**3.1 Insert records from 42\_District\_wise\_crimes\_committed\_against\_women\_2001\_2012.csv into a table.**

First, create a table for the CSV data. Assuming the CSV has columns like State, District, Year, Rape, Kidnapping, etc.:

CREATE TABLE Crimes\_Against\_Women (

State VARCHAR(50),

District VARCHAR(50),

Year INT,

Rape INT,

Kidnapping INT,

Other\_Crimes INT

);

COPY Crimes\_Against\_Women FROM '/path/to/42\_District\_wise\_crimes\_committed\_against\_women\_2001\_2012.csv' CSV HEADER;

**3.2 Write SQL query to find the highest number of rapes & kidnappings that happened in which state, district, and year.**

SELECT

State,

District,

Year,

MAX(Rape) AS Max\_Rapes,

MAX(Kidnapping) AS Max\_Kidnappings

FROM

Crimes\_Against\_Women

GROUP BY

State, District, Year

ORDER BY

Max\_Rapes DESC, Max\_Kidnappings DESC

LIMIT 1;

**3.3Write SQL query to find all the lowest number of rapes & kidnappings that happened in which state, district, and year.**

Sql

SELECT

State,

District,

Year,

MIN(Rape) AS Min\_Rapes,

MIN(Kidnapping) AS Min\_Kidnappings

FROM

Crimes\_Against\_Women

GROUP BY

State, District, Year

ORDER BY

Min\_Rapes ASC, Min\_Kidnappings ASC

LIMIT 1;

3.4**Insert records from 02\_District\_wise\_crimes\_committed\_against\_ST\_2001\_2012.csv into a new table.**

First, create a table for this CSV data:

sql

Copy code

CREATE TABLE Crimes\_Against\_ST (

State VARCHAR(50),

District VARCHAR(50),

Year INT,

Dacoity INT,

Robbery INT,

Other\_Crimes INT

);

COPY Crimes\_Against\_ST FROM '/path/to/02\_District\_wise\_crimes\_committed\_against\_ST\_2001\_2012.csv' CSV HEADER;

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

SELECT

District,

MAX(Dacoity) AS Max\_Dacoity,

MAX(Robbery) AS Max\_Robbery

FROM

Crimes\_Against\_ST

GROUP BY

District ORDER BY

Max\_Dacoity DESC,

Max\_Robbery DESC

LIMIT 1;

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

CREATE TABLE Crimes\_IPC ( State

VARCHAR(50),

District VARCHAR(50),

Year INT,

Murder INT,

Attempt\_to\_Murder INT, Rape INT );

COPY Crimes\_IPC FROM '/path/to/01\_District\_wise\_crimes\_committed\_IPC\_2001\_2012.csv' CSV HEADER;

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

sql

Copy code

SELECT

District,

Year,

Murder

FROM

Crimes\_IPC

ORDER BY

Murder ASC, District ASC, Year ASC;

**3.8.1 Insert 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.**

Create a new table for specific columns:

sql

Copy code

CREATE TABLE Selected\_Crimes\_IPC (

State VARCHAR(50),

District VARCHAR(50),

Year INT,

Murder INT,

Attempt\_to\_Murder INT,

Rape INT

);

Insert the specific records:

sql

Copy code

COPY Selected\_Crimes\_IPC (State, District, Year, Murder, Attempt\_to\_Murder, Rape)

FROM '/path/to/01\_District\_wise\_crimes\_committed\_IPC\_2001\_2012.csv' CSV HEADER;

**3.8.2 Write SQL query to find which district in each state/UT has the highest number of murders year wise.**

sql

Copy code

SELECT

State,

Year,

District,

MAX(Murder) AS Max\_Murders

FROM

Selected\_Crimes\_IPC

GROUP BY

State, Year, District

ORDER BY

State, Year, Max\_Murders DESC;

**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.**

Using Python and Pandas:

python

Copy code

import pandas as pd

import sqlite3

# Connect to the database

conn = sqlite3.connect('crime\_data.db')

# Query to get the data

query = """

SELECT State, Year, District, MAX(Murder) AS Max\_Murders

FROM Selected\_Crimes\_IPC

GROUP BY State, Year, District

ORDER BY State, Year, Max\_Murders DESC;

"""

# Load the data into a DataFrame

df = pd.read\_sql\_query(query, conn)

# Filter districts that appear 3 or more years

district\_counts = df['District'].value\_counts()

frequent\_districts = district\_counts[district\_counts >= 3].index

filtered\_df = df[df['District'].isin(frequent\_districts)]

# Print the filtered DataFrame

print(filtered\_df.sort\_values(by='Max\_Murders', ascending=False))

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

Using Matplotlib or Seaborn in Python:

python

Copy code

import matplotlib.pyplot as plt

import seaborn as sns

# Plot the data

plt.figure(figsize=(14, 7))

sns.barplot(x='District', y='Max\_Murders', hue='Year', data=filtered\_df)

plt.title('Districts with Highest Number of Murders (Appearing 3 or More Years)')

plt.xlabel('District')

plt.ylabel('Number of Murders')

plt.legend(title='Year')

plt.xticks(rotation=90)

plt.show()