

**Bad Landlords II - Councilor Breadon  
Team 4**

Maria Shevchuk ‘23 (maria410@bu.edu)  
Eddie Jones ‘24 (ewj327@bu.edu)  
Aaron Liauw ‘24 (aliau@bu.edu)  
Francesco Ciruolo PhD 1 (fciruolo@bu.edu)

**Table of Contents**

1. Introduction.....2

2. Base Analysis .....2

    2.1. Initial Data Analysis .....2

    2.2. Scofflaw Property Owner List ..... 3

    2.3. Dashboard.....5

    2.4. Challenges and Future Suggestions .....5

3. Extension Analysis.....5

4. Member Contributions.....6

## 1. Introduction

In collaboration with councilor Breadon, we are performing data analysis on Boston's property and renting datasets to determine which landlords in the area are so-called "bad landlords". The stated aim of the project is to build a trackable system for property violations throughout the Boston area and to create a matrix that determines whether a given landlord is a "bad landlord." Finding a specific ordinance that can be used to define a threshold for a "bad landlord" (also called a "scofflaw" landlord) was another stated goal of the project. The overall motivation and impact of the project is to increase transparency and accessibility for the housing development process in the greater Boston area so that the public can make informed decisions about where to live and understand the shifting landscape of their city.

## 2. Base Analysis

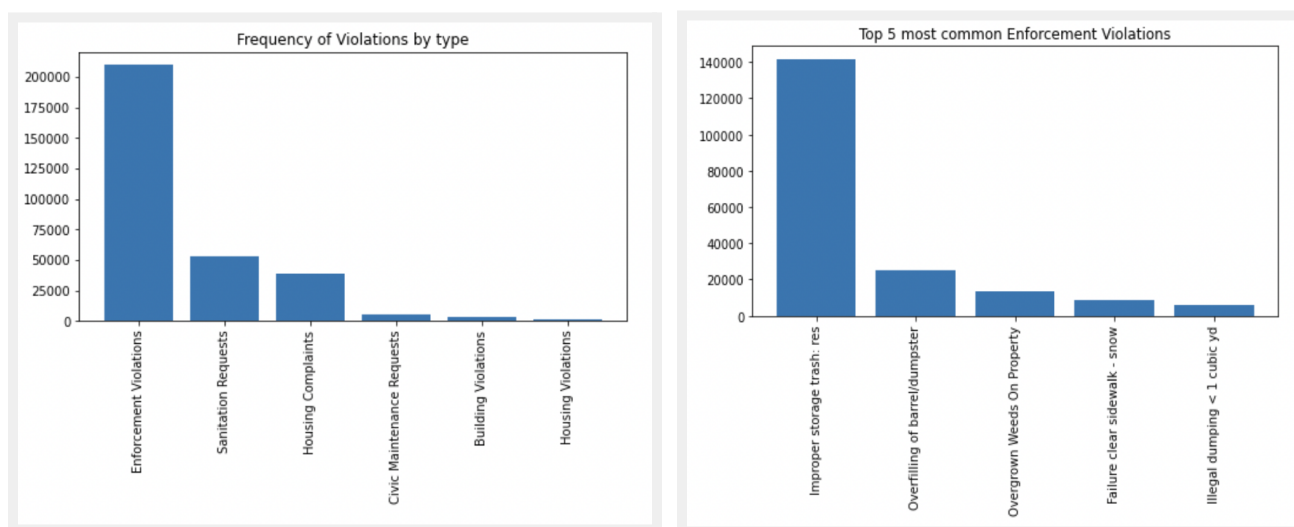
### 2.1. Initial data Analysis

Our initial step when starting the project was to familiarize ourselves with the results and findings from the Fall semester project. At this point, we were posed the following key questions:

1. What are the top 3-5 most common violation types from the fall project?
2. Can we verify these findings with the data available to us?

After consulting some of the deliverables from the Fall semester, we concluded that the most common violation types they outlined were Enforcement violations, followed by Housing complaints and Sanitization requests, with most violations classified in the 'low severity' category. (Source: [team 2 deliverable 4](#), [team 4 deliverable 5](#))

Based on our preliminary analysis, we were able to draw the following conclusions: Our findings from the [RentSmart](#) dataset analysis agree with those of team 4 (deliverable 5). We found Enforcement Violations to be the most common type of violation, followed by Sanitation requests and Housing Complaints (figure on the left). Within the Enforcement Violations, the most common violation is for improper trash storage (figure on the right).



After consulting with the client, we learned that the violations we are interested in do not include any violations enforced by the city/public or not related to a specific unit. Upon receiving this feedback, we changed our strategy, and identified the “Housing Violations” in the RentSmart dataset to be our primary category of interest. In this violation category, the most common violations are “Smoke Detectors & Carbon Monoxide Alarms”, “Owners Responsibility to Maintain Structural Elements”, and “Owners Installation/Maintenance Responsibility”.

According to the [Building and Property Violations](#) dataset, the most common violation for fitness of “human habitation” by the Department of Public Health is violation CMR410.500, “Owner's Responsibility to Maintain Structural Elements;” it is followed by CMR410.351 “Owner's Installation and Maintenance Responsibilities.” Other notable ones are CMR410.550 and CMR410.482, which are “Extermination of Insects, Rodents and Skunks” and “Smoke Detectors and Carbon Monoxide Alarms” respectively. Because the dataset focuses on buildings and property, it makes sense that maintenance takes up a large number of violations, especially in lower-income communities.

	code	case_count
58	CMR410.500	2472
30	CMR410.351 A	1379
70	CMR410.550 B	935
53	CMR410.482 A	794
52	CMR410.481	263
9	CMR410.200 A	214
82	CMR410.602 D	190
11	CMR410.201	174
8	CMR410.190	157
59	CMR410.501 A	149
45	CMR410.451	131
79	CMR410.602 A	125

After completing initial preliminary analysis, we were provided with additional key questions:

1. What is the spectrum of violations and severity in regards to worst landlords classifications?
2. What landlords are non-compliant? Overall volume, severe violations
3. Which landlords are impacting vulnerable communities the most? Is there a correlation between vulnerable communities and violation frequency?

Initially, answering the first question posed some challenges since Boston does not currently have a classification or categorization system for violations and their severity. We brought this concern to the client’s attention and were later provided with an ordinance - a working document highlighting specific violations of interest in regards to worst landlords classification. Please refer to section 2.2 (Scofflaw Property Owner List) for continued exploration and reworking of questions 1 and 2.

As part of our extension project, we plan to tackle the third key question, investigating whether there is a relationship between vulnerable communities (based on social vulnerability index, climate resilience, and college students) and housing violations. It is important to note that, as instructed by the client, all properties owned by the city of Boston are omitted from the data analysis. Please refer to section 3 (Extension Analysis) for further insight into our exploration of key question 3.

## 2.2. Scofflaw Property Owner List

After receiving the [ordinance](#) and scofflaw landlord definition from the city council, we have adjusted our final goal to better fit the given criteria. In summary, the ordinance identified 6 violations as the threshold for a landlord being considered ‘scofflaw’ and outlined 3 specific violation codes that need to be considered: State Building Code (780 CMR), the State Fire Code (527 CMR), and the State Sanitary Code (105 CMR 410).

At this point in our project, new key questions were posed:

1. Since data merging is not possible at this time, which available dataset contains the most relevant data to identify scofflaw property owners based on the ordinance?
2. What features, in addition to the scofflaw property owner list, could provide useful insights for tenants and future renters?

Since we don’t have access to a dataset that contains all violation codes of interest in addition to landlord/owner names, we decided to work with the housing violations in the RentSmart dataset, which seemed to be of most relevance to our new goal. From this dataset, we are able to extract various housing violations including fire safety and sanitation related violations. One of the advantages of this dataset is the fact that it includes the landlord/owner information for each property, while many others don’t. On the other hand, a downside of the RentSmart dataset is the violation category labeling, which is inconsistent with standardized city violation codes and, thus, inconsistent with the codes outlined in the ordinance. We have been able to identify, however, that the housing violations subsections in the dataset are largely representative of all 3 state codes of interest.

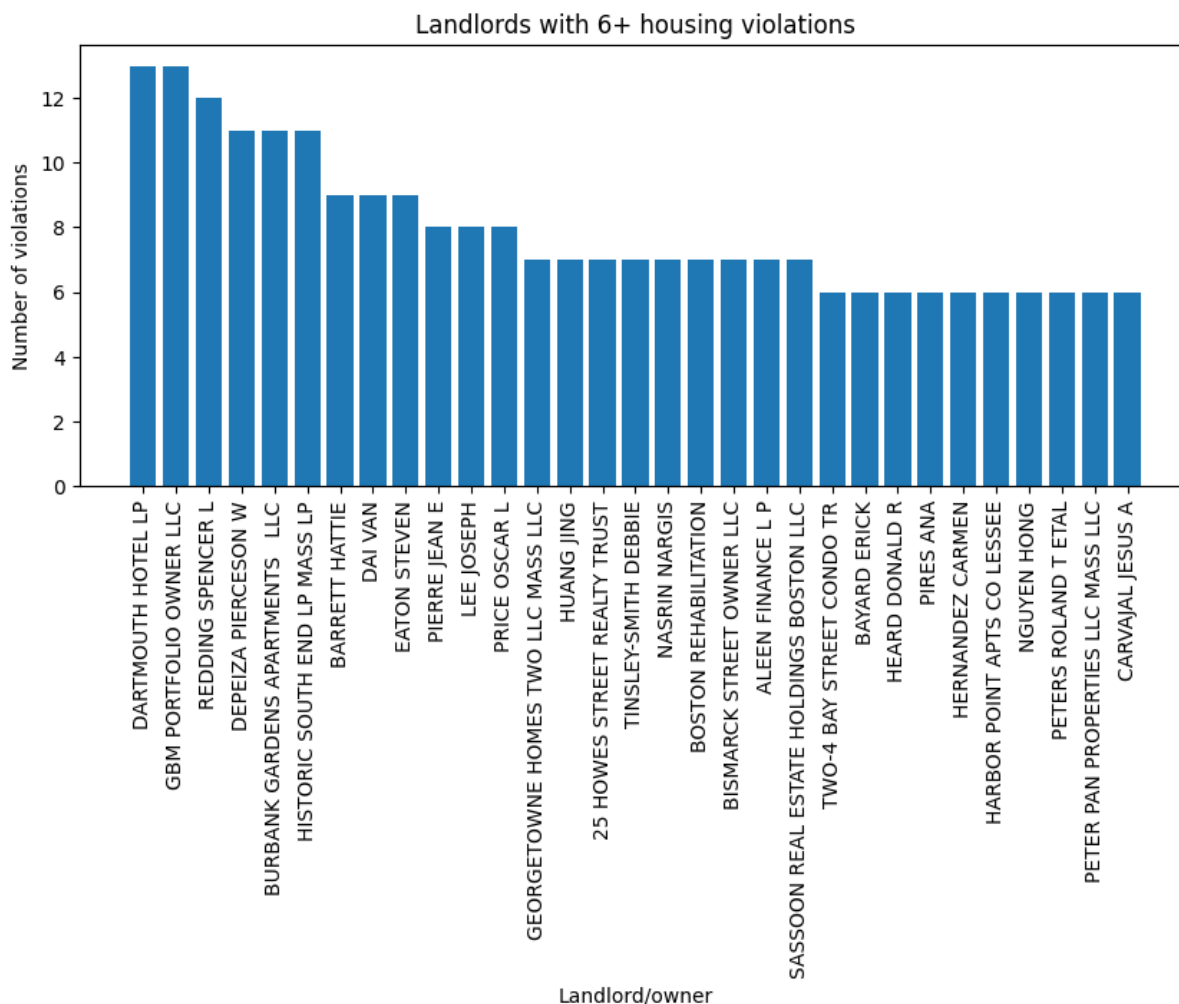
Thus far, we have been able to match each subsection in the Housing Violations table of the RentSmart dataset to a CMR code. We have identified that the housing violations category in the dataset relates specifically to the State Sanitary Code (105 CMR 410). Our code matching can be found in this [spreadsheet](#) and a preview is demonstrated in the figure below.

Subcategory in RentSmart's Housing Violations	Code of Massachusetts Regulations (CMR) Number
Smoke Detectors & Carbon Monoxide Alarms - Owner shall provide, install and maintain in operable condition smoke detectors and carbon monoxide alarms	105 CMR 410.482: Smoke Detectors and Carbon Monoxide Alarms
Owners Responsibility to Maintain Structural Elements - Structural elements shall be maintained free from holes, cracks, loose plaster, or other defects.	105 CMR 410.353: Asbestos Material
Owners Installation/Maintenance Responsibility - All Owner-installed optional equipment	105 CMR 410.351: Owner's Installation and Maintenance Responsibilities

We plan to explore whether any of the other violation categories in the RentSmart dataset correspond to the CMR codes of interest highlighted in the ordinance. We suspect that the Building Violations category may correspond with the State Building Code (780 CMR). We believe that this code matching can be very useful for future work on this project!

Keeping the aforementioned limitations in mind, we proposed a solution utilizing the RentSmart dataset to the client. We clarified that, at this stage, our plan would be to focus exclusively on housing violations, and that we would implement the proposed threshold of 6 violations to compile a list of scofflaw landlords/owners. Our proposal was approved by the client and they are currently working on finding a way to merge existing datasets that could be used to further improve this project.

The figure below shows the landlords/owners that exceed the proposed threshold of 6 housing violations along with the violation count.



Along with this figure and a list of identified scofflaw landlords, we plan to add a search feature to our webpage that will allow the user to select a landlord from a dropdown list and view their violation history along with which specific properties those violations are associated with.

## 2.3. Dashboard

As part of our project, we created an interactive dashboard to allow the users to consult the data from the different sources and in order to allow them to have a bigger picture when seeking a place to rent in Boston.

The dashboard is published, not updated with the last improvements, at the following address <https://francescociraolo.github.io/bad-landlords-team4/>

## 2.4. Challenges and Future Suggestions

As discussed in multiple occasions with the PM, in the first presentation to the client and eventually during the client meeting last week, there are two issues, with no easy solution, that prevent us from achieving relevant results:

- several datasets, curated by the town or in general by the public administrations, are provided but they are not mergeable in any way;
  - this is true also for the ones derived from other ones; as the rents mart based on the 311 complainants
  - the datasets are also inconsistent with each other
- the expected and planned goals of classifying the landlords based on number and relevance of their violations is impossible since no dataset contains the required data

After several calls to support and promised tools, we had to clarify how the goal is not reachable and decide with client and PM how to direct the efforts toward the creation of a reasonable product.

## 3. Extension Analysis

For the extension analysis, we plan to review the Climate Ready Boston Social Vulnerability dataset and see if there's a higher correlation with any of the predictors (disabilities, low income, english proficiency, POC, children, older adults) and the response variable (medical illness). In essence, we'd like to perform an inference task — opposed to prediction — to find the most influential factors to medical illness. The reason we chose medical illness as a response variable is because the Boston council is most interested in CM410 violations, which are public health violations. By finding these factors, we could potentially find what communities are most vulnerable to Bad Landlords and violations.

### Key Questions

1. What are the most influential demographic predictors for medical illness?
2. Which groups are more vulnerable to property violations?

It is important to note that results of the analysis are not prescriptive — a.k.a. they do not perform methods of scientific inquiry and testing in the real world — so we should not infer anything about the results of the analysis or anything about these demographic groups. However, they could shed some light on what groups we might want to consider, for example, when tracking resource allocation or vulnerability.

#### 4. Member Contributions

Although all members have been core contributors to this project throughout the semester, below is a more detailed breakdown of each member's contribution.

Maria Shevchuk - cover page and sections 2.1, 2.2 (report), repo readme(s), any analysis related to *RentSmart* dataset, CMR code matching for *RentSmart* dataset

Eddie Jones - section 1 (report), any analysis related to *Income-Restricted Housing Inventory*

Aaron Liauw - extension project section 3 (report), analysis related to *Building and Property Violations Dataset* and *Climate Ready Boston Social Vulnerability*

Francesco Ciralo - Sections 2.3 and 2.4 of the report, datasets analysis and cleanup, dashboard creation, optimization and improvement, partial PM and client interaction