In [2]:
```
%matplotlib inline
```

Data Analysis & Visualisation of COVID-19 Cases by Adit Kotak

In [3]:
```
import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
```

In [4]:
```
data = pd.read_csv('district.csv')
```

In [ ]:
```
# Describe the statistics of all the the columns.
```

In [5]:
```
data.describe()
```

Out[5]:

	districtData/0/active	districtData/0/confirmed	districtData/0/deceased	districtData/0/recov
count	33.000000	33.000000	33.000000	33.00
mean	249.818182	317.909091	13.878788	54.21
std	994.971936	1238.750034	51.887955	193.10
min	0.000000	1.000000	0.000000	0.00
25%	2.000000	3.000000	0.000000	1.00
50%	14.000000	25.000000	1.000000	5.00
75%	69.000000	79.000000	4.000000	22.00
max	5679.000000	7061.000000	290.000000	1092.00

In [6]:
```
data.head(15)
```

Out[6]:

	districtData/0/district	districtData/0/active	districtData/0/confirmed	districtData/0/deceased
0	Ahmadnagar	17	42	2
1	Yavatmal	69	79	0
2	Washim	1	2	0
3	Solapur	93	99	6
4	Sindhudurg	1	2	0
5	Satara	21	32	2
6	Sangli	3	29	1
7	Ratnagiri	2	8	1
8	Raigarh	44	71	3
9	Parbhani	1	2	0
10	Palghar	119	169	4
11	Osmanabad	0	3	0
12	Nashik	179	197	12
13	Nandurbar	10	11	1
14	Nanded	3	3	0

In [ ]:

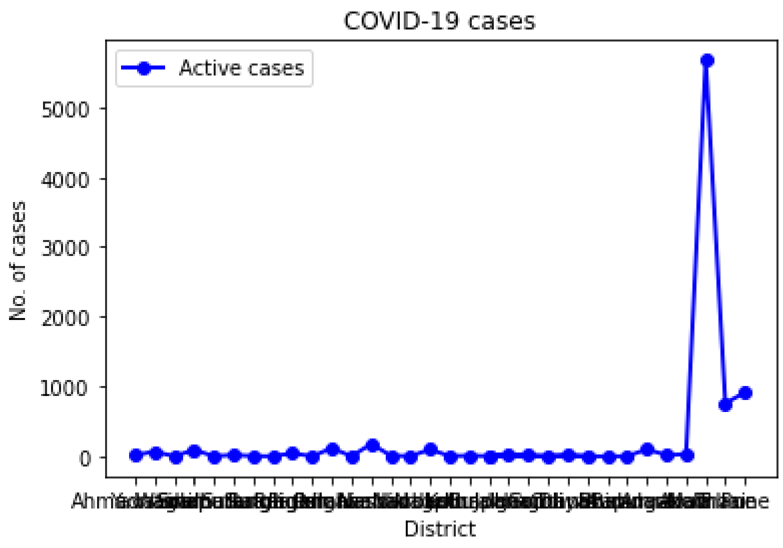
```
# Plot a line diagram including active, confirmed, deceased & recovered cases.
```

In [178…

```
A = data.iloc[0:,1].values
R = data.iloc[0:,4].values
D = data.iloc[0:,3].values
C = data.iloc[0:,2].values
X = data.iloc[0:,0]
```

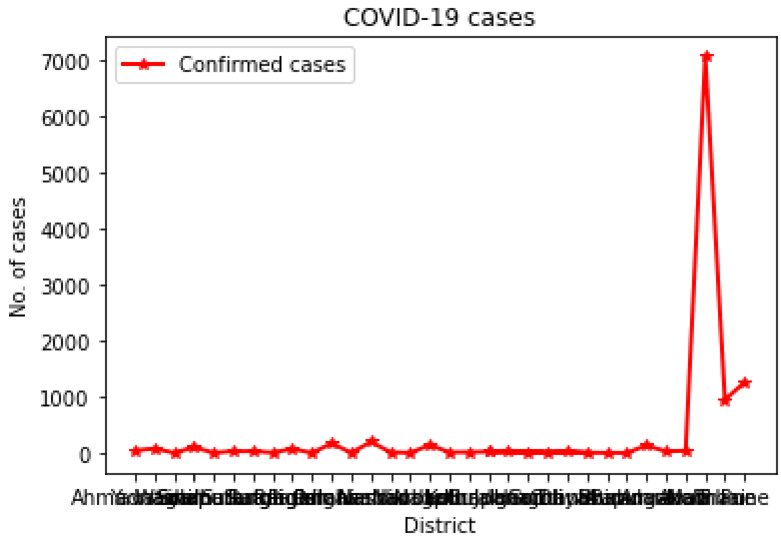
In [173…

```
plt.plot(X,A, label="Active cases", color="b", linewidth=2, marker='o')
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



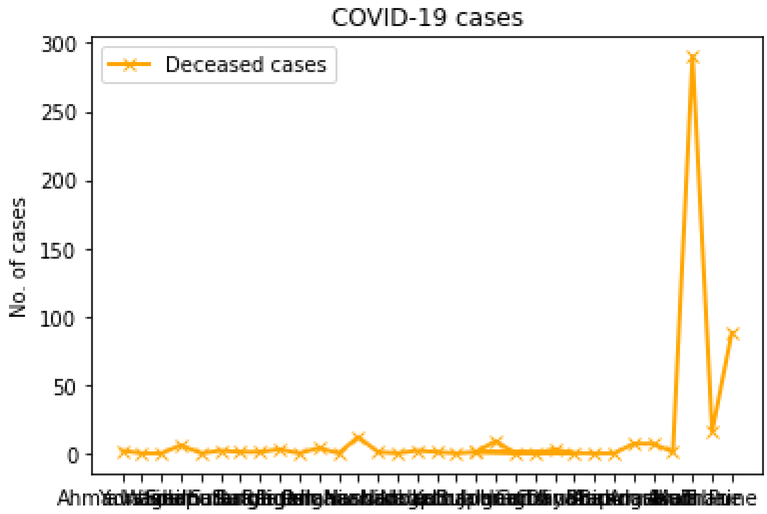
In [174…

```
plt.plot(X,C, label="Confirmed cases", color="r", linewidth=2, marker='*')
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



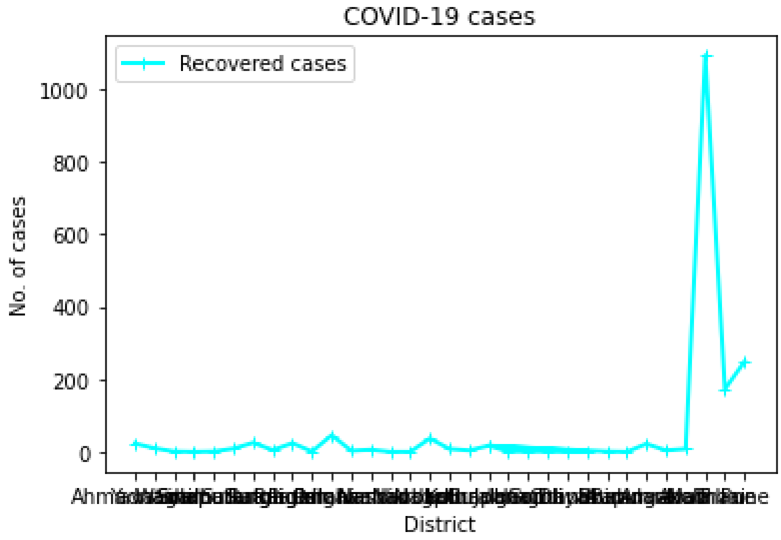
In [175…

```
plt.plot(X,D, label="Deceased cases", color="orange", linewidth=2, marker='x')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



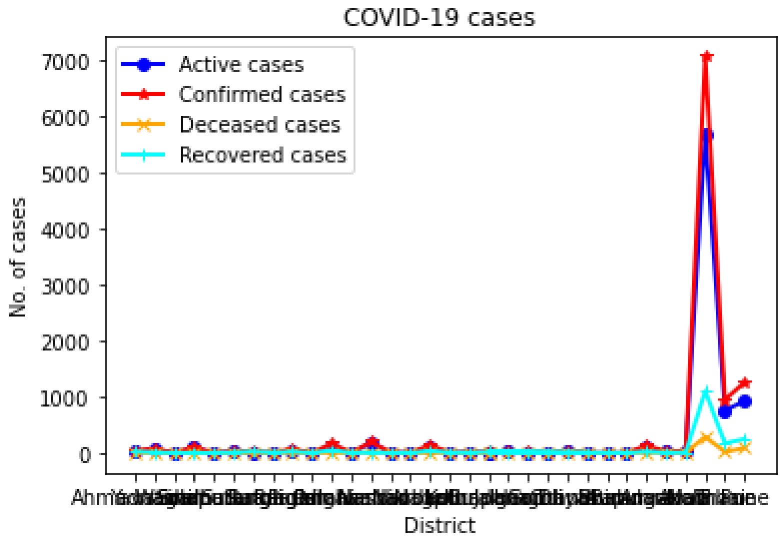
In [176…

```
plt.plot(X,R, label="Recovered cases", color="cyan", linewidth=2, marker='+')
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



In [179…

```
plt.plot(X,A, label="Active cases", color="b", linewidth=2, marker='o')
plt.plot(X,C, label="Confirmed cases", color="r", linewidth=2, marker='*')
plt.plot(X,D, label="Deceased cases", color="orange", linewidth=2, marker='x')
plt.plot(X,R, label="Recovered cases", color="cyan", linewidth=2, marker='+')
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```

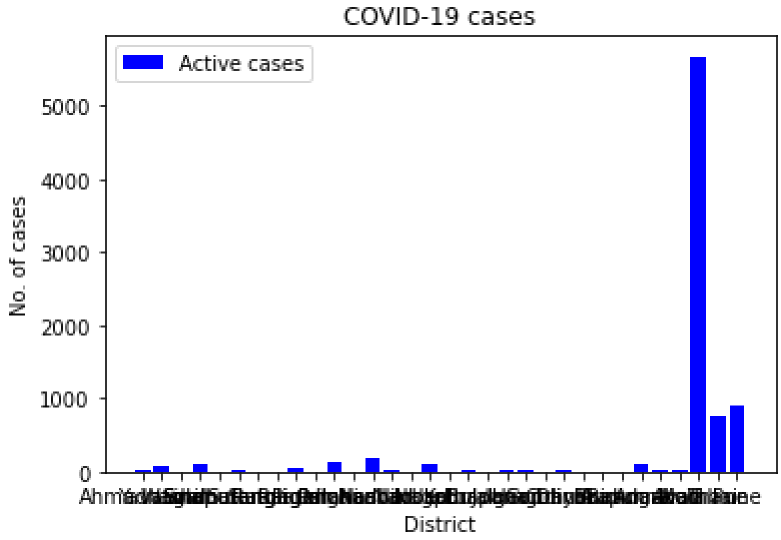


In [ ]:

```
# Plot a bar diagram including active, confirmed, deceased & recovered cases.
```

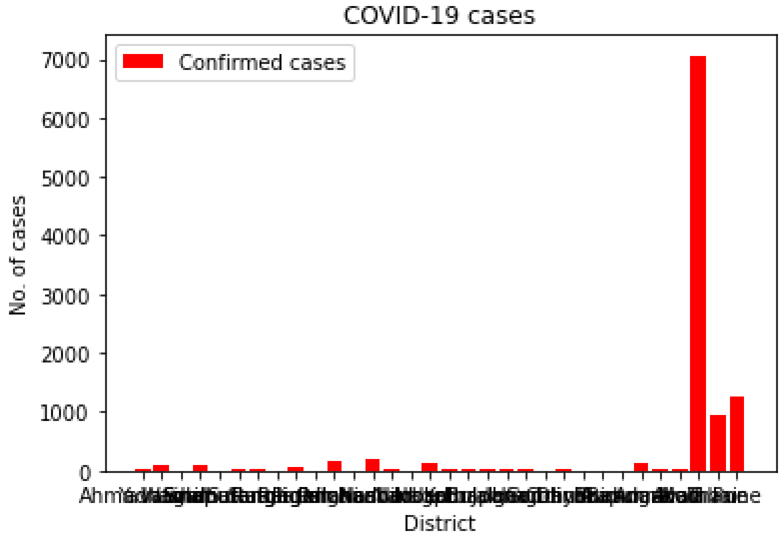
In [476…

```
plt.bar(X, A, label="Active cases", color="b")
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



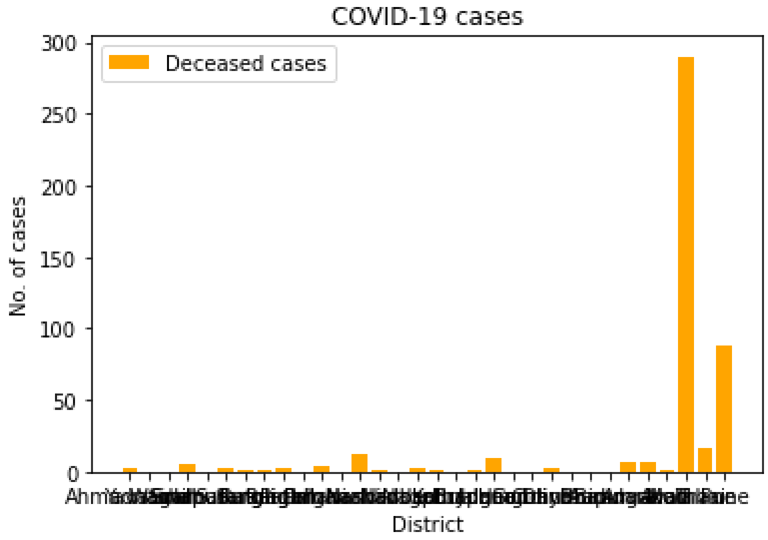
In [160…

```
plt.bar(X, C, label="Confirmed cases", color="r")
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



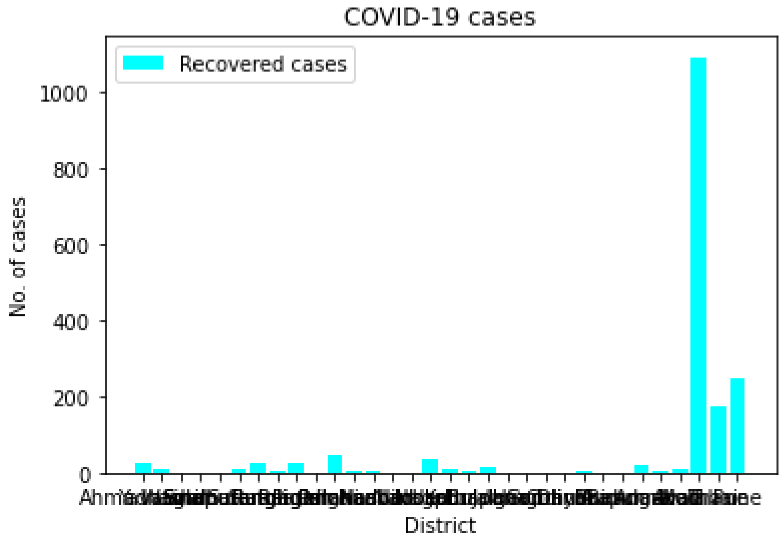
In [161…

```
plt.bar(X, D, label="Deceased cases", color="orange")
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



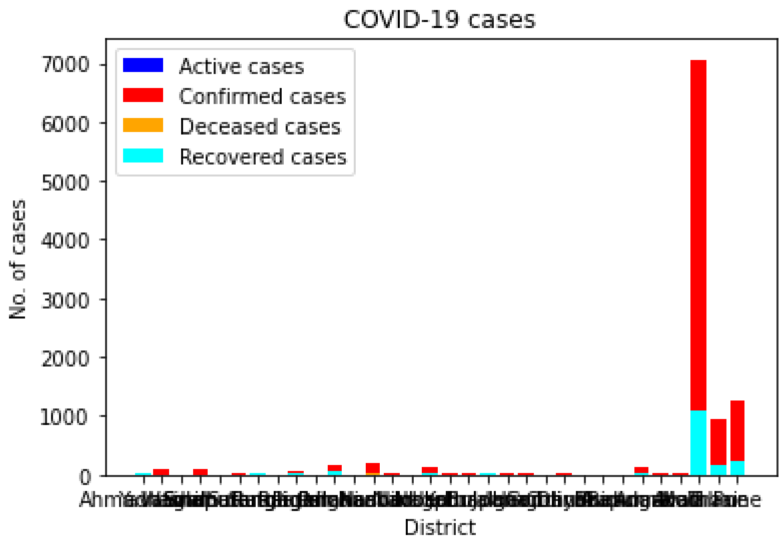
In [162…

```
plt.bar(X, R, label="Recovered cases", color="cyan")
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```

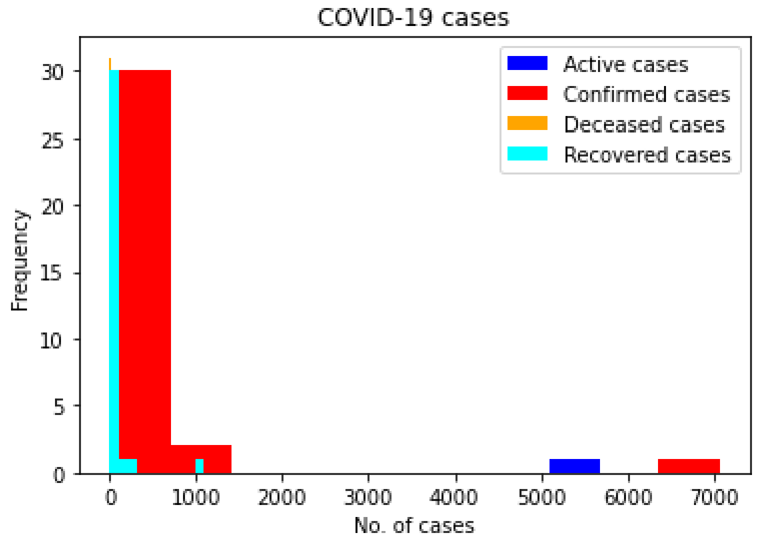


In [163…

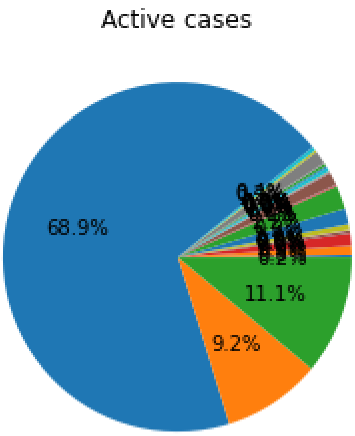
```
plt.bar(X, A, label="Active cases", color="b")
plt.bar(X, C, label="Confirmed cases", color="r")
plt.bar(X, D, label="Deceased cases", color="orange")
plt.bar(X, R, label="Recovered cases", color="cyan")
plt.xlabel('District')
plt.ylabel('No. of cases')
plt.title('COVID-19 cases')
plt.legend()
plt.show()
```



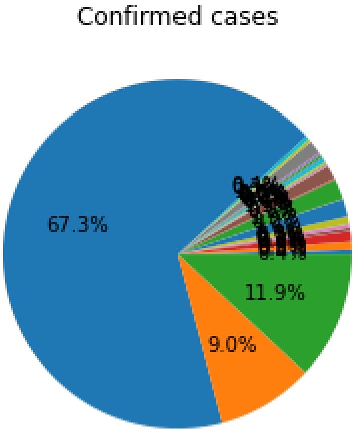
```
In [158…  plt.hist(A, label="Active cases", color="b")
          plt.hist(C, label="Confirmed cases", color="r")
          plt.hist(D, label="Deceased cases", color="orange")
          plt.hist(R, label="Recovered cases", color="cyan")
          plt.xlabel("No. of cases")
          plt.ylabel("Frequency")
          plt.title("COVID-19 cases")
          plt.legend()
          plt.show()
```



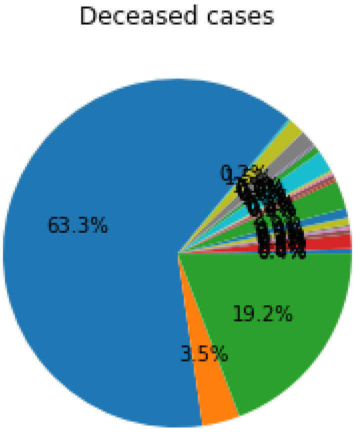
```
In [135…  plt.pie(A, autopct='%1.1f%%')
          plt.title('Active cases')
          plt.show()
```



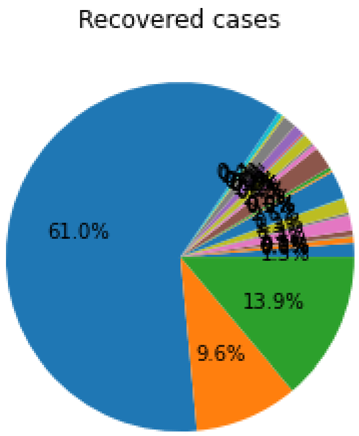
In [171…
```
plt.pie(C, autopct='%1.1f%%')
plt.title('Confirmed cases')
plt.show()
```



In [145…
```
plt.pie(D, autopct='%1.1f%%')
plt.title('Deceased cases')
plt.show()
```

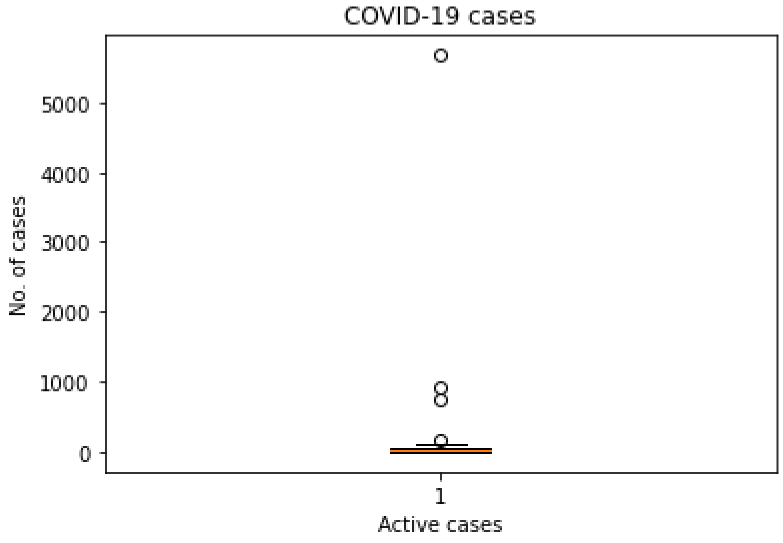


In [146…
```
plt.pie(R, autopct='%1.1f%%')
plt.title('Recovered cases')
plt.show()
```



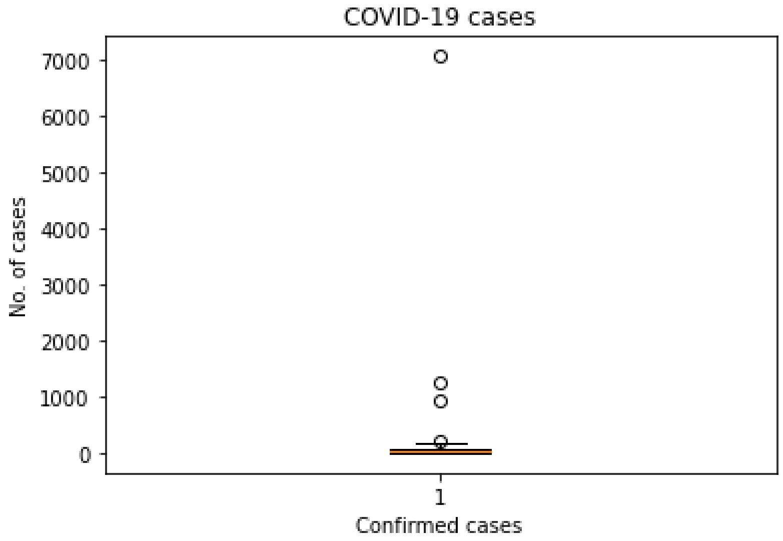
In [427…

```
plt.boxplot(A)
plt.title('COVID-19 cases')
plt.xlabel('Active cases')
plt.ylabel('No. of cases')
plt.show()
```



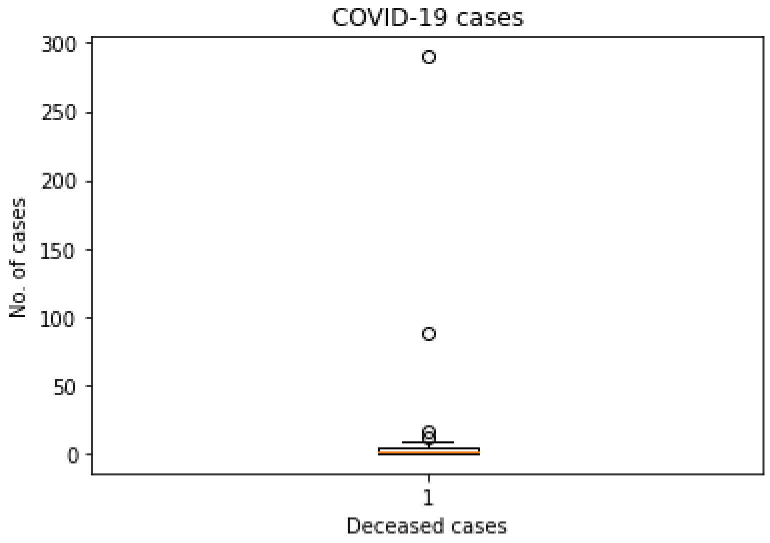
In [428…

```
plt.boxplot(C)
plt.title('COVID-19 cases')
plt.xlabel('Confirmed cases')
plt.ylabel('No. of cases')
plt.show()
```



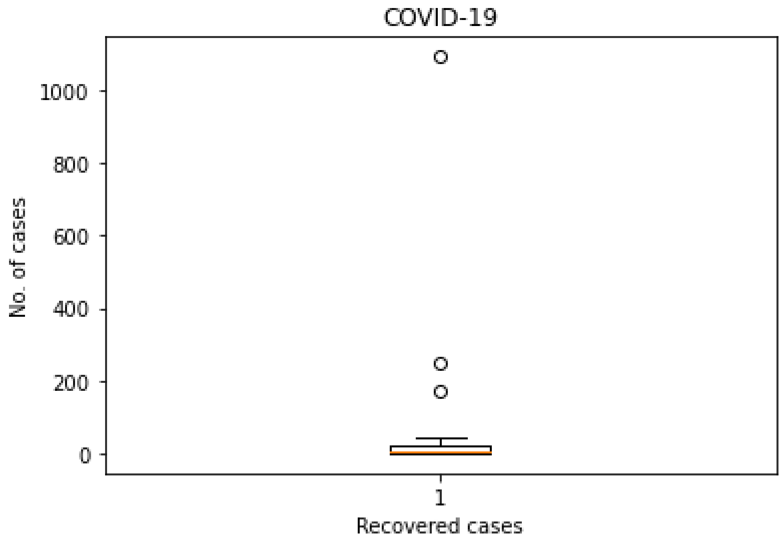
In [429…

```
plt.boxplot(D)
plt.title('COVID-19 cases')
plt.xlabel('Deceased cases')
plt.ylabel('No. of cases')
plt.show()
```



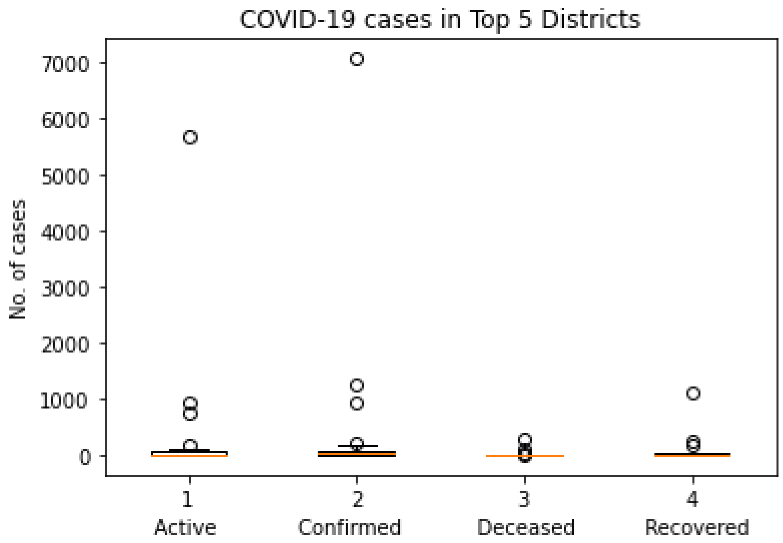
In [430…

```
plt.boxplot(R)
plt.title('COVID-19')
plt.xlabel('Recovered cases')
plt.ylabel('No. of cases')
plt.show()
```



In [475…

```
COVID_cases = [A,C,D,R]
plt.boxplot(COVID_cases)
plt.title('COVID-19 cases')
plt.title('COVID-19 cases in Top 5 Districts')
plt.xlabel('    Active              Confirmed              Deceased              Recovere
plt.ylabel('No. of cases')
plt.show()
```



In [477…

```
# Plot active vs recovered cases for top 5 districts having highest no. of cases.
```

In [94]:

```
data.sort_values(['districtData/0/active', 'districtData/0/district'], ascending =
```

Out[94]:

	districtData/0/district	districtData/0/active	districtData/0/confirmed	districtData/0/deceased
30	Mumbai	5679	7061	290
32	Pune	912	1248	88
31	Thane	755	943	16
12	Nashik	179	197	12
10	Palghar	119	169	4
27	Aurangabad	102	131	7
15	Nagpur	100	139	2
3	Solapur	93	99	6
1	Yavatmal	69	79	0
8	Raigarh	44	71	3
19	Jalgaon	30	40	9
29	Akola	30	39	1
22	Dhule	22	25	3
5	Satara	21	32	2
28	Amravati	17	28	7
0	Ahmadnagar	17	42	2
20	Hingoli	14	15	0
13	Nandurbar	10	11	1
17	Kolhapur	10	14	0
6	Sangli	3	29	1
14	Nanded	3	3	0
16	Latur	3	12	1
18	Buldana	3	21	1
24	Buldana	3	21	1
7	Ratnagiri	2	8	1
2	Washim	1	2	0
4	Sindhudurg	1	2	0
9	Parbhani	1	2	0
26	Bhandara	1	1	0
11	Osmanabad	0	3	0
21	Gondiya	0	1	0
23	Chandrapur	0	2	0
25	Bid	0	1	0

In [15]:

```
Sorted = data.sort_values(['districtData/0/active', 'districtData/0/district'], as
```

In [16]:

```
Sorted.head(5)
```

Out[16]:

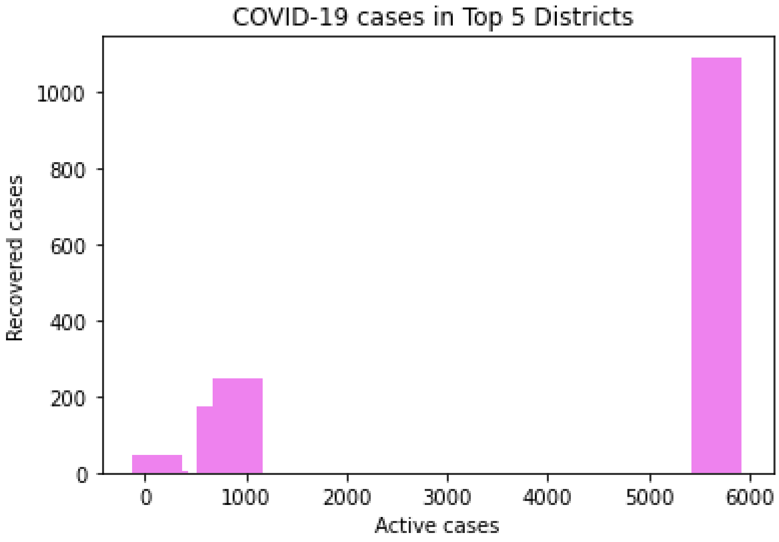
	districtData/0/district	districtData/0/active	districtData/0/confirmed	districtData/0/deceased
30	Mumbai	5679	7061	290
32	Pune	912	1248	88
31	Thane	755	943	16
12	Nashik	179	197	12
10	Palghar	119	169	4

In [228…

```
Sorted_head = Sorted.head(5)
```

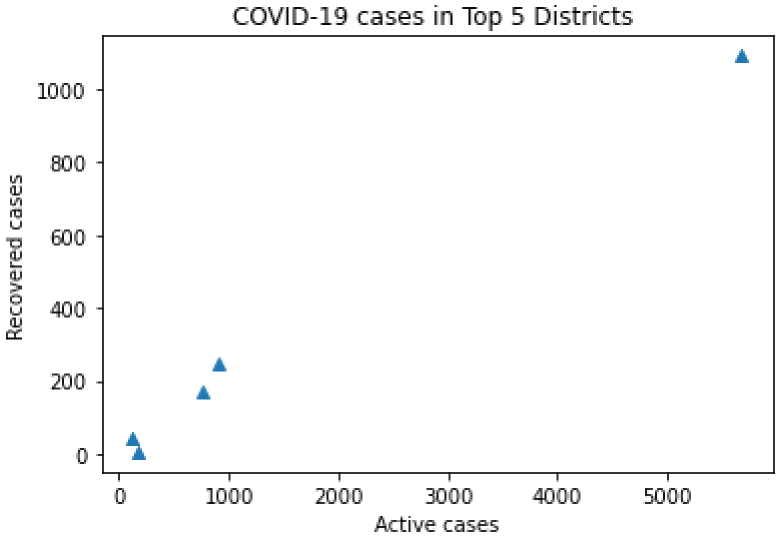
In [325…

```
a = Sorted_head.loc[:,"districtData/0/active"]
r = Sorted_head.loc[:,"districtData/0/recovered"]
plt.bar(x,y, width = 500, color="violet")
plt.xlabel("Active cases")
plt.ylabel("Recovered cases")
plt.title("COVID-19 cases in Top 5 Districts ")
plt.show()
```



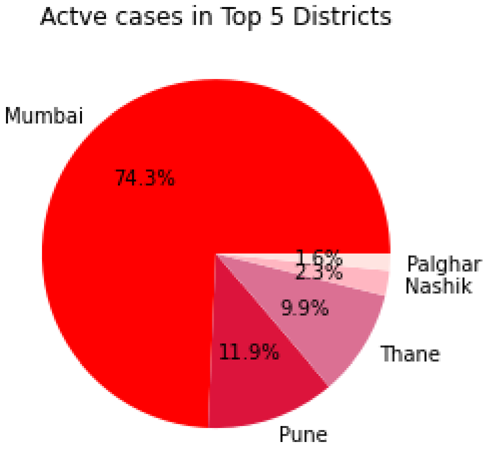
In [326…

```
size = [100,200,100,200,100]
plt.scatter(a,r, marker = '^')
plt.xlabel("Active cases")
plt.ylabel("Recovered cases")
plt.title("COVID-19 cases in Top 5 Districts ")
plt.show()
plt.show()
```



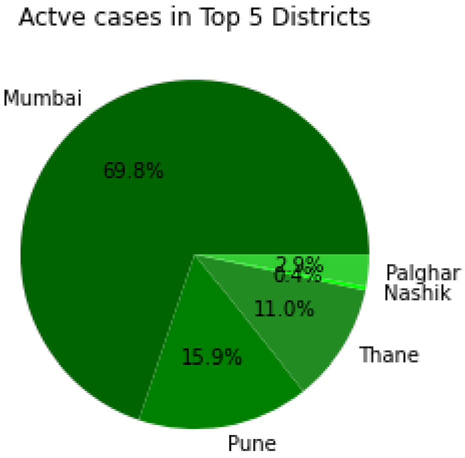
In [380…

```
labels= ['Mumbai', 'Pune', 'Thane', 'Nashik', 'Palghar']
colors = ['red', 'crimson', 'palevioletred', 'lightpink', 'mistyrose']
plt.pie(a, labels=labels, colors=colors, autopct='%1.1f%%')
plt.title('Actve cases in Top 5 Districts')
plt.show()
```



In [382…

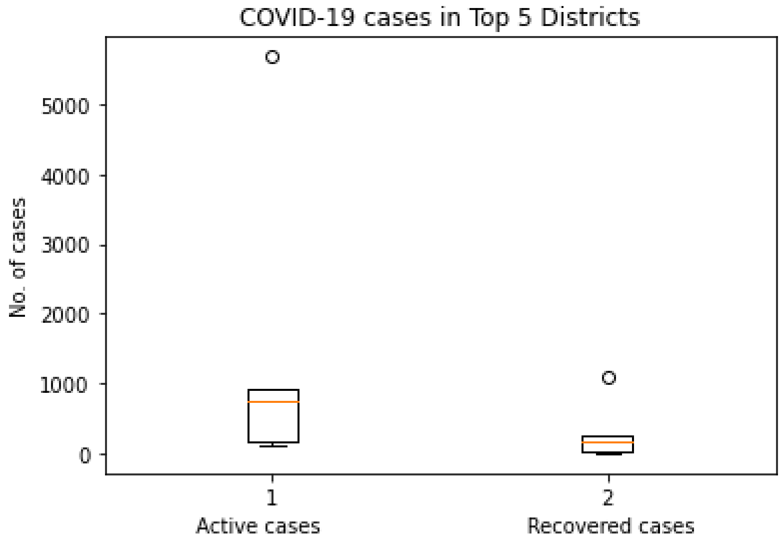
```
labels= ['Mumbai', 'Pune', 'Thane', 'Nashik', 'Palghar']
colors = ['darkgreen', 'green', 'forestgreen', 'lime', 'limegreen']
plt.pie(r, labels=labels, colors=colors, autopct='%1.1f%%')
plt.title('Actve cases in Top 5 Districts')
plt.show()
```



In [421…

```
COVID_cases = [a,r]
```

```
plt.boxplot(COVID_cases)
plt.title('COVID-19 cases in Top 5 Districts')
plt.xlabel('  Active cases                                 Recovered cases')
plt.ylabel('No. of cases')
plt.show()
```


COVID-19 cases in Top 5 Districts
No. of cases
Active cases
Recovered cases