Dividing European Cities into Clusters

Introduction

The number of people emigrating from the EU Member States has fluctuated in recent years. In 2018, 3 million persons emigrated from an EU Member State, including both people emigrating to another EU Member State and those to a country outside the EU, a number which increased by 7 % since 2013. In 2018, a majority (54 %) of those emigrating were nationals, meaning citizens from the reporting country, 25 % were citizens of another EU Member State, while 22 % were non-EU citizens.

In the EU in 2019, the share of other EU citizens (those with a citizenship of another Member State than the one they were living in) in total employment was 4 % and for non-EU citizens 5 %.

These statistics show the increase of people moving within the European union. However, it can be difficult to decide whether to move to another city or which city to choose. A good deciding factor is to look whether a potential destination presents similar venues to another city one might like.

The objective of this project is to divide the most populated European Union cities into clusters based on the most common venues they have. The motivation is to help people moving from a city in Europe find another city that is similar.

Data

The objective of this project is to help people making decision about which city to move into. Since people are more likely to immigrate to big cities, the data will be based on the most populated cities in the European union.

The data will be taken from the Wikipedia page:
https://en.wikipedia.org/wiki/List_of_cities_in_the_European_Union_by_population_within_city_limits.

It contains the 92 most populated cities in the European union. The page will be scraped and a DataFrame containing every city and the country it is located in will be created.

Obtaining the latitude and longitude of each city will be accomplished using Geocode. Once the coordinates of every city are obtained, Foursquare will be used to determine the 20 most common venues of each city.