



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado

Facultad de Ciencias Matemáticas

Unidad de Posgrado

**Análisis exploratorio de datos espaciales de las
enfermedades respiratorias en consulta externa del
Hospital Santa Rosa – 2006**

TESIS

Para optar el Grado Académico de Magíster en Bioestadística

AUTOR

Carla Patricia ZÚÑIGA VILCA

ASESOR

Mg. Emma Norma CABBILLO MOYANO

Lima, Perú

2018

**ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES DE LAS
ENFERMEDADES RESPIRATORIAS EN CONSULTA EXTERNA DEL
HOSPITAL SANTA ROSA – 2006**

Carla Patricia Zuñiga Vilca

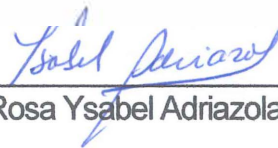
Tesis presentada a consideración del jurado examinador nombrado por la Unidad de Postgrado de la Facultad de Ciencias Matemáticas, de la Universidad Nacional Mayor de San Marcos, como parte de los requisitos para obtener el grado académico de Magíster en Bioestadística.

Aprobada por:



Dra. Ilse Janine Villavicencio Ramírez

Presidente



Mg. Rosa Ysabel Adriazola Cruz

Miembro



Mg. Olga Lidia Solano Dávila

Miembro



Mg. Rosa Fátima Medina Merino

Miembro



Mg. Emma Norma Cambillo Moyano

Miembro Asesor

ACTA DE SUSTENTACIÓN DE TESIS DE GRADO ACADÉMICO DE MAGÍSTER


Siendo las, *17:00* horas del día jueves cuatro de octubre del dos mil dieciocho, en el Auditorio de la Facultad de Ciencias Matemáticas, el Jurado Evaluador de Tesis, Presidido por la Dra. Ilse Janine Villavicencio Ramírez e integrado por los siguientes miembros, Mg. Rosa Ysabel Adriazola Cruz (Jurado Informante), Mg. Olga Lidia Solano Dávila (Jurado Informante), Mg. Rosa Fátima Medina Merino (Jurado Evaluador) y la Mg. Emma Norma Cambillo Moyano como Miembro Asesor, se reunieron para la sustentación de la tesis titulada: «ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES DE LAS ENFERMEDADES RESPIRATORIAS EN CONSULTA EXTERNA DEL HOSPITAL SANTA ROSA - 2006» presentada por la Bachiller Carla Patricia Zúñiga Vilca, para optar el Grado Académico de Magíster en Bioestadística.


Luego de la exposición de la graduanda, los Miembros del Jurado hicieron las preguntas correspondientes, así como las observaciones e inquietudes acerca del trabajo de tesis, a las cuales la Bachiller Carla Patricia Zúñiga Vilca respondió con acierto y solvencia, demostrando pleno conocimiento del tema.


A continuación se realizó la calificación correspondiente, según tabla adjunta, resultando la Bachiller Carla Patricia Zúñiga Vilca aprobada con el calificativo de*18*.....
.....*Dieciocho*.....

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del Grado Académico de **Magíster en Bioestadística** a la **Bachiller Carla Patricia Zúñiga Vilca**.


Siendo las *18:00* horas, se levantó la sesión, firmando para constancia la presente Acta.


Dra. Ilse Janine Villavicencio Ramírez
Presidenta


Mg. Rosa Ysabel Adriazola Cruz
Miembro


Mg. Olga Lidia Solano Dávila
Miembro


Mg. Rosa Fátima Medina Merino
Miembro


Mg. Emma Norma Cambillo Moyano
Miembro Asesor

FICHA CATALOGRÁFICA

ZÚÑIGA VILCA, CARLA PATRICIA

“ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES DE LAS ENFERMEDADES RESPIRATORIAS EN CONSULTA EXTERNA DEL HOSPITAL SANTA ROSA – 2006”

XII, 60p, 29.7 cm (UNMSM, Magister, Bioestadística, 2018)

Tesis de Maestría, Universidad Nacional Mayor de San Marcos.

Facultad de Ciencias Matemáticas.

Unidad de Posgrado, UNMSM/FCM.

DEDICATORIA

Dedicado a Dios, a mi hija Brenda y mis padres quienes son mi principal fuente de inspiración, por su invaluable apoyo y su paciencia durante el desarrollo de la presente tesis.

AGRADECIMIENTOS

Quiero expresar en primer lugar mi más sincero agradecimiento a mi asesora Mg. Emma Cambillo M. por su gran apoyo, predisposición y dedicación en el desarrollo de la presente tesis, sin los que no hubiera sido posible la misma.

También deseo expresar mi agradecimiento a todos los docentes de la Unidad de Posgrado de la Facultad de Ciencias Matemáticas de la UNMSM por sus enseñanzas y a la Universidad Nacional Mayor de San Marcos, por ser mi casa de estudios.

Asimismo, quiero agradecer a mis padres, no tengo palabras suficientes para expresarles mi gratitud y afecto; y a mi prima Clorinda por su apoyo constante.

Finalmente, quiero agradecer al Dr. Cleto de la Torre Dueñas por su ayuda para poder culminar la tesis.

ÍNDICE GENERAL

CAPÍTULO I. INTRODUCCIÓN	1
1.1 SITUACIÓN PROBLEMÁTICA	1
1.2 FORMULACIÓN DEL PROBLEMA	3
1.3 JUSTIFICACIÓN DE LA INVESTIGACIÓN	3
1.4 OBJETIVOS.....	4
1.4.1 OBJETIVO GENERAL.....	4
1.4.2 OBJETIVOS ESPECÍFICOS	4
1.5 HIPÓTESIS.....	4
1.5.1 Hipótesis General.....	4
1.5.2 Hipótesis Específicas	5
CAPÍTULO II. MARCO TEÓRICO	6
2.1 ANTECEDENTES DEL PROBLEMA	6
2.2 CONCEPTOS PREVIOS.....	11
2.2.1 Datos Georreferenciados	11
2.2.2 Tipos de datos.....	12
2.2.3 Variables	13
2.2.4 Calidad de los datos.....	14
2.2.5 Esquema Reticular o Lattice	14
2.3 ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES.....	15
2.3.1 Historia	15
2.3.2 Análisis exploratorio de datos espaciales.....	16
2.3.3 Métodos gráficos del AEDE	18
2.3.4 Medidas de asociación espacial global	27
2.3.5 Mapas de riesgos relativos.....	29
2.3.6 Mapas Probabilísticos y Suavizados con Estimación Bayesiana ..	30
2.3.7 Procesos espacial Poisson no homogéneo.....	31
2.3.8 Pruebas de aleatoriedad espacial completa.....	31
2.4 ENFERMEDADES DEL SISTEMA RESPIRATORIO.....	33
2.4.1 Sistema Respiratorio.....	33
2.4.2 Enfermedades del Sistema Respiratorio	33

2.5	MARCOS CONCEPTUALES.....	35
CAPÍTULO III. METODOLOGÍA		38
3.1	TIPO Y DISEÑO DE INVESTIGACIÓN	38
3.2	UNIDAD DE ANÁLISIS.....	38
3.3	POBLACIÓN DE ESTUDIO	39
3.4	TÉCNICAS DE RECOLECCIÓN DE LOS DATOS.....	39
3.5	IDENTIFICACIÓN DE VARIABLES.....	39
3.7	ANÁLISIS E INTERPRETACIÓN DE LA INFORMACIÓN.....	40
CAPÍTULO IV. ANÁLISIS		
RESULTADOS.....		46
4.1	ANÁLISIS EXPLORATORIO DE DATOS	42
4.2	ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES.....	44
4.2.1	Mapeo de casos	44
4.2.2	Mapas para detectar Outliers	45
4.2.3	Mapas de tasas y suavizado bayesiano empírico	46
4.2.4	Pruebas de aleatoriedad espacial completa.....	49
CONCLUSIONES		54
REFERENCIAS BIBLIOGRÁFICAS		55
ANEXOS.....		58

ÍNDICE DE FIGURAS

Figura 1	Mapa de quintiles.....	20
Figura 2	Mapa de la desviación típica.....	21
Figura 3	Mapa dinámico acumulativo.....	21
Figura 4	Mapa condicional del índice de infancia condicionado a la tasa de paro y tasa de instrucción superior	22
Figura 5	Mapa de dispersión de Morán.....	23
Figura 6	Mapa de dispersión de Morán Multivariante	24
Figura 7	Mapa de cajas.....	25
Figura 8	Mapa de percentiles.....	26
Figura 9	Cartograma	26
Figura 10	Mapa de LISA	27
Figura 11	Procesos de la respiración.....	33
Figura 12	Distribución de casos de infecciones respiratorias del Hospital Santa Rosa – 2006	42
Figura 13	Distribución de los casos de rinofaringitis aguda (RA), Asma (AS) y Bronquitis Aguda (BA) de consulta externa del Hospital Santa Rosa - 2006	43
Figura 14	Diagrama de cajas de los casos de RA, AS y BA de consulta externa del Hospital Santa Rosa – 2006	43
Figura 15	Número de casos de pacientes con Rinofaringitis Aguda (lado izquierdo) y Asma (lado derecho).	44
Figura 16	Número de casos de pacientes con BA (lado izquierdo) y para el total de las tres infecciones infecciones respiratorias (lado derecho)	45
Figura 17	Mapa de Cajas de la distribución de la RA, AS y BA de consulta externa del Hospital Santa Rosa - 2006	46
Figura 18	Mapa de tasa de infecciones respiratorias y Suavizado empírico bayesiano de consulta externa del Hospital Santa Rosa - 2006	47

Figura 19	Mapa de tasa de IRAS y Suavizado empírico bayesiano para los distritos de Lima Metropolitana, 2015.	48
Figura 20	Diagrama de dispersión de la distancia en kilómetros versus la tasa de IRAS atendidas por el Hospital Santa Rosa 2006.....	49
Figura 21	Mapa de conectividades e histograma de vecindades.....	51
Figura 22	Scatterplot de Moran para los casos de IRAS del Hospital Santa Rosa al 2006 y para el total de registrados en Lima Metropolitana al 2015.	52
Figura 23	Detección de clúster espaciales para la tasa de incidencia de infecciones respiratorias de las consultas externas del Hospital Santa Rosa 2006 y para el total de caso de IRAS de Lima Metropolitana al 2015	53

ÍNDICE DE TABLAS

Tabla 1	Esquema de datos georreferenciados	12
Tabla 2	Paquetes informáticos utilizados en el AEDE	16
Tabla 3	Métodos gráficos del AEDE según la tendencia espacial y atípicos espaciales.	18
Tabla 4	Prueba de aleatoriedad espacial completa, para los casos de IRAS atendidos por el Hospital Santa Rosa 2006 y total de casos de IRAS en Lima Metropolitana	50
Tabla 5	Estadísticos descriptivos de rinoфарингитis aguda.....	59
Tabla 6	Estadísticos descriptivos de asma no especificado	59
Tabla 7	Estadísticos descriptivos de bronquitis aguda no especificada	60

RESUMEN

ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES DE LAS ENFERMEDADES RESPIRATORIAS EN CONSULTA EXTERNA DEL HOSPITAL SANTA ROSA – 2006

Carla Patricia Zuñiga Vilca

2018

Asesora : Mg. Emma Norma Cambillo Moyano
Grado obtenido : Magíster en Bioestadística

El presente trabajo de investigación, plantea una aplicación del Análisis Exploratorio de Datos Espaciales (AEDE) para evaluar el tipo de patrón espacial que siguen los casos de pacientes con Infecciones Respiratorias Agudas (IRAS) del año 2006 en la ciudad de Lima - Perú y que fueron atendidos por el Hospital Santa Rosa. De igual forma también se analizan el total de casos de IRAS ocurridos en el año 2015 en Lima Metropolitana.

Los resultados mostraron que los casos de IRAS atendidos por el Hospital Santa Rosa se distribuyen de forma no homogénea a lo largo de Lima Metropolitana, siendo en mayor intensidad en aquellos distritos más cercanos a dicho hospital y específicamente en el Distrito de Pueblo Libre.

A partir del índice de LISA y el índice de Morán se encontraron la presencia de clústeres significativos de distritos con alta demanda de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa. Estos distritos son: Pueblo Libre, Breña, Jesús María, Magdalena del Mar, San Miguel y Lima. A su vez se identificó un patrón espacial distinto para el total de casos de infecciones respiratorias agudas (IRAS) registrados en Lima Metropolitana al 2015. En estos casos son los distritos de San Bartolo, Chaclacayo, Punta Negra y Pucusana las que presentaron las tasas más altas de IRAS por encima de 36% de su población expuesta y representan las zonas de mayor riesgo para contraer infecciones respiratorias.

Palabras Claves: Análisis Exploratorio de Datos Espaciales, Infecciones Respiratorias Agudas, Índice de LISA, Índice de Morán.

ABSTRACT

EXPLORATORY SPATIAL DATA ANALYSIS OF RESPIRATORY DISEASES IN EXTERNAL CONSULTATION OF SANTA ROSA HOSPITAL - 2006

Carla Patricia Zuñiga Vilca

2018

Advisor : Mg. Emma Norma Cambillo Moyano
Obtained degree : Magister in Biostatistics

This research work proposes an application of the Exploratory Spatial Data Analysis (AEDE) to evaluate the type of spatial pattern followed by patients with Acute Respiratory Infections (IRAS) in 2006 in the city of Lima - Peru. They were treated by the Santa Rosa Hospital. Similarly, the total number of IRAS cases that occurred in 2015 in the Lima Metropolitan Area are also analyzed.

The results showed that the cases of IRAS treated by the Santa Rosa Hospital are distributed in an inhomogeneous way throughout the city of Lima, being in greater intensity in those districts closest to that hospital and specifically in the District of Pueblo Libre. From the LISA index and the Morán index, the presence of significant clusters of districts with high demand of the three most prevalent diseases of the respiratory system in the outpatient clinic of Santa Rosa Hospital were found. These districts are: Pueblo Libre, Breña, Jesus Maria, Magdalena del Mar, San Miguel and Lima. In turn, a different spatial pattern was identified for the total number of cases of acute respiratory infections (IRAS) registered in the Lima Metropolitan Area in 2015.

In these cases, the districts of San Bartolo, Chaclacayo, Punta Negra and Pucusana presented the higher rates of IRAS above 36% of its exposed population and represent the highest risk areas for respiratory infections.

Keys Words: Exploratory Analysis of Spatial Data, Acute Respiratory Infections, LISA's Index, Moran's Index.

CAPITULO I

INTRODUCCIÓN

1.1 SITUACIÓN PROBLEMÁTICA

Las enfermedades respiratorias continúan en aumento debido a la exposición del sistema respiratorio a contaminantes presentes en el aire, inhalados de manera voluntaria o involuntaria tal como el humo del tabaco (cigarrillo), combustibles de biomasa como el humo de la leña, sustancias alergénicas y contaminantes del ambiente general y ocupacional.

En Lima Metropolitana las enfermedades asociadas con la contaminación del aire son las que se relacionan con las vías respiratorias.

Las principales causas de atenciones de morbilidad llevadas a cabo por el Ministerio de Salud – MINSA, para el período 2002 – 2005 son las infecciones agudas de las vías respiratorias superiores (que incluye a los resfriados comunes, sinusitis aguda, faringitis, amigdalitis, laringitis aguda, traqueítis aguda, entre otras). Pero, más aún, las otras infecciones agudas de las vías respiratorias inferiores (bronquitis, enfisemas, asma, entre otras) y las enfermedades crónicas de las vías respiratorias inferiores también constituyen otras causas principales de morbilidad tanto para la provincia de Lima como para el Callao.

Según el “Análisis de la Situación de Salud - Guía Para El Análisis Del Proceso Salud – Enfermedad, 2002” el incremento estacional de infecciones respiratorias agudas (IRAS) con el inicio de las temporadas frías es un patrón que resulta familiar para todos. Con el conocimiento de cuando

ocurrirán estos brotes de IRAS, los servicios de salud pueden movilizar sus recursos para prevenir altas incidencias y la aparición de casos complicados.

Según Gonzales y otros (2014) indicaron que uno de los principales contaminantes del aire es el material particulado menor de 2,5 μ (micras), en la ciudad de Lima, anualmente 2300 muertes prematuras son atribuibles a este contaminante. La contaminación del aire en Lima se relaciona directamente con enfermedades respiratorias crónicas en los escolares; así, los escolares de Lima Metropolitana expuestos a un alto tráfico vehicular cerca de casa y/o escuela desarrollan en mayor proporción síntomas de asma y atopía con relación a quienes viven más alejados.

Durante el 2009 en la provincia de Lima, las infecciones de vías respiratorias agudas (23.5%), representaron la principal causa de morbilidad de la consulta externa lo que constituye la cuarta parte de los casos según reportó “El Análisis de Situación de Salud de la Provincia de Lima 2011”. Asimismo, en el periodo 2008-2009, la principal causa de morbilidad en niños fueron las infecciones de las vías respiratorias agudas, a pesar que se observó una disminución de su proporción respecto al periodo 2002-2003. En cambio, en los adolescentes y adultos, las infecciones de las vías respiratorias agudas conservan el primer lugar.

La Oficina Ejecutiva de Planeamiento Estratégico del Hospital Santa Rosa publicó el documento “El Plan Operativo 2008” en el cual indica que el distrito que demanda el mayor porcentaje de atenciones es Pueblo Libre (24.86%), le siguen en orden correlativo los distritos del Cercado de Lima, San Miguel, San Martín de Porres, Jesús María y Magdalena. Así mismo, la principal causa de atenciones que brinda el hospital en consulta externa son las enfermedades del sistema respiratorio (18%) para los pacientes que proceden de los diferentes distritos de Lima Metropolitana.

La Dirección General de Investigación y Asuntos Ambientales del SENAMHI, en el marco del Proyecto “Implementación de un Servicio de Pronóstico de la Calidad de Aire en la zona Metropolitana de Lima y Callao 2011 - 2015”, ha implementado 4 estaciones automáticas para el monitoreo de la calidad del aire, ubicados en los Distritos de Ate, San Borja, Jesús María y Santa Anita,

estas estaciones forman parte de la Red de Monitoreo de la Calidad del Aire para Lima.

El Servicio Nacional de Meteorología e Hidrología (SENAMHI) realiza el monitoreo atmosférico de la calidad del aire a través de la recolección o muestreo de contaminantes del aire ambiental a través de 10 estaciones fijas de monitoreo de calidad del aire las cuales se ubican en Campo de Marte, Ate, San Borja, Santa Anita, Villa María del Triunfo, San Juan de Lurigancho, Huachipa, Carabaylo, San Martín de Porres y Puente Piedra.

De manera internacional Hernández (2000), Romero (2004) y Oryazún (2010) publicaron artículos relacionados a enfermedades respiratorias ocasionadas por la contaminación ambiental.

1.2 FORMULACION DEL PROBLEMA

¿Existirá algún patrón espacial de la morbilidad de enfermedades del sistema respiratorio en consulta externa del Hospital Santa Rosa en el año 2006 y en el año 2015?

1.3 JUSTIFICACIÓN DE LA INVESTIGACIÓN

Las herramientas del Análisis Exploratorio para Datos Espaciales (AEDE) son adecuadas para analizar fenómenos que ocurren en un espacio determinado considerando su ubicación y las relaciones entre estos. A la vez que permiten analizar la distribución espacial de la variable la morbilidad de enfermedades del sistema respiratorio.

De esta forma el AEDE servirá para determinar las zonas con alta frecuencia de enfermedades respiratorias y de esta forma permitirá al Hospital Santa Rosa adoptar medidas de promoción y prevención de salud para elevar la calidad de vida de la población para evitar que la tasa de morbilidad siga creciendo, además podrá minimizar el ingreso hospitalario y la reducción de gastos en medicamentos.

Conocer como están distribuidas en el espacio las infecciones respiratorias ayudarán principalmente a este hospital en temas de focalización de su

demanda potencial y para implementar políticas de monitoreo para controlar la contaminación ambiental de acuerdo al patrón espacial que presenten la morbilidad de las enfermedades del sistema respiratorio.

1.4 OBJETIVOS

1.4.1 OBJETIVOS GENERALES

Determinar el patrón espacial de la morbilidad de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa en el Hospital Santa Rosa, 2006.

Determinar el patrón espacial de la morbilidad del total de casos de las enfermedades del sistema respiratorio en los distritos de Lima Metropolitana, 2015.

1.4.2 OBJETIVOS ESPECÍFICOS

Determinar las zonas con mayor demanda de pacientes con infecciones respiratorias (con Rinofaringitis aguda, Asma no especificado y Bronquitis aguda no especificada) del Hospital Santa Rosa a partir de los indicadores Locales de Autocorrelación Espacial (LISA) para el total de casos de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa.

Comparar los resultados del Hospital Santa Rosa con el total de casos registrados en los distritos de Lima Metropolitana.

1.5 HIPÓTESIS

1.5.1 HIPÓTESIS GENERALES

La morbilidad de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa – 2006 están distribuidas de forma no homogénea a lo largo de Lima Metropolitana generando grupos de distritos vecinos con mayor demanda.

La morbilidad del total de casos de las enfermedades del sistema respiratorio está distribuida de forma no homogénea a lo largo de Lima Metropolitana generando grupos de distritos vecinos con mayor demanda en el 2015.

1.5.2 HIPÓTESIS ESPECÍFICAS

El cálculo del índice de LISA muestra la presencia de clústeres significativos de distritos con alta demanda de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa.

La distribución espacial de los casos de la morbilidad de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa presenta un patrón espacial distinto cuando se analizan el total de casos de IRAS registrados.

CAPÍTULO II

MARCO TEÓRICO

2.1 ANTECEDENTES DEL PROBLEMA

Buzai (2007) en la investigación *“Análisis exploratorio de la estructura espacial de enfermedades en la ciudad de Luján”* cuyo objetivo fue presentar de manera exploratoria los datos de morbilidad asistida a través de los diagnósticos realizados en las salas de atención primaria y las condiciones sociales y habitacionales.

En la investigación Buzai mostró una aplicación que corresponden a la ciudad de Luján en la cual aplica la relación entre diagnósticos de enfermedad, características poblacionales y ambiente urbano para realizar el análisis de la distribución espacial de los datos de salud a través de las enfermedades diagnosticadas en las áreas administrativas o censales y luego generar los respectivos mapas.

Fuenzalida y otros (2013) presentaron la investigación *“Análisis exploratorio de datos espaciales aplicado a MP10 y admisión hospitalaria. Evidencia para áreas urbanas chilenas contaminadas por humo de leña”*, en la cual detalla un protocolo de análisis exploratorio de datos espaciales, para establecer relaciones entre la exposición a la contaminación por material particulado respirable de 10 micras (MP10) y las admisiones hospitalarias por enfermedades respiratorias y cardiovasculares asociadas a este contaminante.

Los autores consideran dos fases, una de preparación de geodatos y otra de Análisis Exploratorio de Datos Espaciales (ESDA), que contempla diagrama

de caja, histograma, tasas, prueba Z, cartografía temática. Los resultados mostraron que en áreas urbanas chilenas presentan condiciones ambientales de contaminación del aire por humo de leña, entre las que se destacan: primero el aumento de las variables estudiadas (registro de niveles de material particulado respirable, admisiones hospitalarias según enfermedades relacionadas a la contaminación por MP10, enfermedad diagnosticada y sistema previsual de salud a la cual pertenece) durante los meses de invierno y primavera; le siguen las enfermedades isquémicas al corazón y la neumonía demandan una mayor cantidad de atención sanitaria y; finalmente los grupos más vulnerables en términos de ingreso, son los que proporcionalmente se enferman más.

Carmona, (2012), presentó la investigación "*Monitoreo de la infección respiratoria aguda en Manizales, mediante sistemas de información geográfica*", utiliza herramientas de recolección, estratificación, ubicación y análisis de datos, para mejorar y prevenir las infecciones respiratorias agudas; se citan aquí ejemplos que relacionan los Sistemas de Información Geográfica (SIG) con estudios aplicados al campo de la salud pública para la ciudad de Manizales. Carmona en la investigación incluye la sectorización de la ciudad de acuerdo a los diferentes niveles de contaminantes atmosféricos detectados por los SIG, el número potencial de personas que se ven afectadas y las características físicas y sociales de las mismas.

En cada una de las cuatro zonas en las que se divide la ciudad es posible recolectar y analizar datos correspondientes a factores contaminantes en la zona; por ejemplo, emisiones en la Zona I, humo generado por automotores en la Zona C, o material particulado suspendido en el aire, por actividades vigentes de construcción en cualquiera de las zonas expuestas. Una vez obtenida la información pertinente a la ubicación y condición de la zona, es posible la recolección de datos para otra capa diferente dentro del SIG. Con la inclusión de una subdivisión correspondiente a la población de cada zona, el SIG cuenta con herramientas para discriminar información importante y datos específicos para niños (de diferentes edades), adultos (en general) y ancianos.

Gómez y otros, (2015), en la investigación "*Análisis espacial de la mortalidad por enfermedades cardiovasculares en la ciudad de Madrid, España*", el objetivo del estudio fue analizar el patrón espacial de la mortalidad por enfermedades cardiovasculares en el área urbana poblada (AUP) del municipio de Madrid, así como identificar agregaciones espaciales. Se realizó un estudio ecológico, por sección censal, para hombres y mujeres durante 2010. Se calculó la Razón de Mortalidad Estandarizada (RME), Riesgo Relativo Suavizado (RRS) y Probabilidad Posterior (PP) de que el RRS fuera mayor que 1. Para identificar clusters espaciales se utilizó el índice de Moran (I Moran) y el Índice Local de Autocorrelación Espacial (LISA). Los resultados fueron representados cartográficamente.

En el patrón espacial de la mortalidad por enfermedades cardiovasculares en Madrid, se observaron diferencias por sexo. Los mapas de RME, RRS y PP mostraron un patrón más heterogéneo en los hombres mientras que en las mujeres se detectó uno más definido, con un riesgo relativamente mayor en zonas periféricas del AUP. El método LISA mostró agrupaciones espaciales similares a los patrones mostrados por el índice de Morán.

Buzai, (2011), en la investigación "*Métodos cuantitativos y Sistemas de Información Geográfica (SIG) aplicados a la Geografía de la salud. Análisis espacial de diagnósticos en la ciudad de Luján (Argentina)*" se centra en el análisis espacial cuantitativo de las distribuciones espaciales de diagnósticos de enfermedades infecto-parasitarias, de la sangre, mentales, neurológicas, del ojo, del oído, circulatorias, respiratorias, digestivas, piel-subcutáneas y osteomusculares con la finalidad de analizar sus distribuciones espaciales.

Se determinó que la distribución espacial de enfermedades de acuerdo al diagnóstico, así como la correlación con una síntesis del mapa social de la ciudad a través de la variable Necesidades Básicas Insatisfechas (NBI). En la investigación se destaca los métodos cuantitativos ligados a la tecnología SIG que presentan gran capacidad para el análisis de la distribución espacial

de enfermedades y con ello servir para la generación de herramientas de apoyo a la planificación socioespacial de la salud.

Robles (2014), presentó la investigación “*Análisis espacial del estudio de los casos de dengue en la Guajira 2013*”, cuyo objetivo fue identificar la distribución del patrón espacial y la localización de áreas de riesgo de los casos de Dengue en La Guajira. La metodología se inició con la determinación de variables relevantes (edad, sexo, antecedentes de dengue, familiar con sintomatología de dengue, ocupación grupo poblacional, área de ocurrencia, desplazamiento o movilidad, clasificación inicial del caso, seguimiento y clasificación final) en los casos, sobre eventos georeferenciados, para luego localizar las áreas de intensidad de riesgo mediante análisis de densidad de casos.

Los municipios de: Riohacha y Dibulla del Departamento de La Guajira presentan un patrón de agrupación de Dengue estadísticamente significativo, focalizado en las áreas pobladas. De igual forma aparecen múltiples núcleos de riesgo a edades tempranas. Con este estudio se aporta a la gestión de control y prevención de esta enfermedad.

Buzai & Baxendale (2009), en la investigación “Análisis Exploratorio de Datos Espaciales” que tiene como objetivo abordar las posibilidades técnico-metodológicas del Análisis Exploratorio de Datos Espaciales (ESDA, Exploratory Spatial Data Analysis) mostrando algunos ejemplos de resultados obtenidos a partir de aplicaciones exploratorias específica.

Chasco (2003), presentó la investigación “*Métodos gráficos del análisis exploratorio de datos espaciales*” tiene por objetivo revisar la literatura existente sobre los principales métodos del análisis exploratorio de datos espaciales reticular o “lattice” univariante, con algunos ejemplos ilustrativos.

Presenta también el concepto de análisis exploratorio de datos espaciales y el desarrollo de los principales métodos gráficos con aplicaciones y uso de paquetes informáticos especializados en el desarrollo de estas técnicas.

La Dirección General de Investigación y Asuntos Ambientales del Servicio Nacional de Meteorología e Hidrología (SENAMHI) (2011) presentaron la investigación “Evaluación de la calidad del aire en Lima Metropolitana 2011”, cuyo objetivo es ejecutar la vigilancia de las condiciones meteorológicas y los contaminantes atmosféricos. En Lima Metropolitana, la vigilancia de la calidad de aire se efectúa en cinco estaciones automáticas ubicadas en los distritos de Ate, San Borja, Jesús María, Santa Anita y Villa María del Triunfo. En estas estaciones se miden los contaminantes criterio, es decir: material particulado menor de 10 micrómetros (PM₁₀), dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂), monóxidos de nitrógeno (NO), óxidos de nitrógeno (NO_x) y ozono superficial (O₃).

Esta investigación registra la concentración media mensual por cada contaminante, identifica dos concentraciones máximas diarias, explica la distribución porcentual de la calidad y finalmente presenta la evolución de la concentración media diaria por tipo de contaminante comparándola con las registradas en las demás estaciones. Entre los principales resultados, se destaca que el mayor problema de Lima Metropolitana sigue siendo las altas concentraciones de material particulado menor a 10 micrómetros, que en el distrito de Ate sobrepasaron 48 veces el Estándar de Calidad Ambiental del Aire (ECA) en 2011. Las estaciones ubicadas en Ate y Santa Anita muestran los mayores índices de contaminación, mientras que San Borja y Jesús María son los distritos más limpios.

Hernández y otros (2000) encontraron asociaciones positivas entre las concentraciones de PM₁₀ y el número de consultas por asma y enfermedades respiratorias aun cuando los niveles alcanzados no excedían las normas ambientales mexicanas. Asimismo, se detectó un efecto sinérgico entre ozono y PM₁₀.

Romero y otros (2004) publicaron el artículo “Contaminación atmosférica, asma bronquial e infecciones respiratorias agudas en menores de edad, de La Habana” sugieren que los niveles de contaminantes atmosféricos en la ciudad de La Habana afectan la salud respiratoria de los niños, por lo que se

requiere de la aplicación de medidas de control, en particular para disminuir las emisiones vehiculares.

Oryazún (2010) indica que los efectos de la contaminación intra y extradomiciliaria sobre la salud humana, poniendo especial énfasis en la situación de los habitantes de ciudades chilenas con niveles críticamente altos de contaminación atmosférica. Los contaminantes atmosféricos riesgosos para la salud humana son el material particulado inhalable (PM10; PM2,5 y PM0,1) y compuestos químicos gaseosos tales como dióxido de nitrógeno, ozono, dióxido de azufre y monóxido de carbono. El aire intradomiciliario contiene una variedad de compuestos nocivos que derivan de múltiples fuentes. Las más importantes son el humo de cigarrillo, artefactos de calefacción y para cocción de alimentos y los agentes biológicos y sus sub-productos. La exposición a contaminantes del aire no solo puede aumentar la tasa de morbilidad sino la tasa de mortalidad como también puede aumentar el número de ingresos hospitalarios de pacientes con síntomas respiratorios y cardiovasculares. La contaminación del aire es importante en la determinación de la calidad de vida de niños menores, ancianos y en pacientes con enfermedades respiratorias y cardiovasculares.

2.2 CONCEPTOS PREVIOS

Previo al análisis exploratorio de datos espaciales (AEDE) se presentan los siguientes conceptos:

2.2.1 Datos Georreferenciados

Según Rocco (2012), un dato espacial es la observación de una variable asociada a una localización del espacio geográfico.

Dado que las mediciones de las características de interés están asociadas a las coordenadas de los sitios donde fueron tomadas; el esquema general de los datos georeferenciados es:

Tabla 1
Esquema de datos georreferenciados

Pto.	Longitud	Latitud	X_1	X_2	...	X_p
1	--	--	X_{11}	X_{12}	...	X_{1p}
2	--	--	X_{21}	X_{22}	...	X_{2p}
.	--	--
.	--	--
.	--	--
n	--	--	X_{n1}	X_{n2}	...	X_{np}

Donde:

n : número de sitios muestreados.

p : número de variables.

X_{ij} : corresponde a la medida de la variable que puede ser cuantitativa o cualitativa

j : 1, 2, ..., p

i : 1, 2, ..., n

2.2.2 Tipos de datos

Existen diferentes tipos de datos espaciales entre los que se mencionan:

- a. Eventos o Patrones puntuales: fenómenos expresados a través de ocurrencias identificadas como puntos localizados en el espacio por ejemplo ocurrencia de dolencia, localización de especies vegetales, etc.
- b. Superficies: son aquellos que se registran de un conjunto de ubicaciones fijas sobre una superficie continua como por ejemplo las características del suelo, la contaminación del aire, profundidad de la nieve o los niveles de precipitación.

- c. Objetos: son aquellos que se refieren a objetos de punto o área que se encuentran en el espacio geográfico, por ejemplo, la ubicación de un individuo en función de su lugar de residencia.

A partir de estos tres tipos de datos Cressie (1991) divide la Estadística Espacial respectivamente en tres ramas:

La primera denominada *análisis espacial de patrones puntuales espaciales*, basado principalmente en la teoría de procesos puntuales espaciales, donde el interés radica principalmente en determinar el tipo de distribución espacial que siguen los eventos registrados en forma de puntos, es decir, determinar si estos están distribuidos de manera aleatoria a lo largo del espacio o se encuentran formando agregaciones espaciales o finalmente presentando separación entre ellos. Así también el interés radica en determinar si los eventos ocurren de forma homogénea a lo largo del espacio o existen determinados lugares con mayores ocurrencias.

La segunda rama referida a la *geoestadística*, donde el interés radica principalmente en predecir el valor de una variable en cuestión, en un lugar donde esta no fue observada.

En último caso está referido principalmente al *análisis de datos laticce* o datos de área. Donde la variable de interés debe ser observada en cada una de las ubicaciones, áreas o territorios. Muchos autores como Anselin (1988), Moreno y Serrano, denominan a esta rama como la Econometría espacial ya que el interés principalmente radica en determinar modelos econométricos que incluyan en su formulación la denominada dependencia espacial o autocorrelación espacial.

2.2.3 Variables

Variable es una característica que hace referencia a los objetos espaciales sujetos a la *conceptualización y representación*, tales como el número de casos de un delito, el número de especies de plantas entre otros.

La conceptualización se refiere a la definición y significado del atributo por ejemplo un delito, una enfermedad, etc. mientras que la representación se refiere a cómo un atributo se operacionaliza en variables con el fin de adquirir y almacenar datos sobre el atributo.

2.2.4 Calidad de los datos

Veregin y Hargitai (1995), indican que se debe cumplir los siguientes aspectos para considerar que los datos son de calidad:

- a. **Precisión:** Según Taylor (1982), precisión de los datos es la inversa de error de datos y se define como la diferencia entre el valor de una variable y el verdadero valor de esa variable.
- b. **Resolución:** se refiere a la precisión con la que los atributos se miden.
- c. **Consistencia:** se define como la ausencia de contradicciones en una base de datos. La definición de consistencia es similar a la utilizada en la estadística clásica.
- d. **Compleción:** Si todas las variables que se han especificado para el estudio son necesarias para realizar los análisis y si la escala espacial y el alcance geográfico son suficientes. Es importante en el contexto del diseño de recolección de datos primarios y para la evaluación de las fuentes de datos secundarios.

2.2.5 Esquema Reticular o Lattice

Llamado así porque es más adecuado para el tratamiento de unidades poligonales (reticulares) continuas, que son las propias de la econometría espacial.

Los datos reticulares o "lattice" pueden definirse como aquéllos referidos a regiones de formas no regulares en el espacio (provincias, distritos, etc.)

2.3 ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES

2.3.1 Historia

El origen del análisis exploratorio de datos espaciales (AEDE) se encuentra en la geografía cuantitativa durante las décadas de 1950 y 1960, convirtiendo la geografía tradicional en una disciplina más científica por el uso de técnicas y métodos geográficos, métodos estadísticos avanzados y el uso de la tecnología.

Tukey en 1977 fue el primero en afianzar y difundir las técnicas del análisis exploratorio de datos como herramientas que permiten realizar un estudio inicial de los mismos en aproximaciones gráficas previas a la realización de procedimientos estadísticos de mayor complejidad.

El trabajo de los estadísticos tales como Ripley (1981), Diggle (1983), Upton y Fingleton (1985 y 1989), Haining (1990), Cressie (1991), Walden y Guttorp, (1992) han contribuido en el desarrollo de la teoría estadística acerca de datos espaciales.

Otros autores que han desarrollado el análisis exploratorio de datos en la década de los noventa son: Agresti y Finlay (1994), Bosque Sendra y Moreno Jiménez (1994) y Hair et al. (1999).

Anselin es uno de los principales desarrolladores del campo de la econometría espacial. En 1990, desarrolló la autocorrelación espacial que es otro aspecto fundamental del análisis exploratorio y la caracterización de la dependencia espacial.

En los últimos años, las investigaciones realizadas en el campo del análisis exploratorio de datos espaciales se han centrado en la conexión de los sistemas de información geográfica (GIS) con paquetes estadísticos tales como SPSS, Matlab, o R project. Muchos de los programas utilizados en el AEDE se pueden encontrar en la Tabla 2.

Tabla 2

Paquetes informáticos utilizados en el AEDE

Paquete informático	Empresa/autor	Página web
cdv, Cartographic Data Visualizer	Jason Dykes	www.geog.le.ac.uk/jad7/cdv/
ClusterSeer	Terraseer	http://www.terraseer.com/products/clusterseer.html
CrimeStat Spatial Statistics Program	Ned Levine & Associates	http://www.icpsr.umich.edu/NACJD/crimestat.html
Flow Mapper	Waldo Tobler & David Jones	http://csiss.ncgia.ucsb.edu/clearinghouse/FlowMapper/
GeoDa	Luc Anselin, Universidad de Illinois (USA)	http://sal.agecon.uiuc.edu/geoda/main.php
SAS/GIS	SAS Institute Inc.	http://www.sas.com/products/gis/
STARS, Space Time Analysis of Regional Systems	Sergio Rey	http://stars-py.sourceforge.net/
STIS, Space Time Intelligence System	Terraseer	http://www.terraseer.com/products/stis/stis_features.html
WinGslib, Geostatistical Software Library	Statios, LLC	http://www.statios.com/WinGslib/

Fuente: Coro Chasco Yrigoyen

2.3.2 Análisis exploratorio de datos espaciales

El AEDE forma parte del más amplio campo del análisis exploratorio de datos. Good (1983) define el AEDE como conjunto de técnicas estadístico gráficas diseñadas para detectar patrones y tendencias en los datos, para identificar características inusuales con objeto de formular hipótesis sobre los mismos.

La estadística espacial se ocupa de la exploración, la descripción, la visualización y el análisis de los datos considerando sus características de distribución en el espacio, que suelen expresarse a través del uso de coordenadas geográficas. Al trabajar con unidades espaciales aparecen dos grandes tipos de efectos espaciales que condicionan la metodología estadística a aplicar: la heterogeneidad espacial y la dependencia espacial.

Cressie (1984) y Haining (1990) definen el AEDE como un conjunto de técnicas para explorar datos espaciales que resume las propiedades espaciales de los datos, la detección de patrones espaciales en los datos, la formulación de hipótesis que se refieren a la geografía de los datos, identificar los casos o subconjuntos de casos que son inusuales dada su ubicación en el mapa.

El análisis exploratorio de datos espaciales según Anselin (1996) es un conjunto de técnicas para describir y visualizar la distribución espacial, descubrir patrones de asociación espacial, identificar las observaciones atípicas espaciales (spatial outliers), descubrir esquemas de asociación espacial, agrupamientos (clusters) o puntos calientes (hot spots) y sugerir otras formas de heterogeneidad espacial.

Un buen método gráfico de AEDE según Chasco (2003) es aquél capaz de analizar y representar dos características fundamentales en toda distribución espacial: tendencia y puntos atípicos.

La tendencia está relacionada con la propiedad de alisado en el campo temporal, la cual está referida a la tendencia central de la variable (mediana) y su dispersión (recorrido intercuartílico), en las variables geográficas incluiría elementos globales, referidos a todo el mapa, como la tendencia espacial y autocorrelación espacial global. Mientras que los puntos atípicos tienen la propiedad de asperezas (“rough”) de la distribución es local y se refiere a aquellos datos localizados a cierta distancia del elemento de alisado, como los llamados atípicos espaciales (“outliers”).

Sin embargo, es preciso mencionar que los gráficos empleados en el AEDE en cada uno de las ramas de la estadística espacial difieren notablemente. Por ejemplo, en el análisis de patrones puntuales espaciales como parte del AEDE, se inicia con el mapeo de puntos, seguido de las funciones basadas entre distancias entre dos eventos o de un punto cualquiera del espacio a un evento, entre otras, los cuales no son adecuados para las otras dos ramas.

En la Tabla 3 se muestran los métodos gráficos del AEDE según: la tendencia espacial y atípicos espaciales.

El AEDE puede ser aplicado para conocer el comportamiento estructural de una variable (análisis univariado), de la relación entre dos variables (análisis bivariado) o desde una perspectiva multivariante.

Cuando no existe un marco formal o teoría previa acerca del fenómeno que se analiza, deben de utilizarse las técnicas del AEDE. Esta situación se plantea cuando se analizan grandes bases de datos geográficos cuya distribución no se conoce a priori.

El AEDE combina el análisis estadístico con el gráfico, dando lugar a lo que podría denominarse una “visualización científica” (Haining *et al.*, 2000) que, a los contrastes estadísticos sobre los efectos espaciales de dependencia y heterogeneidad, une un amplio marco de gráficos o “vistas” múltiples y dinámicas sobre la información geográfica (Unwin, 2000).

2.3.3 Métodos gráficos del AEDE

Tabla 3
Métodos gráficos del AEDE según la tendencia espacial y atípicos espaciales

Tendencia espacial	Atípicos espaciales
Mapas temáticos	Mapa de caja
Mapa dinámico	Mapa de percentiles
Gráficos condicionales	Cartograma
Diagrama de dispersión de Morán	Gráfico de LISA
Diagrama de dispersión de Morán multivariante	Gráficos de LISA multivariantes

Fuente: Coro Chasco Yrigoyen

a) Métodos de representación de la tendencia espacial global

Como la tendencia espacial es de carácter geográfico y tiene un sentido global, es decir, referido a todo el mapa (y no a una parte del mismo). Por eso, las técnicas del AEDE que analizan esta componente son herramientas de representación cartográfica para las que lo fundamental no es tanto el mapa en sí mismo, cuanto la representación gráfica de los estadísticos básicos.

i. Mapas temáticos

Los mapas temáticos son una representación cartográfica de una variable geográfica. Esta representación en un mapa de la variable puede llevarse a cabo mediante símbolos y colores que ubiquen el valor de una variable en cada una de las unidades geográficas consideradas (países, regiones, etc.).

Puede utilizarse un color o símbolo diferente para cada valor o para cada intervalo de valores de la variable. Dentro del AEDE reticular, los mapas temáticos más importantes para la representación de la tendencia espacial de una variable son el mapa de cuantiles y el mapa de la desviación estándar.

i.1 Mapa de cuantiles

En el mapa de cuantiles, los datos se dividen y agrupan en una serie de categorías (cuantiles) con igual número de observaciones.

Los cuantiles dividen a un conjunto de de datos en un determinado número de categorías que en la medida de lo posible cada categoría tenga el mismo número de observaciones. Si el número de categorías es 4, 5 ó 6 toma el nombre de cuartiles, quintiles o sextiles respectivamente.

Por ejemplo, la Figura 1, el número de categorías se ha establecido en 5, para construir un mapa de quintiles de la variable tasa de personas

con título universitario y de bachiller para el conjunto de barrios del municipio de Madrid. Como puede observarse, en la leyenda se indica, entre paréntesis, el número de observaciones de cada categoría, 25 ó 26.

Este tipo de mapa no es útil en casos en los que la variable en cuestión contenga un gran número de observaciones con valores parecidos (por ejemplo, en el caso de variables de sucesos raros, cuando muchas observaciones tienen valor cero).

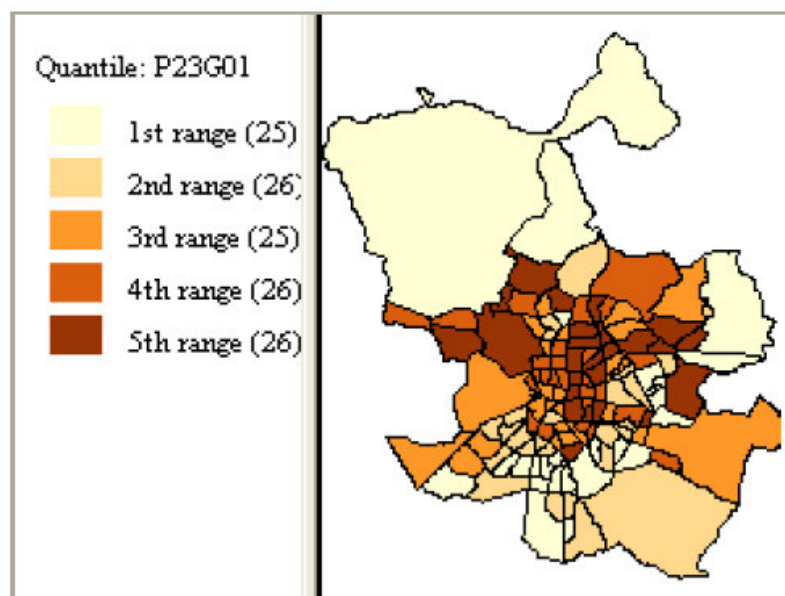


Figura 1. Mapa de quintiles. Coro Chasco Yrigoyen (2003)

i.2 Mapa de la desviación estándar

Este mapa agrupa las observaciones de acuerdo a su pertenencia a un rango estandarizado. Se entiende por rango estandarizado a un número determinado de unidades de la desviación típica a partir de la media.

En este mapa temático, las categorías en las que se divide la variable corresponden con múltiplos de las unidades de la desviación estándar (Coro Chasco Yrigoyen).

En la Figura 2 se ha representado el mapa de la desviación típica de la tasa de paro de los barrios de Madrid.

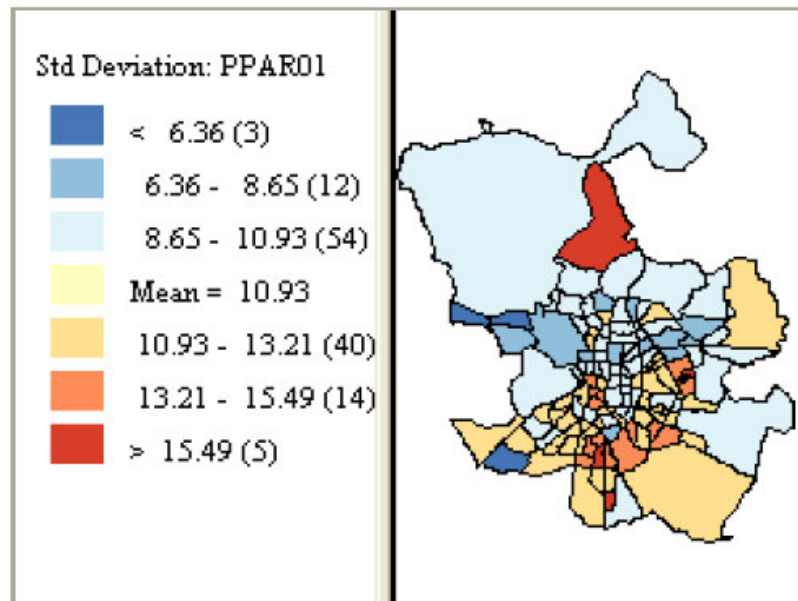


Figura 2. Mapa de la desviación típica. Coro Chasco Yrigoyen (2003)

ii. Mapa Dinámico

El mapa dinámico (“map movie”) es una forma de animación del mapa que consiste en la iluminación, sobre el mismo, de las distintas unidades geográficas según su orden respecto del valor que adoptan en una variable (siempre de menor a mayor).

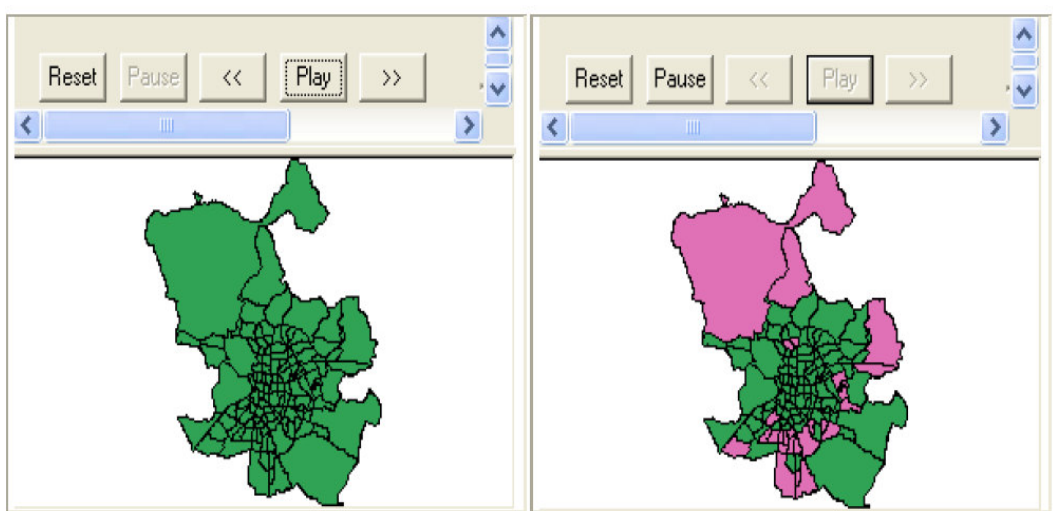


Figura 3. Mapa dinámico acumulativo: inicio (izda.) y pausa (dcha.). Coro Chasco Yrigoyen (2003)

iii. Gráficos condicionales

Este tipo de gráficos utilizan 2 variables condicionales que dividen al conjunto de datos en diferentes grupos o categorías. Se trata de dibujar, para una tercera variable, un gráfico o mapa diferente para cada grupo o categoría.

El programa GeoDa considera, para cada una de las 2 variables condicionales, 3 grupos o categorías, por lo que se producirá un total de 9 gráficos o mapas. Este programa puede diseñar 4 tipos de gráficos condicionales tales como: mapas, diagramas de cajas, histogramas condicionales y diagramas de dispersión condicionales.

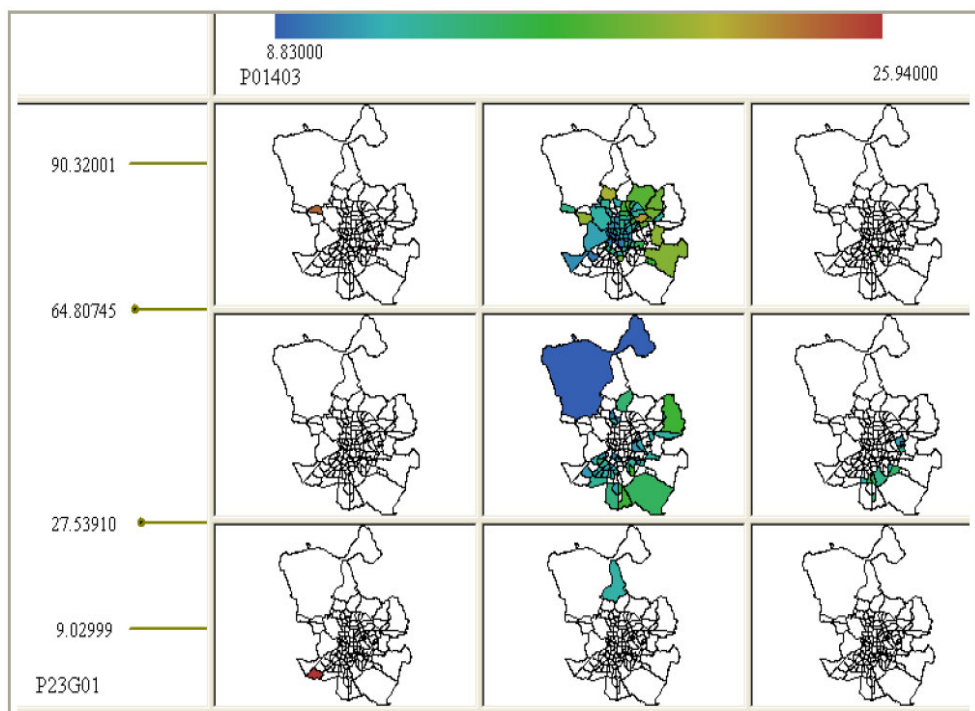


Figura 4. Mapa condicional del índice de infancia condicionado a la tasa de paro y tasa de instrucción superior. Coro Chasco Yrigoyen (2003)

iv. Diagrama de dispersión de Moran

El diagrama de dispersión de Morán es similar a un diagrama de dispersión clásico donde el eje X representa la variable previamente estandarizada y el eje Y representa el retardo espacial de dicha variable estandarizada.

Se entiende por retardo espacial al promedio ponderado de los valores que adopta una variable en el subconjunto de observaciones vecinas a una dada.

Los valores de la variable se encuentran estandarizados en el diagrama de dispersión de Moran, es posible conseguir información tanto de la asociación espacial global (pendiente de la recta de regresión) como de la local. De este modo, los valores en el diagrama de dispersión de Moran situados a más de dos unidades del origen, según el llamado criterio “2-sigma”, pueden considerarse como puntos atípicos en el diagrama de dispersión de Moran.

Este diagrama de dispersión suele dividir el tipo de asociación espacial en cuatro categorías: dos para autocorrelación espacial positiva (valores altos de una variable rodeados de valores altos o valores bajos rodeados de valores bajos) y dos para autocorrelación espacial negativa (valores altos rodeados por valores bajos, y viceversa). Las categorías de asociación espacial positiva se corresponden con los cuadrantes I y III. Por el contrario, las categorías de asociación negativa vienen dadas por los cuadrantes II y IV de este diagrama.

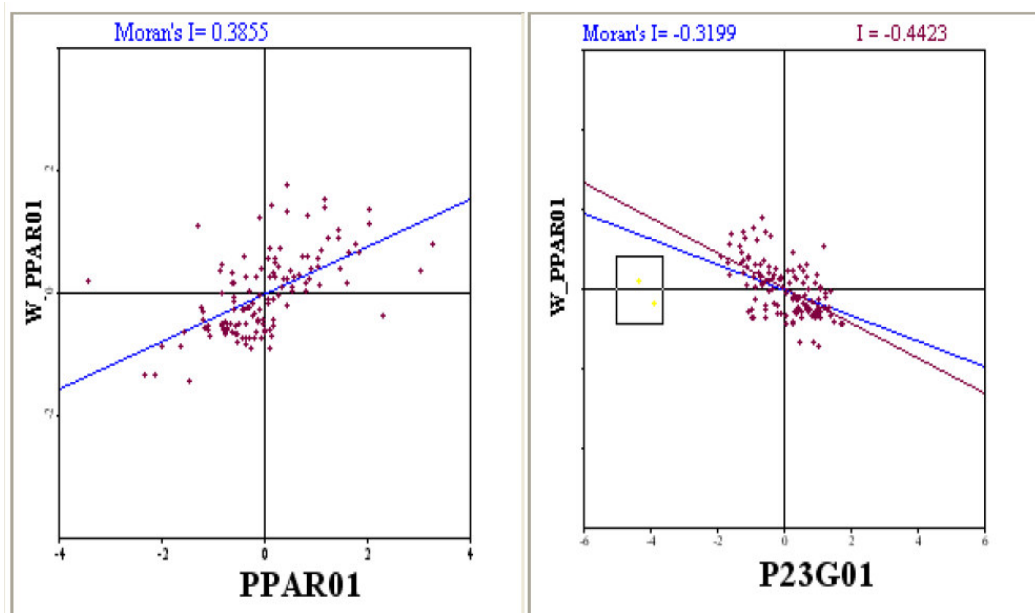


Figura 5. Mapa de dispersión de Morán. Coro Chasco Yrigoyen (2003)

v. Diagrama de dispersión de Moran Multivariante

Este diagrama multivariante es implantado en un entorno dinámico, permite comparar el comportamiento del fenómeno de asociación espacial en varios indicadores, así como analizar el comportamiento de determinadas unidades geográficas que, por ejemplo, se presentan con valores atípicos o extremos.

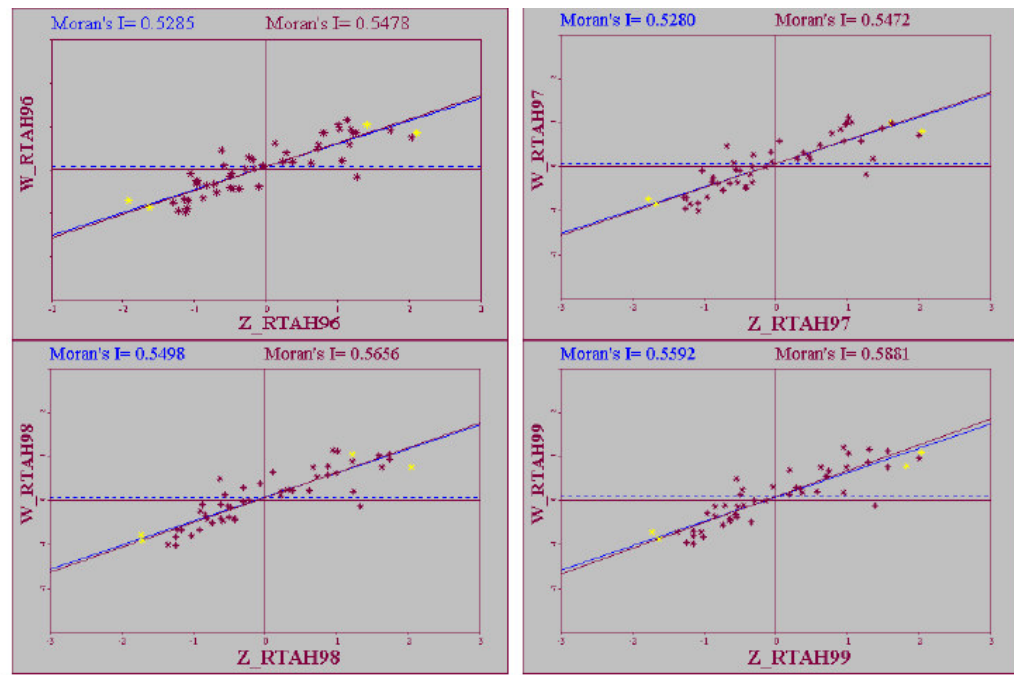


Figura 6. Mapa de dispersión de Morán Multivariante. Coro Chasco Yrigoyen (2003)

b) Métodos de representación de los datos atípicos espaciales

Estos métodos exploratorios analizan la existencia de concentraciones de observaciones cuyo valor se encuentra a cierta distancia de la tendencia general (mediana). Esta propiedad tiene un carácter local debido a que se produce en determinadas zonas del mapa general, normalmente en forma de agrupaciones (o “clusters”).

Estos métodos de representación tienen la propiedad de poder detectar valores atípicamente altos o bajos que se determinan a partir de criterios estadísticos.

i. Mapa de caja

El diagrama de caja, en el análisis exploratorio clásico, es una representación de cinco indicadores correspondientes a las profundidades de Tukey: mínimo, primer cuartil, mediana, tercer cuartil y el máximo. Un gráfico de caja es una herramienta útil para mostrar la forma, la dispersión de los datos y la presencia de datos atípicos si los hubiera.

El mapa de caja del AEDE se realiza a partir del diagrama de caja simple. En este mapa se representan con distintos colores las unidades geográficas cuyos datos en una variable coinciden con la mediana, el rango intercuartílico y los valores atípicos (altos y bajos). El objetivo del mapa de cajas es descubrir valores extremos a partir de un punto de corte superior e inferior.

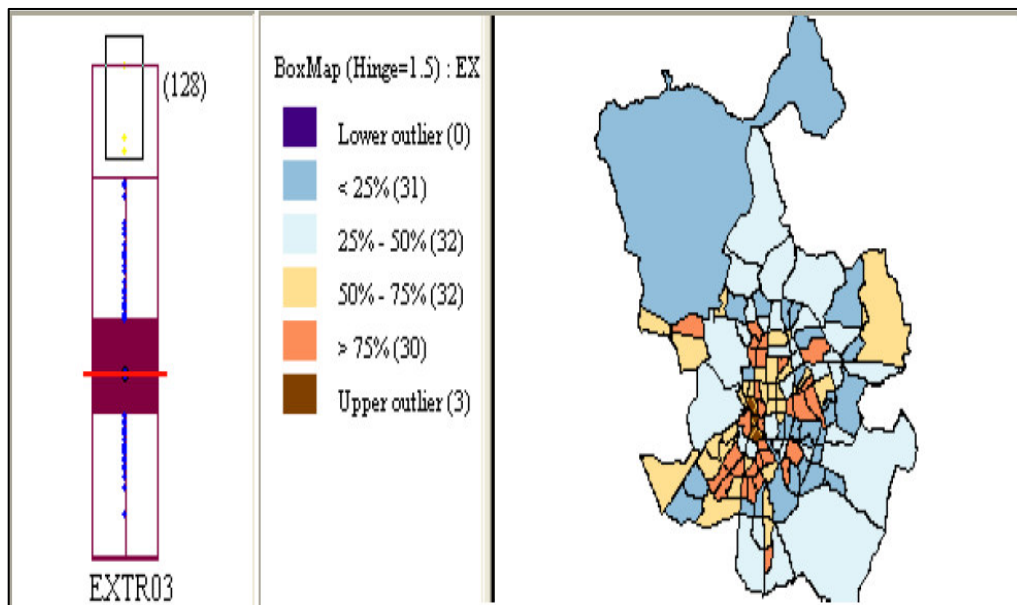


Figura 7. Mapa de cajas. Coro Chasco Yrigoyen (2003)

ii. Mapa de percentiles

El mapa de percentiles es un caso particular del mapa de cuantiles donde los datos se dividen y agrupan en una serie de categorías con igual número de observaciones. Por ejemplo, un mapa de percentiles será aquel cuya distribución se divide en 100 grupos con igual número de unidades territoriales.

Este tipo de mapa no es útil en casos en los que la variable de interés contenga un gran número de observaciones con valores parecidos.

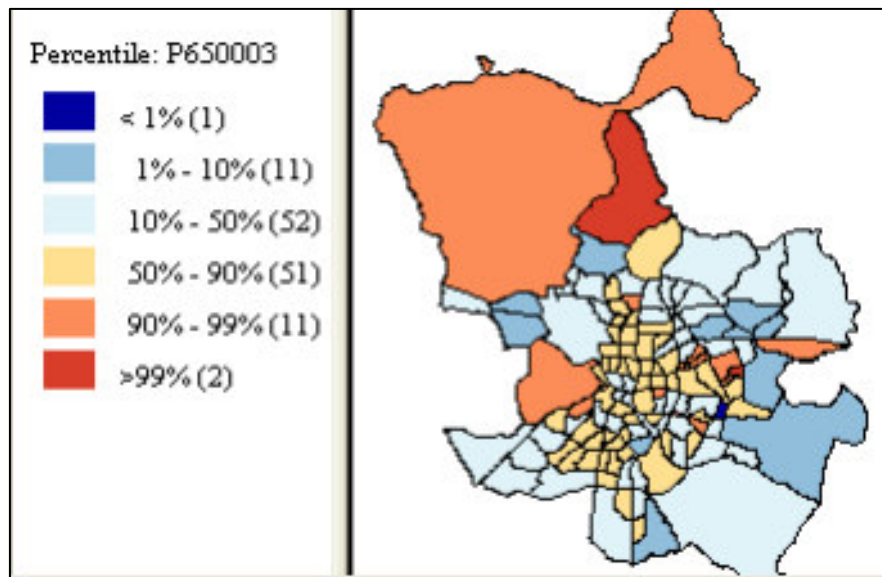


Figura 8. Mapa de percentiles. Coro Chasco Yrigoyen (2003)

iii. Cartograma

El cartograma es un mapa en el que los valores de la variable se representan mediante el tamaño de la superficie de las unidades geográficas (países, regiones, barrios). La representación gráfica de las unidades geográficas se realiza a través de círculos y cuyo tamaño será mayor o menor según sea el valor de la variable.

El cartograma permite la identificación de las unidades geográficas con valores atípicos y se resaltan con colores diferentes del resto de las unidades geográficas.

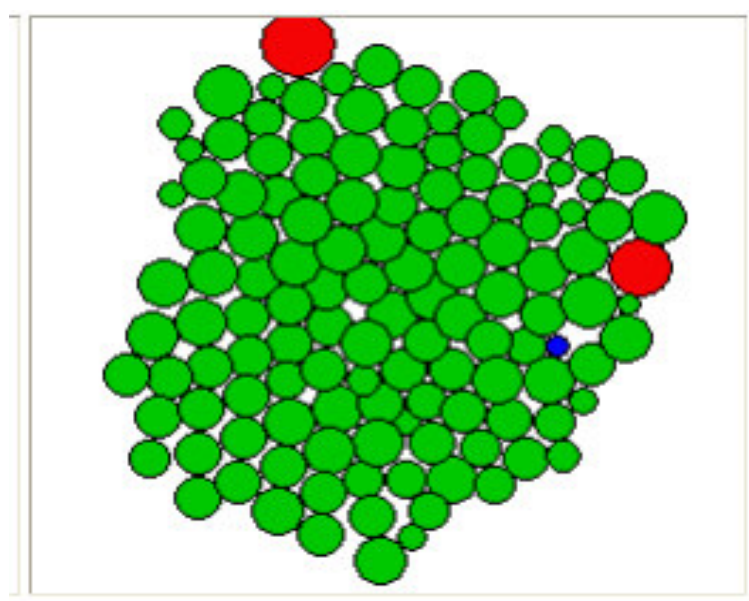


Figura 9. Cartograma. Coro Chasco Yrigoyen (2003)

iv. Mapa de LISA

Las letras LISA significan “Local Indicator of Spatial Association” (Indicador local de asociación espacial). En el mapa LISA se representan las localizaciones con valores significativos en indicadores estadísticos de asociación espacial local permitiendo observar la presencia de puntos calientes (“hot spots”) o atípicos espaciales, cuya mayor o menor intensidad dependerá de las medidas de asociación espacial global.

La presente investigación se realizará utilizando el estadístico de asociación local I de Morán. Este estadístico mide el grado de concentración de valores altos o bajos de una variable en el entorno geográfico de cada una de las observaciones de la muestra para determinar la existencia de heterogeneidad espacial.

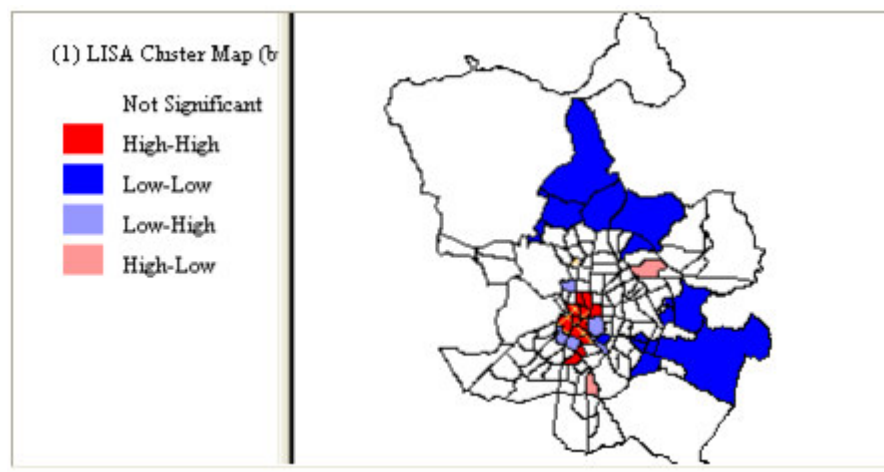


Figura 10. Mapa de LISA. Coro Chasco Yrigoyen (2003)

2.3.4 Medidas de asociación espacial global

Goodchild (1987) define la autocorrelación espacial (AE) como la concentración o dispersión de los valores de una variable en un mapa; es decir refleja el grado en que objetos en una unidad geográfica son similares a otros objetos en unidades geográficas próximas.

El análisis de autocorrelación espacial global realiza un examen conjunto de todas las unidades que componen la muestra para determinar si las

unidades espaciales se encuentran distribuidas aleatoriamente o si, por el contrario, lo hacen conforme a un patrón determinado.

A continuación, se describen dos medidas de asociación espacial:

a. I de Moran: permite averiguar la existencia de *dependencia espacial* de una variable dada en diferentes lugares o, caso contrario, indica una distribución espacialmente aleatoria.

Para determinar la autocorrelación espacial usando el *Índice de Moran* se usa la siguiente fórmula:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n (x_i - \bar{x})} \quad (2.1)$$

donde:

n : número de áreas

x_i : valor de la variable considerada en el área i

\bar{x} : valor medio de la variable en la región de estudio

w_{ij} : pesos de la matriz W

x_j : valor de la variable considerada en el área j

A través de este indicador, podemos averiguar el grado de autocorrelación espacial de una variable x en diferentes áreas i y j , (x_i, x_j) , ponderada por la proximidad geográfica medida por w_{ij} , donde el numerador muestra el promedio de los productos de las desviaciones de las zonas i y j en la media global, y el denominador es una medida de la variabilidad de las desviaciones con respecto a la media global.

El valor del índice I de Moran varía entre -1 y $+1$, donde los valores negativos indican un conglomerado espacial de unidades territoriales con valores de análisis distintos y los valores positivos indican un conglomerado espacial de unidades territoriales con valores de análisis similares, sean estos altos o bajos.

Si el valor de este índice es 0 o cercano a 0 indica que las variables son espacialmente independientes, lo que significa que no hay autocorrelación o no hay dependencia espacial.

b. C de Geary: creado por Roy C. Geary, es una medida de autocorrelación espacial. El índice C Geary está relacionada con la inversa de la índice de I Moran. Esta es una medida global de autocorrelación espacial, mientras que el índice C Geary es más sensible a la autocorrelación espacial local.

El índice de C de Geary se calcula utilizando la siguiente fórmula:

$$C = \frac{(n - 1) \sum_{i,j=1}^n w_{ij} (x_i - x_j)}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n (x_i - \bar{x})} \quad (2.2)$$

donde:

n : número de áreas

x_i : valor de la variable considerada en el área i

\bar{x} : valor medio de la variable en la región de estudio

w_{ij} : pesos de la matriz W

x_j : valor de la variable considerada en el área j

2.3.5 Mapas de riesgos relativos

Los mapas de riesgos relativos suelen trabajar con medidas de riesgo relativo al nivel medio de la región. Estos se utilizan para hacer comparaciones de dos mapas de la misma región sobre la misma variable en periodos de tiempo distintos o en el mismo periodo de tiempo pero con variables distintas.

El riesgo relativo (RR) en el área i se calcula de la siguiente manera:

$$RR_i = \frac{Z_i}{E_i} \quad (2.3)$$

Donde:

z_i : número observado de eventos

E_i : número esperado de eventos

2.3.6 Mapas Probabilísticos y Suavizados con Estimación Bayesiana

Los mapas de probabilidades y el suavizamiento de las tasas a través del método bayesiano empírico se realizan cuando existe alta variabilidad en las tasas.

Para la construcción de los mapas probabilísticos se asume que Z_i sigue un modelo de probabilidad de Poisson con valor esperado E_i . Si X es una variable aleatoria con el mismo modelo de Z_i se define:

$$P_i = \begin{cases} P(X \geq z_i) & \text{Si } z_i \geq E_i \\ P(X \leq z_i) & \text{Si } z_i < E_i \end{cases} \quad (2.4)$$

Un valor de P_i cercano a cero indica que el valor de la tasa en el área i es muy bajo o muy alto respecto a su correspondiente valor esperado.

El suavizamiento bayesiano empírico asume que la tasa en el área i (R_i) es una variable aleatoria que tiene media μ_i y varianza σ_i^2 . Cuando se asume estacionariedad (igual media) se presenta el caso más simple.

El procedimiento consiste en hacer estimaciones iterativas de $(\hat{\theta}_i)$ de la tasa en el área i usando la siguiente expresión:

$$\hat{\theta}_i = w_i r_i + (1 - w_i) r, \quad i = 1, 2, \dots, n. \quad (2.5)$$

Donde:

$$w_i = \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \frac{r}{n_i}}$$
$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n n_i (r_i - r)^2}{\sum_{i=1}^n n_i} - \frac{r}{n}$$
$$r = \frac{\sum_{i=1}^n z_i}{\sum_{i=1}^n n_i}$$

n : número de áreas

2.3.7 Procesos espacial Poisson no homogéneo

El proceso de Poisson no homogéneo con función de intensidad $\lambda(u)$ es definido por las siguientes propiedades:

- Función intensidad: El número esperado de puntos que caen en una región B es la integral $\mu = \int_B \lambda(u) du$ de la función de intensidad $\lambda(u)$ sobre la región B.
- Independencia: Si el espacio se divide en regiones que no se superponen los patrones aleatorios dentro de estas regiones son independientes entre sí.
- El número aleatorio de puntos que caen en una región dada tiene una distribución de probabilidad de Poisson.

2.3.8 Pruebas de aleatoriedad espacial completa

Si se sospecha que la intensidad (número de casos por unidad de área) no es homogénea a lo largo del área de estudio, una prueba estadística no paramétrica útil para corroborar esta suposición, es el método de *conteo por cuadrantes* o también denominadas *pruebas de Aleatoriedad Espacial Completa (CSR, por sus siglas en ingles)*

El método consiste en dividir el área de estudio W en subregiones denominadas cuadrantes, B_1, \dots, B_m de área a_1, \dots, a_m (por simplicidad suponga que son de la misma área) respectivamente tal que $B_1 \cup \dots \cup B_m = W$. El conteo de casos que caen en cada cuadrado, $n_j = n(X \cap B_j)$ para $j = 1, \dots, m$ son estimadores insesgados de los correspondientes valores esperados $E[n(X \cap B_j)]$, estos podrían ser iguales en promedio si el proceso espacial subyacente es un *proceso Poisson con intensidad homogénea* denominado también como *proceso puntual espacial Poisson homogéneo*. Se esperaría que las regiones de igual área tengan aproximadamente igual número de puntos y cualquier tendencia espacial aparente en los conteos n_j sugiere que la intensidad es no homogénea.

De forma general, dejando de asumir que las áreas son iguales, una prueba estadística para homogeneidad, basada en los conteos de los cuadrantes, asumiendo que el proceso puntual espacial es Poisson, consiste en considerar, las siguientes hipótesis:

H_0 : La intensidad es homogénea (CSR) y como H_1 : La intensidad es no homogénea (Es un proceso Poisson no homogéneo). Si la hipótesis nula es verdadera, entonces los n_j son realizaciones de variables aleatorias Poisson independientes, con valores esperado $\mu_j = \rho a_j$.

La prueba χ^2 podría ser usada en dos formas diferentes: como una prueba de bondad de ajuste para la distribución Poisson suponiendo homogeneidad o como una prueba de homogeneidad suponiendo independencia. Enfocándose en la prueba de homogeneidad, se aplica la prueba χ^2 de uniformidad. Dado que el número de puntos $n = \sum_j n_j$ y el área total de la ventana $a = \sum_j a_j$, entonces la intensidad estimada es $\hat{\rho} = n/a$ y el conteo esperado en el cuadrante B_j es $e_j = \hat{\rho} a_j = n a_j / a$. El estadístico ji – cuadrado es:

$$\chi^2 = \sum_j \frac{(n_j - e_j)^2}{e_j} = \sum_j \frac{(n_j - \hat{\rho} a_j)^2}{\hat{\rho} a_j} \sim \chi^2_{(m-1)} \quad (2.6)$$

Si los cuadrantes tienen igual área, entonces los n_j son independientes con igual valor esperado bajo la hipótesis nula. La prueba estadística se reduce a:

$$\chi^2 = \sum_j \frac{(n_j - n/m)^2}{n/m} \sim \chi^2_{(m-1)} \quad (2.7)$$

Bajo la hipótesis nula, la distribución del test estadístico es aproximadamente una χ^2 con $m - 1$ grados de libertad y los residuales Pearson son $r_j = (n_j - e_j) / \sqrt{e_j}$. La aproximación es tradicionalmente aceptable cuando los conteos esperados e_j son mayores que 5 para todos los cuadrantes. Para más detalle consulte (Baddeley et al., 2016).

2.4 Enfermedades del Sistema Respiratorio

2.4.1 Sistema Respiratorio

El sistema respiratorio está formado por un conjunto de órganos que tiene como principal función llevar el oxígeno atmosférico hacia las células del organismo y eliminar del cuerpo el dióxido de carbono producido por el metabolismo celular¹.

La respiración se realiza por medio de dos procesos: **inhalación** que es el proceso de llevar aire a los pulmones y **exhalación** que es el proceso de empujar el aire fuera de los pulmones. La inhalación ayuda al cuerpo a llevar el oxígeno y la exhalación ayuda al cuerpo a expulsar el dióxido de carbono.

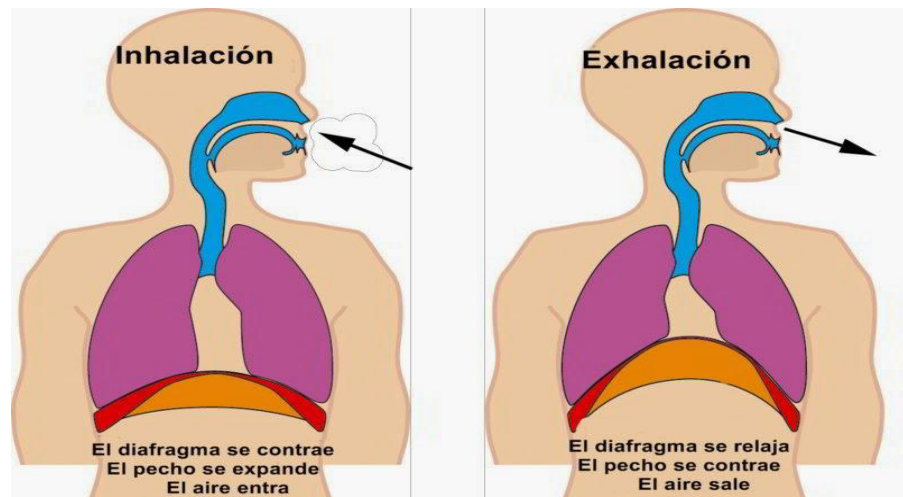


Figura 11. Procesos de la respiración

2.4.2 Enfermedades del Sistema Respiratorio

Las enfermedades respiratorias incluyen todos aquellos padecimientos que afectan el sistema respiratorio desde molestias en la garganta hasta enfermedades pulmonares graves.

¹ El metabolismo celular es el conjunto de reacciones químicas a través de las cuales el organismo intercambia materia y energía con el medio. Entre estos intercambios (por medio de reacciones químicas) está la absorción del oxígeno y su reacción que produce dióxido de carbono que expulsamos al exterior. Pero todo esto se produce gracias a todos los órganos del sistema respiratorio.

Estas enfermedades pueden manifestarse a través de síntomas como obstrucción nasal, dolor de garganta, dolor de cabeza, dolor de oídos, fiebre y malestar general.

Son causadas a través de infecciones bacterianas o víricas, causas ambientales, polución, alta exposición a tóxicos aéreos e incluso tabaquismo como en el caso del cáncer de pulmón.

2.4.2.1 Rinofaringitis aguda: La rinofaringitis aguda es una infección bacteriana de las vías respiratorias, en la faringe y en las cavidades nasales. La rinofaringitis es muy frecuente en los niños pequeños, especialmente en bebés, a partir de los seis meses. De hecho, su sistema inmunitario todavía no es maduro y se defiende con menos eficacia.

En caso de contaminación, la mucosa de la nariz se hincha y produce grandes cantidades de secreciones líquidas, lo que puede taponar la nariz. Se contabiliza que hay más de 200 virus que pueden ser la causa de la rinofaringitis.

2.4.2.2 Asma:

El asma es una enfermedad crónica de los pulmones que inflama y estrecha las vías respiratorias. El asma causa períodos repetidos de sibilancias (silbidos al respirar), presión en el pecho, dificultad para respirar y tos.

2.4.2.3 Bronquitis aguda:

La bronquitis aguda es la inflamación del árbol bronquial. El árbol bronquial se compone de tubos que llevan el aire a los pulmones. Cuando estos tubos se infectan, se hinchan y se forma moco en su interior. Esto hace que usted tenga dificultad para respirar.

2.5 MARCOS CONCEPTUALES

Atopía: es una predisposición genética a padecer alergias.

Material Particulado - PM 10: Son partículas que se encuentran dispersas en el aire y cuyo diámetro aerodinámico es menor a 10 micrómetros (un micrón es la milésima parte de un milímetro).

Estas partículas se encuentran flotando en el aire y pueden ser sólidas o líquidas, orgánicas e inorgánicas siendo su origen mayoritariamente natural las mismas que debido a su pequeño tamaño tienen una velocidad de sedimentación muy baja pudiendo mantenerse en el aire por mucho tiempo y logrando alcanzar distancias de hasta 30 millones con referencia a su lugar de emisión. Las mismas que al ser inhaladas afectan la salud de las personas penetrando en las vías respiratorias pudiendo en su camino adherirse a sus paredes y ocasionar síntomas respiratorios severos (tos, dificultad para respirar, agrava el asma) y llegar hasta los pulmones produciendo enfisemas o cáncer pulmonar.

Material Particulado – PM 2.5: Es el material particulado respirable disperso en la atmosfera con diámetro igual o menor a 2.5 micrómetros (PM_{2.5}).

Estas partículas, ingresan con mayor facilidad a los pulmones, son partículas muchísimo más pequeñas que la pelusa que vemos a la luz del sol y pueden llegar al torrente sanguíneo. Las fuentes de partículas finas incluyen la combustión en vehículos, generadores, quema de madera, procesos industriales; las partículas entre 2.5 y 10 micrómetros de diámetro tienen su origen en polvos que se levanta en las vías de tránsito no pavimentados; las partículas finas pueden acumularse en el sistema respiratorio; la exposición a partículas gruesas puede agravar el asma y bronquitis crónica, mientras que el material fino se asocia con efectos graves.

Enfisemas: El enfisema es un tipo de enfermedad pulmonar obstructiva crónica (EPOC) en donde los alveolos, las bolsitas de aire de los pulmones, se dañan. Como consecuencia, su cuerpo no recibe el oxígeno que necesita. El enfisema hace que sea difícil recuperar el aliento. También le puede causar tos crónica y dificultades para respirar durante el ejercicio.

Dióxido de nitrógeno: Es un gas tóxico, formado por la combinación de oxígeno y nitrógeno.

El bióxido nitroso, el trióxido de nitrógeno (N_2O_3), el pentóxido de nitrógeno (N_2O_5) suelen considerarse en conjunto con la denominación de NO_2 ; que es un gas incoloro e inodoro, tóxico a altas concentraciones y presente en el aire en menos de 0,50 ppm. Los óxidos de nitrógeno son producidos por fenómenos naturales: relámpagos, erupciones volcánicas y la acción bacteriana en el suelo, y además por fuentes antropogénicas como los combustibles de motores de combustión interna.

Calidad del Aire y Contaminación

El aire tiene la función de transmitir el sonido, filtrar y amortiguar los rayos del Sol, dispersar la luz y mantener el equilibrio térmico. La energía solar ayuda en la purificación del aire atmosférico, las plantas a través del proceso de la fotosíntesis absorben el bióxido de carbono y desprenden el oxígeno purificando el aire.

La calidad del aire se mide a través de un índice que es una herramienta usada para proveer al público información oportuna y fácil de comprender sobre la calidad del aire; indica si los niveles de contaminación son perjudiciales a la salud, informa al público si la condición del aire debe preocuparlos por su salud.

La actividad industrial, centrales eléctricas, la actividad de hidrocarburos emite compuestos de azufre, óxidos de nitrógeno, fenoles, la siderúrgica, la

disposición final de residuos sólidos emite óxidos de azufre y nitrógeno, anhídridos y ácidos orgánicos.

El parque automotor formado por vehículos obsoletos de transporte personal y público, emiten gases contaminantes que afectan la salud del habitante en Lima Metropolitana.

La contaminación del aire se está reduciendo en los países industrializados, mientras que se agrava en los países en desarrollo. La contaminación del aire se relaciona con las enfermedades del pulmón, cardiopatía, cáncer del pulmón y daño neurológico en los niños. Desde finales de la década de 1940, el creciente consumo de carbón y petróleo han causado concentraciones mayores de dióxido de carbono en la atmósfera.

La contaminación atmosférica es la presencia en el ambiente de cualquier sustancia química, partículas, microorganismos que alteran la calidad ambiental y la posibilidad de vida. Las causas de la contaminación pueden ser naturales o artificiales, generadas principalmente por la quema de combustibles fósiles (plantas de energía que funcionan a carbón, fábricas y vehículos), partículas y gases industriales.

CAPÍTULO III

METODOLOGÍA

3.1 TIPO Y DISEÑO DE INVESTIGACIÓN

El tipo de estudio es exploratorio porque se va investigar por primera vez el patrón espacial de las tres enfermedades más prevalentes del sistema respiratorio en los pacientes que acuden al Hospital Santa Rosa por consulta externa.

El diseño de investigación es observacional – transversal debido a que observan tal cual se presenta las enfermedades del sistema respiratorio para luego analizar el patrón espacial de estas enfermedades y transversal porque se mide las enfermedades del sistema respiratorio en un momento dado del tiempo.

3.2 UNIDAD DE ANÁLISIS

Cada uno de los distritos de Lima Metropolitana según diagnóstico de las tres enfermedades del sistema respiratorio más prevalentes en consulta externa según CIE - 10 del Hospital Santa Rosa; durante el 2006.

Cada uno de los distritos de Lima Metropolitana según el total de casos de enfermedades del sistema respiratorio en Lima Metropolitana; durante el 2015.

3.3 POBLACIÓN DE ESTUDIO

Todos los distritos de Lima Metropolitana según diagnóstico de las tres enfermedades del sistema respiratorio más prevalentes en consulta externa según CIE - 10 del Hospital Santa Rosa durante el 2006 y del Ministerio de Salud del total de casos de IRAS 2015.

No se trabajará con una muestra debido a que se cuenta con todos los datos de la población de estudio.

3.4 TÉCNICAS DE RECOLECCIÓN DE LOS DATOS

La recolección de los datos se realizó en el Hospital Santa Rosa; durante el 2006 a todos pacientes diagnosticados con las tres enfermedades más prevalentes del sistema respiratorio en consulta externa y del y del Ministerio de Salud el 2015 para el total de casos de IRAS en Lima Metropolitana.

3.5 IDENTIFICACIÓN DE VARIABLES

- **RA: Número de casos de rinofaringitis aguda:** variable cuantitativa discreta.
- **AS: Número de casos de asma no especificada:** variable cuantitativa discreta.
- **BA: Número de casos de bronquitis aguda no especificada:** variable cuantitativa discreta.
- **Total de casos IRAS.** Suma de los tres casos anteriores.
- **Población distrital al 2006 y 2007.** Se consideran la población proyectada por el Instituto Nacional de Estadística e Informática (INEI).
- **Distancia en kilómetros hasta el Hospital Santa Rosa.** Se considera la distancia en línea recta desde el centro de los distritos.

3.7 ANÁLISIS E INTERPRETACIÓN DE LA INFORMACIÓN

Se realizará el análisis exploratorio de datos espaciales que contempla diagrama de caja, histograma, indicadores de agrupamiento (Moran y LISA), pruebas de aleatoriedad espacial completa y cartografía de las enfermedades del sistema respiratorio.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

Para el análisis de los resultados se utilizó los softwares estadísticos SPSS v21 y R Project 3.2.5 para el caso del análisis exploratorio de datos (AED) y para el AEDE se usó QGIS y Geoda 1.8.

Cabe indicar que esta aplicación se realiza con la base de datos de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa en el Hospital Santa Rosa que se encuentra ubicado en el distrito de Pueblo Libre. Estas enfermedades son rinofaringitis aguda, asma no especificada y bronquitis aguda no especificada.

Al 2006, el Hospital Santa Rosa atendió por consulta externa a un total 8'196 pacientes (con algunas de las tres infecciones respiratorias más frecuentes.) de los 43 distritos de Lima Metropolitana. En el lado izquierdo la figura 12, se muestra que la mayor demanda de pacientes era proveniente del distrito de Pueblo Libre en aproximadamente el 27.8% de casos, seguido del distrito de San Miguel con un 11.5%. Asimismo, al lado derecho se observa que la Rinofaringitis Aguda fue la infección más frecuente entre las tres infecciones respiratorias evaluadas, en todos los distritos se observa el mismo comportamiento.

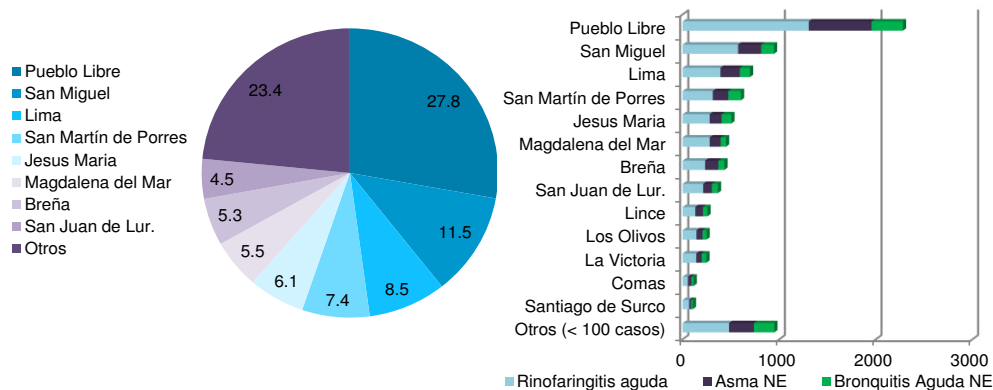


Figura12. Distribución de casos de infecciones respiratorias del Hospital Santa Rosa - 2006

Análisis Exploratorio de datos

Se realizó el análisis exploratorio por separado para cada una de las variables de conteo de las tres enfermedades respiratorias. Se supuso inicialmente que este tipo de variables por su naturaleza deberían ajustarse a una distribución Poisson y por tanto se esperaba que su distribución sea asimétrica.

La figura 13 muestra que la distribución del total de casos en cada una de las enfermedades es semejante y asimétrica positiva propia de variables Poisson con media o tasa de ocurrencia baja. La media distrital para los casos de rinofaringitis aguda (RA), asma (AS) y bronquitis aguda (BA) resultó bajo aproximadamente 107, 52 y 32 casos, respectivamente. En cuanto a la mediana, se obtuvo que el 50% de distritos presentaron menos de 28 casos de RA y el otro 50% de estos fue superior a dicho valor. En el caso de AS y BA las medianas correspondientes fueron de 13 y 12 casos, respectivamente. Así mismo el número máximo de casos registrados a nivel distrital de estas enfermedades (RA, AS, BA) fueron de 1306, 649 y 325 correspondientes al distrito de Pueblo Libre. (Ver Tablas 5, 6 y 7).

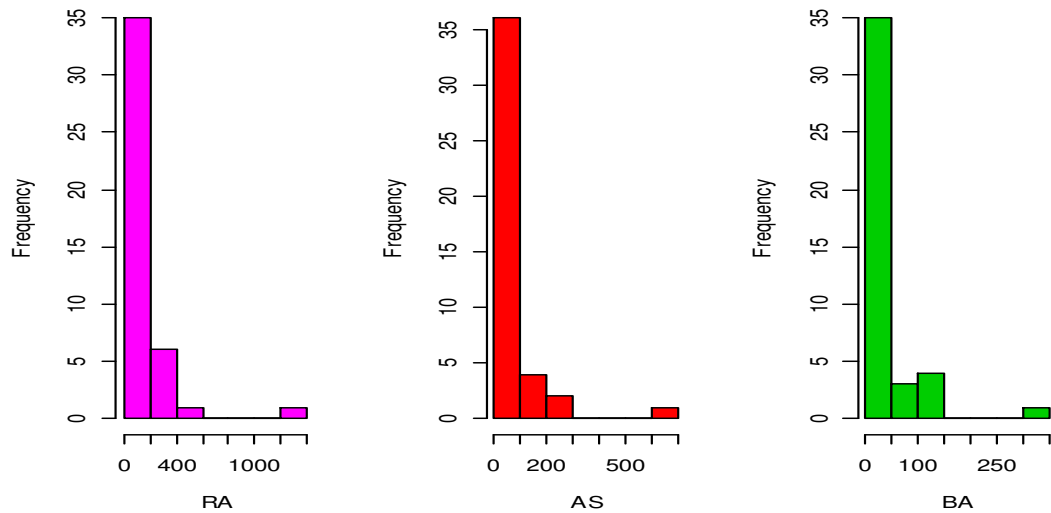


Figura 13. Distribución de los casos de rinofaringitis aguda (RA), Asma (AS) y Bronquitis Aguda (BA) de consulta externa del Hospital Santa Rosa - 2006

El diagrama de cajas de la figura 14 nos indica que existe una alta variabilidad a nivel distrital en los tres tipos de enfermedades, además se observa la presencia entre 5 y 7 distritos outliers con los valores muy altos, los cuales principalmente son: Pueblo Libre, San Miguel, Lima, San Martín de Porres, Magdalena del Mar y Jesús María.

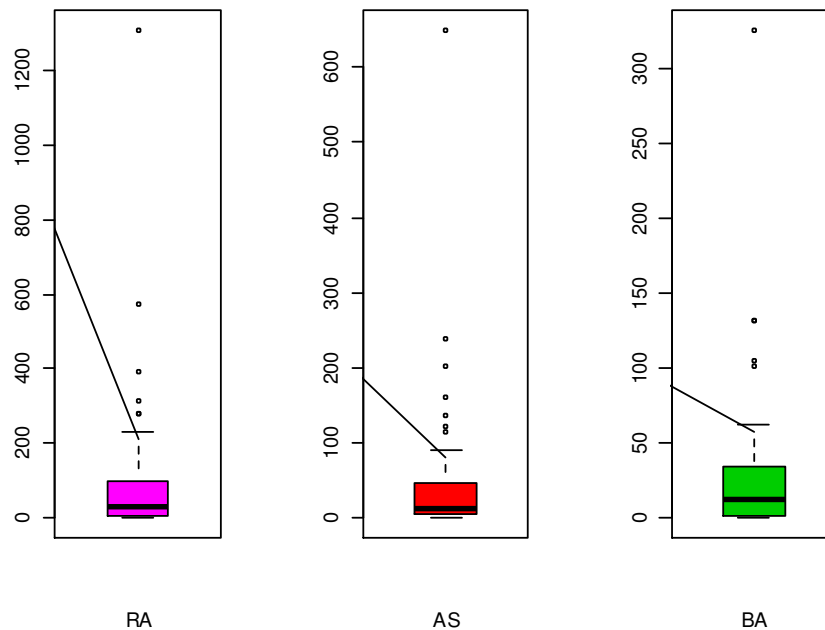


Figura 14. Diagrama de cajas de los casos de RA, AS y BA de consulta externa del Hospital Santa Rosa – 2006

4.1 Análisis Exploratorio de Datos Espaciales

4.2.1 Mapeo de casos

El análisis exploratorio de datos espaciales en muchos casos inicia con la representación de la variable en estudio a través de un mapa coloreado de acuerdo a un criterio de generación de intervalos, con el fin de tener una idea vaga acerca del tipo de distribución espacial de la variable en estudio. En nuestro caso se realizó utilizando el criterio de *cortes naturales*, el cual agrupa a los distritos más similares entre sí, generando intervalos que minimizan la varianza al interior de cada grupo de las variables de conteo.

Los resultados se presentan en la figura 15 y 16, en la cual los distritos con mayor número de casos son coloreados en azul más intenso. En forma de punto y de color rojo se representa a la ubicación del hospital Santa Rosa. Para los tres casos de infecciones respiratorias (Rinofaringitis, Asma y Bronquitis) se observa que existen determinados distritos con un mayor número de casos, en esencia se observa que gran parte de la demanda de pacientes se encuentran en los distritos cercanos a la ubicación del hospital, suponemos que los pacientes prefieren atenderse en un hospital más cercano a estos. Esto conlleva a sospechar que la distribución espacial de los pacientes del hospital Santa Rosa están distribuidos en forma no homogénea a lo largo de la ciudad de Lima y más cercanas a estas.

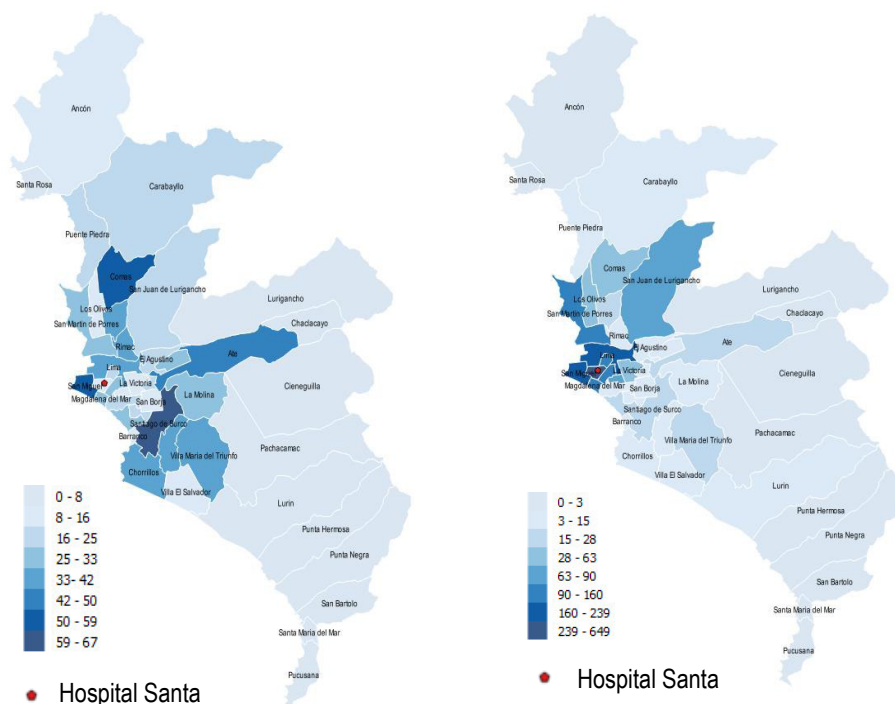


Figura 15. Número de casos de pacientes con Rinofaringitis Aguda (lado izquierdo) y Asma (lado derecho).

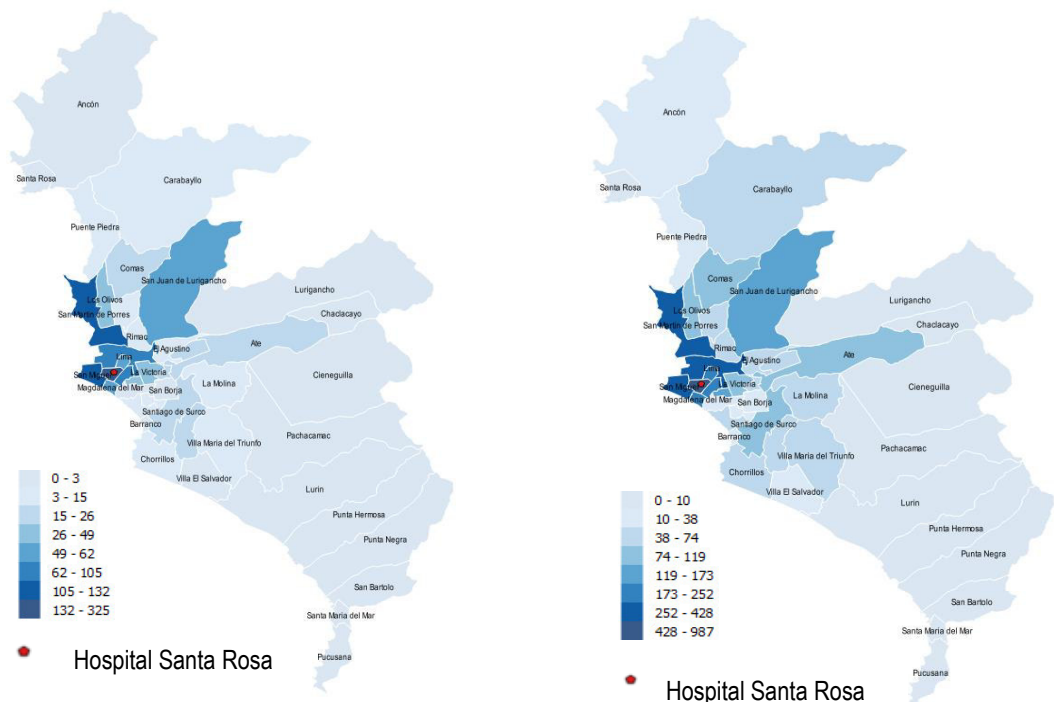


Figura16. Número de casos de pacientes con BA (lado izquierdo) y para el total de las tres infecciones respiratorias (lado derecho).

4.2.2 Mapas para detectar Outliers

Diversos mapas han sido propuestos para identificar los datos outliers y su ubicación en el espacio, sin embargo, no todos resultan útiles para nuestro caso de estudio, así por ejemplo, el mapa de percentiles no es adecuado aquí debido a la cantidad de observaciones o de distritos en análisis. Se requiere por lo menos 100 distritos, para identificar al 1% de estos con valores más altos y bajos. Por lo que en este estudio se usó el mapa de cajas y bigotes (boxmaps). Los resultados se muestran en la figura 17.

Para el caso de RA se encontraron 6 distritos resaltados en color amarillo y rojo, entre ellos: Pueblo Libre, San Miguel, Lima, San Martín de Porres, Jesús María y Magdalena del Mar donde la cantidad de casos con esta enfermedad son muy elevados. En el caso del AS se encontraron 7 distritos, los cuales son los mismos anteriores con la inclusión adicional de Breña. De la misma forma en el caso de BA se encontraron 5 distritos, considerando todos los anteriores a excepción de Breña y Magdalena del Mar.

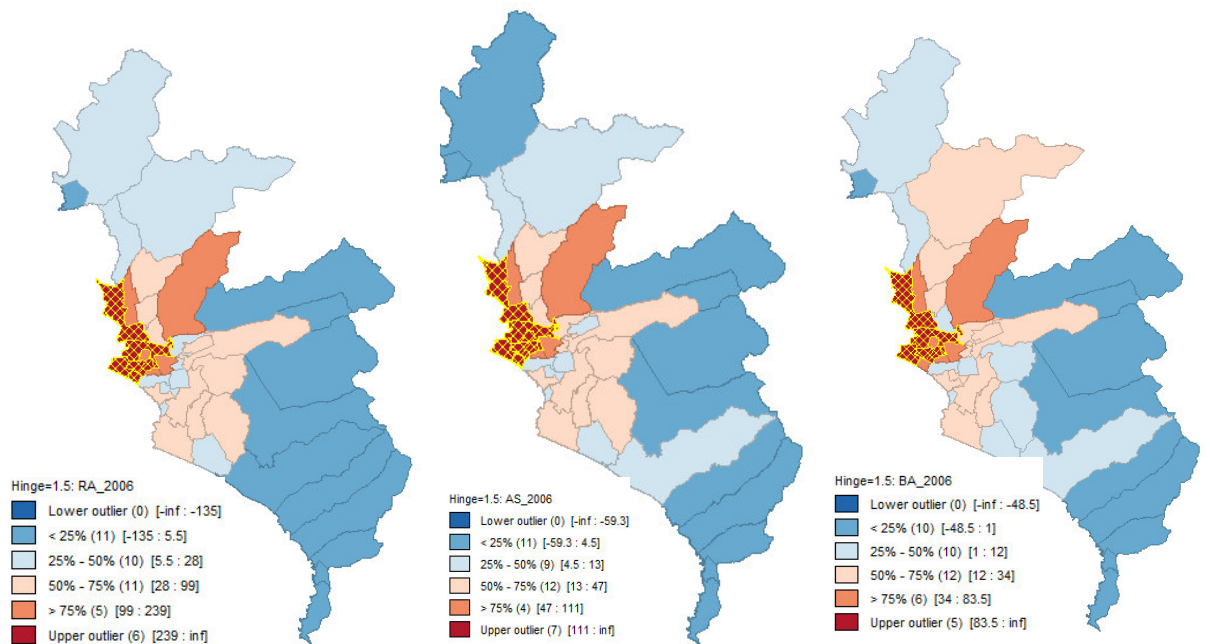


Figura 17. Mapa de Cajas de la distribución de la RA, AS y BA de consulta externa del Hospital Santa Rosa - 2006
Sin embargo, cabe recalcar que estas observaciones no se deben a un error de medición por lo que no son excluidas del análisis.

4.2.3 Mapas de tasas y suavizado bayesiano empírico

En muchos casos analizar variables espaciales de conteo, tal cual se observan, puede llevar a falsas conclusiones principalmente porque la información a nivel distrital no puede ser comparada entre estos, debido a que existen distritos con mayor cantidad de población expuesta a una determinada enfermedad o porque el territorio de algunos de estos es demasiado grande o pequeño. Por lo cual muchos autores sugieren el uso de mapas de tasas o de riesgos relativos, generados a partir de las variables de conteo y la población expuesta.

Para esta tesis, se generó solo una tasa que incluye a las tres enfermedades respiratorias (IRAS). Esta tasa representa al total de casos con enfermedades respiratorias que acudieron al hospital Santa Rosa por consulta externa dividido por la población expuesta al 2006.

El suavizamiento de las tasas a través del **método bayesiano empírico** es una herramienta alternativa cuando se estudian áreas (distritos) con poblaciones en riesgo muy distintas. Sin embargo, este indicador no siempre

resulta útil cuando la cantidad de la población expuesta es bastante alta tal cual ocurre aquí.

El mapeo de esta tasa y su correspondiente suavizado bayesiano empírico, se realizó a partir del método de cuantiles tal cual se observa en el lado izquierdo de la figura 18. Los distritos con mayores tasas de IRAS que acudieron al hospital Santa Rosa fueron en el siguiente orden: Pueblo Libre (3%), Magdalena del Mar (0.8%), Jesús María (0.7%) y San Miguel (0.7%) y otros representados en color rojo oscuro.

Debido a que la población distrital con la que se calcularon las tasas es muy grande, el efecto de suavizamiento bayesiano empírico resultó imperceptible para nuestro caso de estudio, ya que el suavizado resultó similar a los valores de la tasa (Ver lado derecho de la figura 18).

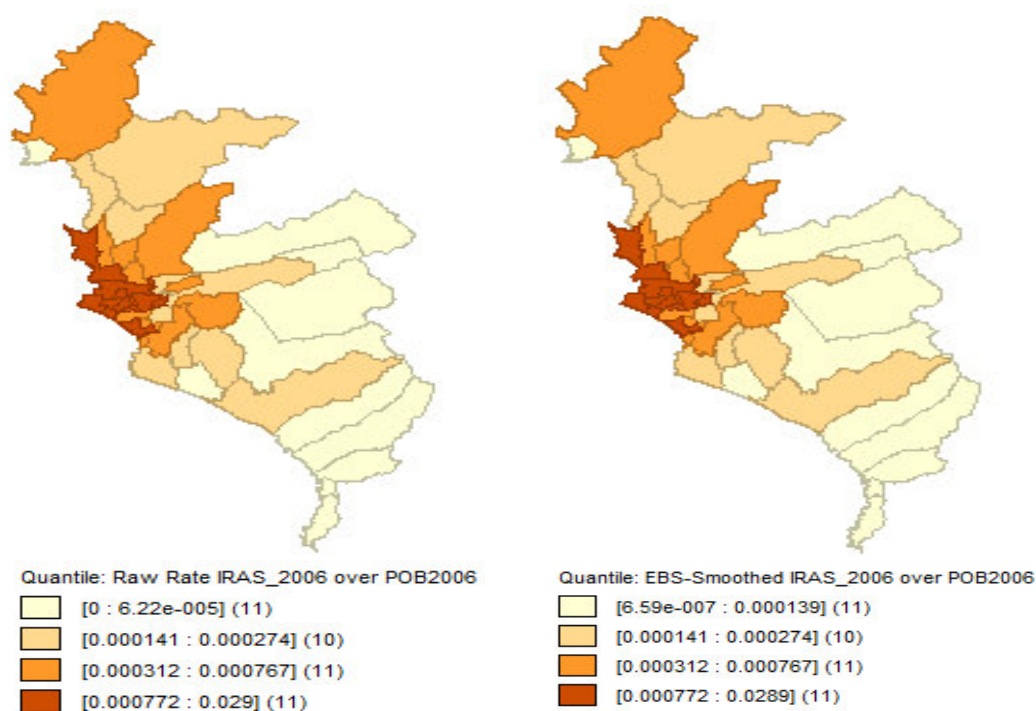


Figura 18. Mapa de tasa de infecciones respiratorias y Suavizado empírico bayesiano de consulta externa del Hospital Santa Rosa - 2006

Adicionalmente se calculó otra tasa correspondiente al total de casos de IRAS reportados por todos los hospitales del MINSA al 2015, en Lima Metropolitana. Para lo cual se consideró como base o población expuesta a la población proyectada por el INEI al 2015.

Los resultados se muestran en el lado izquierdo de la figura 19, para los cuales los distritos de: San Bartolo, Chaclacayo, Punta Negra y Pucusana presentaron las tasas más altas de IRAS por encima de 36% de su población expuesta. De igual forma el efecto de suavizamiento bayesiano empírico resultó imperceptible en este caso.

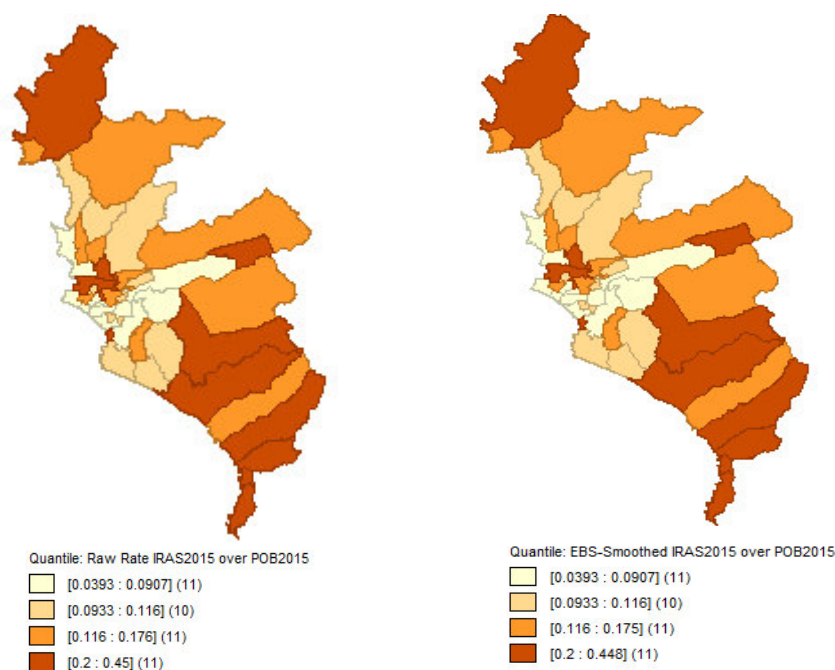


Figura 19. Mapa de tasa de IRAS y Suavizado empírico bayesiano para los distritos de Lima Metropolitana, 2015.

En general, los casos de IRAS atendidos por el hospital Santa Rosa, difieren notablemente si se considera al total de casos registrados en Lima Metropolitana. Se sospecha que en el primer caso se deba principalmente por el hecho que pacientes con enfermedades respiratorias prefieren acudir principalmente al hospital mas cercano a su residencia.

Para corroborar dicha presunción se elaboró un diagrama de dispersión, en la cual la variable dependiente corresponde a la tasa de IRAS atendidas por el hospital Santa Rosa en el 2006 versus la distancia en kilómetros a dicho hospital (DIST_KM). Esta se muestra en la figura 20, en la cual se observa que los distritos que se encuentran más cercanos al hospital Santa Rosa son los mismos que demandan una mayor atención por consulta externa al

hospital Santa Rosa, mientras que, en el caso contrario, son más reducidas la proporción de pacientes que acuden desde distritos lejanos. En general los datos se ajustan de mejor forma a una curva no lineal del tipo exponencial.

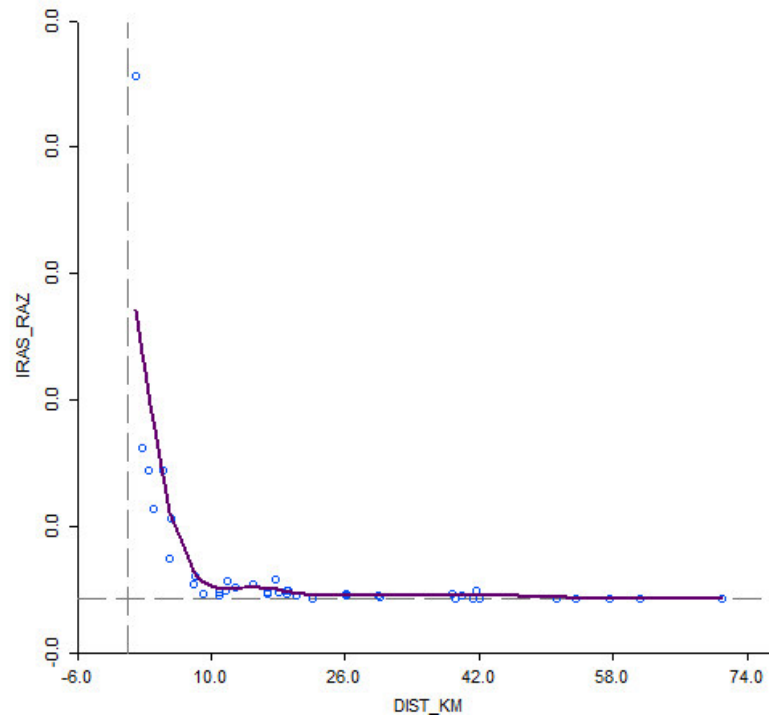


Figura 20. Diagrama de dispersión de la distancia en kilómetros versus la tasa de IRAS atendidas por el hospital Santa Rosa 2006.

4.2.4 Pruebas de aleatoriedad espacial completa

Otra alternativa para estudiar los casos de pacientes con infecciones respiratorias corresponde a analizar el número de casos por unidad de área (km^2), denominada *intensidad* de casos. Esta puede considerarse homogénea u no homogénea a lo largo de la ciudad de Lima.

La prueba de aleatoriedad espacial completa (CSR, por sus siglas en ingles), basada en el estadístico χ^2 , es una alternativa para probar la falta de homogeneidad basada en los conteos de los cuadrantes, asumiendo que el proceso puntual es Poisson, consiste en considerar, la siguiente hipótesis:

H_0 : La intensidad es homogénea (CSR).

H_1 : Es un proceso Poisson con intensidad no homogéneo.

Los resultados se muestran en la Tabla 4, del cual a un nivel de confianza de 99%, se rechaza la hipótesis nula, con lo cual se concluye que los casos de pacientes con infecciones respiratorias que acudieron al hospital Santa Rosa en el 2006, estuvieron distribuidos de forma no homogénea a lo largo de la ciudad de Lima siguiendo un proceso espacial Poisson no homogéneo. A la misma conclusión se llega para el total de casos registrados en Lima Metropolitana en el año 2015.

Tabla 4
Prueba de aleatoriedad espacial completa, para los casos de IRAS
atendidos por el Hospital Santa Rosa 2006 y total de casos de IRAS
en Lima Metropolitana 2015.

Tipo de cuadrantes	Nro. De polígonos	Área total (Km²)	Nro. De casos.	χ^2	Df	p-value
IRAS – Hospital Santa Rosa, 2006	43	2641.4	8196	466878.8	42	0.000
IRAS total Lima Metropolitana, 2015	43	2641.4	1 015218	2382491.3	42	0.000

4.2.5 Autocorrelación espacial

Dado que la distribución espacial de los pacientes con infecciones respiratorias del Hospital Santa Rosa de Lima está distribuido espacialmente de forma no homogénea. Se procedió a Identificar las zonas que presentan altas tasas de ocurrencia (número de casos/ población expuesta) utilizando los Indicadores Locales de Asociación espacial (LISA) aplicadas a la tasa de ocurrencia.

Previamente, se obtuvo la matriz de pesos espaciales considerando las vecindades bajo la regla de frontera, es decir cuando al menos comparten parte de su frontera tal cual se observa en la parte izquierda de la figura 21. Se determinó que en promedio un distrito de Lima Metropolitana tiene aproximadamente 4.37 vecinos y la cantidad máxima y mínima de

vecindades fueron de 11 y uno respectivamente, tal cual se observa en el lado derecho de la figura 21.

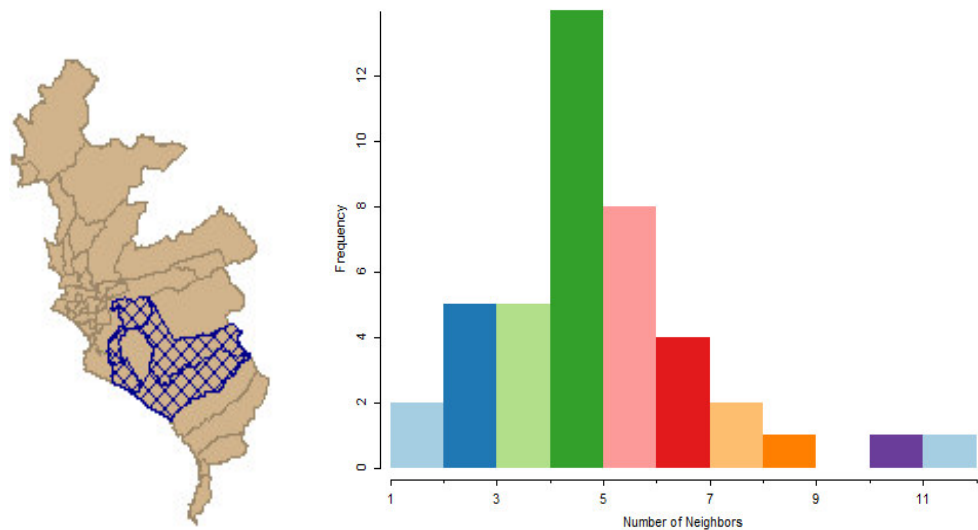


Figura 21. Mapa de conectividades e histograma de vecindades

Se calculó además el índice global de autocorrelación espacial de Moran y su correspondiente Scatterplot como un indicador solo de referencia. Estos se muestran en la figura 22. Su valor varía entre -1 y +1 y un valor cercano a cero indica ausencia de autocorrelación espacial. Este índice para nuestro caso resultó superior a cero lo cual evidencia la dependencia espacial existente en ambos casos, tanto para los atendidos por consulta externa del Hospital Santa Rosa al 2006 (Lado izquierdo) así como para el total de casos de IRAS registrados en Lima Metropolitana al 2015 (Lado derecho).

Además, la figura 22 nos permiten ver a la vez los cambios que puedan producirse de un año para otro tanto en la asociación espacial global (pendiente de la recta de regresión) como la propia asociación espacial local (nubes de puntos).

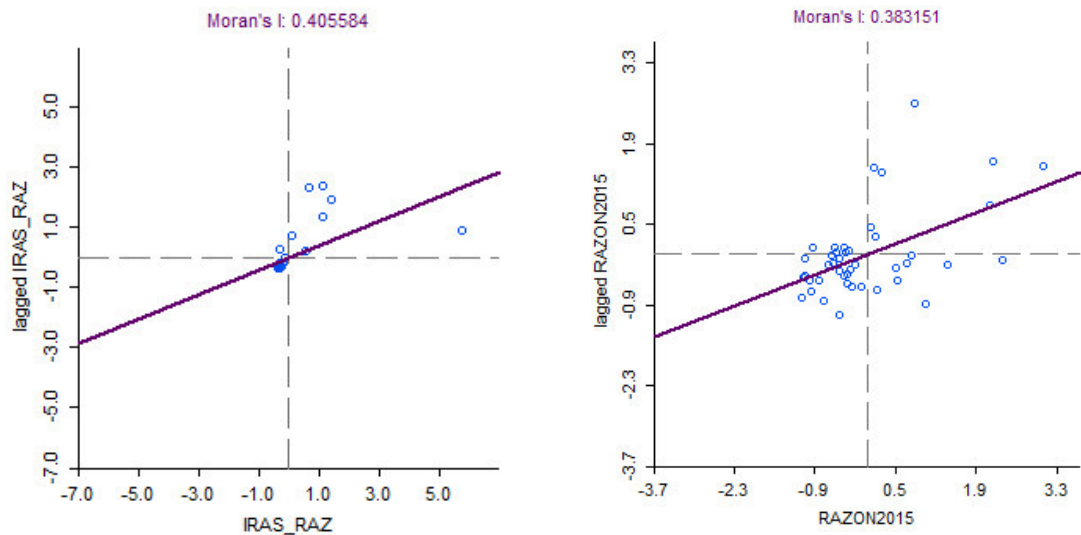


Figura 22. Scatterplot de Moran para los casos de IRAS del hospital Santa Rosa al 2006 y para el total de registrados en Lima Metropolitana al 2015.

Finalmente, en base a los indicadores locales de asociación espacial, se identificó la presencia de clústeres altamente significativo tal cual se muestra en la figura 23.

El lado izquierdo del gráfico corresponde a los casos atendidos por consulta externa del Hospital Santa Rosa en el 2006, para los cuales se identificaron los distritos que presentan altas tasas de pacientes con infecciones respiratorias. Las etiquetas HH, significan que estos distritos tienen valores altos (H) y están rodeados también de vecindades que tienen valores altos (H). Por tanto, estos distritos representan las zonas con alta demanda para el hospital Santa Rosa. Este clúster se encuentra resaltado en color rojo y está formado por los distritos de: Pueblo Libre, Breña, Jesús María, Magdalena del Mar, San Miguel y Lima.

El lado derecho del gráfico corresponde al total de casos de IRAS registrados en Lima Metropolitana en el 2015, para los cual de igual forma se identificaron los distritos que presentan las tasas más altas. Por tanto, estos distritos representan las zonas de Lima Metropolitana con mayor riesgo para contraer infecciones respiratorias. Este clúster se encuentra resaltado en color rojo y está formado por los distritos de: Santa María del Mar, Punta Negra, Punta Hermosa y San Bartolo y se ubican en la parte Sur de Lima

cercanas al mar. En el caso opuesto se identificó un clúster altamente significativo de distritos con riesgo más bajos formados por los distritos de San Isidro, Magdalena del Mar, San Borja y Surquillo resaltados en color azul.

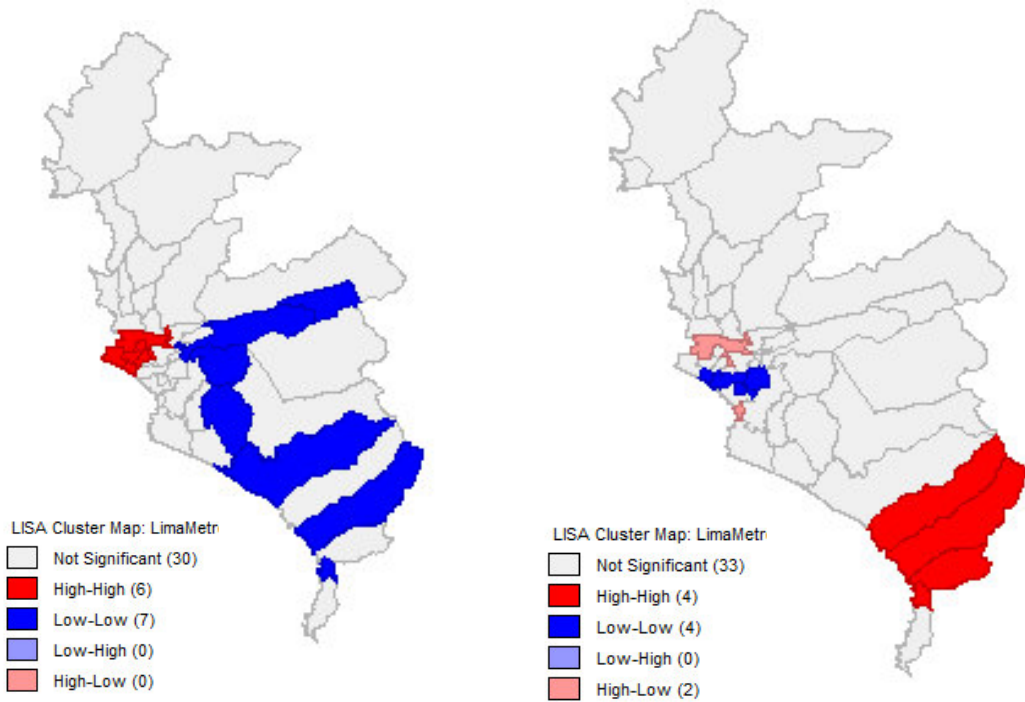


Figura 23. Detección de clúster espaciales para la tasa de incidencia de infecciones respiratorias de las consultas externas del Hospital Santa Rosa 2006 y para el total de caso de IRAS de Lima Metropolitana al 2015.

CONCLUSIONES Y RECOMENDACIONES

La morbilidad de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa – 2006 están distribuidas de forma no homogénea a lo largo de ciudad de Lima en la cual los distritos más cercanos al hospital requieren de mayores consultas externas de dicho Establecimiento de Salud.

La morbilidad para el total de casos de las enfermedades registrados en Lima Metropolitana – 2015 está distribuida de forma no homogénea a lo largo de ciudad de Lima.

A partir del índice de LISA y el índice de Morán se encontraron la presencia de clústeres significativos de distritos con alta demanda las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa. Estos distritos son: Pueblo Libre, Breña, Jesús María, Magdalena del Mar, San Miguel y Lima.

En cambio; en la ciudad de Lima utilizando el índice de Lisa los distritos con alta demanda fueron: San Bartolo, Chaclacayo, Punta Negra y Pucusana las que presentaron las tasas más altas de IRAS por encima de 36% de su población expuesta.

La distribución espacial de los casos de la morbilidad de las tres enfermedades más prevalentes del sistema respiratorio en consulta externa del Hospital Santa Rosa presenta un patrón espacial distinto comparado al total de casos de infecciones respiratorias agudas (IRAS) registrados en la ciudad de Lima.

Las herramientas del análisis exploratorio de datos espaciales proporcionan una gama muy variada de gráficos que permiten analizar distintos fenómenos considerando el lugar donde estos ocurren, lo cual resultan útiles en el análisis de los casos de infecciones respiratorias; razón por la cual se recomiendan utilizar.

REFERENCIAS BIBLIOGRAFICAS

- Anselin, L. (1988). *Spatial Econometric: Methods and Models*. Dordrecht: Kluwer Academic Publisher.
- Baddeley, A; Rubak, E; Turner, R. (2016). *Spatial Point Patterns Methodology and Applications with R*. CRC Press Chapman & Hall CRC BOOK.
- Buzai G. & Baxendale. C. (2009). *Análisis Exploratorio de Datos Espaciales*. Revista digital del Grupo de Estudios sobre Geografía y Análisis Espacial con Sistemas de Información Geográfica (GESIG), Sección Software y Metodología, (1), 1-11.
- Buzai, G. D. (2007). *Análisis exploratorio de la estructura espacial de enfermedades en la Ciudad de Luján*. Métodos Cuantitativos en Geografía de la Salud. Universidad Nacional de Luján.
- Buzai, G. D. (2011). *Métodos cuantitativos y Sistemas de Información Geográfica (SIG) aplicados a la Geografía de la salud. Análisis espacial de diagnósticos en la ciudad de Luján (Argentina)*. La salud ante los cambios globales. Conferencia llevada a cabo en la Universidad Autónoma del Estado de México, Toluca, México.
- Carmona Hernández, J. C. (2012). *Monitoreo de la infección respiratoria aguda en Manizales, mediante sistemas de información geográfica*. Archivos de Medicina (Col), 12 (1), 93-106.
- Chasco Yrigoyen, C. (2003). *Econometría espacial aplicada a la predicción extrapolación de datos microterritoriales*. Recuperado de <http://www.madrid.org/bvirtual/BVCM005618.pdf>.
- Chasco Yrigoyen, C. (2003). *Métodos gráficos del análisis exploratorio de datos espaciales*. Anales de economía aplicada. Recuperado de https://www.researchgate.net/publication/256174755_Metodos_graficos_del_analisis_exploratorio_de_datos_espaciales.
- Cressie, N. (1991). *Statistic for Spatial Data*. Canada: John Wiley & Sons, Inc.
- Dirección General de Investigación y Asuntos Ambientales del Servicio Nacional de Meteorología e Hidrología. (2011). *Evaluación de la calidad del aire en Lima Metropolitana 2011*. Recuperado de [http://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/E926CA7DFE4EE59905257CCA00528B1E/\\$FILE/1_EvalCalidadAireLima2011.pdf](http://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/E926CA7DFE4EE59905257CCA00528B1E/$FILE/1_EvalCalidadAireLima2011.pdf)

- Fuenzalida Díaz Manuel, M. F. (2013). *Análisis exploratorio de datos espaciales aplicado a MP10 y admisión hospitalaria. Evidencia para áreas urbanas chilenas contaminadas por humo de leña*. Revista digital del Grupo de Estudios sobre Geografía y Análisis Espacial con Sistemas de Información Geográfica (GESIG). Número (5), 109-128.
- Gómez-Barroso Diana, P.-F. M.-E. (2015). *Análisis espacial de la mortalidad por enfermedades cardiovasculares en la ciudad de Madrid, España*. Revista Española de Salud Pública, 27-37.
- Gonzales G, Z. Gonzales-Castañeda C, Nuñez D, Gastañaga C y Cabezas C, (2014). *Contaminación ambiental, variabilidad climática y cambio climático: una revisión del impacto en la salud de la población peruana*. Revista Peruana de Medicina Experimental y Salud Publica, 31(3), 547–556.
- Haining R., Wise S. y Signoretta P. (2000). *Providing scientific visualization for spatial data analysis: Criteria and an assessment of SAGE*. Journal of Geographical Systems, (2), 121-140.
- Hernández-Cadena L, Téllez-Rojo MM, Sanín-Aguirre LH, Lacasaña-Navarro M, Campos A, Romieu I. (2000). *Relación entre consultas a urgencias por enfermedad respiratoria y contaminación atmosférica en Ciudad Juárez, Chihuahua*. Rev. salud pública de México, 42 (4), 288-297.
- Lombana, D. F. (2014). *Análisis espacial del estudio de los casos de dengue en la Guajira 2013*. Bogota, Colombia: Universidad Distrital Francisco José de Caldas - Facultad de Ingeniería - Ingeniería Catastral y Geodesa.
- Ministerio de Salud - Dirección General de Epidemiología. (2011). *Análisis de Situación de Salud de la Provincia de Lima 2011*. Lima.
- Ministerio de Salud. (2002). *Análisis de la Situación de Salud - Guía para el análisis del proceso salud - enfermedad*. Lima.
- Moreno, R. y Vayá, E. (2000). *Técnicas econométricas para el tratamiento de datos espaciales: La econometría espacial*. Barcelona: Edicions Universitat de Barcelona, (1), 211-212.
- Oficina Ejecutiva del Planeamiento Estratégico - Hospital Santa Rosa. (2008). *El Plan Operativo Anual del Hospital Santa Rosa*. Lima.
- Oyarzún G, M. (2010). *Contaminación aérea y sus efectos en la salud*. Revista chilena de enfermedades respiratorias, 26(1), 1625. Recuperado de <https://dx.doi.org/10.4067/S0717-73482010000100004>.

- Robles D. (2014). *Análisis espacial del estudio de los casos de dengue en la Guajira 2013* (tesis de pregrado). Universidad Distrital Francisco José de Caldas. Colombia.
- Rocco C. (2012). *Análisis Exploratorio De Datos Espaciales En Estudio De Demanda Eléctrica*. III Congreso Venezolano de Redes y Energía Eléctrica. Venezuela.
- Romero-Placeres M, Más-Bermejo P, Lacasaña-Navarro M, Rojo-Solís M, Téllez MM, Aguilar-Valdés J, y Romieu I. (2004). *Contaminación atmosférica, asma bronquial e infecciones respiratorias agudas en menores de edad, de La Habana*. Salud Pública de México, 46(3), 222-233.
- Unwin, A. (2000), *Using your eyes- making statistics more visible with computers*. Computational Statistics & Data Analysis, 32; pp. 303-312.
- Veregin, H. y Hargitai, P. (1995). *An evaluation matrix for geographical data quality*. In S. C. Guptill & J. L.

ANEXOS

ANEXO 1: BASE DE DATOS UTILIZADA

Código de UBIGEO	Distrito	Rinofaringitis Aguda, 2006	Asma NE, 2006	Bronquitis Aguda, 2006	Total IRAS, 2006	Población proyectada al 2006	IRAS MINSA, 2015	Población proyectada al 2015	Distancia en kilómetros al Hospital Santa Rosa
150119	Lurin	2	7	1	10	62,396	19,255	85,132	39.90
150102	Ancón	10	3	2	15	33,312	12,481	43,382	41.60
150131	San Isidro	18	7	3	28	61,321	2,128	54,206	19.10
150120	Magdalena del Mar	280	114	55	449	53,846	4,208	54,656	1.80
150130	San Borja	9	6	1	16	110,848	6,303	111,928	10.90
150116	Lince	131	79	47	257	58,606	3,422	50,228	5.10
150140	Santiago de Surco	67	20	19	106	293,534	15,035	344,242	16.70
150121	Pueblo Libre	1,306	649	325	2,280	78,723	10,762	76,114	1.00
150139	Santa Rosa	0	0	0	0	10,387	3,291	18,751	42.00
150106	Carabayllo	18	12	14	44	209,901	36,357	301,978	38.70
150113	Jesus Maria	280	123	101	504	70,914	3,665	71,589	2.50
150124	Pucusana	0	0	0	0	10,242	6,163	17,044	70.90
150138	Santa Maria del Mar	0	0	0	0	699	374	1,608	61.00
150108	Chorrillos	34	14	12	60	293,188	33,843	325,547	16.70
150104	Barranco	14	8	1	23	36,138	7,517	29,984	12.90
150143	Villa Maria del Triunfo	40	23	7	70	383,766	49,663	448,545	29.80
150133	San Juan de Miraflores	34	15	16	65	371,620	67,542	404,001	20.20
150122	Miraflores	31	24	14	69	89,371	3,343	81,932	7.80
150141	Surquillo	20	6	12	38	94,223	9,363	91,346	11.70
150127	Punta Negra	0	0	0	0	5,146	2,895	7,934	53.50
150134	San Luis	23	13	20	56	56,763	5,939	57,600	11.80
150136	San Miguel	574	239	132	945	134,246	5,906	135,506	4.20
150142	Villa El Salvador	10	11	3	24	385,609	44,799	463,014	30.00
150126	Punta Hermosa	0	0	0	0	5,736	1,231	7,609	53.40
150129	San Bartolo	0	0	0	0	5,784	3,468	7,699	57.50
150123	Pachacamac	0	0	0	0	64,053	28,426	129,653	51.20
150115	La Victoria	141	56	49	246	205,113	21,191	171,779	8.00
150114	La Molina	28	13	11	52	132,343	14,400	171,646	19.10
150105	Breña	233	137	62	432	86,609	7,086	75,925	3.10
150137	Santa Anita	31	10	26	67	185,828	25,165	228,422	18.10
150109	Cieneguilla	0	0	0	0	25,331	7,334	47,080	41.10
150101	Lima	391	201	105	697	317,755	54,951	271,814	5.00
150111	El Agustino	22	14	15	51	186,246	24,743	191,365	9.00
150128	Rimac	34	14	10	58	185,950	32,978	164,911	11.00
150103	Ate	49	28	20	97	473,918	57,158	630,086	26.00
150107	Chaclacayo	0	0	0	0	42,569	16,562	43,428	39.00
150112	Independencia	36	23	15	74	215,256	26,245	216,822	11.00
150135	San Martín de Porres	312	160	132	604	582,902	40,416	700,177	17.60
150117	Los Olivos	143	63	42	248	323,462	43,212	371,229	15.00
150110	Comas	54	38	22	114	501,837	61,017	524,894	19.00
150118	Lurigancho	0	0	0	0	169,191	36,186	218,976	22.00
150132	San Juan de Lurigancho	213	90	62	365	902,687	125,003	1,091,303	19.00
150125	Puente Piedra	17	7	8	32	227,168	34,102	353,327	26.00

ANEXO 2: TABLAS ESTADÍSTICAS

Tabla 5

Estadísticos descriptivos de rinoфарингитis aguda

		Estadístico
Rinofaringitis aguda	Media	107,09
	Mediana	28,00
	Varianza	50662,610
	Desv. típ.	225,084
	Mínimo	0
	Máximo	1306
	Rango	1306
	Amplitud intercuartil	129
	Asimetría	4,041
	Curtosis	19,529

Tabla 6

Estadísticos descriptivos de asma no especificado

		Estadístico
Asma no especificado	Media	51,79
	Mediana	13,00
	Varianza	11985,598
	Desv. típ.	109,479
	Mínimo	0
	Máximo	649
	Rango	649
	Amplitud intercuartil	53
	Asimetría	4,239
	Curtosis	21,494

Tabla 7
Estadísticos descriptivos de bronquitis aguda no especificada

		Estadístico
Bronquitis aguda no especificada	Media	31,72
	Mediana	12,00
	Varianza	3323,444
	Desv. típ.	57,649
	Mínimo	0
	Máximo	325
	Rango	325
	Amplitud intercuartil	41
	Asimetría	3,581
	Curtosis	15,850