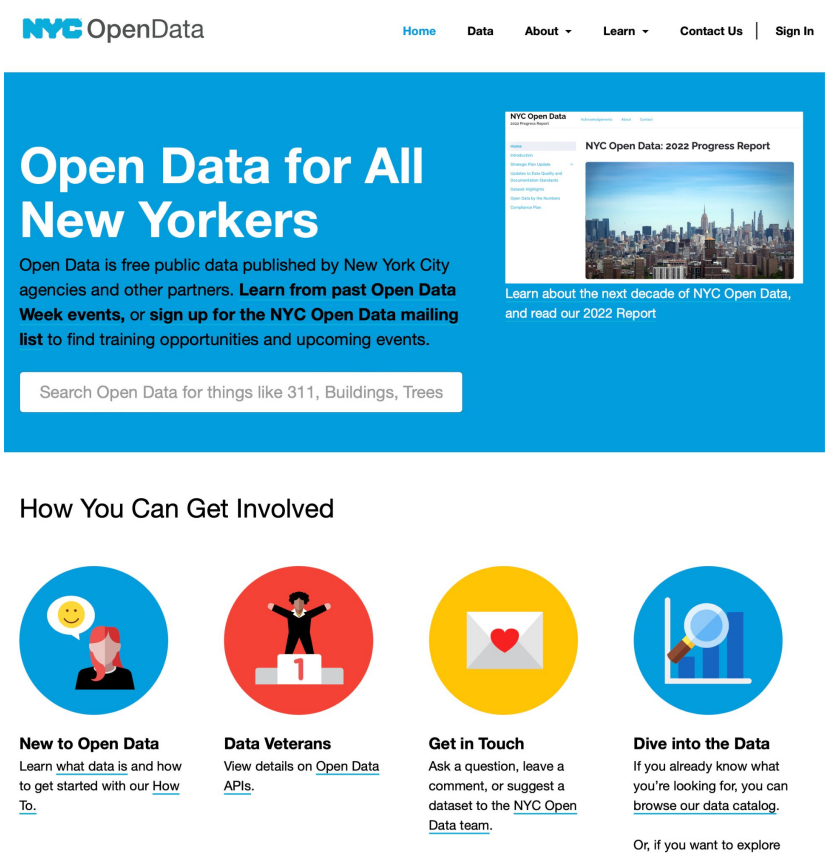
February 2023 NYC Vehicle Collisions



Meet Our Team

Brittni Breese

Yash Kansal

Yeyan Wang

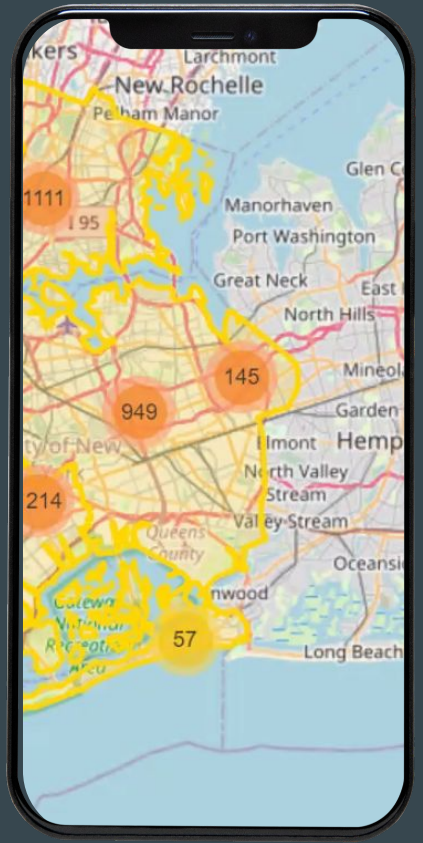
Nicole Campos

Adam Paganini

Who will this inform?

Individuals concerned about traffic safety:

- Police departments monitoring traffic safety
- Urban planning professionals
- Traffic management applications for building smart cities
- Mapping, navigation, and route planning businesses
- Location intelligence companies



Data Discovery

Data Set: NYC Motor Vehicle Collisions

- ~1.98M rows of data, updated daily
  - Only used February 2023 (~6,800)
- Source: NY Police Department (NYPD)

API Request

`requests (.json)` to `Pandas df`

Extract

Dropped Null Rows which didn't have vehicle type defined (~100 of 6800)

Changed Datatypes
Crash Date to "datetime"
Crash Time to "Hour"
Numbers to "int"
Lat/Lon to "float"

Dropped Columns (29 to 19)

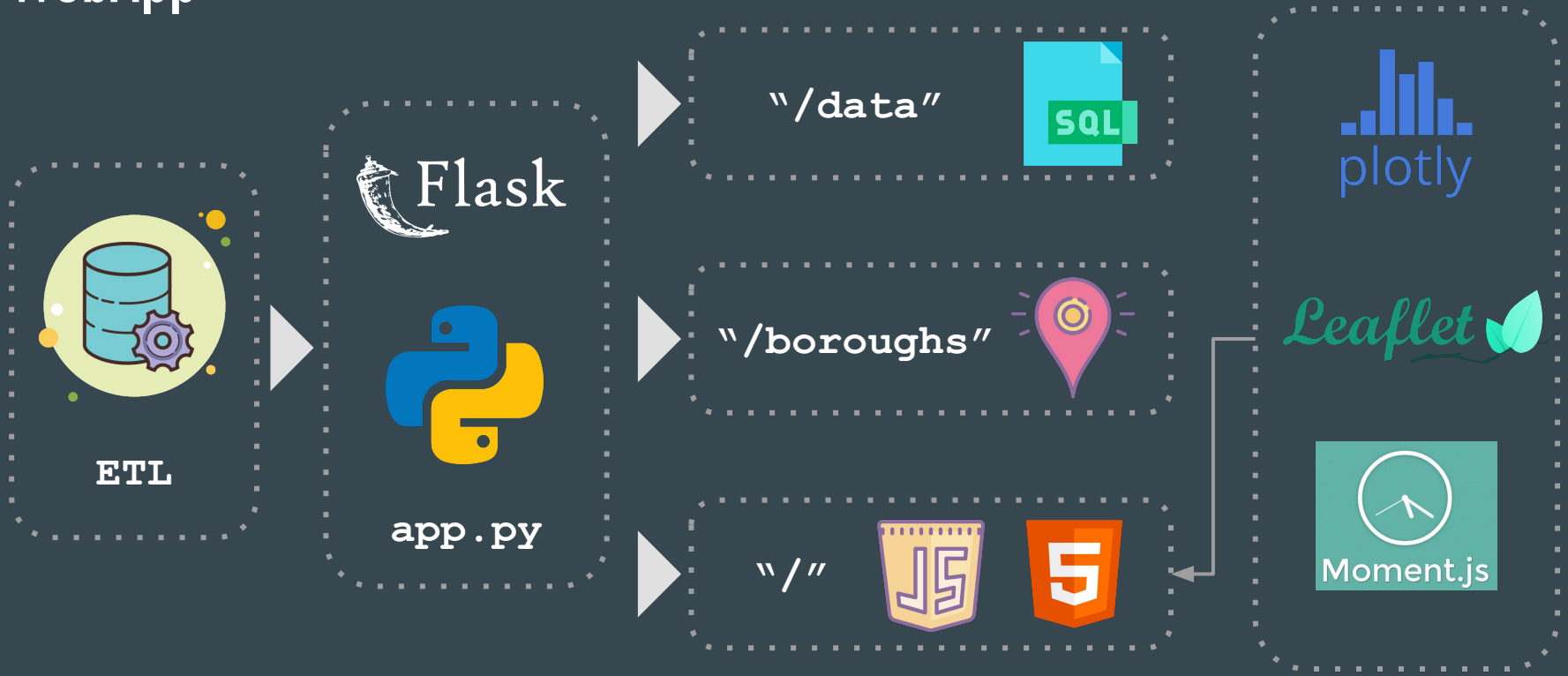
Transform

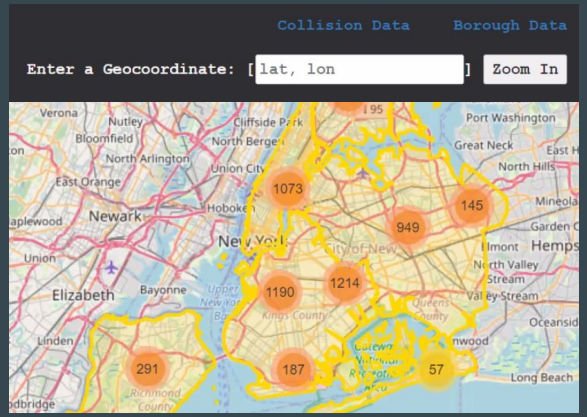
Saved as SQLite DB

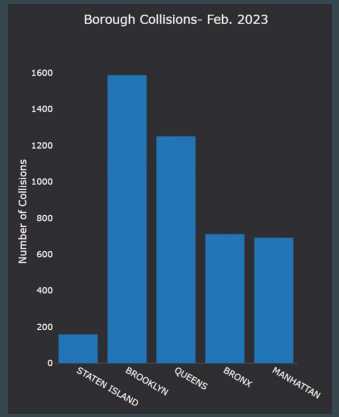
`engine` and `to_sql`

Load

WebApp







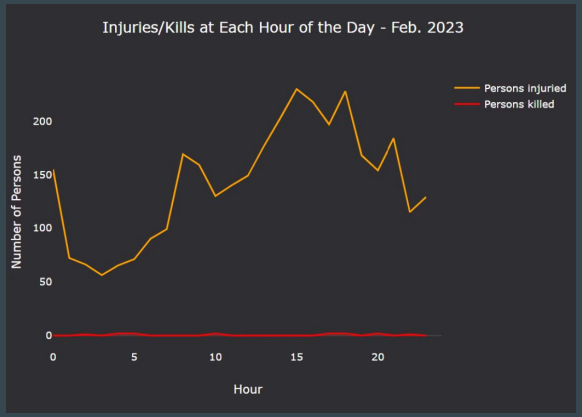


Chart Features

- Marker clusters identifying number of collisions in an area
- Marker pop up displaying more details of each collision
- Polygons displaying NYC Boroughs
- Lat/Lon zoom on specified location
- Scroll over bar/line charts to see data points

Styling

- Tab Buttons
- Tab Contents
- Plotly Chart Layout

DEMO

Key Challenges and Recommendations

- Only used one month of data due to difficulties loading the marker clusters
- Boroughs composed of multiple polygons (e.g., Staten Island)
- Random lat/longs off the coast of Ghana
- Zoom by address vs lat/long

- Efficiently store and display all the data available to enable longitudinal analysis
- Use lat/longs to label all collisions with appropriate Boroughs for popups
- Identify and remove outlier lat/longs from dataset
- Need to use another library that could parse addresses

Thank you!