The Battle of Neighborhoods

Analysis of Chicago Suburbs for A Prospective Restaurant Location

Introduction

Chicago is the third most populous city in the United States and has diverse population. Chicago's 58 million domestic and international visitors in 2018 made it the second most visited city in US when compared to New York City's 65 million visitors. Chicago is famous for large number of regional specialties (cuisines) which reflect the city's ethnic and working-class roots. Large number of world-famous chefs have their restaurants in the City of Chicago.

Business Problem

Our Client is running a fine dining restaurant in downtown Chicago; client wants to expand their operations into the suburbs of Chicago. Top 20 Suburbs of Chicago (Suburbs with population more than 5000) are been considered as prospective locations for the new restaurant. Based on analysis of various data collected from United States Census Bureau and Foursquare API, ideal suburb will be identified for the new restaurant.

Data

1) Top 20 suburbs of Chicago metropolitan area (In the Data set Suburbs with population less than 5000 has been excluded) are based on the information available in www.niche.com
   Ref:https://www.niche.com/places-to-live/search/best-suburbs/m/chicago-metro-area/

2) Suburbs Population data (Population estimates, July 1, 2019, (V2019)) is obtained from United States Census Bureau website for the suburbs obtained from niche website.
   Ref: https://www.census.gov/quickfacts/fact/table/US/PST045219

3) Each suburbs Latitude and Longitude is retrieved using OpenCage Geocoder Python package.

4) Foursquare API is used to explore different venues and frequencies in each suburb, data retrieved will be helpful in clustering the suburbs based on their development.
   Foursquare API Developer Access: https://developer.foursquare.com/

Table 1: Suburb's Top 10 Venues

Suburb	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
Glencoe village	Coffee Shop	Pharmacy	Bank	Cosmetics Shop	Locksmith	Chinese Restaurant	Café	Gourmet Shop	Bistro	Baseball Field
Hinsdale village	Bakery	Yoga Studio	Shipping Store	Massage Studio	Optical Shop	Moving Target	Jewelry Store	Italian Restaurant	Hot Dog Joint	Grocery Store
Long Grove village	Gift Shop	Arts & Crafts Store	Café	American Restaurant	Dessert Shop	Coffee Shop	Brewery	Construction & Landscaping	Pizza Place	Italian Restaurant
Western Springs village	American Restaurant	Hardware Store	Dessert Shop	Farmers Market	Coffee Shop	Convenience Store	Library	Italian Restaurant	Optical Shop	Ice Cream Shop
Wilmette village	Pizza Place	Jewelry Store	Butcher	Convenience Store	Baseball Field	Pet Store	Park	Diner	Antique Shop	Dessert Shop

Methodology

Methodology includes:

1. Data Retrieval, exploration, and wrangling
2. Performing K-mean Clustering
3. Visualizing the venues, median income, and clustering of the suburbs

Data Retrieval, exploration, and wrangling

1) United States Census Bureau website is used to obtain the 2019 estimated dataset for the suburbs shortlisted from niche website. Information in the Census Bureau website can be downloaded for 5 suburbs as an excel file. Multiple excel files have been combined and few of the columns which obsolete and of no importance has been removed.

Table 2: Chicago Suburb's Population Data

Suburb	PE2019	PE2010	PChange	U18	A65	Bdegree	MIncome	PCIncome	Poverty	Larea
Glencoe village	8826	8721	0.012	0.294	0.171	0.864	209143	130020	0.023	3.72
Long Grove village	7905	7921	-0.002	0.241	0.15	0.712	208250	80896	0.023	12.48
Hinsdale village	17637	16845	0.047	0.316	0.14	0.803	188684	92734	0.029	4.6
Western Springs village	13359	12948	0.032	0.318	0.157	0.784	161563	71004	0.032	2.79
Wilmette village	27089	27060	0.001	0.285	0.188	0.815	154738	82805	0.033	5.4
Highland Park city	29515	29745	-0.008	0.245	0.22	0.744	147962	86964	0.05	12.2
Deerfield village	18646	18227	0.023	0.274	0.158	0.786	144229	74334	0.027	5.58
Libertyville village	20205	20405	-0.01	0.246	0.16	0.701	130732	64831	0.031	8.81
Clarendon Hills village	8752	8431	0.038	0.334	0.115	0.762	126500	73245	0.05	1.8
Northbrook village	32958	33200	-0.007	0.219	0.27	0.692	123457	67306	0.03	13.19
La Grange village	15322	15554	-0.015	0.292	0.143	0.63	121425	55516	0.036	2.52
Naperville city	148449	142170	0.044	0.256	0.117	0.678	118187	53061	0.044	38.77
Buffalo Grove village	40494	41503	-0.024	0.226	0.147	0.665	111435	53380	0.044	9.5
Glen Ellyn village	27714	27773	-0.002	0.267	0.159	0.673	110669	59178	0.063	6.61
Wheaton city	52745	53045	-0.006	0.23	0.151	0.633	98544	48160	0.053	11.25
Vernon Hills village	26521	25005	0.061	0.262	0.132	0.634	97366	48541	0.062	7.71
La Grange Park village	13178	13580	-0.03	0.26	0.188	0.564	95466	45513	0.054	2.23
Oak Park village	52381	51878	0.01	0.243	0.144	0.697	91945	53972	0.08	4.7
Lisle village	23270	22552	0.032	0.189	0.151	0.599	87106	49024	0.06	6.84
Evanston city	73473	74483	-0.014	0.2	0.147	0.657	77848	49263	0.133	7.78

Column Description:

1) Suburb – Name of the Suburb in Chicago
2) PE2019 – 2019 Estimated population
3) PE2010 – 2010 Population
4) PChange – Percentage Change in population
5) U18 – Percentage of population below age of 18 years
6) A65 – Percentage of population equal to and above 65 years
7) Bdegree – Percentage of population above 25 years with bachelor's degree
8) MIncome – Median Income
9) PCIncome – Per capita Income
10) Poverty – Percentage of population below poverty line
11) Larea – Area of the suburb in square miles.

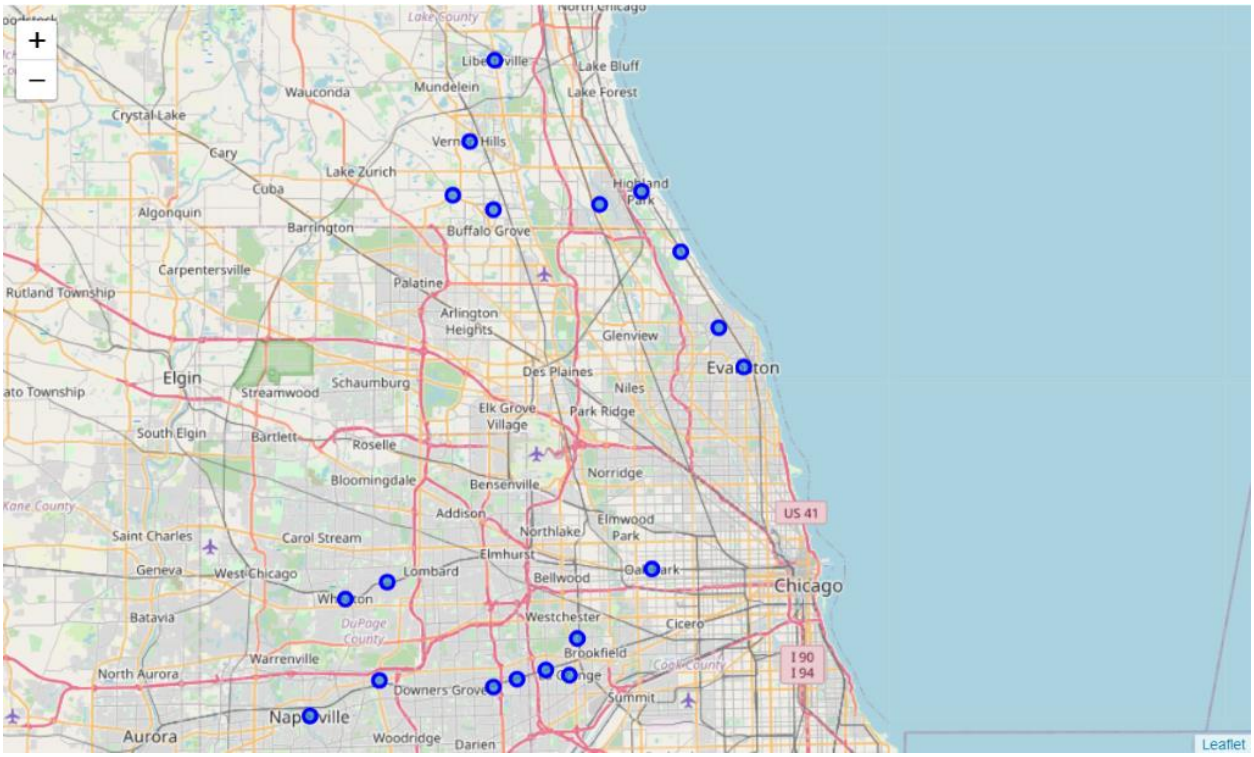
Only the above column has been selected from the census data set, other information has been ignored from the census dataset. Even though, only limited number of columns has been used in this analysis; left out column information can be used to make better decision.

2) Using OpenCage Geocoder Python package Latitude and Longitude of Suburbs will be retrieved.

Table 3: Suburbs Geo coordinates

Suburb	PE2019	PE2010	PChange	U18	A65	Bdegree	MIncome	PCIncome	Poverty	Larea	lat	lon
Clarendon Hills village	8752	8431	0.038	0.334	0.115	0.762	126500	73245	0.05	1.8	41.796124	-87.955143
Buffalo Grove village	40494	41503	-0.024	0.226	0.147	0.665	111435	53380	0.044	9.5	42.167583	-87.955640
Long Grove village	7905	7921	-0.002	0.241	0.15	0.712	208250	80896	0.023	12.48	42.178358	-87.997852
Naperville city	148449	142170	0.044	0.256	0.117	0.678	118187	53061	0.044	38.77	41.772870	-88.147928
Evanston city	73473	74483	-0.014	0.2	0.147	0.657	77848	49263	0.133	7.78	42.044739	-87.693046
Oak Park village	52381	51878	0.01	0.243	0.144	0.697	91945	53972	0.08	4.7	41.887814	-87.788762
Hinsdale village	17637	16845	0.047	0.316	0.14	0.803	188684	92734	0.029	4.6	41.802460	-87.929984
Western Springs village	13359	12948	0.032	0.318	0.157	0.784	161563	71004	0.032	2.79	41.809613	-87.900671
Vernon Hills village	26521	25005	0.061	0.262	0.132	0.634	97366	48541	0.062	7.71	42.220489	-87.980360
Wilmette village	27089	27060	0.001	0.285	0.188	0.815	154738	82805	0.033	5.4	42.075732	-87.719377
Wheaton city	52745	53045	-0.006	0.23	0.151	0.633	98544	48160	0.053	11.25	41.864696	-88.110171
Highland Park city	29515	29745	-0.008	0.245	0.22	0.744	147962	86964	0.05	12.2	42.181692	-87.800344
Libertyville village	20205	20405	-0.01	0.246	0.16	0.701	130732	64831	0.031	8.81	42.283079	-87.953130
La Grange village	15322	15554	-0.015	0.292	0.143	0.63	121425	55516	0.036	2.52	41.805458	-87.874864
Northbrook village	32958	33200	-0.007	0.219	0.27	0.692	123457	67306	0.03	13.19	39.959136	-91.375919
Glencoe village	8826	8721	0.012	0.294	0.171	0.864	209143	130020	0.023	3.72	42.135027	-87.758119
La Grange Park village	13178	13580	-0.03	0.26	0.188	0.564	95466	45513	0.054	2.23	41.833365	-87.867732
Glen Ellyn village	27714	27773	-0.002	0.267	0.159	0.673	110669	59178	0.063	6.61	41.877529	-88.067012
Lisle village	23270	22552	0.032	0.189	0.151	0.599	87106	49024	0.06	6.84	41.801141	-88.074787
Deerfield village	18646	18227	0.023	0.274	0.158	0.786	144229	74334	0.027	5.58	42.171137	-87.844512

Fig 1: Best Suburbs of Chicago



3) Shortlisting Top 5 Suburbs based on their Median Income.

Table 4: Top 5 Median Income Suburbs

Suburb	PE2019	PE2010	PChange	U18	A65	Bdegree	MIncome	PCIncome	Poverty	Larea	lat	lon
Glencoe village	8826	8721	0.012	0.294	0.171	0.864	209143	130020	0.023	3.72	42.135027	-87.758119
Long Grove village	7905	7921	-0.002	0.241	0.15	0.712	208250	80896	0.023	12.48	42.178358	-87.997852
Hinsdale village	17637	16845	0.047	0.316	0.14	0.803	188684	92734	0.029	4.6	41.802460	-87.929984
Western Springs village	13359	12948	0.032	0.318	0.157	0.784	161563	71004	0.032	2.79	41.809613	-87.900671
Wilmette village	27089	27060	0.001	0.285	0.188	0.815	154738	82805	0.033	5.4	42.075732	-87.719377

Performing K-mean Clustering

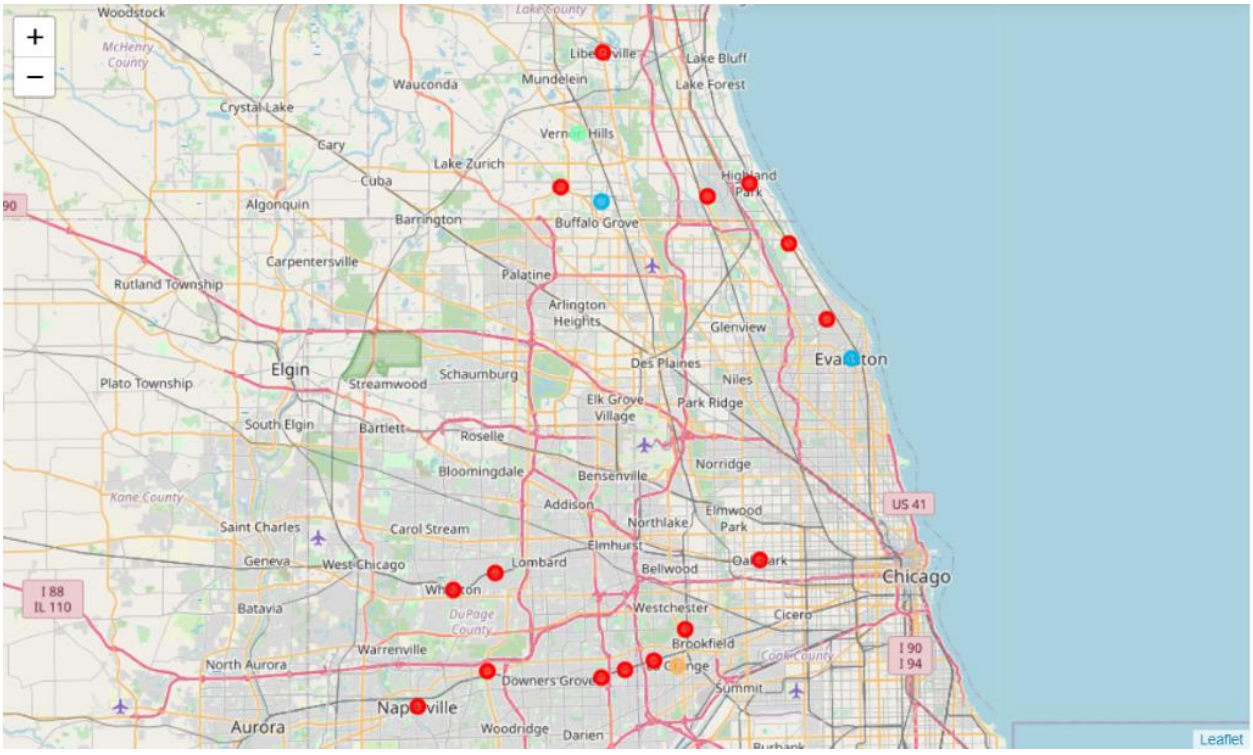
Top 10 Venues of the Suburbs has been selected and considered for K-mean clustering. Since 20 suburbs are been considered for K-mean clustering; clustering size k=5 was assumed.

Table 5: Top 10 Venues for each Suburb

Suburb	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
Glencoe village	Coffee Shop	Pharmacy	Bank	Cosmetics Shop	Locksmith	Chinese Restaurant	Café	Gourmet Shop	Bistro	Baseball Field
Hinsdale village	Bakery	Yoga Studio	Shipping Store	Massage Studio	Optical Shop	Moving Target	Jewelry Store	Italian Restaurant	Hot Dog Joint	Grocery Store
Long Grove village	Gift Shop	Arts & Crafts Store	Café	American Restaurant	Dessert Shop	Coffee Shop	Brewery	Construction & Landscaping	Pizza Place	Italian Restaurant
Western Springs village	American Restaurant	Hardware Store	Dessert Shop	Farmers Market	Coffee Shop	Convenience Store	Library	Italian Restaurant	Optical Shop	Ice Cream Shop
Wilmette village	Pizza Place	Jewelry Store	Butcher	Convenience Store	Baseball Field	Pet Store	Park	Diner	Antique Shop	Dessert Shop

After k-mean clustering, the suburbs are been mapped based on the Cluster. K-mean clustering does not give a proper interpretation and cluster are scattered. A better clustering size needs to be considered for further analysis. Since a clear picture is not obtained from clustering it is better to investigate median income and geological location of each suburbs.

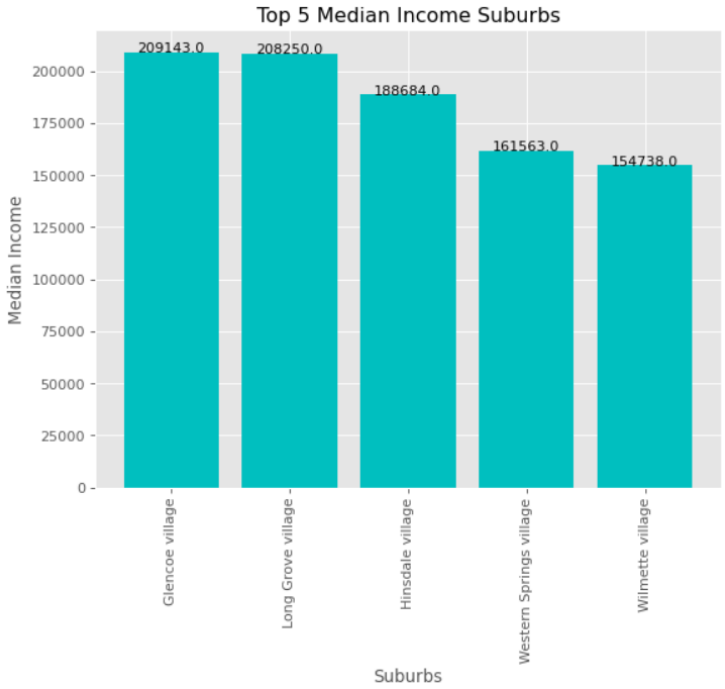
Fig 2: K-Mean Cluster Suburbs



Visualizing the venues, median income, and clustering of the suburbs

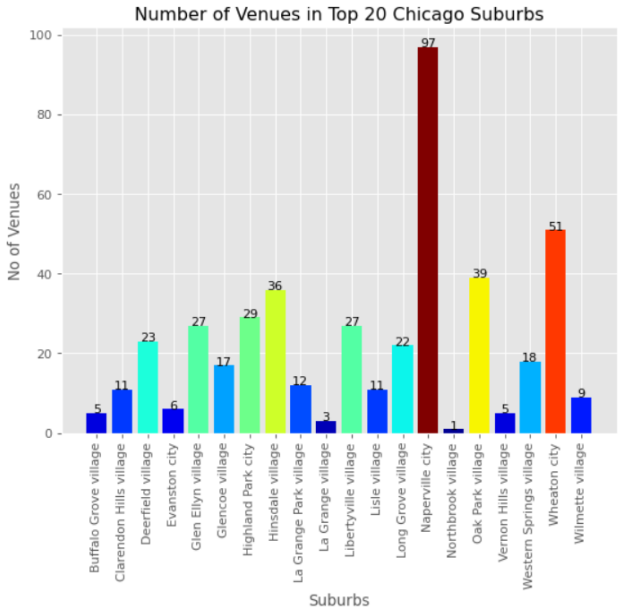
Since K-mean clustering did not led to any decision making; median income of the suburbs is been considered as major decision-making parameter in selecting the prospective location.

Fig 3: Top 5 Median Income Suburbs



Visualizing number of venues in each suburb. Based on the following chart Naperville has the highest number of venues followed by Wheaton city.

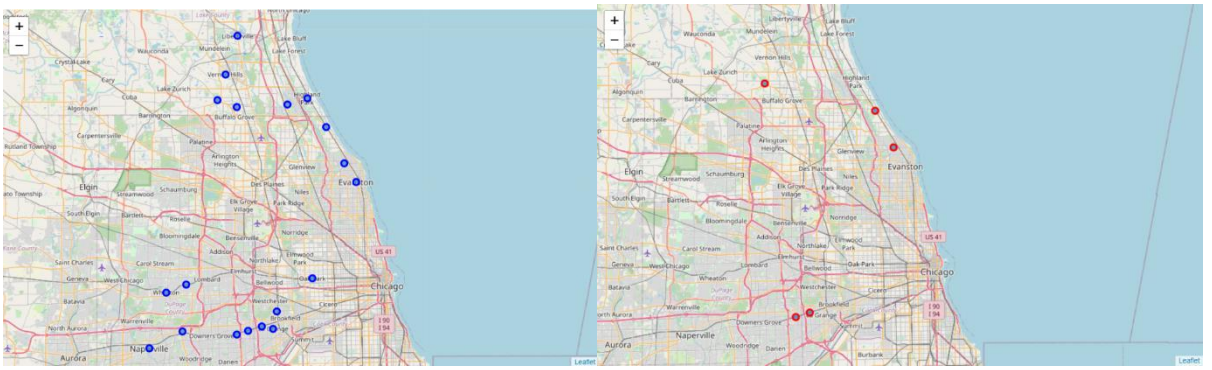
Fig 4: Suburb Venue Count



Results

Western Springs village has been considered as the prospective location for the new fine dining restaurant, this decision was enabled by Analysis of the Visual maps and location pattern. Western Springs Village is located nearer to another 4 suburbs in the list.

Fig 5: Comparison of 20 Suburbs vs Top 5 Median Income Suburbs



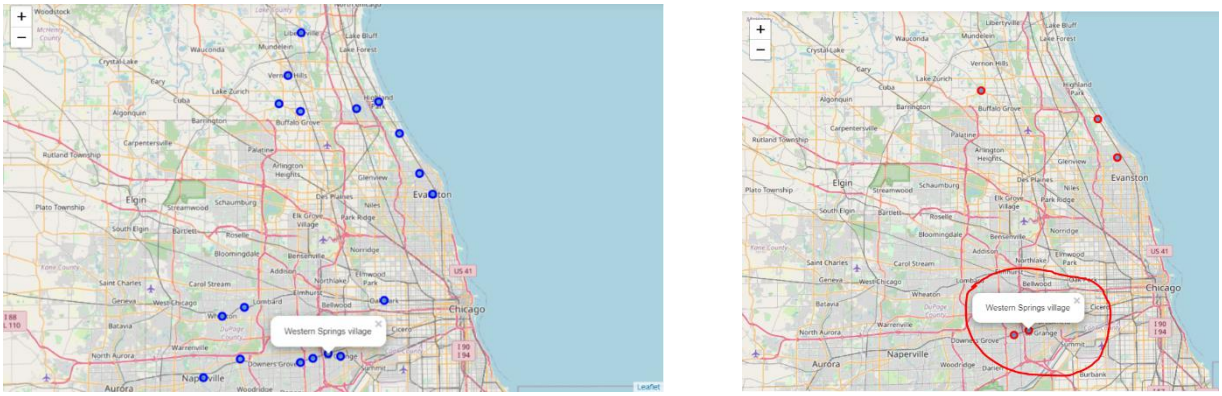
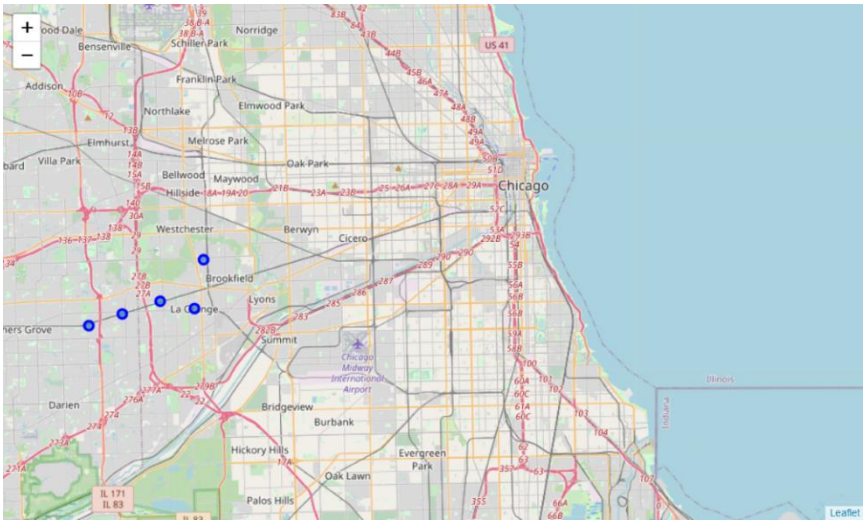


Fig 6: Suburbs near Western Springs Village



Western Springs village is considered as the ideal location to start a new fine dining restaurant due to the following reasons

1. High Median Income of $161,563
2. Location
    1. Western Springs village is located closer to some of the best suburbs like La Grange Park village, Clarendon Hills village, La Grange village, Hinsdale village
    2. These 5 Suburbs has a total population of 68,248
    3. Western Springs village is also located nearer to Tri State Tollway
    4. Hinsdale village is among top 5 richest suburb which is located nearer to Western Springs village
3. 78.4% of the adult population above 25 years holds a bachelor's degree
4. Most of the venues in Western Springs village are located near the railway line. It is highly recommended to open the restaurant in the downtown Western Springs village nearer to the intersection of Rail line and Wolf Road.

Fig 7: Venues in and around Western Springs Village

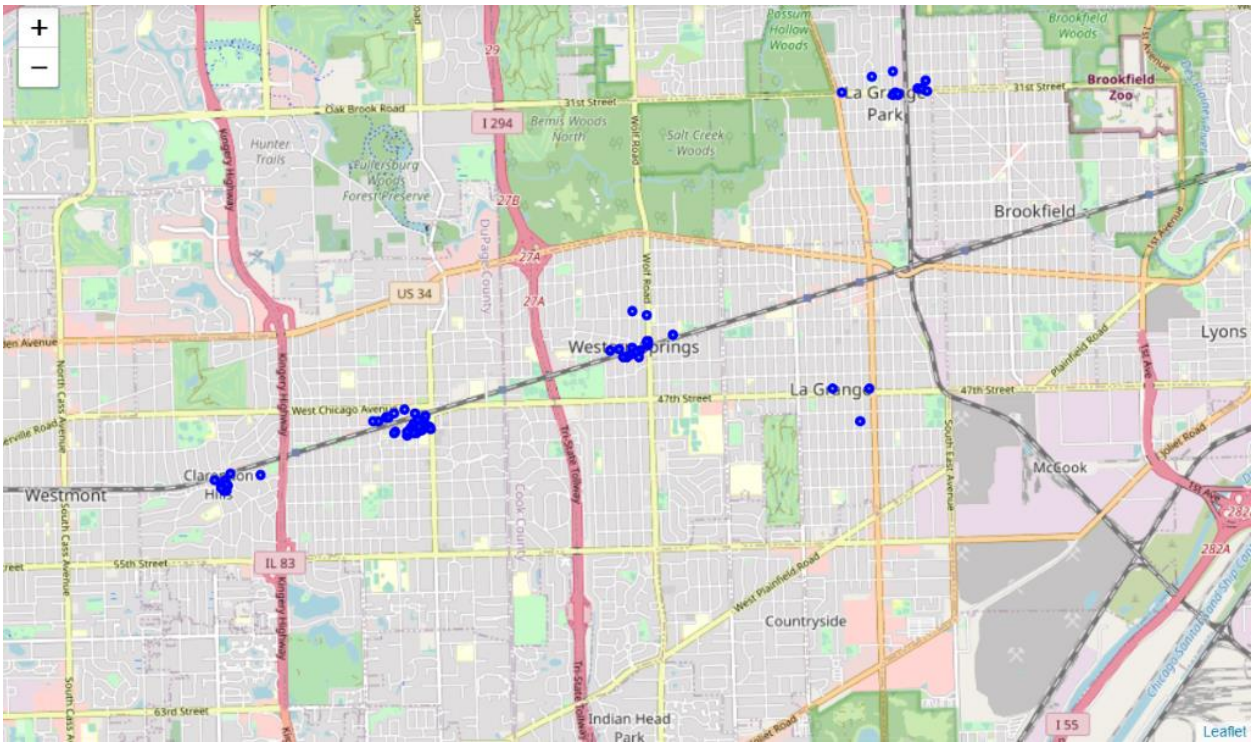
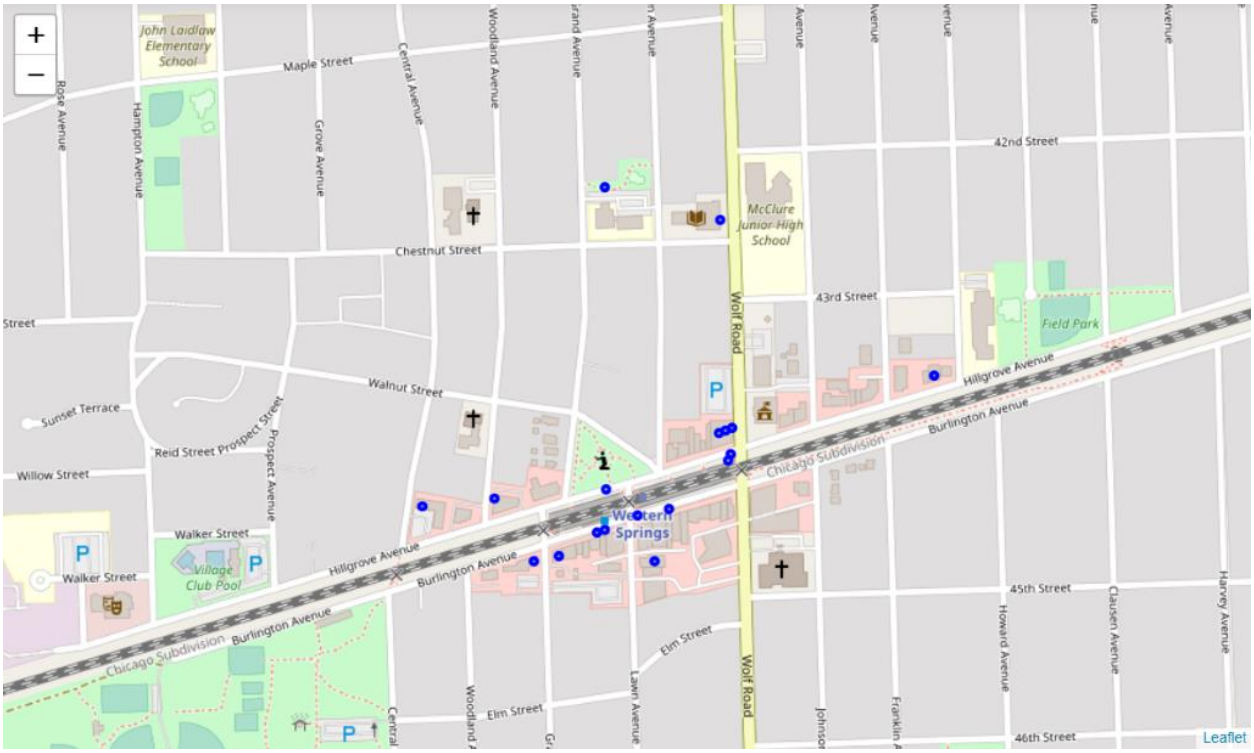


Fig 8: Venues in Western Springs Village



Discussion

Scope of this project was to identity an ideal location for our client to expand their restaurant operation in suburbs of Chicago. It would be interesting to study further parameter available in the census dataset which may lead to different outcome. Parameter like per capita income, poverty can play a vital role in determine the profit and location of the new restaurant. Current, Census dataset can be combined with another dataset like sale in getting better insight of operating a restaurant in that locality.

Conclusion

Even though, the combination of United States Census Bureau data and Foursquare venue data was used for the analysis; combined dataset was not sufficient to determine the prospective location. Standalone dataset from United States Census Bureau website along with geo location coordinates was sufficient to determine the prospective location. Addition information from U.S. Department of Labor can be used to determine the operation cost which would be most valued by the client in predicting the profits.