Team no. 20

Shrushranto Rajbongshi (524110026)

Suraj Pandey (524110023)

Priyanshu Mudgal (524110044)

Nishant Singh (524110035)

## **Aim**

The aim of this project is to study crime data from India to find patterns and trends in crimes involving foreigners (both tourists and residents), cybercrimes, and IPC (Indian Penal Code) crimes from 2018 to 2021. The goal is to better understand how these crimes have changed over time, identify any new trends, and use the data to suggest better ways to prevent these crimes.

## **Why We Chose This Topic**

Crime analysis is crucial for understanding social patterns, improving law enforcement strategies, and enhancing public safety. We chose this topic for the following reasons:

1. **Increasing\_Crimes\_Against\_Foreigners**  
   India is a popular tourist destination and home to many expatriates. However, crimes involving foreigners, including fraud, theft, and assault, have raised concerns. Understanding crime patterns can help improve safety measures and boost international confidence in India's security.
2. **Rising\_Cybercrimes**  
   With digital transformation, cybercrimes such as financial fraud, identity theft, and online harassment have surged. Analyzing trends in cybercrime can help law enforcement agencies develop better strategies to combat digital threats.
3. **Understanding IPC Crimes Over Time**  
   IPC crimes, including theft, assault, and fraud, impact millions of citizens. By studying how these crimes have evolved from 2018 to 2021, we can identify patterns that may be linked to economic shifts, social changes, or law enforcement policies.
4. **Data-Driven Crime Prevention**  
   By merging datasets from multiple years and analyzing trends, we aim to provide insights that can help policymakers, law enforcement agencies, and the public take proactive steps in crime prevention.
5. **Bridging the Research Gap**  
   While general crime statistics are available, a focused study on these specific crime categories can highlight areas that need more attention and resources for better public safety.

This research will not only provide valuable insights into crime trends but also contribute to developing data-driven solutions to enhance security in India.

## **Data Collection**

The dataset is collected from the Indian Government’s official data portal (<https://www.data.gov.in/moreresult/crime%20headwise%20police%20disposal%20cyber%20crime%20cases?domain=data.gov.in&r=%2Fsearch%2F%3Freturn%3Dtrue%26title%3Dcrime>.

## **Data Description**

### Police Disposal of IPC Crime Cases and Cyber Crime

|  |  |
| --- | --- |
| Category (Col.2) | Represents the broad classification of the crime :   * Offences Affecting the Human Body * Offences against the State * Offences against Public Tranquillity * Offences against Property * Offences Relating to Documents & Property Marks * Miscellaneous IPC Crimes * Offences under I.T. Act * Offences under IPC * Offences under SLL |
| Crime Head (Col.3) | The specific type of crime committed. |
| Cases Pending Investigation from Previous Year (Col.4) | Cases carried over from last year that were not closed. |
| Cases Reported During the Year (Col. 5) | Number of new cases reported in the current year. |
| Cases Reopened for Investigation(Col. 6) | Cases that were previously closed but reopened for further investigation. |
| Total Cases for Investigation (Col.7 = Col.4 + Col.5 + Col.6) | Total cases police need to investigate, including previous year, new, and reopened cases.  Formula: **Total Cases = Pending Cases + Reported Cases + Reopened Cases** |
| Cases Not Investigated Under 157(1)(b) CRPC(Col. 8) | Cases where police decide not to investigate due to lack of evidence or legal grounds. If the police find that a reported cybercrime lacks merit, it is recorded here. |
| Cases Transferred to Other State or Agency(Col. 9) | This refers to cybercrime cases that were originally registered in one state or jurisdiction but later transferred to another state or investigative agency for further investigation. Due to   1. Jurisdictional Issues: 2. Central or Specialized Agencies Required: 3. Lack of Resources or Expertise: 4. Legal or Procedural Reasons:   In the **IPC (Indian Penal Code) Disposal Dataset**, the column **"Cases Transferred"** refers to **criminal cases that were originally registered under the IPC in one jurisdiction but later transferred to another state or agency for further investigation or legal proceedings**. |
| Cases Withdrawn by Government During Investi gation (Col 10) | Cases dropped by the government before investigation is completed. |
| **11. Final Report - Cases Ended as FR Non-Cognizable** | Cases deemed **non-cognizable** (minor offences). |
| 12. Final Report - Cases Ended as Final Report False | Cases found to be false complaints after investigation. |
| **13. Final Report - Cases Ended as Mistake of Fact or of Law or Civil Dispute** | Cases dismissed due to **legal misinterpretation**. |
| **14. Final Report - Cases True but Insufficient Evidence or Untraced or No Clue** | Crimes that happened but **lack evidence** to proceed. |
| **15. Final Report - Cases Abated During Investigation** | Cases that **cannot continue** due to accused/suspect being deceased or untraceable. |
| Final Report - Total | (Col.16 = Col.11 + Col.12 + Col.13 + Col.14 + Col.15)   * Sum of all cases closed under the above categories. |
| Chargesheets Submitted (Cases Sent for Prosecution) (  For Column 17-19) | These columns track the number of cases where police found enough evidence to file a chargesheet.  **17. Chargesheets Submitted - Out of Cases from Previous Year**  **18. Chargesheets Submitted - Out of Cases During the Year**  **19. Total Chargesheets Submitted (Col.19 = Col.17 + Col.18)** |
| Total Cases Disposed Off by Police | (Col.20 = Col.8 + Col.9 + Col.16 + Col.19)  Sum of all cases that reached an outcome (whether closed, transferred, or chargesheeted). |
| 21-22. Cases Quashed or Stayed at Investigation Stage | Cases stopped due to legal intervention.  **21. Cases Quashed at Investigation Stage**  **22. Cases Stayed at Investigation Stage** |
| Cases Pending Investigation at End of the Year | (Col.23 = Col.7 - Col.10 - Col.20 - Col.21)   * Cases still open and awaiting resolution. * Pending Cases = Total Cases - Withdrawn - Disposed - Quashed |

Crime against Foreigners

|  |  |
| --- | --- |
| Category (Col 2) | Classifies crimes into two broad types based on the laws under which they are defined –  **A. IPC (Indian Penal Code)**.  **B. SLL (Special and Local Laws)**. |
| Crime Head (Col 3) | The type of crime committed against foreigners.   * **Murder** – Intentional killing of a person. * **Rape** – Sexual assault. * **Kidnapping** **and** **Abduction** – Forceful taking of a person. * **Theft** – Stealing. * **Cheating** – Fraudulent activity. * **Forgery** – Falsification of documents. |
| Cases of Crimes Committed against - Foreign Tourists (Col.4) | The number of reported cases where the victim was a foreign tourist.  ( Min = 0 Max = 240) |
| Cases of Crimes Committed against - Other Foreigners (Col.5) | The number of reported cases where the victim was a foreigner who is not a tourist (e.g.students, business travelers).  (Min = 0 Max = 20) |
| Cases of Crimes Committed against - Total Foreigners (Col.6=Col.4 + Col.5) | The sum of cases against foreign tourists and other foreigners (Col.6 = Col.4 + Col.5).  (Min = 0 Max = 259) |
| Total Rows: | The dataset includes three summary rows:   * **Total (A)** – Sum of all IPC crimes. * **Total (B)** – Sum of all SLL crimes. * **Total (A+B)** – Total of all crimes against foreigners. |

**A. IPC (Indian Penal Code)**.

This category includes general criminal offenses that are covered under the **Indian Penal Code (IPC)**, which is the main criminal code for India. It deals with a wide range of crimes such as **murder**, **rape**, **theft**, and **assault**. These are common crimes and follow the general criminal laws of the country.

* **Examples**: Murder, Kidnapping, Theft, Robbery.

**B. SLL (Special and Local Laws)**.

This category includes crimes that fall under specific or special laws, which are not part of the IPC. These laws are created to address particular issues or situations and are often more focused on particular types of crimes.

* **Examples**: Crimes related to **Immoral Traffic (Prevention) Act**, **Protection of Women from Domestic Violence Act**, or other specialized laws that apply to certain crimes that are not covered under the IPC.

## **Data Cleaning**

1. **Standardization:**

* Ensure consistency across all data fields and units.

1. **Handling Null Values:**

* Identify and deal with missing or null values using appropriate methods (imputation, removal, etc.).

1. **Removing Unnecessary Columns:**

* Eliminate columns not relevant to the analysis to optimize dataset size and focus on key variables.

1. **Data Merging**

* Combine datasets from 2018 to 2021 to create a unified dataset.
* Ensure proper alignment of column names and data formats before merging.
* Handle duplicate entries and inconsistencies that may arise due to merging.

## **Technologies To be Use**

1. **Python:**

* **Libraries**:

1. **Pandas**: For data manipulation, cleaning, and analysis.
2. **Matplotlib/Seaborn**: For data visualization (trend analysis, distributions, comparisons).
3. **SciPy/Statsmodels**: For statistical analysis and hypothesis testing.
4. **Jupyter Notebook/Google Colab:**

* For interactive data analysis and visualizations

## **Objectives**

### Crime Against Foreign:

### Trends Over time from 2018 to 2021:

* **Total Cases of Crimes Committed against - Foreign Tourists (Col.4)** with **Total** **Cases of Crimes Committed against - Other Foreigners (Col.5)**.
* **Cases of Crimes Committed against - Total Foreigners (Col.6=Col.4 + Col.5)** In over the mentioned years**.**
* Which Crime head has the highest rate in three years?

#### Comparative Analysis

Compare Crime heads rate of **foreign tourist** with **foreign**

* Whether tourists are more vulnerable to specific crimes than foreign residents.
* Which crime types are more common among each group.
* Whether different crime prevention strategies are needed for tourists vs. foreign residents**.**

The objective of this analysis is to understand crime trends involving foreign tourists and other foreigners from 2018 to 2021. Overall, the goal is to figure out if crime patterns differ between tourists and foreign residents and how to address these issues more effectively.

## Cyber Crime:

* What are the trends in the number of reported cybercrime cases from 2018 to 2021?
* Which cybercrime categories show the highest number of cases reported during the years 2018-2021?
* What is the percentage of cybercrime cases that remain pending at the end of the year compared to the total number of cases for investigation?
* What is the disposition rate (cases disposed off) for cybercrimes across the years?
* What are the reasons for the closure of cybercrime cases (e.g., cases ended as false, insufficient evidence, mistake of fact)?

The main goal of these questions is to examine trends and patterns in reported cybercrime cases. Overall, the objective is to understand the trends in cybercrime investigations, the outcomes, and how they measure up to other types of crimes in terms of resolution and process efficiency.

## IPC Crime:

* How have the number of IPC crime cases reported and investigated evolved over the years?
* Which IPC crime categories (e.g., theft, assault, murder) show the highest number of cases reported and pending investigations?
* What is the chargesheeting rate for IPC crimes from 2018 to 2021?
* What are the percentages of cases withdrawn by the government during the investigation process?
* How many IPC cases were closed due to insufficient evidence or mistakes during investigation?
* How effective were the final report outcomes in closing IPC cases (e.g., cases ended as non-cognizable, mistake of fact, abated)?

The objective of these questions is to analyze trends and developments in reported IPC crime cases . Overall, the goal is to understand the trends in IPC crime investigations, the reasons behind case closures, and how effectively these cases are handled over time.

## Analysis for Both Datasets Combined (Cybercrime + IPC Crime)

* How do the trends in cybercrime cases reported compare to those in IPC crimes from 2018-2021?
* Which types of crimes (cyber or IPC) show a higher rate of pending investigations at the end of the year?
* What is the chargesheeting rate for cybercrime vs IPC crime cases over the years?
* Which type of crime (cybercrime or IPC crime) has a higher rate of cases transferred to other agencies or states?
* What is the correlation between the increase in cybercrimes and the decrease in IPC crime rates over the years?
* How does the number of cybercrimes and IPC crimes resolved due to insufficient evidence or mistake of fact compare?

The goal is to understand how cybercrimes and IPC crimes have changed over time, how they are investigated and solved, and whether there are any important differences or patterns between the two types of crimes in terms of trends, investigation results, and case closures.

**Timeline:**

|  |  |  |
| --- | --- | --- |
| **Phase** | **Task** | **Timeline** |
| Phase 1: Planning | Define objectives, gather datasets, outline key questions | Week 1-2 |
| Phase 2: Data Preparation | Clean and preprocess dataset for analysis | Week 3-4 |
| Phase 3: Analysis | Develop dashboards and visualizations | Week 5-8 |
| Phase 4: Insights | Interpret findings, document insights, refine visualizations | Week 8-10 |
| Phase 5: Review | Present results and finalize project | Week 10-12 |

**Challenges & Mitigation Strategies:**

* **Challenge:** Incomplete or inconsistent data records.
  + *Mitigation:* Perform data cleaning and imputation techniques.
* **Challenge:** Correlating crime rates with external socio-economic factors.
  + *Mitigation:* Use publicly available government datasets for cross-referencing.
* **Challenge:** Ensuring ethical handling of crime data.
  + *Mitigation:* Follow ethical guidelines and anonymize sensitive data.

**Tools:**

* **Crime Dataset:** Official law enforcement crime records and open-source datasets.
* **Excel & Power BI:** Data cleaning, visualization, and trend analysis.
* **Python & Pandas:** Advanced data processing and statistical analysis.

## Expected Deliverables

1. **Comprehensive Crime Analysis Report:**

* Insights from trends and patterns in crimes against foreigners, cybercrimes, and IPC crimes.
* Comparative analysis between foreign tourists and other foreigners regarding their vulnerability to crimes.
* Detailed examination of cybercrime and IPC case investigation outcomes.
* Visualizations of trends, case distribution, and crime rates over time.
* Data-driven insights on crime prevention strategies targeted at foreign tourists, foreign residents, and cybercrime mitigation.

## Conclusion

This project will closely examine crime trends in India, focusing on crimes involving foreigners, cybercrimes, and IPC offenses from 2018 to 2021. Using Python to analyze the data will help us better understand these crimes, identify new patterns, and suggest ways to improve how law enforcement handles them. The project will also compare cybercrimes and IPC crimes to see if there are any connections or changes in trends between the two over the years.

## **Contributions-**

* **Suraj Pandey** – Managing and analysing Foreign Crime data set
* **Priyanshu Mudgal** – Managing and analysing IPC crime dataset
* **Shrushranto Rajbongshi** – Managing and analysing Cyber crime dataset
* **Nishant Singh** – Planning, Researching and finding datasets, Data cleaning, presentation and report making