Final Report:

The Impact of Remodeling and Zoning Conversions on Boston's Housing Market

April 30, 2024

2024 Spring

CS506: Data Science Tools and Applications

Jingbo Wang: jw6347@bu.edu

Yuchen Cao: caoyc@bu.edu

Jialong Ke: jlke@bu.edu

Zihan Li: lizihan@bu.edu

TABLE OF CONTENTS

1. INTRODUCTION

The rapid evolution of urban landscapes raises complex questions about the balance between development and the preservation of community character and accessibility. In Boston, a city with a rich historical backdrop and a vibrant, diverse populace, the intersection of remodeling, zoning conversions, and housing market dynamics presents a unique opportunity to study these issues in depth. This report stems from an ongoing investigation into the various factors reshaping Boston's housing landscape, focusing on how remodeling and zoning changes are influencing housing availability and community demographics.

The primary concern of this analysis is the examination of how remodeling activities and zoning conversions impact the availability of housing units in Boston. This study specifically addresses the correlation between these urban development activities and the influx of higher-income individuals into the city, exploring whether such trends contribute to a reduction in overall housing stock by transforming multi-unit dwellings into larger, but fewer, single-family homes. The overarching aim is to provide nuanced insights into how these trends affect local communities, particularly concerning housing availability and demographic transformations.

Initial findings suggest that remodeling and zoning activities are not merely physical transformations of property but catalysts for significant social and economic shifts within urban neighborhoods. By analyzing data from various sources, including the Live Street Address Management (SAM) Addresses database and approved building permits, our research delves into the patterns of housing construction, loss, and renovation from 2009 to 2024. This longitudinal approach helps trace the trajectory of development trends and their broader implications on the

city’s socio-economic landscape.

Furthermore, the study engages with the "Property Assessment - Dataset" to assess the outcomes of approved building projects, providing a detailed account of how changes in housing structures align with market demands and regulatory frameworks. The data-driven analysis aims to uncover the underlying motives and consequences of urban remodeling and rezoning, shedding light on how these practices align with or diverge from the community's needs for sustainable and equitable housing.

As this report unfolds, it will explore the annual changes in the number of remodeling activities in various regions, the situation of units lost each year in these areas, the impacts of remodeling activities and unit loss, and the factors that may influence both remodeling activities and unit loss.

In conclusion, this comprehensive study not only highlights the immediate effects of remodeling and zoning on housing availability but also contributes to broader discussions about urban development policies and their implications for community sustainability. This report is dedicated to a comprehensive analysis and detailed exploration of the trends in remodeling activities and unit losses within the Boston area over recent years. Through this in-depth examination, the aim is to provide a nuanced understanding and a robust assessment of the prevailing conditions and dynamics characterizing the Boston region, thereby offering insights that are crucial for informed decision-making and strategic planning in urban development.

2. BASE ANALYSIS

2.1 Background of the Study

The transformation of urban environments through remodeling and zoning conversions is a critical factor in the evolution of cities. As Boston continues to experience significant shifts in its demographic makeup and economic landscape, these changes are directly mirrored in the city's housing market. This study is focused on systematically analyzing the effects of these urban development activities on the availability and configuration of housing units from 2009 to 2024. The aim is to provide a comprehensive overview of the shifting landscape of Boston's housing market, with particular attention to trends in housing unit supply and the types of available dwellings.

Historically, Boston has been a city of diverse neighborhoods, each with its unique character and community dynamics. However, recent years have seen a marked increase in remodeling activities and zoning changes, driven by economic growth and a rising influx of higher-income individuals seeking residence in urban centers. These changes have the potential to significantly alter the availability and nature of housing within the city by potentially leading to a reduction in multi-unit dwellings, which are transformed into fewer, larger single-family homes, thereby impacting the overall housing stock in a way that might cater predominantly to higher-income demographics.

The purpose of this study is not just to track these physical transformations but to understand how they are reshaping the housing availability and meeting the housing needs of Boston's diverse populations. This involves examining how increases through construction and decreases through demolition or conversion affect the city's housing market. This focus is crucial

for city planners and policy-makers as it provides insights critical for designing interventions that ensure a balanced development approach, maintaining sufficient housing supply while fostering an inclusive urban environment.

Additionally, the study investigates the regional distribution of construction and demolition activities, identifying which communities are most affected by these trends. The analysis of building permits and housing unit changes offers insights into the priorities of regional development and highlights the areas most susceptible to rapid changes in housing stock. By providing a detailed backdrop of the city's developmental history and current trends, this study seeks to address the broader questions of how urban development strategies can be aligned with the goals of equitable housing access and community preservation.

By linking the technical data with the human stories behind the numbers, this background sets the stage for subsequent analysis sections and policy recommendations aimed at fostering sustainable urban growth. Through this in-depth examination, this report aims to provide a nuanced understanding and a robust assessment of the prevailing conditions and dynamics characterizing the Boston region, thereby offering insights that are crucial for informed decision-making and strategic planning in urban development.

2.2 Data Preliminary Analysis

During the preliminary data analysis phase, data cleaning was implemented to ensure the accuracy and consistency of the analysis. Missing values in the data were identified and addressed using methods such as filling, deletion, or estimation based on their type and impact to maintain data integrity. At the same time, it was ensured that each data field strictly conformed to the prescribed format and data type, such as unifying data formats, to avoid errors that could

disrupt the analysis results. Additionally, anomalies in the data were identified and evaluated, and based on their potential impact on the analysis and the understanding of the business logic, decisions were made on whether to correct, retain, or delete these values. Through these thorough data cleaning steps, the quality of the data was not only enhanced but also a solid foundation for deeper analysis was laid, ensuring the accuracy and reliability of the analysis results. Besides, all the cleaned data was then stored in a local Excel file.

A thorough analysis was conducted on the changes in each type of housing unit from 2009 to 2024, focusing on trends in utilization and transformation. Additionally, the annual trends in building permits across various regions were analyzed to gauge the dynamic shifts in construction activity and regional development priorities.

After the group's analysis, it was found that the Live Street Address Management (SAM) Addresses table lacks a corresponding key reference and a large number of units are missing, with only 1994 units effectively analyzable. The key 'unit' includes many data types that need to be cleaned up before use. The key 'units' possess a variety of data types that require cleaning before they can be employed for statistical purposes.

2.3 Base Questions

1) **What communities are building more housing units?**

The communities building more housing units, based on the number of permits, are led by Dorchester with 1267 permits, followed by South Boston (696), East Boston (576), Boston (419), and Roxbury (404). Other notable communities include West Roxbury (387), Jamaica Plain (378), Brighton (338), Roslindale (272), and Hyde Park (225).

Dorchester leads in new housing unit construction with 1,267 permits, indicating a significant expansion and potential for attracting a diverse range of residents. South Boston and East Boston follow with 696 and 576 permits, respectively, showcasing these areas as key sites of residential development. This growth reflects the city's evolving housing landscape, catering to the demand for more living spaces amidst urbanization and population growth.

2) Which ones are losing housing units?

Communities experiencing a loss of housing units, indicated by demolition or similar work types, include Dorchester with 304 instances, Boston (222), South Boston (134), East Boston (119), and Roxbury (109). Other areas with notable losses are Jamaica Plain (86), Brighton (84), Roslindale (82), West Roxbury (76), and Hyde Park (56).

Dorchester again stands out, but this time for a concerning reason: it leads to the loss of housing units, with 304 instances of demolition or conversion that potentially reduce available housing. Boston proper and South Boston follow with 222 and 134 instances, respectively. These figures suggest significant changes in the housing stock, where older or multiple-unit homes might be replaced or converted into single-family units or otherwise that cater to a different demographic, potentially displacing existing communities.

3) Where are housing remodels and renovations happening?

The majority of housing remodels and renovations are happening in Boston, with 29,714 instances, indicating a significant amount of work being done to update or change existing housing. Following Boston, Dorchester (9445), Roxbury (6664), Jamaica Plain (4206), and South Boston (3754) are the communities with the highest number of remodels and renovations.

Remodels and renovations are most prevalent in Boston, with a staggering 29,714 instances. This high number reflects the city's dynamic real estate market and the desire to upgrade or repurpose existing structures. Dorchester and Roxbury follow with 9,445 and 6,664 instances, respectively, indicating these communities are also experiencing significant changes. This remodeling trend supports the notion of shifting housing markets towards higher-income individuals, affecting the availability and character of housing units.

4) How many housing units are lost to remodels on average, each year?

In the preliminary analysis presented in the report, a trend examination was conducted through linear regression, and the variations in housing unit numbers from 2009 to 2024 were assessed. This analysis provided insights into the locations experiencing growth in housing units and those witnessing declines. Utilizing this data, a crucial query was addressed: What is the average annual loss of housing units due to remodeling? To calculate this, the changes in unit numbers over the specified period were analyzed and divided by the number of years between 2009 and 2024, yielding the annual average. The initial presentation of this data in *Figure 1*, though informative, lacked clarity in its visual representation.

Zip Code	
02108	6.800000
02109	6.200000
02110	48.733333
02111	-1.466667
02113	7.266667
02114	-6.133333
02115	27.400000
02116	-637.733333
02118	-89.533333
02119	-9.400000
02120	0.866667
02121	7.066667
02122	23.533333
02124	5.266667
02125	34.066667
02126	-1.000000
02127	64.066667

Zip Code	
02128	135.133333
02129	31.066667
02130	-15.333333
02131	39.333333
02132	-12.000000
02134	24.000000
02135	137.200000
02136	24.133333
02137	NaN
02146	NaN
02199	0.000000
02201	NaN
02210	64.466667
02215	39.533333
02445	NaN
02446	0.000000
02467	-13.266667

Figure 1. Average Unit Change in Zip Code

To enhance comprehension and visual appeal, the data visualization was restructured. By assigning zip codes to the x-axis and the changes in unit numbers to the y-axis, a more intuitive graph, *Figure 2*, was created. This graphical representation allows for the immediate identification of regions with increasing or decreasing housing unit trends.

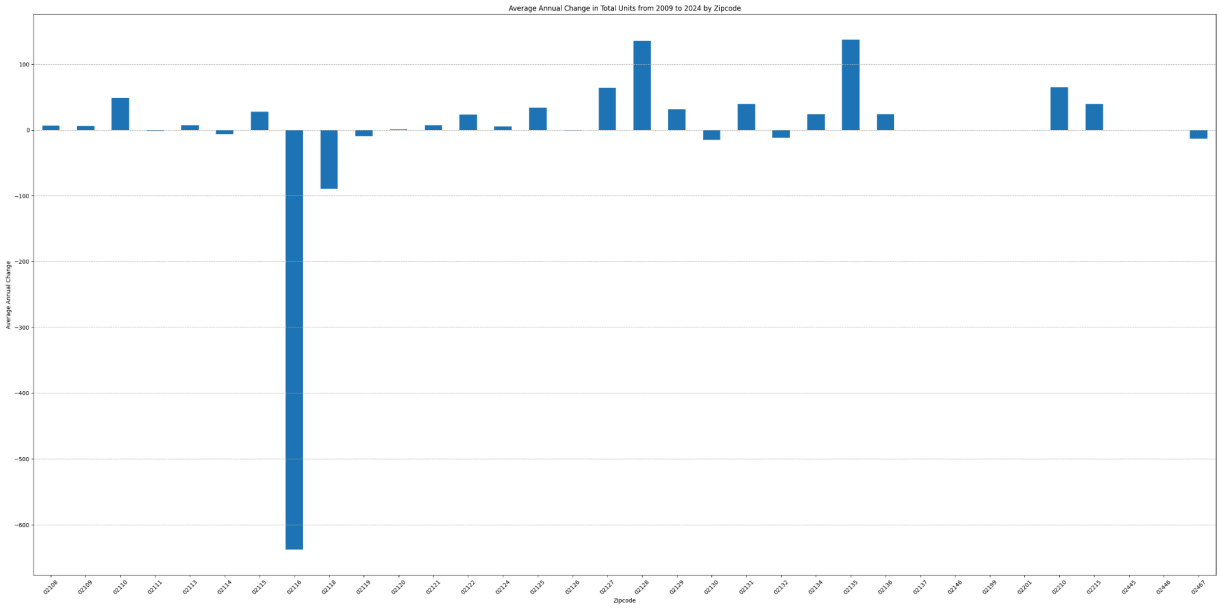

Figure 2. Average Unit Change in Zip Code in Graph

These findings offer valuable insights into the dynamics of housing availability and the impact of remodeling on housing units. Understanding these trends is crucial for policymakers, developers, and community planners in making informed decisions related to housing development and renovation policies.

3. DEEPER/EXTENSION ANALYSIS

3.1 Questions Extension

After addressing all the key questions, some extended questions were proposed to better analyze and investigate the project. Various databases and their data associated with the project were studied and analyzed, and attempts were made to analyze the correlations between different variables, particularly those related to remodeling and changes in units, such as living areas.

In addition, the annual changes of key variables were plotted into line graphs to better analyze trends and generate new questions. This visual representation helped to see changes over time and guided further analysis.

Using the methods above, three new questions have been formulated:

1) How has the remodeling taken place over different years?
2) How does the Living area in each region change every year and is there a relationship between remodeling and the change of Living area?
3) Is there a relationship between remodeling and the change of Units per year?

3.2 Extended Questions Analysis and Answer

1) How has the remodeling taken place over different years?

By using the "APPROVED BUILDING PERMITS" database to explore this, thoroughly reviewing all pertinent fields. Their analysis indicated that the 'issued_date' column specifies when the permit was issued, and the 'status' column shows the current status of the license. These details allowed the team to monitor the trends in remodeling activities across various periods effectively.

Initially, using the 'city' column to identify project locations. However, upon analysis, it is easy to find the data within this column to be inconsistent and poorly formatted, complicating direct usage. As shown in *Figure 3*, there are a lot of different formats.

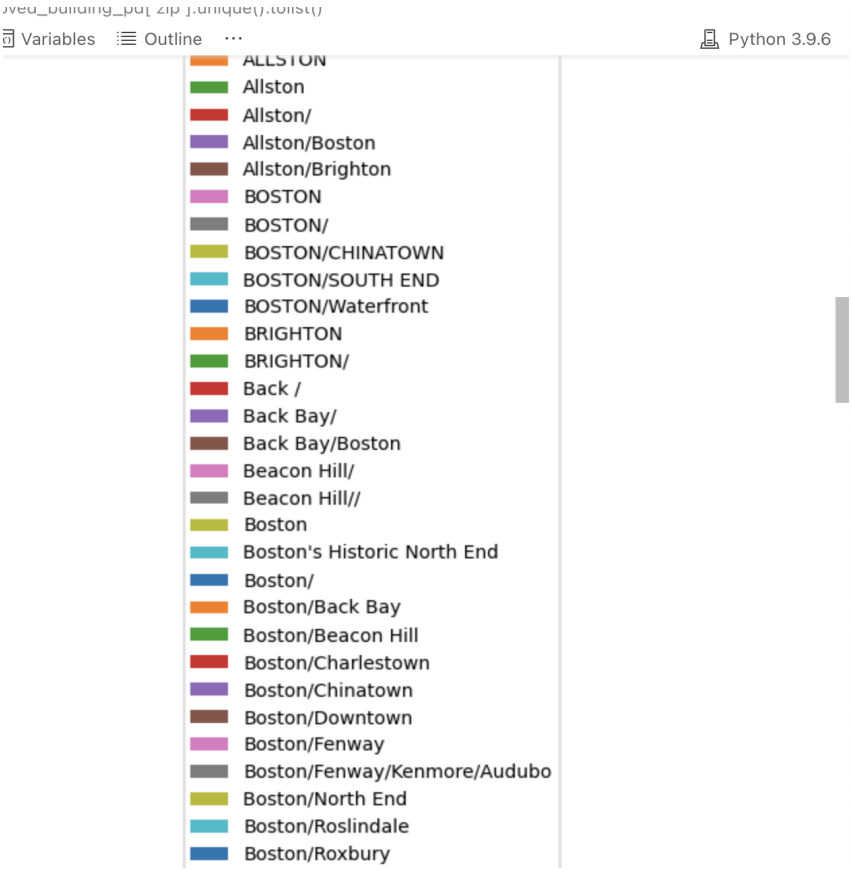

Figure 3. Data in the 'city' Column of the "APPROVED BUILDING PERMITS" Database

Consequently, using the 'zip' code as a more reliable indicator of location for each permit. Proceeding to create a series of graphs displaying the volume of remodeling projects across different areas and years. As shown in *Figure 4*, the horizontal axis represents the year, the vertical axis represents the number of permits, and each line represents a different area (distinguished by zip code).

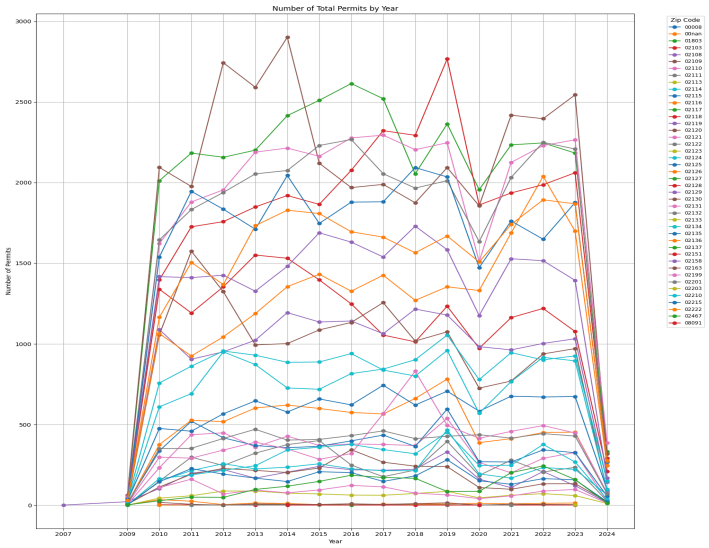

Figure 4. Total Remodeling Projects across Different Areas and Years

Our findings indicated four distinct statuses for permits, including a 'Stop work' status, which signifies permits that were issued but later halted due to non-compliance or other issues. The next step is to exclude permits with this status from the analysis. The remaining statuses—'open' and 'closed'—indicate whether a permit is currently valid or not, respectively. The separate graphs were generated for each status(“Open” and “Close”) to provide clearer insights, which are shown in *Figure 5*.

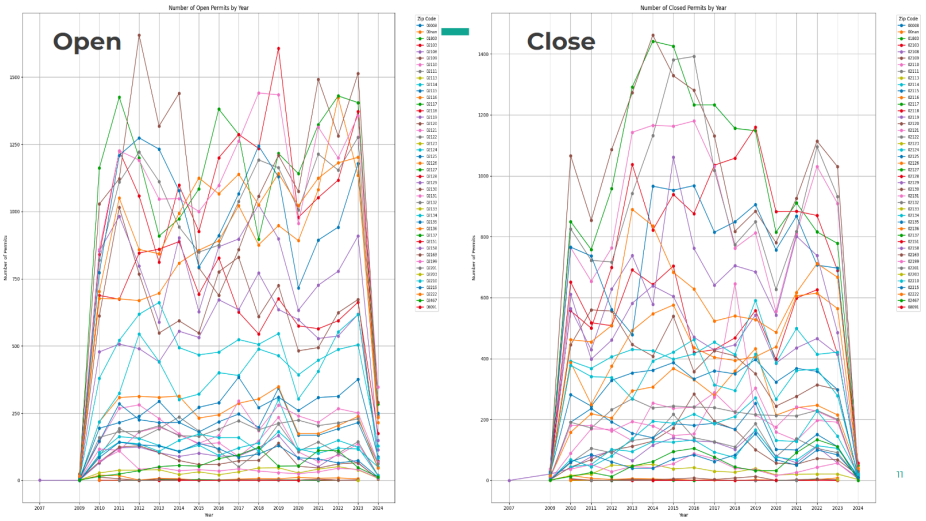

Figure 5. Total Remodeling Projects across Different Areas and Years with Status “Open” and “Close”

However, the initial graphs were overly complex and hindered easy analysis. To remedy this, zip codes were correlated with their corresponding neighborhoods, focusing on the top 10 to simplify the visual presentation. This approach significantly reduced the number of lines in each graph, enhancing clarity. As shown in *Figure 6*, the horizontal axis represents the year, the vertical axis represents the number of permits, and each line represents a different area (distinguished by neighborhoods) and only the top 10 neighborhoods are selected.

Similar to the above, separate graphs were generated for each status("Open" and "Close") to provide clearer insights, which are shown in *Figure 7* and *Figure 8*. Besides, there were some errors in the figures in the midterm report; the figures here are correct.

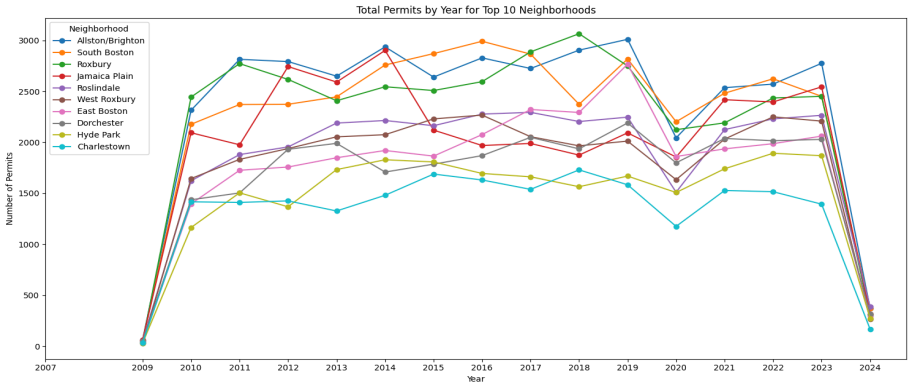

Figure 6. Total Remodeling Projects in the Top 10 Neighborhoods across Different Years

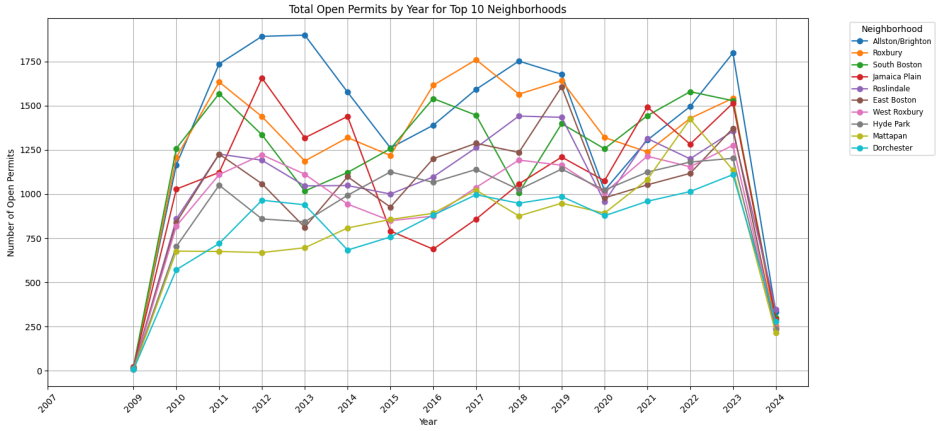

Figure 7. Total Open Remodeling Projects in the Top 10 Neighborhoods across Different Years

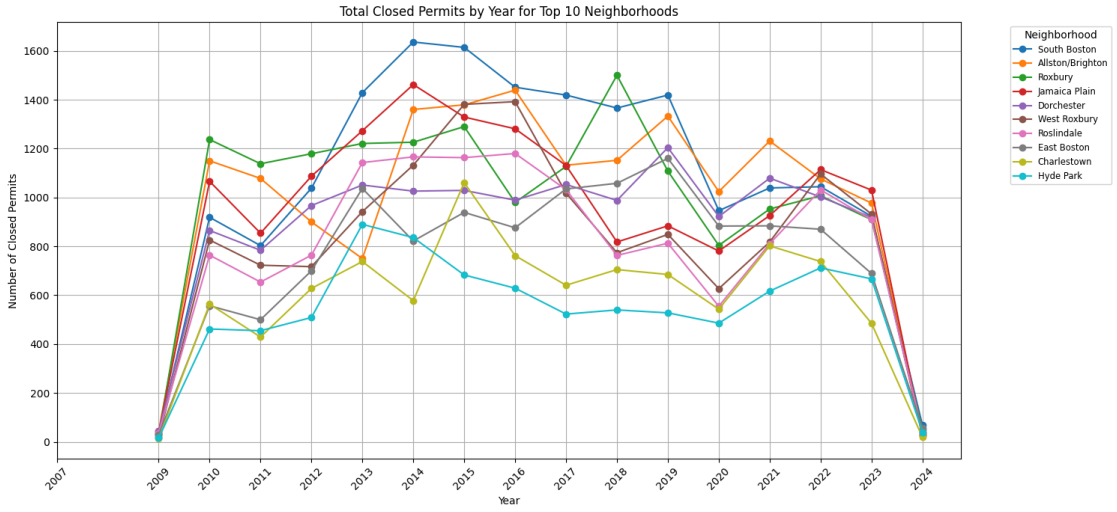

Figure 8. Total Closed Remodeling Projects in the Top 10 Neighborhoods across Different Years

During this process, several incorrect zip codes, such as 01803 and 00008. By cross-referencing the 'address' and 'city' columns, it accurately reassigned these entries to their correct locations, ensuring the integrity of our analysis.

This detailed exploration into remodeling activities sheds light on the construction trends across different neighborhoods, offering insights into the areas experiencing significant renovation activities. Such data is essential for understanding the changing landscape of housing and can guide future development strategies.

2) **How does the Living area in each region change every year and is there a relationship between remodeling and the change of Living area?**

In the analysis, each region is represented by the zip code, the living area for each region is aggregated annually and a line chart is constructed to examine how the living area in each region changes over time. Subsequently, each zip code is mapped to its respective neighborhood, and a graph is created to illustrate the changes in living areas for the top ten neighborhoods. *Figure 9* shows the changes in living areas across the top ten neighborhoods.

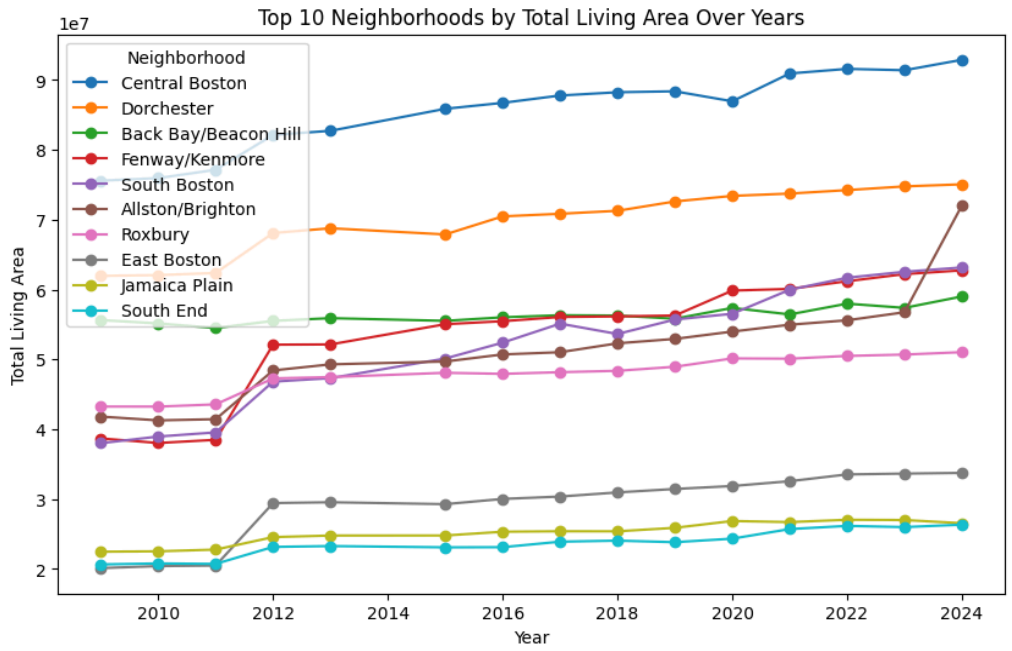

Figure 9. Top 10 Neighborhoods by Total Living Areas Over Years

From these visualizations, it is evident that the living area in each neighborhood is experiencing steady growth. Taking Central Boston as a specific example, *Figure 10* demonstrates that the living area in Central Boston has been consistently expanding.

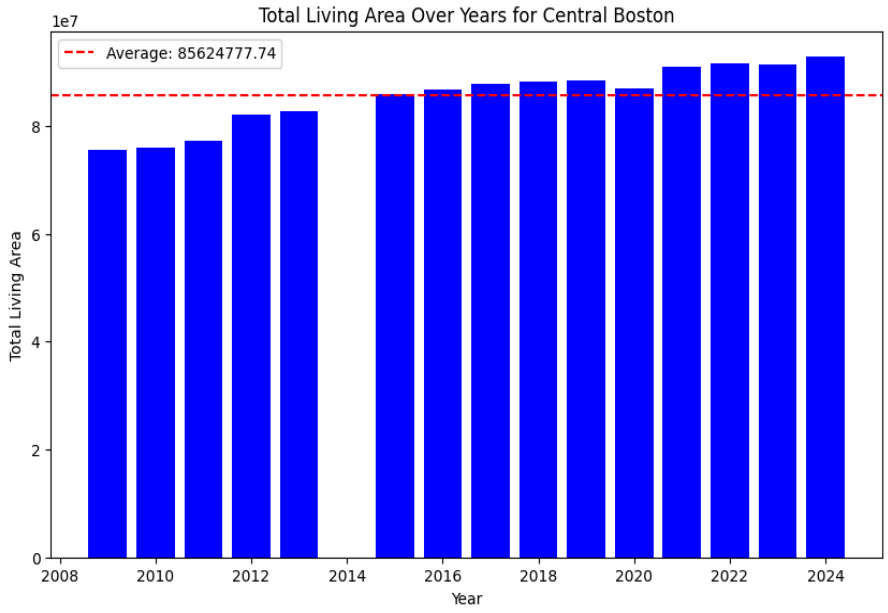

Figure 10. Total Living Area Over Years for Central Boston

Through these data and charts, the first part of the question, "How does the living area in

each region change every year?" is answered. For the latter part, "Is there a relationship between remodeling and the change of living area?" two charts have been created to assist in the analysis. *Figure 11* illustrates the 'top 10 neighborhoods by growth in living area from 2009 to 2024,' while *Figure 12* displays the 'top 10 neighborhoods by total permits from 2009 to 2024.' Since each permit represents one remodeling, *Figure 12* also represents the total remodeling activity.

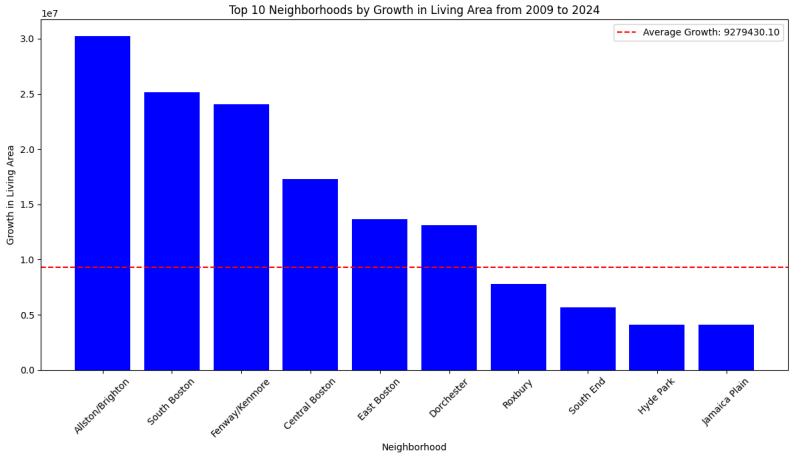

Figure 11. Top 10 Neighborhoods by Growth in Living Areas from 2009 to 2024

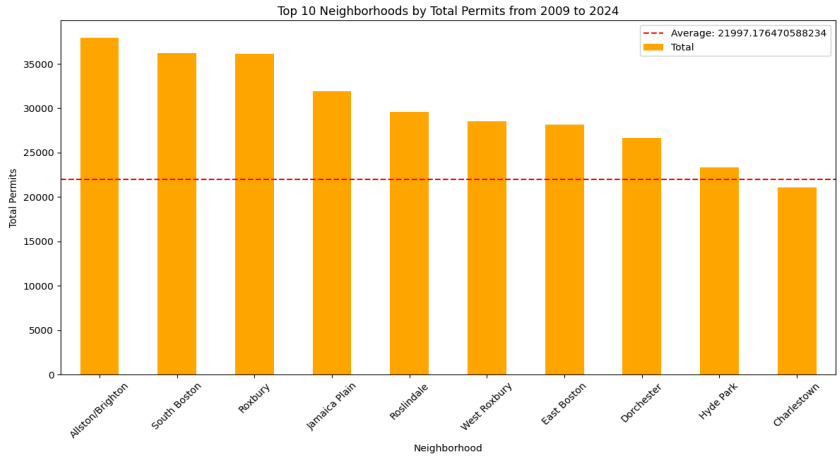

Figure 12. Top 10 Neighborhoods by Total Permits from 2009 to 2024

From *Figure 11* and *Figure 12*, it is observed that Allston/Brighton ranks first in both categories, and South Boston ranks second in both as well. However, the rankings diverge thereafter; the third-ranked areas differ, with Fenway and Roxbury occupying those positions in the respective charts.

Based on this observation, it can be inferred that there may indeed be a positive correlation between remodeling activity and changes in living areas. An increase in the living area could signify heightened remodeling activities in that locality. However, it is crucial to note that they are not determinative factors of each other; other variables also contribute to their respective changes.

3) Is there a relationship between remodeling and the change of Units per year?

This part of the analysis aims to explore whether there's a tangible link between the rate of remodeling activities and the yearly changes in various types of housing units within Boston. The investigation includes residential units, commercial units, and residential-commercial (RC) hybrid units, covering a period from 2009 to 2024. *Figure 13* and *Figure 14* were created to facilitate the analysis due to differing scales in the data set.

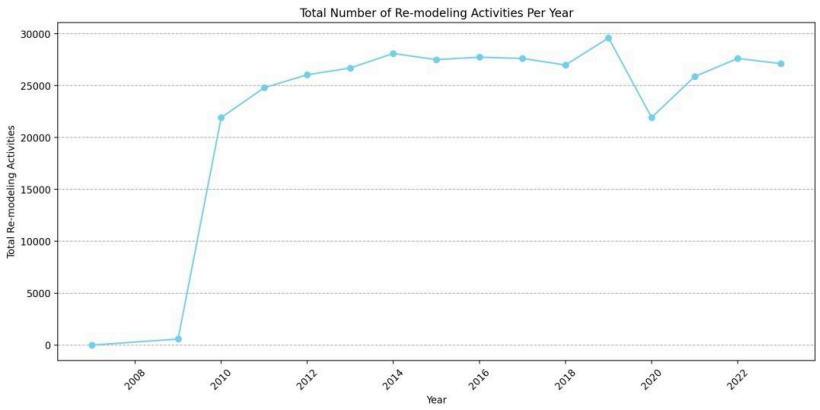

Figure 13. Total Number of Remodeling Activities Per Year

This graph represents the volume of remodeling activities annually, showcasing a sharp increase starting in 2009 with a peak around 2014. Post-peak, the graph indicates fluctuations but maintains a high level relative to the beginning of the period.

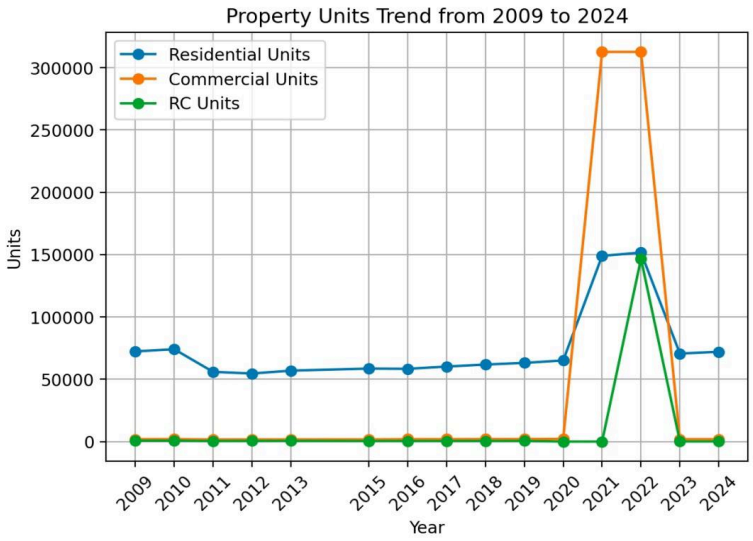

Figure 14. Total Number of Remodeling Activities Per Year

This graph delineates the trends for different types of units over the years. The trends for residential and commercial units display minimal variation, suggesting stability. However, a significant spike was observed in RC units around 2020, which sharply declined in subsequent years.

The analysis of these graphs yields several observations:

- **Stability in Residential and Commercial Units:** The number of residential and commercial units remains relatively stable throughout the analyzed period, indicating that remodeling activities have limited impact on the creation or reduction of these types of units.

- **Impact of the COVID-19 Pandemic:** The spike in RC units during 2020 aligns with the outbreak of the COVID-19 pandemic, suggesting that the pandemic might have influenced housing needs. The increase in RC units could be attributed to a heightened demand for flexible living and working spaces during this period.

- **Post-Pandemic Adjustments:** The sharp decline in RC units post-2023 could indicate a market adjustment or a shift in housing preferences following the end of the pandemic.

The data suggests a correlation between remodeling activities and changes in RC units, particularly around significant events like the COVID-19 pandemic. However, the impact on residential and commercial units is not strongly correlated with remodeling activities, indicating that these unit types are influenced by other factors or that the existing stock is largely maintained despite extensive remodeling.

3.3 Questions from Stakeholders

1) **How accurate are the insights drawn from the available data, considering the missing units and disorganized city column information?**

The accuracy of insights is contingent upon the data's completeness and cleanliness. Rigorous data-cleaning techniques were employed to mitigate these issues. However, stakeholders should note that while trends and patterns identified are indicative, the precision of predictions might be slightly affected by these data limitations.

2) **Can quantify the impact of remodeling on the availability of housing units?**

Yes, by analyzing the net change in housing units from 2009 to 2024 and examining the data on approved building permits, the annual average loss or gain in housing units can be estimated due to remodeling. It's important to note, however, that these figures are averages and may not capture year-on-year volatility.

3) **How can the analysis be improved with the current limitations of the data?**

To address the missing gaps through our ongoing data cleanup efforts and by seeking additional data sources. It would also be helpful to ask for more complete data information during the weekly check-ins with the team rep.

3.4 Conclusion

In conclusion, Boston's housing sector is undergoing a dynamic transformation, characterized by substantial construction and renovation efforts across pivotal neighborhoods including Dorchester, central Boston, South Boston, and East Boston. Dorchester is particularly prominent, leading in the issuance of building permits, which underscores its critical role in the urban development and accommodation of the population. Concurrently, these regions are grappling with the challenge of the loss of housing units, primarily through demolition, which represents a complex interplay of growth and decline within the urban fabric. Moreover, there is a distinct trend towards the renovation and modernization of existing dwellings, with Boston witnessing the highest levels of such activities. This reflects a robust demand for upgrading living spaces to meet contemporary needs and standards, highlighting the ongoing transformation within the city’s housing market.

The stability in the number of residential and commercial units suggests a potential positive correlation with remodeling activities, though this is not a definitive determinant. Furthermore, remodeling activities seem to positively correlate with increases in living space, indicating that areas with expanding living spaces may see more extensive remodeling. However, these relationships are not conclusive, as other variables also play a role in these dynamics.

4. LIMITATIONS

4.1 Potential Limitations

This section outlines the primary limitations encountered during the study of the impact of remodeling and zoning conversions on Boston's housing market. Understanding these limitations is crucial for interpreting the findings and recommendations presented in this report.

1) Data Limitations

The primary dataset utilized, the “Live Street Address Management (SAM) Addresses” database, presented significant challenges due to incomplete records and missing key references. A substantial portion of the data was unusable, with only 1,994 units effectively analyzable out of a larger dataset. This limitation restricts the generalizability of our findings and may not accurately represent the full spectrum of the housing market dynamics in Boston.

2) Inconsistencies in Data Formatting

Data inconsistency, particularly within the 'city' column of the "APPROVED BUILDING PERMITS" database, required substantial cleaning and standardization efforts. This issue was partly mitigated by using ZIP codes as a more reliable indicator of location. However, this approach may introduce its own set of inaccuracies, potentially skewing the geographical analysis and impacting the precision of the results.

3) Methodological Constraints

The methodologies employed in analyzing the data were constrained by the availability and quality of the data. The exclusion of variables such as unrecorded private transactions and

undocumented changes in housing utilization could lead to an incomplete understanding of the trends and impacts associated with remodeling and zoning changes. This limitation underscores the need for a more robust methodological framework that can capture the multifaceted nature of urban development.

4) Potential Stakeholder Bias

The involvement of multiple stakeholders with varying interests and perspectives may influence the focus and interpretation of the research findings. While efforts were made to maintain objectivity, the potential for bias exists and could affect the neutrality and comprehensiveness of the analysis.

4.2 Conclusion

These limitations highlight the challenges faced in conducting comprehensive urban studies. Future research should aim to address these issues through the use of expanded datasets, improved data-cleaning techniques, and the incorporation of more diverse analytical methods. By overcoming these limitations, subsequent studies can provide a more accurate and holistic understanding of the impacts of urban development policies on community sustainability and housing availability in Boston.

5. FUTURE SCOPE

There are several areas for further investigation in this project. Firstly, the database contains a wide range of data, and many columns are related to remodeling and unit loss. In other words, numerous variables are correlated, indicating that changes in remodeling and unit loss are determined by a multitude of factors, some of which may not even be present in the database provided for this project. Secondly, the study focuses on key data and significant variables most likely to impact changes in remodeling and unit loss. Therefore, in the future, additional types of data in the database could be processed and studied to further explore variables and data related to remodeling and unit loss.

Moreover, for data related to remodeling and unit loss that are not present in the database, further data collection efforts can enrich the database, facilitating better research into changes in remodeling and unit loss. Additionally, while some data in the database may hold research value, they may be challenging to analyze directly due to various reasons. Finding ways to further study these data in the future would also be meaningful.

Looking further into the future, it is equally valuable to explore the changes in remodeling and unit loss in the Boston area and their broader implications. The fluctuations in remodeling and unit loss may serve as indicators of local policies, economic conditions, and quality of life, among other factors, and may also be influenced by them. Therefore, a more in-depth investigation of this project would greatly aid in understanding the recent changes in the entire Boston area.

6. INDIVIDUAL CONTRIBUTION

1) **Jingbo Wang (jw6347@bu.edu)**
    a) **Introduction & Background:** Primarily responsible for writing the "Introduction" and "Background of the Study" sections, providing the necessary background and rationale for the study.
    b) **Data Analysis:** In the "Base Analysis" section, responsible for preliminary data analysis and interpretation of results, ensuring the accuracy and clear presentation of the data.
    c) **Proportion:** 25%
2) **Yuchen Cao (caoyc@bu.edu)**
    a) **Data Preparation & Cleaning:** In charge of data cleaning and preparation in the "Data Preliminary Analysis" section, ensuring the quality and reliability of data analysis.
    b) **Deeper Analysis:** Conducts deeper data analysis in the "Deeper/Extension Analysis" section, exploring the relationship between building permit data and housing remodeling trends.
    c) **Proportion:** 25%
3) **Jialong Ke (jlke@bu.edu)**
    a) **Question Formulation & Analysis:** Responsible for formulating and analyzing key questions in the "Base Questions" section, exploring housing construction and loss trends in different communities.
    b) **Extended Questions Analysis:** Handles further data analysis and response to questions in the "Extended Questions Analysis and Answer" section, deepening

understanding of changes in the housing market.

c) **Proportion:** 25%

4) **Zihan Li (lizihan@bu.edu)**

a) **Report Compilation & Editing:** Manages the compilation and editing of the entire report, ensuring the coherence and professionalism of the content.

b) **Limitations & Future Scope:** Write the "Limitations" and "Future Scope" sections, discussing the study's limitations and directions for future research, and proposing improvements and potential future studies.

c) **Proportion:** 25%