Project For Newbieron Technologies

Name of project : - Statewise production of crop in India

Steps for the project

1. Data Scrapping
2. Data Cleaning
3. Data modelling
4. Data Visualization through Python
5. Data Visualization through Tableau
6. Dashboard
7. Conclusion

About the topic

Statewise Production of crop in India.

It is basically about the comparison of states of production of its crop in India with its comparison with its area too. As we all know , India is having different weather conditions in different states for e.g. Jammu and Kashmir is having cold weather and Tamil Nadu is mostly hot while the north east is having most of the monsoon. Also , India is having different tropical condition which also affect the crop production for e.g. Uttarakhand and Himachal Pradesh are having mountainous region while Uttar Pradesh is having plains region . Also , western ghats and eastern ghats are mostly tropical region and having different weather type and hence eventually affect the production of crops .

Hence, in this topic, we have to visualize the data that how different crops are being produced in different states in India and also , which types of crop are mostly producted while which crops are least produced in India.

Data Scrapping

I have searched all over the internet to get the data and finally found kaggle websites the best to provise me the required data. It is very easy to download the required data and the following data was in downloaded in excel format.

The excel data was having 8 columns and 246090 rows . The columns were namely Index, state_name, District name, crop , crop year , season , area and production.

The data shows production of crop with states as well as district wise. The data also shows different season as well as area of different states and crops.

Data cleaning

First , I have opened my Jupyter note and upload the data in csv fprmat and then uploading the data in python workbook to read the data. I have imported required libraries such as Pandas , numpy and matplotlib . Pandas is used to read the data while matplotlib is used to visualize data.

First , I have hea d command to see the required data and found that index column to be the useless one so I have drop the index column. Then , I have checked its shape and data type to execute the data. Then I have checked that is there

any missing value in the data which might change the data and found that there are some missing value in the data and then I have cleaned the data and hence checked again to as to confirm there is no missing data left.

Data Modelling

Data modeling is the process of creating a visual representation of either a whole information system or parts of it to communicate connections between data points and structures. In this , I have checked that how data is related to each other so as to analyse the data at its best and get the conclusion.

I have found that the state are related to district and its area and production of crops. As well as there is another data which tells about the different season with crop type and production.

Hence we found that there can be four relation which can setup i.e. state name and crop production, state name and area, season and production of crop and statewise and crop wise production.

Data Visualization with python

After cleaning and modeling the data , we have to visualize the data that so as to get the best insights from the data . In this , we have used matplotlib pyplot library to make the required the charts of above mentioned graphs.

In this bar charts , we got the relationship between statewise production of crops, states with its area in descending order and season with the production of crops.

Data Visualization with Tableau

The data is then transferred to tableau to get the best visualization. In this we have first uploaded the data in tableau and then make the bar charts of Statewise production of crops in India , state with its area, season and production of crops in India and crop and its production.

I have make another sheet in which I have club the state and crops with its production. I have used dropdown options so as to visualize the data i.e. we can select the select the state and see the production of different crops in that particular state for e.g. we have select Uttar Pradesh as a

state and visualize the production of different crops . We got to know that sugarcane is produced most in uttar pradesh.

Also, we can visualize the crops and its production in different states for e.g. we have select bajra as a crop and visualize its production in different states and we got to know that Bajra is mostly produced in Rajasthan following with Uttar Pradesh.

Dashboard

After visualizing different bar charts of different relation , I have then merged the the following sheets to get the dashboard so as to visualize the whole data in one single sheet .

Conclusion

In this following project, we can conclude that the production of crops varies differently in India . Different states have different climate to produce the different crops.

After visualization , we found that coconut is produced mostly In India following with sugarcane and rice while production of bitter gourd , turnip is least. Also , production of ash gourd , beet root is nil .

Analyzing the states, we conclude that Kerala is best state in production of crops following with Andhra Pradesh and Tamil Nadu while Chandigarh and Mizoram has least production.

If we talk about area then we can say that Uttar Pradesh has the largest area following with Madhya Pradesh and Maharashtra while Chandigarh has the least area.

If we analyze the season then the crops produced in whole year has the best largest production following with Kharif and rabi crops while crops are least produced in Autumn season.

crop_production

September 25, 2023

[1]:
```
import numpy as np
```

[2]:
```
import pandas as pd
```

[6]:
```
crop = pd.read_csv("crop_production.csv")
```

[7]:
```
crop.head()
```

[7]:

	index	State_Name	District_Name	Crop_Year	Season	Crop	Area	Production
0	0	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Arecanut	1254.0	2000.0
1	1	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Other Kharif pulses	2.0	1.0
2	2	Andaman and Nicobar Islands	NICOBARS	2000	Kharif	Rice	102.0	321.0
3	3	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Banana	176.0	641.0
4	4	Andaman and Nicobar Islands	NICOBARS	2000	Whole Year	Cashewnut	720.0	165.0

[9]:
```
crop = crop.drop("index", axis=1)
```

[10]:
```
crop
```

[10]:

	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	279.0	547.0

```
246090                 West Bengal       PURULIA       2014  Winter

                       Crop      Area  Production
0                   Arecanut    1254.0      2000.0
1       Other Kharif pulses       2.0         1.0
2                       Rice     102.0       321.0
3                     Banana     176.0       641.0
4                  Cashewnut     720.0       165.0
…                        …       …          …
246086                  Rice     306.0       801.0
246087               Sesamum     627.0       463.0
246088             Sugarcane     324.0     16250.0
246089                  Rice  279151.0    597899.0
246090               Sesamum     175.0        88.0

[246091 rows x 7 columns]
```

[11]:
```
crop.shape
```

[11]:
```
(246091, 7)
```

[12]:
```
crop.info
```

[12]:
```
<bound method DataFrame.info of                          State_Name
District_Name  Crop_Year       Season  \
0       Andaman and Nicobar Islands      NICOBARS       2000  Kharif
1       Andaman and Nicobar Islands      NICOBARS       2000  Kharif
2       Andaman and Nicobar Islands      NICOBARS       2000  Kharif
3       Andaman and Nicobar Islands      NICOBARS       2000  Whole Year
4       Andaman and Nicobar Islands      NICOBARS       2000  Whole Year
…                              …           …         …          …
246086                 West Bengal       PURULIA       2014  Summer
246087                 West Bengal       PURULIA       2014  Summer
246088                 West Bengal       PURULIA       2014  Whole Year
246089                 West Bengal       PURULIA       2014  Winter
246090                 West Bengal       PURULIA       2014  Winter

                       Crop      Area  Production
0                   Arecanut    1254.0      2000.0
1       Other Kharif pulses       2.0         1.0
2                       Rice     102.0       321.0
3                     Banana     176.0       641.0
4                  Cashewnut     720.0       165.0
…                        …       …          …
246086                  Rice     306.0       801.0
246087               Sesamum     627.0       463.0
246088             Sugarcane     324.0     16250.0
```

```
246089                Rice  279151.0    597899.0
246090             Sesamum     175.0        88.0

[246091 rows x 7 columns]>
```

[13]:
```
crop.describe
```

[13]:
```
<bound method NDFrame.describe of                           State_Name
District_Name  Crop_Year       Season  \
0       Andaman and Nicobar Islands      NICOBARS       2000  Kharif
1       Andaman and Nicobar Islands      NICOBARS       2000  Kharif
2       Andaman and Nicobar Islands      NICOBARS       2000  Kharif
3       Andaman and Nicobar Islands      NICOBARS       2000  Whole Year
4       Andaman and Nicobar Islands      NICOBARS       2000  Whole Year
…                              …           …        …          …
246086                  West Bengal       PURULIA       2014  Summer
246087                  West Bengal       PURULIA       2014  Summer
246088                  West Bengal       PURULIA       2014  Whole Year
246089                  West Bengal       PURULIA       2014  Winter
246090                  West Bengal       PURULIA       2014  Winter

                      Crop      Area  Production
0                  Arecanut    1254.0      2000.0
1       Other Kharif pulses       2.0         1.0
2                      Rice     102.0       321.0
3                    Banana     176.0       641.0
4                 Cashewnut     720.0       165.0
…                       …       …          …
246086                 Rice     306.0       801.0
246087              Sesamum     627.0       463.0
246088           Sugarcane      324.0     16250.0
246089                 Rice  279151.0    597899.0
246090              Sesamum     175.0        88.0

[246091 rows x 7 columns]>
```

[15]:
```
crop.isna().sum(axis=0)
```

[15]:
```
State_Name          0
District_Name       0
Crop_Year           0
Season              0
Crop                0
Area                0
Production       3730
dtype: int64
```

[27]:
```
crop_cleaned = crop.dropna()
```

[28]:
```
crop_cleaned.isna().sum(axis=0)
```

[28]:
```
State_Name       0
District_Name    0
Crop_Year        0
Season           0
Crop             0
Area             0
Production       0
dtype: int64
```

[29]:
```
crop_cleaned.head()
```

[29]:

	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

[31]:
```
crop_cleaned.shape
```

[31]:
```
(242361, 7)
```

[33]:
```
import matplotlib.pyplot as plt
%matplotlib inline
```

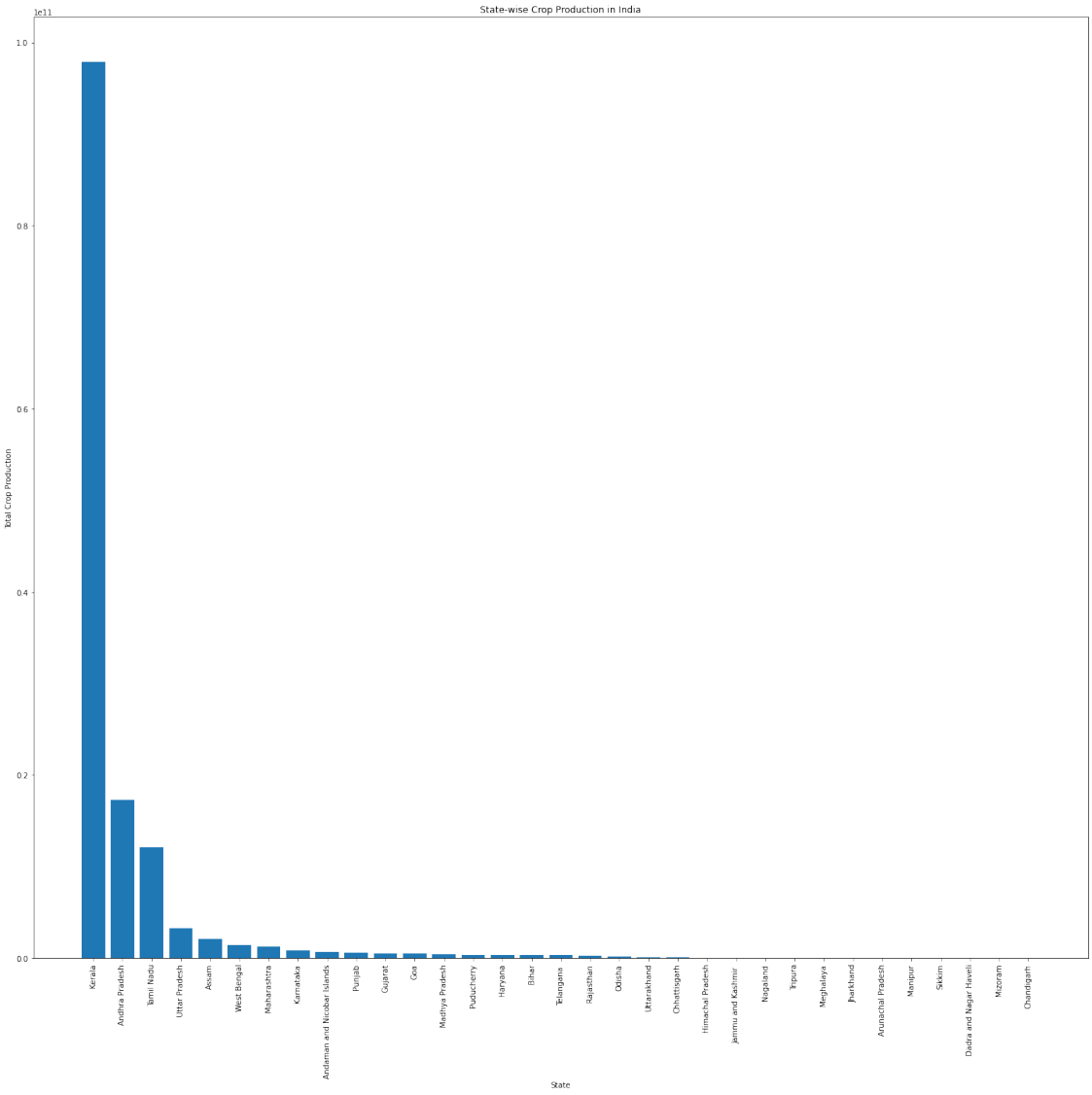
[35]:
```
crop_cleaned.columns
```

[35]:
```
Index(['State_Name', 'District_Name', 'Crop_Year', 'Season', 'Crop', 'Area',
       'Production'],
      dtype='object')
```

[36]:
```
statewise_production = crop_cleaned.groupby('State_Name')['Production'].sum().
↪reset_index()
```

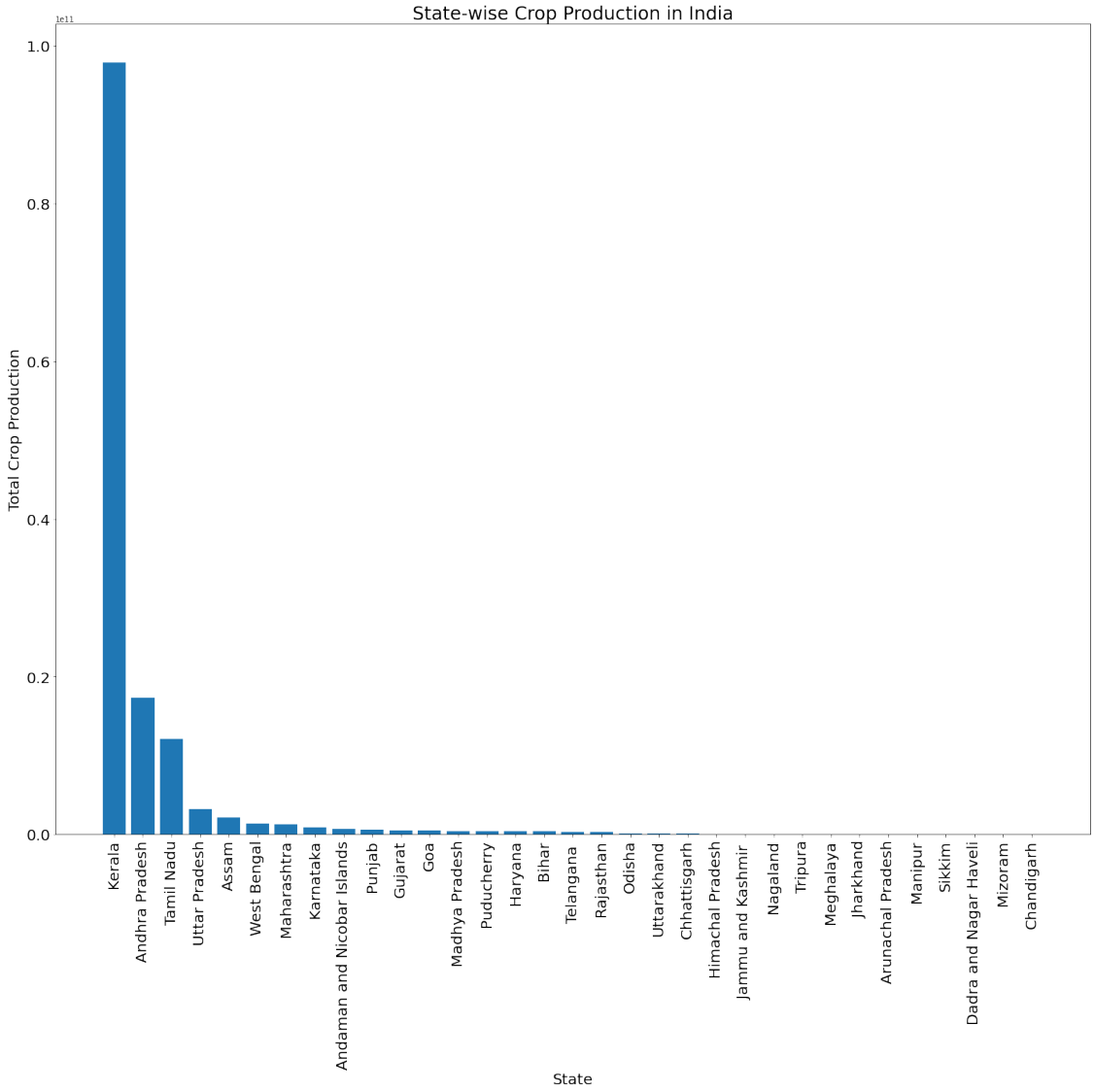
[37]:
```
statewise_production = statewise_production.sort_values(by='Production',␣
↪ascending=False)
```

[40]:
```
plt.figure(figsize=(20,20))
plt.bar(statewise_production['State_Name'], statewise_production['Production'])
plt.xlabel('State')
plt.ylabel('Total Crop Production')
plt.title('State-wise Crop Production in India')
plt.xticks(rotation=90)
plt.tight_layout()
plt.show()
```



[42]:
```
plt.figure(figsize=(20, 20))
plt.bar(statewise_production['State_Name'], statewise_production['Production'])
plt.xlabel('State', fontsize=20)
plt.ylabel('Total Crop Production', fontsize=20)
```

```
plt.title('State-wise Crop Production in India', fontsize=24)
plt.xticks(rotation=90, fontsize=20)
plt.yticks(fontsize=20)
plt.tight_layout()
plt.show()
```

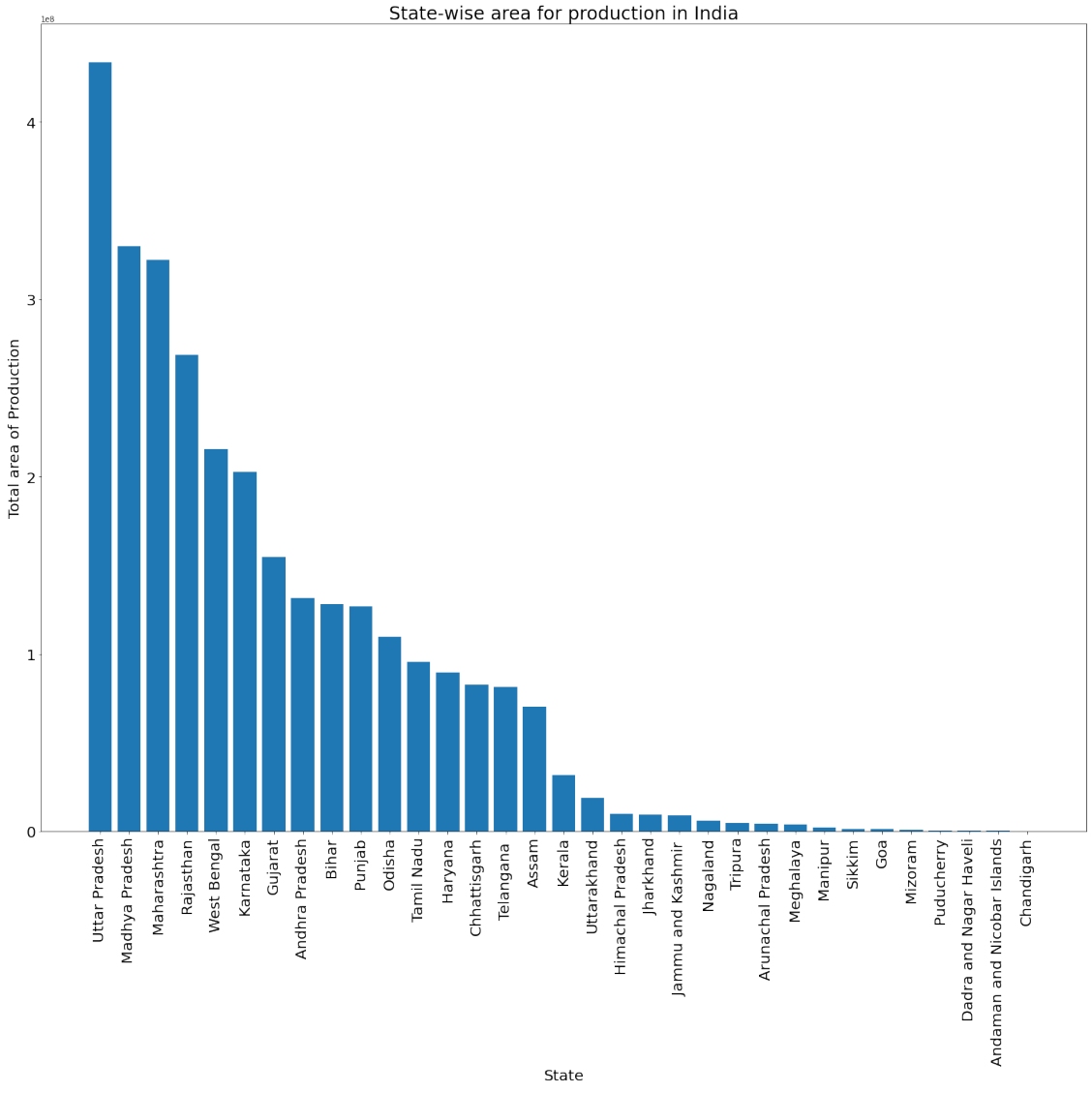


[44]:
```
area_production = crop_cleaned.groupby('State_Name')['Area'].sum().reset_index()
```

[45]:
```
area_production = area_production.sort_values(by='Area', ascending=False)
```

[46]:
```
plt.figure(figsize=(20, 20))
plt.bar(area_production['State_Name'], area_production['Area'])
plt.xlabel('State', fontsize=20)
```

```
plt.ylabel('Total area of Production', fontsize=20)
plt.title('State-wise area for production in India', fontsize=24)
plt.xticks(rotation=90, fontsize=20)
plt.yticks(fontsize=20)
plt.tight_layout()
plt.show()
```



[47]:
```
unique_crops_by_state = crop_cleaned.groupby('State_Name')['Crop'].nunique().
↪reset_index()
```

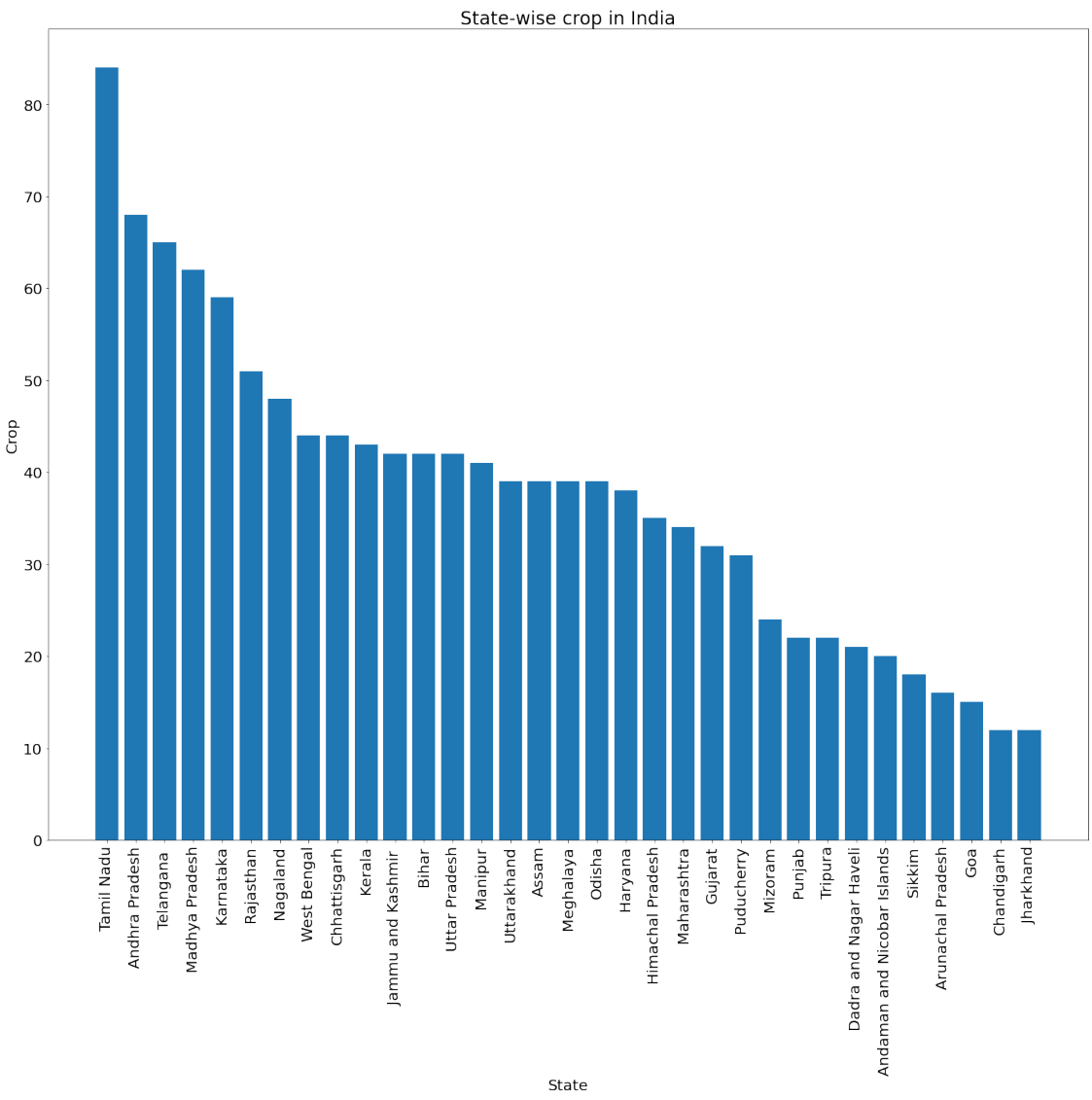
[48]:
```
print(unique_crops_by_state)
```

```
State_Name  Crop
```

```
0   Andaman and Nicobar Islands    20
1                Andhra Pradesh    68
2             Arunachal Pradesh    16
3                         Assam    39
4                         Bihar    42
5                    Chandigarh    12
6                  Chhattisgarh    44
7       Dadra and Nagar Haveli     21
8                           Goa    15
9                       Gujarat    32
10                      Haryana    38
11            Himachal Pradesh     35
12          Jammu and Kashmir      42
13                    Jharkhand    12
14                    Karnataka    59
15                       Kerala    43
16               Madhya Pradesh    62
17                  Maharashtra    34
18                      Manipur    41
19                    Meghalaya    39
20                      Mizoram    24
21                     Nagaland    48
22                       Odisha    39
23                   Puducherry    31
24                       Punjab    22
25                    Rajasthan    51
26                       Sikkim    18
27                   Tamil Nadu    84
28                   Telangana     65
29                      Tripura    22
30                Uttar Pradesh    42
31                  Uttarakhand    39
32                  West Bengal    44
```

[50]:
```
unique_crops_by_state = unique_crops_by_state.sort_values(by='Crop',␣
↪ascending=False)
```

[51]:
```
plt.figure(figsize=(20, 20))
plt.bar(unique_crops_by_state['State_Name'], unique_crops_by_state['Crop'])
plt.xlabel('State', fontsize=20)
plt.ylabel('Crop', fontsize=20)
plt.title('State-wise crop in India', fontsize=24)
plt.xticks(rotation=90, fontsize=20)
plt.yticks(fontsize=20)
plt.tight_layout()
plt.show()
```

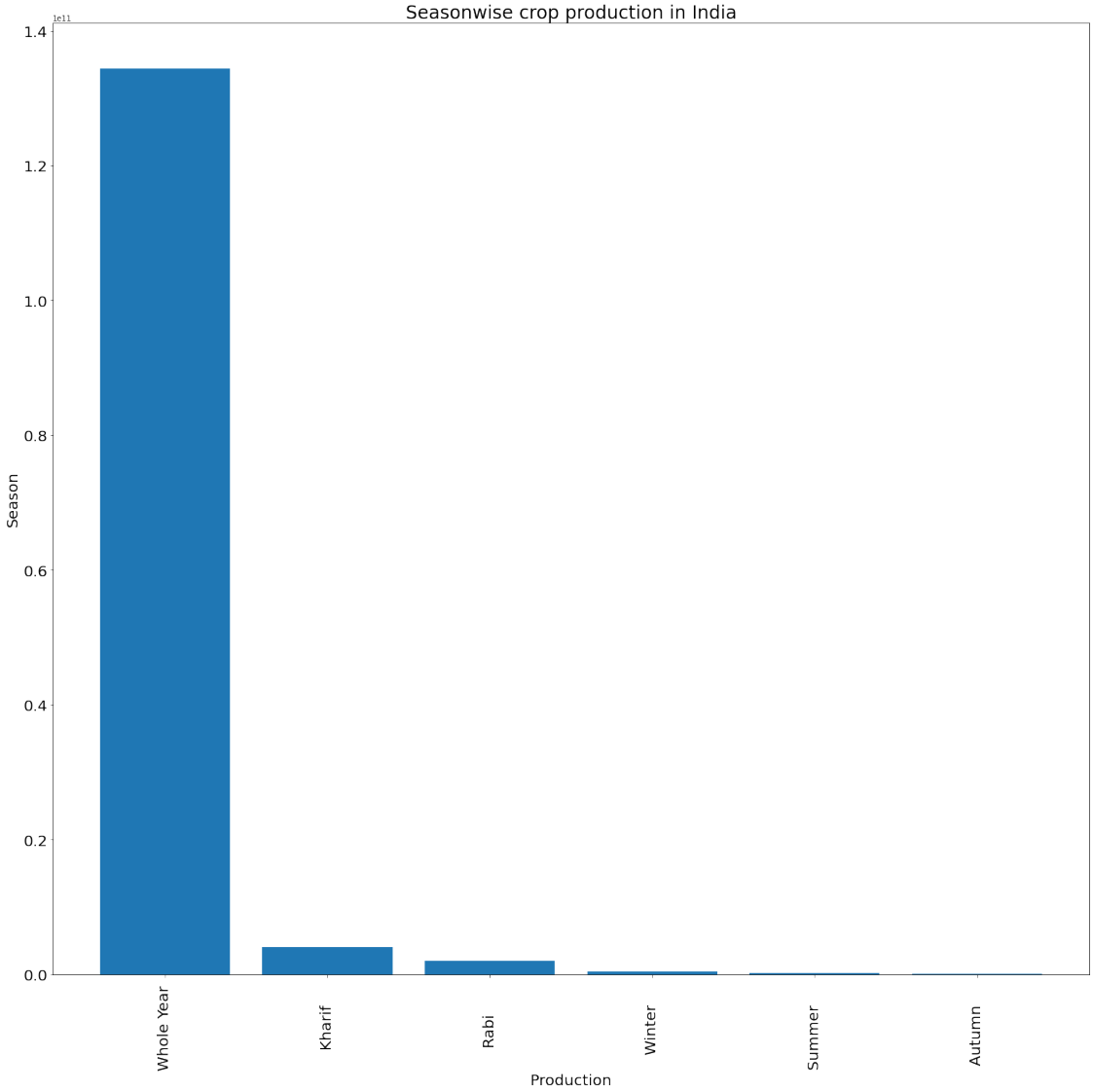


[52]:
```
croptype_production = crop_cleaned.groupby('Season')['Production'].sum().
↪reset_index()
```

[53]:
```
croptype_production = croptype_production.sort_values(by='Production',␣
↪ascending=False)
```

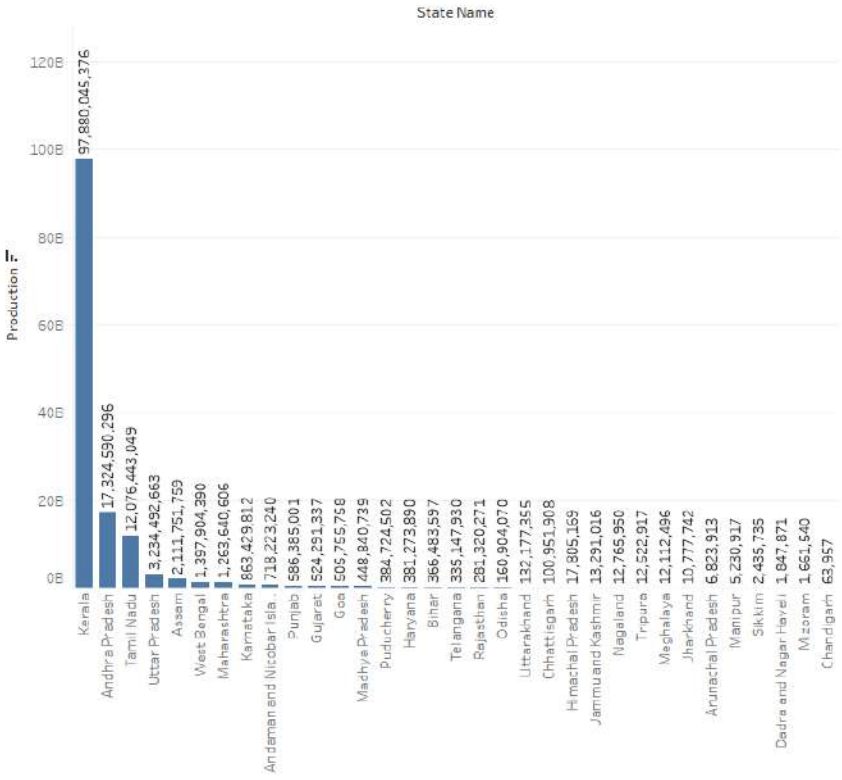
[54]:
```
plt.figure(figsize=(20, 20))
plt.bar(croptype_production['Season'], croptype_production['Production'])
plt.xlabel('Production', fontsize=20)
plt.ylabel('Season', fontsize=20)
plt.title('Seasonwise crop production in India', fontsize=24)
plt.xticks(rotation=90, fontsize=20)
plt.yticks(fontsize=20)
```

```
plt.tight_layout()
plt.show()
```



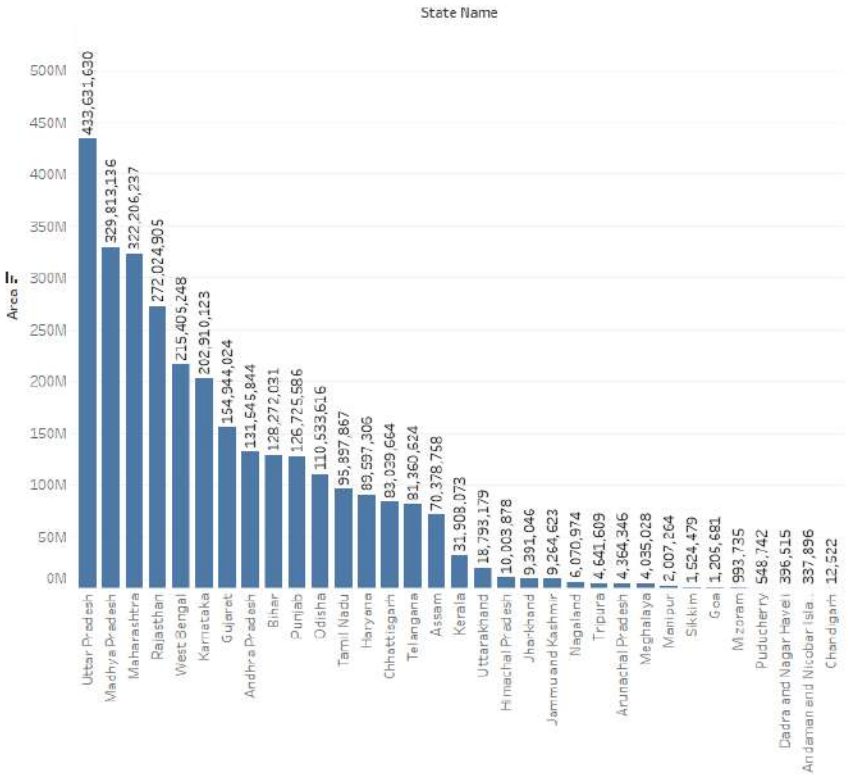
[ ]:

State wise production of crops in India



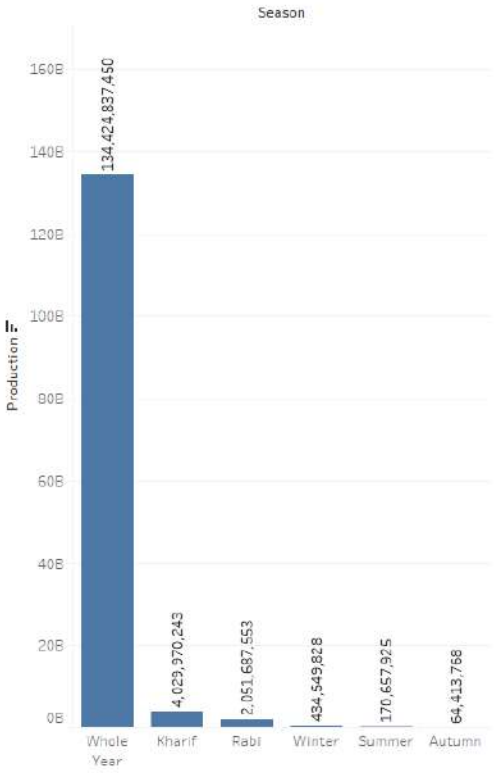
Sheet 1 | Sheet 2 | Sheet 3 | Sheet 4

Statewise area n India

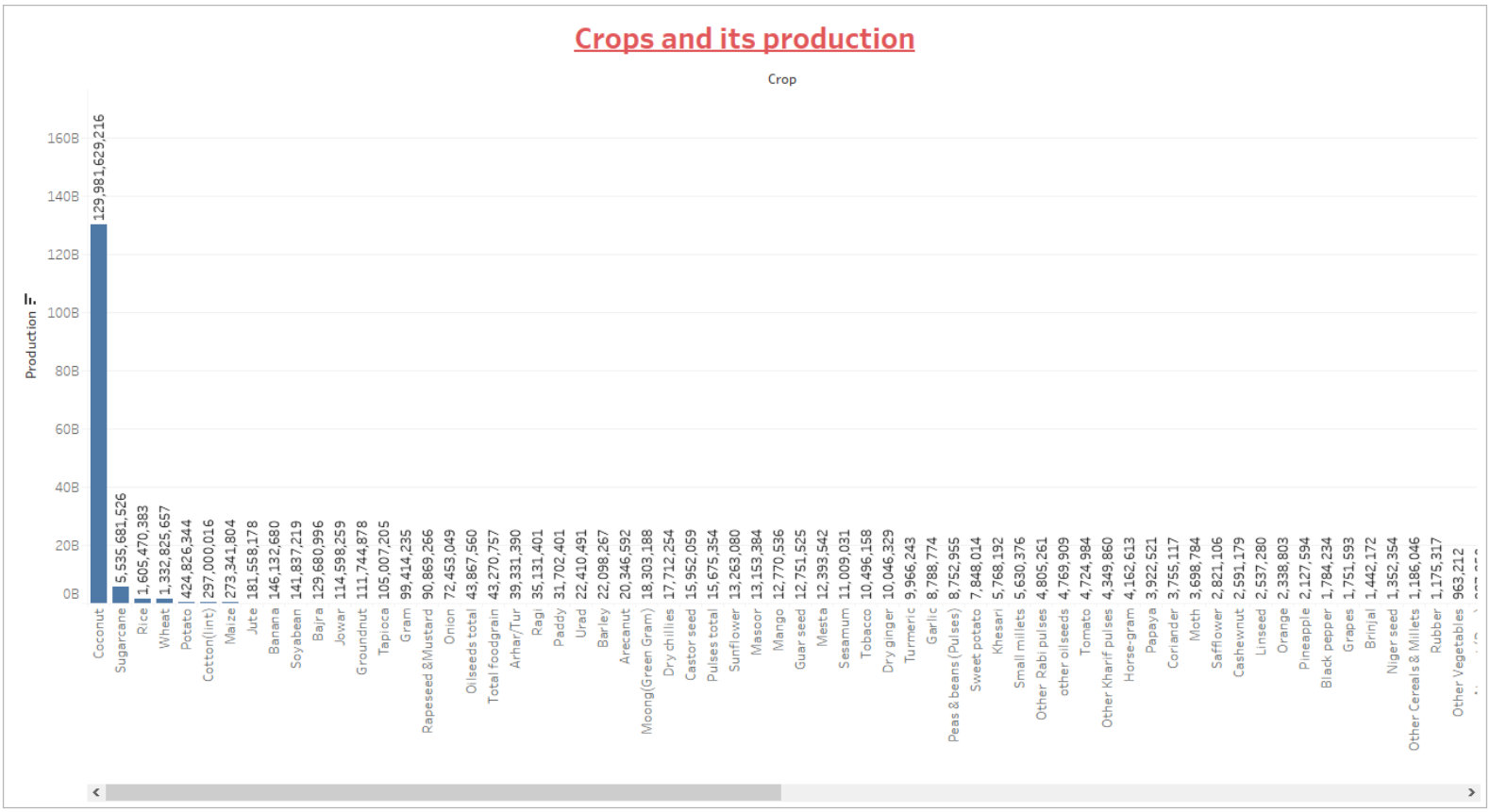


Sheet 1 | Sheet 2 | Sheet 3 | Sheet 4

Seasonwise production of crop

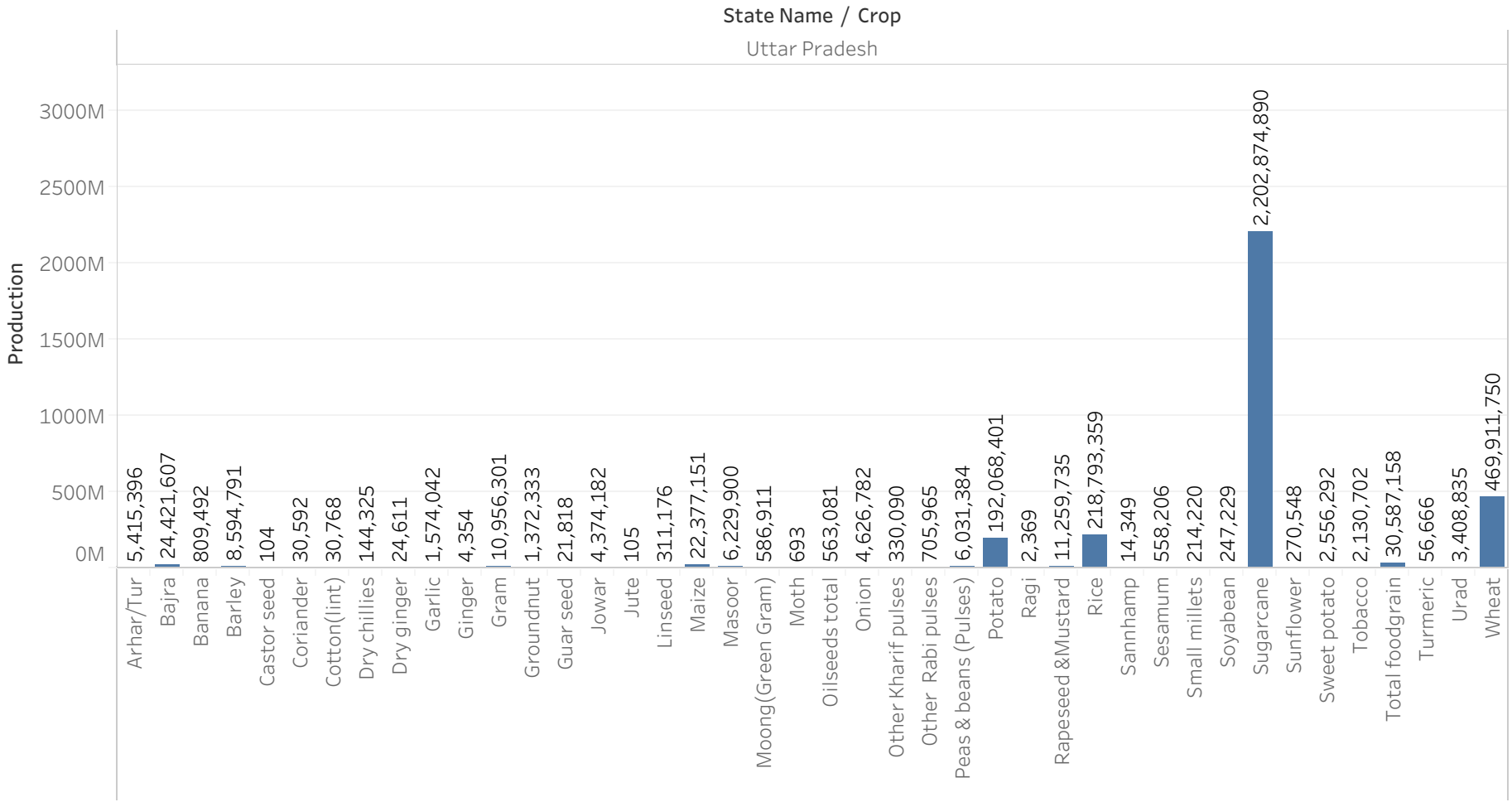


Crops and its production

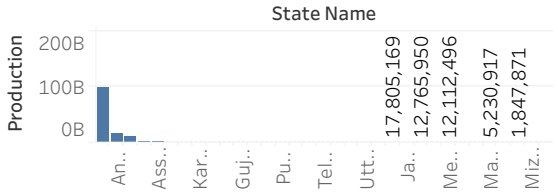


Sheet 1 | Sheet 2 | Sheet 3 | Sheet 4 | **Sheet 5** | Dashboard 1

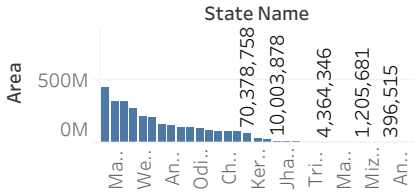
Statewise and Cropwise Prooduction in India



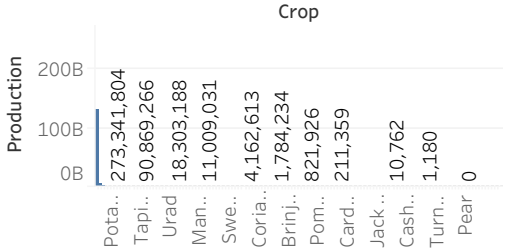
State wise production of crops in India



Statewise area in India



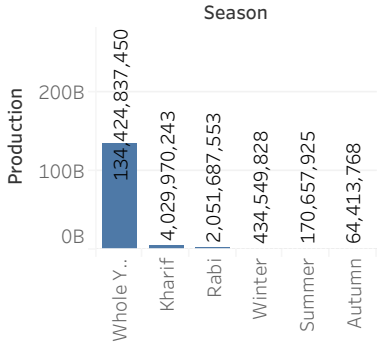
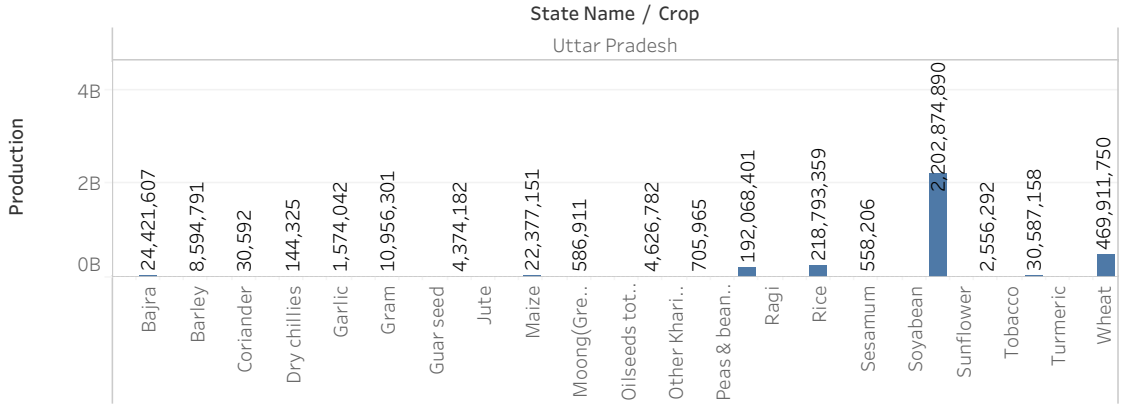
Crops and its production



State Name
Uttar Pradesh

Crop
All

Seasonwise production of crop



Statewise and Cropwise Prooduction in India