

Nick Thompson

PPUA 5262: Big Data for Cities

20 October 2021

First Service Learning Response

Both the city walk and the in-class workshop have influenced and improved my understanding of the practical applications of working with the City of Boston data sets. When I began this semester and started to analyze the property and code violations data set, I had an understanding of what a violation was and how the city would issue such violations and associated fines. However, I did not have a full understanding of how this data set could be used in a practical way to improve the city of Boston.

For my city walk, I decided to explore the area that had the most code violations in my data during 2020. Walking through the streets and seeing the properties that filled my data set gave me a more nuanced understanding of my data. For example, over 70% of all of the records in the data set are related to trash violations. When the semester began and I noticed this trend in the data set, I did not understand why this was the case. Going on the walk gave me the opportunity to see that every property in this Mission Hill neighborhood had trash and recycling cans provided by the City of Boston. It became clear to me that the city had invested in these bins with the hopes that they would be used. Also, because the bins were provided by the city, city officials came each week to empty them. This makes trash violations something that is easy for the city to notice and to cite. Taking a walk through the city provided me an opportunity to notice what the city has invested in (trash removal), which made it obvious to me why this type of violation was so prevalent. The city walk caused me to think deeper about the violations data

and come to the understanding that the prevalence of a violation code might be an indication of the city's priorities, which is not something I had considered before.

The in-class workshop increased my comprehension of the practicality of the violations data set. Discussions during this workshop centered on other data sources that could be joined with our data to provide more useful insights. My group began exploring property assessment data, which had land parcel ID as well as property owner information in the data set. This data could be joined with our violations data to determine which property owners created the most violations in the city. This would allow the city to have a list of "problematic" landlords that could be monitored to ensure renters were not being exploited and to potentially save the city money by preventing future violations and issues before they happen. Discussions with the guest speakers provided my group the opportunity to explore different data that was publically available and join multiple data sources together to create a meaningful and important analysis for the city. We would not have been able to provide this type of insight with only our one data source.

Now that I have a better understanding of the nuances that exist in all data sets, my goal is to explore any new data set that I encounter in a way similar to what I have been able to do with the violations data set. Both by going out into the world and finding data sources with similar columns, I have been able to develop insights and practical applications that did not cross my mind when the semester began. A goal of data analysis is to uncover patterns that are not obvious, and this course has definitely allowed me to improve this skill. For the remainder of this semester, I hope to continue to develop these skills of data exploration that happen outside of a computer or a coding language. I hope that by continuing to explore the violations data, I can continue to uncover interesting and useful ways for the City of Boston to use this data.