

ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCE

SCHOOL OF INFORMATION SCIENCE

Data Science Group Project

GROUP MEMBERS ID

1. KASSAHUN ADDIS UGR/3802/13
2. LAKACHEW DILU UGR/8255/13
3. MESAYIT W/TEKILE UGR/1332/13
4. SIRAW TADESSE UGR/4363/13
5. TIGIST LIKU UGR/4805/13
6. YEKOYE ADMASSUE UGR/3202/13

Submitted to: Instructor Michael Melese(PhD).

Submission date: Jun 1 2024 G.C

**Subject: Summary Report on Traffic Collision Analysis in the City of Los Angeles**

**Introduction:**

Traffic collisions are a significant concern for urban planners, policymakers, and law enforcement agencies worldwide. This report presents an analysis of traffic collision data in the City of Los Angeles from 2010 to the present. The objective is to identify trends, hotspots, and patterns in collision occurrences to inform targeted interventions for improving road safety.

**Data Collection and Preprocessing:**

The dataset used for this analysis comprises 607,855 records of traffic collisions, sourced from <https://catalog.data.gov/dataset/traffic-collision-data-from-2010-to-present> . The dataset includes information such as the date, time, location, severity, and contributing factors of each collision. Several preprocessing steps were undertaken before analysis:

**Date Formatting:** The 'Date Occurred' column was converted to datetime format to facilitate temporal analysis.

**Feature Engineering:** Additional features such as year, month, and area were extracted from the date and location information to enable deeper analysis.

**Analysis and Visualization:**

**Temporal Analysis:**

We analyzed the temporal distribution of collisions over time. A stacked area plot visualized the number of collisions by area across different dates, highlighting potential trends and hotspots.

The number of collisions was aggregated by year. A pie chart visualized the proportion of collisions each year.

**Spatial Analysis:**

A heatmap showcased the distribution of collisions across different areas and months, identifying periods with higher collision frequencies for targeted interventions.

Collisions were grouped by year and area to create bar charts illustrating the total number of collisions for each area over different years, enabling comparison over time.

Bar charts for each year depicted the distribution of collisions across different areas, highlighting areas with the highest frequency of collisions.

Pie charts showcased the proportion of collisions by area for each year, offering a concise representation of the distribution within a given year.

**Yearly Trends:**

Stacked bar plots and box plots illustrated the variation in collision counts over the years, offering insights into long-term patterns and potential areas of concern.

**Top Collision Hotspots:**

Trend analysis identified the top areas with the most collisions over time. Line plots visualized the changes in collision rates over the years for the top 5 areas, aiding in prioritizing resources for intervention strategies.

**Interactive Visualization:**

An interactive bar chart using Altair enabled stakeholders to dynamically explore the total number of collisions by area, enhancing user engagement and understanding of the data.

**Summary of Findings:**

**Key Findings:**

**Temporal Trends:** The analysis revealed fluctuating trends in the number of collisions over the years, with significant increases from 2010 to 2016 and then a decrease from 2016 to the present.

**Spatial Patterns:** Certain areas, such as 77th Street, consistently exhibited the highest collision rates. The top 5 areas with higher collision rates are 77th Street, North Hollywood, Olympic, Southwest, and Wilshire, compared to others like Foothill, which show lower collision rates. Further investigation is warranted to understand the underlying causes and implement targeted interventions.

The line plot revealed fluctuating trends in collision occurrences over the years, with notable peaks and troughs.

Bar charts provided a comprehensive overview of collision distribution across areas in the City of Los Angeles for individual years, highlighting areas of concern.

Pie charts offered insights into the proportional distribution of collisions by area within each year, aiding in identifying hotspots and trends.

The analysis reveals temporal variations in collision rates, suggesting the need for targeted interventions during specific periods.

Spatial analysis identifies areas with high collision frequencies, guiding resource allocation for road safety measures.

Yearly trends highlight areas experiencing consistent increases or decreases in collision rates, informing long-term strategies.

Identifying top collision hotspots facilitates focused interventions to mitigate collision risks effectively.

**Recommendations:**

Based on the findings, several recommendations can be proposed:

**Enhanced Enforcement:** Deploy increased law enforcement patrols and speed cameras in high-risk areas to deter reckless driving behavior.

**Infrastructure Improvements:** Invest in infrastructure upgrades such as improved signage, road markings, and traffic calming measures to mitigate collision risks.

**Public Awareness Campaigns:** Launch public awareness campaigns to educate drivers on safe driving practices and the importance of adhering to traffic regulations.

**Targeted Safety Measures:** Implement targeted safety measures in areas identified as high-risk based on the analysis.

**Community Engagement:** Engage with local communities to raise awareness about road safety and encourage collaborative efforts in reducing collision rates.

**Periodic Assessments:** Conduct periodic assessments of collision data to monitor trends and assess the effectiveness of implemented interventions.

In conclusion, this report provides actionable insights into the traffic collision landscape in the City of Los Angeles, empowering stakeholders to implement evidence-based strategies for enhancing road safety and reducing collision risks.