* 1. OSMnx Network Map

In this section, we wish to examine the crash data through street network on Manhattan to explore spatial patterns of crashes in certain area that might cause by congestions. We used the Motor Vehicle Collisions data from NYC Open data. Each row of the data contains one crash event, with information like location, time, causes, types of vehicles etc. We filtered the time from May 2021 to Dec 2023, in accordance with the date for traffic data in machine learning section.

(graphic)

Through the street network map of crash index, it is obvious to identify areas with high risk of crashes. From the map, areas near Chinatown (Lower East) and Midtown West.

1.2 Crash Density Map by time of the day

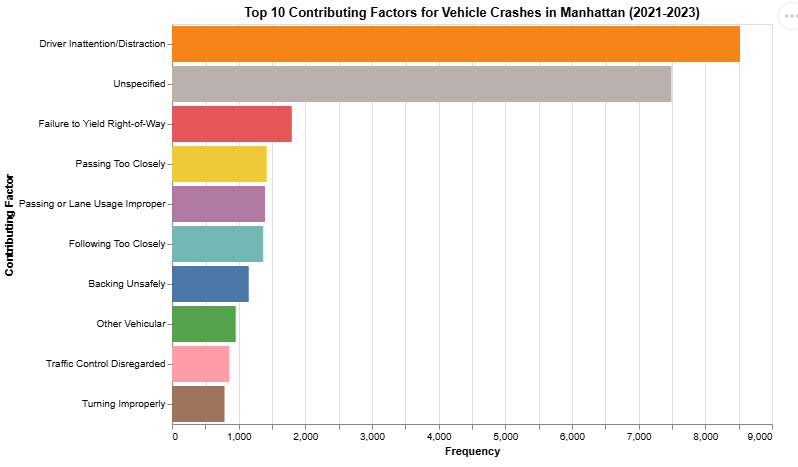
In the following data analysis, we dive deeper into the crash density in different time of the day to see how crash density varies. We choose larger dataset of the crash data ranging from July 2012 to Dec 2024, with 2.14 million rows.

(graph 2 and 3)

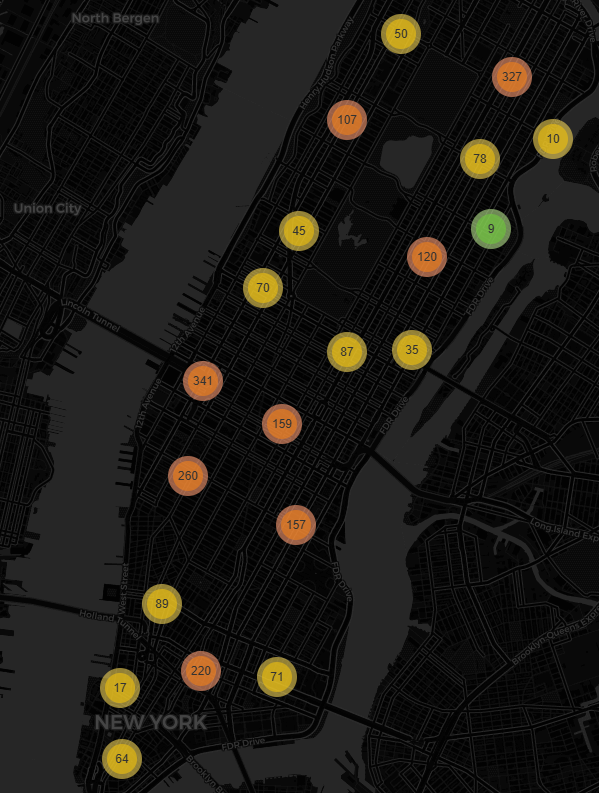
These two graphs follow the similar spatial pattern as the street network for crashes, that most crashes are concentrated in lower east side and midtown west. As for the time of the day, rush hours (8-9 a.m. or 5-6 p.m.) suggest higher density of crashes in total.

…

* 1. cluster map of crashes caused by passing/following too closely

Our research would also like to focus on crashes related to traffic congestion. In the chart below, which are top 10 causes of traffic crashes, passing too closely and following too closely add up to a large number. These two causes are mostly likely to associate with congestion. Therefore, we wish to have a look at the spatial pattern of crashes caused by these two factors as well.

To visualize them, we used the cluster concentration plot to visualize the cluster of these crashes. In the map below, there is also a high concentration of such crashes in lower east near Chinatown and SOHO, and midtown west near Time Square and Lincoln Tunnel. Through such information, we assume these areas in Manhattan may experience higher congestions than others due to large number of crashes caused by following or passing too closely.



## <span style="color: purple;">Intruduction</span>

The purpose of this project is to design a predictive model for traffic congestion using a set of environmental and contextual features, including temperature, precipitation, wind speed, the occurrence of events, and whether it is a weekend. Traffic congestion is a significant issue in urban areas, impacting commute times, air quality, and overall productivity. By leveraging these variables, the model aims to understand the factors influencing traffic patterns and provide accurate predictions of traffic counts. Such a model could be instrumental in improving traffic management systems, informing infrastructure planning, and helping commuters make more informed decisions. The project seeks to demonstrate the feasibility of using readily available data to address real-world urban challenges.

A traffic prediction model has significant potential applications in optimizing traffic light systems to improve traffic flow and reduce congestion. By accurately predicting traffic counts based on environmental factors, events, and time-related variables, the model could serve as a critical input for adaptive traffic light control systems. For instance, the model could help dynamically adjust traffic light timings based on anticipated traffic volumes at specific intersections. During periods of high predicted traffic, longer green light durations could be allocated to heavily congested routes, while during low-traffic periods, shorter cycles could minimize unnecessary delays. This would ensure a more efficient allocation of green time, reducing wait times, fuel consumption, and emissions caused by idling vehicles.

Additionally, the model could be integrated into intelligent traffic management systems that coordinate traffic lights across multiple intersections. By predicting traffic patterns in advance, the system could optimize signal synchronization to create “green waves,” allowing vehicles to travel through a series of intersections without stopping. This approach could be particularly useful in urban areas with high traffic density, where poor signal coordination often exacerbates congestion. Furthermore, during special events or adverse weather conditions, the model could help traffic authorities proactively adjust signal timings to handle anticipated surges in traffic, minimizing disruptions. Overall, the integration of traffic prediction models into traffic light optimization systems has the potential to enhance urban mobility, reduce congestion, and improve the overall efficiency of transportation networks.