# Road to Safety: Traffic Accident Analysis

**Name: gerva3420**

## 1. Introduction

Welcome to the "Road to Safety: Traffic Accident Analysis" challenge, an initiative that blends the prowess of artificial intelligence with the pressing need for safer roads. In this journey, we venture into the realm of data analysis, focusing our efforts on the critical issue of traffic accidents in Catalunya.

As a data scientist participating in this challenge, I am thrilled to delve into a rich dataset, teeming with intricate details of traffic incidents. My objective is clear: to sift through this wealth of information, unearth patterns, and construct predictive models that shed light on the underlying causes of road accidents.

This report is designed to serve two fundamental purposes:

1. **Unveiling Insights**: Here, I aim to present a clear and accessible analysis of the data. By weaving through the complexities of the dataset, I intend to bring forward the primary factors contributing to traffic accidents, any unexpected trends that emerge, and the implications these hold for road safety.
2. **Proposing Solutions**: The insights gleaned are not just for academic interest; they are stepping stones towards tangible change. In this report, I will propose practical, actionable recommendations aimed at mitigating the risks of road accidents. These suggestions may range from policy reforms to modifications in road infrastructure, all geared towards a safer future on the roads of Catalunya.

## 2. Key Findings

The key findings of the "Road to Safety" report can be summarized as follows:

- General Trends (Section 3.1): A consistent decrease in traffic accidents, fatalities, and serious injuries in Catalonia from 2010 to 2021, indicating effective road safety measures.

- Accident Characteristics (Section 3.2): Most severe accidents occur in urban areas and conventional roads, primarily in good weather, suggesting the impact of human factors over environmental ones.

- Geographical Insights (Section 3.3): Barcelona leads in accident numbers, but smaller municipalities have higher per capita rates, highlighting diverse risk profiles across regions.

- Environmental Impact (Section 3.6): Weather conditions influence accident likelihood, with most severe accidents occurring in favorable conditions due to higher traffic volumes.

- Road and Traffic Features (Section 3.7): Roads with 100 km/h speed limits, particularly non-highway urban and conventional roads, are high-risk areas due to a combination of speed and complex traffic patterns.

- Vehicle Types and Accident Severity (Section 3.8): Accidents involving motorcycles or heavy vehicles tend to be more severe, though they are not the majority.

- Temporal Clustering (Section 3.9): Seasonal trends show more accidents in summer months, with a notable decrease in 2020 and 2021 due to COVID-19 impacts.

- Correlation Analysis (Sections 3.5 and 3.10): Detailed studies reveal correlations between lighting conditions and accident severity, suggesting the importance of visibility and illumination in preventing accidents.

# 3. Data visualization and In-depth Analysis

## 3.1 General Trends

(What are the overall trends in traffic accidents, fatalities, and serious injuries in Catalonia from 2010-2021?)

Figure 1, as described, provides a comprehensive view of traffic accident trends in Catalonia from 2010 to 2021, both in absolute and relative terms. The chart is divided into two subplots, with the top depicting absolute numbers and the bottom showing relative changes over time. Let's discuss the key findings from both aspects:

**Absolute Terms (Top Subplot)**

1. **Overall Decrease in Traffic Accidents:** The data indicates a general decline in the total number of traffic accidents over the 11-year period. This could be attributed to various factors, including improvements in road safety, better enforcement of traffic laws, or advances in vehicle safety features.
2. **Serious Injuries as the Most Common Outcome:** Among the three categories (fatalities, serious injuries, and light injuries), serious injuries are reported as the most frequent outcome of traffic accidents. This highlights the severity of many traffic incidents, even if they are not fatal.
3. **Ranking of Categories:** The order of prevalence is serious injuries, followed by light injuries, and lastly fatalities. This distribution suggests that while many accidents result in injury, a smaller proportion are fatal. The relatively lower number of fatalities, compared to injuries, could be indicative of effective emergency response and medical care.

**Relative Terms (Bottom Subplot)**

1. **Decreasing Trend Across All Categories:** All three categories - fatalities, serious injuries, and light injuries - show a decreasing trend in relative terms. This suggests that not only are there fewer accidents, but the severity of these accidents is also diminishing.
2. **Comparative Decrease in Categories:** While all categories show a decrease, light injuries and fatalities have decreased more significantly (around 40%) compared to serious injuries (around 30%). This disparity in reduction rates might reflect different rates of improvement in safety measures or medical care, or it could indicate changes in the nature or reporting of these incidents.
3. **Implications of Relative Trends:** The relative trends are particularly insightful because they provide a context for the absolute numbers. For instance, a decrease in light injuries and fatalities at a higher rate than serious injuries might imply that interventions have been more effective in preventing less severe and fatal accidents.

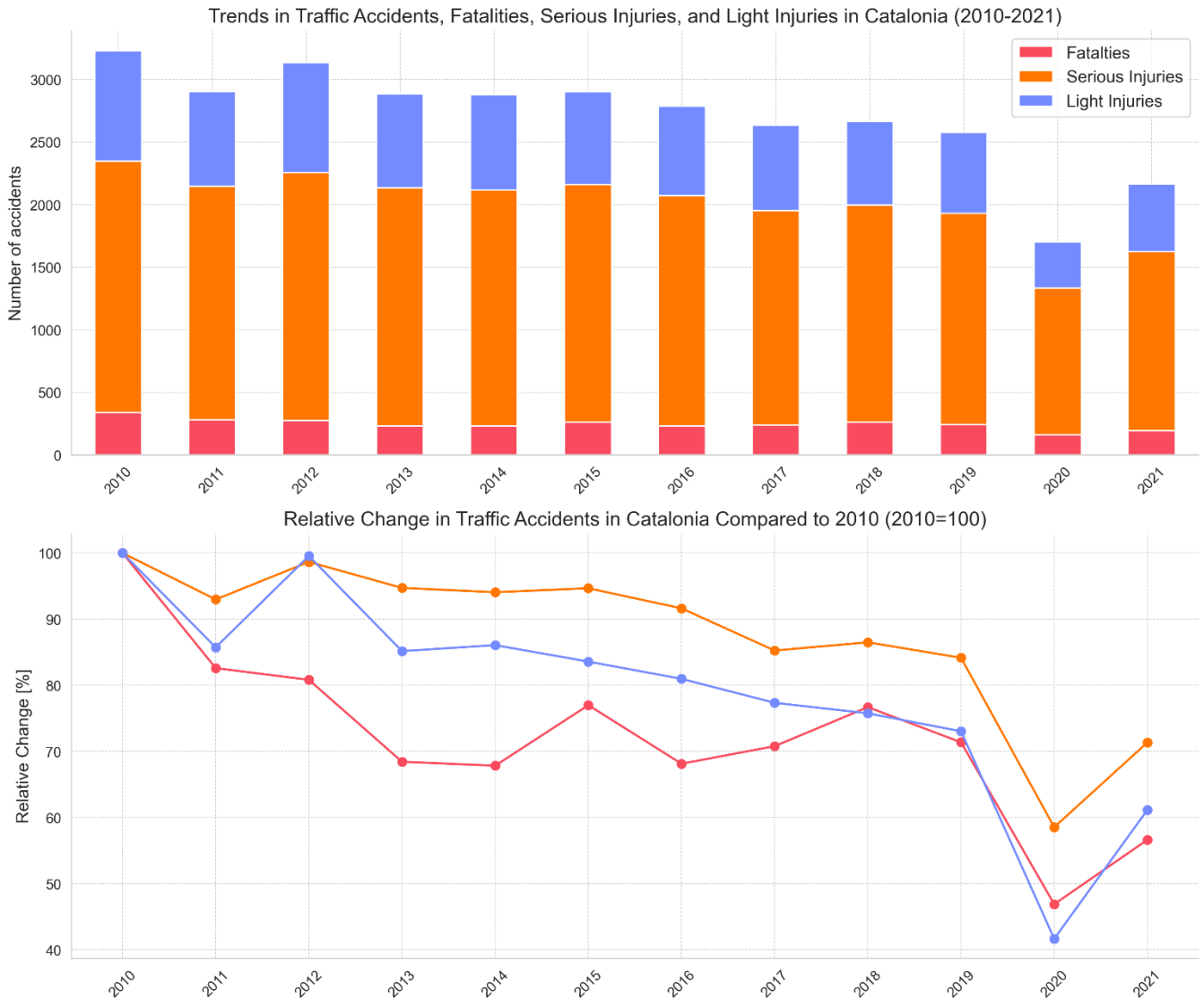


*Figure 1 Trends in traffic accidents, fatalities, and serious injuries in Catalonia from 2010-2021.*

**Overall Interpretation**

The data presents an encouraging picture of improving road safety in Catalonia. The overall reduction in traffic accidents and their severity is a positive sign, possibly reflecting successful road safety policies, public awareness campaigns, advancements in vehicle safety, and medical care improvements. However, the fact that serious injuries still constitute the largest proportion of accidents suggests a need for targeted interventions. It would be beneficial to understand why serious injuries have not decreased at the same rate as other categories and to address this issue specifically in future road safety strategies.

## 3.2 Accident Characteristics

What common characteristics (time of day, type of road, etc.) are observed in the most severe accidents?

Figure 2 provides a comprehensive overview of severe traffic accidents (fatalities and serious injuries) from 2010 to 2021, aggregated to highlight key characteristics. This figure breaks down the accidents based on various conditions such as time of day, type of road, weather conditions, lighting conditions, and road surface conditions. It offers insights into the most common scenarios under which severe accidents have occurred during this period.

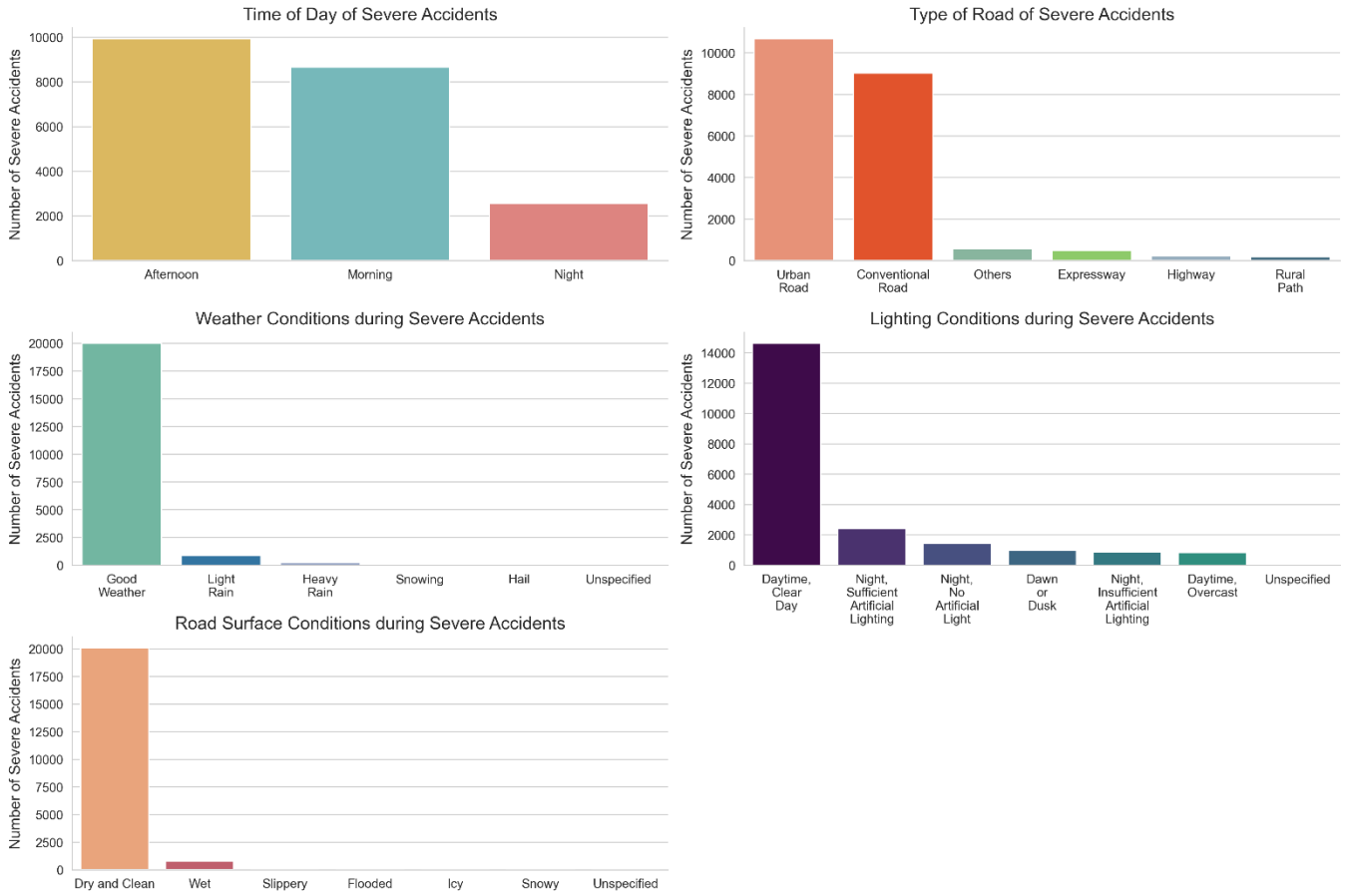


*Figure 2 Characterization of severe accidents conditions: time of day, type of road, Weather conditions, Light conditions and road surface.*

1. **Time of Day:** Most severe accidents occur in the afternoon and morning. The frequency of accidents drops significantly during the night.
2. **Type of Road:** Urban roads and conventional roads see the highest number of severe accidents. Highways have a notably lower incidence of severe accidents.
3. **Weather Conditions:** Good weather or light rain are the most common conditions during severe accidents. Extreme weather like heavy rain and snow are less common, possibly due to less traffic or more cautious driving in such conditions.
4. **Lighting Conditions:** Most accidents happen in clear daylight or at night with sufficient artificial lighting, indicating that poor lighting is not a major factor in severe accidents.
5. **Road Surface Conditions:** Dry and clean road surfaces are predominant in severe accidents, with wet conditions being less common. This might indicate good road maintenance or effective drainage systems.

Figure 3 delves into the trends of severe accidents over the years, from 2010 to 2021, offering a temporal perspective on the data. By examining the same categories as in Figure 2 — time of day, type of road, weather, lighting, and road surface conditions — this figure reveals how the patterns of severe accidents have evolved year by year. It provides a deeper understanding of whether certain types of accidents are becoming more or less common and highlights the effectiveness of road safety measures over time.

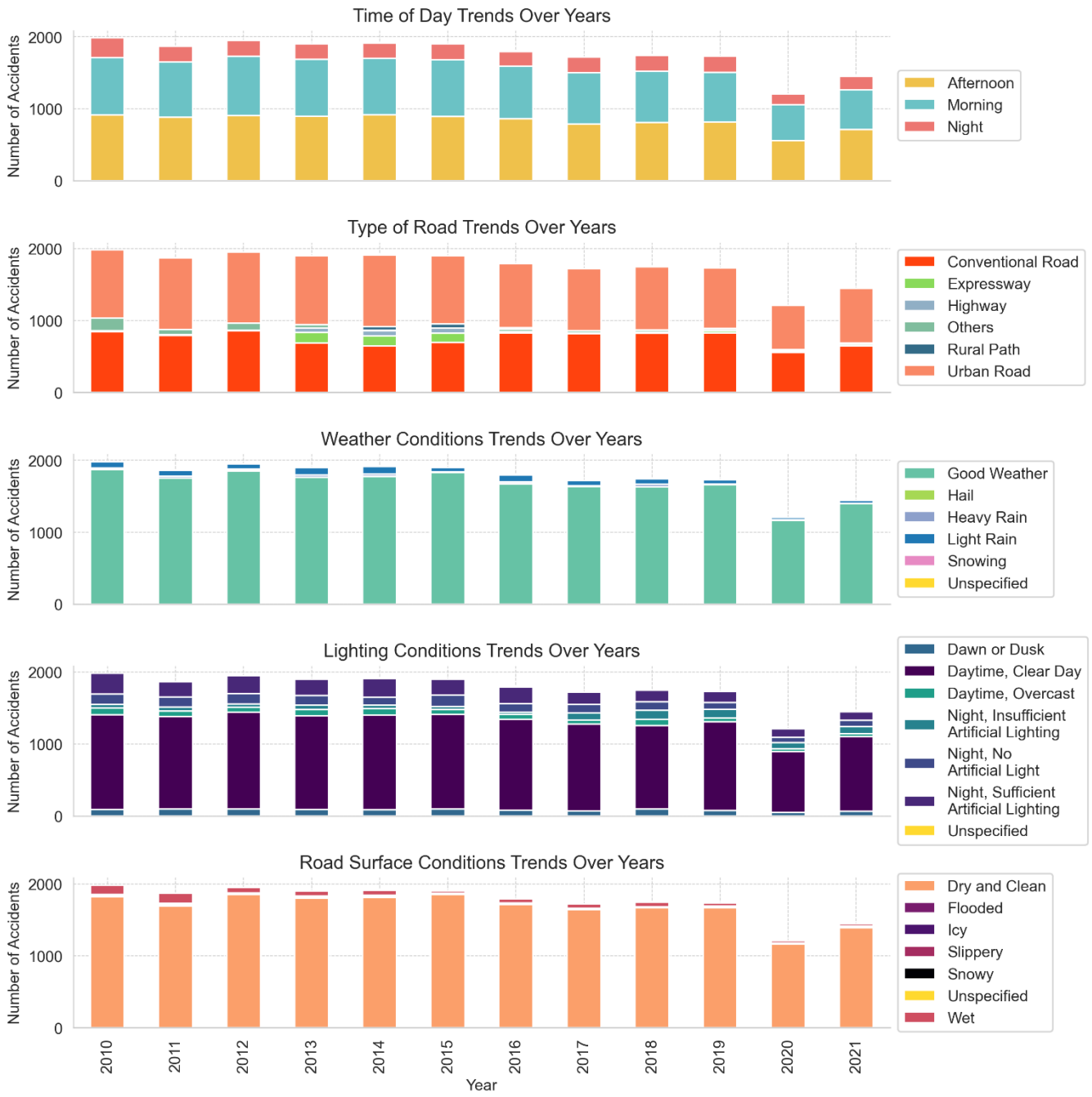


*Figure 3 Trends of severe accidents conditions over the years: time of day, type of road, Weather conditions, Light conditions and road surface.*

1. **Weather-Related Trends:** There is a noticeable decrease in severe accidents occurring during light rain and on wet roads over the years. This could suggest the effectiveness of weather-related safety policies or infrastructure improvements, like better road drainage or more effective public awareness campaigns during rainy conditions.

2. **Stability in Other Conditions:** The rest of the data aligns with the aggregated view in Figure 2, but the year-by-year analysis doesn't show significant shifts in other categories. This consistency might indicate that the overall driving conditions and behaviors have remained relatively stable over this period.

### 3.3 Geographical Insights

Which municipalities or counties in Catalonia have the highest incidence of traffic accidents? How does this correlate with population density or road network characteristics?

Barcelona, as the most populous municipality, unsurprisingly leads in the absolute number of accidents (Figure 4). However, when accounting for population size, Barcelona exhibits a lower rate of accidents per capita, contrasting with less populous municipalities. This finding is critical in understanding the impact of population density on traffic accident rates. Figure 5 provides a per capita perspective, showing that smaller municipalities, despite fewer total accidents, have higher accident rates per capita. This could indicate varying driving behaviors, road conditions, or safety measures in less populated areas compared to a metropolis like Barcelona.

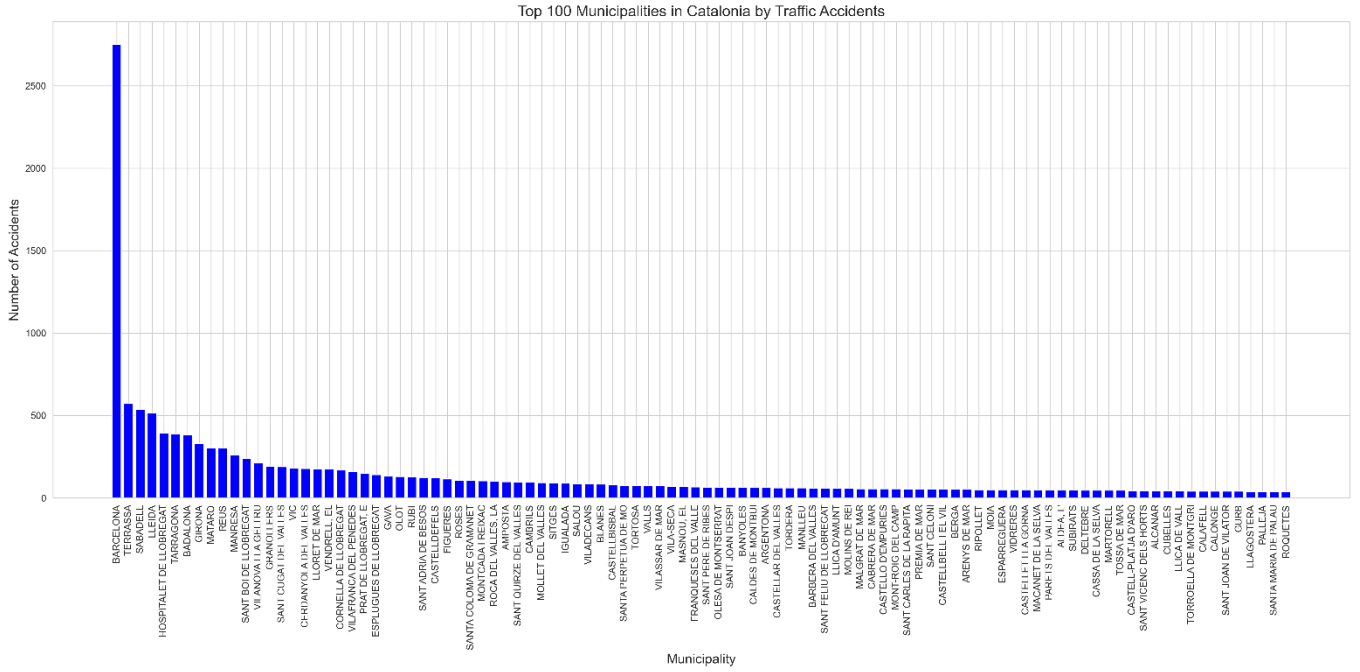


*Figure 4 top 100 municipalities per number of accidents in the period 2010-2021 in Catalonia.*

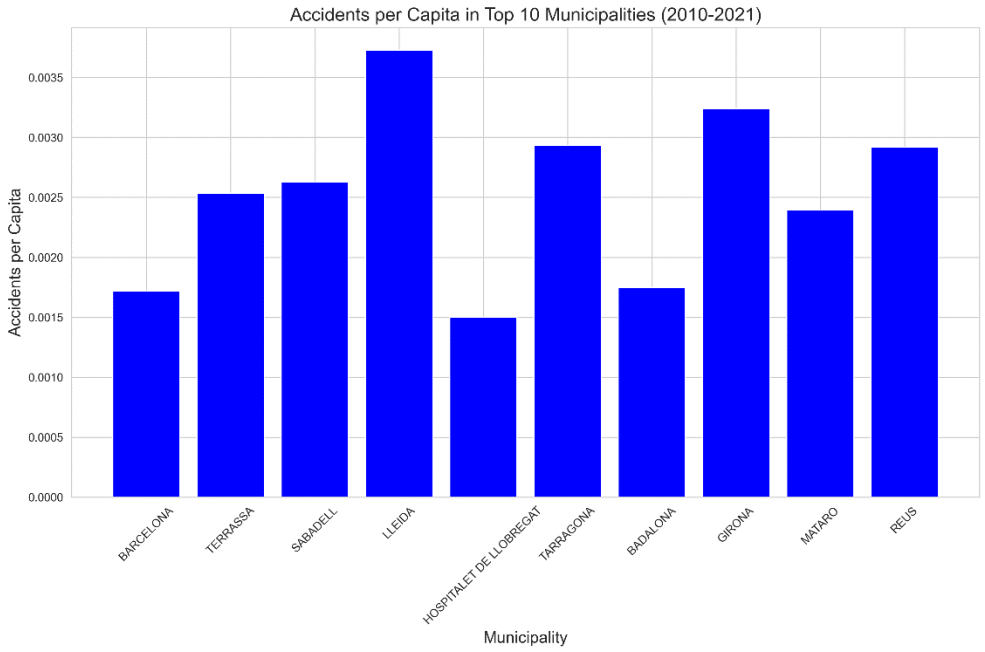


*Figure 5 top 10 municipalities per number of accidents per capita in the period 2010-2021 in Catalonia.*

A decreasing trend in accidents from 2010 to 2021 across the top 10 municipalities (Figure 6) suggests effective safety measures and possibly changes in traffic patterns or road usage over time. Figure 7 reveals a significant correlation between population size and accidents per capita. Municipalities with smaller populations tend to have higher accident rates per capita. This correlation may point to factors like lower traffic volumes, different types of roads, or varying levels of road safety infrastructure and enforcement in these areas.

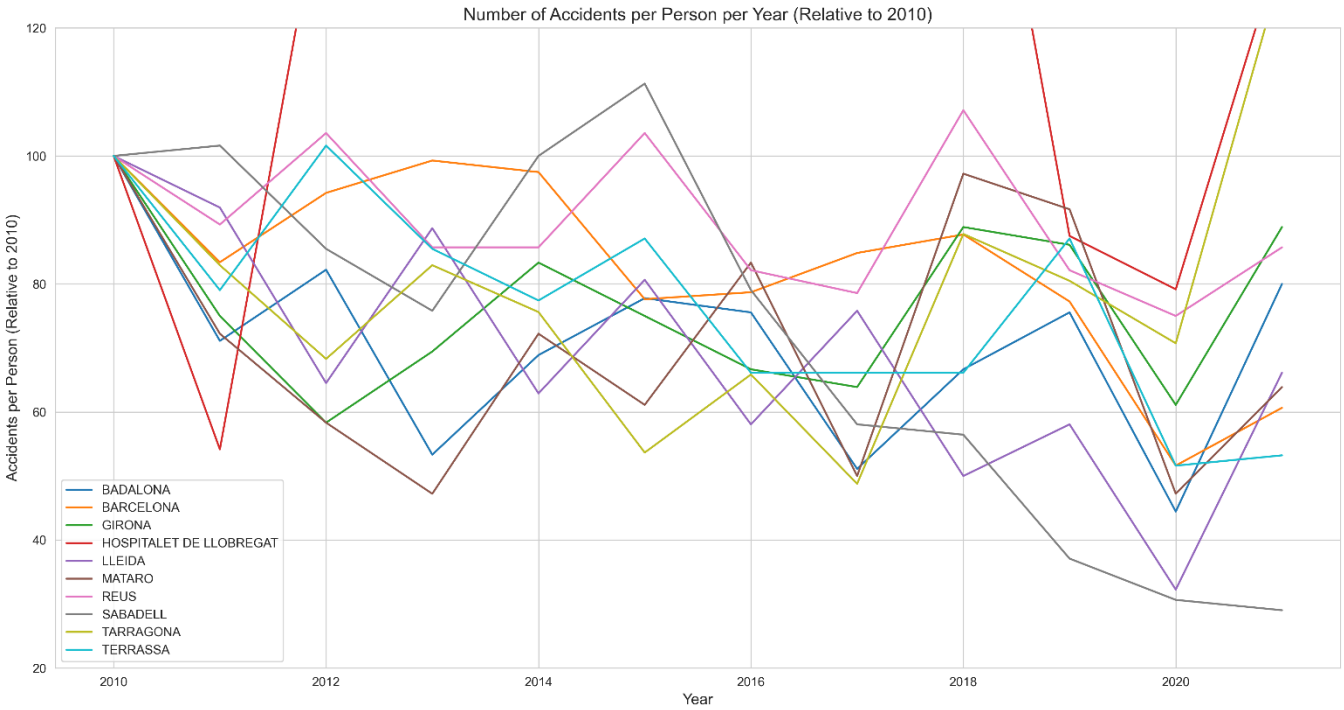


*Figure 6 Traffic Accidents per capita per year in top 10 Catalan Municipalities by number of accidents (2010-2021).*

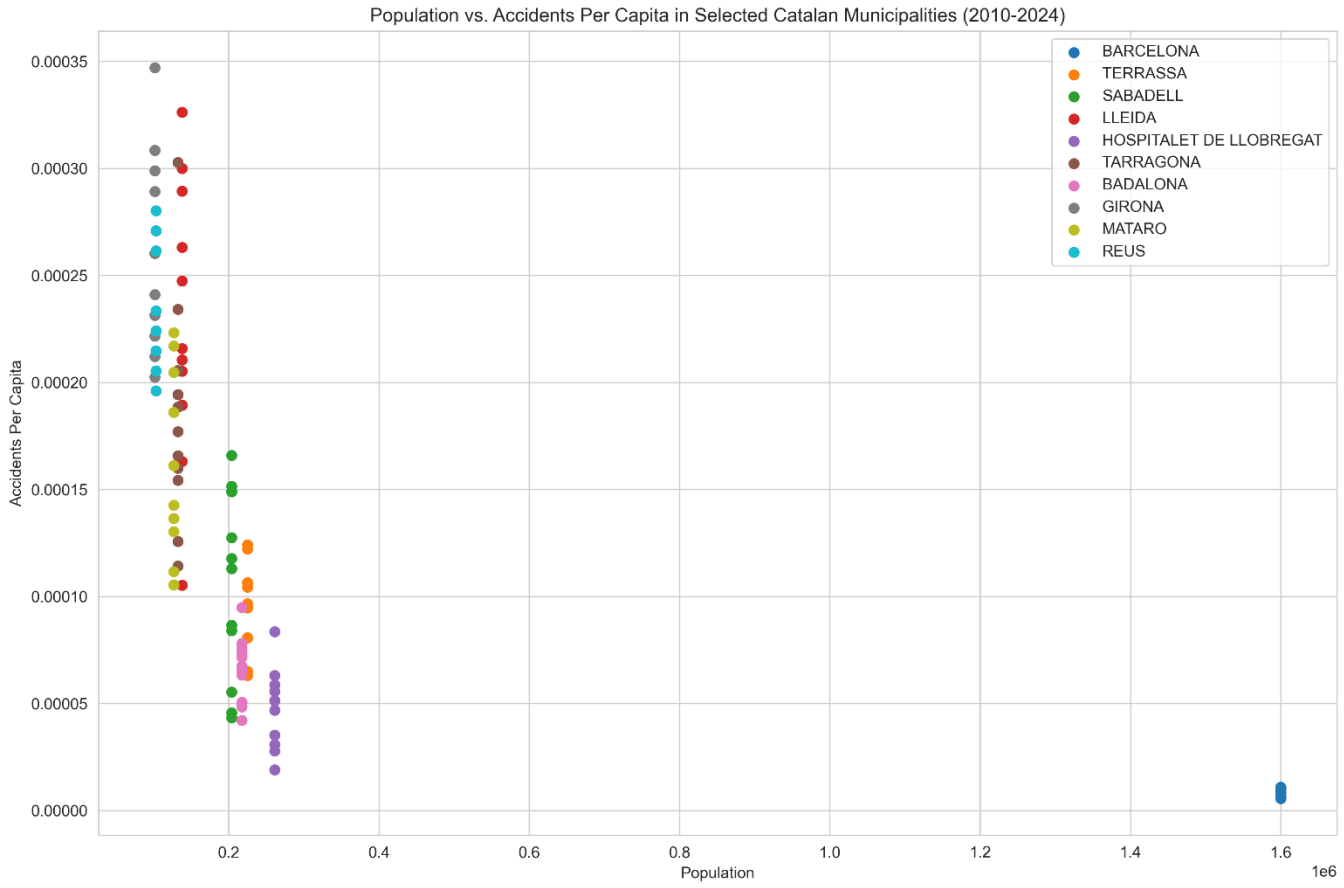


*Figure 7 Correlation between population and accidents per capita in top 10 Catalan Municipalities by number of accidents (2010-2021).*

## 3.4 Yearly Trends

How have traffic accident patterns (frequency, severity) changed yearly from 2010 to 2021?

From 2010 to 2021, traffic accident patterns in Catalonia have evolved significantly, as evidenced by Figures 1, 3, and 6:
Yearly Trends (Figure 1): There's been a general decrease in the frequency of traffic accidents over this period. This trend could be attributed to improved road safety measures, changes in traffic regulations, or advancements in vehicle safety features.
Severity of Accidents (Figure 3): The severity of accidents, measured in terms of injuries or fatalities, has also shown a decreasing trend. This suggests that, alongside the reduction in accident frequency, the accidents that do occur are less severe, possibly due to better emergency response systems and safer vehicle designs.
Top 10 Municipalities (Figure 6): In the top 10 municipalities by the number of accidents, there's a consistent decrease in accident rates from 2010 to 2021. This decline indicates the effectiveness of localized road safety initiatives and potentially shifts in urban traffic patterns.
These patterns highlight the effectiveness of road safety policies and technology improvements in reducing both the frequency and severity of traffic accidents in Catalonia over the past decade.

## 3.5 Day and Time Patterns

On what days of the week and times of day do most accidents occur? Are there notable differences between weekdays and weekends?

The analysis of traffic accidents by day of the week and time of day reveals distinct patterns. Notably, the highest number of accidents occurs on Fridays (see Figure 8), especially in the afternoon and evening hours. This trend might be attributed to increased travel at the end of the workweek. In contrast, the lowest number of accidents is typically observed on Mondays, with a gradual increase throughout the week.

A significant distinction is observed between weekdays and weekends. Saturdays and Sundays show a higher number of accidents during night hours compared to weekdays, possibly due to increased recreational activities

during these times. The afternoon remains a common time for accidents throughout the week, but the evening and night periods see a marked increase during weekends.

These insights suggest that different factors, such as work-related commutes during weekdays and leisure activities during weekends, significantly influence the occurrence of traffic accidents.

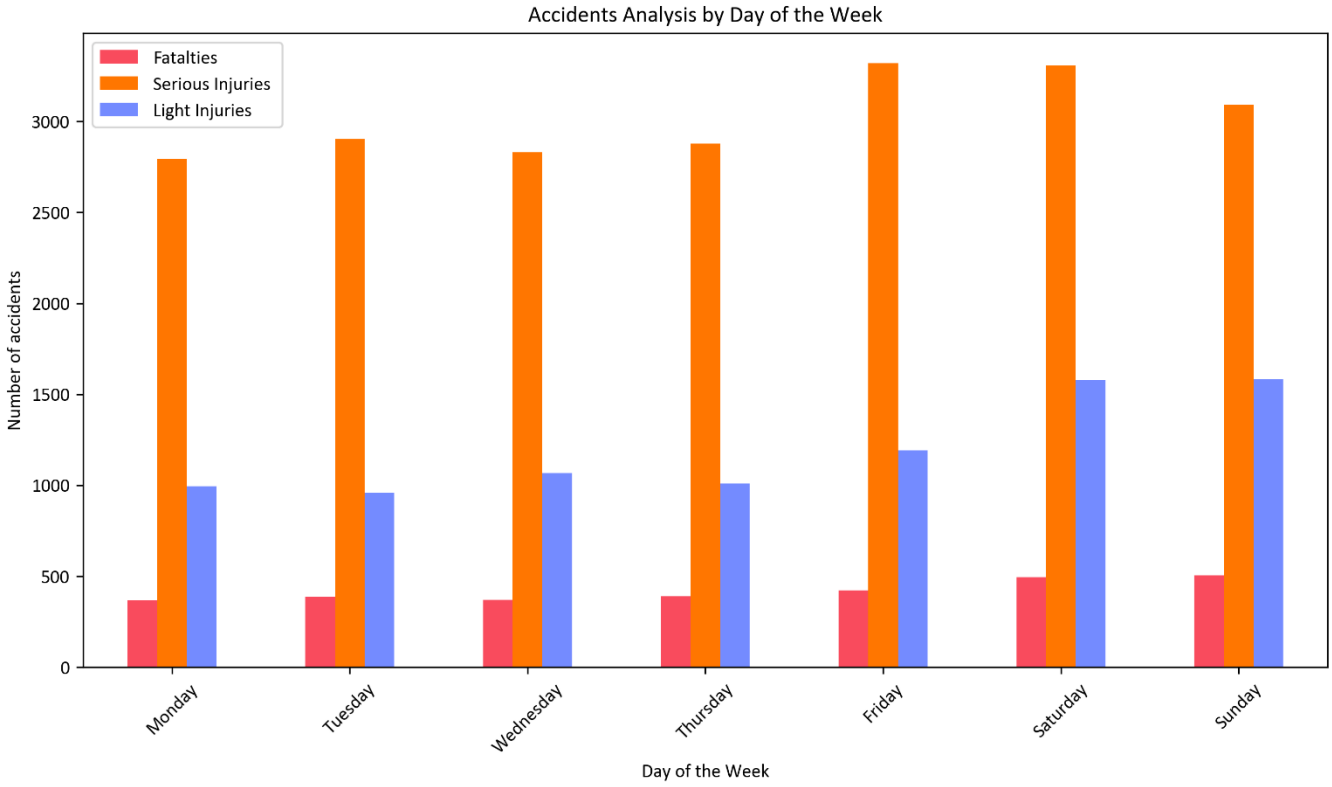


*Figure 8 Weekly Distribution of Road Traffic Accidents: Fatalities, Serious, and Light Injuries.*

### 3.6 Environmental Impact

How do different weather conditions affect the likelihood of accidents? Is there a correlation between visibility, road conditions, and accident severity?

The correlation between different weather conditions and the likelihood of traffic accidents is a significant aspect of road safety analysis. In the context of Catalonia, our analysis, particularly from Figures 2 and 3, indicates noteworthy trends:

1. **General Weather Conditions**: The majority of severe accidents occur in good weather or light rain conditions. This may be counterintuitive but can be explained by the higher volume of traffic during favorable weather conditions.

2. **Extreme Weather Impact**: Severe weather conditions, such as heavy rain and snow, are less common during accidents. This could be due to reduced traffic or more cautious driving behavior in such conditions.

3. **Weather-Related Trends Over Years**: There has been a noticeable decrease in severe accidents during light rain and on wet roads over the years, as shown in Figure 3. This suggests the effectiveness of weather-related safety policies or infrastructure improvements, like better road drainage systems or effective public awareness campaigns during rainy conditions.

Correlation between Visibility, Road Conditions, and Accident Severity

1. **Visibility and Lighting Conditions**: Most severe accidents happen in clear daylight or at night with sufficient artificial lighting. This observation, derived from the data in Figure 2, indicates that poor visibility is not a major contributing factor to severe accidents in the region.

2. **Road Surface Conditions**: Dry and clean road surfaces are most common during severe accidents. The infrequency of accidents on wet or compromised road surfaces might reflect efficient road maintenance or effective drainage systems, reducing the risk of accidents under such conditions.

3. **Yearly Changes in Conditions**: The year-by-year analysis in Figure 3 doesn't show significant shifts in conditions like visibility and road surface, suggesting that interventions and improvements in these areas have been consistently effective over time.

### 3.7 Road and Traffic Features

What impact do road features (such as speed limits and road types) and traffic density have on the occurrence of accidents?

Impact of Road Features on Accidents

1. **Speed Limits**: According to Figure 9 in the report, the majority of traffic accidents occur on roads with a speed limit of 100 km/h. This suggests that roads with this speed limit, typically major roads just below highway speeds, might pose unique risks, potentially due to a combination of high speeds and varying traffic patterns.

2. **Road Types**: Figure 10 provides a fascinating insight: the frequent occurrence of accidents on roads with a 100 km/h speed limit is more common on conventional and urban roads, not on highways. This indicates that such roads, possibly due to factors like intersections, varying traffic flows, and access points, are more accident-prone than highways, which are designed for higher speeds and smoother traffic flow.

Influence of Traffic Density

1. **High Traffic Density**: High-density traffic areas likely contribute to a higher number of accidents due to increased interactions between vehicles, leading to situations like rear-end collisions or merging accidents.

2. **Low Traffic Density**: In contrast, lower traffic densities, especially on high-speed roads, might contribute to more severe accidents due to factors like higher speeds and driver overconfidence in less crowded conditions.

This analysis underscores the complex relationship between road features, like speed limits and road types, and the occurrence of accidents. Roads with a speed limit of 100 km/h, especially conventional and urban roads, are identified as high-risk areas. The impact of traffic density further complicates this relationship, with both high and low densities contributing to accidents in different ways. The data suggests the need for targeted safety measures and road design considerations, particularly on non-highway roads with higher speed limits.

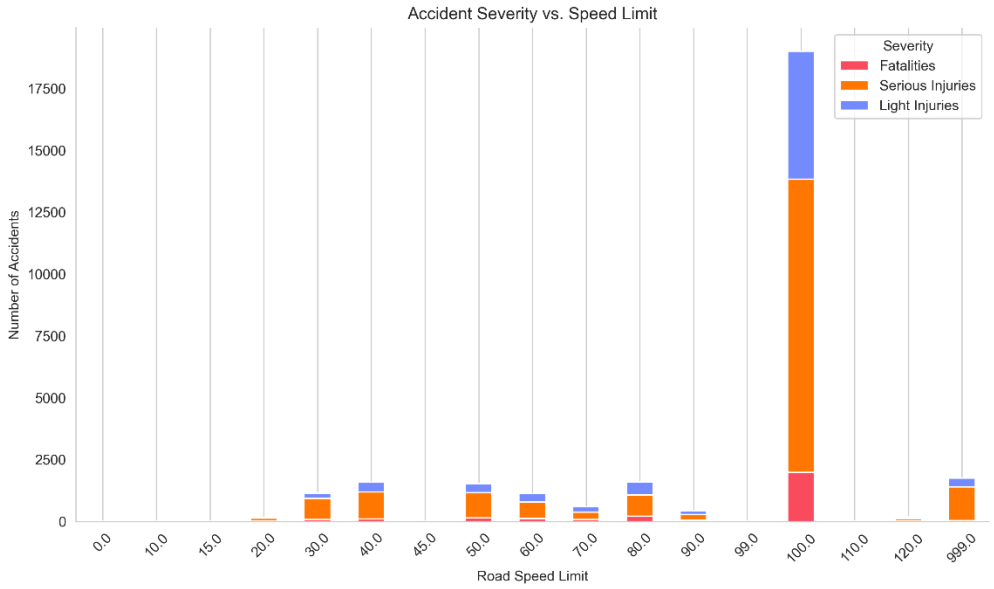


*Figure 9 Accident Frequency and Severity by Road Speed Limit.*

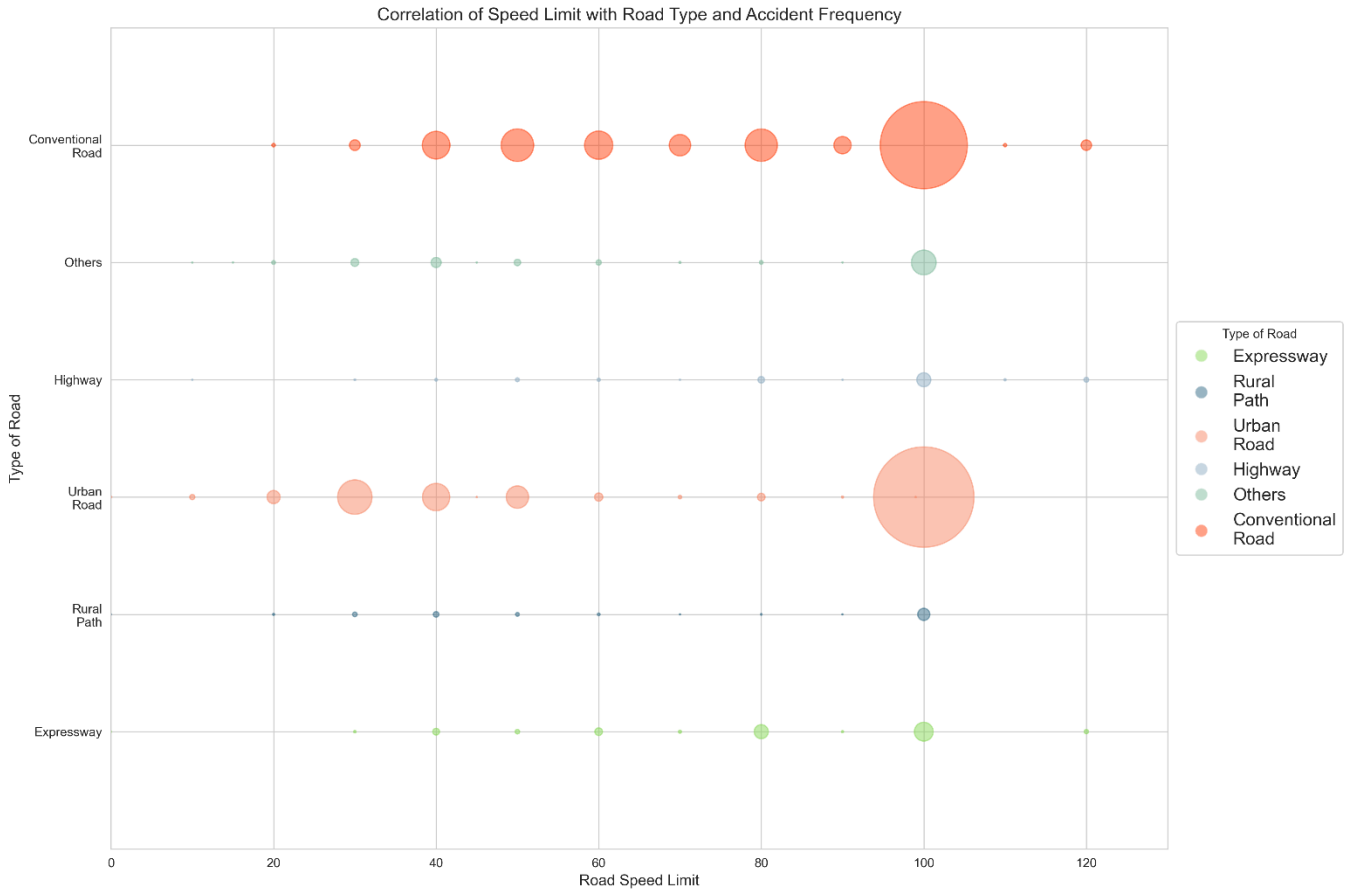


*Figure 10 Accident Distribution by Road Type and Speed Limit*

## 3.8 Vehicle Types and Accident Severity

Does the involvement of specific types of vehicles (like heavy trucks and motorcycles) correlate with more severe accidents?

This section examines how the involvement of specific vehicle types, particularly heavy trucks and motorcycles, correlates with the severity of accidents.

1. **General Observation**: According to Figure 11, a significant majority of accidents do not involve motorcycles or heavy vehicles. This indicates that while these vehicles are a factor in accident dynamics, they are not the predominant ones in road accidents.
2. **Motorcycle Involvement**: The second-largest category of accidents involves motorcycles without the presence of heavy vehicles. Given the vulnerable nature of motorcycles, these accidents are often severe, pointing towards a higher risk associated with motorcycle involvement.
3. **Heavy Vehicle Involvement**: The third category, involving heavy vehicles but no motorcycles, indicates a lesser but still notable frequency of accidents. The size and mass of heavy vehicles can contribute to the severity of accidents, although the frequency is lower compared to motorcycle-involved accidents.
4. **Comparison and Implications**: Comparing the severity and frequency of accidents involving these vehicle types, it is evident that while most accidents do not involve them, the presence of motorcycles and heavy vehicles in accidents tends to correlate with increased severity. This highlights the need for specific road safety measures tailored to these types of vehicles.

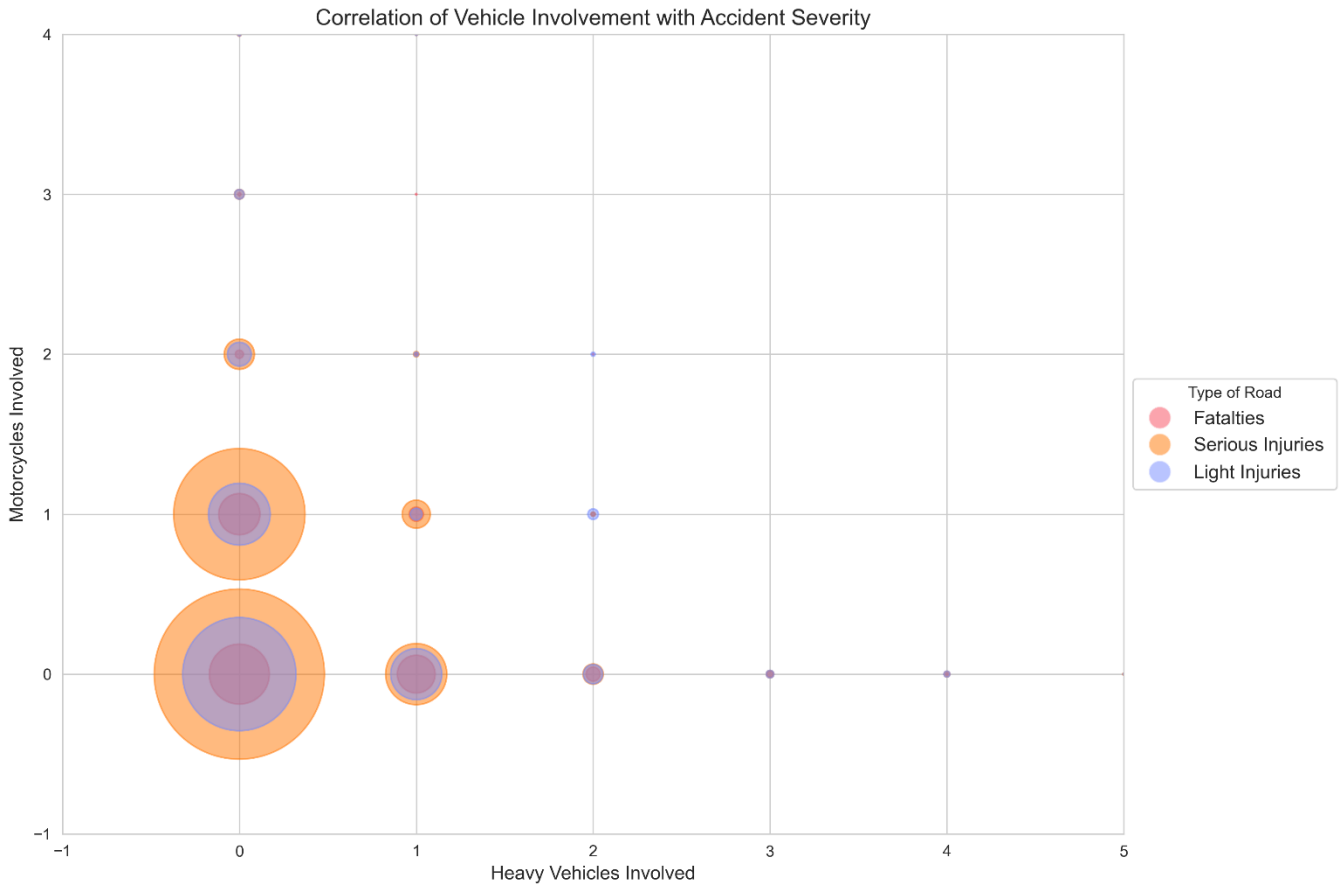


*Figure 11 Vehicle Type Impact on Accident Severity: Analyzing the Role of Motorcycles and Heavy Vehicles*

## 3.9 Temporal Clustering

Are there specific periods (months, years) where accident patterns cluster significantly? What might be the causes for these clusters?

1. **Seasonal Variation**: The heatmap in Figure 12 shows a clear pattern of increased accidents during the summer months. This could be attributed to several factors, such as increased travel and tourism, more daylight hours leading to longer driving times, and possibly more relaxed driving attitudes during vacation periods.
2. **Yearly Changes**: There's a noticeable decrease in overall accident numbers in 2020 and 2021. This trend aligns with the global COVID-19 pandemic, which led to widespread lockdowns and a significant reduction in vehicle usage during parts of these years.

Causes for Clustering

1. **Summer Peak**: The summer increase in accidents could be due to higher traffic volumes, more recreational travel.
2. **COVID-19 Impact**: The decrease in accidents during 2020 and 2021 is likely a direct result of pandemic-related restrictions, which reduced travel and commuting significantly.

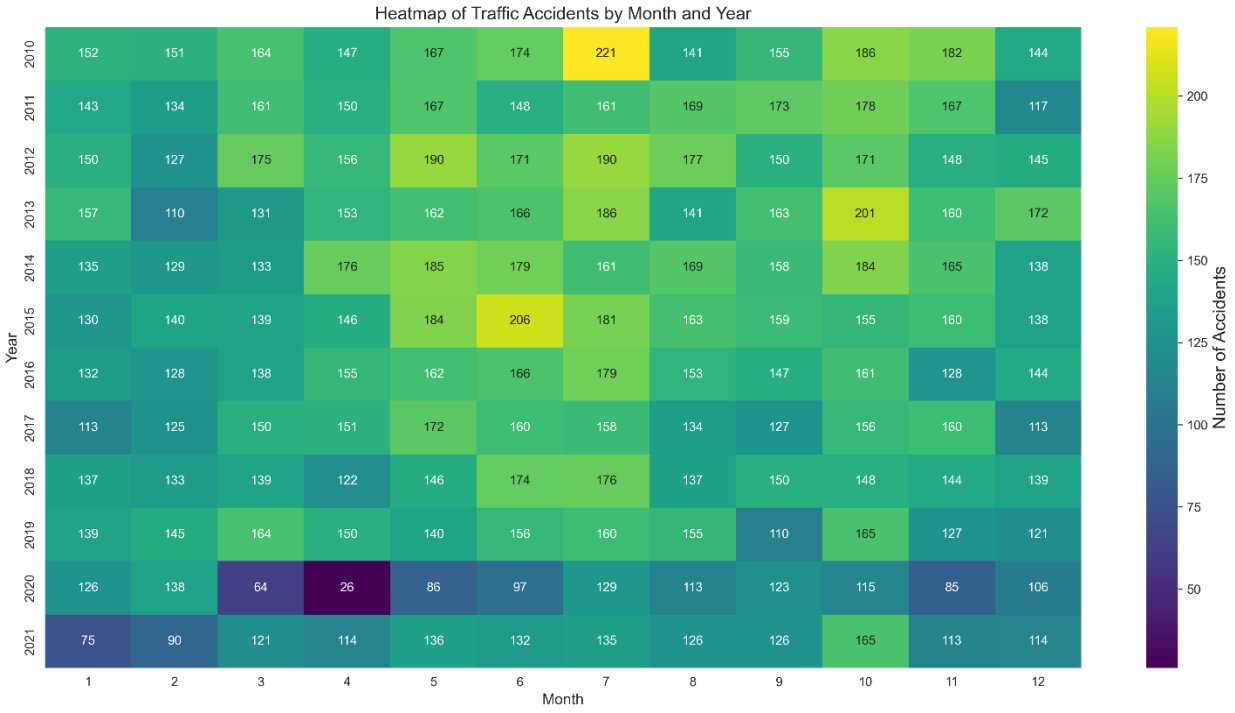


*Figure 12 Seasonal and Pandemic Influences on Road Accidents: A Yearly and Monthly Analysis.*

Figure 13 extends the analysis of seasonal accident patterns, adopting a weekly perspective rather than monthly. This finer granularity reveals more nuanced weekly fluctuations within the broader seasonal trends previously identified.

Comprehensive Correlation Analysis

1. **Overall Correlation Matrix (Figure 14)**: This figure presents a correlation matrix covering all variables in the study, providing a comprehensive overview of how different factors interrelate in the context of road accidents.
2. **Focused Correlation Study (Figure 15)**: Delving deeper into specific correlations, this figure examines the relationship between weather conditions, day of the week, time of day, lighting conditions, road surface conditions, and wind conditions. Notably, certain lighting conditions such as 'De nit', 'sense llum artificial' (without artificial light), and 'de dia, dia fosc' (daytime, dark day) show a strong correlation with accidents. Conversely, 'de dia clear' (clear day) has an inverse correlation with various types of injuries, including fatalities, serious injuries, and light injuries.

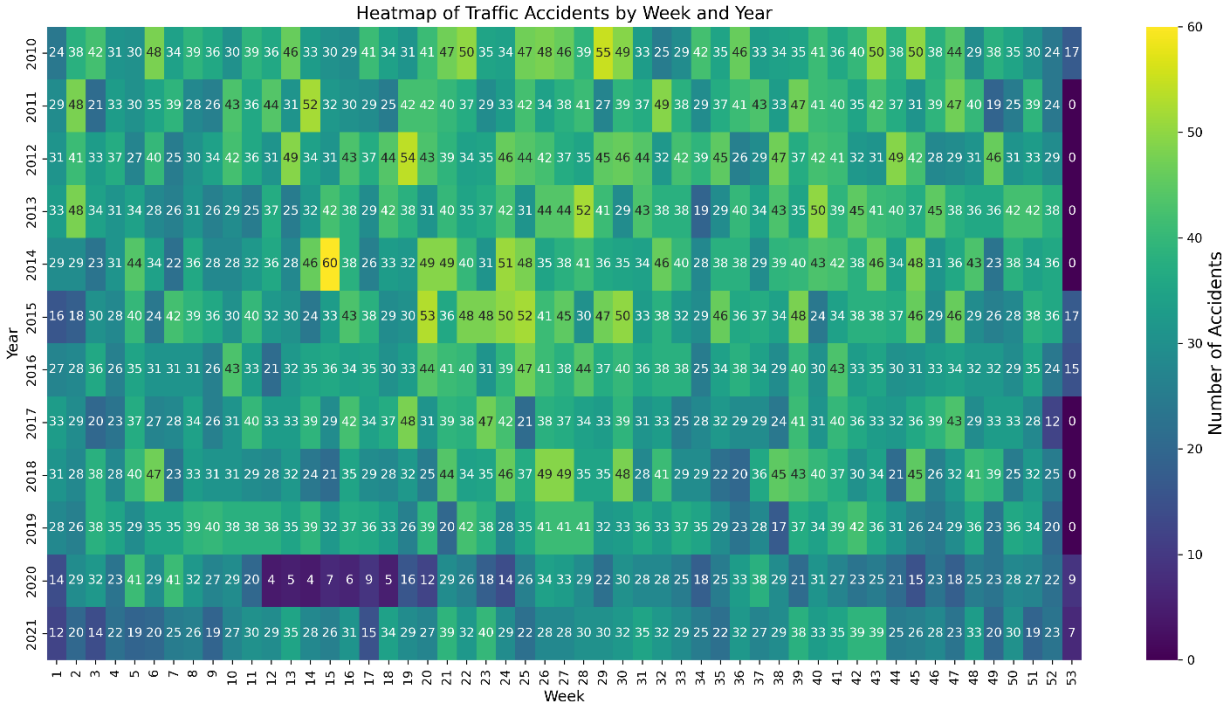


*Figure 13 Weekly Patterns in Road Accidents: Analyzing Seasonal Fluctuations.*

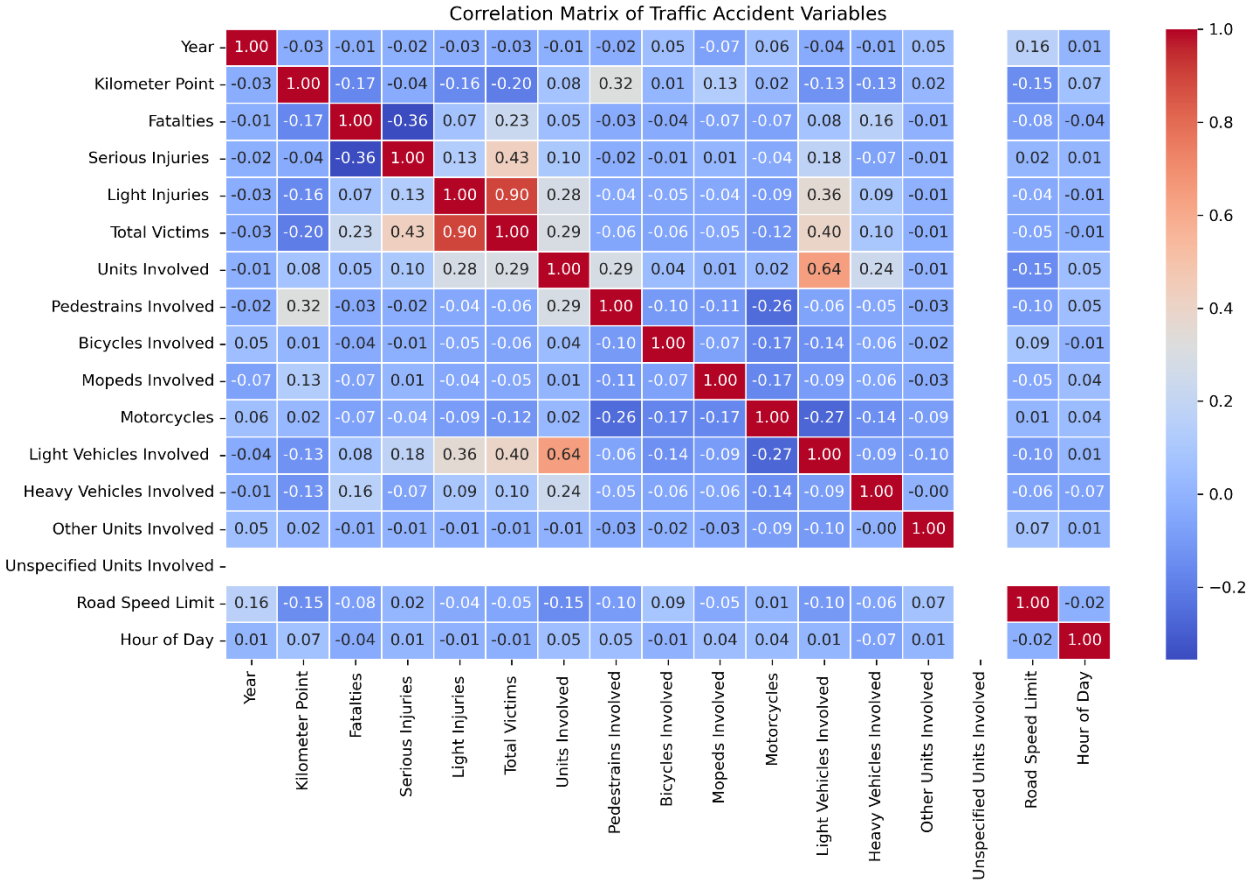


*Figure 14 Comprehensive Correlation Matrix of Road Accident Factors*

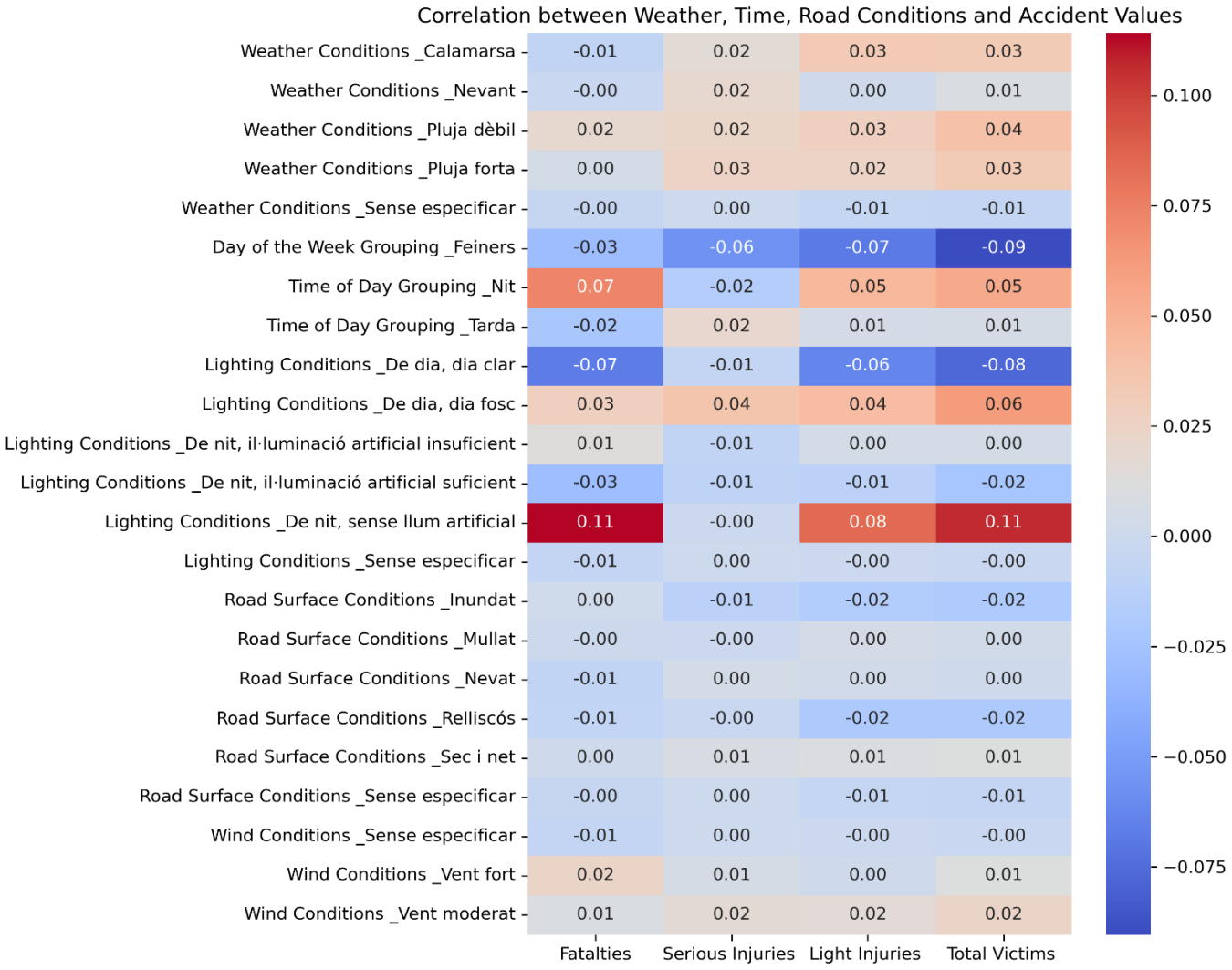


*Figure 15 Detailed Analysis of Correlations: Weather, Lighting, and Injury Severity*

## 3.10 Time-Series Forecasting

Based on past trends, create a model to forecast the number of accidents, fatalities, or serious injuries for the upcoming year.

Clearly describe the forecasting model you have developed. This should include the type of model, its structure, and any specific features or techniques it utilizes. Discuss the factors that influenced your decision, such as the model's accuracy, efficiency, suitability to the data characteristics, or its ability to handle the complexities of the dataset.

1. ARIMA Model

- This is an autoregressive integrated moving average model fit using the pmdarima library's auto_arima function.
- It automatically selects the best p, d, q parameters based on AIC score.
- This captures trends and seasonality in the time series data.

2. Logarithmic Trend Model

- Fits a linear regression model between the log-transformed year and target variables.
- Used to model a long-term logarithmic trend in the data.

- Adjusted to pass through the last known data point in 2019.

The combination of automated ARIMA fitting and a simple logarithmic trend line seems well-suited for this short univariate time series data. The models balance accuracy and efficiency without overfitting on short series. The multiple visualizations allow detailed comparison between different modeling approaches.

For the next phase, time-series regression models incorporating causal factors could improve accuracy. Ensembling or blending models could also help optimize predictive performance. But the current approach provides a robust starting point for forecasting traffic accidents and planning safety interventions.

To evaluate the effectiveness of the forecasting models over the upcoming year, the trends were applied to predict monthly fatalities, serious injuries, and light injuries in upcoming year (Figure 16). The forecasts show a stable encouraging and slightly decreasing trajectory across all three injury categories after an initial increase due to the recovery of situation pre-covid. Though the fall in upcoming year seems marginal, this aligns with the longer-term logarithmic trends projected. It indicates that the ongoing road safety initiatives and interventions are potentially having a sustained positive impact, with accident severity and frequencies on track to either stabilize or reduce further. There are still fluctuations noticeable, however, the general downwards direction is reassuring and should motivate continued vigilance and evidence-based planning towards the road safety goals.

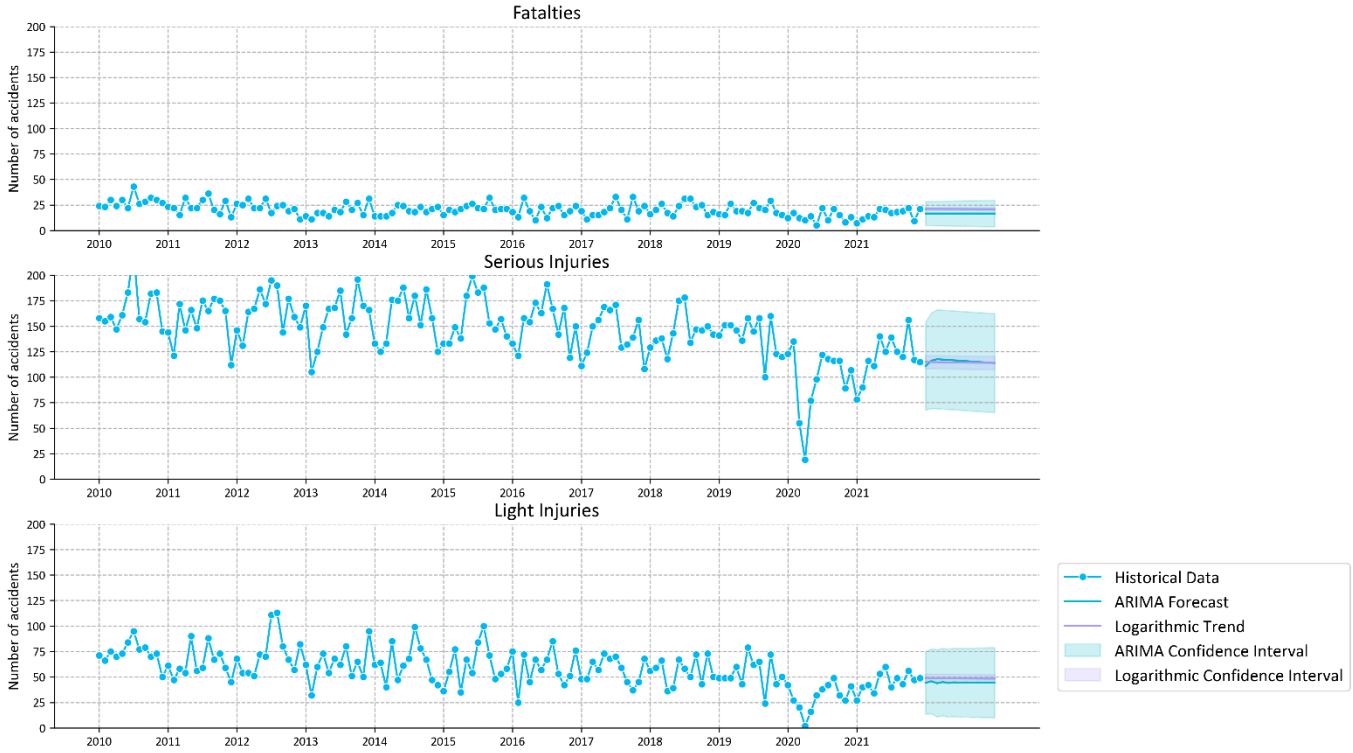


*Figure 16 Forecasted monthly figures for traffic accident fatalities, serious injuries and light injuries for a year in the future based on ARIMA and logarithmic trend models.*

## 4. Conclusion

The comprehensive analysis of traffic accident data in Catalonia over the past decade offers critical insights into improving road safety. The declining trends in traffic accidents, injuries, and fatalities indicate progress through interventions like policy reforms, improved vehicle safety features, and public awareness campaigns. However, the prevalence of severe injuries underscores significant risks that persist, necessitating targeted mitigation strategies.

Key contributing factors identified include high-risk road types (urban and conventional roads with 100 km/h limits), vulnerable vehicle types (motorcycles and heavy vehicles), seasonal and weekend travel peaks, and diminished visibility conditions. These high-risk scenarios underscore the need for customized countermeasures ranging from

enhanced traffic calming methods and dedicated motorcycle/truck lanes on select urban roads to expanded lighting infrastructure and directed driver safety education programs.

The correlations revealed between factors like visibility, road surface conditions, and injury severity further spotlight the impact of environmental and infrastructural conditions on accident outcomes. Investments towards improved lighting, drainage systems, and regular maintenance to ensure optimal road surface conditions could significantly mitigate accident severity.

Overall, a multilayered approach is imperative - combining policy updates, road design improvements, expanded safety infrastructure, customized vehicle and licensing requirements, and continuous public education campaigns based on the evolving landscape of risks. The insights from this report equip decision-makers with the awareness and evidence-base to implement targeted strategies for maximum impact, propelling Catalonia firmly towards the goal of safer travel.