

Introduction	2
<i>Background</i>	2
<i>Problem</i>	2
<i>Interest</i>	2
Data	2
Methodology	2
<i>Exploratory</i>	2
<i>Predictive Model</i>	6
Results	6
Discussion	7
Conclusion	7

Introduction

Background

In cities there are many neighborhoods and in that neighbors different people belonging to different economic lines lives. If an investor has an idea of a product, it is important to put the restaurant in a place where people will understand and appreciate the value of the product. Anybody can have a great product to sell, but you have to be in the right place to sell it even to give a great customer service. This project will be executed in the city of santa cruz de la sierra, Bolivia.

Problem

It is hard to find good places to locate a restaurant, in fact, location is one of the most important variables in a project analysis because is where you market target is or your customers are.

Interest

The person interested in this project will be the owner of a dessert restaurant called “L’ pav”. L’ par is a specialized dessert restaurants that already has 2 stores and what to expand to the third one. The owner is trying to get low cost information to locate her next restaurant.

Data

All the data will come from the Foursquare API in first hand and as a complement, from Google API. This is if there is not enough information in one of them. The search parameters for the search query are: “venues” in a radius of 4000 meters and type “food”. All of this around the location of “Plaza 24 de septiembre” (center part of the city).

Methodology

Exploratory

After applying the search, I get a json file with has many fields as: id, name, location, categories, verified, stats, url, and so. I only need location data, category, price, rating. With some coding I get the following table data:

Data columns (total 14 columns):					categories	
#	Column	Non-Null Count	Dtype			
0	name	30 non-null	object		Fast Food Restaurant	5
1	categories	29 non-null	object		Mexican Restaurant	3
2	location.lat	30 non-null	float64		Burger Joint	3
3	location.lng	30 non-null	float64		Food Truck	3
4	location.labeledLatLngs	30 non-null	object		Food Court	2
5	location.distance	30 non-null	int64		Comfort Food Restaurant	2
6	location.cc	30 non-null	object		Frozen Yogurt Shop	2
7	location.state	26 non-null	object		Argentinian Restaurant	1
8	location.country	30 non-null	object		Salad Place	1
9	location.formattedAddress	30 non-null	object		Mediterranean Restaurant	1
10	location.city	25 non-null	object		Buffet	1
11	location.address	13 non-null	object		Ice Cream Shop	1
12	location.crossStreet	4 non-null	object		Theme Restaurant	1
13	id	30 non-null	object		Thai Restaurant	1
					American Restaurant	1
					BBQ Joint	1

The information is tiny so there is not enough to make any statistical inference but check the data and if it has what I am looking for.

	name	categories	location.lat	location.lng	location.labeledLatLngs	location.distance
0	Las Vegas Buffet Fast Food	Buffet	-17.773001	-63.194445	[{'label': 'display', 'lat': -17.773001, 'lng': ...}	1738
1	Brazilian Fast Food	Fast Food Restaurant	-17.782999	-63.185392	[{'label': 'display', 'lat': -17.782999, 'lng': ...}	348
2	Chapaco's Fast Food Grill	BBQ Joint	-17.766326	-63.164481	[{'label': 'display', 'lat': -17.766326, 'lng': ...}	2657
3	Canyella - Food, Drinks & Music	Mediterranean Restaurant	-17.782080	-63.186061	[{'label': 'display', 'lat': -17.782080, 'lng': ...}	439
4	My Food	Fast Food Restaurant	-17.785425	-63.188457	[{'label': 'display', 'lat': -17.785425, 'lng': ...}	712
5	Garajes Food Trucks	Food Truck	-17.755136	-63.172903	[{'label': 'display', 'lat': -17.755136, 'lng': ...}	3283
6	Mexican Food	Mexican	-17.788620	-63.186354	[{'label': 'display', 'lat': -17.788620, 'lng': ...}	742

The working table (dataframe) remains as:

5d7bd58e2deabf00082c99db	not	rated
4d08be1ce554721e4f66c3f8	not	rated
5349db8f498e54a3178975a3	not	rated
52b1dedd11d2564fa7a312a0	not	rated
55048e00498ee7519a90e643	not	rated
5a419c6a6f0aa276862e5dfd	not	rated
4f95ecdbe4b0d1d5d4540221	not	rated
4fc16687e4b0d516251295af	not	rated
50fb48c3e4b0d0486d09b3f3	8.2	
524af9f611d21590ad9a2786	not	rated
505a71f7e4b065a90b794f8d	not	rated
5a7b7b7ecf72a06f3769c91b	not	rated
5b3ac96a286fda002c4d3427	not	rated
57964237498eaf4cb022e23d	not	rated
52cbf765498ee59c9b700bdb	not	rated
59b459d995da0c3b7f87be4a	not	rated
5de2ae1ba11af9000860180d	not	rated
5a7263015ba046224d084786	not	rated
5a542a2486bc4977f4fdcdc	not	rated
535d4b9d498e55b16385cd13	not	rated
5192c999498e5b1e2ff4fd2a	not	rated
508dc940e4b055a759175128	not	rated
5ad9326eb9a5a84263bd83e4	not	rated
5bd12199088158002c5f92bb	not	rated
5c645dad4a7aae002cbc9ea3	not	rated
5a7ef90cee712049240ab452	not	rated
551727a4498e4a598e2409cf	not	rated

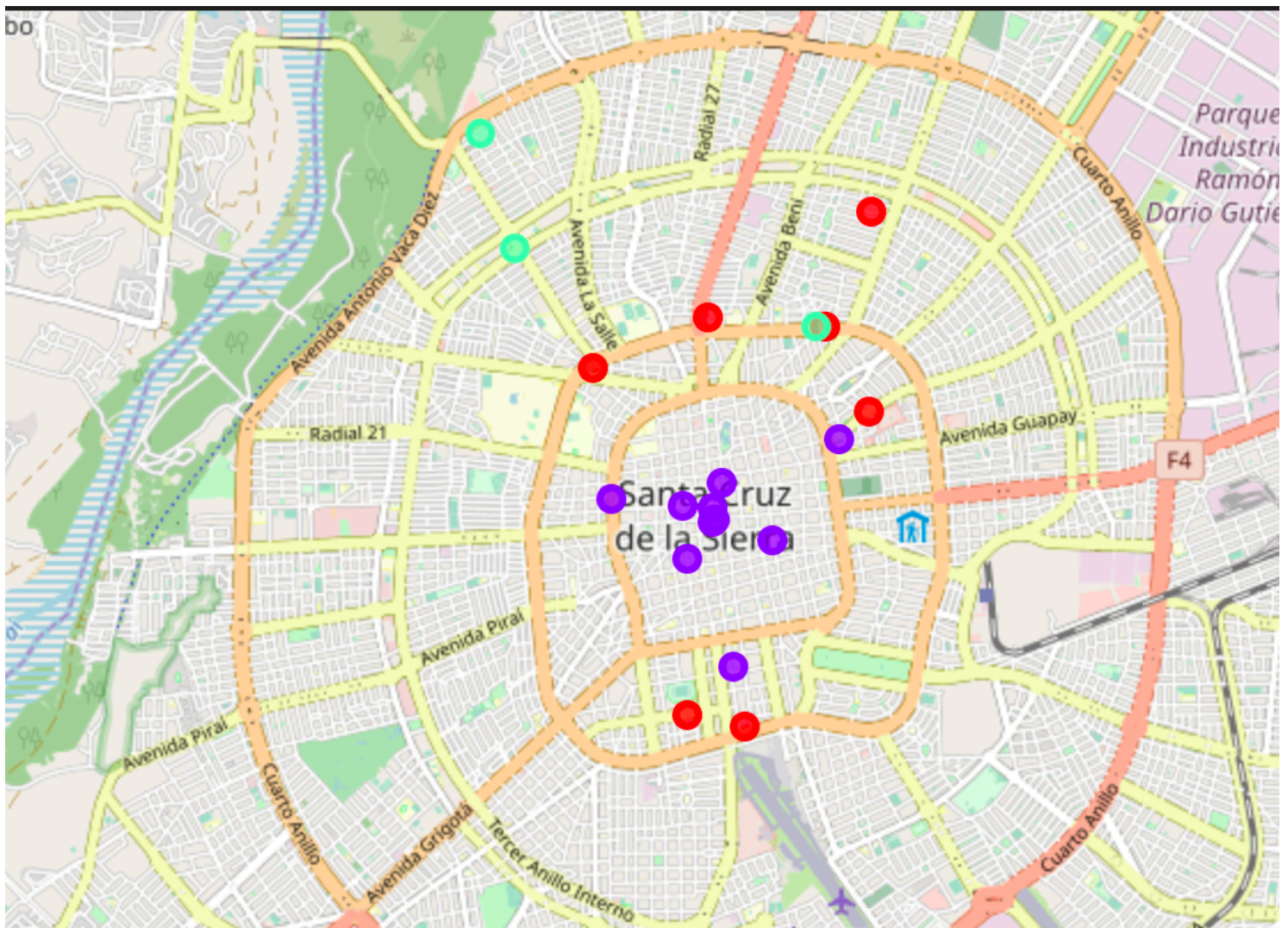
Once i got the data frame, I check if there was a rating column, but it did not, so, I had to check venue by venue if there wasn't a rating column and this was the result:

As we can see above, there is no information about ratings so, that is why I had to use google api to search for them. Google has a lot of information as well and is complementary for the first one. The columns of the data

coming from Google is as follows:

```
places_gmaps_df.columns
```

```
Index(['icon', 'id', 'name', 'photos', 'place_id', 'rating', 'reference',  
      'scope', 'types', 'user_ratings_total', 'vicinity',  
      'geometry.location.lat', 'geometry.location.lng',  
      'geometry.viewport.northeast.lat', 'geometry.viewport.northeast.lng',  
      'geometry.viewport.southwest.lat', 'geometry.viewport.southwest.lng',  
      'opening_hours.open_now', 'plus_code.compound_code',  
      'plus_code.global_code', 'price_level'],  
      dtype='object')
```



```
#getting ratings per location and name of the restaurant and sorting values
places_gmaps_df[["name", "rating", "user_ratings_total", "types"]].sort
```

	name	rating	user_ratings_total	
8	Casa del Camba	4.3	4118	[
1	Los Tajibos Hotel & Convention Center	4.7	3331	[
0	Hotel Camino Real	4.7	2654	[
2	Irish Pub	4.3	1386	[
19	Piccolo	3.6	1134	[
3	Net Pizza	4.4	1091	[
18	Hot Burger	3.9	1066	[
9	El Aljibe	4.4	1046	[
5	Burger King	4.3	763	[
6	Hamburguesas Toby	4.0	694	[
15	Pastelería Fidolin	4.2	596	[
10	La Guitarra	4.0	567	[
4	Hamburguesas Toby	3.8	409	[
13	Lorca Café, Bar y Teatro	4.1	398	[

Predictive Model

The predictive model used for this project was k-means because it will let me know what restaurants are cluster taking into account the distance, likes, ratings and visitors.

The k-means separates the samples into groups of same variance so that's why I use k-means.

Results

After running the k-means algorithms we can see that the following points have common behavior:

The green ones are the top 3 restaurants. To see it with more detail, the following table was built:

Discussion

There are no restaurants specialized in desserts and it will be a good opportunity to sell these products to a cluster of people that appreciate good taste and quality.

I had to use 2 APIs to get the information about likes and ratings. That might be because there aren't much touristic users that give this information to the APIs. It would be a good idea to

Conclusion

locate L' PAV near one of these top 3 restaurants because people that it very often want a really good dessert.