Project 1

Group Slay- Ata Algan, Elizabeth Pauley, CJ Sivak

2022-10-14

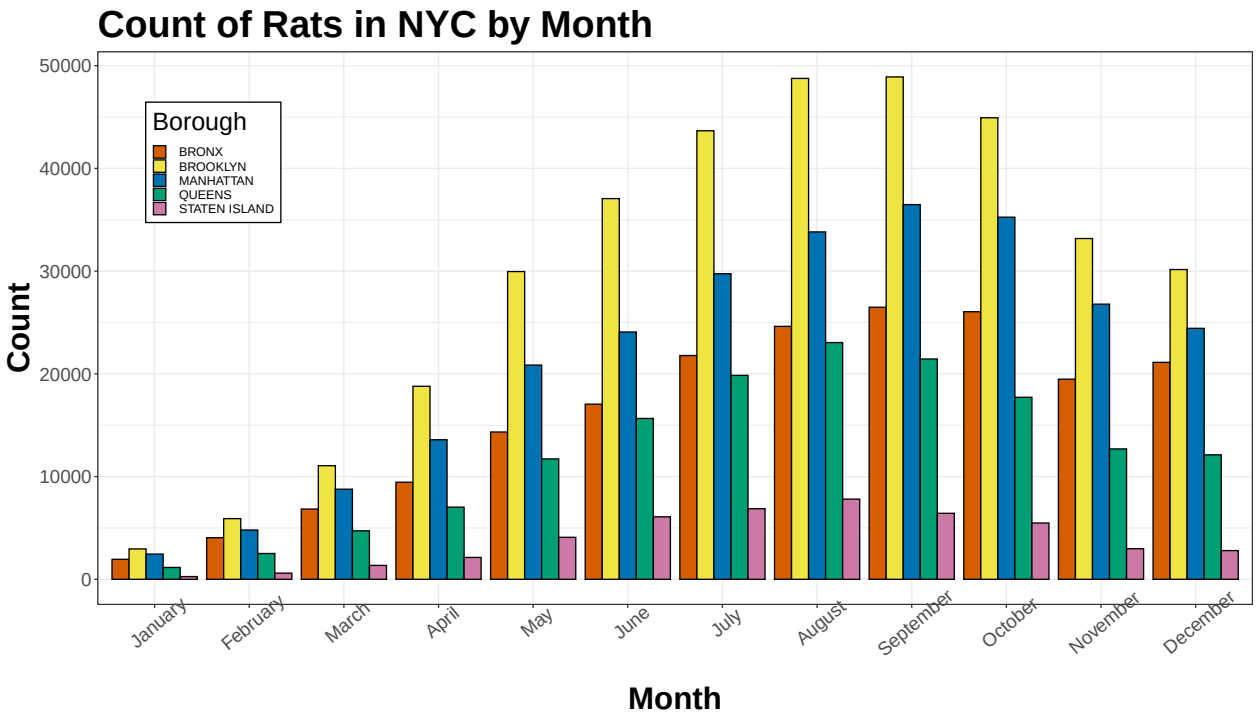
Memo to the Mayor of the City of New York Regarding Rat Sightings

Introduction

Through statistical analysis, we were tasked to investigate past data regarding rat sightings throughout the City of New York in order to find trends that may be of use to city level policy. We found the spread of rat sightings over months to be of interest, as well as segregating the data by borough to allow for efficient allocation of resources to prevent rat populations from growing.

Further, we found the percent change by month for each borough to examine what months we need to allot more resources to defer the rats, and if this may be different by borough.

Rat Sightings Across Boroughs by Month



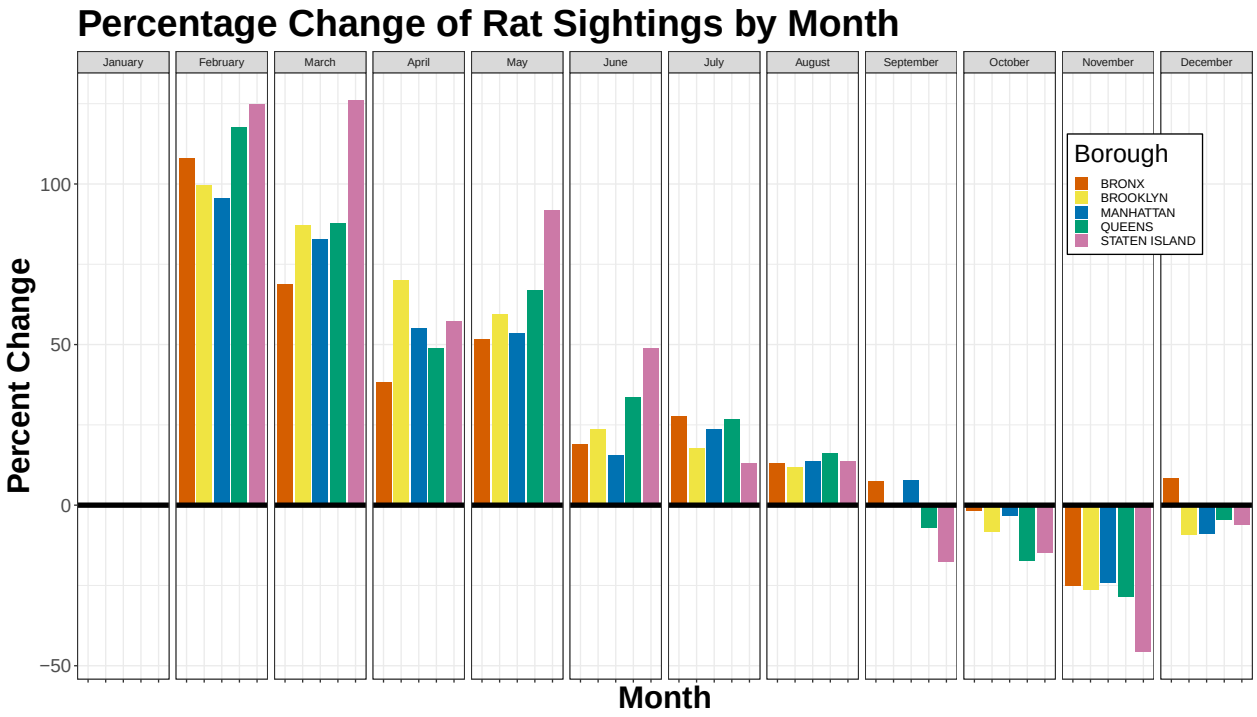
Analysis

The visual above shows that all of the boroughs display similar patterns of rat sightings over the months of the year. Typically, as one burrow sees an increase in rat sightings, so does every other burrow. Rat sightings were higher in the ending summer months (August and September) and lowest in the winter months (January and February). The raw counts of rat sightings for Brooklyn were the highest, but we believe that this can be

explained by the borough having the highest population. We suggest that boroughs with higher rat sightings have more resources to fight the issue.

We further wish to explore whether one borough has significanlty more rat sightings one month compared to the next in their borough. To explore this more, we visualize the percent change of rat sightings for each burrow by month below. By visualizing percent change by month, we can allocate resources not only by total count of rat sightings, but we can prepare better for months where rats increase significantly, especially for boroughs that aren't given a lot of resources to fight the problem. This allows the city to be able to plan for the use of resources better, and devise a plan to eliminate as much rat population as possible before the surge of rat sightings.

Percent Change of Rat Sightings Across Boroughs by Month



Analysis

Above is the visualization for percent change by month sorted by borough. By doing this, we find that while Brooklyn has the highest rat sightings, the percent change for rat sightings follows similar patterns to the rest of the boroughs. Staten Island has the highest percent increase in February and March, but also the lowest percent increase in November. We see the highest percent increase for all boroughs in February and March, and then lower, and negative, increases in the Fall and Winter Months. It is important to note that from our first visualization, the count of rats is lowest in the winter months. Thus, we see really high percentage changes due to the fact that the rat population is nearly doubling for the first few months of the year, since the population started out at its lowest. We predict if we put more resources to the beginning of the year when rat population is low and fight the rats before they have a chance to double in size every month, we could reduce the total number of rats throughout the whole year.

Summary: Our Official Recommendation

We, Group Slay, officially recommend the City of New York to allocate resources by borough by total rat sightings. Thus, Brooklyn and Manhattan should be given the most resources to fight the rat population. We further recommend, by our percentage change visualization, that more resources, for all boroughs (especially Staten Island) be allocated during the first half of the year, as the rat population tends to grow really fast

before tapering off during the winter months, and less resources are needed since the rat population remains stagnant, or decreasing, during the second half of the year after the population has reached its peak.

If we tackle the rat population during the winter months, when the sightings are lowest, but the percentage change is the highest (nearly doubling every month), we forsee great mitigation of the population of rats throughout the entire year.

Part 2: Data Hunting

First dataset: Business openings and closures in the 3 largest cities in America the past 30 years

Link for dataset 1

I found data from the US Census about business openings and closures over the past 30 years. This data goes as far back as 1978 and contains 3 metropolitan areas that I chose (Chicago, New York, and Los Angeles). I think it will be particularly interesting and telling about each city the number of businesses opening and closing, especially in times of the recession in 2008 and covid in 2020. It could be neat to look at what cities bounce back faster than others, as well. Further, the dataset has business type, such as agriculture and manufacturing. This could be interesting to explore the number of certian business types opening and closing over the past 30 years, especially by the cities chosen since they are all in such vast geographical areas. As an economics student, I find particular interest in this dataset.

Second dataset: FBI's National Instant Criminal Background Check System

Link for dataset 2

I have chosen this dataset because I found it interesting to have a detailed dataset of firearm background checks gathered by a government identity, the FBI. Gun control in generally has become an important issue both in politics and social life. Therefore, the popularity of the debate on gun reforms contributed to my interest in this dataset. Another aspect that I like is the applicability of the data to policy proposals. It covers every state in the nation and so it can give us a general understanding of the country. Furthermore, the type of gun was listed on the dataset too, which can explain if certain guns follow a different pattern. Overall, I think gun control is a critical issue of the U.S. affecting everyone in some way or another.

Third dataset: eBird: citizen science data regarding bird sightings across the United States

Link for dataset 3

I have chosen this dataset because I found it to be a possible intersection between public health and the environment. In previous research that I am affiliated with, we record birdsongs and analyze them for species and number. This data is then looked at with regards to the built environment and the physical state of the area in which the birdsong was recorded. This dataset would allow our group to possibly explore the trends of bird-sightings in a not-yet-specified area of Michigan and see if there is a connection to health outcomes. I find the possibility of arguing for mutual benefit in greening areas for both wildlife and human health benefit of great interest as an urban planning major and a health geography researcher. Another point of interest is that this is citizen science data, which would give us unique experience in working with data that was not collected in a research environment, since citizen science is becoming more widely used as a method of data collection.