FINDING CHILD CARE IN ZURICH

Applied Data Science Capstone by IBM/Coursera

Patrik Winiger

TABLE OF CONTENTS

INTRODUCTION

BACKGROUND

In this project I will try to identify places in the city of Zürich where daycare is sparse. Stakeholders for this project are parents who seek to live close to daycare, as well as daycare providers and the city government, which plans to provide day care for working parents.

METHODOLOGY

In this project we will direct our efforts on detecting areas of Zürich that have low child care density. We will limit our analysis to area ~6km around city center.

In the first step I have collected the required **data: location and type (category) of every child care venue within 6km from Zürich center** (Limmatplatz).

Second step in our analysis will be calculation and exploration of '**child care venue density**' across different areas of Zürich - I will use **heatmaps** to identify a few promising areas close to center with low number of Child_Care_Services in general and focus our attention on those areas.

In third and final step I will focus on most promising areas and within those create **clusters of locations that meet some basic requirements** established in discussion with stakeholders: we will take into consideration locations with **no more than two Child_Care_Services in radius of 250 meters**. We will present maps of all such locations but also create clusters (using **k-means clustering**) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

DATA

Based on definition of our problem, factors that will influence our decission are:

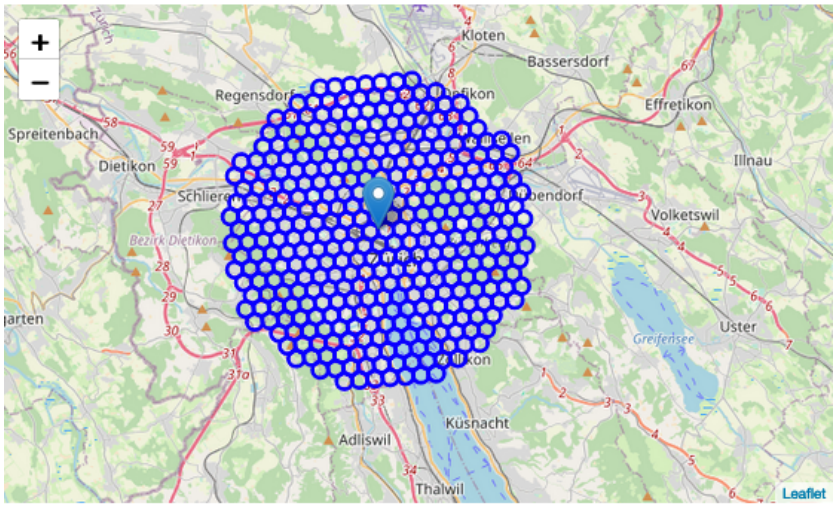
- *number of existing Child_Care_Services in a 300m circle*
- *number of and distance to Child_Care_Services in the circle*
- *I decided to use regularly spaced grid of locations, centered around city center, to define our circles of interest.*

Following data sources will be needed to extract/generate the required information:

- *centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using **Google Maps API reverse geocoding***
- *number of Child_Care_Services and their type and location in every neighborhood will be obtained using **Foursquare API***
- *coordinate of Zürich center will be obtained using **Google Maps API geocoding** of relatively central location (Limmatplatz)*

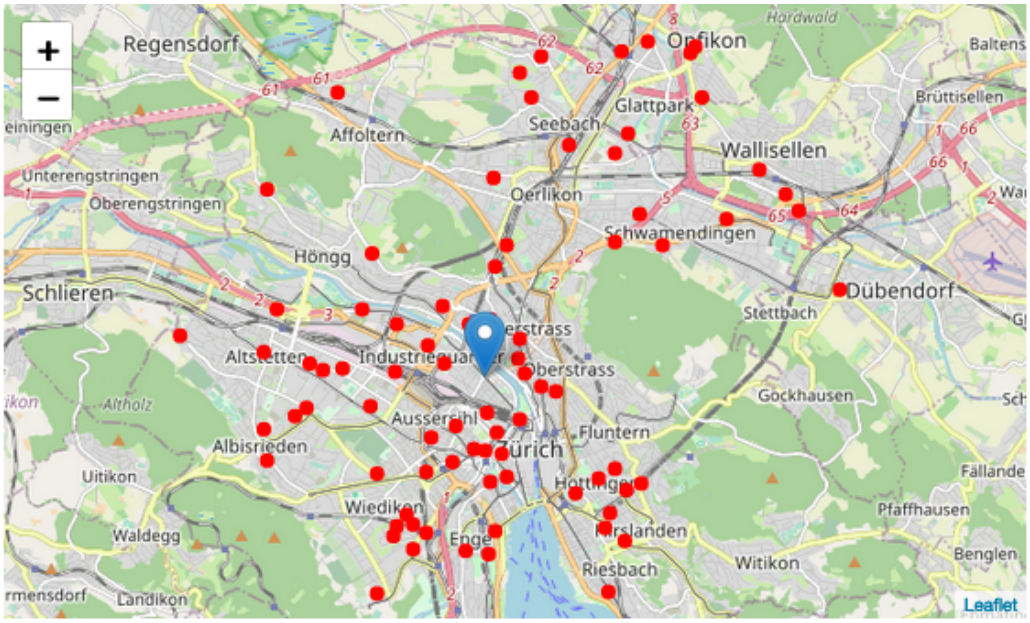
CIRCLES OF INTEREST

I created a grid of area candidates, equaly spaced, centered around city center and within ~6km from Limmatplatz. Our neighborhoods will be defined as circular areas with a radius of 300 meters, so our neighborhood centers will be 600 meters apart.



CHILD CARE SERVICES

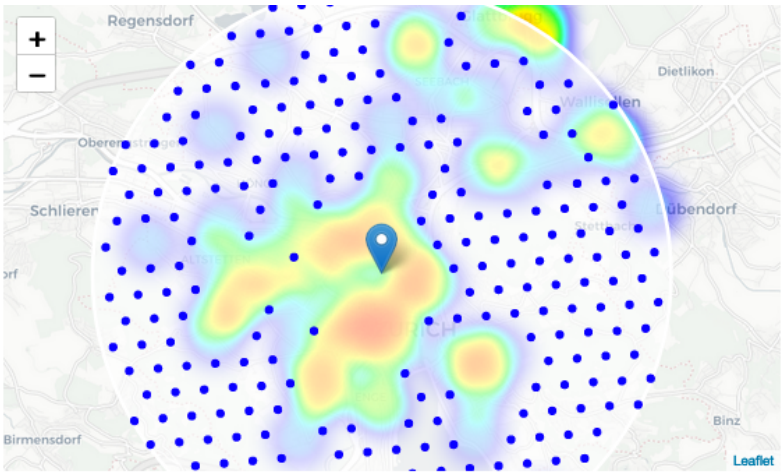
In the circle of interest we search for all child care services.



On average a child care venue can be found within ~900m from every area center candidate. That is quite sparse, so we can use the entire city and don't need to focus on single neighborhoods.

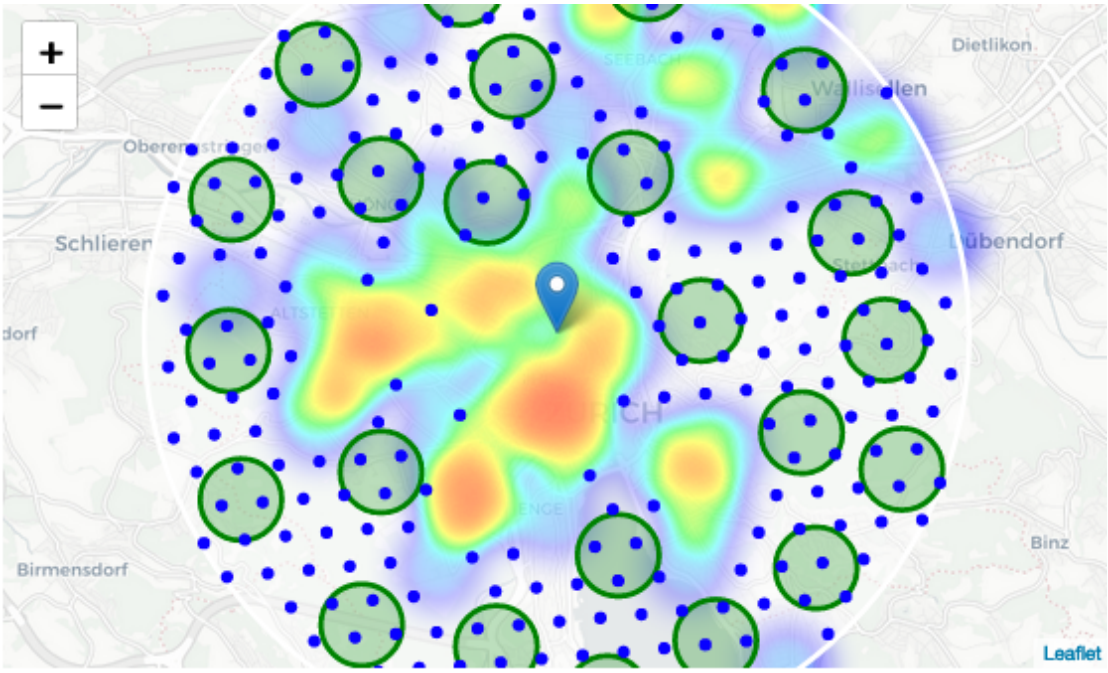
LOCATIONS WITH LOW DENSITY VENUES

Filter those locations: we're interested only in locations with no more than two Child_Care_Services in radius of 250 meters, and no Child_Care_Services in radius of 450 meters.



CLUSTERING

Cluster those locations to create **centers of zones containing good locations**. Those zones, their centers and addresses will be the final result of our analysis.



RESULTS

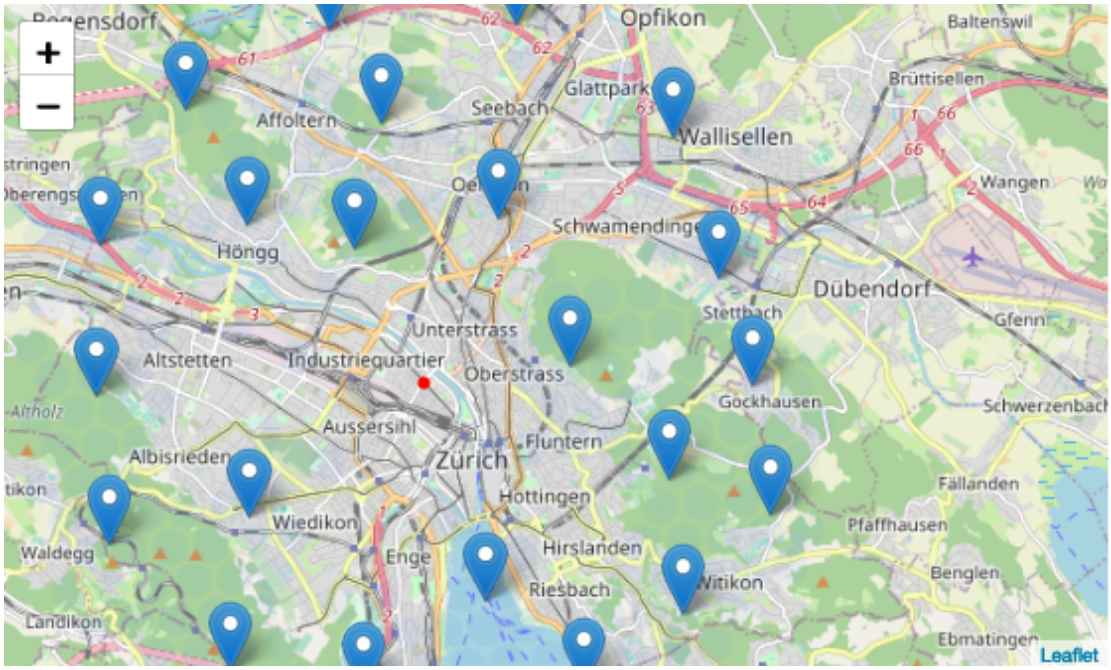
My analysis shows that the number of Child_Care_Services in Zürich is not that large, with 82 in our initial area of interest which was 12x12km around Limmatplatz. There are also pockets of low density fairly close to city center. Highest concentration of Child_Care_Services was detected south from Limmatplatz

I first created a dense grid of location candidates. Those locations were then filtered so that those with more than two Child_Care_Services in radius of 450m, which is half the mean distance of which child care venues are usually away.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 23 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues.

This, of course, does not imply that those zones are actually optimal locations. The k-means clustering did not distinguish between residential areas and those where a venue is logically impossible, like lakes and forests.



Reverse geocode those candidate area centers to get the addresses which can be presented to stakeholders.

```
============================================================
Addresses of centers of areas recommended for further analysis
============================================================

Limmatbrücke, A1H, 8102 Zürich                          => 5.1km from
Limmatplatz
Breitweg, 8032 Zürich                                   => 3.9km from
Limmatplatz
Hospital Helipad, Paul-Clairmont-Strasse 12, 8063 Zürich => 3.3km
from Limmatplatz
Seebacherstrasse, 8046 Zürich                           => 3.8km from
Limmatplatz
Wallisellen, Herti, 8304 Wallisellen                    => 5.1km from
Limmatplatz
Seestrasse 557, 8038 Zürich                             => 5.3km from
Limmatplatz
Bärenbohlstrasse, 8046 Zürich                           => 5.3km from
Limmatplatz
Bergstrasse 16, 8142 Uitikon                            => 5.2km from
Limmatplatz
Stettbacherrain 16, 8051 Zürich                         => 4.5km from
Limmatplatz
Familiengartenverein Susenberg, Hinterbergstrasse, 8044 Zürich =>
2.1km from Limmatplatz
Zürich District, Zürich                                 => 3.3km from
Limmatplatz
A4, 8052 Zürich                                         => 5.3km from
Limmatplatz
```

Gratstrasse 4, 8143 Zürich => 5.1km from Limmatplatz
Unnamed Road, 8049, 8049 Zürich => 5.3km from Limmatplatz
Lyrenweg 300, 8048 Zürich => 4.8km from Limmatplatz
Bellerivestrasse 308, 8008 Zürich => 5.0km from Limmatplatz
Waidbadstrasse 51, 8037 Zürich => 2.2km from Limmatplatz
Rütistrasse 32, 8044 Gockhausen => 4.8km from Limmatplatz
Ligusterstrasse 1, 8057 Zürich => 2.6km from Limmatplatz
Allmendstrasse 93, 8041 Zürich => 4.6km from Limmatplatz
In der Looren 61, 8053 Zürich => 5.4km from Limmatplatz
Ferdinand-Hodler-Strasse 41, 8049 Zürich => 3.4km from Limmatplatz
Rehalpstrasse 67, 8008 Zürich => 5.1km from Limmatplatz

CONCLUSION

Purpose of this project was to identify Zürich areas close to center with low number of Child_Care_Services in order to aid stakeholders, like parents, in narrowing down the search for optimal location for a new child care venue, or in the case of parents, locations to move to within the city. By calculating venue density distribution from Foursquare data we have identified general areas. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decission on optimal location will be made by stakeholders based on specific characteristics of locations in every recommended zone, taking into consideration additional factors like population density of parents of each location, real estate availability, prices, etc.