

Predicting to Open a New Chinese Restaurant in St. Louis

Capstone Project - The Battle of the Neighborhoods

Han Zhang

6/14/2020

Outline

1. Introduction
2. Objective
3. Data Acquisition
4. Methodology
5. Result and Discussion
6. Conclusions

1. Introduction

- **St. Louis locates in the Mid-west of the Unites States, which is the largest metropolitan area in the state of Missouri.**
- **American-style restaurants are most common type of restaurants in St. Louis.**
- **Increases the demand for the diversity of its food culture, like Chinese food.**



2. Objective

- Find optimal location candidates in St. Louis for the stakeholders that interest in opening an Authentic Chinese restaurant

3. Data Acquisition

- **79 neighborhoods in St. Louis**
- **Data needs to be acquired for analysis**
 - **The population of each neighborhood**
 - **The existing restaurants in each neighborhood**
 - **The shops near the neighborhood.**

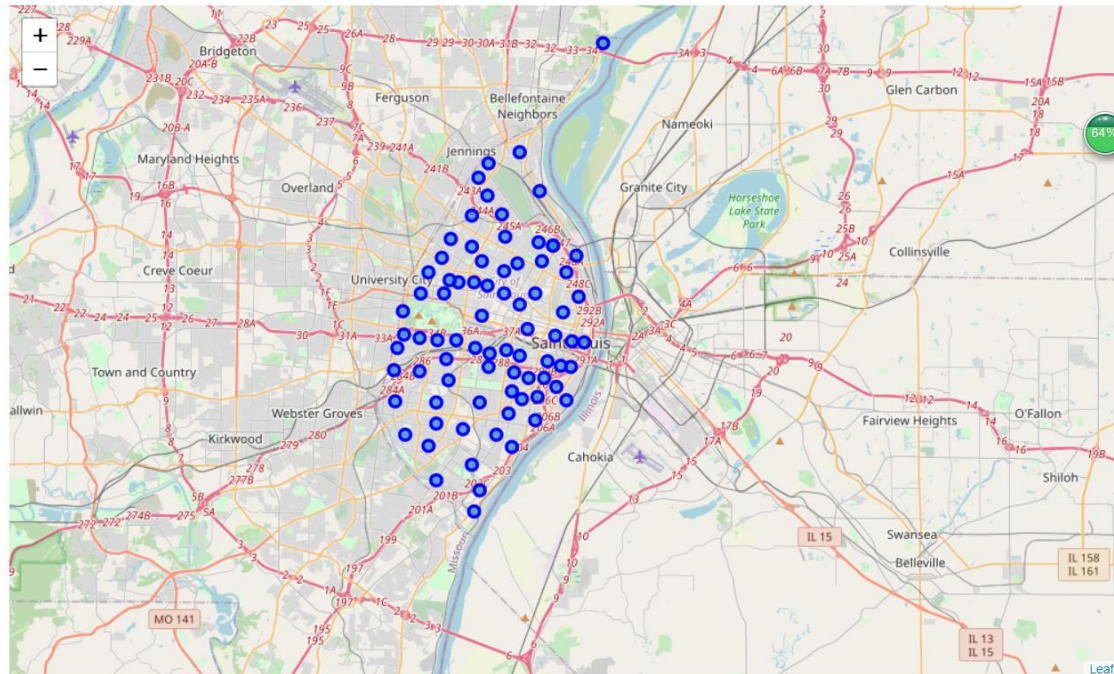
4. Methodology

- **The key methodologies to process data are:**
 - **The list of neighborhoods in St. Louis from Wikipedia**
 - **The address of the neighborhoods will be obtained by using St. Louis Coordinate Datasheet**
 - **The number and types of restaurants will be obtained by using Foursquare API**
 - **The number of shops will be obtained by using Foursquare API**

4. Methodology

❖ Cleaning and Visualizing the Neighborhoods

- **Import the neighborhood information and coordinates of the neighborhoods.**
- **Clear the data and form a new dataframe with neighborhood, population, latitude, and longitude.**
- **Visualize the location of the 79 neighborhoods on the map.**



5. Results and Discussion

❖ Explore the neighborhoods in St. Louis

- **Descending the neighborhoods by the population**
- **The population ranges from 15,770 to 14 people for the 79 neighborhood communities**
- **Mean population is 4044.**

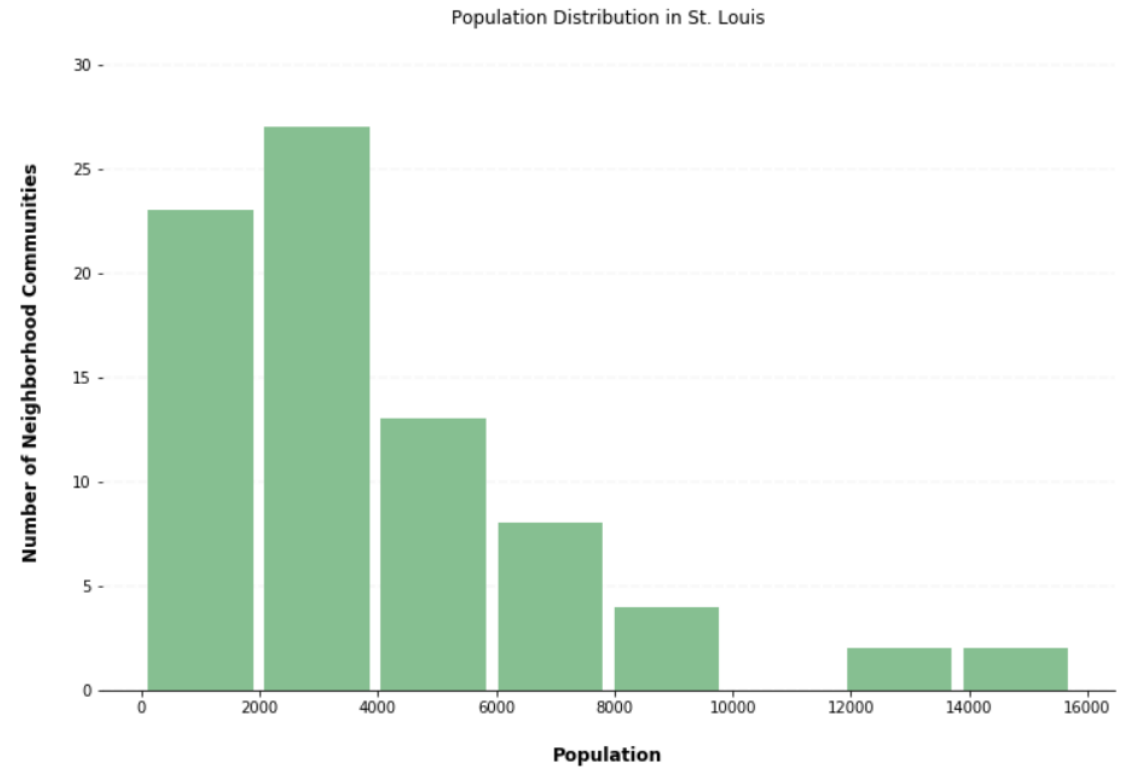
	Neighborhood	Population	Latitude	Longitude
19	Dutchtown	15770	38.5819	-90.2456
9	Central West End	14473	38.6396	-90.2550
70	Tower Grove South	13333	38.5975	-90.2561
4	Bevo Mill	12654	38.5845	-90.2667
43	Lindenwood Park	9486	38.5983	-90.3085

count	79.000000
mean	4044.075949
std	3212.850207
min	14.000000
25%	1868.500000
50%	3349.000000
75%	5604.500000
max	15770.000000

5. Results and Discussion

❖ Explore the neighborhoods in St. Louis

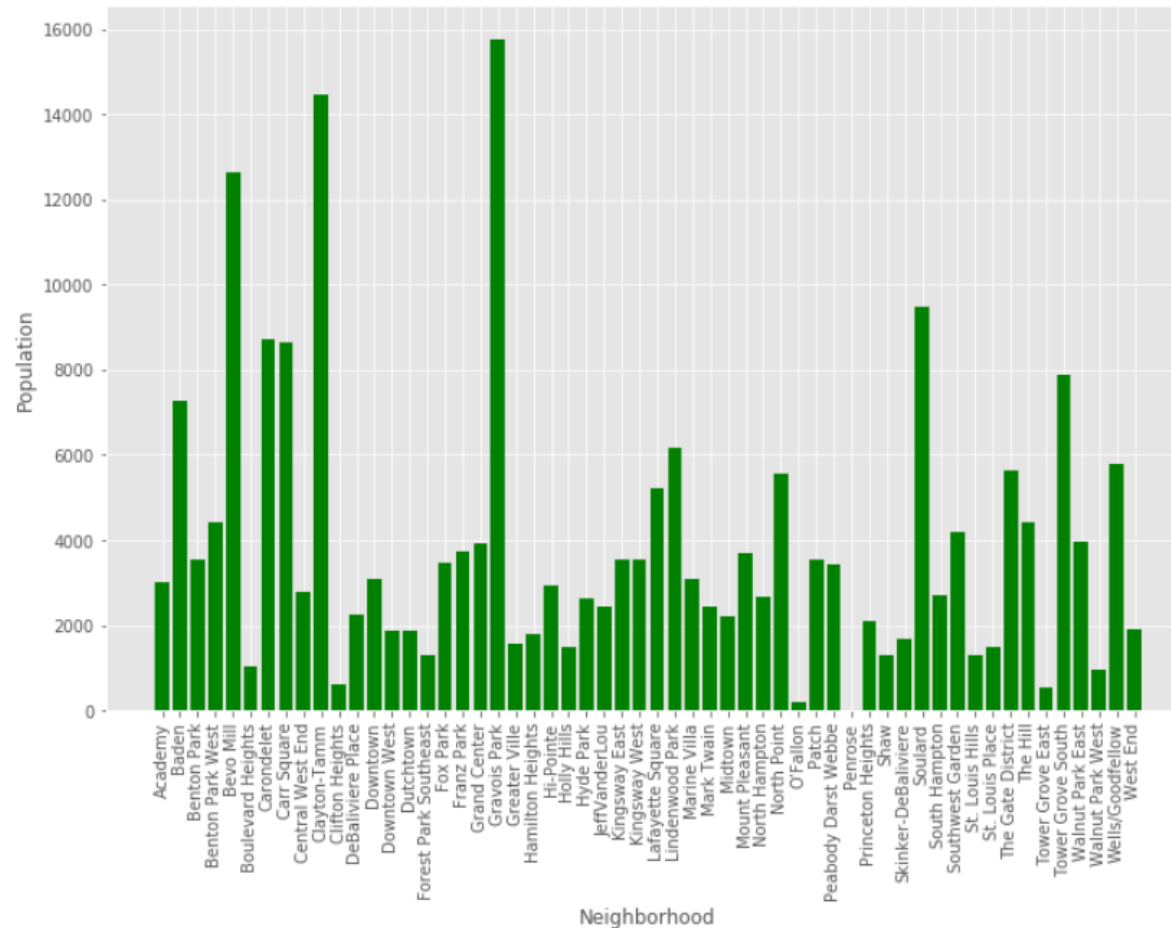
- **Divided the population in 8 ranges**
- **Bar chart to visualize relation between population and the number of neighborhood communities**
- **Neglect population less than 2000**
- **56 communities left**



5. Results and Discussion

❖ Explore the neighborhoods in St. Louis

■ **Visualize the population of each neighborhoods**



5. Results and Discussion

❖ Explore the venues of all the neighborhoods

- **Using Foursquare API to get the venues of each neighborhoods and group them by neighborhood.**

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Dutchtown	38.5819	-90.2456	Ted Drewes Frozen Custard	38.580181	-90.244437	Ice Cream Shop
1	Dutchtown	38.5819	-90.2456	Banh Mi So #1	38.582861	-90.244694	Vietnamese Restaurant
2	Dutchtown	38.5819	-90.2456	Urban Eats	38.580605	-90.241099	Sandwich Place
3	Dutchtown	38.5819	-90.2456	Sister Cities Cajun & Bbq	38.581186	-90.244262	Cajun / Creole Restaurant
4	Dutchtown	38.5819	-90.2456	Feasting Fox Restaurant and Pub	38.580997	-90.244095	German Restaurant

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Academy	6	6	6	6	6	6
Baden	4	4	4	4	4	4
Benton Park	22	22	22	22	22	22
Benton Park West	4	4	4	4	4	4
Bevo Mill	11	11	11	11	11	11

5. Results and Discussion

❖ Explore the shops in all venues

- **Filter the stores in the overall venues with mean frequencies and the summation of frequencies.**

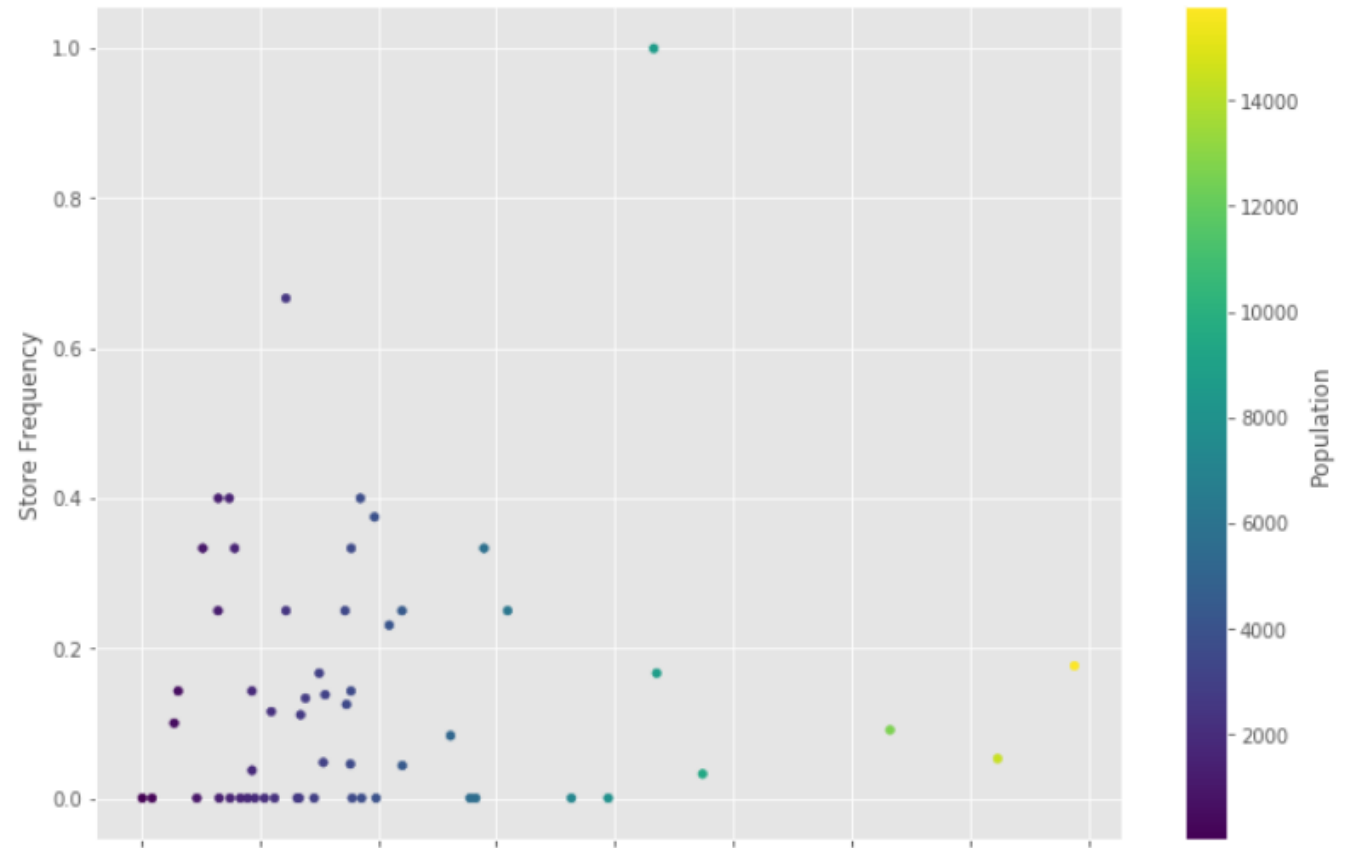
	Neighborhood	Population	Arts & Crafts Store	Candy Store	Clothing Store	Convenience Store	Discount Store	Electronics Store	Fruit & Vegetable Store	Grocery Store	Liquor Store	Men's Store	Music Store	Pet Store	Shoe Store	Thrift / Vintage Store	Toy / Game Store	Video Game Store	Video Store	Sum
0	Academy	3006	0.000000	0.0	0.000000	0.166667	0.000000	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.333333
1	Baden	7268	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000
2	Benton Park	3532	0.000000	0.0	0.045455	0.000000	0.000000	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.090909
3	Benton Park West	4404	0.000000	0.0	0.000000	0.250000	0.000000	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.500000
4	Bevo Mill	12654	0.000000	0.0	0.000000	0.000000	0.090909	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.181818
5	Boulevard Heights	1037	0.166667	0.0	0.000000	0.000000	0.166667	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.666667
6	Carondelet	8708	0.000000	0.0	0.000000	0.166667	0.000000	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.333333
7	Carr Square	8661	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	1.0	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	2.000000
8	Central West End	2774	0.000000	0.0	0.000000	0.066667	0.000000	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.066667	0.0	0.0	0.0	0.266667
9	Clayton-Tamm	14473	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.052632	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.105263

5. Results and Discussion

❖ Explore shop frequency with population

- **Scatter plot to show the relation between shop frequency and population.**
- **The mean shop frequency is 14.6%**

count	56.000000
mean	0.145529
std	0.188385
min	0.000000
25%	0.000000
50%	0.095455
75%	0.250000
max	1.000000



5. Results and Discussion

❖ Explore the restaurants in St. Louis

- **Filter the restaurants in the overall venues with mean frequencies and the summation of frequencies.**
- **Total 26 types of restaurants in St. Louis**

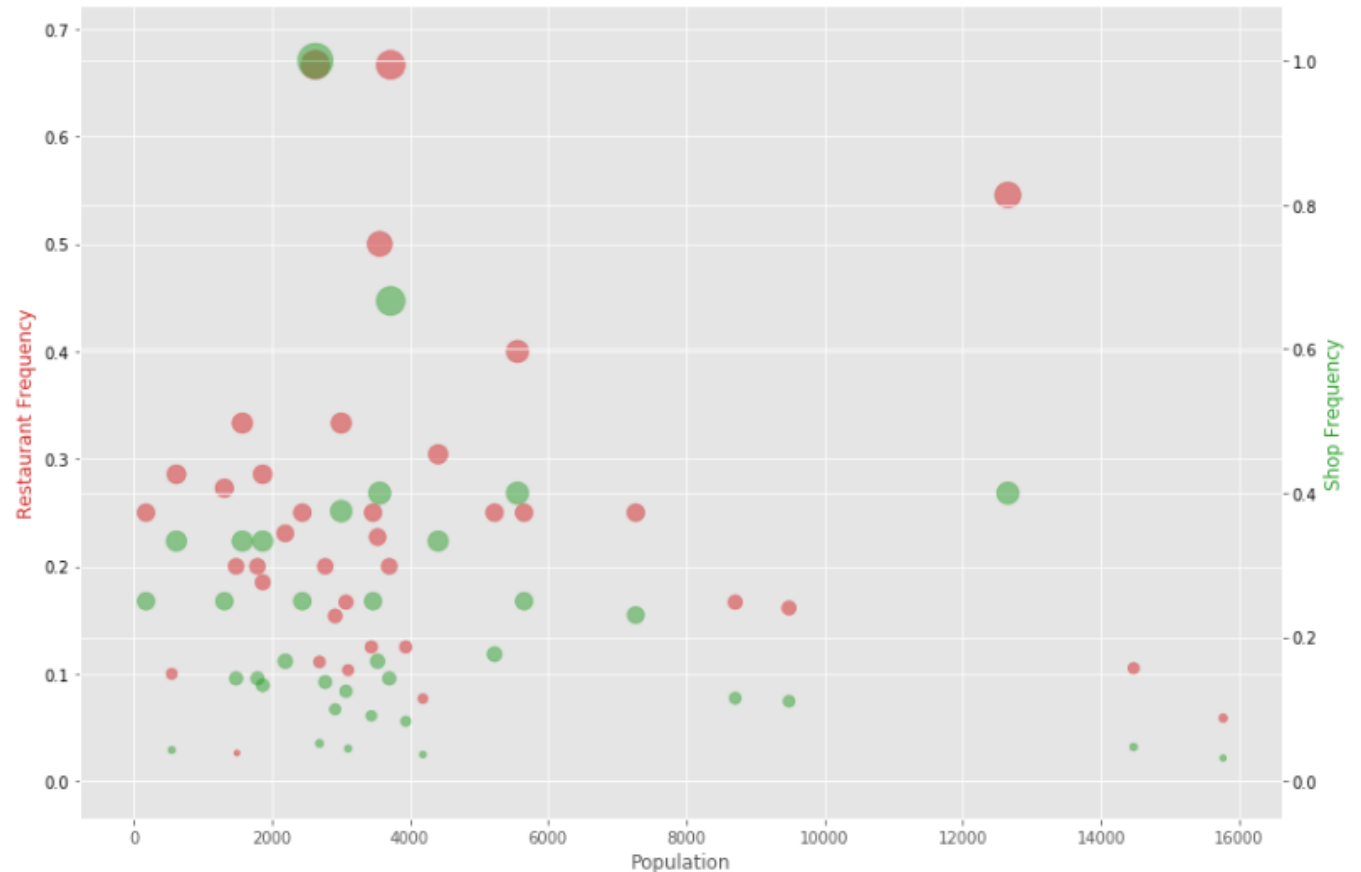
	Neighborhood	Population	American Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	...	Mexican Restaurant	Middle Eastern Restaurant	New American Restaurant	Restaurant	Seafood Restaurant	Southern / Soul Food Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Sum
0	Academy	3006	0.000000	0.0	0.000000	0.0	0.333333	0.000000	0.0	0.000000	...	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.333333
1	Baden	7268	0.000000	0.0	0.000000	0.0	0.000000	0.000000	0.0	0.250000	...	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.250000
2	Benton Park	3532	0.045455	0.0	0.045455	0.0	0.000000	0.000000	0.0	0.045455	...	0.000000	0.0	0.045455	0.000000	0.000000	0.0	0.0	0.0	0.0	0.227273
3	Benton Park West	4404	0.000000	0.0	0.000000	0.0	0.000000	0.000000	0.0	0.000000	...	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000
4	Bevo Mill	12654	0.000000	0.0	0.000000	0.0	0.000000	0.090909	0.0	0.000000	...	0.090909	0.0	0.000000	0.181818	0.000000	0.0	0.0	0.0	0.0	0.545455
5	Boulevard Heights	1037	0.000000	0.0	0.000000	0.0	0.000000	0.000000	0.0	0.000000	...	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000
6	Carondelet	8708	0.000000	0.0	0.000000	0.0	0.000000	0.000000	0.0	0.166667	...	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.166667
7	Carr Square	8661	0.000000	0.0	0.000000	0.0	0.000000	0.000000	0.0	0.000000	...	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000
8	Central West End	2774	0.000000	0.0	0.000000	0.0	0.066667	0.000000	0.0	0.000000	...	0.000000	0.0	0.066667	0.000000	0.066667	0.0	0.0	0.0	0.0	0.200000
9	Clayton-Tamm	14473	0.052632	0.0	0.000000	0.0	0.052632	0.000000	0.0	0.000000	...	0.000000	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.105263

5. Results and Discussion

❖ Explore restaurant frequency with population

- **Bubble chart to show the relation among restaurant frequency, shop frequency and population.**
- **The mean shop frequency is 16.1%**

```
count    56.000000  
mean      0.161028  
std       0.168508  
min       0.000000  
25%      0.000000  
50%      0.139423  
75%      0.250000  
max       0.666667
```



5. Results and Discussion

❖ Cluster the neighborhoods by restaurant venues

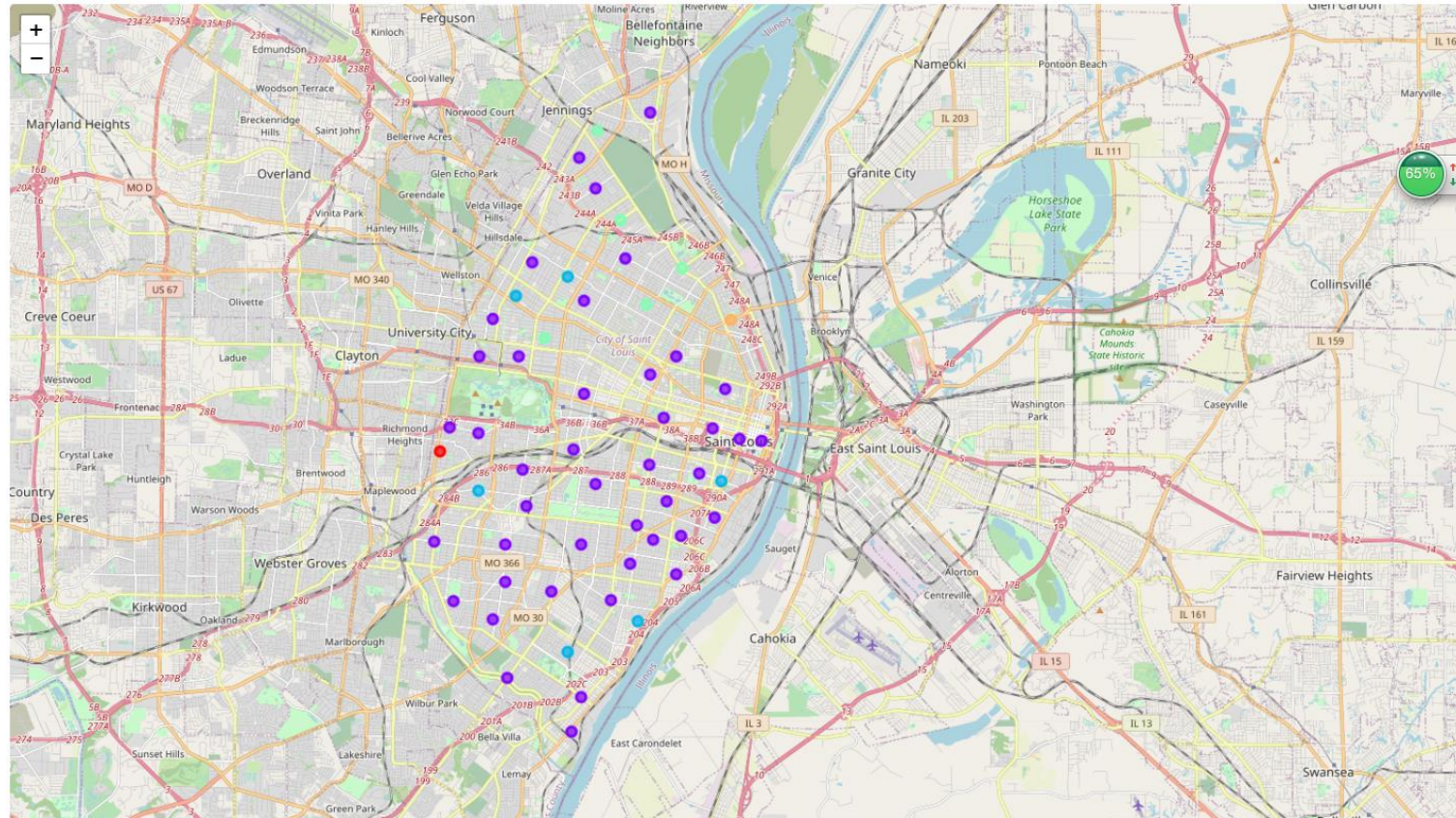
▪ **List most 10 common venues of the restaurants**

	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	Academy	Chinese Restaurant	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
1	Baden	Fast Food Restaurant	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Filipino Restaurant
2	Benton Park	Italian Restaurant	New American Restaurant	Brazilian Restaurant	Fast Food Restaurant	American Restaurant	Restaurant	Seafood Restaurant	Asian Restaurant	Sushi Restaurant	Cajun / Creole Restaurant
3	Benton Park West	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
4	Bevo Mill	Restaurant	Italian Restaurant	German Restaurant	Eastern European Restaurant	Mexican Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Falafel Restaurant

5. Results and Discussion

❖ Cluster the neighborhoods by restaurant venues

■ **Cluster the neighborhoods by restaurant in 5 groups**



5. Results and Discussion

❖ Comprehensively analyze the restaurant information with shops in each cluster

■ Cluster 1

	Neighborhood	Shop Frequency	Population	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	Franz Park	0.0	2442	Greek Restaurant	Mediterranean Restaurant	Vietnamese Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant

■ Cluster 2

- 43 neighborhoods

	Neighborhood	Shop Frequency	Population	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
14	Dutchtown	0.142857	15770	Vietnamese Restaurant	Cajun / Creole Restaurant	German Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant
8	Central West End	0.133333	14473	Seafood Restaurant	New American Restaurant	Chinese Restaurant	Vietnamese Restaurant	German Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Eastern European Restaurant	Falafel Restaurant
51	Tower Grove South	0.000000	13333	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
4	Bevo Mill	0.090909	12654	Restaurant	Italian Restaurant	German Restaurant	Eastern European Restaurant	Mexican Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Falafel Restaurant
29	Lindenwood Park	0.250000	9486	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant

5. Results and Discussion

❖ Comprehensively analyze the restaurant information with shops in each cluster

■ Cluster 2

- **Filter with shop and population larger than mean value**

	Neighborhood	Shop Frequency	Population	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
29	Lindenwood Park	0.250000	9486	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
5	Boulevard Heights	0.333333	8708	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
6	Carondelet	0.166667	8661	Fast Food Restaurant	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Filipino Restaurant
46	St. Louis Hills	0.250000	7373	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
41	Shaw	0.400000	6811	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
54	Wells/Goodfellow	0.333333	5859	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
25	JeffVanderLou	0.666667	5557	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
19	Gravois Park	0.176471	5225	Southern / Soul Food Restaurant	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant
45	Southwest Garden	0.230769	4885	Italian Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
3	Benton Park West	0.250000	4404	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant

5. Results and Discussion

❖ Comprehensively analyze the restaurant information with shops in each cluster

■ Cluster 3

- 6 neighborhoods

	Neighborhood	Shop Frequency	Population	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
33	Mount Pleasant	0.400000	4408	American Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
23	Holly Hills	0.400000	3701	American Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
27	Kingsway West	0.000000	3441	American Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
21	Hamilton Heights	0.000000	3105	American Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
10	Clifton Heights	0.142857	3074	Italian Restaurant	American Restaurant	Sushi Restaurant	German Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant
38	Peabody Darst Webbe	0.250000	2378	American Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant

5. Results and Discussion

❖ Final Decision

- **Final 5 candidates for stockholders to open a new Chinese restaurant**

	Neighborhood	Shop Frequency	Population	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	Dutchtown	0.142857	15770	Vietnamese Restaurant	Cajun / Creole Restaurant	German Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant
1	Lindenwood Park	0.250000	9486	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
2	Boulevard Heights	0.333333	8708	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
3	Shaw	0.400000	6811	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
4	JeffVanderLou	0.666667	5557	Vietnamese Restaurant	Greek Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Eastern European Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant

6. Conclusions

- **The objective of this project is reached.**
 - **Provide 5 candidates for the stockholders to open a new Chinese restaurant in St. Louis**
- **Population is major considered factor; Shop frequency is secondary considered factor.**
- **Five candidate neighborhoods are:**
 - **Dutchtown, Lindenwood Park, Boulevard Heights, Shaw, JeffVanderLou**