

Finding a place for a new food venue in Chicago

Alon Bendor

March 2021

Introduction

The following report is part of the capstone project of *IBM Data Science Professional Certificate* in Coursera.

The goal of the project: *“come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of your choice or to come up with a problem that you can use the Foursquare location data to solve.”*

In the project, I selected to explore Foursquare data to find a potential community area in Chicago for opening a new food venue in Chicago.

Background

With a population of 2.7 million (2019) and 57.7 million visitors to the city during 2018, Chicago is a candidate for leisure business opportunities. Before 2020 and the Coronavirus outbreak, Chicago hosted about 200 annual festivals and parades. It has 250 live music venues, more than 200 theaters and other arts and entertainment venues that attract visitors.

Problem

Due to the diversity of food venues in Chicago, it is difficult to find a place to open a new food venue. This project will look for a different approach for finding a place for a new food venue in Chicago.

Idea

In this project, Foursquare data from the city of Chicago will be used to analyze the city's business opportunities for opening a food venue based on the number of arts and entertainment venues and the corresponding number of food venues in Chicago community areas.

The assumption is that arts and entertainment venues attract visitors (both local and tourists) that are potential clients of food venues. Therefore, the analysis will try to find community areas with potential for new food venue business based on the relation between the number of arts and entertainment venues that attract visitors and number of food venues.

Description of the data

Data that will be used

In this project, the following data will be used for the analysis:

Chicago community areas

The community areas' names will be used to extract their geographic coordinates. Chicago neighbourhoods' names will be extracted from Wikipedia using Beautiful Soup HTML parser Python package.

Chicago neighbourhoods' geographic coordinates

The longitude and latitude coordinates will be used to query Foursquare for venues in Chicago. Geographic coordinates will be extracted using Nominatim geocoding Python API.

Food venues in Chicago

Food venues in Chicago will be extracted from Foursquare and will be used to analyze the count of food venues in each community area.

Arts and entertainment venues in Chicago

Arts and entertainment venues in Chicago will be extracted from Foursquare and will be used to analyze the count of arts and entertainment venues in each community area.

Foursquare data explained

Foursquare - a location data platform that helps people discover and share information about businesses and attractions based on location.

Places by Foursquare - a database of venues with over 900 categories, more than 105 million places worldwide and an API that enables location data. The API will be used to extract venues data using geographic coordinates of the community areas.

Venue categories that will be used

Arts & Entertainment – this category includes venues as: theatres, museums, music venues, performing arts venues, sport stadiums, and more.

Food – this category includes venues as: restaurants, coffee shops, and more.

Data that will be extracted from Foursquare:

The data that will be extracted is:

Venue category as: art gallery, museum, coffee shop, Mexican restaurant.

Venue ID – will be used for removing duplicates from the data.

Data collection

Foursquare data was collected based on the longitude and latitude of each community area. Foursquare limits the search results to 50 venues, so latitude and longitude of Chicago neighbourhoods were used for better search resolution.

Venue ID of each venue was used to remove duplicates from the collected data.

Exploratory data analysis

The data contained the following:

- 2434 art venues
- 4385 food venues

The data was grouped by community area, and for each community area, the following data was grouped:

- Number of food venues
- Number of arts & entertainment venues
- Ratio between the number of food venues and the number of arts & entertainment venues.

The purpose is to select the community area which has business potential for opening a new food venue. A community area with low number of food venues compared to the number of arts and entertainment venues has the potential of high visitors' rate with possibly lower number of food venues to answer the needs of the high visitors count.

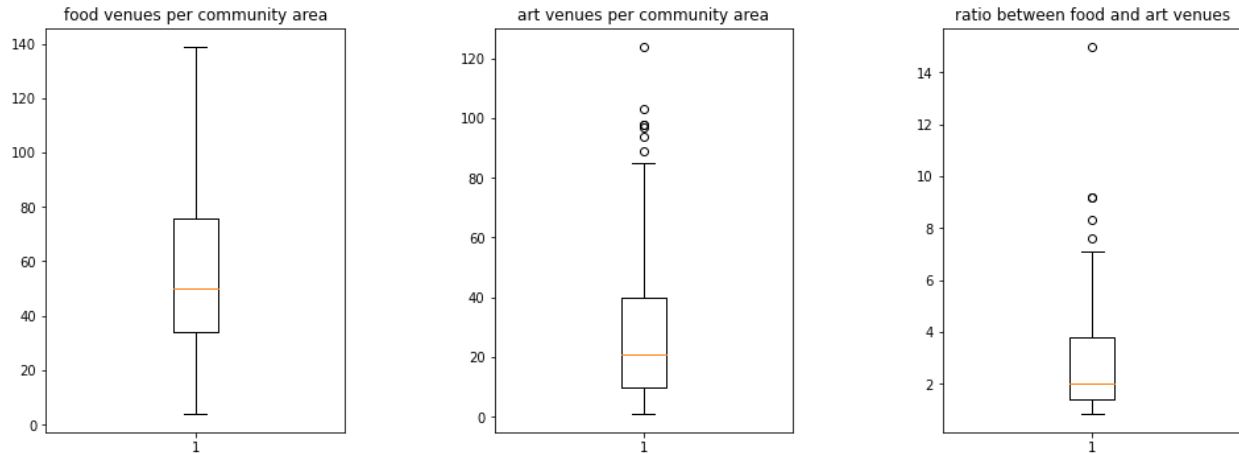
Using descriptive to select the value of the ratio between the number of food venues and the number of arts and entertainment venues:

The mean value of the ratio is 3.04, while 50% of the community areas are with a ratio lower than 2.

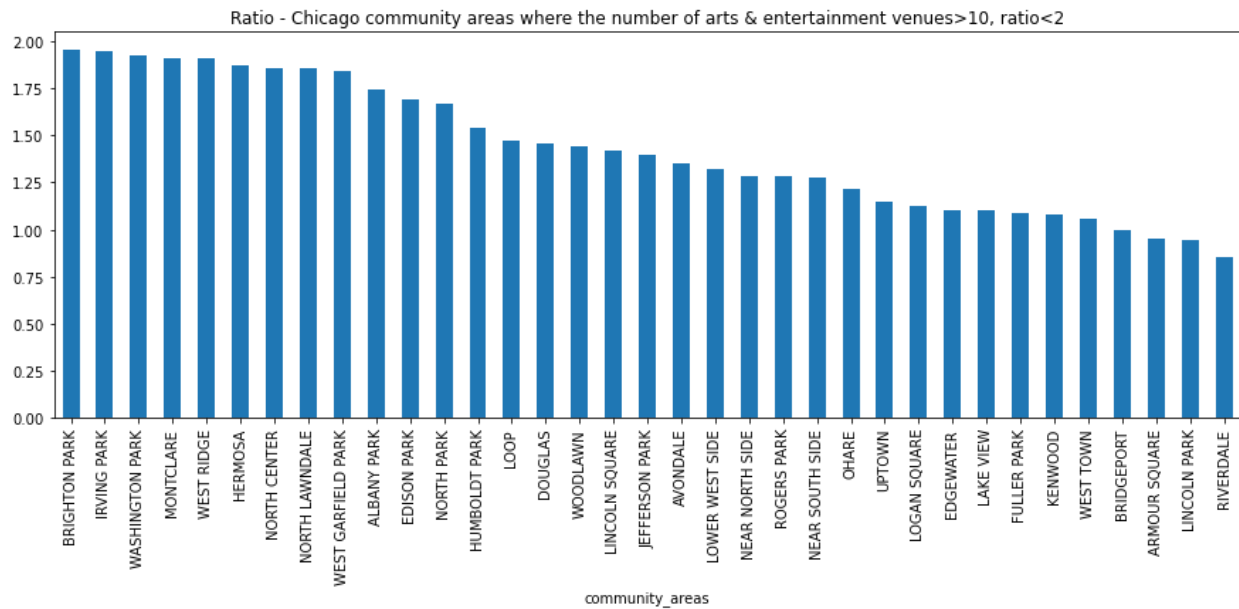
The mean value of the arts and entertainment venues per community area is 31.

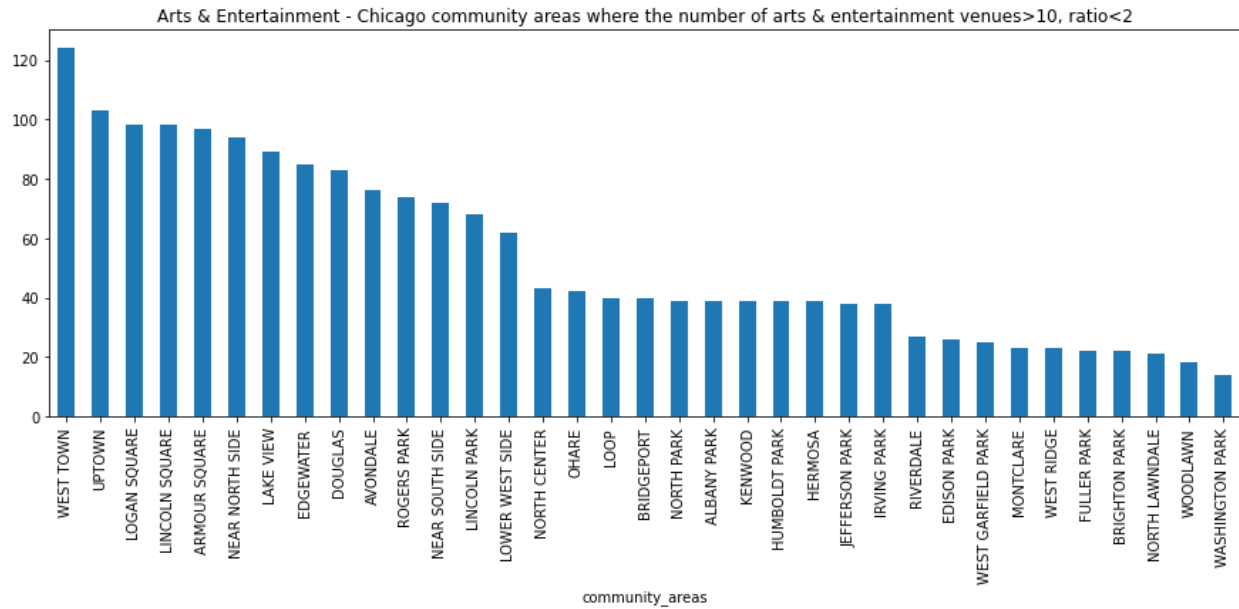
The following table and box plots show the descriptive statistics of the number of food venues, number of arts and entertainment venues of the ratios between them.

	food_count	art_count	ratio
count	77.000000	77.000000	77.000000
mean	56.948052	31.610390	3.044682
std	32.862111	29.978569	2.518193
min	4.000000	1.000000	0.851852
25%	34.000000	10.000000	1.418367
50%	50.000000	21.000000	2.000000
75%	76.000000	40.000000	3.791667
max	139.000000	124.000000	15.000000



Filtering the results by community areas that have a ratio value below 2 and number of arts and entertainment venues above 10 may give a set of community areas with potential for opening a new food business – the ratio below 2 assures that relatively the number of food venues is low compared to the number of arts & entertainment venues, but more than 10 arts & entertainment venues assures a visitors rates of potential customers for the food business.





The following table shows the top 10 community areas with the highest count of arts & entertainment venues and ratio between the number of food venues and the number of arts & entertainment venues below 2.

food_count	art_count	ratio	community_areas
131	124	1.056452	WEST TOWN
118	103	1.145631	UPTOWN
110	98	1.122449	LOGAN SQUARE
139	98	1.418367	LINCOLN SQUARE
92	97	0.948454	ARMOUR SQUARE
121	94	1.287234	NEAR NORTH SIDE
98	89	1.101124	LAKE VIEW
94	85	1.105882	EDGEWATER
121	83	1.457831	DOUGLAS
103	76	1.355263	AVONDALE

Discussion and results:

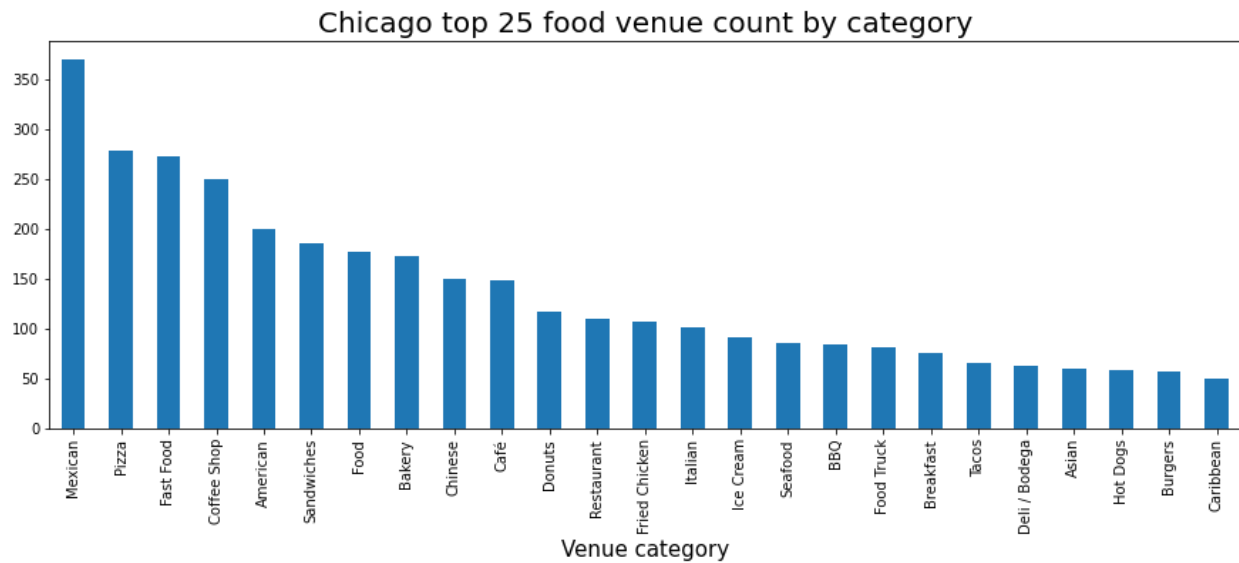
The analysis resulted in a list of community areas with high number of arts & entertainment venues relative to the number of food venues.

Descriptive statistics showed that 50% of the community areas have a ratio below 2, the community areas in the list are with a ratio below 1.5. this shows a potential of high visitors count compared to relatively low count of food venues to answer the desire for visiting a food venue along with an entertainment activity.

Further analysis is needed for better understanding the business potential of opening a new food venue. Suggested analysis for further investigation:

- Analysis of the potential number of visitors and the schedule of visits in the arts & entertainment venues.
- Analysis of the food venues competitors in the relevant community areas.
- Analysis of the food venues types that are preferred in Chicago.
- Financial analysis that considers the financial aspects of the food business, as:
 - o Rental and city taxes
 - o Investment
 - o Potential revenue stream

From the data that was collected, the following are the top 25 categories of food venues in Chicago, this information can be used as part of further investigation.



foodCategory	count
Mexican	370
Pizza	278
Fast Food	273
Coffee Shop	250
American	200
Sandwiches	185
Food	177
Bakery	173
Chinese	149
Café	148
Donuts	117
Restaurant	110
Fried Chicken	107
Italian	101
Ice Cream	91
Seafood	86
BBQ	84
Food Truck	81
Breakfast	75
Tacos	65
Deli / Bodega	62
Asian	59
Hot Dogs	58
Burgers	57
Caribbean	49

Conclusion:

The data that was extracted from Foursquare was sufficient for implementing the idea of selecting community areas that are potential for opening a new food venue in a Chicago community area.

However, the data is a preliminary step for continuing investigation before selecting a location for opening a food venue.

Data sources:

1. Chicago Release 2018 Annual Tourism Report - https://www.chicago.gov/city/en/depts/mayor/press_room/press_releases/2019/april/2018AnnualTourismReport.html
2. Chicago population - <https://www.census.gov/quickfacts/fact/table/chicagocityillinois,US/PST045219>
3. Chicago community areas – https://en.wikipedia.org/wiki/Community_areas_in_Chicago
4. Chicago facts - <https://www.chicago.gov/city/en/about/facts.html>
5. Foursquare Places API - <https://foursquare.com/>