CS591
Lapets

Optimizing an Urban Life

Sean Zhang
Martin Yim

Motivation

How do people make decisions in urban environments regarding time, money, and location? Where people work and live, and where they go in between is often a function of people's preferences and constraints. Workplace location is usually not flexible, but people do have choice in where they live, how much money they spend on a home, and how much time they spend on a commute. We aimed to examine some of the factors and choices people confront when looking for housing. The analyses that follow are a brief venture into what could be a large repository of interesting problems about living in urban areas, and how to optimize for one's own needs.

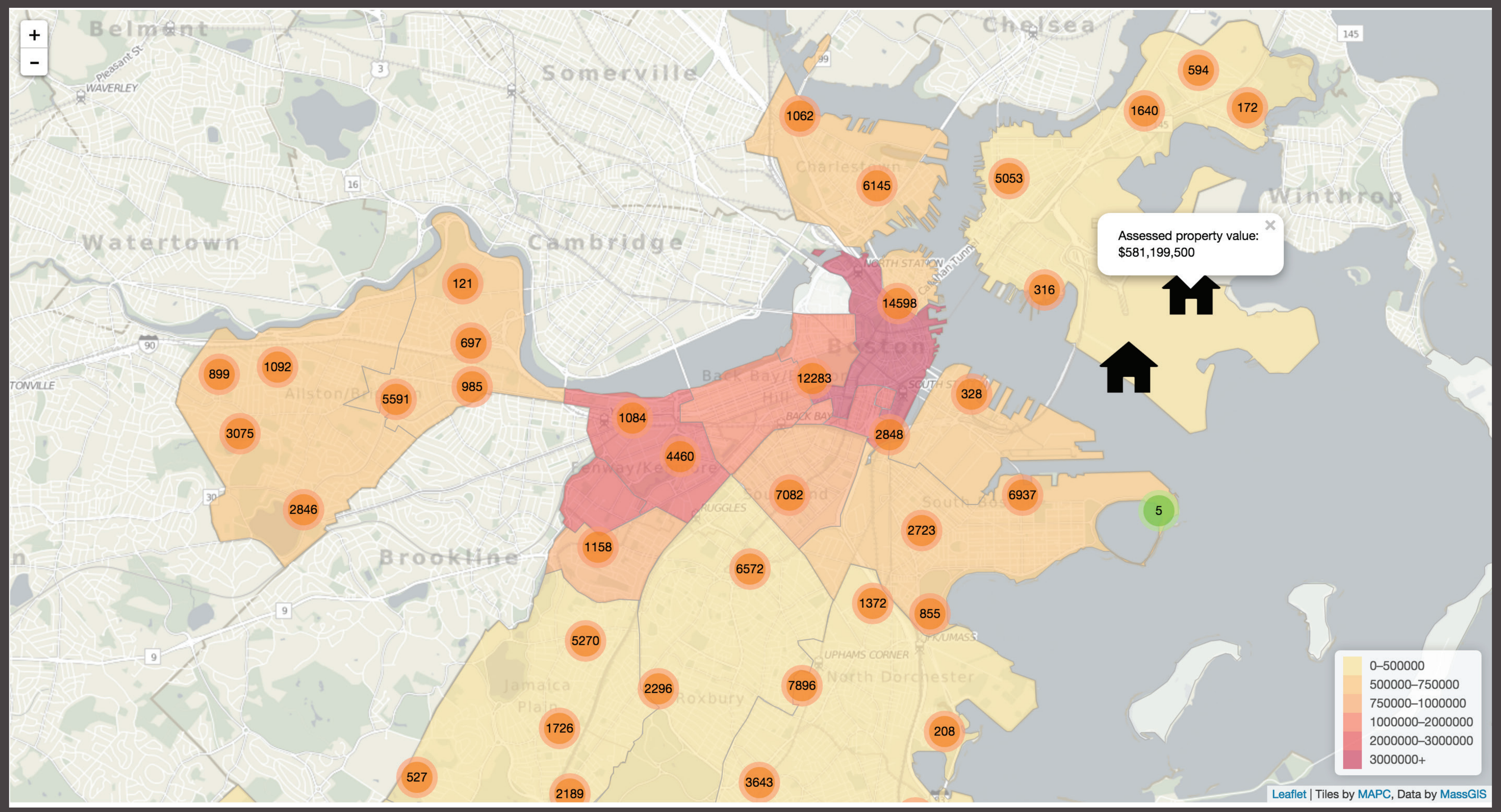
Constraint Satisfaction

We aimed to answer the question: given my residential property costs, walking time to public transit, and commute time to my workplace, does there exist another place in my area that I would be indifferent to living in? To answer this question, we first decided to hard code a company location as a parameter: Kendall Square. Then, we find the closest MBTA stops for a given residential property. We based our walk time and commute time on general averages: a person can walk a kilometer in approximately 10 minutes and public transportation can travel a kilometer in 6 minutes (with traffic).

Once we gathered all of our data, we used the Z3 library to help us determine whether our constraint was satisfiable or not. Our system of equations to solve was:

(w*w_i + t*t_i + r*r_i) <= 1
(w*w_i + t*t_i + r*r_i) > 0.99
For all i within a certain zip code.

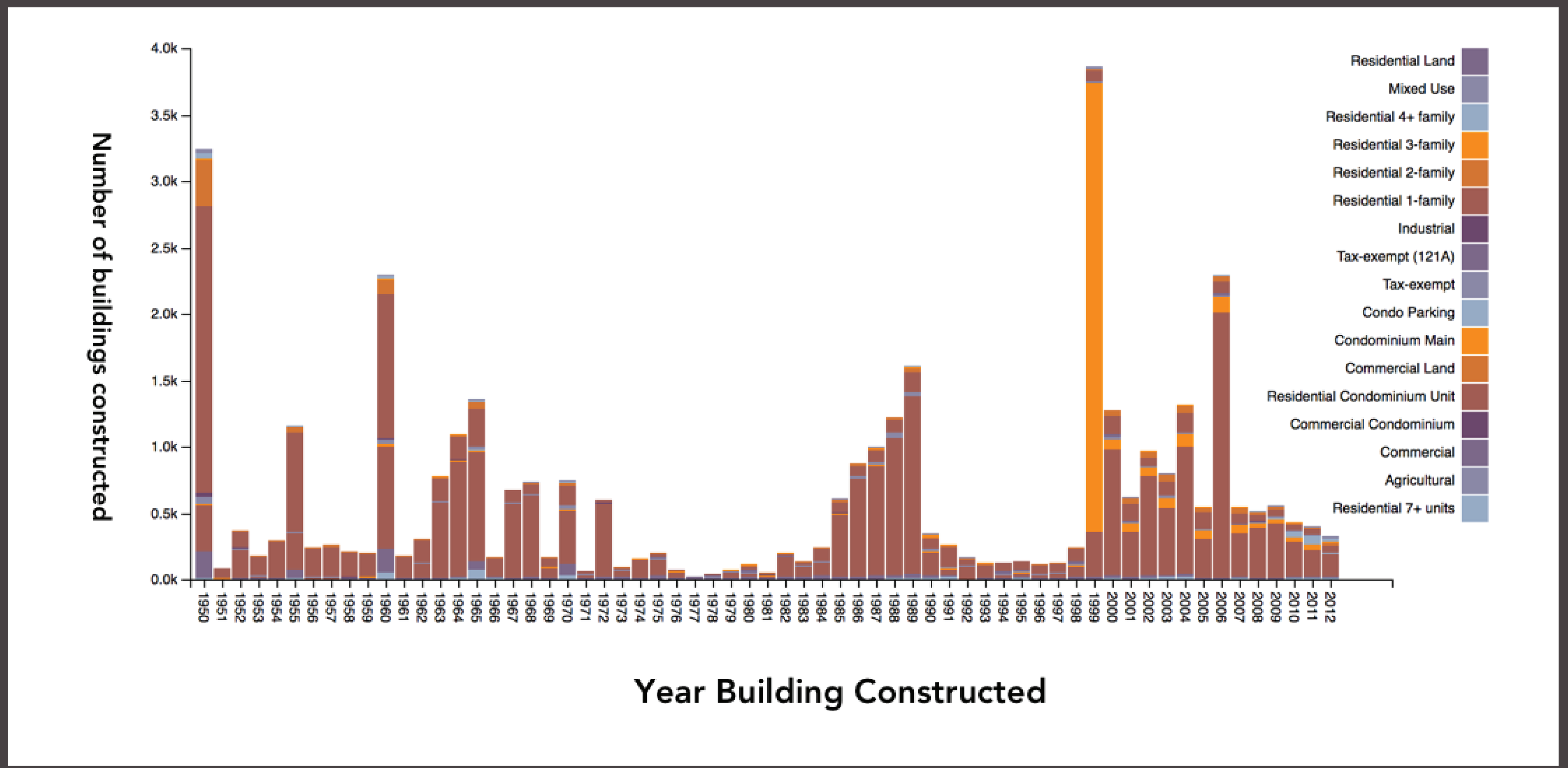
On a small sample size of residential properties, Z3 was already able to find at least five zip code regions where the indifference constraint would be satisfied.



Conclusions

Our work was able to identify interesting sets of questions about convenience, preference, and living space that would relate to most people living in urban areas. The results of our analyses were interesting, however, they were not very intuitive in nature. The constraint satisfaction analysis we conducted would not be easy to interpret for many people, and the regression analysis suffered because of feature choices.

Although we would ask similar questions in future work or expand on what we have learned, we would also work to make the analyses more interpretable and practical for the layman. Designing a life and finding the perfect home in a city is hard, but by asking the right questions, we can conduct meaningful analysis that could then create actionable information for everyday city-dwellers.



Regression Analysis

Using a multiple regression, we built a model to predict assessed property value based on some less traditional factors. Our regression model had an R-squared of 0.635, indicating that roughly 64% of the variance in our dependent variable, assessed property value, could be explained with the five factors we used. Although our model has a fairly high R-squared it is not sufficient on its own to be a reliable predictor of assessed property value.

Our model's explanatory power suffered likely due to feature choices. Problems of omitted variables bias, multicollinearity, and sample bias are likely to blame for the strange coefficients and skew.

	coefficient	std error	t value	P > l t l	95% conf. int.	
x1	-4.345e+04	618.944	-70.195	0.000	-4.345e+04	-4.22+e04
x2	-3.421e+04	966.910	-35.381	0.000	-3.61e+04	-3.23+e04
x3	8.829e+04	1606.694	54.949	0.000	8.51e+04	9.14e+04
x4	243.8660	1.673	145.772	0.000	240.587	247.145
x5	48.2452	1.388	34.748	0.000	45.524	50.967
						R-squared: 0.635

Data Review

Property Assessments 2014

Primary data source for property location (latitude, longitude, zip code), assessed property value, building style

MBTA API

Data source to get locations of public transit stops throughout all of Boston

Google Maps API

Used to query and find the closest public transit stops given a residential location within a set distance of 1 kilometer