# A study of Road Safety

**Some questions to consider:**

1. **What day was of the week was where the most crashes happen?**
2. **Is there a statistically significant relationship between the time of day and the types of vehicles involved in crashes?**
3. **Do the relative proportions of crashes resulting in injuries or fatalities vary significantly across the five boroughs depending on the time of day?**
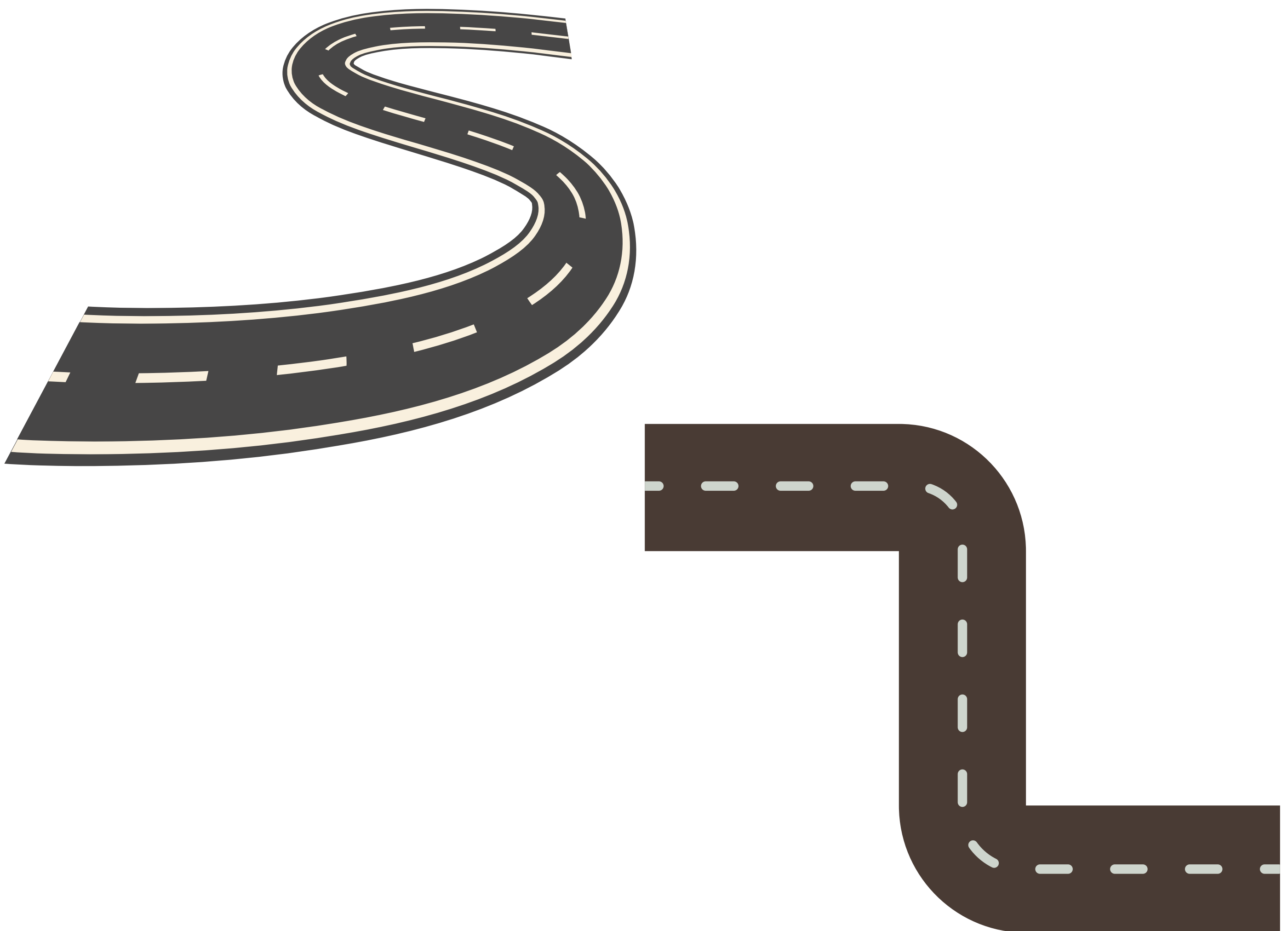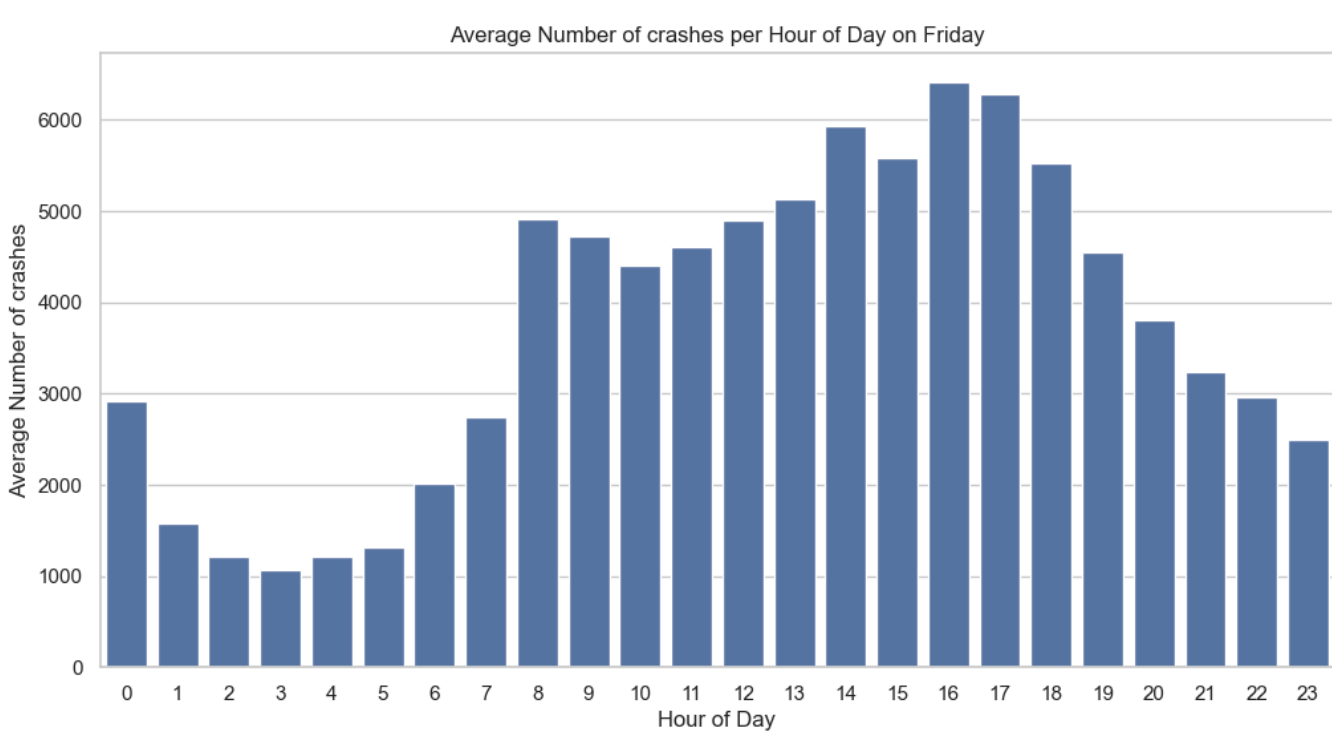
# Q: Is there a statistically significant relationship between the time of day and the types of vehicles involved in crashes?

Before answer this question lets see what day the most car crashes happen:



Number of Crashes by Day of the Week
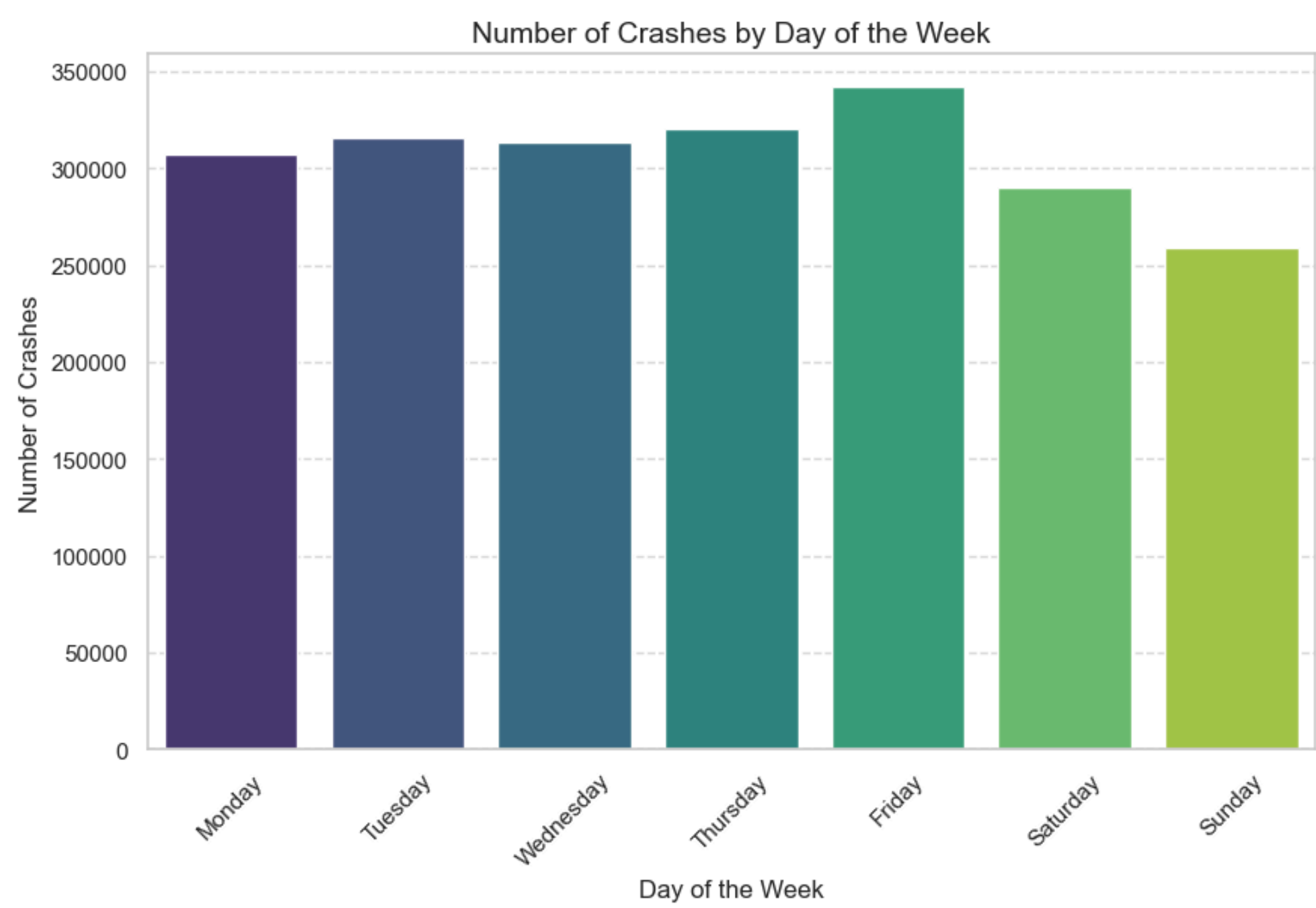
Based on the first question, we can use it to answer the following questions mentioned below. To address these questions, I will focu more in-depth on Friday, as there are more ca crashes reported on that day.

So knowing this we can now focus on Friday to answer both main questions, but first lets see how the number of crashes reflect the time of day:



Average Number of crashes per Hour of Day on Friday

By looking at the time of day we can look at the statistical and the visualization to answer this main question as shown in the python notebook on answer this potential question, we can say that:

There is a statistically significant relationship between the time of day and the types of vehicles involved in crashes, based on Statistical test (Chi-squared test)
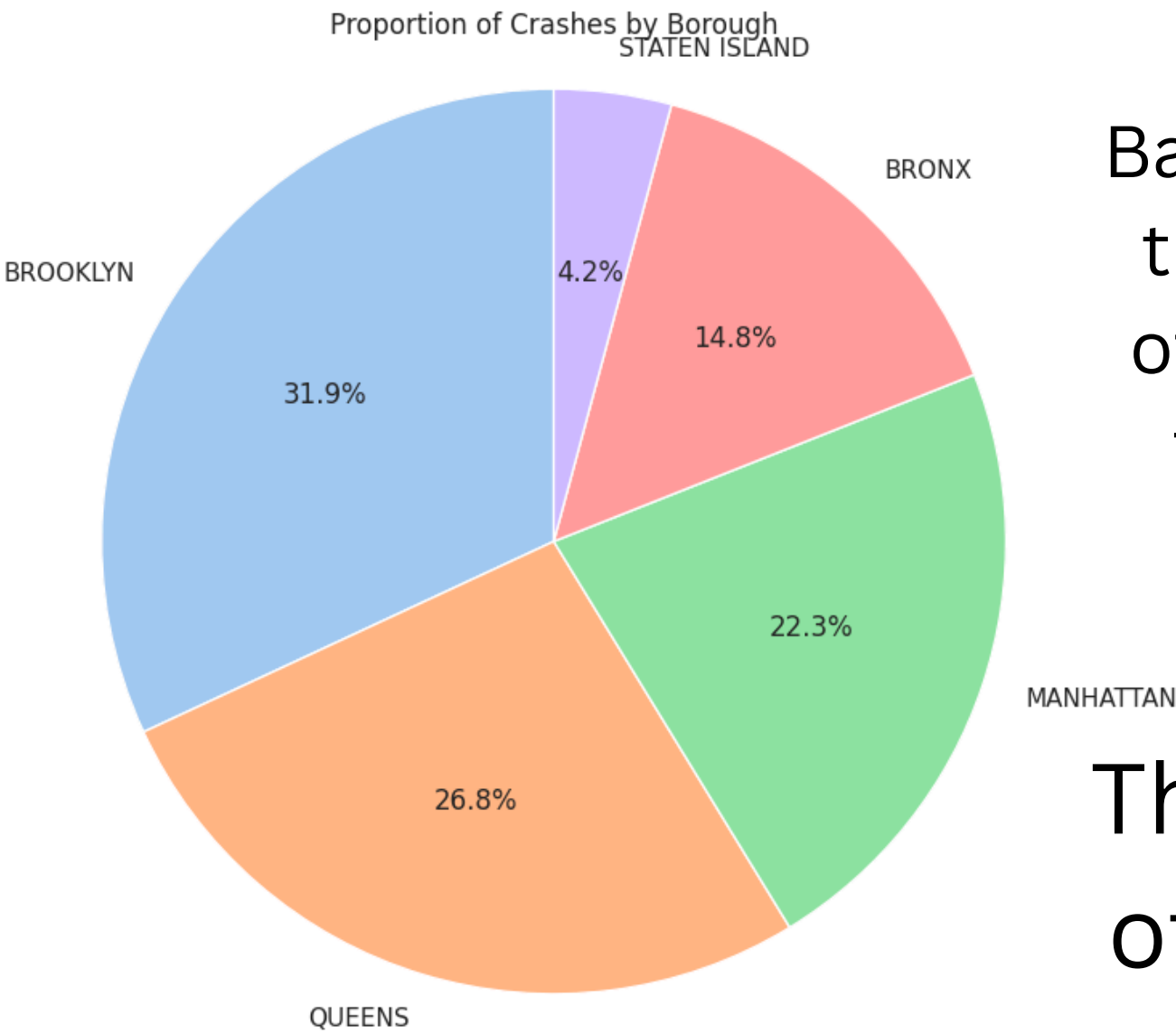
The following two graphs show the severity of the top 10 types of vehicles that were in crashes:



Top 10 Vehicles Involved in Crashes (Fatalities)



Top 10 Vehicles Involved in Crashes (Severity)
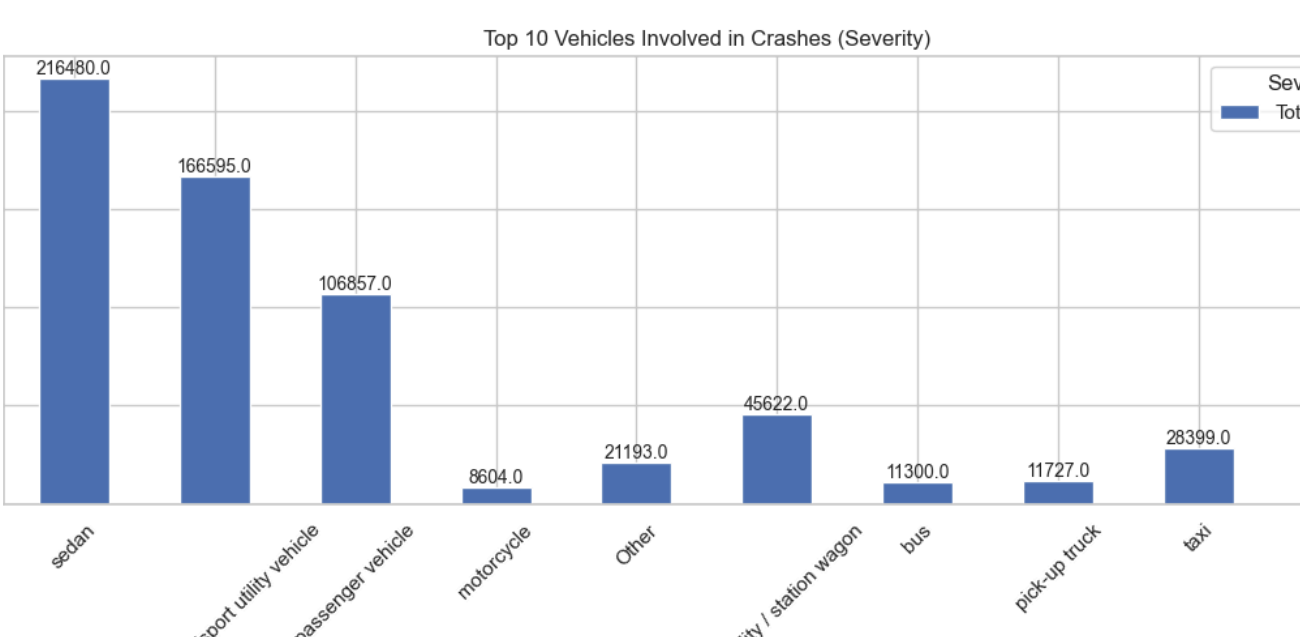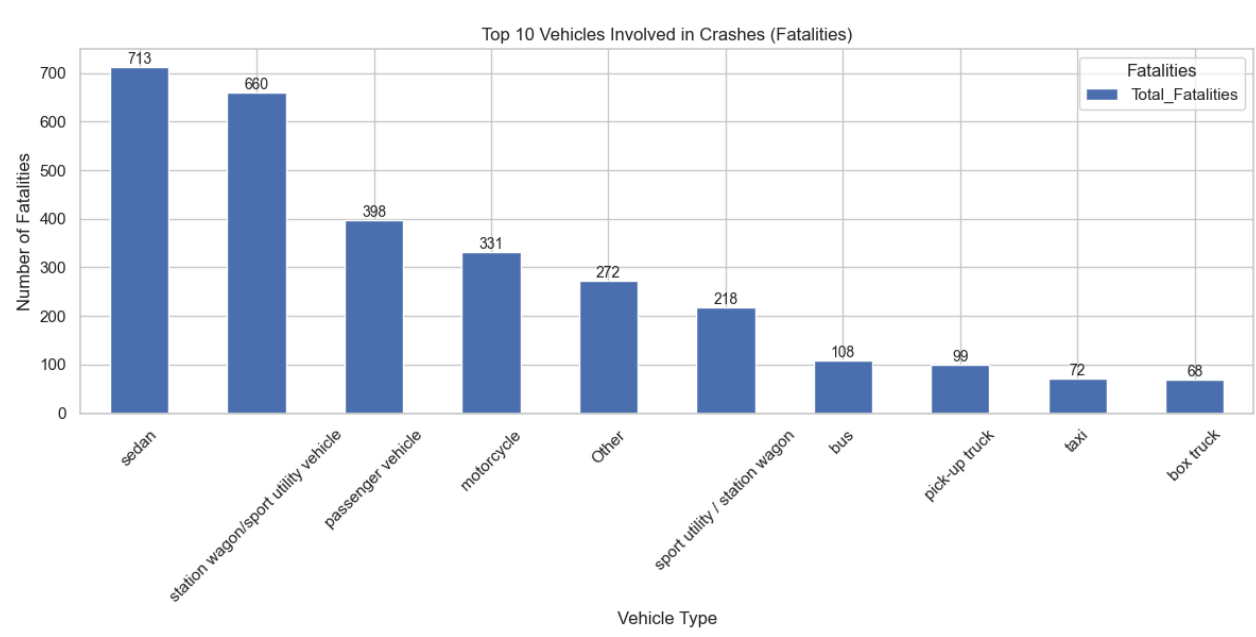
Based on these two graphs, we can observe the severity of each car crash. In both graphs, sedans have the highest severity levels for both Total_Injuries and Total_Fatalities.

# Q: Do the relative proportions of crashes resulting in injuries or fatalities vary significantly across the five boroughs depending on the time of day?

Before answer this main question we got to see where the most occurrence happen:

**Proportion of Crashes by Borough**



Based on this pie chart, we can see that Brooklyn has the highest proportion of crashes compared to the other boroughs, with a leading 5.1% difference over the second-highest percentage, which is Queens.

The following five graphs show the severity of the car crashes in the time of day based on location:



**Crash Severity in BROOKLYN by Hour of Day**







**Crash Severity in STATEN ISLAND by Hour of Day**
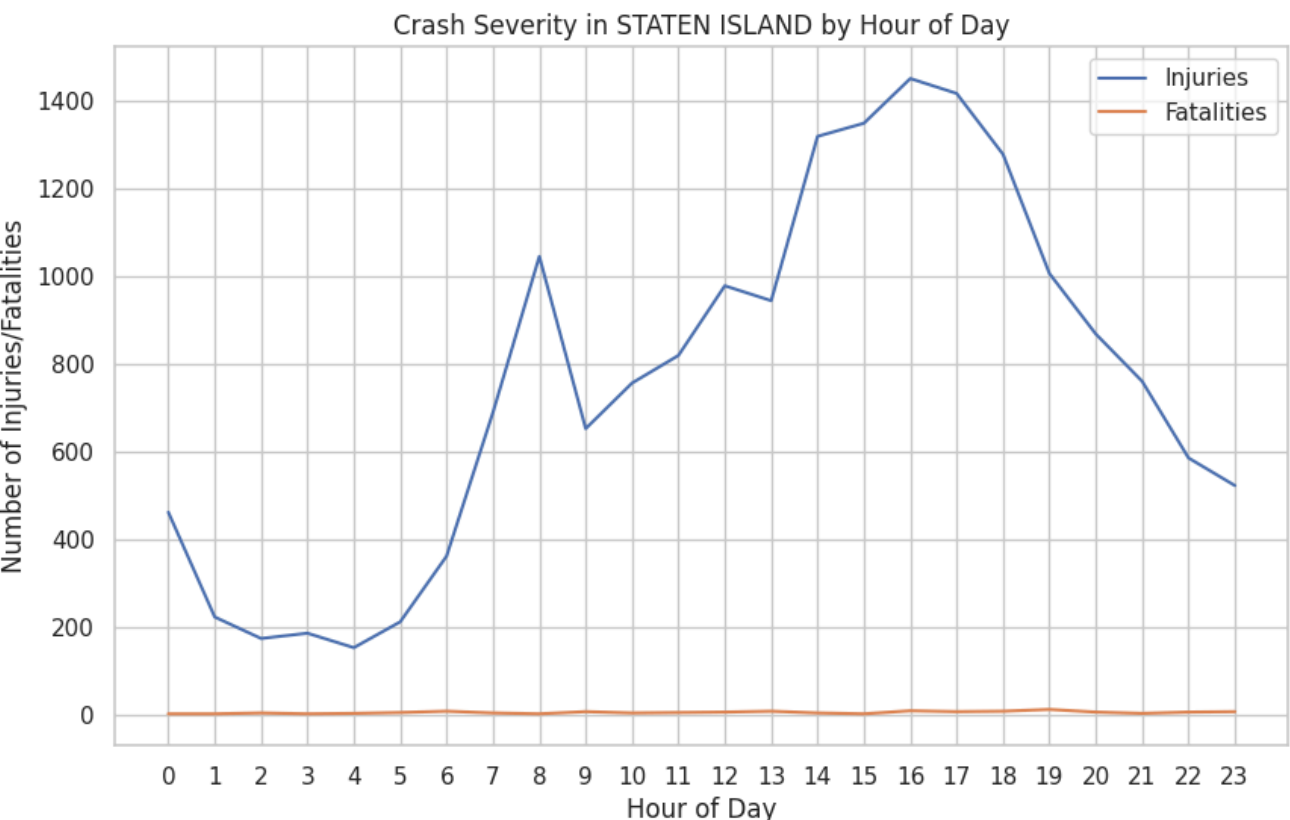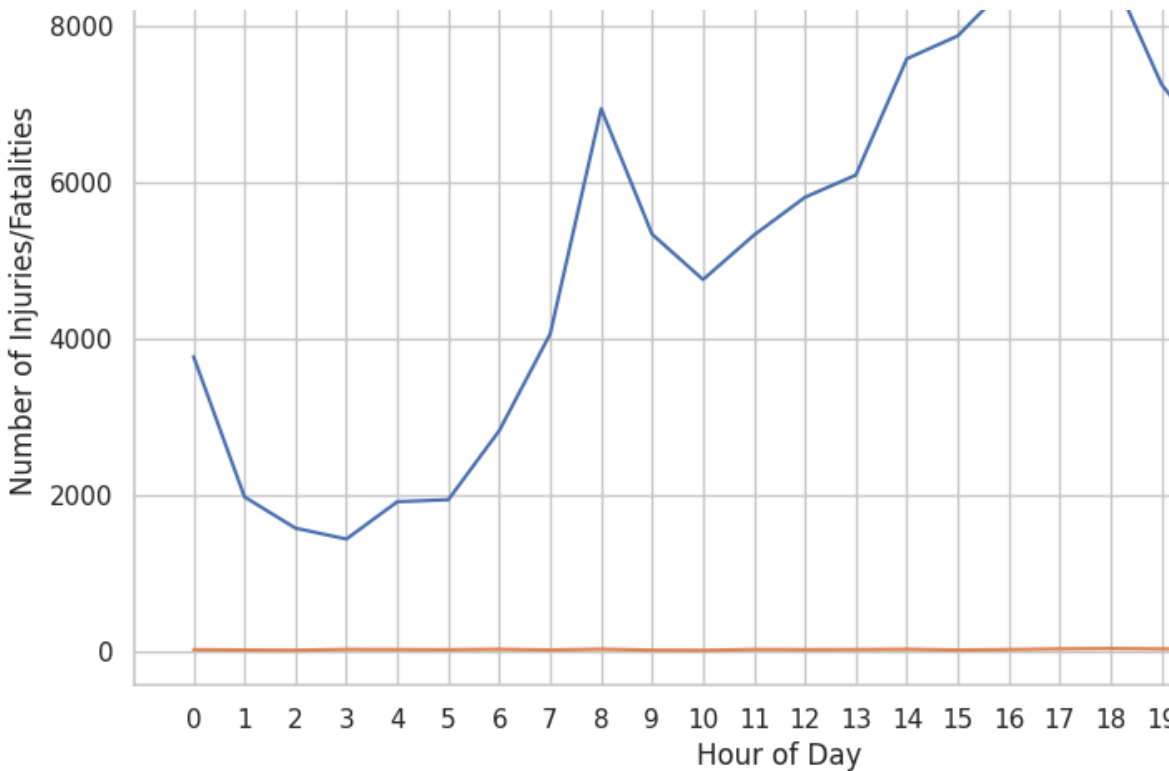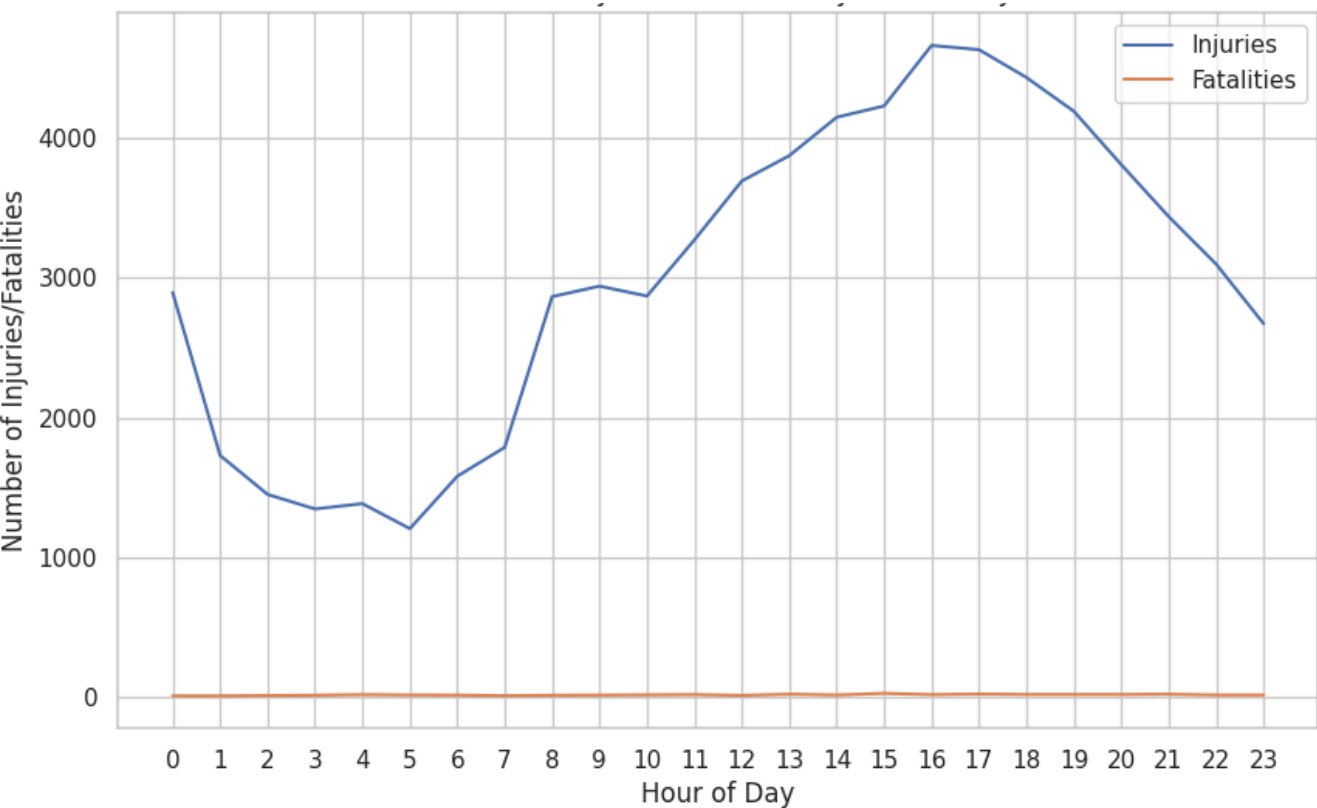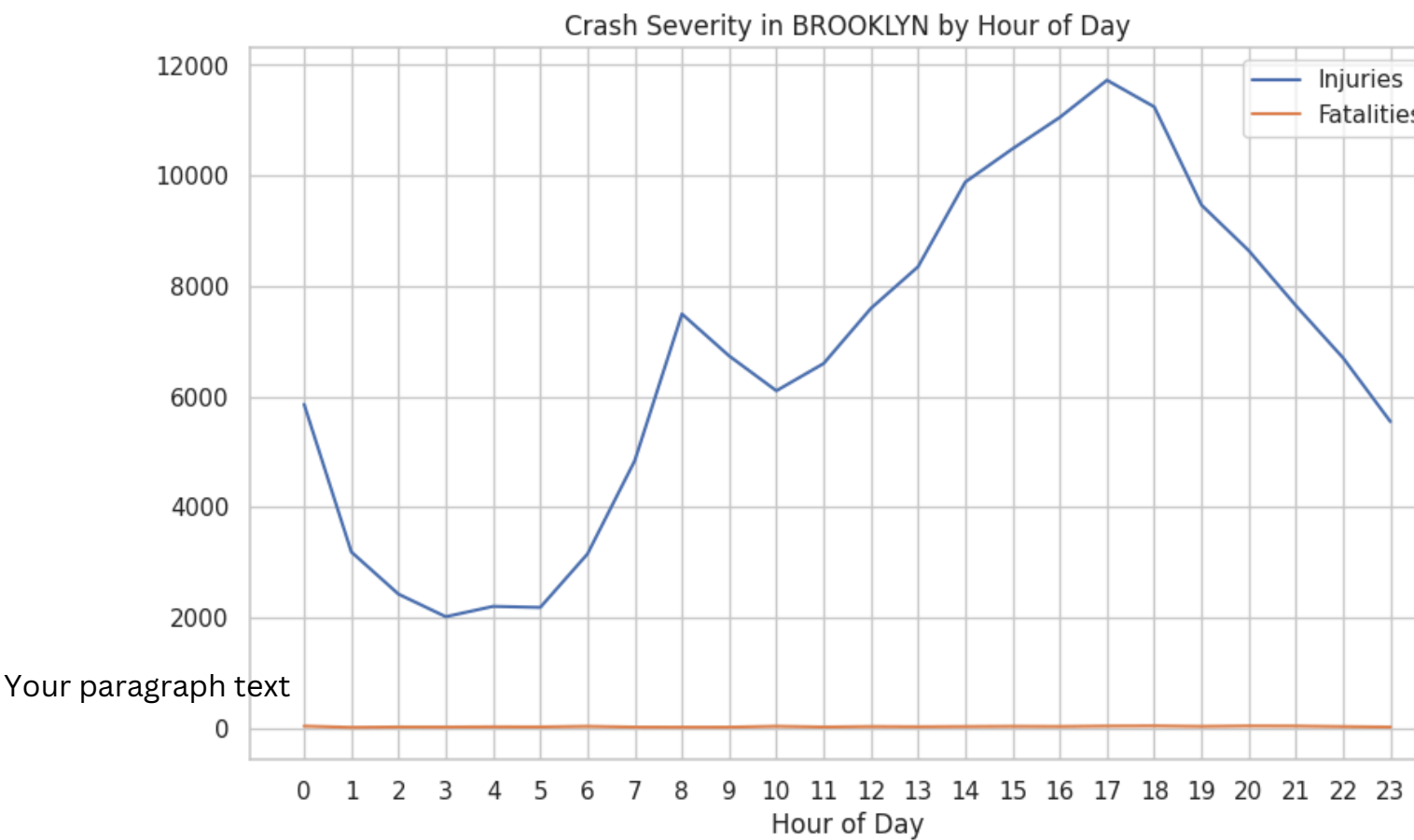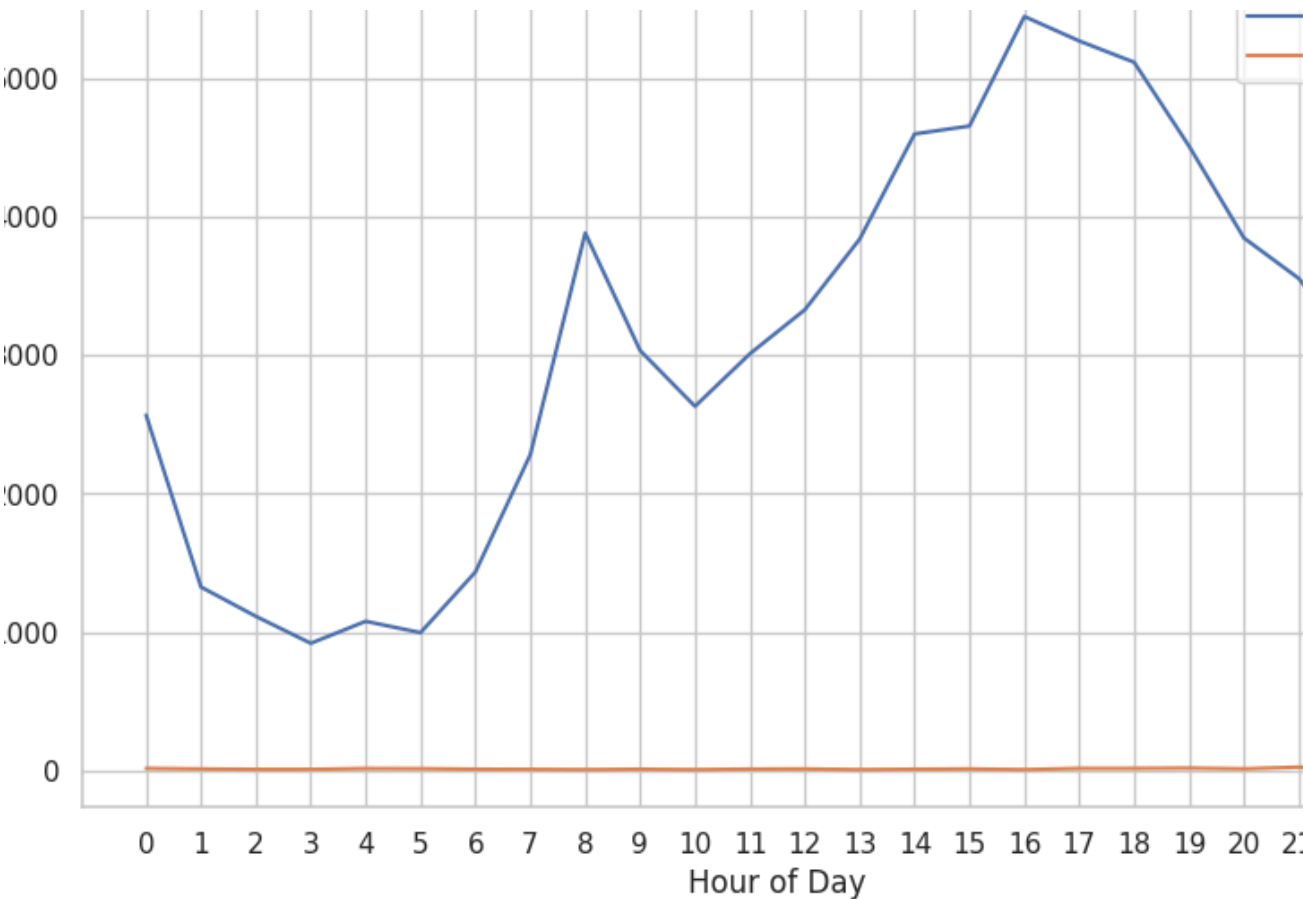


We analyzed two central questions:
1. Is there a relationship between the time of day and the types of vehicles involved in crashes?
2. Do crash proportions vary across boroughs based on time of day?

The data shows a significant link between the time of day and vehicle types involved, with most crashes occurring on Friday afternoons, likely due to fatigue. Brooklyn stands out as a hotspot, with crash severity differing across boroughs and times.

Further investigation could explore whether weather conditions influence crash frequency and timing.

# Sources:

- NSDC Flashcards: Geospatial Analysis, <u>by Northeast Big Data Innovation Hub</u> https://www.youtube.com/playlist?list=PLNs9ZO9jGtUAX_2g1-OJp8VkmVum6DqqP

- NYC OpenData, Motor Vehicle Collisions - Crashes, https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95/about_data

- NSDC Flashcards: Data Science Ethics, <u>by Northeast Big Data Innovation Hub</u>, https://www.youtube.com/playlist?list=PLNs9ZO9jGtUB7XTjXy-ttoo2QSLld9SrV

- Data Ethics, <u>by Moray House School of Education and Sport</u>, https://www.youtube.com/watch?v=Efy8htCDueE