BAD LANDLORDS

Team 2

Harshil Gandhi
Abhishek Tiwari
Longdan Mao
Di Wang

Report

Summary:

This extended analysis aims to provide a comprehensive understanding of housing complaints, evictions, and related factors across neighborhoods by examining data from multiple sources. By identifying the top owners with the most complaints, merging demographic data with complaint information, calculating and visualizing complaint density, incorporating eviction data, and creating focused DataFrames for high eviction rate neighborhoods, the analysis seeks to uncover the underlying factors influencing housing stability and landlord-tenant relationships. This information can be leveraged to develop targeted interventions, inform policy decisions, and ultimately improve housing conditions in affected neighborhoods.

Key Findings:

1. The top 10 owners with the most complaints have been identified and visualized as a horizontal bar graph. This information could be helpful in targeting problematic landlords and understanding the primary reasons behind these complaints.

2. Demographic data has been merged with the number of complaints per neighborhood, providing a richer dataset for analysis. This can be used to examine the relationship between population characteristics and the number of complaints.

3. The number of complaints per population has been calculated for each neighborhood, which can help identify areas with a high density of complaints relative to their population size. This metric can be useful in prioritizing interventions in neighborhoods with the most pressing issues.

4. A bar graph of complaints per population for each neighborhood has been plotted, offering a visual representation of the complaint density across neighborhoods. This visualization can help stakeholders identify problem areas and focus their efforts on addressing these issues.

5. Eviction data has been incorporated into the analysis, providing a deeper understanding of the housing situation in each neighborhood. The annual eviction filing

rate can be used to gauge the stability of housing in each neighborhood and may be correlated with the number of complaints.

6. A new DataFrame called `per_neighborhood_high` has been created with only rows that have an eviction rate greater than 4. This subset of data can be useful in focusing on neighborhoods with the most pressing housing issues and understanding the factors that contribute to high eviction rates.

7. The `per_neighborhood_high` DataFrame has been enriched with demographic data and educational attainment information. The resulting `per_neighborhood_high_edu` DataFrame can be used to explore how factors such as education level influence complaint rates and eviction rates in neighborhoods with high eviction rates.

8. It was found that neighborhoods with a large percentage of population between ages 20-34 have a lower eviction rate, this can be attributed to people in this demographic group having stable jobs and comparatively higher levels of education.

Next Steps:

1. Investigate the relationship between educational attainment levels and complaint rates or eviction rates. This could help determine whether there is a correlation between a population's education level and their likelihood of experiencing housing issues.

2. Identify the most common types of complaints in neighborhoods with high eviction rates to understand the specific challenges faced by residents in these areas. This information could be used to develop targeted interventions and policies to address these issues.

3. Analyze the relationship between income levels or employment status and complaint rates or eviction rates. This could provide insights into how economic factors influence housing stability and the quality of landlord-tenant relationships.

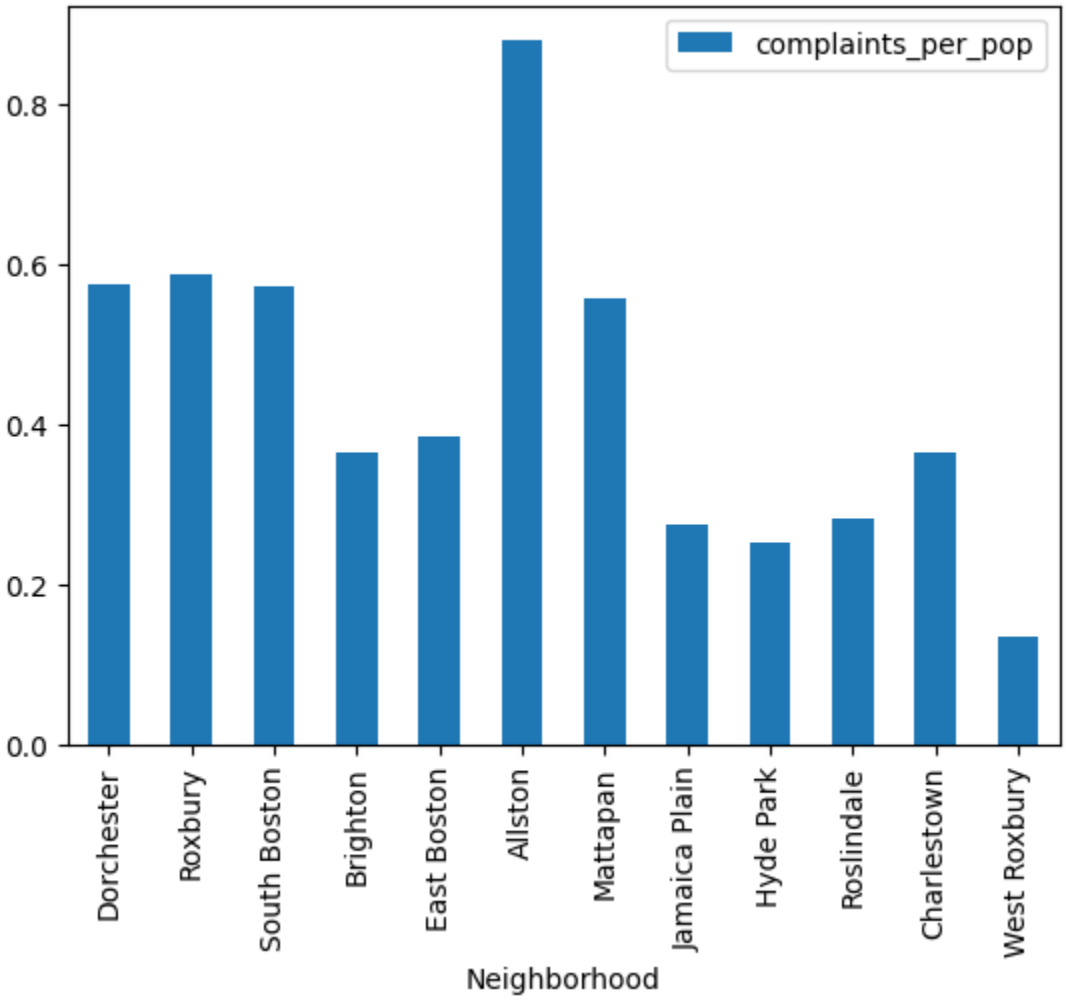
4. Conduct a longitudinal analysis to assess changes in complaint rates, eviction rates, and demographic characteristics over time. This could help identify trends and determine whether interventions have been effective in improving housing conditions in problem neighborhoods.

5. Examine the role of property management companies or housing associations in neighborhoods with high complaint rates or eviction rates. This could help determine

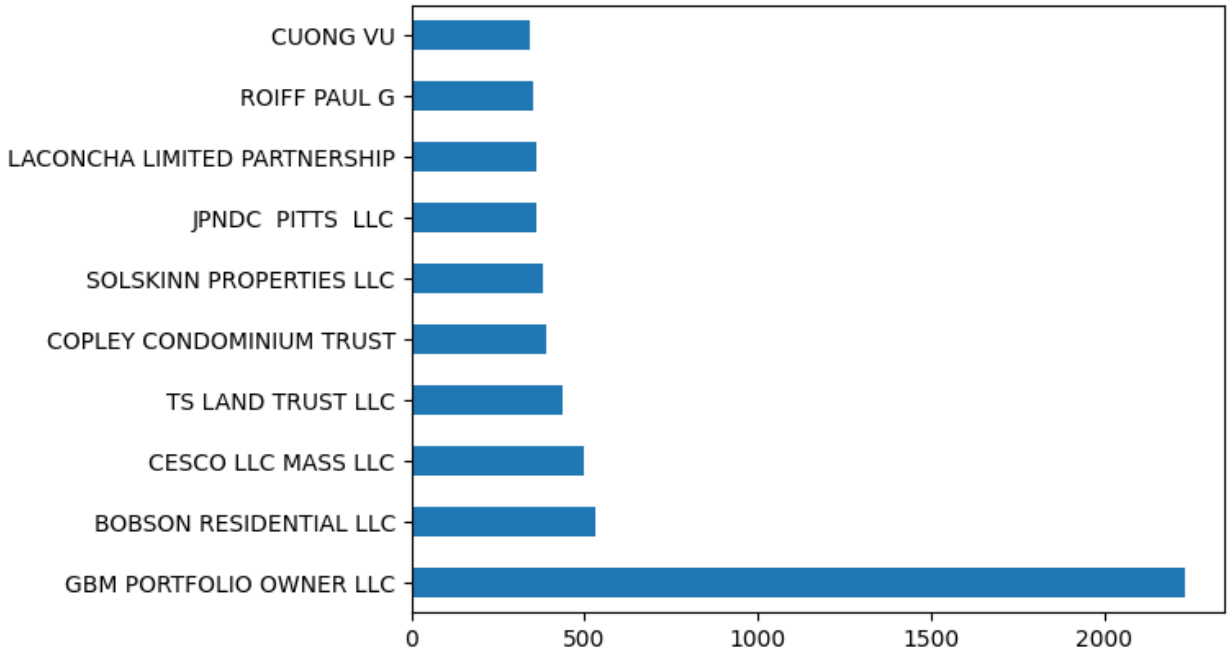
whether certain organizations are more effective at addressing housing issues and could provide examples of best practices for improving landlord-tenant relationships.

6. Develop predictive models to identify neighborhoods at risk of experiencing high complaint rates or eviction rates in the future. This could assist policymakers in proactively addressing housing issues before they become critical problems.

Visualizations



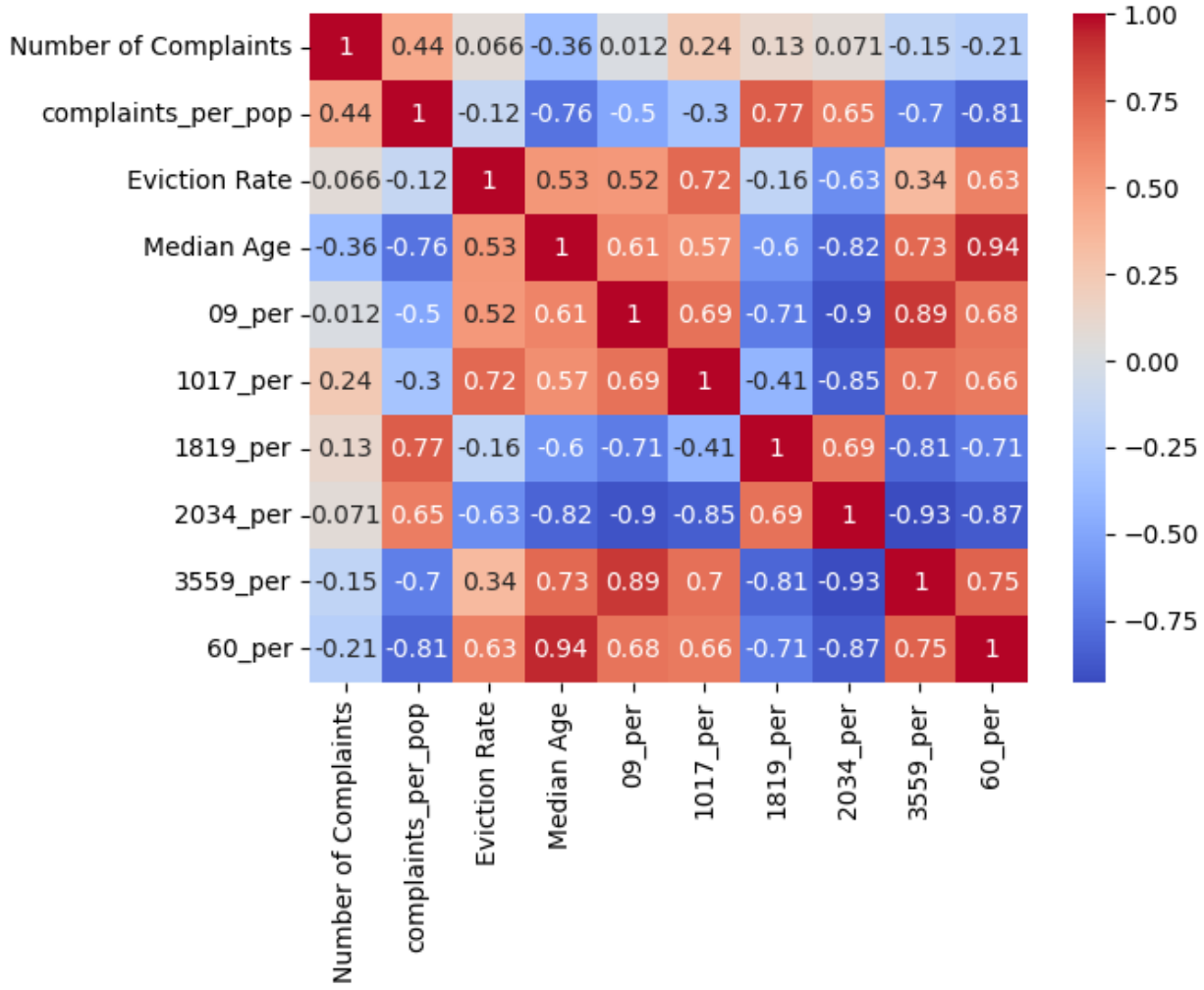
The above graph plots the neighborhoods on the x-axis and complaints per population unit on the y-axis. It can be seen that Allston has the highest complaints per unit of population.



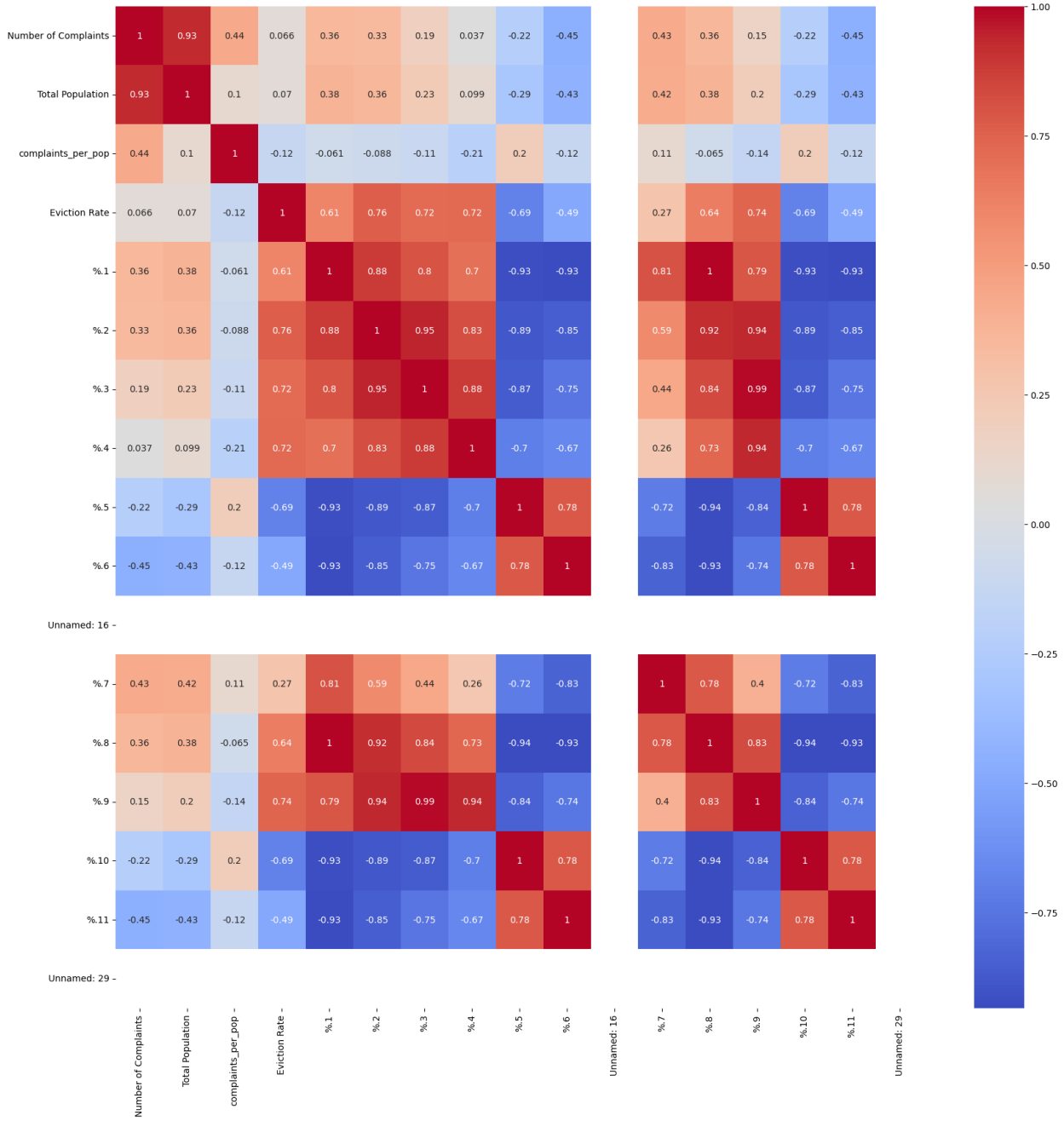
The above graph has number of violations on the x-axis and management companies on the y axis.

	Neighborhood	Number of Complaints	Total Population	complaints_per_pop	Eviction Rate
1	Roxbury	31890	54161	0.588800	6.7
6	Mattapan	14907	26659	0.559173	6.8
11	West Roxbury	4508	33526	0.134463	6.7

As a part of earlier EDA we found that the above 3 neighborhoods have the highest eviction rates, as a part of the extension we are trying to find if certain factors affect eviction rates.



While looking for correlations between eviction rates and demographic parameters, it was found that neighborhoods with a large percentage of population between ages 20-34 have a lower eviction rate, this can be attributed to people in this demographic group having stable jobs and comparatively higher levels of education.



When exploring the relationship between education levels and eviction rates, we see that neighborhoods where the majority of residents have a college degree have a lower rate of eviction and also complaints.

Questions:

1. What is the relationship between demographic factors, such as population characteristics or educational attainment, and the number of complaints or eviction rates in different neighborhoods?

2. Which property owners are associated with the highest number of complaints, and what are the primary reasons behind these complaints?

3. How does the density of complaints per population vary across neighborhoods, and which areas have the most pressing housing issues that may require prioritization in terms of intervention?

4. Are there correlations between income levels or employment status and complaint rates or eviction rates, providing insights into the influence of economic factors on housing stability and the quality of landlord-tenant relationships?

Answers:

1. The report provides insights into the potential relationships between demographic factors and the number of complaints or eviction rates in different neighborhoods. However, further analysis is required to determine the extent and nature of these correlations.

2. The top 10 property owners with the most complaints have been identified, offering valuable information for targeting problematic landlords. The primary reasons behind these complaints need further investigation to develop targeted interventions.

3. The report highlights variations in the density of complaints per population across neighborhoods. Some areas have more pressing housing issues, but the specific neighborhoods would need to be identified and assessed for prioritization in terms of intervention.

4. The report suggests that income levels and employment status may be correlated with complaint rates or eviction rates, but further analysis is needed to explore the influence of these economic factors on housing stability and the quality of landlord-tenant relationships.

Tasks:

Harshil Gandhi:
1. Merge demographic data with the number of complaints per neighborhood and calculate complaints per population for each neighborhood.
2. Plot a bar graph of complaints per population for each neighborhood.
3. Investigate the relationship between educational attainment levels and complaint rates or eviction rates.
4. Analyze the relationship between income levels or employment status and complaint rates or eviction rates.

Abhishek Tiwari:
1. Read and process data from multiple CSV files containing information about neighborhoods, owners, complaints, demographic data, eviction filings, and educational attainment.
2. Identify the top 10 owners with the most complaints and visualize them as a horizontal bar graph.
3. Incorporate eviction data into the analysis and calculate the annual eviction filing rate for each neighborhood.
4. Create a new DataFrame called `per_neighborhood_high` with only rows that have an eviction rate greater than 4 and enrich it with demographic data and educational attainment information.
5. Develop predictive models to identify neighborhoods at risk of experiencing high complaint rates or eviction rates in the future.

Longdan Mao:
1. Identify the most common types of complaints in neighborhoods with high eviction rates.
2. Conduct a longitudinal analysis to assess changes in complaint rates, eviction rates, and demographic characteristics over time.
3. Examine the role of property management companies or housing associations in neighborhoods with high complaint rates or eviction rates.

Di Wang:
1. Create a map visualization of the neighborhoods with the number of open complaints and overdue complaints.
2. Merge the eviction data with the complaints data and create a scatter plot of open complaints vs. eviction rates.
3. Perform additional data cleaning and preprocessing tasks as needed, such as removing duplicates or handling missing data.

4. Assist the team members in their tasks and collaborate on the development of the final comprehensive report.