

# Location analysis to open a Bistro & Wine Bar in the City of Boston

## Overview

1. Introduction
2. Business Problem
3. Data Description
4. Methodology
  - 4.1 Data acquisition
  - 4.2 Exploratory data analysis
    - i. Search parameters
    - ii. Filtering data results
    - iii. Mapping results
    - iv. Venue ratings
    - v. Category analysis
5. Results
6. Discussion section
7. Conclusion

## 1. Introduction

A group of investors is interested in opening a Bistrô & Wine Bar in the City of Boston, and is needing a thorough location analysis before moving ahead with other important business considerations.

The investors are primarily professionals from the banking industry and wine enthusiasts. As experienced business people they have the knowledge and the means to evaluate the opportunities and risks of starting a new business using a structured decision making process. As world travelers they bring extensive experience in visiting and vacationing in top wineries and resorts in Europe, mainly in Italy, France and Spain, and use this knowledge to design and build an unique and outstanding place in the city.

It is important to categorize the bistro from restaurants, grills, sports bars and pubs. Although they serve food and drinks, a bistro is defined as a small, inexpensive restaurant which serves food at moderate prices in a casual or modest setting. The atmosphere inside the premises is relaxed and informal.

The plan is to have a nice Mediterranean interior design, a reduced but top quality food menu and a great selection from major wineries from California, France, Italy and Spain. The goal is to make it a new landmark in the city!

## **2. Business Problem**

One key element of the entrepreneurial success for any brick & mortar business starts with choosing the right location for the venture. The old quote “location, location, location” is still the number one rule in real estate.

The real estate location analysis considers the aspects of the attractiveness of the neighborhood, population density, demographics, average income, built-in traffic, tourism expenditure, economic trends, and public safety indicators among others.

This preliminary study will identify the existence of direct or indirect competition in the same target area or neighborhood.

The favorite areas under consideration by the investors are the Boston Downtown Business District, Back Bay, Beacon Hill and Waterfront, as shown in the city map below within the red squares.

In summary, this analysis will address the following to the client investors:

- Where are the best locations to establish the Bistro & Wine Bar?
- Where and who are the closest competitors?

The areas of interest for the investors are enclosed in red squares.

Area 1: Back Bay, Beacon Hill and Bay Village

Area 2: Waterfront, Downtown / Business District and North End



### 3. Data Description

The location analytics will rely on data supplied by official sources, such as local government planning and development agencies, and data companies offering free access to data sets. All available datasets will be scrutinized for its reliable sources and current data, which includes:

- Neighborhoods and postal codes
- Geo coordinates
- Venue data

The City of Boston data hub ANALYZE BOSTON, offers 161 datasets for public access, which includes a variety of topics such as zoning districts, crime incident reports and economic indicators that will be used in this study.

The location analysis will employ Foursquare for the searching of venues, and other mapping and plotting applications, like Folium.

Reference places in the preferred investor's neighborhoods are Sheraton Boston, Hilton, at the Back Bay and the Marriott Long Wharf and Renaissance Hotel at the Waterfront.

Data sources and uses:

<https://data.boston.gov/>

The City of Boston data hub – ANALYZE BOSTON

<https://datausa.io/profile/geo/boston-ma/>

Created by MIT MediaLab and Deloitte for mining and visualizing US government open data

<https://www.topozone.com/massachusetts/suffolk-ma/>

Boston Topographic maps

<https://www.findlatitudeandlongitude.com/>

Maps and geo coordinates

<https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/table/?>

Postal codes and geo coordinates

<https://www.neighborhoodscout.com/ma/boston/crime>

Public safety and crime incident

<https://www.homesnacks.net/about-us/>

Combines recent data from the Census, FBI, OpenStreetMaps

## 4. Methodology

### 4.1 Data acquisition

1. Initial search to select reliable and current datasets containing the City of Boston neighborhoods, postal codes and geo coordinates. This is the required data to be used to plot and identify locations in the maps and calculate distances between reference points.
2. Acquire the datasets, through downloads or web scraping and combine them in useful format in pandas dataframes. The first built dataframe [neighborhood name, zip postal code, and geo coordinates] was the result of the concatenation of data coming from separate sources.
3. Data cleaning step to verify missing, incomplete, duplicate or mismatched data. A reality check is necessary to ensure a reliable and accurate data table to be used in the subsequent steps.

#### Steps

1. Select the data sources, download the datasets, write them in the local hard drive for back-up and offline work purpose and load them into pandas dataframe.

#### Transform the data into a pandas dataframe

```
In [3]: boston_zip.head()
```

```
Out[3]:
```

	ZIP Code	Neighborhood
0	2134	Allston
1	2115	Back Bay
2	2116	Bay Village
3	2114	Beacon Hill, West End
4	2135	Brighton

2. Sort the ZIP Code column in ascending order, reset the index and check for duplicates. If duplicates, execute the code to “groupby and agg” and combine the neighborhood names separated by comma. Inspect the results.

```
# Inspect the resulting dataframe  
boston_zip.head()
```

	ZIP Code	Neighborhood
0	2110	Downtown
1	2111	Chinatown, Leather District
2	2113	North End
3	2114	Beacon Hill, West End
4	2115	Back Bay

3. Load dataset containing Boston postal codes and geo coordinates

```
In [13]: boston_zip_coord.head()
```

```
Out[13]:
```

	ZIP	Nborhood	Unnamed: 2	Lat	Long	Unnamed: 5	Unnamed: 6
0	2134	Allston	NaN	42.358162	-71.054065	NaN	NaN
1	2115	Back Bay	NaN	42.350274	-71.058768	NaN	NaN
2	2116	Bay Village	NaN	42.365161	-71.055472	NaN	NaN
3	2114	Beacon Hill / West End	NaN	42.362250	-71.067337	NaN	NaN
4	2135	Brighton	NaN	42.342124	-71.096672	NaN	NaN

4. Sort the ZIP Code column in ascending order, delete “Unnamed” columns, reset the index and check for duplicates. If duplicates, execute the code to “groupby and agg” and combine the neighborhood names separated by comma, and concatenate this dataframe with previous dataframe. Inspect the results.

```
Out[22]:
```

	ZIP Code	Neighborhood	Lat	Long
0	2110	Downtown	42.307319	-71.085935
1	2111	Chinatown, Leather District	42.337030	-71.072034
2	2113	North End	42.283214	-71.127027
3	2114	Beacon Hill, West End	42.362250	-71.067337
4	2115	Back Bay	42.350274	-71.058768

## 4.2 Exploratory data analysis

The Foursquare API is the main tool used to locate and identify venues in the selected neighborhoods in the city. It is used to find venues, ratings and number of tips. Venue categories will be grouped and counted. A Developer’s account was open in order to get access rights to the use of the data.

### Search parameters

Interactive search with the keywords ‘wine bar’ and ‘bistro’ is used in the venue category. It is important to differentiate the nature of the business from regular restaurants, pubs, grills and sports bars. In case of any of these categories appear in the search result, additional exploratory analysis is made to determine whether it represents a direct or indirect competition.

The other Foursquare parameter influencing the dataframe results is radius = 1000 meters from the reference points, after trying the value of 500 meters that did not bring enough representative venues.

The parameter “Limit =” that limits the number of return results was set to 30.

As the investors are mainly interested in upscale areas, the reference points selected were Sheraton Boston in the Back Bay and Marriott Long Wharf at the Waterfront.

### Steps

1. Initiate the Foursquare Dev Account with the proper credentials.
2. Acquire the geo coordinates for the reference point = Sheraton Boston Hotel at the Back Bay  
address = '39 Dalton St, Boston, MA 02199 using Nominatim from OpenStreetMapdata  
Sheraton Boston Hotel coordinates Lat:42.3467642 Long:-71.0836627
3. Search for a specific venue category within the given radius of the selected location

```
In [64]: search_query = 'wine bar'
radius = 1000 # expressed in meters of the specified location
print('Search category: ', search_query)

Search category: wine bar
```

4. Execute Foursquare function, generate dataframe with results list
5. Inspect results list and venue categories

```
0          Barcelona Wine Bar
1      Jaho Coffee Roaster & Wine Bar
2          Piattini Wine Cafe
3          Side Bar Grille
4          The Wine Emporium
5      Crazy Dough's Pizza-Artisan Slice Bar
6          Forty Dalton Bar
7      Polenta Bar Pop-Up at La Rosticceria
8          The Wine Emporium
9          Genius Bar
10         Pour House Bar & Grill
11         SideBar & Grille
12         Sheraton Shame Bar
13      Joe's American Bar & Grill
14         Earls Kitchen + Bar
15         Eataly Boston
16         WFM Coffee Bar
17         Harvard Club Grill Bar
18         Bar Boulud
19         The Great Bitter Bar
20         Cello Bar
21         The Paint Bar (Boston)
22         Clarendon Wine Co.
23         The Tie Bar
24         Wine palace
25      Brother's Wine & Spirits
26         The Bar Method
27         Precinct Kitchen & Bar
28         McCarthys Bar and Grill
29      Choice Fine Wine And Spirits
```

## Filtering results

The results table requires sanity check and some cleaning. The words "wine" and "bar" in the search query retrieved venues with the "bar" in the name such as Barnes & Noble (bookstore) and barbershop. The Genius Bar, which is listed as "Electronics Store" category and it is the name used by the Apple for its tech support station located inside Apple's retail stores. Besides that, wine retail stores were also part of the returned list. A visual inspection determined the drop list of all these deviations to create the dataframe filter.

### Steps

1. Create the drop list and delete entries

```
In [34]: sheraton_df = dataframe_filtered

drop_list = [4,5,7,8,12,13,16,21,22,23,25,26,29]
rows = sheraton_df.iloc[drop_list]
rows # review drop_list to confirm deletion
```

2. Clean and arrange the resulting dataframe

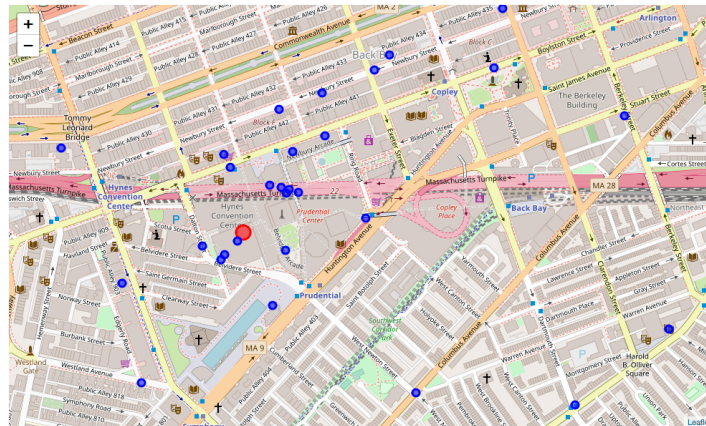
	name	categories	address	distance	lat	lng
0	Barcelona Wine Bar	Wine Bar	525 Tremont St	1075	42.344622	-71.070919
1	Jaho Coffee Roaster & Wine Bar	Coffee Shop	116 Huntington Ave	303	42.347077	-71.079997
2	Piattini Wine Cafe	Italian Restaurant	226-228 Newbury St	394	42.349857	-71.081315
3	Side Bar Grille	Restaurant	39 Dalton St	26	42.346573	-71.083847
4	Forty Dalton Bar	Hotel Bar	40 Dalton St	106	42.346466	-71.084892
5	Pour House Bar & Grill	Bar	907 Boylston St	162	42.348197	-71.084051
6	SideBar & Grille	Bar	39 Dalton St	72	42.346269	-71.084227
7	Sheraton Shame Bar	Cocktail Bar	NaN	86	42.346166	-71.084326
8	Earls Kitchen + Bar	American Restaurant	800 Boylston St	167	42.347656	-71.082018
9	Eataly Boston	Gourmet Shop	800 Boylston St	153	42.347707	-71.082297
10	The Great Bitter Bar	Bar	800 Boylston St	143	42.347641	-71.082390
11	Bar Boulud	French Restaurant	776 Boylston St	310	42.348895	-71.081231
12	Cello Bar	Wine Bar	NaN	144	42.347759	-71.082525
13	McCarthys Bar and Grill	Bar	903 Boylston St	198	42.348495	-71.084251
14	Precinct Kitchen & Bar	American Restaurant	154 Berkeley St	982	42.349350	-71.072251



## Mapping results

Folium is used to map the results loaded into the dataframe obtained from Foursquare, and superimposing the locations over the neighborhood map within a radius of 1000 meters from the reference points.

Reference point: Sheraton Boston Hotel at the Back Bay



## Venue ratings

The venue ratings is a ten-point scale used to provide users with information about the best places around. It is a result of direct users' feedback and one indication how well the venue is performing. This will be used to get us a glimpse of the place in cases where it is believed to be a direct or close competitor.

### Steps

1. Execute the Foursquare Explore capability and get the ratings and number of tips for a selected venue:

```
In [49]: venue_id = '5410e67c498e777b79eb81df' # ID of Bar Boulud
url = 'https://api.foursquare.com/v2/venues/{}?client_id={}&client_secret={}&v={}'.format(venue_id, CLIENT_ID, CLIENT_S
result = requests.get(url).json()
try:
    print('The rating for this venue is: ',result['response']['venue']['rating'])
except:
    print('This venue has not been rated yet.')

The rating for this venue is: 6.3

In [81]: print('Total tips for this venue is: ',result['response']['venue']['tips']['count'])

Total tips for this venue is: 26
```

## Category analysis

Group and plot the returned categories

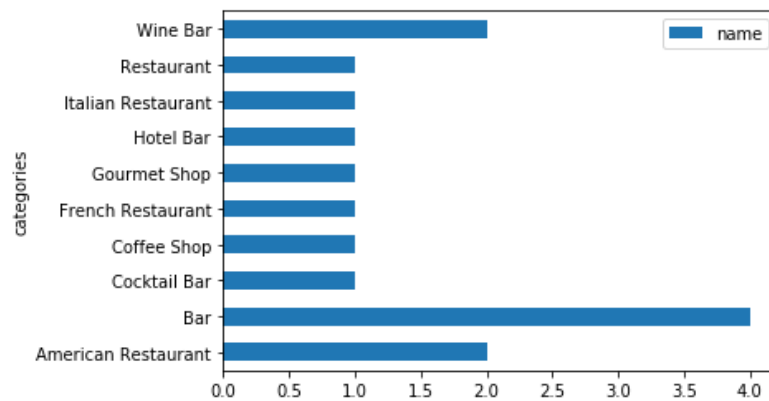
```
In [48]: sheraton_df.groupby(['categories']).count()
```

Out[48]:

categories	name	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	crossStreet	neighborhood	id
American Restaurant	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2
Bar	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4
Cocktail Bar	1	0	1	1	1	1	0	1	1	1	1	1	0	0	1
Coffee Shop	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
French Restaurant	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
Gourmet Shop	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
Hotel Bar	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
Italian Restaurant	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
Restaurant	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
Wine Bar	2	1	2	2	2	2	2	2	2	2	2	2	0	0	2

```
In [49]: sum_cat = sheraton_df.groupby(['categories']).count()
```

```
In [50]: ax = sum_cat.plot.barh( y='name', rot=0)
```



Repeat the above process with Foursquare for another reference point.

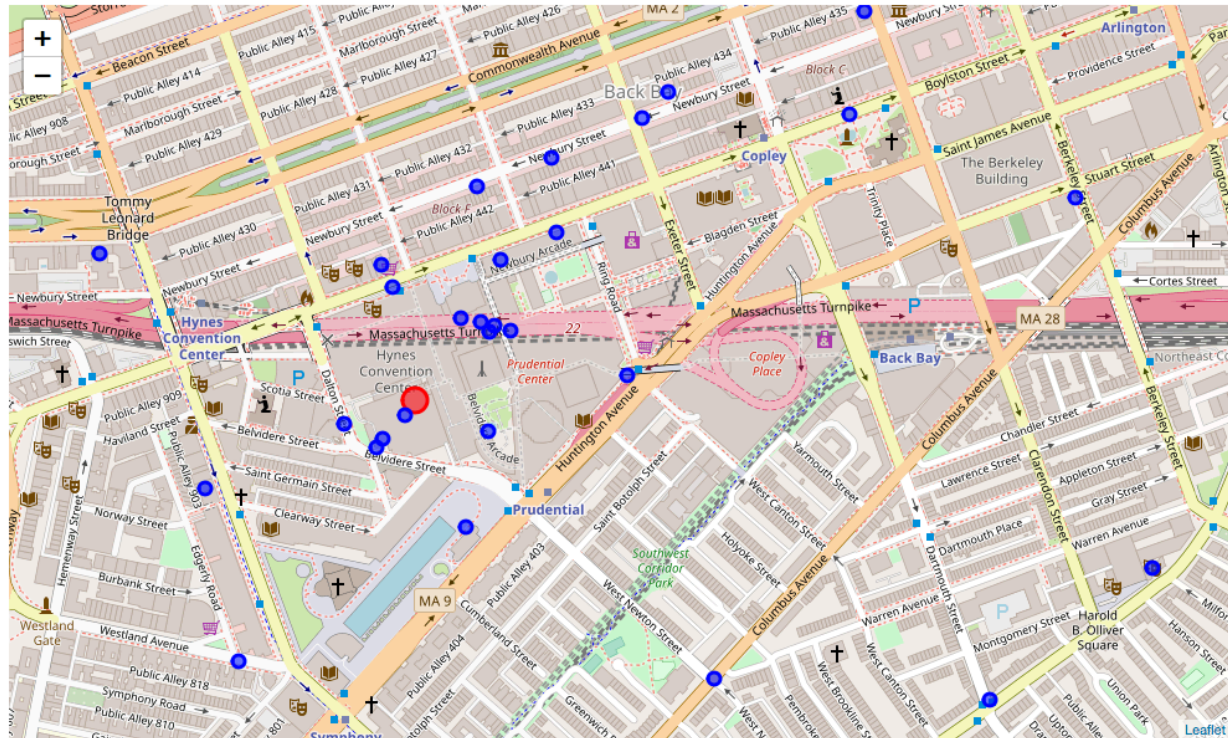
This time is the Boston Marriott Long Wharf at the Waterfront, located at:  
Address = 296 State St, Boston, MA 02109

The results for the Waterfront analysis are presented in the next section.

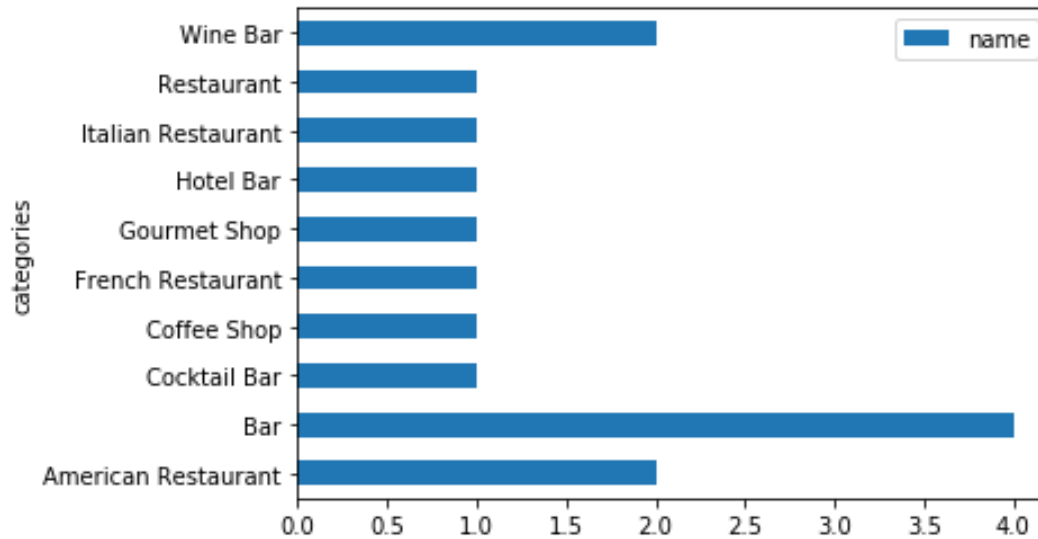
## 5. Results

The generated maps give a good overview of the distribution of the filtered business establishments within a radius of 1000 meters of the reference points.

Reference point: Sheraton Boston at Back Bay – Venues map



## Reference point: Sheraton Boston at Back Bay – Venue categories

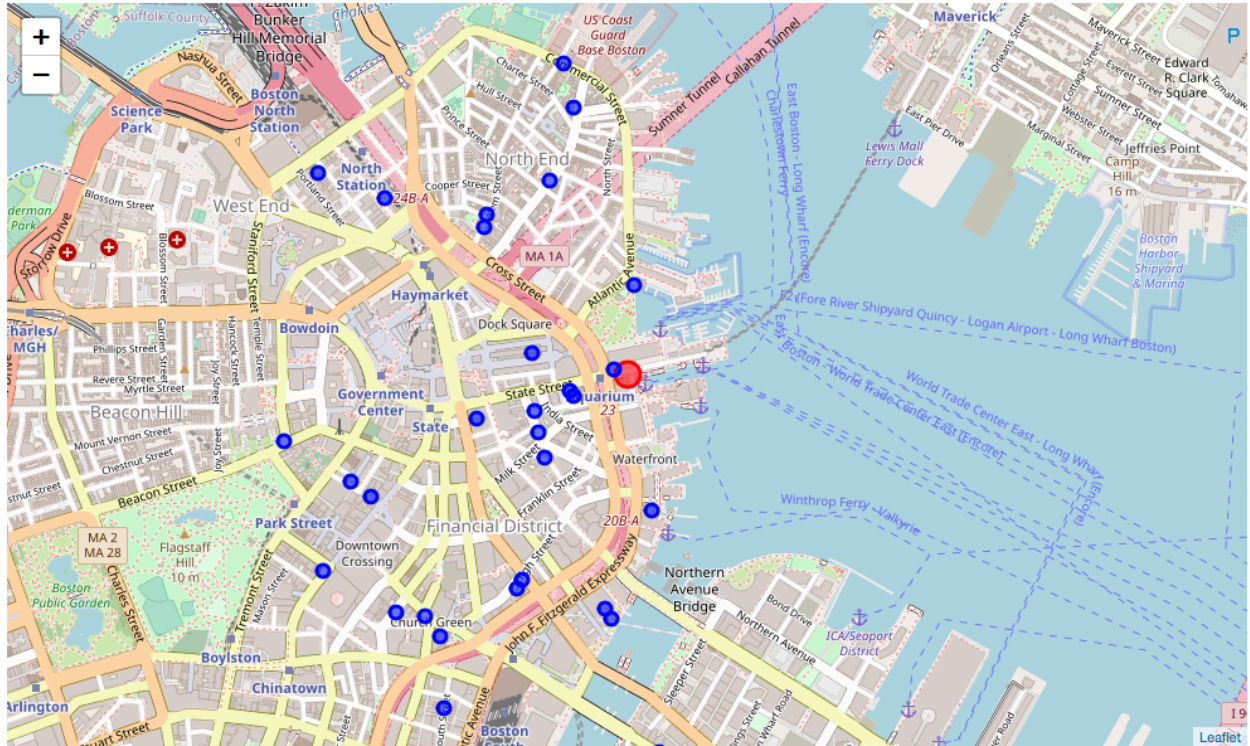


## Reference point: Sheraton Boston at Back Bay – List of filtered venues

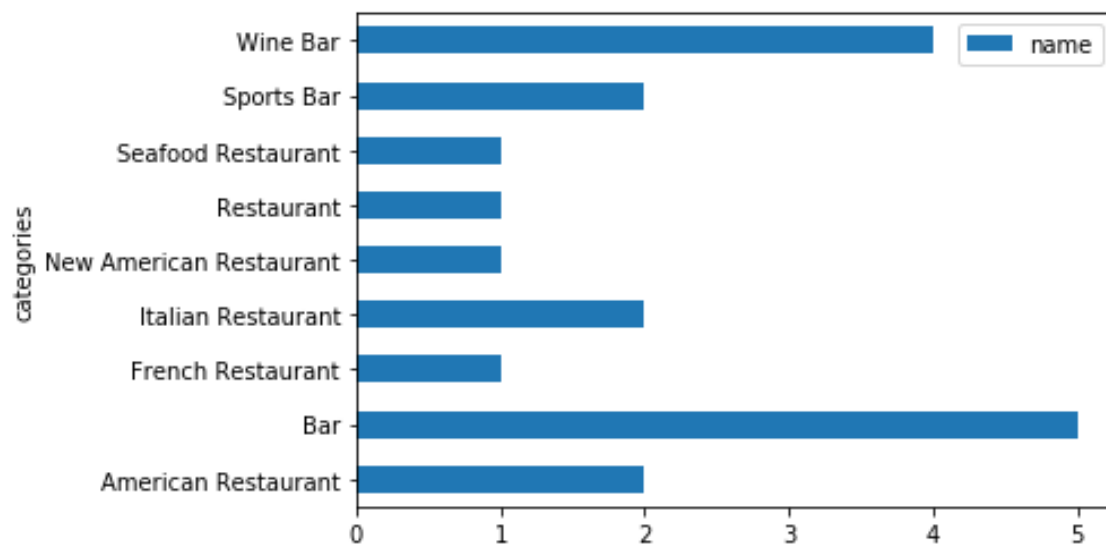
	name	categories	address	distance	lat	lng
0	Barcelona Wine Bar	Wine Bar	525 Tremont St	1075	42.344622	-71.070919
1	Jaho Coffee Roaster & Wine Bar	Coffee Shop	116 Huntington Ave	303	42.347077	-71.079997
2	Plattini Wine Cafe	Italian Restaurant	226-228 Newbury St	394	42.349857	-71.081315
3	Side Bar Grille	Restaurant	39 Dalton St	26	42.346573	-71.083847
4	Forty Dalton Bar	Hotel Bar	40 Dalton St	106	42.346466	-71.084892
5	Pour House Bar & Grill	Bar	907 Boylston St	162	42.348197	-71.084051
6	SideBar & Grille	Bar	39 Dalton St	72	42.346269	-71.084227
7	Sheraton Shame Bar	Cocktail Bar	NaN	86	42.346166	-71.084326
8	Earls Kitchen + Bar	American Restaurant	800 Boylston St	167	42.347656	-71.082018
9	Eataly Boston	Gourmet Shop	800 Boylston St	153	42.347707	-71.082297
10	The Great Bitter Bar	Bar	800 Boylston St	143	42.347641	-71.082390
11	Bar Boulud	French Restaurant	776 Boylston St	310	42.348895	-71.081231
12	Cello Bar	Wine Bar	NaN	144	42.347759	-71.082525
13	McCarthys Bar and Grill	Bar	903 Boylston St	198	42.348495	-71.084251
14	Precinct Kitchen & Bar	American Restaurant	154 Berkeley St	982	42.349350	-71.072251



Reference point: Boston Marriott Long Wharf at the Waterfront – Venues map



Reference point: Boston Marriott Long Wharf at the Waterfront – Venues categories



Reference point: Boston Marriott Long Wharf at the Waterfront – List of filtered venues

	name	categories	address	distance	lat	lng
0	Les Zygomates Wine Bar Bistro	Wine Bar	129 South St	1067	42.351254	-71.057344
1	Taste Wine Bar & Cafe	Wine Bar	NaN	882	42.353578	-71.057971
2	Sarah's Rooftop Wine Bar	Wine Bar	414 Hanover St	762	42.366367	-71.052933
3	The Wine Bike Pop Up Wine Bar	Wine Bar	125 Summer St	904	42.353051	-71.057487
4	Custom House Bar	Bar	NaN	40	42.359791	-71.051529
5	Joe's American Bar And Grill	American Restaurant	100 Atlantic Ave	252	42.361916	-71.050861
6	Warehouse Bar & Grill	American Restaurant	40 Broad Street	297	42.358193	-71.054111
7	Miel Brasserie Provençale	French Restaurant	510 Atlantic Ave	684	42.353517	-71.051640
8	Kingston Grille & Bar (KGB Bar)	New American Restaurant	25 Kingston St	930	42.353659	-71.058965
9	Wet Bar	Bar	183 State St	167	42.359257	-71.053040
10	Rowes Wharf Bar	Bar	70 Rowes Wharf	387	42.356224	-71.050255
11	Blake's Kitchen + Bar	Restaurant	276 Friend St	1034	42.364723	-71.061610
12	Dockside Sports Bar	Sports Bar	183 State St	162	42.359113	-71.052911
13	Rocco's Cucina & Bar	Sports Bar	450 Commercial St	889	42.367481	-71.053261
14	Oyster Bar	Seafood Restaurant	1 Faneuil Hall Market Pl	275	42.360182	-71.054343
15	Pulcinella Mozzarella Bar and Restaurant	Italian Restaurant	78 Salem St	598	42.363693	-71.055872
16	Rabia's Seafood/Oyster Bar & Italian Restaurant	Italian Restaurant	73 Salem St	578	42.363377	-71.055977
17	Side Bar Food & Spirits	Bar	14 Bromfield St	798	42.356574	-71.059848
18	RumBa Rum Bar & Champagne Lounge	Bar	510 Atlantic Ave	658	42.353766	-71.051834

## 6. Discussion Section

Based on the definition of the original problem that is the location for the new business, the factors that will influence the decision are:

- No significant number of direct or indirect competition, ideally no similar type of venue
- Proximity to upscale city landmarks, like luxury shops, major hotel chains

The resulting dataframes and the mapping of the locations in the preferred investor's neighborhoods, Back Bay and Waterfront, present no head-to-head competition for the kind of venue the investors have in mind. There are however other traditional business establishments like restaurants and bars in these areas that should be taken into account when selecting the actual street location for the new venue.

One positive sign is that the chosen areas are not overcrowded with these existing businesses. It can be considered a low restaurant density area making it easier to find a spot at a reasonable distance from these potential competitors.

## 7. Conclusion

The conclusion based on the this preliminary analysis is that there is room for the Bistro & Wine Bar in the two selected areas, without colliding with any existing similar business. Each area has its own attractiveness with luxury shops and upscale entertainment and hospitality venues.

The actual real estate location short list will be decided based on other marketing considerations and the cost of the commercial real estate. These considerations are out of the scope of this present analysis.