

Applied Data Science Capstone Final Project

Idea:

Moving to a new city is always a big step. Choosing the right neighborhood in the new city is difficult because while descriptions can of course help, it's difficult to estimate the real character of a neighborhood. I will try to help people moving from Berlin to Munich or vice versa making that decision by clustering neighborhoods of these cities in one pool to create a groups with a similar character. Knowing the city you are moving from you should then be able to better estimate the character of the different neighborhoods in the new one.

Data:

I will be using lists of neighborhoods in the respective cities from the web. The geo coordinates of these neighborhoods will then be extracted from the geopy module. A collection of venues at these locations will then be requested from the foursquare database to get the character of the neighborhood.

Methodology:

Venues in a 1000 m radius around the centers of the different boroughs were collected. Only venue categories containing 100 or more entries in total for both cities were considered. The boroughs were described by their frequency distribution of these venue categories. A K-means clustering algorithm was then used to cluster into categories. Attempts with different nr for k showed an ideal number of $k=4$. With less, to my knowledge very different boroughs were sorted together. With more, a clear distinction between the groups was lost.

Results:

Found Categories:

0. (dark red) Mellow living boroughs. Characterized by opportunities for grocery shopping but still with the possibilities to eat out or go to a cafe.
1. (dark blue) Similar to 0 but more diverse venue selection
2. (light blue) These regions have the most cafes, bars, and restaurants of all categories. The more expensive boroughs to live in
3. (cyan) Category exclusive to Berlin in this comparison. This has very few venues aside from supermarkets. Appears to be mostly in former eastern Berlin
4. (light green) Very few venues in total. Mostly in the outskirts

Results:

Found Categories:

0

'Adlershof' 'Blankenburg' 'Britz' 'Buch' 'Buckow'
'Charlottenburg-Nord' 'Daglfing' 'Denning' 'Falkenhagener Feld'
'Fennpfuhl' 'Forstenried' 'Freimann' 'Friedrichsfelde' 'Gatow'
'Hadern' 'Hakenfelde' 'Haselhorst' 'Hasenberg' 'Hellersdorf'
'Holzapfelkreuth' 'Johannisthal' 'Karow' 'Kaulsdorf' 'Kladow'
'Konradshöhe' 'Lankwitz' 'Lichtenrade' 'Lichterfelde'
'Mariendorf' 'Marienfelde' 'Marzahn' 'Niederschöneweide'
'Niederschönhausen' 'Pankow' 'Plänterwald' 'Siemensstadt'
'Staaken' 'Tempelhof' 'Untermenzing' 'Waidmannslust'
'Wannsee' 'Wilhelmsruh' 'Wilhelmstadt' 'Wittenau'

Results:

Found Categories:

1

'Allach' 'Alt-Hohenschönhausen' 'Alt-Treptow' 'Am Hart' 'Am Moosfeld' 'Am Riesenfeld' 'Aubing' 'Baumschulenweg' 'Berg am Laim' 'Biesdorf' 'Bogenhausen' 'Borsigwalde' 'Englschalking' 'Fasangarten' 'Feldmoching' 'Fürstenried' 'Gropiusstadt' 'Grünau' 'Halensee' 'Harlaching' 'Johanneskirchen' 'Karlshorst' 'Laim' 'Ludwigsvorstadt' 'Lübars' 'Mahlsdorf' 'Milbertshofen' 'Mitte' 'Mittersendling' 'Moosach' 'Märkisches Viertel' 'Neuhausen' 'Nord Schwabing' 'Oberföhring' 'Obergiesing' 'Obermenzing' 'Obersendling' 'Pasing' 'Perlach' 'Ramersdorf' 'Reinickendorf' 'Riem' 'Rudow' 'Rummelsburg' 'Schmargendorf' 'Schmöckwitz' 'Schwabing' 'Sendling' 'Solln' 'Spandau' 'Steinhausen' 'Tegel' 'Thalkirchen' 'Tiergarten' 'Trudering' 'Untergiesing' 'Untersendling' 'Weißensee' 'West Schwabing' 'Wilmersdorf' 'Zamdorf'

Results:

Found Categories:

2

'Altschwabing' 'Altstadt' 'Au' 'Blankenfelde' 'Charlottenburg'
'Dahlem' 'Friedenau' 'Friedrichshagen' 'Friedrichshain'
'Frohnau' 'Gesundbrunnen' 'Haidhausen' 'Hansaviertel'
'Isarvorstadt' 'Kreuzberg' 'Köpenick' 'Lehel' 'Lichtenberg'
'Maxvorstadt' 'Moabit' 'Müggelheim' 'Neukölln' 'Nikolassee'
'Nymphenburg' 'Oberschöneweide' 'Prenzlauer Berg'
'Schwanthalerhöhe' 'Schöneberg' 'Steglitz' 'Wedding' 'Westend'
'Zehlendorf'

Results:

Found Categories:

3

'Altglienicke' 'Falkenberg' 'Französisch Buchholz' 'Heiligensee'
'Heinersdorf' 'Hermsdorf' 'Neu-Hohenschönhausen' 'Rosenthal'
'Stadtrandsiedlung Malchow' 'Wartenberg'

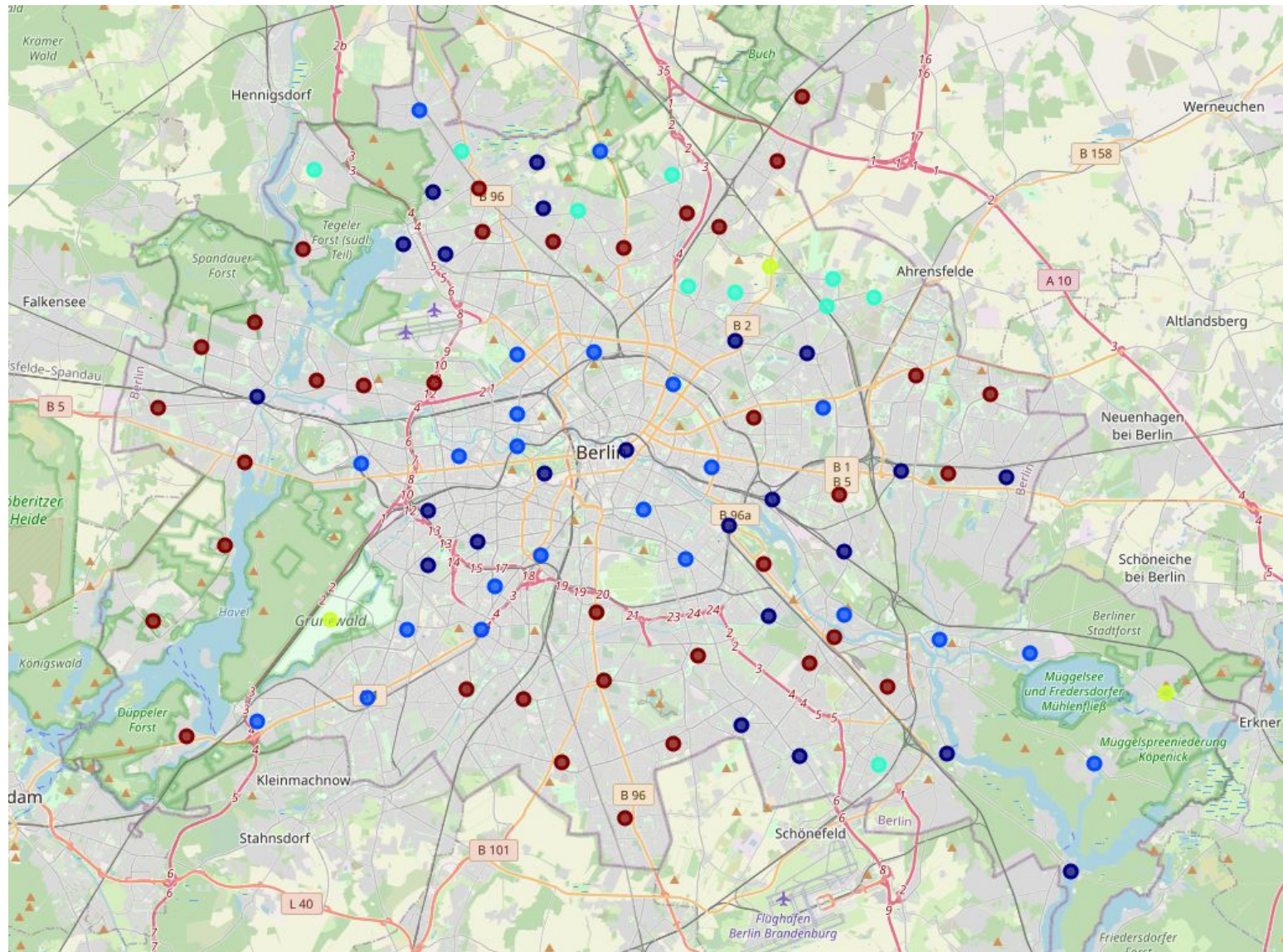
Results:

Found Categories:

4

'Freiham' 'Grunewald' 'Lochhausen' 'Malchow' 'Rahnsdorf'

Berlin



Munich

