

Deliverable 1 - Team Permitting D

Real estate and development significantly contribute to the economic growth of many U.S. cities, including Boston. In Boston, numerous new buildings and renovations are constructed and conducted over days, months, and years. Commencing any development project in Boston requires acquiring an official permit from the city, a process encompassing several steps such as approval stages, community involvement, and public hearings. In this project, the city aims to analyze the aforementioned process in depth.

According to the City of Boston, the permitting process begins with submitting an application and paying the underlying fees. This is followed by a review of the plans against the city's zoning laws. At this stage, applications that do not adhere to the zoning regulations may lead to permit denial. The development project can only start once an official permit card has been issued, which is valid for 6 months and subject to extension.

For rejected applicants, there is an option to pursue an appeal. The appeal process involves filing the appeal, engaging in a community process, undergoing a public hearing, and awaiting the final decision by the Zoning Board of Appeal (ZBA). If the appeal is denied, a new one can be filed after a year.

Furthermore, larger projects exceeding 20,000 square feet must undergo the Article 80 review process, a significant focus of this analysis project.

The scope of the analysis of Boston Real Estate Permitting extends to societal, political, and environmental issue, which are pivotal for the future growth of Boston.

The project utilizes several datasets for comprehensive analysis:

No	Name	File	Source	Details
1	Approved Building Permits	abp.csv	Analyze Boston	Data on approved building permits, including location, date, fees, valuation, work type, etc.
2	Article80 Development Projects	a80.csv		Data on the development projects subject to Article 80 review process, including location, date, status, size, etc.
3	Zoning Board of Appeal Tracker	zba.csv		Data on denied permit applications undergoing the appeal process.
4	Census Data for Demographics	census.csv	US Census Bureau	Demographic data for the Boston locations analyzed in the project

1. Data Cleaning

No	Original Data	Data Cleaning Explained	Cleaned Data
1	abp.csv 27 columns	<ol style="list-style-type: none"> Drop unnecessary columns <ul style="list-style-type: none"> A. applicant, address, state, gpsy, gpsx, geom_2249, geom_4326; Clean numeric columns <ul style="list-style-type: none"> A. Remove \$ signs and convert to float <ul style="list-style-type: none"> a. total_fees, declared_valuation Derive columns from dates (year, month, day) <ul style="list-style-type: none"> A. issued_date, expiration_date Null values <ul style="list-style-type: none"> A. Keep the null values, as some are not mistakes. Keep for future use <ul style="list-style-type: none"> A. Object_id, permitnumber, property_id, parcel_id, lat, long, etc. B. description, comments - wordcloud 	cleaned_abp.csv 25 columns
2	a80.csv 25 columns	<ol style="list-style-type: none"> Drop unnecessary columns <ul style="list-style-type: none"> A. X, Y, Project_Street_Name, Project_Street_Suffix B. contact - Personal information not required Change column name <ul style="list-style-type: none"> A. Lower case letters and simpler names Derive columns from dates (year, month, day) <ul style="list-style-type: none"> A. Filed_Date, BPDA_Board_Approval, First_Building_Permit, COO_Permit_Date, Last_Project_Update_Date Keep for future use <ul style="list-style-type: none"> A. objectid, projectid, lat, lon B. name, description - wordcloud Null values <ul style="list-style-type: none"> A. Keep the null values, as some are not mistakes. 	Cleaned_a80.csv 33 columns
3	zba.csv 18 columns	<ol style="list-style-type: none"> Drop unnecessary columns <ul style="list-style-type: none"> C. address - Use city, zip, and zoning_district variables instead D. contact - Personal information not required Convert categorical variable values to numerical form - Remains categorical <ul style="list-style-type: none"> A. status - 1 to 7 in order of the appeal process - for simplicity B. appeal_type - Zoning : 0, Building: 1 Derive columns from dates (year, month, day) <ul style="list-style-type: none"> B. submitted_date, hearing_date, final_decision_date C. Calculate the differences among the three variables (duration) Correct human error <ul style="list-style-type: none"> A. decision <ul style="list-style-type: none"> 'AppProv' == 'Approved'; 'DeniedPrej' == 'Denied'; '' == 'nan'; Clean up values <ul style="list-style-type: none"> A. zoning_district <ul style="list-style-type: none"> Eliminate the term 'Neighborhood' from the values Keep for future use 	cleaned_zba.csv 35 columns

		C. parent_apno, boa_apno, ever_deferred, num_deferrals, city, zip, ward D. project_description - wordcloud	
		7. Null values B. Keep the null values, as some are not mistakes.	
4	census.csv	1. Imported Census block group data and shapefiles from City of Boston datasets 2. Renamed columns from census identifiers to readable descriptions 3. Performed an attribute join between demographic data and shapefile on GeoID 4. Dropped columns unrelated to demographics (mostly on federal institutions such as prisons, juvenile facilities and military). 5. Dropped columns with unnecessary geographic identifiers (same values for all rows, repeated identifiers due to join) 6. Plotted shapefile geometries for block groups 7. Plotted a heatmap of demographic data for each census group, mapping colors to demographic distributions across block groups 8. Normalized data by total to retrieve proportions of each demographic group for a given census block group and plotted heatmap.	cleaned_census.csv

2. Preliminary Analysis

1. What type of building permits are approved each year by type (worktype), description, valuation (declared valuation), square footage, occupancy type?

A. Work Types & Description

- As shown in Figure 1, the top work types for approved building permits are electrical, plumbing, gas, which are followed by conversion and interior renovation. Due to the frequent need for improvements and maintenance in building facilities, as well as compliance with safety regulations, it's typical for work types like electrical, plumbing, and gas to be common.

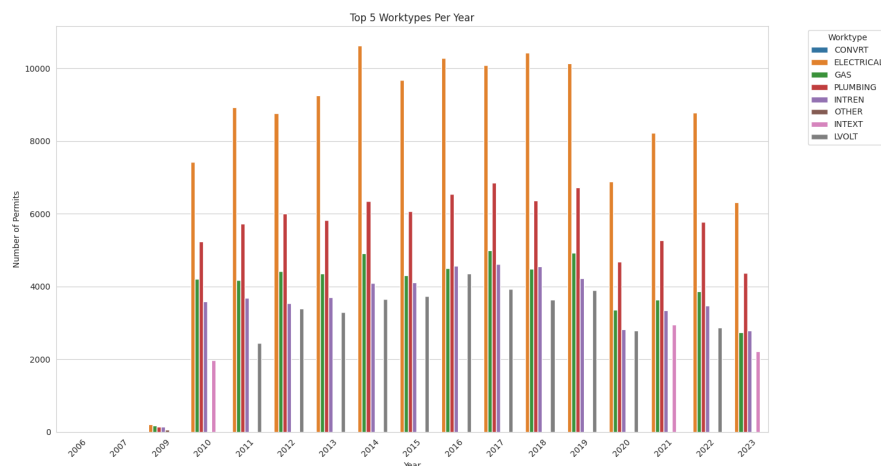


Figure 1. Top 5 Worktypes Per Year

B. Declared Valuation

- As illustrated in Figure 2, the top declared valuations for approved building permits are 0-1,750 USD, 1,750-3,500 USD, and 3,500-5,250 USD. Despite some outliers with significantly high declared valuations, the average declared valuation of the approved permits generally remains lower.

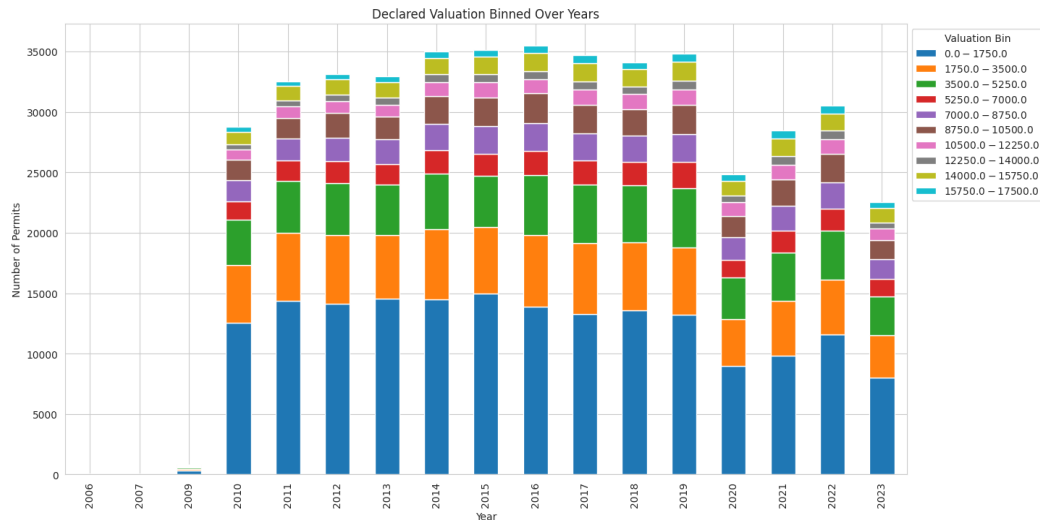


Figure 2. Top 5 Valuations Per Year

C. Occupancy Types

- As depicted in Figure 3, the top occupancy types for approved building permits are 1-3 family residential buildings, multi-family residential buildings, mixed-use developments, and commercial properties. Considering the common types of buildings in Boston, it makes intuitive sense that residential and commercial buildings would have the highest frequency in obtaining approved permits.

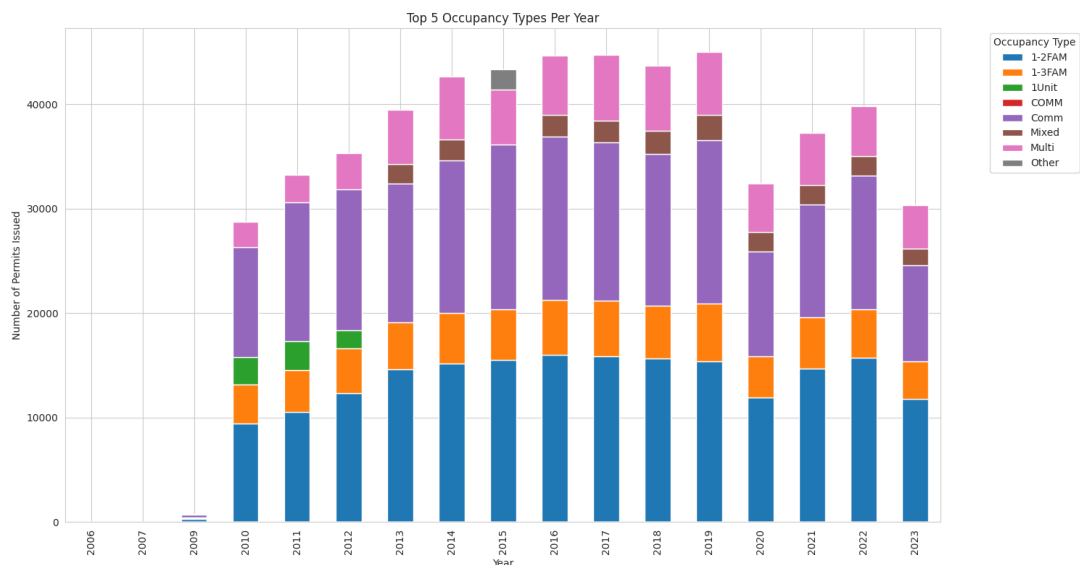


Figure 3. Top 5 Occupancy Types Per Year

2. How have these changed over the past 5 years i.e. a year over year analysis?

A. Work Types & Description

- Reviewing Figure 4, over the last five years, there has generally been a declining trend in the number of approved permits across the work types and descriptions. However, electrical, plumbing, and gas works remain consistently the most commonly approved.

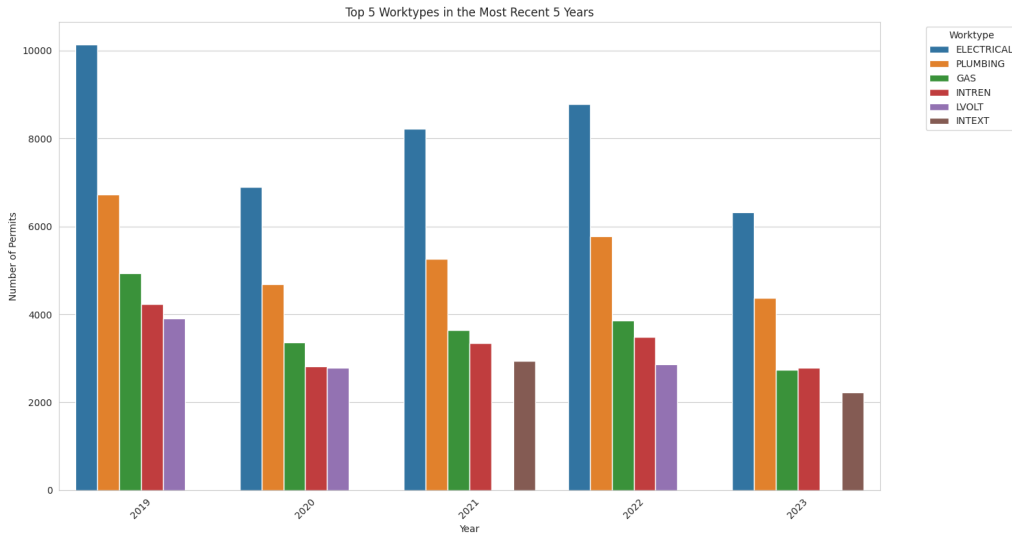


Figure 4. Top 5 Work Types Per Year (Recent Five Years)

B. Declared Valuation

- For declared values for the approved permits, there has been a steady increase each year over the five-year period. While the average declared value had a modest rise, the spike in the values of outliers is substantial.

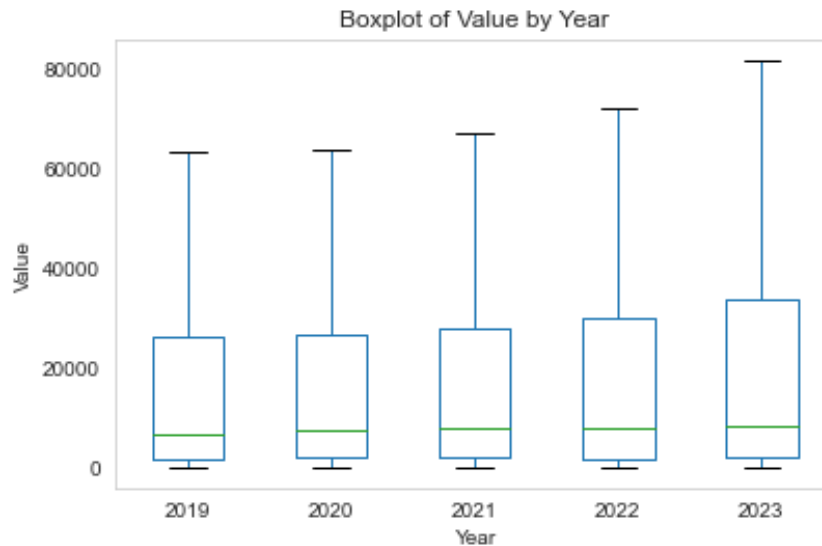


Figure 5. Top 5 Valuations Per Year (Recent Five Years)

C. Occupancy Types

- Looking at Figure 6, which shows the data based on occupancy type, the number of approved permits peaked around 2019. Subsequently, a decline through 2020 and a recovery in 2022 occurred. An assumption can be made that the COVID-19 pandemic was the potential cause of this fluctuation in recent years.

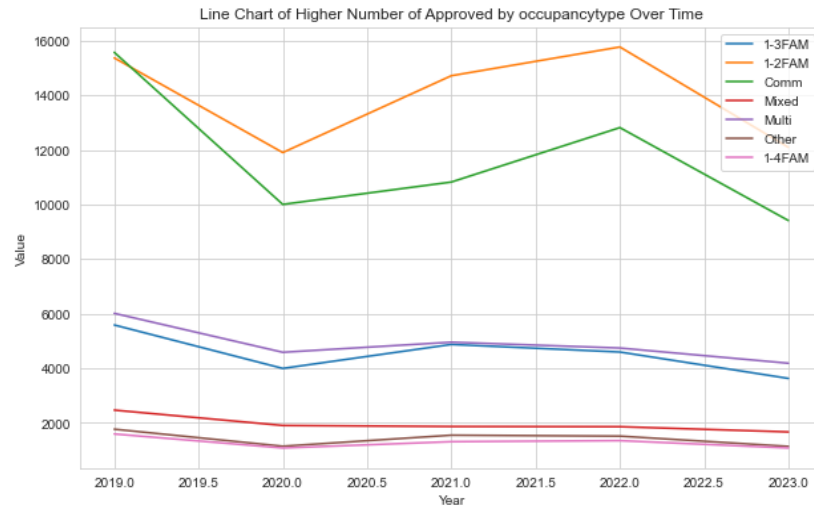
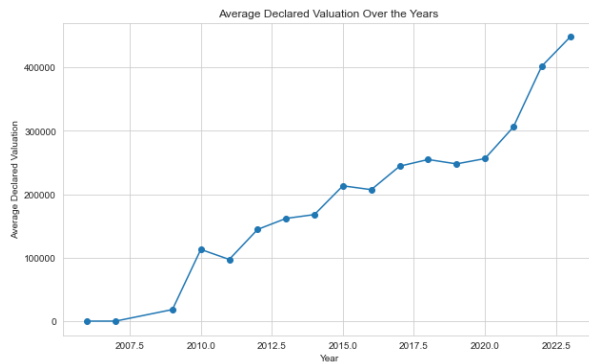


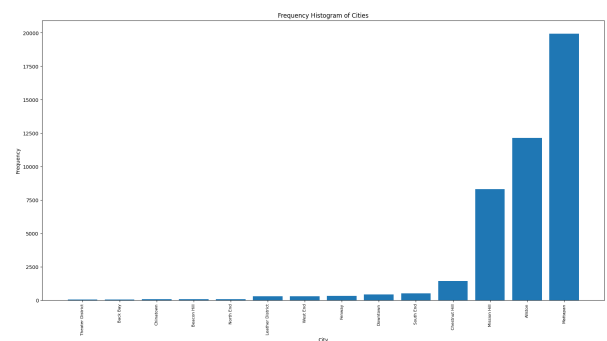
Figure 6. Top Occupancy types for past 5 years

3. Basic Exploratory Data Analysis (To be analyzed in depth in deliverable 2)

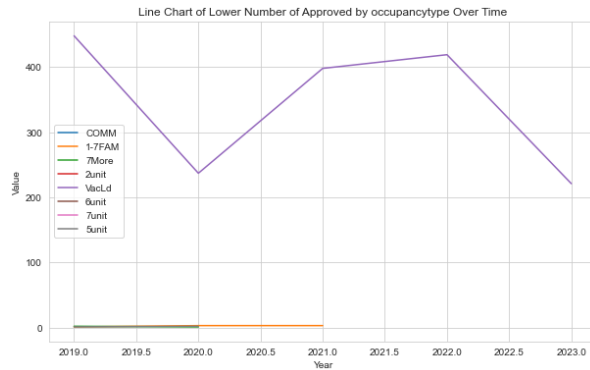
1. Approved Building Permits



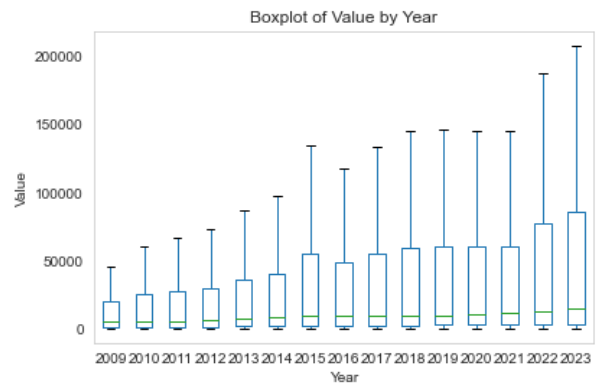
The average declared valuation shows an uptrend over the years in the graph



There are more developments in lower-income areas.

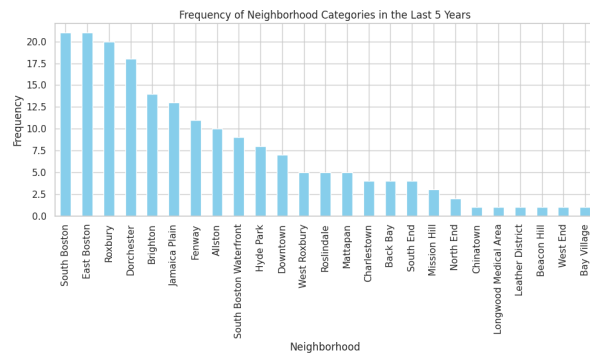


'Vacld,' '1-7FAM,' and '7More' are the lowest occupancy type in approved numbers throughout the years

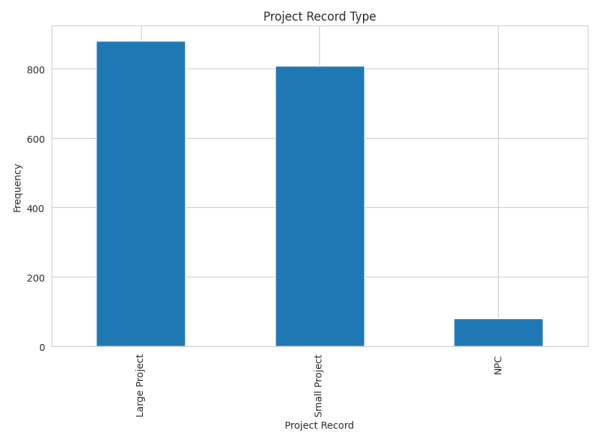


The declared variation seems to vary on a higher level throughout the years

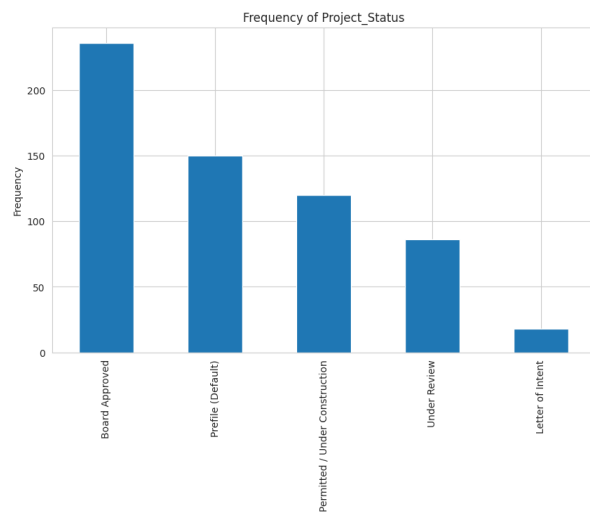
2. Article80 Development Projects



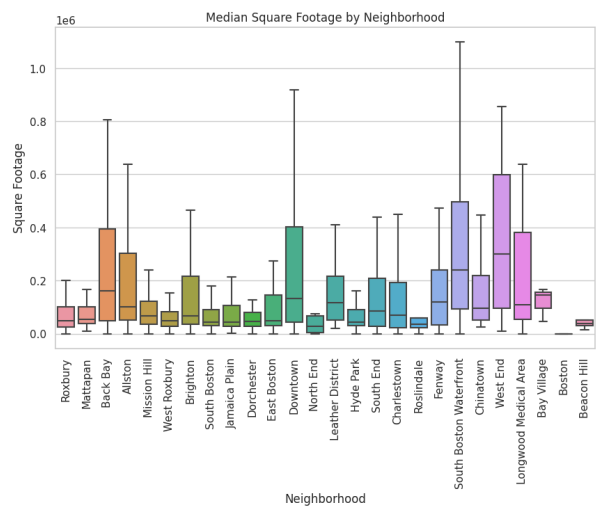
South Boston and East Boston are popular Article 80 development areas



Distribution of large & small projects, and project changes

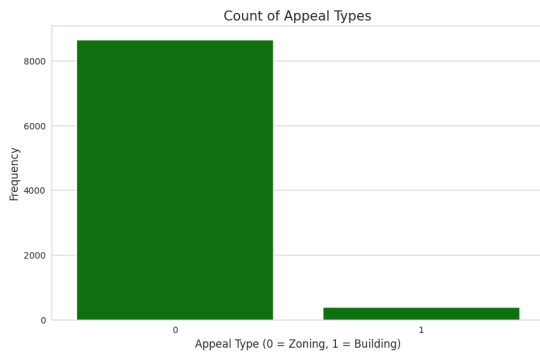


Bottleneck occurs at the beginning (Profile)

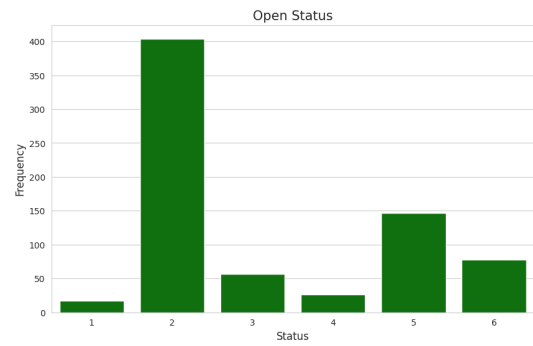


West End has the largest average square footage

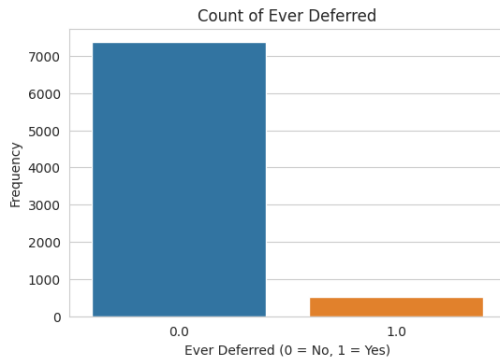
3. Zoning Board of Appeal Tracker



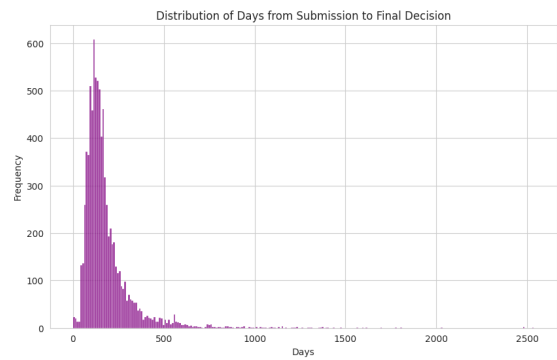
The majority of appeals are on zoning regulations



Bottleneck occurs at the community process

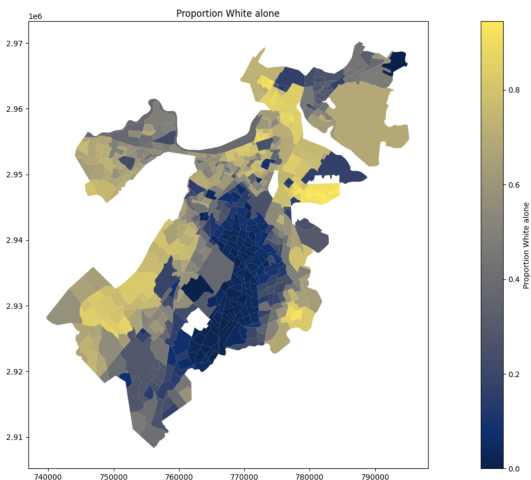


Deferrals are not common (less than 10%)

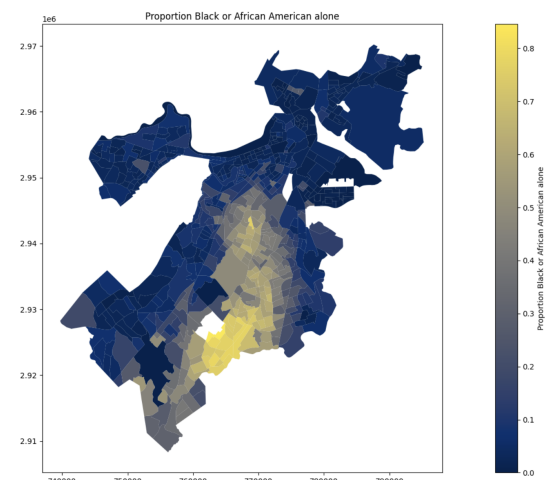


The appeal process usually takes less than 365 days

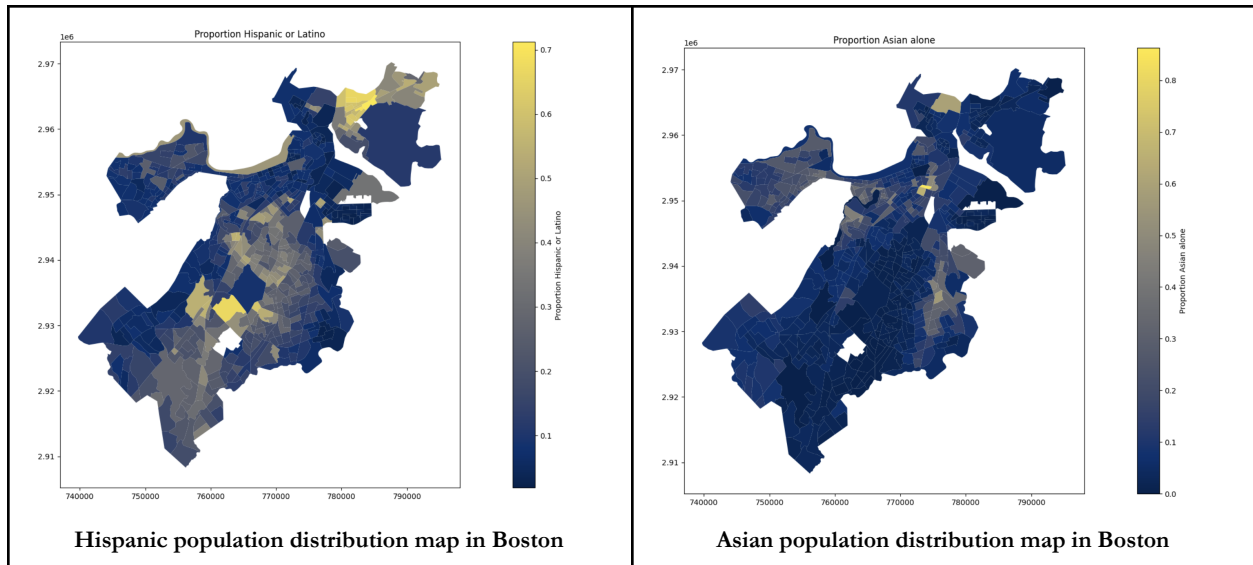
4. Census Data for Demographics



Caucasian population distribution map in Boston



African American population distribution map in Boston



4. Individual Contributions

Name	Contributions
David Euijoon Kim (Team Lead)	Cleaned zba.csv and visualized the zba data on my own. Planned and facilitated the meetings Enforced deadlines and division of labor before the meetings Assisted in cleaning abp.csv and a80.csv, and also visualizations for the data Structuring, organizing, and writing deliverable 1 with the help of teammates
Efim Sokolov	Wrote code to loop through the columns and run analysis such as number of unique, data type, numeric check, duplicates, and etc. Created box plot visualization for each column to understand the mean, spread and outliers. Also worked on the census visualizations, mapping demographic information onto shape files for Boston census block groups.
Zhihuan Hao	Data cleaning for removing signs and editing data type for some columns like issued year and expiration date which will help for time series analysis. Using the dataset Approved to promote Data Visualization and time series analysis for investigating first 2 base project key questions. Investigating the insights of the dataset approved about the permits approved by some features. Transforming the dataset as needed for time series analysis and find trends among the permits approved.
Lukas Werk	General data cleaning, especially on the Approved dataset. Investigating the dataset for optimal data-cleaning decisions and methods, developing text processing (NLP) code to help clean the large amount of typos, misspellings and improper formatings present. Documented dataset insights, context and dictionaries. Identified and started implementing more advanced techniques of cleaning the data for upcoming and exploratory analyses.
Jackson Fisk	Converted article 80 temporal data into useable metrics for graphs to accurately represent Boston's current article 80 permitting state. Helped with the consensus on fields to data clean for each dataset. Helped transform other categorical data into numeric for potential machine learning applications.

5. Reference Files

No.	Directory	File Name	Details
1	/ipynb	zba_clean_final.ipynb	Zoning Board of Appeal Tracker - Data Cleaning
2		zba_visualization.ipynb	Zoning Board of Appeal Tracker - Basic Visualization
3		abp_clean_final.ipynb	Approved Building Permits - Data Cleaning
4		abp_visualization.ipynb	Approved Building Permits - Basic Visualization
5		abp(extra).ipynb	Approved Building Permits - Basic Visualization
6		a80_clean_final.ipynb	Article80 Development Projects - Data Cleaning
7		a80_visualization.ipynb	Article80 Development Projects - Basic Visualization
8	/census	census_plot.ipynb	Census Data for Demographics - Basic Visualization
9		census.csv	Census Data for Demographics - Original CSV File
10		cleaned_census.csv	Census Data for Demographics - Cleaned CSV File
11	/data	zba.csv	Zoning Board of Appeal Tracker - Original CSV File
12		abp.csv	Approved Building Permits - Original CSV File
13		a80.csv	Article80 Development Projects - Original CSV File
14		cleaned_zba.csv	Zoning Board of Appeal Tracker - Cleaned CSV File
15		cleaned_abp.csv	Approved Building Permits - Cleaned CSV File
16		cleaned_a80.csv	Article80 Development Projects - Cleaned CSV File