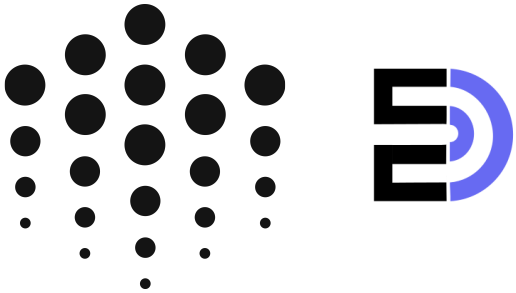
Road to Safety: Traffic Accident Analysis

OCEAN Protocol Data Challenge #1
Desights AI



BY
xMarkitus

January 2024

Contents

1 Introduction

This report delves into a comprehensive analysis of traffic accidents in Catalonia, aiming to shed light on various aspects that may influence their frequency, severity, and patterns over the years. By exploring a dataset encompassing details such as accident severity, weather conditions, road types, and more, our goal is to uncover key insights that can contribute to a better understanding of the dynamics surrounding traffic incidents in the region.
The method we use begins with a thorough examination of the data structure, providing the framework for later analysis. We then focus on specific topics, unravelling the dataset's deep relationships to provide clear, compelling findings. This extensive analysis tries to reduce complex material to simply understood conclusions, allowing for more informed decision-making on safety measures, legislation, and public awareness campaigns.

2 Key Findings

Our exploration into Catalonia's traffic landscape unveils critical insights that resonate with simplicity and clarity. Firstly, we discovered that accidents tend to peak during the morning and evening hours, signaling specific times when heightened vigilance and safety measures may be most effective.

Delving deeper, our analysis spotlights a concerning trend for vulnerable road users—moped and motorcycle riders—who face a higher probability of fatalities in accidents. This emphasizes the urgent need for targeted safety campaigns and tailored interventions to protect these road users.

Furthermore, a promising trend emerges with a downward trajectory in overall accident numbers. This decline may be attributed to the evolving landscape of remote work and the integration of advanced safety features in vehicles. Such positive shifts indicate the potential success of modern approaches to road safety.

In essence, these key findings distill complex data into straightforward insights, providing a glimpse into the nuanced dynamics shaping Catalonia's road safety landscape.

3 In-depth Analysis

The goal in this chapter is to conduct a focused examination of Catalonia's traffic accidents. It is important to analyze the different accidents to spot the main causes that contribute to the occurence of an accident, severity and different trends that may exist. Each question provided acts as a pathway to unravelling the dataset's intricacies and shining light on the root causes of accidents. By taking a non-technical approach, we hope to explain our findings in a clear and entertaining manner, drawing similarities and using analogies to make the complex world of traffic accident investigation more accessible to all.

3.1 General Trends

The first step to address a proper analysis should be to get a general idea of what are we dealing with, in this case traffic data from Catalonia's accidents between 2010 and 2021.

On figure 1 we can see the downside trend that data is showing us. It is important to show that the biggest number of accidents end up with serious injuries, followed by light injuries and fatalities. On the 2020 we can spot the biggest downstep due to Covid quarantine, and then it increased a bit later on 2021, however, due to the latest advance to working remotely there are less accidents on the road. The Fatalities might be more linear than the others, it is true that since 2010 there is a decraese on the number of fatalities per year but is not as appreciative as the other ones. The reason might be on the evolution of safety systems of the cars. However, instead of being involved as a fatality it might be involved as an injury.

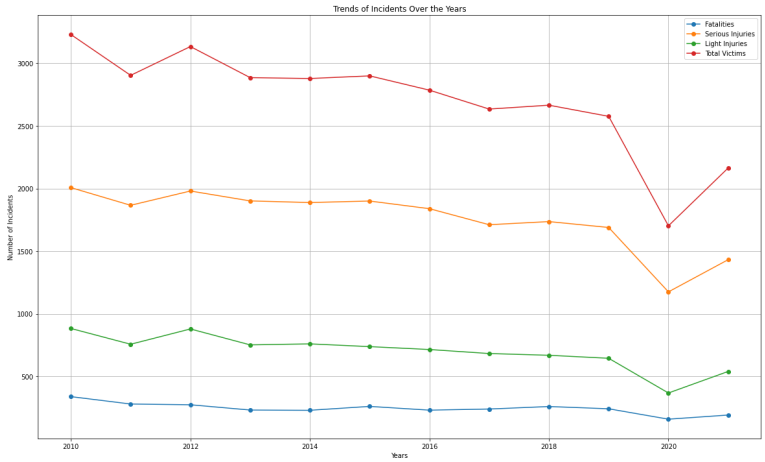


Figure 1: Overall traffic accidents trends by type of injury

3.2 Accident Characteristics

On every accident there is a world around it, which has several characteristics such as time of day, type of road, month, etc. In this case we are analyzing those characteristics that are more common on the most severe accidents, which are considered to be the ones with a death.

After selecting several variables, the most representative characteristics showing of some information were: Type of Road, Time of Day, Day of Week and Month. However, other interesting aspects were Road Ownership where most of the accidents were on autonomical roads, which would indicate that maybe Catalonia's government should put more effort on the maintainance of their own roads and then Type of Accident where the most common case was collision between vehicles, which might indicate lack of regulation on vehicle distance on road.

If we take a look to fig 15 regarding the type of road we can clearly see how most of the accidents are in conventional roads and urban streets. The reason for these kind of accidents might relay on the fact that they are the most used roads since most of the people live in a city or town and there are more distractions there than in a highway. When it comes to conventional roads there might be a lack of control of speed that summed up

with the road state may facilitate accidents.

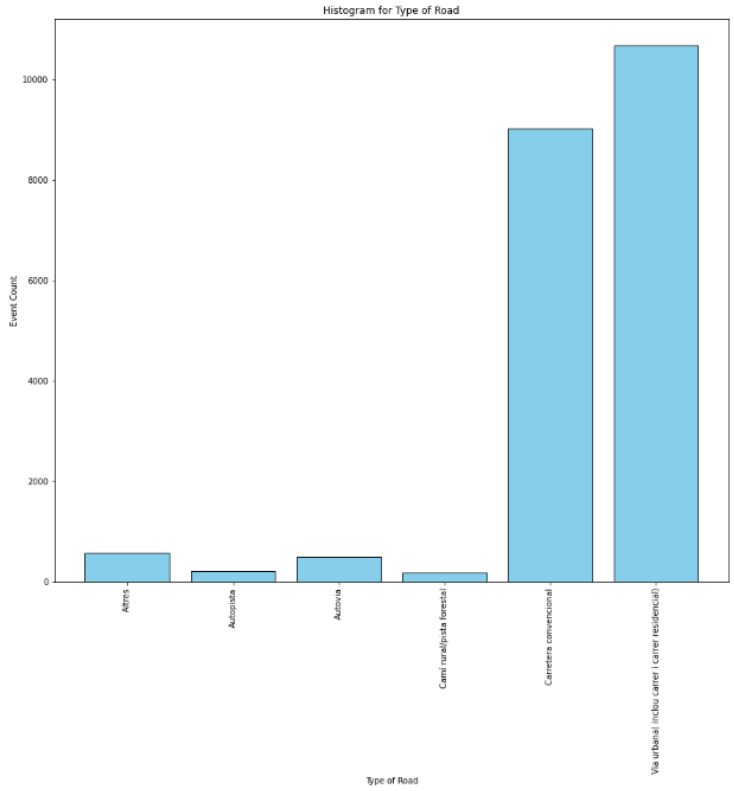


Figure 2: Overall traffic accidents trends by type of injury

On figure 3 we can see that most of the accidents tend to happen on Friday, which adding to figure 4 showing that most accidents occur at evening time may indicate that most accidents occur when people finish their week of work. The main reason behind this fact should be that most people tend to go out on the weekend leaving Friday after work, which increases the level of vehicles on road enhancing the probability of an accident.

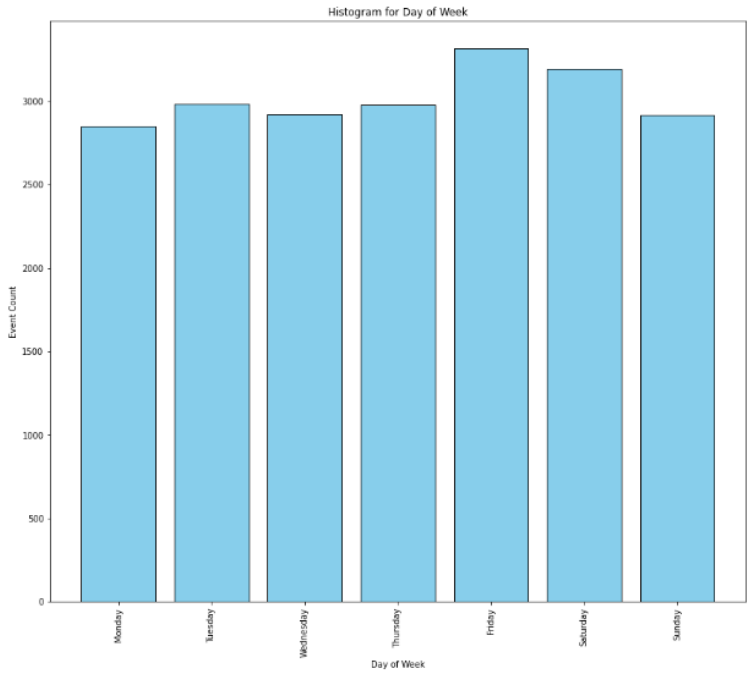


Figure 3: Overall traffic accidents trends by type of injury

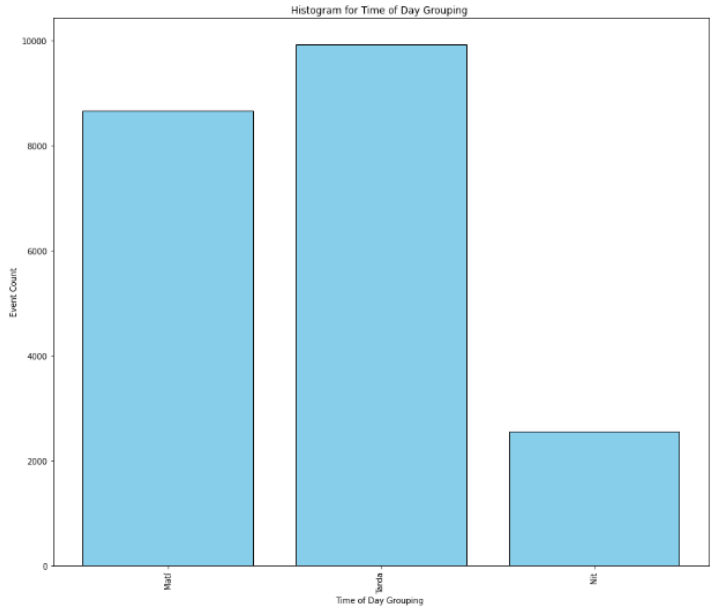


Figure 4: Overall traffic accidents trends by type of injury

Finally, if we take a look at figure 5 we can see that there is a peak on accidents during summer, which is due to the fact that more people is travelling on the road and then there is a peak on October, which may be influenced by the fact that most people is back to work and therefore there are more vehicles travelling daily to the office and leading those to enhancing the probability of a severe accident.

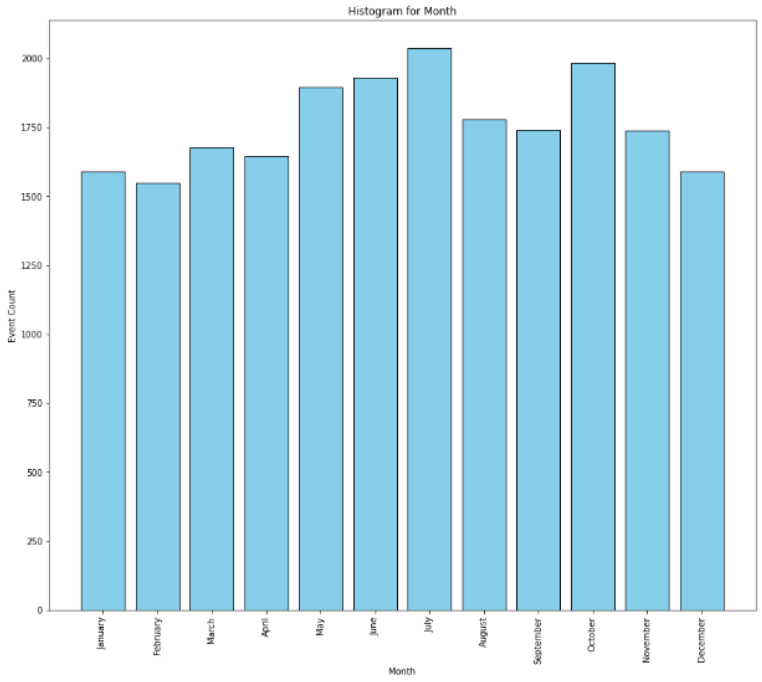


Figure 5: Overall traffic accidents trends by type of injury

3.3 Geographical Insights

Another interesting point is the relation between population and accidents. A geographical analysis has been conducted to determine any kind of relationship between population density and accidents on road.

On figure 6 we can see the relationship between population density and number of accidents. If we take a closer look we will spot that there is not a clear relationship on municipies and counties between higher density leading to more accidents, there are some exceptions like Badalona having less accidents than Sabadell, or Garraf less accidents than Osona. This shows that there are other reasons apart from the population, such as the transport network and the need of taking private vehicles other than public transport. Bigger cities might have easy access to public transport leading to have less vehicles on the street.

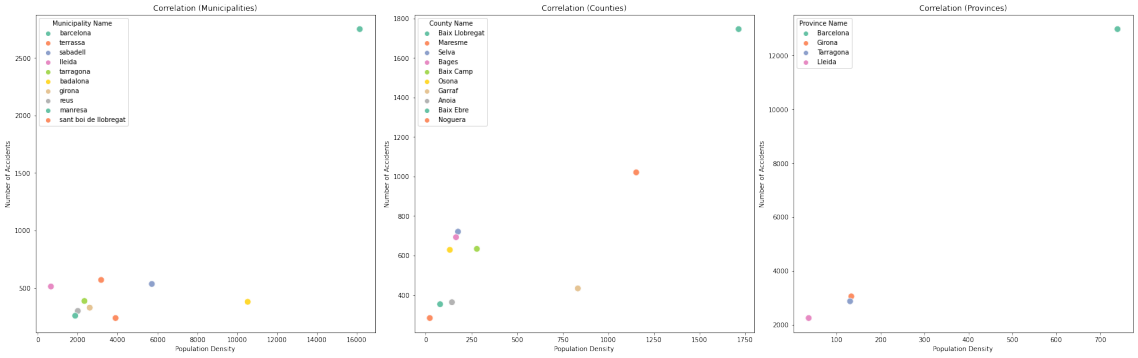


Figure 6: Overall traffic accidents trends by type of injury

If we delve into a comparative analysis of road network characteristics across different provinces in Catalonia, a notable observation emerges: the Province of Barcelona ex-

hibits a significantly higher number of accidents on urban streets compared to any other province. However, this disparity appears to diminish when examining other types of streets. The rationale behind this trend may be attributed to the fact that Barcelona, being the capital and largest city, naturally experiences a substantial portion of its traffic within the urban landscape. In contrast, other provinces may not witness the same level of urban congestion, contributing to a more balanced distribution of accidents across various street types. This insight emphasizes the importance of considering regional characteristics and traffic patterns when evaluating accident data, shedding light on the nuanced dynamics that influence accident rates in Catalonia. If we look at figure 7 we will see the relation.

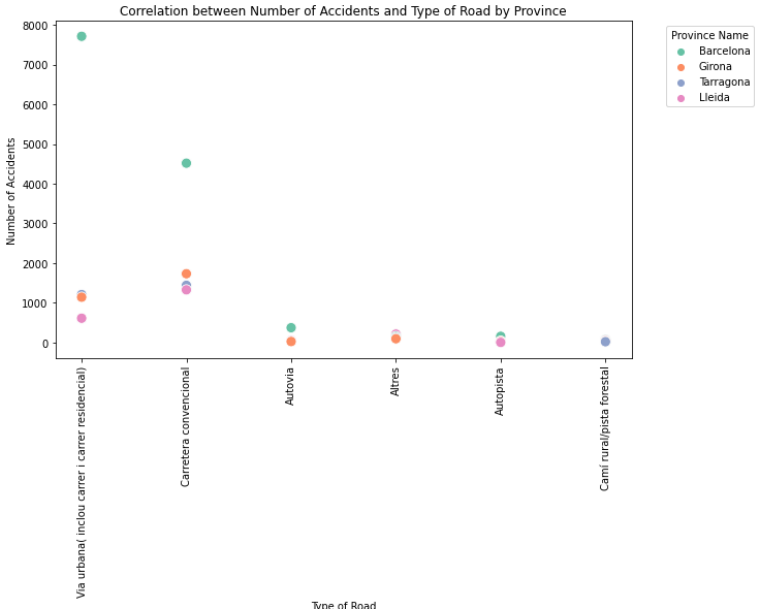


Figure 7: Overall traffic accidents trends by type of injury

3.4 Yearly Trends

Among the time there are changes, and as on figure 8 there is a clear downside trend on the number of accidents. It is true that the most severe accidents have a more flat trend than the less severe ones, which as said previously might indicate that there are less deaths but more injuries.

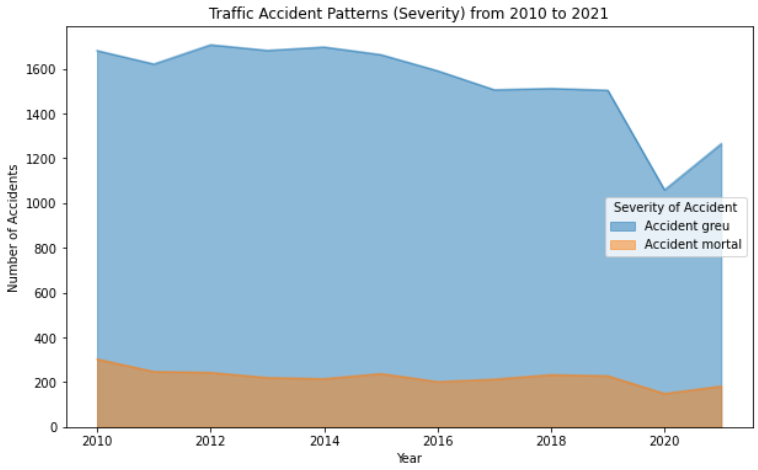


Figure 8: Overall traffic accidents trends by type of injury

3.5 Day and Time Patterns

An important aspect of an accident is the time the accident takes place. The day and time an accident occurs may show off relevant information, and as we spot before on 3 most of the accidents take place on Friday. If we look closer we can see that also weekend days have more accidents than week days, and the reason might be that more people travels during weekends, so there is a higher chance of an accident.

The same reasoning answers at figure 5, where most accidents occur during summer when most people goes out on vacation.

When scrutinizing the time patterns depicted in Figure 4, a discernible trend emerges: the evening hours exhibit a higher frequency of accidents, signifying that the post-working hours period is more prone to incidents compared to the morning or night. A plausible explanation for this observation lies in the volume of cars on the road. It appears that during the evening, there is a heightened presence of vehicles, potentially contributing to an increased likelihood of accidents. To substantiate this hypothesis, a glance at Figure 9 unmistakably reveals a surge in the number of accidents after 15:00h. This reinforcement of accident occurrence in the late afternoon aligns with the notion that accidents tend to rise after typical working hours, underscoring the correlation between traffic density and accident frequency.

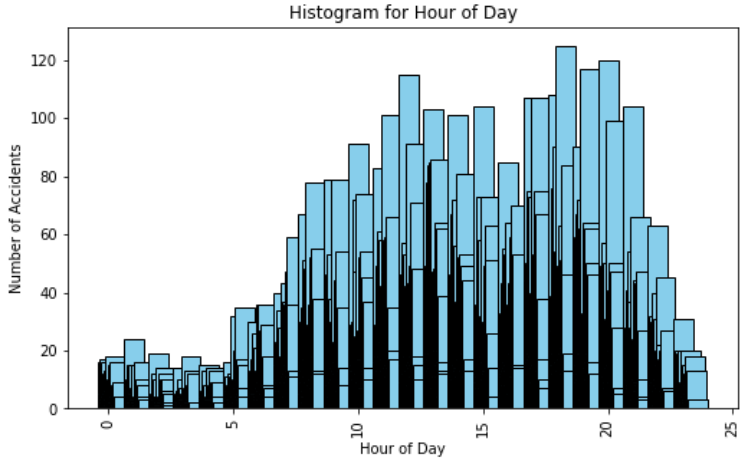


Figure 9: Overall traffic accidents trends by type of injury

3.6 Environmental Impact

The frequency of accidents is significantly influenced by weather, and in Catalonia, where drivers may not be used to extreme weather conditions like heavy rain or snowfall, this factor deserves close examination. The majority of automobiles are not built to manage extreme weather, hence adverse weather conditions may contribute to accidents. The lack of knowledge with negotiating difficult weather conditions may increase the risk on Catalonia's roadways. By investigating the impact of weather on accident rates, we hope to uncover trends and insights that might inform safety measures, infrastructure upgrades, and driver awareness programs, ultimately leading to a better environment for Catalan road users.

In figure 10 we can see that most of the accidents have happened on good weather conditions. This might be biased towards the fact that there are more vehicles on road during good weather days than in other conditions. However, we can see that rain and wind have a little impact on the number of accidents. When looking at lightning conditions we will see that most of the accidents occur during day light or at night with sufficient lightning conditions, although a great number of accidents have happened with insufficient light conditions.

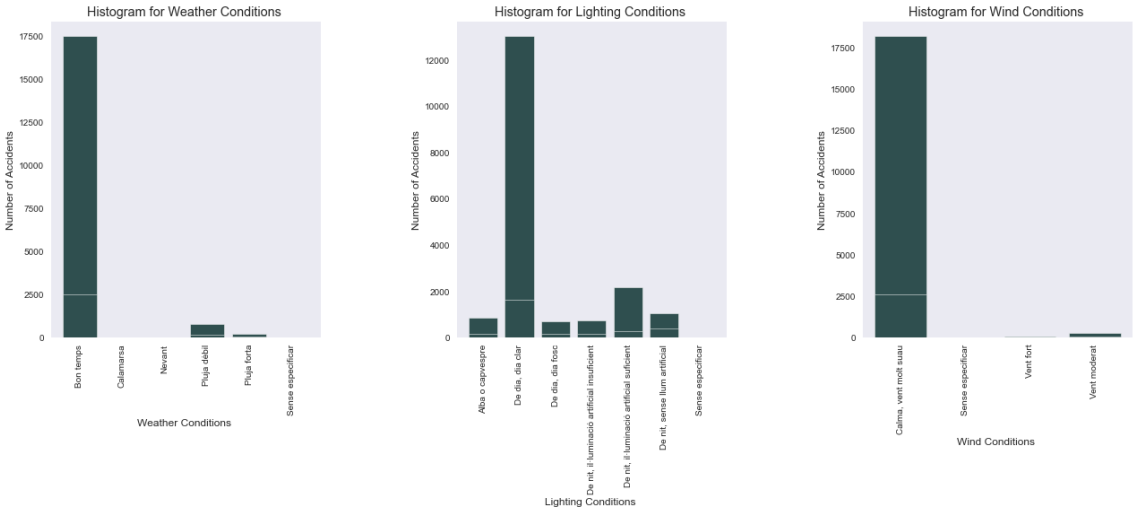


Figure 10: Overall traffic accidents trends by type of injury

In figure 11 and figure 12 we can see that those times where the accidents where influenced by either low visibility or fog there was a higher probability of a fatality accident, so we can see that visibility is a crucial aspect related to the severity of an accident.

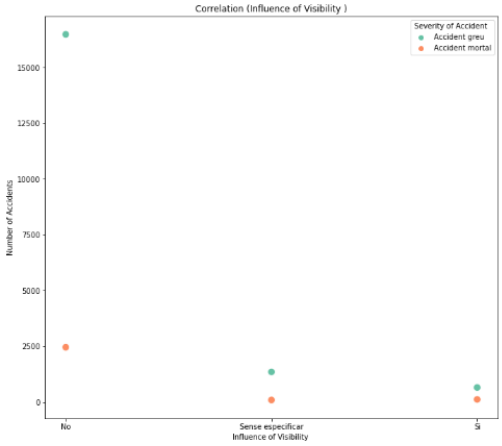


Figure 11: Visibility correlation with accident severity

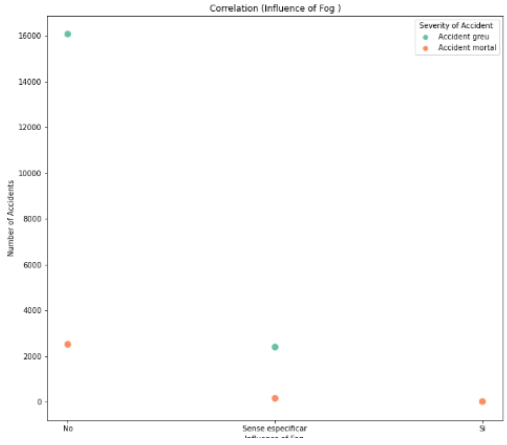


Figure 12: Fog correlation with accident severity

In figure 13 we can see that also surface has some influence on accident severity, meaning that those cases where the accident was influenced by the road conditions there was a higher chance of death. The same applies to figure 14 where those cases where weather was windy the chance of having a fatality was higher.

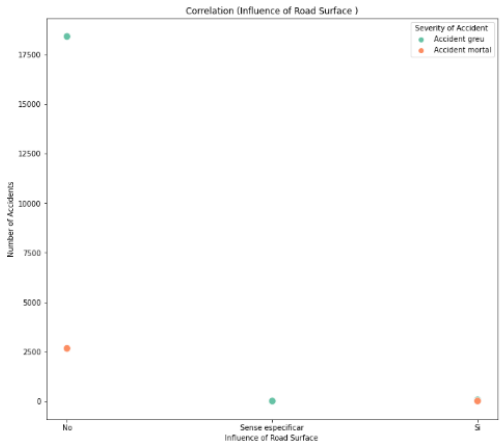


Figure 13: Surface correlation with accident severity

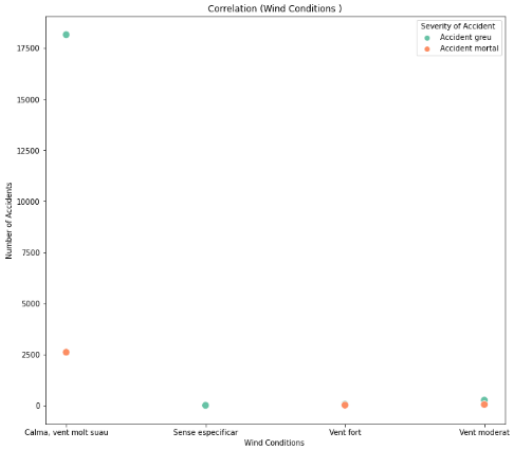


Figure 14: Wind correlation with accident severity

3.7 Road and Traffic Features

The relationship between speed limits and road safety is a critical aspect of our analysis, as adjustments to speed limits often stem from a high number of accidents on specific roads. By scrutinizing road features, we aim to discern any distinctive impact or characteristics that may contribute to accident rates. This examination goes beyond mere speed limits, considering the broader context of road infrastructure, conditions, and configurations. Unraveling these road-specific nuances can provide valuable insights into potential areas for improvement, guiding decisions on speed limit adjustments, road design enhancements, and targeted safety interventions. Through this exploration, we seek to contribute to a more comprehensive understanding of the interplay between road features and accident occurrences in Catalonia.

Data contained some registers as 999.0Km/h, however, there is no such limit in Catalonia and we assumed to dismiss those cases. In figure 15 we can see that most of the accidents are around roads with a speed limit of 100km/h, this are mainly conventional roads which are usually in worse conditions than highways. In the second chart we can see that traffic seems to be an influential factor in a negative way, when there is no influence from traffic it could lead that the driver was speeding and therefore an accident happened. Finally, most of the accidents occur on conventional roads and urban streets due to the fact that are the most common road types in Catalonia.

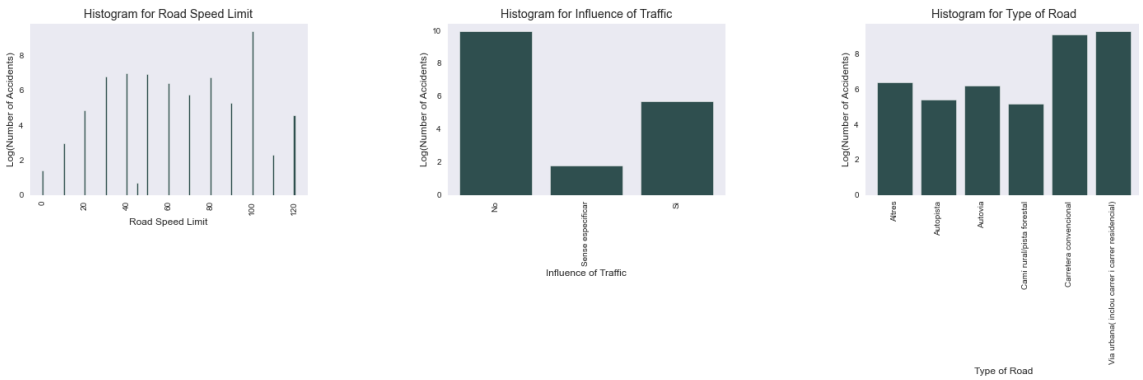


Figure 15: Overall traffic accidents trends by type of injury

3.8 Vehicle Types and Accident Severity

Depending on the vehicle type there is a higher chance of suffering a fatality in an accident. In figure 16 we can see that when there are several mopeds, motorcycles or heavy units involved there is a higher probability of suffering a severe accident. Light vehicles seems to be the safer option, however, when there are more than 5 vehicles involved it seems that the accident is going to be very severe. Therefore we could say there is correlation with the type of vehicle and severity of the accident.

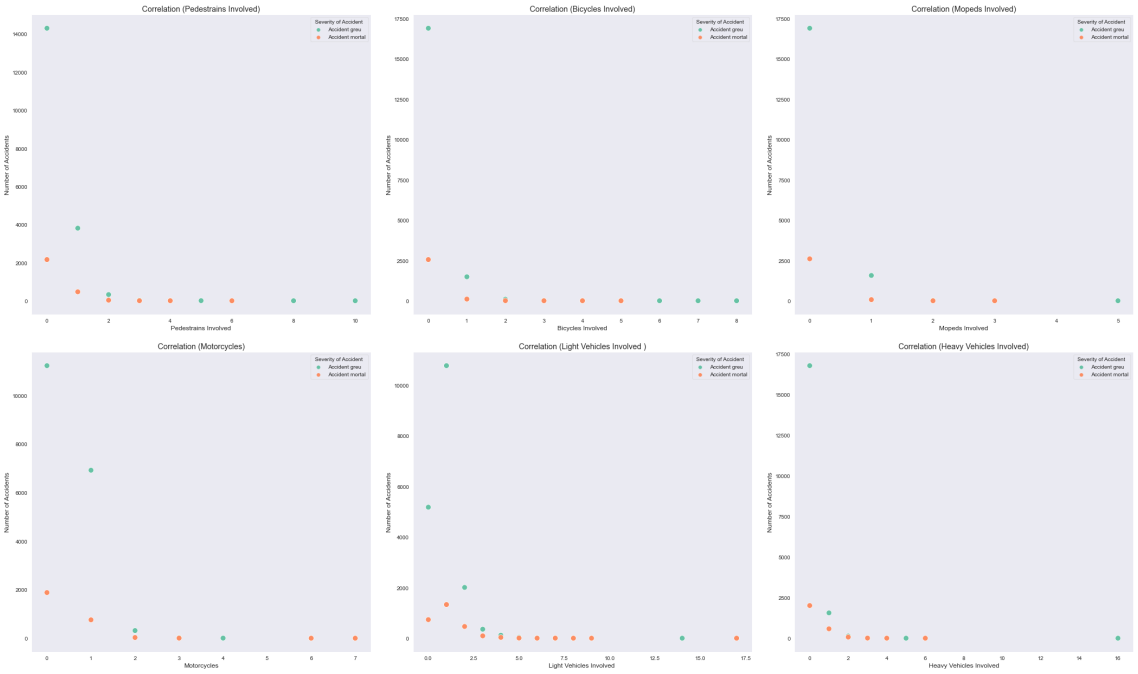


Figure 16: Correlation between units involved and accident severity

3.9 Temporal Clustering

When we try to spot patterns a good technique is to apply clustering to try and spot different relationships between data. In this chapter K-Means algorithm has been applied to detect if there is any temporal cluster which is relevant to provide some insightful information. Figure 17 shows up 3 different clusters. Cluster 1 is the biggest one, and then cluster 0 and 2 are quite similar on size. If we look deeper the cluster 1 refers to the accidents that happened during Day or Evening times, which means that most of the accidents happened during those group of times. The other clusters are divided by years, cluster 0 taking the most recent events and cluster 2 the oldest ones. However, they are quite similar on number of accidents and therefore don't reveal a lot of information.

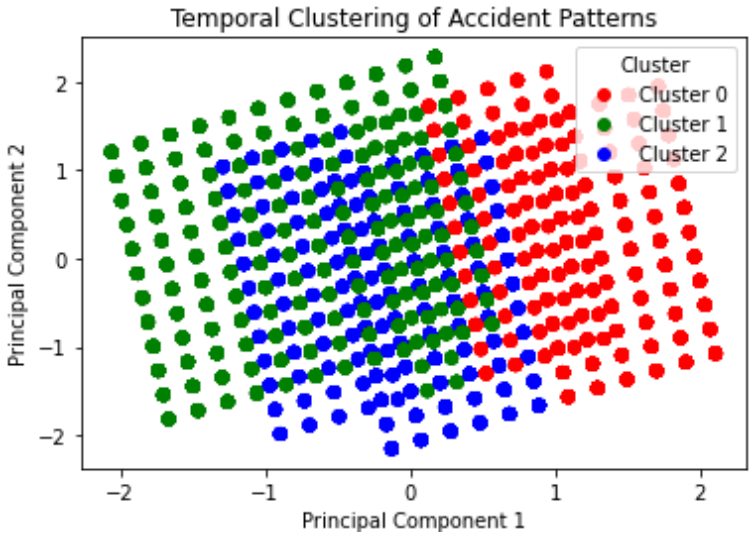


Figure 17: Temporal Clusters by Month, Year and Time of Day

3.10 Time-Series Forecasting

The forecasting model developed for predicting 'Total Victims' in traffic accidents is based on the Prophet model. Prophet is a time-series forecasting model developed by Facebook that is designed to handle datasets with strong seasonality and multiple seasonality components.

Prophet is an additive time-series forecasting model, which decomposes the historical data into several components such as trend, seasonality, and holidays. It is particularly well-suited for datasets that exhibit yearly, monthly, or weekly patterns, which is ideal for our case.

There are several factors that have influenced the decision of the model choice:

1. Seasonality in Traffic Accidents: The dataset likely exhibits strong seasonality due to factors such as weather conditions, holidays, or other external influences. Prophet is well-suited for capturing such seasonality.

2. Ease of Use and Interpretability: Prophet is known for its simplicity and is a very explainable model due to its plots.

3. Additional Regressors: The ability to include additional regressors to the model is a valuable asset.

The model has been tested with 2021 values, obtaining a mean squared-error of 0.0005 which is really low. However, other techniques such Cross-Validation and applying shuffle to data could be applied to improve the model or even used to fine tune the model to enhance its performance to make more accurate predictions.

Figure 18 shows the train and test plus the predictions of the model, which follow the downside trend. This may be due to the fact that nowadays the remote working is becoming more common and then there is less need to take the private vehicle. Therefore, if there are less cars on road there will be less accidents.

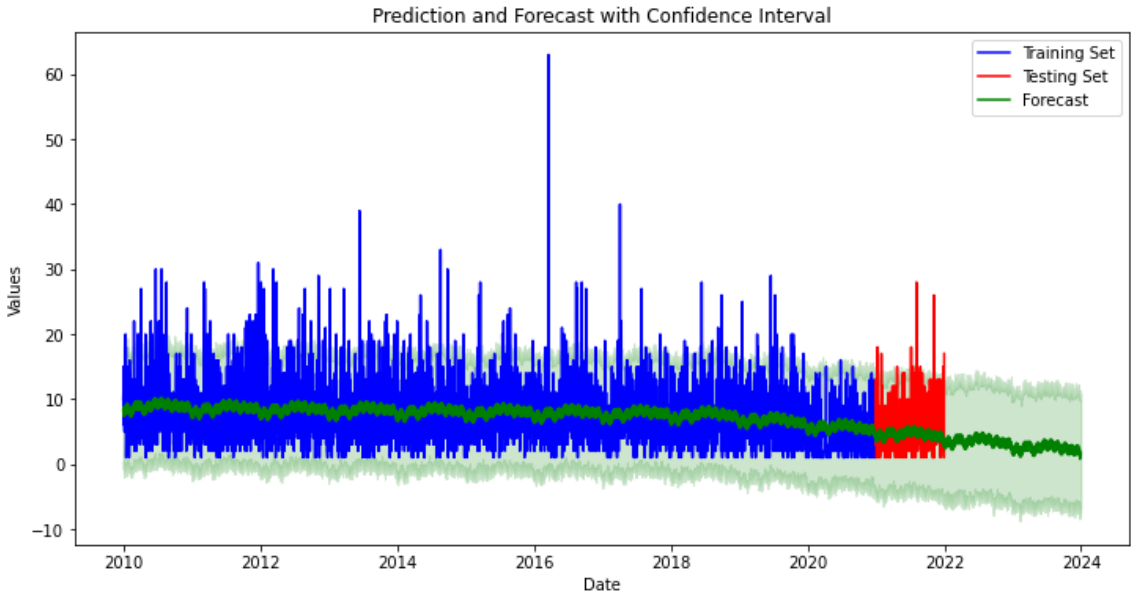


Figure 18: Correlation between units involved and accident severity

If we take a closer look to the model in figure 19 we can understand a little bit more how accidents are distributed and which trends are they following. There is a clear downside trend influenced by the fact that 2020 was a special year and there were less accidents due to Covid. However, if we look to the weekly and monthly chart we can sport a clear trend of having more accidents during the middle week to the weekend, and the same applies to the months, the number of accidents starts raising when summer begins and it stays higher until October. Therefore, we can assume that most people has a higher use of their vehicles during those times.

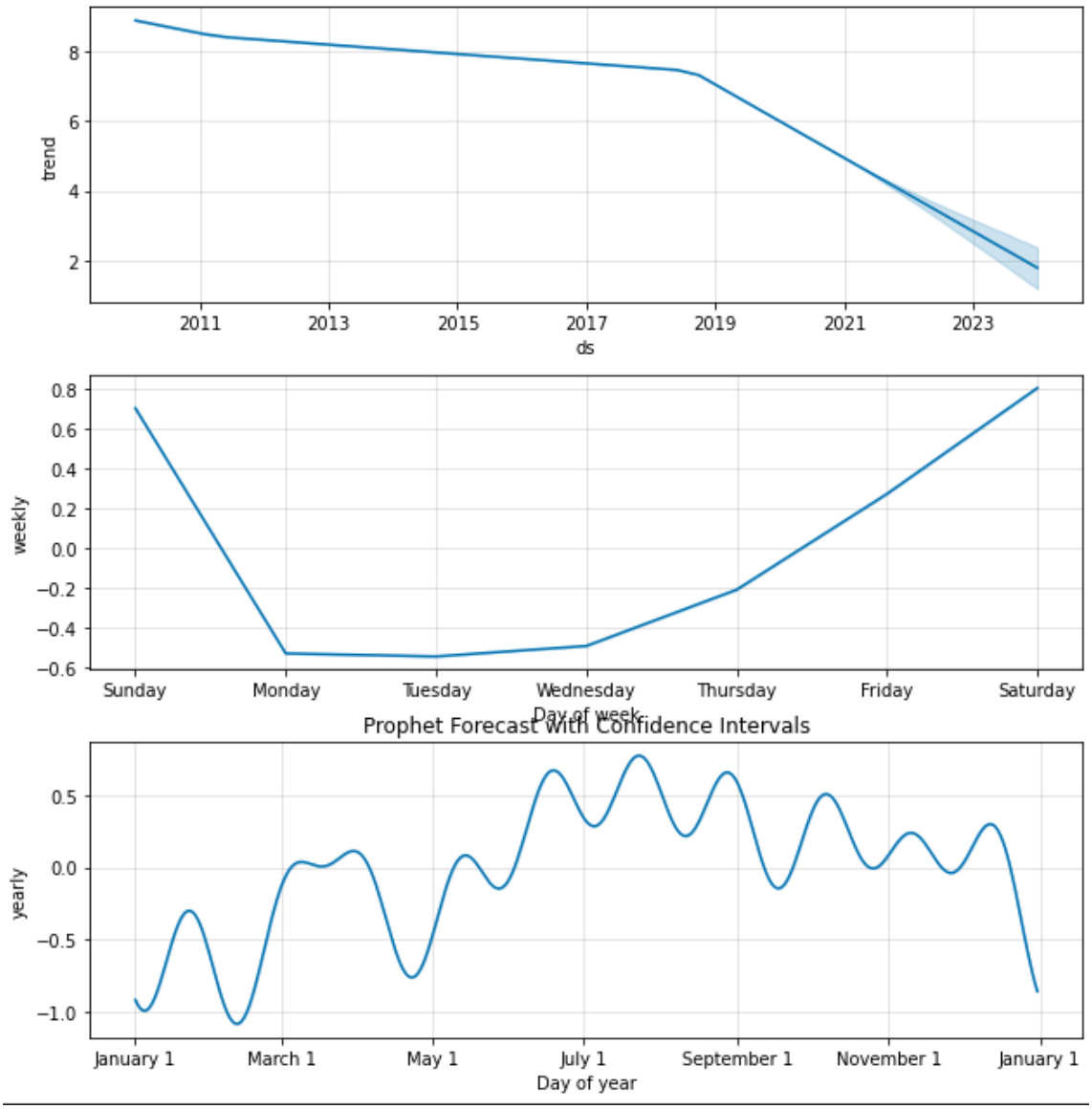


Figure 19: Correlation between units involved and accident severity

4 Conclusions

In drawing conclusions from our comprehensive analysis of traffic accidents in Catalonia, several key findings emerge, shedding light on crucial aspects that influence road safety in the region.

Firstly, the hour of the day emerges as a significant influencing factor, with a notable increase in accidents during the morning and evening hours. This pattern underscores the importance of considering time-specific variables when implementing safety measures and traffic management strategies.

Secondly, the involvement of different types of units, particularly vulnerable vehicles like mopeds and motorcycles, stands out as a significant factor affecting accident outcomes. The higher probability of fatalities associated with these units highlights the need for targeted safety initiatives and awareness campaigns tailored to these specific road users.

Moreover, our analysis reveals a downward trend in the overall number of accidents. This decline may be attributed to various factors, including the rise of remote working arrangements and advancements in vehicle safety technologies. The shift towards remote work could potentially contribute to reduced traffic density, while improved safety features in vehicles might be playing a role in mitigating the severity of accidents.

In essence, these findings underscore the multifaceted nature of traffic accidents, influenced by temporal patterns, vehicle types, and broader societal changes. Recognizing these dynamics is pivotal for devising effective strategies aimed at enhancing road safety, reducing accident rates, and ensuring the well-being of all road users in Catalonia.

4.1 Recommendations and Next Steps

Building on the insights gleaned from our analysis of traffic accidents in Catalonia, several recommendations and next steps can be proposed to enhance road safety and mitigate accident risks.

1. Time-Specific Interventions: Implement time-specific interventions during the morning and evening hours when accidents are more prevalent. Increased law enforcement presence, enhanced traffic management strategies, and public awareness initiatives during these periods can contribute to accident reduction.

2. Road Infrastructure Assessments: Conduct thorough assessments of road infrastructure, especially in areas where accidents are concentrated. Consider adjustments to speed limits, road design enhancements, and the installation of safety features to address specific challenges associated with accident-prone locations.

3. Technological Solutions: Embrace and promote the adoption of advanced safety technologies in vehicles. Encourage the use of vehicles equipped with modern safety features, and explore partnerships with the automotive industry to incentivize the integration of innovative safety technologies.

4. Continous Monitoring and Evaluation: Establish a system for continuous monitoring and evaluation of road safety initiatives. Regularly assess the effectiveness of implemented interventions, refine strategies based on ongoing data analysis, and adapt to emerging trends in traffic behavior.

By systematically addressing these recommendations, Catalonia can take proactive steps toward creating a safer and more resilient road environment, ultimately reducing the frequency and severity of traffic accidents.