

Data Science Capstone Project Report

Introduction:

In order to find the solution for real world problems, data science is the best field to do the data analysis and come to a conclusion. Neighborhood location analysis helps finding best solution for the entrepreneurs who want to start new restaurant business .

Business Problem:

New entrepreneurs who want to start a restaurant business find it difficult to find a best suitable location where there are more potential customers to serve and earn profits.

Problem statement:

Find the best location in Puerto Rico State, USA to start a Burger point.

Methodology.

We need to understand the neighborhood in order to find the best location for starting new Burger point. We should find a location where there are more schools, universities and offices. Utilized Foursquare API call to do the neighborhood analysis.

We will analyze the data to find the location where more offices , schools and universities.

Used Forsquare API calls to perform this analysis.

Data Section:

Data for the analysis is taken from the below URL

https://zipcodedownload.com/lookup/Puerto_Rico

It has Zip code, City Name, State Name columns and links to find the latitude and longitude for each Zip code. I have pulled all these details and saved it in csv file

It has the below columns.

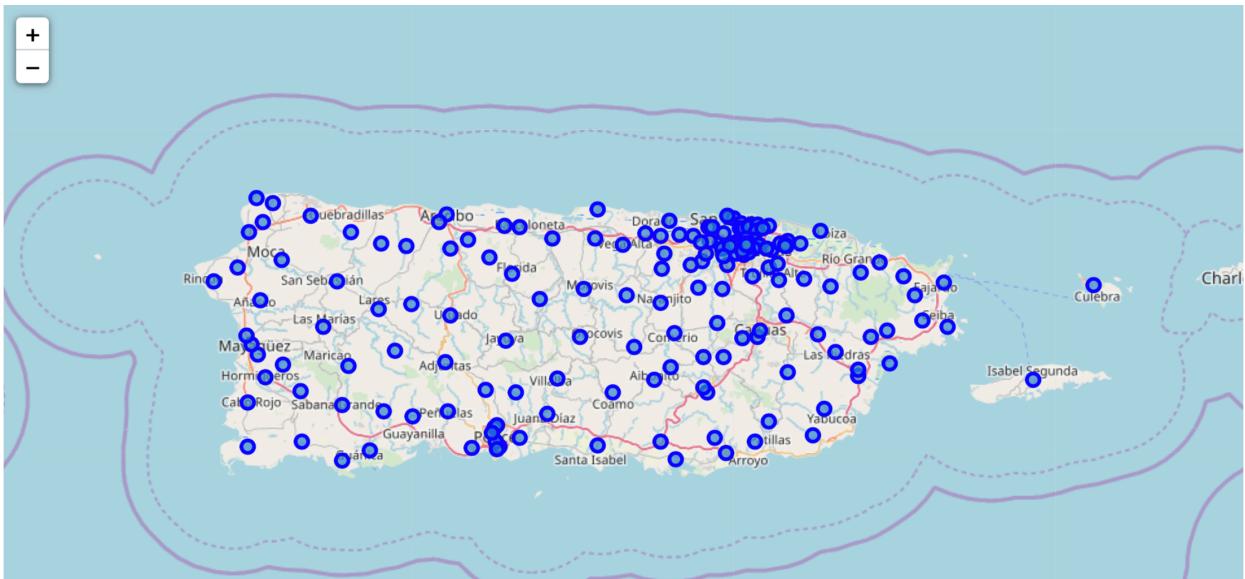
```
df.columns
```

```
Index(['ZIP', 'City Name', 'County Name', 'State Name', 'LAT', 'LNG'], dtype='object')
```

Sample data

	ZIP	City Name	County Name	State Name	LAT	LNG
0	00601	Adjuntas	Adjuntas	Puerto Rico	18.165950	-66.723630
1	00602	Aguada	Aguada	Puerto Rico	18.361945	-67.175597
2	00603	Aguadilla	Aguadilla	Puerto Rico	18.455183	-67.119887
3	00604	Aguadilla	Aguadilla	Puerto Rico	18.505290	-67.135900
4	00605	Aguadilla	Aguadilla	Puerto Rico	18.436150	-67.151340
5	00606	Maricao	Maricao	Puerto Rico	18.158345	-66.932911
6	00610	Anasco	Anasco	Puerto Rico	18.295366	-67.125135
7	00611	Angeles	Utuado	Puerto Rico	18.287720	-66.797580
8	00612	Arecibo	Arecibo	Puerto Rico	18.402253	-66.711397
9	00613	Arecibo	Arecibo	Puerto Rico	18.472740	-66.719280
10	00614	Arecibo	Arecibo	Puerto Rico	18.456900	-66.735890

Map of Neighborhood Zip codes at Puerto Rico(PR)



Results:

Burger points in PR

There are total 216 Burger point venues

```
venues_at_zipcode = getNearbyVenues(names=df['ZIP'], latitudes=df['LAT'], longitudes=df['LNG'], radius=1000,categoryIds=1)
venues_at_zipcode.head()
```

Localidad	Localidad	Latitude	Localidad	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	00601	18.165950		-66.723630	Burger King	18.169486	-66.726676	Fast Food Restaurant
1	00605	18.436150		-67.151340	Joe Spud's	18.439296	-67.147833	Burger Joint
2	00605	18.436150		-67.151340	Wendy's Aguadilla #2	18.443918	-67.146470	Burger Joint
3	00613	18.472740		-66.719280	Burger King arecibo	18.469993	-66.722496	Burger Joint
4	00623	18.083361		-67.153897	Wiliche	18.086051	-67.145678	Burger Joint

```
venues_at_zipcode.shape
```

```
(216, 7)
```

High Schools

There are 149 High schools in PR

High schools

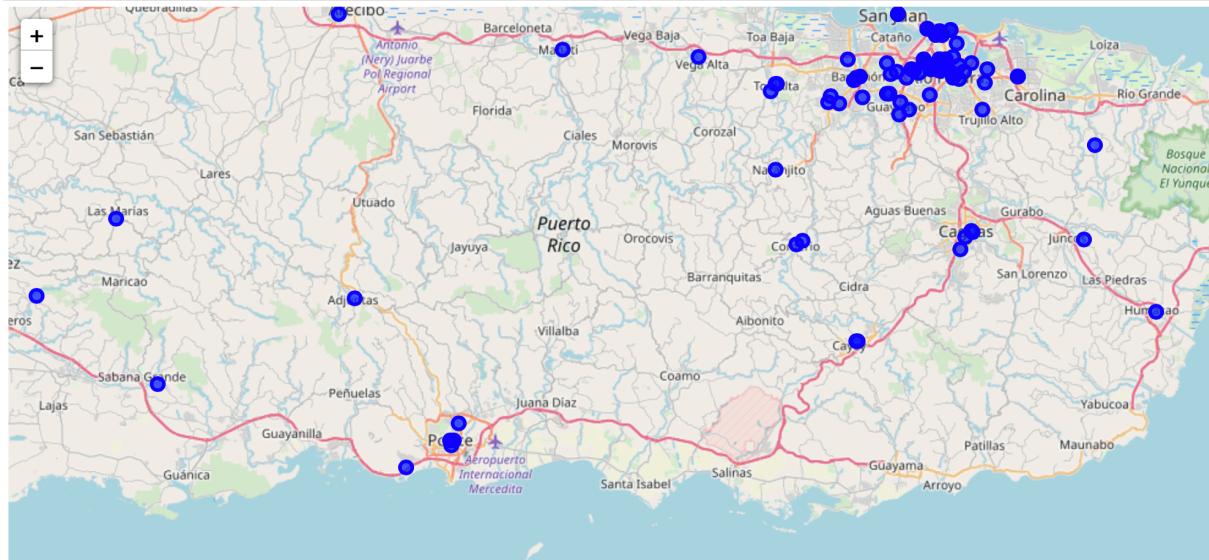
```
: zipcode_venues_highschools = getNearbyVenues(names=df['ZIP'], latitudes=df['LAT'], longitudes=df['LNG'], radius=1000, categoryI  
zipcode_venues_highschools.head()
```

	Localidad	Localidad	Latitude	Localidad	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	00601		18.165950		-66.723630	Escuela José Emilio Lugo	18.162628	-66.721026	High School
1	00604		18.505290		-67.135900	Ramey School	18.498053	-67.139784	High School
2	00614		18.456900		-66.735890	Abelardo Martinez Otero	18.465688	-66.739285	High School
3	00636		18.162130		-67.077930	Laura Mercado High School	18.165284	-67.078158	High School
4	00637		18.076713		-66.947389	Missionary Christian Academy	18.071133	-66.942162	High School

```
: zipcode_venues_highschools.shape
```

```
: (149, 7)
```

```
map_zipcode_schools = folium.Map(location=[latitude, longitude], zoom_start=12)  
addToMap(zipcode_venues_highschools, 'blue', map_zipcode_schools)  
map_zipcode_schools
```



Universities:

There are 187 universities

List of Universities

```
zipcode_venues_uni = getNearbyVenues(names=df['ZIP'], latitudes=df['LAT'], longitudes=df['LNG'], radius=1000, categoryI  
zipcode_venues_uni.head()
```

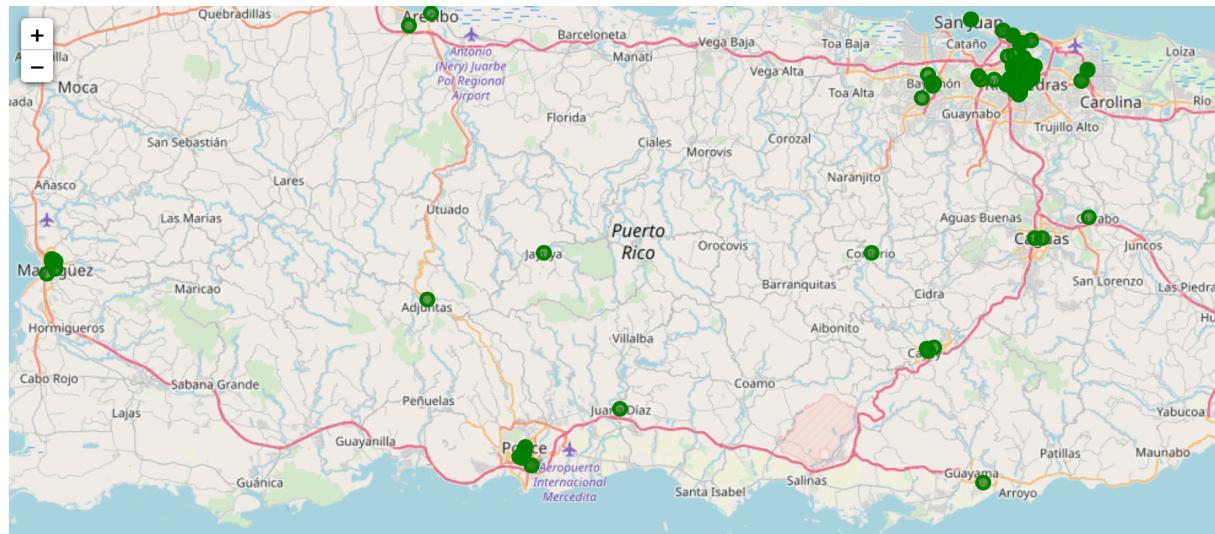
	Localidad	Localidad	Latitude	Localidad	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	00601		18.16595		-66.72363	Esc Jose E Lugo	18.168703	-66.722083	University
1	00604		18.50529		-67.13590	UPR Aguadilla	18.499344	-67.135543	University
2	00604		18.50529		-67.13590	Caribbean Aviation Technical Institute	18.501749	-67.140990	University
3	00613		18.47274		-66.71928	Resepcion	18.473582	-66.718891	University
4	00614		18.45690		-66.73589	UNE. Universidad del Este.	18.459937	-66.743102	University

```
zipcode_venues_uni.shape
```

```
(187, 7)
```

Map showing Universities

```
: map_zipcode_uni = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(zipcode_venues_uni, 'green', map_zipcode_uni)
map_zipcode_uni
```



Offices

There are 3812 Offices

List Of Offices

```
: zipcode_venues_office = getNearbyVenues(names=df['ZIP'], latitudes=df['LAT'], longitudes=df['LNG'], radius=1000, categories='zipcode_venues_office.head()')
```

	Localidad	Localidad	Latitude	Localidad	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	00601		18.16595		-66.72363	Comite PPD Adjuntas	18.165085	-66.723804	Government Building
1	00601		18.16595		-66.72363	Tribunal de Primera Instancia de Adjuntas	18.166603	-66.724940	Courthouse
2	00601		18.16595		-66.72363	COMICIÓN ESTATAL DE ELECCIONES	18.164359	-66.723519	Voting Booth
3	00601		18.16595		-66.72363	Funeraria Del Carmen	18.162893	-66.725784	Funeral Home
4	00601		18.16595		-66.72363	Laboratorio Clinico Adjuntas	18.163314	-66.723996	Office

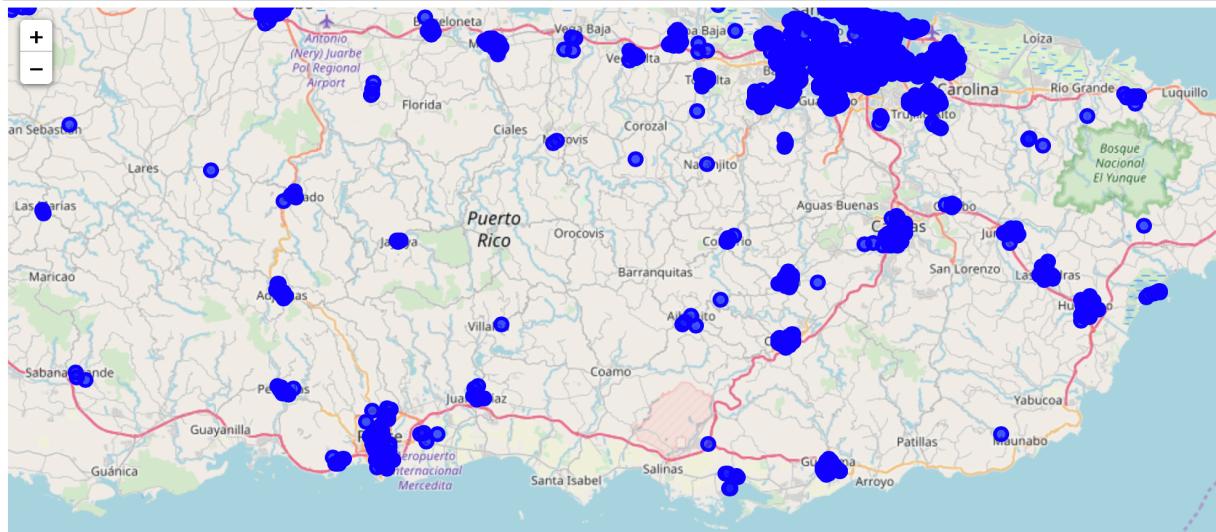
```
: zipcode_venues_office.shape
```

```
: (3812, 7)
```

Map showing Offices

Offices in MAP 1

```
: map_zipcode_office = folium.Map(location=[latitude, longitude], zoom_start=12)
addToMap(zipcode_venues_office, 'blue', map_zipcode_office)
map_zipcode_office
```



In order to find the best score for a location added weights to the venues

Adding weights to the locations based on the existing burger shops

```
# negative weight, to open a burger point we want to avoid concurrence as much as possible
weight_burger = -1

# because high school students are good customers
weight_schools = 1

# university students are good customers so given positive
weight_uni = 2

# employees are even better customers so more weight
weight_offices = 3
```

Applying weights to find the score to find suitable location

```
df_weighted = df_data[['Localidad']].copy()
```

```
df_weighted['Score'] = df_data['Burger'] * weight_burger + df_data['High Schools'] * weight_schools + df_data['Universities'] * weight_uni + df_data['Offices'] * weight_offices
df_weighted = df_weighted.sort_values(by=['Score'], ascending=False)
df_weighted
```

	Localidad	Score
128	00925	184.0
134	00931	176.0
133	00930	175.0
148	00955	168.0
139	00937	168.0
120	00917	167.0
163	00975	166.0
140	00939	166.0
141	00940	166.0

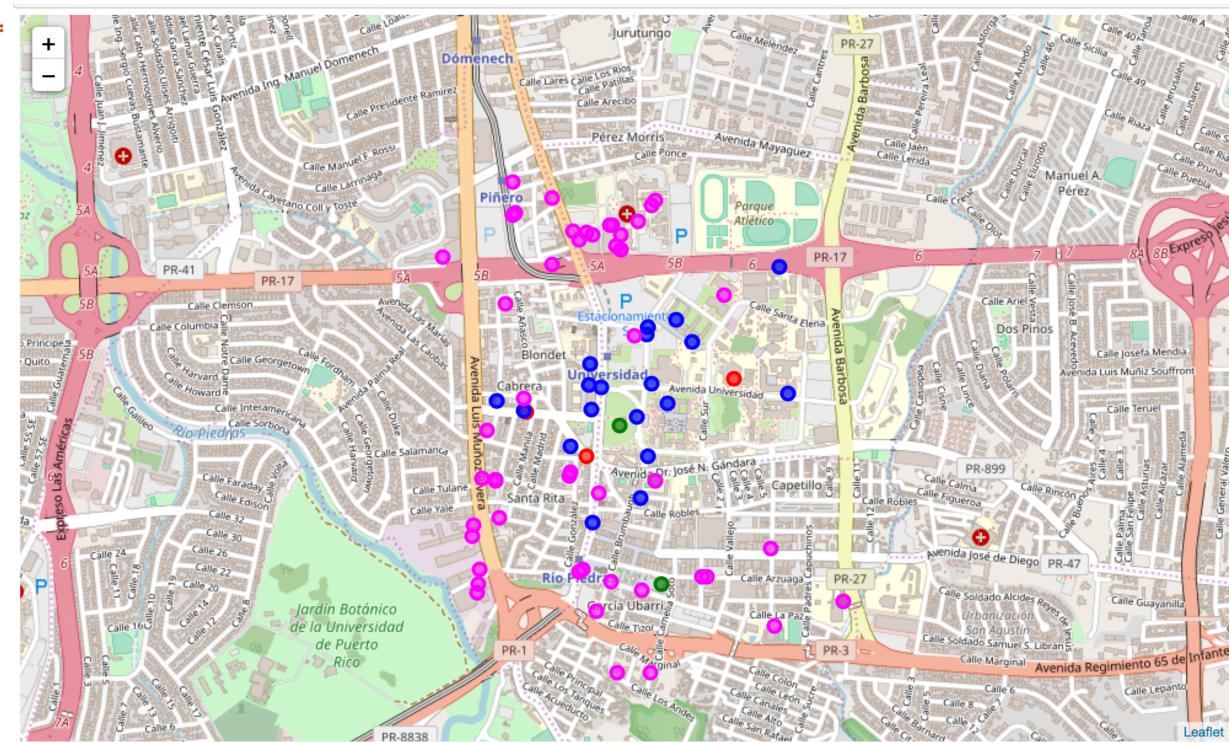
Below map shows the neighborhood of our potential place of our new Burger point

Blue shows Universities

Pink shows Offices

Green shows schools

Red shows Burger Points



zipcode 00925 (SAN JUAN) is the best area to open Burger shop in the state Puerto Rico

Discussion

Based on the entire analysis and applying weights ,

The best score goes to the location zip code: 00925. This location is best for starting a new Burger point.

Details of this zipcode are

```
import zipcodes
zipcodes.matching('00925')

[{'zip_code': '00925',
 'zip_code_type': 'STANDARD',
 'city': 'SAN JUAN',
 'state': 'PR',
 'lat': 18.4,
 'long': -66.06,
 'world_region': 'NA',
 'country': 'US',
 'active': True}]
```

We will examine how many burger pointes are there at this location now serving how many schools, universities and offices at thi location

There are 3 Burger points now

```
venues_at_zipcode[venues_at_zipcode['Localidad']=='00925'].shape

(3, 7)
```

There are 18 schools nearby

```
zipcode_venues_uni[zipcode_venues_uni['Localidad']=='00925'].shape

(18, 7)
```

There are 49 Universities near by.

```
zipcode_venues_office[zipcode_venues_office['Localidad']=='00925'].shape

(49, 7)
```

Conclusion:

We can start a new burger point at SAN JUAN (00925) at this location since we have potential customers to serve and earn profits