1. Introduction

Phoenix is becoming a public city for people to retire. Although it is hot during summer time, the living cost is well balanced. Recently, Alex and Jenny moved from Chicago to Phoenix and are looking for a place to open a restaurant. They would like to know what type of food they want to server. Who are the targeting customers and what is the price range? They also want to pick up a location where there are not having too many competitions but also with reasonable rent.

2. Data Source

Three sources of data, including database from Foursquare, ref.1 for zip code, and ref.2 for population, will be used for the recommending system. Data such as different type of restaurants, traffics, surroundings, etc. will be extracted and used for analysis. In addition, the population distribution will also be used for analysis.

3. Data Cleaning and Exploratory

The data from above mentioned sources were extracted and compiled for predictive model. The venue, specifically different types of restaurants, from Foursquare was merged with zip code, which provided neighborhood longitude and latitude. The population was also merged to form a data frame for predictive model.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	85003	33.452	-112.075	Phoenix Public Market CafÃ©	33.455692	-112.073086	New American Restaurant
1	85003	33.452	-112.075	The Breadfruit	33.455662	-112.072009	Caribbean Restaurant
2	85003	33.452	-112.075	The Vig Fillmore	33.454926	-112.079308	American Restaurant
3	85003	33.452	-112.075	Harumi Sushi	33.449785	-112.075276	Sushi Restaurant
4	85003	33.452	-112.075	The Arrogant Butcher	33.447354	-112.072632	American Restaurant
5	85003	33.452	-112.075	Chick-fil-A	33.453619	-112.071545	Fast Food Restaurant

```
   zipcode  Population  Latitude  Longitude
0   85003     9369.0     33.452   -112.075
1   85004     4965.0     33.451   -112.068
2   85006    25742.0     33.469   -112.046
3   85007    14040.0     33.448   -112.090
4   85008    56145.0     33.470   -111.999
5   85009    52520.0     33.457   -112.123
6   85012     6390.0     33.513   -112.066
```


4. Modeling Results

K-Means classification model was used to classify the neighborhood in terms of the number of restaurants and populations. The results were classified as the most popular restaurants in the area with population.

```
----85003----
                     venue  freq
0        American Restaurant  0.27
1    New American Restaurant  0.20
2        Mexican Restaurant  0.13
3          Sushi Restaurant  0.07
4  Mediterranean Restaurant  0.07


----85004----
                            venue  freq
0             American Restaurant  0.35
1         New American Restaurant  0.18
2  Southern / Soul Food Restaurant  0.12
3            Caribbean Restaurant  0.06
4               Sushi Restaurant  0.06
```

American restaurants:

	Population	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	9369.0	33.452	-112.075	3	American Restaurant	New American Restaurant	Mexican Restaurant	Mediterranean Restaurant	Sushi Restaurant
1	4965.0	33.451	-112.068	3	American Restaurant	New American Restaurant	Southern / Soul Food Restaurant	Mediterranean Restaurant	Sushi Restaurant
3	14040.0	33.448	-112.090	3	Mexican Restaurant	American Restaurant	New American Restaurant	Southern / Soul Food Restaurant	Sushi Restaurant
6	6390.0	33.513	-112.066	3	Mexican Restaurant	American Restaurant	Sushi Restaurant	New American Restaurant	Latin American Restaurant
8	24782.0	33.514	-112.053	3	Mexican Restaurant	American Restaurant	New American Restaurant	Sushi Restaurant	Latin American Restaurant
10	33896.0	33.504	-112.029	3	American Restaurant	New American Restaurant	Mexican Restaurant	Italian Restaurant	Sushi Restaurant

Mexico Restaurants:

	Population	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
2	25742.0	33.469	-112.046	1	Mexican Restaurant	New American Restaurant	American Restaurant	Restaurant
4	56145.0	33.470	-111.999	1	Mexican Restaurant	Italian Restaurant	American Restaurant	Restaurant
5	52520.0	33.457	-112.123	1	Mexican Restaurant	American Restaurant	New American Restaurant	Italian Restaurant
7	19314.0	33.510	-112.080	1	Mexican Restaurant	Italian Restaurant	New American Restaurant	Restaurant
9	37644.0	33.508	-112.101	1	Mexican Restaurant	American Restaurant	Vietnamese Restaurant	New American Restaurant
11	38872.0	33.511	-112.122	1	Mexican Restaurant	Vietnamese Restaurant	Thai Restaurant	American Restaurant

Asia Restaurants:

	Population	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	C
33	7238.0	33.304	-112.114	2	Asian Restaurant	Vietnamese Restaurant	Fondue Restaurant	Hotpot Restaurant	Hawaiian Restaurant	H

In [28]:
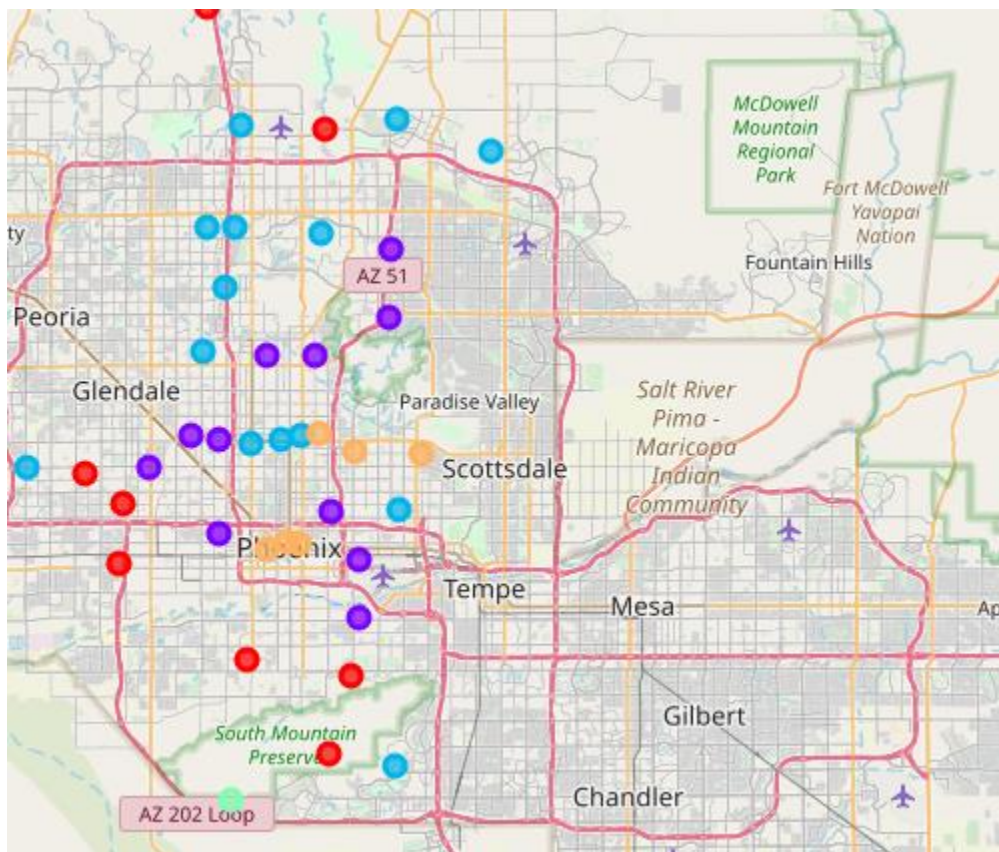
5. Conclusion

If you want to have a new restaurant in the Phoenix area, you can either have the same type of restaurant in the area with most populations or try to open a total different type of restaurant based on the above result.

Ref1. https://phoenix.areaconnect.com/zip2.htm?search=zip
Ref2. https://www.zipdatamaps.com/zipcodes-phoenix-az