



Crime and Neighborhoods of Tallinn, Estonia

BY ANDRI HÖBEMÄGI

Objective

- ▶ During the last 15 or so years, Tallinn has grown into being one of the best cities for start-up businesses, which sets new standards to the safety and security of the city to attract foreign talent.
- ▶ Although the number of crimes in Tallinn has declined considerably over the years, it is still worth to investigate, how the different districts compare to each other in this respect.
- ▶ The objective is to seek to answer the following questions:
 1. What districts have the highest crime rate?
 2. Which neighborhood data correlates to crime level?
 3. Using Foursquare data, what venues are most common in the selected neighborhoods?

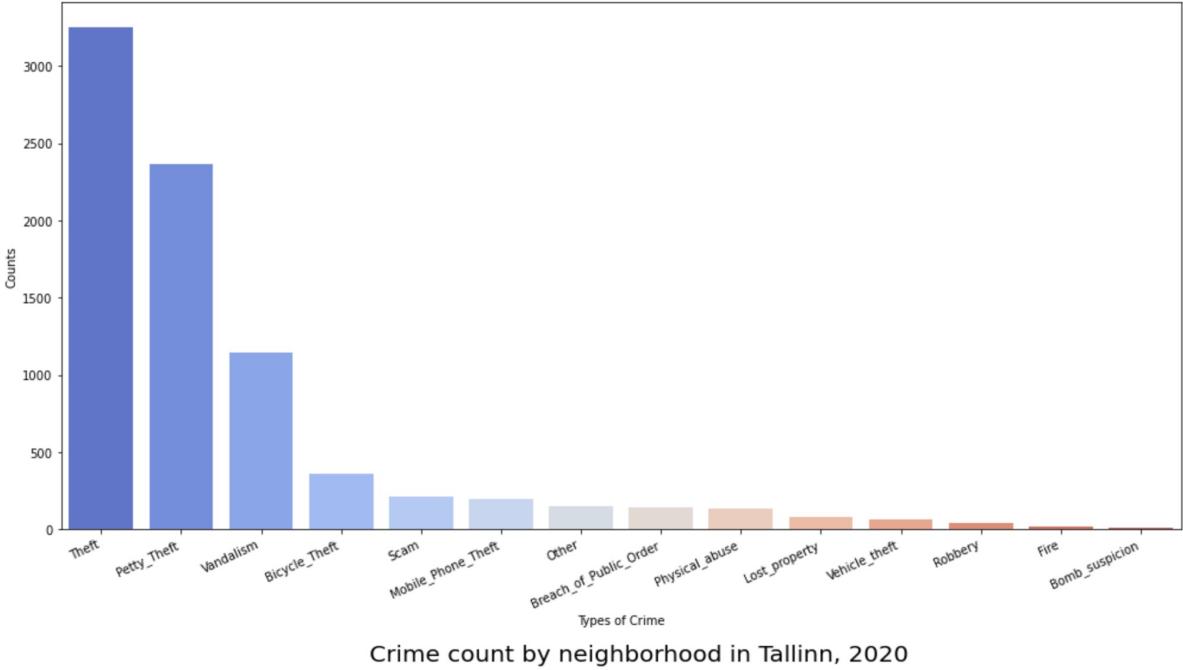
Data

- ▶ Necessary data was obtained from the following sources:
 1. Estonian Open Data (crime data): <https://avaandmed.eesti.ee>
 2. Tallinn in numbers (population data): <https://www.tallinn.ee/eng/g2677s126569>
 3. Tallinn Geospatial Data (neighborhood geometry): <https://www.tallinn.ee/est/geoportaal/Andmed>
 4. Statistics Estonia (income statistics): <https://www.stat.ee/en>
 5. Foursquare Developers Access to venue data: <https://foursquare.com/>

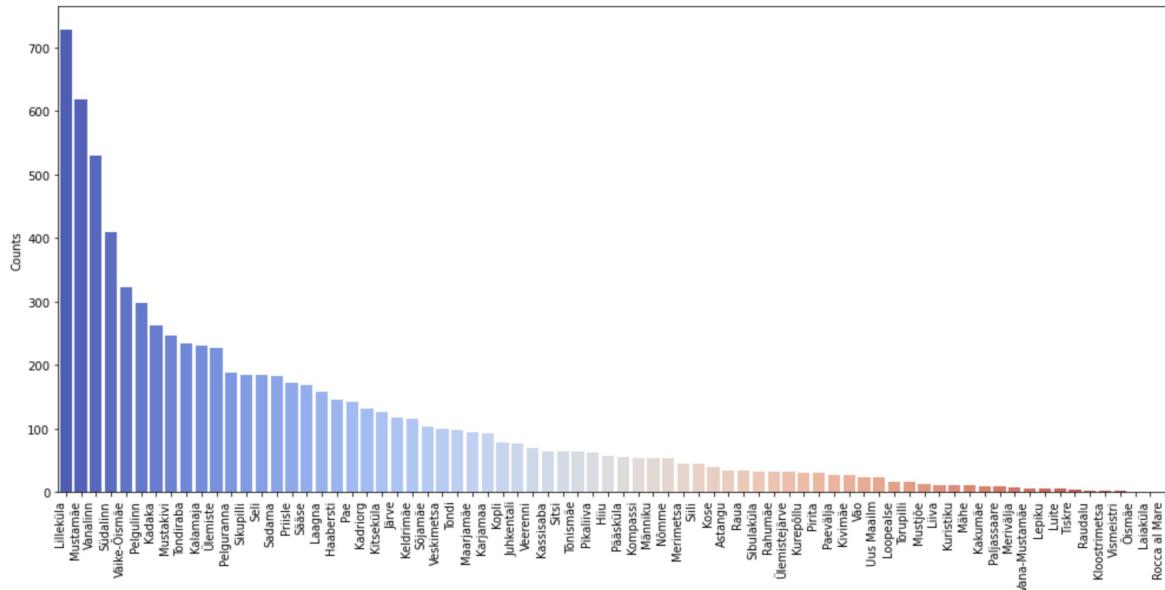
Project approach

- ▶ Prepare the data for analysis and visualization to make the correct selection of neighborhoods of interest
- ▶ After getting the list of the desired neighborhoods sticking to those neighborhoods and further explore them leveraging foursquare API.
- ▶ Then running k-means algorithm and clustering the different neighborhoods based on their similarity index and dissimilarity index.
- ▶ Finally analyze the neighborhoods based on clustering results.

Types of Crime in Tallinn, 2020

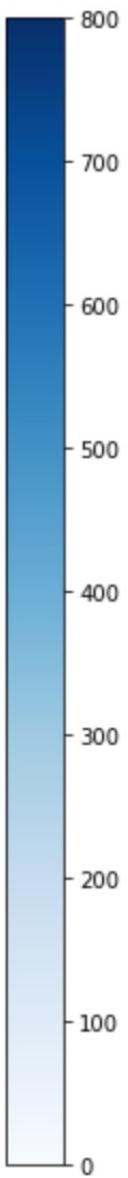
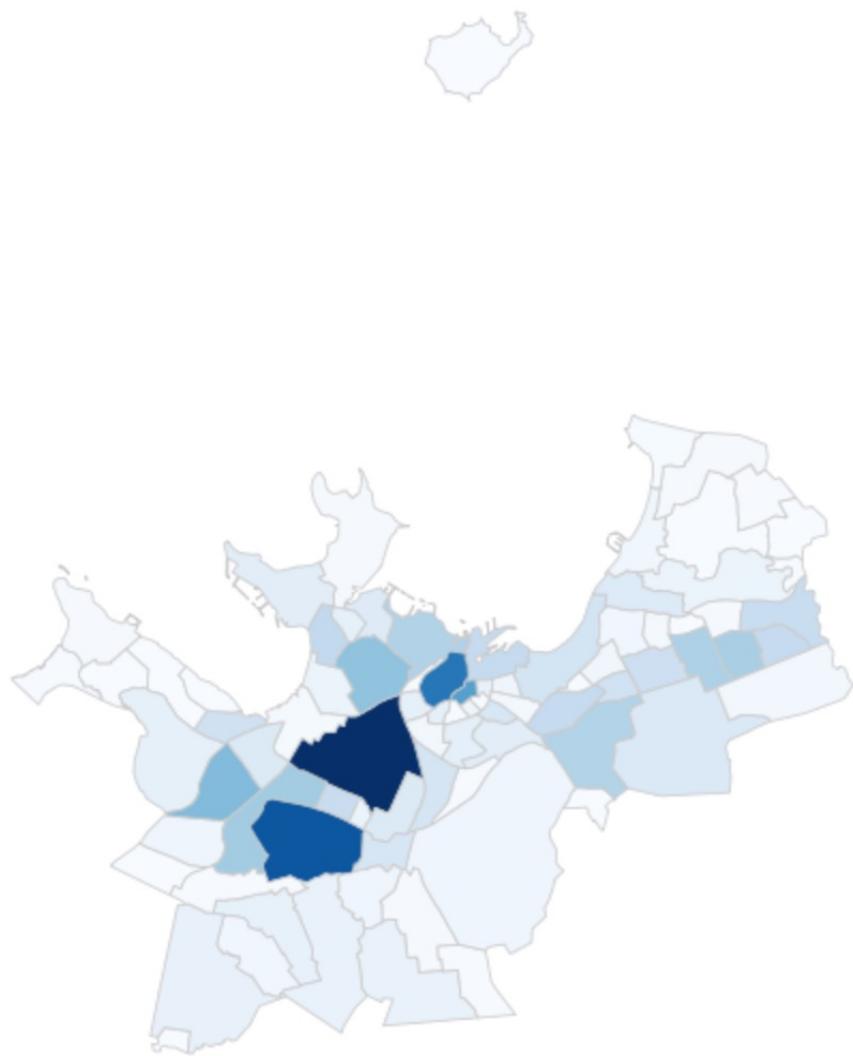


Crime count by neighborhood in Tallinn, 2020



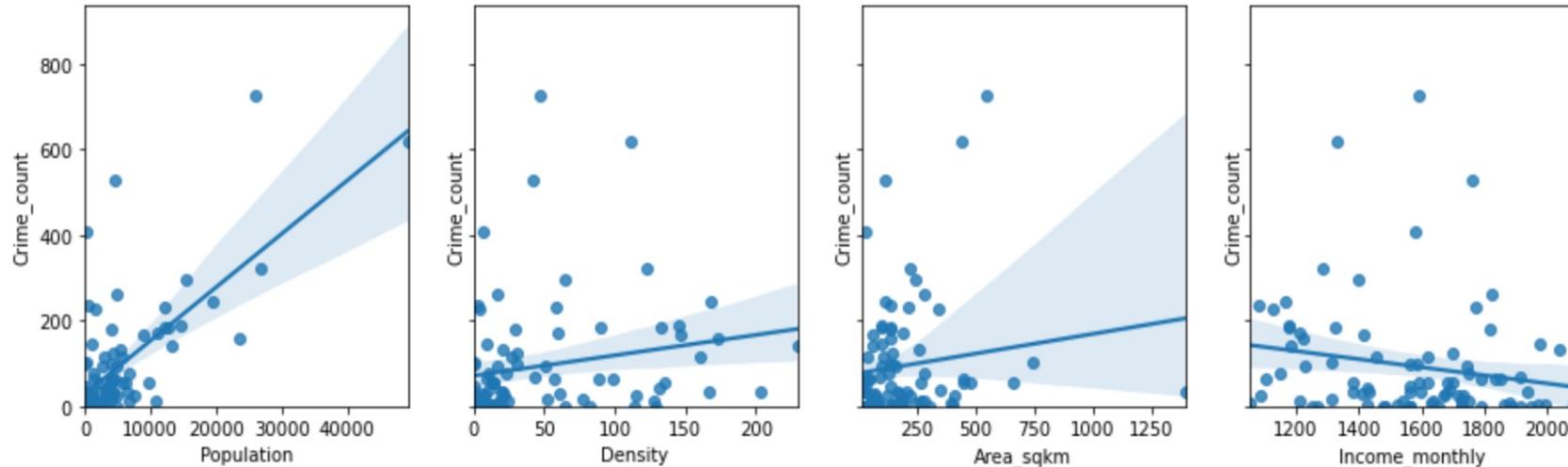
Visualization of
crimes in the city
based on type
and
neighborhood

Number of crimes in neighborhoods of Tallinn



Choropleth
map of crime
count per
neighborhood

Correlations of crime count with general data



Pearson correlation between population and crime count is 0.6889690288839252 and P-value P = 4.3390895837573414e-13
Pearson correlation between density and crime count is 0.20008458035572396 and P-value P = 0.06802697271631018
Pearson correlation between area and crime count is 0.13534320319764231 and P-value P = 0.21963040166596942
Pearson correlation between income and crime count is -0.18390607700334088 and P-value P = 0.09401528006109083

Top ten neighborhoods with highest crime count...

	Neighborhood	District	Latitude	Longitude	Population	Crime_count
29	Lilleküla	Kristiine	59.420076	24.708247	25940	728.0
41	Mustamäe	Mustamäe	59.400405	24.681088	49345	619.0
79	Vanalinn	Kesklinn	59.437876	24.743519	4658	530.0
66	Südalinn	Kesklinn	59.434423	24.752541	168	410.0
77	Väike-Õismäe	Haabersti	59.412688	24.645559	26769	322.0
49	Pelgulinna	Põhja-Tallinn	59.440319	24.711536	15336	297.0
7	Kadaka	Mustamäe	59.407763	24.656631	4781	263.0
40	Mustakivi	Lasnamäe	59.443644	24.874310	19434	246.0
70	Tondiraba	Lasnamäe	59.442285	24.858399	354	235.0
10	Kalamaja	Põhja-Tallinn	59.447065	24.733997	12179	230.0

... and
their
location
on map

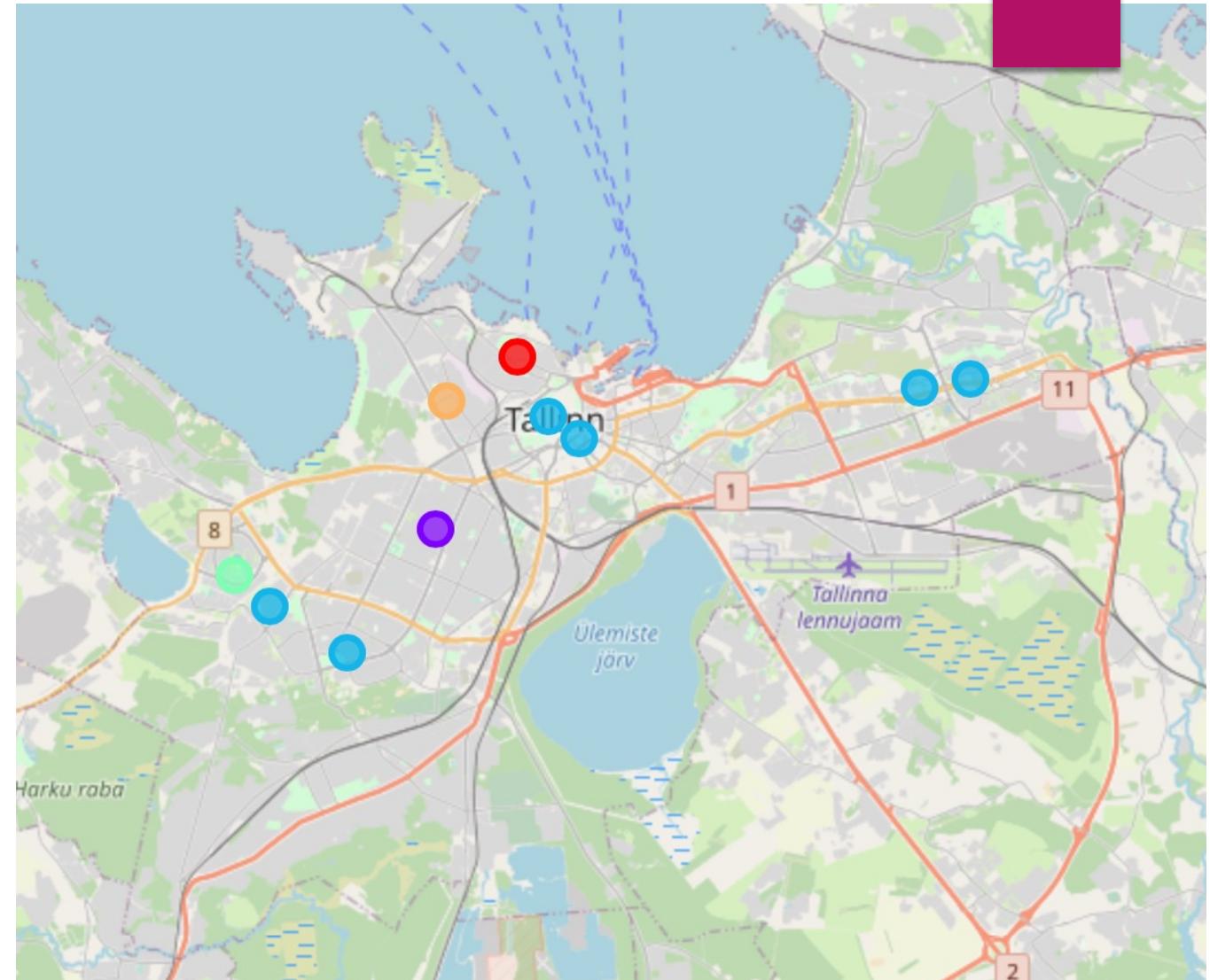


Popular venues of neighborhoods

	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	Kadaka	Bakery	Arcade	Stadium	Grocery Store	Electronics Store	Department Store	Coffee Shop	Pizza Place	Camera Store	Bus Station
1	Kalamaja	Bakery	Playground	Market	Trail	Park	Indie Theater	Theater	Dance Studio	Bar	Wine Shop
2	Lilleküla	Outdoors & Recreation	Café	Wine Shop	Electronics Store	Concert Hall	Convenience Store	Cosmetics Shop	Creperie	Cupcake Shop	Dance Studio
3	Mustakivi	Grocery Store	Playground	Restaurant	Park	Fast Food Restaurant	Bar	Indian Restaurant	Comfort Food Restaurant	Gym	Pool Hall
4	Mustamäe	Electronics Store	Auto Garage	Doner Restaurant	Farmers Market	Restaurant	Food & Drink Shop	Basketball Court	Pub	Bus Station	Athletics & Sports
5	Pelgulinna	Bus Stop	Bus Line	Café	Sports Club	Beer Garden	Supermarket	Fast Food Restaurant	Bus Station	Skate Park	Pizza Place
6	Südalinn	Hotel	Café	Restaurant	Cocktail Bar	Gym / Fitness Center	Cosmetics Shop	Italian Restaurant	Hostel	Coffee Shop	Movie Theater
7	Tondiraba	Furniture / Home Store	Motorcycle Shop	Bus Station	Supermarket	Home Service	Shopping Mall	Gym	Flower Shop	Electronics Store	Eastern European Restaurant
8	Vanalinn	Eastern European Restaurant	Restaurant	Scenic Lookout	Hotel	Theater	Coffee Shop	Cocktail Bar	Park	Plaza	Modern European Restaurant
9	Väike-Õismäe	Convenience Store	Bus Line	Grocery Store	Moving Target	Bus Station	Liquor Store	Wine Shop	Eastern European Restaurant	Cosmetics Shop	Creperie

Results of k-means clustering

- ▶ The neighborhoods are divided into 5 clusters of different color - red, purple, light blue, light green and yellow - based on their similarity measure.
- ▶ The above is result of k-means unsupervised ML algorithm deployed on the dataset.
- ▶ It's basically a lot of public space vs a lot of shopping vs central location.



Conclusion – a "high crime" neighborhood is whether ...

- ▶ A large neighborhood population wise with a lot of public space with pastime amenities (bakeries, cafes etc) and/or shopping facilities for assumingly local people (as they are lower ranked in popularity); or
- ▶ A large neighborhood with significant amount of shopping, attracting therefore assumingly crowds also from other neighborhoods; or
- ▶ City centre / Old Town which are always and typically attractive places for shops and services for general public.



THANK YOU!

