

Coursera Capstone Project

“Battle of the Neighborhoods”

DOMINIK FREISINGER
APRIL 2020

Introduction

The topic of this project will be city tourism in Europe. Usually people tend to travel to the commonly known and popular places like Paris, London or Rome. Our goal will be to identify similar cities in Europe, based primarily on the Foursquare data. As a result a recommendation engine for travelers could be established. When a traveller likes or dislikes a city, he can find a similar city that he may also like. This way also lesser known cities could be recommended.

Data

The data used for this project will consist of the Foursquare Places dataset, that gives information on the number of restaurants, parks, cultural venues, etc. The feature set of a given city will be based on Foursquare's venue categories:

- Arts & Entertainment
- College & University
- Event
- Food
- Nightlife Spot
- Outdoors & Recreation
- Professional & Other Places
- Shop & Service
- Travel & Transport

Therefore a city will primarily be defined by the total number of venues in the given category.

Additionally climate data (average temperature per year) will be pulled from https://en.wikipedia.org/wiki/List_of_cities_by_average_temperature#Europe. Another feature that we will add to our dataset will be the population (i.e. the size of a city), which can be obtained here: <https://worldpopulationreview.com/continents/cities-in-europe/>.

The final feature set will thus contain of the Foursquare categories listed above, the average temperature and the population number. What is more, we will restrict our dataset to the 51 European capitals that can be found here: https://www.nationsonline.org/oneworld/capitals_europe.htm.