

INSIGHTS PREDICTION THROUGH EXPLORATORY DATA ANALYSIS (CRIME DATA ANALYSIS)

A PROJECT REPORT

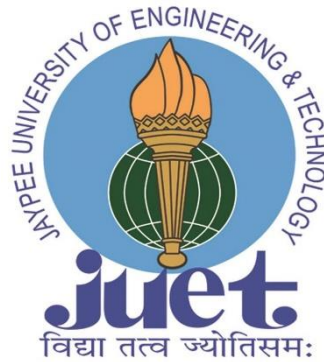
Submitted By:

Devanshi Mathan - 221B142

Lakshya Jha - 221B218

Yashasvi Grover - 221B460

Under the guidance of: Dr. Gaurav Saxena



August 2024 – December 2024

Submitted in partial fulfillment for the award of the degree

of

Bachelor of Technology

IN

Computer Science and Engineering

Department of Computer Science & Engineering

**JAYPEE UNIVERSITY OF ENGINEERING &
TECHNOLOGY, AB ROAD, RAGHOGARH, DT. GUNA-
473226 MP, INDIA**

Declaration by the Student

I hereby declare that the work reported in the B. Tech. the project entitled **“Insights Prediction through Exploratory Data Analysis (Crime Data Analysis)”**, in partial fulfillment for the award of the degree of Bachelor in Technology in Computer Science & Engineering submitted at Jaypee University of Engineering and Technology, Guna, as per best of my knowledge and belief there is no infringement of the intellectual property right and copyright. In case of any violation, I will solely be responsible.

Devanshi Mathan(221B142)

Lakshya Jha(221B218)

Yashasvi Grover(221B460)

Department of Computer Science and Engineering
Jaypee University of Engineering and Technology
Guna, M.P., India

Date - 19/11/2024



JAYPEE UNIVERSITY OF ENGINEERING & TECHNOLOGY

Grade 'A+' Accredited with by NAAC & Approved U/S 2(f) of the UGC Act, 1956
A.B. Road, Raghogarh, Dist: Guna (M.P.) India, Pin-473226
Phone: 07544 267051, 267310-14, Fax: 07544 267011
Website: www.juet.ac.in

CERTIFICATE

This is to certify that the work titled **“Insights Prediction through Exploratory Data Analysis (Crime Data Analysis)”** submitted by **“Devanshi Mathan (221B142), Lakshya Jha (221B218), Yashasvi Grover (221B460)”** fulfillment for the award of degree of Bachelors in Technology in Computer Science & Engineering of Jaypee University of Engineering & Technology, Guna has been carried out under my supervision. As per best of my knowledge and belief there is no infringement of intellectual property right and copyright. Also, this work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma. In case of any violation concern student will solely be responsible.

Signature of Supervisor

Dr. Gaurav Saxena

Assistant Professor, Dept. of CSE

ACKNOWLEDGEMENT

Any endeavor cannot lead to success unless and until a proper platform is provided for the same. This is the reason we find ourselves very fortunate to have undergone our major project work under the supervision of **Dr. Gaurav Saxena**. Our sincere gratitude to him, for having faith in us and thus allowing us to carry out a project on a technology completely new to us, for which we had to research and learn many new things, which will help us deal with advanced work in future.

Secondly, we would like to thank the Department of Computer Science and Engineering that created this opportunity. Last but not the least we want to thank our friends and family who continuously encouraged and helped us in any of the possible ways they could.

Devanshi Mathan(221B142)

Lakshya Jha(221B218)

Yashasvi Grover(221B460)

SUMMARY

Project Objective:

The "Insights Prediction through Exploratory Data Analysis (Crime Data Analysis)" project aims to provide comprehensive insights into crime patterns and trends by utilizing advanced data analysis and machine learning techniques. This project focuses on leveraging diverse crime datasets to identify patterns, hotspots, correlations, and fluctuations in crime frequency, enabling better understanding and support for public safety strategies and crime prevention initiatives.

Key Achievements:

The project successfully employed statistical methods and data visualization techniques to uncover complex crime patterns, including temporal, spatial, and demographic dimensions. By analyzing these facets, it produced valuable insights that inform policy-making, enhance law enforcement strategies, and aid in effective resource allocation. Through systematic exploration, it transcends basic statistics to deliver actionable insights into crime trends and their underlying factors.

Conclusion:

While the project demonstrates the potential of data-driven crime analysis to contribute to community safety and urban planning, challenges such as data quality, and model interpretability were recognized. This initiative advances crime analysis methodologies, contributing meaningfully to a data-informed approach for addressing crime.

LIST OF FIGURES

S.No.	Title	Page No.
1	Fig 1.1: Crime Data Analysis: Visual Representation	1
2	Fig 1.2: Overview of Data Analysis Tools and Workflow	4
3	Fig 2.1: Python Libraries for Data Analysis and Visualization	11
4	Fig 2.2: Power BI Trademark	14
5	Fig 3.1: Data Analysis Tools	18
6	Fig 4.1: Work-flow of System architecture	24
7	Fig 4.2: Steps for performing EDA	27
8	Fig. 4.3: Code Snippet for yearly analysis	27
9	Fig. 4.4: Code Snippet for monthly analysis	28
10	Fig. 4.5: Code Snippet for hourly analysis	28
11	Fig. 4.6: Bar graph for top offence code group	29
12	Fig. 4.7: Visual representation of crime incidents across districts	30
13	Fig. 4.8: Visual representation of top crime prone streets	30
14	Fig. 4.9: Visual representation of monthly shooting incidents	31
15	Fig. 4.10: Visual representation of crime incidents across UCR parts	32

16	Fig. 4.11: Correlation Matrix of selected variables	33
17	Fig. 5.1 Representation of Incidents by Street	38
18	Fig. 5.2 Representation of Incidents by Hours of Day	39
19	Fig. 5.3 Representation of Incidents by Day of Week	40
20	Fig. 5.4 Representation of Incidents by District	41
21	Fig. 5.5 Representation of Incidents by UCR Part	42
22	Fig. 5.6 Representation of Shooting on Street	43
23	Fig. 5.7 HeatMap for Offence Group	44
24	Fig. 5.8 Representation of Incidents by Month	45

TABLE OF CONTENTS

Declaration by the Students	i
Certificate	ii
Acknowledgment	iii
Executive Summary	iv
List Of Figures	v

1. INTRODUCTION

1.1 Project Overview	1
1.2 Objectives and Scope	2
1.3 Significance of the Project	3
1.4 Methodology	3
1.5 Project Timeline	4
1.6 Team Composition and Roles	5
1.7 Target Audience	6
1.8 Project Constraints	6
1.9 Innovation and Contribution	7
1.10 Summary of Chapters	8

2. RELATED WORK / LITERATURE SURVEY

2.1 Overview of Crime Data Analysis	9
2.2 Traditional Crime Analysis Techniques	9
2.3 Modern Data Analysis Techniques in Crime Prediction	9

2.4 Python and Data Analysis	10
2.5 Existing Tools and Technologies for Crime Data Analysis	10
2.6 Libraries for Data Analysis and Visualization	11
2.7 Data Preprocessing Techniques	12
2.8 Exploratory Data Analysis (EDA)	13
2.9 Role of MS Power BI in Crime Data Visualization	13
2.10 Use of MS Excel in Data Management	14
2.11 Challenges in Crime Data Analysis	15
2.12 Data Security and Privacy in Crime Data Analysis	16
2.13 Ethical Considerations in Crime Analysis	16
2.14 Gaps in Existing Literature and Future Directions	16
2.15 Summary of Theoretical Background and Technology	17

3. REQUIREMENT ANALYSIS

3.1 Hardware Requirements	18
3.2 Software Dependencies	18
3.3 Network Requirements	20
3.4 Analyst Skill-set and Expertise	20
3.5 Data Privacy and Compliance	20
3.6 Data Collection and Quality Analysis	21
3.7 Risk Analysis	21
3.8 Project Feasibility Study	21
3.9 Resource Allocation	21
3.10 Budget Analysis	22

3.11 Ethical Considerations	22
3.12 Summary of Requirement Analysis	22
4. DESIGN AND IMPLEMENTATION	
4.1 Implementation Strategy	23
4.2 System Architecture Overview	23
4.3 Data Collection and Preprocessing	25
4.4 Data Analysis and Insights Generation	25
4.5 Data Visualization Using Power BI	26
4.6 Exploratory Data Analysis	26
4.6.1 Distribution of Incidents Over Time	27
4.6.2 Analysis by Offense Code and Group	28
4.6.3 Spatial Analysis of Crime Incidents	29
4.6.4 Shooting Incidents Analysis	31
4.6.5 Categorization Based on UCR Part	31
4.6.6 Correlation Analysis	32
4.7 Performance Optimization	33
4.8 Risk Analysis and Mitigation	34
4.9 Testing and Validation	34
4.10 Summary of Design and Implementation	35
5. RESULT AND CONCLUSION	
5.1 Introduction	36
5.2 Key Findings from the Exploratory Data Analysis	36

5.3 Correlation Between Variables	37
5.4 Power BI Dashboard	38
5.5 Summary of Observations	46
5.6 Implications for Law Enforcement and Public Policy	47
5.7 Proposed Solutions	48
5.8 Action Plan for Implementation	49
5.9 Challenges and Limitations	50
5.10 Summary of Results and Insights	51
 APPENDICES	
Appendix A	52
Appendix B	56
Appendix C	58
 REFERENCES	
	59
 PERSONAL DETAILS	
	60