

The Battle of Neighborhoods

Brendan Hargrove

A large, dark blue, diagonal shape that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.

Problem Statement

Many people are forced to relocate for their careers and/or choose to move to a new city for better opportunities. However, most still have a home neighborhood they are fond of, or a favorite neighborhood from their hometown.

This notebook will attempt to characterize the venues of a chosen neighborhood from one city (Houston, TX) and compare it to all neighborhoods of a different city (Seattle, WA), returning the neighborhoods with similar businesses, restaurants, etc.



Data and Libraries

Data Sources

Neighborhood names for Houston:

https://en.wikipedia.org/wiki/List_of_Houston_neighborhoods.

Neighborhood names for Seattle:

https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Seattle.

Location data obtained from GeoPy Geocoder:

<https://geopy.readthedocs.io/en/stable/>.

Venue data obtained from Foursquare:

<https://developer.foursquare.com/>.

Libraries used in analysis:

- Pandas
- Numpy
- Urllib
- BeautifulSoup
- GeoPy
- Json
- Matplotlib
- SKLearn
- Folium.

Exploration of Data

Input Neighborhood: Houston Heights, Houston, TX

Queried City: Seattle, WA

Total neighborhoods/districts in Seattle (from wikipedia): 127

Total neighborhoods/districts with location data: 111

Neighborhoods dropped: 16

Final neighborhood count (including input neighborhood): 112

Location information for North Beach / Blue Ridge not unavailable
Location information for North College Park
(Lifton Springs) not unavailable
Location information for Portage Bay / Roanoke not unavailable
Location information for Pike-Pine Corridor / Pike/Pine not unavailable
Location information for International District ("ID") not unavailable
Location information for Central Area / Central District ("CD") not unavailable
Location information for Cherry Hill & Squire Park not unavailable
Location information for South End not unavailable
Location information for Dunlap / Othello not unavailable
Location information for Rainier Beach / Atlantic City Beach not unavailable
Location information for Mid Beacon Hill (Maplewood) not unavailable
Location information for South Beacon Hill / Van Asselt not unavailable
Location information for Industrial District not unavailable
Location information for North Admiral / Admiral District not unavailable
Location information for Junction / West Seattle Junction / Alaska Junction not unavailable
Location information for Seaview / Mee-Kwa-Mooks not unavailable

	Neighborhood	Latitude	Longitude
0	North Seattle	47.660773	-122.291497
1	Broadview	47.722320	-122.360407
2	Bitter Lake	47.726236	-122.348764
3	Crown Hill	47.694715	-122.371459
4	Greenwood	47.690981	-122.354877
5	Northgate	47.713153	-122.321231
6	Haller Lake	47.719748	-122.333751
7	Pinehurst	47.603832	-122.330062
8	Maple Leaf	47.693987	-122.322905
9	Lake City	47.719162	-122.295494

Venue Acquisition

Venue data obtained for each neighborhood using Foursquare's Venues/Explore endpoint.

Venues were then grouped by neighborhood, one-hot encoded, and averaged by venue category.

[illegible]

Most Common Venues

To characterize each neighborhood for clustering, the 20 most frequently occurring venue categories were calculated and ranked.

Venue category occurrence and rank become the primary metric with which to group similar neighborhoods.

The top 20 venue categories for Houston Heights are noted to the right.

Neighborhood	Houston Heights, Houston, TX
1st Most Common Venue	Coffee Shop
2nd Most Common Venue	Burger Joint
3rd Most Common Venue	Park
4th Most Common Venue	Pharmacy
5th Most Common Venue	Flower Shop
6th Most Common Venue	Furniture / Home Store
7th Most Common Venue	New American Restaurant
8th Most Common Venue	Thrift / Vintage Store
9th Most Common Venue	Mexican Restaurant
10th Most Common Venue	Trail
11th Most Common Venue	Diner
12th Most Common Venue	Pet Store
13th Most Common Venue	Sandwich Place
14th Most Common Venue	Gift Shop
15th Most Common Venue	Cosmetics Shop
16th Most Common Venue	Spa
17th Most Common Venue	Pizza Place
18th Most Common Venue	Italian Restaurant
19th Most Common Venue	Movie Theater
20th Most Common Venue	Indian Chinese Restaurant

K-Means Clustering

K-Means clustering was employed to determine neighborhood similarity. In order to limit the number of neighborhoods within each cluster and thereby narrow the number of matches with our input neighborhood, a large cluster count was used (25).

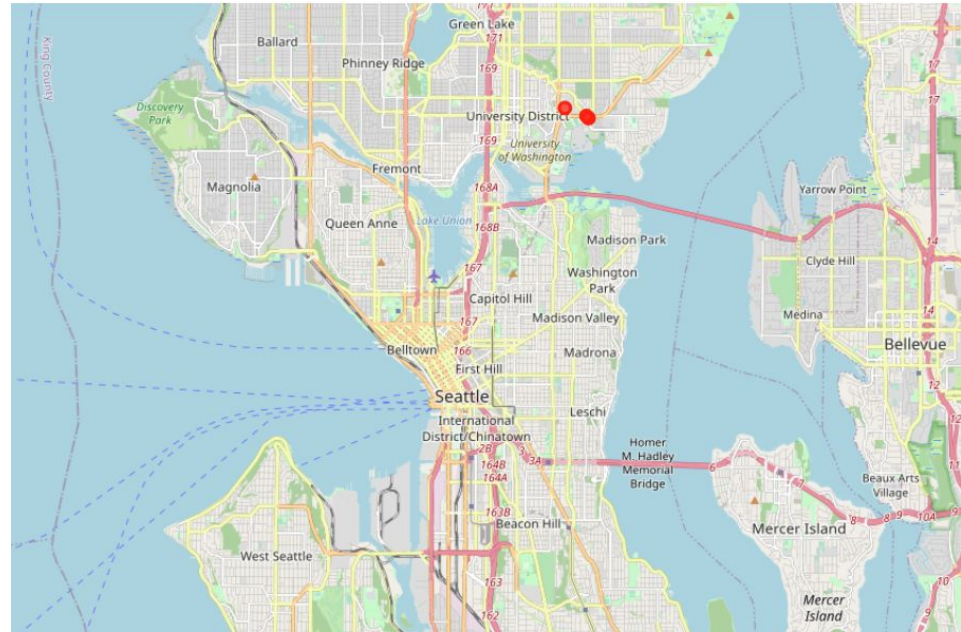
The resulting dataframe was then sliced to find all neighborhoods within the same cluster as Houston Heights (cluster 24).

	Neighborhood	Latitude	Longitude	Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	...
0	Houston Heights, Houston, TX	29.797687	-95.398446	24	Coffee Shop	Burger Joint	Park	Pharmacy	Flower Shop	Furniture / Home Store	...
1	North Seattle	47.660773	-122.291497	24	Coffee Shop	Furniture / Home Store	Clothing Store	Women's Store	Arts & Crafts Store	Pizza Place	...
25	University District (U District)	47.661191	-122.292083	24	Coffee Shop	Furniture / Home Store	Clothing Store	Arts & Crafts Store	Pizza Place	Italian Restaurant	...
26	University Village	47.662740	-122.298925	24	Pizza Place	Arts & Crafts Store	Burger Joint	Thai Restaurant	Coffee Shop	Italian Restaurant	...

Mapping & Results

The neighborhoods that fall within Cluster 24, the same cluster as Houston Heights, all occur within a small area of Seattle centered around the University of Washington.

Although Houston Heights is not near a university itself, it shares many qualities typical of neighborhoods near college campuses. These qualities are well illustrated in the frequency occurrence of venues within Cluster 24. Coffee shops, parks, casual shopping, and American dining venues all feature prominently within these neighborhoods.



Conclusion

In conclusion, I have quantitatively demonstrated that the neighborhoods in Seattle most similar to Houston Heights are those located immediately around the University of Washington, within University Village and University District. A former resident of Houston Heights, if wishing to find neighborhoods with familiar amenities, would want to focus their search for housing within this area. Similarly, if a business with a successful location within Houston Heights was looking to expand into a new market, I might suggest these neighborhoods as likely contenders.

Future work for this notebook:

- Statistically determine ideal cluster count to minimize inter-cluster error
- Additional data cleaning to exclude any neighborhoods that also occur as recognized districts
- Add additional cities into the analysis to look for better matches in other cities, or cities near Seattle within commuting distance.