Phase 3- SQL operations

In [13]:
```
import sqlite3
import csv
import pandas as pd
```

In [3]:
```
db=sqlite3.connect("crime_database")
```

In [4]:
```
cursor=db.cursor()
```

In [5]:
```
cursor.execute("CREATE TABLE crime_against_women (state UT TEXT,DISTRICT TEXT,Y
db.commit()
```

3.1 Insert records from 42_District_wise_crimes_committed_against_women_2001_2012.csv into a table

In [ ]:
```
with open('42_District_wise_crimes_committed_against_women_2001_2012.csv','r')
    #here r is read as
    no_records=0
    for row in file:
        cursor.execute("INSERT INTO crimes_against_women VALUES(?,?,?,?,?,?,?,?
        db.commit()
        no_records += 1
```

In [14]:
```
print(no_records,'Records Inserted')
```

```
9017 Records Inserted
```

3.2 Write SQL query to find the highest number of rapes & Kidnappings that happened in which state, District, and year

In [119]:
```
result=cursor.execute("SELECT  state UT, DISTRICT ,Year ,MAX(Rape), MAX(Kindanp
for row in result:
    print(row)
```

```
('WEST BENGAL', 'MURSHIDABAD', 2011, 433, 492)
('WEST BENGAL', 'MURSHIDABAD', 2012, 257, 464)
('WEST BENGAL', 'MURSHIDABAD', 2010, 526, 441)
('DELHI', 'NORTH-WEST', 2005, 236, 349)
('WEST BENGAL', 'MURSHIDABAD', 2009, 568, 342)
('UTTAR PRADESH', 'LUCKNOW', 2008, 334, 331)
('DELHI', 'NORTH WEST', 2001, 145, 298)
('DELHI', 'NORTH-WEST', 2006, 224, 287)
('DELHI', 'NORTH WEST', 2002, 159, 259)
('DELHI', 'OUTER', 2007, 369, 237)
('ASSAM', 'GUWAHATI CITY', 2004, 197, 228)
('DELHI', 'NORTH WEST', 2003, 137, 191)
```

The state, district and year with district with Highest Rape and Kidnapping i.e MURSHIDABAD District in WEST BENGAL in year 2011 with 568 rape and 492 Kidnapping cases

3.3 Write SQL query to find All the lowest number of rapes & Kidnappings that happened in which state, District, and year

In [122]:
```
result=cursor.execute("SELECT  state UT, DISTRICT ,Year ,MIN(Rape), MIN(Kindanp
for row in result:
    print(row)
```

```
('A & N ISLANDS', 'NICOBAR', 2001, 0, 0)
('ANDHRA PRADESH', 'GUNTAKAL RLY.', 2001, 0, 0)
('ARUNACHAL PRADESH', 'TAWANG', 2001, 0, 0)
('ASSAM', 'C.I.D.', 2001, 0, 0)
('BIHAR', 'ARWAL', 2001, 0, 0)
('CHHATTISGARH', 'BIZAPUR', 2001, 0, 0)
('DAMAN & DIU', 'DIU', 2001, 0, 0)
('DELHI', 'S.T.F.', 2001, 0, 0)
('GUJARAT', 'W.RLY', 2001, 0, 0)
('HARYANA', 'GRP', 2002, 0, 0)
('HIMACHAL PRADESH', 'G.R.P.', 2001, 0, 0)
('JAMMU & KASHMIR', 'CRIME JAMMU', 2001, 0, 0)
('JHARKHAND', 'DEOGHAR', 2001, 0, 0)
('KARNATAKA', 'CHITRADURGA', 2001, 0, 0)
('KERALA', 'CBCID', 2001, 0, 0)
('LAKSHADWEEP', 'LAKSHADWEEP', 2001, 0, 0)
('MADHYA PRADESH', 'INDORE RLY.', 2001, 0, 0)
('MAHARASHTRA', 'MUMBAI RLY.', 2001, 0, 0)
('MANIPUR', 'CHURACHANDPUR', 2001, 0, 0)
('MEGHALAYA', 'GARO HILLS EAST', 2001, 0, 0)
('MIZORAM', 'AIZAWL', 2001, 0, 0)
('NAGALAND', 'KIPHIRE', 2001, 0, 0)
('ODISHA', 'RAYAGADA', 2001, 0, 0)
('PUDUCHERRY', 'KARAIKAL', 2011, 0, 0)
('PUNJAB', 'G.R.P', 2001, 0, 0)
('RAJASTHAN', 'G.R.P.', 2003, 0, 0)
('SIKKIM', 'NORTH', 2001, 0, 0)
('TAMIL NADU', 'CHENNAI', 2001, 0, 0)
('TRIPURA', 'DHALAI', 2001, 0, 0)
('UTTAR PRADESH', 'MAHARAJGANJ', 2002, 0, 0)
('UTTARAKHAND', 'NAINITAL', 2001, 0, 0)
('WEST BENGAL', 'ASANSOL', 2001, 0, 0)
('GOA', 'NORTH GOA', 2001, 3, 1)
('D & N HAVELI', 'D and N HAVELI', 2001, 1, 2)
('CHANDIGARH', 'CHANDIGARH', 2003, 18, 28)
```

All the above result showing Minimum of Rape And Kldnapping in each state And year.

3.4 Insert records from 02_District_wise_crimes_committed_against_ST_2001_2012.csv into a new table

In [ ]:
```
cursor.execute("CREATE TABLE crime_against_st (state UT TEXT,DISTRICT TEXT,Year
db.commit()
```

In [14]:
```
with open('02_District_wise_crimes_committed_against_ST_2001_2012.csv','r') as
    #here r is read as
    no_records=0
    for row in file:
        cursor.execute("INSERT INTO crime_against_st VALUES(?,?,?,?,?,?,?,?,?,?
        db.commit()
        no_records += 1
```

In [15]:
```
print(no_records,'Records Inserted')
```

```
9018 Records Inserted
```

3.5 Write SQL query to find the highest number of dacoity/robbery in which district.

In [21]:
```
result=cursor.execute("SELECT State UT,DISTRICT ,Year, MAX(Dacoity), MAX(Robber
for row in result:
    print(row)
```

```
('GUJARAT', 'DAHOD', 2001, 29, 32)
```

Dahod District of Gujarat has highest number of Daocity and Robbery in Year 2001

3.6 Write SQL query to find in which districts(All) the lowest number of murders happened

In [130]:
```
result=cursor.execute("SELECT State UT,DISTRICT, Murder FROM crime_against_st W
for row in result:
    print(row)
```

```
('A & N ISLANDS', 'ANDAMAN', 0)
('ANDHRA PRADESH', 'ADILABAD', 0)
('ARUNACHAL PRADESH', 'CHANGLANG', 0)
('ASSAM', 'BARPETA', 0)
('BIHAR', 'ARWAL', 0)
('CHANDIGARH', 'CHANDIGARH', 0)
('CHHATTISGARH', 'BALRAMPUR', 0)
('D & N HAVELI', 'D and N HAVELI', 0)
('DAMAN & DIU', 'DAMAN', 0)
('DELHI', 'CENTRAL', 0)
('GOA', 'NORTH GOA', 0)
('GUJARAT', 'AHMEDABAD COMMR.', 0)
('HARYANA', 'AMBALA', 0)
('HIMACHAL PRADESH', 'BILASPUR', 0)
('JAMMU & KASHMIR', 'ANANTNAG', 0)
('JHARKHAND', 'BOKARO', 0)
('KARNATAKA', 'BAGALKOT', 0)
('KERALA', 'ALAPUZHA', 0)
('LAKSHADWEEP', 'LAKSHADWEEP', 0)
('MADHYA PRADESH', 'BETUL', 0)
('MAHARASHTRA', 'AKOLA', 0)
('MANIPUR', 'BISHNUPUR', 0)
('MEGHALAYA', 'GARO HILLS EAST', 0)
('MIZORAM', 'AIZAWL', 0)
('NAGALAND', 'DIMAPUR', 0)
('ODISHA', 'ANGUL', 0)
('PUDUCHERRY', 'PONDICHERRY', 0)
('PUNJAB', 'AMRITSAR', 0)
('RAJASTHAN', 'AJMER', 0)
('SIKKIM', 'NORTH', 0)
('TAMIL NADU', 'ARIYALUR', 0)
('TRIPURA', 'DHALAI', 0)
('UTTAR PRADESH', 'AGRA', 0)
('UTTARAKHAND', 'ALMORA', 0)
('WEST BENGAL', '24 PARGANAS NORTH', 0)
```

The Above result of query shows the minimum number of Murdrs happened in eah Ditsrict

3.7 Write SQL query to find the number of murders in ascending order in district and yearwise.

In [134]:
```
result=cursor.execute("SELECT  state, DISTRICT, Year,Murder FROM crime_against_
for row in result:
    print(row)
```

```
('A & N ISLANDS', 'ANDAMAN', 2001, 0)
('ANDHRA PRADESH', 'ADILABAD', 2001, 0)
('ARUNACHAL PRADESH', 'CHANGLANG', 2001, 0)
('ASSAM', 'BARPETA', 2001, 0)
('CHANDIGARH', 'CHANDIGARH', 2001, 0)
('CHHATTISGARH', 'BALRAMPUR', 2001, 0)
('D & N HAVELI', 'D and N HAVELI', 2001, 0)
('DAMAN & DIU', 'DAMAN', 2001, 0)
('DELHI', 'CENTRAL', 2001, 0)
('GOA', 'NORTH GOA', 2001, 0)
('GUJARAT', 'AHMEDABAD COMMR.', 2001, 0)
('HARYANA', 'AMBALA', 2001, 0)
('HIMACHAL PRADESH', 'BILASPUR', 2001, 0)
('JAMMU & KASHMIR', 'ANANTNAG', 2001, 0)
('JHARKHAND', 'BOKARO', 2001, 0)
('KARNATAKA', 'BAGALKOT', 2001, 0)
('KERALA', 'ALAPUZHA', 2001, 0)
('LAKSHADWEEP', 'LAKSHADWEEP', 2001, 0)
('MANIPUR', 'BISHNUPUR', 2001, 0)
('MEGHALAYA', 'GARO HILLS EAST', 2001, 0)
('MIZORAM', 'AIZAWL', 2001, 0)
('NAGALAND', 'DIMAPUR', 2001, 0)
('ODISHA', 'ANGUL', 2001, 0)
('PUDUCHERRY', 'PONDICHERRY', 2001, 0)
('PUNJAB', 'AMRITSAR', 2001, 0)
('RAJASTHAN', 'AJMER', 2001, 0)
('TAMIL NADU', 'ARIYALUR', 2001, 0)
('TRIPURA', 'DHALAI', 2001, 0)
('UTTAR PRADESH', 'AGRA', 2001, 0)
('UTTARAKHAND', 'ALMORA', 2001, 0)
('WEST BENGAL', '24 PARGANAS NORTH', 2001, 0)
('BIHAR', 'ARARIA', 2001, 1)
('MADHYA PRADESH', 'BALAGHAT', 2001, 1)
('MAHARASHTRA', 'AHMEDNAGAR', 2001, 1)
('SIKKIM', 'EAST', 2001, 4)
```

District ,Year and Murders in Ascending orders

3.8.1Insert records of STATE/UT, DISTRICT, YEAR, MURDER, ATTEMPT TO MURDER, and RAPE columns only from 01_District_wise_crimes_committed_IPC_2001_2012.csv into a new table

In [18]:
```
import pandas as pd
```

In [78]:
```
cursor.execute ("CREATE TABLE crime_ipc(STATE UT TEXT,DISTRICT TEXT,YEAR INT, M
```

Out[78]:
```
<sqlite3.Cursor at 0x1c69e1ad260>
```

In [79]:
```

db.commit()
```

In [80]:
```
with open('01_District_wise_crimes_committed_IPC_2001_2012.csv','r') as file :
    #here r is read as
    no_records=0
    for row in file:
        cursor.execute("INSERT INTO crime_ipc VALUES(?,?,?,?,?,?)",row.split(",
        db.commit()
        no_records += 1
```

In [81]:
```
print(no_records,'Records Inserted')
```

```
9017 Records Inserted
```

3.8.2 Write SQL query to find which District in each state/ut has the highest number of murders yearwise. Your output should show STATE/UT, YEAR, DISTRICT, and MURDERS.

In [32]:
```
result=cursor.execute("SELECT  state UT, Year,DISTRICT ,MAX(MURDER) FROM crime_
for row in result:
    print(row)
```

```
('BIHAR', 2004, 'PATNA', 542)
('UTTAR PRADESH', 2001, 'MUZAFFARNAGAR', 324)
('MAHARASHTRA', 2001, 'MUMBAI', 295)
('JHARKHAND', 2002, 'RANCHI', 270)
('WEST BENGAL', 2010, 'PASCHIM MIDNAPUR', 268)
('KARNATAKA', 2010, 'BANGALORE COMMR.', 266)
('GUJARAT', 2002, 'AHMEDABAD COMMR.', 222)
('ANDHRA PRADESH', 2001, 'RANGA REDDY', 214)
('TAMIL NADU', 2012, 'CHENNAI', 180)
('JAMMU & KASHMIR', 2001, 'BARAMULLA', 144)
('ASSAM', 2007, 'TINSUKIA', 143)
('MADHYA PRADESH', 2004, 'JHABUA', 139)
('CHHATTISGARH', 2006, 'BIZAPUR', 139)
('TRIPURA', 2003, 'WEST', 136)
('DELHI UT', 2002, 'NORTH WEST', 135)
('ODISHA', 2009, 'KEONJHAR', 109)
('RAJASTHAN', 2011, 'ALWAR', 104)
('UTTARAKHAND', 2001, 'HARIDWAR', 94)
('HARYANA', 2006, 'GURGAON', 93)
('PUNJAB', 2010, 'CP LUDHIANA', 90)
('NAGALAND', 2008, 'DIMAPUR', 64)
('MANIPUR', 2007, 'IMPHAL WEST', 61)
('KERALA', 2002, 'TRIVANDRUM', 58)
('MEGHALAYA', 2003, 'JAINTIA HILLS', 52)
('HIMACHAL PRADESH', 2003, 'KANGRA', 38)
('GOA', 2009, 'NORTH GOA', 33)
('PUDUCHERRY', 2011, 'PUDUCHERRY', 32)
('CHANDIGARH', 2002, 'CHANDIGARH', 31)
('MIZORAM', 2010, 'AIZAWL', 23)
('ARUNACHAL PRADESH', 2007, 'LOHIT', 17)
('A & N ISLANDS', 2003, 'ANDAMAN', 16)
('D & N HAVELI', 2011, 'D and N HAVELI', 14)
('SIKKIM', 2001, 'EAST', 13)
('DAMAN & DIU', 2007, 'DAMAN', 12)
('LAKSHADWEEP', 2001, 'LAKSHADWEEP', 1)
```

In [33]:
```
dfsql=pd.read_sql("SELECT  state UT, Year,DISTRICT ,MAX(MURDER) AS Murder FROM
```

In [34]:
```
#converting into dataframe
dfsql
```

Out[34]:

	UT	YEAR	DISTRICT	Murder
0	BIHAR	2004	PATNA	542
1	UTTAR PRADESH	2001	MUZAFFARNAGAR	324
2	MAHARASHTRA	2001	MUMBAI	295
3	JHARKHAND	2002	RANCHI	270
4	WEST BENGAL	2010	PASCHIM MIDNAPUR	268
5	KARNATAKA	2010	BANGALORE COMMR.	266
6	GUJARAT	2002	AHMEDABAD COMMR.	222
7	ANDHRA PRADESH	2001	RANGA REDDY	214
8	TAMIL NADU	2012	CHENNAI	180
9	JAMMU & KASHMIR	2001	BARAMULLA	144
10	ASSAM	2007	TINSUKIA	143
11	MADHYA PRADESH	2004	JHABUA	139
12	CHHATTISGARH	2006	BIZAPUR	139
13	TRIPURA	2003	WEST	136
14	DELHI UT	2002	NORTH WEST	135
15	ODISHA	2009	KEONJHAR	109
16	RAJASTHAN	2011	ALWAR	104
17	UTTARAKHAND	2001	HARIDWAR	94
18	HARYANA	2006	GURGAON	93
19	PUNJAB	2010	CP LUDHIANA	90
20	NAGALAND	2008	DIMAPUR	64
21	MANIPUR	2007	IMPHAL WEST	61
22	KERALA	2002	TRIVANDRUM	58
23	MEGHALAYA	2003	JAINTIA HILLS	52
24	HIMACHAL PRADESH	2003	KANGRA	38
25	GOA	2009	NORTH GOA	33
26	PUDUCHERRY	2011	PUDUCHERRY	32
27	CHANDIGARH	2002	CHANDIGARH	31
28	MIZORAM	2010	AIZAWL	23
29	ARUNACHAL PRADESH	2007	LOHIT	17
30	A & N ISLANDS	2003	ANDAMAN	16
31	D & N HAVELI	2011	D and N HAVELI	14
32	SIKKIM	2001	EAST	13
33	DAMAN & DIU	2007	DAMAN	12
34	LAKSHADWEEP	2001	LAKSHADWEEP	1

3.8.3 Store the above data (the result of 3.2) in DataFrame and analyze districts that appear 3 or more than 3 years and print the corresponding state/ut, district, murders, and year in descending order.

In [31]:
```
#.3.2 Write SQL query to find the highest number of rapes & Kidnappings that ha
data=cursor.execute("SELECT  state UT, DISTRICT ,Year ,MAX(Rape), MAX(Kindanppi
for row in data:
    print(row)
```

```
('WEST BENGAL', 'MURSHIDABAD', 2011, 433, 492)
('WEST BENGAL', 'MURSHIDABAD', 2012, 257, 464)
('WEST BENGAL', 'MURSHIDABAD', 2010, 526, 441)
('DELHI', 'NORTH-WEST', 2005, 236, 349)
('WEST BENGAL', 'MURSHIDABAD', 2009, 568, 342)
('UTTAR PRADESH', 'LUCKNOW', 2008, 334, 331)
('DELHI', 'NORTH WEST', 2001, 145, 298)
('DELHI', 'NORTH-WEST', 2006, 224, 287)
('DELHI', 'NORTH WEST', 2002, 159, 259)
('DELHI', 'OUTER', 2007, 369, 237)
('ASSAM', 'GUWAHATI CITY', 2004, 197, 228)
('DELHI', 'NORTH WEST', 2003, 137, 191)
```

Converting into DataFrame

In [80]:
```
dataframe=pd.read_sql("SELECT  state UT, DISTRICT ,Year ,MAX(Rape) AS Rape, MAX
```

In [81]:
```
dataframe
```

Out[81]:

	UT	DISTRICT	Year	Rape	Kidnapping
0	WEST BENGAL	MURSHIDABAD	2011	433	492
1	WEST BENGAL	MURSHIDABAD	2012	257	464
2	WEST BENGAL	MURSHIDABAD	2010	526	441
3	DELHI	NORTH-WEST	2005	236	349
4	WEST BENGAL	MURSHIDABAD	2009	568	342
5	UTTAR PRADESH	LUCKNOW	2008	334	331
6	DELHI	NORTH WEST	2001	145	298
7	DELHI	NORTH-WEST	2006	224	287
8	DELHI	NORTH WEST	2002	159	259
9	DELHI	OUTER	2007	369	237
10	ASSAM	GUWAHATI CITY	2004	197	228
11	DELHI	NORTH WEST	2003	137	191

In [82]:
```
dataframe['DISTRICT'].value_counts()
```

Out[82]:
```
MURSHIDABAD      4
NORTH WEST       3
NORTH-WEST       2
LUCKNOW          1
OUTER            1
GUWAHATI CITY    1
Name: DISTRICT, dtype: int64
```

From above We can Clearly see that Murshidabad is the state which appreared more than three years.

In [36]:
```
from sqlalchemy import create_engine
```

In [37]:
```
engine = create_engine("sqlite+pysqlite:///:memory:")
```

In [38]:
```
dataframe.to_sql('district_data', engine)
```

Out[38]:
```
12
```

In [40]:
```
from sqlalchemy import text
```

Analyze districts that appear 3 or more than 3 years and print the corresponding state/ut, district, murders, and year in descending order.

In [100]:
```
with engine.connect() as conn:
    result = conn.execute(text("select UT, DISTRICT ,Year FROM district_data OR
    for row in result:
        print(row)
```

```
('WEST BENGAL', 'MURSHIDABAD', 2011)
('WEST BENGAL', 'MURSHIDABAD', 2012)
('WEST BENGAL', 'MURSHIDABAD', 2010)
```

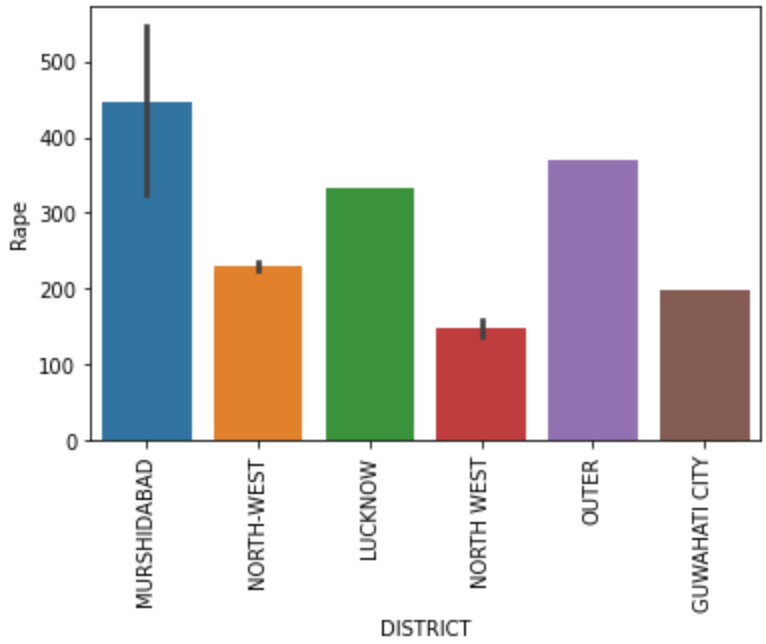
From Above result West Bengal, Murshidabas is the District which appeared more than 3 times year wise

3.8.4 Use appropriate graphs to show your data (the result of 3.8.3)

In [104]:
```
import matplotlib.pyplot as plt
import seaborn as sns
```

In [114]:

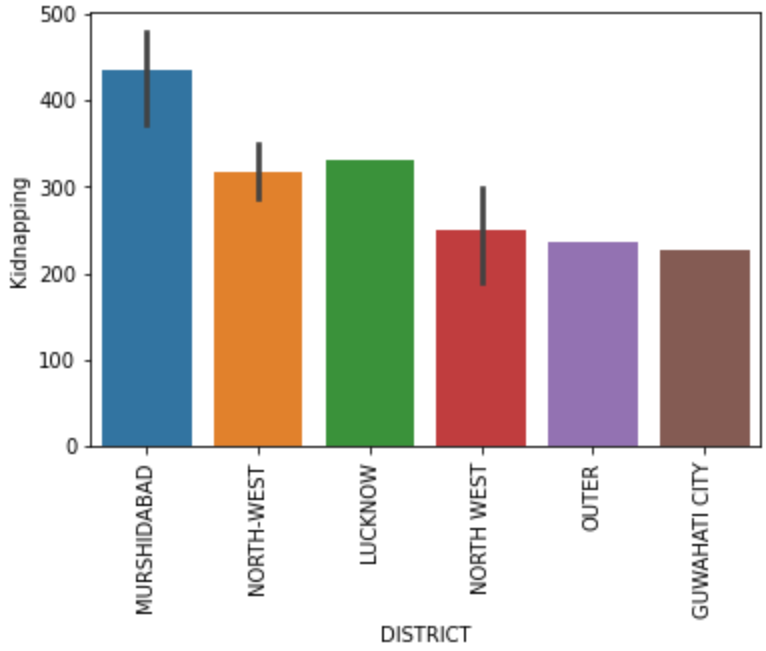
```
plt.figure(figsize=(6,4))
sns.barplot(x=dataframe['DISTRICT'], y=dataframe['Rape'])
plt.xticks(rotation='90')
plt.show()
```



__From Above BarPlot its clear that Murshidabas is the Ditsrict of West Bengal which has highest cases of Rape and Guwahati City District of Assam has lowest cases of Rape as per the dataframe

Anlaysis of District vs Kidnapping

In [115]:
```
plt.figure(figsize=(6,4))
sns.barplot(x=dataframe['DISTRICT'], y=dataframe['Kidnapping'])
plt.xticks(rotation='90')
plt.show()
```



Again Murshidabad District of West Bengal has highest cases of Kidnapping and Guwahati City Ditrict Assam has Lowest Kidnapping cases as per the resulted Dataframe