

Capstone Project - The Battle of Neighborhoods

Clustering the Neighborhoods of Amsterdam, The
Netherlands

Kinga A. Lőrincz

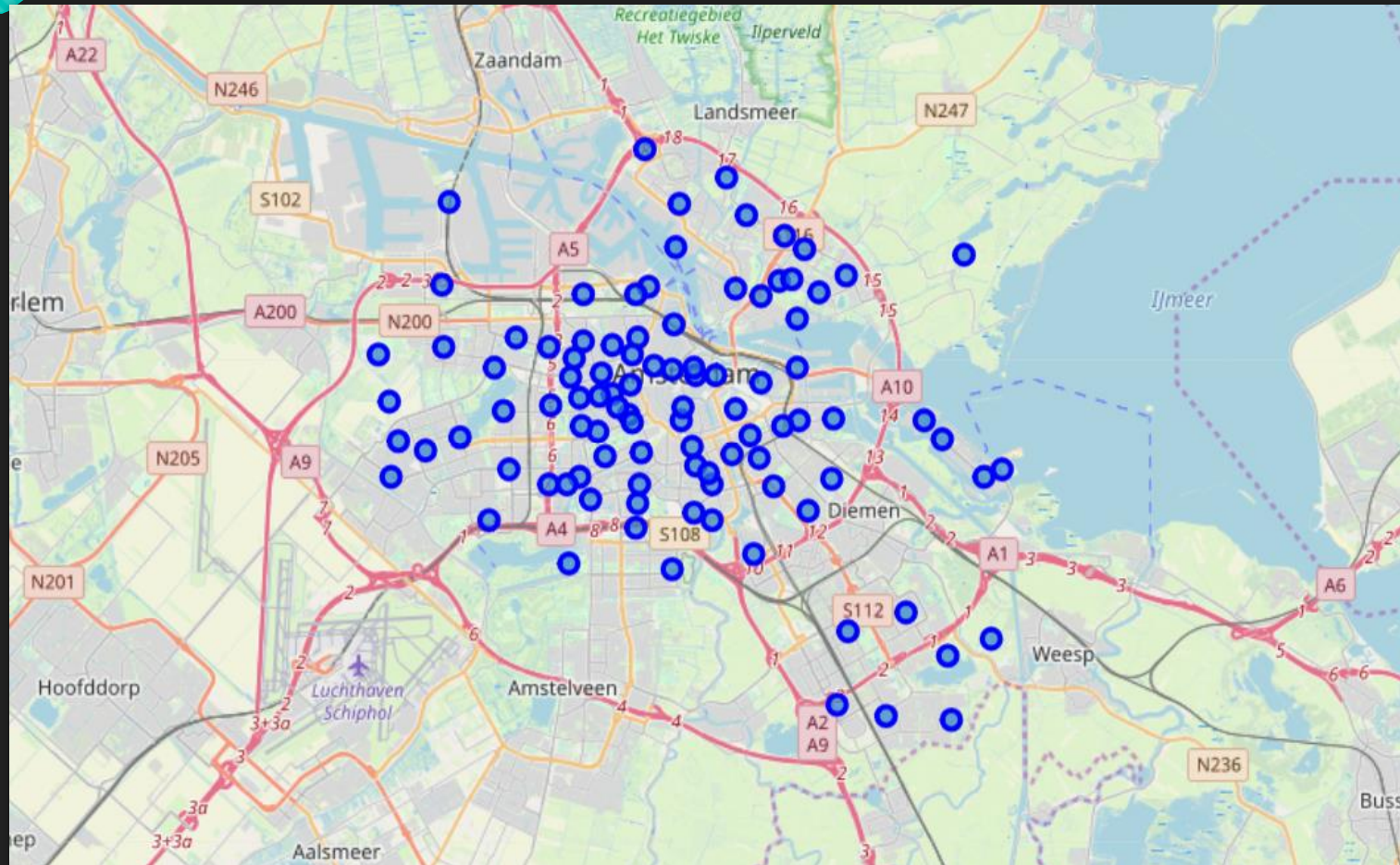
Business problem

- The Netherlands is one of the world's top expat countries
- The expat housing company wants to enhance the experience of their customers by recommending neighborhoods on individual preferences

Data

- Foursquare location data: using the API called *explore* obtain the most common venues in each neighborhood
- The geo-coordinates (latitude and longitude) of the neighborhoods of Amsterdam
- Clean the data by removing unnecessary features and transform it to be suitable for further analysis

The 99 neighborhoods of Amsterdam



Explore the neighborhoods

- Use the Foursquare API explore all neighborhoods → 6557 venues

```
amsterdam_venues.head()
```

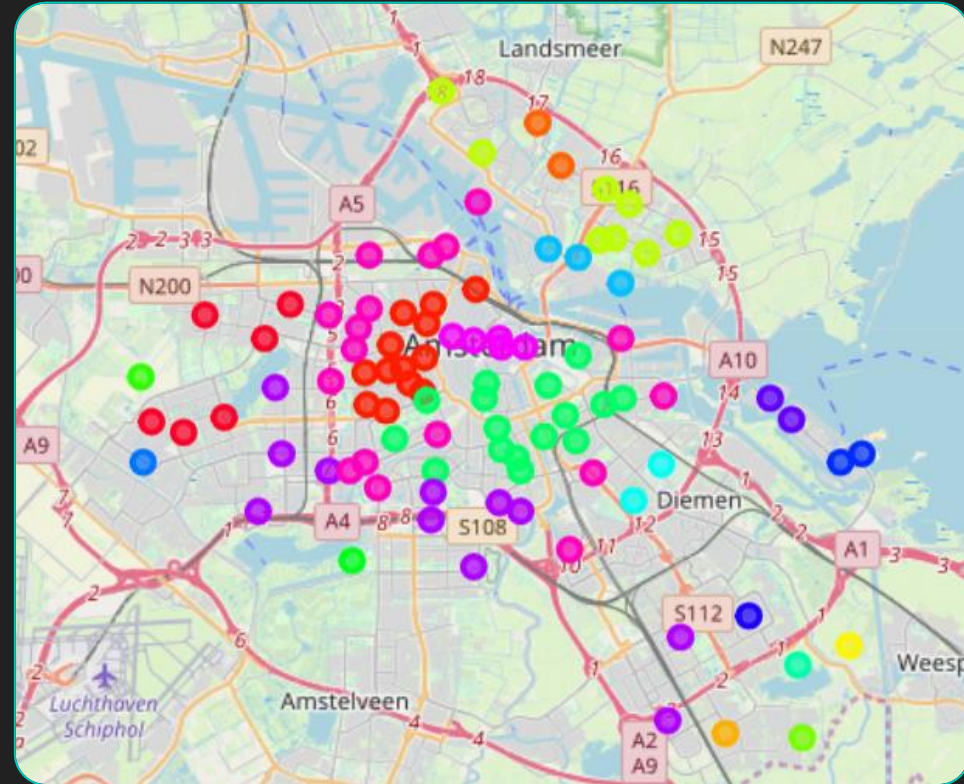
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Amstel III/Bullewijk	52.298741	4.950592	Hotel Lowell	52.304526	4.950671	Hotel
1	Amstel III/Bullewijk	52.298741	4.950592	De Proefzaak	52.303585	4.950890	Café
2	Amstel III/Bullewijk	52.298741	4.950592	Fitness365	52.298421	4.945692	Gym
3	Amstel III/Bullewijk	52.298741	4.950592	Museum Vrolik	52.293968	4.960014	Science Museum
4	Amstel III/Bullewijk	52.298741	4.950592	Joy	52.305268	4.950310	Hotel

Explore the neighborhoods

- Group the venues by neighborhood and remove all neighborhoods with less than 10 venues – 4 neighborhoods removed
- 315 unique venue categories in the remaining 95 neighborhoods
- Group the venues by neighborhoods by taking the mean frequency of occurrence of each venue category

Clustering of the neighborhoods

- Kmeans clustering algorithm, number of clusters=20



Clustering of the neighborhoods- examples

Cluster 1 - 6 neighborhoods

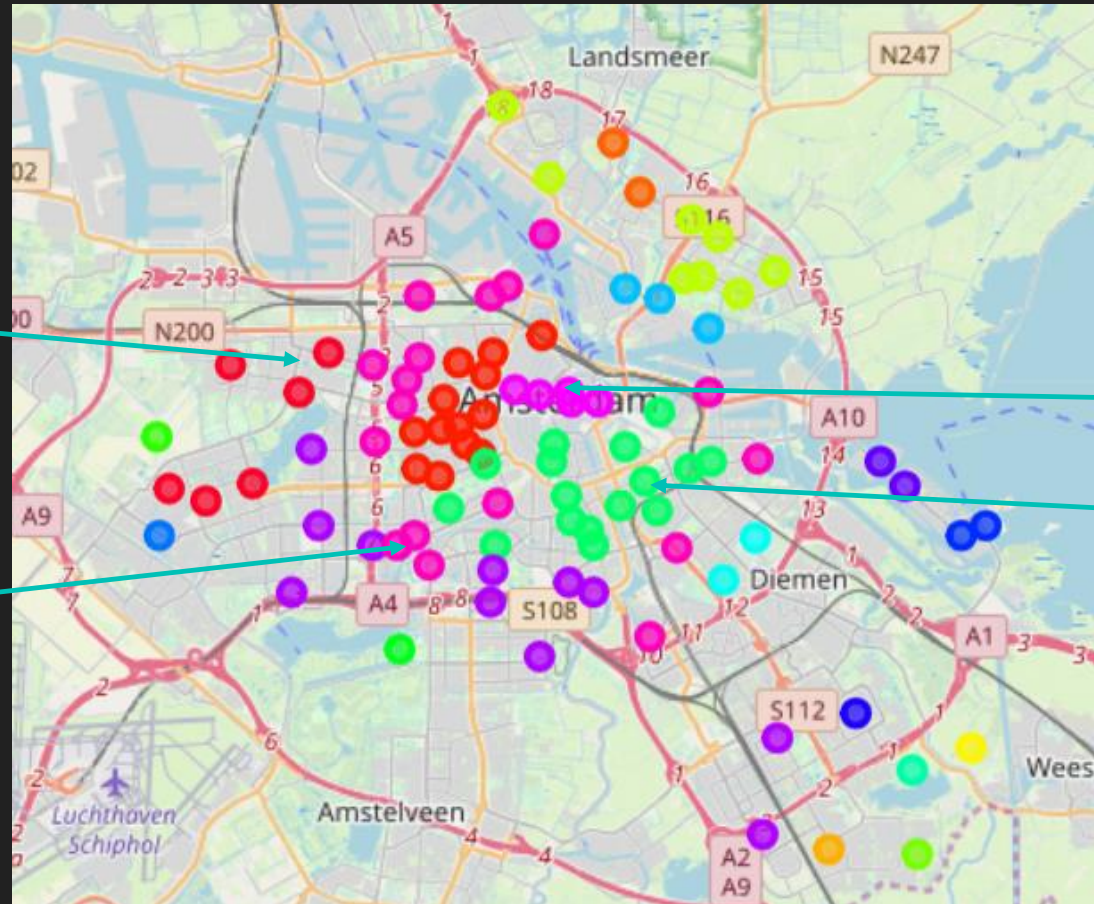
Venues: Turkish restaurants

Target: Turkish background

Cluster 0 - 17 neighborhoods

Venues: restaurants, cafes, entertainment (theater, museum, music venues).

Target: young professionals



Cluster 19 - 5 neighborhoods

Venues: hotels, bars and an marijuana dispensaries

Target: not suitable for families with young children

Cluster 10 - 16 neighborhoods

Venues: zoo, large variety of restaurants: Italian, French, Japanese etc.

Target: culinary adventurers, families with young children

Conclusion

- Using geocoordinates of Amsterdam and the Foursquare location data I have obtained a detailed clustering of the neighborhoods of Amsterdam.
- These clusters can be used to recommend neighborhoods to the customers of the expat housing company. These recommendations can be made using the preferences of the customers for different venues or based on the general preference of customers with the same age range, family status, education, work etc.
- This analysis can be improved using better location data and can be expended to other cities.