Where to open a Chinese restaurant

1. **Introduction**

   a. **Background**

   Nowadays, a lot of Chinese immigrants come to the United States to seek better education for their children. Some of them do not have advanced skills to find a job in the United States, so opening a local Chinese restaurant is a good option for them to make a living while their children go to school.

   On the other hand, Chinese cuisine has been an important part of Chinese culture, and Chinese food has been so popular among Chinese and non-Chinese people that a good Chinese restaurant would normally bring better dinning experience for the local people as well as more job opportunities.

   In this project we would like to develop a way to help new Chinese immigrants who want to open Chinese restaurants to find out the best city and place to live in and start their business.

   b. **Problem**

   In order to find the best city and place to live in and open a Chinese restaurant, we would normally choose a city with a large enough Chinese population. Although Chinese food are also popular among non-Chinese people, local Chinese people would still most likely be the main consumers in Chinese restaurants.

   We could definitely choose metropolitan cities such as New York City or Los Angeles for simplicity (actually many people do that), but the disadvantages with big cities like these are that they are too crowded (thus not the best choice for living), and there are already many Chinese restaurants in them so the competition would be severe (thus not the best choice for business).

   The best way to choose the city might be to combine a large enough Chinese population with a small enough number of available Chinese restaurants, plus a consideration of livability. After we choose a specific city, we could further check its neighborhoods and decide a more detailed location to open a restaurant.

   c. **Interest**

   This project would definitely be interesting to new Chinese immigrants who want to open Chinese restaurants to make a living, and the immigration lawyers or consultants who provide advises for these new immigrants. It has been shown that

This project would also potentially benefit immigrants from other nationalities in a sense that they could easily use a similar approach to find out the best place to start their own featured food business.

2. Data Description

a. Data sources

We could find Chinese-American population data at different cities in this Wikipedia page (https://en.wikipedia.org/wiki/List_of_U.S._cities_with_significant_Chinese-American_populations) It has listed all the large-sized, medium-sized and small-sized cities with significant Chinese-American populations. I only used the large-sized and medium-sized cities, since for the small-sized cities (cities with a population fewer than 100,000) although the percentage of Chinese-American population might be relatively high, the absolute number would not be big enough.

After getting a list of all the large-sized and medium-sized cities with significant Chinese-American populations, I used geolocator to get the geographical coordinates of each city, and then used Foursquare to get the total number of Chinese restaurants in each city. Besides the number of available Chinese restaurants, I also considered livability in each city, and I got the AreaVibes Livability Score from https://www.areavibes.com .

After I decide that Cary in North Carolina is the best city to open a Chinese restaurant, I went on to find Cary's neighborhoods information on its city-data webpage http://www.city-data.com/city/Cary-North-Carolina.html and did more analysis about the neighborhoods using Foursquare data, one hot encoding and k-means clustering.

b. Data cleaning

As mentioned in the "Data Source" section, I used the population tables for large-sized and medium-sized cities in the Wikipedia page and combined them into one table. There was one city "Irvine" duplicated in both tables, so I removed it before combining the two tables. I then sort the cities according to the absolute numbers of Chinese-American population. The table (with geographical coordinates) looks like this:

	City	State	Chinese-Americans	Percentage	latitude	longitude
0	New York	New York	562205	6.6	40.712728	-74.006015
1	San Francisco	California	180372	21.2	37.779281	-122.419236
2	Los Angeles	California	77073	2.0	34.053691	-118.242767
3	San Jose	California	75582	7.5	37.336191	-121.890583
4	Chicago	Illinois	52917	1.9	41.875562	-87.624421

When dealing with the neighborhood information of Cary, I could only find a pure text containing all the neighborhood names in Cary on its city-data webpage http://www.city-data.com/city/Cary-North-Carolina.html. I replaced each 'neighborhood' word with symbols '","' (double quotes and comma), so that the pure text could be transformed to a list and then converted into a Pandas Dataframe, which looks like this:

	Neighborhood
0	Adams Park
1	Allenbrook
2	Ambience Place
3	Andover
4	Applecross Townhomes

c. **Feature selection**

After getting the final city data table with geographical coordinates, I used Foursquare to get the total number of Chinese restaurants in each city and sort the cities according to the available restaurants number ascendingly:

	City	State	Chinese-Americans	Percentage	latitude	longitude	number
16	El Monte	California	20190	17.5	36.701463	-118.755997	0
40	Enterprise	Nevada	5879	4.5	36.002834	-115.201299	12
90	Simi Valley	California	1606	1.3	34.269447	-118.781482	14
98	Pearland	Texas	1112	1.1	29.563976	-95.286430	18
58	Thousand Oaks	California	3889	3.0	34.171427	-118.910588	22

I chose the cities that have relatively small number of Chinese restaurants ('number' < 50) but still a large Chinese-American population (index < 50, that is, the first fifty cities that have the largest number of Chinese-American population), then added the AreaVibes Livability Score for each city, and the final table looks like this:

	City	State	Chinese-Americans	Percentage	latitude	longitude	number	livability score
16	El Monte	California	20190	17.5	36.701463	-118.755997	0	63
40	Enterprise	Nevada	5879	4.5	36.002834	-115.201299	12	75
25	Elk Grove	California	10758	6.6	38.408799	-121.371618	27	75
44	Cary	North Carolina	5283	3.4	35.788305	-78.781196	32	86
14	Fairfax	Virginia	28806	2.5	38.846224	-77.306373	34	76
42	Richmond	California	5523	5.1	37.935758	-122.347749	38	57
31	Ann Arbor	Michigan	7998	6.8	42.268157	-83.731229	40	79
37	Naperville	Illinois	6584	4.5	41.772870	-88.147928	42	75

Since "Cary, North Carolina" has the highest livability score among these cities, I chose Cary as our target city and continued to explore its neighborhoods.

3. Methodology

a. Exploratory data analysis

We now start to explore the neighborhoods of Cary, North Carolina. I have already converted Cary's neighborhood names into a Pandas Dataframe:
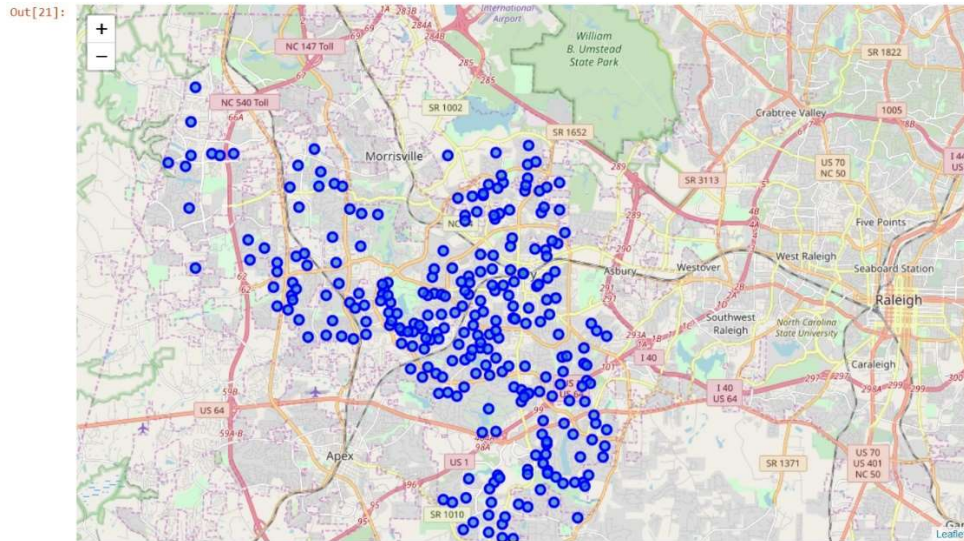
Out[17]:

	Neighborhood
0	Adams Park
1	Allenbrook
2	Ambience Place
3	Andover
4	Applecross Townhomes

I went on to get the latitude and longitude values of each neighborhood, removed any NaN values and get a table containing 275 neighborhoods:

Out[19]:

	Neighborhood	latitude	longitude
0	Allenbrook	35.775758	-78.771380
1	Ambience Place	35.739008	-78.773885
2	Andover	35.775421	-78.817316
3	Applecross Townhomes	35.757609	-78.814131
4	Appleton Acres	35.812176	-78.796356
5	Ardmore at Reedy Creek	35.800821	-78.765561
6	Arlington Park at Amberly	35.824618	-78.906770
7	Ashcroft	35.788305	-78.781196
8	Ashley Woods	35.784135	-78.875776
9	Balmoral	35.766613	-78.805773
10	Battery at High House	35.782228	-78.814795

Here is a map of Cary with all the 275 neighborhoods labeled:

We can now start to use data from Foursquare to explore these neighborhoods. I defined a function to get the information of up to 100 venues within 500 meters of a neighborhood. I then apply this function to each neighborhood of Cary and get a table containing all the top venues in Cary:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Allenbrook	35.775758	-78.77138	DQ Grill & Chill	35.776055	-78.768110	Fast Food Restaurant
1	Allenbrook	35.775758	-78.77138	Your Pie	35.775563	-78.767042	Pizza Place
2	Allenbrook	35.775758	-78.77138	Penn Station East Coast Subs	35.775265	-78.766391	Sandwich Place
3	Allenbrook	35.775758	-78.77138	Banana Leaf	35.776748	-78.767983	Chinese Restaurant
4	Allenbrook	35.775758	-78.77138	Primo Pizza NY Style	35.775195	-78.766781	Pizza Place

I can group these venues by neighborhood to get the total number of venues in each neighborhood, and I can also get the number of unique categories in all the venues (there are 205 unique categories).

In order to analyze each neighborhood and decide a more detailed location to open a Chinese restaurant, I did one hot encoding, grouped the rows by neighborhood and took the mean of the occurrence frequency of venues in each category. I then get the top 10 most common venues for each neighborhood and sort the venues in descending order:

	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	Allenbrook	Baseball Field	Pizza Place	Chinese Restaurant	Park	Wings Joint	Bike Shop	Cosmetics Shop	Discount Store	Fast Food Restaurant	Sports Bar
1	Ambience Place	Gourmet Shop	Yoga Studio	Café	Sushi Restaurant	Mediterranean Restaurant	Lake	Mexican Restaurant	Grocery Store	Automotive Shop	Gym / Fitness Center
2	Andover	Trail	Performing Arts Venue	Home Service	Pool	Harbor / Marina	Lake	Food Court	Food & Drink Shop	Food	Flower Shop
3	Applecross Townhomes	Pool	Park	Yoga Studio	Farm	Food Truck	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop
4	Appleton Acres	Pool	Fast Food Restaurant	Yoga Studio	Farm	Football Stadium	Food Truck	Food Stand	Food Court	Food & Drink Shop	Food

b. K-means clustering

Now we can run k-means to cluster the neighborhood into 10 clusters. Because we need to find out within the city of Cary, what might be the better/best location(s) to open the restaurant, we can group all the neighborhood into a few clusters (for example, 10 clusters) and see the character and distinction of each cluster and choose the one that would best fit the needs of opening a Chinese restaurant.
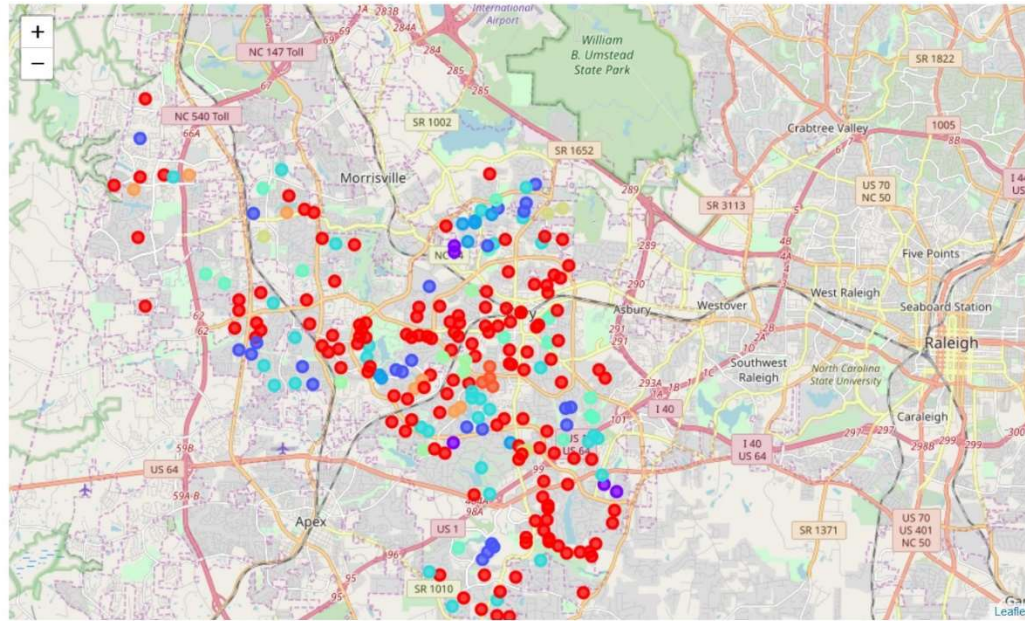
Here is a new dataframe that includes the cluster (labeled as 0, 1, 2, 3, 4, 5, 6, 7, 8, 9) as well as the top 10 venues for each neighborhood:

	Neighborhood	latitude	longitude	Cluster Labels	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	9 C V
0	Allenbrook	35.775758	-78.771380	0.0	Pizza Place	Chinese Restaurant	Baseball Field	Cosmetics Shop	Fast Food Restaurant	Sports Bar	Liquor Store	Sandwich Place	F
1	Ambience Place	35.739008	-78.773885	0.0	Gourmet Shop	Yoga Studio	Grocery Store	Café	Fast Food Restaurant	Burger Joint	Shopping Mall	Multiplex	M R
2	Andover	35.775421	-78.817316	2.0	Lake	Pool	Home Service	Performing Arts Venue	Trail	Harbor / Marina	Farm	Ethiopian Restaurant	E D
3	Applecross Townhomes	35.757609	-78.814131	4.0	Pool	Park	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	F R
4	Appleton Acres	35.812176	-78.796356	3.0	Concert Hall	Pool	Fast Food Restaurant	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	F S
5	Ardmore at Reedy Creek	35.800821	-78.765561	0.0	Indian Restaurant	Supermarket	Pizza Place	Discount Store	Chinese Restaurant	Music Store	Ethiopian Restaurant	Yoga Studio	F S
6	Arlington Park at Amberly	35.824618	-78.906770	0.0	Athletics & Sports	Pizza Place	Bubble Tea Shop	Pharmacy	Grocery Store	Eye Doctor	Food & Drink Shop	Food	F S
7	Ashcroft	35.788305	-78.781196	0.0	Food Truck	Brewery	Bakery	Pakistani Restaurant	Bridal Shop	Bus Station	Restaurant	Café	C R
8	Ashley Woods	35.784135	-78.875776	0.0	Chinese Restaurant	Mediterranean Restaurant	Ice Cream Shop	Jewelry Store	Pizza Place	Pharmacy	Donut Shop	Shoe Store	F R
										Outdoors			

Here is a visualization of the ten clusters (different color means different cluster):

4. Results

By considering Chinese-American populations, the number of available Chinese restaurants and livability in each city, we have determined that Cary, North Carolina would be the best city for new Chinese immigrants to live in and open a Chinese restaurant to make a living.

We then went on to analyze the neighborhoods in Cary, North Carolina and grouped all 275 neighborhoods into 10 clusters according to the top 10 most common venues for each neighborhood. Here are the tables of neighborhood information in each cluster, including the neighborhood names and the top 10 most common venues in each neighborhood:

Cluster 1:

Out[57]:

	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
0	Allenbrook	Pizza Place	Chinese Restaurant	Baseball Field	Cosmetics Shop	Fast Food Restaurant	Sports Bar	Liquor Store	Sandwich Place	Food
1	Ambience Place	Gourmet Shop	Yoga Studio	Grocery Store	Café	Fast Food Restaurant	Burger Joint	Shopping Mall	Multiplex	Mexican Restaurant
5	Ardmore at Reedy Creek	Indian Restaurant	Supermarket	Pizza Place	Discount Store	Chinese Restaurant	Music Store	Ethiopian Restaurant	Yoga Studio	Fabric Shop
6	Arlington Park at Amberly	Athletics & Sports	Pizza Place	Bubble Tea Shop	Pharmacy	Grocery Store	Eye Doctor	Food & Drink Shop	Food	Flower Sho
7	Ashcroft	Food Truck	Brewery	Bakery	Pakistani Restaurant	Bridal Shop	Bus Station	Restaurant	Café	Chinese Restaurant
8	Ashley Woods	Chinese Restaurant	Mediterranean Restaurant	Ice Cream Shop	Jewelry Store	Pizza Place	Pharmacy	Donut Shop	Shoe Store	Fast Food Restaurant
9	Balmoral	Pizza Place	Hardware Store	Smoke Shop	Outdoors & Recreation	Yoga Studio	Electronics Store	Flower Shop	Fishing Spot	Fast Food Restaurant
10	Battery at High House	Supermarket	Home Service	Gym / Fitness Center	Pharmacy	Pet Store	Chinese Restaurant	Salon / Barbershop	Business Service	Bus Station

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
54	Coventry Woods	Breakfast Spot	Yoga Studio	Eye Doctor	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
75	Franklin Heights	Insurance Office	Breakfast Spot	Yoga Studio	Eye Doctor	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant
90	Hazelwood	Construction & Landscaping	Insurance Office	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
93	Hermitage Park	Construction & Landscaping	Insurance Office	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
176	Royal Ridge	Construction & Landscaping	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm

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
2	Andover	Lake	Pool	Home Service	Performing Arts Venue	Trail	Harbor / Marina	Farm	Ethiopian Restaurant	Eye Doctor	Fabric Shop
13	Beechtree	Pool	Playground	Trail	Yoga Studio	Electronics Store	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
19	Bonnell Patio Homes	Construction & Landscaping	Trail	Pool	IT Services	Home Service	History Museum	Food	Flower Shop	Fishing Spot	Fast Food Restaurant
40	Chesney Glen	Pool	Health Food Store	Playground	Farm	Trail	Yoga Studio	Electronics Store	Food	Flower Shop	Fishing Spot
67	Falconers Ridge	Construction & Landscaping	IT Services	Arcade	Trail	Cafeteria	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot
79	Glenkirk	Arts & Entertainment	Trail	Automotive Shop	Yoga Studio	Eye Doctor	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot
88	Harrison Place	Pool	Trail	Yoga Studio	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market

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
4	Appleton Acres	Concert Hall	Pool	Fast Food Restaurant	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Farmers Market
43	Churchill Downs	Concert Hall	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
55	Crown Pointe	Concert Hall	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
71	Fishers Creek	Concert Hall	Pool	Fast Food Restaurant	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Farmers Market
143	Melbourne	Concert Hall	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
185	Silver Oaks	Concert Hall	Pool	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
187	Silvercreek	Concert Hall	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
195	Springbrook	Concert Hall	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market

Cluster 5:

	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	Applecross Townhomes	Pool	Park	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
14	Benedum Place	Auto Workshop	Pool	Bar	Yoga Studio	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant
15	Berkeley	Pool	Yoga Studio	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
18	Bond Pointe	Gym	Pool	Yoga Studio	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
24	Brookstone	Pool	Pizza Place	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
34	Cary Park	Pool	Supermarket	Pharmacy	Yoga Studio	Electronics Store	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
36	Catherwood Place	Auto Workshop	Pool	Bar	Yoga Studio	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant
38	Center Court Cluster Homes	Pool	IT Services	Trail	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market

Cluster 6:

	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
12	Beaver Pond	Park	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
33	Carramore	IT Services	Park	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
35	Cary Pines	Park	Soccer Field	Trail	Yoga Studio	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
49	Coronado Village	Scenic Lookout	Park	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
73	Forest Park	Park	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
84	Greenwood Acres	Spa	Park	Yoga Studio	Convenience Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
94	Highcroft	Karaoke Bar	Park	Yoga Studio	Eye Doctor	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant
99	Hillsdale Forest	Park	Auto Workshop	Yoga Studio	Food Truck	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant

Cluster 7:

	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
77	Giverny	Arcade	Yoga Studio	Eye Doctor	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant
133	Madison Oaks	Gym / Fitness Center	Gym	Yoga Studio	Food Stand	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
161	Picardy Pointe	Gym / Fitness Center	Tennis Court	Food Stand	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
180	Savon Heights	Print Shop	Gym / Fitness Center	Brewery	Arcade	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
202	Strathmore at Cary	Print Shop	Gym / Fitness Center	Brewery	Arcade	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
233	Westfield	Print Shop	Gym / Fitness Center	Brewery	Arcade	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market

Cluster 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
58	Downing Place	Yoga Studio	Athletics & Sports	Deli / Bodega	Department Store	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot
59	Dublin Woods	Athletics & Sports	Bus Station	Food Truck	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
65	Fairview Townhomes	Yoga Studio	Dog Run	Athletics & Sports	Department Store	Fabric Shop	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop
92	Heritage Pines	Athletics & Sports	Fried Chicken Joint	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
114	Kistler Court	Athletics & Sports	Fried Chicken Joint	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
174	Rosewall Townhomes at Wimbledon	Yoga Studio	Dog Run	Athletics & Sports	Department Store	Fabric Shop	Food Stand	Food Court	Food & Drink Shop	Food	Flower Shop

Cluster 9:

	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
17	Blackstone at Amberly	Lake	Bagel Shop	Food Truck	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
28	Cameron Pond	Playground	Yoga Studio	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
41	Cheverly	Lake	Playground	Yoga Studio	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market
42	Chimney Rise Townhomes	Lake	Food Truck	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
118	Lakeridge Townhomes	Lake	Resort	Playground	Martial Arts Dojo	Yoga Studio	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant
146	Normandale	Lake	Food Truck	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
206	Sussex	Lake	Food Truck	Food Court	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm

Cluster 10

	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
21	Briarcliff	Resort	Yoga Studio	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
51	Courtney Pines	Resort	Yoga Studio	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
139	Maynard Woods	Resort	Yoga Studio	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm
244	Wishing Well Village	Resort	Yoga Studio	Electronics Store	Food & Drink Shop	Food	Flower Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm

By examining the venue categories of each cluster, we see that:
Cluster 1 has food-related venue listed as 1st most common venue
Cluster 2 has food-related venue listed as 3rd most common venue or after
Cluster 3 has food-related venue listed as 5th most common venue or after
Cluster 4 has food-related venue listed as 4th most common venue or after
Cluster 5 has food-related venue listed as 4th most common venue or after
Cluster 6 has food-related venue listed as 3rd most common venue or after
Cluster 7 has food-related venue listed as 3rd most common venue or after
Cluster 8 has food-related venue listed as 3rd most common venue or after
Cluster 9 has food-related venue listed as 2nd most common venue or after
Cluster 10 has food-related venue listed as 4th most common venue or after

It seems that the difference between popularities of food-related venues in these 10 clusters are not so big. We could either choose Cluster 1 since food-related venue is the most common venue in the neighborhoods (they even have Chinese restaurants) so that more people would go there for foods, or we could choose Cluster 3 because food-related venues are the least common venue among all clusters, so that the competition would not be severe if one opens a Chinese restaurant there.

If we choose Cluster 3, we could go on to check each neighborhood within in, for example, the neighborhood "Andover":

	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
2	Andover	Lake	Pool	Home Service	Performing Arts Venue	Trail	Harbor / Marina	Farm	Ethiopian Restaurant	Eye Doctor	Fabric Shop

We see that it has lake, pool home service, performing arts venue etc., and only "Ethiopian restaurant" is listed as 8th most common venue in this neighborhood. This would be an ideal place to open a Chinese restaurant if we don't want too much competition.

5. Discussion

We have chosen Cary, North Carolina as the target city to open a Chinese restaurant due to its large Chinese-American population, small number of available Chinese restaurants, and highest livability score. Actually we could also choose some other cities in the table, for example, Enterprise in Nevada, Fairfax in Virginia or Ann Arbor in Michigan as our target cities to live in and open a

Chinese restaurant, since the livability score in these cities are also good. Practically, new immigrants may also want to consider many other factors such as whether they have relatives nearby or how good is the local education (if they come to the United States mainly because they want a better education for their children).

Out[15]:

	City	State	Chinese-Americans	Percentage	latitude	longitude	number	livability score
16	El Monte	California	20190	17.5	36.701463	-118.755997	0	63
40	Enterprise	Nevada	5879	4.5	36.002834	-115.201299	12	75
25	Elk Grove	California	10758	6.6	38.408799	-121.371618	27	75
44	Cary	North Carolina	5283	3.4	35.788305	-78.781196	32	86
14	Fairfax	Virginia	28806	2.5	38.846224	-77.306373	34	76
42	Richmond	California	5523	5.1	37.935758	-122.347749	38	57
31	Ann Arbor	Michigan	7998	6.8	42.268157	-83.731229	40	79
37	Naperville	Illinois	6584	4.5	41.772870	-88.147928	42	75

My discussion here just gives a start for choosing the best place and things could be done further to improve our model (also makes it more complicated). For example, we could add consideration of local education quality if the new Chinese immigrants are particularly interested in living in a city / neighborhood with higher-quality local education. We could even have different factors to choose from depending on the new immigrants' needs.

6. Conclusion

In this study, I analyzed the Chinese-American population data among different large-sized and medium sized cities, and combined them with Foursquare data of available Chinese restaurants and AreaVibes Livability Score to determine that Cary in North Carolina could be the best city for new Chinese immigrants to live in and start their food-feature business (open a Chinese restaurant). I have also analyzed the neighborhoods data in Cary and provided a few suggestions about the detailed location / neighborhood to open the restaurant. We could also explore our search to other cities if the new immigrants have other needs such as local education quality.