

Capston Project

Crime Report in India

Navin Singh

Phase: 1

State Wise from 2001 to 2012

INTRODUCTION:

This project divides into 4 parts:

This is 1st Part of Crime Report Project:

Crime in Indian is a significant concern that demands attention at both societal and governmental levels. The analysis presented here aims to shed light on the state of crimes against people in whole India over the year from 2001 to 2012.

Data Collection:

For this project I needed all district's data of whole India so that, Some Important data provided by my educational Institute (Data Trained Academy) like all types of crime details and data of whole Indian state and some data I collected from different – different website.

I collected data as follow:

- 1) Population of each state
- 2) Average literacy rate in % of each state
- 3) Area(KM2) of each state
- 4) Sex Ratio of each state

Click the link to find the data sources

1. https://en.wikipedia.org/wiki/List_of_states_and_union_territories_of_India_by_sex_ratio

2. <http://cyberjournalist.org.in/census/cenindia.html>

3. <https://www.indiacensus.net/>

Overview:

The data collected and analyzed for this report includes various types of Crimes in Each district of whole Indian State including Rape, Murder, Kidnapping, abduction, dowry deaths and assault.

Now I have all types of crime data in entire Indian state and union territory with district.

For this 1st Phase, Firstly I collect all data from different-different website and save in single excel file name as 'collected_data', in this single file I save all types of important data which is provided by my institute and also which I collected from website.

I created one new jupyter notebook name with Project 'Capstone_Crimes in India'. I import some important python library like

- Pandas as pd
- Numpy as np
- Matplotlib.pyplot as plt
- Seaborn as sns
- Warning for avoid the warnings

Then I load with the help of python library. I also mention all links from where I got the data.

In this file data available from 2001 to 2012 including all district but For analysis of this phase 1, I only need state wise data with yearly basis not all district wise data, so that I divide this data in 12 part by yearly base name as

k1(2001), k2(2002), k3,k4, k5, k6, k7, k8, k9, k10, k11, k12(2012) now it's easy to excess yearly wise data.

Now I start Comparison between two columns to get the knowledge what relationship between two tables is like AVERAGE LITERACY RATE and TOTAL IPC CRIMES using different different plot (bar plot, count plot , strip plot, lamp plot)

After analysis by different-different graph, I can clearly say that, from 2001 to 2012 as the literacy rate is increasing, the crime rate is decreasing. So that if we want to decrease the crime rate, firstly we have to work on to increasing the Literacy rate in whole India.

I also observe:

- 1) MADHYA PRADESH state has the maximum number of RAPE cases come in light from year 2001 to 2012
- 2) UTTAR PRADESH state has the maximum number of CULPABLE HOMICIDE NOT AMOUNTING TO MURDER cases come in light from year 2001 to 2012
- 3) ASSAM state has the maximum number of CUSTODIAL RAPE from 2001 to 2012
- 4) MADHYA PRADESH has Maximum number of OTHER RAPE cases come in light from 2001 to 2012
- 5) UTTAR PRASDESH state has the maximum number of KIDNAPPING & ABDUCTION case were registered
- 6) Uttar Pradesh has the maximum number of Kidnapping and Abduction of Women and Girls in India from 2001 to 2012
- 7) DELHI UT state is the maximum KIDNAPPING AND ABDUCTION OF OTHER from 2001 to 2012
- 8) Bihar states has the maximum number of DACOITY from 2001 to 2012
- 9) MAHARASHTRA states has the maximum number of ROBBERY from 2001 to 2012

- 10) Bihar state has the maximum number of RIOTS from 2001 to 2012
- 11) Rajasthan state has the maximum number of CHEATING from 2001 to 2012
- 12) Utter Pradesh state has the maximum number of Dowry Deaths from 2001 to 2012
- 13) Madhya Pradesh state has the maximum number of ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY from 2001 to 2012
- 14) West Bengal state has the maximum number of CRUELTY BY HUSBAND OR HIS RELATIVES from 2001 to 2012
- 15) Bihar State has the maximum number of IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES from 2001 to 2012
- 16) Tamil Nadu state has the maximum number of CAUSING DEATH BY NEGLIGENCE from 2001 to 2012
- 17) Tamil Nadu state has the maximum number of OTHER IPC CRIMES from 2001 to 2012
- 18) As per as the state's area increasing then total crime rate is also increasing from 2001 to 2012
- 19) As per as the Population increasing then total crime rate is also increasing from 2001 to 2012

Phase: 2

Topic: SQL Operations

This is completely SQL based Phase, as per the instructions firstly I insert all record from the file name is

'42_District_wise_crimes_committed_against_women_2001_2012.csv' on my MYSQL workbench, and I follow all the instructions and solved all question section by section.

Firstly I studied on Rape Cases data.

From 2001 to 2012, the maximum number of rape cases is 568 happened in West Bengal's Murshidabad district.

YEAR,	TOTAL NO. OF RAPE CASE,	STATE/UT,	DISTRICT
• 2001:	total rape case: 145	Chhattisgarh,	Saguna
• 2002:	total rape case: 159	Chhattisgarh,	Raipur
• 2003:	total rape case: 137	M.P,	Betul
• 2004:	total rape case: 197	West Bengal,	Murshidabad
• 2005:	total rape case: 236	West Bengal,	Murshidabad
• 2006:	total rape case: 224	West Bengal,	Murshidabad
• 2007:	total rape case: 369	West Bengal,	Murshidabad
• 2008:	total rape case: 334	West Bengal,	Jalpaiguri
• 2009:	total rape case: 568	West Bengal,	Murshidabad
• 2010:	total rape case: 526	West Bengal,	Murshidabad
• 2011:	total rape case: 433	West Bengal,	Murshidabad
• 2012:	total rape case: 257	West Bengal,	Murshidabad

Then I studied on Dacoity Data from 2001 to 2012.

I got, from 2001 to 2012, the maximum number of Dacoity is 29 in 2004 in Gujarat's Dahod district.

Then I studied on Robbery Data set:

I got, from 2001 to 2012, the maximum number of Robbery is 32 in 2001 in Gujarat's Dahod district.

Then I start work to find which district has the lowest number of Murders Happened.

I get, from 2001 to 2012 more then 600+ rows, where 0 number of murder happened.

I also observe:

- Maximum number of murder is 28 in Nagaland's Kohima district in 2006.
- 640 rows where 1 murder happened
- 204 rows where 2 murder happened
- 86 rows where 3 murder happened
- 31 rows where 4 murder happened
- 22 rows where 5 murder happened
- 10 rows where 6 murder happened
- 1 rows where 7 murder happened
- 3 rows where 8 murder happened
- 2 rows where 9 murder happened
- 1 rows where 10 murder happened
- 1 rows where 11 murder happened
- 1 rows where 12 murder happened
- and there are 8 rows where more than 14 murders happened
- The maximum number of murder in single year is 565 in Delhi in 2010.

For Section: 3.8.2,
I create a PDF file and solve all the Questions,
Because I have to upload multiple Screenshot of
MySQL workbench. After that I draw multiple
graphs to get more knowledge. And then I save
this data as 'result_3.2.csv'

Phase: 4

In this phase I work on Machine Learning
Algorithm for Clustering.

In this phase I have to make 3 clusters based on
crimes types

- 1) Sensitive Area
- 2) Moderate Area
- 3) Peaceful Area

I divide this data in three cluster based on columns:
'MURDER', 'RAPE', 'KIDNAPPING & ABDUCTION',
'DACOITY', 'ROBBERY' using KMeans Unsupervised
ML algorithms.

Observations:

1) **Sensitive Area:**

- Cluster 0 is characterized by high crime rates across multiple categories, including murder, rape, kidnapping, dacoity, and robbery.
- These areas exhibit a high level of criminal activity and may require immediate attention from law enforcement agencies.
- Factors contributing to the high crime rates in these areas need to be further investigated, such as socioeconomic conditions, law enforcement presence, etc.

2 **Moderate Areas:**

- Cluster 1 represents areas with moderate crime rates compared to sensitive areas.
- While crime rates are relatively lower than in sensitive areas, there's still a significant presence of various crimes such as murder, rape, and kidnapping.
- Strategies for crime prevention and community engagement could help maintain or further reduce crime rates in these areas.

3 Peaceful Areas:

- Cluster 2 comprises areas with the lowest crime rates among the three clusters.
- Crime rates in these areas are significantly lower across all categories, indicating a relatively safe environment.
- These areas may serve as examples of effective crime prevention strategies or positive community dynamics.

Then I once again draw multiple graph to get more knowledge.

Recommendations:

Addressing Sensitive Areas:

Allocate additional resources and law enforcement personnel to tackle high crime rates in sensitive areas. Implement targeted interventions to address underlying factors contributing to crime, such as poverty, unemployment, literacy rate and social inequality.

Maintaining Moderate Areas:

Implement community policing initiatives to enhance trust and collaboration between law enforcement agencies and local communities. Focus on crime prevention programs and youth engagement activities to deter criminal activities in moderate areas.

Sustaining Peaceful Areas:

Continue proactive policing efforts and community engagement initiatives to maintain the low crime rates in peaceful areas. Invest in infrastructure development and social programs to further enhance the quality of life and safety in these areas.

Conclusion:

The clustering analysis provides valuable insights into the spatial distribution of crime and helps identify areas requiring targeted interventions. By understanding the distinct characteristics of each cluster, policymakers and law enforcement agencies

can develop tailored strategies to address crime effectively and promote safer communities.

Phase 4 concludes the capstone project, demonstrating the application of data analysis, SQL operations, and machine learning techniques to gain insights from crime datasets and inform decision-making processes.

Overall, the project highlights the importance of data-driven approaches in addressing complex societal challenges and underscores the potential of data analysis in generating actionable insights for public safety and crime prevention efforts.



Thanks

Navin Singh

DS2309