crcs-portal

June 28, 2023

AICTE Ministry of Corporation HACKATHON : CRCS portal dashboard.

DATA COLLECTION

[85]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline

# reading the dataset
df1  = pd.read_csv('table1.csv')
df2  = pd.read_csv('table2.csv')
df3  = pd.read_csv('table3.csv')
df4  = pd.read_csv('table4.csv')
```

[3]:
```
df1.head()
```

[3]:
```
   Sr. No.                                    Name of Society  \
0        1     saharayn universal multipurose society limited
1        2      JMJ Multi State Agro Cooperative Society Ltd.
2        3  Devbhoomi Agro Multi State Cooperative Society…
3        4  Aarya Multi Specialty Cooperative Hospital and…
4        5     Agriculture Development Cooperative Federation

                                             Address          State  \
0  At & P.O. Railway Colony, Gorakkhpur, District…  UTTAR PRADESH
1  TC 48/1442-1, Sreeniva, Parekkati House, Chath…         KERALA
2  S/o Sateesh Kumar, Tejupur, Haridwar, Uttarakh…    UTTARAKHAND
3  Gate No 165, Nimgaon (T), Madha, Solapur, 4132…    MAHARASHTRA
4  A/P 344, Lower Ground Floor, Sector 86, Preet …         PUNJAB

      District Date of Registration                       Area of Operation  \
0  KUSHI NAGAR                 NaN  Haryana, UttarPradesh, Himachal, Punjab
1          NaN          13/10/2022              Bihar, West bengal, Odisha
2          NaN           4/10/2022                   Tamil Nadu, karnataks
3          NaN           4/10/2022                         Haryana, Punajb
4          NaN          22/09/2022                    Manipur, Meghalaya

       Sector Type
```

```
0           Credit
1             Agro
2             Agro
3  Health/Hospital
4       Federation
```

[4]:
```
df2.head()
```

[4]:
```
   Sr. No.                                    Name of Society  \
0       25  Jewan Kothi Multi State Agro Cooperative Socie…
1       26  Bharathimanass Agriculture\nCooperative Ltd. (…
2       27  Youth Employability Skill Training\nCooperativ…
3       28  Jai Ratnagarbha Agriculture Multi State Cooper…
4       29  Shikharji Multi State Agro Processing and Mark…

                     Address          State   District Date of Registration  \
0              Uttar Pradesh  UTTAR PRADESH      BASTI              6/6/2022
1                    Manipur        MANIPUR  BISHNUPUR              3/6/2022
2             PATHANAMTHITTA         KERALA  ERNAKULAM              2/6/2022
3  Prayag Raj, uttar Pradesh  UTTAR PRADESH   BAREILLY            28/05/2022
4       Buldana, Maharashtra    MAHARASHTRA   BULDHANA            27/05/2022

                         Area of Operation Sector Type Unnamed: 8
0  Haryana, UttarPradesh, Himachal, Punjab        Agro        NaN
1               Bihar, West bengal, Odisha        Agro        NaN
2                    Tamil Nadu, karnataks      Others        NaN
3                          Haryana, Punajb        Agro        NaN
4                     Manipur, Meghalaya          Agro        NaN
```

[5]:
```
df3.head()
```

[5]:
```
   Sr. No.                                    Name of Society  \
0     51.0  Ramchandra Multi Speciality Cooperative Hospit…
1     52.0  Mudra Agriculture & Skill Development Multista…
2     53.0                The Bhuj Mercantile Coop. Bank Ltd
3     54.0   Heaven Multi State Agro Cooperative Society Ltd
4     55.0                     The Sutex Cooperative Bank Ltd

                                          Address          State  \
0  S.No. 2/2, Near SBI bank, Savakar Colony, Isla…    MAHARASHTRA
1  15-A, 3-4-757/22, APHB Building Near Raghavend…      TELANGANA
2  Vyapar Bhawan, Mithakhali 6 Roads, Ahmedabad-3…        GUJARAT
3  Kh No 403, Babhani Nagar, Andar, Sominath Naga…  UTTAR PRADESH
4  Surjaram, Bachkaniwala Bhawan, Near Navjivan C…        GUJARAT

    District Date of Registration                        Area of Operation  \
0     SANGLI           10/11/2021  Haryana, UttarPradesh, Himachal, Punjab
```

```
1   ADILABAD          12/04/2017              Bihar, West bengal, Odisha
2  AHMADABAD          14/09/2021                   Tamil Nadu, karnataks
3  MORADABAD          14/09/2021                         Haryana, Punajb
4     VALSAD          14/09/2021                    Manipur, Meghalaya

        Sector Type
0   Health/Hospital
1             Agro
2  Cooperative Bank
3             Agro
4  Cooperative Bank
```

[6]:
```
df4.head()
```

[6]:
```
   Sr. No.                                    Name of Society  \
0       73  Farm to Foreign Exports Entrepreneur Developme…
1       74  Dharitri Jute and Eco Friendly Products Multi …
2       75  The Malabar Multi State Agro Cooperative Socie…
3       76  Sai Raam Multi State Agri Cooperative Society Ltd
4       77  Prabhavana Multi State Womens' Jute and Allied…

                                              Address           State  \
0  H.No. 6-9, Gundugolanu(P.O), Bhimadole (Mandal…  ANDHRA PRADESH
1  D.No.12-2-42111, Alapati Nagar, Gudimalkapur, …       TELANGANA
2  1st Floor, Aiswarya Complex, Thavakkara, Kannu…          KERALA
3  No.6. (D-18), 6th Cross West Extension, Thilla…      TAMIL NADU
4  , House. No 2-2-647/G/18/1, Near Shivam Road, …       TELANGANA

        District Date of Registration  \
0  WEST GODAVARI           16/01/2020
1       ADILABAD           12/12/2019
2         KANNUR           20/10/2019
3     DHARMAPURI           06/12/2019
4       ADILABAD           27/09/2019

                         Area of Operation Sector Type
0  Haryana, UttarPradesh, Himachal, Punjab        Agro
1              Bihar, West bengal, Odisha        Agro
2                         Haryana, Punajb        Agro
3                       Manipur, Meghalaya        Agro
4         Maharashtra, Gujarat, Rajasthan        Agro
```

Exploratory Data Analysis

[7]:
```
# understanding the statistical information
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 24 entries, 0 to 23
Data columns (total 8 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Sr. No.               24 non-null     int64
 1   Name of Society       24 non-null     object
 2   Address               24 non-null     object
 3   State                 24 non-null     object
 4   District              15 non-null     object
 5   Date of Registration  23 non-null     object
 6   Area of Operation     24 non-null     object
 7   Sector Type           24 non-null     object
dtypes: int64(1), object(7)
memory usage: 1.6+ KB
```

[8]:
```
df2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 26 entries, 0 to 25
Data columns (total 9 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Sr. No.               26 non-null     int64
 1   Name of Society       26 non-null     object
 2   Address               26 non-null     object
 3   State                 26 non-null     object
 4   District              26 non-null     object
 5   Date of Registration  26 non-null     object
 6   Area of Operation     26 non-null     object
 7   Sector Type           26 non-null     object
 8   Unnamed: 8            1 non-null      object
dtypes: int64(1), object(8)
memory usage: 2.0+ KB
```

[9]:
```
df3.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 24 entries, 0 to 23
Data columns (total 8 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Sr. No.               22 non-null     float64
 1   Name of Society       22 non-null     object
 2   Address               22 non-null     object
 3   State                 22 non-null     object
 4   District              22 non-null     object
 5   Date of Registration  22 non-null     object
 6   Area of Operation     24 non-null     object
```

```
 7   Sector Type           22 non-null     object
dtypes: float64(1), object(7)
memory usage: 1.6+ KB
```

[10]:
```
df4.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 28 entries, 0 to 27
Data columns (total 8 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Sr. No.               28 non-null     int64
 1   Name of Society       28 non-null     object
 2   Address               28 non-null     object
 3   State                 28 non-null     object
 4   District              27 non-null     object
 5   Date of Registration  28 non-null     object
 6   Area of Operation     28 non-null     object
 7   Sector Type           27 non-null     object
dtypes: int64(1), object(7)
memory usage: 1.9+ KB
```

[11]:
```
# dimensions of the dataset
df1.shape
```

[11]:
```
(24, 8)
```

[12]:
```
df2.shape
```

[12]:
```
(26, 9)
```

[13]:
```
df3.shape
```

[13]:
```
(24, 8)
```

[14]:
```
df4.shape
```

[14]:
```
(28, 8)
```

[15]:
```
# observing the distribution of the response variables
round(df1['Sector Type'].value_counts(normalize = True)*100,2)
```

[15]:
```
Agro               50.00
Health/Hospital    20.83
Federation          8.33
Credit              4.17
Housing             4.17
Tourism             4.17
```

```
Fisheries           4.17
Construction        4.17
Name: Sector Type, dtype: float64
```

[16]:
```
round(df2['Sector Type'].value_counts(normalize = True)*100,2)
```

[16]:
```
Agro                  65.38
Others                11.54
Health/Hospital       11.54
Fisheries              3.85
Cooperative Bank       3.85
Industrial/Textile     3.85
Name: Sector Type, dtype: float64
```

[17]:
```
round(df3['Sector Type'].value_counts(normalize = True)*100,2)
```

[17]:
```
Agro                  54.55
Fisheries             13.64
Cooperative Bank       9.09
Marketing              9.09
Health/Hospital        4.55
Credit                 4.55
Industrial/Textile     4.55
Name: Sector Type, dtype: float64
```

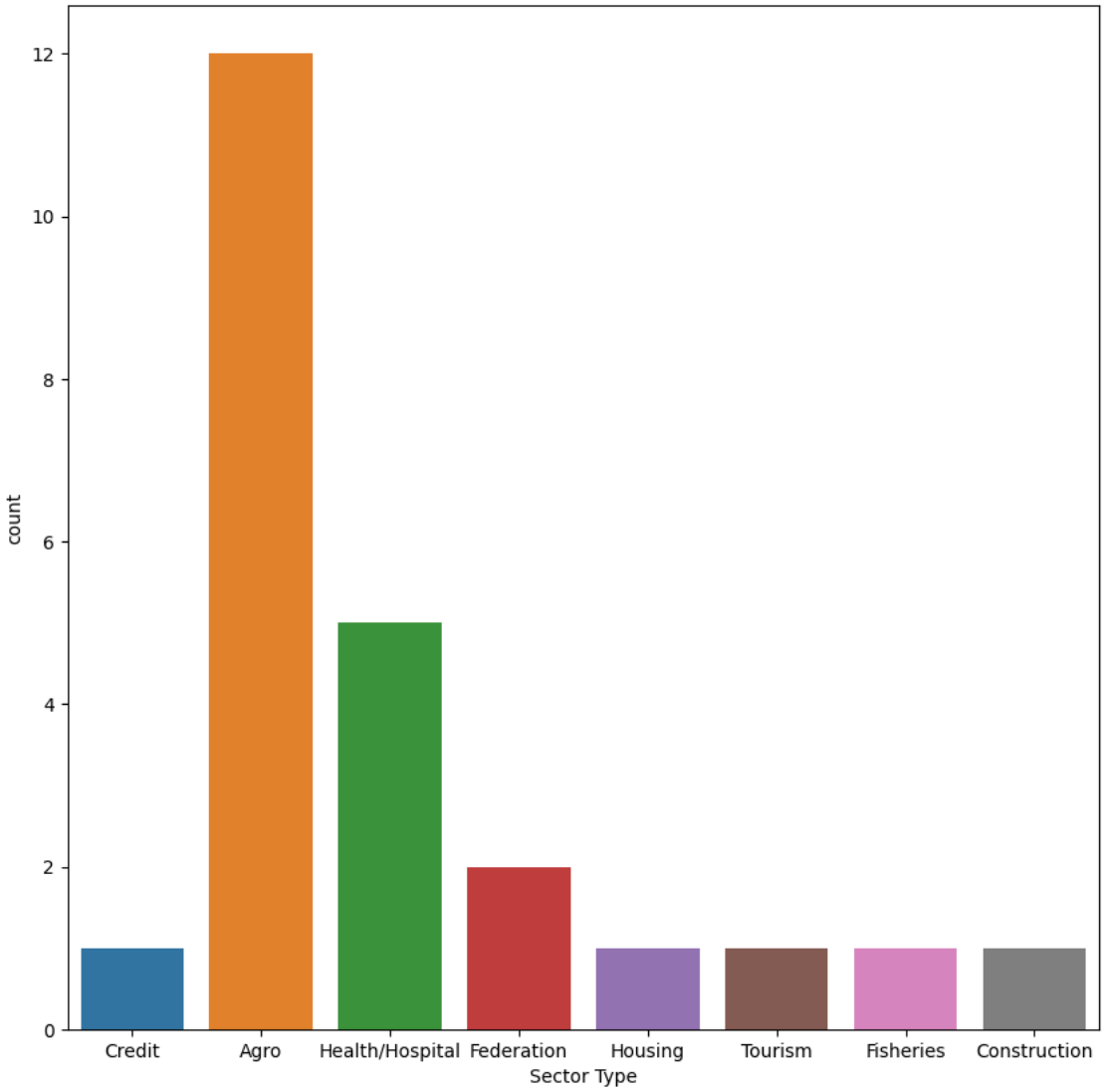
[18]:
```
round(df4['Sector Type'].value_counts(normalize = True)*100,2)
```

[18]:
```
Agro                40.74
Cooperative Bank    25.93
Housing             14.81
Credit              14.81
Dairy                3.70
Name: Sector Type, dtype: float64
```

Feature Engineering

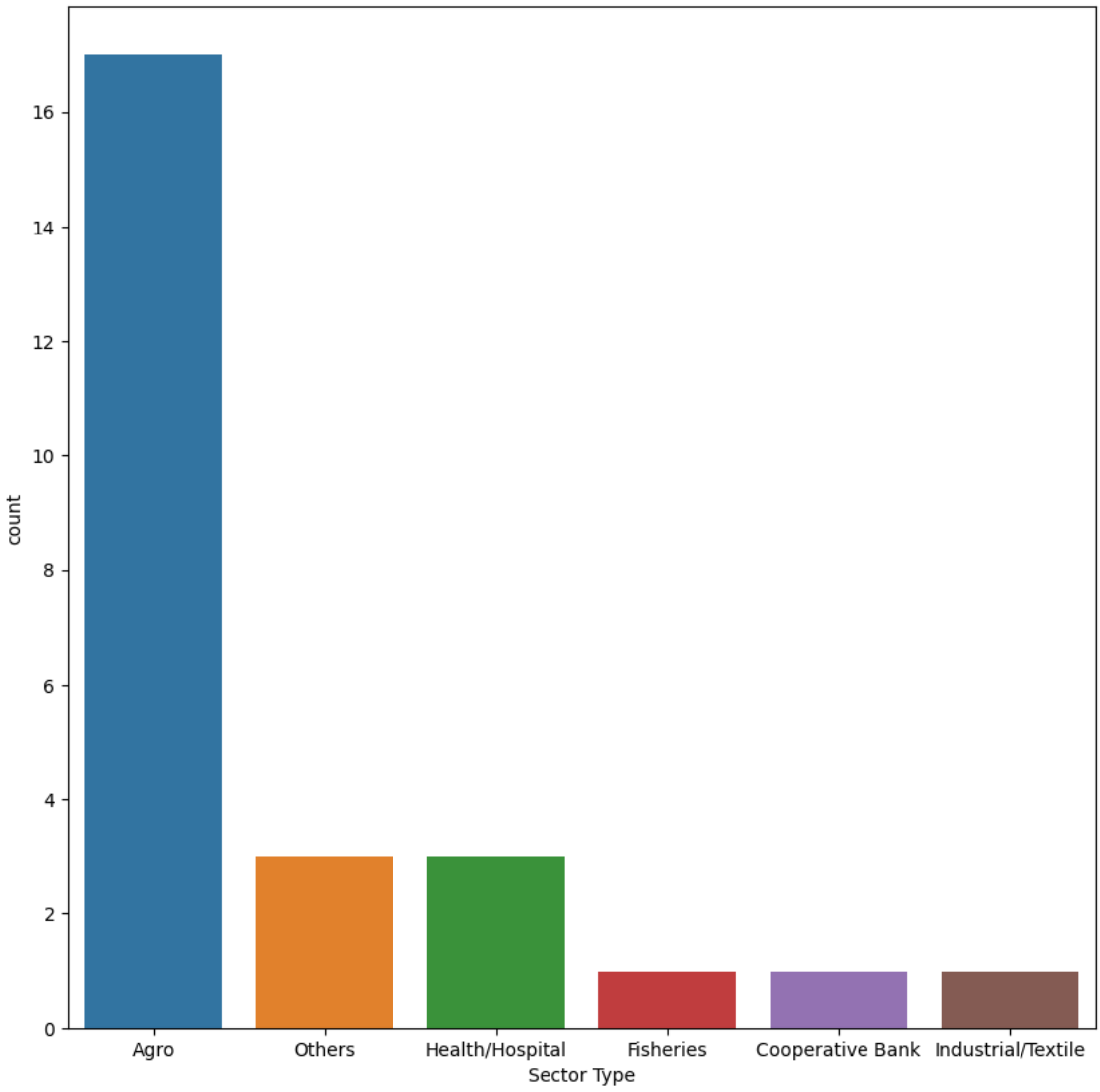
[21]:
```
import seaborn as sns
fig, ax = plt.subplots(figsize=(10, 10))
sns.countplot(x='Sector Type',data = df1 , ax=ax)
```

[21]:
```
<Axes: xlabel='Sector Type', ylabel='count'>
```



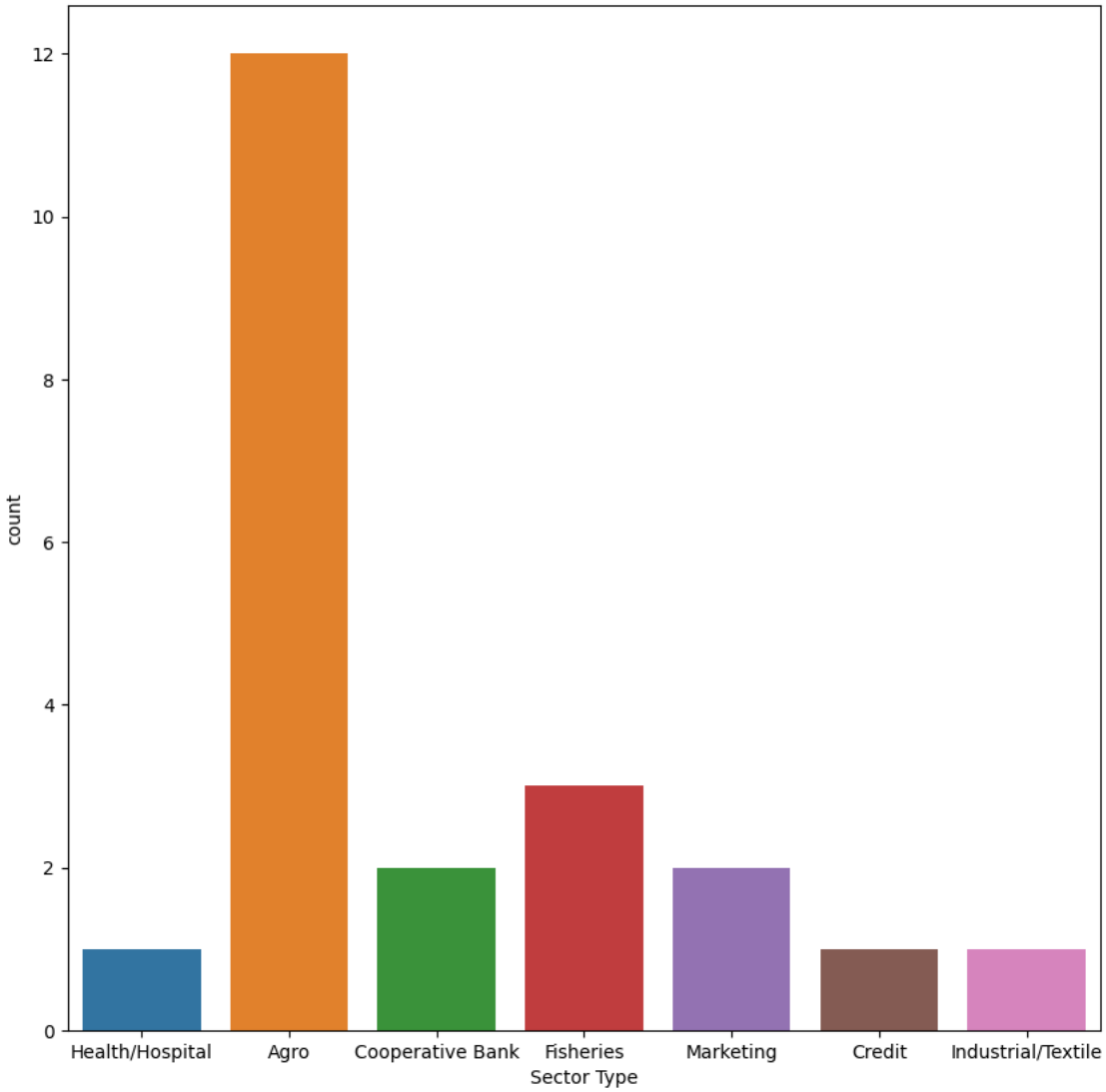
[23]:
```
fig, ax = plt.subplots(figsize=(10, 10))
sns.countplot(x='Sector Type',data = df2 , ax=ax)
```

[23]: <Axes: xlabel='Sector Type', ylabel='count'>



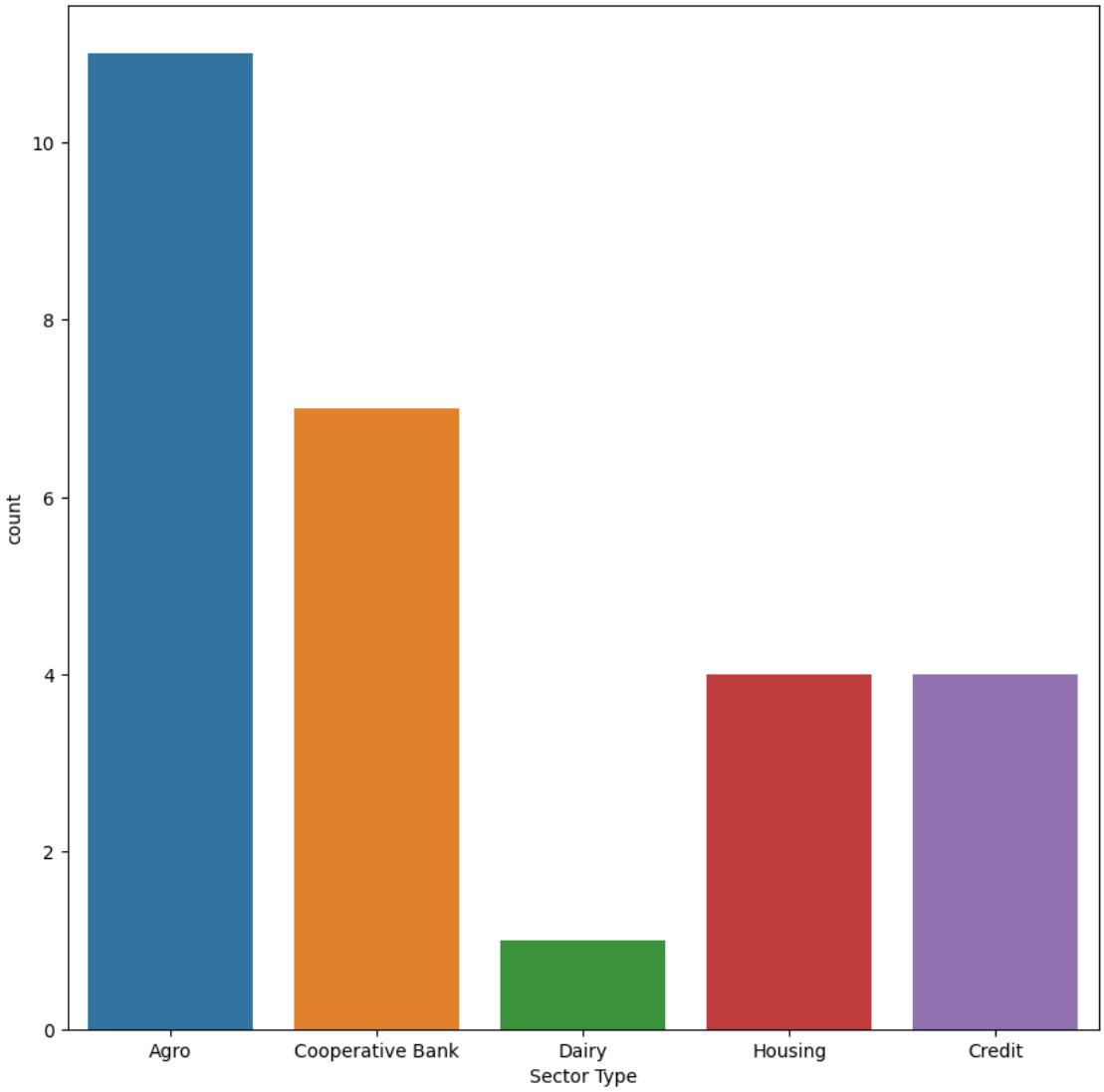
[24]:
```
fig, ax = plt.subplots(figsize=(10, 10))
sns.countplot(x='Sector Type',data = df3, ax=ax)
```

[24]: <Axes: xlabel='Sector Type', ylabel='count'>



[25]:
```
fig, ax = plt.subplots(figsize=(10, 10))
sns.countplot(x='Sector Type',data = df4, ax=ax)
```

[25]: <Axes: xlabel='Sector Type', ylabel='count'>



Data Cleaning by finding any missing values , null values.

[26]:
```
df1.isna()
```

[26]:

	Sr. No.	Name of Society	Address	State	District	Date of Registration	\
0	False	False	False	False	False	True	
1	False	False	False	False	True	False	
2	False	False	False	False	True	False	
3	False	False	False	False	True	False	
4	False	False	False	False	True	False	
5	False	False	False	False	True	False	
6	False	False	False	False	True	False	
7	False	False	False	False	True	False	

```
8     False            False    False  False      True                 False
9     False            False    False  False     False                 False
10    False            False    False  False     False                 False
11    False            False    False  False      True                 False
12    False            False    False  False     False                 False
13    False            False    False  False     False                 False
14    False            False    False  False     False                 False
15    False            False    False  False     False                 False
16    False            False    False  False     False                 False
17    False            False    False  False     False                 False
18    False            False    False  False     False                 False
19    False            False    False  False     False                 False
20    False            False    False  False     False                 False
21    False            False    False  False     False                 False
22    False            False    False  False     False                 False
23    False            False    False  False     False                 False

    Area of Operation  Sector Type
0               False        False
1               False        False
2               False        False
3               False        False
4               False        False
5               False        False
6               False        False
7               False        False
8               False        False
9               False        False
10              False        False
11              False        False
12              False        False
13              False        False
14              False        False
15              False        False
16              False        False
17              False        False
18              False        False
19              False        False
20              False        False
21              False        False
22              False        False
23              False        False
```

[27]:
```
df2.isna()
```

[27]:
```
    Sr. No.  Name of Society  Address  State  District  Date of Registration  \
0     False            False    False  False     False                 False
```

1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
5	False	False	False	False	False	False
6	False	False	False	False	False	False
7	False	False	False	False	False	False
8	False	False	False	False	False	False
9	False	False	False	False	False	False
10	False	False	False	False	False	False
11	False	False	False	False	False	False
12	False	False	False	False	False	False
13	False	False	False	False	False	False
14	False	False	False	False	False	False
15	False	False	False	False	False	False
16	False	False	False	False	False	False
17	False	False	False	False	False	False
18	False	False	False	False	False	False
19	False	False	False	False	False	False
20	False	False	False	False	False	False
21	False	False	False	False	False	False
22	False	False	False	False	False	False
23	False	False	False	False	False	False
24	False	False	False	False	False	False
25	False	False	False	False	False	False

	Area of Operation	Sector Type	Unnamed: 8
0	False	False	True
1	False	False	True
2	False	False	True
3	False	False	True
4	False	False	True
5	False	False	True
6	False	False	True
7	False	False	True
8	False	False	True
9	False	False	True
10	False	False	True
11	False	False	True
12	False	False	True
13	False	False	True
14	False	False	True
15	False	False	True
16	False	False	True
17	False	False	True
18	False	False	True
19	False	False	True

```
20              False        False       False
21              False        False        True
22              False        False        True
23              False        False        True
24              False        False        True
25              False        False        True
```

[28]:
```
df3.isna()
```

[28]:
```
    Sr. No.  Name of Society  Address  State  District  Date of Registration  \
0     False            False    False  False     False                 False
1     False            False    False  False     False                 False
2     False            False    False  False     False                 False
3     False            False    False  False     False                 False
4     False            False    False  False     False                 False
5     False            False    False  False     False                 False
6     False            False    False  False     False                 False
7     False            False    False  False     False                 False
8     False            False    False  False     False                 False
9     False            False    False  False     False                 False
10    False            False    False  False     False                 False
11    False            False    False  False     False                 False
12    False            False    False  False     False                 False
13    False            False    False  False     False                 False
14    False            False    False  False     False                 False
15    False            False    False  False     False                 False
16    False            False    False  False     False                 False
17    False            False    False  False     False                 False
18    False            False    False  False     False                 False
19    False            False    False  False     False                 False
20    False            False    False  False     False                 False
21    False            False    False  False     False                 False
22     True             True     True   True      True                  True
23     True             True     True   True      True                  True

    Area of Operation  Sector Type
0               False        False
1               False        False
2               False        False
3               False        False
4               False        False
5               False        False
6               False        False
7               False        False
8               False        False
9               False        False
10              False        False
```

```
11              False        False
12              False        False
13              False        False
14              False        False
15              False        False
16              False        False
17              False        False
18              False        False
19              False        False
20              False        False
21              False        False
22              False         True
23              False         True
```

[29]:
```
df4.isna()
```

[29]:
```
    Sr. No.  Name of Society  Address  State  District  Date of Registration  \
0     False            False    False  False     False                 False
1     False            False    False  False     False                 False
2     False            False    False  False     False                 False
3     False            False    False  False     False                 False
4     False            False    False  False     False                 False
5     False            False    False  False     False                 False
6     False            False    False  False     False                 False
7     False            False    False  False     False                 False
8     False            False    False  False     False                 False
9     False            False    False  False     False                 False
10    False            False    False  False     False                 False
11    False            False    False  False     False                 False
12    False            False    False  False      True                 False
13    False            False    False  False     False                 False
14    False            False    False  False     False                 False
15    False            False    False  False     False                 False
16    False            False    False  False     False                 False
17    False            False    False  False     False                 False
18    False            False    False  False     False                 False
19    False            False    False  False     False                 False
20    False            False    False  False     False                 False
21    False            False    False  False     False                 False
22    False            False    False  False     False                 False
23    False            False    False  False     False                 False
24    False            False    False  False     False                 False
25    False            False    False  False     False                 False
26    False            False    False  False     False                 False
27    False            False    False  False     False                 False

    Area of Operation  Sector Type
```

```
0               False       False
1               False       False
2               False       False
3               False       False
4               False       False
5               False       False
6               False       False
7               False       False
8               False       False
9               False       False
10              False       False
11              False       False
12              False       False
13              False       False
14              False       False
15              False       False
16              False       False
17              False       False
18              False       False
19              False       False
20              False       False
21              False       False
22              False        True
23              False       False
24              False       False
25              False       False
26              False       False
27              False       False
```

ONE-HOT ENCODING for converting the categorical data into numberic type for data visualization and model evaluation.

[30]:
```
df2['State']=df2['State'].replace({'UTTAR PRADESH' : 1 ,'PATHANAMITHITTA':2,␣
 ↪'DELHI        ':3, 'MAHARASHTRA':4 , 'KASHMIR':5 , 'MADHYA PRADESH':6 ,␣
 ↪'TAMIL NADU':7 , 'HARYANA':8,'JAMMU AND KASHMIR':9, 'NEW DELHI':10,'MANIPUR':
 ↪11,'KERALA':12})
df2
df2['Sector Type']=df2['Sector Type'].replace({'Agro' : 1 ,'Health/Hospital':2,␣
 ↪'Others':3, 'Fisheries':4 , 'Cooperative Bank':5 , 'Industrial/Textile':6})
df2
df2['District']=df2['District'].replace({'BASTI':1, 'BISHNUPUR':2,'ERNAKULAM':
 ↪3,'BAREILLY':4, 'BULDHANA':5, 'NEW DELHI':6,'BEED':7, 'OSMANABAD':8,␣
 ↪'KOLHAPUR':9, 'SANGLI':10 , 'COIMBATORE':11, 'BADGAM':12, 'BULANDSHAHR':13,␣
 ↪'BHANDARA':14, 'NASIK':15, 'AHMEDNAGAR':16, 'NAGPUR':17, 'BHOPAL':18,␣
 ↪'SONIPAT':19 })
df2
```

[30]:

	Sr. No.	Name of Society	Address	State	District
0	25	Jewan Kothi Multi State Agro Cooperative Socie…	Uttar Pradesh	1	1
1	26	Bharathimanass Agriculture\nCooperative Ltd. (…	Manipur	11	2
2	27	Youth Employability Skill Training\nCooperativ…	PATHANAMTHITTA	12	3
3	28	Jai Ratnagarbha Agriculture Multi State Cooper…	Prayag Raj, uttar Pradesh	1	4
4	29	Shikharji Multi State Agro Processing and Mark…	Buldana, Maharashtra	4	5
5	30	News Agriculture Export & Marketing Cooperativ…	Kalkaji, Delhi	10	6
6	31	Krashnadi Multi State Vehicle Services Coopera…	Beed, maharashtra	4	7
7	32	Lotus Multi Specialty Cooperative Hospital and…	Osmanabad, Maharashtra	4	7
8	33	Keshavraj Multi Speciality Cooperative Hospita…	Osmanabad, Maharashtra	4	8
9	34	Keshavraj Multi Speciality Cooperative Hospita…	Kolhaour, Maharashtra	4	9
10	35	Vanashri Cooperative Power & Agro Progressive ltd	Sangli , Maharashtra	4	10
11	36	Samrudh Marathwada Multi State\nVehicle Servic…	Beed, maharashtra	4	7
12	37	Bharathiyajkl Machuwara Samudaay Cooperative L…	Baramulla, Kashmir	9	12
13	38	Prestige Agro Multi State Cooperative Society Ltd	Uttar Pradesh	1	13
14	39	Vidhata Hamal Multi State Cooperative Society Ltd	Bhandara, Maharashtra	4	14
15	40	Dr Appasahed Urf Sa. Re. Patil\nJayysinghpur U…	Kolhapur, maharashtra	4	9
16	41	Krushisadhna Mahila Multi State Agro Cooperati…	Nashik maharashtra	4	15
17	42	Seva Sadan Multi Specialty\nCooperative Hospit…	Gate no. 14, Chaitnya Complex, behind Tahsil k…	4	16
18	43	Bulandshahr Agro MultiState cooperative Societ…			
19	44	B K Dhanlakshmi Multi State Agro Cooperative S…			
20	45	The Walmik Multi State Cooperative Handloom We…			
21	46	Ente Naadu Multi State Agro Co- operative Soci…			
22	47	Rajsidhi Cooperative Cotton & Agro Industries Ltd			
23	48	Natural Agro Farming Multi State Cooperative S…			
24	49	Mansa Development Cooperative Society Ltd			
25	50	Om Agro Multi State Cooperative society Ltd			

```
18  Mohalla Mahadev, Galimpur Road, Dibai, Bulands…      1        13
19  Mayflower signature, Office 7B, 7th Floor, Avi…      7        11
20  Plot No 63, Aoodumbae, Gurudev nagar, Nandanav…      4        17
21  ST George High School, Aluva- Munnar Road, Kot…     12         3
22                    Ahmednagar 414401, Maharashtra      4        16
23       D77, S/F, East of Kailash, New delhi 110065     10         6
24  41 Sukh Sagar, phase 4 colony, near Mittal Col…      6        18
25            Ahulana (Gohana) Sonipat Haryana 131301      8        19

   Date of Registration                            Area of Operation  \
0              6/6/2022          Haryana, UttarPradesh, Himachal, Punjab
1              3/6/2022                       Bihar, West bengal, Odisha
2              2/6/2022                            Tamil Nadu, karnataks
3             28/05/2022                                  Haryana, Punajb
4             27/05/2022                               Manipur, Meghalaya
5             21/05/2022                  Maharashtra, Gujarat, Rajasthan
6              12/5/2022                               Gujarat, Rajasthan
7             27/04/2022        Maharashtra,Goa.Karnataka, Andhra Pradesh
8             27/04/2022                          Andhra Pradesh, Telangna
9             27/04/2022     Madhya Pradesh, Chhatisgarh, Jharkhan, Bihar
10            26/04/2022                  Maharashtra, Gujarat, Rajasthan
11            20/04/2022                               Gujarat, Rajasthan
12            18/04/2022                                   Haryana, Delhi
13             1/4/2022                    Haryana, Delhi, UttarPradesh,
14             1/4/2022    Uttrakhand, Himachal Pradesh, Jammu & Kashmir
15             1/4/2022                  Maharashtra, Gujarat, Rajasthan
16             1/4/2022                       Haryana, Punjab, Rajasthan
17            21/03/2022                               Haryana, Rajasthan
18            21/03/2022              Gujarat, Rajasthan, Haryana, Delhi
19             8/3/2022          Kerala, karnataka, Maharashtra, Gujarat
20            21/02/2022               Maharashtra, Karnataka, Tamil nadu
21             3/2/2022             Haryana, Gujarat, Maharashtra, Delhi
22             12/1/2022                Maharashtra, Gujarat, Rajasthan
23             12/1/2022          Kerala, karnataka, Maharashtra, Gujarat
24            12/10/2021                                   Haryana, Delhi
25            22/11/2021                                   Haryana, Delhi

    Sector Type Unnamed: 8
0             1        NaN
1             1        NaN
2             3        NaN
3             1        NaN
4             1        NaN
5             1        NaN
6             3        NaN
7             2        NaN
8             2        NaN
```

```
9             2        NaN
10            1        NaN
11            3        NaN
12            4        NaN
13            1        NaN
14            1        NaN
15            5        NaN
16            1        NaN
17            1        NaN
18            1        NaN
19            1        NaN
20            6          e
21            1        NaN
22            1        NaN
23            1        NaN
24            1        NaN
25            1        NaN
```

[31]:
```
df3['State']=df2['State'].replace({'UTTAR PRADESH' : 1 ,'GUJARAT':2,␣
↪'TELANGANA':3, 'MAHARASHTRA':4 , 'ASSAM':5 , 'MADHYA PRADESH':6 , 'TAMIL␣
↪NADU':7 , 'HARYANA':8,'JAMMU AND KASHMIR':9, 'NEW DELHI':10,'MANIPUR':
↪11,'BIHAR':12})
df3
df3['Sector Type']=df2['Sector Type'].replace({'Agro' : 1 ,'Health/Hospital':2,␣
↪'Marketing':3, 'Fisheries':4 , 'Cooperative Bank':5 , 'Industrial/Textile':
↪6, 'Credit':7})
df3
df3['District']=df2['District'].replace({'SANGLI':1, 'ADILABAD':2,'ADILABAD':
↪3,'AHMEDABAD':4, 'VALSAD':5, 'BATOD':6,'BISHNUPUR':7, 'DEORIA':8, 'BANDA':9,␣
↪'MUZAFFARNAGAR':10 , 'CHIRANG':11, 'NEW DELHI':12, 'BULANDSHAHR':13, 'BHIND':
↪14, 'BHOPAL':15, 'BAREILLY':16, 'JHANSI':17, 'SOLAPUR':18, 'CHENNAI':19,␣
↪'LUCKNOW':20, 'GURGAON':21, 'PATNA':22 })
df3
```

[31]:
```
    Sr. No.                                  Name of Society  \
0      51.0  Ramchandra Multi Speciality Cooperative Hospit…
1      52.0  Mudra Agriculture & Skill Development Multista…
2      53.0                 The Bhuj Mercantile Coop. Bank Ltd
3      54.0   Heaven Multi State Agro Cooperative Society Ltd
4      55.0                     The Sutex Cooperative Bank Ltd
5      56.0  The Navbharat Multi State Agro Farming and Mar…
6      57.0  BharathiMana Fishermen development Cooperative…
7      58.0  Agrovision Farmer Multi State Cooperative Soci…
8      59.0  Chitrakoot Multi State Agriculture Marketing C…
9      60.0   CP Agrotech Multi State Cooperative Society Ltd.
10     61.0  Simplydesi Selfhelp Products, Processing and M…
11     62.0  Bharati North East Fisherman Co-operative Ltd…
```

```
12     63.0      Chambal Agriculture Marketing Cooperative Ltd
13     64.0    Sharda Agro Multi State Cooperative Society Ltd
14     65.0        N E Railway Cooperative Credit Society Ltd,
15     66.0  Ramraja Multi state Agro Cooperative Society ltd.
16     67.0    Vitthal Rao Shinde Sahakari Sahkar Karkhana Ltd
17     68.0  Ambika Multi State Health Service for the Weak…
18     69.0   The UP Kisan Development cooperative Society Ltd
19     70.0  Farmtrade Agro Multi state Marketing Cooperati…
20     71.0  Bihar State handloom Weavers Cooperative Union…
21     72.0       Radhika Agricultural Cooperative Society Ltd
22      NaN                                               NaN
23      NaN                                               NaN

                                            Address  State  District  \
0   S.No. 2/2, Near SBI bank, Savakar Colony, Isla…      1         1
1   15-A, 3-4-757/22, APHB Building Near Raghavend…     11         2
2   Vyapar Bhawan, Mithakhali 6 Roads, Ahmedabad-3…     12         3
3   Kh No 403, Babhani Nagar, Andar, Sominath Naga…      1         4
4   Surjaram, Bachkaniwala Bhawan, Near Navjivan C…      4         5
5   Shop No 112, Floor-I, Sharda Arcade, Paliyad R…     10         6
6   Kha- Potshangham Maning, Leikai, PO & PS Bishn…      4         7
7   H.No 3, Gate No 1, Krishna Colony, Near Univer…      4         7
8   H.No 104, Village & Post Sahewa, thana-\nGirwa…      4         8
9   C/O Manoj Kumar, Mohalla-Kannon Goyan, Kasba- …      4         9
10  52, Ananda Park Extension, East Punjabi Bagh, …      4        10
11  C/O Swambha Basumatary, near Simlaguri Post Of…      4         7
12  525, rajput Nagar, Bharouli road, Bhind, Madhy…      9        12
13                            Bhopal, Madhya Pradesh      1        13
14  Chief Worshop Manager Office, NE Railway, Izza…      4        14
15  102, Eligent Tower, Royal City, Shivpuri Road,…      4         9
16  Gangamainagar, Post Pimpalner, Taluka Madha, S…      4        15
17  Sai Illam, Floor-2, 53/2, Bharatheshwarar,\n3r…      4        16
18  65, Bheem Nagar, Vijay Khand 2, Gomti Nagar, L…      1        13
19  B-92. Mayfiels Garden Sector 48, Gurugram- 122002      7        11
20      Handloom Bhawan, Rajendra Place, Patna, Bihar      4        17
21  Swami Puran Colony, Nai Basti, Jhansi 284002, …     12         3
22                                             NaN      4        16
23                                             NaN     10         6

   Date of Registration                          Area of Operation  \
0            10/11/2021       Haryana, UttarPradesh, Himachal, Punjab
1            12/04/2017                   Bihar, West bengal, Odisha
2            14/09/2021                        Tamil Nadu, karnataks
3            14/09/2021                              Haryana, Punajb
4            14/09/2021                          Manipur, Meghalaya
5            09/08/2021              Maharashtra, Gujarat, Rajasthan
6            06/07/2021                          Gujarat, Rajasthan
```

7	27/05/2021	Maharashtra,Goa.Karnataka, Andhra Pradesh
8	29/04/2021	Andhra Pradesh, Telangna
9	16/03/2021	Madhya Pradesh, Chhatisgarh, Jharkhan, Bihar
10	24/02/2021	Maharashtra, Gujarat, Rajasthan
11	15/02/2021	Gujarat, Rajasthan
12	15/01/2021	Haryana, Delhi
13	14/12/2020	Haryana, Delhi, UttarPradesh,
14	24/11/2020	Uttrakhand, Himachal Pradesh, Jammu & Kashmir
15	12/11/2020	Maharashtra, Gujarat, Rajasthan
16	05/11/2020	Haryana, Punjab, Rajasthan
17	14/09/2020	Haryana, Rajasthan
18	10/08/2020	Gujarat, Rajasthan, Haryana, Delhi
19	10/08/2020	Kerala, karnataka, Maharashtra, Gujarat
20	14/07/2020	Maharashtra, Karnataka, Tamil nadu
21	24/04/2020	Haryana, Gujarat, Maharashtra, Delhi
22	NaN	West bengal, Andhra Pradesh
23	NaN	Maharashtra, Gujarat, Rajasthan

	Sector Type
0	1
1	1
2	3
3	1
4	1
5	1
6	3
7	2
8	2
9	2
10	1
11	3
12	4
13	1
14	1
15	5
16	1
17	1
18	1
19	1
20	6
21	1
22	1
23	1

[32]:

```
df4['State']=df2['State'].replace({'UTTAR PRADESH' : 1 ,'GUJARAT':2,␣
 ↪'TELANGANA':3, 'MAHARASHTRA':4 , 'ASSAM':5 , 'MADHYA PRADESH':6 , 'TAMIL␣
 ↪NADU':7 , 'HARYANA':8,'JAMMU AND KASHMIR':9, 'NEW DELHI':10,'MANIPUR':
 ↪11,'BIHAR':12})
df4
df4['Sector Type']=df2['Sector Type'].replace({'Agro' : 1 ,'Health/Hospital':2,␣
 ↪'Marketing':3, 'Fisheries':4 , 'Cooperative Bank':5 , 'Industrial/Textile':
 ↪6, 'Credit':7, 'Dairy':8})
df4
df4['District']=df2['District'].replace({'SANGLI':1, 'ADILABAD':2,'ADILABAD':
 ↪3,'AHMEDABAD':4, 'VALSAD':5, 'BATOD':6,'BISHNUPUR':7, 'DEORIA':8, 'BANDA':9,␣
 ↪'MUZAFFARNAGAR':10 , 'CHIRANG':11, 'NEW DELHI':12, 'BULANDSHAHR':13, 'BHIND':
 ↪14, 'BHOPAL':15, 'BAREILLY':16, 'JHANSI':17, 'SOLAPUR':18, 'CHENNAI':19,␣
 ↪'LUCKNOW':20, 'GURGAON':21, 'PATNA':22 })
df4
```

[32]:

	Sr. No.	Name of Society	\
0	73	Farm to Foreign Exports Entrepreneur Developme…	
1	74	Dharitri Jute and Eco Friendly Products Multi …	
2	75	The Malabar Multi State Agro Cooperative Socie…	
3	76	Sai Raam Multi State Agri Cooperative Society Ltd	
4	77	Prabhavana Multi State Womens' Jute and Allied…	
5	78	The Gayatri Cooperative Urban Bank Ltd	
6	79	Shri Saibaba Multi State Cooperative Dudh Utpa…	
7	80	Subhiksha Organic Farmers Multi State Cooperat…	
8	81	The Panipat Urban Cooperative Bank Ltd (PUCB)	
9	82	Green Earth Agro Cooperative Society Ltd	
10	83	Bihar Cooperative Federation Ltd	
11	84	Bharathi Cooperative Housing Society Ltd	
12	85	The Surat People's Cooperative Bank Ltd	
13	86	Sanmati Sahakari Bank Ltd	
14	87	Reserve Bank Staff & Officers cooperative Cred…	
15	88	Dev Lok Housing Cooperative Ltd	
16	89	Maa Vaishno Agro Multi State Cooperative Socie…	
17	90	Sree Balki Multi State Cooperative Housing Soc…	
18	91	Southern Multi State Agro Cooperative Society …	
19	92	Dombivli Nagari Sahakari Bank Ltd	
20	93	Lucknow Division Insurance Employees Cooperati…	
21	94	The Kalyan Janata Sahakari Bank Ltd	
22	95	The Eastern Railway Employees\nCooperative Ban…	
23	96	Purvanchal Krishi Avam Kisan kalyan Multi Stat…	
24	97	The Ajara Urban Cooperative Bank Limited	
25	98	Bihar State Housing Cooperative Federation Ltd	
26	99	The Navodit Cooperative Urban Thrift & Credit …	
27	100	Lokseva Multi State Credit Cooperative Society…	

	Address	State	District	\

0	H.No. 6-9, Gundugolanu(P.O), Bhimadole (Mandal…	1.0	1.0
1	D.No.12-2-42111, Alapati Nagar, Gudimalkapur, …	11.0	2.0
2	1st Floor, Aiswarya Complex, Thavakkara, Kannu…	12.0	3.0
3	No.6. (D-18), 6th Cross West Extension, Thilla…	1.0	4.0
4	, House. No 2-2-647/G/18/1, Near Shivam Road, …	4.0	5.0
5	7-1,72, 73,74 Gayatri Towers, Tehsil Chowrasta…	10.0	6.0
6	A/p Lohgaon, Tal- Rahata, Dist. Ahmednagar, PI…	4.0	7.0
7	Krushi Nivasa, At Kuruvalli,Thirthahalli, Shim…	4.0	7.0
8	510/8, G T Road,Panipat, Haryana 132103	4.0	8.0
9	3/170, Vishwas Khand, Gomti Society , Lucknow …	4.0	9.0
10	Budh Marg, Patna 800001, Bihar	4.0	10.0
11	87-A, Perambur High Road, Perambur Post, Chenn…	4.0	7.0
12	Vasudhara Bhavan, Timalyawad, Nanpura, Surat, …	9.0	12.0
13	9/114, Laxmi Market, Ichalkaranji Kolhapur, Ma…	1.0	13.0
14	C/ Reserve Bank of India, Amar Building, 2nd F…	4.0	14.0
15	A-55, Dwarka, Sector-19, New Delhi	4.0	9.0
16	Shop No. 25, Block-A, Old LIC Building, Near B…	4.0	15.0
17	No. 9, Elango Street, Kabilar, Nagar,\nManaval…	4.0	16.0
18	27 Gokulam, Mount Pleasant Road, Conoor , Ooty…	1.0	13.0
19	Plot No P-52, MIDC Phase-II, Kalyan Shil Road,…	7.0	11.0
20	Jeevan Prakash Building, 30, Hazratganj , Luck…	4.0	17.0
21	Kalyanam astu, Om Vijaykrishna, Apt, Adharwadi…	12.0	3.0
22	17, Neta Ji Subhash Road, Kolkata 700001	4.0	16.0
23	House No. 799, Vill: Taraya Sujan, Post : Tary…	10.0	6.0
24	393-B, Main Road Ajara, Kolhapur, 416505, Maha…	6.0	18.0
25	Lalit Bhawan, Bailey Road, Patna , Bihar 800014	8.0	19.0
26	147, South Amarkali, Delhi 110051	NaN	NaN
27	A/p Tisgaon Pravara, Rahata, Ahmednagar 413712…	NaN	NaN

	Date of Registration	Area of Operation \
0	16/01/2020	Haryana, UttarPradesh, Himachal, Punjab
1	12/12/2019	Bihar, West bengal, Odisha
2	20/10/2019	Haryana, Punajb
3	06/12/2019	Manipur, Meghalaya
4	27/09/2019	Maharashtra, Gujarat, Rajasthan
5	23/08/2019	Gujarat, Rajasthan
6	06/08/2019	Maharashtra,Goa.Karnataka, Andhra Pradesh
7	03/07/2019	Andhra Pradesh, Telangna
8	14/05/2019	Madhya Pradesh, Chhatisgarh, Jharkhan, Bihar
9	03/05/2019	Maharashtra, Gujarat, Rajasthan
10	03/10/2018	Gujarat, Rajasthan
11	18/05/2018	Haryana, Delhi
12	28/03/2018	Haryana, Delhi, UttarPradesh,
13	27/03/2018	Uttrakhand, Himachal Pradesh, Jammu & Kashmir
14	13/02/2018	Maharashtra, Gujarat, Rajasthan
15	02/02/2018	Haryana, Punjab, Rajasthan
16	19/01/2018	Haryana, Rajasthan

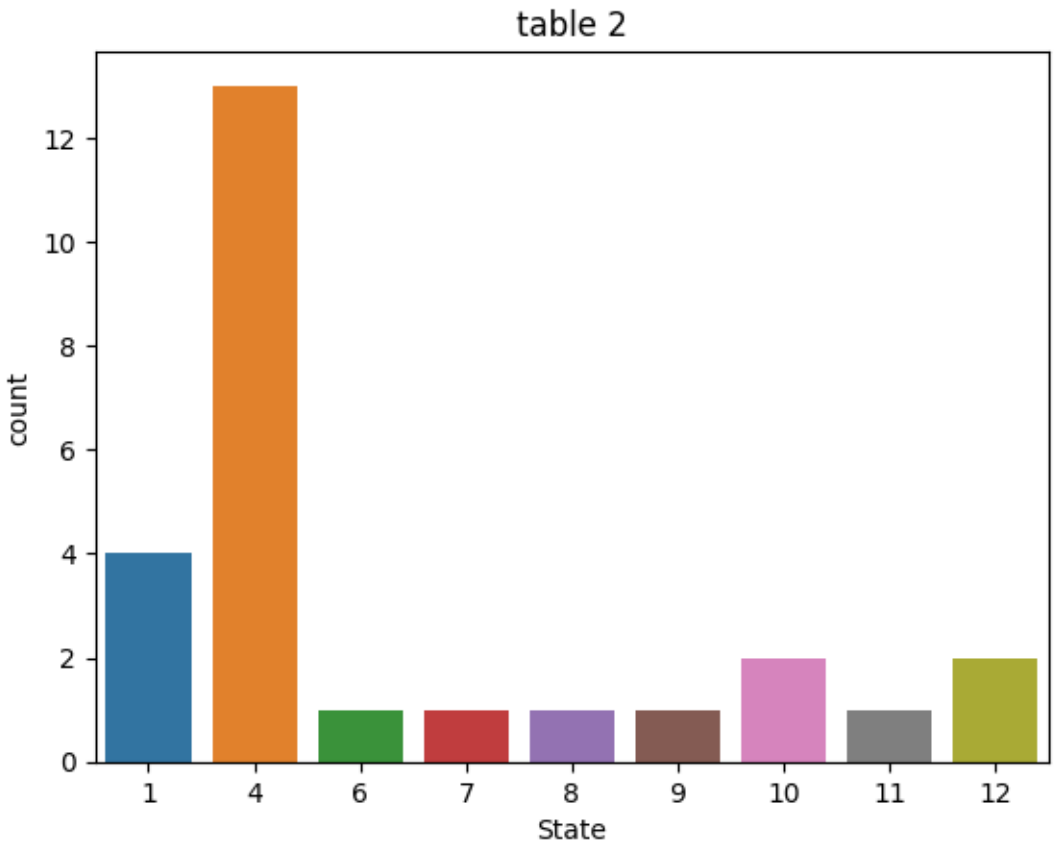
```
17          05/01/2018            Gujarat, Rajasthan, Haryana, Delhi
18          19/12/2017       Kerala, karnataka, Maharashtra, Gujarat
19          28/07/2017            Maharashtra, Karnataka, Tamil nadu
20          11/07/2017          Haryana, Gujarat, Maharashtra, Delhi
21          24/03/2017                    West bengal, Andhra Pradesh
22          22/03/2017               Maharashtra, Gujarat, Rajasthan
23          06/03/2017            Gujarat, Rajasthan, Haryana, Delhi
24          17/02/2017       Kerala, karnataka, Maharashtra, Gujarat
25          19/10/2016            Maharashtra, Karnataka, Tamil nadu
26          27/09/2016          Haryana, Gujarat, Maharashtra, Delhi
27          26/09/2016                    West bengal, Andhra Pradesh
```

```
    Sector Type
0           1.0
1           1.0
2           3.0
3           1.0
4           1.0
5           1.0
6           3.0
7           2.0
8           2.0
9           2.0
10          1.0
11          3.0
12          4.0
13          1.0
14          1.0
15          5.0
16          1.0
17          1.0
18          1.0
19          1.0
20          6.0
21          1.0
22          1.0
23          1.0
24          1.0
25          1.0
26          NaN
27          NaN
```

dATA VISUALIZATION

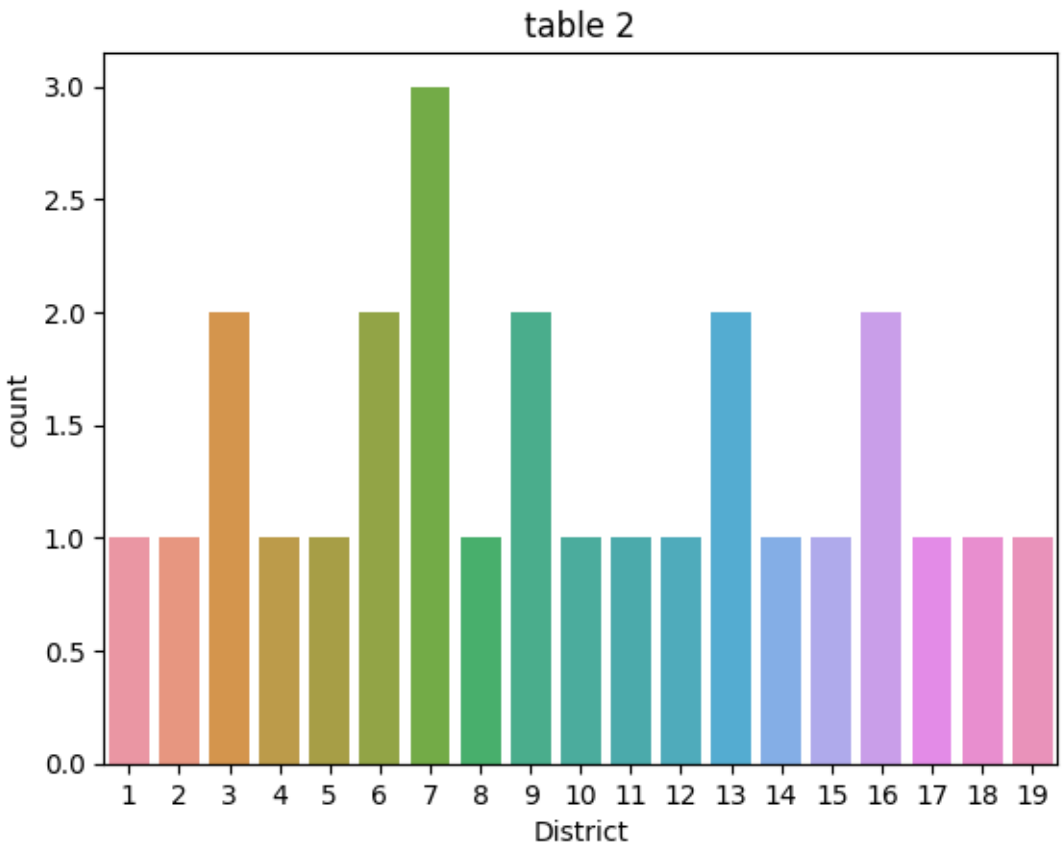
[33]:
```
sns.countplot(x='State', data= df2)
plt.title('table 2')
```

[33]: Text(0.5, 1.0, 'table 2')



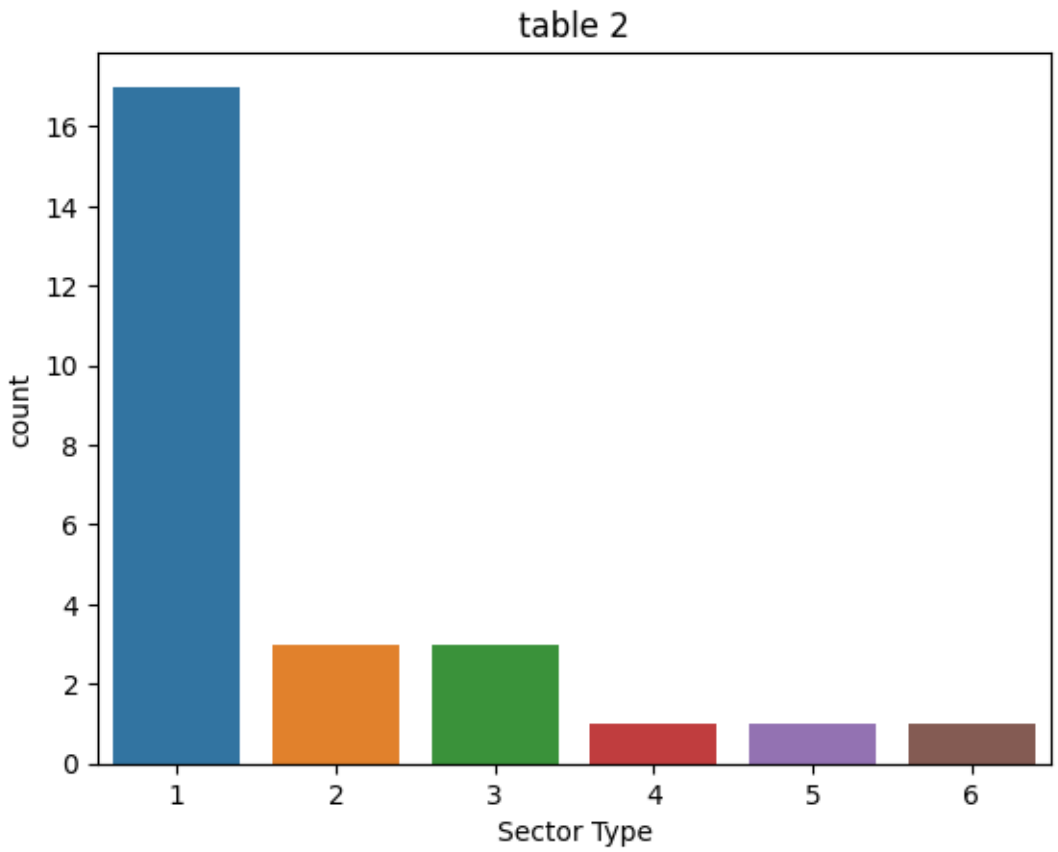
[34]:
```
sns.countplot(x='District', data= df2)
plt.title('table 2')
```

[34]: Text(0.5, 1.0, 'table 2')



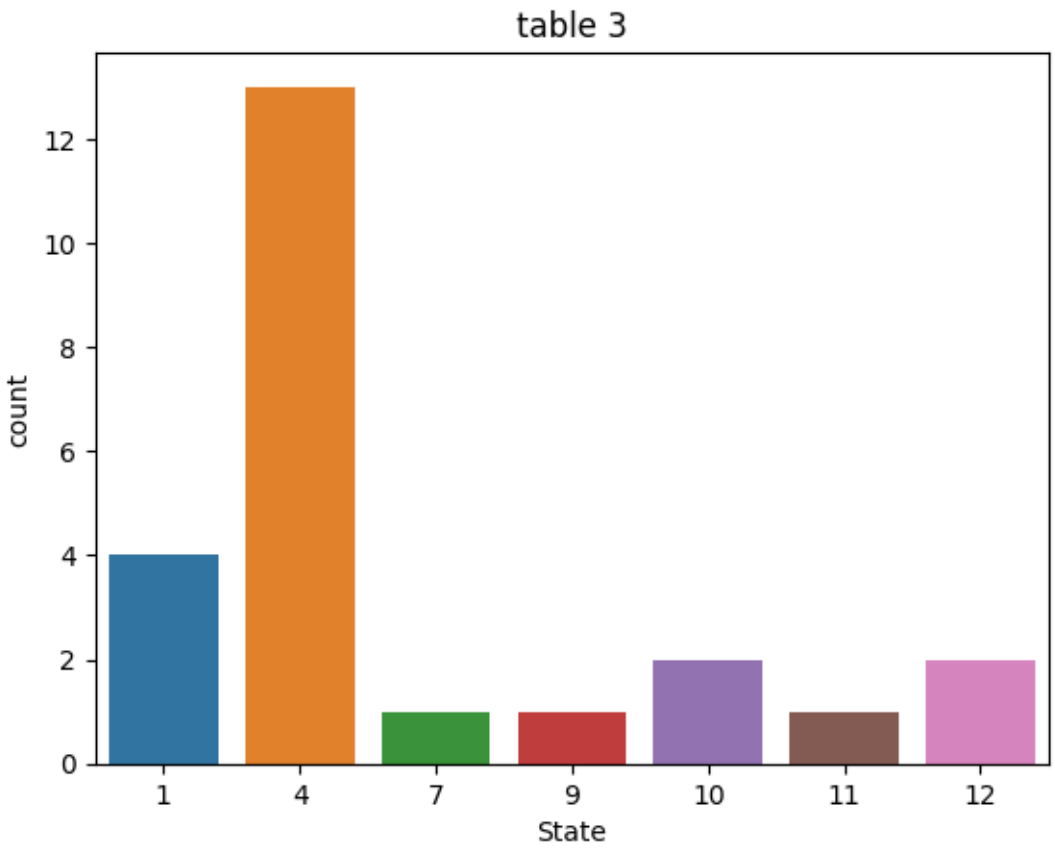
[35]:
```
sns.countplot(x='Sector Type', data= df2)
plt.title('table 2')
```

[35]: Text(0.5, 1.0, 'table 2')



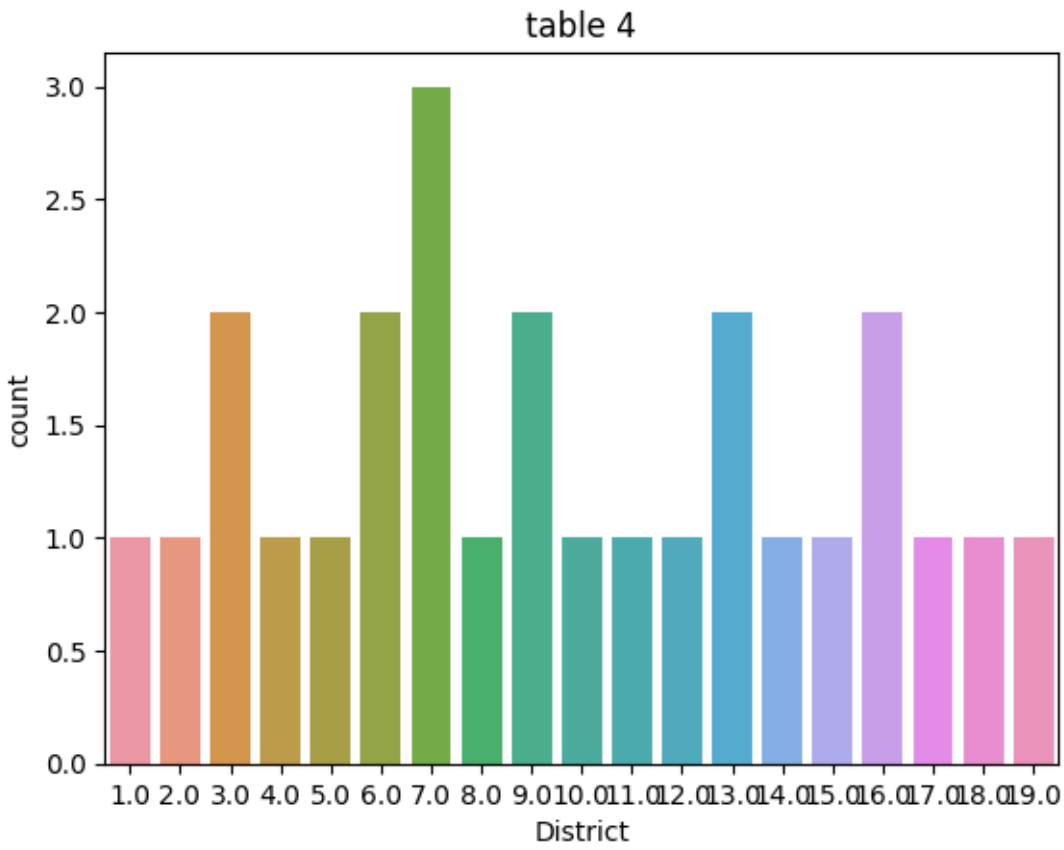
[36]:
```
sns.countplot(x='State', data= df3)
plt.title('table 3')
```

[36]: Text(0.5, 1.0, 'table 3')



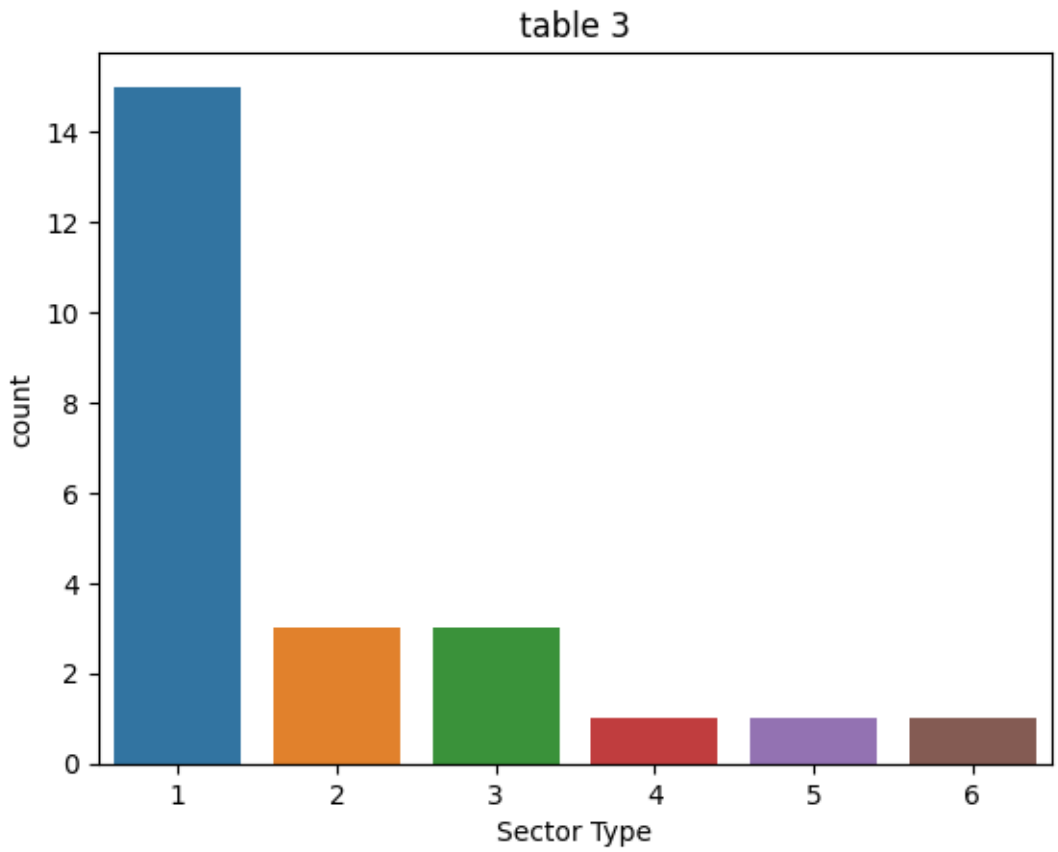
[37]:
```
sns.countplot(x='District', data= df4)
plt.title('table 4')
```

[37]: Text(0.5, 1.0, 'table 4')



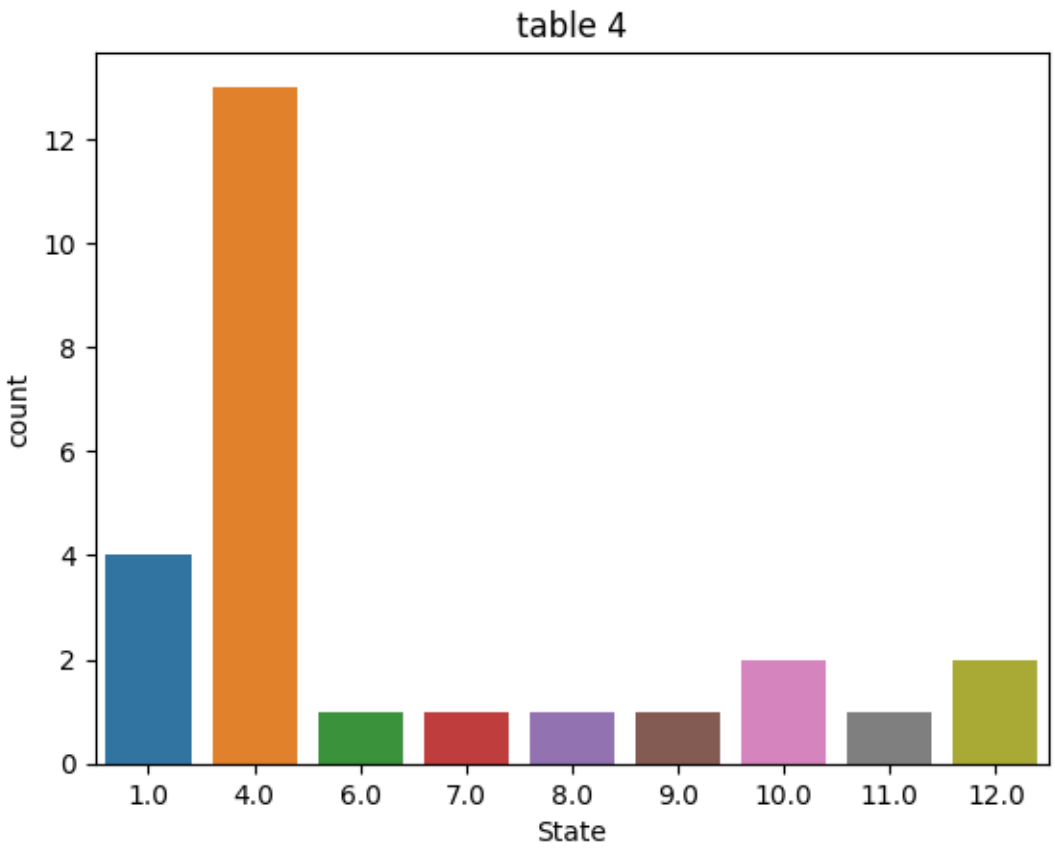
[38]:
```
sns.countplot(x='Sector Type', data= df3)
plt.title('table 3')
```

[38]: Text(0.5, 1.0, 'table 3')



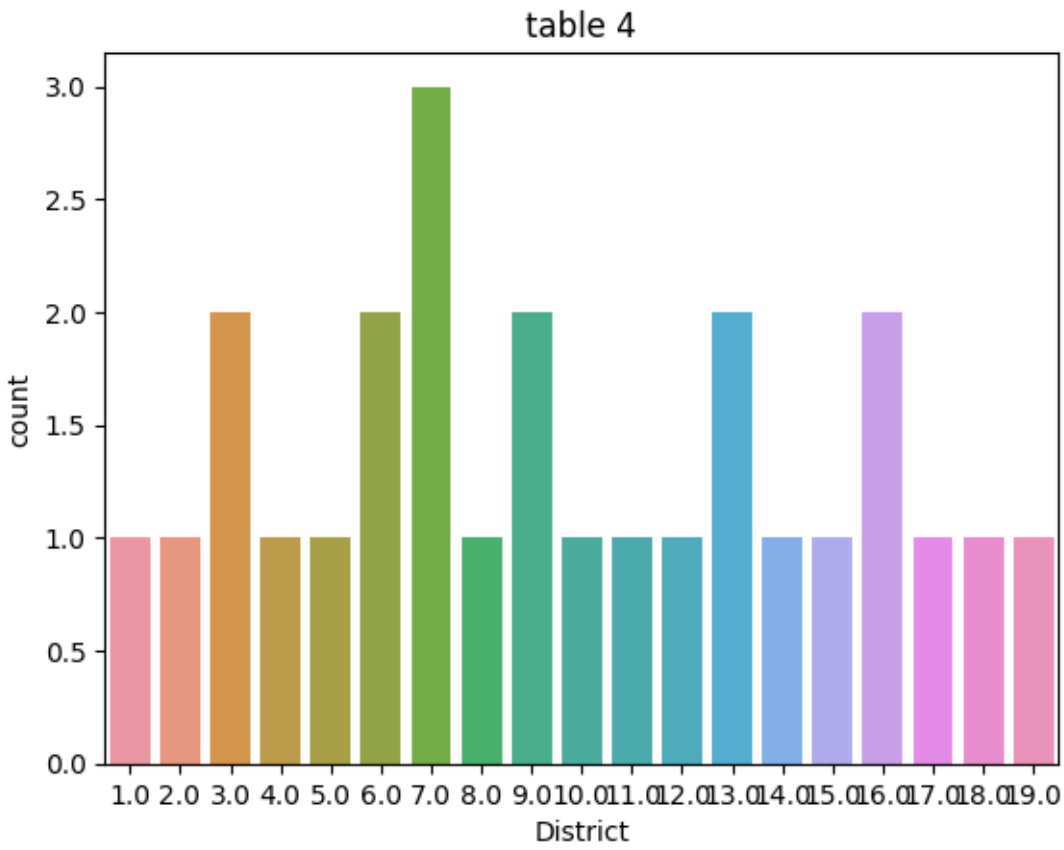
[39]:
```
sns.countplot(x='State', data= df4)
plt.title('table 4')
```

[39]: Text(0.5, 1.0, 'table 4')



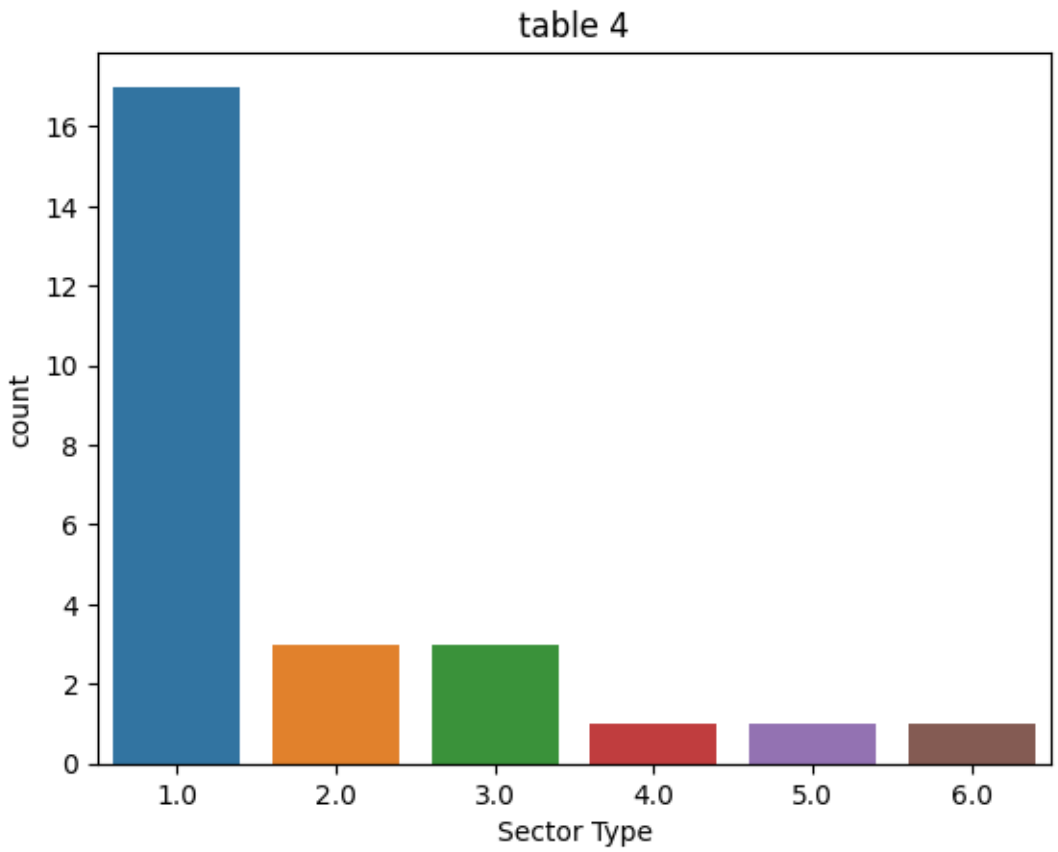
[40]:
```
sns.countplot(x='District', data= df4)
plt.title('table 4')
```

[40]: Text(0.5, 1.0, 'table 4')



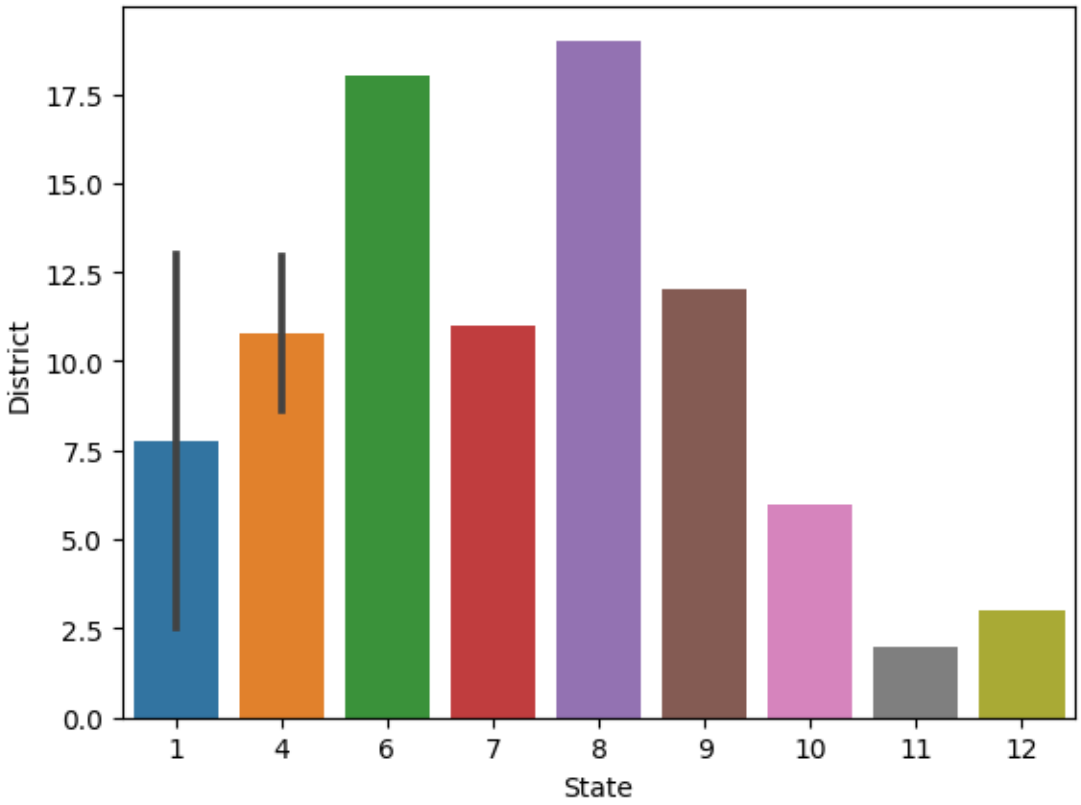
[41]:
```
sns.countplot(x='Sector Type', data= df4)
plt.title('table 4')
```

[41]: Text(0.5, 1.0, 'table 4')



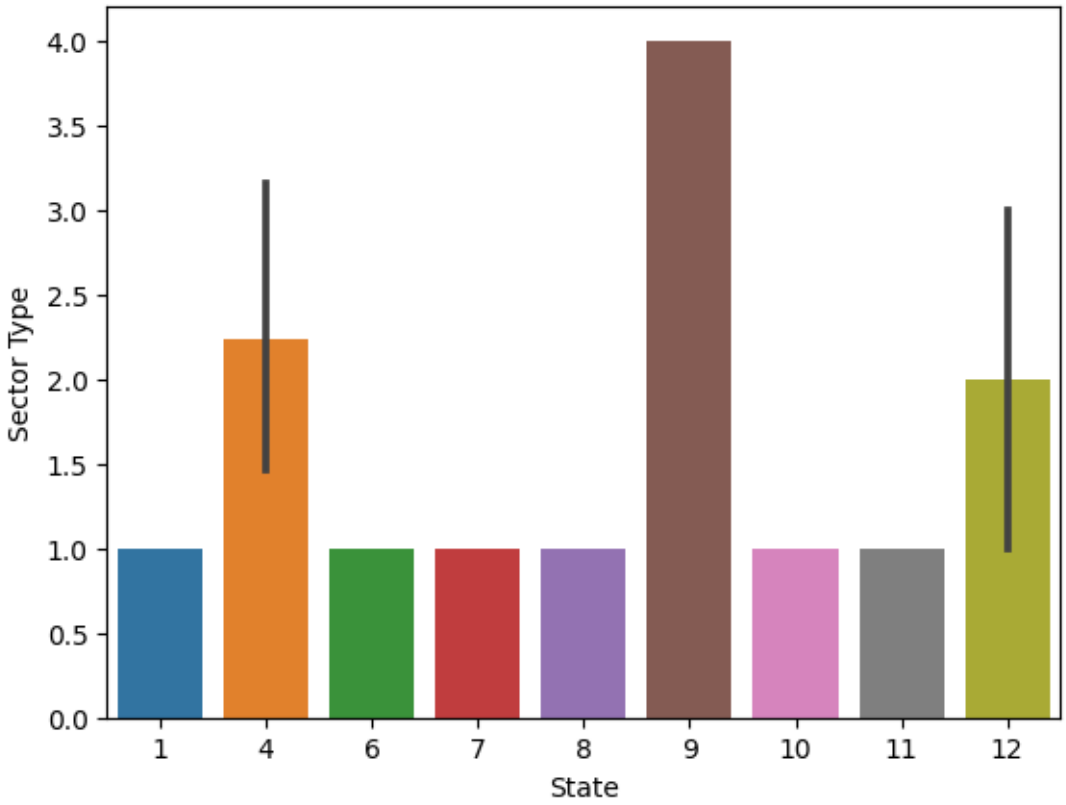
[42]:
```
sns.barplot(x='State', y='District', data=df2)
```

[42]:
```
<Axes: xlabel='State', ylabel='District'>
```



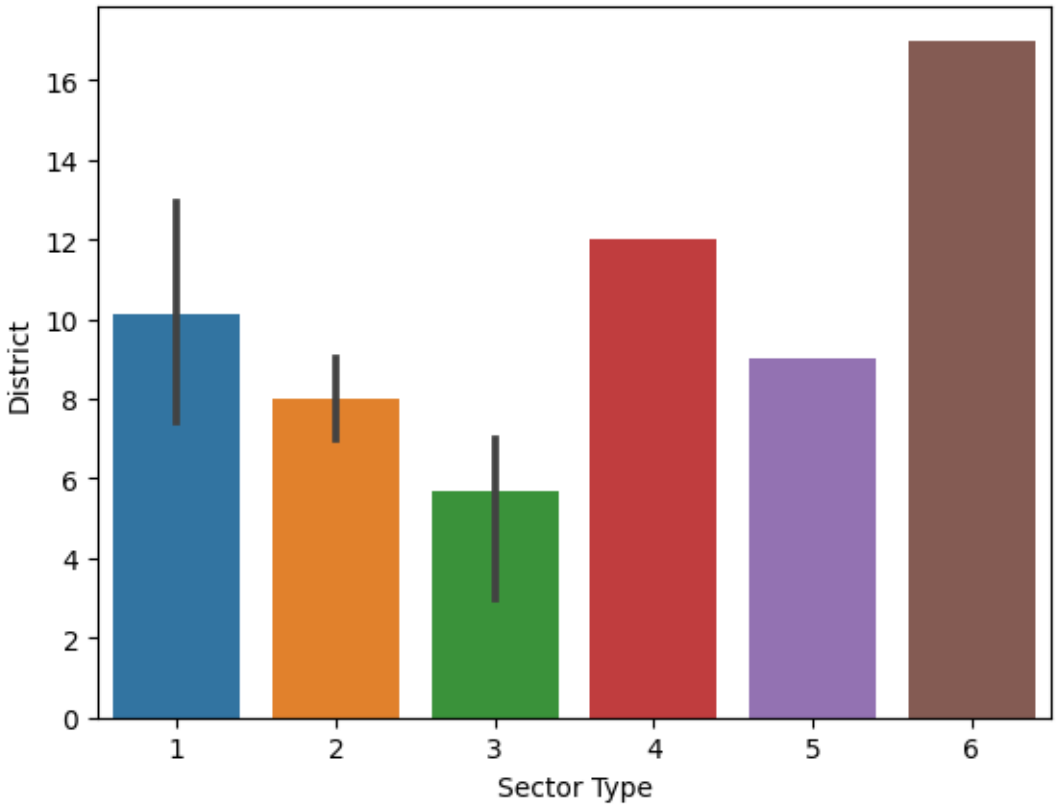
[43]:
```
sns.barplot(x='State', y='Sector Type', data=df2)
```

[43]:
```
<Axes: xlabel='State', ylabel='Sector Type'>
```



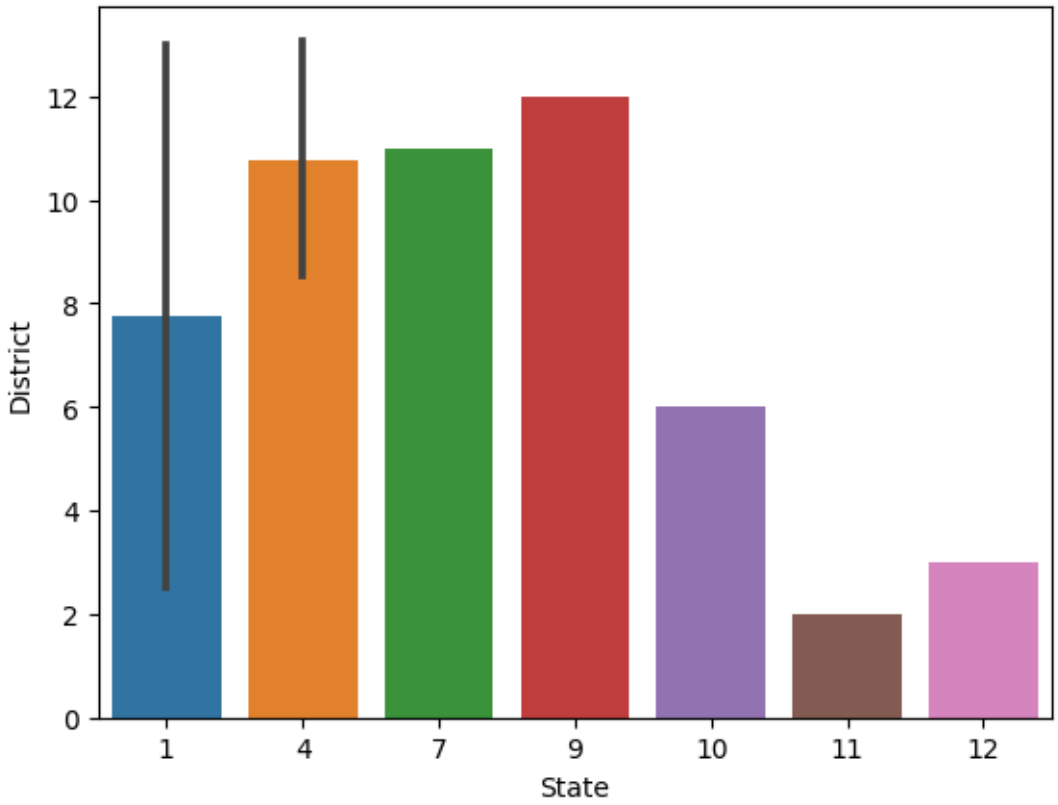
[44]:
```
sns.barplot(x='Sector Type', y='District', data=df2)
```

[44]:
```
<Axes: xlabel='Sector Type', ylabel='District'>
```



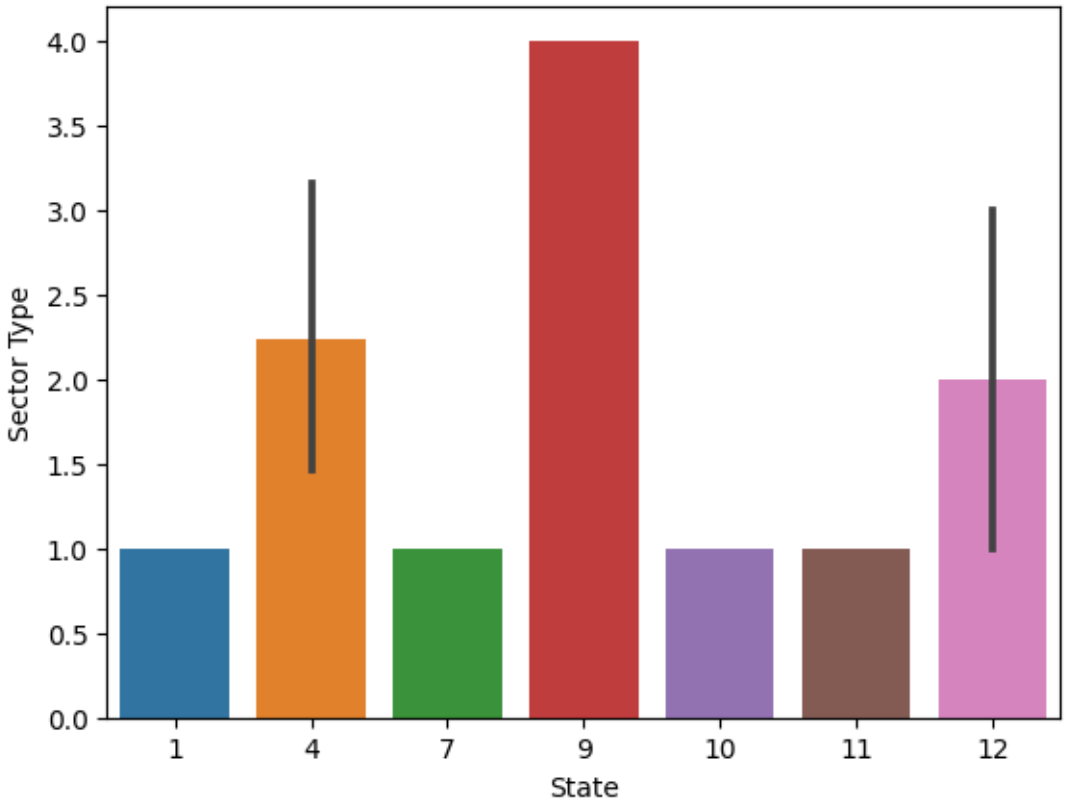
[45]:
```
sns.barplot(x='State', y='District', data=df3)
```

[45]: <Axes: xlabel='State', ylabel='District'>



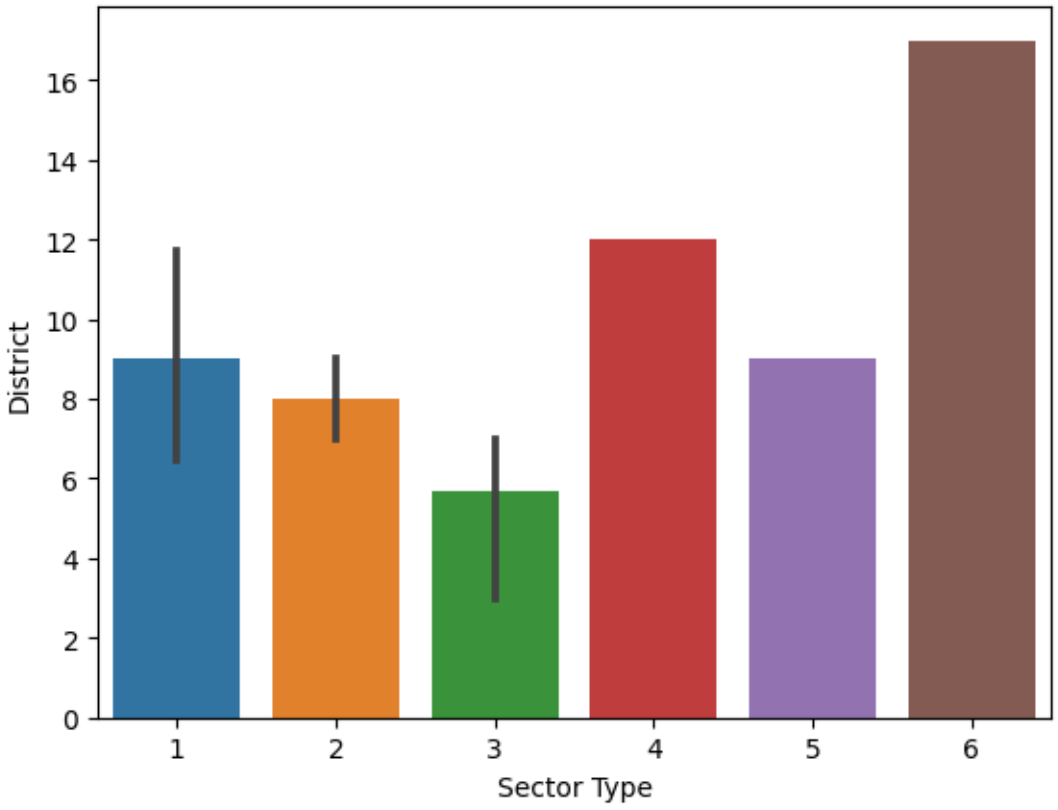
[46]:
```
sns.barplot(x='State', y='Sector Type', data=df3)
```

[46]: <Axes: xlabel='State', ylabel='Sector Type'>



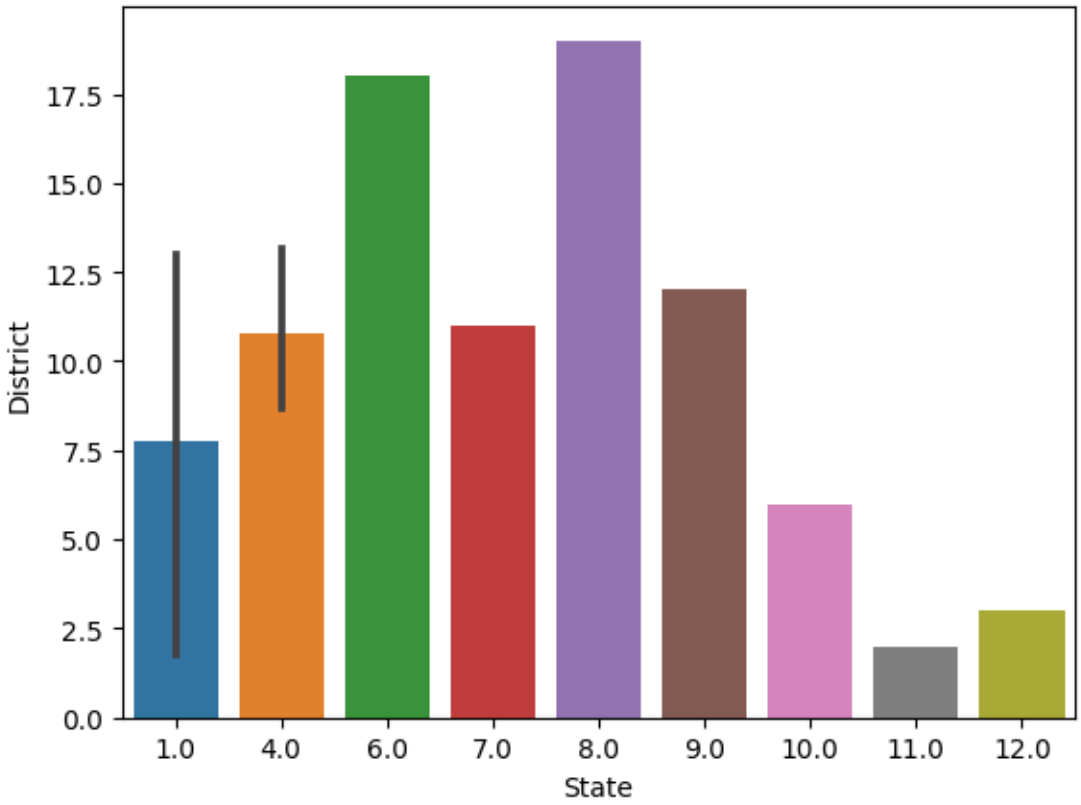
[47]:
```
sns.barplot(x='Sector Type', y='District', data=df3)
```

[47]:
```
<Axes: xlabel='Sector Type', ylabel='District'>
```



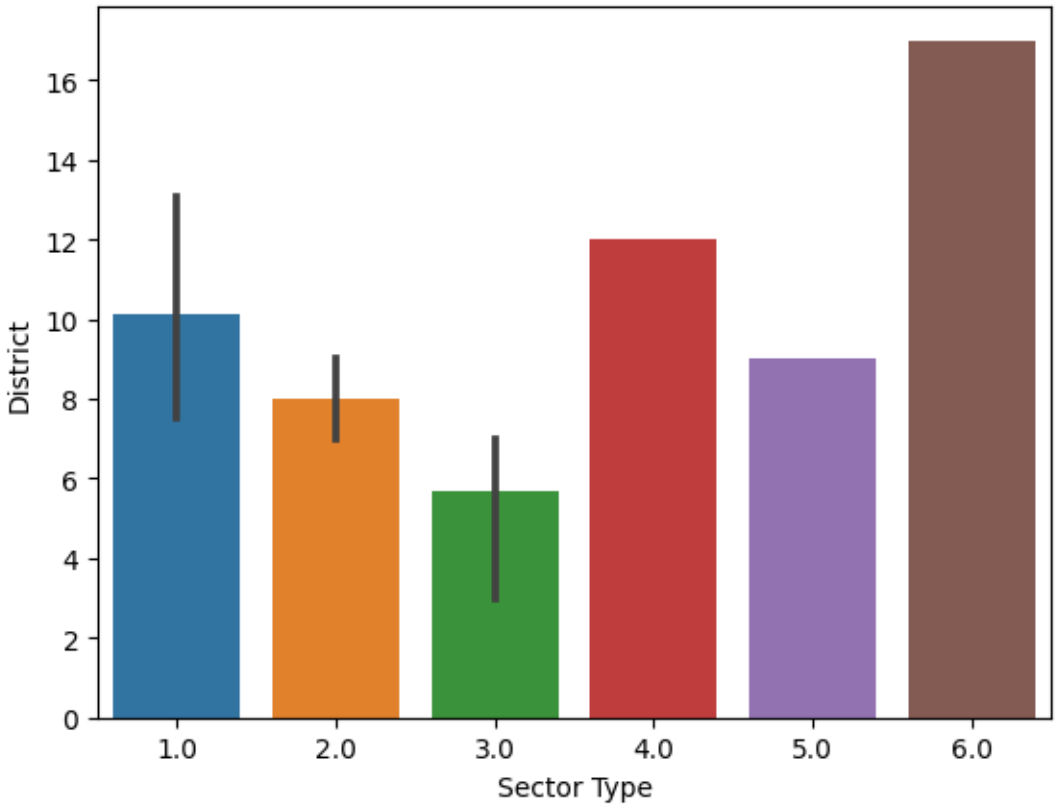
[48]:
```
sns.barplot(x='State', y='District', data=df4)
```

[48]: <Axes: xlabel='State', ylabel='District'>



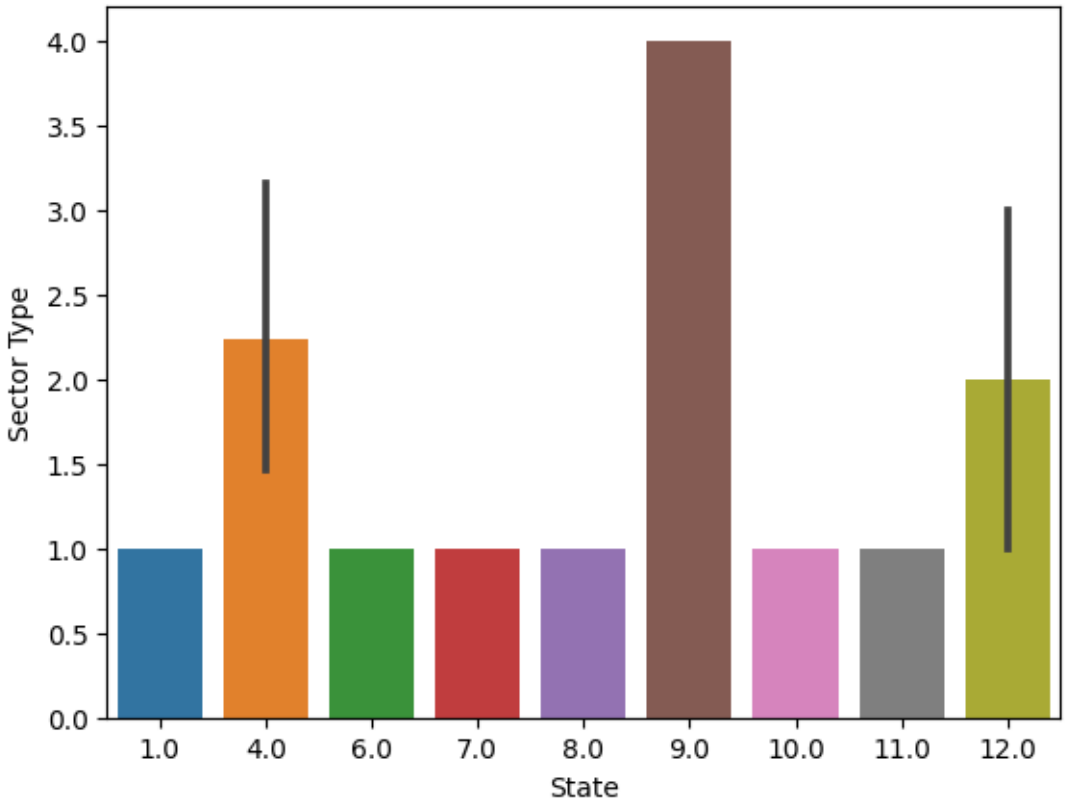
[49]:
```
sns.barplot(x='Sector Type', y='District', data=df4)
```

[49]:
```
<Axes: xlabel='Sector Type', ylabel='District'>
```



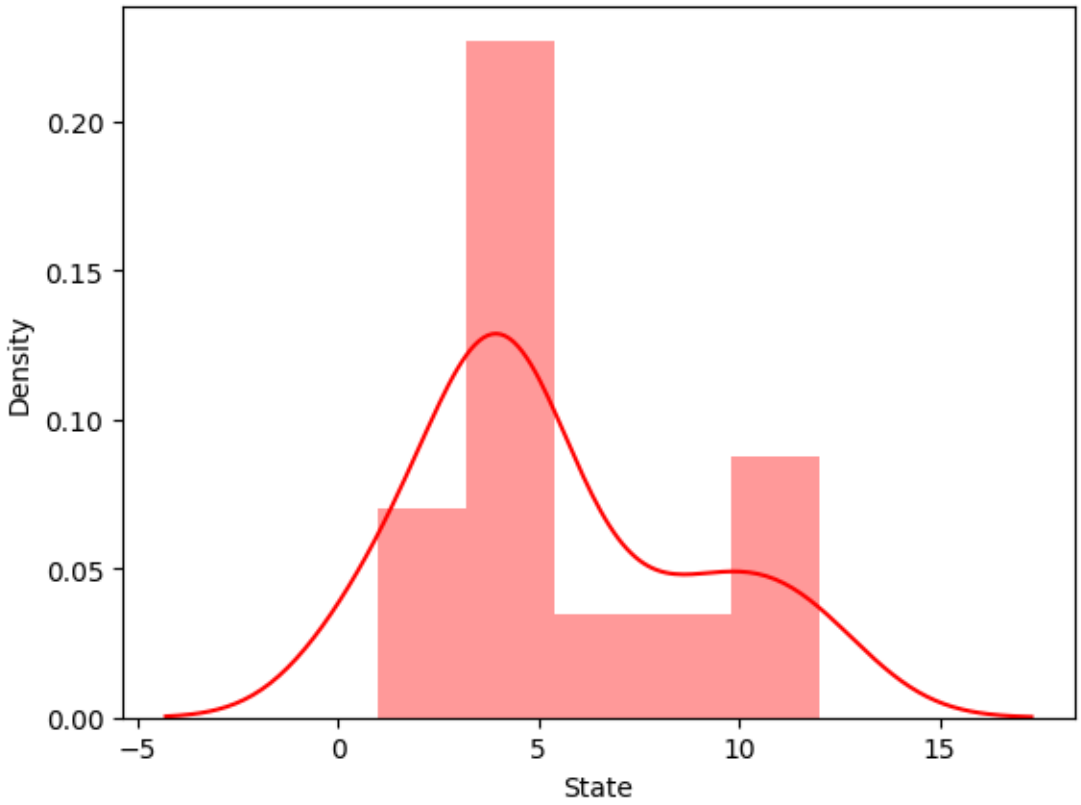
[50]:
```
sns.barplot(x='State', y='Sector Type', data=df4)
```

[50]:
```
<Axes: xlabel='State', ylabel='Sector Type'>
```



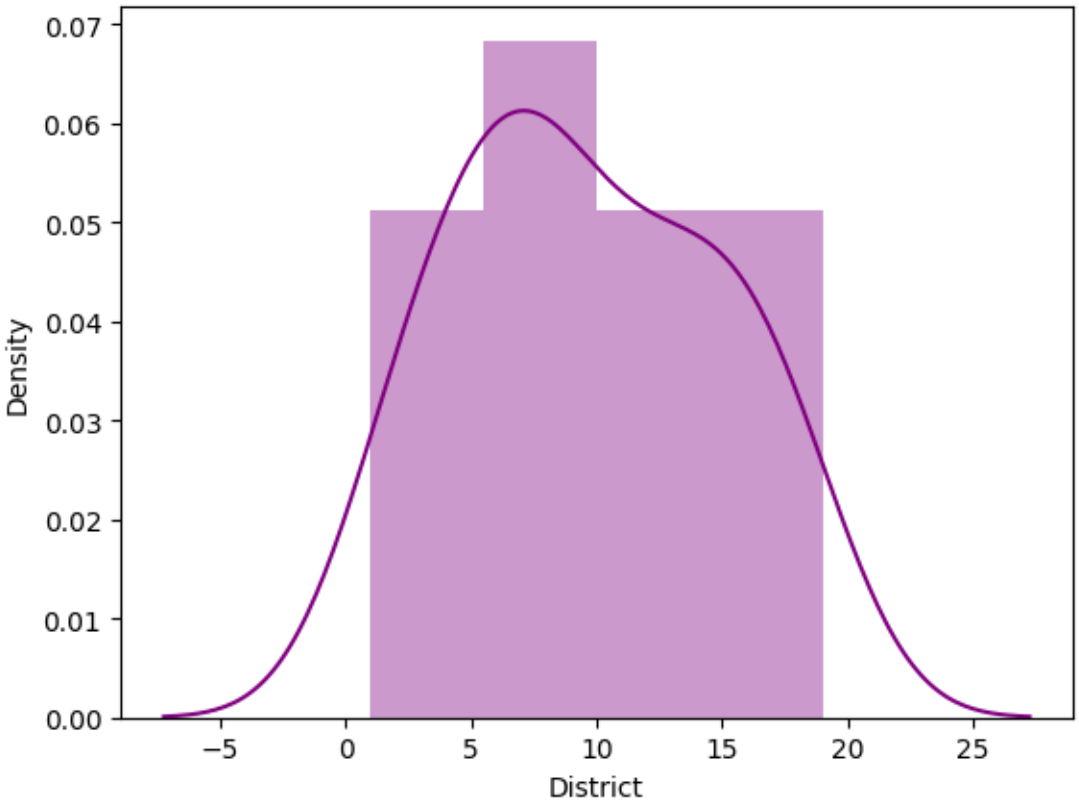
[55]:
```
sns.distplot(df2['State'] , color= 'red');
```

```
<ipython-input-55-faa3ff954315>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df2['State'] , color= 'red');
```



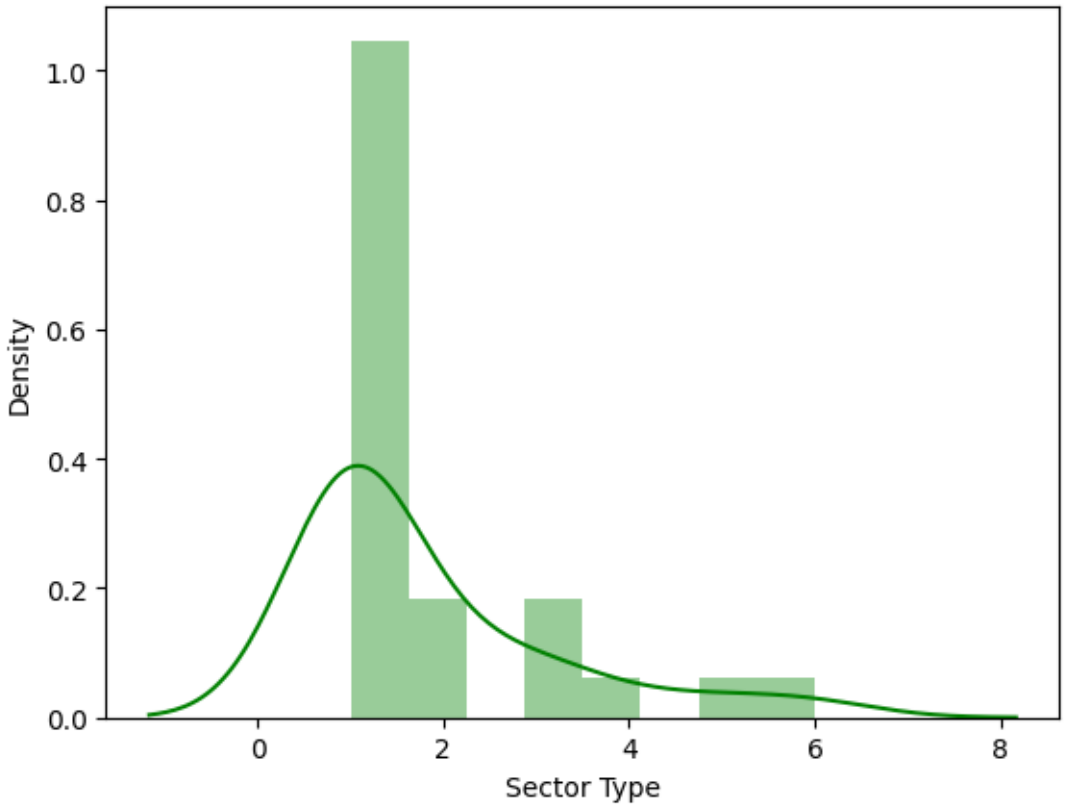
[56]:
```
sns.distplot(df2['District'], color = 'purple') ;
```

```
<ipython-input-56-7420f788a856>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df2['District'], color = 'purple') ;
```



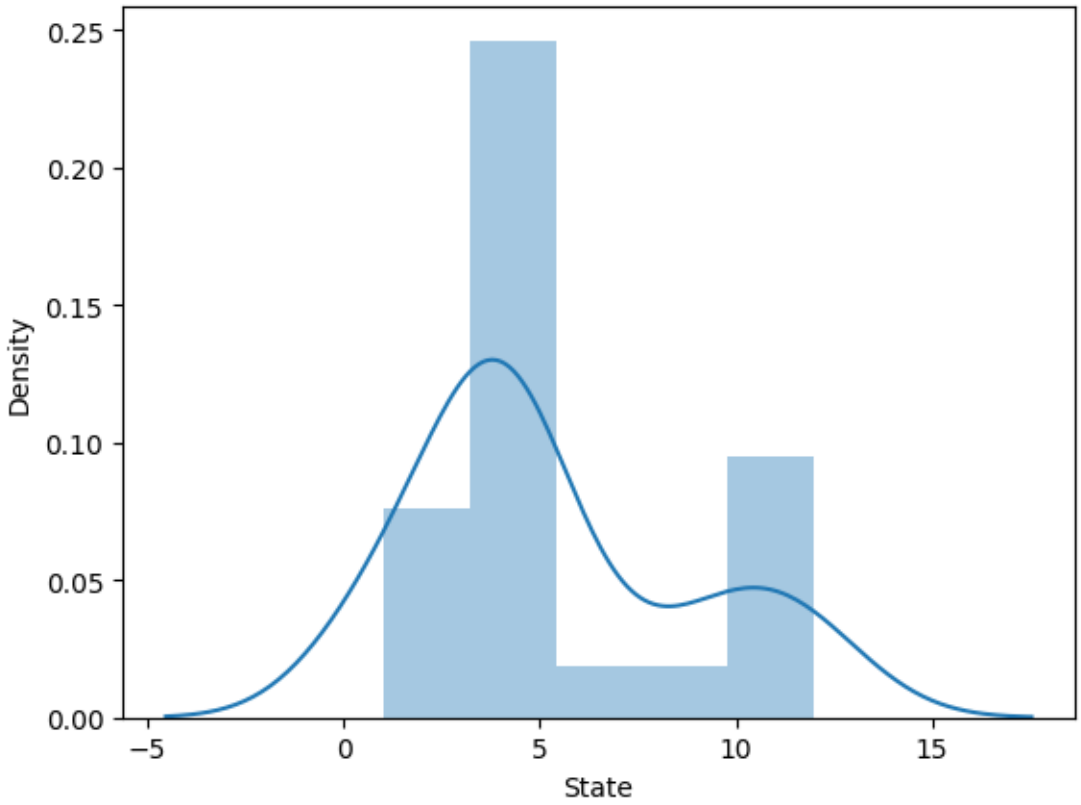
[57]:
```
sns.distplot(df2['Sector Type'] , color ='green');
```

```
<ipython-input-57-8cbca7c8588b>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df2['Sector Type'] , color ='green');
```



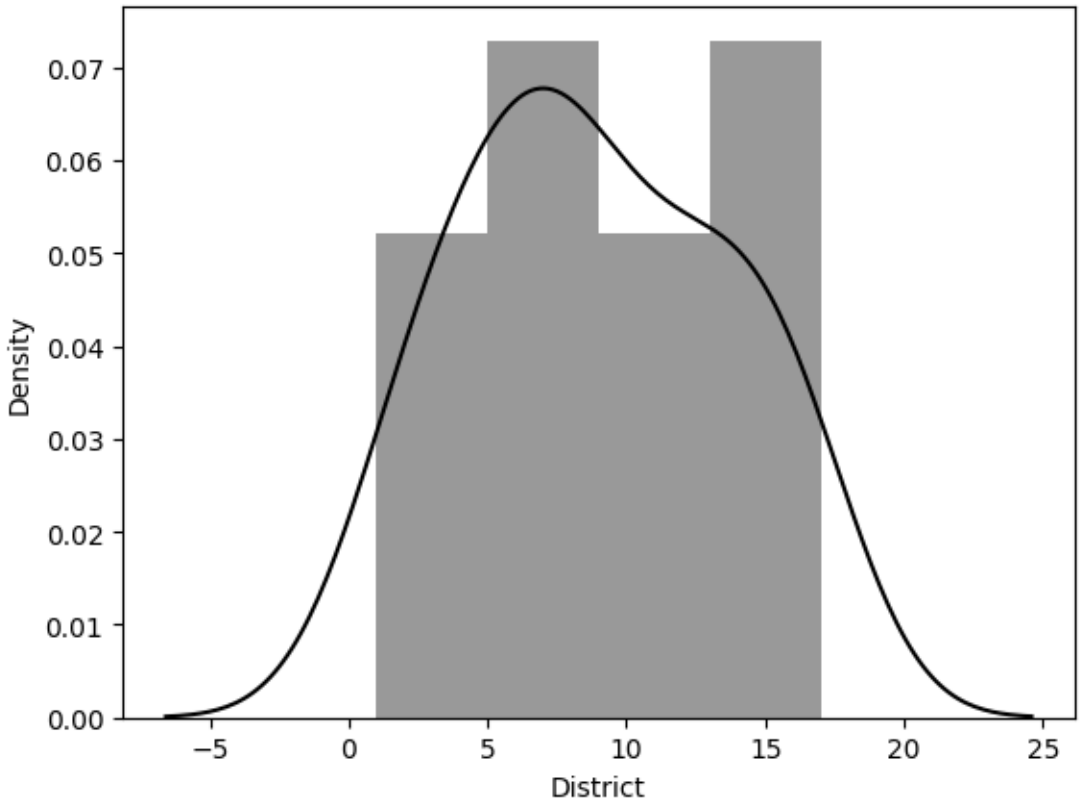
[ ]:
```
sns.distplot(df3['State']);
```

```
<ipython-input-51-0fe6015bb31a>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df3['State']);
```



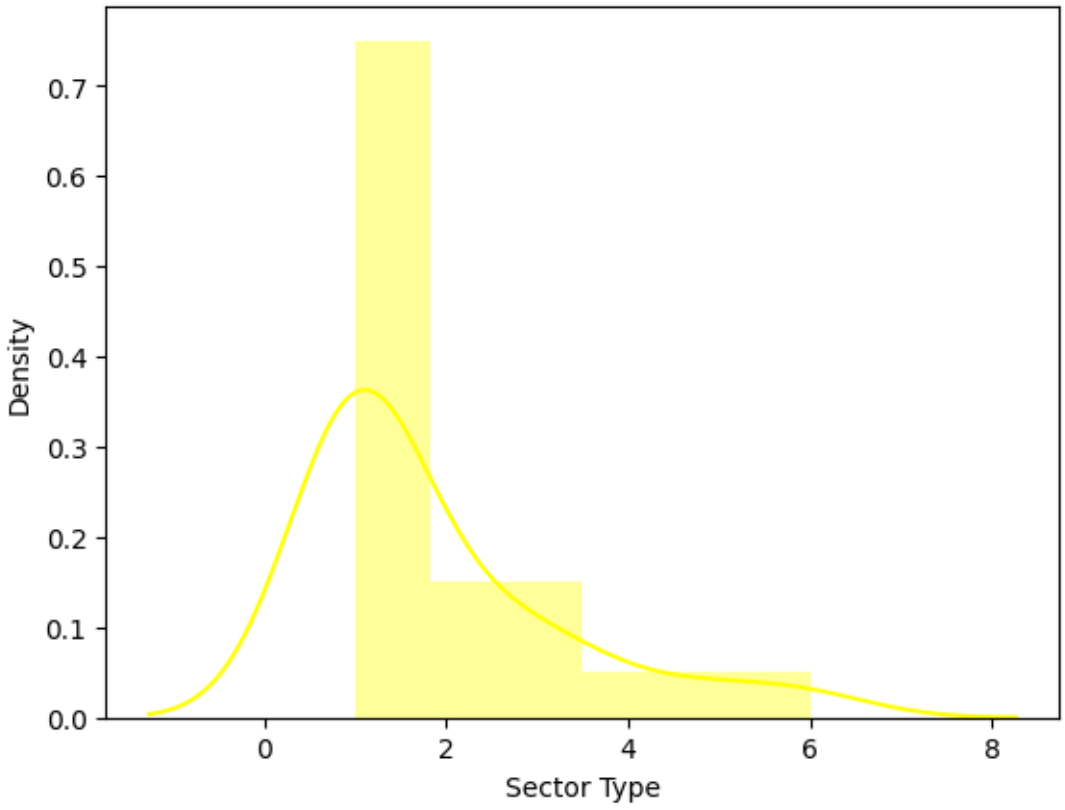
[58]:
```
sns.distplot(df3['District'], color = 'black');
```

```
<ipython-input-58-7b366ff2aaff>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df3['District'], color = 'black');
```



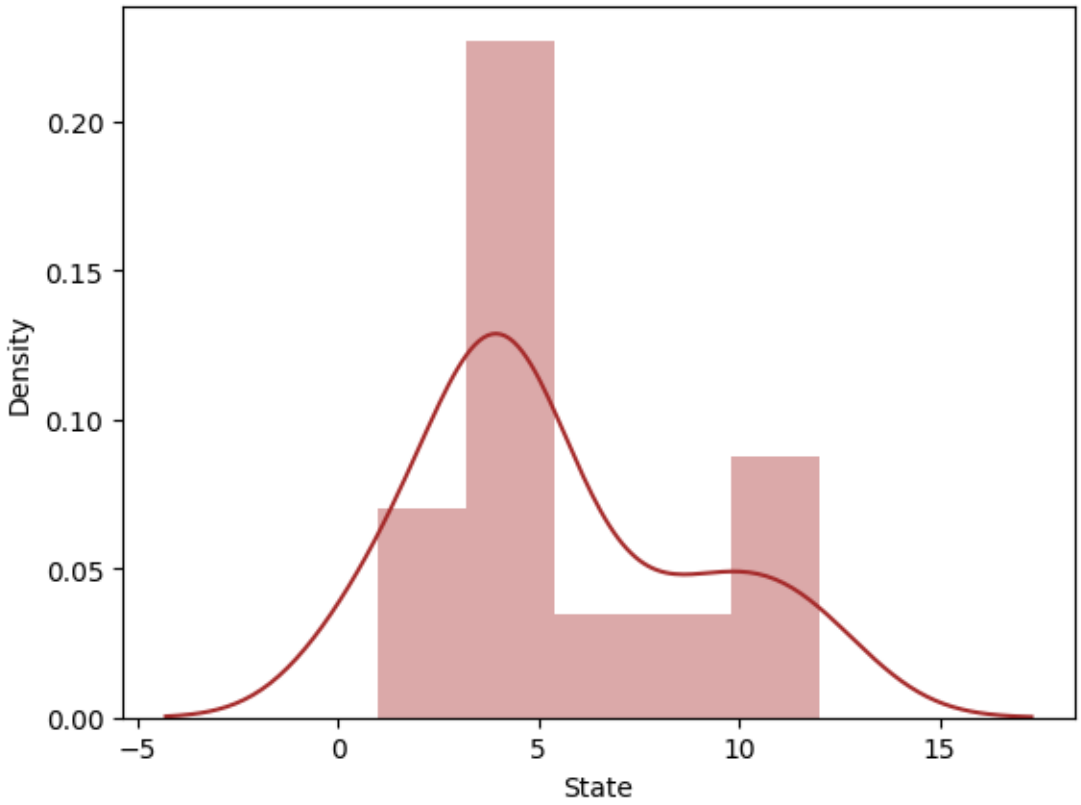
[59]:
```
sns.distplot(df3['Sector Type'],color='yellow');
```

```
<ipython-input-59-5fe4104001b6>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df3['Sector Type'],color='yellow');
```



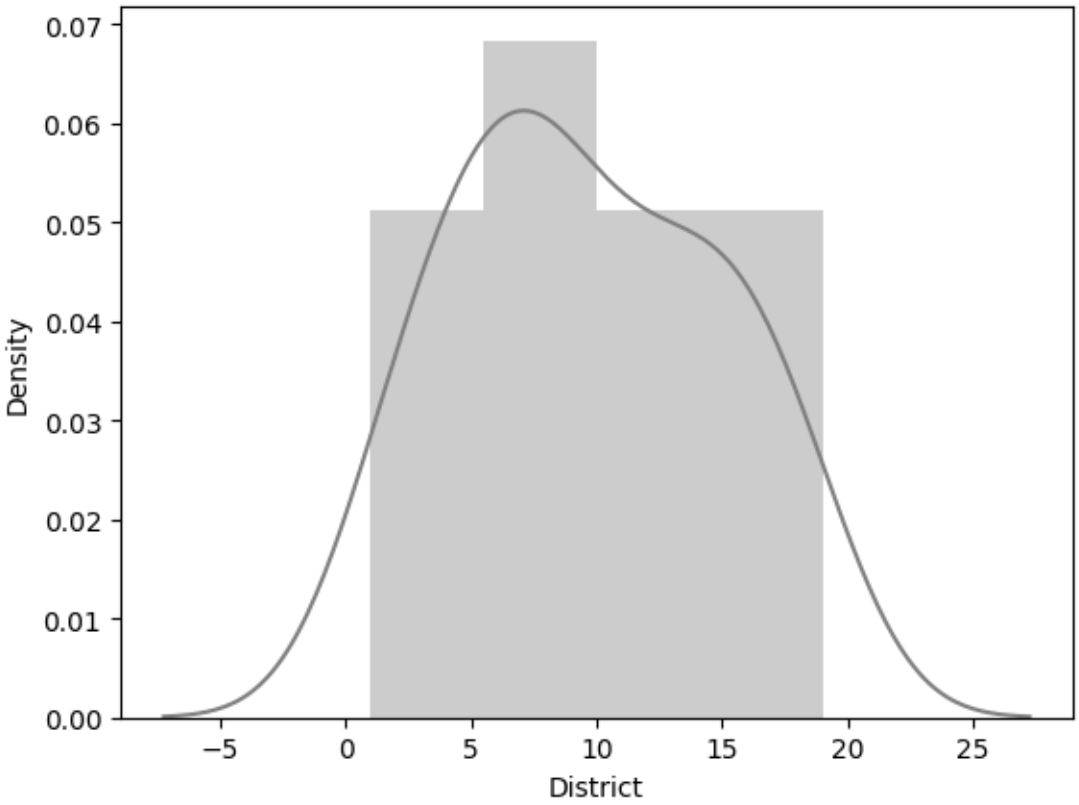
[60]:
```
sns.distplot(df4['State'],color='brown');
```

```
<ipython-input-60-64a5e62a20de>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df4['State'],color='brown');
```



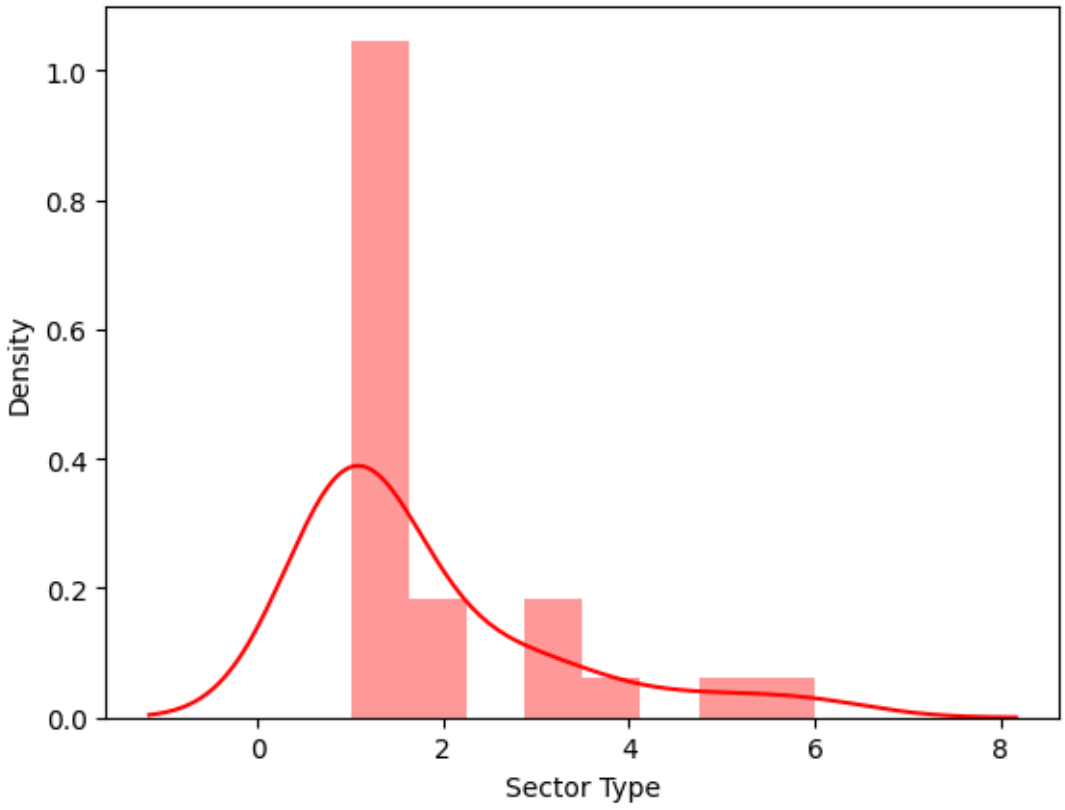
[61]:
```
sns.distplot(df4['District'], color='grey');
```

```
<ipython-input-61-704292e2cc5e>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df4['District'], color='grey');
```



[62]:
```
sns.distplot(df4['Sector Type'],color='red');
```

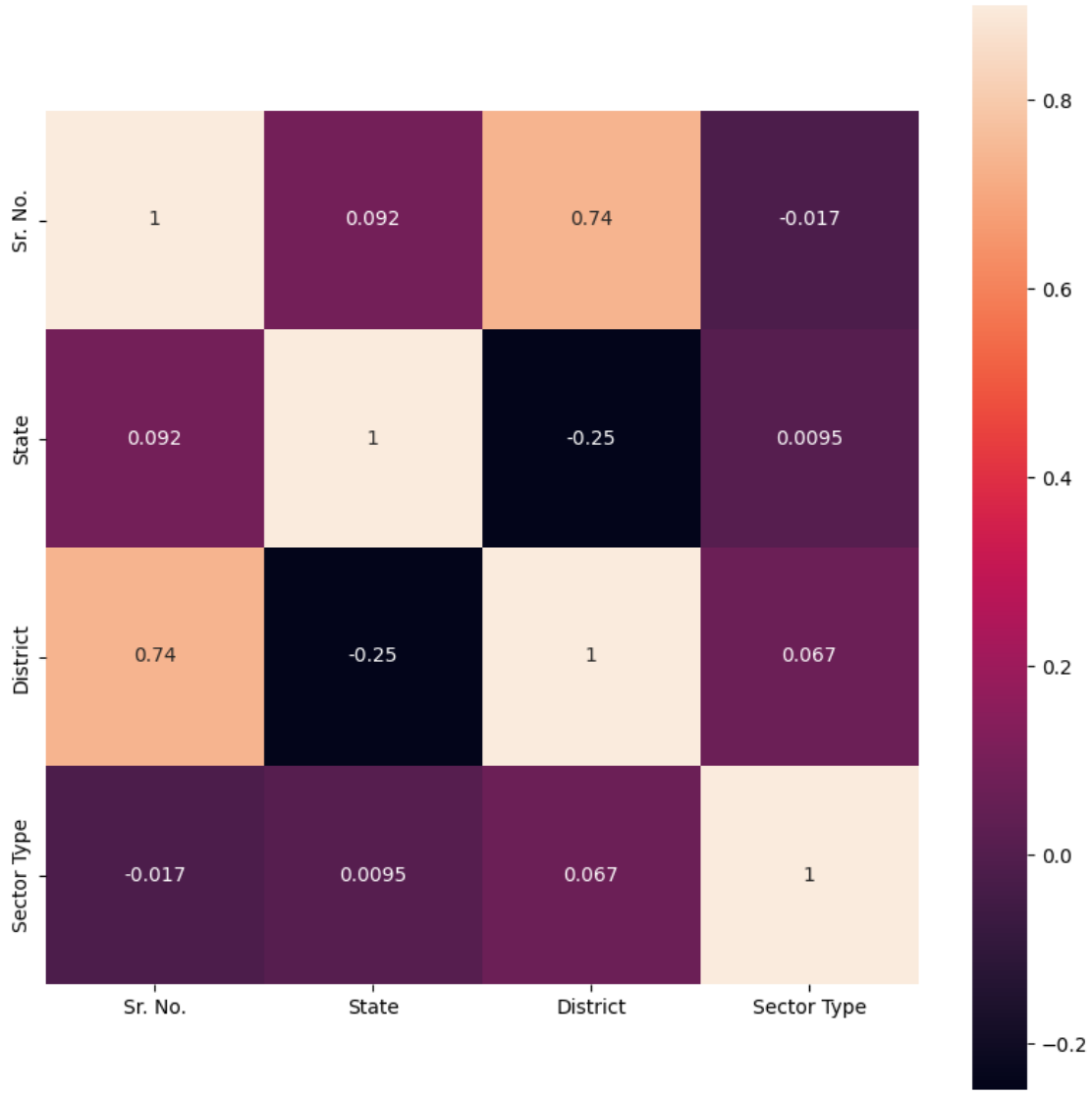
```
<ipython-input-62-fc9e19218782>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(df4['Sector Type'],color='red');
```



Data Analysis

[63]:
```
# correlation matrix
from sklearn.metrics import confusion_matrix
fig, ax = plt.subplots(figsize=(10, 10))
correlation_metrics = df2.corr()
fig = plt.figure(figsize = (6,6))
sns.heatmap(correlation_metrics , vmax = 0.9 , square = True, annot = True ,␣
 ↪ax=ax)
plt.show()
```

```
<ipython-input-63-64bfc29b0dae>:4: FutureWarning: The default value of
numeric_only in DataFrame.corr is deprecated. In a future version, it will
default to False. Select only valid columns or specify the value of numeric_only
to silence this warning.
  correlation_metrics = df2.corr()
```

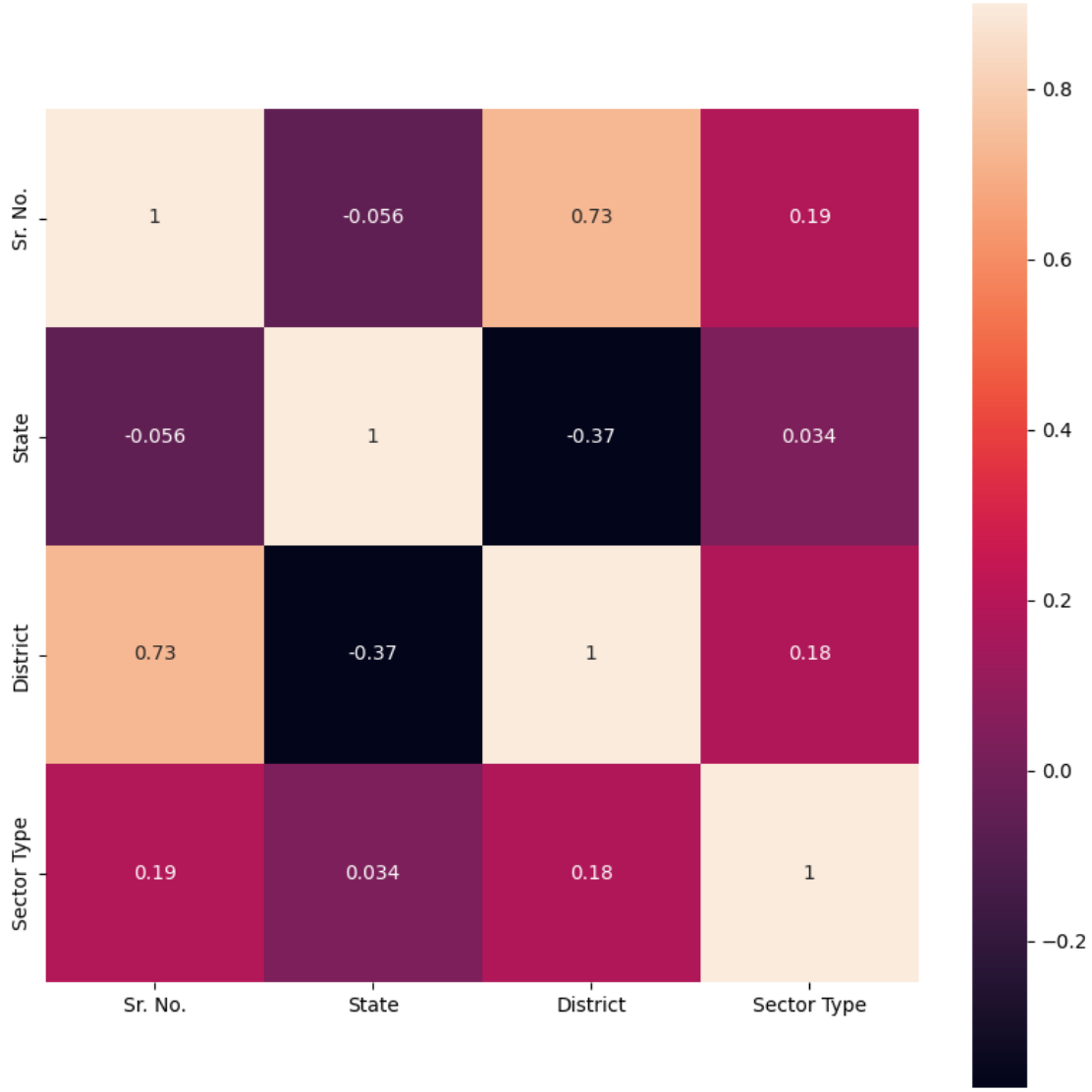



```
<Figure size 600x600 with 0 Axes>
```

[64]:
```
from sklearn.metrics import confusion_matrix
correlation_metrics = df3.corr()
fig, ax = plt.subplots(figsize=(10, 10))
fig = plt.figure(figsize = (6,6))
sns.heatmap(correlation_metrics , vmax = 0.9 , square = True , annot = True,␣
↪ax=ax)
plt.show()
```

```
<ipython-input-64-debb6f466864>:2: FutureWarning: The default value of
numeric_only in DataFrame.corr is deprecated. In a future version, it will
default to False. Select only valid columns or specify the value of numeric_only
```

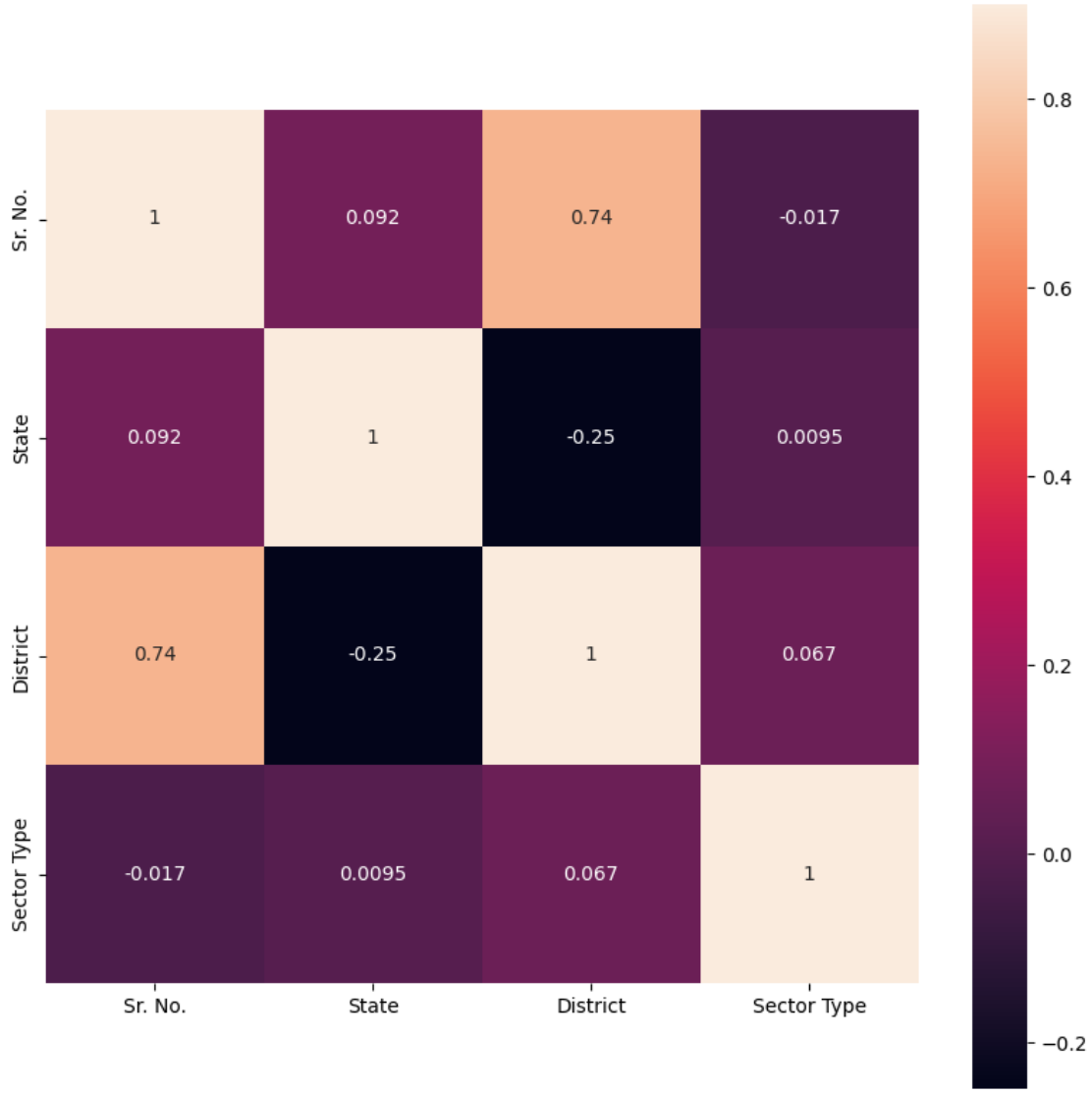
```
to silence this warning.
  correlation_metrics = df3.corr()
```




```
<Figure size 600x600 with 0 Axes>
```

[65]:
```
from sklearn.metrics import confusion_matrix
fig, ax = plt.subplots(figsize=(10, 10))
correlation_metrics = df4.corr()
fig = plt.figure(figsize = (6,6))
sns.heatmap(correlation_metrics , vmax = 0.9 , square = True ,  annot = True ,␣
 ↪ax=ax)
plt.show()
```

```
<ipython-input-65-f257e7c44604>:3: FutureWarning: The default value of
numeric_only in DataFrame.corr is deprecated. In a future version, it will
default to False. Select only valid columns or specify the value of numeric_only
to silence this warning.
  correlation_metrics = df4.corr()
```




```
<Figure size 600x600 with 0 Axes>
```

[ ]:
```
df2.groupby('State').mean()
```

```
<ipython-input-65-c5aa51dd9e70>:1: FutureWarning: The default value of
numeric_only in DataFrameGroupBy.mean is deprecated. In a future version,
numeric_only will default to False. Either specify numeric_only or select only
```

```
columns which should be valid for the function.
```

[ ]:
```
          Sr. No.   District  Sector Type
State
1       33.500000   7.750000     1.000000
4       37.230769  10.769231     2.230769
6       49.000000  18.000000     1.000000
7       44.000000  11.000000     1.000000
8       50.000000  19.000000     1.000000
9       37.000000  12.000000     4.000000
10      39.000000   6.000000     1.000000
11      26.000000   2.000000     1.000000
12      36.500000   3.000000     2.000000
```

[ ]:
```
df2.groupby('District').mean()
```

```
<ipython-input-66-b550094563e8>:1: FutureWarning: The default value of
numeric_only in DataFrameGroupBy.mean is deprecated. In a future version,
numeric_only will default to False. Either specify numeric_only or select only
columns which should be valid for the function.
```

[ ]:
```
           Sr. No.  State  Sector Type
District
1             25.0    1.0     1.000000
2             26.0   11.0     1.000000
3             36.5   12.0     2.000000
4             28.0    1.0     1.000000
5             29.0    4.0     1.000000
6             39.0   10.0     1.000000
7             33.0    4.0     2.666667
8             33.0    4.0     2.000000
9             37.0    4.0     3.500000
10            35.0    4.0     1.000000
11            44.0    7.0     1.000000
12            37.0    9.0     4.000000
13            40.5    1.0     1.000000
14            39.0    4.0     1.000000
15            41.0    4.0     1.000000
16            44.5    4.0     1.000000
17            45.0    4.0     6.000000
18            49.0    6.0     1.000000
19            50.0    8.0     1.000000
```

[ ]:
```
df2.groupby('Sector Type').mean()
```

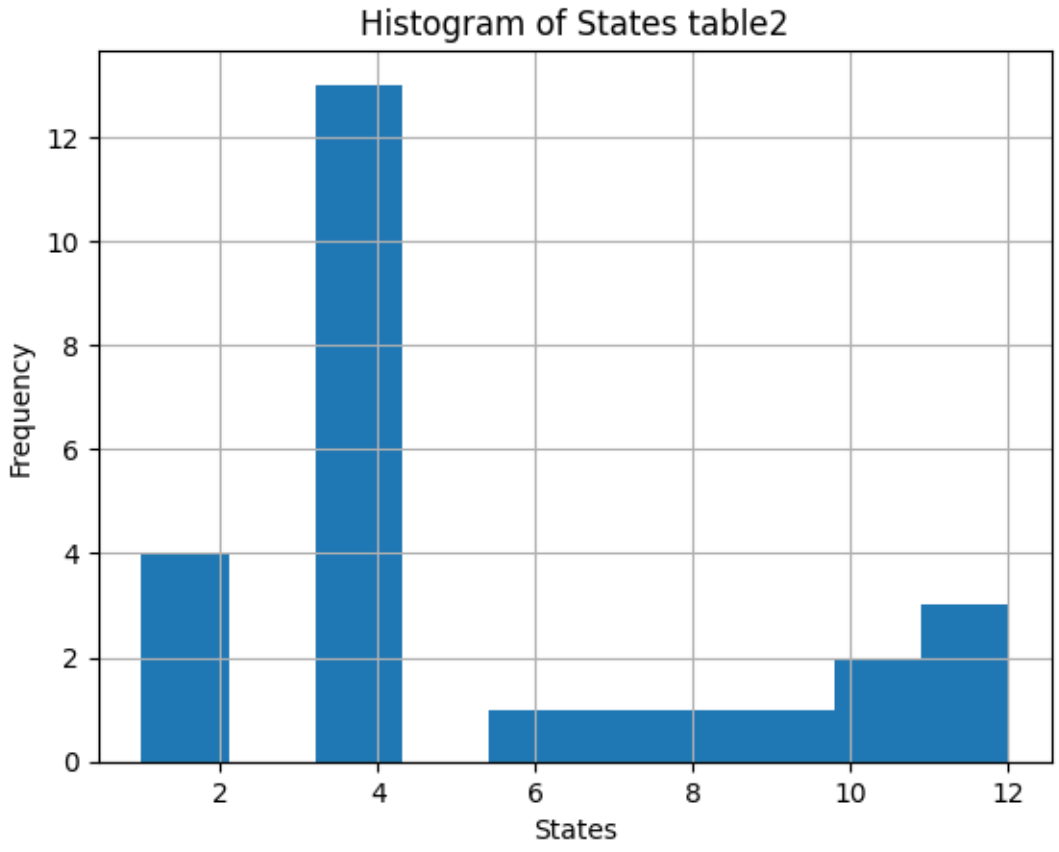
```
<ipython-input-67-1267dd2598cc>:1: FutureWarning: The default value of
numeric_only in DataFrameGroupBy.mean is deprecated. In a future version,
numeric_only will default to False. Either specify numeric_only or select only
```

columns which should be valid for the function.

[ ]:
```
                Sr. No.     State   District
Sector Type
1             38.823529  5.411765  10.117647
2             33.000000  4.000000   8.000000
3             31.333333  6.666667   5.666667
4             37.000000  9.000000  12.000000
5             40.000000  4.000000   9.000000
6             45.000000  4.000000  17.000000
```

[72]:
```
df2.State.hist()
plt.title('Histogram of States table2')
plt.xlabel('States')
plt.ylabel('Frequency')
```

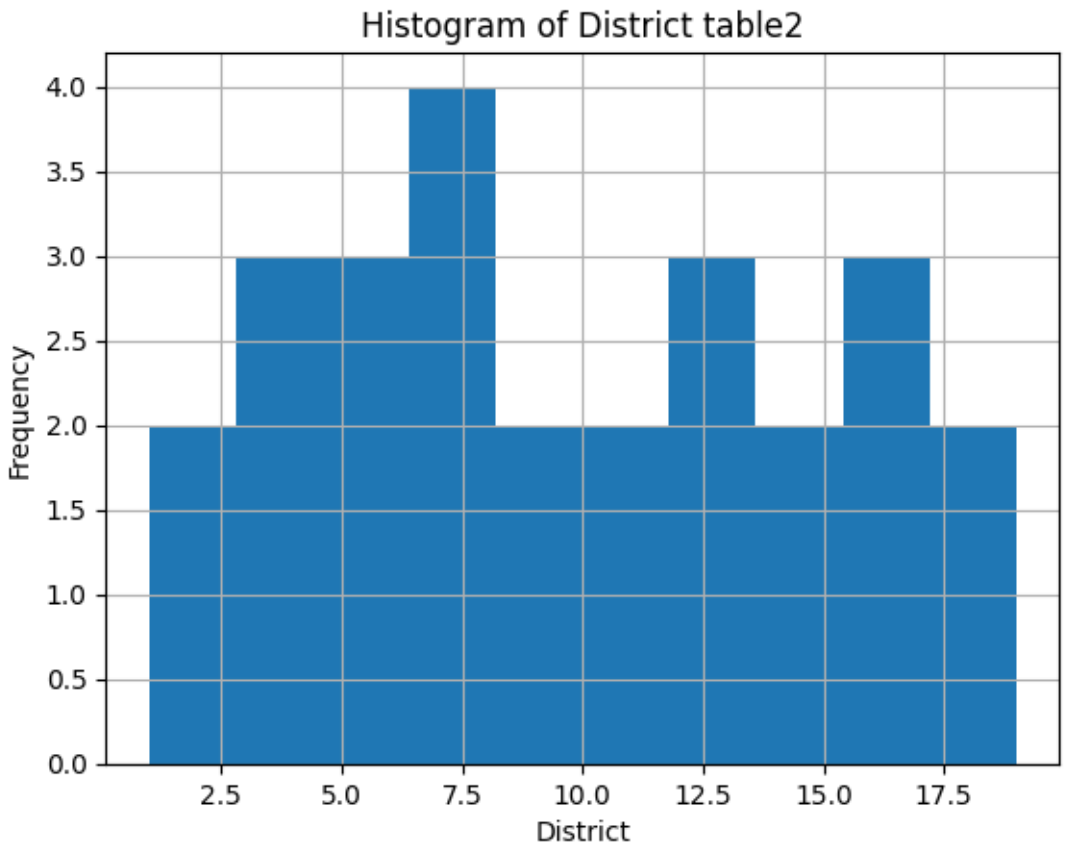
[72]:
```
Text(0, 0.5, 'Frequency')
```



[ ]:
```
df2.District.hist()
plt.title('Histogram of District table2')
```

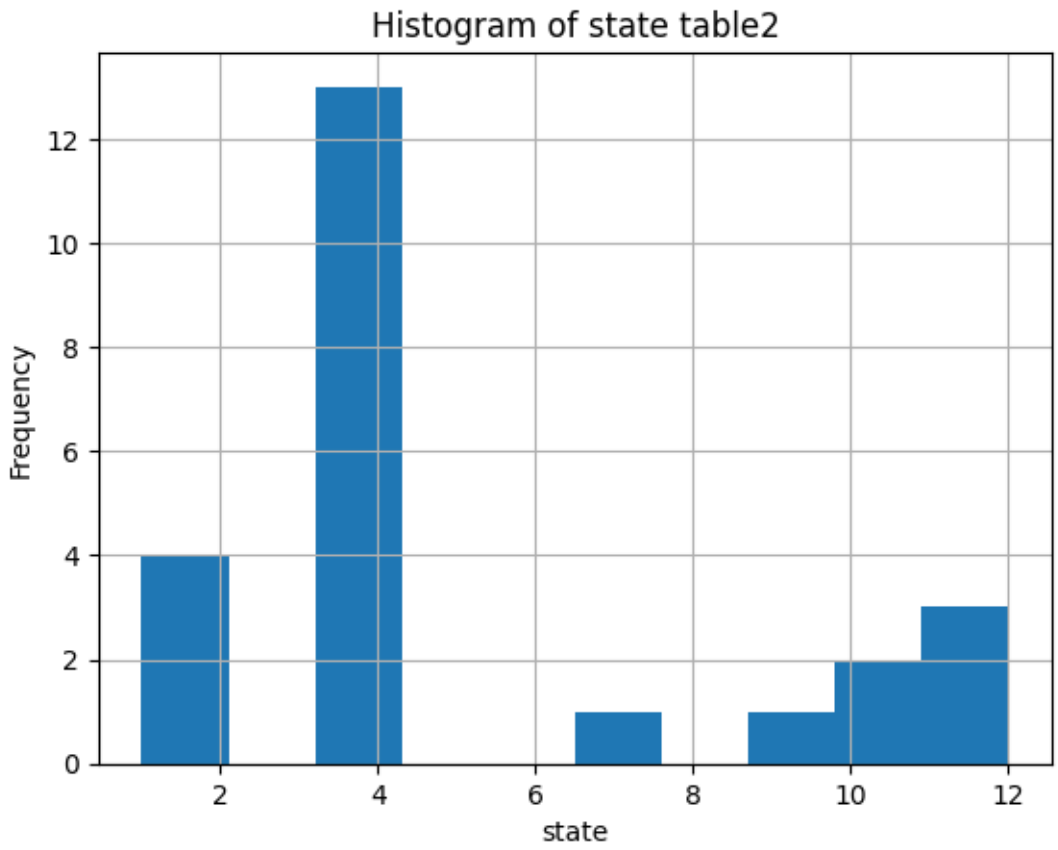
```
plt.xlabel('District')
plt.ylabel('Frequency')
```

[ ]: Text(0, 0.5, 'Frequency')



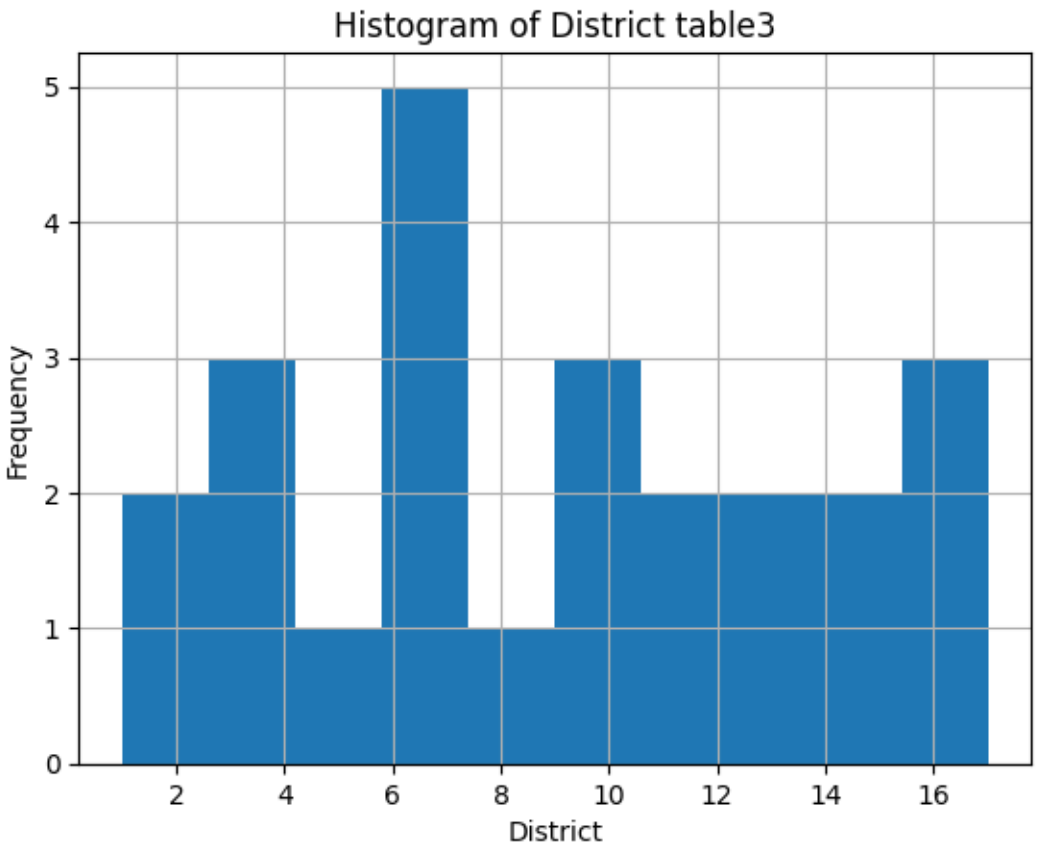
[ ]:
```
df3.State.hist()
plt.title('Histogram of state table2')
plt.xlabel('state')
plt.ylabel('Frequency')
```

[ ]: Text(0, 0.5, 'Frequency')



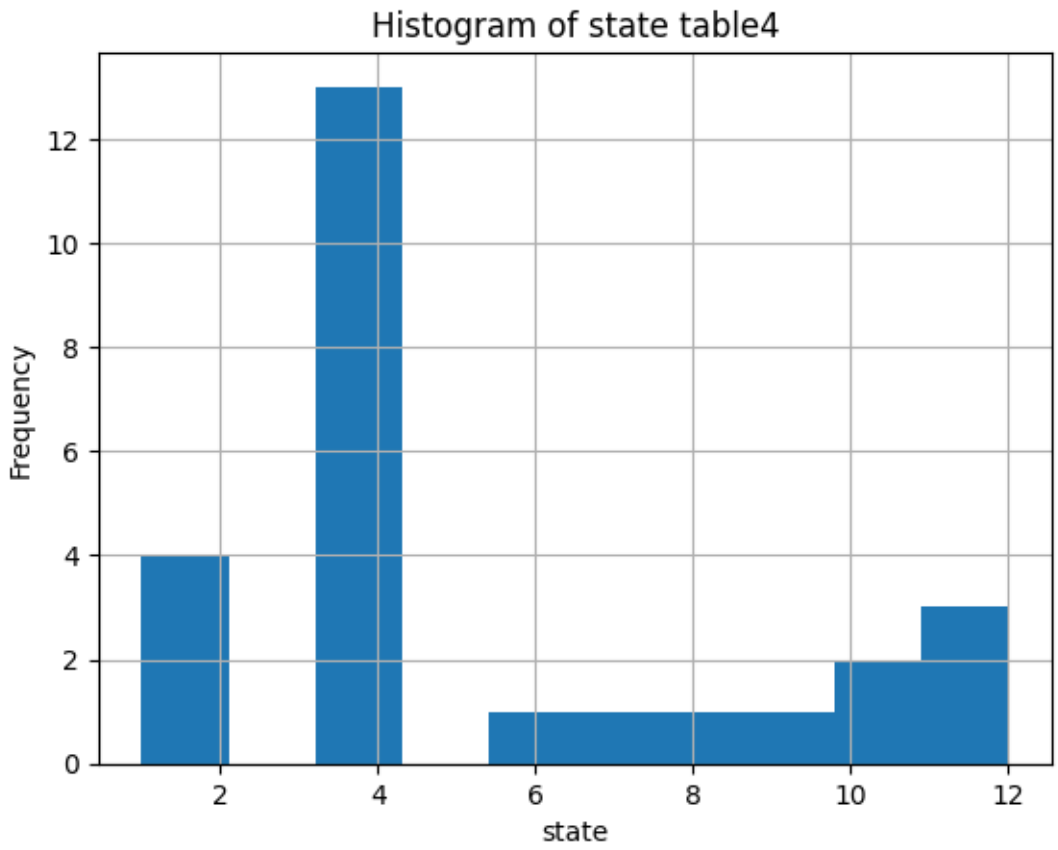
[ ]:
```
df3.District.hist()
plt.title('Histogram of District table3')
plt.xlabel('District')
plt.ylabel('Frequency')
```

[ ]: Text(0, 0.5, 'Frequency')



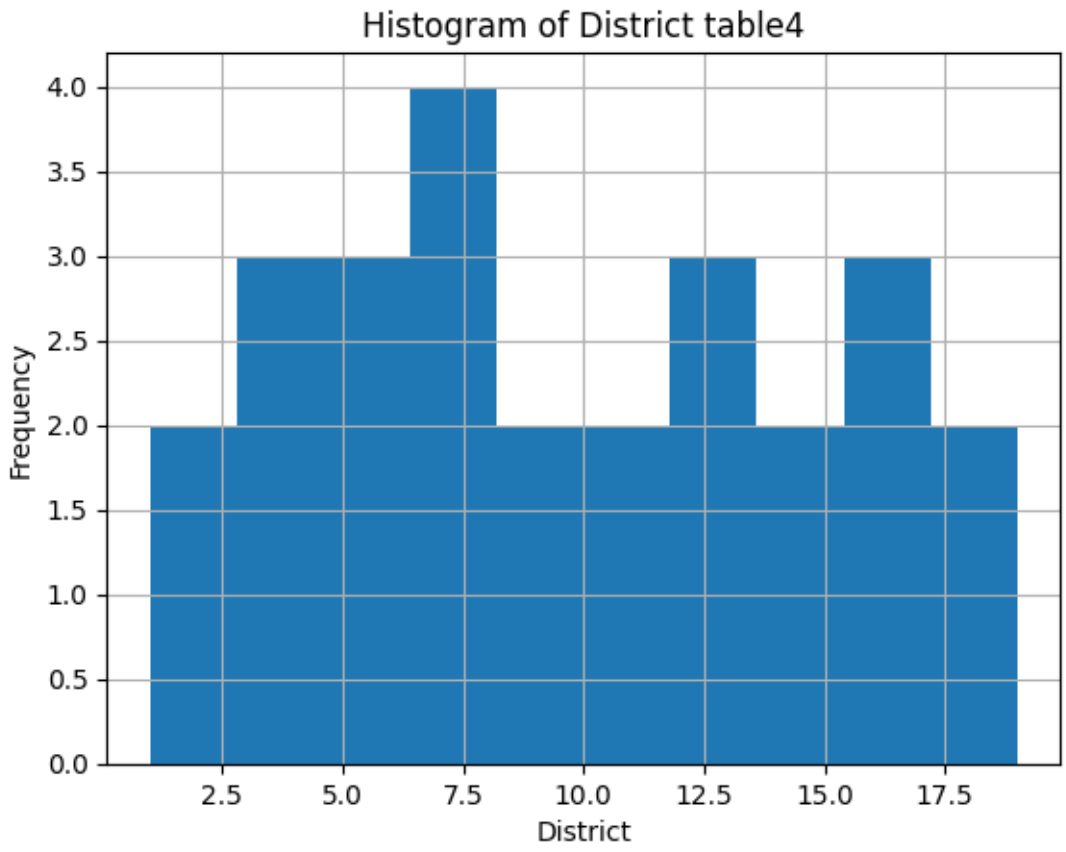
[ ]:
```
df4.State.hist()
plt.title('Histogram of state table4')
plt.xlabel('state')
plt.ylabel('Frequency')
```

[ ]:
```
Text(0, 0.5, 'Frequency')
```

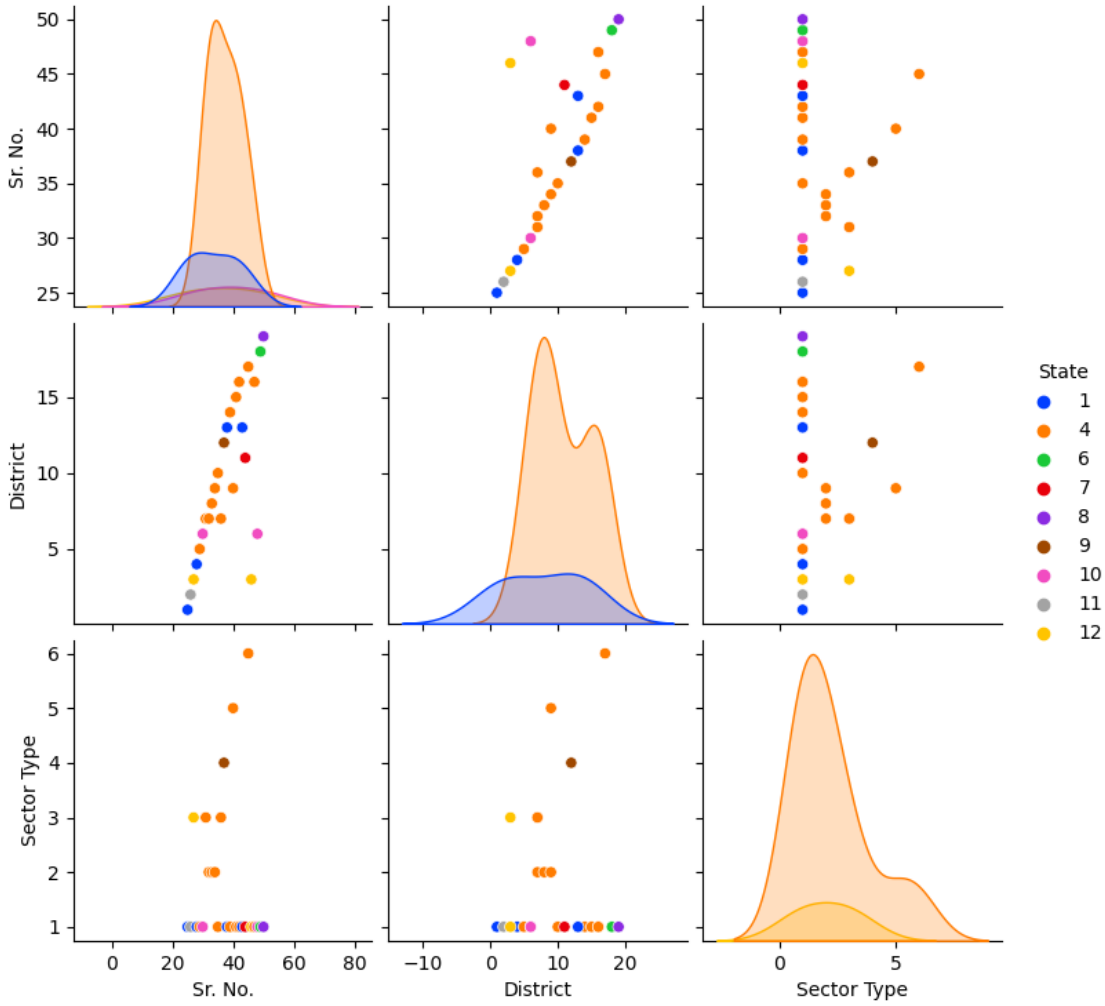


[ ]:
```
df4.District.hist()
plt.title('Histogram of District table4')
plt.xlabel('District')
plt.ylabel('Frequency')
```

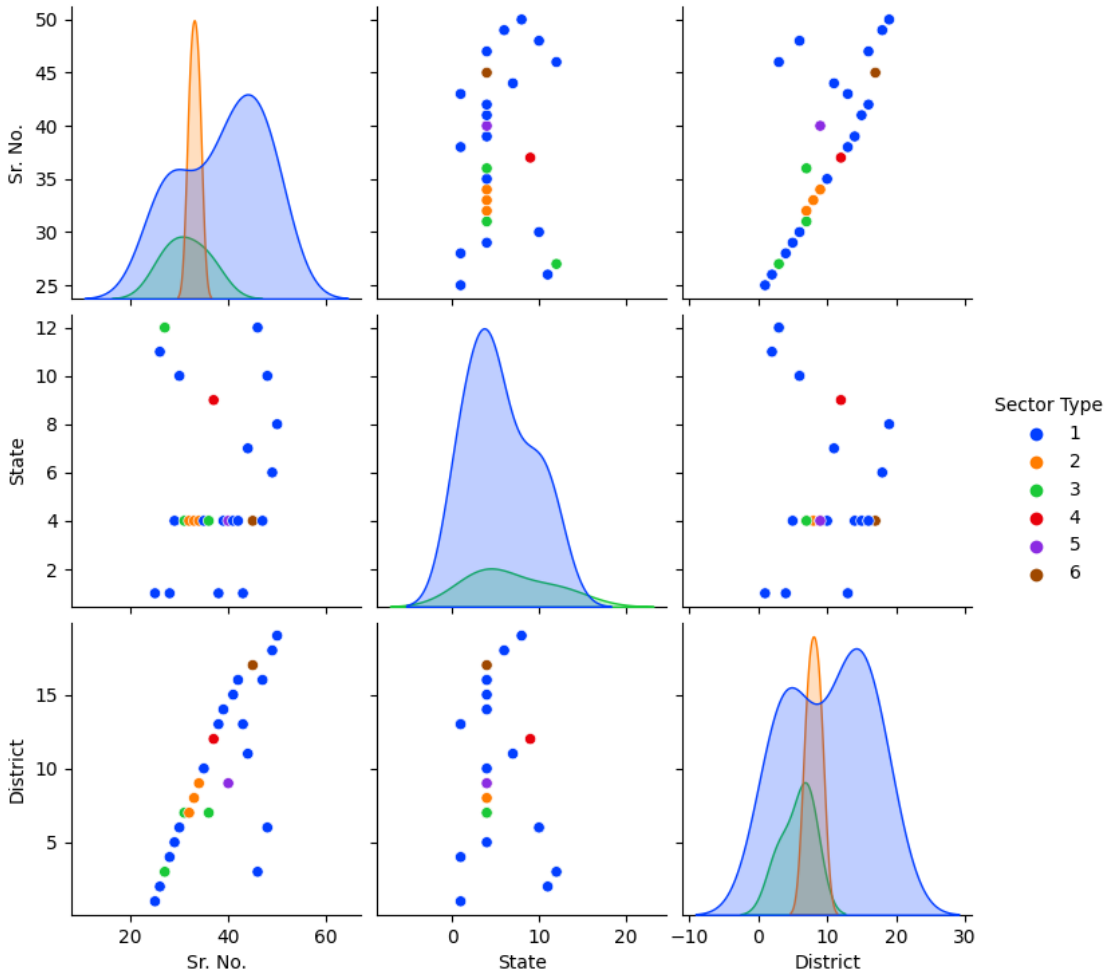
[ ]:
```
Text(0, 0.5, 'Frequency')
```



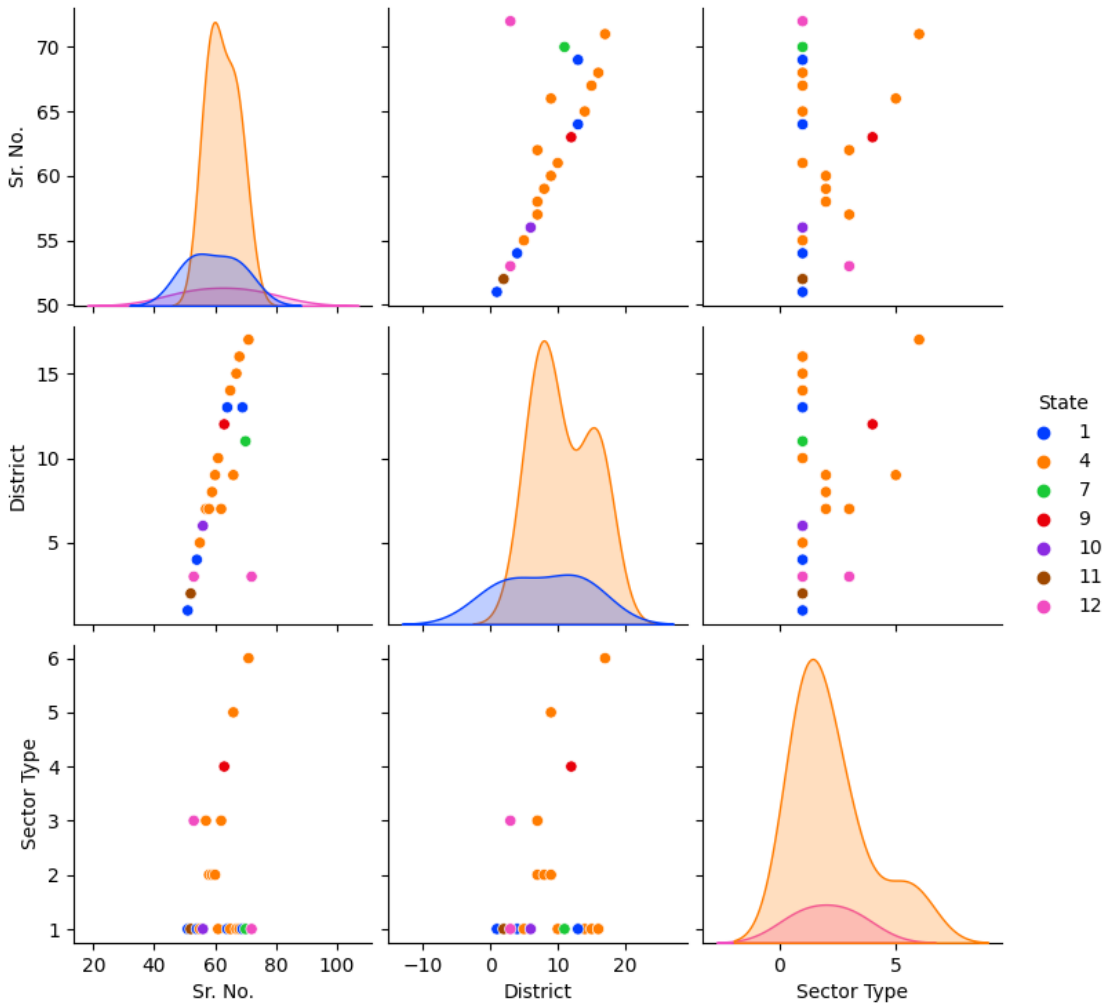
[ ]:
```
#plotting variable distribution with seaborn
sns.pairplot(df2, hue="State", palette='bright');
```



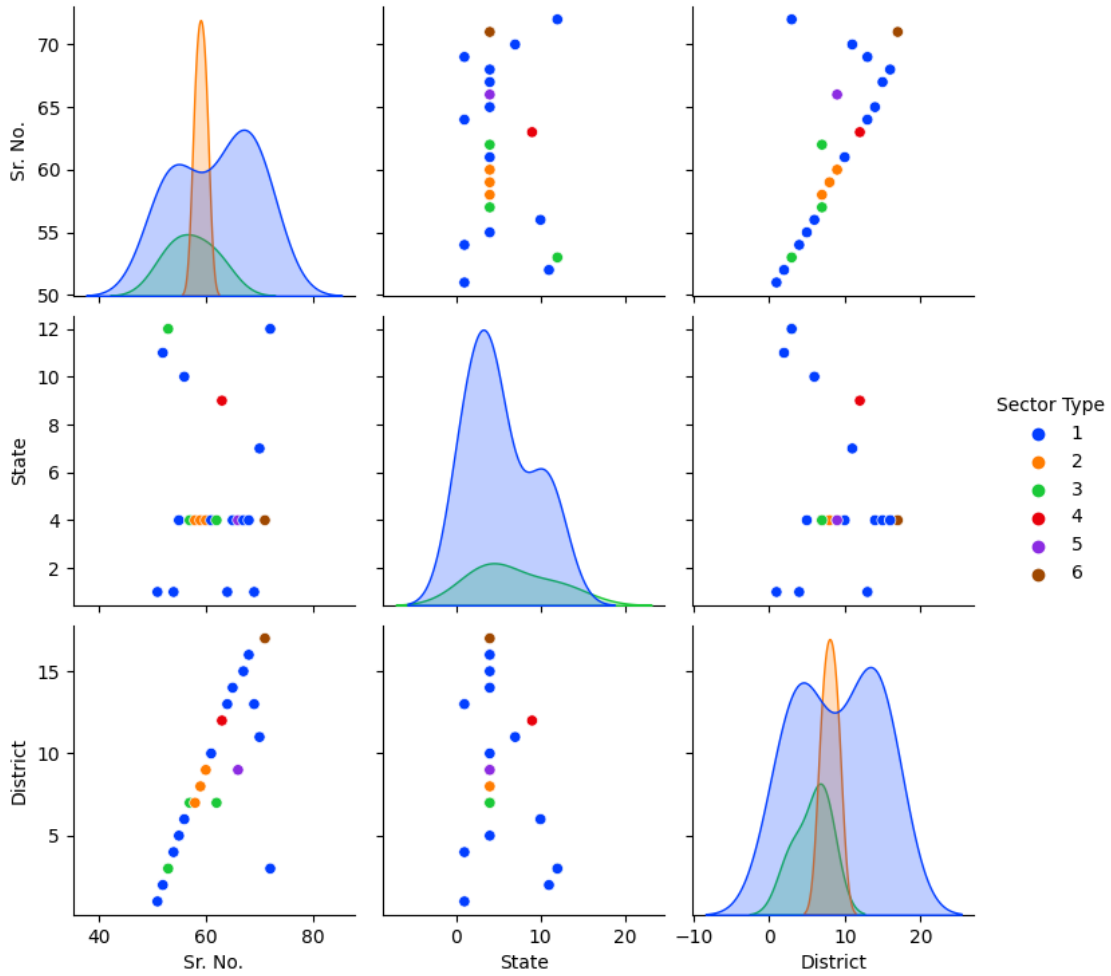
[ ]:
```
sns.pairplot(df2, hue="Sector Type", palette='bright');
```



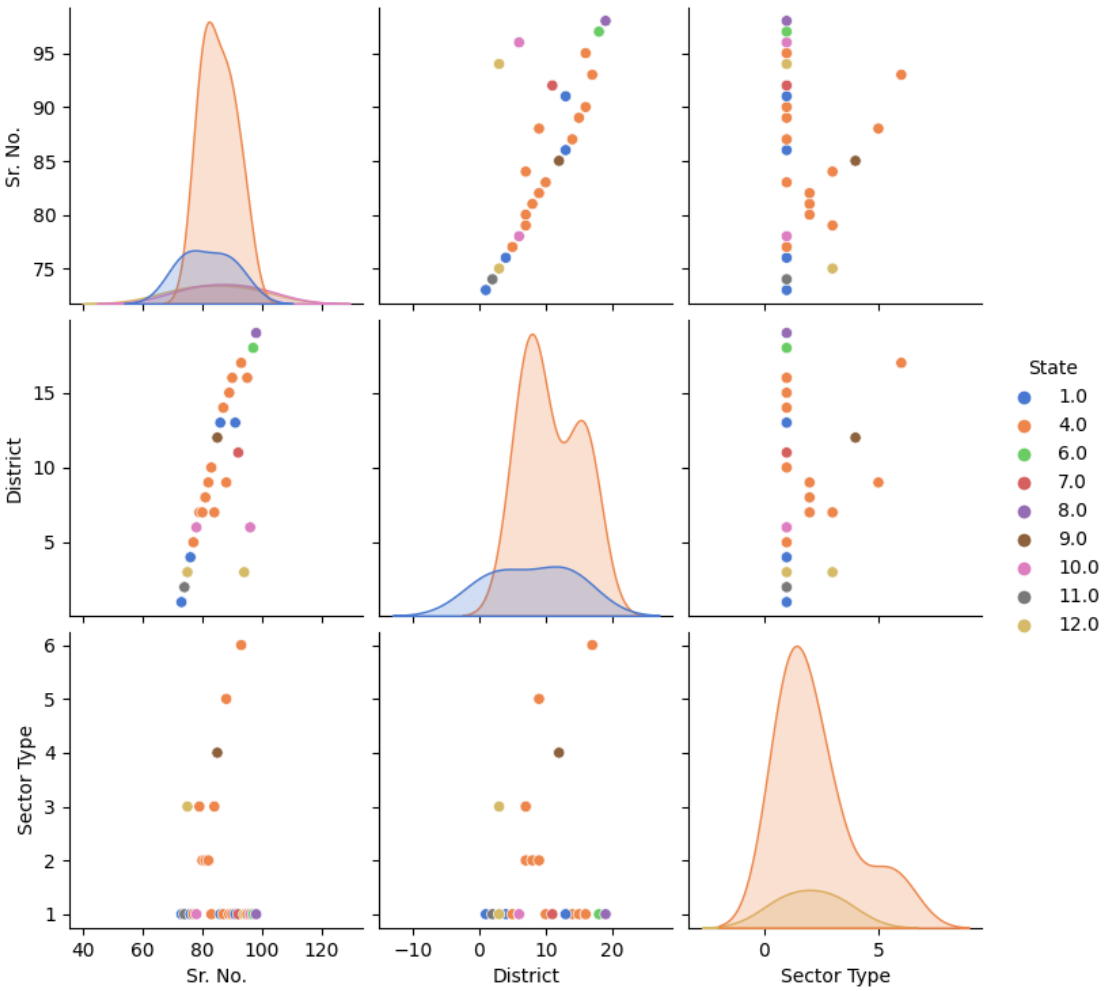
[ ]:
```
sns.pairplot(df3, hue="State", palette='bright');
```



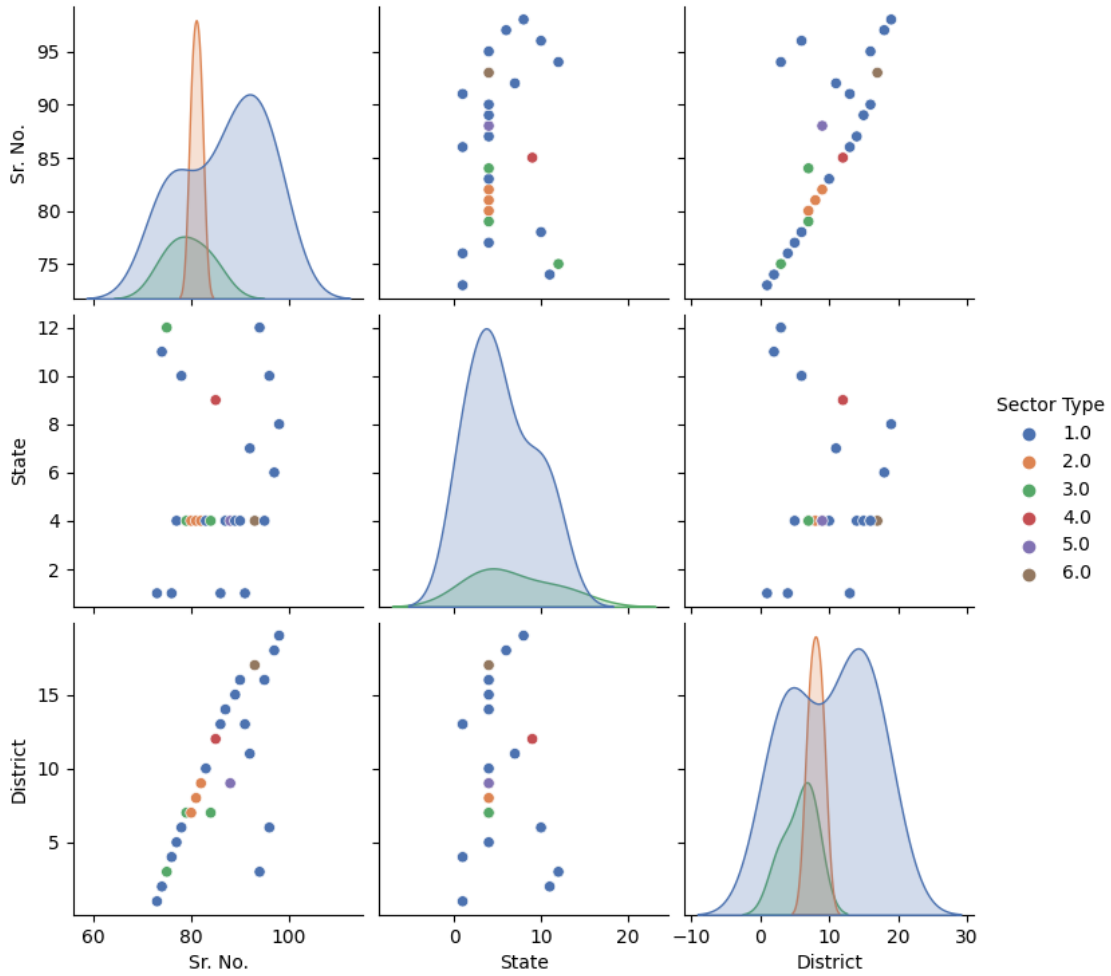
[ ]:
```
sns.pairplot(df3, hue="Sector Type", palette='bright');
```



[ ]:
```
sns.pairplot(df4, hue="State", palette='muted');
```



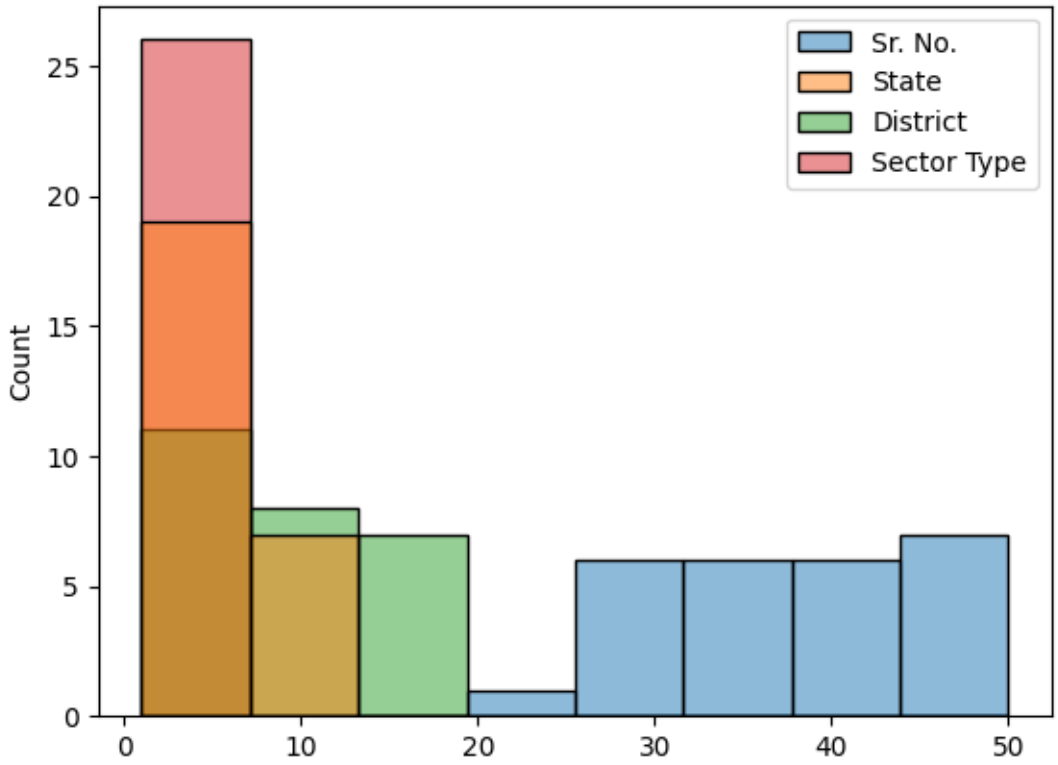
[ ]:
```
sns.pairplot(df4, hue="Sector Type", palette='deep');
```



[ ]:
```
#https://thedatafrog.com/en/articles/visualizing-datasets/
```

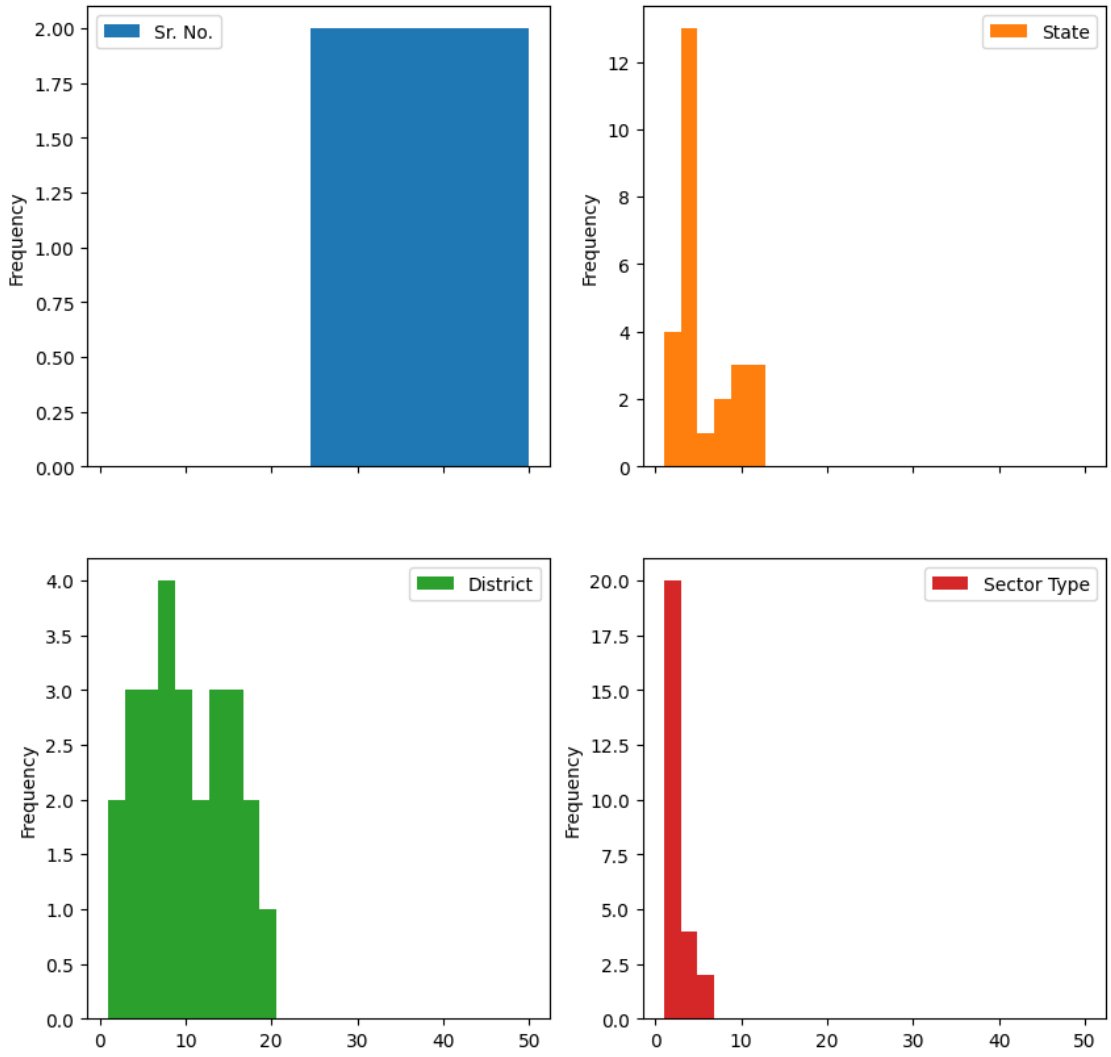
[76]:
```
sns.histplot(data = df2)
```

[76]:
```
<Axes: ylabel='Count'>
```



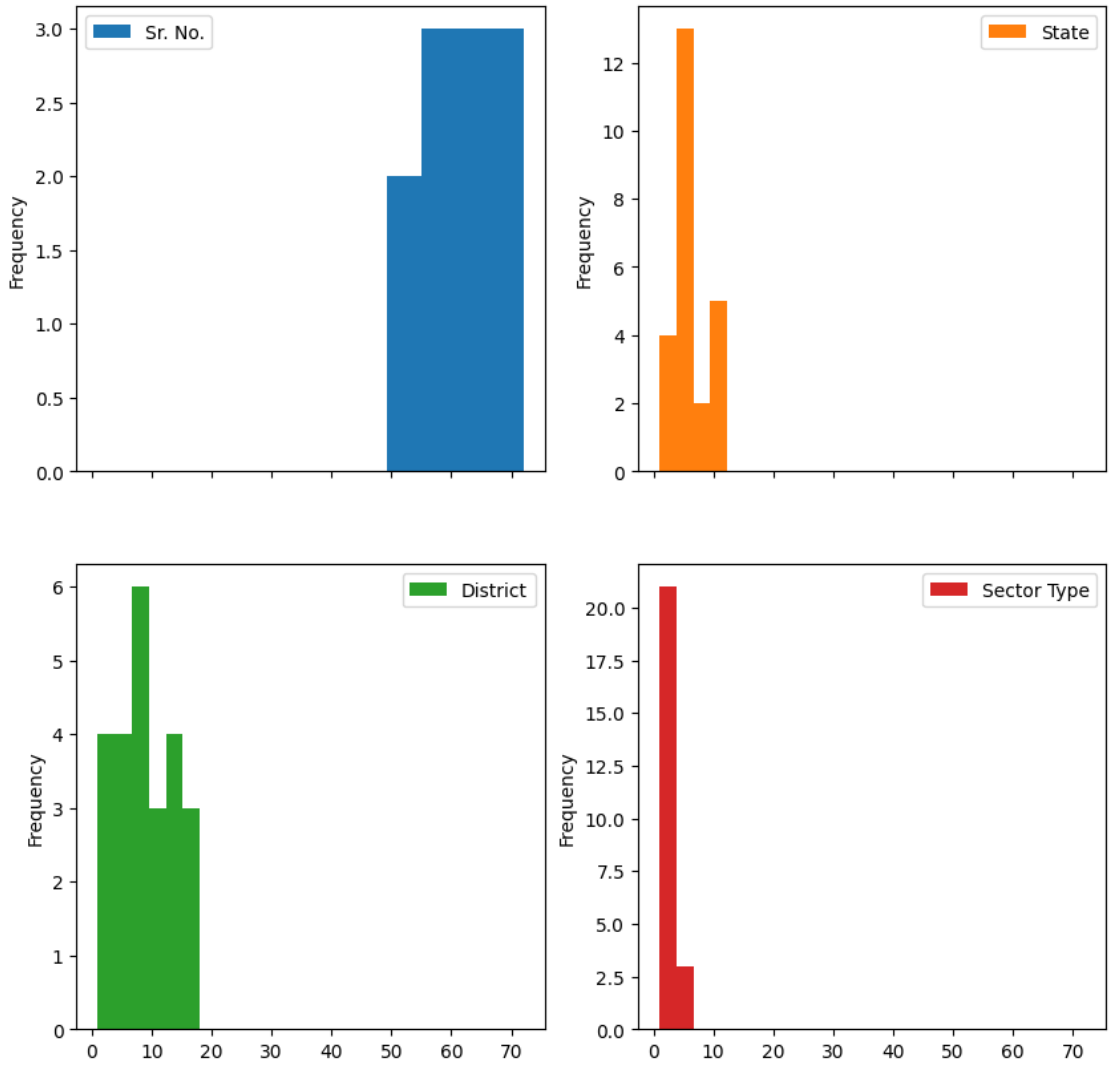
[81]:
```
df2.plot.hist(subplots=True, layout=(2,2), figsize=(10, 10), bins = 25)
```

[81]:
```
array([[<Axes: ylabel='Frequency'>, <Axes: ylabel='Frequency'>],
       [<Axes: ylabel='Frequency'>, <Axes: ylabel='Frequency'>]],
      dtype=object)
```



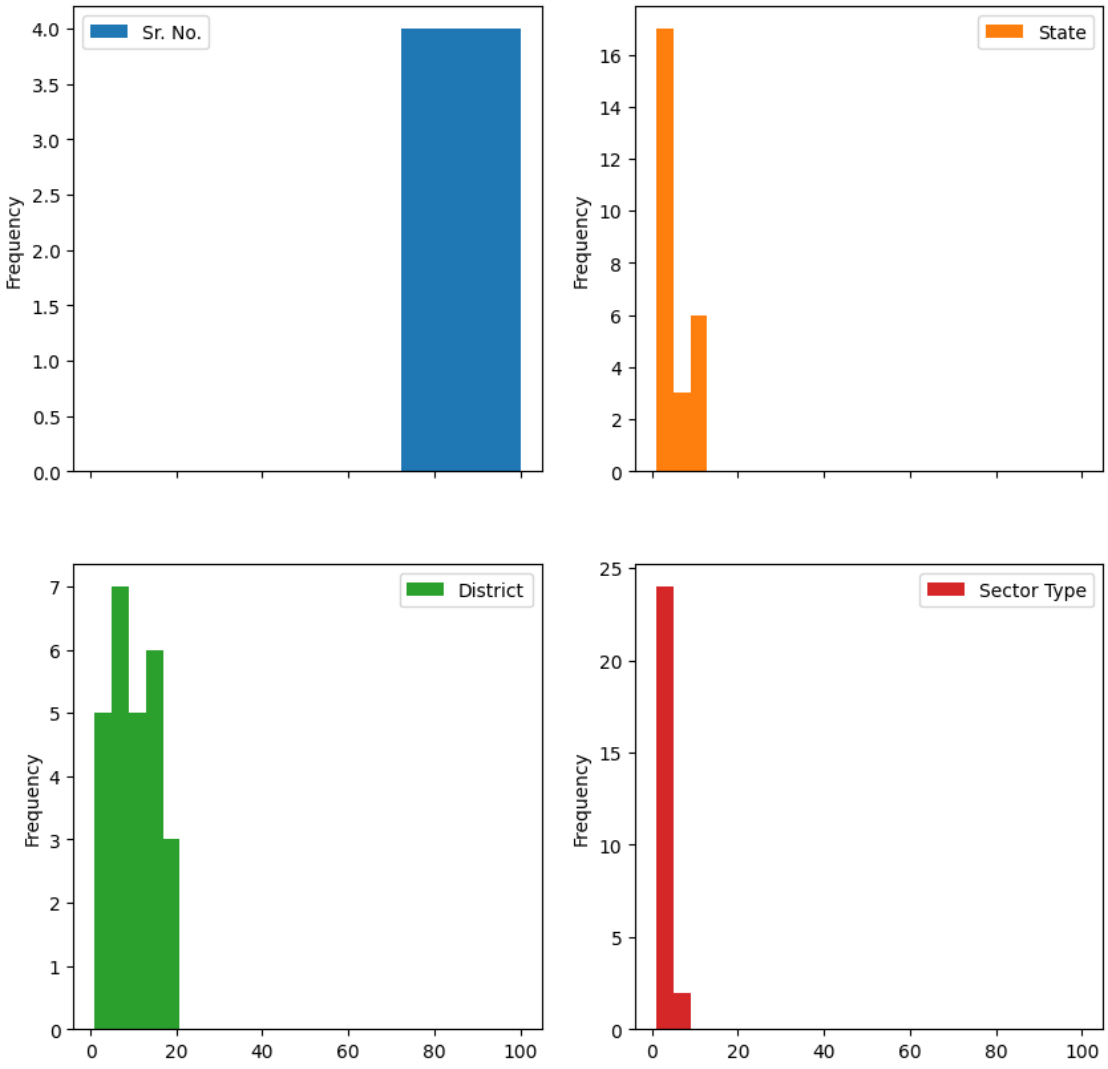
[82]:
```
df3.plot.hist(subplots=True, layout=(2,2), figsize=(10, 10), bins = 25)
```

[82]:
```
array([[<Axes: ylabel='Frequency'>, <Axes: ylabel='Frequency'>],
       [<Axes: ylabel='Frequency'>, <Axes: ylabel='Frequency'>]],
      dtype=object)
```



[83]:
```
df4.plot.hist(subplots=True, layout=(2,2), figsize=(10, 10), bins = 25)
```

[83]:
```
array([[<Axes: ylabel='Frequency'>, <Axes: ylabel='Frequency'>],
       [<Axes: ylabel='Frequency'>, <Axes: ylabel='Frequency'>]],
      dtype=object)
```



Analysis of one hot encoded columns namely state, district and sector type is performed for finding patterns and relationships between then for ease of interpretation.