Identify location candidates for new tourist accommodation objects in Ljubljana, Slovenia

Capstone project report

Marko Kastelic
March 23, 2020

Business problem

The city of Ljubljana, capital of Slovenia, has been a very popular tourist destination in the recent years and it hosts more tourists each year. Also, more and more tourists decide to stay in Ljubljana for more than 1 day, and group visits are popular. Thus, the accommodation offering (hotels, motels, inns, bed & breakfasts etc.) needs to grow with the demand.

One of the most important factors when looking for a location for tourist accommodation is location. The goal of this project is to identify parts of the city of Ljubljana that could be location candidates for new tourist accommodation. The vicinity to the city center, vicinity to the existing food venues and vicinity to the public transport grid are the main factors that are considered in this project.

The report is targeting investors/stakeholders/contractors that are interested in creating new accommodation offerings in Ljubljana. The results should be interesting for either small sized (boutique) accommodations as well as higher capacity accommodation.

Data

The following requirements were considered when trying to find a potential location of tourist accommodation:

- No other tourist accommodation venue in vicinity.
- Gravitating towards the city center. Most of Ljubljana's attractions are in the city center.
- Close to existing food & drink venues.
- Close to public transportation grid.

Based on the requirements the main source of data was the Foursquare database for the following information:

- Location, name and category of existing accommodation venues (any subtype) in an area.
- Location, name and category of existing food venues (any subtype) in an area.
- Location, name and category of existing bus stations/stops in an area.

Transformations between addresses and geographic coordinates were performed using the Nomatim API.

All geographic coordinates calculations were performed in the WGS84 geodetic system using Pyproj API.

All the map visualizations were created using Folium API.

Methodology

I) Data collection

The venue information was extracted from Foursquare API. City's administrative regions (neighborhoods/boroughs/districts) could be used as source points for Foursquare API requests. However, those regions have incorrect shapes and very different sizes which can create gaps in the data. To avoid such search gaps a square grid was placed over the city of Ljubljana, spanning 10 kilometers and centered at the one of most popular landmarks, "Ljubljansko tromostovje" which can be considered as the city's center.

Each grid's unit (area) is a square with sides 1 kilometer in length. Every area was used as a bounding box for retrieving venues from Foursquare using the search API. The extracted data for each found venue (accommodation, bus station, food) was geographic location (latitude and longitude), name and category. In addition, any duplicates were discarded and distance to city's center was calculated for each venue.

In the end the grid contained 189 existing accommodation venues, 206 bus stations/stops and 798 food venues.

II) Analysis

I used mostly maps to visualize the data and to get a sense of accommodation venue density around the city.

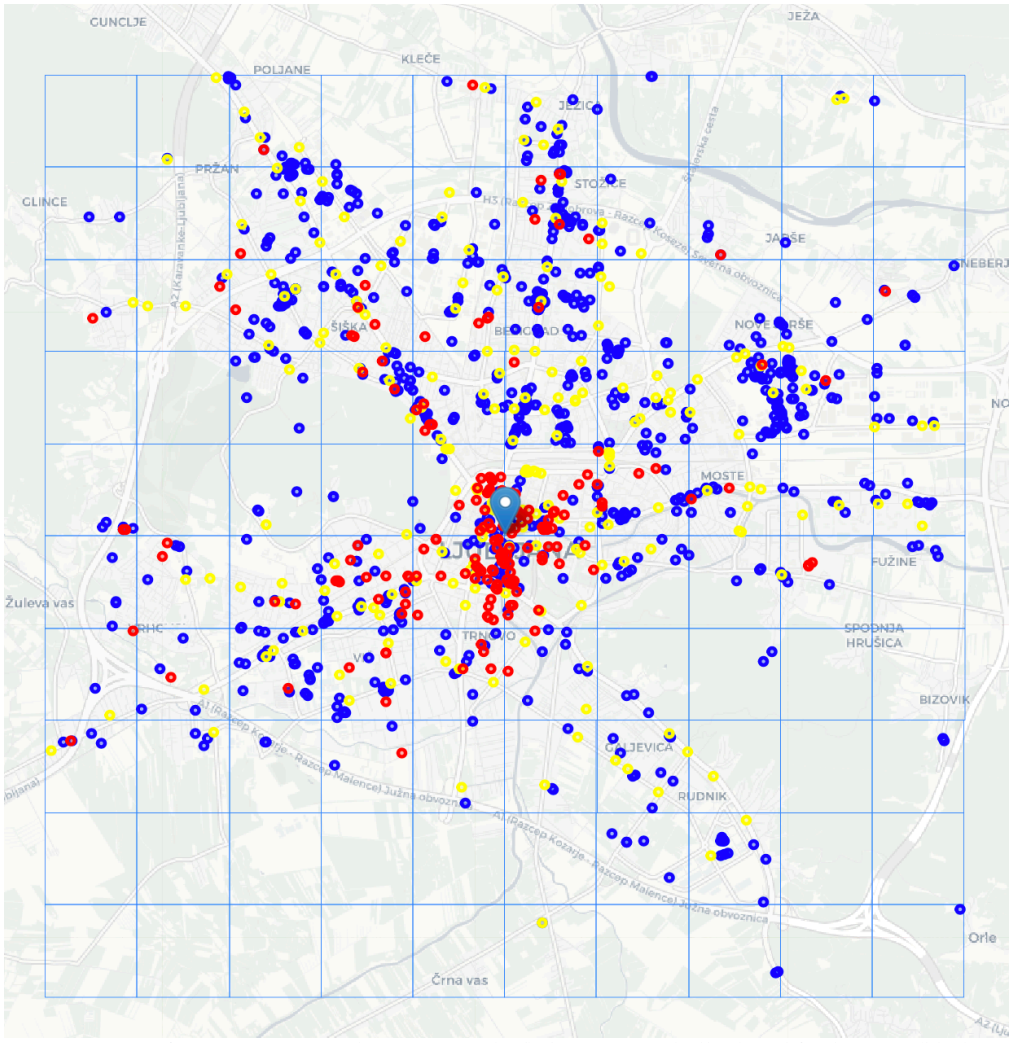

Figure 1. Map of existing accommodation venues (red), bus stations (yellow) and food venues (blue).

Figure 1 already shows that accommodation venues are grouping around the city center with most of the venues in the 1 km radius. The image is quite crowded, so a heap map showcases this better (Figure 2).

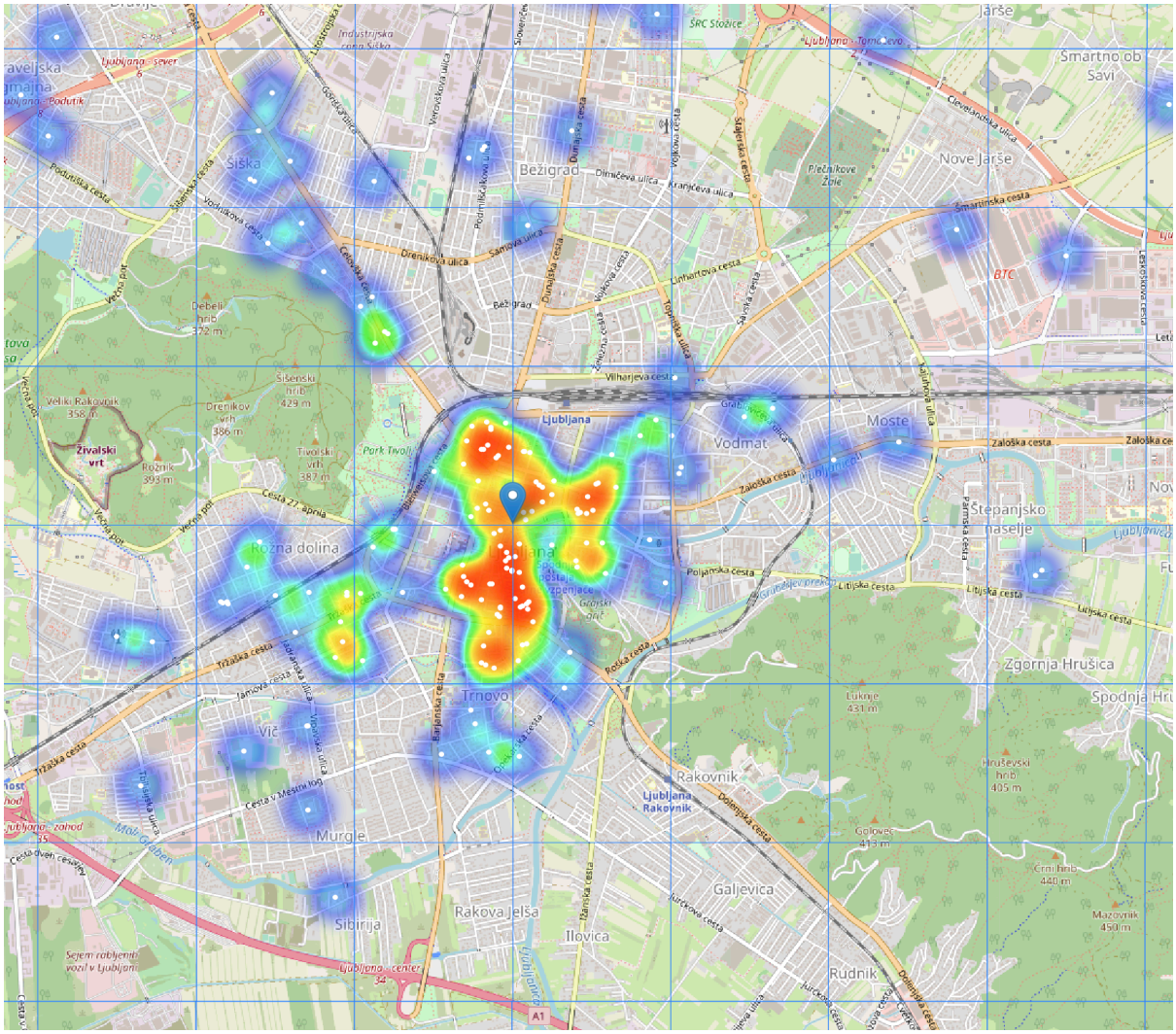

Figure 2. Heatmap of current accommodation density.

Accommodations include hotels and hostels (129 venues), bed & breakfasts (27 venues), motels, apartments, inns etc. So, most accommodation venues can be considered as high capacity and certainly more suitable for groups.

It is also interesting to see the relationship between the accommodation density and the distance from the city center:

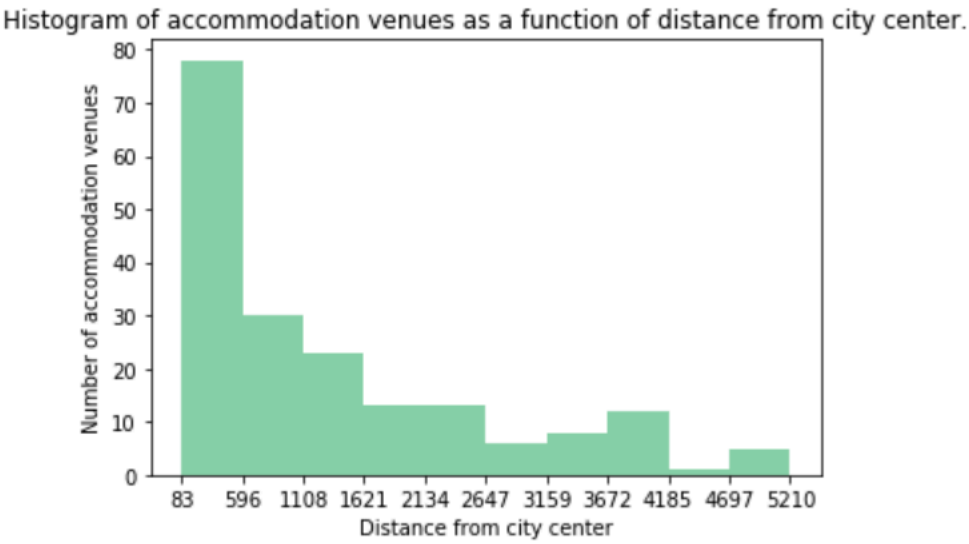

Figure 3. Relationship between the accommodation density and distance from the city center.

If Figure 3 one can see that 78 accommodation venues are within the 600-meter radius which is quite crowded already. Additionally, 66 accommodation venues are between 600 and 2100 meters.

So, the existing accommodation venues group mostly around the city center and 76% of existing accommodation venues are within the 2.1 km radius. This is expected as most of the tourist attractions are focused in the city center - the old and historic Ljubljana. Thus, I expected that the location candidates will be mostly outside of the 2km radius.

III) Location candidates

For the purposes of this project I selected the walking distance as the boundary of search from each location candidate. A walking distance is a distance up to 400 meters. This corresponds to a 5-minute walk (80 meters for 1 minute of walking time).

Accommodation location candidates were calculated in the following way. I used a finer grid within the existing grid areas. Each area was split into 100 squares (subareas), each 100 m x 100 m (100m equals to a good minute of walking time). The geographic centers of these subareas were used as accommodation location candidates. This generated 10000 potential location candidates.

For each location candidate I calculated:

- Distance to nearest existing accommodation venue, bus station/stop and food venue.
- Number of existing accommodation venues, bus stations and food venues within a walking distance from the location candidate.

When the data was prepared, I extracted the location candidates that matched the prerequisites defined at the beginning:

- No other accommodation object within the walking distance from the location candidate.
- At least one bus station within the walking distance from the location candidate.

- At least 3 food venues within the walking distance from the location candidate.
- Focusing on the area around the city center – up to radius of 3.5 km.

The results are showed in Figure 4 where we can clearly see pockets of location candidates.

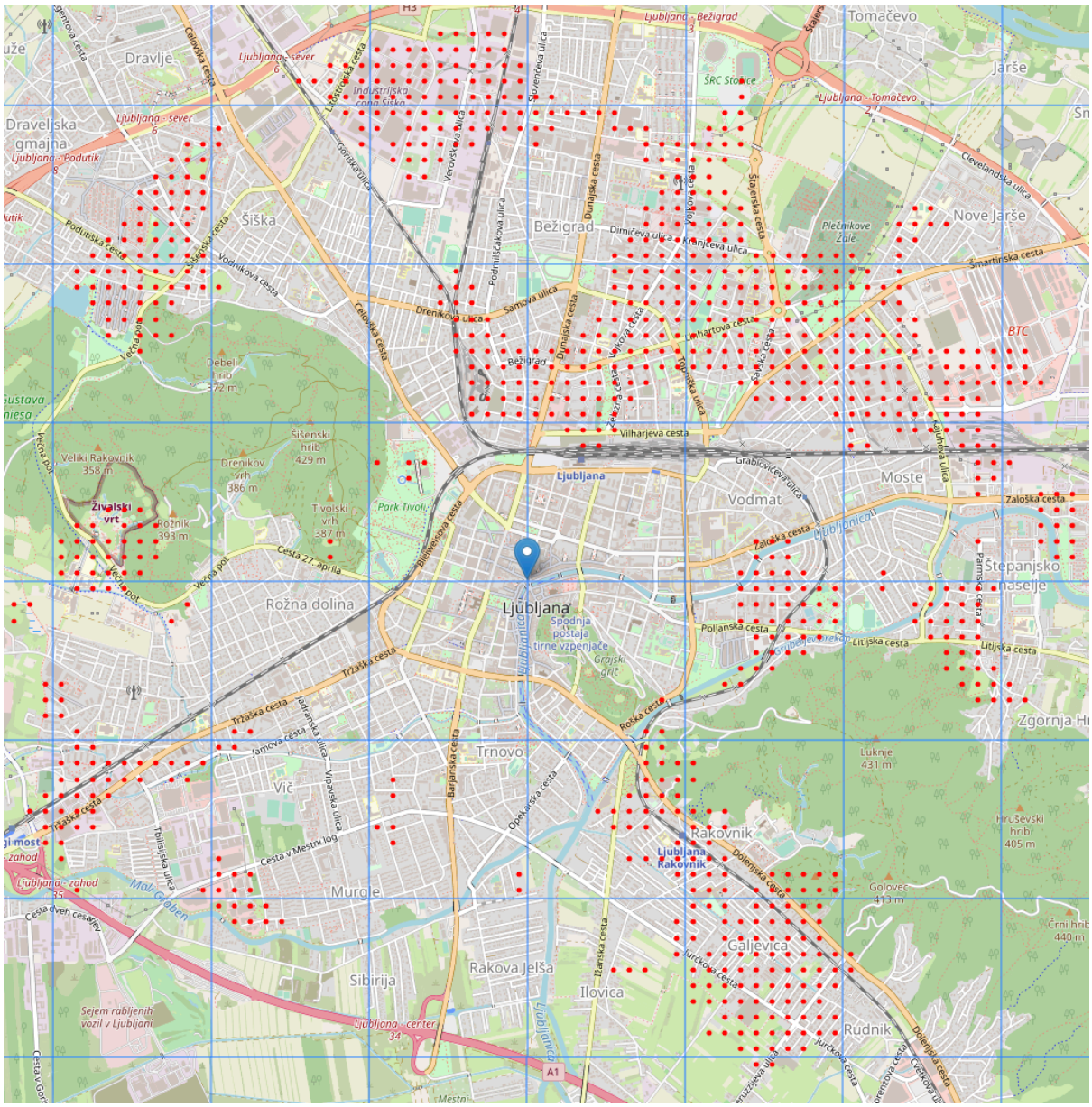

Figure 4. Accommodation location candidates that satisfy the requirements.

Finally, I used K-means clustering method to find groupings of location candidates. The center points of the clusters can be used as a starting point for further research of the neighborhood of the cluster centers. With some trial experiments I found that 25 clusters cover most of the location candidates (Figure 5).

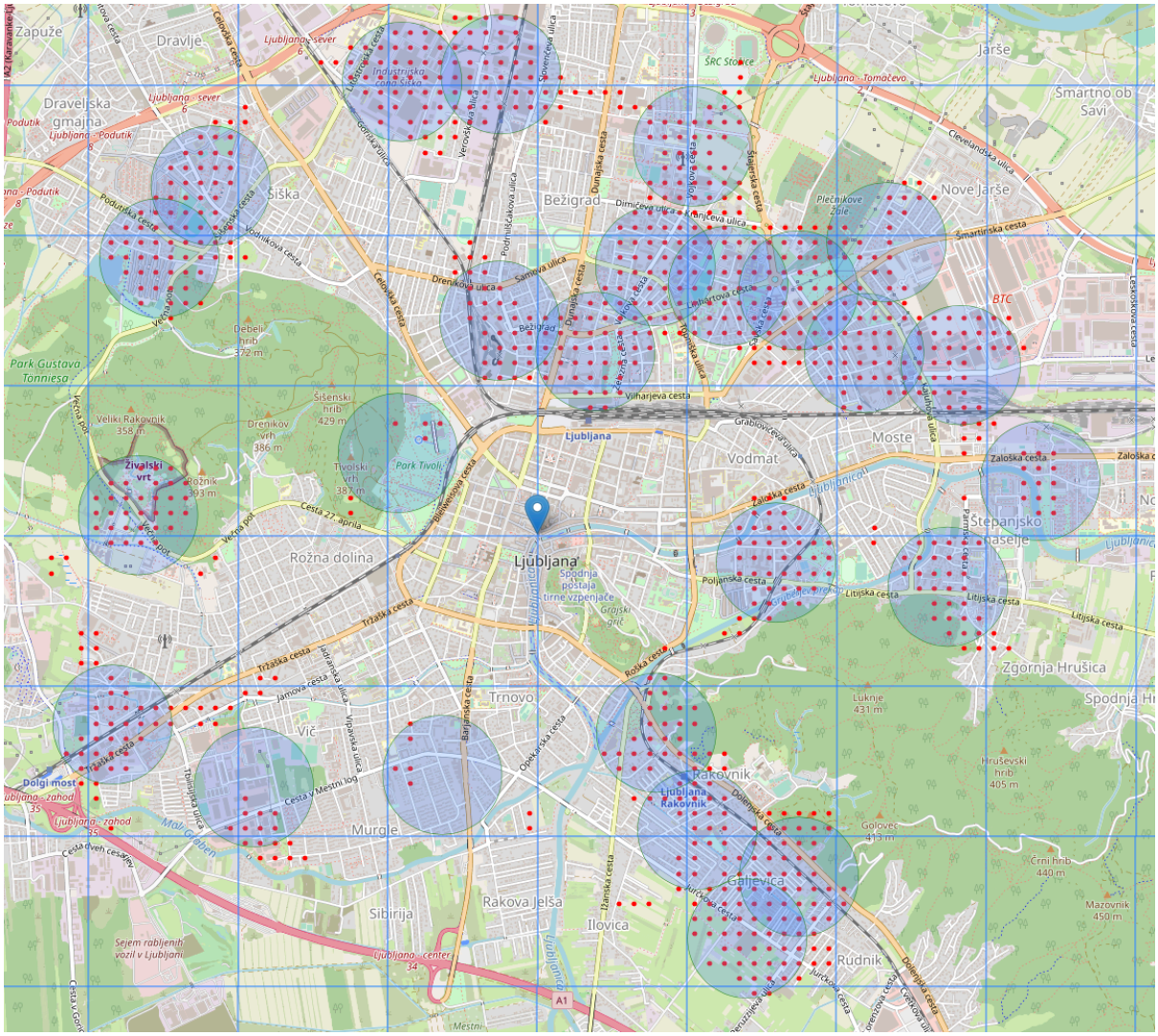

Figure 5. Clusters of accommodation location candidates. The circles represent radius of 400 meters.

One can observe that there are pockets of location candidates all around Ljubljana. The biggest areas seem to be to the north, north-east, east and south-east from the city center, and some start already at 1.3 km away from the center. The location candidates to the west and nearest to the city center are not appropriate since that area is the city park.

From the geographic coordinates of cluster centers, I collected the addresses and distance to the city center (Figure 6). 6 clusters are centered within the 2km radius from the city center which is a good result considering the existing accommodation density. The closest cluster is centered in the city's main park so likely not relevant in this case but leaving it in the results for completeness.

With that we have the final list.

	Latitude	Longitude	Address	Distance to center
22	46.056335	14.494089	Gozdne uganke, Jesenkova pot, Vrtača, Rožna do...	1078.602961
7	46.062412	14.511122	2, Detelova ulica, Zupančičeva Jama, Bežigrad	1283.518224
14	46.064266	14.502860	Mercator, 51, Parmova ulica, Zupančičeva Jama,...	1450.979315
24	46.039581	14.516396	Osnovna Šola Oskarja Kovačiča, 20, Dolenjska c...	1537.383623
0	46.049781	14.526728	4, Glonarjeva ulica, Kodeljevo, Vodmat	1606.994295
16	46.037056	14.497916	16, V Murglah, Krakovo, Murgle	1716.268588
2	46.067527	14.516306	Zweckform, Vojkova cesta, Zupančičeva Jama, Be...	1958.198158
21	46.066358	14.522592	Soča, Linhartova cesta, Savsko naselje, Vodmat	2095.436943
9	46.033767	14.519936	23, Galjevica, London, Rakovnik	2234.391755
12	46.066076	14.528667	Pokopališka ulica, Atrium, Vodmat	2389.336529
6	46.062330	14.534337	66, Kavčičeva ulica, Zelena jama, Moste	2500.001266
11	46.036327	14.481661	74, Koprska ulica, Vrtača, Vič	2527.223775
10	46.052616	14.471662	učilnica, uprava, 70, Večna pot, Vrtača, Rožna...	2668.278437
8	46.048333	14.541563	Petrol - Ljubljana - Litijska, 31, Litijska ce...	2766.090504
20	46.074694	14.519557	Vojkova cesta, Zupančičeva Jama, Bežigrad	2790.047943
4	46.030955	14.528695	London, Galjevica	2868.488496
23	46.027258	14.524245	94, Jurčkova cesta, London, Galjevica	3029.640059
19	46.068965	14.536266	8, Marje Boršnikove, Atrium, Nove Jarše	3042.829788
15	46.061605	14.542686	Harvey Norman, Španska ulica, BTC, Moste	3049.895633
5	46.079013	14.502968	106, Vodovodna cesta, Bratovševa ploščad, Beži...	3077.910813
3	46.040075	14.469444	Cesta Dolomitskega odreda, Kosovo Polje, Vrhovci	3103.948370
13	46.068007	14.473485	2, Karingerjeva ulica, Draveljska gmajna	3126.016372
17	46.078572	14.494457	eurolux, 44A, Litostrojska cesta, Industrijska...	3150.832928
1	46.072360	14.477925	4, Adamičeva ulica, Dravlje	3189.844760
18	46.054661	14.549527	11, Ulica Mire Miheličeve, Kodeljevo, Štepanjs...	3380.510979

Figure 6. Final 25 geographic coordinates and addresses, sorted by distance to the city's center.

Results

The analysis shows there is an existing high accommodation venue density within the 2 km radius of the Ljubljana's center, with highest concentration within the 600 meters of the center. This is expected as many of Ljubljana's landmarks are concentrated there as well. The same holds for existing food venues and public transportation grid.

Most of the existing accommodation venues (129 out of 189) are likely higher capacity venues and more suitable for groups. There are certainly some boutique hotels with a low capacity, however with available data we cannot differentiate.

By placing a fine grid (with areas of 100 m x 100 m) over the Ljubljana map and by filtering out the data to fit the set requirements - no other accommodation venue within the walking distance, public transport access and a number of food venues also within the walking distance (where walking distance is 400 m) - I was able to find pockets of location candidates all around the city of Ljubljana. The number is just over 900 locations.
Since it does not make sense to deal with each individual candidate location, with the help of K-means clustering I exposed 25 zones of interest with an attached address.

Most zones of interest reside to the north, north-east, east and south-east. Some of those are the very close to the city center (already about 1.3 km from the city center).

Discussion

In the project I was able to provide a set of 25 addresses that would be of interest to the stakeholders. The addresses can be used a starting point for further neighborhood research.

Despite the expectations that the location candidates for new accommodation object will not appear within the 2 km radius from the city center, I was able to find 6 pockets of candidate locations that are within the 2 km radius. There are 18 that spread out to the defined limit of 3.5 km. This is a good result.

Further out there are even more options since the accommodation density is less, but the transportation grid and food venues are still plenty and cover the whole city.

By increasing the grid resolution, we would get more precise results, with the expense of the higher computational and memory demands.

The quality of the results depends on the quality of the data. Foursquare data can have gaps in some categories especially for cities that are not big or known. So additional data sources should be used in parallel for any further analysis.

Conclusion

The goal of this project was to find a set of location candidates for new accommodation object in the city of Ljubljana. Preferred locations are those that have low accommodation venue density, are near to the existing public transportation grid and have accessible food venues. This information can aid stakeholders/investors with searching for a good location for new accommodation venues.

The acquired results show that there are options for a good accommodation location based on set prerequisites all around Ljubljana, some starting already at 1.3 km from the center, especially to the north and north-east.

There are certainly other location characteristics that were not taken into account in this project but can impact the decision for a location, for example the tourist numbers, their behaviors and expectations, real estate availability and prices, existing accommodation capacities, proximity to major roads and with this the noise and pollution levels, proximity to entertainment venues and city's landmarks, specificities of the district etc.