

CRIME FORECASTING

MINOR PROJECT I

Submitted by:

Shubham Garg (9919103057)

Pranjal Tiwari (9919103010)

Rahul Sharma (9919103173)

Under the supervision of:

Dr. Shilpa Budhkar



Department of CSE/IT

Jaypee Institute of Information Technology University, Noida

December 2021

ACKNOWLEDGEMENT

We would like to place on record my deep sense of gratitude to Dr. Shilpa Budhkar, Assistant Professor of the Department of Computer Science & Engineering and Information Technology, Jaypee Institute of Information Technology, India for her generous guidance, help and useful suggestions.

We also express our thanks Dr. Shikha Mehta, Dr. Himanshu Mittal and other classmates for their insightful comments and constructive suggestions to improve the quality of this project work.



Signatures of Students

Shubham Garg (9919103057)

Pranjal Tiwari (9919103010)

Rahul Sharma (9919103173)

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and beliefs, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma from a university or other institute of higher learning, except where due acknowledgment has been made in the text.

Place: **Jaypee Institute of Information Technology, Noida, India**

Date: **1st December 2021**

Name: **Shubham Garg**

Enrolment No.: **9919103057**

Name: **Pranjal Tiwari**

Enrolment No.: **9919103010**

Name: **Rahul Sharma**

Enrolment No.: **9919103173**

CERTIFICATE

This is to certify that the work titled “**Crime Forecasting**” submitted by **Shubham Garg, Pranjal Tiwari** and **Rahul Sharma** of B. Tech of Jaypee Institute of Information Technology, Noida has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of any other degree or diploma.

Digital Signature of Supervisor:

Dr. Shilpa Budhkar

Assistant Professor (Grade-II)

Date: **1st December 2021**

ABSTRACT

In the article the Business Insider, suggests that the crimes in India have cost economic growth. According to an international think tank, the Institute for Economics and Peace, **crimes have cost the country, 9% of its GDP in 2017**. This equates to roughly \$600 per person in India(Dhillon, 2018). Analysis of crime and taking prevention to ensure the safety of the public is the necessary step for the government and law enforcement. This study will **mainly focus on the top crimes that are being committed in India**. Top crimes in India will be identified on the chart and the worst affected states for each crime will be plotted on the map of India. The trend of total IPC crimes in every state for the next **9 years** will be identified in this study using a machine learning technique. The seasonal trend graph will analyze the observed and trend crime of every state in India. This forecasting technique will **help law enforcement to examine the trend of crimes every state in India and so that they can take minimize the crime as much as possible**. This forecasting model will be dependent on the count of past crimes that are committed. Cities of the top worst state affected will be clustered according to the crimes committed this will help the government to identify the highly vulnerable cities. Cities that are highly affected in top states will be analyzed.

TABLE OF CONTENTS

<i>Abstract</i>	<i>i</i>
<i>List of Tables</i>	<i>iv</i>
<i>List of Figures</i>	<i>iv</i>
<i>List of Abbreviations</i>	<i>v</i>
<i>List of Nomenclature</i>	<i>v</i>
Chapter 1: INTRODUCTION.....	1
1.1 INTRODUCTION.....	1
1.2 AIMS AND OBJECTIVES.....	2
1.2.1 STUDY QUESTION.....	2
1.2.2 OBJECTIVES.....	2
Chapter 2: BACKGROUND STUDY.....	3
Chapter 3: REQUIREMENT ANALYSIS.....	6
3.1 MOTIVATION.....	6
Chapter 4: DETAILED DESIGN.....	7
4.1 SELECTION OF MODEL.....	7
4.2 TOOLS AND TECHNIQUES USED.....	7
4.3 LIBRARIES USED.....	8
Chapter 5: IMPLEMENTATION.....	9
5.1 DATA EXPLORATION.....	9
5.1.1 STEPS INVOLVED.....	9
5.2 DATA VISUALISATION.....	9
5.3 MODELING APPROACH.....	10
5.3.1 ARIMA MODEL.....	10
5.4 MEASURING ACCURACY.....	11
5.5 K-MEANS CLUSTERING MODEL.....	11

Chapter 6: EXPERIMENTAL RESULTS AND ANALYSIS.....	12
6.1 CRIMES IN INDIA.....	12
6.2 TOP 10 CRIMES IN INDIA.....	13
6.3 THEFT IN INDIA.....	14
6.4 FORECASTING.....	14
6.4.1 UTTRAKHAND.....	15
6.5 CLUSTERING.....	16
6.5.1 UTTAR PRADESH.....	16
Chapter 7: CONCLUSION AND FUTERE SCOPE.....	18
7.1 CONCLUSION.....	18
7.2 FUTURE SCOPE.....	19

LIST OF TABLES

S.No.	TITLE	PAGES
1	Abbreviations and Nomenclature.....	v
2	Library Used.....	8

LIST OF FIGURES

FIGURES	Page
Figure 2.1: Age of Women.....	4
Figure 2.2: Nature of Crime.....	4
Figure 2.3: Percentage distribution on nature of crime.....	5
Figure 4.2.1: Languages Used.....	7
Figure 6.1.1: Distribution of crimes.....	12
Figure 6.2.1: Top 10 crimes India.....	13
Figure 6.3.1: Theft in India.....	14
Figure 6.4.1.1: Trend in Uttarakhand.....	15
Figure 6.4.1.2: Forecast in Uttarakhand.....	15
Figure 6.5.1.1: Crimes happening in Uttar Pradesh.....	16
Figure 6.5.1.2: Cities affected in Uttar Pradesh.....	16
Figure 6.5.1.3: Clustered cities in Uttar Pradesh.....	17

ABBREVIATIONS AND NOMENCLATURE

ML	Machine Learning
-----------	-------------------------

ACF	Autocorrelation function
------------	--------------------------

IDE	Integrated Development Environments
------------	-------------------------------------

ARIMA	Auto Regressive Integrated Moving Average
--------------	---

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

In 1947, India got its independence from the British Colony, with the largest democracy and second most populated country in the world. Crime is still the major concern in the country until now; day by day it keeps on increasing. **In 2018, according to the NCRB report (livemint, 2020), they found that crime rates in India have increased by 1.3%.** There had been a major surge in relation to abduction and kidnapping across the country; this increased by 10.3% over the previous years. Therefore, it is important to analyze the major crimes (abduction, kidnapping, rape, theft, etc) happening in every major state and their cities. This can help Law and Enforcement to analyze the trend behind the crime and in the future, they can prevent or stop it from happening. The crimes in urban areas have been rising, national crime records bureau released the data where cities in the northern state's records two times the higher crimes rates than the southern urban agglomerations of India(Jha, 2019). Crimes against women are the major concern for the government, violent crimes against women including rape are steadily rising every passing year India's GINI coefficient that has increased from 0.32 to 0.38 in the last two decades(Himabindu, Arora and Prashanth, 2015). In 2012, the crimes against women reported by official statistics increased by 24.7%, compared to those reported in 2008.

White-collar crime has acquired new dimensions. Political institutions have changed very rapidly and cultural norms have not kept pace with them. Hence, there is a 'cultural lag' in today's India, Power has also become a source of crime for the privileged sections of society. There is a tendency among powerful persons to abuse their influence and weity. Several cases of rape and murder have been reported by wards of influential persons and political heavyweights in the recent past. White-collar crime is a phenomenon found among educated people engaged in trade, professions, and government services(Mondal, n.d.).

One of the oldest civilizations in the world, India is one of the most popular destinations on the planet earth, In June 2019 with over 720 thousand tourists visiting the country(Madhumitha Jaganmohan, 5AD). According to the statistics the share of foreigners spending in the year 2017 was around 87% and it is expected to increase by 1% till 2028(Madhumita Jaganmohan, 2020). Safety of the tourist is a major concern for the government, this analysis will help them to understand that which state or city safer for the tourist, and which are unsafe, areas where the government needs to improve the situation. Top cities will be clustered according to most crimes committed in India this will help to identify the cities within the cluster. The defined cluster will divide the top cities with respect to the count of crimes committed, these

subgroups will help the police and the government to take the necessary measure to ensure the safety of the public. We will also forecast the total IPC crimes committed in India for the next 10 years. Forecasting or prediction of the crime is necessary for every state because it will help the local law enforcement to manage the crimes according to it.

. Another dataset will be used to plot the map of India, and that was also used to merge with the original dataset. All the trend graphs of the states will be discussed with the machine learning model. The graph made to identify the top crimes in India. The worst affected states with respect to the count of crimes will be plotted on the map of India. The worst affected cities by considering a threshold value for the top crimes and those cities will be plotted on the graph for better understanding. The last chapter of this study will discuss the conclusions and reflections, in this chapter we will write all relative objectives and their points and limitation of the project with future scope to make this analysis better.

1.2 AIMS AND OBJECTIVES

1.2.1 STUDY QUESTION

Can historical crime data for India be used to forecast the trend of crimes in each state of India for the next ten years?

1.2.2 OBJECTIVES

- To provide aggregate statistics from the dataset such as the highest crime areas.
- To represent key data and findings using a suitable visualization method and tool.
- To cluster applicable data.
- To forecast the trend of crimes for all the states for the next 9 years

CHAPTER 2: BACKGROUND STUDY

The total number of crimes and the crime rate has increased, Major reason for the increase in crime due to the population of Indian society has been increased and there is a rapid expansion in industrialization and urbanization as well over the 40 years.

Cases of Burglary has been decreased by 79.84% over the 53 years, but the murder rate has been increased by 7.39%. Robbery and riots have been declined by 28.85% and 10.58% respectively, but kidnapping has been increased by 47.8%.

The location has played a significant role in crimes in India:

- In 2006, the highest crime rate was reported in Puducherry (447.7%) for crimes under the Indian Penal Code which is 2.7 times the national crime rate of 167.7%.
- Kerala reported the highest crime rate at 312.5% among states.
- Delhi, Mumbai, and Bengaluru have accounted for 16.2%, 9.5%, and 8.1% crime rate among the megacities in India followed by Madhya Pradesh (Bhopal) 719.5% and Rajasthan (Jaipur) 597.1%.
- Uttar Pradesh reported the highest incidence of violent crimes accounting for 12.1% of total violent crimes in India (24,851 out of 2,05,656) followed by Bihar with 11.8% (24,271 out of 2,05,656).
- Among 35 megacities, Delhi reported 31.2% (533 out of 1,706) of total rape cases. Madhya Pradesh has reported the highest number of rape cases (2,900) accounting for 15.0% of total such cases reported in the country.
- Uttar Pradesh reported 10% (5,480 out of 54,800) of total murder cases in the country and 18.4% (4,997 out of 27,230) total attempts to murder cases. Today in India, crimes cases against Elderly women are on a rise mainly women are facing problem such as murder, theft, hurt, bag snatching.

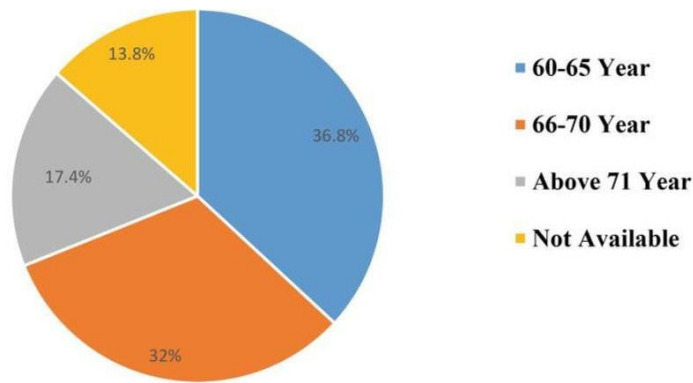


Figure 2.1 Age of Women

Women of age more 60 to 65 have been victimized the most, the chart suggests that 36.8% of women are the most affected.

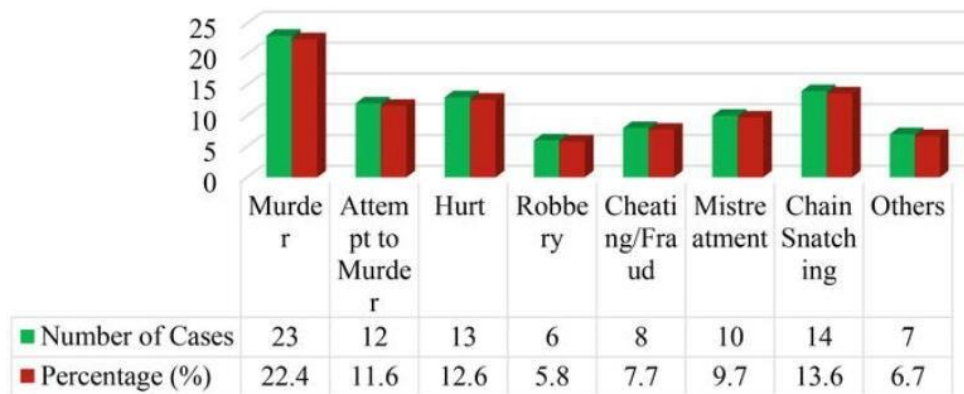


Figure 2.2: Nature of Crime

The nature of the crimes against the women is mostly murder with 22.4% followed by chain snatching and hurt with 13.6% and 12.6% respectively.

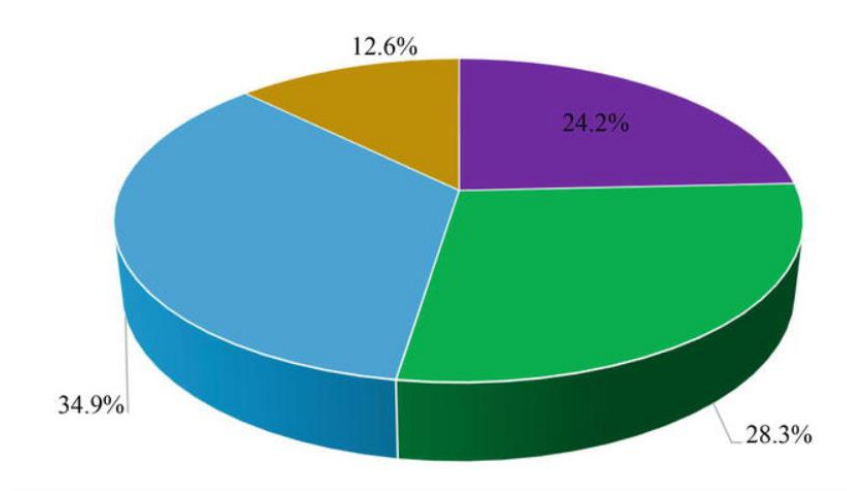


Figure 2.3: Percentage distribution on nature of crime

24.2% of crimes against women are committed by a known person such as family or relatives, 34.9% of crimes offenders are neighbors to the women. Hence it can be said that elderly women are most insecure in their families and neighborhood.

CHAPTER 3: REQUIREMENT ANALYSIS

3.1 MOTIVATION

High or increase crime levels make communities decline, as crimes reduce house prices, neighborhood satisfaction, and the desire to move in a negative manner. To reduce and prevent crimes, and prescribe solutions, Due to large volumes of data and the number of algorithms needed to be applied on crime data, it is unrealistic to do a manual analysis, Therefore, it is necessary to have a platform which is capable of applying any algorithm required to a descriptive, predictive and prescriptive analysis on large volume crime data

Through those three methodologies law-enforcement weities will be able to take suitable actions to prevent the crimes. Moreover, by predicting the highly likely targets to be attacked, during a specific period of time and specific geographical location, price will be able to identify better ways to deploy the limited resources and also to find and fix the problems leading to crimes.

Several applications are already developed for crime analysis. Most of these tools are developed to help the police to identify difference crime patterns and even to predict criminal activities. They are complex software which needs a lot of training before use. Designing a tool which is easy to use with minimum training would help law-enforcing bodies all around the world to reduce crimes.

CHAPTER 4: DETAILED DESIGN

4.1 SELECTION OF MODEL

The output from model training might be utilized for the deduction, which means making expectations on new data. A model is a refined portrayal of what a machine learning framework has learned. Machine learning models are likened to mathematical capacities they take permission in the form of input data, make a forecast or prediction on that input data, and then serve a reaction. In supervised and unsupervised machine learning, the model depicts the sign in the noise or the example distinguished from the training data.

A machine learning model that was used to forecast the crimes in India for the next 6 years for every state was ARIMA. The auto-regressive integrated moving average (ARIMA) is a time series-based model, the ARIMA model uses its variable on itself to predict the outcome. In the study, we used total IPC crimes committed and years to forecast the crime rates for the future.

4.2 TOOLS AND TECHNIQUES USED

- We used the most common programming language that is used for building the machine learning model i.e., Python.

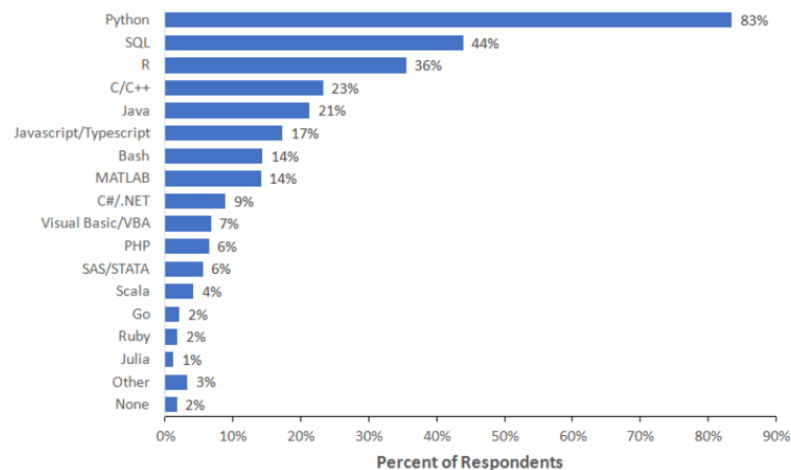


Figure 4.2.1 Languages Used

- The IDE (Integrated Development Environments) that was used to build the forecasting model was Jupyter Notebook. It was used with the help of Anaconda Command Prompt.
- We further deployed our project using flask framework.

4.3 LIBRARIES USED

LIBRARY	DESCRIPTION
Pandas	The Pandas is one of the most important libraries while dealing with data, all the data analysis, and data manipulation can be done with the help of pandas
Numpy	All the mathematical functions can be computed through the NumPy library.
Seaborn	Seaborn library is used for visualization.
Seasonal_decompose	This library was used to find the trend of crimes in every state.
Statsmodels.api	Statsmodels.api is used for modelling purpose, to use the forecasting machine learning algorithms, statsmodels.api provides all the necessary libraries in it.
Sklearn	The Scikitlearning (Sklearn) module contains all the machine learning algorithms e.g., Linear Regression Classification and clustering, in this project K-Means library was used to derive the cluster
Plotly	Plotly library provides an in-depth visualization of any data
Matplotlib	Matplotlib was used for visualization
Geopandas	Geopandas was used to draw the map of India. All the worst affected states in India were plotted on the map
ARIMA	ARIMA machine learning model was used for forecasting, this model was explained in the above section
Plot_acf, plot_pacf	The lags of every state were derived from the ACF. ACF stands for autocorrelation function.

Table 4.3.1 Libraries Used

CHAPTER 5: IMPLEMENTATION

5.1 DATA EXPLORATION

5.1.1 Steps involved:

- Importing the data, from data.goi website, as a CSV file.
- Importing all the necessary libraries for the study.
- Checking for all the information such as the number of columns and rows present in this dataset there are 33 columns present.
- Types of variables: 2 categorical and 30 numerical.
- Checking the count of all the crimes and states.
- Checking for the NULL or NAN value in the dataset.
- Merging the crime data with geographical data that was collected from the website GisDataCollection.

5.2 DATA VISUALISATION

Data visualization is the demonstration of taking (data) and setting it into a visual setting, for example, a guide or chart. Data visualizations make of all shapes and size data simpler for the human mind to understand, and visualization additionally makes it simpler to detect patterns, trends, and outliers in gatherings of data. Data visualization techniques will help to find in-depth insights from the data such as the worst affected state in India, the count of crimes that are being committed in India by criminals, and which is the most frequent crime, etc.

A pie chart is a method of showing information where a circle is partitioned into sections (or cuts") that mirror the relative size or recurrence of the classifications. Now, we're trying to identify the top crimes that are being committed in India by criminals. This help us to classify the count or frequency of each crime so that it can be further analyzed which district or which state is worst affected concerning the top crimes in India.

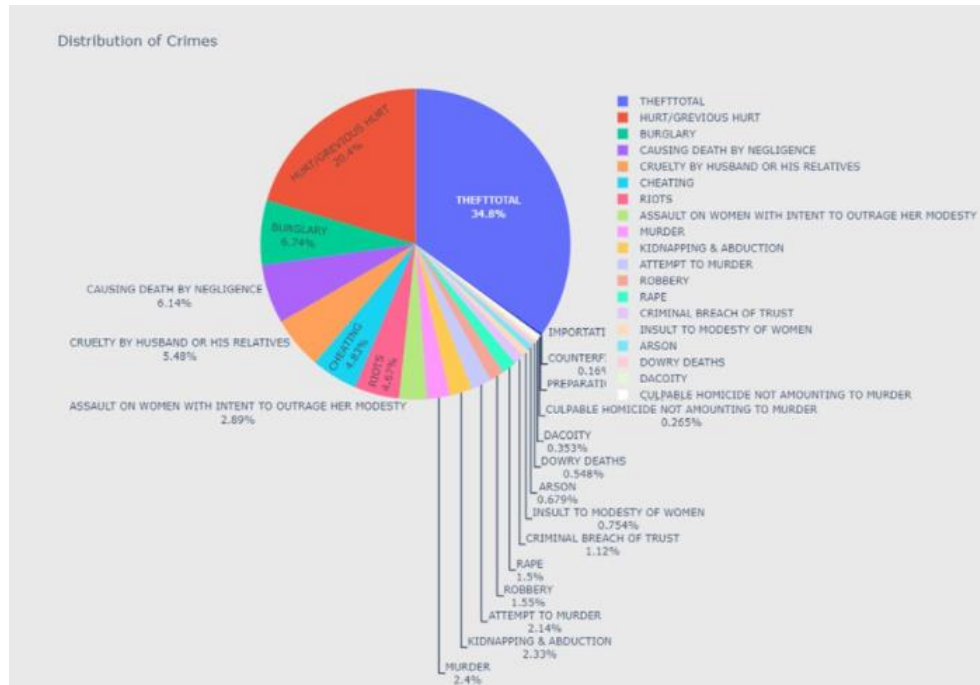


Figure 5.2.1: Distribution of Crimes

5.3. MODELLING APPROACH

5.3.1. ARIMA MODEL

An ARIMA model is a class of statistical models for analysing and forecasting time series data.

It explicitly caters to a suite of standard structures in time-series data, and as such provides a simple yet powerful method for making skilful time-series forecasts. ARIMA is an acronym that stands for Autoregressive Integrated Moving Average. It is a generalization of the simpler Autoregressive Moving Average and adds the notion of integration. This acronym is descriptive, capturing the key aspects of the model itself. Briefly, they are:

- **AR: AUTOREGRESSION.** A model that utilizes the dependent relationship between perception and some number of lagged perceptions.
- **I: INTEGRATED.** The utilization of differencing of crude perceptions (for example deducting an observation from the previous perception at the past time step) to make the time series stationary.
- **MA: MOVING AVERAGE.** A model that utilizes the reliance between a perception and a remaining mistake from a moving normal model applied to slack perceptions.

Standard notation is used of ARIMA (p, d, q) where the parameters are substituted with integer values to quickly implies the specific ARIMA machine learning model being used. Each of these parts of the model are explicitly specified in the model as a parameter.

- AR is a component of p lags
- Integration component(d)
- MA with q lags, MA(q)

5.4 MEASURING ACCURACY

The purpose of the machine learning models is to predict the outcome with the best accuracy score. In this study, we tried to forecast the total IPC crimes committed in India for the next 9 years. It was assumed that the prediction might vary, but finding the accuracy of the model with a good score was an important step. The two most popular techniques that were used to calculate the forecast accuracy are

- 1) Root Mean Square Error (RMSE)
- 2) Mean Absolute Percentage Error (MAPE).

5.5 K-MEANS CLUSTERING MODEL

For better management and to create the homogenous groups of entities most frequently is used application us Clustering. The clustering is a divide and conquer strategy which divides the dataset into a homogenous group. Clustering algorithms are unsupervised learning algorithms in which labels classes are not defined, unlike supervised machine learning.

The function we used to define the clustering technique was K-Means. The K-Means uses the Euclidean distance to find the distance between the two or more observation. It takes the mean(average) of every sample in the dataset and fit them accordingly.

Flask

Flask is a lightweight WSGI web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It began as a simple wrapper around Werkzeug and Jinja and has become one of the most popular Python web application frameworks.

Flask offers suggestions, but doesn't enforce any dependencies or project layout. It is up to the developer to choose the tools and libraries they want to use. There are many extensions provided by the community that make adding new functionality easy.

CHAPTER 6: EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we will discuss all the analysis and modeling that was done. Top crimes that were committed in India from 2001 to 2012 and to forecast the trend of crimes in every state in India.

6.1 CRIMES IN INDIA

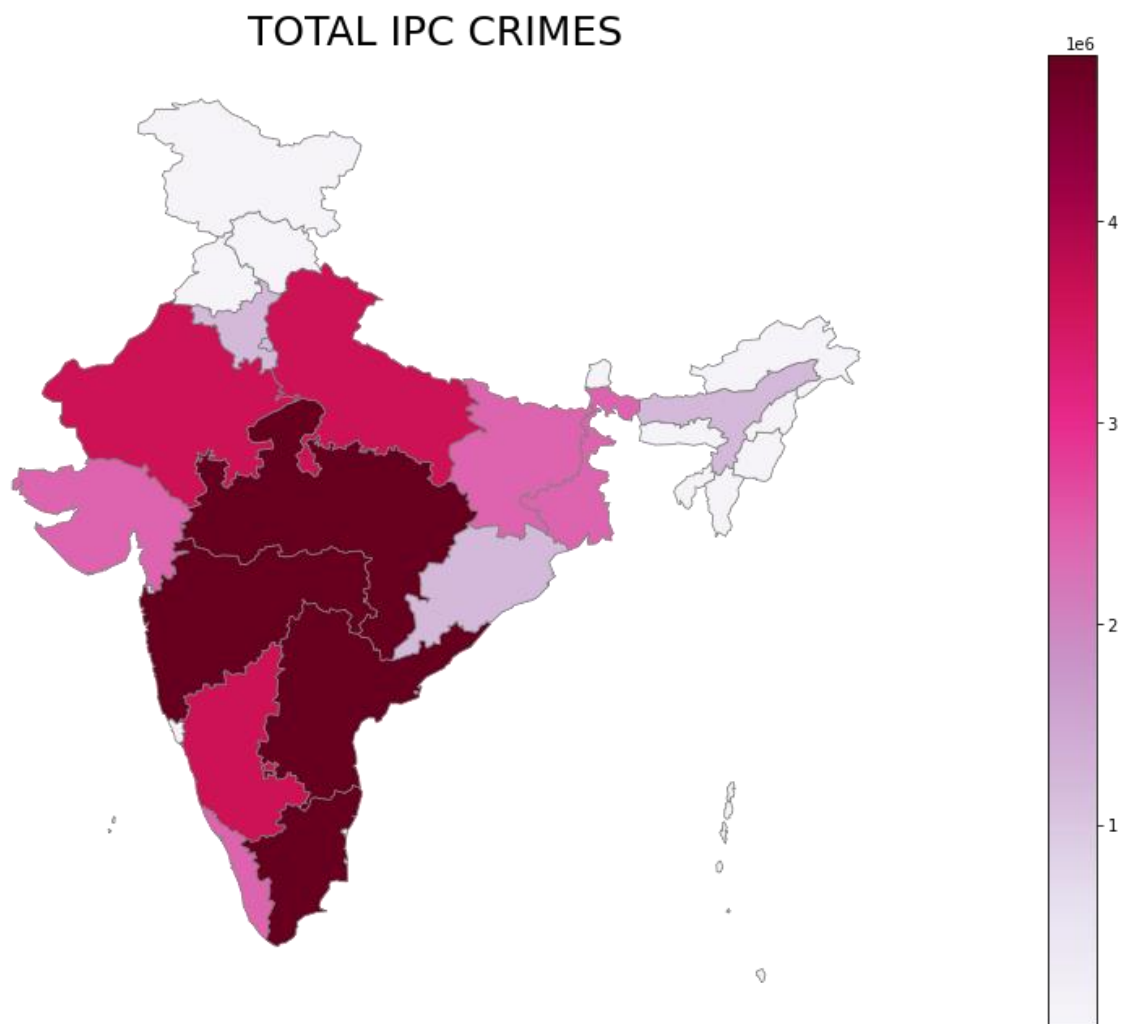


Figure 6.1.1: Distribution of crimes

The above figure shows the distribution of crimes that were committed in India from 2001 to 2013.

6.2 TOP 10 CRIMES IN INDIA

This pie chart helped We to identify the top 10 crimes that were committed in India. All the analysis further in the section was done using top 10 crimes in India

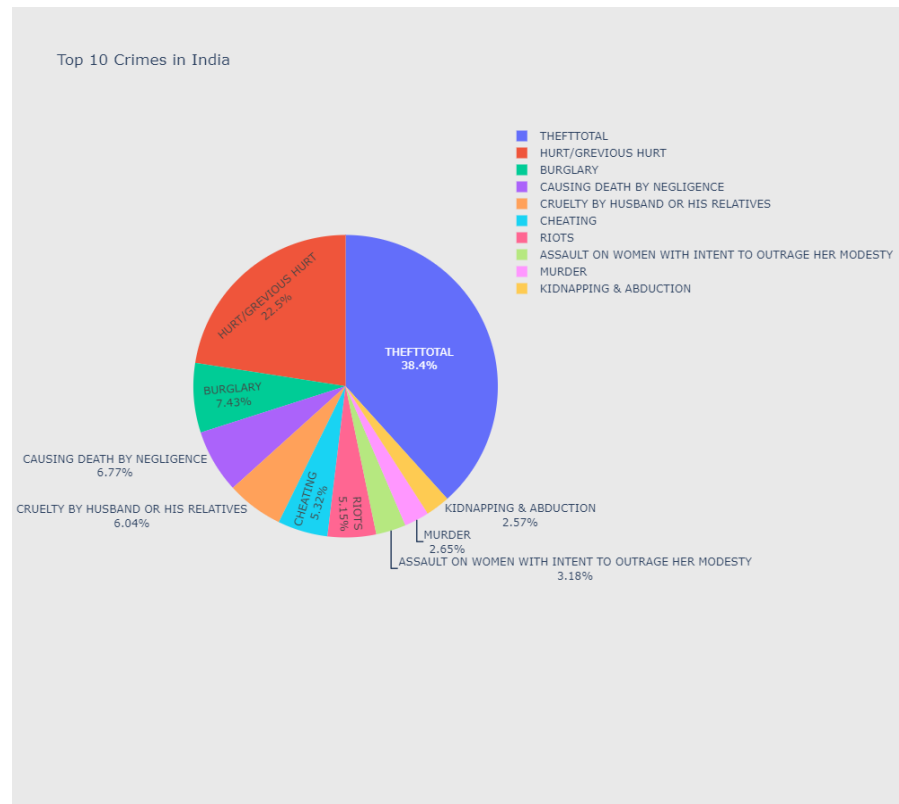


Figure 6.2.1: Top 10 crimes India

Identifying crimes plays an important role in keeping the country safe. In India the top 10 crimes that were committed between 2001 and 2013 were:

MURDER, KIDNAPPING & ABDUCTION, BURGLARY, RIOTS, CHEATING, HURT/GRIEVOUS HURT, ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY, CAUSING DEATH BY NEGLIGENCE, THEFTTOTAL, CRUELTY BY HUSBAND OR HIS RELATIVES.

We combined theft, auto theft, and other theft as the Theft total. It can be observed from the pie chart that the most committed crime in India was **THEFT** with 38.4% followed by **HURT** with 22.5%. The ratio of these two crimes was the most from 2001 to 2013. Crimes whose percentage was less than 2% were **MURDER**, **KIDNAPPING & ABDUCTION** with 2.65% and 2.57% respectively. **BURGLARY** with

7.43% and **CAUSING DEATH BY NEGLIGENCE** with 6.77%, and were third and fourth respectively.

At fifth position recoded crimes was **CRUELTY BY HUSBAND OR HIS RELATIVES** with 6.04%. Crimes that were below 6% was **CHEATING, RIOTS**, and **ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY** with 5.32%, 5.15%, and 3.18% respectively.

6.3 THEFT IN INDIA

One of the crime-category is plotted below:

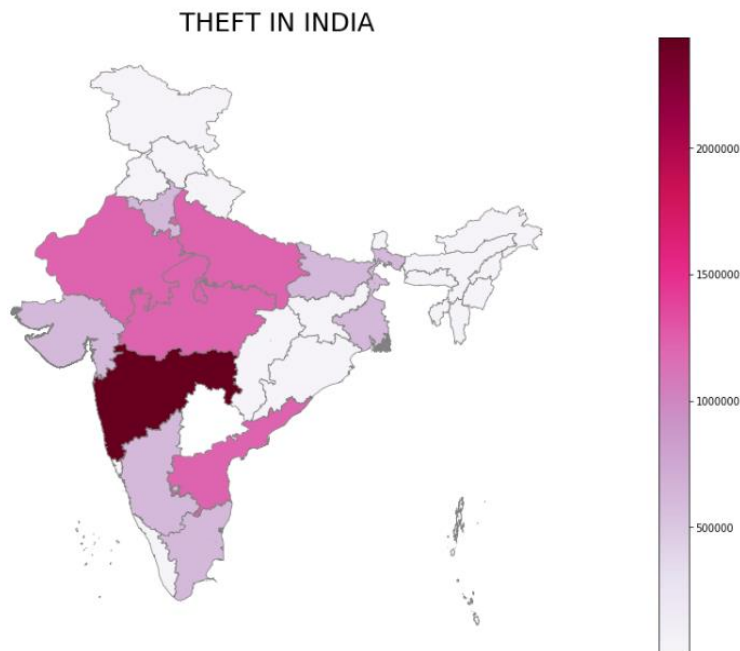


Figure 6.3.1: Theft in India

The theft was the most committed crime in India with a 38.4% rate. The worst affected state in India was Maharashtra in terms of theft with more than 2000000 was recorded. States between 1500000 and 2000000 were Madhya Pradesh, Rajasthan, Uttar Pradesh, and Andhra Pradesh.

6.4 FORECASTING

In this section, we build the machine learning model that was used to forecast the trend in every state for the next nine years. We used the ARIMA model to forecast the trend.

Forecasting of one of the states is shown below:

6.4.1 UTTARAKHAND

Uttarakhand is like the finest Chivalric Romance poetry when narrated, it melts the heart of the listener(Tourmyindia, n.d.). The population of Uttarakhand in 2019 was around 11,250,858(Statistics Times, 2020).

The figure below shows the trend of Uttarakhand from 2001 to 2013.

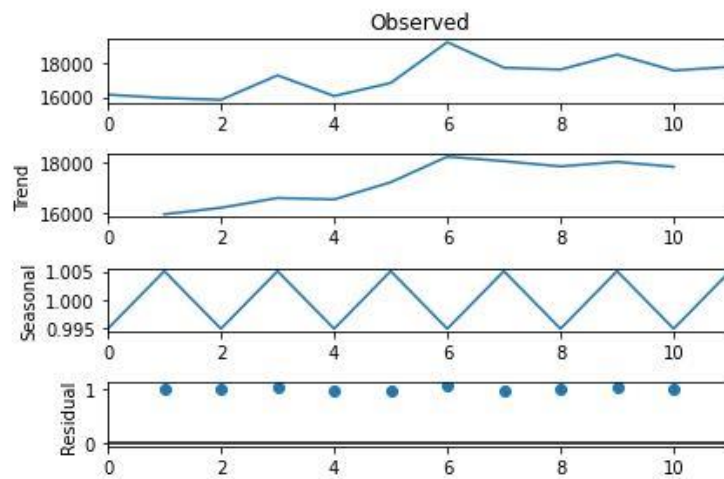


Figure 6.4.1.1: Trend in Uttarakhand

Crime in Uttarakhand followed the trend as shown in the above figure. The observed crime in this state was increasing from 2011. The trend line of Uttarakhand shows crimes will remain constant or might increase. In 2006 the state recorded the highest-amount of cases.

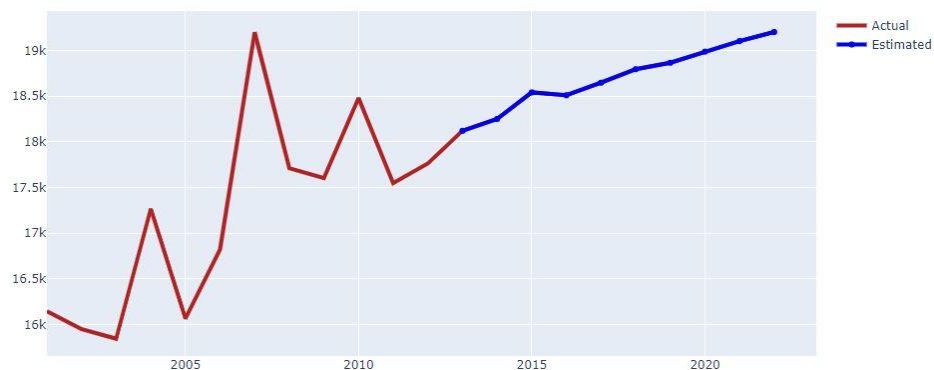


Figure 6.4.1.2: Forecast in Uttarakhand

6.5 CLUSTERING

Clustering of one of the states is shown below:

6.5.1 UTTAR PRADESH

From the given figure below, we tried to find the top crimes that were happening in the state Uttar Pradesh.

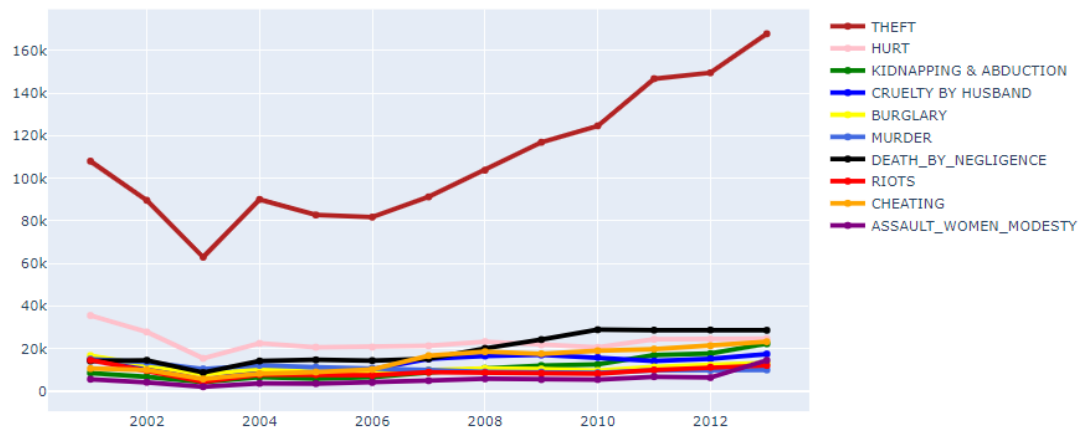


Figure 6.5.1.1: Crimes happening in Uttar Pradesh

From the graph, it can be observed that top crimes in this state were Theft, Hurt, and Death by Negligence. The theft was the most committed crime in this state with over 60000 cases reported every year followed by death by negligence and hurt with more than 15000 cases.

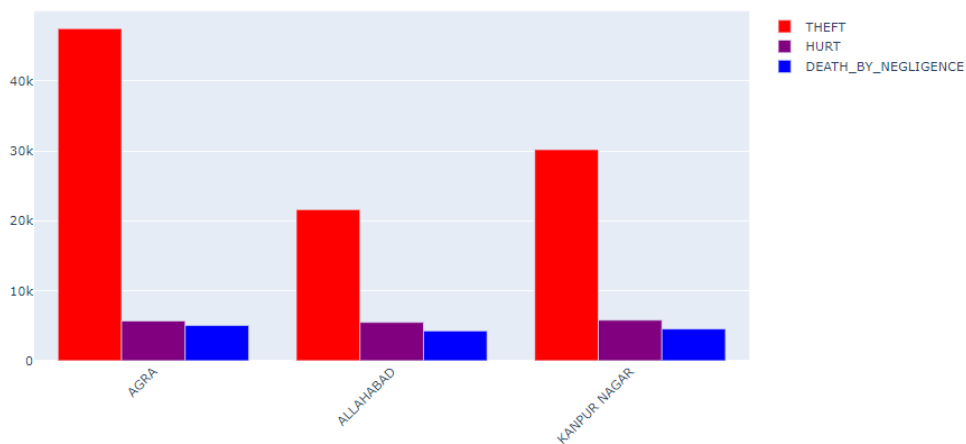


Figure 6.5.1.2: Cities Affected in Uttar Pradesh

Top three worst-affected cities in Uttar Pradesh were Agra, Allahabad, Kanpur based on the count of crime for theft and hurt.

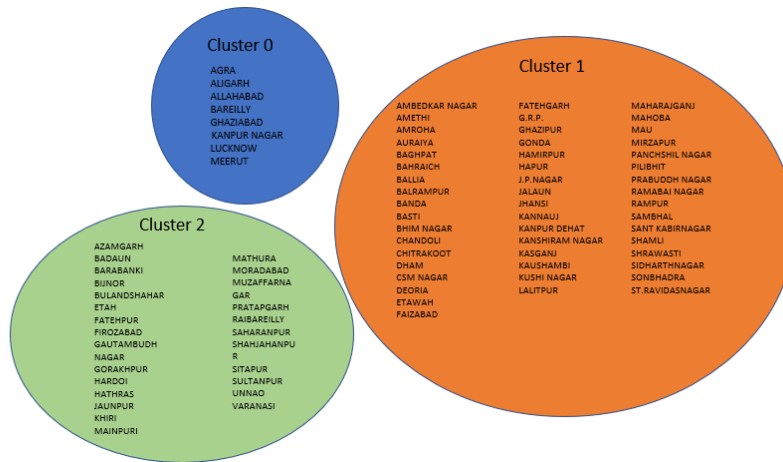


Figure 6.5.1.3: Clustered cities in Uttar Pradesh

This figure shows the cities that were divided into different clusters according to the crime rates that were happening in Uttar Pradesh. Cities in Cluster 0 were the worst affected with the count of other IPC crimes ranging more than 15000 and total-theft committed in these cities were more than 30000. Cities in Cluster 1 are less affected than cities in cluster 0. The count of crimes in these districts was less than cities in cluster 0. The least affected cities were sub-grouped in cluster 2.

CHAPTER 7: CONCLUSION AND FUTURE SCOPE

7.1 CONCLUSION

This study aimed to identify the following questions: Can historical crime data for India be used to forecast the trend of crimes in each state of India for the next nine years. This work did help to forecast the trend of crimes in each state of India with the error rate. Data used for the study was available from 2001 to 2013.

The **first objective** was to provide aggregate statistics from the dataset such as the highest crime areas. We achieved this objective by plotting the pie chart. Pie chart helped us to identify the top 10 crimes in India and plot every crime on the map of India. Frequency of crime distribution was obtained from this chart.

The **second objective** was to represent key data and findings using a suitable visualization method and tool. In this objective, we identified the worst top 5 states in India according to the total IPC crimes committed. Top 10 crimes in India were found. For the top 5 affected states.

The **third objective** was to cluster applicable data. This helped us to divide the district of top 5 states into similar groups. These groups helped in identifying the cities from the worst affected too least affected. Elbow curve method was used to identify the centroids of the cluster of every state and using that cluster was defined. The machine learning function that was used was the K-Means technique. Clustering helped to identify the districts according to the crimes.

The **fourth objective** was to forecast the trend of crimes for all the states for the next 9 years. The work carried out helped us to forecast the trend of crimes in each state of India with the error rate. We built the ARIMA model. Most of the states showed that the trend of crimes will increase over the nine years and some states showed that the crimes will decrease but not rapidly.

In the end, the study was successfully able to achieve all the objectives which were mentioned. Many factors affect the reason behind the crime like an increasing percentage of the unemployment rate or a steady rise in poverty. The clustering for the worst top 5 states in India was achieved successfully.

ARIMA model that was built for forecasting was working well for most of the states but some states showed a high error rate. States with high crime rates were also identified and were plotted on the map of India.

7.2 FUTURE SCOPE

An extension for future work for this task is further analyzing the examination utilizing the information from different pieces around the world to check if the finding from this project can be applied to that. If the author starts a fresh project again, he wished to have a category or any output variable to check what are the reasons behind crimes? Clustering Algorithm can be applied to all the states and union territories to gain more information about the crimes in Indian cities. The cluster will help to identify the in-depth worst-affected cities.

REFERENCES

<https://indianexpress.com/article/opinion/columns/crime-data-in-india-6035032/>

<https://machinelearningmastery.com/arma-for-time-series-forecasting-with-python/>

<https://www.lawctopus.com/academike/crime-crime-rates/>

https://data.gov.in/catalog/district-wise-crimes-under-various-sections-indian-penal-code-ipc-crimes?filters%5Bfield_catalog_reference%5D=87615&format=json&offset=0&limit=6&sort%5Bcreated%5D=desc/

<https://chartio.com/learn/charts/line-chart-complete-guide/>

<https://vciba.springeropen.com/articles/10.1186/s42492-021-00075-z/>