Predicting the location to start a new lunch-restaurant around Parsippany Township, New Jersey

Savyata Regmi

June 16, 2020

1. Introduction

1.1 Background

Parsippany-Troy Hills Township as known as Parsippany is one of the towns in Morris County New Jersey, United States. The name Parsippany comes from the Lenape Native American sub-tribe, which comes from the word parsipanong, which means "the place where the river winds through the valley". Since 2006, Parsippany-Troy Hills has been consistently recognized by Money magazine as one of the Best Places to Live in the United States. It is only 29 miles away from New York City.
The town is working destination of thousands of people as diversified manufacturing, polytechnic and Tech companies and tax firms are located here. Tiffany and Co., IBM, Deloitte, HNTB, ADP, Langan, Evonik, Allergan, B&G Foods, GAF Materials Corporation, IQVIA, Sun Chemical, Zoetis, Kings Food Markets, Lexus and Toyota Financial Services, American Financial Resources, Wyndham Worldwide and PNY Technologies, a manufacturer of computer memory devices, etc. are some popular working destination of people which are located in Parsippany-Troy Hills. The U.S. operations of Cadbury Adams, Reckitt Benckiser, Ricola and Safilo are located here. Cendant Corporation moved its headquarters to Parsippany-Troy Hills in 2001; in 2006 Cendant separated into several different companies, including Avis Budget Group, parent company of Avis Rent a Car System and Budget Rent a Car. This town is working destination for 1000s of people from New Jersey, New York and Pennsylvania states who commute every day for work.

1.2 Discussion of the business problem

This town is a heart for working class-people from New Jersey and its neighboring states, New York, and Pennsylvania. There are 100s of restaurants in the town that feeds 1000s of those working-class people quick and healthy breakfast-lunch and even dinner. Today I am going to explore the restaurants which is within 4 miles of center of the town which typically serves breakfast and lunch. New Jersey is also home of "Diners" which are classical American restaurants who serves delicious breakfast-lunch-dinner in a pleasant environment (some of them are open for 24 hours operation).

The main purposes of this project are:

- To explore the best place for lunch within 4 miles of the search-point location and find out what categories of restaurants are available.

- To predict the best location to open a new restaurant within that radius which can be easily accessible from the heart of the town where there are majority of offices and companies.

1.3 Audience who would be interested in this project

This project is targeted to the following group of people:

- This project could catch the attention from working-class people who is not local to the town and yet is seeking for the best place to grab breakfast and lunch.
- This project could also benefit the people who have newly moved to live in the town and have no idea on where to go to enjoy the delicious breakfast, lunch and dinner.
- Small entrepreneurs who are willing to open a new restaurant in the town could also be benefitted from this project because I am going to predict the best spot to start a new food business within the covered radius.

2. Data Section

2.1 Data Sources

I utilized the **Foursquare API** to explore the restaurants by using the keywords "diner" and "restaurant" for the search query. I designed the limit as **500 venues** and the **radius 6000 meters** (3.7 miles) from their given latitude and longitude information. Then I sent the GET request of my search query to obtain the **JSON file** and converted the JSON file into pandas data frame to make the data readable. Here is the list of first few rows of the data frame:

```
diners_df
```

7]:

	name	categories	distance	lat	lng	id
0	Empire Diner	Diner	1585	40.863335	-74.393717	4c52ec369d642d7f9030d4df
1	The Spa Diner Resraurant	Diner	2976	40.876426	-74.382928	4d50351f9d4937045e40bf39
2	Brookside Diner	Diner	4106	40.825175	-74.419509	4b6a199bf964a52097c72be3
3	Whippany Diner	Diner	4566	40.820793	-74.407954	4c0e46f6336220a10ae6cb77
4	Rockaway Townsquare Diner	Diner	1344	40.856714	-74.426975	4ebdc6c2e5fae164638f5a46
5	Paul's Family Diner	Diner	4714	40.883985	-74.460021	4c1c24178b3aa5931b79985f
6	Ridgedale Diner	Diner	6025	40.827605	-74.356820	4bd1a16fa8b3a593ad04665f
7	Whippany Diner	Diner	4583	40.820630	-74.408104	4f94b48be4b0f639262f3f26
8	West Side Diner	Diner	5332	40.887753	-74.465561	4c3e377d0596c92897fb8378
9	Frank's Diner	Diner	4331	40.823037	-74.418554	4ca736adb0b8236a43aba6e6
10	Ajp Boonton Diner	Food	4962	40.906250	-74.413170	4f3249f319836c91c7c8fc58

2.2 Data Cleaning

When I observed the data frame, I found out that one restaurant was repeated twice so I removed it from the data frame. I also dropped 3 more rows from the data frame which had a restaurant program office, restaurant depot, and Sheraton Hotel, the search-point location.

2.3 Feature Selection

For this analysis, the restaurants and diners are chosen from within 4 miles of radius from the core location. Later, we will be sorting and find the top 5 closest restaurants and top five best rated restaurants. Since, this project is targeted for especially working-class people who work in companies and firm within that location who are seeking for good and decent restaurant where they can get quick and healthy lunch. Hence, we have only the list of classical diners, pizzerias, and cuisine restaurants in our analysis.

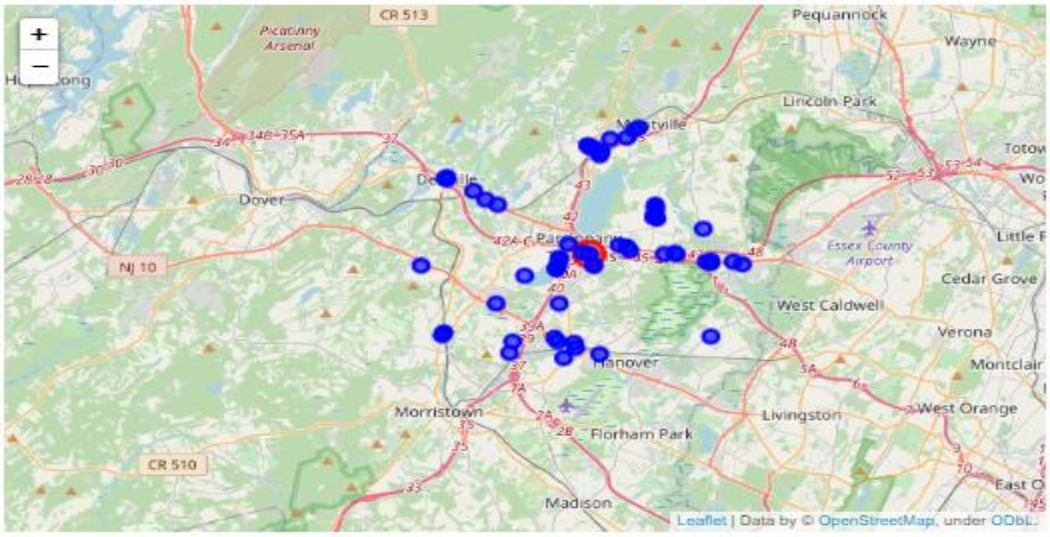
Note: Fast-food chain restaurants are excluded from the analysis.

3. Methodology

In this section I performed some data analysis and EDA to find insight from data. We will try to understand the current stats of all given data. We will perform k means clustering to find the spot to open a new restaurant.

3.1 Exploratory Data Analysis

I used the python folium library to visualize geographic details of the venues that I had, and I generated a map by passing the latitude and longitude of my core location to spot the nearby restaurants. The red mark is the core location, or our search point and the blue mark is for the restaurants nearby, which is given below:



For finding the best restaurants closest to the core location, at first I sorted the restaurants according to the closest distance from the search point and got the following list of restaurants:

```
In [12]: distance_diners_df = diners_df[['name','categories','distance', 'id']].sort_values('distance')
         distance_diners_df
```

Out[12]:

	name	categories	distance	id
4	Annabella's Pizza & Restaurant	Pizza Place	553	4bb15136f964a520e18b3ce3
5	Jumbo LII Chinese Restaurant	Food	564	4f324eff19836c91c7caf3ba
6	Pearl Banquet & Restaurant	Indian Restaurant	922	59b09d17018cbb060fec80ea
8	Mizu Japanese Restaurant	Japanese Restaurant	1194	4f32693219836c91c7d52850
1	Jimmy's Pizzeria & Family Restaurant	Pizza Place	1217	57638a44498e91e51205a67c
9	Peking Restaurant	Food	1220	4f44d92c19836ed00196f43e
4	Rockaway Townsquare Diner	Diner	1344	4ebdc6c2e5fae164638f5a46
23	Carrot Indo Chinese	Chinese Restaurant	1409	59ef7d64c36588289144acfc
3	Marino's Pizzeria and Restaurant	Pizza Place	1460	55034e0b498e843f6363d1f2
7	Baldwin Pizzeria & Restaurant	Pizza Place	1475	4bb52d536edc76b0e9622f1c
18	Ann Farrell's Restaurant	Food	2559	4f32790619836c91c7db2fc7
19	Una Faccia Pizzeria and Restaurant	Pizza Place	2690	4f32453e19836c91c7c71a8b

From the above data frame, I found out the following facts:

- Five closest restaurants are Annabella's Pizza & Restaurant, Jumbo LII Chinese Restaurant, Pearl Banquet & Restaurant, Mizu Japanese Restaurant, and Jimmy's Pizzeria & Family Restaurant.
- Tony's Pizza and Restaurant is the farthest from the core location.

I also calculated the average location between Diners and core location which is 4051 meters or approximately 2 miles.

I explored the ratings of the restaurants and got the given result after dropping the unrated restaurants:

Sorting the Venues in descending order by Rating

```
In [112]: top_5= only_rated_tips[['Venue','Rating']].sort_values('Rating', ascending= False)
          top_5
```

Out[112]:

	Venue	Rating
25	Guiseppe's Brick Oven Pizza and Restaurant	9
17	Perkins Restaurant and Bakery	8.2
35	Fasıl Mediterranean Restaurant	8
49	Tony's Pizza & Family Restaurant	8
6	Pearl Banquet & Restaurant	7.9
4	Annabella's Pizza & Restaurant	7.7
11	Keo Ku Restaurant	7.7
5	Paul's Family Diner	7.7
40	Thai Ping Restaurant	7.7
13	Il Villagio Restaurant	7.5
5	Jumbo LII Chinese Restaurant	7.4
9	Frank's Diner	7.4

The top five best-rated restaurants are:

Guiseppe's Brick Oven Pizza and Restaurant

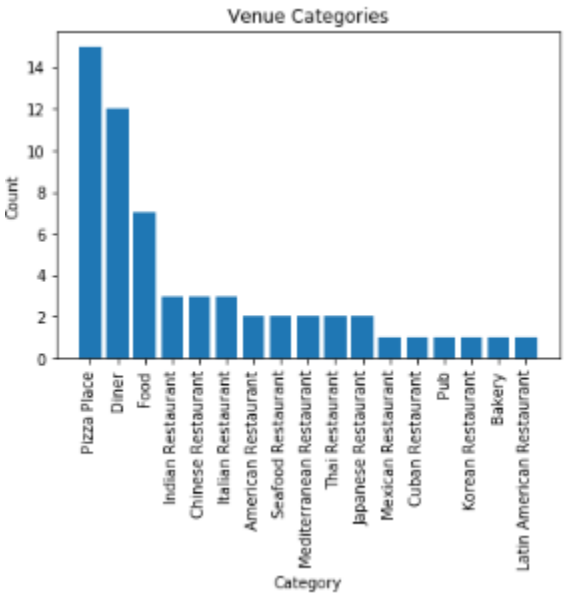
Perkins Restaurant and Bakery

Fasıl Mediterranean Restaurant

Tony's Pizza & Family Restaurant

Pearl Banquet & Restaurant

Counting the number of restaurants that falls under a different category

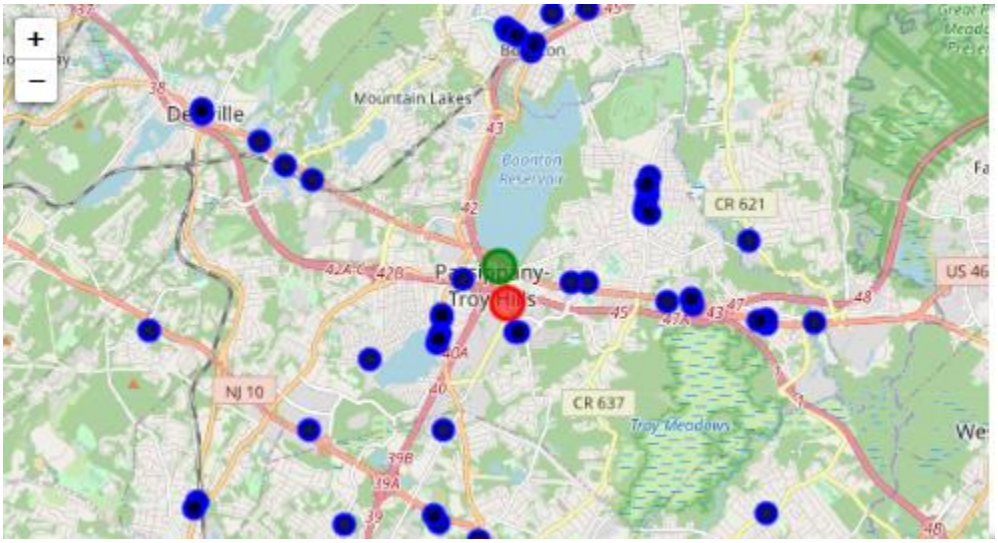


According to the obtained result, there are 15 pizza places and 12 diners within 4 miles, and the rest of all are cuisine restaurants and other food places!

3.2 Predictive Modeling

I used kmean clustering to predict the location to open a new restaurant. I set the number of clusters to 5 and ran kmean algorithm by partitioning the observations into 5clusters and found the center of all clusters and midpoint of all venues to get the longitude and latitude of the spot which will be ideal to open a new restaurant. From the result, the latitude and longitude were 40.86756915210712 and -74.41410530991203, respectively.

The map below shows the core location “red dot”, the restaurants nearby “ blue dots” and the predicted location to start a new food business “green dot”.



My predicted location and core location are very close to each other which is expected. As this lies in the center of all nearby restaurants, the predicted one almost matched with the core.

4. Results

4.1 Top 5 Closest Venues:

- Annabella's Pizza & Restaurant
- Jumbo LII Chinese Restaurant
- Pearl Banquet & Restaurant
- Mizu Japanese Restaurant
- Jimmy’s Pizzeria & Family Restaurant

The average distance between all the restaurants and core location i.e., Sheraton Hotel is approximately 2 miles

4.2 Top Rated Venues:

- Guiseppe's Brick Oven Pizza and Restaurant
- Perkins Restaurant and Bakery
- Fasıl Mediterranean Restaurant
- Tony's Pizza & Family Restaurant
- Pearl Banquet & Restaurant

Here, Pearl Banquet and Restaurant (Category: Indian Restaurant) is the third closest restaurant which is only 900 meters away from the core location and it has rating 7.9/10 which also falls under 5 top-rated restaurants. Therefore, Pearl Banquet and Restaurant could be the best place nearby.

4.3 Most common Category of Restaurants:

There are 15 pizza places and 12 diners within 4 miles and rest of all are cuisine restaurants and other food places!

4.4 Location for new restaurant:

My predicted location and core location are very close to each other which is expected. As this has central attraction, the predicted one almost matched with the core.

5. Discussion

. We could figure out from the above map that the predicted spot is pointed/clustered on the that location as it was in the center of all nearby the center from all nearby companies and workplaces.

We figured out and clustered the venues in different categories with the help of KMean Clustering. The final output was very adjacent to the core location. This proves the accurate spotting of our predicted algorithm.

Despite of the findings, there were some lack in data. Ratings were missing for most of the venues. Also, when I compared foursquare data with google map ,I could see there were many hotels and venues found missing in foursquare.

6. Conclusion

The top-rated restaurants could be the best for the working-class people who work nearby to grab lunch every day as it is easily accessible from their office within 2 miles location.

Entrepreneurs who are interested to start a new food service business could set up a restaurant on a predicted spot. This will bring revenue automatically as the location is spotted in the area where 1000s of people commute to and work all day.