# Research Questions & Answers

## 1. What are the crash severity patterns during weekday rush hours (7–9 AM), and what do they reveal about risk factors?

### Answer:

Weekday rush hours experience **very high crash frequency** (682,879 crashes), but comparatively **moderate injury rates** (40.47%) and **very low fatality rates** (0.12%). Congested traffic leads to slower speeds, reducing crash severity.
**Main risk factors:** heavy traffic density, driver distraction, commuter stress, frequent braking, and stop-and-go movement.

### explanation:

| | |
|---|---|
| **Total Records** | **682,879** |
| **Total Injuries** | **276,392 (40.47%)** |
| **Total Fatalities** | **801 (0.12%)** |

## 2. What are the severity patterns during weekend late-night hours (10 PM–12 AM), and what risk factors explain them?

Answer:

Crash frequency is lower (119,127 records), but **injury rate is extremely high** (63.9%) and **fatality rate increases threefold** (0.40%). These crashes tend to be more violent due to higher speeds and impaired driving.
**Main risk factors:** alcohol use, speeding, fatigue, and poor visibility.

explanation:

| | |
|---|---|
| **Total Records** | **119,127** |
| **Total Injuries** | **76,121 (63.90%)** |
| **Total Fatalities** | **477 (0.40%)** |

## 3. How do crash severity patterns differ between weekdays and weekends?

Answer:

- **Weekdays: High crash volume, lower severity, minimal fatalities.**
- **Weekends: Lower volume but much higher severity and fatality risk.**
  **Conclusion: Weekday congestion suppresses crash severity, while weekend late-night driving is considerably more dangerous**

explaination:

**Weekdays (Rush Hours):**

- **High crash frequency: Morning commutes lead to a large number of crashes.**

- **Moderate injuries (40%): Many crashes result in minor or moderate injuries.**
- **Low fatalities (0.12%): Slower traffic speeds due to congestion reduce risk of death.**
- **Risk factors: Traffic density, driver distraction, stress, stop-and-go conditions.**

**Weekends (Late Nights):**

- **Lower crash frequency: Fewer vehicles on the road late at night.**
- **High injuries (64%): Crashes are more severe, leading to a higher proportion of injuries.**
- **Higher fatalities (0.4%): Although absolute numbers are smaller, the risk of death is significantly higher.**
- **Risk factors: Alcohol/substance use, speeding, fatigue, reduced visibility.**

## 4.Is there evidence of seasonal variation in *motorist-involved* crashes? Which months show the highest motorist fatality rates, and how might weather, road conditions, and daylight patterns explain these seasonal differences?

### Observation:

Crash data shows that the winter months (December–February) have significantly fewer motorist-involved injuries. This decline is most likely driven by reduced traffic volumes during the holiday season, school closures, and generally lower levels of travel. In contrast, the remaining months—spring, summer, and fall—show a marked increase in traffic volume, which corresponds with higher crash counts.

### Conclusion:

**There is clear evidence of seasonal variation in motorist-involved crashes. Winter months consistently experience lower crash frequencies, likely due to:**

- **Reduced traffic volume (fewer daily commuters and fewer long-distance trips).**
- **More cautious driving behavior in response to cold or wet weather conditions.**
- **Overall fewer vehicles are exposed to crash risk.**

**These findings highlight the role of seasonal travel patterns and weather-related factors in shaping motorist crash trends throughout the year.**

# Spatial Risk Analysis

## 5. Which borough has the lowest fatality rate, and what does this indicate about severity vs. frequency?

### Answer:

***Manhattan*** *has the lowest fatality rate despite having high crash frequency.*

***Interpretation:*** *Dense traffic reduces speeds, lowering fatal crash likelihood. High frequency ≠ high severity.*

## 6. Are cyclist crashes geographically concentrated, and where are the hotspots?

### Answer:

*Yes. Cyclist crashes are concentrated mainly in* ***Brooklyn and Manhattan****, which also have the highest cycling activity and infrastructure density.*

***Severity rates****, however, remain similar across boroughs due to consistent vulnerability of cyclists.*

# Vehicle Type & Risk

## 7. How do injury rates differ across vehicle types?

Answer:

- Highest injury rates (75–99%): motorcycles, bicycles, e-bikes, scooters, mopeds.
- Moderate rates (46–48%): sedans, SUVs, passenger vehicles.
- Lowest rates (18–34%): heavy trucks, commercial vehicles.
  Explanation: Vulnerable road users lack physical protection, while larger vehicles protect occupants but can be severe to others.

# Contributing Factor Insights

## 8. Among fatal crashes, what percentage involves driver behavior factors (aggressive driving, distraction, alcohol) vs. environmental factors (weather, road conditions)? How does this inform prevention strategies?

**Answer:**

The majority of fatal crashes involve **driver behavior**, including:

- Driver inattention/distraction (Vehicle 2)
- Unsafe speed (Vehicle 1)
  Environmental factors like weather or road conditions are much less frequent contributors.
  **Conclusion:** Behavioral enforcement is the most effective prevention strategy.

# Trend Analysis & Prediction

## 9. What hour–day combinations should receive traffic enforcement priority to reduce injuries?

### Answer:

The highest-risk windows are:

- Weekdays 8–9 AM
- Weekdays 2–6 PM
- Absolute peak: Friday at 5 PM
  These periods align with commuting surges and elevated driver fatigue and distraction.

## 10. Has there been a statistically significant change in the bicycle crash rate from 2015 to 2023, and can we identify turning points related to NYC safety initiatives?

### Answer :

When filtering for bicycle crashes, the charts show normal values from 2012–2015 but almost no records from 2016–2023. This drop does not reflect a real decline in bicycle crashes—it is caused by missing or inconsistently recorded bicycle vehicle-type data after 2015.
This missingness can be understood in two ways:

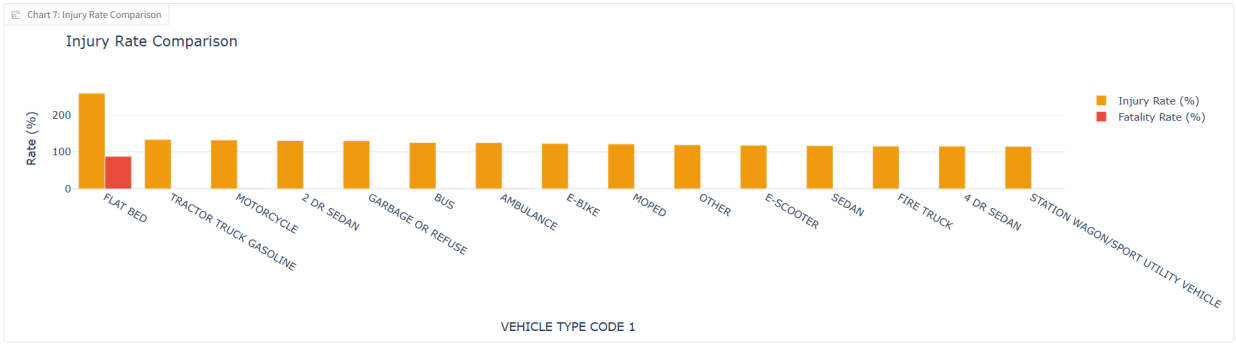
- Ignorable Missingness:
  The missing values are unrelated to the true number of bicycle crashes. The crashes still happened, but the “vehicle type = bicycle” field simply wasn’t recorded.
- Non-Ignorable Missingness:
  The missing values depend on changes in reporting (e.g., schema changes, integration issues). Bicycle crashes are missing because the reporting method changed, not because crashes decreased.

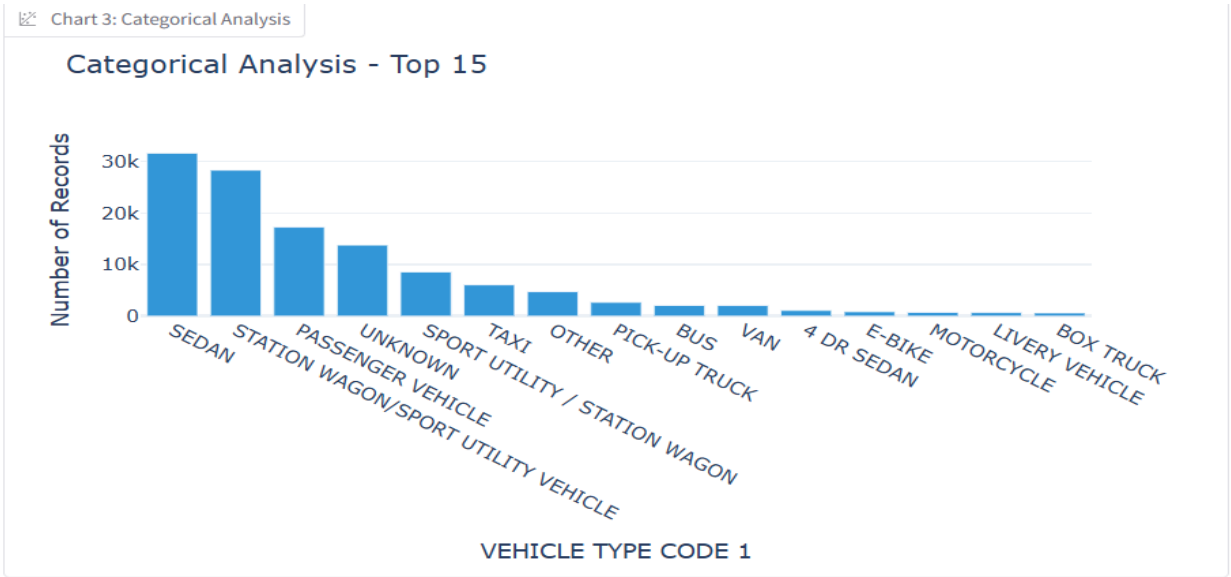
Conclusion:
Because the missing data after 2015 is likely non-ignorable, we cannot evaluate real trends or identify safety-initiative turning points for bicycle crashes between 2015–2023.

# 11.Compare the effect of different vehicle types on injured pedestrians

## Answer

Sedans and SUVs cause the highest number of pedestrian injuries simply because they are the most common vehicles on the road. However, heavier vehicles like flatbeds, trucks, and buses cause more severe injuries per crash due to their larger size and impact force.
Conclusion: Cars injure the most pedestrians, but heavy vehicles are the most dangerous when a collision occurs.



## 12.Do injury rates differ between male, female, and unknown-gender individuals involved in motor vehicle collisions?

## Answer:

Males make up the largest share of crash-involved individuals, followed by females and then unknown gender records. However, when looking at injury rate percentages, the Unknown category shows the highest injury rate (~63%), followed by females (~54%) and males (~47%). This suggests that missing or unclear gender entries may be associated with more severe crash circumstances, while males, despite being the most frequently involved have a lower injury percentage relative to their exposure.