

Capstone Project: Best College Towns

Introduction

As a web-based service that assists students in selecting schools, we accept requests from prospective high school students to compare colleges and universities along a variety of factors.

We receive a request from a students with the schools they are interested in. The ten schools the student is considering are: Northwestern University (Evanston, Illinois); University of Michigan (Ann Arbor, Michigan); Stanford (Palo Alto, California); Auburn (Auburn, Alabama); Florida State (Tallahassee, Florida); Wellesley College (Wellesley, Massachusetts); Tufts (Medford, Massachusetts); Columbia University (New York City, New York); Duke University (Durham, North Carolina); Washington University (St. Louis, Missouri).

Data

We used the Foursquare API and the school address to pull the top venues around with a 1 kilometer radius of the school, easily within walking distance. The Foursquare data is regularly updated, high quality data set for understanding the types of venues that exist whether it be coffee shops, bars, or parks. To use this data, we will take the school addresses and use the geopy.geocoders package to convert them into latitudes and longitudes for each school. We will then retrieve the nearby venues from the API.

	School	Address
0	Northwestern University	633 Clark St, Evanston, IL 60208
1	University of Michigan	500 S State St, Ann Arbor, MI 48109
2	Stanford University	450 Serra Mall, Stanford, CA 94305
3	Auburn University	Quad Center, Auburn, AL 36849
4	Florida State	600 W College Ave, Tallahassee, FL 32306
5	Wellesley College	106 Central St, Wellesley, MA 02481
6	Tufts University	419 Boston Ave, Medford, MA 02155
7	Columbia University	1130 Amsterdam Avenue New York, NY 10027
8	Duke University	2138 Campus Drive Durham, NC 27708
9	Washington University	1 Brookings Dr, St. Louis, MO 63130

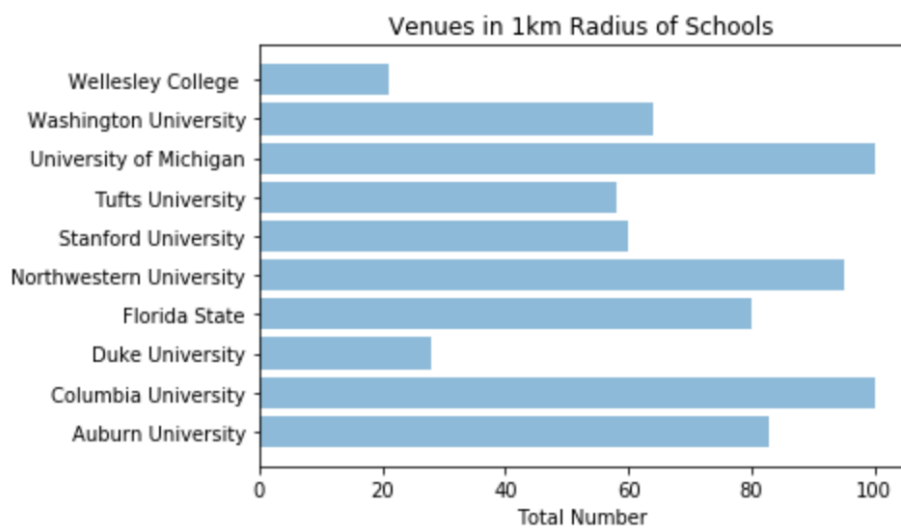


Methodology

As high school students prepare to apply to colleges, they want to consider their college experience. What type of college town are their target schools based in? They can't afford to visit all. Therefore, we will leverage the Foursquare API and develop a K-means clustering to get some sense of what types of environments exist and how similar different schools are.

Results

Using Foursquares API, we generated a count of the top 10 venues within 1 kilometer of each of the schools. We then used K-means with a cluster size of 5 to try to distinguish between different types of campus life that might exist in and around the school



School	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
Northwestern University	Coffee Shop	Park	American Restaurant	Asian Restaurant	Thai Restaurant	Pizza Place	Sandwich Place	Bar	New American Restaurant	Mexican Restaurant
University of Michigan	Coffee Shop	Pizza Place	Bar	Clothing Store	Korean Restaurant	American Restaurant	Noodle House	Mexican Restaurant	Sandwich Place	Museum
Auburn University	Sandwich Place	American Restaurant	Pizza Place	Coffee Shop	Mediterranean Restaurant	Burger Joint	Fast Food Restaurant	College Cafeteria	Sushi Restaurant	Deli / Bodega
Florida State	Sandwich Place	Bar	Coffee Shop	Italian Restaurant	Restaurant	Grocery Store	Pizza Place	Donut Shop	Student Center	Liquor Store
Columbia University	Coffee Shop	Park	Café	American Restaurant	Italian Restaurant	Seafood Restaurant	Grocery Store	Bookstore	Playground	Chinese Restaurant
Washington University	American Restaurant	Candy Store	Thai Restaurant	Ice Cream Shop	Hookah Bar	Coffee Shop	Thrift / Vintage Store	Pizza Place	Music Venue	Golf Course

Cluster 1

College Wellesley	Yoga Studio	Cosmetics Shop	Spa	Cheese Shop	Lake	Fast Food Restaurant	Clothing Store	Coffee Shop	Italian Restaurant	Bus Stop
School	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

Cluster 2

School	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
Wellesley College	Yoga Studio	Cosmetics Shop	Spa	Cheese Shop	Lake	Fast Food Restaurant	Clothing Store	Coffee Shop	Italian Restaurant	Bus Stop

Cluster 3

School	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
Duke University	College Cafeteria	Café	Fast Food Restaurant	Food Court	Pharmacy	College Auditorium	Coffee Shop	Museum	Chinese Restaurant	Eye Doctor

Cluster 4

School	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
Duke University	College Cafeteria	Café	Fast Food Restaurant	Food Court	Pharmacy	College Auditorium	Coffee Shop	Museum	Chinese Restaurant	Eye Doctor

Cluster 5

Discussions

Several colleges form one cluster suggesting they have a similar make-up of venues in the area. These include Northwestern University, University of Michigan, Auburn University, Florida State, Columbia University, and Washington University. They are primarily schools in urban areas:

Four colleges do not fit into that cluster and, instead, fit into their own unique clusters. They present a more unique "college life": Wellesley, Duke, Stanford, and Tufts

Wellesley and Duke have significantly less venues in their area than others suggesting they are in less urban areas.

Wellesley's top venues differ significantly from the other clusters with yoga studio, cosmetic shops, and spas topping the list. Duke University's top venues are mostly places related to school including the cafeteria, auditorium, and food court. While coffee shops rank high for Stanford, there are not many food establishments in the list of nearby venues. The last cluster contains Tufts University. Cafes are its top ranking venue along with a diverse set of restaurants that differ greatly from those in the first cluster.

Conclusion

Additional research should be dedicated to ensuring the quality of Foursquare data in and around college campuses as well as the categorization of venues. Many of the schools had several types of restaurants in the area of interest. It would be useful to look at other venues outside of restaurants. Additionally, Foursquare users can rate venues. It may be useful to look at the most popular venues in the area to get a better sense of the quality of "college life."