

Optimizing Coffee Shop Commutes

Vee Nguyen
Tung-Ho Kang
Christopher Trinh

Purpose

A typical morning commute in Boston often proves to be quite challenging to navigate from day-to-day, especially for those that make it a habit to stop at a coffee shop for their daily caffeine intake before heading to the train station to travel to work. Our goal is to optimize this morning commute by determining the most optimal placement of coffee shops along the MBTA Green Line.

Methodology

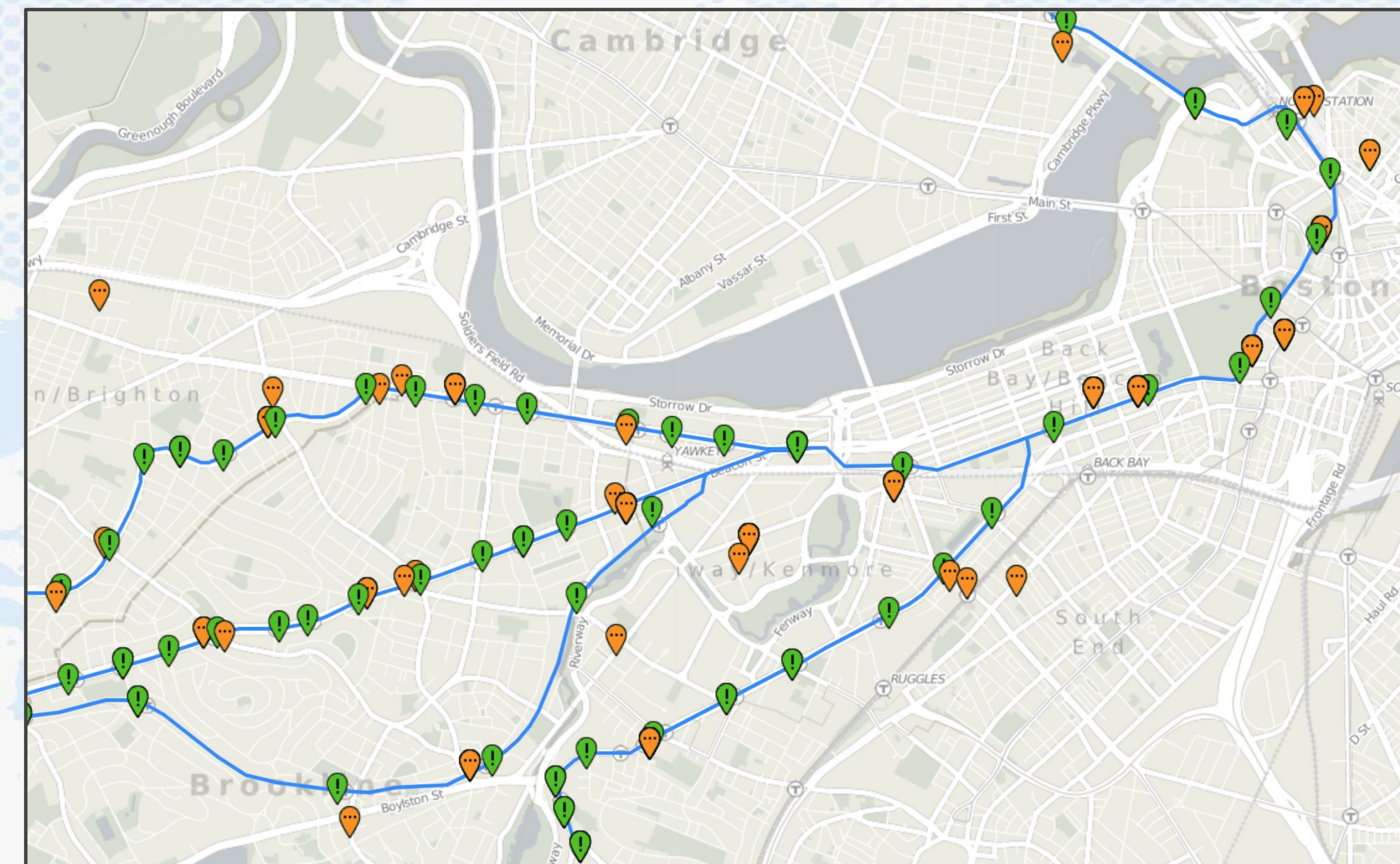
Our first course of action was to identify which MBTA stations along the Green Line were most busy, in order to determine which stations should be closer to a coffee shop; we did this by assigning each station a weight based on number of daily departures from that station. The data sets we used are as follows: Boston coffee shops coordinates, MBTA Green Line station coordinates, and 2018 Uber mean travel times. In order to calculate the new optimized placement of coffee shops along the Green Line, we had to: (1) calculate each coffee shop's distance from their nearest train station, (2) calculate the average distance (in meters) from a coffee shop to their nearest train station, (3) perform the k-means algorithm to calculate new coffee shop locations, and then pick the new closest location to each station, and (4) calculate how much more efficient the new coffee shop locations are in terms of Uber mean travel time.

Assumptions

We will assume morning commute times to be any weekday from 6am to noon. We will also assume that a majority of commuters take Uber rides from a coffee shop to the nearest train station, which would allow us to provide an estimate for how much Uber mean travel times could decrease if coffee shop locations were moved closer to a train station as our experiment. In order to assign a weight to each train station for use in the k-means algorithm, we assumed that the number of departures from a station is reflective of how "busy" it tends to be on an average weekday morning. We calculated the average number of departures of each station and then divided that number by a factor of 100 to determine its weight, represented by assigning an additional number of data points to that location.

Challenges

We struggled the most obtaining our data from the MBTA, due to how limited the API was for our purposes. We could only perform two relevant queries: (1) a query that listed information about every Green Line stop; (2) a query that gave the time and stop ID of every departure on a given day. However, we needed a data set that contained both the stop information and every departure from that stop on a given day, so we joined these queries together and extracted a combined data set, taking every departure that matched a stop ID and placing it in the corresponding stop information data entry.



Map of original coffee shop locations (orange) vs new locations (green).

Conclusion

After running k-means on the data set we generated, we calculated the average distance from the coffee shops to train stations. We then compared the result to the original average distance and found a significant 91.65% decrease in the average distance. To put this result in context, we gathered the aggregated Uber hourly mean travel time of a weekday in Boston. Again assuming that people travel from a coffee shop to the nearest train station using Uber, we can conclude that Uber cars would be driving 91.65% less during the morning weekday commute in Boston. During 2018, Uber cars spent a total of ~199,109 hours driving at 6-7am; with these proposed coffee shop locations, they would instead spend ~16,625 hours driving during 6-7am. In the future, we would like to have a more comprehensive data set. For example, it would be useful if we could find a data set that includes traveling intentions or more precise locations of people riding with Uber, as that would allow us to be more conclusive about how many hours people spend riding from a coffee shop to a train station via Uber, and thus how many hours would be saved with the proposed coffee shop locations.