

# THE BATTLE OF NEIGHBORHOODS

Finding similar neighborhoods in Zagreb and Split, Croatia

Mile Vučković / [milevucko@gmail.com](mailto:milevucko@gmail.com) / January, 2021

# Overview

Introduction

Business Problem

Data

- Neighborhoods in Zagreb and Split

- Geocoordinates of the neighborhoods in Zagreb and Split

- Top venues of neighborhoods in Zagreb and Split

Methodology

Discussion

Conclusion

# Introduction

## Main goal

- This project tries to solve an imaginary, but very common problem on the Croatian real estate market.
- I hope it will help actors at both ends of real estate transactions, especially private traders, but also those who provide real estate services, such as real estate agencies.
- It includes querying Foursquare location data for the purpose of comparing city neighborhoods.

# Business Problem

## About Zagreb and Split

- Zagreb is the capital of Croatia with almost a million inhabitants, and is located in the inland of Croatia. Major companies, government agencies, universities, and other vital institutions are located in Zagreb.
- Split is the second largest town in Croatia, and is situated south, on the Adriatic coast.
- A large number of people from Zagreb want to buy, or trade, a real estate in Split, especially when they reach pension so that they can enjoy in the warmer winter climate of the Adriatic coastline. On the other hand, a large number of people from Split that are not involved in tourism are drawn to Zagreb, either due to larger labor market, number and strength of employers, cheaper real estates, or other reasons.

# Business Problem

## Project goal

The main idea of this project is to create clusters of neighborhoods of these two cities, so that we can compare them and provide insight as a service for actors on the real estate market. Top venues of neighborhoods are going to be used as data for neighborhood comparison.

# Data

List of the neighborhoods in Zagreb and Split.

## **Download and clean neighborhood data**

Neighborhoods of Split are loaded from the excel sheet, picked up from the city's official website. Neighborhoods of Zagreb are collected from Wikipedia.

Geocoordinates of the neighborhoods in Zagreb and Split

## **Geopy library usage**

The latitude and longitude of the neighborhoods are retrieved using a geopy library.

Neighborhood address is used as the input parameter for geocoordinate retrieval.

Top venues of neighborhoods in Zagreb and Split

## **Foursquare REST API**

The venue data is collected using Foursquare. When collecting venues, a 800 meter radius is used around the center coordinates of the neighborhood. Number of venues is limited to 1000 per neighborhood

# Methodology

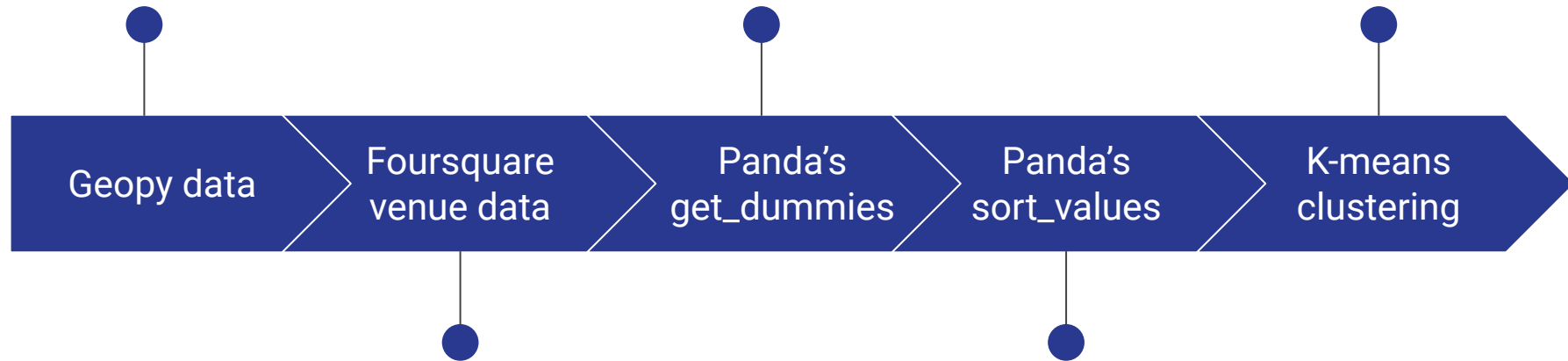
K Means clustering, a form of unsupervised machine learning algorithm, is used for creating clusters of neighborhoods.

---

Geocoordinates of 44 neighborhoods in both cities are inserted into a single pandas dataframe

One-hot encoding technique is applied using venue categories

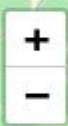
Optimal number of clusters is found using silhouette\_score metrics



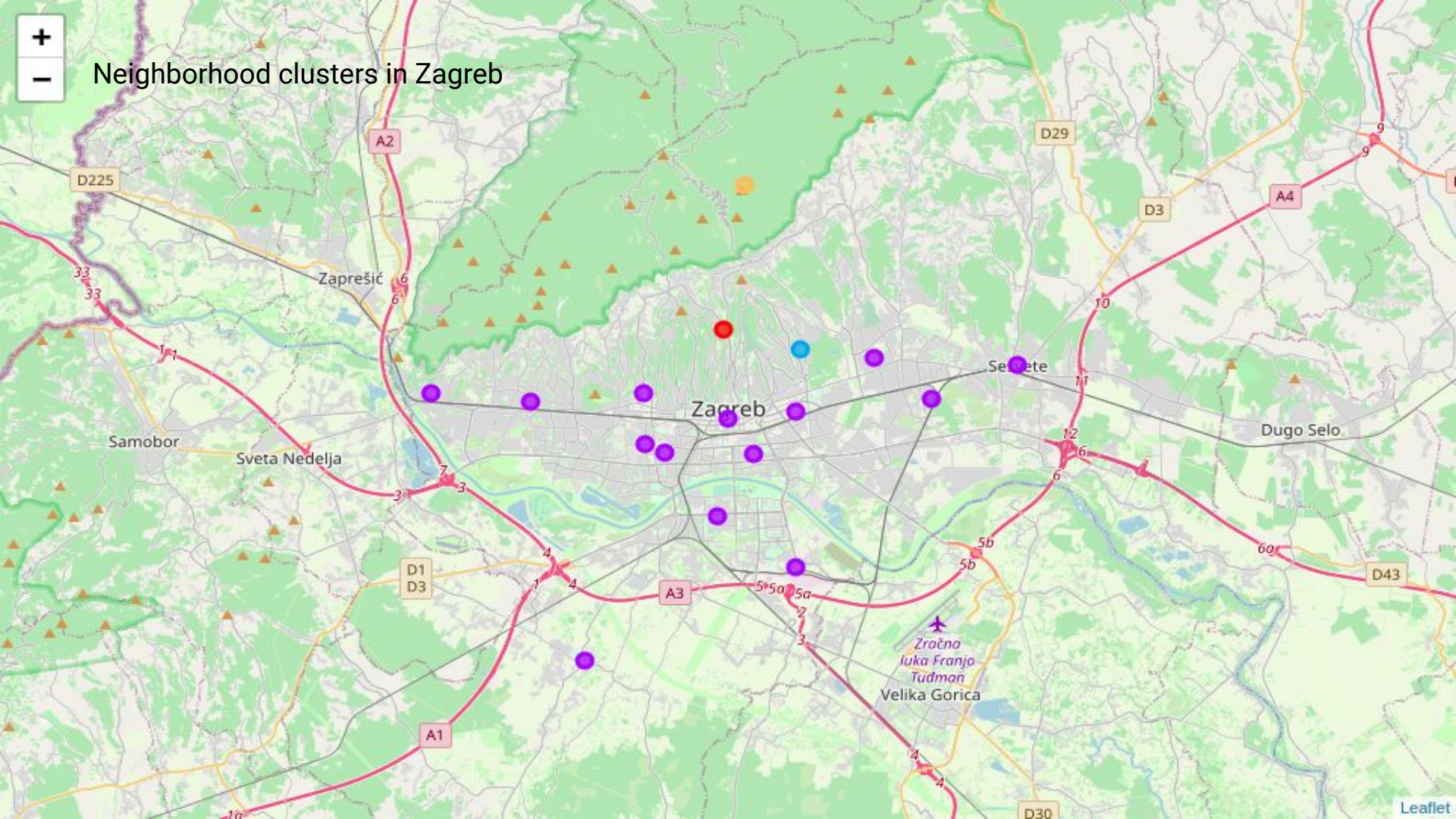
Venue data is collected using Foursquare REST API

Top 10 venues are assigned to each neighborhood by sorting number of occurrence of particular venue category

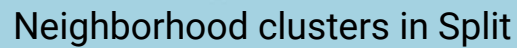




# Neighborhood clusters in Zagreb







# Discussion

## Unique clusters

- Clusters '1' and '5' represent neighborhoods that are unique to Zagreb, with each cluster containing only one neighborhood. Neighborhoods in clusters '1' and '5' lie furthest on the north of Zagreb. In cluster '1', the most common venues fall in the category of 'Light Rail Station', and in cluster '5', in the 'Trail' category. 'Podsljeme' neighborhood, from cluster '5' is a part of Zagreb with family houses, green areas, but with not much social activities going on.
- Cluster '4' represent mostly neighborhoods that are unique to Zagreb, except that it contains a neighborhood 'Sirobuja', which is on the outer ring of Split.
- Cluster '3' represent mostly neighborhoods of Split, with the exception of one neighborhood of Zagreb, 'Maksimir'. Those neighborhoods of Split are part of an industrial area, so it's no wonder that some of the most common venues there are categorized as some kind of stores, etc.

# Discussion

## Cluster '2'

- There is a major overlapping of neighborhoods of Split and Zagreb in cluster '2'. Of all the neighborhoods in both towns, more than half of them belong to this cluster.
- In Split, we can see different types of neighborhoods (not belonging to cluster '2') as we go from the center of the town that is usually occupied by tourists further to north and east.
- Neighborhoods in Zagreb that do not belong to cluster '2' are located in the north, in the hilly part of the city.
- Neighborhoods in cluster '2' contain cafés, restaurants, bar's, etc. In Split, those are areas with lots of tourists, and in Zagreb, those are neighborhoods where lots of people work and socialize.

# Conclusion

Even though this kind of approach is highly sensitive to Foursquare venue categories, there are interesting insights that can be quickly obtained about similarities and dissimilarities of neighborhoods of Zagreb and Split that could potentially serve real estate traders. Since, for instance, venue category ‘restaurant’ and ‘mediterranean restaurant’ are completely different venue types, the project could be upgraded with finer comparison of the venues, and therefore, city neighborhoods.

Thank You!