ZKG INTERNATIONAL ISSN: 2366-1313

CRIME DATA ANALYSIS

P.Nageswara rao, T. Yaswanthsai, V. Sai Ram Nikhil, K. Siddu, S. Janaki

Priyadarshini Institute of Technology and Science, Chintalapudi,Guntur, Andhra Pradesh, India - 522306.

Under Graduate, Priyadarshini Institute of Technology and Science,Chintalapudi, Guntur, Andhra Pradesh, India - 522306.

Abstract:

The data set includes criminal offenses in the City and Country for the previous five calendar years plus the current year to date. The data is based on the National Incident Based Reporting System (NIBRS) which includes all victims of person crimes and all crimes within an incident. The data is dynamic data, which allows for additions, deletions and/or modifications at any time, resulting in more accurate information in the database. The information provided here regarding public safety in city are offered as a courtesy by the City and Country.

The analysis helps the government authorities to have control about the crimes and happening areas with lot more security features. So from the above analysis the government can easily predict the areas where the crimes are happening more and what types of crimes. So all the possible crimes are been listed and the same can be implemented to other data analytics. Here we propose a system that analyzes, detects, and predicts the crime possibilities within a region. This paper explores various types of criminal analysis and crime predictions using various data mining techniques including decision trees, linear regression, k-means clustering and word cloud. crime data analysis plays a crucial role in preventing the criminal activities and ensuring the public safety.

ISSN: 2366-1313

Introduction:

Crimes are the significant threat to the humankind. There are many crimes that happens regular interval of time. Perhaps it is increasing and spreading at a fast and vast rate. Crimes happen from smallvillage, town to big cities. Crimes are of different type – robbery, murder, rape, assault, battery, false imprisonment, kidnapping, homicide. Since crimes are increasing there is a need to solve the casesin a much faster way. The crime activities have been increased at a faster rate and it is the responsibility of police department to control and reduce the crime activities. Crime prediction and criminal identification are the major problems to the police department asthere are tremendous amount of crime data that exist. There is a need of technology through which the case solving could be faster.

The Federal Bureau of Investigation (FBI) defines a violent crime as an offense which involves force or threat. The FBI's Uniform Crime Reporting (UCR) program categorizes these offenses into four categories: murder, forcible rape, robbery, and aggravated assault. The FBI UCR program defines each of the offenses as follows: (i) Murder - The UCR does not include deaths caused by accident, suicide, negligence, justifiable homicides and attempts to murder or assaults to murder (which are scored as aggravated assaults), in this offense classification. (ii) Forcible Rape - Rape is a sexual attack on a female against her will. Though attempts or assaults to commit rape by threat or force are considered crime under this category, statutory rape (without force) and other sex offenses are excluded . (iii) Robbery - The taking or attempting to take anything of value from the care, custody, or control of a person or persons by force or threat of force or violence and/or by putting the victim in fear. (iv) Aggravated Assault It is the unlawful attack conducted by one person upon another to inflict severe or aggravated bodily injury. The UCR

ISSN: 2366-1313

program specifies that an aggravated assault usually involves the use of a weapon

or other means to produce death or great bodily harm. Attempted aggravated assaults that involves the use of guns, knifes and other weapons are considered to belong to this category because if the assault were completed, it would have lead to serious personal injury. An offense that involves both aggravated assault and larceny-theft occurring together, the offense is considered to belong to the category of robbery. Unfortunately, these type of crimes seem to have become common place in the society. Law enforcement officials have turned to data mining and machine learning to aid in the fight of crime prevention and law enforcement. In this research, we implemented the Linear Regression, Additive Regression, and Decision Stump algorithms using the same finite set of features, on the communities and crime un normalized dataset to conduct a comparative study between the violent crime patterns from this particular dataset and actual crime statistical data for the state of Mississippi that has been provided by neighborhoodscout.com. The crime statistics used from this site is data that has been provided by the FBI and had been collected for the year 2013. Some of the statistical data that was provided by neighborhoodscout.com such as the population of Mississippi, population distribution by age, number of violent crimes committed, and the rate of those crimes per 100K people in the population are also features that have been incorporated into the test data to conduct analysis. enforcement officials have turned to data mining and machine learning to aid in the fight of crime prevention and law enforcement.

In this research, we implemented the Linear Regression, Additive Regression, and Decision Stump algorithms using the same finite set of features, on the communities and crime un normalized dataset to conduct a comparative study between the violent crime patterns from this particular dataset and actual crime statistical data for the state of Mississippi that has been provided by neighbourhoodscout.com.

ISSN: 2366-1313

The crime statistics used from this site is data that has been provided by the FBI and had been collected for the year 2013. Some of the statistical data that was provided by neighborhoodscout.com such as the population of Mississippi, population distribution by age, number of violent crimes committed, and the rate of those crimes per 100K people in the population are also features that have been incorporated into the test data to conduct analysis. criminal data, law enforcement agencies can use models that may be important in the crime prevention process.

Criminology is an area where the scientific study of crime and criminal behavior focuses. This is one of the most important areas when applying data mining techniques that can produce significant results. Crime analysis, as part of criminology, is tasked with exploring and discovering crime and its relationship with criminals. Law enforcement is a process that aims to identify the characteristics of crime. Identifying crime characteristics is the first step in developing further analysis. The high volume of crime data and the complexity of the relationships between them have made criminology an appropriate field for applying data mining techniques. Data mining can be used to examine many large datasets involving a large set of variables beyond what a single analyst, or even an analytical team or task force, can consider correct, whereas machine learning uses neural networks, predictive model and automated algorithms to make the decisions. Like any other problem solving method, the task of data mining begins with a problem definition. The identification of the data mining problem enables the determination of the data mining process and the modeling technique. Machine learning is a subfield of data science that deals with algorithms able to learn from data and make accurate predictions. Data mining gives law enforcement agencies the opportunity to learn about crime trends how and why crimes are committed.

Using data mining methods and machine learning improves crime analysis and help reduce and prevent crime.

ISSN: 2366-1313

Related Work:

1. **Crime Prediction Using Data Mining:**
   - Crime prediction involves identifying regions with a high probability of crime occurrences and visualizing crime-prone areas.
   - Data mining techniques extract previously unknown, useful information from unstructured data.
   - Crime data collected from various sources (blogs, news, websites) serves as a record for creating a crime report database.
   - Techniques like decision trees, linear regression, and k-means are used for crime prediction.
2. **Analysis and Prediction of Crime Patterns Using Big Data**:
   - Visualization, supervised learning, statistical approaches, and unsupervised learning play a crucial role in analyzing crime data.
   - Detecting hidden patterns among committed crimes at different times helps law enforcement agencies take proactive measures.
3. **Overview of Crime Analysis**:
   - Crime analysis involves personnel analyzing crime reports, identifying crime hotspots, and studying serial crime data.
   - The analytical function assesses potential threats posed by crime groups or criminal activities and recommends intervention methods.
4. **Machine Learning for Crime Prediction**:
   - Machine learning and data science techniques are applied to crime datasets.
   - Crime information, including location, type, date, time, latitude, and longitude, is extracted from official police portals.

In summary, collection, analysis and prediction of crime-related data

ISSN: 2366-1313

are imperative for securing communities and improving law enforcement practices.

ISSN: 2366-1313

Methodology:

Data Collection:

- Gather relevant data from various sources, including police reports, incident records, witness statements, and surveillance footage.
- Ensure data quality, accuracy, and completeness.

Classification:

- Organize crime data into meaningful categories (e.g., theft, assault, burglary).
- Assign appropriate labels to each incident for further analysis.

Pattern Identification:

- Identify recurring patterns, trends, and anomalies in crime data.
- Use statistical techniques to uncover hidden relationships.

Prediction:

- Employ machine learning algorithms to predict future criminal activities.
- Factors such as time, location, and historical data contribute to predictive models.

Visualization:

- Create visual representations (charts, graphs, maps) to convey insights.
- Geographic crime mapping helps identify hotspots and allocate resources effectively.

GIS (Geographic Information System):

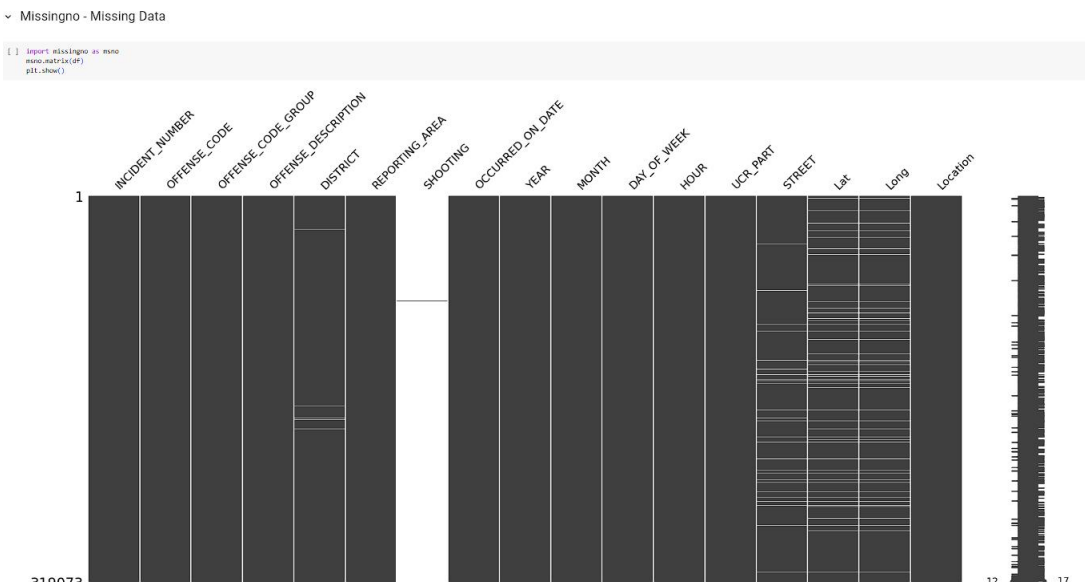
- Utilize GIS tools for spatial analysis and visualization.
- Components include point features (crime incidents), line features (streets), polygon features (neighborhoods), and image features (satellite imager).

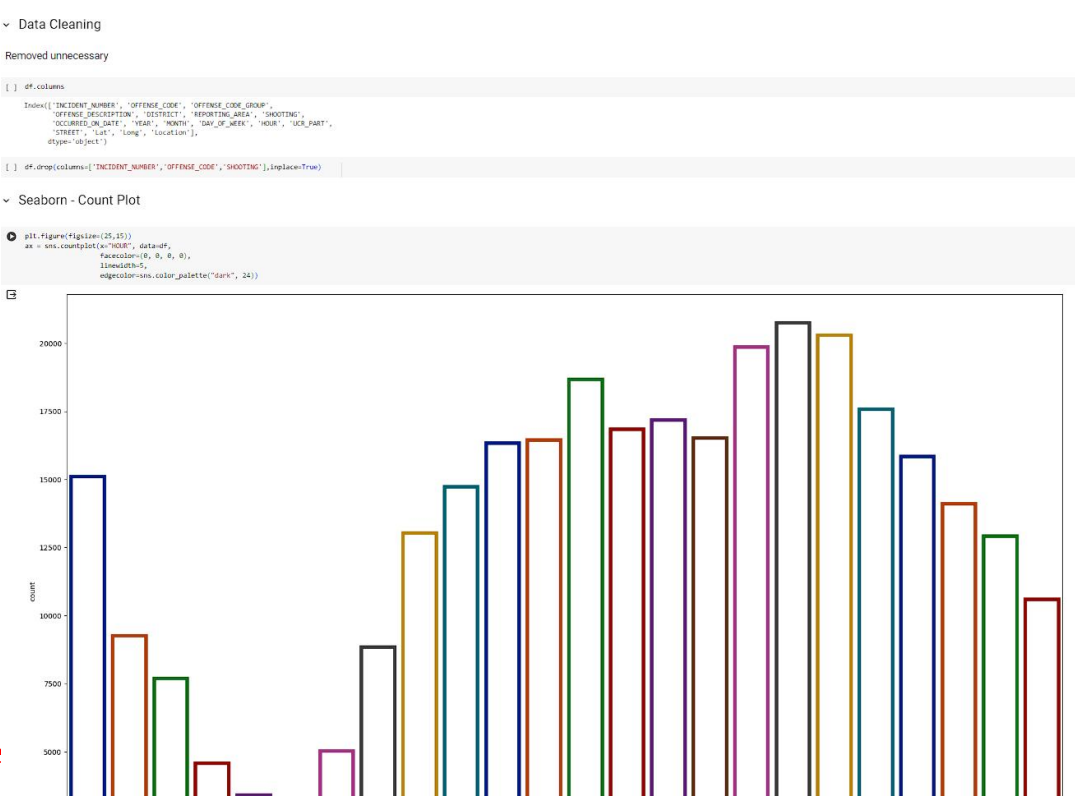
we have to remember that crime analysis combines both qualitative and quantitative approaches to enhance public safety and inform law enforcement strategies.

ISSN: 2366-1313

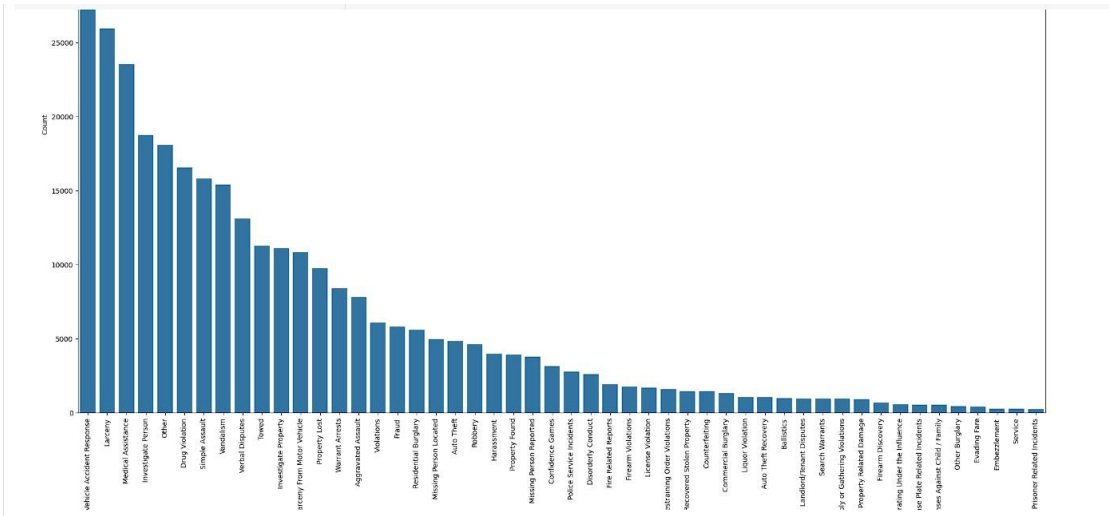
Result:

The result can be obtained by executing the code.The output can be obtained after processing the data in different forms. To understand that visualizations easily, the occurred output patterns will be in different colors based on years or months. The patterns are of any type like bar graphs, pie charts etc





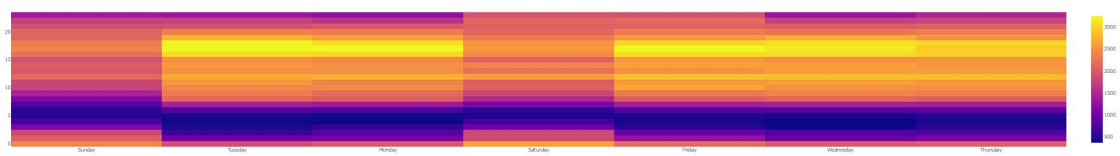
ISSN: 2366-1313



Plotly - 2D Histogram (Interactive)

```
[ ] x = df.DAY_OF_WEEK
    y = df.HOUR

fig = go.Figure(go.Histogram2d(
        x=x,
        y=y
    ))
fig.show()
```



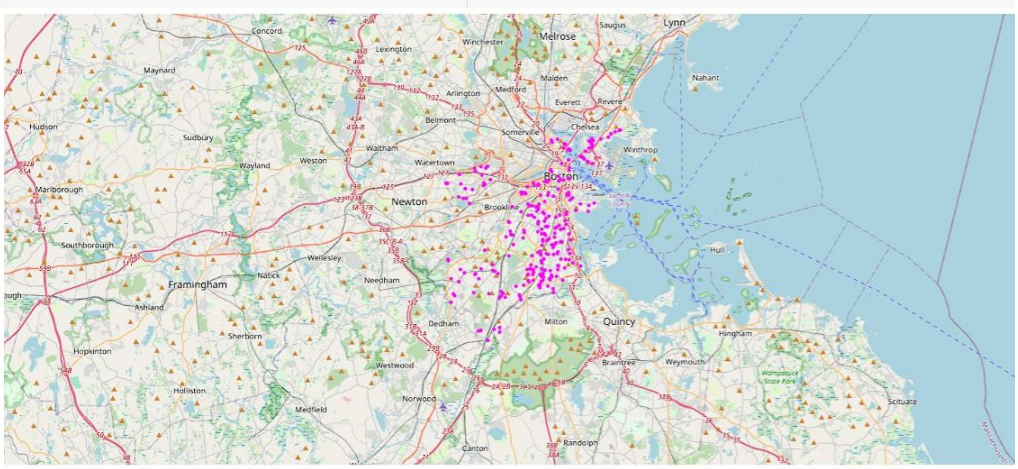
Plotly - Donut Chart (Interactive)

```
[ ] labels = df.DAY_OF_WEEK.unique()
values=[]
for each in labels:
    values.append(len(df[df.DAY_OF_WEEK==each]))

# Use `hole` to create a donut-like pie chart
fig = go.Figure(data=[go.Pie(labels=labels, values=values, hole=.3)])
fig.show()
```

Plotly - Map Box (Interactive)

```
fig = px.scatter_mapbox(df[df["OFFENSE_CODE_GROUP"]=="Service"], lat="Lat", lon="Long", hover_name="HOUR", hover_data=["YEAR", "HOUR"],
                        color_discrete_sequence=["fuchsia"], zoom=10, height=600)
fig.update_layout(mapbox_style="open-street-map")
fig.update_layout(margin={"r":0,"t":0,"l":0,"b":0})
fig.show()
```



Folium - Map

ISSN: 2366-1313

Conclusion:

We built a model that gives us the expected number of crimes that are happened in country in each day in a given time interval. This model can be built for any city or state in the world provided that there is data to analyse. We can predict the estimated time for the crime to take place as a future scope. And the law enforcement agency department have to take the graphical representations then they have to analyse the different crimes, Along with this, one can try to predict the rate of crime. We have to identify the number of problem areas that require additional security attention in the attempt to reduce rates of violent crime in the city.

References:

- Violent-Crime. http://www2.fbi.gov/ucr/cius2009/offenses/violent_crime/.
- Murder.
- http://www2.fbi.gov/ucr/cius2009/offenses/violent_crime/murder_homicide.html.
- Forcible Rape.
- http://www2.fbi.gov/ucr/cius2009/offenses/violent_crime/forcible_rape.html.
- Robbery. http://www2.fbi.gov/ucr/cius2009/offenses/violent_crime/robbery.html
- Assault.
- http://www2.fbi.gov/ucr/cius2009/offenses/violent_crime/aggravated_assault.html.
- Mississippi Crime Rates and Statistics - NeighborhoodScout. Mississippi Crime Rates and Statistics -NeighborhoodScout. Accessed February 17, 2015. http://www.neighborhoodscout.com/ms/crime/.
- S. M. Nirkhi, R.V. Dharaskar and V.M. Thakre. "Data Mining : A Prospective Approach for DigitalForensics," International Journal of

ISSN: 2366-1313

Data Mining & Knowledge Management Process, vol. 2, no. 6 pp.41-48, 2012.

- Weka 3: Data Mining Software in Java. http://www.cs.waikato.ac.nz/ml/weka/.
- E. W. T. Ngai, L. Xiu, and D. C. K. Chau. "Application of Data Mining Techniques in CustomerRelationship Management: A Literature Review and Classification," Expert Systems withApplications, pp. 2592–2602, 2008.
- J. McCarthy, "Arthur Samuel: Pioneer in Machine Learning," AI Magazine, vol. 11, no. 3, pp. 10-11,1990.
- R. Bermudez, B. Gerardo, J. Manalang and B. Tanguilig, III. "Predicting Faculty Performance UsingRegression Model in Data Mining," Proceedings of the 9th International Conference on SoftwareEngineering Research, Management and Applications, pp. 68-72, 2011.
- S. Sathyadevan and S. Gangadharan. "Crime Analysis and Prediction Using Data Mining,"Proceedings of the 1st International Conference on Networks and Soft Computing,pp. 406-412, 2014.
- Communities and Crime Un normalized Dataset. UCI Machine Learning-Repository.https://archive.ics.uci.edu/ml/datasets/Communities and Crime Unnormalized.