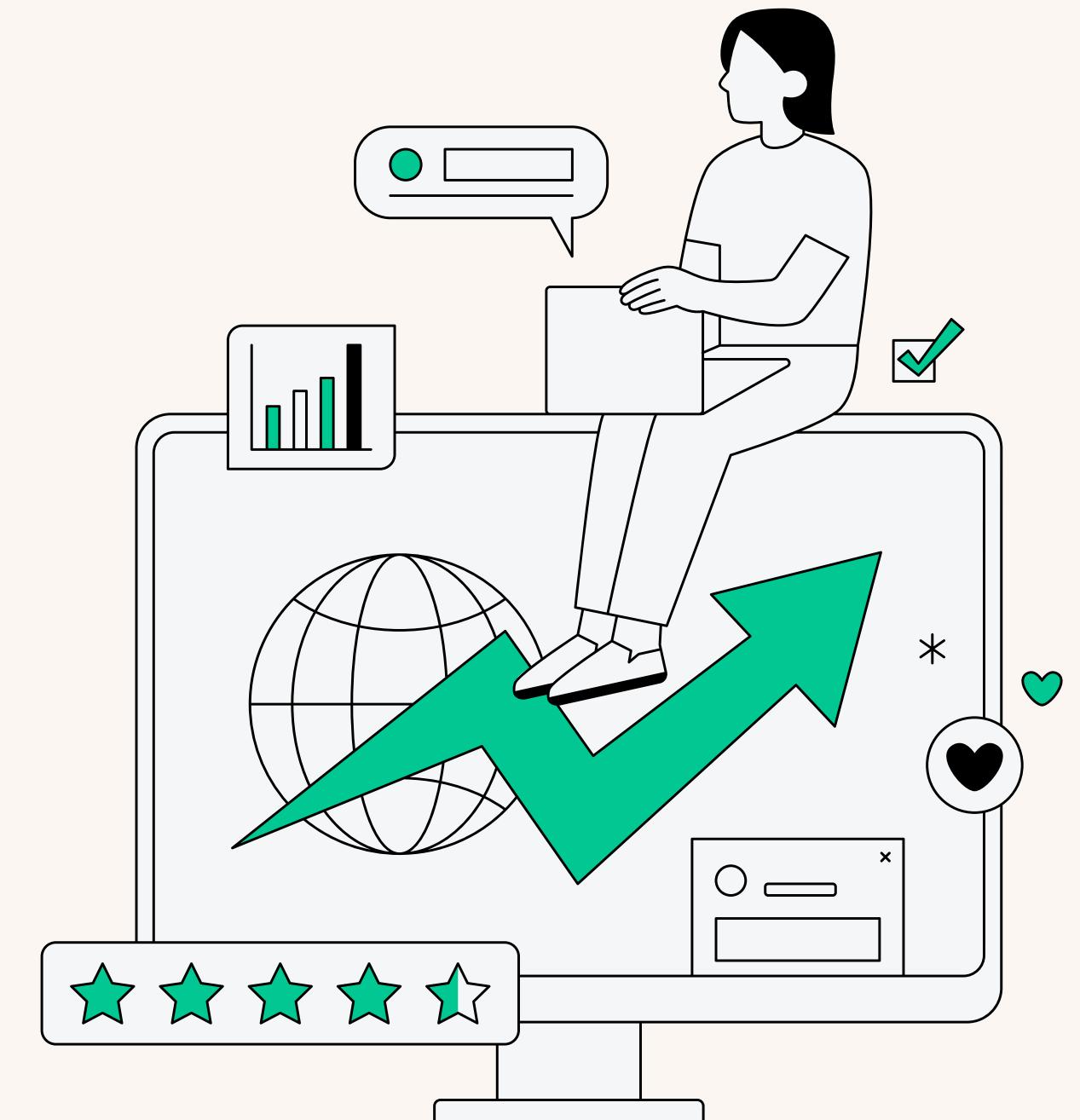


DATA ANALYST INTERNSHIP

BATCH NAME:-MIP-DA-04
BY:- VAISHNAVI PATOLE

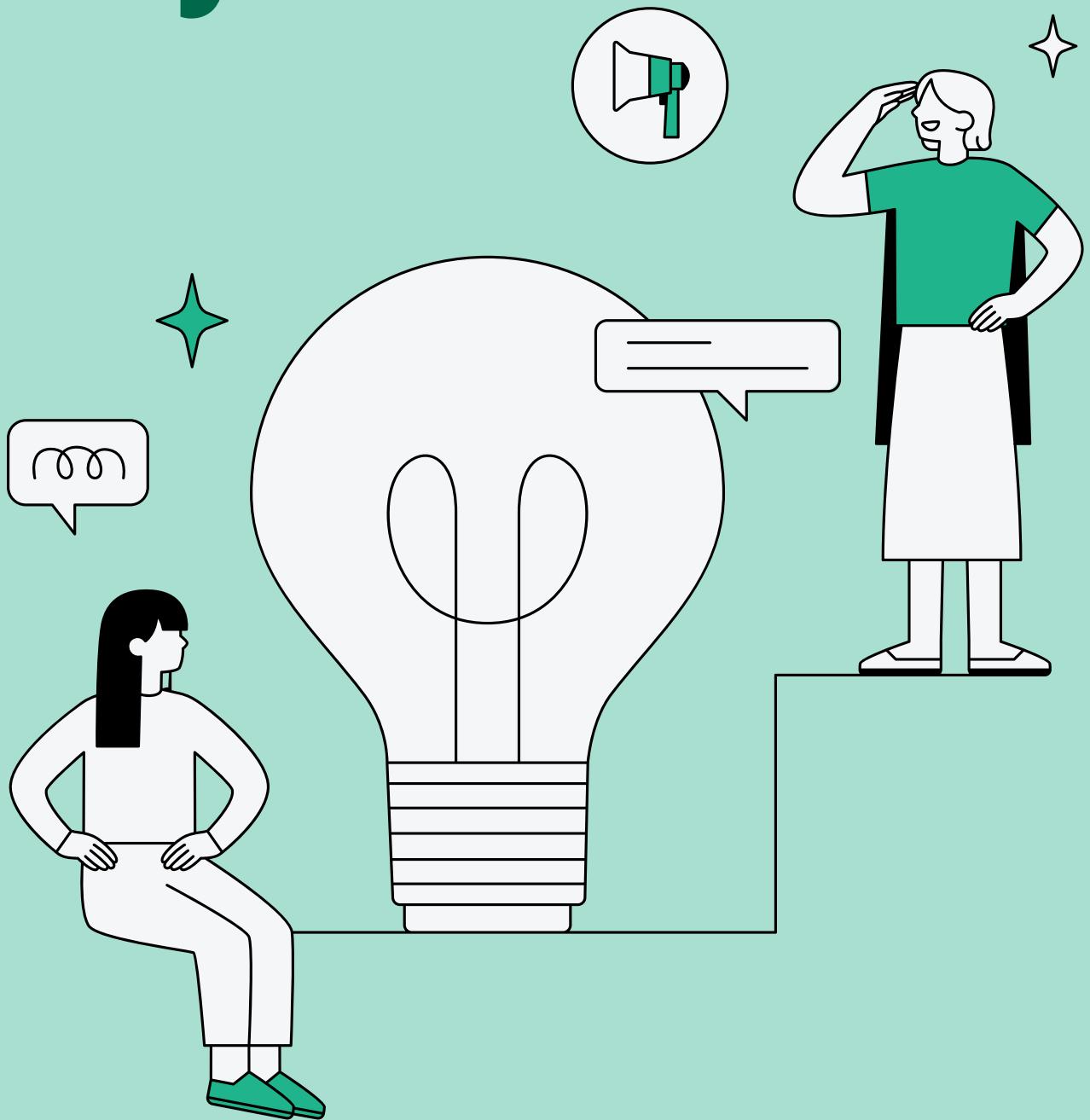


Indian Agriculture Analysis

Problem Statement:

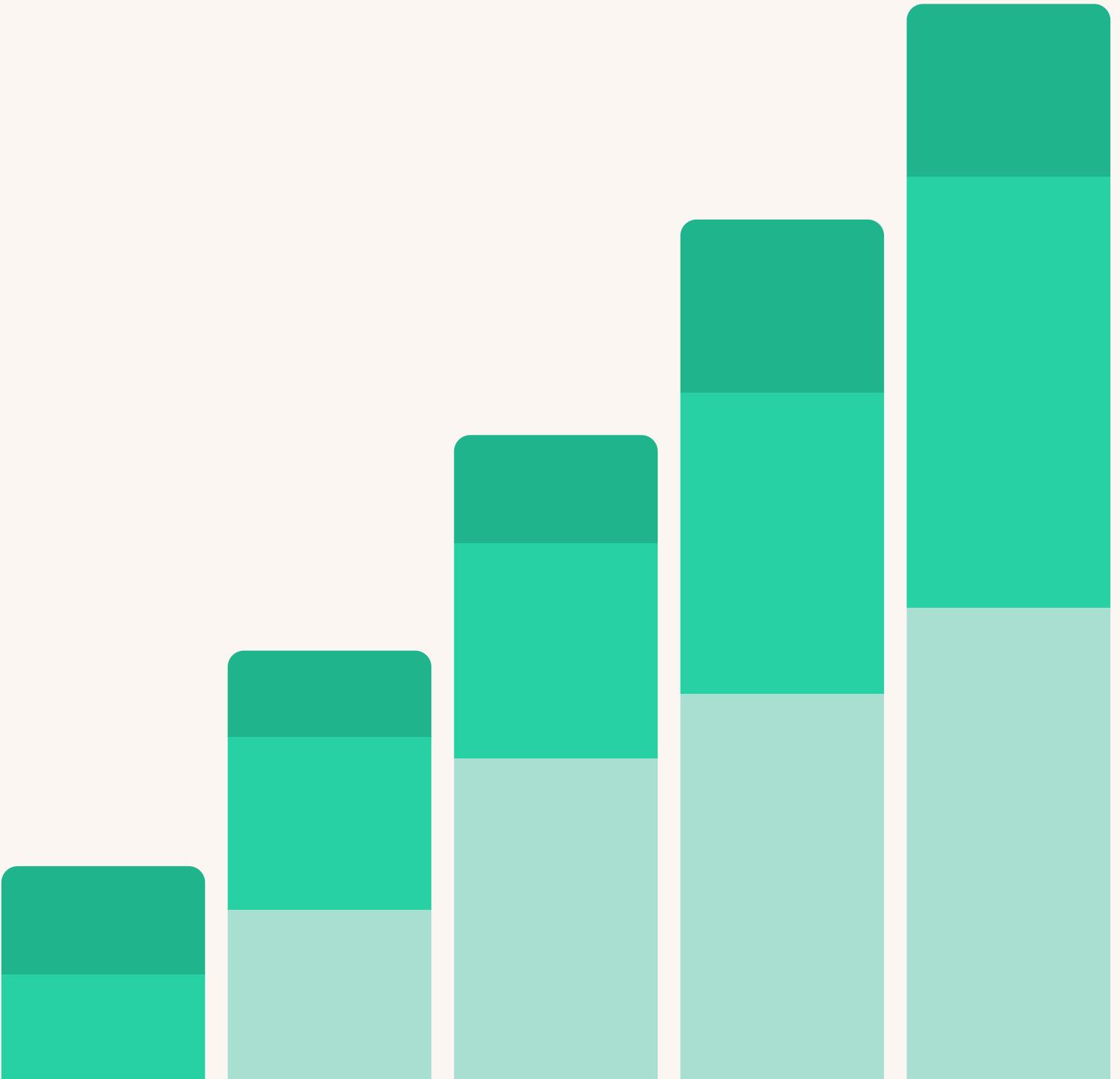
This internship project aims to conduct a comprehensive analysis of Indian agriculture, focusing on district-wise and year-wise data. The dataset provides detailed information on various crops, their areas, production, and yields across different districts and years.

The goal is to leverage Power BI to create interactive visualizations that uncover trends, patterns, and disparities in agricultural practices, enabling stakeholders to make informed decisions for sustainable farming and resource allocation.



Dataset Description:

The dataset encompasses a wide range of agricultural variables, including crop areas, production quantities, and yields for different crops such as rice, wheat, sorghum, millets, pulses, oilseeds, sugarcane, and more.



Overview

The screenshot shows a Jupyter Notebook interface running on a local host. The notebook title is "Vaishnavi Project". The menu bar includes File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. The toolbar below the menu bar includes icons for file operations like Open, Save, and Run, along with a "Code" dropdown.

The code cells show the following sequence:

```
In [1]: import pandas as pd
In [3]: df=pd.read_csv("ICRISAT-District Level Data.csv")
In [4]: df.describe()
```

The output of cell In [4] is a descriptive statistics table for the dataset:

	Dist Code	Year	State Code	RICE AREA (1000 ha)	RICE PRODUCTION (1000 tons)	RICE YIELD (Kg per ha)	WHEAT AREA (1000 ha)	WHEAT PRODUCTION (1000 tons)	WHEAT YIELD (Kg per ha)	KHARIF SORGHUM AREA (1000 ha)	SUC ...
count	16146.000000	16146.000000	16146.000000	16146.000000	16146.000000	16146.000000	16146.000000	16146.000000	16146.000000	16146.000000	16
mean	269.769231	1991.496841	9.568562	128.593192	224.889565	1486.924784	77.057946	182.012746	1492.419859	22.632268	4
std	278.309125	15.011185	4.988538	160.078825	326.629828	956.185281	100.394479	348.834254	1081.255367	45.062714	3
min	1.000000	1966.000000	1.000000	-1.000000	-1.000000	-1.000000	-1.000000	-1.000000	-1.000000	-1.000000	...
25%	78.000000	1978.000000	6.000000	10.400000	9.460000	800.000000	1.770000	2.000000	750.000000	0.000000	2
50%	156.000000	1991.000000	10.000000	66.800000	95.840000	1333.210000	36.800000	42.700000	1347.450000	2.050000	4
75%	241.000000	2005.000000	12.000000	191.390000	315.715000	2113.517500	123.000000	215.192500	2131.580000	20.900000	6
max	917.000000	2017.000000	20.000000	1154.230000	3215.010000	5653.830000	879.490000	4305.500000	5541.520000	334.800000	22

The status bar at the bottom shows the number of rows (8) and columns (78), a search bar, and system icons for date and time (19:34, 28-03-2024).

Home Page - Select or create a Vaishnavi Project - Jupyter Note Vaishnavi Project

localhost:8888/notebooks/Vaishnavi%20Project.ipynb#

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. [Don't show anymore](#)

jupyter Vaishnavi Project Last Checkpoint: an hour ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

In [6]: df.shape
Out[6]: (16146, 80)

In [7]: df.head()
Out[7]:

	Dist Code	Year	State Code	State Name	Dist Name	RICE AREA (1000 ha)	RICE PRODUCTION (1000 tons)	RICE YIELD (Kg per ha)	WHEAT AREA (1000 ha)	WHEAT PRODUCTION (1000 tons)	SUGARCANE YIELD (Kg per ha)	COTTON AREA (1000 ha)	COTTON PRODUCTION (1000 tons)	COTTON YIELD (Kg per ha)	FRUITS AREA (1000 ha)
0	1	1966	14	Chhattisgarh	Durg	548.0	185.0	337.59	44.0	20.0	1777.78	0.0	0.0	0.0	5.95
1	1	1967	14	Chhattisgarh	Durg	547.0	409.0	747.71	50.0	26.0	1500.00	0.0	0.0	0.0	5.77
2	1	1968	14	Chhattisgarh	Durg	556.3	468.0	841.27	53.7	30.0	1000.00	0.0	0.0	0.0	5.41
3	1	1969	14	Chhattisgarh	Durg	563.4	400.8	711.40	49.4	26.5	1900.00	0.0	0.0	0.0	5.52
4	1	1970	14	Chhattisgarh	Durg	571.6	473.6	828.55	44.2	29.0	2000.00	0.0	0.0	0.0	5.45

5 rows x 80 columns

Type here to search 19:34 ENG 28-03-2024

jupyter Vaishnavi Project Last Checkpoint: an hour ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) O

In [8]: df.tail()

Out[8]:

	Dist Code	Year	State Code	State Name	Dist Name	RICE AREA (1000 ha)	RICE PRODUCTION (1000 tons)	RICE YIELD (Kg per ha)	WHEAT AREA (1000 ha)	WHEAT PRODUCTION (1000 tons)	SUGARCANE YIELD (Kg per ha)	COTTON AREA (1000 ha)	COTTON PRODUCTION (1000 tons)	COTTON YIELD (Kg per ha)
16141	917	2013	15	Jharkhand	Singhbhum	267.06	579.70	2170.67	1.53	1.85	0.0	0.0	0.0	0.0
16142	917	2014	15	Jharkhand	Singhbhum	256.33	586.63	2288.57	5.36	6.65	0.0	0.0	0.0	0.0
16143	917	2015	15	Jharkhand	Singhbhum	263.21	264.71	1005.70	1.99	1.82	0.0	0.0	0.0	0.0
16144	917	2016	15	Jharkhand	Singhbhum	224.05	319.01	1423.84	0.38	0.83	0.0	0.0	0.0	0.0
16145	917	2017	15	Jharkhand	Singhbhum	386.91	669.97	1731.62	0.00	0.00	0.0	0.0	0.0	0.0

jupyter Vaishnavi Project Last Checkpoint: an hour ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel) O

File + ⚙️ ↕ Run █ C ➡️ Code

```
MAIZE YIELD (Kg per ha), FINGER MILLET AREA (1000 ha),
'FINGER MILLET PRODUCTION (1000 tons)',
'FINGER MILLET YIELD (Kg per ha)', 'BARLEY AREA (1000 ha)',
'BARLEY PRODUCTION (1000 tons)', 'BARLEY YIELD (Kg per ha)',
'CHICKPEA AREA (1000 ha)', 'CHICKPEA PRODUCTION (1000 tons)',
'CHICKPEA YIELD (Kg per ha)', 'PIGEONPEA AREA (1000 ha)',
'PIGEONPEA PRODUCTION (1000 tons)', 'PIGEONPEA YIELD (Kg per ha)',
'MINOR PULSES AREA (1000 ha)', 'MINOR PULSES PRODUCTION (1000 tons)',
'MINOR PULSES YIELD (Kg per ha)', 'GROUNDNUT AREA (1000 ha)',
'GROUNDNUT PRODUCTION (1000 tons)', 'GROUNDNUT YIELD (Kg per ha)',
'SESAMUM AREA (1000 ha)', 'SESAMUM PRODUCTION (1000 tons)',
'SESAMUM YIELD (Kg per ha)', 'RAPESEED AND MUSTARD AREA (1000 ha)',
'RAPESEED AND MUSTARD PRODUCTION (1000 tons)',
'RAPESEED AND MUSTARD YIELD (Kg per ha)', 'SAFFLOWER AREA (1000 ha)',
'SAFFLOWER PRODUCTION (1000 tons)', 'SAFFLOWER YIELD (Kg per ha)',
'CASTOR AREA (1000 ha)', 'CASTOR PRODUCTION (1000 tons)',
'CASTOR YIELD (Kg per ha)', 'JUTE AREA (1000 ha)',
```

Home Page - Select or create a x Vaishnavi Project - Jupyter Note x Vaishnavi Project x +

localhost:8888/notebooks/Vaishnavi%20Project.ipynb#

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. [Don't show anymore](#)

jupyter Vaishnavi Project Last Checkpoint: an hour ago (autosaved) Logout

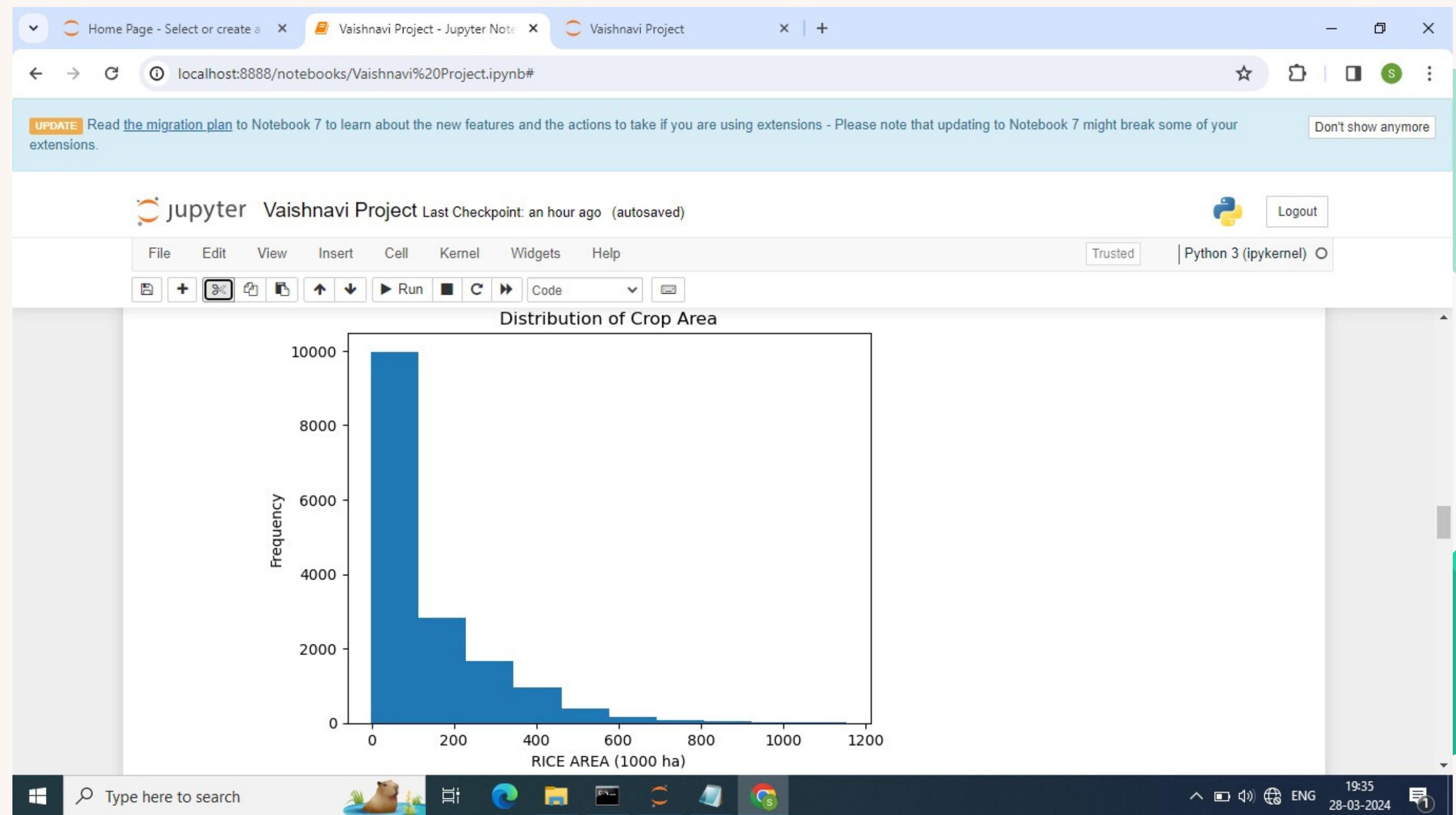
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

OILSEEDS AREA (1000 ha), 'OILSEEDS PRODUCTION (1000 tons)',
'OILSEEDS YIELD (Kg per ha)', 'SUGARCANE AREA (1000 ha)',
'SUGARCANE PRODUCTION (1000 tons)', 'SUGARCANE YIELD (Kg per ha)',
'COTTON AREA (1000 ha)', 'COTTON PRODUCTION (1000 tons)',
'COTTON YIELD (Kg per ha)', 'FRUITS AREA (1000 ha)',
'VEGETABLES AREA (1000 ha)', 'FRUITS AND VEGETABLES AREA (1000 ha)',
'POTATOES AREA (1000 ha)', 'ONION AREA (1000 ha)',
'FODDER AREA (1000 ha)'],
dtype='object')

In [11]: df['RICE AREA (1000 ha)'].value_counts()

Out[11]: RICE AREA (1000 ha)
0.00 1094
0.10 128
0.20 73
0.01 58
2.00 56
...
321.22 1
316.95 1
313.18 1
316.18 1
386.91 1
Name: count, Length: 8447, dtype: int64

Type here to search 19:34 ENG 28-03-2024



Home Page - Select or create a Vaishnavi Project - Jupyter Note Vaishnavi Project

localhost:8888/notebooks/Vaishnavi%20Project.ipynb#

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. [Don't show anymore](#)

jupyter Vaishnavi Project Last Checkpoint: an hour ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

In [19]: #2.Crop-specific Analysis

In [20]: rice_data = df[df['RICE AREA (1000 ha)'] == 'Rice']
wheat_data = df[df['RICE AREA (1000 ha)'] == 'Wheat']
pulses_data = df[df['RICE AREA (1000 ha)'].isin(['Pulses', 'Lentils', 'Chickpeas'])]

In [23]: rice_production_trend = rice_data.groupby('Year')['RICE PRODUCTION (1000 tons)').sum()

In [24]: rice_production_trend

Out[24]: Series([], Name: RICE PRODUCTION (1000 tons), dtype: float64)

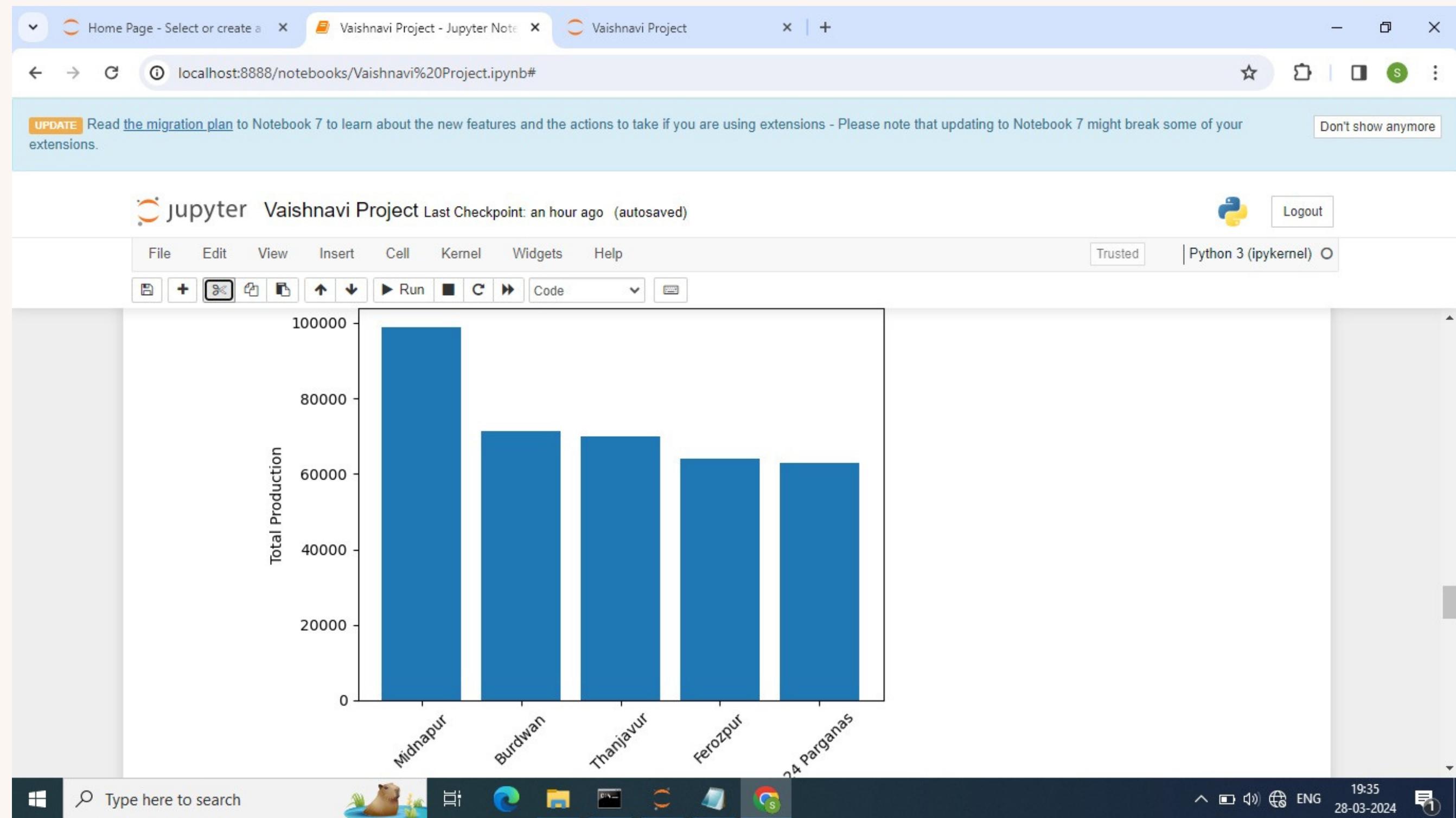
In [26]: district_disparities = df.groupby('Dist Name')['RICE PRODUCTION (1000 tons)').sum().sort_values(ascending=False)
print(district_disparities.head())

Dist Name	RICE PRODUCTION (1000 tons)
Midnapur	98868.28
Burdwan	71422.15
Thanjavur	69852.54
Ferozpur	64075.10
24 Parganas	63011.99

Name: RICE PRODUCTION (1000 tons), dtype: float64

Type here to search

19:35 28-03-2024



Home Page - Select or create a × Vaishnavi Project - Jupyter Note × Vaishnavi Project × +

localhost:8888/notebooks/Vaishnavi%20Project.ipynb#

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. [Don't show anymore](#)

jupyter Vaishnavi Project Last Checkpoint: an hour ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

In [33]:
kharif_data = df[df['KHARIF SORGHUM AREA (1000 ha)'] == 'Kharif']
rabi_data = df[df['RABI SORGHUM AREA (1000 ha)'] == 'Rabi']

In [34]: kharif_data

Out[34]:

Dist Code	Year	State Code	State Name	Dist Name	RICE AREA (1000 ha)	RICE PRODUCTION (1000 tons)	RICE YIELD (Kg per ha)	WHEAT AREA (1000 ha)	WHEAT PRODUCTION (1000 tons)	SUGARCANE ...	COTTON AREA (1000 ha)	COTTON PRODUCTION (1000 tons)	COTTON YIELD (Kg per ha)	FRUITS AREA (1000 ha)	VEGETABLE AREA
0 rows × 80 columns															

In [35]: rabi_data

Out[35]:

Dist Code	Year	State Code	State Name	Dist Name	RICE AREA (1000 ha)	RICE PRODUCTION (1000 tons)	RICE YIELD (Kg per ha)	WHEAT AREA (1000 ha)	WHEAT PRODUCTION (1000 tons)	SUGARCANE ...	COTTON AREA (1000 ha)	COTTON PRODUCTION (1000 tons)	COTTON YIELD (Kg per ha)	FRUITS AREA (1000 ha)	VEGETABLE AREA
0 rows × 80 columns															

Type here to search 19:35 28-03-2024

Home Page - Select or create a Vaishnavi Project - Jupyter Note Vaishnavi Project

localhost:8888/notebooks/Vaishnavi%20Project.ipynb#

UPDATE Read the [migration plan](#) to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. [Don't show anymore](#)

jupyter Vaishnavi Project Last Checkpoint: an hour ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

In [36]:
kharif_production = kharif_data.groupby('RICE AREA (1000 ha)')['RICE PRODUCTION (1000 tons)'].sum()
rabi_production = rabi_data.groupby('RICE AREA (1000 ha)')['RICE PRODUCTION (1000 tons)'].sum()

In [37]: kharif_production

Out[37]: Series([], Name: RICE PRODUCTION (1000 tons), dtype: float64)

In [38]: rabi_production

Out[38]: Series([], Name: RICE PRODUCTION (1000 tons), dtype: float64)

In [39]:
plt.plot(kharif_production, label='Kharif')
plt.plot(rabi_production, label='Rabi')
plt.xlabel('Crop')
plt.ylabel('Production')
plt.title('Seasonal Crop Production')
plt.legend()
plt.xticks(rotation=45)
plt.show()

Seasonal Crop Production

Type here to search

19:36 28-03-2024 ENG

Thank
you very
much!

