

Early Insights Report

1. Collect and pre-process a preliminary batch of data
 - a. Data found in [Github Repo](#)
2. Perform a preliminary analysis of the data
 - a. Maps and Plots found in [Github Repo](#)
3. What questions can and can't be answered with the data you have. What other data do you need and how will you get access to it?

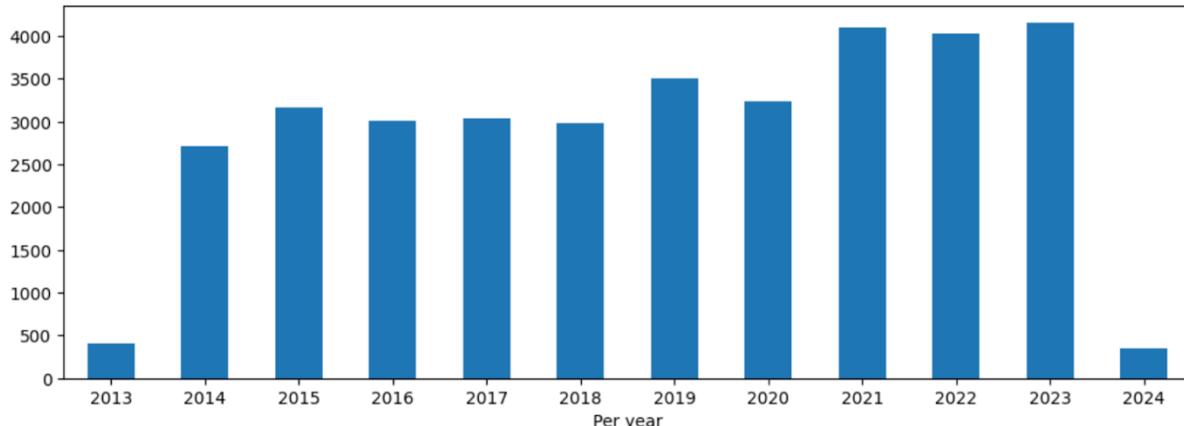
A primary question answered by the preliminary analysis is: "**How many reports were there in the previous 2-3 years? 5-10 years? Are there any positive or negative trends to these reports (decreasing in sum total, etc)? How do other factors such as season impact the nature of the complaints? (ie: are there more complaints in the summer? Winter?)**"

We have collected four distinct graphs that illustrate the number of animal reports per unit of time since 2013. Firstly, there is the number of animal reports per year shown below. The number of reports has generally steadily increased since 2014 (animal reports were only collected October - December 2013 so this data is insufficient for comparison), with around 2500 in 2014 to around 4000 in 2023. The increase of reports can be attributed to two factors:

- Incidents involving animals has increased
- More civilians have been made aware of 311 and utilize it to submit animal reports

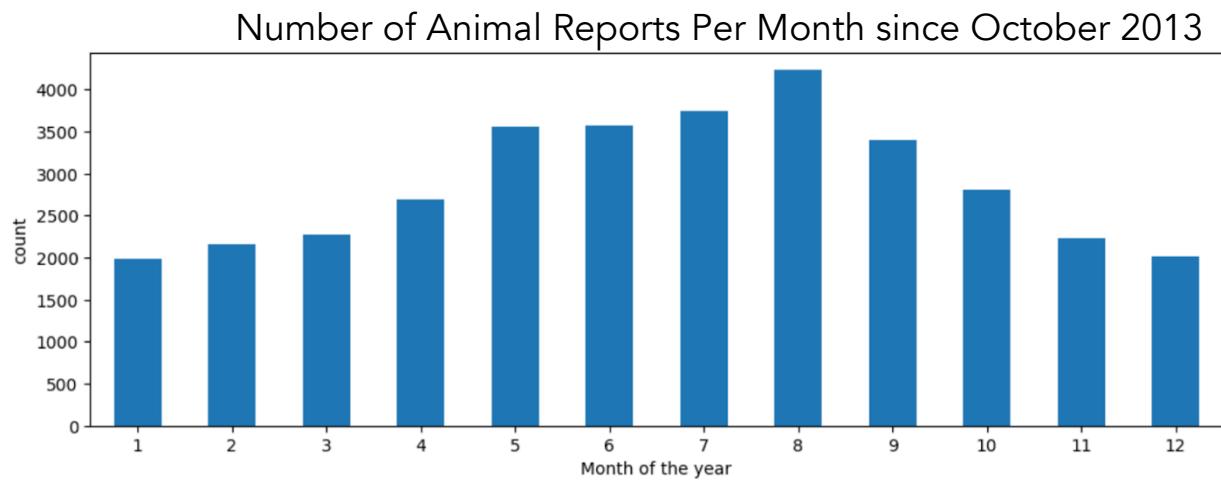
The second point is valuable as it provides positive reinforcement that 311 is successfully publicizing their service to Bostonians and better equipping them to use their city wide resources.

Animal Reports per Year



Other interesting graphs we were able to collect from our preliminary analysis include the number of animals with respect to month of the year and hour of the day, respectively.

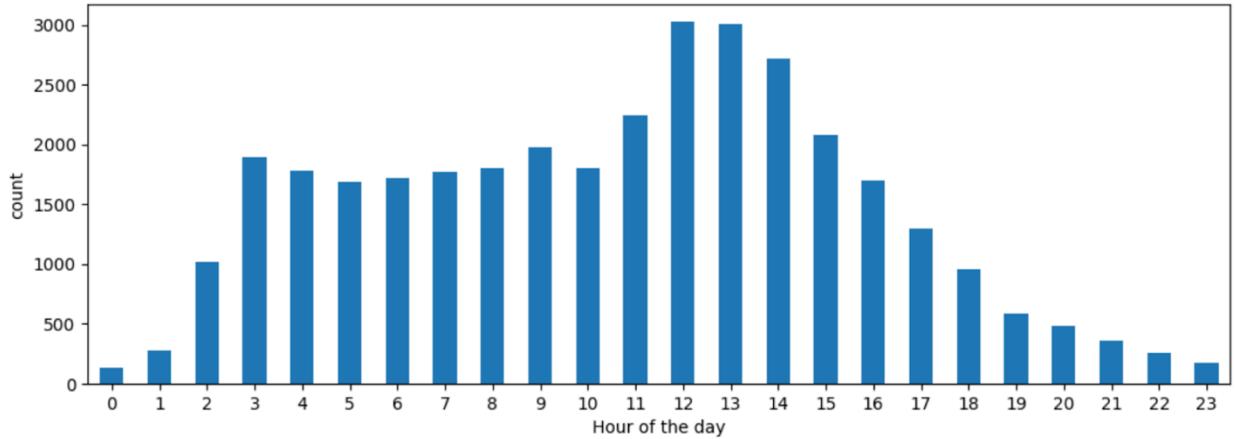
To begin, shown below is the Number of Animal Reports per Month. As expected, animal reports are submitted most frequently during the late-spring/summer months of May through August, with roughly 1,000 more animal reports submitted in May compared to April. During these warmer months, animals are more active and abundant versus the winter months where many animals choose to be dormant as a reprieve from the bitter Boston cold. This insight provides 311 the ability to proactively scale the number of employees working to handle animal reports, with a larger staff necessary during the summer months.



The graph below demonstrates the frequency of animal reports with respect to the hour of the day (in military time). 12PM-2PM is a considerable peak in the number of animal reports. These times are of course very active in a bustling city like Boston, and the number of animal reports starkly decreases during the nighttime hours. Interestingly, the number of hours during the early morning hours (3AM - 6AM) is modest with around 2,000 reports each. This can be attributed to nocturnal animals including raccoons, bats, and opossums causing disturbances to Bostonians. Further analysis on the type of animal per hour of the day will be completed and provide clearer insights about which animals are the cause of reports at various times of the day.

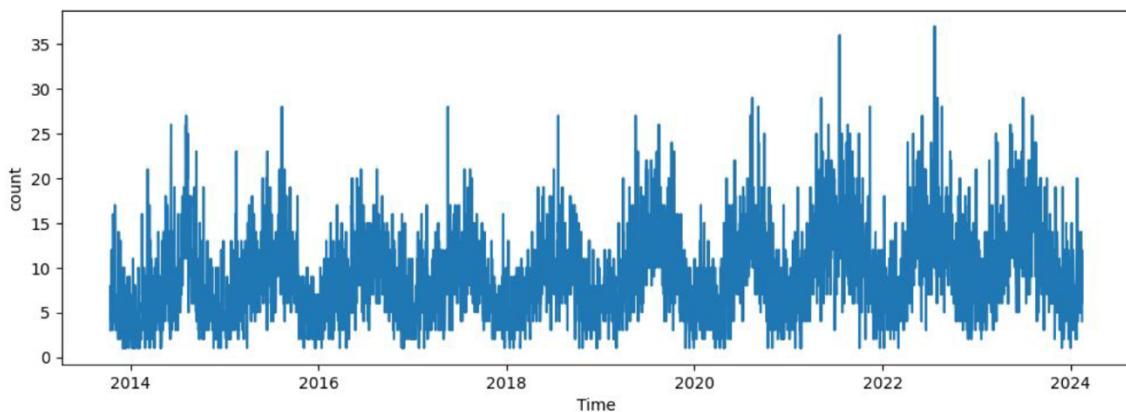
Similar to the months of the year, 311 can use this graph to scale their employees on shift, with a greater number of works needed from 12PM-2PM and few needed from 8PM-2AM.

Number of Animal Reports Per Hour since October 2013



This is a plot of accidents count per the entire period we have reported accidents (Oct-2013 to Feb-2024). The count oscillates periodically in the same pattern every year, where accidents count fall to its lowest value at the beginning of every year, and peak mid-year around June. This pattern repeats for every year, confirming our findings from the count per month of the year plot. What we can see in this plot and not in the count per month of the year plot is what year shows an increase in their peak point, we can see that accidents rates are the same from 2014 to 2019 with only 2017 not having as high of a peak as the rest of years in that interval. 2021 and 2022 have relatively higher peaks than the other years, 2023 also have a higher peak in June than 2014-2019 but not as high as 2021 and 2022. The most interesting finding is the year 2020 which has the lowest peak. The most plausible reason for that one can think of is the pandemic, as there were less people driving and going out. We can argue that because of the drop in accidents perhaps animal populations increased or perhaps since people stayed at home, they got more pets and maybe that's why it increases in 2021 and 2022 then starts to decrease in 2023 (to reach the equilibrium point that 2014-2019 had. This gives us an idea of modeling it is using a logistic model.

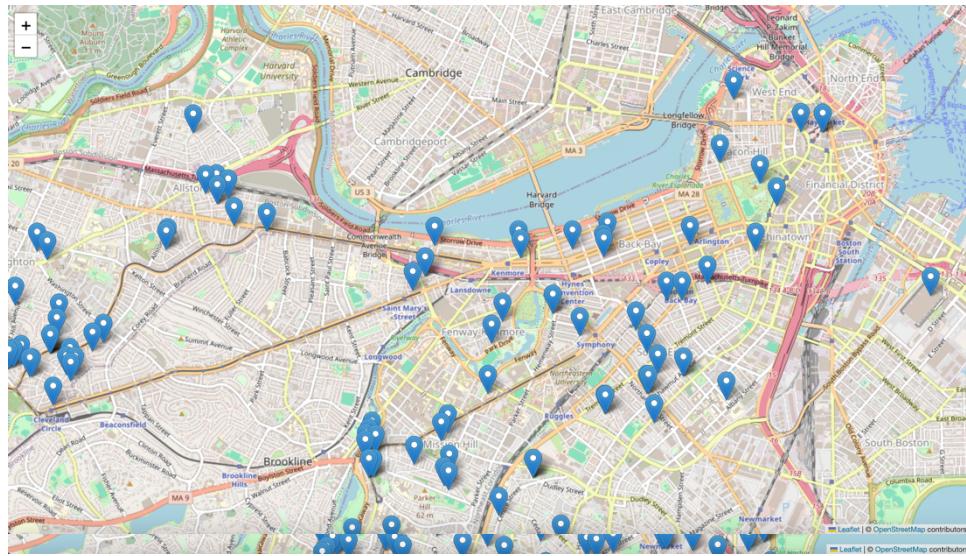
Timeseries for Animal Reports



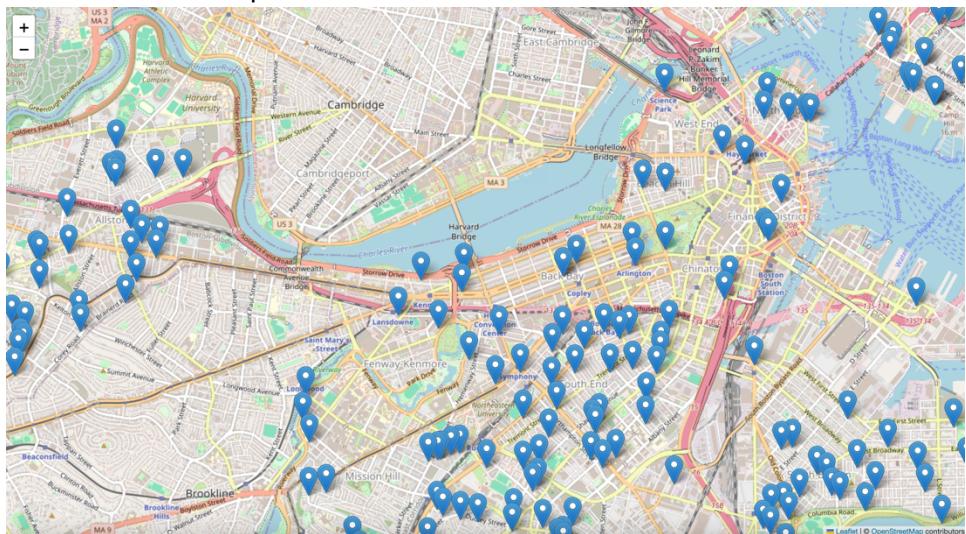
Another crucial questions we were able to answer were, **"What kinds of animals are most common in complaints? (Rats? Domesticated pets? Dogs? Cats?) Where are the complaints coming from geographically? Is there any trend to type of animal complaint and area?"**

We have created five maps representing the location of dog, cat, raccoon, bat, and bird reports. These animals were chosen because they are the animals that are most commonly the subject of submitted animal reports by Bostonian. Understanding the location of where these complaints occur can provide mitigation strategies, i.e. if one area has a high frequency of bat reports, greater measures in reducing the bat presence in said area can be taken. The maps are created In HTML because they are large and cannot be encapsulated in a single screenshot. The HTML files also provide the opportunity to zoom and navigate through the map. Below are some snippets:

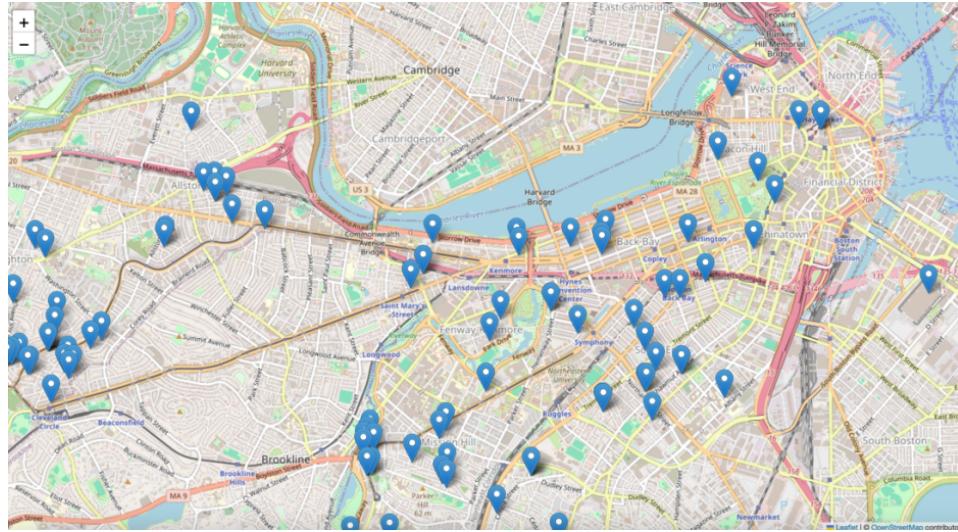
Location of Dog Reports in Downtown Boston since October 2013



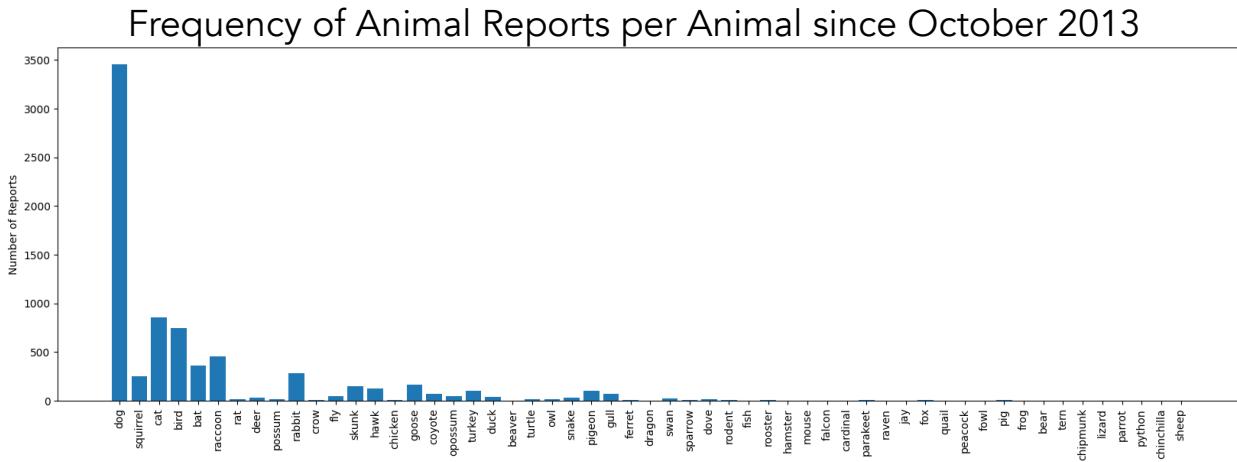
Location of Cat Reports in Downtown Boston since October 2013



Location of Bat Reports in Downtown Boston since October 2013



The following bar chart displays the frequency of the number of animal reports pertaining to each animal.



These maps and most notable the bar chart provide extremely rich insight behind the animal reports. Namely, dogs, cats, bird, raccoons, and bats are most often the cause of civilian report animal incidents. As the former two are common household pets, this is a reasonable piece of data. Moreover, the latter three are common wildlife in New England. Nevertheless, knowing which animals are most frequently reported in civilian incidents allows 311 to allocate resources more effectively, for example campaigns targeting responsible dog ownership can mitigate the number of dog incidents. Furthermore, this data can be useful in urban planning, waste management techniques can be implemented to reduce the frequency of raccoon or other scavenging animal reports.

Further analysis with regards to when each animal is most reported as well as which year they were reported will be conducted in the future to understand trends in animal reports and if mitigation practices followed by 311 have been effective over the past ten years. Clustering is useful data technique that allows us to display the relationship between different data points and is a great use case for this goal.

Unanswered Questions

Existing datasets have been incredibly useful in a preliminary analysis of 311 animal reports. They provide the time and date of the reports as well as the coordinates of where the event took place, so creating time series and mapping out incidents is possible. Therefore, **there are not many questions we cannot answer at this time**. However, to collect the data, our team used a text file of all the animals present in the world. By iterating through the *closure reason* of the publicly available 311 service request data and filtering for “animal issues” as the request reason, we were able to calculate the number of reports for each animal in an effective manner. However, many of the rows in the service request data did not include a closure reason, therefore we could not extract an animal from thousands of rows of animal issues related data. So while we can answer several key questions, the amount of data used to create these insights is limited. Moving forward, if we can somehow discover more archived data that includes more detailed content in closure reason column, we can create graphs, maps, etc. that are backed more strongly with greater amounts of data.

4. Submit all of the following information (code, notebooks, answers to questions) as a PR to your team's branch on GitHub. (Add your TPM and TA as reviewers!)