

# Finding the suitable place for a new hotel

## Introduction

The two key factors for the success of any business are targeted audience and location. The first thing is for sure the people whom a business owner plans to target as potential customers. And the second thing that plays a vital role is the location of the business setup. If the location is suitable and convenient, it attracts more customers.

## Problem statement

A successful business owner plans to open a branch of his hotel in Tallinn, the capital of Estonia. He plans to target travelers as well as local community. For this purpose, he aims to provide accommodation facility and also multicultural food options at single place.

The aim of this project is to find the optimal location for such a business investment in the city. This analysis can be helpful for any business investment.

## Data

For our project, we need the following data:

- List of sub-districts in Tallinn
- Geo-coordinates of each sub-district in Tallinn
- And top venues in each district

The list of sub-districts is obtained from the following Wiki-pedia link:

[https://en.wikipedia.org/wiki/Category:Subdistricts\\_of\\_Tallinn](https://en.wikipedia.org/wiki/Category:Subdistricts_of_Tallinn)

Once we get the list of sub-districts, we can get their Geo-coordinates and also top venues in them.

## Methodology

The process includes two main steps:

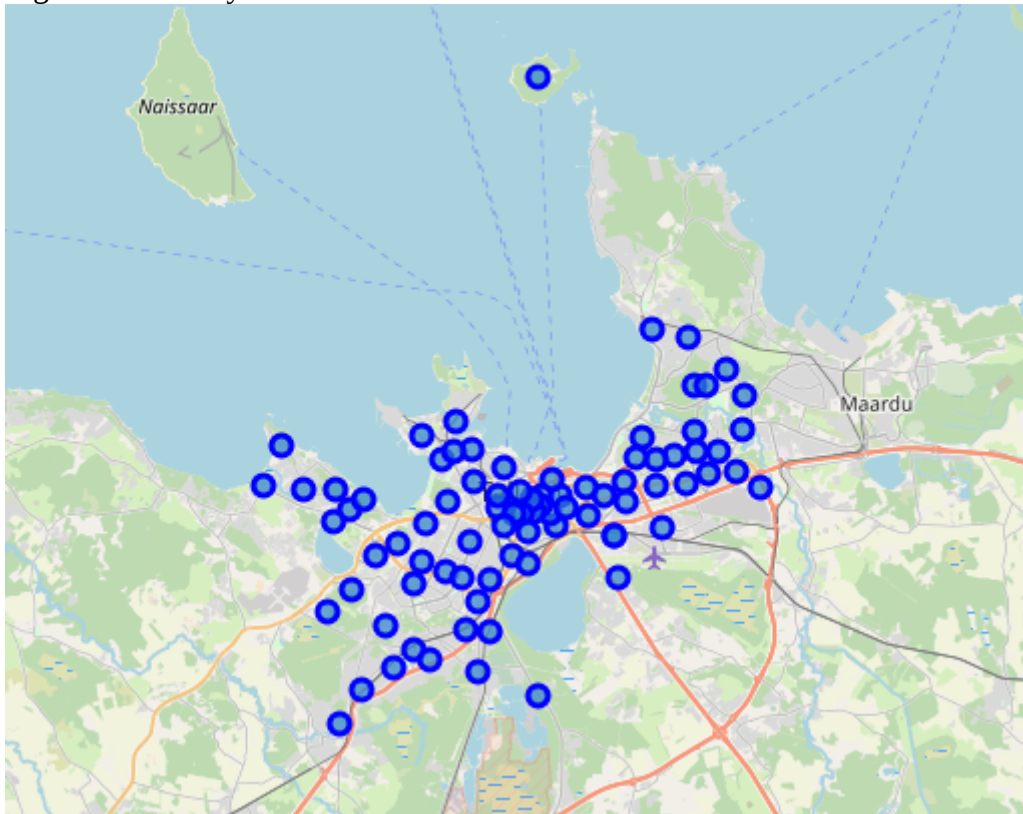
- Data analysis
- Clustering

## Data Analysis

The first thing in data analysis is to get the list of sub-districts which we will get from the source page. Once we have the list of sub-districts, the next step includes finding the Geo-coordinates which we get through Geocode Python library.

	Sub-district	Latitude	Longitude
0	Aegna	59.581319	24.757531
1	Astangu	59.402937	24.630661
2	Haabersti	59.431328	24.629042
3	Hiiu	59.382345	24.672648
4	Iru	59.459188	24.898411
5	Järve	59.398956	24.716329
6	Juhkentali	59.425699	24.770611
7	Kadaka	59.412797	24.677430
8	Kadriorg	59.438466	24.791025
9	Kakumäe	59.453371	24.581699

Tallinn has a total of 83 sub-districts. As we have the Geo-coordinates, we can now visualize them on map using Folium library.



The next step is to get the venues in the radius of each Geo-coordinate. We use Foursquare API for this purpose to get the venues of each sub-district in the radius of 1000 meters. Foursquare API responds and gives us the requested data in a json format file.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Aegna	59.581319	24.757531	Kaluriküla Saun	59.581079	24.759631	Campground
1	Aegna	59.581319	24.757531	House of the Blackheads	59.585157	24.764115	Music Venue
2	Aegna	59.581319	24.757531	Põhjarand	59.585501	24.750149	Beach
3	Aegna	59.581319	24.757531	Aegna rändrahnud	59.581938	24.773533	Scenic Lookout
4	Aegna	59.581319	24.757531	Aegna grill	59.572338	24.756646	BBQ Joint

We can also sort the data to see how many venues each sub-district has, as shown below.

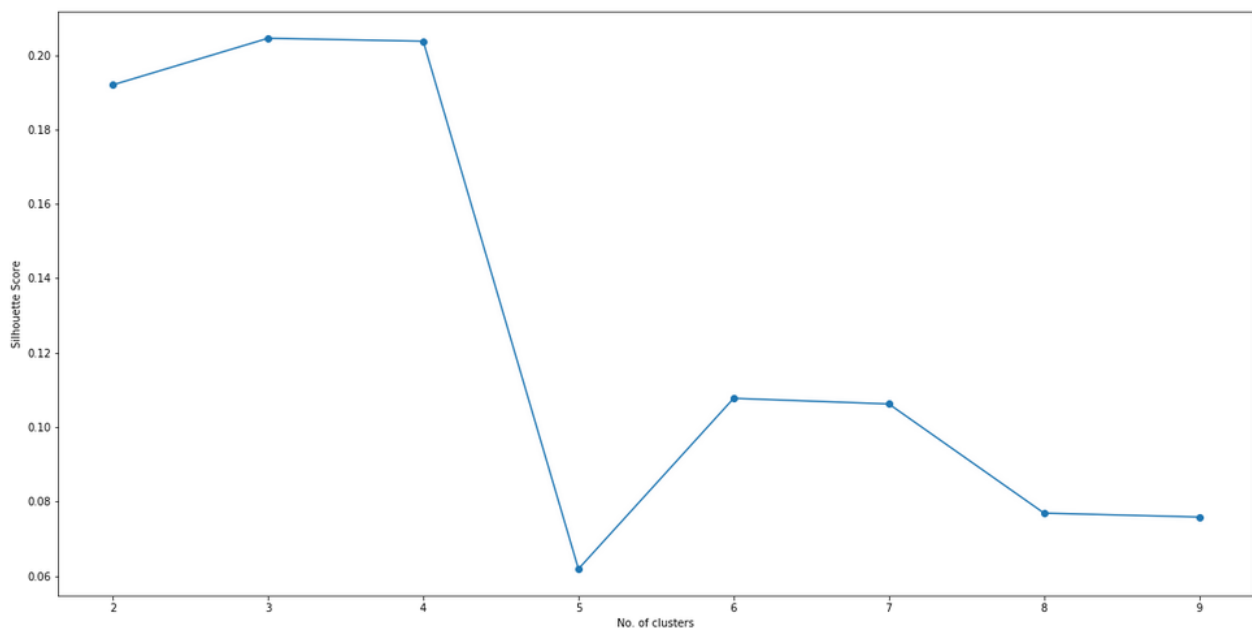
	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Aegna	5	5	5	5	5	5
Astangu	12	12	12	12	12	12
Haabersti	18	18	18	18	18	18
Hiiu	44	44	44	44	44	44
Iru	6	6	6	6	6	6
Juhkentali	80	80	80	80	80	80
Järve	50	50	50	50	50	50
Kadaka	44	44	44	44	44	44
Kadrioriorg	60	60	60	60	60	60

To further analysis, we use one-hot encoding to get the venue categories in each sub-district. This helps us to find the top venue categories in each sub-district. We have got top 10 venue categories for each sub-district in our project.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Aegna	Campground	Beach	Scenic Lookout	Music Venue	BBQ Joint	Fast Food Restaurant	Farmers Market	Fair	Exhibit	Flea Market
1	Astangu	Convenience Store	Trail	Middle Eastern Restaurant	Spa	Motorcycle Shop	Bus Station	Bus Stop	Eastern European Restaurant	Grocery Store	Go Kart Track
2	Haabersti	Bus Station	Trail	Bus Stop	Convenience Store	Eastern European Restaurant	Beach	Supermarket	Furniture / Home Store	Museum	Gym / Fitness Center
3	Hiiu	Café	Bus Station	Bus Stop	Pub	Ski Area	Music Venue	Park	Train Station	Restaurant	Modern European Restaurant
4	Iru	Bus Station	Theme Park	Shopping Mall	Cosmetics Shop	Grocery Store	Driving School	Eastern European Restaurant	Electronics Store	English Restaurant	Escape Room
5	Juhkentali	Sushi Restaurant	Restaurant	Hotel	Lounge	Café	Deli / Bodega	Thai Restaurant	Arts & Crafts Store	Grocery Store	Casino
6	Järve	Furniture / Home Store	Café	Eastern European Restaurant	Supermarket	Bus Station	Sporting Goods Shop	Park	Tennis Court	Bus Stop	Restaurant
7	Kadaka	Supermarket	Café	Shopping Mall	Movie Theater	Pizza Place	Gym	Electronics Store	Sporting Goods Shop	Fast Food Restaurant	Jewelry Store

## Clustering

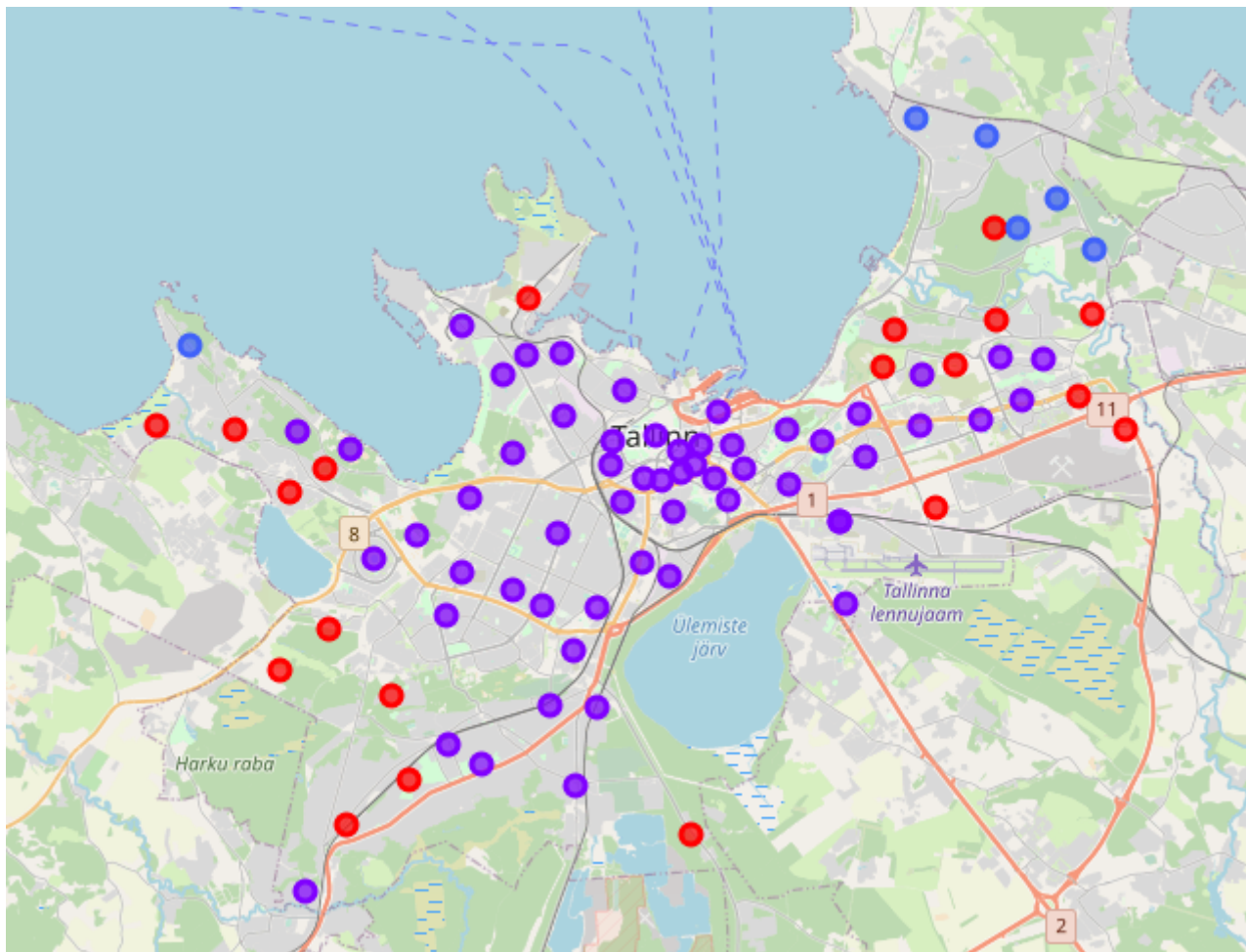
After data analysis, we need clustering to find out the most suitable cluster for this investment. K-means clustering which is an unsupervised machine learning technique is used for this purpose. A question arises how many clusters do we need to get best results. Silhouette score helps us find the optimal number of clusters.



From the graph, we can figure out that the optimal number of clusters is 3. Using 3 as the number of clusters in K-means clustering algorithm parameters, we get the following results.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Aegna	BBQ Joint	Scenic Lookout	Music Venue	Beach	Escape Room	Event Space	Exhibit	Fabric Shop	Fair	Yoga Studio
1	Astangu	Bus Stop	Hookah Bar	Trail	Pedestrian Plaza	Paintball Field	Eastern European Restaurant	Bus Station	Shopping Mall	Motorcycle Shop	Spa
2	Haabersti	Bus Station	Bus Stop	Hardware Store	Eastern European Restaurant	Furniture / Home Store	Beach	Supermarket	Trail	Gym / Fitness Center	Garden Center
3	Hiiu	Bus Stop	Park	Bus Station	Café	Ski Area	Sporting Goods Shop	Pub	Train Station	Restaurant	Farmers Market
4	Iru	Grocery Store	Theme Park	Cosmetics Shop	Shopping Mall	Bus Station	Fair	Event Space	Exhibit	Fabric Shop	Farmers Market
5	Juhkentali	Sushi Restaurant	Restaurant	Hotel	Lounge	Café	Deli / Bodega	Tram Station	Thai Restaurant	Pizza Place	Gym / Fitness Center
6	Järve	Bus Stop	Café	Eastern European Restaurant	Furniture / Home Store	Tennis Court	Trail	Supermarket	Coffee Shop	Sporting Goods Shop	Sushi Restaurant
7	Kadaka	Supermarket	Shopping Mall	Café	Movie Theater	Sporting Goods Shop	Gym	Market	Pizza Place	Pub	Burger Joint

We can also visualize clusters on the map.



## Results

As we have 3 clusters, let's explore each cluster individually.

### Cluster 0

Cluster 0, as shown in red bubbles on map, consists of the outskirts of the city which is definitely not the suitable option. This cluster gives the following results.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	Haabersti	Bus Station	Bus Stop	Hardware Store	Eastern European Restaurant	Furniture / Home Store	Beach	Supermarket	Trail	Gym / Fitness Center	Garden Center
4	Iru	Grocery Store	Theme Park	Cosmetics Shop	Shopping Mall	Bus Station	Fair	Event Space	Exhibit	Fabric Shop	Farmers Market
9	Kakumäe	Dessert Shop	Scenic Lookout	Athletics & Sports	Bus Station	Beach	Farmers Market	Exhibit	Fabric Shop	Fair	Yoga Studio
17	Kivimäe	Bus Station	Pizza Place	Pharmacy	Grocery Store	Toy / Game Store	Park	Music Venue	Trail	Office	Bus Stop
21	Kose	Market	Bus Station	Moving Target	Bus Stop	Trail	Restaurant	Beach	Café	Volleyball Court	Park
34	Mäeküla	Flower Shop	Trail	Yoga Studio	Fair	English Restaurant	Escape Room	Event Space	Exhibit	Fabric Shop	Farmers Market
35	Mähe	Scenic Lookout	Food Service	Moving Target	Outdoors & Recreation	Bus Station	Yoga Studio	Fabric Shop	Escape Room	Event Space	Exhibit
45	Pääsküla	Bus Station	Trail	Train Station	Skate Park	Food & Drink Shop	Cable Car	Grocery Store	Restaurant	Garden Center	Pool
47	Paevälja	Museum	Bus Station	Moving Target	Playground	Soccer Field	Brewery	Shopping Mall	Supermarket	Resort	Fabric Shop
48	Paljassaare	Bus Station	Boat or Ferry	Tram Station	Electronics Store	Bus Stop	Supermarket	Bar	Gun Range	Convenience Store	Health Food Store
51	Pikaliiva	Bus Station	Pharmacy	Trail	Pet Store	Convenience Store	Playground	Motorcycle Shop	Furniture / Home Store	Bus Stop	Supermarket
56	Raudalu	Pizza Place	Gas Station	Forest	Supermarket	Bus Station	Flower Shop	Pet Store	Pharmacy	Cosmetics Shop	Fish & Chips Shop
60	Seli	Bus Station	Bus Stop	Grocery Store	Comfort Food Restaurant	Gym	Flea Market	Fast Food Restaurant	Moving Target	Shopping Mall	Malay Restaurant
68	Tiskre	Italian Restaurant	Athletics & Sports	Moving Target	Steakhouse	Trail	Bus Stop	Bus Station	Farmers Market	Exhibit	Fabric Shop
78	Vana-Mustamäe	Bus Station	Trail	Diner	Athletics & Sports	Hostel	Breakfast Spot	Ski Trail	Ski Area	Moving Target	Cafeteria
80	Vão	Moving Target	Bus Station	Malay Restaurant	Supermarket	Yoga Studio	Farmers Market	Event Space	Exhibit	Fabric Shop	Fair
83	Vismeistri	Bus Station	Hot Spring	Restaurant	Steakhouse	Skating Rink	Escape Room	Event Space	Exhibit	Fabric Shop	Fair

### Cluster 1

Cluster 1, as shown in almost violet on map, is the biggest cluster and covers the suitable locations. If we look deep down in the map, the center of this cluster is denser and is the most suitable place for this investment in this cluster. The cluster gives the following results.



	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Aegna	BBQ Joint	Scenic Lookout	Music Venue	Beach	Escape Room	Event Space	Exhibit	Fabric Shop	Fair	Yoga Studio
1	Astangu	Bus Stop	Hookah Bar	Trail	Pedestrian Plaza	Paintball Field	Eastern European Restaurant	Bus Station	Shopping Mall	Motorcycle Shop	Spa
3	Hiiu	Bus Stop	Park	Bus Station	Café	Ski Area	Sporting Goods Shop	Pub	Train Station	Restaurant	Farmers Market
5	Järve	Bus Stop	Café	Eastern European Restaurant	Furniture / Home Store	Tennis Court	Trail	Supermarket	Coffee Shop	Sporting Goods Shop	Sushi Restaurant
6	Juhkentali	Sushi Restaurant	Restaurant	Hotel	Lounge	Café	Deli / Bodega	Tram Station	Thai Restaurant	Pizza Place	Gym / Fitness Center
7	Kadaka	Supermarket	Shopping Mall	Café	Movie Theater	Sporting Goods Shop	Gym	Market	Pizza Place	Pub	Burger Joint
8	Kadriorg	Café	Restaurant	Park	Beach	Spa	Museum	Cosmetics Shop	Coffee Shop	Grocery Store	Art Museum
10	Kalamaja	Restaurant	Burger Joint	Wine Bar	Bakery	Café	Park	Beer Bar	Bar	Theater	Scenic Lookout
11	Karjamaa	Bus Stop	Bakery	Harbor / Marina	Supermarket	Bar	Bus Station	Convenience Store	Lounge	Restaurant	Shopping Mall
12	Kassisaba	Café	Restaurant	Hotel	Scenic Lookout	Park	Eastern European Restaurant	Coffee Shop	Wine Bar	Pub	Gym / Fitness Center
13	Katleri	Furniture / Home Store	Athletics & Sports	Market	Toy / Game Store	Beach	Middle Eastern Restaurant	Supermarket	Bus Station	Bus Stop	Grocery Store
14	Keldrimäe	Restaurant	Café	Sushi Restaurant	Hotel	Cosmetics Shop	Gym / Fitness	Concert Hall	Coffee Shop	Cocktail Bar	Deli / Bodega

## Cluster 2

Cluster 2, shown in sky-blue on map, also consists of the outskirts of the city and is therefore not a suitable option. The cluster gives the following results.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
26	Laiaküla	Restaurant	Scenic Lookout	Trail	Smoke Shop	Bus Stop	Bus Station	Yoga Studio	English Restaurant	Escape Room	Event Space
27	Lepiku	Restaurant	Scenic Lookout	Trail	Smoke Shop	Bus Stop	Yoga Studio	Electronics Store	English Restaurant	Escape Room	Event Space
52	Pirita	Scenic Lookout	Restaurant	Garden	Bus Stop	Trail	River	Botanical Garden	Smoke Shop	Farmers Market	Fair

## Conclusion

Looking into all clusters, we can figure out that cluster 1 is the most suitable option. In cluster 1, the center is highly dense and the best option to invest for this business plan. Based on this, we can suggest business owner to open his hotel in the area that is the center of cluster 1.

## Source code

The notebook that explains the step by step process of this project can be found here:

[https://github.com/zafarkheshki/Data-Science-Capstone/blob/main/Capstone\\_project\\_tallinn.ipynb](https://github.com/zafarkheshki/Data-Science-Capstone/blob/main/Capstone_project_tallinn.ipynb)