Data Analysis on Cybercrime

Assignment report

Roll Number	Student Name
20201ISE0054	Thejas V
20201ISE0085	Shreyas S
20201ISE0055	Vikram

Objective / Problem Statement

The increasing complexity and volume of cyber threats pose significant challenges to cybersecurity professionals. Traditional methods of analyzing security data often fall short in detecting and responding to evolving threats in real-time. This project addresses the need for effective data visualization tools to enhance threat intelligence, improve situational awareness, and facilitate proactive defense measures against cyber attacks.

Dataset

Our team will be utilizing the dataset provided from “Open Government Data (OGD) Platform India” - Cyber Crime in Indian Cities (2018 & 2019)

Link to the same datasets - https://data.gov.in/catalog/

Data Overview

2018 Cyber Crime Stats:

- Cities included: All major cities of India
- Categories: Personal Revenge, Anger, Fraud, Extortion, etc.
- Metrics: Total crimes and crime rates

2019 Cyber Crime Stats:

- Similar structure to 2018, allowing for year-over-year comparison

Methodology

Data Collection: Gather data from reliable sources, ensuring accuracy and completeness.

Data Cleaning: Handle missing values, correct errors, and ensure data consistency.

Data Analysis: Perform statistical analysis and visualize data for insights.

Reporting: Summarize findings and provide actionable recommendations.

Data Cleaning and Preparation

- Handling Missing Values: Impute or remove missing data to maintain dataset integrity.
- Data Normalization: Ensure consistent formatting and units across datasets.
- Column Standardization: Align column names and data types for seamless analysis.

Visualization Insights

1. Crime Distribution by Category
Visualization: Pie charts or bar charts showing the proportion of each crime category.
Insight: Identify the most prevalent types of cyber crimes.

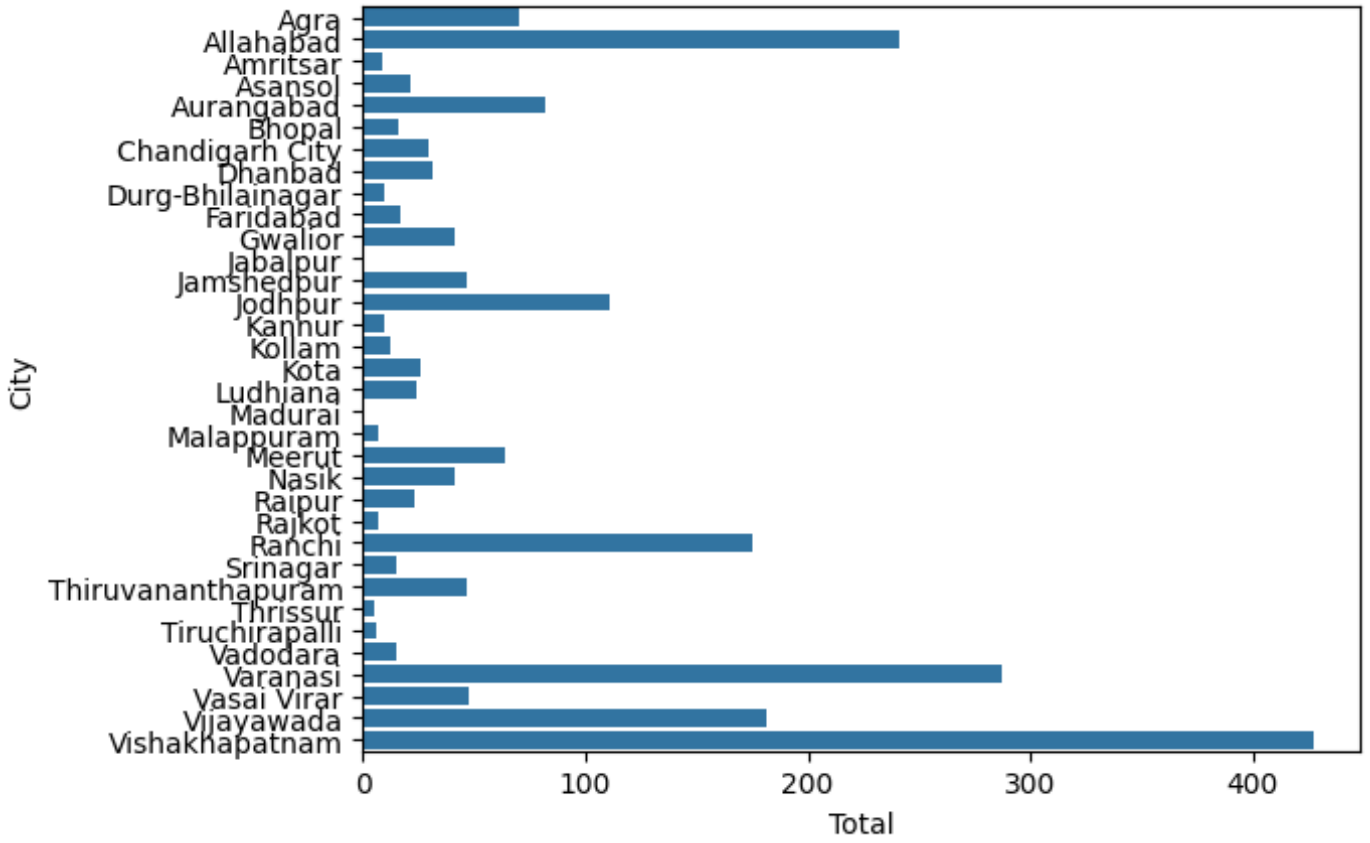
2. Statewise Crime Analysis
Visualization: Choropleth maps highlighting crime intensity in each state.
Insight: Pinpoint states with highest and lowest cyber crime rates.

3. Trends Over Time
Visualization: Line graphs comparing crime rates from 2018 to 2019.
Insight: Identify increasing or decreasing trends in specific crime categories.

4. Correlation Analysis
Visualization: Heatmaps displaying correlations between different crime categories.
Insight: Discover relationships and patterns between various types of cyber crimes.

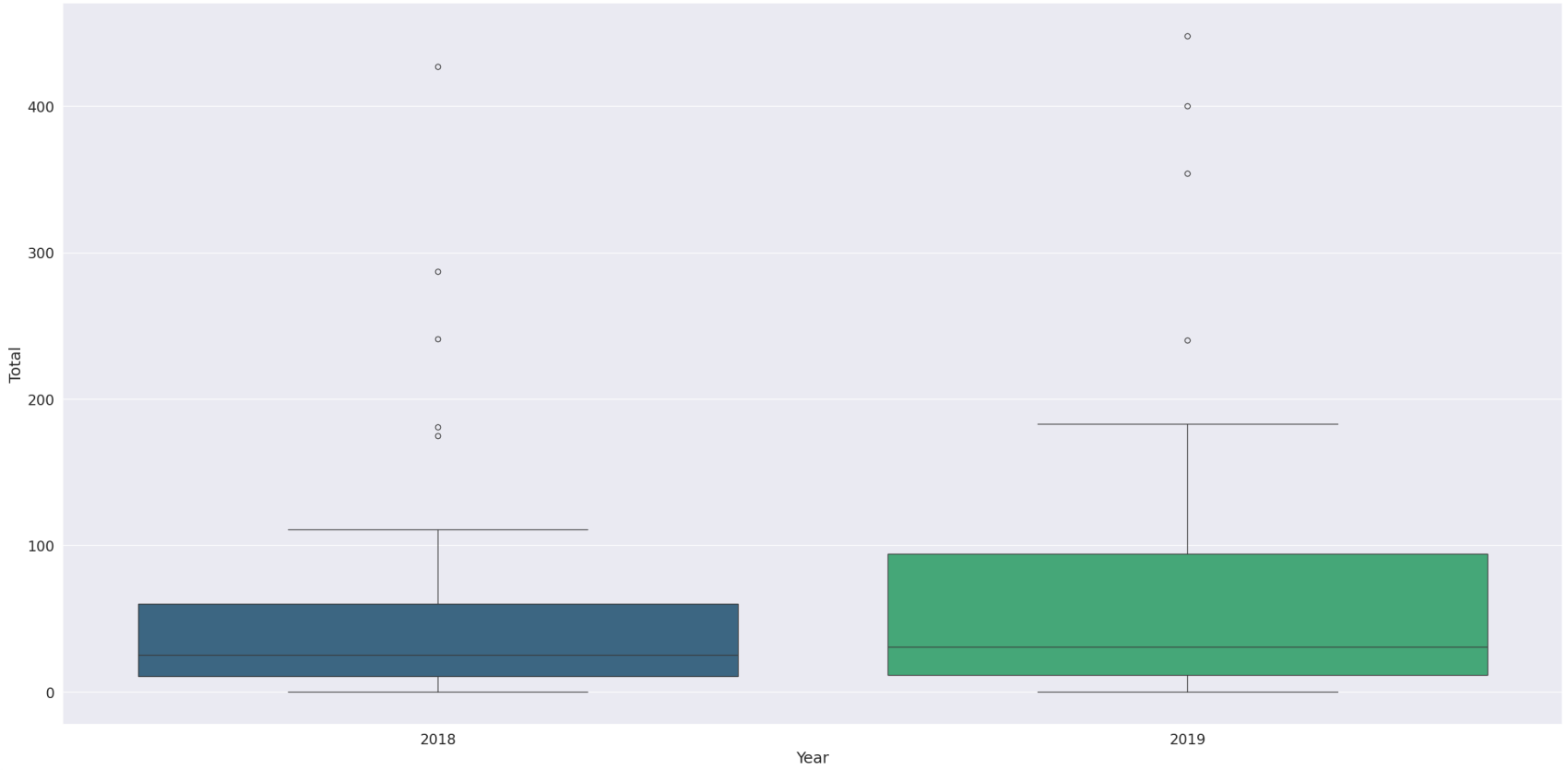
Important Data visualizations

1. State wise horizontal bar plot of crime rates



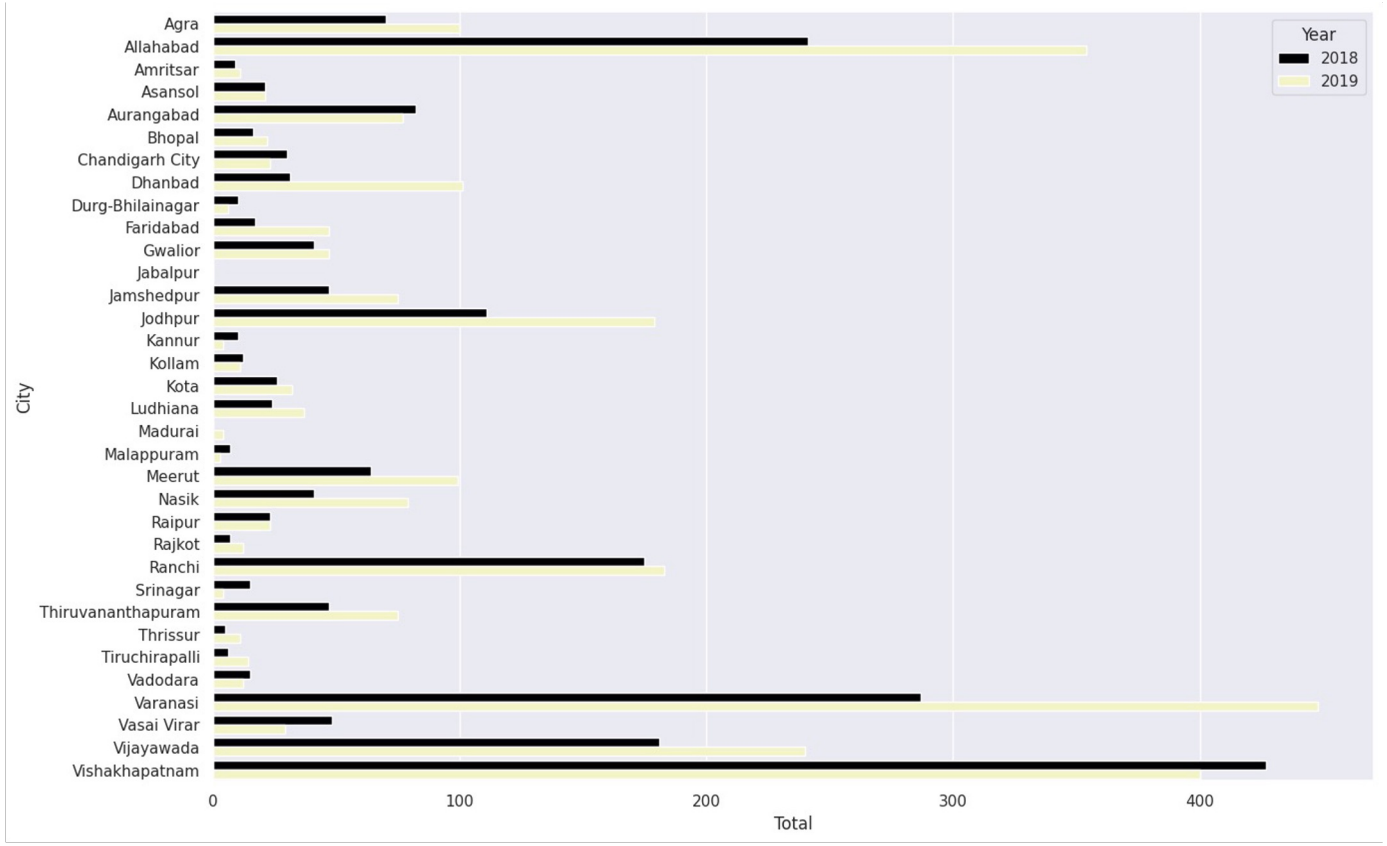
Important Data visualizations

2. Box Plot to visualize total crime rates compared by year



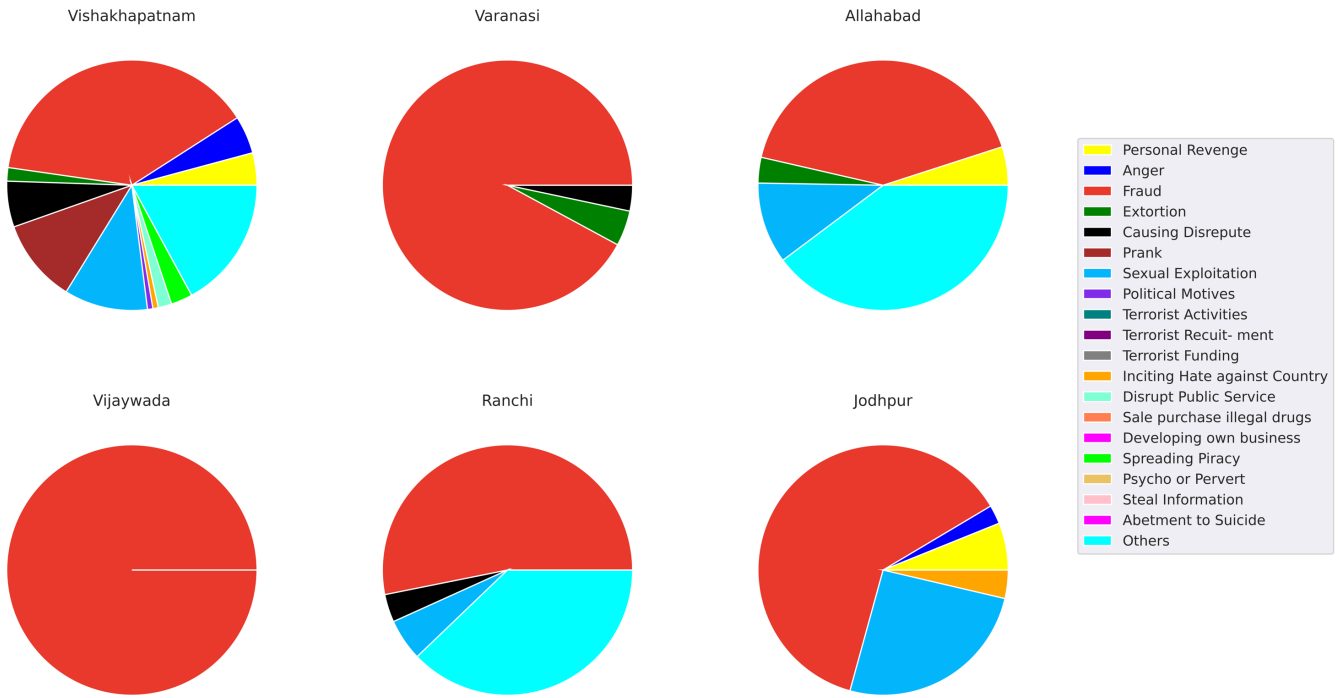
Important Data visualizations

3. Bar plot comparing individual cities over compared by years



Important Data visualizations

3. Pie charts visualizing motivation of crime compared by cities



Summary

- The data visualizations reveal several insights into the nature of cybercrime in India.
- Firstly, the bar charts indicate that the total number of cybercrimes has increased significantly from 2018 to 2019, with cities like Vishakhapatnam, Varanasi, and Allahabad experiencing the highest number of cases.
- Secondly, the pie charts illustrate the distribution of cybercrime motivations, with personal revenge, anger, and fraud being the most prevalent motives. Thirdly, the boxplots highlight the variations in crime rates and total cases across different years, showcasing a general upward trend.
- Lastly, the comparison of top cities based on crime motivations demonstrates the diversity in the reasons behind cybercrimes in different regions. These insights can be valuable for law enforcement agencies and policymakers in understanding the evolving trends and patterns of cybercrime in India.

Conclusion

The analysis of cyber crime data from 2018 and 2019 reveals significant trends and insights into the state of cyber crime in India. Notably, fraud and extortion emerged as the most prevalent categories, indicating a need for stronger preventive measures. The data also highlighted certain states with disproportionately high crime rates, suggesting targeted regional interventions are necessary. Over the two years, while some crime categories showed a decline, others, such as sexual exploitation and terrorism-related activities, unfortunately exhibited an upward trend. These findings underscore the importance of continuous monitoring and adaptive strategies to combat the evolving landscape of cyber crime effectively.

Thank You