

# Northumbria Research Link

Citation: Lengyel, Janka, Roux, Stéphane and Alvanides, Seraphim (2022) Multivariate analysis of socioeconomic profiles in the Ruhr Area, Germany. *Journal of Maps*, 18 (3). pp. 576-584. ISSN 1744-5647

Published by: Taylor & Francis

URL: <https://doi.org/10.1080/17445647.2022.2098839>  
<<https://doi.org/10.1080/17445647.2022.2098839>>

This version was downloaded from Northumbria Research Link:  
<https://nrl.northumbria.ac.uk/id/eprint/49410/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria  
University**  
NEWCASTLE



**UniversityLibrary**

## Multivariate analysis of socioeconomic profiles in the Ruhr Area, Germany

Janka Lengyel<sup>a,b,c</sup> <https://orcid.org/0000-0002-1301-7907>

Stéphane Roux<sup>a</sup> <https://orcid.org/0000-0001-8561-1270>

Seraphim Albanides<sup>d,e</sup> <https://orcid.org/0000-0003-4905-4109>

<sup>a</sup>Université de Lyon, ENS de Lyon, Université Claude Bernard, CNRS, Laboratoire de Physique, F-69342, Lyon, France; <sup>b</sup>Institute for Mobility and Urban Planning, University of Duisburg-Essen, Germany; <sup>c</sup> LVMT, Université Gustave Eiffel, ENPC, Paris, France; <sup>d</sup>Northumbria University, Newcastle upon Tyne, United Kingdom; <sup>e</sup>GESIS Leibniz Institute for the Social Sciences, Cologne, Germany

### ABSTRACT

The aim of this article and associated map is to highlight the social and economic diversity of the Ruhr area in Germany through the use of multivariate analysis and visualisation. To this end we combine two different datasets. Demographic parameters stemming from the 2011 German census and socioeconomic indicators obtained from the *microdialog* of the German post service. Due to the different spatial resolution of the two datasets, we aggregated the data at the neighbourhood (*Stadtteil*) level. The multivariate analysis was carried out at this scale using Self-Organizing Maps (SOM), an artificial neuron network, which uses an unsupervised learning mechanism for projecting multidimensional data in a low (in our case two) dimensional space. First we used a visualization technique to better comprehend the relationship between our observations via reducing the dimensionality or complexity of our input data. At the same time, we established a global statistical relationship between the indicators. Finally, **using these results** we built clusters for revealing the distribution of socioeconomic profiles over the whole region. Our results demonstrate that structural inequalities resulting from the processes of industrialization/deindustrialization in the region are still widely persistent and result in characteristic patterns along the three main rivers, the Lippe, Emscher and the Ruhr. In close connection with this, three types of societal segregation patterns become clearly evident in the Ruhr area, namely nationality, age and economic power.

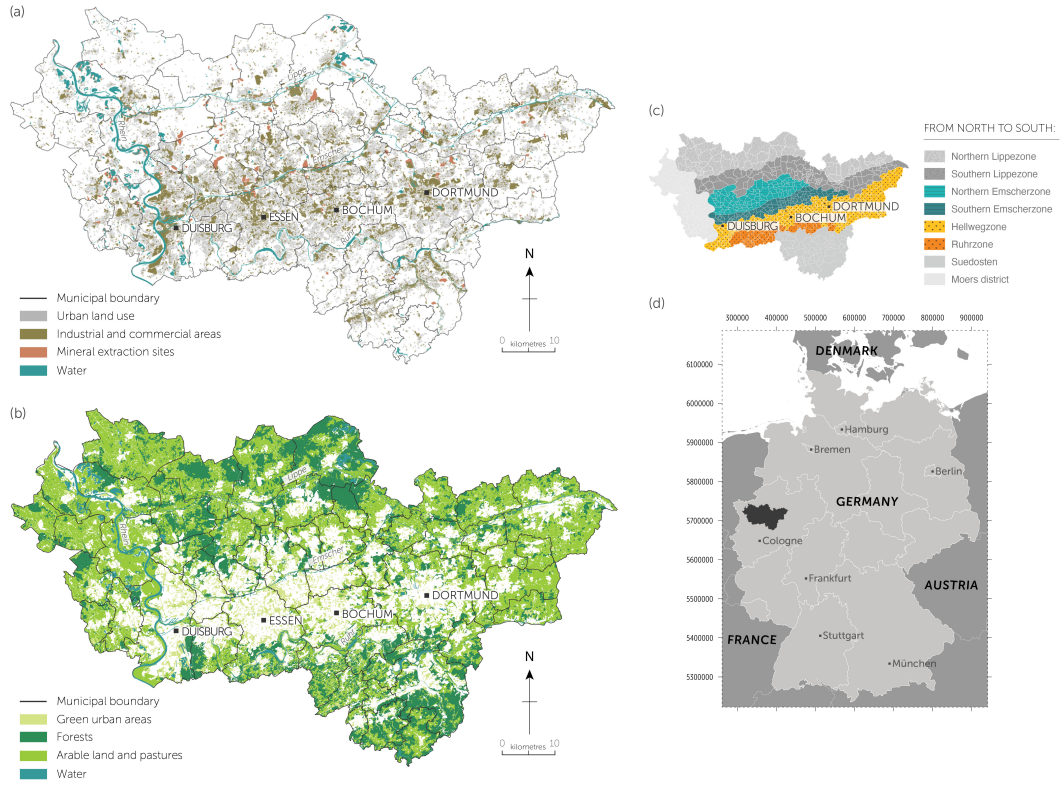
### KEYWORDS

Multivariate analysis; Self-organizing maps; Socioeconomic clusters; Segregation patterns

## 1. Introduction

The Ruhr metropolitan region (*Ruhrgebiet*) in Germany covers an area of 4435 km<sup>2</sup> with a population of just above 5.1 million (2019) and was considered to be the industrial powerhouse of Germany (see also Fig. 1.c). With 53 cities and towns it is also the most densely populated polycentric urban area and exhibits the strongest socioeconomic diversity not only between but also within cities. Due to its strong industrial past the Ruhr area has received multiple waves of foreign workforce since the second half of the nineteenth century resulting in a multi-ethnic population. It also houses 22 Universities and colleges making it the area with the highest density of further and higher education establishments in Germany. This diversity is usually reflected in official statistics as an urban-rural divide between the four major University cities (Dortmund, Essen/Duisburg, Bochum) and the rest of the region, therefore losing sight of the beautiful social mosaic of this region.

For many decades, the economic dynamism of the Ruhr Area was shaped primarily by the production industry. Since the crisis of the coal and steel industries in the 1950s however, the tertiary sector advanced incrementally to the foreground. Between 1957 and 1999 the Ruhr Area lost 87 percent of its employees in the previously imperious coal and steel industries (calculations based on *Regionalkunde Ruhrgebiet* (2019)). Today, the services sector is by far the most dominant in the regional market: In 2015, 78 percent of employees worked here exceeding the national average by four percent. Even though a detailed discussion of latter process that is often referred to as *Strukturwandel* or structural change is beyond the scope of this article, we will uncover its characteristic traces in the results sections.



**Figure 1. Caption** (a) Green and recreational land use. (b) Urban, commercial and industrial land use. (c) Schematic view of the eight historical zones of the Ruhr area (d) The location of the Ruhr area in Germany. Data source: Urban Atlas of the Copernicus Land Monitoring Service (*Urban Atlas*, 2019)

**Figure 1. Alt Text** Industrial, commercial and urban land uses are concentrated mainly in the Emscher and Hellweg zones, located in the central parts of the Ruhr region. Green and recreational areas dominate the outskirts, where industrial production was substantially less extensive throughout the different phases of industrialization.

We should emphasize here that industrialization took place in several consecutive phases (Wehling, 2015) from the southern area along the Ruhr river northwards, leaving definitive territorial and socioeconomic marks behind (Lengyel, 2021; Lengyel & Friedrich, 2019; Wehling, 2015). Their varying profiles can be attributed to resource availability, technological advances and the political and economical whims and interests of the given period in time (Wehling, 2015). The scope of this article is to demonstrate how these zones can be characterized by rather distinctive socioeconomic (and land-use) profiles up until today. The main zones considered here are the fluvial areas next to the Ruhr, Emscher and Lippe rivers, and the so-called Hellwegzone which connects the former large industrial city-centers of Dortmund, Bochum, Essen and Mülheim an der Ruhr. As depicted in Fig. 1.a industrial, commercial and urban land uses are concentrated mainly in the Emscher and Hellweg zones (located in the central parts of the Ruhr region), whereas green and recreational areas dominate the outskirts (Fig. 1.b).

## 2. Datasets

This article brings together two different datasets to characterize the socioeconomic profiles of sub-city areas and identify their spatial distribution. The demographic parameters from the 2011 German census (*Zensus*, 2011) are combined with socioeconomic indicators obtained from the *microdialog* of the German post service (*Microdialog*, 2017). Thereby indices for purchasing power and net rent price are calculated from the latter, whilst the percentage of foreign and German population, residents under 18 and over 65-year-olds as well as household characteristics are all downloaded from the official census website (*Zensus*, 2011). Small scale census data is provided for 100x100 meter grid whilst *microdialog* data is delivered at the building level. Due to the different resolution of the two datasets, we aggregated the data at the neighbourhood or *Stadtteil* scale (in total 732 for the entire region as shown in Fig. 3) and calculated respective mean values for all the variables. Moreover, we used census data from the *Landesdatenbank Nordrhein-Westfalen* (*Landesdatenbank NRW*, 2019) and from the Regional Planning Association Ruhr (*Regionalstatistik Ruhr*, 2019) to identify prevailing societal and economic trends for the entire metropolitan region. The advantage is that at the larger city-level variables are available for longer time periods and enriched by additional indicators such as employment, education and economic performance. Temporal resolution is dependent on the observed feature, e.g. population numbers can be retrieved annually from 1976 onwards, however GDP and employment data is only available from 1990 and 2000 respectively.

## 3. Methodology

The following results section starts with a short summary of metropolitan scale trends that are based on own calculations using data from the official German statistics office (*Federal Statistical Office of Germany*, 2019; *Landesdatenbank NRW*, 2019) and the Regional Planning Association Ruhr (*Regionalstatistik Ruhr*, 2019), attainable mainly for the cities and administrative districts (*Kreis*, e.g. Kreis Recklinghausen or Ennepe-Ruhr-Kreis). The local multivariate analysis is then carried out at the neighbourhood scale using Self-Organizing Maps (SOM), an artificial neuron network (ANN) developed by Teuvo Kohonen (Kohonen, 1990) which uses a competitive learning mechanism for projecting multidimensional data in a low dimensional space. ANNs in general are capable of both machine learning and pattern recognition. They are made up of interconnected nodes - or artificial neurons - that communicate with each other via signal-transmission. These signals are numerical values that are received by the neurons and then further processed according to the learning mechanism in question to reach an output. Furthermore, there are weights assigned to both the nodes (or neurons) and to their connections (or edges) which are incrementally adjusted throughout the learning process. In the case of SOM, data projection happens in a way that similar values in the input space are mapped into the same neuron or to its close vicinity in the output space.

Accordingly, let us consider an  $n$  dimensional vector  $x$  as our input and a set of nodes  $v$  with indices  $i$  of the model space. Each node is assigned a weight vector  $w_i$ . In the training process the node with a weight most similar to the input vector is referred to as the best matching unit (BMU) characterized by its index  $u$  (Agarwal & Skupin, 2008). The BMU and the nodes in its vicinity are incrementally adjusted according to,

$$w_i(t+1) = w_i(t) + \lambda(v, u, t)\alpha(t)[x(t) - w_i(t)] \quad (1)$$

where  $t$  is the current iteration,  $\alpha(t)$  a gradually decreasing learning rate and  $\lambda(v, u, t)$  the so-called neighbourhood function. The latter is dependent on the distance  $d_{u,v}$  between the BMU and the respective node  $v$  at iteration  $t$ . Here, we apply a Gaussian kernel function with an adaptive (Agarwal & Skupin, 2008) kernel environment  $\sigma_u(t)$ , which is incrementally decreasing



with the number of iterations  $t$  as

$$\lambda(v, u, t) = e^{\frac{-d_{u,v}}{2\sigma_u^2(t)}} \quad (2)$$

To sum up, the output of a SOM can be visualised as a two-dimensional spatial representation of the statistical relations between the observations. Accordingly, the axes of the SOM output (see Fig. 2) represent measures of statistical dis/similarity expressed as the distance between observations (Kourtit, Nijkamp, & Arribas, 2012).

In comparison with other dimensionality reduction methods, e.g Principal Component Analysis (PCA) or t-distributed Stochastic Neighbour Embedding (t-SNE), an important advantage of SOM is that it provides information at each node position of the 2D map topology and as a result allows for detection and visualisation of the most characteristic patterns in the dataset (Unglert, Radić, & Jellinek, 2016). It has also been shown that SOM is somewhat less robust than PCA in estimating the optimal number of clusters (Unglert et al., 2016) but performs better when dealing with nonlinear problems (Aguado, Montoya, Borrás, Seco, & Ferrer, 2008). In the context of this study, the above mentioned visual clustering feature is ideal for identifying and representing multilevelled segregation phenomena over the observed territory. SOM thus constitutes a natural choice for the detection and mapping of the complex social economic relationships addressed in Section 6.

#### 4. Software

To generate the maps in this article, the authors used geospatial python programming, which included the following packages: numpy, pandas, geopandas, matplotlib and seaborn. Furthermore, in order to carry out the self-organising maps part of the analysis we used the Python package “Sompy” developed by Moosavi, Packmann, and Vallés (2014) using PCA initialization and a rectangular grid structure. In addition to characterizing urban socioeconomic profiles, the ten here-analysed features (out of a total of 43) were specifically selected after numerous trials, so that their SOM-model would have a topographic error very close to zero (0.072) and a low quantization error (0.7783) according to Sompy-guidelines. Finally, for the layouting and design of our main map file we used the Adobe InDesign software.

#### 5. Current trends in the Ruhr Area

*Societal dynamics today.* As far as current population dynamics are concerned, the single most important trend is the widely prevalent demographic change which is referring to birth rates falling constantly below mortality rates. This culminates in a so-called aging society and decreasing population numbers (Lyberaki, Tinios, Mimis, & Georgiadis, 2013). The difference between live births and deaths in the Ruhr area has been negative since the mid 1970s (own calculations based on data from *Landesdatenbank NRW* (2019)). In the following years and in some recent cases, net population balance was compensated by migration waves, as was the case between 2014-2016, but overall the balance remained negative. The official website of the Regional Planning Association Ruhr estimates an almost 200,000 or four percent decline of inhabitants between 2014 and 2040. At the same time the number of above 65 year-olds will rise by 310,000 (*Regionalstatistik Ruhr*, 2019). Whilst the Ruhr population is shrinking and getting older (Rienow, Stenger, & Menz, 2014), it is also becoming increasingly heterogeneous. The share of foreign inhabitants was 16.2 percent in 2018, which is 3.6 percent higher than the year before. The largest group with foreign nationality (in 2018) comes from Turkey with 24.7% followed by Syria with 10.0% and Poland with 7.6% (*Regionalstatistik Ruhr*, 2019). In comparison, at the national level the proportion of foreign nationals was 12.2% *Federal Statistical Office of Germany*

(2019), which is four percent lower than in the Ruhr area for the same year. In the Ruhr area, the above described general trends of shrinking and aging population are further amplified by the decline of heavy industry (Strohmeier, 2003).

*Economic structural change.* The lengthy restructuring process of old industrial areas may follow rather diverging paths (Birch, MacKinnon, & Cumbers, 2010). In the Ruhr area, a defining characteristic is the diverging timeline of its zonal development: some spatial units faced restructuring challenges earlier than others (Rehfeld & Nordhause-Janz, 2017; Wehling, 2014) which ultimately contributed to the exacerbation of intraregional differences discussed in Section 6. For example, in the mid 1960s mines in Emscher agglomerations were still highly productive, therefore priorities of industrial companies determined land use, land availability and accordingly the direction of urban development. This rendered the settling of new businesses and the implementation of urban regeneration projects almost impossible (Wehling, 2015). Consequently, the regional competitive ability of Emscher settlements was severely restricted compared to the more favourable urban development path of larger Hellweg cities (Duisburg, Mülheim an der Ruhr, Bochum, Essen, Dortmund) where economic restructuring and urban rehabilitation strategies were already taking roots. In comparison to other industrial regions in Europe, population dynamics of the structural change in the Ruhr area can be considered rather stable, losing only 0.4 million inhabitants between 1961 and in 2019 (Rehfeld & Nordhause-Janz, 2017). Although the dynamic interplay of extensive change and structural persistence did incrementally allow for its economic profile to become more diversified (Berndt, 2001), GDP trends in the dusk of the restructuring process between 1996 and 2003 were less favourable than those of most other industrial regions in England and France (Birch et al., 2010). However, this can also be partly attributed to unfavourable performance of the German economy following the German reunification (Birch et al., 2010).

Similar to other post-industrial agglomerations in Europe, the restructuring process in general has led to high (and lengthy) unemployment rates and to a large share of the population working and living without secure long-term perspectives (Netrdová & Nosek, 2016). Latter is strongly indicated for-instance by the high and continuously growing share of people living on social subsidies. The 15.3 percent minimum hedging rate in the Ruhr region in 2015 was more than three percent higher than of its federal state of North Rhine-Westphalia with 12.0 percent (own calculations based on data from the *Federal Statistical Office of Germany* (2019)). Furthermore, these trends may also lead to an increase in small-scale societal polarization and may also have rather diverging zonal- and micro-scale projections, the main topic of the next section.

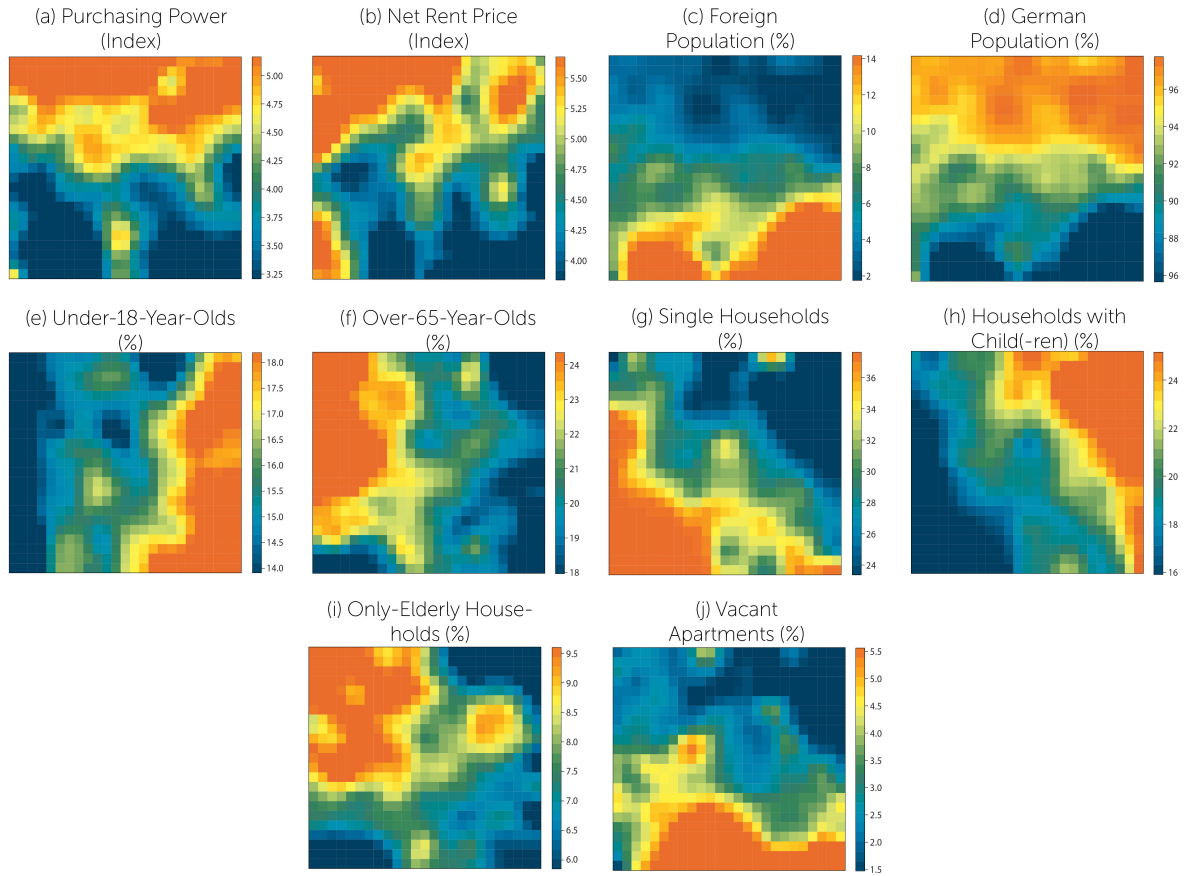
## 6. Distribution of neighbourhood-level socioeconomic clusters over the Ruhr Area

### 6.1. Component Maps

Figure 2 shows the values of vectors for every artificial neuron or node on the 2D map and for each of the ten features/components found to be significant for this analysis. We must stress that there can be multiple geographical neighbourhoods assigned to one node in the 2D model space (Fig. 2 and Fig. 4). The component maps of the neural network may be interpreted in the following way: brighter orange-red correspond to higher values, while darker cyan/blue represent lower values. Therefore, neighbourhoods located in the upper left corner are characterized by high intensities of purchasing power, net rent price, share of German population and percentage of elderly households (having a positive correlation). These have been assigned to clusters 24, 19 and 11 according to the Hit Map in Fig. 4. Furthermore, these latter neighbourhoods are each in a negative correlation with neighbourhoods in the lower right corner (the very opposite side of each SOM map) exhibiting below average rents, purchasing power and elderly, and concurrently very high share of foreign population and under-18-year-olds (Figure 2). These are located in clusters 17, 5 and 19, according to Figure 4. Again, these correlate strongly with one another

and exhibit a negative statistical relationship with features of the previously described upper-left corner. As we already mentioned in the theoretical introduction, statistical (dis-)similarity is expressed as the distance between the nodes of the 2D maps (as well as the Hit Map): the further apart two neurons are located the stronger and more significant the negative correlation and vice versa. It is also important to note here that in the two-dimensional maps (Fig 2), the *three types of societal segregation patterns* (Kersting, Meyer, Strohmeier, & Terpoorten, 2009; Wehling, 2014, 2015) in the Ruhr area become clearly evident:

- > *Nationality*: The inverted appearance of Component Maps 2c and 2d suggests that clusters in the bright orange/red area in the bottom of Component Map 2c which have high foreign population intensities are at the same time in the very dark parts of Component Map 2d characterized by below average percentage of German inhabitants.
- > *Age*: The above observation also holds true for the relationship between Component Map 2e (share of under-18-year-olds) and Component Map 2f (percentage of over-65-year-olds), although these two maps are inverted along the vertical axes.
- > *Economic power*: Here, we observe a strong South-North separation in the model space between high and low purchasing power and net rent price (see Component Maps 2a and 2b). The lower left corner is also of interest, where we observe high net rent prices but at the same time low to very low purchasing power. This area of the Hit Map pertains mostly to cluster 7 and will be further discussed in the next section.



**Figure 2. Caption** The ten SOM component maps. Values are defined as percentages (%) or indices ranging between 1 (very low) and 6 (very high). Data sources: *Microdialog* (2017); *Zensus* (2011).

**Figure 2. Alt Text** In the ten component maps brighter orange/red colors correspond to higher values and darker blue/cyan colors to lower ones, e.g. neighbourhoods located in the upper left corner are characterized by high intensities of purchasing power, net rent price, share of German population and percentage of only-elderly households

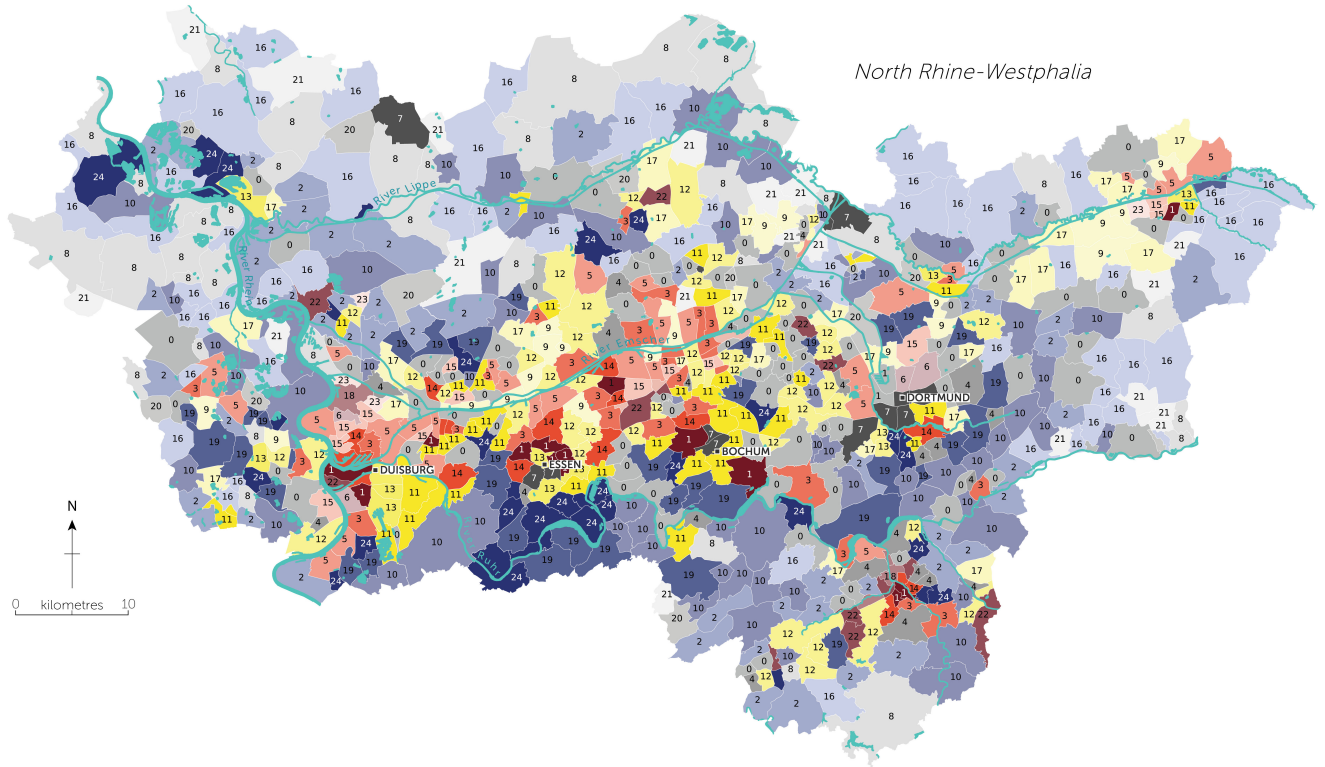
## 6.2. Clusters

Clusters of the Self-Organizing Maps are shown in Fig. 3 and they are constructed in the following way: The ten component maps (Fig. 2) are being juxtaposed and layered upon each other so that *each of the geographical neighbourhoods (remapped into the 2D model-space using the SOM algorithm) are located at the exact same point on all of the Component Maps as well as on the Hit Map*. Therefore, knowing their location, we can then easily deduce the cluster to which they belong. Moving on the surface of the Hit Map showing the clusters in Fig. 4 from right to left (i.e. from less towards more transparency) we observe increasing median age and decreasing household size, while from the bottom towards the top we encounter rising socioeconomic advantage and decreasing foreign population (from red towards blue). We should note here that Cluster 7 (bottom left corner) indicated in black on the Hit Map and not red as one might expect, is somewhat of an outlier. Cluster 7 is characterized by low purchasing power but high rental values and high share of single households, suggesting an interesting and unusual local dynamic, similar to that of commencing gentrification processes. However, justifying this assumption would require further detailed inspection that is beyond the scope of this article.

Focusing now on the distribution of clusters on the regional map shown in Fig. 3, there is an evident pattern of more red and yellow neighbourhoods in the central areas (Emscher and Hellweg zones) with grey and blue colors dominating the outer neighbourhoods. Clusters on the lower half of the output space (Fig. 4) concentrate north and south of the Emscher river (Fig. 3), justifying as well as bringing together our observations from the trend analysis discussed earlier. It is also important to note that these clusters correspond to areas where the prevalence of green and recreational land use is at its scarcest, as depicted in Figure 1, and where industrial production was most intensive regionally, mostly in the end of 19th and the first half of the 20th century.

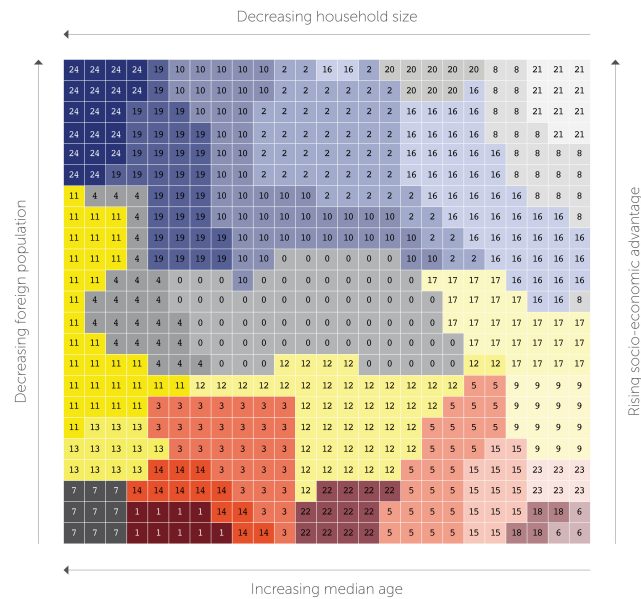
These results correspond to and underline relationships at the zonal scale: socioeconomic factors are demonstrating very clear spatial patterns, whereby more prosperous areas characterized by strong economic profiles in the outskirts surround less affluent zones that are at times gradually “shrinking” (e.g. southern Emscher zone). Thereby, it is especially the neighbourhoods south of the Emscher river that exhibit rather unfavorable values, whilst northern ones seem to perform somewhat better, but still far below the regional average. This trend that appears to start in central Oberhausen continues its course along the Emscher river to the east, up until the city of Castrop-Rauxel. The latter section of the Emscher river covers approximately the length of the “Emscher Island”, a lengthy piece of land between the Emscher and the Rhein-Herne Canal. Based on the SOM analysis, it is evident that from a regional perspective, sustained investment should be targeted in this wider area. The North-South divide along the Emscher river that we discussed so far is then being rotated (at Castrop-Rauxel) to reach the city of Dortmund as the Emscher alters its course and runs in South-North direction. Here, the more disadvantaged areas cluster to the West of the river and those with somewhat more favorable values to the East. Interestingly, the center of Dortmund is one of the few cases containing several neighbourhoods from Cluster 7. Again, the presence of cluster 7 may signal initial phases of gentrification tendencies in the vicinity of Dortmund city centre and of the Technical University Dortmund (Wittowsky, Hoekveld, Welsch, & Steier, 2020). Some final remarks on our previously defined societal segregation patterns; in general we observe more vibrant colors in the South as opposed to more fading colors in the North (with the Emscher being the major dividing element) suggesting younger populations and larger household sizes in the northern parts of the region. It is the Lippe zone in particular that appears to attract a larger share of young families with more affluent backgrounds. Secondly, we can observe that the latter mentioned center-periphery divide (both in the purchasing power and in ethnic distribution) is further intensified by enhanced small-scale polarization at the boundaries of the different zones.

It should be stressed that these findings are representative of the historical and characteristically polycentric development of the study area. Until the beginning of the 19th century,



**Figure 3. Caption** Distribution of the twenty-five socioeconomic SOM clusters over the Ruhr region

**Figure 3. Alt Text** We performed the SOM analysis for the 732 neighbourhoods or *Stadtteile*. There is an evident pattern of more red and yellow neighbourhoods in the central areas (Emscher and Hellweg zones) with grey and blue colors dominating the outer neighbourhoods.



**Figure 4. Caption** SOM Hit Map: Distribution of SOM clusters over the 2D model space.

**Figure 4. Alt Text** Moving on the surface of the Hit Map from right to left (i.e. from less towards more transparency) we observe increasing median age and decreasing household size, while from the bottom towards the top we encounter rising socioeconomic advantage and decreasing foreign population (from red towards blue). There can be multiple geographical neighbourhoods assigned to one node on the Hit Map.



small-scale industrial production took place exclusively south of the Ruhr river. The technological shift from shallow to shaft mining signaled the moment of first northwards industrial expansion from the Ruhr-zone into the Hellweg-zone (see Figure 1.d), starting with 1837. Already since the middle ages, this area was dotted with smaller cities (such as Duisburg, Essen, Bochum, Dortmund) due to the presence of an important trading route “Hellweg” (the eponym of the region). Starting in the 1870s, the unification of Germany and the expansion of the national railway system with its material-demand and its opening up of new markets gave an unprecedented impulse for basic industries. These trends, coupled with increasingly powerful extraction technologies, led to the sprawling of coal mines up until the Emscher Zone. Previously, the area on both sides of the Emscher was rather swampy, experiencing frequent flooding events and it was therefore scarcely inhabited. In the dawn of the first world war, economic recession and the resulting decrease in investment capacity led to the closing of numerous - less technologically efficient - coal mines in the Ruhr-, and Hellweg-zones. However, the reconstruction after WW2 demanded raw material and thus a powerful and efficient industrial Ruhr region, whereby now production mainly concentrated in the Emscher zone’s large scale so-called “combined facilities” (Wehling, 2014), and this remained the case until the end of the decline of basic industries in the region. Finally, from a societal perspective, it is important to mention that the Ruhr area has been attracting foreign workforce throughout all the industrialisation phases. As an example, the 1960s brought a series of recruitment agreements, e.g. with Italy, Spain, Greece, Turkey and Morocco (Schumacher, 2009). Triggered by the economic crisis in Germany and the decline in the importance of heavy industries, foreign recruitments came to a halt in 1973 (Schumacher, 2009). Nevertheless, as non-German workers living in the area had the right to bring along their families, the foreign population continued to rise (Wehling, 2014), mostly in the Emscher and less predominantly in the Hellweg zones (as also seen in Fig. 3). This brief overview demonstrates that throughout history the location of newly emerging settlements was dictated by the geomorphological distribution of coal and thus the economic, social and political influence of basic industries. (Wehling, 2015). The latter described South-North expansion of basic industries and the corresponding time lag for restructuring opportunities - discussed in detail in the paragraph “economic structural change” in Section 5 - can be still observed on the spatial distribution of social economic clusters in Figure 3.

## 7. Conclusions

Regarding data availability, the 100x100 meter cell resolution of the German census from 2011 (*Zensus*, 2011) and its rich number of demographic and household-structural parameters provide an excellent basis for further geographical analysis. However, since indicators for the economic performance of inhabitants are largely missing, we had to retrieve these from an external source (*Microdialog*, 2017). As anticipated, the two accessible *microdialog* features of purchasing power and net rent price, were strongly correlated (see also Fig. 2a and 2b). Therefore, additional information on economic performance, e.g. share of people living on social subsidies would have certainly enriched our analysis, although it is exceptionally hard to obtain such high-resolution data for Germany (Lengyel, 2021). Finally, since methods for data collection and processing differ between the two datasets, we may have introduced some level of inaccuracy to our calculations that could have been avoided with more consistent datasets.

Notwithstanding, our results demonstrate that structural inequalities stemming from the processes of industrialization and deindustrialization in the Ruhr metropolitan region (*Ruhrgebiet*) are still widely persistent and result in characteristic patterns along the three major rivers: the Lippe, Emscher and the Ruhr. For example, areas that are still facing a multiplicity of socioeconomic challenges today concentrate in the mid-reaches of the Emscher river. Moreover, the three types of societal segregation patterns and their distributions over the Ruhr area (ethnic background, age group, economic performance) became conspicuous on our maps. Our global



and local findings combined indicate that new and more diverse lifestyles and living circumstances are shriveling traditional social patterns (Lyberaki et al., 2013), posing new demands and challenges on the whole *Ruhrgebiet* region. As indicated, we expect that in the future, the share of working-age population will both decrease and grow older by a significant amount. Such shortage of labor force, which is already widely apparent, may worsen considerably whilst this particularly applies to highly qualified resources (Butzin, Pahs, & Prey, 2009) as well as to technicians and other skilled workforce. Consequently, the effective integration of young populations with and without migrational background in both education and the labour force is becoming even more pertinent. The effects of demographic change may also have widespread consequences on other urban dimensions. An interlinked loop of population-, job-, purchasing- and tax-loss may lead to difficult-to-solve local and regional problems. Accordingly, the domains of retail and services, labor market, municipal finances and housing, as well as social segregation and urban fragmentation (Butzin et al., 2009; Oueslati, Albanides, & Garrod, 2015) are all to be carefully assessed and targeted with a variety of mitigation and alleviation measures.

## Data Availability Statement

To support our results and conclusions, we provide data containing the following parameters for each of the 732 neighbourhoods analysed in this article (in *.csv* format): The  $x$  and  $y$  coordinates of the centroids (ETRS89 / UTM zone 32N), the SOM cluster-ID (ranging between 0 and 24, see Fig. 4 and Fig. 3) and we also indicate the city to which they belong (out of the total of 53 cities of the Ruhr Area). Moreover, we make the mean values of the ten variables (Fig. 2) available for the twenty-five SOM clusters individually (as *.csv*). Since the German census of 100 x 100 meter resolution follows data protection guidelines (e.g. the precision of the information is reduced by not providing decimal places (*Zensus*, 2011)), a small error margin is to be expected when averaging for the whole metropolitan level.

## References

- Agarwal, P., & Skupin, A. (2008). *Self-organising maps: Applications in geographic information science*. John Wiley & Sons.
- Aguado, D., Montoya, T., Borrás, L., Seco, A., & Ferrer, J. (2008). Using som and pca for analysing and interpreting data from a p-removal sbr. *Engineering Applications of Artificial Intelligence*, 21(6), 919–930.
- Berndt, C. (2001). *Corporate germany between globalization and regional place dependence: business restructuring in the ruhr area*. Springer.
- Birch, K., MacKinnon, D., & Cumbers, A. (2010). Old industrial regions in europe: a comparative assessment of economic performance. *Regional Studies*, 44(1), 35–53.
- Butzin, B., Pahs, R., & Prey, G. (2009). *Sektorale Wirkungen des Demografischen Wandels*. Published Online, <http://www.ruhrgebiet-regionalkunde.de>. Retrieved 2020-05-03, from <http://www.ruhrgebiet-regionalkunde.de/>
- Federal statistical office of germany. (2019). Retrieved 27. Juni 2019, from <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsvorausberechnung/>
- Kersting, V., Meyer, C., Strohmeier, P., & Terpoorten, T. (2009). Atlas der Metropole Ruhr, Vielfalt und Wandel des Ruhrgebiets im Kartenbild. In A. Prossek, H. Schneider, H. A. Wessel, B. Wetterau, & D. Wiktorin (Eds.), (p. 142-145).
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9), 1464–1480.
- Kourtit, K., Nijkamp, P., & Arribas, D. (2012). Smart cities in perspective - A comparative European study by means of self-organizing maps. *Innovation: The European journal of social science research*, 25(2), 229–246.
- Landesdatenbank NRW. (2019). Retrieved from <https://www.landesdatenbank.nrw.de>

- Lengyel, J. (2021). *Multiscale urban analysis and modeling: Trends in the Ruhr area, Germany*. Dissertation, Universität Duisburg-Essen.
- Lengyel, J., & Friedrich, J. (2019). Multiscale urban modeling; A de-urbanization scenario in the Ruhr area. *Neue Dimensionen der Mobilität: Technische und betriebswirtschaftliche Aspekte*, Springer-Verlag.
- Lyberaki, A., Tinios, P., Mimis, A., & Georgiadis, T. (2013). Mapping population aging in Europe: how are similar needs in different countries met by different family structures? *Journal of Maps*, 9(1), 4–9.
- Microdialog. (2017). Retrieved 2017-09-04, from <https://www.deutschepost.de/en/d/deutsche-post-direkt/microdialog.html>
- Moosavi, V., Packmann, S., & Vallés, I. (2014). *Sompy: A python library for self organizing map (som)*. (GitHub.[Online]. Available: <https://github.com/sevamoo/SOMPY>)
- Netrdová, P., & Nosek, V. (2016). Spatial patterns of unemployment in Central Europe: Emerging development axes beyond the Blue Banana. *Journal of Maps*, 12(4), 701–706.
- Oueslati, W., Albanides, S., & Garrod, G. (2015). Determinants of urban sprawl in European cities. *Urban Studies*, 52(9), 1594–1614.
- Regionalkunde Ruhrgebiet. (2019). Retrieved from <http://www.ruhrgebiet-regionalkunde.de/>
- Regionalstatistik Ruhr. (2019). Retrieved 23.07.2019, from <https://www.rvr.ruhr/daten-digitales/regionalstatistik/>
- Rehfeld, D., & Nordhause-Janz, J. (2017). Integration—fragmentation—reintegration? studying cluster evolution, regional path development and cluster policies in the ruhr area. In *The life cycle of clusters*. Edward Elgar Publishing.
- Rienow, A., Stenger, D., & Menz, G. (2014). Sprawling cities and shrinking regions – forecasting urban growth in the Ruhr for 2025 by coupling cells and agents. *Erdkunde*, 85–107.
- Schumacher, J. (2009). *Atlas der metropole ruhr: Vielfalt und wandel des ruhrgebiets im kartenbild*. Emons.
- Strohmeier, K. P. (2003). Bevölkerungsrückgang, Segregation und die Transformation sozialräumlicher Milieus im Ruhrgebiet. *Moving the Social*, 30, 111–122.
- Unglert, K., Radić, V., & Jellinek, A. M. (2016). Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra. *Journal of Volcanology and Geothermal Research*, 320, 58–74.
- Urban Atlas. (2019). Retrieved from <https://land.copernicus.eu/local/urban-atlas>
- Wehling, H.-W. (2014). Organized and disorganized complexities and socio-economic implications in the Northern Ruhr Area. *Understanding complex urban systems: Multidisciplinary approaches to modeling*, Springer, 87–101.
- Wehling, H.-W. (2015). Zollverein und die industrielle Kulturlandschaft Ruhrgebiet. *Auftrag der Stiftung Industriedenkmalspflege und Geschichtskultur, erweiterte Fassung*.
- Wittowsky, D., Hoekveld, J., Welsch, J., & Steier, M. (2020). Residential housing prices: impact of housing characteristics, accessibility and neighbouring apartments—a case study of Dortmund, Germany. *Urban, Planning and Transport Research*, 8(1), 44–70.
- Zensus. (2011). Retrieved 2021-07-23, from <https://www.zensus2011.de>

## Acknowledgments

J.L. acknowledges funding from the Mercator foundation within the NEMO Project and would like to thank both the project’s team and its strategic partners; the Regional Planning Association Ruhr (RVR) and the Emschergerossenschaft (EGLV) for providing extensive data as well as continuous help and feedback. J.L. is grateful for H.-W. Wehling and J. A. Schmidt for numerous fruitful discussions and for their remarks.