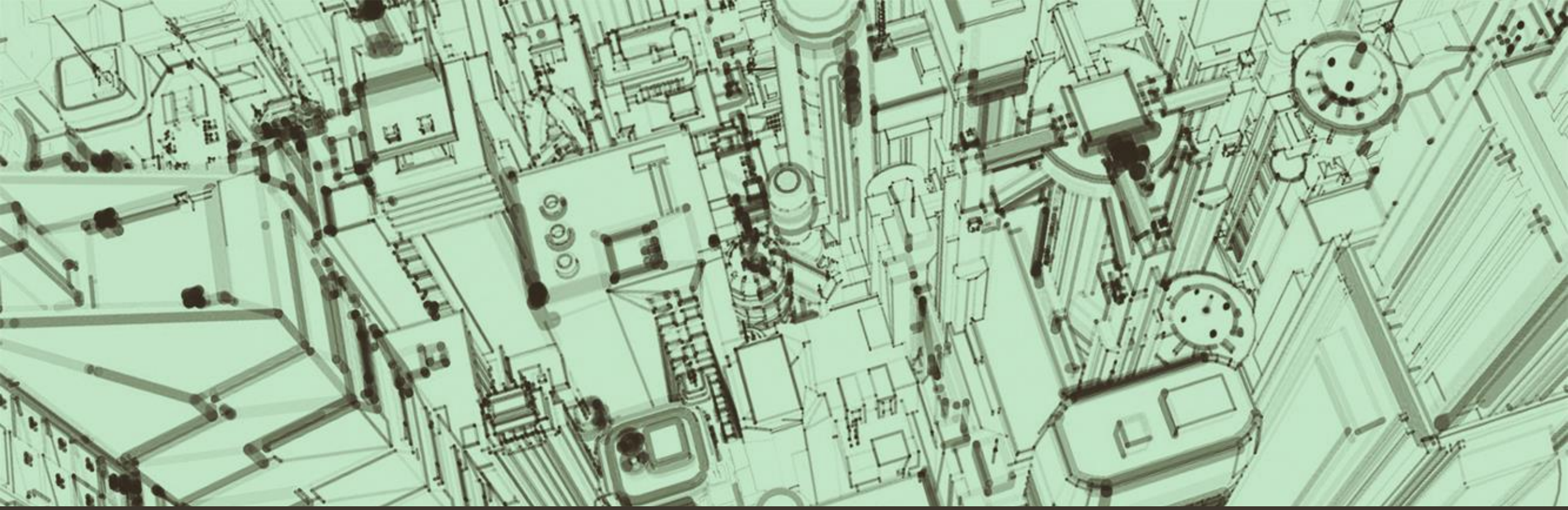


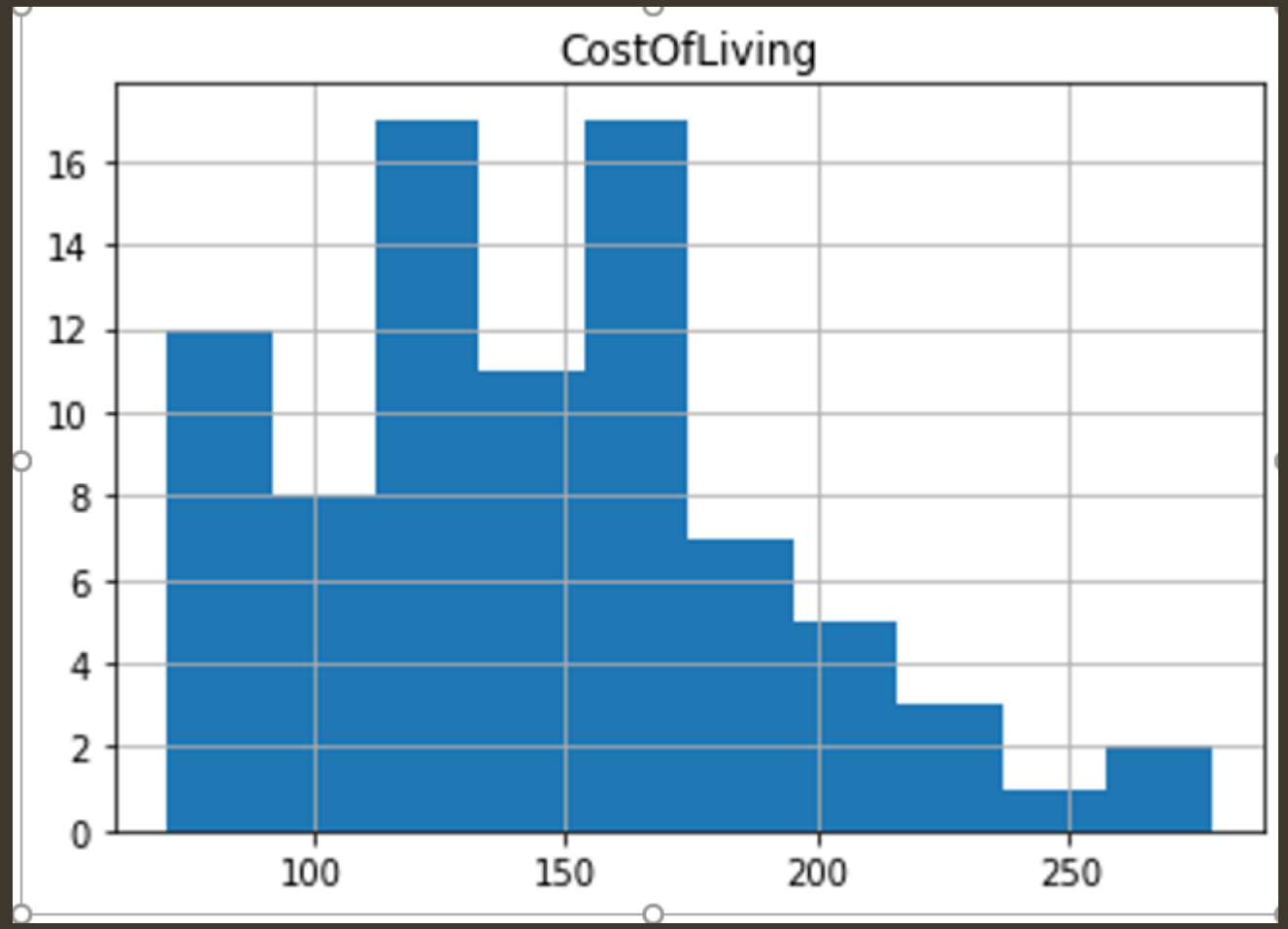
European cities score and mapping

Calculating European cities score

- Travelling is one of the hottest trend nowadays
- Travellers want to save money, and enjoy their time in some beautiful places where they can find all the confort and attractions they need
- For the biggest cities in Europe, we calculate a score number
- This score is calculated using numbers of venues and cost of living parameters.
- The highest this score, the better it is for the final traveller
- High score means low index price and many venues to visit.

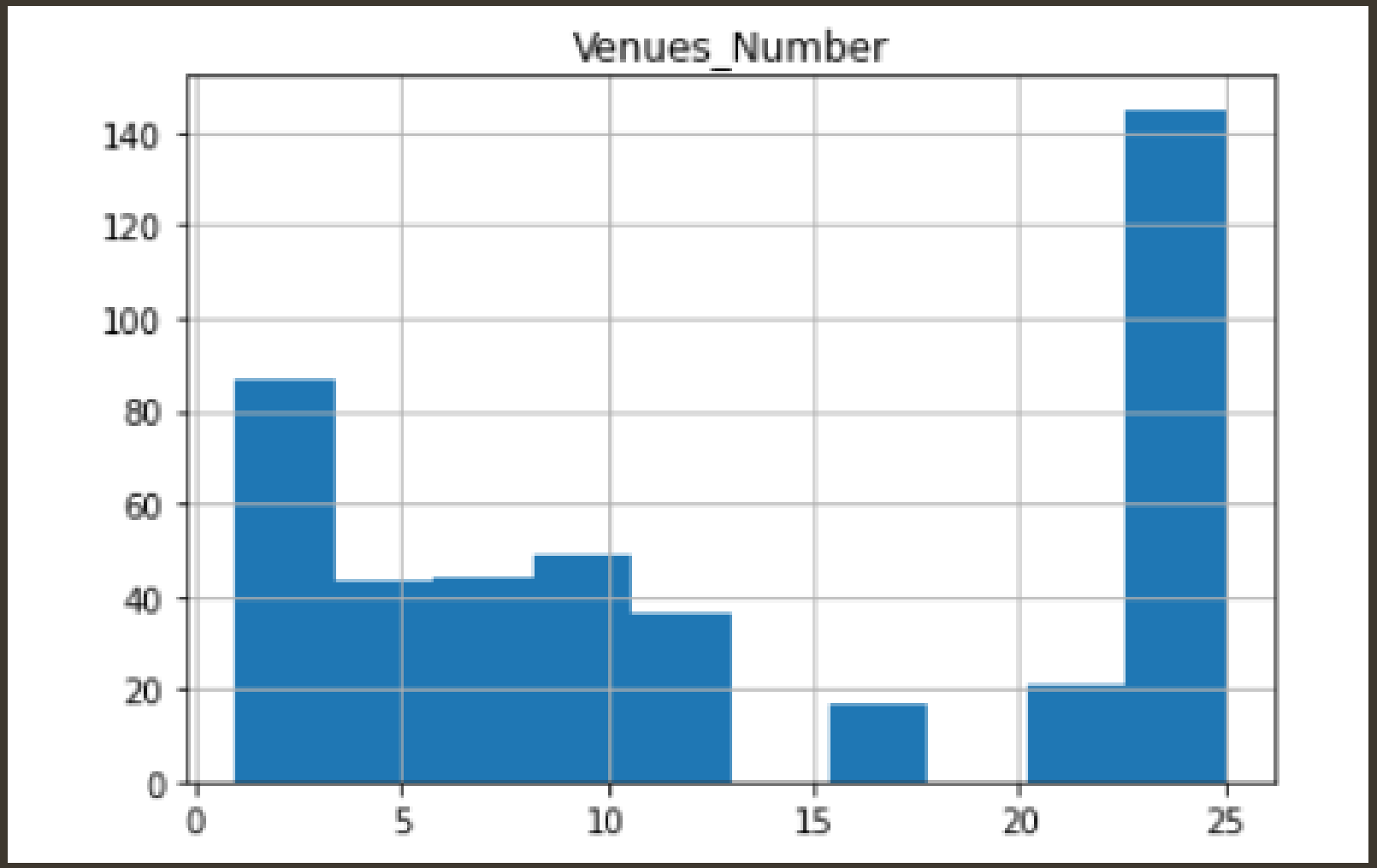
Data acquisition and cleaning

- Costs of living are taken from https://www.expatistan.com/cost-of-living/index/europe
- Venues are taken from Foursquare data provider.
- Latitune and longitudes are calculated using Geolocator.
- Data provided was not always clean. We removed all duplicates and None values from our dataframe.
- In total, we have found 442 venues (considering a distance of 1000 meters) for our cities.



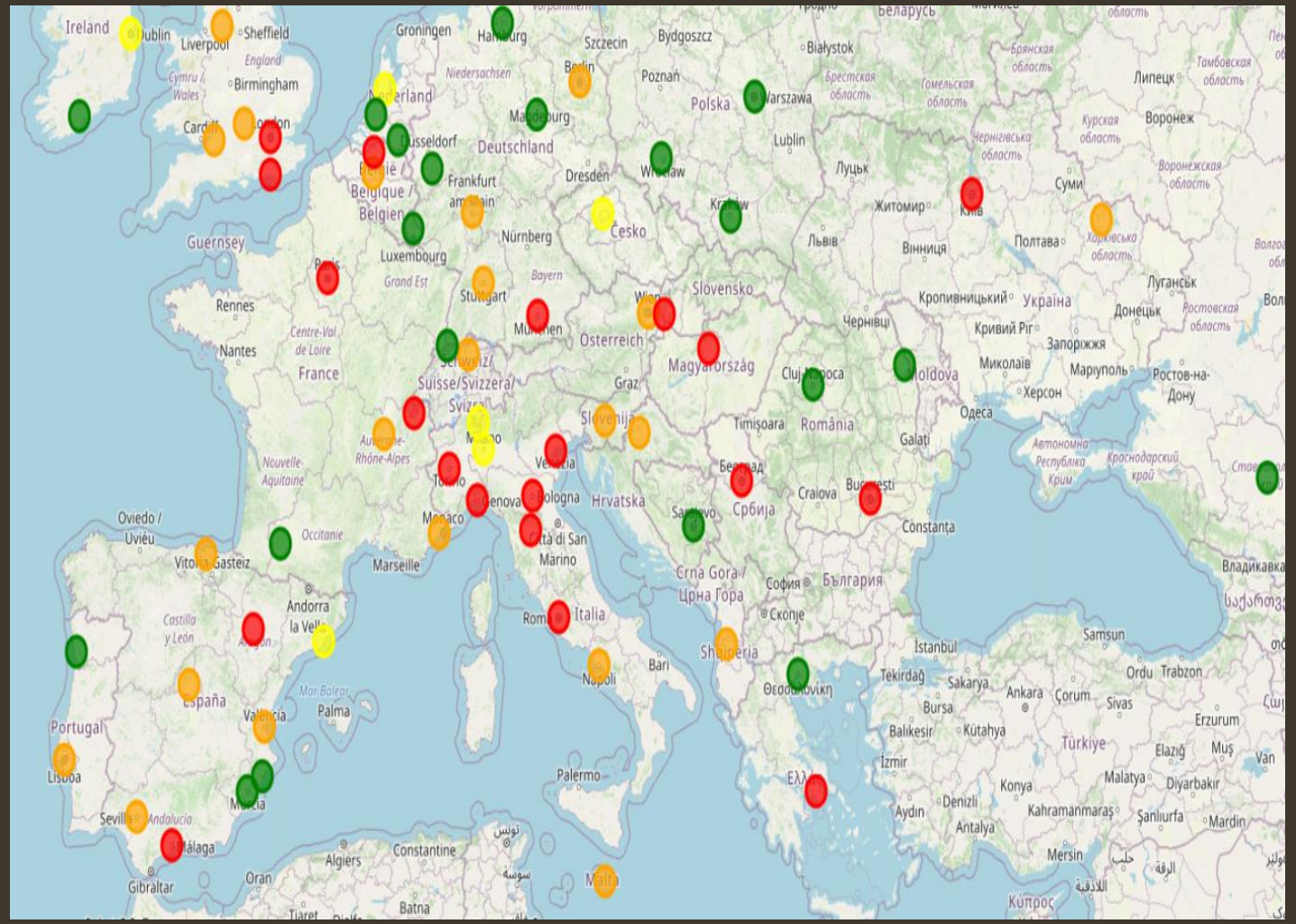
Histogram of costs of living for all European cities

- There are a few cities in Europe with a very high cost of living (above 200). London is one example.
- East European cities usually have low cost of living.



Histogram of venues number for all European cities

- Few cities have many venues nearby (Eg London)
- Many cities in Europe have just a few «Italian related» venues.



Cities mapping using folium

- The algorithm used is k-means
- Green color means good scoring, red color means bad scoring
- Score is calculate considering cheap prices and venues nearby

Conclusions

- From our machine learning (k-means) application, we can say that good places for travelers are in Spain, Germany or even est Europe.
- London for instance has many venues, but because it has high cost of living, the score is low.
- Results showns are calculate using the search keyword «Italian» and a distance of 1000 meters (these are parameters we have passed to Foursquare service). If these parameters were different, also the results shown were different.