Statistics for Social Research III

Derrick Snow

Omega Graduate School

February 17, 2024

Professor

Dr. Taladay

Assignment

**Assignment #2 – Developmental Readings**

1. Create Developmental Readings from seminal sources and scholarly peer-reviewed

journal articles. Review instructions for Assignment #3, the course essential elements, and course readings to identify selections of books and journals to create entries.

1. Refer to the "Student Guide to Developmental Readings" in the General Helps folder for updated information on sample comments, the grading rubric, and key definitions related to developmental readings.

**Source One:** Kent, J. T., & Mardia, K. V. (2022). *Spatial analysis* (Vol. 72). John Wiley & Sons.

**Comment 1:**

**Quote/Paraphrase:** “Spatial analysis involves the analysis of data collected in a spatial region. A key aspect of such data is that observations at nearby sites tend to be highly correlated with one another. Any adequate statistical analysis should take these correlations into account.” (Kent & Mardia, 2022, p.9)

**Essential Element:** Statistical Procedures

**Additive/Variant Analysis:** The concept of spatial autocorrelation in spatial analysis is additive to my understanding of spatial statistics.

**Contextualization:** The issue of spatial autocorrelation is an important factor in spatial analysis, where one may find that there is a correlation based on closeness to geographical areas. There are statistical procedures to examine if spatial autocorrelation exists and the amount, such as Moran’s I. In my research, I am curious if I will have an issue with spatial autocorrelation because land use is somewhat influenced by surrounding land use/land cover. However, there are distinct classifications of land use, and nearby areas may have different classifications. Thus, I don’t see that spatial autocorrelation will be an issue that will need to be addressed in my research.

**Source Two:** Preye Winston Biu, Chinedu Nnamdi Nwasike, Olawe Alaba Tula, Chinedu Alex Ezeigweneme, & Joachim Osheyor Gidiagba. (2024). A review of GIS applications in public health surveillance. *World Journal of Advanced Research and Reviews*, *21*(1), 030–039. <https://doi.org/10.30574/wjarr.2024.21.1.2684>

**Comment 2:**

**Quote/Paraphrase:** “GIS goes beyond mere mapping by offering sophisticated spatial analysis tools that deepen the understanding of health-related phenomena. Spatial analysis allows for identifying spatial clusters, correlations between health outcomes and environmental factors, and assessing accessibility to healthcare resources.” (Biu, et al, 2024, p. 32)

**Essential Element:** Statistical Procedures

**Additive/Variant Analysis:** The concept of statistical capability with in GIS is additive to my understanding of GIS.

**Contextualization:** Geographic information system (GIS) has been a geographic tool for so many fields that have a geographic component (geology, environmental, market analysis, healthcare, etc.) it not only a tool to create maps, but is a whole analysis tool. There are many standard statistical and geostatistical procedures within the tool that can be used to analyze data, spatial and attribute data (non-spatial but attached to geographical with zip codes or county). There are many geographic aspects to health, illness, and healthcare, from areas of poverty, to environmental factors to access to care. Over the years the use of GIS has became more prominent in health and illness research, however, there are so many more aspects of health research could be studied with a geographical perspective. I have had some experience with GIS and have a graduate certificate in Geospatial analysis, but it has been a number of years since I worked in GIS and many advances have been made in GIS. Thus, I am in the process of relearning GIS as a tool I could use in my social research.

**Source Three:** Nedd, R., Light, K., Owens, M., James, N., Johnson, E., & Anandhi, A. (2021). A Synthesis of Land Use/Land Cover Studies: Definitions, Classification Systems, Meta-Studies, Challenges and Knowledge Gaps on a Global Landscape. *Land*, *10*(9), 994. <https://doi.org/10.3390/land10090994>

**Comment 3:**

**Quote/Paraphrase:** “Land-use and land-cover can carry separate definitions, where land-use relates to what purpose the land is utilized, e.g., agricultural or recreational use. In contrast, land-cover states specific landscape patterns and characteristics. While the terminologies for LULC may be used interchangeably, the concept remains the same for any particular region. It focuses on man’s utilization in time and space of the various physical, chemical, and cultural factors of the land.” (Nedd et al., 2021, p.2)

**Essential Element:** Independent Variables

**Additive/Variant Analysis:** The concept of land use and land coverage is additive to my understanding spatial data.

**Contextualization:** Land use/land cover is a common geographical dataset that is used in a wide variety of fields that use GIS. Land cover is what is the coverage of a given geographical area, such as forests, deserts, croplands, and urban areas. With remote sensing from satellites, there is a high degree of accuracy of land coverage datasets. There is a standard classification system with 20 different classifications that is used to describe what is the land coverage in a given spatial area. Land coverage is what is on the land; land use is the utilization of what is on the land. Land use is a representative of environmental factors of the land use, but it also represents social factors within land use, such as economic opportunities and social structural features. This is why land use studies are important to geohealth analysis, where there has been some health research involving land use, but there is much more to be examined. From my literature review, it seems that health studies utilizing land use/land coverage have been more predominant outside the U.S. this is particular to me because the land is such a predominate feature in the U.S. with its diversity of land coverage and hence the diversity of land use. Therefore, I find that there is a gap in geohealth research in the U.S. with land use analysis.

**Source Four:** Zaldo-Aubanell, Q., Serra, I., Sardanyés, J., Alsedà, L., & Maneja, R. (2021). Reviewing the reliability of Land Use and Land Cover data in studies relating human health to the environment. *Environmental Research*, *194*, 110578. <https://doi.org/10.1016/j.envres.2020.110578>

**Comment 4:**

**Quote/Paraphrase:** “The aim of the present work is to analyse the reliability of LULC data as a suitable describer of the environment in studies relating human health to the environment. Particularly, we endeavour to examine several methodologies using LULC data, as well as the study designs and analytical methods that have been commonly performed. As a secondary aim, we also compile some general characteristics (number of assessed individuals, country of origin, individual and area level variables ...) and human health measurements of the articles reviewed. Finally, we propose alternative ways to analyse the relation between LULC and human health with a view to shedding light on cause effect associations.” (Zaldo-Aubanell, et al., 2021, p.2)

**Essential Element:** Post-hoc Analysis

**Additive/Variant Analysis:** The use of land use/land coverage in healthcare research is additive to my understanding of analyzing and understanding the geographical dimensions of health and illness.

**Contextualization:** This quote summarizes the benefits of using land use/land coverage (LULC) in health analysis, along with the different considerations researchers need to be aware of in working with land use/land coverage data. The idea of using land use data in research is to overlay demographic, social, economic, or health over land use to examine the correlations of different land use and non-spatial data. In conducting geo analysis, the standard statistical procedures remain, but with spatial data, there are additional statistical considerations and factors the researchers must deal with, such as the correlation of geographical areas that are near to another area. From my literature review, it does not appear that a correlation study has been done with land use and illnesses, given the different land uses and different illnesses. However, there has been research done using land use and particular illnesses/diseases, and I am learning the different methods researchers are using and what may be the best method to use to measure the correlation between different land use and different illnesses.

**Comment 5:**

**Quote/Paraphrase:** “Particularly, from a geographic and territorial point of view, Electronic Health Records (EHR) seem to be an excellent source of health data creating good synergy with LULC data since they both serve a new trend in studies which are focused on massive information and broader scopes, highlighting territorial realities.” (Zaldo-Aubanell, et al., 2021, p.8)

**Essential Element:** Independent Variables

**Additive/Variant Analysis:** The use of EHR data in healthcare research as a prime source of health data is additive to my understanding.

**Contextualization:** There may be various sources of health data, from surveys to surveillance methods, but I think health data directly come out of the EHR (electronic health records). Where providers input health data about the patient’s visit, such as the reason for the visit, diagnosis, and procedures, in addition to the billing information. The EHR data could then be associated with the patient’s or the facility's zip code, which then provides a geo-location that could be used to associate health information with geography. Electronic health records are a protected source of data, and not many people have direct access to EHR data. However, there are sources that collect, de-identify, and summarize EHR data, such as government agencies such as the Health Resources and Services Administration and the Center for Medicare and Medicaid Services. Often, the available secondary data from these organizations are at a higher geographic level, such as the county level. The intent of my research is to obtain diagnosis EHR data at the zip code level. Which may require a request for that level of data from the agency for research purposes with a data use agreement in place.

**Source Five:** Beltrán, S., Arenas, D. J., López-Hinojosa, I. J., Tung, E. L., & Cronholm, P. F. (2021). Associations of Race, Insurance, and Zip Code-Level Income with Nonadherence Diagnoses in Primary and Specialty Diabetes Care. *The Journal of the American Board of Family Medicine*, *34*(5), 891–897. <https://doi.org/10.3122/jabfm.2021.05.200639>

**Comment 6:**

**Quote/Paraphrase:** “Patient zip codes were also retrieved from the EHR and paired with publicly available data on median household income of Philadelphia zip codes. The effect of categorical variables on the probability

of having at least 1 nonadherence diagnosis was calculated using risk ratios (RR). Logistic regression models were used to adjust for age, HbA1c, body mass index (BMI), sex, race, ethnicity, and insurance b eta (b ) coefficients from a generalized linear model were used for numeric variables. Significance was estimated using 95% confidence intervals (95% CI). Multiple hypothesis corrections were not applied to prevent low statistical power.21 Document S1 offers additional methodological details.” (Beltrán et al., 2021, p. 892)

**Essential Element:** Statistical Procedures

**Additive/Variant Analysis:** The use of logistic regression in spatial analysis is additive to my understanding.

**Contextualization:** Zip Code Tabulation Areas (ZCTA) is a geographic model the U.S. Census Bureau developed to yield spatial feature data at a lower level than counties. Since Zip Codes are not necessarily geographic areas, they have a representation of geographic areas where the postal service delivers, the ZTCA are not exactly equal to Zip Codes, but one could use zip codes to attribute to a given ZTCA, thus allowing non-spatial data to be joined to a geographic area, such as EHR diagnoses data by zip code, which could then be placed on a map that could be analyzed in relation to spatial data such as land use. To analyze spatial relationships, correlational procedures are done, and like general correlational analysis, depending on your data informs what statistical procedure should be done. In this article, the researchers used logistic regression models to analyze the correlations between their variables. Logistic regression provides probability or odds that there is a given relationship between variables by way of the odd an event will occur given the input from the independent variables. Logistic regression may be a method I could use to examine the relationship between land use and illnesses.

**Source Six:** Akindahunsi, T., Olulaja, O., Ajayi, O., Onyenegecha, I. P., Hanson, U., & Fadojutimi, B. (2024). Analytical tools in diseases epidemiology and surveillance: A review of literature. *International Journal of Applied Research*, *10*(9), 155–161. <https://doi.org/10.22271/allresearch.2024.v10.i9c.12018>

**Comment 7:**

**Quote/Paraphrase:** “A significant strand of research in epidemiology explores the application of GIS in disease surveillance. GIS has been extensively reviewed for its ability to track the spatial distribution of diseases. The most common understanding of the term can be summed up as ‘a powerful set of tools for collecting, retrieving at will, transforming and displaying spatial data from the real world’ as proposed by Burrough [20]. Classical geographic maps present a limited set of data, such as roads and cities, however, GIS offers dynamic, digital mapping with integrated data processing and analysis [21]. GIS has the ability to display various map types, including qualitative distribution maps, point-maps, and quantitative maps with proportionate data, such as disease cases, populations at risk, prevalence, and infection intensity [22]. It also supports methods like transect sampling and Kriging for predicting values at unsampled points based on known data [23]. GIS visualizes data using vectors (storing points, lines, and areas) or rasters (using adjacent polygons to represent continuous phenomena like land cover and climate)” (Akindahunsi et al., 2024, p. 157)

**Essential Element:** Results Interpretation

**Additive/Variant Analysis:** The use of GIS in healthcare research is additive to my understanding of the importance of understanding health and illness in its spatial context.

**Contextualization:** There are many spatial and statistical analyses one could do within the GIS software through the spatial analysis feature. In order for one to do spatial analysis with health and illness data the data must be able to be geographically associated such as with zip codes. Then the researcher will work within GIS to modify or combine zip codes to a new feature shape to associate on the map. In my case, using EHR data of certain illnesses represented by diagnosis categories and zip codes of the majority of patients served by a given health center becomes the service area of the health center. These zip codes/ZTCA will be transformed into a geographical area. These service areas will then be overlayed on the land use classification, and a process will be used to assign a land use classification to the service area given the majority of the type of land use. Different health centers will have service areas in different land use classifications. Thus, they could then perform a correlation procedure to the illness to the land use classifications. The illness data is a raw data set of patients of a given health center who have a given illness. The raw numbers would not be a good measure for the correlation analysis. Thus, standardization must be performed, which in this case would be a rate of illness given the raw number of patients compared to the patient population of that service area. Then a correlation study can be done analyzing the rate of a given illness to a given classification. The exact statistical procedure to be used is being reviewed given the literature review findings and also are dependent on the assumption of the data.

**Source Seven:** Das, K., Das Chatterjee, N., Jana, D., & Bhattacharya, R. K. (2023). Application of land-use regression model with regularization algorithm to assess PM2.5 and PM10 concentration and health risk in Kolkata Metropolitan. *Urban Climate*, *49*, 101473. <https://doi.org/10.1016/j.uclim.2023.101473>

**Comment 8:**

**Quote/Paraphrase:** “Several geospatial methods including land use regression (LUR) and machine learning algorithms have been used in inter-urban, intra-urban, and even country-level studies to predict air pollution concentrations (Fan et al., 2020; Son et al., 2018; Wong et al., 2021). Four key factors i.e. traffic related data, meteorological variables, land use type, and location of monitoring are considered for preparing the variable factor layers to develop LUR model in order to better understand the pollutant concentration variation (Son et al., 2018; Mo et al., 2021). Few research works have now used the regularization or shrinkage techniques to show a significant improvement in model performance for estimating air pollution concentrations (Ren et al., 2020). On the other hand, some studies revealed similar performance in different statistical algorithms for predicting the annual average of air pollutants in a large number of the training dataset (Chen et al., 2019) yet, both studies are not in favor to make any general recommendation for any particular algorithms to be the best. For this reason, the robustness of the model relies heavily on the scale, data behavior, resolution and study setting (Chen et al., 2019; Ren et al., 2020). Hence, linear stepwise regression methods and shrinkage algorithms both are relevant in order to explore how PM concentrations much more vary at spatial heterogeneity.” (Das, et al, 2024, p. 2)

**Essential Element:** Results Interpretation

**Additive/Variant Analysis:** The use of regression in spatial analysis is additive to my understanding of spatial analysis.

**Contextualization:** This article discusses the use of land use regression (LUR), which is a form of the multiple regression model that identifies the relationship between land use and environmental factors such as pollutants. For the articles I have read on land use regression they are all analyzing types of pollutants. However, it seems that given the use of land use in my study, I am looking into how I may be able to use LUR in examining the relationship between land use and illness. The data I will be working with may be a bit different than pollutant data. Nevertheless, it seems that I may be able to utilize LUR or perhaps a modified version such as logistic land use regression (LLUR). I think what I am going have to do is to use a sample set of illness and land use data and perform the various correlation/regression models I am reading about to see which one would be appropriate for my research given that I have not found an article that conducted a similar correlation study of land use and illness.

**Source Eight:** Tokey, A. I. (2021). Spatial association of mobility and COVID-19 infection rate in the USA: A county-level study using mobile phone location data. *Journal of Transport & Health*, *22*, 101135. <https://doi.org/10.1016/j.jth.2021.101135>

**Comment 9:**

**Quote/Paraphrase:** “We have fitted the models discussed in the methodology section for each of the four phases of Coronavirus. Each variable’s daily records were aggregated into the average for each phase and were used as throughputs of the models. For gauging the relationship of mobility and COVID severity, OLS and SEM tell the global associations, while GWR helps to understand the spatially varying local impacts. In Table 3 to Table 6, we presented the model results with coefficients, standard error (for OLS and SEM), median local GWR estimate, and the percentages of counties with positive GWR estimates. Although all the models have VIF values less than 7, we did not include WFH in GWR for local multicollinearity. The fit of the models (i.e., adjusted R2, or Adj. R2) and the spatial autocorrelation test results are also reported in tables. A significant Moran’s I value more than the expected index (I = 0.0003) shows significant clustering of residual while a Moran’s I significantly less than expected index denotes dispersed pattern. Insignificant Moran’s I shows the randomness of residual. We have also analyzed the mediation effect with several auxiliary regressions, which we did not report in tables. In these regressions, the mediatory variables from the pool of independent variables were identified. One variable from independent variables can be called mediatory variables if the mobility variables can significantly explain it and are significant when dependent variables are regressed without it (Preacher and Kelley, 2011). In other words, in multivariate regression, a mobility, variable, say OCT, can have a different relation to dependent variables depending on the presence of an independent variable, say PSI. If OCT can regress PSI significantly, PSI can be regarded as a mediatory variable that mediated the effect of OCT on the dependent variable.” (Tokey, 2021, p. 12)

**Essential Element:** Statistical Procedures

**Additive/Variant Analysis:** The use of correlation statistics in spatial analysis and particular spatial regression procedures is additive to my understanding.

**Contextualization:** From this article, the researchers were examining the performance of different regression statistics in their research. While doing so they bring up many of the statistics that they examine which helps inform me what I should be using and examining in my research given the use of similar regression procedures. The authors examined different regression models, such as OLS (Ordinary Least Square), SEM (Structural Equation Modeling), GWR (Geographically Weighted Regression). Along with the output of the regression models they examined the statistics of; coefficients, standard error, VIF (Variance Inflation Factor), multicollinearity, fit of the models, spatial autocorrelation, and Moran’s I. Some of the regression models and regression statistics are particular to spatial analysis such as Moran’s I, but many are used in not spatial regression analysis such as multicollinearity. I am working on becoming familiar with these various regression models in applying them to spatial analysis.

**Source Nine:** Embury, J., Tsou, M.-H., Nara, A., & Oren, E. (2022). A Spatio-Demographic Perspective on the Role of Social Determinants of Health and Chronic Disease in Determining a Population’s Vulnerability to COVID-19. *Preventing Chronic Disease*, *19*, 210414. <https://doi.org/10.5888/pcd19.210414>

**Comment 10:**

**Quote/Paraphrase:** “The COVID-19 case rates in the 5 stages, diabetes deaths, and hypertensive disease hospitalizations exhibited significant positive spatial autocorrelation (Global Moran I) indicating that rates geographically nearby tend to be similar. Of note, the strength of spatial autocorrelation decreased for COVID-19 case rates during pandemic Stage 1 (I = 0.561, z = 6.548, P ≤ .001) and Stage 2 (I = 0.485, z = 5.486, P ≤ .001) before stabilizing during Stages 3 through 5 (0.304 ≤ I ≤0.347, 3.511 ≤ z ≤3.934, P ≤ .001). Spatial autocorrelation results for 2017 hypertensive disease hospitalization rates (I = 0.413, z = 4.912, P ≤ .001) were greater than those for the 2017 diabetes death rates (I = 0.345, z = 3.092, P = .002). Subsequent spatial analysis determined the accuracy with which the rate of diabetes deaths or hypertensive disease hospitalizations could be independently used to model COVID-19 case rates, thereby avoiding the multicollinearity problems inherent in the selected socioeconomic variables.” (Embury et al , 2022, p. 5)

**Essential Element:** Statistical Procedures

**Additive/Variant Analysis:** This quote informs about Global Moran I and its connection to spatial regression is additive to my understanding about spatial analysis.

**Contextualization:** The statistics Getis-Ord Gi and Global Moran’s I are used to measuring spatial autoregression, which is the concept that given the spatial aspect of a variable or phenomena, just due to the spatial closeness of features could be an influence of spatial autocorrelation.Spatial Autocorrelation statistics produces; Moran's I index, expected index, variance, z-score, and p-values. These statistics are used to evaluate the impact of spatial autocorrelation, correlation based on closeness within space. The issue with spatial autocorrelation is that variables may be correlated by spatial closeness, but not related to variables. For example, crime statistics in given neighborhoods may be correlated based on the neighborhoods closeness, but these neighborhoods may have different poverty levels or number of criminals, so uncorrelated by different variables. Given my intent in researching land use and illness, I will be examining the spatial autocorrelation along with the correlation between variables land use and illness.

**Source Ten:** Blanford, J. (2024). *Geographic information, geospatial technologies and spatial data science for health*. CRC Press.

**Comment 11:**

**Quote/Paraphrase:** “GeoHealth is about understanding where and when diseases occur in space and time, why they may be prevalent (risk factors), what the relationships are between different risk factors, who may be affected (population and host(s)) and how to respond (intervention).” (Blanford, 2024, p. 4)

**Essential Element:** Results Interpretation

**Additive/Variant Analysis:** The concept of GeoHealth and its benefit to understanding health and illness phenomena is additive to my understanding of health and illness research.

**Contextualization:** This quote I think is a nice summary of the importance of using geohealth as part of healthcare research. Geohealth brings together the different aspects of health and illness, from population to risk factors, to environmental influences, and insight into the implementation of interventions. In my dissertation research, my aim is to help bring about understanding of how land use impacts illness, so that interventions could be introduced within the given land use area and that health facilities may incorporate land use factors into their healthcare. My aim is to help bring about positive social change through geosocialhealth research by using the statistical tools of spatial analysis to come to a deeper understanding of health and illness phenomena.

**Comment 12:**

**Quote/Paraphrase:** “All of the examples are really about communication and informing the public and health officials where something is happening and, in some cases when something is happening. In 1854, John Snow used a map to identify where the potential source of cholera infections may have been, the Broad Street water pump. Using this information he took action to try to reduce further cholera infections by removing the water pump handle and access to the contaminated water source.” (Blanford, 2024, p. 22)

**Essential Element:** Statistical Procedures

**Additive/Variant Analysis:** Using maps and geographic products to visualize and communicate health and illness concerns is additive to my understanding of the role of geohealth and geohealthstatistics.

**Contextualization:** From my experience of incorporating GIS products into healthcare analytics the consumers of the products found it very beneficial to understanding the health and illness issue being displayed. It was a very good communication tool to inform of the larger context surrounding the given phenomena, because it brought together the different aspects of illness, population, area, and factors. Having this visually communicated provides the readers with a logical comprehension of the issue and could look a possible interventions and solutions. A prime historical example has been with Dr. John Snow’s 1854 cholera outbreak spatial analysis. Even though this was a simple map of plotting where people who were getting cholera lived, it shows the communication value of a geographic product, immediately detecting a pattern of the outbreak with human visualization. And thus was able to implement an intervention to stop the outbreak. With today’s computing power, geostatistics could be performed providing us with advanced understanding and insight into health and illness issues. Thus my interest in pursuing social research using statistical and geostatistical tools to understand deeper of health and illness concerns and more over being able to help make changes to the health of people by way of social research.

**Works Cited**

Akindahunsi, T., Olulaja, O., Ajayi, O., Onyenegecha, I. P., Hanson, U., & Fadojutimi, B. (2024). Analytical tools in diseases epidemiology and surveillance: A review of literature. *International Journal of Applied Research*, *10*(9), 155–161. <https://doi.org/10.22271/allresearch.2024.v10.i9c.12018>

Beltrán, S., Arenas, D. J., López-Hinojosa, I. J., Tung, E. L., & Cronholm, P. F. (2021). Associations of Race, Insurance, and Zip Code-Level Income with Nonadherence Diagnoses in Primary and Specialty Diabetes Care. *The Journal of the American Board of Family Medicine*, *34*(5), 891–897. <https://doi.org/10.3122/jabfm.2021.05.200639>

Blanford, J. (2024). *Geographic information, geospatial technologies and spatial data science for health*. CRC Press.

Das, K., Das Chatterjee, N., Jana, D., & Bhattacharya, R. K. (2023). Application of land-use regression model with regularization algorithm to assess PM2.5 and PM10 concentration and health risk in Kolkata Metropolitan. *Urban Climate*, *49*, 101473. <https://doi.org/10.1016/j.uclim.2023.101473>

Embury, J., Tsou, M.-H., Nara, A., & Oren, E. (2022). A Spatio-Demographic Perspective on the Role of Social Determinants of Health and Chronic Disease in Determining a Population’s Vulnerability to COVID-19. *Preventing Chronic Disease*, *19*, 210414. <https://doi.org/10.5888/pcd19.210414>

Kent, J. T., & Mardia, K. V. (2022). *Spatial analysis* (Vol. 72). John Wiley & Sons.

Nedd, R., Light, K., Owens, M., James, N., Johnson, E., & Anandhi, A. (2021). A Synthesis of Land Use/Land Cover Studies: Definitions, Classification Systems, Meta-Studies, Challenges and Knowledge Gaps on a Global Landscape. *Land*, *10*(9), 994. <https://doi.org/10.3390/land10090994>

Preye Winston Biu, Chinedu Nnamdi Nwasike, Olawe Alaba Tula, Chinedu Alex Ezeigweneme, & Joachim Osheyor Gidiagba. (2024). A review of GIS applications in public health surveillance. *World Journal of Advanced Research and Reviews*, *21*(1), 030–039. <https://doi.org/10.30574/wjarr.2024.21.1.2684>

Tokey, A. I. (2021). Spatial association of mobility and COVID-19 infection rate in the USA: A county-level study using mobile phone location data. *Journal of Transport & Health*, *22*, 101135. <https://doi.org/10.1016/j.jth.2021.101135>

Zaldo-Aubanell, Q., Serra, I., Sardanyés, J., Alsedà, L., & Maneja, R. (2021). Reviewing the reliability of Land Use and Land Cover data in studies relating human health to the environment. *Environmental Research*, *194*, 110578. <https://doi.org/10.1016/j.envres.2020.110578>