Capstone Project Report (India Crime Dataset – 2001 – 2012)

This project was built in 4 phases.

1. Data Collection – Population, Area & Literacy Rate
2. Analysis on 9 different crime data which included: crime against children, women, Juvenile crime data, IPC crimes, Crimes against police, Cases & Property Value, Place of occurrence for different crimes, human rights, Crimes against Scheduled Caste &Schedules Tribe.
3. State & District Wise Analysis of crimes on SQL
4. Creating of 3 clusters to identify sensitive, moderate & peaceful areas.

Firstly, the data was collected using government official websites wherein upon analysis, it was revealed that the population has indeed grown for India along with the literacy rates. Most populated state turned out to be Uttar Pradesh. Literacy rates were higher for states such as Chandigarh, Kerala, Goa, Lakshadweep, Mizoram etc.

Initial stages required efforts to cleanse the dataset and individual upload them on Python to begin with simpler analysis. Each dataset were studied to get more insights about the different types of crime, its frequency and the hotspots for each crime. It was a rigorous process which took quite a lot of time before moving on to the unsupervised ML part. The analysis assumed that ‘0’ values were not missing data but instead represented the true values and thus it was kept during the analysis stages.

For crimes against children, the data underscores the alarming prevalence of certain offenses across different states. Notably, states like Uttar Pradesh, Maharashtra, and Madhya Pradesh exhibit a significant frequency of crimes against children, containing a range of offenses including murder, rape, kidnapping, and abduction. Districts such as Meerut, Mumbai, and Outer Delhi emerge as focal points for these acts, necessitating targeted intervention strategies.

The data reveals an upward trajectory in crimes against children since the early 2000s, indicating a pressing need for enhanced child protection measures nationwide. Of particular concern are crimes like foeticide, abetment of suicide, and exposure/abandonment, which exhibit troubling trends in certain regions, necessitating focused attention and intervention from law enforcement agencies and social welfare organizations.

Similarly, crimes against scheduled castes (SC) highlight pervasive social inequalities and discriminatory practices in certain states. Uttar Pradesh emerges as a hotspot for SC-related crimes, including a spectrum of offenses ranging from murder and rape to dacoity and robbery. The prevalence of such crimes underscores the imperative for stringent enforcement of laws aimed at safeguarding the rights and dignity of marginalized communities.

Conversely, crimes against scheduled tribes (ST) exhibit a declining trend in recent years, indicative of targeted efforts aimed at possibly addressing systemic injustices and safeguarding the rights of tribal communities. However, states like Rajasthan, Madhya Pradesh, and Chhattisgarh continue to grapple with significant challenges in combating crimes against ST, necessitating sustained efforts to ensure the equitable protection of vulnerable populations.

The analysis of IPC (Indian Penal Code) crime statistics reveals several noteworthy trends and patterns across various categories of offenses. On average, murder cases are reported at a rate of approximately 89, with occasional spikes reaching as high as 7601 incidents. Similarly, attempts to murder and culpable homicides exhibit an average occurrence of 78 cases each, with maximum instances recorded at 7964 and 1616, respectively.

Crimes against women, particularly rape and kidnapping, are prevalent issues, with an average of 53 reported rape cases per year, accompanied by a maximum of 3425 instances. Kidnapping and abduction of women and girls also pose significant concerns, with an average of 79 cases annually, peaking at 8878 incidents. Furthermore, cases of assault on women, cruelty by husbands, and insult to the modesty of women underscore the pervasive nature of gender-based violence in society.

Instances of property-related crimes such as theft, robbery, and burglary demonstrate a concerning upward trend, mirroring the overall increase in total IPC crimes across districts. However, certain offenses like dacoity and burglary have shown a decline over the years. Notably, states like West Bengal, Andhra Pradesh, and Bihar have witnessed improvements in crime rates, indicative of effective crime prevention measures implemented within these regions.

Among districts, Bangalore, Mumbai, and Ahmedabad emerge as hotspots for total IPC crimes, reflecting the urban centers' vulnerability to criminal activities. Conversely, districts like Mumbai Commissionerate, Cyberabad, and Ernakulam Commissionerate exhibit noteworthy improvements in crime rates, reflecting successful law enforcement strategies and community engagement initiatives.

The analysis of crime (Place of Occurrence Data) pertaining to various categories such as dacoity, robbery, theft, and burglary provides valuable insights into the patterns and trends of criminal activities across different locations. On average, residential premises emerge as a significant hotspot for dacoity, with an average occurrence of 152 cases, while highways witness a notable prevalence of robbery incidents, averaging at 332 cases. Conversely, theft crimes demonstrate a widespread occurrence, with residential premises accounting for an average of 5881 cases, highlighting the vulnerability of residential areas to such offenses. The average cases of dacoity at other places and commercial establishments are 124 and 24, respectively, reflecting a relatively lower occurrence in comparison.

Moreover, burglary cases exhibit a similar trend, with residential premises ranking high in occurrences, averaging at 4694 cases. The average cases of burglary at other places and commercial establishments are 7149 and 1070, respectively, indicating a significant distribution of burglary incidents across different locations. The data also highlights notable disparities across states, with Maharashtra, Madhya Pradesh, and Andhra Pradesh emerging as top-ranking states for burglary and theft crimes, indicating the need for targeted interventions and law enforcement measures within these regions.

Interestingly, certain areas show fluctuations in crime occurrences over the years, with fluctuations observed in dacoity and burglary cases at various locations. However, theft crimes consistently maintain a high occurrence rate across different areas, emphasizing the need for comprehensive strategies to address property-related offenses. The average cases of theft on highways, railways, and banks are 1000, 1097, and 32, respectively, showcasing the diverse nature of theft occurrences across different settings.

Moreover, the analysis of juvenile crimes, crimes against police, and human rights violations also provides valuable insights into various dimensions of law enforcement in India.

Juvenile crimes, particularly among boys aged 12-18 years, show concerning trends with a steady rise observed over the years. Madhya Pradesh, Maharashtra, Gujarat, Haryana, and Rajasthan emerge as the states with the highest number of juvenile crimes. The data underscores the need for targeted interventions to address juvenile delinquency, especially among boys compared to girls.

Crimes against police officers reveal alarming trends, with a significant number of injuries and fatalities reported. Accidents and terrorist operations pose substantial risks to law enforcement personnel, with Maharashtra, Tamil Nadu, and Jammu & Kashmir being the most affected states. These findings emphasize the importance of enhancing safety measures and providing adequate support to police personnel.

Human rights violations remain a pressing concern, with cases reported across various states. Chattisgarh, Assam, and Andhra Pradesh record higher numbers of human rights violations, underscoring the need for robust mechanisms to safeguard human rights and ensure accountability in law enforcement.

Additionally, the correlation analysis sheds light on the relationship between demographic factors and crime rates. While no significant correlation is observed between literacy rates and total crimes, a positive correlation is noted between population size and overall crime rates, indicating higher crime prevalence in densely populated areas. The analysis also reveals varying recovery rates for stolen property across states, with Maharashtra and Delhi recording higher values of stolen property.

For the unsupervised learning part, all the dataset which included ‘District’ columns were merged. It went through a cleansing phase and crimes were grouped together to reduce the number of columns to represent a better analysis. Clustering analysis was performed and 3 areas where identified such as sensitive (higher amount of crimes), moderate (medium amount of crimes) and peaceful (lower amount of crimes)

The clustering analysis of crime data reveals distinct patterns and trends across different states and districts in India. Majority of the districts in Madhya Pradesh, Maharashtra, Uttar Pradesh, Kerala, Bihar, Gujarat, Karnataka, and Tamil Nadu are classified under the moderate areas category, indicating a relatively stable crime situation. Conversely, districts in Punjab, Odisha, Mizoram, Meghalaya, Himachal Pradesh, Jharkhand, Arunachal Pradesh, Andaman & Nicobar Islands, Daman & Diu, Goa, Lakshadweep, Sikkim, Tripura, and Uttarakhand are categorized as sensitive areas, signifying higher levels of crime vulnerability. Notably, districts in Andhra Pradesh, Madhya Pradesh, Maharashtra, Kerala, and Rajasthan exhibit fewer peaceful districts, indicating persistent challenges in maintaining law and order.

The trend analysis further reveals dynamic shifts in crime patterns over time. Moderate areas witness an overall increasing trend, suggesting a rise in the number of districts falling into this category. Conversely, sensitive areas show a downward trend, indicating a decrease in their representation over time. Peaceful areas, though initially few, exhibit a significant increase over the years, indicating a positive trend towards enhanced safety and security.

In terms of specific crimes, Madhya Pradesh ranks high for moderate areas, while Uttar Pradesh ranks prominently for sensitive crimes. Uttar Pradesh tops the list for murder cases, Madhya Pradesh for rape, Maharashtra for robbery, Andhra Pradesh for crimes against women, and Rajasthan for crimes against Scheduled Castes and Scheduled Tribes. Additionally, Andhra Pradesh records the highest number of violent crimes, Rajasthan for frauds, and Tamil Nadu for IPC cases.

The analysis also identifies specific districts with notable crime occurrences. For instance, Thoothukudi records the highest murder cases, Shahdol for rape, Chandigarh for theft and robbery, Kollam for crimes against women, Badaun for crimes against children, and Jalore for crimes against Scheduled Castes and Scheduled Tribes.

In conclusion, this analysis sheds a light on the multifaceted nature of criminal offenses prevalent in India. It is imperative to identify key trends and hotspots, policymakers, law enforcement agencies, and civil organizations to formulate targeted interventions which can curb crime rates, safeguarding vulnerable populations, and promote social justice and inclusivity across the nation.