Research and Innovation

Martí Municoy, Alba Gordó, Jan-Hendrik Niemann
and Andreas Radke

November 14, 2017

Abstract

The aim of this project is to find and visualize the events posted on Meetup.com in different cities in order to analyze the density of planned activities and to make a social study regarding typical meeting times and most usual activities and interests. One aspect may concern different meeting times in Spain and Germany or group size. We will focus our research on cultural activities like languages tandems and sports. The data shall be obtained via the Meetup-API and OpenStreetMap. The most important data source for this project is the Meetup-API where we have to request the events with different Python packages like "requests" or "meetup-api". Our second task is to visualize the locations of the events on a given city map obtained by data resources like OpenStreetMap. To plot the locations we could use "matplotlib" or similar Python plotting libraries. One further aspect could be to look for interest of twitter users to make recommends to activities nearby. The geotag of the tweets provides the central information for the search.

Contents

1 Summary and Goals

The human being is social. Every day there are millions of events. Some thousand of them are posted at the webpage `Meetup.com`. It is a website where people can seek and find other people with the same interest. It helps to organize your events and makes them public. For this reason we can use `Meetup.com` to analyze the behavior and the preferences of cities and its inhabitants all over the world since almost all events posted on `Meetup.com` have a place which can be expressed as latitude and longitude. By using the request-API we extract the dataset "latitude-longitude" from `Meetup.com`and plot it on a map given by Google Maps. In this way we create a density map, a so-called heatmap, where on can see all the acitivities in the selected city.
Furthermore, we also consider different types of events like "food & drinks" of "sports". Combining these information with the geographical and/or political structure of a city one can identify trending districts and neighborhoods in a city. To do so we make use of geojson-files which can contain the polygonal shape of a district and its name. There are several websites where one can get geojson data or create his own data.
A third data source is `Wikipedia.org` where we performe a web scrapping to extract the population number of a city and, if available, the polulation of a district. Having this information we create a factor to evaluate the degree of acitivity. This factor is given as

$$\text{degree of activity} = \frac{\text{events}}{\text{person}}.$$

Our data is available offline and online.
This work is structured as follows: In section 2 we discuss the data life-cycle of our data. In the section 3 we present our used methods and have a closer look at Gmaps-API, Python and Jupyther, geojson and the webscrapping. Afterwards, in section 4 our results are presented. In section 5, we discuss legal and ethical issues. Finally, in section 6 we discuss the limitations of our work and make proposals for further work.

2 Data Life-Cycle

3 Tools and Data

3.1 Tools

The main work for our project was done in the Python programming language in version 3.6 [7]. The most notable used Python libraries where requests [3], gmaps [6], Jupyter [4] and matplotlib [2]. The requests package was used to call and request data from the Meetup API [5].

With Jupyter Notebook and gmaps we were able to plot the obtained data from the Meetup API, Wikipedia and other resources in a browser. In detail Jupyter provides the general browser session functionality whereas gmaps extends Jupyter with an interactive map provided by Google Maps [1]. One can visualize specific points or areas in various colors and densities to create meaningful maps. Matplotlib was used to transform the data from population density into specific color schemes.

3.2 Data

As already noted we used the following data sources:

- Meetup API: To get the events in the examined cities.
- Wikipedia: For obtaining data about the population density of the cities and their districts.
- Various resources for GeoJSON files: To visualize districts onto Google Maps. TODO
- Google Maps: As a general back-end for our visualization.

4 Results

5 Legal and Ethical Issues

6 Limitations and Future Work

References

[1] Google Inc. Google Maps. `https://maps.google.com`, 2017. Interactive online Map Service.

[2] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing In Science & Engineering*, 9(3):90–95, 2007.

[3] Kenneth Reitz. requests - Python HTTP for Humans, version 2.14.2. `https://pypi.python.org/pypi/requests`, 2017. Python Library for sending HTTP requests.

[4] Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, Carol Willing, and Jupyter Development Team. Jupyter notebooks – a publishing format for reproducible computational workflows. pages 87 – 90, 2016.

[5] Meetup Inc. Meetup API. `https://www.meetup.com/meetup_api/`, 2017. [Online; accessed 11-November-2017].

[6] Pascal Buignon. gmaps, version 0.70. `https://github.com/pbugnion/gmaps`, 2017. Plugin for including interactive Google maps in the Jupyter Notebook.

[7] Python Software Foundation. Python Language reference, version 3.6. `https://python.org`, 2017.