# A Comparison of Boston's Neighborhoods to Chicago's

*William Dean*

## Overview

### Motivation

I am currently living in Boston, MA, but plan on relocating next year to Chicago. I have visited a few times but, from the large size in comparison to Boston, I have found it difficult to grasp which areas in Chicago I would enjoy let alone want to live. In general, in my recent travels, I have found difficulty exploring bigger cities. When spending a few days in a certain city, I often end up spending a lot of time traveling within the city, not having enough time to connect with the people and the cultures.

### Goal

The **goal of this project** is to use data from areas and neighborhoods of Boston to see how communities compare to others within it and to Chicago.

When looking each neighborhood, there a few things that I am interested in measuring:

- What are the people like? How are the people that live in the area and what do people who visit think of the area and their time there.
- What are the living situations like? What kind of houses are in the area.
- What is the culture like? What are these communities known for and what are people talking about?

The question I want to answer with this project is: **What neighborhoods in Chicago are similar to the neighborhood I currently live in Boston?** That is, if I were to relocate to Chicago from Boston, what areas would I like based on my current location.

### Data

I wanted data that would reflect all these measure. Here are a few sources I found.

#### Airbnb

The Airbnb Datasets provides insight into many cities in the US and around the world. Not only are there reviews written by people who use the service, but also the listings have information about the neighborhood, the house, and the host.

Using this information will provide information for the people who live in the neighborhood, details of the neighborhoods from the hosts, and the experiences of the people who visit.

#### Twitter

Twitter data gives insight to what is going on within a neighborhood, and what people are talking about. This accomplished trying to capture the neighborhood culture and potentially what people are like.

### Methods

Doing the analysis, I mainly used a bag of words approach to the text analysis. For a neighborhood's set of reviews, I collected all the defining words for a given neighborhood as well as important sets of two words for a neighborhood.

When comparing two or more neighborhoods to each other, I used latent Dirichlet allocation (LDA) to group into a set of $k$ topics.

This approach seemed appropriate since it will group neighborhoods with neighborhoods whose texts are similar and are talking about the same topics.

## Calibration: The Neighborhoods in Boston

### People from each Neighborhood

#### Hosts

Each Airbnb Host provides a description of them self and their household. Using the methods stated above, I grouped terms from each neighborhood. This process appeared to catch terms associated with people and aspects of their lives. Consider the words associated for Brighton, MA below:

| Brighton Neighborhood Bigrams |
|---|
| england france |
| tech company |
| accountant downtown |
| action movies |
| active person |

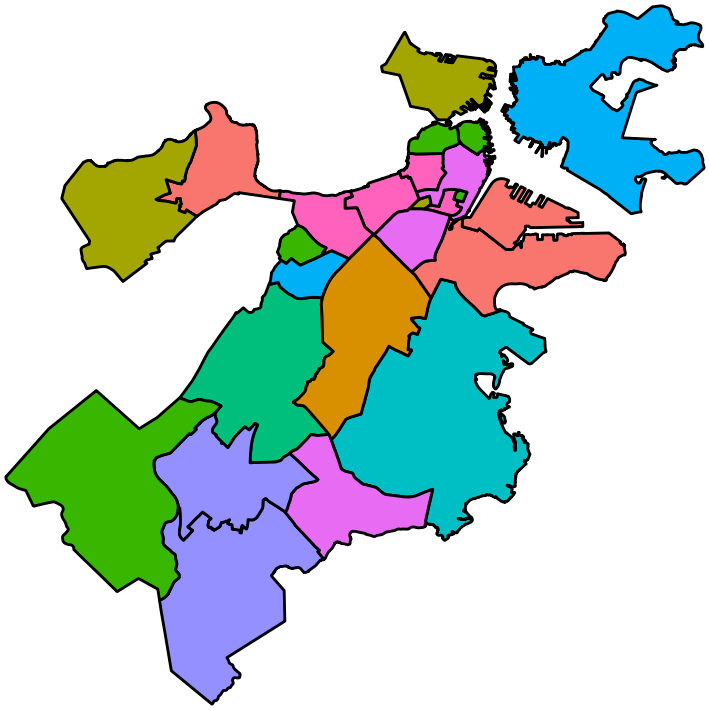
The third row in context of it's host description is: "Lorin works as an **accountant downtown** and takes the expressway bus which is a quick 20 minute ride."

Doing this for all of the neighborhoods, and grouping all 25 Boston neighborhoods into 15 topics, we see the following result:

Regions of the same color (same grouping) appear either next to each other or are in similar proximity to downtown. This seems likely with people living outside of downtown and commuting into the city. For example, the Fenway area is grouped with Longwood area. These areas are located around many hospitals in Boston. Other grouping is the North End and Downtown. These areas are close to many financial jobs and have many people that would be further in a career. This does appear to separate Boston into areas that have similar host.

**People's Reviews**

Considering all the reviews of the Airbnbs for the neighborhoods of Boston and which area's reviews are most similar, we see the results below:

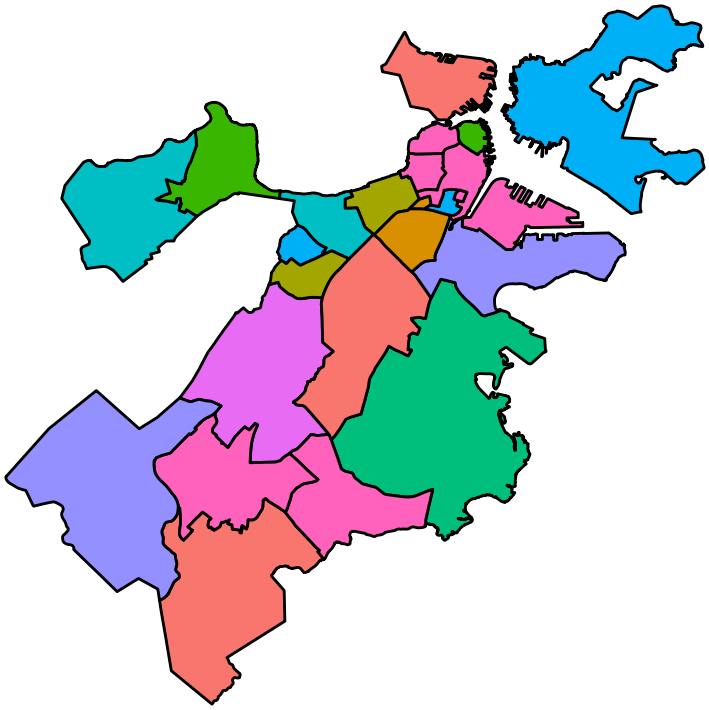


The contents of what people are saying into their reviews seems to depend on what neighborhood they are staying. Again, many neighborhoods far from the heart of Boston look to be grouped together, and as the neighborhoods get closer, there appear to stay grouped together. South Boston and the Waterfront are grouped together, and the Fenway area is grouped with Back Bay. The city appear to be separated into reasonable sections based on shops and attractions throughout the city.

## Boston Neighborhoods

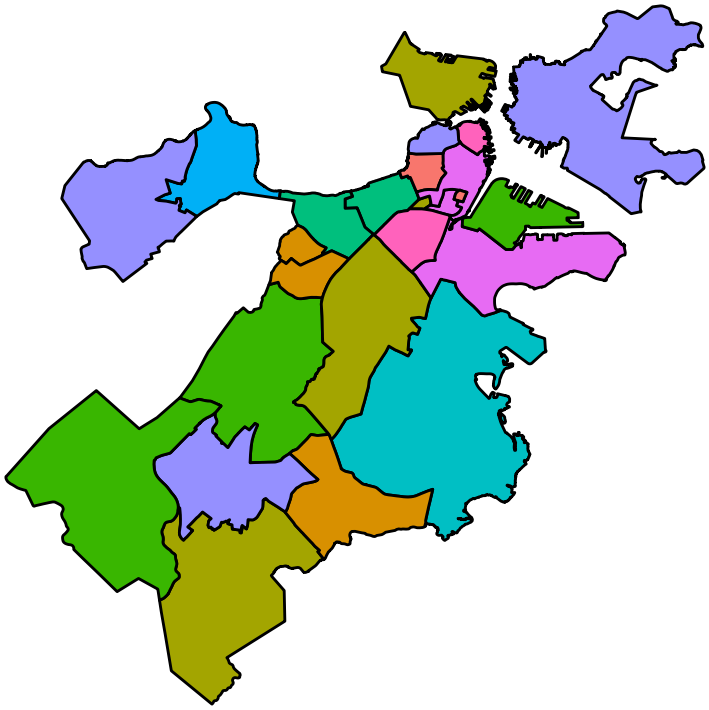
### Neighborhood Descriptions

Every Airbnb host provides a description of the neighborhood where their property. Using this, we can see which neighborhood descriptions are similar. We see that some of the adjacent neighborhoods in Boston are grouped together. For example, Downtown and Seaport are grouped together. Other than that, many of the neighborhood are described differently enough from each other that they are not grouped together.



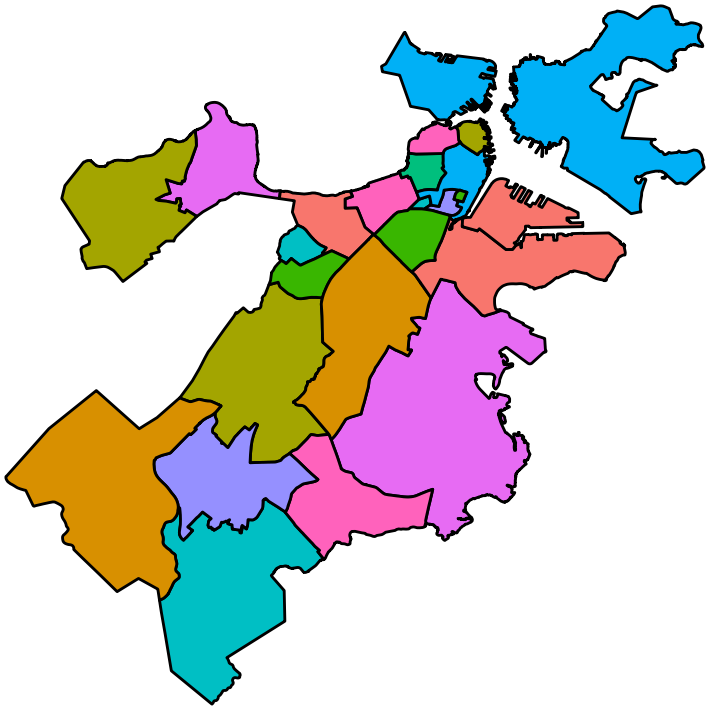
**Housing**

Using each Airbnb listing's detailed description of the property, we gain information about the type of homes in each location are where housing is similar. One interesting grouping is Back Bay with Fenway. This areas is filled with brownstone housing, having most streets with the same style of house.



**Culture**

Looking at tweets associated with each neighborhood either through direct mention or in a hashtag, we can pick up on what people in each area are talking about. The tweets were sample during the week of November, 30th. Here are the clusters:



## Calibration: Comparison of Boston and Chicago

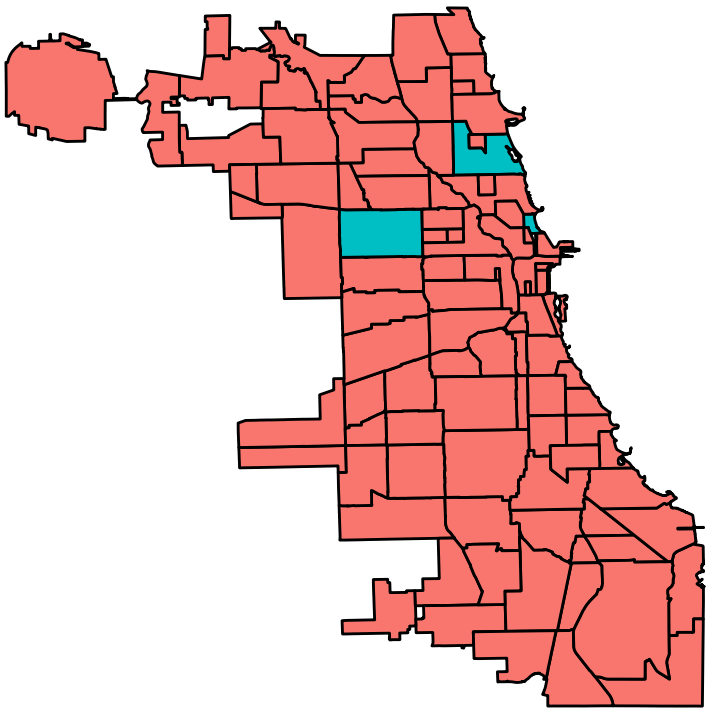
After looking at Boston, there is interpretability within each section above and does appear to give some insight into the similar neighborhoods in Boston. Let's consider how another Boston's neighborhoods compare to the ones in Chicago and see if we can extend this onto another city.

### Tweets

This data will help understand which neighborhoods in Boston are talking about similar topics as the neighborhoods in Chicago. Using LDA with 25 topics for all 25 Boston neighborhoods, we can check if any of of the Boston neighborhoods match with neighborhoods in Chicago and which neighborhoods are not grouped together.

To have some sense of comparison, below are the areas in Chicago that are similar to Downtown Boston based of the twitter data. Interestingly enough, the projected similar locations in Chicago are in fact close to the downtown area of Chicago as well. This worked out well and suggests that this method of comparison is catching similarities between the cities.

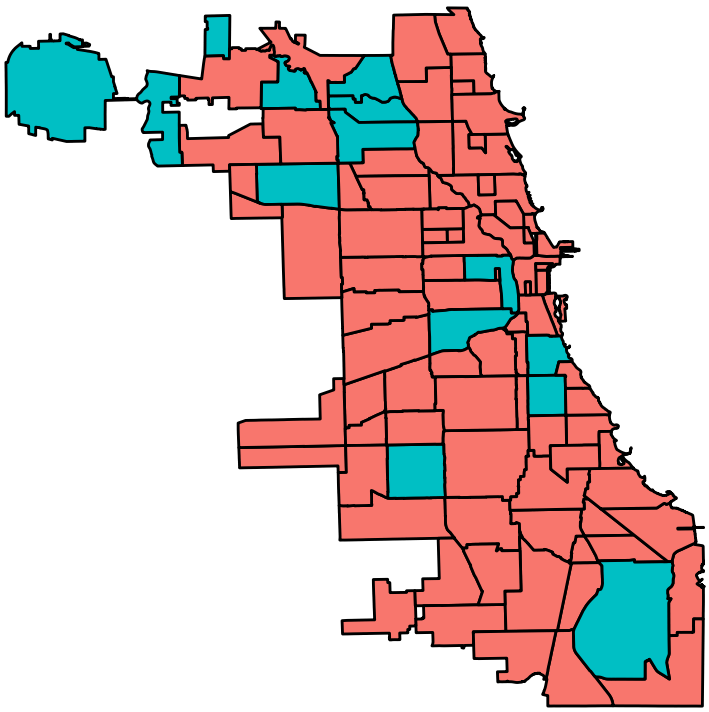
| City | Neighborhood |
|---|---|
| Boston | West End |
| Boston | Downtown |
| Chicago | Gold Coast |
| Chicago | Lake View East |
| Chicago | Humboldt Park |
| Chicago | Lakewood - Balmoral |



## Moving to Chicago

It appears this method does have some grounds in comparing not only neighborhoods in a certain city, but also has some ability to catch where similar places are in different cities as well. With that being said, let's proceed to find which neighborhoods in Chicago are similar to the neighborhood I currently live: Allston, MA.

Below are the results of all neighborhoods in Chicago that were flagged as similar:



Although I am not too familiar with Chicago, there is one area that I am aware of. One of the marked areas in Little Italy and the West Loop. In this area, there's the University of Illinois at Chicago along with its students, many shops and restaurants. Similarly, Allston is close to Boston University and has many of its students live in the area during and after college. There are many strips in Allston that provide many food options and is a big destination for eating.

## Conclusion

The neighborhoods in Boston have a interesting relationship with each other and are similar in as many ways as they are different. There are neighborhoods in Boston that provide many of the same attractions, but differ other aspects like in housing and residents. Using the LDA topic model, these similarities and differences did show with in the clusters which had interpretability that went along with characteristics of the areas in Boston. Extending this model onto Chicago, we saw that it was able to match the downtown area of Boston to areas around Downtown in Chicago. Knowing that this method has some sense of calibration, this process was able to be used to find areas in Chicago that are similar to Allston. Allston was mapped to a college area in Chicago which is quite reassuring. I am unaware of other areas that were marked similar to Allston, but they are on my radar for my next trip to Chicago.

### Next Steps

Even though I was able to find out what locations are similar to Allston, I think this is just one approach to the question. Continuing on, I would like to see what other data I could use to expand on this. If there is data that provides a fine-grain look into the neighborhoods, I may be able separate each neighborhood into further classifications.

## Limitations with the Data

The Boston Airbnb Data set had a lot more information about each listing than the listings from Chicago. When doing the analysis, I was limited to using only the twitter data and the housing listing for the comparison because that was all that was available. If I can get the full listings from the Chicago listings, it would interesting to see how results change.

## What I Learned

With this project, this was the first time where I used only text data to answer questions. I got very familiar with the packages in R for text mining, mainly `tidytext`, `tm`, and `topicmodels`.