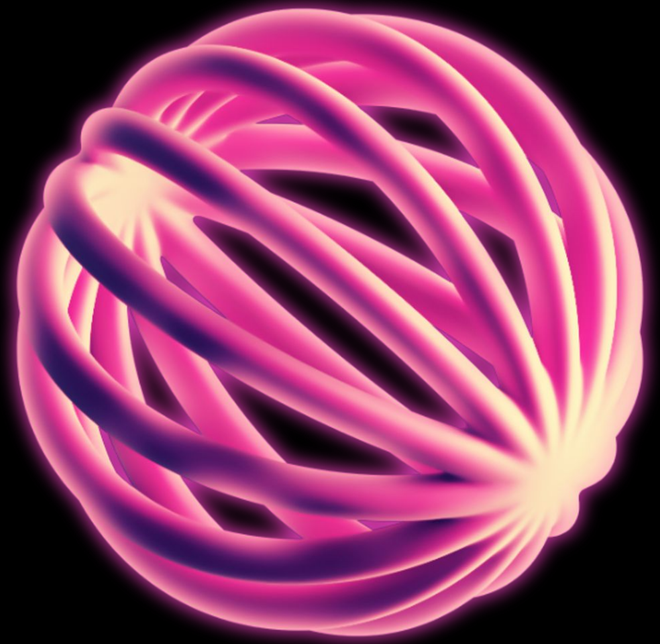
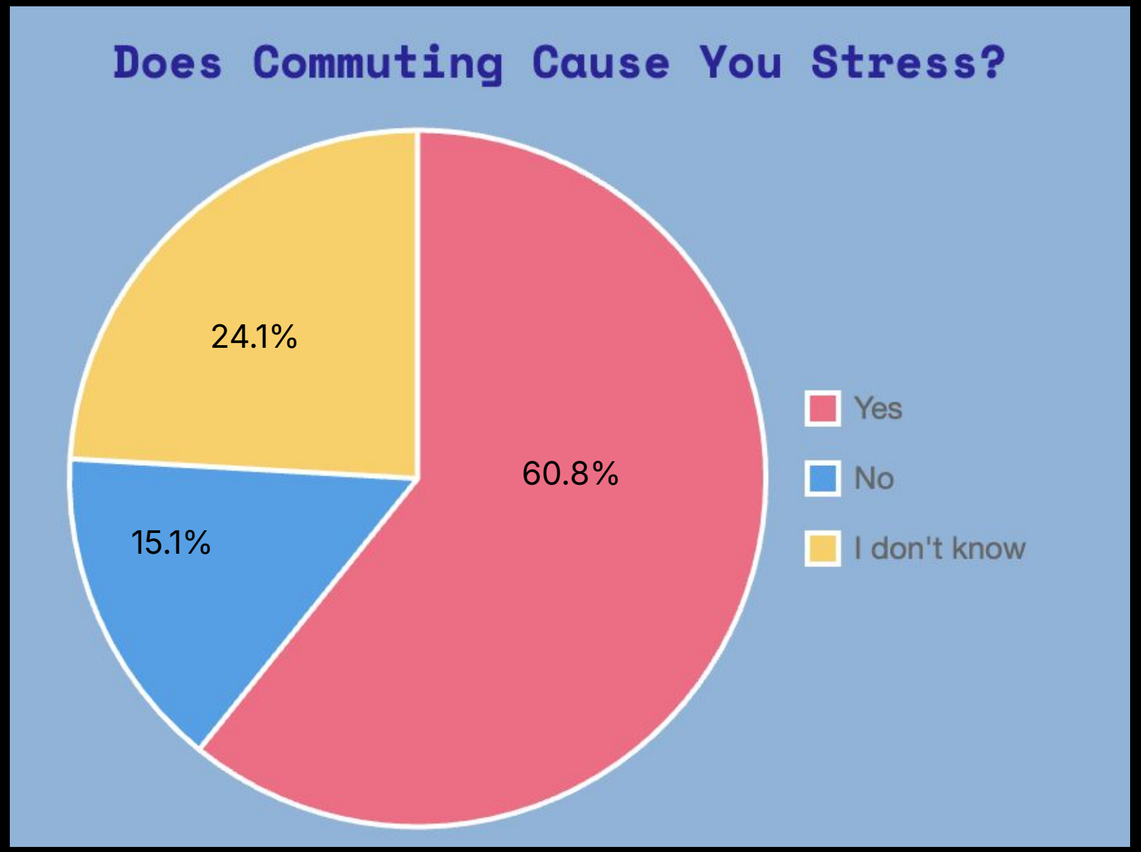


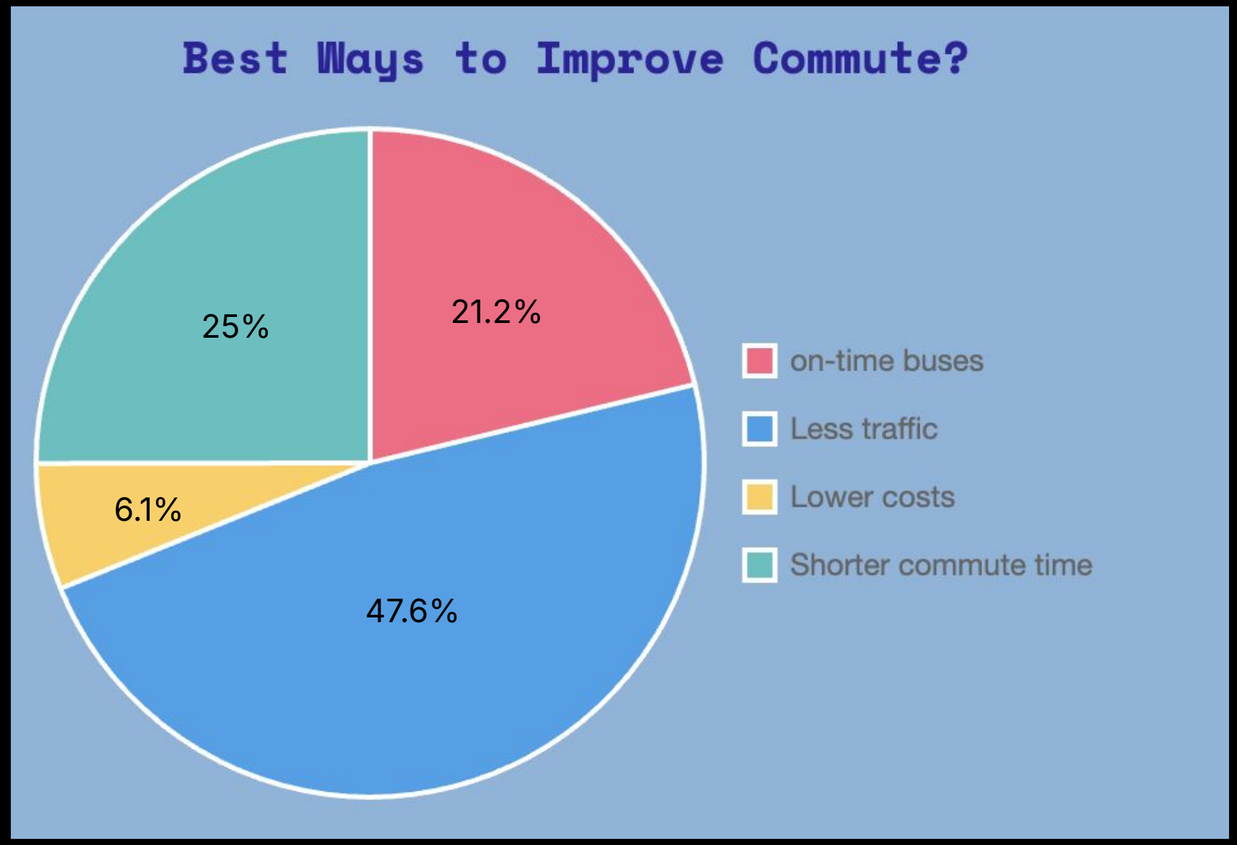
How Do Bus Lane Violations Vary by NYC Neighborhood?

Team GOYAL: Debasree Sen, Addina Rahaman, Sadeq Abubaker Alhanshali

Problem Statement

- Traffic congestion is one of the leading causes of stress in NYC. One of the major contributors to traffic are **bus lane violations**.
- According to DATELINE: CUNY, nearly two thirds of surveyed CUNY students have reported stress about their commute. They have overwhelmingly said that the number one way to improve their commute is by alleviating traffic congestion.
- According to ABC News 7, in Fall 2024, there was a protest held by Brooklyn College students addressing the unreliability of buses in Flatbush.


Does Commuting Cause You Stress?
24.1%
60.8%
15.1%
Yes
No
I don't know


Best Ways to Improve Commute?
25%
21.2%
6.1%
47.6%
on-time buses
Less traffic
Lower costs
Shorter commute time

Bus lanes are necessary for quick and easy bus flow

Bus Lane Violations

- **NYC bus lanes are frequently blocked by private vehicles.**

Bus lane violations impact:

NYC traffic congestion

Bus travel time & delays

Overall transportation efficiency

Understanding **where and when** violations occur helps identify hotspots and improve enforcement strategies.

NYC MTA ACE Violations Dataset

The **ACE (Automated Camera Enforcement)** program expanded over recent years since 2019 with cameras that detect when drivers violate parking laws and block bus lanes.

What does ACE capture? License plate number, bus route violated, exact geolocation the violation, and the time it occurred.

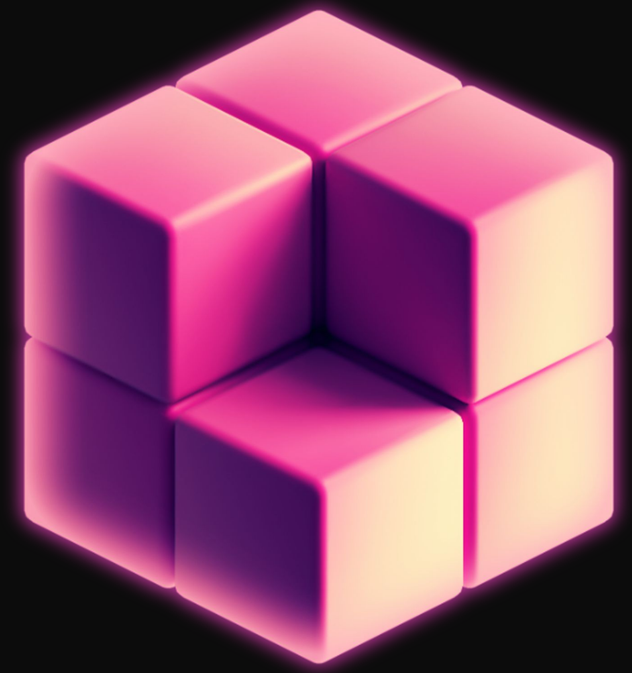
Benefits of this:

- Decreasing blocked lanes increases bus speeds
- Decreases collisions, blockages, and improves safety + accessibility for people
- Lowers recidivism rates, meaning less repeat offenders

Understanding these patterns helps:

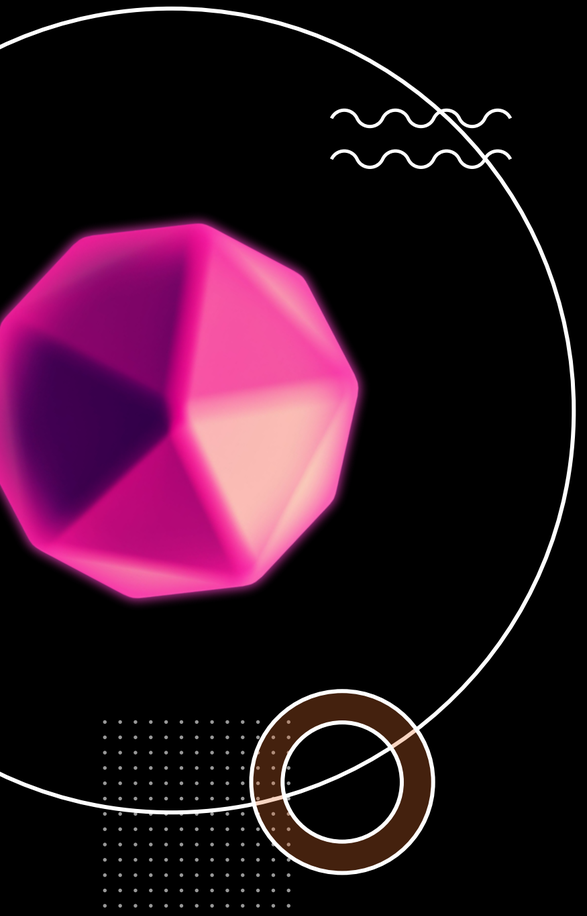
- City planners
- MTA enforcement teams
- Community advocates
- Public transit users
- Legal Teams with identifying recurring offenders.

Dataset Overview



- Our dataset has records from 2019 but for the sake of this project, we are focusing on 2025 data.
- Up to **2.5 million records**
- Fields included:
  - 'Violation ID'
  - 'Vehicle ID'
  - **'First Occurrence'**
  - 'Last Occurrence'
  - 'Violation Status'
  - **'Violation Type'**
  - 'Bus Route ID'
  - **'Violation Latitude'**
  - **'Violation Longitude'**
  - 'Stop ID'
  - 'Stop Name'
  - 'Bus Stop Latitude'
  - 'Bus Stop Longitude'
  - 'Violation Georeference'
  - 'Bus Stop Georeference'

Our Research Topic/Solution



- Our solution: Use visualizations of the ACE violation dataset to ***determine the priority of enforcement resource allocation by neighborhood.***

- We decided to use the ACE Violation records to answer these questions:
  - **How do bus lane violations vary across NYC neighborhoods in 2025?**
  - **Sub-questions:**
    - Which neighborhoods show the highest number of violations?
    - How do violations change month-to-month?
    - Are certain hours or days more problematic?
    - Which violation types are most occurring?

- **WHY?** This helps the city focus enforcement on the right areas at the right times, so they don’t waste resources and can save money.

For this project, we took a look at **30 different neighborhoods** across New York City.

Neighborhoods Analyzed

Queens:

Astoria, Elmhurst, Flushing, Howard Beach, Jamaica Center, Maspeth

Brooklyn:

Bed-Stuy, Bensonhurst, Bushwick, Crown Heights, DUMBO, Downtown Brooklyn, Flatbush, Williamsburg

The Bronx:

Fordham, Hunts Point, Mott Haven, Pelham Parkway, Riverdale, South Bronx

Manhattan:

Harlem, Lower East Side, Lower Manhattan, Midtown, Soho / Greenwich Village, Upper East Side, Upper West Side

Staten Island:

Mid-Island / New Dorp, St. George, Staten Island North Shore

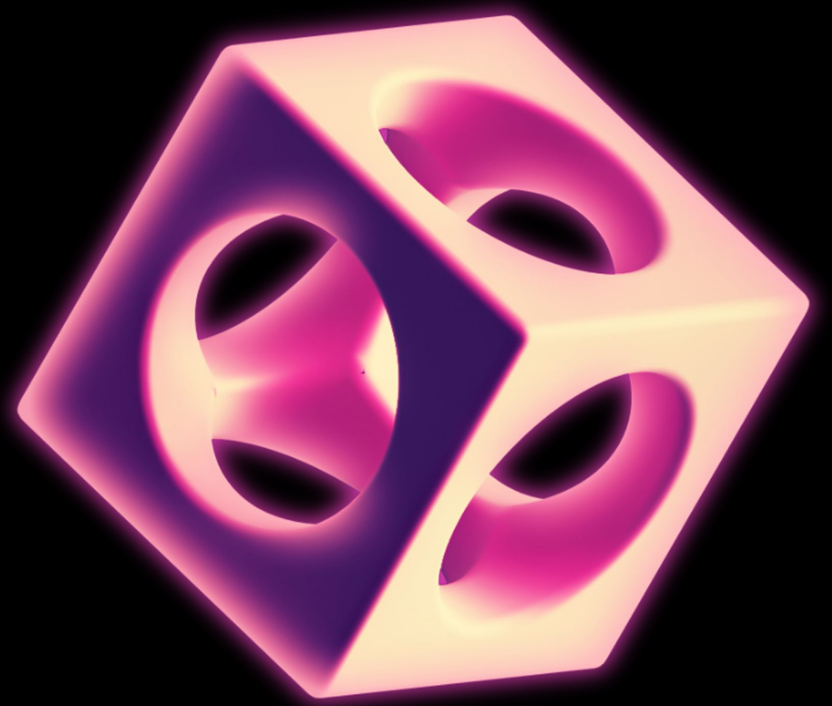
Tools + Methodology

1. Data Cleaning & Processing (pandas, geopandas):

- Filter 2025 violation records
- Clean timestamps
- Define 1-kilometer neighborhood buffer zones.
- Spatial join: which violations from ACE dataset fall into these neighborhood buffer zones?
- Perform aggregations within each buffer zone
- Group by month, borough, and neighborhood

2. Visualization (Kepler.gl + React + Recharts + matplotlib):

- Interactive charts
- Filters (borough, neighborhood, month)
- To be deployed on GitHub Pages



How have violations changed month to month?

- Interactive time series line chart
- One line per borough

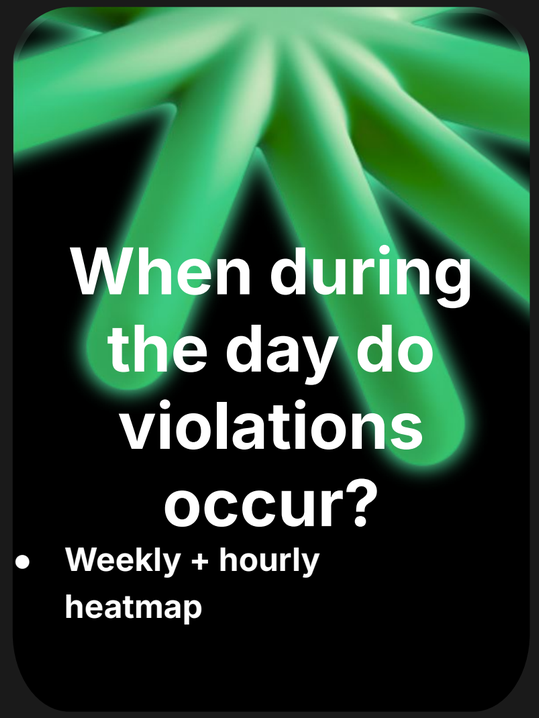
Where are the worst violation hotspots?

- Interactive Hotspot Map
- 4 Hotspot maps, one per quarter of 2025

Dashboards we made...

What are the major violation types per neighborhood?

- Stacked bar chart across neighborhoods



When during the day do violations occur?

- Weekly + hourly heatmap

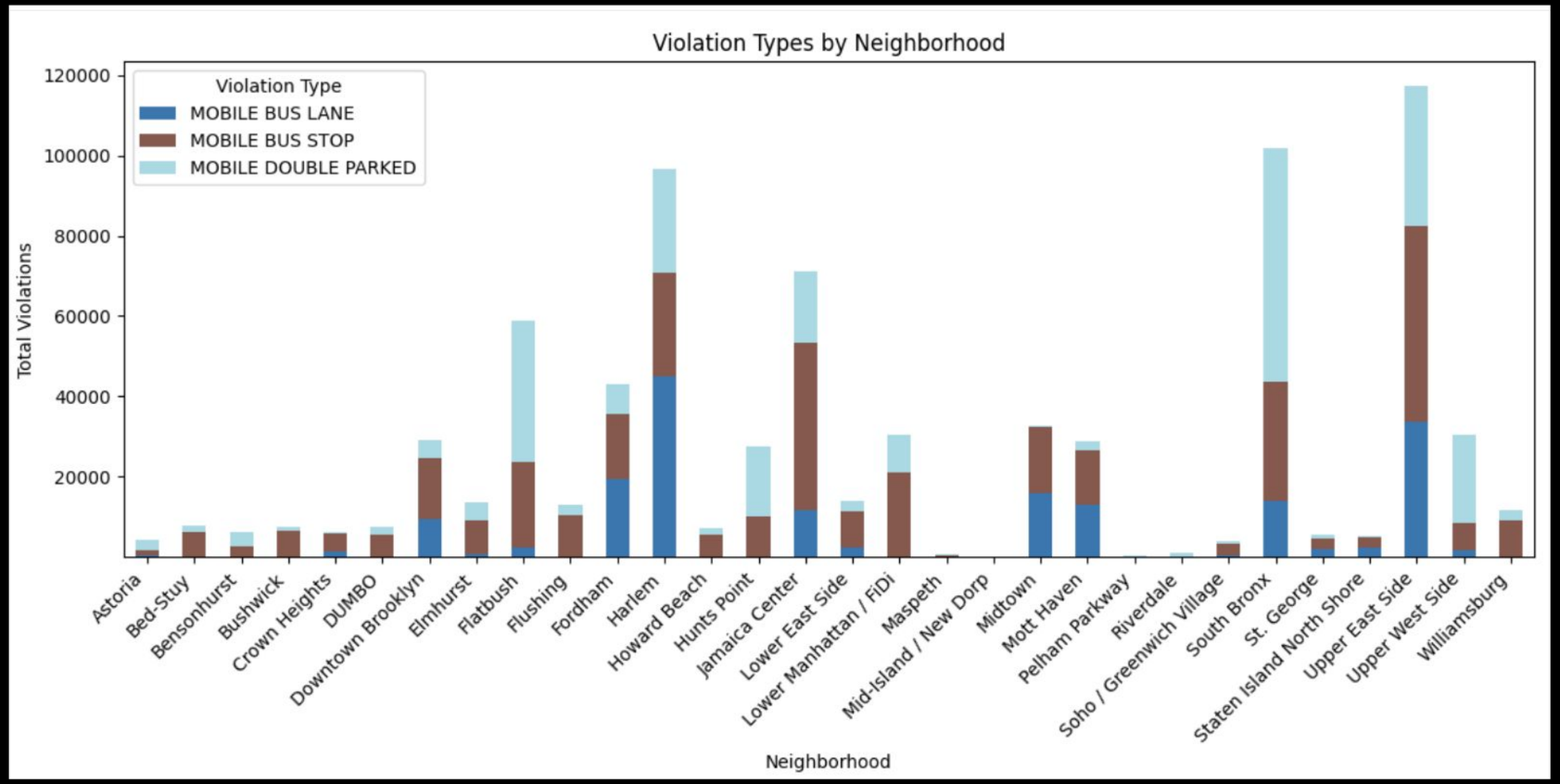
Which neighborhoods deviate from their borough average?

- Interactive time series line chart per borough
- Borough average + quarterly average

We will go over our interactive visualizations live

Interactive Visualizations Demo

Total Violations by Type + Neighborhood



- **Mobile Bus Lane:** A vehicle was recorded traveling **inside a legally restricted bus-only lane** during its hours of operation.
- **Mobile Double Park:** A vehicle was detected **double-parked** (stopped beside another parked vehicle), obstructing a travel lane.
- **Mobile Bus Stop**: A vehicle was recorded **stopping or standing in a designated bus stop zone**, blocking buses from pulling in

Team Roles

- **Addina Rahaman**: Primary EDA, setup React frontend and dashboard components, derived datasets for visualizations, made Kepler.gl hotspot maps
- **Debasree Sen**: Created visualizations borough by borough and formatted them on React interface
- **Sadeq Alhanshali:** Created slides, collaborated with Debasree

Problems We Encountered

1. At first, we tried to look at CUNY campus instead of neighborhood. However, since a lot of CUNY campuses are concentrated in midtown, the interactive hotpot map was very cluttered. So, we decided to switch to neighborhoods.
2. We also wanted to do a heatmap of which lanes had the most violations, however the visual was also very cluttered.
3. Understanding which visualizations would help to represent the data meaningfully.
4. Extra time: with more time, we would have different the visualizations we made.

Thank You!