Crime Analysis in India

Team no. 20

			
Shrushranto Rajbongshi 524110026	Suraj Pandey 524110023	Priyanshu Mudgal 524110044	Nishant Singh 524110035

Introduction

In recent years, the landscape of crime in India has evolved significantly, with emerging challenges posed by crimes involving foreigners, cybercrimes, and violations under the Indian Penal Code (IPC). This report delves into comprehensive analysis spanning the years 2018 to 2021, aiming to uncover underlying patterns and trends in these distinct categories of criminal activities.

The primary objective of this project is to scrutinize crime data to discern temporal shifts, identify emerging trends, and ascertain the prevalence and nature of crimes involving foreigners—both tourists and residents—alongside cybercrimes and IPC offenses. By examining these datasets, we endeavor to gain deeper insights into the dynamics of these crimes, thereby facilitating informed strategies for prevention and mitigation.

Through rigorous data analysis and statistical methods, this report seeks not only to highlight the evolving nature of these crimes but also to propose actionable recommendations aimed at enhancing law enforcement strategies and bolstering public safety measures.

Dataset and Tool:

. 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. Which is official open data platform of the Government of India. It provides access to a wide range of datasets published by various government departments and organizations. The platform aims to promote transparency, innovation, and data-driven decision-making by making government data freely available to the public, researchers, and developers.

The dataset consist of following attributes:

Crime against Foreigners:

- **Category –** Classifies crimes into two broad types : **IPC, SLL**
- **Crime head -** The type of crime committed against foreigners.
- **Cases of Crimes Committed against Foreign Tourist-** The number of reported cases where the victim was a foreign tourist
- **Cases of Crimes Committed against Other Foreigners -** The number of reported cases where the victim was a foreigner who is not a tourist (e.g.students, business travelers).
- **Cases of Crimes Committed against Total Foreigners -** The sum of cases against foreign tourists and other foreigners

IPC Crimes –

- **Year –** The calendar year the data refers to.
- **Category –** Broad classification of the crime type.
- **Crime Head –** Specific nature or type of crime.
- **Cases Pending Investigation from Previous Year –** Unresolved cases carried over from the prior year.
- **Cases Reported during the Year –** New crime cases reported in the current year.
- **Cases Reopened for Investigation –** Closed cases that were reopened for investigation.
- **Total Cases for Investigation –** Total caseload for the year (pending + new + reopened).
- **Cases Not Investigated under 157(1)(b) CRPC –** Cases dropped early due to insufficient grounds.
- **Cases Transferred to other State or Agency –** Cases moved to another jurisdiction or agency.
- **Cases Withdrawn by the Govt during Investigation –** Government-initiated withdrawal before completion.
- **Final Report - Cases Ended as FR Non-Cognizable –** Cases closed as minor (non-cognizable) offences.
- **Final Report - Cases Ended as Final Report False –** Cases proven to be false complaints.
- **Final Report - Cases Ended as Mistake of Fact or of Law or Civil Dispute** – Dismissed due to legal or factual error.
- **Final Report - Cases True but Insufficient Evidence or Untraced or No Clue –** Real crimes but lacked evidence or leads.
- **Final Report - Cases Abated during Investigation –** Dropped due to reasons like accused being deceased or untraceable.
- **Final Report - Total –** Total cases closed through all final report categories.
- **Chargesheets Submitted -** Cases Chargesheeted Out of Cases from Previous Year – Chargesheets filed from prior year's cases.

- **Chargesheets Submitted -** Cases Chargesheeted Out of Cases during the Year – Chargesheets filed from current year's cases.
- **Chargesheets Submitted -** Cases Chargesheeted – Combined total of chargesheets submitted.
- **Total Cases Disposed Off by Police** – Sum of all closed, transferred, withdrawn, or chargesheeted cases.
- **Cases Quashed at Investigation Stage –** Cases terminated by court during investigation.
- **Cases Stayed at Investigation Stage** – Investigation paused due to legal stay.
- **Cases Pending Investigation at End of the Year** – Unresolved cases carried into the next year.
- ***Chargesheeting Rate 100 –** Percentage of cases chargesheeted out of total investigated.
- ***Pendency Percentage 100 –** Percentage of cases still pending investigation.

Cyber Crime Dataset

- **Year** – Indicates the year of the recorded data.
- **Category** – Represents the broad classification of the crime.
- **Crime Head** – Specifies the exact nature or type of crime.
- **Cases Pending Investigation from Previous Year** – Unresolved cases carried forward from the previous year.
- **Cases Reported during the Year** – New crime cases reported in the current year.
- **Cases Reopened for Investigation** – Cases that were previously closed but have been reopened.
- **Total Cases for Investigation** – Total number of cases under investigation (including pending, reported, and reopened).
- **Cases Not Investigated under 157(1)(b) CRPC** – Cases dropped without investigation under legal provision.
- **Cases Transferred to other State or Agency** – Cases moved to a different jurisdiction or investigation agency.
- **Cases Withdrawn by the Govt during Investigation** – Government-withdrawn cases before investigation concluded.
- **Final Report - Cases Ended as FR Non Cognizable** – Minor offences not requiring immediate police action.
- **Final Report - Cases Ended as Final Report False** – Cases closed after being found false or fabricated.
- **Final Report - Cases Ended as Mistake of Fact or of Law or Civil Dispute** – Closed due to misinterpretation of fact/law or being civil in nature.
- **Final Report - Cases True but Insufficient Evidence or Untraced or No Clue** – Genuine cases but lacking sufficient evidence or leads.
- **Final Report - Cases Abated during Investigation** – Cases stopped due to circumstances like death of accused.

- **Final Report - Total** – Total of all final report categories.
- **Chargesheets Submitted - Out of Cases from Previous Year** – Chargesheets filed from previous year's cases.
- **Chargesheets Submitted - Out of Cases during the Year** – Chargesheets filed from current year's cases.
- **Chargesheets Submitted - Total** – Total number of chargesheets filed.
- **Total Cases Disposed Off by Police** – Sum of all cases finalized (closed, transferred, withdrawn, or chargesheeted).
- **Cases Quashed at Investigation Stage** – Cases dismissed by court during investigation.
- **Cases Stayed at Investigation Stage** – Investigations legally halted or stayed.
- **Cases Pending Investigation at End of the Year** – Cases still unresolved at year's end.
- **Chargesheeting Rate 100* – Percentage of total cases chargesheeted.
- **Pendency Percentage 100* – Percentage of total cases still pending investigation.

Tools-

- **Jupyter Notebook –** For Cleaning and analysis
- **MS Excel –** For cleaning and Studying Data.
- **Python -** Programming tool used for the process
  - **Libraries**:
    i. **Pandas**: For data manipulation, cleaning, and analysis.
    ii. **Matplotlib/Seaborn**: For data visualization (trend analysis, distributions, comparisons).
    iii. **SciPy/Statsmodels**: For statistical analysis and hypothesis testing.
- **Power BI-** For visualisation and analysis

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.

Exploration Objectives -

a) 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.

b) 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?

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.

c) 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.

d) 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.

Methodology

1. Data Collection

Gathered raw crime datasets from the official Indian government portal [data.gov.in], ensuring reliable and authentic data sources.

The data included records of IPC crimes, cyber crimes, and crimes against foreigners, covering multiple years and categories.

2. Data Cleaning and Transformation

Cleaned the datasets by addressing missing values, correcting inconsistencies, and standardizing formats.

Renamed columns for better clarity and transformed the data structure to suit analytical objectives.

3. Data Analysis and Visualization

Employed Python libraries such as Pandas, Seaborn, and Matplotlib for data exploration, analysis, and visualization.

Investigated trends in crime reporting, types of offences, chargesheeting efficiency, and investigation backlogs.

4. **Report Creation**

Created a concise and insightful report outlining the major findings, supported by charts and statistics.Focused on answering key research questions regarding crime dynamics and the performance of law enforcement mechanisms.

5. **Dashboard Development with Power BI**

Built an interactive dashboard in Power BI to visualize and explore crime data dynamically.

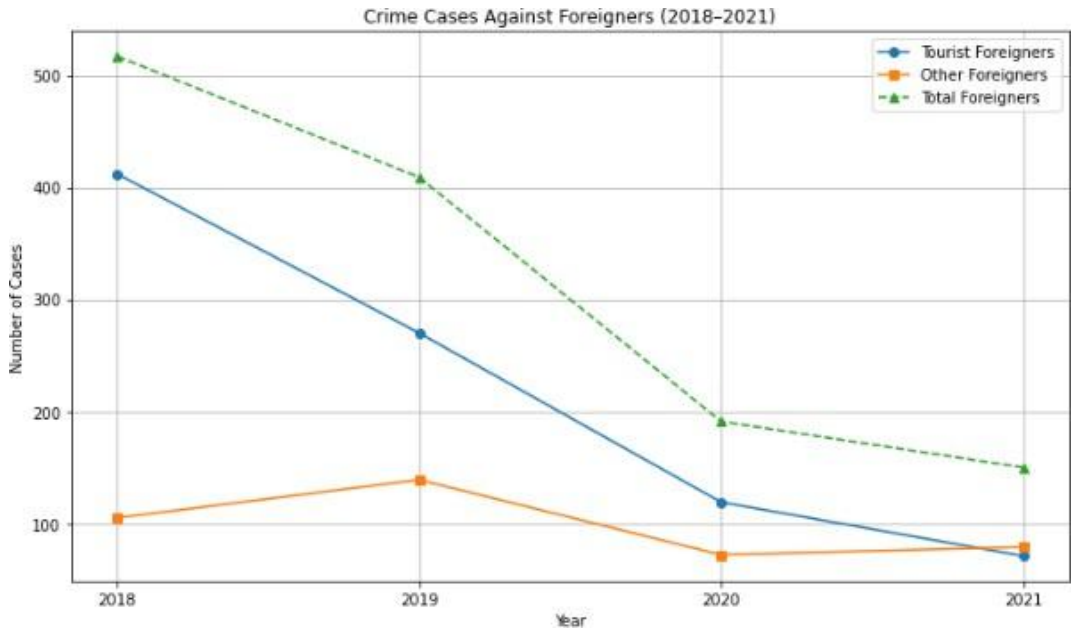
Enabled users to filter insights by year, crime category, and investigation status to support data-driven decisions.

Data Visualisation and Analysis

From the questions above, we will carry out data visualisation to find out answers to these questions as follows:

Analysis of Crime against Foreigners:

Yearly Trends of Crimes Against Foreigners

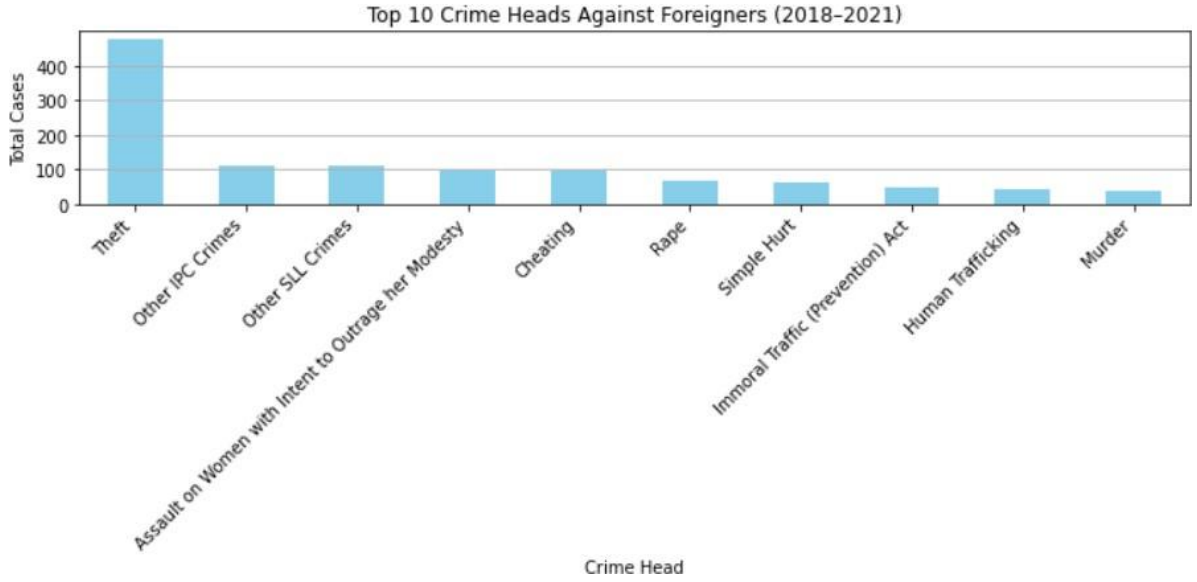


Summary of Trends (2018–2021)

- Crimes against **foreign tourists** gradually declined from 2018 to 2020.
- A **sharp drop** in total crimes was observed in 2020, likely due to the COVID-19 pandemic and global travel restrictions.
- In 2021, cases began to **rise slightly**, indicating recovery in international movement.
- **Other foreigners** (residents, workers) consistently faced **more crime cases** than tourists in all four years.

- Overall, crime trends appear to be **influenced by global and national events**, especially the pandemic. .

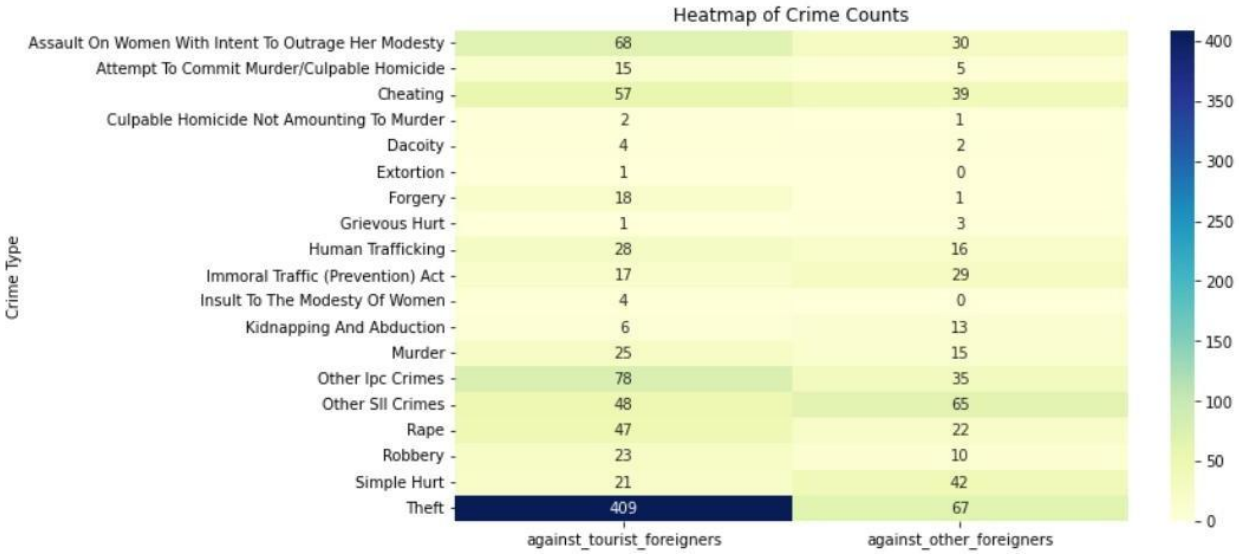
Crime Head with Highest Cases Against Foreigners



Top 10 Crime Heads Against Foreigners (2018–2021)

- The most reported crime head against foreigners was **Theft**, with the highest total cases from 2018 to 2021.
- Other frequently reported crimes include **Other IPC Crimes**, **Other SLL Crimes**, and **Assault on Women with Intent to Outrage her Modesty**, indicating common safety concerns.
- These top 10 crime types account for a significant share of total cases, reflecting areas where foreign nationals are most vulnerable.
- Law enforcement and policymakers may focus on these specific crime types to enhance safety for foreign tourists and residents.

Which crime types are more common among each group.



Insights from Crime Rate Heatmap

1. Foreign Tourists Face Higher Crime Rates in:
    - Molestation, Rape, Outraging Modesty of Women, Theft, Cheating
    - These crimes are often opportunistic or involve personal vulnerability, suggesting that tourists may be targeted due to unfamiliarity with the local environment, lack of local support networks, or language barriers.
2. Foreign Residents Face Higher Crime Rates in:
    - Robbery, Simple Hurt (assault), Grievous Hurt, Forgery, Criminal Breach of Trust
    - These crimes tend to arise in longer-term interactions such as employment, housing, or financial relationships—more applicable to foreign residents who live and work in the country.
3. Balanced or Low Rate Crimes for Both Groups:
    - Burglary, Attempt to Commit Murder, Criminal Trespass
    - These have generally low reported rates across both groups, indicating they are less prevalent or affect both populations similarly.

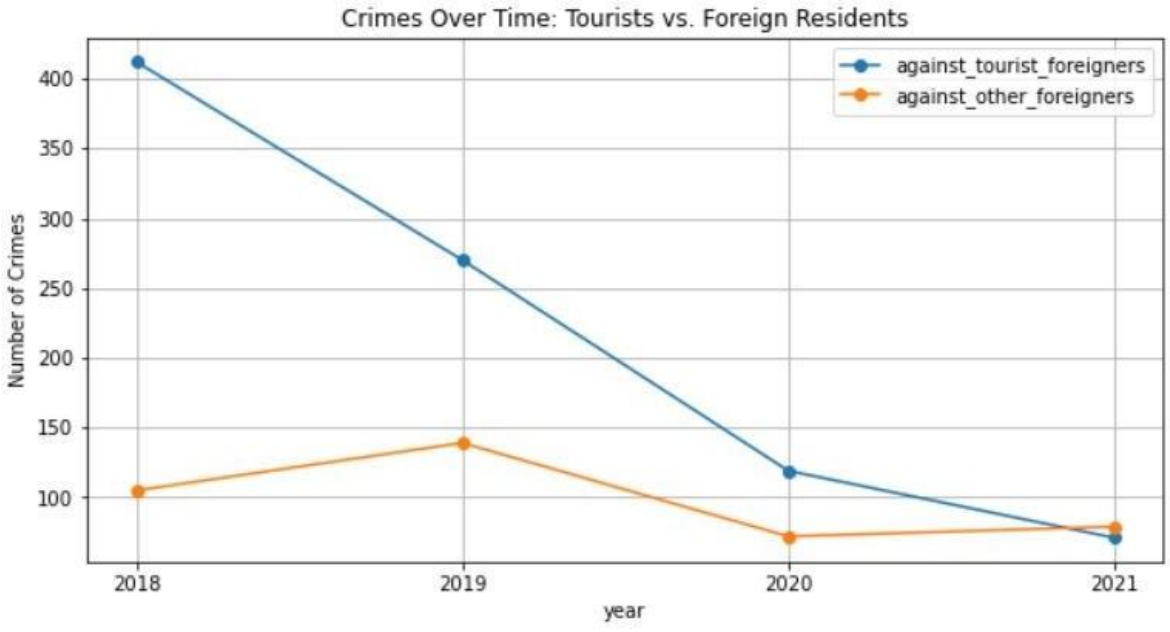
Interpretation & Implications:

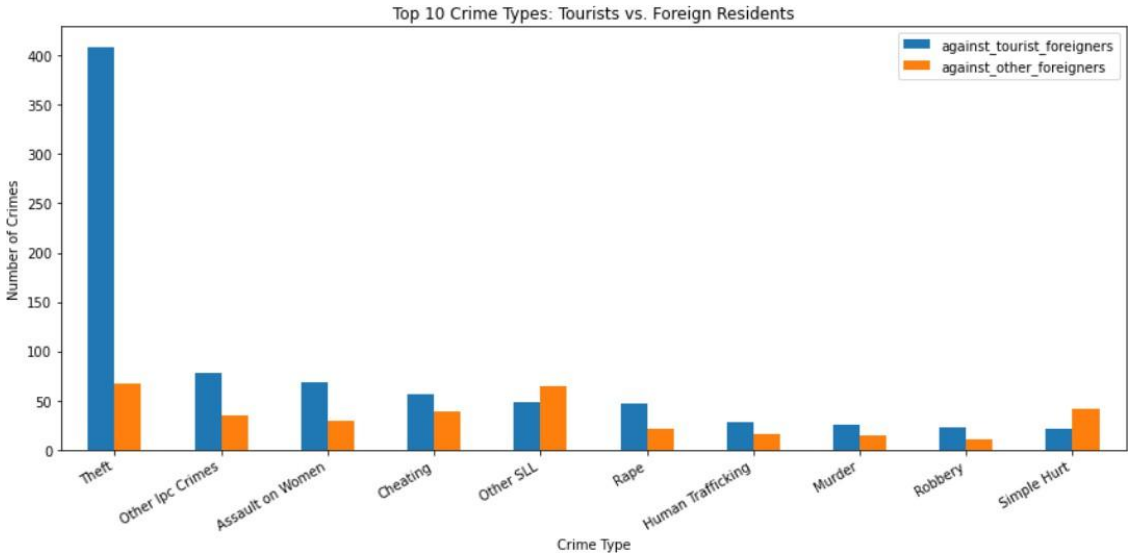
Tourists require protection from street-level and personal crimes, suggesting the need for:

- Increased police visibility in tourist zones
- Tourist helplines and support centres
- Awareness campaigns about scams and harassment

Residents are more exposed to relationship-based or systemic crimes, indicating a need for:

- Legal support systems for foreign workers/residents
- Community policing and housing/employment protections
- Access to justice and mediation for long-term foreigners





Insights -

1. **Overall Crime Volume Differs**

   Tourists experienced significantly more crimes (872 incidents) compared to foreign residents (395 incidents).

   This difference highlights a higher vulnerability or targeting of tourists, warranting more proactive or specialized measures for their protection.

2. **Trends Over Time Differ**

   The line graph (Crimes Over Time) shows a sharp decline in crimes against tourists from 2018 to 2021.

   In contrast, crimes against foreign residents are relatively stable but lower, with a slight decline toward 2021.

   These contrasting trends suggest that strategies effective for one group may not necessarily apply to the other.

3. **Crime Types Show Clear Differences**

   Theft is overwhelmingly the top crime against tourists, far exceeding any other category—over 400 incidents, compared to only about 70 for foreign residents.

   Foreign residents, on the other hand, face relatively higher rates of “Other SLL” crimes and “Simple Hurt”, both of which are nearly equal to or even exceed tourist figures in those categories.

This distinction indicates that tourists are more frequently targeted for opportunistic crimes (like theft), whereas foreign residents are more likely to experience interpersonal or systemic crimes (e.g., simple assault or legal violations).

These differences emphasize the need for context-specific strategies—such as theft prevention and situational awareness for tourists, versus long-term community engagement and protection measures for foreign residents.

The data clearly demonstrates that tourists and foreign residents face different crime patterns, both in volume and nature. Tourists are disproportionately affected by theft and opportunistic crimes, often driven by their transient presence and unfamiliarity with local environments. In contrast, foreign residents are more exposed to interpersonal and systemic crimes, likely influenced by their longer-term integration into the community.

These insights make a compelling case for differentiated crime prevention strategies:

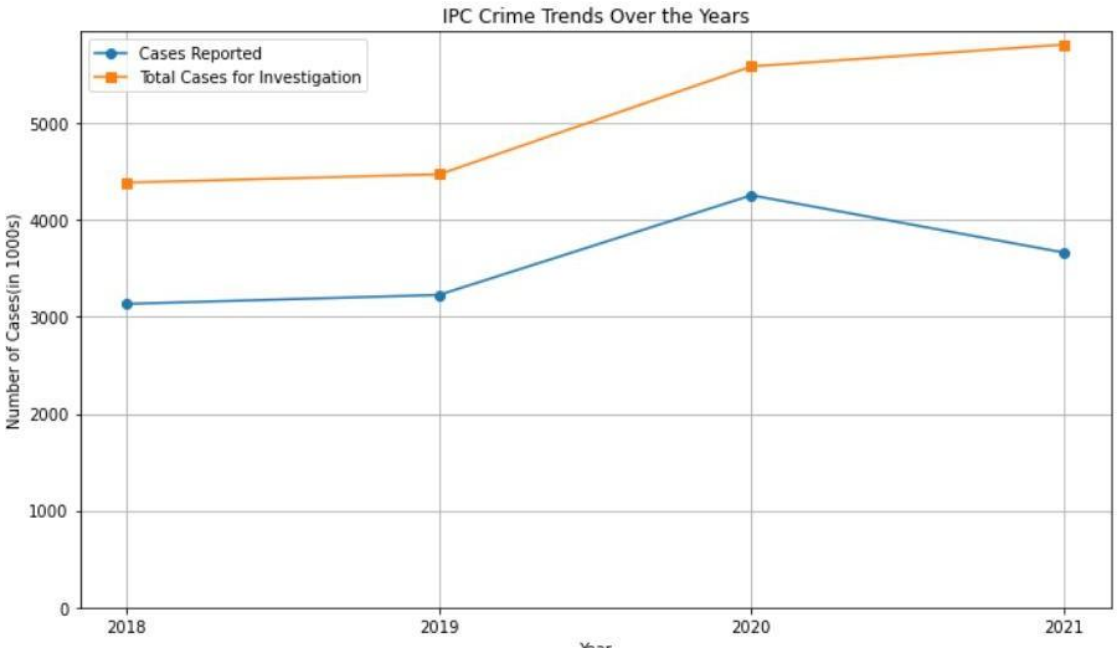
For tourists, focus on awareness campaigns, targeted policing in high-tourist areas, and improved security infrastructure

For foreign residents, strategies should emphasize community policing, legal support services, and mechanisms to report and address interpersonal violence or discrimination.

By tailoring approaches to each group's unique vulnerabilities, authorities can more effectively enhance safety, build trust, and promote a secure environment for all foreigners—temporary or permanent.

Analysis of IPC Crime:

1. How have the number of IPC crime cases reported and investigated evolved over the years?



Insights from the Chart:

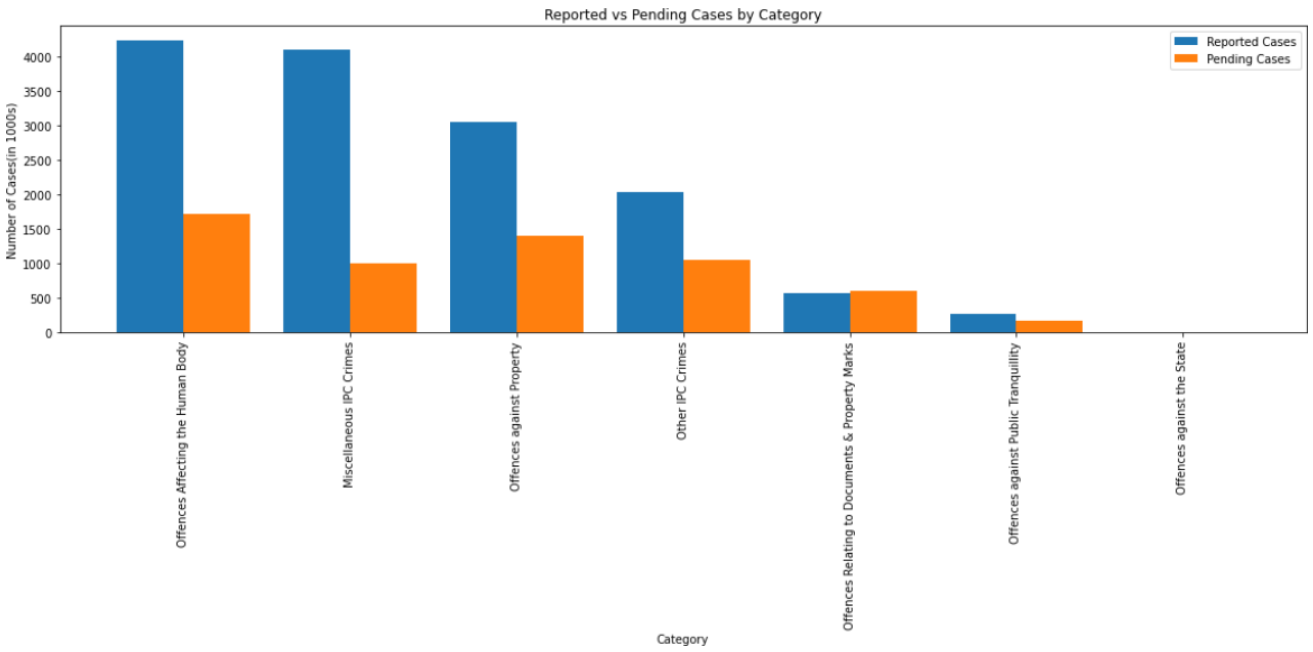
Trend of Crime Reporting:

- If the line for "Cases Reported during the year" shows an upward trend, it suggests an increase in IPC crime cases being reported each year.
- A dip in 2020 might reflect the impact of COVID-19 lockdowns, where public movement was restricted and reporting might have decreased temporarily.

Trend of Investigations:

- The "Total Cases for Investigation" line may follow a similar pattern, but slightly higher, as it also includes cases carried forward from the previous year.
- A growing gap between reported and investigated cases over the years might indicate increasing case backlog.

2. Which IPC crime categories show the highest number of cases reported and pending investigations?



Insights from the Chart:

Category-wise Crime Reporting Trend

- Categories with very high "Cases Reported during the year" (as seen in the top chart) indicate areas where crimes are either more prevalent or more frequently reported.
- This could reflect increased public awareness, improved access to reporting mechanisms, or a genuine rise in criminal activity in those areas.
- The large number of reported cases in certain categories might also be tied to digitally traceable crimes or socially sensitive issues where reporting is more encouraged.

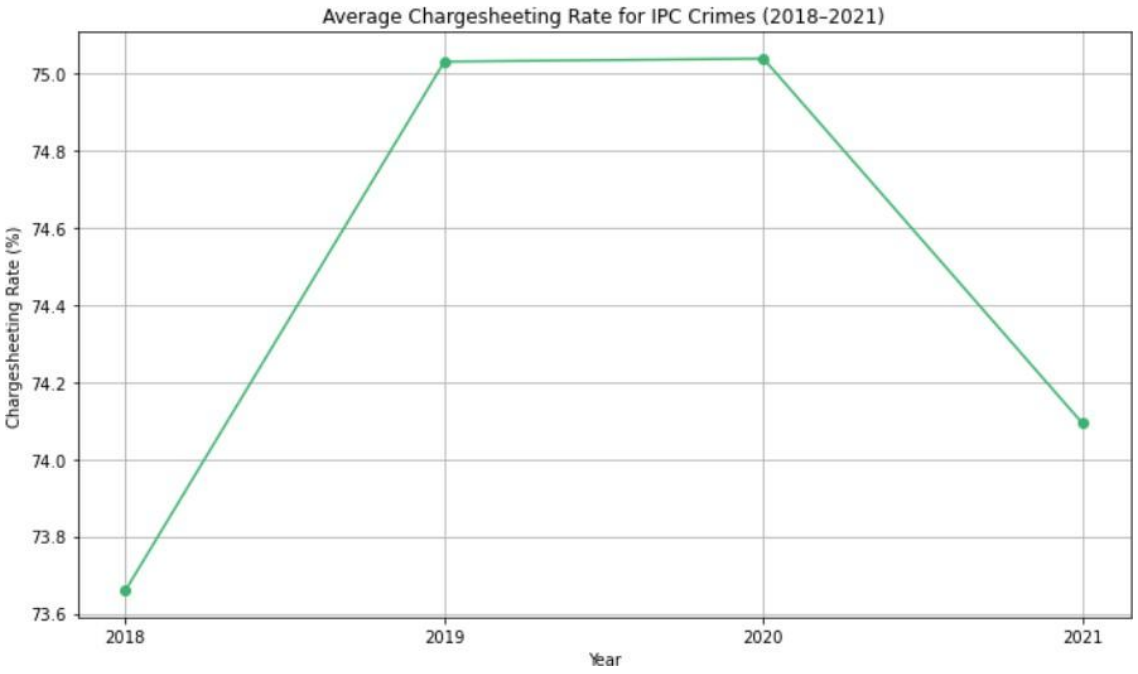
Pending Investigations Reveal System Load

- Categories with high "Cases Pending Investigation from Previous Year" suggest areas where investigative or judicial bottlenecks may exist.
- A large number of pending cases can lead to delayed justice, potentially eroding public trust in the system.
- These categories may require resource allocation, faster case disposal mechanisms, or legal reforms to handle complexity or volume.

Gap Between Reported and Pending Cases

- When a category has high reporting but low pending cases, it indicates efficient follow-up and resolution — possibly a model for other categories.
- Conversely, if pending cases outnumber or closely match reported ones, it could indicate year-over-year accumulation, suggesting an increasing backlog.
- Monitoring this gap helps understand whether the system is keeping up with the pace of crime reporting or falling behind.

3. What is the chargesheeting rate for IPC crimes from 2018 to 2021?



Insights from the Chart:

Chargesheeting Efficiency Over Time

- If the line for “Average Chargesheeting Rate (%)” shows a consistent or upward trend between 2018 and 2021, it suggests that law enforcement agencies have become more efficient in completing investigations and filing chargesheets for IPC crimes.

Possible Impact of COVID-19 (2020)

- A dip in 2020 may indicate the effect of COVID-19 lockdowns, which likely disrupted police investigations, court functioning, and administrative processes, resulting in delayed or fewer chargesheets being filed that year.

Inter-Year Comparison

- A consistent or rising trend from 2018 to 2021 would highlight steady improvements in the justice system's processing capability, possibly due to digitization, legal reforms, or policy-driven performance metrics at state or national levels.

4. What are the percentages of cases withdrawn by the government during the investigation process?

	Year	Total Cases for Investigation	Cases Withdrawn by Govt	% Cases Withdrawn
0	2018	4384597	221	0.0050
1	2019	4470678	53	0.0012
2	2020	5584135	105	0.0019
3	2021	5810088	1747	0.0301

Insights from the table:

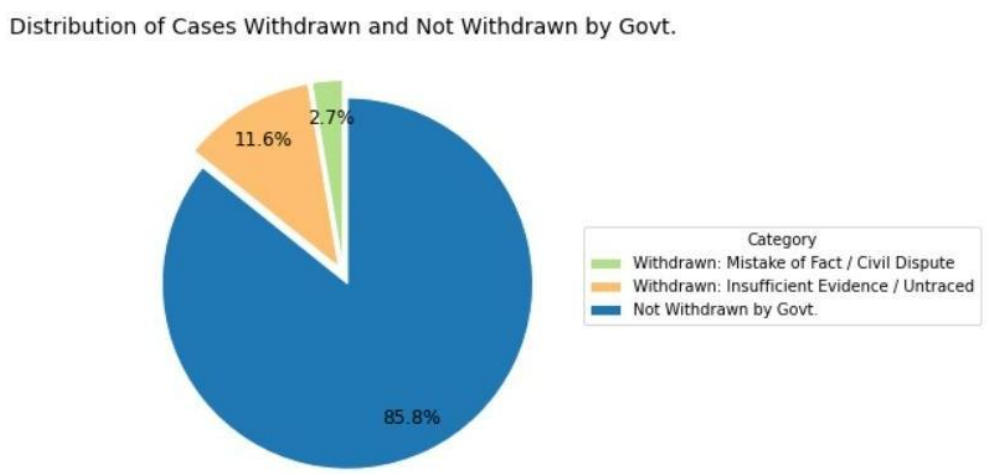
Overall Low Withdrawal Rate

- Across all four years, less than 0.05% of total IPC cases were withdrawn by the government during the investigation stage.
- This suggests that withdrawal is a rare outcome and most cases proceed through other channels (chargesheet, closure reports, pending).

Significant Spike in 2021

- 2021 saw a dramatic increase in the number of withdrawn cases (1,747), over 16 times more than in 2020.
- This caused the withdrawal rate to jump from 0.0019% (2020) to 0.0301% (2021).
- Possible interpretations:
    - Change in government policy or priorities (e.g., backlog clearance).
    - COVID-19 pandemic aftermath might have led to dropping non-priority cases.

5. How many IPC cases were closed due to insufficient evidence or mistakes during investigation?



Insights from the Chart:

Majority of Cases Are Not Withdrawn by the Government

- The largest portion of the pie is composed of cases that remain under investigation or proceed through other legal channels.
- The largest portion of the pie is composed of cases that remain under investigation or proceed through other legal channels.

Withdrawals Due to "Mistake of Fact or Law or Civil Dispute" Form a Significant Share

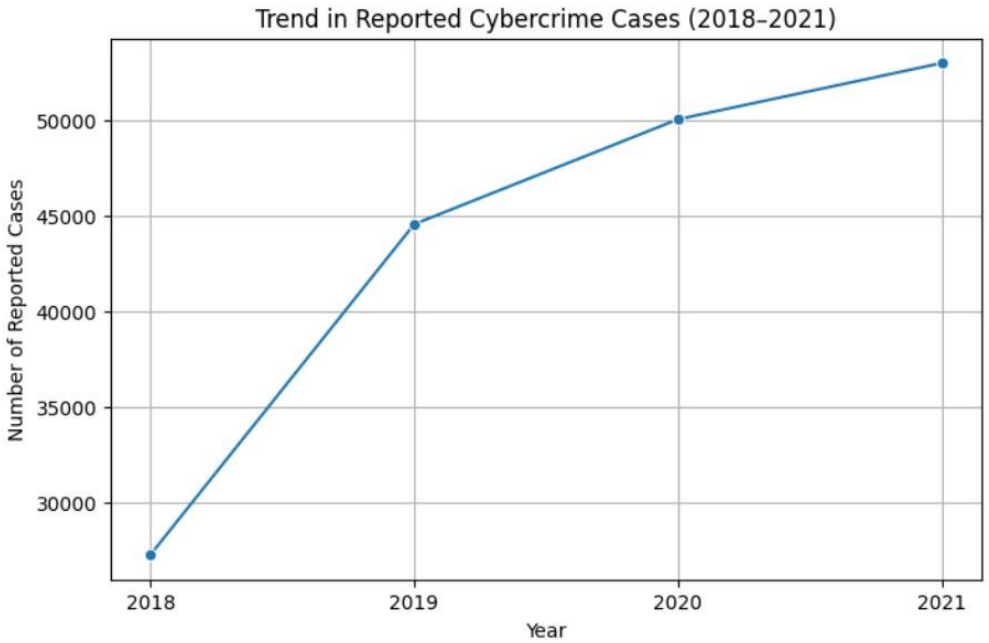
- A considerable portion of withdrawn cases fall under this category, indicating:
  - Cases are sometimes registered as criminal but later found to be civil in nature or legally misclassified.
  - There may be overreach at the FIR stage or unclear distinctions between civil and criminal liabilities.

Withdrawals Due to "True but Insufficient Evidence / Untraced" Also Stand Out

- A substantial number of cases are withdrawn even though the offence is believed to have occurred, due to:
  - Lack of witnesses, forensic evidence, or investigative breakthroughs.

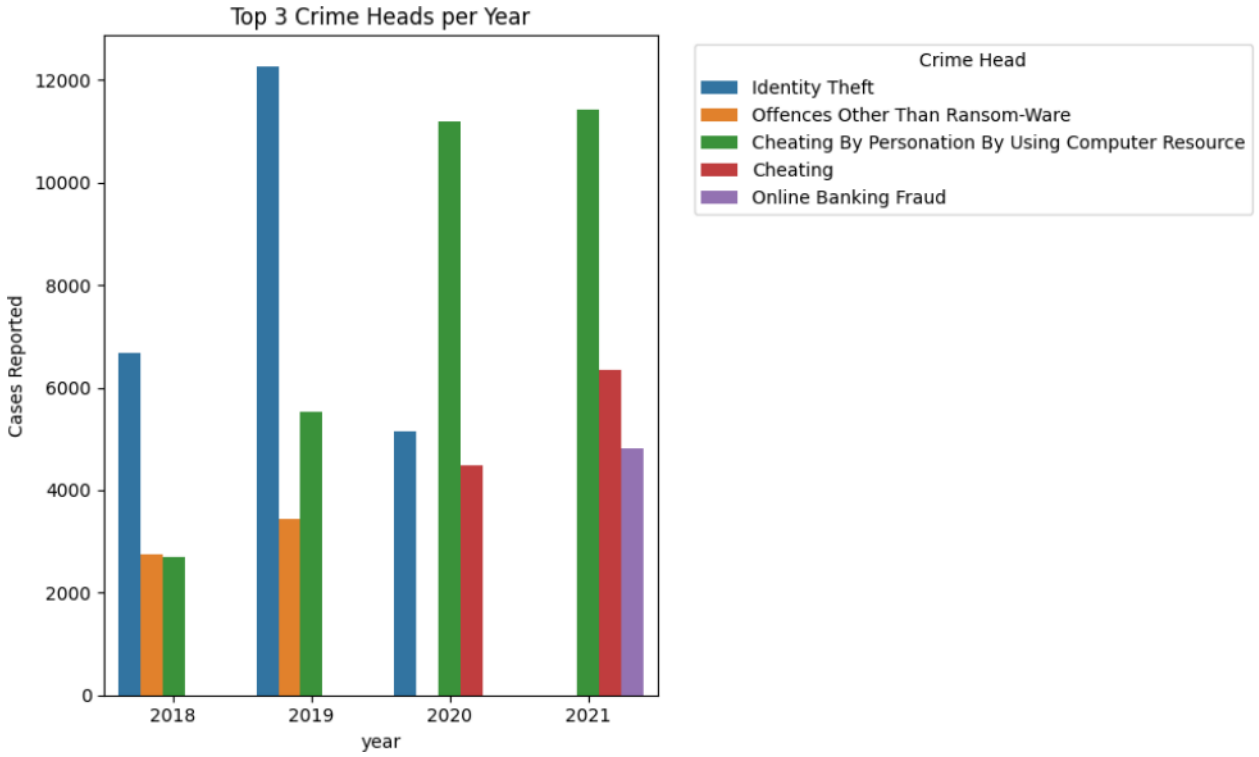
Analysis of Cyber Crimes:

1. Checking the Trend of Cyber Crime



The Chart Shows that the number of reported cases under Cyber Crime are increasing with time.

Top 3 Crime Heads per year



Top 3 Reported Cybercrime Categories (2018–2021) – Insights

2018

Most reported: Identity Theft

Followed by: Offences Other Than Ransomware, Cheating by Personation → Focus on personal data misuse and general cyber offences dominated this year.

2019

Most reported: Identity Theft

Followed by: Cheating by Personation, Offences Other Than Ransomware → Continuation of identity-related crimes, with impersonation-based frauds rising.

2020

Most reported: Cheating by Personation

Followed by: Identity Theft, Cheating → A shift toward social engineering scams, likely influenced by increased online activity during the pandemic.

2021

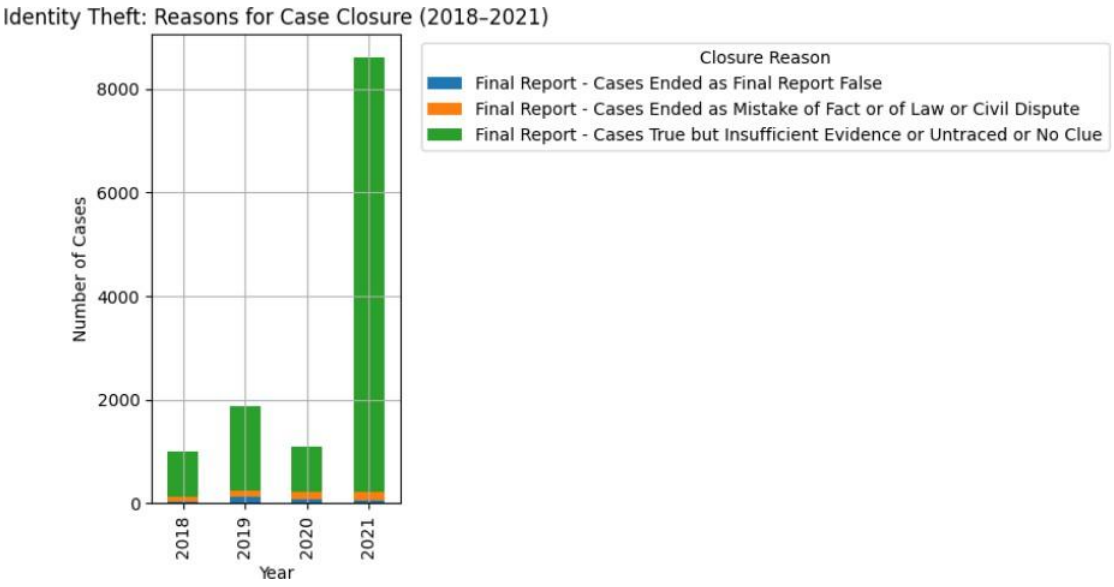
Most reported: Cheating by Personation

Followed by: Cheating, Online Banking Fraud → Financially motivated cybercrimes became more prominent, reflecting a rise in online transactions and banking frauds.

Key Takeaways

- Identity Theft was the leading crime in 2018 and 2019, highlighting concerns around data privacy and misuse of personal information.
- Cheating by Personation emerged as the top threat in 2020 and 2021, showing how impersonation tactics have become increasingly sophisticated.
- Online Banking Fraud appeared in the top 3 in 2021, indicating a shift toward more targeted financial cybercrimes.

If the cases were closed what was the reason behind that ?

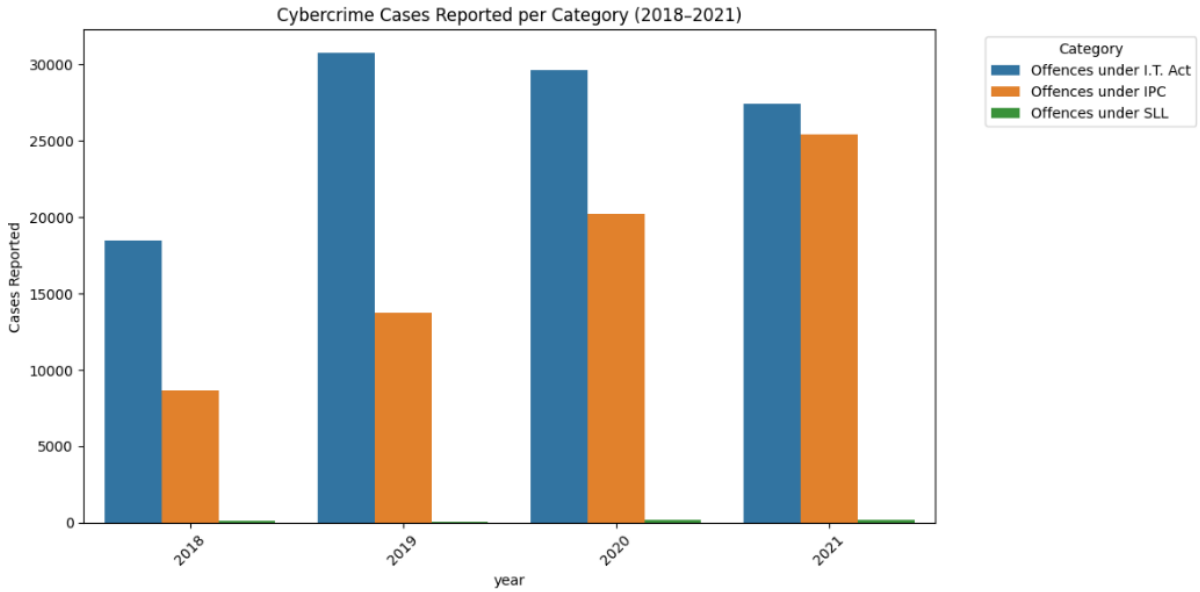


Main Reason for Closure:

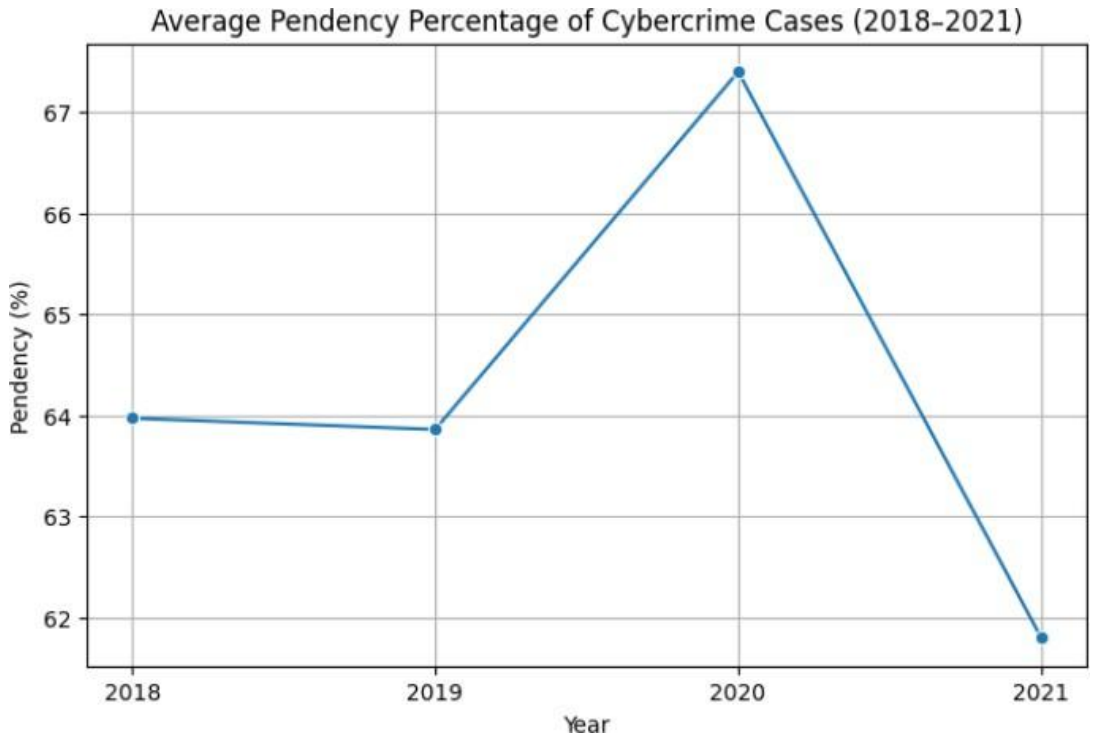
Cases True but Insufficient Evidence / Untraced / No Clue

Majority of cases were genuine but closed due to lack of evidence or inability to trace the offender.

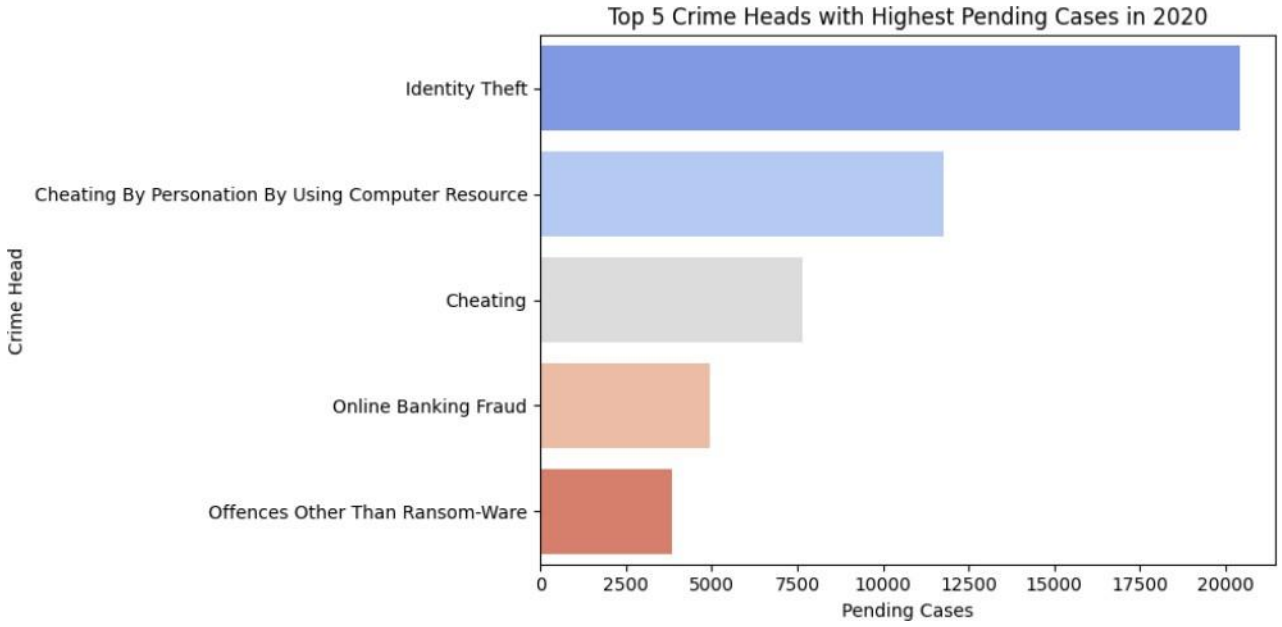
2. Which cybercrime categories show the highest number of cases reported during the years 2018-2021?



3. 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?



As there is a major hike in the year 2020, now let us find that which are the top crime heads affecting to this

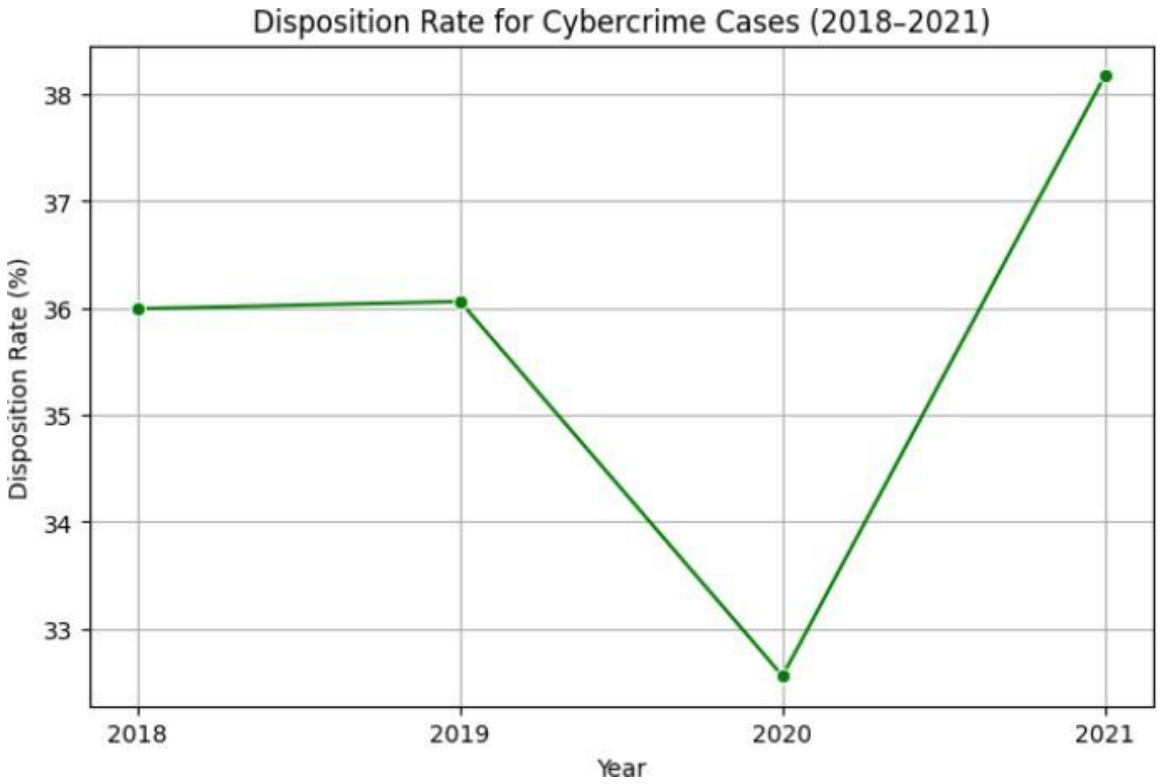


Insight: High Pendency in Identity Theft Cases (2020)

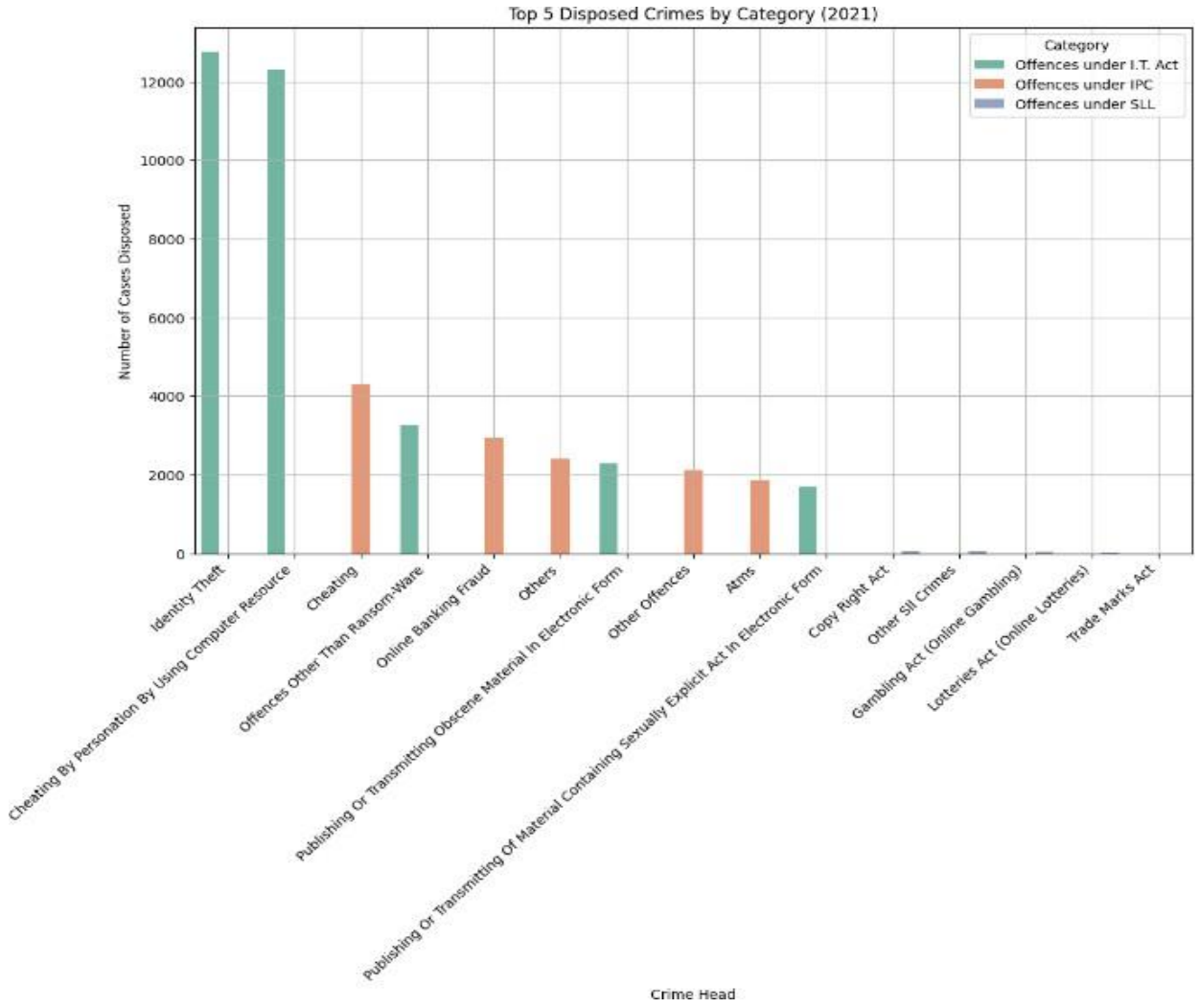
In 2020, **Identity Theft** recorded the highest number of pending cybercrime cases by the end of the year. This reflects a **significant backlog** in resolving such cases.

- **The surge in identity-related offences may be linked to**:
  - **Increased digital dependence** during the COVID-19 pandemic (remote work, online banking, etc.).
  - **Rise in online interactions and data sharing,** increasing vulnerability to identity misuse.
- **The high pendency rate may indicate**:
  - **Challenges in technical investigation** and evidence gathering.
  - **Resource limitations** within law enforcement.
  - **Delays in digital forensic analysis** and case handling.
- **Consequences of high pendency:**
  - **Delayed justice** for victims.
  - **Continued activity of offenders** due to unresolved cases.
  - **Reduced public trust** in the system's ability to address cybercrimes effectively.

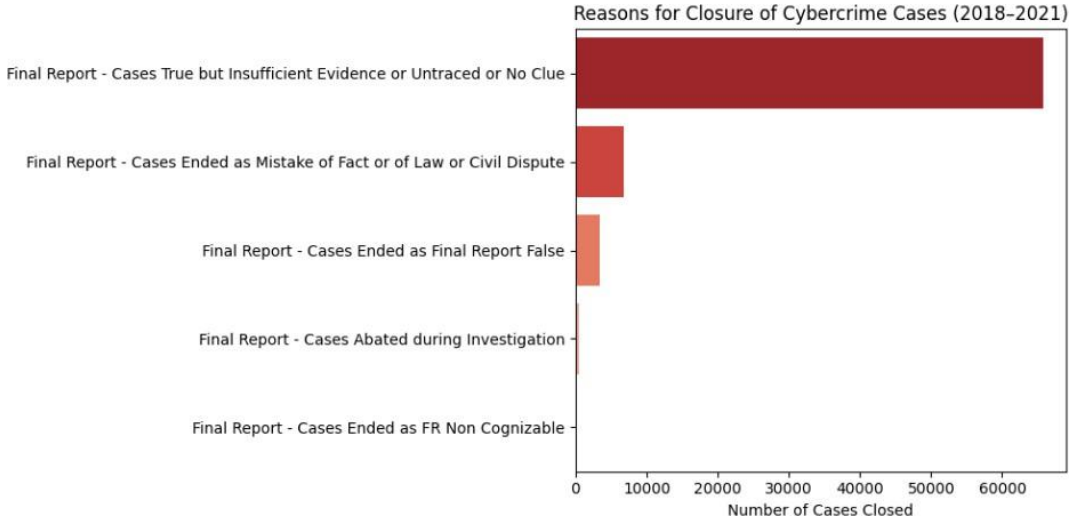
4. What is the disposition rate (cases disposed off) for cybercrimes across the years?

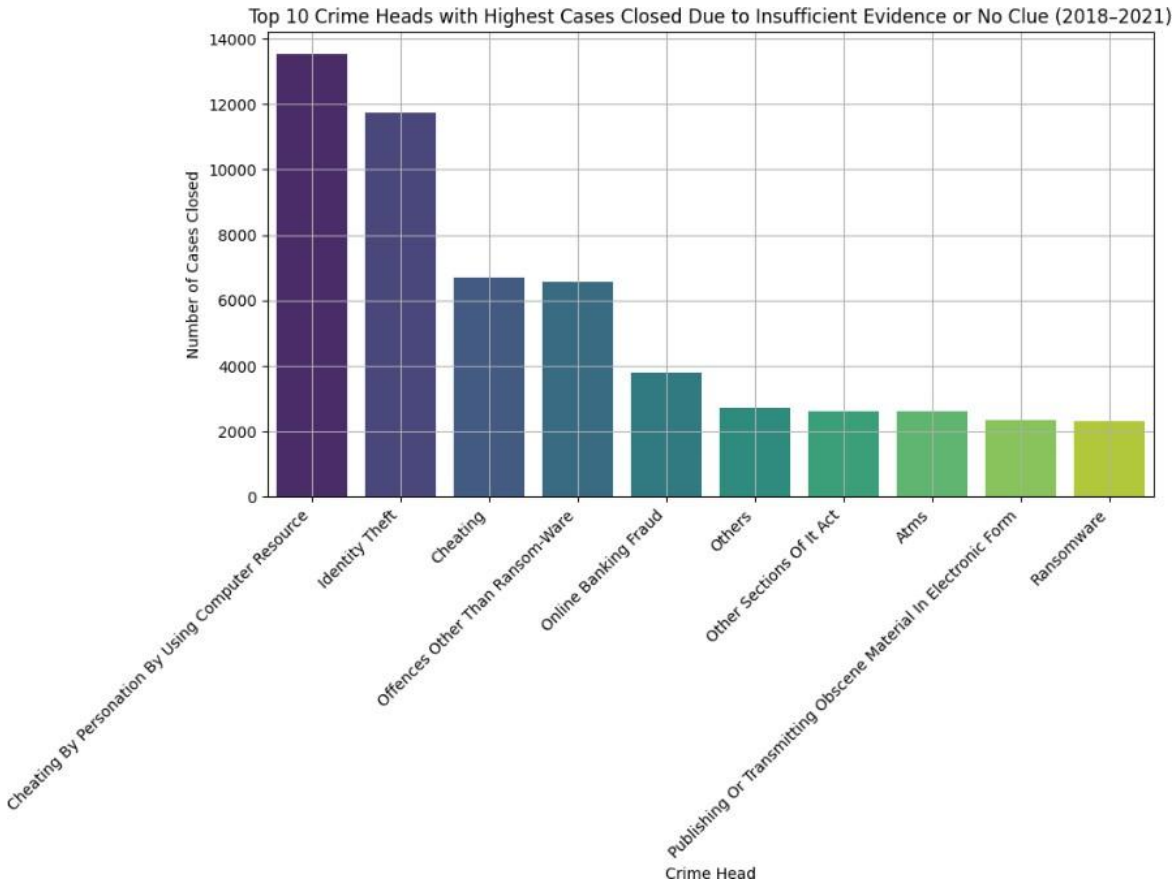


In 2021 higher number of cases was disposed, so the top cases disposed in 2021 are



5. What are the reasons for the closure of cybercrime cases (e.g., cases ended as false, insufficient evidence, mistake of fact)?



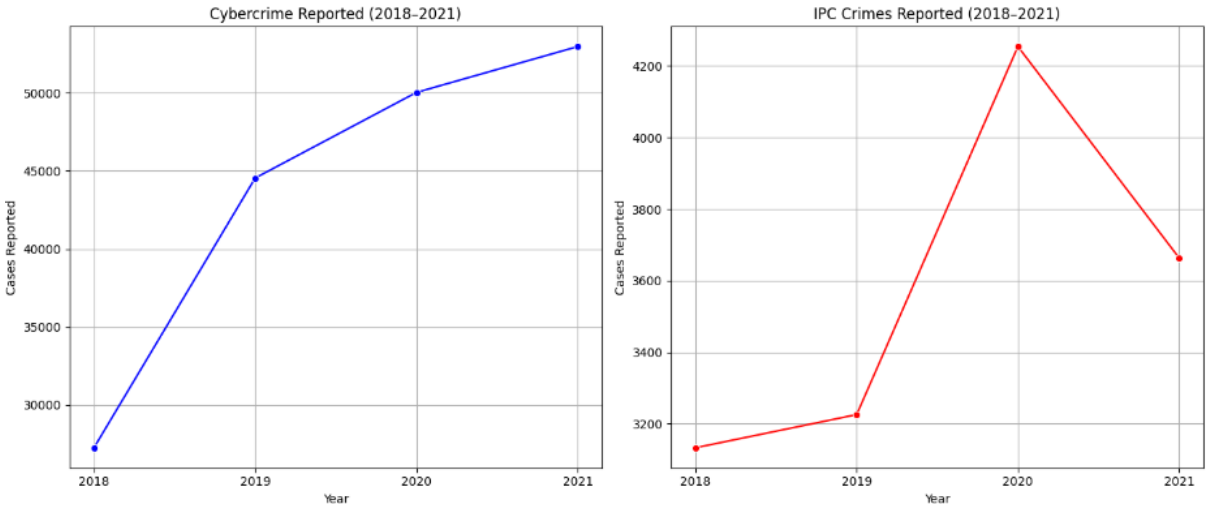


Summary Insight: Cybercrime Trends (2018–2021)

- **Rising Cybercrime Incidents:** Cybercrime, particularly Identity Theft, has shown a consistent increase from 2018 to 2021. This rise correlates with the growing digital dependence during the COVID-19 pandemic, where more people engaged in online banking, remote work, and digital communication, providing more opportunities for cybercriminals.
- **Challenges with Case Resolution:** The pendency rate of cybercrime cases, especially in the Identity Theft category, has remained high. Many of these cases remained unresolved due to factors like insufficient evidence, technical challenges, and limited resources within law enforcement agencies, leading to backlogs in the investigation process.
- **Improvement in Disposition Rates:** In 2021, there was a noticeable increase in the disposition rate, indicating that more cases were closed compared to previous years. However, many of these closures were due to lack of evidence, particularly in Identity Theft and Online Banking Fraud cases, highlighting difficulties in prosecuting complex cybercrimes.
- **Reasons for Case Closures:** A significant number of cases were closed under the reason 'Cases True but Insufficient Evidence or Untraced or No Clue'. This suggests that law enforcement is facing challenges in gathering concrete digital evidence to resolve cases, particularly in categories like Identity Theft and Online Banking Fraud, which are complex and often involve cross-border criminal activities.
- **Looking Ahead:** To address these challenges, there is a pressing need to improve digital forensics capabilities, allocate more resources to cybercrime investigations, and foster greater collaboration among national and international agencies. These steps are crucial for faster resolutions and to deter future cybercriminal activities.

Comparative Analysis on Cyber Crime and IPC Crime

1. How do the trends in cybercrime cases reported compare to those in IPC crimes from 2018-2021?



Cybercrime Trends (2018–2021):

Steady and significant rise over the four years.

The increase is consistently upward, with no dips in the trend

Indicates a growing threat or increased awareness/reporting of cybercrime.

IPC Crimes Trends (2018–2021):

Shows a mild increase from 2018 to 2020.

Peaked in 2020, reaching over 4,200 cases (in 1000s, so ~4.2 million).

Dropped notably in 2021 to ~3,600 cases (in 1000s)

The 2021 drop might be influenced by pandemic-related restrictions, lower mobility, or changes in law enforcement/reporting mechanisms.

Comparative Insights:

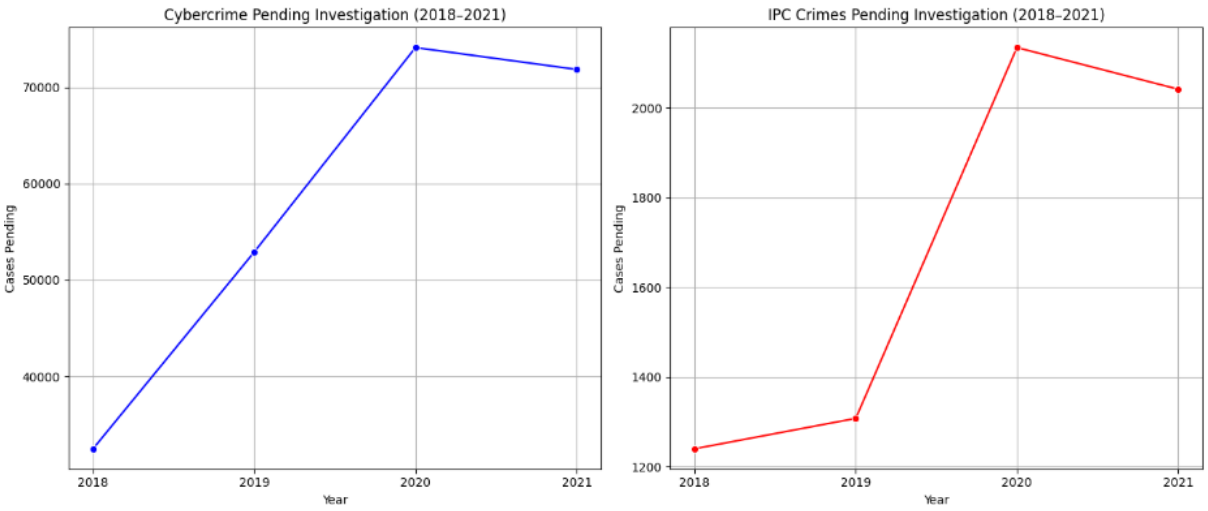
Cybercrime is on a clear upward trajectory, reflecting either more digital activity, better reporting mechanisms, or increasing online threats.

IPC crimes have fluctuated, and the drop in 2021 suggests a temporary or situational impact (possibly due to COVID-19).

While IPC crimes still outnumber cybercrimes overall, the rate of growth in cybercrimes is much higher.

If the trend continues, cybercrime could become a major criminal concern, potentially demanding more resources and policy adjustments.

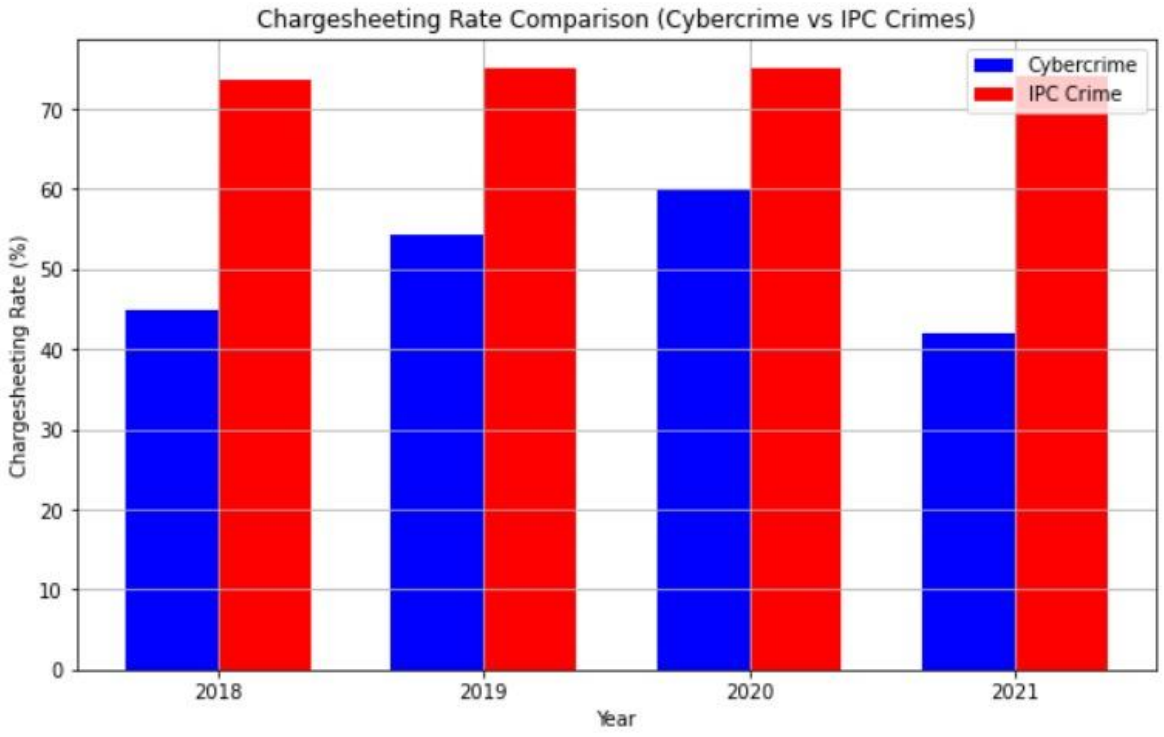
2. Which Types of Crimes (Cyber or IPC) Show a Higher Rate of Pending Investigations at the End of the Year?



Insight: Pending Investigations – Cybercrime vs IPC (2018–2021)

- **Cybercrimes show a significantly higher rate of pending investigations** compared to IPC crimes across all years.
- **Cybercrime Cases:**
  - Slight dip in 2021 but still remains above 70,000
  - Indicates rising cybercrime volume and possible investigation backlog
- **IPC Crimes:**
  - Relatively low and stable
  - Shows better handling or fewer IPC-related tech cases

3. What is the Charge sheeting Rate for Cybercrime vs IPC Crime Cases Over the Years?

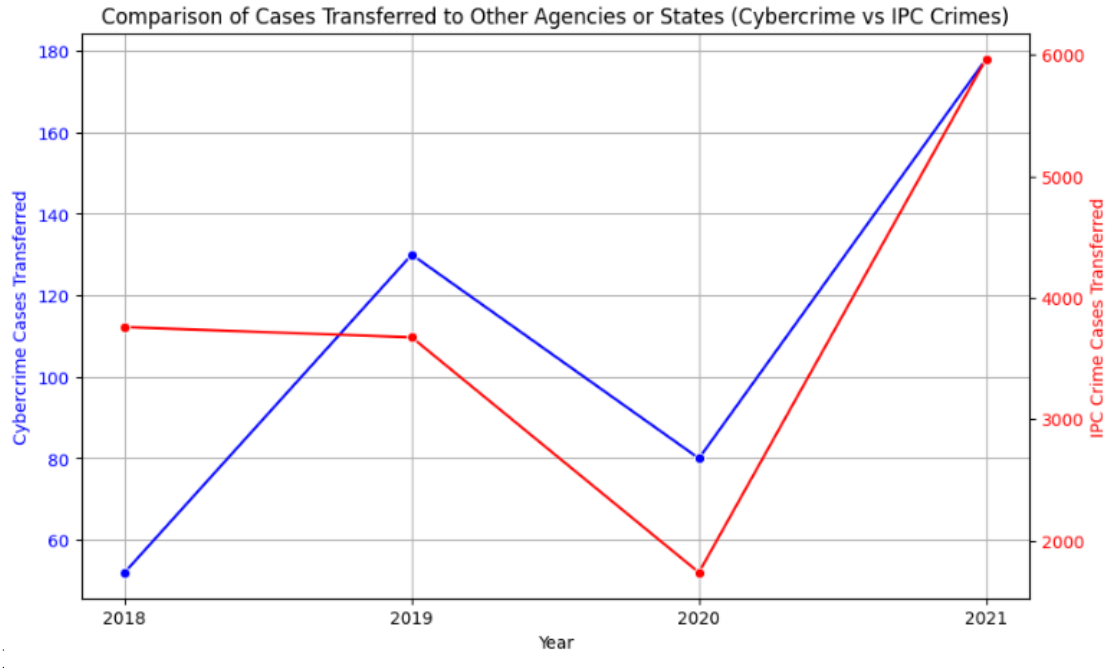


Insights based on Charge Sheeting Rate Comparison (Cybercrimes vs IPC Crimes)

Observation: The charge sheeting rate for IPC crimes consistently outperformed cybercrimes across all years (2018–2021). While IPC crimes maintained a rate above 70%, cybercrimes lingered near 50%.

This disparity highlights challenges in investigating and filing chargesheets for cybercrimes, possibly due to the complexity of digital evidence and tracing perpetrators.

4. Which Type of Crime (Cybercrime or IPC Crime) Has a Higher Rate of Cases Transferred to Other Agencies or States?



Observation:

- **IPC Crimes:** Consistently show a significantly higher number of cases transferred compared to Cybercrimes throughout the years.
- **Cybercrime:** Though the total number is much lower, there is a clear increasing trend from 2018 (lowest) to 2021 (highest).

Year-wise Comparison:

- **2018:** Both cybercrime and IPC cases transferred were relatively low, with IPC still dominant.
- **2019:** Cybercrime cases transferred increased noticeably, while IPC slightly declined.
- **2020:** A dip was observed for both, especially IPC, likely due to disruptions from the pandemic.
- **2021:** Both crime categories showed a sharp rise in transfers, with IPC cases peaking dramatically.

Key Insight:

- **Volume Gap:** IPC crimes have a far larger volume of transferred cases (in thousands) compared to cybercrimes (in hundreds), suggesting that IPC cases often require inter-state or inter-agency coordination.

- **Trend Similarity:** Both crime types follow a similar pattern, with parallel peaks in 2019 and 2021, indicating possible systemic or reporting-driven influences across both domains.

Correlation Matrix

Pearson's Correlation: This method measures the linear relationship between two datasets. If the correlation is close to -1, it indicates a negative correlation (as one increases, the other decreases).
Interpretation of Correlation Coefficient:

- -1: Strong negative correlation (as cybercrimes increase, IPC crimes decrease).
- 0: No correlation (no relationship between the increase in cybercrimes and IPC crimes).
- +1: Strong positive correlation (both increase together).

```
# Group by year and sum the cases reported for both datasets
cyber_total_reported = Cyber.groupby('year')['Cases Reported during the year'].sum().reset_index()
ipc_total_reported = ipc.groupby('Year')['Cases Reported during the year'].sum().reset_index()

merged_data = pd.merge(cyber_total_reported, ipc_total_reported, left_on='year', right_on='Year', suffixes=('_cyber', '_ipc'))

# Calculating Pearson's correlation between Cybercrime and IPC Crimes
correlation = merged_data['Cases Reported during the year_cyber'].corr(merged_data['Cases Reported during the year_ipc'])

print(f"Pearson's Correlation between Cybercrimes and IPC Crimes: {correlation:.2f}")
```

```
Pearson's Correlation between Cybercrimes and IPC Crimes: 0.68
```

Correlation Insight: Cybercrime vs IPC Crime (2018–2021)

Correlation Coefficient: 0.68

Interpretation:

- **Positive Correlation:** Both cybercrimes and IPC crimes tend to increase together. An increase in cybercrime often coincides with an increase in IPC crime.
- **Moderate Strength:** The correlation of 0.68 is moderate—not strong enough to suggest a direct cause, but notable enough to indicate a relationship.

Possible Explanations:

- **Increase in Online & Offline Crimes:** As society becomes more tech-driven, both types of crimes are growing in parallel.
- **Resource Allocation:** With more focus on tackling cybercrime, traditional crime response strategies may be impacted.
- **Improved Awareness & Reporting:** A rise in cybercrime awareness may also lead to increased reporting of other (IPC) crimes.

Conclusion

This project closely examines crime trends in India, with a specific focus on crimes involving foreigners, cybercrimes, and IPC offenses from 2018 to 2021. By leveraging Python for data analysis, we have uncovered key patterns and trends that provide a deeper understanding of the evolving crime landscape. The analysis highlights how the rise in cybercrimes correlates moderately with changes in IPC crime rates, suggesting shared underlying factors such as increased digital adoption and societal changes.

The project also evaluates the effectiveness of law enforcement through metrics such as pendency rates, chargesheeting rates, and case disposition statistics. Our comparative analysis reveals that while IPC crimes still dominate in volume, cybercrimes are rising rapidly, demanding greater attention, resources, and specialized handling.

By identifying data-driven insights, this study contributes toward building more effective crime prevention strategies, enhancing investigative capacities, and shaping public policy to better address both conventional and technology-driven offenses in the future.

Suggestions for Improvements

- Set up more cybercrime units with trained personnel and advanced tools.
- Train police regularly in cyber laws and digital investigation techniques.
- Implement a unified case tracking system to reduce delays.
- Raise public awareness on cyber safety through campaigns and education.
- Establish fast-track courts for speedy resolution of tech-related crimes.
- Improve reporting methods via online FIRs and mobile apps.
- Balance resource allocation between cyber and IPC crimes.
- Use data analytics to shape proactive crime prevention policies.
- Encourage agency collaboration for better cross-border cybercrime handling.

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** – Managing and analysing Combined dataset(IPC + Cyber Crime).