Road Vehicle Collisions



**Student ID:** Daniel Giedraitis

**Programme Title:** Software Development

**Programme Code:** CW\_KCSOF\_B

**Module Title:** Data Visualisation

**Lecturer Name:** Agnes Maciocha

**Submission Date:** 2024-04-17

**Word Count:** 2054

# Introduction

The unsettling reality of road vehicle collisions strikes a chord with many of us. We’ve all experienced those heart-wrenching moments—the sudden blare of sirens, the flashing lights, and the sinking feeling knowing that someone’s been hurt on the road. But the impact of these incidents extend well beyond the immediate scene, sending shockwaves through our communities, disrupting lives, and leaving a lasting imprint on our collective consciousness. It’s a toll that extends beyond individuals and families, affecting the very fabric of our society—our healthcare systems, transportation networks, and economic vitality. Indeed, the burden is heavy, and it demands our full attention.

In embarking on this analysis, I am driven by a profound desire to delve into the data behind these collisions. But this work is more than just crunching numbers, it’s about unraveling the complex patterns of factors that contribute to these incidents. My goal is to uncover patterns, identify trends, and gain insights that can help us navigate the complexities of road safety. Through the lens of data visualization, I seek to present these findings in a way that resonates with policymakers, transportation authorities, and everyday citizens. My goal is simple yet ambitious, to create meaningful conversations and encourage action toward creating safer roads for everyone.

# 

# Research Questions / Problem

As I delve into the complexities of road vehicle collisions, I’m faced with a number of challenging questions:

**1. Primary Factors Driving Collisions:** What are the primary factors propelling road vehicle collisions, and to what extent do they contribute to the overall accident rates? Delving into the root causes of these incidents is essential for developing effective prevention strategies and interventions.

**2. Temporal Influences on Collision Frequencies:** How do temporal factors, such as the time of day, day of the week, and month of the year affect collision frequency and severity? Understanding these temporal patterns is crucial for identifying peak periods and seasonal variations, which will allow for more focused accident prevention strategies.

**3. Distribution of Injury Types:** What insights can be gleaned from the distribution of injury types resulting from collisions, and how do these insights inform emergency response protocols and healthcare resource allocation? Understanding the severity and frequency of different types of injuries allows us to tailor emergency response strategies and allocate healthcare resources more effectively.

**4. Urban vs. Rural Collision Patterns:** Are there noticeable differences in collision patterns between urban and rural areas, and what are the implications of these disparities for infrastructure planning and traffic management? Investigating the differences in collision occurrences between urban and rural environments sheds light on the unique challenges each setting presents, guiding the development of location-specific interventions.

These questions are driving my exploration, guiding me as I navigate through the data to uncover meaningful insights.

# 

# Rationale

The reason I’m focusing on road vehicle collisions is simple: they have a significant and far-reaching impact on our communities, extending well beyond the immediate scene of the accident. Each year, millions of lives are disrupted, families are torn apart, and individuals suffer injuries or fatalities due to road traffic incidents. This isn’t just a statistic—it’s a ugly reality that demands our attention and action.

With road traffic injuries and fatalities imposing such a heavy toll on society, there’s simply no room for carelessness. We cannot afford to turn a blind eye to this critical public health and safety issue. Instead, we must confront it head-on with evidence-based strategies aimed at preventing unnecessary harm and suffering.

Data visualization serves as a powerful tool in this ongoing battle against road vehicle collisions. Its ability to display complex datasets into digestible insights makes it uniquely suited to our purpose. By using visual tools to express trends, patterns, and correlations within collision data, I aim to bridge the gap between raw data and actionable information.

Through compelling visualizations, we can empower decision-makers across various sectors to develop targeted interventions and allocate resources where they are most needed. Whether it’s identifying high-risk areas for infrastructure improvements, implementing public awareness campaigns to promote safe driving behaviors, or enhancing law enforcement efforts to reduce reckless driving, data visualization provides us with the insights necessary to drive meaningful change.

By shining light on the underlying factors contributing to road vehicle collisions, we can inform evidence-based decision-making and policy formulation. This proactive approach not only addresses the immediate challenges posed by road traffic incidents but also lays the groundwork for long-term solutions aimed at creating safer, more resilient communities.

In essence, my interest in road vehicle collisions stems from a sense of duty, recognizing the significant impact these incidents have on individuals, families, and society as a whole. Through data visualization and analysis, I hope to amplify the voices of those affected by road traffic injuries and fatalities, pushing for change and working toward a future in which such tragedies are reduced, if not completely eliminated.

# 

# Data Presentation

My analysis is supported by a rich dataset 53,945 rows and 11 rows providing a detailed compilation of information related to common factors influencing road accidents, such as collision severity, weather conditions, road types, and contributing elements, offering valuable insights for the analysis and enhancement of overall road safety measures. Each offering a unique lens through which to view road vehicle collisions:

* **Year:** Provides temporal context, allowing us to analyze trends and changes over time.
* **Month:** Offers insights into seasonal variations and trends in collision frequencies.
* **Weekend** (Weekday or Weekend): Distinguishes between collisions occurring on weekdays versus weekends, highlighting potential differences in driving behaviors and risk factors.
* **Hour:** Indicates the time of day when collisions occur, facilitating analyses of temporal patterns and peak traffic hours.
* **Collision Type:** Classifies collisions into categories such as rear-end, head-on, and sideswipe, providing insights into the nature and severity of collisions.
* **Injury Type:** Describes the types of injuries sustained by individuals involved in collisions, ranging from minor bruises to severe trauma.
* **Primary Factor:** Identifies the primary contributing factor cited in collision reports, such as speeding, distracted driving, or adverse weather conditions.
* **Reported\_Location, Latitude, Longitude:** Geographic coordinates pinpointing the location of collisions, enabling spatial analyses and visualization of collision hotspots.

With these variables at my disposal, I aim to paint a comprehensive picture of road vehicle collisions, uncovering hidden trends, patterns, and insights that can inform evidence-based decision-making and policy creation.

# Data Manipulation

Before delving into the realm of data visualization and analysis, it is important to preprocess the dataset to ensure its integrity, consistency, and suitability for exploration. This involves a series of data manipulation steps aimed at cleaning, transforming, and aggregating the raw data to facilitate meaningful insights and actionable conclusions.

**Key data manipulation tasks may include:**

* **Duplicate Removal:** Identifying and removing duplicate records to eliminate redundancy and ensure data accuracy.
* **Missing Value Imputation:** Addressing missing or incomplete data points through imputation techniques such as mean substitution, predictive modeling, or deletion based on data characteristics and analysis requirements.
* **Variable Standardization:** Standardizing variable formats and units to enhance comparability and facilitate cross-referencing with external datasets or industry standards.
* **Categorical Encoding:** Converting categorical variables into numerical or ordinal representations to enable quantitative analysis and modeling.

By meticulously preparing the dataset through these manipulation steps, I lay a solid foundation for later visualization and analysis, reducing the risk of incorrect conclusions or misinterpretations caused from data quality issues.

# Data Analysis and Visualization

Armed with a cleansed and preprocessed dataset, I embark on a visual journey through the world of road vehicle collisions, leveraging the expressive power of data visualization to uncover insights, trends, and patterns hidden within the data. Through a diverse array of visualization techniques, I aim to transform raw data into actionable knowledge, empowering stakeholders to make informed decisions and enact meaningful change.

My analysis encompasses a wide spectrum of visualizations, each offering a unique perspective on collision data:

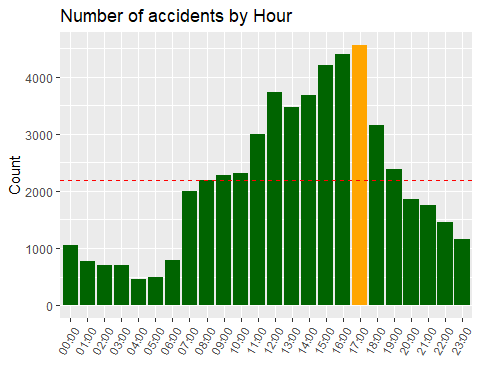
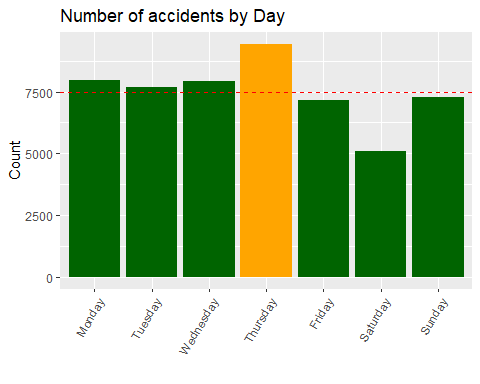
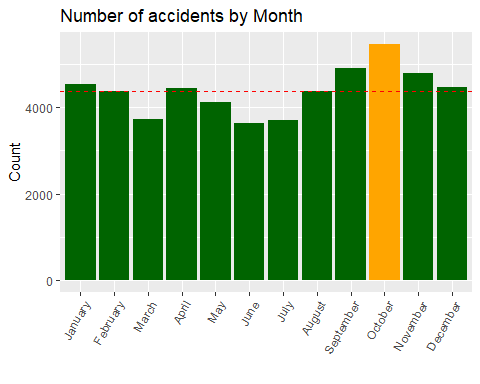
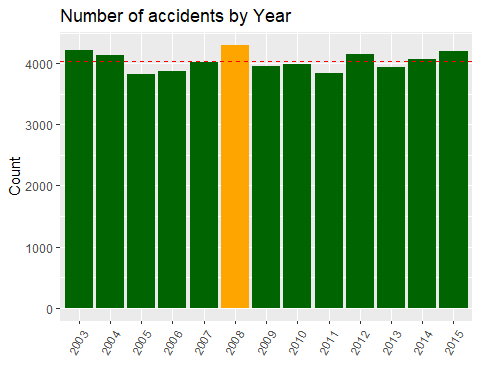
* **Temporal Analysis:** I investigate the temporal element by dissecting collision frequencies over various time intervals, uncovering complex temporal patterns, determining peak times, recognizing seasonal fluctuations, and detecting ongoing patterns in collision frequencies. This analysis is visualized using bar plots that indicate temporal fluctuations in collision occurrences.
* **Factor Analysis:** Engaging in a deep dive into the factors shaping collision occurrences, I investigated the complex interaction of variables, identifying the primary factors driving collisions and unraveling their intricate dynamics. This analysis is visualized through bar plots and pie charts, showcasing the frequency distribution of collision factors.
* **Injury Distribution Analysis:** By examining the distribution of injury types resulting from collisions, we get important insights that drive emergency response methods and healthcare resource allocation, determining the number of different injury types and their implications. This analysis is visualized through bar plots and pie charts, illustrating the distribution of injury types.

Through these visualizations, I aim to engage stakeholders on both an emotional and rational level, fostering empathy, awareness, and a sense of urgency in addressing the many challenges posed by road vehicle collisions.

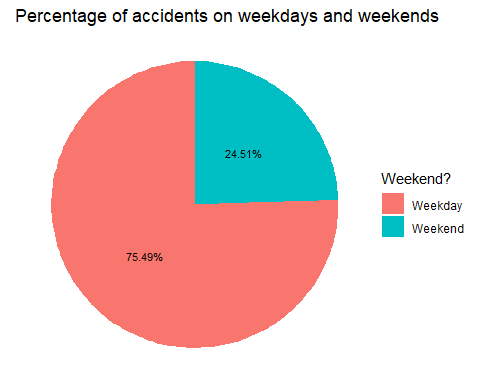
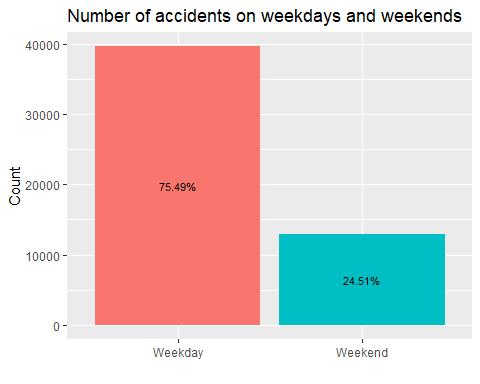
## 

## Number of accidents by years, months, days and hours

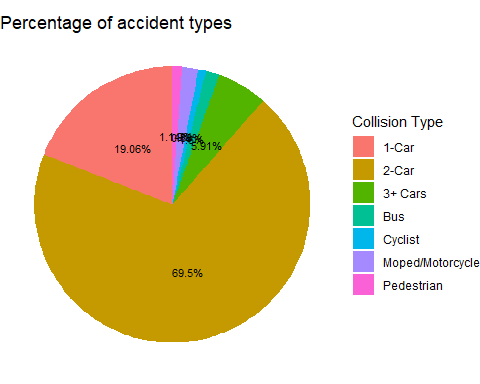
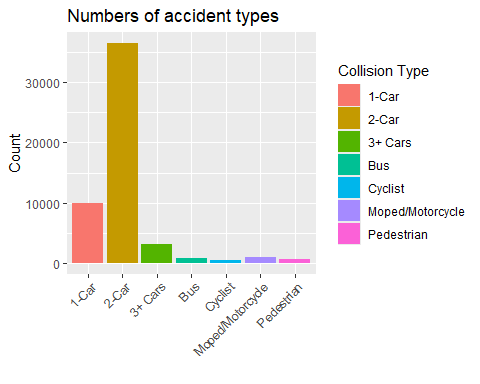
The analysis of collision frequencies over different time intervals provides valuable insights into temporal patterns. By visualizing the data, we can identify peak times, seasonal variations, and long-term trends in collision occurrences.



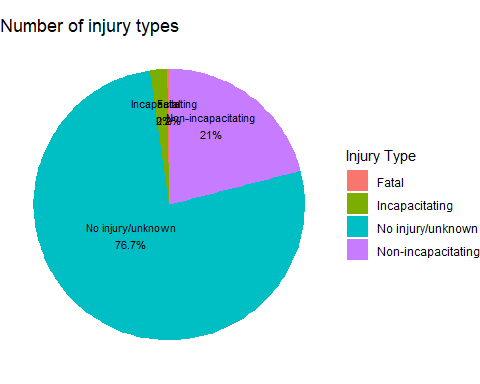
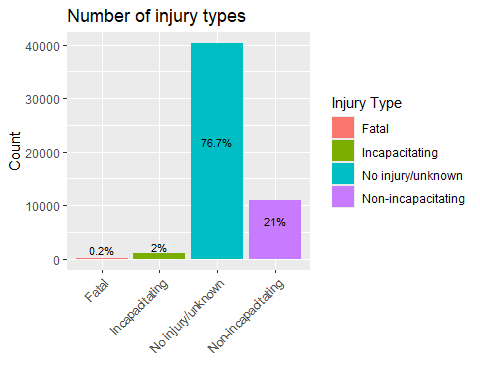
## Number and rates of accidents on weekdays and weekends

Understanding the distribution of accidents between weekdays and weekends is crucial for allocating resources and implementing targeted interventions. By comparing the number and rates of accidents on different days of the week, we can identify potential risk factors and tailor preventive measures accordingly. 

## Numbers and rates of accident types

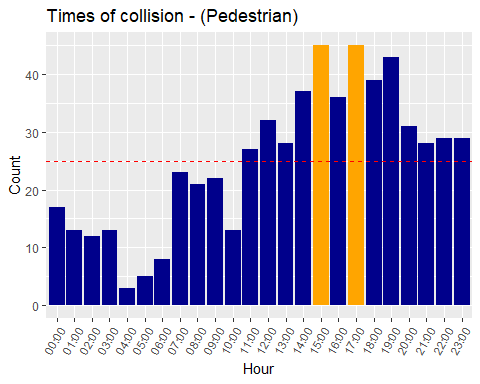
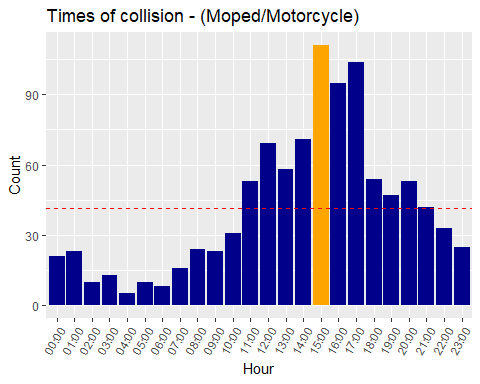
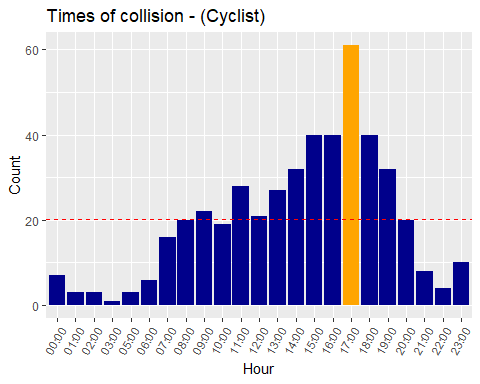
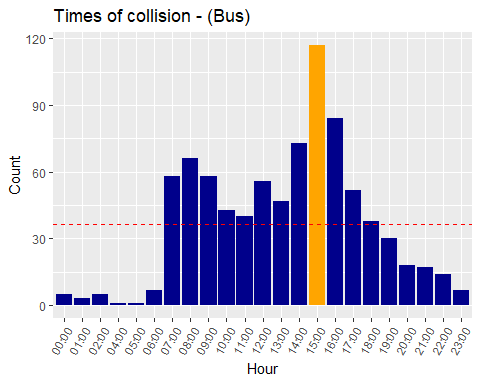
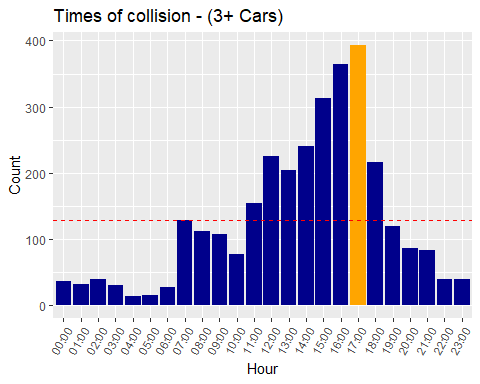
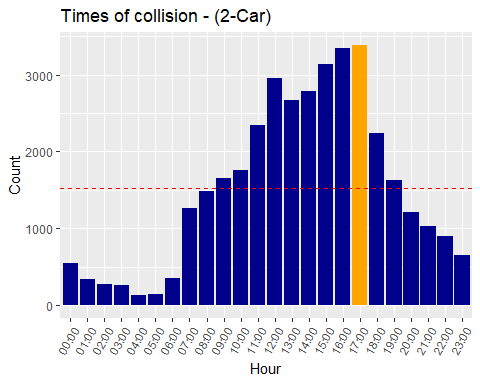
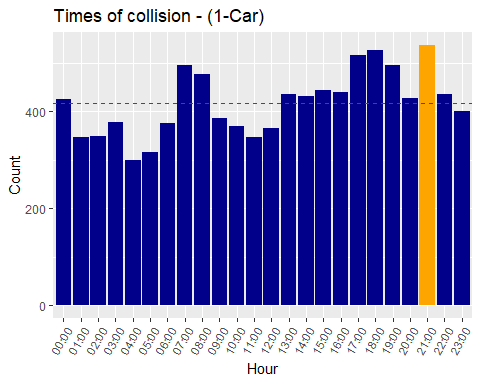
Analyzing the distribution of accident types provides insights into the predominant causes and mechanisms underlying collisions. By visualizing the data, we can identify the most common types of accidents and prioritize interventions aimed at reducing their occurrence. 

## Numbers and rates of injury types in accidents

Examining the distribution of injury types resulting from accidents informs emergency response protocols and healthcare resource allocation. By visualizing the data, we can assess the severity and frequency of different types of injuries and implement measures to mitigate their impact. 

## Times of collision types

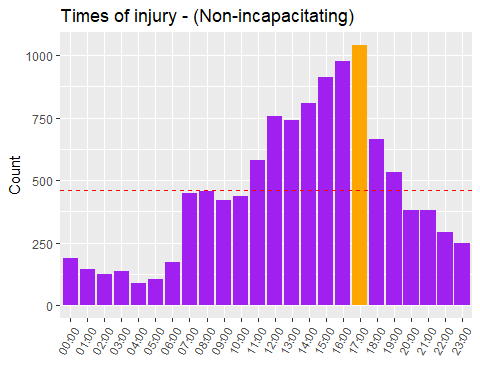
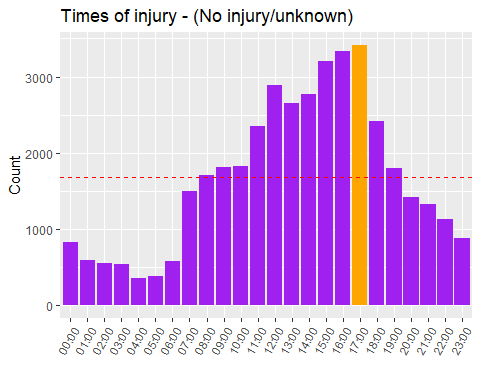
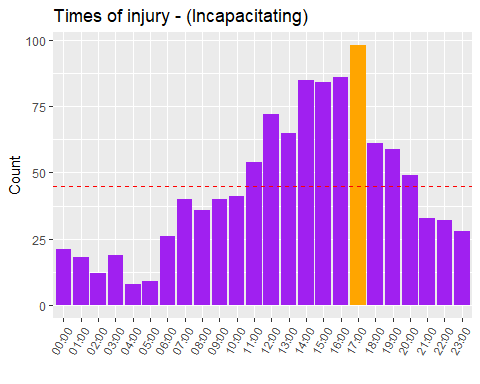
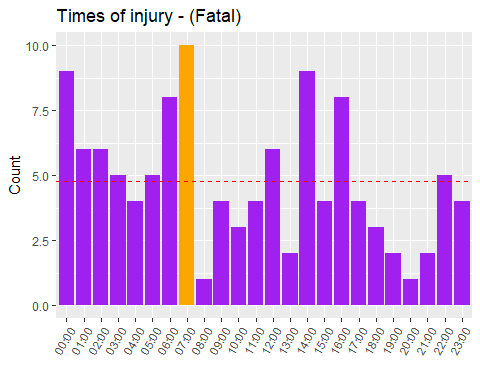
Analyzing the times of collision types helps identify temporal patterns and risk factors associated with different types of collisions. By visualizing the data, we can determine peak hours for specific collision types and implement targeted interventions to reduce their occurrence.



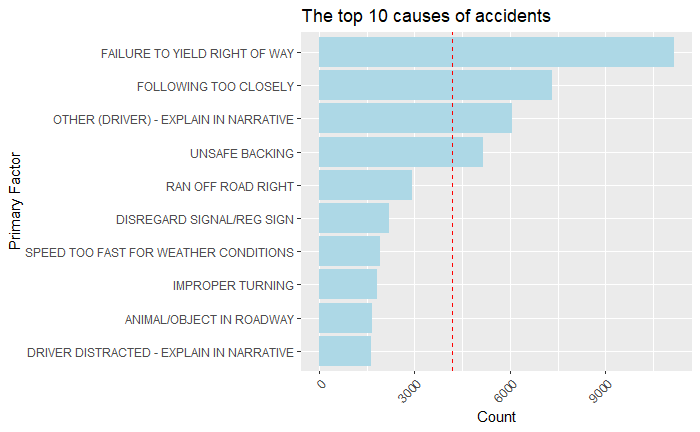
## 

## Times of injury for each injury type

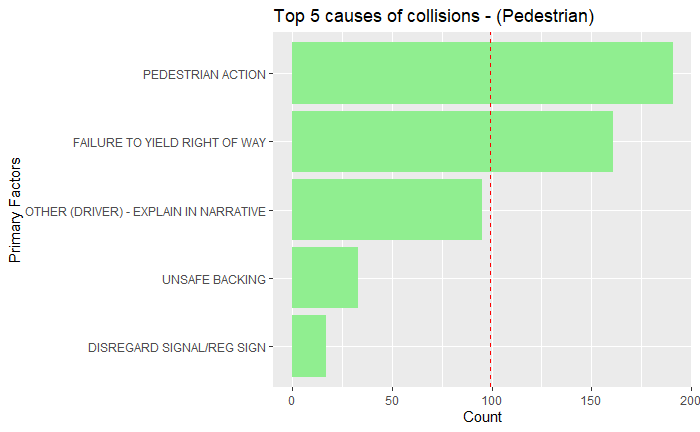
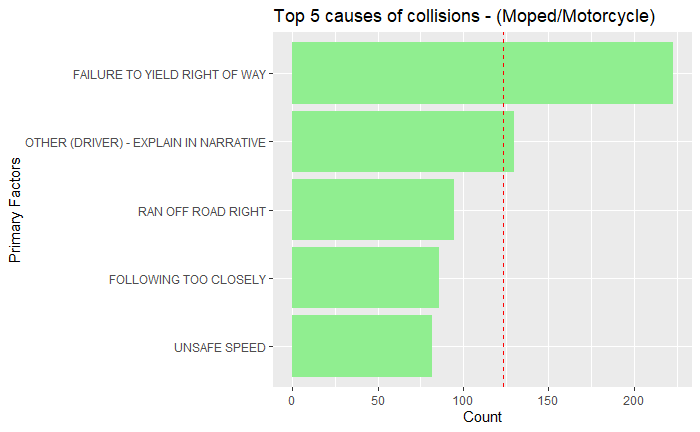
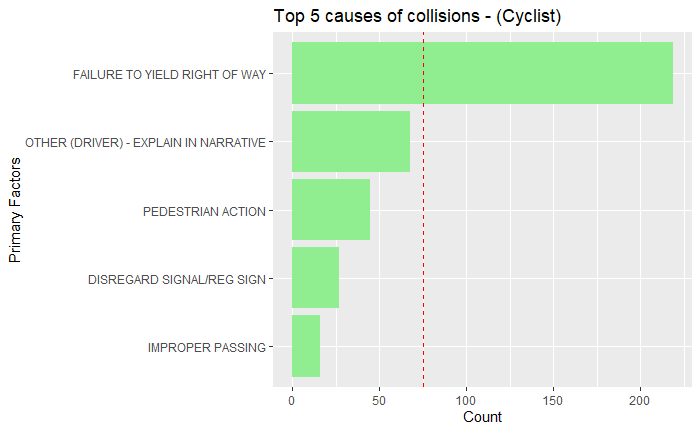
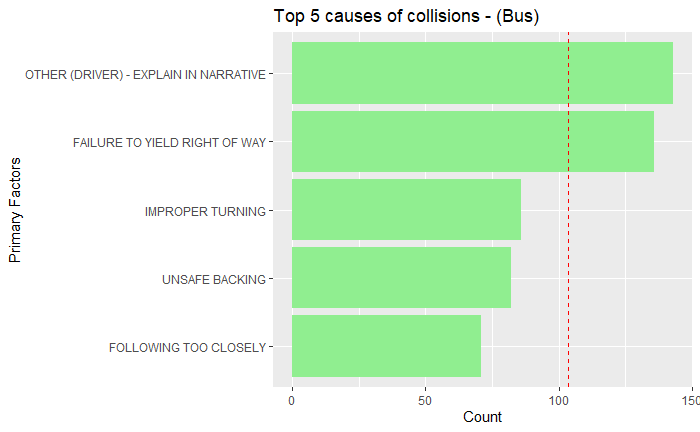
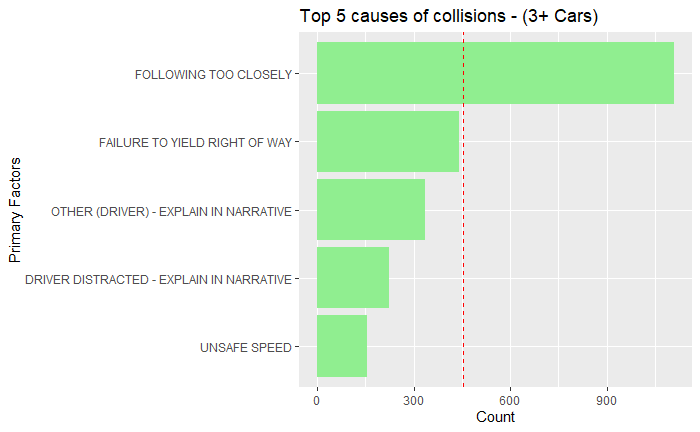
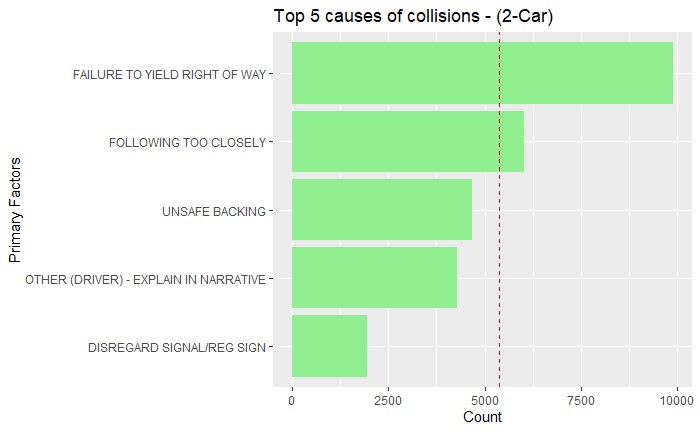
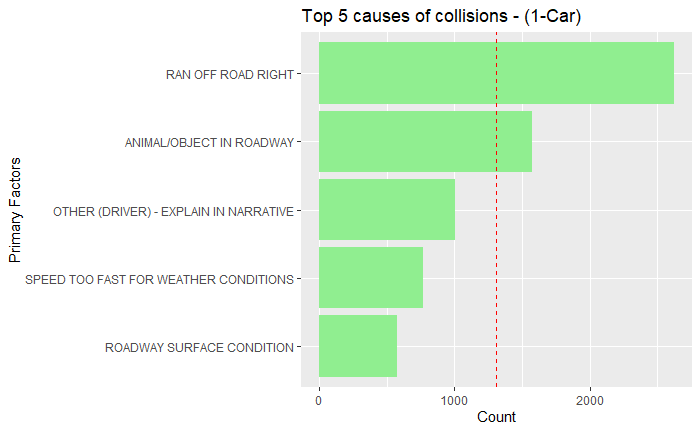
Examining the times of injury for each injury type provides insights into temporal trends and risk factors associated with different types of injuries. By visualizing the data, we can identify peak hours for specific injury types and develop strategies to prevent or minimize their occurrence.



## Causes of accidents

Identifying the primary factors contributing to accidents is essential for developing effective prevention strategies. By analyzing the data, we can pinpoint the most common causes of accidents and implement targeted interventions to address them. 

## Top 5 causes of collisions for each collision type

Analyzing the top causes of collisions for each collision type helps prioritize preventive measures and interventions. By visualizing the data, we can identify the primary factors contributing to different types of collisions and develop strategies to mitigate their impact. 

# Conclusions

In conclusion, my analysis of road vehicle collisions highlights the importance of data-driven approaches in tackling complex societal issues. By harnessing the power of data visualization, I’ve gained valuable insights into the underlying causes, temporal patterns, spatial distributions, and injury outcomes associated with collisions, laying the groundwork for evidence-based interventions and policy reforms.

Moving forward, it is important that we use these insights to drive positive change and develop a culture of road safety and responsible driving behavior. Whether through targeted infrastructure investments, public awareness campaigns, enhanced law enforcement efforts, or technological innovations in vehicle safety systems, there exists a multitude of avenues through which we can collectively work towards reducing the incidence and severity of road vehicle collisions.

As I navigate the road ahead, I remain strong in my commitment to leveraging data and technology as catalysts for progress. Empowering communities, saving lives, and building a safer, more sustainable future for everyone is most important. By continuing to innovate and collaborate across sectors, we can implement comprehensive solutions that address the root causes of road vehicle collisions while also ensuring the safety of all road users.