Crime In India

PHASE 3- SQL OPERATIONS By Priti Varma

CAPSTONE PROJECT

**Project Description:**

Crime in India is prevalent and occurs in various forms. Besides violent crimes such as homicide, robbery, and assault, and property crimes like burglary, theft, motor vehicle theft, and arson, the country also faces significant issues with organized crime, the illegal drug trade, arms trafficking, corruption, and many other criminal activities.

**Objective:**

It is crucial to identify specific details about crime and the data we have, in order to reduce crime rates in particular regions.

**Tech Stack Used:**

My SQL Workbench- It help in extracting data and help us in analysing the area which needs mor focus.

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

A computer screen shot of a computer code

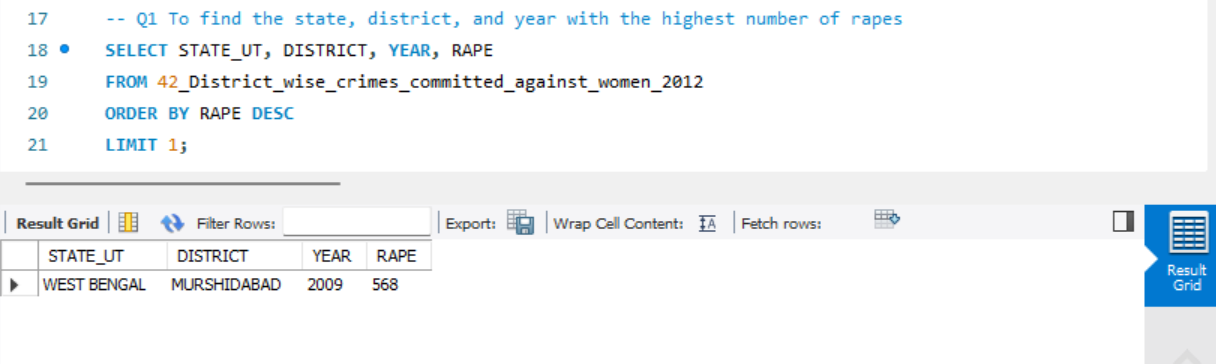
Description automatically generated

For the Categorical columns “VARCHAR” datatype was used, For numerical columns “INT” datatype was used. Constraints are provided to some columns to get proper results.

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

To find the state, district, and year with the highest number of rapes and kidnappings, you can use the following SQL query:

Query to find State, District, Year with the highest number of rapes.



Query to find the State, District, Year with the highest number of kidnappings and abductions.

A screenshot of a computer

Description automatically generated

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

To find all instances of the lowest number of rapes and kidnappings that happened in each state, district, and year, you can use the following SQL queries:

Query for the lowest number of rapes:

A screenshot of a computer

Description automatically generated

Query for the lowest number of kidnappings and abductions:

A screenshot of a computer

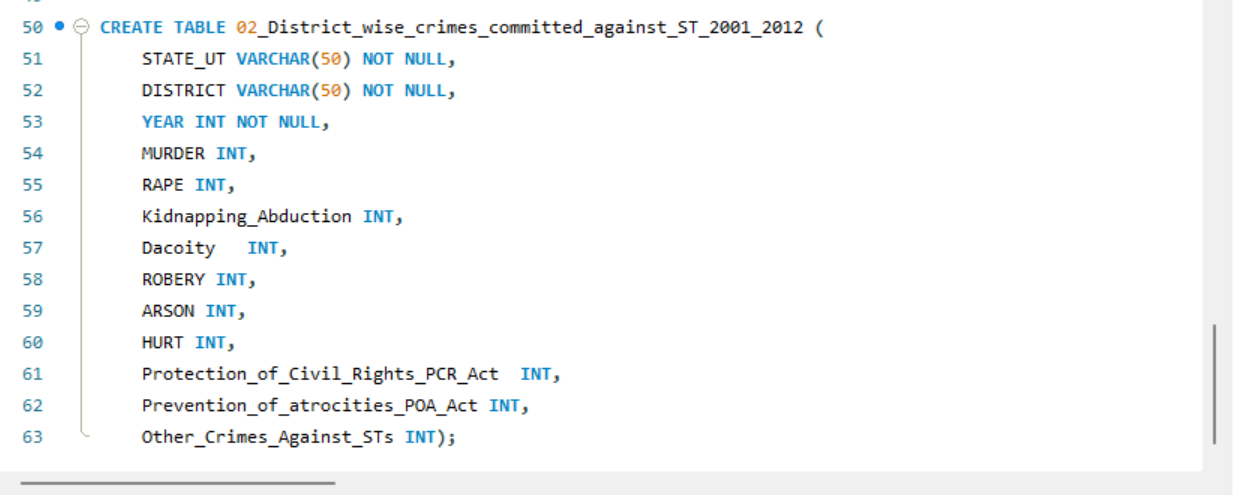
Description automatically generated

**Explanation:**

1. **Finding the minimum value:**
   * The first part of each query finds the minimum value for RAPE and Kidnapping\_and\_Abduction respectively.
2. **Finding all instances with the minimum value:**
   * The second part of each query selects all rows where the number of rapes or kidnappings matches the minimum value found in the first part of the query.

This will give you all the state, district, and year combinations where the lowest number of these crimes occurred.

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



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

To find the district with the highest number of dacoity and robbery, you can use the following SQL query:

Query for the highest number of dacoity:

A screenshot of a computer

Description automatically generated

Query for the highest number of robberies:

A screenshot of a computer

Description automatically generated

**Explanation:**

* The first query selects the STATE\_UT, DISTRICT, YEAR, and Dacoity columns, orders the results by Dacoity in descending order, and limits the result to the top 1, which gives the district with the highest number of dacoity incidents.
* The second query selects the STATE\_UT, DISTRICT, YEAR, and ROBBERY columns, orders the results by ROBBERY in descending order, and limits the result to the top 1, which gives the district with the highest number of robbery incidents.

These queries will provide the specific districts with the highest number of these crimes.

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

To find all districts where the lowest number of murders occurred, you can use the following SQL query:

A screenshot of a computer

Description automatically generated

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

To find the number of murders in ascending order, district-wise and year-wise, you can use the following SQL query:

A screenshot of a computer

Description automatically generated

This will give you a list where murders are sorted first by district and then by year, showing the number of murders in ascending order for each district and year combination.

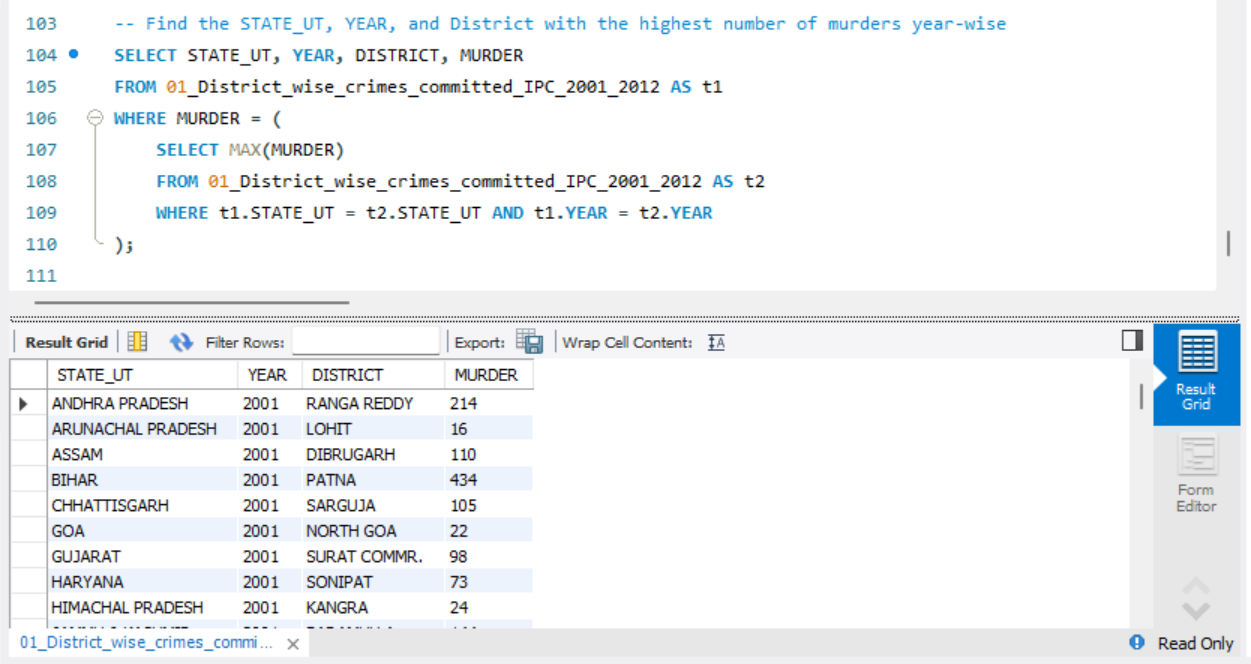
**3.8.1 Insert records of STATE/UT, DISTRICT, YEAR, MURDER, ATTEMPT TO MURDER, and RAPE columns only from 01\_District\_wise\_crimes\_committed\_IPC\_2001\_2012.csv into a new table.**

**A screenshot of a computer

Description automatically generated**

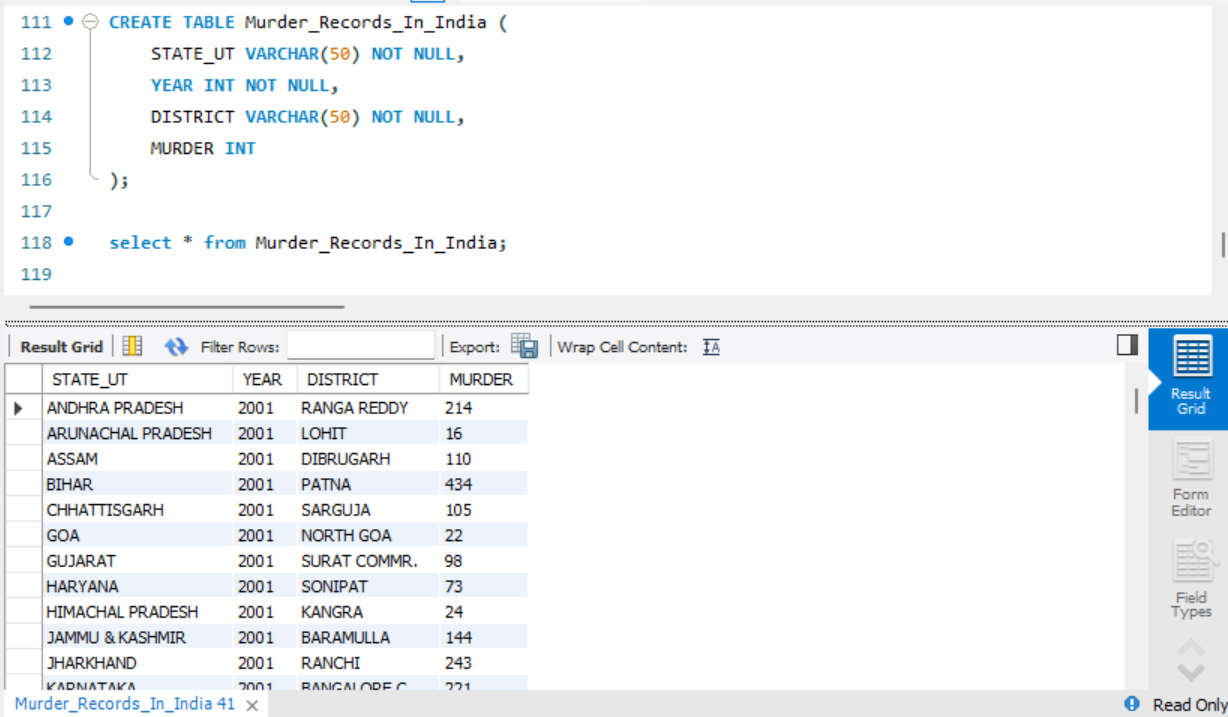
**3.8.2 Write SQL query to find which District in each state/UT has the highest number of murders year wise. Your output should show STATE/UT, YEAR, DISTRICT, and MURDERS.**

To find which STATE\_UT, YEAR, and District within each STATE\_UT has the highest number of murders year-wise, you can use the following SQL query:



**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. [ I believe that the data stored is from question 3.8.2 and not 3.2, as Question 3.2 has only one output, and the columns does not match the requirements]**

The data is first saved in a csv file dataframe and then loaded again with a new table

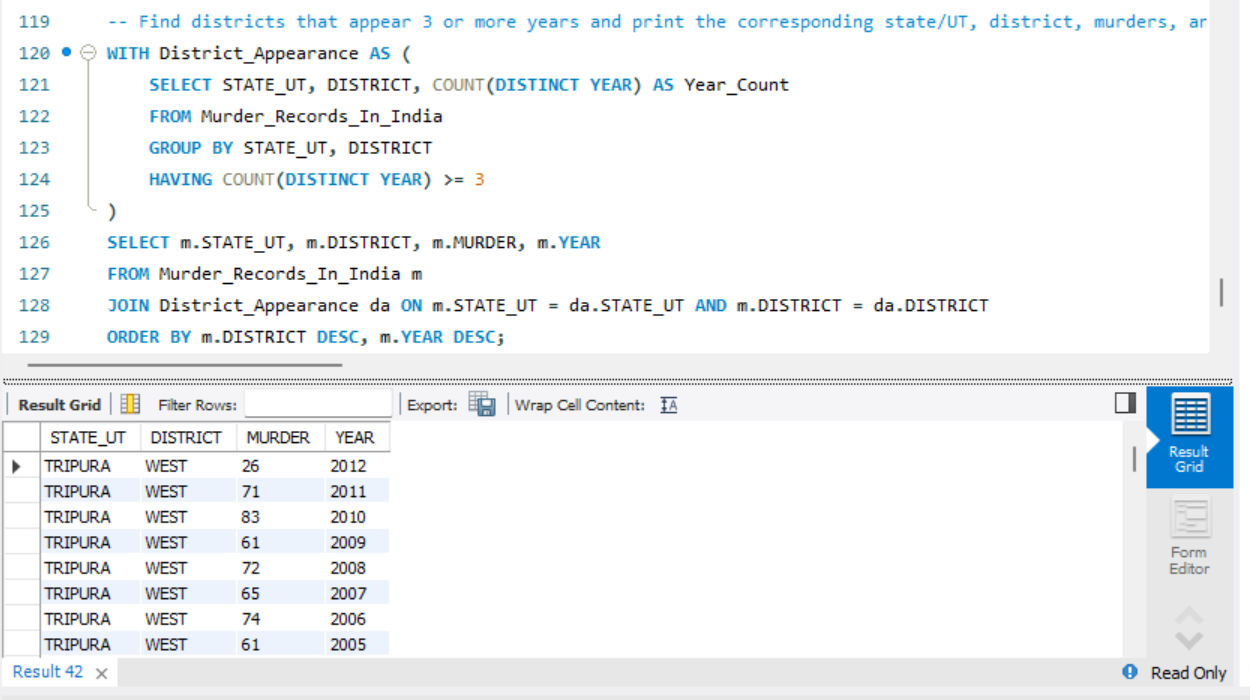


A screenshot of a computer

Description automatically generated

**Explanation:**

1. **Common Table Expression (CTE) District\_Appearance:**
   * This CTE calculates the number of distinct years (Year\_Count) each district appears in the dataset.
   * It groups the results by STATE\_UT and DISTRICT.
   * The HAVING clause filters the results to include only those districts that appear in 3 or more years.



**Explanation of Main Query:**

* The main query selects the STATE\_UT, DISTRICT, MURDER, and YEAR columns from the Murder\_Records\_In\_India table (m alias).
* It joins the Murder\_Records\_In\_India table with the District\_Appearance CTE on STATE\_UT and DISTRICT to include only the districts that appear in 3 or more years.
* The ORDER BY clause sorts the results by DISTRICT and YEAR in descending order.

This will provide a list of districts that have records for 3 or more years, showing the corresponding state/UT, district, number of murders, and year, sorted in descending order by district and year.

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

Following is the Link :- <https://github.com/PritiVasanth/Capstone_Project-DataTrained/blob/main/Districtwise_Murder_Analysis_2.xlsx>

Following is the Graphical representation:-

1. **State Uttar Pradesh:**

* **Insights: - Murder has decreased over a period for the state Uttar Pradesh in the city of Muzaffarnagar.**

1. **State Rajasthan:**

* **Insights: Murder increased over a period of time in Alwar District of Rajasthan.**

1. **State Haryana:**

* **Insights: Murder rate has been almost consistent throughout years in Haryana except the year 2003 when murder rate was very high**.

1. **State West Bengal:**

* **Insights: In the state of West Bengal 24 Praganas North has more Murder cases than 24 Praganas South. In 24 Praganas North Lowest Murder cases were recorded in the year 2006**.

1. **Mumbai:**

* **Insights: It is observed that Over a period of time Murder cases have decreased but there seems to be a minute fall.**

1. **Jammu & Kashmir:**

* **Insights: In the State of Jammu & Kashmir we can observe that the rate off Murder have decreased a lot. Most of the cases were recorded from the district Baramulla, and the rate here too decreased over a period of time.**

1. **State Chhattisgarh:**

* **Insights: In the state most of the cases were recorded from Bilaspur and the highest cases recorded were in the year 2008.**

1. State Meghalaya:

* **Insights: It is observed that in Meghalaya most of the cases were recorded from Jaintia Hills (76%) than Khasi Hills East (24%).**

1. **State Gujarat:**

* **Insights: In Gujarat the highest number of murders were recorded in the year 2002. For the rest of the years the rate seems to be almost the same.**
* **Insights: Rate of Murder recorded in Surat Commr. (68%) is more than Ahmedabad Commr. (32%).**

**Insights: In Ahmedabad the rate of murder cases was recorded less I the year 2003.**

1. **State Assam:**

* **Insights: In the state of Assam Dibrugarh has recorded maximum Murder cases. The rate of murder cases was maximum in the year 2004 and minimum in 2006.**

**--------------------------\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*THANK YOU\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*-----------------------------**