Crime Analysis and Prediction Using K-Means Clustering

By Salvador Mendiola

Table of contents

1
Introduction
Problem Statement, Background, Motivation

2
Contribution
Research Objectives

3
Methodology
Experimental setup

4
Results
Data analysis and visualization

5
Conclusion
Summary of Research, Future work

1

Introduction

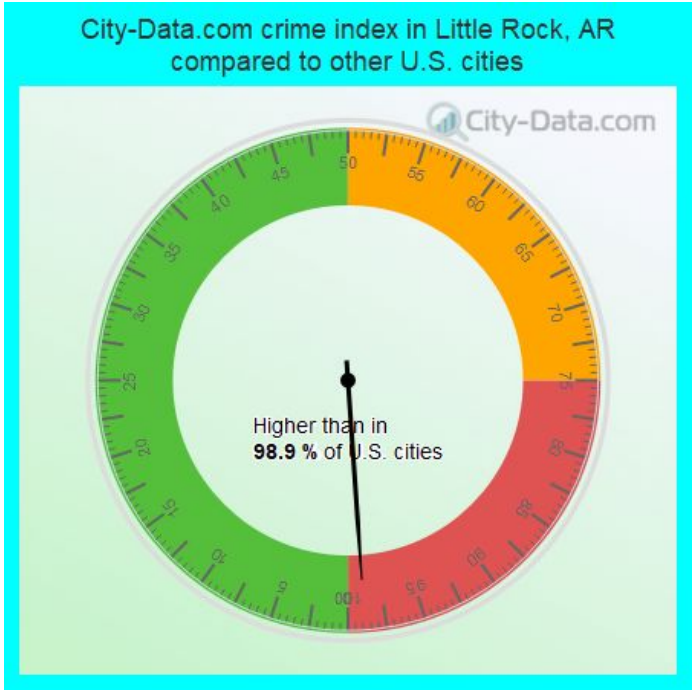
Background & Motivation, Problem Statement

Background & Motivation

28.64%

Increase in crime rate for the year 2020. 2019 only saw a 1.19% increase

Crime in Little Rock



Little Rock crime 2020

Crime Prediction methods

01

Confusion Matrix

02

Geo Mapping

03

Analysis of human mobility flows

04

Clustering

About K-Means clustering

Unlabeled Data

Used to find several hidden patterns in data

Visualization

Previous research lacks good visuals

Problem Statement:

Crime rates are increasing, and the use of K-means clustering in crime prediction as a visualization tool is often overlooked

2

Contribution

Research objectives

Purpose of Research:

Research Objectives

1. To display in detail which areas of Little Rock have the most occurrences of crime with the use of k-means
2. To predict which areas should see an increase in law enforcement.
3. To show which type of crimes are most common in the city of Little Rock.

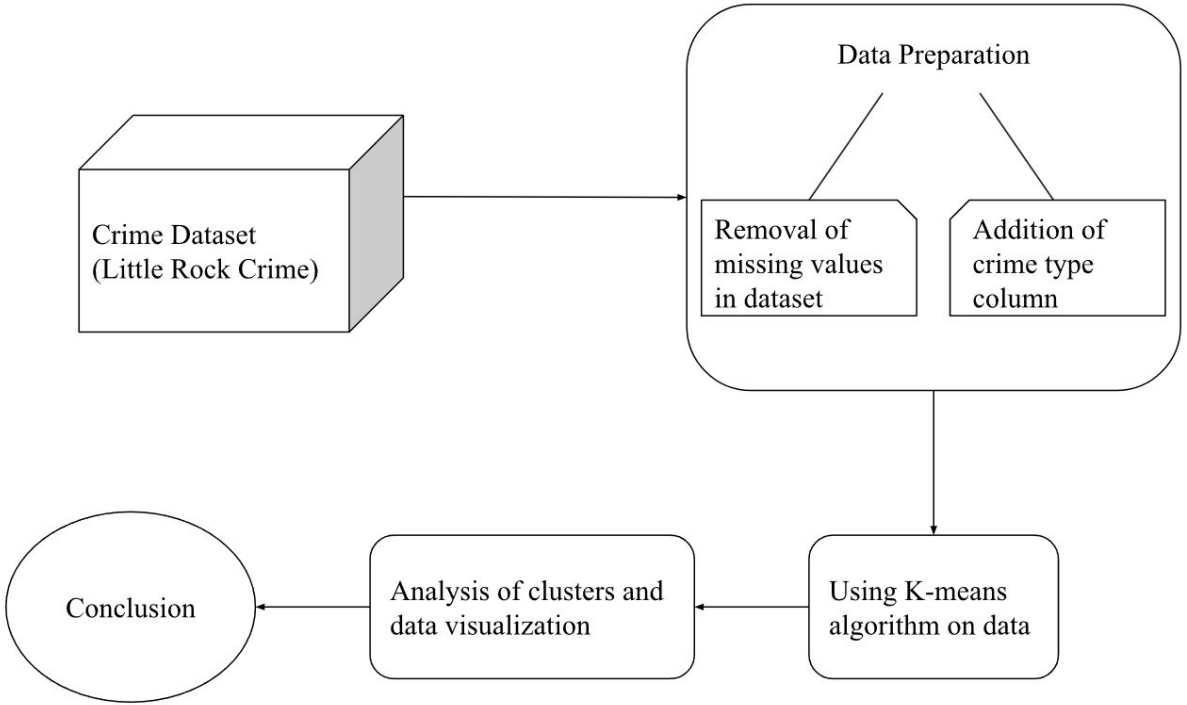
Research Contributions

- A map showcasing which areas in Little Rock have the most occurrences of crime
- Dataset analysis of which type of crimes are most frequent in the city
- Possible prevention of crime and reduced crime rates in the city of Little Rock, AR.

3

Methodology

Experimental Setup


Data Preparation
Crime Dataset
(Little Rock Crime)
Removal of
missing values
in dataset
Addition of
crime type
column
Conclusion
Analysis of clusters and
data visualization
Using K-means
algorithm on data

Experimental Setup

01 Dataset

94584 rows of violent and property crimes

02 Tool Used

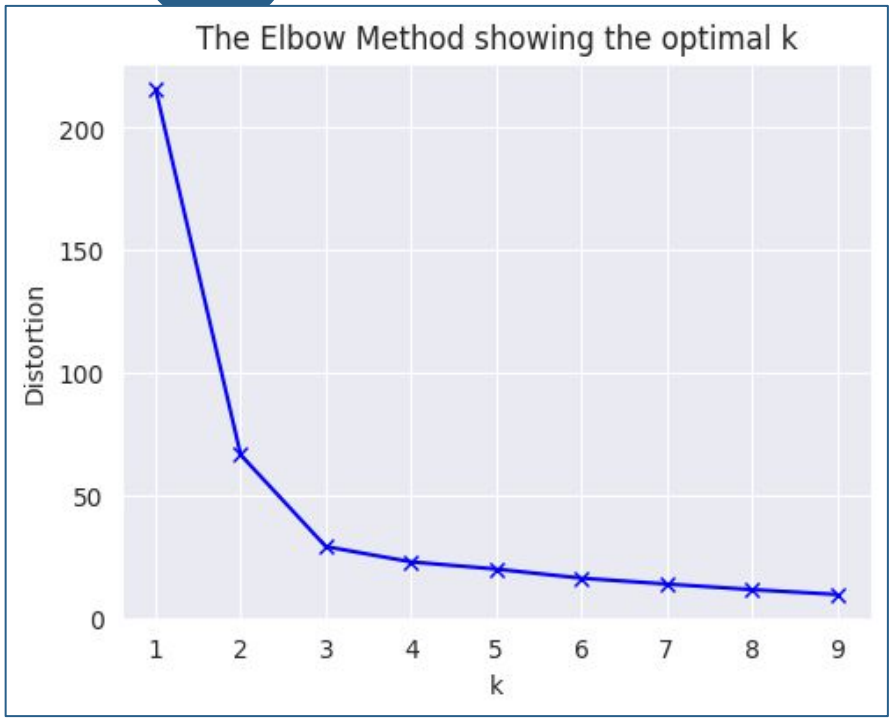
Python language in the Google Colab environment

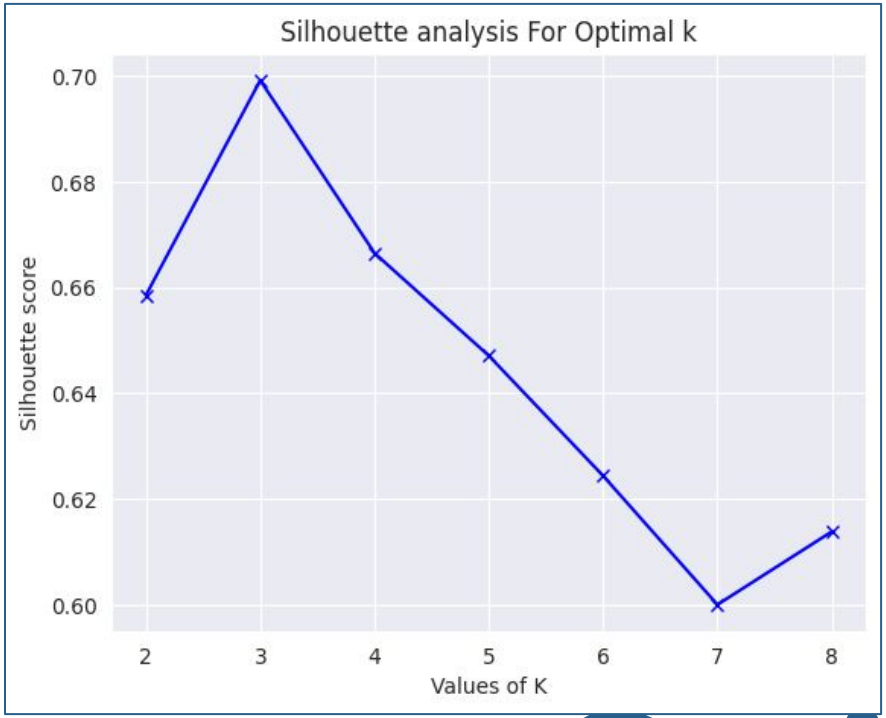
Preprocessing 03

Missing value removal and addition of columns

Validity measure 04

Use of Elbow method and Silhouette analysis to find K

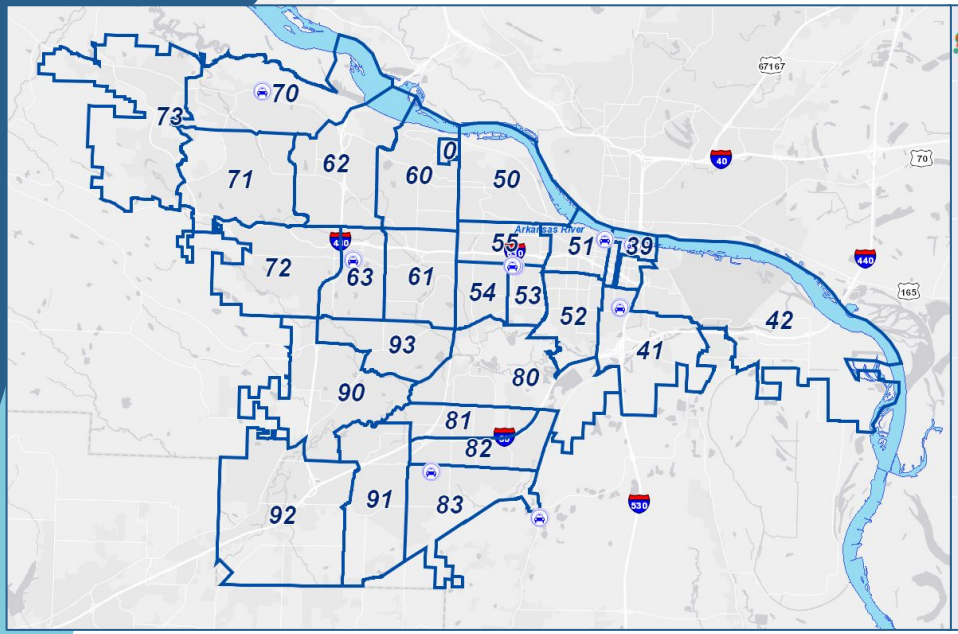

The Elbow Method showing the optimal k
Distortion
0
50
100
150
200
1
2
3
4
5
6
7
8
9
k


Silhouette analysis For Optimal k
Silhouette score
0.60
0.62
0.64
0.66
0.68
0.70
2
3
4
5
6
7
8
Values of K

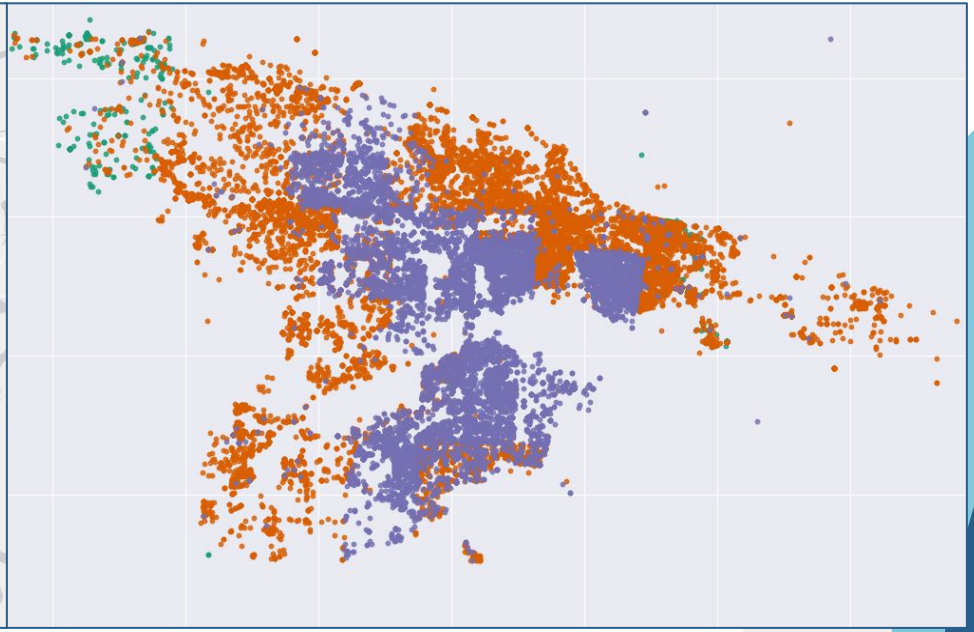
4

Analysis and Results

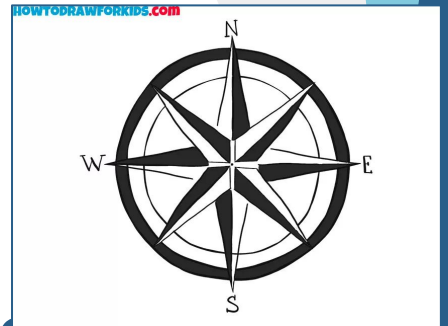
Police Patrol Districts



Clusters



Clusters	Total Entries	Distributed percentages			
		Violent Crime			Non-Violent
		Murder & nonnegligent manslaughter	Theft with weapon	Aggravated Assault	Theft without weapon
Cluster 1	1056	0.1	3	17.5	79.4
Cluster 2	18879	0.3	2.7	19.8	77.1
Cluster 3	19198	0.6	3.7	26.5	69.1



5

Conclusion

Summary & Future Work

Summary

Cluster 3 is the most dangerous (southeast and center) and Cluster 1 is the safest (northwest). An increase in law enforcement in Cluster 3 is likely to prevent crime in the city.

Future Work

Future research should focus on using multiple clustering strategies and on research on different cities/countries.

References

- *U.S. Crime Rate & Statistics 1990-2023*. (n.d.). MacroTrends. https://www.macrotrends.net/countries/USA/united-states/crime-rate-statistics
- Khan, M., Ali, A., & Alharbi, Y. (2022). Predicting and Preventing Crime: A Crime Prediction Model Using San Francisco Crime Data by Classification Techniques. *Complexity*, *2022*, 1–13. https://doi.org/10.1155/2022/4830411
- A Crime Data Analysis of Prediction Based on Classification Approaches. (2022). *Baghdad Science Journal*. https://doi.org/10.21123/bsj.2022.6310
- *Socrata Developer Portal | Socrata*. (n.d.). https://dev.socrata.com/foundry/data.littlerock.gov/bz82-34ep
- Schiller, A. (2022, December 12). *Little Rock, AR Crime Rates*. NeighborhoodScout. https://www.neighborhoodscout.com/ar/little-rock/crime
- Wu, J., Abrar, S. M., Awasthi, N., Frias-Martinez, E., & Frias-Martinez, V. (2022). Enhancing short-term crime prediction with human mobility flows and deep learning architectures. *EPJ Data Science*, *11*(1), 53. https://doi.org/10.1140/epjds/s13688-022-00366-2
- Lepenski, B. (2022, October 13). *Little Rock economist and broker talk about crime's impact on economy*. KATV. https://katv.com/news/local/little-rock-economist-and-broker-talk-about-crimes-impact-on-economy-chief-economist-state-economic-forecaster-university-of-arkansas-at-little-rock-statistics-stacy-hamilton-michael-pakko

References continued

- Krishnendu, S. G., Lakshmi, P. P., & Nitha, L. (2020). Crime Analysis and Prediction using Optimized K-Means Algorithm. *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)*, 915–918. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-000169
- Joshi, A., Sabitha, A. S., & Choudhury, T. (2017). Crime Analysis Using K-Means Clustering. *2017 3rd International Conference on Computational Intelligence and Networks (CINE)*, 33–39. https://doi.org/10.1109/CINE.2017.23
- Mohammedaashiq. (n.d.). *Crime-Rate-Prediction-and-Analysis/analysis-of-crimes.ipynb at main · Mohammedaashiq/Crime-Rate-Prediction-and-Analysis*. GitHub. https://github.com/Mohammedaashiq/Crime-Rate-Prediction-and-Analysis/blob/main/analysis-of-crimes.ipynb
- *Socrata Developer Portal | Socrata*. (n.d.). https://dev.socrata.com/foundry/data.littlerock.gov/bz82-34ep
- https://maps.littlerock.gov/portal/home/webmap/viewer.html?webmap=7cccb13b1b3f4e549a24dce63bee4f2b

Thanks!

Do you have any questions?