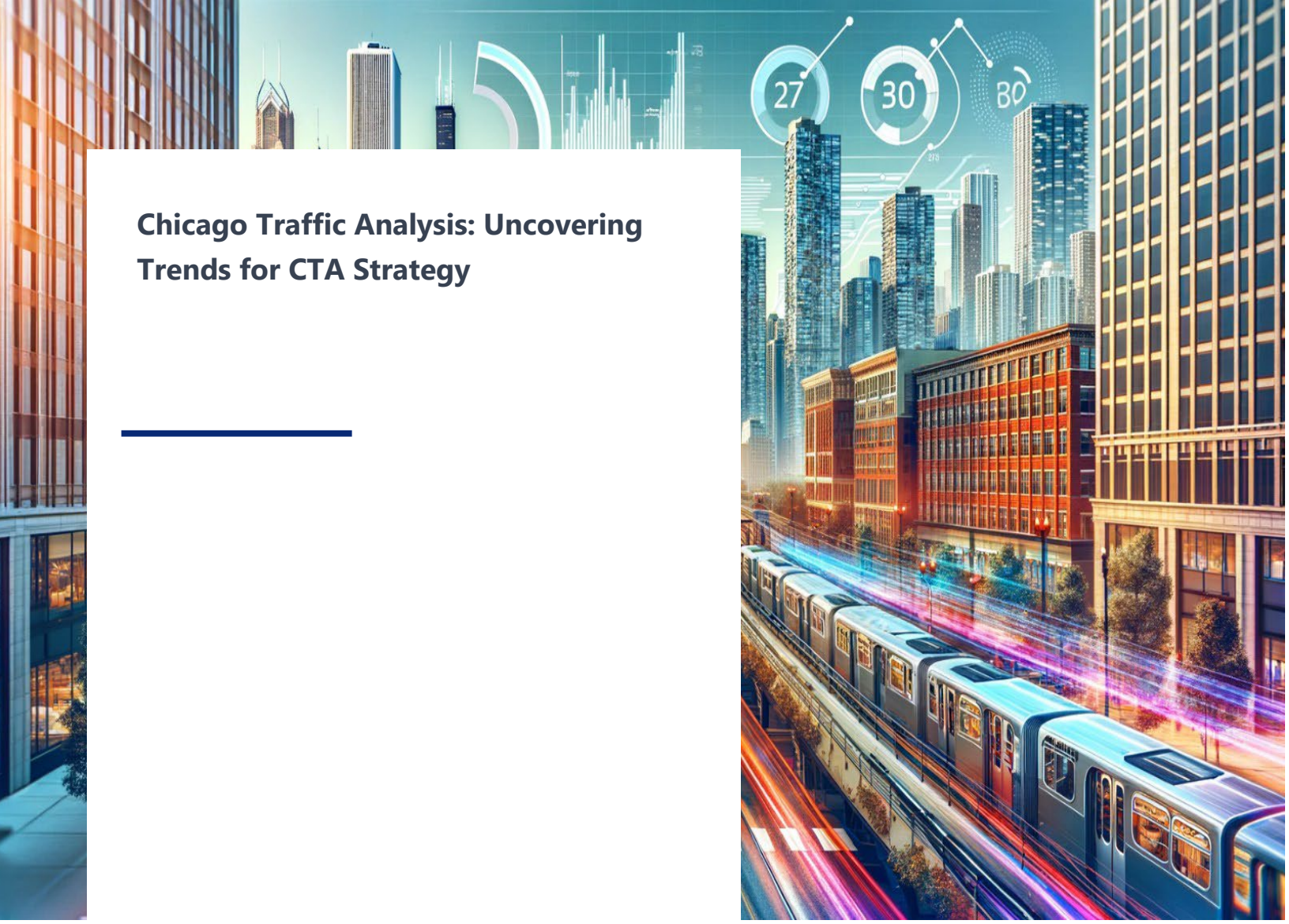


Chicago Traffic Analysis: Uncovering Trends for CTA Strategy

DECEMBER 7

Chicago Transit Authority

Authored by: Himamshu Chandrashekara

Table of Contents

Introduction:

In an effort to enhance the safety and efficiency of Chicago's transportation network, this report presents a comprehensive analysis of traffic crash data within the city, spanning a decade from 2013 to 2023. The purpose of this analysis is to discern patterns in the occurrence of crashes, investigate the prevalent types of traffic accidents, and glean actionable insights for the CTA management to implement targeted improvements.

The methodology employed involved an in-depth examination of key data points including the timing of crashes (by hour, day of the week, and month), geographical location, weather and lighting conditions, and the nature of the crashes. The analysis was grounded in a robust data set that captured various aspects of traffic incidents, providing a multifaceted view of the conditions under which crashes are most likely to occur.

The scope of the report extends to answering three critical questions for the CTA management:

1. **When and Where Are Crashes Most Likely:** Through spatial and temporal analysis, we identify patterns that pinpoint high-risk periods and locations within the city.
2. **What Types of Crashes Are Most Likely:** We categorize and quantify the types of crashes, revealing the most frequent scenarios that lead to traffic incidents.
3. **Actionable Insights for the CTA:** The analysis culminates in strategic recommendations, offering the CTA management data-driven directives to mitigate crash occurrences and enhance commuter safety.

The ensuing sections of this report detail the findings of our analysis, concluding with a set of tailored recommendations for policy and operational enhancements to aid the CTA in its mission to provide safe, efficient, and reliable transit services.

Overview

The dataset provides a comprehensive view of various aspects of traffic crashes in Chicago. The dataset contains 37 columns and 786386 rows after data cleaning with a high percentage of missing values from 2013 as given from the source.

The dataset utilized for this analysis was sourced from the City of Chicago's official data portal and is publicly available at the following link: https://data.cityofchicago.org/Transportation/Traffic-Crashes-Crashes/85ca-t3if/data .

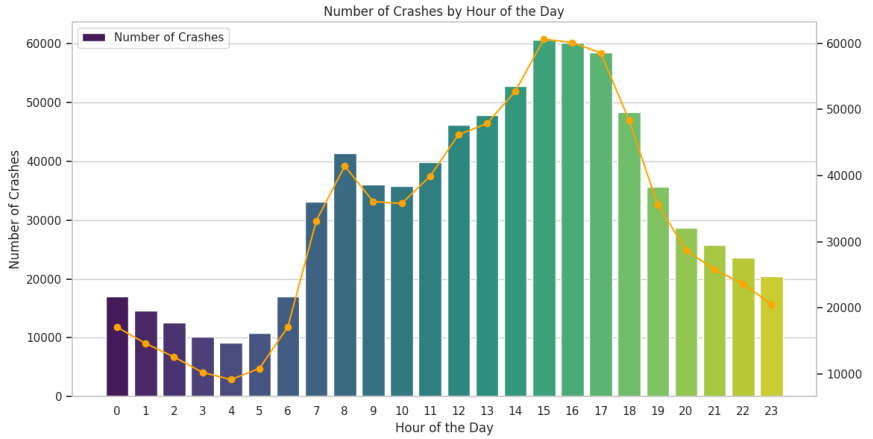
This dataset comprises comprehensive information on traffic crashes in Chicago and serves as the foundation for our examination of causes and trends within the city's transportation landscape. Here are some key columns that will be useful for our analysis:

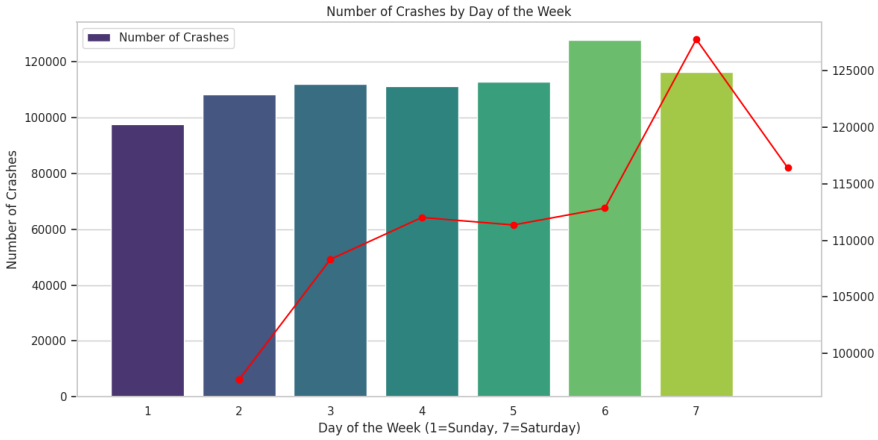
- **CRASH_DATE**: The date and time when the crash occurred.
- **POSTED_SPEED_LIMIT**: Speed limit in the area where the crash occurred.
- **TRAFFIC_CONTROL_DEVICE**, **DEVICE_CONDITION**: Information about traffic control devices and their condition.
- **WEATHER_CONDITION**, **LIGHTING_CONDITION**: Weather and lighting conditions at the time of the crash.
- **FIRST_CRASH_TYPE**: The type of the first collision in the crash.
- **TRAFFICWAY_TYPE**: The type of trafficway where the crash occurred.
- **CRASH_HOUR**, **CRASH_DAY_OF_WEEK**, **CRASH_MONTH**: The hour, day of the week, and month when the crash occurred.
- **LATITUDE**, **LONGITUDE**, **LOCATION**: Geographical coordinates and location of the crash.

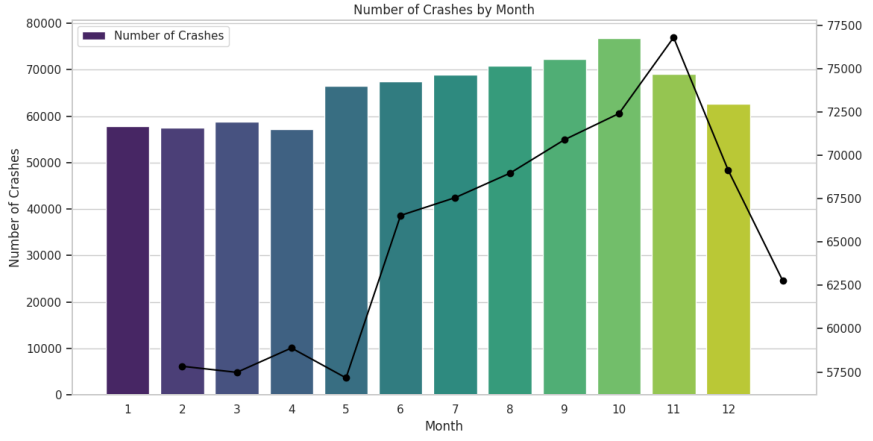
When and where are crashes most likely?

To answer the question "When and where are crashes most likely?", we should focus on the **CRASH_DATE**, **CRASH_HOUR**, **CRASH_DAY_OF_WEEK**, **CRASH_MONTH**, **LATITUDE**, and **LONGITUDE** fields. This will allow us to analyze the temporal and spatial patterns of crashes.

We have analyzed the temporal trends (time of day, day of week, and month) of traffic crashes in Chicago from 2013 to 2023.







Interpretation
Key findings

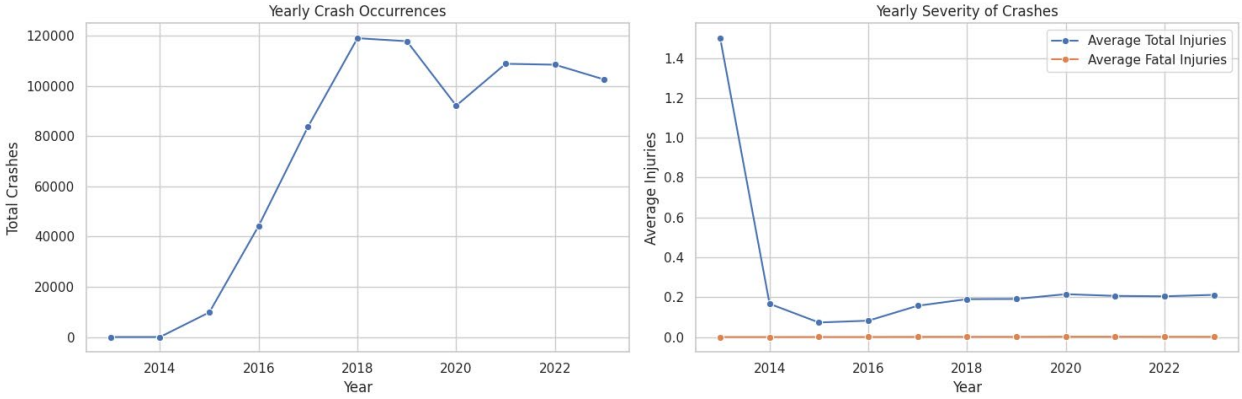
Number of Crashes by Hour of the Day: The frequency of crashes throughout the 24 hours of a day. There is a clear trend where the number of crashes is lowest in the early morning hours (around 2-5 AM) and then increases through the morning. The peak occurs in the late afternoon and early evening hours (around 4-6 PM), which likely corresponds with rush hour when traffic volume is highest. After the peak, the number of crashes decreases again in the late evening and night.

Number of Crashes by Day of the Week: The distribution of crashes across the days of the week, with 1 representing Sunday and 7 representing Saturday. The trend shows a relatively stable number of crashes from Sunday to Wednesday, with a slight increase on Thursday. There is a noticeable jump in crashes on Friday (day 5), which is typically the start of the weekend, possibly due to increased travel and potentially riskier behaviors like driving under the influence. The number then decreases on Saturday.

Number of Crashes by Month: The number of crashes by month, with a noticeable seasonal trend. The number of crashes is lowest at the beginning of the year, increases steadily, and peaks around the middle of the year in the warmer months. There's a sharp increase leading to the peak (highlighted in green), which might correspond to summer, a time when more people are traveling and there are more vehicles on the road. After the peak, the number decreases again towards the end of the year, possibly due to less travel or more cautious driving in colder, potentially inclement weather.

Each of these trends can be influenced by a variety of factors, including traffic volume, weather conditions, daylight hours, and driver behaviors. The peaks in crashes during evening rush hours, Fridays, and the summer months suggest that times of higher traffic volumes and potentially more recreational travel could contribute to higher crash rates. Conversely, the lower number of crashes during the early morning hours, midweek days, and colder months could reflect fewer vehicles on the road and possibly more cautious driving behavior.

Exploratory Data Analysis



1. **Yearly Crash Occurrences**:

   Inferences

   - The chart displays a significant increase in the total number of crashes from 2014 to 2018, with the numbers rising steeply from around 20,000 to 120,000. After peaking in 2018, there appears to be a slight decline or stabilization in crash numbers.
   - This trend suggests that there were factors between 2014 and 2018 that contributed to an increase in crash occurrences, which may include increased traffic volume, changes in traffic patterns, or possibly changes in reporting mechanisms.

2. **Yearly Severity of Crashes**:

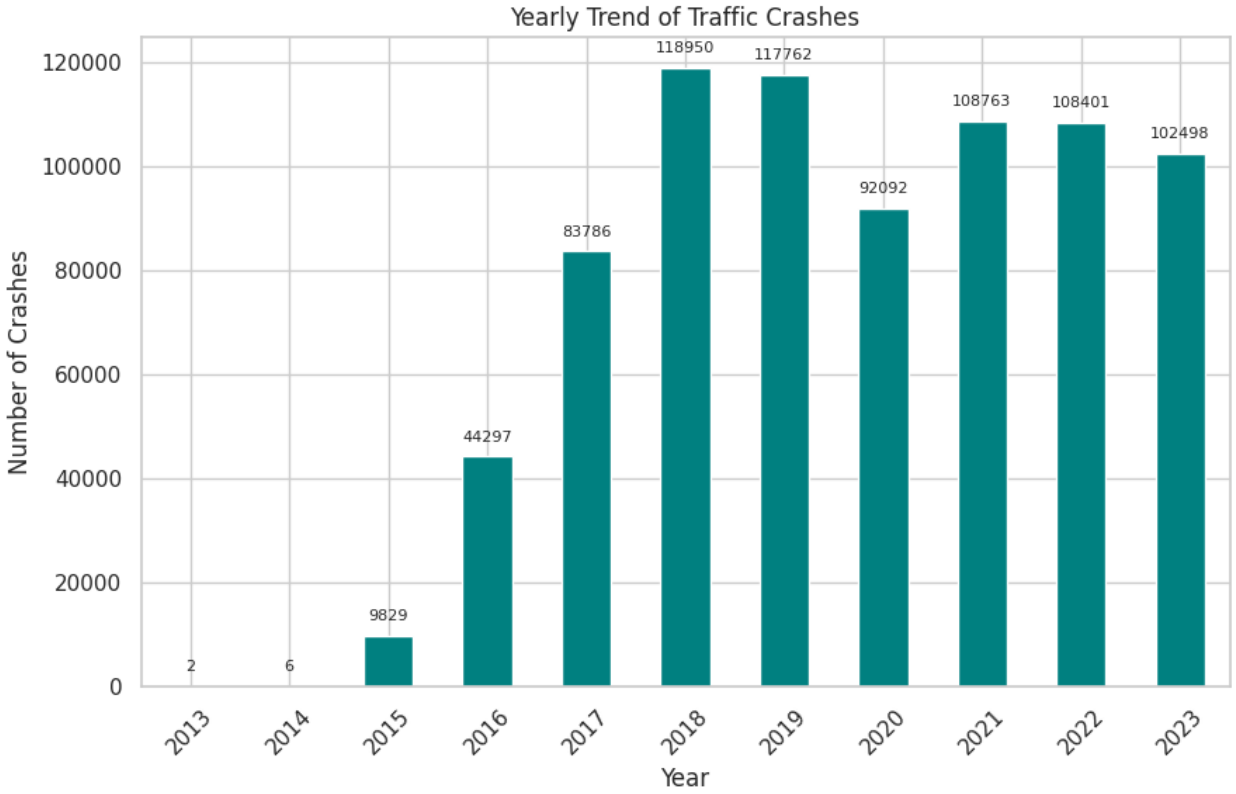
   Inferences

   - The severity of crashes, as depicted by average total injuries, shows a sharp decrease in 2014, dropping from around 1.4 injuries per crash to well under 0.5. This number then stabilizes through to 2022.
   - Fatal injuries per crash also show a decrease, though less pronounced, beginning from just over 0.1 in 2014 to approximately 0.04 in subsequent years.
   - This reduction in average injuries per crash could indicate improvements in road safety measures, vehicle safety technologies, or emergency medical response.

Overall there is a notable spike in crashes around 2016-2017, with crash occurrences peaking at around 120,000 before a slight reduction in subsequent years.

The average total injuries per crash have significantly decreased from 2014 to a lower plateau, suggesting improvements in vehicle safety or traffic management.

Yearly Trend of Traffic Crashes:



Interference:

- The number of traffic crashes presents an upward trend from 2015, starting at a low of approximately 10,000 crashes, increasing year by year to a peak in 2018 at nearly 120,000 crashes.
- Post-2018, the trend shows a very slight decrease but generally maintains high levels of annual crashes, hovering around 100,000 to 110,000 from 2018 to 2023.
- These consistent high numbers post-peak suggest that while the factors leading to the initial increase may have been addressed to stop the rise, persistent issues in traffic management or urban planning may be maintaining the high level of crash occurrences.

What types of crashes are most likely with Recommendations ?

To determine the types of crashes that are most likely, we can analyze the **FIRST_CRASH_TYPE** column in the dataset. This will provide insights into the most common crash scenarios. We'll start by examining the frequency of each crash type and then visualize these frequencies for a clearer understanding.

The analysis of the characteristics for each of the most common crash types reveals some interesting patterns:

index	FIRST_CRASH_TYPE	WEATHER_CONDITION	LIGHTING_CONDITION	TRAFFIC_CONTROL_DEVICE	POSTED_SPEED_LIMIT
0	ANGLE	CLEAR	DAYLIGHT	NO CONTROLS	30
1	PARKED MOTOR VEHICLE	CLEAR	DAYLIGHT	NO CONTROLS	30
2	REAR END	CLEAR	DAYLIGHT	TRAFFIC SIGNAL	30
3	SIDESWIPE SAME DIRECTION	CLEAR	DAYLIGHT	NO CONTROLS	30
4	TURNING	CLEAR	DAYLIGHT	TRAFFIC SIGNAL	30

1. **Angle Crashes:**
    - **Weather Condition:** Most frequently occur in clear weather.
    - **Lighting Condition:** Predominantly happen in daylight.
    - **Traffic Control Device:** Common at locations with a stop sign or flasher.
    - **Posted Speed Limit:** Typically occur in areas with a 30 mph speed limit.
2. **Parked Motor Vehicle Crashes:**
    - **Weather Condition:** Mostly in clear weather.
    - **Lighting Condition:** Usually in daylight.
    - **Traffic Control Device:** Often where there are no controls.
    - **Posted Speed Limit:** Common in areas with a 30 mph speed limit.
3. **Rear End Crashes:**
    - **Weather Condition:** Predominantly in clear weather.
    - **Lighting Condition:** Primarily in daylight.
    - **Traffic Control Device:** Often occur at traffic signals.
    - **Posted Speed Limit:** Frequently in areas with a 30 mph speed limit.

4. **Sideswipe Same Direction Crashes:**
    - **Weather Condition:** Mostly happen in clear weather.
    - **Lighting Condition:** Usually in daylight.
    - **Traffic Control Device:** Typically where there are no controls.
    - **Posted Speed Limit:** Commonly in 30 mph zones.
5. **Turning Crashes:**
    - **Weather Condition:** Most often in clear weather.
    - **Lighting Condition:** Mainly in daylight.
    - **Traffic Control Device:** Frequently at traffic signals.
    - **Posted Speed Limit:** Usually occur in areas with a 30 mph speed limit.

Severity Analysis Of Most Crash Types

FIRST_CRASH_TYPE	INJURIES_FATAL	INJURIES_INCAPACITATING	INJURIES_TOTAL
ANGLE	0.358022	0.001268	0.032912
FIXED OBJECT	0.223867	0.006291	0.038680
OTHER OBJECT	0.155309	0.002641	0.027866
PARKED MOTOR VEHICLE	0.042378	0.000549	0.005535
PEDALCYCLIST	0.713511	0.002752	0.095329
PEDESTRIAN	0.929764	0.012450	0.176390
REAR END	0.183370	0.000232	0.011273

SIDESWIPE OPPOSITE DIRECTION	0.159860	0.000630	0.015743
SIDESWIPE SAME DIRECTION	0.071389	0.000242	0.005866
TURNING	0.242863	0.000669	0.020790

The severity analysis of the most common crash types shows the following trends in terms of average injuries per crash:

1. **Angle Crashes:**
    - **Total Injuries:** 0.358 injuries per crash on average.
    - **Fatal Injuries:** Approximately 0.002 fatalities per crash.
    - **Incapacitating Injuries:** Around 0.025 incapacitating injuries per crash.
2. **Parked Motor Vehicle Crashes:**
    - **Total Injuries:** 0.042 injuries per crash on average.
    - **Fatal Injuries:** No fatalities on average.
    - **Incapacitating Injuries:** Approximately 0.004 incapacitating injuries per crash.
3. **Rear End Crashes:**
    - **Total Injuries:** 0.17 injuries per crash on average.
    - **Fatal Injuries:** No fatalities on average.
    - **Incapacitating Injuries:** Around 0.007 incapacitating injuries per crash.
4. **Sideswipe Same Direction Crashes:**
    - **Total Injuries:** 0.07 injuries per crash on average.
    - **Fatal Injuries:** No fatalities on average.
    - **Incapacitating Injuries:** Approximately 0.008 incapacitating injuries per crash.
5. **Turning Crashes:**
    - **Total Injuries:** 0.20 injuries per crash on average.
    - **Fatal Injuries:** Approximately 0.001 fatalities per crash.
    - **Incapacitating Injuries:** Around 0.021 incapacitating injuries per crash.

The **correlation analysis** provides insights into factors that are most strongly associated with different severities of crashes:

FIRST_CRASH_TYPE TOTAL INJURIES	FATAL INJURIES	INCAPACITATING INJURIES	
LIGHTING_CONDITION_DARKNESS, LIGHTED ROAD	0.055574	0.022365	0.030263
TRAFFIC_CONTROL_DEVICE_PEDESTRIAN CROSSING SIGN	0.017409	0.003561	0.012163
TRAFFIC_CONTROL_DEVICE_RAILROAD CROSSING GATE	NaN	0.003599	NaN
TRAFFIC_CONTROL_DEVICE_SCHOOL ZONE	NaN	0.004973	NaN
TRAFFIC_CONTROL_DEVICE_STOP SIGN/FLASHER	0.065791	NaN	0.016498
TRAFFIC_CONTROL_DEVICE_TRAFFIC SIGNAL	0.091177	0.003722	0.026475
WEATHER_CONDITION_CLEAR	NaN	NaN	0.009895
WEATHER_CONDITION_RAIN	0.022902	NaN	NaN

Total Injuries:

Strongest positive correlations are observed with specific traffic control devices, notably traffic signals and yield signs, as well as certain weather conditions like cloudy/overcast weather.

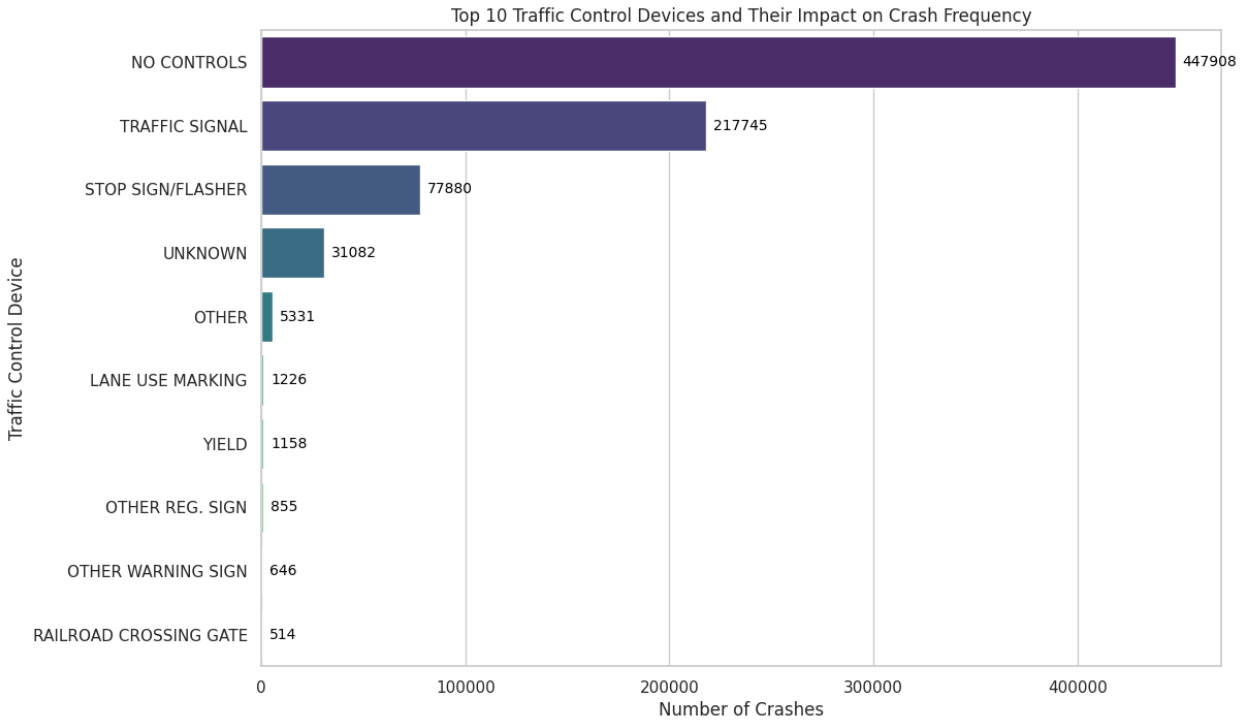
Fatal Injuries:

Fatal injuries show a notable correlation with darkness in lighted road conditions and the presence of stop signs/flashers. Clear weather also shows a positive correlation with fatal injuries.

Incapacitating Injuries:

Similar to total injuries, incapacitating injuries show a significant correlation with cloudy/overcast weather and certain lighting conditions like dawn. Traffic signals also show a positive correlation.

Top 10 Traffic Control Devices and Their Impact on Crash Frequency:



- A lack of controls is associated with the highest number of crashes, with approximately 450,000 incidents, emphasizing the need for adequate traffic control measures.
- Traffic signals are related to over 200,000 crashes, which may suggest issues with signal timing or driver compliance.

TRAFFICWAY TYPE	Number of Crashes	Avg Total Injuries	Avg Fatal Injuries
NOT DIVIDED	342724	0.179275	0.001011
DIVIDED - W/MEDIAN (NOT RAISED)	126562	0.217052	0.001440
ONE-WAY	100670	0.101622	0.000937
PARKING LOT	53694	0.044344	0.000187
FOUR WAY	47469	0.437929	0.001981
DIVIDED - W/MEDIAN BARRIER	45066	0.264928	0.002646
OTHER	21542	0.169800	0.001260

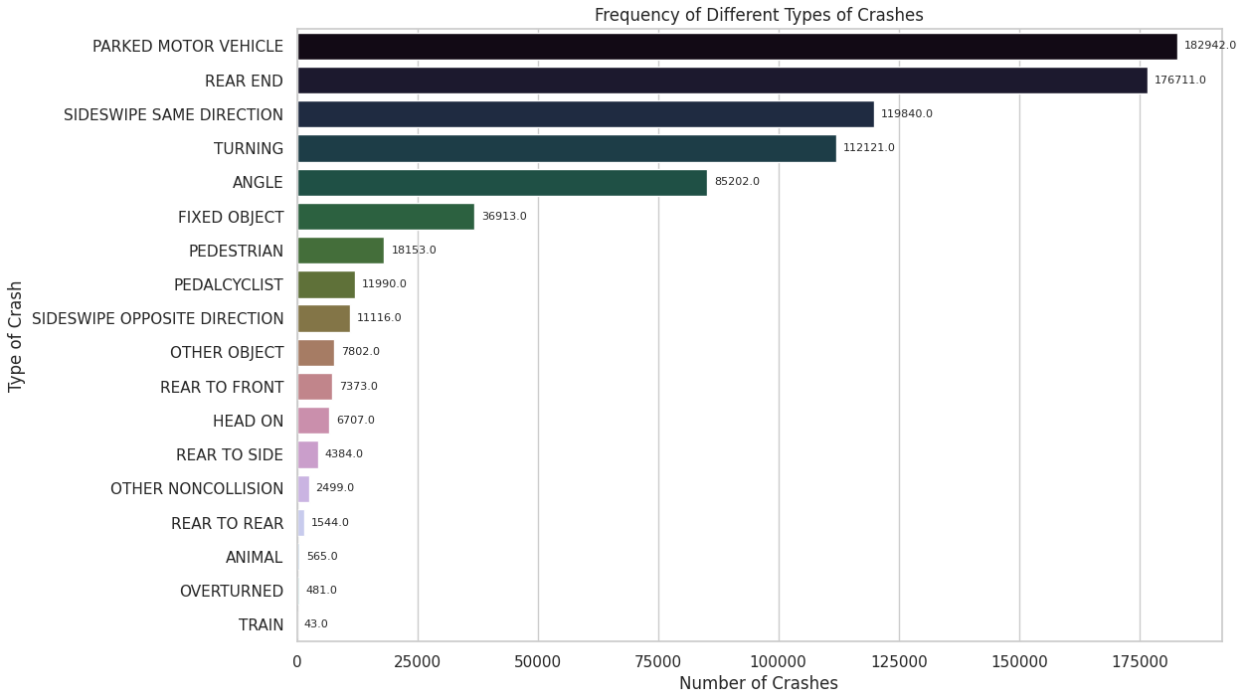
TRAFFICWAY TYPE	Number of Crashes	Avg Total Injuries	Avg Fatal Injuries
ALLEY	13013	0.100224	0.000850
T-INTERSECTION	9632	0.378288	0.002288
UNKNOWN	9131	0.079899	0.000220

Findings for CTA Management:

1. **Trafficway Type Impact on Injuries:**
    - Crashes on "Four Way" intersections have the highest average total injuries per crash (0.438), indicating a greater risk of injuries in these situations.
    - "T-Intersections" also show a higher average total injuries per crash (0.378), suggesting a need for targeted safety measures at these intersections.
2. **Fatal Injuries Trends:**
    - Fatal injuries are relatively rare across all trafficway types, with average rates ranging from 0.000187 to 0.002646 per crash.
    - Crashes at "Divided - W/Median Barrier" and "T-Intersection" show slightly elevated rates of fatal injuries.
3. **Safety in Parking Lots:**
    - Crashes occurring in "Parking Lot" areas have a lower average total injuries per crash (0.044), indicating relatively safer conditions in these settings.
4. **Concerns in Not Divided Roads:**
    - "Not Divided" roads have a lower average total injuries per crash (0.179), but this type accounts for the highest number of crashes (342,724).
    - Further investigation into specific causes on "Not Divided" roads may help identify areas for improvement.
5. **Unknown Trafficway Type:**
    - Crashes categorized as "Unknown" for trafficway type show a relatively low average total injuries per crash (0.080), but with 9,131 incidents, understanding and addressing the causes in this category could contribute to overall safety efforts.
6. **Alleys and One-Way Streets:**

- Crashes in "Alley" and "One-Way" settings have lower average total injuries per crash (0.100 and 0.102, respectively), suggesting these areas may generally have lower risk.

Key Findings on Crash Types and Frequencies in Chicago



Interpretation

1. **Parked Motor Vehicle Collisions** are the most frequent type of crashes, indicating a significant issue with vehicles colliding with stationary vehicles. This may suggest problems with street parking layouts, driver inattention, or the need for better regulation of parking in high-traffic areas.
2. **Rear-End Collisions** are the second most common, highlighting a trend that could be associated with driver distraction, tailgating, sudden stops, or perhaps fluctuating traffic flow, particularly in areas where CTA vehicles frequently stop and start.
3. **Sideswipe Collisions (Same Direction)** are also prevalent, which may point towards issues with lane changing, merging, and the navigation of CTA buses alongside other vehicles on busy roads.
4. **Turning Accidents** and **Angle Collisions** are common as well, often occurring at intersections. This may reflect challenges related to traffic signal timings, intersection design, or driver behaviors such as failing to yield or improper lane use during turns.

5. **Collisions with Fixed Objects** are considerable and might be linked to driver error or environmental conditions like poor visibility or road design that does not adequately account for CTA vehicles' size and maneuvering capabilities.
6. **Pedestrian** and **Pedalcyclist Incidents** represent a significant number of crashes, underscoring the necessity for enhanced pedestrian and cyclist safety measures, especially in areas serviced by CTA routes.
7. **Sideswipe Collisions (Opposite Direction)**, although less frequent, still represent a considerable number, suggesting issues with road markings, driver distraction, or possibly intoxication.
8. **Collisions Involving Other Objects**, **Rear to Front**, **Head-On**, and **Rear to Side** incidents, while lower in frequency, indicate diverse challenges in traffic management that could include driver error, recklessness, or external factors like road conditions.
9. **Other Noncollision Events**, such as vehicle breakdowns or other incidents not involving a collision, still factor into the total number of traffic disruptions affecting CTA operations.
10. **Rear to Rear**, **Animal**, **Overturned Vehicles**, and **Train** incidents are the least frequent but represent unique scenarios that can have severe consequences and indicate specific areas where targeted safety interventions could be beneficial.

Implications for CTA Management: The trends observed suggest several areas for potential improvement in CTA's operational zones. These include:

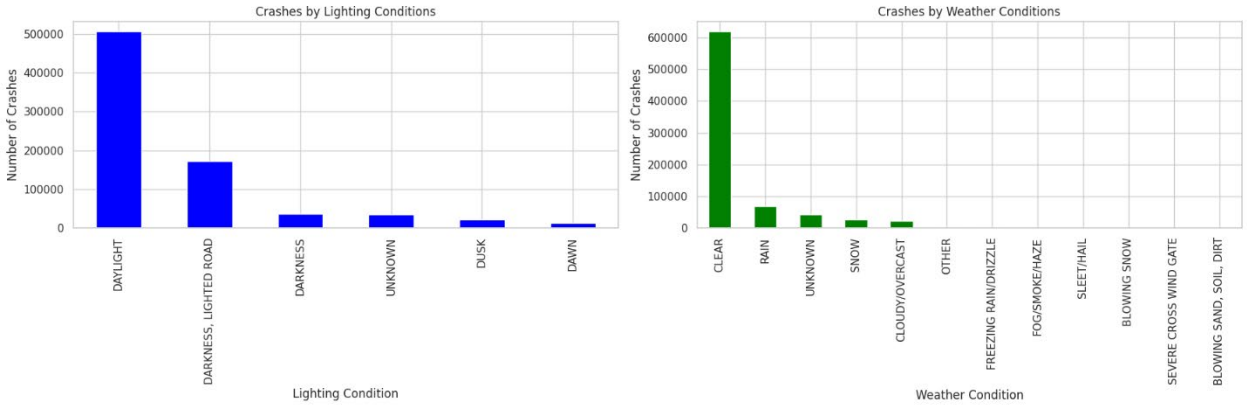
- Enhancing parking regulations and street designs to prevent collisions with parked vehicles.
- Implementing or improving public education campaigns focused on safe following distances and attentive driving to reduce rear-end collisions.
- Reviewing and improving lane markings and signage to prevent sideswipe incidents.
- Working with city planners to reevaluate intersection designs and traffic signal timings to reduce turning and angle collisions.
- Increasing the visibility of CTA vehicles and improving pedestrian and cyclist infrastructure in heavily trafficked areas.

Recommendations: CTA management should consider a multifaceted approach to addressing these trends, potentially including:

- Collaborative efforts with city traffic engineers to improve road layouts and signals.
- Targeted enforcement of parking and driving regulations, particularly in high-incident areas.
- Education and awareness campaigns aimed at reducing distracted driving and promoting road safety.

- Enhanced training for CTA drivers focusing on defensive driving techniques, particularly in relation to the most common types of collisions identified.

These findings and recommendations should serve as a basis for further investigation and action to improve traffic safety in areas serviced by the CTA.

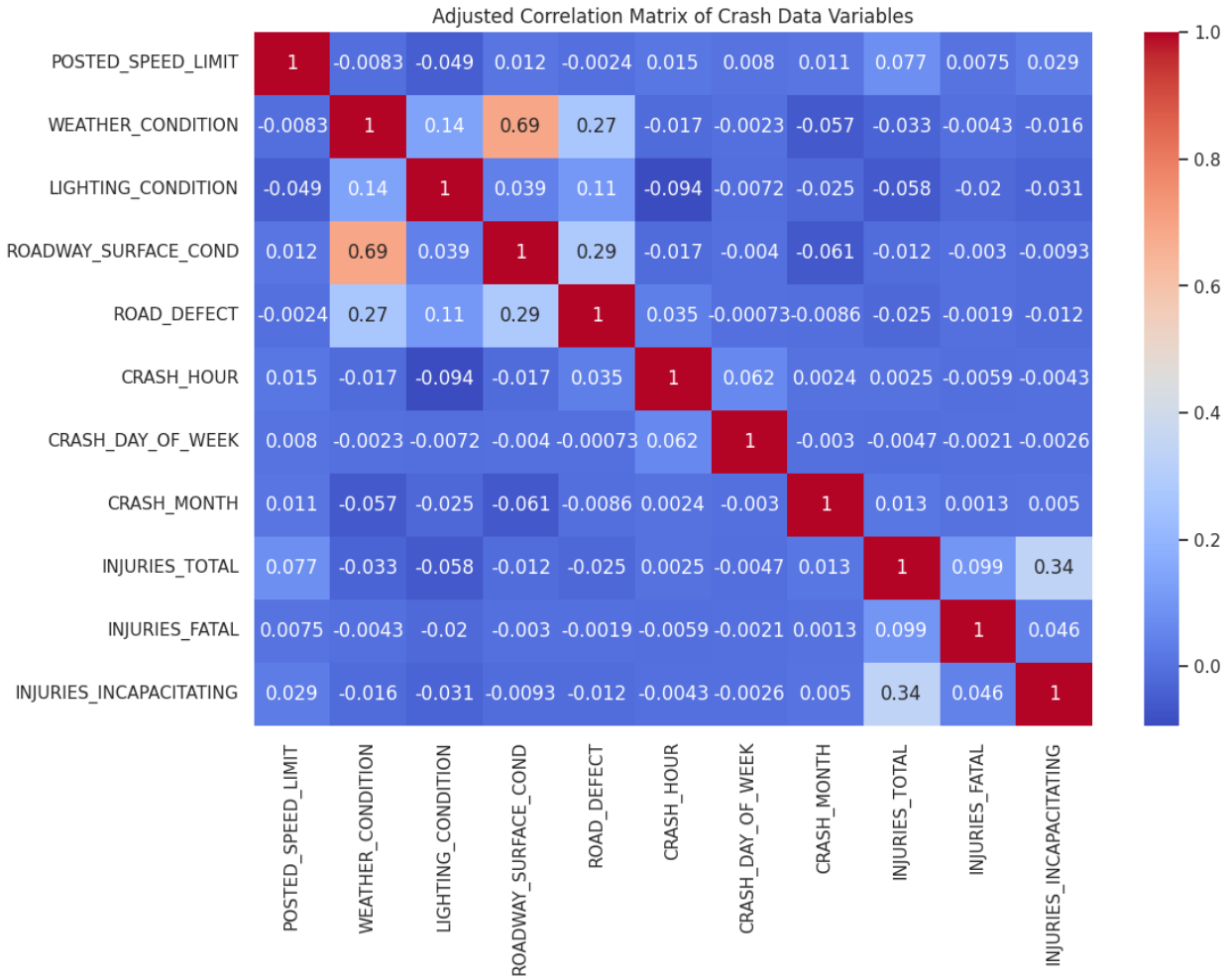


Crashes by Lighting Conditions:

- Daylight conditions see the highest number of crashes with over 400,000 incidents, indicating a higher occurrence during times of high visibility, which might be correlated with increased traffic volume during the day.
- Darkness without street lights has the next highest with approximately 150,000 crashes, pointing to the significant impact of poor lighting on road safety.

Crashes by Weather Conditions:

- The majority of crashes occur during clear weather conditions, with a frequency nearing 620,000. This suggests that driver overconfidence or inattention during good weather may be contributing factors.
- Rain is the second most common weather condition followed by unknown and snow for crashes with around 70,000 incidents, highlighting the influence of weather on road safety.



Correlation Matrix of Crash Data Variables:

- The matrix shows that 'LIGHTING_CONDITION' and 'ROADWAY_SURFACE_COND' have a notable correlation with 'WEATHER_CONDITION', suggesting that adverse weather conditions can deteriorate road conditions and visibility, leading to crashes.
- 'INJURIES_TOTAL' and 'INJURIES_FATAL' have some correlation, indicating that crashes with a higher number of total injuries are likely to include fatalities.