Deliverable 1 | Boston Permitting Project | Team F
Aidan Ruvins, Timothy Evdokimov, Akhil Kokkula, Andre Lesnick, Zachary Gou
October 2023

Introduction:
We are studying the City of Boston Permitting Project, the goal of which is to analyze trends in the approval and denial of the City of Boston building permits. Thus far we have done preliminary data analysis to answer the base questions given in the project outline. We have promising answers in our early analysis to questions 1, 3, and 6.

Our general findings were that the vast majority of applications are there for electrical or fire alarm-related works and that most approved permits are in heavily commercialized areas of Boston (see map on page #). Furthermore, we found that there is a negative correlation between area income and denial rate of permits, but is considerably weaker than we hypothesized and requires further study.

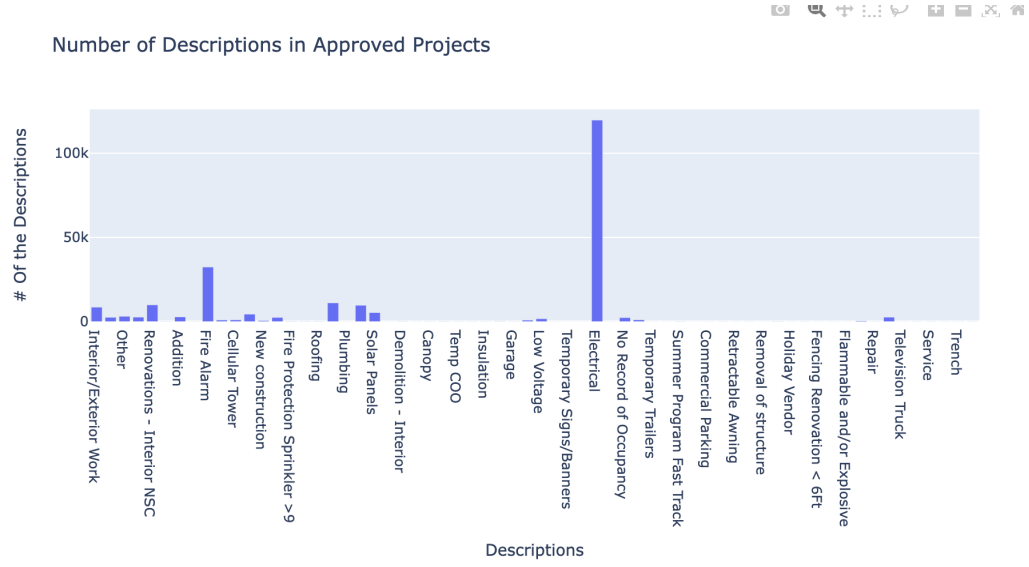
Data Collection:
Our primary data sources were provided by the City of Boston and the United States Census Bureau and is comprised of the following information

- Zoning Board Appeals Data
- Article 80 Development Projects Data
- Approved Building Permit Data
- MA Voter and Census Data

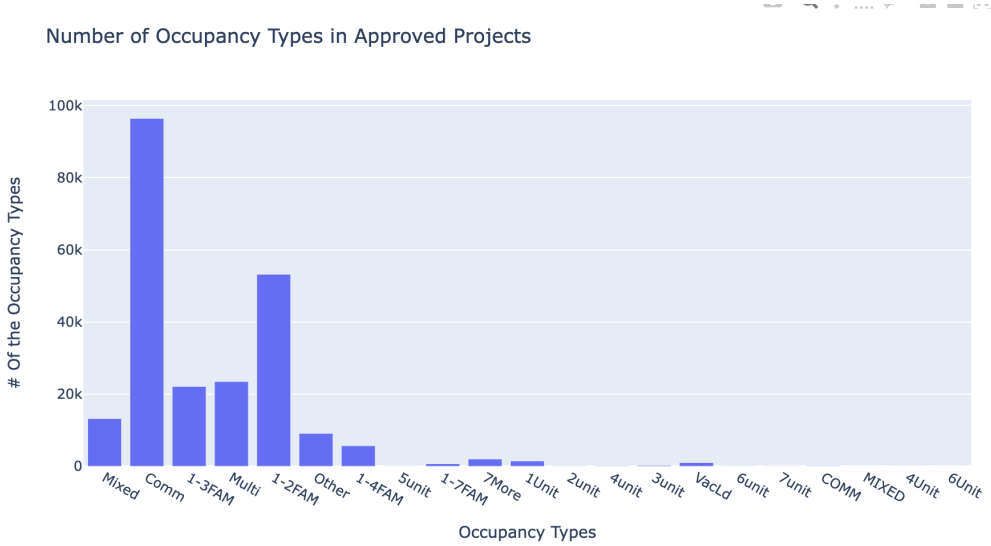
Our data cleaning steps mainly included removing incomplete data points by filtering out various “N/A”, “NaN”, etc values in the data, and filtering out obvious outliers from the datasets to prevent them from skewing means and other measurements of central tendency.

Exploratory Data Analysis:
Question: What type of building permits are approved each year by type (worktype), description, valuation (declared valuation), square footage, occupancy type?

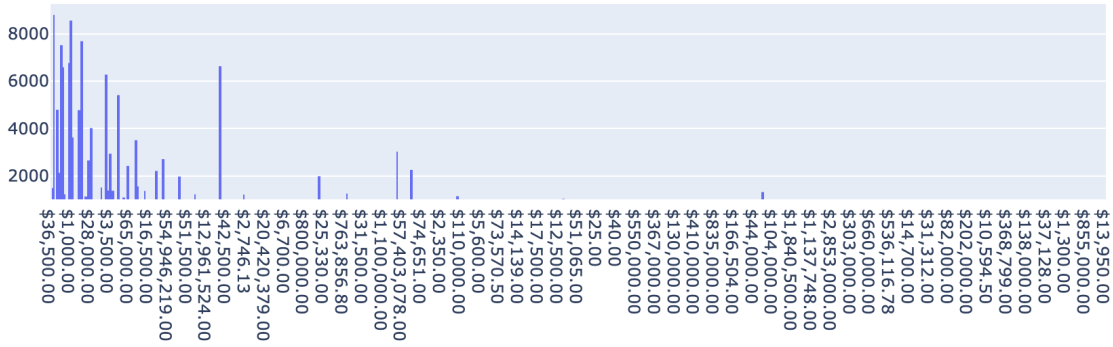
Most of the approved projects are to work on electrical issues. The second most common are fire alarm issues.



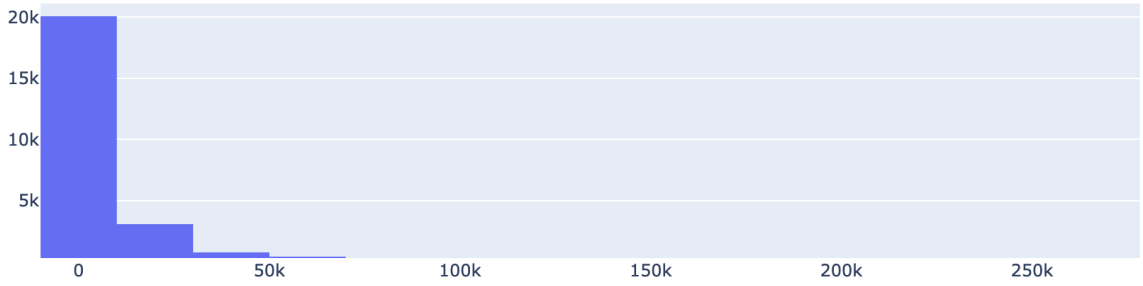
Most of the approved projects are to work on electrical issues. The second most common are fire alarm issues.



The most common project approvals are for commercial properties.

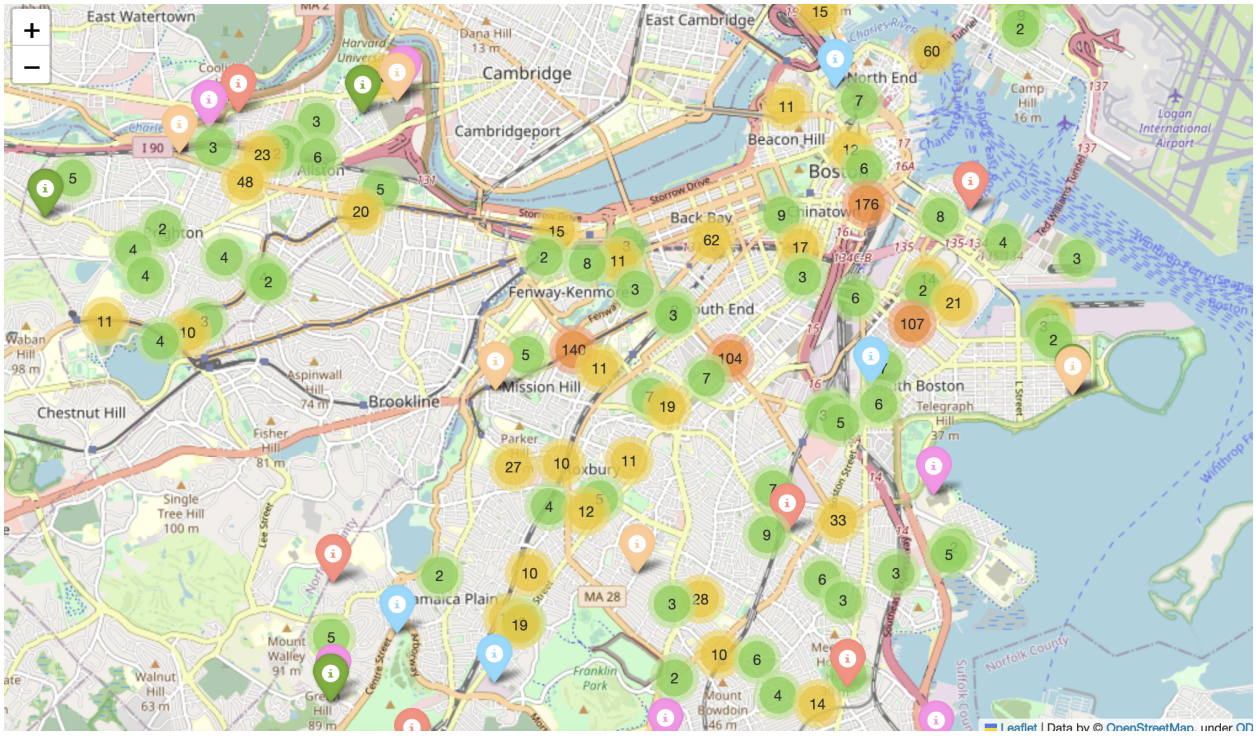


The declared valuation of approved projects tends to fall under $100,000 with a few outliers falling in the millions.



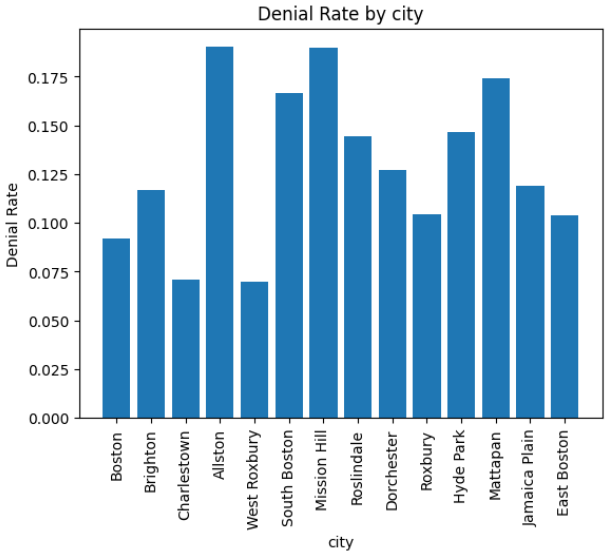
The first bar represents approved projects with square footage from the range of 0-10000 square feet. The second bar represents 10,000 to 30,000 square feet. The third bar represents 30,000 - 50,000 square feet. The fourth bar represents 50,000 to 70,000 square feet. Majority of the approved projects are below 10,000 square feet.

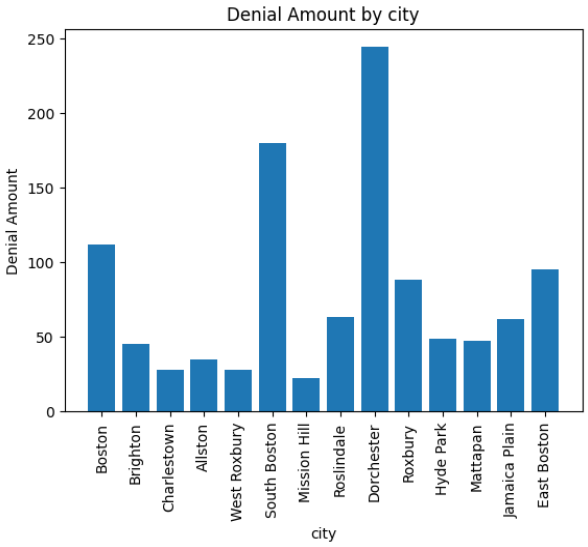
Question: Who is applying for building permits by geography (neighborhood, zip code, zoning district)?



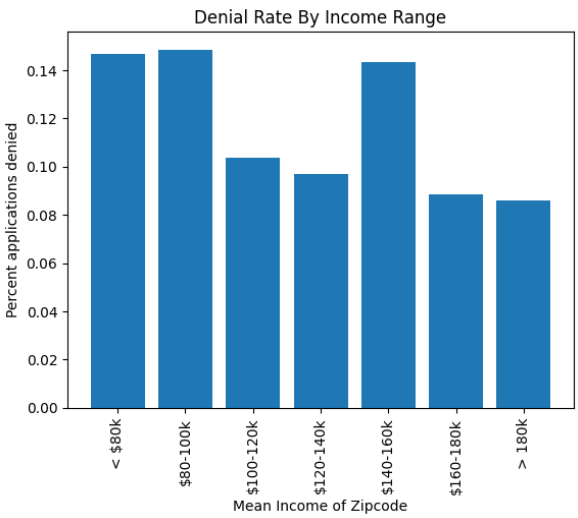
Within the Boston zip code, most approved projects are in the Fenway-Kenmore, South End Mission Hill, Financial District, and Seaport neighborhoods. This is mostly due to commercial properties being located in those areas.

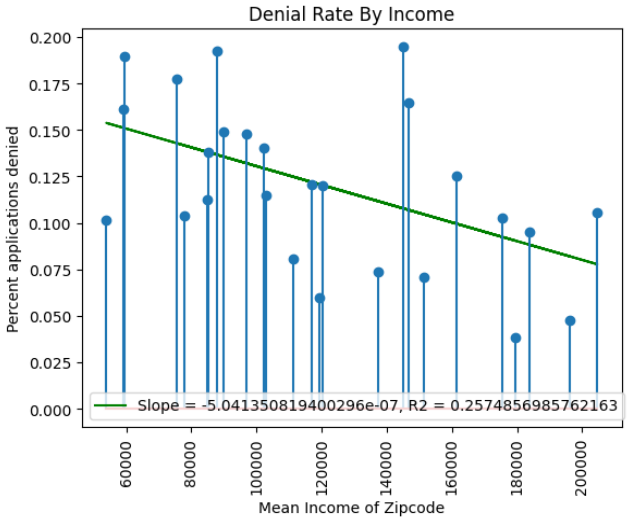
Question: What are the geographic profiles of the census tracts of the addresses for the permits submitted and zoning board approvals and denials (use project address and match to census data)?
There is a clear difference in permit appeal denial rates from one locale to another, Dorchester clearly being the most disadvantaged by number of denials, but Allston and Mission Hill being worse off in terms of denial *rate*.





Analyzing this data in terms of per-zip code data found in the MA voter file, our group hypothesized that there would be a strong negative correlation between income and denial rate, i.e. the poorer an area the more likely it would be to be denied a permit. However, the raw data of income vs denial rate showed a fairly weak correlation of R^2 = 0.25, which is a correlation for sure, but a sketchy one. Filtering out the noise by measuring denial rate vs income *range* proved considerably more useful, and showed a much more visualizable negative correlation between income and denial rate. Note the outlier of the range between 140-160k, which has yet to be adequately explained. It could be hypothesized that it is due to an abnormally high income rate in most non-residential zip codes coupled with businesses getting rejected at a high rate, but ruther research is needed to support or refute this theory.





It is worth noting that while some locations have abnormally high denial rates compared to others, the average denial rate across all locations is 12%, and the number of denials in general is not very high, with no location exceeding 20% denial rate.

Individual Contributions:

Aidan Ruvins: Created exploratory data analysis for questions 1 and 3 (report)

Timothy Evdokimov: Did data analysis, visuals, and data cleanup for question 6 (report).

Akhil Kokkula: Cleaned up the data and helped with the visualizations for questions 1 and 3 (report)

Andre Lesnick: Cleaned up MA voter file data and began exploring potential trends for the extension project.

Zachary Gou: Created exploratory data analysis for questions 1 and 3 (visuals)