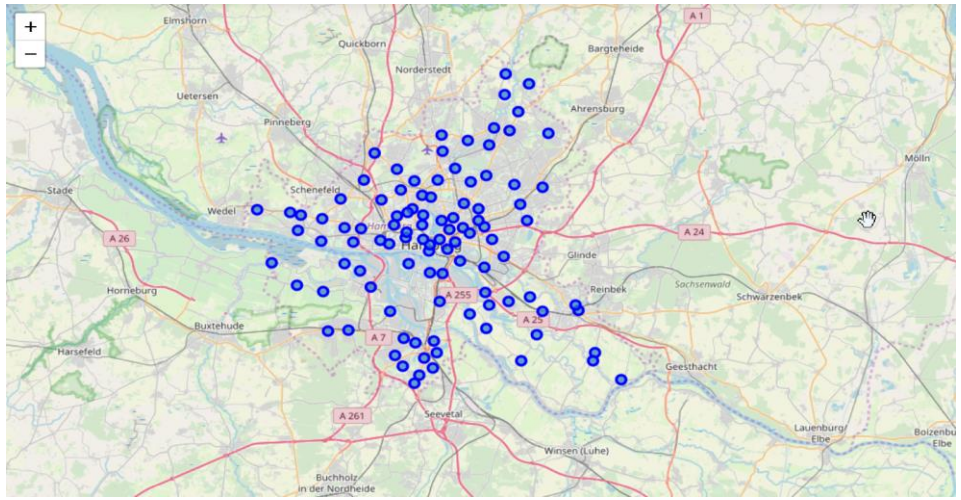
Where would you open a new shopping mall in Hamburg?

Coursera Capstone Final Project

21.01.2020

Objective & Business Problem

- The objective of this project is to use machine-learning methods (e.g. clustering) in Python to analyze and identify the optimal locations to open shopping malls in the city of Hamburg, Germany.
- The project might be useful and interesting for real estate developers and investors looking for solid investment in European market.



Data Requirements

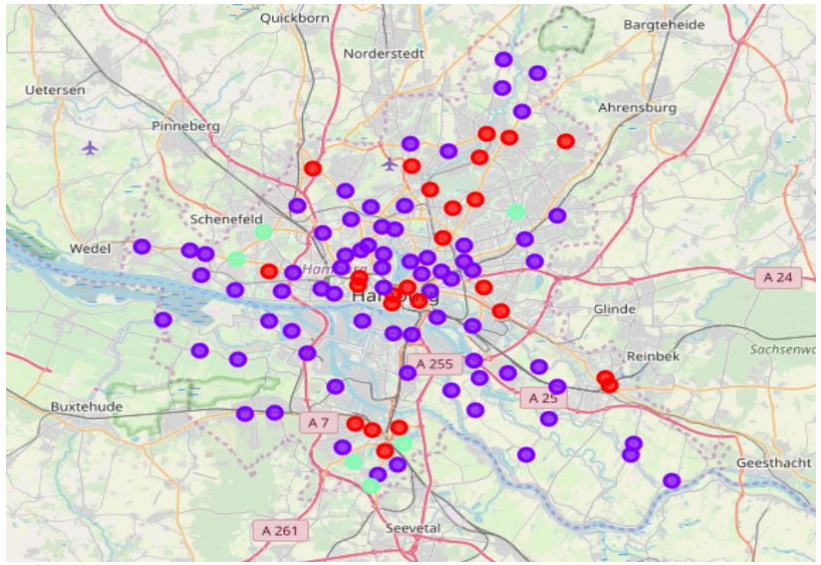
- List of neighborhoods in Hamburg (e.g. Alsterdorf, Blankenese, Hamm, etc) for scope definition: it will be extracted using web scraping from Hamburg Wikipedia pages;
- Latitude and longitude coordinates data of the above-mentioned neighborhoods: Python Geocoder package is used to get latitude and longitude coordinates;
- Venue data for the neighborhoods, which particularly includes shopping mall data, (e.g. location, tips, and categories): venue data is used for clustering and will be extracted with Foursquare API.

Methodology

- Web-scraping: downloading data from websites and extracting valuable information
- Python Geocoder: locating the coordinates of addresses, cities, countries, and landmarks across the globe
- Foursquare API: requesting venue data like name, unique ID, location, and category
- K-means Clustering: an iterative algorithm that tries to partition the dataset into K pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group.

Result

- The neighborhoods are grouped into 3 clusters based on the frequency of occurrence for “Shopping Mall”
  - Cluster 0 (in red): neighborhood with moderate number of shopping malls
  - Cluster 1 (in purple): neighborhood with very few number of shopping malls
  - Cluster 2 (in green): neighborhood with rather high number of shopping malls



Discussion

- As the Hamburg city is getting more and more crowded and expensive, we can expect that more city residents will move further away towards the boundary areas with other cities. Thus, those border areas with less shopping malls could be an interesting option as new shopping mall locations.
- Hamburg city center areas, for example Hafen City are already concentrated with a number of shopping malls; the land is expensive, rent is high and competition is fierce. Therefore, we won't recommend to invest in those "traditional" center areas in cluster 0.
- Along the rivers and lakes areas as depicted in cluster 1, where there are only a few to none shopping malls, we can consider investing in high-end boutique shopping malls. As the residents in those neighborhoods are mostly high-income earners, they would have the desire, as well as the purchasing power to buy more exclusive and unique items.