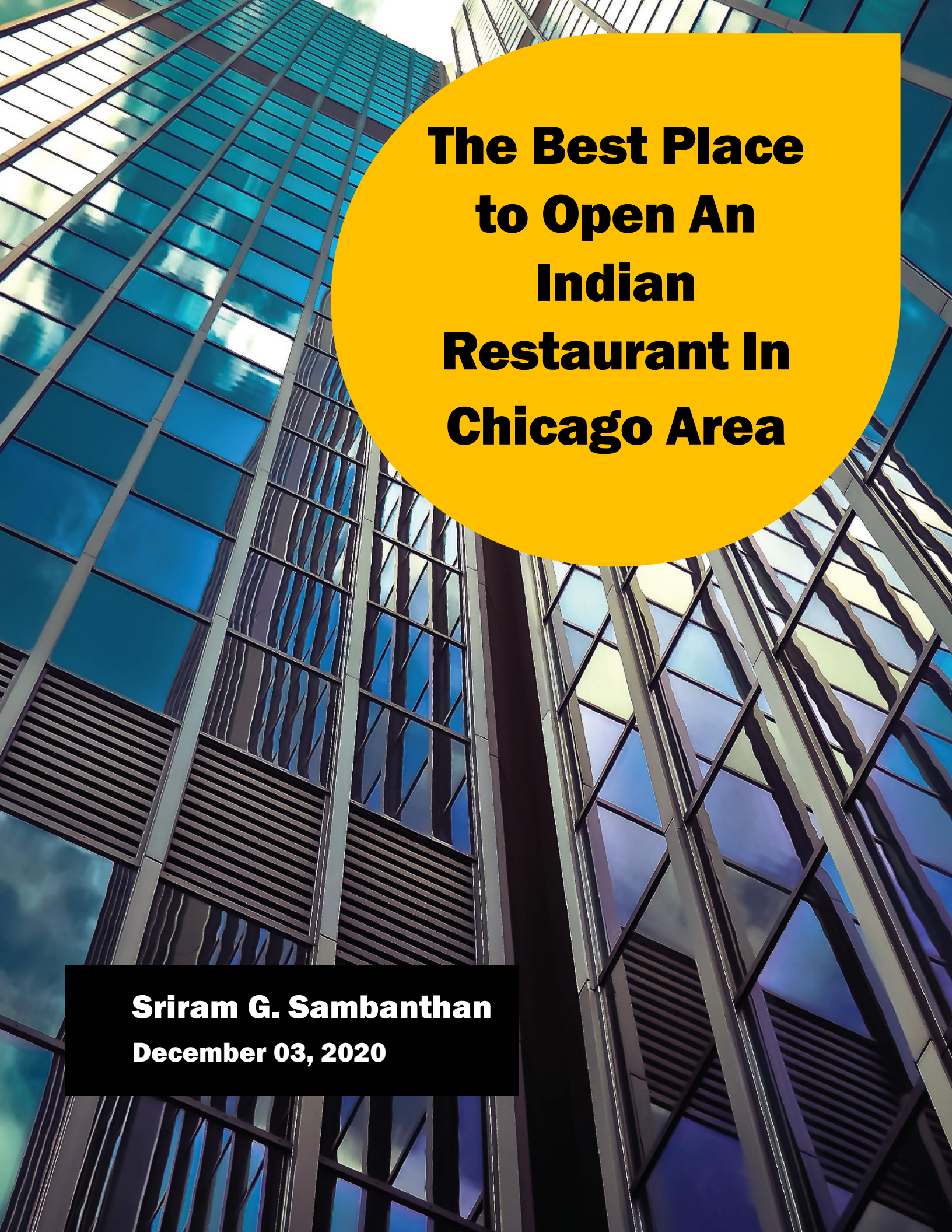


The Best Place to Open An Indian Restaurant In Chicago Area

Sriram G. Sambanthan

December 03, 2020

Contents

With all the skills and knowledge gathered from the course, I take this opportunity to showcase my level of understanding.

Though one can use the data science to a variety of problems and scenarios, for this capstone I take the scenario of finding an apt place around the Greater Chicago Area to open an Indian Restaurant. In this case study, I use the Foursquare API for obtaining the location data and comparing them. I hope this study would open up new findings.

I would like to thank the instructors of the specialisation Dr. Joseph Santarcangelo and Dr. Alex Aklson for extending their knowledge for the courses. Last but not the least, I thank my peers for their support.

Sriram.

1. Introduction

1.1 Problem Description

My Friend, who runs a chain of restaurant all over the US wants to open a new restaurant in the Greater Chicago Area. She has started her preliminary works for the process and now, she needs to find an optimum place in the Chicago area which would attract more customers to her restaurant. Being a data scientist, she wants me to look into this matter and suggest her the best place for her restaurant.

2. Source of Data

In order to start the process, we need to gather the data upon which we can apply our analytics. In this case, I am going to use the Foursquare API to gather the data. The process flow would be,

a. Gathering all the restaurant data around the Chicago Area.

b. Creating the DataFrame.

c. Refining the obtained data.

2.1 Gathering Restaurant Data around the Chicago Area

Using the Foursquare API, the location data are obtained in the JSON format which would be formatted later.


```
{'meta': {'code': 200, 'requestId': '5fc8c42596525d382262974f'},
 'notifications': [{'type': 'notificationTray', 'item': {'unreadCount': 0}}],
 'response': {'venues': [{'id': '4b9ad8c5f964a520ccdb35e3',
    'name': 'Plymouth Restaurant & Bar',
    'location': {'address': '327 S Plymouth Ct',
     'crossStreet': 'btwn W Jackson Blvd & W Van Buren St',
     'lat': 41.87713751627466,
     'lng': -87.62796900098705,
     'labeledLatLngs': [{'label': 'display',
       'lat': 41.87713751627466,
       'lng': -87.62796900098705},
      {'label': 'entrance', 'lat': 41.877473, 'lng': -87.628615}],
     'distance': 342,
     'postalCode': '60604',
     'cc': 'US',
     'city': 'Chicago',
     'state': 'IL',
     'country': 'United States',
     'formattedAddress': ['327 S Plymouth Ct (btwn W Jackson Blvd & W Van Buren St)',
      'Chicago, IL 60604',
      'United States']},
```


1. Gathering restaurant data

2.2 Creating the DataFrame

By using the json_normalize function, the data from the JSON are being transferred into the dataframe using the pandas library.

	id	name	categories	referralId	hasPerk	location.address	location.crossStreet	location.lat	location.lng
0	4b9ad8c5f964a520ccdb35e3	Plymouth Restaurant & Bar	[{'id': '4bf58dd8d48988d116941735', 'name': 'B...	v-1606992933	False	327 S Plymouth Ct	btwn W Jackson Blvd & W Van Buren St	41.877138	-87.627969
1	40b28c80f964a52037fb1ee3	Exchequer Restaurant & Pub	[{'id': '4bf58dd8d48988d11b941735', 'name': 'P...	v-1606992933	False	226 S Wabash Ave	NaN	41.878819	-87.625986
2	4b2bedd4f964a520e6bd24e3	The Berghoff Restaurant	[{'id': '4bf58dd8d48988d10d941735', 'name': 'G...	v-1606992933	False	17 W Adams St	btwn S Dearborn St & S State St	41.879334	-87.628368
3	5d8136d5adf8620008fd8200	Opus Restaurant & Café	[{'id': '4bf58dd8d48988d16d941735', 'name': 'C...	v-1606992933	False	65 E Adams St	Wabash	41.879321	-87.625508
4	4b295853f964a520489d24e3	65 Chinese Restaurant	[{'id': '4bf58dd8d48988d145941735', 'name': 'C...	v-1606992933	False	201 W Madison St	btwn Wells & Franklin St	41.881909	-87.634165

2. Creating the DataFrame

2.3 Refining the obtained data

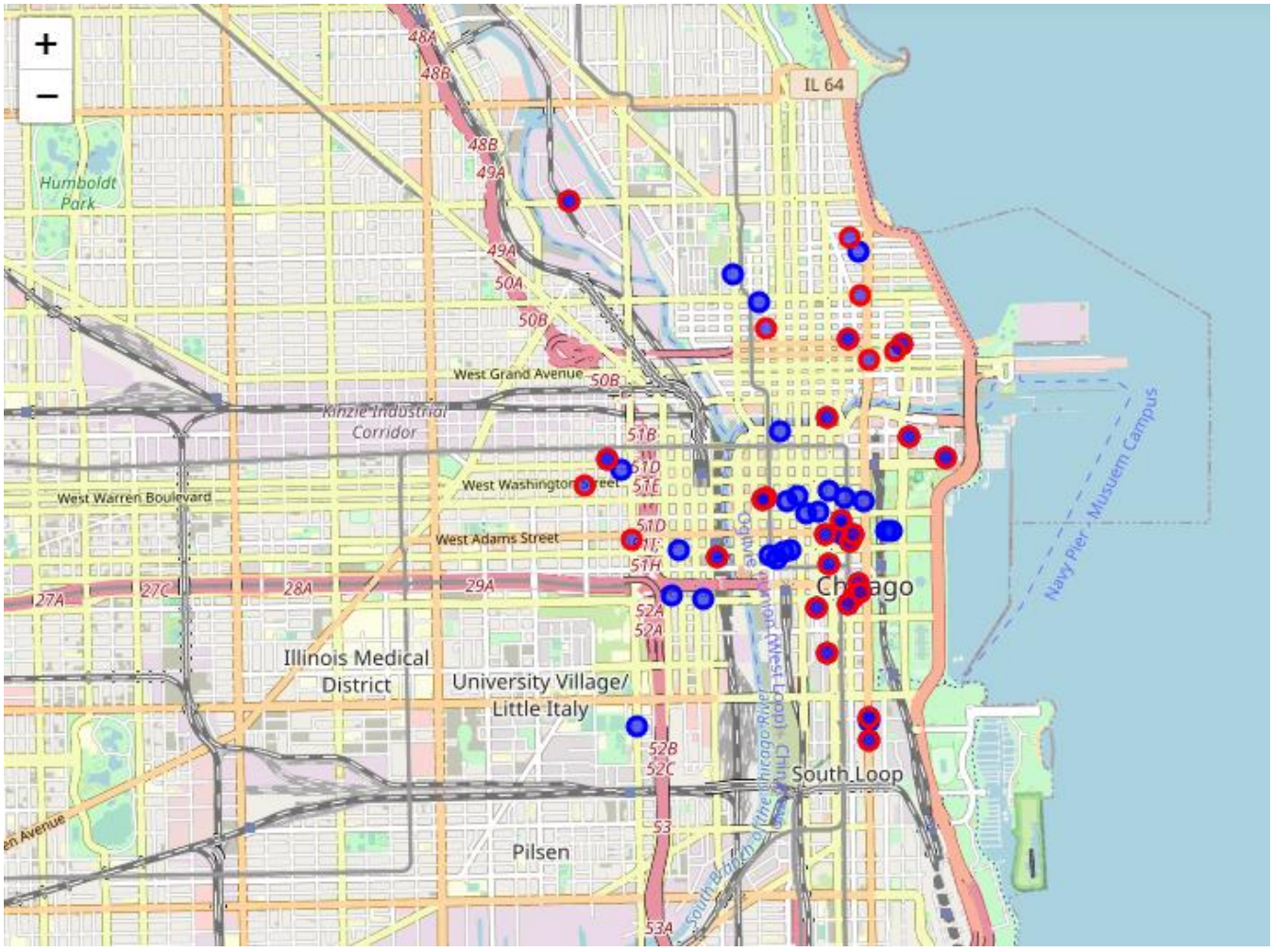
The dataframe contains many columns which often consist of irrelevant information. In this step, the irrelevant data are removed and the refined set of dataframe was created.

	id	name	location.address	location.lat	location.lng	location.distance	location.postalCode	location.cc	location.city	locati
0	4b9ad8c5f964a520ccdb35e3	Plymouth Restaurant & Bar	327 S Plymouth Ct	41.877138	-87.627969	342	60604	US	Chicago	IL
1	40b28c80f964a52037fb1ee3	Exchequer Restaurant & Pub	226 S Wabash Ave	41.878819	-87.625986	385	60604	US	Chicago	IL
2	4b2bedd4f964a520e6bd24e3	The Berghoff Restaurant	17 W Adams St	41.879334	-87.628368	532	60603	US	Chicago	IL
3	5d8136d5adf8620008fd8200	Opus Restaurant & Café	65 E Adams St	41.879321	-87.625508	428	60603	US	Chicago	IL
4	4b295853f964a520489d24e3	65 Chinese Restaurant	201 W Madison St	41.881909	-87.634165	1073	60606	US	Chicago	IL

3. Refined DataFrame

3. Methodology

By using the data in the pandas dataframe, the restaurant info are plotted on the maps using the folium library. The ordinary restaurants are plotted with the red circles and the Indian restaurants are plotted with the blue circles. The maps can be found below.



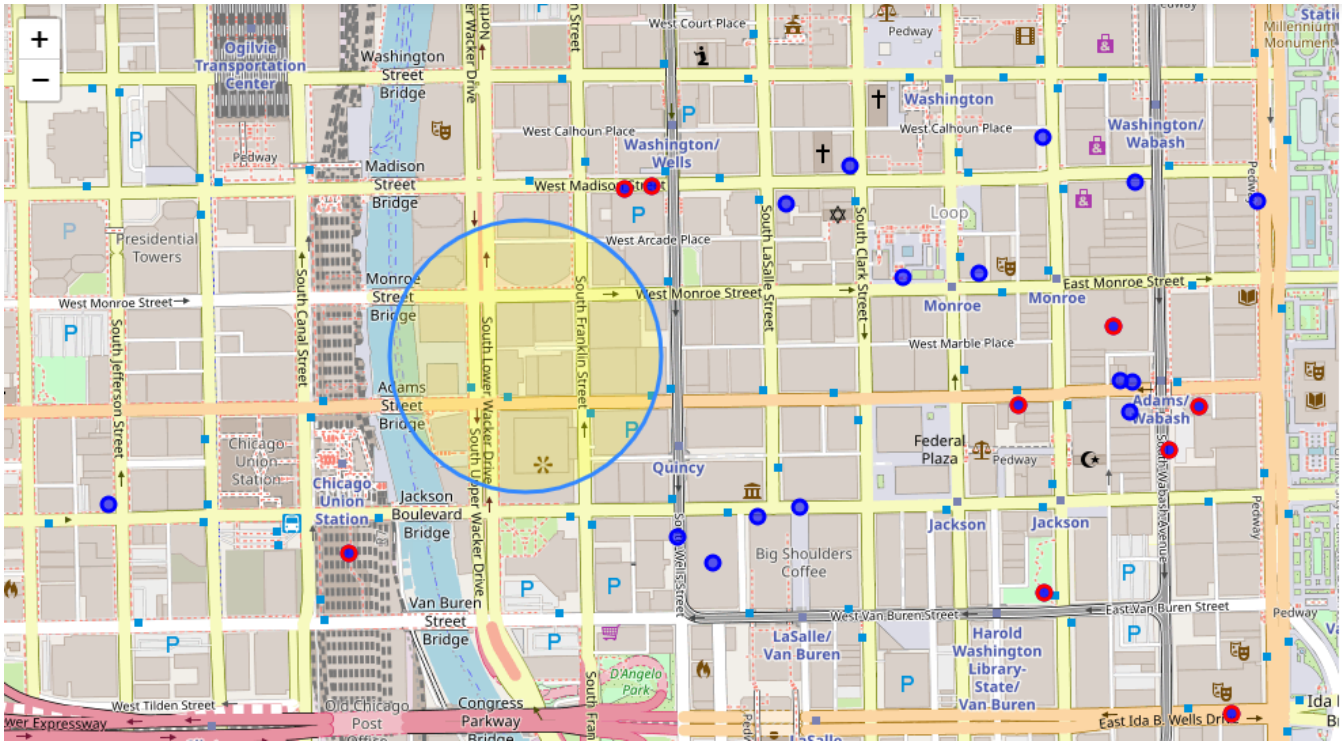
4. Restaurant's Geospatial Data

4. Results & Discussion

Based on the observed data, the restaurants are more common only in the Central Chicago area. If we go further to the west and to the south, the restaurants are much more scattered. Based on this, one can tell that the restaurants are much more dynamic around the central region of the city. So, choosing a site within the proximity of the central region would draw more footfall. The central region is also the place for high number of Asian and Indian restaurants. This makes us to choose a place which is free from Asian and Indian cuisines.

5. Conclusion

The Chicago central station is at the heart of the city. The number of Indian restaurants are very less in that region. So, starting the restaurant near the Postal Region of 60606 would serve us the most benefit. The suitable area can be found below.



5. Optimum space for new restaurant

The area circled by the yellow colour is the optimum space for opening the restaurant.