BAREKENG: Journal of Mathematics and Its Applications
December 2025 Volume 19 Issue 4 Page 2633–2646
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss4pp2633-2646

CLUSTERING WITH SKATER METHODS AND UTILIZATION OF LISA ON UNEMPLOYMENT RATE

Naufal Shela Abdila[1*], **Rahma Fitriani**[2], **Muhamad Liswansyah Pratama**[3]

[1,2] *Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya*
Jln. Veteran, Malang, 65145, Indonesia

[3]*Department of Data Sciences, Faculty of Computer Sciences, UPN "VETERAN" Jawa Timur*
Jln. Rungkut Madya, Surabaya, 60294, Indonesia
*Corresponding author's e-mail: ** ***naufalshela24@ub.ac.id***

Article History:

Received: 6th February 2025
Revised: 21st April 2025
Accepted: 22nd May 2025
Available online: 1st September 2025

Keywords:

LISA;
SKATER;
Unemployment Rate.

ABSTRACT

Spatial cluster analysis is an analysis used to identify a spatial pattern or geographical grouping of data. One method that can be used in spatial cluster analysis is Spatial Cluster Analysis by Tree Edge Removal (SKATER). This research aims to analyze the spatial pattern of the Unemployment Rate in East Java by utilizing the SKATER method. The clustering results are then used to create a weighting matrix, which is used to find local spatial autocorrelation values using the Local Indicators of Spatial Association (LISA) index. The data is taken from BPS East Java with variables including unemployment rate, education level, minimum wage, Human Development Index, and population density. The results show that this approach is able to identify significant local spatial patterns. However, the selection of the number of clusters and input variables proved to be very influential on the results, so care needs to be taken.

How to cite this article:

N. S. Abdila, R. Fitriani, and M. L. Pratama, "CLUSTERING WITH SKATER METHODS AND UTILIZATION OF LISA ON UNEMPLOYMENT RATE," *BAREKENG: J. Math. & App.*, vol. 19, iss. 4, pp. 2633-2646, December, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*
Research Article · Open Access

1. INTRODUCTION

The unemployment rate is an important indicator that reflects the socio-economic conditions of a region. Unemployment is a situation in which a person who wants to work is not able to find a suitable job and is willing to work at the market rate of pay [1]. High unemployment reflects an imbalance between supply and demand for labor, which in the end can hamper economic growth and increase poverty and social inequality in a region. The mismatch between the qualifications of educational graduates and the needs of the job market also exacerbates the problem of unemployment. Graduates from various levels of education often struggle to find jobs that match their skills due to a lack of practical training. In East Java Province, with a population of more than 40 million people, the problem of unemployment has become a crucial issue that must be addressed immediately [2]. Even though economic growth in East Java is relatively stable, the problem of unemployment remains a significant challenge that requires serious attention from the local government and various stakeholders [3]. Open unemployment in East Java Province is a problem that must be addressed immediately by the local government. This challenge is not unique to East Java but across all regions of Indonesia. Therefore, intensive intervention is required to improve the community's access to a decent standard of living and, consequently, to reduce poverty rates to the greatest extent possible [4].

The unemployment rate is the percentage of the number of unemployed to the total workforce. A high percentage of the unemployment rate indicates that there is a large workforce that is not being absorbed. According to data from the Central Bureau of Statistics Agency [5], in February 2023, the unemployment rate in East Java was 4.33% and will increase to 4.88% in August 2023. This means that for a period of one year, East Java Province experienced an increase in the percentage of unemployment rate compared to several provinces in East Java, such as West Java, Central Java, and Banten, which experienced a decrease in the percentage of unemployment rates over a period of one year. This percentage shows that although there has been improvement, the unemployment problem is still not completely resolved.

Various efforts that can be made to deal with the problem of unemployment have received attention from researchers and scientists, many of whom have studied the causes of the high unemployment rate in East Java, as research conducted by [6] shows that there is a negative relationship between the human development index variable and education level on the unemployment rate, which means that when these two variables increase, the unemployment rate will decrease. Another research conducted by [7] concluded that the minimum wage variable had a negative and significant effect on the unemployment rate in Banten Province. Apart from that, research conducted by [8] shows that the level of labor force participation and population density affect the unemployment rate.

Based on several previous studies, it can be concluded that the unemployment rate is not only influenced by economic factors, but also by other variables, such as education, social factors, and population density. East Java Province, which consists of 38 districts/cities, has different unemployment, education, social, and population density levels in each region. To see whether these regions have similar characteristics, the spatial clustering method is very appropriate to use, because this method considers the geographical location of the observations [9].

One such method is Spatial Kluster Analysis by Tree Edge Removal (SKATER), which is very suitable for identifying geographic patterns or groups of areas with similar socioeconomic characteristics. This technique is often used in ecological, epidemiological, and economic studies to detect homogeneous regions in spatial data. The SKATER method operates by calculating the dissimilarity between adjacent regions based on several variables, then pruning the resulting minimum spanning tree (MST) to form coherent spatial clusters. This method was chosen in this study due to its simplicity and computational efficiency compared to other clustering methods [10], [11], [12].

Spatial analysis and regional clustering have become increasingly important methods for analyzing unemployment problems at the regional level. Several previous studies have applied various approaches to analyze spatial patterns of unemployment rates: (1) The traditional spatial cluster analysis approach has been used to identify geographic groupings of unemployment data but has limitations in analyzing relationships between regions. (2) The SKATER method was then developed as a more sophisticated approach for spatial cluster analysis. This method allows a more detailed analysis of how neighboring regions influence each other. (3) The use of Local Indicators of Spatial Association (LISA) has begun to be applied to measure local spatial autocorrelation and identify significant clusters at the local level. (4) Previous studies have examined variables that influence the unemployment rate, such as education level, minimum wage, and population density. Such as research conducted by [13] which analyzes the influence of population, HDI, and minimum wage on the unemployment rate in Bangkalan Regency, then research conducted by [14] which analyzes the influence of education, minimum wage, and employment opportunities on unemployment in 2017-2021, and research conducted by [15] which analyzed population density on unemployment rates. However, not many have integrated spatial analysis with these variables comprehensively. Therefore, this study will combine the SKATER method with LISA to provide a deeper understanding of the spatial pattern of unemployment and the factors that influence it.

This research aims to analyze the spatial pattern of the Unemployment Rate in East Java and examine the relationship with other variables such as education level, Minimum Wage, Human Development Index, and population density. Next, the clustering results will be used to see whether there is spatial autocorrelation between regions by utilizing the Local Indicators of Spatial Association (LISA) index to identify cluster patterns or distribution in geographic data. LISA provides insight into how an individual's location is related to its neighbors in terms of certain attributes [16]. The LISA analysis method, developed by Anselin in 1995, performs spatial autocorrelation analysis based on spatial feature location and feature values [17]. LISA is one of the most commonly used methods to characterize differences among continuous values by considering their spatial arrangement and categorizing the values into spatial groups and outliers through significance testing [18]. Through this approach, it is hoped that significant spatial relationships between regions in East Java can be revealed. This research integrates the Spatial Kluster Analysis by Tree Edge Removal (SKATER) method with Local Indicators of Spatial Association (LISA) to provide a more comprehensive approach in identifying unemployment patterns. This integration has not been widely done, thus providing a new dimension in spatial analysis.

This research is making a new contribution by integrating the SKATER and LISA methods for a more comprehensive analysis; considering broader socio-economic variables (level of education, minimum wage, HDI, and population density) focus on the East Java region which has unique geographical and economic characteristics; provides new findings that variations in the unemployment rate and related variables in East Java are not influenced by spatial factors or inter-regional linkages. This research opens a new perspective in understanding the dynamics of regional unemployment and can help in formulating more effective policies to overcome the unemployment problem in East Java.

2. RESEARCH METHODS

2.1 Data

The data used is secondary data taken from BPS East Java in 2023, namely data regarding the unemployment rate based on regencies/cities in East Java. Several variables used are presented in Table 1.

Table 1. Research Variables

No.	Variables	Definition	Unit
1.	unem	Unemployment Rate	Percent (%)
2.	educ	Education Level	Percent (%)
3.	wage	Minimum Wage	Rupiah (IDR)
4.	HDI	Human Development Index	–
5.	dens	Population Density	Thousand people/km^2

2.2 Data Analysis Methods

The analysis used in this research was carried out using RStudio and Geoda software, with details of the research sequence as follows:

1. Prepare data to be analyzed based on several research variables.
2. Applying the SKATER method:
   a. Create a Minimum Spanning Tree (MST) based on proximity relationships between regions using Prim's algorithm, where MSTs are built iteratively, starting from a single node (points in a graph that represent a particular object or geographic region) and adding edges with the smallest weight that connect existing nodes to nodes that don't yet exist [19]. Prim's Algorithm steps:

i. Randomly select a starting node (v_0) and add it to the MST.

ii. Select the edge (e) with the smallest weight that connects the node (u) in the MST to node (v) outside the MST:

$$\min\big(w(e)\big), \qquad e = (u, v)$$

where:
$u \in MST$: a node that is already included in the current MST
$v \notin MST$: a node that is not yet included in the MST
$w(e)$: weight of edge e

iii. Add node v into the MST and repeat the process until all nodes are connected.

b. Carrying out edge removal aims to cut connections between regions that are considered the most different. The edge-breaking step in the SKATER method is carried out after the MST is formed. To perform edge removal, we will select the edge that has the largest difference between connected regions. This process is often calculated based on dissimilarity, as in Euclidean Distance [20]:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{1}$$

where:
d_{ij}: Euclidean distance between region i and region j
x_i: x-axis coordinate of location i
x_j: x-axis coordinate of location j
y_i: y-axis coordinate of location i
y_j: y-axis coordinate of location j

The edge with the highest dissimilarity will be removed: $\max(d_{ij})$

c. Cluster formation. Once some edges are removed, the still connected regions will form clusters. In a truncated Minimum Spanning Tree, each component is connected to form a cluster.

d. Determining the number of clusters according to the needs. With the SKATER process, it will automatically decide the specified number of clusters produced by MST.

3. Defining the spatial weighting matrix using the grouping results obtained through the second process, applying the SKATER method. In the spatial weighting matrix, the following values will apply.

$$w_{ij} = \begin{cases} 1 & \text{if regencies/cities i and j are in the same cluster} \\ 0 & \text{other} \end{cases}$$

where:
i : observed region
j : neighbors of the observed region
w_{ij} : spatial weight matrix showing the proximity of regions i and j

Next, standardization of the weighting matrix is carried out. The results of the standardization of the weighting matrix will be used to calculate the local index.

4. Calculate the Local Indicators of Spatial Association (I_i) index for each region i, using the formula:

$$I_i = z_i \sum_{j=1}^{n} w_{ij} z_j \tag{2}$$

z is the standard value (z-score) of the variable studied in the region i [21], calculated by:

$$z_i = \frac{x_i - \bar{x}}{s} \tag{3}$$

where:
i : observed region
j : neighbors of the observed region
x_i : variable value in region i
$\bar{x}$: the average of all variable values
s : standard deviation of the variable value

w_{ij} : spatial weight matrix showing the proximity of regions i and j
z_i : standard value of the variable from the i-th region

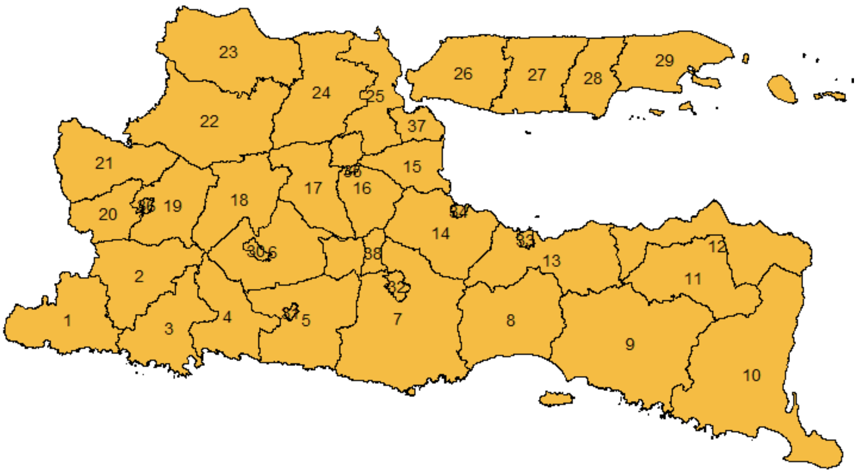
If you get a positive value from I_i, it shows that region i has variable values that are similar to neighboring regions (potential High-High or Low-Low clusters), whereas if you get a negative value from I_i, it shows that region i has different variable values from neighboring regions (potential outliers, for example, High-Low or Low-High).

After calculating LISA for all regions, it can then be classified into four classifications according to spatial relationships based on the LISA values:

- **High-High:** Regions with high variable values are surrounded by regions with high values (positive clusters).
- **Low-Low:** Regions with low variable values are surrounded by regions with low values (negative clusters).
- **High-Low:** Regions with high variable values are surrounded by regions with low values (outliers).
- **Low-High:** Regions with low variable values are surrounded by regions with high values (outliers).

3. RESULTS AND DISCUSSION

East Java Province consists of 38 regencies/cities, where each region has different and unique characteristics. **Figure 1** shows that East Java Province is divided into two large islands, namely Madura Island, which includes the regencies of Bangkalan, Sampang, Pamekasan, and Sumenep Regencies, and the other regencies are on the island of Java.



No	Regency/City	No	Regency/City
1	Pacitan Regency	20	Magetan Regency
2	Ponorogo Regency	21	Ngawi Regency
3	Trenggalek Regency	22	Bojonegoro Regency
4	Tulungagung Regency	23	Tuban Regency
5	Blitar Regency	24	Lamongan Regency
6	Kediri Regency	25	Gresik Regency
7	Malang Regency	26	Bangkalan Regency
8	Lumajang Regency	27	Sampang Regency
9	Jember Regency	28	Pamekasan Regency
10	Banyuwangi Regency	29	Sumenep Regency
11	Bondowoso Regency	30	Kediri City
12	Situbondo Regency	31	Blitar City
13	Probolinggo Regency	32	Malang City
14	Pasuruan Regency	33	Probolinggo City
15	Sidoarjo Regency	34	Pasuruan City
16	Mojokerto Regency	35	Mojokerto City
17	Jombang Regency	36	Madiun City
18	Nganjuk Regency	37	Surabaya City
19	Madiun Regency	38	Batu City

Figure 1. Map of 38 Regencies/Cities in East Java
RStudio Version 4.3.0

This results in problems when forming a Minimum Spanning Tree in the R Studio Software. Therefore, the "*connect_subgraphs*" function from the "*bigDM*" package **[25]-[27]** is used with the aim of forming a line between Surabaya City and Bangkalan Regency, as seen in **Figure 2**.

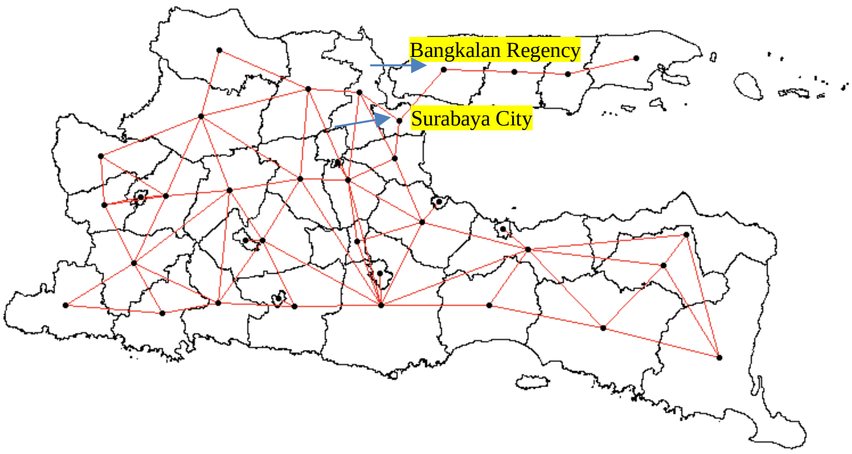


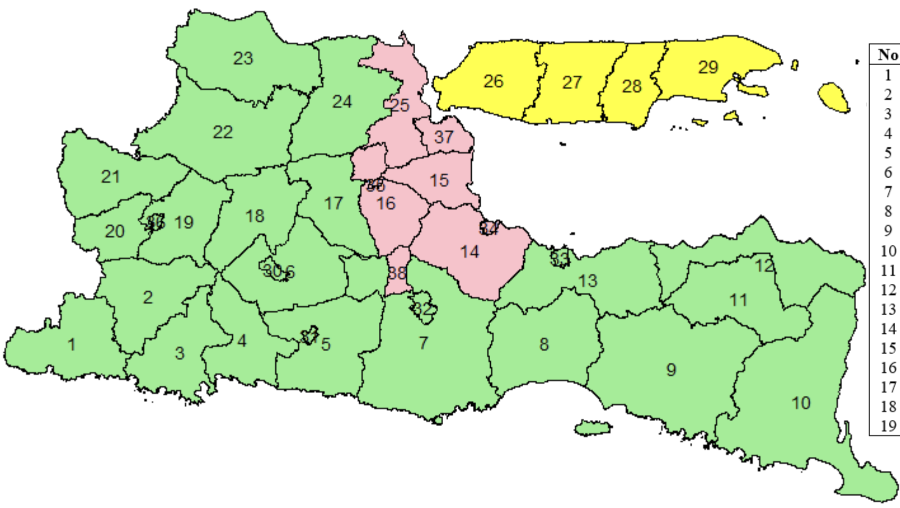
Figure 2. Formation of Minimum Spanning Tree
RStudio Version 4.3.0

Figure 2 shows that it consists of nodes and edges that connect various regions in East Java. Nodes (black dots) represent regencies/cities in East Java, and edges (red lines) connect these nodes, indicating a relationship or connection between these regions. The formation of nodes and edges is then used to visualize spatial relationships in the context of clustering. The minimum spanning tree allows the formation of clusters based on proximity or real spatial relationships, rather than administrative divisions. In addition, MST is also often used in spatial analysis because it respects the geographic structure and spatial distribution of the data, making it a very useful tool in this kind of spatial analysis.

Table 2. Characteristics of Each Cluster based on Average Value

Cluster	*unem*	*educ*	*wage*	*HDI*	*dens*
Cluster 1 (yellow)	3.09	18.18	2,144,315.27	66.61	891.19
Cluster 2 (green)	4.55	26.79	2,402,241.25	73.43	1,582.71
Cluster 3 (pink)	5.83	35.96	3,920,708.63	78.03	3,719.92

Table 2 shows that cluster 1 (yellow color in **Figure 3**) includes Bangkalan Regency, Sampang Regency, Pamekasan Regency, and Sumenep Regency, all of which are located on Madura Island. This cluster has the lowest average values for unemployment rate, education level, minimum wage, HDI, and population density. Cluster 2 (green color in **Figure 3**) consists of regencies and cities with the moderate average values in unemployment rate, education level, minimum wage, HDI, and population density among all clusters. Meanwhile, cluster 3 (pink color in **Figure 3**) consists of Pasuruan Regency, Sidoarjo Regency, Mojokerto Regency, Gresik Regency, Pasuruan City, Mojokerto City, Surabaya City, and Batu City. This cluster has the highest unemployment rate, education level, minimum wage, HDI, and population density compared to the other clusters.



No	Regency/City	No	Regency/City
1	Pacitan Regency	20	Magetan Regency
2	Ponorogo Regency	21	Ngawi Regency
3	Trenggalek Regency	22	Bojonegoro Regency
4	Tulungagung Regency	23	Tuban Regency
5	Blitar Regency	24	Lamongan Regency
6	Kediri Regency	25	Gresik Regency
7	Malang Regency	26	Bangkalan Regency
8	Lumajang Regency	27	Sampang Regency
9	Jember Regency	28	Pamekasan Regency
10	Banyuwangi Regency	29	Sumenep Regency
11	Bondowoso Regency	30	Kediri City
12	Situbondo Regency	31	Blitar City
13	Probolinggo Regency	32	Malang City
14	Pasuruan Regency	33	Probolinggo City
15	Sidoarjo Regency	34	Pasuruan City
16	Mojokerto Regency	35	Mojokerto City
17	Jombang Regency	36	Madiun City
18	Nganjuk Regency	37	Surabaya City
19	Madiun Regency	38	Batu City

Figure 3. Results of the Formation of 3 Clusters
RStudio Version 4.3.0

The clusters formed using the SKATER method are 3 groups, as shown in **Figure 3**. The accuracy level uses the Silhouette Score. This result shows the optimal accuracy value, because if tried for another number of clusters, the accuracy level is smaller. Therefore, these three clusters were chosen with the results shown in **Table 3**.

Table 3. Accuracy of the Clusters

Cluster	Number of Members	Silhouette Score
Cluster 1 (yellow)	4	86.85%
Cluster 2 (green)	26	44.15%
Cluster 3 (pink)	8	23.30%

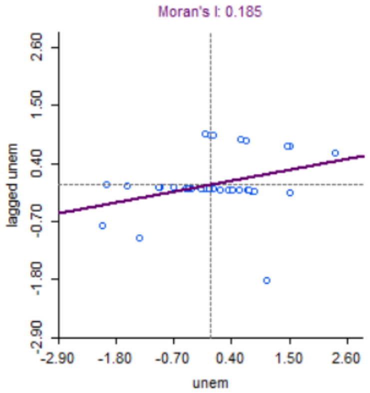
This result is then used to form a spatial weighting matrix, which will be given a value of 1 if the region is in the same cluster, and will be given a value of 0 if the region is not in the same cluster. For example, Malang City with region code number 32 and Bojonegoro Regency with region code number 22 are in the same cluster (green color), so the weighting $w_{32.22} = 1$ and otherwise, namely the weighting $w_{22.32}$ also has a value of 1. Another example, such as Surabaya City (pink color) with region code number 37 and Sampang Regency (yellow color) with region code number 27, are in two different clusters, then the weighting is $w_{37.27} = 0$ and otherwise, namely, the weighting is $w_{27.37} = 0$.

The results of forming a spatial weighting matrix are used to see whether there is local spatial autocorrelation, namely to measure the pattern of relationships between observation locations. If the i-th observation location has similar characteristics to its neighboring locations, then there is positive spatial autocorrelation, whereas if the i-th observation location is not similar to its neighboring locations, then there is negative spatial autocorrelation.

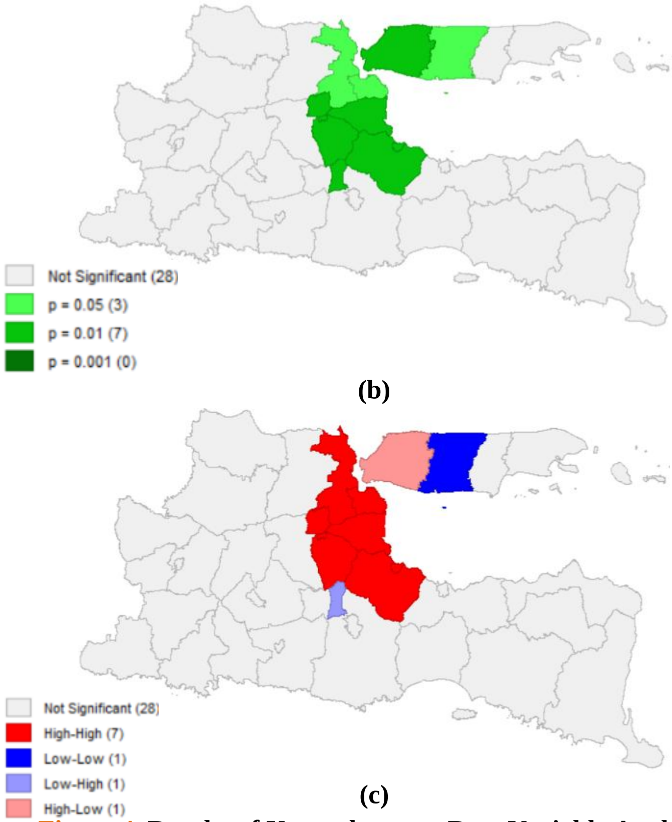
Moran Scatterplot is a visualization tool used to display the relationship between the value of a variable in one location and the average value of the same variable in neighboring locations (lagged values). This plot provides a visual interpretation of the Moran Index and helps identify spatial patterns in the data. **[22]** explains that the Moran Scatterplot consists of four quadrants, which represent types of spatial relationships:

- Quadrant I (High-High): locations with high values are surrounded by locations with high values
- Quadrant II (Low-High): locations with low values are surrounded by locations with high values
- Quadrant III (Low-Low): locations with low values are surrounded by locations with low values
- Quadrant IV (High-Low): locations with high values are surrounded by locations with low values

A study from **[23]** uses Moran Scatterplot to analyze spatial autocorrelation in economic activity. This research shows that the slope of the Moran Scatterplot is equivalent to the value of the Moran's I statistic; points far from the center of the plot represent spatial outliers; and grouping of points in certain quadrants indicates the existence of a spatial cluster.



(a)



(b)

(c)

Figure 4. Results of Unemployment Rate Variable Analysis
GeoDa Version 1.22

A frequently used index in spatial autocorrelation is Moran's I. The global spatial statistic Moran's I measures the global spatial correlation based on different geographic locations and associated feature values [24]. Figure 4 (a) shows that Moran's I value is 0.185, which means that there is a positive but very weak spatial correlation, or in other words, there is a tendency that regions with similar unemployment rate variables (high or low) tend to be close together, but the relationship is not very strong. Figure 4 (b) shows that there are only 10 regencies/cities that are significant, either with a p-value of 5% or with a p-value of 10% for the unemployment rate variable, where the smaller the p-value, the stronger the evidence that the spatial relationship does not happen randomly. Figure 4 (c) shows the 10 regions that are significant with a p-value of 5% or with a p-value of 10%. Seven regions are colored red **(High-High)**, meaning that these 7 regions have high unemployment rates, which are followed by regions that also have high unemployment rates. There is 1 region in blue **(Low-Low)**, namely Sampang Regency, which has a low unemployment rate, and in the surrounding locations the unemployment rate is also low. There is 1 region colored light blue **(Low-High)**, namely Batu City, which has a low unemployment rate but is surrounded by regions with a high unemployment rate. There is 1 region colored pink **(High-Low)**, namely Bangkalan Regency, which has a high unemployment rate and is surrounded by regions with a low unemployment rate.

Most of the areas classified in the High-High cluster are densely populated urban areas, such as Surabaya City and its surroundings, which face challenges in providing employment. Meanwhile, Low-Low regions such as Sampang Regency tend to have an agriculture-based economic structure that is more stable in terms of employment. Areas belonging to the High-High cluster are geographically close and share similar economic characteristics, such as the dominance of the informal sector and high population density.

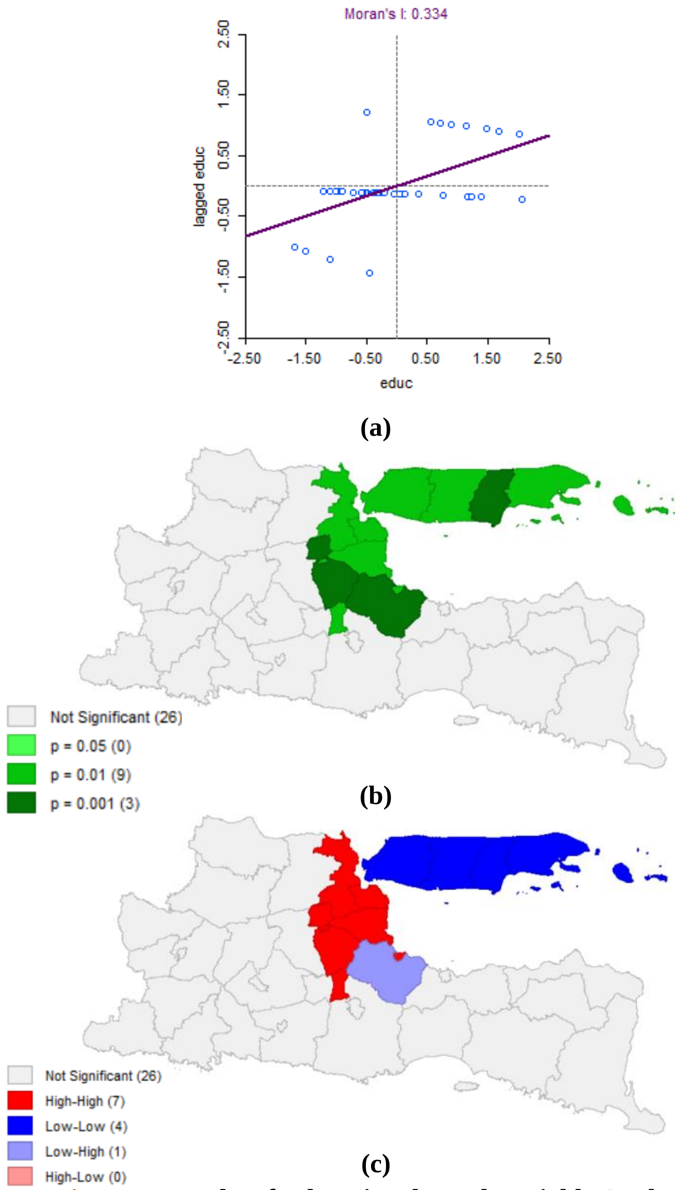


Figure 5. Results of Educational Level Variable Analysis
GeoDa Version 1.22

Figure 5 (a) shows that Moran's I value is 0.334, which means that there is a positive but weak spatial correlation, or in other words, there is a tendency that regions with similar educational level variables (high or low) tend to be close together, but the relationship is not very strong. **Figure 5** (b) shows that there are only 12 regencies/cities that are significant, either with a p-value of 10% or with a p-value of 0.1% for the education level variable. **Figure 5** (c) shows the 12 regions that are significant with a p-value of 10% or with a p-value of 0.1%. Seven regions are colored red **(High-High)**, meaning that these 7 regions have a high level of education, and their surroundings are regions that also have high levels of education. There are 4 regions in blue **(Low-Low)**, namely 4 regencies on Madura Island, which have low levels of education, and the surrounding regions also have low levels of education. There is 1 region colored light blue **(Low-High)**, namely Pasuruan Regency, which has a low level of education but is surrounded by regions with a high level of education.

Areas classified as High-High, such as Surabaya City and Malang City, are known as education centers in East Java with many universities and quality education facilities. In contrast, Low-Low areas on Madura Island generally still face limited access to education and learning support facilities. The regions in the Low-Low cluster in Madura are similar in terms of limited education infrastructure and low high school enrollment, leading to low average years of schooling regionally.

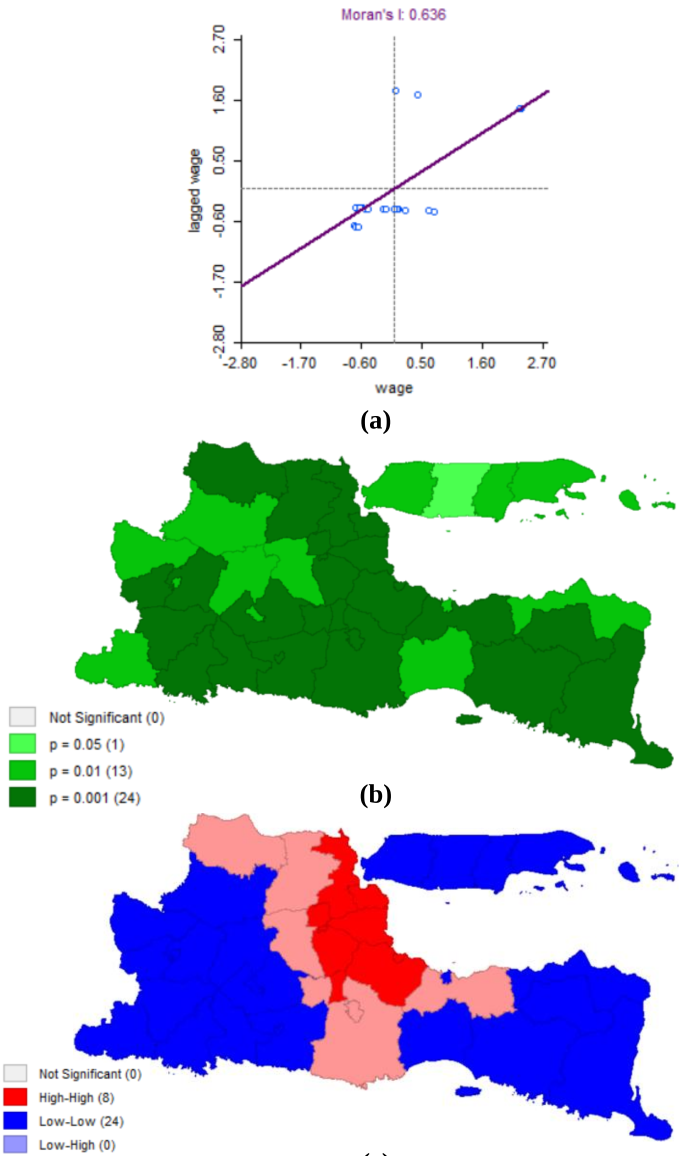


Figure 6. Results of Minimum Wage Variable Analysis
GeoDa Version 1.22

Based on **Figure 6** (a), it can be seen that Moran's I value is 0.636, which means that there is a positive and strong spatial correlation, or in other words, there is a tendency that regions with similar minimum wage variables (high or low) tend to be close together. **Figure 6** (b) shows that all regencies/cities in East Java are significant for the minimum wage variable, either with a p-value of 5%, a p-value of 10% or a p-value of 0.1%. **Figure 6** (c) shows there are 8 regions that are colored red **(High-High)**, meaning that these 8 regions have high minimum wages, and the surrounding regions are regions that also have high minimum wages. There are 24 regions in blue **(Low-Low)**, namely 24 regencies that have low minimum wages, and the surrounding regencies also have low minimum wages. There are 6 regions colored pink **(High-Low)**, which have a high minimum wage but are surrounded by regions with a low minimum wage.

High-High regions such as Surabaya, Gresik, and Sidoarjo are large industrial areas with high economic activity, which contributes to the setting of high minimum wage. In contrast, Low-Low regions such as regencies in the Madura region generally have an agrarian or fishing economic base, with a relatively low cost of living. Most areas in the Low-Low cluster are characterized by geographical proximity, limited industrial infrastructure, and low levels of urbanization. In contrast, the High-High cluster is generally located on major trade routes and has good transportation connectivity.

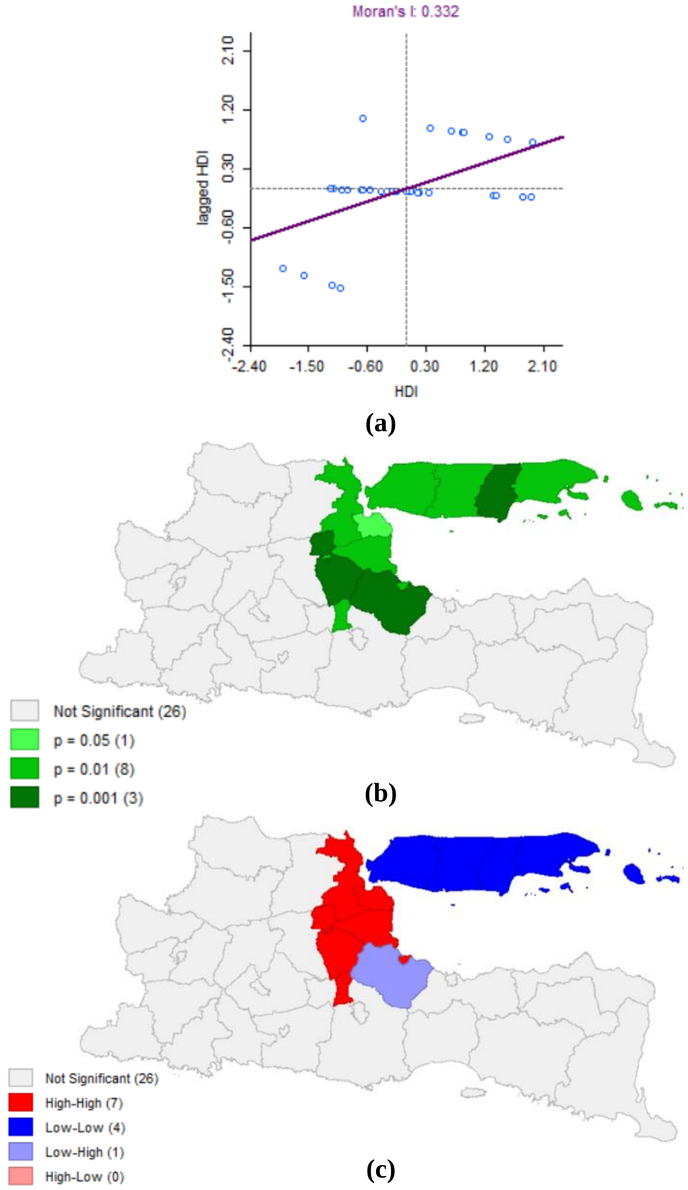


Figure 7. Results of HDI Variable Analysis
GeoDa Version 1.22

Figure 7 (a) shows that Moran's I value is 0.332, which means that there is a positive but weak spatial correlation, or in other words, there is a tendency that regions with similar HDI variables (high or low) tend to be close together, but the relationship is not too strong. Figure 7 (b) shows that there are only 12 regencies/cities that are significant, either with a p-value of 5% or 10% or with a p-value of 0.1% for the HDI variable. Figure 7 (c) shows that of the 12 significant regions, there are 7 regions that are colored red **(High-High)**, meaning that these 7 regions have high HDI and their surroundings are followed by regions that also have high HDI. There are 4 regions in blue **(Low-Low)**, namely 4 regencies on Madura Island, which have low HDI, and the surrounding regions also have low HDI. There is 1 region colored light blue **(Low-High)**, namely Pasuruan Regency, which has a low HDI but is surrounded by regions with a high HDI. The results of this analysis are similar to the results of the analysis for the education level variable.

High-High areas such as Surabaya City and Malang City have excellent education and health infrastructure and high levels of welfare. In contrast, Low-Low areas on Madura Island have limited access to education and health services, as well as lower income levels, which have an impact on low HDI scores. Areas in the High-High cluster are generally located in urban areas with economic centers, higher education facilities, and adequate health services. Meanwhile, the Low-Low cluster generally consists of rural areas with limited infrastructure access and low levels of urbanization. This suggests that spatial clustering in HDI is closely related to the level of development and urbanization of a region.

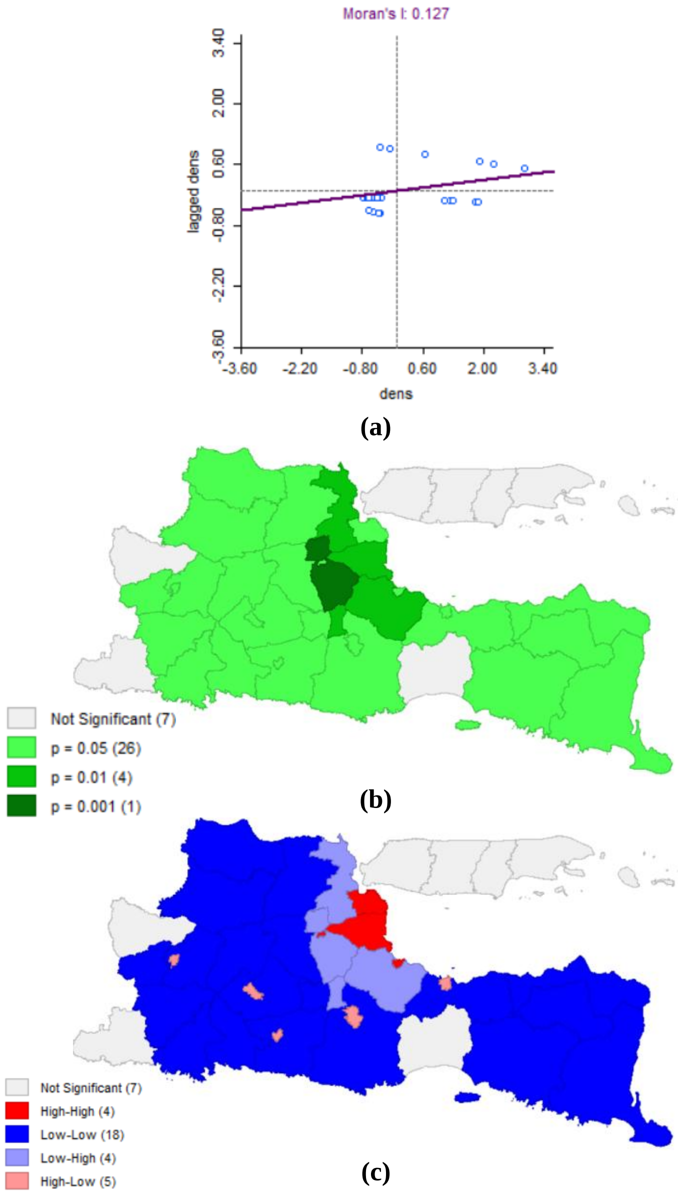


Figure 8. Results of Population Density Variable Analysis
GeoDa Version 1.22

Figure 8 (a) shows that Moran's I value is 0.127, which means that there is a positive but very weak spatial correlation, or in other words, there is a tendency that regions with similar unemployment rate variables (high or low) tend to be close together, but the relationship is not too strong. **Figure 8** (b) shows that 31 regencies/cities are significant, both with a p-value of 5%, 10%, or 0.01%. **Figure 8** (c) shows that of the 31 significant regions, 4 regions are colored red **(High-High)**, namely Surabaya City, Sidoarjo Regency, Mojokerto City, and Pasuruan City, meaning that these 4 regions have high population density, and the surrounding regions also have high population density. This is visible from these regions being industrial regencies and cities, so these four regencies and cities indeed have high population densities, and those in their surroundings are also high. There are 18 regions colored blue **(Low-Low)**, namely regions that have low population density, and the surrounding locations are also low. There are 4 regions colored light blue **(Low-High)**, namely Pasuruan Regency, Batu City, Mojokerto Regency, and Gresik Regency, which have low population density but are surrounded by regions with high population density. There are 5 regions colored pink **(High-Low)**, namely Madiun City, Kediri City, Blitar City, Malang City, Probolinggo City, which have high population density and are surrounded by regions with low population density.

High-density areas such as Surabaya, Sidoarjo, and Mojokerto are centers of industry and trade, with high levels of urbanization and in-migration. In contrast, Low-Low areas are generally agrarian or hilly areas such as Trenggalek or Bondowoso, with limited access to transportation and relatively small populations. High-High clusters tend to be located in urban areas or suburbs of large cities that have adequate access to infrastructure, educational facilities, and the economy. In contrast, the Low-Low cluster is dominated by rural or mountainous areas with scattered settlement patterns and limited development.

4. CONCLUSION

After the analysis was carried out using the SKATER (Spatial Kluster Analysis by Tree Edge Removal) method to form a spatial weighting matrix and continued with the use of the LISA (Local Indicators of Spatial Association) method for the variables of unemployment rate, education level, minimum wage, HDI, and population density to detect local spatial autocorrelation in East Java Province is quite good. However, the selection of the number of clusters to be formed and the variables used in defining the clusters must be done carefully.

AUTHOR CONTRIBUTIONS

Naufal Shela Abdila: Conceptualization, Investigation, Validation, Visualization, Writing - Original Draft, Writing - Review and Editing. Rahma Fitriani: Conceptualization, Formal Analysis, Methodology. Muhamad Liswansyah Pratama: Conceptualization, Data Curation, Investigation, Software. All authors discussed the results and contributed to the final manuscript.

FUNDING STATEMENT

This research did not receive any external funding or financial support from any institution or organization.

ACKNOWLEDGMENT

The authors would like to express their deepest gratitude to all parties who have provided support throughout the research process, and also thank the BICOSDA 2024 committee for organizing the conference and providing the opportunity to present the results of this research.

CONFLICT OF INTEREST

The authors declare no conflicts of interest to report study.

REFERENCES

[1] M. N. Saputri, S. Sifriyani, and W. Wasono, "APPLICATION OF NONPARAMETRIC GEOGRAPHICALLY WEIGHTED REGRESSION METHOD ON OPEN UNEMPLOYMENT RATE DATA IN INDONESIA," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 17, no. 4, pp. 2071–2080, Dec. 2023, doi: https://doi.org/10.30598/barekengvol17iss4pp2071-2080.

[2] Rusman, "DAMPAK PANDEMI COVID-19 TERHADAP ANGKA PENGANGGURAN DI INDONESIA," in *Proceeding Seminar Nasional & Call For Papers*, 2021, pp. 687–693.doi: https://doi.org/10.34123/semnasoffstat.v2021i1.911

[3] Muaidy Yasin, Muhammad Irwan, and Wahyunadi, "ANALISIS PERTUMBUHAN EKONOMI, PENGANGGURAN DAN KEMISKINAN DI KABUPATEN LOMBOK TENGAH," *Ekonobis*, vol. 6, no. 2, pp. 134–164, Sep. 2020.doi: https://doi.org/10.29303/ekonobis.v6i2.52

[4] A. R. D. Prayitno and D. Kusumawardani, "OPEN UNEMPLOYMENT RATE IN THE PROVINCE OF EAST JAVA," *The Winners*, vol. 23, no. 1, pp. 11–18, Jan. 2022, doi: https://doi.org/10.21512/tw.v23i1.7047.

[5] Badan Pusat Statistik, "TINGKAT PENGAGGURAN TERBUKA MENURUT PROVINSI (PERSEN), 2023," Badan Pusat Statistik. Accessed: Oct. 26, 2024. [Online]. Available: https://www.bps.go.id/id/statistics-table/2/NTQzIzI=/tingkat-pengangguran-terbuka--februari-2024.html

[6] O. Siahaan, R. B. S. Pardede, R. Rahim, and D. Desmawan, "ANALISIS PENGARUH INDEKS PEMBANGUNAN MANUSIA DAN PENDIDIKAN TERHADAP TINGKAT PENGANGGURAN TERBUKA DI PROVINSI SUMATERA

UTARA," *JETISH: Journal of Education Technology Information Social Sciences and Health*, vol. 1, no. 1, pp. 55–63, 2022, [Online]. Available: https://data.tempo.co/data/887/tingkat-pengangguran-terbuka-berdasarkan-pendidikan-tertinggi-

[7] U. N. Faizah and N. Woyanti, "ANALISIS PENGARUH PENDIDIKAN, PARTISIPASI KERJA, DAN UPAH MINIMUM TERHADAP PENGANGGURAN DI PROVINSI BANTEN TAHUN 2011-2020," *BISECER (Business Economic Entrepreneurship)*, vol. 6, no. 1, pp. 48–61, Jan. 2023.doi: https://doi.org/10.61689/bisecer.v6i1.386

[8] Nurlaily, L. Aridinanti, and Z. Wildani, "PEMODELAN TINGKAT PENGANGGURAN TERBUKA DI PROVINSI JAMBI MENGGUNAKAN REGRESI DATA PANEL," *Jurnal Sains dan Seni ITS* , vol. 11, no. 1, pp. 157–162, 2022.doi: https://doi.org/10.12962/j23373520.v11i1.69229

[9] L. Anselin, "SPATIAL MODELS IN ECONOMETRIC RESEARCH," 2021 doi: https://doi.org/10.1093/acrefore/9780190625979.013.643.

[10] J. F. Wang, T. L. Zhang, and B. J. Fu, "A MEASURE OF SPATIAL STRATIFIED HETEROGENEITY," *Ecol Indic*, vol. 67, pp. 250–256, Aug. 2016, doi: https://doi.org/10.1016/j.ecolind.2016.02.052.

[11] K. Kopczewska, "SPATIAL MACHINE LEARNING: NEW OPPORTUNITIES FOR REGIONAL SCIENCE," *Ann Reg Sci*, vol. 68, no. 3, pp. 713–755, Jun. 2022, doi: https://doi.org/10.1007/s00168-021-01101-x.

[12] N. Amoroso *et al.*, "PSI SPATIALLY CONSTRAINED CLUSTERING: THE SIBARI AND METAPONTO COASTAL PLAINS," *Remote Sens (Basel)*, vol. 15, no. 10, May 2023, doi: https://doi.org/10.3390/rs15102560.

[13] I. M. Pratiwi, Marseto, and Sishadiyati, "ANALISIS PENGARUH JUMLAH PENDUDUK, INDEKS PEMBANGUNAN MANUSIA DAN UPAH MINIMUM TERHADAP PENGANGGURAN DI KABUPATEN BANGKALAN," *Jurnal Syntax Transformation*, vol. 2, no. 6, pp. 787–796, 2021.doi: https://doi.org/10.46799/jst.v2i6.300

[14] D. Wulandari and N. Woyanti, "PENGARUH PENDIDIKAN, UPAH MINIMUM, DAN KESEMPATAN KERJA SEKTOR FORMAL TERHADAP PENGANGGURAN TERDIDIK DI PROVINSI JAWA BARAT (2017-2021)," *BISECER (Business Economic Entrepreneurship)*, vol. VI, no. 2, pp. 90–104, Jul. 2023.doi: https://doi.org/10.61689/bisecer.v6i2.434

[15] N. Ismayani and L. Mariana Hura, "ANALISIS KEPADATAN PENDUDUK TERHADAP TINGKAT PENGANGGURAN SEBELUM DAN DI SAAT (COVID-19) DI KOTA PADANG," *Jurnal Azimut*, vol. 4, no. 1, pp. 39–50, Jun. 2022, [Online]. Available: https://ojs.unitas-pdg.ac.id/index.php/azimut

[16] Y. Chen, "RECONSTRUCTION AND NORMALIZATION OF LISA FOR SPATIAL ANALYSIS," *PLoS One*, vol. 19, no. 5 May, pp. 1–26, May 2024, doi: https://doi.org/10.1371/journal.pone.0303456.

[17] L. V Pinto, C. S. S. Ferreira, and Pereira P, "USING LOCAL INDICATORS OF SPATIAL ASSOCIATION (LISA) TO ASSESS SPATIAL RELATION BETWEEN URBAN PARKS' STRUCTURE AND USERS' ACTIVITIES IN VILNIUS (LITHUANIA)," in *XVI International Scientific Conference "Monitoring of Geological Processes and Ecological Condition of the Environment,"* Kyiv, Ukraine, Nov. 2022, pp. 1–5. doi: https://doi.org/10.3997/2214-4609.2022580180

[18] P. Kowe, O. Mutanga, J. Odindi, and T. Dube, "EXPLORING THE SPATIAL PATTERNS OF VEGETATION FRAGMENTATION SSING LOCAL SPATIAL AUTOCORRELATION INDICES," *J Appl Remote Sens*, vol. 13, no. 02, pp. 1–14, Jun. 2019, doi: https://doi.org/10.1117/1.JRS.13.024523.

[19] N. M. Huda, F. Fran, Y. Yundari, L. Fikadila, and F. Safitri, "MODIFIED WEIGHT MATRIX USING PRIM'S ALGORITHM IN MINIMUM SPANNING TREE (MST) APPROACH FOR GSTAR(1;1) MODEL," *BAREKENG: Journal of Mathematics and Its Applications*, vol. 17, no. 1, pp. 0263–0274, Apr. 2023, doi: https://doi.org/10.30598/barekengvol17iss1pp0263-0274.

[20] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, "EUCLIDEAN DISTANCE MATRICES: ESSENTIAL THEORY, ALGORITHMS, AND APPLICATIONS," *IEEE Signal Process Mag*, vol. 32, no. 6, pp. 12–30, Nov. 2015, doi: https://doi.org/10.1109/MSP.2015.2398954.

[21] L. Anselin, "LOCAL INDICATORS OF SPATIAL ASSOCIATION-LISA," *Geographically Analysis*, vol. 27, no. 2, pp. 93–115, Apr. 1995.doi: https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

[22] L. Anselin, *THE MORAN SCATTERPLOT AS AN ESDA TOOL TO ASSESS LOCAL INSTABILITY IN SPATIAL ASSOCIATION*. 1996.

[23] R. Guillain and J. Le Gallo, "AGGLOMERATION AND DISPERSION OF ECONOMIC ACTIVITIES IN AND AROUND PARIS: AN EXPLORATORY SPATIAL DATA ANALYSIS," *Environ Plann B Plann Des*, vol. 37, no. 6, pp. 961–981, Jan. 2010, doi: https://doi.org/10.1068/b35038.

[24] A. S. Jaber, A. K. Hussein, N. A. Kadhim, and A. A. Bojassim, "A MORAN'S I AUTOCORRELATION AND SPATIAL CLUSTER ANALYSIS FOR IDENTIFYING CORONAVIRUS DISEASE COVID-19 IN IRAQ USING GIS APPROACH," *Caspian Journal of Environmental Sciences*, vol. 20, no. 1, pp. 55–60, Jan. 2022, doi: https://doi.org/10.22124/CJES.2022.5392.

[25] E. Orozco-Acosta, A. Adin, and M. D. Ugarte, "Scalable Bayesian modeling for smoothing disease risks in large spatial data sets using INLA," Spatial Statistics, vol. 41, p. 100496, 2021.

[26] E. Orozco-Acosta, A. Adin, and M. D. Ugarte, "Big problems in spatio-temporal disease mapping: methods and software," Computer Methods and Programs in Biomedicine, vol. 231, p. 107403, 2023.

[27] G. Vicente, A. Adin, T. Goicoa, and M. D. Ugarte, "High-dimensional order-free multivariate spatial disease mapping," Statistics and Computing, vol. 33, p. 104, 2023.