|  |
| --- |
| OTH Regensburg |
| Spatial Databases |
| Project Report |

|  |
| --- |
| Birgit Huber, Florian Fritz, Florian Fußeder  2.1.2017 |

Inhaltsverzeichnis

[Project Summary 2](#_Toc471138709)

[Decisions 2](#_Toc471138710)

[Questions for the User 2](#_Toc471138711)

# Project Summary

The goal of the project was to create a web application that helps a user to find a suitable region for accommodation in a city. By answering a series of questions about the preferences and living circumstances we aim to suggest a ranked list of the zones of a city.

Furthermore, the user should be able to enter some main locations. With the given data we are able to calculate a time- and cost-efficient route.

After some research we choose the city “New York” and discussed the which datasets we might use. Using the datasets, we defined our questions for the user and started with processing the data for an easy integration in our project. Afterwards we were able to start with the development of the final webpage.

# Decisions

We started off by considering 9 which we thought might be interesting and might have good datasets. Each member of the team had a closer look at the city and its availed datasets as following.

* Birgit: San Francisco, Beijing, London
* Florian Fr: Berlin, New York, Tokyo
* Florian Fu: Munich, Seoul, Singapore

After a week we had a discussion about the results and choose New York. One of the main reasons for that is that the United States has a lot of open geo datasets because of an open data initiative[[1]](#footnote-1) . New York has its own website with data[[2]](#footnote-2) which we used to get the data for our project. New York is also by default divided into 5 Zones with total area of aprox. 790 km² and 194 Neighborhood-Tabulation-Areas (NTA) with an average size of four km² which perfectly suited the needs of our project. With the questions, we wanted to ask the user we choose to use the following dataset.

* Parks
* Play areas
* Restaurants
* Soccer Fields
* School Points
* Parking lots
* Rental Prices
* Colleges and University’s
* Population
* Complaint Data
* Subways

# Questions for the User

Resulting out of our chosen datasets we picked the following questions to determine a suitable NTA for the user.

* Age range
* Has Children
* Is a student
* Owns a car
* Has a dog
* Does outdoor sports
* Uses subway
* Likes nature
* Prefers vibrant or quiet areas
* Importance of low rental prices
* Prefer to live central
* Favors specific zones

# Processing the Data

Each Dataset was processed individuality. The result of each dataset a database view containing a rating value between zero to one for each region in terms of the data. For example, the area with the least parking lot area scores the lowest rating, the area with the most scores one. This allows us to easily weight the different ratings in the web application as they are now all in the same range.

## Neighborhood-Tabulation-Areas

As mentioned we chose the NTA areas as foundation of our ratings and zones which we want to suggest to our user. The NTA could easily be imported from the exiting shapefiles[[3]](#footnote-3). The Shapefile contained a NTA code, the NTA name, the geometry and some other metadata. Nothing had do be processed in this table.

## School Points

The School point data[[4]](#footnote-4) was imported as shapefile containing locations based on the official address. It also includes some basic school information such as Name, Address, Principal, and Principal’s contact information.

The rating of how good a NTA is, in terms of School Points, was determined by the amount of points within the area. This was done by using basic postgis queries[[5]](#footnote-5) ,including ST\_AREA and ST\_CONTAINS which we have learned in our lecture. As final step a normalized view which only contained NTA code and the rating[[6]](#footnote-6).

1. https://www.data.gov/ [↑](#footnote-ref-1)
2. https://nycopendata.socrata.com/ [↑](#footnote-ref-2)
3. https://data.cityofnewyork.us/City-Government/Neighborhood-Tabulation-Areas/cpf4-rkhq [↑](#footnote-ref-3)
4. https://data.cityofnewyork.us/Education/School-Point-Locations/jfju-ynrr/data [↑](#footnote-ref-4)
5. https://github.com/FlorianFusseder/SpatialDatabases/blob/master/FuFlo%20Data/TABLES/4.publicSchoolPoints/SchoolData.sql [↑](#footnote-ref-5)
6. https://github.com/FlorianFusseder/SpatialDatabases/blob/master/FuFlo%20Data/TABLES/4.publicSchoolPoints/view.txt [↑](#footnote-ref-6)