

U.S. Accidents (2016-2025)

Capstone Projects

Zainab Alsaffar - DAB16

Table of Content

Introduction.....	3
Problem Statement.....	3
Objectives	3
Aim.....	4
Target Audience	4
Dataset	4
Data Handling.....	8
Analysis.....	9
Recommendations	19
Limitations	19
References	20

Introduction

Traffic Accidents represent a significant challenge affecting millions of people across the **U.S.** each year. With the continuous increase in vehicle numbers and expansion of the road network, accidents necessitate research to identify the contributing factors to the rising number of injuries and fatalities. Despite advancements in transportation systems, safety technologies, and the efforts of traffic authorities and specialized agencies, national accident trends indicate a persistent rise in crash rates across many regions.

This project analyzes **U.S. traffic accidents** from **2016 to 2023** with the objective of identifying the conditions that contribute to higher crash severity. By examining the interaction between **time patterns, weather conditions, visibility(mi), distance(mi), and road type**, this study provides a comprehensive, data-driven assessment that can inform the development of more effective road-safety policies aimed at reducing accident occurrence and severity.

Problem Statement

The U.S. has reported over **7.7M traffic accidents** between **2016** and **2023**, leading to an increase in **moderate to severe** crashes across several states and a rise in injury rates and road safety risks.

Objectives

- 1- What are the yearly trends in the number of traffic accidents in the U.S. from 2016 to 2023?
- 2- Which U.S. states have the highest number of accidents?
- 3- Which U.S. states record the highest number of severe traffic accidents?
- 4- How does travel distance influence accident severity in the U.S?
- 5- How do different weather conditions influence both the number and severity of traffic accidents across the U.S.?
- 6- Which Months and Seasons Show Higher Accident Severity Levels?
- 7- Does limited visibility in foggy conditions significantly increase crash severity levels?
- 8- Which weather scenarios are most dangerous during the day compared to at night?
- 9- What is the relationship between humidity levels and the severity of road accidents?
- 10- During which hour of the day do most accidents occur across the United States?
- 11- Are traffic accidents more frequent during weekdays or weekends in the United States?
- 12- Which road features become more dangerous during nighttime driving conditions?
- 13- How many crashes occur at each road feature, and how does accident severity differ across them?

Aim

To analyze U.S. traffic accident data (2016–2023) to uncover the main factors behind rising accident frequency and severity, and to provide evidence-based recommendations that reduce severe crashes

Target Audience

1. Federal transportation agencies (**NHTSA & U.S. DOT**) : to understand accidents pattern and weather related risk
2. State Departments of Transportation (**FDOT**): to identify state-level hotspots and improve infrastructure.
3. General Publics (**Drivers**): to increase awareness of dangerous conditions such as rain, fog, nighttime driving, and seasonal risks.

Dataset

Data link: [U.S. Accidents 2016-2023](#)

Data size: 7.7M Columns / 46 Rows

● Original Data:

Column Name	Definition	Format
ID	unique identifier of the accident record.	string
Source	Source of raw accident data	string
Severity	Shows the severity of the accident, a number between 1 and 4, where 1 indicates the least impact on traffic and 4 indicates the highest impact on traffic	Integer
Start Time	show start time of the accident in local time zone	Date
End Time	show end time of the accident in local time zone	Date
Start Lat	shows latitude in GPS coordinate of the start point	Float
Start Lng	shows longitude in GPS	Float

	coordinate of the start point	
End Lat	shows latitude in GPS coordinate of the end point	Float
End Lng	shows longitude in GPS coordinate of the end point	Float
Distance (mi)	The length of the road extent affected by the accident in miles	Float
Description	Show human description of the accident	String
Street	Show the street name in address field	String
City	Show the city in address field	String
State	Show the state in address field	String
ZIPcode	Show the ZIP code in address field	String
Country	Show the country in address field	String
Timezone	Show timezone based on the location of the accident (e.g. Eastern, Central)	String
Airport Code	Denotes an airport-based weather station which is the closest one to location of the accident.	String
Weather_Timestamp	Shows the time-stamp of the weather observation record (in local time).	Timestamp
Temperature (F)	Shows the temperature (in Fahrenheit).	Float
Wind Chill (F)	Shows the wind chill (in Fahrenheit).	Float
Humidity (%)	Shows the humidity (in percentage).	Float
Pressure (in)	Shows the air pressure (in inches).	Float
Visibility (mi)	Shows visibility (in miles).	Float

Wind Direction	Shows wind direction.	Float
Wind Speed (nph)	Shows wind speed (in miles per hour).	Float
Precipitation(in)	Shows precipitation amount in inches, if there is any.	Float
Weather Condition	Shows the weather condition (rain, snow, thunderstorm, fog, etc.)	String
Amenity	A POI annotation which indicates presence of amenity in a nearby location.	Boolean
Bump	A POI annotation which indicates presence of speed bump or hump in a nearby location.	Boolean
Crossing	A POI annotation which indicates presence of crossing in a nearby location.	Boolean
Give Away	A POI annotation which indicates presence of give_way in a nearby location.	Boolean
Junction	A POI annotation which indicates the presence of a junction in a nearby location.	Boolean
No Exit	A POI annotation which indicates presence of no_exit in	Boolean

	a nearby location.	
Railway	A POI annotation which indicates presence of a railway in a nearby location.	Boolean
Roundabout	A POI annotation which indicates the presence of a roundabout in a nearby location.	Boolean
Station	A POI annotation which indicates the presence of a station in a nearby location.	Boolean
Stop	A POI annotation which indicates presence of stop in a nearby location.	Boolean
Traffic Calming	A POI annotation which indicates presence of traffic_calming in a nearby location.	Boolean
Traffic Signal	A POI annotation which indicates presence of traffic_signal in a nearby location.	Boolean
Turning Loop	A POI annotation which indicates presence of turning_loop in a nearby location.	Boolean
Civil Twilight	Shows the period of day (i.e.	String

	day or night) based on civil twilight.	
Nautical Twilight	Shows the period of day (i.e. day or night) based on nautical twilight.	String
Astronomical Twilight	Shows the period of day (i.e. day or night) based on astronomical twilight.	String
Sunrise Sunset	Shows the period of day (i.e. day or night) based on sunrise/sunset.	String

- **Cleaning Data:**

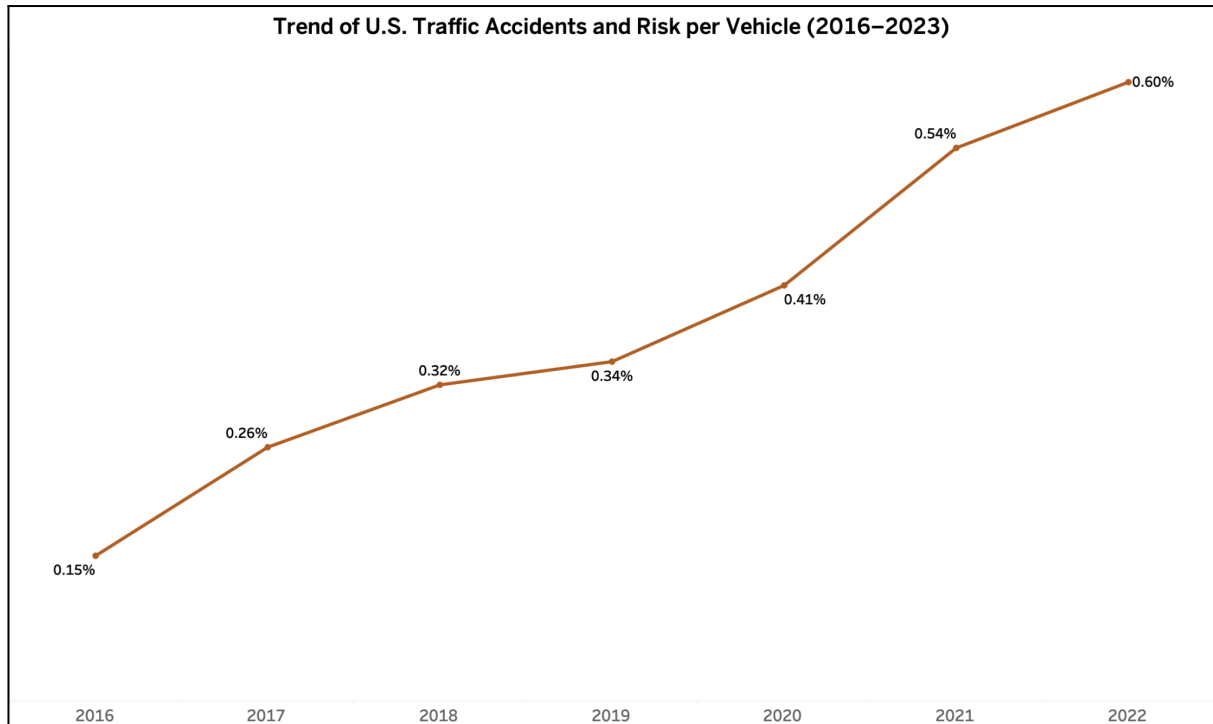
Weather Type	Classify weather type into (rain, cloudy, snow, ice & freezing, fog, hazard, thunderstorm)	String
Number of Cars	Number of cars from 2016-2023	Integer

Data Handling

Please refer to the Jupyter Notebook.

Analysis:

1- What are the yearly trends in the number of traffic accidents in the U.S. from 2016 to 2023?

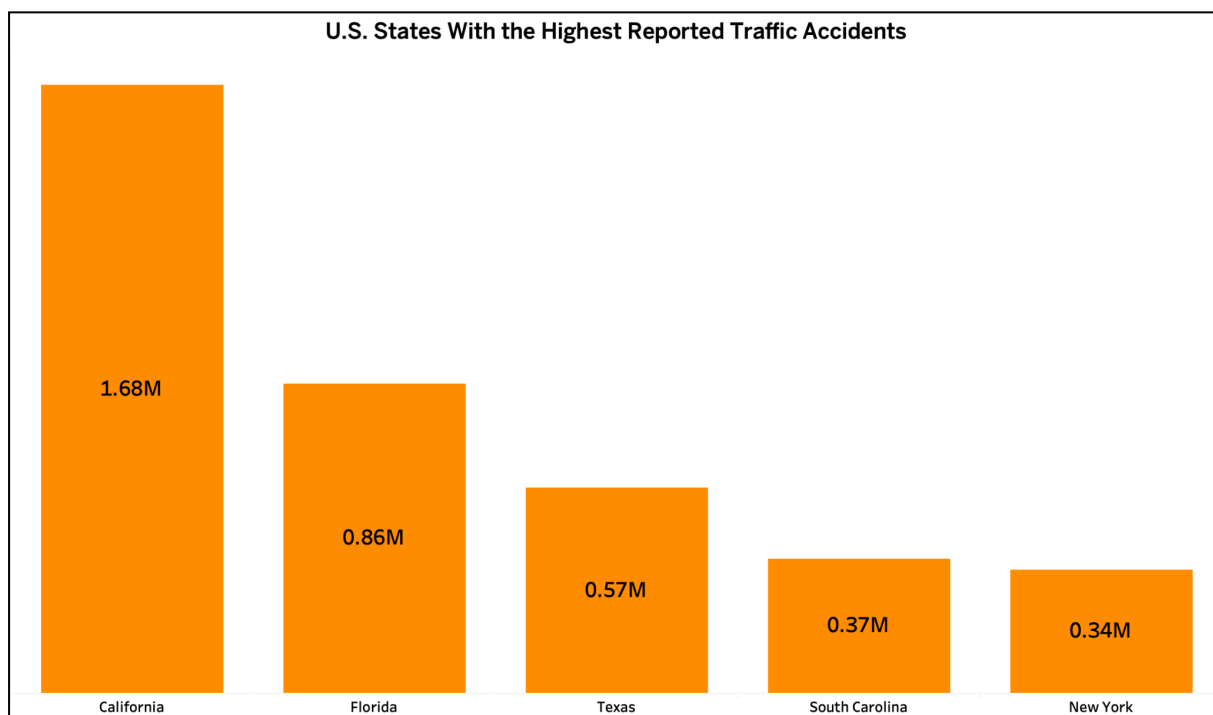


The line chart shows that the number of traffic accidents has been **continuously increasing** along with the **growing number of vehicles**. The highest spike occurred in **2021**, mainly due to **speeding, alcohol, and distracted driving**.

for Number of Cars: I calculate the Ratio of accidents per each car

Ratio: Count [ID] / Avg[Number of car]

2- Which U.S. states have the highest number of accidents?



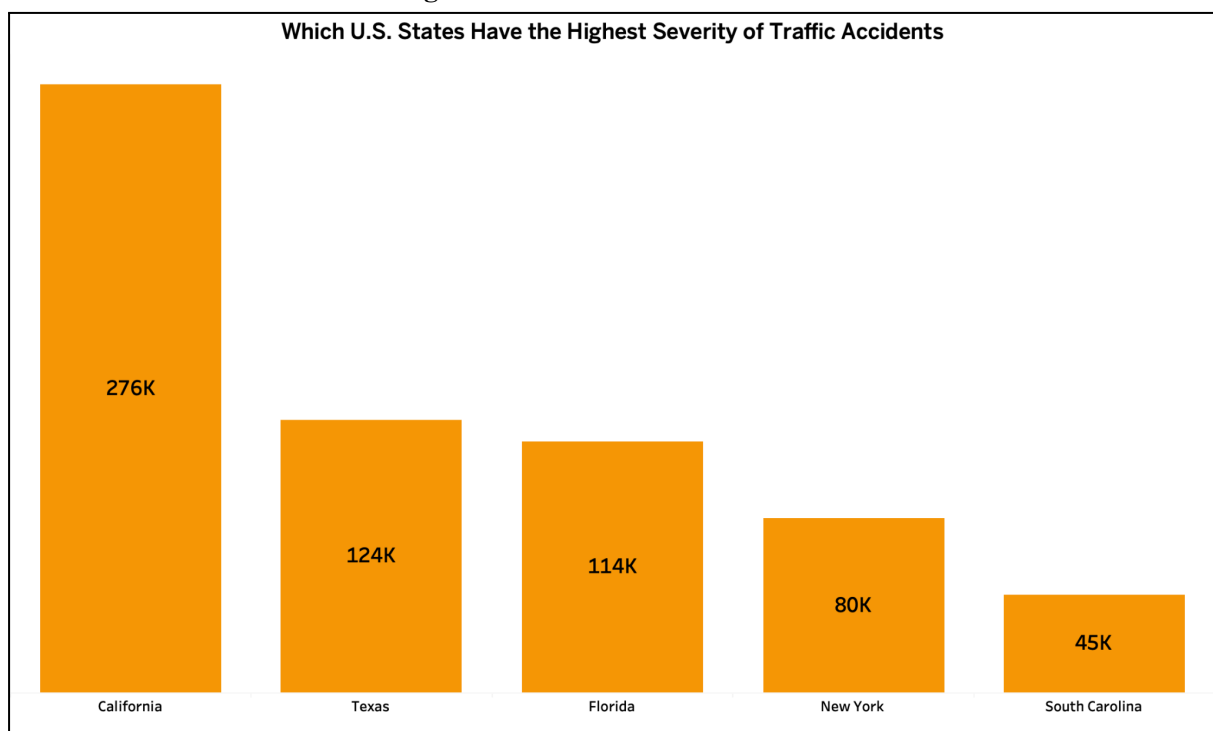
The chart shows that states with high population and heavy road congestion tend to record the highest number of accidents.

The chart indicates that **California (1.6M)** has the greatest number of accidents by a large margin due to its population, crowded highways, and major urban centers such as **Los Angeles** and **San Francisco**.

Florida ranks as the second-highest state in accident counts, largely because of its high tourism activity and frequent rainfall.

As for **South Carolina**, although its population is lower than that of **New York**, it still has a higher number of accidents. This is mainly due to **weaker infrastructure** compared to New York.

3- Which U.S. states record the highest number of severe traffic accidents?

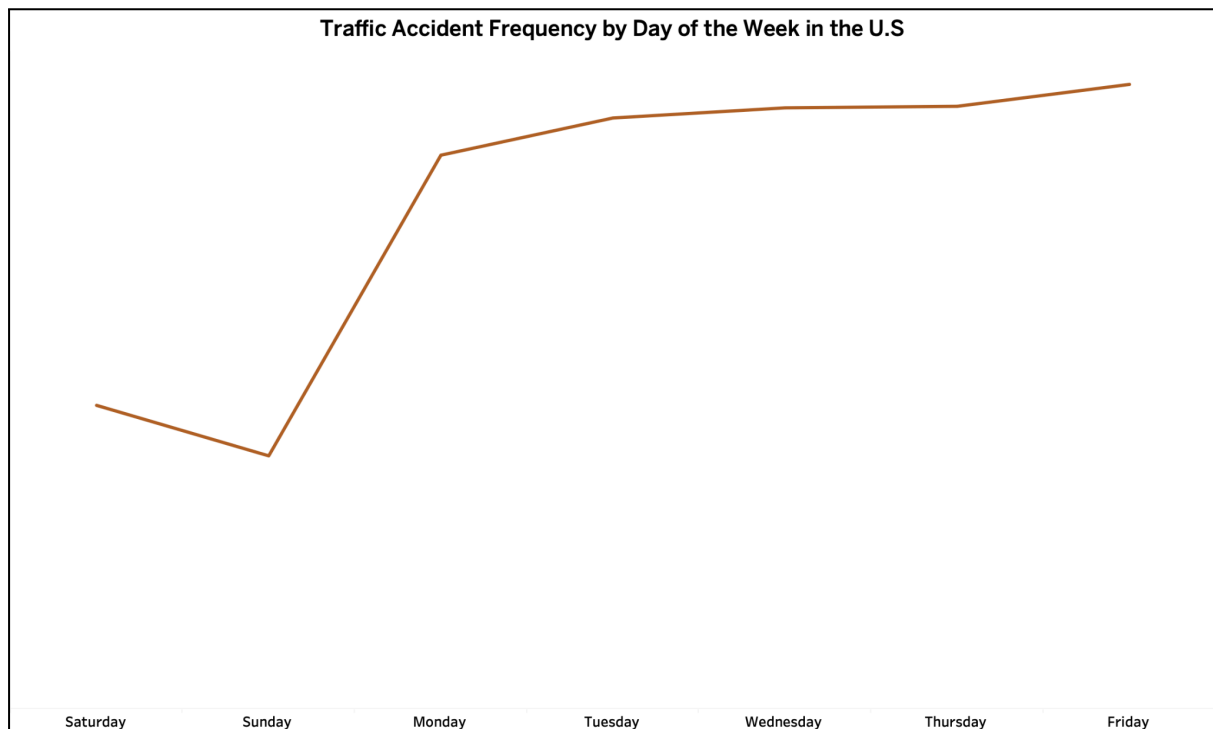


The chart shows that **California** has the highest accident severity (**276K**), which reflects its intense traffic density and heavily congested road networks.

In comparison, **Texas** records higher accident severity (**124K**) than **Florida**, due to several contributing factors, including:

1. A large number of teenage, inexperienced, or unlicensed drivers
2. Extensive rural areas with limited emergency response access
3. Short merge lanes, sudden exits, and busy intersections
4. Distracted or fatigued drivers, especially in high-truck-traffic regions

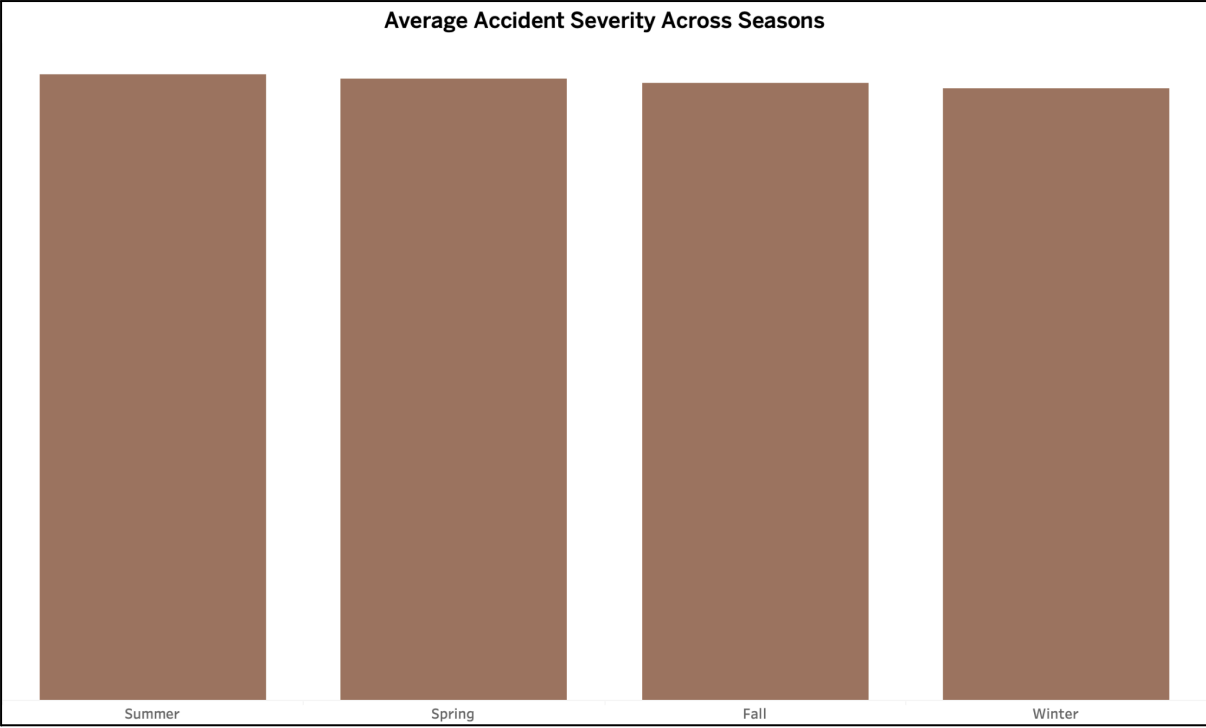
4- Are traffic accidents more frequent during weekdays or weekends in the United States?



The chart shows that **weekends (Saturday and Sunday)** have the **lowest number of accidents**, as schools and workplaces are closed, resulting in lighter traffic and less congestion.

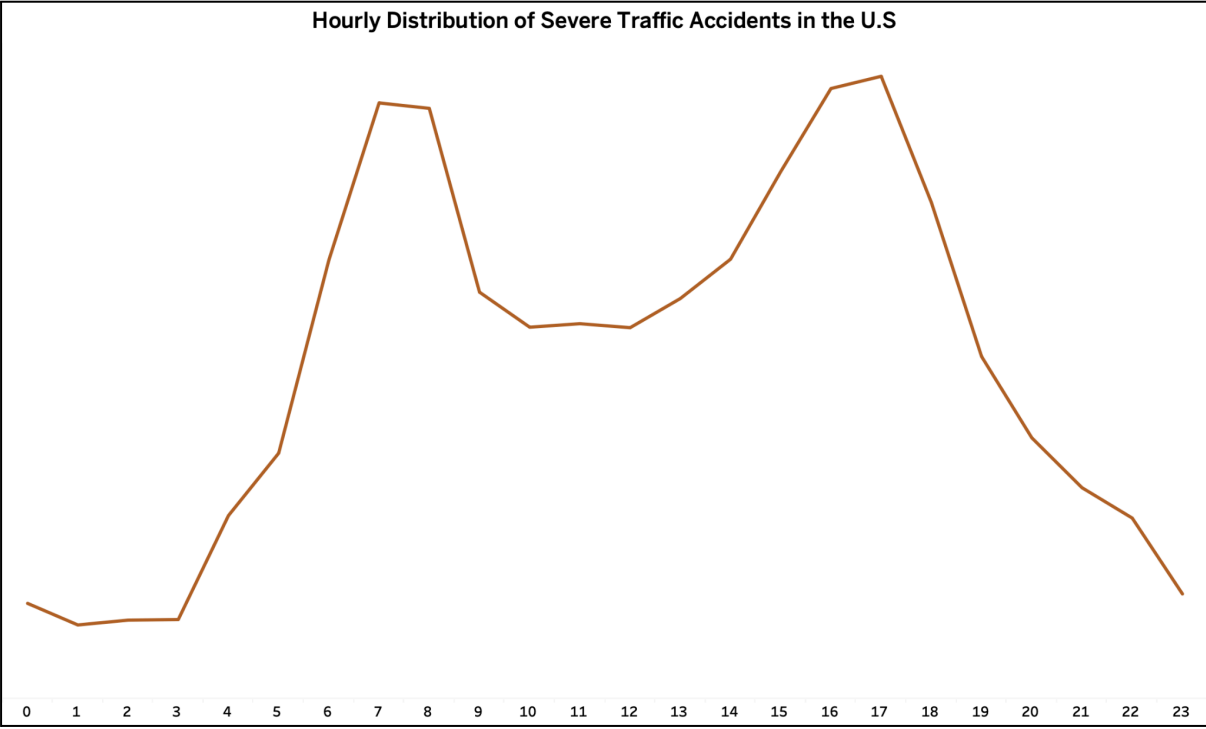
In contrast, **accidents increase sharply at the start of the workweek**, with **Monday showing the highest jump**. This spike reflects the beginning of the week, when road density and congestion rise significantly. Throughout the rest of the weekdays, accident counts **remain consistently high** due to heavy traffic and daily commuting patterns.

5- Which Months and Seasons Show Higher Accident Severity Levels?



While weather patterns vary by season, accident severity remains nearly identical across summer, spring, fall, and winter. This suggests that seasonal changes alone are not a major factor in determining how severe a crash becomes.

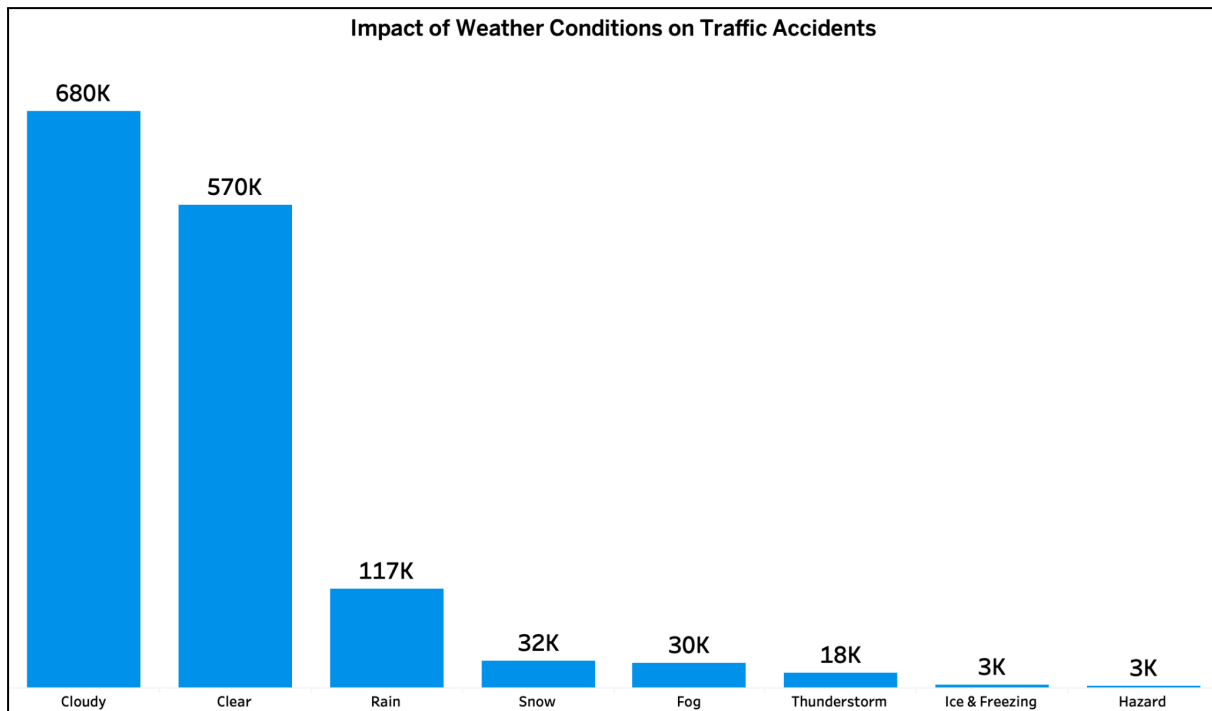
6- During which hour of the day do most accidents occur across the United States?



The figure shows that severe accidents occur most frequently during two peak periods.

The first peak is **7–9 AM**, driven by heavy congestion as people commute to work and school, often accompanied by speeding. The second peak is **4–6 PM**, during the evening rush hour.

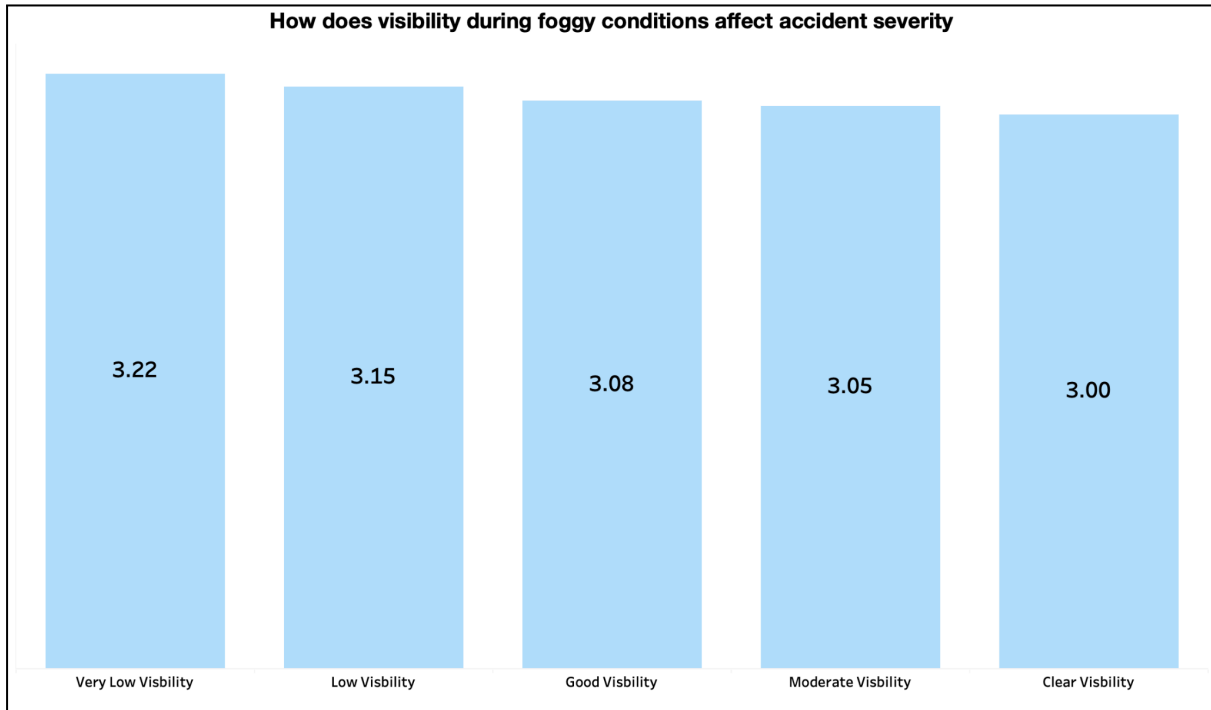
7- How do different weather conditions influence both the number and severity of traffic accidents across the U.S.?



Most accidents occur during **normal weather conditions** (whether cloudy or clear), primarily because these conditions are the most common. In such weather, some drivers tend to become less cautious and drive at higher speeds, which increases the likelihood of accidents.

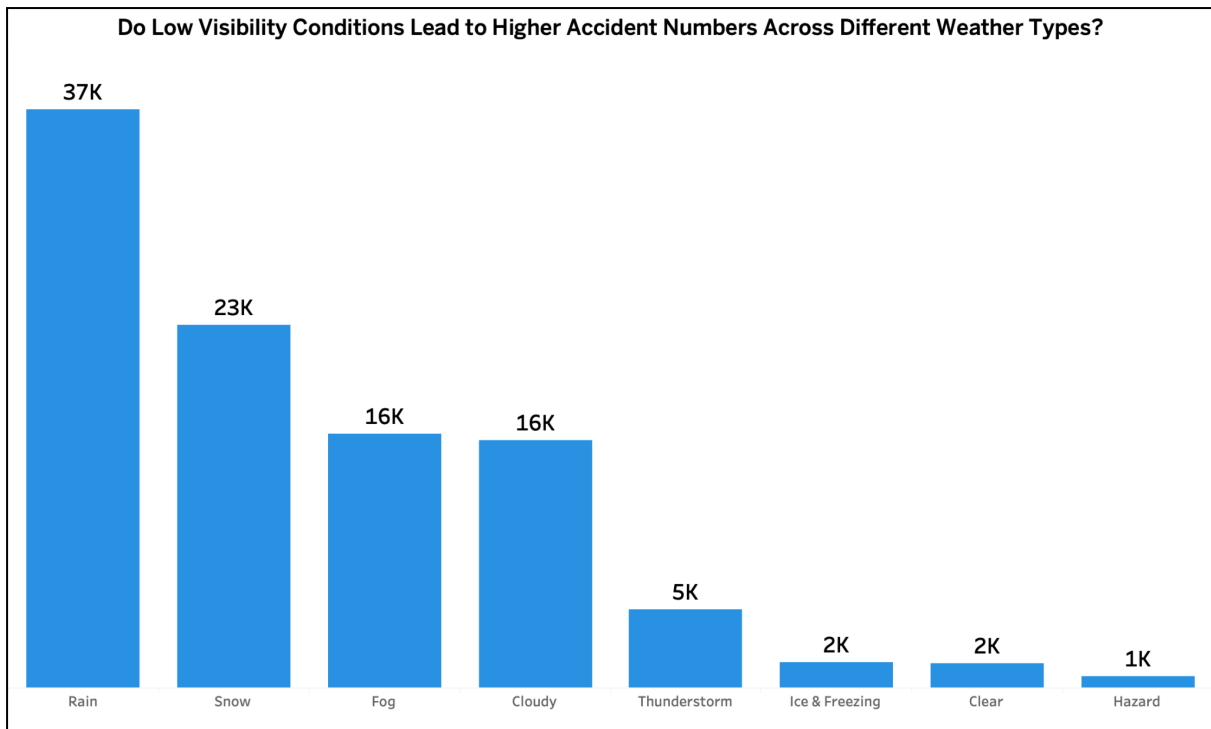
However, **rain, fog, snow, and thunderstorms**, although less frequent, create far more dangerous situations and lead to **more severe accidents**. This indicates that dangerous weather conditions significantly increase accident severity, even if the total number of accidents remains lower.

8- Does limited visibility in foggy conditions significantly increase crash severity levels



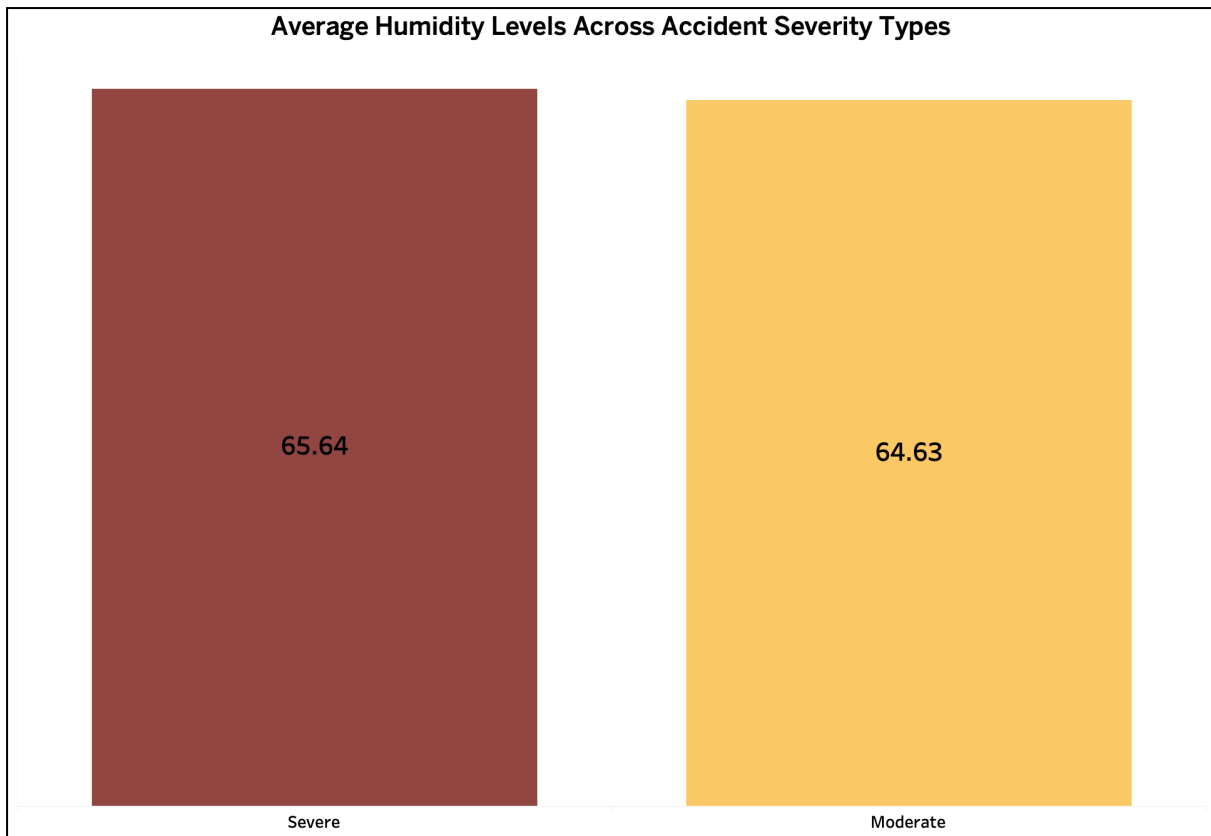
During **foggy conditions**, accident severity increases sharply as visibility decreases. Extremely low visibility leads to highly severe accidents. This highlights that **fog combined with poor visibility is a major risk factor** for high-impact accidents.

9- Which weather scenarios are most dangerous during the day compared to at night?



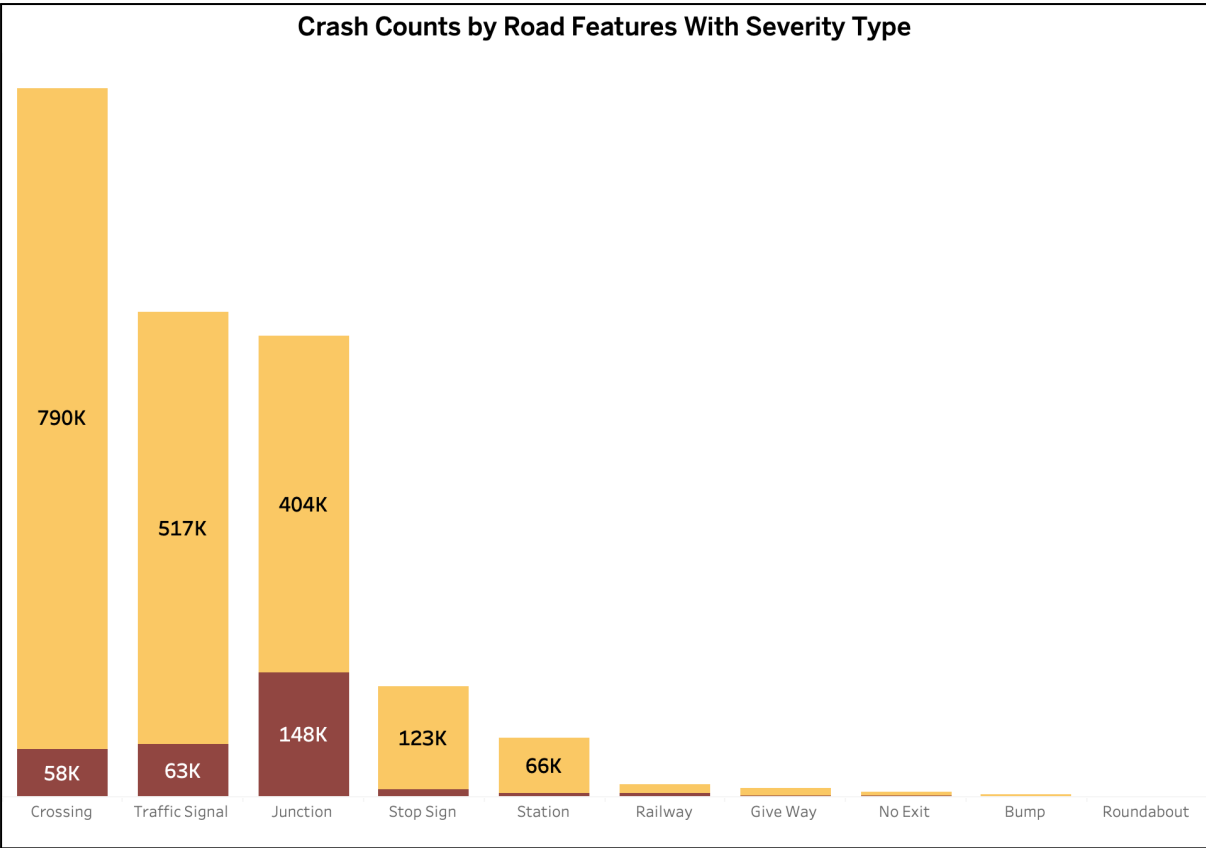
Severe accidents are most frequent during **rainy conditions with reduced visibility**, making rain the most dangerous weather type. It is followed by **snow and fog**, both of which significantly increase accident severity. In contrast, **clear and cloudy weather** contribute very little to severe accidents.

10- What is the relationship between humidity levels and the severity of road accidents?



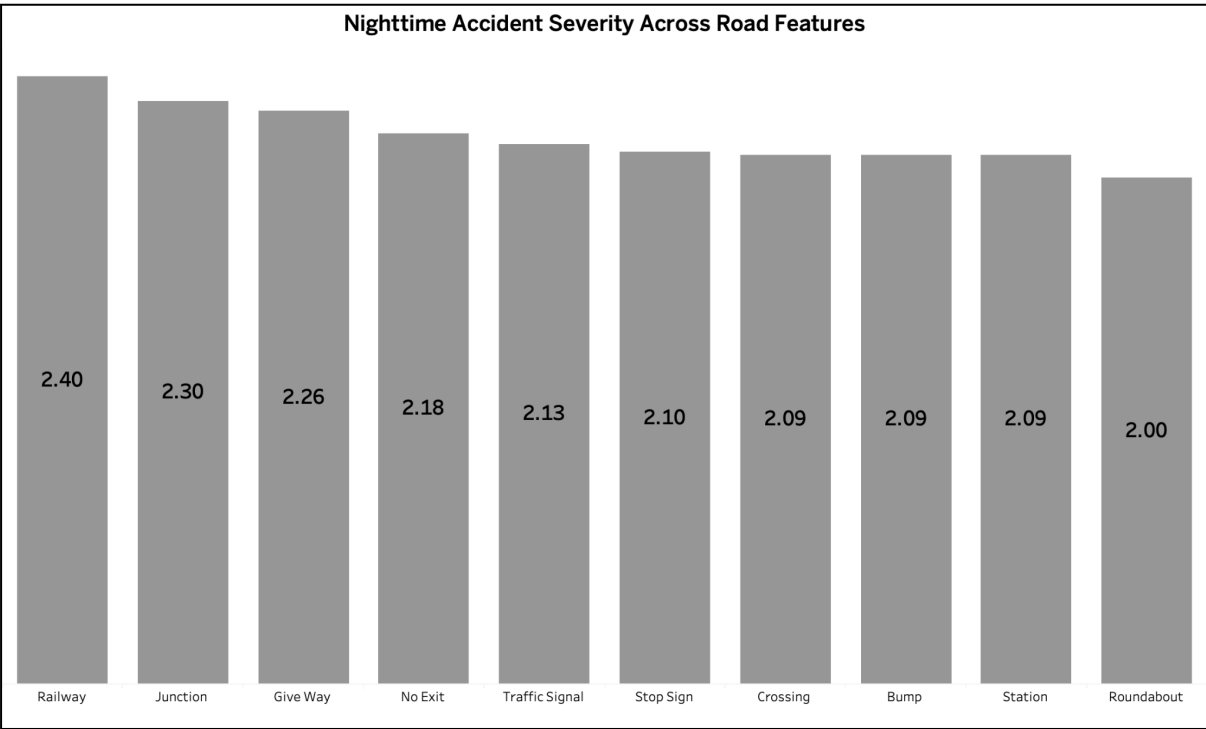
The **average humidity level is nearly identical** in both moderate and severe accidents, with only a slight increase observed in severe cases. This indicates that **humidity alone does not have a strong impact on accident severity**.

11- How many crashes occur at each road feature, and how does accident severity differ across them?



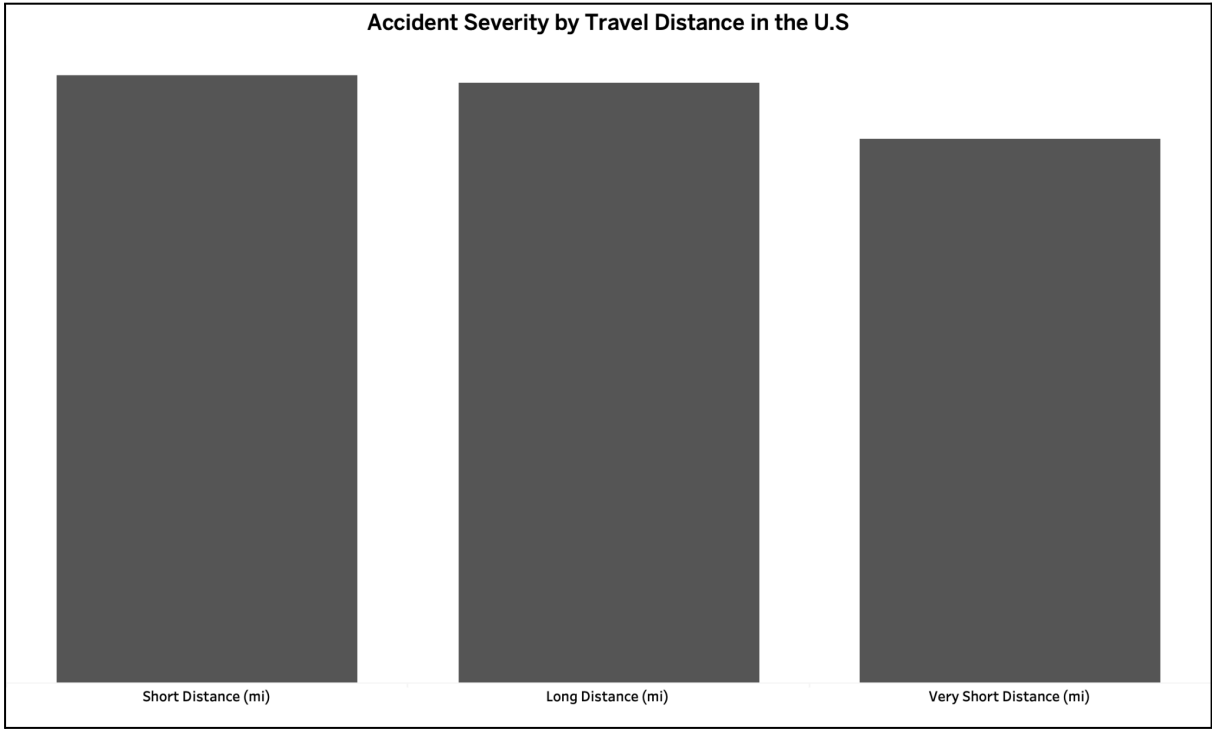
Most accidents occur at **crossing, traffic signals, and junctions**, where vehicle movement overlaps with pedestrian activity. While **crossing and traffic lights** record the highest overall number of accidents, **junctions** show a higher proportion of severe accidents, indicating that they are more dangerous despite having fewer total incidents.

12- Which road features become more dangerous during nighttime driving conditions?



The figure shows that **Railway, Junction, and Giveaway** experience the highest levels of accident severity during nighttime. The danger associated with these road features increases significantly in **low-visibility environments**, where reduced lighting causes slower driver reaction times, making it harder to avoid hazards.

13- How does travel distance influence accident severity in the U.S?



Short and long distance record the highest accident rates, largely due to heavy congestion on urban roads and high-speed conditions on highways. During short distances, drivers often have less control over their speed, which increases the likelihood of accidents in these situations.

Recommendation

1. Improve Infrastructure in High-Risk Areas
2. Launch Awareness Campaigns About Road Safety
3. Send Alerts During Hazardous Weather Conditions
4. Manage Traffic During Rush Hours
5. Improve Nighttime Road Lighting
6. Monitor Roads Using Cameras and Sensor
7. Promote Safe Driving Behavior

Limitation

1. Lack of Accident Type
2. No Clear Cause for Each Accident
3. No Driver Demographic Information
4. Limited Road Information
5. No Vehicle Information
6. Missing data

Reference

<https://www.nhtsa.gov/press-releases/early-estimate-2021-traffic-fatalities>

<https://www.weather.gov/safety/fog>

<https://en.wikipedia.org/wiki/Visibility#:~:text=The%20international%20definition%20of%20fog,be%20of%20smaller%20particle%20size>

Number of Cars

- 2016

<https://www.fhwa.dot.gov/policyinformation/statistics/2016/mv1.cfm>

- 2017

<https://www.fhwa.dot.gov/policyinformation/statistics/2017/mv1.cfm>

- 2018

<https://www.fhwa.dot.gov/policyinformation/statistics/2018/mv1.cfm>

- 2019

<https://www.fhwa.dot.gov/policyinformation/statistics/2019/mv1.cfm>

- 2020

<https://www.fhwa.dot.gov/policyinformation/statistics/2020/mv1.cfm>

- 2021

<https://www.fhwa.dot.gov/policyinformation/statistics/2021/mv1.cfm>

- 2022

<https://www.fhwa.dot.gov/policyinformation/statistics/2022/mv1.cfm>

- 2023

<https://www.fhwa.dot.gov/policyinformation/statistics/2023/mv1.cfm>

Weather

<https://www.visitcalifornia.com/experience/weather-timing-your-visit/#:~:text=Weather%20&%20Seasons,and%20destinations%20you%20are%20considering.>

<https://www.weather.gov/tbw/rainyseason#:~:text=Late%20May%20through%20June%20is,when%20the%20rainy%20season%20peaks.>

https://gov.texas.gov/film/page/agriculture_climate_geography#:~:text=Weather%20Information,Texas%20Ag%20Stats

<https://niceguytours.com/understanding-nycs-weather-what-visitors-can-expect-each-season/#:~:text=expect%20each%20season!-,NYC%20Weather%20Guide,milder%20temperatures%20and%20changing%20foliage>.

<https://chatgpt.com/c/69240aed-10a8-8333-a6bf-d7aa3039f98e>

Taxes

<https://www.pstrialaw.com/legal-news/what-state-has-the-most-car-accidents#:~:text=What%20State%20Has%20the%20Most%20Car%20Crashes%E2%80%94and%20Why%20Is,crash%20rates%20relative%20to%20population>.

Police in Peak Hours

https://international.fhwa.dot.gov/pubs/pl07012/atm_eu07_04.cfm#:~:text=a%20travel%20lane-,Queue%20Warning,focal%20point%20for%20the%20system.