# Estimating Ridership on the Green Line Extension

**Figure 1: Thiessen Polygons**
**Figure 2: 1-Mile Network Buffer Polygons**
**Figure 3: Intersect of Figures 1 and 2**
**Figure 4: Green Line Extension Polygons**
**Figure 5: # of Attractions**
**Figure 6: T Lines at each Station**
**Figure 7: # of Transfers to other Transit**
**Figure 8: End of Line**
**Figure 9: % Area Retail**
**Figure 10: % Area Public Institution/University**

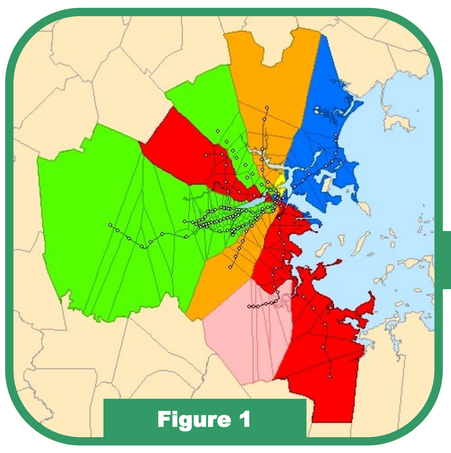
Figure 1

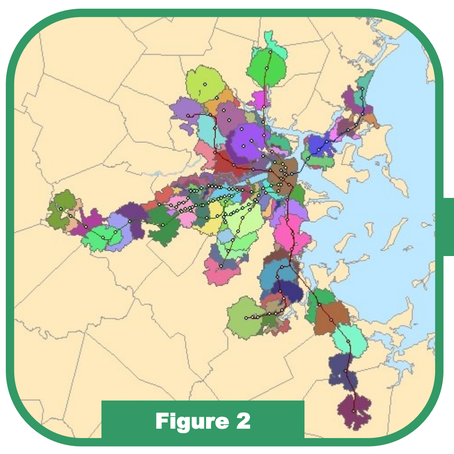
Figure 2

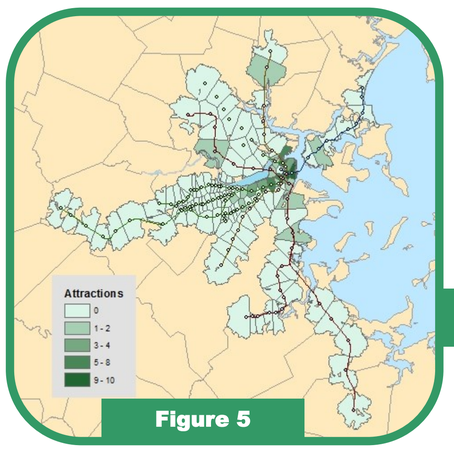

Figure 5

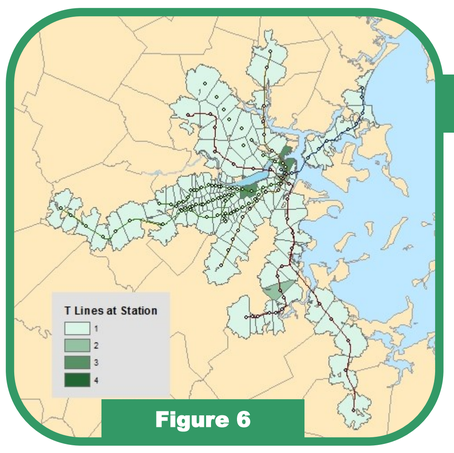

Figure 6

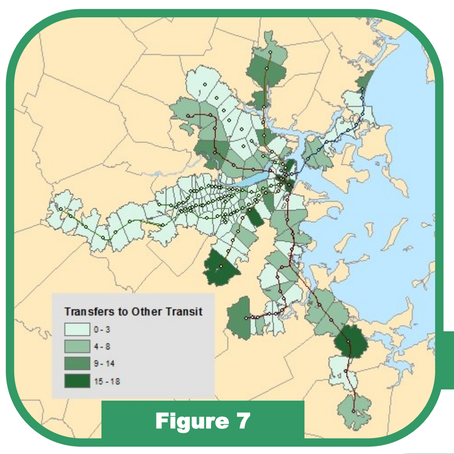

Figure 7

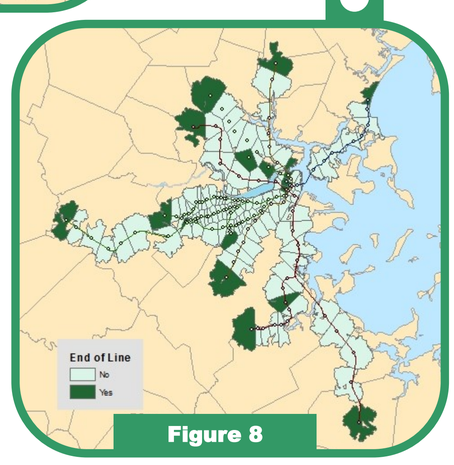

Figure 8

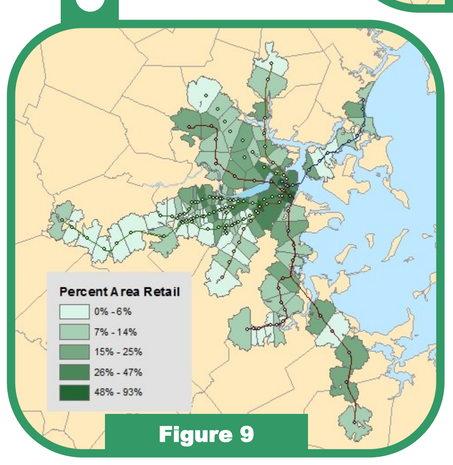

Figure 9

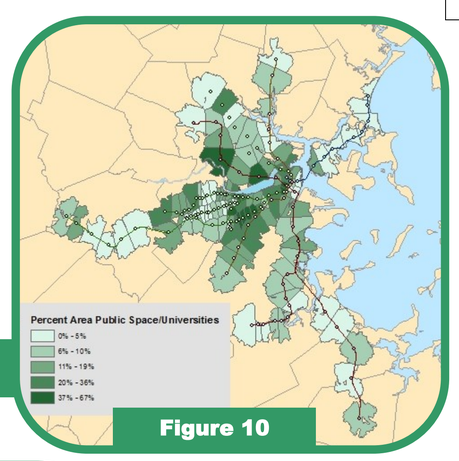

Figure 10

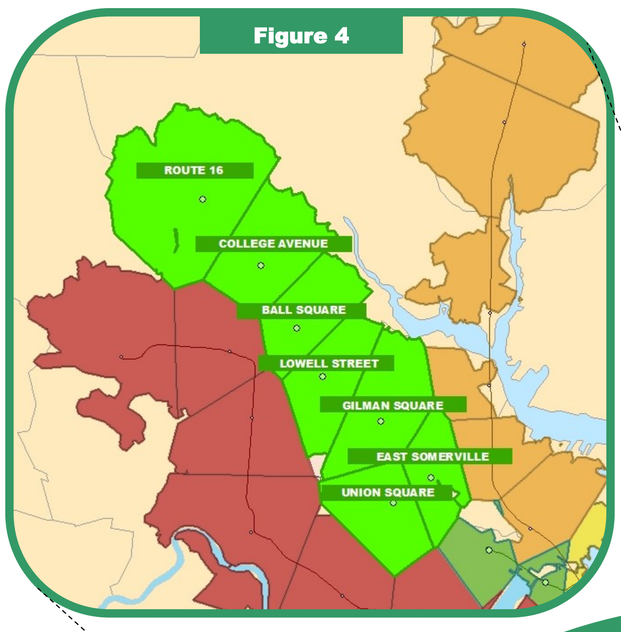
Figure 4

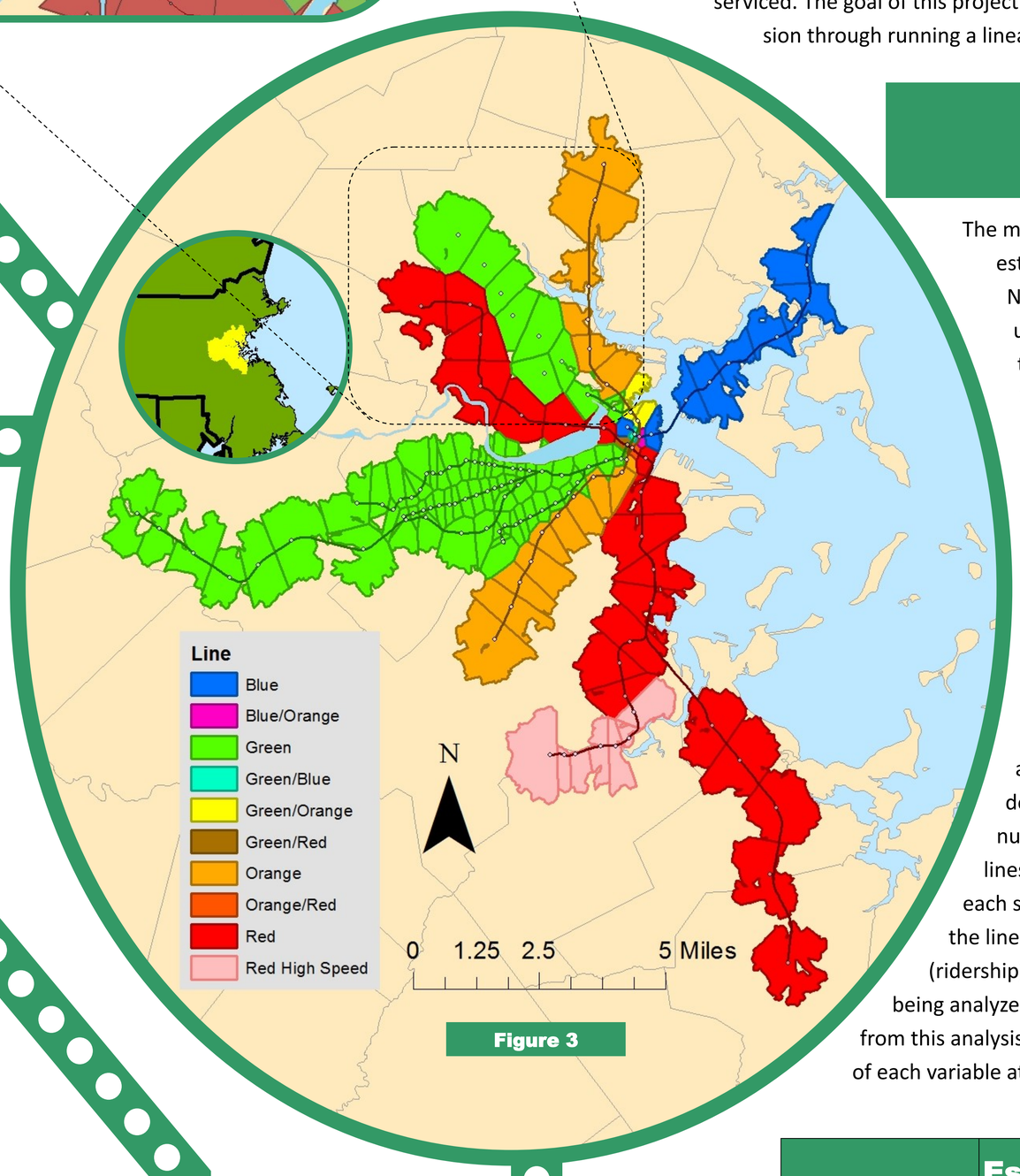

Figure 3

## BACKGROUND

For any college student or resident of the Medford/Somerville area, the Green Line Extension has seemingly been in perpetual development over the years. Currently, the area the extension would cover is poorly serviced by T stations, with long walks needed to reach the closest stations on the Red Line or Orange Line. In recent years, however, progress has been occurring on the Green Line Extension. With approvals for finances making their way through the federal government, the new Green Line is expected to be completed and operational in 2021. The plan establishes two Green Line tracks which will run along current Commuter Rail lines: One line will service Union Square, while the other line will have stations at East Somerville, Gilman Square, Lowell Street/Magoun Square, Ball Square, and College Avenue/Tufts University. Another station located at Route 16 remains in discussion, with advocates believing it will provide greater T access to those located further out in Medford. There is optimism that these stations will service a large population, spurn economic development , and greatly improve traffic and air pollution by reducing the amount of vehicular trips needed by those being serviced. The goal of this project is to attempt to estimate ridership on the Green Line Extension through running a linear regression on several spatially analyzed variables.

## METHODS

The methods of this analysis were based off a similar analysis and estimation of ridership for the new Second Avenue subway in New York City, with some variations made in the variables used. First, expected service areas were created for each T station, including those proposed in the Green Line Extension. These service areas were generated from intersecting two types of polygons which attempt to show service area: Thiessen Polygons and 1-mile network buffer polygons. Once the final service area map was created, block group polygons with population data from the American Community Survey, zip codes polygons with employment information from Community Business Patterns, and land use polygons from MASSGIS were analyzed against service areas in order to calculate weighted variables. Variables utilized for each service area in this process included population, average income, employment, percentage retail area, and percentage public space/university. Other variables were derived from the characteristics of each T station, including the number of popular attractions near each station, number of T lines at the station, other methods of public transport available at each station, and whether or not the station was the terminus for the line. Due to the rightward skew of the dependent variable (ridership), logarithmic values of each variable were calculated, before being analyzed in a regression in Geoda. With the coefficients generated from this analysis, estimated ridership could be calculated based off the values of each variable at the proposed Green Line Extension stations.

| Variable | Coefficient | Probability |
|---|---|---|
| Constant (Daily Ridership) | 5.80599 | 0.00000 |
| Population | 0.0000953942 | 0.00000 |
| Income | 0.00000177891 | 0.36981 |
| Employment | 0.0000157674 | 0.44469 |
| Percent Area Retail | 2.02581 | 0.00050 |
| Percent Area Public Institution/ University | 1.20167 | 0.02988 |
| Attractions | 0.0106607 | 0.85676 |
| T Lines at Station | 0.244831 | 0.19720 |
| Transfer to Other Transportation | 0.0173773 | 0.00001 |
| End Point | 0.361182 | 0.08298 |
| Lamda (Error) | 0.541267 | 0.00001 |

| Station | Estimated Daily Ridership |
|---|---|
| Ball Square | 1808 |
| College Avenue | 2384 |
| East Somerville | 1118 |
| Gilman Square | 3608 |
| Lowell Street | 1814 |
| Route 16 | 4681 |
| Union Square | 4430 |

## RESULTS

The results of this analysis suggest that the Green Line Extension would be well utilized by residents of Medford and Somerville. Stations with the most ridership include Union Square, with a large retail area, and the disputed Route 16 station, which would extend service to further away suburbs. The statistical analysis in Geoda produced several interesting results. Numerous variables were shown to be statistically significant, including population, the percent area of retail and of public land/universities, and transfers available to other forms of public transportation. It can thus be argued from this analysis that those variables have a significant impact on ridership. A station being a terminus is close to but not quite statistically significant. Income, employment, attractions, and T lines at each station were shown to be not statistically significant values. On the whole, when compared to actual ridership numbers of stations (courtesy of the MBTA's 2014 service report), this method seems to consistently underestimate ridership at most stations. It is thus possible that, while this process estimated a fairly healthy ridership for the extension, actual ridership numbers may be even higher when the line opens. Between these 7 stations, nearly 20,000 people are estimated to ride the Green Line Extension daily. This accounts for boarding of the line only. The MBTA and MASSDOT estimate that daily ridership, for both boarding and alighting, of the fully functional extension would be 45,000 by 2030. If it is assumed that half of this daily ridership is boarding, this gives a number of 22,500. This discrepancy is close to the general underestimation of other stations in this analysis, though the Route 16 station may not be included in the estimations of the MBTA and MASSDOT. There is much room to improve this analysis—many other variables can be considered to generate a more accurate estimation of ridership. All of the variables used here are also associated with conditions that were thought to be positively correlated to ridership, a more comprehensive analysis could make use of variables that could potentially be negatively correlated, such as crime rates or percentage of industrial area. In the estimation of the new Green Line stations, there are certain variables that are challenging to account for, mostly based around demographic and land use changes that would occur from the opening of the new stations.

**Map and Analysis by Alex Shimmel 5/9/2017 GIS 102: Advanced GIS**

**Map Coordinates: NAD 1983 (2011) State Plane Massachusetts Mainland (Meters) FIPS 2001**

**Projection System: Lambert Conformal Conic**

**Sources of Data: US Census, MBTA, MASSDOT, MASSGIS**

**Zhang, D. and Wang, X. (2014). Transit ridership estimation with network Kriging: a case study of Second Avenue Subway, NYC. *Journal of Transport Geography, 41*, 107-115.**

**https://doi.org/10.1016/j.jtrangeo.2014.08.021**