Spatial inequalities in neighbourhood socioeconomic status
A case study of six large cities in the Netherlands

Master Thesis Population Studies

Student Zita Bastmeijer
S4122429

Supervisor Zuzana Zilincikova

Date July 12 2024

University University of Groningen
Faculty Faculty of Spatial Sciences
Department Demography

Abstract

This thesis examines the relation between the demographic composition of a neighbourhood and the mean socioeconomic status of the neighbourhood, additionally, the correlation between socioeconomic status and neighbourhood characteristics is examined. This is done for six large cities in the Netherlands; Enschede, Groningen, Eindhoven, Dordrecht, Amersfoort, and Alphen aan den Rijn. The demographic variables used in this study are population density, ethnicity, age composition, and household composition. The neighbourhood characteristics in the analysis are amenities, green spaces, housing, and safety. Theory suggests that certain groups in the population are more likely to live in low SES neighbourhoods. This includes migrants and single people. Socioeconomic segregation, the clustering of people with similar socioeconomic statuses into specific neighbourhoods, largely overlaps with segregation based on demographic characteristics. Residents of a neighbourhood with a low average SES are often worse off than those of high SES neighbourhoods. The multiple linear regression analysis showed that the relation between the share of inhabitants with a non-western migration background and the SES of the neighbourhood is significant in all cities. At the same time the household composition variables are only significant in Enschede and Groningen, the cities with the lowest mean SES. The Pearson's correlation analysis showed that the correlation between mean house value and neighbourhood SES is strongest in all cities, followed by the safety variables, which have negative coefficients. The correlation coefficients are not significantly different between the cities in most cases. Further research could include more cities and use longitudinal data, to get a deeper understanding of the relations between the demographic composition of a neighbourhood, the neighbourhood characteristics, and socioeconomic status.

Keywords: Socioeconomic status, Neighbourhood effects, Demography, Spatial inequality, Residential segregation, Socioeconomic segregation, Netherlands

Table of contents

List of tables and figures

List of abbreviations

CBS - Centraal Bureau voor de Statistiek (Central Statistical Office of the Netherlands)
SES - socioeconomic status

1. Introduction

‘Stop the accumulation of problems in disadvantaged neighbourhoods’ is what twelve directors of the National Programm for Liveability and Safety call for in the newspapers (Schaapman et al., 2024). They argue that problems such as poverty, low quality of education, nuisance, and crime are unequally distributed across neighbourhoods, with neighbourhoods with very low liveability as a result. According to Leidelmeijer et al. (2023), there is a positive correlation between the socioeconomic status of inhabitants of a neighbourhood and the liveability of this neighbourhood. This indicates that people with higher socioeconomic statuses live in the most liveable neighbourhoods, and the other way around. This phenomenon is called residential segregation, or socioeconomic segregation. This form of spatial inequality is growing in the Netherlands as a whole (Boterman et al., 2021), although the welfare state provides some buffer against this, mainly against the income-inequality share. According to Leidelmeijer et al. (2023), despite efforts to decrease this, the strong inequality between neighbourhoods has not decreased between 2014 and 2020, the divide between richer and poorer neighbourhoods in the Netherlands remains. This divide additionally influences the equality of opportunities, people with higher socioeconomic statuses live in more liveable neighbourhoods, which might be safer, greener, and have more amenities. There are several ways in which this divide can have arisen. People with higher socioeconomic statuses could have moved into these liveable neighbourhoods, liveable neighbourhoods could have been developed in places where people with high socioeconomic statuses lived, or neighbourhoods with favourable characteristics could have led to people increasing their socioeconomic status.

Ellen & Turner (1997) concluded that neighbourhood conditions matter in terms of individual outcomes, although not as much as family characteristics. Demographic characteristics can have significant effects on the socioeconomic status of individuals. Simultaneously, neighbourhoods with a low socioeconomic status have different demographic compositions than neighbourhoods with a high socioeconomic status. Certain populations are more often exposed to low socioeconomic status neighbourhoods. Demographic characteristics like ethnicity and household composition can therefore contribute to the existing residential segregation (Coenen et al., 2019). Still, neighbourhood characteristics are often easier to influence, and therefore interesting to policymakers who focus on topics like social and spatial inequality. The clustering of people with low socioeconomic statuses in low socioeconomic status neighbourhoods comes with the risk of reproducing inequality.

These topics are not only relevant to policymakers, but also academically. While there is increasing attention to residential and socioeconomic inequalities, most research on segregation focuses on racial or ethnic segregation, and residential or socioeconomic segregation is strongly understudied. Therefore, this research contributes to filling in this research gap. Additionally, most research on these topics is situated in the United States, this research tries to give insight into the Dutch context.

Rather than focussing on liveability, which can be a rather abstract and subjective concept, often measured with factors relating to quality of life, this research aims to find out how different specific demographic and neighbourhood characteristics relate to the socioeconomic status of a neighbourhood. Additionally, the aim is to compare the differences and similarities of these relations between six cities in the Netherlands, in order to get insight into whether these are location-sensitive or not.

To reach the aim of the research the following research question was formulated: *'To what extent is the demographic composition of the inhabitants related to neighbourhood socioeconomic status and to what extent is there a correlation between socioeconomic status and neighbourhood characteristics in six large cities in the Netherlands?'*. The four subquestions used to answer this research question are: *"Which demographic characteristics (population density, ethnicity, age composition, household composition) are related to neighbourhood socioeconomic status?', 'What are the differences and similarities in this relationship between different cities?', 'Which neighbourhood characteristics (amenities, green spaces, housing, safety) correlate with socioeconomic status?'* and *'What are the differences and similarities in these correlations between different cities?'*. Demographic characteristics are measured through population density, ethnicity, age composition, and household composition. Neighbourhood characteristics are measured through amenities, green spaces, housing, and safety. To answer the research questions a quantitative analysis was performed using regression for the demographic characteristics and correlation for the neighbourhood characteristics. Data was retrieved from the CBS (Central Statistical Office of the Netherlands). The data on socioeconomic status and the demographic and neighbourhood characteristics that is used is on the neighbourhood level (buurten in Dutch) rather than the individual level.

Studying demographic and neighbourhood characteristics in relation to socioeconomic status can give insight into the extent and spatial distribution of inequality between neighbourhoods, cities, or within society as a whole. These insights are vital for policymakers when deciding on policies that aim to reduce inequality. Comparing cities gives further insights, through general and city-specific patterns. The relation between characteristics and SES does not have to be equal across different cities and the relative importance of characteristics might differ per city as well.

In the following theoretical framework, the neighbourhood effects theory will be explored, together with the main demographic and neighbourhood characteristics and the concepts of socioeconomic status and residential segregation. Subsequently, the data and methodology of the study will be discussed extensively. Thereafter, the results and conclusions will be presented.

2. Theoretical framework

2.1 Theory

Neighbourhood effects theory is the study of how various neighbourhood conditions relate to collective or individual outcomes (Roosa & White, 2014). The main idea underlying the neighbourhood effects theory is that living in a deprived neighbourhood negatively affects the chances and opportunities of the residents. Some authors state that these effects are stronger than the residents' individual characteristics, while others say the opposite (Van Ham & Manley, 2010). A large share of the current popularity of the theory can be attributed to the book "The Truly Disadvantaged: The Inner City, The Underclass and Public Policy" by William Julius Wilson (Wilson, 1987). In this book, Wilson examines the effects of living in concentrations of poverty and concludes that social and economic forces on a macro scale shape local conditions in poor neighbourhoods (Darcy & Gwyther, 2011; Van Ham et al., 2011).

Although there is agreement among many scholars that neighbourhood effects exist, the causal mechanisms behind the effects of the residential context on individual outcomes are highly debated (Galster, 2011). Additionally, much is still unclear about the relative importance of neighbourhood characteristics among each other and the context dependency of these processes (Van Ham et al., 2011). Critics of neighbourhood effects theory suggest that there is too little evidence to state that (poor) residents' chances and opportunities are negatively affected by living in poorer or worse-off neighbourhoods, leaving out individual characteristics that are the underlying causes for poverty and low opportunities (Cheshire, 2007; Van Ham et al., 2011).

Research on neighbourhood effects and the neighbourhood effects theory often focuses on social or economic conditions, or characteristics, of a neighbourhood and their relation to individual or collective outcomes. Examples are the level of unemployment or mean income in a neighbourhood. This, however, is often causing the simultaneity problem, in which individuals are included in both the condition and the outcome (Manski, 1993; Van Ham et al., 2011). Neighbourhood effects research can also be performed using other characteristics that might negatively affect individual or collective outcomes, and that do not cause the simultaneity problem. Neighbourhood characteristics, and the deprivation thereof, are an example of this. Six characteristics that are widely used across neighbourhood research in general are amenities, cohesion, green spaces, housing, infrastructure, and safety (Galster, 2001). These characteristics all play an important role in an individual's daily life and can be important contributors to quality of life; living, moving through space, working, shopping, health, social interaction, and being safe. Additionally, these characteristics help to make the neighbourhood as a concept more objectively measurable and quantifiable. George Galster (2001) used these, along with some social and economic characteristics in his paper "On the Nature of Neighbourhood", to define the spatially based attributes relevant to defining a neighbourhood. Each of these characteristics has been correlated to the socioeconomic status of a neighbourhood and both the neighbourhood characteristics and socioeconomic status can thus negatively impact individuals.

Socioeconomic status is one of the concepts that are used to measure the individual or collective outcomes within neighbourhood effects theory (Van Ham et al., 2011). Socioeconomic status (SES) is a measure that is often used to study inequality. It is usually based on an individual's income, education, and employment (history). While it is often used on an individual scale, it can also be used to analyze groups of people, or for instance neighbourhoods. SES is often used instead of income since it gives a more

complete picture of someone's relative position in society, including access to opportunities and resources, which is a vital part of the neighbourhood effects theory (Andersson & Mustard, 2010; Iceland & Wilkes, 2006). Significant inequalities exist between people with different socioeconomic statuses, people with a lower socioeconomic status often have a shorter life expectancy and poorer health (Aamodt et al., 2023; Mackenbach et al., 2008).

Residential segregation, or socioeconomic segregation, entails the separation of specific groups of people into separate areas, in the case of this research, people with similar socioeconomic statuses clustering in specific neighbourhoods. Residential segregation can increase spatial inequality and has been increasing over the last decades (Boterman et al., 2021). Social mixing policies are often implemented to decrease residential segregation and spatial inequality and create a mix of more diverse socio-economic groups (Van Ham et al., 2011). For instance, instead of only building expensive new houses in a neighbourhood where many people live with higher socioeconomic statuses, these policies can force developers to create more affordable housing in a neighbourhood. The potential of negative neighbourhood effects should be removed, or at least decreased by this (Musterd & Andersson, 2005). However, social mix policies can also have the opposite effect and stimulate gentrification processes and increase social inequality (Boterman et al., 2021).

There are multiple theories that substantiate the relationship between the demographic composition of a neighbourhood and its mean socioeconomic status. Within human capital theory, for instance, it can be stated that the age and household composition of a neighbourhood relate to the mean human capital, and thus socioeconomic status (Klevmarken & Quigley, 1976; Simpson et al., 2002). Following the life course approach, life events such as moving in together or having children can influence outcomes like socioeconomic status (Roode et al., 2017). Therefore, changes in the household composition of a neighbourhood will also influence the mean socioeconomic status of the neighbourhood.

The mean socioeconomic status of a neighbourhood can thus be linked to different characteristics of the population. Specific groups of the population are more likely to live in low socioeconomic status neighbourhoods, this includes single-parent families (Flouri et al., 2016) and migrants (Leclerc, 2022; Vaalavuo et al., 2019). As a result, socioeconomic segregation often overlaps with segregation based on other demographic characteristics. The most obvious example is ethnic segregation or segregation based on migration status. On average, migrants have lower incomes than natives, consequently, ethnic minorities or people with a migration background more often live in deprived neighbourhoods (Van Ham et al., 2021; Vogiazides & Chihaya, 2020). Another type of segregation that often overlaps with socioeconomic segregation is age segregation, this is caused for a large share by housing affordability (Sabater & Finney, 2023). Younger people move to more affordable housing, in neighbourhoods with a lower mean socioeconomic status (Lau, 2023).

2.2 Literature review

In the following paragraphs, literature will be discussed that has investigated socioeconomic status in relation to either demographic or neighbourhood characteristics. These studies all focused on one of the variables used in this paper and its relation to either individual or collective socioeconomic status. Most of these authors focused on the United States, Australia, or countries in Europe.

The relationship between demographic characteristics and socioeconomic status has been investigated in many different ways. Meijer et al. (2012) state that high population density is often related to lower socioeconomic status in Denmark. Their study found that living in an urban (high population density) low socioeconomic neighbourhood increases mortality. The relationship between ethnicity and neighbourhood income in the UK was examined by Coulter and Clark (2019), who state that ethnic minorities tend to live in less advantaged neighbourhoods. However, they also state that income is a more important determinant, people with more resources move to more advantaged neighbourhoods, regardless of ethnicity. A positive association between education and income and healthy ageing was found by Wagg et al. (2021) in their review of multiple cross-sectional and longitudinal studies. This means that people with more education and a higher income (and thus a higher socioeconomic status) live longer and healthier lives. This could indicate that a neighbourhood with a larger share of older inhabitants has a higher socioeconomic status. However, socioeconomic status is expected to decrease after retirement age, mainly due to decreasing income. When it comes to household composition, one person households are associated with lower socioeconomic status, compared to couples, mainly due to couples often having a double income. Between households with and without children, those without children often invest more time in education and employment and therefore have higher socioeconomic status in the US (Karney, 2021).

There are some studies that relate socioeconomic status to neighbourhoods and neighbourhood characteristics. Mackenbach et al. (2016), for instance, studied why neighbourhood perceptions of residents of socioeconomically deprived neighbourhoods were less positive than those of residents of wealthier neighbourhoods in Belgium, France, Hungary, the Netherlands, and the UK. Neighbourhood social cohesion explained the largest part of this difference, 52%, followed by traffic safety, aesthetics, and the presence of destinations, explaining 15% of differences in neighbourhood perception (Mackenbach et al., 2016). Social cohesion thus might be an important determinant of the lower neighbourhood perception of residents in socioeconomically deprived neighbourhoods. Erdem et al. (2016) state that the negative effects of low socioeconomic status in the Netherlands can be partly resolved by living in a neighbourhood with high social cohesion, since the distress caused by a lack of income or other financial problems can be reduced. Therefore, higher social cohesion might have a positive effect on subjective socioeconomic status, and increase mental wellbeing. At the same time, living in a neighbourhood with a higher average socioeconomic status increases the likelihood of social cohesion.

In Melbourne, Australia, more amenities are found in neighbourhoods with higher socioeconomic statuses, according to Crawford et al. (2008), although they mainly talk about recreational facilities. Their conclusion that neighbourhoods with a higher socioeconomic status are better off in terms of amenities is supported by Altschuler et al. (2004) and by Powell et al. (2007) when talking about supermarkets in the United States. A lack of supermarkets in low socioeconomic neighbourhoods leads to a lack of healthy, fresh, and lower-priced food. While not all higher socioeconomic neighbourhoods have supermarkets in the neighbourhood, more residents of these neighbourhoods have access to cars, while residents from lower socioeconomic neighbourhoods often have to rely on public transport (Altschuler et al., 2004).

Viinikka et al. (2023) examined the relationship between green spaces and socioeconomic status in Finland, they concluded that areas with a lower socioeconomic status had more access to green spaces

with recreational facilities and routes, while residents of areas with a higher socioeconomic status had more access to larger green spaces and forests. Schüle et al. (2017) state that earlier evidence does not suggest a clear pattern in the correlation between low socioeconomic areas and a lack of environmental resources. Their analysis of a large city in Germany however showed that when including a buffer of 200 meters, low socioeconomic status neighbourhoods had on average 43% less green space, compared to high socioeconomic status neighbourhoods. They concluded that green space availability within and around the neighbourhood was decreasing for neighbourhoods with low socioeconomic status, thus increasing spatial inequalities (Schüle et al., 2017). Many other studies in different contexts like Australia and the United States also found that the availability of green spaces is lower for residents of lower income or SES neighbourhoods (Astell-Burt et al., 2014; Rigolon, 2016).

For housing, the correlation with socioeconomic status is more straightforward. People with higher incomes and education, and thus higher socioeconomic status, are more likely to live in, and buy, more expensive houses in the United States (Li & Wei, 2020). This is therefore one of the main drivers of residential segregation. Like for the other neighbourhood characteristics, people with higher socioeconomic status are better off and have more opportunities when it comes to housing. According to Kang & Seo (2022), subjective socioeconomic status in South Korea is determined for a large part by housing characteristics like home ownership, the type of housing someone lives in and the type of housing in the neighbourhood. This indicates a strong correlation between housing and socioeconomic status. Additionally, across Europe and the United States, housing costs are often the main reason for people to choose a specific location or neighbourhood to live in and people often prefer to live near other people with similar socioeconomic statuses (Schirmer et al., 2014).

Sugiyama et al. (2015) show that low socioeconomic status neighbourhoods have lower scores for walking infrastructure than high socioeconomic status neighbourhoods in Southern Australia. Contrastingly, Cowie et al. (2016) could not find any linear trends between walkability and road density and neighbourhood socioeconomic status in the same country. In theory, residents of low socioeconomic status neighbourhoods are more likely to experience transport disadvantages. However, Rachele et al. (2017) did not find this in their research on Brisbane, Australia. On the contrary, they found that these disadvantaged neighbourhoods had higher connectivity and transit access. However, these neighbourhoods also experienced more exposure to traffic.

Crime rates are not equal across neighbourhoods, there is often significant spatial clustering, as found in research on London by Zhou et al. (2023). Van Wilsem et al. (2006) and Sugiyama et al. (2015) agree that the chance of becoming a victim of crime is larger in disadvantaged, lower SES neighbourhoods in the Netherlands and Australia respectively. However, Van Wilsem et al. (2006) found that this risk is also higher in neighbourhoods that are changing in terms of socioeconomic composition, for instance through gentrification. Therefore, in terms of safety, the dynamics in socioeconomic composition might be just as important as the average SES.

All these findings and conclusions give some insight into the possible relations between demographic and neighbourhood characteristics and socioeconomic status. However, it has to be taken into account that these researches were mostly conducted in different contexts than the Netherlands. Therefore, the results might not translate into the Dutch context. Additionally, these researches focussed on the relationship or

correlation between one of the neighbourhood characteristics and SES, rather than looking at multiple characteristics. This research will give insight into the relative importance of the different demographic characteristics in relation to SES, and the correlations between neighbourhood characteristics and SES, in the Dutch context.

2.3 Conceptual model

Figure 1 shows the conceptual model for this study. It shows the main concepts as discussed in the theoretical framework and their relationships. On the left are the demographic characteristics, measured at the neighbourhood level, these are assumed to contribute to the neighbourhood's socioeconomic status and thus the arrows go in one direction. Four out of the six neighbourhood characteristics (Galster, 2001) are used in the statistical analysis, due to unavailable data for cohesion, and infrastructure being measured only on larger scales than neighbourhoods. The arrows show that the relationships between neighbourhood characteristics and socioeconomic status can go both directions. For instance, more expensive houses in a neighbourhood attract people with higher socioeconomic statuses, while on the other hand, new houses built in a neighbourhood where many people live with higher socioeconomic statuses will most likely also be more expensive houses.

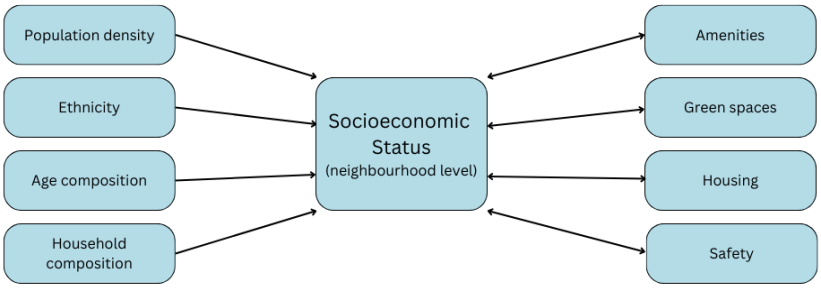

Figure 1: Conceptual model

2.4 Hypotheses and expectations

Based on the theory and existing literature the following hypotheses will be tested for the demographic characteristics.

H1: The population density of a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood.

Earlier research suggests that high population density is associated with low socioeconomic status. Therefore, the population density of a neighbourhood is expected to be negatively associated with the socioeconomic status of the neighbourhood.

H2: The share of inhabitants with a migration background in a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood.

Inhabitants with a migration background on average have lower incomes than natives, resulting in lower socioeconomic status. As a result, people with a migration background have a higher chance to live in a deprived neighbourhood. Based on this knowledge it is hypothesized that a larger share of inhabitants with a migration background is negatively associated with socioeconomic status.

H3: The share of older inhabitants in a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood.
Although people with more education and income live longer and healthier lives, after retirement age socioeconomic status is expected to decrease, according to existing research. Thus, the share of older inhabitants is expected to be negatively associated with socioeconomic status.

H4.1: The share of one person households in a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood.
H4.2: The share of households without children in a neighbourhood is positively associated with the socioeconomic status of the neighbourhood.
H4.3: The share of households with children in a neighbourhood is positively associated with the socioeconomic status of the neighbourhood.
Past research has shown that one person households often have the lowest socioeconomic status, followed by households with children. Households without children in many cases invest more time in education and employment. A negative association is therefore hypothesized for the share of one person households in a neighbourhood. The share of households with children and without children are both hypothesized to give a positive association, since these often have a double income, leading to a higher socioeconomic status.

From the theory and previous studies on different neighbourhood characteristics, it can be hypothesized that residents of neighbourhoods with a low average SES are worse off compared to those of high average SES neighbourhoods, in terms of accessibility of amenities and green spaces, the value of houses, and the number of crimes. Living in these disadvantaged neighbourhoods negatively affects the resident's chances and opportunities. This general pattern is expected in all cities.

H5: The number of accessible amenities (per 1000 inhabitants) of a neighbourhood is negatively correlated with the socioeconomic status of the neighbourhood.
The scale of neighbourhoods in the Netherlands is relatively small and amenities in other neighbourhoods are usually very easily accessible, therefore this correlation could be weak. It is difficult to predict whether the correlation between amenities and socioeconomic status will be positive or negative. More amenities in a neighbourhood might relate to a higher socioeconomic status of the neighbourhood, however, previous studies also showed that high SES neighbourhoods do not always have more amenities. Since most amenities are usually located in city centres, which often have lower socioeconomic status, a negative correlation is hypothesized.

H6: The hectares of green space (per 1000 inhabitants) of a neighbourhood is positively correlated with the socioeconomic status of the neighbourhood.
Although not all scholars agree, most existing studies found that lower SES neighbourhoods have lower availability of green spaces. Therefore, green spaces are expected to have a positive correlation with socioeconomic status, which means that neighbourhoods with more green spaces are expected to have a higher average socioeconomic status.

H7: The mean house value (x1000 euro) of a neighbourhood is positively correlated with the socioeconomic status of the neighbourhood.

Housing is measured as the average home value in the neighbourhood. This variable is expected to have the strongest positive correlation with socioeconomic status. Since theory shows that housing is often the main reason for people to choose to live in a certain neighbourhood. This can cause, or increase, residential segregation, mainly when neighbourhoods have a very homogeneous housing supply. Furthermore, it can be hypothesized with high confidence that neighbourhoods with a high average home value will have a higher average SES, since people with lower SES will not be able to afford these houses. This is less clear for neighbourhoods with lower average housing prices, however, the correlation is still expected to be positive in all cities.

H8: The number of crimes (per 1000 inhabitants) in a neighbourhood is negatively correlated with the socioeconomic status of a neighbourhood.

Theory and earlier research suggest that crime rates are higher in neighbourhoods with a lower socioeconomic status. Therefore, as safety is measured by the number of registered crimes, a negative correlation between this variable and socioeconomic status is expected. This entails that a neighbourhood with many registered crimes is expected to have a low SES, and a neighbourhood with a high SES is expected to have few registered crimes.

3. Data and methodology

3.1 Data and study design

The aim of the analysis is to answer the research questions. To do this, data from the CBS is used. This is open-source data, which is freely available and of high quality. Since not all used variables are published within the same dataset, datasets were combined to create a dataset for the analysis. All the used data is on the neighbourhood level, in the data there is a distinction between two types of neighbourhoods; wijken and buurten. Wijken are larger areas, more like districts, and each wijk includes several buurten. This analysis focusses on buurten. Since this is aggregated data rather than individual data, there are no issues with confidentiality. The analysis focuses on the year 2017, since this is the most recent year for which data is available for all variables.

In order to decide which cities to include in the analysis all municipalities in the Netherlands were sorted based on the number of inhabitants. Thereafter, all municipalities with over 100,000 inhabitants were sorted based on the SES score. These municipalities were then divided into low, medium and high SES, as can be seen in Table 1. From this, two cities from each category were chosen to be included in the analysis. Figure 2 shows the distribution of these cities across the Netherlands.

Table 1: Municipalities with over 100,000 inhabitants, sorted according to SES

Municipality name	Number of inhabitants	SES	Category	Difference
Rotterdam	634660	-0,301	Low	
Enschede	158140	-0,275	Low	0,026
Groningen	202636	-0,271	Low	0,004
Maastricht	122753	-0,255	Low	0,016
Delft	101381	-0,253	Low	0,002
Nijmegen	173557	-0,215	Low	0,038
Arnhem	155699	-0,21	Low	0,005
's-Gravenhage	524882	-0,194	Low	0,016
Leeuwarden	108667	-0,184	Low	0,01
Emmen	107490	-0,149	Medium	0,035
Tilburg	213804	-0,142	Medium	0,007
Amsterdam	844947	-0,139	Medium	0,003
Leiden	123661	-0,122	Medium	0,017
Venlo	101059	-0,12	Medium	0,002
Eindhoven	226868	-0,116	Medium	0,004
Dordrecht	118731	-0,094	Medium	0,022
Utrecht	343038	-0,056	Medium	0,038
Zaanstad	153679	-0,056	Medium	0
Breda	182304	-0,008	High	0,048
Zwolle	125548	-0,006	High	0,002
Almere	200914	-0,003	High	0,003
Alkmaar	108373	0,002	High	0,005
's-Hertogenbosch	152410	0,015	High	0,013
Apeldoorn	160047	0,02	High	0,005
Zoetermeer	124763	0,031	High	0,011
Haarlem	159229	0,058	High	0,027
Ede	113421	0,08	High	0,022
Amersfoort	154337	0,094	High	0,014
Alphen aan den Rijn	108915	0,124	High	0,03
Westland	105632	0,135	High	0,011
Haarlemmermeer	146003	0,173	High	0,038

Figure 2: Map of the six cities in the analysis

3.2 Study population

The study population in this analysis includes neighbourhoods of six large cities in the Netherlands. These cities are Enschede, Groningen, Eindhoven, Dordrecht, Amersfoort, and Alphen aan den Rijn. Unfortunately, not all of the neighbourhoods could be included in the analysis, therefore the sample includes all neighbourhoods in these cities with available data. In general, this means all neighbourhoods with a minimum population of 300 inhabitants. Most neighbourhoods with a smaller population do not have a calculated score for the neighbourhood SES. Additionally, for the neighbourhoods with less than 300 inhabitants that do have a calculated score, this score might be less strong and representable. In total, 26.2 per cent of the neighbourhoods were dropped, leaving 490 neighbourhoods in the study. Of the neighbourhoods that remain, the mean number of inhabitants is 1947, the smallest neighbourhood in the study has 320 inhabitants and is located in Dordrecht, and the largest neighbourhood in the study is located in Groningen and has 8480 inhabitants.

3.3 Operationalisation

The analysis includes nine different variables; socioeconomic status on a neighbourhood level, four demographic variables on a neighbourhood level and four neighbourhood characteristics variables. The SES variable is made up of the mean income, education, and employment history of all residents of the neighbourhood. The CBS first calculates the SES score per household and uses this to calculate the mean SES score per neighbourhood. In the calculation of the SES score per household, income is measured in ten categories, education is measured in five categories, and employment history is measured in seven categories. Each of these categories has been given a specific value, for each household the values for the categories that they belong to are added, which leads to the SES score. The average SES score for the whole Netherlands is zero in 2019 and very close to zero for 2014-2018 and 2020-2021 (CBS, n.d.). The

demographic variables on a neighbourhood level include the population density, ethnicity, the age composition, and the household composition. The neighbourhood characteristics included in the analysis are amenities, green spaces, housing, and safety. The variables ethnicity, age composition, household composition, amenities, and safety have been subdivided into more specific categories to get a deeper understanding of the associations and correlations. All (sub)variables and their operationalisation can be seen in Table 2.

Table 2: Variables and operationalisation

Variable	Operationalisation
SES	SES-WOA mean score
Population Density	Population density (inhabitants/km^2)
West Mig Background	Ethnicity - Share of inhabitants with a western migration background (%)
Non-West Mig Background	Ethnicity - Share of inhabitants with a non-western migration background (%)
Age 0-14	Share of inhabitants aged 0 to 14 (%)
Age 15-24	Share of inhabitants aged 15 to 24 (%)
Age 25-44	Share of inhabitants aged 25 to 44 (%)
Age 45-64	Share of inhabitants aged 45 to 64 (%)
Age 65+	Share of inhabitants aged 65 or older (%)
One Person Households	Share of one person households (%)
Households Without Child	Share of households without children (%)
Households With Children	Share of households with children (%)
Health Wellbeing	Amenities - Health and wellbeing (per 1000 inhabitants)
Retail	Amenities - Retail (per 1000 inhabitants)
Catering Industry	Amenities - Catering industry (per 1000 inhabitants)
Childcare Education	Amenities - Childcare and education (per 1000 inhabitants)
Leisure Culture	Amenities - Leisure and culture (per 1000 inhabitants)
Green Spaces	Hectares of green space (per 1000 inhabitants)
House Value	Mean house value (x1000 euro)
Theft Crimes	Crimes - theft (per 1000 inhabitants)
Vandalism & Public Order Crimes	Crimes - vandalism, crimes against public order (per 1000 inhabitants)
Violent & Sexual Crimes	Crimes - violent and sexual crimes (per 1000 inhabitants)

3.4 Data analysis

Visualization techniques using ArcGIS Online are used to make maps of the cities that show the spatial distribution of SES scores per neighbourhood. The statistical analysis consists of two parts, a regression analysis with the demographic variables and SES, and a correlation analysis with the neighbourhood characteristics and SES. Both analyses are done in Stata. The Stata commands can be found in Appendix 1.

The regression analysis shows the relationship between the population of the neighbourhoods in relation to the neighbourhood average SES. Since there is one continuous dependent variable and multiple continuous independent variables, a multiple linear regression was used. The dependent variable is SES, and the independent variables are population density, ethnicity (in two subcategories), age composition, and household composition (in three subcategories).

The specification of the model is expressed as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \varepsilon$$

Where y is the dependent variable neighbourhood SES, β_0 is the intercept, β_1,..., β_7 are the regression coefficients for x_1,..., x_7, the seven independent variables, and ε is the residual error.

Since the direction of the relationship between neighbourhood characteristics and socioeconomic status is not straightforward, meaning it could work both ways, it is difficult to perform a regression analysis, for which dependent and independent variables need to be specified. Therefore, a correlation analysis was used to analyse the relationship between neighbourhood characteristics and socioeconomic status. The correlation analysis shows something about the context or environment that the inhabitants are living in. The Pearson's correlation coefficient measures the strength of the linear relationship between continuous variables, it also measures the direction of the relationship, ranging from a strong negative to a strong positive correlation. The values of these correlations are between -1 and 1. Correlation coefficients between 0 and 0.1 indicate a very weak correlation, between 0.1 and 0.4 indicates a weak correlation, between 0.4 and 0.7 indicates a moderate correlation, and coefficients between 0.7 and 1 mean that variables have a strong correlation (Costa & Neves, 2020).

3.5 Ethical considerations

The data was retrieved from the CBS, this is open-source data that is freely available online and for everyone. With the CBS being the source of the data, it can be assumed that the data is trustworthy. Since the data is all on a neighbourhood level, individuals can not be identified and the risk of violating confidentiality is non-existent. For very small neighbourhoods, with few inhabitants, the data is often not available, and therefore these neighbourhoods were not used in the analysis. The research is done objectively and the researcher's positionality did not influence the outcomes and conclusions.

3.6 GenAI use

Generative AI such as ChatGPT was exclusively used for ‘general functionalities’, such as gaining inspiration or spelling and grammar checks, following the guidelines of the university. GenAI was not used to produce text and inspiration from GenAI was never copied and used in the thesis.

4. Results

4.1 Descriptive analysis

Before going into the results of the regression and correlation analyses, the descriptive results will be discussed to get a better understanding of the data and the variables. Table 3 shows that the mean SES per neighbourhood is lowest in Enschede on the left side of the table, and increases to Alphen aan den Rijn on the right side of the table. The values for 'all cities' are calculated using the neighbourhoods from all six cities. The mean SES for 'All cities' (-.004) is very close to zero, this corresponds to the nature of the variable, since the average SES score for the whole Netherlands is zero in 2019 and very close to zero for 2017 (CBS, n.d.). Thus, based on the SES score, these six cities are a good reflection of the Netherlands. Figure 3 shows the distribution of the SES scores of all neighbourhoods in the study, a total of 490 cases. The data is mostly normally distributed, with some negative outliers.

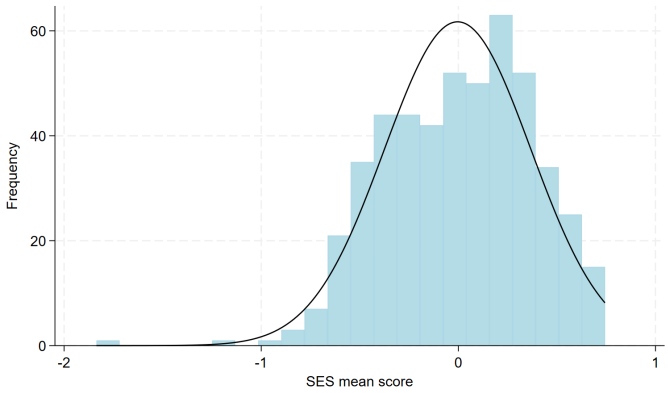

Figure 3: Distribution of SES scores

Figure 4 shows the maps that were created to get a better understanding of the spatial distribution of the neighbourhood SES scores. The large neighbourhoods on the edges of most cities give a bit of a distorted view, the size of neighbourhoods is not part of the SES calculation. The maps show that the cities with a lower mean SES per neighbourhood have a larger share of red neighbourhoods, while the cities with a higher mean SES per neighbourhood have a larger share of (dark) green neighbourhoods. The neighbourhoods with lower SES are concentrated in and around the city centres for most cities in the analysis. For Eindhoven this distribution is not as clearly visible, the higher and lower SES neighbourhoods seem to be distributed throughout the city.

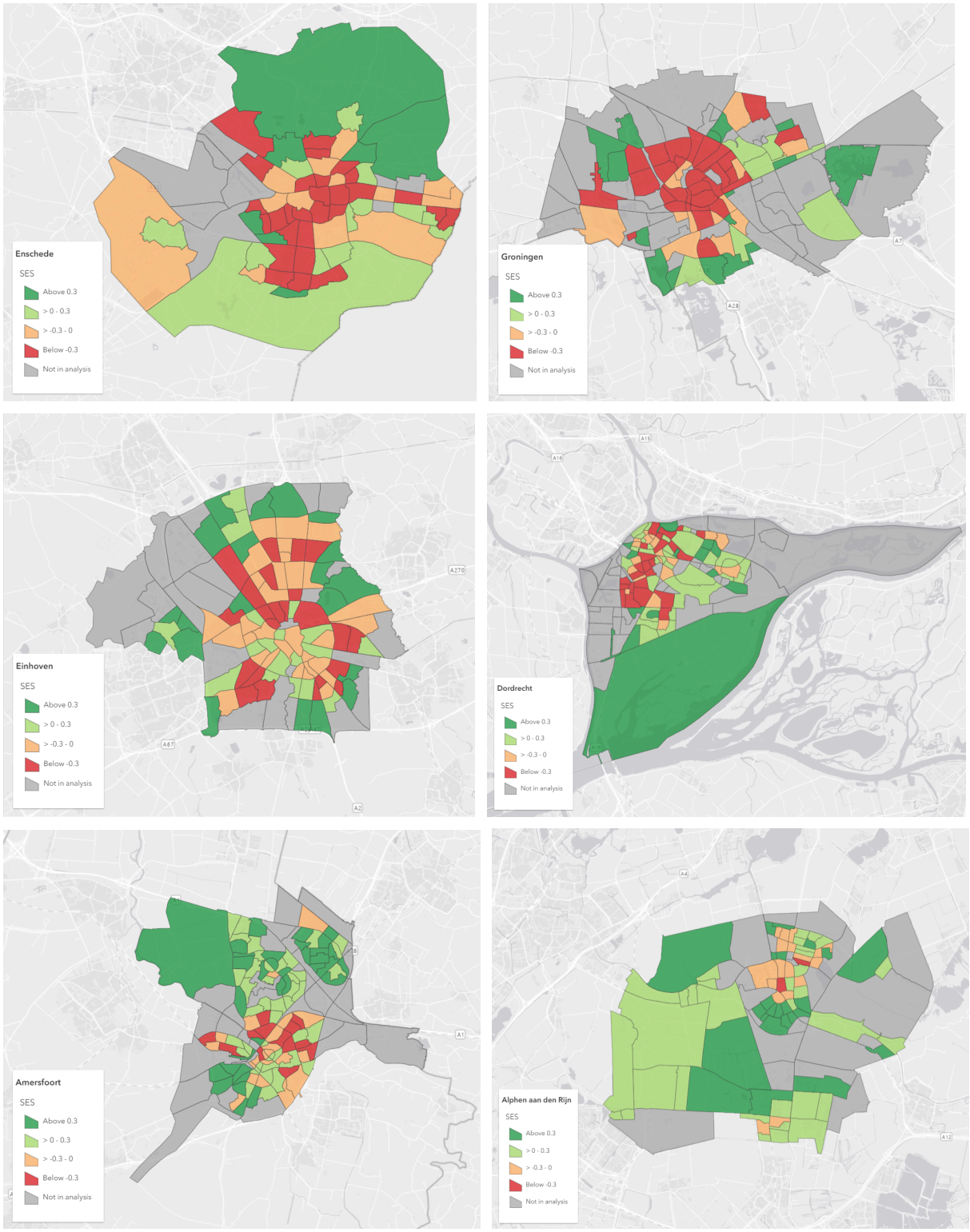

Figure 4: Maps of the SES distribution in the six cities (Author, 2024, CBS data)

As can be seen in Table 3, population density in neighbourhoods is highest in Dordrecht, followed by Groningen. The lowest population density is seen in Enschede and Eindhoven. Dordrecht and Groningen have high standard deviations, and therefore a lot of variation in population density between neighbourhoods.

Groningen excluded, all cities have on average a larger share of people with a non-western migration background per neighbourhood, compared to people with a western migration background. The mean percentage of people with a western migration background living in a neighbourhood ranges from 8.7% to 14.2% between the cities. For people with a non-western migration background, this ranges from 9.1% to 17.8%. For both variables, the lowest mean is seen in Alphen aan den Rijn. The highest means are seen in Eindhoven for both variables. Standard deviations are higher for the percentage of people with a non-western migration background in all models, showing that there is a larger variation between neighbourhoods.

The age composition variables show that Groningen and Eindhoven have younger populations. For these cities, the mean is highest for the share of inhabitants between 25 and 44, while for the other cities, the highest mean is for the share of inhabitants between 45 and 64. The highest mean for the share of inhabitants over 65 years old is seen in Alphen aan den Rijn, with an average of 18.9 per cent of the population of a neighbourhood being this age. The lowest mean for the share of inhabitants over 65 years old is seen in Groningen, with a mean of 12.5 per cent, followed by Amersfoort (15.4%) and Eindhoven (17%). Standard deviations for all age groups and for all cities are relatively low, which indicates that there is little variation between neighbourhoods when it comes to the age composition.

When it comes to household composition, for most cities the mean is highest for the share of one person households, neighbourhoods in Amersfoort and Alphen aan den Rijn on average have a higher share of households with children, these are the cities with the highest mean SES. This could indicate that the share of households with children is positively related to SES, the regression analysis will give more insight into this. In most cities, the share of households without children is the lowest. This variable also has the lowest standard deviations, and thus the least variation between neighbourhoods.

As could be expected, health and well-being amenities are the least present in all cities. The number of catering amenities per 1000 inhabitants gives the largest mean in all cities, however with large differences between these means. Ranging from an average of 8.9 catering amenities per 1000 inhabitants in Alphen aan den Rijn to 52.8 catering amenities per 1000 inhabitants in Dordrecht. The standard deviations for catering amenities are significantly higher than the standard deviations for the other amenities, showing large variation between neighbourhoods. This could be caused by neighbourhoods in city centres that often have a lot more catering amenities but not necessarily more inhabitants.

The hectares of green space per 1000 inhabitants is much higher on average in Enschede and Dordrecht, with mean values of 44.3 and 34.7 respectively. For the other cities, the mean values range from 1.4 (Amersfoort) to 8.9 (Eindhoven). The high means for Enschede and Dordrecht could be caused by neighbourhoods with many hectares of green space and very few inhabitants. The maps in Figure 4 show that these cities have very large neighbourhoods on the edges of the city, these might have many hectares of green. This could also explain the high standard deviations for these two cities.

Amersfoort has the highest mean house value, while Dordrecht has the lowest. Alphen aan den Rijn has rather little variation in mean house prices between neighbourhoods, while Eindhoven has the most variation in mean house prices between neighbourhoods. Enschede and Groningen, the cities with the lowest mean SES have the second and third-lowest mean house values. Alphen aan den Rijn has the second-highest mean house value. This already shows a possible correlation between mean house value and SES, which will be further analyzed with the correlation analysis.

For most cities, the number of violent and sexual crimes per 1000 inhabitants is on average higher than 'thefts' and 'vandalism & public order crimes' per 1000 inhabitants. The mean for the number of thefts per 1000 inhabitants is the lowest out of the crime variables in all models. These low numbers might be due to under-registration of crimes that are considered less serious. At first sight, the descriptive statistics do not show a clear correlation between the mean number of crimes per 1000 inhabitants and the mean SES of the cities. Although the mean values for all three crime variables are lowest for Alphen aan den Rijn, the city with the highest mean SES.

Table 3: Descriptive statistics

Variable	All cities (n=490)		Enschede (n=61)		Groningen (n=66)		Eindhoven (n=92)		Dordrecht (n=87)		Amersfoort (n=107)		Alphen a/d Rijn (n=77)	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
SES	-.004	.371	-.174	.364	-.132	.453	-.048	.395	-.043	.368	.124	.294	.162	.233
PopDensity	5966.3	3309	4192	2477	6909.3	4389.8	5041.2	2572.1	7464.7	3459.7	6580	2450.8	5123	3335.7
WestMig Background	11.1	3.9	11.9	2.8	11.9	5.4	14.2	3.6	11.3	3.2	9.2	2	8.7	3.2
NonWestMig Background	14.1	10.8	13.5	10.2	10.1	6	17.8	9.8	17.6	13.1	14.3	11.8	9.1	7.8
Age 0-14	15.8	5.8	15.7	4.5	13.3	7.5	13.9	5.5	15.9	4.5	18.3	6.5	16.8	4
Age 15-24	14	7.3	14.9	7.9	21.4	12.9	13.8	6.8	12	3.4	12.1	3.2	11.8	2.3
Age 25-44	27.4	8.5	25.5	6	29.6	8.5	31	10.7	26.4	7.6	27.1	9	24	5.2
Age 45-64	26.7	6.7	26.3	6.6	23.1	8.2	24.1	6.4	28.5	6.1	28.3	6.1	29	4.6
Age 65+	16.4	8.6	17.1	6.1	12.5	7.7	17	8.9	17.3	9.4	15.4	9.3	18.9	7.3
Oneperson Households	40.2	17.5	39.3	15.6	51.6	22.9	46.8	17.5	39.2	15.4	35.5	14.2	31.2	11.7
Households WithoutChild	27.6	7.5	27.5	6.6	23.9	7.8	27.2	8.9	27.8	6.8	26.3	6.2	32.6	5.6
Households WithChildren	32.6	14.5	32.5	11.7	24.3	17.2	26.7	12.2	34.9	13.4	37.1	15.2	38	10.7

Table 3: Descriptive statistics continued

Variable	All cities (n=490)		Enschede (n=61)		Groningen (n=66)		Eindhoven (n=92)		Dordrecht (n=87)		Amersfoort (n=107)		Alphen a/d Rijn (n=77)	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Health Wellbeing	3.1	3.1	2.2	1.9	3.2	3.9	3.2	2.6	4.5	3.6	3.3	3.4	1.9	1.4
Retail	12.4	19.4	5.5	5.4	9.4	18.9	9.1	8.2	23	27.7	16.6	24.7	6.1	6.6
Catering	31.5	65	10.1	18.3	34.1	67.2	26.6	26.6	52.8	84.9	45.5	93.4	8.9	13.9
Childcare Education	11.5	12.2	5.5	4.3	8.8	10.6	7.4	6.2	20.5	17.9	15.6	11.9	7.8	7.1
Leisure Culture	8.7	8.2	6.8	6	8.9	9.9	10.4	9.4	12	9	8.4	7.5	4.6	2.9
GreenSpaces	14.3	139.3	44.3	153	2.5	4.5	8.9	24.8	34.7	303	1.4	2.8	2.3	3.4
HouseValue	214.3	97.5	188.5	88.3	183.3	84.9	232.4	112.8	174.7	80.3	250.7	106.4	234	68.4
TheftCrimes	4	3.4	4.3	3	3.7	2.3	6.3	4.3	3.5	3.7	4.2	3.1	1.8	1.6
Vandalism& PublicOrder Crimes	5.8	5.8	5.1	3.4	7.8	5.3	6.6	5.8	6.5	7.1	5.6	6.6	3.3	3.1
Violent& SexualCrimes	6.5	14.4	5.2	8.8	7.5	13.8	10.1	25.1	7.5	8.4	4.6	11.6	4.2	7.4

4.2 Regression analysis

To find answers to the first and second research subquestions: 'Which demographic characteristics are related to neighbourhood socioeconomic status?' and 'What are the differences and similarities in this relationship between different cities?', a multiple linear regression was performed. The results can be found in Table 4.

Earlier research showed that there is an association between high population density and lower socioeconomic status (Meijer et al., 2012). The maps in Figure 4 show that most low SES neighbourhoods are in the central parts of the cities in this analysis, which is also where population density is often higher. This suggests that the first hypothesis, '*The population density of a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood'*, could be confirmed. The variable population density is, however, not significant in any of the seven regression models, which shows that in this analysis, there is no association between the population density of a neighbourhood and the mean socioeconomic status of this neighbourhood. The first hypothesis can thus not be confirmed or denied based on this analysis.

The second hypothesis, '*The share of inhabitants with a migration background in a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood'*, is based on research that states that ethnic minorities and migrants often live in less advantaged neighbourhoods (Van Ham et al., 2021; Coulter & Clark, 2019). These studies however did not make a distinction between people with a western or non-western migration background. The regression coefficients for the share of inhabitants with a western migration background are significant in four out of seven models. These significant coefficients are all positive, indicating a positive association between the share of inhabitants with a western migration and the mean SES of the neighbourhood. The regression coefficients for the share of inhabitants with a non-western migration background are significant in all seven models. These coefficients are all negative, thus the higher the share of inhabitants with a non-western migration background, the lower the SES of the neighbourhood. The second hypothesis can therefore be confirmed for the share of inhabitants with a non-western migration background, but not for the share of inhabitants with a western migration background. The people in this analysis with a western migration background likely have a higher income and education, resulting in a higher socioeconomic status, compared to those with a non-western migration background.

Previous research suggests that SES is expected to decrease after retirement age, due to decreasing income and employment history (which is measured as whether and how much they worked over the last four years). While at the same time, it is suggested that a neighbourhood with a larger share of older inhabitants has a higher socioeconomic status, due to the fact that people with more education and a higher income live longer and healthier lives (Wagg et al., 2021). Since the retirement age in the Netherlands was around 65 years old in 2017 (CBS, 2024), the share of inhabitants aged 65 years and older is expected to have negative coefficients. In the regression analysis, the share of inhabitants aged 65 years and older is significant in four models. All significant coefficients are negative, which means that the third hypothesis, '*The share of older inhabitants in a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood',* can be confirmed based on four models in this study. The higher the share of inhabitants aged 65 years and older in a neighbourhood in the models 'All cities', 'Groningen', 'Eindhoven', and 'Amersfoort', the lower the neighbourhood SES.

One person households are associated with lower socioeconomic status than couples, while at the same time, households without children are associated with higher socioeconomic status than households with children, according to research by Karney (2021). For the share of one person households, this gave the following hypothesis: '*Hypothesis 4.1: The share of one person households in a neighbourhood is negatively associated with the socioeconomic status of the neighbourhood'*. This variable is significant in the models 'Enschede', 'Groningen' and 'Amersfoort'. For Enschede and Groningen, the coefficients are positive and thus do not support the hypothesis. For Amersfoort, the coefficient is negative, which does support the hypothesis. The share of households without children gives significant coefficients in three out of seven models, 'All cities', 'Enschede', and 'Groningen', these coefficients are all positive. Therefore, hypothesis 4.2, *'The share of households without children in a neighbourhood is positively associated with the socioeconomic status of the neighbourhood'*, is supported by these three models. The variable share of households with children is significant in two models, both with positive coefficients. The models 'Enschede' and 'Groningen' therefore support the hypothesis for the share of households with children: '*The share of households with children in a neighbourhood is positively associated with the socioeconomic status of the neighbourhood'*.

The adjusted R-squared, which adjusts for the number of variables, is high in all models, ranging from 0.761 to 0.853. This means that in the highest-scoring model, which is Dordrecht, 85.3% of the variance in the dependent variable SES is explained by the independent variables. Notable is that in the regression results, the household composition coefficients that are significant are mainly found in the models for the two cities with the lowest mean SES (Enschede and Groningen). In cities with a higher mean SES household composition does not seem to have a significant effect on neighbourhood-level SES. The significance of the variables about migration background and age composition does not seem to be associated with the mean SES of the cities.

Table 4: Regression results

	(1) All cities	(2) Enschede	(3) Groningen	(4) Eindhoven	(5) Dordrecht	(6) Amersfoort	(7) Alphen ad Rijn
PopDensity	0.000	-0.000	-0.000	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
WestMigBg	0.014***	-0.012	0.027***	0.024***	0.005	0.028***	0.007
	(0.003)	(0.014)	(0.006)	(0.007)	(0.006)	(0.007)	(0.005)
NonWestMigBg	-0.014***	-0.014***	-0.017**	-0.015***	-0.016***	-0.012***	-0.014***
	(0.001)	(0.003)	(0.005)	(0.002)	(0.002)	(0.001)	(0.002)
Age65plus	-0.005***	-0.008	-0.012*	-0.007*	-0.006	-0.005*	-0.002
	(0.001)	(0.005)	(0.005)	(0.003)	(0.003)	(0.002)	(0.003)
OnePersonHH	-0.011	0.080**	0.089*	0.000	-0.001	-0.034*	-0.007
	(0.006)	(0.027)	(0.037)	(0.026)	(0.012)	(0.016)	(0.011)
HHWithoutChildren	0.012*	0.105***	0.123**	0.026	0.020	-0.016	0.007
	(0.005)	(0.027)	(0.037)	(0.026)	(0.013)	(0.015)	(0.011)
HHWithChildren	0.002	0.089**	0.102**	0.013	0.009	-0.020	0.008
	(0.006)	(0.027)	(0.037)	(0.026)	(0.013)	(0.016)	(0.011)
cons	0.087	-8.618**	-10.119**	-1.089	-0.598	2.496	-0.078
	(0.543)	(2.659)	(3.632)	(2.642)	(1.340)	(1.548)	(1.109)
N	490	61	66	92	87	107	77
adj. R^2	0.761	0.799	0.793	0.811	0.853	0.815	0.834

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3 Correlation analysis

A Pearson's correlation analysis was performed to find answers to the third and fourth research subquestions: 'Which neighbourhood characteristics correlate with socioeconomic status?' and 'What are the differences and similarities in these correlations between different cities?'. Table 5 shows the correlation coefficients of this analysis.

Previous research suggested that neighbourhoods with higher socioeconomic status have more amenities (Crawford et al., 2008; Powell et al., 2007). Altschuler et al. (2004) added to this that high socioeconomic status neighbourhoods do not have more amenities per definition, but residents with higher SES do have more access to cars, while residents with lower SES are more likely to rely on public transport. These studies, however, focused on the United States and Australia. In the Dutch context, more amenities are often located in city centres and surroundings, which are also the areas with the most lower SES neighbourhoods. Therefore, the hypothesis for the correlation between amenities and neighbourhood SES is: '*H5: The number of accessible amenities (per 1000 inhabitants) of a neighbourhood is negatively correlated with the socioeconomic status of the neighbourhood'*. The amenities variable is subdivided into five categories; 'health and wellbeing', 'retail', 'catering industry', 'childcare and education', and 'leisure and culture'. In all models, the correlation between each of these amenities and SES is very weak to weak (lower than 0.4). The coefficients for amenities are largely negative, with an exception for the city of Eindhoven, for which all amenities have positive coefficients. Thus, for most cities, the hypothesis can be confirmed. On average, the correlations are least strong for 'health and wellbeing', and strongest for 'retail'. For all amenities except 'leisure and culture', the strongest correlation is Alphen aan den Rijn, and Eindhoven is the weakest or second to weakest correlation.

Looking at the mean correlation for all amenities per city, Eindhoven and Dordrecht, the two cities that are in the middle when it comes to mean SES, score the lowest correlations. Enschede and Groningen, the two cities with the lowest mean SES, follow after that, and Amersfoort and Alphen aan den Rijn, the two cities with the highest mean SES, score the highest mean correlations for amenities. Since almost all coefficients are negative in these cities, a higher amount of present amenities is correlated with a lower neighbourhood SES, and the other way around, confirming the hypothesis. Thus, from the correlation coefficients, it seems that for cities with a higher mean SES, there is a stronger correlation between neighbourhood SES and the presence of amenities than for cities with a lower mean SES. However, the difference between the correlation coefficients can be statistically tested using the cortesti command (see Appendix 1). The null hypothesis of this test is an equal correlation in both models, which is rejected in the case of significant p-values. Doing this with the correlation coefficients of the four cities with lower mean SES compared to the two cities with the highest mean SES (Amersfoort and Alphen aan den Rijn) shows that a large majority of coefficients are not significantly different from each other (p-values are higher than 0.05). Therefore, the highest mean SES cities do not have significantly stronger correlations between amenities and neighbourhood SES.

The maps in Figure 4 show that for most cities the neighbourhoods with lower SES are concentrated in and around the city centres. These maps thus confirm the negative coefficients, the lower SES neighbourhoods are concentrated around city centres, which is also the location of most amenities. Eindhoven is an exception to this, since in this city the neighbourhoods with higher and lower SES scores seem to be distributed throughout the city. This could explain why for Eindhoven the coefficients for all

amenities are positive, while they are negative for most other cities. 'Leisure and culture' is the only amenity with more positive than negative coefficients. This could be due to the fact that some of these amenities are more frequently located outside the city centres.

Like the correlation coefficients for the different amenities, those for 'green spaces' are mostly very weak to weak, although on average slightly higher than for amenities. For most models the correlation between green spaces and SES is positive, with the exception of a negative correlation for Eindhoven (-0.023), and no correlation for Amersfoort (-0.000), these are also the cities with the least strong correlations. The positive correlation coefficients for the other models support the sixth hypothesis: '*The hectares of green space (per 1000 inhabitants) of a neighbourhood is positively correlated with the socioeconomic status of the neighbourhood'*. This hypothesis is based on multiple previous studies that conclude that the availability of green spaces is lower in low SES neighbourhoods (Astell-Burt et al, 2014; Rigolon, 2016; Schüle et al., 2017). A positive correlation implies that neighbourhoods with a higher amount of green spaces per 1000 inhabitants have a higher SES, and the other way around. The highest correlation coefficients for the correlation between green spaces and SES are found in Groningen and Enschede, the cities with the lowest mean SES scores. This could indicate that Groningen and Enschede have less green in their city centres, since that is where most low SES neighbourhoods are located (Figure 4). The correlation coefficients of Eindhoven, Dordrecht, and Amersfoort are all significantly different (p-value lower than 0.05) from either Enschede or Groningen, or both. Therefore it can be stated that the lowest mean SES cities do have significantly stronger correlations between green spaces and neighbourhood SES, compared to most cities with a higher mean SES.

The correlation between housing and socioeconomic status was most straightforward based on the existing literature. Higher income and education, and thus socioeconomic status, leads to more opportunities when it comes to housing and choosing where to live. More expensive houses in a neighbourhood attract people with higher socioeconomic status, while people with lower socioeconomic status are much more limited in their options (Li & Wei, 2020; Schirmer et al., 2014). The seventh hypothesis was therefore: '*The mean house value (x1000 euro) of a neighbourhood is positively correlated with the socioeconomic status of the neighbourhood'*. This hypothesis can be confirmed since the correlation between mean house value and SES is positive in all models. A neighbourhood with a higher mean house value likely also has a higher SES, and the other way around.

In all models, the correlation between mean house value and SES has the highest correlation coefficients out of all the neighbourhood characteristics, with coefficients ranging from 0.592 (moderate) to 0.774 (strong). The correlation has the lowest coefficients in the models Enschede and Groningen, the cities with the lowest mean SES, and the highest coefficients in the model for Alphen aan den Rijn, the city with the highest mean SES. The coefficients of house value for Enschede and Alphen aan den Rijn are significantly different (p-value 0.046), however, the coefficients of Groningen and Alphen aan den Rijn are not significantly different (p-value 0.122). Therefore, for the city in this study with the highest mean SES per neighbourhood, Alphen aan den Rijn, there is a significantly stronger positive correlation between neighbourhood SES and mean house value. For the city in this study with the lowest mean SES per neighbourhood, Enschede, there is a significantly weaker positive correlation between neighbourhood SES and mean house value, but this does not apply to the cities in between.

In disadvantaged neighbourhoods, with a lower mean socioeconomic status, the chance of becoming a victim of crime is larger, according to Sugiyama et al. (2015) and Van Wilsem et al. (2006). Since the number of crimes is spatially clustered in low SES neighbourhoods according to the literature, the hypothesis for safety is: '*The number of crimes (per 1000 inhabitants) in a neighbourhood is negatively correlated with the socioeconomic status of a neighbourhood'*. The neighbourhood characteristic safety is subdivided into three categories; 'theft', 'vandalism and public order crimes', and 'violent and sexual crimes'. For all three variables, the crime correlation coefficients are negative in all models, which confirms the hypothesis. Since all values are negative, a higher number of crimes per 1000 inhabitants is correlated to a lower neighbourhood SES. While at the same time, a neighbourhood with a higher mean SES is correlated with a lower number of crimes per 1000 inhabitants. Looking at the maps in Figure 4 shows that most crimes (per 1000 inhabitants) are happening in and around the city centres of most cities.

The correlation between the crime variables and SES is mostly weak. However, on average the correlations are stronger than the ones for amenities and green spaces. In most models, the correlation between violent and sexual crimes and SES is the strongest out of the safety variables, and the correlation between theft and SES is the least strong. Looking at the mean correlation for all safety variables per city, Enschede and Eindhoven score the highest correlation coefficients. The lowest mean safety correlation coefficients are for Amersfoort and Dordrecht. Correlations for safety thus seem to be slightly higher for the cities with lower mean SES, but there is no clear distinction. This distinction is there when looking only at vandalism and public order crimes, the three highest correlation coefficients are for the three cities with the lowest mean SES, and the three lowest correlation coefficients are for the three cities with the highest mean SES. However, testing for the significance of the difference between the correlation coefficients, shows that none of these coefficients are significantly different. It is thus impossible to say that the correlations between the number of crimes and neighbourhood SES are significantly stronger in cities with a lower mean SES.

Table 5: Correlation results (shown are correlations with SES)

	(1) All cities	(2) Enschede	(3) Groningen	(4) Eindhoven	(5) Dordrecht	(6) Amersfoort	(7) Alphen ad Rijn
SES	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Health Wellbeing	-0.057	0.028	-0.086	0.002	0.015	-0.126	-0.140
Retail	-0.088	-0.164	-0.192	0.036	-0.078	-0.167	-0.409
Catering Industry	-0.078	-0.171	-0.205	0.034	-0.006	-0.152	-0.288
Childcare Education	-0.019	-0.128	-0.071	0.062	-0.094	-0.032	-0.292
Leisure Culture	-0.033	0.150	0.014	0.165	-0.105	-0.207	0.104
Green Spaces	0.060	0.327	0.455	-0.023	0.102	-0.000	0.160
House Value	0.689	0.592	0.644	0.752	0.707	0.702	0.774
Theft Crimes	-0.187	-0.350	-0.070	-0.144	-0.165	-0.106	-0.131
Vandalism & Public Order Crimes	-0.312	-0.453	-0.347	-0.385	-0.250	-0.167	-0.322
Violent & Sexual Crimes	-0.359	-0.322	-0.349	-0.493	-0.278	-0.277	-0.360

Orange: very weak (0 - 0.1), yellow: weak (0.1 - 0.4), blue: moderate (0.4 - 0.7), green: strong (0.7 - 1)

5. Conclusion and discussion

5.1 Conclusions

Problems such as poverty, low quality of education, and crime are unequally distributed across space, and accumulated in disadvantaged neighbourhoods. This spatial clustering is causing significant differences between neighbourhoods. People with higher socioeconomic status often live in the most liveable neighbourhoods, these tend to be safer, greener, and have more amenities. Consequently, socioeconomic segregation influences the chances and opportunities of residents of both disadvantaged neighbourhoods and those with a high average SES. Certain demographics are more likely to live in these disadvantaged neighbourhoods, creating overlapping patterns of segregation. At the same time, certain characteristics of a neighbourhood are correlated with the mean SES of the neighbourhood.

This research aimed to find out how different demographic (population density, ethnicity, age composition, household composition) and neighbourhood (amenities, green spaces, housing, safety) characteristics relate to the socioeconomic status of a neighbourhood. Starting from the neighbourhood effects theory, the theoretical framework outlined these main demographic and neighbourhood characteristics and their (cor)relation to SES. Although theory suggests relationships between SES and all variables in the analysis, not all of these were found in this study. However, both the regression and correlation analysis gave some meaningful results. All these variables have been studied in relation to SES before, however, never all together. This research therefore gives a better understanding of the relative importance of the demographic and neighbourhood characteristics in the context of six large cities in the Netherlands.

The first research subquestion was: *'Which demographic characteristics (population density, ethnicity, age composition, household composition) are related to neighbourhood socioeconomic status?'*. All variables were significant in two or more out of seven models, except population density, which was not significant in any model. Therefore, no relationship was found between population density and neighbourhood socioeconomic status. Only one variable was significant in all seven models of the regression, the share of inhabitants with a non-western migration background. The coefficients are negative in all models, therefore, a higher share of inhabitants with a non-western migration background is associated with a lower neighbourhood socioeconomic status. It can thus be concluded that in the cities in this analysis, people with a non-western migration background are among the demographic groups that are more likely to live in disadvantaged neighbourhoods, which can significantly influence their opportunities. Knowing this, policymakers could invest in these neighbourhoods by increasing these residents' socioeconomic status, for instance through education and economic development programs.

Comparing the significance of the variables in the different models, in order to answer the second subquestion; *'What are the differences and similarities in this relationship between different cities?'*, shows that the significant household composition variables are almost all found in the two models for the cities with the lowest mean SES, Enschede and Groningen. Thus, in those cities household composition has a significant effect on the neighbourhood SES, while it does not in cities with a higher mean SES. The household composition of a neighbourhood can be partly influenced by the reorganization of the housing

stock. This way, policymakers for instance can try to attract more households with children, who on average have higher socioeconomic status than one person households.

The third subquestion was: *'Which neighbourhood characteristics (amenities, green spaces, housing, safety) correlate with socioeconomic status?'*. The mean house value gives the strongest correlation between any neighbourhood characteristic and socioeconomic status in each model. This is not surprising, since an individual's income (as part of the SES) largely determines their possibilities in terms of housing. This does mean that many people live in very homogenous neighbourhoods. Creating a more diverse housing stock in a neighbourhood might be a way to prevent further socioeconomic segregation.

Safety, measured as the number of registered crimes in three categories, are the next strongest correlations. The correlation coefficients are negative for all three variables in all seven models, and thus a higher number of crimes correlates with lower neighbourhood socioeconomic status. The maps with the spatial distribution of socioeconomic status show that for most cities the neighbourhoods with lower socioeconomic status are located in and around the city centers. Since the safety correlations are negative, most crimes are happening in the central parts of the cities in the analysis. This implies that people living in lower socioeconomic status neighbourhoods, on average experience more crime in their daily living environment. This could influence neighbourhood satisfaction and mental health.

The correlations for the different amenities are all relatively weak, making it difficult to form strong conclusions. The negative correlation coefficients for most amenities are confirmed by the maps with the spatial distribution of SES. Most amenities are located in and around city centres, which is also the location of most low SES neighbourhoods. Living in a low SES neighbourhood thus often means more amenities in the immediate vicinity. For residents with a lower income, this could mean more affordable options and fewer travel expenses.

The fourth subquestion; *'What are the differences and similarities in these correlations between different cities?'*, can be answered by statistically testing the difference between correlation coefficients of different models. The correlation between amenities and SES is highest in Amersfoort and Alphen aan den Rijn, the cities with the highest mean SES. However, these coefficients are not significantly different from those of the other cities, and thus the highest mean SES cities do not have significantly stronger correlations when it comes to amenities. The same applies to the correlations between the safety variables and SES, the correlations seem to be higher for the cities with the lower mean SES, yet the test showed that none of these coefficients are significantly different between models. Thus, although the correlation coefficients are always higher in some models, in most cases they are not significantly different from each other. The correlations for the different cities are very similar. The correlation between green spaces and SES is an exception, the test showed that the lowest mean SES cities do have significantly stronger correlations, compared to most cities with a higher mean SES. Since these correlations are positive, inhabitants of a high SES neighbourhood in Enschede or Groningen have a higher chance of having more green spaces in their neighbourhood. This could also indicate that most green spaces are located outside the central areas of the cities.

With these subquestions the main research question; *'To what extent is the demographic composition of the inhabitants related to neighbourhood socioeconomic status and to what extent is there a correlation*

between socioeconomic status and neighbourhood characteristics in six large cities in the Netherlands?' can be answered. The demographic composition of the inhabitants of a neighbourhood is related to the neighbourhood's socioeconomic status in all models. This relation is most clearly visible for the share of inhabitants with a non-western migration background, while the share of inhabitants with a western migration background and the share of inhabitants aged 65 years or older also give significant coefficients in four out of seven models. The household composition seems more influential on neighbourhood SES in cities with a lower mean SES. The correlation between the neighbourhood characteristic mean house value and neighbourhood SES is strongest in all models, followed by the different safety variables. Although smaller, there is also a correlation between amenities and SES, and between green spaces and SES. In most cases, the coefficients of different models are not significantly different.

5.2 Limitations

This research does have some limitations, the first limitation is that the analysis was done with data from 2017, which was the only year for which data was available for all variables. When it comes to the demographic composition of neighbourhoods (like an increasing share of inhabitants aged 65 years or older), or neighbourhood characteristics (for instance crime rates and house values), a lot can change in seven years. Additionally, some cities have changed the spatial division of their neighbourhoods since 2017. Therefore, more recent data would be useful to formulate suitable policy suggestions.

Another limitation is the relatively small sample size for some of the cities in the analysis. This was attempted to be prevented by choosing cities with at least 100,000 inhabitants, however, after removing the neighbourhoods with too few inhabitants, some cities did not have that many neighbourhoods left. This might have affected the results of the regression and correlation analysis. The number of cities in the analysis also makes it difficult to generalize the findings.

5.3 Recommendations for further research

Further research could include more cities in the analysis, to make the results more generalizable. Additionally, the change in (cor)relations over time could be interesting. With longitudinal data, the changes and trends over time could be investigated, which would give a deeper understanding of the relations between the demographic composition of a neighbourhood, the neighbourhood characteristics, and SES. This could give insight into the way neighbourhoods change over time, and the development of socioeconomic segregation.

Finally, further research could try to find out whether people with higher socioeconomic statuses move to more liveable neighbourhoods, whether these liveable neighbourhoods are created in places where people live with high socioeconomic statuses, or whether neighbourhoods with favourable characteristics lead to people increasing their socioeconomic status. This would give a better understanding of the processes behind socioeconomic segregation.

References

Aamodt, G., Nordh, H., & Nordbø, E. C. (2023). Relationships between socio-demographic/socio-economic characteristics and neighborhood green space in four Nordic municipalities–Results from NORDGREEN. *Urban Forestry & Urban Greening*, 82, 127894.

Altschuler, A., Somkin, C. P., & Adler, N. E. (2004). Local services and amenities, neighborhood social capital, and health. *Social science & medicine*, 59(6), 1219-1229.

Andersson, R., & Musterd, S. (2010). What scale matters? Exploring the relationships between individuals' social position, neighbourhood context and the scale of neighbourhood. *Geografiska Annaler: Series B, Human Geography*, 92(1), 23-43.

Astell-Burt, T., Feng, X., Mavoa, S., Badland, H. M., & Giles-Corti, B. (2014). Do low-income neighbourhoods have the least green space? A cross-sectional study of Australia's most populous cities. *BMC Public Health*, 14, 1-11.

Boterman, W. R., Musterd, S., & Manting, D. (2021). Multiple dimensions of residential segregation. The case of the metropolitan area of Amsterdam. *Urban Geography*, 42(4), 481-506.

CBS (2024). *Pensioenleeftijd werknemers nadert 66 jaar*. Retrieved on 27-06-2024 from https://www.cbs.nl/nl-nl/nieuws/2024/18/pensioenleeftijd-werknemers-nadert-66-jaar. Centraal Bureau voor de Statistiek.

CBS (n.d.). *Hoe interpreteer je de SES-WOA-scores en hoe zijn deze bepaald?* Retrieved on 01-05-2024 from https://www.cbs.nl/nl-nl/faq/infoservice/hoe-interpreteer-je-de-ses-woa-scores-en-hoe-zijn-deze-bepaald-. Centraal Bureau voor de Statistiek.

Cheshire, P. (2007). Are mixed communities the answer to segregation and poverty? *Viewpoint: Informing debate*. York: Joseph Rowntree Foundation

Coenen, A., Verhaeghe, P. P., & Van de Putte, B. (2019). Ethnic residential segregation: A family matter? An integration of household composition characteristics into the residential segregation literature. *European Journal of Population*, 35(5), 1023-1052.

Costa, M. & Neves, J. (2020). Insertion of Distributed Photovoltaic Generation in Brazil: A Correlation Analysis between Socioeconomic and Geographic Aspects. *International Journal of Energy Economics and Policy*. 10. 102-111.

Coulter, R., & Clark, W. A. (2019). Ethnic disparities in neighbourhood selection: Understanding the role of income. *International Journal of Urban and Regional Research*, 43(5), 947-962.

Cowie, C. T., Ding, D., Rolfe, M. I., Mayne, D. J., Jalaludin, B., Bauman, A., & Morgan, G. G. (2016). Neighbourhood walkability, road density and socio-economic status in Sydney, Australia. *Environmental health*, 15, 1-12.

Crawford, D., Timperio, A., Giles-Corti, B., Ball, K., Hume, C., Roberts, R., Andrianopoulos, N. & Salmon, J. (2008). Do features of public open spaces vary according to neighbourhood socio-economic status?. *Health & place*, 14(4), 889-893.

Darcy, M., & Gwyther, G. (2011). Recasting research on 'neighbourhood effects': A collaborative, participatory, trans-national approach. In M. van Ham, D. Manley, N. Bailey, L. Simpson, & D. Maclennan (Eds.), *Neighbourhood effects research: New perspectives*. Dordrecht: Springer

Ellen, I. G., & Turner, M. A. (1997). Does neighborhood matter? Assessing recent evidence. *Housing policy debate*, 8(4), 833-866.

Erdem, Ö., Van Lenthe, F. J., Prins, R. G., Voorham, T. A., & Burdorf, A. (2016). Socioeconomic inequalities in psychological distress among urban adults: the moderating role of neighborhood social cohesion. *PloS one*, 11(6).

Flouri, E., Midouhas, E., & Ruddy, A. (2016). Socio–economic status and family structure differences in early trajectories of child adjustment: Individual and neighbourhood effects. *Health & place*, 37, 8-15.

Galster, G. (2001). On the nature of neighbourhood. *Urban Studies*, 38(12), 2111-2124.

Galster, G. (2011). The mechanism (s) of neighbourhood effects: Theory, evidence, and policy implications. In *Neighbourhood effects research: New perspectives* (pp. 23-56). Dordrecht: Springer Netherlands.

Iceland, J., & Wilkes, R. (2006). Does socioeconomic status matter? Race, class, and residential segregation. *Social problems*, 53(2), 248-273.

Kang, S. J., & Seo, W. (2022). Do people determine their subjective socioeconomic status based on the housing type and residential neighborhood? Empirical evidence from Seoul. *Land*, 11(11), 2036.

Karney, B. R. (2021). Socioeconomic status and intimate relationships. *Annual review of psychology*, 72(1), 391-414.

Klevmarken, A., & Quigley, J. M. (1976). Age, experience, earnings, and investments in human capital. *Journal of Political Economy*, 84(1), 47-72.

Lau, M. H. (2023). Residential age segregation: Evidence from a rapidly ageing Asian city. *Journal of Population Ageing*, 1-21.

Leclerc, C. (2022). Immigrants' earnings and neighbourhood economic wealth: the conditioning role of citizenship. *Journal of Ethnic and Migration Studies*, 48(7), 1591-1609.

Leidelmeijer, K., Burema, F., Mandemakers, J. (2023). *Leefbarometer 2020 - Verdiepende analyses*. Retrieved on 05-03-2024 from https://atlasresearch.nl/leefbaarometer-2020-verdiepende-analyses/. Atlas research.

Li, H., & Wei, Y. D. (2020). Spatial inequality of housing value changes since the financial crisis. *Applied Geography*, 115, 102141.

Mackenbach, J. D., Lakerveld, J., van Lenthe, F. J., Bárdos, H., Glonti, K., Compernolle, S., De Bourdeaudhuij, I., Oppert, J.M., Roda, C., Rutter, H., Brug, J. & Nijpels, G. (2016). Exploring why residents of socioeconomically deprived neighbourhoods have less favourable perceptions of their neighbourhood environment than residents of wealthy neighbourhoods. *Obesity Reviews*, 17, 42-52.

Mackenbach, J. P., Stirbu, I., Roskam, A. J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 European countries. *New England Journal of Medicine*, 358(23), 2468-2481.

Meijer, M., Kejs, A. M., Stock, C., Bloomfield, K., Ejstrud, B., & Schlattmann, P. (2012). Population density, socioeconomic environment and all-cause mortality: a multilevel survival analysis of 2.7 million individuals in Denmark. *Health & place*, 18(2), 391-399.

Musterd, S., & Andersson, R. (2005). Housing mix, social mix and social opportunities. *Urban Affairs Review*, 40 (6), 761–790.

Powell, L. M., Slater, S., Mirtcheva, D., Bao, Y., & Chaloupka, F. J. (2007). Food store availability and neighborhood characteristics in the United States. *Preventive medicine*, 44(3), 189-195.

Rachele, J. N., Learnihan, V., Badland, H. M., Mavoa, S., Turrell, G., & Giles-Corti, B. (2017). Neighbourhood socioeconomic and transport disadvantage: The potential to reduce social inequities in health through transport. *Journal of Transport & Health*, 7, 256-263.

Rigolon, A. (2016). A complex landscape of inequity in access to urban parks: A literature review. *Landscape and urban planning*, 153, 160-169.

Roode, T., Sharples, K., Dickson, N., & Paul, C. (2017). Life-Course Relationship between Socioeconomic Circumstances and Timing of First Birth in a Birth Cohort. *PLoS ONE*, 12.

Roosa, M.W., White, R.M.B. (2014). Neighborhood Effects. In: Michalos, A.C. (eds) *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Dordrecht.

Sabater, A., & Finney, N. (2023). Age segregation and housing unaffordability: Generational divides in housing opportunities and spatial polarisation in England and Wales. *Urban Studies*, 60(5), 941-961.

Schaapman, C., Pastors, M., Slob, A. (2024). *'Stop de stapeling van problemen in achterstandswijken'*. AD/Algemeen Dagblad.nl.

Schirmer, P. M., Van Eggermond, M. A., & Axhausen, K. W. (2014). The role of location in residential location choice models: a review of literature. *Journal of Transport and Land Use*, 7(2), 3-21.

Schüle, S. A., Gabriel, K. M., & Bolte, G. (2017). Relationship between neighbourhood socioeconomic position and neighbourhood public green space availability: An environmental inequality analysis in a large German city applying generalized linear models. *International journal of hygiene and environmental health*, 220(4), 711-718.

Simpson, P. A., Greller, M. M., & Stroh, L. K. (2002). Variations in human capital investment activity by age. *Journal of Vocational Behavior*, 61(1), 109-138.

Sugiyama, T., Howard, N. J., Paquet, C., Coffee, N. T., Taylor, A. W., & Daniel, M. (2015). Do relationships between environmental attributes and recreational walking vary according to area-level socioeconomic status?. *Journal of Urban Health*, 92, 253-264.

Vaalavuo, M., Van Ham, M., & Kauppinen, T. M. (2019). Income mobility and moving to a better neighbourhood: An enquiry into ethnic differences in Finland. *European Sociological Review*, 35(4), 538-551.

Van Ham, M. & Manley, D. (2010). The effect of neighbourhood housing tenure mix on labour market outcomes: A longitudinal investigation of neighbourhood effects. *Journal of Economic Geography*, 10 (2), 257–282.

Van Ham, M., Manley, D., Bailey, N., Simpson, L., & Maclennan, D. (2011). Neighbourhood effects research: New perspectives. In *Neighbourhood effects research: New perspectives* (pp. 1-21). Dordrecht: Springer Netherlands.

Van Ham, M., Tammaru, T., Ubarevičienė, R., & Janssen, H. (2021). Urban socio-economic segregation and income inequality: A global perspective. Springer Nature.

Van Wilsem, J., Wittebrood, K., & De Graaf, N. D. (2006). Socioeconomic dynamics of neighborhoods and the risk of crime victimization: A multilevel study of improving, declining, and stable areas in the Netherlands. *Social Problems*, 53(2), 226-247.

Viinikka, A., Tiitu, M., Heikinheimo, V., Halonen, J. I., Nyberg, E., & Vierikko, K. (2023). Associations of neighborhood-level socioeconomic status, accessibility, and quality of green spaces in Finnish urban regions. *Applied Geography*, 157, 102973.

Vogiazides, L., & Chihaya, G. K. (2020). Migrants' long-term residential trajectories in Sweden: persistent neighbourhood deprivation or spatial assimilation?. *Housing Studies*, 35(5), 875-902.

Wagg, E., Blyth, F. M., Cumming, R. G., & Khalatbari-Soltani, S. (2021). Socioeconomic position and healthy ageing: A systematic review of cross-sectional and longitudinal studies. *Ageing Research Reviews*, 69, 101365.

Wilson, W. J. (1987). The truly disadvantaged: the inner city, the underclass and public policy. Chicago: University of Chicago Press.

Zhou, Y., Wang, F., & Zhou, S. (2023). The Spatial Patterns of the Crime Rate in London and Its Socio-Economic Influence Factors. *Social Sciences*, 12(6), 340.

Appendix 1: Stata commands

Histogram distribution of SES scores (Figure 3)
"hist SES, frequency normal"

Descriptive statistics, for all seven datasets (Table 3)
"sum SES PopDensity WestMigBg NonWestMigBg Age0to14 Age15to24 Age25to44 Age45to64 Age65plus OnePersonHH HHWithoutChildren HHWithChildren HealthWellbeing Retail CateringIndustry ChildcareEducation LeisureCulture GreenSpaces HouseValue Theft VandalismPublicOrder ViolenceSexual"

"ssc install asdoc"

"asdoc sum SES PopDensity WestMigBg NonWestMigBg Age0to14 Age15to24 Age25to44 Age45to64 Age65plus OnePersonHH HHWithoutChildren HHWithChildren HealthWellbeing Retail CateringIndustry ChildcareEducation LeisureCulture GreenSpaces HouseValue Theft VandalismPublicOrder ViolenceSexual, save(sum_[city name]) dec(3)"

Regression, for all seven datasets (Table 4)
"ssc install estout"

"reg SES PopDensity WestMigBg NonWestMigBg Age65plus OnePersonHH HHWithoutChildren HHWithChildren"

"est store [city name]"

"esttab allcities enschede groningen eindhoven dordrecht amersfoort alphenadrijn, se ar2 b(%9.3f) se(%9.3f) mti(allcities enschede groningen eindhoven dordrecht amersfoort alphenadrijn)"

"esttab allcities enschede groningen eindhoven dordrecht amersfoort alphenadrijn using "results65plus.rtf", rtf se ar2 b(%9.3f) se(%9.3f) mti(allcities enschede groningen eindhoven dordrecht amersfoort alphenadrijn) compress"

Correlation, for all seven datasets (Table 5)
"corr SES HealthWellbeing Retail CateringIndustry ChildcareEducation LeisureCulture GreenSpaces HouseValue Theft VandalismPublicOrder ViolenceSexual"

Statistically test the difference between correlation coefficients
"cortesti (coefficient1) (n1) (coefficient2) (n2)"