

# Exploring for potential district in Barcelona for building a new hotel

## 1. Introduction

### a. Background

This report provides a suggestion on what would be the best venue to start a new hotel in a highly visited city with already many good hotels. Barcelona is one of the most popular tourist destinations in the world. Currently ranked in the Top 10 of most visited cities worldwide and Top 3 in Europe. Barcelona is definitely one of the best places to start up a new hotel business. In 2018, Barcelona received over 15,8 million visitors, especially in the summer and Christmas time, tourist areas in Barcelona provide huge opportunities for hotels. Even the most expensive hotels, like W Hotel, have 113% of their occupancy rated year in year out, given this scenario, we will go through the benefits and pitfalls of opening a new Hotel in a highly visited city with already many hotels offering their services. The core of Barcelona is made of 10 districts but, I will later concentrate on only 3 districts, where the busiest venues of Barcelona can be found to target the tourists visiting the city. With that in mind, I will be able to find the Top 3 districts to open a brand new hotel in the city.

### b. Business Problem

This report focusses on the issue of where to open a new hotel in a city like Barcelona, once one has decided to go ahead.

Now we will try to find the answers to the following questions:

- Which info that STARHOTEL should use to choose the district of their new hotel? Which specific district is that?
- What is the district in the 2nd place that can replace the 1st one we find out (if needed)
- The related hotels in the same area those could affect and be the competitors of our new hotel

## 2. Data Preparation

### a. Scrapping Barcelona Districts Table from Wikipedia

Extracting cells into a dataframe

Calculating Latitude and Longitude per District

Clustering Barcelona's districts

## Clustering Barcelona's districts

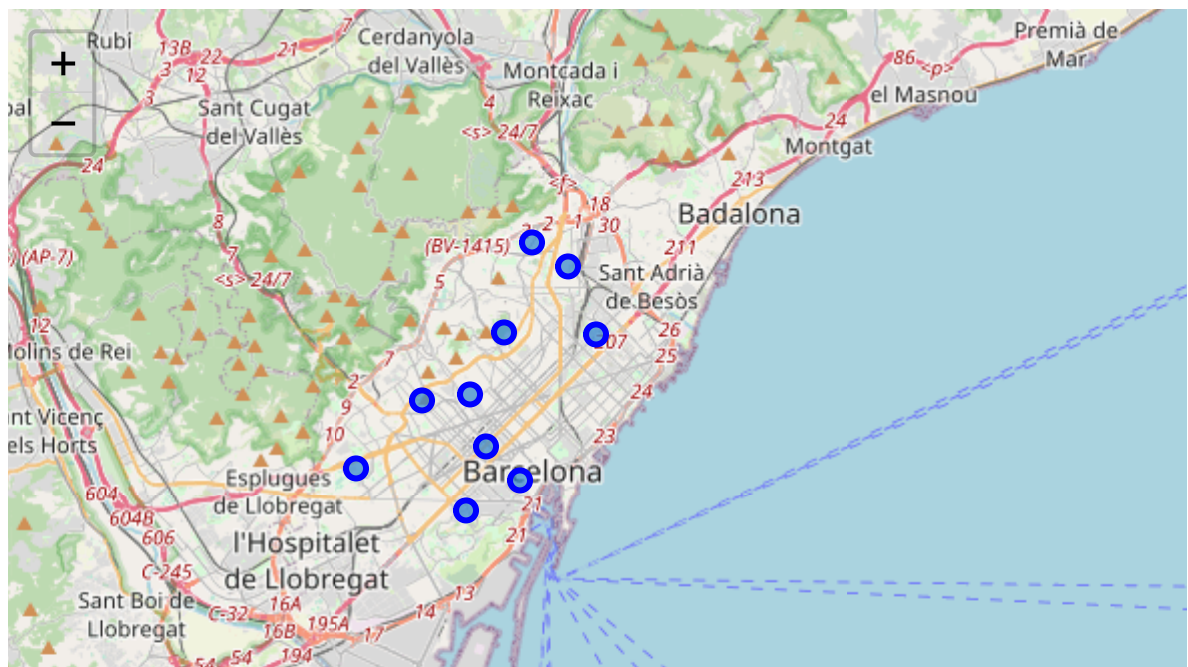
### Mapping Barcelona

```
In [183]: # create map of BARCELONA using Latitude and Longitude values above:
map_barcelona = folium.Map(location=[latitude, longitude], zoom_start=11)

# add markers to map
for lat, lng, label in zip(districts['Latitude'], districts['Longitude'], districts['Label']):
    popup = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=popup,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_barcelona)

map_barcelona
```

Out[183]:



### Foursquare API to explore and segment

**Writing the code to run the function on each neighborhood and create a new dataframe called *Barcelona\_venues*.**

### Analyzing Districts

**Grouping rows by district and by the mean of the frequency of occurrence of each category**

## Printing districts along with the top 3 most common venues

```
In [290]: Barcelona_t = Barcelona_grouped.T
for row in Barcelona_grouped.iterrows():
    print("-----")
    x= (row[1]).sort_values(ascending=False)
    print x.iloc[0]
    print("-----")
    print x.index[1]
    print x.iloc[1]
    print x.index[2]
    print x.iloc[2]
    print x.index[3]
    print x.iloc[3]
```

```
-----
Ciutat Vella
-----
Spanish Restaurant
0.11
Tapas Restaurant
0.07
Hotel
0.05
-----
Eixample
-----
Hotel
0.16
Spanish Restaurant
0.06
Bookstore
0.05
-----
Gràcia
-----
Mediterranean Restaurant
0.06
Pizza Place
0.05
Tapas Restaurant
0.05
-----
Horta-Guinardó
-----
Bar
0.222222222222
Tapas Restaurant
0.111111111111
Playground
0.111111111111
-----
Les Corts
-----
Soccer Field
0.103448275862
Tram Station
0.0689655172414
```

Nightclub  
0.0689655172414  
-----  
Nou Barris  
-----  
Spanish Restaurant  
0.172413793103  
Grocery Store  
0.103448275862  
Tapas Restaurant  
0.0689655172414  
-----  
Sant Andreu  
-----  
Spanish Restaurant  
0.0975609756098  
Tapas Restaurant  
0.0975609756098  
Park  
0.0731707317073  
-----  
Sant Martí  
-----  
Coffee Shop  
0.1111111111111  
Tapas Restaurant  
0.1111111111111  
Pizza Place  
0.0740740740741  
-----  
Sants-Montjuïc  
-----  
Tapas Restaurant  
0.07  
Spanish Restaurant  
0.07  
Restaurant  
0.06  
-----  
Sarrià-Sant Gervasi  
-----  
Bar  
0.0864197530864  
Hotel  
0.0864197530864  
Bakery  
0.0864197530864

```
In [292]: num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['District']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
districts_venues_sorted = pd.DataFrame(columns=columns)
districts_venues_sorted['District'] = Barcelona_grouped['District']

for ind in np.arange(Barcelona_grouped.shape[0]):
    districts_venues_sorted.iloc[ind, 1:] = return_most_common_venues(Barcelona_grouped, ind+1, num_top_venues)

districts_venues_sorted
```

Out[292]:

	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
0	Ciutat Vella	Spanish Restaurant	Tapas Restaurant	Hotel	Mediterranean Restaurant	Ice Cream Shop	Plaza	Italian Restaurant
1	Eixample	Hotel	Spanish Restaurant	Bookstore	Tapas Restaurant	Mediterranean Restaurant	Coffee Shop	Cosmetic Shop
2	Gràcia	Mediterranean Restaurant	Pizza Place	Tapas Restaurant	Plaza	Restaurant	Bakery	Bar
3	Horta-Guinardó	Bar	Park	Playground	Café	Tapas Restaurant	Gym	Chinese Restaurant
4	Les Corts	Soccer Field	Tram Station	Nightclub	Hockey Arena	Museum	Garden	Bowling Alley
5	Nou Barris	Spanish Restaurant	Grocery Store	Plaza	Tapas Restaurant	Park	Café	Pizza Place
6	Sant Andreu	Spanish Restaurant	Tapas Restaurant	Park	Café	Grocery Store	Coffee Shop	Bar
7	Sant Martí	Tapas Restaurant	Coffee Shop	Soccer Field	Pizza Place	Supermarket	Bakery	Grocery Store
8	Sants-Montjuïc	Spanish Restaurant	Tapas Restaurant	Restaurant	Cocktail Bar	Mediterranean Restaurant	Pizza Place	Plaza
9	Sarrià-Sant Gervasi	Bar	Hotel	Bakery	Japanese Restaurant	Italian Restaurant	Café	Grocery Store

Clustering districts

**New dataframe for clusters and top 10 venues for each district****Clustering Districts**

In [242]: `from sklearn.cluster import KMeans`

In [293]: `Barcelona_merged = districts`  
*# add clustering labels*  
`Barcelona_merged['Cluster Labels'] = kmeans.labels_`  
*# merge Barcelona\_grouped with Barcelona\_data to add latitude/longitude for each*  
`Barcelona_merged = Barcelona_merged.join(districts_venues_sorted.set_index('District', inplace=True))`  
`Barcelona_merged.head(11) # check the last columns!`

Out[293]:

	District	Size km <sup>2</sup>	Population	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Ciutat Vella	4.49	111,290	41.380218	2.17319	0	Spanish Restaurant	Tapas Restaurant	Hotel
1	Eixample	7.46	262,485	41.388960	2.16179	0	Hotel	Spanish Restaurant	Bookstore
2	Gràcia	4.19	120,087	41.402371	2.15641	0	Mediterranean Restaurant	Pizza Place	Tapas Restaurant
3	Horta-Guinardó	11.96	169,920	41.418491	2.16770	1	Bar	Park	Playground
4	Les Corts	6.08	82,588	41.383500	2.11680	0	Soccer Field	Tram Station	Nightclub
5	Nou Barris	8.04	164,981	41.441631	2.17727	2	Spanish Restaurant	Grocery Store	Plaza
6	Sant Andreu	6.56	142,598	41.435409	2.18982	2	Spanish Restaurant	Tapas Restaurant	Park
7	Sant Martí	10.80	221,029	41.418140	2.19933	2	Tapas Restaurant	Coffee Shop	Soccer Field
8	Sants-Montjuïc	21.35	177,636	41.372631	2.15460	0	Spanish Restaurant	Tapas Restaurant	Restaurant
9	Sarrià-Sant Gervasi	20.09	140,461	41.401039	2.13940	0	Bar	Hotel	Bakery

```
In [295]: # set number of clusters
kclusters = 5

Barcelona_grouped_clustering = Barcelona_grouped.drop('District', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(Barcelona_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[1:10]
```

```
Out[295]: array([2, 2, 3, 0, 1, 1, 4, 2, 2])
```

### Clusters visualization of Barcelona's Districts



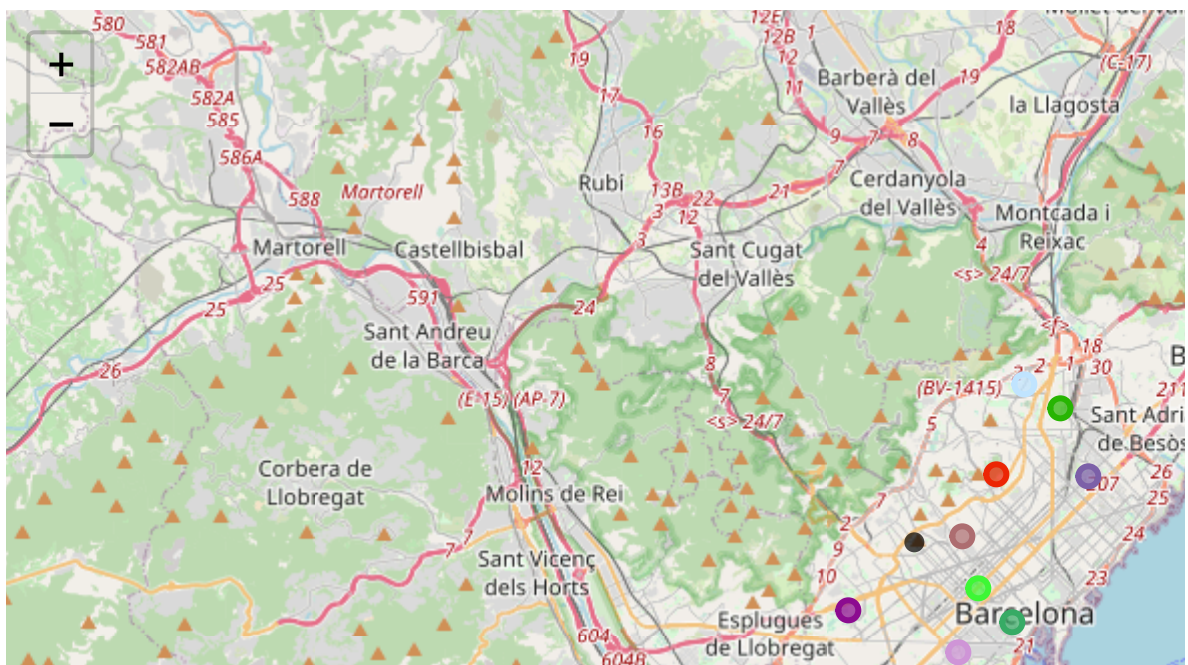
```

In [328]: map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i+x+(i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = ['#37AB65', '#3DF735', '#AD6D70', '#EC2504', '#8C0B90', '#C0E4FF', '#271
#print rainbow
# add markers to the map
markers_colors = []
i = 0
for lat, lon, poi in zip(Barcelona_merged['Latitude'], Barcelona_merged['Longitud
# #label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        color=rainbow[i],
        fill=True,
        fill_color=rainbow[i],
        fill_opacity=0.7).add_to(map_clusters)
    i = i + 1
map_clusters

```

Out[328]:



```
In [337]: Barcelona_merged = districts

# add clustering labels
Barcelona_merged['Cluster Labels'] = kmeans.labels_

Barcelona_merged = Barcelona_merged.join(districts_venues_sorted.set_index('District', level=0))

Barcelona_merged.head(11)
```

Out[337]:

	District	Size km <sup>2</sup>	Population	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Ciutat Vella	4.49	111,290	41.380218	2.17319	2	Spanish Restaurant	Tapas Restaurant	Hotel
1	Eixample	7.46	262,485	41.388960	2.16179	2	Hotel	Spanish Restaurant	Bookstore
2	Gràcia	4.19	120,087	41.402371	2.15641	2	Mediterranean Restaurant	Pizza Place	Tapas Restaurant
3	Horta-Guinardó	11.96	169,920	41.418491	2.16770	3	Bar	Park	Playground
4	Les Corts	6.08	82,588	41.383500	2.11680	0	Soccer Field	Tram Station	Nightclub
5	Nou Barris	8.04	164,981	41.441631	2.17727	1	Spanish Restaurant	Grocery Store	Plaza
6	Sant Andreu	6.56	142,598	41.435409	2.18982	1	Spanish Restaurant	Tapas Restaurant	Park
7	Sant Martí	10.80	221,029	41.418140	2.19933	4	Tapas Restaurant	Coffee Shop	Soccer Field
8	Sants-Montjuïc	21.35	177,636	41.372631	2.15460	2	Spanish Restaurant	Tapas Restaurant	Restaurant
9	Sarrià-Sant Gervasi	20.09	140,461	41.401039	2.13940	2	Bar	Hotel	Bakery

## Cluter result

**Most common venues: Restaurants, Hotel, Plaza and Drinking places**

```
In [343]: Barcelona_merged.loc[Barcelona_merged['Cluster Labels'] == 2, Barcelona_merged.co
```

Out[343]:

	District	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
0	Ciutat Vella	2	Spanish Restaurant	Tapas Restaurant	Hotel	Mediterranean Restaurant	Ice Cream Shop	Plaza
1	Eixample	2	Hotel	Spanish Restaurant	Bookstore	Tapas Restaurant	Mediterranean Restaurant	Coffee Shop
2	Gràcia	2	Mediterranean Restaurant	Pizza Place	Tapas Restaurant	Plaza	Restaurant	Bakery
8	Sants-Montjuïc	2	Spanish Restaurant	Tapas Restaurant	Restaurant	Cocktail Bar	Mediterranean Restaurant	Pizza Place
9	Sarrià-Sant Gervasi	2	Bar	Hotel	Bakery	Japanese Restaurant	Italian Restaurant	Cafe

Most common venues: Restaurants, Cafe and Park - less common in Hotel

```
In [339]: Barcelona_merged.loc[Barcelona_merged['Cluster Labels'] == 1, Barcelona_merged.co
```

Out[339]:

	District	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
5	Nou Barris	1	Spanish Restaurant	Grocery Store	Plaza	Tapas Restaurant	Park	Café	Pizza Place	Restaurant
6	Sant Andreu	1	Spanish Restaurant	Tapas Restaurant	Park	Café	Grocery Store	Coffee Shop	Bar	

Most common venues: Bar , Park, Playground - less common in Hotel

```
In [344]: Barcelona_merged.loc[Barcelona_merged['Cluster Labels'] == 3, Barcelona_merged.co
```

Out[344]:

	District	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
3	Horta-Guinardó	3	Bar	Park	Playground	Café	Tapas Restaurant	Gym	Chinese Restaurant	B

**Most common venues: Restaurants, Coffee Grocery stores with less common in Hotel**

```
In [345]: Barcelona_merged.loc[Barcelona_merged['Cluster Labels'] == 4, Barcelona_merged.co
```

Out[345]:

	District	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
7	Sant Martí	4	Tapas Restaurant	Coffee Shop	Soccer Field	Pizza Place	Supermarket	Bakery	Grocery Store	Ba

**District with less common in restaurant, hotel and bar**

```
In [338]: Barcelona_merged.loc[Barcelona_merged['Cluster Labels'] == 0, Barcelona_merged.co
```

Out[338]:

	District	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
4	Les Corts	0	Soccer Field	Tram Station	Nightclub	Hockey Arena	Museum	Garden	Bowling Alley	Scien Museum

## Results

## **With the clustering result, we can divide the 10 district of Barcelona into 5 different group:**

1. "The crowded districts" those are popular with a lot of attraction like Restaurants, Hotel, Plaza and Drinking places
2. The districts with Restaurants, Cafe and Park but less common in Hotel
3. "The peaceful" one with Bar, Park, Playground but less common in Hotel
4. "The busy district" which Restaurant, Cafe, Sport area is popular but Hotel is not in the top 10
5. "The student/sporty district" which Hotel and Restaurant don't appear in top 10 but most are Sport areas

## **Discussion and Conclusion**

Base on the type of the Hotel, the business should choose the group that suitable for them. Let have some discussion about these groups:

1. The first group seems to be the best choice if the business are confident to compete the other hotels. The common places are Restaurant, Plaza and Drinking places, which makes the district in this group become potential place with lots of people
2. The 2nd and 3rd group have less common in Hotel. Our business should consider to build a hotel here and it can be special here.
3. The 2 last groups seem to be places for young people, especially students, so may be building hotel there is not a good idea