UCL DEPARTMENT OF GEOGRAPHY

YEAR 2022-23

EXAM CANDIDATE ID:	**ZJXV1**
MODULE CODE:	**GEOG0114**
MODULE NAME:	**Principles of Spatial Analysis**
COURSE PAPER TITLE:	**Understanding relationships between cultural accessibility and educational deprivation in Inner London**
WORD COUNT (excluding title and figure annotations):	**2960**
CODE REPOSITORY LINK:	**N/A**

Your essay, appropriately anonymised, may be used to help future students prepare for assessment. Double click this box to opt out of this ☐

<u>Understanding relationships between cultural accessibility and educational deprivation in Inner London</u>

Introduction

Cultural venues play a significant role in an urban agglomeration, affecting tourism and individual interaction but support urban revival (Zhang et al., 2017). Accessibility to cultural venues has been found to be strongly influenced by their proximity. Indeed, individuals with better access to cultural venues have been recorded to elicit behavioural changes, affecting their social inclusion, and increasing their cultural capital (Brook, 2016). We believe that understanding population interaction to these cultural venues is important. UNESCO led initiatives for younger people to provide them with exposure and active participation in culture (UNESCO, 2016) are an example of these positive effects.

This article aims to dive deeper in understanding the influence of cultural proximity in London, considering the city's role as a cultural epicentre of the west since the industrial revolution. We define cultural venues to include museums, art galleries, grass-root music venues, theatres, and libraries. Our study aims to understand the interaction between cultural accessibility and measures of deprivation. Specifically, we want to focus on education as we want to provide policy indications as to where additional cultural venue implementation would be necessary. Indeed, understanding educational deprivation and how to target it is primordial. Access to cultural institutions provides soft skills (Majid et al., 2012) and the Bourdieusian concept of cultural capital (Erel, 2010) which are an important factor for individuals aiming for financial development, especially from higher socio-economic deprivation backgrounds. We carry a particular interest in the relationship between cultural proximity and educational deprivation.

To do this, we create a proxy measure of cultural accessibility through weighted distance calculations to the three closest venues. Once this measure was created, the study uses four spatial models and 2019 UK deprivation variables to understand relationships between educational deprivation and access to cultural venues. We ran a simple linear model, and then provided more precise global overview with the Spatial Lag and Spatial error models, which account for spatial autocorrelation. We continue by providing local estimates and analysis with the use of Geographically Weighted Regression (GWR). The Spatial Lag and GWR providing the most successful results, we use these to make policy proposals for venue implementation in educationally deprived areas, such as west Wandsworth, south Lambeth and Greenwich. Finally, we acknowledge certain limitations in these statistical methods, yet providing a clear rational behind our choice of method and framework. We also provide indications for future research and the impact of COVID-19 lockdowns and sanitary regulations on current and future interactions to live culture and entertainment.

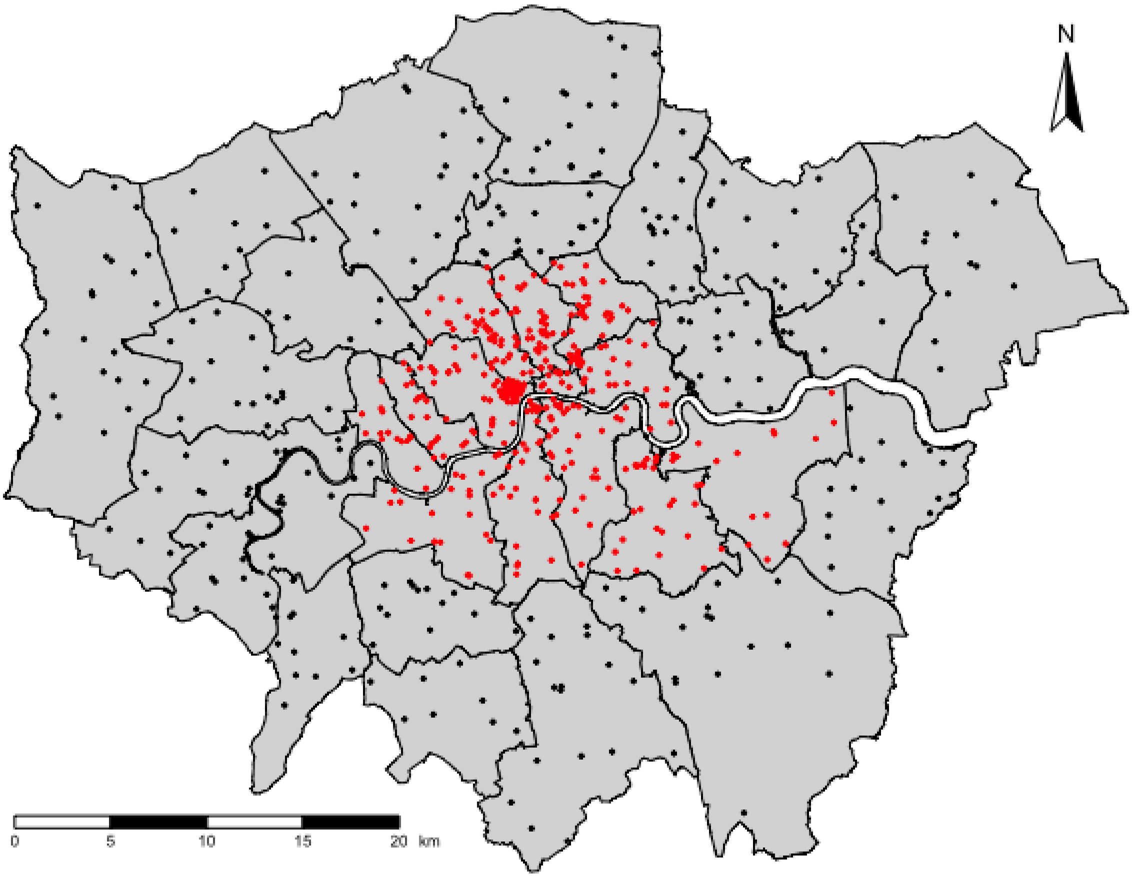
Our study is an ecological study with a cross-sectional framework, we study groups of individuals, at Lower Super Output Areas (LSOA) level, over a single point in time. This is best fit for this situation as it allows for the inclusion of a large area and amount of people (all inner London LSOA's) whilst also being quicker and more efficient. We consider the methodology used in the study as an example of how to understand the relationship between cultural access and population deprivation, which can be applied to different locations when relevant or new data is made available.

Data and Methods

We used a file containing the 4968 LSOA's in Greater London and 1737 in Inner London. We decided to use LSOA's as they currently provide the highest granularity when reporting on small areas in the United Kingdom. For additional visibility, we also include borough shapes and names on maps presented in this report.

We proceeded to collect information on the location of various cultural amenities around London, available from a cultural infrastructure initiative led by the Greater London Authority in 2019. In total, our cultural venues dataset contains 703 locations in greater London and 398 in inner London.

Figure 1: distribution of museum around Inner London (red) and Greater London (black)



Our next aim was to create a measure representing cultural accessibility for individuals living in different LSOA's. We used average walking distance from the three closest cultural venues as a proxy measure for accessibility. To do so, we considered each LSOA centroid as the departure point for its inhabitants when going to these locations. We acknowledge that taking a single point as a representation of an area can cause falsified results. In our case, we do not consider the size of the LSOA to have major effects on the walking access of an individual to their closest cultural locations. We calculated that the average LSOA size was approximately 416 meters in diameter[1] meaning that individuals should not find themselves further than 200 meters away from their closest centroid. 144 seconds if we consider an average walking pace of 5km per hour.

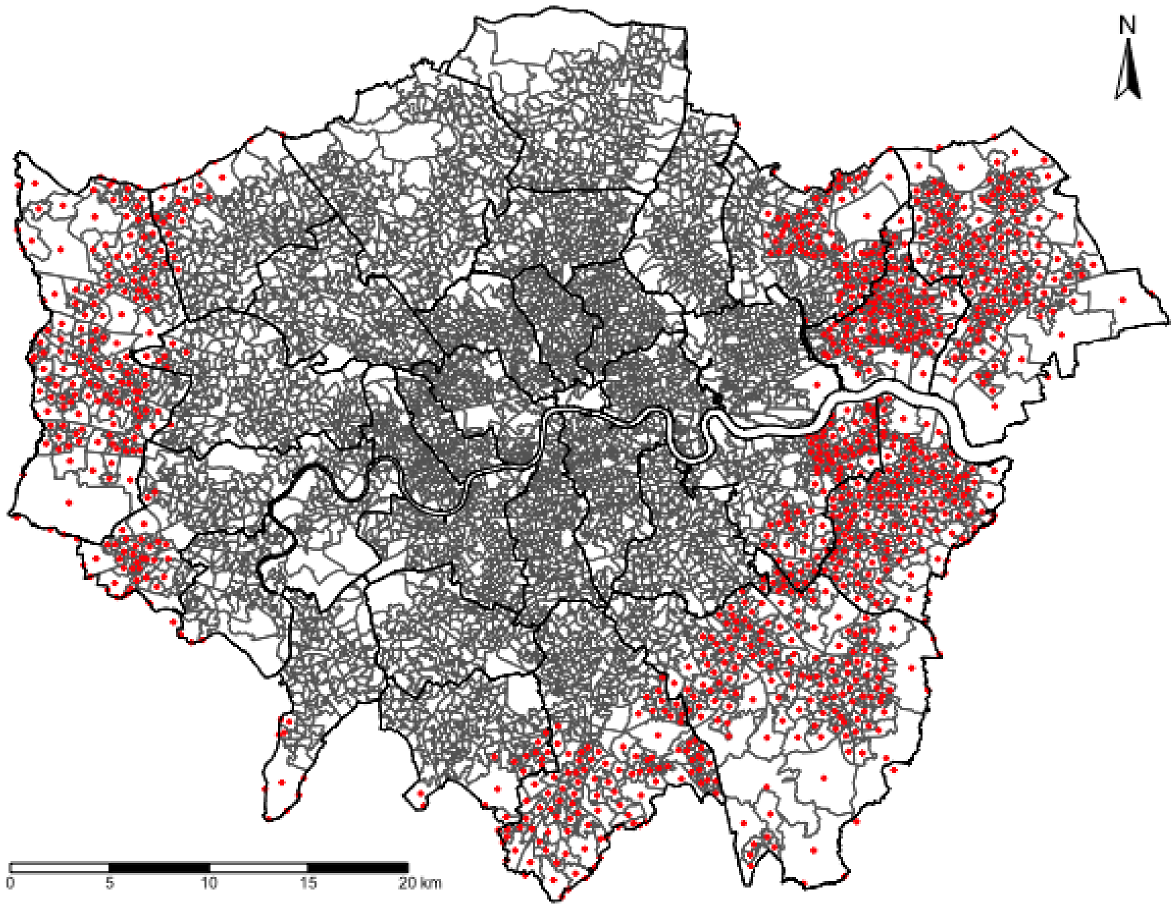
We then proceeded to extract the road network for Greater London from Open Street Maps (OSM). Once constructed, we created a weighted network graph which modelled walking accessibility.

[1] Also acknowledging that LSOA shapes are not perfectly circular or square, but rather random polygons

With this method, we got representative distances that the individuals would have to travel, including *cul-de-sacs* they need to avoid or major highway that they cannot cross. This allows us to have more realistic results compared to Euclidean distances, which do not account for layout of the environment.

Once in place, we calculated the distances between every origin (LSOA centroid) and destination (Cultural Amenities) in greater London using the *dodgr* package in R. Our initial results found that that 1094 centroids did not have direct walking access to three or more museums (illustrated in figure 2). For more precise results we ran the analysis again on a new study area, inner London, which was successful. We will be focusing our analysis on 1737 LSOA's and 398 cultural venues of inner London. Inner London is composed of twelve boroughs[2] and the 'City of London'.

Figure 2: LSOA centroids where distance to cultural amenities could not be calculated (in red) – as a result we decide to focus on inner London rather than Greater London



The initial results were returned in meters, but we decided to convert them into seconds of walking. We considered the average walking pace to be at 5km per hour (Bohannon, 1997), accounting for stops that need to be made (i.e., crossing a street) and is representative of general population characteristics such as age and gender. With the distance from LSOA centroids to the three closest cultural venues by walking distance in seconds, we were able to create a proxy measure for cultural accessibly.

[2] These include: "Camden", "Greenwich", "Hackney", "Hammersmith and Fulham", "Islington", "Kensington and Chelsea", "Lambeth", "Lewisham", "Southwark", "Tower Hamlets", "Wandsworth", "Westminster", "City of London"

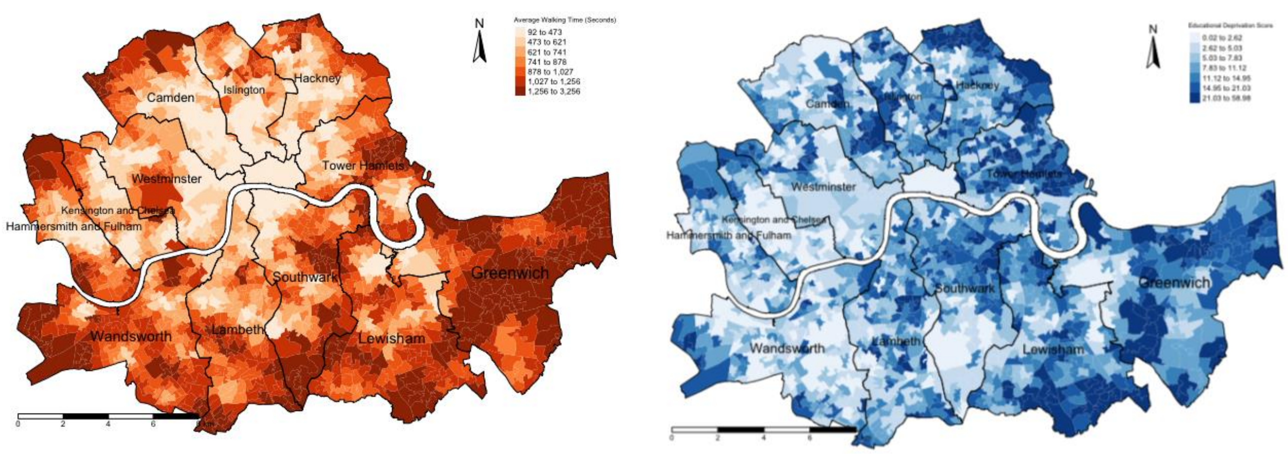
With the proxy measure and area delimitation defined, we then proceeded to add deprivation scores from the 'English indices of deprivation 2019' (UK Government, 2019), which had scores for each LSOA.

We decided to normalise the data distribution for both independent and dependant variables. As much of the data is negative, we add constants to the variables which have values under 0, to make all values positive, and then apply square root normalisation. This normalisation is beneficial as it limits the skewedness of the distributions of the variables used although making precise interpretability of the results more difficult.

Results:

We begin by looking at the spatial distribution of the average walking time and education across inner London LSOA's.

Figure 3: Descriptive Statistics for Average Walking Time (Left) and Educational Deprivation (Right)



A linear regression was run, with cultural accessibility as the outcome variable and education deprivation as the main independent variable whilst also including Income, Health and Disability and Living Environment attributes to benefit the accuracy model. The latter three variables, Income Score, Health and Deprivation score and Living environment score, will not be discussed but rather served in model construction and allow for better representativeness. The linear model has an Adjusted R-squared of 0.1513, meaning that approximately 15% of cultural accessibility in London can be explained by our variables, of which Education will be our main interest. We decide to continue with this model, as an R-squared value above 10% is accepted in the social-sciences (Ozili, 2022).

We proceed to extract the residuals to see if spatial autocorrelation is present and if the results provided by the linear regression are reliable.

Figure 4: Residuals from the Simple Linear Model, indicating high spatial autocorrelation

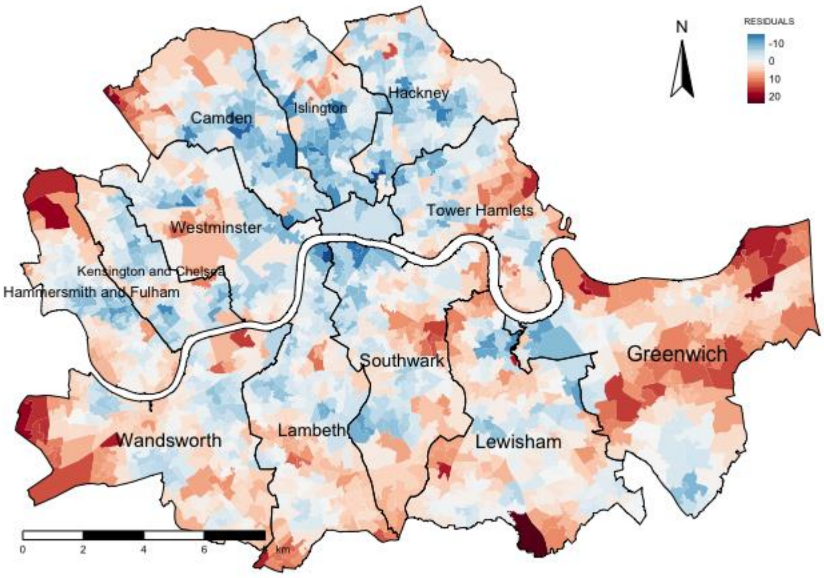


Figure 4 illustrates that these residuals are separated into two parts, over-prediction in blue and under-prediction in red. These results do not seem out of place as we are interested in the relationship between deprivation scores in educational attainment and cultural accessibility. This is partly explained by figure 1, as most cultural venues are clustered in central London, and figure 3 showing that higher deprivation scores tend to be in outskirt LSOA's, which have high residuals.

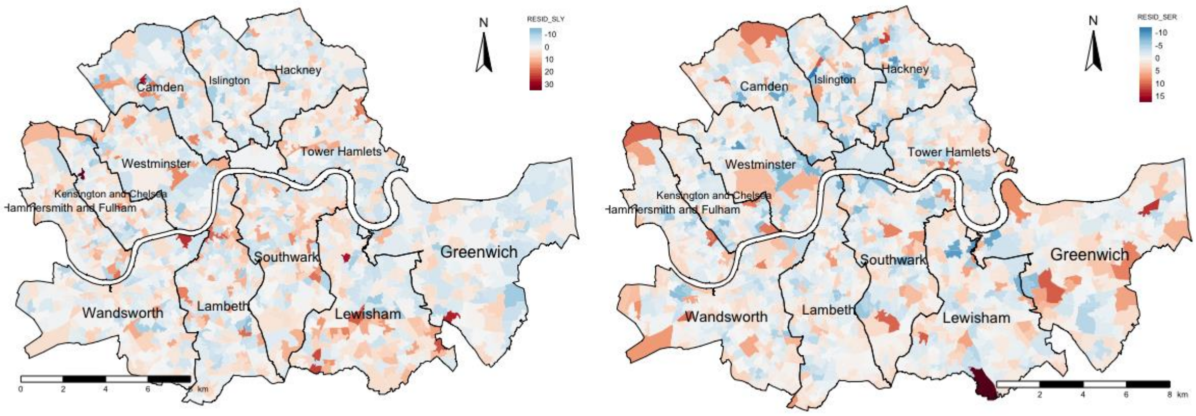
From visual interpretation of figure 4, we believe that spatial autocorrelation is likely, and Moran's I estimate for this model (Table 1) confirms this. A statistically significant ($p<0.05$) Moran's I value of 0.685 indicates positive spatial autocorrelation, meaning neighbouring values tend to be similar. This is in line with current literature as small geographical scale areas tend not to have direct changes in values, but rather a gradual increase or decrease across several of these areas (Brattbakk, 2014).

Table 1: Summarising Moran's I statistic for all three models

	Linear Regression	**Spatial Lag**	**Spatial Error**
Moran's I	0.685	-0.032	0.005
P-value	$<2.2e^{-16}$	0.98	0.346

We continue our analysis with two additional spatial models addressing spatial autocorrelation: the spatial lag model (SLM) and the spatial error model (SER). Figure 6 illustrates the residuals for both models.

Figure 5: residuals for the Spatial Lag (Left) and Spatial Error (right) models



Visually, we see less apparent bright red or blue residual values compared to figure 4, indicating that these models have successfully accounted for spatial autocorrelation. SLM and SER have respective Moran I values of *0.005* and *-0.032* with p-values of *0.98* and *0.35* indicating no evidence of spatial autocorrelation in these models.

Table 2 provides evaluation metrics for all three different models presented above whilst also including Geographically weighted regression (GWR). GWR has the advantage to overcome certain limitations in standard linear regression, SLM and SER as it aims to understand the evolution of relationships between the dependent and independent variables at a local level. Furthermore, GWR is also argued to better address spatial heterogeneity (Gao et al., 2020).

Table 2: Model evaluation through AIC comparison

	Linear Regression (LM)	**Spatial Lag (SLM)**	**Spatial Error (SER)**	**Geographically weighted regression (GWR)**
AIC	11212	8939.60	8961.20	8363.48
R-squared	0.151	0.726	0.718	0.886[3]

Our SLM model has a rho value of 0.911, indicating that neighbouring LSOA's have a statistically significant (p-value<0.05) positive effect. Furthermore, the SLM provides a considerable improvement on the AIC compared to that of the simple linear regression. It decreases from 11212 for the LM to 8940 for the SLM.

The Lambda estimate for the Spatial error model indicates whether a sudden change in the error term is related to spatially correlated errors, how it impacts the error term of our cultural accessibility measurements in each LSOA. Our SER has a statically significant (p-value<0.05) Lambda value of 0.923 meaning that the effect of neighbouring LSOA's is positive. The AIC value from the SER is at 8961, which is again a considerable improvement compared to that of the simple linear model, but slightly higher than the SLM. In addition, both SLM and SER provide a greatly increased R-squared value from the initial 0.151 in the LM to 0.726 and 0.718.

From these results, we consider SLM to be a better model than the LM and SER for accounting for spatial autocorrelation with the data and variables we have used in its construction. SLM performs

[3] This is a quasi-global R-squared value meaning that it is the average r-squared value in each of the individual LSOA's in inner London

better as it has a better AIC and R-squared than the other two models and has a greater p-value when calculating Moran's I, making this more reliable and correctly account for spatial autocorrelation.

Table 3 illustrates the impact coefficients for the SLM model. Our variable of interest, education, has a 0.315 direct effect and a 2.33 total effect. Most importantly, these impacts are positive. This means that an increase in the distance to cultural venues, is positively associated to an increase in education deprivation score. In other words, there is a correlation between educational deprivation and cultural accessibility in the Inner London LSOA's.

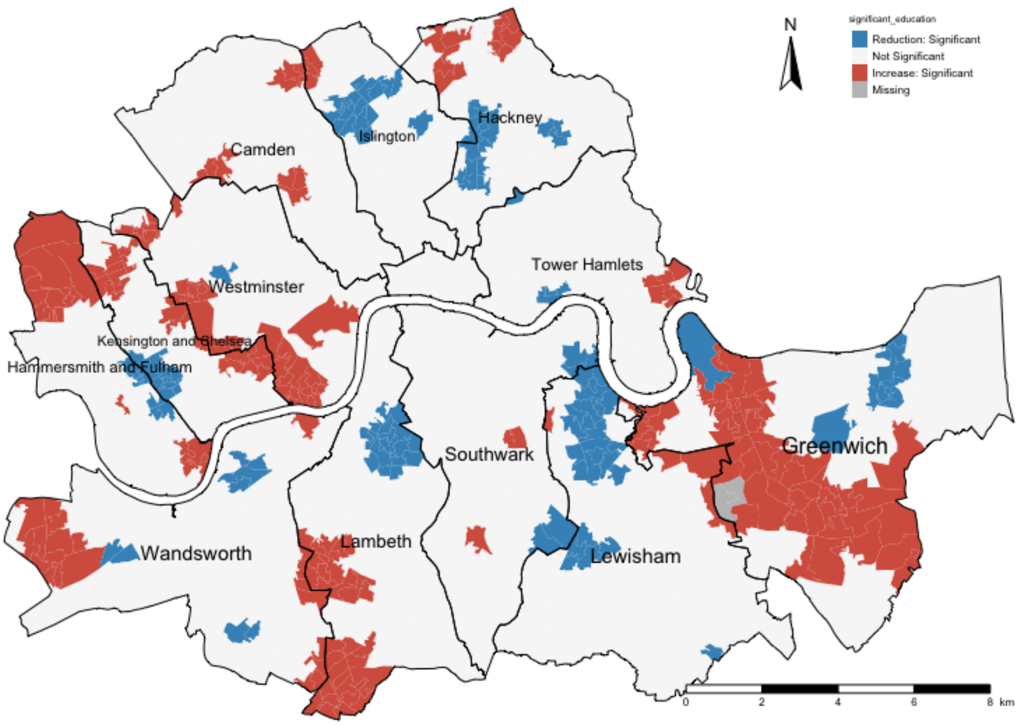
The SLM accounts for spatial autocorrelation and, even without precise interpretation of the impact values, the positive impact illustrates a significant relationship. We are confident that these results are representative of unequal levels of cultural accessibility in London.

Table 3: Summary of impacts from the Spatial Lag model[4]

	Direct	Indirect	Total
Sqrt (Employment)	-4.2866505	-27.4429773	-31.729628
Sqrt (Education)	**0.3151842**	**2.0177974**	**2.332982**
Sqrt (Living Environment)	-0.5108181	-3.2702385	-3.781057
Sqrt (Health and Disability)	0.1487538	-0.9523162	-1.101070

With initial results showing relationships between cultural accessibility and educational deprivation, we now look at the results through the Geographically weighted regression which provides local estimates across individual LSOAs. Figure 6 highlights both significant increase and reduction in educational deprivation with regards to cultural proximity.

Figure 6: LSOA's with significant relationships between levels of educations and cultural proximity



[4] Summary impacts include the results from the other variables used in the analysis, but their effects will not be analysed, rather they are used to support model accuracy

Figure 6 provides a succinct overview of the GWR results, aiming to pinpoint the relationship between cultural proximity and educational deprivation. With this information, we propose policy measures regarding future implementation of cultural venues to lower inequalities in access. West Wandsworth, south Lambeth, north Hammersmith and parts of Greenwich show a significant increase in educational deprivation as cultural proximity increases. In the aforementioned areas, cultural accessibility is hindered (i.e., further away from cultural venues) and associated with an increase in educational deprivation, indicating geographical inequalities. We believe government and council agencies, such as the Greater London Authority (GLA) should implement more venues such as libraries, museums and galleries in that area. This would provide the local population, especially the youth, with increased access to cultural and educative tools, important for their development.

Figure 7: R-squared values across all inner London LSOA's

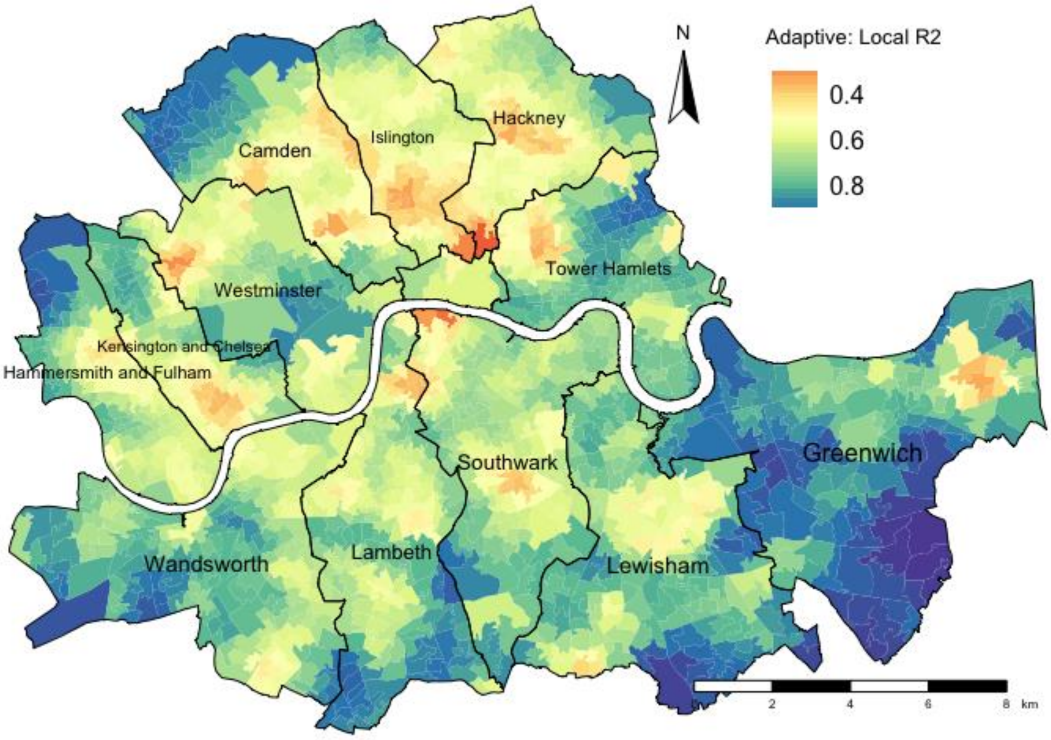


Figure 7 shows the distribution of R-squared values across the LSOA's. The Global R-squared value for the GWR model being the highest of the four tested at 0.88. We see that certain areas witness significant increases (figure 6) but low local R-squared values (figure 7). For this reason, we include neither Westminster nor central Lambeth in our areas for policy recommendation. We also notice that R-squares tend to be higher on the outskirts, possibly prone to more externalities in more central areas.

Limitations

We acknowledge that certain limitations exist, and we attempt to cover these below.

The impact of COVID-19 government regulations, repeated lockdowns and economic difficulties and how it affected cultural venues is yet unknown. We acknowledge that an important proportion of cultural venues were closed during the pandemic, with effects especially strong for independent music grass-root venues and theatres (Drury, 2022). These effects were not only economic, as many of these businesses had to close for extended periods of time, but also due to a psychological shift in consumer behaviour (Zwanka and Buff, 2021). After the government imposed sanitary restrictions, digital entertainment consumption was privileged and the return to physical shows didn't return to pre-COVID levels. Research argues that risk averse consumer preferences did not help with the mitigation of the effect to live entertainment consumption (Taylor, Raine and Hamilton, 2020). We understand that this affects the cross-sectional framework, making the results less representative of the current situation.

We currently do not have any available data on the extent of the effect of COVID-19 on venue closures and loss of revenue, but we aim to investigate this once further data is collected. Nonetheless, many of the libraries, museum and galleries are government or council run, therefore we expect they are less impacted by the COVID crisis. Our study serves as a methodological framework and example which can be reutilised once new data is available.

The use of granular level data allows us to provide representative data on London population, having high internal validity. On the other hand, the external validity of the results are limited, as the data is highly specific to the London population. These results would not be applicable to other population groups as their characteristics are likely different. Yet, the way in which the study is constructed allows for high external validity. Indeed, taking representative population and cultural accessibility information for other areas, we can apply a similar methods and models.

We acknowledge the SER and SLM models have limitations, as they provide global overviews of the study area. GWR limits this, providing local estimates. GWR is also prone to certain limitations such as multi-collinearity (Bivand et al, 2008) and unsmooth surfaces due to potential variations in mean parameter estimates. In further research, we aim to investigate other versions of GWR, such as Conditional GWR (CGWR), which considers the varying bandwidths that influence variations in parameter estimates (Leong and Yue, 2017). Nonetheless, GWR is efficient in exploring patterns in spatial data and that why it was selected for this research. Using inputs from these models, our research provides both global estimate directions, with positive impacts from the SLM, and LSOA level information through GWR results, visualised in figure 7.

Conclusion

In our analysis, we used four spatial models to try and understand relationships between cultural proximity and educational deprivation in Inner London at LSOA level. The analysis of residuals from the simple Linear Model reveals that there is existence of spatial autocorrelation. To overcome this, we implement both the Spatial Lag and Spatial Error models. The Spatial Lag model provides the biggest improvement and gives confident estimates that positive impacts exist between cultural distance and educational deprivation. London individuals with lower levels of education tend to have more difficult access to cultural venues, which further emphasises geo-demographic inequalities. As these latter models provide global estimates, we also implemented a Geographically weighted regression to provide Local LSOA estimates. Whilst acknowledging limitations in our models and data selection, we propose indications for future research using new methods and applying the study framework to new data or areas.

This study presents its results as policy recommendation for the Greater London Authority. We argue the necessary the implementation of cultural venues in Inner London, notably south Lambeth, west Wandsworth and Greenwich to counteract geographical inequalities in cultural accessibility for areas with higher educational deprivation.

References

Bivand, R.S., Pebesma, E.J., Gómez-Rubio, V. and Pebesma, E.J., 2008. *Applied spatial data analysis with R* (Vol. 747248717, pp. 237-268). New York: Springer.

Bohannon, R.W., 1997. Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants. *Age and ageing*, *26*(1), pp.15-19.

Brattbakk, I., 2014. Block, neighbourhood or district? The importance of geographical scale for area effects on educational attainment. *Geografiska Annaler: Series B, Human Geography*, *96*(2), pp.109-125.

Brook, O., 2016. Spatial equity and cultural participation: how access influences attendance at museums and galleries in London. *Cultural Trends*, *25*(1), pp.21-34.

Drury, J., Rogers, M.B., Marteau, T.M., Yardley, L., Reicher, S. and Stott, C., 2021. Re-opening live events and large venues after Covid-19 'lockdown': Behavioural risks and their mitigations. *Safety science*, *139*, p.105243.

Erel, U., 2010. Migrating cultural capital: Bourdieu in migration studies. *Sociology*, *44*(4), pp.642-660.

Gao, C., Feng, Y., Tong, X., Lei, Z., Chen, S. and Zhai, S., 2020. Modeling urban growth using spatially heterogeneous cellular automata models: Comparison of spatial lag, spatial error and GWR. *Computers, Environment and Urban Systems*, *81*, p.101459.

Khlystova, O., Kalyuzhnova, Y. and Belitski, M., 2022. The impact of the COVID-19 pandemic on the creative industries: A literature review and future research agenda. *Journal of Business Research*, *139*, pp.1192-1210.

Leong, Y.Y. and Yue, J.C., 2017. A modification to geographically weighted regression. *International journal of health geographics*, *16*(1), pp.1-18.

Majid, S., Liming, Z., Tong, S. and Raihana, S., 2012. Importance of soft skills for education and career success. *International Journal for Cross-Disciplinary Subjects in Education*, *2*(2), pp.1037-1042.

Ozili, P.K., 2022. The acceptable R-square in empirical modelling for social science research. *Available at SSRN 4128165*.

Taylor, I.A., Raine, S. and Hamilton, C., 2020. COVID-19 and the UK live music industry: a crisis of spatial materiality. *The Journal of Media Art Study and Theory*, *1*(2), pp.219-241.

UNESCO, 2016, "Cultural Education Projects, Art Events Designated for and Attended by Children and Young People." *Diversity of Cultural Expressions*, https://en.unesco.org/creativity/policy-monitoring-platform/cultural-education-projects-art.

UNESCO, 2022, "Cutting Edge: Culture & Education: A Strategic Investment for Inclusive and Sustainable Development." *UNESCO.org*, https://www.unesco.org/en/articles/cutting-edge-culture-education-strategic-investment-inclusive-and-sustainable-development.

Zhang, H., Xu, F., Lu, L. and Yu, P., 2017. The spatial agglomeration of museums, a case study in London. *Journal of Heritage Tourism*, *12*(2), pp.172-190.

Zwanka, R.J. and Buff, C., 2021. COVID-19 generation: A conceptual framework of the consumer behavioral shifts to be caused by the COVID-19 pandemic. *Journal of International Consumer Marketing*, *33*(1), pp.58-67.