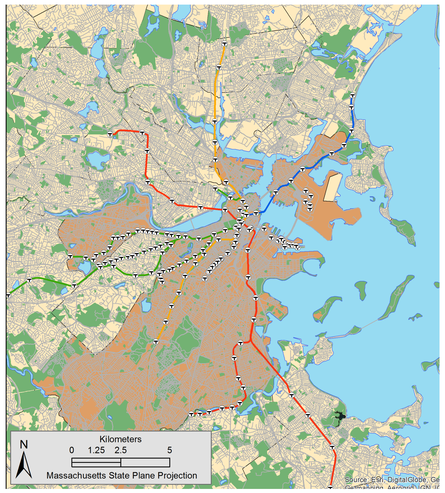
MBTA Rapid Transit Crime Analysis

Jennifer Straitz | Tufts University, Undergraduate School of Arts and Sciences | GIS101: Introduction to GIS

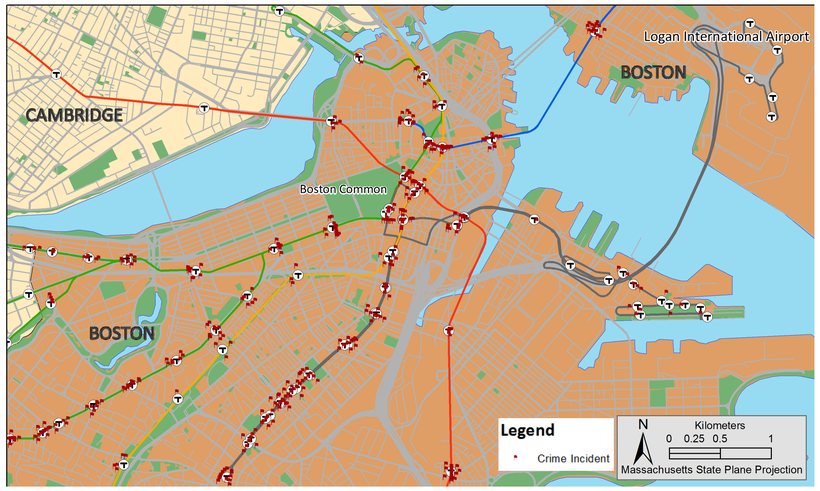
Introduction

For those people living without a car in Boston, MA, proximity to a T station is usually a top priority in determining where to buy or rent a house or an apartment, in order to be able to commute to work and access various places for entertainment throughout the city. For this reason, it can be useful to see which T stops have the greatest incidence of crime in their vicinity in order to better inform Bostonians in this decision.

In addition to determining the relative safety of T stops within Boston, from an urban planning standpoint it also seemed useful to know if the proximity of a police station to a T stop influenced the crime within the vicinity of the T stop. An inverse relationship between the number of crime incidents at a T stop and close proximity to a police station could indicate the effectiveness of the presence of police stations in deterring crime and could signal to the city government those places where a police station may be welcome in order to better deter crime.



Context Map of Boston MBTA Rapid Transit System



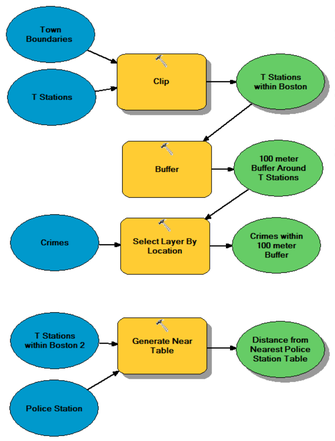
Crime Incidents Within 100 meters of T Stops in Boston (2013)

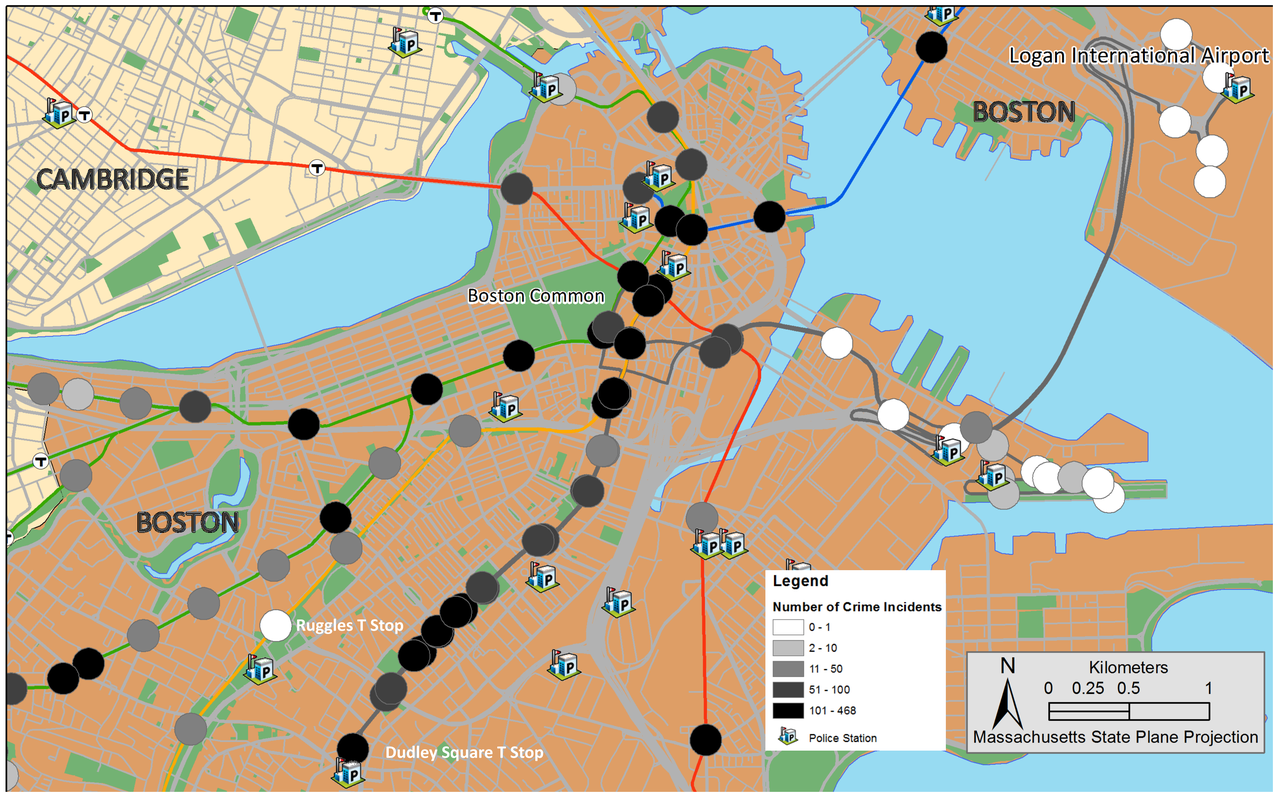
Methodology

First, the MBTA T Stations dataset is clipped to only include those T stations within the town boundary of Boston, since the 2013 crime data only covers Boston proper as well. Second, a 100 meter buffer is created around each T stop within the clipped dataset. Third, the "Select Layer By Location" tool is used to select only those incidences of crime that occurred within the 100 meter buffers around each individual T stop. Finally, outside of the model the count of crime incidences with the 100 meter buffers around each individual T stop is summarized in order to see which T stops have the highest and lowest number of crime incidents in their vicinity.

The second part of the model takes in as parameters the clipped dataset only including T stations within Boston, as well as the information on police stations. Using the "Generate Near Table" tool results in a table including the distance in meters of each T station from the closest police station.

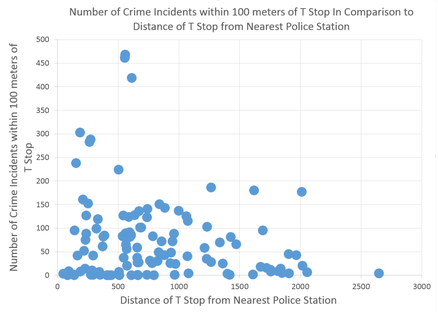
The most critical procedure used to model a relationship in the model is the step where the "Select Layer By Location" tool is used, allowing me to only select those crimes within the 100 meter buffer of each T stop. Before, I used the Point Statistics and Sample tools to attempt to do this, but spatially joining the crime information to the buffer zones was critical in speeding up the processing time as crimes outside of these zones were not counted or subsequently summarized.





Crime Incidents Within 100 meters of T Stops in Boston Frequency Visualization (2013)

This final map visualizes the frequency of crime incidents within a 100 meter radius of each T stop in Boston: the darker each buffer, which surrounds each T stop, the higher the incidence of crime. The bottom left area of the graph gives a good example of the results of the data that I found, which indicate that there does not seem to be a strong relationship between further distance from a police station and a higher number of crime incidents. The Ruggles T Stop only had 1 crime incident recorded, while the Dudley Square stop at the end of the Silver Line was in the highest range of recorded crime incidents. However, these results could be partially due to limitations in the data addressed in the section "Limitations and Results of Data".



Limitations and Results of Data

There is a clear limitation in the 2013 dataset used to map crime data. Due to privacy issues, the geocoded address locations of crimes are "obscured to the street segment centroid for privacy"

(Curt Savoie, Principal Data Scientist, Department of Innovation & Technology, City of Boston).

Unfortunately, when trying to analyze crime incidence in relation to a very specific location- in the instance of this study, a T stop- instead of a general area, some of this data will be misrepresented. Thus, the model will underestimate or over-estimate instances of crime fairly arbitrarily, depending on if a street segment's centroid falls within the 100 meter radius buffer zone of the T stop. However, if the crime incidence information provided by the model was to be used in a decision-making context, considering the fact that street segments are likely relatively short I would still be confident that the model could generally represent the relative safety of T stops to one another.

This is a hard problem to improve upon as it is not indicative of a flaw in the capabilities of the City of Boston to accurately report crimes, but is instead based off of the need for privacy. One solution is to continue using the street centroid method in documenting the specific details of crimes, but to also provide the actual location of the crime without any additional identifying information that those involved would not want exposed to the public.

Finally, the results in the "Distance from Nearest Police Station Table" likely overestimate the distance of a T Stop to a Police Station, as university campus police stations are not included in this dataset.

Data Sources

Crime Incidents, April 2015, Boston Police Department; published by the City of Boston On-Line Link, accessed April 13 2015.

Police Stations, December 2014, Massachusetts Emergency Management Agency (MEMA) GIS Program/ Regional Planning Agencies/ participating communities; published by MassGIS On-Line Link, accessed April 13 2015.

MBTA Rapid Transit, September 2014, Central Transportation Training Staff/ MassGIS; published by MassGIS On-Line Link, accessed April 7 2015.

Roads, The Massachusetts Department of Transportation- Office of Transportation Planning , June 2014.

Hydrography, UMass Water Resources Research Center, February 2013.

Open Space, MassGIS Executive Office of Energy and Environmental Affairs (EOEEA), January 2015.

Tufts UNIVERSITY