ComputerScience

BOSTON UNIVERSITY

Boston Hubway Trips in Relation to Average Neighborhood Income and Population

Ricardo Ballesteros | Jake Bloomfeld | Daniel Medina
rfballes@bu.edu | jtbloom@bu.edu | medinad@bu.edu
CS 591 L1 Fall 2017

INTRODUCTION

Hubway is a public bike-share system serving the people of Boston, Brookline, Cambridge, and Somerville. With roughly 1800 bikes in practically 200 stations, the Hubway serves as a fun, affordable, and convenient transportation option for quick trips around Boston and its surrounding municipalities. With some further thought, we proposed our final questions: is there a correlation between income per capita in Boston's neighborhoods with the number of Hubway trips taken in that neighborhood? Also, is there a correlation between neighborhood population with Hubway trips taken in that neighborhood?

DATASETS

Below are the datasets and data we used from each:

Hubway Station Locations
- This dataset consists of the name of each Hubway station, along with its longitude and latitude, number of bike docks, and municipality.
- Source: Analyze Boston.

Hubway Trip History
- This dataset consists of trip histories of a given month, including the start station name & ID and end station name & ID
- Source: Hubway System Data.

Per Capita Income by Boston Neighborhood
- This dataset consists of each Boston neighborhood with its average income per capita.
- Source: Census Bureau

Population by Boston Neighborhood
- This dataset consists of populations by Boston neighborhoods.
- Source: Census Bureau

Boston Neighborhoods
- This dataset GeoJSON polygons, which are enclosed areas within certain geographical coordinates that represents the different neighborhoods in Boston.
- Source: Boston OpenDataSoft

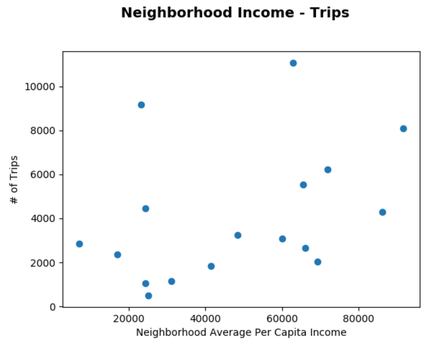
PROBLEM

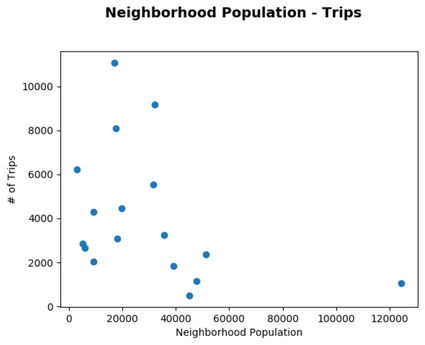
To answer our questions and perform the necessary analysis, we use and manipulate several datasets. In the Hubway Trip History dataset, given the coordinate points of the station, we check if the station is within a certain Boston neighborhood given by the Boston Neighborhoods dataset. Then, once the Hubway stations are mapped to a certain neighborhood, using the Per Capita Income and Population datasets, we see if there is a correlation across the following dimensions: Incoming/outgoing Hubway trips per neighborhood, average per capita income by neighborhood, and population by neighborhood. Seeing if such a correlation exists can pose further questions for discussion, such as, are there more trips being taken to/from wealthy neighborhoods than poor neighborhoods? Are wealthy/poor neighborhoods or high/low populated neighborhoods over/underutilizing available Hubway bikes?

RESULTS

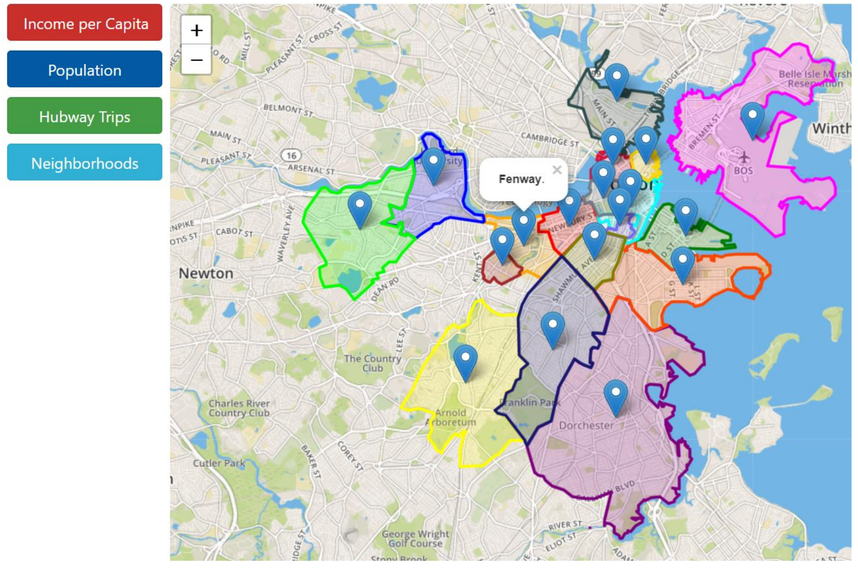
In order to perform statistical analysis, we used the Pearson correlation coefficient to determine the strength and direction of the linear relationship between the number of Hubway trips per neighborhood with its average income and population. Using a sample of trips taken in one month, we were able to obtain the following results:

		Correlation coefficient	P-value
Average Per-Capita Income by Neighborhood	# of trips	0.376	0.133
		Correlation coefficient	P-value
Average Population by Neighborhood	# of trips	-0.377	1





VISUALIZATION



CONCLUSION

Based on our calculations, there is a moderate positive correlation between the number of Hubway trips per neighborhood and average neighborhood income.

There is also a moderate negative correlation with the average neighborhood population with a high p-value, meaning that this observation is a non-significant result. Based on these results, we can learn that neighborhoods with higher frequencies of trips tend to have a higher average per-capita income.

FUTURE WORK

For expanding on this project in the future, there are several aspects to consider:
- Combining Hubway trip history data sets from past several years to get an even more accurate correlation data and to observe the system's behavior over time.
- Using datasets of income per capita and population in municipalities outside of Boston to get a bigger picture.