Rate how touristic a City Is

Michael Schmitt

2020-02-10

1 Introduction

1.1. Background

I want to move to another City on the Continent Europe. An important factor is “how touristy is a city”. There are cities like Munich that are touristic but you can have a good life. One extreme is for example Venice: The city is very nice but extremly touristy - it is almost impossible to life a regular live here.

1.2 Problem

How can I determine how touristy a city is?

1.3 Interest

The factor is very interesting for people that want to change their residence.

2 Data acquisition and cleaning

2.1 Data Sources

Wikidata is a good source for informations about cities.
I will retrieve informations about venues from Squarespace

2.2 Data cleaning

The API responses are very good. One Problem is the limitation of the SquareSpace Free Plan.

2.3 Feature Selection

For the Touristic score is the amount of touristic / non touristic venues relevant. I just cluster venue categories to relevant for tourists and non relevant data
 'Furniture / Home Store' or 'Department Store' are not relevant for tourists but a Restaurant or Gift Shop is.

3 Exploratory Data Analysis

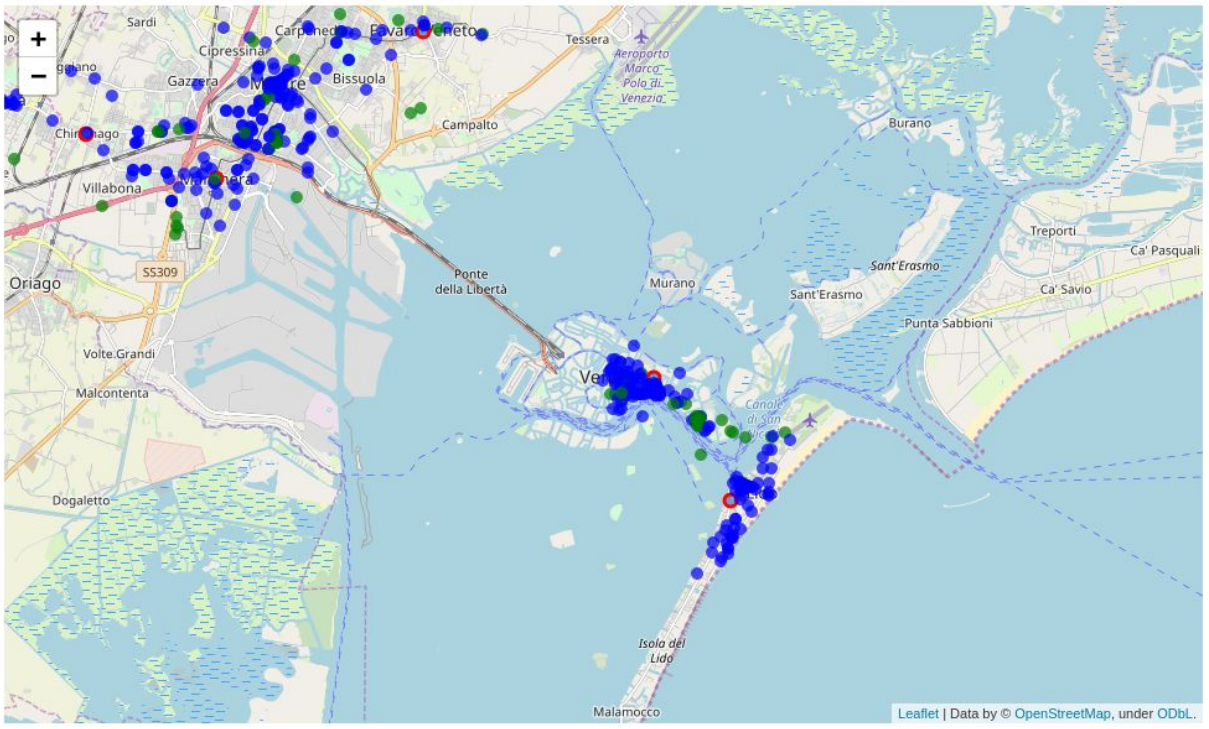
Wikidata provides a lot of relevant Information about cities:
Example Venice
Venice
Population of the City 261321.0
Area 415.9 km^2
Population Density 628.33 Humans / km^2

EntityID	Area	Bouroughs	Image	Lat	Lon	Name	Population	PopulationDensity
Q641	415.9	[{'Name': 'Municipality 1 Venezia-Murano-Buran...	https://upload.wikimedia.org/wikipedia/commons...	45.439722	12.331944	Venice	261321.0	628.326521

We get all relevant Geo Data:

- Name, Lat , Lon of the City
- Name, Lat , Lon of all Bouroughs

We can label Venues with a touristic interest (blue) and a non touristic interest (green)



Now we fetch the Squarespace Data for all Cities and all Bouroughs, calculate the Relation Touristic, Non Touristic, apply the Z Score

$$Z = \frac{X - \mu}{\sigma}$$

and sort the Table

]:

	Name	zscore
11	Venice	3.086572
4	Amsterdam	0.419817
5	Jerusalem	0.347567
10	Las Vegas	-0.109458
7	Essen	-0.198324
3	Barcelona	-0.298467
1	Saarbrücken	-0.317738
0	Trier	-0.408492
8	New York City	-0.487549
6	Bochum	-0.501531
2	Munich	-0.537674

4 Conclusion

The Question “How touristy is a city?” is hard to answer. I tried to do this with a simple Analyze of Venue Data. At this point I’m not very happy with the Analyze Part. Als the Data returned from Squarespace is not the best. My data also does not represent the real world - there are only places and venues that are entered by regular humans and no guarantee, that all venues of a city are represented.
None the less I can say, that my Table of touristyness of cities that I built by emotion is represented well in the Table that relies on facts from the API.

5 Future Directions

I would like to get more Information like “Visitors by year” and correlate them with the population of the city. And put more Features in the Equation for the Touristy Score.