

CRIME DATA ANALYSIS



Jeffrey L. Ritter and John Bookunich

INTRODUCTION

A composite image on the left side of the slide. The top half shows an open metal evidence kit with various tools and evidence bags inside. The bottom half shows a silver metal evidence marker with a small white card labeled '1' standing upright in a dark, textured ground. A small white ruler is visible next to it.

This project focuses on analyzing crime data spanning over 20 years across different states in India, with a special focus on crimes against women.

Using SQL, the project delves into key crime trends, state-wise crime patterns, and comparisons between regions to uncover valuable insights.

Objective:

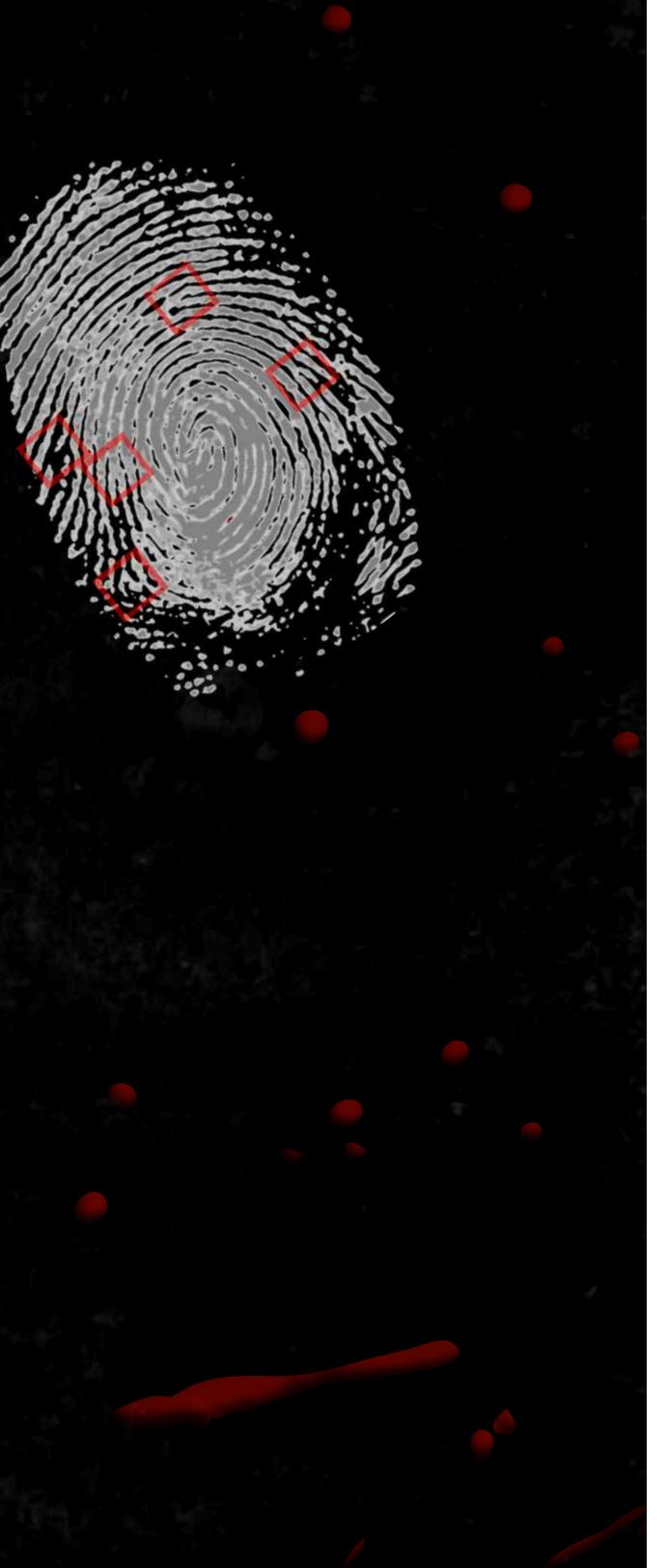
- To leverage SQL for deep data exploration and analysis.
- Identify important crime trends, regional disparities, and annual fluctuations.
- Provide actionable insights through data-driven analysis.

DATASET

DATASET NAME: Crimes Against Women in India: A 20-Year Analysis

SOURCE: Kaggle Dataset ([Link to dataset](#))

TIME PERIOD: 2001 - 2020 (20 Years)



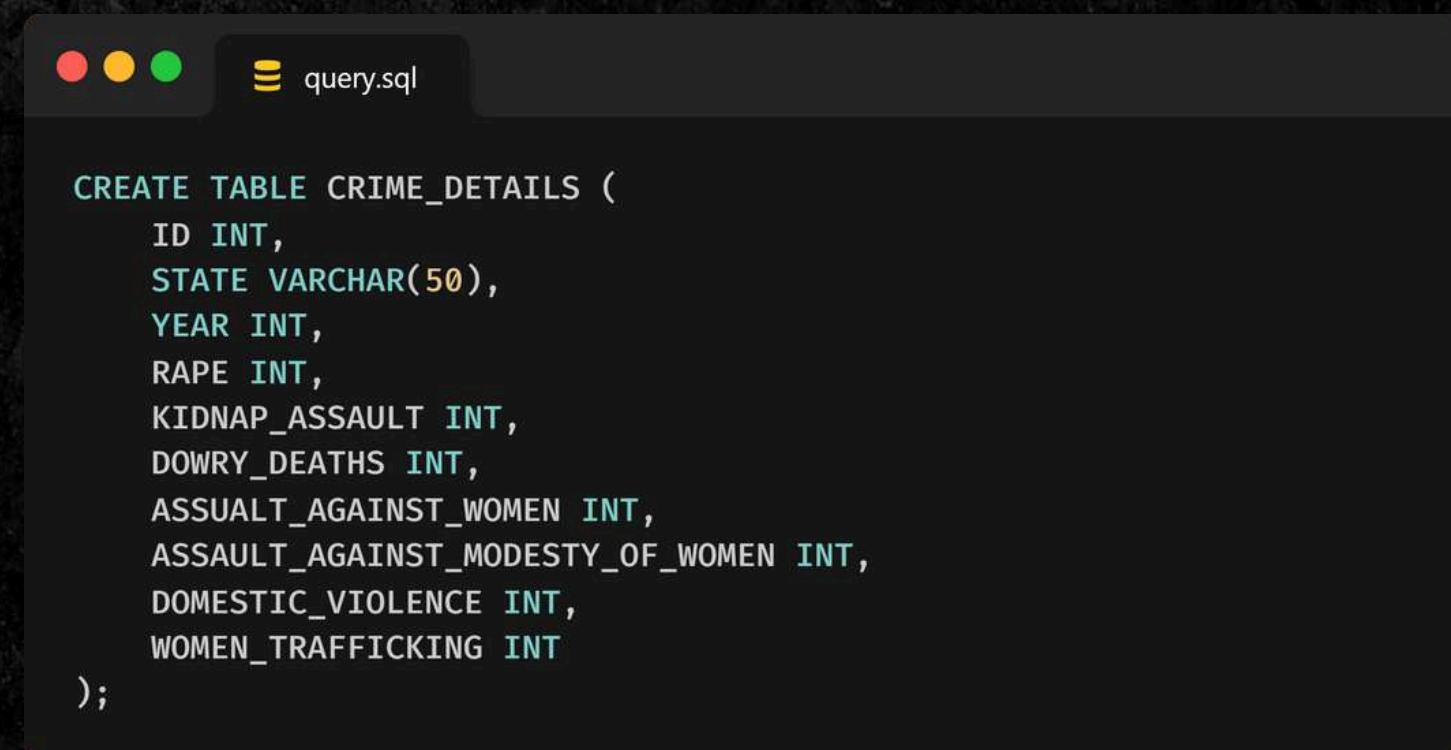
ATTRIBUTES

- **State:** The name of the state or union territory.
- **Year:** The year in which the crimes were reported.
- **Rape:** Number of rape cases reported.
- **Kidnap_Assault:** Number of cases involving kidnapping and assault.
- **Dowry_Deaths:** Number of dowry-related deaths.
- **Assault_Against_Women:** Physical assault cases involving women.
- **Assault_Against_Modesty_Of_Women:** Cases involving sexual assault or harassment of women.
- **Domestic_Violence:** Number of domestic violence cases reported.
- **Women_Trafficking:** Number of human trafficking cases involving women.



QUERIES

CREATING TABLE



A screenshot of a terminal window titled "query.sql". The window contains the following SQL code:

```
CREATE TABLE CRIME_DETAILS (
    ID INT,
    STATE VARCHAR(50),
    YEAR INT,
    RAPE INT,
    KIDNAP_ASSAULT INT,
    DOWRY_DEATHS INT,
    ASSUALT AGAINST WOMEN INT,
    ASSUALT AGAINST MODESTY_OF WOMEN INT,
    DOMESTIC_VIOLENCE INT,
    WOMEN_TRAFFICKING INT
);
```

TOTAL NUMBER OF CRIMES REPORTED FOR EACH TYPE OF CRIME ACROSS ALL STATES?

```
query.sql
```

```
SELECT
    STATE,
    SUM(RAPE) AS TOTAL_RAPES,
    SUM(KIDNAP_ASSAULT) AS TOTAL_KIDNAP_ASSAULT,
    SUM(DOWRY_DEATHS) AS TOTAL_DOWRY_DEATHS,
    SUM(ASSAULT AGAINST WOMEN) AS TOTAL_ASSAULTS,
    SUM(ASSAULT AGAINST MODESTY_OF_WOMEN) AS TOTAL_ASSAULTS_MODESTY,
    SUM(DOMESTIC_VIOLENCE) AS TOTAL_DOMESTIC_VIOELENCE,
    SUM(WOMEN_TRAFFICKING) AS TOTAL_TRAFFICKING
FROM CRIME_DETAILS
GROUP BY 1
ORDER BY 1;
```

STATE	TOTAL_RAPES	TOTAL_KIDNAP_ASSAULT	TOTAL_DOWRY_DEATHS	TOTAL_ASSAULTS	TOTAL_ASSAULTS_MODESTY	TOTAL_DOMESTIC_VIOELENCE	TOTAL_TRAFFICKING
A & N ISLANDS	424	363	13	558	135	365	10
ANDHRA PRADESH	23424	21707	8165	92084	54684	188511	2548
ARUNACHAL PRADESH	1153	1272	4	1584	97	705	7
ASSAM	32190	71462	2757	2757	47115	2698	135415
BIHAR	21006	77687	24428	9953	800	55187	825
CHANDIGARH	795	1547	68	954	281	1689	7
CHHATTISGARH	25178	16536	1797	33310	4382	16381	103
D&N Haveli	2343	7148	256	3977	840	7349	12
DAMAN & DIU	619	1725	30	3642	58	912	29
Delhi UT	13065	31251	1106	26039	6046	27119	103
GOA	1019	720	23	1475	475	356	111

ALL THE STATES WHERE THE TOTAL NUMBER OF CRIMES EXCEEDED 10,000 IN ANY GIVEN YEAR.

```
query.sql

SELECT
    STATE,
    SUM(RAPE) AS TOTAL_RAPES,
    SUM(KIDNAP_ASSAULT) AS TOTAL_KIDNAP_ASSAULT,
    SUM(DOWRY_DEATHS) AS TOTAL_DOWRY_DEATHS,
    SUM(ASSAULT AGAINST WOMEN) AS TOTAL_ASSAULTS,
    SUM(ASSAULT AGAINST MODESTY_OF_WOMEN) AS TOTAL_ASSAULTS_MODESTY,
    SUM(DOMESTIC_VIOLENCE) AS TOTAL_DOMESTIC_VIOELENCE,
    SUM(WOMEN_TRAFFICKING) AS TOTAL_TRAFFICKING
FROM CRIME_DETAILS
GROUP BY 1
ORDER BY 1;
```

	STATE	YEAR	TOTAL_CRIMES
▶	Uttar Pradesh	2016	52246
	Uttar Pradesh	2018	48603
	Uttar Pradesh	2019	47480
	Uttar Pradesh	2017	47139
	Tripura	2021	43448
	Uttar Pradesh	2014	39774
	Rajasthan	2019	39699
	Punjab	2021	38894
	Tripura	2020	38523
	Uttar Pradesh	2015	37998
	West Bengal	2014	36516
	Maharashtra	2015	33993
	Maharashtra	2016	33578
	Punjab	2020	33106

NUMBER MANY CRIMES WERE REPORTED IN EACH YEAR ACROSS ALL STATES

```
query.sql
```

```
SELECT
    STATE,
    YEAR,
    SUM(RAPE) + SUM(KIDNAP_ASSAULT) + SUM(DOWRY_DEATHS) + SUM(ASSAULT AGAINST WOMEN) +
    SUM(ASSAULT AGAINST MODESTY_OF WOMEN) + SUM(DOMESTIC_VIOLENCE) + SUM(WOMEN_TRAFFICKING)
    AS TOTAL_CRIMES
FROM CRIME_DETAILS
GROUP BY 1,2;
```

STATE	YEAR	TOTAL_CRIMES
UTTARAKHAND	2001	744
WEST BENGAL	2001	6533
A & N ISLANDS	2001	34
CHANDIGARH	2001	146
D&N Haveli	2001	19
DAMAN & DIU	2001	7
LAKSHADWEEP	2001	0
PUDUCHERRY	2001	78
ANDHRA PRADESH	2002	15146
ARUNACHAL PRADESH	2002	159
ASSAM	2002	5001
BIHAR	2002	4953
CHHATTISGARH	2002	3514
GOA	2002	51

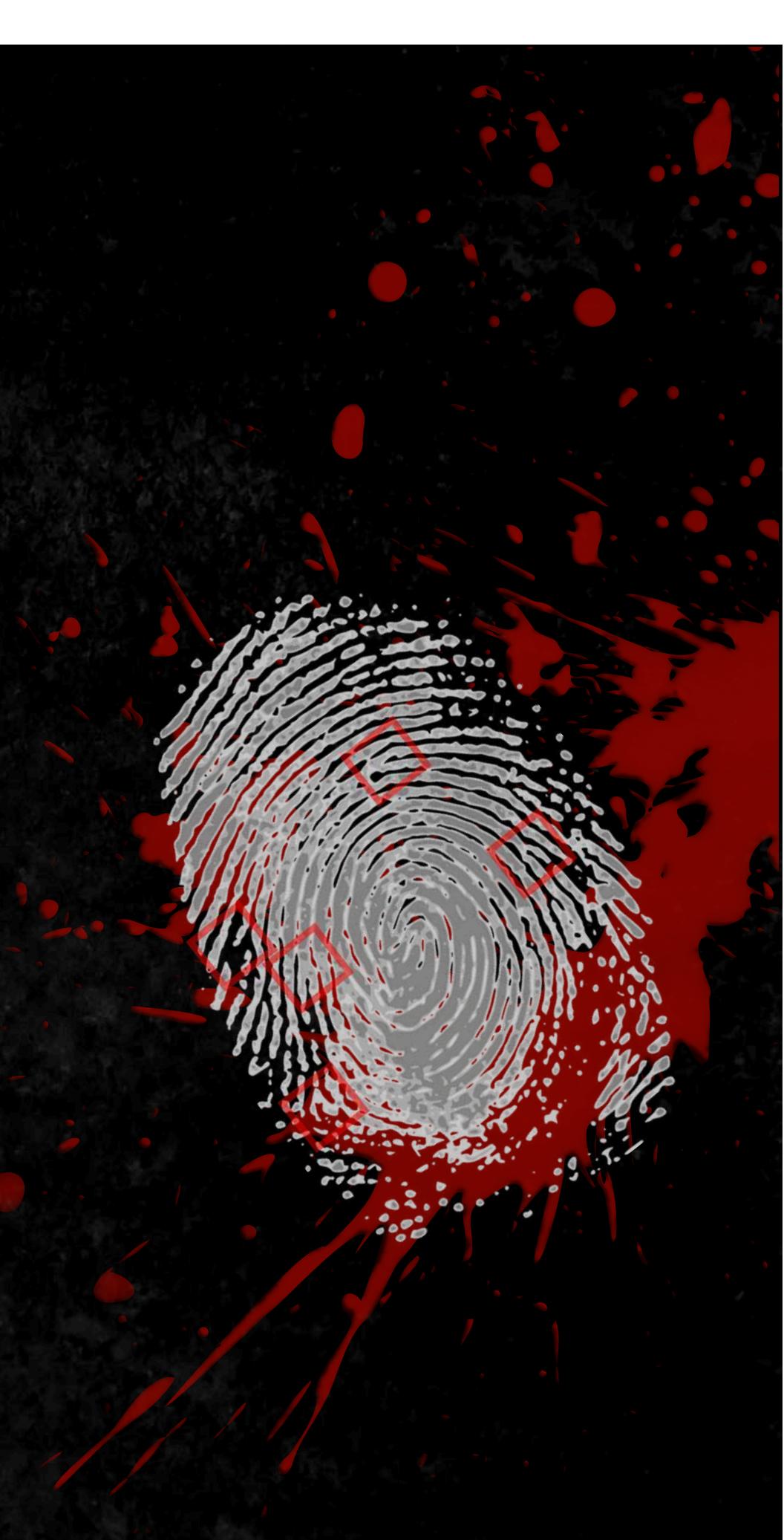


THE TOTAL NUMBER OF CRIMES REPORTED IN DELHI OVER THE 20-YEAR PERIOD

```
query.sql
```

```
SELECT
    STATE,
    YEAR,
    SUM(RAPE) + SUM(KIDNAP_ASSAULT) + SUM(DOWRY_DEATHS) + SUM(ASSAULT AGAINST WOMEN) +
    SUM(ASSAULT AGAINST MODESTY_OF WOMEN) + SUM(DOMESTIC_VIOLENCE) + SUM(WOMEN_TRAFFICKING)
    AS TOTAL_CRIMES
FROM CRIME_DETAILS
WHERE STATE = 'Delhi UT'
GROUP BY 1,2;
```

	STATE	YEAR	TOTAL_CRIME
▶	Delhi UT	2011	4569
	Delhi UT	2012	5944
	Delhi UT	2013	12873
	Delhi UT	2014	14832
	Delhi UT	2015	16663
	Delhi UT	2016	15050
	Delhi UT	2017	11366
	Delhi UT	2018	11757
	Delhi UT	2019	11648
	Delhi UT	2020	0



DIFFERENCE BETWEEN TOTAL NUMBER OF CRIMES BETWEEN THE YEAR 2019 AND 2020

```
query.sql

WITH TOTAL_CRIMES AS (
    SELECT
        STATE,
        YEAR,
        SUM(RAPE) + SUM(KIDNAP_ASSAULT) + SUM(DOWRY_DEATHS) + SUM(ASSAULT AGAINST WOMEN) +
        SUM(ASSAULT AGAINST MODESTY_OF_WOMEN) + SUM(DOMESTIC_VIOLENCE) + SUM(WOMEN_TRAFFICKING) AS TOTAL_CRIMES
    FROM CRIME_DETAILS
    WHERE YEAR BETWEEN 2019 AND 2020
    GROUP BY 1,2
)
SELECT *,
    LAG(TOTAL_CRIMES, 1, 0) OVER(PARTITION BY STATE ORDER BY YEAR) AS PREVIOUS_YEAR_CRIMES,
    TOTAL_CRIMES - LAG(TOTAL_CRIMES, 1, 0) OVER(PARTITION BY STATE ORDER BY YEAR) AS CRIME_DECREASE
FROM TOTAL_CRIMES;
```

	STATE	YEAR	TOTAL_CRIMES	PREVIOUS_YEAR_CRIMES	CRIME_DECREASE
▶	A & N Islands	2019	42	0	42
	A & N Islands	2020	293	42	251
	Andhra Pradesh	2019	16062	0	16062
	Andhra Pradesh	2020	15562	16062	-500
	Arunachal Pradesh	2019	272	0	272
	Arunachal Pradesh	2020	250	272	-22
	Assam	2019	25877	0	25877
	Assam	2020	23327	25877	-2550
	Bihar	2019	13635	0	13635
	Bihar	2020	11078	13635	-2557
	Chandigarh	2019	505	0	505
	Chandigarh	2020	18	505	-487
	Chhattisgarh	2019	5437	0	5437
	Chhattisgarh	2020	5006	5437	-431

COMPARE THE TOTAL NUMBER OF CRIMES BETWEEN NORTHERN AND SOUTHERN STATES FOR EACH YEAR

query.sql

```
SELECT
CASE
    WHEN STATE IN ('Delhi UT', 'Haryana', 'Himachal Pradesh', 'Jammu & Kashmir', 'West Bengal',
                    'Punjab', 'Rajasthan', 'Uttar Pradesh', 'Uttarakhand', 'Jharkhand') THEN "Northern"
    WHEN STATE IN ('Andhra Pradesh', 'Goa', 'Karnataka', 'Kerala',
                    'Tamil Nadu', 'Telangana', 'Puducherry', 'Lakshadweep') THEN 'Southern'
    WHEN STATE IN ('Arunachal Pradesh', 'Assam', 'Manipur',
                    'Meghalaya', 'Mizoram', 'Nagaland', 'Sikkim', 'Tripura') THEN 'Northeastern'
    WHEN STATE IN ('Bihar', 'Chhattisgarh', 'Gujarat',
                    'Madhya Pradesh', 'Maharashtra', 'Odisha') THEN 'Central'
    WHEN STATE IN ('A & N Islands', 'Chandigarh', 'Daman & Diu',
                    'Lakshadweep', 'Puducherry', 'Delhi UT') OR STATE LIKE '%Haveli%' THEN 'Union Territories'
    END AS STATE_CATEGORY,
    SUM(RAPE) + SUM(KIDNAP_ASSAULT) + SUM(DOWRY_DEATHS) + SUM(ASSAULT AGAINST WOMEN) +
    SUM(ASSAULT AGAINST MODESTY_OF WOMEN) + SUM(DOMESTIC_VIOLENCE) + SUM(WOMEN_TRAFFICKING) AS TOTAL_CRIMES
FROM CRIME_DETAILS
GROUP BY 1
ORDER BY 2;
```

STATE_CATEGORY	TOTAL_CRIMES
Union Territories	36149
Northeastern	468562
Southern	982522
Central	1415475
Northern	1965014

STATES THAT HAVE CONSISTENTLY RANKED IN THE TOP 5 FOR THE HIGHEST NUMBER OF CRIMES REPORTED OVER THE YEARS

query.sql

```
WITH YEAR_WISE_CRIMES AS (
    SELECT
        STATE,
        YEAR,
        SUM(RAPE) + SUM(KIDNAP_ASSAULT) + SUM(DOWRY_DEATHS) + SUM(ASSAULT AGAINST WOMEN) +
        SUM(ASSAULT AGAINST MODESTY_OF_WOMEN) + SUM(DOMESTIC_VIOLENCE) + SUM(WOMEN_TRAFFICKING) AS TOTAL_CRIMES
    FROM CRIME_DETAILS
    GROUP BY 1,2
)
SELECT STATE,
    COUNT(*) AS TOP_5_APPEARENCE_COUNT
FROM (
    SELECT *,
        DENSE_RANK() OVER(PARTITION BY YEAR ORDER BY TOTAL_CRIMES DESC) AS CRIME_RANKINGS
    FROM YEAR_WISE_CRIMES
) AS SUBQUERY
WHERE CRIME_RANKINGS BETWEEN 1 AND 5
GROUP BY 1;
```

STATE	TOP_5_APPEARENCE_COUNT
UTTAR PRADESH	19
MADHYA PRADESH	18
ANDHRA PRADESH	13
MAHARASHTRA	16
RAJASTHAN	15
WEST BENGAL	15
Assam	2
Tripura	2
Punjab	2
Uttarakhand	2
Nagaland	1

CONCLUSION

- SQL was used to analyze 20 years of crime data across Indian states, focusing on crimes against women.
- Identified states with the highest and lowest crime rates, highlighting regions with significant crime disparities.
- Analyzed yearly crime trends, revealing patterns of increase and decrease in various types of crimes.
- Compared Northern and Southern states, showing geographic differences in crime rates.
- Provided a data-driven foundation for making informed decisions in law enforcement and public policy.