        0
538   0.003756          1         0              0       0       0        0
2336  0.014880          0         0              1       0       0        0
2172  0.001397          0         0              1       0       0        0
...        ...        ...       ...            ...     ...     ...      ...
6278  0.001999          0         0              0       0       0        0
8796  0.000129          0         0              0       0       0        0
8831  0.000385          0         0              0       0       0        0
8995  0.000344          0         0              0       0       0        0
1644  0.000268          0         1              0       0       0        0

      KURNOOL  PRAKASAM  SPSR NELLORE  ...  2006  2007  2008  2009  2010  \
1711        0         0             0  ...     0     0     0     0     0
5452        1         0             0  ...     0     0     0     0     0
538         0         0             0  ...     0     0     0     0     0
```

0s completed at 11:30 PM

Type here to search

26°C Mostly clear

11:56 PM 11/19/2022

CONCLUSION

- Weather aberration can cause physical damage to crop.With help of this project we can predict in certain environmental conduction which crop should be taken.

- This data will continue to enhance farmer efficiency by further enabling them to monitor each plot of land and determine the precise input needed for their crops.

THANK YOU