City of Boston: Building Violations

By
Julissa Mijares and Tauriel Duan

Analysis Paper
Continuation of CS506 Spring 2024 at Boston University
BU Spark!

Introduction

Data Cleaning:

There are four main datasets used for this analysis. The first dataset is *Building and Property Violations*[1], which provides violations issued by inspectors from the Building and Structures Division of the Inspectional Services Department. The original dataset has 15963 rows and 24 columns. From this dataset, we dropped the columns 'value', 'contact_addr2', 'ap_case_defn_key', 'violation_sthigh', because they had many missing values or a repeated value. All rows not within the year range 2019-2024 were dropped as well. This leaves the cleaned dataset with 4431 rows and 21 columns.

Similar to the previous dataset, we also used *Public Works Violation*[2] dataset. The original dataset has 749294 rows and 23 columns. We dropped similar columns 'value', 'contact_addr2', 'violation_sthigh' and all rows not within the year range 2019-2024 were dropped as well. This leaves the cleaned dataset with 295331 rows and 21 columns.

The next dataset is *Live Street Address Management (SAM) Addresses*.[3] The original dataset has 400244 rows and 27 columns. From this dataset, we dropped the columns 'IS_RANGE', 'RANGE_FROM', 'RANGE_TO','UNIT', 'STREET_PREFIX', 'STREET_SUFFIX_DIR', because they had many missing values or a repeated value. This leaves the cleaned dataset with the same number of rows and 21 columns.

Finally, the last main dataset is *Property Assessment*[4]. This data was split up by fiscal year, and it was determined that we would keep the years 2019 through 2024.The columns keep, that could be found across all datasets, would include 'PID', 'OWNER', 'LAND_SF', 'GROSS_AREA', 'LIVING_AREA', 'LAND_VALUE', 'BLDG_VALUE', 'TOTAL_VALUE', 'YR_BUILT', 'YR_REMODEL', 'STRUCTURE_CLASS', 'OVERALL_COND', 'BED_RMS', 'TT_RMS', 'BTHRM_STYLE1', 'HEAT_TYPE', 'AC_TYPE', 'NUM_PARKING'. Each dataset had a unique number of rows and columns, but the combined dataset had a total of 1068278 rows and 18 columns.

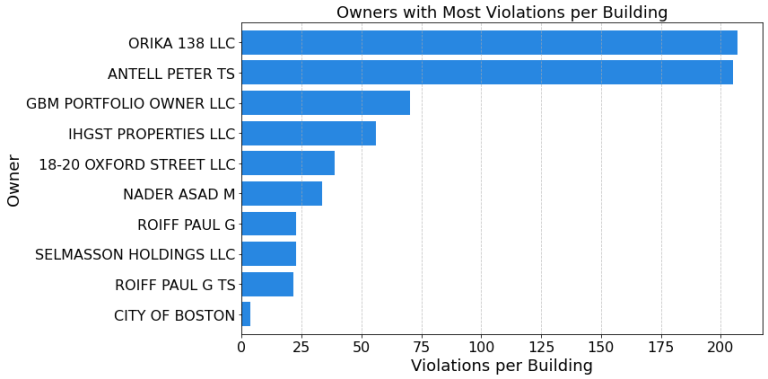
To answer the following base questions, the datasets would need to be merged together. The *SAM* dataset would act as the binder between the other two datasets, as it had the column 'SAM_ADDRESS_ID' which could be matched with 'sam_id' from the *Building Violations* dataset as well as the column 'PARCEL' which could be matched with 'PID' from the *Property Assessment* dataset. After merging all three datasets, repeating columns were dropped.

[1] https://data.boston.gov/dataset/building-and-property-violations1
[2] https://data.boston.gov/dataset/public-works-violations/resource/90ed3816-5e70-443c-803d-9a71f44470be
[3] https://data.boston.gov/dataset/live-street-address-management-sam-addresses
[4] https://data.boston.gov/dataset/property-assessment

Base Questions

Are there certain landlords/management companies that have repeated violations?

The range of unique violation counts by owner ranges between 0 and 1262 between 2019-2024. The management company with the greatest value by a significant amount is GBM PORTFOLIO OWNER LLC with a violation count of 1262. Approximately 99.70% of landlords/management companies have less than 100 violations. Another 0.25% have between 100 and 200 violations, leaving only 0.05% with over 200 violations. This corresponds to 10 unique owners with a high violation count within the last 5 years.



Information about many of these landlords is limited due to their status as individual owners rather than larger management companies. Additionally, LLCs represent five of the top 10 companies with the highest violations per building. Their popularity among small businesses for tax advantages and liability protection complicates pinpointing responsibility for these violations.

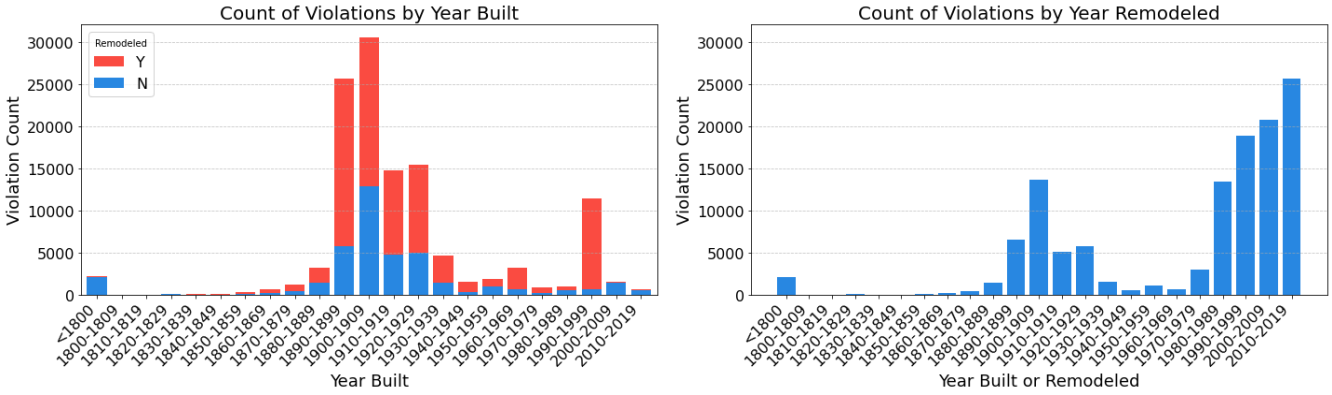
Peter Antell, who had a history of violations as an individual landlord, owned a building located at 140 Newbury St. He recently passed away in 2021, so some maintenance oversights may have resulted due to his age. Being situated on a busy street, any minor issues with the property were likely to be noticed and reported. The majority of Antell's violations consisted of improper trash storage. The current ownership status of the property is unknown publicly.

Another notable individual is Paul Roiff, who stands out with his appearance in past news articles. Roiff was targeted by housing-rights groups in a protest over what was perceived as an unfair eviction notice due to a rent increase in a multi-family home in East Boston.[5] He owns upscale properties such as the XV Beacon Hotel and the Teatro and Mistral restaurants. We also see Paul Roiff appear twice in the chart above, once under his name and another under TS, which we believe represents a trust.

[5] https://www.bostonherald.com/2010/02/23/tenants-lash-out-at-hub-business-owner/

Are there common features of certain buildings?

Two initial perspectives can be taken when examining the characteristics of buildings with the highest violation counts. The first involves considering the year of construction. As illustrated in the graph below, there is a significant spike in violations among buildings constructed in the early 1900s, indicating that these structures are approximately a century old. Nearly half of these buildings underwent renovations in the late 1900s or early 2000s, as depicted by the red section in the left graph.



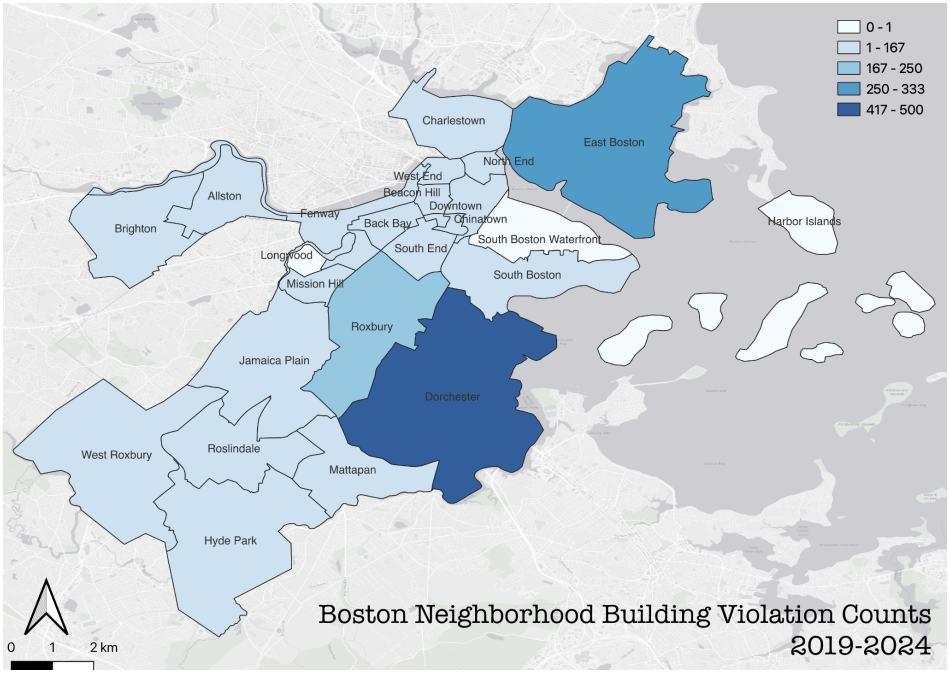
Considering both the year of construction and the timing of renovations, it appears that addressing the complaints associated with early 1900s buildings may necessitate more than renovation alone, potentially requiring demolition. However, with Boston being a significant historical hub, it is important to preserve as many of the older buildings as possible. This emphasizes the need for increased transparency with tenants regarding common complaints before their commitment. While many tenants value the historical architecture of these buildings, they may not fully consider the drawbacks inherent in older constructions.

Additionally, building regulations are continually evolving as the government recognizes the importance of accessibility and appropriate living conditions, whereas older buildings were constructed under different standards. With that said, it is difficult to decipher if it is a small landlord that is having difficulty keeping up with modern renovations versus a larger investment firm. Approximately 80.99% of the overall condition of these buildings are reported as *A* for Average.

During the period from 1990 to 1999, there was a notable increase in both new constructions and remodeling activities. A key factor behind this trend was the volatile housing market of the previous decade. In Boston, the housing boom of the 1980s drove demand for new constructions and remodeling to increase property values. Following the boom, the bust led homeowners to remodel rather than sell at

a loss. As the market stabilized in the 1990s, these activities continued, fueled by the earlier boom's momentum and improving market conditions.[6]

What neighborhoods/ communities are affected most?

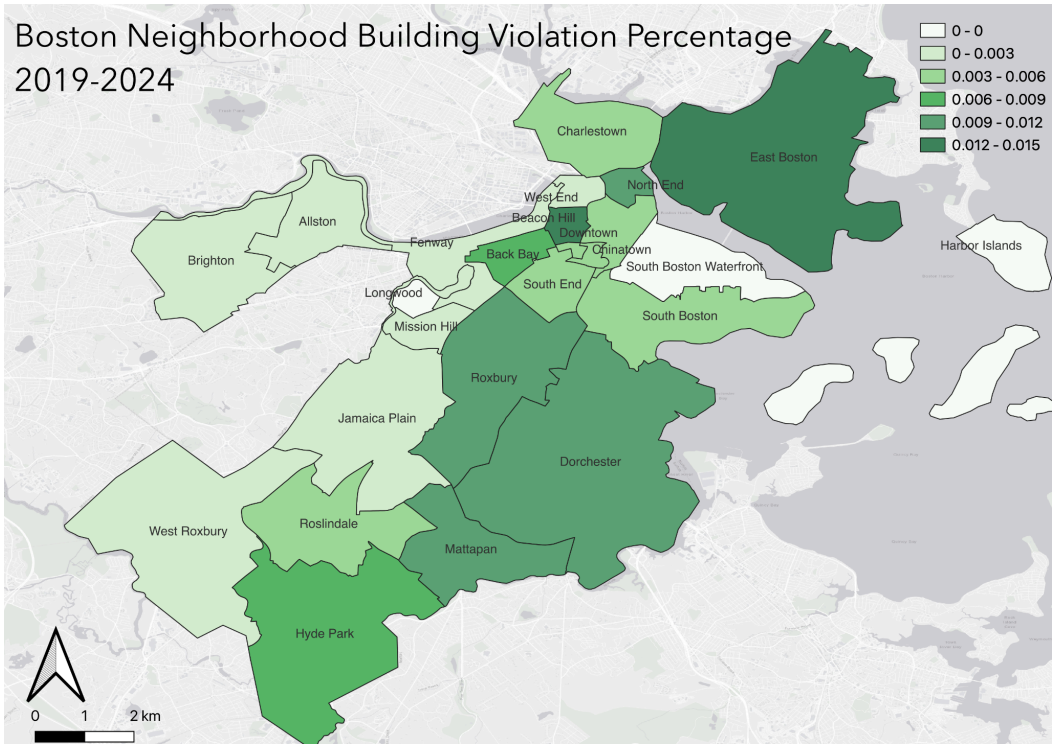


An equal interval method with 7 intervals to best provide visualization for the building violation counts in each neighborhood. Based on the reported violation counts, Dorchester is the neighborhood that is affected the most, with a building violation count of 500, followed by East Boston and Roxbury with violation counts of 270 and 236. The other neighborhoods all have counts below 100.

However, Harbor Islands, South Boston Waterfront, and Longwood are having counts of zero. Instead of using the neighborhood column from the building violation dataset, the X and Y coordinates of the reported buildings are extracted and put into the actual neighborhood polygons grouped by census blocks to ensure the accuracy of the map.

These results can not tell the whole story as Dorchester and East Boston, the two neighborhoods with the highest count are bigger and might contain more household buildings which can cost a larger number of reports. Therefore, a normalized map is needed to provide more insights to answer this question.

[6]https://www.bostonfed.org/publications/new-england-economic-review/1994-issues/issue-march-april-1994/a-decade-of-boom-and-bust-in-the-prices-of-singlefamily-homes-boston-and-los-angeles-1983-to-1993.aspx

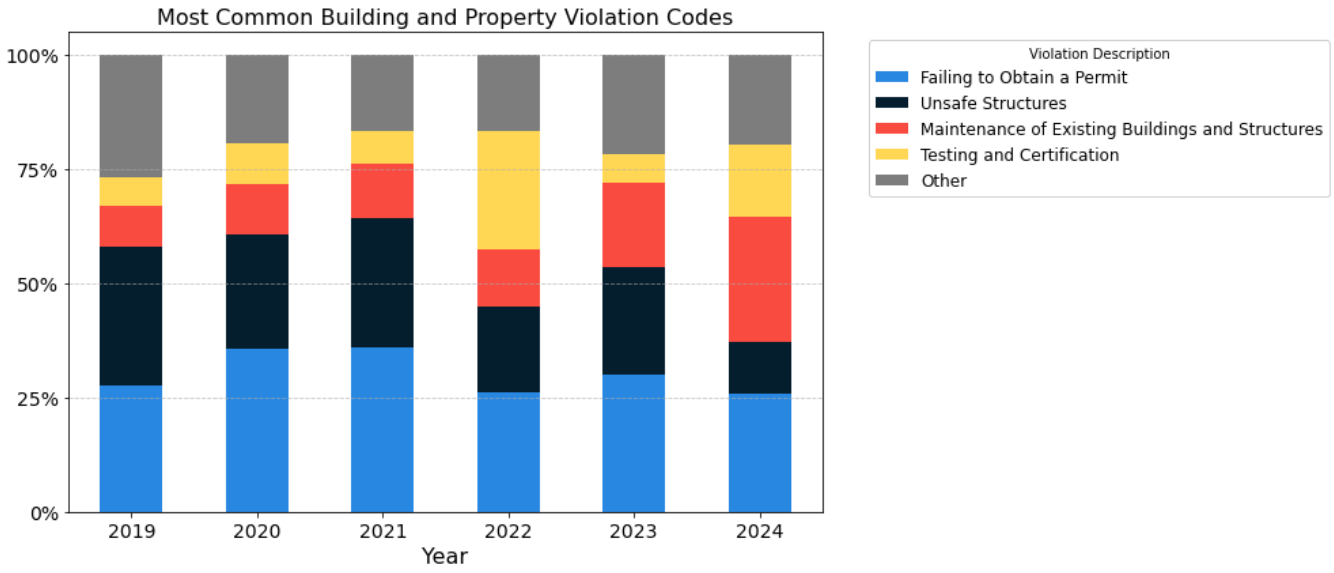


The normalization is done using the reported building violation count divided by the total household numbers in each neighborhood. With an equal interval method with 5 intervals, East Boston and Beacon Hill neighborhoods stand out with the highest percent building violation rates of 1.5% and 1.3%. Followed by North End, Roxbury, Mattapan, and Dorchester with 1.2%, 1.1%, 1.1% and 1.0%. The other neighborhoods all have a building violation rate under 1.0%.

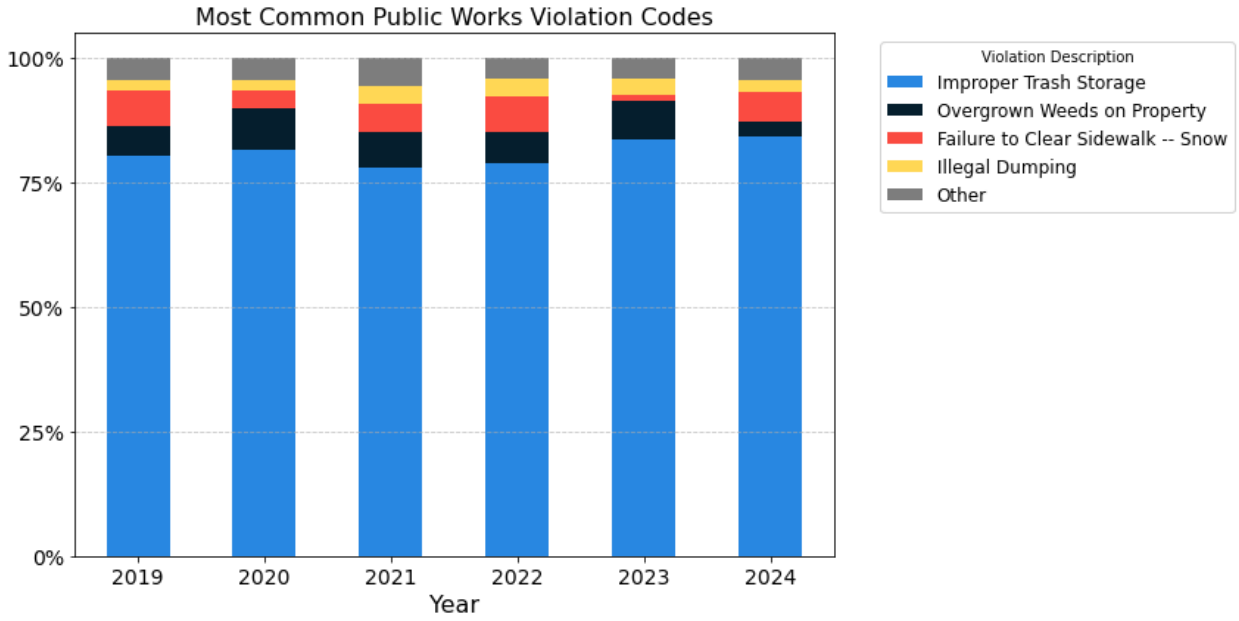
The results here confirmed the problem in East Boston but proved the innocence of Dorchester. East Boston has similar sizes but contains only 1/3 of the households compared to Dorchester, causing the higher violation rate. Beacon Hill and North End come up on the ranking probably because they have some of the oldest houses in the Boston region that are prone to having violations and need more attention, which corresponds with the answer to Question 2.

What kinds of building complaints are people making around the city?

The majority of building violations can be represented by 4 main categories of codes. The most frequent violation, accounting for the largest share, is code 105, which pertains to failing to obtain a permit. Next is code 116, concerning unsafe structures. Together, these two categories represent, on average, 53.77% of all violations. Additionally, code 102, which deals with the maintenance of existing buildings and structures, and code 1001, which covers testing and certification, are another 26.05%. The remaining 20.18% of violations are spread across all other categories.



The overall trend in violation types has remained relatively consistent over time, except for a notable spike in reports for violation code 1001 (Testing and Certification) in 2022. The largest concern that should be addressed is violation code 116 as it pertains to unsafe structures, which can lead to further health and safety issues.



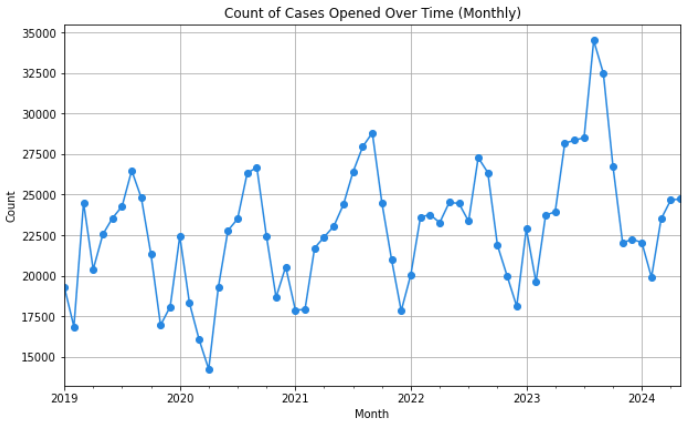
There is significantly less variation among the Public Works Violations. The majority of Public Works Violations can be explained by violation code 1 (Improper storage trash: res), which makes up, on average, 63.55% of the total violations made between 2019 and 2024. Other significant violations include codes 3 (Overfilling of barrel/dumpster: 12.26%), and 2 (Improper storage trash: com: 5.21%).

In the graph above codes 1, 2, and 3 are combined to represent improper trash storage. The second most common violation is 24 (Overgrown Weeds On Property) at 6.60%. The method of collection of violations might differentiate, which may help explain the lack of variation in the Public Works violations. It is easier for street teams to report trash and outside violations compared to a walk in building inspection that would need to be completed more thoroughly.

Each of the remaining violation codes individually contributes less than 5.00% to the total. However, when combining like groups, we also decided that codes 17a, 17b, and 17c were significant and represented a failure to clear the sidewalk of snow. This violation appears less during years that snow was not as prominent. And finally codes 9a, 9b, and 9c were combined to represent illegal dumping.

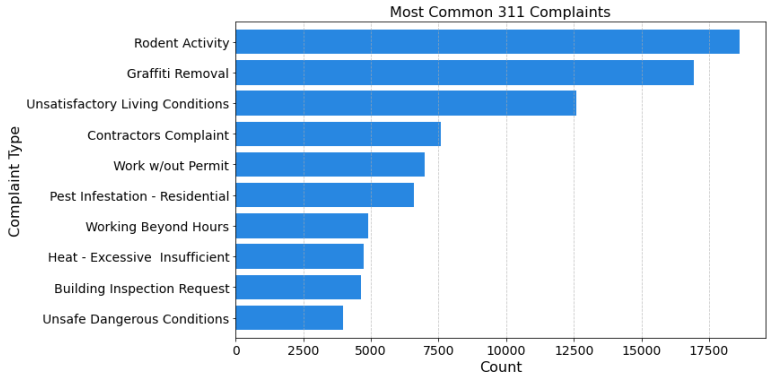
311 data

In our effort to better understand resident concerns, we decided to examine the Boston 311 data. It's important to realize that this data might not be a perfect reflection of what is happening because some things might not get reported with other complaints may be reported more often. Factors like how easy it is to report problems, how involved different communities are, and if certain groups report more than others can all affect what ends up in the data. By recognizing these factors, we're looking at the data with a clear understanding that it might not tell the whole story.



The majority of cases open during August and September, correlating with the return of students to the city and the prevalence of moving activities during the warmer seasons. In addition, the majority of Boston apartments begin their new lease on September 1st. Alternatively, the lowest volume of cases arises during the colder months, from November to February. This trend is likely influenced by individuals having settled into their chosen residences for the year during this time.

In order to find the most relevant complaints, we decided to filter on ISD (Inspectional Services), HS_O (Housing-Office of Civil Rights), and PROP (Property Management).



The graph above indicates that the top three complaints are rodent activity, graffiti removal, and unsatisfactory living conditions. Among these, unsatisfactory living conditions are the most concerning regarding building violations. This category, which includes various issues, has around 12500 self-reported complaints. Therefore, these complaints should be closely monitored to enhance overall living conditions in Boston.

Conclusion

In conclusion, this analysis of building violations in Boston from 2019 to 2024 has revealed several insights that need attention.

First, a small number of landlords and management companies, such as LLCs, are responsible for a large number of violations. More investigations should be done on these parties to improve the situation further. Second, buildings constructed in the early 1990s and renewed in the late 1900s and early 2000s are more prone to having violations, indicating the potential problem of managing older buildings. Third, East Boston, Beacon Hill and North End neighborhoods have the highest building violation complaints rate, which could be due to them having the oldest buildings that need more maintenance. Fourth, the most common types of violations relate to failure to obtain permits, unsafe structures, and improper maintenance.

Finally, an examination of Boston 311 data shows that rodent activity, graffiti removal, and unsatisfactory living conditions are the top complaints from residents. By understanding these patterns and trends, the city can better target its enforcement efforts and work with property owners to improve living conditions for all Bostonians.