Random Forest Variable Importance Measures for Spatial Dynamics

Case Studies from Urban Demography

Georgati, Marina; Hansen, Henning Sten; Keßler, Carsten

Published in:
ISPRS International Journal of Geo-Information

DOI (link to publication from Publisher):
10.3390/ijgi12110460

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

International Journal of
Geo-Information

MDPI

Article

Random Forest Variable Importance Measures for Spatial Dynamics: Case Studies from Urban Demography

Marina Georgati [1,*], Henning Sten Hansen [1] and Carsten Keßler [1,2]

[1] Department of Sustainability and Planning, Aalborg University, 2450 København, Denmark; hsh@plan.aau.dk (H.S.H.); carsten.kessler@hs-bochum.de (C.K.)
[2] Department of Geodesy, Bochum University of Applied Sciences, 44801 Bochum, Germany
* Correspondence: marinag@plan.aau.dk

Abstract: Population growth in urban centres and the intensification of segregation phenomena associated with international mobility require improved urban planning and decision-making. More effective planning in turn requires better analysis and geospatial modelling of residential locations, along with a deeper understanding of the factors that drive the spatial distribution of various migrant groups. This study examines the factors that influence the distribution of migrants at the local level and evaluates their importance using machine learning, specifically the variable importance measures produced by the random forest algorithm. It is conducted on high spatial resolution (100 × 100 grid cells) register data in Amsterdam and Copenhagen, using demographic, housing and neighbourhood attributes for 2018. The results distinguish the ethnic and demographic composition of a location as an important factor in the residential distribution of migrants in both cities. We also examine whether certain migrant groups pay higher prices in the most attractive areas, using spatial statistics and mapping for 2008 and 2018. We find evidence of segregation in both cities, with Western migrants having higher purchasing power than non-Western migrants in both years. The method sheds light on the determinants of migrant distribution in destination cities and advances our understanding of the application of geospatial artificial intelligence to urban dynamics and population movements.

Keywords: residential distribution; machine learning; population dynamics; gridded data; GeoAI

check for updates

Citation: Georgati, M.; Hansen, H.S.; Keßler, C. Random Forest Variable Importance Measures for Spatial Dynamics: Case Studies from Urban Demography. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 460. https://doi.org/10.3390/ijgi12110460

Academic Editors: Wolfgang Kainz and Maria Antonia Brovelli

Received: 6 September 2023
Revised: 24 October 2023
Accepted: 31 October 2023
Published: 9 November 2023

1. Introduction

Residential location modelling plays an important role in urban planning and policy-making [1,2]. As the world's population grows and becomes more concentrated in urban centres, effective urban planning requires improved forecasting and a deeper understanding of the factors that influence residential locations. Inner city movements and residential relocations trigger urban dynamics with implications for economic development, socio-demographic structures, spatial segregation and inequalities [3]. Housing availability and prices, commuting time to work, proximity to services or recreation, and the quality of the urban environment are some of the determinants that influence our next residential choices [3]. However, which of these factors are the most important? And for whom?

Several studies have examined the interaction between purchase prices or rental costs and household income [2,4]. Other studies have analysed the importance of transport accessibility in the choice of where households without cars (or bikes) live, or in relation to access to work and commuting time [5]. Chen et al. [6] found interactions between households with children and neighbourhoods with good schools, while Mulder [7] and Mulder and Cooke [8] highlighted the role of family ties in residential location choices. Research has also identified proximity to various amenities as an important attraction factor. Examples include educational [9,10] and recreational facilities [10,11], services and retail [5], and transport systems [12,13]. Others have emphasised the role of previous residential location when it comes to moving out of or staying in the same neighbourhood [6].

However, little is known about the factors that specifically influence migrant's location choices [14,15], particularly when it comes to identifying 'place-specific attributes' [16] in destination cities and assessing their importance for certain migrant groups. This challenge is complex and multidimensional, with limited empirical research to date [14,17,18]. Its complexity is amplified in modern multicultural urban centres with increased population density and diversity [19–21].

Thus, several studies have focused on the residential distribution of migrants, examining spatial patterns of representation, segregation or integration [14,22–29], and have highlighted the need for novel theories to explain the residential outcomes of migrants [29]. They have also emphasised the influence of cultural background or diversity on migrants' residential distribution [23–25], but the specific role of the presence of co-ethnics has not yet been quantified. We use the term *co-ethnics* to refer to individuals who may belong to the same ethnic group but come from different countries and share a common cultural background, language and/or religion.

To address the above research gaps and to gain a deeper understanding of the role of ethnic composition of neighbourhoods in migrant residential location modelling, we conducted a data-driven study that related several determinants of residential preferences, including the effect of migration background. Quantifying the influence of co-ethnics involved examining how the presence of individuals from the same ethnic background affected, among other factors, the distribution of migrants.

We assessed the importance of these factors using novel machine learning (ML) techniques, in particular random forest (RF; [30]) regression on gridded population and topographic data at high spatial granularity. RF variable importance measures (VIMs; [31,32])—an innovative method previously used mainly in biological studies [33,34]—assess the importance of the examined variables and quantify their impact on model performance. While they have been incorporated into geospatial studies, their use has typically been limited to certain contexts, including disaggregation methods [35,36], poverty [37] or flood zone modelling [38]. Our analysis focused on the factors influencing the residential location of migrants in two case studies in Europe with a rich migration history (i.e., Amsterdam and Copenhagen) in 2018, and unfolded in two stages, for aggregated groups and individual countries of origin (CoOs). We also explored the distribution of migrants in relation to house prices in 2008 and 2018, and whether certain migrant communities paid more in the most desirable locations.

2. Background

For households, residential choices describe the decision-making process of choosing between housing and location options [39]. Modelling residential choices refers to maximising utility for each of its members [6]. It provides the means to explain and predict discrete choices after weighing the contributions of various measures that define the decision, such as housing and location attributes [40]. It is well known that trade-offs are usually made where the utility of one attribute is reduced to gain the utility of another (e.g., longer commuting times for lower house prices).

Several variables determine households' residential location choices [3], and many models have been developed to explain the relationships among them [41]. Although most studies provide a broad overview of the characteristics that influence household decisions, they do not address differences among migrant groups and natives. These differences in housing and location choices reflect spatial, social and economic conditions, including segregation and discrimination. They are the result of complex urban dynamics and are explained by theories such as spatial assimilation [42–44], place stratification [45,46], ethnic preferences [47], housing career and structural changes in the economy and labour market [48].

Spatial assimilation theory suggests that recently arrived migrants prefer affordable, diverse, potentially disadvantaged locations that already accommodate a high proportion of migrants. These new migrants adopt local attitudes, integrate over time, establish their socio-economic position and move to more desirable areas with larger native populations [48–50]. In contrast, according to place stratification theory, migrant groups face challenges in assimilating due to discrimination, especially in neighbourhoods dominated by the host population [46,47]. The theory of ethnic preference suggests that not all migrants seek spatial assimilation; they may instead prefer to live close to co-ethnics, even if they could afford more desirable neighbourhoods [47,51]. This decision is influenced by the benefits of maintaining cultural traditions, forming strong social connections and having access to ethnic shops and services.

In studies across northern European countries, findings based on different types of regression models are consistent with spatial assimilation theory [50–53]. While these studies highlight the importance of ethnic composition in shaping migrants' residential decisions, they primarily use discrete choice models to measure the impact of different factors on choice probabilities. This modelling approach is confined to the use of survey data in which respondents navigate hypothetical situations and has limitations related to survey design, respondent honesty and the strong assumptions they often make. It also requires data over multiple time periods and is not suitable for capturing complex interactions between factors.

Recently, there has been a growing trend in the use of geocoded, register and census population data in segregation analysis. These applications cover a wide methodological range, including sequential studies [47], longitudinal analyses [28,54], the use of indices [55,56] and the adoption of the individualised neighbourhood concept [57–60]. Notably, the latter examples simultaneously address issues related to the modifiable aerial unit problem (MAUP; [61,62]), which has raised difficulties because the spatial units for which segregation is measured influence segregation outcomes [62,63].

Given the emerging demand for new theories and models to explain migrant residential outcomes [29], it is essential to quantify the specific impact of ethnic composition on residential modelling. This requires a deeper understanding of which ethnic groups primarily influence each other. To this end, our study departed from examining residential trajectories and changes in segregation over time. Instead, we focused on a specific time period to explore the variables in detail for the current situation, as a result of years of residential mobility up to that point. We proposed the use of innovative ML modelling on fine-grained register data, which can effectively handle large datasets and capture complex relationships between multiple factors. This allowed us to analyse, in depth, the factors involved and the population dynamics.

3. Materials and Methods

3.1. Case Studies and Population Datasets

The analysis focused on the foreign populations residing in the municipality of Amsterdam and the capital region of Copenhagen. The population datasets describe the distribution of the population by CoO in 2008 and 2018 at a resolution of 100 m. The data were provided by the municipality of Amsterdam (Onderzoek, Informatie en Statistiek (OIS; https://data.amsterdam.nl) and Statistics Denmark (DST; https://www.dst.dk) and are not publicly available as they contain sensitive information in sparsely populated areas. The spatial representation of migrants varies between the two cities (Table A1). In 2018, migrants occupied 90% of the inhabited cells in Amsterdam and 78% in Copenhagen. In both cities, only a negligible proportion of cells was inhabited exclusively by migrants.

We examined the residential distribution of both aggregated groups and individual CoOs. In total, 304 CoOs in Amsterdam and 178 CoOs in Copenhagen were aggregated into seven regions of origin (RoOs). RoOs refer to broader geographical areas from which individual CoOs were grouped on the basis of shared characteristics, such as geography or culture. The classification (Table A2) differs between the cities because of their different histories and settings as migration destinations; Amsterdam, with its colonial past, has long been an established migration hub for many nationalities; Copenhagen, following restrictive policies, has been a less accessible destination. The aggregation into groups considered these differences, the social discourse and the largest migrant groups in each case. For reference, the lower part of Tables A3 and A4 in the appendix show the statistical summary of the migrant groups by RoO and CoO used in Amsterdam in 2018.

3.2. Selected Determinants

We adopted the classification of determinants proposed by Schirmer et al. [3] into two broad categories related to building characteristics (e.g., value of residence, size) and location/neighbourhood attributes (e.g., proximity to services, points of interest). In addition, we included socio-demographic attributes describing the population by age, income and education, where data were available from the corresponding statistical offices for 2018.

Case-specific layers were produced at a fine-grained resolution, considering the main factors influencing the residential choices of different demographic groups. Most of these layers were common to both cases, but additional socio-demographic variables were included for Copenhagen. For the analysis at CoO level, we included data on the distribution of migrants by CoO where the population was greater than 50 persons. We used the same socio-demographic, housing and neighbourhood factors but replaced the two generic layers for places of worship (i.e., Christian, non-Christian) with religion-specific layers. In Amsterdam, there were 110 places of worship—2 for Buddhism; 7 for Islam; 3 for Sikhism; 6 for Judaism; 7 for Hinduism; and 86 for Christianity (5 classified as Catholic, 1 as Lutheran and 27 as Protestant churches [64]). In Copenhagen, there were 217 churches [65], 1 place of worship for Buddhists and 2 for Muslims [64]. Table 1 describes the layers, their sources and their processing details, while Table A3 shows the statistical summary of the corresponding layers in Amsterdam in 2018.

Table 1. The selected determinants with their data sources and description of the pre-processing. The data are provided for 2018, with real estate prices also provided for 2008.

Determinants	Description	Data Source
Demographic attributes		
Age	The total population divided into five groups: children, students, mobile adults, non-mobile adults and the elderly.	Amsterdam: OIS, Copenhagen: DST
Marriages	The number of marriages.	Copenhagen: DST
Educational attainment	The share of people with tertiary education.	Copenhagen: DST
High and low income	The population in the highest and lowest decile of the yearly equivalent disposable income.	Copenhagen: DST
Building attributes		
Building height and construction year	The average height and year of construction of the buildings in the corresponding grid cell.	Amsterdam: PDOK [66], Copenhagen: Bygnings-og Boligregistret (BBR) [65]
Dimensions, usage and ownership status	The total number of residential dwellings and rooms, the size of residential area in m^2, the share of unoccupied or rented units and the share of private or public year-round dwellings.	Amsterdam: OIS, Copenhagen: DST
Neighbourhood attributes		
Real estate prices	For Copenhagen, the layer represents the average purchase price per square metre of the dwellings sold. For Amsterdam, the property value map is the output of an interpolation process disaggregated to grid cells from the neighbourhood level, taking the average value in the provided range.	Amsterdam: Gemeente Amsterdam [67], Copenhagen: Bygnings-og Boligregistret (BBR) [65]
Proximity to bus stops, train stations, schools, universities, leisure and cultural facilities (e.g., restaurants, cinemas), and places of worship by religion	The total number of accessible services/facilities within walking or cycling distance.	Amsterdam: OpenStreetMap contributors [64], Copenhagen: OpenStreetMap contributors [64], Bygnings-og Boligregistret (BBR) [65], Movia Trafik [68]

3.3. Random Forest Variable Importance Measures

Different demographic groups are influenced by a range of socio-economic and urban variables. We aimed to assess the relationships between migrants from different CoOs/RoOs and the distribution of one group as a determinant of the residential distribution of others, together with population and topographic characteristics. The RF algorithm was applied because it is more suitable for capturing complex relationships between dependent variables, such as different population groups, than the linear and logit models used in previous studies [50,52,53]. These models are oversimplified and unsuitable for detecting such patterns, especially when a much wider range of factors is examined; they are prone to underfitting, sensitive to outliers and assume that the data are independent—which is not the case for the data at hand.

Furthermore, the RF regression algorithm can capture non-linear relationships among the investigated factors. The inclusion of non-linear associations is essential for our study as migration patterns and residential locations are often the result of complex interactions of socio-economic and cultural factors, which may not always follow a strictly linear trajectory. In this way, we examine how strong the links among various migrant groups are in supporting the relocation of others. Unlike other studies that examine mobility patterns in multiple time steps, we take a snapshot of the year in question (2018) and conduct a data-driven study to examine the outcomes of years of spatial mobility. This can serve as a baseline for a future longitudinal study, allowing us to highlight existing patterns and explore, in depth, the variables at that particular point in time.

The RF, an '*ensemble learning*' technique, combines multiple decision trees and aggregates their predictions for enhanced accuracy [30,69,70]. It supports the interpretability and identification of informative variables [31,71] through VIMs—built-in functions that rank the training features in terms of prediction accuracy [32]. By ranking the contribution of the training variables, they capture the strength of the dependency between features and predictions [32,72].

Mean decrease in impurity (MDI) and mean decrease accuracy (MDA) are two commonly used RF VIMs supported in Python and the Pedregosa et al. [73] implementation. While MDA is typically more reliable [74], being unbiased to variables with numerous split points, it can overvalue highly correlated variables and, for high-dimensional data, its rankings are less robust to perturbations in the data than those obtained with the MDI [72,75]. For our analysis, given continuous data layers with potential correlations, we chose the MDI score.

Since the distribution of population groups was available, we employed a supervised learning framework to make predictions. The training data consisted of the layers of demographic, building and neighbourhood attributes, as well as data on the distribution of other migrant groups. These inputs were paired with the target values representing the specific migrant group under consideration. For model training, we only used data from 2018, which provided a more concise and detailed perspective on the variables of interest. During training, we retrieved the importance of each layer in predicting the target variable and distinguished the stronger relationships between migrants based on their RoO and then by individual CoO.

Our goal was not necessarily to develop the most accurate predictive model but to use it as a tool to explain a complicated phenomenon. Therefore, after experimenting with different hyperparameters (i.e., the number of trees, the maximum depth) to ensure the model's robustness and consistency, we adopted a simple RF implementation based on Python frameworks (i.e., scikit-learn; [73]). We used a small number of trees as the relative value of feature importance did not change with the depth of the model. We allocated 25% of the data for testing and used the R-squared value to evaluate the prediction performance. The variable importance scores were extracted into EXCEL tables and ranked by determinant category. The CoO scores were further related to the corresponding population sizes and visualised using the Matplotlib [76] and Seaborn [77] libraries after eliminating scores lower than 7%.

3.4. Migrant Distribution and Real Estate Prices

We also investigated the relationship between the distribution of migrants and house prices. We examined whether migrants bought more expensive properties in areas that were more attractive to them or where their residential density was higher. We estimated the sum of purchase prices by migrant group per grid cell along the inhabited cells as described in Equation (1), where i is the individual grid cell, N_{im} is the migrant population in cell i, N_m is the total migrant population in the examined area and P_i is the real estate price in the individual grid cell. The bivariate maps show the distribution of the share of migrants in the total population in relation to house prices (see Figures 4 and 5). That way, we explored the proportions of migrants living in zones of expensive/inexpensive housing and how these tendencies were represented spatially in each city.

For validation purposes, we compared two different time periods from 2008 and 2018. As the Danish property value layer represents the sales prices of only those properties sold in the respective year, we aggregated the gridded layer to the district level and selected the average value among the purchased properties. In Amsterdam, we used the mean real estate prices interpolated to contours. Due to the differences between the original datasets and the studied areas—in Amsterdam we focused on only one central municipality, while in Copenhagen we included two central and 15 suburban municipalities—we expected variations, with central Amsterdam being more expensive and Copenhagen having lower average house prices due to the cheaper suburbs. For this reason, the comparison between the two cities cannot be direct.

$$P_m = \sum_{i=1}^{N_I} \left(\frac{N_{im}}{N_m} \times P_i \right) \tag{1}$$

4. Results

4.1. Results on Aggregated Groups

Tables 2 and 3 show the results of the aggregated groups by RoO in Amsterdam and Copenhagen with 29 and 34 determining factors, respectively. Each column refers to a target RoO with the importance of the factor as a percentage, and each row contains the factor. The italicised headings group the factors into classes (i.e., age groups, building and neighbourhood features, and the presence of native/migrant groups) and summarise their individual importance. For example, in Table 2, the distribution of children receives a high importance score for predicting the distribution of Turks and Moroccans (66.12%) but less than 1% for predicting the distribution of natives and of Western and Eastern EU migrants. The R-squared value in the last row assesses the performance of the regression model.

In the Dutch case, dwelling features are most important for natives and migrants from the Western EU (WEU) and the rest of the world. For migrants from the Eastern EU (EEU), Turkey and Morocco, the distribution by age is the most important. For migrants from the former colonies, the Middle East and Africa (MEA), the presence of other migrant groups ranks higher; each of these groups plays a significant role in the distribution of the other. The relationship between WEU migrants, natives and the rest of the world population is noteworthy. More details on specific groups by CoO follow in Section 4.2. Neighbourhood attributes are less important for migrants from former colonies, Turkey and Morocco.

In the Danish case, building characteristics are again the most important features for the native and WEU population. Age groups are the most influential factors for migrants from EEU, MENAP, non-EU European and other non-Western countries. Other Western migrants are significantly influenced by the presence of WEU migrants. The distribution of non-EU European and MENAP migrants is closely related to the presence of other non-Western migrants and vice versa. Information on the distribution of the population by social status has a limited impact on the residential locations of MENAP and other non-Western migrants. Finally, neighbourhood characteristics have little effect on other Western migrants, who may be more attracted to city centre leisure activities, such as restaurants and bars, than other groups.

Table 2. Importance of determinants of migrants' residential locations by RoO in percent ratio in Amsterdam.

Determinants	**Natives**	**Former Colonies**	**EU West**	**EU East**	**Middle East + Africa**	**Turkey + Morocco**	**Rest**
Age groups	*4.11*	*12.63*	*9.19*	***78.31***	*24.85*	***69.61***	*22.44*
Children	0.25	1.60	0.35	0.96	11.00	66.12	1.90
Young adults	0.30	0.75	6.35	51.59	8.18	1.72	11.91
Mobile adults	0.35	1.51	2.34	24.69	1.35	0.18	8.23
Non-mobile adults	3.12	8.45	0.09	0.31	3.77	0.96	0.30
Elderly	0.09	0.32	0.06	0.76	0.55	0.63	0.10
Building features	***92.63***	*10.17*	***68.02***	*5.92*	*7.25*	*7.02*	*62.22*
Residential area	72.57	1.06	61.88	0.41	0.50	0.17	45.89
Number of dwellings	4.58	4.34	3.62	1.25	0.80	0.10	14.34
Number of rooms	15.13	3.56	0.15	0.83	0.48	1.23	0.34
Building height	0.01	0.01	0.15	0.90	2.15	0.03	0.63
Building oldness	0.24	0.31	0.25	0.43	0.55	0.26	0.54
Not used	0.00	0.06	0.47	0.49	0.09	0.01	0.07
Rented	0.01	0.83	0.22	0.27	1.61	4.84	0.23
Private	0.09	0.00	1.28	1.34	1.07	0.38	0.18
Neighbourhood features	*0.87*	***11.37***	*2.84*	*3.78*	*1.85*	***15.38***	*0.67*
Real estate prices	0.65	1.25	0.16	0.88	0.20	0.11	0.02
Prox. to supermarkets	0.01	0.02	0.12	0.17	0.13	0.40	0.13
Prox. to restaurants and bars	0.02	0.08	0.49	0.47	0.18	0.50	0.18
Prox. to cultural spaces	0.00	0.01	0.02	0.03	0.03	0.02	0.04
Prox. to bus stops	0.00	0.00	0.03	0.20	0.21	0.08	0.06
Prox. to train stations	0.02	5.03	0.02	0.44	0.07	11.28	0.02
Prox. to schools	0.03	0.11	0.54	0.66	0.61	0.44	0.10
Prox. to university	0.00	0.00	0.22	0.02	0.00	0.00	0.02
Prox. to a place of Christian worship	0.04	0.03	1.19	0.20	0.08	0.16	0.06
Prox. to a place of non-Christian worship	0.10	4.84	0.05	0.71	0.34	2.39	0.04
Presence of natives/migrant groups	*2.39*	***65.83***	*19.96*	*11.98*	***66.1***	*7.95*	*14.67*
Natives		0.44	8.60	1.05	1.47	6.73	0.48
Former Colonies	0.45		0.26	0.14	61.40	0.63	0.21
EU West	1.06	0.09		3.14	0.97	0.10	13.19
EU East	0.16	0.05	0.63		0.18	0.05	0.36
Middle East + Africa	0.40	62.15	0.05	0.16		0.30	0.24
Turkey + Morocco	0.21	2.06	0.09	1.87	1.44		0.19
Other Europe etc	0.11	1.04	10.33	5.62	0.64	0.14	
R-squared	0.938	0.804	0.825	0.526	0.736	0.792	0.663

The italicised text indicates the classes of factors, such as age groups, building/neighbourhood characteristics and the presence of native/migrant groups, summarising their importance and highlighting the two highest scores per class in bold.

Table 3. Importance of determinants of migrants' residential locations by RoO in percent ratio in Copenhagen.

Determinants	**Natives**	**EEU**	**WEU**	**MENAP**	**Non-EU Europe**	**Other Non-Western**	**Other Western**
Age groups	*1.31*	***61.79***	*27.92*	*52.05*	***55.25***	*47.69*	*6.17*
Children	0.25	0.56	1.76	46.05	8.57	2.80	1.09
Young adults	0.32	42.84	15.91	3.67	5.55	7.07	2.30
Mobile adults	0.30	18.00	7.45	0.78	40.61	35.40	1.05
Non-mobile adults	0.44	0.26	2.72	1.54	0.46	2.39	1.28
Elderly	0.00	0.13	0.08	0.01	0.06	0.03	0.45

Table 3. *Cont.*

Determinants	Natives	EEU	WEU	MENAP	Non-EU Europe	Other Non-Western	Other Western
Social features	*0.45*	*3.57*	*11.4*	***25.6***	*4.2*	***25.87***	*6.23*
Marriages	0.11	0.14	1.44	0.52	0.16	2.48	1.05
Higher education	0.14	0.16	4.75	18.58	0.57	0.78	2.67
High income	0.01	0.11	0.78	0.43	0.35	0.01	1.00
Low income	0.19	3.16	4.43	6.07	3.12	22.60	1.51
Building features	***96.42***	*2.8*	***43.85***	*1.59*	*4.74*	*0.6*	*9*
Residential area	66.55	0.13	42.47	0.10	0.34	0.04	2.69
Number of dwellings	0.75	1.34	0.22	0.07	1.50	0.21	1.14
Number of rooms	28.98	0.08	0.34	0.01	1.44	0.03	1.36
Number of floors	0.00	0.04	0.03	0.07	0.16	0.06	0.35
Building oldness	0.01	0.57	0.38	0.95	0.37	0.09	0.80
Not used	0.10	0.17	0.16	0.07	0.08	0.05	1.36
Rented	0.00	0.21	0.14	0.12	0.49	0.08	0.50
Private	0.03	0.26	0.10	0.20	0.36	0.04	0.64
Public	0.00	0.00	0.01	0.00	0.00	0.00	0.16
Neighbourhood features	*0.02*	***4.98***	*1.07*	*0.61*	*2.63*	*1.02*	***8.65***
Real estate prices	0.00	0.05	0.02	0.07	0.03	0.01	0.29
Prox. to supermarkets	0.00	1.03	0.12	0.04	0.08	0.04	0.65
Prox. to bus stops	0.00	0.03	0.04	0.02	0.13	0.01	0.91
Prox. to train stations	0.00	0.60	0.07	0.05	0.98	0.06	0.53
Prox. to cultural spaces	0.00	0.67	0.21	0.06	0.20	0.18	0.28
Prox. to schools	0.00	0.98	0.09	0.13	0.46	0.14	0.65
Prox. to places of leisure	0.00	0.15	0.06	0.09	0.40	0.17	0.27
Prox. to bars and restaurants	0.02	0.25	0.38	0.09	0.15	0.13	4.56
Prox. to a place of Christian worship	0.00	1.18	0.05	0.05	0.18	0.28	0.48
Prox. to a place of non-Christian worship	0.00	0.04	0.03	0.01	0.02	0.00	0.03
Presence of natives/migrant groups	*1.8*	*26.84*	*15.76*	*20.15*	***33.17***	*24.83*	***69.94***
DNK		1.39	2.04	1.76	0.67	0.77	2.14
EEU	0.26		3.38	0.56	0.94	5.32	0.60
WEU	0.31	2.30		0.61	0.88	3.44	63.93
MENAP	0.48	0.28	2.06		25.47	14.52	0.14
Non-EU Europe	0.30	0.77	2.07	0.91		0.66	0.61
Other Non-Western	0.26	22.03	0.98	16.21	5.12		2.52
Other Western	0.19	0.07	5.23	0.10	0.09	0.12	
R-squared	0.961	0.517	0.682	0.785	0.457	0.666	0.552

The italicised text indicates the classes of factors, such as age groups, building/neighbourhood characteristics and the presence of native/migrant groups, summarising their importance and highlighting the two highest scores per class in bold.

4.2. Results on Individual CoO

The previous section showed that in Amsterdam migrants from the former colonies are influenced by the presence of MEA migrants (62.15%) and vice versa with a similarly high weight (61.40%). In Copenhagen, the presence of WEU migrants is primarily important (69.94%) for other Western migrants. This subsection focuses on individual CoOs; we examine whether specific migrant groups explain the above-mentioned trends and how binational relations are expressed in the destination cities.

4.2.1. The Impact of Natives and Other Migrant Groups

Most groups examined by CoO (≈85%) in both cities have an accumulated influence of more than 50% from the presence of other migrant groups. The strongest interactions (>7% importance) between migrants by CoO are shown in Figures 1 and 2. The figures reveal important relationships among migrant groups and patterns that are not clear from the previous tables. They show relationships among nations in geographical proximity and with shared cultural backgrounds, traditions and languages. These links are manifested beyond the borders of their geographical region through their residential locations in the destination cities to which they migrated in their current country of residence, regardless of whether their choices were made consciously or not.

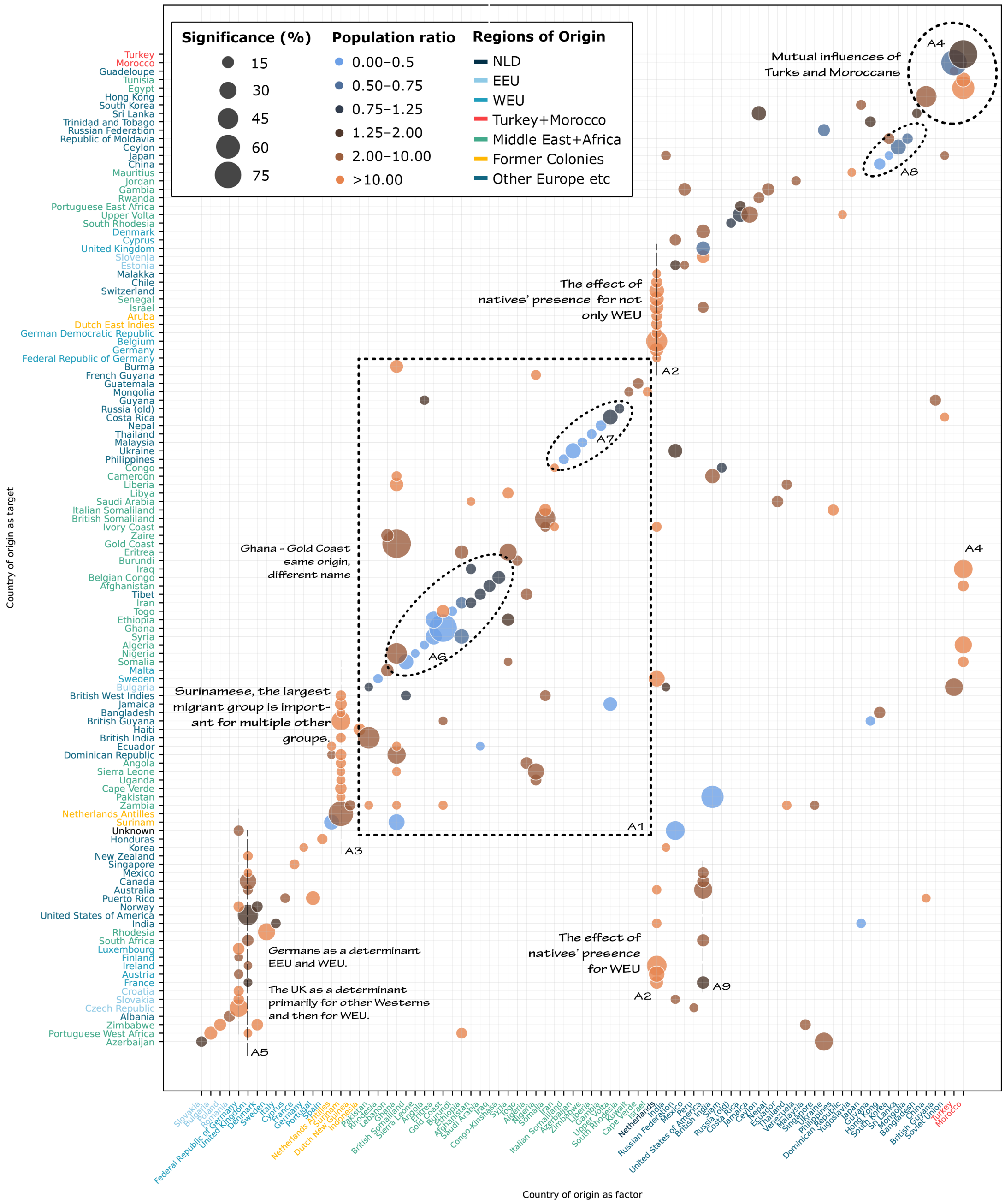


Figure 1. Feature significance by CoO and proportional population size in Amsterdam (>7%). The figure shows the influence of each source-CoO on the horizontal axis on the target-country on the vertical axis in relation to the ratio of their population sizes. The size and colour of the bubble indicate the value of the influence and the relative size of the two populations. We notice that i. the most influential pairs are migrants from Ghana–Gold Coast, Turkey–Morocco, Suriname–Netherlands Antilles, and ii. larger migrant groups from Western and EU countries have a large influence on smaller ones from the same regions.

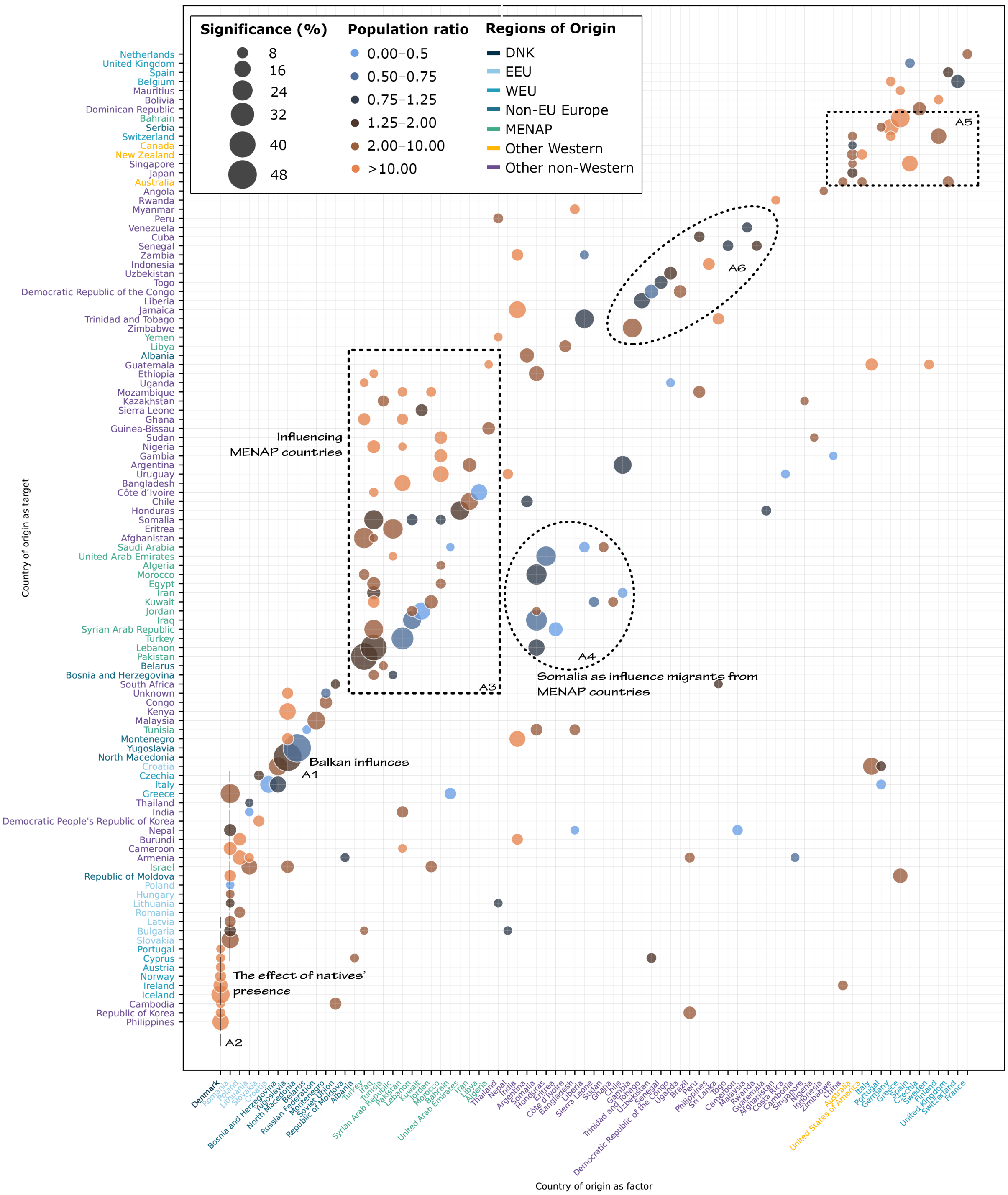


Figure 2. Feature significance by CoO and proportional population size in Copenhagen (>7%). As in the previous figure, we notice that i. the most influential pairs are migrants from Yugoslavia–North Macedonia, Iran–Lebanon, Turkey–Pakistan, and ii. migrants from the MENAP region have a high influence on migrants from other MENAP and non-Western CoO.

Specifically, the figures illustrate the influence of one CoO on another as a function of their population size ratios. The factor-country and the target-country are plotted on the horizontal and vertical axes, respectively. On both axes, the countries are coloured by their RoO. A high vertical concentration of bubbles indicates the influence of the factor on several other groups. The size of the bubbles illustrates the importance of the factor-country on the horizontal axis in predicting the population of the target-country on the vertical axis; the larger the bubble, the greater the importance. Their colour indicates the ratio of the population of the factor-country to the target-country; oranges indicate that the factor-population is larger than the target-population, and blues indicate that the factor-population is significantly smaller than the target-population.

The size of the population is important in understanding whether large groups have a large or low impact on smaller groups. The presence of large groups indicates established centres and results in diversity, safety or familiarity, and ready-made facilities that consequently attract multiple smaller groups. Conversely, if small groups are more influential on equal-sized or larger groups, family ties, developed bonds or cultural relationships are more likely reasons to explain these residential locations.

I. Amsterdam.

In Figure 1, migrants from Ghana or its predecessor state (Gold Coast) have the strongest influence on each other, followed by another African state. Ghana is also an important determinant for other Africans, Latin Americans and Caribbeans (e.g., migrants from Nigeria, Dominican Republic, Suriname, Ecuador, Sierra Leone, Liberia and Cameroon; A1). The largest in population groups—the natives, Surinamese and Moroccans—have a medium to high impact on several groups. Specifically, the presence of the natives is an attractive factor primarily for migrants from the EU (e.g., Belgians, Irish, Swedes), Senegal and Israel (A2); the Surinamese are important for migrants from other former colonies, the Caribbean and Africa (A3); the Moroccans are important for other MEA migrants (A4).

In the highlighted rectangle of the graph, a list of MEA countries appears to influence migrant groups from the same region or from countries in the rest of the world (A6, A7), with interesting interactions between populations of the same size (e.g., Upper Volta–Costa Rica, Iraq–Afghanistan), or with origins with geographical or cultural proximity (e.g., Somalia–British Somalia, Azerbaijan–Ukraine, and UK–USA–Canada). Although the populations are small, the model identifies links between migrants from Sri Lanka and its predecessor Ceylon—A8.

II. Copenhagen.

The links between migrant groups in Copenhagen are much smaller in number and strength than in Amsterdam (Figure 2). Yugoslavs are influenced by the presence of North Macedonians and vice versa (51–53%; A1). Another Balkan relationship is noticed between Bosnians and Croats (A1), whose distribution is influenced by the presence of Italians. Natives are again an influential factor for WEU (e.g., Icelanders, Irish, Norwegians and Austrians; A2) and some non-Western migrants (e.g., Filipinos). Romanians are important for other EEU Europeans (e.g., Slovakians, Bulgarians), and neighbouring Moldovans. Apart from their strong internal impact, non-EU Europeans are important for the distribution of MENAP migrants (e.g., Yugoslavs–Israelis, Belarusians–Tunisians, and Bosnians–Syrians).

In MENAP relations, reciprocal links are noted both internally and with non-Western migrants (A3). The presence of Iraqis and Somalis influences the largest number of target-CoO with smaller or almost equal populations (A4). Pakistanis and Afghans are influenced by the presence of Turks (30–48%); Lebanese, Somalis and Syrians by Iraqis; Moroccans, Lebanese and Iraqis by Somalis; Eritreans by Syrians; and Jordanians by Kuwaitis. There are strong ties between African and Caribbean migrants, with the following indicative examples: Zimbabwe–Gambia, Libya–Côte d'Ivoire, and Trinidad and Tobago–Liberia (A6).

Given the high impact of the WEU on other Western migrants in Table 3, the Australians confirm the findings (A5). Although the residential area receives the highest score, it is followed by the presence of migrants from the UK, the USA, China and Belgium. The presence of Australians is significant for migrants from Switzerland, Canada, Singapore and Japan. Similar trends can be seen for Canadians, who are influenced by the presence of Australians and French.

4.2.2. The Impact of Other Factors

Regarding the influence of other factors, in Amsterdam age-related variables have a strong influence (30–76%) on migrants from another EU country (e.g., Poland, Germany, Romania, Portugal, Italy, Greece and Spain—PIGS), Indonesia, Dutch East Indies, Brazil, Vietnam, India, Yugoslavia and Czechoslovakia. The distributions of young and mobile adults are the most important.

In Copenhagen, the impact of age is low, apart from Portugal, where the non-mobile adults layer gains the highest score. The importance of higher education for the Turkish population—the largest migrant group—explains the high significance in the MENAP group in the aggregated analysis of Section 4.1. High educational attainment also impacts migrants from New Zealand, Montenegro and France, while the low-income layer has impact on people coming from Azerbaijan, Angola and the Philippines.

In addition to the native Dutch population, migrants from the Soviet Union, the UK, France, Brazil and Belgium are strongly influenced by the size of the residential area; the number of dwellings is important for migrants from Iran, Colombia, Thailand and Chile. Similarly, in Copenhagen, the residential area is significant for the native, British, Swedish, German, Russian, French, Brazilian, Dutch and American populations, while the number of dwellings influences migrants from Thailand, Poland, Iran, China and Vietnam.

Bearing in mind the housing career theory and the fact that some studies have linked migrants' housing choices and segregation to housing markets and tenure [78–81], we find no evidence that ownership status is exceptionally important for any of the examined groups. The rented properties layer has low importance for some MEA and non-EU European groups in Amsterdam and only for Yugoslavs in Copenhagen.

The impact of neighbourhood attributes on migrants from individual CoOs is negligible in both cases. House prices have low significance for people from Israel, the USA, Bolivia, Suriname and Burundi in Amsterdam and from Trinidad and Tobago, Taiwan, Mozambique and Luxembourg in Copenhagen.

4.3. Real Estate Prices by Migrant Group

The fact that the neighbourhood attributes, and especially the value of the residence, rank so low in their influence for all migrant groups is unanticipated and raises questions. We further investigate whether migrants pay higher prices to buy property in the most appealing areas where diversity or the migrant concentration of co-ethnics is higher.

Table 4 presents the sum of purchase prices by migrant groups in the occupied cells, as estimated by Equation (1). It also shows the difference and the percentage change between the two time periods for each group in the right columns. We can immediately recognise the dramatic increase in the property market in both cities over the decade under consideration, ranging from 1600 to 2230 €/m^2 group and constituting a 41–56.5% increase in Amsterdam. The percentage increase in the Danish capital was lower: 32–42.5% and 7500–11,100 DKK/m^2 group (≈1000–1500 €/m^2 group). We highlight in bold the two groups paying the highest prices for property.

In Amsterdam, these groups are, in order, the WEU, the natives, the group coming from the rest of the world, and the EEU migrants in both examined years. Meanwhile, migrants from the former colonies and the MEA region seem to purchased property at the lowest prices. Although Turks and Moroccans had slightly higher purchasing power than the above groups, they, along with migrants from the rest of the world and the MEA region, showed the largest percentage change over the decade.

In the Danish capital, the patterns are similar; WEU and other Western migrants had the highest purchasing power, followed by the native population and non-Western migrants. EEU and non-EU European migrants paid similar amounts. MENAP migrants paid the least and, together with EEU migrants, showed the smallest percentage change over the decade.

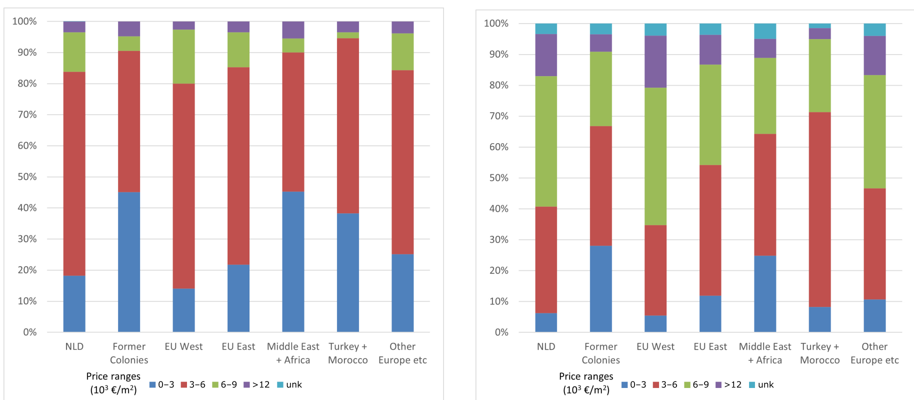
Table 4. Total purchase prices by migrant group in 2008 and 2018 in Amsterdam and Copenhagen.

City	RoO	Price per Square Meter 2008	Price per Square Meter 2018	Difference	Percentage Change
Amsterdam (€/m^2)	NLD	**3999.15**	**6198.26**	**2199.11**	54.99
	Former Colonies	3087.10	4687.16	1600.06	51.83
	EU West	**4292.56**	**6521.69**	**2229.13**	51.93
	EU East	3875.81	5501.21	1625.40	41.94
	Middle East + Africa	3025.40	4728.88	1703.45	**56.31**
	Turkey + Morocco	3112.10	4857.10	1745.00	**56.07**
	Other Europe etc	3801.01	5898.65	2097.64	55.19
Copenhagen (DKK/m^2)	DNK	23,963.24	33,144.63	9181.39	38.31
	EEU	23,895.33	31,706.37	7811.04	32.69
	WEU	**25,494.21**	**36,254.49**	**10,760.28**	**42.21**
	MENAP	21,687.94	29,275.37	7587.43	34.98
	Non-EU Europe	22,473.48	31,272.98	8799.50	39.16
	Other Non-Western	23,595.73	32,188.45	8592.72	36.42
	Other Western	**26,554.56**	**37,619.51**	**11,064.95**	**41.67**

The bold indicates the two highest values per column.

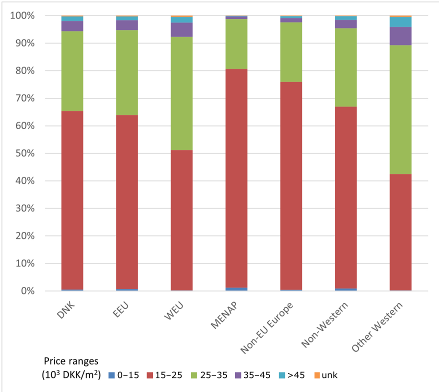
The stacked bar charts in Figure 3 display the shares of migrants and house prices by migrant group. On the horizontal axis, the graphs show the migrant groups, and each bar is scaled to 100% for the total population of the group. The colours illustrate the house price categories. The price ranges are measured in thousands €/m^2 and DKK/m^2 in Amsterdam and Copenhagen, respectively. For reference, the average sales price of existing residential property in 2021 was 3250 €/m^2 in the Netherlands and 20,180 DKK/m^2 in Denmark [82].

The charts illustrate that large proportions of all groups moved upwards at least one price category. In both cases, most natives and migrants lived in areas with house prices in the two lowest price intervals in 2008 (<6000 €/m^2 and <25,000 DKK/m^2). On the contrary, in Copenhagen in 2018 the vast majority of the population belonged to the three highest intervals. The shares of the population paying more than 35,000 DKK/m^2 increased exponentially (purple and light blue bars in Figure 3c,d). In 2018, half of the Western and WEU migrants lived in areas with house prices between 35,000 and 45,000 DKK/m^2, and 10% of them in even more expensive areas. Most migrants from MENAP and non-EU European countries lived in the areas with the lowest house prices.

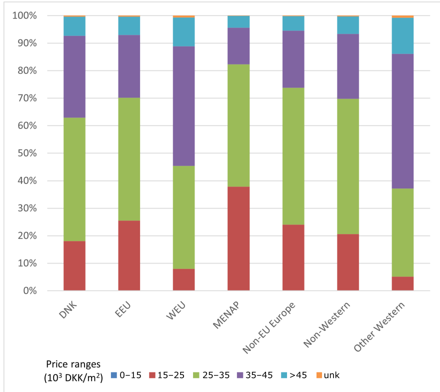


(**a**) Amsterdam, 2008. (**b**) Amsterdam, 2018.

Figure 3. *Cont.*



(**c**) Copenhagen, 2008.



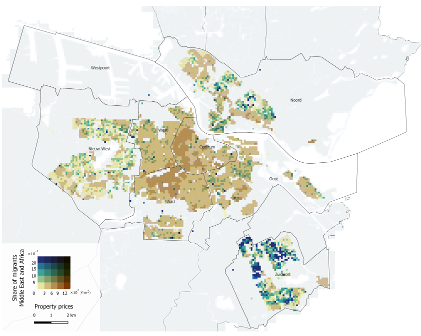
(**d**) Copenhagen, 2018.

Figure 3. Shares of migrants and house prices by migrant group in Amsterdam (**top**) and in Copenhagen (**bottom**) in 2008 (**left**) and 2018 (**right**). The graphs show the share of migrants in the corresponding group residing in areas with different ranges of house prices, indicated by the different colours. Prices are given in €/m^2 for Amsterdam and in DKK/m^2 for Copenhagen.

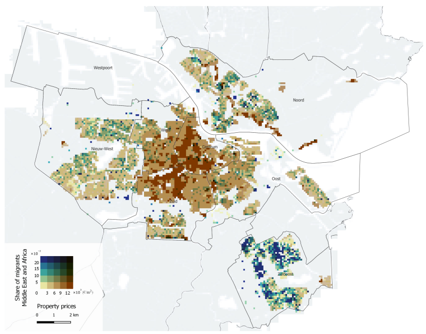
In Amsterdam, the changes were not as radical, but variations are observed among the groups. Two-thirds of migrants from the former colonies, the MEA region, Turkey and Morocco would still pay less than 6000 €/m^2 in 2018, while at the top of the scale, most natives and migrants from the WEU and the rest of the world would pay more than 6000 €/m^2.

The bivariate maps in Figures 4 and 5 illustrate these patterns and trends spatially. Each figure shows the distribution of migrants with the highest and lowest purchasing power, as reported in Table 4, as a percentage of the total population. The expensive areas are shown in dark brown in central Amsterdam and central and north-eastern Copenhagen. Dark blue cells display areas with high proportions of the corresponding migrant group and low real estate prices. Similarly, light yellow cells show areas with low house prices and low shares of migrants.

Figure 4a,b show that only a small share of MEA migrants lived in the central part of Amsterdam in both 2008 and 2018. The majority of this population lived in Zuidoost—the south-eastern exclave of the municipality with the largest community of migrants from Suriname [83,84] and apparently from MEA—Noord and Nieuw-West. In these regions, house prices remained at the lower end of the scale, while the concentration of MEA migrants is the main determinant of the distribution of migrants from the former colonies and vice versa (Table 4). It might not be clear whether the high concentration of migrants maintained the stable price trend or whether the cheap housing has attracted more migrants from these regions, but their combination does not attract WEU Europeans, who prefer the city centre, where, quaintly, prices are considerably higher (Figure 4c,d).

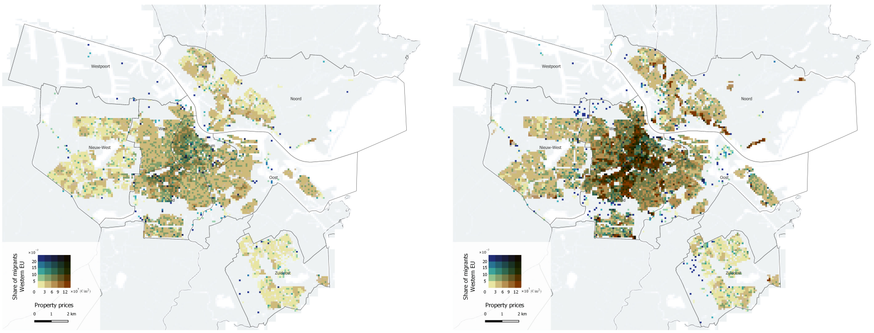


(**a**)



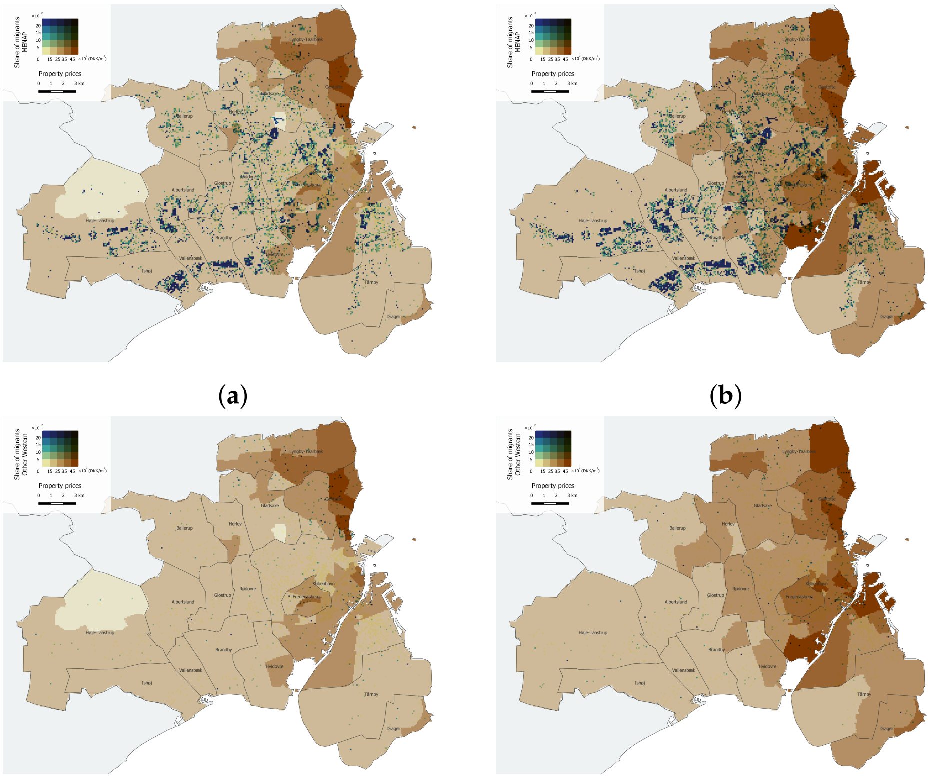
(**b**)

Figure 4. *Cont.*



(**c**) (**d**)

Figure 4. Distribution of migrants from MEA (**top**) and WEU countries (**bottom**) in relation to average house prices in Amsterdam in 2008 (**left**) and 2018 (**right**). (**a**) Migrants from MEA countries and average house prices in 2008. (**b**) Migrants from MEA countries and average house prices in 2018. (**c**) Migrants from WEU countries and average house prices in 2008. (**d**) Migrants from WEU countries and average house prices in 2018.



(**a**) (**b**)

(**c**) (**d**)

Figure 5. Distribution of migrants from MENAP (**top**) and other Western countries (**bottom**) in relation to average house prices at district level in Copenhagen in 2008 (**left**) and 2018 (**right**). (**a**) Migrants from MENAP countries and average house prices in 2008. (**b**) Migrants from MENAP countries and average house prices in 2018. (**c**) Migrants from other Western countries and average house prices in 2008. (**d**) Migrants from other Western countries and average house prices in 2018.

In Copenhagen, the patterns are much clearer; despite the small number of western migrants, their shares are high in the north-eastern municipalities, where the house prices are substantially higher (Figure 5c,d). These areas are primarily chosen by natives and high-income migrants. Lower shares are located in the municipalities of Copenhagen and Frederiksberg (light yellow), where most WEU migrants are located [85], explaining their high influence on other Westerners in the modelling (Table 3). On the contrary, the majority of MENAP migrants resided in cheaper districts.

5. Discussion

The results of this work highlight the diverse determinants of residential location among people from different ethnic and cultural backgrounds. They underline the complexity behind residential decisions and the fact that no single factor determines the preferences of all groups. Migrants are influenced by a number of factors, but the ethnic composition of the neighbourhood is an important one.

The study introduced a novel methodology using ML techniques and high spatial resolution data to improve our understanding of complex urban dynamics at the local level. We examined how different demographic and socio-economic factors influence the residential distribution of migrants and proposed a data-driven approach to quantify the weight of their importance. Quantifying the weight of neighbourhoods' ethnic composition is the main contribution of this study. Although the percentages of the weights may be spurious, the methodology highlighted connections between ethnic minorities and captured their relative importance. We have attempted to interpret complex urban phenomena and cultural bonds among migrant groups and historical patterns of migration, and to contribute to the emerging field of explainable artificial intelligence (XAI; [86,87]) by relating the obtained importance scores to the geographical or cultural proximity of the CoOs. A major strength of the study is its detailed geographical scale, using 100 metre grid cells, and its extensive examination of a large number of CoOs that have not been studied before. To date, most studies have focused on broad regions of origin or racial composition and various aggregated administrative units [24,28,55,88].

Potential interpretations based on the above observations and the similarities of the patterns of the two case studies were generalised and placed into the following categories:

I. Influences of co-ethnics. *High importance scores were found among migrants from the MEA region and the former colonies in Amsterdam, as well as among WEU and other Westerns in Copenhagen.*

The findings suggested a relationship between the presence of co-ethnics or minorities with similar cultural backgrounds and the residential patterns of various migrant groups. This observation was evident for migrants originating from regions that are geographically adjacent or historically united (e.g., Ghana–Gold Coast, Yugoslavia–North Macedonia, and Bosnia and Herzegovina–Croatia). These groups may share ethnic identities and display interconnected residential patterns influenced by factors such as the formation of social connections or access to ethnic shops.

Apart from the geographical neighbourhood of the CoOs, the data also suggested patterns of common residential distributions across broader cultural identities in the two cities under study. For example, there were associations in the residential distributions between the WEU and other Western groups, as well as between several Asian and African groups. The MENAP groups also showed a correlation in their residential patterns, which appeared to influence the distribution of other non-Western migrants from Africa and Latin America. These observations may contribute to our understanding of residential segregation and socio-economic inequalities within these communities.

Moreover, the observed association between the presence of Danes in Copenhagen and the residential patterns of Filipinos—a potential result of the significant number of Filipinos working and residing in the city's more affluent areas [85,89]—may reveal a link between housing pathways and labour market patterns. However, it is crucial to treat these findings with caution and avoid drawing direct causal relationships without further investigation.

II. Influences of dwelling type. *The residential area was more important for predicting the distribution of natives and WEU migrants; the number of dwellings was more important for non-Western migrants.*

It is a truism that housing features such as the number of dwellings and rooms, or the residential area and house prices, play a significant role in population density and in the distribution of different migrant groups. Although we cannot directly explain whether the concentration of certain migrant groups drove house prices or vice versa, the model accounted for variations in building types rather than property values and migrant groups.

Although the residential area and number of dwellings were correlated, the variations in their influence on migrants and natives provided clues to the groups' locations and may not be coincidental. Large numbers of dwellings, usually associated with large residential areas, indicated small dwelling units, more affordable housing and compact urban development. Conversely, small residential areas with a low number of dwellings per grid cell indicated large residential units, detached houses and more open space. The model related Western migrants and natives—more affluent groups—to the residential area layer, and non-Western migrant groups—less prosperous groups—to the number of dwellings, regardless of the group's population size. Indeed, Section 4.3 showed that Western migrants lived in expensive areas, taking advantage of a central location in Amsterdam or enjoying spacious dwellings and greener areas in Copenhagen. This allows for careful assumptions to be made about the economic status of the groups, their preferences and their quality of life, even in areas where the economic prospects of the groups are not known.

III. Influences of age as an indicator of migrants' initial settlement. *Distribution by age was the most important factor for migrants derived from EEU and the MEA region.*

Before or in the early 2000s, large migration waves from Turkey and Morocco arrived in Amsterdam, consisting mainly of young adults. We can safely assume that these groups had families and children of their own by 2018 and indeed maintained high fertility rates [90], thus establishing major groups of non-mobile adults with children. The same can be assumed for migrants from the MENAP region in Copenhagen. More recent waves of migrants have arrived in both cities from EEU and Mediterranean countries. Motivated by unrestricted mobility, financial and educational incentives after their countries joined the EU or after the 2008 financial crisis, these waves consisted of young and mobile adults who migrated to study and work. These observations may be useful in interpreting a link between migration and the period of initial settlement of the corresponding group.

Despite the valuable insights and methodological contributions of this study, some limitations should also be discussed. Our study captured a specific snapshot in time and focused on the distribution of migrants at that moment. This approach provided a deeper understanding of the underlying factors. However, it is essential to examine several time steps to gain a comprehensive overview of whether the data and results are consistent with any of the theories discussed in Section 2. Such an analysis may challenge existing theories and open up possibilities for highlighting unexplored dimensions and developing new conceptual frameworks. It is also important to note that the observed correlations do not necessarily imply direct causality or explicit residential preferences.

The study is also limited in allowing a direct comparison between the two case studies, particularly in relation to house prices. The datasets for these regions differ significantly, affecting the conclusions of the comparisons. In future research, consistent datasets would allow for more accurate cross-study comparisons. Additionally, more detailed data on the age and gender of migrants could help in understanding individual preferences, while combining geographically weighted regression with the RF model or incorporating geographic weights for the variables would improve the analyses. Further quantitative research should also include the role of previous location and social networks in residential location modelling. The visual representation of the distribution of selected migrant groups, along with their key determinants, opens up new potential avenues for research at the local level in these destination cities.

6. Conclusions

The growing population in urban centres and the exacerbation of inequalities and segregation associated with international mobility require better local planning; better urban planning and decision-making require more efficient residential location modelling. Therefore, a deeper understanding of the factors that influence migrants' residential locations and the reasons that enable them is crucial. This study provided data-driven evidence on the importance of the determinants of migrants' residential locations. We showed that by using a straightforward modelling approach and a wide range of geospatial data, we can extract rich information about the distribution of migrant groups and the factors that drive their residential distributions.

The study contributed to estimating the impact of cultural background and ethnic diversity on migrants' residential locations. We ranked which migrant groups are the most influential for other minorities and evaluated the relationship between the distribution of co-ethnics and the cultural and geographical proximity of their CoOs in 2018. We found indicative migrant groups from neighbouring CoOs with mutually high importance scores, with examples from the Balkans in Copenhagen and from Africa and Latin America in Amsterdam. Analysing the distribution of migrant groups by housing prices, we found separate clusters of WEU and non-EU migrants in both cases, with the groups with the financial means living in more desirable areas with higher housing quality and a low concentration of non-Western migrants.

To the best of our knowledge, this is the first study to perform ML-based analysis for residential location modelling using such a variety of detailed geospatial variables. Although RF and VIMs have primarily been adopted in bioinformatics, they can also be used in other research fields, such as geography and demography, to identify features that are significant in improving the explainability of complicated urban phenomena that cannot be explained by simple regression models.

The present work contributed to research in XAI, where we interpreted the decisions of the ML model by confirming existing knowledge about the residential distribution of migrants. The model captured patterns across groups of various sizes, revealing hidden relationships that might otherwise be assumed. It presented a relative association of factors and ordered them, identifying extraneous noise. The results allowed us to infer the economic status of these groups and their quality of life, and potentially find evidence of residential segregation for the year in question. We highlighted the role of diversity and ethnic composition for different minorities and their residential locations, along with their ability to acquire property in the most desirable areas.

Such an approach provides important insights for decision-making and drawing guidelines for inclusive urban development, with safeguards to limit the exploitation of foreign populations. Examples include locating not only safe but also desirable spaces for the initial settlement of new waves of migrants or refugees (e.g., Ukrainians moving to Europe), thus preventing the phenomena of segregation and ghettoisation.

Author Contributions: Conceptualization, Marina Georgati, Henning Sten Hansen and Carsten Keßler; methodology, Marina Georgati, Henning Sten Hansen and Carsten Keßler; software, Marina Georgati; validation, Marina Georgati; formal analysis, Marina Georgati; investigation, Marina Georgati; data curation, Marina Georgati; writing—original draft preparation, Marina Georgati; writing—review and editing, Marina Georgati, Henning Sten Hansen and Carsten Keßler; visualization, Marina Georgati; and supervision, Carsten Keßler and Henning Sten Hansen. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by EU Horizon 2020 Programme grant number 870649.

Data Availability Statement: The supporting computational workflow is available at Figshare (https://figshare.com/s/78580b2609f00216b34c) (https://doi/org/10.6084/m9.figshare.21983807) and can be openly re-used for similar analyses. Several of the used datasets are publicly available.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Details on the Case Studies

Table A1. Variations between the case studies areas in 2018.

Feature	Amsterdam	Copenhagen
Spatial extent	219.5km^2	527.9km^2
Cells	42,244	114,367
Inhabited cells	7512	21,403
Inhabited cells by natives	7288	21,233
Inhabited cells by migrants	6782	16,818

Table A2. Classification of CoO to RoO.

Case Study	Group	Countries
Amsterdam	NLD	The Netherlands.
	EEU	Bulgaria, Hungary, Czech Republic, Poland, Roumania, Slovakia, Estonia, Latvia, Lithuania, Croatia and Slovenia.
	WEU	The rest of EU and the United Kingdom (UK).
	Turkey + Morocco	Turkey and Morocco.
	Middle East + Africa	All countries of Middle East and Africa.
	Former Colonies	Aruba, Bonaire, Curaçao, Indonesia, Saba, Saint Maarten-Dutch part, Sint Eustatius and Suriname.
	Other Europe etc.	Non-EU European and European Free Trade Association (EFTA) countries, Asia, Oceania and the Americas.
Copenhagen	DNK	Denmark.
	EEU	Bulgaria, Hungary, Poland, Roumania, Slovakia, Estonia, Latvia, Lithuania, Croatia, Slovenia and Czechoslovakia.
	WEU	The rest EU and the EFTA countries, the UK, the Faroe Islands and Greenland.
	Non-EU Europe	Yugoslavia, Serbia and Montenegro, and the Soviet Union.
	MENAP	Turkey, Algeria, Egypt, Libya, Morocco, Tunisia, Iraq, Iran, Bahrain, Israel, Palestine, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, United Arab Emirates, Syria, Yemen and Pakistan.
	Other Western	Canada, United States of America (USA), Australia and New Zealand.
	Non-Western	All other countries.

Table A3. Summary statistics of determinants for Amsterdam in 2018, including age groups, building and neighbourhood features, and native/migrant presence by region of origin. Abbreviations: Min.—minimum, Max—maximum, Std.—standard deviation and RoO—region of origin.

Variables	Min.	Max.	Mean	Std.	Range
Age groups					
Children	0	295	3.94	12.70	295
Young adults	0	494	3.93	15.72	494
Mobile adults	0	608	5.04	16.70	608
Non-mobile adults	0	583	4.97	15.05	583
Elderly	0	236	2.56	9.01	236
Building features					
Not used	0	185	0.59	3.20	185
Rented	0	576	7.21	25.18	576
Private	0	164	3.12	11.06	164
Residential area	0	26,681	631.35	1854.15	26,681
Number of dwellings	0	576	10.34	32.39	576
Number of rooms	0	1145	33	98	1145
Building height	0	89.32	3.45	5.57	89.32
Building oldness	0	2016	782.88	958.99	2016
Real estate prices	0	12,064.05	891.34	2339.12	12,064.05

Table A3. *Cont.*

Variables	Min.	Max.	Mean	Std.	Range
Neighbourhood features					
Proximity to restaurants and bars	0	1046	14.86	75.41	1046
Proximity to bus stops	0	29	1.37	3.02	29
Proximity to culture	0	42	0.57	2.83	42
Proximity to supermarkets	0	64	1.99	6.70	64
Proximity to a Christian place of worship	0	31	3.39	6.19	31
Proximity to a non-Christian place of worship	0	21	1.61	3.70	21
Proximity to a Buddhist place of worship	0	2	0.13	0.40	2
Proximity to a Christian place of worship	0	20	2.12	4.03	20
Proximity to a Christian place of worship (Catholic)	0	3	0.20	0.57	3
Proximity to a Christian place of worship (Lutheran)	0	1	0.07	0.25	1
Proximity to a Christian place of worship (Protestant)	0	9	1	1.71	9
Proximity to a Hindu place of worship	0	3	0.28	0.72	3
Proximity to a Jewish place of worship	0	4	0.28	0.86	4
Proximity to a Muslim place of worhsip	0	9	0.69	1.55	9
Proximity to a Sunni Muslim place of worship	0	1	0.06	0.24	1
Proximity to a Sikh place of worship	0	3	0.17	0.49	3
Proximity to schools	0	26	2.47	4.17	26
Proximity to train stations	0	7	0.74	1.44	7
Proximity to university	0	4	0.22	0.83	4
RoO					
Natives	0	796	20.76	63.49	796
Former Colonies	0	499	2.04	8.42	499
EU West	0	140	1.59	6.02	140
EU East	0	68	0.47	2.03	68
Middle East + Africa	0	355	1.37	6.56	355
Turkey + Morocco	0	434	2.65	13.40	434
Other Europe etc	0	377	1.90	7.07	377

Table A4. Summary statistics of determinants for Amsterdam in 2018, including the presence of natives/migrants by country of origin. Abbreviations: Min.—minimum, Max.—maximum, Std.—standard deviation and CoO—Country of origin.

Variables	Min.	Max.	Mean	Std.	Range	Variables	Min.	Max.	Mean	Std.	Range
CoO											
Afghanistan	0	38	0.07	0.66	38	Kroatia	0	2	0.00	0.04	2
Albania	0	9	0.01	0.15	9	Letland	0	4	0.01	0.13	4
Algerije	0	9	0.03	0.27	9	Libanon	0	7	0.02	0.19	7
Angola	0	6	0.01	0.13	6	Liberia	0	6	0.01	0.12	6
Argentinia	0	6	0.03	0.26	6	Libia	0	11	0.00	0.11	11
Aruba	0	3	0.00	0.05	3	Litouwen	0	6	0.02	0.18	6
Australia	0	14	0.05	0.34	14	Luxemburg	0	4	0.00	0.09	4
Azerbeidzjan	0	11	0.00	0.08	11	Malakka	0	5	0.01	0.09	5
Bangladesh	0	8	0.01	0.14	8	Maleisia	0	5	0.01	0.11	5
Belgisch-Congo	0	2	0.00	0.06	2	Malta	0	3	0.00	0.04	3
Belgia	0	13	0.10	0.52	13	Marokko	0	349	1.68	9.02	349
Bolivia	0	4	0.01	0.10	4	Mauritius	0	3	0.00	0.05	3
Bondsrepubliek Duitsland	0	44	0.21	1.08	44	Mexico	0	7	0.03	0.23	7
Brazilia	0	17	0.12	0.65	17	Moldavia	0	3	0.00	0.06	3
Brits Oost-Afrika	0	3	0.00	0.05	3	Mongolia	0	4	0.00	0.06	4
Brits West-India	0	4	0.00	0.07	4	Nederland	0	796	20.76	63.49	796
Brits-Guyana	0	8	0.02	0.20	8	Nederlands Nieuw-Guinea	0	4	0.01	0.12	4
Brits-India	0	8	0.04	0.28	8	Nederlands-India	0	29	0.29	1.08	29
Brits-Somaliland	0	5	0.00	0.08	5	Nederlandse Antillen	0	88	0.22	1.29	88
Bulgarije	0	55	0.11	0.99	55	Nepal	0	10	0.01	0.23	10
Burma	0	6	0.00	0.10	6	Nicaragua	0	3	0.00	0.04	3
Burundi	0	5	0.00	0.07	5	Nieuw-Zeeland	0	6	0.02	0.19	6

Table A4. *Cont.*

Variables	Min.	Max.	Mean	Std.	Range	Variables	Min.	Max.	Mean	Std.	Range
Cambodja	0	7	0.00	0.06	7	Nigeria	0	36	0.05	0.50	36
Canada	0	16	0.04	0.33	16	Noorwegen	0	6	0.01	0.16	6
Canarische Eilanden	0	3	0.00	0.05	3	Oekraine	0	23	0.01	0.16	23
Ceylon	0	5	0.01	0.13	5	Onbekend	0	38	0.04	0.50	38
Chili	0	6	0.02	0.21	6	Oostenrijk	0	8	0.03	0.25	8
China	0	67	0.16	1.05	67	Opper-Volta	0	6	0.00	0.07	6
Colombia	0	10	0.06	0.39	10	Pakistan	0	38	0.11	0.81	38
Congo-Kinshasa	0	6	0.00	0.07	6	Palestina	0	3	0.00	0.07	3
Congo	0	6	0.00	0.08	6	Panama	0	3	0.00	0.05	3
Costa Rica	0	4	0.00	0.06	4	Peru	0	7	0.02	0.20	7
Cuba	0	5	0.01	0.10	5	Polen	0	12	0.13	0.67	12
Cyprus	0	4	0.00	0.09	4	Portugal	0	13	0.08	0.49	13
Denemarken	0	8	0.02	0.21	8	Portugees Oost-Afrika	0	3	0.00	0.05	3
Dominicaanse Republiek	0	49	0.06	0.62	49	Portugees West-Afrika	0	4	0.00	0.06	4
Duitse Democratische Republiek	0	6	0.01	0.16	6	Puerto Rico	0	3	0.00	0.04	3
Duitsland	0	15	0.15	0.63	15	Rhodesia	0	3	0.00	0.04	3
Ecuador	0	30	0.02	0.29	30	Roemenia	0	17	0.09	0.56	17
Egypte	0	36	0.13	0.88	36	Rusland (oud)	0	2	0.00	0.05	2
El Salvador	0	4	0.00	0.05	4	Rusland	0	16	0.01	0.16	16
Eritrea	0	23	0.02	0.34	23	Rwanda	0	5	0.00	0.06	5
Estland	0	4	0.01	0.10	4	Saoedi-Arabia	0	6	0.01	0.11	6
Ethiopia	0	26	0.06	0.51	26	Senegal	0	7	0.01	0.12	7
Filipijnen	0	55	0.06	0.54	55	Siam	0	2	0.00	0.05	2
Finland	0	7	0.02	0.22	7	Sierra Leone	0	9	0.01	0.16	9
Frankrijk	0	18	0.17	0.85	18	Singapore	0	7	0.01	0.15	7
Frans West-Afrika	0	5	0.00	0.06	5	Slovenia	0	3	0.00	0.04	3
Frans-Guyana	0	4	0.00	0.08	4	Slowakije	0	5	0.00	0.06	5
Gambia	0	5	0.00	0.09	5	Soedan	0	13	0.02	0.23	13
Ghana	0	104	0.21	2.22	104	Somalia	0	20	0.03	0.39	20
Goudkust	0	39	0.08	0.82	39	Sovjet-Unie	0	30	0.19	0.94	30
Griekenland	0	23	0.06	0.45	23	Spanje	0	26	0.14	0.71	26
Groot-Brittannia	0	35	0.28	1.29	35	Sri Lanka	0	5	0.00	0.09	5
Guadeloupe	0	3	0.00	0.04	3	Suriname	0	392	1.39	6.61	392
Guatemala	0	3	0.00	0.06	3	Syria	0	43	0.09	0.79	43
Guinee	0	8	0.01	0.14	8	Taiwan	0	7	0.01	0.16	7
Guyana	0	6	0.01	0.11	6	Tanzania	0	4	0.00	0.06	4
Haiti	0	5	0.00	0.07	5	Thailand	0	10	0.03	0.26	10
Honduras	0	3	0.00	0.05	3	Tibet	0	6	0.01	0.14	6
Hongarije	0	10	0.05	0.35	10	Togo	0	9	0.01	0.14	9
Hongkong	0	14	0.05	0.36	14	Trinidad en Tobago	0	6	0.00	0.08	6
Ierland	0	8	0.04	0.29	8	Tsjechia	0	5	0.00	0.06	5
IJsland	0	5	0.00	0.07	5	Tsjecho-Slowakije	0	7	0.04	0.29	7
India	0	326	0.17	1.98	326	Tunesia	0	10	0.02	0.26	10
Indonesia	0	17	0.12	0.59	17	Turkije	0	119	0.97	5.03	119
Irak	0	22	0.07	0.53	22	Uganda	0	14	0.01	0.19	14
Iran	0	47	0.09	0.53	47	Uruguay	0	4	0.01	0.10	4
Israal	0	11	0.05	0.37	11	Venezuela	0	7	0.02	0.20	7
Italiaans-Somaliland	0	3	0.00	0.06	3	Verenigde Arabische Emiraten	0	3	0.00	0.06	3
Italia	0	28	0.21	1.02	28	Verenigde Staten van Amerika	0	63	0.21	1.18	63
Ivoorkust	0	9	0.01	0.12	9	Vietnam	0	6	0.02	0.20	6
Jamaica	0	6	0.00	0.09	6	Zambia	0	4	0.00	0.06	4
Japan	0	26	0.04	0.42	26	Zaire	0	6	0.00	0.10	6
Joegoslavia	0	15	0.15	0.74	15	Zimbabwe	0	5	0.00	0.05	5
Jordania	0	7	0.00	0.11	7	Zuid-Afrika	0	18	0.05	0.36	18
Kaapverdia	0	7	0.01	0.17	7	Zuid-Korea	0	10	0.01	0.18	10
Kameroen	0	7	0.01	0.12	7	Zuid-Rhodesia	0	4	0.00	0.05	4
Kenya	0	6	0.01	0.11	6	Zweden	0	10	0.03	0.27	10
Koeweit	0	5	0.00	0.08	5	Zwitserland	0	8	0.03	0.24	8
Korea	0	6	0.00	0.09	6						

References

1. Pagliara, F.; Wilson, A. The State-of-the-Art in Building Residential Location Models. In *Residential Location Choice*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 1–20. [CrossRef]
2. Zolfaghari, A.; Sivakumar, A.; Polak, J.W. Choice set pruning in residential location choice modelling: A comparison of sampling and choice set generation approaches in greater London. *Transp. Plan. Technol.* **2012**, *35*, 87–106. [CrossRef]
3. Schirmer, P.M.; Eggermond, M.A.B.V.; Axhausen, K.W. The role of location in residential location choice models. *J. Transp. Land Use* **2014**, *7*, 3–21. [CrossRef]
4. Walker, J.L.; Li, J. Latent lifestyle preferences and household location decisions. *J. Geogr. Syst.* **2007**, *9*, 77–101. [CrossRef]
5. Zondag, B.; Pieters, M. Influence of Accessibility on Residential Location Choice. *Transp. Res. Rec.* **2005**, *1902*, 63–70. [CrossRef]
6. Chen, J.; Chen, C.; Timmermans, H.J.P. Accessibility trade-offs in household residential location decisions. *Transp. Res. Rec.* **2008**, *2077*, 71–79. [CrossRef]
7. Mulder, C.H. The family context and residential choice: A challenge for new research. *Popul. Space Place* **2007**, *13*, 265–278. [CrossRef]
8. Mulder, C.H.; Cooke, T.J. Family ties and residential locations. *Popul. Space Place* **2009**, *15*, 299–304. [CrossRef]
9. Axhausen, K.W.; Scott, D.M.; König, A.; Jürgens, C. Locations, commitments and activity spaces. In *Human Behaviour and Traffic Networks*; Springer: Berlin/Heidelberg, Germany, 2001. [CrossRef]
10. Pinjari, A.R.; Bhat, C.R.; Hensher, D.A. Residential self-selection effects in an activity time-use behavior model. *Transp. Res. Part B Methodol.* **2009**, *43*, 729–748. [CrossRef]
11. Pinjari, A.R.; Pendyala, R.M.; Bhat, C.R.; Waddell, P.A. Modeling the choice continuum: An integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. *Transportation* **2011**, *38*, 933–958. [CrossRef]
12. de Palma, A.; Motamedi, K.; Picard, N.; Waddell, P. A Model of Residential Location Choice with Endogenous Housing Prices and Traffic for the Paris Region. *Eur. Transp.* **2005**, *31*, 67–82.
13. Muhammad, A.H.; Miller, E.J. Reference-Dependent Residential Location Choice Model within a Relocation Context. *Transp. Res. Rec.* **2009**, *2133* , 92–99.
14. Beckers, P.; Boschman, S. Residential choices of foreign highly skilled workers in the Netherlands and the role of neighbourhood and urban regional characteristics. *Urban Stud.* **2019**, *56*, 760–777. [CrossRef]
15. Frenkel, A.; Bendit, E.; Kaplan, S. Residential location choice of knowledge-workers: The role of amenities, workplace and lifestyle. *Cities* **2013**, *35*, 33–41. [CrossRef]
16. Florida, R. *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life*; BasicBooks: New York, NY, USA, 2002.
17. Dimou, M.; Ettouati, S.; Schaffar, A. From dusk till dawn: The residential mobility and location preferences of immigrants in France. *Ann. Reg. Sci.* **2020**, *65*, 253–280. [CrossRef]
18. Richardson, K. Attracting and Retaining Foreign Highly Skilled Staff in Times of Global Crisis: A Case Study of Vancouver, British Columbia's Biotechnology Sector. *Popul. Space Place* **2016**, *22*, 428–440. [CrossRef]
19. International Organization for Migration IOM. World Migration Report 2022. 2021 . Available online: https://publications.iom.int/books/world-migration-report-2022 (accessed on 23 July 2023).
20. World Bank. Urban Development. 2023. Available online: https://www.worldbank.org/en/topic/urbandevelopment/overview (accessed on 23 July 2023).
21. United Nations Human Settlements Programme UN-Habitat. *World Cities Report 2022: Envisaging the Future of Cities*; Technical Report; United Nations Human Settlements Programme: Nairobi, Kenya, 2022.
22. Aslund, O. Now and forever? Initial and subsequent location choices of immigrants. *Reg. Sci. Urban Econ.* **2005**, *35*, 141–165. [CrossRef]
23. Perez, L.; Dragicevic, S.; Gaudreau, J. A geospatial agent-based model of the spatial urban dynamics of immigrant population: A study of the island of Montreal, Canada. *PLoS ONE* **2019**, *14*, e0219188. [CrossRef] [PubMed]
24. Søholt, S.; Lynnebakke, B. Do Immigrants' Preferences for Neighbourhood Qualities Contribute to Segregation? The Case of Oslo. *J. Ethn. Migr. Stud.* **2015**, *41*, 2314–2335. [CrossRef]
25. Ibraimovic, T.; Hess, S. A latent class model of residential choice behaviour and ethnic segregation preferences. *Hous. Stud.* **2018**, *33*, 544–564. [CrossRef]
26. Robinson, D.; Reeve, K.; Casey, R. *The Housing Pathways of New Immigrants*; Joseph Rowntree Foundation: London, UK, 2007 .
27. Wang, C.; Li, C.; Wang, M.; Yang, S. Transnational migrants in Shanghai: Residential spatial patterns and the underlying driving forces. *Popul. Space Place* **2020**, *26*, e2272. [CrossRef]
28. Andersen, H.S. Spatial assimilation? The development in immigrants' residential career with duration of stay in Denmark. *J. Hous. Built Environ.* **2016**, *31*, 297–320. [CrossRef]
29. Chihaya, G.K.; Marcińczak, S.; Strömgren, M.; Lindgren, U.; Tammaru, T. Trajectories of Spatial Assimilation or Place Stratification? A Typology of Residence and Workplace Histories of Newly Arrived Migrants in Sweden. *Int. Migr. Rev.* **2022**, *56*, 433–462. [CrossRef]
30. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. :1010933404324 [CrossRef]

31. Hapfelmeier, A.; Ulm, K. A new variable selection approach using Random Forests. *Comput. Stat. Data Anal.* **2013**, *60*, 50–69. [CrossRef]
32. Couronné, R.; Probst, P.; Boulesteix, A.L. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinform.* **2018**, *19*, 270. [CrossRef]
33. Boulesteix, A.L.; Janitza, S.; Kruppa, J.; König, I.R. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2012**, *2*, 493–507. [CrossRef]
34. Voges, L.F.; Jarren, L.C.; Seifert, S. Exploitation of surrogate variables in random forests for unbiased analysis of mutual impact and importance of features. *Bioinformatics* **2023**, *39*, btad471. [CrossRef]
35. Stevens, F.R.; Gaughan, A.E.; Linard, C.; Tatem, A.J. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLoS ONE* **2015**, *10*, e0107042. [CrossRef] [PubMed]
36. Georgati, M.; Monteiro, J.; Martins, B.; Keßler, C. Spatial Disaggregation of Population Subgroups Leveraging Self-Trained Multi-Output Gradient Boosting Regression Trees. *AGILE Gisci. Ser.* **2022**, *3*, 5. [CrossRef]
37. Luo, Y.; Yan, J.; McClure, S.C.; Li, F. Socioeconomic and environmental factors of poverty in China using geographically weighted random forest regression model. *Environ. Sci. Pollut. Res. Int.* **2022**, *29*, 33205–33217. [CrossRef]
38. Kopeć, A.; Trybała, P.; Głąbicki, D.; Buczyńska, A.; Owczarz, K.; Bugajska, N.; Kozińska, P.; Chojwa, M.; Gattner, A. Application of remote sensing, gis and machine learning with geographically weighted regression in assessing the impact of hard coal mining on the natural environment. *Sustainability* **2020**, *12*, 9338. [CrossRef]
39. Cosacov, N. Residential Choices. In *The Wiley Blackwell Encyclopedia of Urban and Regional Studies*; John Wiley & Sons: Hoboken, NJ, USA, 2019; pp. 1–5. [CrossRef]
40. Hunt, L.M.; Boots, B.; Kanaroglou, P.S. Spatial choice modelling: New opportunities to incorporate space into substitution patterns. *Prog. Hum. Geogr.* **2004**, *28*, 746–766. [CrossRef]
41. Pagliara, F.; Simmonds, D.; Preston, J. *Residential Location Choice: Models and Applications*; Springer: Berlin/Heidelberg, Germany, 2010. [CrossRef]
42. Wen, T. Spatial Assimilation. In *The Wiley Blackwell Encyclopedia of Urban and Regional Studies*; John Wiley & Sons: Hoboken, NJ, USA, 2019. [CrossRef]
43. Massey, D.S.; Mullan, B.P. Processes of Hispanic and Black Spatial Assimilation. *Am. J. Sociol.* **1984**, *89*, 836–873. [CrossRef]
44. Massey, D.S.; Denton, N.A. Spatial Assimilation as a Socioeconomic Outcome. *Am. Sociol. Rev.* **1985**, *50*, 94–106. [CrossRef]
45. Alba, R.D.; Logan, J. Minority proximity to whites in suburbs: An individual-level analysis of segregation. *Am. J. Sociol.* **1993**, *98*, 1388–1427. [CrossRef]
46. Charles, C.Z. The Dynamics of Racial Residential Segregation. *Annu. Rev. Sociol.* **2003**, *29*, 167–207. [CrossRef]
47. Vogiazides, L. Exiting distressed neighbourhoods: The timing of spatial assimilation among international migrants in Sweden. *Popul. Space Place* **2018**, *24*, e2169. [CrossRef]
48. Thomas, R. Viewing immigrants' neighbourhood and housing choices through the lens of community resilience. *Sapiens* **2013**, *6*.
49. Burgess, E.W. The Growth of the City: An Introduction to a Research Project. In *Urban Ecology*; Springer: Boston, MA, USA, 2008; pp. 71–78. [CrossRef]
50. Zorlu, A.; Mulder, C.H. Initial and Subsequent Location Choices of Immigrants to the Netherlands. *Reg. Stud.* **2008**, *42*, 245–264. [CrossRef]
51. Bolt, G.; van Kempen, R. Ethnic Segregation and Residential Mobility: Relocations of Minority Ethnic Groups in the Netherlands. *J. Ethn. Migr. Stud.* **2010**, *36*, 333–354. [CrossRef]
52. Hedman, L. Moving Near Family? The Influence of Extended Family on Neighbourhood Choice in an Intra-urban Context. *Popul. Space Place* **2013**, *19*, 32–45. [CrossRef]
53. Coenen, A.; Verhaeghe, P.; de Putte, B.V. Ethnic residential segregation: A matter of ethnic minority household characteristics? *Popul. Space Place* **2019**, *25*, e2244. [CrossRef]
54. Pais, J.; South, S.J.; Crowder, K. Metropolitan Heterogeneity and Minority Neighborhood Attainment. *Soc. Probl.* **2012**, *59*, 258–281. [CrossRef] [PubMed]
55. Martori, J.C.; Madariaga, R. Residential segregation by nationalities: A global and multilevel approach to Barcelona and Madrid (2008–2018). *Popul. Space Place* **2023**, *29*, e2655. [CrossRef]
56. Catney, G.; Lloyd, C.D. Population Grids for Analysing Long-Term Change in Ethnic Diversity and Segregation. *Spat. Demogr.* **2020**, *8*, 215–249. [CrossRef]
57. Malmberg, B.; Andersson, E.K.; Nielsen, M.M.; Haandrikman, K. Residential Segregation of European and Non-European Migrants in Sweden: 1990–2012. *Eur. J. Popul.* **2018**, *34*, 169–193. [CrossRef]
58. Sleutjes, B.; de Valk, H.A.G.; Ooijevaar, J. The Measurement of Ethnic Segregation in the Netherlands: Differences between Administrative and Individualized Neighbourhoods. *Eur. J. Popul.* **2018**, *34*, 195–224. [CrossRef] [PubMed]
59. Imeraj, L.; Willaert, D.; de Valk, H.A.G. A comparative approach towards ethnic segregation patterns in Belgian cities using multiscalar individualized neighborhoods. *Urban Geogr.* **2018**, *39*, 1221–1246. [CrossRef]
60. Rogne, A.F.; Andersson, E.K.; Malmberg, B.; Lyngstad, T.H. Neighbourhood Concentration and Representation of Non-European Migrants: New Results from Norway. *Eur. J. Popul.* **2020**, *36*, 71–83. [CrossRef] [PubMed]
61. Openshaw, S. *The Modifiable Areal Unit Problem*; CATMOG, 38; Geo: Norwich, UK, 1980.

62. Andersson, E.K.; Lyngstad, T.H.; Sleutjes, B. Comparing Patterns of Segregation in North-Western Europe: A Multiscalar Approach. *Eur. J. Popul.* **2018**, *34*, 151–168. [CrossRef] [PubMed]
63. Jones, K.; Johnston, R.; Manley, D.; Owen, D.; Charlton, C. Ethnic Residential Segregation: A Multilevel, Multigroup, Multiscale Approach Exemplified by London in 2011. *Demography* **2015**, *52*, 1995–2019. [CrossRef] [PubMed]
64. OpenStreetMap Contributors. 2022. Available online: http://download.geofabrik.de/europe.html (accessed on 3 February 2022).
65. Bygnings-og Boligregistret (BBR). Forside—BBR Teknik. 2022. Available online: https://teknik.bbr.dk/forside (accessed on 28 April 2021).
66. PDOK. PDOK—3D Downloads. 2022. Available online: https://3d.kadaster.nl/basisvoorziening-3d/ (accessed on 28 April 2021).
67. Gemeente Amsterdam. Maps Amsterdam. Available online: https://maps.amsterdam.nl/open_geodata/?LANG=en (accessed on 3 February 2022).
68. Movia Trafik. Movia. 2022. Available online: https://www.moviatrafik.dk/ (accessed on 28 April 2021).
69. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2009. [CrossRef]
70. Géron, A. *Hands-on Machine Learning with Scikit-Learn and TensorFlow*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2009.
71. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
72. Nembrini, S.; König, I.R.; Wright, M.N. The revival of the Gini importance? *Bioinformatics* **2018**, *34*, 3711–3718. [CrossRef]
73. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
74. Nicodemus, K.K. Letter to the Editor: On the stability and ranking of predictors from random forest variable importance measures. *Brief. Bioinform.* **2011**, *12*, 369–373. [CrossRef]
75. Calle, M.L.; Urrea, V. Letter to the Editor: Stability of Random Forest importance measures. *Brief. Bioinform.* **2011**, *12*, 86–89. [CrossRef]
76. Matplotlib—Visualization with Python. 2023. Available online: https://matplotlib.org/ (accessed on 5 September 2023).
77. Seaborn: Statistical Data Visualization. 2023. Available online: https://seaborn.pydata.org/ (accessed on 5 September 2023).
78. Andersen, H.S.; Andersson, R.; Wessel, T.; Vilkama, K. The impact of housing policies and housing markets on ethnic spatial segregation: Comparing the capital cities of four Nordic welfare states. *Eur. J. Hous. Policy* **2016**, *16*, 1–30. [CrossRef]
79. Andersson, E.K.; Malmberg, B.; Costa, R.; Sleutjes, B.; Stonawski, M.J.; de Valk, H.A.G. A Comparative Study of Segregation Patterns in Belgium, Denmark, the Netherlands and Sweden: Neighbourhood Concentration and Representation of Non-European Migrants. *Eur. J. Popul.* **2018**, *34*, 251–275. [CrossRef]
80. Arbaci, S. Ethnic segregation, housing systems and welfare regimes in Europe. *Eur. J. Hous. Policy* **2007**, *7*, 401–433. [CrossRef]
81. Tammaru, T.; Marcińczak, S.; Aunap, R.; van Ham, M. *Inequalities and Segregation across the Long-Term Economic Cycle: An Analysis of South and North European Cities*; IDEAS Working Paper Series from RePEc; IZA—Institute of Labor Economics: Bonn, Germany, 2007.
82. Deloitte. Average Sales Price of New and Existing Residential Property in Europe in 2021, per Country (In Euros Per Square Meter). 2022. Available online: https://www.statista.com/statistics/722905/average-residential-square-meter-prices-in-eu-28-per-country/ (accessed on 5 September 2023).
83. Kveladze, I.; Georgati, M.; Kessler, C.; Hansen, H.S. Analytics of historical human migration patterns: Use cases of Amsterdam and Copenhagen. *J. Locat. Based Serv.* **2023**. [CrossRef]
84. Gemeente Amsterdam, Bureau Onderzoek en Statistiek. *Stadsdelen in Cijfers 2013*; Technical Report; Gemeente Amsterdam, Bureau Onderzoek en Statistiek: Amsterdam, The Netherlands, 2013.
85. Georgati, M.; Keßler, C. Spatially Explicit Population Projections: The case of Copenhagen, Denmark. *AGILE GISci. Ser.* **2021**, *2*, 1–6. [CrossRef]
86. Lent, M.V.; Fisher, W.; Mancuso, M. An explainable artificial intelligence system for small-unit tactical behavior. In Proceedings of the National Conference on Artificial Intelligence, San Jose, CA, USA, 25–29 July 2004.
87. Carvalho, D.V.; Pereira, E.M.; Cardoso, J.S. Machine learning interpretability: A survey on methods and metrics. *Electronics* **2019**, *8*, 832. [CrossRef]
88. Erickson, B.O.; Esser, D.E.; Turner, K.L. Spatial Dynamics of Racialized Income Inequality in the Washington, DC, Metropolitan Area. *Prof. Geogr.* **2021**, *73*, 72–84. [CrossRef]
89. Andersen, N.T. Organizing against all Odds: Filipina chambermaids in Copenhagen, 1970s–1990s. *Women's Hist. Rev.* **2021**, *30*, 241–255. [CrossRef]
90. Garssen, J.; Nicolaas, H. Fertility of Turkish and Moroccan women in the Netherlands. *Demogr. Res.* **2008**, *19*, 1249–1280. [CrossRef]