
Bad Landlords Deliverable 1

Spring 2023 CS506 Data Science

1 Data Analysis

This part shows what we have analyzed about the datasets.

1.1 Violation types

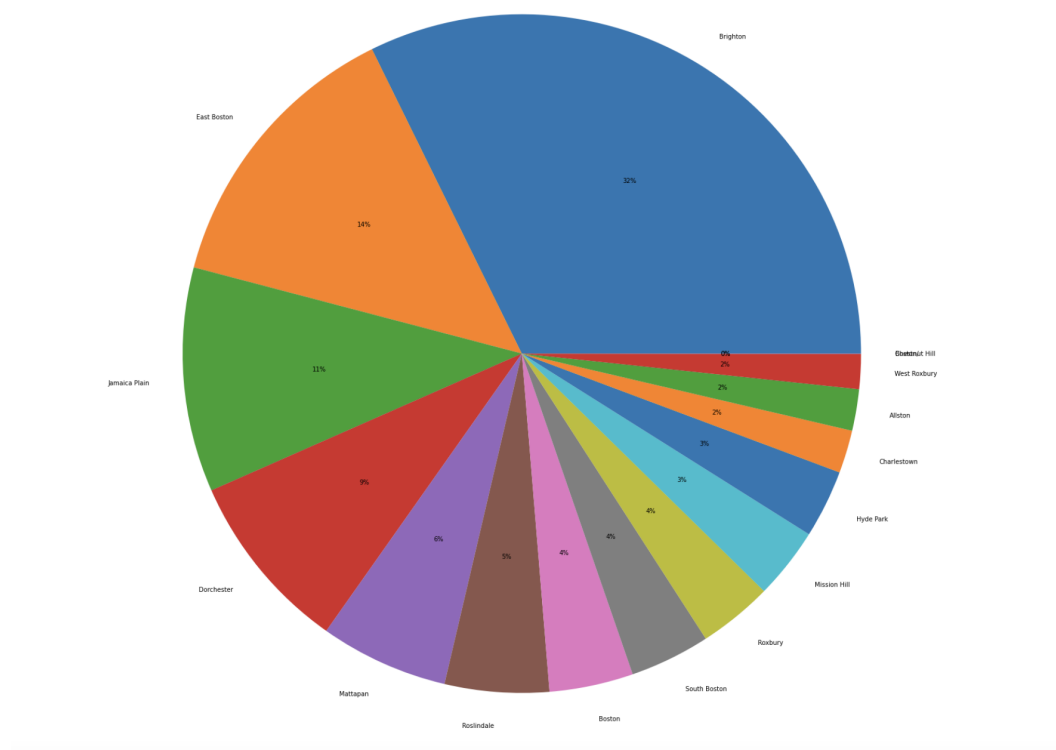
We've analyzed the [Building and Property Violations Dataset](#) and give a preliminary violation types of all landlords violations.

The columns of this dataset are as follows.

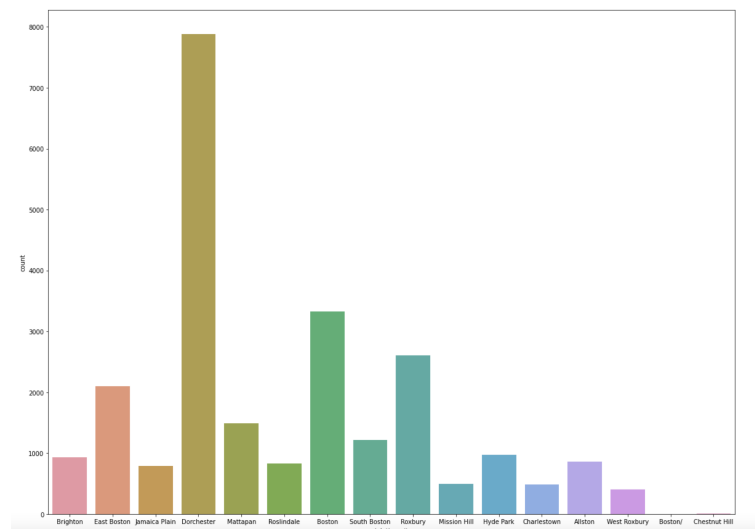
```
Index(['case_no', 'status_dttm', 'status', 'code', 'value', 'description',  
      'violation_stno', 'violation_shigh', 'violation_street',  
      'violation_suffix', 'violation_city', 'violation_state',  
      'violation_zip', 'ward', 'contact_addr1', 'contact_addr2',  
      'contact_city', 'contact_state', 'contact_zip', 'sam_id', 'latitude',  
      'longitude', 'location'],  
      dtype='object')
```

2 Preliminary analysis of data (Visualization)

The preliminary batch of data that was selected to be pre-processed was the building and property violations dataset. To find the main locations in and surrounding Boston with the most number of violations, the 'violation_city' column of the CSV file was cleaned by removing all of the null values. Then the results were plotted as a pie chart indicating violation frequency by location shown as follows:



In addition to the frequency, the violation counts by location were also plotted as a bar graph shown as follows:



Additionally, the most common types of violations were extracted and counted from the dataset which yielded the following visual:

```

Failure to Obtain Permit
3431
Unsafe and Dangerous
3385
Owners Responsibility to Maintain Structural Elements - Structural elements shall be maintained free from holes, cracks, loose plaster, or other defects.
2505
Owners Installation/Maintenance Responsibility - All facilities and equipment which are required by owner including but not limited to water, gas,
electrical and heating, shall be installed in accordance with all accepted codes. 1386
Maintenance
1123

```

The failure to obtain a permit seems to be the most common type of violation followed very closely by unsafe and dangerous housing conditions.

3 Key Questions

Q: What is the spectrum of violations and severity in regards to worst landlords classifications?

A:

Now we are targeting at sorting the buildings' violations number to get the worst buildings. The next step we'll take is to combine all the buildings to their owner landlords. Here are the results.

	ST_NAME	violation_city	violation_zip	violation_counts
474	Chelsea ST	East Boston	02128	182
2043	Saratoga ST	East Boston	02128	132
10	Adams ST	Dorchester	02122	114
448	Centre ST	Jamaica Plain	02130	114
693	Devon ST	Dorchester	02121	110
2190	Sumner ST	East Boston	02128	100
1862	Princeton ST	East Boston	02128	96
330	Broadway	South Boston	02127	94
166	Beacon ST	Boston	02116	92
210	Bennington ST	East Boston	02128	92

Q: What is the number of eviction filings based on location?

A: Answering this question was difficult as there was no obvious place to find a reliable and easily accessible dataset. Most datasets and reports we find are about the whole massachusetts. And it's hard to find the detailed data of each neighborhood in Boston. The most recent data we could find was during the 2020-2021 year and the result of average eviction filings by location was as follows:

```
GEOD
25025082100    655.75
25025130406    580.50
25025100500    420.00
Name: filings_avg, dtype: float64
```

by referring to the map we find that the top3 are 821, Suffolk County, 1304.06, Suffolk County 1005, Suffolk County

We will try to do some more research in terms of evictions and housing court cases in the coming weeks as that data seems to be scarcely available.