



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

**Aplicación de modelos matemáticos mediante estadística espacial para
identificar zonas con alta incidencia en el delito del robo, en la ciudad
de Cuenca, Ecuador en el año 2019**

EDY SAMUEL VELASTEGUÍ CASCO

Trabajo de Titulación modalidad: Proyectos de Investigación y Desarrollo, presentado
ante el Instituto de Posgrado y Educación Continua de la ESPOCH, como requisito
parcial para la obtención del grado de:

**MAGÍSTER EN MATEMÁTICA, MENCIÓN MODELACIÓN Y
DOCENCIA**

Riobamba-Ecuador

Julio 2023

DECLARACIÓN DE AUTENTICIDAD

Yo, EDY SAMUEL VELASTEGUI CASCO, declaro que el presente **Trabajo de Titulación Proyectos de Investigación y Desarrollo**, es de mi autoría y que los resultados del mismo son auténticos y originales. Los textos constantes en el documento que provienen de otra fuente están debidamente citados y referenciados.

Como autor, asumo la responsabilidad legal y académica de los contenidos de este proyecto de investigación de maestría.

Riobamba, julio de 2023

Edy Samuel Velasteguí Casco

No. Cédula: 060379076-7

©2023, Edy Samuel Velasteguí Casco

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

CERTIFICACIÓN:

EL TRIBUNAL DEL TRABAJO DE TITULACIÓN CERTIFICA QUE:

El trabajo de titulación modalidad proyectos de Investigación y Desarrollo, titulado: Aplicación de modelos matemáticos mediante estadística espacial para identificar zonas con alta incidencia en el delito del robo, en la ciudad de Cuenca, Ecuador en el año 2019, de responsabilidad del señor Edy Samuel Velasteguí Casco, ha sido minuciosamente revisado por los Miembros del Tribunal del trabajo de titulación, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal autoriza su presentación.

Dra. Jenny Margoth Villamarín Padilla Mgtr. _____

PRESIDENTA

Dra. Silvia Mariana Haro Rivera Mgtr. _____

DIRECTORA

Mat. Luís Marcelo Cortéz Bonilla Mgtr. _____

MIEMBRO

Dr. Franklin Marcelo Coronel Maji Mgtr. _____

MIEMBRO

Riobamba, julio 2023

DEDICATORIA

Mi formación académica tuvo sus inicios en el hogar en el que nací y crecí, donde obtuve una formación y un carácter de superación constante basado en principios honestos y buenas costumbres practicadas por mis padres y mi familia. Por esta razón dedico este trabajo principalmente a mis padres, a toda mi familia y a todas las personas que me han acompañado desde temprana edad en este camino de aprendizaje, también a las personas que fueron integrándose en el camino de mi vida y se han convertido en parte de mi familia.

Edy

AGRADECIMIENTO

Mi eterna y absoluta gratitud a Dios, dueño y dador de mi vida, quien me trajo a este mundo a través de mis padres, para aprender e ir superándome día a día en cada una de las dificultades que se han ido presentando en el camino, permitiéndome vencer todos los escollos con su bendición y alcanzar muchos de los objetivos que me he planteado en mi vida tanto espiritual como personal, académica y profesional. Mi gratitud imperecedera a todos mis maestros porque durante toda mi trayectoria educativa han sido pilares de mi formación con sus experiencias, enseñanzas y conocimientos, lo que me ha permitido avanzar siempre. Mi agradecimiento infinito a mis padres que me dieron todo, más de lo que les fue posible y, a mi familia porque me han sabido acompañar en todo momento brindándome su apoyo incondicional de todas las formas posibles.

Edy

TABLA DE CONTENIDO

	Páginas
RESUMEN	xx
SUMMARY	xxi
CAPITULO I	1
1 INTRODUCCIÓN	1
1.1 OBJETIVO GENERAL	4
1.2 OBJETIVOS ESPECÍFICOS	4
1.3 HIPÓTESIS	4
CAPITULO II	5
2 MARCO TEÓRICO	5
2.1 ANTECEDENTES DEL ESTUDIO	5
2.2 BASES TEÓRICAS	6
<i>2.2.1 Estadística Espacial</i>	<i>6</i>
<i>2.2.1.1 Datos Espaciales</i>	<i>7</i>
<i>2.2.1.2 Visualización de datos Espaciales</i>	<i>8</i>
<i>2.2.1.3 Matriz de datos Espaciales</i>	<i>11</i>
<i>2.2.1.4 Estadístico general de Producto Cruzado</i>	<i>12</i>
<i>2.2.1.5 Autocorrelación Espacial</i>	<i>12</i>
<i>2.2.2 Modelado Espacial</i>	<i>20</i>
<i>2.2.2.1 Tipos de Modelos de Regresión Espacial</i>	<i>20</i>
<i>2.2.2.2 Modelo de Retardo Espacial</i>	<i>21</i>
<i>2.2.2.3 Modelo de Error Espacial</i>	<i>22</i>
<i>2.2.2.4 Modelo espacial de Durbin (SDM)</i>	<i>23</i>
<i>2.2.3 Prueba para Dependencia Espacial</i>	<i>25</i>
<i>2.2.3.1 Test de Moran</i>	<i>25</i>
<i>2.2.3.2 Test de multiplicador de LaGrange (LM)</i>	<i>26</i>
<i>2.2.4 Prueba de dependencia espacial sustantiva</i>	<i>28</i>
<i>2.2.5 Estimación de un Modelo Espacial</i>	<i>28</i>
<i>2.2.6 Índice de Akaike (AIC)</i>	<i>31</i>
CAPITULO III	32
3 METODOLOGÍA DE INVESTIGACIÓN	32

3.1 SELECCIÓN DE LOS MODELOS ESPACIALES	32
3.1.1 <i>Obtención de los datos espaciales para aplicación de los modelos espaciales.....</i>	32
3.1.2 <i>Análisis de las variables utilizadas para aplicar los modelos espaciales.....</i>	33
3.2 ESTIMACIÓN DE PARÁMETROS E INDICADORES	34
3.2.1 <i>Estimación de indicadores a nivel de zona urbana censal.....</i>	34
3.2.2 <i>Paquetes y librerías de estadística espacial aplicadas en el software estadístico R</i>	38
3.2.2.1 <i>Paquete rgdal</i>	38
3.2.2.2 <i>Paquete raster.....</i>	38
3.2.2.3 <i>Paquete sp.....</i>	39
3.2.2.4 <i>Paquete spdep</i>	40
3.2.3 <i>Código de R aplicado para los modelos espaciales.....</i>	44
CAPITULO IV	46
4 RESULTADOS Y DISCUSIÓN.....	46
4.1 ANÁLISIS EXPLORATORIO DE LAS CARACTERÍSTICAS DE LAS VARIABLES.....	46
4.1.1 <i>Análisis de las características de robo en sus diferentes modalidades.....</i>	47
4.1.2 <i>Variables de características sociodemográficas población ciudad de Cuenca.....</i>	51
4.2 ANÁLISIS DE RESULTADOS MODELOS DE RETARDO ESPACIAL, ERROR ESPACIAL Y DE DURBIN	64
4.2.1 <i>Resultados del modelo de retardo espacial.....</i>	66
4.2.2 <i>Resultados del modelo de error espacial</i>	75
4.2.3 <i>Resultados del modelo de Durbin</i>	84
4.3 IDENTIFICACIÓN DE LAS ZONAS URBANAS DE ALTA INCIDENCIA DELICTIVAS DEL ROBO Y VARIABLES SOCIODEMOGRÁFICAS EN LA CIUDAD DE CUENCA	94
4.3.1 <i>Identificación de las zonas urbanas de alta incidencia de las variables de robo en sus diferentes modalidades.....</i>	95
4.3.2 <i>Análisis inferencial espacial de las variables sociodemográficas</i>	106
4.3.3 <i>Identificación de las zonas urbanas de alta incidencia delictiva de robos explicadas mediante las variables sociodemográficas.....</i>	134
4.4 COMPROBACIÓN DE LA HIPÓTESIS.....	149
4.4.1 <i>Prueba del índice estadístico de Aikake (AIC), para la selección del mejor modelo matemático aplicando la estadística espacial.....</i>	149
CONCLUSIONES.....	152
RECOMENDACIONES.....	153
GLOSARIO	
BIBLIOGRAFÍA	

ANEXOS

ÍNDICE DE TABLAS

Tabla 1-2: Esquema general de la estructura de una base de datos georreferenciada.....	8
Tabla 1-4: Resultado del modelo de retardo espacial de la tasa de robo a personas año 2019..	66
Tabla 2-4: Resultado del modelo de retardo espacial ajustado de la tasa de robo a personas año 2019.....	67
Tabla 3-4: Resultado del modelo de retardo espacial de la tasa de robo a domicilios año 2019.....	69
Tabla 4-4: Resultado del modelo de retardo espacial ajustado de la tasa de robo a domicilios año 2019.....	70
Tabla 5-4: Resultado del modelo de retardo espacial de unidades económicas robadas año 2019.....	71
Tabla 6-4: Resultado del modelo de retardo espacial ajustado de unidades económicas robadas año 2019.....	72
Tabla 7-4: Resultado del modelo de retardo espacial de vehículos robados año 2019.....	73
Tabla 8-4: Resultado del modelo de retardo espacial ajustado de vehículos robados año 2019.	74
Tabla 9-4: Resultado del modelo de error espacial de la tasa de robo a personas año 2019.....	75
Tabla 10-4: Resultado del modelo de error espacial ajustado de la tasa de robo a personas año 2019.....	77
Tabla 11-4: Resultado del modelo de error espacial ajustado de la tasa de robo a personas año 2019.....	77
Tabla 12-4: Resultado modelo de error espacial de la tasa de robo a domicilios año 2019.....	78
Tabla 13-4: Resultado modelo de error espacial ajustado de la tasa de robo a domicilios año 2019.....	79
Tabla 14-4: Resultado del modelo de error espacial del número de unidades económicas robadas año 2019.....	80
Tabla 15-4: Resultado del modelo de error espacial ajustado del número de unidades económicas robadas año 2019.....	81
Tabla 16-4: Resultado modelos de error espacial del número de vehículos robados año 2019.	82
Tabla 17-4: Resultado modelos de error espacial ajustado del número de vehículos robados año 2019.....	83
Tabla 18-4: Resultado del modelo espacial Durbin de la tasa de robo a personas año 2019.....	84
Tabla 19-4: Resultado del modelo espacial Durbin ajustado para la tasa de robo a personas año 2019.....	86
Tabla 20-4: Resultado del modelo espacial Durbin de la tasa de robo a domicilios año 2019..	87

Tabla 21-4: Resultado del modelo espacial Durbin ajustado de la tasa de robo a domicilios año 2019.	88
Tabla 22-4: Resultado del modelo espacial Durbin del número de unidades económicas robadas robo a personas año 2019.	90
Tabla 23-4: Resultado del modelo espacial Durbin ajustado del número de unidades económicas robadas robo a personas año 2019.	91
Tabla 24-4: Resultado del modelo espacial de Durbin de vehículos robados año 2019.	92
Tabla 25-4: Resultado del modelo espacial de Durbin ajustado de vehículos robados año 2019.	93
Tabla 26-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa de robo a personas a nivel de zona censal, ciudad Cuenca.	95
Tabla 27-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa de robo a domicilios a nivel de zona censal, ciudad Cuenca.	98
Tabla 27-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de robos a unidades económicas a nivel de zona censal, ciudad Cuenca.	101
Tabla 29-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de robos a vehículos a nivel de zona censal, ciudad Cuenca.	103
Tabla 30-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de habitantes (población) a nivel de zona censal, ciudad Cuenca.	106
Tabla 31-4: Test de <i>I Moran</i> , matriz de pesos (W) utilizando simulación de Montecarlo para el número de Emigrantes a escala internacional a nivel de zona censal, ciudad Cuenca.	108
Tabla 32-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para tasa de inmigrantes extranjeros presentes en el censo a nivel de zona censal, ciudad Cuenca.	110
Tabla 33-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el promedio de personas por hogar a nivel de zona censal, ciudad Cuenca.	112
Tabla 34-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa de desplazamiento temporal de personas a nivel de zona censal, ciudad Cuenca.	114
Tabla 35-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de inmigrantes extranjeros a nivel de zona censal, ciudad Cuenca.	116
Tabla 36-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de emigrantes internacionales a nivel de zona censal, ciudad Cuenca. ...	118
Tabla 37-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de personas en edad de trabajar (PET) a nivel de zona censal, ciudad Cuenca.	120

Tabla 38-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de personas cesantes a nivel de zona censal, ciudad Cuenca.	122
Tabla 39-4: Test de I Moran, matriz de pesos (W) simulación de Montecarlo para el número de personas ocupadas que realizó su actividad económica dentro del hogar por zona en Cuenca.	124
Tabla 40-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de personas económicamente inactiva (PEI) a nivel de zona, ciudad Cuenca, 2019.	126
Tabla 41-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa de analfabetismo a nivel de zona censal, ciudad Cuenca.	128
Tabla 42-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa neta de asistencia escolar nivel de zona censal, ciudad Cuenca.	130
Tabla 43-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de personas que tienen vivienda propia y pagada a nivel de zona en Cuenca.	132
Tabla 44-4: Indicador AIC, para seleccionar el mejor modelo especial para identificar las zonas de mayor incidencia de la tasa de robo a personas en la ciudad de Cuenca año 2019.	150
Tabla 45-4: Indicador estadístico AIC, para seleccionar el mejor modelo que permite identificar las zonas de mayor incidencia de la tasa de robo a domicilios en Cuenca año 2019.	150
Tabla 46-4: Indicador estadístico AIC, para seleccionar el mejor modelo que permite identificar las zonas de mayor incidencia de robo a unidades económicas en Cuenca año 2019.	151
Tabla 47-4: Indicador estadístico AIC, para seleccionar el mejor modelo que permite identificar las zonas de mayor incidencia de robo a vehículos en la ciudad de Cuenca año 2019.	151

ÍNDICE DE FIGURAS

Figura 1-2: Representación de un mapa de coropletas para un intervalo de clases naturales.....	9
Figura 2-2: Representación de un mapa de coropletas para un intervalo de clases cuartiles.....	10
Figura 3-2: Representación de un mapa de coropletas para un intervalo de clases iguales.....	10
Figura 4-2: Diferentes definiciones de casos de relación de vecindad.	13
Figura 1-4: Mapa de coropletas división político administrativa de Cuenca con sus parroquias urbanas con sus respectivas zonas censales.....	47
Figura 2-4: Mapa de coropletas, distribución por zonas de robo a personas, ciudad de Cuenca 2019.	48
Figura 3-4: Mapa de coropletas, distribución por zonas de robo a domicilios, ciudad de Cuenca 2019.	49
Figura 4-4: Mapa de coropletas, distribución por zonas de robo a unidades económicas, ciudad de Cuenca 2019.	50
Figura 5-4: Mapa de coropletas, distribución por zonas de robo a vehículos, ciudad de Cuenca 2019.	51
Figura 6-4: Mapa de coropletas, distribución por zonas de la población, ciudad de Cuenca 2010.	52
Figura 7-4: Mapa de coropletas, distribución por zonas de la población inmigrante a escala internacional, ciudad de Cuenca 2010.	53
Figura 8-4: Mapa de coropletas, distribución por zonas tasa de población inmigrante extranjero, ciudad de Cuenca 2010.	54
Figura 9-4: Mapa de coropletas, distribución por zonas promedio de personas por hogar, ciudad de Cuenca 2010.	55
Figura 10-4: Mapa de coropletas, distribución por zonas del número de inmigrantes extranjeros presentes en el censo, ciudad de Cuenca 2010.	56
Figura 11-4: Mapa de coropletas, distribución por zonas de la tasa de desplazamiento temporal de personas por trabajo o estudio, ciudad de Cuenca 2010.	57
Figura 12-4: Mapa de coropletas, distribución por zonas del número de emigrantes internacional, ciudad de Cuenca 2010.	58
Figura 13-4: Mapa de coropletas, distribución por zonas de población en edad de trabajar, ciudad de Cuenca 2010.	59
Figura 14-4: Mapa de coropletas, distribución por zonas promedio de personas por hogar, ciudad de Cuenca 2010.	60
Figura 15-4: Mapa de coropletas, distribución por zonas población que realiza su actividad económica dentro del hogar, ciudad de Cuenca 2010.	60

Figura 16-4: Mapa de coropletas, distribución por zonas PEI, ciudad de Cuenca 2010.....	61
Figura 17-4: Mapa de coropletas, distribución por zonas tasa de analfabetismo personas de 10 años y más, ciudad de Cuenca 2010.	62
Figura 18-4: Mapa de coropletas, distribución por zonas tasa de tasa de asistencia escolar población de 24 años y más, ciudad de Cuenca 2010.....	62
Figura 19-4: Mapa de coropletas, distribución por zonas hogares según tenencia de la vivienda propia y totalmente pagada, ciudad de Cuenca 2010.	63
Figura 20-4: Mapa de la ciudad de Cuenca por zonas urbanas, modelo de vecindad, criterio de Rook.	65
Figura 21-4: Mapa de la ciudad de Cuenca por zonas urbanas, modelo de vecindad, criterio de Queen.....	65
Figura 22-4: Mapa de la ciudad de Cuenca por zonas urbanas, modelo de vecindad, criterio de K-cercanos.....	66
Figura 23-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> global, tasa de robo a personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	96
Figura 24-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de la tasa de robo a personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	97
Figura 25-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	97
Figura 26-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , tasa de robo a domicilios por zonas censales urbanas de la ciudad de Cuenca, 2019.	99
Figura 27-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de la tasa de robo a domicilios por zonas censales urbanas de la ciudad de Cuenca, 2019.	99
Figura 28-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a domicilios por zonas censales urbanas de la ciudad de Cuenca, 2019.	100
Figura 29-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , número de robo a unidades económicas por zonas censales urbanas de la ciudad de Cuenca, 2019.	101
Figura 30-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores del número de robo a unidades económicas por zonas censales urbanas de la ciudad de Cuenca, 2019.	102
Figura 31-4: Mapa de patrones geográficos de agrupación, para los diferentes valores del número de robo a unidades económicas por zonas censales urbanas de la ciudad de Cuenca, 2019.	102
Figura 32-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , número de robo a vehículos por zonas censales urbanas de la ciudad de Cuenca, 2019.	104
Figura 33-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores del número robo a vehículos por zonas censales urbanas de la ciudad de Cuenca, 2019.	105

Figura 34-4: Mapa de patrones geográficos de agrupación, para los diferentes valores del número de robo a vehículos por zonas censales urbanas de la ciudad de Cuenca, 2019.	105
Figura 35-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , número de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	106
Figura 36-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de población por zonas censales urbanas de la ciudad de Cuenca, 2019.	107
Figura 37-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la población por zonas censales urbanas de la ciudad de Cuenca, 2019.	107
Figura 38-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , emigrantes a escala internacional por zonas censales urbanas de la ciudad de Cuenca, 2019.	108
Figura 39-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de emigración a escala internacional por zonas censales urbanas de la ciudad de Cuenca, 2019... ..	109
Figura 40-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de emigración a escala internacional por zonas censales urbanas de la ciudad de Cuenca, 2019.	109
Figura 41-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , tasa de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.	110
Figura 42-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de la tasa de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.	111
Figura 43-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.	111
Figura 44-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , promedio de personas por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.	112
Figura 45-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de promedio de personas por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019....	113
Figura 46-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de promedio de personas por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.	113
Figura 47-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	114
Figura 48-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de la tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	115

Figura 49-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	115
Figura 50-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , número de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.	116
Figura 51-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores del número de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.	117
Figura 52-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.	117
Figura 53-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , número de emigrantes internacionales por zonas censales urbanas de la ciudad de Cuenca, 2019.	118
Figura 54-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores del número de emigrantes internacionales por zonas censales urbanas de la ciudad de Cuenca, 2019.	119
Figura 55-4: Mapa de patrones geográficos de agrupación, para los diferentes valores del número de emigrantes internacionales por zonas censales urbanas de la ciudad de Cuenca, 2019.	119
Figura 56-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , número de personas en edad de trabajar por zonas censales urbanas de la ciudad de Cuenca, 2019.	120
Figura 57-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de población en edad de trabajar por zonas censales urbanas de la ciudad de Cuenca, 2010.	121
Figura 58-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población en edad de trabajar por zonas censales urbanas de la ciudad de Cuenca, 2019.	121
Figura 59-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , población cesante por zonas censales urbanas de la ciudad de Cuenca, 2019.	122
Figura 60-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de población cesante por zonas censales urbanas de la ciudad de Cuenca, 2010.	123
Figura 61-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población cesante por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.	123
Figura 62-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , población ocupada que realizó actividad económica dentro del hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.	124

Figura 63-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de población ocupada que realice actividad económica en el hogar por zonas de la ciudad de Cuenca, 2019.	125
Figura 64-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población ocupada que realizó actividad económica en el hogar por zonas urbanas en Cuenca.	125
Figura 65-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , población económicamente inactiva (PEI) por zonas censales urbanas de la ciudad de Cuenca, 2019.	126
Figura 66-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de población económicamente inactiva por zonas censales urbanas de la ciudad de Cuenca, 2010.	127
Figura 67-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población económicamente inactiva (PEI) por zonas urbanas de la ciudad de Cuenca, 2019.	127
Figura 68-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , tasa de analfabetismo por zonas censales urbanas de la ciudad de Cuenca, 2010.	128
Figura 69-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de la tasa de analfabetismo por zonas censales urbanas de la ciudad de Cuenca, 2010.	129
Figura 70-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de analfabetismo por zonas censales urbanas de la ciudad de Cuenca, 2019.	129
Figura 71-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , tasa neta de asistencia escolar por zonas censales urbanas de la ciudad de Cuenca, 2010.	130
Figura 72-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de la tasa neta de asistencia escolar por zonas censales urbanas de la ciudad de Cuenca, 2010.	131
Figura 73-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa neta de asistencia escolar por zonas censales urbanas de la ciudad de Cuenca, 2019.	131
Figura 74-4: Gráfico de correlación del índice <i>I</i> de <i>Moran</i> , hogares que tienen vivienda propia y totalmente pagada por zonas censales urbanas de la ciudad de Cuenca, 2019. .	132
Figura 75-4: Mapa del test de <i>Moran</i> a nivel local, para los diferentes valores de hogares que tienen vivienda propia y totalmente pagada por zonas urbanas de la ciudad de Cuenca, 2010.	133
Figura 76-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de hogares que tienen vivienda propia y pagada por zonas de la ciudad de Cuenca, 2019.	133

Figura 77-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población inmigrante extranjera por zonas de la ciudad de Cuenca, 2019.	134
Figura 78-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs promedio de personas por hogar, zonas urbanas de la ciudad de Cuenca, 2019.	135
Figura 79-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas vs población cesante, zonas urbanas de la ciudad de Cuenca, 2019.	136
Figura 80-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs hogares que tienen vivienda propia y pagada por zonas de Cuenca, 2019.	137
Figura 81-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs población inmigrante a escala internacional por zonas de Cuenca, 2019.	138
Figura 82-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs tasa de inmigrantes internacionales por zonas de la ciudad de Cuenca, 2019.	139
Figura 83-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a domicilios vs población cesante por zonas urbanas de la ciudad de Cuenca, 2019.	140
Figura 84-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs población económicamente inactiva por zonas de la ciudad de Cuenca, 2019.	141
Figura 85-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs hogares que tienen vivienda propia y pagada por zonas de Cuenca, 2019.	142
Figura 86-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población inmigrante extranjera por zonas de la ciudad de Cuenca, 2019.	143
Figura 87-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas vs promedio de personas por hogar, zonas de la ciudad de Cuenca, 2019.	144
Figura 88-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población que realiza actividad económica el hogar por zonas de Cuenca, 2019.	145

Figura 89-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs tasa de desplazamiento temporal por estudio o trabajo por zonas de Cuenca, 2019.	146
Figura 90-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población migrante internacional por zonas de la ciudad de Cuenca, 2019.	147
Figura 91-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población en edad de trabajar por zonas de la ciudad de Cuenca, 2019.	148
Figura 92-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas vs población cesante por zonas urbanas de la ciudad de Cuenca, 2019.	149

RESUMEN

El objetivo fue identificar zonas geográficas con alta incidencia delictual, específicamente el delito de robo en sus diferentes modalidades como a: personas, domicilios, unidades económicas y vehículos; registrados en la ciudad de Cuenca, Ecuador en el año 2019, a través de la aplicación de modelos matemáticos, mediante estadística espacial para datos de área. Los modelos matemáticos espaciales que se eligieron fueron de Retardo Espacial, Error Espacial y Durbin. Estos se aplicaron mediante el *software* estadístico R mediante su entorno RStudio, utilizando los paquetes espaciales correspondientes, mismo que permitieron identificar las zonas censales urbanas que tienen alta incidencia. Además, se analizó la relación con las variables sociodemográficas al mismo nivel geográfico donde se calcularon estos indicadores, para relacionarlos con el delito de robo en las modalidades antes mencionadas; analizando, además, la auto correlación espacial y el grado de relación que existe en las variables correspondientes a indicadores sociodemográficos con los delitos. Se determinó que, efectivamente, estos presentan una estrecha relación con los delitos ocurridos en cada una de las zonas de la ciudad; Los diferentes modelos espaciales para estimar adecuadamente este tipo de concentración en el área geográfica de estudio. Se concluyó que, el mejor modelo para estimar esta problemática es el Modelo de error espacial (SEM) en los delitos de robo a personas y a vehículos, el Modelo espacial autorregresivo (SAR) domicilios y el Modelo espacial de Durbin (SDM) unidades económicas ya que estos son los de mejor ajuste, los cuales se comprobó mediante el indicador estadístico de calidad Akaike (AIC). Se recomienda a los investigadores o analistas de las diferentes áreas, que se utilice este tipo de modelo espacial para determinar concentraciones de este delito en los distintos períodos de tiempo en cualquier área geográfica empleada en este estudio.

Palabras claves: <MATEMÁTICAS>, >AUTOCORRELACIÓN ESPACIAL>, <DATOS DE ÁREA>, <ÍNDICE AIC>, <MODELO DE RETARDO ESPACIAL (SAR)>, <MODELO DE ERROR ESPACIAL (SEM)>, <MODELO ESPACIAL DE DURBIN (SDM)>, <RSPATIAL>.



28-06-2023

0067-DBRA-UPT-IPEC-2023

SUMMARY

The objective was to identify geographical areas with a high incidence of crime, specifically the crime of robbery in its different modalities such as people, homes, economic units and vehicles, registered in the city of Cuenca, Ecuador in 2019, through the application of mathematical models using spatial statistics for area data. The spatial mathematical models chosen were Spatial Delay, Spatial Error and Durbin. These were applied using the R statistical software through its RStudio environment, using the appropriate spatial packages, which allowed the identification of urban census areas with high incidence. In addition, the relationship with socio-demographic variables was analyzed at the same geographical level where these indicators were calculated, in order to relate them to the crime of robbery in the above-mentioned modalities, analyzing also the spatial autocorrelation and the degree of relationship that exists in the variables corresponding to the socio-demographic indicators with the crimes. It was found that, in fact, these present a close relationship with the crimes that occurred in each of the zones of the city; the different spatial models to adequately estimate this type of concentration in the geographic area of study. It was concluded that the best model to estimate this problem is the Spatial Error Model (SEM) for the crimes of robbery of persons and vehicles, the Spatial Autoregressive Model (SAR) for the residences and the Durbin Spatial Model (SDM) for the economic units, since these are the best fit, which was verified by means of the Akaike Statistical Quality Indicator (AIC). It is recommended that researchers or analysts in different fields use this type of spatial model to determine concentrations of this crime in different time periods in each geographical area used in this study.

Key words: <MATEMATICS>, <SPATIAL AUTOCORRELATION>, <AREA DATA>, <AIC INDEX>, <SPATIAL LAG MODEL (SAR)>, <SPATIAL ERROR MODEL (SEM)>, <SPATIAL DURBIN (SDM)>, <SPATIAL <RSPATIAL>.

CAPITULO I

1 INTRODUCCIÓN

Situación problemática

En la actualidad, en las diferentes ciudades, que tienen un considerable crecimiento poblacional, es importante y necesario aplicar métodos y modelos matemáticos mediante la estadística espacial, los mismos que permitirán explicar de alguna manera estos fenómenos sociales que ocurren en estas y como se relacionan.

En el Observatorio de Seguridad Ciudadana del cantón Cuenca, se realizan estudios asociados a este tipo de problemática, sin embargo, el uso de modelos matemáticos y herramientas de *software* específico para el análisis y aplicación de modelos espaciales, aún no han logrado desarrollarse de manera adecuada. De esta forma se busca realizar su implementación, de acuerdo a las exigencias tecnológicas actuales lo cual hace que estas sean imprescindibles.

En este trabajo, se pretende aplicar modelos matemáticos mediante estadística espacial, que permitan analizar de manera confiable datos con estas características, para identificar las diferentes zonas urbanas de alta incidencia delictiva. Particularmente el problema a tratar es el delito de robo cuyos eventos se han registrado en la ciudad de Cuenca, Ecuador. La información a utilizar en esta investigación corresponde a la recolectada en el año 2019. Ésta permitirá identificar cambios en las concentraciones delictivas de robos en las diferentes zonas urbanas en la ciudad de Cuenca, mismas que se mostrarán mediante mapas de coropletas a un nivel geográfico de zona censal mostrando lugares específicos de la ciudad. Los mapas son herramientas muy útiles, que facilitan el análisis visual del espacio, o de lo que está ocurriendo espacialmente en las diferentes zonas urbanas de la ciudad. Se debe tener presente que un mapa de coropletas, siempre se convierte en una herramienta poderosa al momento de analizar un problema y además los valores numéricos obtenidos o calculados confirman lo que se refleja en los mapas.

Las unidades geográficas que se emplearon en el estudio corresponden a zonas censales urbanas de la ciudad de Cuenca, de acuerdo al Instituto Nacional de Estadística y Censos (INEC), los mismos que han homologado los límites de las mismas en relación al CONALI y el Municipio de Cuenca, de acuerdo a la delimitación cantonal y parroquial en el área urbana de la ciudad que es la unidad de desagregación geográfica más pequeña que manejan estas entidades; mientras que el INEC, maneja una desagregación geográfica de zonas o aún más pequeñas como sectores y manzanas censales, respetando las delimitaciones parroquiales.

Formulación del problema

¿Qué modelo matemático mediante Estadística Espacial, permitirá determinar las zonas urbanas de alta incidencia en el delito del robo en la ciudad de Cuenca, Ecuador en el año 2019 de una manera más confiable?

Preguntas directrices

¿Cómo ayudará la modelación matemática mediante estadística espacial en la identificación de las características del delito de robo en la ciudad de Cuenca durante el año 2019?

¿Permitirán los indicadores del delito del robo en la ciudad de Cuenca, medir cuantitativamente las concentraciones en el año 2019?

¿Cómo ayudará, la aplicación de modelos matemáticos, mediante estadística espacial a identificar zonas urbanas de la ciudad de Cuenca, con la alta concentración de robos durante el año 2019?

Justificación de la investigación

Los modelos matemáticos mediante estadística espacial que existen, se han aplicado en diferentes investigaciones y distintas áreas. Estos son instrumentos válidos para dar una explicación estadística, por ejemplo, la concentración de la delincuencia. En tal contexto, esta investigación se enfocará en el delito de robo en determinadas zonas urbanas de la ciudad de Cuenca, siendo ésta una alternativa para analizar la incidencia de este fenómeno, con el fin de mejorar y generar avances en los estudios relacionados con la seguridad ciudadana.

El aporte que brindará este trabajo, es proporcionar un instrumento mediante la aplicación de este tipo de modelos matemáticos, que permita identificar las zonas geográficas con mayor incidencia en el delito del robo, los mismos que periódicamente podrán ser actualizados, para mostrar, si este delito se mantiene o varía espacialmente en las diferentes zonas urbanas de la ciudad de Cuenca, Ecuador.

Para cumplir el objetivo planteado, se han aplicado métodos y modelos matemáticos (estadísticos espaciales), realizando previamente un análisis de las características de los robos mediante la base de datos que posee actualmente el Observatorio de Seguridad Ciudadana del cantón Cuenca. Las bases de datos son retroalimentadas de manera periódica de los expedientes de las denuncias registradas por la Fiscalía General del Estado de la provincia del Azuay. Se consideran las distintas zonas urbanas de la ciudad, además de establecer indicadores sobre el delito de robo; dichos indicadores permitirán medir de manera cuantitativa su concentración. Es importante indicar que existen algunos trabajos realizados a nivel nacional como internacional sobre análisis espacial delincuencia; los mismos que servirán como referencia. En estos trabajos mediante la aplicación de los diferentes métodos y modelos estadísticos espaciales, ha

sido posible llegar a determinar concentraciones, con el objetivo de aportar en la generación de explicaciones más claras de un evento o problema en cualquier ámbito; en este caso se lo aplicará en el campo de la seguridad ciudadana.

La presente investigación genera un aporte que es muy importante, ya que se aplicarán los modelos matemáticos de regresión espacial como: retardo espacial, error espacial y de Durbin. Al aplicar este tipo de modelos matemáticos se determinarán concentraciones del delito en las zonas urbanas de la ciudad. El periodo de la información usada en este estudio, para el análisis y aplicación de modelos es 2019. Se espera que este estudio sirva como referente para que el Observatorio de Seguridad Ciudadana del cantón Cuenca lo replique en los demás delitos restantes registrados en la ciudad y, además, que sea un instrumento útil para contribuir en el análisis de información de datos espaciales sobre la delincuencia, para de esta manera aportar con este tipo de métodos de análisis en beneficio de la ciudad.

1.1 Objetivo general

Aplicar modelos matemáticos mediante estadística espacial para identificar zonas de alta incidencia en el delito del robo, en la ciudad de Cuenca, Ecuador en el año 2019.

1.2 Objetivos específicos

- a) Identificar las características de delito del robo distribuidos en las zonas urbanas en la ciudad de Cuenca durante el año 2019, analizando la base de datos proporcionada por el Consejo de Seguridad Ciudadana del cantón Cuenca, Ecuador.
- b) Establecer indicadores del delito del robo en la ciudad de Cuenca del año 2019, para medir cuantitativamente su concentración a nivel de zonas urbanas.
- c) Aplicar modelos matemáticos, mediante estadística espacial para identificar las zonas urbanas de la ciudad de Cuenca, que permitan visualizar la alta concentración de los robos en el año 2019.
- d) Determinar, cual es el mejor modelo matemático, mediante estadística espacial que permita visualizar la alta concentración de los robos en las zonas urbanas de la ciudad de Cuenca en el año 2019.

1.3 Hipótesis

La Estadística Espacial mediante el modelo más idóneo permitirá identificar con mayor exactitud las zonas urbanas de alta incidencia de robos, registrados en el año 2019, en la ciudad de Cuenca-Ecuador.

CAPITULO II

2 MARCO TEÓRICO

2.1 Antecedentes del estudio

La aplicación de los modelos matemáticos de estadística espacial es un tema bastante amplio para abordarlo en la actualidad y se ha convertido en uno de los instrumentos de mayor utilidad en todos los campos; específicamente en tratar de observar o visualizar de manera objetiva los comportamientos que existe en las diferentes áreas geográficas en un país, cantón, ciudad, etc. Morales Oñate. & Morales (2019) en el trabajo denominado “Innovación en Ecuador – un enfoque espacial”, (p. 327 - 329), plantean explorar la distribución espacial del éxito de las empresas en el Ecuador, su objetivo principal fue ver si existen relaciones espaciales entre sus provincias y además resaltar las políticas estatales que contribuyan a la innovación, en el que efectivamente al aplicar la técnica de espacial autoregresivo, con error espacial y espacial de Durbin, se obtuvo un resultado, donde claramente existe una relación espacial en el éxito de innovación mediante apoyo del gobierno y la banca privada.

Para realizar la aplicación de este tipo de modelos, es muy importante el tipo de dato que se utilizará, para lo cual Cabrero Ortega & García Pérez (2015) mencionan que los datos deben ser de tipo punto, línea y polígono. El libro publicado “Análisis estadístico de datos espaciales con QGIS y R”, contiene una explicación más amplia y clara con las definiciones de la estructura de los datos georreferenciados aplicado al análisis y del modelado espacial, en cada uno de los diferentes casos y, específicamente en el tipo de datos de área. (Cabrero Ortega & García Pérez, 2015). También es importante mencionar la cantidad de datos que se puede manejar. En este sentido Aguayo Téllez & Medellín Mendoza (2014, p. 71), realizan la aplicación de modelos de regresión espacial, con un periodo de un año, como lo muestran en el estudio publicado “Dependencia espacial de la delincuencia en Monterrey, México”; en el cual presentan un análisis espacial para el año 2010 de los vecindarios en el Área Metropolitana de Monterrey, México. Se investigó el efecto de las características del entorno en la delincuencia y, al aplicar la técnica de modelo de regresión espacial, se observó, que existe una alta probabilidad de ser víctima y además, de que los jóvenes se conviertan en delincuentes, ya que existe un alto índice de jóvenes desempleados y está directamente relacionado con la existencia de centros comerciales en las áreas vecinas. (Aguayo Téllez & Medellín Mendoza, 2014).

De manera general es conocido que, la aplicación de las diferentes técnicas y modelos espaciales, permiten detectar asociaciones espaciales. En el artículo de Herrera (2015, p. 4), denominado “Econometría Espacial usando Stata. Breve Guía aplicada para datos de corte

transversal”, se manifiesta que, al aplicar técnicas y modelos espaciales, efectivamente existen asociaciones espaciales, concentraciones locales, detección de datos anómalos (*outliers*). Al aplicar el análisis exploratorio de datos espaciales, estos hablan por sí mismos, sin imponer una estructura conceptual previa, ya que, al tratar de identificar una incidencia relacional de las variables, estas por su naturaleza generan autocorrelación espacial, por lo que es importante aplicar ciertos indicadores y mostrar mediante mapas de coropletas, la existencia de patrones o concentraciones. (Herrera, 2015). La aplicación de modelos de regresión espacial lineal y múltiple, se demuestra que las dos variables utilizadas siguen una estructura espacial y que estas se deben determinar de manera explícita, dado que puede existir una autocorrelación espacial, por lo que es importante utilizar otras variables explicativas para quitar la existencia de dependencia espacial o se la anule, con respecto a la variable dependiente.

La modelación de estadística espacial no se emplea únicamente en el sector social, minero, entre otros, como se muestra en el trabajo realizado por Franco Galeano (2018, p. 15), el mismo que se denomina “Análisis Espacial de la Criminalidad a nivel Intra-Urbano-El Caso de Medellín, Colombia”. En este trabajo se presentan modelos matemáticos mediante estadística espacial para determinar las zonas de mayor criminalidad en la ciudad de Medellín. Para determinar estas zonas trabajan con la información de la Secretaria de Seguridad de Medellín del año 2016, utilizando como variable dependiente la tasa de homicidios distribuidos en las diferentes zonas y como independientes las variables socioeconómicas. También trabaja con tres clases de modelos espaciales, estos son: de regresión espacial por mínimos cuadrados ordinarios (MCO), modelo autorregresivo (SAR ó spatial lag) y de error espacial (SEM ó spatial error). Estos modelos parten de la correlación entre las regiones analíticas a través del espacio y como los datos espaciales pueden mostrar dependencia en las variables y en el error.

Mediante la variedad de estudios realizados sobre el identificar las concentraciones en diferentes zonas geográficas de un determinado lugar, en los distintos problemas presentados como por ejemplo en la minería, hidrología, cultural, socioeconómico y específicamente la delincuencia, se puede observar que se conoce un amplio conjunto de técnicas y estructuras para modelos de estadística espacial bajo el criterio de mínimos cuadrados, error espacial y retardo espacial, entre otros que es el objetivo del presente trabajo de titulación.

2.2 Bases Teóricas

2.2.1 Estadística Espacial

La estadística espacial es una rama de la estadística general, que trata del análisis de datos georreferenciados, es decir datos disponibles en coordenadas espaciales. La estadística espacial aplica diferentes técnicas de estadística descriptiva e inferencial en el espacio; las mismas que

son mencionadas por un conjunto de investigadores, en el documento titulado: “Nuevas técnicas de estadística espacial para la detección de clusters residenciales de población inmigrante”. En el cual generalizan la importancia de la aplicación en diferentes campos de la Estadística espacial, donde ya se introduce el espacio como un elemento fundamental para el análisis. (Martori & Hoberg, 2008).

Formalizando la definición de estadística espacial; corresponde al análisis de ejecuciones de un proceso estocástico $\{Z(s) = s \in D\}$, en el que $s \in R^d$ representa una ubicación en el espacio euclidiano d-dimensional, $Z(s)$ es una variable aleatoria en la ubicación s y s varía sobre un conjunto de índices $D \subset R^d$. (Giraldo Henao, 2015).

Por lo tanto; la estadística espacial en la actualidad, es una de las metodologías más innovadoras que se están utilizando para visualizar y analizar en el campo de la seguridad ciudadana a nivel mundial y continental (por ejemplo Latinoamérica), existiendo también algunos estudios realizados en nuestro país como en las ciudades de Quito y Guayaquil, lo cual permite contar con una referencia para aplicarla en este trabajo y así determinar la concentración del delito de robo con la aplicación de modelos matemáticos mediante estadística espacial en las diferentes zonas urbanas de la ciudad de Cuenca; para lo cual es de suma importancia disponer de datos que cuenten con las propiedades y características adecuadas con el fin de aplicar correctamente este conjunto de herramientas y técnicas.

2.2.1.1 Datos Espaciales

Para realizar análisis y aplicación de modelos espaciales, son muy importantes los datos que se van a utilizar. En este sentido, por la naturaleza de los datos espaciales es imprescindible diferenciar los campos discretos o continuos de las variables o mediciones. Si existe un espacio continuo (modelo basado en campo) los valores de las variables o mediciones se tienen que interpretar de forma continua dado que la continuidad del campo no se puede preservar en valores discretos. En el caso de que si existe un espacio discreto (modelo de objetos) o si un espacio continuo se ha convertido en discreto, estos valores de las variables se las puede valorar de forma discreta o continua. (Borrego Sánchez, 2018).

Bajo la definición mencionada anteriormente, se puede decir que existen cuatro tipos de datos: de patrones puntuales, de campo, de área y de interacción espacial.

Para el desarrollo de este trabajo nos enfocaremos en los datos de tipo área donde los valores de las variables (observaciones) se relacionan con unidades de área (modelos basados en entidades). Por lo que; para esto, también es necesario definir como está estructurada una base de datos georreferenciados, lo que quiere decir que sus mediciones tienen características en cada zona, implícitamente asociadas las coordenadas geográficas de la ubicación de los lugares o

sitios donde estas fueron tomadas u ocurrieron uno o más eventos. Cuando el área de estudio es considerablemente grande se usa un geoposicionador para establecer dichas coordenadas. En otros casos es necesario por ejemplo en diseños experimentales como parcelas, hacer asignaciones según el plano cartesiano. Las coordenadas pueden ser de diferentes tipos: planas, geográficas (grados, minutos y segundos) o cartesianas. Hay que aclarar que esto no depende del tipo de software que se vaya a utilizar para realizar el análisis. A continuación, se muestra un esquema general de una base de datos georreferenciados. Ver Tabla 1-2. (Giraldo Henao, 2015).

Tabla 1-2: Esquema general de la estructura de una base de datos georreferenciada.

Lugar	Latitud Norte	Longitud Este	X ₁	X ₂	.	.	.	X _p
1	-	-	X ₁₁	X ₁₂	.	.	.	X _{1p}
2	-	-	X ₂₁	X ₂₂	.	.	.	X _{2p}
3	-	-	X ₃₁	X ₃₂	.	.	.	X _{3p}
4	-	-	X ₄₁	X ₄₂	.	.	.	X _{4p}
.	-	-
.	-	-
.	-	-
n	-	-	X _{n1}	X _{n2}	.	.	.	X _{np}

Fuente: Base de datos georreferenciada CSC, 2019.

Realizado por: Velastegui Casco, Edy, 2021.

donde, n es el número de lugares o sitios, p es el número de variables medidas en cada uno de los lugares o sitios, X_{ij} , corresponde a la medida de la variable X_j ($j = 1, 2, 3, \dots, p$) en el lugar o sitio i ($i = 1, 2, 3, \dots, n$), este puede ser cuantitativo o categórico.

Definición de datos de área: Este tipo de valores de datos se asocian con observaciones que tienen un número fijo de unidades de área (objetos de área) que pueden formar una red regular, como con imágenes de detección remota, o un conjunto de zonas irregulares, como por ejemplo; países, provincias o departamentos, cantones, parroquias o condados, zonas censales, hasta una desagregación geográfica más pequeña que pueda existir. (Borrego Sánchez, 2018).

2.2.1.2 Visualización de datos Espaciales

En el análisis exploratorio de datos espaciales, el mapa es el mejor medio o instrumento establecido y convencional para visualizar los datos de área. El más utilizado es el mapa de coropletas, el mismo que tiene una característica, esta permite que cada una de las áreas muestre un color de relleno o sombreado, de acuerdo a una escala directa mediante un valor o un indicador de las variables de interés (atributos) dentro de cada área definida. El número de clases y los intervalos de las clases pueden estar fundamentados en varios o diferentes criterios. Para calcular el número de clases de una manera formal, algunos investigadores recomiendan

aplicar la regla descrita mediante la ecuación (1) de forma generalizada, dado que el número de clases depende de la cantidad de datos existentes. (Borrego Sánchez, 2018).

$$K = 1 + 3.3\ln(n) \quad (1)$$

Dónde: K es el número de clases calculadas, \ln es el logaritmo natural en base 10, y n es el número de áreas.

En referencia a lo mencionado anteriormente con respecto a la selección de intervalos, existen cuatro tipos los cuales se mencionan a continuación:

Divisiones naturales: En este caso el número de clases se define según algunas agrupaciones naturales propias de los valores de las variables. Las diferentes divisiones que se generan siempre se imponen a partir de puntos de ruptura que de antemano se conoce son relevantes (asignación deductiva) o también se lo puede realizar mediante la utilización de herramientas de SIG que buscan saltos importantes en los valores de las variables. Ver Figura 1-2.

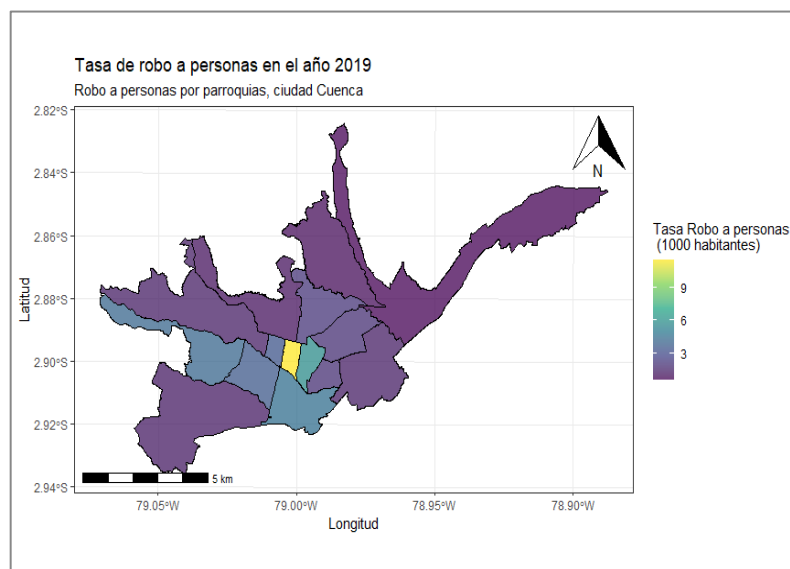


Figura 1-2: Representación de un mapa de coropletas para un intervalo de clases naturales.

Fuente: Base de datos Observatorio de Seguridad Ciudadana cantón Cuenca 2019, CSC y Base de datos VII Censo de Población y VI de Vivienda 2010, INEC.

Realizado por: Velasteguí Casco, Edy, 2021.

Divisiones por cuartiles: En este caso cada clase debe tener un número igual de observaciones. La clasificación de cuartiles (cuatro categorías) y quintiles (cinco categorías), son los que comúnmente se utilizan por la mayoría de investigadores para representar y analizar los valores de una variable en estadística. Ver Figura 2-2.

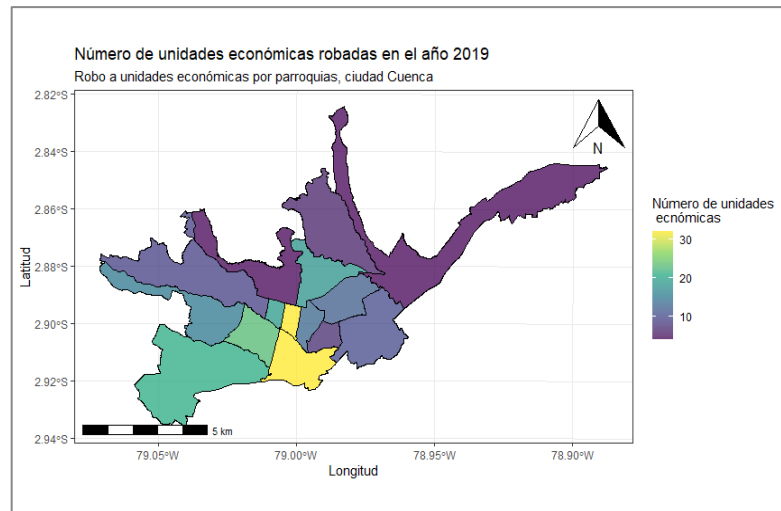


Figura 2-2: Representación de un mapa de coropletas para un intervalo de clases cuartiles.

Fuente: Base de datos Observatorio de Seguridad Ciudadana cantón Cuenca 2019, CSC.

Realizado por: Velasteguí Casco, Edy, 2021.

División de intervalos iguales: En este caso las clases tienen una validez cuando las observaciones se distribuyen de manera uniforme en su rango. Pero si se presenta el caso de que los datos son sesgados, esto ocasionará que un gran número de observaciones se introduzcan en unas pocas clases. Ver Figura 3-2.

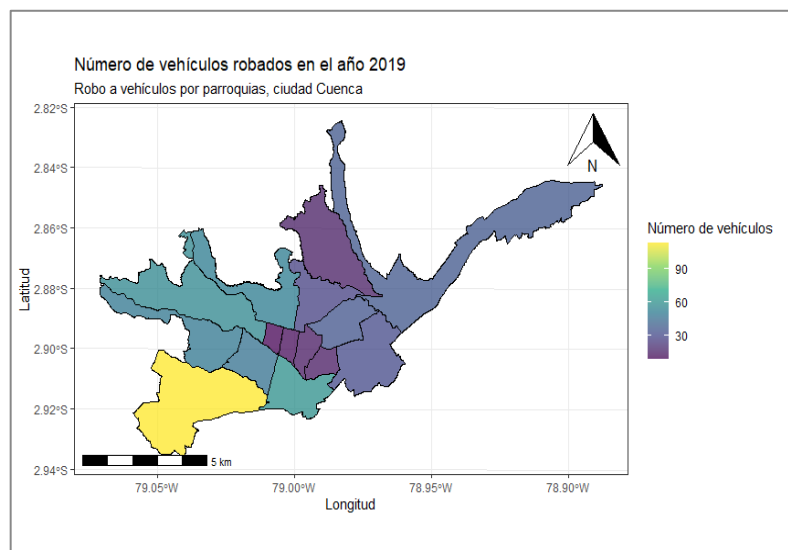


Figura 3-2: Representación de un mapa de coropletas para un intervalo de clases iguales.

Fuente: Base de datos Observatorio de Seguridad Ciudadana cantón Cuenca 2019, CSC.

Realizado por: Velasteguí Casco, Edy, 2021.

División según desviación estándar: En este caso el número de clases se basan en intervalos que se distribuyen alrededor de la media en unidades de desviación estándar.

Es importante aclarar que los mapas de coropletas tienen algunas características que ocasionan problemas que deben ser mencionados antes de su construcción; mismos que se mencionan a continuación según, (Borrego Sánchez, 2018).

Principales problemas de los mapas de coropletas antes de su construcción:

- a) Estos mapas traen una implicación visual de uniformidad dentro del área de los valores variables.
- b) También el mapa convencional permite que cualquier área físicamente grande domine la pantalla, de una manera que puede ser inapropiada para el tipo de datos que se esté mapeando. Por ejemplo, al realizar el mapeo de datos socioeconómicos, las áreas rurales más grandes y bastantes dispersas en población pueden dominar el mapa, pero el interés real puede estar en las áreas más pequeñas físicamente, como las áreas urbanas que generalmente son las más pobladas.
- c) Se debe tener mucho cuidado con el problema de unidad de área modificable (MAUP) que surge a menudo en este tipo de mapas. Esto se presenta cuando las variables de interés surgen de la agregación de datos individuales de forma arbitraria a las áreas y se observa cualquier patrón; por un lado, en los límites de área elegidos como en la distribución espacial. Esto proviene de dicha arbitrariedad de los valores de las variables analizadas, no teniendo ninguna certeza ni formalidad al estudiar estos datos con la hipótesis de seguir el patrón encontrado.
- d) Es de suma importancia tener en cuenta que los resultados estadísticos que se obtengan de cualquier análisis de concentraciones o patrones dependerán inevitablemente de la configuración de las áreas particulares que se estén utilizando, como se mencionó en el punto anterior. En general, para llegar a tener un buen análisis y resultados bastantes fiables de lo que estemos estudiando, los datos se deben analizar sobre las bases de unidades de áreas más pequeñas para las cuales esté disponibles, y de esta manera se debe evitar y descartar la agregación de áreas más grandes arbitrarias, a menos que exista una buena justificación para utilizarlas, y por último es muy importante verificar cualquier inferencia extraída de los datos mediante el uso de diferentes configuraciones de área de los mismos datos.

2.2.1.3 Matriz de datos Espaciales

Para la realización de este trabajo, las técnicas que se aplicarán, se basan en la matriz que recopila datos espaciales necesarios para realizar el análisis.

Definición: Sean $Z_1, Z_2, Z_3, \dots, Z_K$ variables aleatorias y sea S la ubicación del punto o área. Entonces la matriz de datos espaciales esta generada por:

$$\left(\begin{array}{cccc|c} z_1(1) & z_2(1) & \dots & z_K(1) & s(1) \\ z_1(2) & z_2(2) & \dots & z_K(2) & s(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1(n) & z_2(n) & \dots & z_K(n) & s(n) \end{array} \right)$$

$$\{z_1(i), z_2(i), z_3(i), \dots, z_K(i) | s(i)\}_{i=1 \dots n} \quad (2)$$

donde, z_K denota una realización muestral (valores de los datos real) de una variable Z_K ($K = 1, \dots, n$). El símbolo i dentro de los paréntesis hace referencia al caso u observación particular (punto, área, polígono o zona espacial). Para cada caso, $i = 1, \dots, n$; $s(i)$ representa la ubicación del objeto espacial. La referencia implicará dos coordenadas geográficas. Generalmente se referirá a un espacio bidimensional, por lo tanto, $s(i) = (s_1(i), s_2(i))^t$. En el caso de datos referentes a objetos de puntos en un espacio bidimensional, la ubicación del punto i -ésimo puede estar dado por un par de coordenadas cartesianas (ortogonales). Los ejes del sistema de coordenadas; generalmente se construye para el conjunto de datos particular, pero se puede usar cualquier sistema de referencia. En el caso de datos referentes a objetos de área de forma irregular, una opción para presentar la ubicación es seleccionar un punto representativo de cada área denominada centroide y luego usar el mismo procedimiento para que un objeto de puntos sea identificado mediante $s(i)$. Existen situaciones en las que la información de georreferenciación proporcionada por $\{s(i)\}$ en la ecuación (2) tiene que completarse con información de vecindad, la misma que define no solo qué pares de áreas son adyacentes entre sí, sino que también puede cuantificar la cercanía de esa adyacencia. Esta información es necesaria y muy importante para la especificación de muchos modelos estadísticos espaciales, como los modelos de regresión espacial. (Borrego Sánchez, 2018).

2.2.1.4 Estadístico general de Producto Cruzado

Borrego Sánchez (2018, p. 23) afirma que, en la mayoría de medidas de autocorrelación espacial siempre utiliza el producto cruzado de matrices que se define de la siguiente manera:

$$\sum_{i=1}^n \sum_{j=1}^n M_{ij} W_{ij}$$

donde, W_{ij} son las entradas de la matriz de pesos espaciales definidas en la sección anterior y M_{ij} es una medida de proximidad entre las áreas i y j en alguna otra dimensión, la distancia euclídea, distancia esférica o la distancia Manhattan.

2.2.1.5 Autocorrelación Espacial

Es muy importante definir lo que significa la autocorrelación espacial (AE), Vilalta y Perdamo (2005) menciona que la AE, es la concentración o dispersión de los valores de una variable en una determinada área geográfica. Además, menciona que otros autores la definen a la AE como la medida que muestra el grado en que objetos o actividades en una unidad geográfica son

iguales a otros objetos o actividades en unidades geográficas próximas. Lo que valida que este tipo de autocorrelación espacial prueba la primera ley de Tobler (1970), el mismo que afirma que, todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las distantes. Es importante mencionar que (Bravo López, 2021), dice que la autocorrelación espacial existe cuando una variable presenta un patrón en el espacio en el que sus valores en un conjunto de ubicaciones dependen de los valores de la misma variable en otras ubicaciones.

Es importante mencionar que ante las diferentes definiciones donde proporcionan un enfoque general y validan la ley de Tobler, como se mencionó anteriormente; Borrego Sánchez (2018) concluye que, si bien es cierto que una o más variables en una ubicación geográfica son similares a los valores de estas variables en ubicaciones geográficas cercanas, a esta relación se la denomina autocorrelación espacial positiva; ahora también se dice que existe una autocorrelación negativa, cuando las observaciones que están cerca en el espacio tienden a ser más diferentes en los valores de las variables que las observaciones que están más separadas, contradiciendo a la ley de Tobler. También es importante mencionar que cuando se presenta una autocorrelación cero, los datos se distribuyen aleatoriamente en el espacio y por lo tanto no se puede aplicar las técnicas de estadística espacial mencionadas en este trabajo.

Para realizar el análisis de la autocorrelación espacial es necesario apoyarse en las medidas de contigüidad, que se define de manera general como una relación de vecindad. Esta relación se puede presentar de tres maneras diferentes, el primero que puede presentarse, es el caso de torre (Rook's case), el segundo es el caso de alfil (Bishop's case) y el último caso de la reina (Queen's case), estos nombres se deben a los movimientos del ajedrez. En el caso de tipo torre está rodeado de un vecindario de cuatro ubicaciones adyacentes a cada celda (posiciones N, S, E y O), en el caso tipo alfil únicamente considera las diagonales de cada celda (posiciones NE, SE, NO y SO), mientras que en el tercer caso considera un vecindario de ocho celdas (posiciones N, S, E, O, NE, NO, SO y SE). En la Figura 4-2, se muestra de manera visual para entender de mejor manera. (Sawada, 2009).

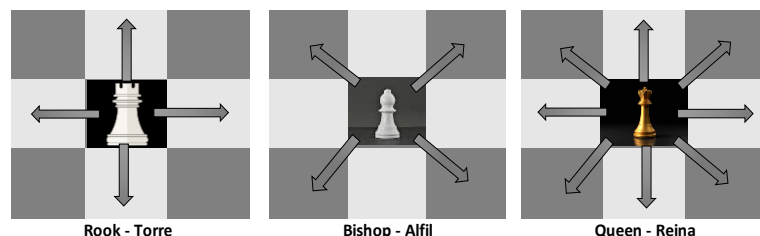


Figura 4-2: Diferentes definiciones de casos de relación de vecindad.

Elaborado por: Velasteguí Casco, Edy, 2021.

Entonces de manera formal, se puede decir que; la autocorrelación espacial es la medición de las distancias de las zonas más cercanas, es decir, aquellas zonas que rodean un punto de datos

dado que podría considerarse que influyen en la observación en ese punto de datos. (Borrego Sánchez, 2018). Pero hay que aclarar que la determinación de este vecindario tiene un cierto grado de arbitrariedad. Entonces el número de observaciones en el vecindario determinado para cada ubicación, se puede expresar mediante una matriz de ponderaciones W :

$$W = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{pmatrix}$$

donde, n representa el número de ubicaciones (observaciones), $i (i = 1, \dots, n)$ es la entrada en la fila y $j (j = 1, \dots, n)$ es la entrada en la columna, a estas dos se las denota como W_{ij} que corresponde al par (i, j) de ubicaciones.

También es importante conocer que los valores de la diagonal de la matriz W_{ij} son iguales a cero, por convenio; mientras que los demás valores que no pertenecen a la diagonal $W_{ij} (i \neq j)$ toman valores diferentes de cero (uno, para una matriz binaria) cuando las ubicaciones i y j se consideran vecinas.

Entonces podemos decir que en los datos de área, n representa el número de áreas o zonas geográficas donde cada una de estas se identifica con un punto (centroide) del que se conoce sus coordenadas cartesianas, y cada elemento de la matriz W_{ij} corresponde a la relación de dos áreas o zonas. Es importante saber que la matriz W_{ij} por lo general está estandarizada por filas, esto quiere decir que la suma de cada fila de la matriz es igual a uno, lo que permite visualizar que los valores individuales W_{ij} están proporcionalmente representados. Esto permite que cada vecino de un área o zona tenga el mismo peso y la suma de todos los W_{ij} (sobre j) sea igual a uno ($\sum_{j=1}^n W_{ij} = 1$).

Podemos decir que una de las formas de representar las relaciones espaciales con datos de áreas es aplicando el concepto de contigüidad. A los vecinos contiguos de primer orden se define con áreas que tienen un límite común, de manera formal se dice que:

$$W_{ij} = \begin{cases} 1, & \text{si el área } j \text{ comparte un límite común con el área } i \\ 0, & \text{caso contrario} \end{cases}$$

De manera alternativa, dos áreas i y j pueden definirse como vecinas cuando la distancia d_{ij} entre sus centroides es menor que un valor crítico dado, consideremos que éste es d , es así que las distancias se calculan a partir de la información sobre latitud y longitud $s(i)$ de las ubicaciones del centroide como:

$$W_{ij} = \begin{cases} 1, & \text{si } d_{ij} < d, \quad (d < 0) \\ 0, & \text{caso contrario} \end{cases} \quad (3)$$

La definición basada en la distancia de la ecuación (1) de la matriz de ponderaciones W_{ij} , depende de un valor de distancia crítica dada d . Sin embargo, cuando se presenta un alto grado de heterogeneidad en el tamaño de las unidades de área, puede presentar dificultad para encontrar una distancia crítica satisfactoria. En estas circunstancias, una distancia pequeña podrá conducir a generar muchas “islas”, mientras que, si se elige una distancia para garantizar que cada unidad de área tenga al menos un vecino, puede producir un tamaño inaceptablemente grande de número de vecinos para las unidades de áreas que son más pequeñas. Una buena solución a este problema es restringir la estructura contigua a los k -vecinos más cercanos, y por tanto excluir las “islas” (las áreas que por no estar a una distancia d de otras áreas se podría decir que no tiene vecinos) y forzar a cada unidad de área a tener el mismo número k de vecinos. De manera formal la definimos:

$$W_{ij} = \begin{cases} 1, & \text{si el centroide de } j \text{ es uno de los } k \text{ centroides más cercanos al de } i \\ 0, & \text{caso contrario} \end{cases}$$

En este caso k , es el parámetro de este esquema de ponderaciones y es el número de vecinos. También se puede cambiar la ponderación para que los vecinos más distantes obtengan menos peso introduciendo un parámetro θ que permita indicar la tasa de disminución de los pesos.

Entonces un esquema de ponderación continuo muy conocido y utilizado está fundamentado en la función de distancia inversa, lo que quiere decir que los pesos están inversamente relacionados con el área de separación de distancia i y distancia j donde el parámetro θ se estima o se establece a priori:

$$W_{ij} = \begin{cases} d_{ij}^{-\theta}, & \text{si la distancia entre centroides } d_{ij} < d \text{ (} d > 0, \theta > 0 \text{)} \\ 0, & \text{caso contrario} \end{cases}$$

Otro de los esquemas de ponderación continua se deriva de la función exponencial negativa que está dada por:

$$W_{ij} = \begin{cases} \exp(-\theta d_{ij}), & \text{si la distancia entre centroides } d_{ij} < d \text{ (} d > 0, \theta > 0 \text{)} \\ 0, & \text{caso contrario} \end{cases}$$

donde, θ es un parámetro que puede ser estimado, pero generalmente es elegido a priori por el investigador. Un parámetro elegido comúnmente por los investigadores es $\theta = 2$.

Es importante destacar que existe evidencia de una gran cantidad de métodos de matrices de ponderaciones espaciales para el mismo modelo espacial; por lo que es muy importante considerar que los resultados que arroje cualquier análisis estadístico espacial, siempre están condicionados a la matriz de pesos espaciales que elijamos. Lo que siempre resaltan varios investigadores es que se debe hacer una buena práctica aplicando varios métodos de matrices de ponderaciones espaciales, y ver si los resultados permiten concluir si se sigue cumpliendo para la mayoría de estos métodos las matrices de ponderaciones seleccionadas. Es importante

considerar que solo en el caso en el que exista una razón de conveniencia sobre bases teóricas se debe considere solo un método de matriz de ponderaciones. (Borrego Sánchez, 2018).

Para medir la autocorrelación espacial existen medidas globales y medidas locales. Las medidas globales significan, que todos los elementos de la matriz W se aplica a una evaluación de la autocorrelación; es decir, todas las asociaciones espaciales de áreas se incluyen en el cálculo de la autocorrelación espacial. De manera más sencilla y entendible lo que quiere decir es que, el análisis espacial global es un estadístico que ayuda a resumir toda la zona de estudio, por lo tanto, asume homogeneidad; lo que permite que se produzca un valor para cualquier matriz de pesos espaciales. Por otro lado, las medidas locales permiten evaluar la autocorrelación espacial asociada con una o algunas unidades de área particulares. (Borrego Sánchez, 2018).

Medidas globales

Las dos medidas que más se utilizan para el caso de las unidades de área, son los estadísticos I de Moran y c de Geary. Morán (1950) creó una de las primeras medidas para analizar la autocorrelación espacial, con el fin de estudiar fenómenos estocásticos que se distribuyen en el espacio en dos o más dimensiones, esta es la medida I de Moran que está dada por la ecuación (4):

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n (z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z})^2} \quad (4)$$

donde, n es el número de áreas que se considerará, W_{ij} es una matriz de pesos espaciales definida anteriormente. En esta medida concreta, la distancia que utiliza para crear la matriz M , que viene dada por $(z_i - \bar{z})(z_j - \bar{z})$, con z_i y z_j el valor de la variable Z del área i y j respectivamente, y siendo \bar{z} la media aritmética de todos los valores de las áreas por estudiar.

Es una de las medidas más utilizadas para medir la autocorrelación espacial comparando el valor de Z con el área i con el valor de Z en todas las demás áreas $j (j \neq i)$. Al observar la ecuación (4), se puede ver que la sumatoria $\sum_{i=1}^n (z_i - \bar{z})^2$ tiene que ser diferente de cero (0), caso contrario la ecuación no estaría bien definida. Mientras tanto, en el caso de un campo “*casi sin varianza*”, donde los valores serían iguales; el I de Moran no serviría para medir la autocorrelación espacial. La otra sumatoria que en un principio podría dar problemas de definición es $\sum_{i=1}^n \sum_{j=1}^n W_{ij}$ pero este nos muestra la suma de las entradas de la matriz de adyacencia; lo que, para aplicar las técnicas descritas se asume que existe relación de adyacencias entre las distintas áreas y, por tanto, el término no podrá anularse.

Entonces, por lo tanto, la única forma para que el estadístico de I de Moran sea cero (0) es que $\sum_{i=1}^n \sum_{j=1}^n (z_i - \bar{z})(z_j - \bar{z}) = 0$. Este estadístico varía entre -1 y 1. Un valor de I cercano a 1,

quiere decir que existe una autocorrelación espacial positiva, mientras que un valor de I que se aproxime a menos -1 nos indica una autocorrelación espacial negativa. En el caso de que el valor de I se aproxime a cero (0), se rechazaría la hipótesis de que exista autocorrelación espacial y, por tanto, los modelos y técnicas mencionadas en este trabajo no serían posible aplicarlas, dado que los datos observados serían poco fiables sobre la información que proporcionan.

Los estadísticos antes mencionados son los que se utilizan para medir la autocorrelación espacial, son reglas de decisión que permiten evaluar en qué medida la distribución espacial observada de los valores de la información; misma que es parte de la hipótesis nula, indica que el espacio no importa, es decir, las áreas cercanas no se afectan entre sí de forma tal que haya independencia y aleatoriedad espacial; mientras que la hipótesis alternativa implica que las áreas se agrupan mediante algún tipo de autocorrelación espacial positiva o negativa.

Se puede considerar que la autocorrelación espacial está presente cuando el estadístico de autocorrelación espacial calculado para un patrón particular adquiere un valor distinto, en comparación con el que se esperaría bajo la hipótesis nula de no existencias de asociación espacial. Esta variación depende de la distribución del estadístico que se utilice, por lo general sigue una distribución normal o una aproximación por permutación aleatoria. Como se explicó en el numeral de productos cruzados, si se tiene n observaciones z_i , para n unidades de áreas dadas, entonces $n!$ permutaciones de esos valores son posibles y para cada una de esas permutaciones se puede calcular el valor del estadístico. Pero hay que considerar que computacionalmente esto no sería apropiado ya que a medida que aumente las áreas, las permutaciones crecerán de manera exponencial. Por lo que es necesario utilizar un método de muestreo, por ejemplo, Monte Carlo, con el fin de escoger al azar un número representativo de permutaciones para calcular el estadístico. Entonces la varianza del estadístico I permutación aleatoria está dada como se muestra a continuación:

$$Var p(I) = \frac{(n[(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2])}{(n-1)(n-2)(n-3)S_0^2} - \frac{k[(n^2 - n + 3)S_1 - 2nS_2 + 6S_0^2]}{(n-1)(n-2)(n-3)S_0^2} - 2E(I)$$

donde, los parámetros desconocidos se definen como:

$n = \text{número de áreas}$

$$E(I) = \frac{-1}{(n-1)}, \text{ la esperanza de } I.$$

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n W_{ij}, \text{ la suma de las entradas de la matriz de pesos.}$$

$$S_1 = \sum_{i=1}^n \sum_{j=1}^n (W_{ij} + W_{ji})^2, \text{ si la matriz de pesos es simétrica entonces } S_1 = 2S_2.$$

$$S_1 = \sum_{i=1}^n \sum_{j=1}^n (W_i + W_j)^2, \text{ 2 veces la suma de la } i - \text{ésima columna y } i - \text{ésima fila.}$$

$$k = \frac{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^4}{\left[\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2 \right]^2}$$

Bajo ciertas condiciones de normalidad y el número apropiado de áreas, podemos utilizar una aproximación normal de los valores, es decir, cada valor z_i es el que se selecciona de una variable Z_i que sigue una distribución normal, entonces la varianza del estadístico I bajo aproximación normal es como se indica a continuación:

$$\text{Var } N(I) = \frac{1}{S_0(2n-1)} (n^2 S_1 - n S_2 + 3 S_0^2) - E(I)^2$$

Entonces se puede concluir que, el estadístico I de Moran como el estadístico local, que se describirá en la siguiente sección, sirve para realizar la autocorrelación espacial de los datos a partir del siguiente contraste de hipótesis dada en la ecuación (5):

$$\begin{cases} H_0: \text{"Autocorrelación espacial nula"} \\ H_1: \text{"Autocorrelación espacial no nula"} \end{cases} \quad (5)$$

Anteriormente se indicó que los estadísticos se aproximan de manera general, a una distribución normal o una aproximación por permutaciones aleatoria. Esta última también se la llama prueba de asignación al azar, reconoce que los valores observados se tratan como una población y por este motivo no tomaremos valores en las distintas áreas. Sin embargo, la prueba de distribución de muestreo aproximada hace la suposición de que las observaciones z_i son sobre observaciones de variables normales Z_i , es decir, son una realización de un proceso aleatorio y pueden darse otras relaciones posibles. La prueba de distribución de muestreo aproximado es por tanto, una autocorrelación espacial, que permite suponer que la distribución de las variables aleatorias Z_i puede ser normal.

En conclusión, en las medidas de estadísticos globales hay que destacar que las fórmulas de las varianzas de los estadísticos son más sencillas en el caso de aproximar los valores a una variable normal, por lo tanto, este método de aproximación es el que se intenta utilizar siempre que se den las condiciones adecuadas.

Medidas locales

Con la existencia de grandes cantidades de información o datos con características ideales para trabajar con SIG, se han implementado varios estadísticos, denominados estadísticos locales que son adecuados para identificar y explicar la diferencia de puntos “calientes” (grupos locales de valores altos) o puntos “fríos” (grupos locales de valores bajos) y para identificar distancias más grandes sobre la información que se tiene sobre asociación de áreas, esta sería más confusa e irrelevante.

Si se considera que de cada área $i = (i, \dots, n)$ se asocia el valor z_i mismo que representa una observación sobre la variable aleatoria Z_i , se supone que estas tienen distribuciones marginales idénticas, pero no son independientes ya que, como se ha visto, se dará autocorrelación espacial cero y esto implicaría ausencia de autocorrelación espacial. La base para las pruebas locales y medidas de autocorrelación espacial proviene de una variación del estadístico de productos cruzados:

$$\sum_{j=1}^n W_{ij} M_{ij}$$

A continuación, se menciona de manera breve cuatro estadísticos locales para medir la autocorrelación espacial: Los estadísticos locales de Getis y Ord G_i y G_i^* y las versiones locales de I de Moran y c de Geary. Para este trabajo se utilizará el estadístico local I de Moran, el mismo que se explica a continuación.

El estadístico local I_i de Moran para cada área i se define mediante la ecuación (6):

$$I_i = (z_i - \bar{z}) \sum_{j \in J_i}^n W_{ij} (z_j - \bar{z})^2 \quad (6)$$

donde, J_i es el conjunto de vecinos del área i y \bar{z} es la media de las observaciones que son vecinas al área i .

Se puede comprobar que la suma de los I_i para todas las observaciones i es proporcional al estadístico I de Moran dado por la ecuación (6).

$$\sum_{i=1}^n I_i = \sum_{i=1}^n (z_i - \bar{z}) \sum_{j \in J_i}^n W_{ij} (z_j - \bar{z})$$

El estadístico de *LISA* I_i , tiene dos propósitos. Por una parte, puede ser visto como indicador de puntos calientes y, por otro lado, se puede usar para evaluar la influencia de ubicaciones individuales (observaciones) en la magnitud del estadístico de autocorrelación espacial global correspondiente *I de Moran*. (Borrego Sánchez, 2018).

2.2.2 Modelado Espacial

De acuerdo a lo descrito anteriormente, el presente trabajo se basó en el modelado de los datos de área. Para lo cual se consideró el análisis exploratorio; enunciado en el punto 2.2.1, como base fundamental e inicial para plantear el modelo; mismo que busca establecer relaciones entre las observaciones de una variable y las observaciones de otras variables, insertadas en cada unidad de área. Para la aplicación de los modelos de regresión espacial se utilizaron datos almacenados a partir de un diseño transversal simple, aclarando que no se consideraron datos de panel.

De acuerdo con Borrego Sánchez (2018) en su trabajo de modelos de regresión para datos espaciales; establece que, para aplicar este tipo de modelos espaciales; es muy importante realizar una muestra de datos de la población objetivo, con el fin de obtener información solo por única vez, además se debe considerar que los datos obtenidos de la muestra se distribuyan aproximadamente como una normal; entonces se espera que para los datos en los que la variable de interés sea un recuento o una proporción, estos modelos sigan una distribución binomial o de Poisson, pues a medida que el número de áreas crezca, estas distribuciones tienden asintóticamente a una distribución normal. Para este trabajo se partió del modelo de regresión lineal múltiple, y se consideraron diferentes modelos de regresión espacial como: de retardo espacial, error espacial, de orden superior y el modelo espacial de Durbin, además se realizó el estudio de la estimación de parámetros de algunos modelos enunciados.

2.2.2.1 Tipos de Modelos de Regresión Espacial

Para hablar y definir los tipos de modelos espaciales, es importante recordar el modelo de regresión lineal, el cual será el punto de partida para estudiar los principales modelos de regresión espacial. Por lo que, el modelo de regresión lineal múltiple está dado por la ecuación (7):

$$Y = \sum_{q=1}^Q X_q \beta_q + \varepsilon \quad (7)$$

y su versión muestral está dada por la ecuación (8):

$$y_i = \sum_{q=1}^Q x_{iq} \beta_q + \varepsilon_i \quad i = 1, \dots, n \quad (8)$$

donde, y_i es una observación de la variable dependiente o de interés; x_{iq} es una observación en una variable explicativa con $q = 1, \dots, Q$; β_q es el coeficiente de regresión que mide la influencia por sí sola de la q -variable explicativa en la variable dependiente, es decir, que mide el cambio en Y por cada cambio unitario en X_q manteniendo las demás variables explicativas constantes; ε_i es el error aleatorio, que puede presentar debido a variables no controladas y a la variabilidad muestral. Para estos términos de error se debe asumir que ε_i son variables independientes e idénticamente distribuidas a una variable normal con media cero y varianza constante σ^2 , esto es, $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ donde $E(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i) \cdot E(\varepsilon_j) = 0$. En notación matricial este modelo puede ser enunciado mediante la ecuación (9):

$$Y = X\beta + \varepsilon \quad (9)$$

donde, Y es un vector $n \times 1$ en el que se encuentran las n observaciones de la variable dependiente; X es una matriz $n \times Q$ que muestra las observaciones de las variables explicativas; β es el vector $Q \times 1$ de parámetros de regresión asociados a dichas variables explicativas y, ε es el vector de dimensiones $n \times 1$ de términos de error.

Este modelo es muy importante y válido para comprender la relación funcional entre la variable dependiente y las variables explicativa, analiza cuáles pueden ser las causas de la variación de Y . Entonces al realizar este análisis; a partir de la información muestral, se obtiene un estimador de β ($\hat{\beta}$) para así obtener una predicción de Y a partir de las variables explicativas X_1, \dots, X_Q .

Es importante mencionar que al basarse en la suposición de que las observaciones son independientes, el modelo se simplifica notablemente; mientras tanto si se estudia los datos espaciales, esta simplificación puede arrojar resultados parciales inconsistentes debido a la dependencia espacial; misma que puede presentarse en las variables explicativas, en la dependiente o en los residuos (términos de error), entonces bajo estos criterios se puede decir que, cuando la dependencia espacial se encuentra en la variable dependiente los modelos se denominan “modelos de retardo espacial” mientras que si está en los residuos se denominan “modelos de error espacial”; en tanto que, si está presente en las variables explicativas se denominan “modelos de regresión cruzada o modelos X espacialmente retardados”, pero en comparación con los otros dos modelos, no tienen procedimientos espaciales precisos para su estimación. (Borrego Sánchez, 2018).

2.2.2.2 Modelo de Retardo Espacial

Los modelos de retardo espacial, presentan la correlación espacial (dependencia) en la variable dependiente (Y). Este tipo de modelos son extensiones de modelos de regresión dado por la ecuación (8), estos permiten a las observaciones de la variable dependiente Y en el área

$i(i = 1, \dots, n)$ depender de observaciones en áreas vecinas. El modelo de retardo espacial básico, al cual también se lo conoce como modelo autorregresivo espacial de primer orden (SAR), está definido de forma muestral, como se muestra en la ecuación (10):

$$y_i = \rho \sum_{j=1}^n W_{ij} y_j + \sum_{q=1}^Q x_{iq} \beta_q + \varepsilon_i \quad i = 1, \dots, n \quad (10)$$

donde, W_{ij} es el elemento (i, j) th de la matriz de pesos $n \times n$; W ver en el ítem (2.2.1.4); ρ es un escalar, dicho parámetro a estimar que determinará el nivel de relación autoregresiva espacial entre y_i y $\sum_{j=1}^n W_{ij} y_j$ (combinación lineal de observaciones espacialmente relacionadas en elementos distintos de cero en la i -ésima fila de W).

En notación matricial se escribe mediante la ecuación (11):

$$Y = \rho W Y + X \beta + \varepsilon \quad (11)$$

donde, $W Y$ se conoce como una variable dependiente espacialmente retardada; El dominio de ρ , está definido por $(w_{min}^{-1}, w_{max}^{-1})$, donde w_{min}^{-1} y w_{max}^{-1} representan el máximo y mínimo autovalor de la matriz W . Si la matriz de pesos W es estocástica, se puede probar que los autovalores de dicha matriz están en el intervalo $[-1, 1]$. La demostración de esta prueba se encuentra en el capítulo 2 del libro de Baris y Mate (2012). Cuando W es estocástica por filas, una entrada (i, j) distinta de cero, nos indica que la j -ésima observación se usará para ajustar la predicción de la fila i .

Este es un modelo estructural, que se lo enuncia en forma reducida como se muestra en la ecuación (12):

$$Y = (I - \rho W)^{-1} (X \beta + \varepsilon) \quad (12)$$

Su valor esperado está dado por la siguiente expresión:

$$E[Y] = (I - \rho W)^{-1} (X \beta)$$

donde, el término $(I - \rho W)^{-1}$ se denomina multiplicador especial e indica que el valor esperado de cada observación y_i dependerá de una combinación lineal de valores X tomados por observaciones vecinas, escalado por el parámetro de dependencia ρ .

2.2.2.3 Modelo de Error Espacial

Los modelos de error espacial, explican la dependencia espacial en el término del error, esta dependencia del error espacial puede suceder o aparecer de variables latentes no observables

que están correlacionadas espacialmente; pero también pueden surgir de los límites del área que no reflejan con precisión la vecindad que dan lugar a las variables recopiladas para el análisis; entonces un proceso autorregresivo espacial de primer orden esta dado mediante la ecuación (13):

$$\varepsilon_i = \lambda \sum_{j=1}^n W_{ij} \varepsilon_j + u_i \quad i = 1, \dots, n \quad (13)$$

donde, λ es el parámetro autorregresivo; u_i es un termino aleatorio de error, que se asume es independiente e idénticamente distribuido.

En el caso: si $|\lambda| < 1$, entonces se tendría:

$$\varepsilon = (I - \lambda \mathbf{W})^{-1} u$$

reemplazando esta ecuación en el modelo de regresión estándar se obtiene la ecuación (14):

$$Y = X\beta + (I - \lambda \mathbf{W})^{-1} u \quad (14)$$

Con $E[uu'] = \sigma^2 I$, por lo tanto la matriz de varianza-covarianza del vector de errores aleatorio está dado por:

$$E[\varepsilon \varepsilon^t] = \sigma^2 (I - \lambda \mathbf{W})^{-1} (I - \lambda \mathbf{W})^{-1}$$

El modelo definido mediante la ecuación (13), se denomina modelo de error espacial (SEM) y se lo puede ver como una combinación entre el modelo de regresión estándar con un modelo autorregresivo espacial en el término del error.

2.2.2.4 Modelo espacial de Durbin (SDM)

Del modelo espacial de Durbin empleado en algunos estudios, se puede decir que la dependencia influye sobre la variable objetivo o dependiente, a través de esta variable objetivo y las variables explicativas o predictoras; se puede realizar un modelo tipo SAR, el cual fue definido mediante el modelo presentado en la ecuación (10), aumentando variables explicativas espacialmente retardadas, mediante la ecuación (15):

$$Y = \rho \mathbf{W}Y + \mathbf{X}\beta + \mathbf{W}\bar{\mathbf{X}}\gamma + \varepsilon \quad (15)$$

donde, $\bar{\mathbf{X}}$, es la matriz $n \times (Q - 1)$ no constantes de variables explicativas, este modelo puede reducirse como se muestra a continuación mediante la ecuación (16):

$$Y = (I - \rho \mathbf{W})^{-1} (\mathbf{X}\beta + \mathbf{W}\bar{\mathbf{X}}\gamma + \varepsilon) \quad (16)$$

con

$$\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$$

donde, γ es un vector $(Q - 1) \times 1$ de parámetros que miden el impacto marginal de las variables explicativas de observaciones vecinas (áreas) sobre la variable dependiente Y ; y $\mathbf{W}\bar{\mathbf{X}}$ genera variables explicativas retardadas espacialmente que arrojan un promedio de observaciones vecinas.

Ahora para el uso de un modelo SDM, este se basa en dos circunstancias que suelen surgir cuando se aplica un modelo de regresión espacial a datos de área. Una de estas dos circunstancias es la dependencia espacial en las perturbaciones de un modelo de MCO (Mínimos Cuadrados Ordinarios); la segunda circunstancia es cuando existe una variable explicativa omitida que muestra una covarianza no nula con una variable incluida en el modelo, y las variables omitidas son probables cuando se trata muestras de datos de área. Por lo que también es muy importante el modelo espacial de Durbin ya que, generaliza la mayoría de los modelos que se utilizan pueden presentarse cuando ocurren los siguientes casos:

- a) Si se impone que $\gamma = 0$ esto lleva a un modelo de tipo SAR como se definió mediante la ecuación (11).
- b) Si $\gamma = -\rho\beta$ esto lleva a un modelo de tipo SEM definido por la ecuación (13).
- c) Ahora, si se tiene la restricción $\rho = 0$ entonces el resultado es un modelo de regresión \mathbf{X} espacialmente retardado; que asume la independencia entre las observaciones de la variable dependiente, pero este incluye características de las áreas vecinas, en la forma de variables explicativas espacialmente retardadas.
- d) Y cuando se impone las restricciones $\rho = 0$ y $\gamma = 0$, se obtiene como resultado el modelo de regresión estándar definido mediante la ecuación (9).

Ahora, si se define $\mathbf{Z} = [\mathbf{X} \quad \mathbf{W}\bar{\mathbf{X}}]$ y $\delta = \begin{bmatrix} \beta \\ \gamma \end{bmatrix}$ se puede observar de manera más evidente que es un modelo tipo SAR (11), quedando como resultado el modelo definido por la ecuación (17):

$$Y = \rho\mathbf{W}Y + \mathbf{Z}\delta + \varepsilon \quad (17)$$

Que si se agrupa se puede expresar mediante la ecuación (18)

$$Y = \rho\mathbf{W}Y + \mathbf{X}\beta + \mathbf{W}\bar{\mathbf{X}}\gamma + \varepsilon \quad (18)$$

Entonces finalmente, se puede generalizar el modelo espacial de Durbin mediante la ecuación (18):

$$Y = \rho\mathbf{W}_1Y + \mathbf{X}\beta + \mathbf{W}_1\bar{\mathbf{X}}\gamma + \varepsilon \quad (19)$$

$$\varepsilon = \lambda\mathbf{W}_2\varepsilon + u$$

$$u \sim \mathcal{N}(0, \sigma_u^2 I)$$

donde, W_1 y W_2 son matrices de $n \times n$ de pesos espaciales que pueden ser iguales o distintos.

2.2.3 Prueba para Dependencia Espacial

En los modelos de regresión, cuando existe la presencia de dependencia espacial, esta se la puede detectar mediante pruebas de diagnóstico. A continuación, se enuncian tres pruebas diferentes que sirven para detectar la presencia de dependencia espacial en términos de error:

- a) El test de Moran y,
- b) Los test de los multiplicadores de LaGrange para los modelos espaciales de retardo definido en la ecuación (10) y de error definido mediante la ecuación (13).

2.2.3.1 Test de Moran

El estadístico para realizar este contraste de hipótesis para la existencia de autocorrelación espacial, es el estadístico I de Moran definido anteriormente mediante la ecuación (4), aplicada a los residuos de regresión. Ahora si denotamos a e como el vector $n \times 1$ de mínimos cuadrados ordinarios (MCO) definido como: $e = Y - X\hat{\beta}$, el estadístico I de Moran queda redefinido de la siguiente forma como se muestra en la ecuación (20):

$$I = \frac{n}{W_0} \frac{e'W_e}{e'e} \quad (20)$$

donde, W_0 es el vector normalizado que aparece en el estadístico I de Moran; $W_0 = \sum_{i=1}^n \sum_{j=1}^n W_{ij}$ y $e'e$ es la suma de los cuadrados de los residuos. El estadístico I se interpreta como el coeficiente de una regresión de MCO de W_e sobre e .

Aplicado a la práctica, la inferencia mediante la prueba I de Moran, se basa en una aproximación normal utilizando un valor estandarizado obtenido restando la media, bajo la hipótesis nula de no dependencia espacial, y dividido por la raíz cuadrada de la varianza. El test estadístico está dado por la siguiente ecuación (21):

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \sim \mathcal{N}(0,1) \quad (21)$$

donde, el valor esperado y la varianza del estadístico I de Moran están expresados con detalle en la sección donde se definió mediante la ecuación (4).

Es importante tener en cuenta que la causa de dependencia espacial bajo hipótesis alternativa no está especificada, por lo tanto, la prueba de Moran es una prueba general que permite detectar autocorrelación espacial. Cuando se aplica esta prueba a los residuos se debe tener mucho cuidado, ya que si se estiman los Q coeficientes de regresión, entonces los residuos observados están sujetos a Q restricciones lineales, lo cual quiere decir que estos residuos estarán correlacionados y el procedimiento de prueba mediante I de Moran no servirá en este caso; mientras que en el caso de que el número de coeficientes Q fuese pequeño en comparación con n (número de áreas) se podría ignorar esta correlación y de este modo validar la prueba.

2.2.3.2 Test de multiplicador de LaGrange (LM)

Una de las pruebas alternativas es justamente el test de multiplicador de LaGrange. Burrige P. (1980), recomienda que se debe calcular mediante los residuos de MCO y que es diferente según qué tipo de dependencia espacial se quiera probar. Para una dependencia de tipo error dado por la ecuación (13), el estadístico LM (error) viene dado como se muestra a continuación, mediante la ecuación (22):

$$LM(error) = \left(\frac{e^t W_e}{\frac{1}{n} e^t e} \right)^2 \frac{1}{tr[W^t W + W^2]} \quad (22)$$

donde, tr es el operador traza y $\frac{1}{n} e^t e$ es la varianza muestral del error.

Con el factor de normalización incluido se consigue una distribución asintótica chi-cuadrado con un grado de libertad (χ_1^2) bajo la hipótesis nula y de no dependencia espacial. Esto, es a partir del siguiente contraste de hipótesis:

$$\begin{cases} H_0: \lambda = 0 \\ H_1: \lambda \neq 0 \end{cases}$$

Si se toma la decisión de rechazar H_0 , cuando:

$$LM(error) > \chi_{1,1}^2(\alpha, \text{nivel de significancia}).$$

Si la hipótesis alternativa, el logaritmo de la función de verosimilitud de los datos viene dado mediante la siguiente ecuación (32):

$$\begin{aligned} (y \setminus x, \lambda, \beta, \sigma^2) = & -\frac{n}{2} \ln(2\pi\sigma^2) + \ln|I - \lambda W| \\ & - \frac{1}{2\sigma^2} (y - x\beta)^t (I - \lambda W)^t (I - \lambda W) (y - x\beta) \end{aligned} \quad (23)$$

Cuando le imponemos un valor de cero a λ este produce la función de verosimilitud restringida como se muestra a continuación mediante la ecuación (24):

$$L^R() = L() - \alpha\lambda \quad (24)$$

donde, α es un parámetro que debe ser estimado.

Las condiciones de primer orden para obtener el máximo de la función de verosimilitud restringida están dadas por las ecuaciones (25), (26) y (27) :

$$\partial L^R / \partial \beta^t = \frac{1}{\sigma^2} x^t (I - \lambda \mathbf{W})(y - x\beta) = 0 \quad (25)$$

$$\partial L^R / \partial \sigma^2 = \frac{1}{2\sigma^2} [n + \sigma^2 (y - x\beta)^t (I - \lambda \mathbf{W})^t (I - \lambda \mathbf{W})(y - x\beta)] = 0 \quad (26)$$

$$\begin{aligned} \partial L^R / \partial \lambda &= -\alpha - \sum_{j=1}^n \delta_j (1 - \lambda \delta_j)^{-1} \\ &+ \frac{1}{2\sigma^2} [n + \sigma^2 (y - x\beta)^t (I - \lambda \mathbf{W})^t (I - \lambda \mathbf{W})(y - x\beta)] \\ &= 0 \end{aligned} \quad (27)$$

donde, $\delta_j, j = 1, \dots, n$ son los autovectores de la matriz \mathbf{W} y de las ecuaciones anteriores sacamos los estimadores para α y σ^2 , insertando $\hat{\beta}$ el estimador de MCO queda de la siguiente forma:

$$\begin{aligned} \sigma^2 &= \frac{1}{n} (y - x\hat{\beta})^t (y - x\hat{\beta}) = \frac{e^t e}{n} \\ \hat{\alpha} &= -tr(\mathbf{W}) + \frac{1}{2\sigma^2} (y - x\hat{\beta})^t (\mathbf{W} + \mathbf{W}^t)(y - x\hat{\beta}) = \frac{ne^t (\mathbf{W} + \mathbf{W}^t)e}{2e^t e} \end{aligned}$$

Ahora se puede demostrar de manera simple que bajo la hipótesis H_0 , la matriz de información es diagonal expresada de la siguiente manera:

$$-E \left(\frac{\partial^2 L}{\partial \lambda \partial \lambda} \right) \Big|_{\lambda=0} = tr(\mathbf{W}^2 + \mathbf{W}^t \mathbf{W})$$

Por lo tanto, el test puede ser construido considerando a $\frac{\hat{\alpha}}{(tr(\mathbf{W}^2 + \mathbf{W}^t \mathbf{W}))^{-\frac{1}{2}}}$ como una desviación normal estándar.

2.2.4 Prueba de dependencia espacial sustantiva

Esta prueba es aplicada para un modelo de retardo espacial omitido, la prueba toma la siguiente forma y se la representa mediante la ecuación (28):

$$LM(Lag) = \left(\frac{e^t W e}{\frac{1}{n} e^t e} \right)^2 \frac{1}{H} \quad (28)$$

Quedando de la siguiente forma:

$$H = \left\{ (W)^t [I - X(X^t X)^{-1} X]^t (WX\hat{\beta}) \hat{\sigma}^{-2} \right\} + tr(W^t W + W^2)$$

donde, $\hat{\beta}$ y $\hat{\sigma}$ denotan los MCO estimados; $LM(Lag)$ es la prueba que sigue una distribución asintótica chi-cuadrado con un grado de libertad (χ_1^2) bajo la hipótesis nula de no dependencia espacial ($H_0: \rho = 0$).

Dadas estas condiciones la prueba se realiza mediante el siguiente contraste de hipótesis el cual se muestra mediante la siguiente expresión (29):

$$\begin{cases} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{cases} \quad (29)$$

Se toma la decisión de rechazar H_0 si $LM(Lag) > \chi_{1,1-\alpha}^2(\alpha, \text{nivel de significancia})$. (Borrego Sánchez, 2018, p. 40-44).

2.2.5 Estimación de un Modelo Espacial

El método de máxima verosimilitud (MV) es el que se utiliza para estimar los modelos de regresión espacial, este se basa en maximizar la probabilidad de distribución conjunta en relación a ciertos parámetros relevantes. Es importante mencionar que el método de MV tiene propiedades teóricas asintóticas muy interesantes tales como: consistencia, eficiencia o normalidad asintótica, y además se considera fuerte para pequeñas variaciones bajo la suposición de normalidad.

En primer lugar, vemos como se aplica el método de MV en los modelos tipo SAR, para esto se debe partir de la solución de la variable dependiente Y , que tiene como ecuación a la expresión (12).

Para resumir el procedimiento de las sucesivas ecuaciones definiremos que $A = I - \rho W$. Pero de antemano se sabe que los ε_i son independientes e idénticamente distribuidos mediante una

distribución normal, $\mathcal{N}(0, \sigma^2)$, por lo tanto, la función de densidad quedaría como se muestra a continuación:

$$\phi(s) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}s^2\right)$$

En este caso, el vector ε sigue una distribución normal multivalente y denotando la matriz de covarianza $\Sigma = \sigma^2 I$; ahora podemos decir que: $|\Sigma|^{-\frac{1}{2}} = \sigma^{-n}$; $\Sigma^{-1} = \frac{1}{\sigma^2} I$ y $|\Sigma| = \sigma^{2n}$ y por tanto la distribución de ε será:

$$\varepsilon \sim \mathcal{N}_n(0, \sigma^2 I)$$

Por lo tanto, dado que:

$$Y = \mathbf{A}^{-1} \mathbf{X} \beta + \mathbf{A}^{-1} \varepsilon$$

Entonces

$$Y \sim \mathcal{N}_n\left(\mathbf{A}^{-1} \mathbf{X} \beta, \sigma^2 \mathbf{A}^{-1} (\mathbf{A}^{-1})^t\right)$$

Por lo que la función de máxima verosimilitud quedaría de la siguiente forma:

$$\begin{aligned} L(\rho, \beta, \sigma^2) &= (2\pi)^{-\frac{n}{2}} |\sigma^2 \mathbf{A}^{-1} (\mathbf{A}^{-1})^t|^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2} ((Y - \mathbf{A}^{-1} \mathbf{X} \beta)^t \mathbf{A}^t \mathbf{A}) (Y - \mathbf{X} \beta)\right) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} |\mathbf{A}| \exp\left(\frac{1}{2} (\mathbf{A} Y - \mathbf{X} \beta) \Sigma^{-1} (\mathbf{A} Y - \mathbf{X} \beta)\right) \end{aligned}$$

Pero ahora, se la puede representar como un modelo SAR definido por la ecuación (30).

$$Y = \rho \mathbf{W} Y + \mathbf{Z} \delta + \varepsilon \quad (30)$$

Sea $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, entonces el logaritmo de la función de probabilidad de la ecuación del modelo SDM se expresa como se muestra en la ecuación (31):

$$\ln L(\rho, \delta, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) + \ln |\mathbf{A}^*| - \frac{1}{\sigma^2} (\mathbf{A}^* Y - \mathbf{Z} \delta)^t (\mathbf{A}^* Y - \mathbf{Z} \delta) \quad (31)$$

donde, n es el número de observaciones (áreas); a la cual se asigna la expresión $\mathbf{A}^* = I - \rho \mathbf{W}$. Entonces para maximizar esta probabilidad se utilizan los parámetros ρ, δ y σ^2 . Ahora, los estimadores de los coeficientes regresivos δ y σ^2 se expresan en función de ρ como se muestra a continuación:

$$\begin{aligned} \hat{\delta} &= \delta_0 - \rho \delta_L \\ \widehat{\sigma^2} &= (e_0 - \rho e_L)^t (e_0 - \rho e_L) \frac{1}{n} \end{aligned}$$

donde, δ_0 y δ_L , son los coeficientes de regresión de MCO en una regresión de \mathbf{Z} y \mathbf{WY} , mientras que e_0 y δ_L son los vectores residuales en la regresión de e_0 y δ_L . En forma de ecuación se expresa como:

$$\delta_0 = (\mathbf{Z}^t \mathbf{Z})^{-1} \mathbf{Z}^t Y, \delta_L = (\mathbf{Z}^t \mathbf{Z})^{-1} \mathbf{Z}^t \mathbf{WY}$$

$$e_0 = Y - \mathbf{Z}\delta_0 \text{ y } e_L = \mathbf{WY} - \mathbf{Z}\delta_L.$$

Esto permite usar la función de verosimilitud restringida que proporcionará el mismo estimador para ρ que maximiza la función de verosimilitud, y por tanto también determinará el mismo estimador para β y σ^2 . La importancia de utilizar este mecanismo permite simplificar el problema de optimizar disminuyendo el problema de una multivariable a uno univariable; de la misma manera que en el modelo SAR definido mediante la ecuación (10). Ahora, sustituyendo estas ecuaciones en la (31), se determina el valor del logaritmo de la función de verosimilitud concentrada como se muestra a continuación:

$$\ln L_{res}(\rho) = k + \ln|I - \rho W| - \frac{n}{2} \ln[(e_0 - \rho e_L)^t (e_0 - \rho e_L)]$$

donde, k es una constante que no depende de ρ ; ahora entonces sea $S(\rho)$ el término de la expresión de la función $\ln L_{con}(\rho)$, $(e_0 - \rho e_L)^t (e_0 - \rho e_L) = e_0^t e_0 - 2\rho e_0^t e_L + \rho^2 e_L^t e_L$. Entonces para simplificar la optimización de $\ln L_{con}(\rho)$ con respecto al parámetro ρ , Pace y Barry (1997) recomiendan evaluar dicha función usando un vector de valores para ρ , y de la misma manera lo ratifica Borrego Sánchez (2018):

$$\begin{pmatrix} \ln L(\rho_1) \\ \ln L(\rho_2) \\ \vdots \\ \ln L(\rho_n) \end{pmatrix} = k + \begin{pmatrix} \ln|I - \rho_1 W| \\ \ln|I - \rho_2 W| \\ \vdots \\ \ln|I - \rho_q W| \end{pmatrix} - \frac{n}{2} \begin{pmatrix} \ln L(S(\rho_1)) \\ \ln L(S(\rho_2)) \\ \vdots \\ \ln L(S(\rho_q)) \end{pmatrix}$$

Se puede ver que la matriz de pesos \mathbf{W} es uno de los argumentos que en un principio mayor dificultad puede ocasionar al realizar los cálculos de la optimización; al ser una matriz $n \times n$, sin embargo, al ser una matriz de pesos construida a partir de la contigüidad de las áreas, se supone que tendrá muchas entradas nulas. Entonces esto significa que se realizan los cálculos con una matriz dispersa, a ésta también se la conoce como hueca, lo que permite saber que su tiempo de operación será lineal ($O(n)$), mientras tanto la matriz densa es ($O(n^2)$).

A continuación, se determina la estimación de los parámetros para los modelos SEM cuya ecuación viene dada por (13). Partiendo del método de máxima verosimilitud y asumiendo normalidad en los términos de error (ε) y usando el Jacobiano, entonces el logaritmo de la función de probabilidad se determina mediante la ecuación (32):

$$\begin{aligned} \ln L(\lambda, \beta, \sigma^2) = & -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) + \ln|I - \rho W| \\ & - \frac{1}{2\sigma^2} (Y - \mathbf{X}\beta)^t (I - \lambda W)^t (I - \rho W) (Y - \mathbf{X}\beta) \end{aligned} \quad (32)$$

Se puede ver que el último término de la ecuación (32), que el máximo de la función de máxima verosimilitud es equivalente al mínimo de la suma de los residuos al cuadrado de una regresión, donde la variable espacialmente dependiente es $Y^* = Y - \lambda WY$ en un conjunto de variables explicativas dada por $X^* = X - \lambda WX$.

Replicando el proceso de estimación que se utilizó para los modelos SAR, estimación β y σ^2 y, aplicando el método de MV se tiene:

$$\begin{aligned} \hat{\beta}_{MV} = & [\mathbf{X}^t (I - \lambda W)^t (I - \lambda W) \mathbf{X}]^{-1} \mathbf{X}^t (I - \lambda W)^t (I - \lambda W) Y \\ \widehat{\sigma^2}_{MV} = & \frac{1}{n} (e - \lambda W e)^t (e - \lambda W e) \end{aligned}$$

donde, $e = Y - X\hat{\beta}_{MV}$. Es muy importante mencionar que, si se quiere conseguir un estimador de mayor consistencia; se lo podría hacer mediante la aplicación de un método numérico. Existen varios autores que aplican estos métodos donde detallan sobre este tipo de estimación para los parámetros en los diferentes modelos de regresión espacial, como referencia se puede citar (LeSage & Pace, 2009).

2.2.6 Índice de Akaike (AIC)

AIC es una medida de calidad relativa de un modelo estadístico, el cual permite medir el grado de efectividad de ajuste a los datos que se quiere ajustar un determinado modelo. Su fórmula está dada por:

$$AIC = 2k - 2\ln(L) \quad (33)$$

donde,

k es el número de parámetros del modelo, y

L es el máximo de la función de verosimilitud.

La regla de decisión para seleccionar el mejor modelo estadístico se cumple para el de menor medida de AIC. En este sentido se aplicará esta medida de calidad para seleccionar el mejor modelo de regresión espacial en este trabajo.

CAPITULO III

3 METODOLOGÍA DE INVESTIGACIÓN

En este capítulo se realiza la identificación de los delitos como: robo a personas, robo a domicilios, robo a unidades económicas y robo a vehículos, mediante las bases de datos del año 2019; esta identificación permitirá determinar los indicadores sobre la tasa de robos a personas y domicilios, número de robos a unidades económicas y vehículos, definición de variables e indicadores sociodemográficos y su concentración en las diferentes zonas censales urbanas de la ciudad de Cuenca; lo cual permitirá la aplicación tanto para medir la autocorrelación espacial y de modelos espaciales como: Retardo espacial, Error espacial y Durbin en cada uno de los delitos, y finalmente se detallará la aplicación de los paquetes y librerías más importantes en el lenguaje estadístico R, mediante su entorno de desarrollo RStudio.

3.1 Selección de los modelos espaciales

Para la determinación de las zonas urbanas con mayor concentración de robos en la ciudad de Cuenca durante el año 2019, se aplicaron los modelos espaciales: *I* de Moran, Retardo espacial, Error espacial y Durbin los mismo que se estimaron mediante los datos espaciales de los delitos de robos en las cuatro diferentes modalidades, como son: robo a personas, domicilios, unidades económicas y robo a vehículos, aplicando los métodos matemáticos planteados, en las secciones 2.2.1.5, 2.2.2.2, 2.2.2.3 y 2.2.2.4.

Estos modelos, son los más aplicados para determinar este tipo de comportamiento en diferentes áreas; de acuerdo a la revisión bibliográfica realizada en el capítulo 2, ya que los resultados que arrojan conllevan a ajustes acordes a la problemática de estudio, en este sentido se aplicaron estos modelos matemáticos espaciales para medir la alta incidencia que existe en las diferentes zonas urbanas, para lo cual se aprovecharon los indicadores sociodemográficos de población y vivienda desagregados a nivel de zona, para determinar su influencia en este tipo de delito en la ciudad de Cuenca. Para obtener estos resultados se describe de manera detallada los paquetes y librerías empleados en el software estadístico R, mediante su entorno de desarrollo RStudio.

3.1.1 *Obtención de los datos espaciales para aplicación de los modelos espaciales*

Para esta investigación se obtuvieron datos de fuentes secundarias, los mismos que son recopilados mediante registros administrativos de las denuncias realizadas en la Fiscalía General del Estado en la provincia del Azuay. Estos datos, son procesados por el Observatorio de Seguridad Ciudadana del Cantón Cuenca y homologados con la Policía Nacional; están

estructurados como se muestra en la sección 2.2.1.1, además se especifica de manera espacial las herramientas en las que se puede presentar para visualizar y analizar este fenómeno delincencial como se muestra en la sección 2.2.1.2.

También es importante mencionar que se utilizaron datos que corresponden a las características de la población y hogares de la ciudad de Cuenca; los cuales se obtuvieron de una fuente secundaria del Instituto Nacional de Estadística y Censos (INEC), que corresponden al 7mo. Censo de Población y 6to. de Vivienda 2010, ya que estos son datos oficiales y se encuentran homologados a un nivel de desagregación geográfica de zona censal, lo cual se ajusta a lo que se quiere realizar en este trabajo.

3.1.2 Análisis de las variables utilizadas para aplicar los modelos espaciales

Para la aplicación de los modelos descritos en las secciones 2.2.1.5, 2.2.2.2, 2.2.2.3 y 2.2.2.4 se tomaron las variables: número de robo a personas y robos a: domicilios, establecimientos económicos y vehículos; ocurridos por zonas urbanas censales en la ciudad de Cuenca. Para medir la concentración de robos y además conocer si existe alguna relación o influencia de las características sociodemográficas con respecto a este delito en las diferentes zonas geográficas, se han seleccionado las siguientes variables:

Delito de Robos

BD_Tasa_R_: Tasa de robo a personas

BD_Tasa__1: Tasa de robo a domicilios

BD_Robo_U_: Robo a unidades económicas

BD_Robo_ve: Robo a vehículos

Demografía y distribución espacial

BD_DPA_ZON: Identificador de zona censal urbana

BD_N: Número de personas.

BD_PMI: Inmigrantes a escala internacional

BD_TIEpc: Tasa de inmigrantes extranjeros

BD_PPH: Promedio de personas por hogar

BD_Tasa_Dt: Tasa de desplazamiento temporal

BD_Inmi_Ex: Inmigrantes extranjeros

BD_NEI: Número de emigrantes internacional

Actividades Económicas

BD_PET: Población en edad de trabajar

BD_PCes: Población de 10 años y más que son cesantes

BD_POdh: Población que realizan su actividad económica dentro del hogar

BD_PEI: Población económicamente inactiva

Educación

BD_TA: Tasa de analfabetismo - población 10 años y más

BD_Tasa_Na: Tasa neta de asistencia a la educación superior - población de 24 años y más

Vivienda y servicios

BD_HTV: Total de hogares según cada forma de tenencia de la vivienda: propia y totalmente pagada

3.2 Estimación de parámetros e indicadores

La estimación de parámetros e indicadores se aplica a 81 zonas censales existentes en la ciudad de Cuenca, específicamente la zona urbana, que está dividida geográficamente mediante la metodología empleada por el INEC, donde se establece que el parámetro para conformar una zona censal urbana, es el sector censal y esta zona censal está conformada por 10 sectores censales, y cada sector censal lo integran 150 viviendas en promedio, para lo cual se realiza el cálculo de indicadores bajo estos parámetros de medida geográfica conocida como zona censal.

3.2.1 Estimación de indicadores a nivel de zona urbana censal

Para la estimación de indicadores sociodemográficos se aplicó la metodología empleada por el INEC, cuyos cálculos se detallan a continuación, y se emplean en diferentes áreas como: demografía y distribución espacial, educación, vivienda y servicio:

Demografía y distribución espacial

Nombre del Indicador: Población (habitantes)

Definición Operativa:

$$H = \sum_{1}^{N} N$$

donde, N es el número de personas

Descripción del indicador: Número total de habitantes según la División Político Administrativa.

Nombre del Indicador: Emigrantes a escala internacional

Definición Operativa:

$$PMI = \sum_{i=1}^{IPME} PME$$

donde, $\sum_{i=1}^{IPME} PME$ es la sumatoria de personas que emigraron al extranjero.

Descripción del indicador: Número de personas de todas las edades que emigraron al extranjero en el transcurso de noviembre de 2001 a noviembre de 2010. Nota: Datos derivados del censo.

Nombre del Indicador: Tasa de inmigrantes extranjeros presentes en el censo

Definición Operativa:

$$TIE_{pc} = \frac{TIE_{pc}}{TP_i} * 1000$$

donde, TIE_{pc} es el total de inmigrantes extranjeros presentes en el censo y TP_i es el total de población en el año 2010 y i es la zona geográfica determinada.

Descripción del indicador: Es el total de inmigrantes extranjeros que estuvieron presentes en el censo en una zona geográfica determinada por 1000 habitantes de dicha zona en el año 2010.

Nota: Datos derivados del censo.

Nombre del Indicador: Promedio de personas por hogar

Definición Operativa:

$$PPH = \frac{\sum_{i=1}^N personas}{TH}$$

donde, $\sum_{i=1}^N personas$ es la suma de personas y TH es el total de hogares de dicha zona en el año 2010.

Descripción del indicador: Fija el promedio de personas que habitaron en el momento censal en los hogares del Ecuador.

Actividades Económicas

Nombre del Indicador: Población en edad de trabajar (PET)

Definición Operativa:

$$PET = NP_{\geq 10 t}$$

donde, $NP_{\geq 10 t}$ es la población de 10 años y más.

Descripción del indicador: Está constituida por las personas de 10 años y más.

Nombre del Indicador: Población Cesante

Definición Operativa:

$$PCes = PCes_{\geq 10}$$

donde, $PCes_{\geq 10}$ es la población de 10 años y más que buscó trabajo habiendo trabajado antes y está disponible para trabajar.

Descripción del indicador. - Personas en edad de trabajar (10 años y más) que en la semana de referencia presentaron las siguientes características:

- No ocupado/a, trabajó anteriormente.
- Buscando trabajo.
- Disponible para trabajar.

Nombre del Indicador: Población ocupada que realizó su actividad económica dentro del hogar

Definición Operativa:

$$POdh = \frac{PO_{dh \geq 10}}{TPO_{\geq 10}}$$

donde, $PO_{dh \geq 10}$ es la población de 10 años y más, ocupada que realiza su actividad económica dentro del hogar y $TPO_{\geq 10}$ es el total de personas ocupadas de 10 años y más.

Descripción del indicador. - Mide el porcentaje de la población de 10 años y más ocupada que realiza su actividad económica dentro del hogar.

Nombre del Indicador: Población Económicamente Inactiva (PEI)

Definición Operativa:

$$PEI = P_{\geq 10 r} + P_{\geq 10 j} + P_{\geq 10 e} + P_{\geq 10 qh} + P_{\geq 10 d} + P_{\geq 10 o}$$

donde; $P_{\geq 10 r}$ es la población de 10 años y más catalogadas como rentistas, $P_{\geq 10 j}$ es la población de 10 años y más catalogadas como jubilados o pensionistas, $P_{\geq 10 e}$ es la población de 10 años y más catalogadas como estudiante, $P_{\geq 10 qh}$ es la población de 10 años y más que realizan quehaceres domésticos, $P_{\geq 10 d}$ es la población de 10 años y más a las que su discapacidad le impide realizar alguna actividad y $P_{\geq 10 o}$ son las demás personas dentro de esta categoría.

Descripción del indicador: Personas en edad de trabajar (10 años y más) que en la semana de referencia presentan las siguientes características:

- Es rentista.
- Es jubilado o pensionista.
- Es estudiante.
- Realiza quehaceres en el hogar.
- Le impide su discapacidad.
- Otro.

Educación

Nombre del Indicador: Tasa de Analfabetismo

Definición Operativa:

$$TA = \frac{TP_{lne}}{TP} * 1000$$
$$i = 1, 2$$

donde, TP_{lne} es el total de población que no sabe leer ni escribir, en el año 2010, $i = 1 =$ población de 10 años y más $i = 2 =$ población de 15 años y más y TP es el total de población en la edad de referencia, en el año 2010.

Descripción del indicador: Porcentaje de población analfabeta de una edad determinada. Este indicador se calculará para la población de 10 años y más para mantener comparabilidad con los datos del censo 2001 y de 15 años y más de edad por comparabilidad internacional.

Nombre del indicador: Tasa neta de asistencia a la educación superior – personas de 24 años y más.

Definición Operativa:

Descripción del indicador:

Vivienda y servicios

Nombre del Indicador: Tenencia de la vivienda de los hogares.

Definición Operativa:

$$HTV = \frac{THS_{tviv i}}{TH} * 100$$
$$i = 1, 2$$

donde, $THS_{tviv i}$ son los Total de hogares para cada i . $i =$ cada forma de tenencia de la vivienda, TH es el total de hogares.

Descripción del indicador: Porcentaje de hogares según las diferentes formas de tenencia de la vivienda (propia y totalmente pagada; propia y la está pagando; propia (regalada, donada, heredada o por posesión); prestada o cedida (no paga); por servicios; arrendada; anticresis); respecto al total de hogares.

3.2.2 Paquetes y librerías de estadística espacial aplicadas en el software estadístico R

Para el presente trabajo se utilizó en el lenguaje de programación estadístico R mediante su entorno de desarrollo RStudio, el cual permite mediante la aplicación de paquetes y librerías de SIG tipo área; llegar a determinar estas concentraciones aplicando los modelos de estadística espacial explicados en las secciones 2.2.2.2, 2.2.2.3 y 2.2.2.4 del capítulo 2. Se analizaron paquetes y funciones de mayor importancia con el fin de realizar la aplicación, para datos de los delitos de robos y sociodemográficos. Para esto se revisaron diferentes referencias bibliográficas; como (Cabrero Ortega & Gaecía Pérez, 2015), y otros autores importantes que se citan en la sección de citas bibliográficas; además se revisó información de la web relacionada con Modelos para datos de áreas y Econometría espacial; así como el documento de (Borrego Sánchez, 2018). Los paquetes más esenciales y usados para datos espaciales de tipo área son: *rgdal*, *raster*, *sp* y *spdep*. Los tres primeros paquetes son necesarios para utilizar las funciones del paquete *spdep*; el cual, como se verá a continuación, es el que visualmente se emplea para el análisis espacial. Es decir, para utilizar las funciones que analizarán la autocorrelación espacial y los distintos modelos vistos en el capítulo 2, respectivamente, es necesario cargar los demás paquetes y librerías. A continuación, se detallan cada uno de estos.

3.2.2.1 Paquete *rgdal*

Este paquete presta enlaces a la librería de abstracción de datos geoespaciales (“GDAL”) y da acceso a las operaciones de proyección/transformación desde la librería “PROJ.4”. Su uso más general es leer los datos para devolver un objeto espacial y proporcionar la proyección de coordenadas o matrices de coordenadas. Para conocer a detalle sobre el uso y funcionamiento del paquete revisar la web y la referencia (Borrego Sánchez, 2018).

3.2.2.2 Paquete *raster*

Este paquete facilita clases y funciones para manipular datos geográficos (espaciales) en formato “raster”. Los datos ráster dividen el espacio en celdas (rectángulos, píxeles) del mismo tamaño (en unidades del sistema de referencia de coordenadas). Estos datos espaciales continuos también se conocen como datos de “cuadrícula” y se contrastan con datos espaciales discretos (basados en objetos como puntos, líneas o polígonos). Este paquete es útil cuando se

emplean conjuntos de datos muy grandes; que no se pueden cargar en la memoria del computador. Las funciones se ejecutan correctamente, ya que procesan archivos grandes en trozos; se leen, computan y escriben bloques de datos, sin cargar todos los valores en la memoria al mismo tiempo.

El paquete implementa clases para datos raster y su uso es amplio, por ejemplo, en tareas como:

- Creación de objetos raster desde cero o desde archivo.
- Manejo de archivos raster computacionalmente grandes.
- Álgebra raster y funciones de superposición.
- Funciones de distancia y vecindad.
- Conversión de polígonos, líneas y puntos a datos raster.
- Modelo de predicciones.
- Resumir valores raster.
- Fácil acceso a valores de celda rasterizados.
- Graficar (hacer mapas).
- Cálculo de los números de la fila, columna y celda a coordenadas y viceversa.
- Leer y escribir varios tipos de archivos raster.

Para mayor detalle y ayuda se puede revisar en la página web de econometría espacial de (Oñate, 2022) en la sección 3 (Objetos Espaciales) y (Borrego Sánchez, 2018), donde se pueden visualizar las diferentes funciones y sus usos con mayor detalle.

3.2.2.3 Paquete *sp*

Clases y métodos para datos espaciales. Es un paquete de clases, donde se encuentra la información de ubicación espacial, para datos 2D o 3D. Dispone de funciones de utilidad, por ejemplo, para crear datos como mapas y selección espacial, de la misma manera existen métodos para recuperar coordenadas, subconjuntos, impresión, resumen, entre otros. Este paquete no tiene muchas funciones por modificar o analizar para datos espaciales, pero las clases que se definen son muy utilizadas por otros paquetes de R. El paquete *sp* introduce una serie de clases: Para datos vectoriales, los paquetes más básicos son `SpatialPoints`, `SpatialLines` y `SpatialPolygons`. Para almacenar en los atributos, las clases están disponibles con el nombre `DataFrame`, por ejemplo, `SpatialPolygonsDataFrame` y `SpatialPointsDataFrame`. Cuando se hace referencia a cualquier objeto con un nombre que comienza con `Spatial`, es común escribir `Spatial'`. Cuando se hace referencia a un objeto `SpatialPolygons` o `SpatialPolygonsDataFrame`, es común escribir `SpatialPolygons'`. (Borrego Sánchez, 2018).

Para mayor detalle sobre este paquete y su utilidad de las funciones mencionadas se sugiere ver la página web de econometría espacial de (Oñate, 2022) en la sección 3 (Objetos Espaciales), (Borrego Sánchez, 2018) y (Morales, 2012), que han servido de guía para realizar este trabajo.

3.2.2.4 Paquete *spdep*

Este paquete es uno de los más importantes para realizar análisis exploratorio e inferencia estadística de datos espaciales. Tiene un conjunto de funciones para crear matrices de pesos espaciales a partir de la contigüidad de los polígonos, para el estadístico general de productos cruzados, para pruebas como el contraste de existencia de autocorrelación espacial; así como el estadístico I de Moran o c de Geary y sus versiones locales (observar capítulos 2). También contiene funciones para estimar modelos del tipo SAR o SEM (observar capítulo 2) (Borrego Sánchez, 2018).

Para mayor conocimiento de las funciones a continuación descrita y otras funciones que incluye el paquete, se recomienda revisar la página web econometría espacial sección 3, Objetos Espaciales y también (Borrego Sánchez, 2018), guías para este trabajo.

3.2.2.4.1 Funciones y paquetes para vecindad de datos de área

Como se ha definido en el capítulo 2, en los apartados 2.2.3, 2.2.4 y 2.2.5, antes de poder hacer un contraste de hipótesis de autocorrelación espacial o modelar los datos, se debe disponer de una organización de los datos; es decir, se debe tener en cuenta la contigüidad y los datos para a partir de ello crear una matriz de pesos, misma que se emplea posteriormente en la inferencia estadística de los datos de área; para conseguir estas matrices se empleó el paquete *spdep* y se utilizó la clase *nb*. A continuación, se verán las funciones que se usan más frecuentemente en R de esta clase; para crear matrices de pesos a partir de un conjunto de datos:

- **poly2nb**: La función crea una lista de vecinos basada en regiones con límites contiguos, que comparte uno o más puntos de límite.

Sintaxis

```
poly2nb (pl, row.names = NULL, snap=sqrt (. Machine$double.eps), queen=TRUE,)
```

Argumentos

pl: Lista de polígonos de la clase SpatialPolygons.

row.names: Vector de identidad de las regiones para ser añadidas a la lista de los vecinos.

snap: Los puntos de límites inferiores a la distancia de separación para indicar la contigüidad.

queen: Si TRUE el tipo de contigüidad que se utiliza es la de tipo reina. Si FALSE, la contigüidad utilizada es la de tipo torre.

Valor

Devuelve una lista de vecinos de la clase nb.

- **knearneigh:** La función devuelve una matriz con los índices de puntos que pertenecen al conjunto de los k vecinos más cercanos entre sí.

Sintaxis

Knearneigh (x , $k = 1$, longlat = NULL,...)

Argumentos

x : Matriz de puntos coordinados o un objeto de SpatialPoints.

k : Número de vecinos más cercanos para cada área.

longlat: TRUE si las coordenadas del punto son grados decimales de longitud latitud, en cuyo caso las distancias se miden en kilómetros; si x es un objeto SpatialPoints, el valor se toma del propio objeto.

Valor

Una lista con los siguientes elementos:

nn : Matriz entera de identificadores de número de región.

np : Número de puntos de entrada.

k : Valor de entrada requerido.

dimensión: Número de columnas de x

x : Coordenadas de entrada

- **knn2nb:** La función convierte un objeto knn devuelto por knearneigh en una lista de vecinos de clase nb con una lista de vectores de números enteros que contienen identificadores de número de región vecina.

Sintaxis

knn2nb (knn, row.names = NULL, sym = FALSE)

Argumentos

knn: Un objeto knn devuelto por la función *knearneigh*.

sym: Fuerza a la lista de regiones vecinas de salida a ser simétrica.

- **nbdists:** Dada una lista de enlaces vecinos adyacentes (una lista de vecinos de tipo de objeto *nb*), la función devuelve las distancias euclídeas a lo largo de los enlaces en una lista de la misma forma que la lista de vecinos. Si `longlat = TRUE`, se usan las distancias geodésicas (great-circle distance).

Sintaxis

`Nbdists (nb, coords, longlat = NULL)`

- **nb2listw:** Esta función toma un objeto de lista de vecinos y lo convierte en un objeto o matriz de pesos. El estilo de conversión predeterminado es *W*, donde los pesos para cada entidad de área se estandarizan en sumar por filas igual a la unidad, esto se revisó en el capítulo 2.

Sintaxis

`nb2listw (neighbours, glist=NULL, style="W",...)`

Argumentos

`neighbours:` Un objeto de clase *nb*.

`glist:` Lista de pesos generales correspondientes a vecinos.

`style:` A partir de una lista binaria de vecinos, en la que las regiones están listadas como vecinas o están ausentes (por lo tanto, no están en el conjunto de vecinos para alguna definición), la función agrega una lista de ponderaciones con valores dados por el estilo de esquema de codificación elegido. *B* es la codificación binaria básica mientras que *W* está estandarizada por filas. Hay más opciones para el esquema de codificación que no utilizaremos puesto que no lo hemos visto en teoría.

3.2.2.4.2 Funciones y paquetes para dependencia espacial para datos de área

En la parte del marco teórico de este trabajo en el capítulo 2, se han detallado las pruebas mediante los estadísticos globales *I* de Moran. En el paquete de R; este estadístico es el más usado ya que es efectivo en la mayoría de situaciones. Por este motivo, se analiza detalladamente las funciones del paquete *spdep* que se utilizan en este estadístico para la prueba de dependencia espacial. También se indican algunas funciones de otros estadísticos, pero sin entrar en mayor detalle ya que su aplicación sería similar al del estadístico *I*.

- **moran.test:** Prueba de Moran para la autocorrelación espacial utilizando una matriz de ponderaciones espaciales en forma de lista de ponderaciones. Las hipótesis que subyacen a la prueba son sensibles a la forma de la gráfica de relaciones entre vecinos y otros factores, y los resultados pueden compararse con la función de “`moran.mc`” basada en permutaciones.

Sintaxis

`moran.test(x, listw, randomisation=TRUE, alternative="greater", rank = FALSE,...)`

Argumentos

x: Vector numérico de la misma longitud que la lista de vecinas de *listw*.

litw: Un objeto tipo *listw* creado por ejemplo por `nb2listw`.

Randomisation: Varianza de *I* calculada bajo la hipótesis de aleatoriedad. Si es `FALSE`, la hipótesis utilizada será normalidad.

alternative: Es una cadena de caracteres que especifique la hipótesis alternativa, debe ser una de mayor (por defecto), menor o doble.

rank: Valor lógico. Si es `FALSE` utiliza variables continuas. Si es `TRUE` utiliza la adaptación de *I* de Moran para los rangos sugeridos por Cliff y Ord.

Valor

La función devuelve el nombre del método y el de los datos usados, así como una descripción de la hipótesis alternativa. Devuelve el valor del estadístico *I* y el *p-valor* del test para saber cuál es el resultado del contraste.

- **lm.morantest**: La prueba *I* de Moran para autocorrelación espacial en residuos de un modelo lineal estimado.

Sintaxis

`lm.morantest(model, listw, alternative = "greater", resfun=weighted.residuals,...)`

Argumentos

model: Un objeto de la clase `lm` devuelto por `lm`. Los pesos pueden especificarse en el ajuste de `lm`, pero no se deben usar las compensaciones.

resfun: Por defecto: `weighted.residuals`. La función se utiliza para extraer los residuos del objeto `lm`.

Valor

Al igual que la función anterior, esta función devolverá los nombres de los métodos y datos utilizados, el *p-valor*, el estadístico *I*, su esperanza y varianza.

- **moran.mc**: Una prueba de permutación para el estadístico *I* de Moran calculado usando *nsim* permutaciones aleatorias de *x* para el esquema de ponderación espacial dado, para establecer el rango de la estadística observada en relación con los valores simulados *nsim*.

Sintaxis

`moran.mc (x, listw, nsim, alternative="greater",...)`

Argumentos

Son los mismos argumentos de la función *moran.test*.

- **localmoran:** El estadístico espacial local I de Moran se calcula para cada zona según el objeto de ponderaciones espaciales utilizado.

Sintaxis

`Localmoran (x, listw, alternative = "greater", p.adjust.method="none", mlvar=TRUE)`

Argumentos

`mlvar:` Predeterminado TRUE: los valores de la I de Moran local se informa utilizando la varianza de la variable interés (suma de desviaciones al cuadrado sobre n), pero se pueden informar como la varianza de la muestra, dividiendo por $(n - 1)$ en su lugar.

`Adjust.x:` Si TRUE, los valores de las observaciones de x que no tengan vecinos se omiten para calcular la media de x .

Cuando se va a utilizar cualquiera de estas funciones, no es necesario usar todos los argumentos, es decir, en cualquiera de estos lo fundamental son los datos a utilizar para hacer la inferencia y el tipo de matriz de pesos por utilizar, x y *listw* respectivamente. (Borrego Sánchez, 2018).

3.2.3 Código de R aplicado para los modelos espaciales

Para la estimación de los modelos de estadística espacial se aplica el siguiente código en R considerando las cuatro diferentes modalidades de robo: a personas, domicilios, unidades económicas y vehículos. A continuación, se detalla el código en cada modelo.

Para los modelos de regresión espacial, estudiado en el capítulo 2 del marco teórico, los datos se trabajaron en el lenguaje R tomando en cuenta que dentro del paquete *spdep* se puede encontrar la función *lagsarlm*, misma que es utilizada tanto para el modelo tipo SAR dada por la ecuación (10); como para el modelo espacial de Durbin, ecuación (15), ya que el modelo SDM se puede expresar como un modelo tipo SAR. También se estudia con detalle la función *errorsarlm* que se emplea en el modelo tipo SEM dado por la ecuación (14).

- **lagsarlm:** Estimación del modelo de regresión de retardo espacial.

Sintaxis

Lagsarlm (formula, data = list(), listw, Durbin, method= "---", interval=NULL, tol.solve=1.0e-10,...)

Argumentos

formula: Una descripción simbólica del modelo a ser ajustado. Los detalles de la especificación del modelo se dan para *lm()*.

data: Un marco de datos opcional que contiene las variables en el modelo. Por defecto, las variables se toman del entorno al que se llama la función.

listw: Una lista de datos creados por ejemplo por *nb2listw*.

Durbin: Si es FALSE, el modelo utilizado sería el de retraso. Si es TRUE, el modelo que utiliza es el de Durbin.

method: El método predeterminado de *method = "eigen"* usa valores propios y, por lo tanto, también puede establecer los límites inferior y superior para la búsqueda de líneas para ρ pero es no es factible para grandes n . El otro método utilizado es *method = "Matrix"* para calcular directamente el determinante de la matriz $I - \rho W$.

interval: Intervalo de búsqueda para el parámetro autorregresivo ρ .

tol.solve: Tolerancia para detectar dependencias lineales en las columnas de matrices que se van a invertir (por defecto = 1.0e-10).

- **errorsarlm:** Estimación del modelo de regresión espacial de error.

Sintaxis

Errorsarlm (formula, data=list(), listw, method="---", interval = NULL, tol.solve=1.0e-10,...)

Argumentos

Para este modelo es el mismo caso de retardo espacial y espacial de Durbin, por lo cual se omite su explicación y detalle.

Como se ha especificado anteriormente, no es necesario utilizar todos los argumentos descritos en ambos modelos espaciales. En la aplicación que se realizará a continuación y en la presentación de resultados, se observará que a la hora de estudiar los modelos y estimar sus coeficientes utilizaremos los siguientes argumentos: *formula*, *data*, *listw* y *tol.solve*. (Borrego Sánchez, 2018).

CAPITULO IV

4 RESULTADOS Y DISCUSIÓN

En este capítulo se presentan breves análisis de las variables de la base de datos georreferenciada, descrita a detalle en los capítulos 2 y 3; los resultados y la discusión de la aplicación de las diferentes pruebas de autocorrelación mediante mapas de coropletas y los modelos de regresión espacial para: tasa de robo a personas y a domicilios; robo a unidades económicas y vehículos que fueron registrados en el año 2019 en la ciudad de Cuenca, Ecuador. También se describen los modelos de regresión espaciales implementados mediante el lenguaje estadístico R a través del entorno RStudio, los cuales fueron explicados en los capítulos 2 y 3 respectivamente, para cada uno de los delitos, a fin de identificar las zonas censales urbanas de la ciudad en estudio que presentan mayor concentración. Finalmente, mediante un índice estadístico denominado Akaike (AIC), se determinará cuál de los modelos aplicados es el mejor para estimar la concentración de los diferentes tipos de robos en las distintas zonas censales urbanas; de manera espacial en la ciudad de Cuenca, Ecuador en el año 2019.

4.1 Análisis exploratorio de las características de las variables

Para analizar las características de las variables de los diferentes delitos del robo en la ciudad de Cuenca; ocurridos en el año 2019, en cada una de sus modalidades; se realizó un breve análisis exploratorio de las variables descritas en el capítulo 3; y se verificó su distribución geográfica utilizando mapas de coropletas aplicando el lenguaje R, mediante su entorno de desarrollo RStudio. De igual forma se analizaron las variables sociodemográficas del censo 2010 a un nivel de desagregación de zona censal, para de esta manera observar su distribución y concentración en cada una de estas.

Se puede observar en la Figura 1-4, que la ciudad de Cuenca, está dividida políticamente de acuerdo a parroquias urbanas y zonas censales en: Bellavista está conformada por 6, Cañaribamba tiene 3, El Batán cuenta con 6, El Sagrario 2, El Vecino consta de 8, Gil Ramírez Dávalos tiene 2, Hermano Miguel 5; al igual que Huayna Cápac, Monay y Totoracocha, Machángara tiene 4, San Blas 2, San Sebastián consta de 11, Sucre 4, y Yanuncay tiene 13; zonas censales. Lo anterior muestra un total global de 81 zonas censales distribuidas en todas las parroquias urbanas de la ciudad.

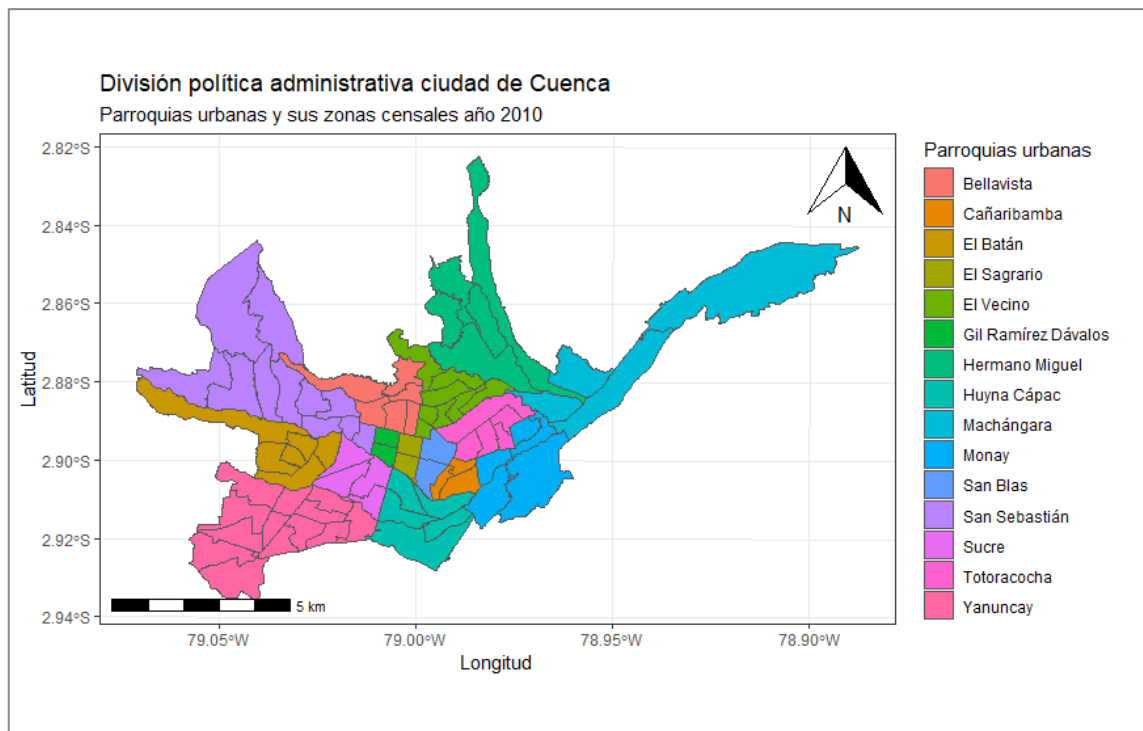


Figura 1-4: Mapa de coropletas división político administrativa de Cuenca con sus parroquias urbanas con sus respectivas zonas censales.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

4.1.1 *Análisis de las características de robo en sus diferentes modalidades*

Robo a Personas

Al realizar el análisis de este delito, se puede decir que existe una mayor incidencia en las zonas urbanas del centro de la ciudad; en las parroquias: El Batán, El Sagrario y El Vecino. Esto quiere decir que entre 6 a 12 personas de cada 100 habitantes es robado. Se puede decir que no es muy alarmante, sin embargo, se puede ver que esta incidencia se concentra más en estas zonas de la ciudad tal como podemos observar en la Figura 2-4, la cual muestra la distribución de la tasa de robo a personas denunciadas en estos lugares por cada 100 habitantes.

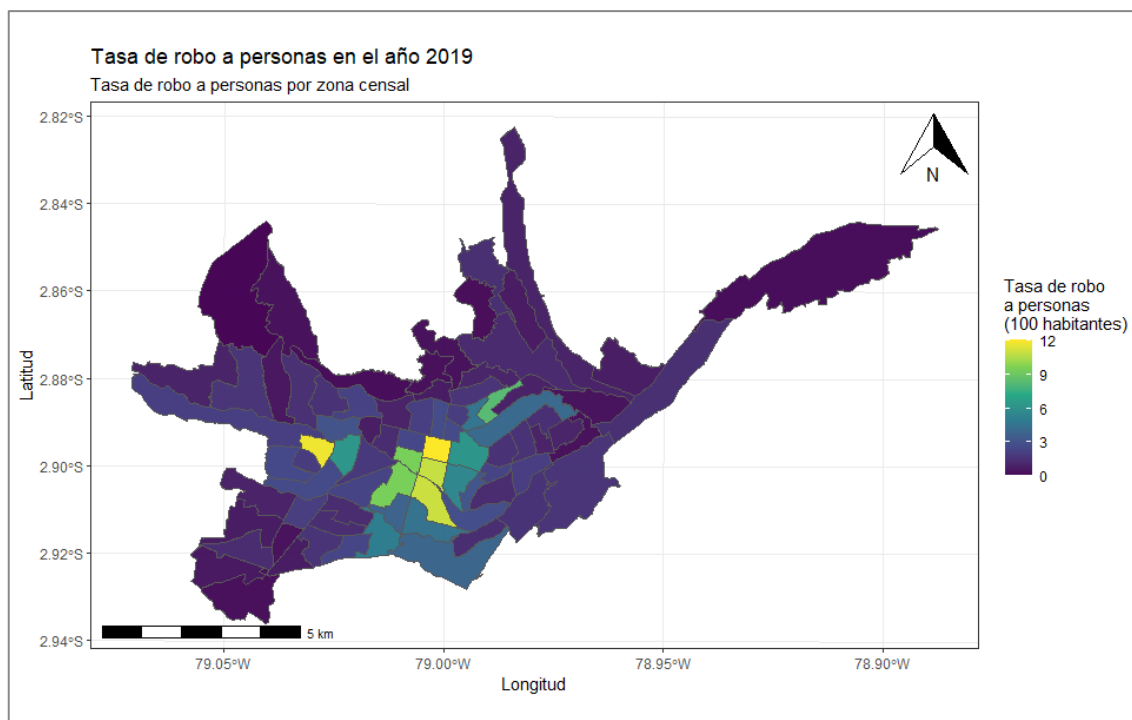


Figura 2-4: Mapa de coropletas, distribución por zonas de robo a personas, ciudad de Cuenca 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a domicilios

Al realizar el análisis de este delito, se puede observar que existe una mayor incidencia en las zonas urbanas de las parroquias: San Sebastián, El Vecino y Yanuncay, lo cual indica que de cada 100 domicilios uno de ellos sufre un robo; esto no es muy alarmante, sin embargo se puede ver que esta incidencia se concentra más en estas zonas de la ciudad tal como se observa en la Figura 3-4. Además, con base en lo mostrado en la figura antes mencionada, se puede decir que más del 90% de las zonas de las parroquias restantes tienen una tasa menor a un domicilio robado en el año 2019.

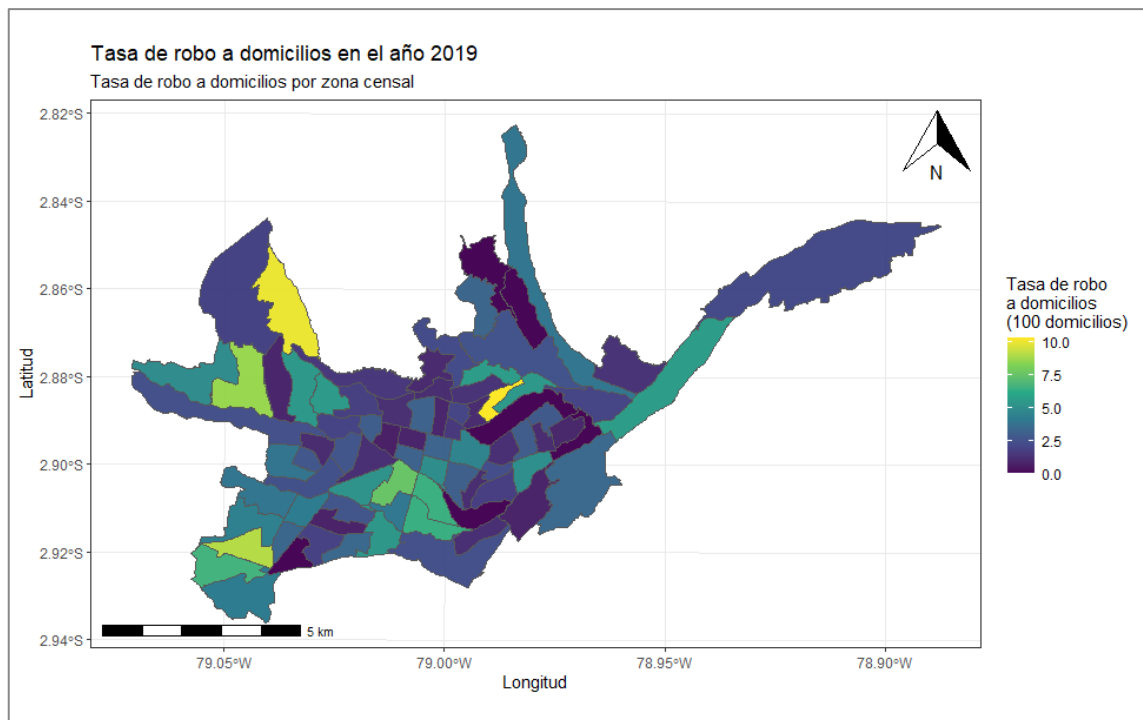


Figura 3-4: Mapa de coropletas, distribución por zonas de robo a domicilios, ciudad de Cuenca 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a unidades económicas

Al realizar el análisis de este delito se determina que existe una mayor incidencia en las zonas urbanas del centro de la ciudad, en las parroquias: El Sagrario y Huayna Cápac de Cuenca, con un promedio de 15 a 20 unidades económicas robadas durante el año 2019. Se puede decir que la concentración es de pocos casos, sin embargo, se puede ver que esta incidencia se concentra más en estas zonas de la ciudad como se observa en la Figura 4-4. Además, se puede establecer que la mayor parte de las zonas de la ciudad no se ha presentado con tanta incidencia este tipo de delito, como se puede observar en la figura en mención.

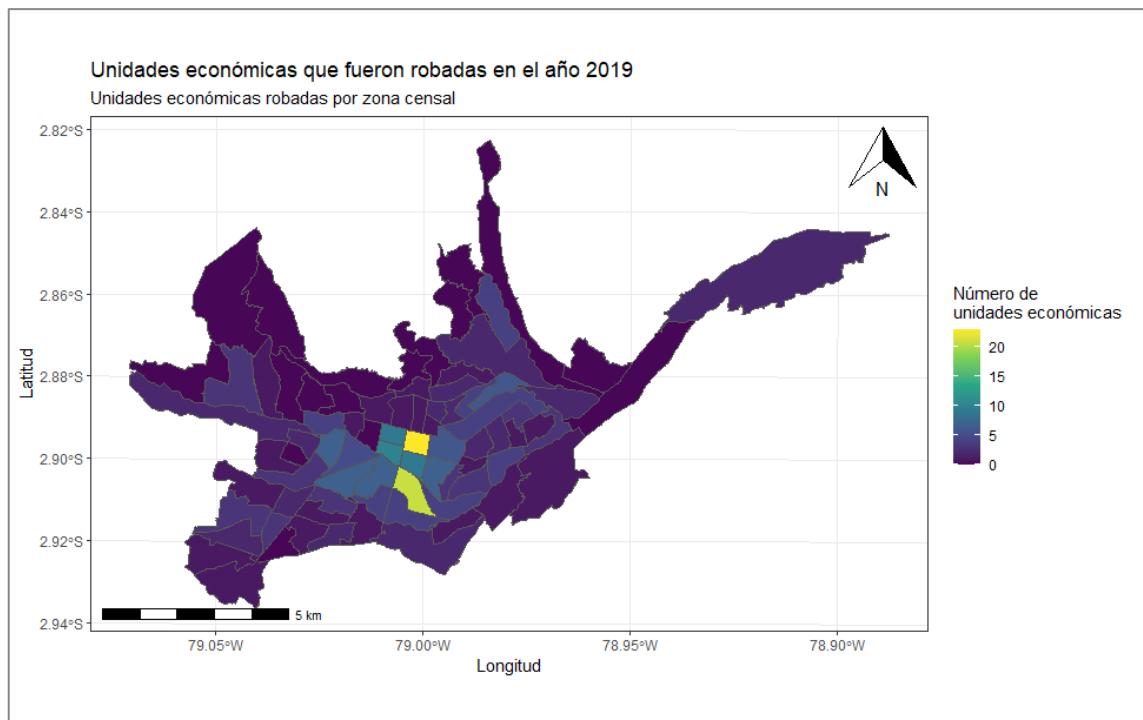


Figura 4-4: Mapa de coropletas, distribución por zonas de robo a unidades económicas, ciudad de Cuenca 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velastegui Casco, Edy, 2022.

Robo de vehículos

Al realizar el análisis de este delito, se puede decir que existe una mayor incidencia en las zonas urbanas del centro de la ciudad, con un promedio de 15 a 25 vehículos que fueron robados, en el año 2019, se puede decir que a pesar que la concentración es de pocos casos, sin embargo se puede ver que esta incidencia se concentra más en estas zonas de la ciudad como podemos observar en la Figura 5-4.

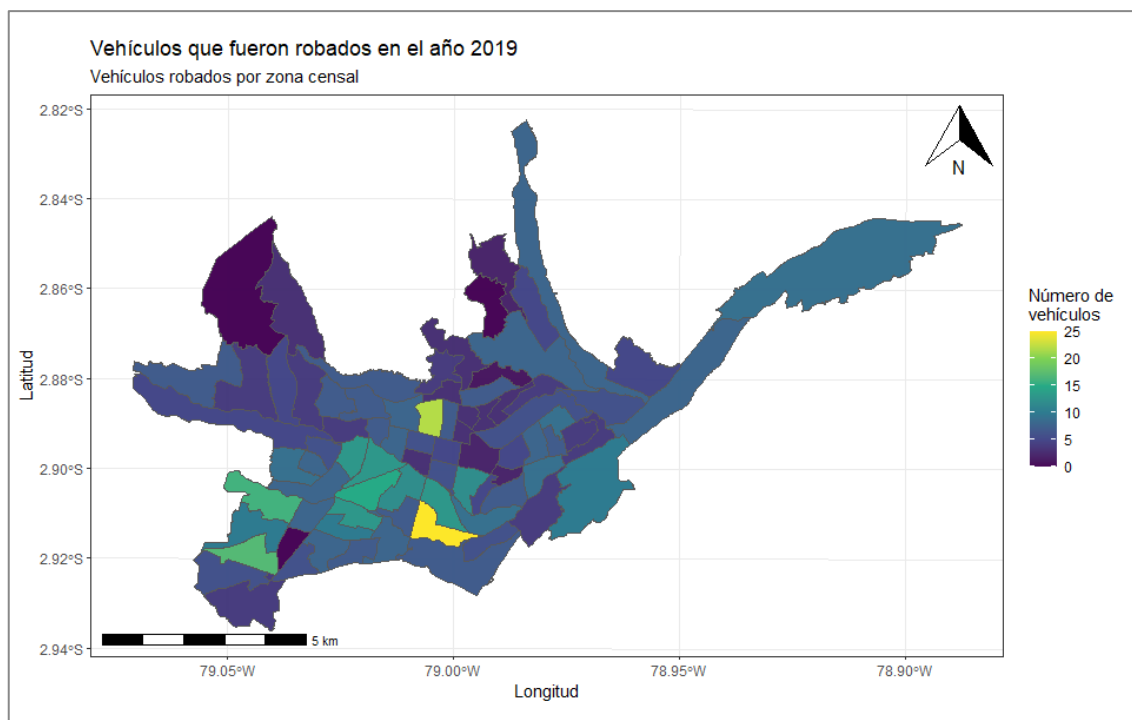


Figura 5-4: Mapa de coropletas, distribución por zonas de robo a vehículos, ciudad de Cuenca 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Eddy, 2022.

4.1.2 Variables de características sociodemográficas población ciudad de Cuenca

Demografía y distribución espacial

Población (habitantes)

De acuerdo Figura 6-4; existe una mayor concentración poblacional en las zonas urbanas de las parroquias: Hermano Miguel, Machángara, Totoracocha, Cañaribamba, Bellavista, Sucre, El Batán y San Sebastián; con un rango de población de entre 5000 a 6000 habitantes, durante el año 2019; mientras que en las parroquias: El Vecino, Monay, Huyna Cápac, San Blas, El Sagrario, Gil Ramírez Dávalos y Yanuncay, existe una distribución de 3000 a 5000 habitantes, y en las zonas restantes existen concentraciones menores a 3000 habitantes. Se determinó que en este año la distribución de la población en las diferentes zonas censales es menor a la emigración internacional como podemos observar en la Figura 7-4.

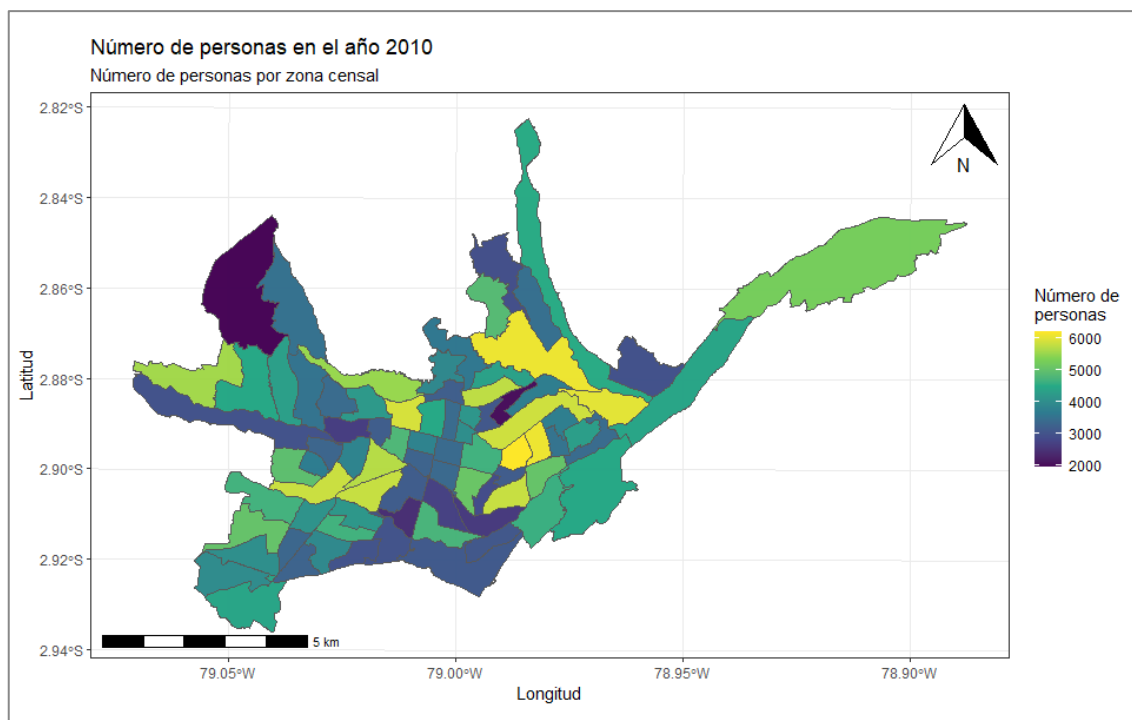


Figura 6-4: Mapa de coropletas, distribución por zonas de la población, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Emigrantes a escala internacional

Se observa la distribución de personas emigrantes a escala internacional en las diferentes zonas urbanas de la ciudad de Cuenca, como se observa en la Figura 7-4 la escala se describe entre menos de 2500 a más de 10000 personas con esta característica. Las parroquias en las que más se presenta este fenómeno son: San Sebastián, Yanuncay, Huyna Cápac, El Batán, Sucre, Gil Ramírez Dávalos, El Sagrario, San Blas, Cañaribamba, Monay, Machángara, El Vecino y Hermano Miguel; y existen zonas con una menor concentración nuevamente en las parroquias de San Sebastián y Bellavista, durante el año 2010. Es importante mencionar que existe una emigración de las personas a escala internacional importante en la mayoría de zonas urbanas de la ciudad de Cuenca.

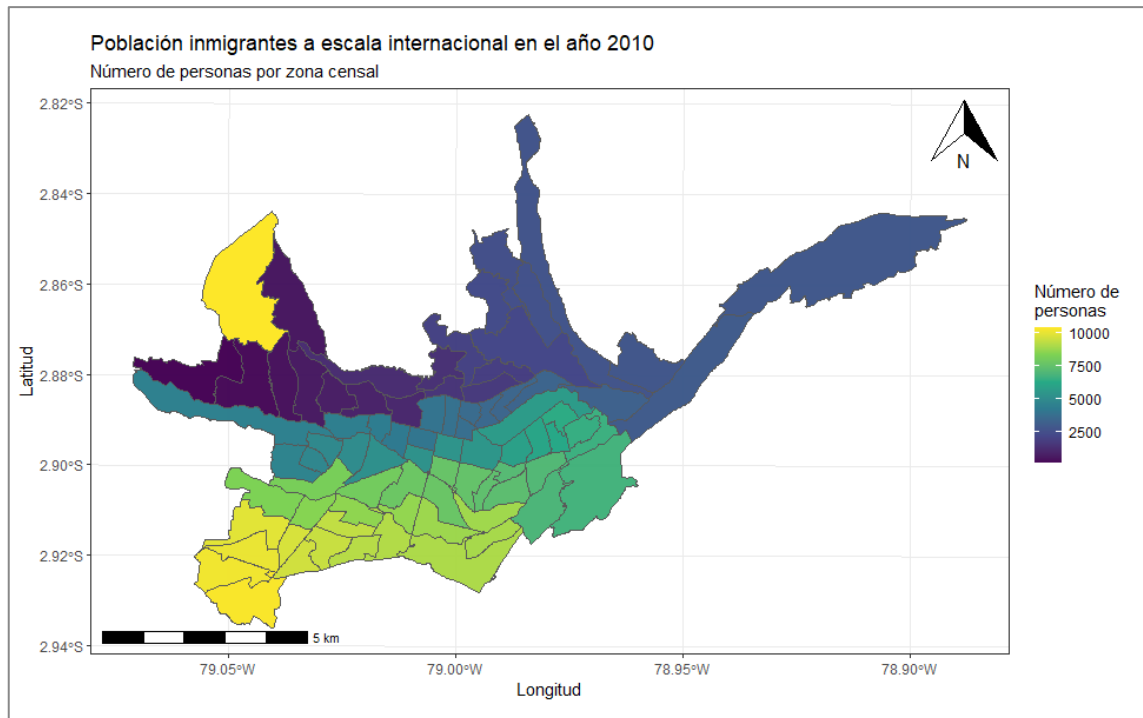


Figura 7-4: Mapa de coropletas, distribución por zonas de la población inmigrante a escala internacional, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Tasa de inmigrantes extranjeros presentes en el censo

Se puede decir que existe una mayor incidencia en las zonas urbanas de las parroquias Monay y San Sebastián, determinándose que de cada 1000 habitantes más de 80 de ellos fueron inmigrantes extranjeros en el año 2010. También se puede identificar que en las zonas de las parroquias Totoracocha y Sucre, hay entre 40 y 80 migrantes extranjeros por cada 1000 habitantes, mientras que en las parroquias: Totoracocha, Sucre, Monay, Cañaribamba, El Vecino, Yanuncay, El Batán, San Sebastián, Bellavista y Hermano Miguel, se presenta con menor intensidad este fenómeno social como se observa en la Figura 8-4.

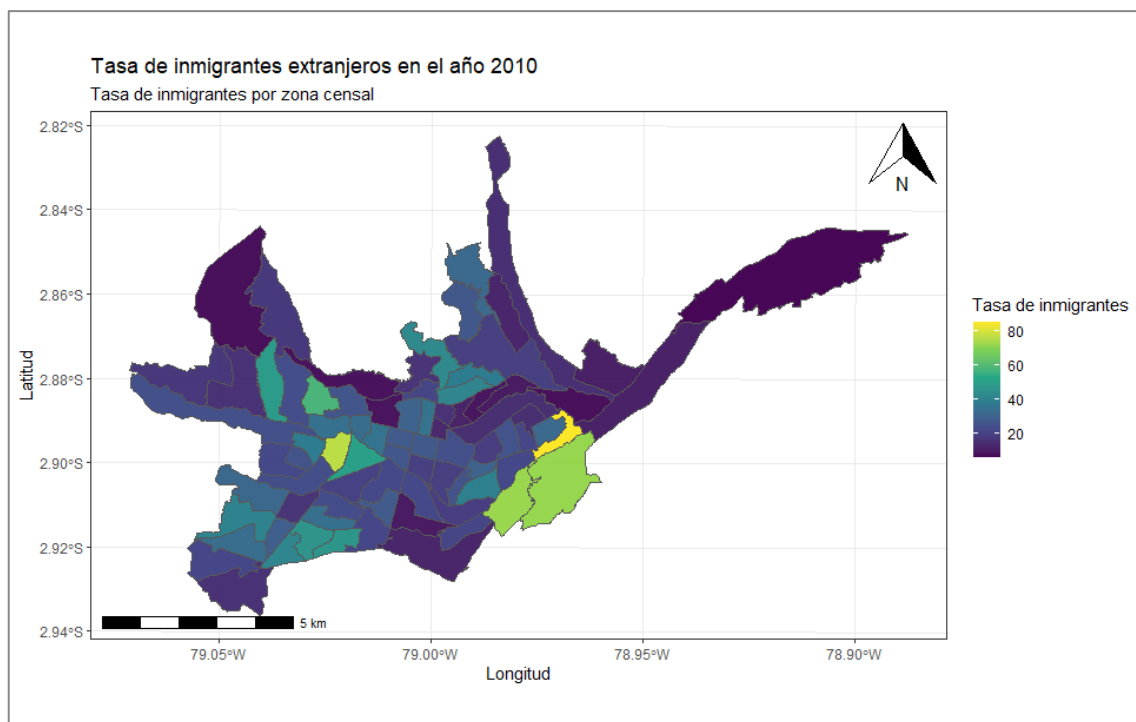


Figura 8-4: Mapa de coropletas, distribución por zonas tasa de población inmigrante extranjero, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Promedio de personas por hogar

En las diferentes zonas urbanas de Cuenca existió un promedio de entre tres a cuatro personas que conformaron un hogar durante el año 2010. También se puede decir que las zonas de las parroquias: San Sebastián, Bellavista y El Vecino, presentaron un mayor número de integrantes por hogar (mayor a cuatro), y en la mayor parte de las zonas de la ciudad, existe una cantidad menor o igual a tres integrantes por hogar como se observa en la Figura 9-4.

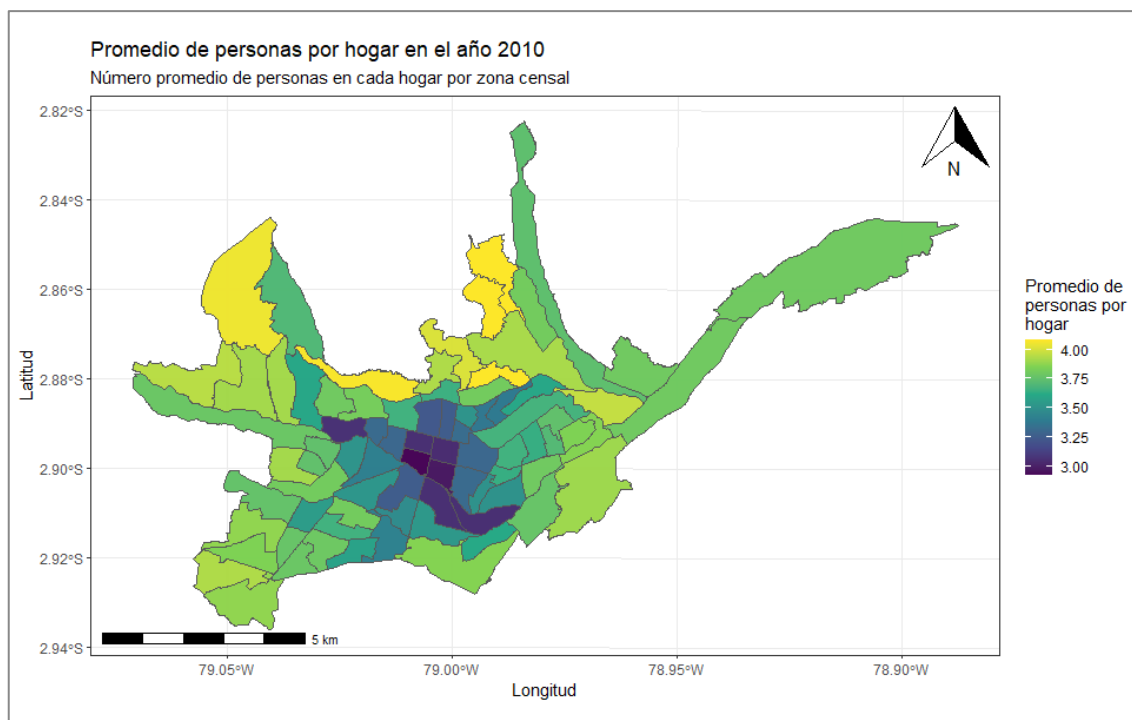


Figura 9-4: Mapa de coropletas, distribución por zonas promedio de personas por hogar, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Inmigrantes extranjeros presentes en el año 2010

En las diferentes zonas urbanas de Cuenca, existió un promedio de entre 100 a 200 personas que fueron inmigrantes extranjeros durante el año 2010. También se puede decir que en las zonas de las parroquias: San Blas, El Sagrario, Gil Ramírez Dávalos y Yanuncay, existió un mayor número de inmigrantes extranjeros, y la mayor parte de las zonas de las afueras de la ciudad existe menor o igual a 100 integrantes extranjeros como podemos observar en la Figura 10-4.

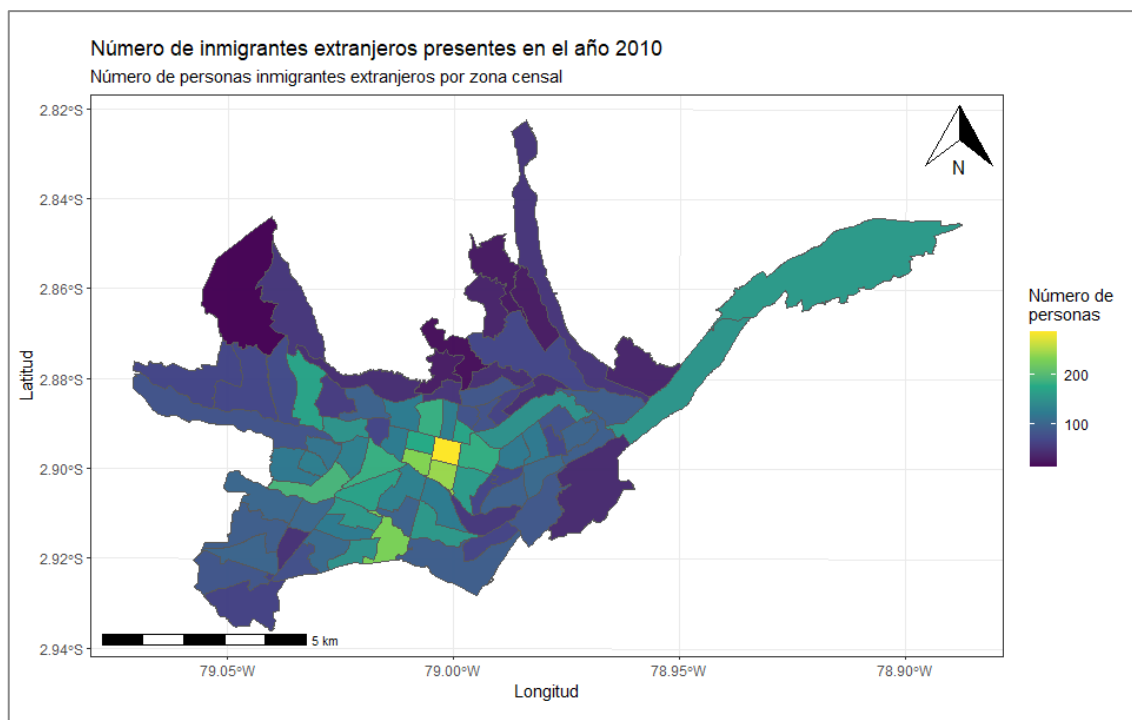


Figura 10-4: Mapa de coropletas, distribución por zonas del número de inmigrantes extranjeros presentes en el censo, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velastegui Casco, Edy, 2022.

Tasa de desplazamiento temporal por trabajo o estudio

Se establece que en las diferentes zonas urbanas de Cuenca existe una tasa de desplazamiento temporal de las diferentes personas que lo hacen por trabajo o por estudio Figura 11-4. Se observa claramente que en las parroquias Machángara, El Vecino, y Bellavista se refleja la mayor tasa de desplazamiento temporal; esto quiere decir que entre 30 a 35 personas por cada 1000 se desplazan de manera temporal por estudio o trabajo, este fenómeno sucede ya sea con personas dentro de la misma ciudad; con personas que salen de la ciudad o con aquellos que llegan de cantones cercanos, de cualquier parte del país o de otros lugares del mundo.

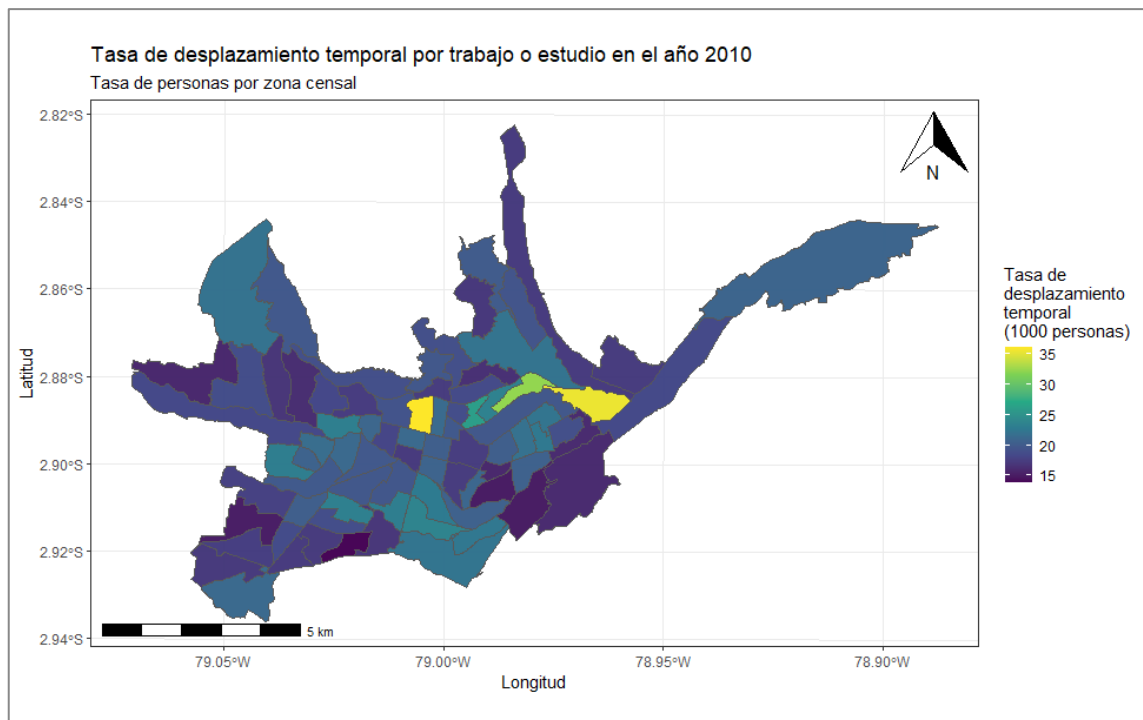


Figura 11-4: Mapa de coropletas, distribución por zonas de la tasa de desplazamiento temporal de personas por trabajo o estudio, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Número de emigrantes internacional

Durante el año 2010 en la ciudad de Cuenca se identificaron zonas urbanas con presencia de 50 a 200 personas que fueron emigrantes internacionales y se distribuyeron en sus 15 parroquias urbanas. También se debe mencionar que en las parroquias: Hermano Miguel, Sucre, Yanuncay, Monay, Cañaribamba y San Sebastián, al menos 150 personas de diferentes hogares fueron inmigrantes internacionales, y en las demás zonas urbanas del resto de parroquias existió un número menor de personas que migraron (Figura 12-4).

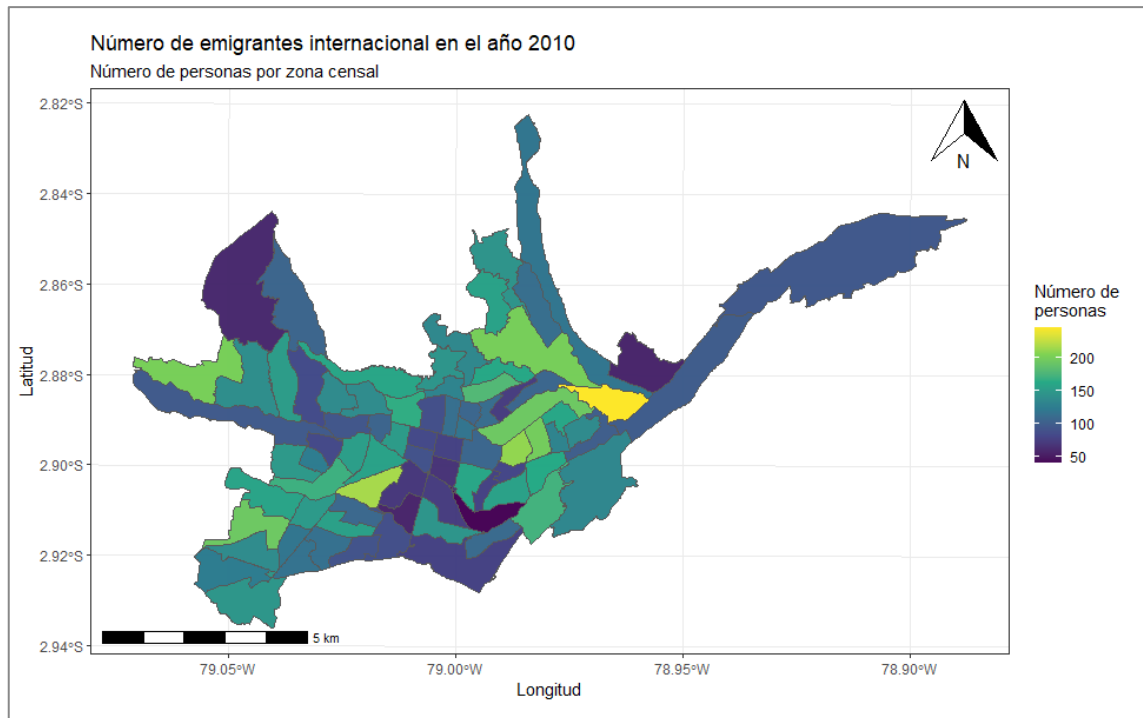


Figura 12-4: Mapa de coropletas, distribución por zonas del número de emigrantes internacional, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Eddy, 2022.

Actividades Económicas

Población en edad de trabajar (PET)

Se puede decir que existe mayor población en edad de trabajar en las zonas urbanas de las parroquias: Machángara, Hermano Miguel, Cañaribamba, Bellavista, San Blas, Huayna Cápac, Yanuncay, El Batán y Gil Ramírez Dávalos; lo que quiere decir que entre 4000 o más de 5000 personas estuvieron en edad de trabajar, durante el año 2010. Además, en las zonas urbanas de las parroquias: Machángara, Monay, Hermano Miguel, Yanuncay, Sucre, San Sebastián, Bellavista, San Blas y El Vecino, existieron entre 3000 a 4000 personas en edad de trabajar, tal como se observa en la Figura 13-4. En las demás zonas urbanas de la ciudad existió una cantidad menor a 3000 personas en edad de trabajar.

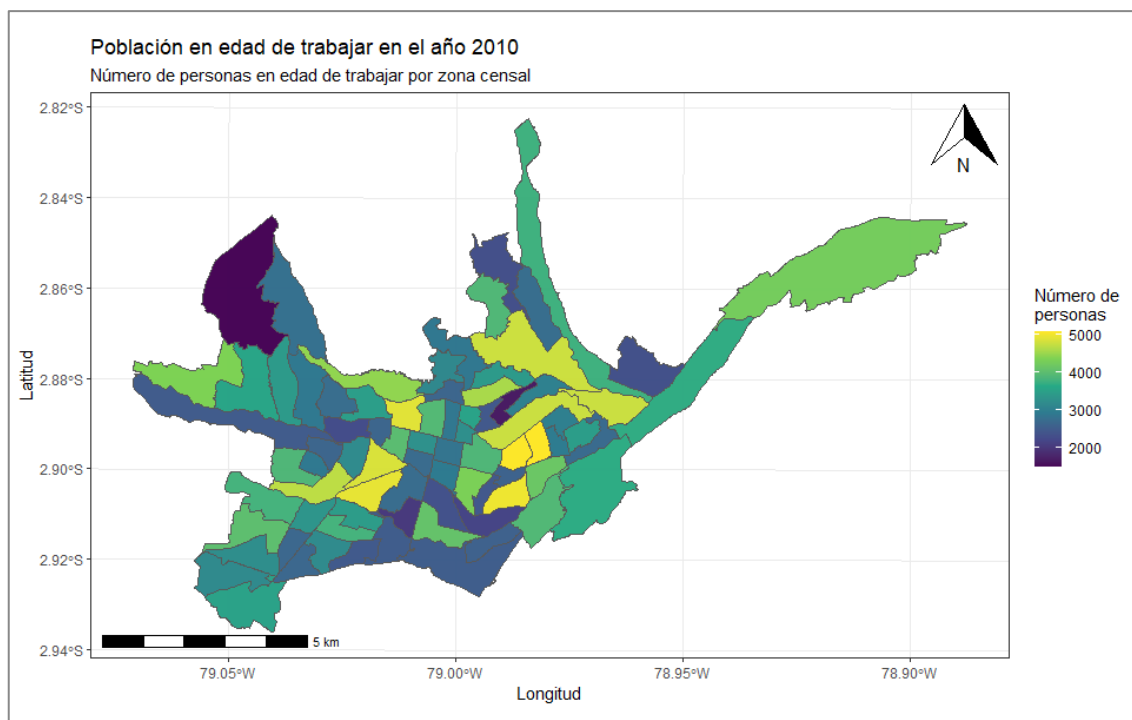


Figura 13-4: Mapa de coropletas, distribución por zonas de población en edad de trabajar, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velastegui Casco, Edy, 2022.

Población Cesante

La zona urbana con mayor concentración de personas que están desocupadas; que trabajó anteriormente; que está buscando trabajo y que está disponible para trabajar; es la parroquia San Blas con más de 50 personas cesantes, en el año 2010; mientras que en las demás zonas urbanas de la ciudad existen menos de 50 personas con estas características, como se observa en la Figura 14-4., y en referencia a las personas en edad de trabajar Figura 13-4.

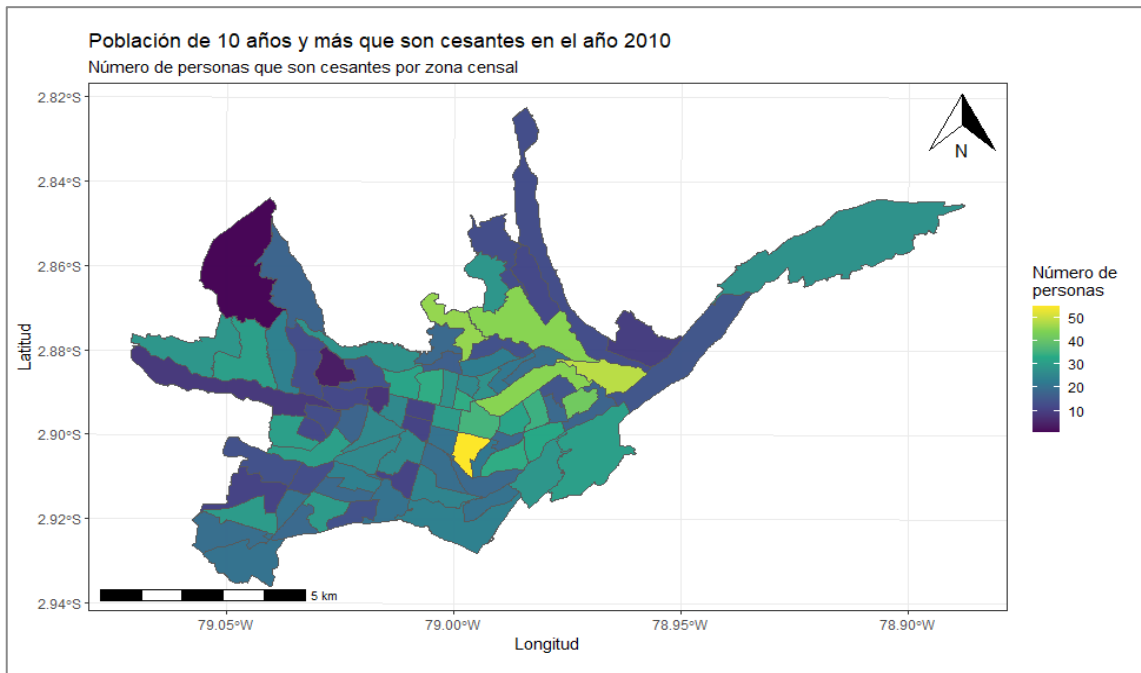


Figura 14-4: Mapa de coropletas, distribución por zonas promedio de personas por hogar, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Población ocupada que realizó su actividad económica dentro del hogar

Existieron pocas personas que realizaron actividades económicas dentro de sus hogares en las zonas urbanas de todas las parroquias de la ciudad de Cuenca; distribuidas entre 100 y más de 400 personas, durante el año 2010, como se observa en la Figura 15-4.

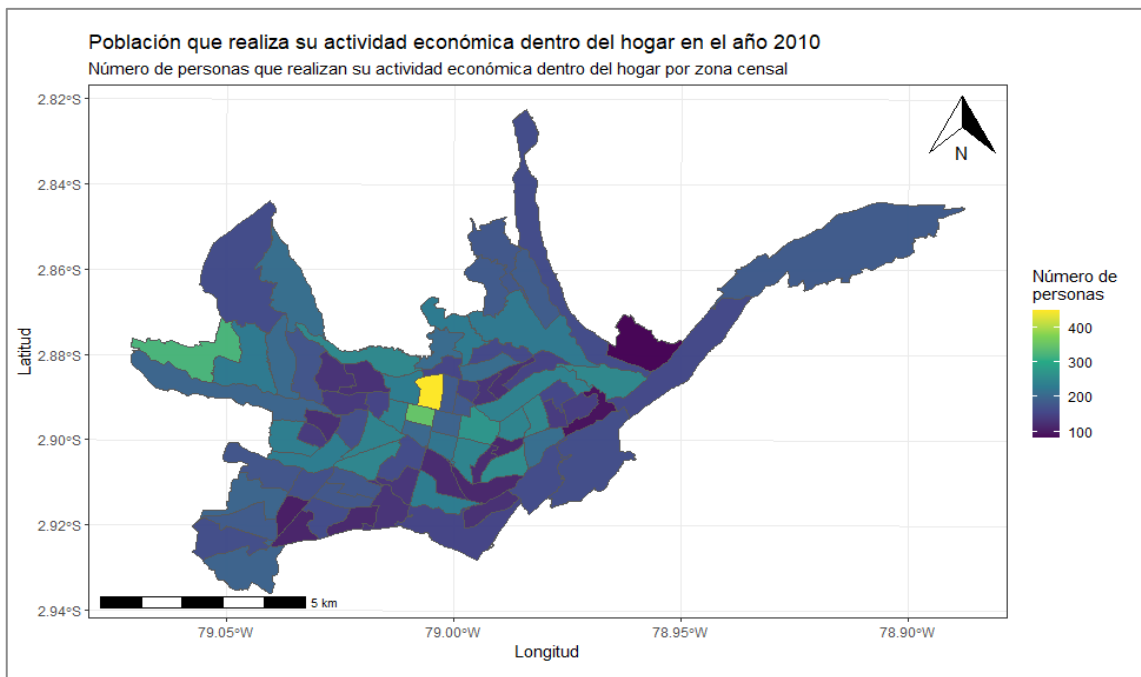


Figura 15-4: Mapa de coropletas, distribución por zonas población que realiza su actividad económica dentro del hogar, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Población Económicamente Inactiva (PEI)

Existió una distribución de personas en edad de trabajar que no contaba con un empleo durante el año 2010 en las diferentes zonas urbanas de la ciudad de Cuenca, con mayor intensidad entre 1000 a más de 2000 como se observa en la Figura 16-4.

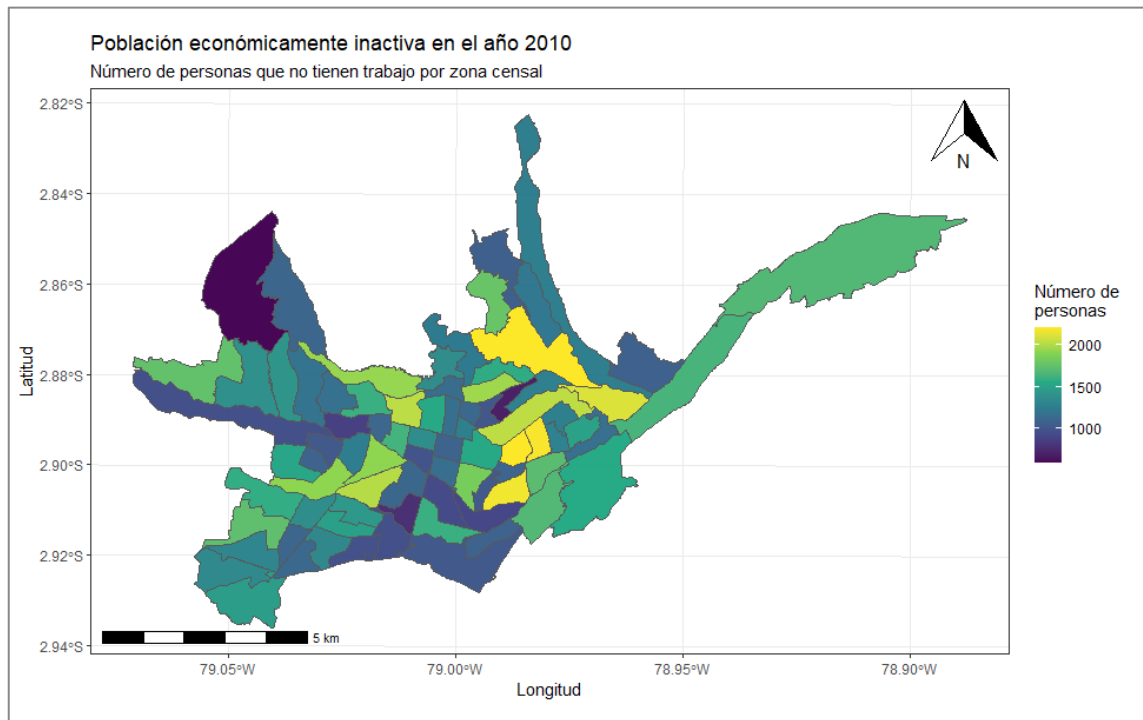


Figura 16-4: Mapa de coropletas, distribución por zonas PEI, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Educación

Tasa de Analfabetismo

Se puede decir que se presentó una mayor tasa de analfabetismo en las zonas urbanas de la parroquia San Sebastián; lo que indica que de cada 1000 habitantes, seis o más personas presentaron un grado de analfabetismo durante el año 2010; mientras que en las demás zonas urbanas de la ciudad de Cuenca, menos de seis personas de cada 1000 habitantes tuvieron este problema, como se observa en la Figura 17-4.

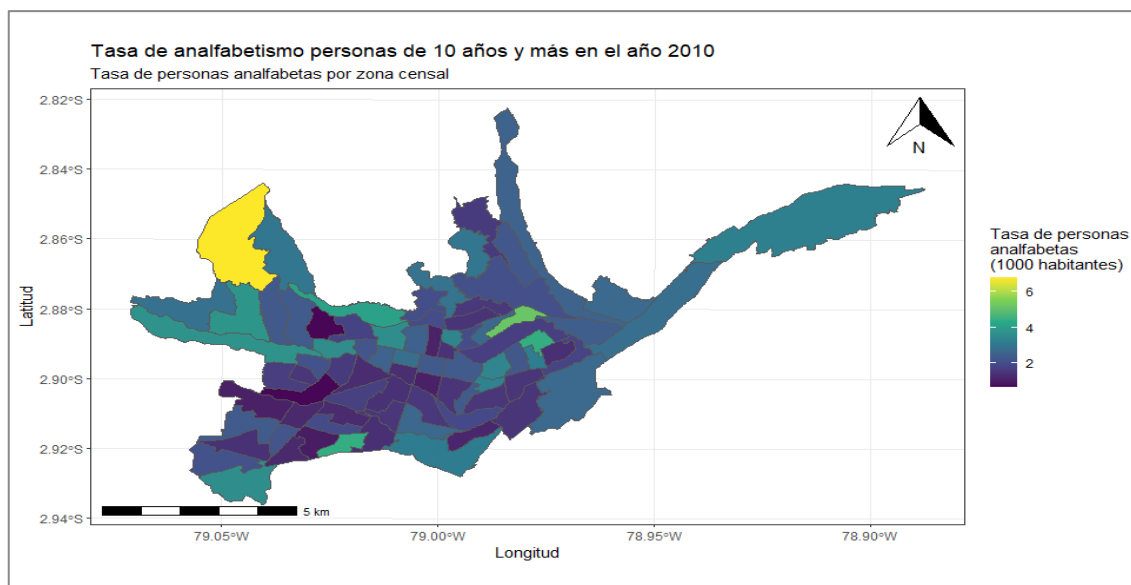


Figura 17-4: Mapa de coropletas, distribución por zonas tasa de analfabetismo personas de 10 años y más, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Tasa neta de asistencia a la educación superior – población de 24 años y más

En el año 2010 existió una mayor tasa neta de asistencia a la educación superior en personas de 24 años y más en las zonas urbanas de las parroquias: San Sebastián, Machángara, Totoracocha, Huayna Cápac, Cañaribamba, San Blas, El Vecino, Gil Ramírez Dávalos, El Sagrario, Sucre, Yanuncay, El Batán y Bellavista, lo que quiere decir que de cada 100 habitantes, de 70 a 90 personas de 24 años y más tuvieron acceso a la educación superior; mientras que en las demás zonas urbanas de la ciudad, menos de 70 personas de cada 100 habitantes asistieron a una institución de educación superior, como se observa en la Figura 18-4.

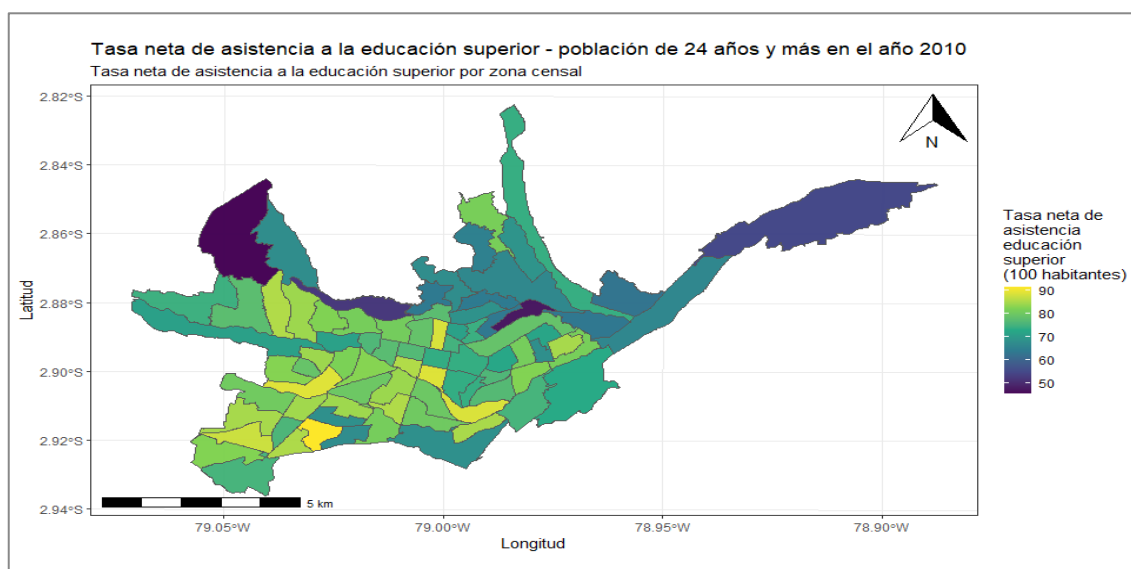


Figura 18-4: Mapa de coropletas, distribución por zonas tasa de tasa de asistencia escolar población de 24 años y más, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Vivienda y servicios

Tenencia de la vivienda propia y totalmente pagada de los hogares.

Se puede decir que existieron pocos hogares que contaron con viviendas propias y totalmente pagadas en las diferentes zonas urbanas de la ciudad de Cuenca, que van de 200 a 800, durante el año 2010. También se puede observar que las zonas donde se concentraron estos hogares en mayor parte son las parroquias: Yanuncay, Sucre, Huyna Cápac, Cañaribamba, Bellavista y Machángara (Figura 19-4). Además, se puede decir que existen 11 zonas urbanas distribuidas en las parroquias de: San Sebastián, Hermano Miguel, El Vecino, Yanuncay y Monay donde hubo una concentración menor a 200 hogares con vivienda propia y totalmente pagada.

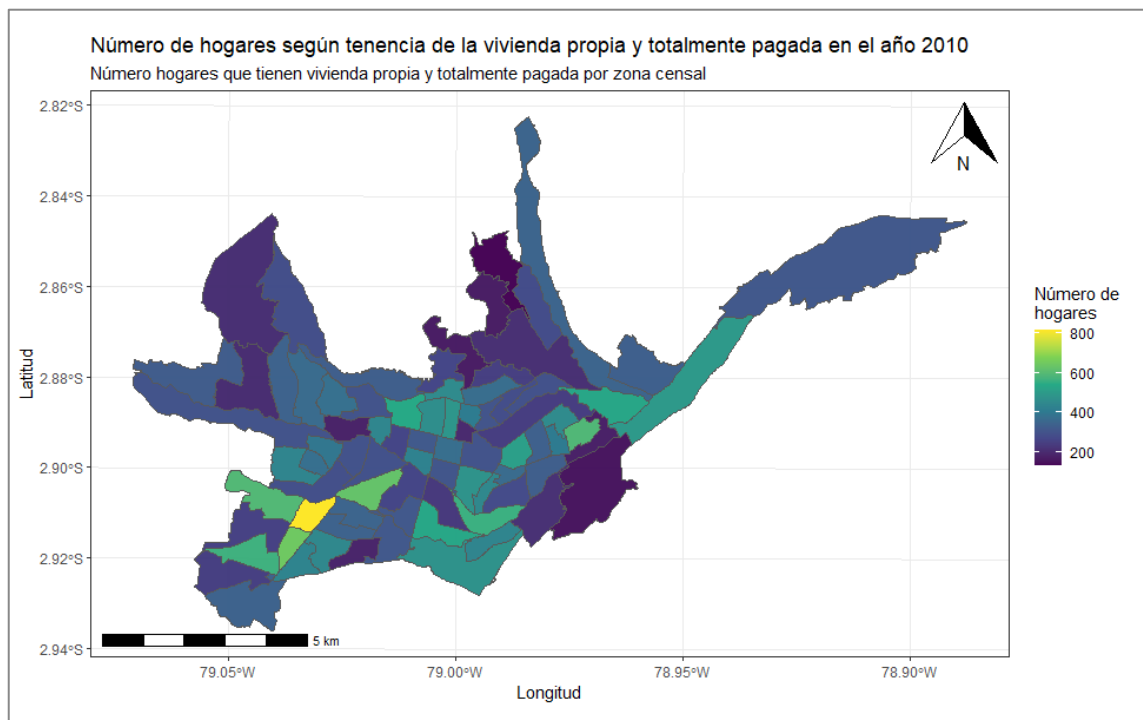


Figura 19-4: Mapa de coropletas, distribución por zonas hogares según tenencia de la vivienda propia y totalmente pagada, ciudad de Cuenca 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

4.2 Análisis de resultados modelos de retardo espacial, error espacial y de Durbin

Antes de entrar a la presentación y discusión de los resultados de los modelos espaciales aplicados para identificar las zonas de mayor concentración en el delito de robo a personas, domicilios, unidades económicas y vehículos, también se analizó su relación con las variables sociodemográficas descritas y analizadas en el capítulo 3 y sección 4.1.2 del capítulo 4. Para el análisis de la vecindad se implementó código mediante el lenguaje estadístico R, con su entorno de desarrollo RStudio, en el que se aplicaron los tres criterios: *Rook*, *Queen* y *k-cercanos*. A continuación, se presentan los resultados obtenidos y su respectivo análisis.

Se puede observar que al aplicar el criterio de *Rook*, las zonas que comparten un borde son las más cercanas o vecinas, pero también, es importante mencionar que al aplicar el criterio de *Queen*, se tiene mucha similitud al criterio de *Rook* aunque el primero tiene una relación no solamente con las zonas vecinas que están divididas por un borde; sino también por un vértice, como se observa en las Figura 20-4 y Figura 21-4, respectivamente. Mientras tanto al aplicar el criterio de *k-cercanos*, método que permite seleccionar el número de vecinos más cercanos de cada zona censal y se aplica debido a que las diferentes zonas geográficas o sus polígonos son de diferente tamaño; se obtuvo el mismo número de vecinos sin importar el tamaño que tiene cada zona; para este caso se tomaron 6 vecinos más cercanos; donde k es el número de vecinos, es así que cuando $k = 6$, se pudo determinar que la vecindad espacial se genera entre las zonas más cercanas espacialmente (Figura 22-4). Se puede ver claramente que los métodos *Rook* y *Queen* en este caso generan resultados similares en términos de vecindad en el mapa de la ciudad de Cuenca de acuerdo a las diferentes zonas censales. El modelo *k-cercanos* realiza más relaciones entre zonas vecinas, ya que este método tiene una condición importante, el cual considera el número de relaciones, sin considerar el tamaño de cada zona geográfica como se puede observar en la Figura 22-4. En este sentido y de acuerdo a la bibliografía revisada y citada en los capítulos 2 y 3 en el que aplican el método de *Queen* para la implementación de modelos de regresión espacial; para este trabajo se utilizó el método de vecindad de *Queen* para la aplicación de los diferentes modelos espaciales.

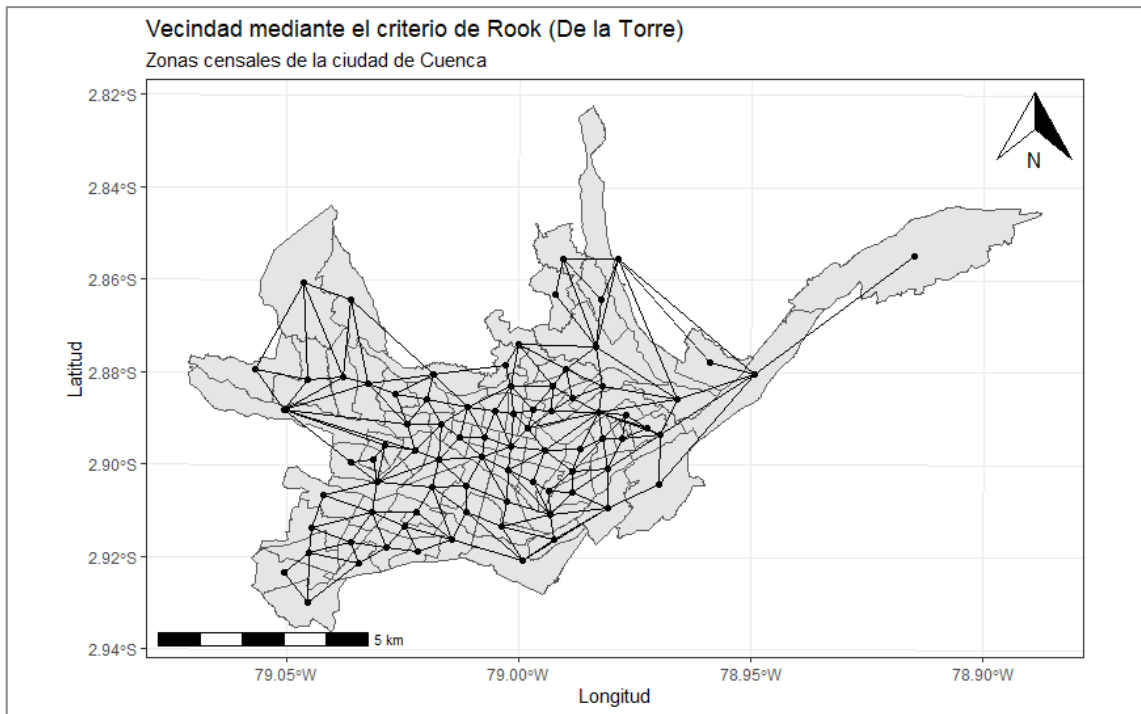


Figura 20-4: Mapa de la ciudad de Cuenca por zonas urbanas, modelo de vecindad, criterio de Rook.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

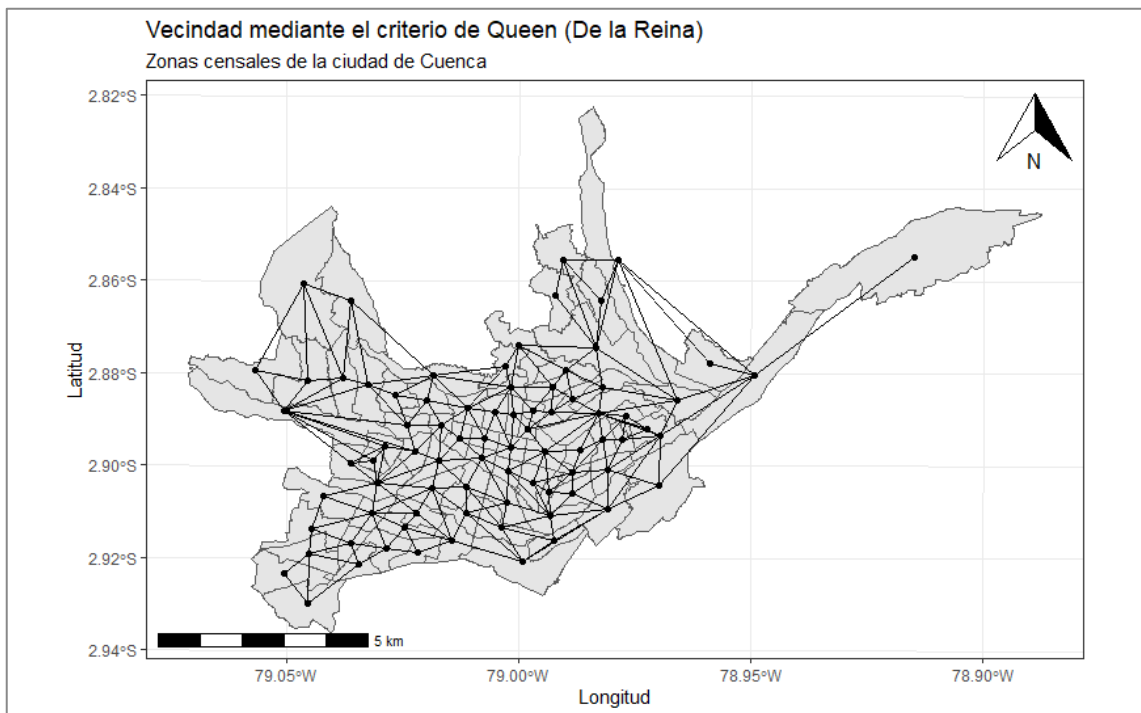


Figura 21-4: Mapa de la ciudad de Cuenca por zonas urbanas, modelo de vecindad, criterio de Queen.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

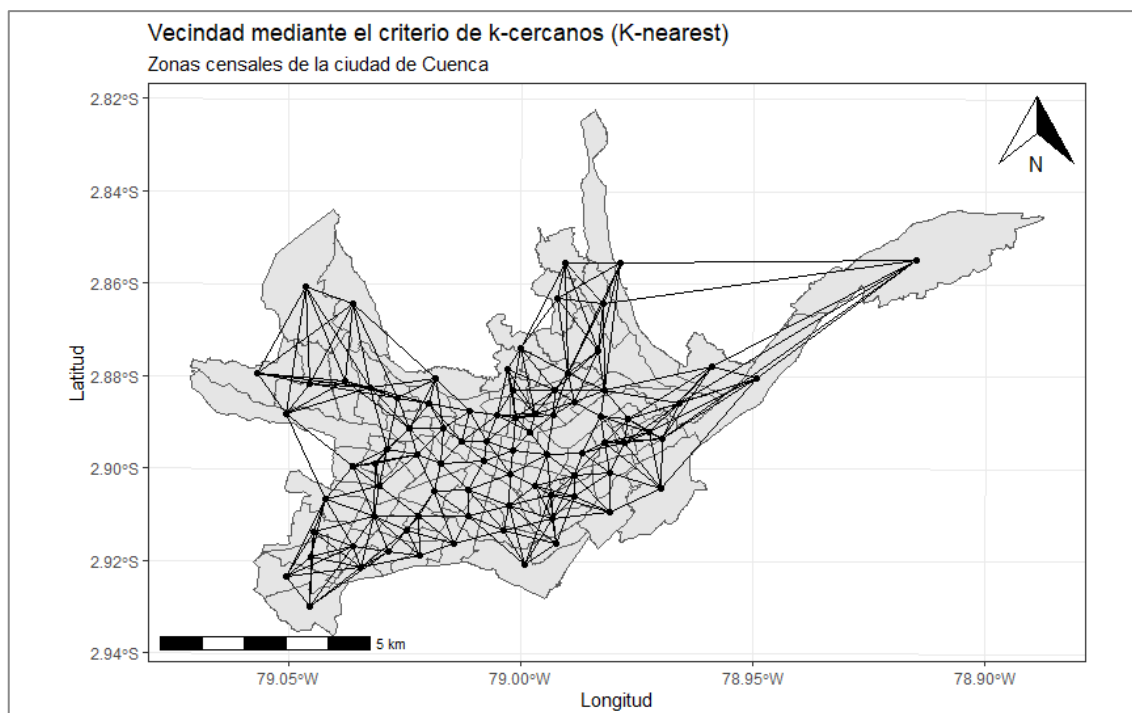


Figura 22-4: Mapa de la ciudad de Cuenca por zonas urbanas, modelo de vecindad, criterio de K-cercanos.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

4.2.1 Resultados del modelo de retardo espacial

A continuación, se presentan los resultados del modelo de retardo espacial, en el cual se expresa la dependencia espacial de las variables en función de los distintos robos (personas, domicilios, unidades económicas y vehículos), es decir, permite intuir si existe una correlación netamente por los valores, o si ésta también se da por el comportamiento en el espacio.

Robo a personas: El valor de dependencia espacial (ρ_0) es 0.436, lo cual indica una mediana dependencia espacial (cercano a 0.5) de las distintas variables en función del comportamiento espacial de la tasa de robo a personas, este valor es estadísticamente significativo, pues el valor $-p$ es 0.000320.

Tabla 1-4: Resultado del modelo de retardo espacial de la tasa de robo a personas año 2019.

Call: lagsarlm(formula = frm, data = shpf, listw = seaw)

```
Residuals:
    Min       1Q   Median       3Q      Max
-4.463116 -0.928934 -0.081929  0.548046  8.682790

Type: lag
Coefficients: (asymptotic standard errors)
            Estimate Std. Error z value Pr(>|z|)
(Intercept) 10.457727517  5.562930850  1.8799 0.0601223
BD_TA       -0.144483736  0.286383895 -0.5045 0.6139026
BD_Tasa_Na  0.005611462  0.034970477  0.1605 0.8725165
BD_Tasa_Dt -0.003338050  0.062092104 -0.0538 0.9571267
BD_Inmi_Ex  0.019741732  0.005585043  3.5347 0.0004082
```

```

BD_PMI      -0.000051054  0.000074735 -0.6831  0.4945181
BD_NEI      0.007007440  0.009549411  0.7338  0.4630654
BD_TIEpc    -0.013523274  0.014944960 -0.9049  0.3655333
BD_PPH      -1.75009879  1.104188634 -1.5849  0.1129929
BD_PET      -0.000861416  0.001182283 -0.7286  0.4662439
BD_PCes     0.046138552  0.024632488  1.8731  0.0610577
BD_POdh     -0.006650199  0.005011022 -1.3271  0.1844709
BD_PEI      -0.000825281  0.002600021 -0.3174  0.7509301
BD_HTV      -0.002495321  0.001802434 -1.3844  0.1662307

```

```

Rho: 0.43586, LR test value: 8.4513, p-value: 0.0036477
Asymptotic standard error: 0.12112
Z-value: 3.5987, p-value: 0.00031983
Wald statistic: 12.951, p-value: 0.00031983

```

```

Log likelihood: -160.8086 for lag model
ML residual variance (sigma squared): 2.9811, (sigma: 1.7266)
Number of observations: 81
Number of parameters estimated: 16
AIC: 353.62, (AIC for lm: 360.07)
LM test for residual autocorrelation
test value: 0.11488, p-value: 0.73465

```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Una vez que se identificó las variables sociodemográficas que tienen una correlación espacial significativa en el modelo espacial de SAR y explican la alta incidencia de personas que fueron robadas.

Por lo cual, el modelo me permite concluir que existe una dependencia espacial estadísticamente significativa para la variable de migración extranjera ya que el p-valor es menor al 5%.

Entonces el modelo de Retardo Espacial ajustado para la Tasa de robo a personas a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$Tasa\ robo\ a\ personas_i = \rho * lag.Tasa_R_{-i} + \beta_0 + \beta_1 * BD_Inmi_Ex_i + \varepsilon_i$$

donde,

$$\varepsilon_i \sim N(0, \sigma^2)$$

$$lag.Tasa_R_{-i} = \sum_j^n W_{ij} * Tasa_R_{-i}$$

Una vez que se han identificado las variables sociodemográficas con una correlación espacial significativa en la mayor incidencia de las personas que fueron robadas; entonces el modelo de retardo espacial que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 2-4.

Tabla 2-4: Resultado del modelo de retardo espacial ajustado de la tasa de robo a personas año 2019.

```
Call: lagsarlm (formula = fr, data = bdta, listw = seaw)
```



```

Residuals:
    Min       1Q   Median       3Q      Max
-3.17347 -0.95653 -0.27869  0.33771  8.84760

Type: lag
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.9297349  0.4882251 -1.9043   0.05687
BD_Inmi_Ex   0.0175167  0.0043697  4.0087  0.00006106

Rho: 0.59562, LR test value: 18.717, p-value: 0.000015165
Asymptotic standard error: 0.10408
      Z-value: 5.7229, p-value: 0.000000010475
Wald statistic: 32.751, p-value: 0.000000010475

Log likelihood: -174.0106 for lag model
ML residual variance (sigma squared): 3.9596, (sigma: 1.9899)
Number of observations: 81
Number of parameters estimated: 4
AIC: 356.02, (AIC for lm: 372.74)
LM test for residual autocorrelation
test value: 0.9322, p-value: 0.33429

```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Como se observa en la Tabla 2-4., se muestran cinco parámetros estimados que son los que se sustentan en el capítulo 2, los mismos que se necesitan para obtener el modelo. Estos parámetros son: el parámetro ρ el mismo que mide y determina el nivel de relación autorregresiva espacial que es estimado por Rho igual a 0.5956, coeficientes β_q dados por $\beta_0 = -0.9297349$, $\beta_1 = 0.0175167$ y por último la desviación típica de los residuos generados por este método, es decir $\sigma = 1.9899$. Además, se obtuvo el valor-p de un test realizado a los residuos que generan, cuyo valor es 0.334, con lo cual mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial. Lo anterior quiere decir que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SAR, la dependencia de los residuos desaparece.

Se presenta la ecuación del modelo de retardo espacial para la tasa de robo a personas por zona urbana en la ciudad de Cuenca año 2019:

$$Tasa\ robo\ a\ personas = 0.5956 * lag.Tasa_R_ - 0.9297349 + 0.0175167 * BD_Inmi_Ex$$

Esto quiere decir que, por cada inmigrante extranjero presente en una determinada zona de ciudad de Cuenca, la tasa de robo a personas disminuye en una unidad.

Robo a domicilios: El valor de dependencia espacial (ρ_0) es 0.20, lo cual indica que no existe una correlación espacial dependiente entre las variables sociodemográficas analizadas; debido a que su valor-p no es estadísticamente significativo, sumado a que las variables de forma individual tampoco presentan un valor estadísticamente significativo.

Tabla 3-4: Resultado del modelo de retardo espacial de la tasa de robo a domicilios año 2019.
 Call:lagsarlm(formula = frm, data = shpf, listw = seaw)

```

Residuals:
      Min       1Q   Median       3Q      Max
-3.38052 -1.53054 -0.45040  0.97194  6.64038

Type: lag
Coefficients: (asymptotic standard errors)
              Estimate      Std. Error z value Pr(>|z|)
(Intercept)  1.9895553146  6.6836659170  0.2977  0.76595
BD_TA        -0.1338143063  0.3604401314 -0.3713  0.71045
BD_Tasa_Na   0.0012175658   0.0440186024  0.0277  0.97793
BD_Tasa_Dt  -0.0466747236   0.0783624615 -0.5956  0.55143
BD_Inmi_Ex   0.0045472186   0.0069744275  0.6520  0.51441
BD_PMI       0.0000040489   0.0000939435  0.0431  0.96562
BD_NEI       0.0074569635   0.0120567009  0.6185  0.53625
BD_TIEpc    -0.0349971123   0.0188674356 -1.8549  0.06361
BD_PPH       0.7995650806   1.3085169194  0.6110  0.54117
BD_PET       0.0010937438   0.0014887735  0.7347  0.46255
BD_PCes      0.0050238757   0.0309898431  0.1621  0.87122
BD_Podh      0.0055094583   0.0063495625  0.8677  0.38556
BD_PEI      -0.0045957992   0.0032795970 -1.4013  0.16112
BD_HTV      -0.0007069934   0.0022743698 -0.3109  0.75591

Rho: 0.20247, LR test value: 1.3772, p-value: 0.24058
Asymptotic standard error: 0.15753
      Z-value: 1.2853, p-value: 0.19869
Wald statistic: 1.652, p-value: 0.19869

Log likelihood: -178.3324 for lag model
ML residual variance (sigma squared): 4.7466, (sigma: 2.1787)
Number of observations: 81
Number of parameters estimated: 16
AIC: 388.66, (AIC for lm: 388.04)
LM test for residual autocorrelation
test value: 1.7942, p-value: 0.18042
  
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

En este caso, tal como se ilustró, no se presenta ninguna covariable que sea estadísticamente significativa, pues en todos los casos el valor-p está por encima de 0.05; sin embargo, como se puede ver la variable de tasa de inmigrantes de extranjeros es la que presenta el valor-p más cercano a 0 (propiamente ~ 0.06). El modelo fue nuevamente ajustado con esta única variable, dando como resultado un valor rho de 0.21 el cual tiene un valor-p de 0.19, es decir, no es estadísticamente significativo. Por otro lado, el valor del intercepto es de 3.15 con un valor-p muy cercano a 0 (propiamente ~0.0000009), y la variable de tasa de inmigrantes extranjeros presento un valor-p significativo de 0.07, es decir, una confiabilidad del 93%.

El modelo de retardo espacial ajustado para la tasa de robo a domicilios a nivel de zonas urbanas en la ciudad de Cuenca año 2019, queda representada de la siguiente manera:

$$Tasa\ robo\ a\ domicilios_i = \rho * lag.Tasa_1_i + \beta_0 + \beta_1 * BD_TIEpc_i + \varepsilon_i$$

donde,

$$\varepsilon_i \sim N(0, \sigma^2)$$

$$lag.Tasa_1_i = \sum_j^n W_{ij} * Tasa_1_i$$

Una vez identificadas las variables sociodemográficas que tienen una correlación espacial significativa en la mayor incidencia de las personas que fueron robadas; entonces el modelo de retardo espacial que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación mediante la Tabla 4-4.

Tabla 4-4: Resultado del modelo de retardo espacial ajustado de la tasa de robo a domicilios año 2019.

```
Call:lagsarlm(formula = fr, data = bdta, listw = seaw)

Residuals:
    Min       1Q   Median       3Q      Max
-3.35873 -1.79142 -0.50866  1.22565  6.82304

Type: lag
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  3.151156   0.712769  4.4210 0.000009824
BD_TIEpc    -0.027930   0.015801 -1.7676  0.07713

Rho: 0.21203, LR test value: 1.7152, p-value: 0.19031
Asymptotic standard error: 0.15977
      Z-value: 1.3271, p-value: 0.18448
Wald statistic: 1.7612, p-value: 0.18448

Log likelihood: -180.9642 for lag model
ML residual variance (sigma squared): 5.0613, (sigma: 2.2497)
Number of observations: 81
Number of parameters estimated: 4
AIC: 369.93, (AIC for lm: 369.64)
LM test for residual autocorrelation
test value: 0.008297, p-value: 0.92742
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Como se observa en la Tabla 4-., se muestran cinco parámetros estimados mismos que se sustentan en el capítulo 2, los cuales se necesitan para obtener el modelo y son los siguientes: el parámetro ρ que mide y determina el nivel de relación autorregresiva espacial que es estimado por Rho igual a 0.212, coeficientes β_q dados por $\beta_0 = 3.151156$, $\beta_1 = -0.02793$ y por último la desviación típica de los residuos generados por este método, es decir $\sigma = 2.2497$. Además, se obtiene el valor-p de un test realizado a los residuos que se generan cuyo valor-p es 0.9274, con lo cual mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial, lo que quiere decir, que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SAR, la dependencia de los residuos desaparece.

A continuación, se presenta la ecuación del modelo de retardo espacial para la tasa de robo a domicilios por zona urbana en la ciudad de Cuenca año 2019:

$$Tasa\ robo\ a\ domicilios = 0.212 * lag.Tasa_1 + 3.151156 - 0.02793 * BD_TIEpc$$

Esto quiere decir que, por cada unidad de disminución en la tasa de inmigrantes extranjeros presente en una determinada zona de ciudad de Cuenca, la tasa de robo a domicilios aumenta en tres unidades.

Robo a unidades económicas: el valor de la dependencia espacial, en este caso, es de 0.08, un valor bastante bajo, sumado a un valor-p por encima de 0.05; precisamente 0.592, lo cual indica que no existe suficiente evidencia estadística para revelar una autocorrelación espacial global entre el robo a unidades económicas y las variables sociodemográficas. Ahora bien, existe una fuerte autocorrelación espacial entre el promedio de personas por hogar y la tasa de robo a unidades económicas, siendo esta una relación inicial lógica. Esta relación se puede constatar mediante el índice de retardo espacial, el dato es, por cada 6 personas que habiten el hogar la tasa de robo a unidades económicas baja un punto.

Tabla 5-4: Resultado del modelo de retardo espacial de unidades económicas robadas año 2019.

```
Call:lagsarlm(formula = frm, data = shpf, listw = seaw)

Residuals:
    Min       1Q   Median       3Q      Max
-6.287254 -0.929781 -0.040055  0.805714 13.206368

Type: lag
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) 29.68162663 9.02307332 3.2895 0.0010036
BD_TA         -0.07406248  0.46276410 -0.1600 0.8728466
BD_Tasa_Na    -0.06671639  0.05649153 -1.1810 0.2376035
BD_Tasa_Dt    0.02078047  0.10049570  0.2068 0.8361819
BD_Inmi_Ex   0.02177275 0.00899595 2.4203 0.0155084
BD_PMI        0.00012979  0.00012110  1.0718 0.2838168
BD_NEI        0.01730613  0.01545916  1.1195 0.2629378
BD_TIEpc      0.00037853  0.02421131  0.0156 0.9875260
BD_PPH       -6.47055461 1.81840887 -3.5584 0.0003732
BD_PET        -0.00139944  0.00190818 -0.7334 0.4633223
BD_PCes       0.02963580  0.03988304  0.7431 0.4574406
BD_POdh       -0.00385119  0.00811059 -0.4748 0.6349046
BD_PEI        0.00132986  0.00419332  0.3171 0.7511391
BD_HTV        -0.00319001  0.00291685 -1.0936 0.2741087

Rho: 0.083544, LR test value: 0.19831, p-value: 0.65609
Asymptotic standard error: 0.15576
Z-value: 0.53636, p-value: 0.59171
Wald statistic: 0.28768, p-value: 0.59171

Log likelihood: -198.2118 for lag model
ML residual variance (sigma squared): 7.8063, (sigma: 2.794)
Number of observations: 81
Number of parameters estimated: 16
AIC: 428.42, (AIC for lm: 426.62)
LM test for residual autocorrelation
test value: 0.063742, p-value: 0.80068
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Ahora una vez que se han identificado las variables sociodemográficas que tienen una correlación espacial significativa en el modelo espacial de SAR y explican la alta incidencia de

unidades económicas que fueron robadas; el modelo de retardo espacial ajustado para el número de robo a unidades económicas a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$\text{Robo unidades económicas} = \rho * \text{lag. Robo}_{u_i} + \beta_0 + \beta_1 * \text{BD_Inmi_Ex}_i + \beta_1 * \text{BD_PPH}_i + \varepsilon_i$$

donde,

$$\varepsilon_i \sim N(0, \sigma^2)$$

$$\text{lag. Robo}_{u_i} = \sum_j^n W_{ij} * \text{Robo}_{u_i}$$

Identificadas las variables sociodemográficas con una correlación espacial significativa en la alta incidencia de las unidades económicas que fueron robadas; entonces el modelo de retardo espacial que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan en la Tabla 6-4.

Tabla 6-4: Resultado del modelo de retardo espacial ajustado de unidades económicas robadas año 2019.

```
Call: lagsarlm(formula = fr, data = bdta, listw = seaw)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-5.76913	-1.27105	-0.29370	0.99945	14.23632

Type: lag

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	22.0285561	6.4726711	3.4033	0.0006657
BD_Inmi_Ex	0.0153480	0.0076865	1.9967	0.0458527
BD_PPH	-5.8098882	1.6061438	-3.6173	0.0002977

Rho: 0.093113, LR test value: 0.28303, p-value: 0.59472

Asymptotic standard error: 0.15537

Z-value: 0.59929, p-value: 0.54898

Wald statistic: 0.35914, p-value: 0.54898

Log likelihood: -201.9361 for lag model

ML residual variance (sigma squared): 8.5555, (sigma: 2.925)

Number of observations: 81

Number of parameters estimated: 5

AIC: 413.87, (AIC for lm: 412.16)

LM test for residual autocorrelation

test value: 0.051398, p-value: 0.82065

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Mediante la Tabla 6-4 se pueden determinar cinco parámetros estimados que son los que se sustentan en el capítulo 2, los mismos que se necesitan para obtener el modelo y son: el parámetro ρ mismo que mide y determina el nivel de relación autorregresiva espacial que es estimado por Rho e igual a 0.093113, coeficientes β_q dados por $\beta_0 = 22.03$, $\beta_1 = 0.0153480$,

$\beta_2 = -5.81$ y por último la desviación típica de los residuos generados por este método, es decir $\sigma = 2.925$. Además, se obtiene el valor-p de un test realizado a los residuos que se generan, cuyo valor-p es 0.821, mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial, lo que quiere decir, que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SAR, la dependencia de los residuos desaparece.

A continuación, se presenta la ecuación del modelo de retardo espacial para el robo a unidades económicas por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Robo unidades económicas} = 0.093 * \text{lag. Robo}_u + 22.03 + 0.015348 * \text{BD_Inmi_Ex} - 5.81 * \text{BD_PPH}$$

Esto quiere decir que, por cada persona más en promedio; que viva en el hogar el número de robo a unidades económicas disminuye en cinco unidades.

Robo a vehículos: En este caso el valor de dependencia espacial (rho) es 0.06 (bastante bajo), sumado a un valor estadísticamente no significativo, pues su valor-p es de 0.7. Ahora bien, al observar el comportamiento de las variables de forma individual dentro del modelo se tiene una correlación estadísticamente significativa entre la cantidad de robo a vehículos y la tasa de desplazamiento de personas de forma temporal dentro y fuera de la ciudad, otras variables con la que existe una correlación estadísticamente significativa entre la cantidad de robo a vehículos, son las personas en edad de trabajar y la población de inmigrantes internacionales.

Tabla 7-4: Resultado del modelo de retardo espacial de vehículos robados año 2019.

```
Call: lagsarlm(formula = frm, data = shpf, listw = seaw)

Residuals:
    Min       1Q   Median       3Q      Max
-7.3464 -2.0126 -0.2531  1.6653 12.5949

Type: lag
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -17.45387112  9.71771696 -1.7961 0.072481
BD_TA        -0.59932701  0.52377886 -1.1442 0.252525
BD_Tasa_Na   0.03933933  0.06415859  0.6132 0.539772
BD_Tasa_Dt  0.35989172  0.11409278  3.1544 0.001608
BD_Inmi_Ex  -0.00038632  0.01019621 -0.0379 0.969777
BD_PMI     0.00037975  0.00014636  2.5946 0.009471
BD_NEI      -0.02568159  0.01750638 -1.4670 0.142380
BD_TIEpc    0.04110807  0.02736079  1.5024 0.132982
BD_PPH      0.88491004  1.90551545  0.4644 0.642365
BD_PET     0.00535277  0.00216210  2.4757 0.013297
BD_PCes     -0.08128442  0.04536410 -1.7918 0.073161
BD_POdh     0.01421203  0.00920903  1.5433 0.122765
BD_PEI      -0.00609185  0.00475397 -1.2814 0.200045
BD_HTV      0.00504802  0.00331392  1.5233 0.127690

Rho: 0.067795, LR test value: 0.14433, p-value: 0.70401
Asymptotic standard error: 0.15511
Z-value: 0.43708, p-value: 0.66205
Wald statistic: 0.19104, p-value: 0.66205
```

Log likelihood: -208.3122 for lag model
 ML residual variance (sigma squared): 10.022, (sigma: 3.1657)
 Number of observations: 81
 Number of parameters estimated: 16
 AIC: 448.62, (AIC for lm: 446.77)
 LM test for residual autocorrelation
 test value: 0.068388, p-value: 0.7937

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Identificadas las variables sociodemográficas con una correlación espacial significativa en el modelo espacial de SAR y que explican la alta incidencia de vehículos que fueron robadas, el modelo de retardo espacial ajustado para el número de robo a vehículos a nivel de zonas urbanas en la ciudad de Cuenca año 2019, queda de la siguiente manera:

$$Robo\ a\ veh\acute{u}culos = \rho * lag.Robo_ve_i + \beta_0 + \beta_1 * BD_Tasa_Dt_i + \beta_1 * BD_PMI_i + \beta_1 * BD_PET_i + \varepsilon_i$$

donde,

$$\varepsilon_i \sim N(0, \sigma^2)$$

$$lag.Robo_ve_i = \sum_j^n W_{ij} * Robo_ve_i$$

Reconocidas las variables sociodemográficas que tienen una correlación espacial significativa en la alta incidencia de los vehículos que fueron robados; entonces el modelo de retardo espacial que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 8-4.

Tabla 8-4: Resultado del modelo de retardo espacial ajustado de vehículos robados año 2019.

Call: lagsarlm(formula = fr, data = bdta, listw = seaw)

Residuals:

	Min	1Q	Median	3Q	Max
	-7.66265	-2.36727	-0.22613	1.74724	13.49747

Type: lag

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-9.87623566	2.92500189	-3.3765	0.0007342
BD_Tasa_Dt	0.29687310	0.10614357	2.7969	0.0051595
BD_PMI	0.00041507	0.00014502	2.8621	0.0042086
BD_PET	0.00206385	0.00047312	4.3622	0.00001287

Rho: 0.24841, LR test value: 2.2009, p-value: 0.13793

Asymptotic standard error: 0.14623

Z-value: 1.6987, p-value: 0.089367

Wald statistic: 2.8857, p-value: 0.089367

Log likelihood: -218.8785 for lag model

ML residual variance (sigma squared): 12.863, (sigma: 3.5865)

Number of observations: 81

Number of parameters estimated: 6

AIC: 449.76, (AIC for lm: 449.96)

LM test for residual autocorrelation

test value: 0.97737, p-value: 0.32285

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

En la Tabla 8-4., se muestran los cinco parámetros estimados mismos que se sustentan en el capítulo 2, los cuales se necesitan para obtener el modelo y son: el parámetro ρ el que mide y determina el nivel de relación autorregresiva espacial y que es estimado por Rho igual a 0.2484, los coeficientes β_q dados por $\beta_0 = -9.876$, $\beta_1 = 0.29687$, $\beta_2 = 0.000415$, $\beta_3 = 0.00206385$; y por último la desviación típica de los residuos generados por este método, es decir $\sigma = 3.5865$. Además, se obtiene el valor-p de un test realizado a los residuos que generan, cuyo valor-p es 0.32285, con lo cual mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial, lo que indica, que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SAR, la dependencia de los residuos desaparece.

A continuación, se presenta la ecuación del modelo de retardo espacial para el robo a vehículos por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Robo a vehículos} = 0.2484 * \text{lag.Robo_ve} - 9.876 + 0.29687 * \text{BD_Tasa_Dt} + \\ 0.000415 * \text{BD_PMI} + 0.00206385 * \text{BD_PET}$$

Esto quiere decir que, por cada unidad que incremente la tasa de desplazamiento temporal por trabajo o estudio, las personas inmigrantes internacionales y la población en edad de trabajo, el número de robo a vehículos disminuye en nueve unidades.

4.2.2 *Resultados del modelo de error espacial*

El modelo de error espacial incorpora dependencia espacial en los errores, si se está produciendo un proceso de error espacial y si se ajusta un modelo OLS, los coeficientes serán imparciales pero ineficientes; es decir, el tamaño y el signo del coeficiente son asintóticamente correctos pero sus errores estándar están subestimados.

Robo a personas: El modelo dio con un valor estadísticamente significativo del 99%, su valor p registra 0.01, siendo el valor del intercepto 15. Sumado a esto, se tiene suficiente evidencia estadística para afirmar que las variables inmigración extranjera y el promedio de personas por hogar tienen relación con la tasa de robo a personas, desde una mirada del error espacial. Para ver esto en detalle, se amplía en la sección del análisis de Moran bivariado. En adición, el valor λ es positivo y estadísticamente significativo (valor-p de 0.017), lo que indica la necesidad de controlar la autocorrelación espacial en el error.

Tabla 9-4: Resultado del modelo de error espacial de la tasa de robo a personas año 2019.

Call: errorsarlm(formula = frm, data = shpf, listw = seaw)


```

Residuals:
      Min       1Q   Median       3Q      Max
-3.80432 -1.01110 -0.15641  0.48361  8.73252

Type: error
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) 14.490891896 6.037568509 2.4001 0.01639
BD_TA        -0.122069654  0.298511052 -0.4089  0.68259
BD_Tasa_Na   -0.009847743  0.036660458 -0.2686  0.78822
BD_Tasa_Dt   -0.013161037  0.067464469 -0.1951  0.84533
BD_Inmi_Ex  0.024355920 0.006167987 3.9488 0.00007856
BD_PMI       -0.000040959  0.000104130 -0.3933  0.69406
BD_NEI       0.005472099  0.009876375  0.5541  0.57954
BD_TIEpc     -0.004471756  0.016049044 -0.2786  0.78053
BD_PPH     -2.554554306 1.269878420 -2.0117 0.04426
BD_PET       -0.000717479  0.001152261 -0.6227  0.53350
BD_PCes    0.052568170 0.025947182 2.0260 0.04277
BD_POdh      -0.007122422  0.005637050 -1.2635  0.20641
BD_PEI       -0.001046987  0.002616686 -0.4001  0.68907
BD_HTV       -0.001179408  0.001849029 -0.6379  0.52357

Lambda: 0.48387, LR test value: 5.6661, p-value: 0.017296
Asymptotic standard error: 0.12922
Z-value: 3.7446, p-value: 0.00018067
Wald statistic: 14.022, p-value: 0.00018067

Log likelihood: -162.2012 for error model
ML residual variance (sigma squared): 3.0528, (sigma: 1.7472)
Number of observations: 81
Number of parameters estimated: 16
AIC: 356.4, (AIC for lm: 360.07)

```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Ahora una vez que se han identificado las variables sociodemográficas con una correlación espacial significativa en el modelo espacial de SEM y que explican la alta incidencia de personas que fueron robadas; el modelo de retardo espacial ajustado para la tasa robo a personas a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$Tasa\ robo\ a\ personas = \beta_0 + \beta_1 * BD_Inmi_Ex_i + \beta_2 * BD_PPH_i + \beta_3 * BD_PCes_i + \lambda * lag.resid_i + \mu_i$$

donde,

$$\mu_i \sim N(0, \sigma^2)$$

$$lag.resid_i = \sum_j^n W_{ij} * \varepsilon_j$$

Identificadas las variables sociodemográficas que tienen una correlación espacial significativa en la alta incidencia de la tasa de robo a personas; entonces el modelo de error espacial que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 10-4.

Tabla 10-4: Resultado del modelo de error espacial ajustado de la tasa de robo a personas año 2019.

Call: errorsarlm(formula = fr, data = bdta, listw = seaw)

Residuals:

	Min	1Q	Median	3Q	Max
	-3.47588	-0.98886	-0.33616	0.47391	9.22330

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	15.1152890	4.6571151	3.2456	0.001172
BD_Inmi_Ex	0.0151511	0.0058346	2.5968	0.009411
BD_PPH	-3.8778801	1.2157888	-3.1896	0.001425
BD_PCes	-0.0063523	0.0223501	-0.2842	0.776243

Lambda: 0.48541, LR test value: 8.1894, p-value: 0.0042135

Asymptotic standard error: 0.129

Z-value: 3.763, p-value: 0.00016788

Wald statistic: 14.16, p-value: 0.00016788

Log likelihood: -170.543 for error model

ML residual variance (sigma squared): 3.7497, (sigma: 1.9364)

Number of observations: 81

Number of parameters estimated: 6

AIC: 353.09, (AIC for lm: 359.28)

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se observa en la Tabla 10-4, que la variable población cesante (BD_PCes) tiene un valor-p mayor a 0,05 (5%), por lo tanto esta no influye en la tasa de robo a personas lo que quiere decir que no existe una autocorrelación significativa en términos espaciales, por lo tanto a continuación se presenta nuevamente el modelo ajustado.

Tabla 11-4: Resultado del modelo de error espacial ajustado de la tasa de robo a personas año 2019.

Call: errorsarlm(formula = fr, data = bdta, listw = seaw)

Residuals:

	Min	1Q	Median	3Q	Max
	-3.40923	-0.96943	-0.37214	0.53374	9.28372

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	15.3819635	4.5684054	3.3670	0.0007598
BD_Inmi_Ex	0.0145774	0.0054763	2.6619	0.0077701
BD_PPH	-3.9743645	1.1698621	-3.3973	0.0006806

Lambda: 0.48397, LR test value: 8.1171, p-value: 0.0043849

Asymptotic standard error: 0.1292

Z-value: 3.7458, p-value: 0.00017979

Wald statistic: 14.031, p-value: 0.00017979

Log likelihood: -170.5834 for error model

ML residual variance (sigma squared): 3.7547, (sigma: 1.9377)

Number of observations: 81

Number of parameters estimated: 5

AIC: 351.17, (AIC for lm: 357.28)

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Como se observa en la Tabla 11-4, se muestra los mismos cinco parámetros estimados, donde el parámetro autorregresivo λ tiene un valor de 0.48397; los coeficientes β_q son: $\beta_0 = 15.38$, $\beta_1 = 0.014577$, $\beta_2 = -3.9743645$, y por último la desviación típica de los residuos generados por este método μ_i , es $\sigma = 1.9377$. Mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial, lo que quiere decir, que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SEM, la dependencia de los residuos desaparece.

A continuación se presenta la ecuación del modelo de error espacial para la tasa de robo a personas por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Tasa robo a personas} = 15.38 + 0.014577 * BD_Inmi_Ex - 3.9743645 * BD_PPH + 0.48541 * lag.resid$$

Esto quiere decir que, por cada tres unidades que disminuya el promedio de personas por hogar y aumente en una unidad los inmigrantes extranjeros, el número de robo a personas aumenta en 15 unidades al año.

Robo a domicilios: El modelo de error espacial generó un valor de intercepto en el modelo espacial de 5.6, sin embargo, este valor no es estadísticamente significativo, ya que, su valor-p es 0.44; por otro lado, una variable estadísticamente significativa dentro del modelo, es la población económicamente inactiva, con lo cual se puede decir que, dentro del modelo de error espacial, esta variable tiene relación con la tasa de robos a domicilios.

Tabla 12-4: Resultado modelo de error espacial de la tasa de robo a domicilios año 2019.

Call: errorsarlm(formula = frm, data = shpf, listw = seaw)

```
Residuals:
    Min       1Q   Median       3Q      Max
-3.3847 -1.6250 -0.3334  0.9985  6.1024

Type: error
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.60794096  7.26658460  0.7717  0.44027
BD_TA        -0.11954362  0.35855260 -0.3334  0.73883
BD_Tasa_Na   0.00083799  0.04405368  0.0190  0.98482
BD_Tasa_Dt  -0.07572145  0.08110230 -0.9337  0.35048
BD_Inmi_Ex   0.00606815  0.00741833  0.8180  0.41336
BD_PMI       -0.00016495  0.00012602 -1.3089  0.19057
BD_NEI       0.00562520  0.01186051  0.4743  0.63530
BD_TIEpc    -0.03202193  0.01928915 -1.6601  0.09689
BD_PPH       0.16636606  1.53041278  0.1087  0.91344
BD_PET       0.00163679  0.00138254  1.1839  0.23645
BD_PCes     0.02230247  0.03116624  0.7156  0.47424
BD_POdh     0.00816442  0.00678225  1.2038  0.22867
BD_PEI      -0.00632690  0.00314112 -2.0142  0.04399
BD_HTV      0.00071015  0.00222001  0.3199  0.74906

Lambda: 0.49258, LR test value: 3.824, p-value: 0.050522
Asymptotic standard error: 0.12797
Z-value: 3.8493, p-value: 0.00011846
Wald statistic: 14.817, p-value: 0.00011846
```

Log likelihood: -177.1089 for error model
 ML residual variance (sigma squared): 4.402, (sigma: 2.0981)
 Number of observations: 81
 Number of parameters estimated: 16
 AIC: 386.22, (AIC for lm: 388.04)

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Identificadas las variables sociodemográficas con una correlación espacial significativa en el modelo espacial de SEM y que explican la alta incidencia de domicilios que fueron robados; el modelo de retardo espacial ajustado para la tasa robo a domicilios a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$Tasa\ robo\ a\ domicilios = \beta_0 + \beta_1 * BD_PEI_i + \lambda * lag.resid_i + \mu_i$$

donde,

$$\mu_i \sim N(0, \sigma^2)$$

$$lag.resid_i = \sum_j^n W_{ij} * \varepsilon_j$$

Conocidas las variables sociodemográficas con una correlación espacial significativa en la alta incidencia de la tasa de robo a domicilios entonces el modelo de error espacial que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 13-4.

Tabla 13-4: Resultado modelo de error espacial ajustado de la tasa de robo a domicilios año 2019.

```
Call:errorsarlm(formula = fr, data = bdta, listw = seaw)

Residuals:
    Min       1Q   Median       3Q      Max
-3.49959 -1.76229 -0.39624  1.07798  6.92803

Type: error
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  3.86365568  0.98095628  3.9387 0.00008194
BD_PEI       -0.00059582  0.00066598 -0.8947  0.371

Lambda: 0.20477, LR test value: 1.4836, p-value: 0.22322
Asymptotic standard error: 0.16151
Z-value: 1.2678, p-value: 0.20485
Wald statistic: 1.6074, p-value: 0.20485

Log likelihood: -182.1096 for error model
ML residual variance (sigma squared): 5.2096, (sigma: 2.2825)
Number of observations: 81
Number of parameters estimated: 4
AIC: 372.22, (AIC for lm: 371.7)
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

En la Tabla 13-4, se muestra los mismos cinco parámetros estimados, donde el parámetro autorregresivo λ tiene un valor de 0.20477, los coeficientes β_q son: $\beta_0 = 3.864$, $\beta_1 = -0.0005958$, y por último la desviación típica de los residuos generados por este método μ_i , es decir $\sigma = 2.2825$. Mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial, lo que quiere decir, que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SEM, la dependencia de los residuos desaparece.

A continuación, se presenta la ecuación del modelo de error espacial para la tasa de robo a domicilios por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Tasa robo a domicilios} = 3.864 + 0.0005958 * BD_PEI + 0.20477 * lag.resid$$

Esto quiere decir que, por cada unidad que aumenta en el número de personas desempleadas, el número de robo a domicilios aumenta en cuatro unidades al año

Robo a unidades económicas: El modelo referente a la variable de robo a unidades económicas generó un valor de intercepto de 31.16, el cual se presenta como estadísticamente significativo, pues su valor-p fue cercano a 0, precisamente 0.00032. Ahora bien, las variables que dentro del modelo fueron estadísticamente significativas son el número de inmigrantes extranjeros y el promedio de personas por hogar, es decir, que estas variables aportan importancia al modelo de error espacial.

Tabla 14-4: Resultado del modelo de error espacial del número de unidades económicas robadas año 2019.

```
Call:errorsarlm(formula = frm, data = shpf, listw = seaw)

Residuals:
      Min       1Q   Median       3Q      Max
-6.414206 -1.003610 -0.028518  0.884052 13.194671

Type: error
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) 31.16457647  8.66856741  3.5951 0.0003242
BD_TA       -0.08318706  0.46507229 -0.1789 0.8580405
BD_Tasa_Na  -0.06838527  0.05673404 -1.2054 0.2280621
BD_Tasa_Dt   0.02186679  0.10135944  0.2157 0.8291942
BD_Inmi_Ex   0.02249404  0.00904120  2.4879 0.0128482
BD_PMI       0.00013496  0.00012360  1.0919 0.2748842
BD_NEI       0.01713049  0.01552967  1.1031 0.2699918
BD_TIEpc     0.00175601  0.02434934  0.0721 0.9425085
BD_PPH      -6.83524600  1.70997118 -3.9973 0.00006407
BD_PET      -0.00140109  0.00190580 -0.7352 0.4622339
BD_PCes      0.03160166  0.04004764  0.7891 0.4300526
BD_POdh     -0.00391103  0.00820304 -0.4768 0.6335198
BD_PEI       0.00135227  0.00420313  0.3217 0.7476574
BD_HTV      -0.00310698  0.00293037 -1.0603 0.2890223

Lambda: 0.048963, LR test value: 0.030965, p-value: 0.86032
Asymptotic standard error: 0.17367
Z-value: 0.28193, p-value: 0.778
Wald statistic: 0.079485, p-value: 0.778
```

Log likelihood: -198.2954 for error model
 ML residual variance (sigma squared): 7.8292, (sigma: 2.7981)
 Number of observations: 81
 Number of parameters estimated: 16
 AIC: 428.59, (AIC for lm: 426.62)

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Identificadas las variables sociodemográficas que tienen una correlación espacial significativa en el modelo espacial de SEM y que explican la alta incidencia de unidades económicas que fueron robadas; entonces el modelo de retardo espacial ajustado para este delito a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$\text{Robo a unidades económicas} = \beta_0 + \beta_1 * BD_Inmi_Ex_i + \beta_2 * BD_PPH_i + \lambda * lag.resid_i + \mu_i$$

donde,

$$\mu_i \sim N(0, \sigma^2)$$

$$lag.resid_i = \sum_j^n W_{ij} * \varepsilon_j$$

Descritas las variables sociodemográficas que tienen una correlación espacial significativa en la alta incidencia del robo a unidades económicas entonces el modelo de error espacial que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 15-4.

Tabla 15-4: Resultado del modelo de error espacial ajustado del número de unidades económicas robadas año 2019.

```
Call:errorsarlm(formula = fr, data = bdta, listw = seaw)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-5.70867	-1.31506	-0.29535	0.98207	14.24660

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	23.6054952	5.9615062	3.9597	0.00007506
BD_Inmi_Ex	0.0163220	0.0076632	2.1299	0.03318
BD_PPH	-6.1977388	1.4868056	-4.1685	0.00003066

Lambda: 0.05733, LR test value: 0.09152, p-value: 0.76225

Asymptotic standard error: 0.17311

Z-value: 0.33117, p-value: 0.74052

Wald statistic: 0.10967, p-value: 0.74052

Log likelihood: -202.0319 for error model

ML residual variance (sigma squared): 8.5844, (sigma: 2.9299)

Number of observations: 81

Number of parameters estimated: 5

AIC: 414.06, (AIC for lm: 412.16)

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Como podemos observar en la Tabla 15-4., se muestran cinco parámetros estimados, donde el parámetro autorregresivo λ es 0.057, los coeficientes β_q son: $\beta_0 = 23.605495$, $\beta_1 = 0.016$, $\beta_2 = -6.19774$, y por último la desviación típica de los residuos generados por este método μ_i , es decir $\sigma = 2.9299$. Mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial, lo que quiere decir, que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SEM, la dependencia de los residuos desaparece.

A continuación, se presenta la ecuación del modelo de error espacial para el robo a unidades económicas por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Robo a unidades económicas} = 23.61 + 0.016 * \text{BD_Inmi_Ex} - 6.19774 * \text{BD_PPH} + 0.057 * \text{lag.resid}$$

Esto quiere decir que, por cada seis unidades que disminuya en el promedio de personas por hogar y aumente en una unidad los inmigrantes extranjeros, el número de robo a unidades económicas aumenta en 23 unidades al año.

Robo a vehículos: Este modelo presenta un valor de intercepto negativo de -17.1; sin embargo, su valor-p es superior a 0.05, lo que indica que no existe evidencia estadística para afirmar una relación entre el valor del intercepto con los datos de robo a vehículos. Ahora bien, la tasa de desplazamiento temporal presenta suficiente evidencia estadística para indicar un comportamiento importante dentro del modelo, con un valor de intercepto de 0.363. Por otro lado, la población en edad de trabajo resulta ser una variable importante dentro del modelo de error espacial, pues su valor-p es por cercano a 0, propiamente 0.0137.

Tabla 16-4: Resultado modelos de error espacial del número de vehículos robados año 2019.

Call: errorsarlm(formula = frm, data = shpf, listw = seaw)

```
Residuals:
      Min       1Q   Median       3Q      Max
-7.27626 -2.02350 -0.21501  1.61141 12.53771

Type: error
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -17.13085531  9.79849592 -1.7483 0.080410
BD_TA        -0.60187005  0.52630950 -1.1436 0.252803
BD_Tasa_Na   0.04134838  0.06421295  0.6439 0.519623
BD_Tasa_Dt  0.36261751  0.11464185  3.1630 0.001561
BD_Inmi_Ex  0.00011263  0.01022087  0.0110 0.991208
BD_PMI      0.00040247  0.00013938  2.8876 0.003882
BD_NEI      -0.02549435  0.01757757 -1.4504 0.146950
BD_TIEpc    0.04137760  0.02754360  1.5023 0.133030
BD_PPH      0.81783291  1.93052081  0.4236 0.671833
BD_PET      0.00532257  0.00215950  2.4647 0.013712
BD_PCes     -0.08232954  0.04530760 -1.8171 0.069198
BD_POdh     0.01401750  0.00927405  1.5115 0.130667
BD_PEI      -0.00599949  0.00475960 -1.2605 0.207488
BD_HTV      0.00512079  0.00331704  1.5438 0.122641

Lambda: 0.040419, LR test value: 0.02234, p-value: 0.88119
```

Asymptotic standard error: 0.17422
 Z-value: 0.232, p-value: 0.81654
 Wald statistic: 0.053822, p-value: 0.81654

Log likelihood: -208.3732 for error model
 ML residual variance (sigma squared): 10.043, (sigma: 3.169)
 Number of observations: 81
 Number of parameters estimated: 16
 AIC: 448.75, (AIC for lm: 446.77)

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Establecidas las variables sociodemográficas que tienen una correlación espacial significativa en el modelo espacial de SEM y que explican la alta incidencia de vehículos que fueron robados, el modelo de retardo espacial ajustado a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$Robo\ a\ veh\iculos = \beta_0 + \beta_1 * BD_Tasa_Dt_i + \beta_2 * BD_PMI_i + \beta_3 * BD_PET_i + \lambda * lag.resid_i + \mu_i$$

donde,

$$\mu_i \sim N(0, \sigma^2)$$

$$lag.resid_i = \sum_j^n W_{ij} * \varepsilon_j$$

Una vez que se han identificado las variables sociodemográficas que tienen una correlación espacial significativa en la alta incidencia del robo a vehículos; entonces el modelo de error espacial que arroja el lenguaje estadística R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 17-4.

Tabla 17-4: Resultado modelos de error espacial ajustado del número de vehículos robados año 2019.

```
Call:errorsarlm(formula = fr, data = bdta, listw = seaw)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-7.55575 -2.05902 -0.14223  1.76096 13.01269
```

```
Type: error
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -8.93284543  3.00441737 -2.9732  0.002947
BD_Tasa_Dt   0.30361516  0.10805102  2.8099  0.004955
BD_PMI       0.00042987  0.00017560  2.4480  0.014366
BD_PET       0.00224915  0.00046194  4.8689 0.000001122
```

```
Lambda: 0.3293, LR test value: 3.3178, p-value: 0.068535
Asymptotic standard error: 0.14888
Z-value: 2.2119, p-value: 0.026976
Wald statistic: 4.8923, p-value: 0.026976
```

```
Log likelihood: -218.32 for error model
ML residual variance (sigma squared): 12.562, (sigma: 3.5444)
Number of observations: 81
```


Number of parameters estimated: 6
 AIC: 448.64, (AIC for lm: 449.96)
Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Eddy, 2022.

Como se observa en la Tabla 17-4., se muestran cinco parámetros estimados, donde el parámetro autorregresivo λ es 0.3293, los coeficientes β_q son: $\beta_0 = -8.933$, $\beta_1 = 0.3036$, $\beta_2 = 0.00042987$, $\beta_3 = 0.002249$ y por último la desviación típica de los residuos generados por este método μ_i , es decir $\sigma = 3.54$. Mediante este resultado no se puede rechazar la hipótesis nula de no autocorrelación espacial, lo que quiere decir, que los residuos se pueden considerar independientes y, por lo tanto, tras la aplicación del modelo SEM, la dependencia de los residuos desaparece.

A continuación, se presenta la ecuación del modelo de error espacial para robo a vehículos por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Robo a vehículos} = -0.8933 + 0.3036 * \text{BD_Tasa_Dt} + 0.00042987 * \text{BD_PMI} + 0.002249 * \text{BD_PET} + 0.3293 * \text{lag.resid}$$

Esto quiere decir que, por cada unidad que aumente la tasa de desplazamiento temporal por trabajo o estudio, las personas inmigrantes internacionales y la población en edad de trabajo, el número de robo a vehículos disminuirá en una unidad.

4.2.3 Resultados del modelo de Durbin

El modelo espacial de Durbin tiene una gran ventaja, pues permite identificar la relación que existe directa o indirectamente por el vecino de un vecino, es decir, los polígonos de las zonas de interés en el presente caso.

Robo a personas: El modelo espacial de Durbin genera un valor de intercepto de 14.19, con un valor-p por encima de lo estadísticamente significativo, propiamente 0.27; sin embargo, la variable de hogares que tienen vivienda propia resulta ser estadísticamente significativa dentro de este modelo espacial, con un nivel de significancia de 99%. Ahora bien, el valor de intercepto es -0.01, lo que indica que la variable bajo rezago de más de un vecino reduce en pequeña medida la tasa de robo a personas.

Tabla 18-4: Resultado del modelo espacial Durbin de la tasa de robo a personas año 2019.

Call: lagsarlm(formula = fl, data = shpf, listw = seaw, type = "mixed")

```
Residuals:
    Min       1Q   Median       3Q      Max
-3.53877 -1.03060 -0.19010  0.77391  8.27312

Type: mixed
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  14.194804420 12.923654744  1.0984  0.27205
BD_TA       -0.218368580  0.299320955 -0.7295  0.46567
```

BD_Tasa_Na	-0.016541470	0.039056818	-0.4235	0.67191
BD_Tasa_Dt	0.012551843	0.067108463	0.1870	0.85163
BD_Inmi_Ex	0.020980280	0.006418321	3.2688	0.00108
BD_PMI	-0.000085069	0.000169149	-0.5029	0.61502
BD_NEI	-0.000185197	0.009748904	-0.0190	0.98484
BD_TIEpc	-0.009991945	0.016391298	-0.6096	0.54213
BD_PPH	-0.404297049	1.503422867	-0.2689	0.78799
BD_PET	-0.000939911	0.001236528	-0.7601	0.44718
BD_PCes	0.044425665	0.025979642	1.7100	0.08726
BD_POdh	-0.006355511	0.005862098	-1.0842	0.27829
BD_PEI	-0.000397353	0.002748895	-0.1445	0.88507
BD_HTV	-0.002332945	0.001798132	-1.2974	0.19448
lag.BD_TA	-0.568049386	0.600451964	-0.9460	0.34413
lag.BD_Tasa_Na	0.056399311	0.091486724	0.6165	0.53758
lag.BD_Tasa_Dt	0.254434660	0.184220777	1.3811	0.16724
lag.BD_Inmi_Ex	-0.013032656	0.014626191	-0.8910	0.37290
lag.BD_PMI	0.000249672	0.000194235	1.2854	0.19865
lag.BD_NEI	-0.014782156	0.020643967	-0.7161	0.47396
lag.BD_TIEpc	-0.021104812	0.041321875	-0.5107	0.60953
lag.BD_PPH	-2.544179058	2.702130104	-0.9415	0.34643
lag.BD_PET	-0.000727568	0.002817735	-0.2582	0.79624
lag.BD_PCes	-0.074326929	0.054454467	-1.3649	0.17227
lag.BD_POdh	0.011570472	0.013537496	0.8547	0.39272
lag.BD_PEI	0.002342845	0.005400949	0.4338	0.66445
lag.BD_HTV	-0.010130488	0.004084637	-2.4801	0.01313

Rho: 0.24427, LR test value: 1.6, p-value: 0.2059
Asymptotic standard error: 0.15298
Z-value: 1.5968, p-value: 0.11032
Wald statistic: 2.5497, p-value: 0.11032

Log likelihood: -153.7496 for mixed model
ML residual variance (sigma squared): 2.5771, (sigma: 1.6053)
Number of observations: 81
Number of parameters estimated: 29
AIC: 365.5, (AIC for lm: 365.1)
LM test for residual autocorrelation
test value: 3.0829, p-value: 0.079118

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Establecidas las variables sociodemográficas que tienen una correlación espacial significativa en el modelo espacial de SDM y explican la alta incidencia de la tasa de personas que fueron robadas; entonces el modelo de espacial de Durbin ajustado para este delito a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$Tasa\ robo\ a\ personas = \rho * lag.BD_Tasa_{R_i} + \beta_0 + \beta_1 * BD_{HTV_i} + \gamma_1 * lag.BD_{HTV_i} + \varepsilon_i$$

donde,

$$lag.BD_{HTV_i} = \sum_j^n W_{ij} * BD_{HTV_j},$$

y las variables $lag.BD_{Tasa_{R_i}}$ y ε_i son las mismas que para el modelo SAR.

Una vez que se han determinado las variables sociodemográficas con una correlación espacial significativa en la alta incidencia de la tasa de robo a personas; entonces el modelo de SDM que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 19-4.

Tabla 19-4: Resultado del modelo espacial Durbin ajustado para la tasa de robo a personas año 2019.

```
Call: lagsarlm(formula = fr, data = bdta, listw = seaw, type = "mixed")

Residuals:
    Min       1Q   Median       3Q      Max
-2.20572 -1.13971 -0.55535  0.32439  8.96719

Type: mixed
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.5364052  1.3539843  1.1347  0.2565
BD_HTV       -0.0010630  0.0018923 -0.5618  0.5743
lag.BD_HTV   -0.0017238  0.0036366 -0.4740  0.6355

Rho: 0.75917, LR test value: 35.897, p-value: 0.0000000020798
Asymptotic standard error: 0.080061
      Z-value: 9.4824, p-value: < 0.000000000000000222
Wald statistic: 89.915, p-value: < 0.000000000000000222

Log likelihood: -180.98 for mixed model
ML residual variance (sigma squared): 4.3835, (sigma: 2.0937)
Number of observations: 81
Number of parameters estimated: 5
AIC: 371.96, (AIC for lm: 405.86)
LM test for residual autocorrelation
test value: 2.0942, p-value: 0.14786
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Como se observa en la Tabla 19-4., en este modelo se estimaron siete parámetros que son, el parámetro autorregresivo ρ cuyo valor es 0.759, los coeficientes β_q son: $\beta_0 = 1.5364$, $\beta_1 = -0.001063$, el coeficiente $\gamma_1 = -0.0017238$ y por último la desviación típica de los residuos generados por este método ε_i , es decir $\sigma = 2.09$. De igual manera que el modelo SAR se obtuvo un valor-p para un test de dependencia espacial para los residuos generados, este valor es, 0.14786, el cual indica que los nuevos residuos son independientes desapareciendo de esta manera la dependencia espacial.

A continuación, se presenta la ecuación del modelo espacial de Durbin para la tasa de robo a personas por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Tasa robo a personas} = 0.759 * \text{lag.BD_Tasa}_R_ + 1.5364 - 0.001063 * \text{BD_HTV} - 0.0017238 * \text{lag.BD_HTV}_i$$

Esto quiere decir que, por cada unidad que disminuye el número de hogares que tienen una vivienda propia y totalmente pagada, la tasa de robo a personas aumenta en dos unidades al año.

Robo a domicilios: El modelo espacial de Durbin para el caso de relacionar las variables analizadas con el robo a domicilios genera un valor de intercepto de 23.82, con un valor-p por debajo de 0.05, lo que da suficiente evidencia estadística para afirmar que existe una relación de las variables rezagadas con el robo a domicilios de manera conjunta, ahora bien, evaluando las variables de forma independiente, se observa la existencia de suficiente evidencia estadística

para identificar una relación rezagada para la variables de población inmigrante a escala internacional, con una confiabilidad de más de 99%.

Por otro lado, la variable de población cesante también resulta importante dentro del modelo, entre tanto, existe suficiente evidencia estadística con más de 99% de confiabilidad. La variable de población económicamente inactiva resulta ser de igual manera importante en la relación del error espacial con variable rezagada, esto con una confiabilidad de un poco más del 97%. Por último, la variable de hogares con vivienda propia y totalmente pagada presenta una fuerte relación con la variable de robos a domicilios, con una confiabilidad de más de 98%; quizás esta resulta ser la relación más clara existente, pues ambas variables tienen relación con los hogares (robo a hogares y vivienda propia). Los resultados obtenidos, se presentan a continuación:

Tabla 20-4: Resultado del modelo espacial Durbin de la tasa de robo a domicilios año 2019.

Call: lagsarlm(formula = frm, data = shpf, listw = seaw, type = "mixed")

```

Residuals:
    Min       1Q   Median       3Q      Max
-3.659890 -1.115385  0.032098  0.893619  4.114603

Type: mixed
Coefficients: (asymptotic standard errors)
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  23.82752803 13.84531143  1.7210  0.08525
BD_TA        -0.16157434  0.32492976 -0.4973  0.61901
BD_Tasa_Na   -0.06049534  0.04235037 -1.4284  0.15316
BD_Tasa_Dt   -0.02621789  0.07280401 -0.3601  0.71876
BD_Inmi_Ex   0.00896471  0.00694011  1.2917  0.19645
BD_PMI       -0.00019640  0.00018798 -1.0448  0.29612
BD_NEI       0.00053360  0.01058482  0.0504  0.95979
BD_TIEpc     -0.03090089  0.01772345 -1.7435  0.08125
BD_PPH       -0.30612209  1.63113033 -0.1877  0.85113
BD_PET       0.00062840  0.00136895  0.4590  0.64621
BD_PCes     -0.01864918  0.02877707 -0.6481  0.51695
BD_POdh     0.01135602  0.00635979  1.7856  0.07416
BD_PEI      -0.00407723  0.00305175 -1.3360  0.18154
BD_HTV      -0.00036933  0.00194438 -0.1899  0.84935
lag.BD_TA    0.39855509  0.64635329  0.6166  0.53748
lag.BD_Tasa_Na -0.08059909  0.10022272 -0.8042  0.42128
lag.BD_Tasa_Dt 0.23205899  0.19819629  1.1709  0.24166
lag.BD_Inmi_Ex -0.01209941  0.01563095 -0.7741  0.43889
lag.BD_PMI    0.00083668  0.00021087  3.9678  0.00007254
lag.BD_NEI    0.02380843  0.02230806  1.0673  0.28586
lag.BD_TIEpc  0.01193092  0.04526715  0.2636  0.79211
lag.BD_PPH   -2.57437553  2.86729691 -0.8978  0.36927
lag.BD_PET   -0.00567421  0.00305899 -1.8549  0.06361
lag.BD_PCes  -0.18682659  0.05994796 -3.1165  0.00183
lag.BD_POdh  -0.00035130  0.01474301 -0.0238  0.98099
lag.BD_PEI   0.01321828  0.00586181  2.2550  0.02413
lag.BD_HTV   -0.00903208  0.00434058 -2.0808  0.03745

Rho: 0.14921, LR test value: 0.68138, p-value: 0.40911
Asymptotic standard error: 0.15908
    Z-value: 0.93794, p-value: 0.34827
Wald statistic: 0.87974, p-value: 0.34827

Log likelihood: -160.0836 for mixed model
ML residual variance (sigma squared): 3.036, (sigma: 1.7424)
Number of observations: 81
Number of parameters estimated: 29

```

AIC: 378.17, (AIC for lm: 376.85)
 LM test for residual autocorrelation
 test value: 2.4673, p-value: 0.11624

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Descritas las variables sociodemográficas que tienen una correlación espacial significativa en el modelo espacial de SDM y que explican la alta incidencia de domicilios que fueron robados; entonces el modelo de espacial de Durbin ajustado para este delito a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$Tasa\ robo\ a\ domicilios = \rho * lag.BD_Tasa_1_i + \beta_0 - \beta_1 * BD_PMI_i + \beta_2 * BD_PCes_i - \beta_3 * BD_PEI_i + \beta_4 * BD_HTV_i + \gamma_1 * lag.BD_PMI_i - \gamma_2 * lag.BD_PCes_i + \gamma_3 * lag.BD_PEI_i - \gamma_4 * lag.BD_HTV_i + \varepsilon_i$$

donde,

$$lag.BD_PMI_i = \sum_j^n W_{ij} * BD_PMI_j,$$

$$lag.BD_PCes_i = \sum_j^n W_{ij} * BD_PCes_j,$$

$$lag.BD_PEI_i = \sum_j^n W_{ij} * BD_PEI_j,$$

$$lag.BD_HTV_i = \sum_j^n W_{ij} * BD_HTV_j.$$

y las variables $lag.BD_Tasa_1_i$ y ε_i son las mismas que para el modelo SAR.

Conocidas las variables sociodemográficas que tienen una correlación espacial significativa en la alta incidencia de la tasa de robo a domicilios; entonces el modelo de SDM que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 21-4.

Tabla 21-4: Resultado del modelo espacial Durbin ajustado de la tasa de robo a domicilios año 2019.

```
Call:lagsarlm(formula = fr, data = bdta, listw = seaw, type = "mixed")
```

Residuals:

	Min	1Q	Median	3Q	Max
	-4.33390	-1.37312	-0.24168	1.23314	6.47506

Type: mixed

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	7.168179692	2.639029040	2.7162	0.006603
BD_PMI	-0.000262853	0.000172274	-1.5258	0.127064
BD_PCes	0.015734795	0.029437403	0.5345	0.592984
BD_PEI	-0.001278009	0.000803807	-1.5899	0.111847
BD_HTV	0.000031481	0.001940937	0.0162	0.987059
lag.BD_PMI	0.000545821	0.000195411	2.7932	0.005219
lag.BD_PCes	-0.108268083	0.052567319	-2.0596	0.039436
lag.BD_PEI	0.001134614	0.001633944	0.6944	0.487430
lag.BD_HTV	-0.010898966	0.004518705	-2.4120	0.015867

Rho: 0.20495, LR test value: 1.4502, p-value: 0.2285
 Asymptotic standard error: 0.15797
 z-value: 1.2974, p-value: 0.1945
 Wald statistic: 1.6832, p-value: 0.1945

Log likelihood: -174.1482 for mixed model
 ML residual variance (sigma squared): 4.2799, (sigma: 2.0688)
 Number of observations: 81
 Number of parameters estimated: 11
 AIC: 370.3, (AIC for lm: 369.75)
 LM test for residual autocorrelation
 test value: 0.50318, p-value: 0.47811

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Como se observa en la Tabla 21-4., en este modelo se estimaron siete parámetros que son: el parámetro autorregresivo ρ cuyo valor es 0.20495; los coeficientes β_q son: $\beta_0 = 7.168$, $\beta_1 = -0.00026$, $\beta_2 = 0.0157$, $\beta_3 = -0.001278$, $\beta_4 = 0.00003148$ el coeficiente $\gamma_1 = 0.0005458$, $\gamma_2 = -0.108$, $\gamma_3 = 0.001134$, $\gamma_4 = -0.01089$ y por último la desviación típica de los residuos generados por este método ε_i , es decir $\sigma = 2.0688$. De igual manera que el modelo SAR, devuelve un valor-p para un test de dependencia espacial para los residuos generados; este es 0.47811, el cual indica que los nuevos residuos son independientes desapareciendo de esta manera la dependencia espacial.

A continuación, se presenta la ecuación del modelo espacial de Durbin para la tasa de robo a domicilios por zona urbana en la ciudad de Cuenca año 2019:

$$\begin{aligned} \text{Tasa robo a domicilios} = & 0.20495 * \text{lag.BD_Tasa}_1 + 7.168 - 0.00026 * \text{BD_PMI} + \\ & 0.0157 * \text{BD_PCes} + 0.00003148 * \text{BD_PEI} + 0.00003148 * \text{BD_HTV} - 0.0005458 * \\ & \text{lag.BD_PMI}_i - 0.108 * \text{lag.BD_PCes}_i - 0.001134 * \text{lag.BD_PEI}_i - 0.01089 * \\ & \text{lag.BD_HTV}_i \end{aligned}$$

Esto quiere decir que, por cada unidad que disminuye el número de la población inmigrante a escala internacional, la población cesante, población desempleada y los hogares que tienen vivienda propia y totalmente pagas, la tasa de robo a domicilios aumenta en siete unidades al año.

Robo a unidades económicas: El modelo espacial de Durbin para encontrar relaciones en las variables rezagadas respecto a la variable de robo a establecimientos económicos presenta un valor de intercepto de 20.5; sin embargo, este valor no es estadísticamente significativo, pues su valor-p es 0.29. Ahora bien, al revisar las variables de forma individual dentro del modelo, las tipo “lag” (rezagadas), se tiene que la variable de población que realiza su actividad económica dentro del hogar tiene una fuerte relación con el robo a establecimientos económicos. Desde el punto de vista lógico esto tiene mucho sentido, ya que se darían más robos en los hogares donde las personas trabajan allí mismo. Esta relación presenta suficiente evidencia estadística para ser aseverada, pues su valor-p es muy cercano a 0.

Los resultados obtenidos en el lenguaje R se presentan a continuación.

Tabla 22-4: Resultado del modelo espacial Durbin del número de unidades económicas robadas robo a personas año 2019.

```
Call: lagsarlm(formula = frm, data = shpf, listw = seaw, type = "mixed")
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-5.89284 -1.45900  0.10346  0.93627  8.65776
```

Type: mixed

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	20.548563705	19.735885498	1.0412	0.297793
BD_TA	-0.342875877	0.445643290	-0.7694	0.441659
BD_Tasa_Na	-0.117506764	0.058362372	-2.0134	0.044073
BD_Tasa_Dt	0.082147230	0.099707168	0.8239	0.410005
BD_Inmi_Ex	0.025959707	0.009606052	2.7024	0.006883
BD_PMI	-0.000011692	0.000252679	-0.0463	0.963094
BD_NEI	0.005252841	0.014515848	0.3619	0.717450
BD_TIEpc	0.017183344	0.024640384	0.6974	0.485574
BD_PPH	-3.993060615	2.238717487	-1.7836	0.074483
BD_PET	-0.002015557	0.001855249	-1.0864	0.277299
BD_PCes	0.036472271	0.038703159	0.9424	0.346009
BD_POdh	-0.007467865	0.008877632	-0.8412	0.400236
BD_PEI	0.003125826	0.004118090	0.7590	0.447824
BD_HTV	-0.003227780	0.002659298	-1.2138	0.224835
lag.BD_TA	-0.730885933	0.891682158	-0.8197	0.412404
lag.BD_Tasa_Na	0.186606103	0.137276989	1.3593	0.174039
lag.BD_Tasa_Dt	0.440431678	0.273026030	1.6131	0.106712
lag.BD_Inmi_Ex	-0.028006468	0.021638798	-1.2943	0.195572
lag.BD_PMI	0.000533774	0.000289927	1.8411	0.065613
lag.BD_NEI	-0.039520664	0.030599392	-1.2916	0.196513
lag.BD_TIEpc	-0.091519048	0.061576106	-1.4863	0.137206
lag.BD_PPH	-3.146718654	4.156412825	-0.7571	0.449005
lag.BD_PET	-0.006130865	0.004221114	-1.4524	0.146383
lag.BD_PCes	-0.113543908	0.081051193	-1.4009	0.161247
lag.BD_POdh	0.052706682	0.020180973	2.6117	0.009009
lag.BD_PEI	0.014454871	0.008077491	1.7895	0.073530
lag.BD_HTV	-0.009250153	0.006002225	-1.5411	0.123287

Rho: -0.17478, LR test value: 0.69524, p-value: 0.40439

Asymptotic standard error: 0.17899

Z-value: -0.97648, p-value: 0.32883

Wald statistic: 0.95351, p-value: 0.32883

Log likelihood: -185.7342 for mixed model

ML residual variance (sigma squared): 5.7129, (sigma: 2.3902)

Number of observations: 81

Number of parameters estimated: 29

AIC: 429.47, (AIC for lm: 428.16)

LM test for residual autocorrelation

test value: 0.26773, p-value: 0.60486

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Identificadas las variables sociodemográficas que tienen una correlación espacial significativa en el modelo espacial de SDM y explican la alta incidencia de unidades económicas que fueron robadas; el modelo espacial de Durbin ajustado para esta infracción a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$\text{Robo a unidades económicas} = \rho * \text{lag.BD_Robo}_{u_i} + \beta_0 + \beta_1 * \text{BD_POdh}_i + \gamma_1 * \text{lag.BD_POdh}_i + \varepsilon_i$$

donde,

$$\text{lag.BD_POdh}_i = \sum_j^n W_{ij} * \text{BD_POdh}_j,$$

y las variables lag.BD_Robo_{u_i} y ε_i son las mismas que para el modelo SAR.

determinadas las variables sociodemográficas con una correlación espacial significativa en la alta incidencia de la tasa de robo a domicilios; entonces el modelo de SDM que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presentan a continuación en la Tabla 23-4.

Tabla 23-4: Resultado del modelo espacial Durbin ajustado del número de unidades económicas robadas robo a personas año 2019.

```
Call:lagsarlm(formula = fr, data = bdta, listw = seaw, type = "mixed")
```

Residuals:

	Min	1Q	Median	3Q	Max
	-5.02332	-1.73922	-0.64028	0.76957	17.00475

Type: mixed

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.4937903	2.6186380	-0.9523	0.3409
BD_POdh	-0.0039318	0.0063936	-0.6150	0.5386
lag.BD_POdh	0.0232094	0.0142603	1.6275	0.1036

Rho: 0.56761, LR test value: 15.789, p-value: 0.000070802

Asymptotic standard error: 0.11553

Z-value: 4.9129, p-value: 0.00000089759

Wald statistic: 24.136, p-value: 0.00000089759

Log likelihood: -213.5423 for mixed model

ML residual variance (sigma squared): 10.603, (sigma: 3.2562)

Number of observations: 81

Number of parameters estimated: 5

AIC: 437.08, (AIC for lm: 450.87)

LM test for residual autocorrelation

test value: 0.45985, p-value: 0.49769

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Como se observa en la Tabla 23-4., en este modelo se estimaron siete parámetros, el parámetro autorregresivo ρ cuyo valor es 0.5676, los coeficientes β_q son: $\beta_0 = -2.49379$, $\beta_1 = -0.0232094$, el coeficiente $\gamma_1 = 0.0232094$ y por último la desviación típica de los residuos generados por este método ε_i , es decir $\sigma = 3.2562$. De igual manera que el modelo SAR devuelve un valor-p para un test de dependencia espacial para los residuos generados, este es, 0.49769, el cual indica que los nuevos residuos son independientes desapareciendo de esta manera la dependencia espacial.

A continuación, se presenta la ecuación del modelo espacial de Durbin para el robo a unidades económicas por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Robo a unidades económicas} = 0.558 * \text{lag.BD_Robo_u_} - 2.49379 - 0.0232094 * \text{BD_POdh} - 0.0232094 * \text{lag.BD_POdh}_i$$

Esto quiere decir que, por cada unidad que disminuye en el número de personas que realiza su actividad económica dentro de su hogar, el número de robo a unidades económicas disminuye en tres unidades al año.

Robo a vehículos: El modelo espacial de Durbin relacionado con el robo a vehículos genera un valor de intercepto negativo de -11, siendo este no estadísticamente significativo, entre tanto su valor-p es de 0.58, muy por encima de lo deseado. Ahora bien, dentro de las variables de tipo rezago, se tiene que existe una relación estadísticamente significativa con la variable de robo a vehículos y la tasa de desplazamiento temporal por trabajo o estudio, siendo esto una relación lógica, entre tanto, es posible tener más robos para la población que se desplaza de forma temporal. Esto se puede aseverar entre tanto, existe suficiente evidencia estadística, con un valor p muy cercano a 0, propiamente (0.006648). Por otro lado, la variable de población cesante desde el punto de vista rezago (relación del vecino del vecino) tiene suficiente evidencia estadística para afirmar que existe una relación, pues su valor-p es muy cercano a 0 (valor p igual a 0.000005288), es decir, con una confiabilidad de más del 99%.

Los resultados obtenidos en el lenguaje R se presentan a continuación.

Tabla 24-4: Resultado del modelo espacial de Durbin de vehículos robados año 2019.

```
Call: lagsarlm(formula = frm, data = shpf, listw = seaw, type = "mixed")
```

```
Residuals:
      Min       1Q   Median       3Q      Max
-6.67887 -1.86109 -0.13772  1.70849  6.73481

Type: mixed
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -11.22581430  20.37504710 -0.5510  0.581662
BD_TA        0.13449295   0.48008574  0.2801  0.779367
BD_Tasa_Na   0.03428963   0.06268828  0.5470  0.584388
BD_Tasa_Dt   0.30445131   0.10835660  2.8097  0.004959
BD_Inmi_Ex   0.00259990   0.01027536  0.2530  0.800250
BD_PMI       0.00030081   0.00027168  1.1072  0.268194
BD_NEI      -0.02105862   0.01571454 -1.3401  0.180222
BD_TIEpc     0.05663584   0.02645579  2.1408  0.032292
BD_PPH       2.26189161   2.41800499  0.9354  0.349563
BD_PET       0.00261673   0.00198554  1.3179  0.187538
BD_PCes     -0.06284627   0.04332324 -1.4506  0.146881
BD_POdh      0.02373769   0.00945173  2.5115  0.012023
BD_PEI      -0.00283402   0.00443426 -0.6391  0.522746
BD_HTV       0.00511527   0.00286243  1.7870  0.073931
lag.BD_TA   -0.31644677   0.95754146 -0.3305  0.741039
lag.BD_Tasa_Na 0.16186637   0.14682224  1.1025  0.270260
lag.BD_Tasa_Dt 0.82680044  0.30464495  2.7140  0.006648
lag.BD_Inmi_Ex -0.03644156   0.02304238 -1.5815  0.113763
lag.BD_PMI   0.00045176   0.00032210  1.4026  0.160747
```

```

lag.BD_NEI      -0.05668349   0.03344330 -1.6949   0.090092
lag.BD_TIEpc   -0.11375252   0.06731695 -1.6898   0.091065
lag.BD_PPH     -6.47318409   4.23886972 -1.5271   0.126736
lag.BD_PET     0.00355849   0.00451587  0.7880   0.430699
lag.BD_PCes    -0.41311271   0.09073367 -4.5530 0.000005288
lag.BD_Podh    -0.03943890   0.02219196 -1.7772   0.075540
lag.BD_PEI     0.00475532   0.00866856  0.5486   0.583300
lag.BD_HTV     -0.00416515   0.00640303 -0.6505   0.515371

```

```

Rho: -0.31677, LR test value: 2.2997, p-value: 0.1294
Asymptotic standard error: 0.17596
Z-value: -1.8002, p-value: 0.071823
Wald statistic: 3.2409, p-value: 0.071823

```

```

Log likelihood: -192.2806 for mixed model
ML residual variance (sigma squared): 6.6346, (sigma: 2.5758)
Number of observations: 81
Number of parameters estimated: 29
AIC: 442.56, (AIC for lm: 442.86)
LM test for residual autocorrelation
test value: 4.5942, p-value: 0.03208

```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Identificadas las variables sociodemográficas que tienen una correlación espacial significativa en el modelo de SDM y que explican la alta incidencia de vehículos que fueron robados; el modelo de espacial de Durbin ajustado para el robo a vehículos a nivel de zonas urbanas en la ciudad de Cuenca año 2019, quedaría de la siguiente manera:

$$\begin{aligned}
\text{Robo a vehículos} = & \rho * \text{lag.BD_Robo_ve}_i + \beta_0 + \beta_1 * \text{BD_Tasa_Dt}_i + \beta_2 * \text{BD_PCes}_i + \\
& \gamma_1 * \text{lag.BD_Tasa_Dt}_i - \gamma_2 * \text{lag.BD_PCes}_i + \varepsilon_i
\end{aligned}$$

donde,

$$\text{lag.BD_PCes}_i = \sum_j^n W_{ij} * \text{BD_PCes}_j$$

$$\text{lag.BD_Tasa_Dt}_i = \sum_j^n W_{ij} * \text{BD_Tasa_Dt}_j,$$

y las variables lag.BD_Robo_ve_i y ε_i son las mismas que para el modelo SAR.

Una vez que intifadas las variables sociodemográficas con una correlación espacial significativa en la alta incidencia de robo a vehículos; entonces el modelo de SDM que arroja el lenguaje estadístico R mediante su entorno de desarrollo RStudio, se presenta en la Tabla 25-4.

Tabla 25-4: Resultado del modelo espacial de Durbin ajustado de vehículos robados año 2019.

```

Call:lagsarlm(formula = fr, data = bdta, listw = seaw, type = "mixed")

Residuals:
    Min       1Q   Median       3Q      Max
-7.85490 -2.41652 -0.50013  1.99784 14.79440

Type: mixed
Coefficients: (asymptotic standard errors)
              Estimate Std. Error z value Pr(>|z|)

```

(Intercept)	0.309430	5.665071	0.0546	0.9564407
BD_Tasa_Dt	0.284116	0.116708	2.4344	0.0149158
BD_PCes	0.085327	0.041847	2.0390	0.0414464
lag.BD_Tasa_Dt	0.173920	0.316377	0.5497	0.5825087
lag.BD_PCes	-0.292204	0.081504	-3.5851	0.0003369

Rho: 0.34875, LR test value: 5.1949, p-value: 0.022653
 Asymptotic standard error: 0.14198
 z-value: 2.4564, p-value: 0.014036
 Wald statistic: 6.0337, p-value: 0.014036

Log likelihood: -222.2684 for mixed model
 ML residual variance (sigma squared): 13.809, (sigma: 3.7161)
 Number of observations: 81
 Number of parameters estimated: 7
 AIC: 458.54, (AIC for lm: 461.73)
 LM test for residual autocorrelation
 test value: 0.000013534, p-value: 0.99706

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Como se observa en la Tabla 25-4., en este modelo se estimaron siete parámetros que son, el parámetro autorregresivo ρ cuyo valor es 0.348. Los coeficientes β_q son: $\beta_0 = 0.309$, $\beta_1 = 0.284$, $\beta_2 = 0.0853$ el coeficiente $\gamma_1 = 0.1739$, $\gamma_2 = -0.2922$ y por último la desviación típica de los residuos generados por este método ε_i , es decir $\sigma = 3.71$. De igual manera que el modelo SAR nos devuelve un valor-p para un test de dependencia espacial para los residuos generados; este es 0.997, el cual indica que los nuevos residuos son independientes desapareciendo de esta manera la dependencia espacial.

A continuación, se presenta la ecuación del modelo espacial de Durbin para el robo a vehículos por zona urbana en la ciudad de Cuenca año 2019:

$$\text{Robo a vehículos} = 0.348 * \text{lag.BD_Robo_ve_} + 0.309 + 0.28 * \text{BD_Tasa_Dt} + 0.0853 * \text{lag.BD_PCes}_i + 0.1739 * \text{lag.BD_Tasa_Dt} - 0.2922 * \text{lag.BD_PCes}$$

Esto quiere decir que, por cada unidad que aumenta la tasa de desplazamiento temporal por trabajo o estudio y la población cesante disminuye, el número de robo a vehículos aumenta en una unidad al año.

4.3 Identificación de las zonas urbanas de alta incidencia delictivas del robo y variables sociodemográficas en la ciudad de Cuenca

Para analizar las variables se utilizaron los datos espaciales del año 2019 correspondientes a los delitos y del 7mo. Censo de Población y 6to. Vivienda a un nivel de zona censal (2010), aplicando un análisis inferencial espacial como se detalla en los capítulos 2 y 3, para medir e identificar las zonas censales urbanas de mayor concentración y su comportamiento mediante el *I* de Moran, como se presenta a continuación.

4.3.1 Identificación de las zonas urbanas de alta incidencia de las variables de robo en sus diferentes modalidades

A continuación se presentan los resultados y análisis mediante mapas de coropletas aplicando los métodos de estadística espacial antes mencionados los mismos que están enunciados en los capítulos 2 y 3 como, análisis de vecindad, y autocorrelación espacial mediante el índice *I* de Moran.

Robo a personas

Análisis de autocorrelación espacial con el índice *I* de Moran

Para este análisis se utilizó el criterio de *Queen* y calculó el *I* de Moran global, con el fin de determinar si existe autocorrelación espacial en la tasa de robo a personas en las diferentes zonas censales de la ciudad de Cuenca. Se observa claramente la existencia de una alta relación positiva con una ligera expectativa de relación negativa, sin embargo, el test no resulta ser estadísticamente significativo, ya que tiene un valor-p menor que 0.05, por lo que se puede decir que la tasa de robo a personas no tiene una alta autocorrelación espacial al analizar estos datos a nivel de zona censal en la ciudad de Cuenca. Puede ser que al realizar este análisis a un nivel de unidad geográfica más pequeña se obtenga una autocorrelación espacial. (Tabla 26-4). En la Figura 23-4; se observar el gráfico de dispersión, entonces como se mencionó anteriormente, se puede identificar que existe una autocorrelación positiva alta. También se observa la existencia de relación de vecindad según los cuadrantes. En este sentido, la mayor parte de los datos de robo a personas caen en el II y III cuadrante, lo que quiere decir que tienen una relación de vecindad Alto-Alto y Bajo-Alto. Adicionalmente se puede ver que existen datos atípicos en las zonas de la ciudad de Cuenca.

Tabla 26-4: Test de *I* Moran, matriz de pesos (*W*) utilizando simulación de Montecarlo para la tasa de robo a personas a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_Tasa_R_
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 6.8255, p-value = 4.38e-12
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.436995198      -0.012500000      0.004336868
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

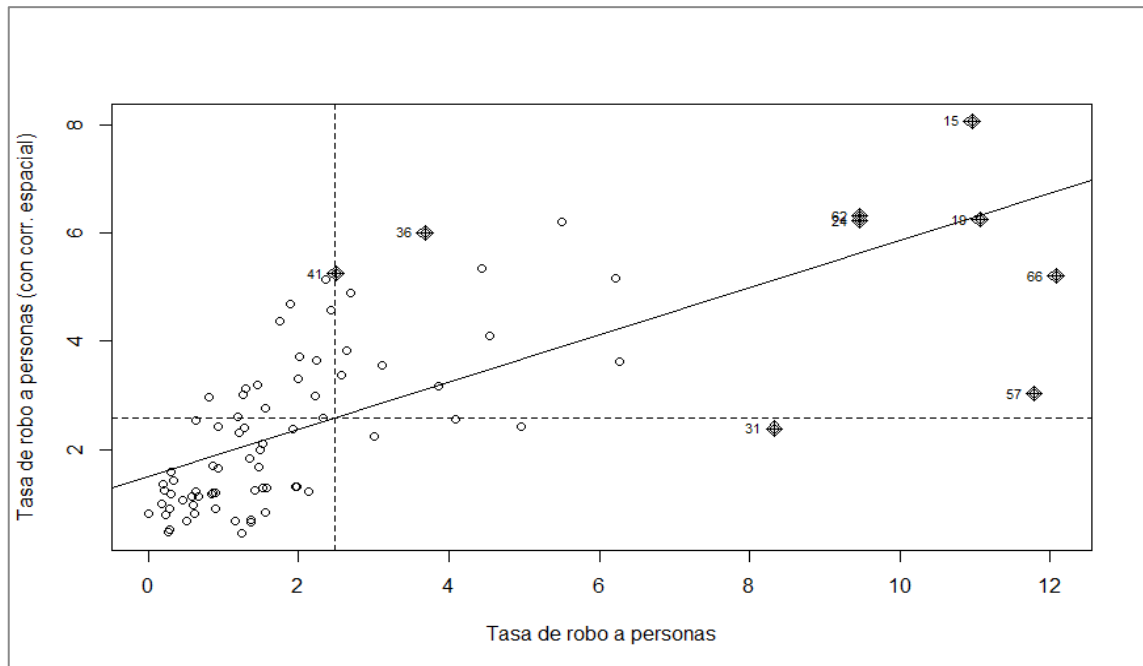


Figura 23-4: Gráfico de correlación del índice I de *Moran* global, tasa de robo a personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Para analizar la autocorrelación local o para medir la concentración o patrones existentes de las zonas censales de la tasa de robo a personas a nivel espacial geográficamente, se calculó el índice I de *Moran* a nivel local, y se puede observar en la Figura 24-4, que existe una autocorrelación significativa local en las zonas censales para la tasa de robo a personas en el centro de la ciudad; en las parroquias: El Sagrario, Gil Ramírez Dávalos, San Blas, Sucre y Huayna Cápac. Para ver esta alta concentración se muestra en la Figura 25-4, con el análisis LISA, donde claramente se observa una alta concentración de la tasa de robo a personas alta-alta distribuidas en todas las zonas de las parroquias antes mencionadas; ante esto existe una alta concentración de este delito al igual que sus zonas vecinas y bajo-bajo en una de las zonas del centro de la ciudad de Cuenca, lo que indica que hay una baja concentración de este delito al igual que sus zonas vecinas. La existencia de dos zonas Alto-Bajo, quiere decir que estas zonas tienen una alta concentración de este delito pero que sus zonas vecinas tienen una concentración baja; mientras que las demás zonas que están distribuidas en la ciudad de Cuenca como insignificantes no son significativas estadísticamente, por lo que no tienen una ocurrencia importante de robo a personas.

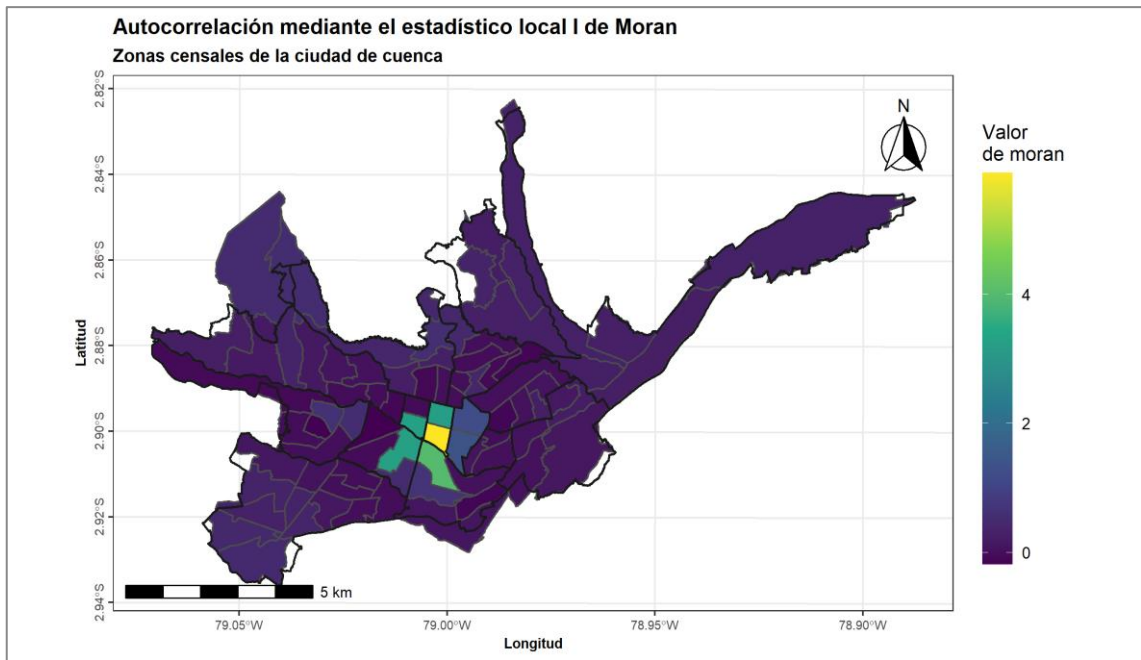


Figura 24-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de la tasa de robo a personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

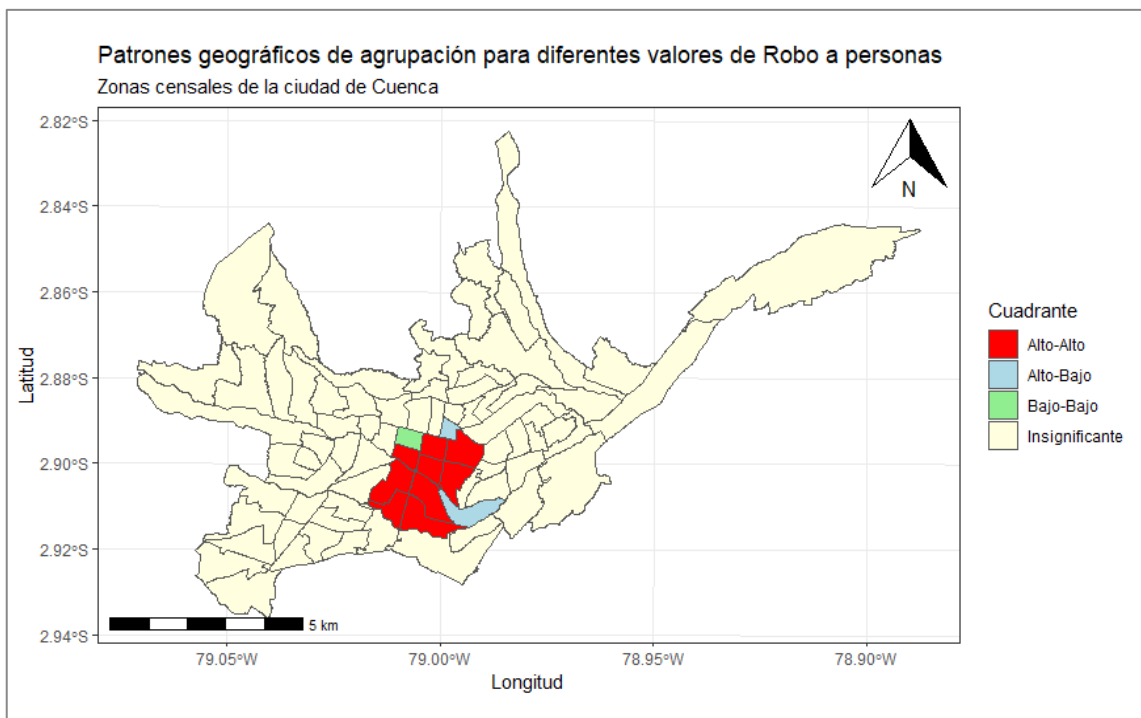


Figura 25-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a domicilios

Análisis de autocorrelación índice *I* de Moran

Utilizando el criterio de *Queen* se calculó el *I* de *Moran* global, para determinar si existe autocorrelación espacial de la tasa de robo a domicilios en las diferentes zonas censales de la ciudad de Cuenca. Se observa claramente que existe una leve relación positiva con una ligera expectativa de relación negativa, sin embargo, el test no resulta ser estadísticamente significativo, ya que tiene un valor p menor que 0.05, por lo que se puede decir que la tasa de robo a domicilios no tiene una alta autocorrelación espacial al analizar estos datos a nivel de zona censal en la ciudad de Cuenca. Es posible que al analizar a un nivel de unidad geográfica más pequeña esta tenga autocorrelación espacial (Tabla 27-4). Ahora en la Figura 26-4; se observa el gráfico de dispersión, en el que como se mencionó anteriormente, se puede ver la existencia de una autocorrelación positiva leve. También se observa que la existencia de relación de vecindad según los cuadrantes muestra que la mayor parte de los datos de robo a personas caen en el II y III cuadrante, lo que indica que tienen una relación de vecindad Alto-Alto, Bajo-Bajo, Alto-Bajo y Bajo-Alto. También se puede ver que existen datos atípicos en las zonas de la ciudad de Cuenca.

Tabla 27-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa de robo a domicilios a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_Tasa__1
## weights: queen_Cuenca2_lw
## Moran I statistic standard deviate = 1.7376, p-value = 0.04114
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.103934355      -0.012500000      0.004490004
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

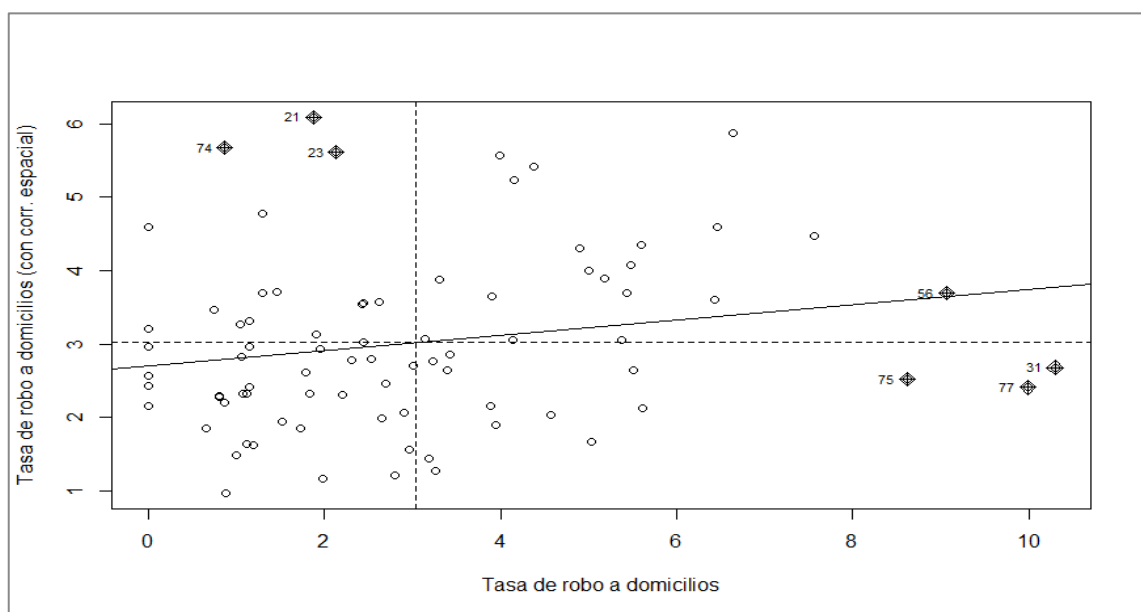


Figura 26-4: Gráfico de correlación del índice *I* de Moran, tasa de robo a domicilios por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Para analizar la autocorrelación local, medir la concentración o patrones existentes de las zonas censales de la tasa de robo a domicilios a nivel espacial geográficamente, se calculó el índice *I* de Moran a nivel local. Al observar la Figura 27-4, se identifica que existe una autocorrelación significativa local en las zonas censales para el número de robos a domicilios en el centro de la ciudad y también en otras zonas en las parroquias alejadas del centro; estas son: Sucre, Huayna Cápac, El Vecino, Bellavista, Yanuncay y Mony. Esta alta concentración se muestra en la Figura 28-4, con el análisis LISA, donde claramente se observa una alta concentración de los domicilios que fueron robados, Alto-Alto en las zonas del centro de la ciudad, lo cual indica que existe una alta concentración de domicilios donde se cometió este delito al igual que en las zonas vecinas, y Bajo-Bajo en dos zonas del centro de la ciudad lo que indica que hay una baja concentración al igual que sus zonas vecinas, mientras que existen dos zonas Alto-Bajo y Bajo-Alto, por lo que estas zonas tienen una alta concentración, pero sus zonas vecinas tienen una baja concentración, pues sus zonas vecinas tienen una alta concentración donde ocurrió este delito; mientras que las demás zonas que están distribuidas en toda la ciudad tienen un nivel insignificante, lo cual indica que no son significativas estadísticamente para la ocurrencia de este delito.

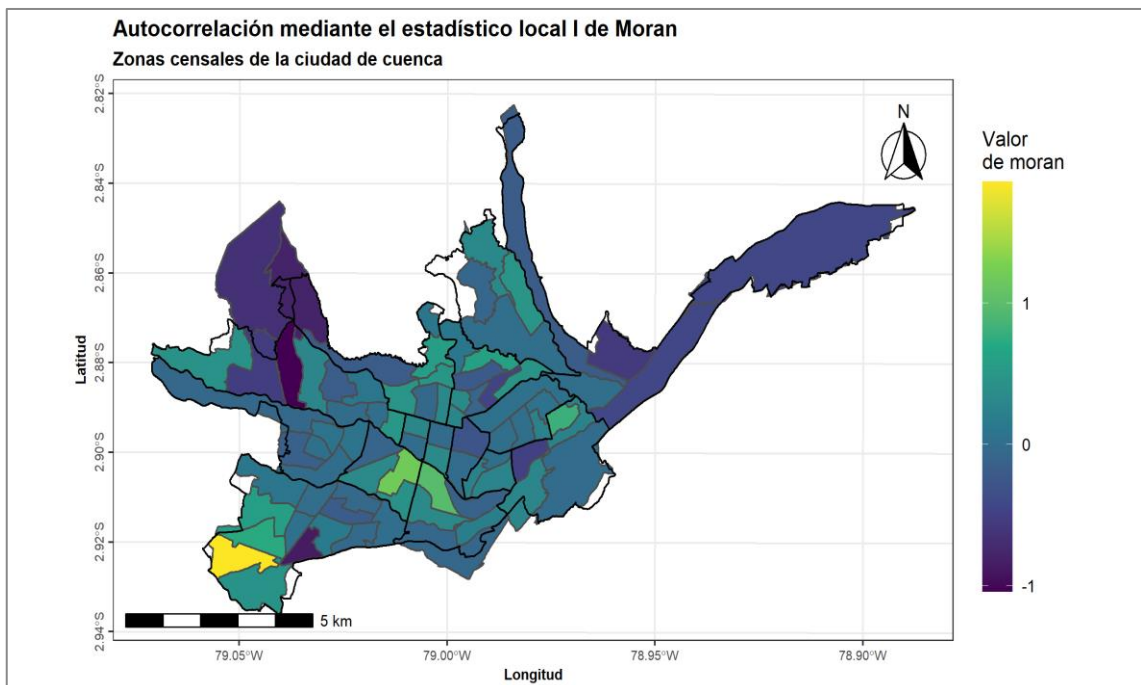


Figura 27-4: Mapa del test de Moran a nivel local, para los diferentes valores de la tasa de robo a domicilios por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

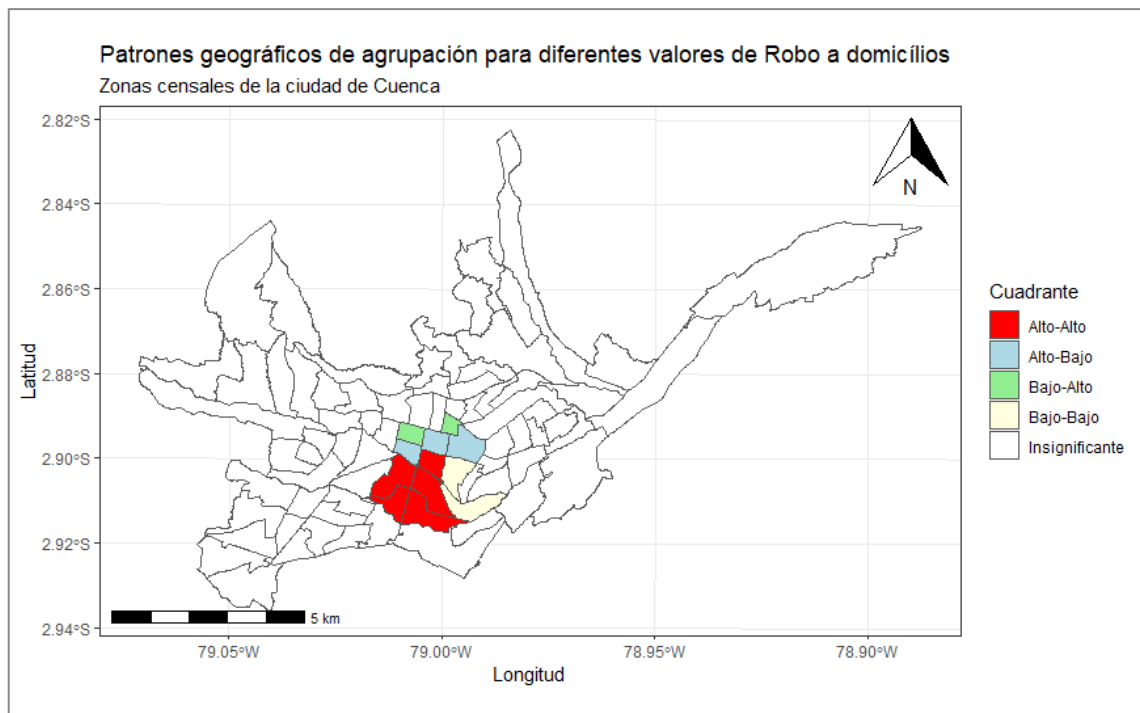


Figura 28-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a domicilios por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a unidades económicas

Análisis de autocorrelación índice *I* de Moran

Utilizando el criterio de *Queen* se calculó el *I* de *Moran* global, para determinar si existe autocorrelación espacial del número de robo a unidades económicas en las diferentes zonas censales de la ciudad de Cuenca. Se observa claramente la existencia de una moderada relación positiva con una ligera expectativa de relación negativa, sin embargo, el test no resulta ser estadísticamente significativo, ya que tiene un valor-p menor que 0.05, por lo que se puede decir que la tasa de robo a unidades económicas no tiene una alta autocorrelación espacial, al analizar estos datos a nivel de zona censal en la ciudad de Cuenca. Es posible que, al analizar a un nivel de unidad geográfica más pequeña, esta tenga autocorrelación espacial (Tabla 28-4). En la Figura 29-4; se puede observar en el gráfico de dispersión la existencia de una autocorrelación positiva moderada. También se observa la existencia de relación de vecindad según los cuadrantes, ya que la mayor parte de los datos de robo a unidades económicas caen en los cuatro cuadrantes, lo que quiere decir que tienen una relación de vecindad Alto-Alto, Bajo-Bajo y Alto-Bajo y Bajo-Alto. También se puede ver que existen datos atípicos en las zonas de la ciudad de Cuenca, esto se visualizará de mejor manera al aplicar el *I* de *Moran*.

Tabla 28-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de robos a unidades económicas a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_Robo_U_
## weights: queen_Cuenca3_lw
## Moran I statistic standard deviate = 5.3942, p-value = 3.441e-08
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.317622408      -0.012500000      0.003745369
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

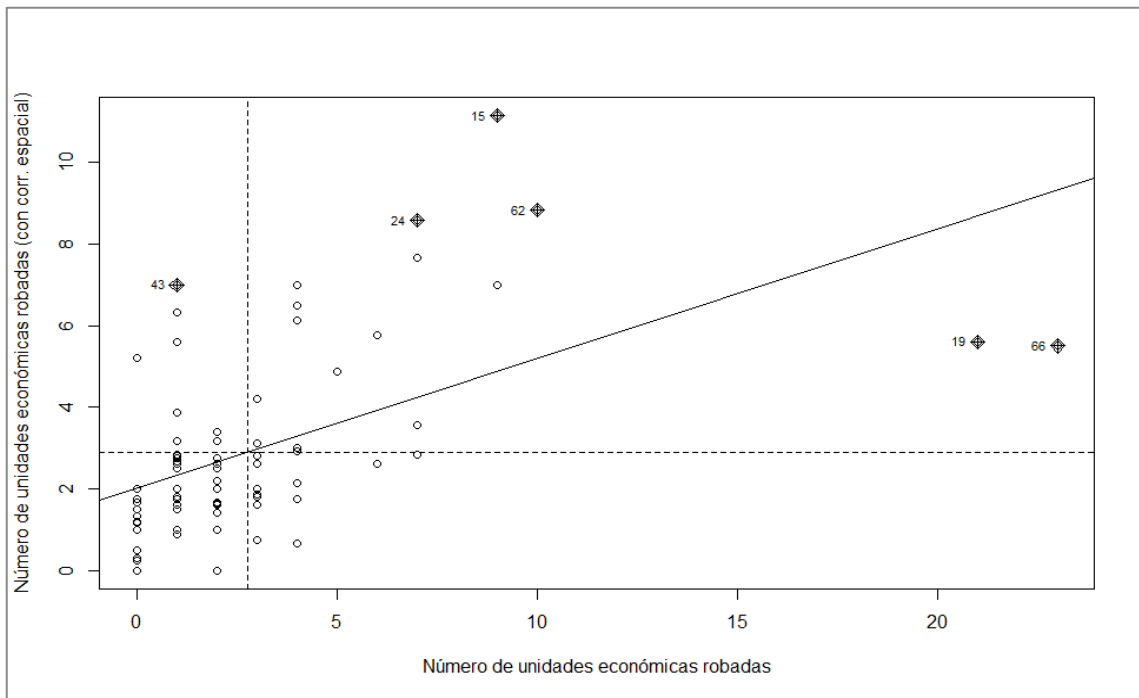


Figura 29-4: Gráfico de correlación del índice *I* de Moran, número de robo a unidades económicas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Para analizar la autocorrelación local, medir la concentración o patrones existentes de las zonas censales del número de robo a unidades económicas a nivel espacial geográficamente, se calculó el índice *I* de Moran a nivel local, y se puede observar en la Figura 30-4, que existe una autocorrelación significativa local en las zonas censales para el número de robo a unidades económicas en el centro de la ciudad, esta alta concentración se muestra en la Figura 31-4, con el análisis LISA, donde claramente se identifica una alta concentración en el número de unidades económicas que fueron robadas. En el caso Alto-Alto, existió una alta concentración de robos en estas zonas urbanas al igual que sus vecinos; mientras que en Bajo-Bajo, existe una baja concentración al igual que sus vecinos, Alto-Bajo; indica que existe una alta concentración en estas zonas pero sus vecinos tienen una baja concentración en las zonas del centro de la ciudad de Cuenca; mientras que las demás zonas son insignificantes, lo cual indica que no son

zonas de alta incidencia y estadísticamente implica que no son significativas; por lo que la cantidad de sucesos de robo a unidades económicas no tiene un alto riesgo.

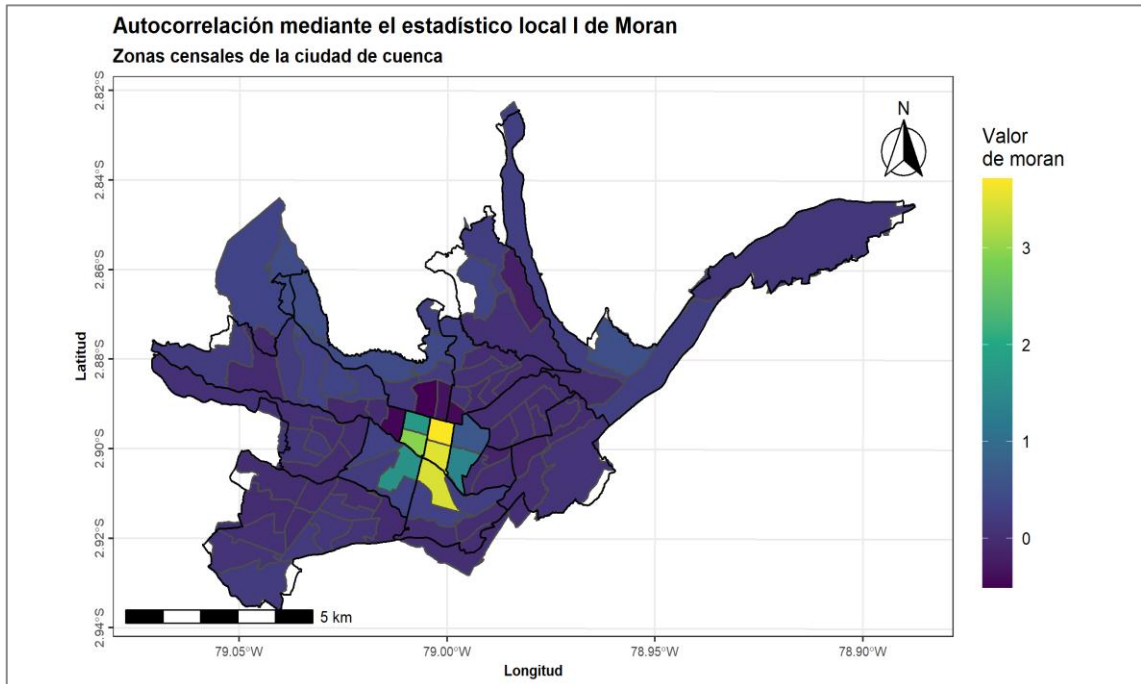


Figura 30-4: Mapa del test de *Moran* a nivel local, para los diferentes valores del número de robo a unidades económicas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

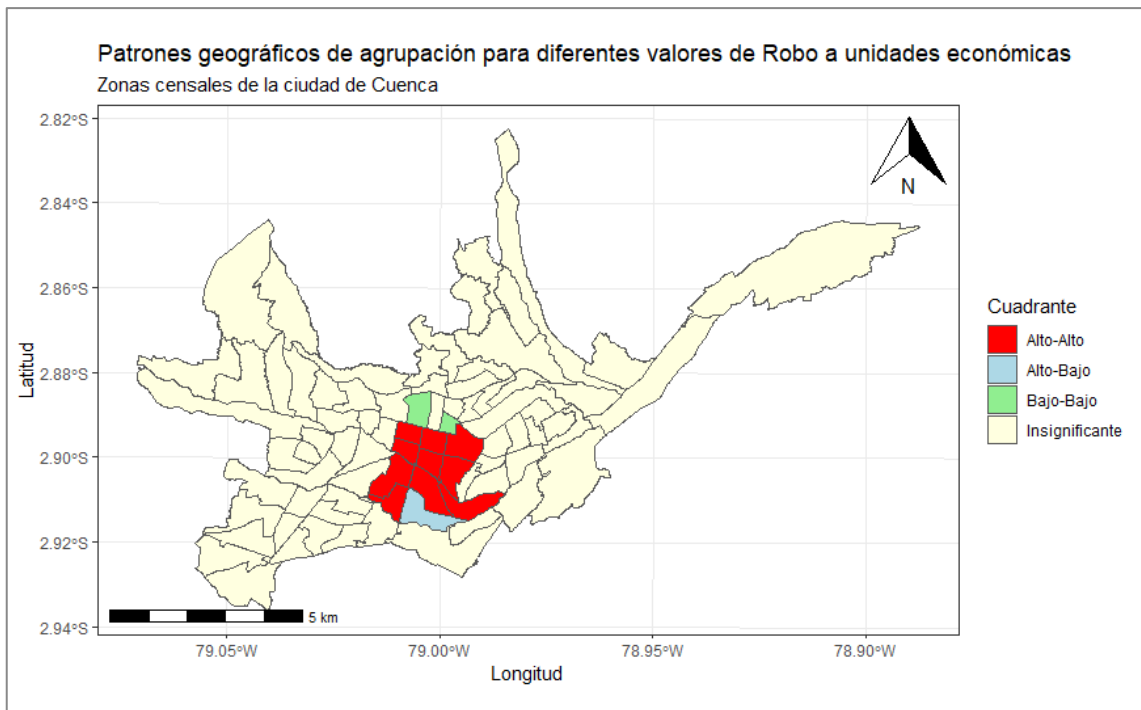


Figura 31-4: Mapa de patrones geográficos de agrupación, para los diferentes valores del número de robo a unidades económicas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo de vehículos

Análisis de autocorrelación índice *I* de Moran

Utilizando el criterio de *Queen* se calculó el *I* de Moran global para determinar si existe autocorrelación espacial de robo a vehículos en las diferentes zonas censales de la ciudad de Cuenca. Se observa claramente la existencia de una leve relación positiva con una ligera expectativa de relación negativa, sin embargo, el test no resulta ser estadísticamente significativo, ya que tiene un valor-p menor que 0.05, por lo que el número de robo de vehículos no tiene una alta autocorrelación espacial al analizar estos datos a nivel de zona censal en la ciudad de Cuenca. Es posible que al analizar a un nivel de unidad geográfica más pequeña esta tenga autocorrelación espacial (Tabla 29-4). En la Figura 32-4; se puede observar el gráfico de dispersión, y como se mencionó anteriormente, claramente se puede ver que existe una autocorrelación positiva leve. También se observa la existencia de relación de vecindad según los cuadrantes, en la que la mayor parte de los datos de robo a vehículos caen en los cuatro cuadrantes, lo que quiere decir que tienen una relación de vecindad Alto-Alto, Bajo-Bajo y Alto-Bajo, Bajo-Alto. También se puede ver que existen datos atípicos en las zonas de la ciudad de Cuenca.

Tabla 29-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de robos a vehículos a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_Robo_ve
## weights: queen_Cuenca4_lw
## Moran I statistic standard deviate = 2.5317, p-value = 0.005675
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.154133789      -0.012500000      0.004332033
```

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

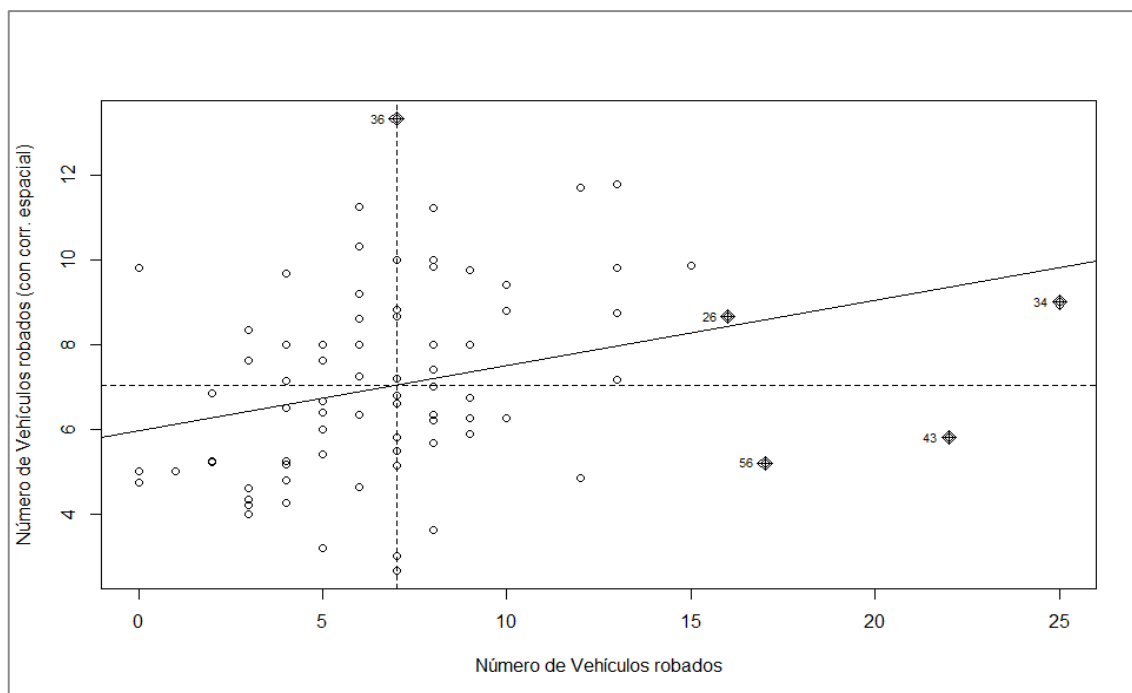


Figura 32-4: Gráfico de correlación del índice *I* de *Moran*, número de robo a vehículos por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Para analizar la autocorrelación local, medir la concentración o patrones existentes de las zonas censales del número de robo a vehículos a nivel espacial geográficamente, se calculó el índice *I* de *Moran* a nivel local, y se puede observar en la Figura 33-4, que existe una autocorrelación significativa local en las zonas censales para el número de robo a vehículos distribuido en las zonas censales de las parroquias urbanas de Huayna Cápac y Sucre, esta alta concentración se muestra en la Figura 34-1. Con el análisis LISA se observa una alta concentración de robo a vehículos Alto-Alto en las zonas censales de las parroquias urbanas de Huayna Cápac, Sucre, Yanuncay y El Batán, lo que quiere decir que en estas zonas existe una alta concentración al igual que sus vecinos; Bajo-Bajo en las zonas censales de la parroquia El Vecino lo cual indica que hay una baja concentración al igual que sus vecinos; Alto-Bajo en una zona de la parroquia Hermano Miguel, existiendo una alta concentración pero en sus vecinos existe una baja concentración de este delito y Bajo-Alto en la parroquia San Sebastián, con zonas donde existe una baja concentración pero en sus vecinos existe una concentración alta de este delito. Las demás zonas distribuidas en la ciudad de Cuenca son insignificantes, es decir que no son significativas estadísticamente, por lo que no existe una concentración importante de este delito al igual que en sus vecinos.

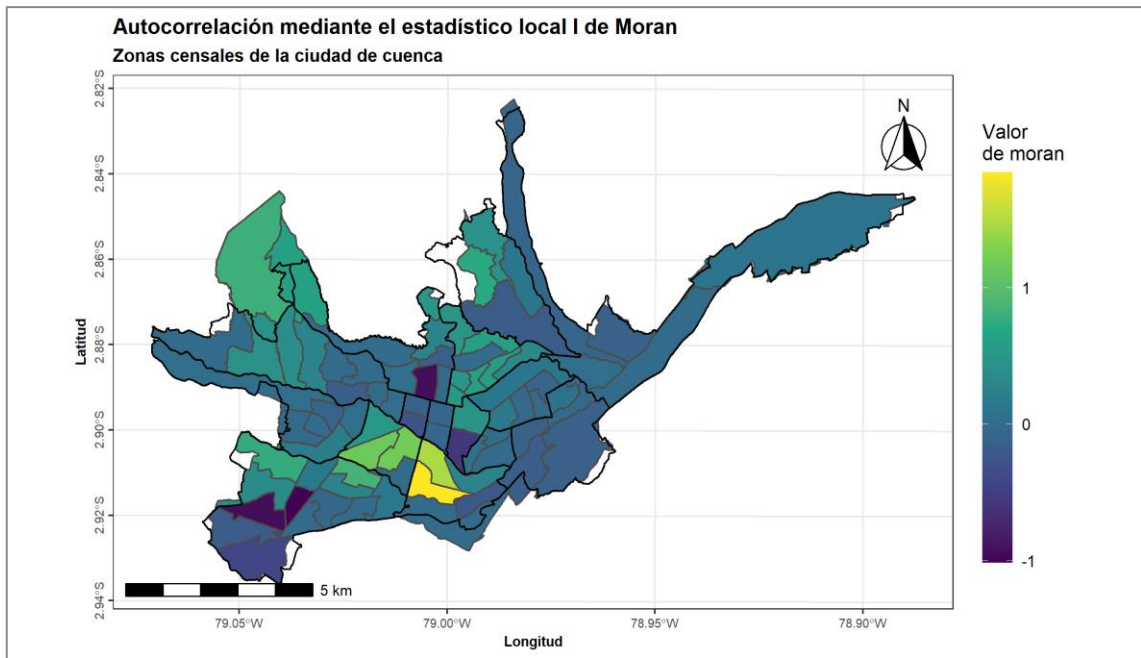


Figura 33-4: Mapa del test de *Moran* a nivel local, para los diferentes valores del número robo a vehículos por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

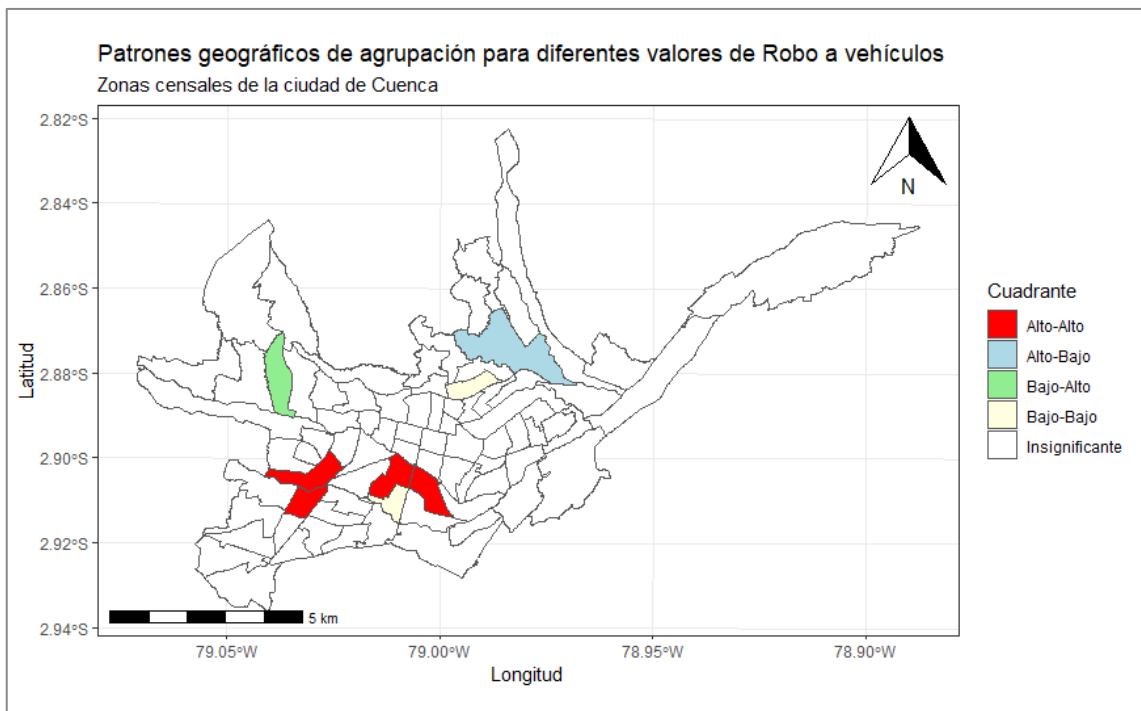


Figura 34-1: Mapa de patrones geográficos de agrupación, para los diferentes valores del número de robo a vehículos por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

4.3.2 Análisis inferencial espacial de las variables sociodemográficas

Para el análisis de las variables sociodemográficas que se presenta a continuación, se realizó el mismo análisis elaborado en el apartado de las variables de las diferentes modalidades del delito de robos, sección 4.3.1. A continuación, se presentan las correspondientes figuras y tablas, ya que el análisis del modelo de autocorrelación *I* de Moran tiene el mismo criterio, ya sea de manera global o local para identificar la concentración en las diferentes zonas de la ciudad de Cuenca.

Demografía y distribución espacial

Población (habitantes)

Análisis de autocorrelación índice *I* de Moran

Tabla 30-4: Test de *I* Moran, matriz de pesos (*W*) utilizando simulación de Montecarlo para el número de habitantes (población) a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:   shp_Cuenca$bd_TOTPOB
## weights: queen_Cuenca_lw
## Moran I statistic standard deviate = 0.13414, p-value = 0.4466
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      -0.003315065      -0.012500000      0.004688216
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velastegui Casco, Edy, 2022.

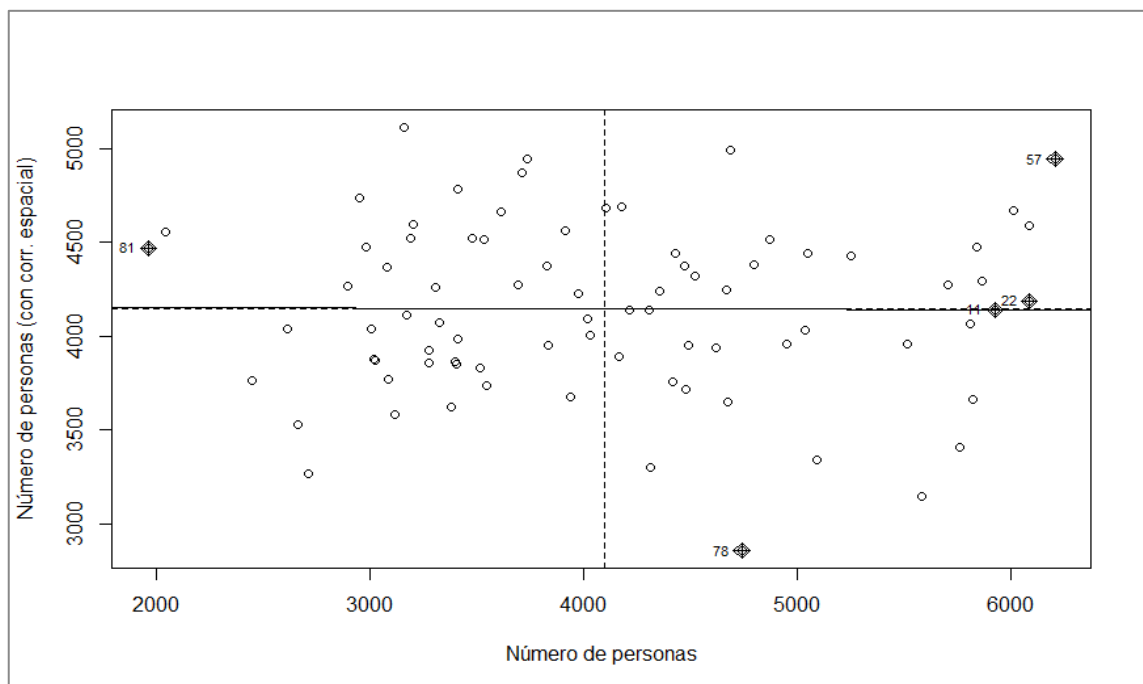


Figura 35-4: Gráfico de correlación del índice *I* de Moran, número de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

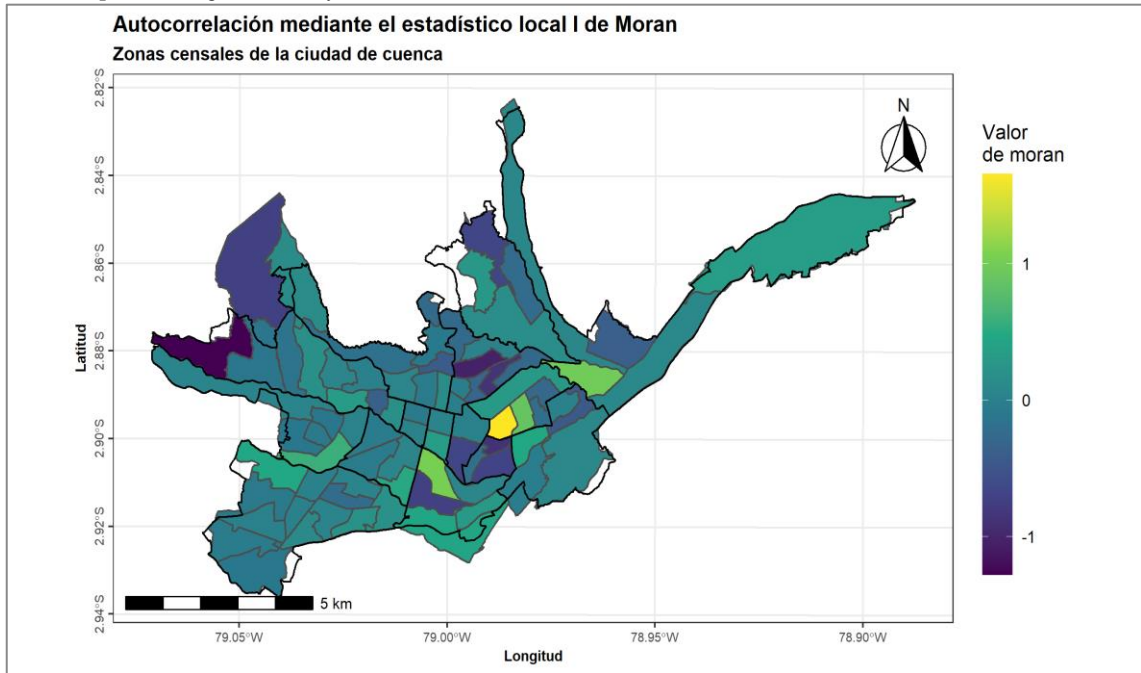


Figura 36-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de población por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

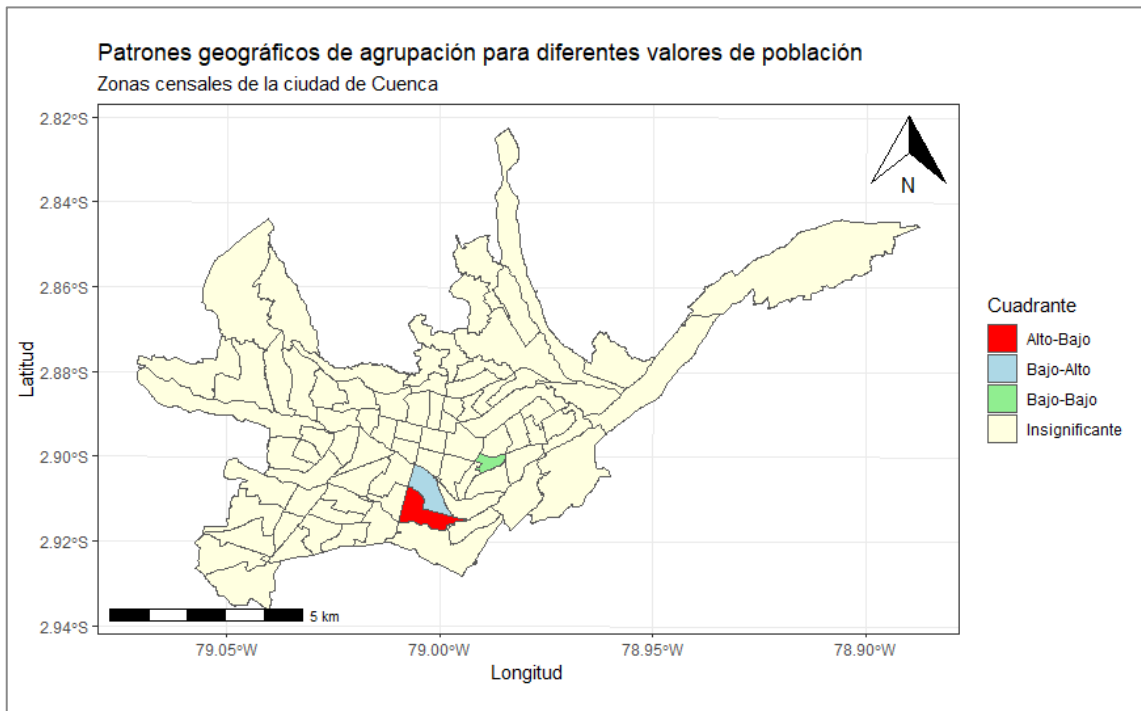


Figura 37-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la población por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Emigrantes a escala internacional

Análisis de autocorrelación índice *I* de Moran

Tabla 31-4: Test de *I* Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de Emigrantes a escala internacional a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_PMI
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 11.248, p-value < 2.2e-16
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.751842935      -0.012500000      0.004618022
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

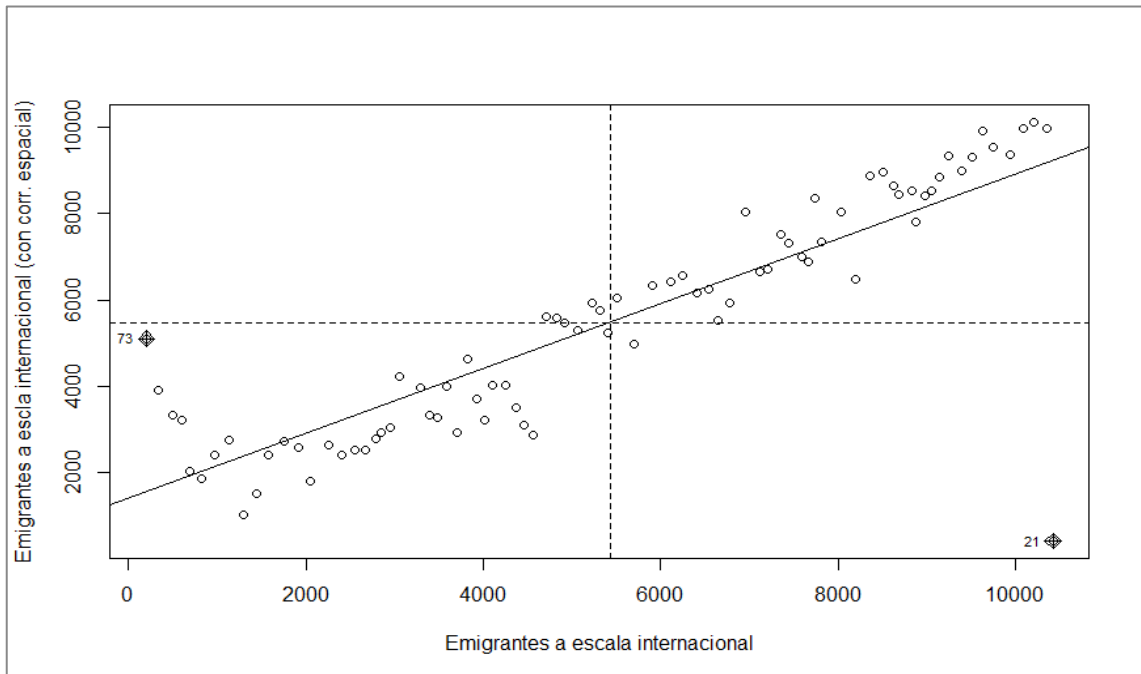


Figura 38-4: Gráfico de correlación del índice *I* de Moran, emigrantes a escala internacional por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

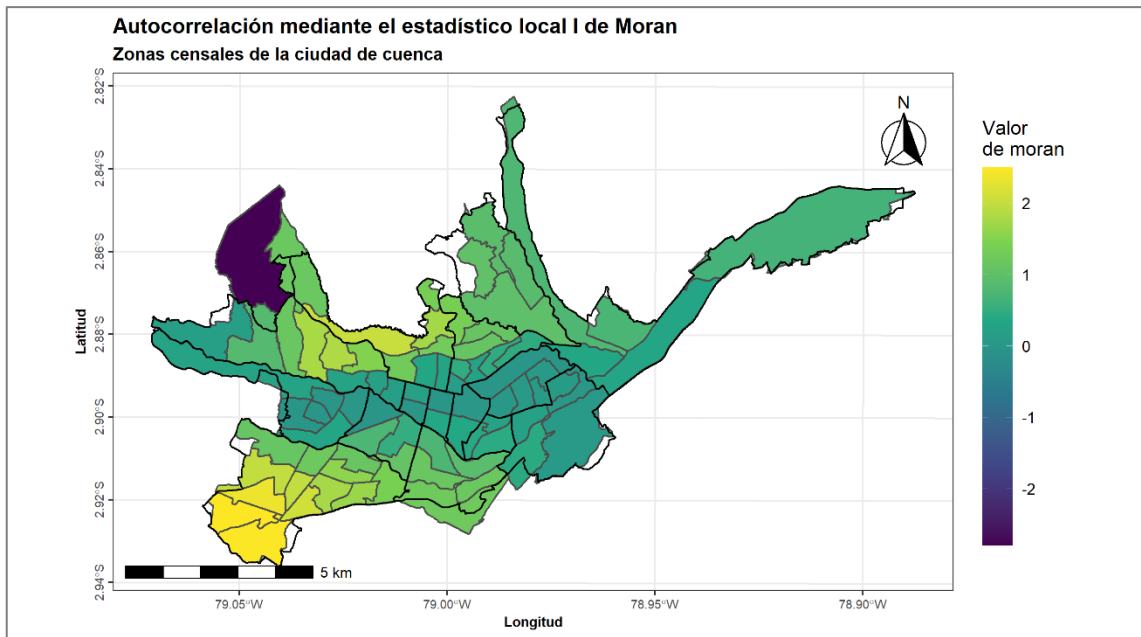


Figura 39-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de emigración a escala internacional por zonas censales urbanas de la ciudad de Cuenca, 2019.
Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

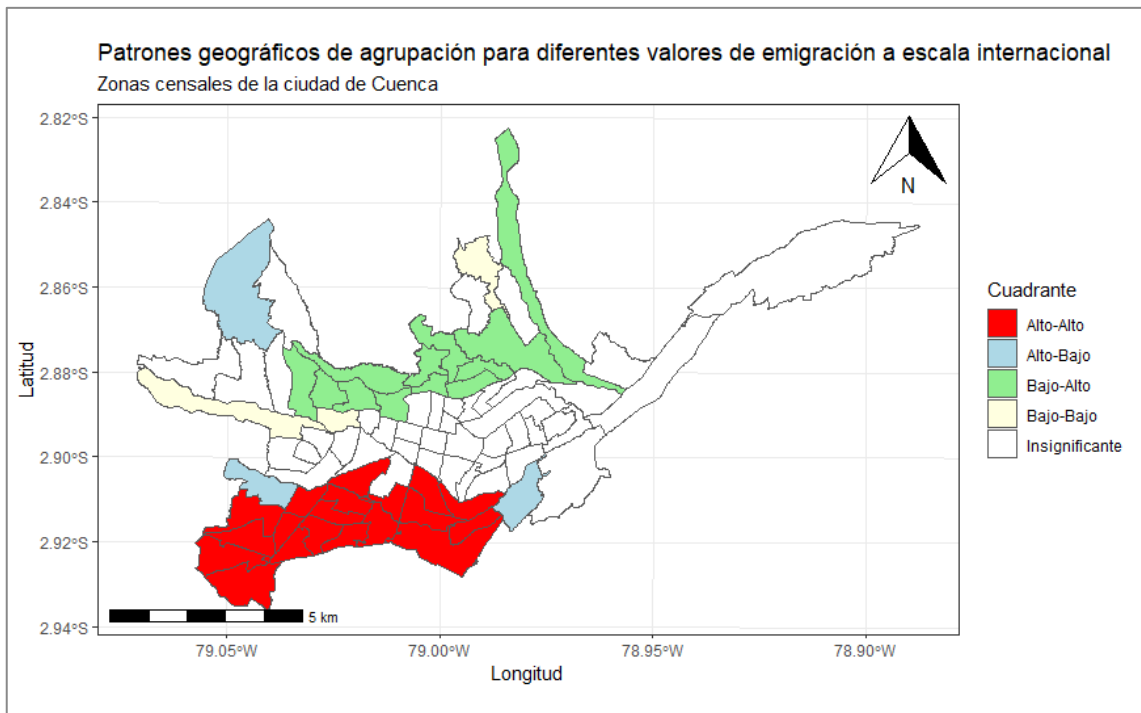


Figura 40-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de emigración a escala internacional por zonas censales urbanas de la ciudad de Cuenca, 2019.
Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

Tasa de inmigrantes extranjeros presentes en el censo

Análisis de autocorrelación índice *I* de Moran

Tabla 32-4: Test de *I* Moran, matriz de pesos (*W*) utilizando simulación de Montecarlo para tasa de inmigrantes extranjeros presentes en el censo a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_TIEpc
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 2.6888, p-value = 0.003586
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.165653896      -0.012500000      0.004390103
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

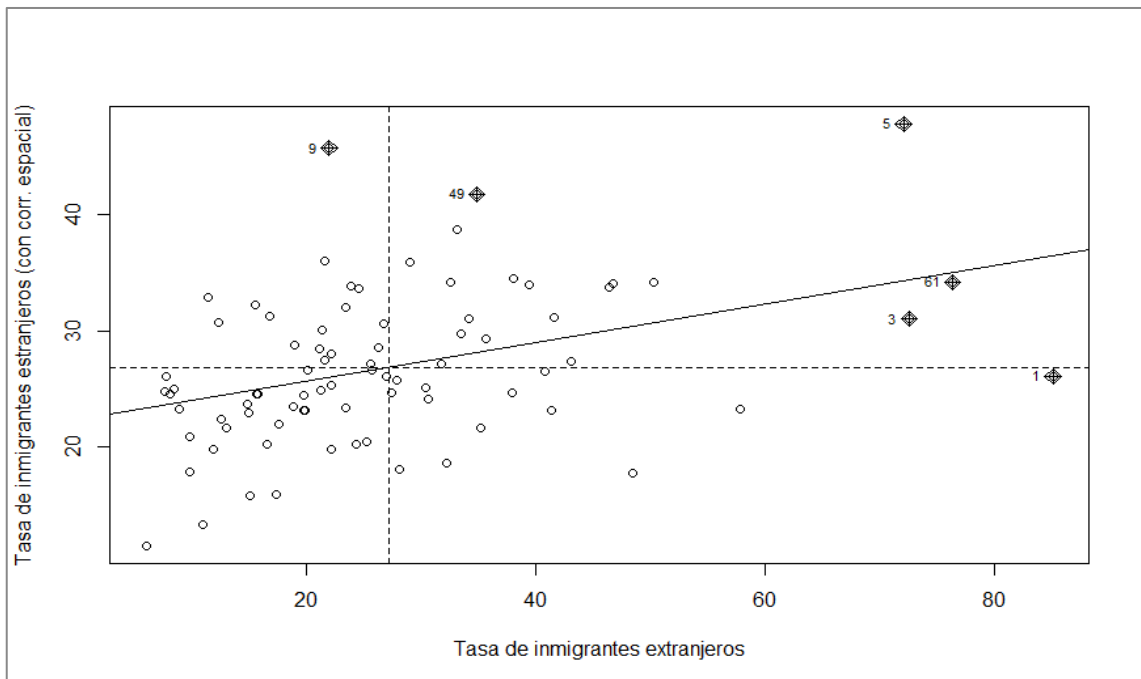


Figura 41-4: Gráfico de correlación del índice *I* de Moran, tasa de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

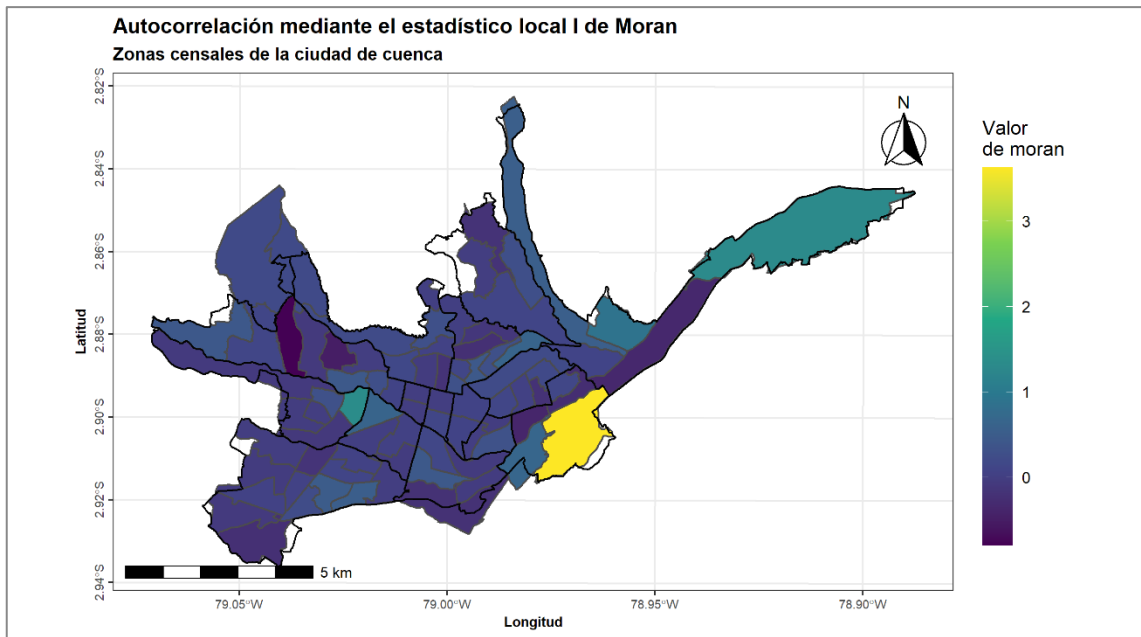


Figura 42-2: Mapa del test de *Moran* a nivel local, para los diferentes valores de la tasa de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

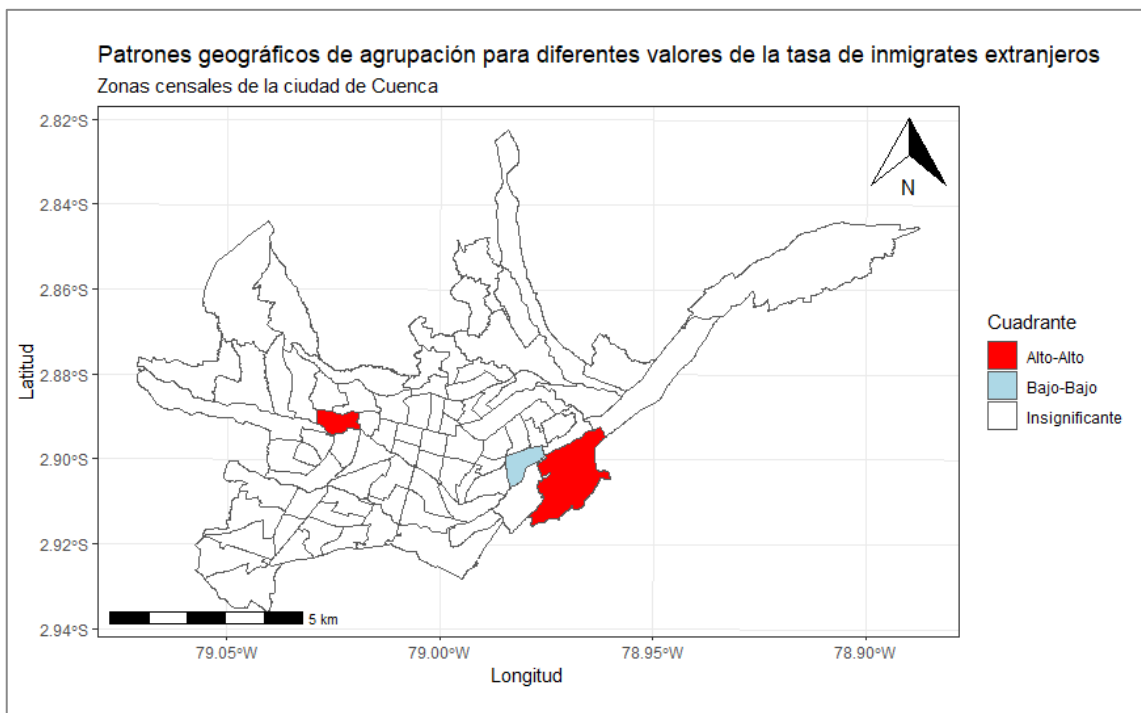


Figura 43-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Promedio de personas por hogar

Análisis de autocorrelación índice *I* de Moran

Tabla 33-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el promedio de personas por hogar a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_PPH
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 9.5631, p-value < 2.2e-16
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.63328872      -0.01250000      0.00456019
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

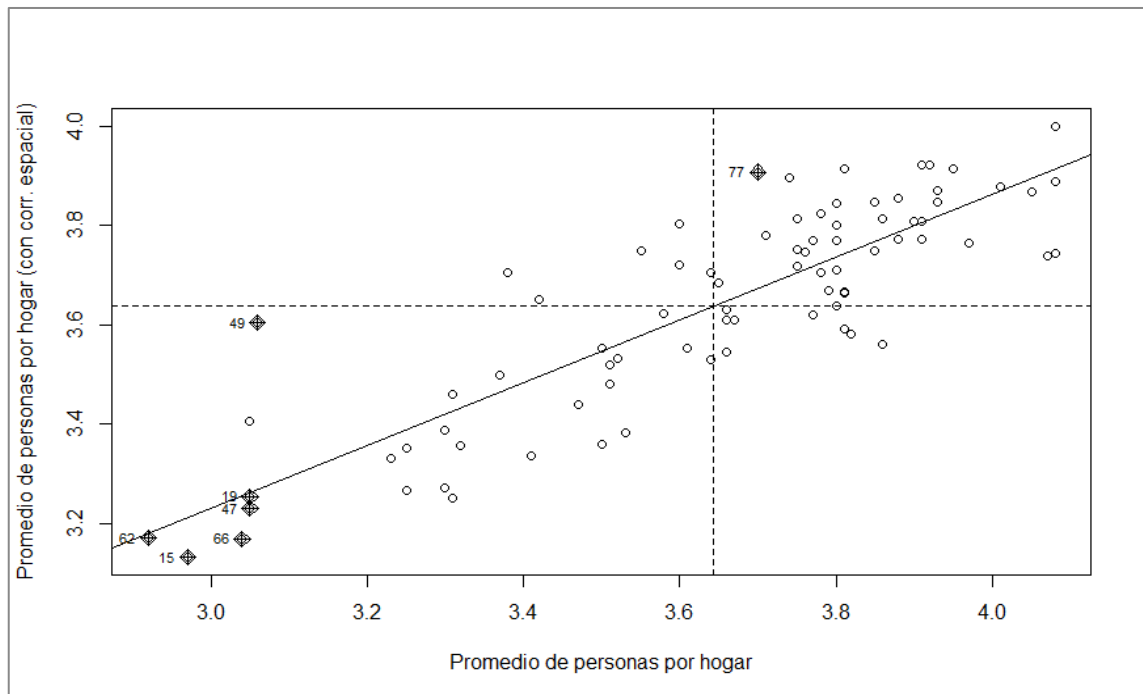


Figura 44-3: Gráfico de correlación del índice *I* de Moran, promedio de personas por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

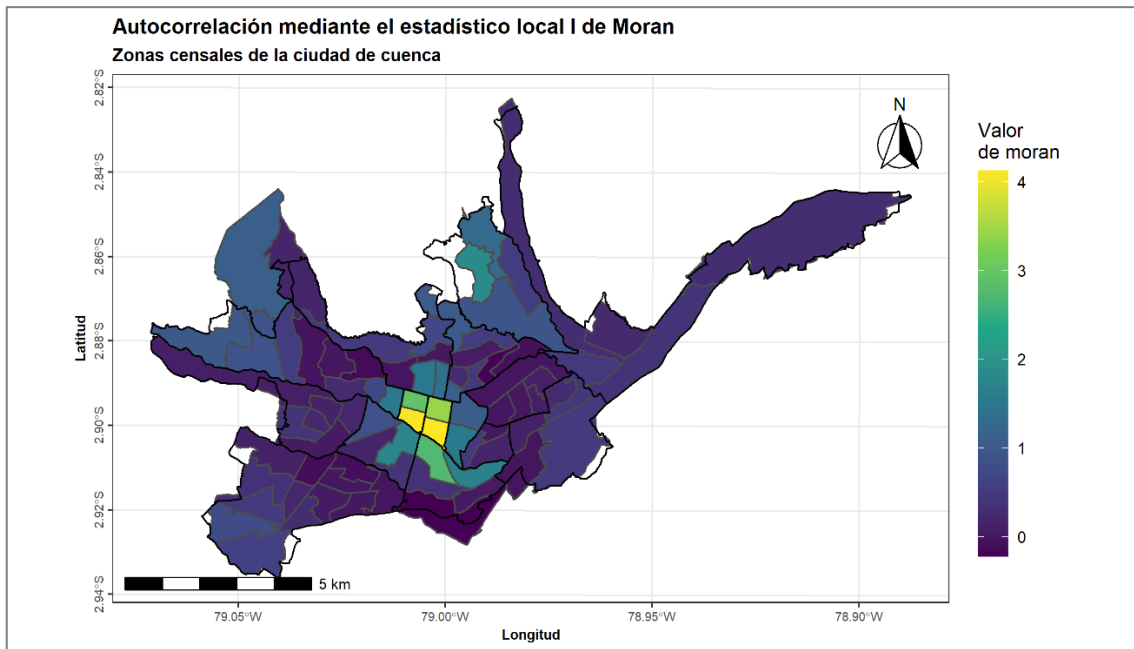


Figura 45-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de promedio de personas por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

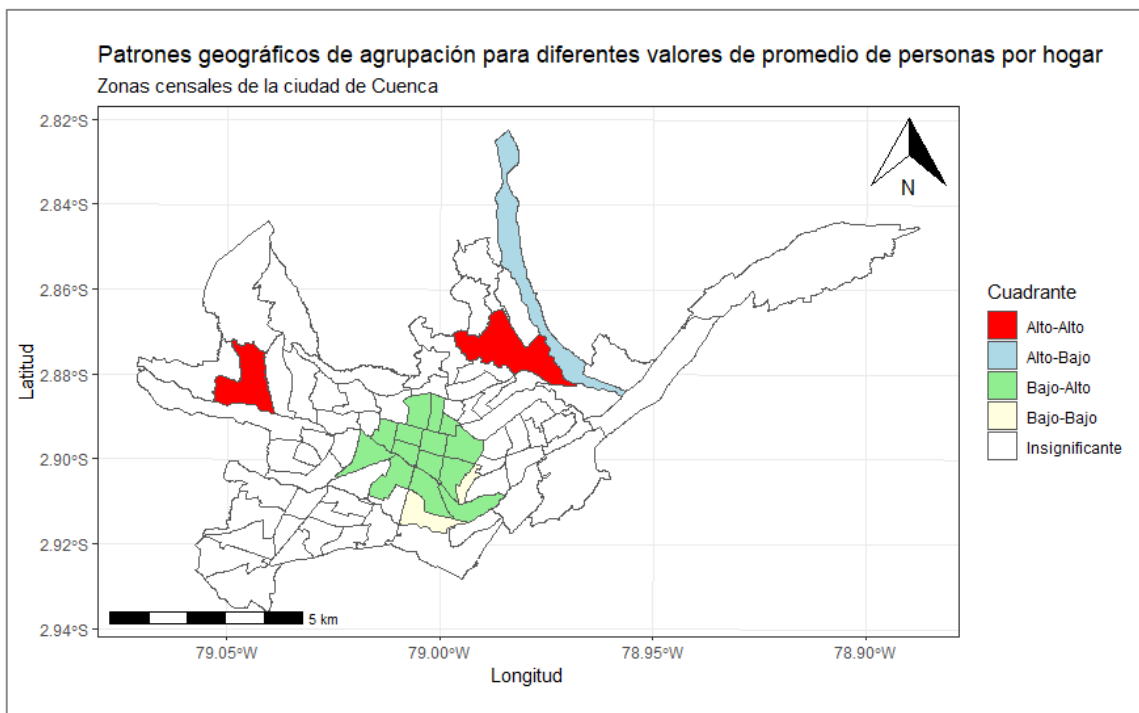


Figura 46-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de promedio de personas por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Tasa de desplazamiento temporal

Análisis de autocorrelación índice *I* de Moran

Tabla 34-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa de desplazamiento temporal de personas a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_Tasa_Dt
## weights: queen_Cuenca_lw
## Moran I statistic standard deviate = -0.57263, p-value = 0.7166
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      -0.049776391      -0.012500000      0.004237522
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

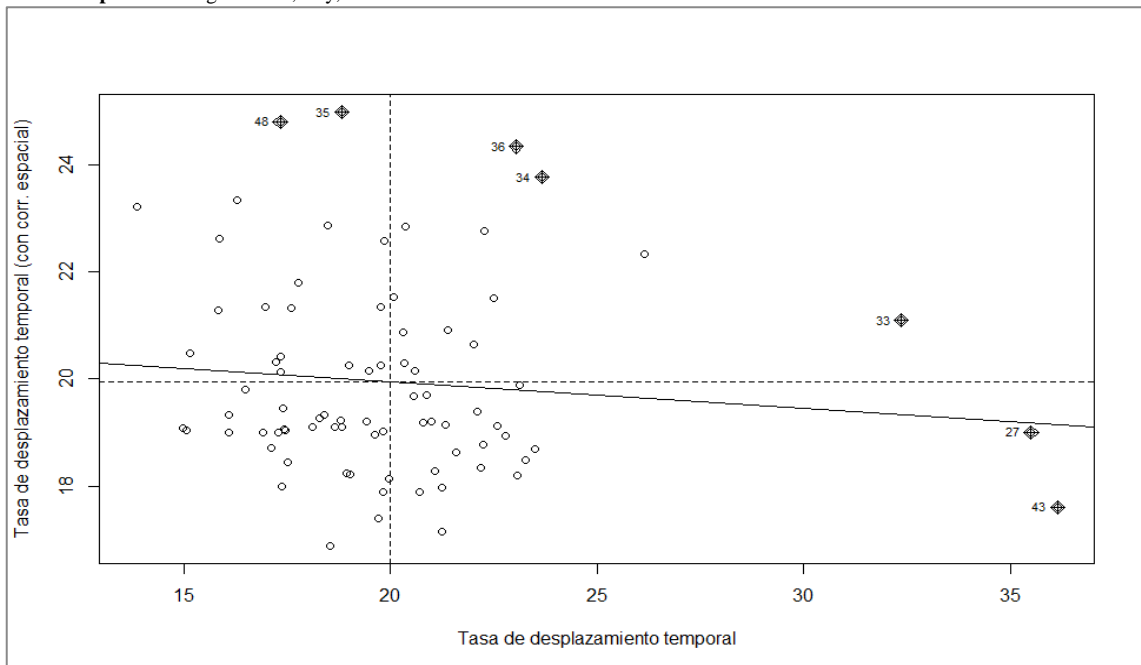


Figura 47-4: Gráfico de correlación del índice *I* de Moran, tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

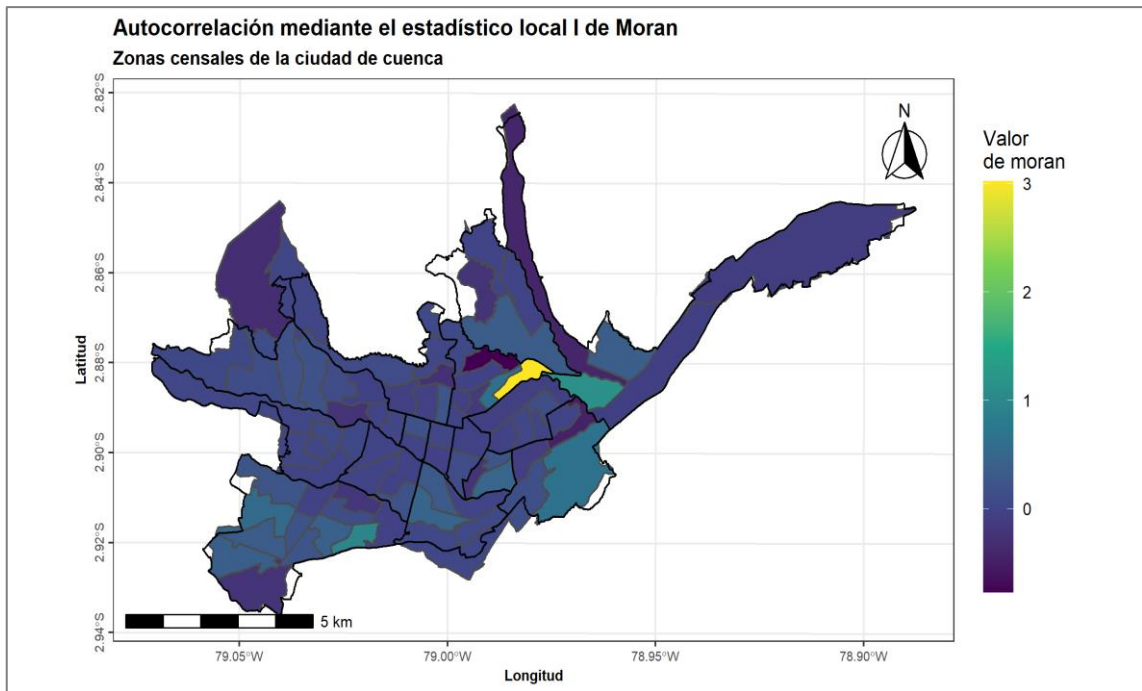


Figura 48-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de la tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

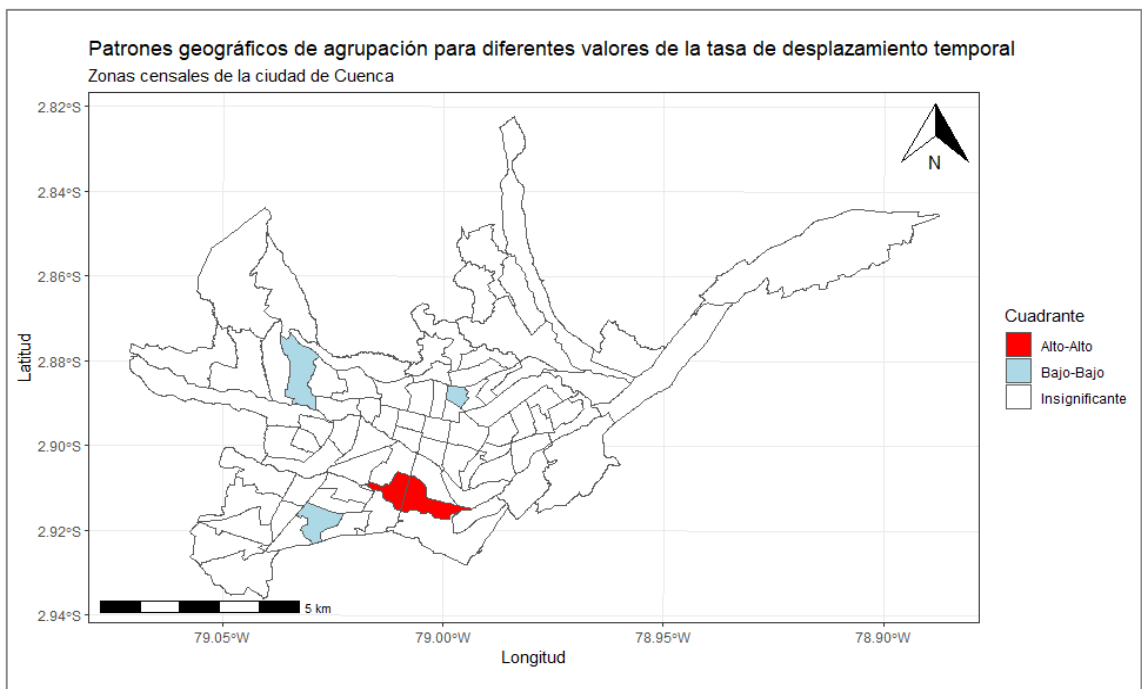


Figura 49-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Inmigrante extranjero

Análisis de autocorrelación índice *I* de Moran

Tabla 35-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de inmigrantes extranjeros a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_Inmi_Ex
## weights: queen_Cuenca_lw
## Moran I statistic standard deviate = 2.0634, p-value = 0.01954
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.127597362      -0.012500000      0.004609886
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

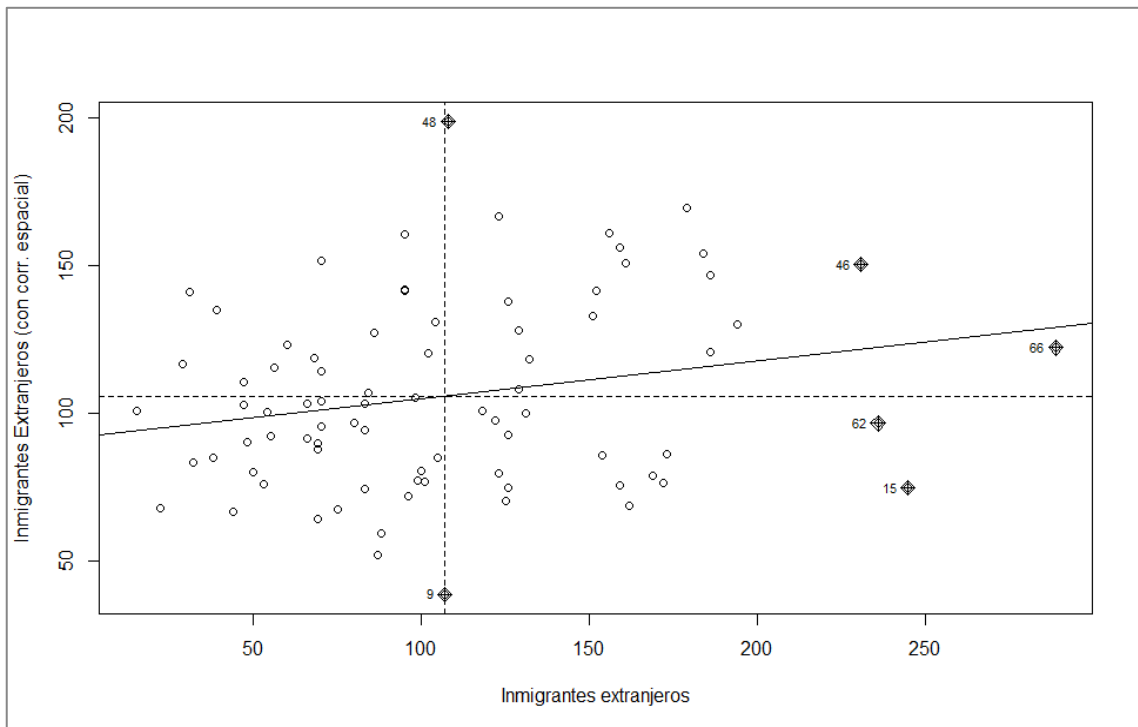


Figura 50-4: Gráfico de correlación del índice *I* de Moran, número de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

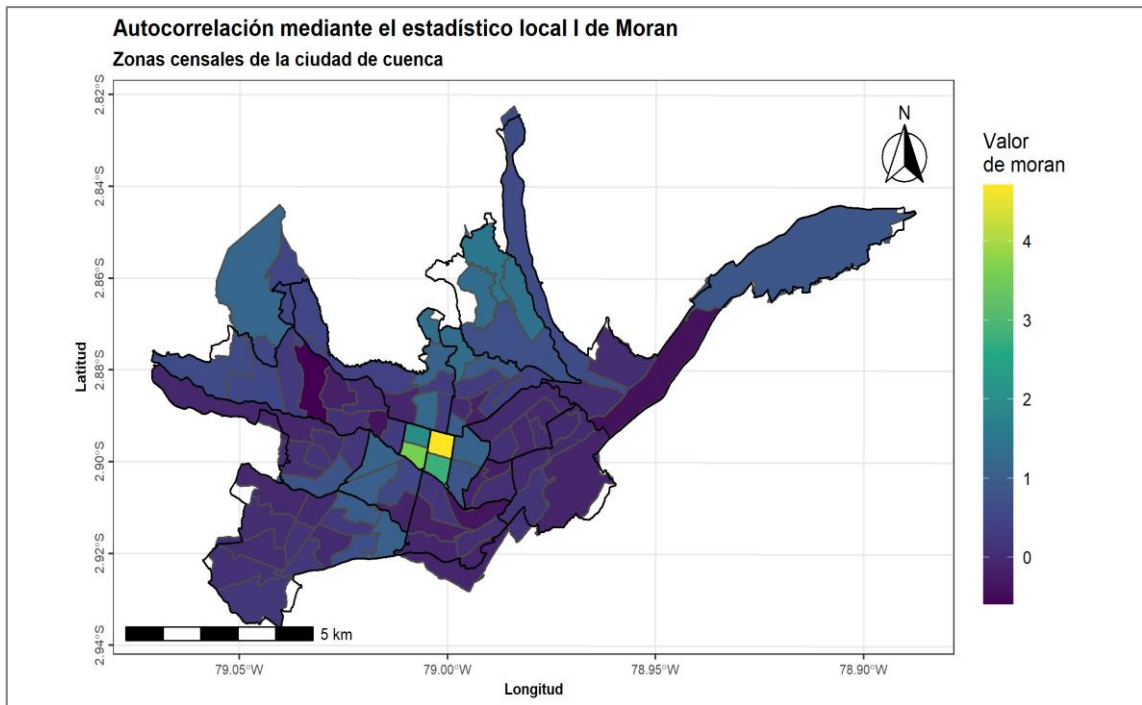


Figura 51-4: Mapa del test de *Moran* a nivel local, para los diferentes valores del número de inmigrantes extranjeros por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

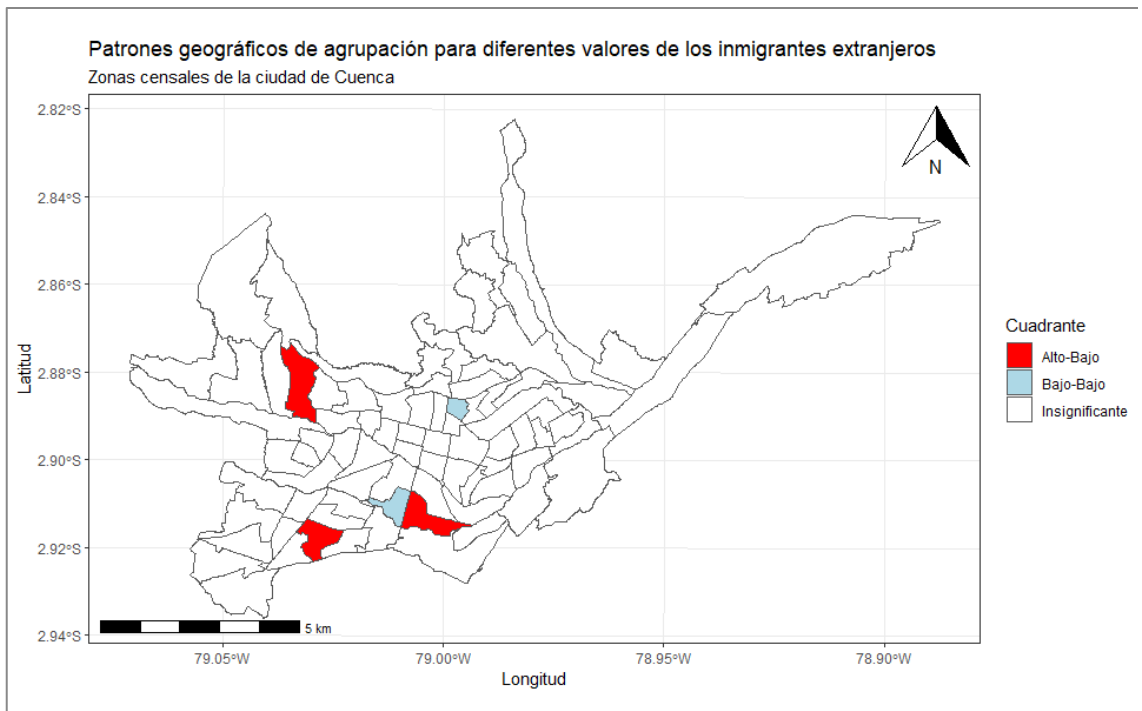


Figura 52-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de desplazamiento temporal de personas por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Número de emigrantes internacional

Análisis de autocorrelación índice *I* de Moran

Tabla 36-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de emigrantes internacionales a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_NEI
## weights: queen_Cuenca_lw
## Moran I statistic standard deviate = 0.80025, p-value = 0.2118
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.042124494      -0.012500000      0.004659364
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

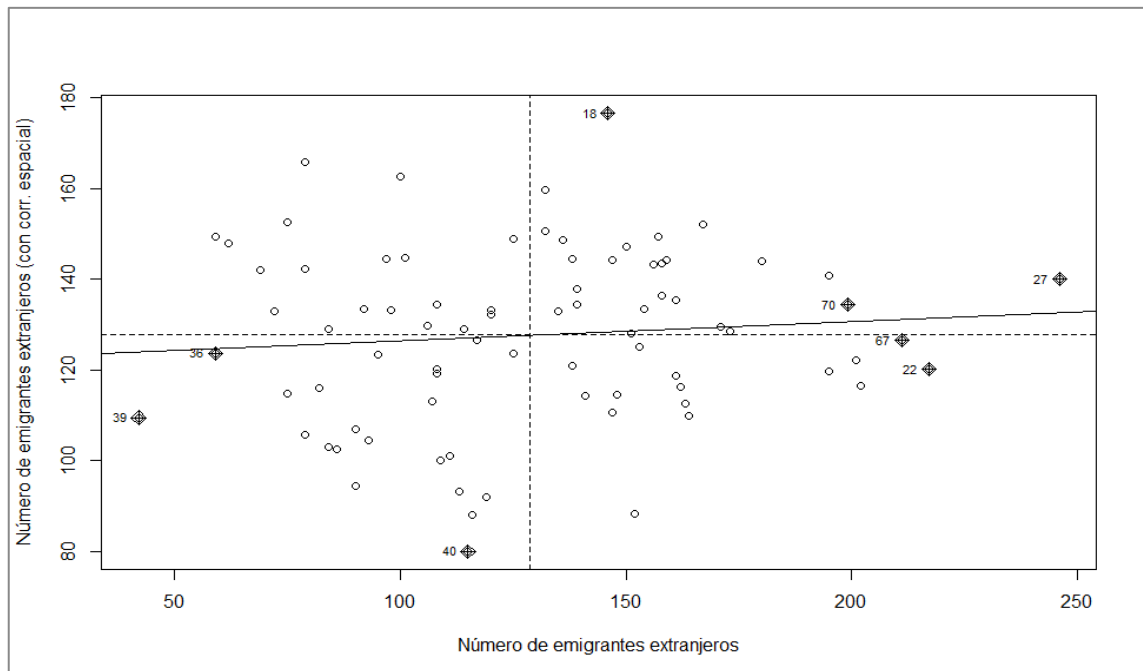


Figura 53-4: Gráfico de correlación del índice *I* de Moran, número de emigrantes internacionales por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

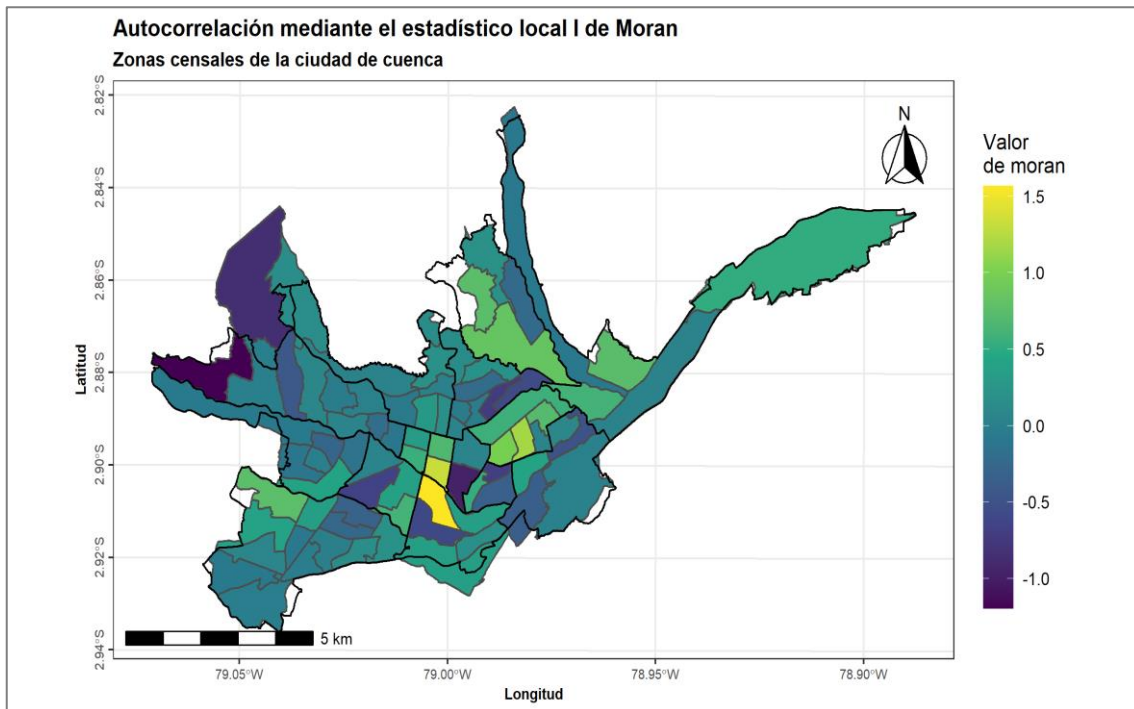


Figura 54-4: Mapa del test de *Moran* a nivel local, para los diferentes valores del número de emigrantes internacionales por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

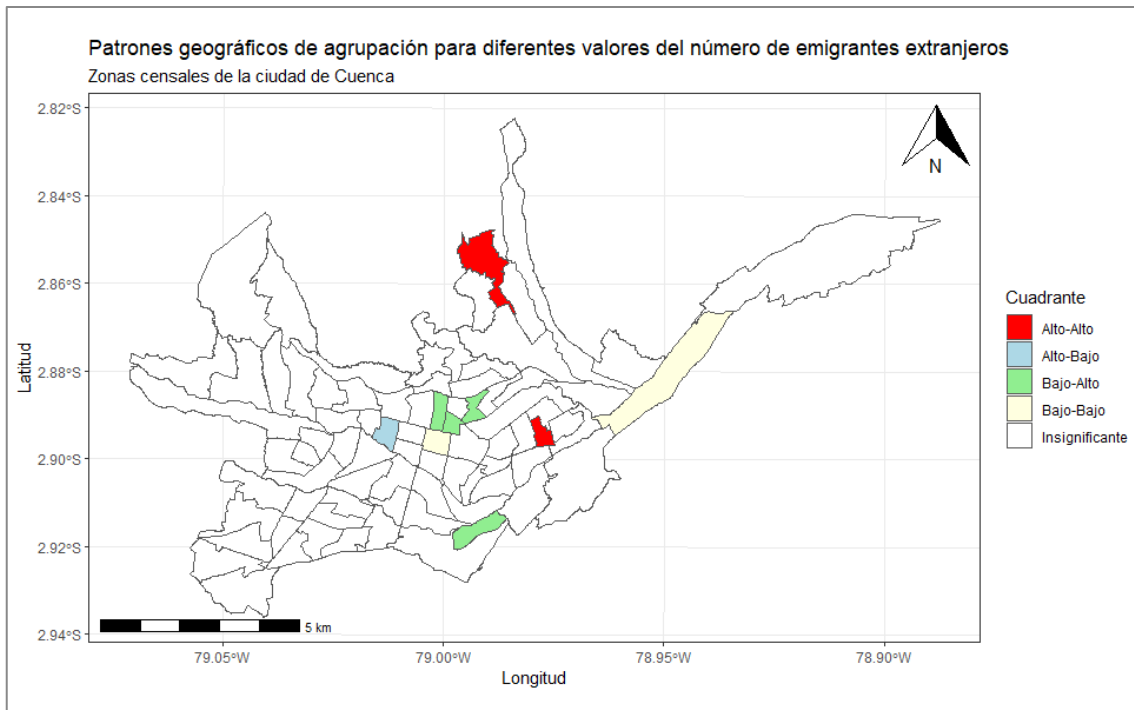


Figura 55-4: Mapa de patrones geográficos de agrupación, para los diferentes valores del número de emigrantes internacionales por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Actividades Económicas

Población en edad de trabajar (PET)

Análisis de autocorrelación índice *I* de Moran

Tabla 37-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de personas en edad de trabajar (PET) a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_PET
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = -0.19551, p-value = 0.5775
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
## -0.025742435      -0.012500000      0.004587762
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

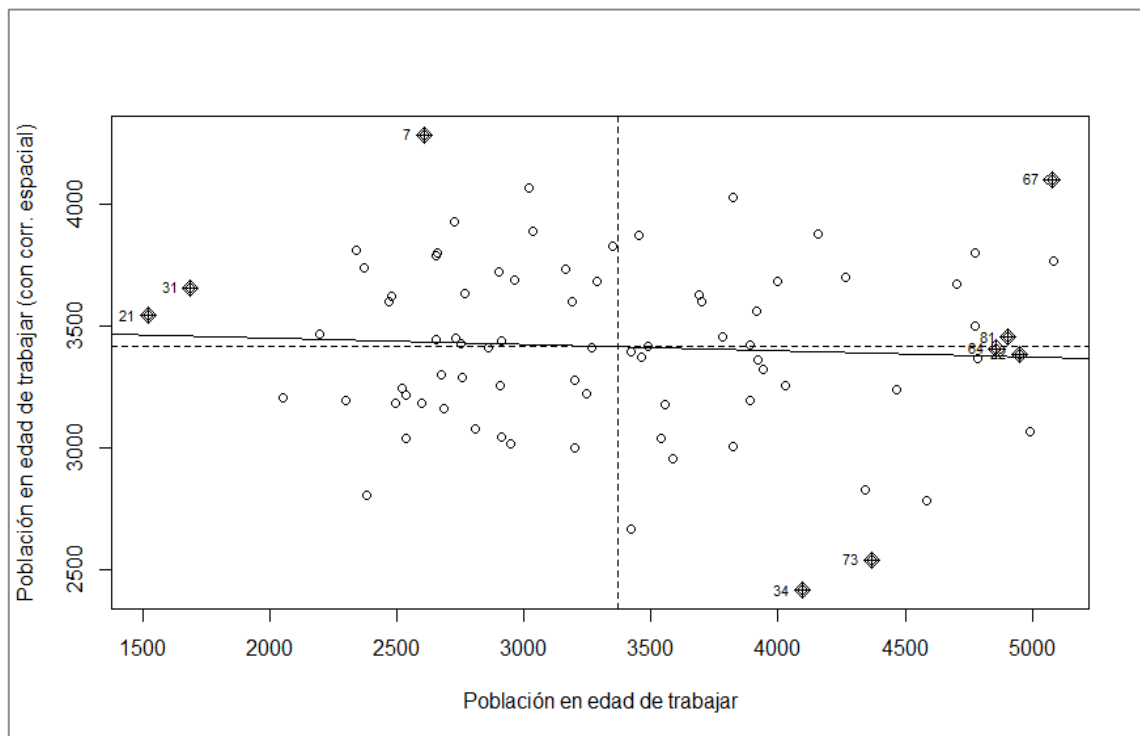


Figura 56-4: Gráfico de correlación del índice *I* de Moran, número de personas en edad de trabajar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

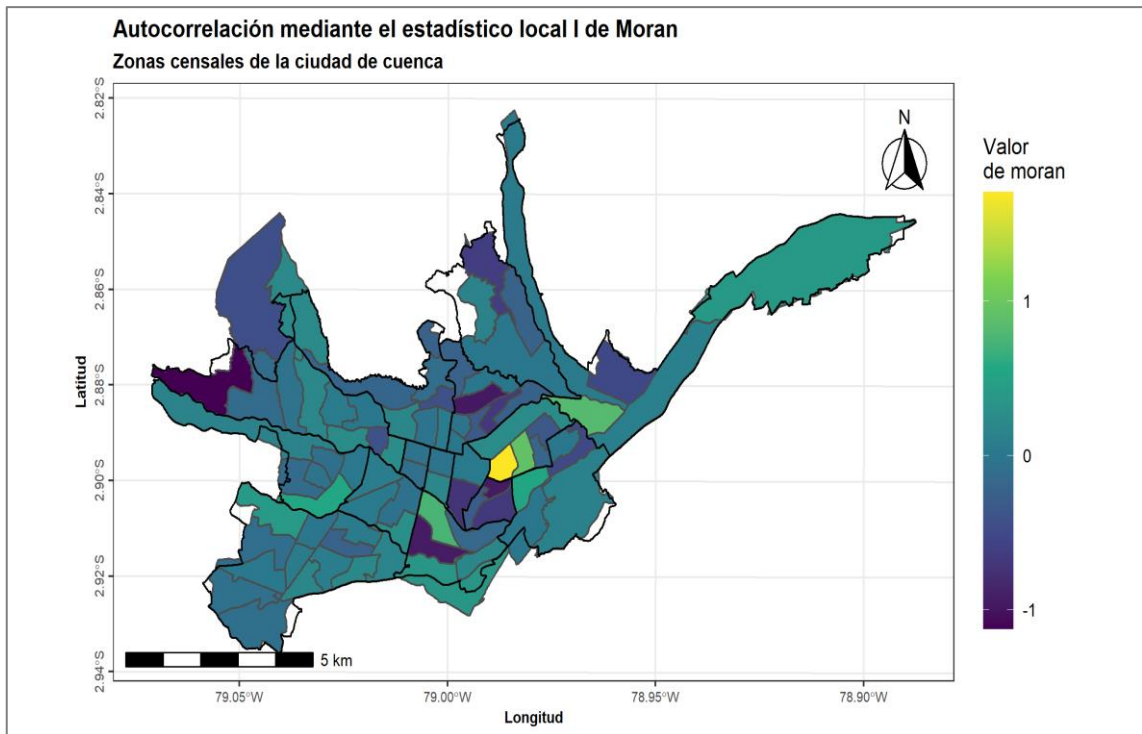


Figura 57-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de población en edad de trabajar por zonas censales urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

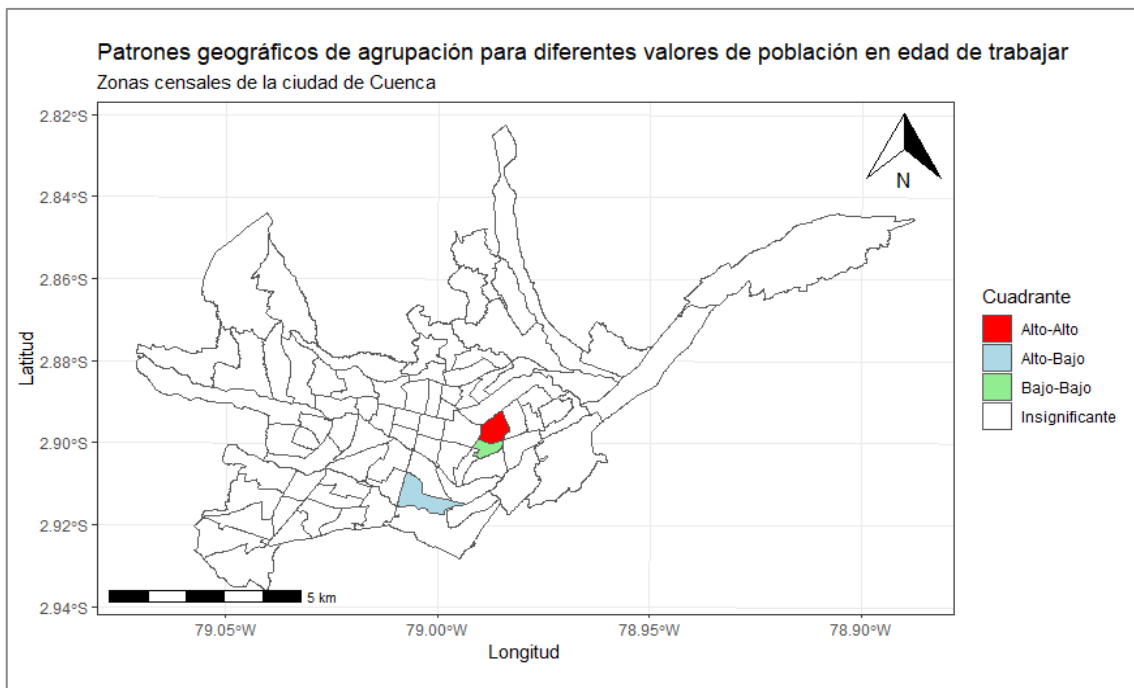


Figura 58-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población en edad de trabajar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Población Cesante

Análisis de autocorrelación índice *I* de Moran

Tabla 38-4: Test de *I* Moran, matriz de pesos (*W*) utilizando simulación de Montecarlo para el número de personas cesantes a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_PCes
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 1.3521, p-value = 0.08818
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.078432650      -0.012500000      0.004523204
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

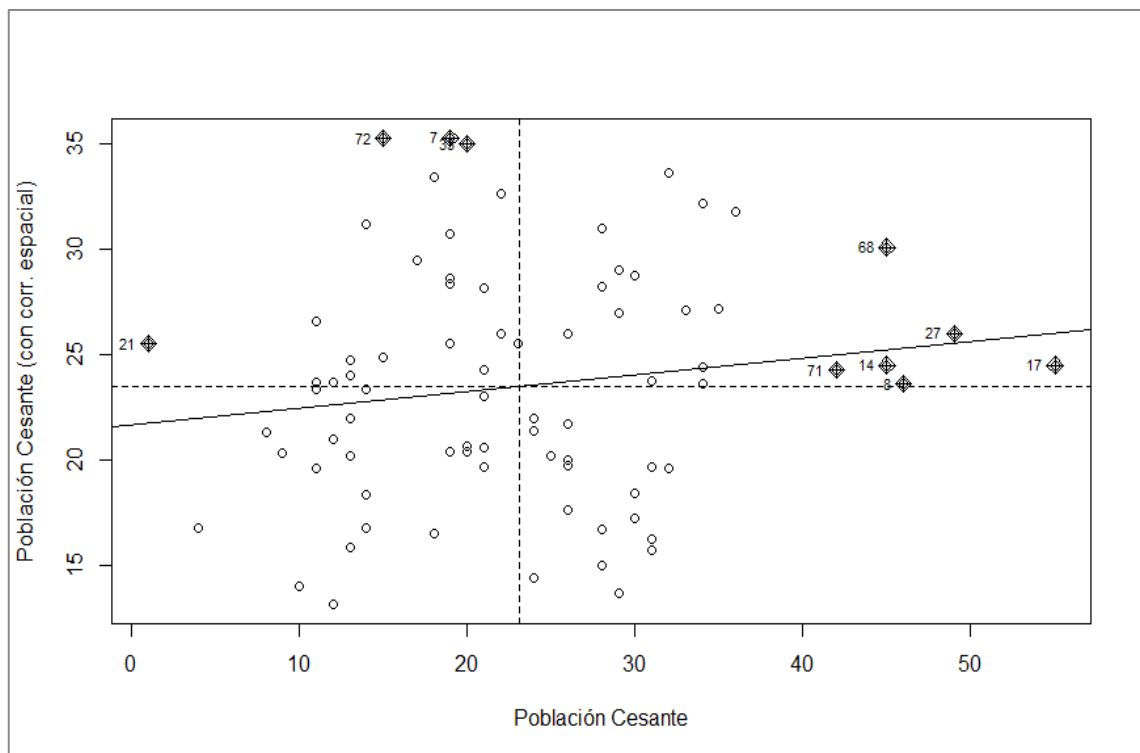


Figura 59-4: Gráfico de correlación del índice *I* de Moran, población cesante por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

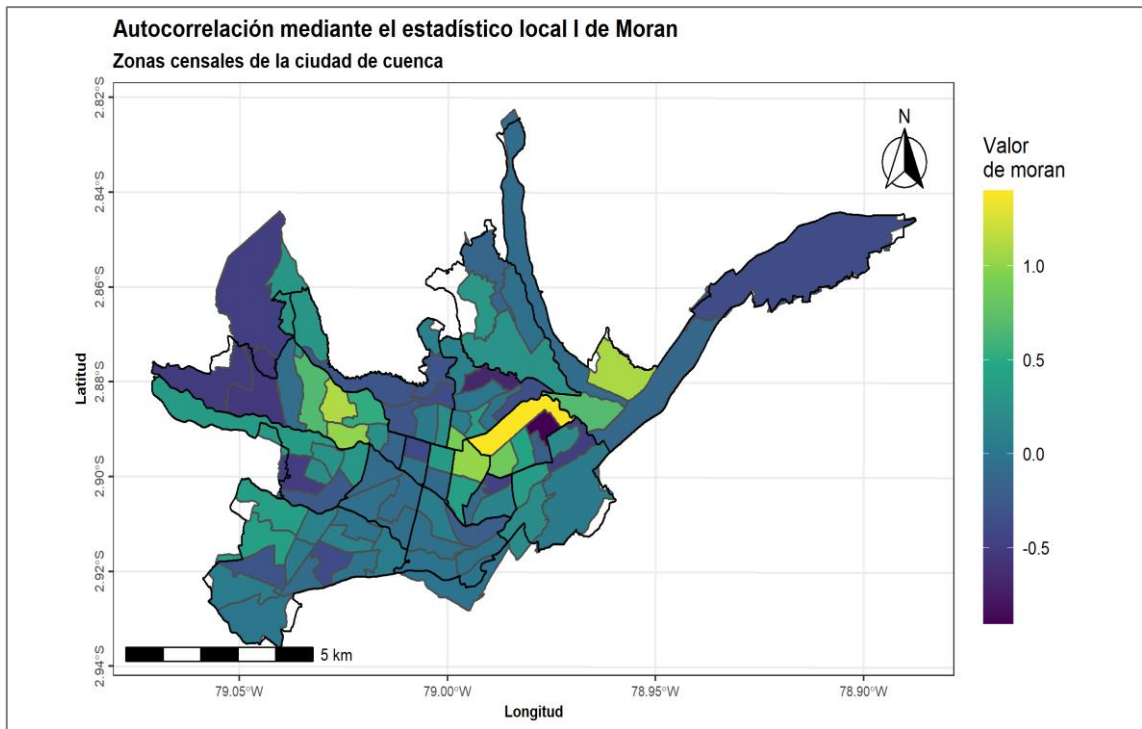


Figura 60-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de población cesante por zonas censales urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

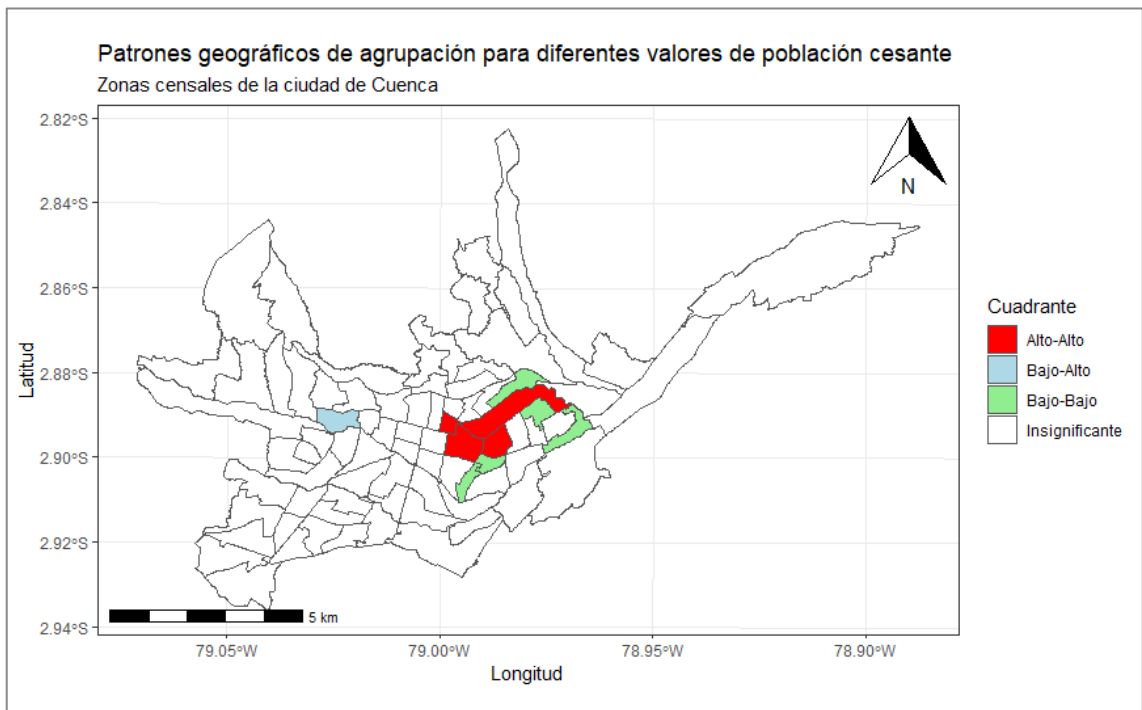


Figura 61-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población cesante por hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Población ocupada que realizó su actividad económica dentro del hogar

Análisis de autocorrelación índice *I* de Moran

Tabla 39-4: Test de I Moran, matriz de pesos (W) simulación de Montecarlo para el número de personas ocupadas que realizó su actividad económica dentro del hogar por zona en Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_POdh
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 2.132, p-value = 0.0165
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.127617615      -0.012500000      0.004319408
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

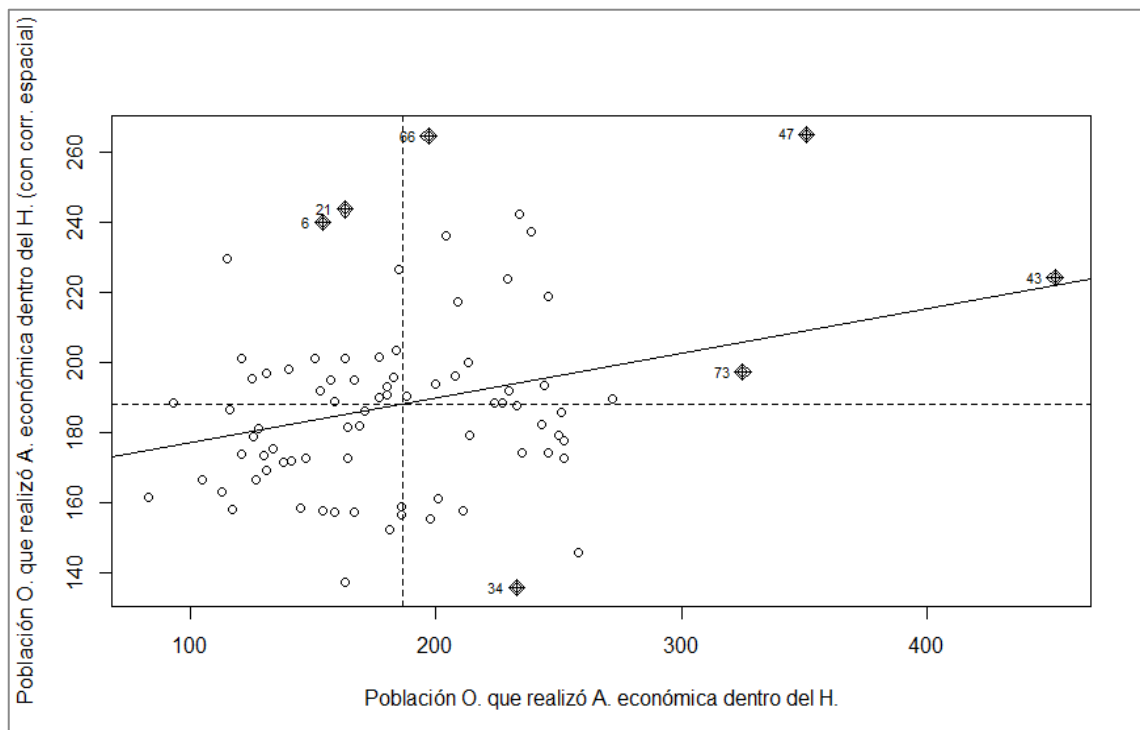


Figura 62-4: Gráfico de correlación del índice *I* de Moran, población ocupada que realizó actividad económica dentro del hogar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

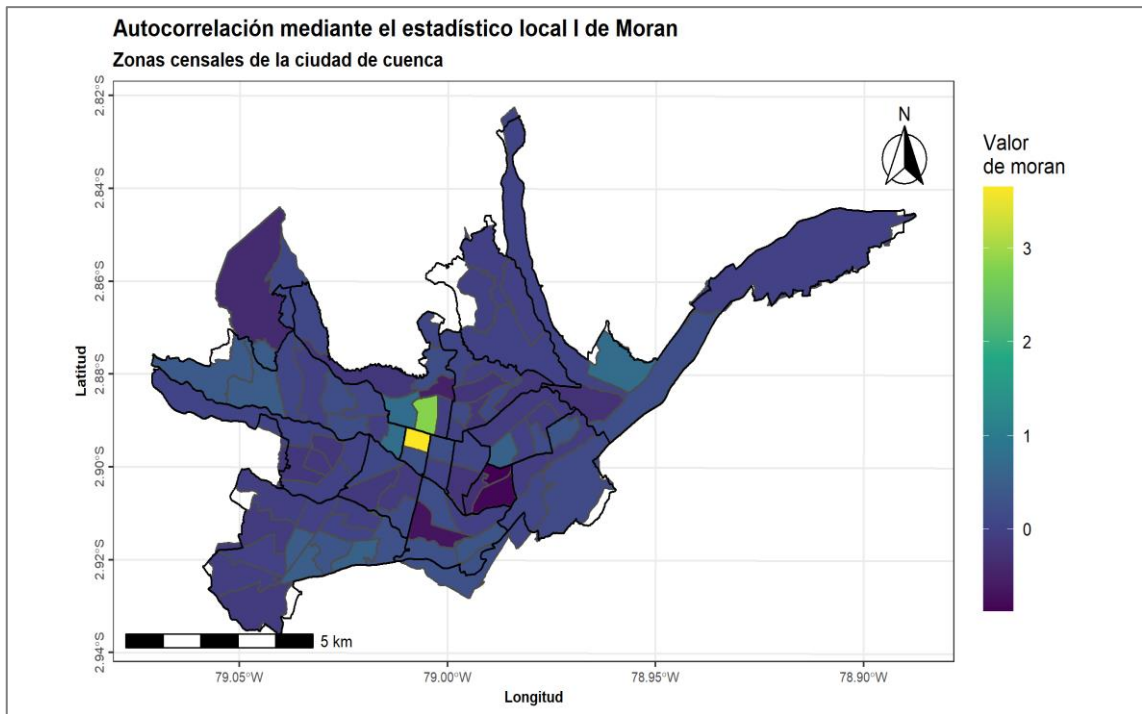


Figura 63-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de población ocupada que realice actividad económica en el hogar por zonas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

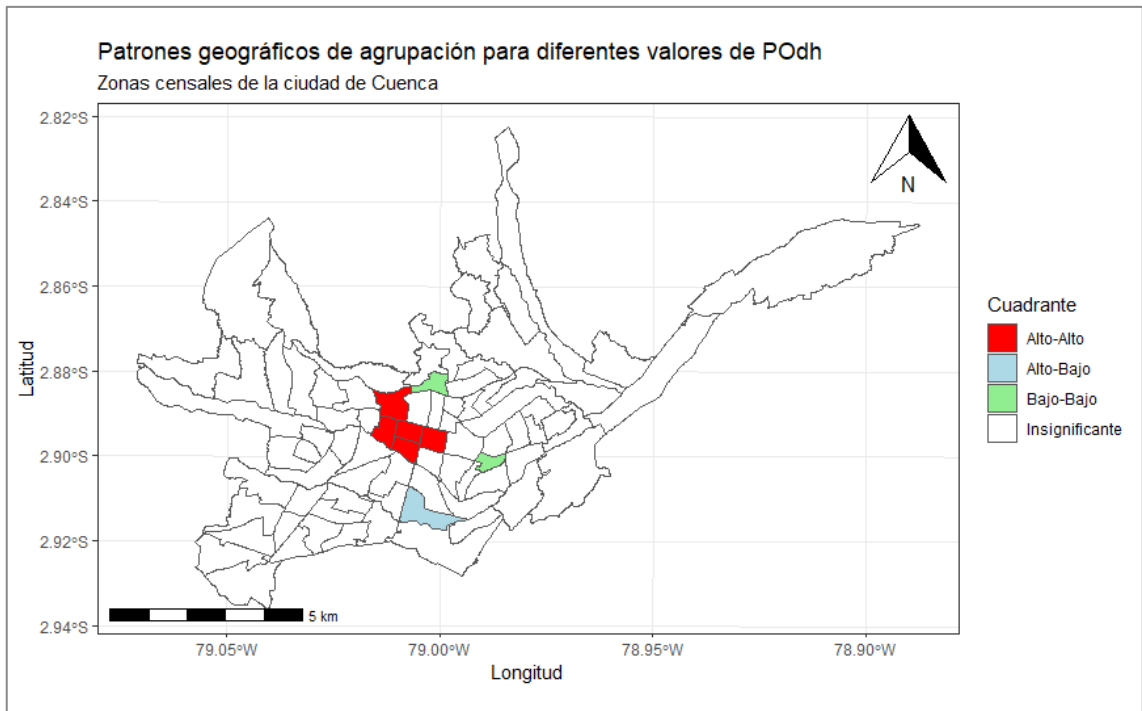


Figura 64-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población ocupada que realizó actividad económica en el hogar por zonas urbanas en Cuenca.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Población Económicamente Inactiva (PEI)

Análisis de autocorrelación índice *I* de Moran

Tabla 40-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de personas económicamente inactiva (PEI) a nivel de zona, ciudad Cuenca, 2019.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_PEI
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 0.74408, p-value = 0.2284
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
## 0.037849574      -0.012500000      0.004578776
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

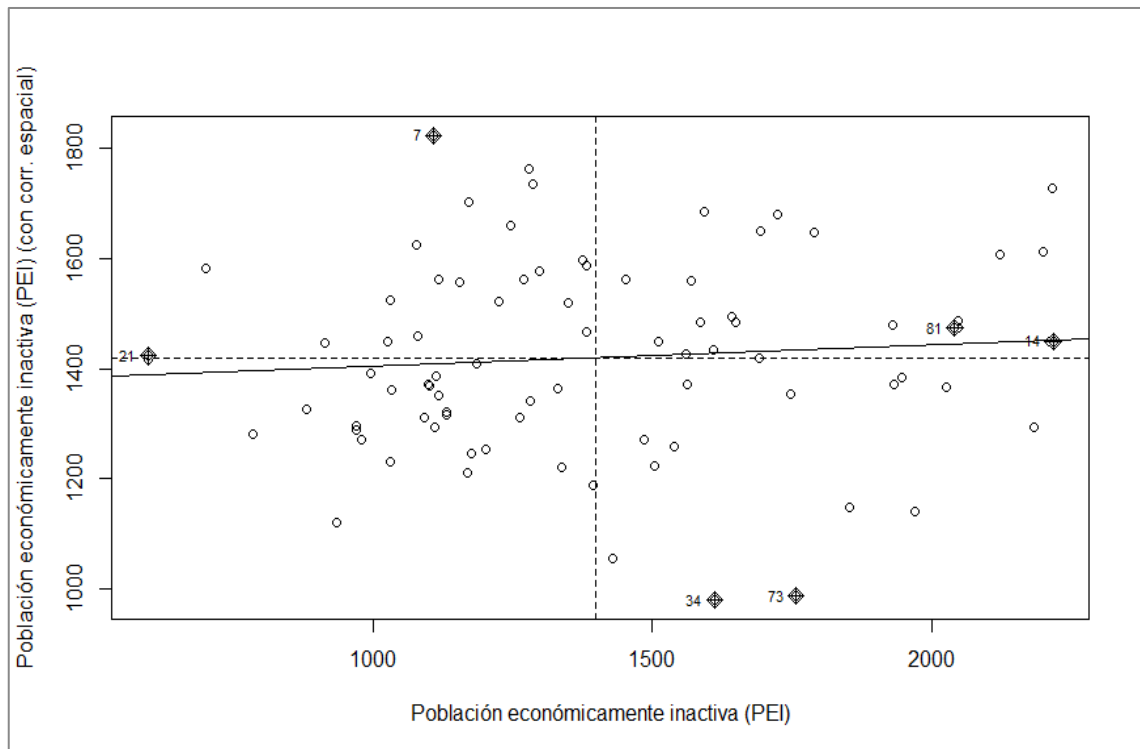


Figura 65-4: Gráfico de correlación del índice *I* de Moran, población económicamente inactiva (PEI) por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

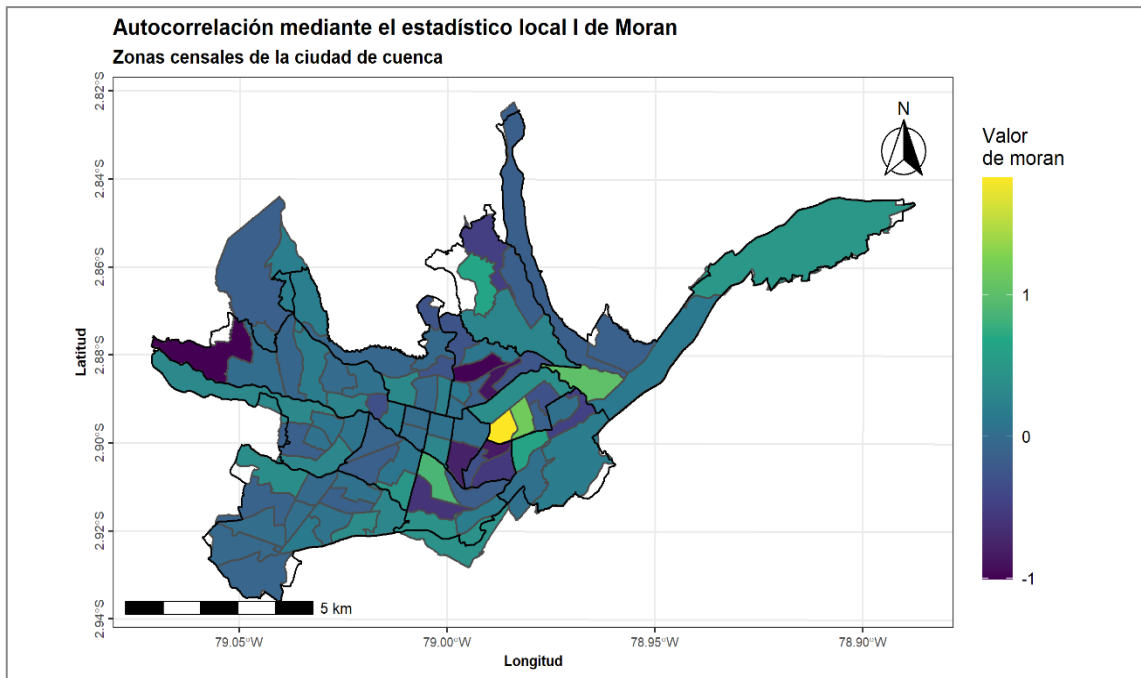


Figura 66-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de población económicamente inactiva por zonas censales urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

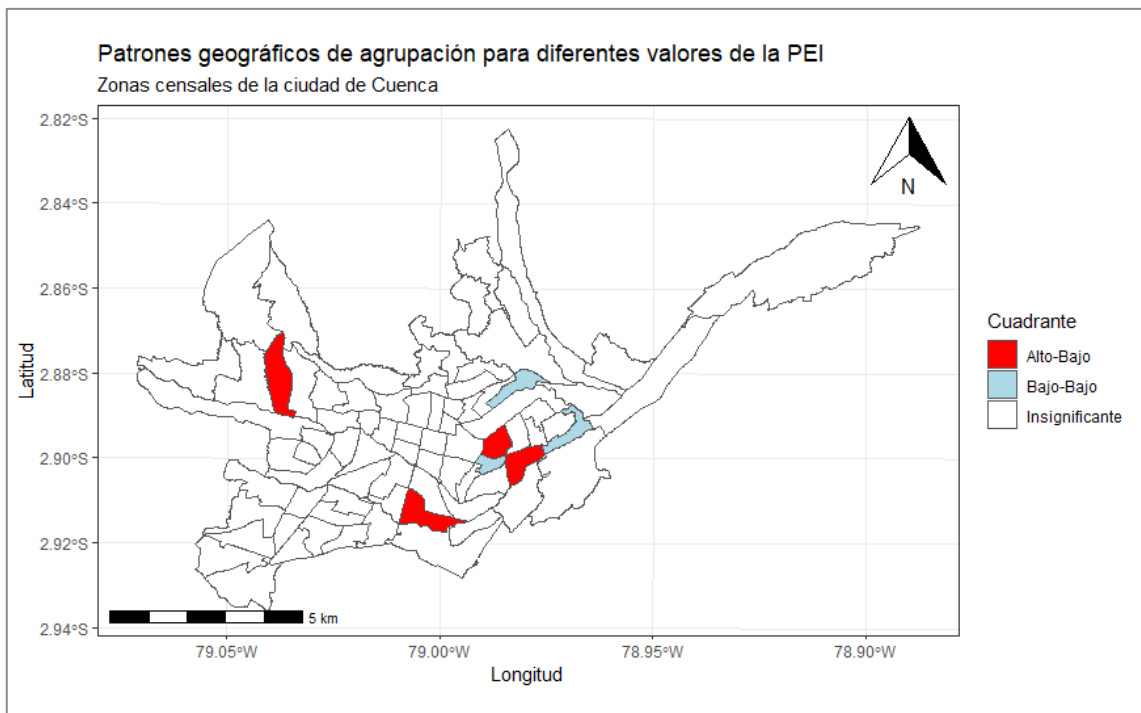


Figura 67-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de población económicamente inactiva (PEI) por zonas urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Educación

Tasa de Analfabetismo

Análisis de autocorrelación índice *I* de Moran

Tabla 41-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa de analfabetismo a nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca$bd_TA_10YM
## weights: queen_Cuenca_lw
## Moran I statistic standard deviate = 5.1247, p-value = 1.49e-07
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
##      0.330168190      -0.012500000      0.004471002
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

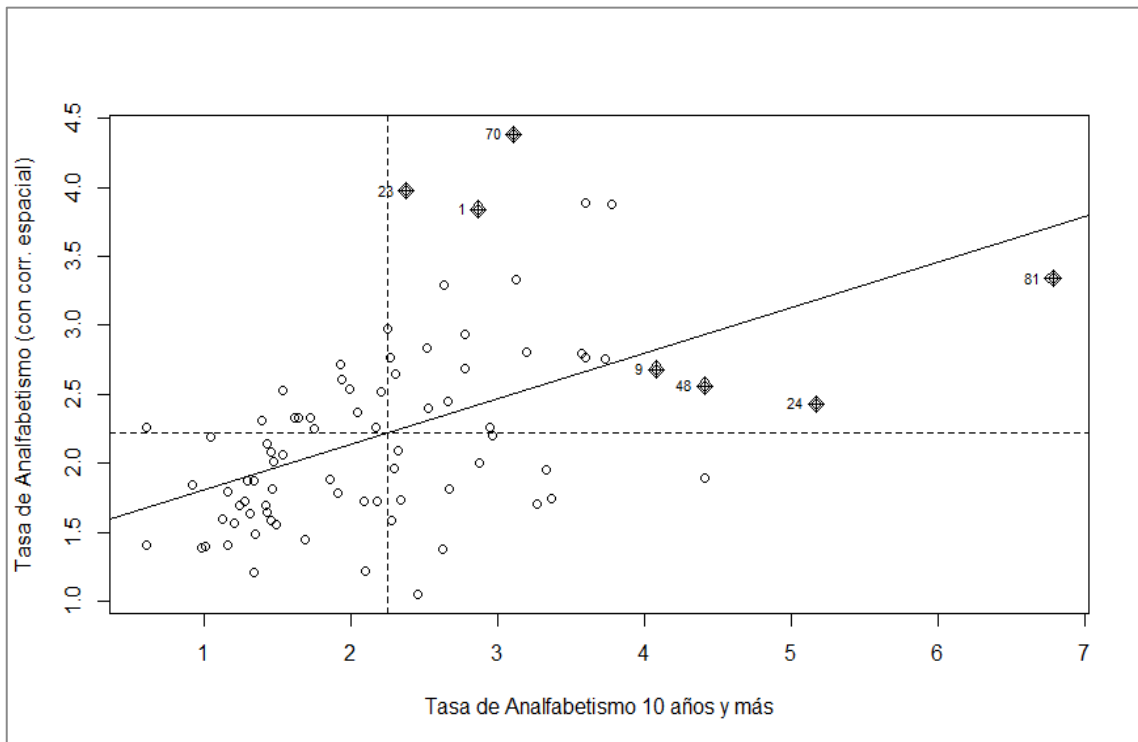


Figura 68-4: Gráfico de correlación del índice *I* de Moran, tasa de analfabetismo por zonas censales urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

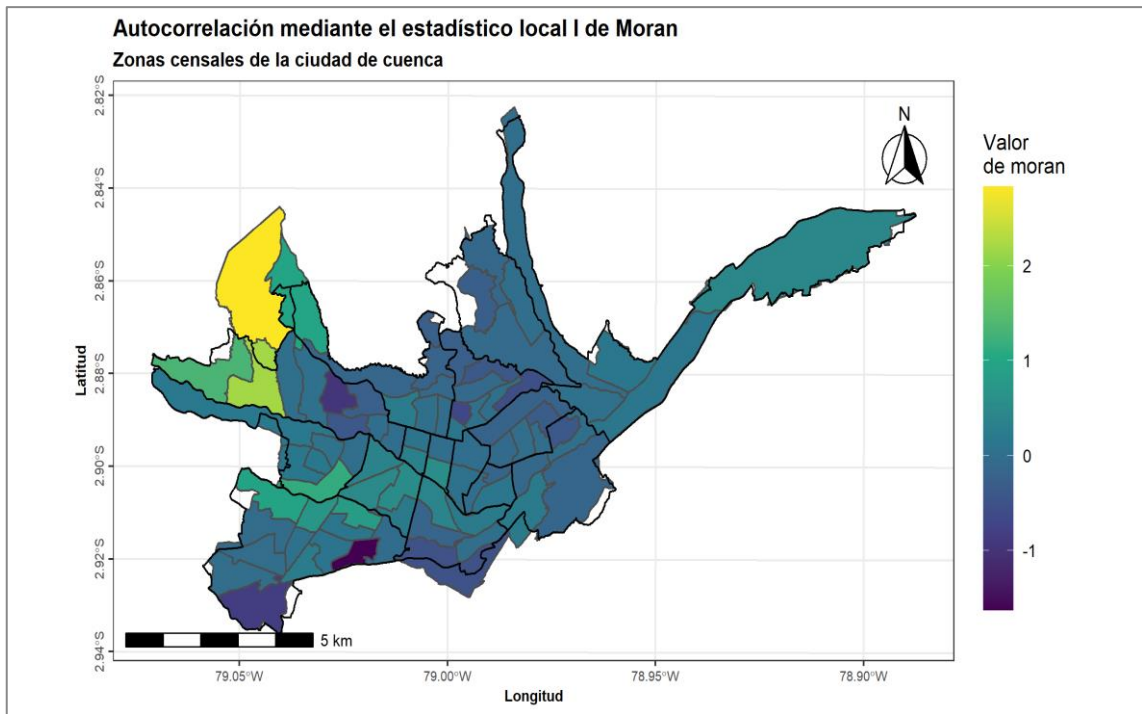


Figura 69-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de la tasa de analfabetismo por zonas censales urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

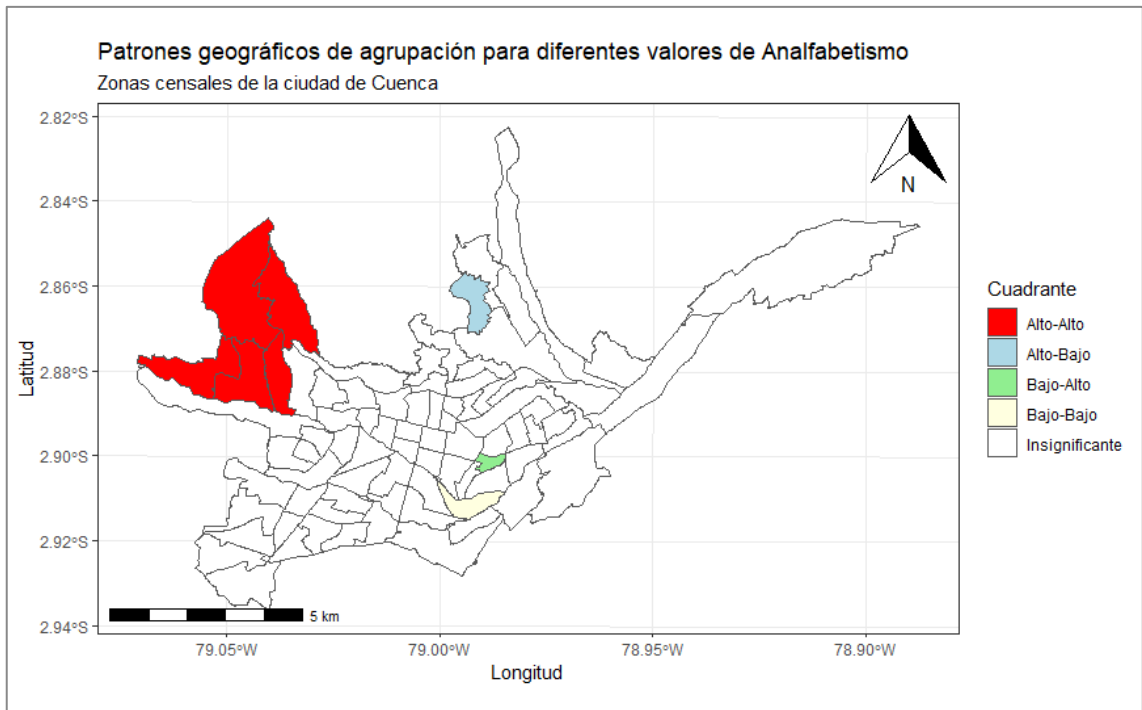


Figura 70-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de analfabetismo por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Tasa de tasa neta de asistencia escolar

Análisis de autocorrelación índice *I* de Moran

Tabla 42-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para la tasa neta de asistencia escolar nivel de zona censal, ciudad Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_Tasa_Na
## weights: queen_Cuenca_lw
## Moran I statistic standard deviate = -1.1341, p-value = 0.8716
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
## -0.089220213          -0.012500000      0.004576187
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

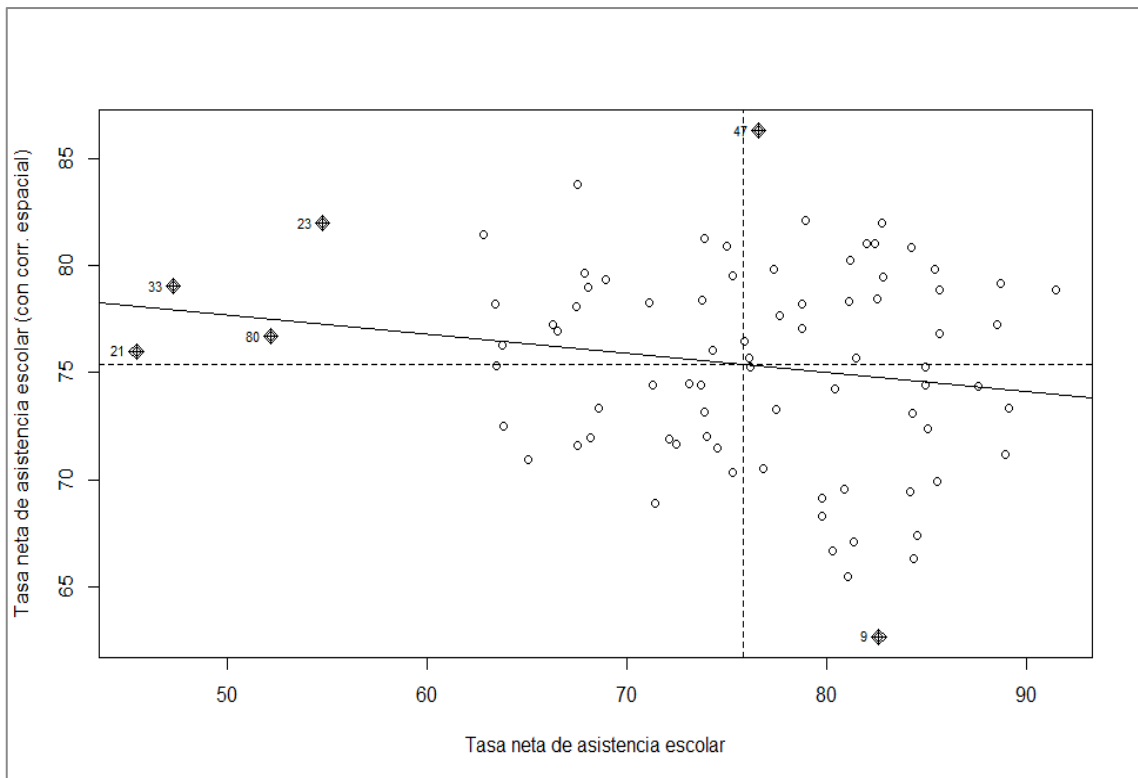


Figura 71-4: Gráfico de correlación del índice *I* de Moran, tasa neta de asistencia escolar por zonas censales urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.
Realizado por: Velasteguí Casco, Edy, 2022.

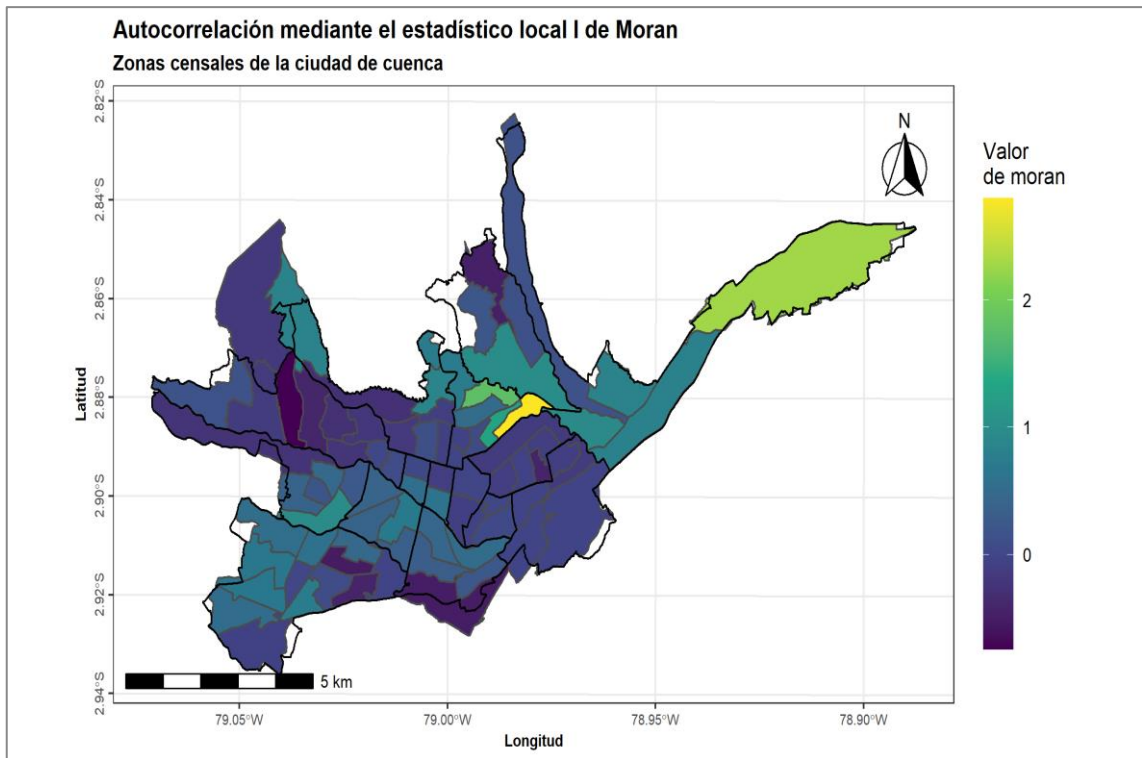


Figura 72-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de la tasa neta de asistencia escolar por zonas censales urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

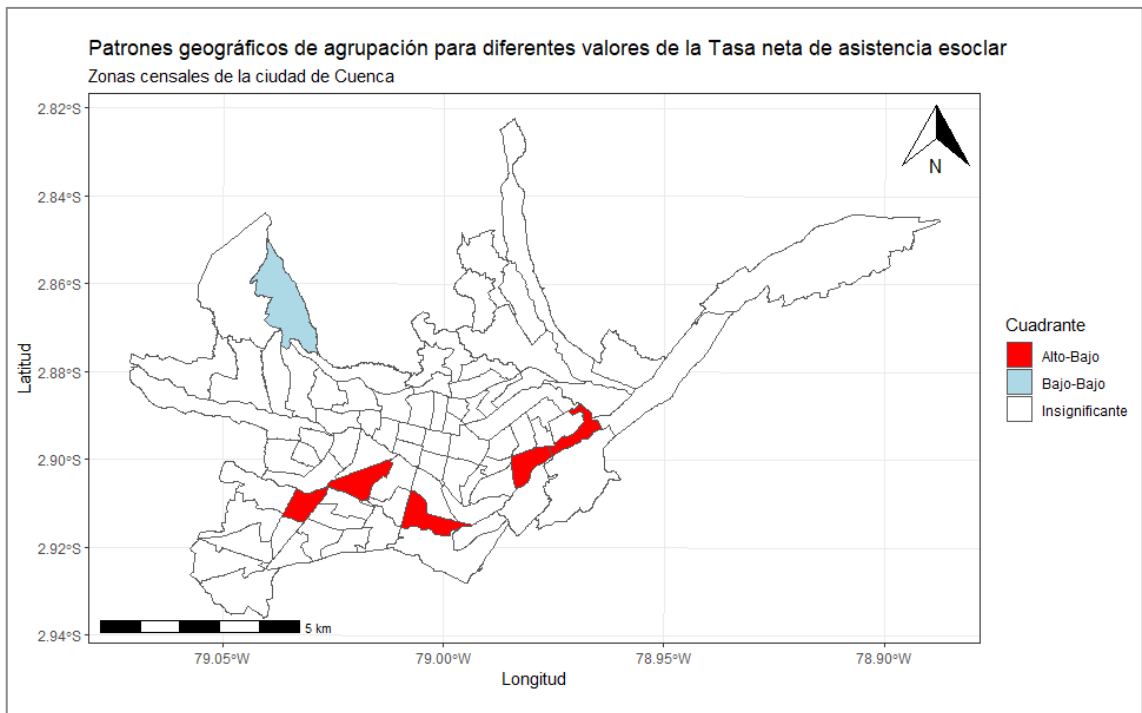


Figura 73-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa neta de asistencia escolar por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

Vivienda y servicios

Tenencia de la vivienda de los hogares.

Análisis de autocorrelación índice *I* de Moran

Tabla 43-4: Test de I Moran, matriz de pesos (W) utilizando simulación de Montecarlo para el número de personas que tienen vivienda propia y pagada a nivel de zona en Cuenca.

```
##                               Moran I test under randomisation
## data:  shp_Cuenca1$BD_HTV
## weights: queen_Cuenca1_lw
## Moran I statistic standard deviate = 2.5606, p-value = 0.005225
## alternative hypothesis: greater
##                               sample estimates:
## Moran I statistic      Expectation      Variance
## 0.159294981      -0.012500000      0.004501326
```

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

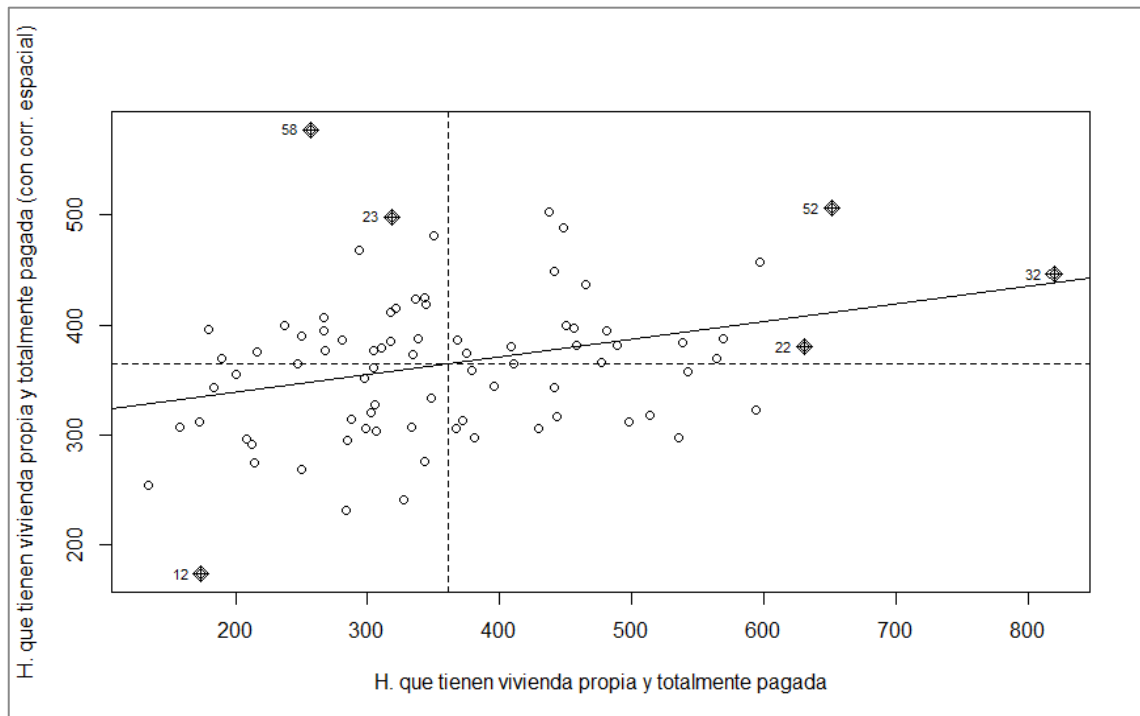


Figura 74-4: Gráfico de correlación del índice *I* de Moran, hogares que tienen vivienda propia y totalmente pagada por zonas censales urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

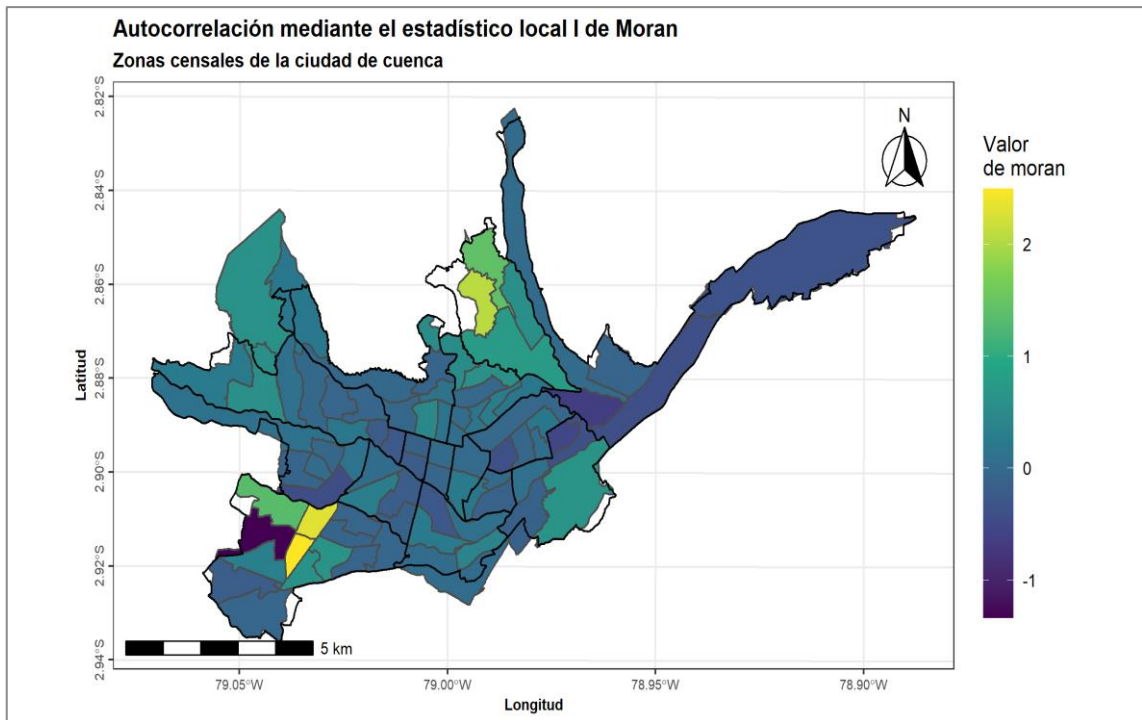


Figura 75-4: Mapa del test de *Moran* a nivel local, para los diferentes valores de hogares que tienen vivienda propia y totalmente pagada por zonas urbanas de la ciudad de Cuenca, 2010.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

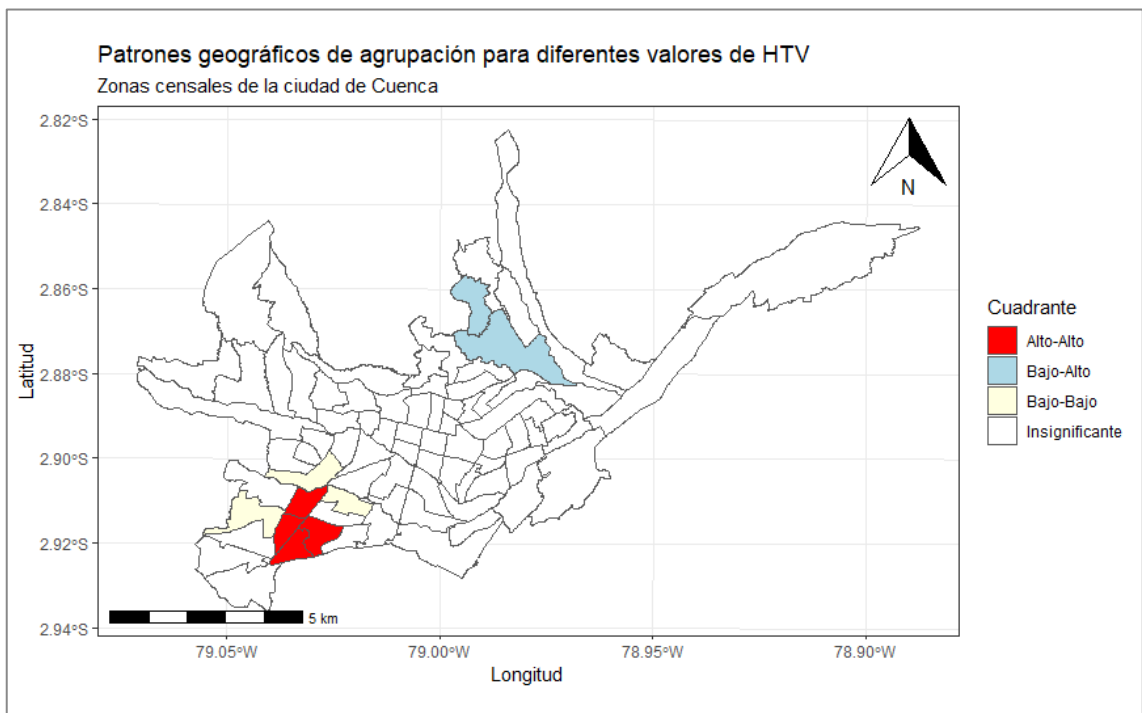


Figura 76-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de hogares que tienen vivienda propia y pagada por zonas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC.

Realizado por: Velasteguí Casco, Edy, 2022.

4.3.3 Identificación de las zonas urbanas de alta incidencia delictiva de robos explicadas mediante las variables sociodemográficas.

A continuación, mediante el *I* de Moran a nivel local (*LISA*) y los mapas respectivos, se muestran las zonas urbanas de alta incidencia en la ciudad de Cuenca, una vez que se ha identificado la autocorrelación existente entre las variables dependientes de los robos a personas, domicilios, unidades económicas y vehículos; con cada una de las variables sociodemográficas analizadas, se determina su influencia y explicación espacial para este fenómeno social durante el año 2019.

Robo a Personas

En la Figura 77-4., se observa que, en nueve zonas urbanas existe una alta concentración de la tasa de robo a personas; explicadas por un alto número de inmigrantes extranjeros (Alto – Alto) en las parroquias de: San Blas, El Sagrario, Sucre, El Vecino y El Batán. Esto quiere decir que se puede tener el mismo comportamiento en cada una de las zonas vecinas más cercanas. Mientras que existen nueve zonas con una baja concentración de la tasa de robo a personas y que sus vecinos más cercanos tienen una alta concentración explicada por el número de inmigrantes internacionales presentes en estas zonas vecinas (Bajo – Alto). Corroborando la identificación de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante los tres modelos aplicados.

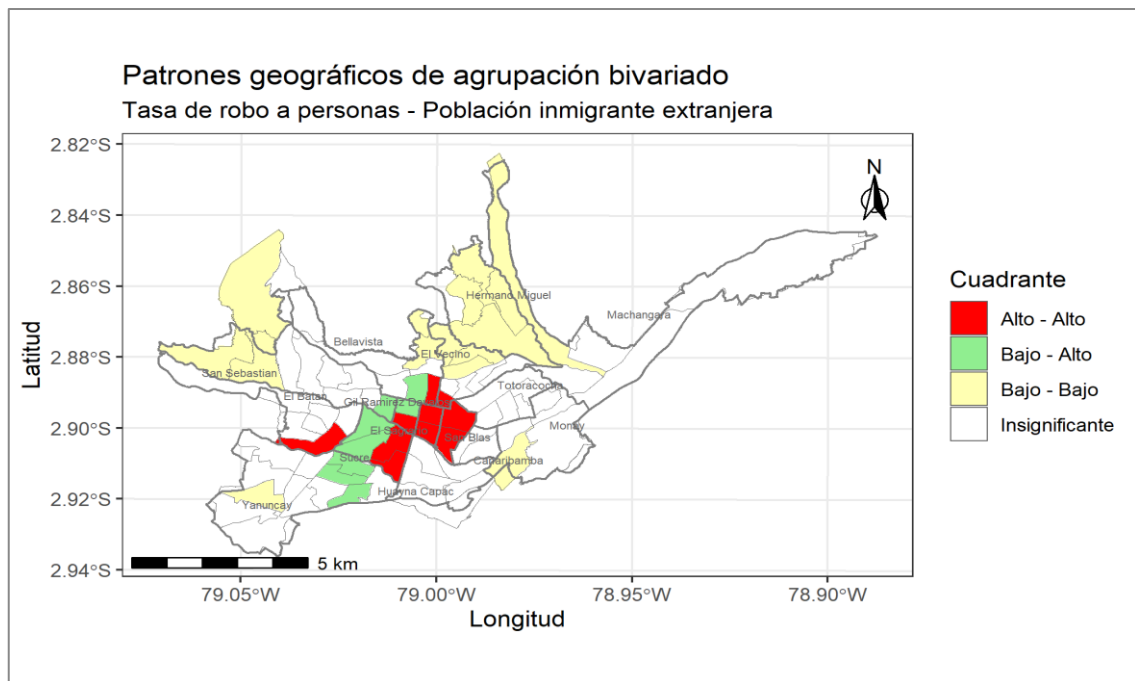


Figura 77-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población inmigrante extranjera por zonas de la ciudad de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 78-4, que en 12 zonas urbanas existe una alta concentración de la tasa de robo a personas explicadas por un alto número de personas promedio por hogar (Alto – Bajo) en las parroquias: San Blas, El Sagrario, Sucre, El Vecino, Cañaribamba, Gil Ramírez Dávalos y Bellavista, mientras que el comportamiento de sus zonas vecinas más cercanas tendrá una concentración baja en el delito mencionado, explicado por el promedio de personas por hogar. Sucede lo contrario en 20 zonas urbanas que tienen una concentración (Bajo – Alto), lo cual indica que estas zonas urbanas tienen una baja concentración de la tasa de robo a personas; las mismas que están explicadas por un bajo promedio de personas por hogar en las parroquias: Yanuncay, San Sebastián, Machángara, Monay. Hermano Miguel, El Vecino y Bellavista, las mismas que tienen vecinos que si presentan una alta concentración de la tasa de robo a personas explicadas por el promedio de personas por hogar. Corroborando la identificación de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante los tres modelos aplicados.

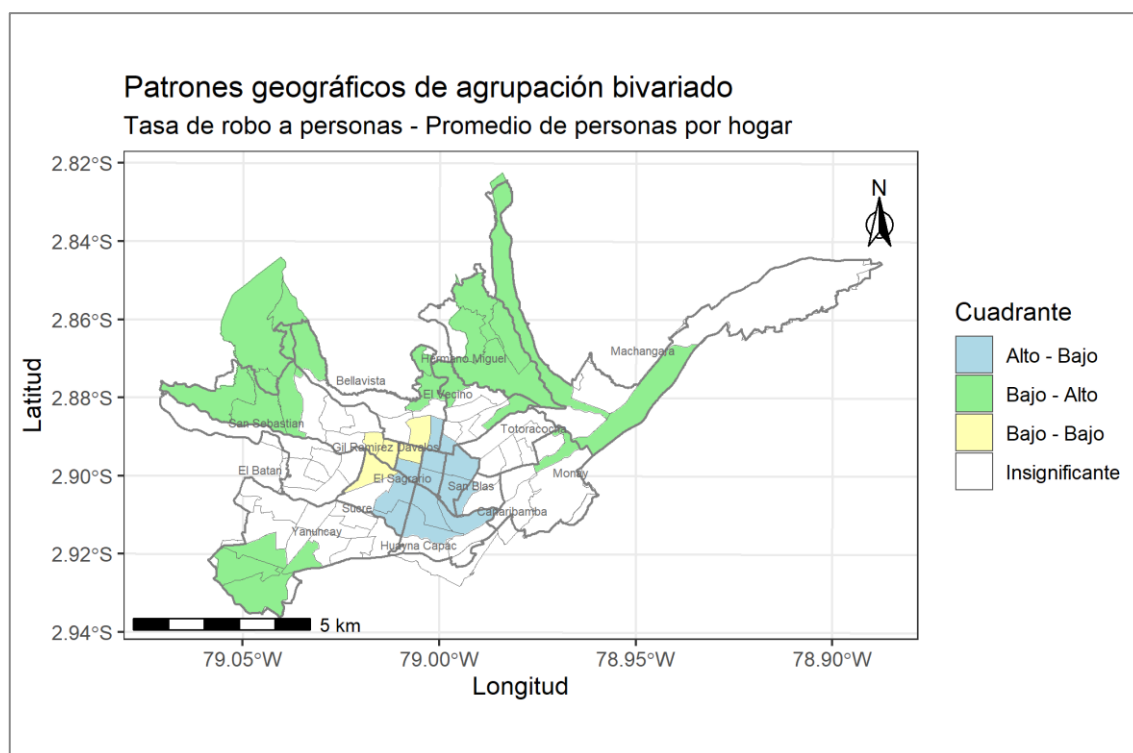


Figura 78-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs promedio de personas por hogar, zonas urbanas de la ciudad de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 79-4, que en una zona urbana existe una alta concentración de la tasa de robo a personas, explicadas por un alto número de personas que hace una semana dejaron de trabajar y se encuentran buscando empleo (Alto – Alto) en las parroquias: San Blas, El Sagrario, El Vecino, Cañaribamba, Totoracocha y Huayna Cápac. Esto quiere decir que el mismo comportamiento se puede presentar en zonas vecinas más cercanas. Mientras que

también existen dos zonas con una alta concentración de la tasa de robo a personas; y que sus vecinos más cercanos presentan una baja concentración explicada por el número de personas cesantes (Alto – Bajo), presentes en estas zonas vecinas. También se observa que existen ocho zonas urbanas que tienen una baja concentración de robo a personas con una baja concentración de personas cesantes y que el efecto de alta incidencia de sus zonas vecinas más cercanas es alto (Bajo –Alto). Corroborando la identificación de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019.

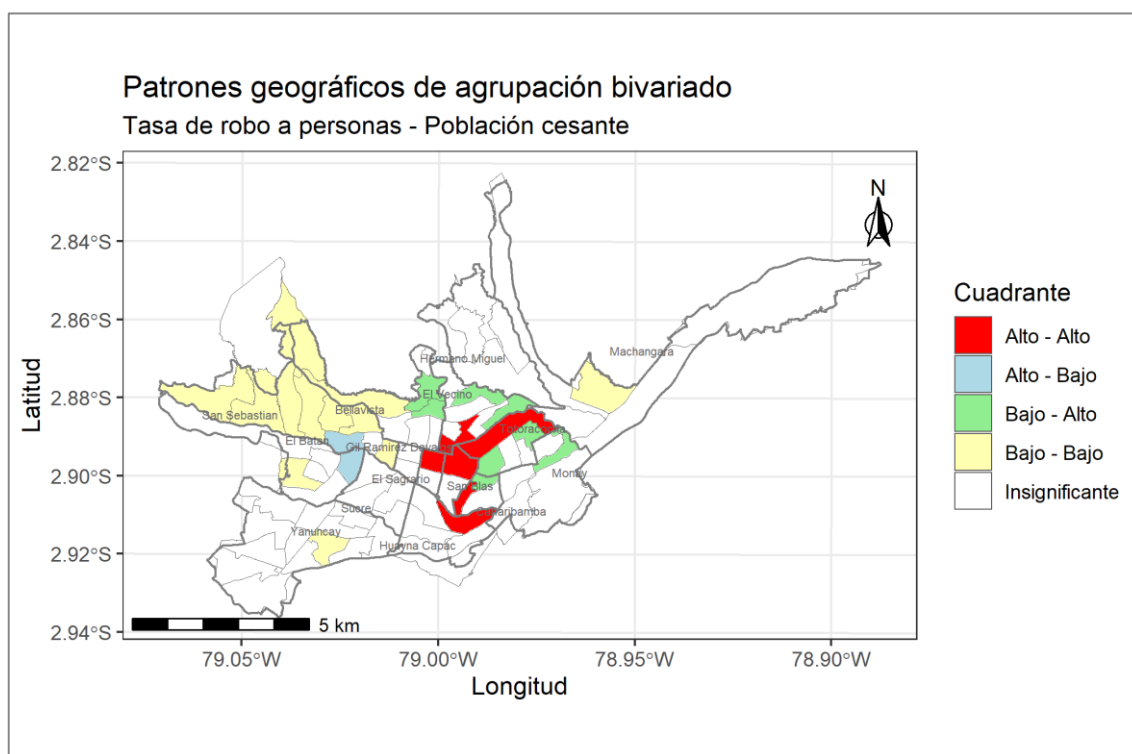


Figura 79-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas vs población cesante, zonas urbanas de la ciudad de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 80-4, que en una zona urbana existe una alta concentración de la tasa de robo a personas, explicadas por un alto número de hogares que tienen una vivienda propia y totalmente pagada (Alto – Alto) en la parroquia El Batán. Esto indica que el mismo comportamiento se puede presentar en cada una de las zonas vecinas más cercanas. También existen nueve zonas con una baja concentración de la tasa de robo a personas y sus vecinos más cercanos tienen una alta concentración; explicada por el número hogares que tienen una vivienda propia y totalmente pagada presentes en estas zonas vecinas. Corroborando la identificación de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019.

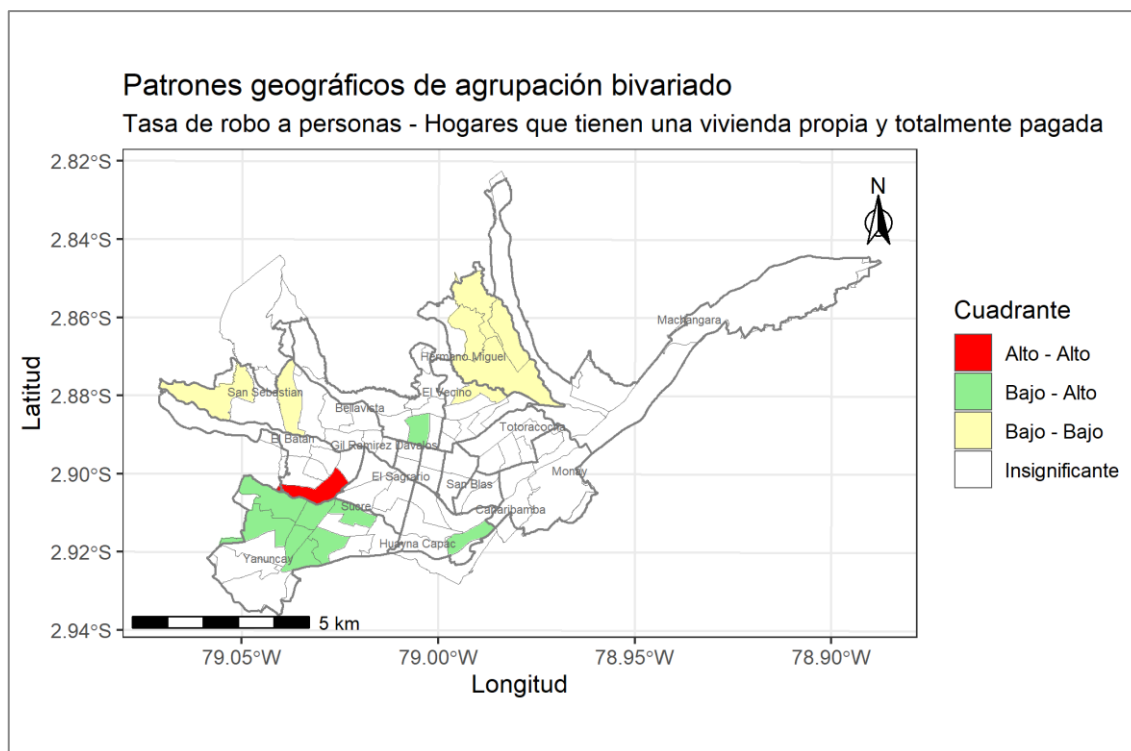


Figura 80-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs hogares que tienen vivienda propia y pagada por zonas de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a domicilios:

En Figura 81-4, se puede observar que en 14 zonas urbanas existe una alta concentración de la tasa de robo a personas explicadas por un alto número de migrantes internacionales (Alto – Alto); en las parroquias: Yanuncay, Sucre, El Sagrario y Huayna Cápac. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen 11 zonas con una alta concentración de la tasa de robo a personas y que sus vecinos más cercanos tienen una baja concentración explicada por el número de migrantes internacionales presentes en estas zonas vecinas. También se tiene 11 zonas urbanas que tienen una baja concentración de la tasa de robos explicadas por el número de migrantes internacionales (Alto – Bajo) y sus vecinos en cambio tienen una alta concentración; corroborando del delito mencionado anteriormente explicado por el número de migrantes internacionales (Bajo – Alto) la identificación de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019.

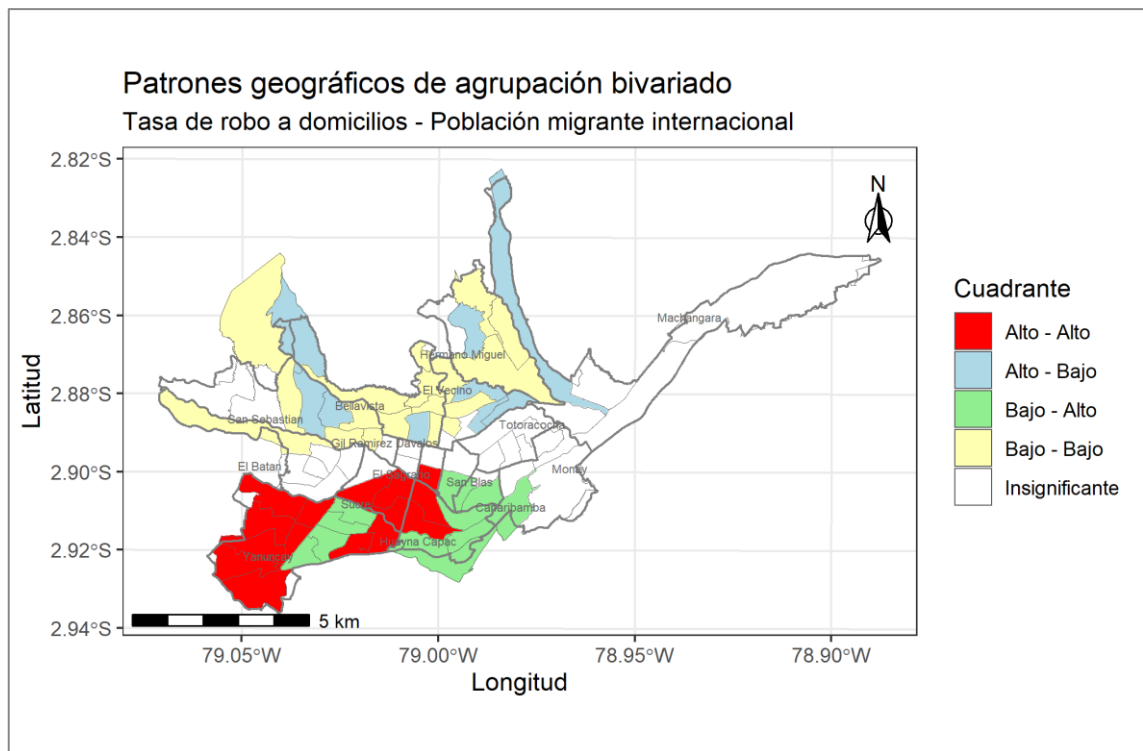


Figura 81-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs población inmigrante a escala internacional por zonas de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 82-4, que en cuatro zonas urbanas existe una alta concentración de la tasa de robo a personas explicadas por una alta tasa de inmigrantes extranjeros (Alto – Alto) en las parroquias Monay y Cañaribamba. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen cinco zonas con una alta concentración de la tasa de robo a personas y que sus vecinos más cercanos tienen una alta concentración explicada por tasa de inmigrantes extranjeros presentes en estas zonas vecinas y esto es corroborado con el cuadrante (Bajo – Alto), identificando las zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante los modelos espaciales aplicados anteriormente.

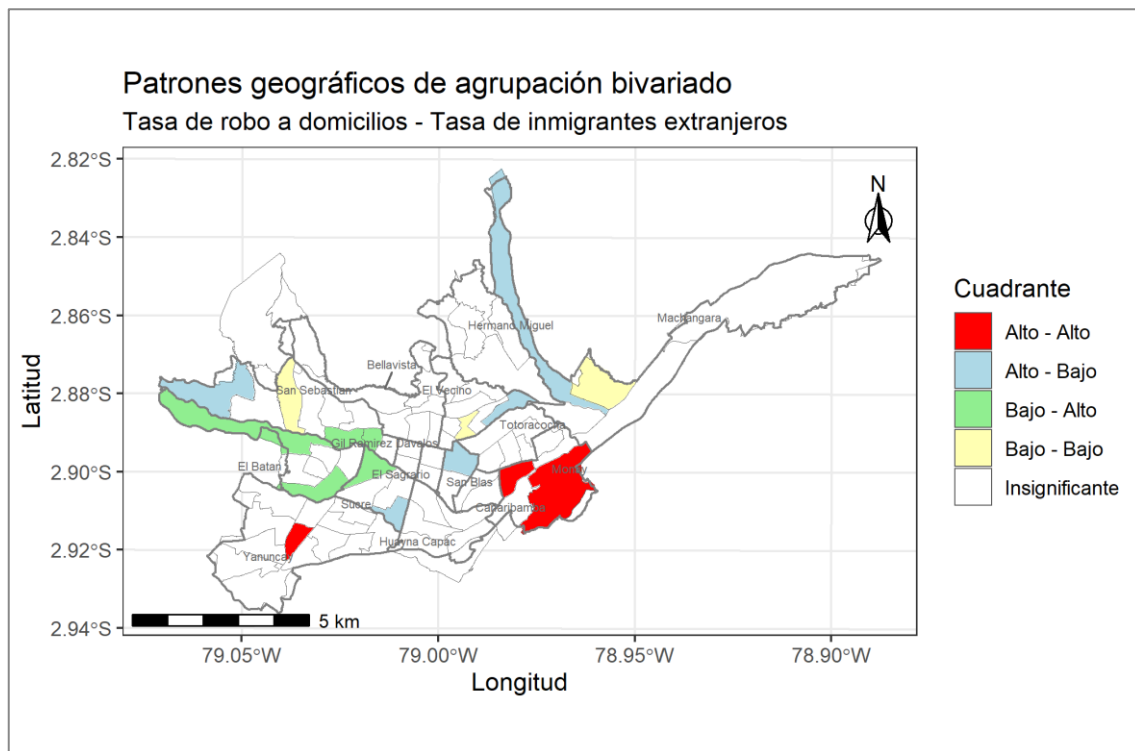


Figura 82-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs tasa de inmigrantes internacionales por zonas de la ciudad de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Al observar la Figura 83-4, se puede identificar que en cuatro zonas urbanas existe una alta concentración de la tasa de robo a personas explicadas por un alto número de población cesante (Alto – Alto) en las parroquias: San Blas, El Sagrario y El Vecino. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que; existen siete zonas con una alta concentración de la tasa de robo a personas y que sus vecinos más cercanos tienen una alta concentración explicada por la población cesante en estas zonas vecinas (Alto – Bajo), y el cuadrante (Bajo – Alto) corrobora este comportamiento, identificando estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante los modelos espaciales aplicados anteriormente.

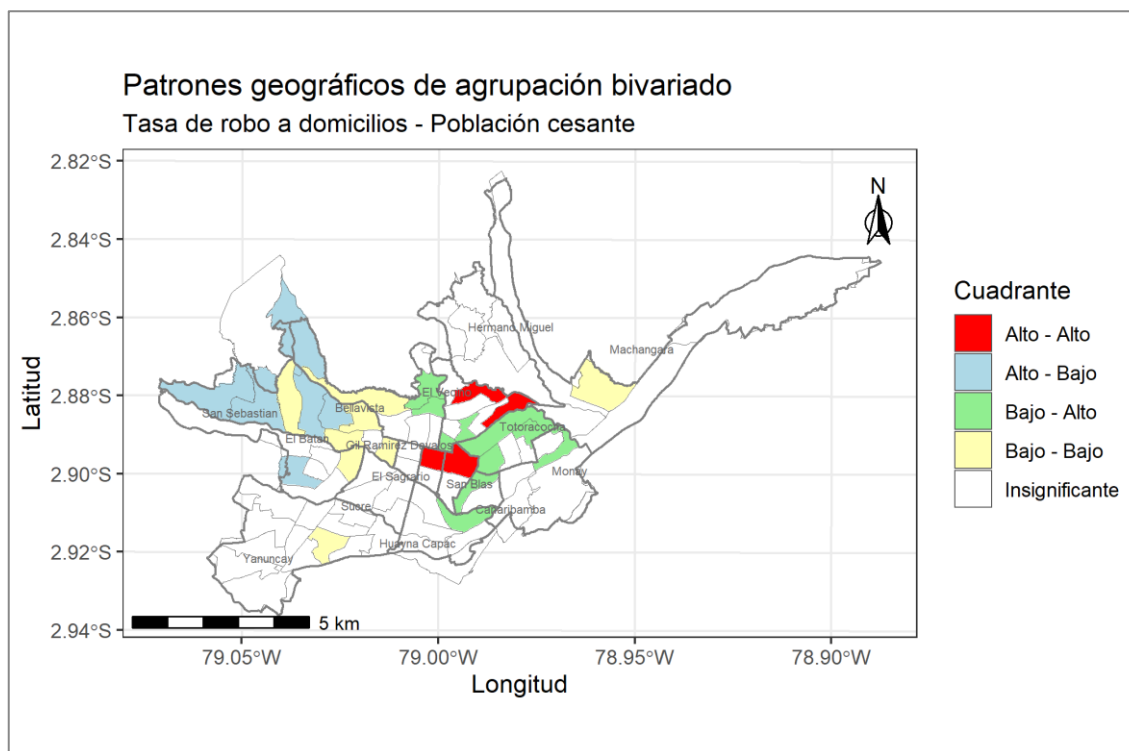


Figura 83-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a domicilios vs población cesante por zonas urbanas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC - Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 84-4, que en cuatro zonas urbanas existe una alta concentración de la tasa de robo a personas explicadas por un alto número de población económicamente inactiva (Alto – Alto) en las parroquias: Yanuncay, El Vecino y Cañaribamba. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Existen también cuatro zonas con una alta concentración de la tasa de robo a personas y que sus vecinos más cercanos tienen una baja concentración explicada por el número de población económicamente inactiva presentes en estas zonas vecinas, mismo que corroborando el cuadrante (Bajo – Alto). La identificación de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 es posible mediante los modelos espaciales aplicados anteriormente.

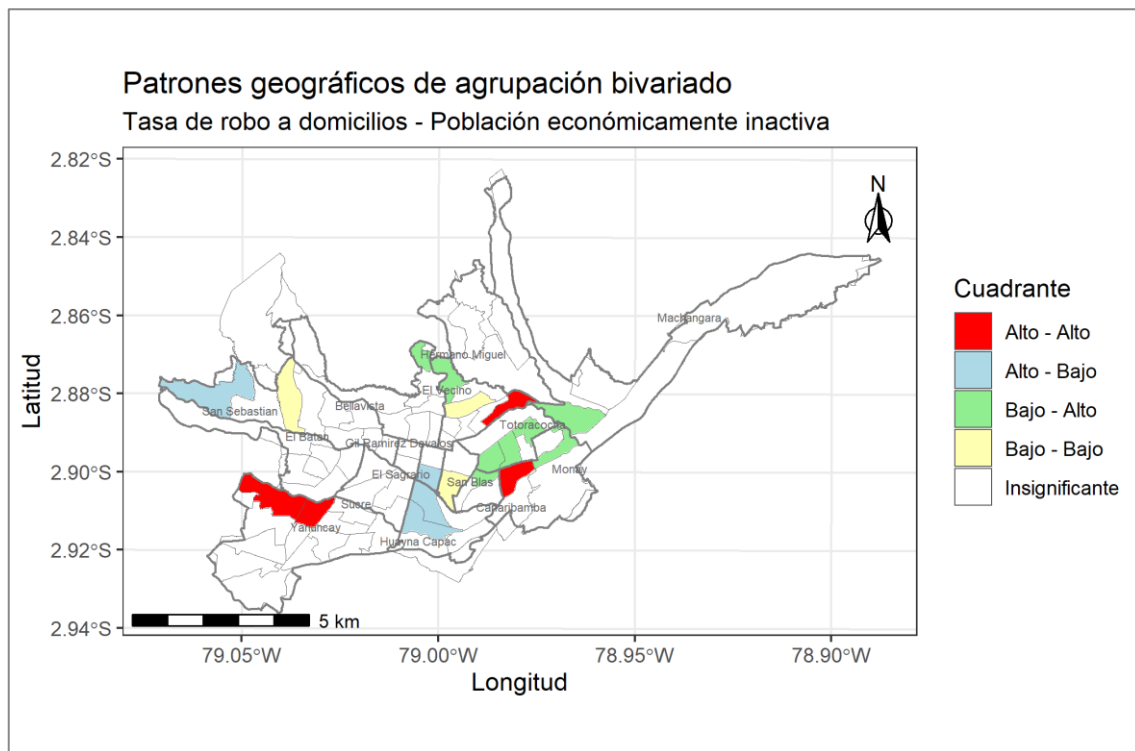


Figura 84-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs población económicamente inactiva por zonas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC - Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

En la Figura 85-4, se identifican cinco zonas urbanas con una alta concentración de la tasa de robo a domicilios; explicadas por un alto número de hogares que tienen vivienda propia y totalmente pagada (Alto – Alto) en las parroquias Bellavista y Yanuncay. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen tres zonas con una alta concentración de la tasa de robo a domicilios y que sus vecinos más cercanos tienen una baja concentración explicada por el número de hogares que tienen una vivienda propia y totalmente pagada presentes en estas zonas vecinas y esto lo corrobora el cuadrante (Bajo – Alto), para la identificación de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante los modelos espaciales aplicados.

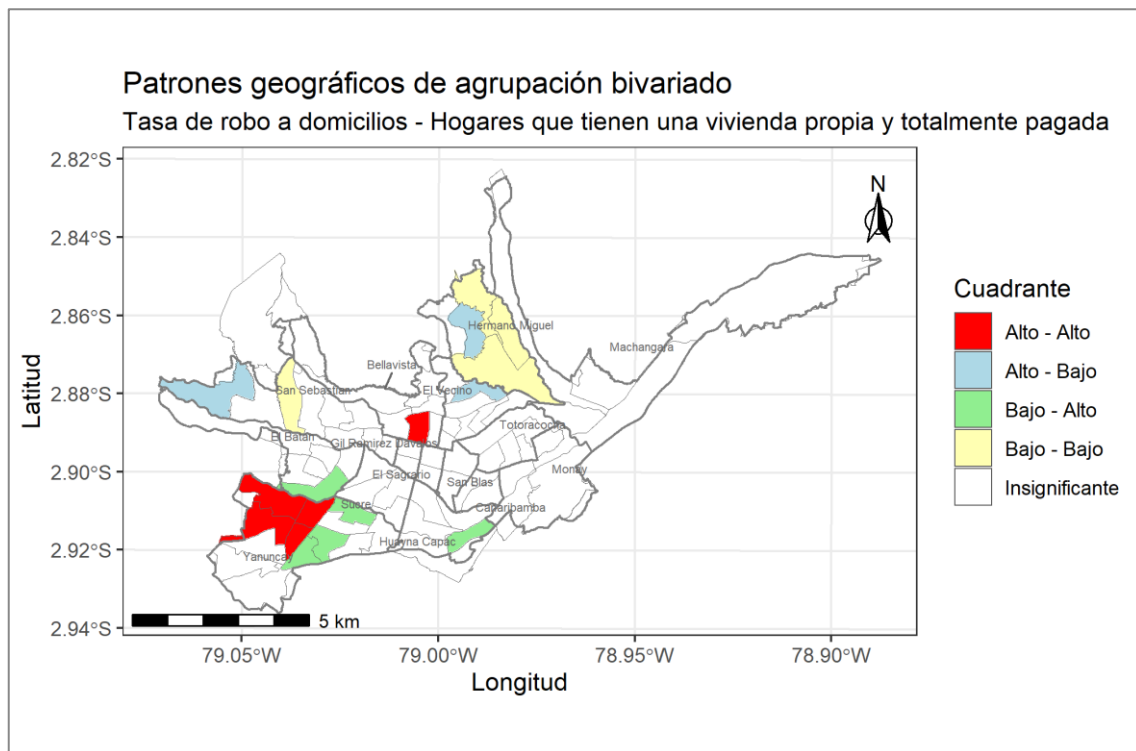


Figura 85-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a domicilios vs hogares que tienen vivienda propia y pagada por zonas de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC - Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a unidades económicas:

Se puede observar en la Figura 86-4, que en 10 zonas urbanas existe una alta concentración de robo a unidades económicas explicadas por un alto número de inmigrantes extranjeros (Alto – Alto) en las parroquias: San Blas, El Sagrario, Sucre, Gil Ramírez Dávalos y El Batán. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. También existen tres zonas con una alta concentración de robo a unidades económicas y que sus vecinos más cercanos tienen una baja concentración explicada por el número de inmigrantes internacionales presentes en estas zonas vecinas y esto los corrobora el cuadrante (Bajo – Alto). Se identificaron estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante los modelos espaciales aplicados.

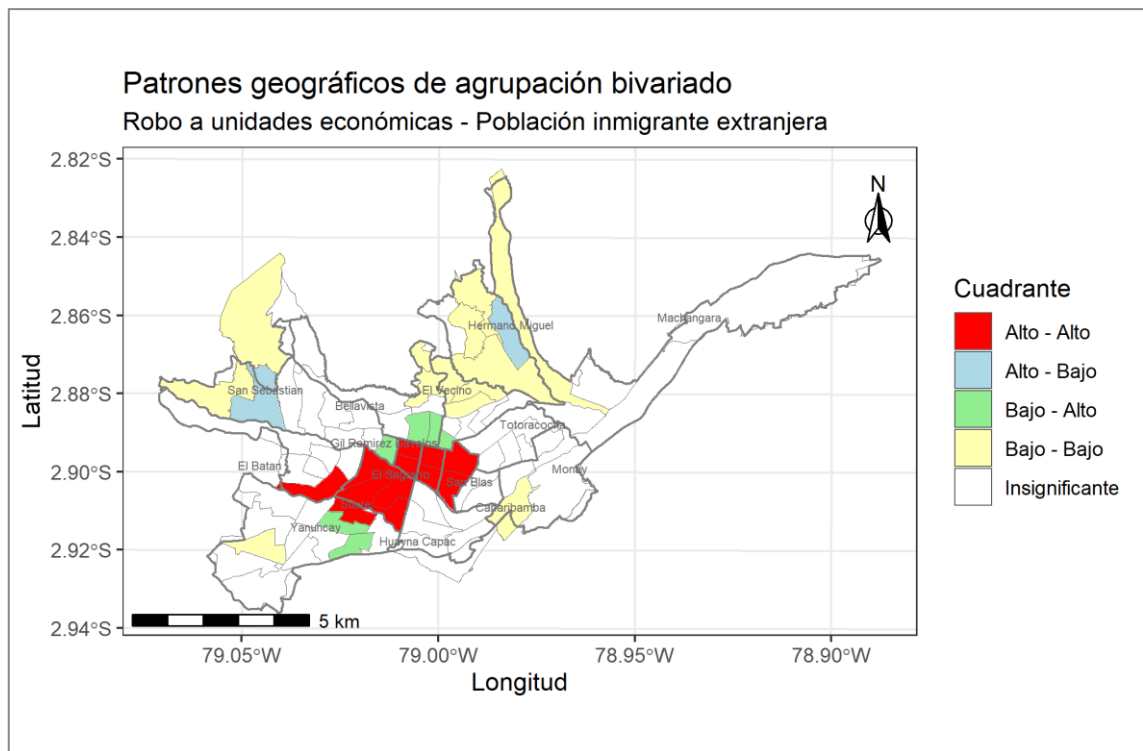


Figura 86-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población inmigrante extranjera por zonas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC - Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 87-4, que en cuatro zonas urbanas existe una alta concentración de robo a unidades económicas explicadas por un alto promedio de personas por hogar (Alto – Alto) en las parroquias: Monay, El Vecino, Hermano Miguel y San Sebastián. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen 10 zonas con una alta concentración de robo a unidades económicas y que sus vecinos más cercanos tienen una alta concentración explicada por el promedio de personas por hogar presentes en estas zonas vecinas y esto lo corrobora el cuadrante (Bajo – Alto). Se identificaron estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante la aplicación de los modelos espaciales anteriormente.

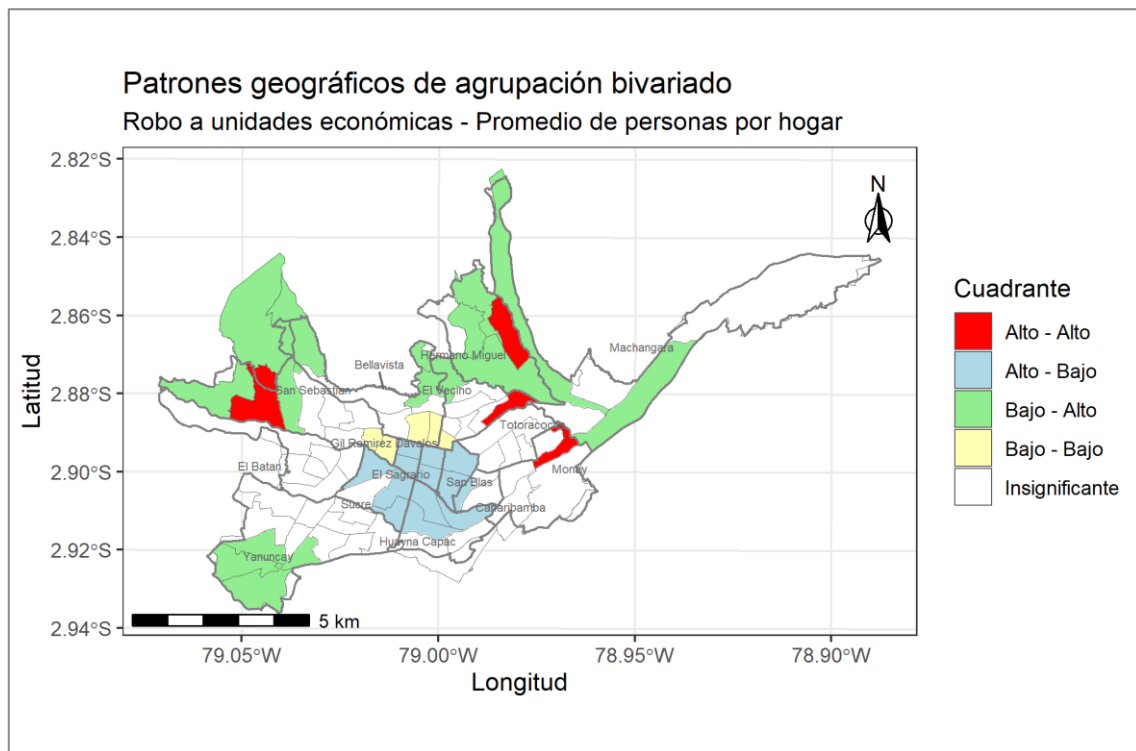


Figura 87-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas vs promedio de personas por hogar, zonas de la ciudad de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC - Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 88-4, que en cinco zonas urbanas existe una alta concentración de robo a unidades económicas explicadas por la población que realiza su actividad dentro de su hogar (Alto – Alto) en las parroquias: El Sagrario, Gil Ramírez Dávalos y San Sebastián. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen dos zonas con una alta concentración de robo a unidades económicas y que sus vecinos más cercanos tienen una alta concentración explicada por la población que realiza su actividad dentro de su hogar presentes en estas zonas vecinas (Alto – Bajo) y esto lo corrobora el cuadrante (Bajo – Alto), identificando estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante la aplicación de los modelos espaciales.

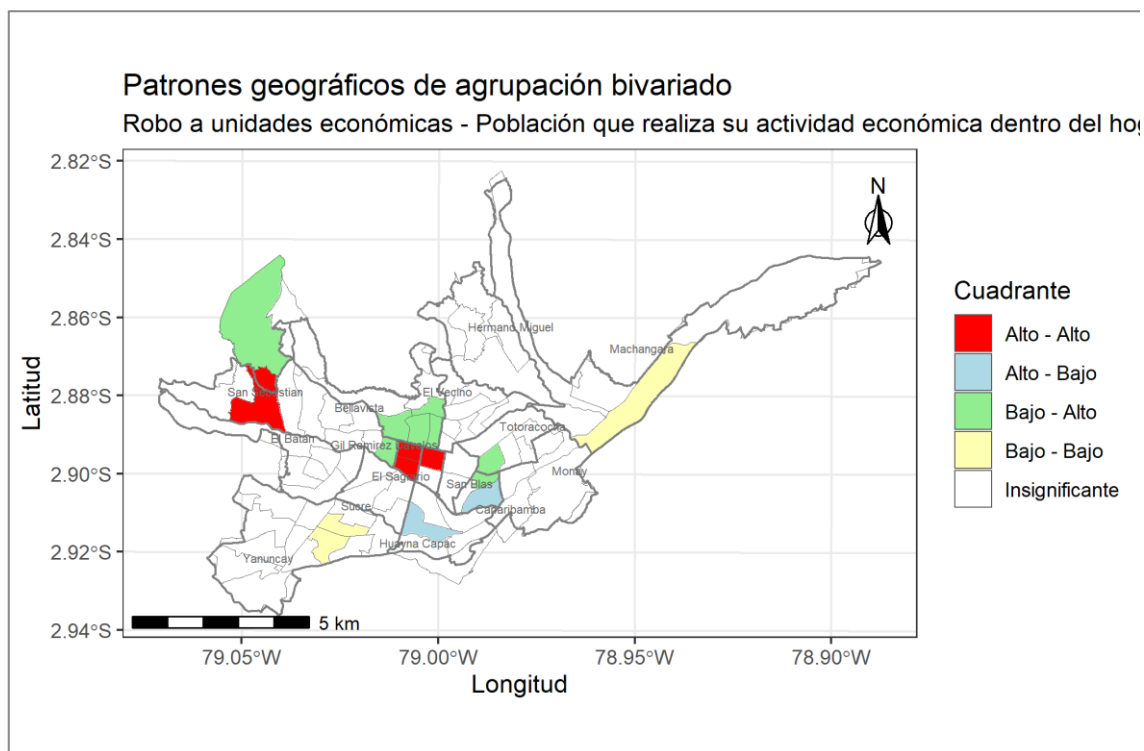


Figura 88-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población que realiza actividad económica el hogar por zonas de Cuenca, 2019.

Fuente: VII Censo de población y VI de vivienda año 2010. INEC - Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a vehículos:

Se puede observar en la Figura 89-4, que en dos zonas urbanas existe una alta concentración de robo a vehículos explicadas por una alta tasa de desplazamiento de personas por trabajo o estudio (Alto – Alto) en las parroquias Hermano Miguel y Machángara. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen cuatro zonas con una alta concentración de robo a vehículos y que sus vecinos más cercanos tienen una baja concentración explicada por la tasa de desplazamiento de personas por trabajo o estudio presentes en estas zonas vecinas (Alto – Bajo) en la parroquias de Monay, Cañaribamba y Yanuncay; esto lo corrobora el cuadrante (Bajo – Alto), identificando estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante la aplicación de los modelos espaciales.

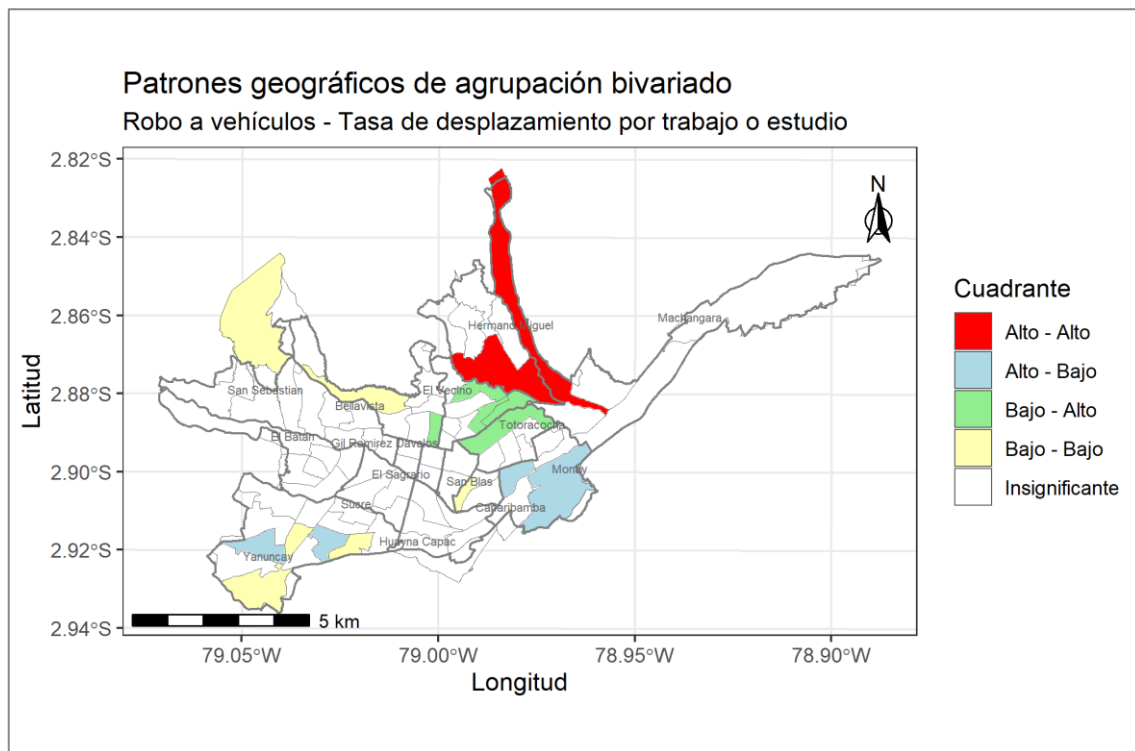


Figura 89-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs tasa de desplazamiento temporal por estudio o trabajo por zonas de Cuenca, 2019.
Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.
Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 90-4, que en 14 zonas urbanas existe una alta concentración de robo a vehículos explicadas por un alto número de población migrante internacional (Alto – Alto) en las parroquias: San Blas, Huayna Cápac, Sucre y Yanuncay. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen cuatro zonas con una alta concentración de robo a vehículos y que sus vecinos más cercanos tienen una alta concentración explicada por de número de población migrante internacional presentes en estas zonas vecinas (Alto – Bajo) y esto lo corrobora el cuadrante (Bajo – Alto), identificando estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante la aplicación de los modelos espaciales.

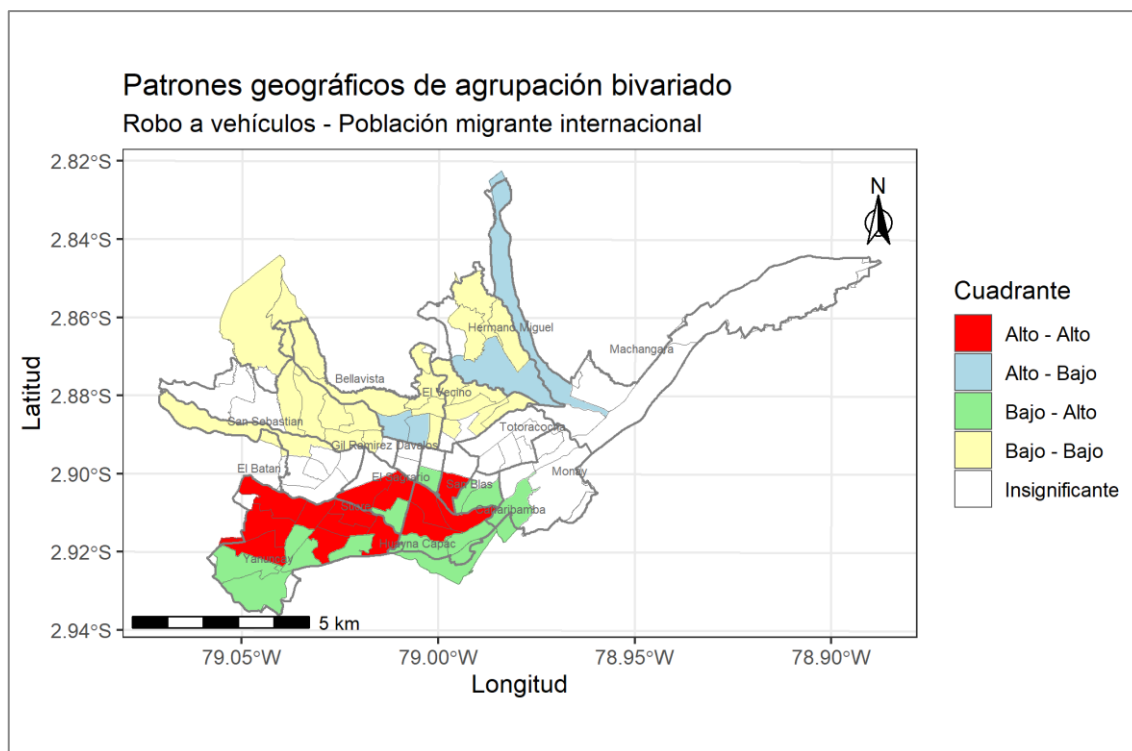


Figura 90-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población migrante internacional por zonas de la ciudad de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Se puede observar en la Figura 91-4, que en seis zonas urbanas existe una alta concentración de robo a vehículos explicadas por un alto número de personas en edad de trabajar (Alto – Alto) en las parroquias: Totoracocha, Cañaribamba y El Batán. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen tres zonas con una alta concentración de robo a vehículos y que sus vecinos más cercanos tienen una baja concentración explicada por el número de personas en edad de trabajar presentes en estas zonas vecinas (Alto – Bajo) y esto lo corrobora el cuadrante (Bajo – Alto), identificando estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante la aplicación de los modelos espaciales.

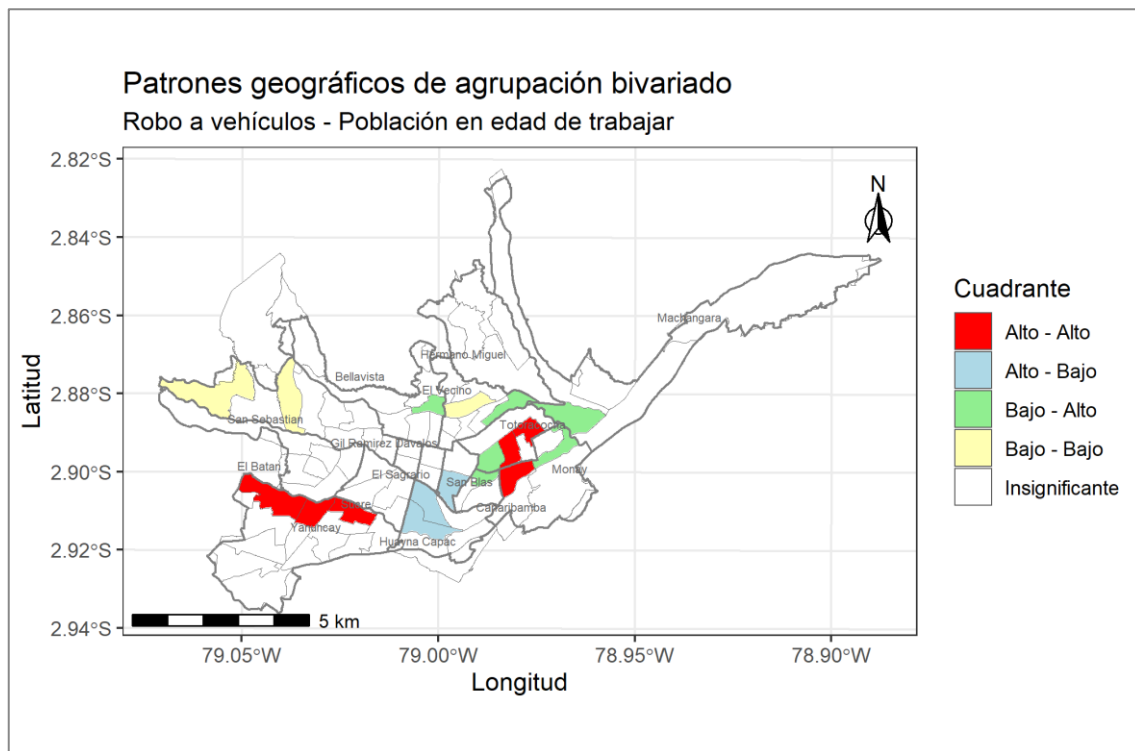


Figura 91-4: Mapa de patrones geográficos de agrupación, para los valores de la tasa de robo a personas vs población en edad de trabajar por zonas de la ciudad de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Eddy, 2022.

Se puede observar en la Figura 92-4, que en dos zonas urbanas existe una alta concentración de robo a vehículos explicadas por un alto número de población cesante (Alto – Alto) en las parroquias: San Blas, Totoracocha y Huayna Cápac. Esto quiere decir que se puede tener el mismo comportamiento con cada una de las zonas vecinas más cercanas. Mientras que también existen cuatro zonas con una alta concentración de robo a vehículos y que sus vecinos más cercanos tienen una baja concentración explicada por el número de población cesante presentes en estas zonas vecinas (Alto – Bajo) y esto corrobora el cuadrante (Bajo – Alto), identificando de estas zonas urbanas de alta incidencia dentro de la ciudad de Cuenca en el año 2019 mediante la aplicación de los modelos espaciales.

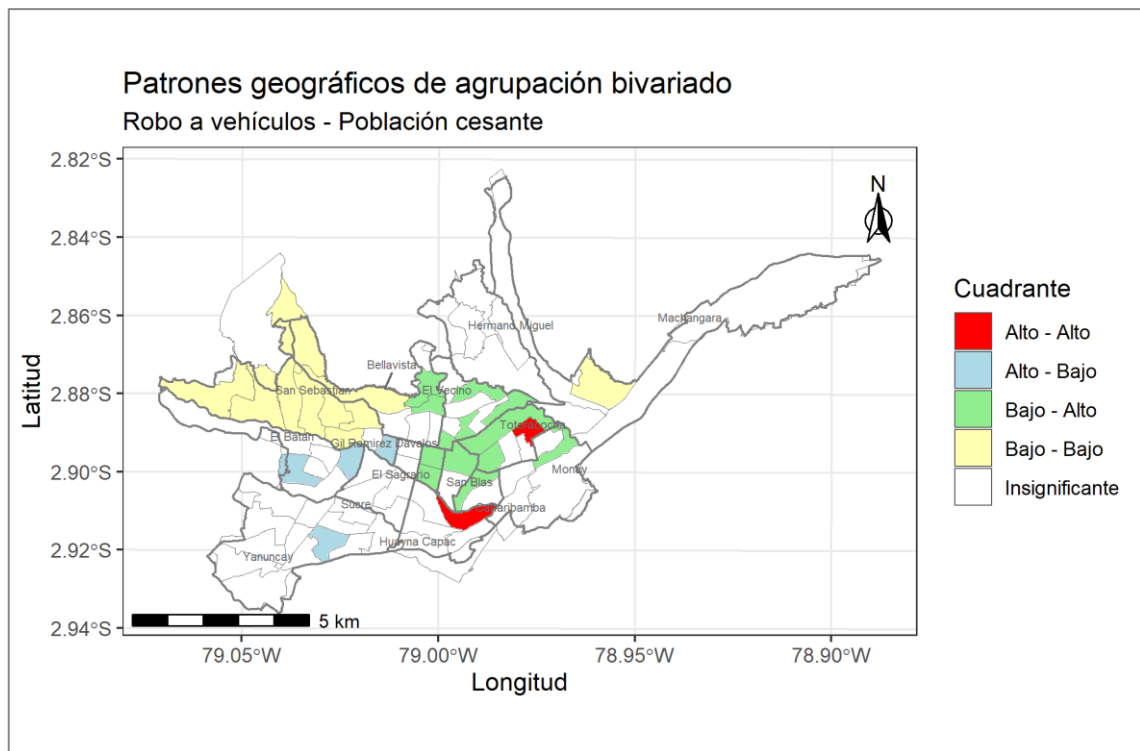


Figura 92-4: Mapa de patrones geográficos de agrupación, para los diferentes valores de la tasa de robo a personas vs población cesante por zonas urbanas de la ciudad de Cuenca, 2019.

Fuente: INEC - Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

4.4 Comprobación de la hipótesis

Para la comprobación de la hipótesis para este trabajo de investigación; Aplicación de modelos matemáticos mediante estadística espacial, se empleó el índice de Akaike (AIC), el cual se explica en el capítulo 2. Con ello, se procede a comprobar cuál de los diferentes modelos matemáticos de regresión espacial permite identificar las zonas urbanas de mayor incidencia con mayor exactitud y, que se detallan a continuación.

4.4.1 Prueba del índice estadístico de Aikake (AIC), para la selección del mejor modelo matemático aplicando la estadística espacial.

Para todos los modelos presentados se estimó la métrica AIC, la cual indica el desempeño del modelo en función de un indicador adimensional. Este indicador entre menor sea el valor es mejor. En las siguientes tablas se ilustran los resultados de este indicador para todos los modelos espaciales para cada uno de los delitos de robo aplicados con sus respectivas ecuaciones.

Robo a personas: Se puede observar en la Tabla 44-4, y siguiendo el criterio de Akaike (AIC) los modelos SAR y SEM son los que mejor se ajustan a los datos espaciales de robo a personas por zona, pero el modelo de Error Espacial (SEM) tiene el valor menor de AIC igual a 351.17,

lo que indica que este modelo es el más adecuado para medir la concentración e identificar las diferentes zonas urbanas de la ciudad de Cuenca para la tasa de robo a personas en el año 2019.

Tabla 44-4: Indicador AIC, para seleccionar el mejor modelo espacial para identificar las zonas de mayor incidencia de la tasa de robo a personas en la ciudad de Cuenca año 2019.

Modelo	Ecuación	AIC
Retardo Espacial (SAR)	$Tasa\ robo\ a\ personas = 0.5956 * lag.Tasa_R_ - 0.9297349 + 0.0175167 * BD_Inmi_Ex$	356.02
Error Espacial (SEM)	$Tasa\ robo\ a\ personas = 15.38 + 0.014577 * BD_Inmi_Ex - 3.9743645 * BD_PPH + 0.48541 * lag.resid$	351.17
Espacial de Durbin (SDM)	$Tasa\ robo\ a\ personas = 0.759 * lag.BD_Tasa_R_ + 1.5364 - 0.001063 * BD_HTV - 0.0017238 * lag.BD_HTV_i$	371.96

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a domicilios: Se puede observar en la Tabla 45-4, y siguiendo el criterio de Akaike (AIC) que los modelos SDM y SEM son los que mejor se ajustan a los datos espaciales de robo a domicilios por zona. El modelo espacial de Durbin (SDM) tiene el valor menor de AIC igual a 369.75, lo que indica que este modelo es el más adecuado para medir la concentración e identificar las diferentes zonas urbanas de la ciudad de Cuenca para la tasa de robo a domicilios en el año 2019.

Tabla 45-4: Indicador estadístico AIC, para seleccionar el mejor modelo que permite identificar las zonas de mayor incidencia de la tasa de robo a domicilios en Cuenca año 2019.

Modelos	Ecuación	AIC
Retardo Espacial (SAR)	$Tasa\ robo\ a\ domicilios = 0.212 * lag.Tasa_1 + 3.151156 - 0.02793 * BD_TIEpc$	369.93
Error Espacial (SEM)	$Tasa\ robo\ a\ domicilios = 3.864 + 0.0005958 * BD_PEI + 0.20477 * lag.resid$	372.22
Espacial de Durbin (SDM)	$Tasa\ robo\ a\ domicilios = 0.20495 * lag.BD_Tasa_1 + 7.168 - 0.00026 * BD_PMI + 0.0157 * BD_PCes + 0.00003148 * BD_PEI + 0.00003148 * BD_HTV - 0.0005458 * lag.BD_PMI_i - 0.108 * lag.BD_PCes_i - 0.001134 * lag.BD_PEI_i - 0.01089 * lag.BD_HTV_i$	369.75

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a unidades económicas: Se puede observar en la Tabla 46-4, y siguiendo el criterio de Akaike (AIC) que los modelos SAR y SEM son los que mejor se ajustan a los datos espaciales de robo a unidades económicas por zona, pero el modelo de Retardo Espacial (SAR) tiene el valor menor de AIC igual a 413.87, lo que indica que este modelo es el más adecuado para medir la concentración e identificar las diferentes zonas urbanas de la ciudad de Cuenca para la tasa de robo a personas en el año 2019.

Tabla 46-4: Indicador estadístico AIC, para seleccionar el mejor modelo que permite identificar las zonas de mayor incidencia de robo a unidades económicas en Cuenca año 2019.

Modelos	Ecuación	AIC
Retardo Espacial (SAR)	$Robo\ unidades\ económicas = 0.093 * lag.Robo_U_ + 22.03 + 0.015348 * BD_Inmi_Ex - 5.81 * BD_PPH$	413.87
Error Espacial (SEM)	$Robo\ a\ unidades\ económicas = 23.61 + 0.016 * BD_Inmi_Ex - 6.19774 * BD_PPH + 0.057 * lag.resid$	414.06
Espacial de Durbin (SDM)	$Robo\ a\ unidades\ económicas = 0.558 * lag.BD_Robo_u_ - 2.49379 - 0.0232094 * BD_POdh - 0.0232094 * lag.BD_POdh_i$	437.08

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Robo a vehículos: Se puede observar en la Tabla 47-4, y siguiendo el criterio de Akaike (AIC) que los modelos SAR y SEM son los que mejor se ajustan a los datos espaciales de robo a vehículos por zona, pero el modelo de Error Espacial (SEM) tiene el valor menor de AIC igual a 448.64, lo que indica que este modelo es el más adecuado para medir la concentración e identificar las diferentes zonas urbanas de la ciudad de Cuenca para el número de robo a vehículos en el año 2019.

Tabla 47-4: Indicador estadístico AIC, para seleccionar el mejor modelo que permite identificar las zonas de mayor incidencia de robo a vehículos en la ciudad de Cuenca año 2019.

Modelos	Ecuación	AIC
Retardo Espacial (SAR)	$Robo\ a\ vehículos = 0.2484 * lag.Robo_ve - 9.876 + 0.29687 * BD_Tasa_Dt + 0.000415 * BD_PMI + 0.00206385 * BD_PET$	449.76
Error Espacial (SEM)	$Robo\ a\ vehículos = -0.8933 + 0.3036 * BD_Tasa_Dt + 0.00042987 * BD_PMI + 0.002249 * BD_PET + 0.3293 * lag.resid$	448.64
Espacial de Durbin (SDM)	$Robo\ a\ vehículos = 0.348 * lag.BD_Robo_ve_ + 0.309 + 0.28 * BD_Tasa_Dt + 0.0853 * lag.BD_PCes_i + 0.1739 * lag.BD_Tasa_Dt - 0.2922 * lag.BD_PCes$	458.54

Fuente: Base de datos delitos, Observatorio de seguridad Ciudadana Cuenca 2019.

Realizado por: Velasteguí Casco, Edy, 2022.

Por lo tanto, de esta manera se comprueba la hipótesis de cuáles son los modelos matemáticos aplicando estadística espacial y mediante el índice de AIC, que determinan su utilidad para identificar las zonas de mayor incidencia delictiva con mayor exactitud en cada una de las modalidades de robos, obteniendo que el modelo de SEM es el mejor en los casos de robo a personas y vehículos; el modelo SAR en el caso de robo a unidades económicas y el modelo SDM para el caso de robo a domicilios.

CONCLUSIONES

- Las bases teóricas descritas en el capítulo 2 (marco teórico), contribuyeron a sustentar el presente trabajo de investigación, técnicos e investigadores en temas de explicar la seguridad ciudadana.
- Los procedimientos detallados en el capítulo 3 sobre la metodología para emplear datos de área para aplicar los modelos matemáticos mediante estadística espacial, con el fin de identificar las zonas de mayor concentración en la ciudad de Cuenca, permitieron mejorar el análisis de manera sistematizada; además la utilización de herramientas tecnológicas actuales como en este caso el lenguaje estadístico R, mediante su entorno de desarrollo RStudio, los paquetes espaciales: *rgdal*, *raster*, *sp* y *spdep*; los cuales contienen un conjunto de librerías y funciones, permitieron obtener los parámetros de los modelos aplicados en este trabajo. Es importante también concluir que en el país aún no se reconoce la importancia del análisis espacial dado de que existe muy poca bibliografía local sobre este tipo de estudios.
- Los modelos y los análisis de vecindad investigados y aplicados en este trabajo; que el más idóneo es el Queen (La Reina), dado que es el más completo para determinar las distancias y la correlación espacial que las zonas tienen con otras cercanas. También es importante mencionar que los demás modelos de vecindad descritos y estudiados en este trabajo no se utilizaron debido a que funcionan en un solo sentido; tal es el caso del método Keep (La Torre). De manera similar funciona el método de Boshot (El Alfil) a diferencia del método de Keep este se mueve de manera diagonal en el sentido geográfico Noroeste a Sureste y de Noreste a Suroeste, mientras que el método de Queen utiliza la funcionalidad de estos dos. Es importante decir que existe el método K-vecinos el mismo que tiene la condición de no considerar el tamaño del área de las zonas analizadas ubicando el centroide según el número de vecinos que el investigador desee.
- Los resultados de modelos de regresión espacial obtenidos con respecto a los robos a: personas, domicilios, unidades económicas y vehículos; Se determinó que estas zonas en su mayoría están en el centro de la ciudad con respecto a la tasa de robo a personas, domicilios y unidades económicas mientras que el robo a vehículos presentó otro comportamiento los mismos que se concentran en el área rural de la ciudad.
- Al analizar y discutir los resultados, se comprobó que para poder identificar las zonas de alta incidencia en la ciudad de Cuenca los modelos matemáticos mediante la estadística espacial son: El modelo SEM para robo a personas y vehículos, el modelo SAR para robo a domicilios y el modelo SDM para robo a unidades económicas, llegando a probar de esta manera mediante el índice AIC la validez de este trabajo, cumpliendo con los objetivos planteados tanto general como específicos en este trabajo.

RECOMENDACIONES

- Se recomienda que, para este tipo de estudios o trabajos sobre análisis estadístico espacial, se utilicen programas estadísticos que cuenten con funcionalidades orientadas a SIG y análisis espacial, ya que permiten analizar los diferentes fenómenos que se presenten dentro de una zona o área geográfica.
- Es importante recomendar el aprendizaje continuo sobre el análisis estadístico espacial ya que en la actualidad es una ciencia bastante eficiente para analizar fenómenos que ocurren en los diferentes niveles geográficos. Por ello es importante investigar y aplicar este tipo de modelos espaciales en las diferentes áreas, para identificar los puntos o zonas críticas, con lo cual se aporta para una mejor toma de decisiones, ya que es posible visualizar de manera objetiva el conflicto delictivo.
- Se recomienda la aplicación del método *Queen* para determinar las distancias entre zonas vecinas más peligrosas o de alta incidencia para la ubicación de los diferentes dispositivos de disuasión; como cámaras de video vigilancia, puntos de auxilio y ubicación de Unidades de Policía Comunitaria (UPC). Con ello es posible mejorar de cierta manera la asistencia y atención a este tipo de problemáticas como el delito del robo en sus diferentes modalidades.
- También se recomienda la aplicación de este tipo de análisis matemático mediante estadística espacial para analizar los demás delitos existentes. Es importante destacar la necesidad de que las instituciones correspondientes cuenten con bases de datos georreferenciadas, ya que esto permitirá identificar las zonas conflictivas donde se presenten eventos delictivos y también otros fenómenos ya sean sociales o naturales.

GLOSARIO

AE: Autocorrelación espacial

CONALI: Consejo Nacional De Límites Internos del Ecuador

CSC: Consejo de Seguridad Ciudadana del cantón Cuenca

INEC: Instituto Nacional de Estadística y Censos

MAUP: Problema de unidad de área modificable

MCO: Mínimos cuadrados ordinarios

MV: Método de máxima verosimilitud

SAR: Modelo espacial autorregresivo ó spatial lag

SDM: Modelos espacial de Durbin

SEM: Modelo de error espacial ó spatial error

SIG: Sistema de Información Geográfica

BIBLIOGRAFÍA

- Aguayo Téllez, E., & Medellín Mendoza, S. E. (2014). Dependencia espacial de la delincuencia en Monterrey, México. *Ecos de Economía*, 71.
- Alcívar Tejo, C., & Calderón Cisneros, J. (2013). *La Delincuencia en la ciudad de Guayaquil, Un Análisis Espacial de su Distribución por Delito*. Guayaquil: s/f.
- Andrade Quizhpe, J. L. (2017). *Análisis espacial de la distribución del delito de robo en el distrito metropolitano de Quito-Ecuador*. Quito: s/f.
- Borrego Sánchez, J. Á. (2018). Modelos de Regresión para Datos Espaciales. 11. s/f.
- Bravo López, P. E. (2021). Autocorrelación espacial - Índices para determinar su presencia en datos geográficos breve revisión de la literatura. *Universidad Verdad Revista Científica de Ciencias Sociales y Humanas*, 51-59.
- Buzai, G. D., Fuenzalida, M., Moreno Jiménez, A., & García De León, A. (2015). *Gepgrafía, Geotecnología y Análisis Espacial: Tendencias , Métodos y Aplicaciones*. Santiago: TRIANGULO.
- Cabrero Ortega, Y., & Gaecía Pérez, A. (2015). *Análisis Estadístico de Datos Espaciales con QGIS y R*. UNED.
- Carles Martori, J., & Hoberg, K. (2008). Nuevas Técnicas de Estadística Espacial para la detección de Clusters Residenciales de Población Inmigrante. *Scripta Nova*.
- Castro Toledo, F. J., Gómez Bellvís, A. B., & Buil Gil, D. (2019). *La Criminología que Viene*. REJIC.
- Coba Rubio, E., & Balseca Carrera, C. E. (2015). *Análisis Geo-estadístico de la delincuencia para Quito Urbano usando GeoDa*. Quito: s/f.
- Coro Chasco, Y. (2003). *Econometría Espacial Aplicada a la Predicción-Extrapolación de datos Microterritoriales*. Madrid: Consejería de Economía e Innovación Tecnológica.
- Corral De Whitt, D., Carrera, E. V., Muñoz Romero, S., & Rojo Álvarez, J. L. (2018). Mapeo Estadístico, Espacial y Temporal de Emergencias 911 en Ecuador. *MDPI*.
- David, J., Gallegos, J., López, A., Nolasco, M., Ramírez, D., & Vilalta, C. (2019). *Análisis Espacial para la reducción de la Violencia Letal*. Open Society Fundations.
- Díaz Viera, M. A. (2002). *Geosestadística Aplicada*. UNAM, CITMA.
- Fischer, M. M., & Wang, J. (2015). *Spatial Data Analysis Models, Methods and Thechniques*. Springer.

- Franco Galeano, V. (2018). *Análisis Espacial de la Criminalidad a Nivel Intra-urbano el caso de Medellín, Colombia*. Medellín: s/f.
- Fundación PAZ CIUDADANA. (03 de 2019). <https://pazciudadana.cl/>. Obtenido de <https://pazciudadana.cl/biblioteca/analisis-criminal/buenas-practicas-en-analisis-criminal-en-america-latina-2018/>
- Garabito, D. (11 de Abril de 2019). *\$ Pubs por RStudio*. Obtenido de Modelos para datos por áreas: <https://rpubs.com/bogotan/AMESP12ModelosDatosareas>
- Giraldo Henao, R. (2015). *Introducción a la Geoestadística*. Colombia: Departamento de Estadística Universidad Nacional de Colombia.
- Herrera Gómez, M., & Paz, J. (2012). Introducción a la Econometría Espacial. Una Aplicación al Estudio de la Fecundidad en la Argentina usando R. *ResearchGate*.
- Herrera, M. (2015). *Econometría espacial usando Stata. Breve guía aplicada para datos de corte trasversal*. Salta: UNSa.
- LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.
- Lind, D. A., Mason, R. D., & Marchal, W. G. (2001). *Estadística para Administración y Economía*. McGraw-Hill Interamericana .
- Martori, J. C., & Hoberg, K. (15 de Abril de 2008). Nuevas técnicas de Estadística Espacial para la detección de clusters residenciales de población inmigrante. *Scripta Nova Rervista electrónica de geografía y ciencias sociales*, 12(263), 13-14.
- MINISTERIO DE GOBERNACIÓN Guatemala C.A. (2014). *Estudio Espacial de los Homicidios y Femicidios Departamento de Escuintla, Guatemala*. s/f.
- Morales Oñate, V., & Morales, B. (2019). Innovación en Ecuador: Un Enfoque Espacial. *VI Congreso Internacional de Ciencia, Tecnología, Emprendimiento e Innovación*, 327-329.
- Morales, A. (2012). *MmappingGIS*. Obtenido de MmappingGIS: https://mappinggis.com/2019/12/los-paquetes-de-r-para-gis-mas-utilizados/#2_sp_Clases_y_metodos_para_datos_espaciales
- Moreno López, J. L. (2009). *Introducción a la Geoestadística y su Aplicación*. México: s/f.
- Oñate, V. M. (29 de marzo de 2022). *Econometría Espacial*. Obtenido de Econometría Espacial: https://bookdown.org/victor_morales/SpatialEconometrics/regresi%C3%B3n-lineal.html

- Rosero Santana, M. F. (2015). *Aplicación de Técnicas Espaciales para el análisis de patrones y modelación del delito en la provincia de Pichincha, utilizando herramientas Geo estadísticas, aplicaciones para el SIS-Ecu-911*. Quito: s/f.
- Sánchez Martín, J. M., Sánchez Rivero, M., & Rengifo Gallego, J. I. (2017). *Análisis de Equilibrio entre el potencial Turístico y la oferta de Alojamientos en Turismo Rural Mediante técnicas de Estadística Espacial. Una Aplicación a la Provincia de Cáceres (España)*. Universidad de Murcia.
- Santa , F., & Fuentes López, J. (2012). Modelos de Regresión Espacial para el Comportamiento de las Enfermedades Infecciosas Dengue y Malaria en Colombia para los años 2000, 2005 y 2010. *ResearchGate*.
- Sawada, M. (2009). *Universidad de Ottawa*. Obtenido de Departamento de Geografía:
<http://www.lpc.uottawa.ca/publications/moransi/moran.htm>
- Vilalta y Perdamo, C. J. (2005). *Cómo Enseñar Autocorrelación Espacial*. *Redalyc*.

ANEXOS

Anexo A.

Código R de análisis exploratorio de variables e indicadores mediante mapas de coropletas:

```
-----

library(tidyverse)
library(raster)
library(sf)
library(sp)
library(ggspatial)
library(spdep)
library(rgdal)
library(rspatial)
library(rgeos)
library(readxl)
library(ggplot2)
library(car)

## leer la base de datos formato Shape##

setwd("E:/Tesis Edy/Datos_Tesis/Delitos_Zonas")
shp <- st_read("indicadores_zonas.shp")

setwd("E:/Tesis Edy/Datos_Tesis/Parro_ur_Cueca")
shp1 <- st_read("purbanas.shp")

## leer la base de datos formato Excel##

tbl <- read_excel("E:/Tesis Edy/Datos_Tesis/Datos_Zonas_Cuenca1.xlsx")
View(tbl)
tbl1 <- read_excel("E:/Tesis Edy/Datos_Tesis/Datos_Robos.xlsx")

tbl2 <- read_excel("E:/Tesis Edy/Datos_Tesis/Parro_ur_Cueca/Parroquias_Cuenca.xlsx")

## Gráficar la ciudad de Cuenca por zonas censales ##

shp %>%
  ggplot() +
  geom_sf()

## Gráficar la ciudad de Cuenca por zonas censales ##

shp1 %>%
  ggplot() +
  geom_sf()

## Análisis exploratorio de variables censales por zonas ##

pan_map_roboue %>%
  # usamos el aesthetic fill para indicar la columna de casos
  ggplot(aes(fill = Robo_U_economicas)) +
  geom_sf() +
  annotation_north_arrow(location='tr') +
  scale_fill_viridis_c(alpha = 0.75) +
  ggplot2::theme_bw() + #hasta aqui es sin titulo y sin tanto chechere
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Unidades económicas que fueron robadas en el año 2019",
    subtitle = "Unidades económicas robadas por zona censal, ciudad Cuenca",
    fill = "Número de unidades\n económicas")+
  annotation_scale()

pan_map_robov <- shp %>%
  # unir tablas
  left_join(tbl1,
    # indicar explícitamente las columnas índice,
    # necesario cuando no tienen el mismo nombre
    by = "DPA_ZONA")

pan_map_robov %>%
  # usamos el aesthetic fill para indicar la columna de casos
  ggplot(aes(fill = Robo_vehiculos)) +
```

```

geom_sf() +
annotation_north_arrow(location='tr') +
scale_fill_viridis_c(alpha = 0.75) +
ggplot2::theme_bw() + #hasta aqui es sin titulo y sin tanto chechere
ggplot2::xlab("Longitud") +
ggplot2::ylab("Latitud") +
ggplot2::labs(
  title = "Vehículos que fueron robados en el año 2019",
  subtitle = "Vehículos robados por zona censal, ciudad Cuenca",
  fill = "Número de vehículos")+
annotation_scale()

pan_map_tasarp <- shp %>%
# unir tablas
left_join(tbl1,
  # indicar explícitamente las columnas índice,
  # necesario cuando no tienen el mismo nombre
  by = "DPA_ZONA")

pan_map_tasarp %>%
# usamos el aesthetic fill para indicar la columna de casos
ggplot(aes(fill = Tasa_R_pers)) +
geom_sf() +
annotation_north_arrow(location='tr') +
scale_fill_viridis_c(alpha = 0.75) +
ggplot2::theme_bw() + #hasta aqui es sin titulo y sin tanto chechere
ggplot2::xlab("Longitud") +
ggplot2::ylab("Latitud") +
ggplot2::labs(
  title = "Tasa de robo a personas en el año 2019",
  subtitle = "Tasa de robo a personas por zona censal, ciudad Cuenca",
  fill = "Tasa de robo personas\n (100 habitantes)")+
annotation_scale()

pan_map_tasard <- shp %>%
# unir tablas
left_join(tbl1,
  # indicar explícitamente las columnas índice,
  # necesario cuando no tienen el mismo nombre
  by = "DPA_ZONA")

pan_map_tasard %>%
# usamos el aesthetic fill para indicar la columna de casos
ggplot(aes(fill = Tasa_R_domi)) +
geom_sf() +
annotation_north_arrow(location='tr') +
scale_fill_viridis_c(alpha = 0.75) +
ggplot2::theme_bw() + #hasta aqui es sin titulo y sin tanto chechere
ggplot2::xlab("Longitud") +
ggplot2::ylab("Latitud") +
ggplot2::labs(
  title = "Tasa de robo a domicilios en el año 2019",
  subtitle = "Tasa de robo a domicilios por zona censal, ciudad Cuenca",
  fill = "Tasa de robo domicilios\n (100 domicilios)")+
annotation_scale()

```

Código R modelos de SAR, SEM y SDM:

```

# Leer librerías -----
require(pacman)
pacman::p_load(terra, spdep, tidyr, spatialreg, broom, viridis, raster, sf, ggrepel, ggspatial,
hablar, fs, car, tidyverse, geodata, ggpubr, RColorBrewer, rgeoda, gtools, rgeos, stringr, glue)

g <- gc(reset = T)
rm(list = ls())
options(scipen = 999, warn = -1)

# Leer la base de data -----
bdta <- st_read('gpkg/spatial_bdata.gpkg')

# ecuaciones de los modelos -----

# Personas, domicilios, unidades economicas y vehiculos
f1 <- BD_Tasa_R_ ~ BD_TA + BD_Tasa_Na + BD_Tasa_Dt + BD_Inmi_Ex + BD_PMI + BD_NEI + BD_TIEpc + BD_PPH
+ BD_PET + BD_PCes + BD_POdH + BD_PEI + BD_HTV
f2 <- BD_Tasa_1 ~ BD_TA + BD_Tasa_Na + BD_Tasa_Dt + BD_Inmi_Ex + BD_PMI + BD_NEI + BD_TIEpc + BD_PPH
+ BD_PET + BD_PCes + BD_POdH + BD_PEI + BD_HTV
f3 <- BD_Robo_U_ ~ BD_TA + BD_Tasa_Na + BD_Tasa_Dt + BD_Inmi_Ex + BD_PMI + BD_NEI + BD_TIEpc + BD_PPH
+ BD_PET + BD_PCes + BD_POdH + BD_PEI + BD_HTV
f4 <- BD_Robo_ve ~ BD_TA + BD_Tasa_Na + BD_Tasa_Dt + BD_Inmi_Ex + BD_PMI + BD_NEI + BD_TIEpc + BD_PPH
+ BD_PET + BD_PCes + BD_POdH + BD_PEI + BD_HTV
fs <- list(f1, f2, f3, f4)

```

```

# Cálculo del método de vecindad -----
seab <- poly2nb(bdta, queen = T)
seaw <- nb2listw(seab, style = 'W', zero.policy = T)

# Índice de AIC -----
clss <- c('Personas', 'Domicilios', 'Unidades economicas', 'Vehiculos')

# Spatial Lag model -----
lagm <- purrr::map(.x = 1:4, .f = function(i){

  cat(i, '\n')
  fr <- fs[[i]]
  ft <- lagsarlm(fr, data = bdta, listw = seaw)
  ai <- AIC(ft)
  rs <- tibble(type = clss[i], typemodel = 'lagmodel', aic = ai)
  return(rs)

}) %>%
  bind_rows()

# Spatial error model -----
sper <- purrr::map(.x = 1:4, .f = function(i){

  cat(i, '\n')
  fr <- fs[[i]]
  ft <- errorsarlm(fr, data = bdta, listw = seaw)
  ai <- AIC(ft)
  rs <- tibble(type = clss[i], typemodel = 'Spatial error model', aic = ai)
  return(rs)

}) %>%
  bind_rows()

# Durbin spatial model -----
drbn <- purrr::map(.x = 1:4, .f = function(i){

  cat(i, '\n')
  fr <- fs[[i]]
  ft <- lagsarlm(fr, data = bdta, listw = seaw, type = 'mixed')
  ai <- AIC(ft)
  rs <- tibble(type = clss[i], typemodel = 'Durbin spatial model', aic = ai)
  return(rs)

}) %>%
  bind_rows()

# Unir las tablas en una sola -----
allr <- rbind(lagm, sper, drbn)
minm <- allr %>% group_by(type) %>% top_n(n = -1, wt = aic) %>% ungroup()
allr <- spread(allr, typemodel, aic)
colnames(allr) <- gsub(' ', '_', colnames(allr))
allr <- mutate(allr, select_model = pull(minm, typemodel))

write.csv(allr, 'tbl/results/aic_models.csv', row.names = FALSE)

```

Código R identificación de zonas de alta incidencia:

```

# Leer librerías de los paquetes para estadística espacial -----
library(tidyverse)
library(sf)
library(spdep)
library(ggplot2)

# Leer datos espaciales -----

shp_Cuenca <- read_sf("C:/Users/Edy/Desktop/Datos_Tesis/Indicadores_Zonas")
#class(shp_Cuenca) # muestra que la base de datos es una base geométrica y al mismo tiempo un data
frame
#st_geometry(shp_Cuenca) # muestra la geometría de la base de datos espacial

shp_Cuenca1 <- read_sf("C:/Users/Edy/Desktop/Datos_Tesis/BD_Robo")

# TASA DE ROBO A PERSONAS -----
-

# GENERAMOS EL MAPA DE CUENCA Y SU DPA

ggplot(data = shp_Cuenca1) +
  geom_sf()

## GENERACIÓN DE MATRIZ DE PESOS ESPACIALES #

```

```

coords <- coordinates(as(shp_Cuenca1), 'Spatial')

## METODO DE CRITERIO ROOK (DE LA TORRE) DE VECINDAD #

rook_Cuenca1 <- poly2nb(as(shp_Cuenca1, 'Spatial'), queen = FALSE)

nb_to_df <- function(nb, coords){
  x <- coords[, 1]
  y <- coords[, 2]
  n <- length(nb)

  cardnb <- card(nb)
  i <- rep(1:n, cardnb)
  j <- unlist(nb)
  return(data.frame(x = x[i], xend = x[j],
                    y = y[i], yend = y[j]))
}

# GENERAMOS EL DATAFRAME #

rook_Cuenca1_df <- nb_to_df(rook_Cuenca1, coords)

# GRAFICAMOS LA VECINDAD MEDIANTE EL CRITERIO DE ROOK (DE LA TORRE) #

ggplot(shp_Cuenca1) +
  geom_sf()+
  geom_point(data = rook_Cuenca1_df,
             mapping = aes(x = x, y = y))+
  geom_segment(data = rook_Cuenca1_df,
              mapping = aes(x = x, xend = xend, y = y, yend = yend))+
  labs(x = "", y = "") +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Vecindad mediante el criterio de Rook (De la Torre)",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# VECINDAD MEDIANTE CRITERIO QUEEN (DE LA REINA)

queen_Cuenca1 <- poly2nb(as(shp_Cuenca1, 'Spatial'), queen = T)

queen_Cuenca1_df <- nb_to_df(queen_Cuenca1, coords)

# GRAFICAMOS LA VECINDAD MEDIANTE EL CRITERIO QUEEN (DE LA REINA)

ggplot(shp_Cuenca1) +
  geom_sf()+
  geom_point(data = queen_Cuenca1_df,
             mapping = aes(x = x, y = y))+
  geom_segment(data = queen_Cuenca1_df,
              mapping = aes(x = x, xend = xend, y = y, yend = yend))+
  labs(x = "", y = "") +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Vecindad mediante el criterio de Queen (De la Reina)",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# VECINDAD MEDIANTE CRITERIO k-cercanos (K-nearest) #

kn_Cuenca1 <- knn2nb(knearneigh(coords, k = 6))

kn_Cuenca1_df <- nb_to_df(kn_Cuenca1, coords)

# GRAFICAMOS LA VECINDAD MEDIANTE EL CRITERIO k-cercanos (K-nearest)

ggplot(shp_Cuenca1) +
  geom_sf()+
  geom_point(data = kn_Cuenca1_df, mapping = aes(x = x, y = y))+
  geom_segment(data = kn_Cuenca1_df,
              mapping = aes(x = x, xend = xend, y = y, yend = yend))+
  labs(x = "", y = "") +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Vecindad mediante el criterio de k-cercanos (K-nearest)",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

```

```

# AUTOCORRELACION ESPACIAL GLOBAL -----
# INDICE I DE MORAN #

queen_Cuenca1_lw <- nb2listw(queen_Cuenca1)
moran.test(shp_Cuenca1$BD_Tasa_R_, listw = queen_Cuenca1_lw)

# GRAFICO DE DISPERSION PARA VISUALIZAR EL SIGNO Y LA FUERZA DE CORRELACION ESPACIAL#

moran.plot(shp_Cuenca1$BD_Tasa_R_, listw = queen_Cuenca1_lw,
           xlab = "Tasa de robo a personas",
           ylab = "Tasa de robo a personas (con corr. espacial)")

# PATRONES GLOBALES DE AUTOCORRELACION ESPACIAL I DE MORAN -----

queen_Cuenca1_b_lw <- nb2listw(queen_Cuenca1, style = "B")

shp_Cuenca1 <- shp_Cuenca1 %>%
  mutate(lmoran = localmoran(x = BD_Tasa_R_, listw = queen_Cuenca1_b_lw)[, 1],
         lmoran_pval = localmoran(x = BD_Tasa_R_, listw = queen_Cuenca1_b_lw)[, 5]
        )

# MAPA ESTADISTICO LOCAL DE MORAN -----

ggplot (shp_Cuenca1) +
  geom_sf(aes(fill = lmoran))+
  labs(fill = " Estadística \n local I de Moran") +
  scale_fill_viridis_c() +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Autocorrelacion mediante el estadístico local I de Moran",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# MAPA CLUSTER DE LISA PARA VER A DONDE SE AGRUPA LA AUCORRELACION -----

shp_Cuenca1 <- shp_Cuenca1 %>%
  mutate(
    # Estandarizar la tasa de analfabetismo y el Moran local a sus valores medios:
    st_BD_Tasa_R_ = BD_Tasa_R_ - mean(BD_Tasa_R_),
    st_lmoran = lmoran - mean(lmoran),
    # Crear la nueva variable categórica:
    cuadrante = case_when(
      lmoran_pval > 0.05 ~ "Insignificante",
      st_BD_Tasa_R_ > 0 & st_lmoran > 0 ~ "Alto-Alto",
      st_BD_Tasa_R_ < 0 & st_lmoran < 0 ~ "Bajo-Bajo",
      st_BD_Tasa_R_ < 0 & st_lmoran > 0 ~ "Bajo-Alto",
      st_BD_Tasa_R_ > 0 & st_lmoran < 0 ~ "Alto-Bajo"
    )
  )

# GRAFICAMOS MAPA -----

ggplot(shp_Cuenca1, aes(fill = cuadrante)) +
  geom_sf() +
  labs(fill = "Cuadrante") +
  scale_fill_manual(values = c("red", "lightblue", "lightgreen", "lightyellow", "white")) +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Patrones geográficos de agrupación para diferentes valores de Robo a personas",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# TASA DE ROBO A DOMICILIOS -----

# GENERAMOS EL MAPA DE CUENCA Y SU DPA

ggplot(data = shp_Cuenca1) +
  geom_sf()

## GENERACIÓN DE MATRIZ DE PESOS ESPACIALES #

coords <- coordinates(as((shp_Cuenca1), 'Spatial'))

## METODO DE CRITERIO ROOK (DE LA TORRE) DE VECINDAD #

rook_Cuenca2 <- poly2nb(as(shp_Cuenca1, 'Spatial'), queen = FALSE)

nb_to_df <- function(nb, coords){
  x <- coords[, 1]
  y <- coords[, 2]
  n <- length(nb)

```

```

cardnb <- card(nb)
i <- rep(1:n, cardnb)
j <- unlist(nb)
return(data.frame(x = x[i], xend = x[j],
                  y = y[i], yend = y[j]))
}

# GENERAMOS EL DATAFRAME #

rook_Cuenca2_df <- nb_to_df(rook_Cuenca2, coords)

# GRAFICAMOS LA VECINDAD MEDIANTE EL CRITERIO DE ROOK (DE LA TORRE)#

ggplot(shp_Cuenca1) +
  geom_sf()+
  geom_point(data = rook_Cuenca2_df,
            mapping = aes(x = x, y = y))+
  geom_segment(data = rook_Cuenca2_df,
            mapping = aes(x = x, xend = xend, y = y, yend = yend))+
  labs(x = "", y = "") +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Vecindad mediante el criterio de Rook (De la Torre)",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# VECINDAD MEDIANTE CRITERIO QUEEN (DE LA REINA)

queen_Cuenca2 <- poly2nb(as(shp_Cuenca1, 'Spatial'), queen = T)

queen_Cuenca2_df <- nb_to_df(queen_Cuenca2, coords)

# GRAFICAMOS LA VECINDAD MEDIANTE EL CRITERIO QUEEN (DE LA REINA)

ggplot(shp_Cuenca1) +
  geom_sf()+
  geom_point(data = queen_Cuenca2_df,
            mapping = aes(x = x, y = y))+
  geom_segment(data = queen_Cuenca2_df,
            mapping = aes(x = x, xend = xend, y = y, yend = yend))+
  labs(x = "", y = "") +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Vecindad mediante el criterio de Queen (De la Reina)",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# VECINDAD MEDIANTE CRITERIO k-cercanos (K-nearest) #

kn_Cuenca2 <- knn2nb(knearneigh(coords, k = 6))

kn_Cuenca2_df <- nb_to_df(kn_Cuenca2, coords)

# GRAFICAMOS LA VECINDAD MEDIANTE EL CRITERIO k-cercanos (K-nearest)

ggplot(shp_Cuenca1) +
  geom_sf()+
  geom_point(data = kn_Cuenca2_df, mapping = aes(x = x, y = y))+
  geom_segment(data = kn_Cuenca2_df,
            mapping = aes(x = x, xend = xend, y = y, yend = yend))+
  labs(x = "", y = "") +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "K-nearest)",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# AUTOCORRELACIÃN ESPACIAL GLOBAL -----
# INDICE I DE MORAN #

queen_Cuenca2_lw <- nb2listw(queen_Cuenca2)
moran.test(shp_Cuenca1$BD_Tasa_1, listw = queen_Cuenca2_lw)

# GRAFICO DE DISPERSION PARA VISUALIZAR EL SIGNO Y LA FUERZA DE CORRELACION ESPACIAL#

moran.plot(shp_Cuenca1$BD_Tasa_1, listw = queen_Cuenca2_lw,
           xlab = "Tasa de robo a domicilios",

```



```

        ylab = "Tasa de robo a domicilios (con corr. espacial)")

# PATRONES GLOBALES DE AUTOCORRELACION ESPACIAL I DE MORAN #
queen_Cuenca2_b_lw <- nb2listw(queen_Cuenca2, style = "B")

shp_Cuenca1 <- shp_Cuenca1 %>%
  mutate(lmororan = localmororan(x = BD_Tasa_1, listw = queen_Cuenca2_b_lw)[, 1],
         l Moran_pval = localmororan(x = BD_Tasa_R, listw = queen_Cuenca2_b_lw)[, 5]
  )

# MAPA ESTADISTICO LOCAL DE MORAN#
ggplot(shp_Cuenca1) +
  geom_sf(aes(fill = l Moran)) +
  labs(fill = " Estadística \n local I de Moran") +
  scale_fill_viridis_c() +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Autocorrelacion mediante el estadístico local I de Moran",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

# MAPA CLUSTER DE LISA PARA VER A DONDE SE AGRUPA LA AUCORRELACION#
shp_Cuenca1 <- shp_Cuenca1 %>%
  mutate(
    # Estandarizar la tasa de analfabetismo y el Moran local a sus valores medios:
    st_BD_Tasa_1 = BD_Tasa_1 - mean(BD_Tasa_1),
    st_l Moran = l Moran - mean(l Moran),
    # Crear la nueva variable categórica:
    cuadrante = case_when(
      l Moran_pval > 0.05 ~ "Insignificante",
      st_BD_Tasa_1 > 0 & st_l Moran > 0 ~ "Alto-Alto",
      st_BD_Tasa_1 < 0 & st_l Moran < 0 ~ "Bajo-Bajo",
      st_BD_Tasa_1 < 0 & st_l Moran > 0 ~ "Bajo-Alto",
      st_BD_Tasa_1 > 0 & st_l Moran < 0 ~ "Alto-Bajo"
    )
  )

# GRAFICAMOS MAPA #
ggplot(shp_Cuenca1, aes(fill = cuadrante)) +
  geom_sf() +
  labs(fill = "Cuadrante") +
  scale_fill_manual(values = c("red", "lightblue", "lightgreen", "lightyellow", "white")) +
  ggplot2::theme_bw() + # pinta de color transparente el fondo del mapa
  ggplot2::xlab("Longitud") +
  ggplot2::ylab("Latitud") +
  ggplot2::labs(
    title = "Patrones geográficos de agrupación para diferentes valores de Robo a domicilios",
    subtitle = "Zonas censales de la ciudad de Cuenca") +
  annotation_scale() +
  annotation_north_arrow(location='tr')

-----
# LEER LIBRARIES -----
require(pacman)
pacman::p_load(terra, spdep, tidyr, readxl, spatialreg, broom, viridis, raster, sf,
               ggrepel, ggspatial, hablar, fs, tidyverse, geodata, ggpubr,
               RColorBrewer, rgeoda, gtools, stringr, rgeos, glue)

g <- gc(reset = T)
rm(list = ls())
options(scipen = 999, warn = -1)

# FUNCION PARA CREAR TABLA DE CRUCE DE VARIABLES -----
make_bimoran <- function(v1, v2){

  cat(v1, ' ', v2, '\n')# v1 <- 'BD_Tasa_1'; v2 <- 'BD_PMI'
  bdt <- dplyr::select(bdta, ZON, v1, v2)
  colnames(bdt)[2:3] <- c('var1', 'var2')

  # To make the bimoran
  mrn <- local_bimoran(w = qnwg, df = bdt[c('var1', 'var2')], permutations = 999, significance_cutoff
= 0.1)
  mrn_lbls <- lisa_labels(mrn)
  mrn_clr <- setNames(lisa_colors(mrn), mrn_lbls)

  mrn_vle <- mean(mrn$lisa_vals)
  mrn_pvl <- mean(mrn$p_vals)
  mrn_rsl <- tibble(var1 = v1, var2 = v2, value = mrn_vle, pvalue = mrn_pvl)

  # To add the category column to the shapefile

```

```

    bdt <- mutate(bdt, cluster_num = lisa_clusters(mrn) + 1, cluster = factor(mrn_lbls[cluster_num],
levels = mrn_lbls))
    lbl <- tibble(cluster = mrn_lbls, clase = c('Insignificante', 'Alto - Alto', 'Bajo - Bajo', 'Bajo -
Alto', 'Alto - Bajo', 'Indefinido', 'Isolado'))
    bdt <- inner_join(bdt, lbl, by = 'cluster')

    # Column as a factor (reorder)
    bdt <- mutate(bdt, clase = factor(clase, levels = c('Alto - Alto', 'Alto - Bajo', 'Bajo - Alto',
'Bajo - Bajo', 'Insignificante', 'Isolado')))
    lbl <- mutate(lbl, color = c('#ffffff', '#ff0000', '#ffffb2', '#90ee90', '#add8e6', 'grey30',
'grey60'))
    lbl <- mutate(lbl, clase = factor(clase, levels = c('Alto - Alto', 'Alto - Bajo', 'Bajo - Alto',
'Bajo - Bajo', 'Insignificante', 'Isolado')))
    lbl <- drop_na(lbl)
    lbl <- lbl %>% arrange(clase)

    # COLORES UICOS -----
    unq <- bdt %>% pull(clase) %>% unique() %>% as.character()
    clr_unq <- lbl %>% filter(clase %in% unq) %>% pull(3)

    # Get the right names of each variable
    nm1 <- filter(vars, Tipo == v1) %>% pull(4) %>% unique()
    nm2 <- filter(vars, Tipo == v2) %>% pull(4) %>% unique()

    # GRAFICAR EL MAPA I DE MORAN LISA -----
    gmp <- ggplot() +
    geom_sf(data = bdt, aes(fill = clase, col = clase), lwd = 0.3) +
    scale_fill_manual(values = clr_unq) +
    scale_color_manual(values = clr_unq, guide = F) +
    geom_sf(data = bdt, fill = NA, col = 'grey40', lwd = 0.2) +
    geom_sf(data = prqs, fill = NA, col = 'grey50', lwd = 0.5) +
    geom_text_repel(data = prqs_lbls, aes(x = X, y = Y, label = nombre), size = 1.85, col = 'grey40')
+
    ggtitle(label = glue('Patrones geogrÁficos de agrupaci3n bivariado'),
    subtitle = glue('{nm1} - {nm2}')) +
    labs(x = 'Longitud', y = 'Latitud', fill = 'Cuadrante') +
    coord_sf() +
    theme_bw() +
    theme(panel.grid.minor = element_blank() +
    annotation_scale(height = unit(0.18, 'cm'),
    pad_x = unit(0.15, "cm"),
    pad_y = unit(0.15, "cm")) +
    annotation_north_arrow(location = 'tr',
    height = unit(1.1, "cm"), width = unit(0.7, 'cm'),
    pad_x = unit(0.1, "in"), pad_y = unit(0.2, "in"), # 0.2 # 0.3
    style = north_arrow_fancy_orientteering())

    # ESCRIBIR EL GEOPAQUETE FINAL -----
    dout <- glue('gpkg/moran_bivariado/{v1}_{v2}.gpkg')
    st_write(bdt, dout)

    # Finish
    cat('Done!\n')
    return(list(gmp, mrn_rsl))
}

# LEER LA BASE DE DATOS -----
vars <- read_excel('tbl/results/summary_resultado_modelos.xlsx')
bdta <- st_read('gpkg/spatial_bdata.gpkg')
prqs <- st_read('shp/purbanas.shp')

# CRUCE ENTRE PARROQUIAS Y BASES DE DATOS DE ZONAS URBANAS -----
bdta_prqs <- dplyr::select(bdta, gid, FUENTE, ZON, geom) %>% st_intersection(., prqs) %>%
dplyr::select(PARROQUIAS, ZON, geom)
bdta_prqs_tble <- st_drop_geometry(bdta_prqs) %>% as_tibble()

# PONER ETIQUETAS (NOMBRES) DE LAS PARROQUIAS -----
prqs_lbls <- prqs %>% st_centroid %>% st_coordinates %>% as_tibble()
prqs_lbls <- mutate(prqs_lbls, nombre = str_to_title(prqs$PARROQUIAS))

# CONTRUCCI3N DE I DE MORAN LISA -----
qnrwg <- queen_weights(bdta, order = 1)

# DOMICILIOS -----
domi <- filter(vars, Robo == 'Domicilios')
domi_vars <- domi %>% pull(Tipo) %>% unique()
domi_vars

domi_bmrn <- map2(.x = rep('BD_Tasa_1', length(domi_vars)), .y = domi_vars, .f = make_bimoran)
domi_bmrn_tble <- map(domi_bmrn, 2) %>% bind_rows()
domi_bmrn_grph <- map(domi_bmrn, 1)

# To save
out <- 'png/maps/bimoran'
ggsave(plot = domi_bmrn_grph[[1]], filename = glue('{out}/domi_BD_PMI.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)

```

```

ggsave(plot = domi_bmrn_grph[[2]], filename = glue('{out}/domi_BD_PCes.png'), units = 'in', width =
7, height = 4.6, dpi = 300)
ggsave(plot = domi_bmrn_grph[[3]], filename = glue('{out}/domi_BD_PEI.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)
ggsave(plot = domi_bmrn_grph[[4]], filename = glue('{out}/domi_BD_HTV.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)
ggsave(plot = domi_bmrn_grph[[5]], filename = glue('{out}/domi_BD_TIEpc.png'), units = 'in', width =
7, height = 4.6, dpi = 300)

out <- 'tbl/results'
write.csv(domi_bmrn_tble, glue('{out}/moran_bivariado_domicilios.csv'), row.names = FALSE)

# PERSONAS -----
prsn <- filter(vars, Robo == 'Personas')
prsn_vars <- prsn %>% pull(Tipo) %>% unique()

prsn_bmrn <- map2(.x = rep('BD_Tasa_R_', length(prsn_vars)), .y = prsn_vars, .f = make_bimoran)
prsn_bmrn_tble <- map(prsn_bmrn, 2) %>% bind_rows()
prsn_bmrn_grph <- map(prsn_bmrn, 1)

# To save
out <- 'png/maps/bimoran'
ggsave(plot = prsn_bmrn_grph[[1]], filename = glue('{out}/prsn_BD_HTV.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)
ggsave(plot = prsn_bmrn_grph[[2]], filename = glue('{out}/prsn_BD_Inmi_Ex.png'), units = 'in', width
= 7, height = 4.6, dpi = 300)
ggsave(plot = prsn_bmrn_grph[[3]], filename = glue('{out}/prsn_BD_PPH.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)
ggsave(plot = prsn_bmrn_grph[[4]], filename = glue('{out}/prsn_BD_PCes.png'), units = 'in', width =
7, height = 4.6, dpi = 300)

out <- 'tbl/results'
write.csv(prsn_bmrn_tble, glue('{out}/moran_bivariado_personas.csv'), row.names = FALSE)

# UNIDADES ECONOMICAS -----
undd <- filter(vars, Robo == 'Unidades economicas')
undd_vars <- undd %>% pull(Tipo) %>% unique()

undd_bmrn <- map2(.x = rep('BD_Robo_U_', length(undd_vars)), .y = undd_vars, .f = make_bimoran)
undd_bmrn_tble <- map(undd_bmrn, 2) %>% bind_rows()
undd_bmrn_grph <- map(undd_bmrn, 1)

# To save
out <- 'png/maps/bimoran'

ggsave(plot = undd_bmrn_grph[[1]], filename = glue('{out}/undd_BD_POdh.png'), units = 'in', width =
7, height = 4.6, dpi = 300)
ggsave(plot = undd_bmrn_grph[[2]], filename = glue('{out}/undd_BD_Inmi_Ex.png'), units = 'in', width
= 7, height = 4.6, dpi = 300)
ggsave(plot = undd_bmrn_grph[[3]], filename = glue('{out}/undd_BD_PPH.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)

out <- 'tbl/results'
write.csv(undd_bmrn_tble, glue('{out}/moran_bivariado_unidadeseconomicas.csv'), row.names = FALSE)

# VEHICULOS -----
vehi <- filter(vars, Robo == 'Vehiculos')
vehi_vars <- vehi %>% pull(Tipo) %>% unique()

vehi_bmrn <- map2(.x = rep('BD_Robo_ve', length(vehi_vars)), .y = vehi_vars, .f = make_bimoran)
vehi_bmrn_tble <- map(vehi_bmrn, 2) %>% bind_rows()
vehi_bmrn_grph <- map(vehi_bmrn, 1)

# To save
out <- 'png/maps/bimoran'

ggsave(plot = vehi_bmrn_grph[[1]], filename = glue('{out}/vehi_BD_Tasa_Dt.png'), units = 'in', width
= 7, height = 4.6, dpi = 300)
ggsave(plot = vehi_bmrn_grph[[2]], filename = glue('{out}/vehi_BD_PCes.png'), units = 'in', width =
7, height = 4.6, dpi = 300)
ggsave(plot = vehi_bmrn_grph[[3]], filename = glue('{out}/vehi_BD_PMI.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)
ggsave(plot = vehi_bmrn_grph[[4]], filename = glue('{out}/vehi_BD_PET.png'), units = 'in', width = 7,
height = 4.6, dpi = 300)

out <- 'tbl/results'
write.csv(vehi_bmrn_tble, glue('{out}/moran_bivariado_vehiculos.csv'), row.names = FALSE)
-----

```



epoch

Dirección de Bibliotecas y
Recursos del Aprendizaje

UNIDAD DE PROCESOS TÉCNICOS Y ANÁLISIS BIBLIOGRÁFICO Y
DOCUMENTAL

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 17 / 07 / 2023

INFORMACIÓN DEL AUTOR/A (S)

Nombres – Apellidos: *Edy Samuel Velasteguí Casco*

INFORMACIÓN INSTITUCIONAL

Instituto de Posgrado y Educación Continua

Título a optar: *Magíster en Matemática mención Modelación y Docencia*

f. Analista de Biblioteca responsable: Lic. Luis Caminos Vargas Mgs.



0067-DBRA-UTP-IPEC-2023