Price of living in Helsinki, Finland – Neighborhood characteristics or geographical location?

Emil Kronberg, 2021

Introduction

- Prices in Helsinki are steadily rising
- It is difficult to find a good but affordable neighborhood when moving to Helsinki
- Prices are generally considered to be affected by location, not neighborhood characteristics
- This research was aimed to give more insight on what affects the price and if there are neighborhoods cheaper than others that are similar in characteristic

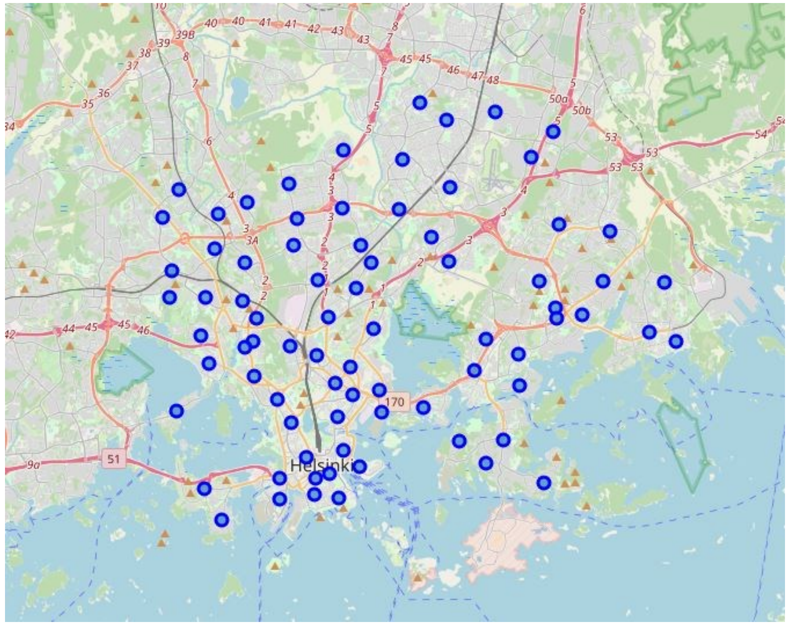
Description of data

- The data was gathered both from webscraping as well as using API's and even a little bit of simple google maps searching.
- Data included:
  - Name, Postcode, Coordinates of neighborhood
  - Average price (€) per square meter
  - Boundary coordinates for respective neighborhood
  - Venues located in each neighborhood

Methodology

- After extraction of the initial neighborhood characteristic data, the entries was plotted on a map for general overview

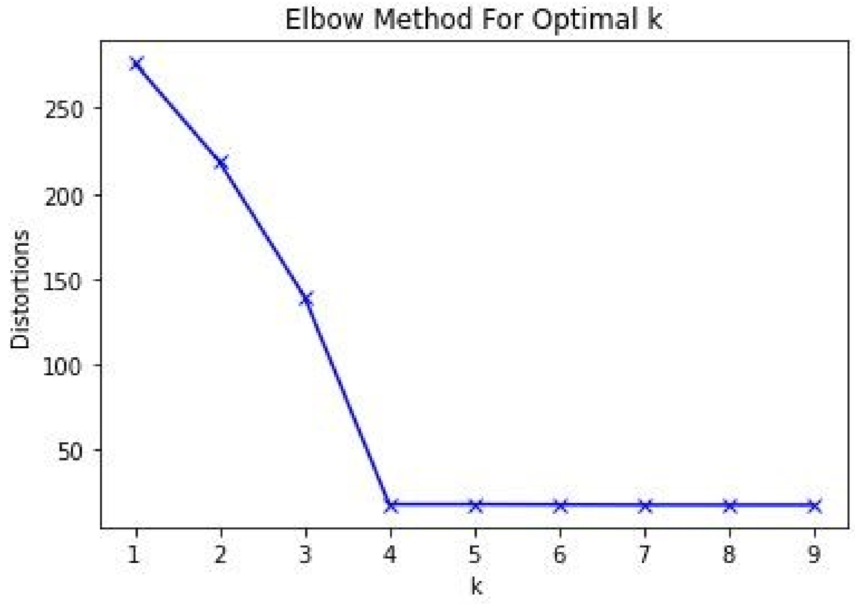
- After gathering venue data through the neighborhood coordinates, several iterations of data transformation was made to quantify the data for further analysis



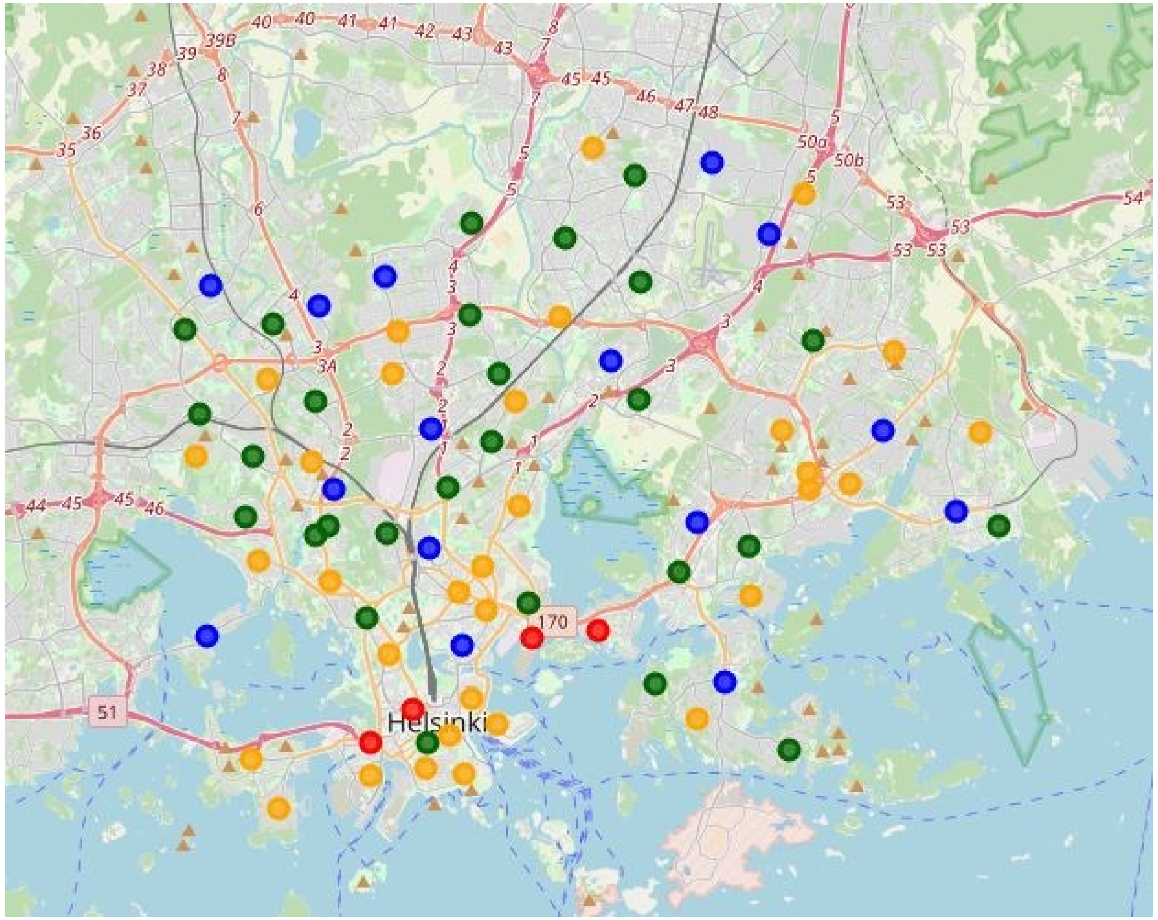
	Neighborhood	ATM	Accessories Store	American Restaurant	Antique Shop	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Auditorium	Auto Dealership	Auto Garage	Auto Workshop	Automotive Shop	BBQ Joint	Badminton Court	Bagel Shop	Bakery	Bar	Baseball Field
0	Aurinkolahti	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.0	0.0
1	Eira - Hernesaari	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.022727	0.068182	0.0	0.0
2	Etelä-Haaga	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.0	0.0
3	Etelä-Laajasalo	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.000000	0.0	0.0
4	Etelä-Vuosaari	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000000	0.040000	0.0	0.0

K-Means

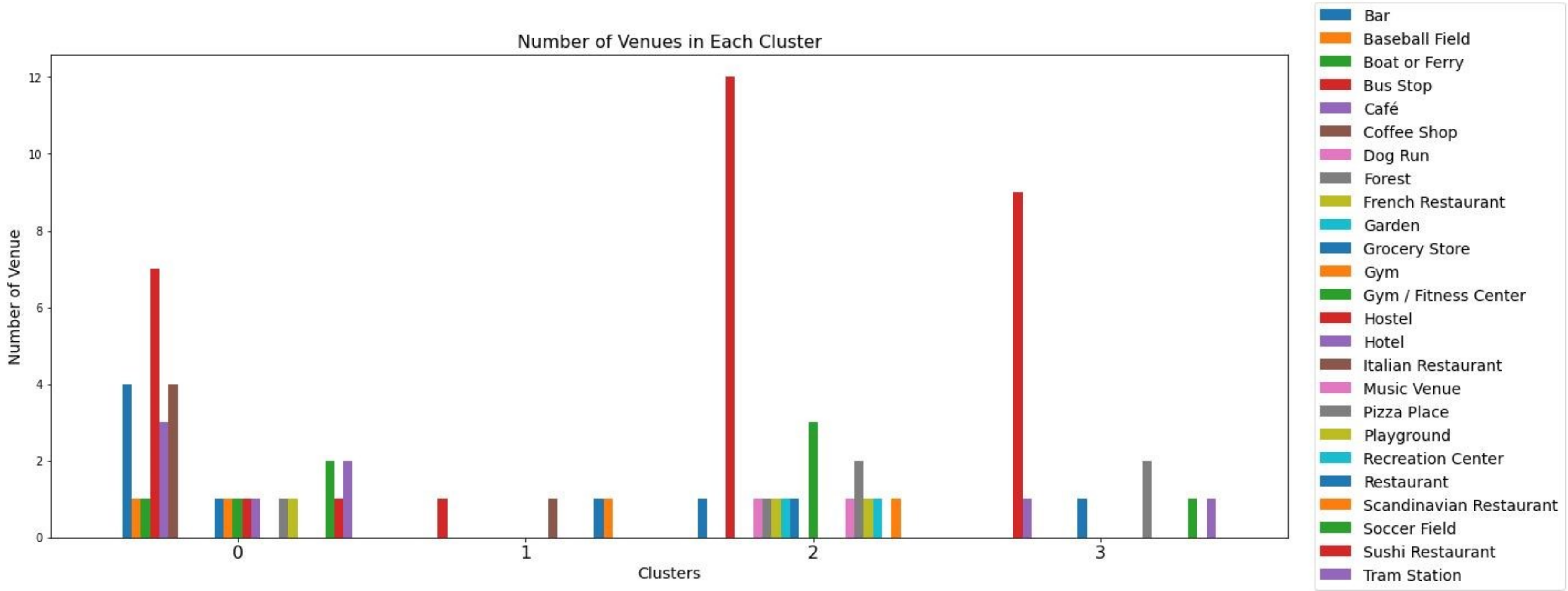
- The unsupervised learning method K-means was utilized to cluster neighborhoods based on similarities in popular venues.
- The elbow method was primarily used to find the optimal number of clusters (k)
- The optimal number was 4, which was also confirmed with the silhouette score method



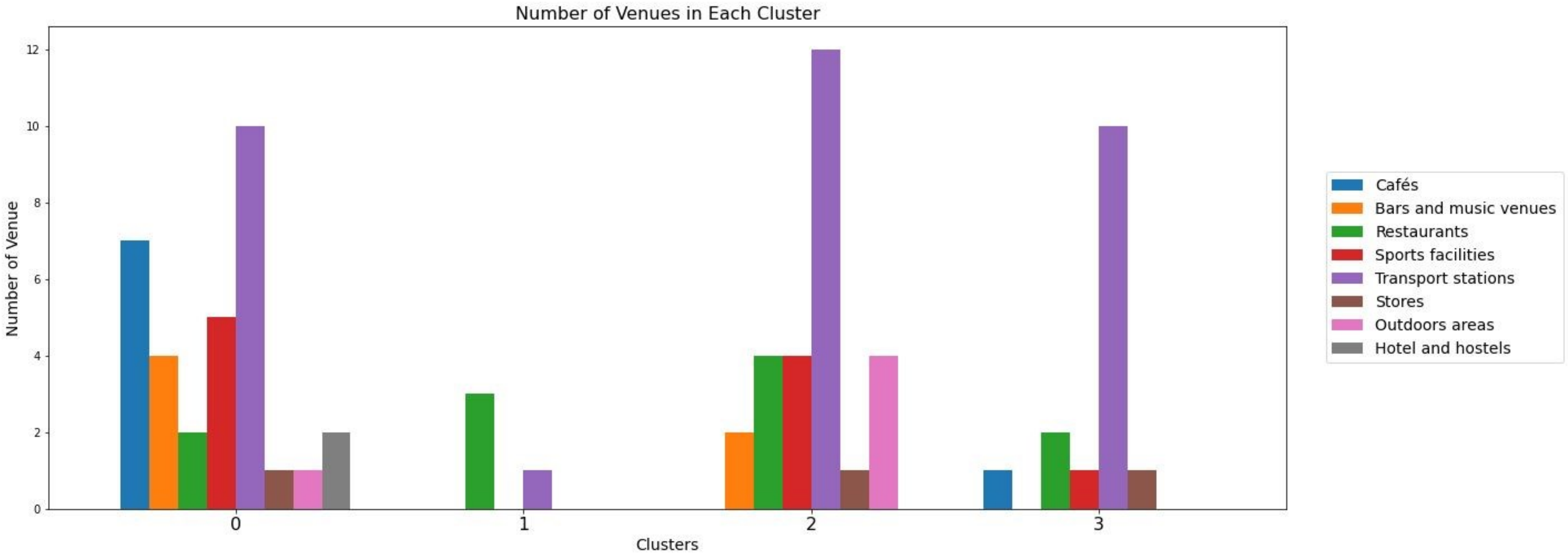
Mapping clusters



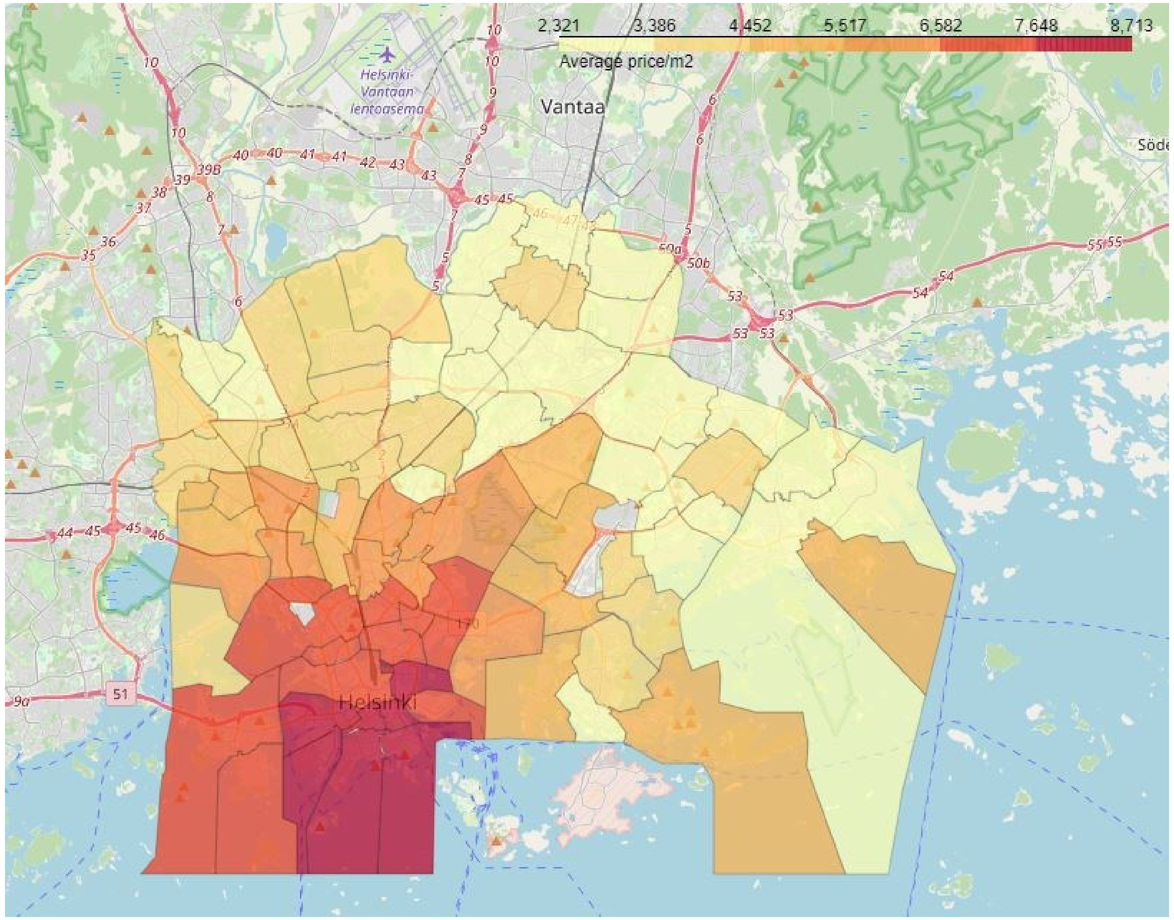
Differences in clusters



After merging similar venues



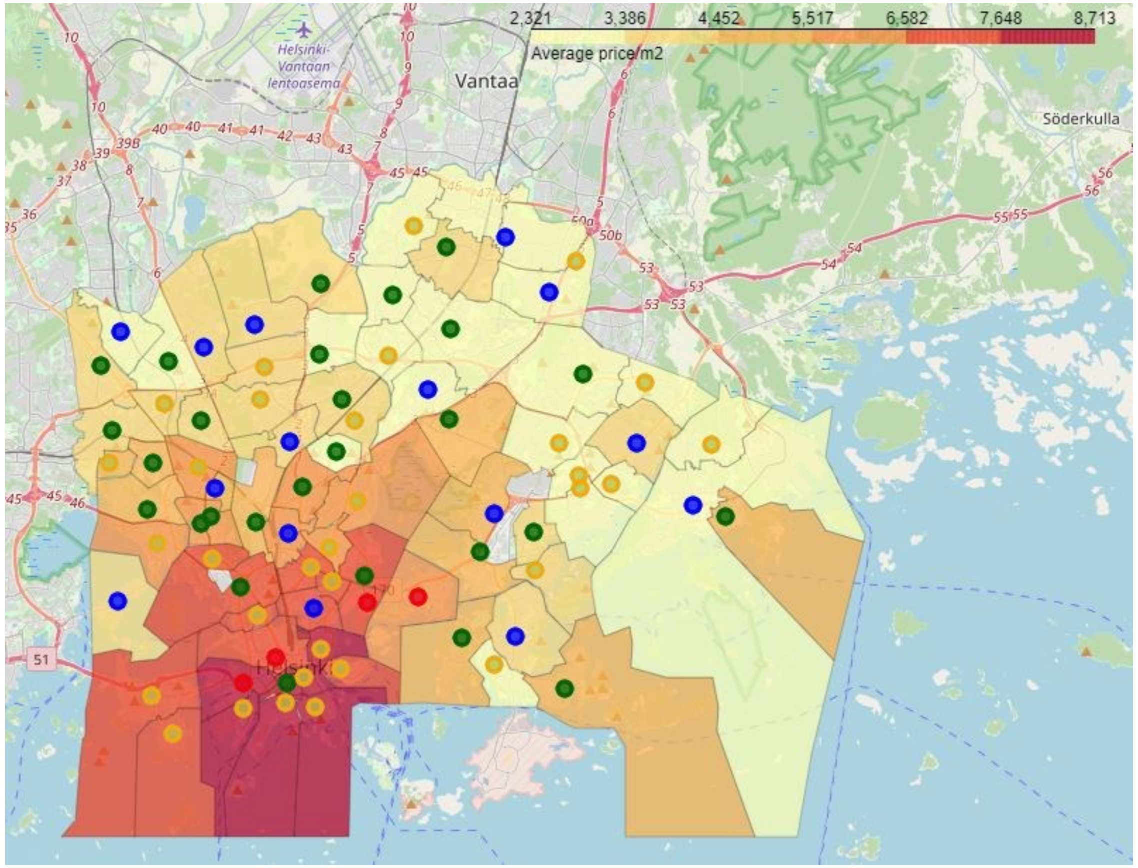
Geographical position and price



	Average price
Cluster Labels	
0	5196.43750
1	7138.50000
2	4640.62963
3	3998.80000

Results

- Neighborhood characteristics not as impactful on price as location
- Some cheaper neighborhoods may therefore be similar to those in expensive areas



Discussion

- Price is heavily tied to location, but some ‘trendy’ neighborhoods in expensive areas are still similar in characteristic to those far from the city center
- The distribution of venues in the data was quite large and could have been merged immediately for more accurate analysis, but would have been very time consuming
- Foursquare API may not have data on all locations
- Other variables should be combined in future research for a more robust result that better defines characteristics of neighborhoods

Conclusion

- Data may help find neighborhoods to live in based on price and what nearby venues they may offer
- It is ultimately a personal choice if characteristics is worth more than location, for example having a short distance to work
- Data is not deterministic but can assist decision making