

IBM Data Science Capstone Project

Operating a rickshaw in Stockholm

Contents

Introduction	2
Background	2
Problem	2
Data	3
Methodology	4
Results	5
Discussion	6
Conclusion	6

Introduction

Background

In Stockholm, Sweden's capital, you find cobblestone streets and ochre-colored buildings. The greater area contains about 50 bridges stretching over about 14 islands which are connected also by ferries and sightseeing boats shuttle passengers. To this mode of transport, it shall be evaluated to add transport by rickshaw.

Problem

More and more traffic congestions exist in Stockholm due to the number of cars using the limited resource streets.



Figure 1 Potential rickshaw model for Stockholm

A potential solution for that is me operating a rickshaw to transport people in between Stockholm's hotspots. This project shall evaluate data available from foursquare.com to figure out where to best position my rickshaw.

Data

For this project the foursquare API will be used to fetch, analyze and use the available data.

reasons.count	reasons.items	referrerId	venue.categories	venue.id	venue.location.address	venue.location.cc	venue.location.city	venue.location.country	venue.location.crossStreet	venue.location.labelId	venue.location.lat	venue.location.lng	venue
0	[[{"summary": "This spot is popular", "type": "..."}]]	e-0-4b0f8b9964a520fe622e3-0	4b0f8b9964a520fe622e3-0	4b0f8b9964a520fe622e3-0	Vasterloggatan 40	SE	Stockholm	Sverige	Gamla Stan	[[{"label": "display", "lat": 59.324047, "lng": 18.070682}]]	59.324047	18.070682	
1	[[{"summary": "This spot is popular", "type": "..."}]]	e-0-514c068e4b07a3bda27ca3-1	4b0f8b9964a520fe622e3-0	514c068e4b07a3bda27ca3-1	Lilla Nygatan 16	SE	Stockholm	Sverige	NaH	[[{"label": "display", "lat": 59.323342, "lng": 18.069431}]]	59.323342	18.069431	
2	[[{"summary": "This spot is popular", "type": "..."}]]	e-0-549447b498ed7c31980c2-2	4b0f8b9964a520fe622e3-0	549447b498ed7c31980c2-2	Stora Nygatan 20	SE	Stockholm	Sverige	Yxmedsgränd	[[{"label": "display", "lat": 59.324655, "lng": 18.065513}]]	59.324655	18.065513	
3	[[{"summary": "This spot is popular", "type": "..."}]]	e-0-534d0c3449ef702b62ac29c-3	4b0f8b9964a520fe622e3-0	534d0c3449ef702b62ac29c-3	Osterloggatan 1	SE	Stockholm	Sverige	NaH	[[{"label": "display", "lat": 59.32594, "lng": 18.073710}]]	59.32594	18.073710	

Figure 2: Foursquare data

We try to utilize the following fields.

Field	Explanation	Usage
id	A unique string identifier for this location (foursquare calls venue)	To be evaluated
name	Name of the location	To be evaluated
location	An object containing coordinates of the location	To be evaluated
categories	An array containing descriptive information about the location	To be evaluated
reason	An array containing a summary of the place	To be evaluated

To evaluate and identify potential rickshaw locations, we re-classify and plot the foursquare data onto maps views to visualize the results.

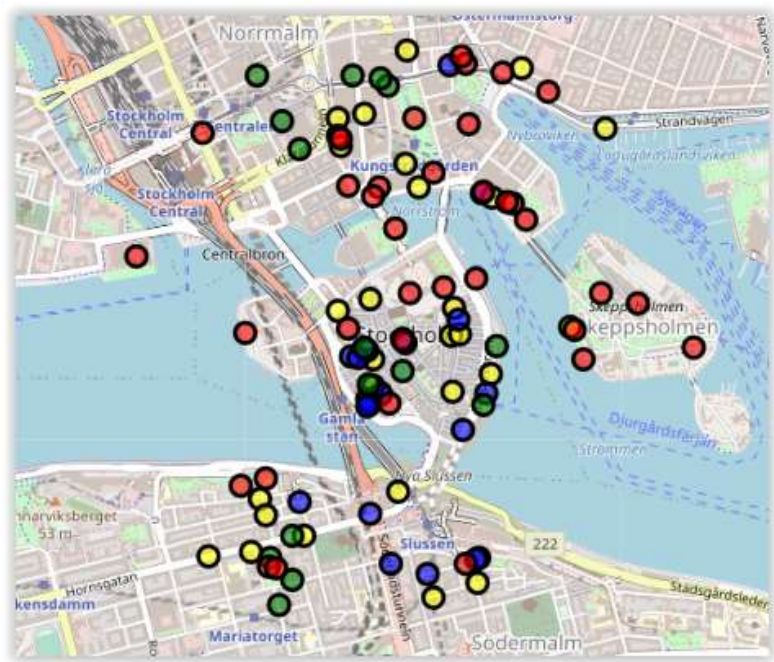


Figure 3: Preview of location analysis

Methodology

To identify a good place for positioning a rickshaw, the surroundings in Stockholm shall be analyzed. First the data set was more deeply analyzed to understand if we could figure out popular places using the summary information provided.

	name	items	lat	lng
0	Science Fiction Bokhandeln	This spot is popular	59.324047	18.070682
1	Corner Club	This spot is popular	59.323342	18.069431
2	Barrels Burgers & Beer	This spot is popular	59.324685	18.068513
3	Tradition	This spot is popular	59.325940	18.073710
4	Stortorget	This spot is popular	59.324973	18.070727

Figure 4: Summary fields

As this did not turn out the expected useful result to identify specific popular locations, the categories information was further explored.

	id	name	categories	lat	lng
0	4b06fe89f964a520fef422e3	Science Fiction Bokhandeln	Bookstore	59.324047	18.070682
1	514c0868e4b07a3d6af2fca3	Corner Club	Cocktail Bar	59.323342	18.069431
2	549447fb498ed7d33f190dc2	Barrels Burgers & Beer	Burger Joint	59.324685	18.068513
3	534d0a34498ef782062ac89c	Tradition	Scandinavian Restaurant	59.325940	18.073710
4	4b0e3607f964a520fa5523e3	Stortorget	Plaza	59.324973	18.070727

```
print('There are {} uniques categories.'.format(len(normalized_neighborhood['categories'].unique())))
```

There are 49 uniques categories.

Figure 5: Exploring category information

This kind of categorization does not lead to allow a proper location based analysis as the categories are to precise, therefore the data set was harmonized into 4 categories.

```
9]: categories
Bar      18
Recreation  36
Restaurant  29
Shop      15
Name: name, dtype: int64
```

```
print('There are {} uniques categories.'.format(len(normalized_neighborhood['categories'].unique())))
```

There are 4 uniques categories.

Figure 6: Harmonizing categories

Results

Based on the re-categorization the locations, the average latitude/longitude method is used to finding a midpoint in between the categories.

This method finds a simple average latitude and longitude for the locations in 'Your Places'. This is equivalent to finding a midpoint on a flat rectangular projection map. When the distance between locations is less than 250 miles (400 km), this method gives a close approximation to the true geographic midpoint in Method A. ⁱ

The color coding used is yellow for restaurants, blue for bars, green for shops and red for recreational places.

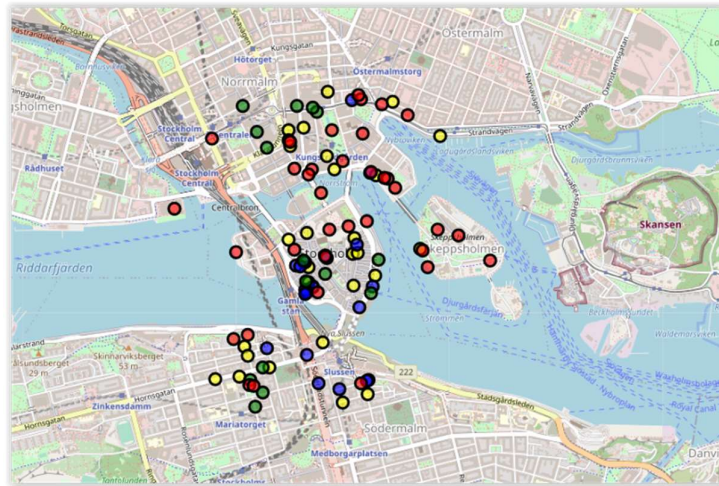


Figure 7: Starting point per category

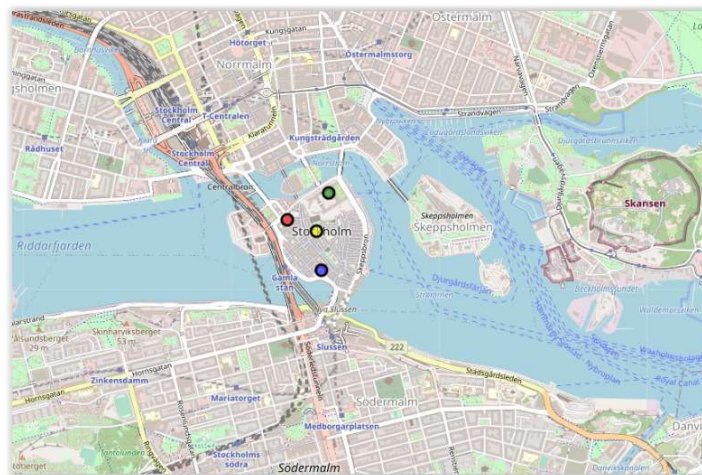


Figure 8: Midpoints per category

Discussion

Based on this preliminary analysis we can confirm that the best place to position a rickshaw in Stockholm is Gamla Stan. It is the centre of most restaurant, bars, shops and recreational places in Stockholm. This is purely based on location data of foursquare venue information which was categorized.

Certain locations will have different opening hours / hours in which these are frequented by people which this analysis did not include. This could be the next step of the project, including i.e. opening hours assuming that also people will frequent these spots.

For example one could argue that shops will have more population frequenting these places between 10:00hrs to 20:00hrs where as bars and restaurants a frequented between the hours of 11:00hrs and 01:00hrs.

Conclusion

As the centres of the locations are quite close to each other, for regular operating hours of my rickshaw, Gamla Stan is the place to be.

ⁱ <http://www.geomidpoint.com/calculation.html>