



**ANÁLISIS ESPACIAL DE LOS MUNICIPIOS EN
CONDICIÓN DE POBREZA, PARA LOS
DEPARTAMENTOS DE MAGDALENA, BOLÍVAR
Y SUCRE, DURANTE EL AÑO 2018**

**SPATIAL ANALYSIS OF PEOPLE IN POVERTY CONDITION
AT MUNICIPAL LEVEL, IN THE DEPARTMENTS OF
MAGDALENA, BOLIVAR AND SUCRE DURING YEAR 2018**

Pablo Andrés Lombana Zorro
U3101509
Ingeniero Catastral y Geodesta

Director:
Ing. Freddy León Reyes M.Ed

**UNIVERSIDAD MILITAR NUEVA GRANADA
ESPECIALIZACIÓN EN GEOMÁTICA
NOVIEMBRE DE 2020
BOGOTÁ-COLOMBIA**

ANÁLISIS ESPACIAL DE LOS MUNICIPIOS EN CONDICIÓN DE POBREZA, PARA LOS DEPARTAMENTOS DE MAGDALENA, BOLÍVAR Y SUCRE, DURANTE EL AÑO 2018

SPATIAL ANALYSIS OF PEOPLE IN POVERTY CONDITION AT MUNICIPAL LEVEL, IN THE DEPARTMENTS OF MAGDALENA, BOLIVAR AND SUCRE DURING YEAR 2018

Pablo Andrés Lombana Zorro
Ingeniero Catastral y Geodesta.
Especialización en Geomática
Universidad Militar Nueva Granada.
Bogotá, Colombia
u3101509@unimilitar.edu.co

RESUMEN

La pobreza es un fenómeno que afecta las distintas ramas de la sociedad colombiana, tanto desde el aspecto económico como social, y el impacto que ejerce sobre la comunidad es de alto coste, por ende, se hace necesario investigar cómo se comporta esta problemática y cuáles son los posibles determinantes de esta situación. La presente investigación, conlleva a realizar un proceso de ciencia de datos a nivel espacial, utilizando análisis exploratorio de datos espacial y análisis confirmatorio de datos espaciales, para los municipios de los departamentos de Sucre, Bolívar y Magdalena durante el año 2018. Los resultados muestran cómo se distribuye y aglomeran los municipios con altos registros de personas en condición de pobreza, determinando que la relación espacial es de tipo positivo y el modelo espacial autorregresivo, estableció que los principales factores que inciden sobre el fenómeno corresponden a analfabetismo y problemas de hacinamiento.

Palabras Clave: Análisis espacial de Datos, Pobreza, Modelos Espaciales, Regresión

ABSTRACT

Poverty is a phenomenon that affects the different branches of Colombian society, both from the economic and social aspects, and the impact it has on the community is costly, therefore, it is necessary to investigate how this problem behaves and which ones are the possible determinants of this situation. This research, entailed carrying out a data science process at a spatial level, using exploratory analysis of spatial data and confirmatory analysis of spatial data for the municipalities of the departments of Sucre, Bolívar and Magdalena during 2018. The results show how distributes and agglomerates the municipalities with high records of people living in poverty, determining that the spatial relationship is positive and autoregressive spatial model established that the main factors that affect the phenomenon correspond to illiteracy and overcrowding problems

Keywords: Spatial Data Analysis, Poverty, Spatial Models, Regression.

INTRODUCCIÓN

La pobreza es un concepto que define a una población humana que se encuentra bajo un esquema de inaccesibilidad de bienes y servicios, donde fundamentalmente no se posee el acceso a los elementos básicos de consumo, por causas asociadas a la baja generación de ingresos, conllevando a que se establezca la condición de pobreza monetaria (Haughton G & Khandker, 2009).

Esta problemática es un reflejo de políticas económicas y sociales atrasadas que se presentan a lo largo de décadas en las diferentes regiones del mundo, ya que por una parte frena el desarrollo de la economía y posterga el rompimiento de la brecha social que divide las sociedades con base a sus ingresos, y por otra, degrada la condición de la vida humana llevándola a problemas de hambruna y desnutrición (Vargas, 2010). Las cifras de pobreza a nivel mundial, indican que para el año 2015, aproximadamente 734 millones de personas se encontraban en condición de pobreza, relacionando un total del 10 % de la población mundial, que si bien, es un valor reducido a lo que se presentaba en el año 1990 (36% de la población se consideró pobre) aún es un dato significativo en la lucha contra este fenómeno (Banco Mundial, 2020).

En Colombia, los esfuerzos para mitigar la pobreza han venido aumentando desde la década de los 90, creando políticas económicas de apertura al mercado global, descentralización fiscal, reforma laboral entre otros, esto ha significado que se disminuya los niveles de pobreza sustancialmente, pero aún así las cifras del año 2000, la población pobre se ubicó en el 59% del total de población (Núñez Méndez et al., 2002) y para el año 2011 el porcentaje de personas en pobreza correspondió a 34.1% (Universidad Católica de Oriente, 2013). En cifras más reciente la población monetariamente pobre se ubicó en el 27,0% (DANE, 2018b).

Según el Departamento Administrativo Nacional de Estadística (DANE), la medición de la pobreza en Colombia se realiza evaluando métodos directos, los cuales muestran la satisfacción que posee un individuo con base a características fundamentales como la salud, educación, entre otros aspectos críticos en el desarrollo del individuo, por otra parte, esa misma entidad relaciona que el índice de pobreza multidimensional es una aproximación utilizando datos censales y cuya metodología de elaboración es un proceso diseñado por el Departamento Nacional de Planeación DNP (DANE, 2018), el cual busca clasificar la población en condición de pobreza.

Identificar cuáles son las causales de la pobreza, permiten que exista un planteamiento lógico y secuencial, sobre el cual se pueden implementar políticas que permitan ayudar a reducir este fenómeno, de allí que el uso de modelos de regresión son una

herramienta fundamental para el estudio de la pobreza (Cruz-Martinez, 2015). La investigación realizada por Liu et al., 2020, se enfocó en establecer dos metodologías para analizar las variables que influyen sobre la pobreza, la primera utilizó mediante algoritmos de machine learning, más específicamente random forest, un análisis de la importancia relativa de las diferentes características. La segunda metodología, implicó el uso de modelos de regresión y la implementación de los resultados del modelo de random forest que determinó las características más importantes dentro del proceso. Por una parte, el modelo de regresión logró capturar mejor la importancia de cada característica, pero el modelo de random forest estableció que la accesibilidad de las personas a la industria, bancos, ciudades y escuelas son los factores de mayor relevancia. Adicionalmente, el estudio presencié la implementación de estadística espacial, sobre la cual se validó la autocorrelación espacial a nivel global y local, donde se permitió encontrar puntos calientes y fríos de concentración de pobreza en China.

El estudio de la pobreza a nivel de regiones, permite realizar un análisis sobre cómo se comporta el fenómeno a nivel espacial, también, existe la posibilidad de asociar las causas que conllevan a la condición de pobreza y a la explicación del fenómeno dentro del espacio. Sobre el enfoque geográfico y espacial, un estudio desarrollado en Sabzevar, Iran implementó una metodología de encuestas para 18 vecindarios, en donde se evaluaron un total 17 indicadores de la pobreza a nivel urbano (zandi et al., 2019). Según el estudio, los datos se asociaron a los 18 vecindarios, y la metodología se basó en procesos de red analítica, análisis relacional gris y se implementaron métodos de estadísticas espaciales. Los resultados obtenidos determinaron que las zonas de mayor pobreza corresponden, a aquellos lugares donde se asentaban barrios no oficiales y áreas de expansión urbana, adicionalmente, se identificaron cluster de agrupación espacial de la pobreza en los vecindarios mediante técnicas de medición de la autocorrelación espacial.

Otro estudio evocado en Estados Unidos, desarrolla y analiza la pobreza de una manera estructural mediante variables de tipo económico, social y político. La metodología usada corresponde a la generación de modelos de regresión y usando técnicas de análisis de datos espaciales, tales como: diagrama de dispersión de moran, mapa de cuantiles e indicadores autocorrelación espacial. Los resultados son concluyentes, ya que se establece que la desigualdad étnica, los ingresos, capital social, población nacida en el extranjero son efectos causantes de pobreza a nivel espacial para los condados en el país (Rupasingha & Goetz, 2007).

El crecimiento económico refleja una manera en la cual una población puede aumentar y mejorar sus condiciones de vida en una determinada región, ya que al acceder a artículos de consumo básicos, no solo contribuye a la economía de la zona, sino también a mejorar el bienestar propio del individuo; pero a su vez, está ligado a que se presente un aumento en la población, puesto que mejoran sus condiciones de vida y son más atractivos y habitables dichas zonas, conllevando a que ese crecimiento económico disminuya por la ampliación desmedida de la población (Kuznets, 1974).

Esto es un fenómeno que se presenta regularmente, como lo demuestra Galvis (2015), el cual realizó una investigación que reunió en conjunto el crecimiento económico junto con la densidad de la población. El autor, determina el uso de cadenas de Markov a nivel espacial para su análisis y también buscan identificar aglomeraciones de características usando mapas de autocorrelación espacial. Los resultados son bastante explicativos ya que, hace énfasis en ciudades como Bogotá, la cual tuvo un crecimiento económico y demográfico considerable, pero deja en claro la reflexión de que el aumento de la población genera que la calidad de vida se deteriore y que también los índices de delincuencia aumentan, como también los indicadores de pobreza.

La región del caribe colombiano es una zona que se encuentra altamente afectada por el fenómeno de la pobreza y es que según la encuesta de calidad de vida, la distribución geográfica de este índice, demostró que aún existen desigualdades bastantes significativas en la región y a su vez que existen aglomeraciones significativas de municipios bajo una condición de desfavorabilidad principalmente en el departamento del magdalena (Tapias Ortega, 2017). Por otra parte, según el comunicado de prensa emitido por el DANE en el año 2015, los departamentos de Magdalena, Sucre y Bolívar presentaron 50,5 %, 47,3 % y 41,8 % de población en condición de pobreza (DANE, 2015, p. 24).

En consideración a la importancia que conlleva realizar un análisis econométrico espacial sobre la pobreza y principalmente en la región de la costa colombiana, se escogieron los departamentos de Magdalena, Bolívar y Sucre para ser objeto de un estudio a nivel espacial sobre el esquema de datos de área lattice, donde se relaciona las características propias de cada municipio de los departamentos y se representó las asociaciones espaciales. Consecutivamente se desarrolló un modelo espacial autorregresivo con los atributos seleccionados para una muestra de 120 municipios en los tres departamentos para el año 2018.

1. MATERIALES Y MÉTODOS

1.1 MATERIALES

1.1.1 Área de estudio

Los departamentos de Sucre, Bolívar y Magdalena se encuentran en la zona norte de Colombia, específicamente en la región Caribe. Sucre cuenta con un total de 26 municipios y una extensión de 10.587 km², limita al norte con el Mar Caribe, al sur con el departamento de Antioquia, al occidente con Córdoba y al oriente con Bolívar. Magdalena posee un total de 30 municipios con una superficie de 23.134 km² y cuyos límites corresponden a; por el norte el Mar Caribe, al sur con Bolívar, al occidente con

Atlántico y Bolívar, al oriente con Cesar y La Guajira. Por otra parte, Bolívar alberga 46 municipios con un área de 26.685 km² y sus límites geográficos corresponde al norte con el Mar Caribe y Atlántico, al sur con Santander y Antioquia, al occidente con Sucre y al Oriente con Magdalena. En la Figura 1 se representan los departamentos objeto de estudio con sus respectivos municipios:

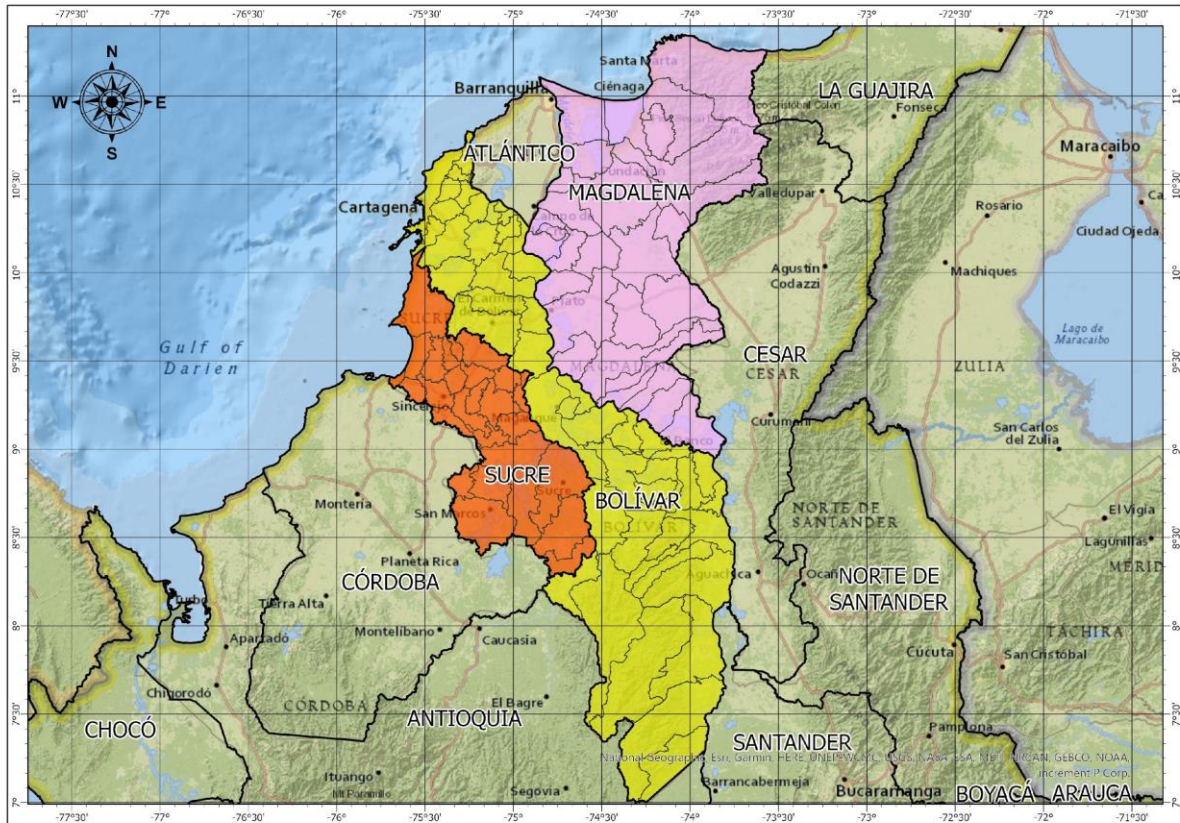


Figura 1: Área de Estudio, Departamentos de Sucre, Bolívar y Magdalena

1.1.2 Insumos

A partir de información oficial de diferentes fuentes, se recopilamos datos asociados a factores incidentes en la pobreza. La tabla 1, contiene la descripción de las variables utilizadas dentro del estudio con su debida descripción.

Tabla 1: Recopilación y descripción de las variables de estudio

Tipo	Variable	Temporalidad	Medida	Formato	Fuente	Descripción
Endógena	POP	2018	Unidad	Excel	DANE (Encuesta Calidad de Vida)	Personas que se encuentran en condición de Pobreza según Índice de Pobreza Multidimensional

	AFABE	2018	Porcentaje	Excel	DANE (Encuesta Calidad de Vida)	Índice de analfabetismo
	HAC	2018	Porcentaje	Excel	DANE (Encuesta Calidad de Vida)	Índice de hacinamiento
	TOT ESC	2018 - 2020	Unidad	SHP*	API Google Places	Total de escuelas
	KM VIA	2017	Kilómetro	SHP	IGAC	Total de Kilómetros de vía construidos
	DEND POB	2018	Porcentaje	Excel	DANE (Encuesta Calidad de Vida)	Total de habitantes por unidad de área en km ²
Exógen as	TOT SAL	2018 – 2020	Unidad	SHP	API Google Places	Total entidades de salud
	INASIS	2018	Porcentaje	Excel	DANE (Encuesta Calidad de Vida)	Índice de inasistencia escolar
	DEP ECO	2018	Porcentaje	Excel	DANE (Encuesta Calidad de Vida)	Índice de dependencia económica
	BSCI TO	2018	Porcentaje	Excel	DANE (Encuesta Calidad de Vida)	Índice de barrera de servicios de primera infancia
	SAS TO	2018	Porcentaje	Excel	DANE (Encuesta Calidad de Vida)	Índice de sin aseguramiento en salud
	IITEM	2018	Porcentaje	Excel	DANE	Índice que describe la importancia económica en el departamento.

SHP: Información recopilada en formato shapefile nativo de ESRI.

DANE: Departamento Administrativo Nacional de Estadística.

IGAC: Instituto geográfico Agustín Codazzi.

1.1.3 Software

Para el desarrollo de este estudio se utilizó principalmente el IDE (Entorno de Desarrollo Integrado) llamado Pycharm en versión 2019, a su vez se implementó bajo Python 3.7. Las librerías o paquetes con licencia open source utilizados de Python correspondieron a: Geopandas, numpy, pysal, seaborn, pandas, matplotlib, os, chart_studio, cufflinks, plotly, scypy, esda, mgwr. Por otra parte, la elaboración de los mapas se generó mediante Arcgis Pro 2.6 con licencia estudiantil.

1.2 MÉTODOS

A continuación, se presentan los métodos utilizados para el desarrollo del presente estudio, donde en primera instancia se realizan procesos de estadística descriptiva, luego, se ejecuta el análisis exploratorio de datos espaciales (AEDE) y por último se ejecuta el análisis confirmatorio de datos espaciales (ACDE).

1.2.1 Proceso estadística descriptiva

La estadística descriptiva es el conjunto de técnicas tanto gráficas como numéricas que permite describir información a priori de un conjunto de datos, para analizar y extraer información relevante sobre la población total (Faraldo & Pateiro, 2012), con base a esto, se utilizaron algunas medidas de forma, tendencia central y dispersión para conocer el estado actual de la información. Además, se insertaron gráficos de dispersión para las variables de mayor relación (4 en total) y mapa de calor asociados a la correlación

1.2.2 Validación de la Normalidad de la variable Endógena

Los modelos de regresión tanto espaciales como no espaciales, funcionan bajo el cumplimiento de los supuestos y uno de ellos corresponde a la normalidad en la variable respuesta, esto especifica que la endógena tiene una distribución como la campana de Gauss (Bivand et al., 2013). Esto implicó que se aplicara la transformación Box – Cox, donde se busca obtener un valor de $\lambda = 0$ para que cambie la forma de la variable y se logre corregir errores en el supuesto de normalidad en los datos (Castaño, 2011).

1.2.3 Análisis Econométrico Espacial.

Según Anselin, 1988 la econometría espacial se ocupa de la dependencia espacial y la heterogeneidad espacial, encontrando aspectos críticos de los datos utilizados por los científicos regionales, por ende, esta metodología se divide en dos fundamentales procesos: 1) Análisis Exploratorio de Datos Espaciales y 2) Análisis Confirmatorio de Datos Espaciales.

1.2.3.1 Análisis Exploratorio de Datos Espaciales (AEDE)

Dentro del campo de la econometría espacial se encuentra análisis exploratorio de datos espaciales, que es una disciplina en donde se busca extraer información de datos georeferenciados, generalizando conceptos propios del análisis de datos que se da en muchos aspectos desde la minería de datos, aquí se busca fundamentalmente conocer cómo son las relaciones de las variables pero desde un entorno espacial (Yrigoyen, 2011). El objetivo principal del AEDE corresponde a identificar excepciones locales o tendencias que se presenten en los datos georeferenciados (Acevedo & Velásquez, 2008). Para ello ejecutaron análisis de autocorrelación global y local, donde se utilizan el I de Moran (Moran, 1948) y C de Geary (Cliff & Ord, 1973) para validar las pruebas de hipótesis, adicionalmente, se construyen mapas LISA (Local Index Spatial Autocorrelation) para verificar agrupaciones de valores de personas en condición de pobreza entre municipios (Anselin, 1995).

1.2.3.2 Análisis Confirmatorio de Datos Espaciales (ACDE).

Según Moreno & Vaya, 2000 citados en Lombana Zorro & Jiménez Joven, 2016, con la adición de un componente espacial sobre los modelos de regresión clásicos, se hace la necesidad de construir modelos que capturen tanto la relación entre las variables independientes con la dependiente, como las asociaciones espaciales (bien sea que se encuentren correlacionadas espacialmente o que exista una dependencia espacial en el término de perturbación). Este es un modo en el cual los datos son analizados mediante métodos estadísticos con el objeto de confirmar o descartar hipótesis hechas a priori sobre los datos recopilados (Parra, 2002). En primera instancia y con la inclusión de un parámetro que captura las relaciones espaciales, se realizó un modelo de regresión lineal simple para validar la relación de la variable dependiente con las independientes, en un paso siguiente, se desarrollaron dos modelos espaciales autorregresivos. Moreno & Vaya (2000) describen las ecuaciones de los modelos como se presenta a continuación:

Se desarrollaron dos modelos autorregresivos espaciales, el primero contiene el término de que está rezagado espacialmente y se describe de la siguiente manera

$$y = \rho W y + X \beta + u$$
$$u \sim N(0, \sigma^2 I)$$

Fuente: Tomado de Moreno & Vaya (2000) Econometría Espacial; nuevas técnicas para el análisis regional

Donde ρ corresponde al parámetro autorregresivo, $W y$ la matriz de retardo espacial de la variable y , X es la matriz de las regresoras y β son los coeficientes estimados. u es el término de perturbación y se distribuye normalmente con media 0 y variancia sigma cuadrado por la matriz identidad.

El otro modelo presentado contempla que la autocorrelación espacial se encuentra en el término de perturbación y se describe de la siguiente manera:

$$y = X \beta + \varepsilon$$
$$\varepsilon = \lambda W e + u$$
$$u \sim N(0, \sigma^2 I)$$

Fuente: Tomado de Moreno & Vaya (2000) Econometría Espacial; nuevas técnicas para el análisis regional

$W e$ el rezago espacial en el término del error, ε corresponde al término de perturbación y u es un término de perturbación de la dependencia espacial residual.

2. RESULTADOS Y DISCUSIONES

2.1 ANÁLISIS SOBRE LAS ESTADÍSTICAS DESCRIPTIVAS

Las tabla 2 y tabla 3, muestran los resultados de las estadísticas descriptivas las cuales reflejan una completitud en la totalidad de los datos ya que todas las variables cuentan con registro no vacíos. Por otra parte, en el caso de la variable POP, el promedio la población económicamente en condición de pobreza se ubica en 2.328 personas, también se observa que el valor máximo ronda los 2.927 personas y en el caso de las medidas de forma, la kurtosis y el coeficiente de asimetría (Skewness) presentan valores de 4,3 y 1,9 respectivamente, esto muestra que la distribución de la variable es asimétrica de cola derecha y el apuntamiento es de tipo leptocúrtica. En general las variables que refieren a índices, muestran valores promedio cercanos a 0, pero variables como AFABE y SAS tienen valores significativamente altos.

Tabla 2: Estadísticas descriptivas para el primer grupo de variables

Tipo	Ítems	POP	AFABE	HAC	TOT_ESC	KM_VIA	DEND_POB
Tendencia Central	Cantidad	102,0	102,0	102,0	102,0	102,0	102,0
	Promedio	2328,9	0,3	0,1	9,7	1007,0	86,9
	Desv Estándar	2195,8	0,1	0,0	15,6	2463,1	179,6
Distribución	Mínimo	56,0	0,1	0,0	0,0	11,0	6,4
	25%	876,0	0,3	0,0	2,0	237,3	24,5
	50%	1710,0	0,3	0,1	5,0	424,5	48,9
	75%	2927,3	0,4	0,1	11,8	762,8	80,7
Forma	Máximo	12362,0	0,5	0,2	110,0	20729,0	1486,5
	Skewness	1,9	-0,2	0,6	4,4	6,3	6,2
	Kurtosis	4,3	0,0	-0,2	23,9	45,0	42,5

Tabla 3: Estadísticas descriptivas para el segundo grupo de variables

Tipo	Ítems	TOC_SAL	INASIS	DEP_ECO	BSCI_TO	SAS	IITEM
Tendencia Central	Cantidad	102,0	102,0	102,0	102,0	102,0	102,0
	Promedio	4,5	0,0	0,1	0,0	0,2	478,1
	Desv Estándar	9,1	0,0	0,0	0,0	0,0	2042,8
Distribución	Mínimo	0,0	0,0	0,0	0,0	0,1	25,9
	25%	1,0	0,0	0,1	0,0	0,1	92,1
	50%	2,0	0,0	0,1	0,0	0,2	132,6
	75%	4,0	0,0	0,1	0,0	0,2	252,3
Forma	Máximo	66,0	0,1	0,2	0,1	0,2	19943,7
	Skewness	4,9	2,8	0,5	1,2	0,4	8,9
	Kurtosis	27,2	9,9	1,0	1,4	0,0	83,9

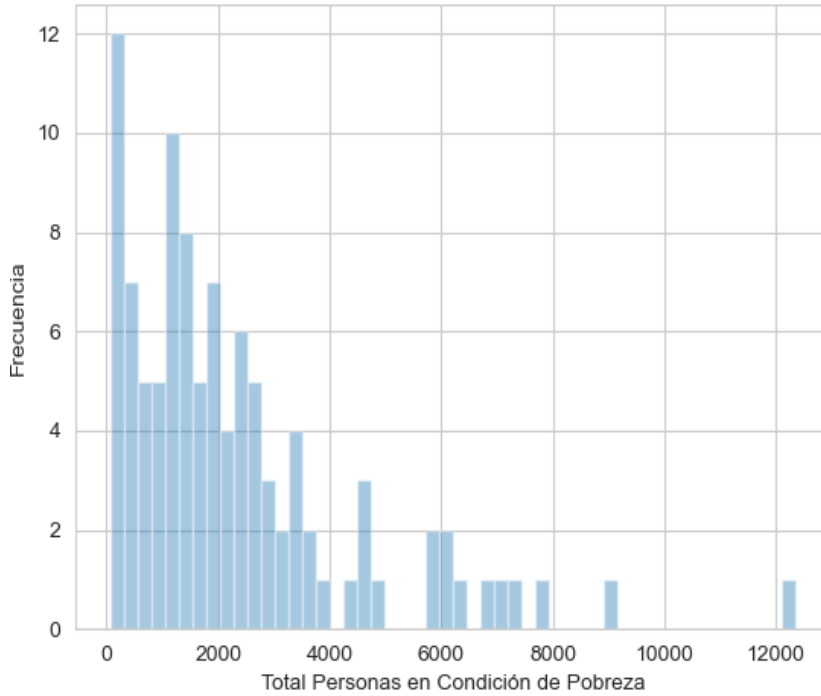


Figura 2: Histograma de la distribución de la variable respuesta

El histograma anterior, mostró que la distribución de la variable respuesta no presenta un comportamiento normal y se encuentra distribuido a la izquierda del conjunto de datos, corroborando lo encontrado el coeficiente de asimetría.

El análisis de correlación de las variables (tanto endógena como exógenas) arrojó valores significativamente bajos para variables como DEP_ECO, INASIS, SAS y EIIM, en caso contrario valores como 0.26, 0.327, 0.322 y 0.32, que corresponden a variables HAC, TOC_ESC, TOC_SAL y KM_VIA respectivamente, presentaron un grado de relación mayor, tal y como se muestra en la siguiente figura:

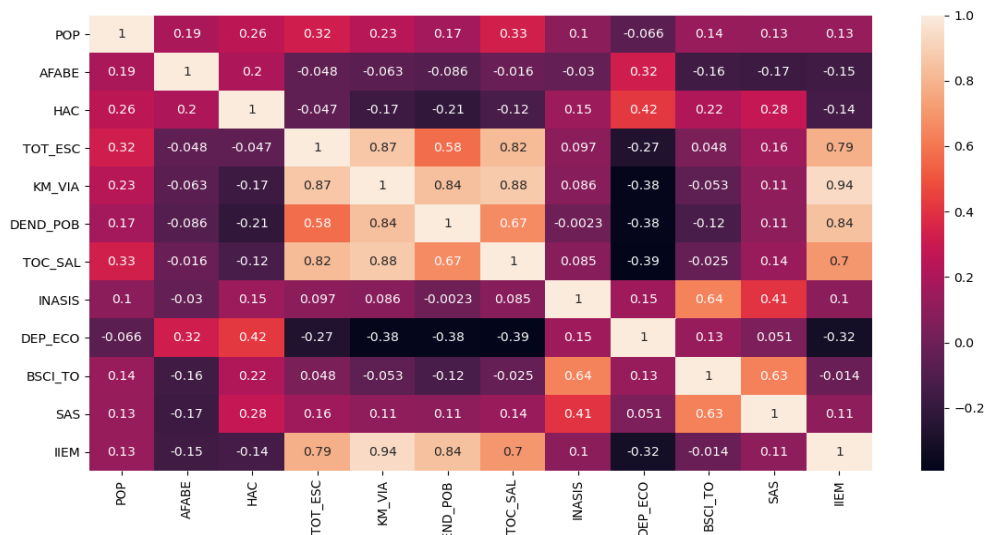


Figura 3: Mapa Correlación de Calor de las variables de estudio.

Los resultados de los diagramas de dispersión (figura 4), para las variables en términos de porcentaje que significaron un mayor valor de correlación, determinaron un aumento no tan pronunciado, en cuanto a la medida de la variable y su relación con POP. Para el caso de HAC, muestra que el aumento es lento y progresivo y que las poblaciones en condiciones de pobreza presentan altos índices de hacinamiento, por otra parte, en el caso del analfabetismo, la relación es positiva pero no muy fuerte tal y como se validó en el mapa de calor. En general, las relaciones de la variable respuesta con sus respectivas independientes, no muestra una asociación lineal fuerte y se presentan concentraciones de valores bajos en relación al total de personas en condición de pobreza por cada municipio.

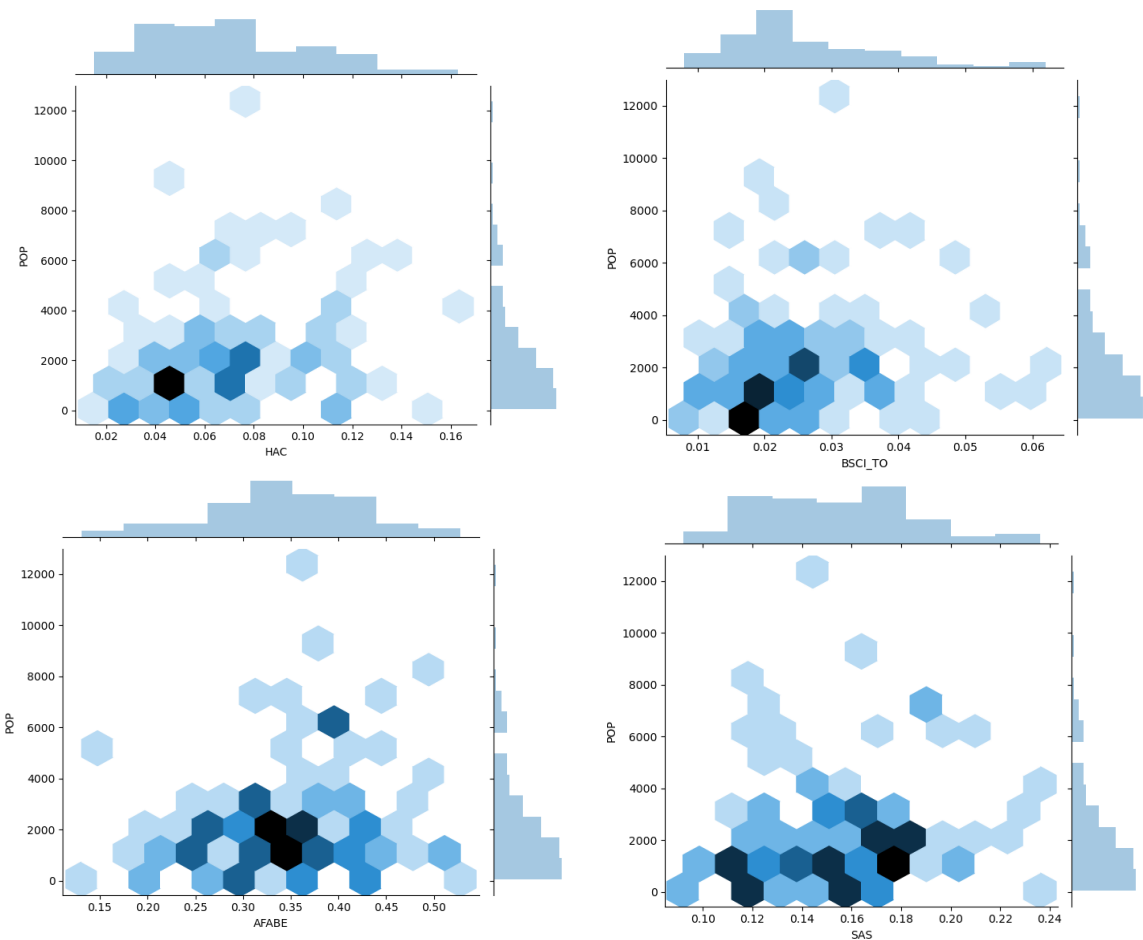


Figura 4: Diagramas de Dispersión para las variables en términos porcentuales que presentaron mayor correlación

El proceso de la estadística descriptiva arrojó resultados que son significativos dentro el presente estudio, ya que, por una parte, mostró una baja relación de la variable respuesta con sus respectiva independiente en mayor cuantía, pero también aquellas variables que mostraron mejor correlación (aun siendo baja) son las que se asocian directamente con la problemática de pobreza, como la tasa de analfabetismo, el hacinamiento y la falta de infraestructura educativa. En el caso de AFABE, esto es una

clara evidencia que la educación es un punto fundamental en el desarrollo de las personas y que su valor de correlación, presenta que los municipios con altos valores de analfabetismo son los que se encuentran mayor cantidad de personas en condición de pobreza, estableciendo que existe una brecha muy pequeña entre las variables educativas y la pobreza (Mendoza, 2003). El hacinamiento demostró estar por encima de las demás variables independientes en cuanto a su explicación sobre la variable respuesta, tanto en el dispersograma como en el mapa de calor de correlación, eso refleja que los municipios en condición de pobreza tienden una tendencia al alza en cuanto mayor es el hacinamiento más serán las personas en condición de pobreza y que es un factor determinante dentro del análisis del fenómeno.

2.1 PRUEBAS DE NORMALIZACIÓN Y TRANSFORMACIÓN VARIABLE RESPUESTA

Los resultados que se obtuvieron del proceso de la estadística descriptiva, dieron a conocer visualmente que es necesario realizar una transformación a la variable respuesta para que cumpliera el supuesto de normalidad, por ende, se ejecutó el test de Shapiro - Wilk para verificar estadísticamente el cumplimiento o no del supuesto. La gráfica 5 muestra que al aplicar una transformación Box-Cox, se rechaza la hipótesis nula de no normalidad en los datos y que está se distribuye normalmente.

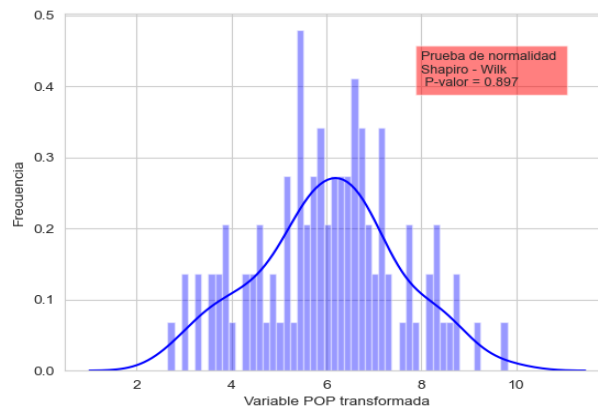


Figura 5: Histograma de la variable transformada

2.1 ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES

La estructuración de los criterios de contigüidad para la determinación de la matriz de pesos espaciales, determinaron que utilizar una matriz con cuatro vecinos más cercanos (criterio de contigüidad física) recopila en mayor grado las asociaciones espaciales en las unidades geográficas para la variable respuesta. Posteriormente,

con la obtención de la matriz de pesos espaciales, se ejecutaron el I de Moran y el C de Geary para determinar la no heterogeneidad espacial o autocorrelación espacial.

La tabla 4, mostró que efectivamente la variable endógena se encuentra correlacionada espacialmente en los municipios de los tres departamentos con un nivel de significancia al 95% y a su vez que es de tipo positivo.

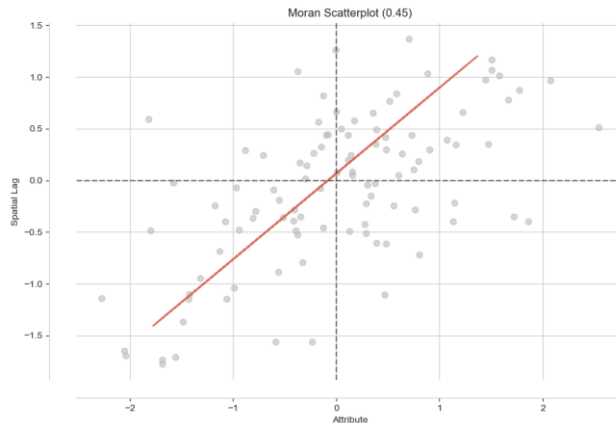


Tabla 4: Pruebas Autocorrelación Espacial

	Valor Estadístico	P-valor
I Moran Test	0,451	0,001
C Geary test	0,548	0,001

Figura 6: Scatter Plot de Moran

Por otra parte, el scatter plot de Moran válida lo obtenido por las pruebas de autocorrelación espacial, donde la pendiente es positiva e indicó que municipios con valores altos de personas en condiciones de pobreza se encuentran rodeados de municipios con esta misma característica, a su vez la mayor concentración de los datos se encuentra sobre el primer y tercer cuadrante. El siguiente mapa, relaciona espacialmente los resultados de la prueba y también visualizó los cluster de asociación espacial que se presentan en el área de estudio.

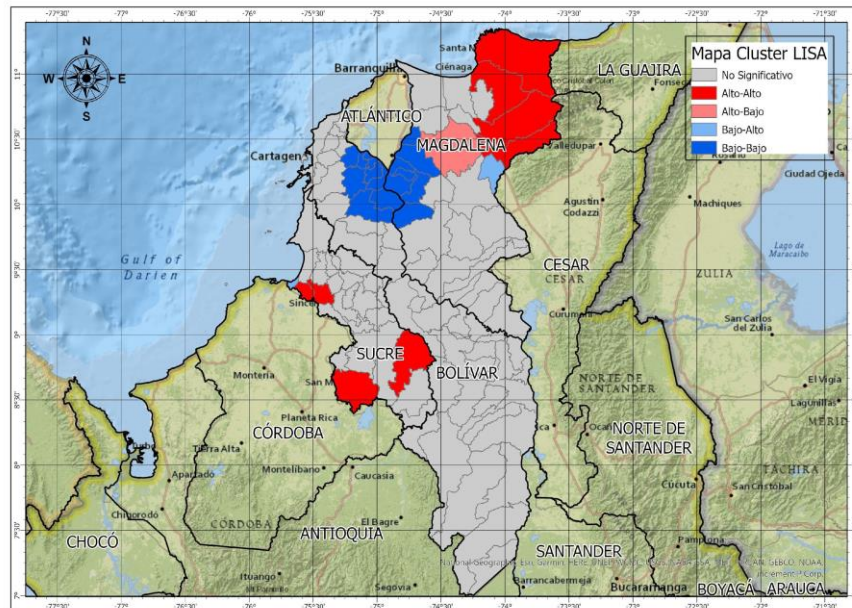


Figura 7: Mapa Autocorrelación Espacial Local (LISA)

Con el mapa LISA, se encontraron puntos calientes y fríos a nivel espacial que resultaron ser significativos estadísticamente, de allí se resalta que en la parte norte del departamento de Magdalena en los municipios de Santa Marta, Ciénaga, Aracataca y Fundación se presentan asociaciones de Alto – Alto, por otra parte, se presentaron agrupaciones de Bajo – Bajo en los límites del departamento de Magdalena y Bolívar, acogiendo municipios como Cerro de San Antonio, El Piñon, Concordia, Pedraza, Zapayán, Tenerife por los límites del Magdalena y por parte de Bolívar, San Estanislao, Soplaviento, San Cristóbal, Arroyohondo, Mahates, Calamar, San Juan de Nepomuceno y El Guamo. Sucre presentó tres cluster de agrupaciones, en los municipios de San Marcos, Sucre, Palmito y Sincelejo. El resultado también relaciona que hay un cluster de asociación de tipo Alto – Bajo en el municipio de Pivijay y Bajo – Alto en Algarrobo.

La relación de la pobreza a nivel espacial, determinó que este fenómeno se puede asociar espacialmente entre municipios y que aquellas regiones con valores altos de personas en condición de pobreza se encuentran rodeados de zonas con esta misma característica (Rupasingha & Goetz, 2007), indicando autocorrelación espacial de la variable que relaciona la pobreza. Consecuentemente, los resultados asociados a mayor cantidad de personas en condición de pobreza en los municipios del área de estudio, determinaron que estos municipios tienen valores altos asociados a factores como el índice de analfabetismo o el hacinamiento, esto se ve estructurado y especificado por una realidad que afecta al país donde las bajas oportunidades educativas y la cobertura negada en ciertas zonas, obliga a que las personas no tengan oportunidades de acceder a un empleo formal (Buitrago & Cabrera, 2014). Adicionalmente, Las principales actividades económicas desarrolladas en el departamento del Magdalena (El que mayor presentó municipios en condición de pobreza), corresponden a tipo sociales (26,6%), seguido de comercio, restaurantes, hoteles y reparaciones (15,4%) y la menor participación proviene de actividades mineras (0,4%), por otra parte, el PIB per cápita del departamento se ubica muy por debajo del promedio del país (CORPOMAG & Universidad del Atlántico, 2017) conllevando a que se presente en mayor cuantía municipios valores de pobreza por encima del promedio en el área de estudio.

2.1 ANÁLISIS CONFIRMATORIO DE DATOS ESPACIALES

Según los resultados que se obtuvieron del análisis exploratorio, afirman que existe una relación entre la variable dependiente y las independientes, por ende, se ejecutó un modelo de regresión lineal por el método de mínimos cuadrados ordinarios, de aquí se validó el porcentaje de variabilidad explicado de la variable respuesta por parte de las variables independientes, ratificación del cumplimiento del supuesto de normalidad mediante el test de Jarque – Bera en los residuales del modelo, así como los parámetros que son significativos estadísticamente y su relación positiva o negativa con la explicativa. Para este caso, se utilizó POP transformado para hacer cumplimiento de los supuestos.

Tabla 5: Ajuste modelo de regresión lineal

Mínimos Cuadrados Ordinarios		
Coefficiente	Parametros	P-Valor
Constante	3,5947	0
AFABE	3,9216	0,04
HAC	14,1871	0,005
TOT_ESC	0,0494	0,023
KM_VIA	-8,14E-05	0,826
DEND_POB	0,0027	0,12
TOC_SAL	0,0043	0,924
INASIS	-8,5906	0,381
DEP_ECO	-8,0151	0,05
BSCI_TO	39,6163	0,041
SAS	-2,1235	0,715
IEM	-0,0003	0,313
AIC	355,0	
R2	0,321	
R2 Ajustado	0,237	
Jarque – Bera	0,129	
Probabilidad (JB)	0,937	

La Tabla 5, permite establecer que la mayoría de las variables no son significativas estadísticamente, a excepción de AFABE, HAC, TOT_ESC, BSCI_TO y la constante aun nivel de significancia del 95%. También que apenas el 23% del fenómeno de estudio se logró explicar por parte de las variables recopiladas y que no existen problemas de normalidad en los residuos según lo obtenido en el test de Jarque – Bera. Es importante aclarar, que fue necesario incluir el componente espacial para validar si este parámetro ayuda en la explicación del fenómeno de la pobreza. Con base a los resultados obtenidos del modelo de regresión lineal, se desarrollaron dos modelos adicionales, pero con la inclusión del término que recoge la dependencia espacial que son los modelos de tipo autorregresivo. La tabla 6 recopiló los rendimientos del proceso de modelado. El producto de los modelos espaciales autorregresivos (Tabla 6), arrojan diferentes resultados en cuanto a los contrastes de bondad ajuste. En el caso del modelo de rezago espacial, variables como AFABE, HAC, TOC_ESC fueron significativas estadísticamente al 95% de significancia (a un 90% se puede incluir DEP_ECO), el valor del R cuadrado aumentó a 45%, la probabilidad del parámetro Rho resultó ser significativa. Por otro lado, el modelo de rezago en el error presentó dos variables significativas (AFABE y HAC) y el intercepto, también el valor del R2 no aumentó y se mantuvo según lo que se obtuvo en el modelo lineal, pero presentó un menor valor de AIC en comparación con el que se obtuvo en el modelo de rezago espacial. Estos modelos indican, que las variables como la tasa de analfabetismo y el hacinamiento son claros causantes de personas en condición de pobreza para los municipios del área de estudio.

Tabla 6: Resultados de Modelos con dependencia espacial sustantiva y de rezago en el error

Modelo de Rezago Espacial			Modelo de Rezago en el Error		
Coficiente	Parámetros	P-Valor	Coficiente	Parámetros	P-Valor
CONSTANT	1,340	0,166	CONSTANT	4,593	0,000
AFABE	3,419	0,022	AFABE	2,610	0,083
HAC	5,030	0,034	HAC	0,875	0,0442
TOT_ESC	0,022	0,044	TOT_ESC	0,019	0,133
TOC_SAL	0,001	0,980	TOC_SAL	0,003	0,905
DEP_ECO	-5,857	0,082	DEP_ECO	-5,418	0,102
BSCI_TO	18,994	0,144	BSCI_TO	17,712	0,194
SAS	1,121	0,812	SAS	3,525	0,478
W_POP_TRS	0,499	0,000	lambda	0,581	0,000
AIC	355,0		AIC	332,594	
R2 Espacial	0,45		R2 Espacial	0,2101	
Rho	0,49		lambda	0,581	
Probabilidad Rho	0,000		Probabilidad Lambda	0,000	
LikeliHood	-157,701		LikeliHood	-158,297	

La inclusión del término que recoge la dependencia espacial en una variable dentro del modelo, presentó un mejor ajuste y mayor relación en la explicación de la variabilidad de la variable respuesta, pero el modelo que mejor describe el fenómeno de la pobreza a nivel espacial corresponde al de rezago espacial. Este resultado también presenta una similitud en cuanto a lo encontrado por Chen et al., (2015), en el cual la estimación del modelo espacial autorregresivo presenta un mejor ajuste que en el modelo lineal clásico. Por otra parte, los resultados que se obtuvieron de los parámetros de regresión en cuanto al valor de los signos es concluyente tanto para AFABE como para HAC, ya que un aumento en la tasa de analfabetismo y en el hacinamiento, elevará considerablemente la cantidad de personas por municipio, pero en el caso de TOT_ESC, pasa lo contrario al ser de tipo positivo muestra que si se construye una nueva escuela la pobreza aumentará, esto podría verse debido que no solo implica el construir un nuevo asentamiento educativo si no que hace falta reforzar la estructura organizacional de la educación para el correcto desarrollo de los individuos.

3. CONCLUSIONES

Los resultados obtenidos del proceso de modelado espacial de características, determinaron que el fenómeno de la pobreza presenta clúster de agrupaciones por municipio y entre departamentos, resolviendo que la pobreza se asocia espacialmente y afecta a sus colindantes entre regiones, también, que los factores más influyentes sobre la pobreza corresponden a analfabetismo y hacinamiento principalmente. Este

análisis permitió relacionar mediante ecuaciones matemáticas con la inclusión de un componente espacial, como es la dinámica de la pobreza a nivel municipal y que variables influyen en la explicación de este fenómeno, resaltando que el acceso a educación es un factor influyente y determinante en la explicación del problema, así como las condiciones de vivienda también son un factor asociado a la pobreza, permitiendo explorar e inferir cómo fue el desarrollo económico de cada municipio y de las personas durante el año 2018.

El análisis exploratorio de datos espaciales determinó que se presentan fundamentalmente dos grupos determinados como puntos fríos y calientes, en la zona norte – oriente del departamento del Magdalena y en la parte centro – occidente de Sucre, estableciendo que las relaciones de interdependencia entre municipios reflejan la problemática de la pobreza sectorizada y agrupada en los departamentos de estudio. A su vez, la relación de la pobreza se da debido a la falta de educación y problemas de calidad de vida, permitiendo establecer que municipios con valores similares de personas en condición de pobreza se encuentran rodeados de municipios con cifras semejantes.

Por otra parte, el proceso de modelado espacial representó una mejor manera de analizar el fenómeno de estudio y el modelo más adecuado correspondió al espacialmente rezagado. Esta metodología, permitió relacionar las variables no solo dentro del contexto de la regresión lineal clásica, si no que a su vez introdujo el componente de la posición espacial y las relaciones de entre regiones, lo cual conlleva a una aproximación en el estudio de esta problemática que afecta al país.

REFERENCIAS BIBLIOGRAFICAS

1. Acevedo, I., & Velásquez, E. (2008). *Algunos Conceptos de la Econometría Espacial y el Análisis Exploratorio de Datos Espaciales*. 27, 9-34.
2. Anselin, L. (1988). *Spatial Econometrics: Methods and Models* (Vol. 4).
3. Anselin, L. (1995). *Local Indicators of Spatial Association*.
4. Banco Mundial. (2020). *Pobreza: Panorama general* [Text/HTML]. World Bank. <https://www.bancomundial.org/es/topic/poverty/overview>
5. Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied Spatial Data Analysis with R*. Springer New York. <https://doi.org/10.1007/978-1-4614-7618-4>
6. Buitrago, H., & Cabrera, M. (2014). *EDUCACIÓN PARA EL TRABAJO Y DESARROLLO HUMANO EN LOS INICIOS DEL SIGLO XXI: INCLUSIÓN SOCIAL, EMPRENDIMIENTO Y AUTOGESTIÓN* [Universidad Pedagógica Nacional].

<http://repositorio.pedagogica.edu.co/bitstream/handle/20.500.12209/710/TO-17324.PDF?sequence=1&isAllowed=y>

7. Castaño, E. (2011). Una estimación no paramétrica y robusta de la transformación Box-Cox para el modelo de regresión. *Lécturas de Economía*, 89-106.
8. Chen, X., Pei, Z., Chen, A. L., Wang, F., Shen, K., Zhou, Q., & Sun, L. (2015). Spatial Distribution Patterns and Influencing Factors of Poverty—A Case Study on Key Country From National Contiguous Special Poverty-stricken Areas in China. *Procedia Environmental Sciences*, 26, 82-90. <https://doi.org/10.1016/j.proenv.2015.05.005>
9. Cliff, A., & Ord, K. (1973). *Spatial Autocorrelation, Monographs in Spatial Environmental Systems Analysis*.
10. CORPOMAG, & Universidad del Atlántico. (2017). *Caracterización, Diagnóstico y Análisis de Vulnerabilidades y Amenazas en el Departamento del Magdalena*. 31.
11. Cruz-Martinez, G. (2015). *Esfuerzo de bienestar y pobreza desde el enfoque monetarista y de capacidades: Análisis transnacional en América Latina y el Caribe (1990-2010)*. 29.
12. DANE. (2015). *20 de los 24 Departamentos estudiados redujeron la pobreza en 2014*. https://www.dane.gov.co/files/investigaciones/condiciones_vida/pobreza/cp_pobreza_departamentos_R2014.pdf
13. DANE. (2018a). *Boletín Técnico Medida de Pobreza Multidimensional Municipal CNPV2018*. https://www.dane.gov.co/files/investigaciones/condiciones_vida/pobreza/2018/informacion-censal/bt-censal-pobreza-municipal-2018.pdf
14. DANE. (2018b). *Boletín Técnico Pobreza Monetaria Departamental 2018*. https://www.dane.gov.co/files/investigaciones/condiciones_vida/pobreza/2018/bt_pobreza_monetaria_18_departamentos.pdf
15. Faraldo, P., & Pateiro, B. (2012). *Estadística y Metodología de la Investigación*. http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP-DPTO/MATERIALES/Mat_G2021103104_EstadisticaTema1.pdf
16. Galvis, L. A. (2015, diciembre 11). *Crecimiento económico y demográfico regional en Colombia, 1985-2012*. *Revista de Economía Institucional*. <https://doi.org/10.18601/01245996.v17n33.08>
17. Haughton G, J., & Khandker, S. R. (2009). *HANDBOOK ON POVERTY + INEQUALITY*.

<http://documents1.worldbank.org/curated/en/488081468157174849/pdf/483380PUB0Pove101OFFICIAL0USE0ONLY1.pdf>

18. Kuznets, S. (1974). *Población y Crecimiento Económico*. https://repositorio.cepal.org/bitstream/handle/11362/7950/S7400430_es.pdf?sequence=1&isAllowed=y
19. Liu, M., Hu, S., Ge, Y., Heuvelink, G. B. M., Ren, Z., & Huang, X. (2020). Using multiple linear regression and random forests to identify spatial poverty determinants in rural China. *Spatial Statistics*, 100461. <https://doi.org/10.1016/j.spasta.2020.100461>
20. Lombana Zorro, P. A., & Jiménez Joven, L. (2016). *Análisis Espacial de la Deforestación a Escala Municipal, Estudio de Caso: Departamento del Chocó 2005-2010*. <http://repository.udistrital.edu.co/handle/11349/4488>
21. Mendoza, J. L. (2003). EDUCACIÓN Y POBREZA EN EL PERÚ. *Pensamiento Crítico*, 2, 069-077. <https://doi.org/10.15381/pc.v2i0.9255>
22. Moran, P. (1948). *The Interpretation of Statistical Maps* (Vol. 10).
23. Moreno, R. M., & Vaya, E. (2000). Econometría Espacial; nuevas técnicas para el análisis regional: Una aplicación a las regiones europeas. *Investigaciones Regionales*, 83-106.
24. Núñez Méndez, J. A., Ramírez Jaramillo, J. C., Naciones Unidas, & Comisión Económica para América Latina y El Caribe (Bogotá). (2002). *Determinantes de la pobreza en Colombia: Años recientes*. CEPAL.
25. Parra, J. (2002). Análisis Exploratorio y Análisis Confirmatorio de Datos. 1, 11, 115-124.
26. Rupasingha, A., & Goetz, S. J. (2007). Social and political forces as determinants of poverty: A spatial analysis. *The Journal of Socio-Economics*, 36(4), 650-671. <https://doi.org/10.1016/j.socec.2006.12.021>
27. Tapias Ortega, J. M. (2017). Pobreza y violencia en la Región Caribe colombiana: Un enfoque espacial. *Ensayos sobre Política Económica*, 35(83), 139-153. <https://doi.org/10.1016/j.espe.2016.12.002>
28. Universidad Católica de Oriente. (2013). *Pobreza y Desigualdad*. <https://www.uco.edu.co/ova/OVA%20Economia%20Colombiana/Objetos%20informativos/Unidad%202/4.%20pobreza%20y%20desigualdad.pdf>
29. Vargas, G. (2010). *Introducción a la Teoría Económica*. <http://herzog.economia.unam.mx/profesores/gvargas/libro1/cp23pob.pdf>
30. Yrigoyen, C. C. (2011). *ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES AL SERVICIO DEL GEOMARKETING*. 21.

31. zandi, R., Zanganeh, M., & Akbari, E. (2019). Zoning and spatial analysis of poverty in urban areas (Case Study: Sabzevar City-Iran). *Journal of Urban Management*, 8(3), 342-354. <https://doi.org/10.1016/j.jum.2019.09.002>