Selecting the best location to live in Berlin, Germany

Ian Carrier

3 July 2021

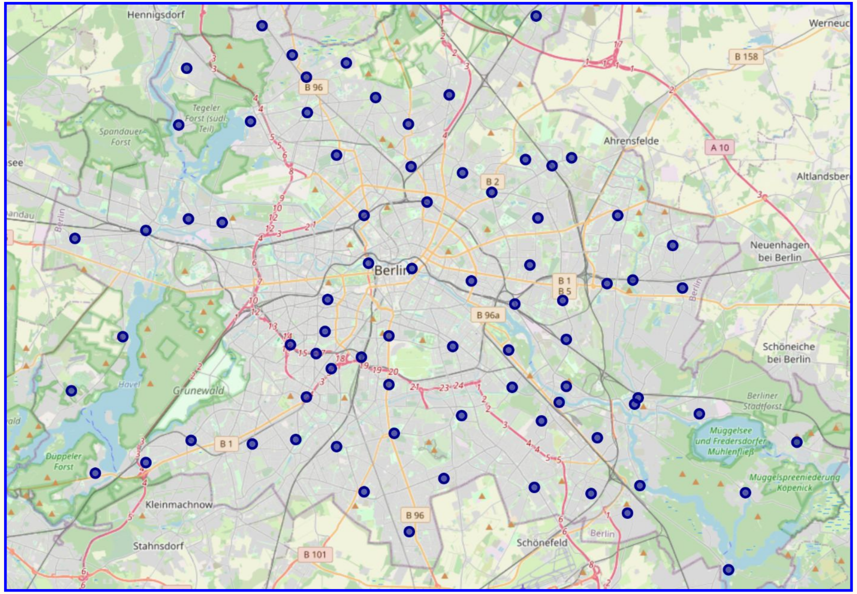
Overview

- my wife and I are moving to Berlin soon, we are unfamiliar with the city and have to select a location for long-term living
- we have the following list of requirements for the potential area:
  - vicinity of eateries, gym, cinema, supermarket
  - rent of €1,500 per month maximum
- the purpose of my project is to use various datasets for each Berlin borough / neighbourhood to decide upon an optimum living location
- the analysis will be used for my personal purposes but it is also of value to other future Berlin expats

Data

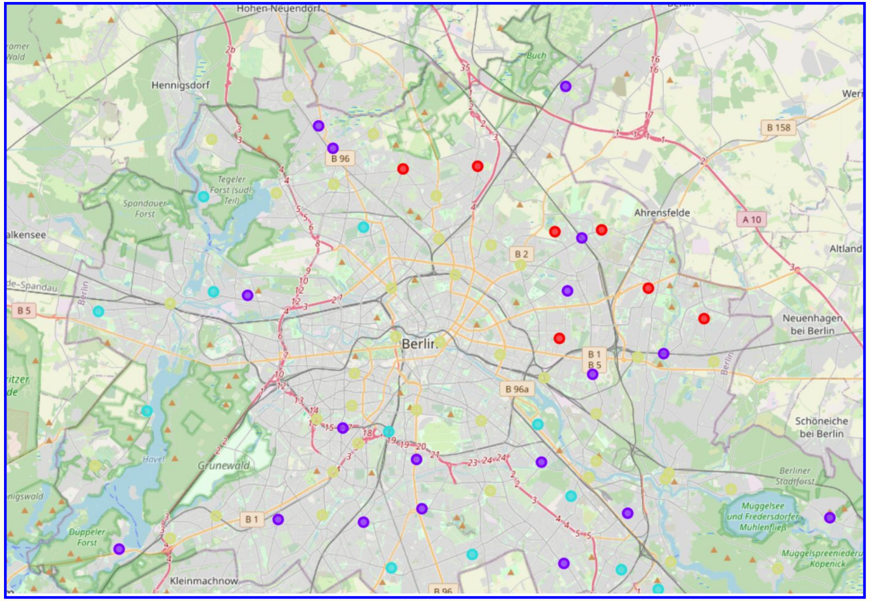
- I used the following data sources for the analysis:
  - Foursquare's geolocation of Berlin's amenities
  - Immo Scout24 rental prices for cities across Germany
  - geospatial coordinates for the boroughs and neighbourhoods of Berlin - this was generated within Python using the geopy library
  - my wife's and my preferences for local amenities and monthly rent of no more than €1,500

Methodology



- ❑ I generated a map using longitude and latitude coordinates for boroughs / neighbourhoods of Berlin
    - ❑ the challenge in this case was that Google longitude and latitude information is not free of charge any more
    - ❑ I used a **Python geopy library** to overcome this
- ❑ I then applied functions from a previous lab to extract venues from Foursquare
    - ❑ The venue data type was categorical so I used **one hot encoding** to convert the data into quantities enabling me to see most popular venues within each neighbourhood
- ❑ I decided to narrow down the analysis to top six most popular venues within each neighbourhood

Methodology



- ❏ I grouped all the neighbourhoods by common amenities via a **k-means clustering algorithm**
- ❏ I opted for 4 clusters as it provided the most balanced number of neighbourhoods within each group

Results

- ❑ Cluster 1 - purple dots
- ❑ Cluster 2 - blue dots
- ❑ Cluster 3 - yellow dots
- ❑ Cluster 4 - red dots

Cluster 1

```
[33]: cluster1 = berlin_merged.loc[berlin_merged['Cluster Labels'] == 0, berlin_merged.columns[[4] + list(range(5, berlin_merged.shape[1]))]]
cluster1
```

[33]:

	No.1 Most Common Venue	No.2 Most Common Venue	No.3 Most Common Venue	No.4 Most Common Venue	No.5 Most Common Venue	No.6 Most Common Venue
8	Supermarket	Drugstore	Gym / Fitness Center	Department Store	Seafood Restaurant	Electronics Store
9	Supermarket	Miscellaneous Shop	Light Rail Station	Trail	Plaza	Museum
13	Supermarket	German Restaurant	Drugstore	Fast Food Restaurant	Greek Restaurant	Chinese Restaurant
17	Supermarket	Furniture / Home Store	Fast Food Restaurant	Metro Station	Playground	Hardware Store
19	Supermarket	Drugstore	Bus Stop	Art Gallery	Bakery	Asian Restaurant
21	Supermarket	Hotel	Greek Restaurant	Café	Mexican Restaurant	Italian Restaurant
26	Drugstore	Supermarket	Italian Restaurant	Trattoria/Osteria	Shopping Mall	Steakhouse
28	Hotel	Drugstore	Vietnamese Restaurant	Supermarket	Memorial Site	Liquor Store
30	Drugstore	Supermarket	Metro Station	Zoo Exhibit	Market	Restaurant
38	Italian Restaurant	Supermarket	Drugstore	Light Rail Station	Gastropub	Fast Food Restaurant
44	Supermarket	Bakery	Drugstore	Sushi Restaurant	Movie Theater	Soccer Field
49	Supermarket	German Restaurant	Light Rail Station	Gas Station	Storage Facility	Museum
50	Supermarket	Park	Café	Italian Restaurant	Drugstore	Fried Chicken Joint
55	Supermarket	Wine Shop	Seafood Restaurant	Snack Place	Drugstore	German Restaurant
58						
61						
62						
65						

Cluster 2

```
[34]: cluster2 = berlin_merged.loc[berlin_merged['Cluster Labels'] == 1, berlin_merged.columns[[4] + list(range(5, berlin_merged.shape[1]))]]
cluster2
```

[34]:

	No.1 Most Common Venue	No.2 Most Common Venue	No.3 Most Common Venue	No.4 Most Common Venue	No.5 Most Common Venue	No.6 Most Common Venue
0	Eastern European Restaurant	Italian Restaurant	Supermarket	Restaurant	Bus Stop	Newsstand
22	Pool	Soccer Field	Bus Stop	German Restaurant	Metro Station	Supermarket
29	Bus Stop	Furniture / Home Store	Drugstore	Café	Gym / Fitness Center	Bakery
34	Supermarket	Sushi Restaurant	Pizza Place	Movie Theater	Taverna	Burger Joint
36	Italian Restaurant	Supermarket	Restaurant	Bus Stop	ATM	Newsstand
39	Soccer Field	Supermarket	Bus Stop	Doner Restaurant	Other Repair Shop	Museum
40	Supermarket	Pizza Place	German Restaurant	Bus Stop	Miscellaneous Shop	Music Venue
47	Bus Stop	Bakery	Supermarket	Park	Farmers Market	Fast Food Restaurant
60	Forest	Bus Stop	Home Service	Outdoor Sculpture	Museum	Music Store
67	Bank	Gym / Fitness Center	Park	Metro Station	Automotive Shop	Beer Store
69	Italian Restaurant	Insurance Office	Supermarket	Bus Stop	ATM	Newsstand
75	Bus Stop	Supermarket	Liquor Store	Light Rail Station	ATM	Nightclub
	Italian Restaurant	Hotel	Bus Stop	Bathing Area	Trattoria/Osteria	Harbor / Marina

Cluster 3

```
[79]: cluster3 = berlin_merged.loc[berlin_merged['Cluster Labels'] == 2, berlin_merged.columns[[4] + list(range(5, berlin_merged.shape[1]))]]
cluster3
```

[79]:

	No.1 Most Common Venue	No.2 Most Common Venue	No.3 Most Common Venue	No.4 Most Common Venue	No.5 Mo	
1	German Restaurant	Beach	Park	Indian Restaurant		
2	Hotel	History Museum	Art Museum	Scenic Lookout		
3	Italian Restaurant	Café	Nightclub	Bar		
4	Hotel	Sandwich Place	Bakery	Coffee Shop		
5	Tram Station	Supermarket	Smoke Shop	Snack Place		
7	Clothing Store	Drugstore	Tram Station	Bank		
10	Café	Falafel Restaurant	Bar	Ice Cream Shop	Vie	
12	Hotel	Zoo Exhibit	Clothing Store	French Restaurant		
14	Bar	Café	Coffee Shop	Cocktail Bar		
15	Pool	Trattoria/Osteria	Pet Store	Fast Food Restaurant		
16	Bar	Bakery	Café	Hotel	Vie	
18	Palace	Historic Site	Eastern European Restaurant	Bus Stop		
20	Vegetarian / Vegan Restaurant	Café	Coffee Shop	Middle Eastern Restaurant		
23	Harbor / Marina	Kids Store	Supermarket	Gas Station	German Restaurant	Tram Station
24	Italian Restaurant	Concert Hall	German Restaurant	Eastern European Restaurant	Church	Park
25	Mobile Phone Shop	Clothing Store	Drugstore	Restaurant	Bakery	Bus Stop
27	Café	Supermarket	Bistro	Organic Grocery	Bar	Trattoria/Osteria
31	Supermarket	Tram Station	Bakery	Playground	Organic Grocery	Café
32	Supermarket	Burger Joint	Asian Restaurant	Café	Falafel Restaurant	Park
33	Café	Drugstore	Supermarket	Gym / Fitness Center	Eastern European Restaurant	Pharmacy

Cluster 4 (the red dots)

```
[86]: cluster4 = berlin_merged.loc[berlin_merged['Cluster Labels'] == 3, berlin_merged.columns[[4] + list(range(5, berlin_merged.shape[1]))]]
cluster4
```

[86]:

	No.1 Most Common Venue	No.2 Most Common Venue	No.3 Most Common Venue	No.4 Most Common Venue	No.5 Most Common Venue	No.6 Most Common Venue
6	Tram Station	Asian Restaurant	Plaza	German Restaurant	Windmill	Supermarket
11	Tram Station	Supermarket	Asian Restaurant	Multiplex	Museum	Music Store
41	Supermarket	Tram Station	Escape Room	Gym / Fitness Center	Italian Restaurant	Bowling Alley
45	Tram Station	Supermarket	Automotive Shop	Hotel	Bus Stop	German Restaurant
70	Tram Station	Ice Cream Shop	ATM	Museum	Music Store	Music Venue
73	Supermarket	Tram Station	Nature Preserve	Bus Stop	Auto Garage	Nightclub
74	Tram Station	Supermarket	Chinese Restaurant	Shopping Mall	Lake	Gas Station

- ❑ Cluster 1 (purple dots) - is a location where supermarkets are the most common venues and it also has the most frequent occurrence of gyms and eateries

most appropriate match

Discussion

[85]:

	Neighbourhood	avgRent
11	Dahlem, Zehlendorf, Berlin	2116.109091
16	Wannsee, Zehlendorf, Berlin	1927.432973
5	Wilmersdorf, Wilmersdorf, Berlin	1908.607033
14	Rahnsdorf, Köpenick, Berlin	1525.583636
13	Hermsdorf, Reinickendorf, Berlin	1254.286452
12	Tempelhof, Tempelhof, Berlin	1129.882989
6	Adlershof, Treptow, Berlin	1063.875556
10	Lankwitz, Steglitz, Berlin	1056.022456
17	Rudow, Neukölln, Berlin	1036.339032
4	Buch, Pankow, Berlin	1007.536087
2	Mariendorf, Tempelhof, Berlin	995.839333
15	Baumschulenweg, Treptow, Berlin	928.417692
8	Friedrichsfelde, Lichtenberg, Berlin	867.931125
9	Waidmannslust, Reinickendorf, Berlin	867.675517
3	Siemensstadt, Spandau, Berlin	864.853800
7	Alt, Hohenschönhausen, Hohenschönhausen, Berlin	818.420982
0	Neu, Hohenschönhausen, Hohenschönhausen, Berlin	776.001750
1	Hellersdorf, Hellersdorf, Berlin	717.573784

- after determining Cluster 1 (purple dots) to be the most appropriate group of neighbourhoods for us, the next step was to narrow the search down to locations offering an average rent of no more than €1,500 a month
- I generated a table showing locations within Cluster one in a descending order by rental value
- since our rental cap is €1,500 per month, the areas we would be interested in start at line 13 and go down to the very bottom (line 1).

Conclusions

- ❑ the k-means clustering allowed me to narrow down the search into unfamiliar neighbourhoods and to select a location best matching our requirements
- ❑ given the budget constraint, I was then able to eliminate neighbourhoods outside of our budget
- ❑ eventually, based on the algorithm and the average rental information, we have decided to chose Hermsdorf, Reinickendorf as our long-term living location in Berlin
- ❑ should our preferences change, then once again the k-means clustering could help us by suggesting an alternative cluster