

A Foursquared European Trip

1 Business Understanding

Problem: There are many people who keep saying that the inner cities in Europe basically all look the same: a big square with a cathedral, nearby mostly the same types of restaurants, bars, shops – which mostly belong to a bigger chain; a castle, a big river, the cool trendy quarter, the not so cool run-down quarter, the train station. Thus, a european trip can be boring for *tourists (our audience)*.

Idea: Why not using open data to avoid boring holidays? Why not using e.g. *Foursquare data* to determine a trip with the "most diverse" cities or city centers in Europe?

The goals are:

- To **compare** about 70 **cities** or their city centers in Italy, France, Spain, Portugal
- To **select** 3 italian, 3 french, 3 spanish and 1 portugese **cities** as such so that they are as **diverse** as possible.
- Optionally, **recommend** some **venues** in the **selected cities**

The results of this analysis might be useful for tourists planning to visit European cities.

2 Data

To conquer this problem we need:

- A **list of cities** from Italy (20 cities), France (20), Spain (20), Portugal (10). Please don't feel discriminated right now if you are Portuguese. The list of cities shall be retrieved from **Wikipedia**
- The **coordinates** of each city (which should also roughly be the city center) based on **geopy** / **Nomatim** data
- Access to **Foursquare** to determine venues of interest
- For visualization purposes we need the **Folium library**

The data will be processed using **Python** in a Jupyter environment.

A **clustering** method (most likely k-means) will be used to compare the cities. The algorithm shall be optimized by using different evaluation methods.