Finding the best place to open a bar in Vitória/ES – Brazil

Roberto Carlos do Nascimento Siqueira

November 4, 2020

1. **Introduction**
   ## 1.1. Background
   Vitória over the years is in the top 10 cities to live. The city is an island and also has a continental part. It's an old city with 469 years old. The city population is small compared with other cities and capitals. With only 363 thousand habitants, Vitória is being considered as the city with the best public services as education and health in Brazil. Some business are not allowed to have in specifics regions due the Urban Director Plan (PDU, in portuguese). This plan limit some areas according to specifics rules. It's not part of the scope of this work. This work is a final project of the IBM Data Science Professional certification. As the city has some beach, the first place to consider to open a bar would be there. Although, the litoral has some kiosk controlled by the city hall and it's not allowed to construct anything there. This report details the process to finding a best place to have a bar considering main factors as per capita income, others bars and so on.

   ## 1.2. Target Audience
   The target audience is people interested in open a bar in Vitória/ES-Brazil.

2. **Data**

   The data to be used in this analysis will be provided by three sources:

   1. **Foursquare**
   2. **"Vitória Bairro a Bairro – 2013" report**
      (http://legado.vitoria.es.gov.br/regionais/geral/publicacoes/Vitoria_bairro_bairro/Vit%C3%B3ria_bairro_%20a_bairro.pdf)
   3. **"Administrative regions – Vitória" report**
      (http://legado.vitoria.es.gov.br/regionais/geral/dados/reg_adm_ger.pdf)

   ## 2.1. Foursquare
   The Foursquare will be responsible to show others main bars in Vitória.

   ## 2.2. "Vitória Bairro a Bairro – 2013" report
   This report will provide some informations as Male and Female concentration per region, age range and income class.

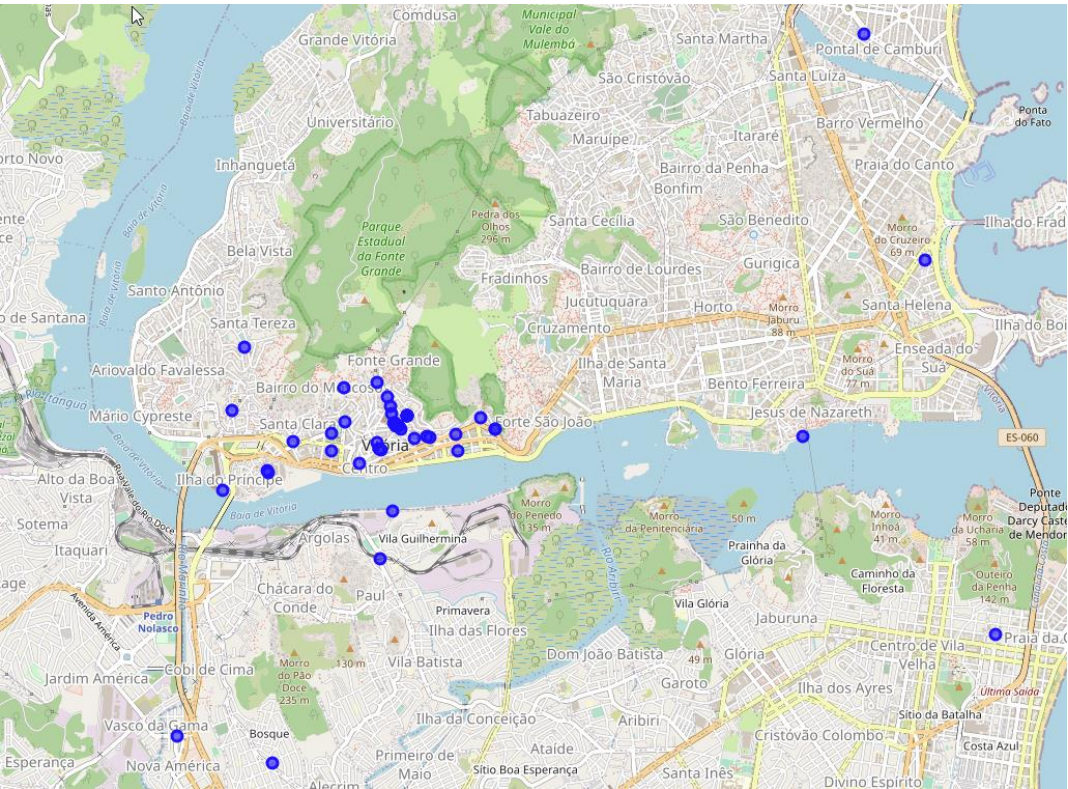   ## 2.3. "Administrative regions – Vitória" report
   Vitória is divided in seven administrative regions. This information will be acquired from this report.

3. **Methodology**

   ## 3.1. Data collection, preparation, exploration and visualization
   ### 3.1.1. Foursquare

Using the jupyter notebook available in the Git repository is possible to get all kind of bars in Vitória and theirs locations. The most relevant bars are more concentrate in the towncity. This information is available as a panda dataframe and a graph shown below:

```
        'country': 'Brasil',
        'distance': 1446,
        'formattedAddress': ['Brasil'],
        'labeledLatLngs': [{'label': 'display',
          'lat': -20.312522725912157,
          'lng': -40.3493830273724}],
        'lat': -20.312522725912157,
        'lng': -40.3493830273724},
        'name': 'Bar do Zeca',
        'referralId': 'v-1604537566'}]}}
```

```
[53] # assign relevant part of JSON to venues
     venues = results['response']['venues']

     # tranform venues into a dataframe
     dataframe = json_normalize(venues)
     dataframe

/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:5: FutureWarning: pandas.io.json.json_normalize is deprecated, use pandas.json_normalize instead
     """
```

	id	name	categories	referralId	hasPerk	location.address	location.lat	location.lng	location.labeledLatLngs	location.distance	location.c
0	4ee2b58561aff5a340194235	Buery Bar	[{'id': '4bf58dd8d48988d1c7941735', 'name': 'S...	v-1604537566	False	Rua 7	-20.318546	-40.336738	[{'label': 'display', 'lat': -20.3185460006729...	455	Bi

3.1.2. "Vitória Bairro a Bairro – 2013" report

The data was collected and treated as the image below:

Before:

```
[88] vitoria_data = pd.read_csv("vitoria.csv")
```


	Homens	Mulheres	0-14 anos	15-24 anos	25-64 anos	acima de 65	Ate 1/2 salario minimo	Mais de 1/2 a 1 salario minimo	Mais de 1 a 2 salarios minimos	Mais de 2 a 5 salarios minimos	Mais de 5 a 10 salarios minimos	Mais de 10 a 20 salarios minimos	Mais de 20 salarios minimos	Sem rendimento
Centro	4416	5422	1274	1441	5673	1450	8	116	340	1035	1155	772	262	79
Santo Antonio	343	444	174	134	407	72	1	26	46	89	57	13	2	7
Bento Ferreira/Jucutuquara	2741	2904	833	799	3292	645	4	32	80	320	532	604	361	37
Maruipe	977	1054	420	389	1098	124	2	64	161	274	109	8	1	33
Praia do Canto	2634	2995	851	849	3448	481	2	14	40	169	385	606	744	48
Continental	599	748	180	230	737	200	2	14	30	119	136	86	30	9
Sao Pedro	438	456	228	178	443	45	2	38	84	90	13	0	0	56

After:

```
[105] data.sort_values(by='People',ascending=False)
```

	People	Teen	Adult	Older	Low	Medium	High
Centro	9838	1274	7114	1450	124	1375	2189
Bento Ferreira/Jucutuquara	5645	833	4091	645	36	400	1497
Praia do Canto	5629	851	4297	481	16	209	1735
Maruípe	2031	420	1487	124	66	435	118
Continental	1347	180	967	200	16	149	252
Sao Pedro	894	228	621	45	40	174	13
Santo Antonio	787	174	541	72	27	135	72

As can be noted the "Centro" is the place with more concentration of people, followed by "Bento Ferreira/Jucutuquara" and "Praia do Canto".

It's possible to noted in the figure below that the "Centro" is the region with more income concentration, followed by "Praia do Canto" and "Bento Ferreira/Jucutuquara":

```
[106] data.sort_values(by='High',ascending=False)
```

	People	Teen	Adult	Older	Low	Medium	High
Centro	9838	1274	7114	1450	124	1375	2189
Praia do Canto	5629	851	4297	481	16	209	1735
Bento Ferreira/Jucutuquara	5645	833	4091	645	36	400	1497
Continental	1347	180	967	200	16	149	252
Maruípe	2031	420	1487	124	66	435	118
Santo Antonio	787	174	541	72	27	135	72
Sao Pedro	894	228	621	45	40	174	13

Using K-means we can note that "Bento Ferreira/Jucutuquara" and "Praia do Canto" are in the same cluster. We can choose one or other:

```
[116] data["Clus_km"] = labels
      data
```

```
/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:1: SettingWithCopyWarning
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/use
  """Entry point for launching an IPython kernel.
```

	People	Teen	Adult	Older	Low	Medium	High	Clus_km
Centro	9838	1274	7114	1450	124	1375	2189	0
Santo Antonio	787	174	541	72	27	135	72	1
Bento Ferreira/Jucutuquara	5645	833	4091	645	36	400	1497	2
Maruípe	2031	420	1487	124	66	435	118	1
Praia do Canto	5629	851	4297	481	16	209	1735	2
Continental	1347	180	967	200	16	149	252	1
Sao Pedro	894	228	621	45	40	174	13	1

4. **Results**

Crossing with data from Foursquare it's possible to see that there is no significant places there. Due this, the best place to open a bar in Vitória, according to the data available is "Praia do Canto" that has the highest income.