

Segmenting and Clustering Towns in Massachusetts Based on Venues

Vivian Xia

January 4, 2021

1. Introduction

1.1 Background

Massachusetts is the 15th most populous state in the United States. It is also home to over one hundred colleges and universities including Harvard University, MIT, Tufts University, and my own undergraduate college, Bentley University. Many college students end up residing in Massachusetts after completing their education. Therefore, it would be useful information to understand the many towns of Massachusetts.

1.2 Problem

This project aims to segment and cluster similar towns based on venues in Massachusetts. The types of venues that are frequented the most in each town are used to cluster towns with one another. The project will locate towns that are similar to other towns and their amenities to allow for a more informed decision on where to move to or stay.

2. Data acquisition and cleaning

2.1 Data sources

The list of towns in Massachusetts was scraped from a [Wikipedia page](#) for a total of 351 towns. The list of latitude and longitude points of the state and each town were found through the Google Maps Geocoding API. The venues within a 500-meter radius of each town were found through the Foursquare API.

2.2 Data cleaning

The data scraped from the Wikipedia page was prepared by removing missing values. Those missing values were associated with rows that were not classified as towns or cities, thus insignificant to the goal of the project and removed. The column heading containing towns was changed from “Municipality” to “Neighborhood” to allow for more ease in merging the venue information from Foursquare to the main dataframe. Additional columns in the dataframe are added including the state, to be used in the creation of a map, as well as the latitude and longitude, to be used to get the nearby venues of each town.

3. Methodology

3.1 Exploratory data analysis

The location of each town is denoted by its latitude and longitude. Using its latitude and longitude, a map of Massachusetts, Figure 1, is created and each town is marked.

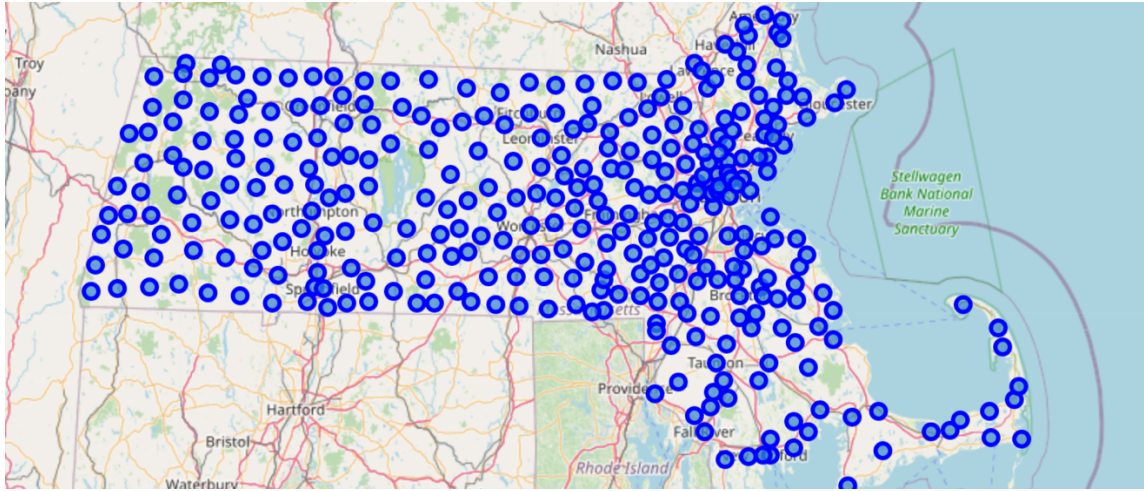


Figure 1. Map of the towns.

3.2 K-means clustering

To figure out the optimal number of clusters to use, the elbow method is used to visualize and evaluate the best value of k . The k is chosen where the graph bends at the “elbow” where the distortion and inertia decrease. As seen from figures 2, and 3, the optimal number of clusters to use is 5.

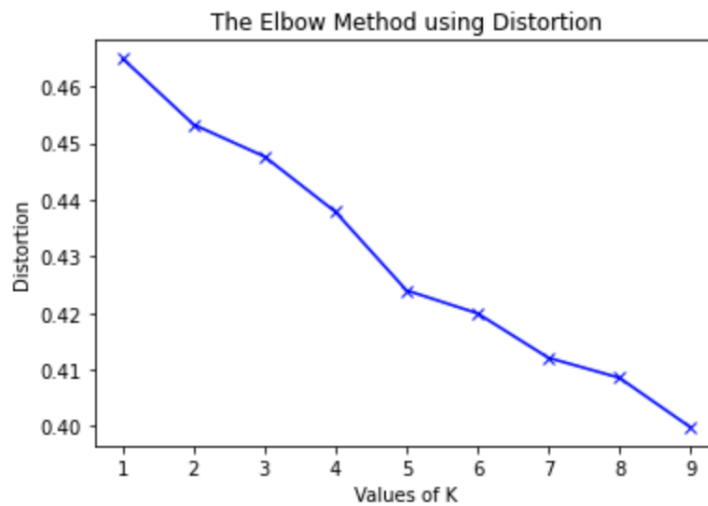


Figure 2. Elbow method using distortion visualization.

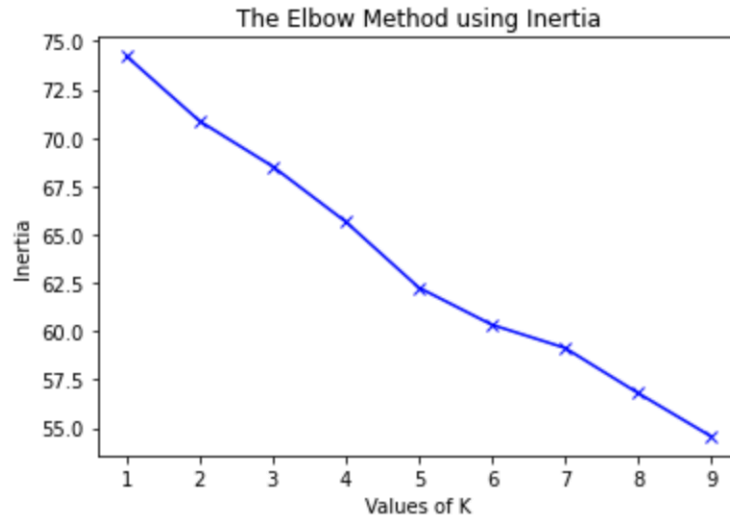


Figure 3. Elbow method using inertia visualization.

The k-means clustering method is used as an unsupervised machine learning to segment towns based on their top 10 most frequented venues. The venues are organized by their venue category and the frequency of occurrence in each category by town is averaged. Using the venues, each town is assigned one of the five clusters.

4. Results

The resulting five clusters are shown in Figures 4-8.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
32	Boylston	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
39	Pelham	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
40	Tyringham	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
44	Mount Washington	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
49	Dunstable	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
334	Rowe	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
335	Monroe	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
339	Athol	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
346	Orange	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
350	Amherst	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

96 rows × 11 columns

Figure 4. Cluster 1

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
27	Bolton	Park	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Fast Food Restaurant
36	Stow	Pizza Place	American Restaurant	Park	Women's Store	Fast Food Restaurant	Event Space	Exhibit	Eye Doctor	Fair	Farm
53	Reading	Pool	Playground	Park	Farm	Ethiopian Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair
90	Newbury	Butcher	Park	Home Service	Women's Store	Farm	Event Service	Event Space	Exhibit	Eye Doctor	Fair
95	Burlington	Park	Gas Station	Department Store	Fast Food Restaurant	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market
159	Hampden	Post Office	Park	Italian Restaurant	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair
165	Charlton	Park	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Fast Food Restaurant
189	Townsend	Golf Course	Women's Store	Ethiopian Restaurant	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market	Fast Food Restaurant
209	Bourne	Bar	Post Office	Pharmacy	Park	Farm	Ethiopian Restaurant	Event Service	Event Space	Exhibit	Eye Doctor
225	Swansea	Dessert Shop	Park	Gift Shop	Farmers Market	Event Space	Exhibit	Eye Doctor	Fair	Farm	Women's Store
254	Somerset	Baseball Field	Park	Harbor / Marina	River	Women's Store	Fair	Event Service	Event Space	Exhibit	Eye Doctor
275	Bernardston	Park	Ski Chalet	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm
290	Leominster	Business Service	Park	Dance Studio	Women's Store	Farmers Market	Event Space	Exhibit	Eye Doctor	Fair	Farm
294	Hinsdale	Golf Course	Women's Store	Ethiopian Restaurant	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market	Fast Food Restaurant
327	Palmer	Park	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Fast Food Restaurant

Figure 5. Cluster 2

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
213	Hubbardston	Construction & Landscaping	Women's Store	Fast Food Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market
300	Dudley	Construction & Landscaping	Women's Store	Fast Food Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market
343	Southbridge	Construction & Landscaping	Women's Store	Fast Food Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market

Figure 6. Cluster 3

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	Sherborn	Food	Convenience Store	Sandwich Place	Pizza Place	Donut Shop	Farm	Event Service	Event Space	Exhibit	Eye Doctor
4	Sudbury	Mediterranean Restaurant	Donut Shop	Pizza Place	Italian Restaurant	Fast Food Restaurant	Event Space	Exhibit	Eye Doctor	Fair	Farm
17	Southborough	Convenience Store	Lawyer	Pizza Place	Pharmacy	Print Shop	Breakfast Spot	Golf Course	Baseball Field	Spa	Historic Site
26	Topsfield	Lawyer	Pizza Place	Hardware Store	Deli / Bodega	Bank	Convenience Store	Fishing Store	Farm	Ethiopian Restaurant	Event Service
31	Acton	Historic Site	Playground	Pizza Place	Trail	Farmers Market	Ethiopian Restaurant	Event Service	Event Space	Exhibit	Eye Doctor
58	Milton	Convenience Store	Home Service	Construction & Landscaping	Park	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor
68	Wrentham	Gastropub	Convenience Store	Breakfast Spot	Donut Shop	Bank	Fast Food Restaurant	Exhibit	Eye Doctor	Fair	Farm
74	Medway	Playground	Convenience Store	Pizza Place	Nightclub	Business Service	Women's Store	Farm	Event Space	Exhibit	Eye Doctor
75	Georgetown	Pizza Place	Sandwich Place	Baseball Field	Pharmacy	Grocery Store	American Restaurant	Bank	Bagel Shop	Donut Shop	Fair

83	Upton	Food	Pizza Place	Business Service	Martial Arts School	Restaurant	Fair	Event Service	Event Space	Exhibit	Eye Doctor
94	Hanover	Pizza Place	Grocery Store	Coffee Shop	Thai Restaurant	Donut Shop	Gift Shop	Gym / Fitness Center	Liquor Store	Farmers Market	Fast Food Restaurant
107	Tewksbury	Bank	Convenience Store	Italian Restaurant	Liquor Store	Park	Greek Restaurant	Spa	Fish & Chips Shop	Fish Market	Film Studio
114	Plainville	Diner	Video Store	Electronics Store	Park	Bookstore	Liquor Store	Breakfast Spot	Cosmetics Shop	Pizza Place	Convenience Store
116	Middleton	Grocery Store	Pizza Place	Donut Shop	Event Space	Business Service	Farmers Market	Construction & Landscaping	Post Office	Convenience Store	Liquor Store
117	Hopedale	Convenience Store	Construction & Landscaping	Home Service	Women's Store	Fast Food Restaurant	Event Space	Exhibit	Eye Doctor	Fair	Farm
118	Wilbraham	Pizza Place	Restaurant	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm
131	Beverly	Pizza Place	Donut Shop	Convenience Store	Automotive Shop	Pharmacy	Print Shop	Sandwich Place	Breakfast Spot	Furniture / Home Store	Gym
134	Rochester	Baseball Field	Convenience Store	Pizza Place	American Restaurant	Women's Store	Fast Food Restaurant	Event Space	Exhibit	Eye Doctor	Fair
139	Pembroke	Convenience Store	Donut Shop	Pharmacy	Bank	Women's Store	Farmers Market	Event Space	Exhibit	Eye Doctor	Fair
146	Brimfield	Business Service	Farmers Market	Convenience Store	Flea Market	Pizza Place	Women's Store	Event Space	Exhibit	Eye Doctor	Fair
148	Eastham	Liquor Store	Convenience Store	Antique Shop	Park	Grocery Store	Farm	Ethiopian Restaurant	Event Service	Event Space	Exhibit
173	East Bridgewater	Pizza Place	Convenience Store	Breakfast Spot	Mexican Restaurant	Grocery Store	Dessert Shop	Coffee Shop	Sushi Restaurant	Park	Women's Store
184	Dighton	Post Office	Pizza Place	Harbor / Marina	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair
187	Sheffield	Convenience Store	Bookstore	Bed & Breakfast	Women's Store	Film Studio	Exhibit	Eye Doctor	Fair	Farm	Farmers Market
195	Abington	Convenience Store	American Restaurant	Coffee Shop	Chinese Restaurant	Sandwich Place	Brewery	Liquor Store	Pizza Place	Video Store	Home Service
205	Saugus	Historic Site	Diner	Convenience Store	Pizza Place	Donut Shop	Italian Restaurant	Music Venue	Event Space	Exhibit	Eye Doctor
215	Frammingham	Coffee Shop	Pizza Place	Fast Food Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market
216	Holbrook	Pizza Place	Breakfast Spot	Cosmetics Shop	Donut Shop	Pharmacy	Liquor Store	Playground	Fishing Store	Fishing Spot	Fish Market
217	Hanson	Ice Cream Shop	Convenience Store	New American Restaurant	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm
224	Ayer	Pizza Place	Italian Restaurant	Liquor Store	Sandwich Place	Donut Shop	Pub	Convenience Store	Trail	Chinese Restaurant	Dive Bar
227	Avon	Convenience Store	American Restaurant	Bar	Pizza Place	Post Office	Liquor Store	Lake	Farm	Event Service	Event Space
235	Dracut	Golf Course	Pizza Place	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm
242	Freetown	Convenience Store	Women's Store	Fast Food Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market
247	Petersham	Scenic Lookout	Post Office	Farmers Market	Convenience Store	Cricket Ground	Creperie	Event Service	Event Space	Exhibit	Eye Doctor
251	Barre	Pizza Place	Breakfast Spot	Flower Shop	Bistro	Italian Restaurant	Restaurant	Museum	Women's Store	Farmers Market	Exhibit
252	Middleborough	Sandwich Place	Convenience Store	Pizza Place	American Restaurant	Donut Shop	Gift Shop	Farmers Market	Event Space	Exhibit	Eye Doctor
268	West Boylston	Bakery	Playground	Pizza Place	Antique Shop	Fast Food Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair
293	Spencer	Convenience Store	Restaurant	Pizza Place	Cricket Ground	Pharmacy	Video Store	Mexican Restaurant	Electronics Store	Chinese Restaurant	Supermarket
321	Taunton	Convenience Store	Women's Store	Fast Food Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm	Farmers Market
342	Brockton	Pizza Place	Lounge	Chinese Restaurant	Pub	Automotive Shop	Asian Restaurant	Women's Store	Farm	Event Space	Exhibit

Figure 7. Cluster 4

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Weston	Bagel Shop	Coffee Shop	Grocery Store	Market	Diner	Peking Duck Restaurant	Pharmacy	Pizza Place	Fish Market	Fish & Chips Shop
1	Dover	Grocery Store	Post Office	Donut Shop	Deli / Bodega	Farm	Ethiopian Restaurant	Event Service	Event Space	Exhibit	Eye Doctor
2	Carlisle	Playground	Sandwich Place	Women's Store	Farmers Market	Event Service	Event Space	Exhibit	Eye Doctor	Fair	Farm
5	Wellesley	Pizza Place	Clothing Store	Coffee Shop	Arts & Crafts Store	Thai Restaurant	Mediterranean Restaurant	Italian Restaurant	Spa	Restaurant	Fast Food Restaurant
6	Winchester	New American Restaurant	Italian Restaurant	Coffee Shop	Plaza	Jewelry Store	Soccer Field	Farmers Market	Fair	Spa	Café
...
347	New Bedford	Gym	Breakfast Spot	Fast Food Restaurant	Burger Joint	Botanical Garden	Flower Shop	Flea Market	Fishing Store	Fishing Spot	Fish Market
348	Chelsea	Plaza	Dessert Shop	Mexican Restaurant	Fast Food Restaurant	Electronics Store	Bank	Donut Shop	Thai Restaurant	Grocery Store	Discount Store
349	Holyoke	Pizza Place	Spanish Restaurant	Bar	Bakery	Science Museum	Storage Facility	Trail	Art Gallery	Fast Food Restaurant	American Restaurant
351	Springfield	Hotel	Hockey Arena	Shipping Store	American Restaurant	Donut Shop	Café	Cocktail Bar	BBQ Joint	Casino	Strip Club
352	Lawrence	Caribbean Restaurant	Donut Shop	Pub	Bank	Café	Fried Chicken Joint	Mexican Restaurant	Fast Food Restaurant	Pharmacy	Convenience Store

192 rows x 11 columns

Figure 8. Cluster 5

A map of the resulting clusters is shown in Figure 9.

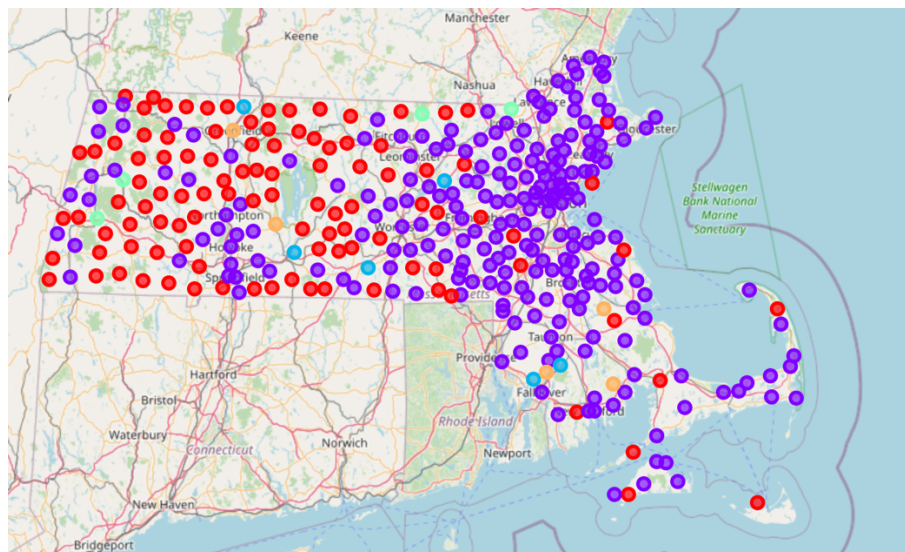


Figure 9. Map of clusters.

5. Discussion

The towns within each cluster have similar frequented venues. Cluster 1 has the most towns in their cluster which is denoted by the purple markers in Figure 9. It can be easily seen from the map that Cluster 1, which includes Boston and Brookline, is the largest cluster and highly concentrated in the east. Cluster 2, denoted by the red markers, is mainly in the west. Cluster 3, 4, and 5 are more sporadically scattered in the southeast or west.

6. Conclusion

For Massachusetts residents who are looking to move, it may be helpful to see which town is similar in venues to their current location, especially if they are looking to go somewhere similar. The results can also be used to find towns that are not as similar to their current location as well. With other factors that may associated with choosing a town to live in, it may also be beneficial for another project to assess cost of living, job prospects, safety, and more for each town.