Applying data for urban neighborhood development

the case of Jernbanebyen

Bøje-Kovács, Bence János; Eriksson, Adam; Stenfeldt, Liselott; Thestrup, Sixten Maximillian

BUILD REPORT

Gehl

2023:04

Applying data for urban neighborhood development

the case of Jernbanebyen

Bence Boje-Kovacs, Adam Eriksson, Liselott Stenfeldt & Sixten Maximillian Thestrup

Bence Boje-Kovacs
Adam Eriksson
Liselott Stenfeldt
Sixten Maximillian Thestrup

Institut for Byggeri, By og Miljø, Aalborg Universitet
2023

TITLE	Applying data for urban neighborhood development
SUBTITLE	the case of Jernbanebyen
SERIAL TITLE	BUILD Report 2023:04
FORMAT	Digital
EDITION	1
UDGIVELSESÅR	2023
DIGITAL PUBLICATION DATE	March 2023
AUTHOR	Bence Boje-Kovacs, Adam Eriksson, Liselott Stenfeldt, Sixten Maximillian Thestrup
LANGUAGE	English
PAGES	41
REFERENCES	28
KEY WORDS	Jernbanebyen; micro-data; neighborhood creation; urban development
ISBN	978-87-563-2095-5
ISSN	2597-3118
DRAWINGS	jernbanebyen.dk
ILLUSTRATIONS	Sixten Maximillian Thestrup
PUBLISHER	Department of the Built Environment, Aalborg University A.C. Meyers Vænge 15, 2450 Copenhagen SV E-mail build@build.aau.dk www.build.auu.dk

CONTENT

1

INTRODUCTION

1 INTRODUCTION

As cities increase in population, new areas develop to make room for more residents. In a complex process, neighborhoods change and develop, balancing current residents' needs, densification, creating livable neighborhoods with accessible public spaces, and future residents' needs. As neighborhoods develop, their social identities change; production sites become residential areas, and arable land becomes new suburbs. Contemporary theories promote mixed use neighborhoods, proximity, and perimeter block density, such as in Soft City (Sim, 2019), and the concept of the 15-minute city (Moreno et al., 2021). Finding a balance in developing neighborhoods is a complex process and will affect the residential composition of neighboring areas.

One way to develop new neighborhoods more efficiently in urban contexts is to have accessible data on the social interactions in and the demographic and socioeconomic profile of neighborhoods. With access to this type of information, decision makers can make more evidence-based decisions. However, access to the necessary information may be limited. Reasons for this might be lack of access to up-to-date data or expertise to process the data. Finally, the most granular data is often represented either within administrative boundaries or arbitrary 100x100 meter grid cells in a way that does not represent urban dynamics.

One of the challenges, even when having access to data, is the clustering of arbitrary areas into neighborhoods. It is in the interest of the decision-makers that aggregated statistics are represented in spatially meaningful areas. Within the city, neighborhoods can be defined using the existing boundaries of the built environment, such as roads and buildings. However, establishing such neighborhoods in an automated manner is not only complex, but also requires striking a balance between granularity and concerns regarding the privacy of the individual.

Historically, data use in cities has existed as long as cities have, albeit in different forms and amounts (Williams, 2022). Data use increased together with cities during the industrial revolution (Williams, 2022) and new data collection techniques have been developed. From a social movement caring for people's activities in cities in the 1960s (Gehl & Svarre, 2013), Whyte (1980) and Gehl (2011) developed observational data collection techniques to measure people's use of public space and public life in cities (Gehl & Svarre, 2013). Some of these techniques were later adopted by cities as part of their development goals (Gehl & Svarre, 2013). By contributing techniques for visualizing socio-demographic data in a socially meaningful way, better representing residents in neighborhoods, we hope urban planners and decision-makers can adopt and use the data more effectively, and in conjunction with observational data, as part of developing neighborhoods.

Denmark's capital, Copenhagen, is growing and has had many new neighborhood developments lately; Carlsbergbyen, Nordhavn, Ørestad, just to name a few. A previous train and logistics terminal, Jernbanebyen, is another neighborhood in Copenhagen that will be developed, with the first construction commencing in 2024 (jernbanebyen.dk, 2023 [1].). The area is centrally located in Copenhagen, in a previously blocked-off area in the middle of the city, surrounded by historically and socially different neighborhoods.

With a centrally located new development, the new neighborhood will shift the urban dynamics in that part of the city by connecting the different surrounding neighborhoods. In order to understand how it will shift, we first need to know what the residential composition of the surrounding areas looks like. Therefore, in this paper, we explore the demographic and socioeconomic dynamics in the neighboring areas to Jernbanebyen. By constructing meaningful neighborhoods combined with register microdata, we can visualize relevant demographic and socio-economic profiles of the surrounding neighborhoods. The socio-demographic data represented in a meaningful way allows for creating compositions of user profiles in the different neighborhoods. With this data, we find great potential in combining this data with observational data, to further understand the social dynamics and movements in the area.

2

BACKGROUND - JERNBANEBYEN

2 BACKGROUND - JERNBANEBYEN

Jernbanebyen has historically been a freight station and train workshop, inaugurated in the beginning of and then developed throughout the 20th century (jernbanebyen.dk, 2023 [2]). A small worker's village, *Den Gule By*, houses the railway company DSB employees and operators throughout the century (jernbanebyen.dk, 2023 [2]). As workshops are relocated and freight operations are discontinued, the area is vacated from train activity in the beginning of the 21st century. In 2009, the area opened up for other businesses to use the existing spaces, and in 2020 new student housing welcomes the first new residents to the area (jernbanebyen.dk, 2023 [2]).

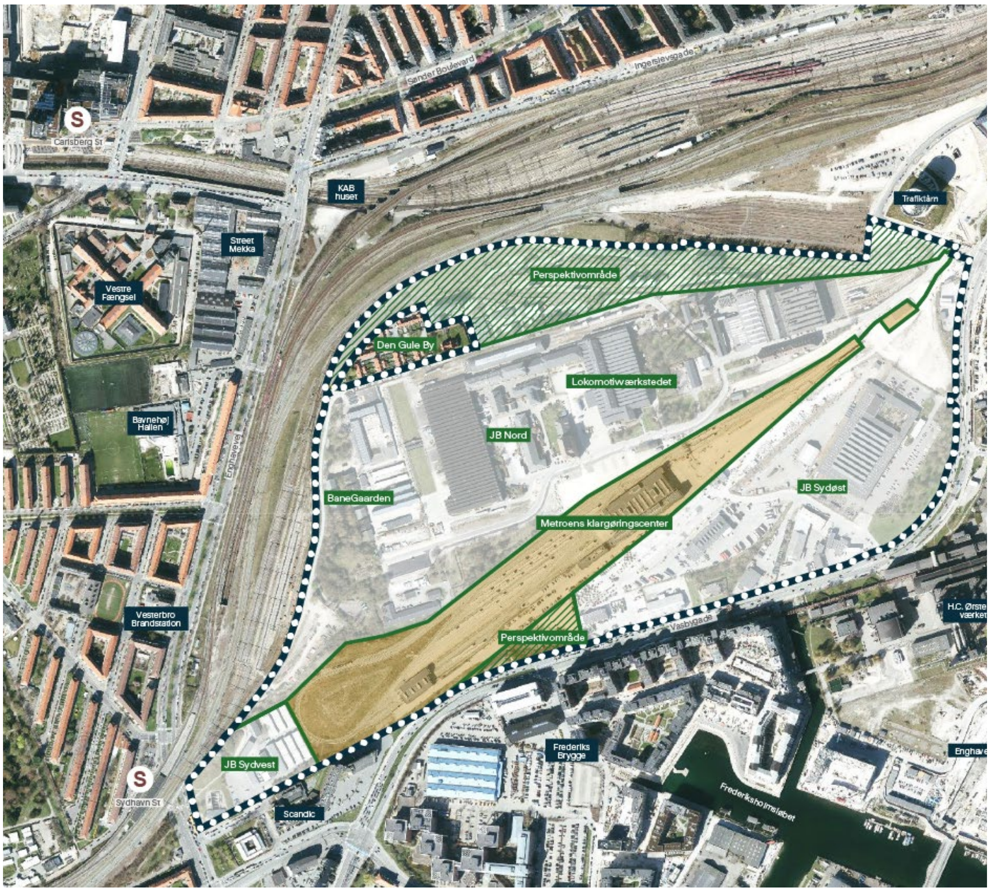

FIGURE 1. Overview of Jernbanebyen (jernbanebyen.dk, 2023 [3])

The area is owned by and is going to be developed by the state-owned DSB Ejendomme and Freja Ejendomme together with a consortium of private developers which joined in 2023 (jernbanebyen.dk, 2023 [4]). The new area is estimated to be 365.000 square meters and will consist of private and public housing as well as local businesses, with an emphasis on being a green and sustainable neighborhood (jernbanebyen.dk, 2020). The landowner's ambitions are to develop Jernbanebyen together with Copenhageners and want to over time create a lively neighborhood for residents and recreation (jernbanebyen.dk, 2023 [5]).

The 2020 master plan competition brief addresses that Jernbanebyen needs to develop from an isolated island to an integrated part of the city (jernbanebyen.dk, 2020). With a limited number of connections to the surrounding areas, with tracks and roads enclosing the space, there are clear physical barriers for integrating the space (jernbanebyen.dk, 2020).

To be an integrated part of the city, Jernbanebyen needs to attract residents in the neighboring areas, where the surrounding areas have different characteristics; the closest parts of Vesterbro with a prominent café culture, a green Sønder Boulevard, the largest shopping mall in Denmark, Fisketorvet, on Kalvebod Brygge, new developments around Holmene and historical working quarters from the 1930s in Bavnehøj (jernbanebyen.dk, 2020).

3

CREATING NEIGHBORHOODS

3 CREATING NEIGHBORHOODS

One of the main challenges in creating neighborhoods is achieving spatially meaningful neighborhoods. Being spatially meaningful refers to each neighborhood following the existing boundaries of the built environment, such as roads and buildings. At the same time, because the purpose of creating neighborhoods is to aggregate individual-level descriptive statistics at a fine-grained geospatial level, they should also adhere to national privacy regulations, such as minimum population counts.

We have developed a hierarchical approach to algorithmically create on-average compact districts that adhere to privacy constraints since 1990, using population and household counts from Statistics Denmark alongside open-source road and municipality boundary data. The algorithm uses three inputs: a geographical area (polygon) to be split into neighborhoods, OpenStreetMap linestring road data (Open Street Map, 2023), and a vector of constraints. An overview of the proposed algorithm is presented in Figure 2.

The starting point was 98 Danish municipalities. The hierarchical aspect of the algorithm is to group road data according to their importance and perform a sequential iterative split-union of an initial polygon. Our chosen grouping of road types is (based on the assigned OpenStreetMap *fclass*)

1) Major roads; *highway, trunk, primary, secondary*
2) Minor roads; *tertiary, unclassified*
3) Residential; *residential, living street, service*
4) Paths; *pedestrian, footway, bridleway, cycleway, track, path*

To illustrate, the first iteration splits each municipality according to the major roads and then unions the resulting split, adhering to our chosen vector of constraints. Next, each resulting polygon was split according to the second level of roads (minor roads).

Constraints determine whether a polygon resulting from a split is unioned. They refer to both the spatial characteristics of the polygon and interpolated population counts. In summary, the chosen constraints are as follows:

1) Minimum interpolated population count of 100
2) Minimum interpolated household count of 50
3) Minimum area of 1 hectare
4) Not enclosed
5) Minimum compactness as a function of interpolated population count

The population and household counts were the minimum required according to the privacy regulations set out by Statistics Denmark. The population and household counts for a given polygon will always be an estimation because the counts are provided on a 100 × 100 m grid and not geolocated within each polygon. To determine whether a polygon is enclosed, a given set of polygons was first converted into a graph. Next, we check whether any bi-connected component shares its total border with its corresponding articulation point. The compactness constraint for a polygon is modeled as an exponential decay function of the population count, where a higher population count leads to a decrease in the required compactness score. This is because of the expectation that a polygon with a high population count will be further split in future iterations.

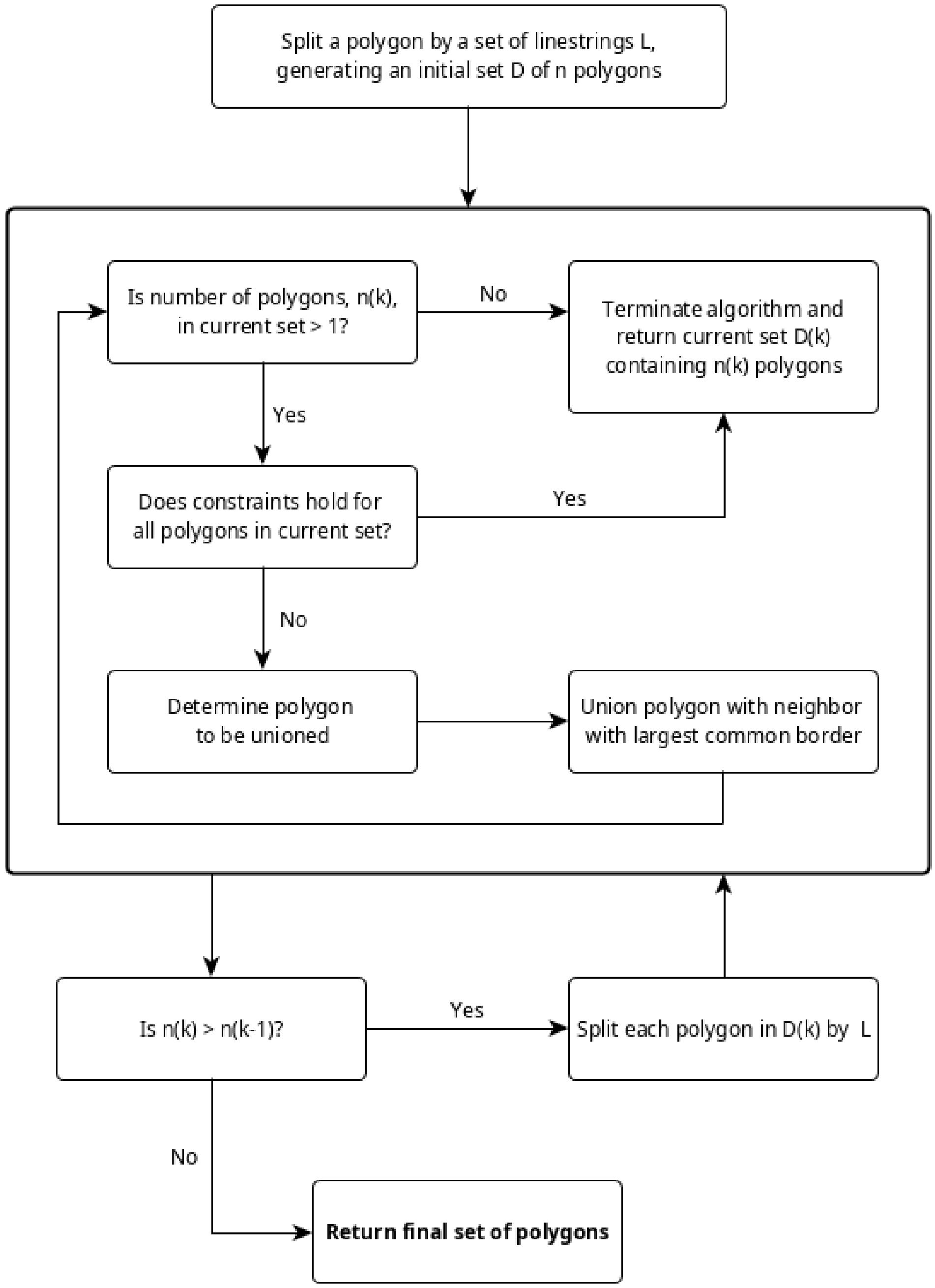

FIGURE 2. Algorithm flow-chart for one iteration

Following this procedure, we created 15.000 neighborhoods across 98 municipalities in Denmark. Our approach has several noteworthy characteristics. First, it is deterministic because it does not rely on random seeds for selecting the district to union during a given iteration. This implies that the same input set of districts results in the same output districts. Compared with existing algorithms, such as max-p (Duque et al., 2012) and Sequential Monte Carlo (McCartan and Imai, 2023), one advantage of this is that we do not need to provide a fixed number of output districts. However, we can easily supply a fixed number of districts to the constraint vector. Second, districts that already adhere to the vector of constraints prior to unioning are not guaranteed to remain intact, which differs from existing algorithmic redistricting approaches. Consequently, we do not attempt to maximize the number of districts during clustering, but rather strike a balance between having compact districts and meeting imposed constraints. Third, the algorithm allows great flexibility in including any constraints of choice. However, the average expected number of output neighborhoods drops drastically with the introduction of additional constraints or by setting constraint values that are too severe.

Finally, to ensure that the privacy constraints are satisfied, we map the final set of polygons back onto the grid that contains the population and household counts. Any neighborhood that potentially falls below the threshold is unioned with the neighbor that it shares the largest common border with.

It is worth noting that the constraint of having at least 100 people and 50 households since 1990 within each neighborhood has the effect of making the areas larger compared to using only population counts since, for example, 2020. This is illustrated in Figure 3, where it is clear that some of the surrounding areas have experienced new residential developments in the last 30 years, leading to a notable increase in population.

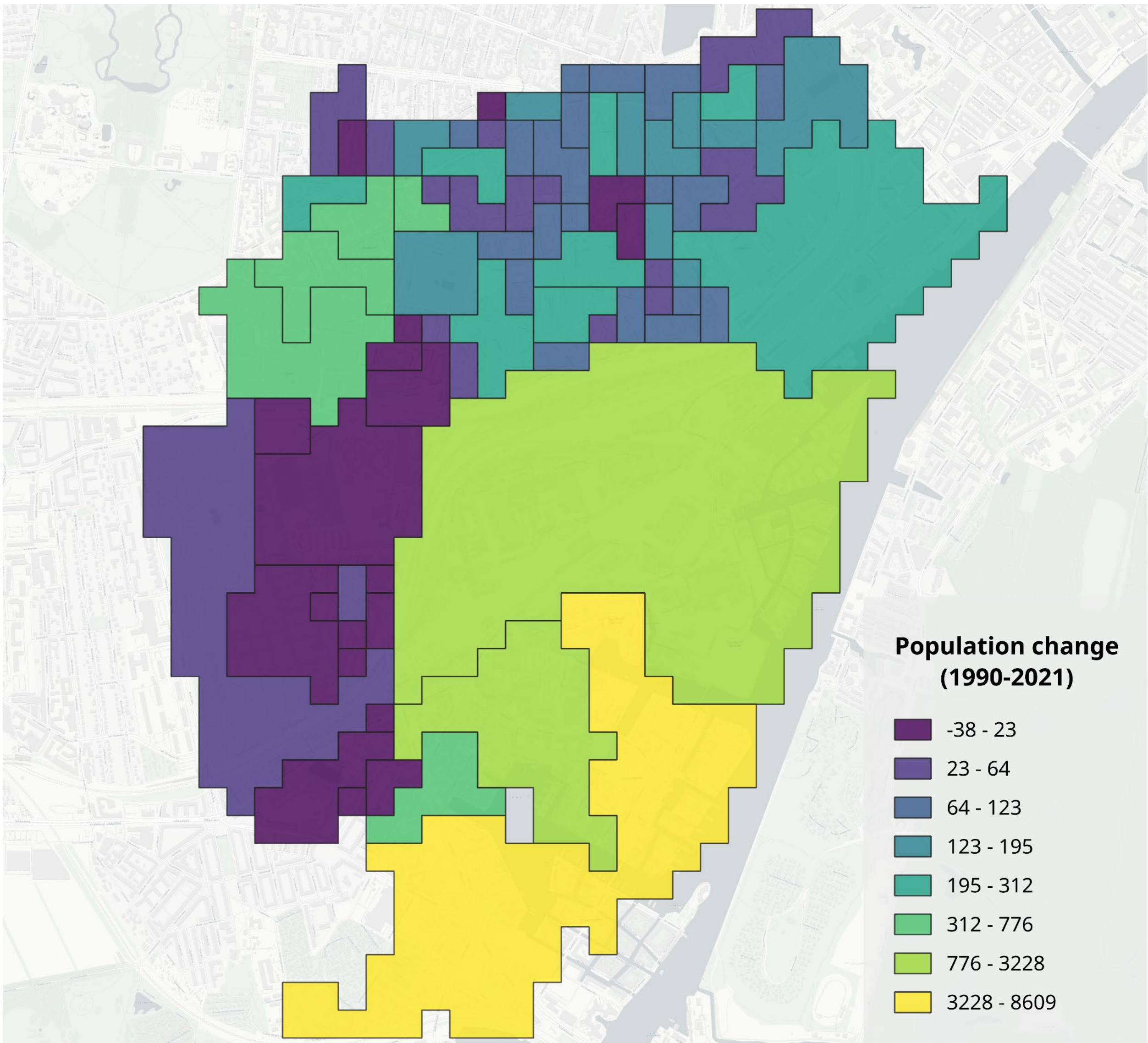

FIGURE 3. Change in neighborhood population from 1990 to 2021

4

DATA

4 DATA

This paper combines small-scale neighborhoods with Danish administrative register data from 2020. We obtain from the register data yearly information on each individual's address, socioeconomic, demographic, as well as housing characteristics. We link individuals to neighborhoods by using individuals' uniquely identified addresses.

To understand the socio-demographic profile of the surrounding neighborhoods of Jernbanebyen, we use the following variables: age, age of the children in the household, income, employment, education, crime, non-Western ethnic background[1], and living in public housing. In particular, we divide the age of the children into three categories: up to six years, between six and 16 years, and finally 16 years and above. Moreover, we distinguish between residents being employed and outside of the labor market. The latter, we define as individuals aged 16-65 receiving either social aid or a disability pension and not being enrolled at an educational institution. Likewise, we look at the education level of the neighborhood, especially the share of residents with high education. A resident is highly educated if she or he aged 25-65 years has at least a bachelor's degree. We define criminal residents if individuals aged 16 or above have been convicted of any type of crime.

[1] If a country does not belong to one of the following countries: All EU countries as well as Andorra, Iceland, Liechtenstein, Monaco, Norway, San Marino, Switzerland, Great Britain, Vatican City, Canada, USA, Australia, and New Zealand, we define it as non-Western.

5

RESULTS

5 RESULTS

Table 1 shows the descriptive statistics for each above-mentioned characteristics for the surrounding areas in 2020. The Table reveals that there is a large variation in all characteristics among the 80 neighborhoods surrounding Jernbanebyen. For instance, there are neighborhoods without any public housing compared to neighborhoods where the housing stock only consists of public housing. Similarly, there are neighborhoods with an average share of residents out of the labor of 4 percent compared to neighborhoods where two out of five residents are out of the labor market.

To highlight the variation in the residential composition across neighborhoods around Jernbanebyen, we plot the average characteristics of residents on several maps. Figure 4 shows the share of public housing in the neighborhoods and reveals a clear pattern. Public housing is dominant in the residential area of Sydhavn[2] and negligible in the area of Vesterbro[3]. Nevertheless, there are few neighborhoods in Vesterbro with public housing above 10 percent of the overall housing stock. When moving on to the demographic variables, Figure 5 shows that the share of residents aged 16-67 is relatively evenly spread across neighborhoods. Next, we look at the residential composition in relation to children. As mentioned in the data section, we created three groups of children based on age. Figure 6 reveals that neighborhoods in Vesterbro seem to be more heterogeneous compared to the surrounding neighborhoods of Sydhavn. Consequently, the share varies between 2- 4 and 17-25 percent. Overall, Figure 6 suggests that the concentration of children under age 6 is less pronounced in the surrounding neighborhoods of Sydhavn compared to the surrounding neighborhoods in Vesterbro. We find a similar picture among children aged 6-16 as shown in Figure 7. Looking at the distribution of non-Western immigrants we find that even though there is variation in the share of non-Western immigrants across neighborhoods in Sydhavn the overall concentration of this group is significantly higher compared to Vesterbro.

TABLE 1. Descriptive statistics of neighborhoods in 2020

	Obs	Mean	Std. dev.	Min	Max
Population aged 16-67 years	80	77.596	5.869	57.666	95.122
Children under 6 years	80	7.651	2.910	2.091	24.646
Children aged 16-16 years	80	7.372	2.654	0.697	14.537
Non-Western immigrants	80	11.890	7.568	1.980	31.632
Public housing	80	17.752	29.892	0	100
Employed	80	72.989	8.752	45.625	85.241
Convicted of crime	80	3.194	1.628	0.621	11.211
Personal Gross income (in DKK)	80	338,357	75,194	173,033	609,899
High education	80	58.892	12.576	23.077	79.251

Source: authors' own calculation based on register data from Statistics Denmark

[2] By Sydhavn we refer to the following residential areas: Bavnehøj, Gl Sydhavn, and Holmene (Københavns Kommune, 2023)

[3] By Vesterbro we refer to the following residential areas: Vesterbro Vest, Vesterbro Central, and Vesterbro Øst. (Københavns Kommune, 2023)

Finally, we describe in Figures 9-12 the socioeconomic composition of the surrounding neighborhoods. First, Figure 9 shows that the share of employed residents aged 16-67 does not go below 60 percent. On the contrary, several neighborhoods in the area of Sydhavn have a share of employed residents below 60 percent. The same pattern holds for personal gross income and the level of education. Consequently, neighborhoods in Vesterbro are richer and more educated than neighborhoods in Sydhavn. Likewise, it is worth mentioning that the residents in immediate neighborhoods of Jernbanebyen in Vesterbro on average earn more, are slightly more employed and commit less crime than the other neighborhoods in Vesterbro[4]. Comparing Figure 4 with Figures 9-12 shows a clear correlation between housing stock and socioeconomic resources. In particular, neighborhoods with a high share of public housing tend also to have a relatively low contraction of residents with employment, as well as high education level[5].

[4] On the contrary, they have the same concentration of residents with high education

[5] The raw correlations between the share of public housing in the neighbourhood and socioeconomic characteristics such as employment, personal gross income, high education, and crime are -0.84, -0.54, -0.86, and -0.52.

6

DISCUSSION & CONCLUSIONS

6 DISCUSSION & CONCLUSIONS

In this paper, we aim to understand the socio-demographics of the surrounding neighborhoods of the developing area of Jernbanebyen by using detailed spatial data combined with register microdata. Our overall goal is to demonstrate how we can inform and improve current and future developments in cities by using data.

One of the major advantages of having access to neighborhoods at a fine-grained geospatial level, is the potential to carry out more detailed analyses across time and space. It enables, for example, the possibility of evaluating the effect of interventions in the built environment within neighborhoods through the change in spatial autocorrelation for core socio-demographic indicators. Due to the granularity of the neighborhoods and yearly updated data, it will be possible to separate intervention-specific effects from effects which apply to a greater area within the city.

We see the socio-demographic data represented in socially meaningful neighborhoods show nuances and shifts on a very granular level, but also overall patterns between neighborhoods.

For example, the Figures show a clear division between the residential areas of Sydhavn and Vesterbro. Overall, Vesterbro consists of residents who are to a larger extent employed, have a higher education, earn more and commit less crime compared to Sydhavn. Another example is that the neighborhoods with a larger share of public housing, Sydhavn and Bavnehøj, also on average earn less and have lower education than other neighborhoods surrounding Jernbanebyen.

However, the main takeaway is that neighborhoods within these relatively small residential areas, which urban planners would treat as homogenous areas, have a large variation both in the demographic and socioeconomic composition of the residents. For example, parts of Vesterbro have the same share of income and non-Western immigrants as other parts of Bavnehøj and Sydhavn. Another example in Vesterbro is in an otherwise child-dense neighborhood, two blocks have the lowest share of children aged 6 to 16. One of the blocks has an above average large share of children younger than 6, whereas the other also has the lowest share of children younger than 6, implying that one of the blocks consists of many families with small children, and the other has a comparatively low amount of children.

With the granular socio-demographic data, different resident profiles can be attributed within one neighborhood, among multiple neighborhoods, highlighting nuances within each area. These profiles can be used to observe current use of the space as well as planning for future use. Instead of having one, too general and usually not useful, profile of a neighborhood, the neighborhood can be understood from the composition of multiple profiles. Understanding the current dynamics in neighboring areas allows for a more detailed analysis of potential dynamics in the developing neighborhood. It allows for urban planners to tailor activities and develop the built environment to cater to the compositions of user profiles from different areas.

When temporarily activating developing neighborhoods, to attract locals, activations can be tailored to the surrounding social composition. For example, activating the area for families or for daytime activities will attract different user groups from different neighborhoods and blocks. Further, to understand the impact of activation programs, and development, observational methods as suggested in Gehl (2011), and Gehl and Svarre (2013) can be used. Combining observations on who is using the space, what activities are done and where they use the space with the neighborhood compositions has potential to further enhance understanding the social dynamics in the neighborhoods. The combination of data enables for understanding resident profiles and social activity in the space. Comparing these granular socio-demographic with observational data can highlight which resident groups are present in the area, and which aren't. This combination could help urban planners guide their actions in a new development.

For the development of Jernbanebyen to become a lively neighborhood and an integrated part of Copenhagen, it is important to connect to the current neighboring areas. With socio-demographic data represented in a socially meaningful way, urban planners and decision makers have a better understanding of the social composition of the area. We see that the granularity enables for understanding neighborhoods as a composition of multiple profiles. The profiles show a lot of promise to be combined with observational data to further understand the social dynamics in neighborhoods and evaluate early activations. In future work, we wish to combine use of the presented socio-demographic data with eye-level observations to further enhance our understanding of the social dynamics in the area.

7

REFERENCES

7 REFERENCES

Duque, J. C., Anselin, L., & Rey, S. J. (2012). The max-p-regions problem. *Journal of Regional Science*, *52(3)*, 397-419. https://doi.org/10.1111/j.1467-9787.2011.00743.x

Gehl, J. (2011). Life Between Buildings. Island Press, Washington D.C.

Gehl, J. & Svarre, B. (2013), How to study public life. Island Press, Washington D.C.

jernbanebyen.dk (2020). Program for parallelopdrag om byudvikling i Jernbanebyen. jernbanebyen.dk, accessed 2023-03-01. URL: https://indd.adobe.com/view/0277ac5c-497a-4dee-96f0-53c16d0711a3

jernbanebyen.dk (2023) [1]. Tidsplan for Jernbanebyen. jernbanebyen.dk, 2023-03-01. URL: https://jernbanebyen.dk/artikel/tidsplan-jernbanebyen

jernbanebyen.dk (2023) [2]. Historisk tidslinje. jernbanebyen.dk, 2023-03-01. URL: https://jernbanebyen.dk/artikel/historisk-tidslinje

jernbanebyen.dk (2023) [3]. Kort over Jernbanebyen. jernbanebyen.dk, 2023-03-01. URL: https://jernbanebyen.dk/artikel/kort-over-jernbanebyen

jernbanebyen.dk (2023) [4]. Fire parter går sammen om den samlede udvikling af Jernbanebyen. jernbanebyen.dk, 2023-03-01. URL: https://jernbanebyen.dk/nyhed/fire-parter-gaar-sammen-om-den-samlede-udvikling-af-jernbanebyen

jernbanebyen.dk (2023) [5]. Grundejernes vision for Jernbanebyen. jernbanebyen.dk, 2023-03-01. URL: https://jernbanebyen.dk/artikel/grundejernes-vision-jernbanebyen

Københavns Kommune (2023). City of Copenhagen Quarters. URL: https://www.opendata.dk/ city-of-copenhagen/kvarter. Accessed 2023-03-05.

McCartan, C., Imai, K. (2023). Sequential Monte Carlo for sampling balanced and compact redistricting plans. arXiv. arXiv:2008.06131v5

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the “15-Minute City”: Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. *Smart Cities, 4(1),* 93–111. https://doi.org/10.3390/smartcities4010006

OpenStreetMap contributors (2023), Planet dump [Data file from 2023-02-08]. Retrieved from https://planet.openstreetmap.org

Sim, D. (2019). Soft City. Island Press, Washington D.C.

Williams, S. (2022). Data action: Using data for public good. MIT Press.

Whyte, W.H. (1980). The social life of small urban spaces, The Conservation Foundation, Washington, D.C.

8

APPENDIX

8 APPENDIX

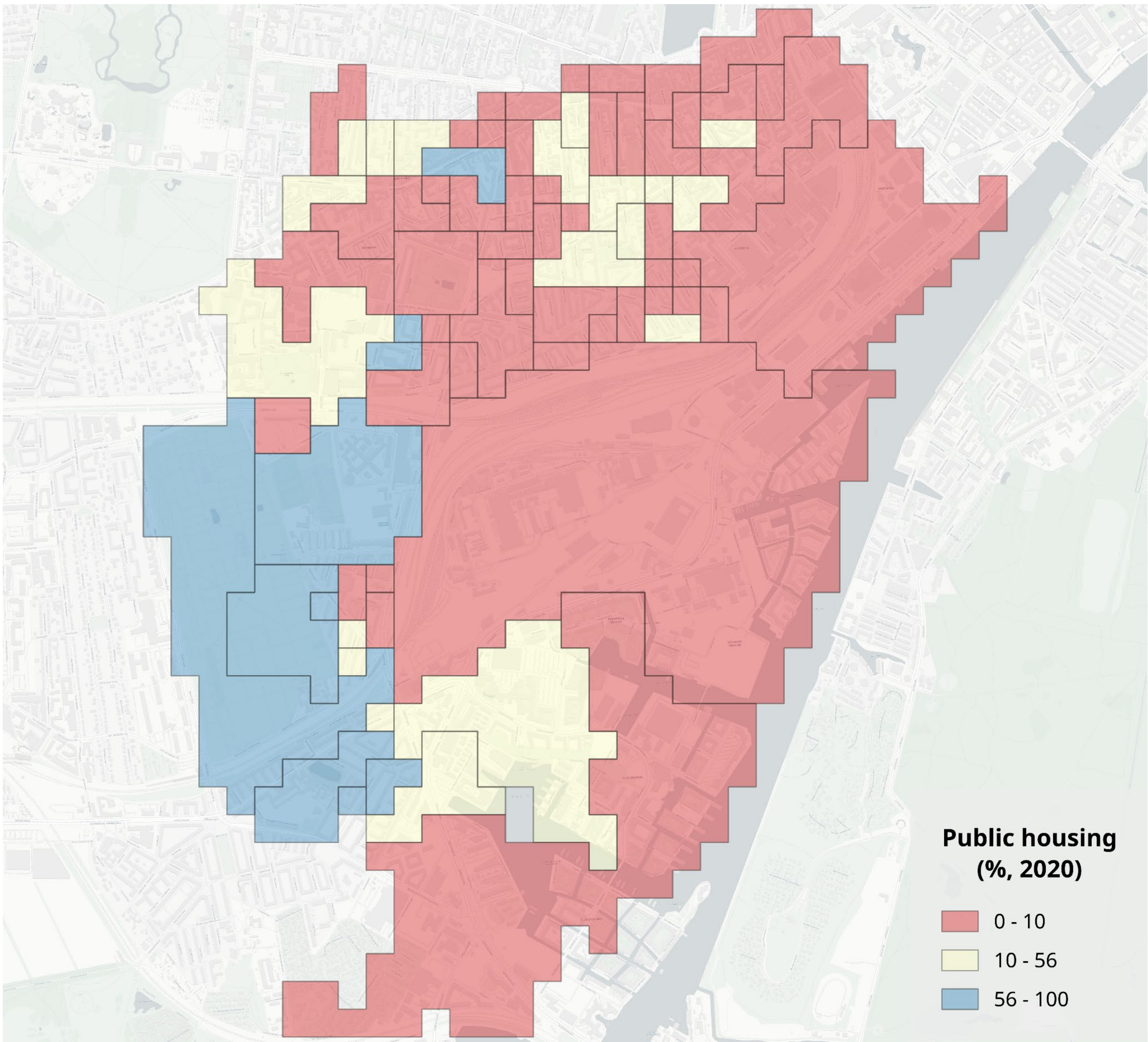

FIGURE 4. Neighborhood percentage of public housing

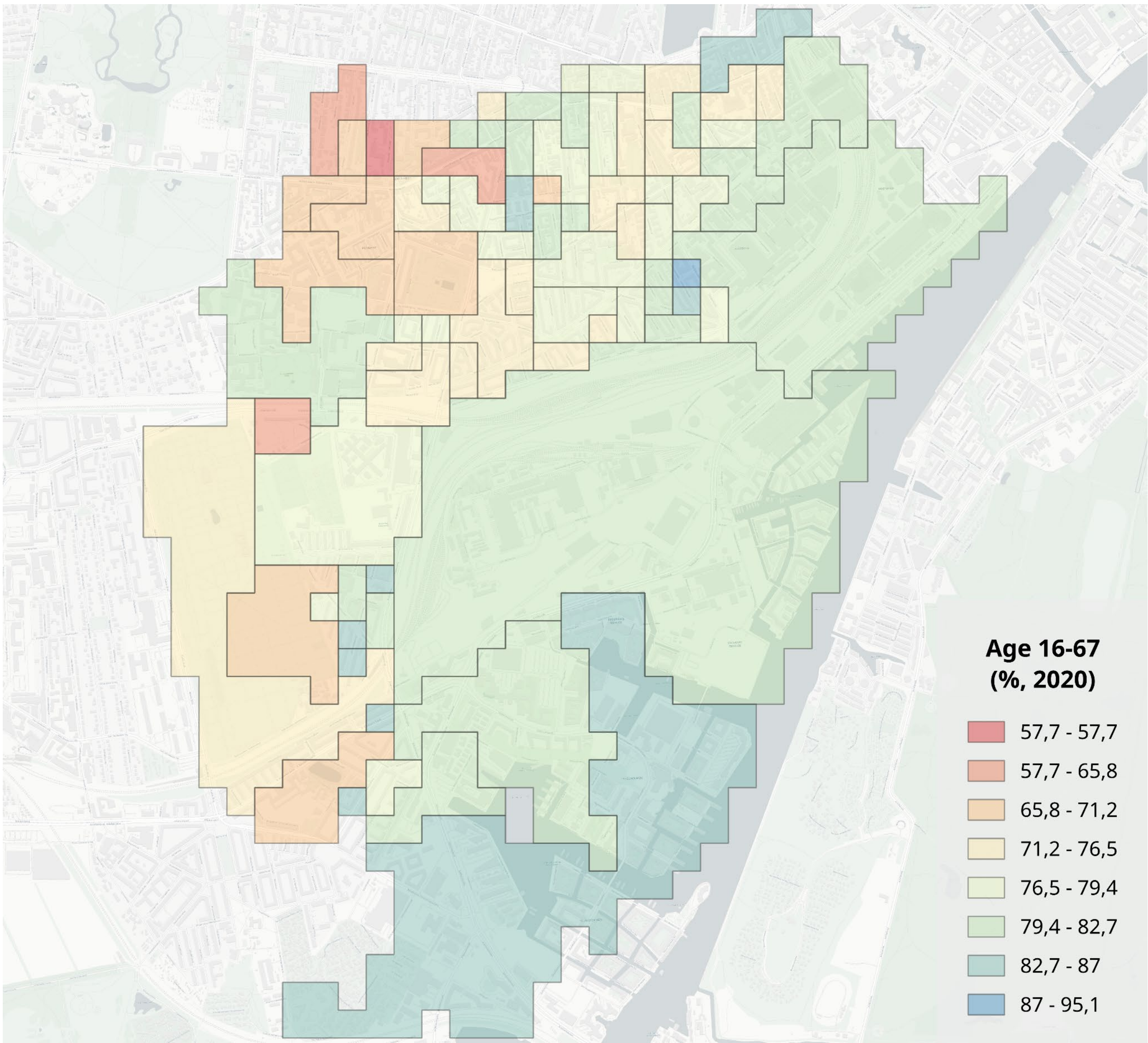

FIGURE 5. Neighborhood percentage of individuals aged 16-67 years

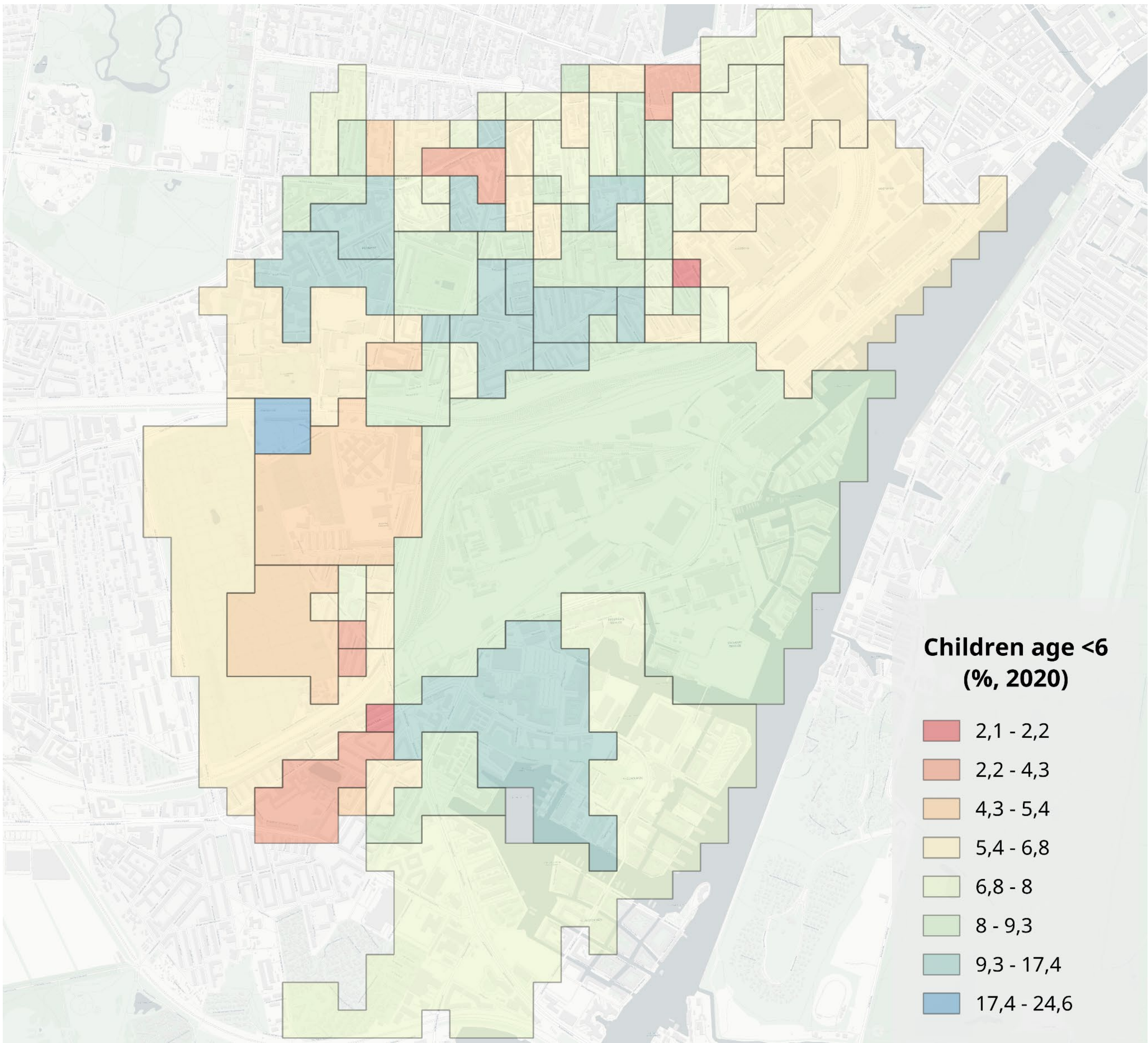

FIGURE 6. Neighborhood percentage of children below age 6 years.

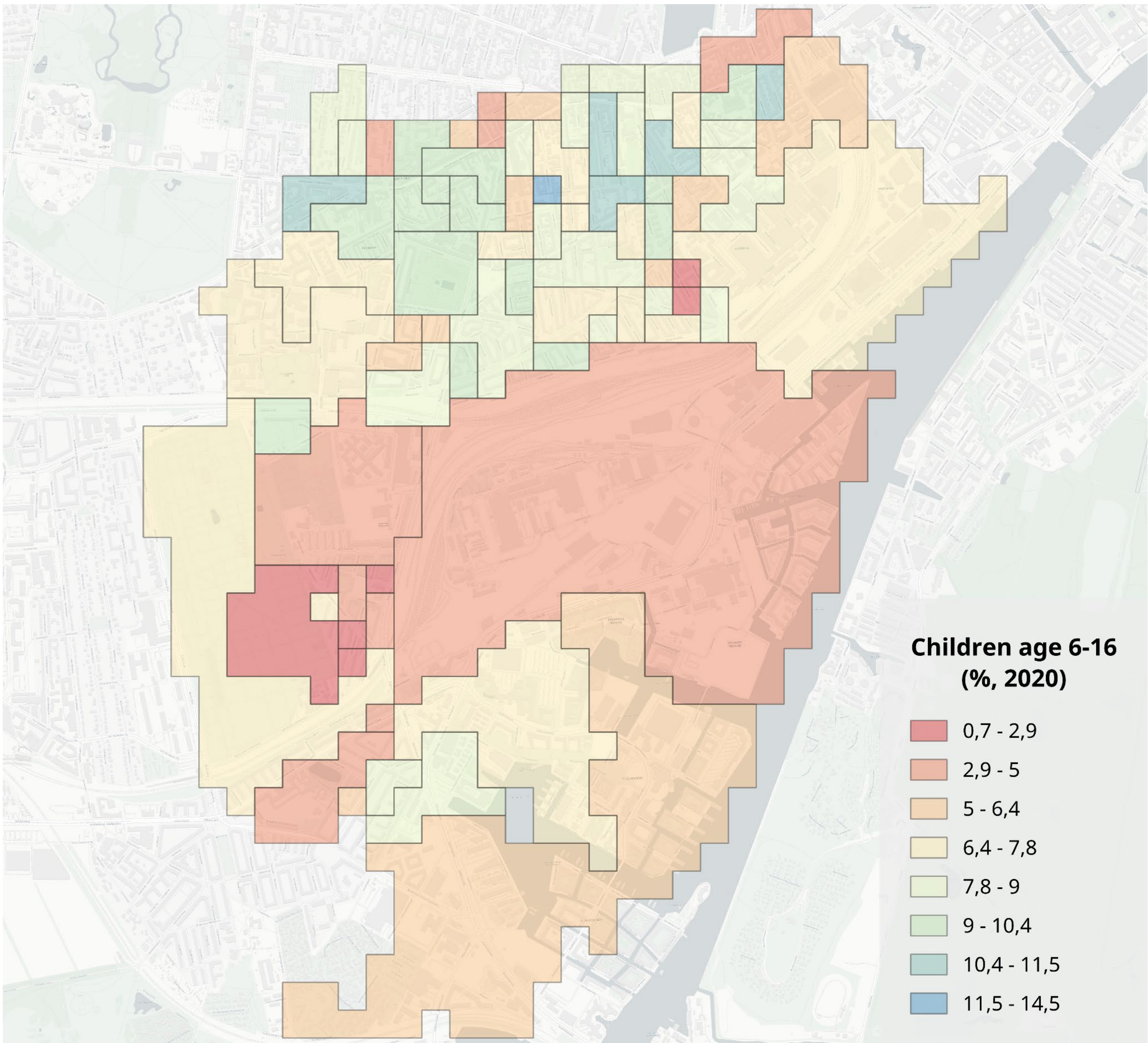

FIGURE 7. Neighborhood percentage of children aged 6-16 years

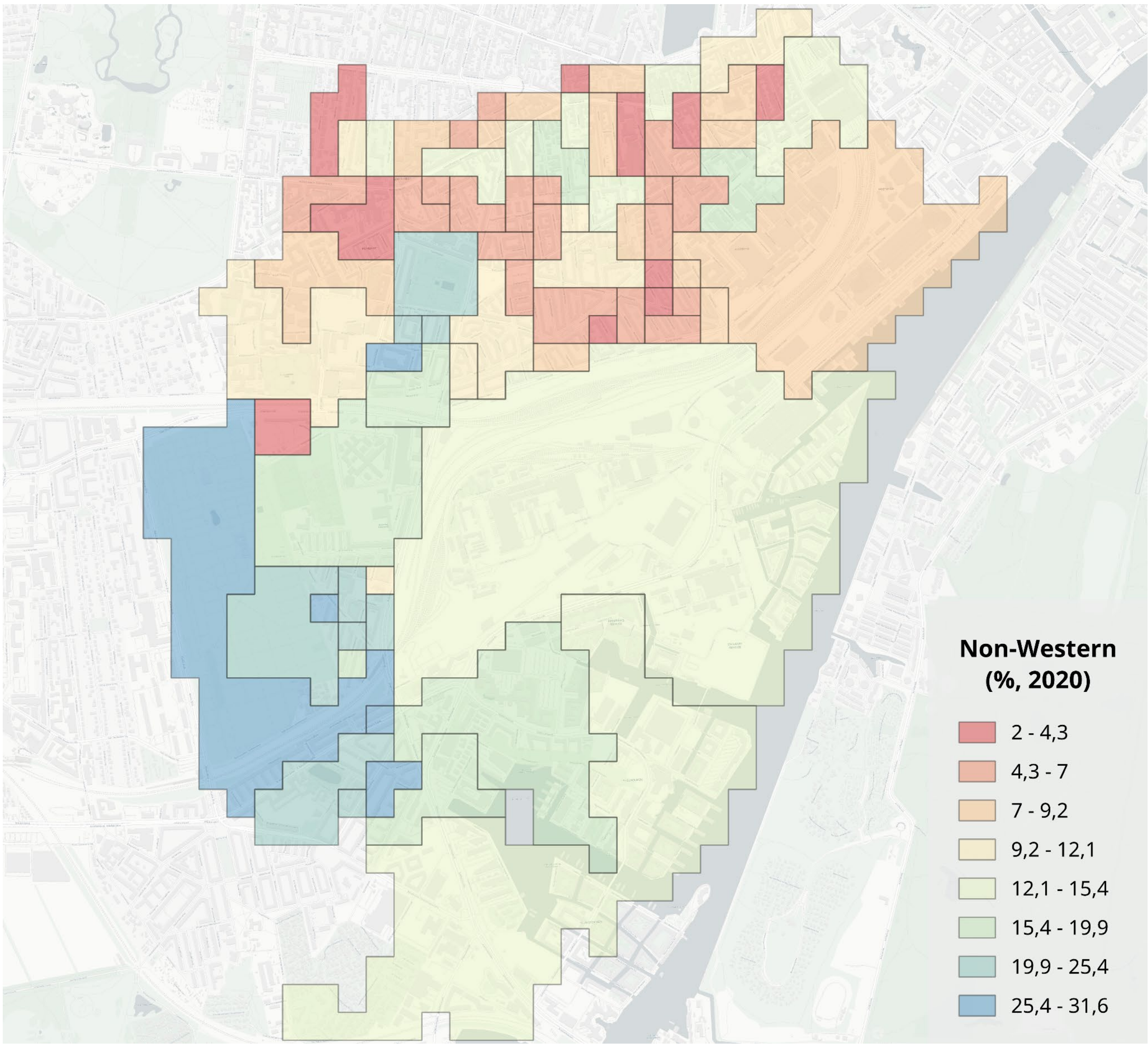

FIGURE 8. Neighborhood percentage of non-Western residents

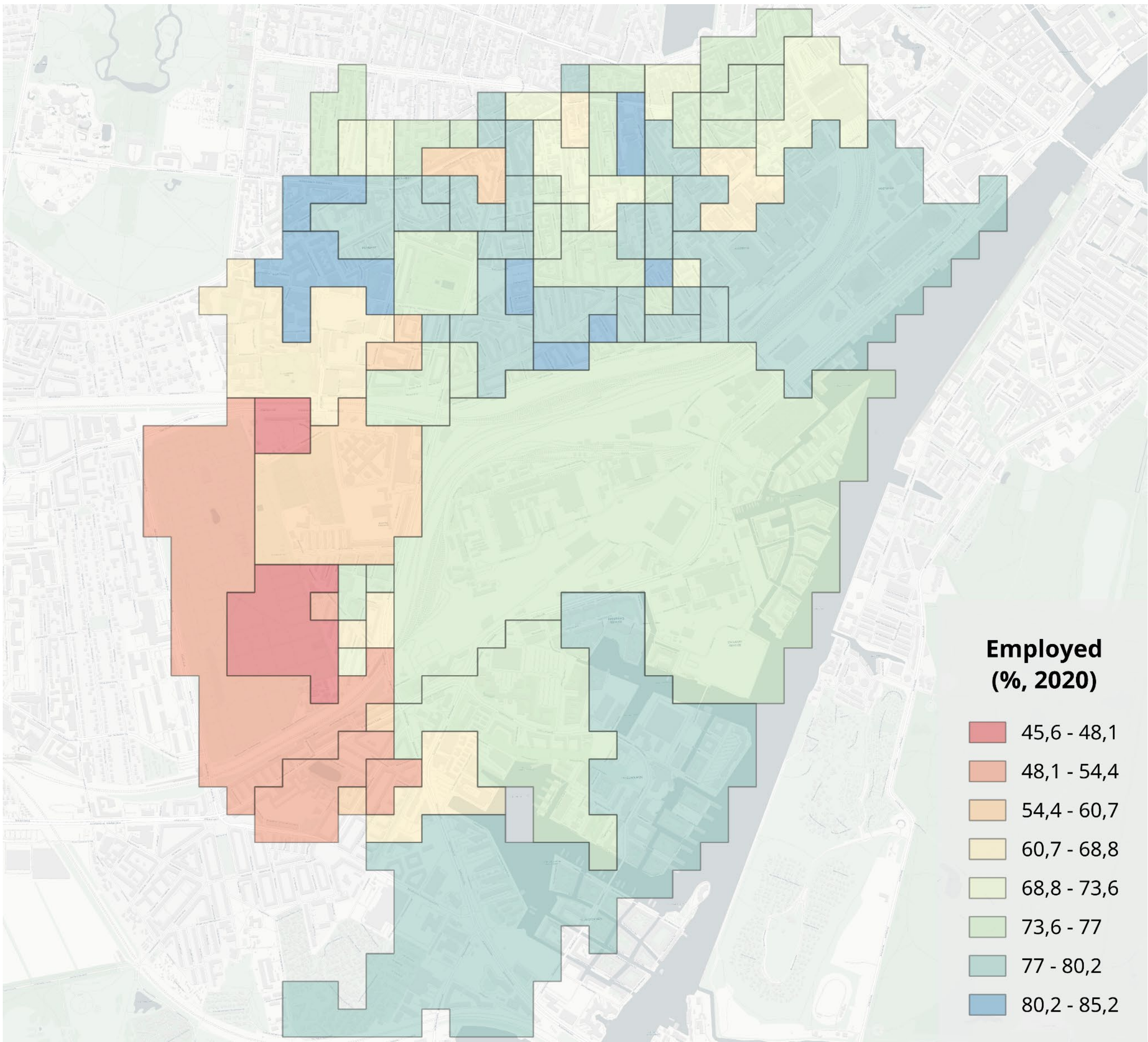

FIGURE 9. Neighborhood percentage of employed residents

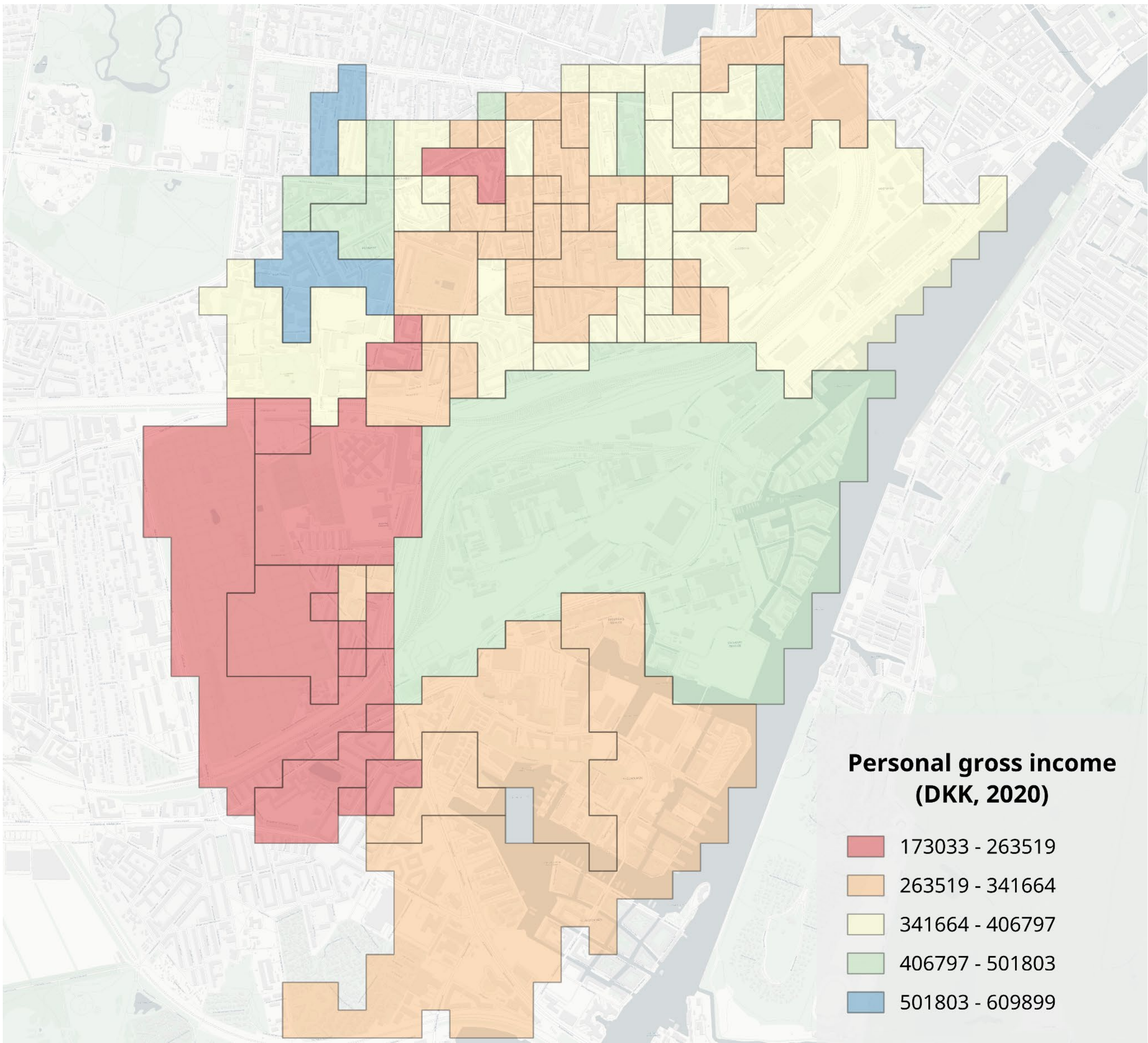

FIGURE 10. Neighborhood average personal gross income

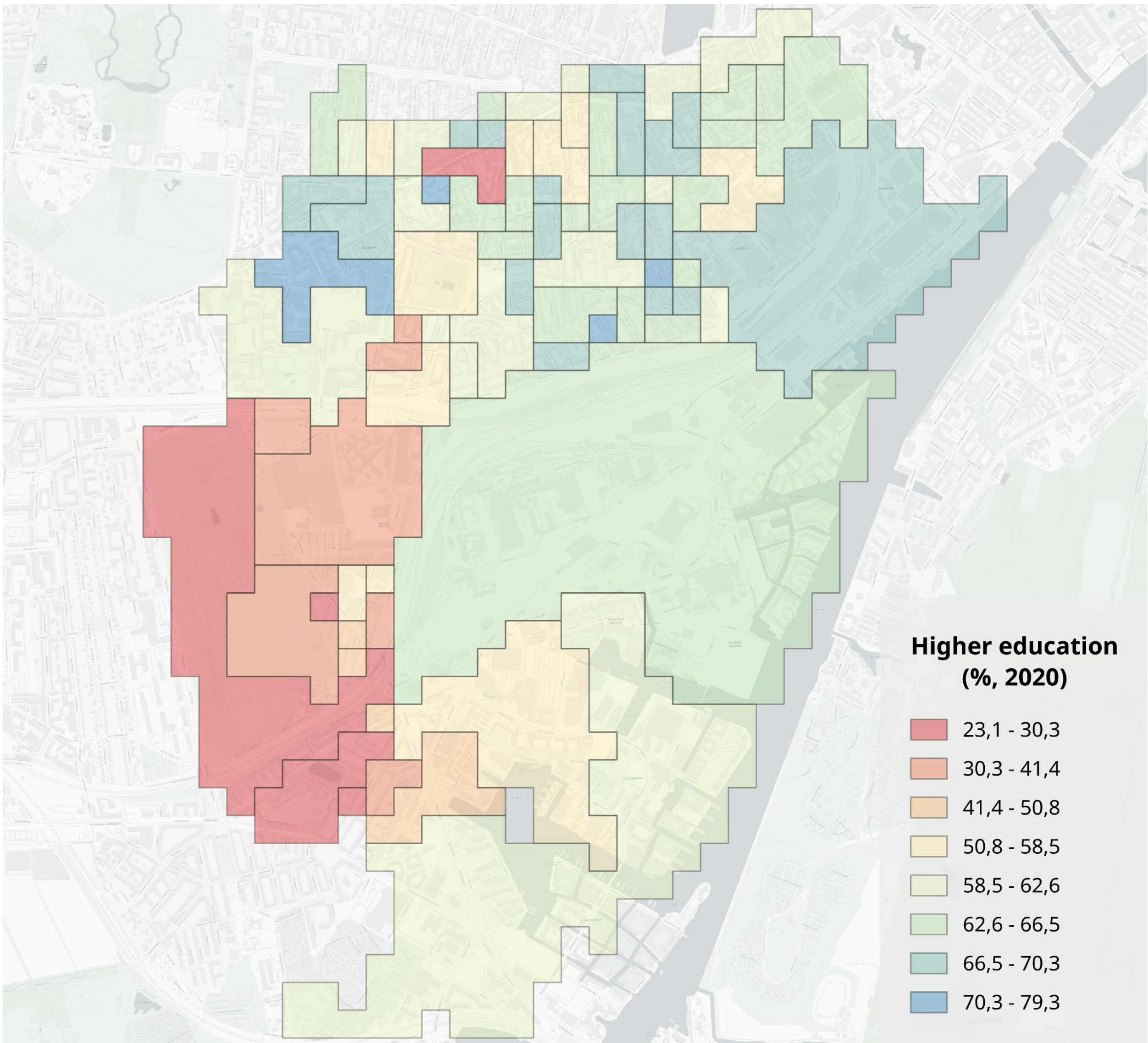

FIGURE 11. Neighborhood average of residents with higher education

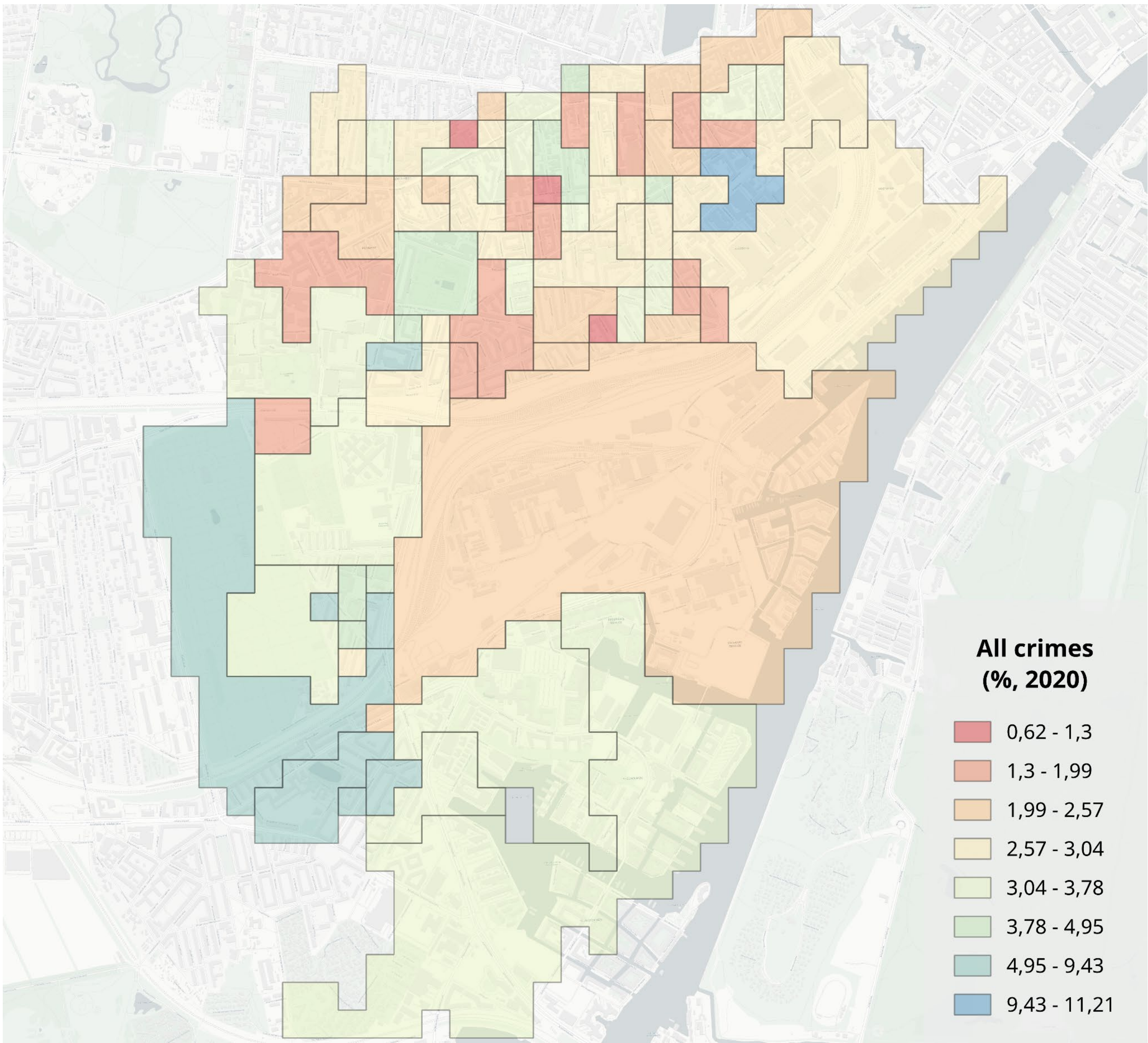

FIGURE 12. Neighborhood average of residents convicted of crime

Applying data for urban neighborhood development; the case of Jernbanebyen

Jernbanebyen – a previous train and logistics terminal - is a neighborhood in Copenhagen that will be developed. The area is centrally located in Copenhagen, in a previously blocked-off area in the middle of the city, surrounded by his-torically and socially different neighborhoods. The new area is estimated to be 365.000 square meters and will consist of private and public housing as well as local businesses, with an emphasis on being a green and sustainable neighborhood.

With a centrally located new development, the new neighborhood will shift the urban dynamics in that part of the city by connecting the different surrounding neighborhoods. To understand how it will shift, we first need to know what the residential composition of the surrounding areas looks like. Therefore, in this paper, we explore the demographic and socioeconomic dynamics in the neighboring areas to Jernbanebyen. By constructing meaningful neighborhoods combined with register microdata, we can visualize relevant demographic and socioeconomic profiles of the surrounding neighborhoods.