Finding the best location for a new hotel in Amsterdam

Gijs Hendrix

IBM Data Science certificate – Capstone Project

Contents

Introduction

Background and problem

I want to open a new hotel in Amsterdam, my hometown. It historic center, museums, restaurants and bars attract a large amount of tourist from both The Netherlands as abroad. As a 30-something-year old myself I will look for locations that would fit a tourist like myself. My interests include restaurants and bars and in this study I will focus on these. Additionally I want to make sure my hotel is unique and attracts customers.

Therefore there are two criteria I will use to select the location:

- Location versus competition. As the hotel market in Amsterdam is rather saturated I want to make sure to find a location without much competition.
- Vicinity to restaurants and bars. While the city is rather small and accessible my hotel needs to have a lot of restaurants and bars within walking distance.

Audience

The target audience is (potential) investors in hotels or existing hotel chains. This report will provide them with options for neighborhoods to open a new hotel in.

Data and methodology

Data sources

In order to answer the problem posed before I need data on the current hotel market in Amsterdam as well as information on restaurants and bars. After searching online I found two data sources that fulfill my requirements: Foursquare data on restaurants and bars and datasets from the city of Amsterdam, found on the following location: https://data.amsterdam.nl/. I found two sets:

- https://data.amsterdam.nl/datasets/smyIEaiLR6uRlg/. The dataset contains information on hotels within Amsterdam (name, address, star rating, amount of rooms and amount of beds). The data is unfortunately from 2014. I was not able to find a newer dataset. While in a real-life case a newer dataset might be required for this project I will continue with this one.
- https://data.amsterdam.nl/data/bag/adressen/?modus=volledig. This set also contains geographical data such as latitude and longitude.

Data cleaning

In order to use the data I import and use several libraries for this analysis:

- Pandas
- Matplotlib
- Seaborn
- Folium
- Requests
- Json and JSON normalize
- Numpy

Exploratory data analysis

Let's explore the different data sources. The hotels dataset looks as follows:

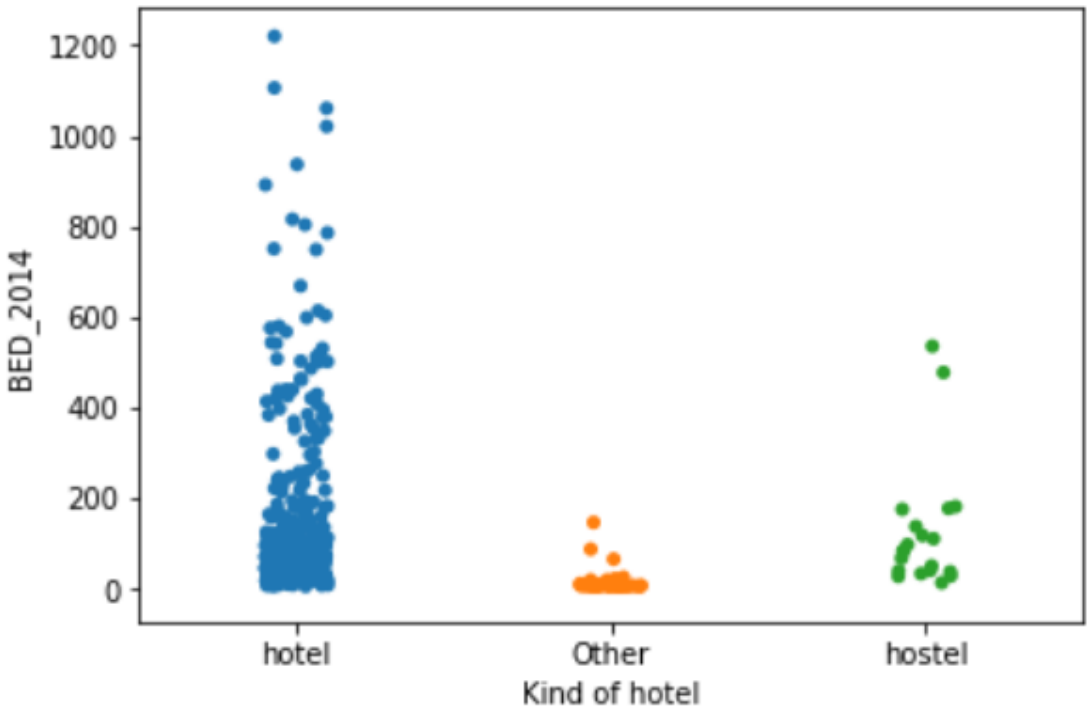
	HOTELNAAM_2014	NHC_2014	KAM_2014	BED_2014	STRAAT_2014	HUISID_2014	POSTCODE_2014	BRTK10	unesco	Soort_hotel_hostel/ jeugdherberg
417	B&B Apostrophe	0	3	6	Geerdinkhof	258	1103RA	T94j	NaN	B&B, pension, appartamenten, guesthouse, boere...
418	Hampton by Hilton Amsterdam Arena Boulevard	3	260	520	Hoekenrode	1	1102BR	T93e	NaN	hotel
419	Reijgersbroeck Bed & Breakfast	0	4	8	Provincialeweg	30	1108AB	T95b	NaN	B&B, pension, appartamenten, guesthouse, boere...
420	Brooklyn hotel	4	88	160	NDSM - Plein	28	1033WB	N71d	NaN	hotel
421	totaal	421 hotels	26287	56718	NaN	NaN	NaN	NaN	NaN	NaN

Let's see some descriptive statistics for the amount of rooms (KAM_2014) and beds (BED_2014):

	KAM_2014	BED_2014
count	421.000000	421.000000
mean	62.439430	134.722090
std	91.568082	188.811915
min	2.000000	5.000000
25%	13.000000	28.000000
50%	27.000000	65.000000
75%	67.000000	140.000000
max	610.000000	1220.000000

```
hotel      362
Other       39
hostel      20
```

There are 421 items in the database, 362 hotels, 20 hostels and 39 other venues. On average there are 62 rooms per item and 135 beds. The smallest venue has 5 beds, the largest 1220. The different venues seem to differ quite a lot. Let’s look at this visually:



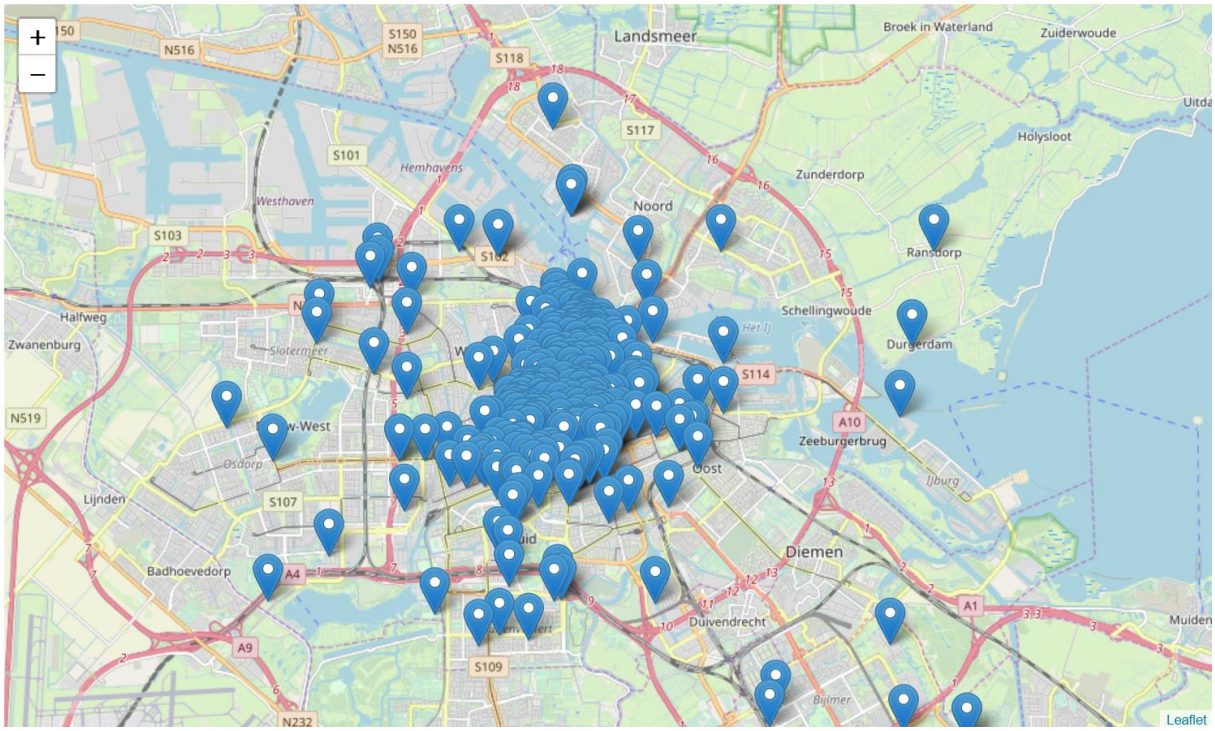
The visualization confirms the numbers and shows a couple of things:

- Hotels come in many sizes. There seems to be a large concentration of hotels with between 0 and 200 beds.
- Other all score below 200 beds. All rather small it seems.
- Hostel shows most items between 50 and 200 beds, and two outliers of around 500-600.

Geographical location

The hotels dataset unfortunately does not contain geographical data. The City of Amsterdam luckily also has a table online including this data. I will now import this database and add the geographical data to the hotels dataframe. I have used the postal code to link the two tables.

Based on some data cleansing I was able to plot the hotels on a map, a great way to visualize the location data:



The map shows some interesting results. The city centre of Amsterdam is usually seen as the part within the (right) ring road. In there, the hotels are mostly spread out in the middle of the city. In neighborhoods West, Noord and Oost there are few hotels.

Finding the neighborhoods with least hotels

Next I have analyzed the hotels per neighborhood to find the neighborhoods with the least current hotels. I have selected the bottom 10 for further analysis:

	Naam Wijk	HOTELNAAM_2014_x
0	Amstel III/Bullewijk	1
1	Volewijck	1
2	Tuindorp Oostzaan	1
3	Sloterdijk	1
4	Prinses Irenebuurt e.o.	1
5	Overtoomse Sluis	1
6	Osdorp-Oost	1
7	Osdorp-Midden	1
8	Omval/Overamstel	1
9	Kinkerbuurt	1

The 10 neighborhoods above have the least amount of hotels: great for my new hotel as there will not be a lot of competition around allowing my hotel to provide an unique experience for its customers!

Analyzing the bottom 10 neighborhoods

The 10 neighborhoods with the least hotels have been selected for further analysis. These neighborhoods I want to analyze on attractiveness. This is measured by the venues in the neighborhood. Foursquare is used to get this data. A maximum of hundred venues per neighborhood have been retrieved:

Neighborhood	
Amstel III/Bullewijk	12
Kinkerbuurt	100
Omval/Overamstel	25
Osdorp-Midden	8
Osdorp-Oost	29
Overtoomse Sluis	65
Prinses Irenebuurt e.o.	24
Sloterdijk	11
Tuindorp Oostzaan	7
Volewijck	28

The results show that there is only one neighborhood with at least 100 venues within 500 meters distance: Kinkerbuurt. Let’s dive into more details, what kind of venues are this?

	Neighborhood	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	9th Most Common Venue	10th Most Common Venue
0	Amstel III/Bullewijk	Restaurant	Food Court	Trail	Sandwich Place	Fried Chicken Joint	Furniture / Home Store	Hotel	Hotel Bar	Asian Restaurant	Gym
1	Kinkerbuurt	Bar	Coffee Shop	Italian Restaurant	Café	Restaurant	Burger Joint	Yoga Studio	Vegetarian / Vegan Restaurant	Ice Cream Shop	Mexican Restaurant
2	Omval/Overamstel	Hotel	Restaurant	Dance Studio	Garden	Gym / Fitness Center	Indoor Play Area	Nightclub	Metro Station	Men's Store	Sandwich Place
3	Osdorp-Midden	Bakery	Supermarket	Plaza	Baseball Field	Chinese Restaurant	Breakfast Spot	Hotel	Discount Store	Diner	Fast Food Restaurant
4	Osdorp-Oost	Clothing Store	Turkish Restaurant	Bus Stop	Supermarket	Drugstore	Shopping Mall	Gym	Indonesian Restaurant	Electronics Store	Fast Food Restaurant
5	Overtoomse Sluis	Café	Bar	Restaurant	Italian Restaurant	French Restaurant	Hotel	Vegetarian / Vegan Restaurant	Supermarket	Tram Station	Diner
6	Prinses Irenebuurt e.o.	Salad Place	Cocktail Bar	Hotel	Italian Restaurant	Lawyer	Food Court	Buffet	Breakfast Spot	Plaza	Snack Place
7	Sloterdijk	Hostel	Supermarket	Hotel	Italian Restaurant	Restaurant	Cafeteria	Soccer Field	Chinese Restaurant	Garden	Gastropub
8	Tuindorp Oostzaan	Business Service	Motorcycle Shop	Snack Place	Arts & Entertainment	Plaza	Theater	Bus Stop	Yoga Studio	Electronics Store	Drugstore
9	Volewijck	Bakery	Café	Bar	Pizza Place	Clothing Store	Dance Studio	Modern European Restaurant	Market	Doner Restaurant	Coffee Shop

Results and conclusions

The table above shows the analysis of venues of the bottom 10 neighborhoods (in terms of other hotels). The table show the most popular venues per neighborhood. I am interested in restaurants and bars, as I made the assumption in this research that this is what the type of tourist I am looking for longs for. The table shows neighborhood Amstel III/Bullewijk, Kinkerbuurt and Overtoomse Sluis as strong candidates for my hotel as this neighbourhood contains at least 6 categories of food/drinks venues as most common venues. Kinkerbuurt is the clear favorite with 9 and qualifies as the neighbourhood with the most common venues related to food/drinks.

The previous table showed the amount of venues found within a 500 meter radius of the neighborhood. Kinkerbuurt is the only neighborhood with at least 100 (I put a limit on 100) venues.

Therefore my hotel will be built in the Kinkerbuurt neighborhood. It has the most venues and the most restaurants/bars venues. Additionally there are almost no competitors in this neighborhood!

Discussion

In this report I analyzed the best spot for a new hotel in Amsterdam, keeping certain criteria in mind (away from competition yet close to restaurants and bars). There are several limitations that could be improved upon in a further study:

- More up-to-date information on currents hotels (this dataset is from 2014)
- Including additional data on neighborhoods, for example distance to museums or other interesting places.
- Including more neighborhoods in the analysis.