Chicago Traffic Safety Analysis: Unraveling Patterns and Predicting Crash Severity

by Korir Brian

Summary

In the bustling city of Chicago, traffic safety is a critical concern impacting the lives of residents every day. The "Chicago Traffic Safety Analysis" project seeks to delve into the patterns and dynamics of traffic crashes within the city, leveraging a comprehensive dataset provided by the City of Chicago. By scrutinizing this dataset, we aim to uncover valuable insights that can contribute to a safer and more informed urban environment.

Outline

1. Business Understanding
2. Modeling
3. Evaluation
4. Conclusions

Business Understanding

Primary Causes: Identify the primary causes of car crashes in Chicago. Vehicle Types: Explore the types of vehicles involved in crashes. Time and Location Patterns: Analyze if there are specific patterns related to time and location. Severity: Understand the severity of crashes and factors influencing severity.

Stakeholder Audience:

The stakeholder audience for the Chicago Car Crashes dataset could include various entities such as:

- City Authorities,Insurance Companies, General Public

Modeling

Predict Severity: Using features from the dataset to predict the severity of a crash (e.g., minor, moderate, severe).

Explore Clusters: Identifying clusters of crashes with similar characteristics.

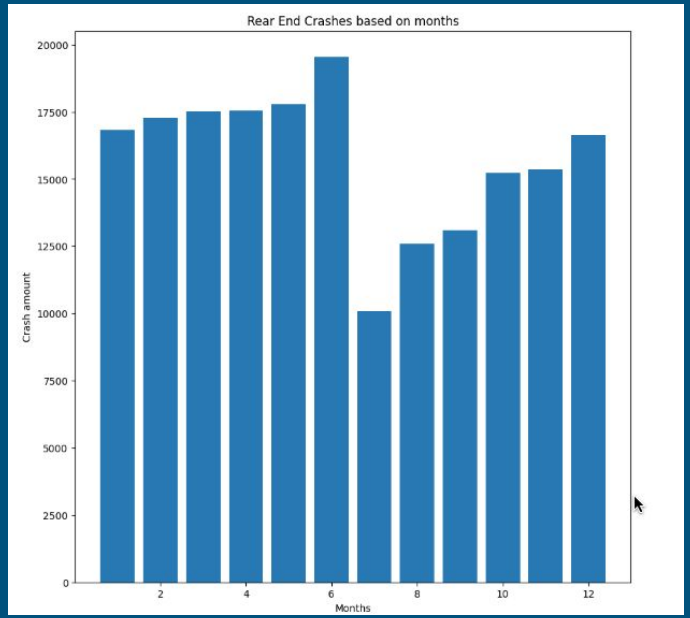
Techniques:

Classification Models: Train models to predict severity (e.g., Decision Trees, Random Forest).

Clustering Algorithms: Applying algorithms like K-Means to identify groups with similar characteristics.

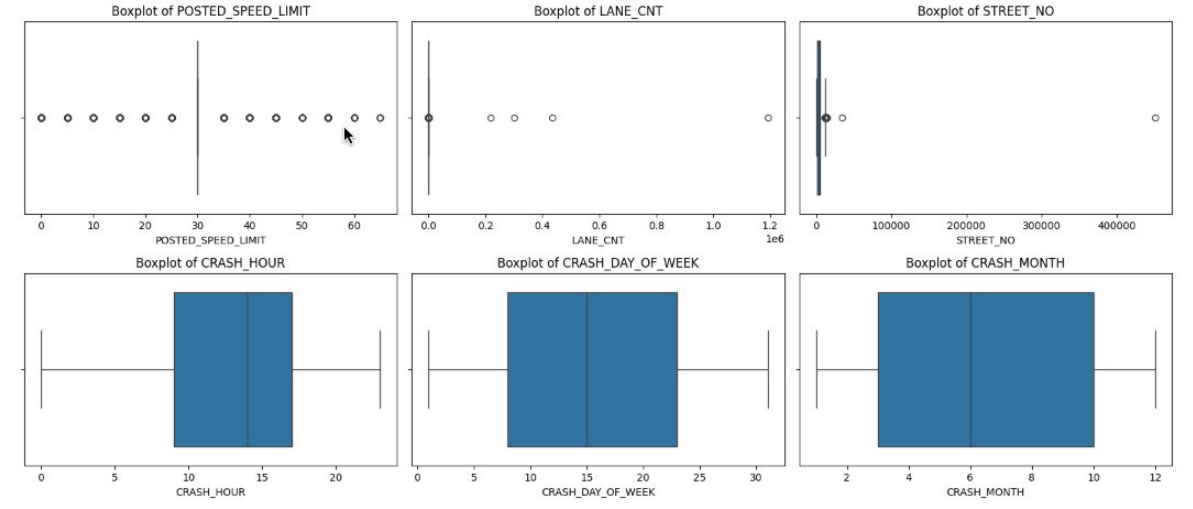
Data Analysis

Comparing rear end crashes occurrence in months.



Outliers For Selected Data

Detecting outliers in a dataset involves analyzing the distribution of values and identifying data points that deviate significantly from the norm. Below demonstrates distribution of Crash Occurrence.



Correlation

correlations close to 0 do not necessarily mean no relationship; it might indicate a nonlinear relationship or the need for more advanced statistical techniques.

Evaluation

1. Accuracy: Evaluating the accuracy of the severity prediction model.
2. Cluster Cohesion: Measuring the cohesion within identified clusters.

Conclusion

In conclusion, the analysis of the Chicago Car Crashes dataset has provided valuable insights into the dynamics of traffic incidents within the city. The investigation covered various aspects, including crash patterns, high-risk areas, temporal influences, contributing factors, vehicle involvement, geospatial distribution, weather impacts, and the development of predictive models.

In summary, this comprehensive analysis equips stakeholders with evidence-based insights and actionable information to address and mitigate traffic incidents in Chicago. By leveraging these findings, the city can make informed decisions, implement targeted interventions, and work towards creating safer road environments for its residents.

Thank You