

# AAMC Medical Schools for Coffee-Drinking, Pizza-Eating, and Park-Loving Medical Students

Bill Cash  
April 11, 2021

## 1. Introduction

### 1.1 Background

Every year pre-med students at colleges and universities in the United States and Canada are researching and applying to medical schools. There is a publication every year by the U.S. News & World Report that ranks the best medical schools and can be used to be more specific regarding medical programs and specialties, and it also includes information such as tuition amounts and enrollment counts. The pre-med students know from their undergraduate careers that it takes a lot of both energy and stress management to be able to succeed as medical students. For energy, many medical students have learned to use coffee and pizza to fuel themselves. In addition to pizza there are plenty of other types of foods that different students may enjoy. For stress management, many students will take their coffee or their food outside to a park when weather is nice. This project is aimed at identifying which medical schools in the United States and Canada are near a good mix of places to get coffee, pizza places (and other types of restaurants), and parks at which to recreate and enjoy the weather when it is nice.

### 1.2 Interest

Many pre-med students who are considering many different medical schools would be interested in knowing what sorts of venues are near the medical schools that offer the following:

- Coffee
- Pizza
- Other restaurants
- Parks

## 2. Data

### 2.1 Data Sources

For this project, a list of medical schools for the United States and Canada will be needed. The list will need to provide the name and the address of the medical school. The Association of American Medical Colleges (AAMC) has the data on a webpage at:

<https://members.aamc.org/eweb/DynamicPage.aspx?site=AAMC&webcode=AAMCOrgSearchResult&orgtype=Medical%20School>

The Python library 'geocoder' will be used to get the latitude and longitude data by using the address data that will be obtained from the AAMC webpage.

The Foursquare developers Places API will be used to identify venues for coffee, pizza (and other food), and parks that are near the medical schools in the United States and Canada. The Foursquare developers Places API can be found at this webpage:

<https://foursquare.com/developers/apps>

### 3. Methodology

In this project I will use **Pandas DataFrames** to show the top 25 AAMC member medical schools with the most coffee shops, pizza places, and parks nearby. I will also use **Folium maps** to show all the AAMC member medical schools with color coding that indicates which ones have more or less coffee shops, pizza places, and parks nearby. For the other restaurants I will use **k-means clustering** to group the AAMC member medical schools together by what they have in common, and I will use the **elbow method** using distortion and inertia to identify the best value for k. I will display the clusters as Pandas DataFrames and simply look at them to describe the clusters.

### 4. Analysis

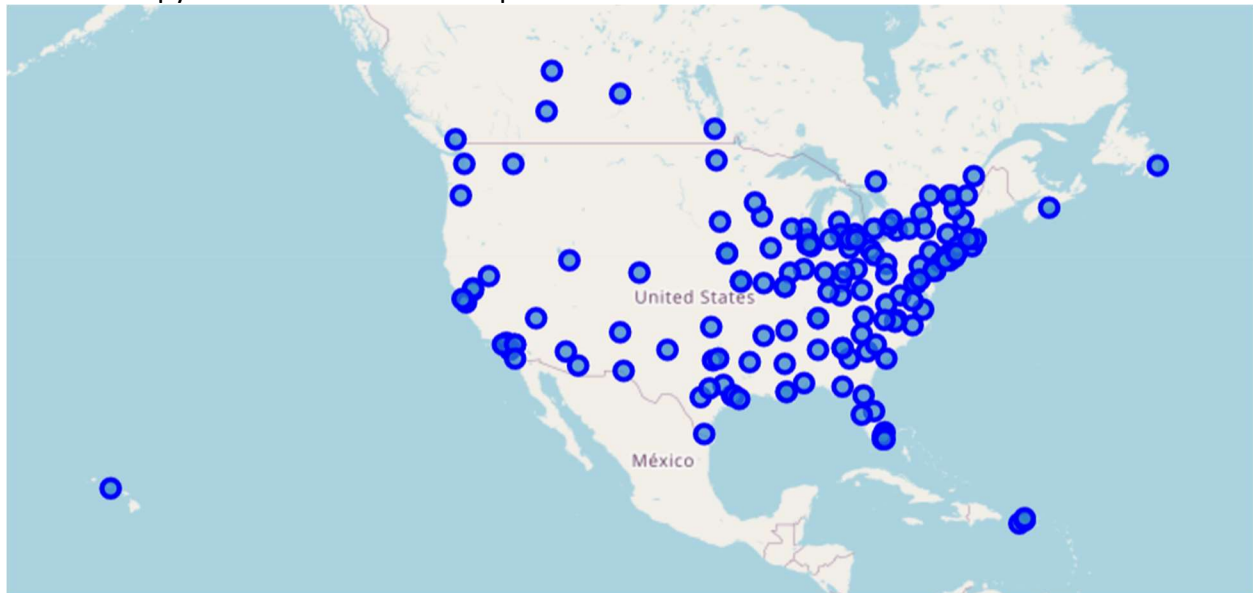
A total of 172 member medical schools were extracted from the html code on the AAMC member medical school webpage:

	Medical_School_Name	City_State_Postal_Code	Country
0	Albany Medical College	Albany, NY 12208-3412	UNITED STATES
1	Albert Einstein College of Medicine	Bronx, NY 10461-1900	UNITED STATES
2	Baylor College of Medicine	Houston, TX 77030-3411	UNITED STATES
3	Boston University School of Medicine	Boston, MA 02118-2307	UNITED STATES
4	Brody School of Medicine at East Carolina Univ...	Greenville, NC 27834-4300	UNITED STATES
...	...	...	...
167	Weill Cornell Medicine	New York, NY 10065-4805	UNITED STATES
168	West Virginia University School of Medicine	Morgantown, WV 26506-9100	UNITED STATES
169	Western Michigan University Homer Stryker M.D....	Kalamazoo, MI 49008-1282	UNITED STATES
170	Wright State University Boonshoft School of Me...	Dayton, OH 45401-0927	UNITED STATES
171	Yale School of Medicine	New Haven, CT 06510-3206	UNITED STATES

172 rows × 3 columns

The 172 medical schools are as far north as Edmonton, Alberta as far west as Honolulu, Hawaii as far south as Ponce, Puerto Rico, and as far east as St. John's Newfoundland.

Here is a copy of a non-interactive map of the AAMC member medical schools:



*See the Jupyter notebook on nbviewer for an interactive Folium map.*

A function was defined that separately looked for coffee shops, pizza places, parks, and all food venue types. It was expected that this function would return some duplicates because the category ID that was used to identify all other fee venue types also identifies coffee shops and pizza places. It is assumed that the Foursquare API calls do not always produce the same results, and I am convinced that there was probably a better way that I could have handled errors. Ultimately many iterations of the Places API request were required to get to radius values that did not produce errors, and the radius values I used were 500 meters for the food venues and 200 meters for the parks. I know that early on when I was working on that function that I was able to run it with larger radius values, but for some reason the function started returning errors and I the only way I could get the function to work without returning errors was to reduce the radius values. This resulted in fewer venues to analyze. The function did return 4,299 select nearby venues for the 172 AAMC member medical schools. After removing duplicates there were 3,786 unique select nearby venues, and there were 132 unique venue categories.

The venue categories were then grouped by coffee shops and cafes, pizza places, and parks. Those groupings were counted and ranked most to least. Pandas DataFrames were created to show the top 25 AAMC member medical schools for each of the three venue category types, and Folium maps were created with circle markers that have darker shading for the AAMC member medical schools with more of the selected venue type nearby. The following 3 pages show the top 25 tables for coffee shops, pizza places, and parks. The interactive Folium maps can be seen in the Jupyter notebook on nbviewer.

Top 25 AAMC member medical schools with the most coffee shops:

Medical_School_Name	Coffee_Shop_Count
University of Washington School of Medicine	23
Sidney Kimmel Medical College at Thomas Jefferson University	22
Tufts University School of Medicine	22
University of California, Los Angeles David Geffen School of Medicine	22
Northwestern University The Feinberg School of Medicine	21
Harvard Medical School	21
University of British Columbia Faculty of Medicine	19
Weill Cornell Medicine	18
Washington State University Elson S. Floyd College of Medicine	18
University of Michigan Medical School	17
Emory University School of Medicine	16
Columbia University Vagelos College of Physicians and Surgeons	16
George Washington University School of Medicine and Health Sciences	16
Carle Illinois College of Medicine	15
University of Toronto Temerty Faculty of Medicine	15
University of Chicago Division of the Biological Sciences The Pritzker School of Medicine	14
Kaiser Permanente Bernard J. Tyson School of Medicine	13
University of Pittsburgh School of Medicine	13
University of Hawaii, John A. Burns School of Medicine	13
Universite de Montreal Faculty of Medicine	13
Baylor College of Medicine	13
University of Texas at Austin Dell Medical School	12
Perelman School of Medicine at the University of Pennsylvania	11
McGill University Faculty of Medicine and Health Sciences	11
Mayo Clinic Alix School of Medicine	11

Top 25 AAMC member medical schools with the most pizza places:

<b>Medical_School_Name</b>	<b>Pizza_Place_Count</b>
<b>New York University Grossman School of Medicine</b>	21
<b>Sidney Kimmel Medical College at Thomas Jefferson University</b>	18
<b>Weill Cornell Medicine</b>	12
<b>Columbia University Vagelos College of Physicians and Surgeons</b>	11
<b>CUNY School of Medicine</b>	10
<b>University of California, Los Angeles David Geffen School of Medicine</b>	9
<b>Icahn School of Medicine at Mount Sinai</b>	8
<b>Howard University College of Medicine</b>	7
<b>State University of New York Upstate Medical University</b>	7
<b>University of New Mexico School of Medicine</b>	7
<b>Northwestern University The Feinberg School of Medicine</b>	7
<b>University of Virginia School of Medicine</b>	7
<b>University of Texas at Austin Dell Medical School</b>	6
<b>University of Michigan Medical School</b>	6
<b>TCU and UNTHSC School of Medicine</b>	6
<b>The Warren Alpert Medical School of Brown University</b>	6
<b>Lewis Katz School of Medicine at Temple University</b>	6
<b>Kaiser Permanente Bernard J. Tyson School of Medicine</b>	6
<b>George Washington University School of Medicine and Health Sciences</b>	6
<b>Yale School of Medicine</b>	5
<b>University of Pittsburgh School of Medicine</b>	5
<b>New York University Long Island School of Medicine</b>	5
<b>University of Maryland School of Medicine</b>	5
<b>University of Texas Rio Grande Valley School of Medicine</b>	5
<b>Washington State University Elson S. Floyd College of Medicine</b>	5

Top 25 AAMC member medical schools with the most parks:

<b>Park_Count</b>	<b>Medical_School_Name</b>
4	Columbia University Vagelos College of Physicians and Surgeons
2	George Washington University School of Medicine and Health Sciences
2	Northwestern University The Feinberg School of Medicine
2	Ohio State University College of Medicine
2	State University of New York Downstate Medical Center College of Medicine
2	Tufts University School of Medicine
1	CUNY School of Medicine
1	University of Hawaii, John A. Burns School of Medicine
1	Wake Forest School of Medicine of Wake Forest Baptist Medical Center
1	University of New Mexico School of Medicine
1	University of Minnesota Medical School
1	University of Michigan Medical School
1	University of Massachusetts Medical School
1	The Warren Alpert Medical School of Brown University
1	University of California, Davis, School of Medicine
1	Universite de Montreal Faculty of Medicine
1	Sidney Kimmel Medical College at Thomas Jefferson University
1	Rutgers, Robert Wood Johnson Medical School
1	Icahn School of Medicine at Mount Sinai
1	Cooper Medical School of Rowan University
1	Washington University in St. Louis School of Medicine

After the AAMC member medical schools were ranked by most coffee shops, most pizza places, and most parks the AAMC member medical schools were analyzed for the rest of the venue category types. The other restaurant category types ranged from Afghan restaurants to themed restaurants and there were 2,407 venues for 167 of the 172 AAMC member medical schools. One hot encoding was used to translate the categorical venue type data into numerical values that could be used in calculations. Means were calculated for each AAMC member medical school for those categorical venue type values and then for each AAMC member medical school the mean values for the venue category types were used to rank the venue category types and the top ten venue category types for each AAMC medical school were used for grouping the AAMC member medical schools using k-means clustering. The elbow method was used to determine the best value for k for the k-means clustering. The distortion and inertia values

used for the elbow method did not reveal an obvious value for  $k$ . I ended up choosing 14 for the value for  $k$  because after 14 the distortion and inertia values seem to level off. I probably could have made the project easier by choosing a smaller value for  $k$ , but I chose 14 and I am sticking with it. After grouping the AAMC member medical schools for the other restaurant types into 14 clusters there were many clusters that identify AAMC member medical schools with food trucks, deli/bodegas, bakeries, sandwich places, fast food restaurants, food courts, American restaurants, Indian restaurants, wings joints, snack places, Czech restaurants, cafeterias, dim sum restaurants, diners, Eastern European restaurants, falafel restaurants, Ethiopian restaurants, empanada restaurants, and dumpling restaurants.

## 5. Results and Discussion

The analysis shows that the top five AAMC member medical schools to go to for coffee lovers are:

1. University of Washington School of Medicine
2. Sidney Kimmel Medical College at Thomas Jefferson University
3. Tufts University School of Medicine
4. University of California, Los Angeles David Geffen School of Medicine
5. Northwestern University The Feinberg School of Medicine

The analysis shows that the top five AAMC member medical schools to go to for pizza lovers are:

1. New York University Grossman School of Medicine
2. Sidney Kimmel Medical College at Thomas Jefferson University
3. Weill Cornell Medicine
4. Columbia University Vagelos College of Physicians and Surgeons
5. CUNY School of Medicine

The analysis shows that the top five AAMC member medical schools to go to for park lovers are:

1. Columbia University Vagelos College of Physicians and Surgeons
2. George Washington University School of Medicine and Health Sciences
3. Northwestern University The Feinberg School of Medicine
4. Ohio State University College of Medicine
5. State University of New York Downstate Medical Center College of Medicine

There are also many clusters that identify AAMC member medical schools with food trucks, deli/bodegas, bakeries, sandwich places, fast food restaurants, food courts, American restaurants, Indian restaurants, wings joints, snack places, Czech restaurants, cafeterias, dim sum restaurants, diners, Eastern European restaurants, falafel restaurants, Ethiopian restaurants, empanada restaurants, and dumpling restaurants.

## 6. Conclusion

The purpose of this project was to show AAMC member medical schools with rankings for most coffee venues, most pizza venues, most parks, and to identify other types of restaurants near AAMC member medical schools so that pre-med students who are looking at medical schools in the United States and Canada could see which medical schools have the most venues for some popular types of food and beverage for students.

While the maps did not show up as well as I would have hoped, the interactive features that the Folium maps have will allow users to zoom in and click on the markers to pull up more information about the medical schools that rank highly in geographic areas where they may be considering going to medical school. The Pandas DataFrames showing the top 25 AAMC member medical schools for the top-ranked schools for coffee, pizza, and parks will make it easy for pre-med students to identify medical schools with lots of coffee shops, pizza places, and parks.

The clusters did not work out like I had hoped they would. There were a few clusters of one medical school, and a couple of clusters were exceptionally large groups (one of which did not seem to have many venues in common). However, most of the clusters identified a fair number of schools, and they did have some other restaurant categories in common. The clusters will help pre-med students identify medical schools with nearby food trucks, deli/bodegas, bakeries, sandwich places, fast food restaurants, food courts, American restaurants, Indian restaurants, wings joints, snack places, Czech restaurants, cafeterias, dim sum restaurants, diners, Eastern European restaurants, falafel restaurants, Ethiopian restaurants, empanada restaurants, and dumpling restaurants. I thought that the elbow method would reveal a more obvious value for  $k$ . I suppose I could have simply chosen a smaller value for  $k$  and forced some of the clusters to be larger, but I do not know that that would have resulted in better clusters.



## Project Notebook Links:

The Coursera IBM Data Science Capstone project notebook can be found on GitHub here:

[https://github.com/Cash17/Coursera\\_Capstone/blob/master/Coursera IBM Data Science Capstone Working NB.ipynb](https://github.com/Cash17/Coursera_Capstone/blob/master/Coursera%20IBM%20Data%20Science%20Capstone%20Working%20NB.ipynb)

*Note: this version does not allow the interactive Folium maps. See the nbviewer version for the interactive Folium maps.*

The Coursera IBM Data Science Capstone project notebook can be found on nbviewer here:

[https://nbviewer.jupyter.org/github/Cash17/Coursera\\_Capstone/blob/master/Coursera IBM Data Science Capstone Working NB.ipynb](https://nbviewer.jupyter.org/github/Cash17/Coursera_Capstone/blob/master/Coursera%20IBM%20Data%20Science%20Capstone%20Working%20NB.ipynb)