

# **Deciding if the location for a congress is suitable or not**

Georgina Company

## **1. Introduction**

### **1.1. Background**

Every year and every month many congresses are celebrated around the world. Congresses cover different aspects about medicine, science, engineering, social science, etc. They represent a great opportunity to share knowledge about a specific topic. Moreover, congresses represent an opportunity for young researchers who want to be known in the field of expertise for other people of the same area of work. A good congress, not only needs professional speakers but also needs people wanting to attend to the congress. And what makes people want to attend a congress? There are different aspects to be taken into consideration. One essential aspect is that good accommodations near the congress are a must.

### **1.2. Problem**

The Polytechnic University of Catalonia wants to host an international well-known congress. It will suppose the opportunity for the university to get known around the world. Another university interested in hosting the congress is the University of Barcelona. The organization committee of the congress needs to know what location is more suitable for celebrating the congress. So, **“which is the most suitable university to host the congress?”**.

### **1.3. Interest**

This methodology of analysis is interesting for the organization committee of the congress in order to decide which location is more suitable for the congress, which in turn will represent more money for the congress as more people will want to attend to it.

## 2. Data acquisition and cleaning

First, we import and install all the libraries that we will need in order to perform our project.

```
import requests # library to handle requests
import pandas as pd # library for data analysis
import numpy as np # library to handle data in a vectorized manner
import random # library for random number generation

!pip install geopy
# module to convert an address into latitude and longitude values
from geopy.geocoders import Nominatim

# libraries for displaying images
from IPython.display import Image
from IPython.core.display import HTML

# transforming json file into a pandas dataframe library
from pandas.io.json import json_normalize

! pip install folium==0.5.0
import folium # plotting library

print('Folium installed')
print('Libraries imported.')
```

### 2.1.Data sources

In order to solve the problem, I am going to use Foursquare data, as I have learnt from previous assignments and weeks. In order to retrieve data, I will define my credentials.

```
CLIENT_ID = 'CKDWP5VJMC3VBZKCUEV4NUQHKRIQ4WL1HK0CHNGCGGY12SXL' # your Foursquare ID
CLIENT_SECRET = 'Q2W2QUMC5FQHZYOICKFXL0ZE1IJHFVRJ4YWAL3FAA4YNXIBN' # your Foursquare Secret
ACCESS_TOKEN = 'LNRTSWVI122MIV03PRDURET2WMHQDN0TOJBCZF11A0OI4TKB' # your FourSquare Access Token
VERSION = '20180604'
LIMIT = 30
print('Your credentials:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET: ' + CLIENT_SECRET)

Your credentials:
CLIENT_ID: CKDWP5VJMC3VBZKCUEV4NUQHKRIQ4WL1HK0CHNGCGGY12SXL
CLIENT_SECRET: Q2W2QUMC5FQHZYOICKFXL0ZE1IJHFVRJ4YWAL3FAA4YNXIBN
```

Then I extract the location information of the two possible places to attend the congress, Polytechnic University of Catalonia and University of Barcelona.

```
address1 = 'C. Jordi Girona, 3208034 Barcelona Cataluña España'
address2 = 'Gran Via de les Corts Catalanes, 585, 08007 Barcelona Cataluña España'

geolocator1 = Nominatim(user_agent="foursquare_agent")
location1 = geolocator1.geocode(address1)
latitude1 = location1.latitude
longitude1 = location1.longitude
print('Location of Polytechnic University of Catalonia:',latitude1, longitude1)

geolocator2 = Nominatim(user_agent="foursquare_agent")
location2 = geolocator2.geocode(address2)
latitude2 = location2.latitude
longitude2 = location2.longitude
print('Location of University of Barcelona:',latitude2, longitude2)
```

```
Location of Polytechnic University of Catalonia: 41.3877523 2.1147272
Location of University of Barcelona: 41.3866985 2.163828091253939
```

After obtaining the location, we extract information about the hotels that are near the two possible locations.

```

search_query = 'Hotel'
radius = 1000

url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&oauth_token={}&v={}&query={}&radius={}'
url = url.format(search_query, radius)
results = requests.get(url).json()
hotels1 = results['response']['venues']
dataframe1 = json_normalize(hotels1)

url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&oauth_token={}&v={}&query={}&radius={}'
url = url.format(search_query, radius)
results = requests.get(url).json()
hotels2 = results['response']['venues']
dataframe2 = json_normalize(hotels2)

```

/srv/conda/envs/notebook/lib/python3.6/site-packages/ipykernel\_launcher.py:8: FutureWarning: pandas.io.json.json\_normalize is deprecated, use pandas.json\_normalize instead  
/srv/conda/envs/notebook/lib/python3.6/site-packages/ipykernel\_launcher.py:14: FutureWarning: pandas.io.json.json\_normalize is deprecated, use pandas.json\_normalize instead

In order to see the results, we visualize the first hotels in the list of the two possible locations.

For the Polytechnic University of Catalonia:

dataframe1.head()									
	id	name	categories	referralId	hasPerk	location.address	location.lat	location.lng	location.labeledLatLngs
0	4bc5a24b0a30d13a94385a9c	AC Hotel Victoria Suites by Marriott	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Beltran I Ròzpide, 7-9	41.389324	2.121338	[{'label': 'display', 'lat': 41.389324, 'lng': ...}]
1	4adcd4cf964a520b93f21e3	Gran Hotel Princesa Sofia	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Plaça de Pius XII, 4	41.386517	2.123076	[{'label': 'display', 'lat': 41.38651744622853, 'lng': ...}]
2	4bcc5f5baaeae1a1f23d6d	Hotel Sansi Pedralbes	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Av. Pearson, 1-3	41.393435	2.111722	[{'label': 'display', 'lat': 41.39343506028381, 'lng': ...}]
3	5d30def44318c90008b07cba	Hotel Sofia Barcelona - In the Unbound Collect...	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Plaça De Pius XII 4	41.385980	2.123314	[{'label': 'display', 'lat': 41.38598, 'lng': ...}]
4	4e38aa161fc7e413a51a4788	Fairmont Hotel Rey Juan Carlos I (Hotel Rey Ju...	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Av Diagonal 661 671	41.381291	2.109191	[{'label': 'display', 'lat': 41.381291, 'lng': ...}]

For the University of Barcelona:

dataframe2.head()									
	id	name	categories	referralId	hasPerk	location.address	location.lat	location.lng	location.labeledLatLngs
0	4adcd4cf964a520673f21e3	Hotel NH Collection Barcelona Gran Hotel Calderón	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Rbla. de Catalunya, 26	41.389264	2.166129	[{'label': 'display', 'lat': 41.38926386465471, 'lng': ...}]
1	4b39cb3df964a520015f25e3	Hotel Pulitzer	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	C. Bergara, 8	41.386245	2.168272	[{'label': 'display', 'lat': 41.38624508838907, 'lng': ...}]
2	4b127c3af964a5202c8a23e3	Axel Hotel	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Aribau, 33	41.387351	2.160403	[{'label': 'display', 'lat': 41.38735095471836, 'lng': ...}]
3	4adcd4cf964a5208e3f21e3	Hotel Catalonia Plaza Catalunya	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	C. de Bergara, 11	41.386406	2.168230	[{'label': 'display', 'lat': 41.38640606572736, 'lng': ...}]
4	4dced5db183899ddfadd35eb	Hotel Condestable	['id': '4bf58dd8d48988d1fa931735', 'name': 'H...']	v-1625649193	False	Plaça Universitat, 1	41.386140	2.164967	[{'label': 'display', 'lat': 41.38614013259076, 'lng': ...}]

## 2.2.Data cleaning

We are going to keep only the data that is associated with the location of the different hotels as well as the hotel's names for both locations.

```
filtered_columns1 = ['name', 'categories'] + [col for col in dataframe1.columns if col.startswith('location.')] + ['id']
dataframe_filtered1 = dataframe1.loc[:, filtered_columns1]
def get_category_type(row):
    try:
        categories_list1 = row['categories']
    except:
        categories_list1 = row['hotels1.categories']

    if len(categories_list1) == 0:
        return None
    else:
        return categories_list1[0]['name']
dataframe_filtered1['categories'] = dataframe_filtered1.apply(get_category_type, axis=1)
dataframe_filtered1.columns = [column.split('.')[0] for column in dataframe_filtered1.columns]

filtered_columns2 = ['name', 'categories'] + [col for col in dataframe2.columns if col.startswith('location.')] + ['id']
dataframe_filtered2 = dataframe2.loc[:, filtered_columns2]
def get_category_type(row):
    try:
        categories_list2 = row['categories']
    except:
        categories_list2 = row['hotels2.categories']

    if len(categories_list2) == 0:
        return None
    else:
        return categories_list2[0]['name']
dataframe_filtered2['categories'] = dataframe_filtered2.apply(get_category_type, axis=1)
dataframe_filtered2.columns = [column.split('.')[0] for column in dataframe_filtered2.columns]
```

The results are shown below:

For the Polytechnic University of Catalonia:

dataframe_filtered1.head()														
	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	crossSt
0	AC Hotel Victoria Suites by Marriott	Hotel	Beltran I Rózpide, 7-9	41.389324	2.121338	[{"label": "display", "lat": 41.389324, "lng": ...}	579	08034	ES	Barcelona	Cataluña	España	[Beltran I Rózpide, 7-9, 08034 Barcelona Cataluña]	
1	Gran Hotel Princesa Sofia	Hotel	Plaça de Pius XII, 4	41.386517	2.123076	[{"label": "display", "lat": 41.386517, "lng": ...}	710	08028	ES	Barcelona	Cataluña	España	[Plaça de Pius XII, 4 (Av. Diagonal), 08028 Barcelona Cataluña]	Diag
2	Hotel Sansi Pedralbes	Hotel	Av. Pearson, 1-3	41.393435	2.111722	[{"label": "display", "lat": 41.393435, "lng": ...}	680	NaN	ES	Barcelona	Cataluña	España	[Av. Pearson, 1-3, Barcelona Cataluña]	
3	Hotel Sofia Barcelona - In the Unbound Collect...	Hotel	Plaça De Pius XII 4	41.385980	2.123314	[{"label": "display", "lat": 41.38598, "lng": ...}	743	08028	ES	Barcelona	Cataluña	España	[Plaça De Pius XII 4, 08028 Barcelona Cataluña]	
4	Fairmont Hotel Rey Juan Carlos I (Hotel Rey Ju...	Hotel	Av Diagonal 661 671	41.381291	2.109191	[{"label": "display", "lat": 41.381291, "lng": ...}	855	08028	ES	Barcelona	Cataluña	España	[Av Diagonal 661 671, 08028 Barcelona Cataluña]	

For the University of Barcelona:

```
dataframe_filtered2.head()
```

	name	categories	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	crc
0	Hotel NH Collection Barcelona Gran Hotel Calderón	Hotel	Rbla. de Catalunya, 26	41.389264	2.166129	[{"label": "display", "lat": 41.38926386465471...	344	08007	ES	Barcelona	Cataluña	España	[Rbla. de Catalunya, 26, 08007 Barcelona Catal...	
1	Hotel Pulitzer	Hotel	C. Bergara, 8	41.386245	2.168272	[{"label": "display", "lat": 41.38624508838907...	374	08002	ES	Barcelona	Cataluña	España	[C. Bergara, 8, 08002 Barcelona Cataluña]	
2	Axel Hotel	Hotel	Aribau, 33	41.387351	2.160403	[{"label": "display", "lat": 41.38735095471836...	295	08011	ES	Barcelona	Cataluña	España	[Aribau, 33, 08011 Barcelona Cataluña]	
3	Hotel Catalonia Plaza Catalunya	Hotel	C. de Bergara, 11	41.386406	2.168230	[{"label": "display", "lat": 41.38640606572736...	369	08002	ES	Barcelona	Cataluña	España	[C. de Bergara, 11 (Pl. de Catalunya), 08002 B...	
4	Hotel Condestable	Hotel	Plaça Universitat, 1	41.386140	2.164967	[{"label": "display", "lat": 41.38614013259076...	113	08007	ES	Barcelona	Cataluña	España	[Plaça Universitat, 1, 08007 Barcelona Cataluña]	

As we can see the cleaning has succeeded as only information related with location is maintained.

We can also extract the list of the hotels that are near the two locations.

For the Polytechnic University of Catalonia:

```
dataframe_filtered1.name
```

0	AC Hotel Victoria Suites by Marriott
1	Gran Hotel Princesa Sofia
2	Hotel Sansi Pedralbes
3	Hotel Sofia Barcelona - In the Unbound Collect...
4	Fairmont Hotel Rey Juan Carlos I (Hotel Rey Ju...
5	Gimnasio del Gran Hotel Princesa Sofia
6	Hotel Catalonia Rigoletto ****
7	Pool Hotel Rey Juan Carlos I (Piscina Hotel Re...
8	We Barcelona Hotel
9	abba Garden hotel 4*
10	Senator Barcelona Spa Hotel
11	Hotel Upper Diagonal
12	Hotel Arenas
13	Hotel Ginosi Pedralbes
14	Hotel Alguer Camp Nou
15	Hotel L'Alguer
16	Junior Suite @Gran Hotel Princesa
17	Salon de Peluquería Hotel Rey Juan Carlos I **...
18	Husa Hotel
19	Swimming Pool Hotel Abba Garden
20	33rd Gastro-bar
21	Espinaler Experience Tapas Bar
22	5o. Piso Senator Hoteles

Name: name, dtype: object

For the University of Barcelona:

```
dataframe_filtered2.name
0      Hotel NH Collection Barcelona Gran Hotel Calderón
1                                          Hotel Pulitzer
2                                  Axel Hotel
3      Hotel Catalonia Plaza Catalunya
4                                  Hotel Condestable
5      Hotel HLG CityPark Pelayo**
6                                  Hotel Regina
7      Hotel Atlantis by Atbcn
8                                  Hotel Lleó
9                                  Hotel Reding
10     Hotel Petit Palace Museum
11     Hotel Ciutat Vella Barcelona
12     Hotel Acta Splendid
13     Hotel Catalonia Ramblas
14     Hotel Praktik Rambla
15     Hotel Caledonian
16     Hotel Room Mate Anna
17     Atlantis Hotel Barcelona
18     Hotel Cram
19     Hotel Gran Ronda
20     Hotel Casanova by RafaelHotels
21     Hotel Praktik Garden
22     Hotel América
23     Hotel Advance
24     Hotel Eurostars Cristal Palace
25     Hotel Ohla
26     Hotel Acta BCN 40
27     Olivia Plaza Hotel
28     Hotel Medium Monegal
29     Hotel Neri
Name: name, dtype: object
```

In order to make us an idea of the possible hotels nearer the two locations we can display them in a map.

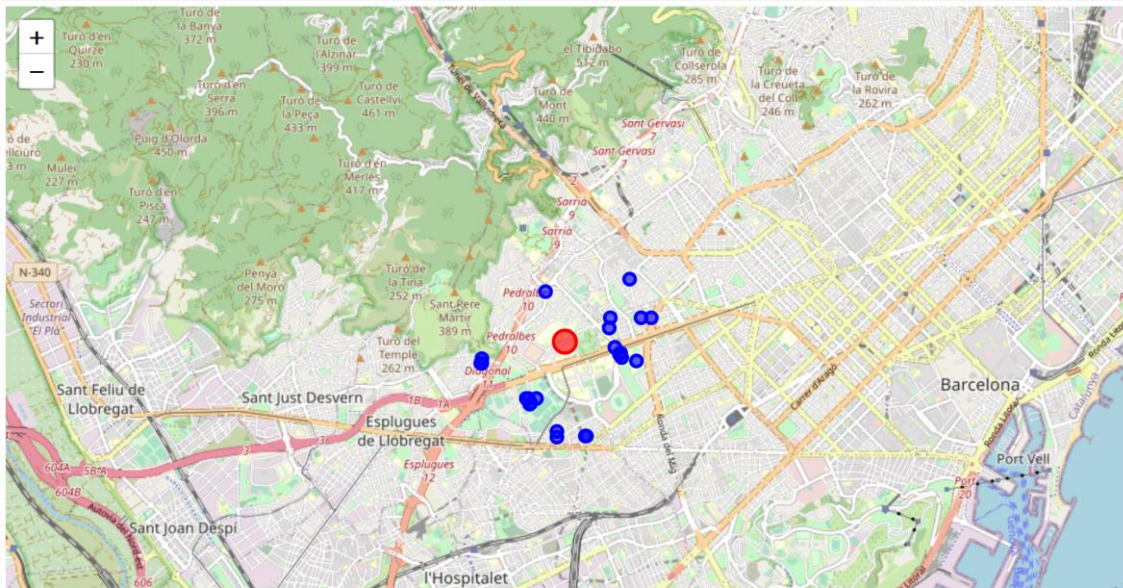
For the Polytechnic University of Catalonia:

```
hotels_map1 = folium.Map(location=[latitude1, longitude1], zoom_start=13)
folium.CircleMarker(
    [latitude1, longitude1],
    radius=10,
    color='red',
    popup='Conrad Hotel',
    fill = True,
    fill_color = 'red',
    fill_opacity = 0.6
).add_to(hotels_map1)
for lat, lng, label in zip(dataframe_filtered1.lat, dataframe_filtered1.lng, dataframe_filtered1.categories):
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        color='blue',
        popup=label,
        fill = True,
        fill_color='blue',
        fill_opacity=0.6
    ).add_to(hotels_map1)

hotels_map1
```



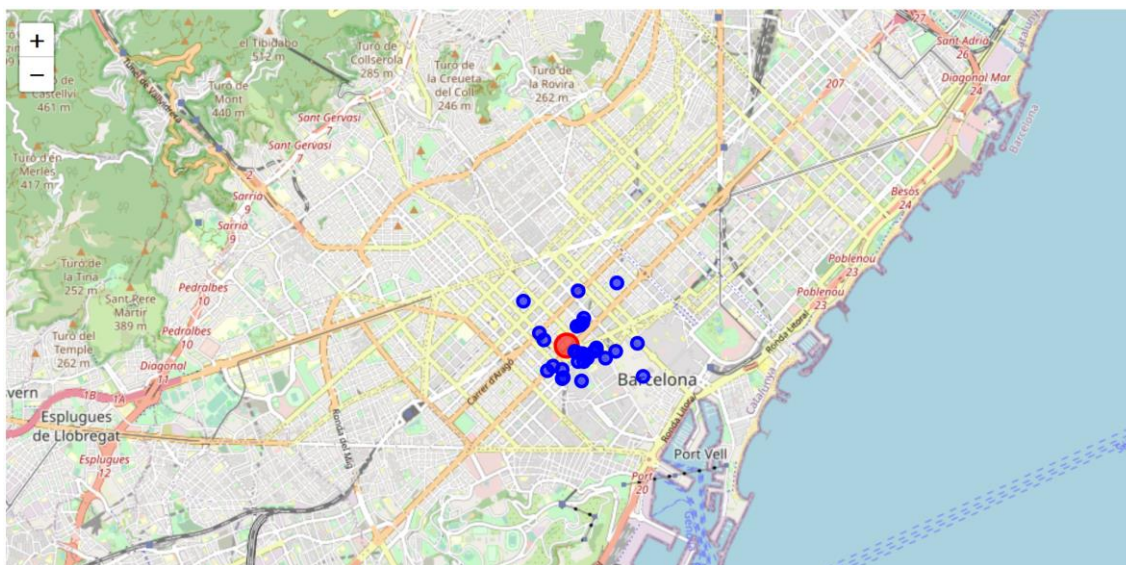
hotels\_map1



For the University of Barcelona:

```
hotels_map2 = folium.Map(location=[latitude2, longitude2], zoom_start=13)
folium.CircleMarker(
    [latitude2, longitude2],
    radius=10,
    color='red',
    popup='Conrad Hotel',
    fill = True,
    fill_color = 'red',
    fill_opacity = 0.6
).add_to(hotels_map2)
for lat, lng, label in zip(dataframe_filtered2.lat, dataframe_filtered2.lng, dataframe_filtered2.categories):
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        color='blue',
        popup=label,
        fill = True,
        fill_color='blue',
        fill_opacity=0.6
    ).add_to(hotels_map2)
```

hotels\_map2



As we can see in the map, the location of the University of Barcelona seems to be better as it is more centric than the other possible location.

### 3. Exploratory data analysis and feature selection

We start by comparing the two nearest hotels to each of the possible locations.

```
distance_1 = min(dataframe_filtered1.distance)
closest_hotel_1 = dataframe_filtered1.name[dataframe_filtered1.distance == distance_1]
distance_2 = min(dataframe_filtered2.distance)
closest_hotel_2 = dataframe_filtered2.name[dataframe_filtered2.distance == distance_2]

print('The closest hotel at the Polytechnic University of Catalonia is',closest_hotel_1, 'at a distance of',distance_1)
print('The closest hotel at the University of Barcelona is',closest_hotel_2, 'at a distance of',distance_2)

The closest hotel at the Polytechnic University of Catalonia is 0    AC Hotel Victoria Suites by Marriott
Name: name, dtype: object at a distance of 579
The closest hotel at the University of Barcelona is 4    Hotel Condestable
Name: name, dtype: object at a distance of 113
```

As we can see the nearest hotel in the Polytechnic University of Catalonia is the AC Hotel while the nearest in the University of Barcelona is the Condestable Hotel. Moreover we can see that the distance is less in the second case. It seems that there are more hotels nearer the university of Barcelona than the Polytechnic.

**We aim to extract the distances and make a comparison. Moreover, we aim to find the ratings of the hotels in order to conclude which hotels are more suitable for the attendants.**

### 4. Results

In the figure below, we can see the distances for both locations of the hotels.

```
from matplotlib import pyplot
import matplotlib.pyplot as plt

distances_1 = dataframe_filtered1.distance
distances_2 = dataframe_filtered2.distance
x1=dataframe_filtered1.name
x2=dataframe_filtered2.name

plt.xticks(rotation='vertical')
matplotlib.pyplot.bar(x1,distances_1)

plt.xticks(rotation='vertical')
matplotlib.pyplot.bar(x2,distances_2)

plt.title('Distance of the hotels for both locations')
plt.ylabel('Distance')
```





Another aspect to consider is the quality of the hotels. I extract the “**rating**” feature of each of the hotels for both locations following the code below and changing each time the rating1 by rating2 and the dataframe\_filtered1 by dataframe\_filtered2.

```
rating1=[]
idl =dataframe_filtered1.id
for i in range (len(idl)):
    url1 = 'https://api.foursquare.com/v2/venues/{}?client_id={}&client_secret={}&oauth_token={}&v={}'.format(idl[i], C
    result = requests.get(url1).json()
    result['response']['venue']
    try:
        print(result['response']['venue']['rating'])
        rating1.append(result['response']['venue']['rating'])
    except:
        print('This hotel has not been rated yet.')
```

After that we print the ratings for each of the locations. It has to be said that some of the hotels have not been valued. We take into consideration all the hotels that has a rating parameter.

```
print(rating1)
```

```
[6.4, 6.4, 6.0, 6.1, 5.9, 5.9, 5.4, 5.2]
```

```
print(rating2)
```

```
[7.4, 8.4, 6.8, 7.3, 5.8, 6.8, 6.4, 6.6, 6.9, 6.4, 5.9, 6.3, 7.5, 6.3, 8.5, 5.4, 7.9, 5.2, 7.0, 7.7, 7.5, 6.9]
```

The mean of the ratings for each of the locations is:

```
mean_r1= np.mean(rating1)
mean_r2= np.mean(rating2)
print('Mean of the rating of the hotels of the Polytechnic University of Catalonia is:', mean_r1)
print('Mean of the rating of the hotels of the University of Barcelona is:', mean_r2)
```

```
Mean of the rating of the hotels of the Polytechnic University of Catalonia is: 5.9125
Mean of the rating of the hotels of the University of Barcelona is: 6.859090909090909
```

## 5. Discussion

After performing the analysis we can see that both locations offer different options of hotels near the universities. However, if we look at the maps, the location of the University of Barcelona is more centric than the location of the Polytechnic University of Catalonia. Moreover, we can see in the bar graphic that the distances from the hotels to each of the locations are smaller in the case of the University of Barcelona. Furthermore, the rating of the hotels is higher in the case of the University of Barcelona. The committee in charge of organizing the congress has to value all these aspects.

## 6. Conclusions

In conclusion, with all the results presented I should **consider to perform the congress in the University of Barcelona** as all the aspects valued are better in this case.