



Rekha Remadevi
09.03.2021

Berlin or Munich:

Clustering the Neighbourhoods



Introduction

Background and objective

Background

- Munich and Berlin, two of the popular cities in Germany, have a lot to offer.
- Berlin is not only Germany's capital and largest city; it is also the cultural hub of the nation.
- Berlin, known as the techno capital of the world, is famous for its impressive nightlife.
- Munich, the wealthy capital of Bavaria is known as gateway to the Alps.
- Known to be one of the most beautiful and charming cities in all of Germany, it is filled with museums and beautiful architecture.

Business Problem

- A person looking to migrate or to visit Germany would like to evaluate the neighborhoods of Berlin or Munich in terms of what the cities have to offer, the standard of living in terms of housing and other amenities, the social aspects etc.
- Such information is intended to help the person in making a decision in terms of which city to choose, be it for a visit or relocation.
- This study will help stakeholders make informed decisions and address any concerns they have including the different kinds of cuisines, provision stores and other offerings.



Data Description

Acquisition of data

Data Acquisition

- In terms of data, it is of paramount importance to obtain geographical location data for both Berlin and Munich. Postal codes in each city serve as a starting point.
- Berlin : To derive our solution, We scrape our data from <http://www.places-in-germany.com/14356-places-within-a-radius-of-15km-around-berlin.html>
- Munich : The data published at <https://www.muenchen.de/int/en/living/postal-codes.html> is used in order to fetch the necessary data
- The *python geopy* library is used for getting the latitude and longitude values.
- *Four square API* is used to explore the venue data



Methodology

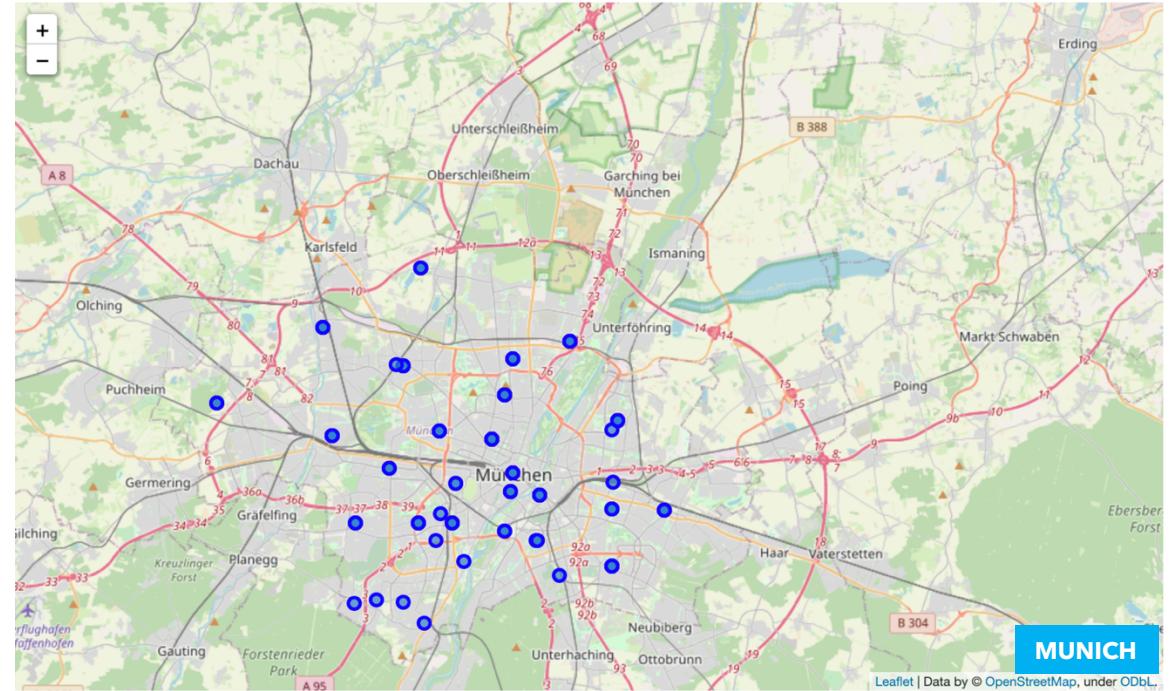
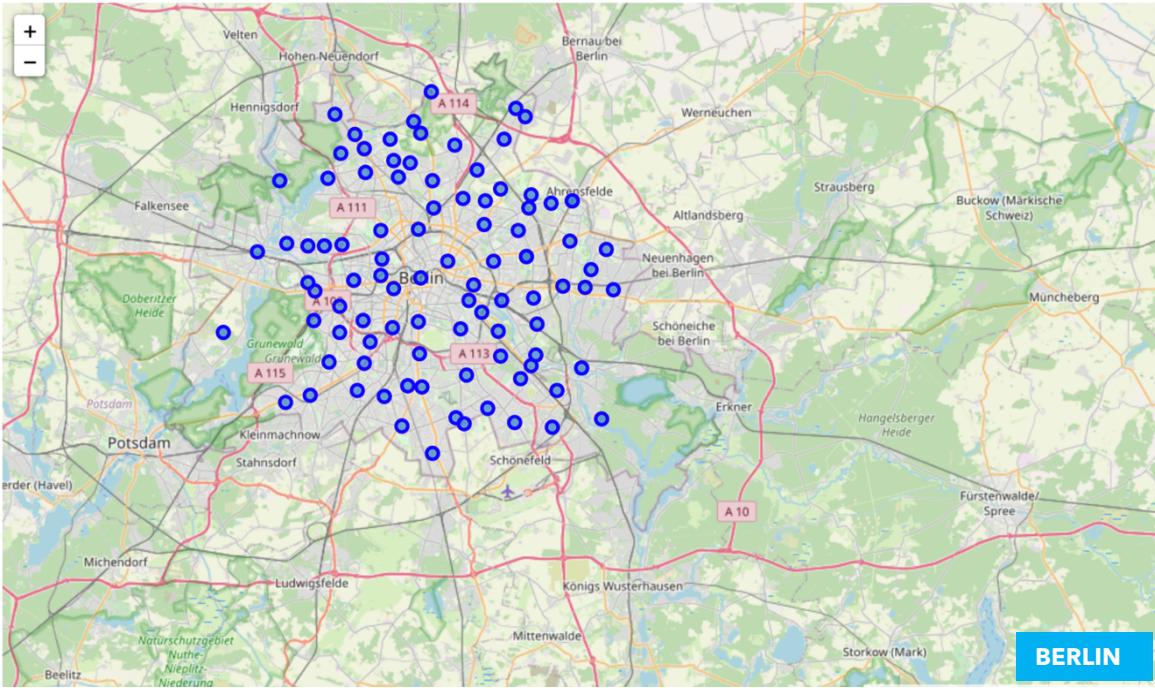
Approach and framework

Methodology

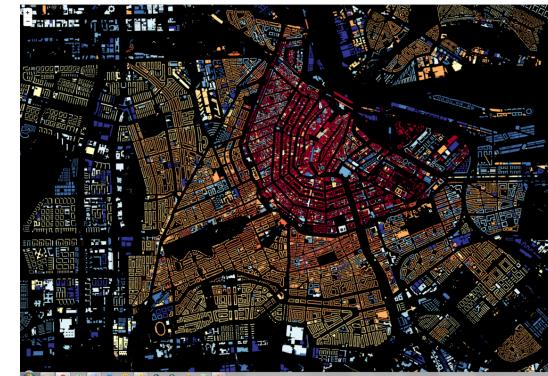
- By using *Python libraries*, we create our model. The required packages are the given to the right.

```
import pandas as pd
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)
import numpy as np
import json
from geopy.geocoders import Nominatim
from pandas.io.json import json_normalize
from bs4 import BeautifulSoup
import requests
import geopy as geo
import geopandas as gpd
import folium
from sklearn.cluster import KMeans

import matplotlib.cm as cm
import matplotlib.colors as colors|
```



Visualizing the Neighborhoods of Berlin and Munich



One-hot encoding

- Performed *one-hot encoding* to find out what are the different kinds of venue categories present in each neighbourhood and then calculate the top 10 common venues.
- Defined a function to get the topmost common venue categories.
- Used *K-means algorithm* to cluster similar neighbourhoods together.



Results

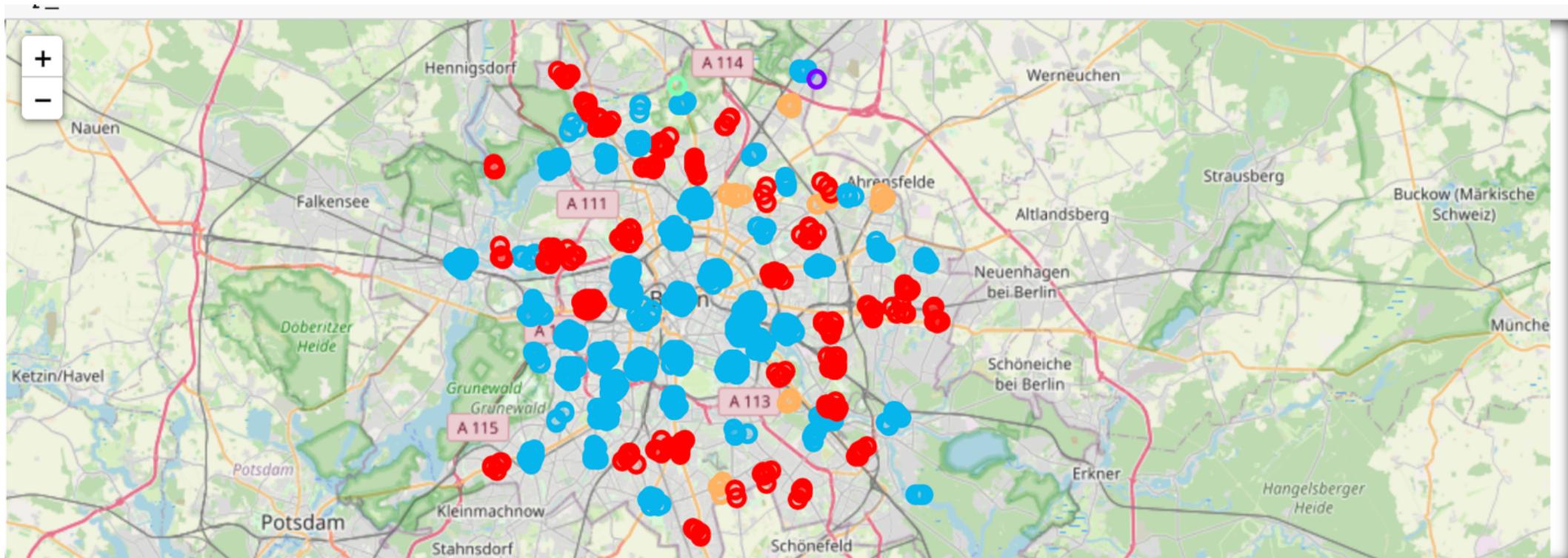
Key observations and findings

Overview of top Venues in Berlin

Café	453
Hotel	105
Bar	74
Bakery	73
Italian Restaurant	69
Coffee Shop	55
Turkish Restaurant	49
Park	43
Drugstore	42
German Restaurant	31
Sushi Restaurant	29
Restaurant	29
Furniture / Home Store	27
Clothing Store	16
Gym / Fitness Center	15
Lounge	10
Tram Station	7
Nature Preserve	4
Greek Restaurant	4
Zoo Exhibit	4
Sporting Goods Shop	4
Lake	3

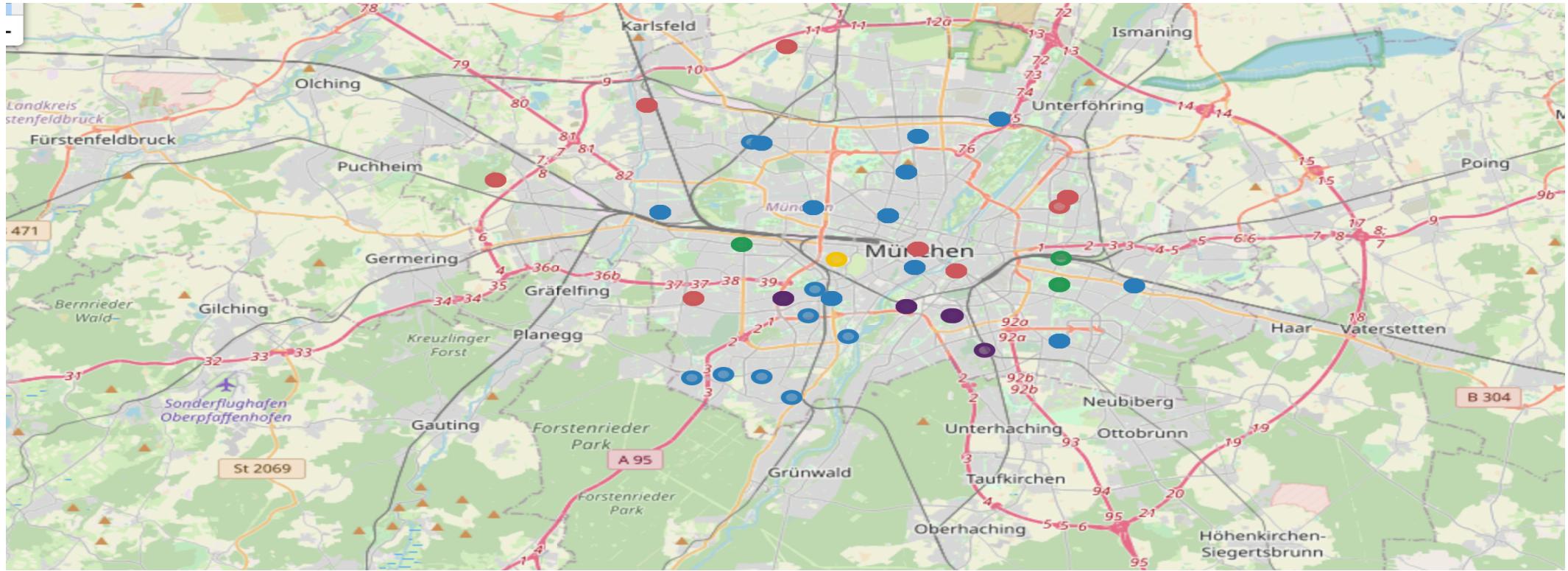
Visualizing the clustered Neighborhoods

- Map of the clustered Neighbourhoods in Berlin



Visualizing the clustered Neighborhoods

- Map of the clustered Neighborhoods in Munich





Discussions

Key facts

Berlin

The neighbourhoods of Berlin are very multicultural with a variety of supermarkets, restaurants, bars, coffee shops, ice cream shops, drug store and clothing store.

With a variety of shopping options in terms of fish markets, garden centre, gaming café, dessert shop, bookstores and sporting goods shop, the main modes of transport seem to be trams and buses.

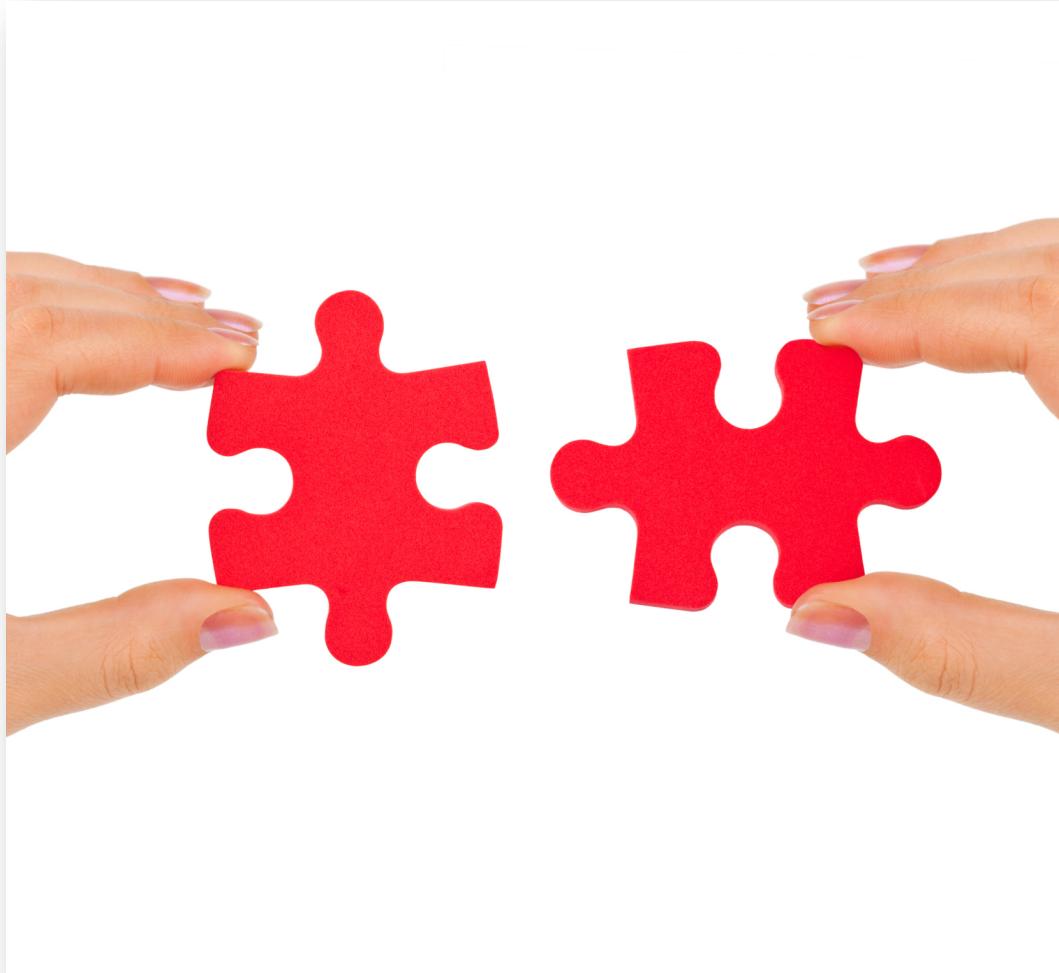
The cluster number zero indicates that the most common venues are Supermarket and Bakery. It also has the greatest number of tram stations, metro stations and bus stops. This cluster also has café, soccer fields, climbing gym and German restaurants. This caters well to families having kids.

Munich

The neighbourhoods of Munich offers a wide variety of cuisines and eateries including Italian, Currywurst joint, Asian, Chinese etc. with a lot of Sporting goods shop.

The blue cluster, the most common cluster in Munich, seems to have a lot of similar districts in the city. In cluster number one, drugstores and supermarkets are the most common venues and it's therefore the 'Drugstore + Supermarket' Cluster.

The cluster number two, with lots of hotels and plazas, also have a lot of coffee shops making this my favorite.



Conclusion

Conclusion

- The purpose of this project was to explore the cities of Berlin and Munich and see how attractive it is to potential migrants and tourists.
- The neighbourhoods of Berlin and Munich have more like and similar venues .
- While the neighbourhoods around both the cities have a wide variety of experiences, each of them unique, to offer; the dissimilarity exists in terms of different venues and facilities but with minor variations.
- While the cultural diversity is quite evident, it also gives one a feeling of a sense of inclusion. Overall, it's up to the stakeholder's preference as to what each wants to experience, and what suits their tastes.



Thank you