Final Project - Proposal
Team members: Pratishta Yerakala and Brendan Mapes

Title: NYC Rats

Abstract: We're interested in rat data. Since NYC is known for being so populated and dense, it's inevitable that there will be vermin sightings and nuisances. We'd like to take a look at the geographic nature of these sightings with relation to other maps with geographic features (i.e. sewage systems, neighborhoods with many restaurants, etc).

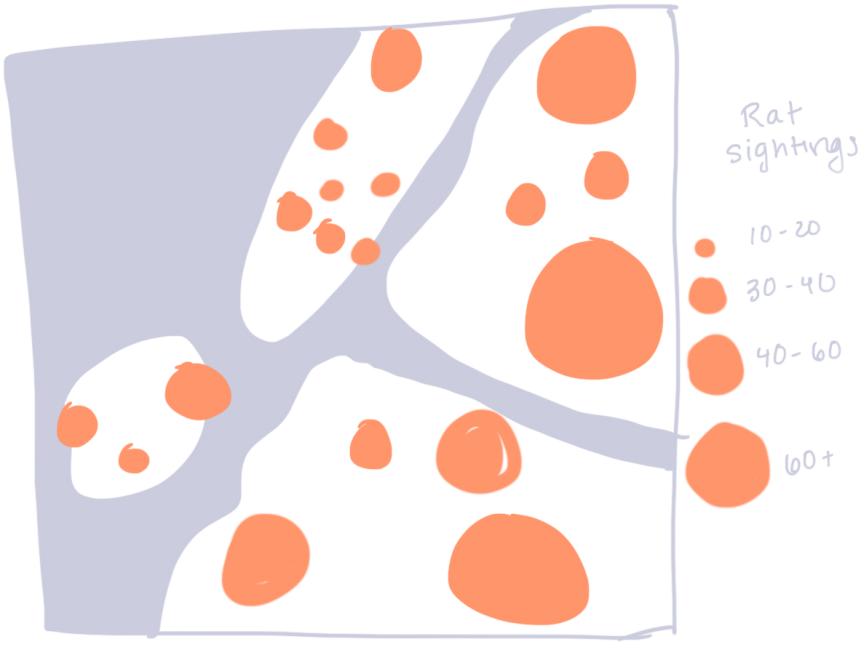
- What is the distribution of rat sightings across New York City and its boroughs?
- How do rat sightings compare when overlaid with sanitation pick up, neighborhoods with densely packed restaurants, neighborhoods with high or low income, etc
- Current status of rat inspections, where they fail or pass or are in process the most
- Compare data from 311 rat sighting calls and rodent inspections to see if there is a correlation between the complaint and the response from the city
- Rat sightings and proximity to parks
- Rat sightings in relation to dog ownership density (because dogs have to be licensed, and this data is available)

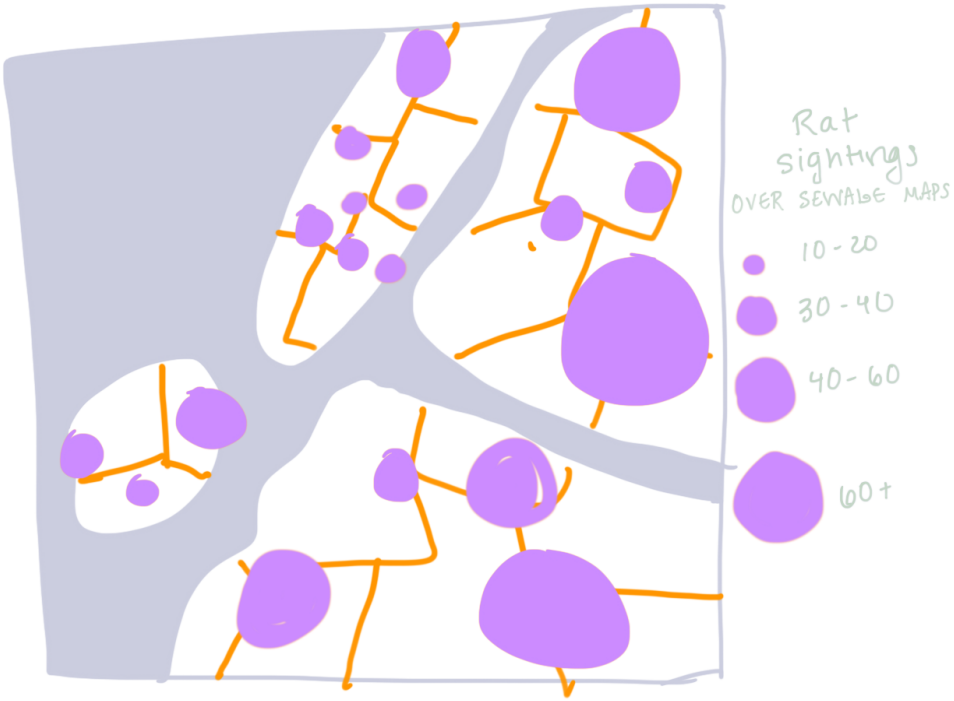
Data:

- Rodent Inspection - NYC OpenData
- OPEN SEWER ATLAS NYC (tumblr.com)
- DOB Stalled Construction Sites | NYC Open Data (cityofnewyork.us)
- NYC Dog Licensing Dataset | NYC Open Data (cityofnewyork.us)
- NYC OpenData - 311 calls filtered by rat sightings
- Housing Maintenance Code Violations | NYC Open Data (cityofnewyork.us)
- Potential others to include:
    - City parks location / shapefile data
    - Restaurant data
    - Sanitation data
    - [some demographic / census data]

Visualization brainstorming:

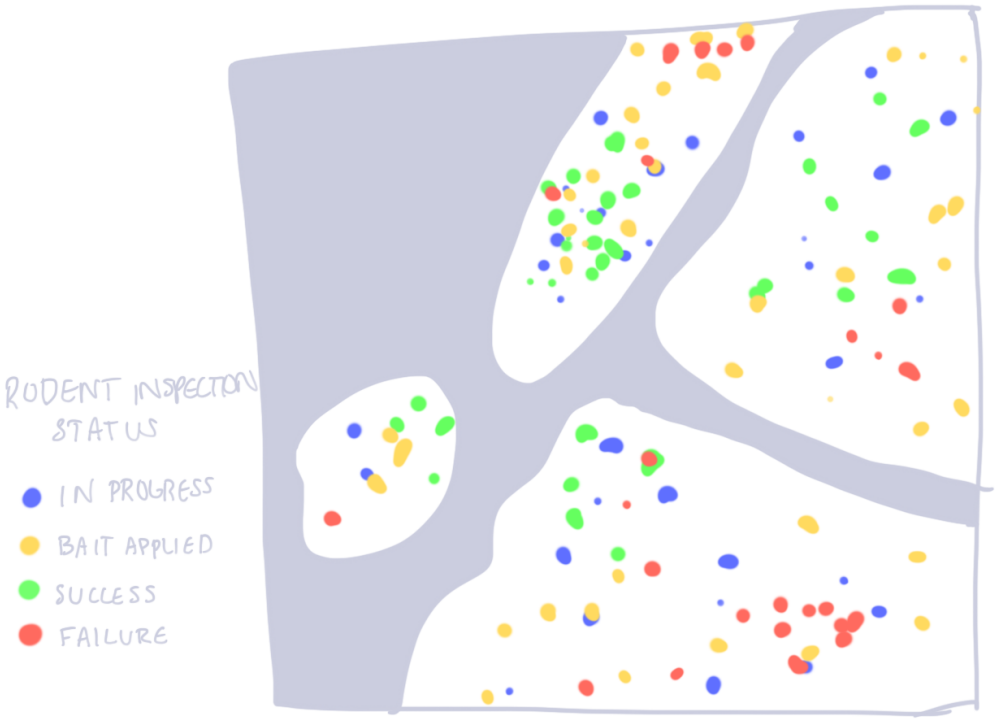
- Primary visualization style: **maps**
    - Rat sightings data
        - Centroid data, size of each centroid represents frequency of rat sightings
        - Heatmap with rat sightings data
        - Heatmap / choropleth with zip code data to see if there is a significantly rat dense neighborhood
          Could look something like this:





- Rodent inspection status data
  - A dot on the map will represent an inspection, each dot will have a different color based on inspection status.

- Possibility of this map being shown over time (we can look at columns inspection date and approved date)
- Can be possibly used for text analysis
  Could look something like this:



- Timeline / frequency over time visualization
  - Bar plot with time on the x axis
  - Line graph with frequency over time, can explore pre-COVID-19 and COVID-19 data