Health and Obesity in Boston

Ann Ming Samborski, Amalia Safer, Viviana Yee, Max Mesirow | Supervised by Andrei Lapets | Boston University Computer Science Department

Introduction

- Does a neighborhood's lack of access to healthy locations manifest in higher
- Aim to find the correlation between access to healthy locations and obesity percentages
- Utilize MBTA data to determine how accessible healthy locations are from obesity locations
- Determine whether these healthy programs/locations are located in neighborhoods with higher levels of obesity
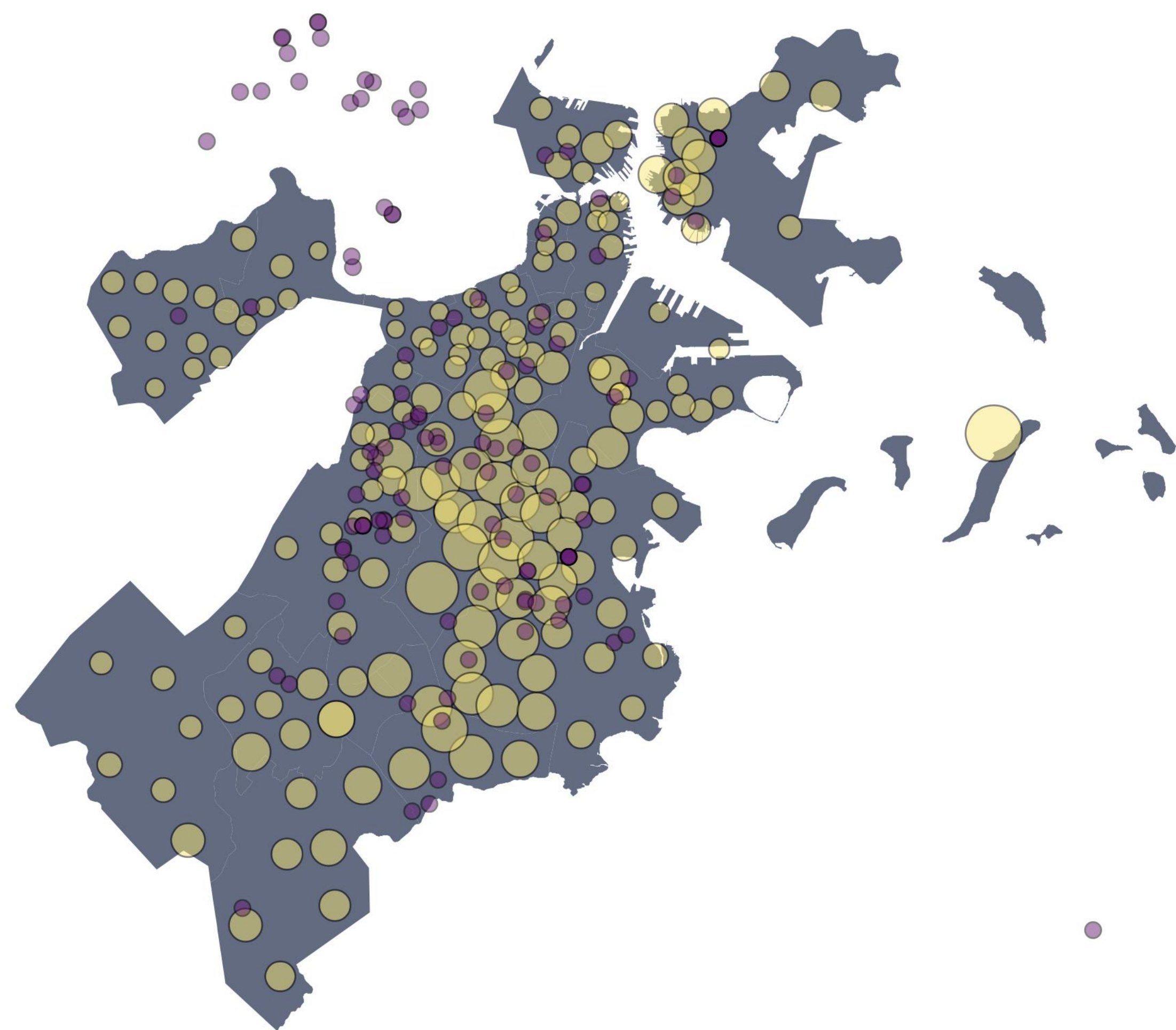

Figure 1: A map of Boston marked with purple dots denoting healthy locations and yellow dots showing an obese area. Larger dots indicates a higher percentage of obesity in the area's population.

Datasets

- Urban orchard locations
- Healthy corner stores
- Community culinary and nutrition programs
- Locations of high adult obesity
- MBTA bus and subway Stops
- Big Belly trash receptacle locations (control)
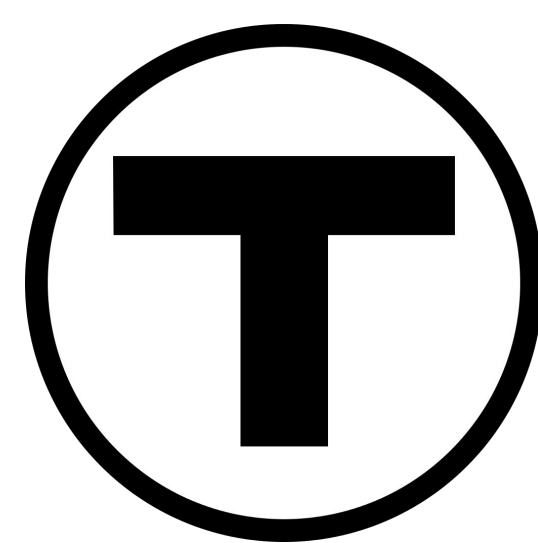
Transformations

- **Preprocessed data**
  - Aggregate all healthy locations into a single dataset
  - Identify the list of MBTA stops within a mile of each location in our database
  - For each obesity location, matched the closest health location
- **Shortest Path**
  - Finds minimum travel time (walking/bus/T) from obese locations to a healthy location
  - Use networkx to create network of weighted subway routes
  - Utilize Google Maps API to calculate times for walking to the nearest T stop and solely using the bus
- **Linear Regression**
  - Calculates the best fit slope for the data points (minimum travel time vs percent obesity) from shortest path

Results

- Found positive correlation between time it takes to get to a healthy location and obesity
- For every 5 minutes it takes to travel to a healthy location, rate of obesity increases by 1%
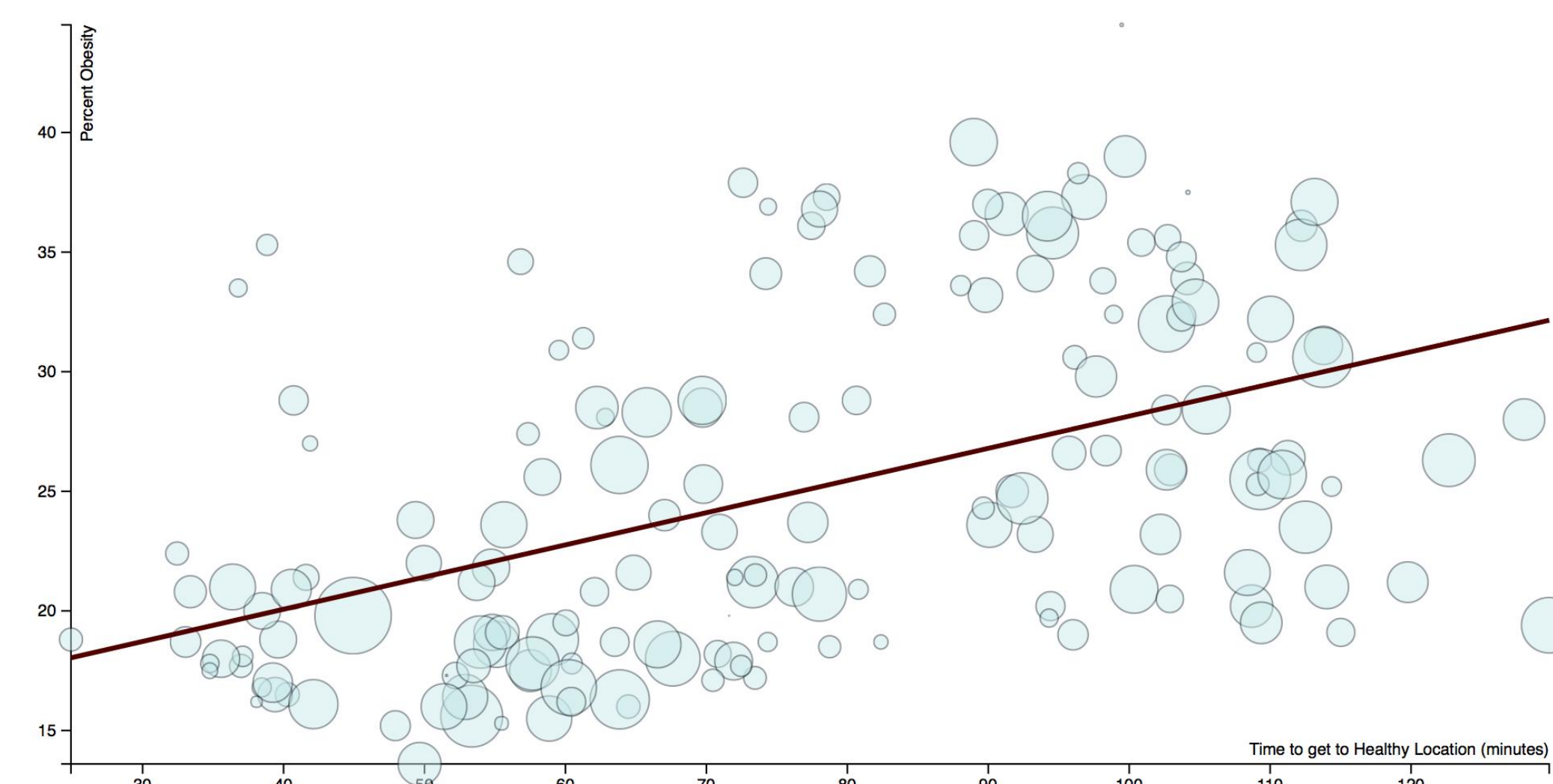

Figure 2: A graph of each locations' estimated travel time from a healthy location vs the obesity rate in that location. Larger dots represent larger populations. The best fit line calculated in the Linear Regression transformation is drawn to show our results.

Lingering Questions

- Would the introduction of a healthy location into an obese area lower the obesity rate over time?
- Could we predict an optimal point or points to place a healthy location based on our data?