

IBM Data Science Capstone Project

Problem Definition and Data

Problem Description:

For this project I have chosen to investigate the districts in Stuttgart to find out where there are good places to place an open air restaurant.

Background:

A good place to place a open air restaurant (german Biergarten) might be a location with high throughput of people locals as well as tourists and might be marked as a place with several parks in urban area and some other outdoor takeaway stuff like cafes or icecream shops etc. Another option could be also that there are several metro stations nearby.

Data:

1. Wikipedia for getting the Districts of Stuttgart
2. Geocoder to get Locationdata of Districts
3. Foursquare API to get venues in that area

Idea to solve the Problem:

Fetching the top venues per district and then clustering the districts by applying K-Means algorithm. Then interpret the clusters and draw a conclusion which cluster might fit best