

# Spatial Analysis of Traffic Accidents in New York City: Identifying Hotspots and Contributing Factors

Muhammad Amin Yahya (BSCS21049)

Session: Spring 2024

Supervisor: Dr. Adnan Siddique

Course: Spatial Data Science

## 1. Introduction

- Traffic accidents are a significant public safety concern in urban areas like New York City.
- NYC, the most populated city in the US, experiences some of the worst traffic, ranking 3rd nationally.
- In 2018, NYC reported 228,047 traffic accidents, including 44,300 injury-causing collisions.
- Traffic accidents result in severe economic and social impacts, including medical costs, property damage, and stress.

## 2. Problem Statement

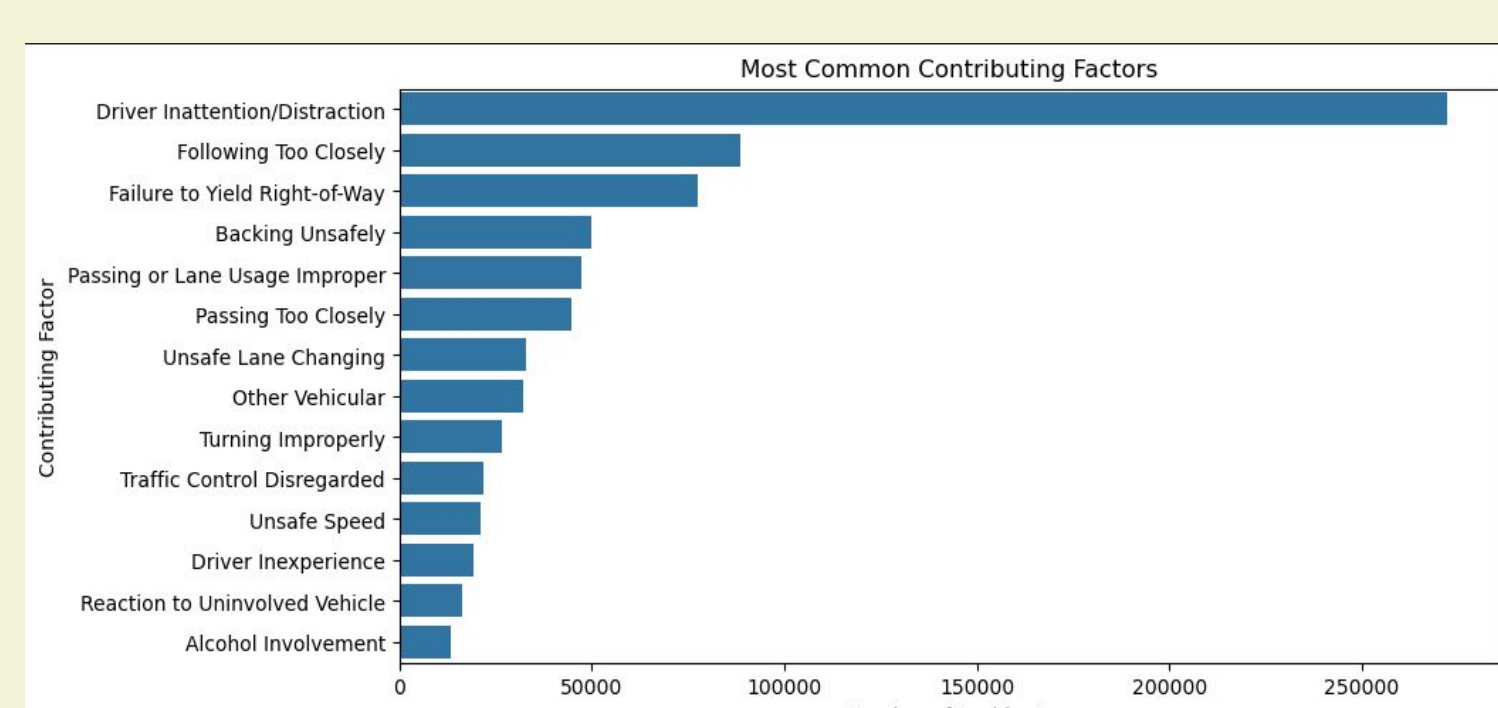
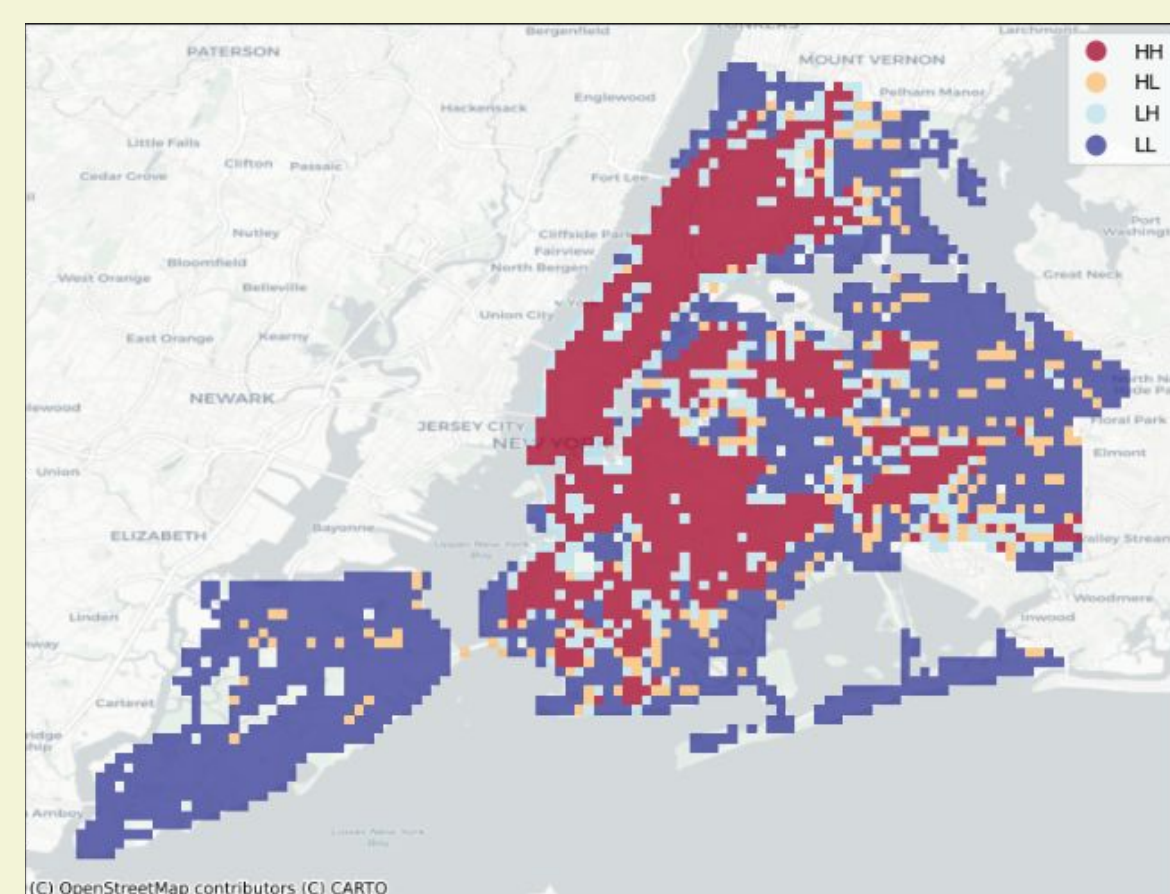
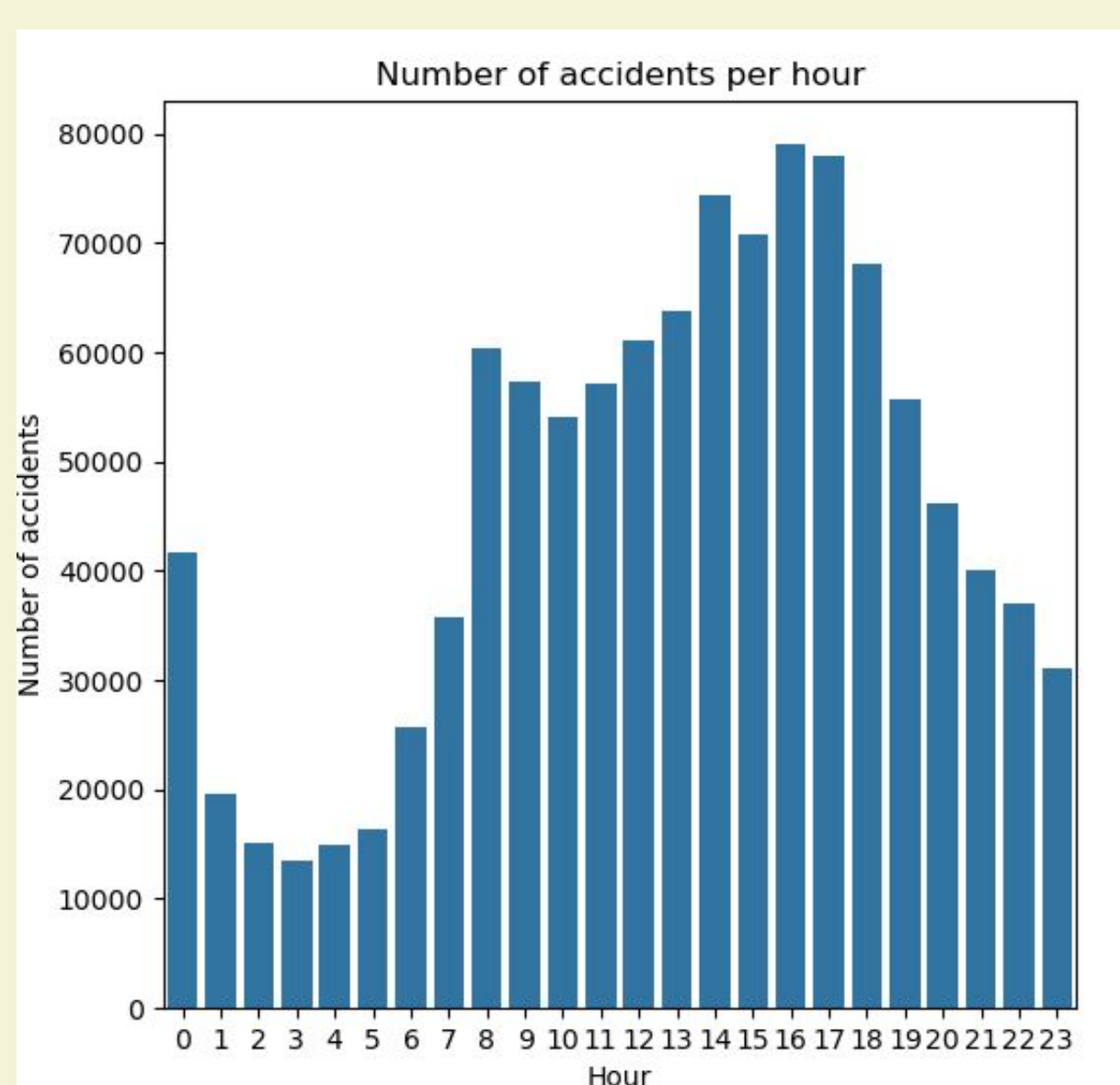
- This study aims to conduct a detailed spatial analysis of traffic accidents in New York City from 2016 to 2022. By identifying high-risk areas and contributing factors, we seek to provide actionable insights for enhancing road safety and reducing the incidence of traffic accidents.

## 3. Data

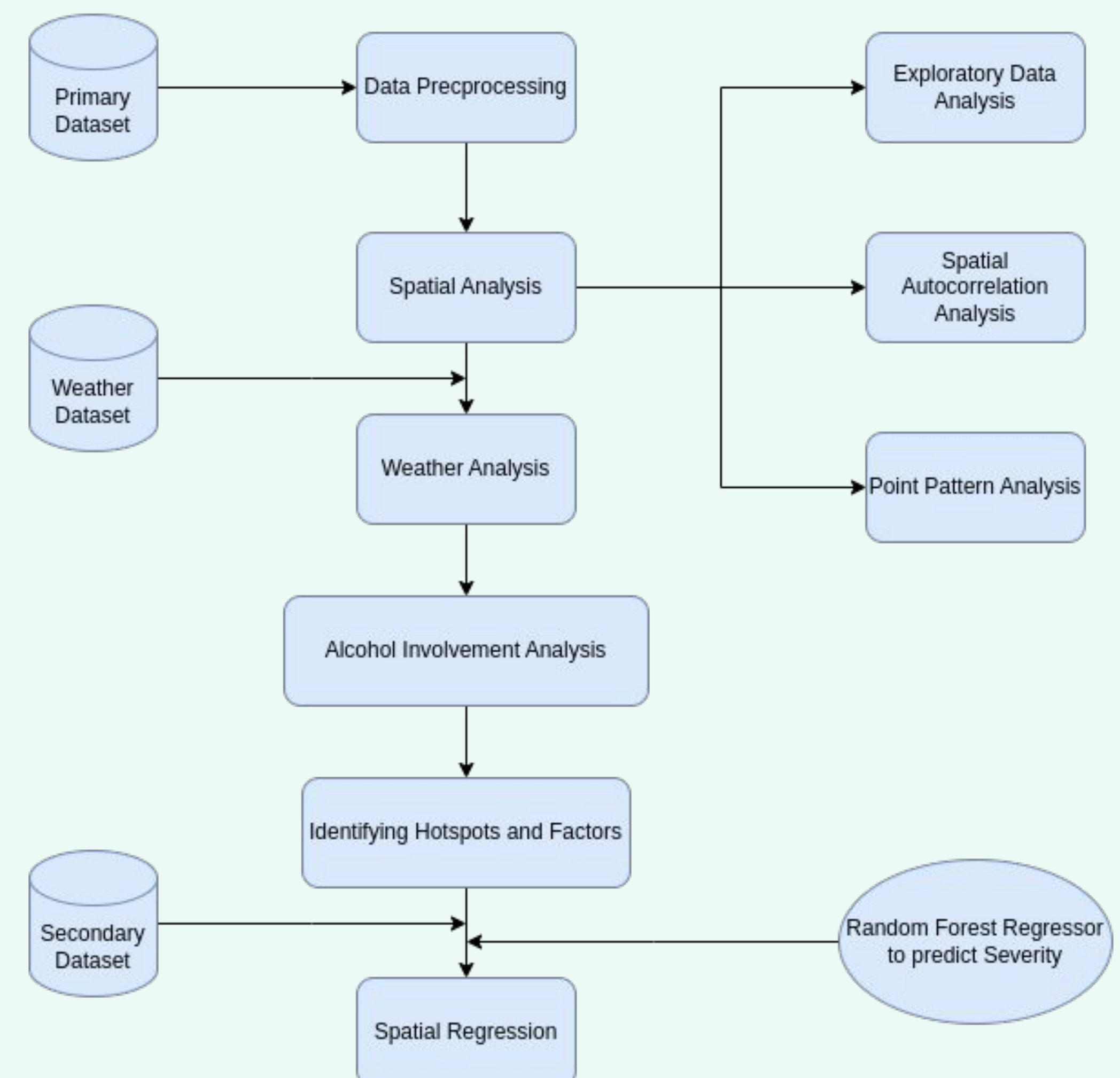
- Dataset retrieved from official US government repository, "Motor Vehicle Collisions - Crashes," comprises over 1M incidents spanning 2016 to 2022
- Contains geographical data and contributing factors
- Supplementary weather and secondary traffic crash data obtained from Kaggle provides annotations for nearby Points of Interest (POI)

### Data Preprocessing involved:

- Dropping irrelevant columns such as vehicle code and cyclist injuries
- Using spatial interpolation to address missing weather data
- Imputing missing numerical data with mean and missing string data with mode
- Filtering out data points beyond NYC's geographical boundary

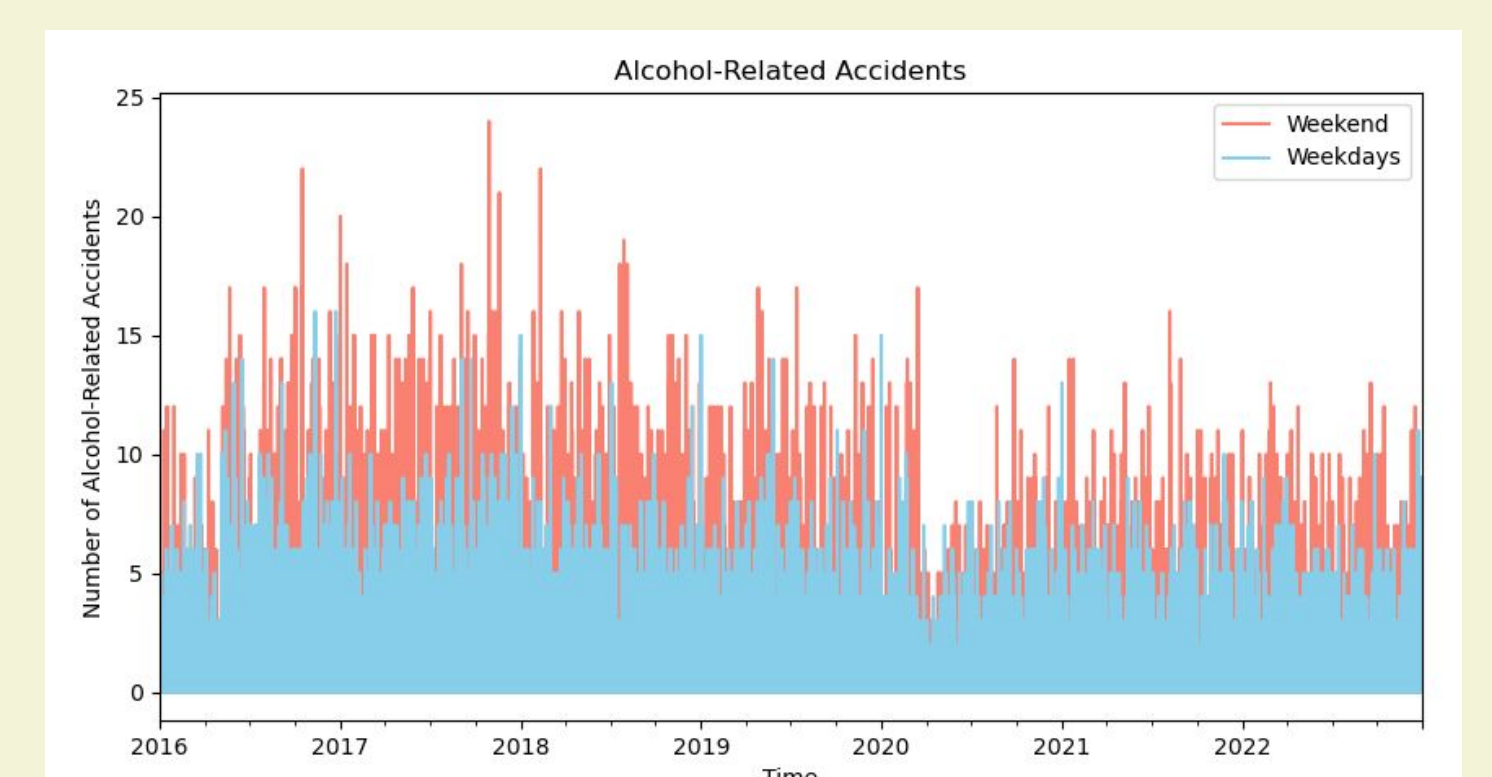
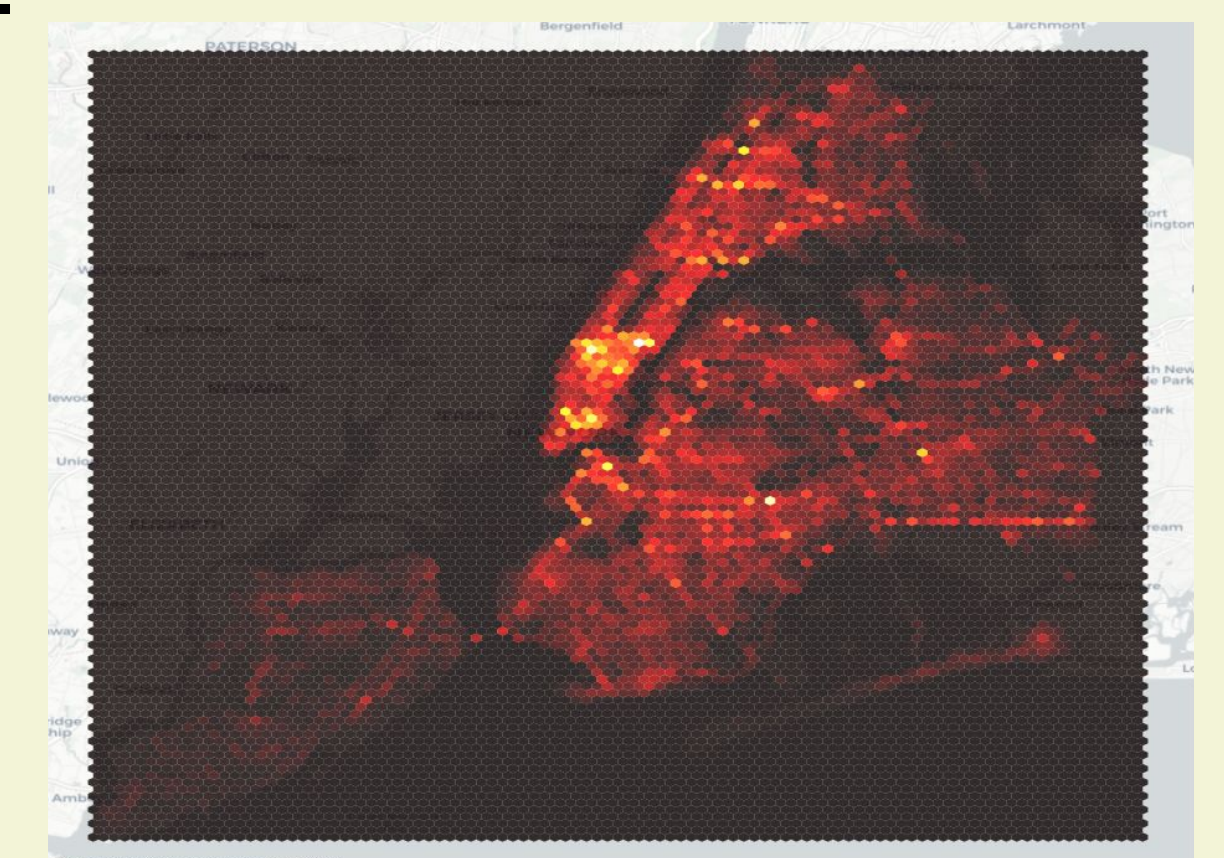
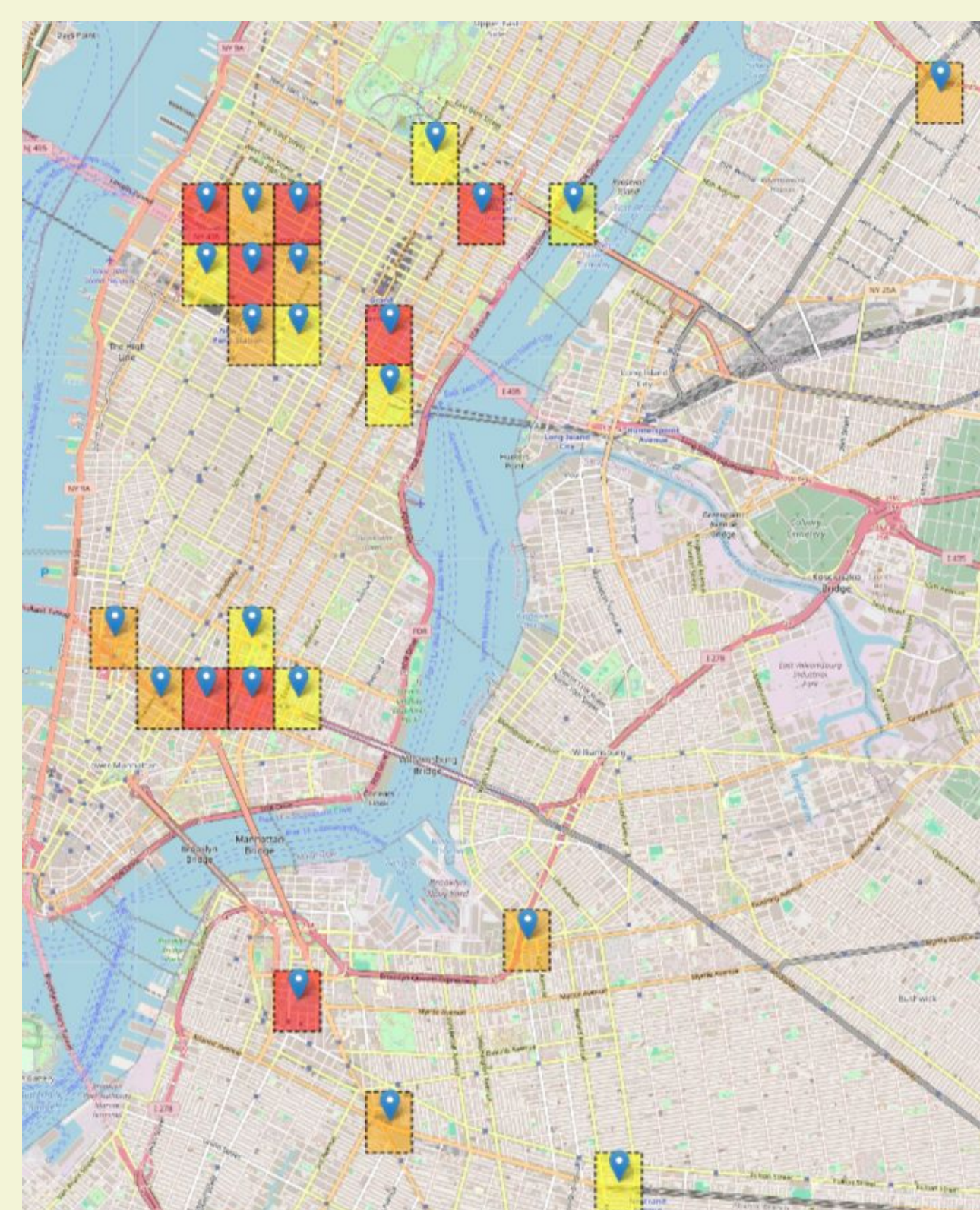


## 4. Methodology



## 5. Results & Discussion

- Traffic crashes in NYC showed an upward trend from 2016 to 2019 but sharply declined with the onset of Covid-19 in late 2019.
- The peak hours for accidents are between **2-6 pm**, coinciding with office and school dismissal times,
- Driver inattention or distraction is identified as the leading cause of most traffic accidents in the city.
- Strong positive spatial autocorrelation was observed indicating traffic crashes tend to aggregate in specific locations across the city.
- Rainy conditions marginally increase accidents by **3.6%**, while alcohol-related crashes surge by **122%** on weekends.
- Hotspots for accidents are concentrated in Manhattan, particularly along the **1st to 10th Avenue** corridor.



Acknowledgments: [Data.gov Motor Vehicle Collisions - Crashes](#), [Weather Data](#), [Secondary Crash Dataset](#)