Predicting Rat Activity in New York City using Machine Learning

Ali Bridgers

Problem: New York City faces a significant challenge in controlling rat populations, especially in densely populated urban areas. Rats are a problem for residents, property owners, businesses, and entire neighborhoods. Their gnawing and burrowing can cause damage, contaminate food, spread disease, and reduce the overall quality of life. Accurate prediction of rat activity is crucial for addressing the rat problem in NYC. By identifying areas with high rat activity, the city can target its rat control measures more effectively, including rat-resistant garbage containers, building and sidewalk maintenance, and public awareness campaigns on proper waste disposal and not feeding wildlife. Furthermore, by predicting future rat activity, the city can proactively address potential rat infestations and optimize the allocation of resources to rat control measures, such as the Rat Mitigation Zones (RMZs). By using machine learning techniques, I aim to build a predictive model to help NYC identify areas with high rat activity and potential hotspots for future rat infestations. This predictive model will be valuable for guiding future rat control efforts, optimizing the allocation of resources to RMZs, and improving the quality of life for the residents of NYC.

Data Source: https://data.cityofnewyork.us/Health/Rodent-Inspection/p937-wjvj
This webpage provides information on rat inspections in New York City and is part of the Rat Information Portal (RIP), a web-based mapping application where users can view rat inspection data. The data provided on this webpage comes from the NYC Department of Health and Mental Hygiene (DOHMH), Division of Environmental Health Pest Control Database and is owned by NYC OpenData. It also includes notes on data limitations, noting that the absence of a property/taxlot from the file does not indicate an absence of rats, but rather that it has not been inspected. It also notes that neighborhoods with higher numbers of properties showing active rat signs may not necessarily have higher rat populations but could simply have more inspections. This dataset is updated daily, however the data we are using is from August 17, 2023.

Expected Outcome: By the end of this project, I aim to have a robust machine learning model capable of predicting rat activity in NYC with high accuracy. The model's insights will be valuable for guiding future rat control efforts and optimizing the allocation of resources to RMZs.