

IBM Coursera Capstone: Week 4 Assignment

Problem description:

Berlin is the capital of Germany and a very fast growing [city](#). Due to this constant growth prices for houses and flats have reached new heights and it gets more and more expensive to live in the [city](#). The city is also a hotspot for creative internet startups like Delivery Hero, Zalando and N26. All these companies offer interesting jobs for Data Scientists.

The goal of this project is to combine data about the housing market with data about venues from Foursquare to see how homogeneous the city.

Data:

The sets used in this project will be data from <https://berlin.kauperts.de/Strassenverzeichnis> to get information about the boroughs and neighborhoods of Berlin. After this the geodata of the neighborhoods will be gathered together with the Foursquare data of the venues in the neighborhood. At the end KMeans algorithm will be used to see how similar the different neighborhoods are. This will be combined with data from <https://de.statista.com/statistik/daten/studie/252943/umfrage/kaufpreise-fuer-eigentumswohnungen-in-berlin-nach-bezirken/> which shows prices for new homes in Berlin from 2011 to 2019.