Introduction to Data Management

PROJECT REPORT

(Project Semester January-April 2025)

Trends in Crimes Against Women in India (2001–2014)

Submitted by:

Om Gupta

Registration No: 12315354

Section: K23GX

Program: B. Tech - Computer Science and Engineering

Course Code: INT217

Under the Guidance of

Savleen Kaur (UID: 18306)

Discipline of CSE/IT

Lovely School of Computer Science and Engineering

Lovely Professional University, Phagwara

CERTIFICATE

This is to certify that Om Gupta bearing Registration no. 12315354 has completed INT217 (Introduction to Data Management) project titled, “Trends in Crimes Against Women in India (2001–2014)” under my guidance and supervision. To the best of my knowledge, the present work is the result of his original development, effort and study.

Signature
Savleen Kaur (UID: 18306)
Designation of the Supervisor
School of Computer Science and Engineering
Lovely Professional University
Phagwara, Punjab.

Date:11/04/2024

DECLARATION

I, Om Gupta, student of B.Tech Computer Science and Engineering under CSE/IT Discipline at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Om Gupta

12315354 – K23GX

B. Tech - Computer Science and Engineering Lovely Professional University, Phagwara

22/04/2025

Acknowledgement

The opportunity of attaining a course based on Data Management using Excel at Lovely Professional University was worth learning. It was a prestige for me to be part of it. During the period of my course, I received tremendous knowledge related to Microsoft Excel and Data Management.

Pre-eminently, I would like to express my deep gratitude and special thanks to my course teacher Jaffar Amin Chacket for her theoretical knowledge and encouragement on this project and for her valuable guidance and affection for the successful completion of this project.

Secondly, I would like to thank Lovely Professional University for giving me an opportunity to learn this course.

Lastly, I would like to thank the almighty and my parents for their constant encouragement, moral support, personal attention, and care.

Om Gupta
12315354 – K23GX
B. Tech - Computer Science and Engineering
Lovely Professional University, Phagwara

Contents

Abstract

Excel is a powerful software program developed by Microsoft that uses spreadsheets to organize, analyze, and visualize data through formulas, functions, pivot tables, charts, and more. Its versatility and analytical capabilities make it a widely used tool for data analysis across industries.

In this project, I have built an Excel Dashboard to analyze and visualize district-wise crime data against women in India from 2001 to 2014. The dashboard includes tables, charts, and other visual tools to highlight key insights and trends across various crime categories such as rape, kidnapping, dowry deaths, and domestic violence. By bringing together the most critical aspects of the dataset into one interactive view, the dashboard supports better understanding and informed decision-making.

This report outlines the steps taken during the analysis, the structure of the dataset, and my learnings throughout the project and coursework. It demonstrates how Excel can be used effectively to handle and interpret large-scale social data in a meaningful way.

.

List of Tables and Figures

Chapter 1 - Introduction

Crimes against women have been a persistent and pressing issue in India, drawing the attention of policy makers, researchers, law enforcement agencies, and civil society. With increasing awareness and evolving legal frameworks, the need for robust data analysis to understand the nature, distribution, and trends of such crimes has become more critical than ever. The dataset titled "Crimes Against Women in India (2001–2014)" serves as a vital resource in this context, offering granular, district-level information on various categories of crimes committed against women over a span of 14 years.

This dataset, compiled from official sources, covers reported incidents across all Indian states and union territories from the year 2001 to 2014. Each record in the dataset contains information about the state, district, year, and the number of reported cases under specific crime categories. These include rape, kidnapping and abduction, dowry deaths, assault on women with intent to outrage her modesty, cruelty by husband or relatives, insult to modesty, and importation of girls. By capturing data at the district level, the dataset allows for a micro-level examination of regional variations in crime prevalence, while the temporal coverage enables the study of changes and emerging patterns over time.

One of the key strengths of this dataset lies in its ability to highlight not only the magnitude of crimes but also the geographical disparities in how crimes against women are reported and recorded. Some regions show consistently high numbers, while others may reveal sudden spikes or downward trends, possibly reflecting social, legal, or administrative developments during the time period. This makes the dataset highly useful for developing evidence-based insights, assessing the effectiveness of policies and interventions, and identifying priority areas for action.

This project utilizes Microsoft Excel to analyze and visualize the data through dashboards, charts, and tables. By leveraging Excel’s powerful features—such as pivot tables, slicers, conditional formatting, and visualization complex data can be transformed into interactive and insightful dashboards. These dashboards make it easier to spot patterns, compare crime categories across regions, and draw meaningful conclusions that can support further research or public policy formulation.

Overall, this dataset not only sheds light on the grim realities of gender-based violence in India but also provides a solid foundation for data-driven decision-making. It reinforces the importance of ongoing monitoring, reporting, and targeted strategies in creating a safer environment for women across all parts of the country.

Chapter 2 - Objectives

This project on Crimes Against Women Analysis in India (2001–2014) explores the district-wise and state-wise crime statistics reported across India over a 14-year period. The dataset includes multiple categories of gender-based violence and enables detailed insights into regional and temporal patterns. The project makes use of Excel dashboards to visualize and analyze trends, enabling better understanding, awareness, and data-driven decision-making. The key objectives of this project are:

- To display the total number of reported crimes against women and their distribution across different states, districts, and years.
- To highlight crime-specific trends such as incidents of rape, dowry deaths, domestic violence, abduction, and others, and analyze how they vary geographically and over time.
- To compare the performance of states and districts in terms of crime rates, identifying regions with high or low incidence of crimes against women.
- To present a year-wise trend analysis, showcasing how reported crimes have evolved over the years and detecting any significant fluctuations or patterns.
- To visualize the most prevalent crime categories in each state and district, allowing for focused analysis of dominant issues in different areas.
- To identify potential hotspots and areas of concern, aiding authorities and researchers in targeting intervention efforts more effectively.
- To provide an interactive dashboard that supports comparative analysis, enabling users to filter and explore data by year, state, and crime type.
- To support awareness and policy-making efforts by offering clear, data-backed insights into the challenges faced by women in different parts of India.
- This project combines data analysis and visualization to turn raw crime statistics into meaningful insights that can inform public discourse and governmental responses.

Chapter 3 - Source of Dataset

The dataset is taken from Kaggle, a popular platform for data science and machine learning. It provides access to various datasets uploaded by individuals and organizations for public use.

I have selected a Border Crossing dataset that includes detailed information on crossing rates across the United States, categorized by factors such as gender, mode of transport, age, latitudes/longitudes, and year.

Here are the details of my chosen dataset:

Name -Crime Against Women(2001-2014)

Link - https://www.kaggle.com/datasets/greeshmagirish/crime-against-women-20012014-india

Format – CSV

Data Fields:

- State: The state or union territory of India where the crime occurred.
- District: The specific district within the state.
- Year: The year in which the crime was reported (ranging from 2001 to 2014).
- Crime Type: The type of crime (e.g., Rape, Kidnapping, Domestic Violence, etc.).
- Total Cases: The total number of cases reported for that particular crime type in that year and state/district.
- Crimes Under IPC (Indian Penal Code): Number of crimes reported under specific sections of the Indian Penal Code.
- Rate of Crime: The crime rate for that specific crime type, per 100,000 women (can sometimes be calculated based on population).

Chapter 4 - Dataset Preprocessing

In this project, the dataset preprocessing was an important step to ensure accurate and meaningful analysis. The dataset was downloaded from the Kaggle. The raw dataset contained information about border crossing crossing counts in the United States categorized by gender, race, transportation mode, and year.

The following preprocessing steps were performed:

1. Data Cleaning:

- Removed unnecessary columns that were not required for analysis.
- Identified and handled missing values in important columns.
- Replaced blank cells or missing entries with appropriate values where applicable.

2. Data Transformation:

- Renamed column headers for easy understanding and readability.
- Organized data into a clean and structured format suitable for analysis.
- Filtered relevant records and removed inconsistent data entries.

3. Creating Clean Dataset Sheet:

- After cleaning and transforming the data, a new sheet named "Clean Dataset" was created within the Excel workbook.
- The following screenshots represent:

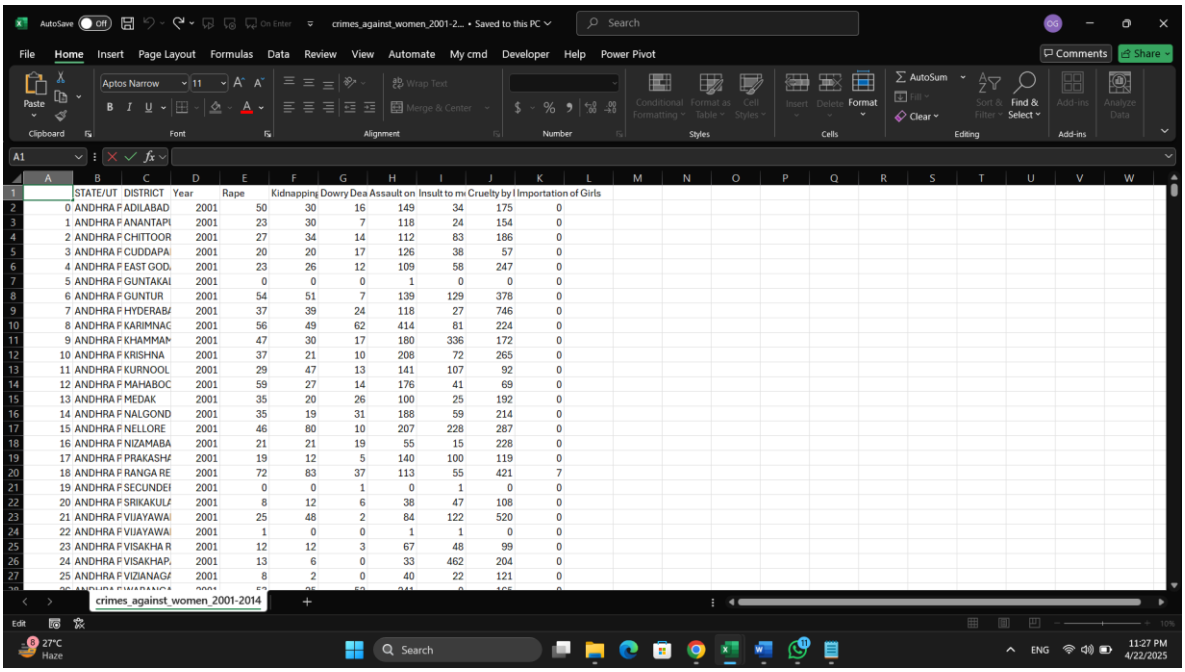


Figure1: The original / raw dataset

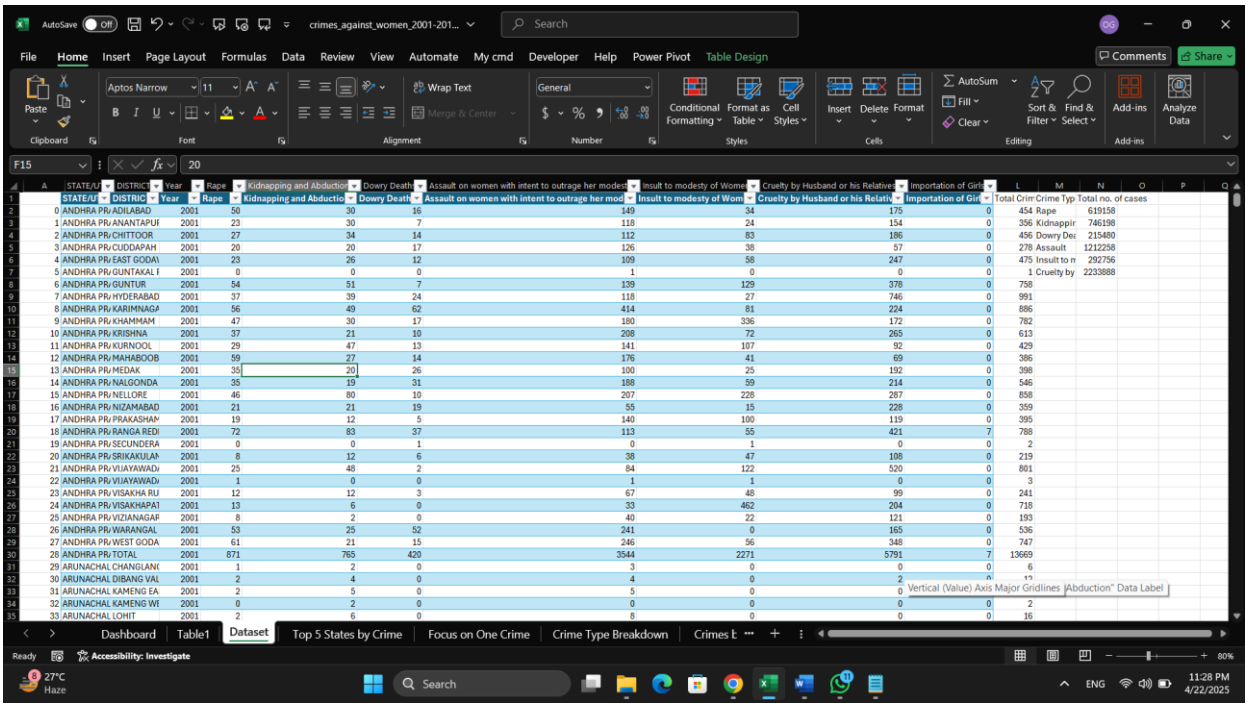
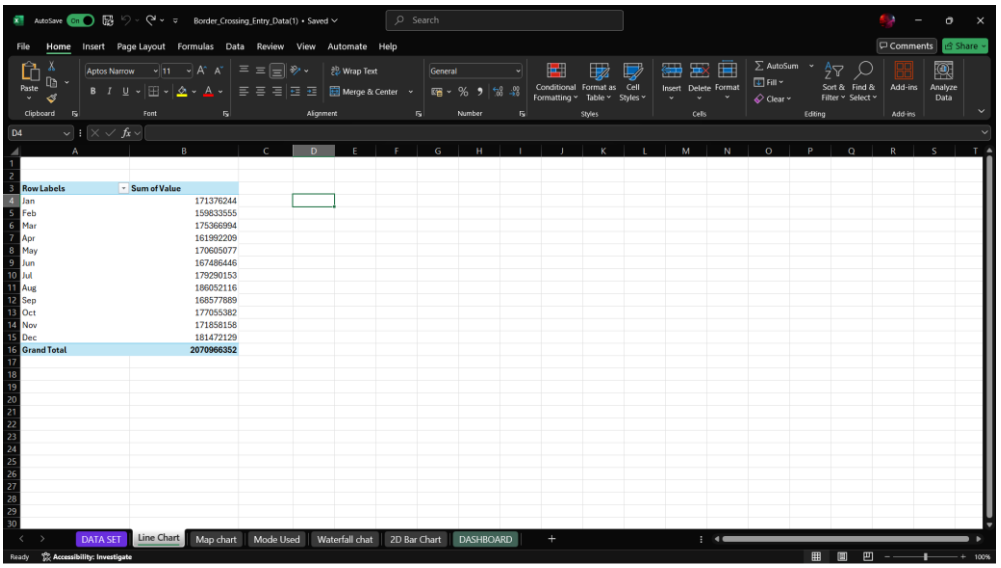


Figure 2:

- This clean dataset sheet was used as the primary source for performing analysis and creating visualizations in the dashboard.



Row Labels	Sum of Value
Jan	171376244
Feb	159833555
Mar	175366994
Apr	161992209
May	170605077
Jun	167486446
Jul	179290153
Aug	186052116
Sep	168577889
Oct	177055382
Nov	171858158
Dec	181472129
Grand Total	2070966352

Figure 3:

- Used sorting and filtering techniques to analyze specific data points.

This preprocessing phase ensured that the dataset was accurate, consistent, and ready for further analysis using Excel tools like Pivot Tables, Charts, Slicers, and Conditional Formatting.

Chapter 5 - Analysis of dataset

General Description :

The dataset provides detailed information about crimes against women across India from 2001 to 2014. It covers various aspects such as the type of crime, geographical distribution, victim demographics (age group and gender), and year-wise data. This comprehensive dataset allows for the observation of patterns and analysis of trends in crime statistics over different time periods and across various states and districts.

This project focuses on analyzing how crime rates vary between different groups of people. The data helps in identifying which states have higher crime rates, comparing the rates between male and female victims (where applicable), identifying the most vulnerable age groups, and examining the differences in crime statistics among various regions and states. The dataset also enables the study of changing crime patterns over the years, highlighting any increases or decreases based on several factors. The clean dataset was created to ensure that the data was accurate, consistent, and structured properly for effective analysis. Various data analysis tools were used to generate insights and represent them visually through dynamic dashboards and charts.

This analysis provides meaningful information that can help researchers, policymakers, and law enforcement agencies understand the trends in crimes against women in India and make data-driven decisions to address this serious issue. The clean dataset was used to generate insights and create various visualizations to support better decision-making and policy formulation.

- Specific Requirements :

The specific requirements of this project were to analyze the Crime Against Women (2001-2014) India dataset and provide meaningful insights using Excel. The project aimed to leverage the capabilities of Excel to transform raw data into understandable and visually engaging insights through a dashboard. The analysis and dashboard creation were focused on enabling easier decision-making and presenting key data in an effective way.

The main requirements included:

- Cleaning and transforming the raw dataset for accurate and meaningful analysis.
- Creating a separate clean dataset sheet within the Excel workbook for structured analysis and easy accessibility.
- Generating pivot tables to analyze crime statistics state-wise, district-wise, year-wise, and crime-type-wise.
- Designing various charts like bar charts, line charts, pie charts, and heatmaps to visualize different aspects of the data (e.g., crime trends, comparison of crime rates between states, etc.).
- Incorporating slicers to filter the dataset interactively based on parameters like year, state, crime type, and age group.
- Highlighting key performance indicators (KPIs) like states with the highest and lowest crime rates, average crime statistics, and year-over-year crime changes.
- Ensuring the dashboard is dynamic, user-friendly, and provides a comprehensive overview of the data.
- Utilizing conditional formatting to visually emphasize significant values, trends, and anomalies in the data.

These requirements were fulfilled using Excel tools and functions, ensuring an effective analysis of the dataset and the creation of a dynamic and insightful dashboard. This dashboard helps users better understand trends in crimes against women in India and facilitates data-driven decision-making for researchers, policymakers, and law enforcement agencies.

- Analysis results:

  Based on the analysis of the Crime Against Women (2001-2014) India dataset using Excel, the following key results and insights were obtained:

  - State-wise analysis revealed that certain states had significantly higher crime rates against women, particularly in categories such as rape and domestic violence. This highlights the need for state-specific interventions and policies to address the particular challenges faced in these regions.

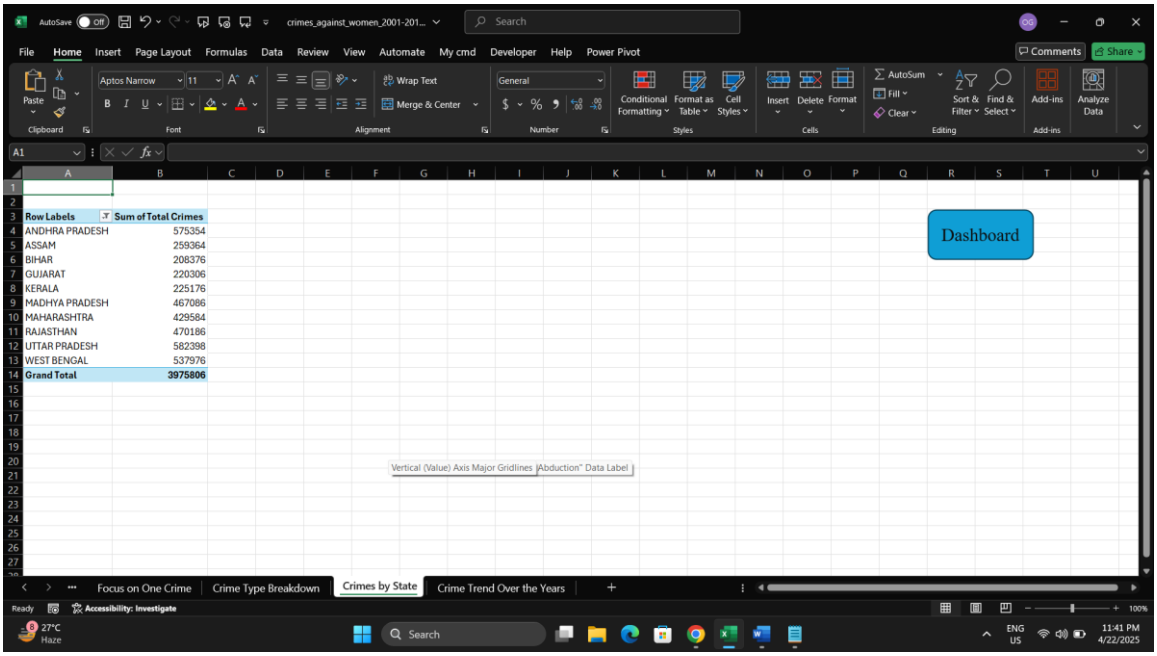
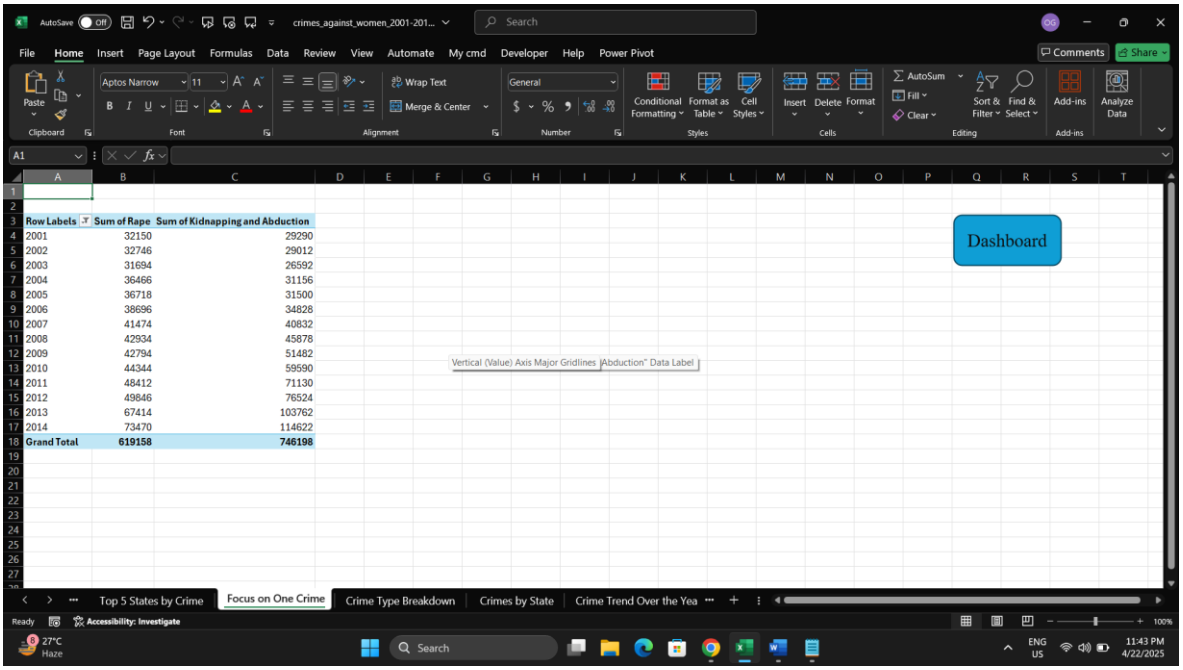


Figure 5:

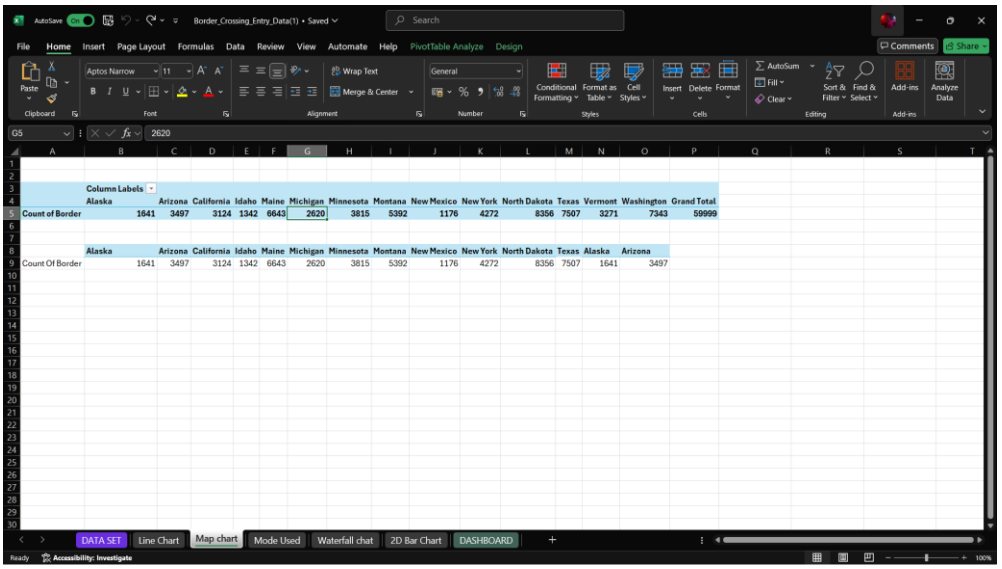
- Crime type comparison showed that rape consistently had the highest number of reported cases across most years and states, underlining the importance of targeted measures for prevention and support for victims.



Row Labels	Sum of Rape	Sum of Kidnapping and Abduction
2001	32150	29290
2002	32746	29012
2003	31694	26592
2004	36466	31156
2005	36718	31500
2006	38696	34828
2007	41474	40832
2008	42934	45878
2009	42794	51482
2010	44344	59590
2011	48412	71130
2012	49846	76524
2013	67414	103762
2014	73470	114622
Grand Total	619158	746198

Figure 6:

- Year-wise trends revealed a gradual increase in crime rates over the years, which may indicate either a rise in actual crime or better reporting and awareness. Further investigation is required to understand the underlying causes.



Column Labels	Alaska	Arizona	California	Idaho	Maine	Michigan	Minnesota	Montana	New Mexico	New York	North Dakota	Texas	Vermont	Washington	Grand Total
Count of Border	1641	3497	3124	1342	6643	2620	3815	5392	1176	4272	8356	7507	3271	7343	59999

	Alaska	Arizona	California	Idaho	Maine	Michigan	Minnesota	Montana	New Mexico	New York	North Dakota	Texas	Alaska	Arizona
Count Of Border	1641	3497	3124	1342	6643	2620	3815	5392	1176	4272	8356	7507	1641	3497

Figure 7:

- The use of slicers and KPIs in the dashboard allowed users to interactively filter the data and focus on specific aspects such as crime type, state/UT, and year, enhancing the overall analytical experience. This interactivity empowers users to quickly identify crime trends, compare states, and analyze the rise or fall of specific crimes over time, thereby facilitating quick and informed decision-making based on different segments of the dataset.

These results offer a clear understanding of crime patterns against women across India from 2001 to 2014, enabling data-driven decision-making for awareness campaigns, resource allocation, and policy formulation. The analysis highlights the effectiveness of Excel's powerful tools—such as PivotTables, Pivot Charts, slicers, and KPIs—in handling multi-year crime data and transforming it into meaningful, actionable insights. This approach not only aids in identifying high-risk regions and crime trends but also supports practical interventions aimed at reducing crimes against women.

.

- Visualization:

  This section presents the visualizations developed during the analysis of the Crimes Against Women dataset using Excel. These interactive charts and dashboards offer a clear and comprehensive view of crime trends across states and years, enabling better understanding and informed decision-making. The visual representation of the data enhances pattern recognition and allows users to explore key insights with ease.

  The following visualizations were created:

  - **Bar Chart:** Comparison of total crimes against women across different States/UTs.

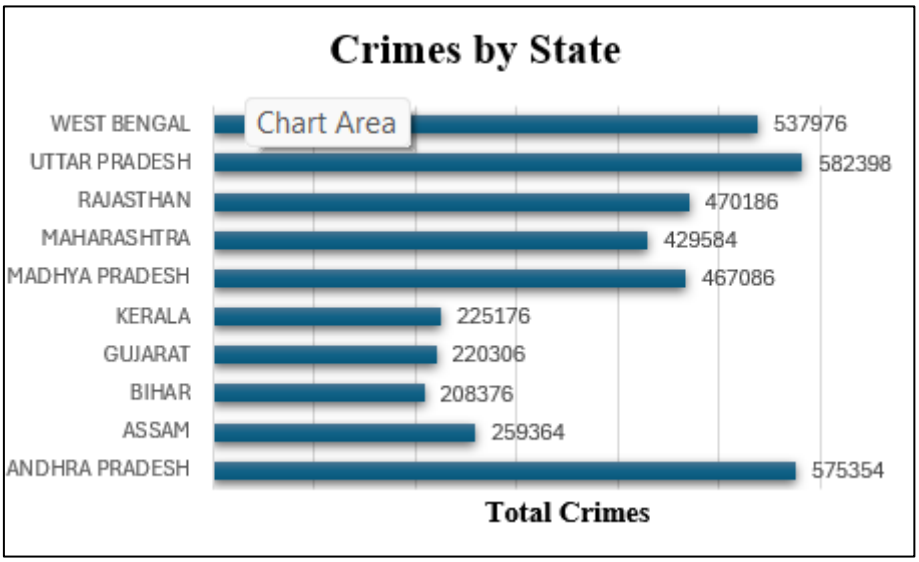


Figure 8:

Pie Chart: Distribution of crimes based on different crime types against women.

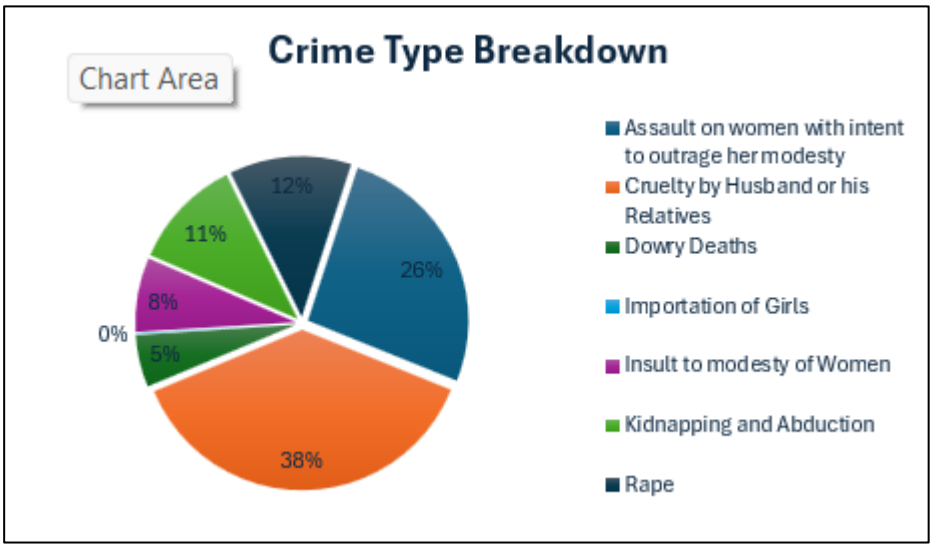


Figure 9:

- **Line Chart:** Year-wise trend analysis of total crimes against women from 2001 to 2014.

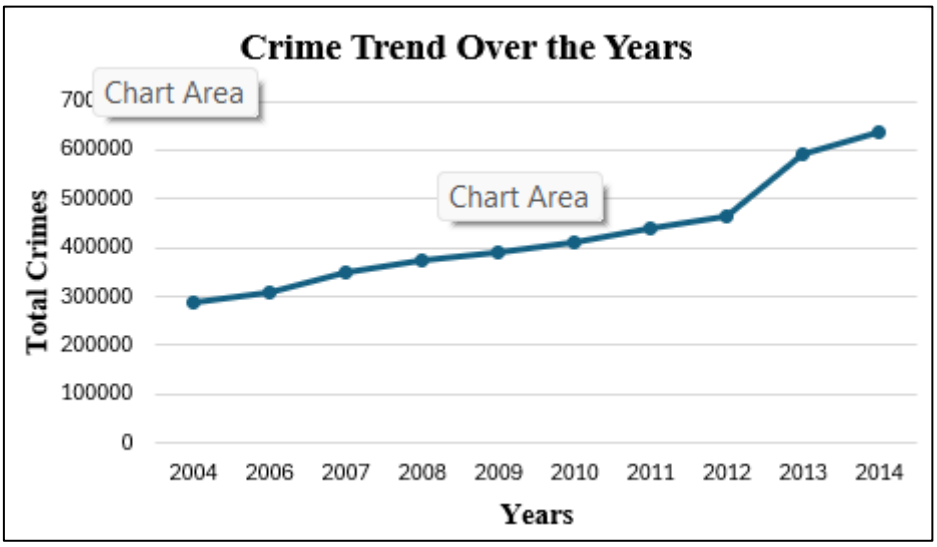


Figure 10:

These charts and visuals helped to represent the key insights obtained from the dataset in a visually appealing and user-friendly manner. They allow users to quickly interpret the results and focus on critical areas for better analysis and reporting.

Chapter 6 - Border Crossing Dashboard

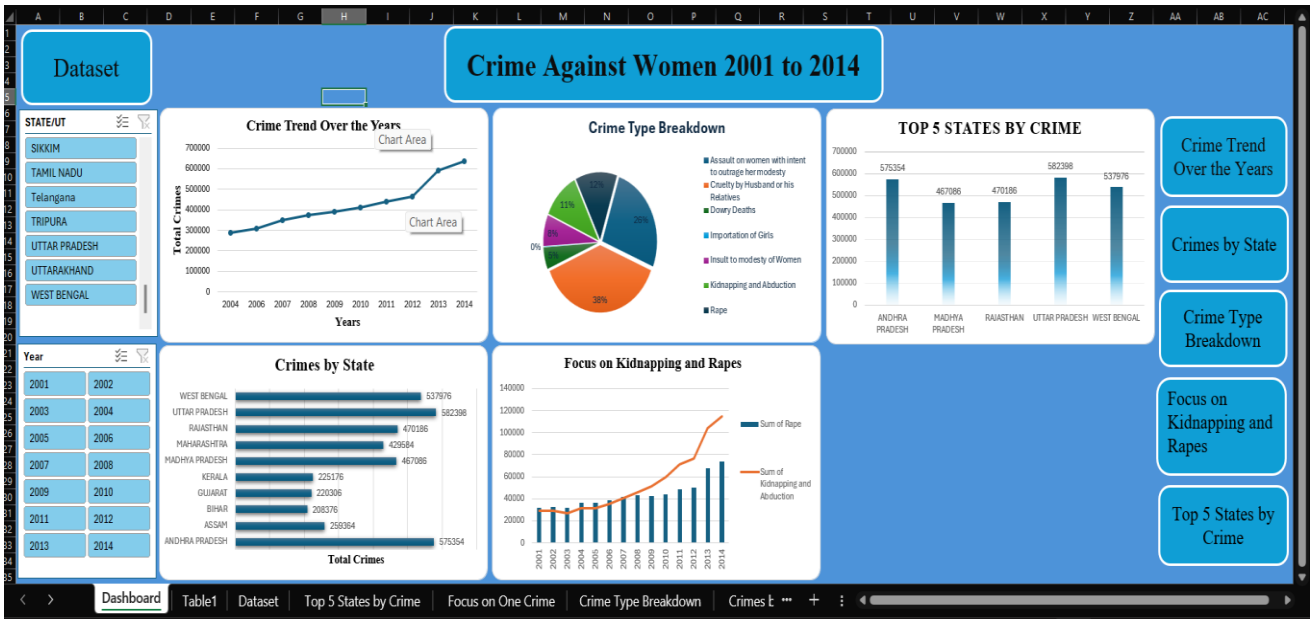


Figure 13:

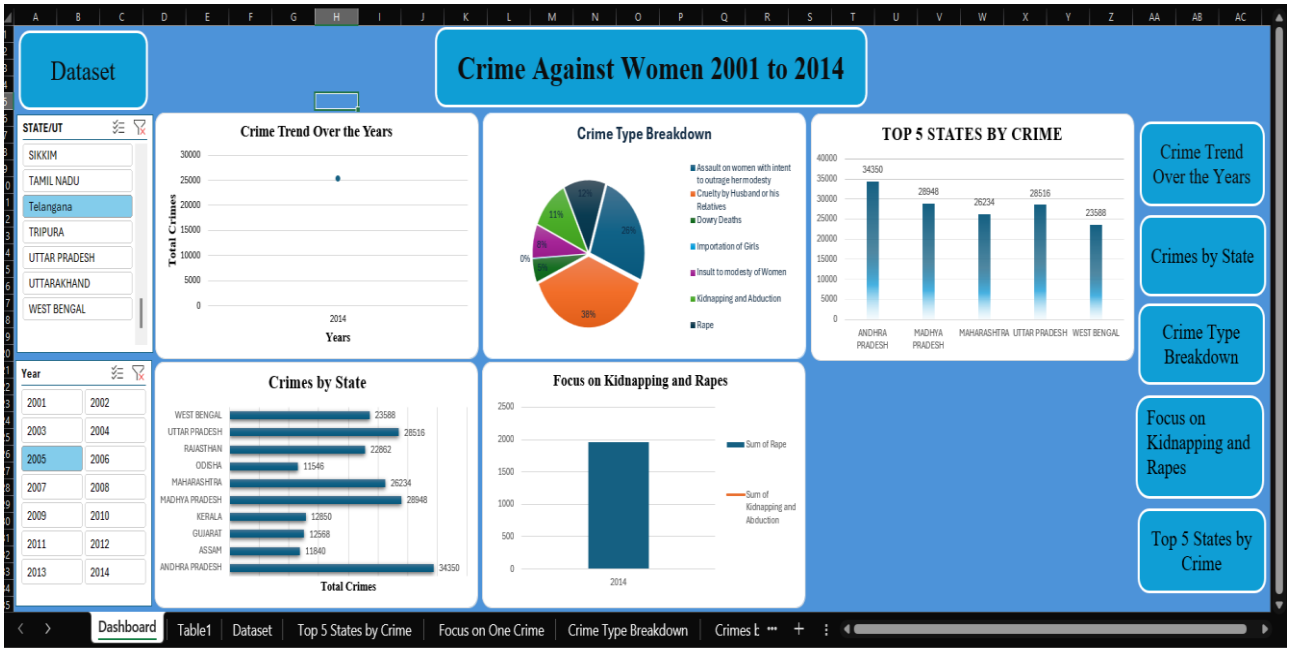


Figure 14: Analyzing data using slicers

Chapter 7 - Conclusion

This project provided valuable insights into the patterns of **crimes against women** across different states and districts in India. Using Excel tools and techniques, I was able to clean, organize, and analyze the dataset efficiently. The visualizations created using pivot tables and charts helped in representing the data in a user-friendly and interactive way.

The analysis revealed important trends and patterns, such as the higher crime rates in certain states and districts, the vulnerability of specific age groups (particularly women aged 19-30), and the variations in crime statistics across different regions. The year-wise trend analysis provided an overview of changes in crime rates over time, helping to understand the effectiveness of past social and law enforcement initiatives.

This project also allowed for a detailed examination of how regional and demographic factors influence crime rates. It was identified that certain states and age groups are more susceptible to higher crime statistics, particularly in urban areas. These findings are essential for the development of targeted intervention strategies, improved law enforcement policies, and greater awareness of women's safety issues.

Overall, this project enhanced my skills in data cleaning, analysis, and visualization using Excel. It demonstrated how large datasets can be transformed into meaningful insights that can aid in decision-making and the development of public awareness campaigns addressing crimes against women. It also helped me understand the importance of using data to tackle real-life social issues and the value of presenting data visually to communicate insights effectively.

Working on this project provided me with practical experience in handling real-world data, performing exploratory data analysis, and generating an interactive dashboard to support data-driven decisions. This project not only strengthened my technical skills but also highlighted the societal impact of data analysis in addressing crucial social issues like crimes against women.

Chapter 8 - Future scope

This project has successfully demonstrated the ability to use Excel as a powerful tool for data cleaning, analysis, and visualization. However, there is significant scope for future enhancements and further research to expand the effectiveness and applicability of this project to a wider audience and real-world use cases.

In the future, the dataset can be expanded to include more recent data, covering years beyond 2014, for a more up-to-date analysis. Additional parameters such as the education level, employment status, socio-economic background, law enforcement resources, access to support services, and regional cultural factors could be incorporated. Including these parameters would allow for a more comprehensive and multi-dimensional analysis of the factors contributing to crimes against women across different regions.

Furthermore, advanced data visualization tools like Power BI, Tableau, or Python libraries such as Matplotlib and Seaborn can be integrated for creating more dynamic and visually appealing dashboards. These tools can offer better interactivity, deeper analysis, and real-time data visualization capabilities. Implementing these advanced tools would make the analysis more scalable and user-friendly for a broader audience, including law enforcement agencies, policymakers, and social welfare organizations.

Additionally, machine learning techniques and predictive modeling could be applied to forecast future trends in crimes against women based on historical data patterns. Predictive analytics could help identify high-risk regions or communities, enabling policymakers to implement targeted prevention strategies before issues escalate.

Another important future enhancement could involve integrating real-time data from verified sources such as law enforcement agencies, government databases, and NGOs working in the women's safety domain. This would ensure that the data remains up-to-date, accurate, and relevant for continuous monitoring and proactive analysis.

Collaboration with government bodies, women's advocacy groups, and law enforcement agencies can further enhance the utility of this project. These partnerships can assist in developing targeted awareness campaigns, intervention programs, and preventive strategies based on the insights derived from the data analysis.

Chapter 9 - References

Trends in Crimes Against Women in India (2001–2014)

Kaggle [Online]. Available: [Accessed: Apr. 07, 2025].

Lovely Professional University, "Course Material - INT 217 Introduction to Data Management," School of Computer Science and Engineering, Phagwara, Punjab, 2025.

LinkedIn Link : https://www.linkedin.com/posts/om-gupta03_datavisualization-womensafety-dashboard-activity-7320855964927016961-wbRW?utm_source=share&utm_medium=member_desktop&rcm=ACoAAEU1RU4BW27HtTUbtT7Yr9kqDr73KnD-NzI

GitHub Link :