

Overview

This dataset contains information about attributes involved in causing accidents.

The dataset contains no missing values.

We removed 31 duplicate rows from the dataset.

Questions to Answer

- Top 5 accidents with most injuries and fatalities
- Cars involved in crashes vs. total injuries
- Month with most crashes
- Weather and road conditions causing most accidents
- Most common traffic devices present at accident sites
- Year with highest crash rate
- Proportion of intersection-related accidents
- Primary causes of accidents
- Daytime vs. night-time accidents
- Crash frequency in optimal vs. worst conditions

Data Overview

- Dataset contains 209,306 rows (samples) and 24 columns (features)
- No missing values present
- 31 duplicate rows were removed
- Statistical information of numerical columns follows below

Top 5 Accidents with Most Injuries and Fatalities

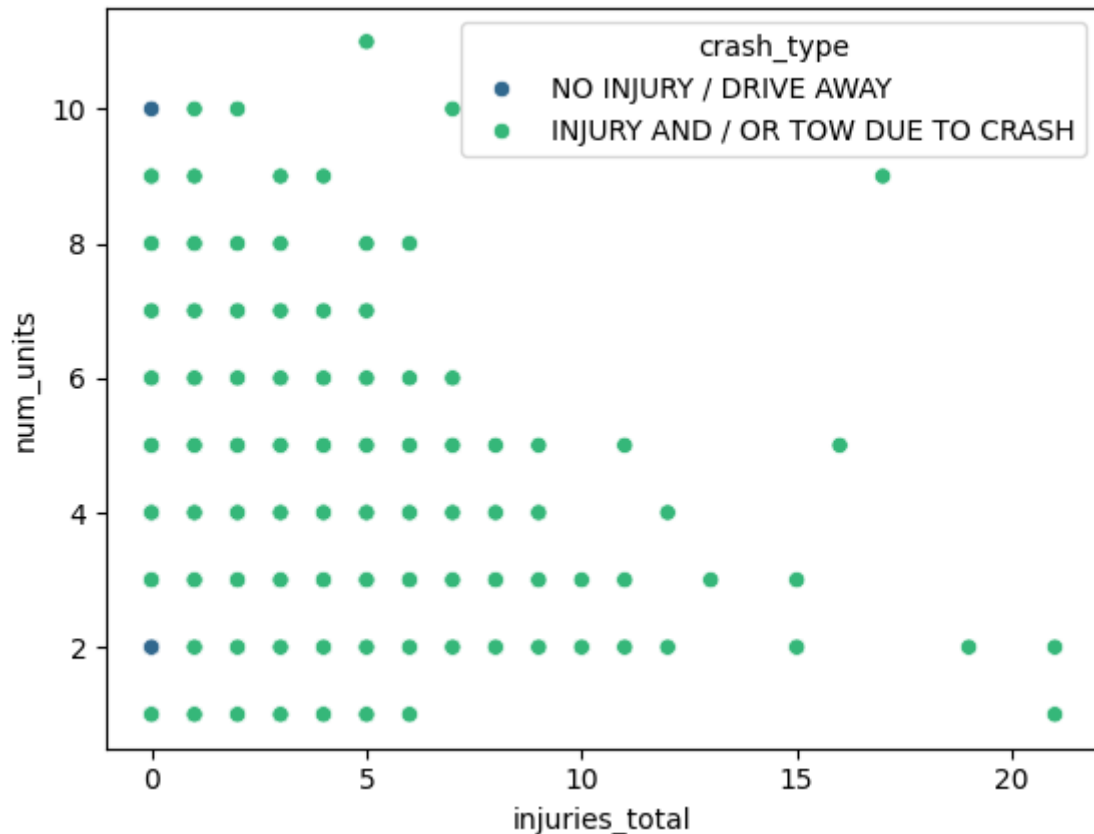
The highest number of injuries recorded in a single accident is 21, with the highest number of fatalities being 3. Among the five most severe injuries, two were incapacitating, two were non-incapacitating, and one was fatal.

	most_severe_injury	injuries_total
91225	NONINCAPACITATING INJURY	21.0
50879	INCAPACITATING INJURY	21.0
58313	INCAPACITATING INJURY	19.0
10830	FATAL	17.0
192581	NONINCAPACITATING INJURY	16.0

	most_severe_injury	injuries_fatal
54213	FATAL	3.0
160678	FATAL	3.0
189366	FATAL	3.0
194369	FATAL	3.0
79014	FATAL	2.0

Cars Involved in Crashes vs. Total Injuries

Accidents involving 2–6 vehicles resulted in the most injuries, with some notable outliers—such as one two-car crash resulting in 20 or more injuries.



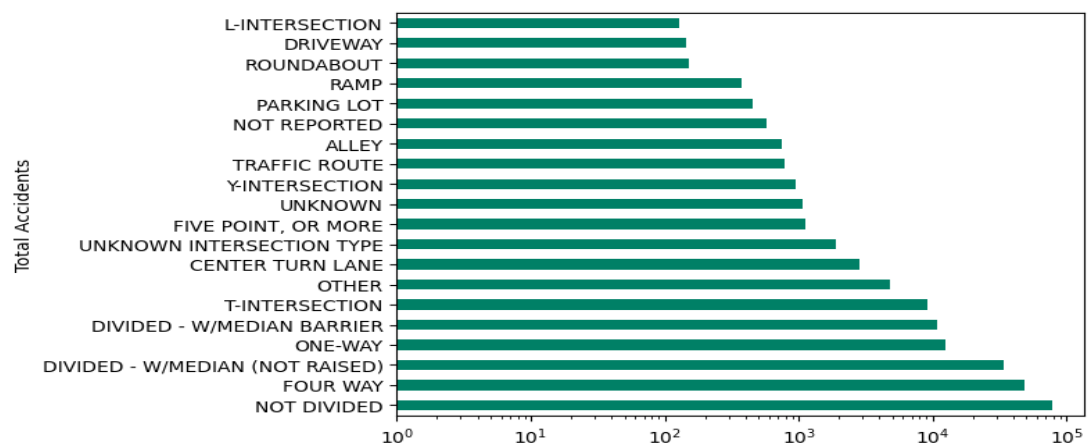
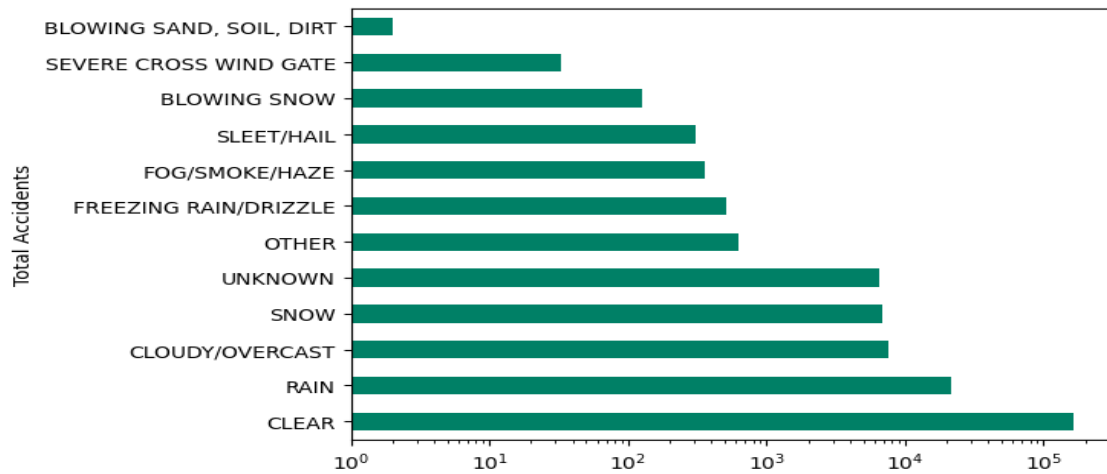
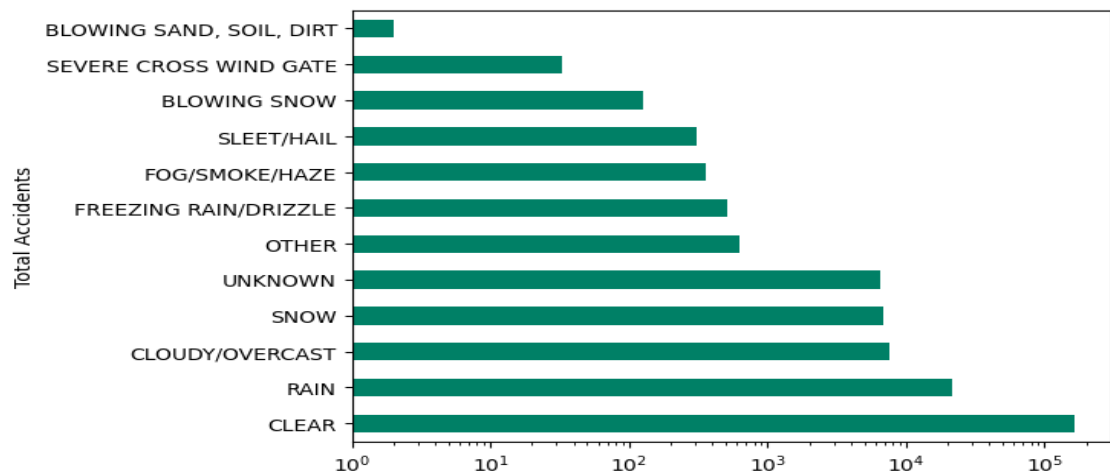
Weather and Road Conditions Causing Most Accidents

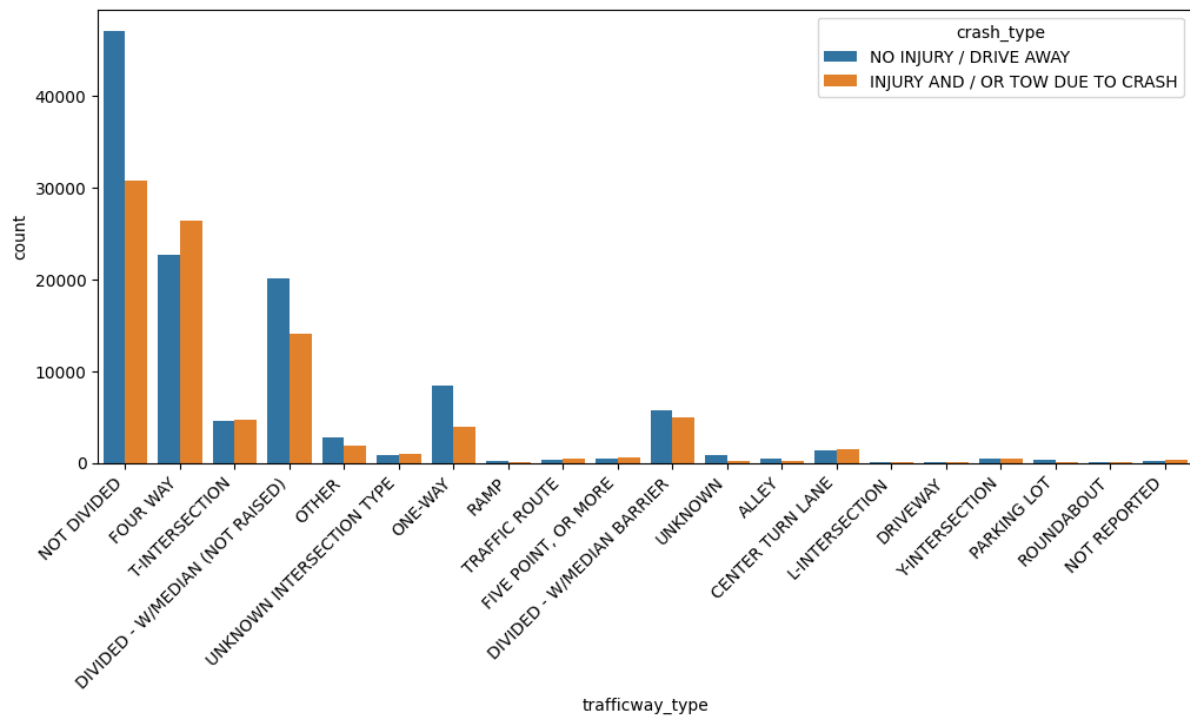
While one might expect more accidents during adverse weather conditions like rain, snow, or hailstorms, most accidents actually occur in clear weather, followed by rain and overcast/cloudy conditions.

Most accidents occur on undivided roads, followed by four-way intersections and roads with unraised medians.

On undivided roads, there are more "no injury/drive away" cases than "injury and/or tow" cases. This might be due to lower-speed collisions during overtaking manoeuvres.

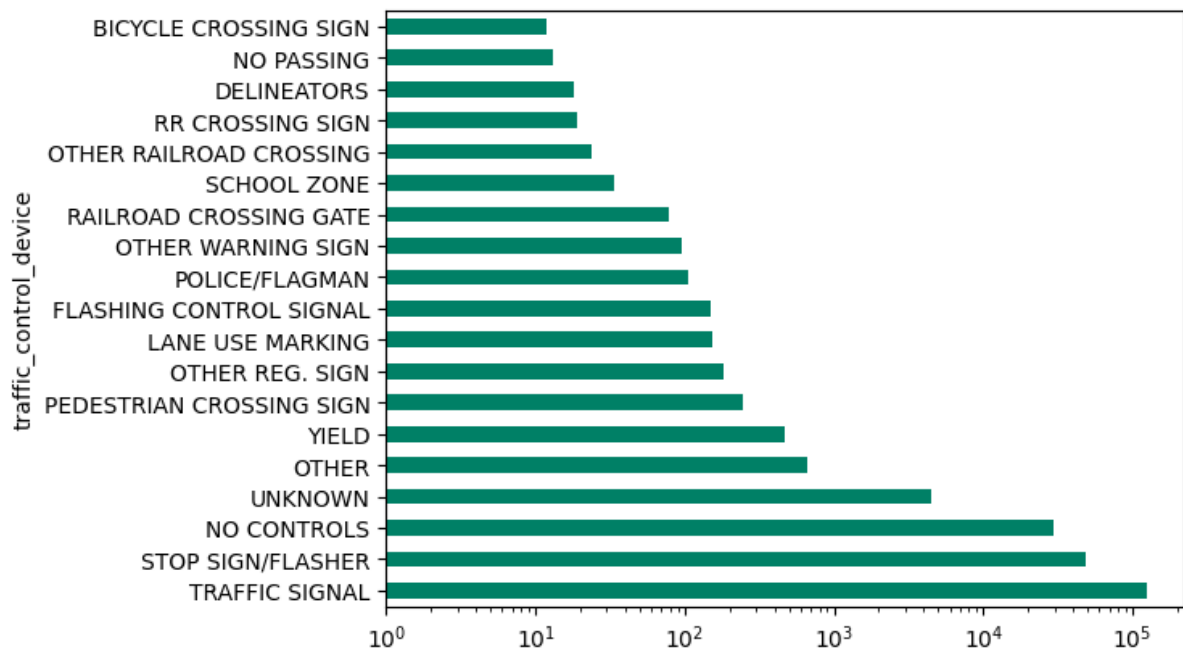
At four-way intersections, there are more "no injury/drive away" cases than "injury and/or tow" cases. When traffic lights turn green, multiple vehicles often accelerate simultaneously, making incorrect crossings particularly dangerous.



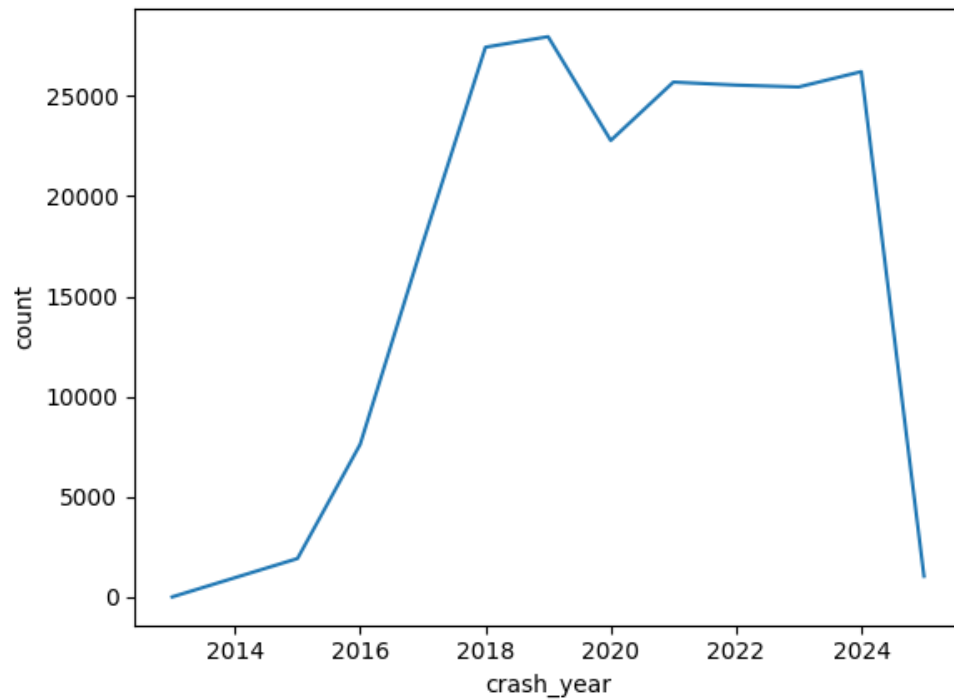


Traffic Devices Present at Accident Sites

Most accidents occur at traffic signals, followed by stop signs and locations without traffic control devices.

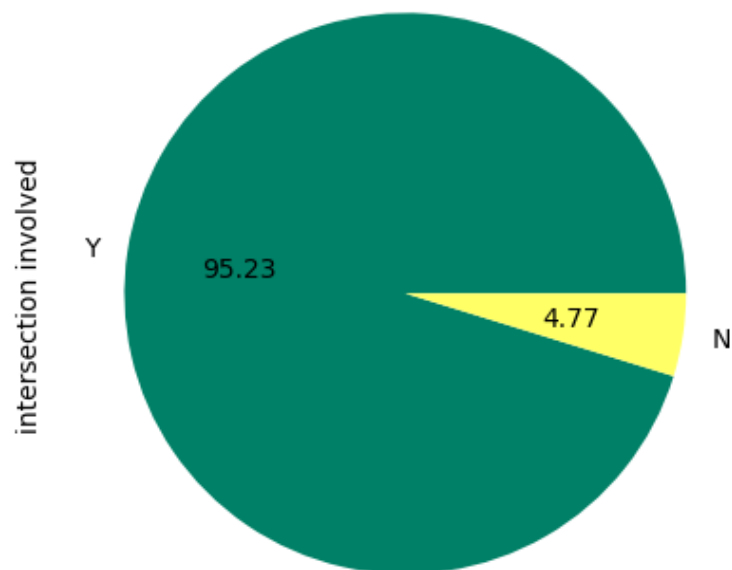


Year with Most Crashes



Intersection-Related Accidents

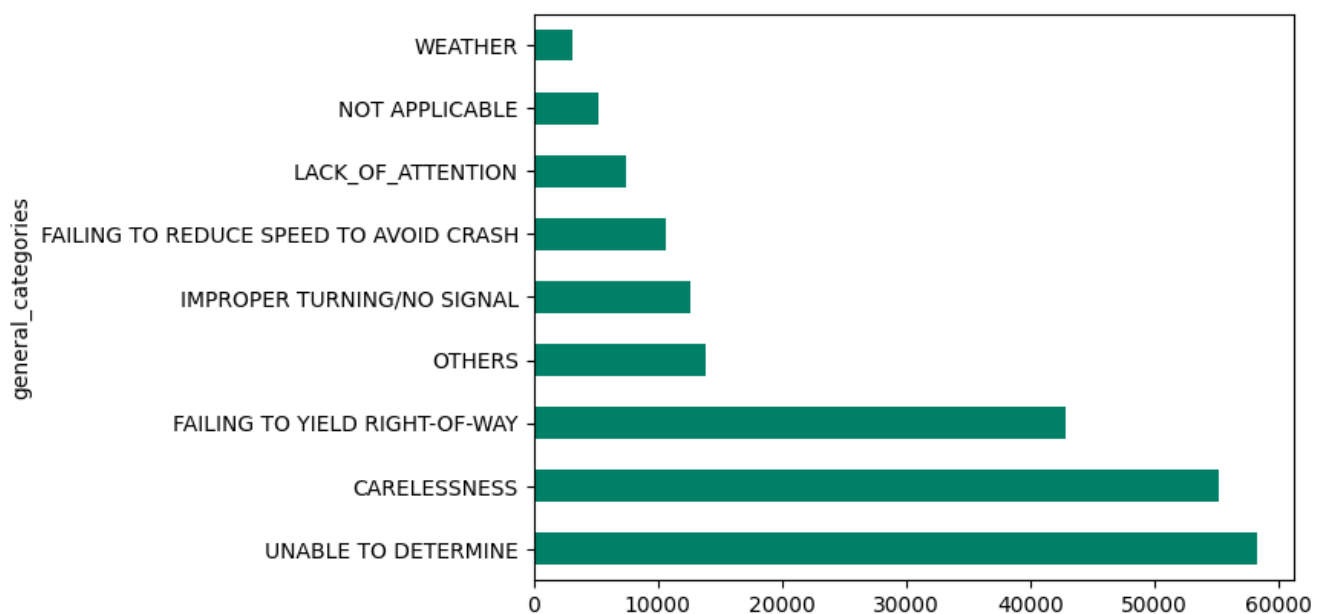
Approximately 95% of accidents are intersection-related.



Primary Causes of Accidents

The "primary contributory cause" feature contained over 40 categories, which we consolidated into broader categories: CARELESSNESS and LACK_OF_ATTENTION. Categories with fewer than 2000 occurrences were grouped as OTHERS due to their specificity and rarity.

In most cases, the cause was undetermined, possibly due to data collection issues. Carelessness is the second most common cause, including behaviours like disregarding traffic signals, internal and external distractions, improper overtaking, and incorrect lane usage.

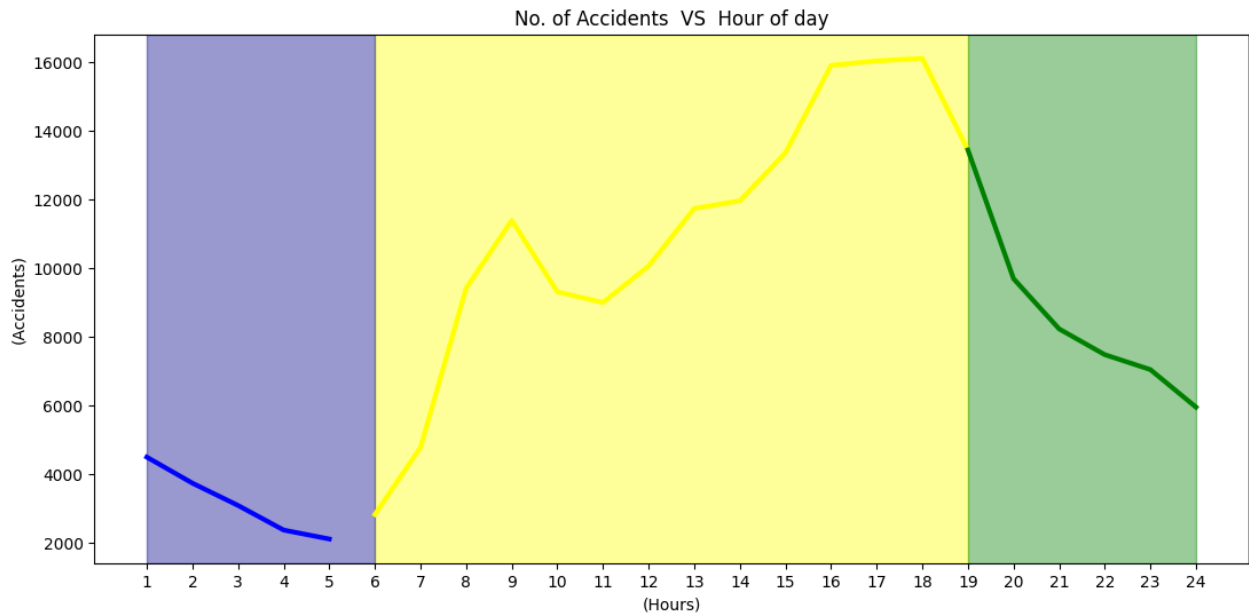


Daytime vs. Night-time Accidents

Accidents occur more frequently during daylight hours than at night.

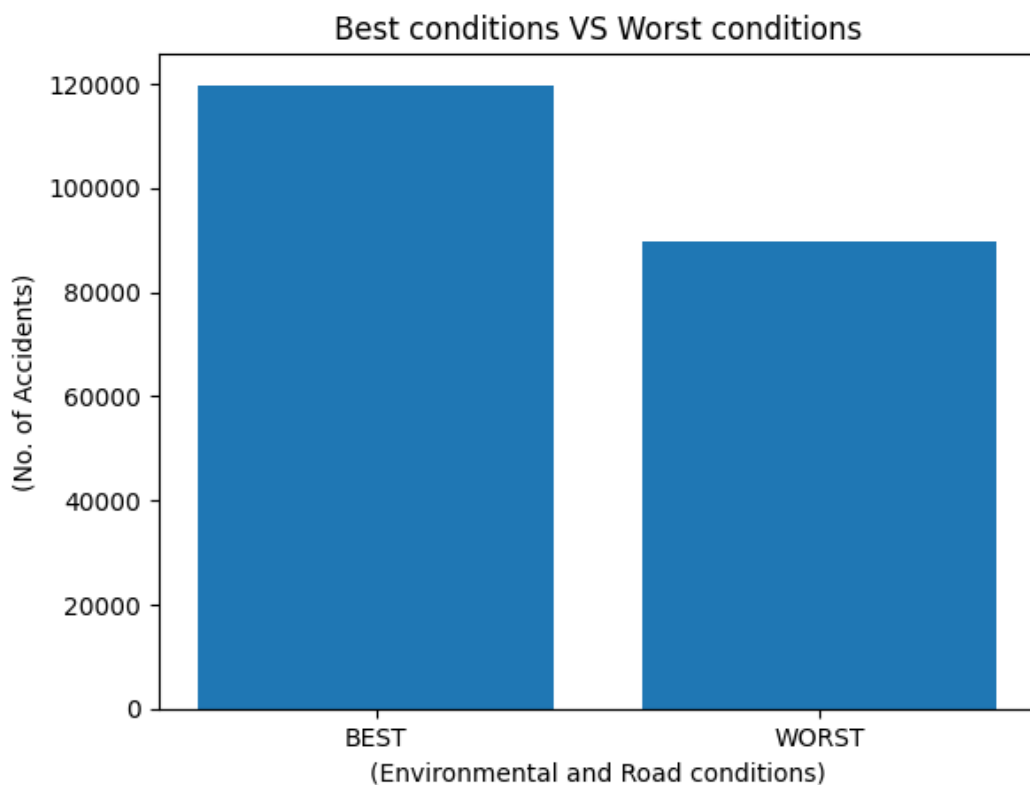
During the day, accidents peak between 7 AM and 10 AM during the morning commute, with a second peak between 4 PM and 6 PM during the evening commute.

Night-time accidents peak between 7 PM and 8 PM, then gradually decrease until morning.



Crashes in Optimal vs. Worst Conditions

Surprisingly, more crashes occur in optimal conditions than in adverse conditions, suggesting that environmental and road conditions are not the only significant factors in accidents.



Conclusions

Based on the comprehensive analysis of the traffic accident dataset containing over 209,000 records, several key insights emerge:

- Most severe accidents resulted in up to 21 injuries and 3 fatalities, with accidents involving 2-6 vehicles causing the highest number of injuries
- Contrary to expectations, most accidents occur in clear weather conditions rather than adverse weather
- Undivided roads and four-way intersections are the most common accident sites, with traffic signals being the most frequent traffic control device present
- An overwhelming 95% of accidents are intersection-related, highlighting the critical nature of these road features
- Carelessness and lack of attention emerge as major contributory factors, though many cases remain undetermined
- Accidents peak during morning (7-10 AM) and evening (4-6 PM) commute hours, with more incidents occurring in daylight than at night

These findings suggest that human factors, rather than environmental conditions, play a crucial role in accident occurrence. This underscores the importance of focusing on driver education, attention, and behaviour in addition to infrastructure improvements. Future safety measures should particularly target intersection safety and peak commuting hours, while also investigating ways to reduce accidents caused by carelessness and inattention.