In [1]:
```
# Importing libraraies
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
```

In [2]:
```
crime=pd.read_csv("NCRB-2021_Table_19B.6.csv")
crime
crime_city_wise1=crime[["Sl. No.","City","Persons Arrested - Total","Persons Charge sheeted - To
crime_city_wise=crime_city_wise1.iloc[0:19]
crime_city_wise
#crime.T
```

Out[2]:

| | Sl. No. | City | Persons Arrested - Total | Persons Charge sheeted - Total | Persons Convicted - Total | Persons Discharged - Total | Persons Acquitted - Total |
|---|---|---|---|---|---|---|---|
| **0** | 1 | Ahmedabad (Gujarat) | 34237 | 34237 | 4787 | 13 | 27225 |
| **1** | 2 | Bengaluru (Karnataka) | 24970 | 25509 | 4910 | 151 | 4610 |
| **2** | 3 | Chennai (Tamil Nadu) | 85946 | 19908 | 3653 | 0 | 8929 |
| **3** | 4 | Coimbatore (Tamil Nadu) | 7755 | 7644 | 1416 | 0 | 1859 |
| **4** | 5 | Delhi City | 129925 | 105905 | 38204 | 2274 | 4486 |
| **5** | 6 | Ghaziabad (Uttar Pradesh) | 2840 | 7437 | 327 | 0 | 363 |
| **6** | 7 | Hyderabad (Telangana) | 5056 | 14885 | 3711 | 0 | 1391 |
| **7** | 8 | Indore (Madhya Pradesh) | 6016 | 19616 | 3563 | 0 | 1453 |
| **8** | 9 | Jaipur (Rajasthan) | 15542 | 16068 | 3375 | 573 | 3698 |
| **9** | 10 | Kanpur (Uttar Pradesh) | 10931 | 8323 | 4607 | 0 | 3576 |
| **10** | 11 | Kochi (Kerala) | 6113 | 7054 | 7003 | 0 | 749 |
| **11** | 12 | Kolkata (West Bengal) | 14642 | 13031 | 449 | 0 | 342 |
| **12** | 13 | Kozhikode (Kerala) | 2367 | 5053 | 2989 | 7 | 617 |
| **13** | 14 | Lucknow (Uttar Pradesh) | 5738 | 20761 | 48 | 3 | 0 |
| **14** | 15 | Mumbai (Maharashtra) | 78052 | 62678 | 8904 | 10900 | 4284 |
| **15** | 16 | Nagpur (Maharashtra) | 16001 | 17767 | 3058 | 0 | 3910 |
| **16** | 17 | Patna (Bihar) | 12515 | 8507 | 16 | 0 | 20 |
| **17** | 18 | Pune (Maharashtra) | 13675 | 18716 | 674 | 276 | 1607 |
| **18** | 19 | Surat (Gujarat) | 43605 | 43601 | 936 | 2779 | 2560 |

In [ ]:

```
```

In [3]:

```
# Sorting values
crime_city_wise_sort=crime_city_wise.sort_values(by="Persons Arrested - Total",ascending = False
crime_city_wise_sort
```

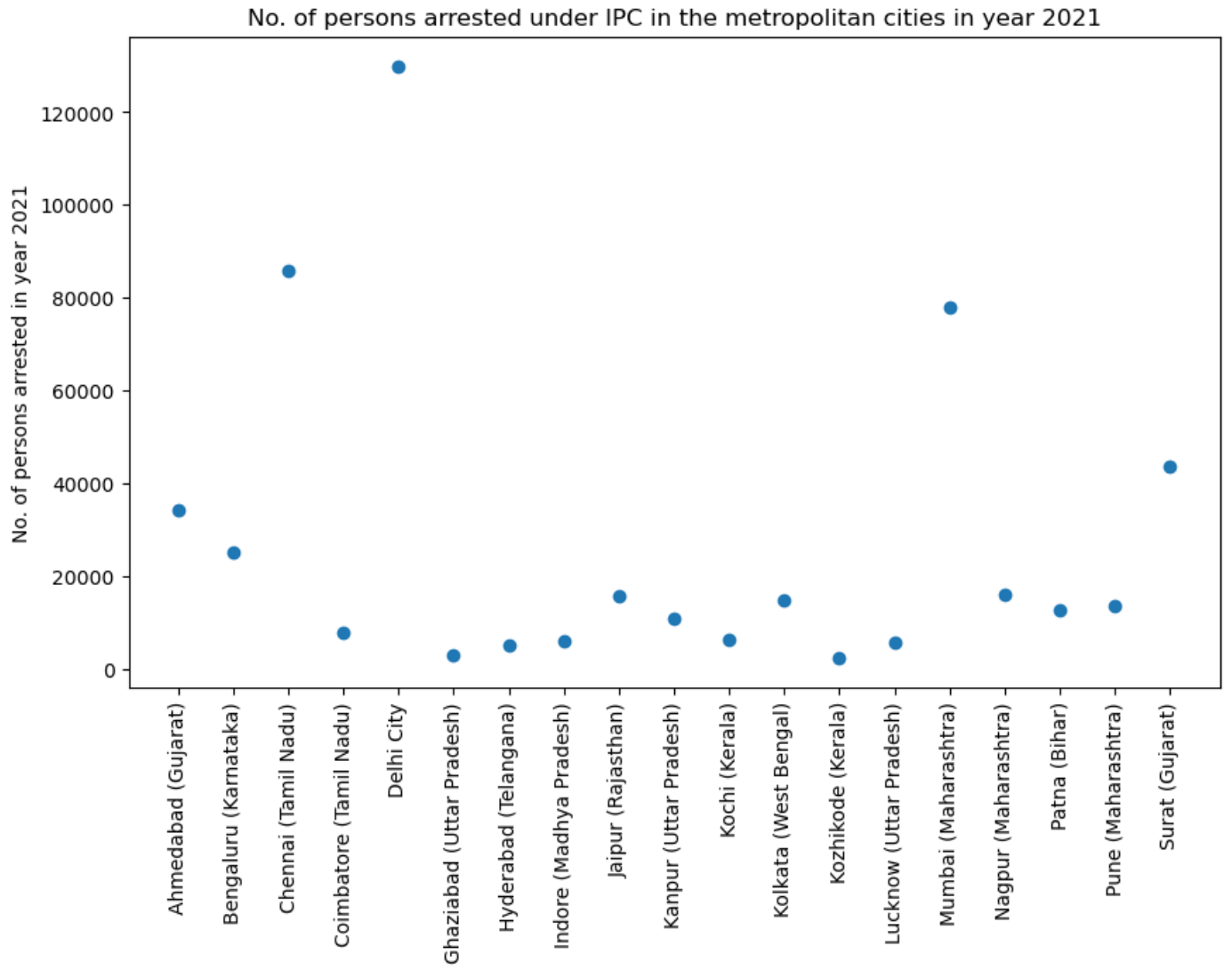
Out[3]:

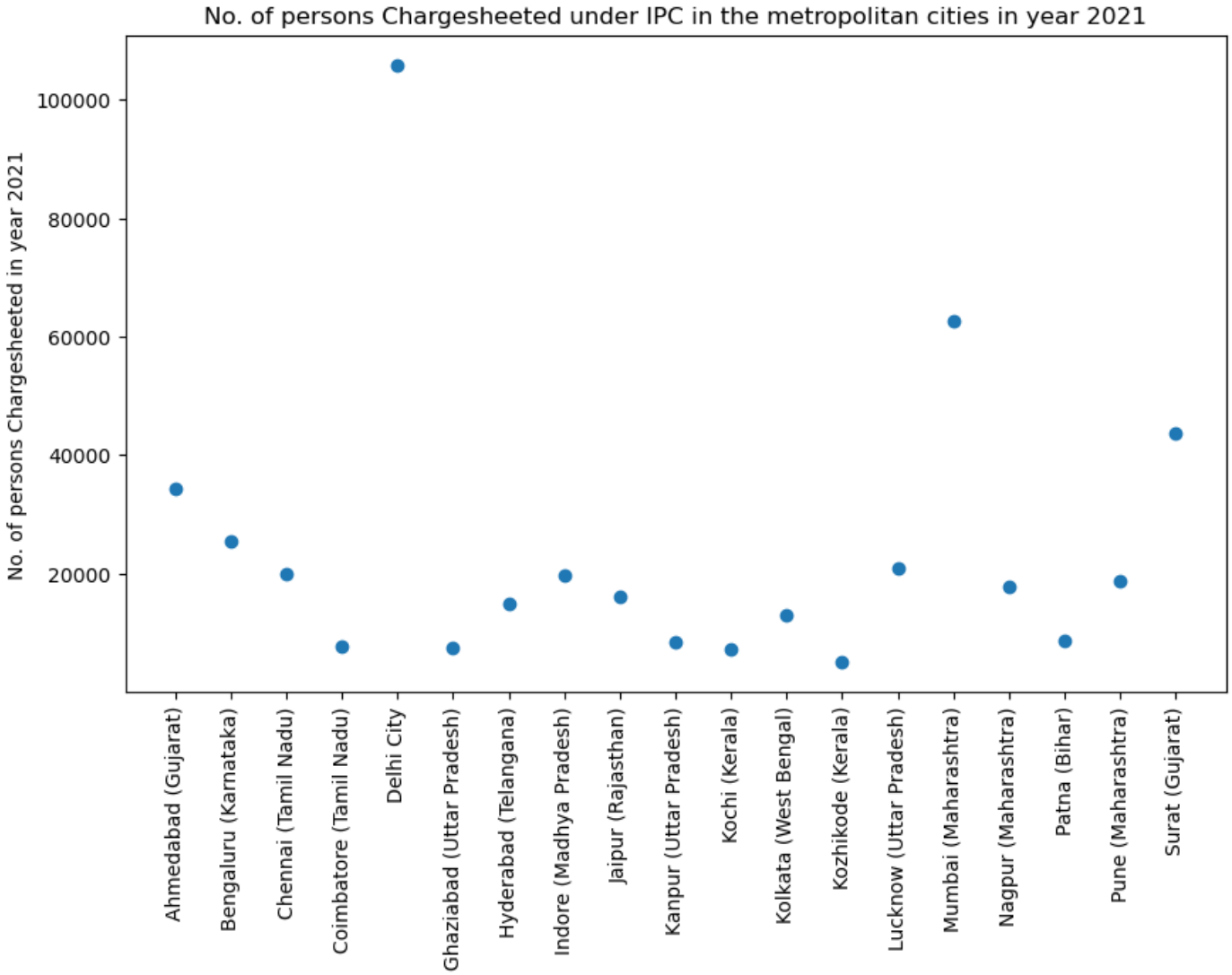
| | Sl. No. | City | Persons Arrested - Total | Persons Charge sheeted - Total | Persons Convicted - Total | Persons Discharged - Total | Persons Acquitted - Total |
|---|---|---|---|---|---|---|---|
| **4** | 5 | Delhi City | 129925 | 105905 | 38204 | 2274 | 4486 |
| **2** | 3 | Chennai (Tamil Nadu) | 85946 | 19908 | 3653 | 0 | 8929 |
| **14** | 15 | Mumbai (Maharashtra) | 78052 | 62678 | 8904 | 10900 | 4284 |
| **18** | 19 | Surat (Gujarat) | 43605 | 43601 | 936 | 2779 | 2560 |
| **0** | 1 | Ahmedabad (Gujarat) | 34237 | 34237 | 4787 | 13 | 27225 |
| **1** | 2 | Bengaluru (Karnataka) | 24970 | 25509 | 4910 | 151 | 4610 |
| **15** | 16 | Nagpur (Maharashtra) | 16001 | 17767 | 3058 | 0 | 3910 |
| **8** | 9 | Jaipur (Rajasthan) | 15542 | 16068 | 3375 | 573 | 3698 |
| **11** | 12 | Kolkata (West Bengal) | 14642 | 13031 | 449 | 0 | 342 |
| **17** | 18 | Pune (Maharashtra) | 13675 | 18716 | 674 | 276 | 1607 |
| **16** | 17 | Patna (Bihar) | 12515 | 8507 | 16 | 0 | 20 |
| **9** | 10 | Kanpur (Uttar Pradesh) | 10931 | 8323 | 4607 | 0 | 3576 |
| **3** | 4 | Coimbatore (Tamil Nadu) | 7755 | 7644 | 1416 | 0 | 1859 |
| **10** | 11 | Kochi (Kerala) | 6113 | 7054 | 7003 | 0 | 749 |
| **7** | 8 | Indore (Madhya Pradesh) | 6016 | 19616 | 3563 | 0 | 1453 |
| **13** | 14 | Lucknow (Uttar Pradesh) | 5738 | 20761 | 48 | 3 | 0 |
| **6** | 7 | Hyderabad (Telangana) | 5056 | 14885 | 3711 | 0 | 1391 |
| **5** | 6 | Ghaziabad (Uttar Pradesh) | 2840 | 7437 | 327 | 0 | 363 |
| **12** | 13 | Kozhikode (Kerala) | 2367 | 5053 | 2989 | 7 | 617 |

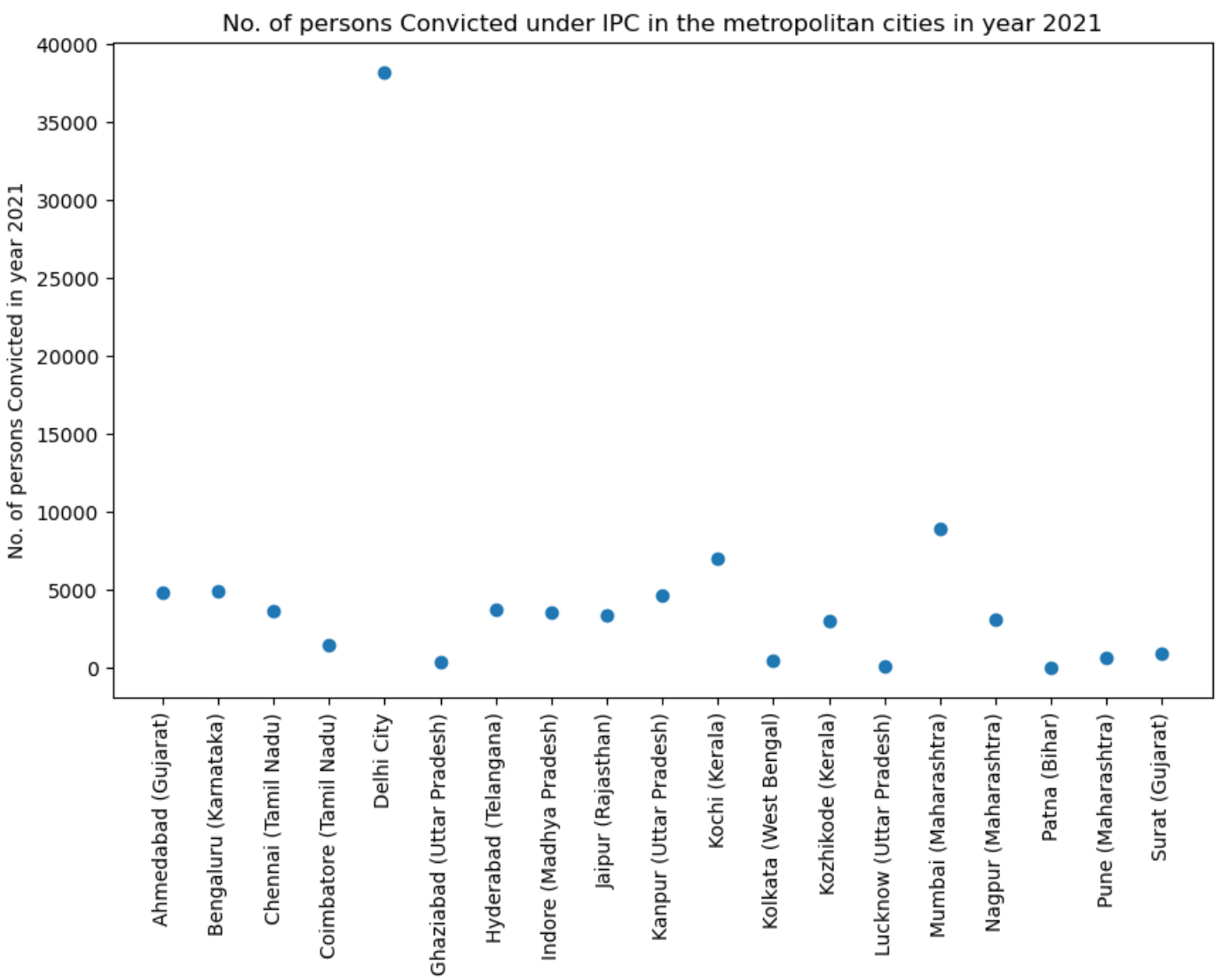
In [4]:

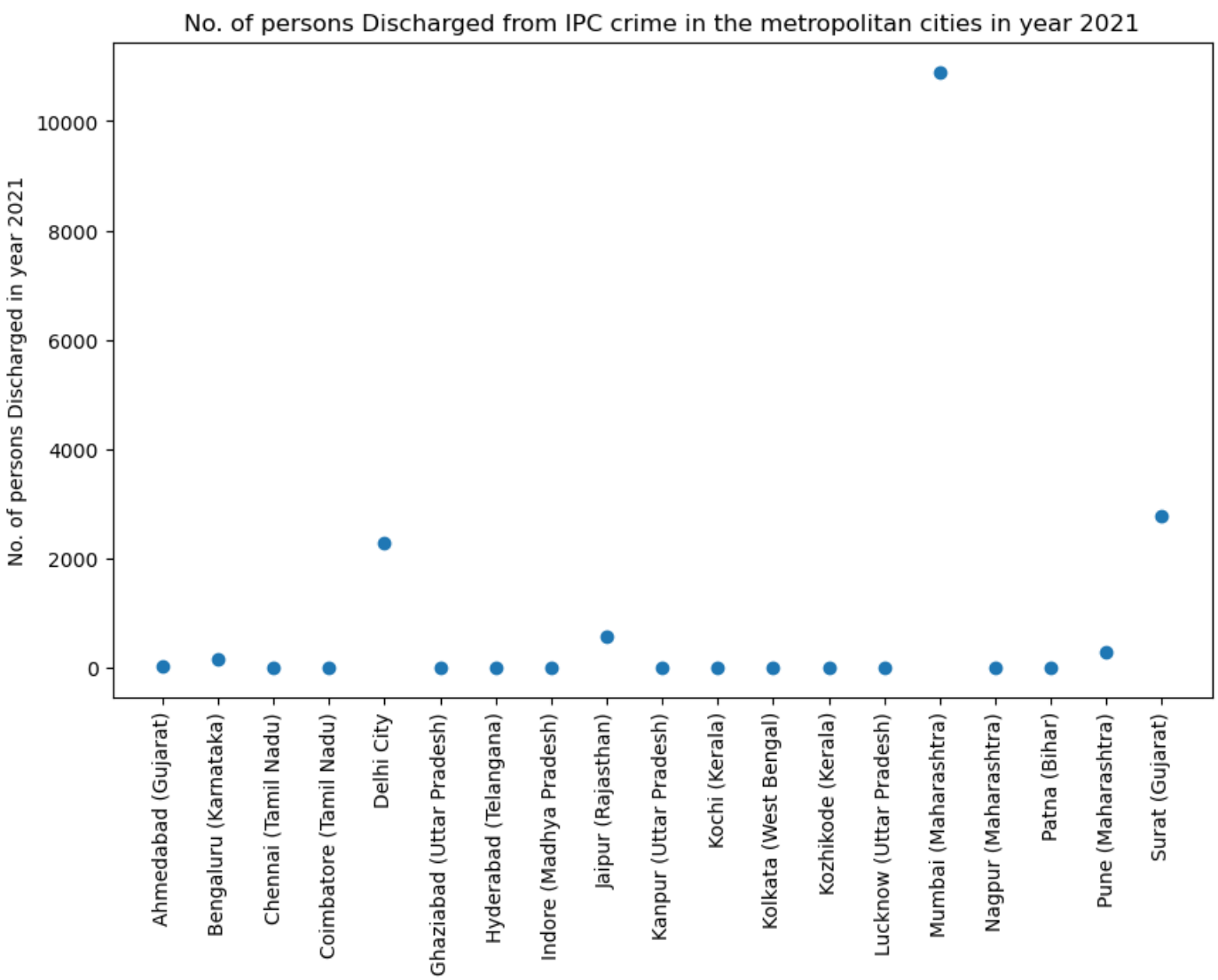
```
# Scatter Plot
plt.figure(figsize=(10,6),dpi=100)
plt.scatter(crime_city_wise["City"],crime_city_wise["Persons Arrested - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons arrested in year 2021')
plt.title('No. of persons arrested under IPC in the metropolitan cities in year 2021')
plt.show()

plt.figure(figsize=(10,6),dpi=100)
plt.scatter(crime_city_wise["City"],crime_city_wise["Persons Charge sheeted - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
```

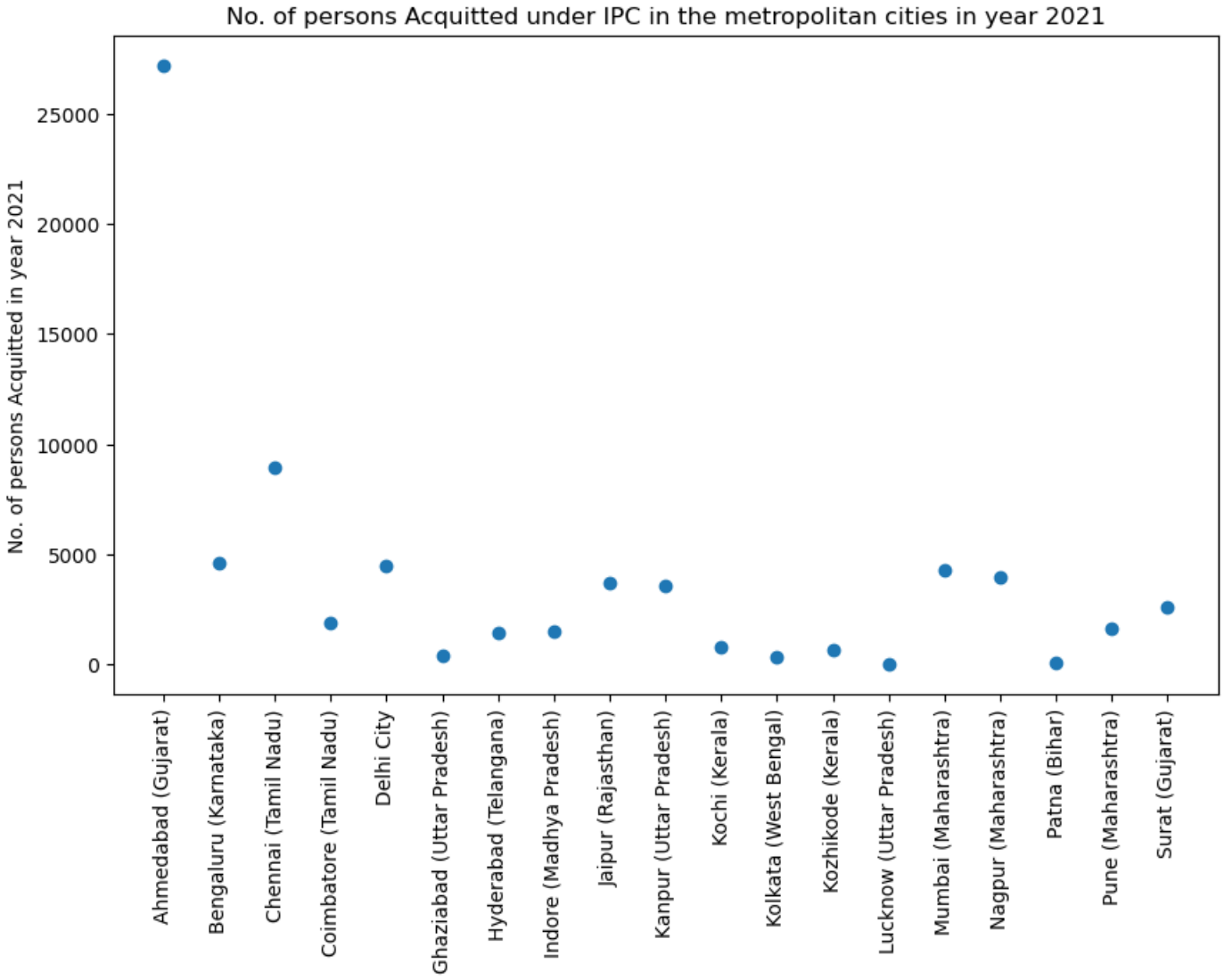
```
plt.ylabel('No. of persons Chargesheeted in year 2021')
plt.title('No. of persons Chargesheeted under IPC in the metropolitan cities in year 2021')
plt.show()

plt.figure(figsize=(10,6),dpi=100)
plt.scatter(crime_city_wise["City"],crime_city_wise["Persons Convicted - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons Convicted in year 2021')
plt.title('No. of persons Convicted under IPC in the metropolitan cities in year 2021')
plt.show()

plt.figure(figsize=(10,6),dpi=100)
plt.scatter(crime_city_wise["City"],crime_city_wise["Persons Discharged - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons Discharged in year 2021')
plt.title('No. of persons Discharged from IPC crime in the metropolitan cities in year 2021')
plt.show()

plt.figure(figsize=(10,6),dpi=100)
plt.scatter(crime_city_wise["City"],crime_city_wise["Persons Acquitted - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons Acquitted in year 2021')
plt.title('No. of persons Acquitted under IPC in the metropolitan cities in year 2021')
plt.show()
```




No. of persons Chargesheeted under IPC in the metropolitan cities in year 2021
No. of persons Chargesheeted in year 2021
100000
80000
60000
40000
20000
Ahmedabad (Gujarat)
Bengaluru (Karnataka)
Chennai (Tamil Nadu)
Coimbatore (Tamil Nadu)
Delhi City
Ghaziabad (Uttar Pradesh)
Hyderabad (Telangana)
Indore (Madhya Pradesh)
Jaipur (Rajasthan)
Kanpur (Uttar Pradesh)
Kochi (Kerala)
Kolkata (West Bengal)
Kozhikode (Kerala)
Lucknow (Uttar Pradesh)
Mumbai (Maharashtra)
Nagpur (Maharashtra)
Patna (Bihar)
Pune (Maharashtra)
Surat (Gujarat)

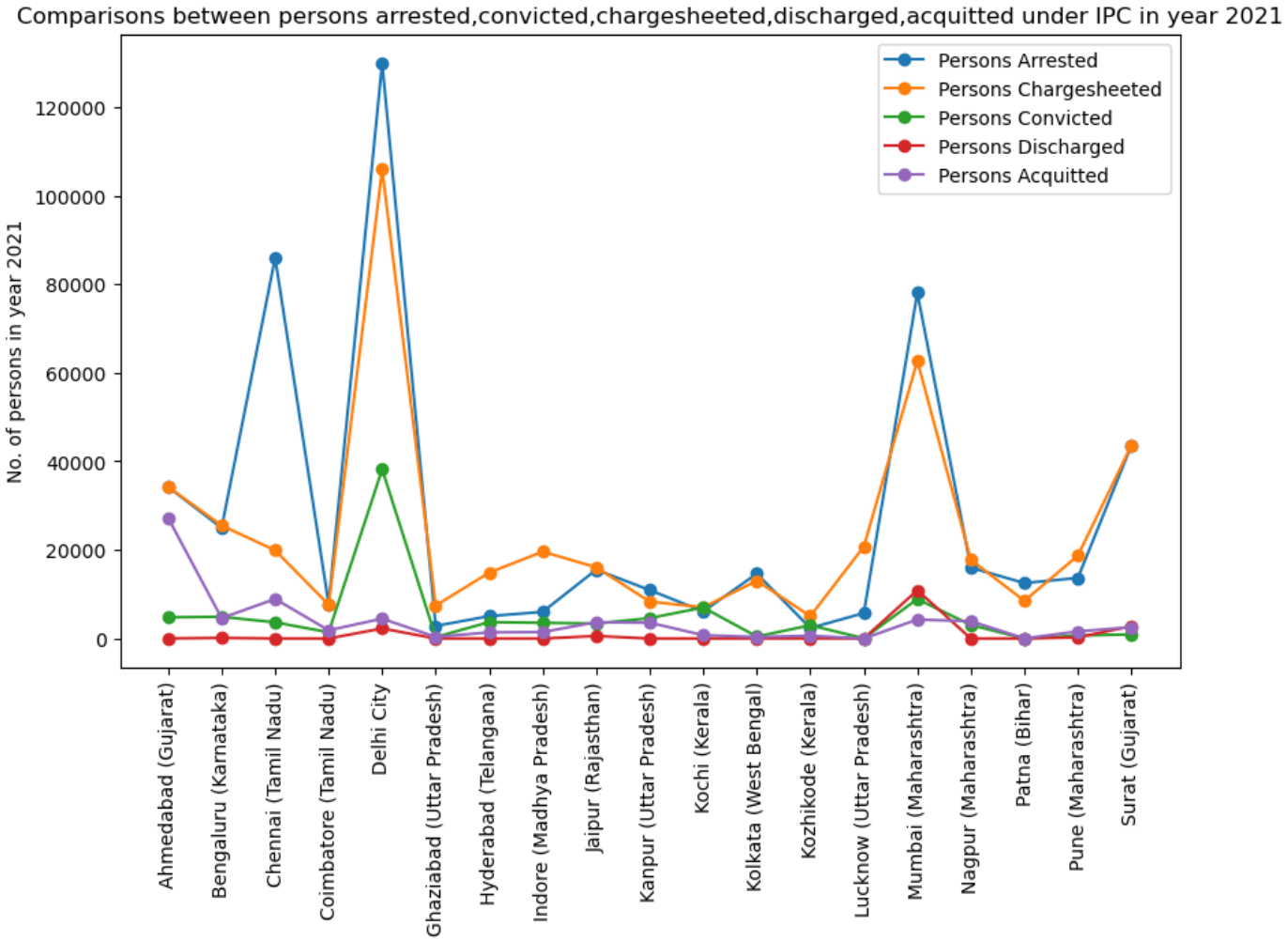

No. of persons Convicted under IPC in the metropolitan cities in year 2021
No. of persons Convicted in year 2021
40000
35000
30000
25000
20000
15000
10000
5000
0
Ahmedabad (Gujarat)
Bengaluru (Karnataka)
Chennai (Tamil Nadu)
Coimbatore (Tamil Nadu)
Delhi City
Ghaziabad (Uttar Pradesh)
Hyderabad (Telangana)
Indore (Madhya Pradesh)
Jaipur (Rajasthan)
Kanpur (Uttar Pradesh)
Kochi (Kerala)
Kolkata (West Bengal)
Kozhikode (Kerala)
Lucknow (Uttar Pradesh)
Mumbai (Maharashtra)
Nagpur (Maharashtra)
Patna (Bihar)
Pune (Maharashtra)
Surat (Gujarat)


No. of persons Discharged from IPC crime in the metropolitan cities in year 2021
No. of persons Discharged in year 2021
0
2000
4000
6000
8000
10000
Ahmedabad (Gujarat)
Bengaluru (Karnataka)
Chennai (Tamil Nadu)
Coimbatore (Tamil Nadu)
Delhi City
Ghaziabad (Uttar Pradesh)
Hyderabad (Telangana)
Indore (Madhya Pradesh)
Jaipur (Rajasthan)
Kanpur (Uttar Pradesh)
Kochi (Kerala)
Kolkata (West Bengal)
Kozhikode (Kerala)
Lucknow (Uttar Pradesh)
Mumbai (Maharashtra)
Nagpur (Maharashtra)
Patna (Bihar)
Pune (Maharashtra)
Surat (Gujarat)



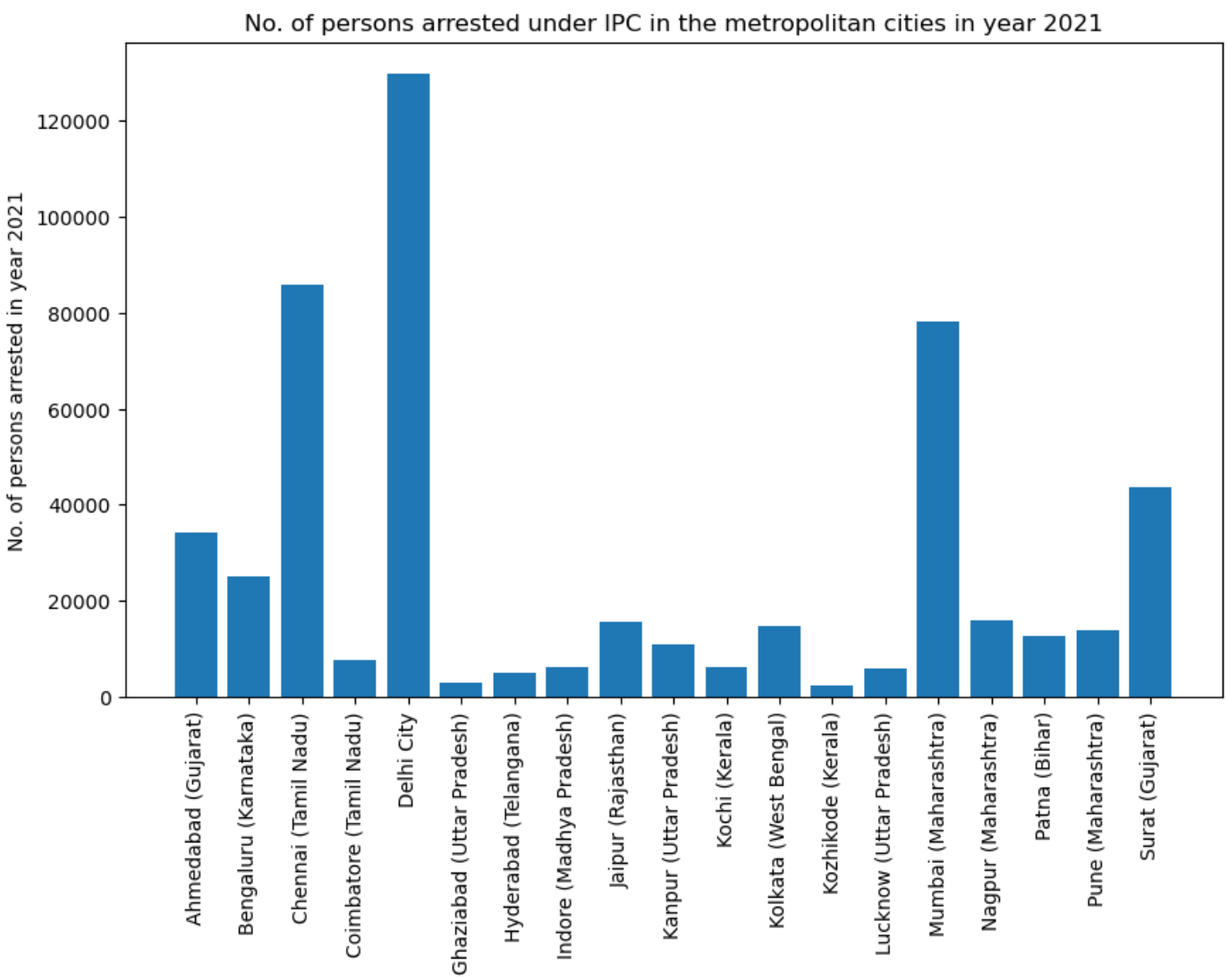

Scatter plots are important because they can show the extent of correlation, if any, between the values of observed quantities or phenomena (called variables). If no correlation exists between the variables, the points appear randomly scattered on the coordinate plane. Here we see that in the city Delhi,Mumbai,Chennai,Surat the large number of people is involved in IPC crime. In the other city the data lies near the x axis so this data is un-correlated.

In [5]:
```
#Line Plot
plt.figure(figsize=(10,6),dpi=100)
plt.plot(crime_city_wise["City"],crime_city_wise["Persons Arrested - Total"],'-o',label="Persons
plt.plot(crime_city_wise["City"],crime_city_wise["Persons Charge sheeted - Total"],'-o',label="P
plt.plot(crime_city_wise["City"],crime_city_wise["Persons Convicted - Total"],'-o',label="Person
plt.plot(crime_city_wise["City"],crime_city_wise["Persons Discharged - Total"],'-o',label="Perso
plt.plot(crime_city_wise["City"],crime_city_wise["Persons Acquitted - Total"],'-o',label="Person
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons in year 2021')
plt.title('Comparisons between persons arrested,convicted,chargesheeted,discharged,acquitted unde
plt.legend()
plt.show()
```



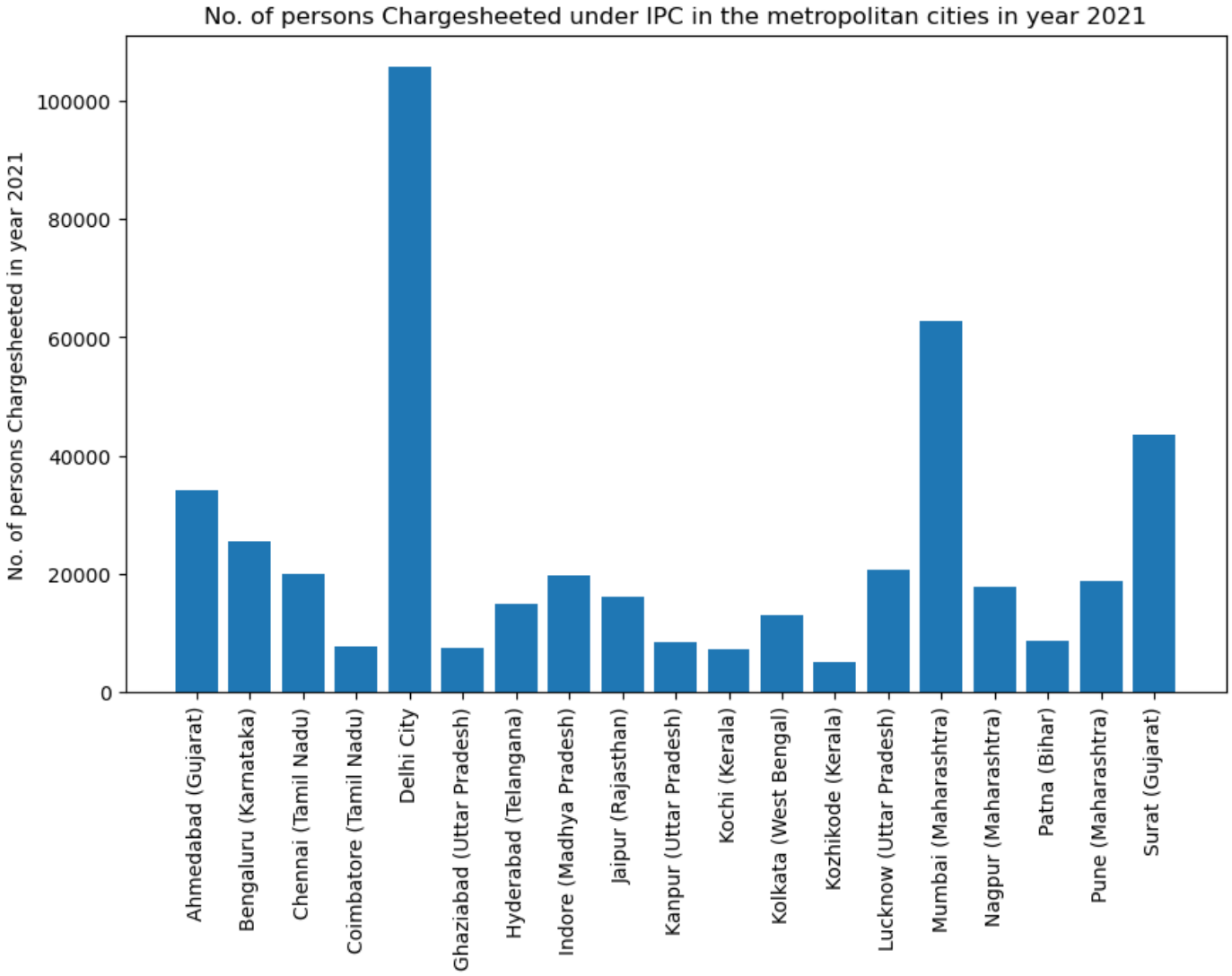

Line plot are used to express a relationship between two variables.Here we also see that largest number of people involved in IPC crime from the city Delhi,Chennai,Surat.The persons Discharged from case is zero in most of the cities.

In [6]:
```
#Bar Plot
plt.figure(figsize=(10,6),dpi=100)
plt.bar(crime_city_wise["City"],crime_city_wise["Persons Arrested - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons arrested in year 2021')
plt.title('No. of persons arrested under IPC in the metropolitan cities in year 2021')
plt.show()
```



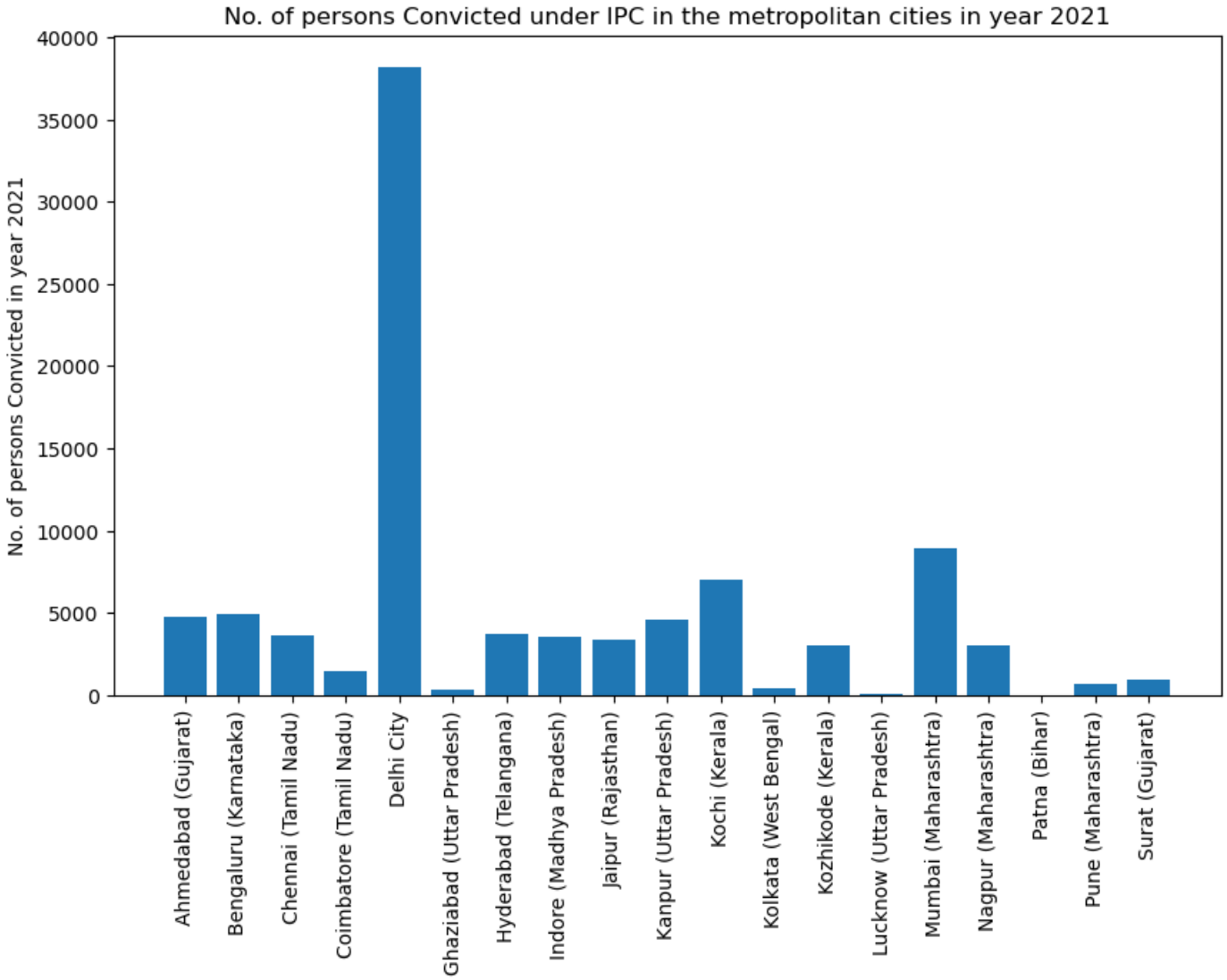

The bar graph helps to compare the different sets of data among different groups easily. As seen from the bar plot, number of the people arrested(under IPC) in the Delhi city is highest city.Mumbai and Chennai are close to each other in the number of people arrested.The lowest number of people is arrested in Kozhikode (Kerala) .

In [7]:
```
plt.figure(figsize=(10,6),dpi=100)
plt.bar(crime_city_wise["City"],crime_city_wise["Persons Charge sheeted - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons Chargesheeted in year 2021')
plt.title('No. of persons Chargesheeted under IPC in the metropolitan cities in year 2021')
plt.show()
```



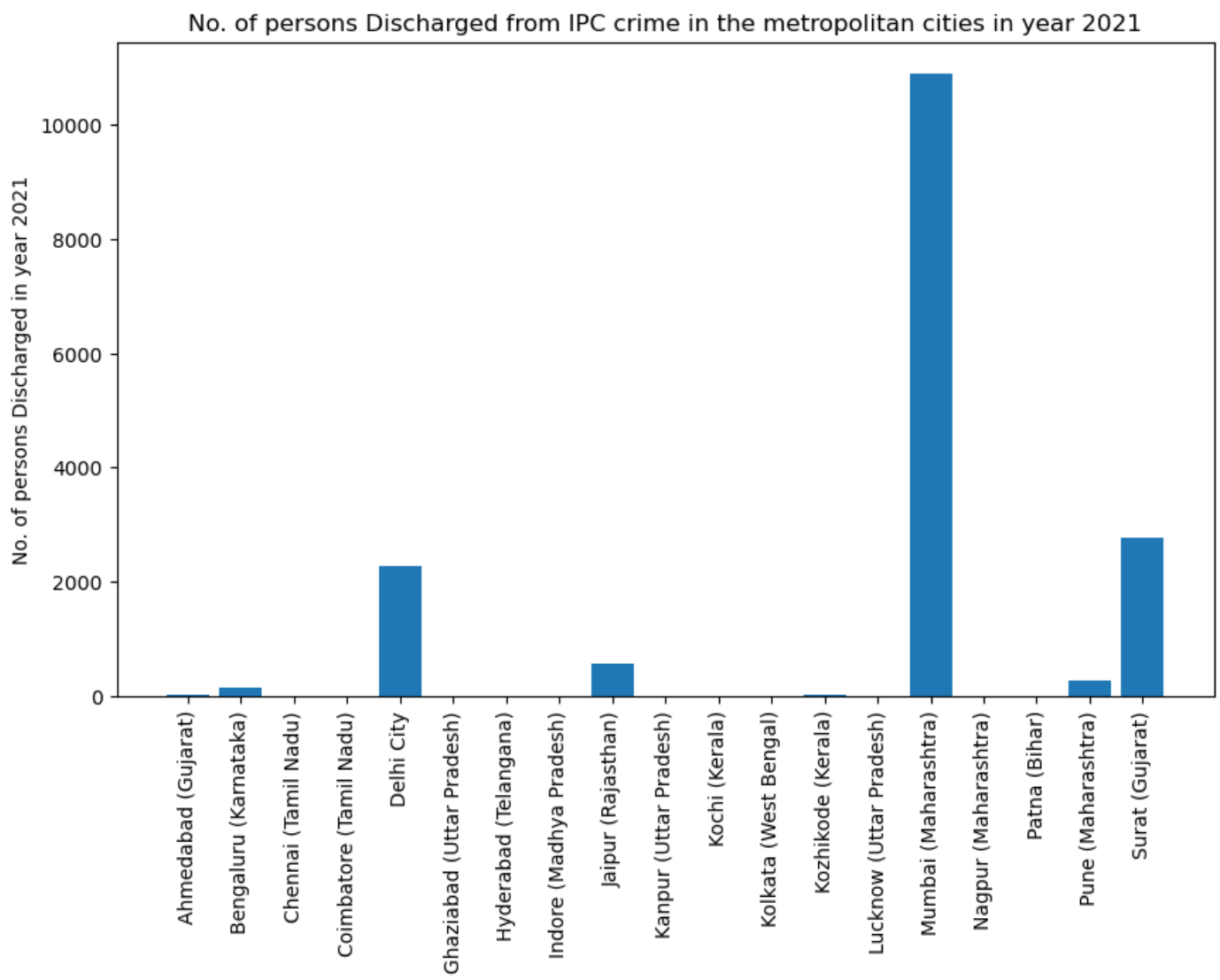

Here also Delhi city has the highest number of persons Chargesheeted and some large amount of people in mumbai and suratalso chargesheeted under IPC, while the lowest is for Kozhikode (Kerala).

In [8]:
```
plt.figure(figsize=(10,6),dpi=100)
plt.bar(crime_city_wise["City"],crime_city_wise["Persons Convicted - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons Convicted in year 2021')
plt.title('No. of persons Convicted under IPC in the metropolitan cities in year 2021')
plt.show()
```

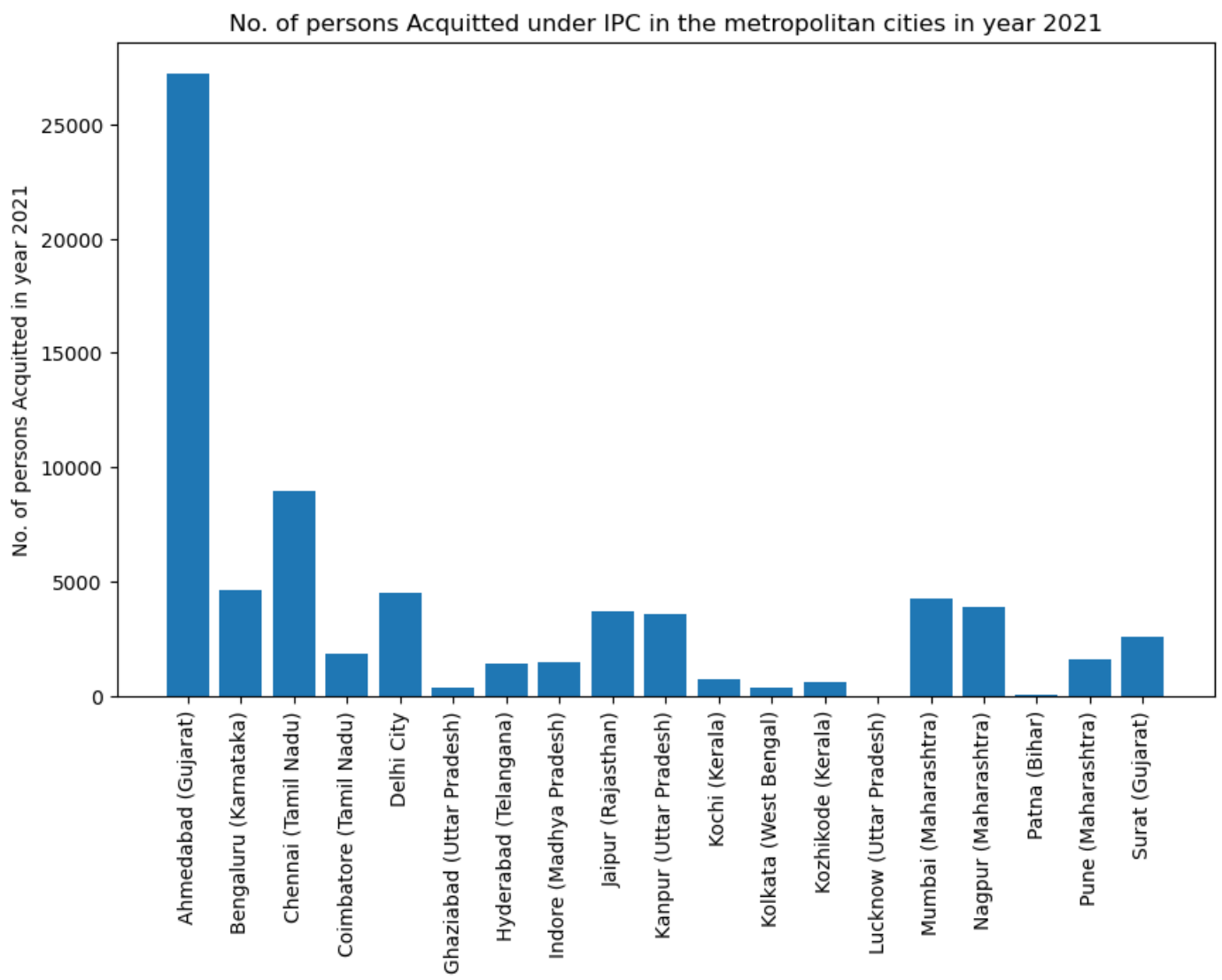



From the plot we see,number of convicted persons under IPC in the delhi city is largest and has approximately 7 times than the next city.The lowest number of people convicted in Patna(16 people) and in Lucknow(48 People).

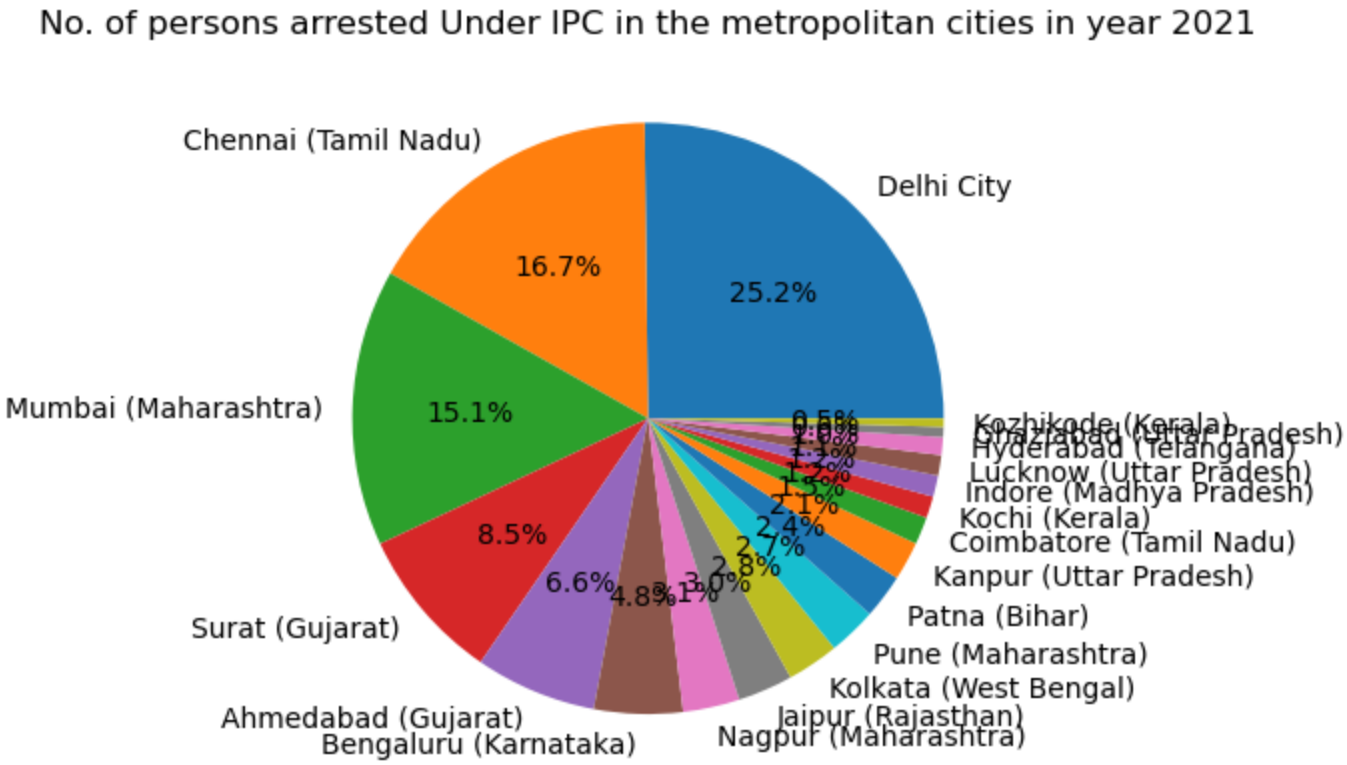
In [9]:
```
plt.figure(figsize=(10,6),dpi=100)
plt.bar(crime_city_wise["City"],crime_city_wise["Persons Discharged - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons Discharged in year 2021')
plt.title('No. of persons Discharged from IPC crime in the metropolitan cities in year 2021')
plt.show()
```




In [10]:
```
plt.figure(figsize=(10,6),dpi=100)
plt.bar(crime_city_wise["City"],crime_city_wise["Persons Acquitted - Total"])
plt.xticks(crime_city_wise["City"],rotation=90)
plt.ylabel('No. of persons Acquitted in year 2021')
plt.title('No. of persons Acquitted under IPC in the metropolitan cities in year 2021')
plt.show()
```




In [11]:
```
#Pie Chart
plt.pie(crime_city_wise_sort["Persons Arrested - Total"], labels=crime_city_wise_sort["City"], au
plt.title('No. of persons arrested Under IPC in the metropolitan cities in year 2021')
plt.show()
```




The overall persons arrested irrespective of their age duirng the year 2021, is highest for Delhi city and lowest for Kozhikode.

As seen from pie chart, the contribution of total number of arrested person in the year 2021 is 25.2% (highest) for the Delhi city. In fact the top three metropolitan cities (Delhi-25.2%, Chennai-16.7%, and Mumbai-15.1%) contributes to a total of 57% of arrested persons. The rest 16 of 19 cities contributes only 43% of the total.

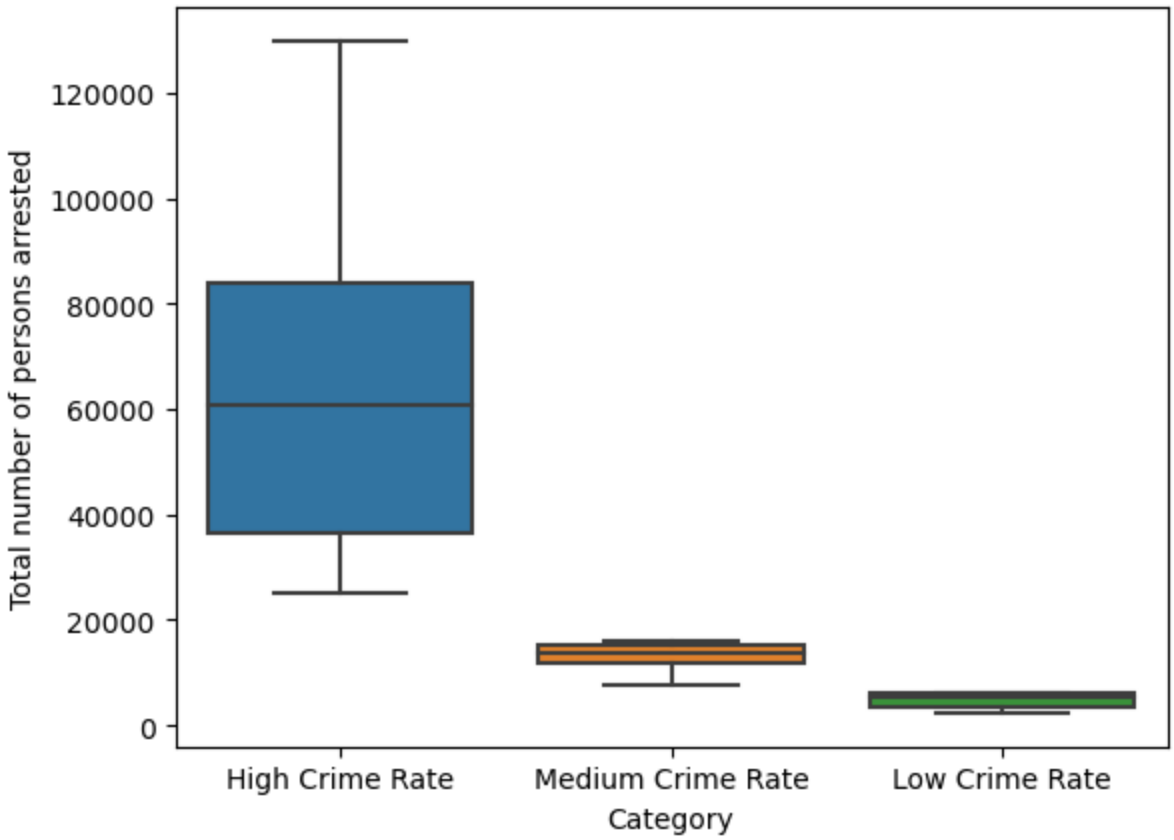
In [12]:
```
#Box Plot

#  Adding new row to dataframe
x=[]            # defining new column
for row in crime_city_wise_sort["Persons Arrested - Total"]:
    if row>24500:
        x.append('High Crime Rate')
    elif row<24500 and row>6200:
        x.append('Medium Crime Rate')
    else:
        x.append('Low Crime Rate')
crime_city_wise_sort['Category']=x
crime_city_wise_sort
```

Out[12]:

| | Sl. No. | City | Persons Arrested - Total | Persons Charge sheeted - Total | Persons Convicted - Total | Persons Discharged - Total | Persons Acquitted - Total | Category |
|---|---|---|---|---|---|---|---|---|
| **4** | 5 | Delhi City | 129925 | 105905 | 38204 | 2274 | 4486 | High Crime Rate |
| **2** | 3 | Chennai (Tamil Nadu) | 85946 | 19908 | 3653 | 0 | 8929 | High Crime Rate |
| **14** | 15 | Mumbai (Maharashtra) | 78052 | 62678 | 8904 | 10900 | 4284 | High Crime Rate |
| **18** | 19 | Surat (Gujarat) | 43605 | 43601 | 936 | 2779 | 2560 | High Crime Rate |
| **0** | 1 | Ahmedabad (Gujarat) | 34237 | 34237 | 4787 | 13 | 27225 | High Crime Rate |
| **1** | 2 | Bengaluru (Karnataka) | 24970 | 25509 | 4910 | 151 | 4610 | High Crime Rate |
| **15** | 16 | Nagpur (Maharashtra) | 16001 | 17767 | 3058 | 0 | 3910 | Medium Crime Rate |
| **8** | 9 | Jaipur (Rajasthan) | 15542 | 16068 | 3375 | 573 | 3698 | Medium Crime Rate |
| **11** | 12 | Kolkata (West Bengal) | 14642 | 13031 | 449 | 0 | 342 | Medium Crime Rate |
| **17** | 18 | Pune (Maharashtra) | 13675 | 18716 | 674 | 276 | 1607 | Medium Crime Rate |
| **16** | 17 | Patna (Bihar) | 12515 | 8507 | 16 | 0 | 20 | Medium Crime Rate |
| **9** | 10 | Kanpur (Uttar Pradesh) | 10931 | 8323 | 4607 | 0 | 3576 | Medium Crime Rate |
| **3** | 4 | Coimbatore (Tamil Nadu) | 7755 | 7644 | 1416 | 0 | 1859 | Medium Crime Rate |
| **10** | 11 | Kochi (Kerala) | 6113 | 7054 | 7003 | 0 | 749 | Low Crime Rate |
| **7** | 8 | Indore (Madhya Pradesh) | 6016 | 19616 | 3563 | 0 | 1453 | Low Crime Rate |
| **13** | 14 | Lucknow (Uttar Pradesh) | 5738 | 20761 | 48 | 3 | 0 | Low Crime Rate |
| **6** | 7 | Hyderabad (Telangana) | 5056 | 14885 | 3711 | 0 | 1391 | Low Crime Rate |

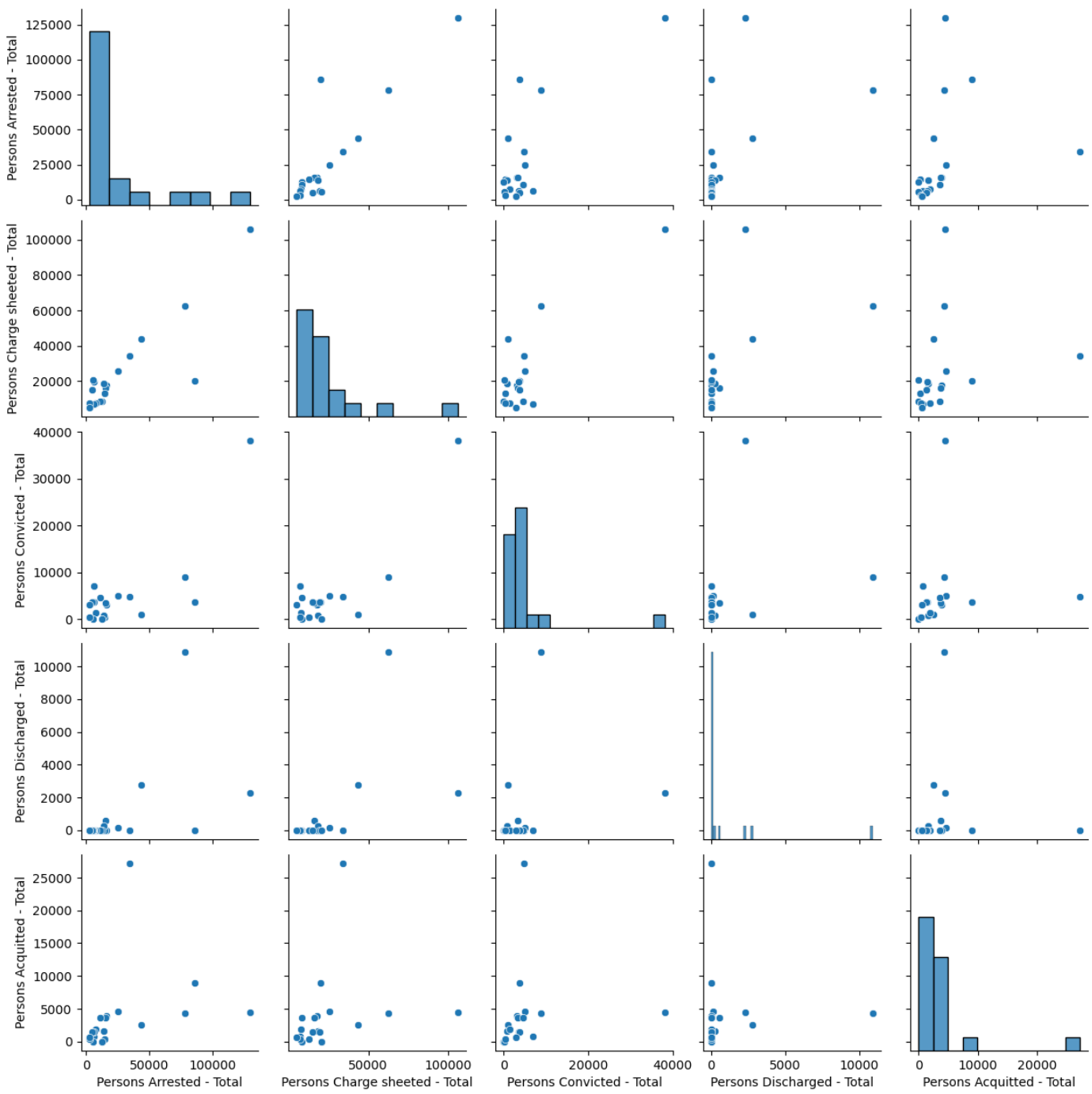
| | Sl. No. | City | Persons Arrested - Total | Persons Charge sheeted - Total | Persons Convicted - Total | Persons Discharged - Total | Persons Acquitted - Total | Category |
|---|---|---|---|---|---|---|---|---|
| **5** | 6 | Ghaziabad (Uttar Pradesh) | 2840 | 7437 | 327 | 0 | 363 | Low Crime Rate |
| **12** | 13 | Kozhikode (Kerala) | 2367 | 5053 | 2989 | 7 | 617 | Low Crime Rate |

In [13]:
```
sns.boxplot(data=crime_city_wise_sort, x="Category", y="Persons Arrested - Total")
plt.ylabel('Total number of persons arrested')
plt.show()
```




It can be seen from above box plot that "High Crime Rate" category has a very large range. While the "Medium Crime Rate" and "Low Crime Rate" categories are scattered very close to the median.

In [14]:
```
#Pair plot
sns.pairplot(crime_city_wise_sort)
plt.show()
```




Pairplot is very usefull for performing EDA and identifying corelation between the parameters. As seen from the pair plots, there is a weak correlation between arrested persons and person person discharged,acquitted,But there is strong correlation with person charsheeted.The correlation is almost linear.

We also see that the person chargesheeted is correlated with person convicted but uncorrelated with person arrested,person discharged. In the person discharged ,if we removed 2-3 outliers then it has strong correlation with person arrested,chargesheeted,convicted,otherwise it is uncorrelated. Also see that person acquitted is uncorrelated with every person . From this data we know that when there is large nubmer of people arrested,then there is higher probability of existence of people with chargesheeted .Also see that number of people discharged from the case (except 2-3 city)is linear proportional to the person arrested,chargesheeted,convicted.