

**INFORME FINAL
PRACTICA EMPRESARIAL
PROGRAMA DE PRÁCTICAS UNIVERSITARIAS
PONTIFICIA UNIVERSIDAD JAVERIANA**

Nombre del estudiante: Santiago Nicolás García Herrera Número de cédula: 1019125749

Carrera: Economía

Nombre de la empresa: Departamento Nacional de Planeación Área de trabajo: Dirección de Desarrollo Digital

Fecha de presentación del informe:

Fecha de inicio de la práctica: 01/14/2020 Fecha de terminación de la práctica: 07/14/2020

TÉRMINOS DE REFERENCIA DEL INFORME

Para el desarrollo de este informe utilice las herramientas de apoyo anexas en el blackboard. Por favor seguir las instrucciones contenidas en cada punto, y presente toda la información relevante en este documento; incluyendo el "INFORME MENSUAL CUMPLIMIENTO PLAN DE PRÁCTICA / PLAN DE TRABAJO".

1. PROYECTO LÍDER

Haga un informe sobre el desarrollo y resultado final de su Proyecto Líder. Debe desarrollar cada uno de los siguientes puntos siguiendo las indicaciones descritas en el blackboard:

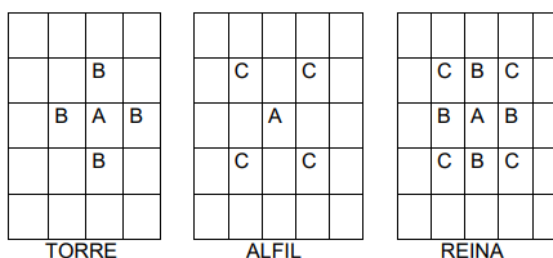
- a. **Objetivo General:**
Entender las dinámicas espaciales del crimen en la ciudad de Bucaramanga para diferentes periodos de tiempo dados
- b. **Objetivos específicos:**
Desarrollar un modelo de Econometría espacial que logre determinar si existen relaciones espaciales en la tasa de delitos de diferentes secciones de Bucaramanga y con ello lograr un entendimiento de las dinámicas del crimen para esta zona.
- c. **Metodología utilizada:**
Para el desarrollo de este proyecto se utilizó un mapa de Bucaramanga dividido en secciones DANE y cada una de estas secciones se le calculo su tasa de hurtos, homicidios y lesiones siendo los principales delitos acontecidos en la ciudad, Toda esta información ha sido provista por la Policía Nacional a través de la Dirección de Justicia y Seguridad puesta en disposición para el desarrollo de un modelo de predicción del crimen. La metodología utilizada se basa en los métodos de la estadística espacial (LeSage, 1999) los cuales basan su análisis en las vecindades de una zona geográfica y otros en patrones de puntos los cuales se escapan del alcance de este trabajo, esta se define como el área de la estadística que estudia fenómenos que se encuentren georreferenciados, esto puede ser a través de coordenadas o teniendo en cuenta algún tipo de desagregación como barrios o localidades. Dado esto la econometría espacial se desarrolla a partir de estos métodos y nace en la década de los ochenta, la

intención o necesidad de hacer esto fue para lograr incorporar las relaciones espaciales como fuente de explicación de las variables económicas tales como el nivel de empleo, salarios o el PIB per cápita.

Los modelos espaciales en econometría en su forma matemática tienen similitudes con los de series de tiempo y es que al igual que en estos se estudia una variable sobre misma, en series de tiempo la variable se estudia sobre su pasado, en los modelos espaciales tenemos que se estudia el resultado de la variable contra resultados de esta teniendo en cuenta vecinos cercanos de la misma lo cual genera una matriz de adyacencia muy similar a un modelo de datos panel. Por esto los modelos son autorregresivos solo que en los modelos espaciales a diferencia de los de series de tiempo se tiene en cuenta la cercanía geográfica de las observaciones, lo que se define como los vecinos cercanos.

En la Figura 1 se observa la matriz de vecinos, todo mapa se puede dividir en una grilla donde se puede observar los vecinos de un polígono específico. En un modelo de econometría espacial se consideran tres tipos de vecinos: En forma de torre cuando se considera buscar vecinos en las horizontales de la zona de interés, de alfil cuando se busca en las diagonales y en forma de reina cuando estos se buscan en toda la periferia de la zona de interés.

Figura 1 : Tipos de vecinos



En estos vecinos solo se considera el inmediatamente más cercano o de primer orden ya que no se interesa por buscar vecinos que estén más allá de una grilla esto se puede alterar si se toman otros modelos de distancia como la euclidiana, pero esto sale del alcance de este proyecto y se tomara como una recomendación. Para el modelo se debe escoger cuál de estos tres tipos de vecinos es el más apropiado, es decir que se le puede considerar como un hiperparámetro que en la literatura son aquellos parámetros que no son estimados por el modelo, sino que sirven como instrumento de afinación de este.

En la Figura 2 se observa después de haberle asignado a cada una de las localidades un número que se consideran vecinos tipo torre.

Figura 2: Matriz de vecinos ejemplo

	1	2B	3	
	4B	5A	6B	
	7	8B	9	

En la Figura 3 se genera la matriz correspondiente, se observa que la diagonal se encuentra llena de ceros esto se debe a que una observación “no puede ser vecindad de sí misma” pero también es porque en el sentido de la futura regresión una observación no puede depender estrictamente de sí misma, si una localidad es vecina de otro tipo torre a este se le coloca un uno.

Figura 3: Matriz de vecinos ejemplo

Localidad	1	2	3	4	5	6	7	8	9
1	0	1	0	1	0	0	0	0	0
2	1	0	1	0	1	0	0	0	0
3	0	1	0	0	0	1	0	0	0
4	1	0	0	0	1	0	1	0	0
5	0	1	0	1	0	1	0	1	0
6	0	0	1	0	1	0	0	0	1
7	0	0	0	1	0	0	0	1	0
8	0	0	0	0	1	0	1	0	1
9	0	0	0	0	0	1	0	1	0

Ya teniendo todo esto se procede a la construcción de la matriz de vecinos W la cual dependiendo de la distancia y de la distribución de vecinos presentada designa unos pesos a cada observación, entre más cerca estén dos observaciones, mayor será la ponderación.

La matriz W es fundamental para todos los modelos de econometría espacial ya que esta se puede incorporar tanto a los variables explicativas como a los términos de perturbación. El modelo escogido para el desarrollo de este trabajo se encuentra entre los más simples y es el que se muestra en la Ecuación 1

Ecuación 1: Regresión espacial

$$y_i = \rho W y_i + \beta x_i + e_i$$

Donde ρ captura la relación espacial entre las observaciones y es el parámetro a Estimar las demás partes de la ecuación siguen teniendo las interpretaciones de una regresión normal.

Para el modelo de Bucaramanga en el año 2016 se escogió una matriz tipo reina (siendo la más recomendada para este tipo de análisis de crimen) (al, 2016) y en la Figura 4 se observan el tipo de conexiones que esto genera entre las distintas secciones del mapa.

Figura 4: Bucaramanga matriz tipo reina

Bucaramanga vecinos conexiones



d. Herramientas informáticas y/o tecnología utilizada

Para el desarrollo de este proyecto se utilizó el lenguaje R y como librerías de apoyo de uso el paquete **tidyverse** junto con **spdep**, **maptools** y **tmap** que incluyen métodos para el manejo de datos georreferenciados.

e. Indicadores

Como todo modelo econométrico los espaciales tienen sus propias métricas que permiten dar un diagnóstico inicial de los datos. En este caso y dado que lo que se busca estudiar es la existencia de autocorrelaciones espaciales se usará el índice de Moran que se construye de la siguiente manera como lo expresa la Ecuación 2:

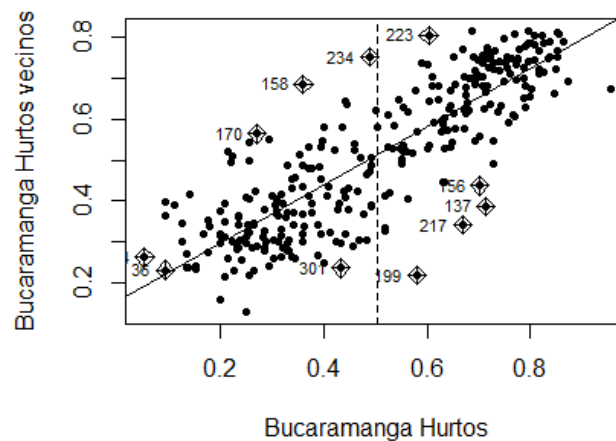
Ecuación 2: Índice de Moran

$$I = \frac{R}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{X})(x_j - \bar{X})}{\sum_i (x_i - \bar{X})^2}$$

Donde R representa el número de regiones y w son las ponderaciones dadas por la matriz W bajo este estadístico se puede construir la hipótesis nula de no autocorrelación espacial.

En la Figura 5 se muestra el índice de Moran para la tasa de Hurtos en las distintas secciones de Bucaramanga.

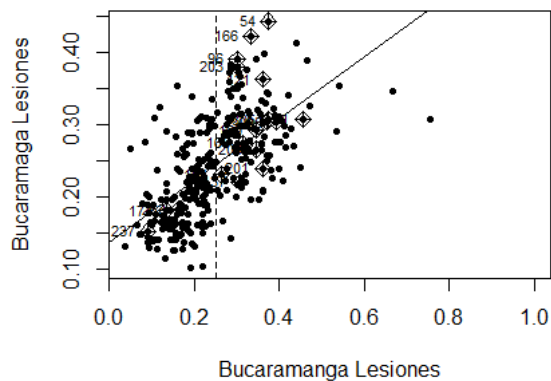
Figura 5: Índice de Moran Hurtos



La Figura 5 se puede leer en cuadrantes y en sentido antihorario en el primero nos muestra que zonas donde la tasa de hurtos es alta sus vecinos también presentan una tasa de hurtos alta, el siguiente en sentido antihorario muestra las zonas donde los hurtos son bajos presentan vecinos donde los hurtos tienen un nivel alto, luego estaría el tercero donde zonas que presentan una baja tasa de hurtos sus vecinos también presentan una baja tasa de hurtos por último el cuadrante cuatro muestra las zonas donde la tasa de hurtos es alta los vecinos presentan una tasa de hurtos baja. Como se observa la mayoría de las observaciones se encuentran en los extremos lo cual da indicios de que la tasa de hurtos si presentan una correlación espacial.

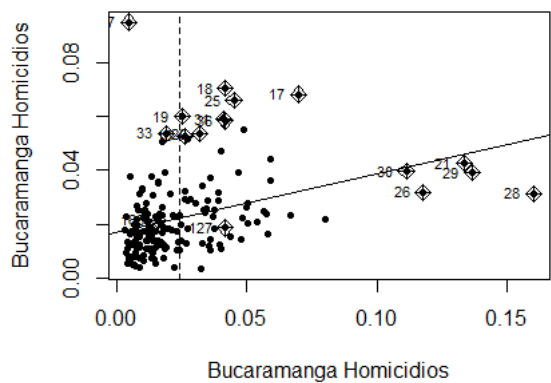
En la Figura 6 observamos el mismo comportamiento para la tasa de Lesiones. Donde incluso la relación espacial es mayor que en la tasa de hurtos.

Figura 6: Índice de Moran Lesiones



En la Figura 7 se observa el índice de Moran para la tasa de Homicidios, dado que en la muestra esto son los que menos se presentan el índice de Moran solo permite concluir que zonas donde la tasa de homicidios es baja también presentan una baja tasa de homicidios.

Figura 7: Índice de Moran Homicidios



En la Tabla 1 se resumen el índice de Moran para cada tasa de delitos junto con su valor p con ello se permite concluir a todos los niveles de significancia que existe evidencia estadística de que existe una relación espacial en las distintas tasas de hurtos en la ciudad de Bucaramanga

Tabla 1: Resumen índice de Moran

	Índice de Moran	Valor P
Tasa de Hurtos	0.705081499	p-value < 2.2e-16
Tasa de Lesiones	0.317639726	p-value < 2.2e-17
Tasa de Homicidios	0.216163379	p-value < 2.2e-18

f. Resultados logrados

Ya con esto se puede especificar un modelo de regresión como el mostrado en la Ecuación 1 donde como variable dependiente tenemos la tasa de hurtos, como variables explicativas tenemos los vecinos de la tasa de hurtos y la tasa de lesiones.

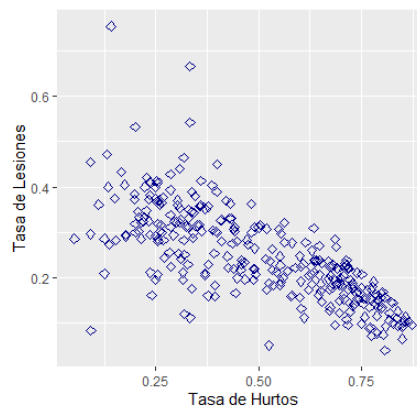
En la Tabla 2 se muestra la salida de la regresión, el parámetro de vecinos tasa de hurtos tiene un valor de 0.67 esto implica que la tasa de hurtos es explicada en un 67 % por sus vecinos.

Tabla 2: Salida de la Regresión

	Estimación de coeficiente	Error estandar	Valor z	Valor P
Intercepto	0.355096	0.030403	11.680	0
Tasa de Lesiones	-0.810520	0.065030	-12.464	0
Vecinos tasa de Hurtos	0.67835			

En la Tabla 2 también se observa que existe una relación negativa con respecto la tasa de lesiones la cual se muestra en la Figura 8

Figura 8: relación lesiones hurto



g. Impacto

El modelo puede usarse como una primera aproximación para ayudar a encontrar una mejor para la administración de las fuerzas policial o como diagnóstico de recomendación de políticas públicas enfocadas en Seguridad

h. Conclusiones

Dados los resultados del índice de Moran y el modelo de regresión se encuentra que el crimen no se distribuye aleatoriamente en la ciudad y encontramos patrones o centros donde se concentra. También se encuentra que la tasa de hurtos depende espacialmente de los resultados de esta en otras zonas geográficas y que esta tiene una relación negativa con la tasa de delitos

i. Recomendaciones

Se puede considerar otros modelos espaciales o incluso lograr incorporar temporalidad a los modelos o más variables explicativas que logren capturar mejor las relaciones que puedan existir a nivel espacial.

Bibliografía

- al, L. Q. (2016). *ECONOMETRÍA APLICADA UTILIZANDO R*. Mexico: Universidad Nacional Autónoma de México.
- LeSage, J. P. (1999). *The Theory and Practice of Spatial Econometrics*. Toledo: University of Toledo.

2. DIAGNOSTICO DEL ÁREA

a. Realice un diagnostico DOFA (debilidades, oportunidades, fortalezas y amenazas) acerca del área o departamento donde realizó la práctica.

Fortalezas: El equipo es muy variado y altamente calificado para el desarrollo de los distintos proyectos junto a que va acompañado con un buen sistema de organización interno.

Oportunidades: El equipo se encuentra muy especializado en ciertas habilidades por lo tanto una oportunidad seria ampliar su rango para así abarcar más proyectos que requiera el DNP.

Debilidades: Puede que se generen proyectos que requieran más tiempo de desarrollo que el estipulado en un primer momento debido a su complejidad y debido al tiempo dado sea difícil llevarlos a buen término.

Amenazas: La llegada de proyectos de rápida ejecución que puedan exigir una mayor dedicación de la esperada y que por lo tanto termine afectando el cronograma y desarrollo de otros proyectos.

b. Plantee las acciones de mejoramiento y recomendaciones, que considere necesarias con base en el diagnostico anterior, para que el área o departamento pueda perfeccionar algunos aspectos y logre resultados superiores.

Para solucionar las debilidades en las primeras semanas del proyecto se haría de nuevo para considerar eventos nuevos que se pudieron desarrollar durante ese periodo de tiempo para saber si existe la posibilidad de cambiar los plazos de entrega del proyecto en virtud de este

3. COMPETENCIAS

Mencione los principales aportes que obtuvo como resultado de su práctica, a nivel:

a. Personal

Un entendimiento mayor de cuáles son las principales fortalezas y aspectos para mejorar como por ejemplo ser mucho más riguroso en la presentación de informes. También se desarrollaron habilidades de comunicación con el equipo aspecto que debe seguir en mejora. En general fue una muy buena experiencia que aporte muchas cualidades y herramientas que serán de utilidad para el futuro laboral.

b. Académico

Aprender nuevos modelos estadísticos ya sean el área del aprendizaje de maquina o de la econometría en sí que no se alcanzaron a aprender durante la carrera y que insista a la profundización de estos

c. Profesional

Aparte de una profundización en herramientas como R o Python y demás herramientas tecnológicas también se hizo un fuerte énfasis en cómo se deben elaborar informes y presentaciones. Se mostro nuevas herramientas que pueden ser muy útiles para un futuro

INFORME MENSUAL CUMPLIMIENTO PLAN DE PRÁCTICA / PLAN DE TRABAJO

Explique en detalle las actividades y/o funciones desarrolladas durante el sexto mes de práctica, completando el siguiente cuadro.

SEXTO MES

fecha corte:

Espacio para ser diligenciado por el estudiante	No.	Funciones de práctica registradas en plan de práctica / o de trabajo.	Actividades, avances o resultados ejecutados en el último corte.
	1	Apoyo en el desarrollo de procesos, herramientas, métodos y técnicas, bajo entornos de software libre, que procesen Información estructurada y no estructurada.	Se ayudo en la elaboración de un tablero de visualización bajo la librería Shiny de R para mostrar los modelos de crimen
	2	Apoyo en la aplicación de metodologías convencionales y no convencionales para el procesamiento de datos que permita la extracción de información relevante de los datos	Se continuo en el desarrollo de funciones para explotar las capacidades de la librería twint de Python que permite hacer scrapping en twitter
	3	Apoyo en el análisis de los resultados de los proyectos de analítica de datos finalizados y en desarrollo, para la elaboración de informes, presentaciones y/o reportes	
	4	Apoyo en la elaboración y revisión de documentos relacionados con las metodologías empleadas por el equipo de científicos de datos.	Se ayudo en los documentos referentes al proyecto de crimen donde se analizan los resultados obtenidos hasta el momento
	5		<i>(inserte tantas líneas como sea necesario)</i>
	No.	Otras funciones de prácticas adicionales a las registradas en plan de práctica / o de trabajo.	Otras actividades, avances o resultados ejecutados el último corte
	6		
	7		
	8		
	9		
	10		<i>(inserte tantas líneas como sea necesario)</i>

Tutor:	Observaciones generales sobre el cumplimiento y desempeño del estudiante en el último corte.
---------------	-----------------------------------------------------------------------------------------------------

Espacio para ser diligenciado por el Tutor responsable de la práctica del estudiante en la Entidad.	Santiago en general tuvo una práctica bastante buena con algunos puntos concretos en los cuales puede mejorar, tales como la presentación de documentos, presentación de resultados y focalizar el trabajo a un objetivo concreto.		
	Evaluación del desempeño y ejecución de actividades esperadas en el último corte.		
	Escala de Evaluación de Desempeño	Excelente: 90% - 100% Bueno: 75% - 89% Aceptable: 60% - 74% Deficiente: 0% - 59%	<u>87</u> %

Nombre del Tutor/Jefe: Iván Mauricio Durán Pabón

Firma del tutor/jefe: _____



Nombre del monitor/docente: Oscar Yovanny Barrera

Firma del monitor/docente: _____

Firma del estudiante: _____

