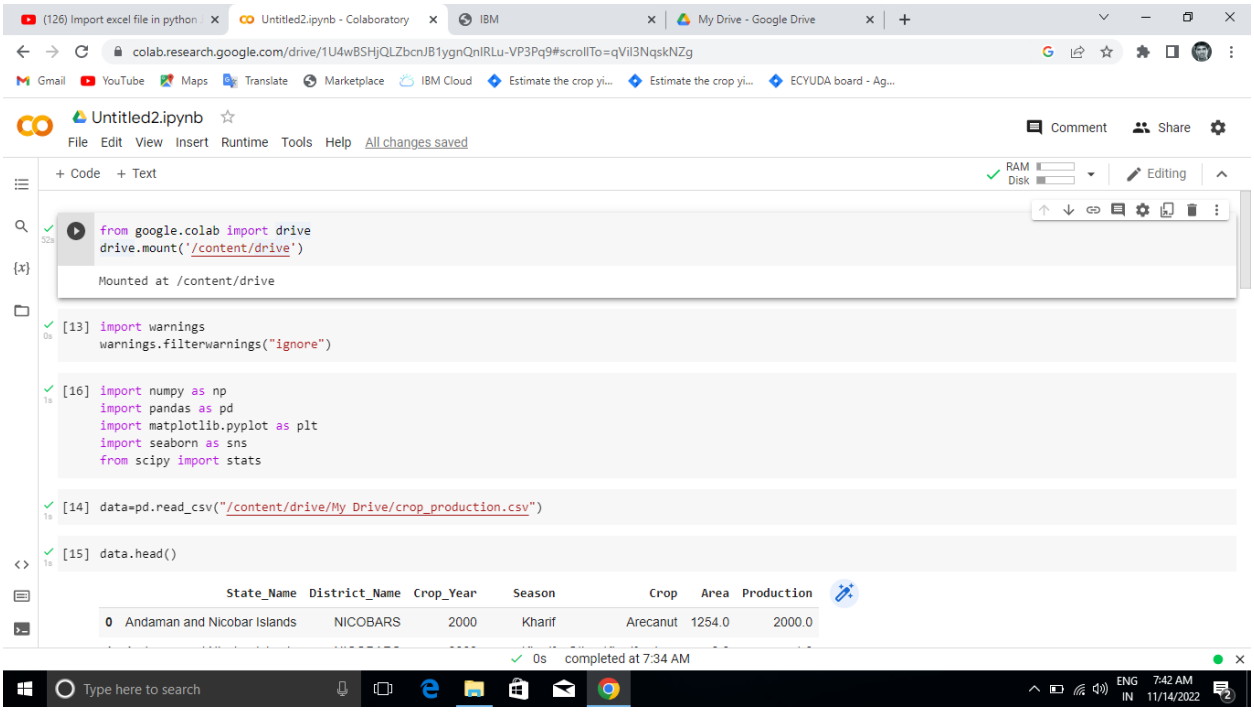
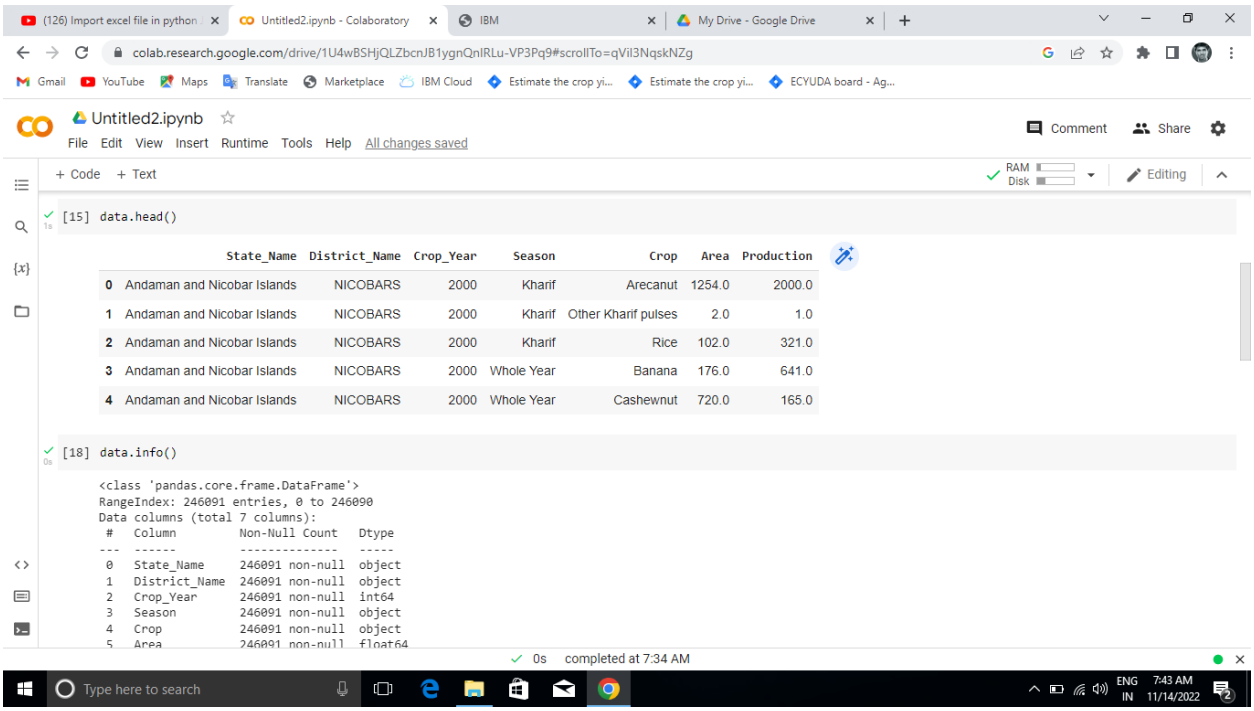
PROJECT DEVELOPMENT PHASE

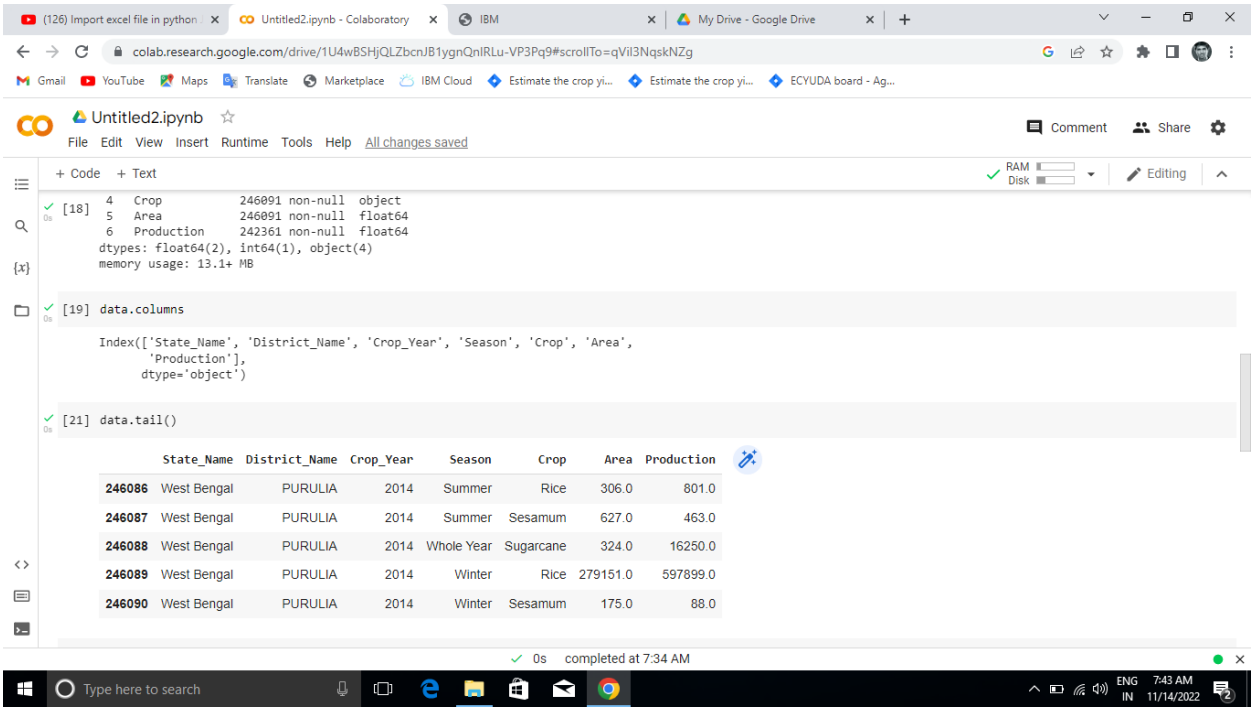
SPRINT – 2

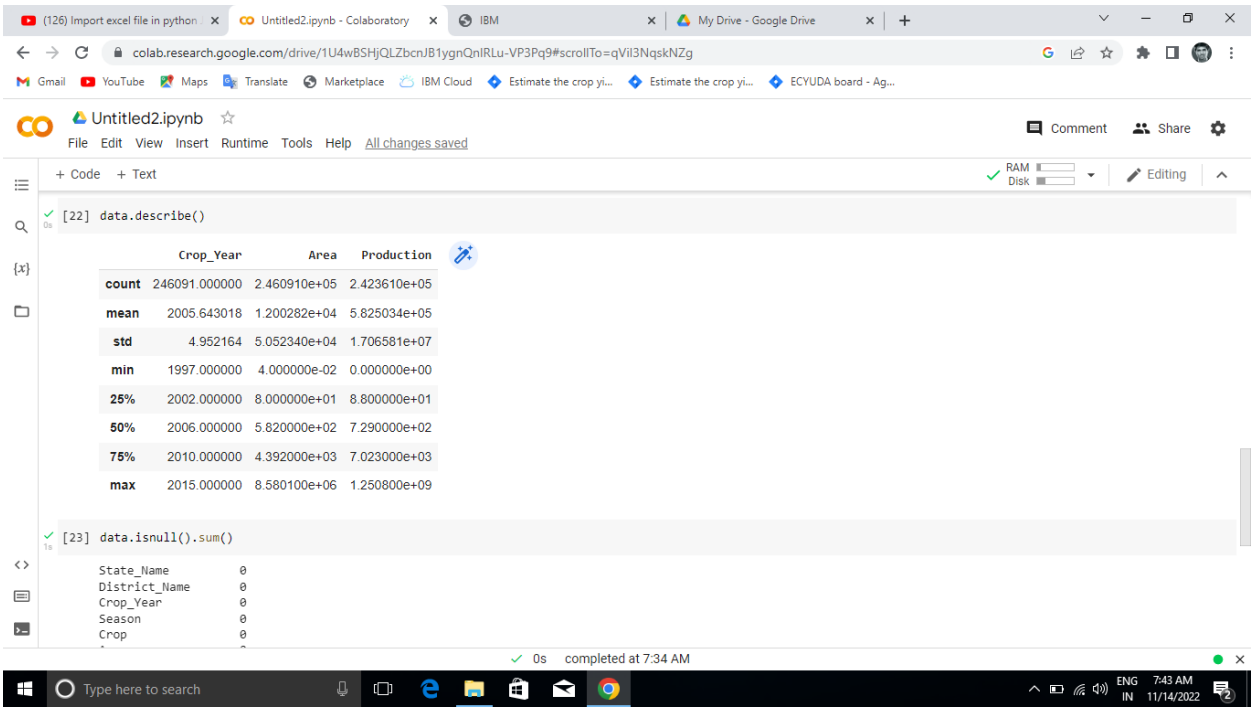
Date	31 October 2022
Team Id	PNT2022TMID07696
Project Name	Estimate the Crop Yield Using Data Analytics

DATA EXPLORATION INPUT & OUTPUT:




(126) Import excel file in python
Untitled2.ipynb - Colaboratory
IBM
My Drive - Google Drive
colab.research.google.com/drive/1U4wBSHjQLZbcnJB1ygnQnIRLu-VP3Pq9#scrollTo=qViI3NqskNZg
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
[15] data.head()
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
[18] data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 246091 entries, 0 to 246090
Data columns (total 7 columns):
Column Non-Null Count Dtype
0 State_Name 246091 non-null object
1 District_Name 246091 non-null object
2 Crop_Year 246091 non-null int64
3 Season 246091 non-null object
4 Crop 246091 non-null object
5 Area 246091 non-null float64
0s completed at 7:34 AM
Type here to search
ENG IN 7:43 AM 11/14/2022


(126) Import excel file in python
Untitled2.ipynb - Colaboratory
IBM
My Drive - Google Drive
colab.research.google.com/drive/1U4wBSHjQLZbcnJB1ygnQnIRLu-VP3Pq9#scrollTo=qViI3NqskNZg
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Editing
[18] 4 Crop 246091 non-null object
5 Area 246091 non-null float64
6 Production 242361 non-null float64
dtypes: float64(2), int64(1), object(4)
memory usage: 13.1+ MB
[19] data.columns
Index(['State_Name', 'District_Name', 'Crop_Year', 'Season', 'Crop', 'Area',
'Production'],
dtype='object')
[21] data.tail()
State_Name District_Name Crop_Year Season Crop Area Production
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
0s completed at 7:34 AM
Type here to search
ENG IN 7:43 AM 11/14/2022



(126) Import excel file in python | Untitled2.ipynb - Colaboratory | IBM | My Drive - Google Drive

colab.research.google.com/drive/1U4wBSHjQLZbcnJB1ygnQnlRLu-VP3Pq9#scrollTo=qViI3NqskNZg

Gmail YouTube Maps Translate Marketplace IBM Cloud Estimate the crop yi... Estimate the crop yi... ECYUDA board - Ag...

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk Editing

```
[22] data.describe()
```

	Crop_Year	Area	Production
count	246091.000000	2.460910e+05	2.423610e+05
mean	2005.643018	1.200282e+04	5.825034e+05
std	4.952164	5.052340e+04	1.706581e+07
min	1997.000000	4.000000e-02	0.000000e+00
25%	2002.000000	8.000000e+01	8.800000e+01
50%	2006.000000	5.820000e+02	7.290000e+02
75%	2010.000000	4.392000e+03	7.023000e+03
max	2015.000000	8.580100e+06	1.250800e+09

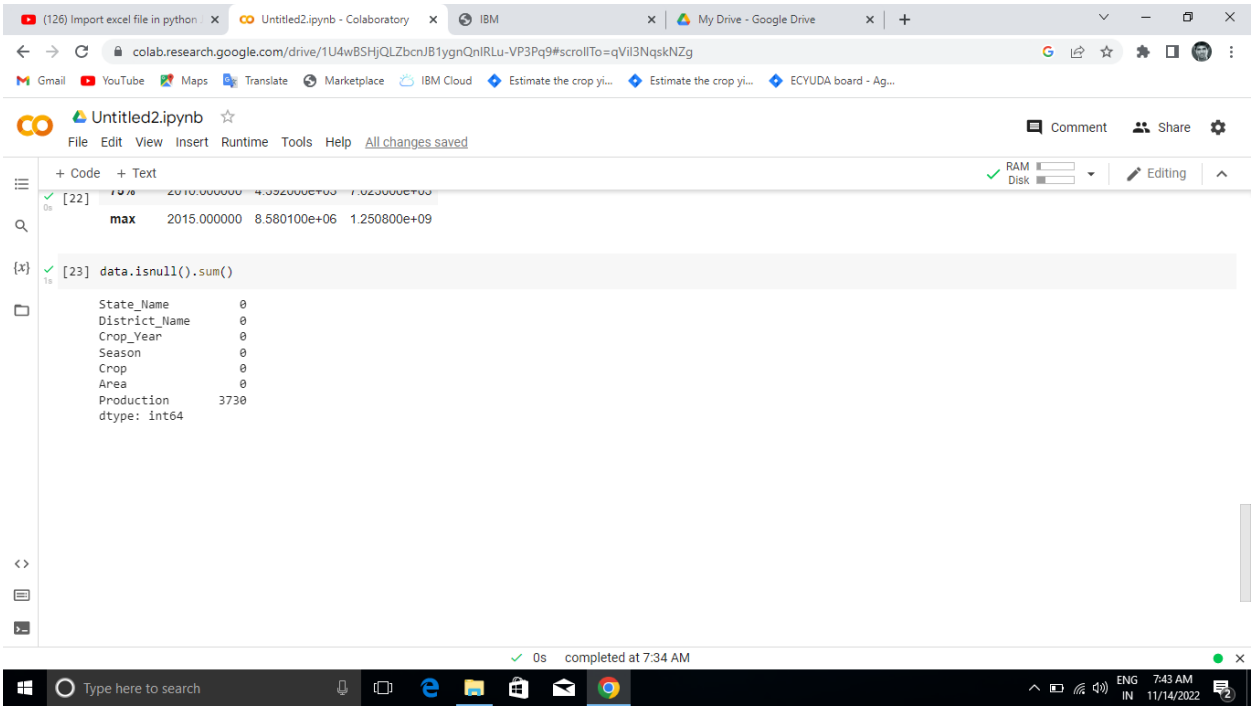
```
[23] data.isnull().sum()
```

```
State_Name          0
District_Name       0
Crop_Year           0
Season              0
Crop                0
```

0s completed at 7:34 AM

Type here to search

ENG IN 7:43 AM 11/14/2022



(126) Import excel file in python | Untitled2.ipynb - Colaboratory | IBM | My Drive - Google Drive

colab.research.google.com/drive/1U4wBSHjQLZbcnJB1ygnQnlRLu-VP3Pq9#scrollTo=qViI3NqskNZg

Gmail YouTube Maps Translate Marketplace IBM Cloud Estimate the crop yi... Estimate the crop yi... ECYUDA board - Ag...

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk Editing

```
[22]
```

max	2015.000000	8.580100e+06	1.250800e+09

```
[23] data.isnull().sum()
```

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

0s completed at 7:34 AM

Type here to search

ENG IN 7:43 AM 11/14/2022

(126) Import excel file in pyth | Untitled2.ipynb - Colaborato | IBM | My Drive - Google Drive | seabon displot tag - Google

colab.research.google.com/drive/1U4wBSHjQLZbcnJB1ygnQnIRLu-VP3Pq9#scrollTo=-MrMNq_wNkuv

Gmail YouTube Maps Translate Marketplace IBM Cloud Estimate the crop yi... Estimate the crop yi... ECYUDA board - Ag...

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

+ Code + Text

RAM Disk Editing

```
[ ] data.corr()
```

	Crop_Year	Area	Production
Crop_Year	1.000000	-0.026022	0.006989
Area	-0.026022	1.000000	0.040587
Production	0.006989	0.040587	1.000000

```
[ ] data.cov()
```

	Crop_Year	Area	Production
Crop_Year	24.523927	-6.510591e+03	5.914148e+05
Area	-6510.590664	2.552614e+09	3.522683e+10
Production	591414.831146	3.522683e+10	2.912420e+14

```
[ ] data.dtypes

State_Name        object
District_Name     object
Crop_Year          int64
```

0s completed at 9:11 AM

Type here to search

ENG IN 9:12 AM 11/14/2022

(126) Import excel file in pyth | Untitled2.ipynb - Colaborato | IBM | My Drive - Google Drive | seabon displot tag - Google

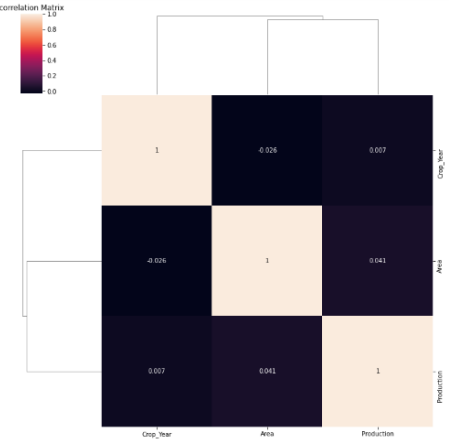
colab.research.google.com/drive/1U4wBSHjQLZbcnJB1ygnQnIRLu-VP3Pq9#scrollTo=-MrMNq_wNkuv

Gmail YouTube Maps Translate Marketplace IBM Cloud Estimate the crop yi... Estimate the crop yi... ECYUDA board - Ag...

+ Code + Text

RAM Disk

```
[ ] correlation Matrix
```




```
[ ] %matplotlib inline

[ ] data['Crop'].hist(bins=10)

    <matplotlib.axes._subplots.AxesSubplot at 0x7f597fc08a90>
```

0s completed at 9:11 AM

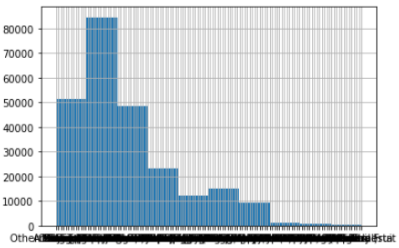
Type here to search

ENG IN 9:13 AM 11/14/2022

```
[ ] %matplotlib inline
```

```
[ ] data['Crop'].hist(bins=10)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f597fc08a90>
```




```
[ ] data['District_Name'].hist(bins=10)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f597fb5b290>
```

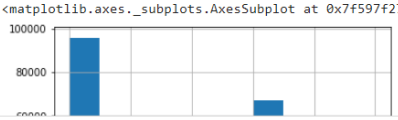
0s completed at 9:11 AM


```
[ ] data['District_Name'].hist(bins=10)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f597fb5b290>
```

```
[ ] data['Season'].hist(bins=10)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f597f27db50>
```



0s completed at 9:11 AM

(126) Import excel file in pyth | Untitled2.ipynb - Colaborato | IBM | My Drive - Google Drive | seabon displot tag - Google

colab.research.google.com/drive/1U4wBSHjQLZbcnJB1ygnQnIRLu-VP3Pq9#scrollTo=-MrMNq_wNkuv

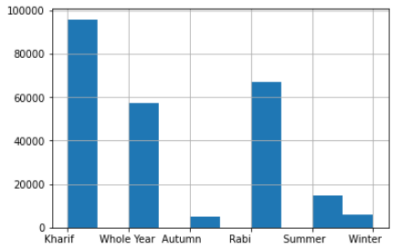
Gmail YouTube Maps Translate Marketplace IBM Cloud Estimate the crop yi... Estimate the crop yi... ECYUDA board - Ag...

+ Code + Text

RAM Disk

```
[ ] data['Season'].hist(bins=10)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f597f27db50>
```




```
[ ] data.columns
```

```
Index(['State_Name', 'District_Name', 'Crop_Year', 'Season', 'Crop', 'Area',
       'Production'],
      dtype='object')
```

```
[ ]
```

0s completed at 9:11 AM

Type here to search

ENG IN 9:14 AM 11/14/2022