

Revised proposal for deliverable 1

City of Boston: 15-minute city (Spring 2021, Team 2)	
Contact	<p>Nayeli Rodriguez Technologist, Mayor's Office of New Urban Mechanics City of Boston nayeli.rodriguez@boston.gov</p> <p>(secondary contact) Kat Eshel Carbon Neutrality Program Manager City of Boston katherine.eshel@boston.gov</p> <p>(project manager) Kamran Arif kamran55@bu.edu</p>
Organization	City of Boston
Organization Description	The municipal government for the city of Boston, overseeing all public works and governance for Boston.
Team Members	<p>Katelyn Lee (ktilee@bu.edu) Darren Liu (darrenjl@bu.edu) Khaled Sulayman (khaledsh@bu.edu) Sandra Zhen (szhen@bu.edu)</p>
Meeting Times	<p>Bi-weekly with clients: Thursday, 11-11:30am ET, beginning (02/25/21)</p> <p>Weekly with PM: Friday, 4-4:30pm ET</p>
Project Type	Data Science
Project Description	<p>We would like to research the features that can make Boston a “15-minute city”, where essential services and resources are all within 15 minutes of each parcel in the city. Types of essential services within the city include commercial, education, food access, healthcare, recreation, and social.</p> <p>We will target dense urban areas for the 15-minute city concept. This will be done by identifying parcels of land in the city nearby or within areas with the highest concentration</p>

	<p>of essential services and then finding the shortest paths between them. Then, the areas identified as 15-minute city sections will be displayed on a map to allow for identification of areas in need of more development of essential services.</p> <p>Many cities such as Melbourne that have implemented the plan seem to heavily emphasize bike lanes and access to parks in the scheme, so a worthwhile addition to the project might be identifying locations that can easily be integrated into wide bike lanes and potential public lands to be marked out for green spaces/parks. Further, the C40 mayor's report emphasizes the use of bike lanes as a post-COVID economic recovery idea to build more closely knit cities and improve accessibility and reduce the reliance on public transit and cars that posed a COVID risk for many.</p>
Data Sets	<p>Data Sets:</p> <p>Found or from client:</p> <ul style="list-style-type: none"> • Greenspace data: https://data.boston.gov/dataset/open-space1 • Map and parcel data: https://data.boston.gov/ • Census data <p>From previous Spark! Projects:</p> <ul style="list-style-type: none"> • Boston Master businesses and establishments dataset • Grocers dataset from MAPC <ul style="list-style-type: none"> ○ Farmers ○ Food retailers ○ Schools • Environmental justice demographics
Steps	<p>Step One: Combine the datasets and process them into clear categories of essential amenities. Amenities should be categorized into the following categories: commercial, education, food access, healthcare, recreation, and social . Use the business_type fields within the Boston Master businesses and establishments dataset. Base geographic unit will be land parcels (parcel data from Analyze Boston).</p> <p>Step Two: Identify the 15-minute city areas. Which of Boston's neighborhoods/zip codes contain parcels with essential amenities within a 15-minute commute? Note that commute refers to walking or biking. Employ the CLARANS algorithm to cluster areas with dense essential amenities.</p> <p>Step Three: Analyze the missing essential amenities within zip codes that do not meet the requirements for a 15-minute city. Which areas are missing which amenities?</p> <p>Step Four: Conduct demographic analysis on the parcels of land and compare the residents in 15-minute communities to non 15-minute communities. What are the demographic</p>

	<p>features of residents living within the 15-minute communities and those not living in them? Features to include: race, income, language spoken, housing density, etc.</p> <p>Step Five: Create a map of the city displaying the 15-minute compliant areas, and also the areas which need more access to essential amenities and the amenities they require. Final product should be at least usable by clients to display needs throughout Boston—.kml file should be considered to create a clickable map.</p>
Questions to be answered in Analysis	<ol style="list-style-type: none"> 1. What amenities in Boston belong to each of the six categories of essential amenities necessary for a 15-minute city? 2. What percentage of parcels are within 15 minutes of essential amenities? 3. Which areas of the city are underserved in terms of a lack of essential amenities, and which amenities are they missing? 4. What are the demographics of the underserved (non-15-minute) communities?
Limitations	<ul style="list-style-type: none"> • The businesses and amenities generally are not necessarily clearly in one category or another—some manual decision-making may be required to determine which amenities are in which categories. Some datasets have clearly defined amenity types while others require a categorization algorithm to determine amenity type. • Walking and biking speeds can vary dramatically per person. For biking time specifically, there may not be complete data for safe bike routes, which would impact the distances defined as “within 15-minutes” of any given parcel. • Lack of specific and consistent location data in each dataset could prevent accurate geospatial analysis.
Data-specific Limitations	<ul style="list-style-type: none"> • Boston Master businesses and establishments dataset <ul style="list-style-type: none"> ○ This dataset lacks street address and has major inconsistencies in business type. It may not be necessary to include street address, but the inconsistencies in the business type classification must be resolved. • Grocers dataset from MAPC <ul style="list-style-type: none"> ○ MAPC contains a large number of datasets, which must be sorted through to determine usefulness. Not all datasets contain necessary data. • Greenspace data <ul style="list-style-type: none"> ○ The greenspace data lacks longitude and latitude. This may be a significant problem, because clustering and distance analysis cannot be carried out with street addresses alone. Units/parcels and amenity type in this data are also unclear.

Additional Information	<p><u>Tools and Methods</u></p> <p><i>Data processing:</i> Pandas and NumPy will be used to clean and process the datasets into one master dataset.</p> <p><i>Categorization:</i> Some categorization algorithm must be used to categorize the data into the appropriate amenities.</p> <p><i>Clustering & Graph search:</i></p> <p><u>CLARANS algorithm</u> - Works much better for high-dimension spatial clustering than DBScan that becomes inefficient and inaccurate when multiple attributes are added. We have multiple essential service types so it would be best to use CLARANS.</p> <p><u>Kruskal's algorithm implementation in Python</u> - To find the shortest distance between multiple sets of points, this will be done to find the parcels that are within a given short distance from all necessary amenities.</p> <p><i>Data Visualization:</i> Matplotlib, Seaborn, Tableau, Folium for visualizing the geospatial data as needed.</p> <p><i>Final product:</i> Note that clients have no data processing ability. Maybe Simplekml to create an interactive map.</p> <p><u>Optional Additional Steps with Extra Time</u></p> <p>Examine the variance in type/density of amenities in each area. For example, there may be a difference between 15-minute areas that include only one park or grocery store and areas that have a large variety of each type of amenity.</p> <p>Analyze the paths between essential amenities to determine where additional bike lanes and other paths may be necessary. This could be done by using Kruskal's or Dijkstra's algorithm to find shortest routes between different areas, and analyzing junctions of paths to find which could be easily interconnected to most benefit.</p> <p><u>Further Background Information</u></p> <p>An example of the type of final product desired:</p> <p><u>https://app.developer.here.com/15-min-city-map/</u></p>
------------------------	---