British Journal of Environmental Studies
ISSN: 2755-0982
DOI: 10.32996/bjes
Journal Homepage: www.al-kindipublisher.com/index.php/bjes

BJES
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Socio-spatial Disparities of Tobacco Outlets in Edinburgh

Yiqun Wang
University of Edinburgh, Old College, South Bridge, Edinburgh, United Kingdom

Corresponding Author: Yiqun Wang, **E-mail**: s2073600@ed.ac.uk

| ABSTRACT

This paper joins spatial autocorrelation models (Global Moran's I index, Local Moran I index) to analyse the distribution of cigarette vending outlets in Edinburgh and utilises global regression (OLS) models and Geographically Weighted Regression models (GWR) to examine the spatial heterogeneity of fiscal and social factors that affect the distribution of retail outlets in Edinburgh. The results demonstrate that Edinburgh has a considerable spatial agglomeration of cigarette retail outlets, with a clear trend towards clustering in high and low-value areas of retail outlet distribution, showing significant geographical characteristics. The impact of each element on the distribution of cigarette retail outlets was analysed utilising the ranking of the elements in SIMD (The Scottish Index of Multiple Deprivation) as explanatory variables. It was established that data regions with a high ranking in education, skills, and training and those with a high ranking in geographic access to services revealed a positive trend in the density distribution of cigarette retail outlets, while data regions with a high ranking in housing and those with a high ranking in crime revealed a negative trend on the density distribution of cigarette retail outlets. The outcomes are different from prior observations, so this paper specifically analyses the differences in the extent to which each influencing element affects cigarette retail outlets in Edinburgh across data regions, which will supply a scientific basis for decreasing the prevalence of smoking, decreasing socio-economic inequalities in smoking, and increasing diverse local tobacco control intervention strategies.

| KEYWORDS

Tobacco outlets, SIMD, Spatial Autocorrelation, GWR, Edinburgh

| ARTICLE DOI: 10.32996/bjes.2022.2.1.7

1. Introduction

Smoking and its accompanying health problems have long been criticized (Mayers et al., 2012), but outlets contribute significantly to the availability of cigarettes and many other tobacco products, which is a potential causal factor in promoting smoking and is a barrier to cessation, especially among lower socio-economic status groups (Wheeler et al., 2020). The rising emphases on controlling tobacco use and reducing public health disparities have made the location of tobacco outlets critical (Yu et al., 2010).

This paper investigates whether there is variability in the distribution of tobacco outlets in Edinburgh by analysing their spatial distribution patterns combined with key socio-economic factors. It is hypothesized that the distribution varies in communities with different socio-economic characteristics.

This study will use global and local Moran indices to express spatial patterns of cigarette outlets and examine such associations in terms of spatial heterogeneity, i.e., Geographically Weighted Regression (GWR), taking into account the spatial influences present in the distribution of tobacco retailing.

This work found that there is spatial autocorrelation (i.e., aggregation) in the distribution of tobacco outlets and that their distribution is also influenced by different socio-economic factors, which will provide evidence for further interventions to reduce tobacco consumption and smoking rates.

2. Literature Review

Research on the correlation between the distribution of cigarette outlets and socio-economic factors has a long history. However, early studies (prior to 2011) (Campbell et al., 2009; Connor et al., 2011; Romley et al., 2007) put their focus on the analysis of alcohol outlet distribution and its correlates rather than cigarette outlets. Earlier studies of cigarette outlets have focused on the locations such as US cities and other regions (McCarthy et al., 2009; Novak et al., 2006), but the attention to the UK, particularly to Scotland, has only emerged after the enactment of the Tobacco and Primary Medical Services (Scotland) Act 2010.

Besides, the focus of tobacco sales in Scotland has taken a more diverse direction. For example, Ford (Ford et al., 2020) explored the influence of the Tobacco and Primary Medical Services (Scotland) Act on young people before, during, and after its implementation. Shortt et al. (2016) discussed the relationship between the smoking behaviour of young people in Scotland and the density of tobacco outlets around homes and schools. In addition, Shortt et al. (2021) proposed to respond to community income differences by adjusting the retail price of cigarettes. However, few have analysed the relationship between different socio-economic characteristics and the distribution of cigarette outlets.

Furthermore, to date, analysis of data collected in geographical areas in the Scottish region has rarely considered spatial influences (Pearce et al., 2016; Shortt et al., 2015), which may lead to potentially misleading analytical results.

Finally, the area selected for analysing the association between cigarette outlets and related factors in the Scottish region is relatively homogeneous, i.e., dominated by Shortt et al. (2015) and Macdonald et al. (2018). The studies carried out on other cities are not known.

To fill these gaps, this paper, therefore, selects Edinburgh (the capital of Scotland) as the study area and uses the GWR (Yu et al., 2010) to conduct a more comprehensive correlation analysis, with a view to reducing tobacco consumption and smoking rates and promoting social equality.

3. Methodology

3.1 Research Framework

This paper seeks to examine whether there is a spatial pattern of tobacco retail outlets in Edinburgh and to explore their relationship with the socioeconomic characteristics of neighbourhoods.

To this end, this paper first analyses the distribution pattern at a macro level using a global spatial autocorrelation (Moran's I index) approach (Cao & Xu, 2018). The local spatial autocorrelation (Anselin Local Moran's I) method in ArcGIS Pro was used to analyse the spatial aggregation at each observed data zone. The Moran scatter diagram was drawn to clearly represent the correlation between each observation unit and its neighbouring units. Secondly, the socioeconomic characteristics of the community were studied by constructing a density of tobacco outlets, selecting explanatory variables, and analysing the influences at both the global and spatial levels using Ordinary Least Squares (OLS) and GWR to obtain spatial heterogeneity of the influences (Figure 1) (Cao & Xu, 2018).

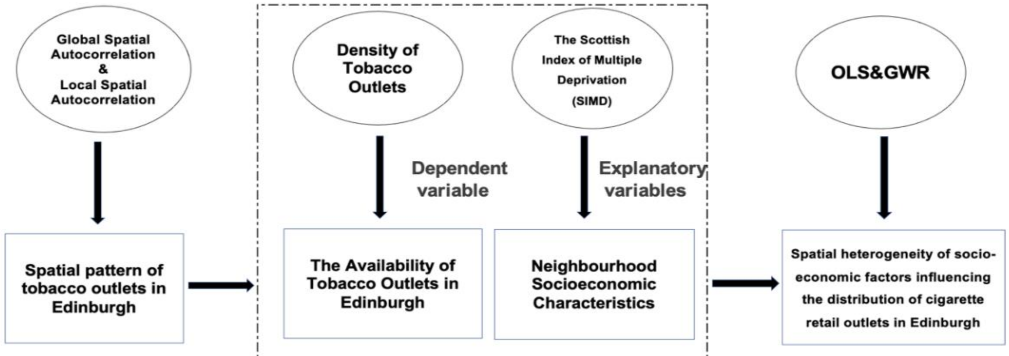


Figure 1: Analytical framework

3.2 Data Sources

There are three main sources of research data.

1. Point data for tobacco outlets in Edinburgh as of 2019

The Health (Tobacco, nicotine etc. And Care) (Scotland) Act 2016 was approved by the Parliament on 3 March 2016 (Tobacco Register Scotland, 2017). The Act states that retailers who sell nicotine vapour products, tobacco, or both are going to be illegal if they're not registered by 1 October 2017 (Tobacco Register Scotland, 2017). The data in this article come from the addresses and postcodes of all premises registered on the Edinburgh Tobacco Retailers Register as of 2019, 894 in total (n = 894) (Shortt et al., 2015).

2. Edinburgh Data Zone

In 2010, in order to keep the existing census data relevant, the boundaries of the 2001 Data Zone were updated, and a new version of the 2011 Data Zone was established to reflect the results of the 2011 census (Scottish Government, 2014). Each of the updated data areas has a unique 9-digit identifier starting with 'S01' followed by 6 digits. 2011 data zones are labelled from S01006506 to S01013482 (6976 separate data zones) (Scottish Government, 2015). This paper separates the data areas contained in Edinburgh to obtain 597 Data Zones as data sources (from S011008417 to S01009013).

3. The Scottish Index of Multiple Deprivation (SIMD)

In line with the research direction of this paper, an indicator of socioeconomic deprivation needs to be collected for each data zone in Edinburgh. The most appropriate indicator is SIMD, as it is the Scottish Government's standard tool for identifying concentrations of deprivation in Scotland (Scottish Government, 2021; Shortt et al., 2015). SIMD combines seven different domains or aspects of deprivation, including Income (28%), Employment (28%), Health (14%), Education, skills and training (14%), Geographic access to services (9%), Crime (5%), Housing (2%) (Scottish Government, 2021). The data in this paper uses SIMD 2020 for the Edinburgh region, the Scottish Government's sixth edition since 2004 (Scottish Government, 2021).

3.3 Variable Selection and Pre-Processing

This paper plans to use the Spatial Join tool in ArcGIS Pro to integrate a point map of tobacco outlets into the data zone map and to calculate the density of cigarette outlets (pcs/km2) in each data zone, and the data obtained will be used as the dependent variable for this study. The independent variables are focused on the seven dimensions of SIMD and are based on their weighting relationships. Finally, seven economic and social factors that could be related to the density of cigarette outlets were ultimately selected.

However, preliminary tests revealed that three explanatory variables had Variance Inflation Factors (VIF) greater than 7.5, with the presence of significant multicollinearity problems, so the removal of these three variables resulted in only four independent variables (Table 1). In addition, the results of the White's test showed that there were no significant heteroskedasticity issues in the model.

Table 1: Descriptions and expected effects of factors

Variable type		Variable name	Variable Description [19]	Expected effect	Adjusted variables
Dependent variable		density	Using the Spatial Join tool in ArcGIS Pro to integrate a point map of tobacco outlets into the data zone map and calculating the density of cigarette outlets (pcs/km2) in each data zone.		√
Independent variable	Income	INCRANKV2	Domain score is constructed by adding all five income indicators and dividing by the 2017 mid-year total population from SAPE. Domain score is ranked to create domain rank.	+	x
	Employment	EMPRANK	Domain score is constructed by adding the three employment indicators and dividing by the 2017 mid-year working age population estimates taken from SAPE. Domain score is ranked to create domain rank.	–	x
	Health	HLTHRANK	Indicators are ranked, transformed to a normal distribution and then combined using weights generated by factor analysis to create the domain score. Domain score is ranked to create domain rank.	–	x
	Education, skills and training	EDURANK	Indicators are ranked, transformed to a normal distribution and then combined using weights generated by factor analysis to create the domain score. Domain score is ranked to create domain rank.	–	√
	Geographic access to services	GACCRANK	Indicators are ranked, transformed to a normal distribution and then combined using weights generated by factor analysis to create the domain score. Domain score is ranked to create domain rank.	+	√
	Crime	CRIMERANK	Domain score is a count of selected recorded crimes, divided by the 2017 mid-year population estimates from SAPE. Domain score is ranked to create domain rank.	+	√
	Housing	HOUSERANK	Domain score is the sum of the two housing indicators, divided by the total household population from the 2011 Census. Domain score is ranked to create domain rank.	–	√

Notes:
1. SAPE: Small Area Population Estimates
2. Refer to *Appendix 1* for specific indicators and ranking calculations for each domain.

3.4 Research Methodology

This paper will use the following three research methods. Please see Appendix II for calculations and descriptions of the above three methods.

1. Global Spatial Autocorrelation

Global spatial autocorrelation is mainly used to describe the overall spatial distribution of a phenomenon in a regional unit and to determine whether the phenomenon is aggregated in space (Su et al., 2019). In this paper, the commonly used Moran's I index is used for spatial autocorrelation analysis, reflecting the distribution similarity of cigarette outlets in spatially adjacent or similar data zones.

2. Local Spatial Autocorrelation

Global Moran's I index can only study the spatial clustering of cigarette outlets in the Edinburgh region, but it is not possible to measure the local distribution, so it is important to consider the global spatial autocorrelation alongside the local one (Zhou, 2020). Local spatial autocorrelation can be used to analyse the clustering in each data zone and is useful for identifying local differences, which are often characterised by the Local Moran's I index (Liu et al., 2021). This study plans to calculate the standardised Local Moran's I index for each study unit and to construct a Moran scatter plot.

3. Geographically Weighted Regression (GWR)

Traditional linear regression models only estimate all samples and parameters globally, without considering spatial patterns and other factors, making it impossible for the independent variables to meet the assumption of independence of residual terms in traditional regression models (OLS models) in the presence of spatial autocorrelation. Besides, the OLS method of parameter estimation is no longer applicable (Cao & Xu, 2018). The GWR model introduces the estimation of the effects of different regions, and it is able to reflect the spatial non-stationarity of the parameters in different spaces, allowing the parameters to vary by spatial location, which means it can offer a more reliable result (Nakaya, 2001). This paper uses GWR analysis.

4. Results and Analysis

4.1 Spatial pattern analysis

4.1.1 Global Characteristics

The spatial correlation of cigarette outlets distribution in Edinburgh was investigated with the software ArcGIS Pro. The retail point map was imported into the data area map through the spatial join tool to obtain a global Moran's I index equal to 0.65. The z-score of 7.6 > 2.58 and p-value of 0 < 0.01 mean that there is < 1% likelihood of that clustered pattern happening at random (Figure 2). The results suggest that the distribution of cigarette outlets in Edinburgh exhibits positive spatial correlation and significant spatial clustering characteristics, i.e., data zones with the same or a similar number of tobacco outlets are spatially adjacent to each other.

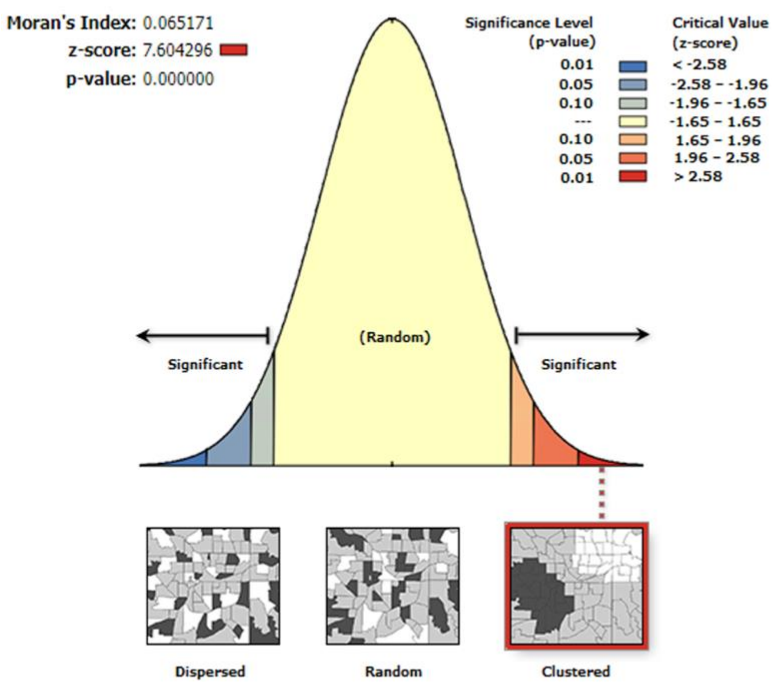


Figure 2: Graphical output from Global Moran's I tool

4.1.2 Local characteristics

The retail outlet map was imported into the data zone map through the Spatial Join tool, and the local spatial autocorrelation was analysed through ArcGIS Pro to derive the local Moran's I index. The Moran's I scatter plot (Figure 3) was drawn using the global Moran's I index as a coefficient. The Moran's I distribution for each data zone (Figure 4) shows significant spatial differentiation, and most of the data areas are in the HL and HH quadrants. HH clusters are mainly located in the city centre with high population mobility and in commercial areas with high accessibility to goods. In addition, restaurants, supermarkets, bars, and other establishments with a high association with cigarettes also contribute to the clustering of these outlets. The LH clusters are mainly located around the HH quadrant and are spread outwards from the city centre. The exceptions to this include the LH clusters located in the west of Edinburgh, which have distributions similar to those around Edinburgh Airport; LL clusters located in the suburbs of Edinburgh, which are formed to serve the surrounding community, and the HL clusters, which are mainly large supermarkets, restaurants or convenience stores that serve the surrounding community. The results of the Moran's I scatter calculations show that the East of Edinburgh area is dominated by HL and LL types of association, but neither of them have statistically significant characteristics and are shown as null in the graph.

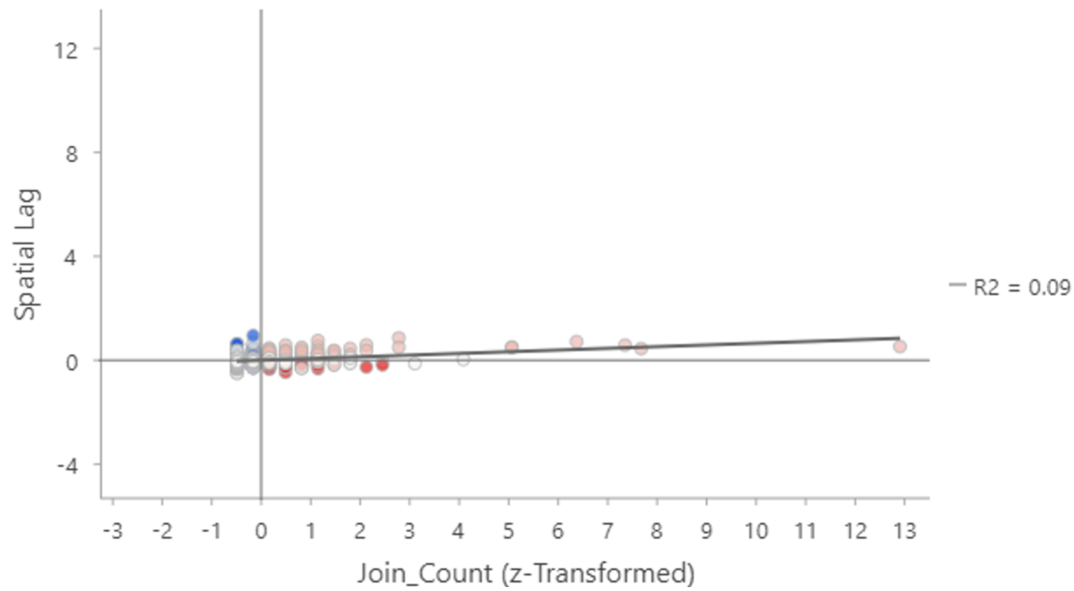


Figure 3: Local Moran's Scatterplot

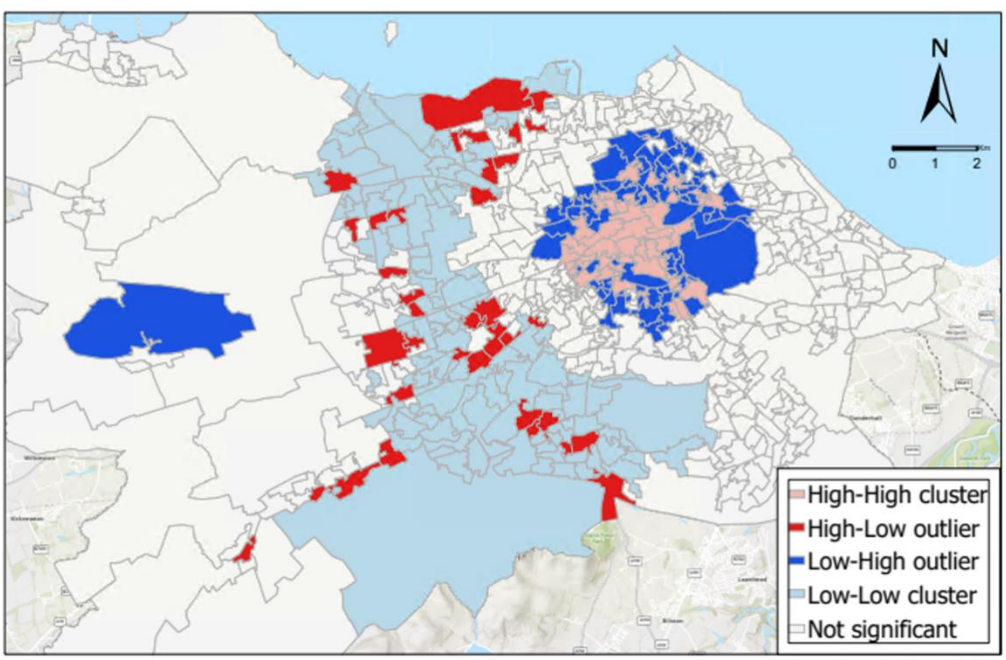


Figure 4: Distribution of spatial aggregation types in the study area

4.2 Spatial Heterogeneity Analysis

4.2.1 Analysis of Influencing Factors Based on the OLS Model

Before using the GWR model to investigate the spatial heterogeneity of cigarette retail density and its socio-economic influences, an OLS model (global regression) was used to test the average relationship with each explanatory variable (Table 2). The coefficient of determination R2 and the corrected coefficient of determination adjusted R2 reached 0.2167 and 0.2065 respectively, i.e., the model explained 20.65% of the Edinburgh cigarette outlet variation in density. The results show that the density of cigarette outlets has a negative relation with the rank of crime and housing, but positive with the rank of education, skills, and training and geographic access to services explanatory variables. A Moran index analysis of the residuals from the OLS revealed that the residuals were autocorrelated and the Koenker (BP) statistic shows that the OLS model varies across the study area, i.e., there is no geospatial consistency between the density of cigarette outlets and the explanatory variables, and therefore spatial instability, i.e., spatial heterogeneity, needs to be addressed by the GWR model.

Table 2: Calculation results of OLS model

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	15.346992	8.238837	1.862762	0.063458	6.501688	2.360463	0.018868*	--------
EDURANK	0.003689	0.001210	3.048210	0.002513*	0.001330	2.773031	0.005896*	1.849638
GACCRANK	0.004047	0.001384	2.924649	0.003713*	0.001014	3.991845	0.000089*	1.326491
CRIMERANK	-0.002818	0.001434	-1.965303	0.050282	0.001732	-1.626589	0.104868	1.464101
HOUSERANK	-0.007243	0.001298	-5.579165	0.000000*	0.001186	-6.106517	0.000000*	1.837655

OLS Diagnostics

Input Features:	SG_SIMD_2020_Clip_Spati	Dependent Variable:	DENSITY
Number of Observations:	310	Akaike's Information Criterion (AICc) [d]:	3080.005043
Multiple R-Squared [d]:	0.216736	Adjusted R-Squared [d]:	0.206463
Joint F-Statistic [e]:	21.099007	Prob(>F), (4,305) degrees of freedom:	0.000000*
Joint Wald Statistic [e]:	84.253410	Prob(>chi-squared), (4) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	17.960319	Prob(>chi-squared), (4) degrees of freedom:	0.001256*
Jarque-Bera Statistic [g]:	5639.517937	Prob(>chi-squared), (2) degrees of freedom:	0.000000*

4.2.2 Analysis of GWR Model Regression Results

The results of the GWR model operation are shown in Table 3. The coefficient of determination R^2 and the corrected coefficient of determination of the model are 0.4911 and 0.3640, respectively, higher than the OLS and the corrected coefficient. Besides, the goodness of fit is improved; the AICc value is lower than that of the OLS model; and the fitting performance of the GWR model has improved compared to the OLS model. The condition numbers of the models are all less than 30, and there is no local multicollinearity between the variables. The median of the parameter estimates from the GWR regressions showed that rank of geographic access to services was positively correlated with cigarette outlet density, and the correlation is the strongest, followed by factors rank of education, skills, and training. Rank of crime and housing showed a negative correlation with cigarette outlet density. During regression, there were some data areas with zero retail points, for which the density could not be calculated, so these data areas were excluded from the regression analysis. In addition, a Moran index analysis of the residuals of GWR revealed that the residuals were not autocorrelated and became discrete in distribution.

Table 3: Calculation results of GWR model

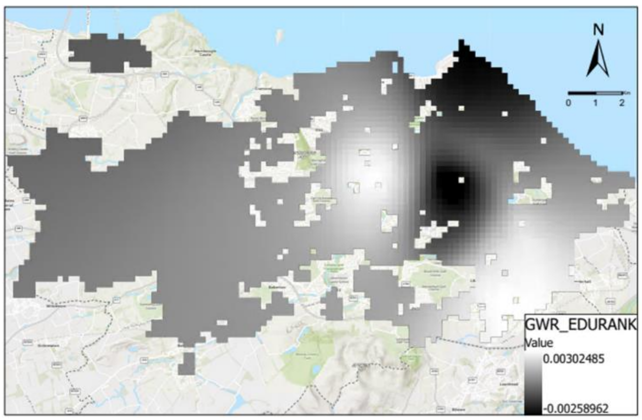
---------- Model Diagnostics ----------

R2:	0.4911
AdjR2:	0.3640
AICc:	3048.3746
Sigma-Squared:	942.8581
Sigma-Squared MLE:	755.1201
Effective Degrees of Freedom:	248.2741

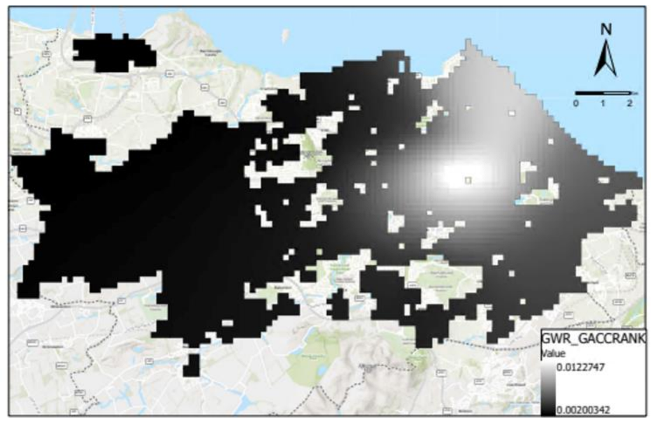
	Minimum	Median	Maximum	Mean
Rank of Education, skills and training	-0.0078	0.00083	0.0078	0.00055
Rank of Geographic access to services	0.0004	0.0045	0.026	0.00619
Rank of Crime	-0.0192	0.00013	0.0058	-0.00233
Rank of Housing	-0.0312	-0.00601	-0.0019	-0.00737

4.2.3 Spatial Heterogeneity Analysis of the Influencing Factors (Figure 5)

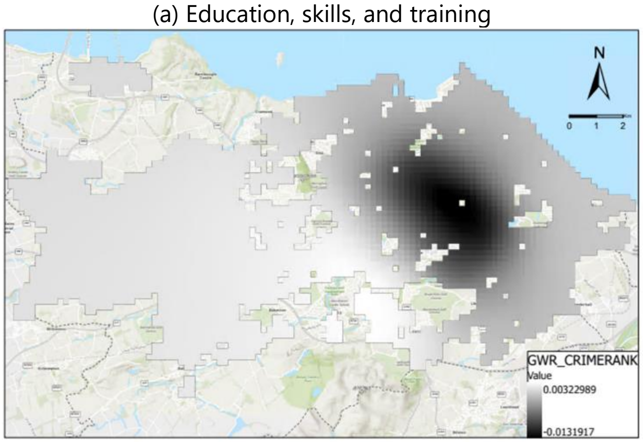
The coefficients of each explanatory variable in the regression outcomes of the GWR model were spatially visualised and expressed to acquire the spatial distribution of each explanatory variable. There was a noteworthy spatial variation in the effect of each explanatory variable on the distribution of cigarette trade outlets.



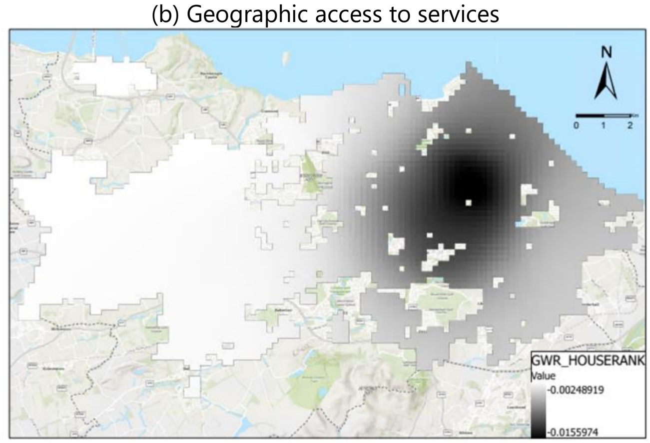
(a) Education, skills, and training



(b) Geographic access to services



(c) Crime



(d) Housing

Figure 5: Spatially divergent patterns of socio-economic factors influencing the distribution of tobacco outlets in Edinburgh

1. Education, skills, and training factors

The education, skills and training element is a varied indicator, judged in terms of the high weighting component; the larger the group with a high level of education, skills, and training, the higher the concentration of cigarette trade outlets in white areas of the graph and the lower the density of trade outlets in black areas. Comparison with the map shows that golf courses, sports centres and pitches are located in white areas, which often gather elite or young people who are studying, with more open spaces and social occasions and a more homogeneous demographic profile. The black areas tend to overlap with school locations, and the Health (Tobacco, nicotine etc. And Care) (Scotland) Act 2016, which places greater restrictions on underage tobacco, therefore shows a negative correlation for the black areas.

2. Geographic access to services factors

Geographic access to services aspects is positively connected to the distribution of cigarette trade outlets in Edinburgh, with the distribution of the degree of influence increasing with distance from the city centre. That is, the nearer the proximity to public services, the more cigarette retail outlets.

The coefficients of each explanatory variable in the regression results of the GWR model were spatially visualised and expressed to obtain the spatial distribution of each explanatory variable. There was significant spatial variation in the effect of each explanatory variable on the distribution of cigarette outlets.

3. Crime and Housing factors

As with the aspect of geographic access to services, the distribution of crime and housing aspects increases with distance from the city centre.

The higher the crime ranking, i.e., the more places with a history of crime, the fewer cigarette trade outlets there are in the city centre, and the density of cigarette outlets rises as the distance from the city centre rises. The lower the housing ranking, i.e., the worse the living environment was, the greater the density of cigarette trade outlets was.

Overall, the GWR model fits better than the OLS model, and its locally based regression of explanatory variables can provide a certain quantitative basis for developing differentiated regional equality strategies.

5. Discussion and Conclusion

This paper provides a global and spatial analyses of the socio-economic factors that offer insight into the differences in the density of cigarette outlets through spatial autocorrelation and GWR models, which provide a scientific basis for the government to design intervention programmes to reduce smoking prevalence and reduce inequalities among social groups. The main conclusions follow.

1. Cigarette outlets in Edinburgh show significant spatial clustering, with a clear trend towards clustering in the high- and low-value retail outlet distribution areas.

2. The overall results of the OLS model regression indicated a positive trend in the distribution of the density of cigarette outlets in data areas with higher rankings for education, skills training, and accessibility of facilities. On the other hand, the OLS model regression indicated a negative trend in the distribution of the density of cigarette outlets in data areas with higher housing and crime rankings.

3. The results of the GWR model regressions indicate that different data areas with different factors have different levels of impact on the distribution of cigarette outlets in Edinburgh.

Therefore, the development of appropriate tobacco control interventions can be tailored to the degree of influence of the various factors in each area. For instance, to effectively reduce smoking rates and socio-economic inequalities, measures can include: a) increased policing in some areas to reduce the incidence of crime, b) targeted smoking cessation support or increased tobacco prices in places with more public facilities, c) advertising bans near junior and senior schools, and d) mass media interventions for young populations.

While this paper only focuses on how some socio-economic factors influence the distribution of cigarette outlets, this distribution is also influenced by various factors that are not included in this paper. Therefore, to supplement these findings, subsequent studies can be carried out. Furthermore, due to the limitations of the panel data, spatial variation in the temporal dimension of the influencing factors is not addressed. For this reason, the inclusion of spatial heterogeneity as an explanatory variable in the long time series in future studies will help to provide a more comprehensive analysis of the variation in cigarette outlet density with socio-economic factors as well as the spatial and temporal heterogeneities. In addition, the data provided for cigarette outlets were from 2019 while the data for socio-economic influences were from 2020. Because of the prevalence of covid-19, it is difficult to be certain that outlets are unaffected, so year differences in data should also be avoided in subsequent studies.

Funding: This research received no external funding
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: https://orcid.org/0000-0003-4243-3245

References

[1] Campbell, C. A., Hahn, R. A., Elder, R., Brewer, R., Chattopadhyay, S., Fielding, J., Naimi, T. S., Toomey, T., Lawrence, B., & Middleton, J. C. (2009). The effectiveness of limiting alcohol outlet density as a means of reducing excessive alcohol consumption and alcohol-related harms. *American Journal of Preventive Medicine, 37*(6), 556-569. https://doi.org/10.1016/J.AMEPRE.2009.09.028

[2] Cao, X., & Xu, J. (2018). Spatial heterogeneity analysis of regional economic development and driving factors in China's provincial border counties. *Acta Geographica Sinica, 73*, 1065-1075.

[3] Connor, J. L., Kypri, K., Bell, M. L., & Cousins, K. (2011). Alcohol outlet density, levels of drinking and alcohol-related harm in New Zealand: A national study. *Journal of Epidemiology and Community Health, 65*(10), 841-846. https://doi.org/10.1136/JECH.2009.104935

[4] Ford, A., MacKintosh, A. M., Moodie, C., Kuipers, M. A. G., Hastings, G. B., & Bauld, L. (2020). Impact of a ban on the open display of tobacco products in retail outlets on never smoking youth in the UK: Findings from a repeat cross-sectional survey before, during and after implementation. *Tobacco Control, 29*(3), 282-288. https://doi.org/10.1136/TOBACCOCONTROL-2018-054831

[5] Liu, T., Sun, P., Ge, D., & Jiang, Y. (2021). Spatial features of economic development and classification of the provincial boundary areas of Jiangsu, Shandong, Henan and Anhui. *Geography and Geo-Information Science, 37*(4), 122-127.

[6] Macdonald, L., Olsen, J. R., Shortt, N. K., & Ellaway, A. (2018). Do 'environmental bads' such as alcohol, fast food, tobacco, and gambling outlets cluster and co-locate in more deprived areas in Glasgow City, Scotland? *Health and Place, 51*, 224-231. https://doi.org/10.1016/J.HEALTHPLACE.2018.04.008

[7] Mayers, R. S., Wiggins, L. L., Fulghum, F. H., & Peterson, N. A. (2012). Tobacco outlet density and demographics: A geographically weighted regression analysis. *Prevention Science, 13*(5), 462-471. https://doi.org/10.1007/S11121-011-0273-Y

[8] McCarthy, W. J., Mistry, R., Lu, Y., Patel, M., Zheng, H., & Dietsch, B. (2009). Density of tobacco retailers near schools: Effects on tobacco use among students. *American Journal of Public Health, 99*(11), 2006-2013. https://doi.org/10.2105/AJPH.2008.145128

[9] Nakaya, T. (2001). Local spatial interaction modelling based on the geographically weighted regression approach. *GeoJournal, 53*(4), 347-358. https://doi.org/10.1023/A:1020149315435

[10] Novak, S. P., Reardon, S. F., Raudenbush, S. W., & Buka, S. L. (2006). Retail tobacco outlet density and youth cigarette smoking: A propensity-modeling approach. *American Journal of Public Health, 96*(4), 670-676. https://doi.org/10.2105/AJPH.2004.061622

[11] Pearce, J., Rind, E., Shortt, N., Tisch, C., & Mitchell, R. (2016). Tobacco retail environments and social inequalities in individual-level smoking and cessation among Scottish adults. *Nicotine and Tobacco Research, 18*(2), 138-146. https://doi.org/10.1093/NTR/NTV089

[12] Romley, J. A., Cohen, D., Ringel, J., & Sturm, R. (2007). Alcohol and environmental justice: The density of liquor stores and bars in urban neighborhoods in the United States. *Journal of Studies on Alcohol and Drugs, 68*(1), 48-55. https://doi.org/10.15288/JSAD.2007.68.48

[13] Rogerson, P. (2006). *Statistical methods for geography: A student guide*. Sage.

[14] Scottish Government. (2014). *Data zone boundaries 2011*. https://www.spatialdata.gov.scot/geonetwork/srv/api/records/7d3e8709-98fa-4d71-867c-d5c8293823f2

[15] Scottish Government. (2015). *Comparing population and area of data zones 2001 and 2011 methodology report*. https://www.gov.scot/binaries/content/documents/govscot/publications/statistics/2020/01/small-area-statistics-reference-materials/documents/data-zone-matching---quick-guide-2001-2011/data-zone-matching---quick-guide-2001-2011/govscot%3Adocument/Data%2BZone%2BMatching%2B-%2BQuick%2BGuide%2B%25282001-2011%2529.pdf

[16] Scottish Government. (2021). *SIMD 2020 technical notes*. https://www.gov.scot/publications/simd-2020-technical-notes/

[17] Shortt, N., Tisch, C., Pearce, J., Richardson, E., & Mitchell, R. (2016). The density of tobacco retailers in home and school environments and relationship with adolescent smoking behaviours in Scotland. *Tobacco control, 25*(1), 75-82. https://doi.org/10.1136/tobaccocontrol-2013-051473

[18] Shortt, N. K., Tisch, C., Pearce, J., Mitchell, R., Richardson, E. A., Hill, S., & Collin, J. (2015). A cross-sectional analysis of the relationship between tobacco and alcohol outlet density and neighbourhood deprivation. *BMC Public Health, 15*(1), 75-82. https://doi.org/10.1186/S12889-015-2321-1

[19] Shortt, N. K., Tunstall, H., Mitchell, R., Coombes, E., Jones, A., Reid, G., & Pearce, J. (2021). Using point-of-sale data to examine tobacco pricing across neighbourhoods in Scotland. *Tobacco Control, 30*(2), 168-176. https://doi.org/10.1136/TOBACCOCONTROL-2019-055484

[20] Su, S., Li, L., & Weng, M. (2019). *Spatial data analysis*. Science Press.

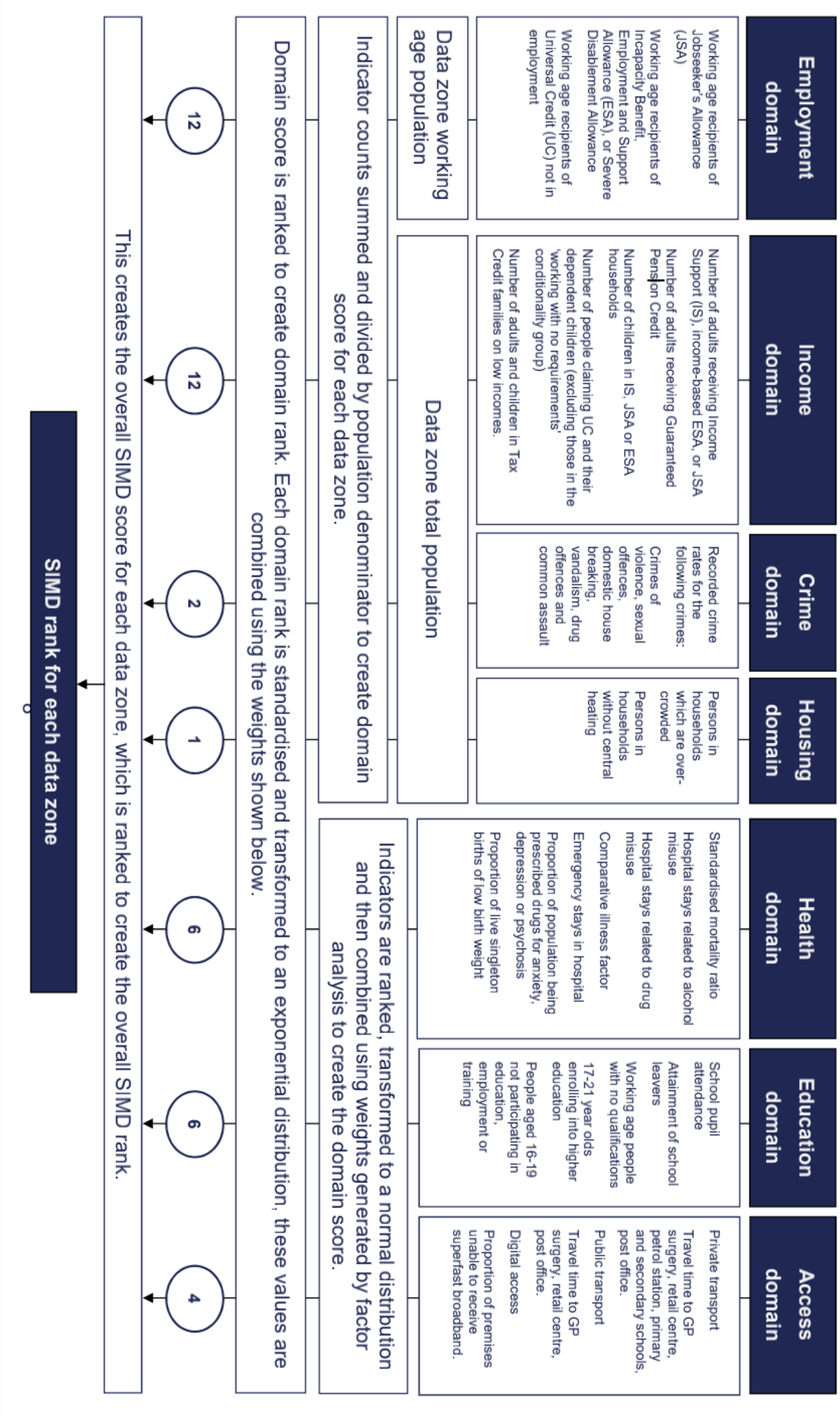
[21] Tobacco Register Scotland. (2017). *About the register - STRR*.

[22] Wheeler, D. C., Do, E. K., Hayes, R. B., Fugate-Laus, K., Fallavollita, W. L., Hughes, C., & Fuemmeler, B. F. (2020). Neighborhood disadvantage and tobacco retail outlet and vape shop outlet rates. *International Journal of Environmental Research and Public Health, 17*(8), Article 2864. https://doi.org/10.3390/IJERPH17082864

[23] Yu, D., Peterson, N. A., Sheffer, M. A., Reid, R. J., & Schnieder, J. E. (2010). Tobacco outlet density and demographics: Analysing the relationships with a spatial regression approach. *Public Health, 124*(7), 412-416. https://doi.org/10.1016/J.PUHE.2010.03.024

[24] Zhou, F. (2020). *Research on spatial differences and influencing factors of urban-rural integration in the Yangtze river delta region* (Publication Number 10295-1020346249) [Master's thesis, Jiangnan University]. CNKI.

Appendices I SIMD 2020 Methodology (Scottish Government, 2021)



Appendices II Calculation and description of research methods

1. Global Spatial Autocorrelation (Global Moran's I)

The Moran's I index ranges from [-1, 1], with I > 0 indicating that the distribution of cigarette outlets in the study area is positively correlated (The closer the value is to 1, the stronger the spatial clustering); I < 0 indicates that the distribution in the study area is negatively correlated (The closer the value is to -1, the stronger the dispersed distribution trend); I = 0 indicates that the cigarette outlets have a random spatial distribution (Liu et al., 2021).

2. Local Spatial Autocorrelation (Local Moran's I)

The standardised observations and spatial lags were analysed to classify all study units into four categories i.e., High values clustered in the first quadrant (H-H type), Low and medium values surrounded by high values in the second quadrant (L-H type), Low values clustered in the third quadrant (L-L type), High values surrounded by low values in the fourth quadrant (H-L type) (Zhou, 2020).

The data area in quadrants first and third indicates that the spatial distribution has positive spatial correlation and that in quadrants second and fourth means a negative spatial correlation (Zhou, 2020).

3. Geographically Weighted Regression (GWR) (Rogerson, 2006)

with the following model structure.

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i$$

Where: y_i is the observed value, $\beta_0(u_i, v_i)$ is the regression coefficient at point i indicating the degree of influence of the independent variable on the dependent variable, (u_i, v_i) is the coordinates of the geographical centre of the ith sample spatial unit, $\beta_k(u_i, v_i)$is the value of the continuous function $\beta_k(u, v)$ in the i sample spatial unit; x_{ik} denotes the value of the independent variable x_k at point i; ε_i is a normally distributed function with a constant variance, representing the random error term.