**PROJECT REPORT**

**FACTORS INFLUENCING TRAFFIC CRASHES**

Towards Safer Roads

Team-11

Pavani Chella - G41309434

Yasaswitha Gaddam - G31728688

Introduction

Traffic crashes pose a substantial threat to public safety, public health, and economic stability. This project delves into understanding the multifaceted factors that contribute to traffic crashes, aiming to uncover insights that can drive the development of effective strategies for crash prevention and mitigation. By examining a diverse range of variables such as road infrastructure, weather conditions, driver behavior, vehicle characteristics, and demographic factors, we seek to unravel the complex interplay between these elements and their impact on crash occurrences. Through rigorous data analysis, statistical modeling, and visualization techniques, we aim to identify patterns, trends, and risk factors associated with traffic crashes. The ultimate goal is to provide actionable insights to policymakers, urban planners, law enforcement agencies, and stakeholders, facilitating evidence-based decision-making and interventions aimed at creating safer roads and communities for all.

Problem Statement

The problem at hand is to analyze traffic crash data and to understand the underlying factors contributing to traffic accidents. By examining variables such as road conditions, weather, and contributing causes, the goal is to identify patterns and risk factors associated with accidents. This analysis will provide valuable insights to improve road safety and reduce the frequency and severity of accidents.

### Description of the Data

The crash data provided includes information on traffic crashes of Chicago from 2013 to 2024 occurring within the City of Chicago limits and under the jurisdiction of the Chicago Police Department (CPD)., providing insights into various aspects such as crash dates and times, locations, contributing factors, and injury severity. Each record represents a single traffic crash incident, allowing for comprehensive analysis of road safety trends and contributing factors. Dataset has 8,21,122 rows and 48 Columns.

Dataset Source Link: <https://catalog.data.gov/dataset/traffic-crashes-crashes>

CRASH\_RECORD\_ID: Unique identifier for each crash record.

CRASH\_DATE: Date of the crash.

TRAFFIC\_CONTROL\_DEVICE: Type of traffic control device at the crash location (stop sign, traffic signal).

WEATHER\_CONDITION: Weather conditions at the time of the crash.

LIGHTING\_CONDITION: Lighting conditions at the time of the crash.

ROADWAY\_SURFACE\_COND: Condition of the roadway surface at the time of the crash

CRASH\_TYPE: Type of crash (no injury, injury, fatal).

HIT\_AND\_RUN\_I: Indicator for whether the crash was a hit-and-run (Y/N). PRIM\_CONTRIBUTORY\_CAUSE: Primary contributing cause of the crash. SEC\_CONTRIBUTORY\_CAUSE: Secondary contributing cause of the crash.

INJURIES\_FATAL: Number of fatal injuries reported in the crash.

CRASH\_HOUR: Hour of the day when the crash occurred.

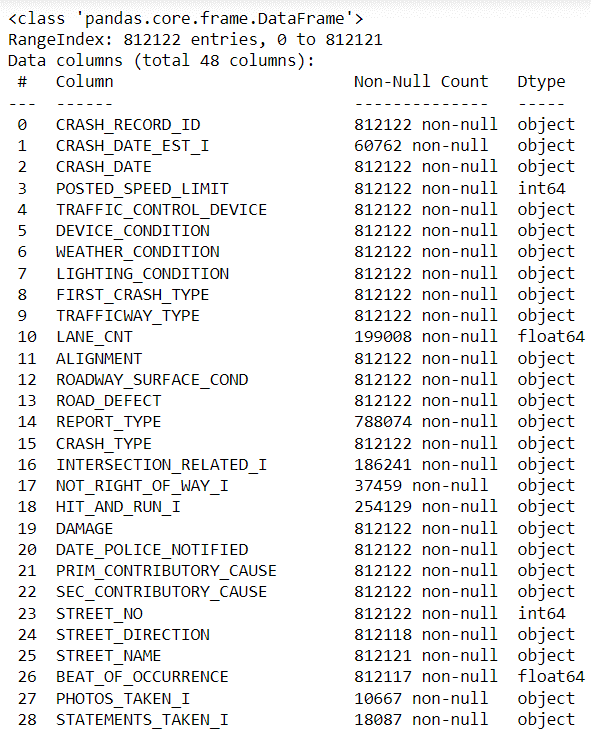
CRASH\_DAY\_OF\_WEEK: Day of the week when the crash occurred.

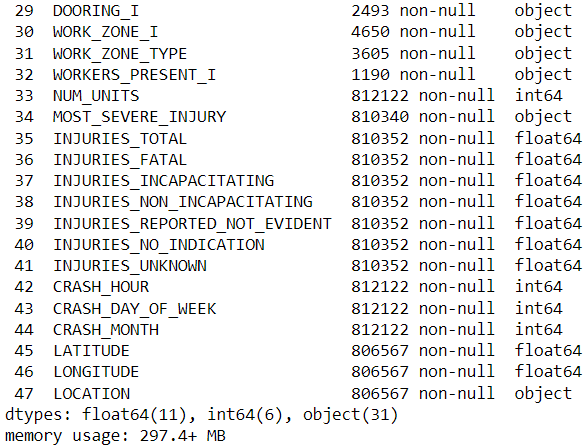
CRASH\_MONTH: Month when the crash occurred.

LATITUDE: Latitude coordinate of the crash location.

LONGITUDE: Longitude coordinate of the crash location.

LOCATION: Geographic location of the crash (latitude and longitude).





Business Questions

1. Is there any day of the week or time of day when crashes are more frequent?

2. Do certain types of traffic control devices (e.g., stop signs, traffic lights) correlate with fewer accidents?

3. What are the primary contributory causes of crashes?

4. What are the most common causes of traffic crashes based on the dataset?

5. How do weather conditions, time of day, lightning condition influence crash frequency and severity?

Data Quality Analysis

Completeness and Integrity

Our dataset underwent a comprehensive preprocessing stage where missing values were addressed, duplicates were removed, and data types were standardized. This ensured that the dataset was complete and robust, suitable for conducting a reliable analysis. The treatment of outliers further refined the dataset, making it representative of typical traffic conditions without extreme values skewing the results.

Consistency

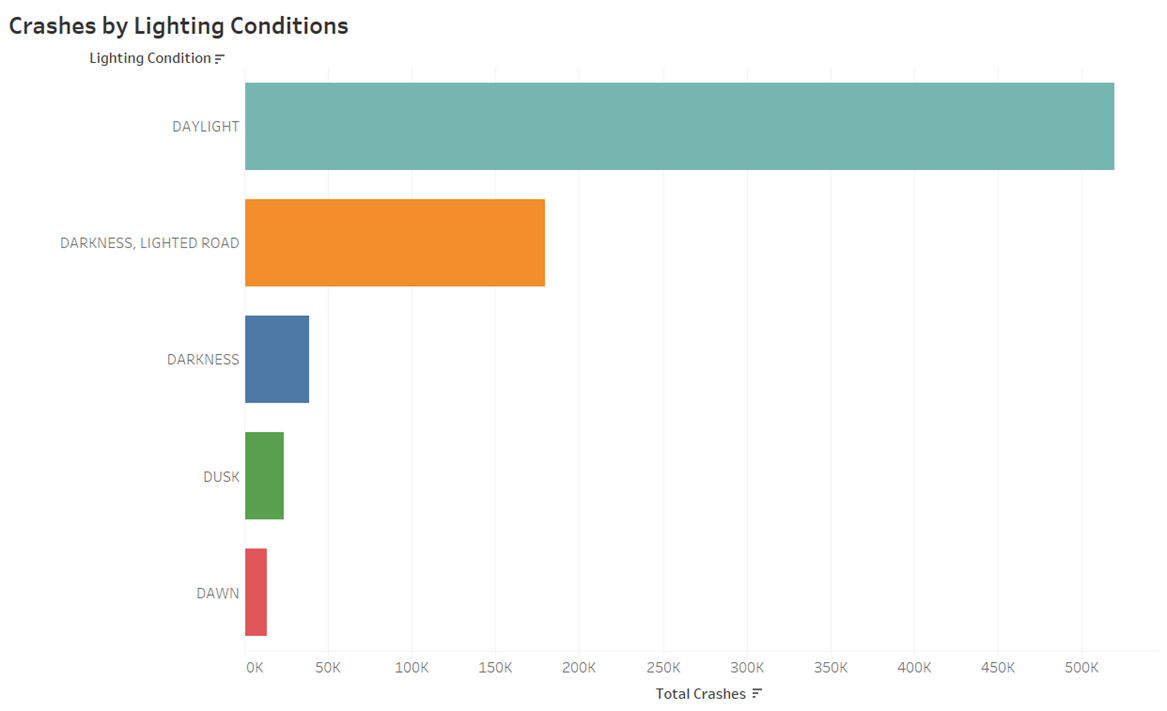
We maintained consistency across data entries by standardizing date and numeric formats, which was crucial for temporal and quantitative analyses. The transformations applied, such as scaling and categorization, helped in normalizing the data distribution, making it ready for comparison and aggregation.

Main Analysis

Descriptive Statistics and Trends

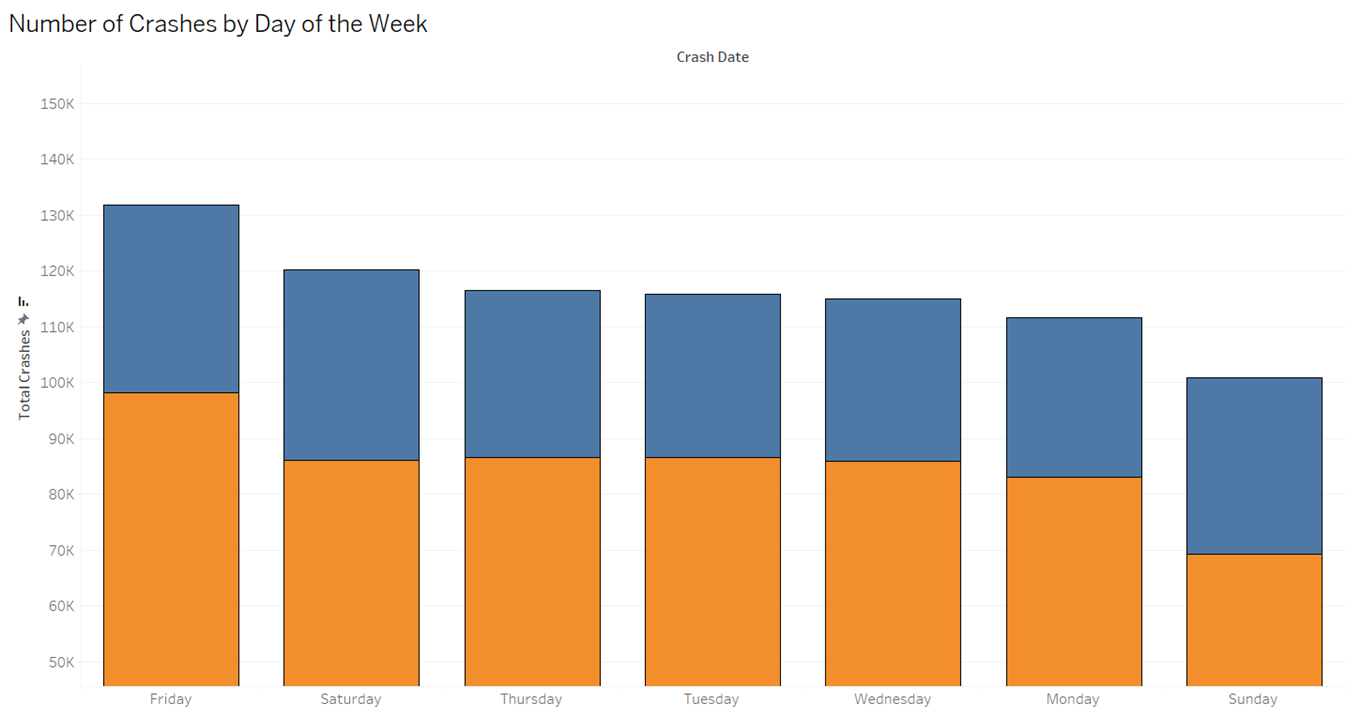
1.Crashes by Lighting Conditions:

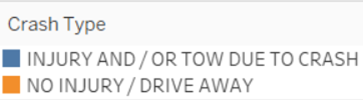
This bar graph illustrates the total number of vehicular crashes classified by different lighting conditions. The categories include Daylight, Darkness on Lighted Roads, Darkness, Dusk, and Dawn. Daylight conditions account for the highest number of crashes, followed by crashes on lighted roads during darkness.



2. Number of Crashes by Day of the Week:

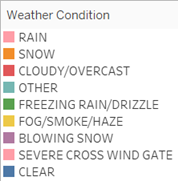
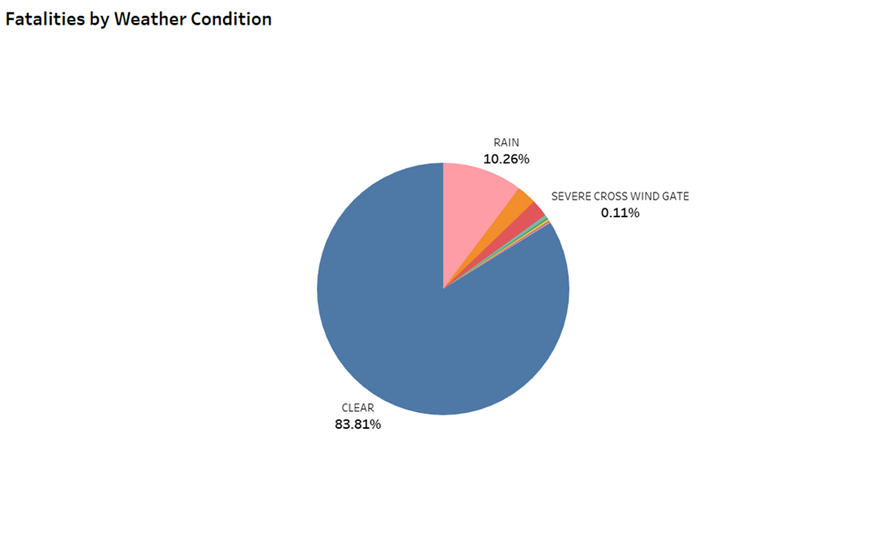
This bar graph displays the total number of crashes divided by the day of the week, with a distinction between crashes involving injuries or towing and those with no injuries where drivers could drive away. Fridays and Saturdays show the highest total number of crashes.





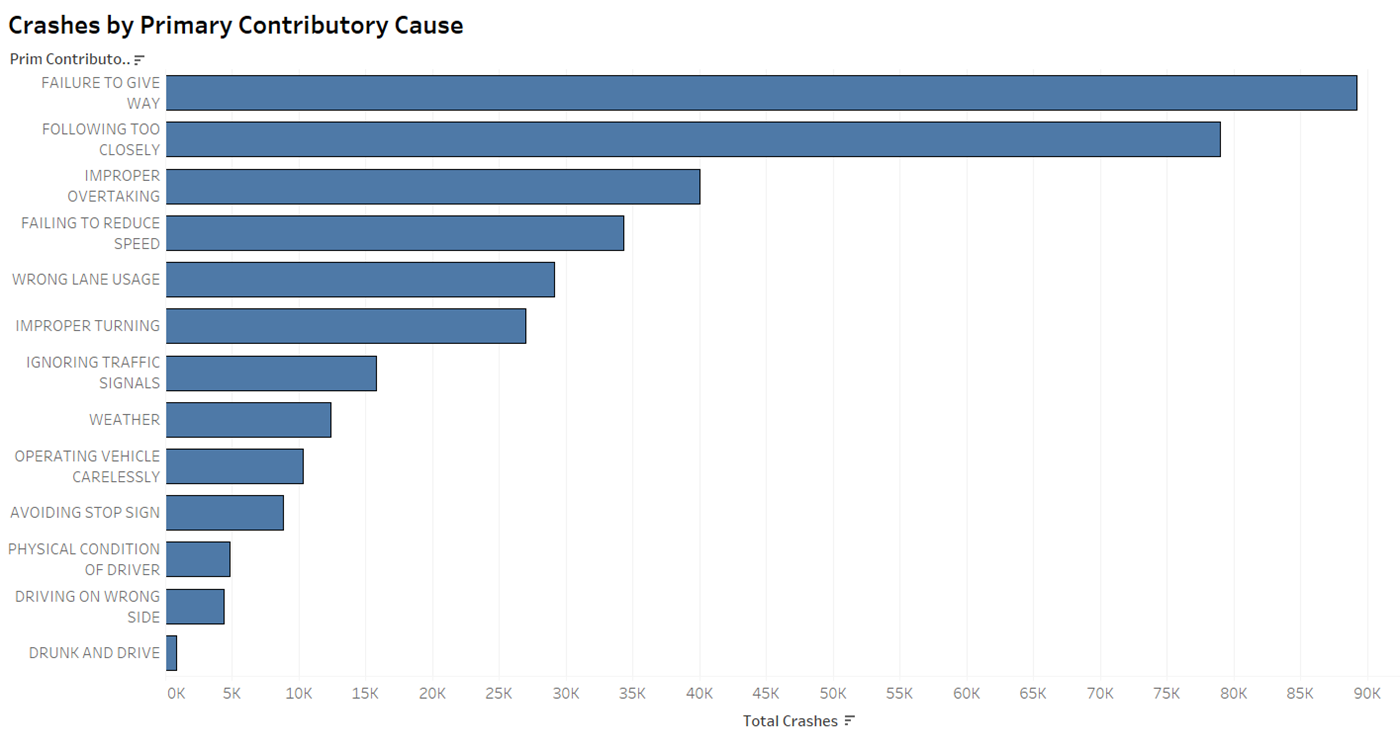
3. Fatalities by Weather Condition:

This pie chart represents the percentage of fatalities occurring under various weather conditions. The vast majority of fatalities occur under clear conditions, followed by a smaller percentage during rain, with a negligible percentage attributed to severe crosswind gate conditions.



4. Crashes by Primary Contributory Cause:

This bar graph ranks the primary causes contributing to crashes, with 'Failure to Give Way' being the most common cause. Other significant causes include 'Following Too Closely' and 'Improper Overtaking.' Weather-related crashes also appear as a contributory factor but less significant compared to others.



5. Crashes by Traffic Control Device:

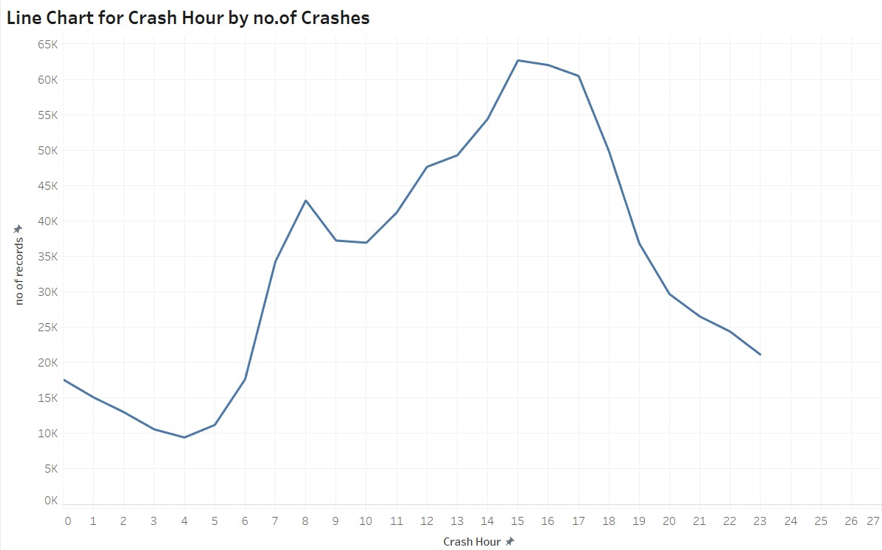
This bubble chart visualizes the total number of crashes in relation to the type of traffic control devices present, if any. The largest proportion of crashes occurs at locations with no traffic controls, followed by areas with traffic signals and stop signs.

6. Weather Conditions by Number of Crashes:

This bubble chart presents the total number of crashes in different weather conditions. The largest number of crashes occur in clear weather, followed by smaller but significant numbers in rain and snow. Other conditions like fog and sleet appear to have a lower incidence of crashes.

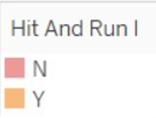
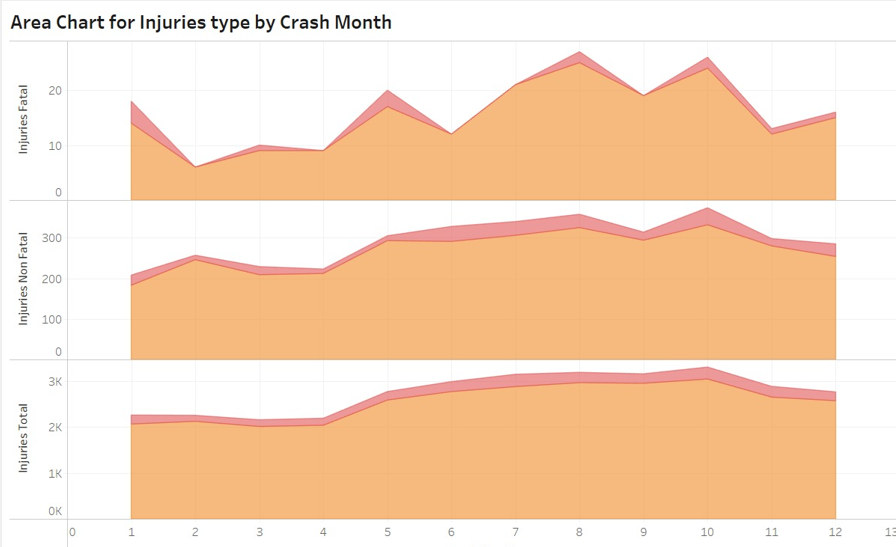
7. Crash Hour by Number of Crashes:

This chart shows the number of crashes occurring at different hours of the day. There is a notable peak during the afternoon hours, particularly around 3 PM to 5 PM, indicating these are the most common hours for crashes.



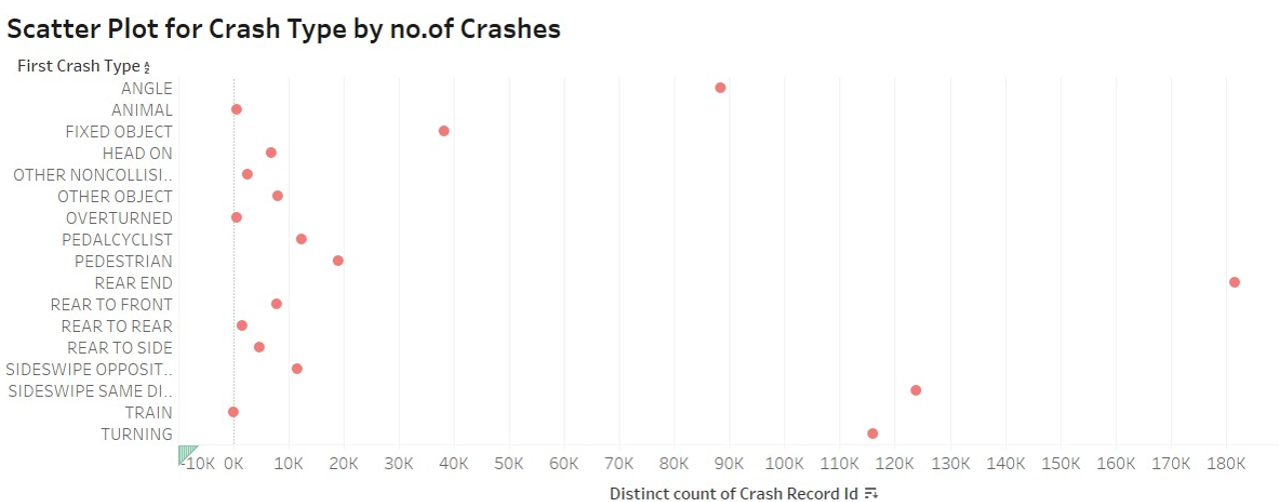
8. Area Chart for Injuries Type by Crash Month:

This chart displays the number of injuries, both fatal and non-fatal, across different months. Both types of injuries tend to rise during the middle of the year, with noticeable peaks around the summer months.



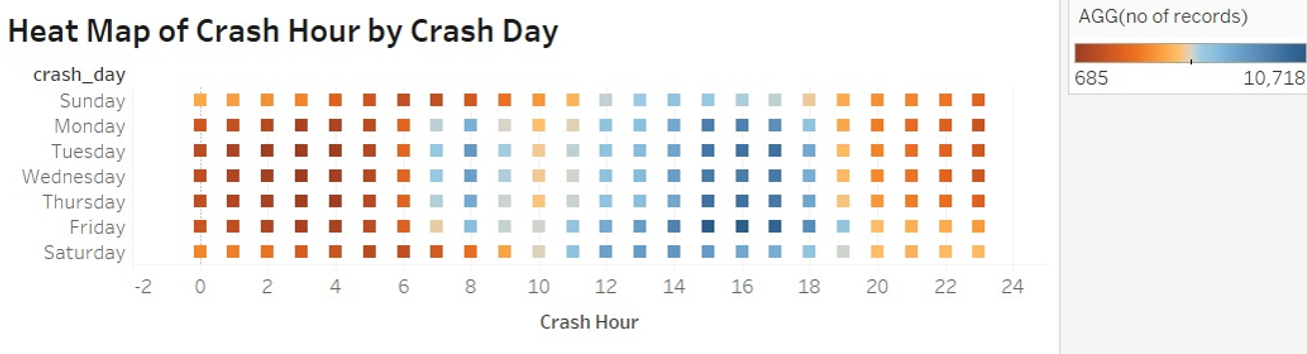
9. Crash Type by Number of Crashes:

This scatter plot categorizes crashes by type, such as rear-end, sideswipe, or involving pedestrians, and shows the frequency of each type. Rear-end and sideswipe crashes appear to be the most common.



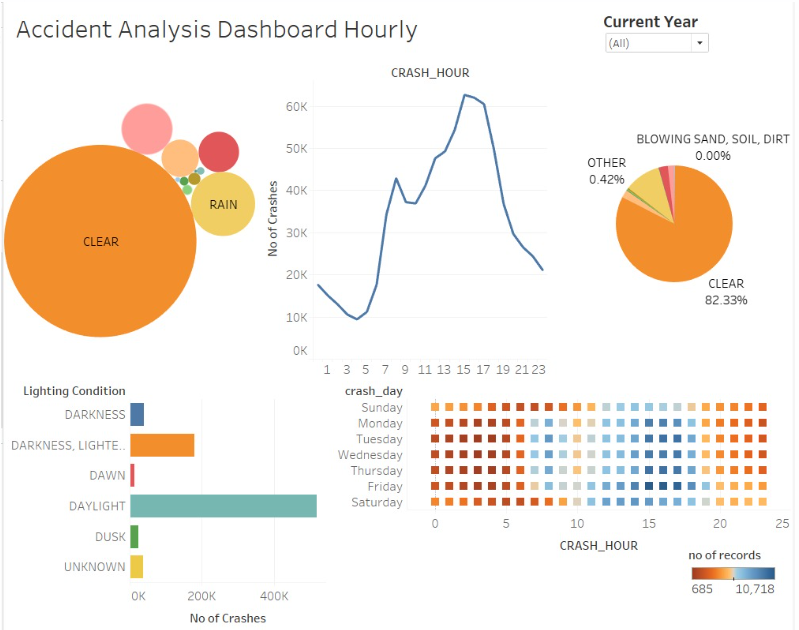
10. Heat Map of Crash Hour by Crash Day:

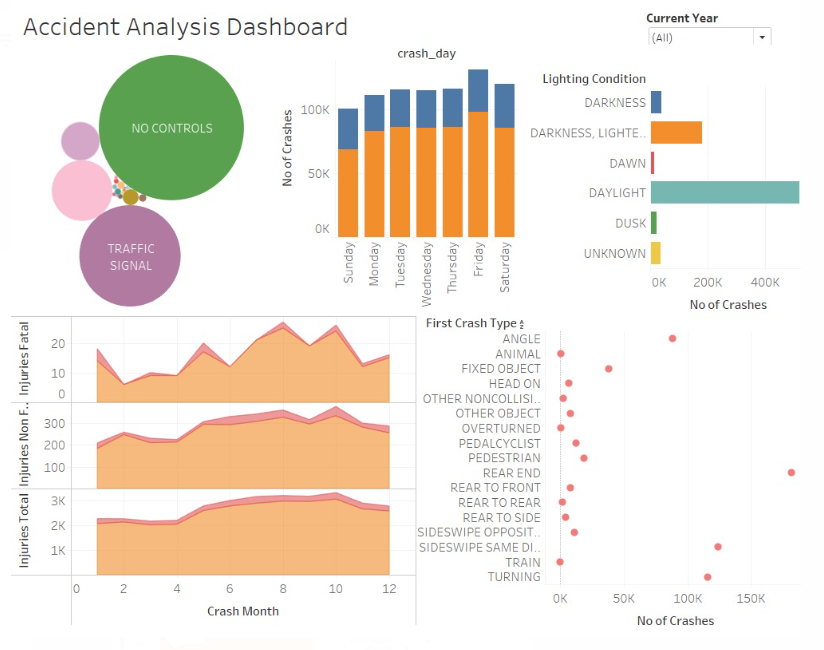
The heat map displays the frequency of crashes across different days of the week and hours of the day, with color intensity representing the number of crashes. Weekday afternoons, especially from Wednesday to Friday, appear to have higher crash rates.



Comparative Analysis

We compared crash rates across different times of the day and days of the week, which pointed to higher incidents during evening rush hours and weekends. This temporal distribution suggests targeted times for enforcing traffic regulations and enhancing road safety measures.





Key Findings

Majority of crashes occur during daylight hours, especially on weekends (Friday and Saturday).

"Failure to give way" is a significant contributing factor to crashes, highlighting the importance of driver awareness and adherence to traffic regulations.

Crashes are more prevalent in areas lacking traffic control devices, emphasizing the role of infrastructure in accident prevention.

Clear weather conditions are associated with higher crash rates, suggesting potential factors such as driver complacency or increased traffic volume.

Fatality rates are higher in clear weather conditions, followed by rainy weather, indicating the need for enhanced safety measures and awareness campaigns to address weather-related risks.

Behavioural Insights: The significant occurrences of crashes due to failure to yield and speeding underscore the importance of addressing driver behaviour through education and stricter law enforcement.

Conclusion

The visualization-driven exploration of traffic crash data provided us with actionable insights, highlighting the crucial interplay of various factors like environmental conditions, infrastructure, and human behaviour in traffic safety. By leveraging these insights, stakeholders can develop focused interventions to enhance road safety, ultimately contributing to the reduction of traffic-related incidents and injuries.

Overall, these findings underscore the multifaceted nature of road safety challenges, highlighting the need for comprehensive strategies that include enforcement of traffic laws, improvements in road and traffic control infrastructure, public awareness campaigns on safe driving practices, and perhaps adjustments in road use regulations during peak times and in specific weather conditions. These strategies could significantly reduce the number of vehicular crashes and enhance overall road safety.