

C>ONSTRUCTOR UNIVERSITY

School of Business, Social & Decision Sciences

Master Thesis

Ethnic Segregation in England's Selected Areas

Supervisor: Prof. Dr. Jan Lorenz

Handed in by: Date: 30/06/2024
Niloufar Shokri
Finkenschlag 7
28759 Bremen
30006412

Table of Contents

Table of Contents	II
List of figures.....	IV
List of tables.....	VI
List of abbreviations.....	VII
Executive Summary	VIII
1 Introduction.....	1
2 Data Source.....	1
3 Data Preprocessing	3
3.1 Data of 2001	3
3.2 Data of 2011	3
3.3 Data of 2021	4
4 Population Trends in England.....	4
5 Dissimilarity Index	5
5.1 Dissimilarity Index at Country Level	7
5.2 Dissimilarity Index at LAD Level.....	9
6 Simpson Index	13
6.1 Simpson Index at Country Level.....	13
6.2 Simpson Index at LAD Level.....	14
7 Moran's I Index.....	17
7.1 Moran's I at country level.....	18
7.2 Moran's I at LAD level.....	19
8 Detailed Examination of Segregation in Selected English Cities	22
8.1 Birmingham.....	23
8.1.1 Birmingham Map – OA Level	24
8.1.2 Birmingham Map – MSOA Level.....	25
8.2 Leicester	26
8.2.1 Leicester Map – OA Level.....	27
8.2.2 Leicester Map – MSOA Level.....	28
8.3 Bradford.....	29
8.3.1 Bradford Map – OA Level	30
8.3.2 Bradford Map – MSOA Level	31
8.4 Blackburn with Darwen	32
8.4.1 Blackburn with Darwen Map – OA Level	33
8.4.2 Blackburn with Darwen Map – MSOA Level.....	34
8.5 Oldham	35
8.5.1 Oldham Map – OA Level.....	36
8.5.2 Oldham Map – MSOA Level.....	37
8.6 Pendle	38
8.6.1 Pendle Map – OA Level.....	39
8.6.2 Pendle Map – MSOA Level.....	40
8.7 London.....	41
8.7.1 London Map – OA Level	42
8.7.2 London Map – MSOA Level	43
8.7.3 London Map – LAD Level	44
9 Comparing the cities	45

10	Conclusion	48
11	Table of references.....	50
12	Appendix	51
	Affidavit.....	52

List of figures

Figure 1. Ethnic Group Fractions over Time.....	5
Figure 2. Ethnicity population Line Chart	5
Figure 3. Dissimilarity Index Variation Across Geographical Levels	8
Figure 4. Dissimilarity Index Variation Over the Time Across Geographic Levels	8
Figure 5. Dissimilarity Index Box-Plot by Ethnicity and Area	11
Figure 6. OA-based Dissimilarity Index for all ethnic groups in LADs	12
Figure 8. Simpson Index Line Chart at Country Level	14
Figure 9. Simpson Index Box-Plot based on different areas across the years.....	15
Figure 10. Simpson Index Differences between LAD and other areas across the years	16
Figure 11. LADs' Simpson Index on Map of England	17
Figure 12. Difference of LAD-OA Simpson Indices on Map of England.....	17
Figure 13. Moran Index Variation Across the Year	19
Figure 14. LSOA-based Moran's I in LADs for Ethnicities across the years.....	21
Figure 15. OA-based Moran's I in LADs on Map of England	22
Figure 16. Ethnicity and Total Population Line Chart – Birmingham	24
Figure 17. Ethnic Group Fractions Over Time - Birmingham	24
Figure 18. OA Simpson Index, White and Asian Fraction on Map - Birmingham OAs	25
Figure 19. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Birmingham MSOAs	26
Figure 20. Ethnicity and Total Population Line Chart - Leicester.....	27
Figure 21. Ethnic Group Fractions Over Time - Leicester.....	27
Figure 22. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map - Leicester OAs	28
Figure 23. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Leicester MSOAs.....	29
Figure 24. Ethnicity and Total Population Line Chart - Bradford	30
Figure 25. Ethnic Group Fractions Over Time - Bradford	30
Figure 26. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map - Bradford OAs	31
Figure 27. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Bradford MSOAs	32
Figure 28. Ethnicity and Total Population Line Chart – Blackburn	33
Figure 29. Ethnic Group Fractions Over Time – Blackburn	33
Figure 30. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map – Blackburn OAs	34
Figure 31. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Blackburn MSOAs.....	35
Figure 32. Ethnicity and Total Population Line Chart - Oldham	36
Figure 33. Ethnic Group Fractions Over Time - Oldham.....	36
Figure 34. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map – Oldham OAs	37

Figure 35. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Oldham MSOAs.....	38
Figure 36. Ethnicity and Total Population Line Chart – Pendle.....	39
Figure 37. Ethnic Group Fractions Over Time - Pendle.....	39
Figure 38. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map - Pendle OAs	40
Figure 39. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Pendle MSOAs.....	41
Figure 40. Ethnicity and Total Population Line Chart - London.....	42
Figure 41. Ethnic Group Fractions and Total Population in London Over Time	42
Figure 42. OA Simpson Index, OA-London Simpson difference, White and Asian Fraction on Map - London OAs	43
Figure 43. MSOA Simpson Index, MSOA-London Simpson difference, Asian and White Moran's I on Map - London MSOAs	44
Figure 44. LAD Simpson Index, LAD-London Simpson difference, Asian Fraction and Asian and White Moran's I on Map - London LADs.....	45
Figure 45. Indices Trends in Selected English Cities (2001-2021).....	47

List of tables

Table 1. Dissimilarity index for all ethnicities at country level.....	7
Table 2. Dissimilarity Index for All Ethnicities at LAD Level	10
Table 3. Simpson Index at Country Level.....	14
Table 4. Simpson Index at LAD Level.....	15
Table 5. Country Level Moran Index in 3 years.....	19
Table 6. Dissimilarity Index at LAD level	20
Table 7. Correlation Coefficient between OA-based Simpson Index and White Fraction	45

List of abbreviations

OA	Output Areas
LSOA	Lower Layer Super Output Areas
MSOA	Middle Layer Super Output Areas
LAD	Local Authority District

Executive Summary

This study presents a detailed analysis of ethnic segregation in England, utilizing census data from 2001, 2011, and 2021. The analysis covers the country level, Local Authority District (LAD) level, and specific cities, including Birmingham, Leicester, Bradford, Blackburn with Darwen, Oldham, Pendle, and London. Three key segregation indices were calculated: the Dissimilarity Index, the Simpson Index, and Moran's I.

There is a notable decrease in the dissimilarity index across the study period (2001-2021), indicating a reduction in ethnic segregation at both the country and LAD levels. This trend suggests improved spatial integration of ethnic groups within England.

The Simpson Index, which measures diversity within neighbourhoods, shows a general decline over the study period. This decrease suggests that neighbourhoods are becoming more ethnically diverse. At the Output Area (OA) level, some neighbourhoods with high ethnic concentrations have higher Simpson Index values, indicating lower diversity within those specific areas. However, when these values are aggregated across entire cities or LADs, the overall Simpson Index is lower, reflecting a balance of diverse neighbourhoods within larger geographic units.

Moran's I, which measures spatial autocorrelation, indicates a decrease at the country level for White, Asian, and Black ethnic groups, suggesting reduced clustering and greater dispersion of these populations.

Conversely, at the OA level within each LAD, Moran's I is increasing for all ethnic groups, indicating that within these smaller areas, ethnic groups are becoming more clustered over time.

In the next phase, Birmingham, Leicester, Bradford, Blackburn with Darwen, Oldham, Pendle, and London were selected for detailed analysis due to significant differences between their town-wide and OA-based Simpson Index values.

In these cities, localized ethnic concentrations result in higher Simpson Index values at the OA level, highlighting pockets of low diversity. However, at the city-wide level, the diversity is higher, reflecting a mix of different neighbourhoods.

1 Introduction

Segregation, particularly residential segregation, refers to the spatial separation of different social groups within a city or metropolitan area. This phenomenon often results in the clustering of individuals from similar socioeconomic backgrounds, ethnicities, or races in distinct neighbourhoods or regions.

Research on segregation is traditionally associated with studies on residential segregation and the patterns of spatial clustering of minority ethnic groups or residents with immigrant backgrounds.

Spatial segregation might happen in various types of spaces, in living places, work, leisure, means of transport and daily encounters. Spatial separation mirrors the way societies are organized, produce and distribute resources. Segregation reflects a preference for living in more homogenous areas and is usually connected with ethnic minority or immigrant communities (Piekut, 2021).

In this thesis, the ethnic segregation the UK, will be computed and its changes during the time will be observed. Although English neighbourhoods have become more ethnically diverse, ethnic segregation is still prevalent and the extent of segregation varies greatly by ethnicity. Immigrants from particular ethnic groups tend to co-locate in space forming segregated communities because of their limited opportunities in an urban area. The segregation of minority populations in England varies among ethnic groups (Patias, Rowe and Arribas-Bel, 2023)

2 Data Source

This analysis uses the aggregated Census data (2001, 2011 and 2021) which is available in geographical areas used in National Statistics work. The data of ethnic groups are available in OA, LSOA and MSOA.

Output Areas (OA) were created for Census data, specifically for the output of census estimates. The OA is the lowest geographical level at which census estimates are provided, typically encompassing approximately 300 people.

Super Output Areas (SOA) are a geography hierarchy designed to improve the reporting of small-area statistics. In England and Wales **Lower Layer SOAs (LSOA)** with a minimum population of 1,000 and **Middle Layer SOAs (MSOA)** with a minimum population of 5,000 were introduced in 2004. LSOAs and MSOAs are of consistent size across the country and will not be subject to regular boundary change.

Local Authority District (LAD): or also known as local authority districts or local government districts (to distinguish from unofficial city districts) are a level of subnational division of England used for the purposes of local government ('Districts of England', 2024).

Spatially, the Output Areas (OAs) exhibit a fine-grained nature, with a median radius of 105 meters. To enhance social homogeneity, the OAs were combined into LSOAs. Typically, each LSOA encompasses five OAs. The units at each scale are nested hierarchically within one another. Within this spatial framework, we examine the distributions of four ethnic groups. These ethnic categories stem from a larger pool of 16 to 19 groups, each representing distinct classifications based on various criteria such as country of ancestry and skin colour (Jones *et al.*, 2015).

The datasets used in this analysis were gotten from:

- **Census Data:** The dataset provides information about the ethnic group of the usual resident population of England and Wales as at census day. It contains the OA codes and the observation of ethnic groups in each OA(*Census - Office for National Statistics*, no date).
- **Lookup:** This file links and aggregate data at different geographic levels (e.g., from OA to larger areas like LSOAs or MSOAs) for spatial analysis (*Open Geography Portal*, no date).
- **Census boundaries (Shape Files):** This file contains the polygons of boundaries of each OA(*Open Geography Portal*, no date).

In this section, an analysis of segregation patterns in the UK at both the country and Local Authority District (LAD) levels is conducted across three different years: 2001, 2011, and 2021.

3 Data Preprocessing

The census data available is available for years of 2001, 2011, 2021 without any missing values. But in order to perform the analysis, the data should be manipulated. in this section, the preprocessing of each year is briefly explained.

3.1 Data of 2001

In 2001, the primary census data exhibited the following structure: There were 175,434 rows and 18 columns, without any missing values. The initial two columns contained OA data, while the subsequent 16 columns represented the count of 16 ethnic groups within each OA. These groups were later categorized into four main ethnicities: White, Asian, Black, and Other.

To extract the higher levels data from OA level, a lookup file was employed. This file contained OA, LSOA and MSOA codes, as well as names for each LSOA and MSOA. However, there was a challenge in merging census dataset and lookup file as the census data utilized old OA codes, while the lookup file employed new OA codes. To resolve this disparity, a connection file was used to establish the relationship between the two datasets.

Additionally, in 2001, the lookup file lacked LAD codes and names. Since the LAD codes were necessary for further analysis, the 2011 lookup file was utilized to append LAD codes and names to the 2001 lookup data.

At the largest scale (LAD), which ranges in population from 2,151 to 977,057, there are 348 areas. At the smallest areal scale, there are 175,434 output areas (OAs), with population ranging from 96 to 4,157 and an average of 296.64. These OAs were aggregated into 34,378 lower-level super output areas (LSOAs) and 7,194 middle-level super output areas (MSOAs), with mean populations of 1,513.08 (SD = 202.7) and 7,234.04 (SD = 1,319.98), respectively. The total population was 52,041,655.

3.2 Data of 2011

In 2011, the primary census data exhibited the following structure: There were 181,408 rows and 19 columns, without any missing values. The first column contained OA data, while the

subsequent 18 columns represented the count of 18 ethnic groups within each OA. The same as 2001, these groups were later categorized into four main ethnicities: White, Asian, Black, and Other.

The lookup file in 2011 contained the whole codes and names for OA, LSOA, MSOA and LAD so these 2 datasets were combined to extract the higher levels data from OA level.

At the largest scale (LAD), which ranges in population from 2,203 to 1,073,045, there are 348 areas. At the smallest areal scale, there are 181,408 OAs, with population ranging from 91 to 4,140 and an average of 309.11. These OAs were aggregated into 34,753 LSOAs and 7,201 MSOAs, with mean populations of 1,613.55 ($SD = 303.38$) and 7,787.24 ($SD = 1,599.63$), respectively. The total population was 56,075,912.

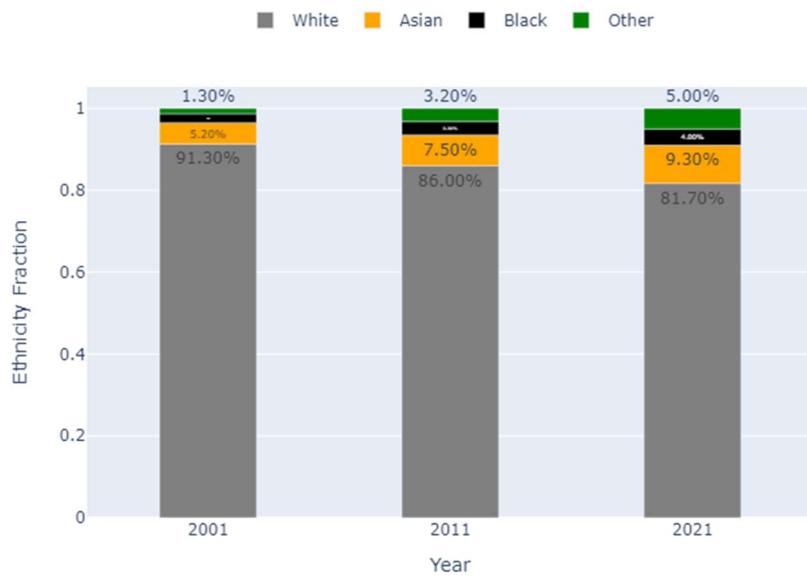
3.3 Data of 2021

In 2021, the primary census data exhibited the following structure: There were 3,777,600 rows and 5 columns, with no missing values. The first 2 columns contained OA data. However, the structure of this dataset differed from previous datasets. The ethnicity group, which consisted of 20 categories, was consolidated into one column, with the count of each ethnicity group in each OA recorded in the observation column. To maintain consistency, this dataset was converted to the format of previous datasets and merged with the lookup file.

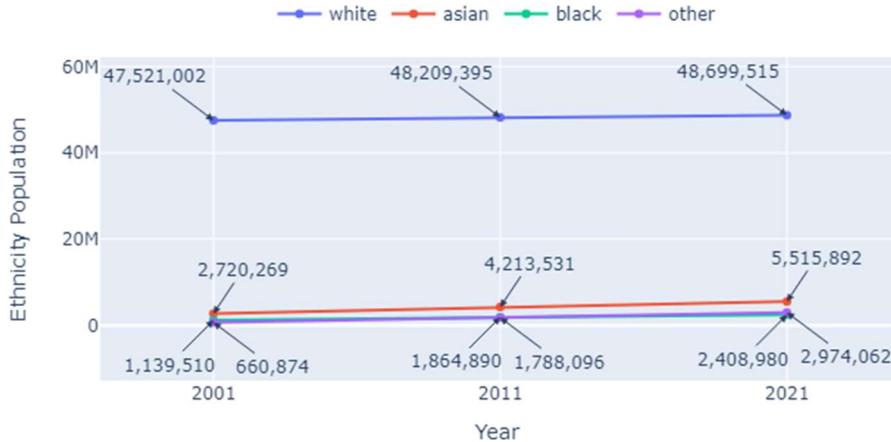
At the largest scale (LAD), which ranges in population from 2,054 to 1,144,951, there are 331 areas. At the smallest areal scale, there are 188,880 OAs, with population ranging from 97 to 4,500 and an average of 315.53. These OAs were aggregated into 35,672 LSOAs and 7,264 MSOAs, with mean populations of 1,670.73 ($SD = 354.21$) and 8,204.63 ($SD = 1,815.62$), respectively. The total population was 59,598,449.

4 Population Trends in England

In Figure 1, which illustrates the ethnic fractions within the population, it is noticeable that the proportion of the White ethnicity is declining, whereas the fractions of Asian, Black, and Other ethnic groups are on the rise.

**Figure 1. Ethnic Group Fractions over Time**

In Figure 2, which illustrates the population changes across different ethnic groups, it is evident that the Asian ethnic group exhibits the highest growth rate compared to other ethnicities. The White and Black ethnic groups demonstrate the lowest growth rates alongside the population growth,

**Figure 2. Ethnicity population Line Chart**

5 Dissimilarity Index

Dissimilarity index is a demographic measure of the evenness with which two groups are distributed across component geographic areas that make up a larger area. A group is evenly distributed when each geographic unit has the same percentage of group members as the total population. A score of zero reflects a fully integrated environment; a score of 1 reflects full

segregation. In terms of black–white segregation, a score of 0.60 means that 60 percent of blacks would have to exchange places with whites in other units to achieve an even geographic distribution ('Index of dissimilarity', 2023).

This index is calculated using the following formula:

$$D^X = \frac{\sum_i \frac{N_i}{N_{Area}} |P_i^X - P_{Area}^X|}{2P_{Area}^X(1 - P_{Area}^X)}$$

where P_i^X indicates the fraction of individuals for ethnic group X in sub_area_i , P_{Area}^X the fraction in the whole area, N_i is the number of individuals in sub_area_i and N_{Area} is the number of individuals in the whole geographical area. The denominator is to normalize the Dissimilarity index to lie between zero and one (Zuccotti *et al.*, 2023)

In this analysis, the index is calculated for each ethnic group at both the country and LAD levels.

At the country level, the country serves as the main area, with OA, LSOA, MSOA, and LAD considered as sub-areas.

At the LAD level, each LAD is treated as the main area, with OA, LSOA, and MSOA considered as sub-areas within each LAD.

The D value is a function of how the population in each of the four groups is distributed among areal units. If all

areas have the same proportion among the four groups, then there will be no segregation, and D will be zero. If one group exclusively occupies each areal unit, then D will be one, indicating a perfectly segregated situation.

When D is calculated for a given region but with data at different aggregation levels, the results are usually not consistent over scale levels. When the areal units are smaller, more variation in the population distribution is captured by the detailed data

and thus, D tends to be higher. At the more aggregated level, less variation is found in the data and therefore D is lower (Wong, 2003).

In the next part of the study, the dissimilarity index is calculated at country level and LAD level.

5.1 Dissimilarity Index at Country Level

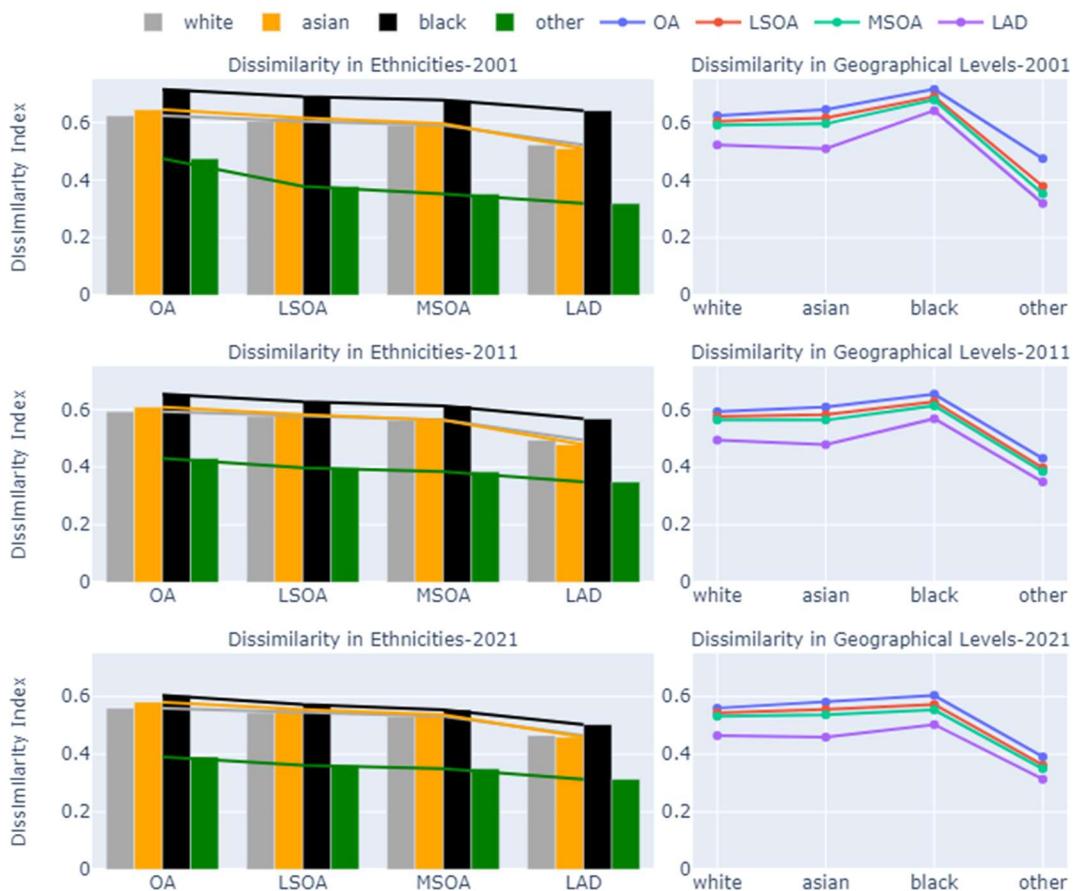
For country level, the dissimilarity index is calculated four times, each time considering OA, LSOA, MSOA, and LADs as the subareas. The results are shown in Table 1.

Table 1. Dissimilarity index for all ethnicities at country level

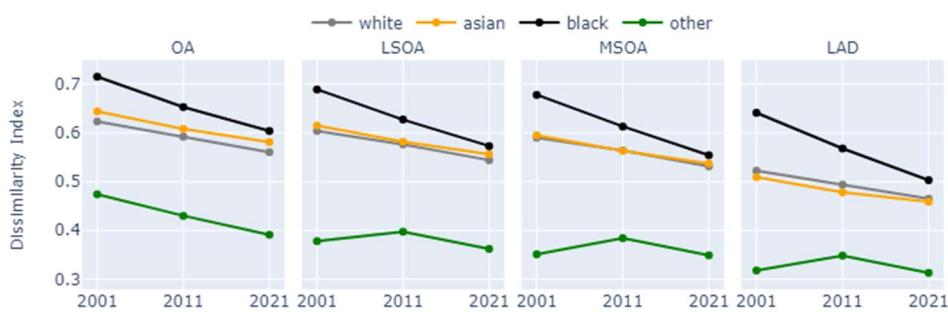
year	white_frac	OA_white	LSOA_white	MSOA_white	LAD_white
2001	0.913	0.623	0.604	0.59	0.522
2011	0.86	0.592	0.576	0.564	0.494
2021	0.817	0.56	0.544	0.531	0.465
year	asian_frac	OA_asian	LSOA_asian	MSOA_asian	LAD_asian
2001	0.052	0.644	0.615	0.595	0.509
2011	0.075	0.608	0.582	0.563	0.478
2021	0.093	0.581	0.556	0.537	0.459
year	black_frac	OA_black	LSOA_black	MSOA_black	LAD_black
2001	0.022	0.715	0.689	0.678	0.641
2011	0.033	0.653	0.627	0.613	0.568
2021	0.04	0.604	0.573	0.554	0.503
year	other_frac	OA_other	LSOA_other	MSOA_other	LAD_other
2001	0.013	0.474	0.378	0.351	0.318
2011	0.032	0.43	0.397	0.384	0.348
2021	0.05	0.391	0.362	0.349	0.313

In Figure 3, the Dissimilarity index for all ethnic groups at the country level is presented. As depicted in the bar chart segment of the plot, across all geographical levels and for each year, the "Other" ethnic group consistently exhibits the lowest dissimilarity index, while the "Black" ethnic group consistently demonstrates the highest dissimilarity index. Notably, the disparity in dissimilarity index between the "Asian" and "White" ethnic groups is not significant, and in some instances, the dissimilarity index of the "White" ethnic group surpassed that of the "Asian" ethnic group. This trend is particularly evident at the LAD Level in all years of the study and at MSOA levels in 2011.

In the line chart of Figure 3, it is evident that the dissimilarity index decreases across all ethnic groups with increasing geographical area. This trend underscores the correlation between population size and dissimilarity, demonstrating that as population density increases, dissimilarity tends to decrease.

**Figure 3. Dissimilarity Index Variation Across Geographical Levels**

The line chart shown in Figure 4 illustrates a decreasing trend in the Dissimilarity index for the ethnic groups of "white", "Asian" and "black" over time across each geographical level. The dissimilarity index for the black ethnic group, which initially has the highest value, demonstrates a steeper decline compared to the white and Asian ethnic groups. For the "Other" ethnic group, a decreasing pattern is observed solely at the OA level. However, at other geographical levels, an increase is noted in 2011, followed by a decrease in 2021.

**Figure 4. Dissimilarity Index Variation Over the Time Across Geographic Levels**

5.2 Dissimilarity Index at LAD Level

For the LAD level, OAs, LSOAs, and MSOAs were considered as the subareas of interest. The results are presented in Table 2. This comprehensive approach allowed for a detailed examination of segregation trends over time and across different geographic scales. The table contains the LADs with the highest dissimilarity index for 'White' ethnic group considering OA as the sub-area in each year. The full table can be found in the appendix.

When the sub-district is equivalent to the district division (e.g., MSOA and LAD, or LSOA and LAD), the dissimilarity index of the LAD regarding the sub-area (LSOA or MSOA) becomes 0.

For example, the LAD of Isles of Scilly, characterized by a high proportion of the white ethnic group, exhibits the highest dissimilarity index in 2001 regarding its Output Areas (OAs). However, the dissimilarity index of the LAD regarding its LSOAs and MSOAs is zero. This means that each LSOA or MSOA perfectly fits within a single LAD. When the administrative areas are the same, there will be no dissimilarity between them.

This phenomenon is also observed in the LAD of City of London at the MSOA level, indicating that this LAD comprises only one MSOA.

In addition, when the population of an ethnic group in a district is zero, the value of the dissimilarity index is NaN, as the denominator of the dissimilarity formula would be zero.

Table 2. Dissimilarity Index for All Ethnicities at LAD Level

year	LADNM	LAD_population	white_frac	OA_white	LSOA_white	MSOA_white	asian_frac	OA_asian	LSOA_asian	MSOA_asian
2001	Isles of Scilly	2151	0.996	0.79	0	0	0	NaN	NaN	NaN
	Blackburn with Darwen	137446	0.779	0.727	0.7	0.635	0.209	0.746	0.719	0.652
	Oldham	217266	0.862	0.706	0.682	0.663	0.122	0.772	0.742	0.719
	Bradford	467633	0.783	0.7	0.677	0.631	0.193	0.739	0.711	0.662
	Pendle	89321	0.849	0.695	0.636	0.591	0.143	0.712	0.653	0.605
	LADNM	LAD_population	black_frac	OA_black	LSOA_black	MSOA_black	other_frac	OA_other	LSOA_other	MSOA_other
	Isles of Scilly	2151	0	NaN	NaN	NaN	0.004	0.79	0	0
	Blackburn with Darwen	137446	0.002	0.787	0.435	0.275	0.009	0.427	0.269	0.15
	Oldham	217266	0.006	0.602	0.377	0.282	0.011	0.405	0.273	0.204
	Bradford	467633	0.009	0.518	0.396	0.351	0.015	0.389	0.275	0.214
	Pendle	89321	0.001	0.864	0.604	0.371	0.007	0.494	0.268	0.224
year	LADNM	LAD_population	white_frac	OA_white	LSOA_white	MSOA_white	asian_frac	OA_asian	LSOA_asian	MSOA_asian
2011	Blackburn with Darwen	147489	0.692	0.709	0.677	0.618	0.281	0.726	0.692	0.631
	Oldham	224897	0.775	0.703	0.688	0.646	0.192	0.768	0.75	0.7
	Burnley	87059	0.874	0.7	0.682	0.636	0.11	0.754	0.735	0.688
	Bradford	522452	0.674	0.667	0.645	0.607	0.268	0.692	0.667	0.626
	Pendle	89452	0.799	0.665	0.625	0.588	0.188	0.687	0.643	0.605
	LADNM	LAD_population	black_frac	OA_black	LSOA_black	MSOA_black	other_frac	OA_other	LSOA_other	MSOA_other
	Blackburn with Darwen	147489	0.006	0.538	0.356	0.254	0.021	0.336	0.251	0.182
	Oldham	224897	0.012	0.49	0.387	0.289	0.02	0.314	0.228	0.168
	Burnley	87059	0.002	0.644	0.401	0.25	0.014	0.392	0.29	0.246
	Bradford	522452	0.018	0.471	0.401	0.354	0.04	0.33	0.265	0.228
	Pendle	89452	0.001	0.755	0.447	0.297	0.012	0.345	0.22	0.192
year	LADNM	LAD_population	white_frac	OA_white	LSOA_white	MSOA_white	asian_frac	OA_asian	LSOA_asian	MSOA_asian
2021	Blackburn with Darwen	154761	0.604	0.712	0.683	0.618	0.357	0.722	0.695	0.625
	Burnley	94673	0.825	0.712	0.698	0.635	0.147	0.767	0.751	0.682
	Pendle	95804	0.706	0.671	0.646	0.627	0.267	0.688	0.658	0.636
	Bradford	546416	0.611	0.654	0.632	0.594	0.321	0.661	0.636	0.596
	Oldham	242056	0.681	0.653	0.637	0.588	0.246	0.71	0.686	0.63
	LADNM	LAD_population	black_frac	OA_black	LSOA_black	MSOA_black	other_frac	OA_other	LSOA_other	MSOA_other
	Blackburn with Darwen	154761	0.009	0.554	0.389	0.315	0.031	0.301	0.214	0.167
	Burnley	94673	0.003	0.624	0.411	0.26	0.024	0.364	0.28	0.239
	Pendle	95804	0.003	0.641	0.352	0.227	0.025	0.323	0.242	0.217
	Bradford	546416	0.02	0.479	0.403	0.36	0.047	0.295	0.23	0.197
	Oldham	242056	0.034	0.491	0.419	0.333	0.039	0.299	0.234	0.199

To better understand the dissimilarity index, a box plot is used. In the box plot represented by Figure 5, it is evident that moving to larger geographic areas, the dissimilarity index decreases

for all ethnic groups. Additionally, areas with a predominant fraction of a single ethnic group tend to exhibit higher dissimilarity indices.

By comparing OAs with other levels, it is observed that the changes in the dissimilarity index for OAs from 2001 to 2011 were more significant compared to the changes for LSOAs and MSOAs.



Figure 5. Dissimilarity Index Box-Plot by Ethnicity and Area

The following maps are visual representations of the Dissimilarity Index on the map of England, based on OA Dissimilarity Index. The dissimilarity index quantifies the degree to

which two groups are distributed across neighbourhoods. This mapping highlights geographic areas with higher or lower levels of segregation. In Figure 6, it can be observed that, generally for all ethnic groups, from 2001 to 2011, the dissimilarity index is decreasing, indicating more even distribution in LADs. However, from 2011 to 2021, this change is not very significant. There are some areas in the northern part of England (including Bradford, Blackburn with Darwen, Pendle, Oldham, etc.) that have higher dissimilarity indices, indicating that in the OAs corresponding to these LADs, the selected ethnicity is not evenly distributed.

The Black ethnicity has higher dissimilarity values in most LADs. In some areas, especially around London, the dissimilarity value even for black ethnic group is low, indicating that ethnicities are more evenly distributed in the LSOAs corresponding to the LADs.

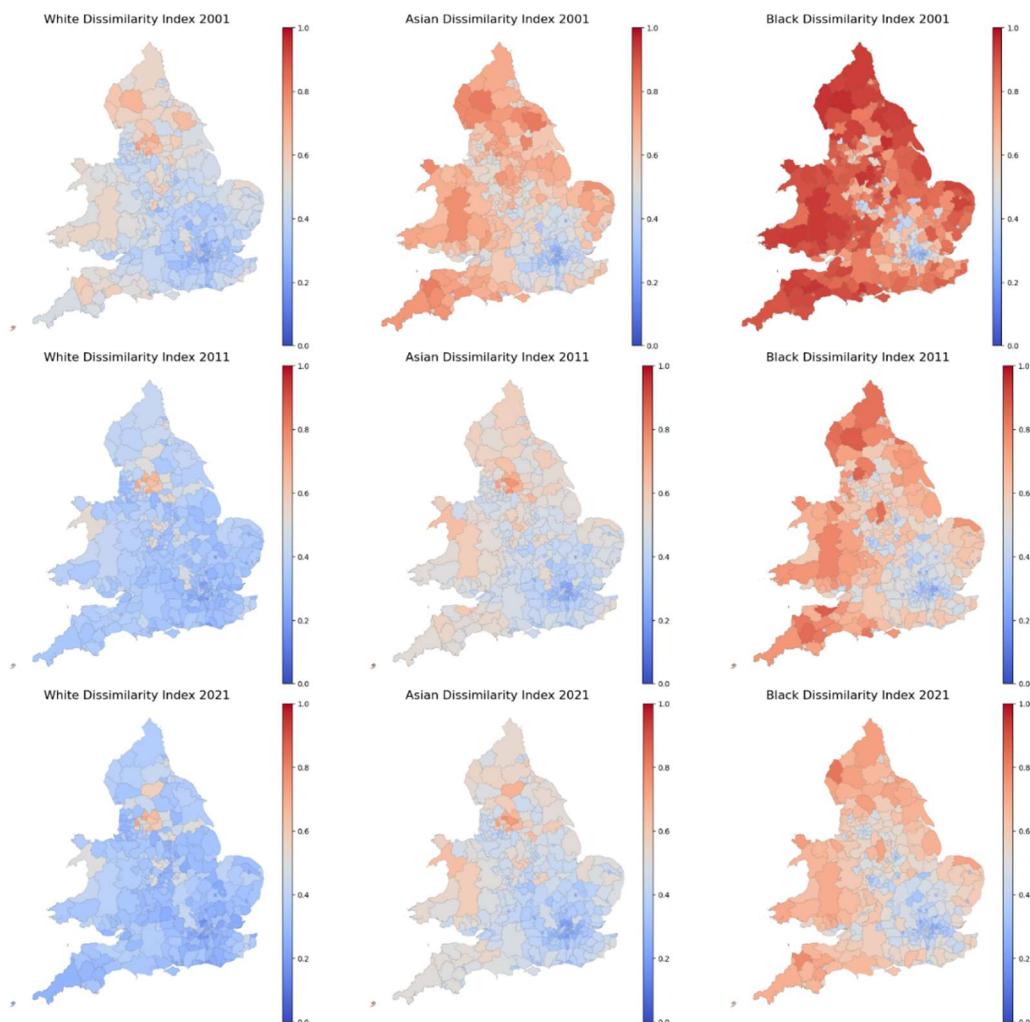


Figure 6. OA-based Dissimilarity Index for all ethnic groups in LADs

6 Simpson Index

Simpson Index, which is a measure of ethnic concentration, can be calculated for each OA, LSOA, MSOA, LAD and for the whole country. It expresses the probability that two randomly selected individuals from the geographic area have the same ethnicity.

This Average Local Simpson Index is calculated using the following formula:

$$S = \sum_i \left(\frac{N_i}{N_{Area}} \sum_X (P_i^X)^2 \right)$$

Calculated for each sub_area_i as $S_i = \sum_X (P_i^X)^2$. The lower S_i , the more ethnically diverse is the sub_area_i . The maximal value is achieved when the district is populated exclusively by one ethnic group. To calculate this index, only White British, Asians, and Blacks are considered, and “Other” ethnic groups are left outside. It has been done upon the consideration that individuals from this group are less likely to recognize each other as similar in terms of ethnic preferences. The Simpson Index is mathematically closely related to entropy, a measure of disorder which is also sometimes used in segregation studies to quantify the unequal distribution of ethnic groups between neighbourhoods or between the global and the local level. (Zuccotti *et al.*, 2023)

In this analysis, Simpson Index is also calculated at both LAD and Country level.

At the country level, the country serves as the main area, with each of the OA, LSOA, MSOA, and LAD considered as sub-areas.

At the LAD level, each LAD is treated as the main area, with OA, LSOA, and MSOA considered as sub-areas within each LAD.

6.1 Simpson Index at Country Level

In Table 3, the Simpson index at the country level for each year is presented.

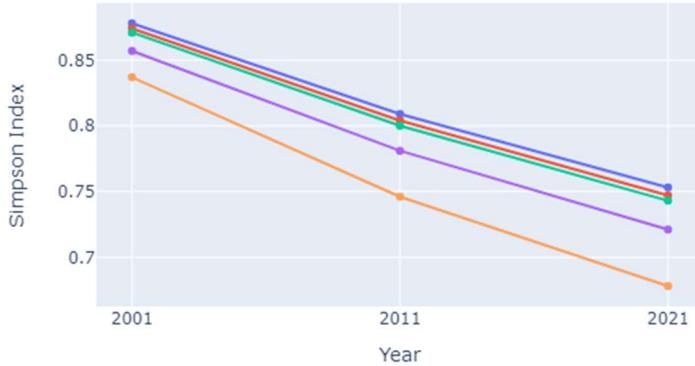
As depicted in Figure 7, the Simpson index is decreasing over time at all geographical levels.

In addition, this index decreases as it moves to larger geographical areas.

Table 3. Simpson Index at Country Level

year	total_population	white_frac	asian_frac	black_frac	other_frac	OA	LSOA	MSOA	LAD	country
2001	52041655	0.913	0.052	0.022	0.013	0.878	0.874	0.871	0.857	0.837
2011	56075912	0.86	0.075	0.033	0.032	0.809	0.804	0.8	0.781	0.746
2021	59598449	0.817	0.093	0.04	0.05	0.753	0.747	0.743	0.721	0.678

— OA — LSOA — MSOA — LAD — country

**Figure 7. Simpson Index Line Chart at Country Level**

6.2 Simpson Index at LAD Level

In this part of the study, the Simpson index is calculated at LAD level. Therefore, OAs, LSOAs, and MSOAs were considered as the subareas of interest. The results are presented in Table 4.

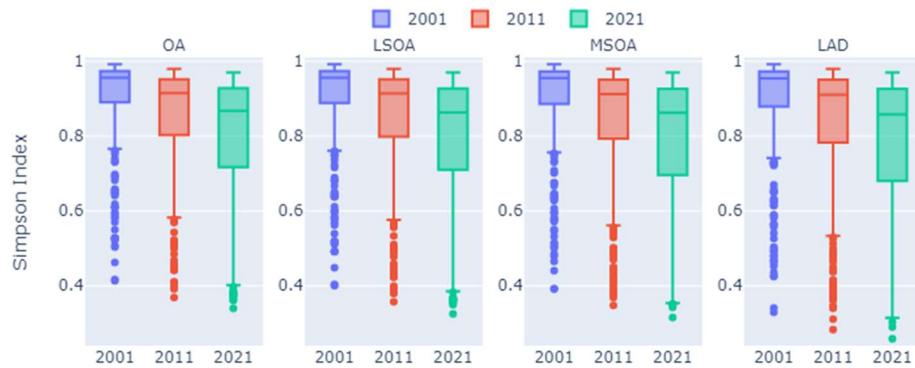
The table contains the LADs with the highest Simpson index considering OA as the sub-area in each year. The full table can be found in the appendix.

By sorting the table based on the lowest Simpson index, it is noticeable that the LADs included as part of London (LAD codes between E09000001 and E09000033) have the lowest Simpson index in England. This suggests that these areas can be considered less segregated.

Table 4. Simpson Index at LAD Level

year	LADNM	LADCD	LAD_population	white_frac	asian_frac	black_frac	other_frac	OA	LSOA	MSOA	LAD	OA_LAD_diff	LSOA_LAD_diff	MSOA_LAD_diff
2001	Blackburn with Darwen	E06000008	137446	0.779	0.209	0.002	0.009	0.832	0.815	0.779	0.651	0.181	0.164	0.128
	Leicester	E06000016	279832	0.639	0.307	0.031	0.023	0.682	0.667	0.645	0.503	0.179	0.164	0.142
	Bradford	E08000032	467633	0.783	0.193	0.009	0.015	0.817	0.8	0.775	0.65	0.167	0.15	0.125
	Birmingham	E08000025	977057	0.704	0.207	0.061	0.029	0.698	0.689	0.676	0.541	0.157	0.148	0.135
	Oldham	E08000004	217266	0.862	0.122	0.006	0.011	0.873	0.859	0.845	0.757	0.116	0.102	0.088
2011	Blackburn with Darwen	E06000008	147489	0.692	0.281	0.006	0.021	0.783	0.765	0.728	0.558	0.225	0.207	0.17
	Bradford	E08000032	522452	0.674	0.268	0.018	0.04	0.733	0.717	0.691	0.527	0.206	0.19	0.164
	Leicester	E06000016	329839	0.505	0.371	0.062	0.061	0.582	0.567	0.548	0.397	0.185	0.17	0.151
	Oldham	E08000004	224897	0.775	0.192	0.012	0.02	0.823	0.808	0.791	0.638	0.185	0.17	0.153
	Birmingham	E08000025	1073045	0.579	0.266	0.09	0.065	0.577	0.566	0.553	0.415	0.162	0.151	0.138
2021	Blackburn with Darwen	E06000008	154761	0.604	0.357	0.009	0.031	0.755	0.736	0.696	0.491	0.264	0.245	0.205
	Bradford	E08000032	546416	0.611	0.321	0.02	0.047	0.697	0.681	0.656	0.477	0.22	0.204	0.179
	Oldham	E08000004	242056	0.681	0.246	0.034	0.039	0.733	0.716	0.696	0.526	0.207	0.19	0.17
	Pendle	E07000122	95804	0.706	0.267	0.003	0.025	0.758	0.733	0.716	0.569	0.189	0.164	0.147
	Leicester	E06000016	368531	0.409	0.434	0.078	0.079	0.537	0.523	0.509	0.362	0.175	0.161	0.147

In Figure 8, the boxplot of the Simpson index is displayed. As shown, moving from 2001 to 2021, the Simpson indices for each geographical area are decreasing. Additionally, it is noticeable that the LAD-based Simpson index has slight differences compared to other subareas.

**Figure 8. Simpson Index Box-Plot based on different areas across the years**

In Figure 9, the box plot of these differences is displayed. The differences between the LAD-based Simpson index of LADs and other subareas (OA, MSOA and LSOA) highlight variations in segregation patterns across different geographic scales. Comparing these indices

provides insights into how segregation patterns differ within larger administrative areas (LAD) versus smaller geographic units (OA or LSOA). As observed, the difference between LADs and other geographical areas is mostly less than 0.04, decreasing as the geographic area expands and increasing in recent years. However, notable outliers in 2001 include Birmingham, Leicester, Bradford, and Blackburn with Darwen, where the difference is relatively high. Oldham is added to this list in 2011, and Pendle joins in 2021.

This indicates that the disparities between the LAD and other subareas are most pronounced in these cities. Consequently, these cities have been selected for deeper analysis in the subsequent steps of the project. Additionally, London, being the capital and inherently interesting, is included in the study. Therefore, the 33 LADs corresponding to London have been selected for further examination.



Figure 9. Simpson Index Differences between LAD and other areas across the years

To better understand these numbers, a map visualization is used. As observed in Figure 10, from 2001 to 2021, the Simpson index has declined. In the border areas of England, the Simpson index is relatively high compared to the central areas. Conversely, some central parts adjacent to Bradford, Birmingham, and London demonstrate a lower Simpson index. Figure 11 depicts the difference between the LAD-based Simpson index and the OA-based Simpson index for each LAD. By comparing these two figures, it is evident that in these areas, the difference between the LAD-based Simpson index and the OA-based Simpson index is quite significant.

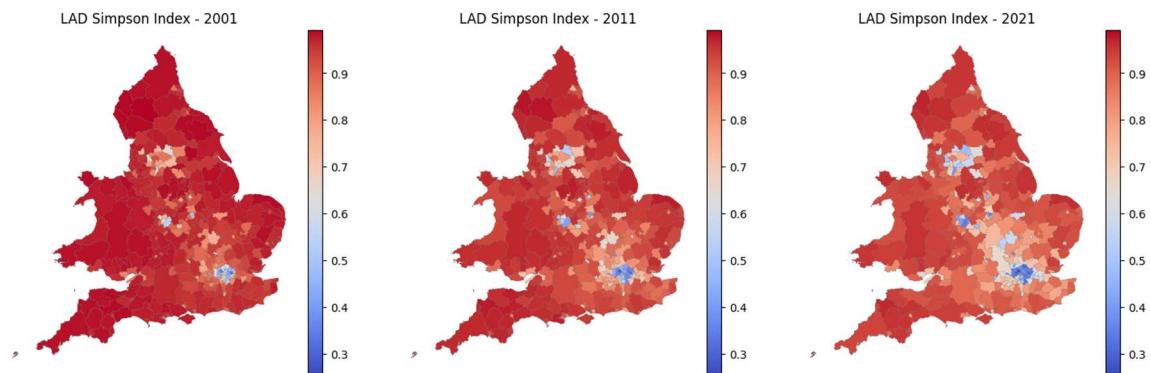


Figure 10. LADs' Simpson Index on Map of England

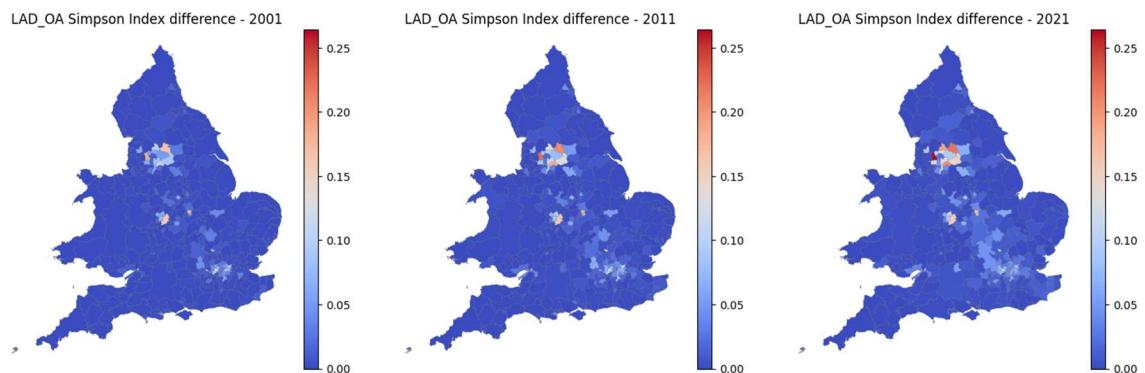


Figure 11. Difference of LAD-OA Simpson Indices on Map of England

7 Moran's I Index

Moran's I index is a statistical measure used to assess spatial autocorrelation. Spatial Autocorrelation means looking at how similar or dissimilar values of a variable are in neighbouring locations. Spatial autocorrelation measures this similarity or dissimilarity in geographic space. It quantifies the degree of similarity between values of a variable at different locations in a spatial dataset. Moran's I index determines whether the observed pattern of spatial autocorrelation indicates clustering (similar values close together), dispersion (dissimilar values close together), or randomness (no discernible pattern). This information is valuable for understanding the underlying spatial processes and making informed decisions in various fields such as urban planning, epidemiology, and environmental science.

This Moran's I Index is calculated using the following formula:

$$Moran - I^X = \frac{L}{W} \frac{\sum_i \sum_j w_{ij} (P_i^X - \bar{P}^X)(P_j^X - \bar{P}^X)}{\sum_i (P_i^X - \bar{P}^X)^2}$$

Where P_i^X is fraction of individuals for ethnic group X in $district_i$ and \bar{P}^X is the average of fraction of the ethnic group X in all $district_i$. The weights w_{ij} are one when $district_i$ and j share a border and zero otherwise. Finally, L is the number of $district$ and W the number of positive w_{ij} (Zuccotti *et al.*, 2023).

Similar to the Dissimilarity index, Moran's I is computed independently for each ethnic group at both the national and Local Authority District (LAD) levels. To compute this index, spatial data is required to delineate the shared boundaries between each subordinate unit.

At the country level, the entire country is regarded as the primary unit, while LADs are treated as subordinate units.

Conversely, at the LAD level, each LAD serves as the principal unit, with OAs, LSOAs, and MSOAs considered as subordinate units within each LAD.

7.1 Moran's I at country level

To compute the Moran's I index, the shared border between each subarea must be determined. computing the shared borders of 175,000 to 180,000 OA areas is computationally intensive, so at the country level, LADs are considered as subareas, and other geographical areas were ignored. Table 5 shows the Moran's I index for each ethnic group across three years. It reflects trends in how these populations are spatially organized within the studied area.

The index for the white population shows a slight decrease from 0.630 in 2001 to 0.617 in 2021. This indicates that the spatial clustering of the white population has slightly decreased over time, suggesting a slight move towards a more random distribution.

The index for the asian population decreases from 0.502 in 2001 to 0.478 in 2021. This suggests a slight decrease in spatial clustering, indicating that the asian population has become slightly more evenly distributed over the years.

The Moran's I index for the black population decreases from 0.616 in 2001 to 0.580 in 2021. This indicates a reduction in spatial clustering, suggesting that the black population has become more evenly spread out over time.

The Moran's I index for the "other" category increases from 0.669 in 2001 to 0.750 in 2021. This indicates an increase in spatial clustering, suggesting that the "other" population has become more clustered together over the years.

Table 5. Country Level Moran Index in 3 years

year	white_moran	asian_moran	black_moran	other_moran
2001	0.63	0.502	0.616	0.669
2011	0.621	0.483	0.602	0.703
2021	0.617	0.478	0.58	0.75

As it is shown in Figure 12, for the white, Asian, and black populations, the Moran's I index values are all positive and show a slight decreasing trend, indicating that while there is still a moderate degree of clustering, the populations are becoming slightly more dispersed over time.

For the "other" category, the Moran's I index values are positive and show an increasing trend, indicating that the spatial clustering of this population is becoming stronger over time.

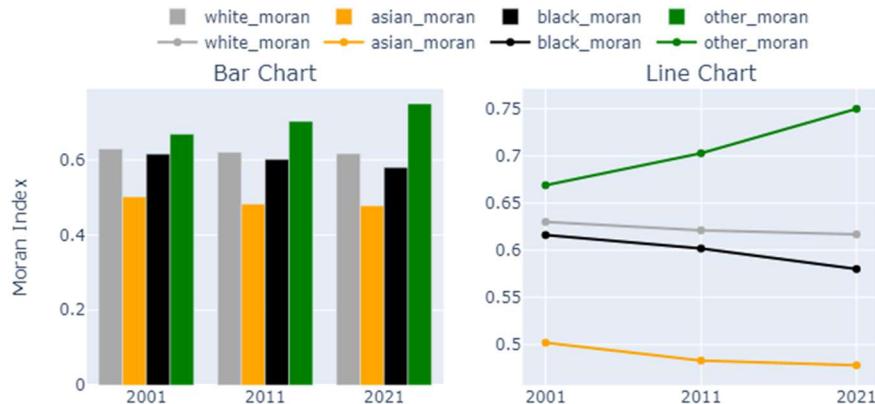


Figure 12. Moran Index Variation Across the Year

7.2 Moran's I at LAD level

In this part of the study, Moran's I index is calculated for each LAD. For LADs, Moran's I index is computable for OA, LSOA and MSOA levels.

The table contains the LADs with the highest Moran's I index for 'White' ethnic group considering LSOA as the sub-area in each year. The full table can be found in the appendix.

Table 6. Dissimilarity Index at LAD level

year	LADNM	LADCD	LAD_population	white_frac	asian_frac	black_frac	other_frac	LSOA_white_mor	LSOA_asian_mor	LSOA_black_mor	LSOA_other_mor
2001	Birmingham	E08000025	977057	0.704	0.207	0.061	0.029	0.92	0.878	0.835	0.557
	Croydon	E09000008	330688	0.701	0.128	0.133	0.037	0.919	0.801	0.907	0.488
	Redbridge	E09000026	238511	0.636	0.265	0.076	0.024	0.85	0.819	0.641	0.144
	Enfield	E09000010	273616	0.771	0.094	0.105	0.03	0.849	0.717	0.857	0.461
	Ealing	E09000009	300886	0.587	0.289	0.088	0.036	0.846	0.876	0.364	0.347
2011	Birmingham	E08000025	1073045	0.579	0.266	0.09	0.065	0.919	0.874	0.75	0.676
	Croydon	E09000008	363378	0.551	0.164	0.202	0.083	0.902	0.753	0.882	0.64
	Derby	E06000015	248752	0.803	0.125	0.029	0.043	0.866	0.819	0.787	0.738
	Enfield	E09000010	312466	0.61	0.112	0.172	0.106	0.861	0.677	0.87	0.592
	Hounslow	E09000018	253957	0.514	0.344	0.066	0.076	0.844	0.863	0.175	0.211
2021	Birmingham	E08000025	1144951	0.486	0.31	0.11	0.094	0.905	0.87	0.713	0.666
	Derby	E06000015	261338	0.738	0.157	0.04	0.065	0.875	0.821	0.676	0.743
	Leicester	E06000016	368531	0.409	0.434	0.078	0.079	0.855	0.866	0.471	0.437
	Bradford	E08000032	546416	0.611	0.321	0.02	0.047	0.849	0.82	0.671	0.654
	Croydon	E09000008	390701	0.484	0.175	0.226	0.115	0.841	0.708	0.773	0.572

The NaN value for each subarea's Moran's I indicates subareas with a population count of zero for a specific ethnicity.

In Figure 13, the boxplot of LSOA-based Moran's I index is displayed. LSOA is selected as it provides a good balance between detailed and aggregated data. As shown, moving from 2001 to 2021 for each ethnic group, the Moran's I index is increasing, indicating that the spatial autocorrelation within each Local Authority District (LAD) is strengthening. For all ethnic groups, fewer than 25% of LADs exhibit a negative Moran's I index, suggesting that the majority of LADs have positive spatial autocorrelation. This occurs when similar values (e.g., high or low percentages of a particular ethnic group) are clustered together in a geographic space. In other words, areas with high fraction of each ethnic group are surrounded by other areas with high fractions, and areas with low fractions are surrounded by other areas with low fractions. A positive Moran's I index indicates this pattern of clustering.

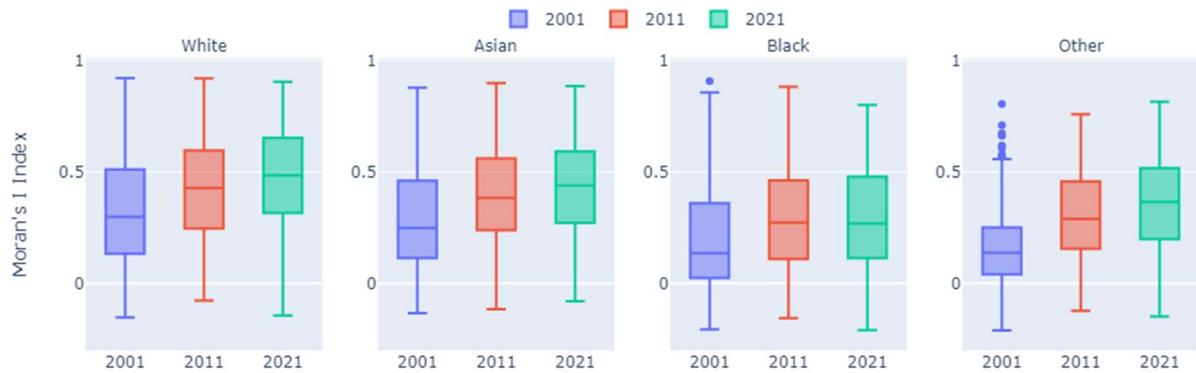


Figure 13. LSOA-based Moran's I in LADs for Ethnicities across the years

To get a better overview, these values are plotted on the map of England in in Figure 14. As it can be observed, in 2001, the central parts of England exhibit higher positive values, indicating that within these LADs, OAs with similar fractions of White, Asian and Black ethnicities are adjacent. Moving to 2021, more LADs show higher positive values. As an example, in 2021, almost 50% of LADs have a Moran's I index greater than +0.5 for White ethnicity (percentage based on box plot), indicating increased clustering and segregation of these ethnic groups.



Figure 14. OA-based Moran's I in LADs on Map of England

8 Detailed Examination of Segregation in Selected English Cities

In this part of the study, a deeper analysis of segregation in selected cities across England is conducted. These cities were chosen based on significant differences between the Simpson Index at the Local Authority District (LAD) level and the Output Area (OA) level, indicating notable variations in the spatial distribution of ethnic groups within these regions. Additionally, London is included in this analysis due to its prominence as the capital city and its diverse demographic landscape.

As mentioned earlier, the "other" ethnic group was excluded from the calculation of the Simpson Index in all cities and at all levels.

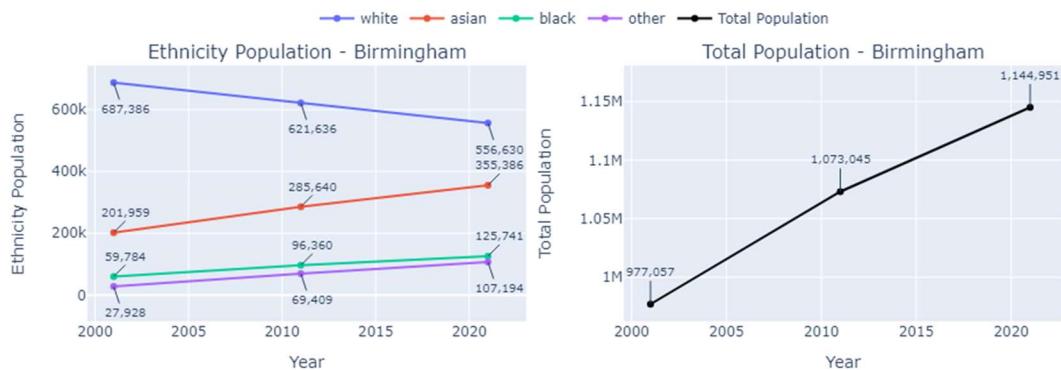
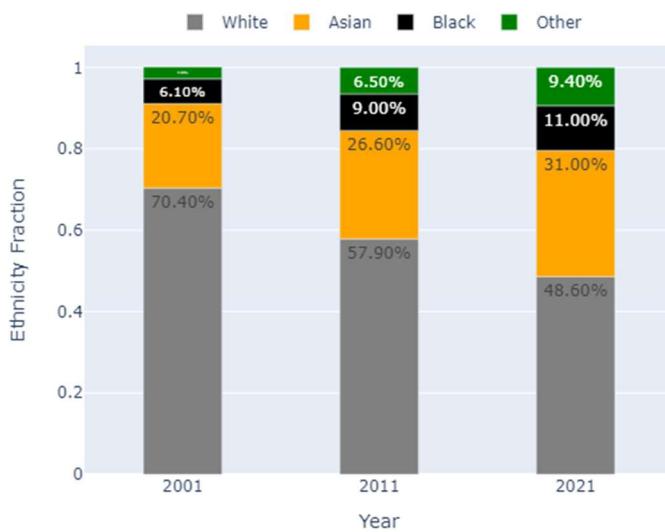
To understand how diverse the population is distributed within the cities, the data in different levels are plotted on a map.

In addition, in each LAD, the OA-based Moran's I index is calculated and reflected on the city map. For calculating the Moran's I index for each MSOA, the OAs inside each MSOA were considered as the sub-areas. The choice of OAs over LSOAs for evaluating the Moran's I index within MSOAs stems from their differing spatial scales and levels of aggregation. OAs, being the smallest geographic units in census data, provide a finer resolution of spatial analysis with each unit typically encompassing a population of around 100-150 households or roughly 300 people. This finer granularity allows for more precise detection of spatial clustering phenomena. For instance, the positive OA-based Asian Moran's I index indicates a tendency for ethnic groups to cluster together within neighbouring OAs, reflecting localized spatial autocorrelation where similar ethnic compositions are grouped together.

In contrast, LSOAs are larger units aggregating multiple OAs, typically containing about 1,000-3,000 people or 400-1,200 households. The aggregation of diverse OAs within LSOAs results in a broader spatial scope but can lead to a dilution of localized patterns observed at the OA level. At the LSOA level, the spatial analysis may show more dispersed patterns or spatial dissimilarity because of the mixing of different local compositions within each LSOA. This is reflected in the negative LSOA-based Moran's I values, indicating a lack of clustering or dispersion of high and low ethnic composition values within neighbouring LSOAs. Therefore, by focusing on OAs within MSOAs, we gain a more detailed understanding of local spatial patterns and clustering tendencies within the finer geographic units.

8.1 Birmingham

Birmingham, the second largest city in England, has a population where less than 50% is composed of the white ethnic group. According to Figure 15, Birmingham's total population increased from 977,057 in 2001 to 1,144,951 in 2021. During this period, the population of the white ethnic group decreased from 687,386 to 556,630. As shown in Figure 16, the fraction of the white ethnic group decreased from 70% of the population in 2001 to 48% in 2021. Among all ethnicities, the Asian population has experienced the steepest growth rate.

**Figure 15. Ethnicity and Total Population Line Chart – Birmingham****Figure 16. Ethnic Group Fractions Over Time - Birmingham**

8.1.1 Birmingham Map – OA Level

Figure 17 shows the Simpson Index for each OA and its difference from the town-wide Simpson Index in Birmingham. In the first column, which depicts the Simpson Index of OAs, the overall Simpson Index is decreasing (more OAs are moving toward blue).

In the second column which shows the Simpson Index difference with town-wide Simpson Index, this difference increased in some OAs while decreased in some others. However, especially in central areas, this difference increased.

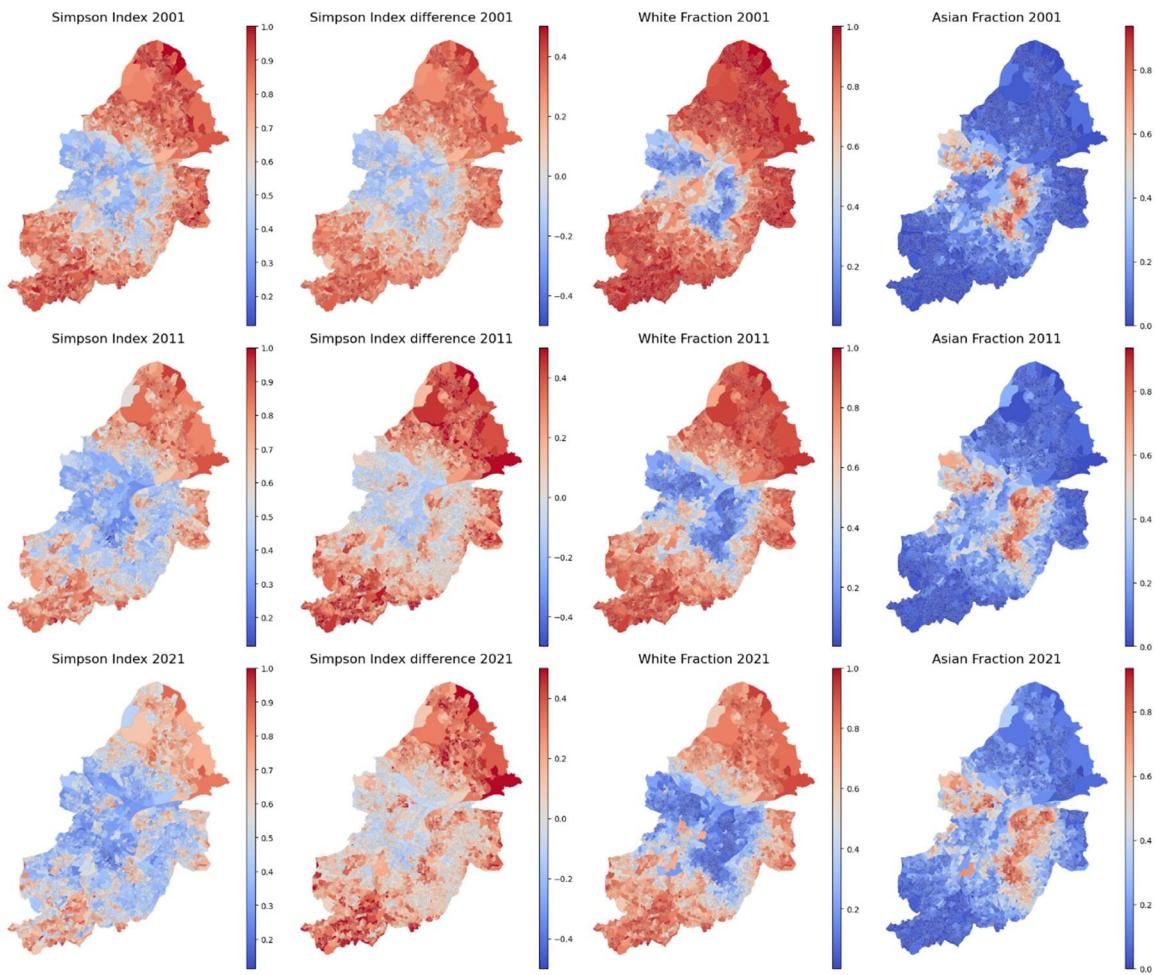


Figure 17. OA Simpson Index, White and Asian Fraction on Map - Birmingham OAs

8.1.2 Birmingham Map – MSOA Level

Figure 18 shows the Simpson Index for each MSOA and its difference from the town-wide Simpson Index in Birmingham. The same patterns observed in OAs and LSOAs can be seen in MSOAs. This plot includes the White and Asian Moran's I index of each MSOA, both calculated based on the OAs inside them. It is observed that in most MSOAs, especially in central part of the map, the Moran's I index is positive for both White and Asian ethnicities. This means a significant clustering of both White and Asian ethnicities. A positive Moran's I index for both groups indicates that similar values (high percentages or low percentages of the respective ethnicities) are geographically clustered together. This suggests that neighbourhoods or areas within these MSOAs tend to have higher or lower concentrations of

White and Asian populations, rather than these populations being evenly distributed or randomly dispersed.

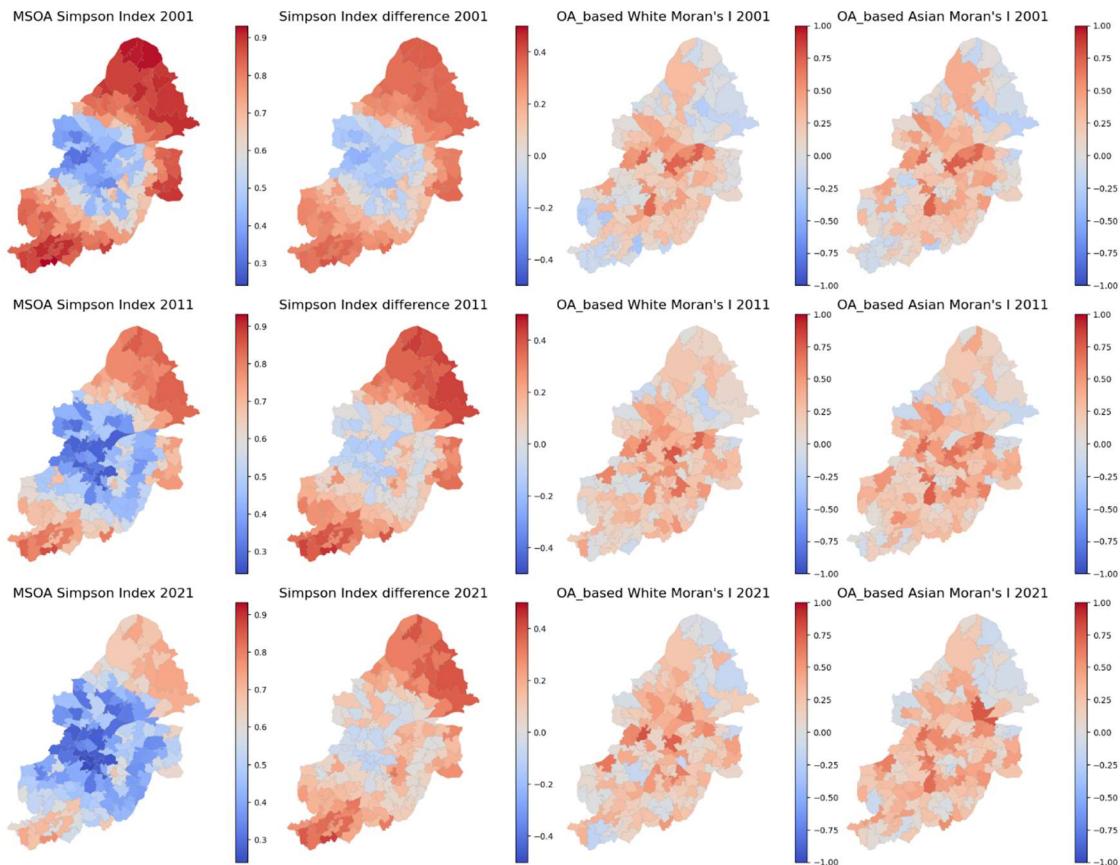


Figure 18. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Birmingham MSOAs

8.2 Leicester

Leicester is one of the most diverse cities in England. As shown in Figure 19, Leicester's total population increased from 279,832 in 2001 to 368,531 in 2021. During this period, the population of the white ethnic group decreased from 178,693 to 150,676. In contrast, the population of the Asian ethnic group increased and surpassed the white ethnic group population. The growth rate of the Asian population is considerably higher than that of the black and other ethnic groups. As depicted in Figure 20, the fraction of the white ethnic group decreased from 63.9% in 2001 to 40.9% in 2021. In 2021, Asian fraction surpassed White fraction, reaching 43.4%.

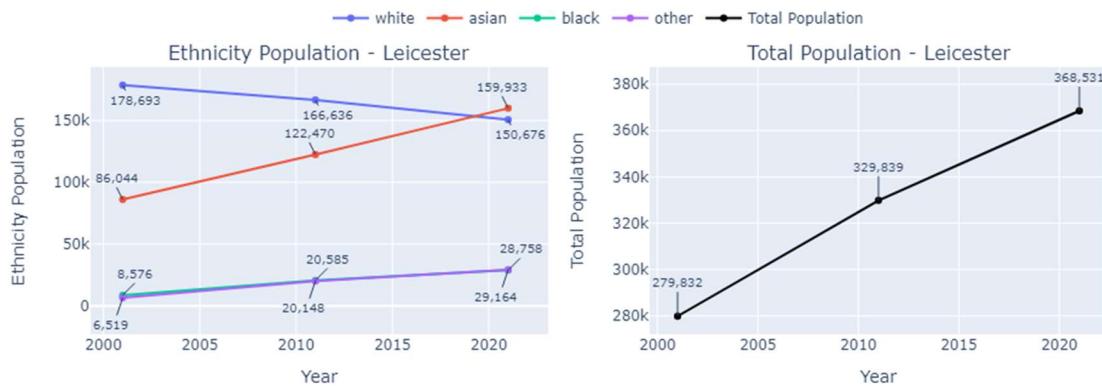


Figure 19. Ethnicity and Total Population Line Chart - Leicester

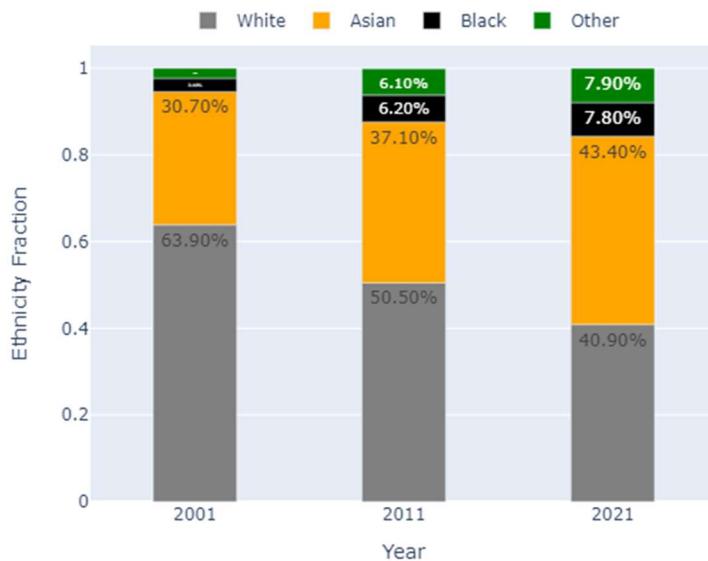


Figure 20. Ethnic Group Fractions Over Time - Leicester

8.2.1 Leicester Map – OA Level

Figure 21 shows the Simpson Index for each OA and its difference from the town-wide Simpson Index in Leicester. In the first column, which depicts the Simpson Index of OAs, the overall Simpson Index is decreasing (more OAs are becoming blue). However, there are some areas that had a lower Simpson Index in 2001 but higher in 2011 and 2021. By focusing on the second column of the plot, this becomes more apparent. In these areas, the difference between their Simpson Index and the town-wide Simpson Index increases. By comparing it with the fraction plots, it is noticeable that in areas with each of White and Asian as predominant ethnicity, the difference between the OA Simpson Index and the town-wide Simpson Index is significant.

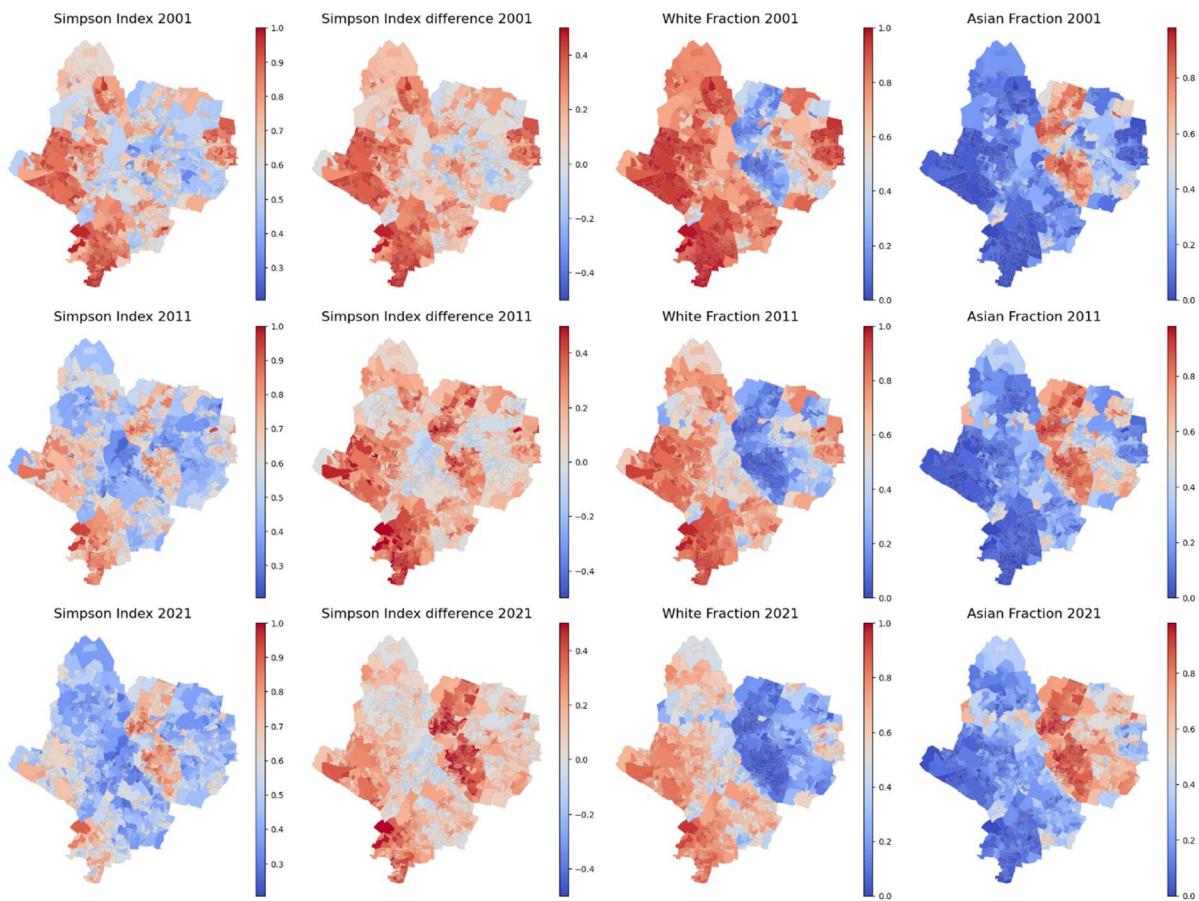


Figure 21. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map - Leicester OAs

8.2.2 Leicester Map – MSOA Level

Figure 22 shows the Simpson Index for each MSOA and its difference from the town-wide Simpson Index in Leicester. The same patterns observed in OAs and LSOAs are evident in MSOAs. This plot also includes the Asian and White Moran's I index for each MSOA, both calculated based on the OAs within them. By comparing the OA-based Moran's I index and the Asian/White fractions in Figure 21, it is apparent that areas with a higher fraction of Asians or a lower fraction of White ethnicity have a positive and relatively high Moran's I index. This suggests that neighborhoods or areas within these MSOAs tend to have higher or lower concentrations of White and Asian populations, indicating significant clustering of these ethnicities.

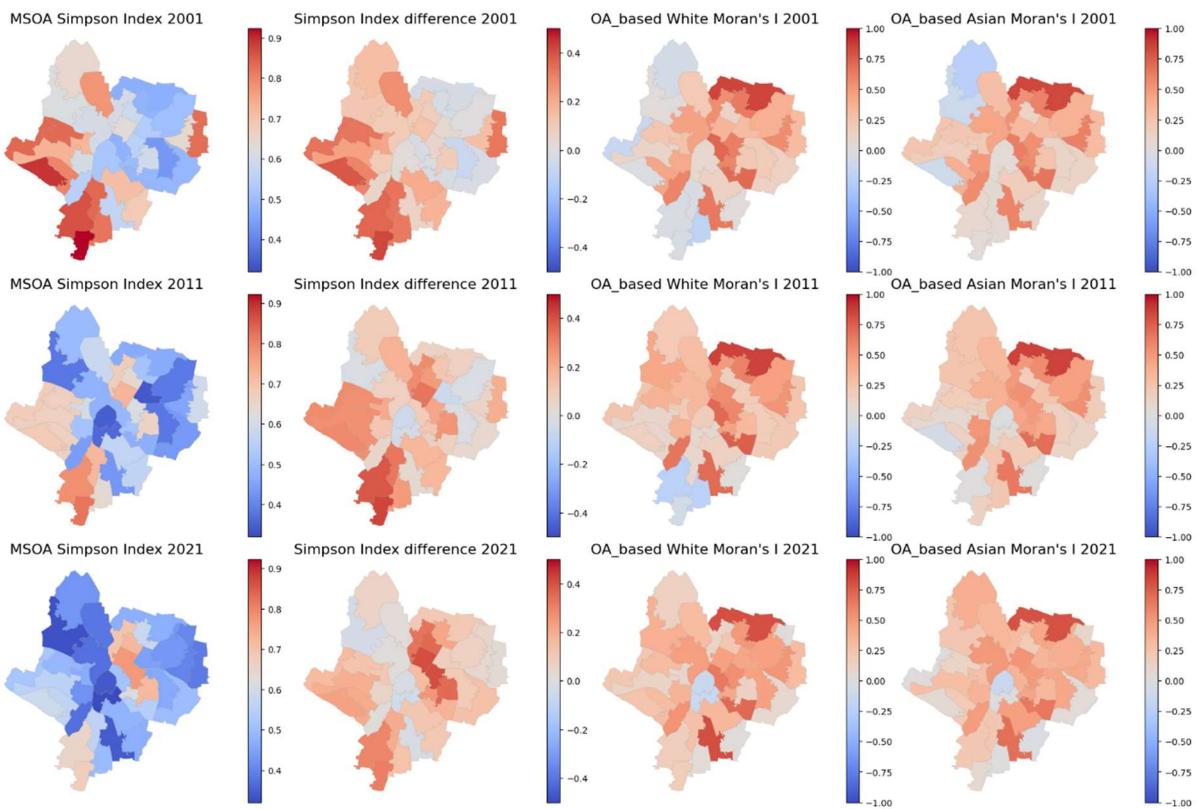
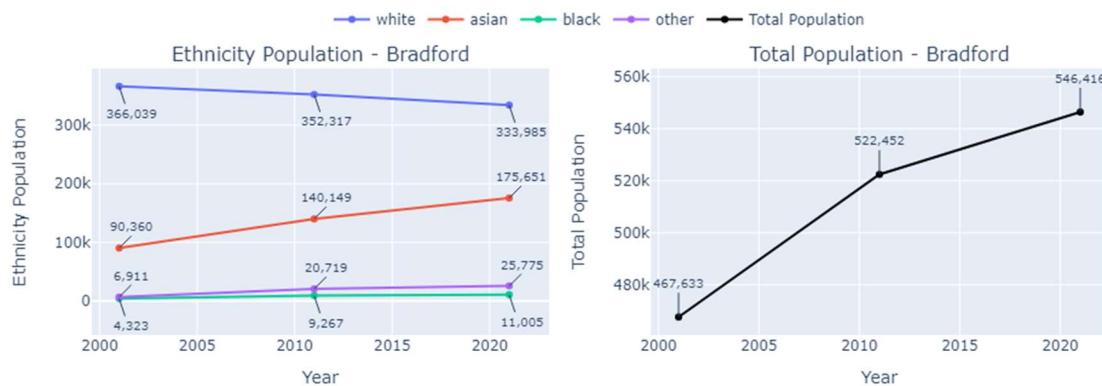
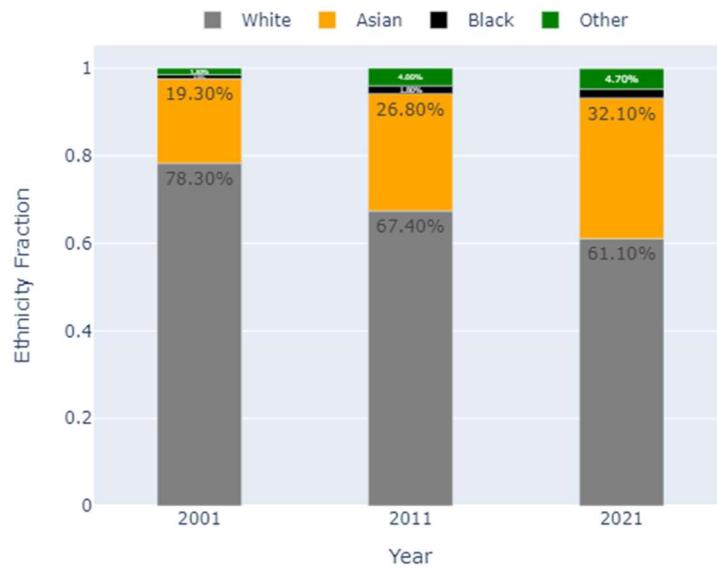


Figure 22. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Leicester MSOAs

8.3 Bradford

Bradford is the second city with largest population of Pakistanis (categorized in Asian ethnicity in this project) in England. As shown in **Error! Reference source not found.**, Bradford's total population increased from 467,633 in 2001 to 546,416 in 2021. During this period, the population of the white ethnic group decreased from 366,039 to 333,985. This decrease is not noticeable compared to the increase of Asian Ethnic group, which increased from 90,360 in 2001 to 175,651 in 2021. In Bradford like Leicester, the growth rate of the Asian population is higher than that of the black and other ethnic groups. As depicted in Figure 24, the fraction of the white ethnic group decreased from 78.30% in 2001 to 40.10% in 2021 and the Asian fraction increases from 19.30% in 2001 to 32.10% in 2021.

**Figure 23. Ethnicity and Total Population Line Chart - Bradford****Figure 24. Ethnic Group Fractions Over Time - Bradford**

8.3.1 Bradford Map – OA Level

Figure 25 shows the Simpson Index for each OA and its difference from the town-wide Simpson Index in Bradford. In the first column, which depicts the Simpson Index of OAs, the overall Simpson Index is decreasing, particularly around two points one in the center and another in the southeast of the city. As observed from the White and Asian fractions map, these points consist of less White fraction and more Asian fraction. By focusing on the second column of the plot, it is evident that the difference between the OA Simpson Index and the

town-wide Simpson Index is increasing in Bradford. However, in areas where the OA Simpson Index was relatively low, this difference is negative or close to zero.

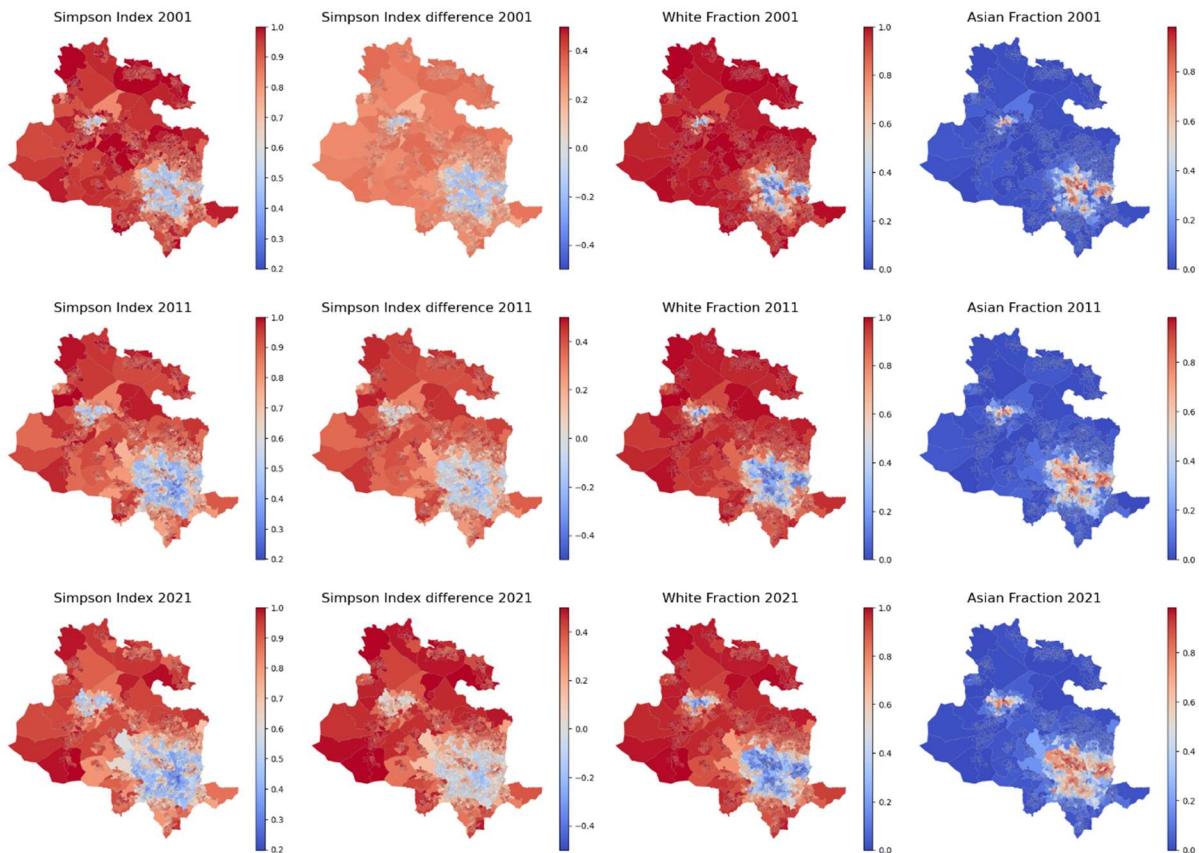


Figure 25. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map - Bradford OAs

8.3.2 Bradford Map – MSOA Level

Figure 26 shows the Simpson Index for each MSOA and its difference from the town-wide Simpson Index in Bradford. The same patterns observed in OAs and LSOAs can be seen in MSOAs. Regarding the white and Asian Moran's I index of each MSOA, calculated both based on the OAs within them, the two areas with more Asian fraction and less white fraction have positive indices, indicating significant clustering of these ethnicities.

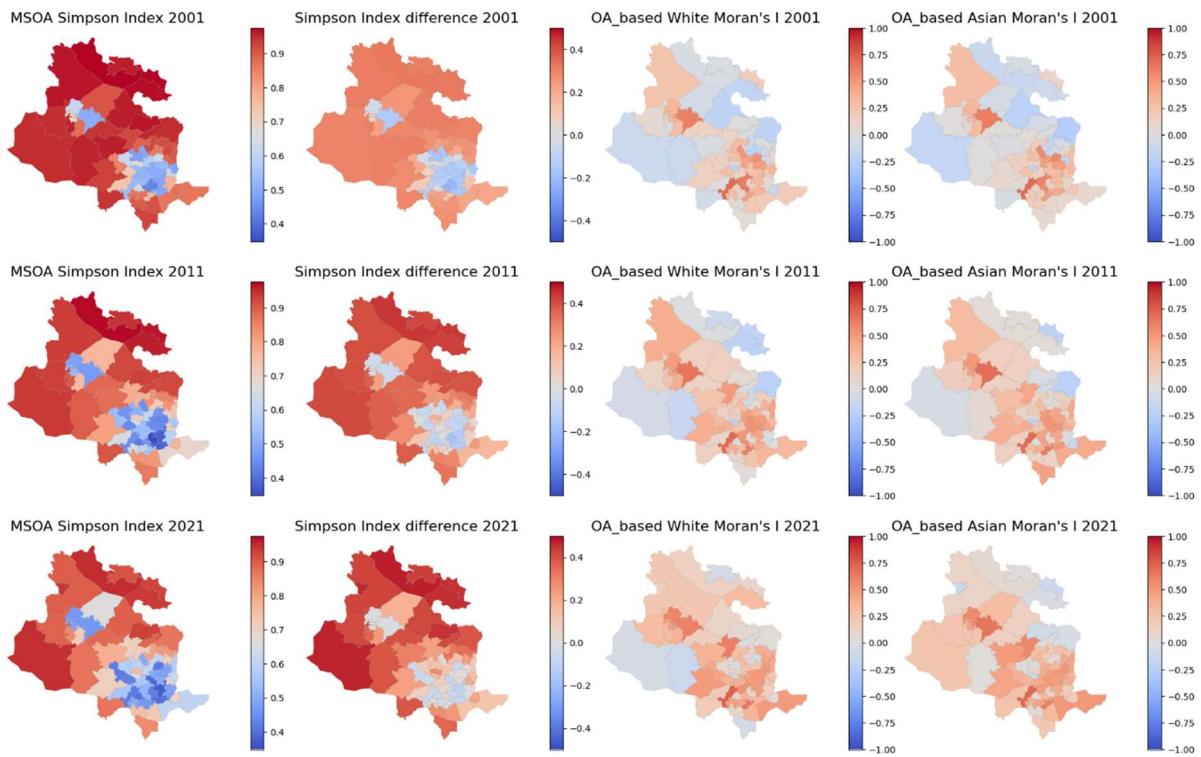
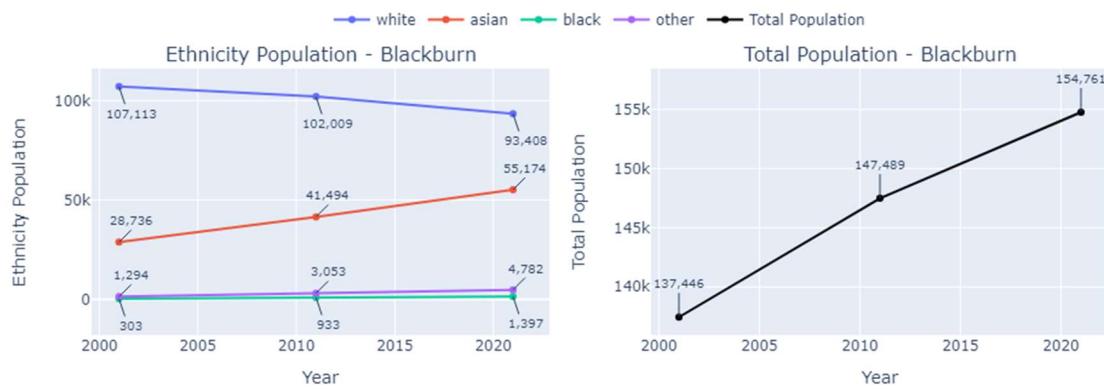
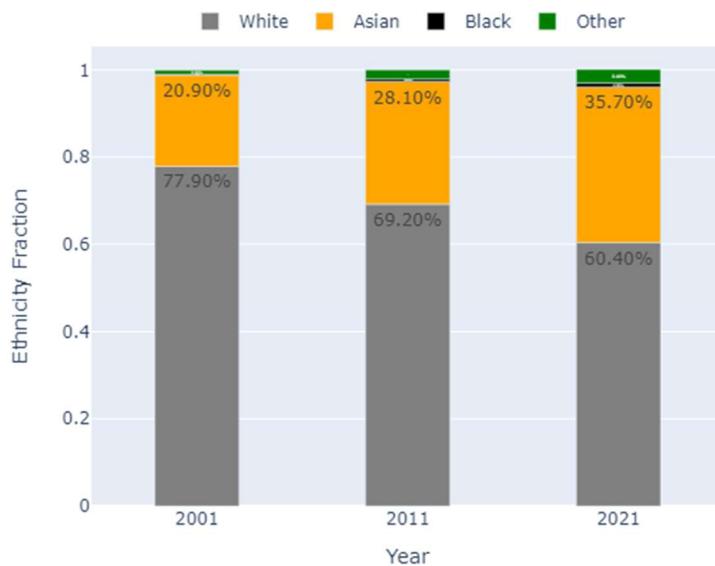


Figure 26. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Bradford MSOAs

8.4 Blackburn with Darwen

Blackburn is an industrial town in the west of England. As shown in Figure 27, Blackburn's total population increased from 137,446 in 2011 to 154,761 in 2021. During this period, the population of the white ethnic group decreased 107,113 in 2001 to 93,408 while the Asian Ethnic group population increased from 28,736 in 2001 to 55,174 in 2021. In Blackburn like Leicester, the growth rate of the Asian population is higher than that of the black and other ethnic groups. As depicted in Figure 28, in 2021, 60.4% of people in Blackburn with Darwen identified their ethnic group within the White category (compared with 69.2% in 2011), while 35.7% identified their ethnic group within the Asian category (compared with 28.1% the previous decade).

**Figure 27. Ethnicity and Total Population Line Chart – Blackburn****Figure 28.Ethnic Group Fractions Over Time – Blackburn**

8.4.1 Blackburn with Darwen Map – OA Level

Figure 29 shows the Simpson Index for each LSOA in Blackburn and its difference from the town-wide Simpson Index. The same pattern observed in OAs can be seen in LSOAs. While the changes in the LSOA Simpson Index are slightly tangible, the changes in their differences from the town-wide Simpson Index are more evident. This can be inferred to mean that the rate of decrease in the town-wide Simpson Index is greater than that of the LSOAs. Again, there are no significant changes in the fractions of White and Asian ethnicities in LSOAs. However, regarding the Moran's I index on map, their spatial distribution is changing within the OAs of each LSOA, as indicated by meaningful changes in the Moran's I index for some LSOAs.

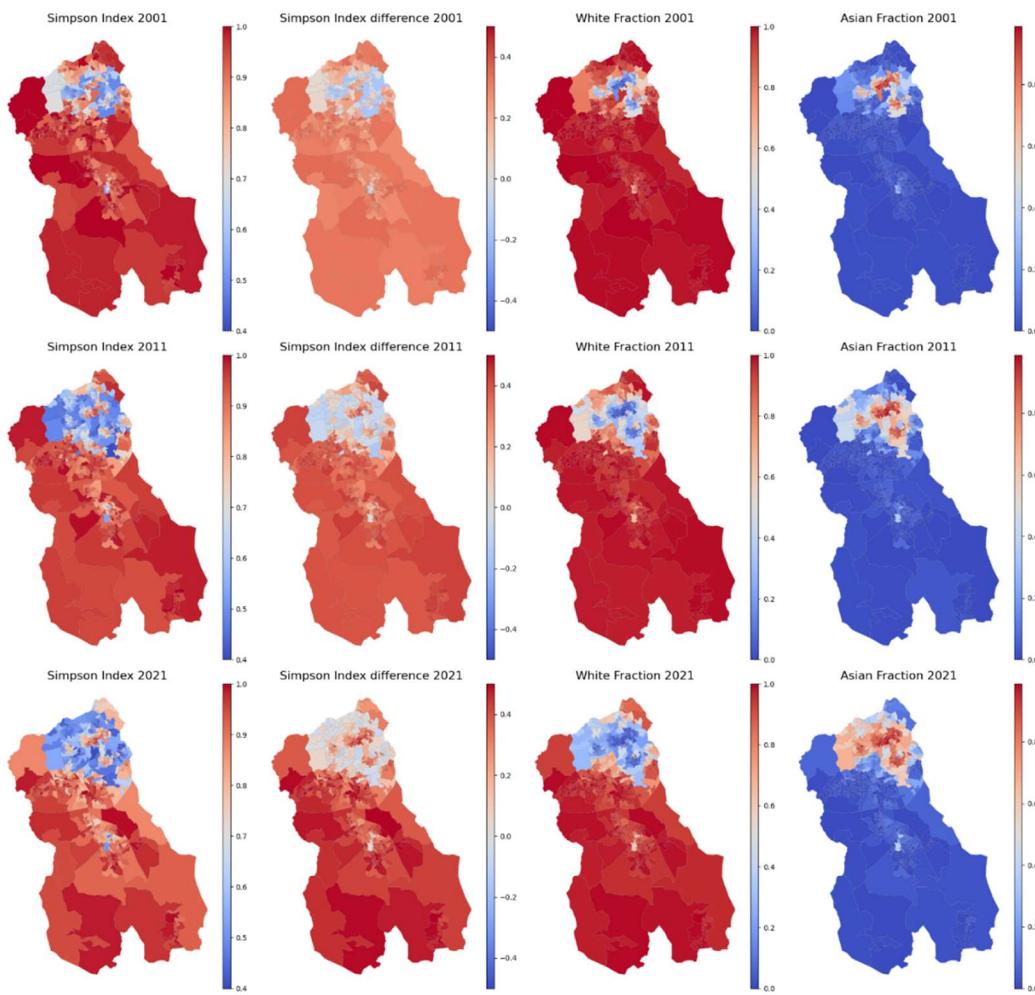


Figure 29. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map – Blackburn OAs

8.4.2 Blackburn with Darwen Map – MSOA Level

Figure 30 shows the Simpson Index for each MSOA and its difference from the town-wide Simpson Index in Blackburn with Darwen. At this level, changes in the MSOA Simpson Index are more evident. By examining the OA-based White and Asian Moran's I index on the map, a positive value is noticeable in the northern part of the city, indicating a significant clustering of the Asian population in that area.

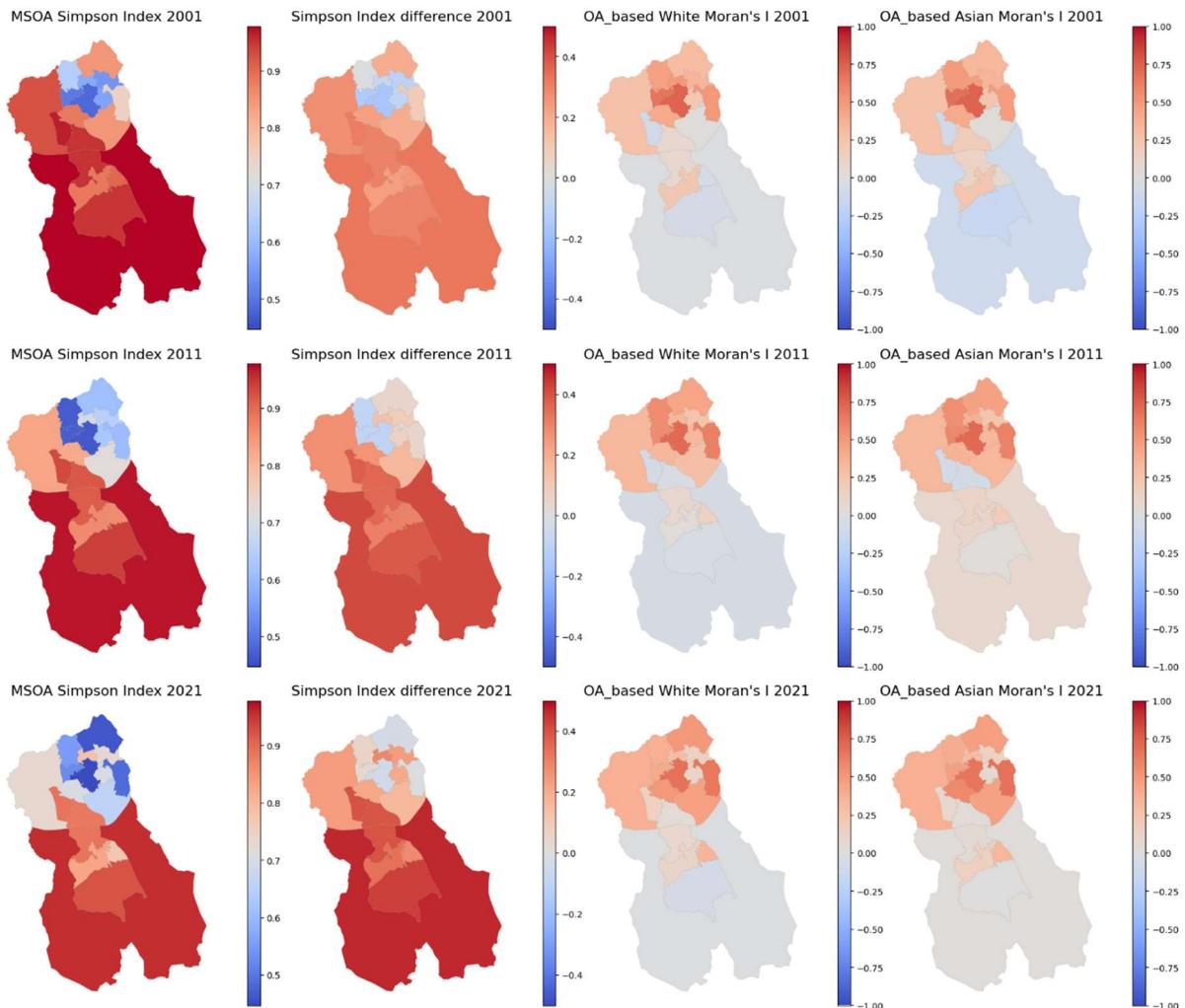


Figure 30. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Blackburn MSOAs

8.5 Oldham

Oldham is a town in Greater Manchester, located in the west of England. As shown in Figure 31, Oldham's total population increased from 217,266 in 2001 to 242,256 in 2021. During this period, the population of the White ethnic group decreased from 187,187 to 164,900, while the population of the Asian ethnic group more than doubled, increasing from 26,406 in 2001 to 59,558 in 2021. In Oldham, similar to other cities, the growth rate of the Asian population is higher than that of the Black and other ethnic groups. As depicted in Figure 32, the fraction of the White ethnic group decreased from 86.20% in 2001 to 68.1% in 2021, while the Asian fraction increased from 12.2% in 2001 to 24.60% in 2021.

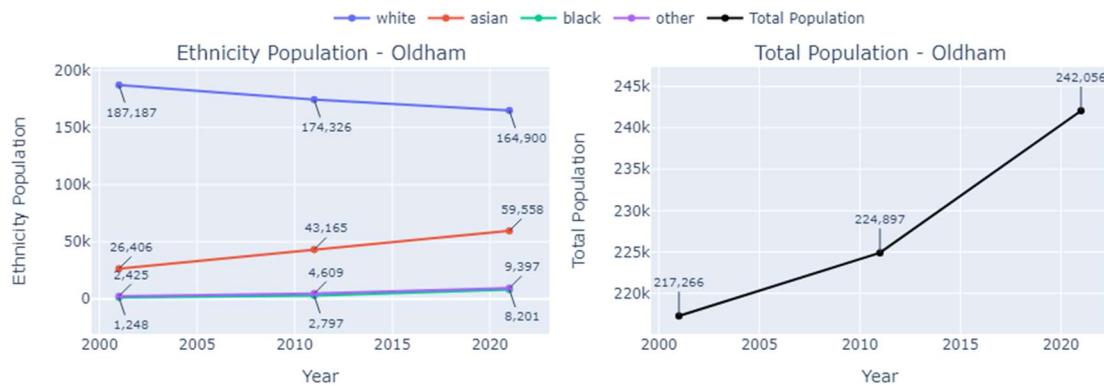


Figure 31. Ethnicity and Total Population Line Chart - Oldham

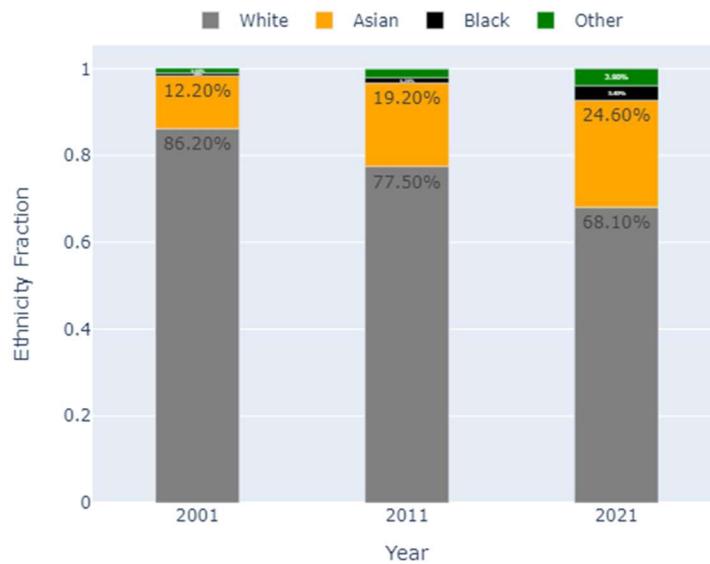


Figure 32. Ethnic Group Fractions Over Time - Oldham

8.5.1 Oldham Map – OA Level

Figure 33 shows the Simpson Index for each OA and its difference from the town-wide Simpson Index in Oldham. In the first column, the Simpson Index of the western OAs has the lowest values, and the overall Simpson Index is decreasing. On the other hand, the Simpson Index difference plot shows an increase, indicating that the town-wide Simpson Index is decreasing at a higher rate compared to the OA Simpson Indexes. As observed from the White and Asian fractions map, the OAs in the southwest areas have a lower White fraction and a higher Asian fraction.

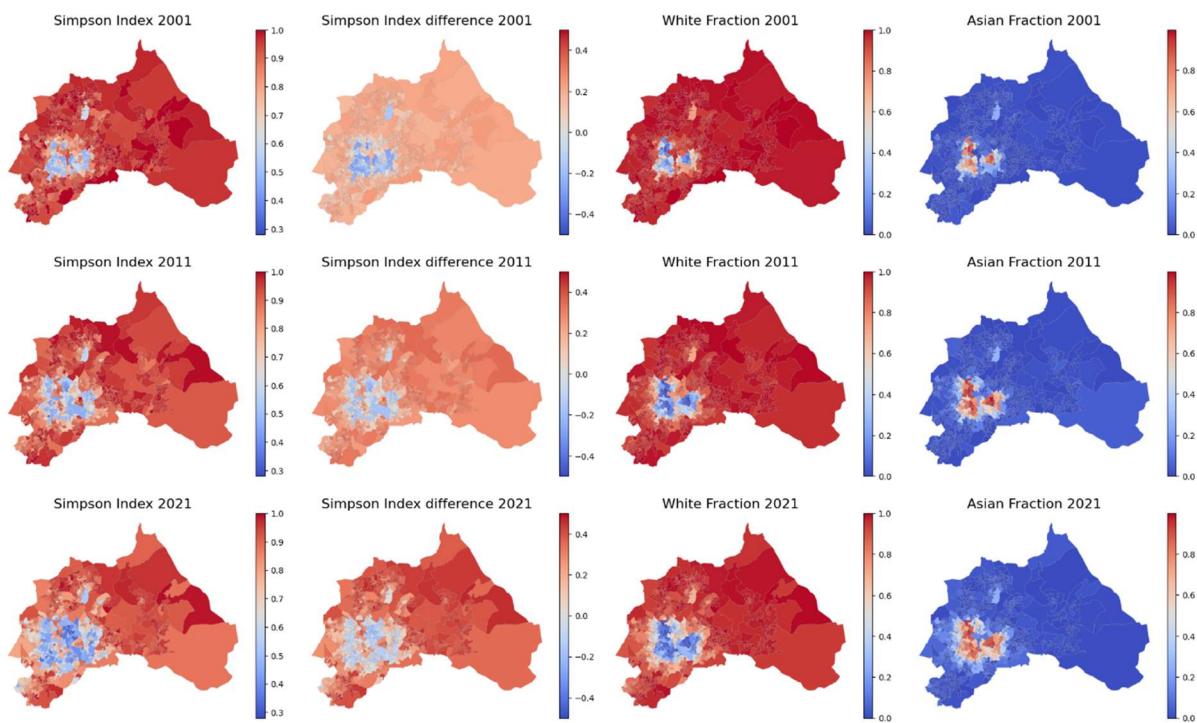


Figure 33. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map – Oldham OAs

8.5.2 Oldham Map – MSOA Level

Figure 34 shows the Simpson Index for each MSOA and its difference from the town-wide Simpson Index in Blackburn with Darwen. At this level, changes in the MSOA Simpson Index are more evident, especially in the western part of the city. By examining the OA-based Asian Moran's I index on the map, a positive value is noticeable in the western part of the city in 2001 and 2011, and in the western and southern parts in 2021, indicating a clustering of the Asian population in those areas. However, the LSOA-based Asian Moran's I index is mostly negative, with some positive values in western MSOAs.

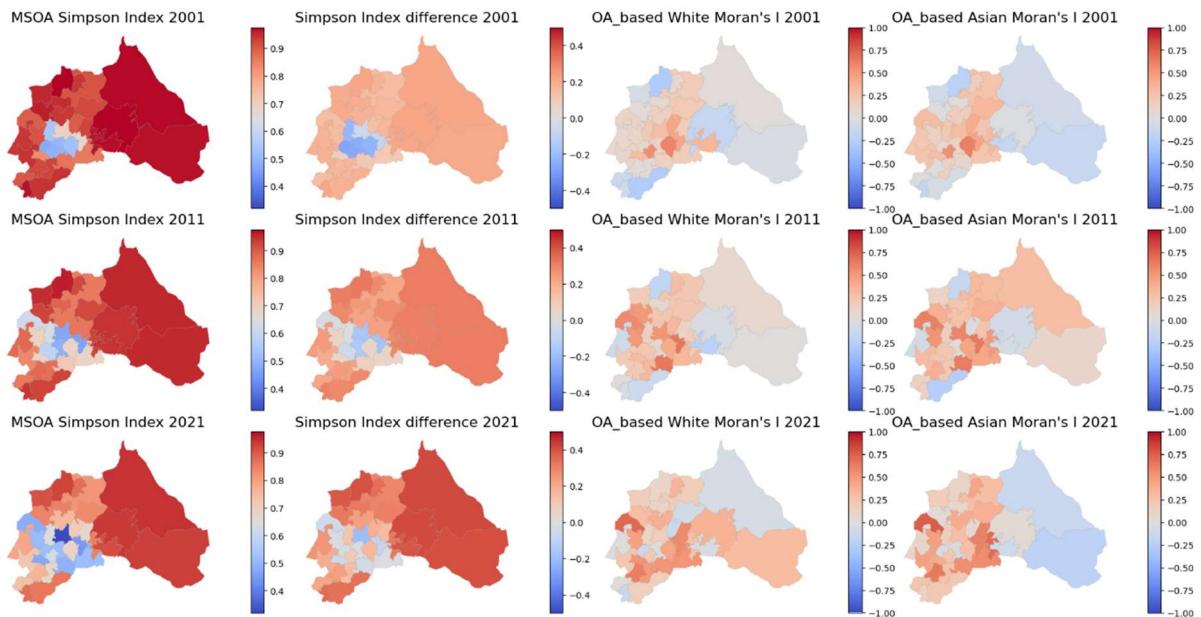


Figure 34. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Oldham MSOAs

8.6 Pendle

Pendle is a local government district in Northwest England. As shown in Figure 35, between the last two censuses (held in 2011 and 2021), the population of Pendle increased by 7.0%, from around 89,321 in 2011 to around 95,804 in 2021. During this period, the population of the white ethnic group decreased from 75,814 to 67,590 while the population of Asian ethnic group doubled from 12,765 in 2001 to 25,600 in 2021. As depicted in Figure 36 the fraction of the white ethnic group decreased from 84.90% in 2001 to 70.60% in 2021 and the Asian fraction increased from 14.30% in 2001 to 26.70% in 2021.

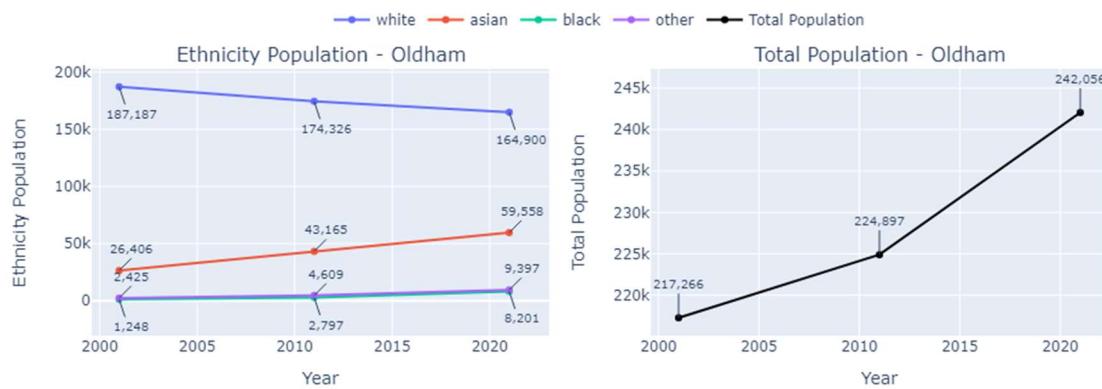


Figure 35. Ethnicity and Total Population Line Chart – Pendle

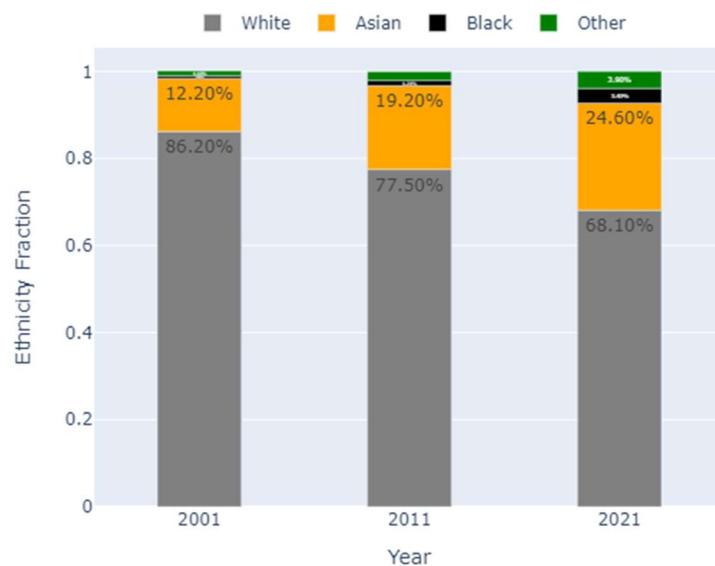


Figure 36. Ethnic Group Fractions Over Time - Pendle

8.6.1 Pendle Map – OA Level

Figure 37 shows the Simpson Index for each OA and its difference from the town-wide Simpson Index in Pendle. In the first column, OAs Simpson Index is decreasing in northern and southern areas, while increasing in some central areas expanded from east to west. However, the Simpson Index difference plot shows an increase in all areas. As observed from the White and Asian fractions map, the OAs in the southern areas have a lower White fraction and a higher Asian fraction.

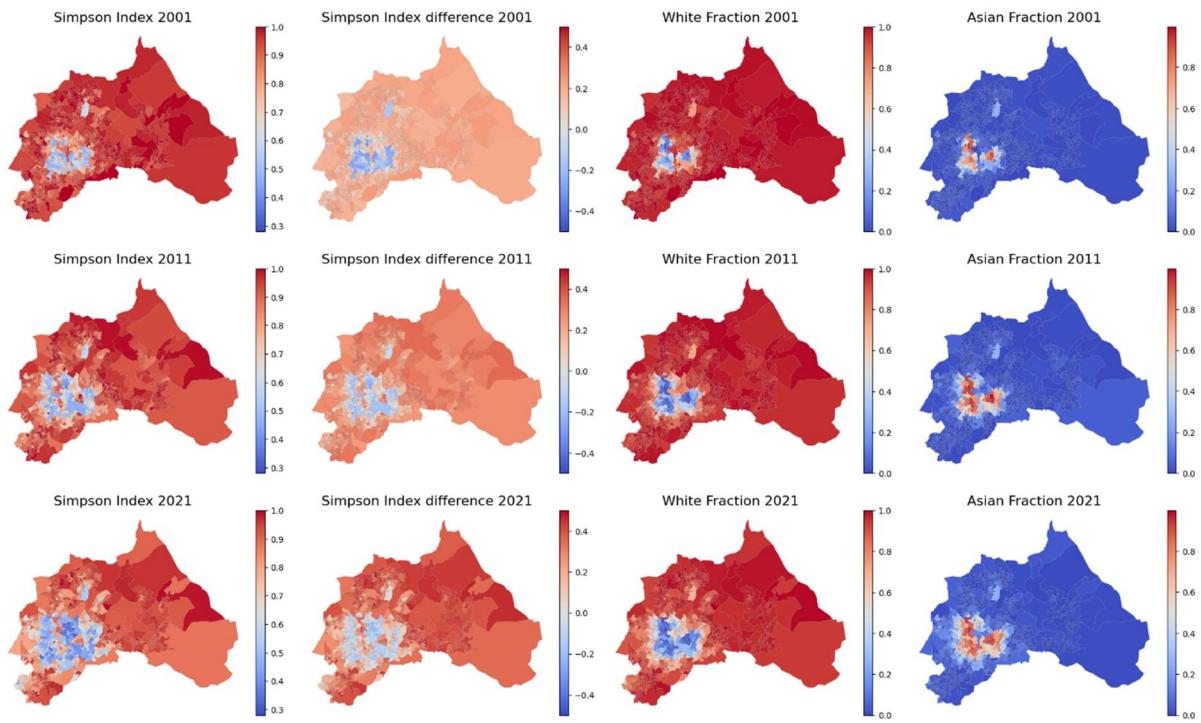


Figure 37. OA Simpson Index, OA-LAD Simpson difference, White and Asian Fraction on Map - Pendle OAs

8.6.2 Pendle Map – MSOA Level

Figure 38 shows the Simpson Index for each MSOA and its difference from the town-wide Simpson Index in Pendle. In the first column of the plot, it can be observed that the Simpson Index is decreasing in the western MSOAs, while the decrease in the eastern parts is less noticeable. However, based on the second column, an increase in the difference between the MSOA Simpson Index and the town-wide Simpson Index is noticeable in these areas. On the map based on the OA-based Moran's I index, positive values are observed in the eastern part of the city in 2001 and 2011, as well as in the southeastern MSOA, indicating a significant clustering of the Asian population in those areas.

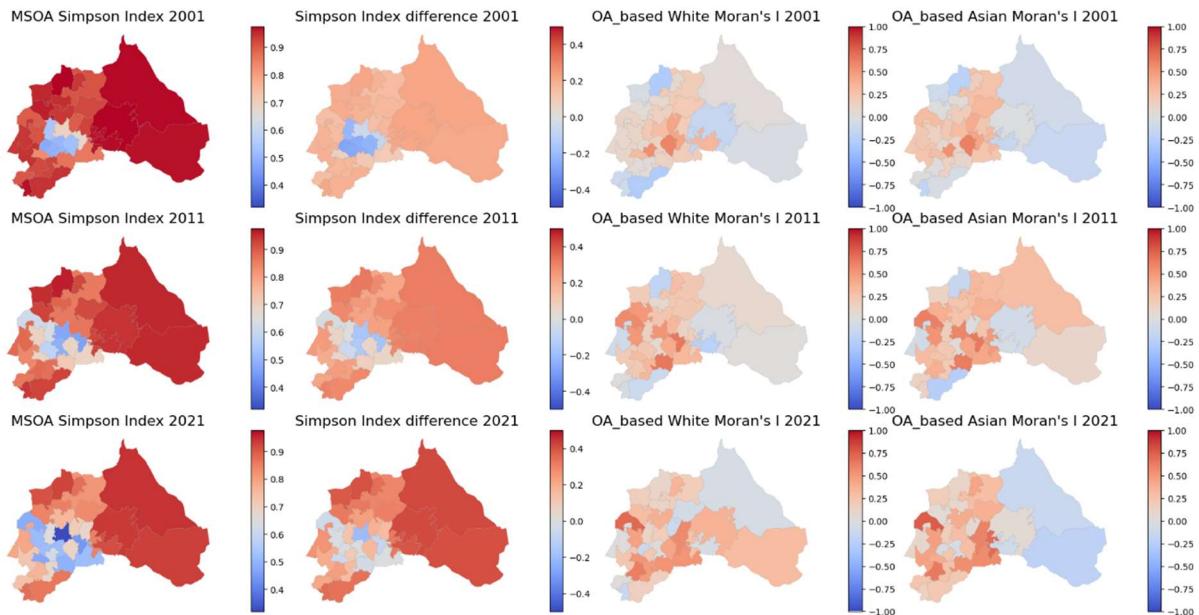
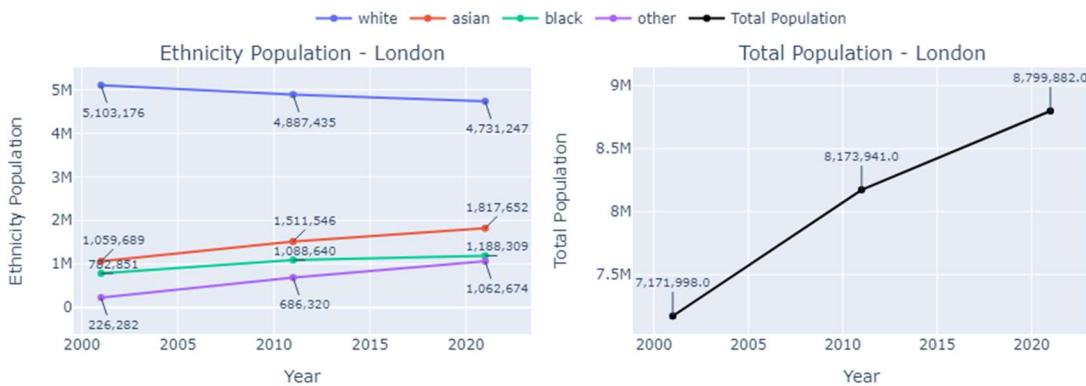
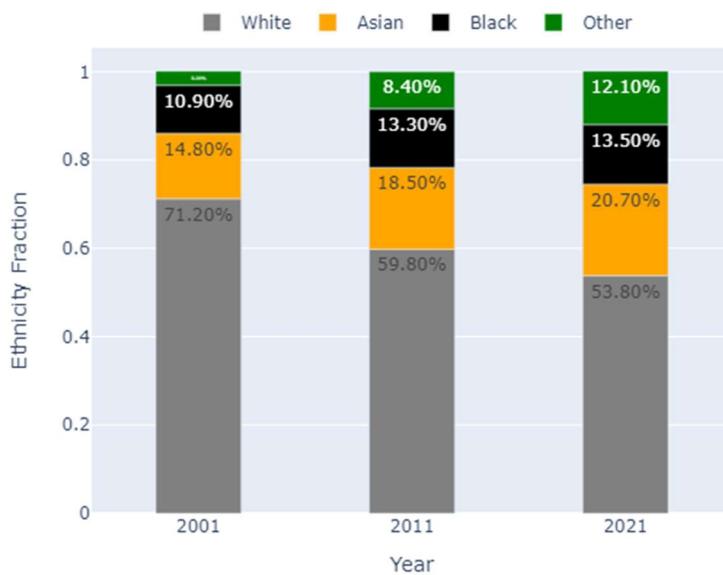


Figure 38. MSOA Simpson Index, MSOA-LAD Simpson difference, Asian and White Moran's I on Map - Pendle MSOAs

8.7 London

London, the capital of England and the United Kingdom, has become one of the most ethnically diverse and multicultural cities in the world. London is significantly more diverse than England as a whole. London consists of 33 LADs.

As shown in Figure 39, between the last two censuses (held in 2011 and 2021), the population of Pendle increased by 7.0%, from 7,177,998 in 2001 to 8,799,882 in 2021. During this period, the population of the white ethnic group decreased from 5,103,176 to 4,731,247 while the population of Asian ethnic group increased from 1,059,689 in 2001 to 1,817,652 in 2021. London has the highest population pf black ethnicity (13.5% of its population) in the whole England. Figure 40, 46.2% of Londoners are Minority Ethnic, compared to 18.3% of England as a whole. the fraction of the white ethnic group decreased from 84.90% in 2001 to 70.60% in 2021 and the Asian fraction increased from 14.30% in 2001 to 26.70% in 2021.

**Figure 39. Ethnicity and Total Population Line Chart - London****Figure 40. Ethnic Group Fractions and Total Population in London Over Time**

8.7.1 London Map – OA Level

London consists of 25,053 OAs in 2021. Figure 41 shows the Simpson Index for each OA and its difference from the town-wide Simpson Index in London. In the first column, the Simpson Index for OAs is decreasing, especially in OAs close to the borders of the city. Unlike other studied cities, the Simpson Index and Simpson Index differences show the same trend. As observed from the White and Asian fractions map, except for three areas, the White fraction is rather high in 2001 but decreases in 2011 and 2021. The Asian fraction did not have a drastic change and remains high in the same three areas where the Simpson Index is low.

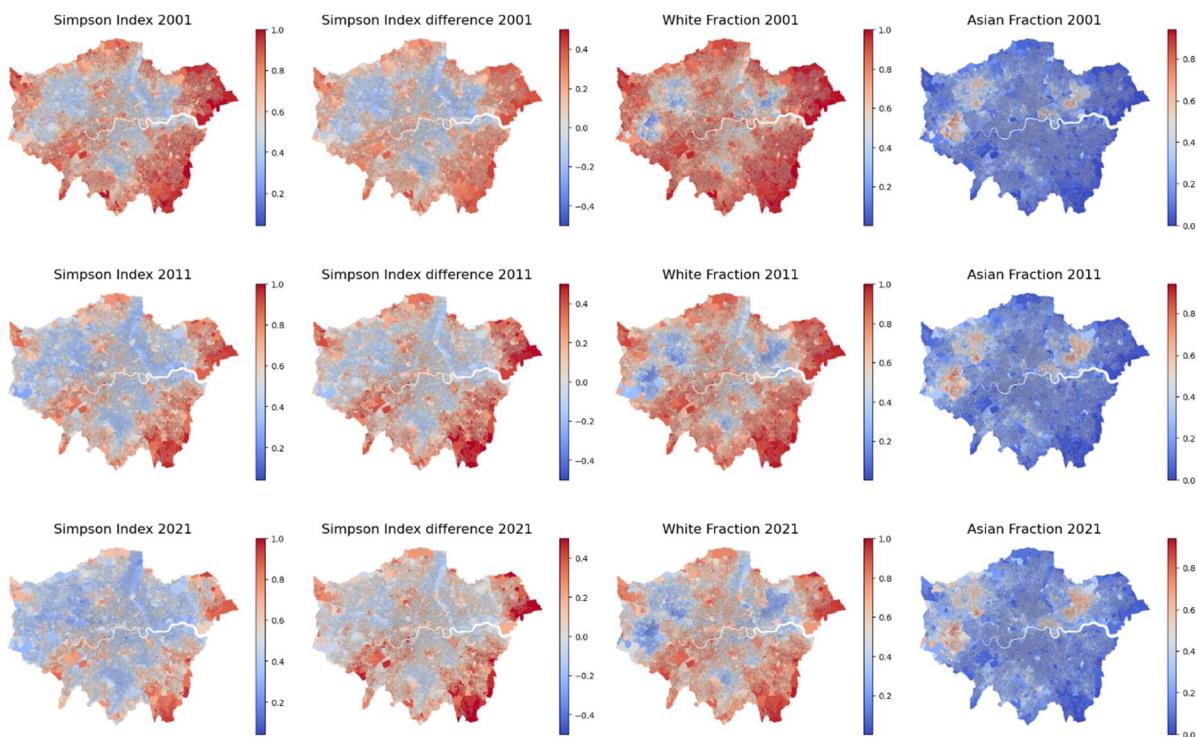


Figure 41. OA Simpson Index, OA-London Simpson difference, White and Asian Fraction on Map - London OAs

8.7.2 London Map – MSOA Level

Figure 42 displays the Simpson Index for each MSOA and its deviation from the city-wide Simpson Index in London. In the first column of the plot, the Simpson Index shows a decreasing trend. However, according to the second column, the disparity between the MSOA Simpson Index and the city-wide Simpson Index approaches from negative values to zero where the Simpson Index is low but appears to increase at some edges.

On the map of the OA-based Asian Moran's I index, positive values predominate, primarily in areas where the Simpson Index is low. Regarding the White Moran's I index, the values appear higher than for Asians, with no discernible pattern observed.

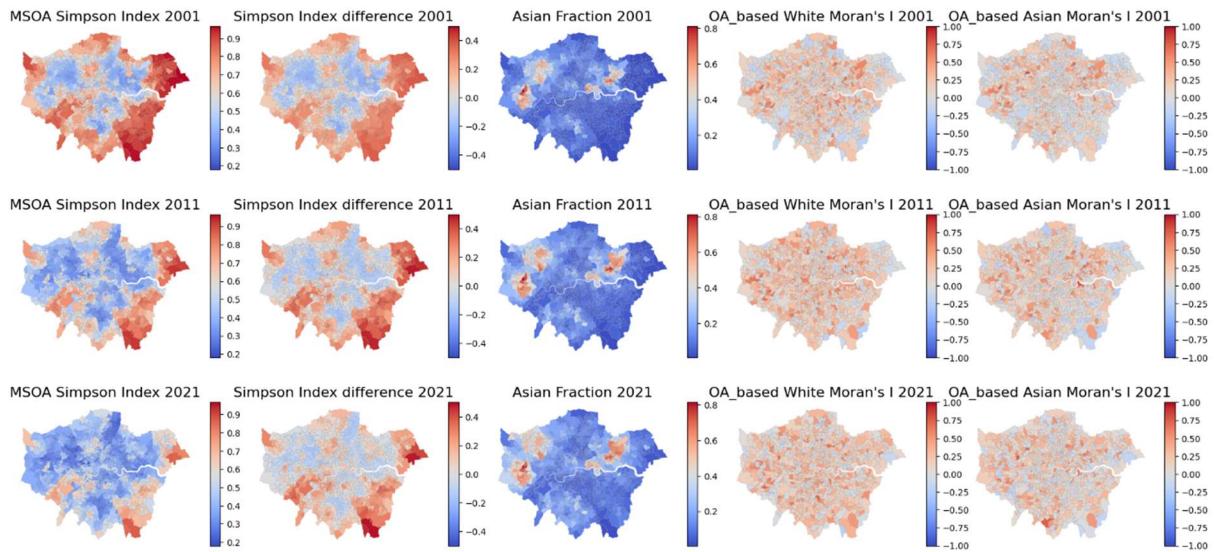


Figure 42. MSOA Simpson Index, MSOA-London Simpson difference, Asian and White Moran's I on Map - London MSOAs

8.7.3 London Map – LAD Level

Figure 43 depicts the 33 LADs of London. In the first column of the plot, it is observed that the Simpson Index in the eastern areas of London was rather high in 2001 but decreased in the subsequent years. The difference between the LADs Simpson Index and the city-wide Simpson Index did not show a significant change, especially from 2011 to 2021. The Asian fraction is considerably increasing in the western parts of the city, as well as in three LADs in the northeast. The White and Asian Moran's I indices did not experience a drastic change; generally, the White Moran's I appears to be higher than the Asian Moran's I.

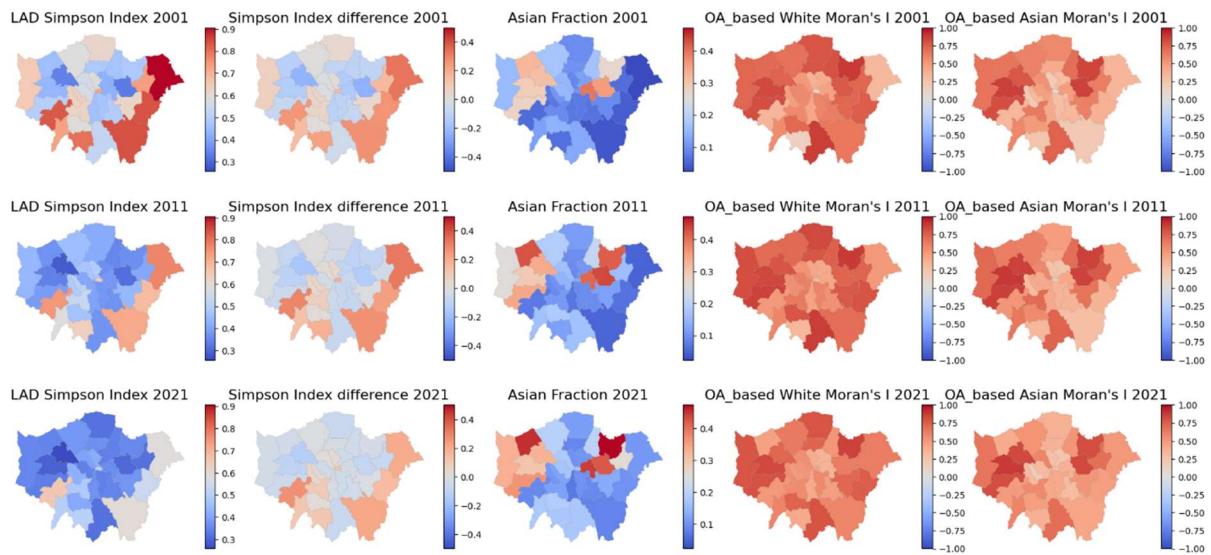


Figure 43. LAD Simpson Index, LAD-London Simpson difference, Asian Fraction and Asian and White Moran's I on Map - London LADs

9 Comparing the cities

In this part of the study, all the aforementioned cities are compared based on various aspects. Table 7 presents the correlation coefficient between the OA-based Simpson index and the White fraction in each city. The table indicates that this correlation is positive and rather strong. However, the correlation of the OA-based Simpson index with other ethnicity fractions is negative across all cities. By 2021, as the White fraction in all the selected cities decreases, this correlation also diminishes.

In 2011, in Leicester, the White ethnicity remains the largest among other ethnicities, but the Asian fraction has increased, reducing the difference between the White and Asian fractions to less than 15%. Consequently, the Simpson Index does not have a strong correlation with the White fraction, though its value is still positive. By 2021, this correlation in Leicester turns negative as the Asian ethnic fraction surpasses the White ethnic fraction, resulting in a positive correlation (0.45) with the Asian fraction.

Table 7. Correlation Coefficient between OA-based Simpson Index and White Fraction

year	Birmingham	Leicester	Bradford	Blackburn	Oldham	Pendle	London
2001	0.75	0.54	0.73	0.7	0.74	0.87	0.88
2011	0.66	0.2	0.65	0.61	0.58	0.81	0.82
2021	0.53	-0.27*	0.6	0.52	0.59	0.78	0.75

* The correlation with Asian Fraction was 0.45

Figure 44 illustrates the changes in various indices over the years for each city, along with a comparison to other selected cities.

In the first row of the figure, it is evident that in all cities, the White fraction is decreasing while the Asian fraction is increasing. Leicester has the lowest White fraction and the highest Asian fraction.

In the second row, it can be observed that London has the lowest dissimilarity index for both the White and Asian fractions, indicating a more even distribution of these ethnicities across OAs. However, Blackburn and Pendle show a slight increase in the White dissimilarity index in 2021, while other cities exhibit a decreasing trend. The Asian dissimilarity index shows an increasing trend in Pendle in 2021.

The third row shows that the Simpson index, based on both OA and LAD levels, exhibits a decreasing trend across all cities.

In the fourth row, which depicts Moran's I index, it is clear that, except for London and Birmingham, the White and Asian Moran's I index values are increasing. Pendle is experiencing the highest rate of increase. This indicates that in these cities, OAs with similar values are becoming more clustered.

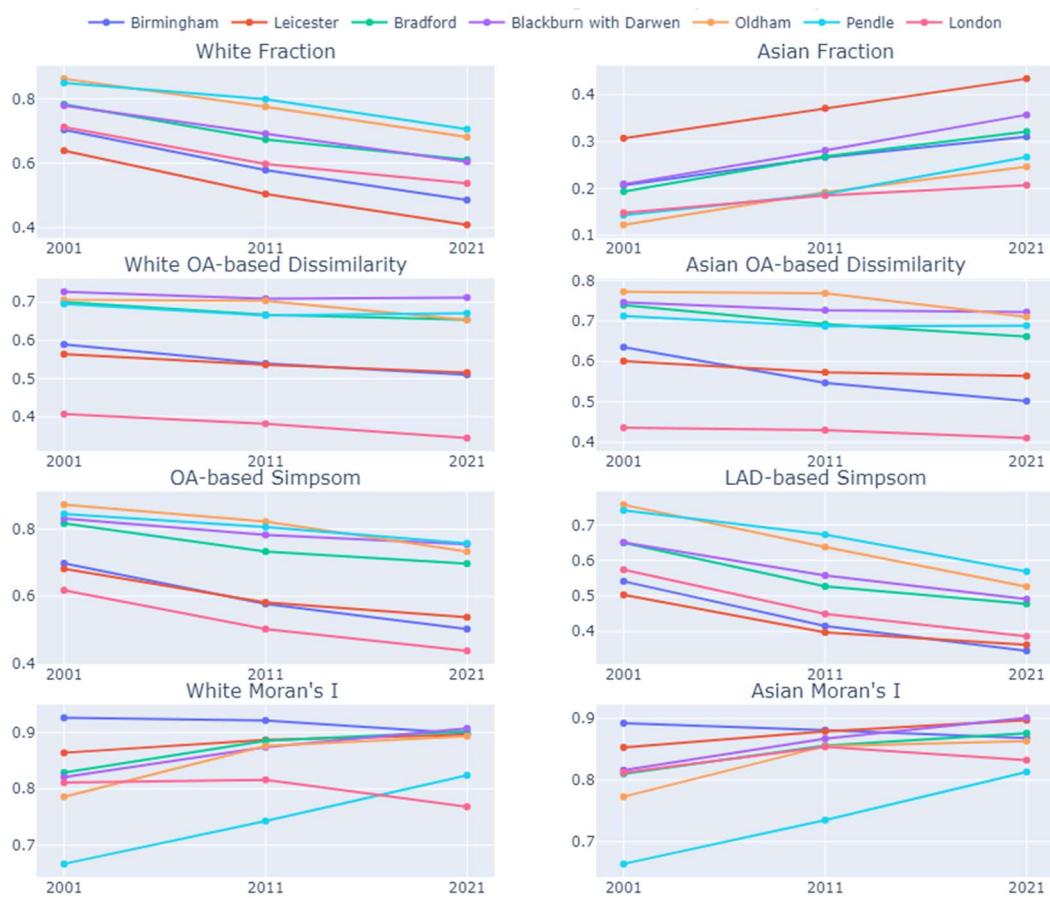


Figure 44. Indices Trends in Selected English Cities (2001-2021)

10 Conclusion

In this study, a comprehensive analysis of segregation indexes was performed at the country level, LAD level, and in selected cities of England. The results show that in Output Areas (OAs), which are small and designed to be relatively homogeneous, the Simpson Index reflects micro-level variations within neighborhoods. Some OAs, predominantly composed of a single ethnic group, have higher Simpson Index values, indicating lower diversity within those OAs. However, aggregating across all OAs balances out these variations, providing a more generalized measure of diversity for the entire urban area.

Individual neighborhoods within a city might show higher Simpson Indices due to localized ethnic concentrations. However, when considering the city as a whole, the town-wide Simpson Index is lower (or equal) because the overall composition includes a mix of diverse neighborhoods. LAD-wise indices, which aggregate data across a larger sample size, offer a broader view compared to OAs, smoothing out local variations and providing a more comprehensive measure of diversity.

A higher Simpson Index indicates lower diversity, suggesting a dominant ethnic group within the area. This dominance means individuals are more likely to be from the same ethnic group when chosen randomly. The maps show that higher White ethnicity fractions correspond to higher Simpson Index values. When the OA Simpson Index is much higher than the town-wide Simpson Index, it indicates strong local concentration. For example, a high OA-based Simpson Index, which is the probability that a randomly selected pair of individuals from the same OA belongs to one ethnic group, shows that the OA is less diverse compared to an OA with a lower Simpson Index. In addition, the results show that moving from 2001 to 2021, Simpson index at country and LAD level is generally decreasing.

Based on the analysis of the dissimilarity index at various geographic levels, the findings indicate a general decrease in the index values for both White and Asian ethnicities over the studied period and by expanding the area of study. This decline in the dissimilarity index suggests increasing spatial integration and a reduction in ethnic segregation within these communities. These findings reflect positive strides toward achieving greater ethnic cohesion and reducing segregation in urban settings.

The different spatial scales of analysis—OAs, LSOAs, and LADs—affect the interpretation of diversity and spatial patterns. OAs are smaller and more homogeneous, leading to higher Moran's I values, indicating stronger spatial clustering. LSOAs are larger and encompass a greater diversity of neighborhoods, which can dilute spatial autocorrelation and lead to lower Moran's I values. This trend is especially observed when comparing data from 2011 to 2021.

The spatial scale effect is evident: at the local scale (OAs), Moran's I captures fine-scale variations and local clusters within neighborhoods. At the aggregated scale (LSOAs), Moran's I smooths out local variations, offering a broader view that may reduce spatial autocorrelation. This highlights the importance of considering the scale of analysis when interpreting spatial patterns and autocorrelation.

The overall trends from 2001 to 2021 indicate a positive movement towards greater ethnic cohesion and reduced segregation in England. The decrease in both the dissimilarity and Simpson indices at the country and LAD levels points to increasing diversity and integration within urban areas. However, the increase in Moran's I at the OA level within LADs suggests that while broader integration is occurring, localized clustering of ethnic groups persists. By understanding both macro and micro-level trends, more effective and targeted interventions can be developed to enhance social cohesion and integration in England's urban landscapes.

11 Table of references

- Census - Office for National Statistics (no date). Available at: <https://www.ons.gov.uk/census>.
- 'Districts of England' (2024) Wikipedia. Available at: https://en.wikipedia.org/w/index.php?title=Districts_of_England&oldid=1215596947.
- 'Index of dissimilarity' (2023) Wikipedia. Available at: https://en.wikipedia.org/w/index.php?title=Index_of_dissimilarity&oldid=1143568055.
- Jones, K. et al. (2015) 'Ethnic Residential Segregation: A Multilevel, Multigroup, Multiscale Approach Exemplified by London in 2011', *Demography*, 52(6), pp. 1995–2019. Available at: <https://doi.org/10.1007/s13524-015-0430-1>.
- Open Geography Portal* (no date). Available at: <https://geoportal.statistics.gov.uk/>.
- Patias, N., Rowe, F. and Arribas-Bel, D. (2023) 'Local urban attributes defining ethnically segregated areas across English cities: A multilevel approach', *Cities*, 132, p. 103967. Available at: <https://doi.org/10.1016/j.cities.2022.103967>.
- Piekut, A. (2021) 'Re-Theorising Spatial Segregation: A European Perspective', in G. Pryce et al. (eds) *Urban Inequality and Segregation in Europe and China*. Cham: Springer International Publishing (The Urban Book Series), pp. 13–38. Available at: https://doi.org/10.1007/978-3-030-74544-8_2.
- Wong, D.W.S. (2003) 'Spatial Decomposition of Segregation Indices: A Framework Toward Measuring Segregation at Multiple Levels', *Geographical Analysis*, 35(3), pp. 179–194. Available at: <https://doi.org/10.1111/j.1538-4632.2003.tb01109.x>.
- Zuccotti, C.V. et al. (2023) 'Exploring the dynamics of neighbourhood ethnic segregation with agent-based modelling: an empirical application to Bradford, UK', *Journal of Ethnic and Migration Studies*, 49(2), pp. 554–575. Available at: <https://doi.org/10.1080/1369183X.2022.2100554>.

12 Appendix

Supplementary Material such as preprocessing codes, the main code and HTML file of this thesis is provided in the following link:

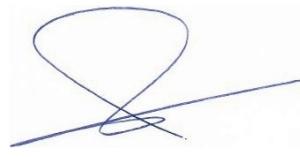
https://github.com/Niillooff/Master_Thesis

Affidavit

The present thesis has, in the same or in similar form, never been handed in to another board of examination.

Bremen 30.06.2024

Place, Date



Signature