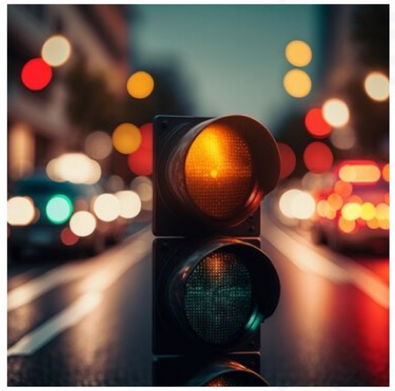


Project Description

This Project is based on Data over Traffic Accidents that was processed for about 10 years from an unknown location.

Traffic Accidents have been prime reason to bring upon unnecessary deaths and loss to more than many % of World's population.

Understanding this data might help us to decide what factor should we give importance to in order to save fatalities and to bring safety to public.

Is it Road Defects or Weather or the Reckless behavior of Driver that caused accidents?

Lets Find out ....

DATA DICTIONARY

CRASHES TABLE:
SERIALNO/CASEID : Primary Key with Unique ID for each incident recorded
DAMAGERANGE: The extent of the damage caused by the accident.
INJURYTYPE : Kind of Injury Caused(Incapacitating, Non-Incapacitating etc)
PRIMARYCONTRIBUTORYCAUSE: The primary cause contributing to the crash.
TOTALINJURIES: The total number of injuries reported.
NUMBEROFVEHICLES: The number of vehicles involved in the accident.

ROADTYPES TABLE:
ALIGNMENT: The alignment of the road where the accident occurred (e.g., straight, curved).
CRASHTYPE: The type of the crash (e.g., head-on, rear-end).
TRAFFICWAYTYPE: The type of roadway involved in the accident (e.g., highway, local road).

WEATHERROADCONDITIONS TABLE :
WEATHERCONDITION: The weather conditions at the time of the accident.
LIGHTINGCONDITION: The lighting conditions at the time of the accident.
ROADWAYSURFACE The condition of the roadway surface (e.g., dry, wet, icy).
ROADDEFECT: Any defects present on the road surface.

YEARTIMEMONTH TABLE :
CRASHHOUR: The hour the accident occurred.
CRASHDAY: The day of the week the accident occurred.
CRASHMONTHNumber: The month the accident occurred.

ADDED AN INDEX COLUMN SINCE THERE WAS NO PRIMAY KEY.

Removed Multiple columns with repeated data.

For Instance, the INJURYTYPE has been mentioned in 5 different columns with their respective category names filled with the number of accidents reported in that category.

Replaced the multi-valued names :

For example: 'DARKNESS, LIGHTED ROAD' to 'LIGHTED' in the LIGHTINGCONDITION to make the analysis much easier.

Changed the Column names of multiple attributes that were unclear and had underscore characters.

Removed Columns with Unneeded data.

For instance, I disregarded the Column with repeated data. Date and Time of the Accident which was mentioned in separate columns. Used the Time detail and classified them as either 'AM'/'PM' , to make data intuitive.

When Are Pedestrian Crashes Most Likely and Most Dangerous?
Exploring timing, locations, and behaviors behind pedestrian injuries.

Data Relationships used : Time Series, Part-to-whole, Nominal Comparison

Report over how Pedestrians were the cause of Injuries

Top 6 Injury Hotspots: Choose ROAD Type:

T-INTERSECTION	NOT DIVIDED	DIVIDED - W/MEDIAN BARRIER
ONE-WAY	FOUR WAY	DIVIDED - W/MEDIAN (NOT RAISED)

SEVERITY of Pedestrian-Involved Crashes:

Total Vehicles Damaged

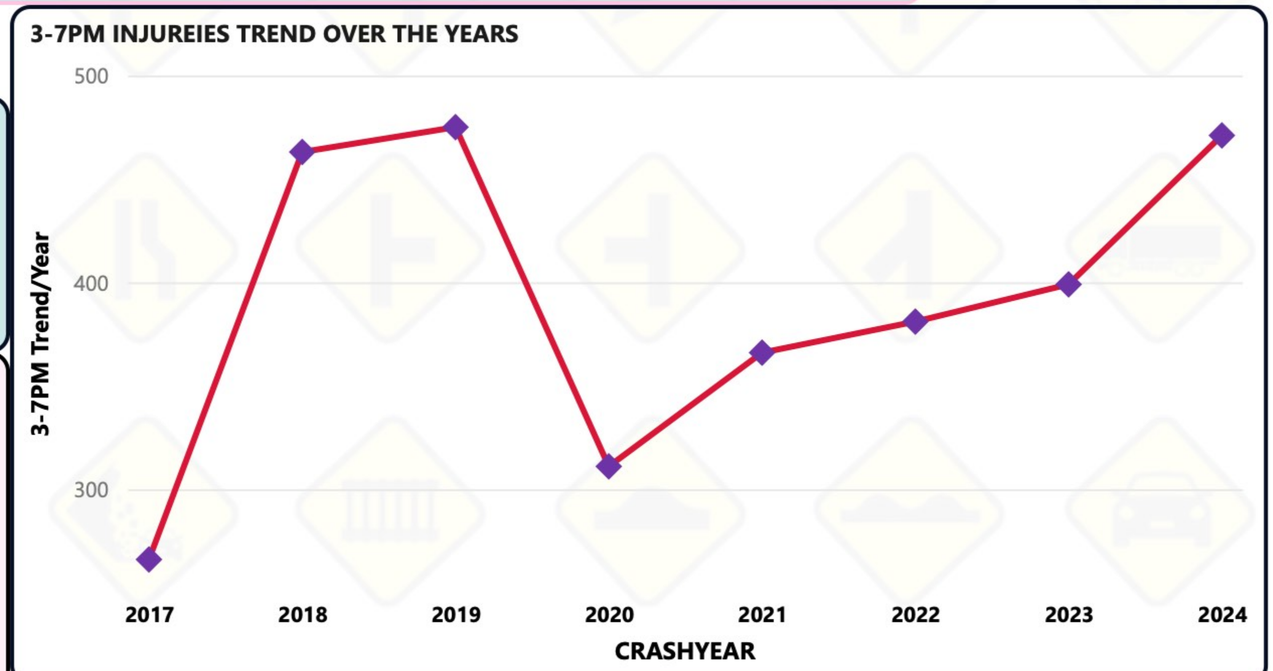
17733

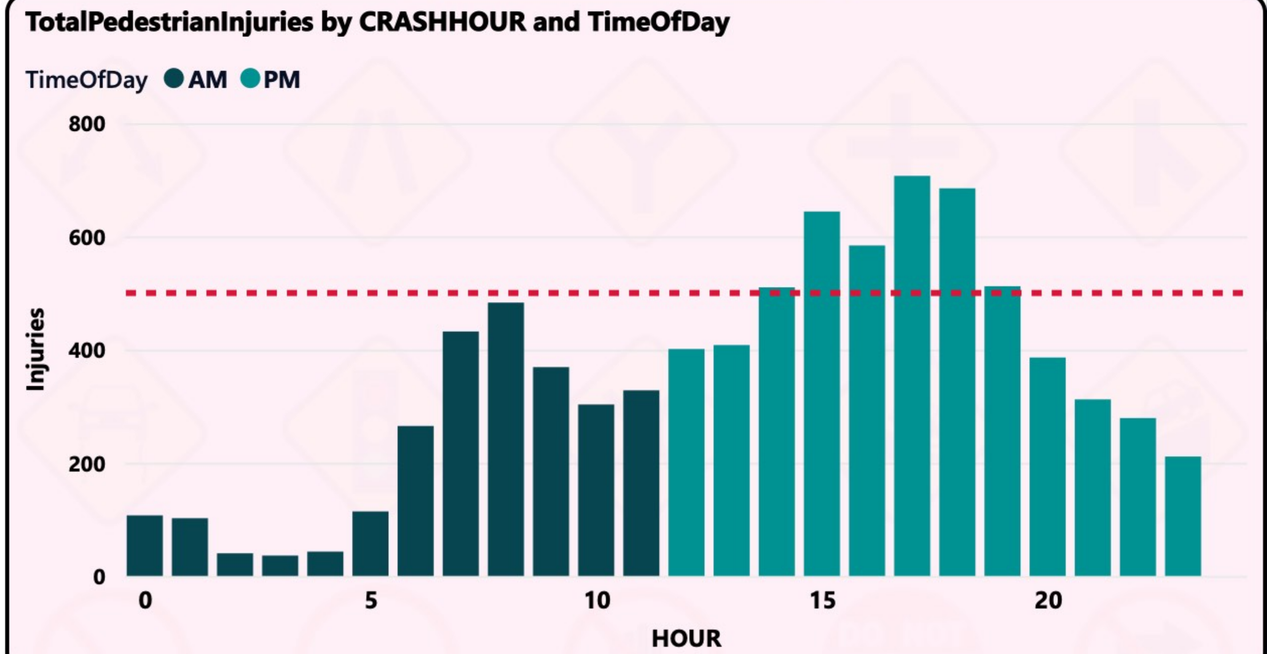
Fatal Injuries (PEDESTRIANS)

117

Injuries Per One Crash (PEDESTRIANS)

0.96





PRIMARYCONTRIBUTORYCAUSE	%GT Sum of TOTALINJURIES
FAILING TO YIELD RIGHT-OF-WAY	49.93%
UNABLE TO DETERMINE	35.74%
FAILING TO REDUCE SPEED TO AVOID CRASH	5.08%
NOT APPLICABLE	4.75%
VISION OBSCURED (SIGNS, TREE LIMBS, BUILDINGS, ETC.)	4.50%

Between the years 2017 and 2024, Pedestrian-related crashes accounted for over 17000 damage to Vehicles, with almost 1 injury recorded for every Case, suggesting that pedestrian related accidents are not minor.
The Hours pattern reveals that the **Injuries spike between 3PM and 7PM**, a window that is even similar to Schools dismissals etc.

The peak hours are especially notable from 2017 to 2019 , but a dip in 2020(probably due to COVID-19 Pandemic resulted in lockdowns), and they gradually peaked until 2024. About **41%**(48 out of 117)of **Fatal** injuries were recorded in the **years 2021 and 2022**. Although the primary cause of such injuries are Drivers not slowing down in time Failure to turn right or recognize pedestrian presence, these both years accounted for **Recklessness** by drivers through Disregarding Traffic signals at **3.68%**
The Traffic way or the environment where most of the injuries had occured are mostly at Intersections, Divided with median or non raised barrier ,at One way or

Road defect & the surface conditions (wet, mud, snow etc) can make the driving for vehicles harder & their visibility impaired. These conditions make the overall situation either 'GOOD' or 'BAD'.

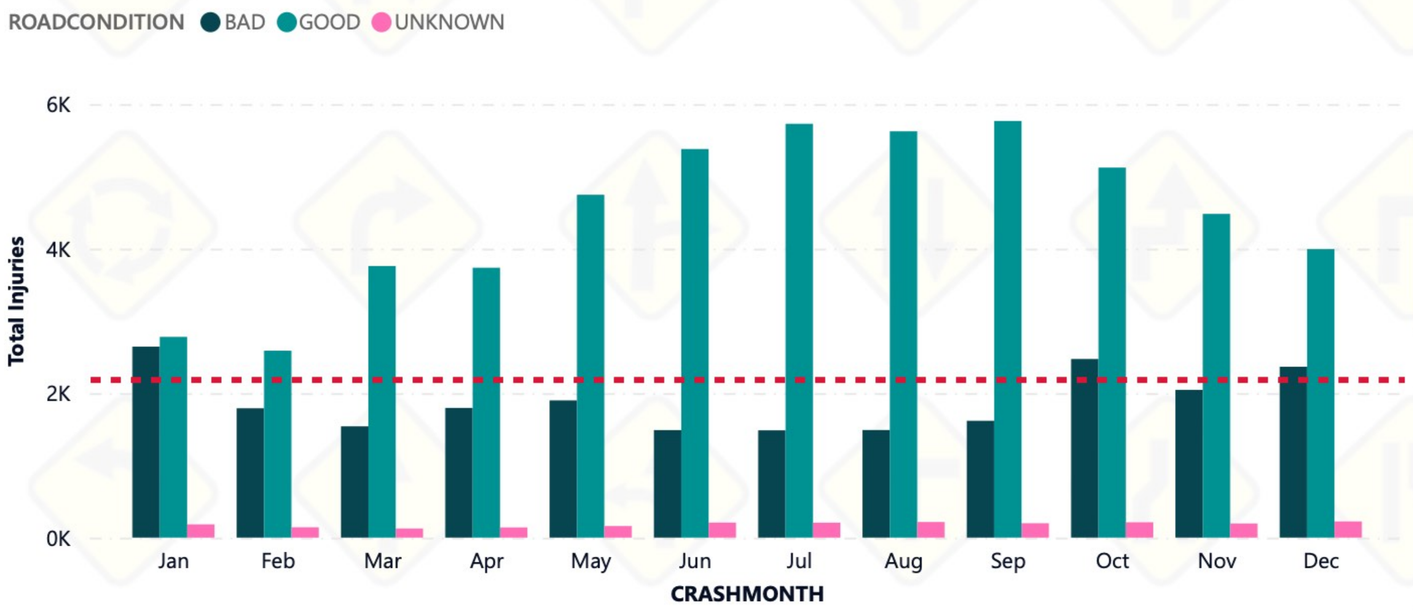
Do they cause more accidents, injuries? Are these injuries fatal or incapacitating? Or are we having a wrong intuition about 'BAD' situations?

Lets find out....

Data Relationships used : Time Series, Distribution, Deviation, Nominal Comparison

The Surprising Risk of 'Good' Roads:

Monthly Injury Trends by Road Condition



Injury Severity by Road Condition:

Average Injuries / Bad Crash

0.39

Average Injuries / Good Crash

0.40

Avg Injury Ratio:

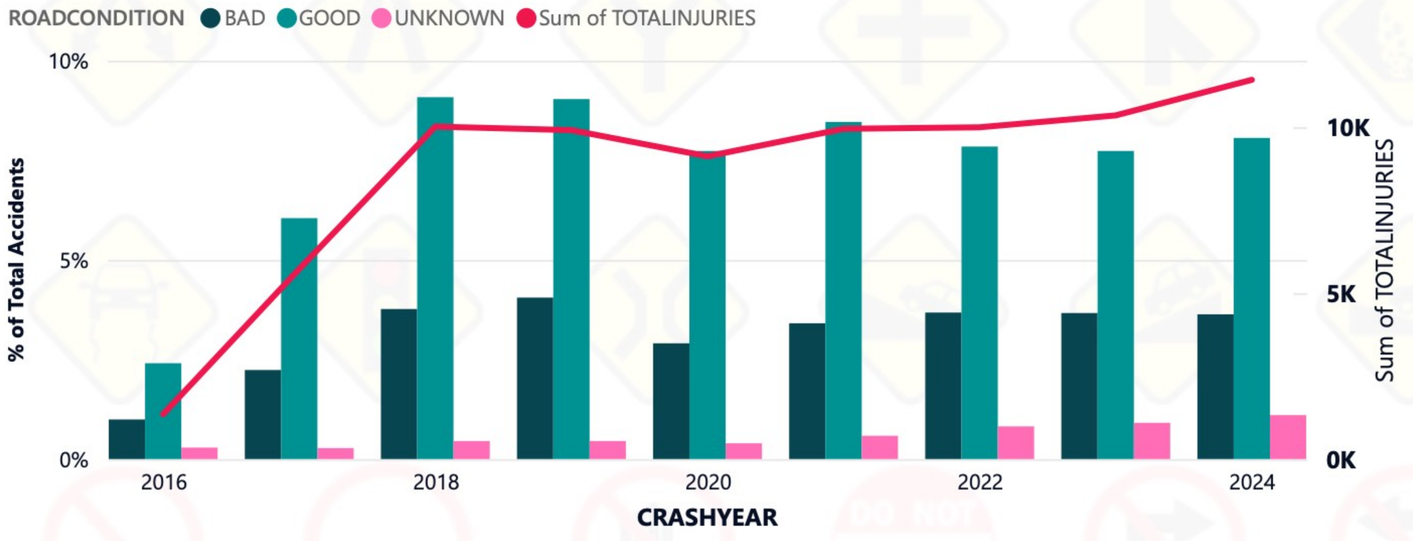
GOOD SEEM RISKY

Top 5 Contributing Causes on GOOD Roads

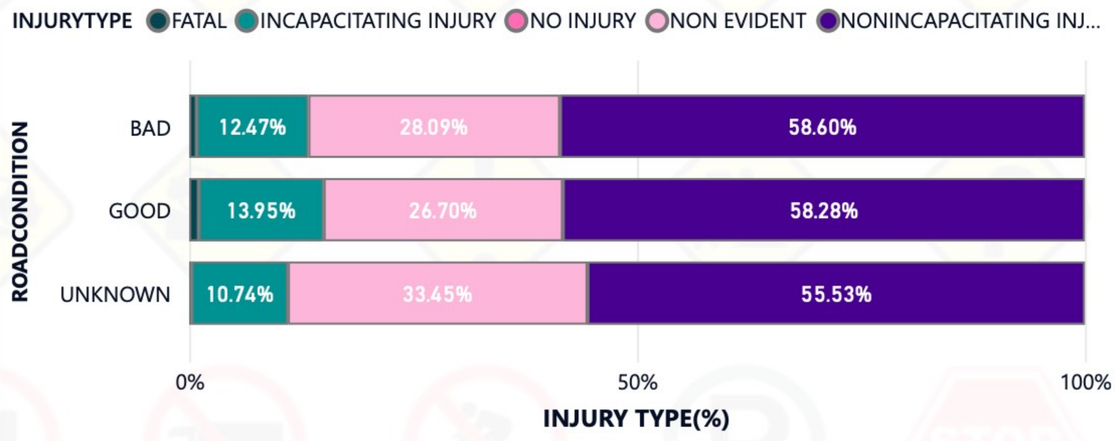
Cause Detail	Cases No(%)	Injuries(%)
FAILING TO YIELD RIGHT-OF-WAY	30.74%	37.76%
UNABLE TO DETERMINE	36.25%	29.78%
DISREGARDING TRAFFIC SIGNALS	10.32%	19.89%
FOLLOWING TOO CLOSELY	13.53%	6.52%
IMPROPER TURNING/NO SIGNAL	9.15%	6.06%

Accidents by Road Condition

For about 10 Years



Share of Each Injury Type per Road Condition



Over the span of 8 years, our report reveals that more injuries have occurred under **GOOD** road conditions than in **BAD** ones. Despite the situation of road surface (mud, gravel etc.) and defects, data shows that **GOOD** conditions may actually **induce riskier driving behavior**, resulting in more injuries. On closer look, the data indicates a rise in accident rates from 2016 to 2019, a dip in 2020 by around 2% (possibly due to the pandemic), followed by a slight rebound in 2021, post which the trend has mostly stabilized. However, injuries have not followed the same pattern. Instead, they show a consistent annual rise of about 1,000 injuries per year, indicating severe consequences even if crash counts stay steady.

Through the months, injuries during **BAD** conditions stayed about halfway lower than its counterpart, which showed consistent peaks all through.

The **BAD** conditions are a bit on a **rise** in the months **Jan, Oct, Nov, and Dec**, indicating higher severity during winter months (possibly). Interestingly, in Jan and Feb, both conditions shared nearly equal injury contributions. (While few of the records were of **Unknown** conditions, which were consistent.)

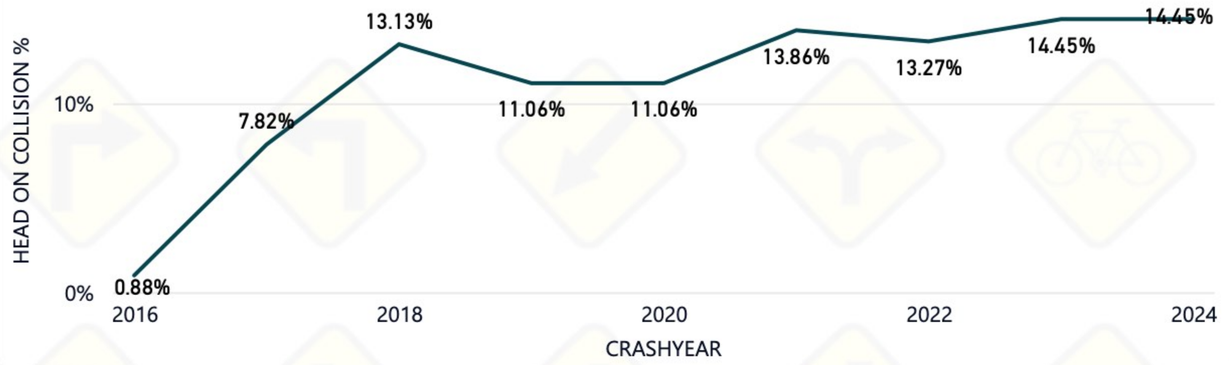
Exploring the Hidden Risks: How Much More Dangerous Are Head-On Collisions Compared to Other Accidents? Uncovering the Trends, Injury Severity, and Financial Impact

Data Relationships used : Time Series, Distribution, Deviation, Nominal Comparison, Ranking, Part-to-whole

Head-On Collisions: A Risk Analysis of Injuries, Financial Impact, and Contributing Factors

Hourly Trend of Head-On Collisions & Injuries

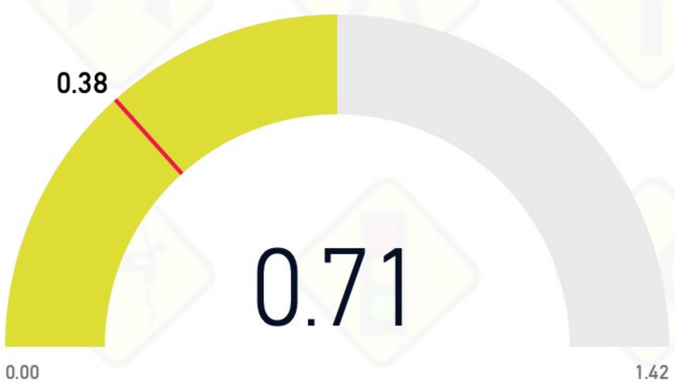
Yearly Share of Head-On Collisions



Severity & Financial Impact of Head-On Collisions

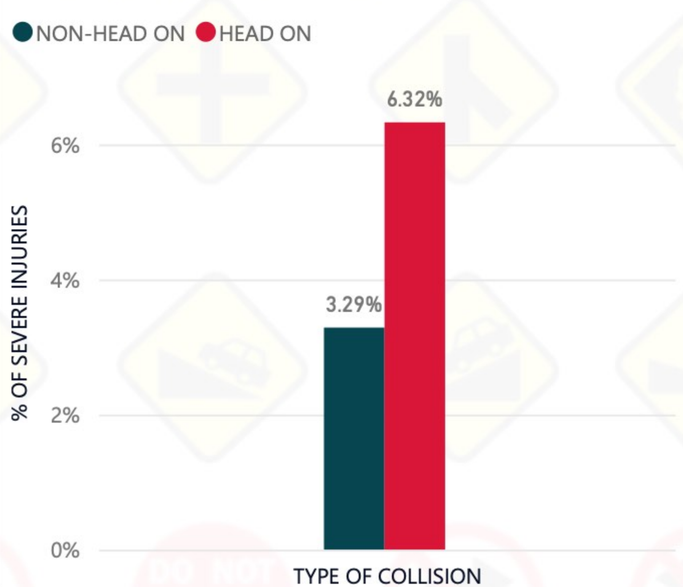
Severity Index:

Head-On Collisions VS Typical Accidents

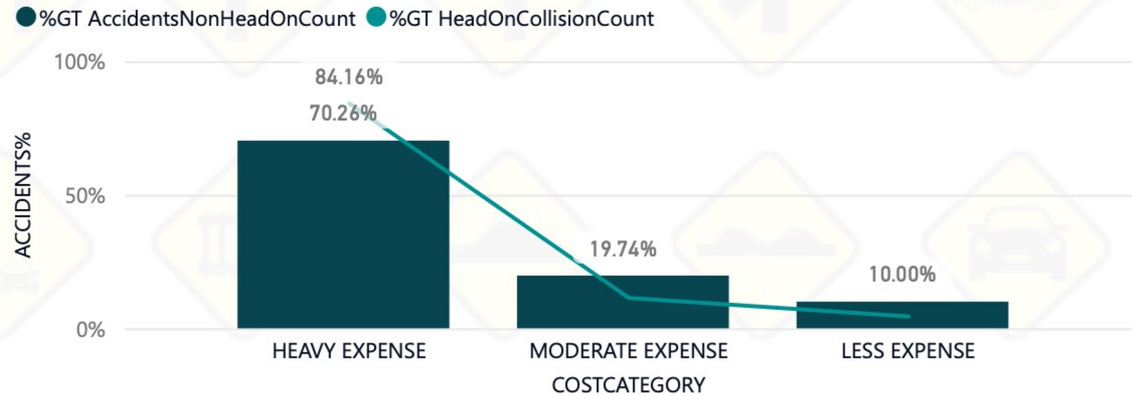


Injury Severity: Head-On vs Other Collisions

FATAL+INCAPACITATING INJURIES



EXPENSE COMPARISION(HEAD ON , NON HEAD ON)



Top Causes of Head-On Collisions

Top Causes by Collision Frequency	% Contribution of Top8
UNABLE TO DETERMINE	39.93%
FAILING TO YIELD RIGHT-OF-WAY	15.56%
DRIVING ON WRONG SIDE/WRONG WAY	13.85%
DISREGARDING TRAFFIC SIGNALS	8.59%
FAILING TO REDUCE SPEED TO AVOID CRASH	7.11%
IMPROPER LANE USAGE	6.96%

The **hourly trend** shows that head-on collision injuries **peak** from about **3 PM** and then see a gentle **decline** into the **night**. Interestingly, the injury trend typically remains **below the collision count**, except for **two spikes at 2 AM and 9 AM**. The **yearly trend** reveals a significant spike from **2016 (0.88%) to 2018 (13.15%)**, followed by a **gradual decline until 2020**, and then a **gentle rise again through 2024**, reaching **14.75%**. This shows a volatile but upward trend over the long term.

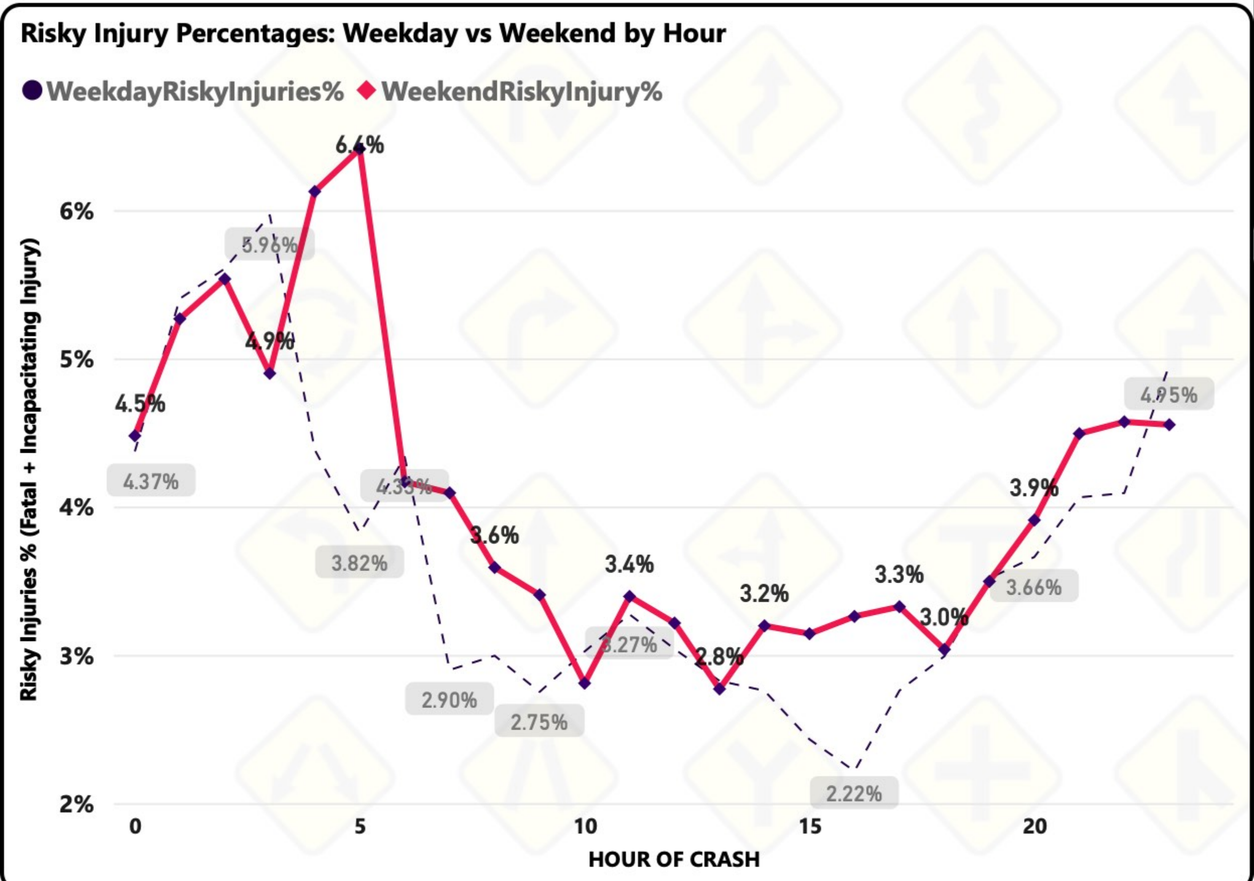
Investigating further to understand its **severity**, head-on collisions account for **double the percentage of fatal and incapacitating injuries** compared to the non-head-on category highlighting their significantly higher impact. Additionally, the **injury rate for head-on collisions stands at 0.71**, **twice that of Non-Head on crashes**, explaining their risk level.

When analyzing **costs**, head-on collisions are associated with **heavier expenses at about 84%** compared to non-head-on collisions, where the cost was about **70%**. However, in other expense categories, **head-on was not the dominant type**, suggesting that while they are more severe, they’re not always the most costly

Severe injuries — both fatal and incapacitating — are among the most concerning outcomes in any crash scenario. Do weekends contribute more to these injuries compared to weekdays? Or is there another primary cause behind them? Do holidays play a role in elevating this risk? And finally, does the time of day or type of day significantly influence when and why severe injuries occur?

Data Relationships used : Distribution, Deviation, Nominal Comparison, Part to whole, Ranking

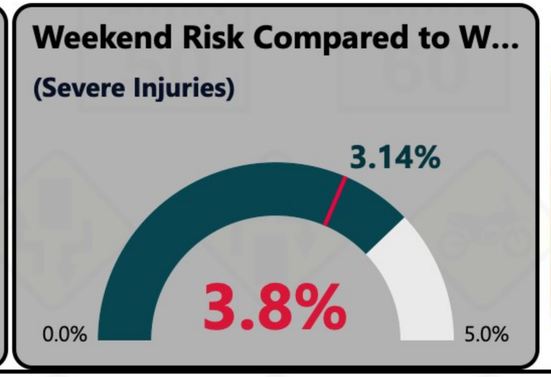
Exploring Timing & Causes Behind Severe Accidents

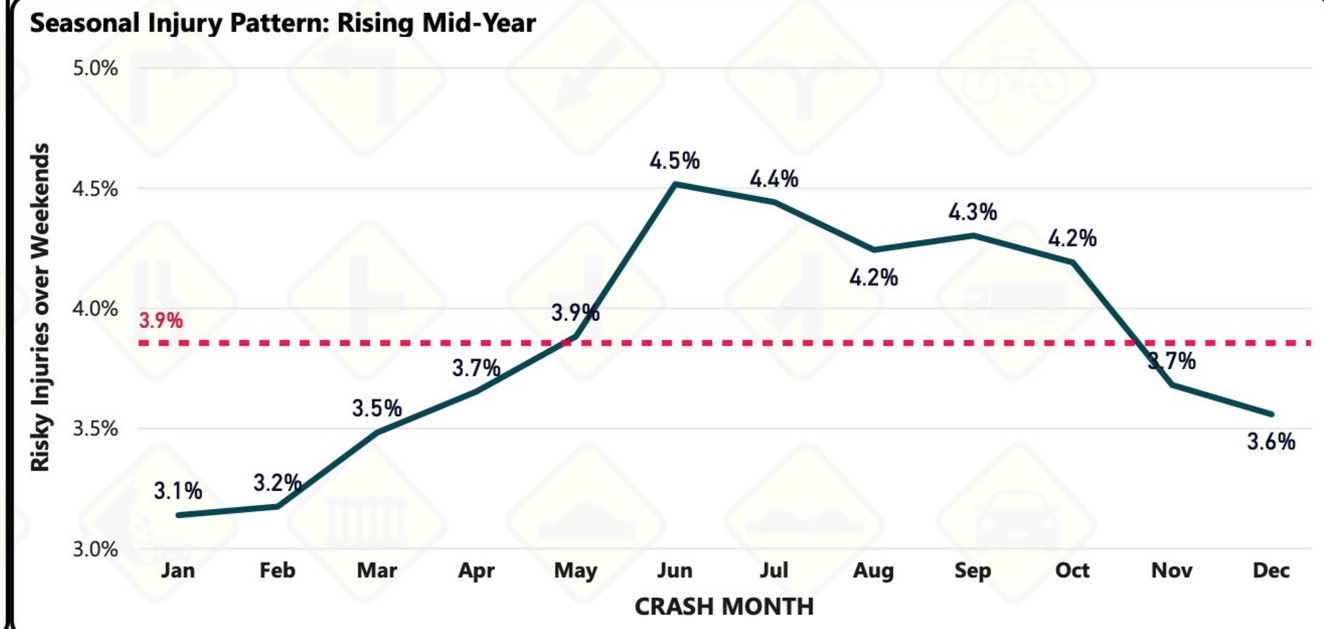


Riskier Weekends:

+20% Serious Injury Rate

20.54%





Top WEEKEND Causes Behind Severe Injuries

PRIMARYCONTRIBUTORYCAUSE	WeekendRiskyInjury%
EXCEEDING AUTHORIZED SPEED LIMIT	16.3%
RELATED TO BUS STOP	14.3%
HAD BEEN DRINKING (USE WHEN ARREST IS NOT MADE)	13.8%
UNDER THE INFLUENCE OF ALCOHOL/DRUGS (USE WHEN ARREST IS EFFECTED)	11.2%
DISREGARDING YIELD SIGN	9.7%
BICYCLE ADVANCING LEGALLY ON RED LIGHT	9.1%
PHYSICAL CONDITION OF DRIVER	8.8%

Top WEEKDAY Causes Behind Severe Injuries

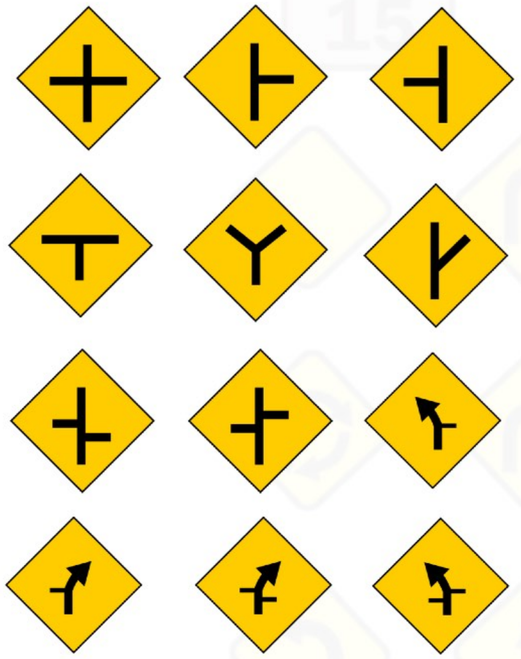
PRIMARYCONTRIBUTORYCAUSE	WeekdayRiskyInjuries%
BICYCLE ADVANCING LEGALLY ON RED LIGHT	25.00%
PHYSICAL CONDITION OF DRIVER	13.52%
UNDER THE INFLUENCE OF ALCOHOL/DRUGS (USE WHEN ARREST IS EFFECTED)	9.69%
EXCEEDING AUTHORIZED SPEED LIMIT	9.45%
DISREGARDING YIELD SIGN	7.92%
OBSTRUCTED CROSSWALKS	7.32%
RECKLESS OR AGGRESSIVE	6.75%

Overall, the data shows that **weekends are about 20.5% riskier than weekdays** in terms of causing **severe injury rate**. While the number of cases recorded during **weekdays** accounted for **3.14%** of severe injuries, **weekends** accounted for **3.8%**, suggesting a **notable increase in risk during weekends**.

As far as **time of day** is concerned, severe injuries appear to occur at a **higher rate between 12 AM and 5 AM**, and start to **rise** again from **6 PM**. The trend for weekdays appears almost the same as the weekend trend, but **stays slightly below the weekend's trend almost every time**, except for small spikes at **3 AM and 12 AM**. The **lowest Severe injury rate is observed at 10 AM**, with fluctuations continuing between **10 AM and 6 PM**.

Looking at the **monthly trend**, most risky injuries **peaked between May and October**, with a **steady rise from January to May(May being heavy)**. There's a slight dip among **victims between June and September**, but the overall stretch points to **warmer months being more injurysome** which might even include **vacation time**.

The **top primary causes of severe injuries** differ slightly between weekdays and

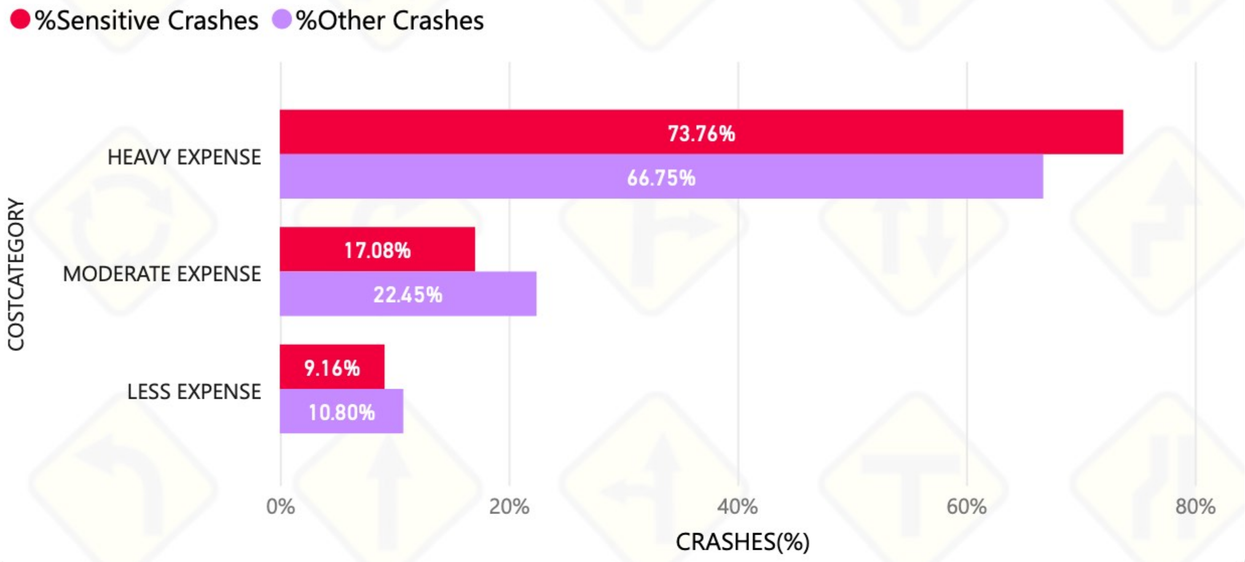


Do Road Dividers and Major Junctions Contribute to More Injuries and High-Cost Crashes?

This following report(REP5) explores how specific road types like medians, T-intersections, and roundabouts relate to crash volume, injury severity, and cost burden.

Are Sensitive Roads More Dangerous or Just More Costly?

Sensitive Roads Have More High-Cost Crashes Than Others

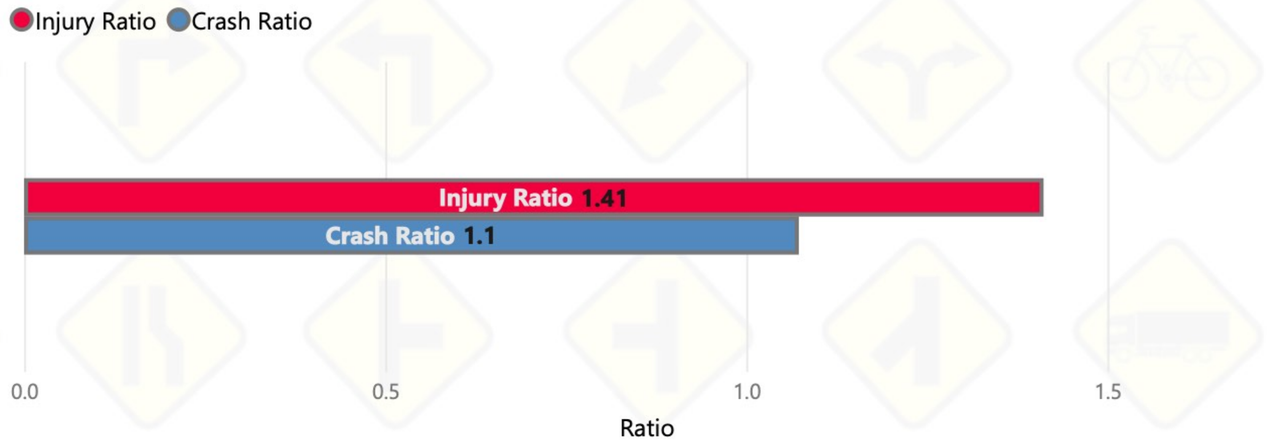


Injuries per 100 Crashes (Sensitive Roads)

43 Injuries/100 Crashes

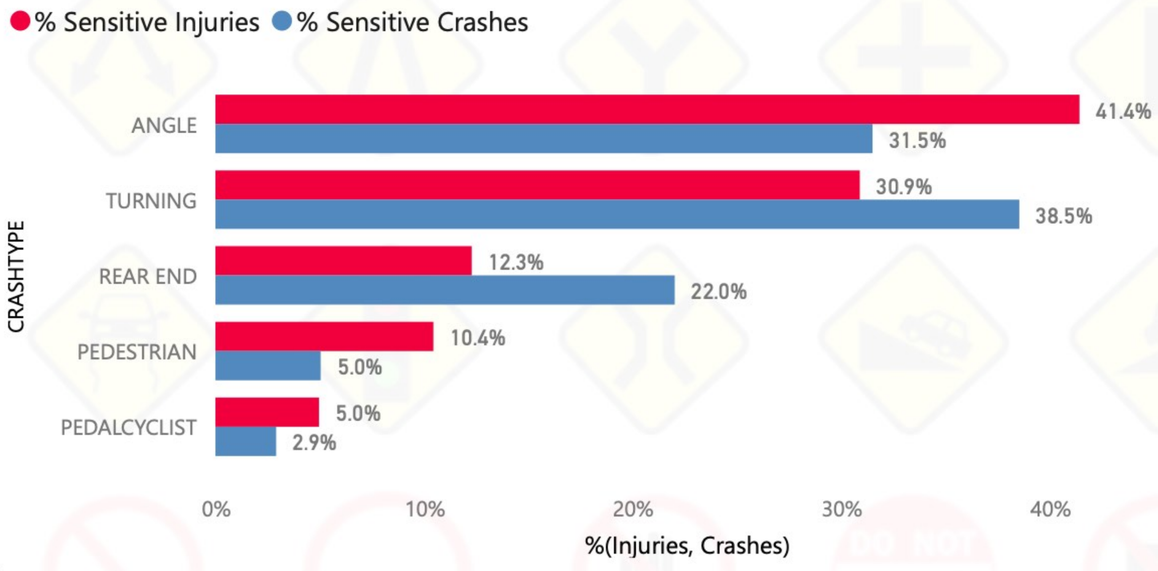
Sensitive Roads: Crash Count Increases Slightly

But Injuries Rise Sharply by 41%



Top 5 Crash Types by % of Injuries & Crashes (Sensitive Roads)

Based on Trafficway-defined sensitivity



Injury Severity Distribution by Road Type

The publisher of Dumbbell Bar Chart by Nova Silva requires a license. Email the report or dashboard contact. Or get a license now.

Learn about licensed visuals

Email contact | Get a license

Sensitive road categories include locations usually believed to be more prone to accidents such as intersections, four-way junctions, or divided roads with improper barriers. These areas are considered higher-risk due to complex vehicle movement. The data shows that about **73% of crashes on sensitive roads resulted in heavy expenses**, while the **non-sensitive category accounted for about 66%**, a figure still relatively close. In the moderate expense category, non-sensitive roads slightly dominated by about 5%, but for **lower expense**, both categories were similarly represented. Overall, **cost is a common factor** across all road types.

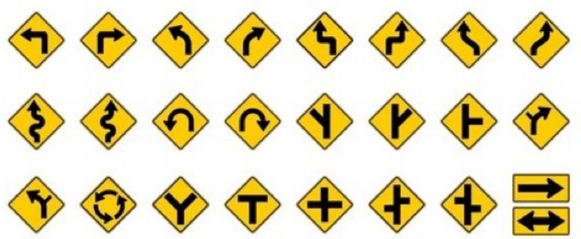
From an injury perspective, **every 100 crashes** on sensitive roads resulted in **43 injuries**, which means nearly **1 injury for every 2 crashes** — a high ratio compared to typical traffic averages.

Additionally, sensitive roads accounted for **10% more crashes** than other roads, but the injuries for the same are 41% higher. Although the crashes are slightly higher, the injuries are way higher posing serious risk on sensitive roads.

Looking closer at crash scenarios, most incidents on sensitive roads occurred at

Which months see the highest crash volumes, and are weather or road alignment key contributors? How does injury severity change in these peak months?

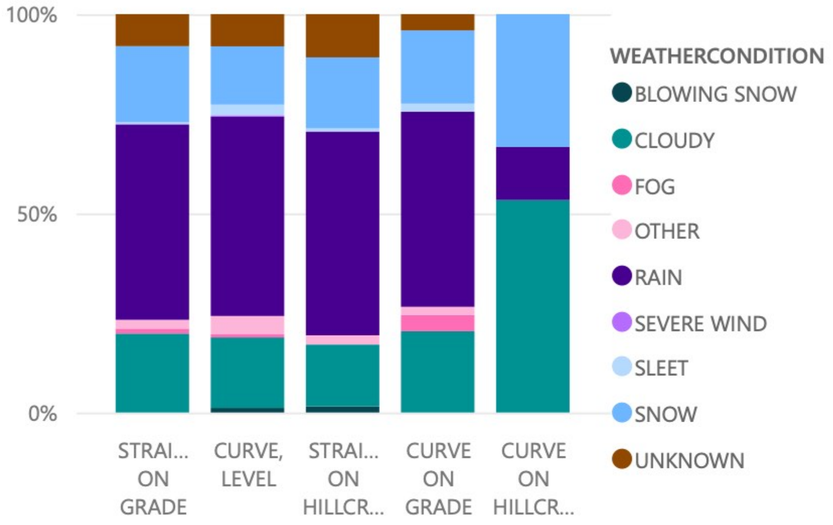
Data Relationships used : Distribution, Nominal Comparison, Part to Whole



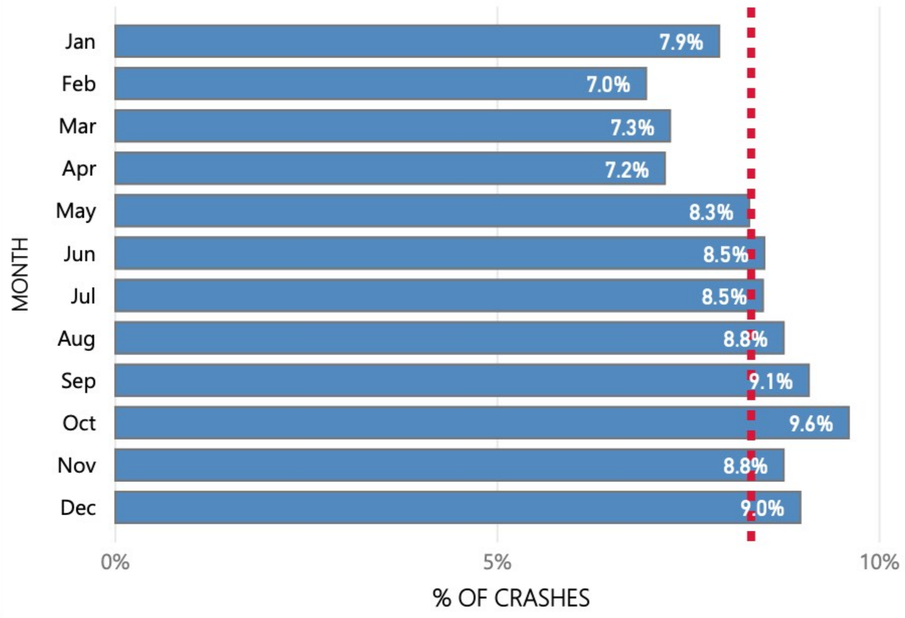
When & Why Do Crashes Peak? A Monthly and Weather-Based Breakdown

Crash Patterns on Different Alignments

Under Adverse Weather

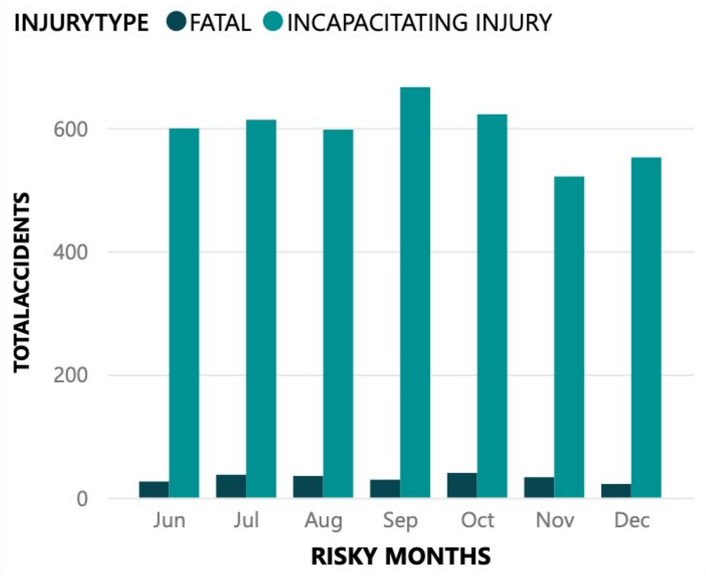


Crash Volume by Month – When Are Roads Riskiest?



Injury Severity in Peak Crash Month

(Excludes Non-Injury Cases)



Top 6 Injury-Heavy Months – What's Driving the Severity?

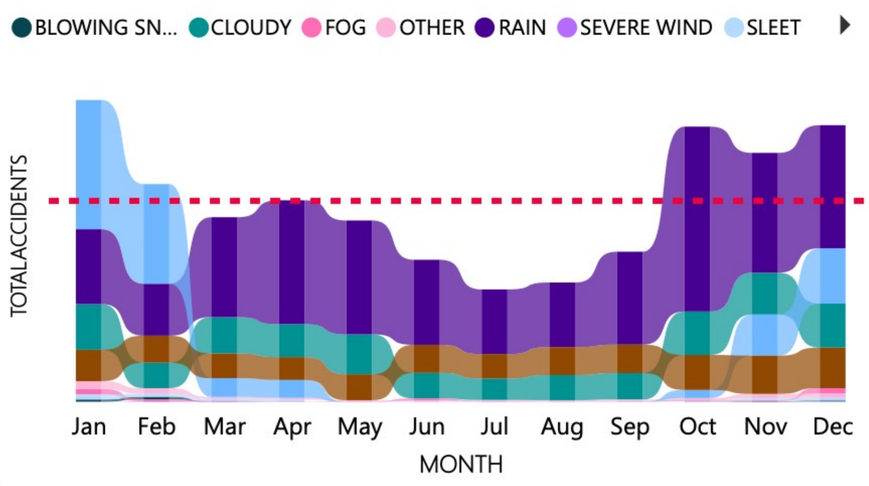
WEATHERCONDITION	Jun	Jul	Sep	Oct	Nov	Dec	Total
RAIN	8.60%	7.33%	9.30%	19.13%	12.09%	12.83%	**69.28%**
CLOUDY	2.44%	2.30%	2.93%	3.92%	3.70%	3.50%	**18.80%**
SNOW	0.01%	0.10%	0.00%	0.93%	2.80%	3.95%	**7.79%**
OTHER	0.19%	0.14%	0.15%	0.23%	0.53%	0.73%	**1.97%**
FOG	0.16%	0.11%	0.07%	0.15%	0.10%	0.71%	**1.30%**
SLEET			0.00%	0.04%	0.25%	0.34%	**0.63%**
BLOWING SNOW				0.04%	0.03%	0.12%	**0.19%**
SEVERE WIND	0.01%			0.00%	0.01%	0.01%	**0.04%**
%Contribution/Month	**11.42%**	**9.98%**	**12.45%**	**24.44%**	**19.51%**	**22.20%**	**100.00%**

Inside TOP6: What Drove the Crash Spike?

ⓘ ***Please Click the button to explore crash causes and risky road types.***

Monthly Crash Trends by Weather Condition

(Excluding 'CLEAR' Conditions))



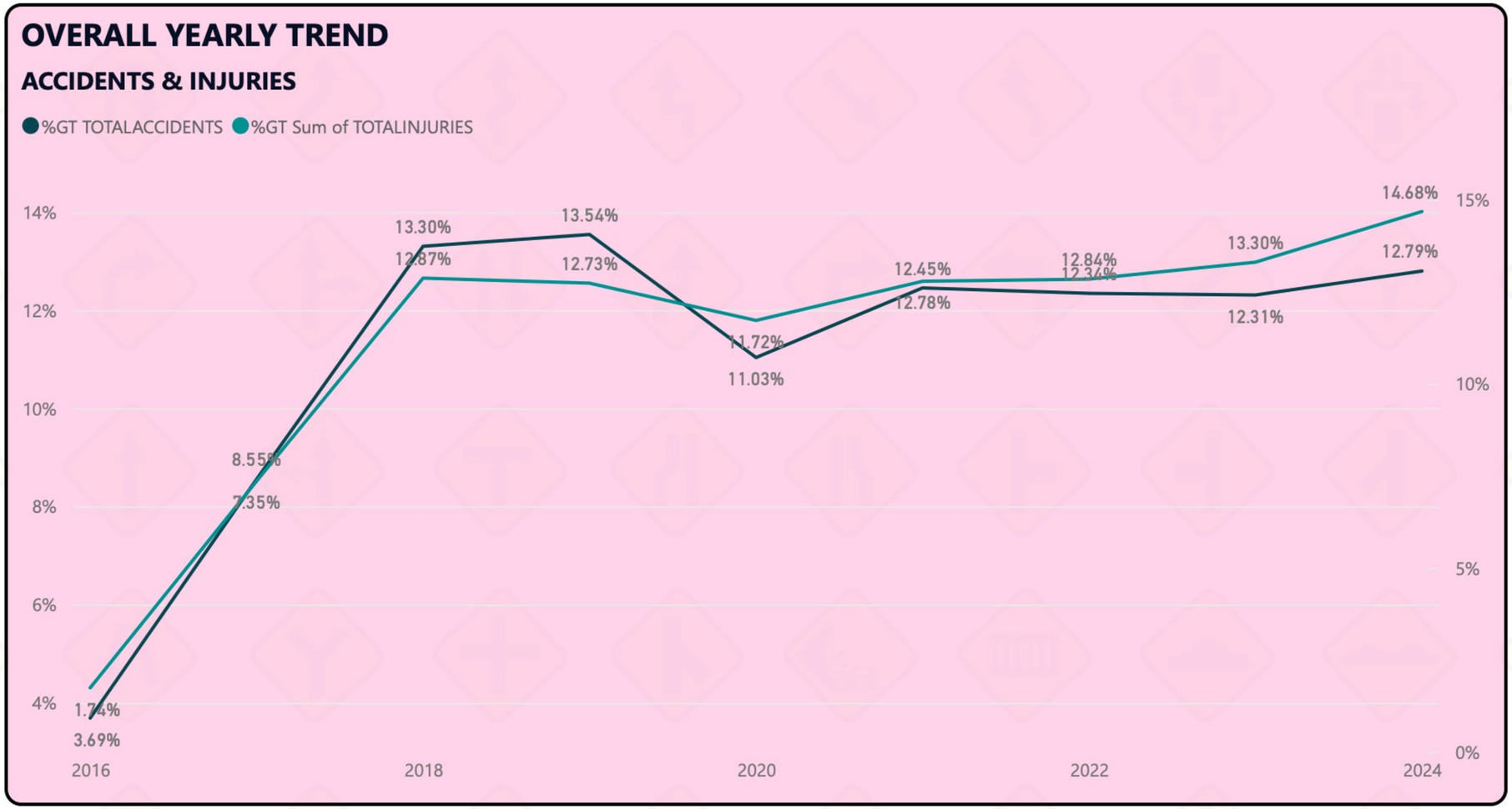
The report reveals that **the second half** of the year consistently recorded **higher crash volumes**, with **October** being the most **riskiest** month overall. These months, **summer** into early **winter** also correlate with a spike in injury severity. In fact, each month from **July to December** reports at **least 600 incapacitating** injuries, along with an average of **27 fatal** injuries per month, highlighting a seasonal shift toward more dangerous conditions.

From a **weather** perspective, there's strong evidence pointing to **adverse weather** as a **major** crash **contributor**. The **last quarter** of the year is dominated by **rain**, aligning with the peak in crash and injury data, while the **first two months** are led by **snow**, with total crash counts exceeding **4,000**. Notably, **around 70%** of injuries in the **Top 6 injury-heavy** months occurred during **rainy** conditions, posing serious risk associated with wet weather.

When **road alignment** is taken into the analysis, Most crashes on **curves and straight** alignments occur during rain, but **"curve on hillcrest"** conditions stand out, with **over 50% of those crashes happening in cloudy weather**, and an additional **18% during**

What Primary cause Contributed to the most Vehicle Damage ?

PRIMARYCONTRIBUTORYCAUSE	Sum of NUMBEROFVEHICLES	%Vehicle Damage	TOTALVEHICLEDAMAGEALLCAUSES
UNABLE TO DETERMINE	116208	27.49%	422733
FAILING TO YIELD RIGHT-OF-WAY	87154	20.62%	422733
FOLLOWING TOO CLOSELY	38899	9.20%	422733
DISREGARDING TRAFFIC SIGNALS	30564	7.23%	422733
IMPROPER TURNING/NO SIGNAL	25067	5.93%	422733
FAILING TO REDUCE SPEED TO AVOID CRASH	22257	5.27%	422733
IMPROPER OVERTAKING/PASSING	16701	3.95%	422733
DISREGARDING STOP SIGN	14081	3.33%	422733
IMPROPER LANE USAGE	12862	3.04%	422733
NOT APPLICABLE	10120	2.39%	422733
DRIVING SKILLS/KNOWLEDGE/EXPERIENCE	10018	2.37%	422733
WEATHER	5949	1.41%	422733
IMPROPER BACKING	4594	1.09%	422733
RECKLESS OR AGGRESSIVE	4036	0.95%	422733
VISION OBSCURED (SIGNS, TREE LIMBS, BUILDINGS, ETC.)	3603	0.85%	422733
DISTRACTION - FROM INSIDE VEHICLE	2653	0.63%	422733
DRIVING ON WRONG SIDE/WRONG WAY	2502	0.59%	422733
DISREGARDING OTHER TRAFFIC SIGNS	2257	0.53%	422733
EQUIPMENT - VEHICLE CONDITION	1912	0.45%	422733
UNDER THE INFLUENCE OF ALCOHOL/DRUGS (USE WHEN ARREST IS EFFECTED)	1840	0.44%	422733
PHYSICAL CONDITION OF DRIVER	1590	0.38%	422733
DISTRACTION - FROM OUTSIDE VEHICLE	1488	0.35%	422733
Total	422733	100.00%	422733



This Time Series Line chart is plotted to just show that the severity of number of Accidents or Injuries did not take a positive turn for about 8 years containing rows of about 2,00,000.

MY OVERALL TAKE

Designing smarter barriers, signals, or speed breakers in areas such as curves might reduce alignment risks, especially when the weather is not entirely suitable for driving.
Rains can now be categorized as heavy-risk times(from our analysis), and investing in illuminated borders on the roads and reducing speed limits by about half could give positive outcomes. Good weather ironically leads to higher injury rates, likely due to overconfidence. This points to the need for awareness campaigns and stricter enforcement of laws against reckless driving, including signal violations and drunk driving.

Considering the heavy risk due to pedestrians, increasing the crosswalk sign frequency during 3–7 PM at intersections and four-way roads might reduce pedestrian-related injuries, which are significantly higher than in any other analysis I've done. Enhancing the vetting process for driving licenses and shortening renewal intervals could help maintain driver accountability.

The yearly plot clearly shows that neither injuries nor accidents have shown any positive trend, making it subtly clear that transportation authorities might not have introduced many changes — nor have drivers. Most analyses point to recklessness as a prime factor. Increasing night-time checks could help address this issue further, especially considering the late-night spikes in severe injuries.