

MODELO ESTADÍSTICO PARA POSIBLE PUNTO DE ATENCIÓN EN UNA ENTIDAD FINANCIERA

MABEL CAROLINA GALINDRES BERNAL

MARÍA CAMILA FERNÁNDEZ BOLAÑOS

SANTIAGO OSORIO SIERRA

Nota de Aceptación

Certificamos que el presente Trabajo de Grado Satisface, en alcances y calidad, todos los requisitos que demanda un Trabajo de Grado de Maestría.

David Arango Londoño

Director

Mauricio B.

Jurado

[Signature]

Jurado

Aprobado en cumplimiento de los requisitos exigidos por la Pontificia Universidad Javeriana Cali, para optar el título de Magister en Ciencia de Datos.

Camilo Rocha

HERNÁN CAMILO ROCHA NIÑO Ph. D.
Decano Facultad de Ingeniería y Ciencias

[Signature]

JUAN CARLOS MARTÍNEZ ARIAS
Director Posgrados de Ingeniería y Ciencias

Santiago de Cali, 29_de Mayo del 2023

Doctora

Gloría Inés Alvarez V.

Directora Maestría en Ciencia de Datos

Facultad de Ingeniería y Ciencias

Pontificia Universidad Javeriana de Cali

Con el fin de cumplir con los requisitos exigidos por la Universidad para optar por el título de Magíster en Ciencia de Datos, nos permitimos presentar a su consideración el proyecto denominado “Modelo Estadístico para posible punto de atención en una entidad financiera”, el cual fue realizado por el (los) estudiantes Carolina Galindres, María Camila Fernández y Santiago Osorio con códigos · 0225388, 971545 y 8926997 pertenecientes a la Maestría en Ciencia de Datos, bajo la dirección de David Arango Londoño.

El suscrito director del Trabajo de Grado autoriza para que se proceda a hacer la evaluación de este Proyecto ante el Tribunal que para el efecto se designe, toda vez que ha revisado cuidadosamente el documento y avala que ya se encuentra listo para ser presentado oficialmente.


Atentamente,



Carolina Galindres 1130671230 de Cali



David Arango Londoño 1130586950 de Cali



Camila Fernández 1061088201 de Florencia



Santiago Osorio 1144029865 de Cali

Documentación anexa:

Resumen del Proyecto Aplicado en formato digital (máximo 1 página).

Una copia digital (PDF) del documento del proyecto aplicado



Bogotá D. C., 04 de julio de 2023

Señores
PONTIFICIA UNIVERSIDAD JAVERIANA
Atn, Juan Carlos Martínez Arias, Director Posgrados de Ingeniería y Ciencias
Cali

Cordial saludo

BANCOLOMBIA S.A. concede autorización al funcionario Santiago Osorio Sierra para hacer uso de la información en LZ tabla maestro_transaccional_canales donde se extrajo las transacciones financieras y exitosas realizadas en sucursal física de los clientes pertenecientes al segmento Empresarial de gerencias basadas en la ciudad de Cali , de igual manera autoriza a ser compartida con el fin de hacer uso de la misma en la investigación y para tratar sus datos con fines académicos.

Estamos a su disposición para confirmar la anterior información, a través de los siguientes números:
Bogotá 4886000 Ext 16779 celular 3105747365.

Atentamente,



Lina Marcela Guzmán Vásquez
Jefe de Sección
Gerencia de Servicios Regionales Especializados

**Maestría en Ciencia de Datos
Facultad de Ingeniería y Ciencias**

FICHA RESUMEN
PROYECTO DE TRABAJO DE GRADO

TÍTULO: MODELO ESTADÍSTICO PARA POSIBLE PUNTO DE ATENCIÓN EN UNA ENTIDAD FINANCIERA

1. ÉNFASIS: Ciencia de Datos
2. TIPO DE PROYECTO: Aplicado
3. ÁREA DE TRABAJO: EMAP, Estadística y Matemática Aplicada
4. ESTUDIANTE(S):
 - 0225388 Carolina Galindres
 - 8971545 María Camila Fernández
 - 8926997 Santiago Osorio
5. CORREO ELECTRÓNICO:
 - mcgb1130@javerianacali.edu.co
 - mcfernandezb@javerianacali.edu.co
 - sosorio25@javerianacali.edu.co
6. DIRECCIÓN Y TELEFONO:
 - Calle 12ª #31-63 Colseguros, Cali - 3043772450
 - Calle 34 BN #4b-30 casa 359 Popayán Cauca - 3113540217
 - Calle 32AN # 2B – 55 AP 901A, Cali - 3176612342
7. DIRECTOR: David Arango Londoño
8. VINCULACIÓN DEL DIRECTOR: Docente Universidad Javeriana Cali
9. CORREO ELECTRÓNICO DEL DIRECTOR:
 - david.arango@javerianacali.edu.co
10. GRUPO O EMPRESA QUE LO AVALA: Bancolombia
11. PALABRAS CLAVE (al menos 5): Modelo Estadístico, Georreferenciación, Patrones puntuales, indicadores, imágenes raster, Análisis de Componentes Principales.
12. . ODS QUE APLICA EL PROYECTO (Agenda 2030): 8. Trabajo decente y crecimiento económico
13. FECHA DE INICIO: 4 de julio 2022

14. **RESUMEN:** Para el desarrollo del presente trabajo, se realizó un estudio basado en los datos recolectados de las transacciones efectuadas por clientes empresariales y corporativos durante el primer semestre de 2022 en una entidad financiera de la ciudad de Santiago de Cali. El objetivo de este estudio fue desarrollar un modelo estadístico mediante la generación de indicadores, con el fin de identificar posibles zonas óptimas para abrir un punto de servicio adicional al existente (Oficina de Atención Empresarial), utilizando algún punto de atención asignados para atender a personas naturales.

En este sentido, se llevó a cabo un proceso inicial de comprensión y análisis de los datos, seguido de un estudio exploratorio espacial de los datos. Posteriormente, se definieron cinco variables de estudio, a partir de las cuales se generaron imágenes de densidad y raster. Estas imágenes permitieron obtener cuatro indicadores que delimitarían las zonas óptimas.

Finalmente, con base al análisis y comparativo de los gráficos de los indicadores, se seleccionó la ubicación del punto de servicio para clientes empresariales mediante la metodología de Análisis de Componentes Principales, generada en el indicador 4. El análisis previo determinó que la Oficina de San Nicolás Cali, situada en la Cl. 20 #5-49, COMUNA 3, Cali, Valle del Cauca, sería la zona elegida para establecer el posible punto de atención.



Pontificia Universidad
JAVERIANA
Cali

**MODELO ESTADÍSTICO PARA POSIBLES PUNTOS DE ATENCIÓN EN UNA ENTIDAD
FINANCIERA**

*0225388 Carolina Galindres
8971545 María Camila Fernández
8926997 Santiago Osorio*

*Proyecto Aplicado para optar al título de
Magíster en Ciencia de Datos*

Director(a)
David Arango Londoño

FACULTAD DE INGENIERÍA Y CIENCIAS
MAESTRÍA EN CIENCIA DE DATOS
SANTIAGO DE CALI, JUNIO 1 DE 2023

TABLA DE CONTENIDO

	Pág.
INTRODUCCIÓN	9
1. DEFINICIÓN DEL PROBLEMA.....	11
1.1. PLANTEAMIENTO DEL PROBLEMA.....	11
1.2. FORMULACIÓN DEL PROBLEMA	11
2. OBJETIVOS DEL PROYECTO.....	12
2.1. OBJETIVO GENERAL	12
2.2. OBJETIVOS ESPECÍFICOS	12
3. MARCO DE REFERENCIA	13
3.1. MARCO TEÓRICO	13
3.2. ANTECEDENTES	17
4. ENTENDIMIENTO DE LOS DATOS	20
4.1. DESCRIPCIÓN DE VARIABLES.....	20
5. ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES	25
5.1 DISTRIBUCIÓN DE LOS DATOS	25
5.2 CONTABILIZACIÓN MEDIANTE MAPAS INTERACTIVOS	28
5.3 CONTEO TRANSACCIONES Y NUMERO DE CLIENTES POR COMUNAS.....	30
6. IMÁGENES DE DENSIDAD Y ZONAS ÓPTIMAS	32
6.1. VARIABLES DE ANÁLISIS.....	32
6.2. IMÁGENES DE DENSIDAD	32
6.3. IMÁGENES RASTER DE VARIABLES DE ESTUDIO	34
6.4. ESCENARIOS PROPUESTOS Y CÁLCULO DE INDICADORES	36
7. ANÁLISIS COMPARATIVO DE INDICADORES Y SELECCIÓN DE LA MEJOR OPCIÓN PARA LA DELIMITACIÓN DE LA ZONA ÓPTIMA	45
7.1. INDICADOR 1 y 4 – Punto de servicio San Nicolas Cali (Of. 62).....	45
7.2. UBICACIÓN DEL POSIBLE PUNTO DE ATENCIÓN	46
8. CONCLUSIÓN Y TRABAJOS FUTUROS	48
8.1. CONCLUSIONES	48
8.2. TRABAJOS FUTUROS.....	49
BIBLIOGRAFÍA.....	51

TABLA DE ILUSTRACIONES

	Pag.
Ilustración 1. PPP – Puntos de servicio	26
Ilustración 2. PPP – Ubicaciones de clientes	26
Ilustración 3. Prueba Chi2 – Puntos de servicio	26
Ilustración 4. Prueba Chi2 – Ubicación Clientes	27
Ilustración 5. Función de distribución Kest – Puntos de servicio.....	27
Ilustración 6. Función de distribución Kest - Ubicación de los clientes	28
Ilustración 7. Clusters Ubicación de Clientes	29
Ilustración 8. Clusters Transacciones	30
Ilustración 9. Conteo por comunas – Ubicaciones de clientes	31
Ilustración 10. Cantidad de Transacciones	33
Ilustración 11. Ubicación de los clientes	33
Ilustración 12. Vías	33
Ilustración 13. Uso de Suelo	33
Ilustración 14. Imagen raster de las variables	35
Ilustración 15. Imagen raster de las variables escaladas	36
Ilustración 16. Imagen estática indicador 1	38
Ilustración 17. Imagen dinámica indicador 1	38
Ilustración 18. Imagen estática indicador 2	39
Ilustración 19. Imagen dinámica indicador 2	39
Ilustración 20. Imagen estática indicador 3	41
Ilustración 21. Imagen dinámica indicador 3	41
Ilustración 22. Distribución gráfica de componentes principales – Indicador 4	42
Ilustración 23. Imagen estática indicador 4 - ACP	44
Ilustración 24. Imagen dinámico indicador 4 -ACP	44
Ilustración 25. Zona óptima Indicador 1	45
Ilustración 26. Puntos de atención zona óptima	45
Ilustración 27. Punto de servicio – San Nicolas (Of. 62)	47

LISTA DE TABLAS

	Pag.
Tabla 1. Variables de base de datos	20
Tabla 2. Variables de base de datos	21
Tabla 3. Características de Monto de transacciones	22
Tabla 4. Librerías y funciones utilizadas	23
Tabla 5. Variables analizadas	32
Tabla 6. Variable distancia Oficina de servicio especializado	34
Tabla 7. Escenarios de análisis de indicadores	37
Tabla 8. Resultado de la encuesta	40
Tabla 9. Porcentajes de resultado de encuesta	40
Tabla 10. Matriz de Correlaciones de Variables	42
Tabla 11. Componentes principales con la mayor parte de la variación en los datos	43
Tabla 12. Porcentajes de prioridad componentes principales – Indicador 4	43

INTRODUCCIÓN

El sistema financiero es muy importante para el desarrollo y crecimiento de una economía, este es definido como el conjunto de instituciones y operaciones por medio de las que se transfieren recursos de un sector superavitario a uno deficitario (empresas o personas que requieren financiación). Este sistema comprende el mercado bancario (intermediario) y el de valores (no intermediados). Para este primero, está compuesto por instituciones que captan recursos del público, principalmente en forma de depósitos y su estructura está formada por establecimientos de crédito y demás entidades que realizan operaciones similares; dentro de la primera categoría mencionada se encuentran los bancos comerciales, corporaciones financieras y las compañías de financiamiento [1]. Por ejemplo, a través los bancos comerciales los clientes pueden acceder a diversos productos tales como cuentas de ahorro, tarjeta de crédito y diferentes líneas de financiamiento para vivienda, vehículo y estudio.

Dentro de las tendencias actuales del mercado y las demandas orientadas hacia la sostenibilidad a largo plazo, se destaca que un excelente servicio al cliente se fundamenta en el comportamiento y el cumplimiento de la organización frente a las necesidades de quienes reciben dicho servicio [2]. Con base en estas exigencias, se reconoce la necesidad que tienen las entidades financieras de establecer puntos estratégicos de atención para sus clientes. En el caso específico de este estudio, la entidad analizada se distingue por contar únicamente con una sucursal especializada para los clientes de los segmentos Empresarial y Corporativo en la ciudad de Cali. En esta sucursal se atiende exclusivamente a clientes de este tipo y a sus apoderados, quienes realizan solicitudes de servicios de diferentes tipos. No obstante, muchos de estos clientes se encuentran dispersos por toda la ciudad, lo que dificulta su acceso al punto asignado y los lleva a dirigirse a las sucursales destinadas asignadas para personas naturales, donde pueden ser atendidos, pero sin recibir un trato exclusivo y preferencial. En este contexto, el presente estudio propone una posible solución para mejorar la atención a los clientes de dicho segmento, desarrollando un modelo de indicadores que permita identificar la ubicación óptima en la ciudad de Cali para establecer un punto estratégico de atención a clientes empresariales, aprovechando los puntos de atención disponibles asignados para personas naturales.

Para lograr esto, se llevó a cabo una serie de etapas. En primer lugar, se realizó la consolidación de los datos necesarios, seguida de su georreferenciación y organización. Posteriormente, se llevó a cabo un

análisis exploratorio de los datos, seguido del desarrollo de un modelo estadístico no paramétrico. Este modelo fue analizado mediante la generación de cuatro indicadores diferentes, con el objetivo de seleccionar el indicador más adecuado que permitiera definir la zona óptima para la apertura de los posibles puntos de atención. Durante el análisis, se tomaron en cuenta variables como las ubicaciones de los clientes, el nivel de transaccionalidad realizado en los diferentes puntos de servicio, la ubicación de vías principales y secundaria, el uso del suelo y la distancia al punto de atención principal

En este sentido, el documento está estructurado en la siguiente manera: se inicia con la definición del problema, seguido de los objetivos generales y específicos, la justificación, el marco de referencia, solución a la problemática, conclusiones, trabajos futuros y referencias bibliográficas.

1. DEFINICIÓN DEL PROBLEMA

1.1. PLANTEAMIENTO DEL PROBLEMA

El servicio al cliente en el sector financiero se ha convertido en uno de los diferenciales más importantes para ser competitivos, sin embargo, estos se han visto marcados por los bajos índices de satisfacción en relación al servicio prestado. Esto ha llevado a que las entidades financieras se reinventen y busquen maneras de mejorar sus procesos y la prestación de servicios.

En un estudio realizado por la empresa Customer Index Value en 2019, se presentó un informe de satisfacción y cercanía emocional de los clientes hacia los bancos. Este informe indicó que los consumidores financieros calificaron la banca tradicional colombiana con 73 sobre 100, de acuerdo con el índice ACSI (American Customer Satisfaction Index). Esto representó una fuerte caída en comparación con 2018 y alcanzó niveles bajos similares a los presentados en 2015 [3].

Un punto importante en cuanto al servicio y experiencia es estar presente en los canales en los que los clientes solicitan servicios. Es más probable que los consumidores se queden con los bancos e instituciones financieras que facilitan el proceso de comunicación, haciéndolo más rápido, fácil, de calidad y oportuno. Por tanto, comprender los canales más utilizados por los clientes y sus preferencias al solicitar un servicio se convierte en un punto de partida para buscar formas de crear experiencias más atractivas y duraderas, lo que también puede resultar en una mayor fidelización.

Por ello, en este proyecto se desarrolló un modelo de ciencia de datos que analiza diferentes variables con el fin de determinar los posibles puntos de atención a este segmento. Se aprovechan los puntos de servicio ofrecidos a personas naturales para mejorar el servicio y la experiencia de los clientes.

1.2. FORMULACIÓN DEL PROBLEMA

¿Es posible identificar, a través del uso de herramientas de ciencia de datos y el análisis de la información, puntos estratégicos de atención para los segmentos corporativo y empresarial, con el objetivo de mejorar el servicio prestado?

2. OBJETIVOS DEL PROYECTO

2.1. OBJETIVO GENERAL

Desarrollar un análisis estadístico no paramétrico que permita identificar un posible punto de atención para clientes empresariales en las sucursales de servicio para persona natural de una entidad financiera con base en el análisis e integración de diferentes variables.

2.2. OBJETIVOS ESPECÍFICOS

- Realizar entendimiento de los datos, identificando y definiendo variables de interés relacionadas con la base de datos de las transacciones realizadas por los clientes empresariales en el primer periodo de 2022 en la Ciudad de Cali.
- Realizar un análisis exploratorio de datos utilizando las variables definidas en el punto anterior, con el fin de identificar tendencias y comportamientos en la distribución de la información geográficamente en el mapa de Cali.
- Generar imágenes de densidad o mapas de calor basados en las variables estudiadas, y combinarlas entre sí para crear diferentes escenarios e indicadores que permitan delimitar zonas óptimas para la habilitación del punto de servicio a clientes empresariales.
- Seleccionar el punto de servicio más adecuado dentro de las opciones de las zonas óptimas identificadas, tomando en cuenta los indicadores generados y las características actuales del punto de servicio ubicado en estas zonas.

3. MARCO DE REFERENCIA

3.1. MARCO TEÓRICO

3.1.1 Patrones de Puntos

Un patrón espacial de puntos es un conjunto finito de eventos $[U_i: i=1, \dots, n]$ para alguna región plana $W \subset \mathbb{R}^2$. Muy a menudo, W es una ventana de muestreo dentro de una región mucho más grande y es razonable considerar el patrón de puntos como una realización parcial del proceso puntual en el plano, los eventos consisten en todos los puntos del proceso que se encuentran dentro de W . Bajo la suposición de estacionariedad (el proceso es homogéneo o invariante a la translación) e isotropía (el proceso es invariante a la rotación), las características principales de un proceso de puntos pueden ser resumidas por su propiedad de primer y por su propiedad de segundo orden, que describe las relaciones entre pares de puntos (la probabilidad de encontrar un punto en las inmediaciones de otro). En el caso de patrones uniformes o regulares, la probabilidad de encontrar un punto en las inmediaciones de otro es menor de la que tendría un patrón aleatorio mientras que en los patrones agrupados la probabilidad es mayor.

La aleatoriedad espacial completa (AEC), de un proceso puntual espacial, se caracteriza por la ausencia de estructura en los datos. Por tanto, habitualmente es la hipótesis nula de un test estadístico para determinar cuando hay estructura espacial en un patrón puntual dado. El principal interés del AEC es que representa un modelo ideal que, a pesar de no alcanzarse en la práctica, resulta muy útil como primera aproximación de algunos patrones observados.

La función de intensidad es conocida como mapa de calor o mapa de zonas calientes. Su interpretación se suele basar en una escala de colores tal que las zonas más intensas representan el área donde se agrupan y conglomeran un alto número de eventos. Los mapas de calor son adecuados para visualizar grandes cantidades de datos espaciales y pueden usarse para identificar clústeres [4].

3.1.2 Análisis Exploratorio de Datos Espaciales

El análisis exploratorio de datos espaciales tiene la finalidad de examinar los datos previamente a la aplicación de cualquier técnica estadística, de tal forma que se entienden con facilidad los datos y las relaciones entre variables a analizar; proporcionando métodos ágiles para organizar, preparar y detectar fallas en el diseño o casos atípicos de los datos.

Inicialmente se debe estudiar la información georreferenciada con el fin de identificar localización, variabilidad, observaciones extremas, cálculo de medidas de localización variabilidad y correlación, para definir el procedimiento de predicción más pertinente. De igual manera se debe realizar un estudio exploratorio clásico, cuando se dispone de información georreferenciada, mediante histogramas, diagramas de tallos y hojas y/o diagramas de caja y bigotes con el propósito de identificar localización, variabilidad, forma y observaciones extremas.

La estadística espacial como una de las principales herramientas de análisis exploratorio de los datos ha proporcionado una serie de herramientas con el fin de analizar cualquier conjunto de datos de puntos en el plano y en particular, los procesos puntuales espaciales proporcionan una gran variedad de herramientas para modelar comportamientos particulares de fenómenos espaciales discretos identificados en el Sistema de Información Geográfica (GIS) [5].

El sistema de información geográfica GIS, es una herramienta imprescindible para el manejo de los datos asociados al terreno, ya que la gran cantidad de variables que interactúan en el medio natural hacen necesario el empleo de programas informáticos en los que se recoja simultáneamente información gráfica y bases de datos asociadas a ésta (información no gráfica o alfanumérica). Así, permiten trabajar con gran cantidad de información y realizar complejas operaciones sobre un determinado escenario [5].

3.1.3 Estadística Espacial

La estadística espacial es aquella variante de la estadística que se dedica a analizar la información de las variables de interés, pero haciendo énfasis en su ubicación espacial. En este caso, los atributos medidos que son área de estudio por parte del investigador pueden ser georreferenciados ubicándolos en el espacio o mejor contextualizado, en un mapa. Uno de los fundamentos más importantes es la determinación de la estructura de la correlación del componente espacial para poder realizar predicciones de la variable de interés en determinado lugar.

La estadística espacial analiza los eventos o situaciones de un proceso estocástico $[Y(u): u \in D]$, donde $u \in R^d$ y representa una ubicación en el espacio euclidiano d -dimensional, $Y(u)$ es una variable aleatoria en la

ubicación u la cual, a su vez, varía sobre un conjunto de índices $D \subset \mathbb{R}^d$. Dentro de la estadística espacial como macro tema y rama importante de la estadística, se encuentran varias áreas de estudio diferentes que tienen, a su vez, propiedades inherentes a su campo. Tres de esas importantes áreas corresponden a la geoestadística, lattices o áreas y a los patrones espaciales.

En la geoestadística tenemos que las ubicaciones provienen de un conjunto D continuo y deben ser particularmente seleccionadas bajo el juicio del investigador. Se tiene por ejemplo que esta metodología puede tratar datos como los niveles de un contaminante en diferentes sitios de una parcela o los valores de precipitación en una ciudad medida en diferentes estaciones meteorológicas. Nótese que en los ejemplos anteriores se puede identificar claramente que hay una continuidad espacial.

Por otro lado, en la metodología de lattices o áreas, las ubicaciones u pertenecen a un conjunto D discreto y, al igual que su predecesora, son seleccionadas por el investigador. Esta técnica se puede utilizar con datos como por ejemplo la tasa de morbilidad de hepatitis en una ciudad mediada por distritos censales o la distribución espacial de la segregación residencial en la capital de un país.

A diferencia del anterior tema, se observa que el conjunto de ubicaciones es claramente discreto. Por su parte, en los patrones espaciales, las ubicaciones u pertenecen a un conjunto D discreto y su selección es independiente del investigador, es decir, ahora se cuenta con un D aleatorio. Fenómenos tales como terremotos, incendios y contagios de enfermedades se registran utilizando una ubicación cartográfica con los elementos del fenómeno que ocurre en la región geográfica correspondiente [6].

3.1.4 Modelo Raster:

Los modelos raster se utilizan comúnmente en sistemas de información geográfica (SIG) y en análisis espaciales. Proporcionan una forma eficiente de almacenar y analizar grandes conjuntos de datos espaciales, ya que los datos se organizan en una estructura de cuadrícula que es fácil de manejar y procesar.

A diferencia de los modelos vectoriales, que representan características geográficas como puntos, líneas y polígonos, los modelos raster representan características como superficies continuas. Esto los hace especialmente adecuados para el análisis de fenómenos que varían de manera continua en el espacio, como

la topografía, el clima y la vegetación.

El modelo Raster, hace que el espacio esté dividido en celdas (o píxeles), de manera que cualquier información del mundo real se codifica sobre un formato espacial, al sobreponer la retícula o malla sobre el territorio. A través de un sistema de referencia apropiado, se establece una relación biyectiva entre las celdas de una cuadrícula y un conjunto de áreas elementales de la superficie terrestre. La información temática queda reflejada sobre cada una de las celdas resultantes de la división del espacio, en elementos discretos, regulares, en general, de forma cuadrada. Otros sistemas utilizan retículas triangulares o hexagonales, donde cada elemento espacial sería un triángulo equilátero o un hexágono regular [7].

3.1.5 Indicadores Multivariados

Los indicadores multivariados son medidas estadísticas utilizadas para representar y resumir la relación entre múltiples variables en un conjunto de datos. A diferencia de los indicadores univariados, que se enfocan en una sola variable, los indicadores multivariados consideran múltiples variables simultáneamente.

Estos indicadores permiten analizar la estructura y las interacciones entre variables en un conjunto de datos, lo que proporciona una visión más completa y detallada de las relaciones entre ellas. Entre estos se encuentran: Covarianza, correlación, Análisis de Componentes Principales (PCA) y Análisis de Conglomerados (Cluster análisis).

Estos indicadores permiten identificar distintas situaciones de alejamiento del producto respecto de las especificaciones. Las comparaciones del comportamiento de cada índice se llevan a cabo frente a diferentes escenarios que combinan número de variables, intensidad de las asociaciones entre ellas, tamaños de muestra y capacidad del proceso [8].

3.1.6 Georreferenciación

Se entiende que una información está georreferenciada cuando se conocen las coordenadas geográficas, latitud y longitud, de un punto [9].

3.1.7 Modelo estadístico

Un modelo estadístico es una representación matemática o formal de un fenómeno o proceso real que se basa en la

teoría y los métodos estadísticos. Estos modelos se utilizan para describir, explicar, predecir y tomar decisiones sobre datos observados o generados a partir de un experimento. Un modelo estadístico se compone de tres elementos principales: Variables, Parámetros, Distribución de Probabilidad.

El objetivo de un modelo estadístico es utilizar los datos observados para estimar los parámetros desconocidos y realizar inferencias sobre fenómenos o procesos subyacentes. Lo cual implica que se deba realizar análisis de regresión, pruebas de hipótesis, análisis de varianza, entre otros procedimientos estadísticos.

Un modelo estadístico también se puede ver como una representación formal de un sistema real, con el que se pretende hacer predicciones y ayudar a el control [10].

3.1.8 Clientes empresariales

Para la entidad financiera en estudio, los clientes empresariales corresponden a una Persona Natural o Persona Jurídica con Ventas Anuales mayores a 13.000 millones y menores a 100.000 millones de pesos, realizadas en el establecimiento de crédito.

3.2. ANTECEDENTES

3.2.1. Análisis Geoespacial Para Ubicar Cajeros Automáticos y Oficinas Bancarias Como Herramienta para Mejorar la Cobertura de Servicios de las Entidades Financieras en Bogotá

Sandra Bibiana Muñoz Díaz [11] de la Universidad Militar Nueva Granada realizó el trabajo de grado referente al “Análisis Geoespacial Para Ubicar Cajeros Automáticos y Oficinas Bancarias Como Herramienta para Mejorar la Cobertura de Servicios de las Entidades Financieras en Bogotá”. Este trabajo se llevó a cabo debido al inconformismo de los servicios prestados por las entidades bancarias. Una alternativa consistió en mejorar los desplazamientos a puntos de distribución del servicio como son oficinas y cajeros automáticos en diferentes horarios, la necesidad de ubicar estratégicamente sus puntos de atención se convirtió en un factor diferenciador de gran importancia. El trabajo se llevó a cabo porque según las estadísticas de la Superintendencia Financiera de Colombia se evidencia inconformismo frente a los servicios prestados por las entidades bancarias, sin embargo, una alternativa de mejora era que los servicios ofrecidos por las entidades financieras contengan características de accesibilidad física para los usuarios, donde éstos deben desplazarse a puntos de distribución del servicio como son oficinas y cajeros

automáticos en diferentes horarios. La necesidad de ubicar estratégicamente sus puntos de atención se convierte en un factor diferenciador de gran importancia.

Inicialmente se realizó la caracterización del sector financiero en la ciudad de Bogotá, posteriormente se identificaron las herramientas de análisis geoespacial como Sistemas de Información Geográfica SIG, aplicables al sector financiero y se analizaron las diferentes variables que se tuvieron en cuenta en procesos decisorios acerca de la ubicación de puntos de atención financiera.

Como resultado, obtuvieron, mapas digitales con la ubicación de clientes, oficinas y cajeros, analizando variables como ingresos, familia, educación, situación económica, patrimonio, entre otras, lo que les permitió generar mapas temáticos que visualizan y clasifican los clientes por estas variables.

3.2.2. Cobertura Hospitalaria en la ciudad de Santiago de Cali

Este proyecto se realiza en el año 2020 bajo el contexto del programa de Data Science For All, donde se analiza la accesibilidad a los servicios de salud en Colombia utilizando como fuente de información encuestas quinquenales desarrolladas por el Ministerio de Salud, matrices de origen y destino y análisis indirectos.

Estos análisis que se realizaban sobre la accesibilidad a los servicios de salud representaban muchos contratiempos, específicamente en 3 aspectos: son costosos, no son sensibles a los cambios en el tráfico y se demoraban mucho. Un análisis geoespacial utilizando ciencias de datos y Big Data ofrecería una alternativa mucho más costo eficiente que sería más rápido de analizar, más preciso y sería sensible a las variaciones del tráfico. La idea del proyecto era aumentar la cobertura de los hospitales analizando su ubicación geográfica y diferentes factores importantes como el tráfico en horas pico, locales cercanos como colegios, sitios de entretenimiento, estaciones de bomberos y otras variables socioeconómicas.

Durante la exploración de la información inicial se determinó que la cobertura de hospitales de alta complejidad para casos de accidentes era del 48,2% específicamente en horas de alto tráfico que representen un trayecto de 20 minutos. Por otro lado, se detectó que los hospitales estaban ubicados geográficamente con una tendencia a la cercanía a estratos socioeconómicos más altos. Esto significó que la cobertura para zonas de la ciudad en condiciones más precarias es del 24,6%.

Entre los resultados se encontró que solamente con ubicar un nuevo hospital de alta complejidad puede incrementar la cobertura hasta un 73%. De igual forma se encontró la mejor ubicación para establecer 2 nuevos hospitales de alta complejidad que llevarían la tasa de cobertura de la ciudad hasta un 85%, dándole así una oportunidad de acceso a una mayor parte de la población y una mayor tasa de supervivencia a los accidentes [12].

3.2.3. A multi-attribute data mining model for rule extraction and service operations benchmarking

En este estudio realizado en el 2022 por estudiantes de la Facultad de Administración en la Universidad de Teherán en Irán, los autores, buscaban comprender el patrón de comportamiento de los clientes en la industria bancaria y propusieron un enfoque híbrido de minería de datos con extracción de reglas y evaluación comparativa de la operación.

Para ello se analizaron los datos de los clientes identificando patrones y utilizando un modelo modificado de actualidad, frecuencia y dinero (RFM) y agrupamiento de K-Means. El número de conglomerados se determinó con un análisis de calidad de K-means de dos pasos basado en los índices de Silhouette, Davies-Bouldin y Calinski-Harabasz y la evaluación basada en la solución de distancia desde el promedio (EDAS). El método mejor-peor (BWM) y el área total basada en vectores ortogonales (TAOV) se utilizó para clasificar los conglomerados. Finalmente, se utilizó las reglas asociativas y el algoritmo A priori para derivar los patrones de comportamiento de los clientes.

Como resultado se segmentó a 20.000 clientes de una entidad financiera y las características demográficas y las transacciones financieras de los clientes se encontraron entre los factores analizados en este estudio. En este estudio se identificaron seis tipos de clientes y se propusieron estrategias adecuadas para interactuar con cada tipo con base a las opiniones de expertos bancarios. A demás se diseñaron estrategias adecuadas para aumentar la satisfacción de cada grupo de clientes en función de su valor y preferencias [13].

4. ENTENDIMIENTO DE LOS DATOS

La entidad financiera de estudio ha implementado un modelo de atención especializado para sus clientes empresariales y corporativos, conocido como Oficina de Servicios Especializados. Sin embargo, este modelo solo está disponible en una sucursal ubicada en el noroccidente de la ciudad de Cali, específicamente en el barrio Granada. Debido al alto número de clientes y transacciones requeridas por estos, esta única ubicación resulta insuficiente.

Por este motivo, este proyecto explora la posibilidad de habilitar un punto de atención remoto para clientes empresariales y corporativos dentro de las sucursales existentes destinadas a personas naturales. Este enfoque no supondría un costo o inversión significativa para la organización y permitiría ofrecer un modelo de atención exclusivo para clientes empresariales, sin afectar la atención a los demás clientes naturales. En este proyecto se llevó a cabo un análisis geoespacial utilizando diferentes tipos de variables de análisis, para determinar la zona óptima para la ubicación de un punto de atención remoto dentro de las sucursales de atención a personas naturales, empleando técnicas de ciencia de datos.

Se realizó un estudio basado en los datos recolectados de las transacciones realizadas por los clientes empresariales y corporativos durante el primer semestre de 2022 en la ciudad de Santiago de Cali. La Tabla 1 muestra las variables incluidas en la base de datos.

Tabla 1. Variables de base de datos

Variable	Tipo de Variable
Tipo de transacción	Categorica
Ciudad de transacción	Categorica
Ubicación de los clientes	Categorica
Monto de la transacción	Cuantitativa
Ubicación de puntos de servicio	Cuantitativa

Fuente: Elaboración propia

4.1. DESCRIPCIÓN DE VARIABLES

4.1.1. Tipo de transacción:

La variable tipo de transacción corresponde al tipo de acción que el cliente está realizando. Los datos

cuentan con 36 diferentes tipos de transacciones, siendo las siguientes 5 las de mayor volumen:

Tabla 2. Variables de base de datos

TIPO DE TRANSACCIÓN	DESCRIPCIÓN
RECAUDOS MASIVOS	Corresponde a la transacción donde el cliente realiza un pago a un tercero a través de un convenio de recaudo, presentando una factura o código de convenio junto con el pago en efectivo, cheque o débito a cuenta de ahorros o corriente.
PAGO CHEQUES	Corresponde a la transacción donde un cliente lleva un cheque del cual es titular y reclama el pago al que tiene derecho, siendo el cheque de la misma entidad financiera. Este pago puede ser en efectivo o mediante abono a cuenta de ahorros o corriente.
CONS CUENTAS CTES	Corresponde a la transacción donde un cliente realiza una consignación a su cuenta corriente mediante un cheque o dinero en efectivo.
DEPÓSITOS AHORROS	Esta transacción se trata de una consignación que el cliente realiza a su cuenta de ahorros mediante un cheque o dinero en efectivo.
ABONOS TARJETA CRÉDITO CAJA	Esta transacción corresponde a un pago dirigido a una obligación de tarjeta de crédito que se realiza en la sucursal física en la caja utilizando un medio de pago como cheque, efectivo o débito a cuenta de ahorros o corriente.

Fuente: Elaboración propia

Los anteriores 5 tipos de transacciones componen el 82% de todas las operaciones registradas en la base de datos que se utilizó para el estudio.

4.1.2. Ciudad de transacción

Corresponde a la ciudad donde se ubica la sede principal o domicilio del cliente. En la base de datos completa se tienen los clientes que son gerenciados por los ejecutivos de relación de la ciudad de Cali, por lo que algunos de ellos pueden ser clientes de municipios aledaños. Sin embargo, para el estudio realizado, únicamente se consideró aquellos que tenían su sede principal dentro del perímetro urbano.

4.1.3. Ubicación de los clientes:

La variable ‘Ubicación de los clientes’ se obtiene de la información que almacena la entidad como dirección principal de cada uno de dichos clientes y la cual se geolocalizó utilizando diferentes técnicas que más adelante se describen.

4.1.4. Monto de la transacción:

La variable ‘Monto de la transacción’ es de tipo cuantitativa y corresponde al valor que se ejecutó en cada una de las operaciones registradas en la base de datos. En las consignaciones corresponde al valor del dinero que se depositó en las cuentas o el pago realizado al convenio de recaudo. En retiros se muestra el valor en dinero en efectivo que los clientes retiraron en la caja. De igual forma en la base de datos completa se tienen algunas operaciones sin valor o monto cero, ya que corresponden a operaciones no financieras tales como ‘cambio de clave’ o ‘actualización de datos’.

A continuación, en la tabla 3 se muestra un breve resumen descriptivo de los montos de transacción:

Tabla 3. Características de Monto de transacciones

Métrica	Descripción	Valor (COP)
Media	Valor promedio de transacciones	\$13.759.559
Error típico	Dispersión de una muestra en relación con la media.	\$333.540
Mediana	Valor medio de las transacciones	\$2.400.000
Moda	Valor de la transacción más se presentó	\$68.439
Desviación estándar	Dispersión de los valores con respecto a la media.	\$62.057.284
Varianza de la muestra	Variabilidad de los valores con respecto a la media	3.8
Curtosis	Forma de la distribución	\$995.5266
Coficiente de asimetría	Simetría de la distribución	22.7
Rango	Diferencia entre el valor máximo y el valor mínimo	4.4
Mínimo	Valor más pequeño	0
Máximo	Valor más grande	4.4
Suma	Adición de todos los valores	\$476.314.663,6
Cuenta	Cantidad de transacciones realiza.	34.6

Fuente: Elaboración propia

4.1.5. Ubicación de puntos de servicio:

Esta variable corresponde a la ubicación en latitud y longitud de los puntos de servicio de la entidad financiera donde se ejecutaron las operaciones. Esta información se utilizó para georreferenciar la ubicación dentro del mapa de la ciudad y los volúmenes de operaciones ejecutadas en cada uno de los puntos de servicio.

4.2. BASE DE DATOS Y HERRAMIENTA UTILIZADA

La base de datos inicial tenía un total de 72,922 registros. Sin embargo, dado que el objetivo del proyecto era analizar las operaciones realizadas en las sucursales de servicio de la entidad financiera ubicadas en la Ciudad de Santiago de Cali durante el primer semestre del 2022, se depuró la base de datos eliminando los registros de operaciones realizadas fuera de la ciudad, así como aquellos registros en los que no se especificaba la dirección de los clientes. Al final, la base de datos utilizada para el análisis consistió en un total de 34,617 registros que representaban diferentes tipos de transacciones realizadas por 369 clientes en las 25 sucursales ubicadas en la ciudad.

Para llevar a cabo el análisis exploratorio de los datos y otros análisis en este proyecto, se empleó la herramienta de programación R-studio. Esta herramienta es un entorno de desarrollo integrado diseñado específicamente para el lenguaje de programación R, el cual se enfoca en el tratamiento de datos con fines estadísticos. En la tabla 4 se describen las librerías y funciones que se utilizaron dentro de esta herramienta:

Tabla 4. Librerías y funciones utilizadas

Librería	Descripción
tidygeocoder	Facilita la obtención de datos de los servicios de geocodificación.
tidyverse	Tiene como finalidad estructurar los datos.
Spatstat	Sirve para el análisis estadístico de datos espaciales. Su enfoque principal es el análisis de patrones espaciales de puntos en el espacio bidimensional.
Leaflet	Dispone de varios métodos y opciones para controlar el estado del mapa y configurarlo.
rasterVis	Métodos avanzados para visualización e interacción con datos ráster, encargado de ejecutar métodos de visualización para datos cuantitativos y categorías
Plotly	Crea gráficos interactivos con calidad de publicación.
Raster	Análisis y modelado de datos geográficos
Maps	Dibuja mapa geográfico

Sp	Clases y métodos para datos espaciales
ade4	Análisis de Datos Ecológicos: Métodos Exploratorios y Euclidianos en Ciencias Ambientales.
Función	Descripción
quadrat.tes	Permite determinar si existe asociación entre dos variables discretas.
Kest	Estima la función de segundo momento reducido a partir de un patrón de puntos en una ventana de forma arbitraria.
Over	Encuentra uniones espaciales o intersecciones.
Spplot	Métodos de trazado para datos espaciales con atributos
Density	Crea gráficos de densidad en base R o ggplot2
rasterFromXYZ	Crea un objeto Raster* a partir de valores x, y, z. En ráster: análisis y modelado de datos geográficos
scale	Es una técnica para comparar datos que no se miden de la misma manera.
mean	Cálculo del promedio o la media aritmética de un conjunto de valores normalmente definidos como un vector como parámetro
levelplot	Paquete que permite construir mapas de calor
ppp	Representa un conjunto de datos de patrón de puntos en el plano bidimensional.
eigen	Se usa para calcular valores propios y vectores propios de una array.
dudi.pca	Realiza un análisis de componentes principales de un marco de datos y devuelve los resultados como objetos de clase
inertia.dudi	Se utiliza para calcular la inercia de un análisis de correspondencia múltiple (ACM) bidimensional
cor	Muestra la similitud entre variables estudiadas
distanceFromPoints	La función calcula la distancia desde un conjunto de puntos hasta todas las celdas de un objeto Raster.

Fuente: Elaboración propia

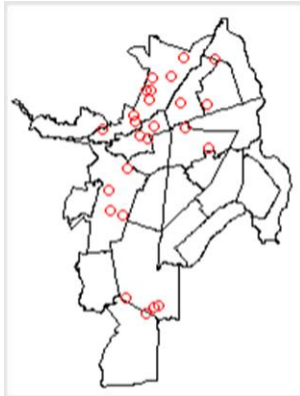
5. ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES

5.1 DISTRIBUCIÓN DE LOS DATOS

Para el análisis realizado en este proyecto se utilizó las variables: Ubicación de los clientes, Ubicación de puntos de servicio y monto de la transacción. Es importante resaltar que, para realizar el análisis geoespacial, estas variables debían estar definidas en formato de longitud y latitud. Sin embargo, solo la variable “Ubicación de puntos de servicio” contaba con esta definición desde la base de datos original suministrada por la entidad, por tal motivo para la variable “Ubicación de los clientes” que además de no contar con este formato, en algunos casos presentaban variaciones en la nomenclatura registrada, lo que implicó realizar previamente una estandarización a la nomenclatura urbana definida acorde al municipio de Cali. Posteriormente se realizó el proceso de geo-codificación para así obtener la latitud y longitud de las ubicaciones de cada cliente. Este proceso de geocodificación se llevó a cabo utilizando las librerías “*tidygeocoder*” y “*tidyverse*” y el sistema de información geográfico denominado “*arcgis*”.

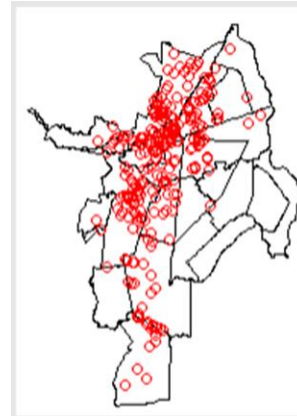
Una vez se obtuvo las latitudes y longitudes de las variables, se procedió a ubicar cada uno de los puntos en el mapa de Cali. Para realizar este análisis, primero se debió cambiar el formato de los datos que se encontraban en latitud y longitud al formato Planar Point Patterns – PPP para crear un patrón de puntos a partir de coordenadas (x, y). Este formato es utilizado para identificar y examinar la estructura y características de los puntos en relación con su distribución espacial para comprender la relación entre estos.

Ilustración 1. PPP – Puntos de servicio



Fuente: Elaboración propia

Ilustración 2. PPP – Ubicaciones de clientes



Fuente: Elaboración propia

Para determinar la distribución de los puntos (aleatoria, regular o agregado) dentro de la región espacial de Cali, se utilizó la librería “*spatstat*”, también se realizó la estimación de los estadísticos de resumen de segundo orden de acuerdo con la función “*quadrat.test*” de la prueba Chi2, que permitió testear el comportamiento de los datos de las variables de interés.

Pariendo de la hipótesis de aleatoriedad en los datos, la prueba Chi2, determinó un P-value menor a $2.2e-16$ que es inferior al nivel de significancia establecido, comúnmente 0.05, esto indica que los datos proporcionan evidencia suficiente para rechazar la hipótesis nula de aleatoriedad en los puntos de servicios.

Ilustración 3. Prueba Chi2 – Puntos de servicio

```
Chi-squared test of CSR using quadrat counts  
data: patron_punt_servicio  
X2 = 153256, df = 24, p-value < 2.2e-16  
alternative hypothesis: two.sided  
  
Quadrats: 5 by 5 grid of tiles  
  
Chi-squared test of CSR using quadrat counts
```

Fuente: Herramienta R

En el caso de la ubicación de los clientes, de acuerdo con la prueba Chi2 también se determina que se rechaza la hipótesis de aleatoriedad con un P-value menor a 2.2e-16.

Ilustración 4. Prueba Chi2 – Ubicación Clientes

```
Chi-squared test of CSR using quadrat counts

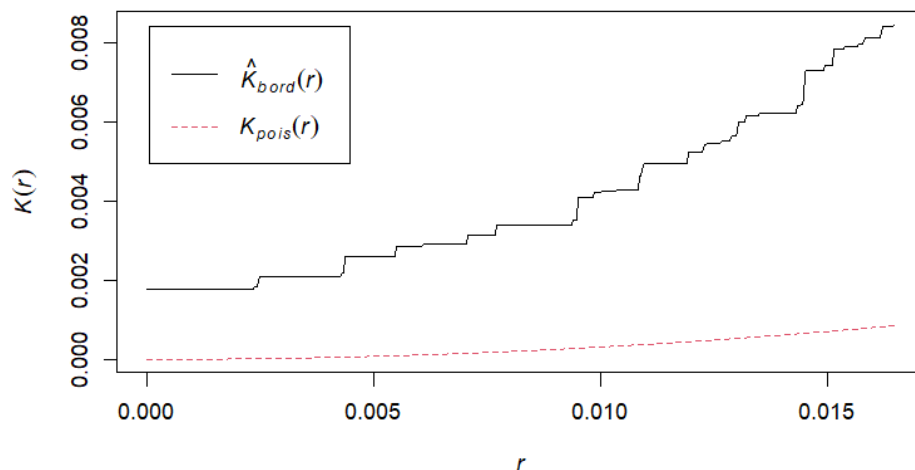
data: patron_clie
x2 = 1315.5, df = 24, p-value < 2.2e-16
alternative hypothesis: two.sided

Quadrats: 5 by 5 grid of tiles
```

Fuente: Herramienta R

La función Kest(r) determina el patrón de distribución espacial y calcula la función K de Ripley corregida por la intensidad, que define la proporción de puntos dentro de un radio r de distancia promedio de cualquier otro punto en el conjunto de datos y como se observa en la ilustración 5, la distribución de los datos para las ubicaciones de los puntos de servicio en total 25, tienen una distribución agregada, ya que la función $K_{bord}(r)$ se encuentra por encima de la función del modelo $K_{pois}(r)$ la cual tiene un patrón aleatorio con una probabilidad de ocurrencia media. Además, se observa que a distancias pequeñas se presenta una alta frecuencia de los datos, lo que implica que los puntos de servicio a nivel espacial se agregan por sectores, es decir que existen nichos donde se encuentran más cercanos entre ellos.

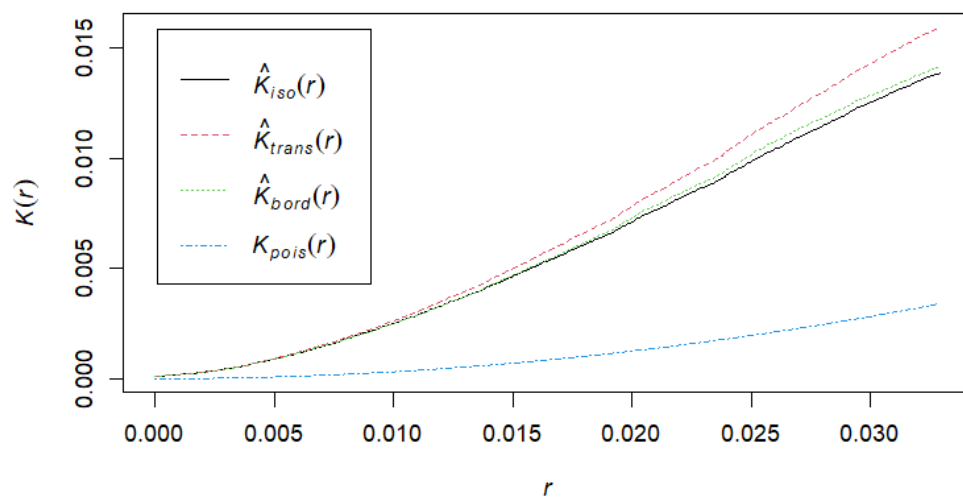
Ilustración 5. Función de distribución Kest – Puntos de servicio



Fuente: Elaboración propia

Al igual que para los puntos de servicio que se distribuyen de forma agregada, en las ubicaciones de los clientes también se observan una distribución agregada de acuerdo con la función aleatoria $K_{pois}(r)$, la cual se ubica por debajo de la función $K_{bord}(r)$ lo cual indica que indica un patrón agregado. Las funciones $K_{iso}(r)$ y $K_{trans}(r)$, son correcciones o transformaciones que se presentan porque existe una mayor cantidad de datos que corresponde a las ubicaciones de los 369 clientes en comparación con las ubicaciones de los 25 puntos de servicio.

Ilustración 6. Función de distribución Kest - Ubicación de los clientes

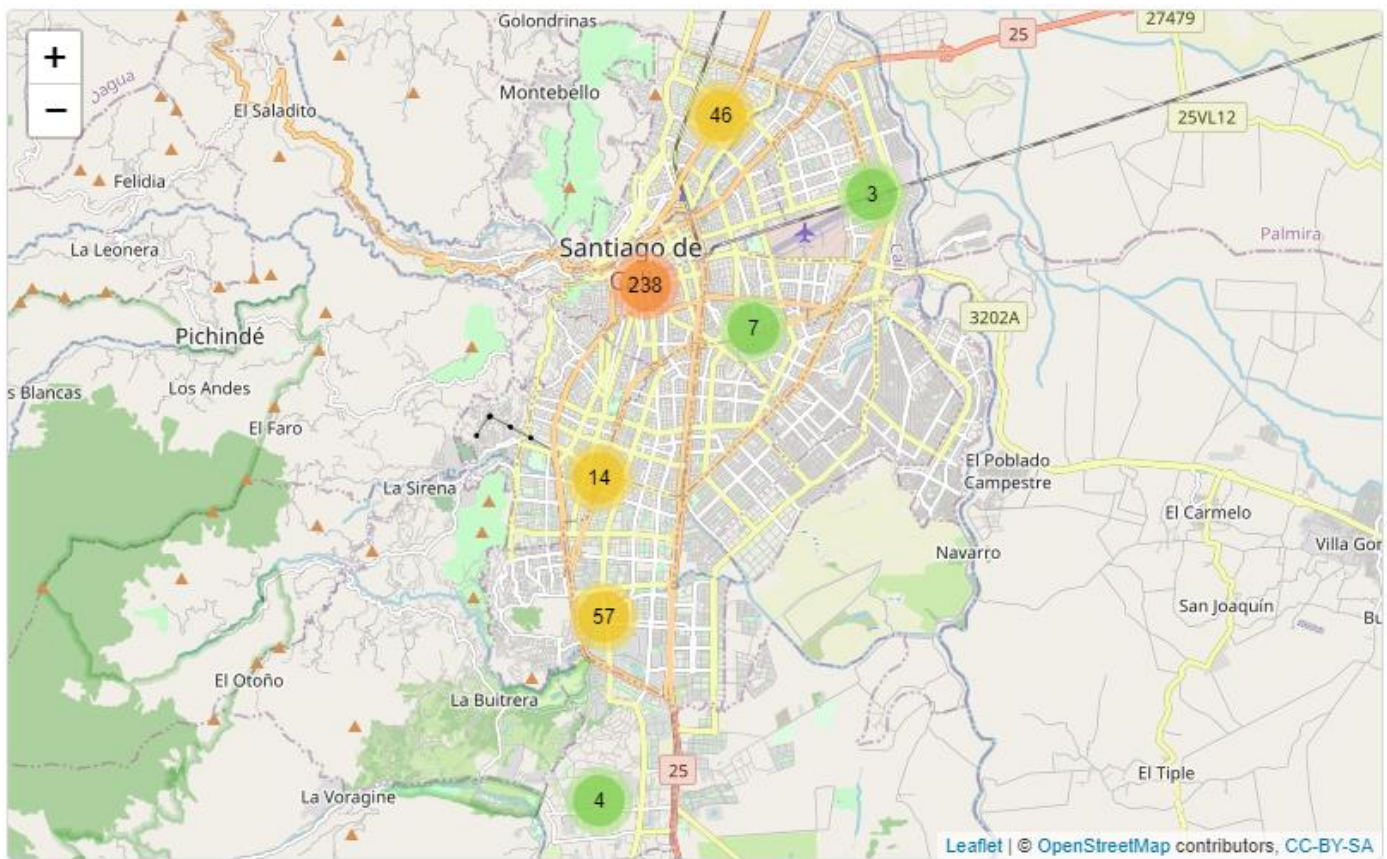


Fuente: Elaboración propia

5.2 CONTABILIZACIÓN MEDIANTE MAPAS INTERACTIVOS

Para visualizar los datos a través de mapas interactivos que utilizan por defecto la herramienta de OpenStreetMap, se hace uso de la librería “leaflet”, que permite generar Polígonos de Thiessen, los cuales muestran las concentraciones de los datos por sectores y los proyecta sobre el mapa de Cali. A continuación, en la ilustración 7, se observa que, para las ubicaciones de los clientes, la mayor concentración de clientes representada en color naranja, se encuentra en la zona centro occidental de Cali, con un total de 238 clientes en ese sector, seguido de los representados de color amarillo que se encuentran en la zona norte oeste con 48 clientes, en la zona sur centro con 14 y 57 clientes y finalmente los representados por los círculos de color verde en menor cantidad ubicándose 3 en el norte oriente, 7 en la parte centro oriente y 4 en la zona sur casi a las afueras de la ciudad.

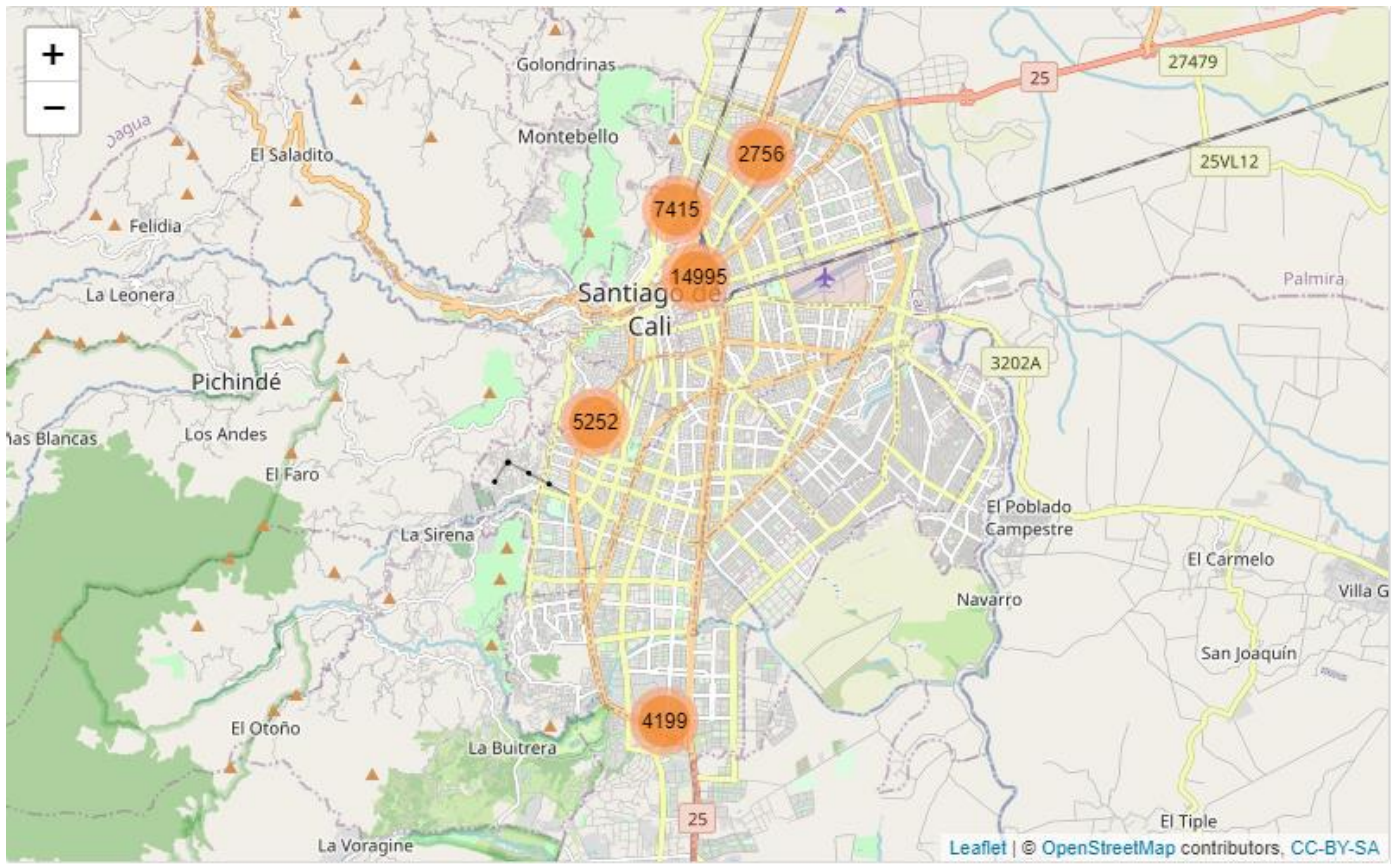
Ilustración 7. Clusters Ubicación de Clientes



Fuente: Elaboración propia

Por otro lado, para el caso de los puntos de servicio, al contabilizar el número de transacciones realizadas en cada uno de estos, se observa que el mayor número de transacciones se hicieron cerca de la zona en donde se encuentran ubicados la mayor parte de los clientes. En esta zona se muestran por círculos de color naranja debido a que representan un número significativo de eventos. En la ilustración 8 se observa que se realizaron 14.995 transacciones en la zona centro oeste de un total de 34.617, siendo esta la zona de mayor número, seguido de las 7415 ubicadas en la parte norte oeste y muy cercana a los 2756 que se ubican más al norte centro. En el centro sur se observa 5252 transacciones y en la zona sur, casi a las afueras de la ciudad se realizaron 4199 transacciones.

Ilustración 8. Clusters Transacciones



Fuente: Elaboración propia

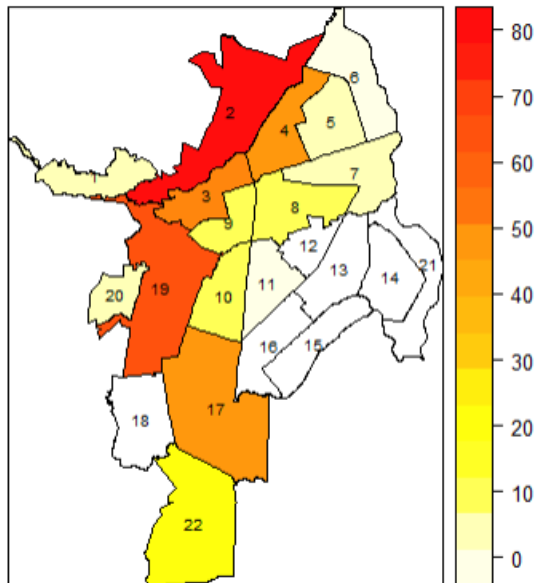
5.3 CONTEO TRANSACCIONES Y NUMERO DE CLIENTES POR COMUNAS

Para visualizar el conteo por comunas se hace uso de la librería “rasterVis” y la función “spplot” y “over”. Se realiza conteo por comunas de la cantidad de clientes existentes en cada una según su ubicación. En la Ilustración 9, se observa que en las comunas 2 y 19 se concentra el mayor número de clientes con un rango de 50-80 clientes. Para el caso de la comuna 19 se encuentra ubicada en el oeste de la ciudad y se caracteriza por ser un sector de clases socio económicas media y alta, mientras que la comuna 2 está ubicada en el noroeste del casco urbano y se caracteriza por ser una zona financiera importante por su alta infraestructura de hoteles, centros educativos, sitios comerciales, de servicios, de entretenimiento, parques, etc.

Por otro lado, en la Ilustración 10 se observa que, al realizar el conteo de transacciones en cada uno de los

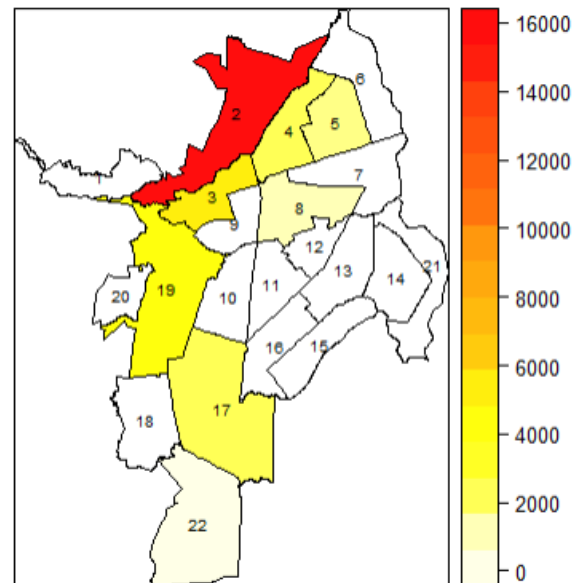
puntos de servicio, la mayoría se realizaron en la comuna 2, con un rango de 14.000 a 16.000 transacciones, lo cual tiene sentido debido a que en esta comuna se encuentra ubicada la Oficina de Servicios Especializados.

Ilustración 9. *Conteo por comunas – Ubicaciones de clientes*



Fuente: Elaboración propia

Ilustración 10. *Conteo por comunas – Transacciones*



Fuente: Elaboración propia

El análisis exploratorio de datos reveló tendencias y patrones en la distribución espacial de los puntos de servicio y la ubicación de los clientes en el mapa de la ciudad de Santiago de Cali. Además, se examinó la cantidad de transacciones realizadas en dichos puntos de servicio para identificar las áreas con mayor concentración de transacciones. Asimismo, se analizó la concentración de clientes en las distintas comunas de la ciudad. En el siguiente capítulo, se profundizará en imágenes de densidad de las variables estudiadas hasta el momento y se integraran a otras importantes para el desarrollo del proyecto como ubicación de vías principales, uso del suelo y la distancias al punto de la Oficina de Servicios Especializados.

6. IMÁGENES DE DENSIDAD Y ZONAS ÓPTIMAS

6.1. VARIABLES DE ANÁLISIS

En este punto se realizó un análisis multicriterio y para ello, además de las variables vistas en el capítulo 5, se utilizaron variables que ayudarían a determinar el posible punto de atención para clientes empresariales. En la tabla 3 se muestran las variables utilizadas en el análisis multicriterio.

Tabla 5. Variables analizadas

Variable	Definición
Cantidad de transacciones	Contiene el número de transacciones realizadas en cada punto de servicio
Ubicación de los clientes	Ubicación espacial de las oficinas registradas por los clientes
Vías principales y secundarias	Ubicación de vías principales y secundarias en la ciudad de Santiago de Cali
Uso del Suelo	Ubicación de uso de suelo comercial en la ciudad de Santiago de Cali

Fuente: Elaboración propia

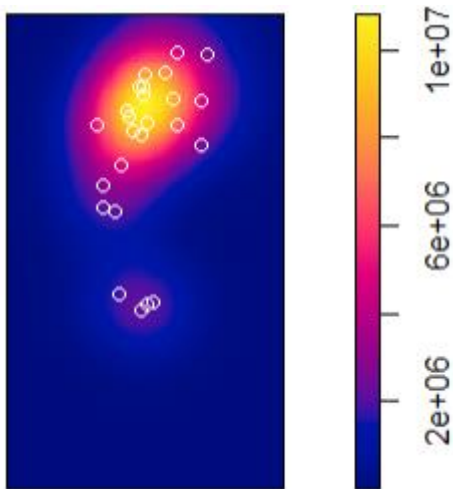
6.2. IMÁGENES DE DENSIDAD

Las variables de interés, se transformaron en formato de imágenes de densidad que permiten visualizar la distribución espacial de los puntos de variables analizadas. Esto hace que, al asignar valores a cada celda o píxel en un formato de rejilla, se puede representar de manera efectiva la información geoespacial según sea el mayor o menor grado de concentración de puntos. Para generar las imágenes de densidad se utilizaron las librerías “plotly”, “raster” y “rastervis”, así como también la función de intensidad “density”. En las ilustraciones 10, 11, 12 y 13 se observan las imágenes de densidades generadas. En estas imágenes se muestra las zonas donde se presenta mayor densidad o agrupación espacial de los puntos de las variables en estudio.

Las imágenes de densidad, posteriormente se integrarán en un solo análisis. Esto permitió examinar las relaciones y patrones espaciales entre las variables, lo que puede conducir a la identificación de

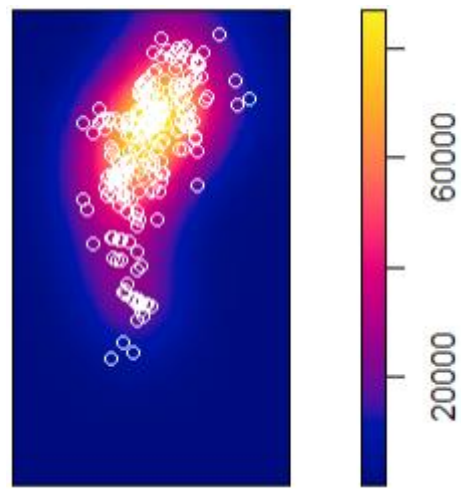
interacciones complejas e identificación de las zonas optimas.

Ilustración 10. Cantidad de Transacciones



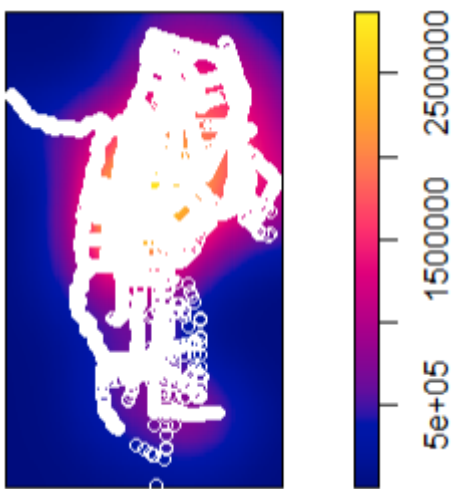
Fuente: Elaboración propia

Ilustración 11. Ubicación de los clientes



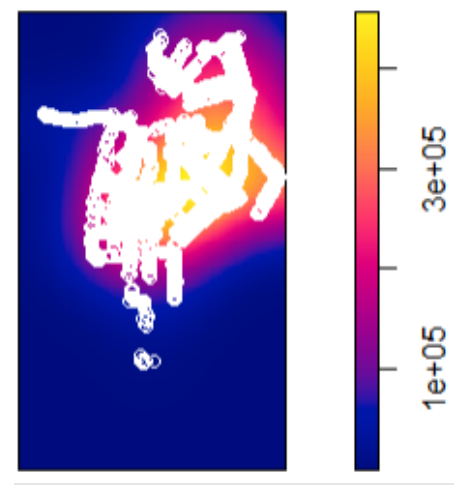
Fuente: Elaboración propia

Ilustración 12. Vías



Fuente: Elaboración propia

Ilustración 13. Uso de Suelo



Fuente: Elaboración propia

6.3. IMÁGENES RASTER DE VARIABLES DE ESTUDIO

Posteriormente las imágenes de densidad se convirtieron a un tipo de imágenes “raster” que transforma todas las variables a tipo numéricas, ya que se discretiza las variables continuas en una cuadrícula regular de píxeles, donde cada píxel representa un valor numérico. Cada píxel tiene asignado un valor que indica la intensidad en ese punto específico de la imagen. Además, este tipo de imágenes permiten operar con otro tipo de información como la cartografía de Cali. Las imágenes raster se generaron a partir de las imágenes de densidad generadas previamente y utilizando la función “rasterFromXYZ”.

En el análisis, se incluyó una variable llamada "Distancia a la Oficina de Servicios Especializados". Esta variable se calculó teniendo en cuenta la ubicación de la Oficina de Atención Empresarial, la cual se encuentra ubicada en las coordenadas de longitud y latitud (-76.53574, 3.46048). Se consideró importante tener en cuenta esta distancia, debido a que se buscaba que el posible punto de atención no quedara cerca de la Oficina de Atención Empresarial.

Para calcular la variable de distancia se utilizó la función "distanceFromPoints", las coordenadas de la Oficina de Atención Principal se utilizaron como punto de referencia para calcular las distancias desde cada punto en la imagen raster "Transacciones" a la Oficina de Atención Principal. La información detallada de la variable generada se encuentra en la tabla 6.

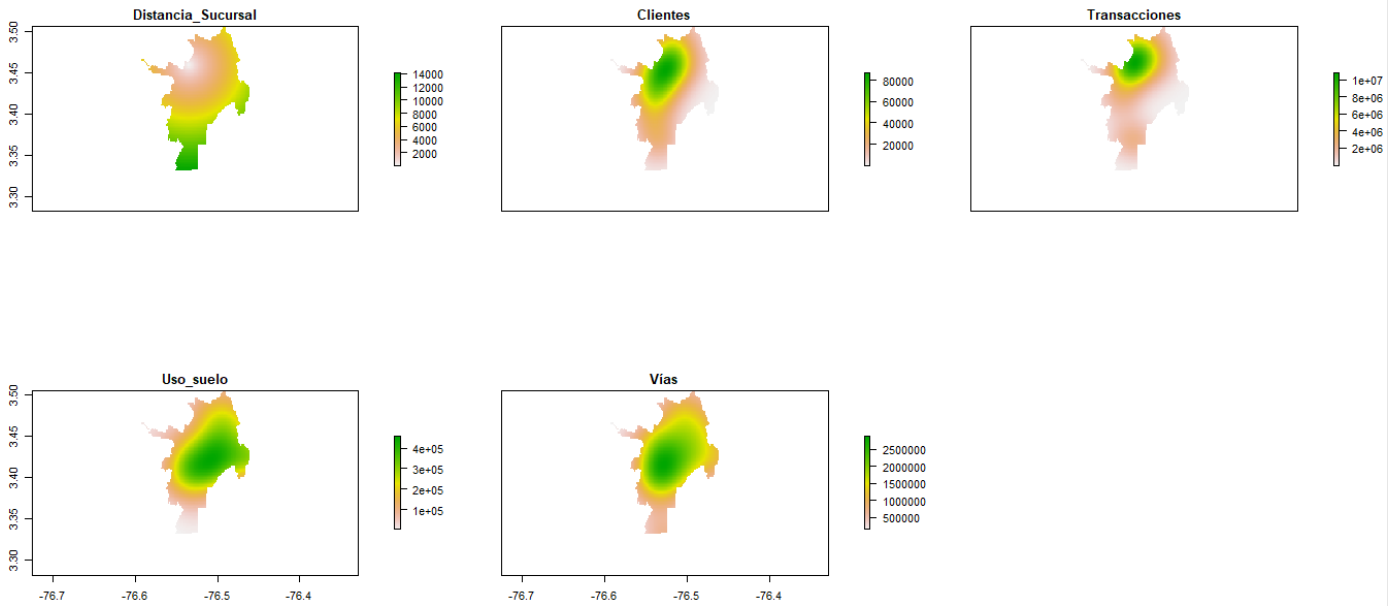
Tabla 6. Variable distancia Oficina de servicio especializado

Variable	Definición
Distancia a la Oficina de Servicios Especializados	Corresponde a las distancias existentes entre las ubicaciones de los clientes y la Oficina de Servicios Especializados

Fuente: Elaboración propia

Al generar las imágenes raster de cada variable, estas presentan una escala diferente, lo cual significa que los valores numéricos asociados a los píxeles representan diferentes rangos de magnitudes, por lo tanto, se deben estandarizar para que se encuentren en un rango comparable. Al estandarizar las imágenes raster, se eliminan las diferencias de escala y se logra una comparabilidad más adecuada, lo que facilita el análisis. En la ilustración 14, se muestra las imágenes raster de cada una de las variables estudiadas.

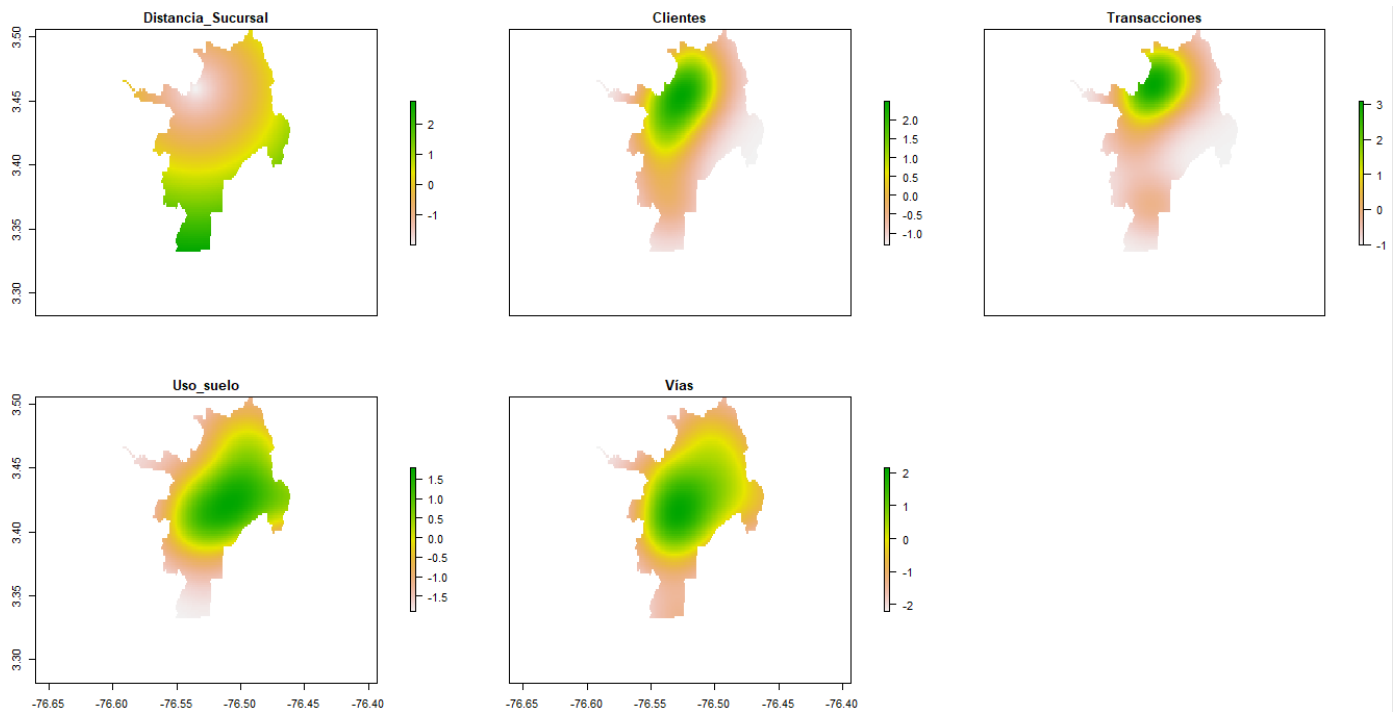
Ilustración 14. Imagen raster de las variables



Fuente: Elaboración propia

Para estandarizar las imágenes se utilizó la función “scale” que calcula la media y la desviación estándar de los valores de píxeles en la imagen original y posteriormente resta la media de cada valor de píxel, lo que hace que se desplace la distribución de los valores de píxeles para que su media sea cero y divide cada valor de píxel en la imagen resultante por la desviación estándar. En la ilustración 15 se presenta las imágenes raster escaladas.

Ilustración 15. Imagen raster de las variables escaladas



Fuente: Elaboración propia

Una vez que todas las imágenes están en la misma escala, se pueden utilizar diversos métodos para realizar comparaciones y determinar la mejor combinación, lo cual permite tomar decisiones basadas en esas comparaciones. Para el caso de este proyecto, esto sirve para realizar un análisis multicriterio variable y generar indicadores a partir de diferentes escenarios propuestos que delimiten zonas óptimas para la ubicación del posible punto de atención para clientes empresariales.

6.4. ESCENARIOS PROPUESTOS Y CÁLCULO DE INDICADORES

Para generar las zonas óptimas a través de la integración de las 5 variables de estudio, se plantearon 4 escenarios que resultaron en 4 indicadores diferentes. En cada escenario, se propone asignar pesos de ponderación a las variables de análisis según criterios distintos. Para el primer escenario, se asignó una ponderación uniforme a todas las imágenes raster, lo que implica que todas las variables tuvieran la misma importancia en la generación del indicador 1.

En el segundo escenario, se asignan las ponderaciones a las imágenes raster en función de su relevancia en el análisis. En este caso, se consideró que la variable más relevante era la que más distante al punto de

servicio de la Oficina de Servicios Especializados, ya que se dio mayor peso a aquellas que se encuentren ubicaciones alejadas de la Oficina de Servicios Especializados.

En el tercer escenario, se realizó una asignación de ponderaciones basada en el juicio de expertos, los cuales son colaboradores de la entidad financiera. Estos expertos evaluaron la relevancia y contribución de cada variable en función de sus necesidades y experiencia.

Finalmente, en el cuarto escenario, se propone asignar las ponderaciones utilizando la técnica de análisis de componentes principales (ACP). Esta técnica asigna mayor importancia a las variables en función de la variabilidad presente en el conjunto de datos. En la tabla 5, se muestra lo anterior.

Tabla 7. Escenarios de análisis de indicadores

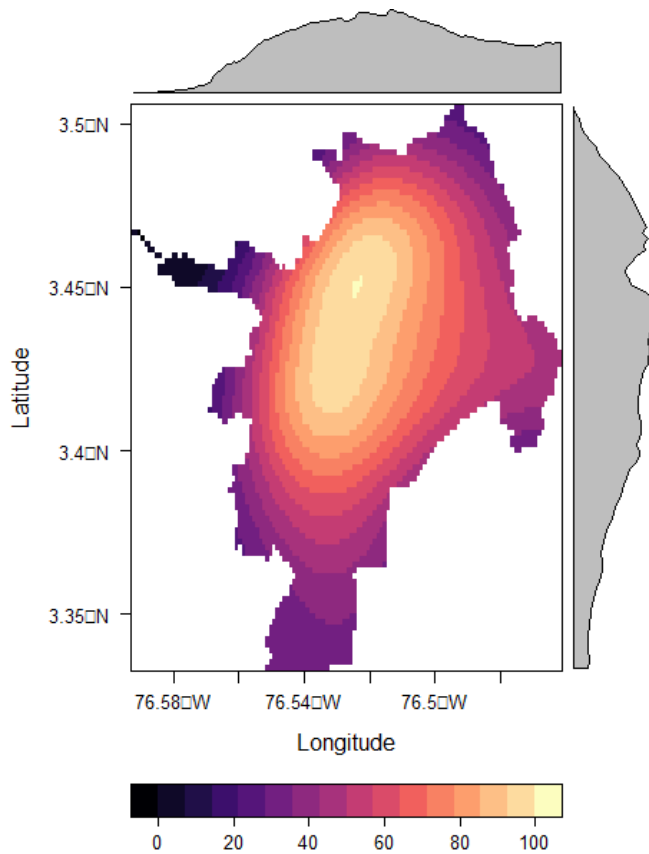
Indicadores	Distante a la Oficina de Servicios Especializados	Ubicación de los clientes	No de transacciones	Uso del Suelo	Vías principales y secundarias
Indicador 1	20%	20%	20%	20%	20%
Indicador 2	40%	15%	15%	15%	15%
Indicador 3	21%	20%	17%	19%	23%
Indicador 4	28.9%	28.7%	21.8%	6.3%	14.2%

Fuente: Elaboración propia

6.4.1. Indicador 1 – Ponderación de acuerdo a la media

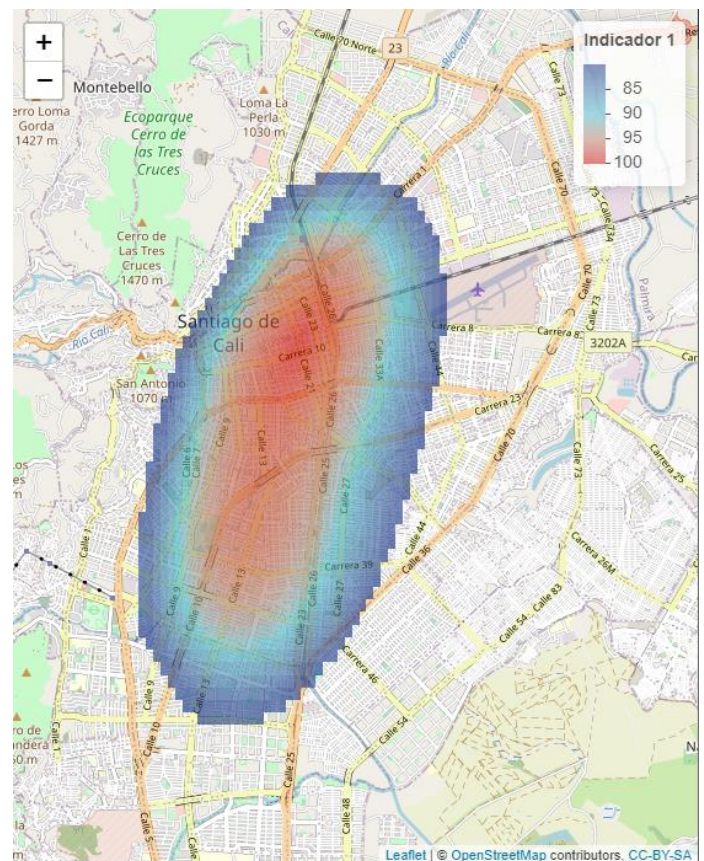
La asignación de importancia de cada variable para este indicador, se realizó de acuerdo con la media ponderada de los valores de píxeles de las imágenes de cada variable en su forma escalada, utilizando la función “mean”. Al trabajar sobre la media, se está asumiendo de que todas las variables tienen el mismo valor de ponderación, es decir se les da igual importancia a todas las variables. En las ilustraciones 16 y 17 se aprecia una mayor densidad por unidad de área en la zona centro-norte de la ciudad de Santiago de Cali, específicamente a lo largo de la carrera 10 con calle 21, abarcando las comunas 3 y 9. En esta zona particular, se encuentran ubicadas tres sedes importantes: Cali Principal (Of. 60), San Nicolás Cali (Of. 62) y Calle Catorce (Of. 803). Históricamente, esta área se destaca por albergar sectores industriales de la ciudad de Santiago de Cali.

Ilustración 16. Imagen estática indicador 1



Fuente: Elaboración propia

Ilustración 17. Imagen dinámica indicador 1

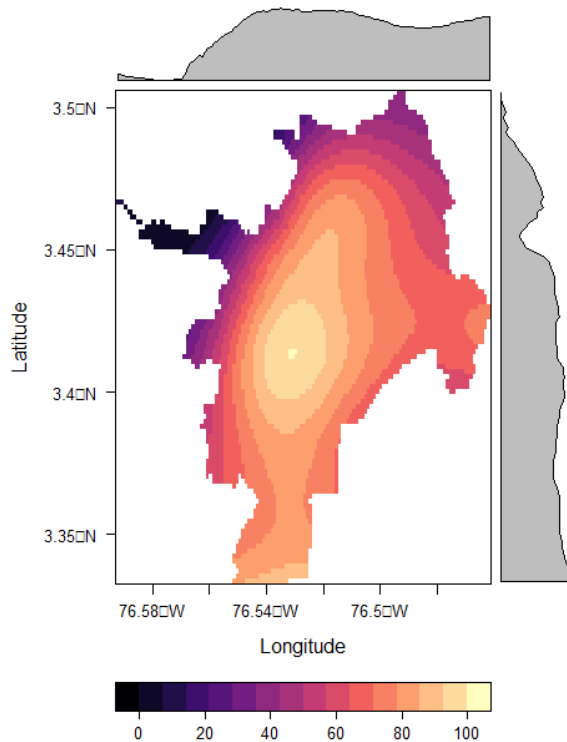


Fuente: Elaboración propia

6.4.2. Indicador 2 – Mayor prioridad a las distancias alejadas a la oficina de servicios especializados

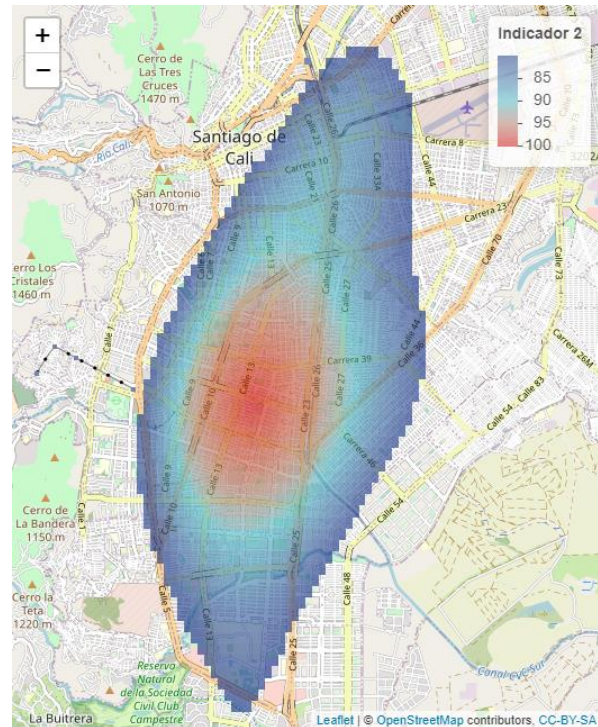
Para el indicador 2, se desarrolló el segundo escenario en el cual se dio prioridad a la variable " Distancia a la Oficina de Servicios Especializados ". En este caso, se buscó asignar un mayor peso de ponderación a los píxeles que se encuentran más alejados de la Oficina de Servicios Especializados, con el objetivo de evitar la proximidad del punto de atención propuesta con el ya existente. Sin embargo, cabe destacar que se tuvo en cuenta la importancia de las demás variables en el análisis, asegurándose de no descuidar su relevancia para la toma de decisiones. En la ilustración 18 y 19, se observa que la zona sugerida cambia completamente el panorama del indicador 1, ya que se ubica en entre la calle 13 con carrera 44. En esta zona en particular el punto de servicio más cercano se encuentra ubicado en el Centro Comercial Palmetto Plaza (Of. 808).

Ilustración 18. Imagen estática indicador 2



Fuente: Elaboración propia

Ilustración 19. Imagen dinámica indicador 2



Fuente: Elaboración propia

6.4.3. Indicador 3 – Juicio de expertos

Para generar este indicador, se llevó a cabo una consulta mediante una encuesta dirigida a diversos colaboradores de la entidad financiera, incluyendo tanto directivos como personal operativo. Entre los encuestados se encontraron colaboradores con experiencia previa o actual en el cargo de jefe de Centro, el cual es el responsable de las oficinas de servicios empresariales.

El objetivo principal de la encuesta fue determinar la importancia relativa de cada variable y obtener la opinión de los participantes acerca de la calificación o puntaje que debería asignarse a cada una de ellas al momento de habilitar un punto de servicio como oficina de servicios empresariales. La escala de calificación utilizada en la encuesta fue del 1 al 4, donde 1 representó el puntaje más bajo y 4 el más alto.

Con esta encuesta, se buscó recopilar la información necesaria para establecer una base sólida en la generación del indicador que permita identificar la ubicación óptima para una oficina de servicios empresariales en función de las variables consideradas.

Tras analizar los datos recopilados, se obtuvieron los siguientes resultados

Tabla 8. Resultado de la encuesta

ID	Distante de la Oficina de Servicios Especializados	Ubicación de los clientes	No de transacciones	Uso del Suelo	Vías principales y secundarias	Tota
1	1	2	4	3	4	14
2	1	4	4	4	3	16
3	2	4	3	1	4	14
4	1	3	2	4	4	14
5	4	4	4	4	1	17
6	4	1	3	4	2	14
7	3	3	4	4	4	18
8	4	4	1	2	3	14
9	3	2	4	4	3	16
10	3	1	1	4	4	13
11	3	4	1	4	2	14
12	3	4	3	4	3	17
13	2	3	4	4	4	17

Fuente: Elaboración propia

Las calificaciones obtenidas de los encuestados revelan diversas opiniones en relación a este tema, sin embargo, la variable que recibió la puntuación más alta fue la "Cercanía a vías principales y secundarias".

Para determinar las ponderaciones de cada una de las variables, se llevó a cabo una suma de los puntos obtenidos en cada una de ellas. Posteriormente, se calculó el puntaje total combinando todas las variables y se obtuvieron las siguientes ponderaciones

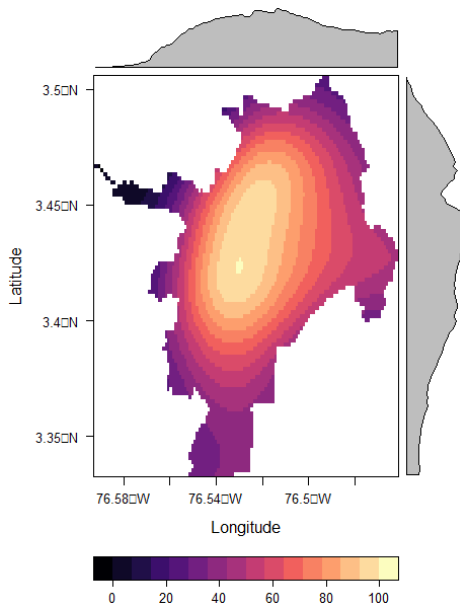
Tabla 9. Porcentajes de resultado de encuesta

ID	Distante de la Oficina de Servicios Especializados	Ubicación de los clientes	No de transacciones	Uso del Suelo	Vías principales y secundarias	Total
Suma de puntajes	41	39	34	38	46	198
Porcentaje	21%	20%	17%	19%	23%	100%

Fuente: Elaboración propia

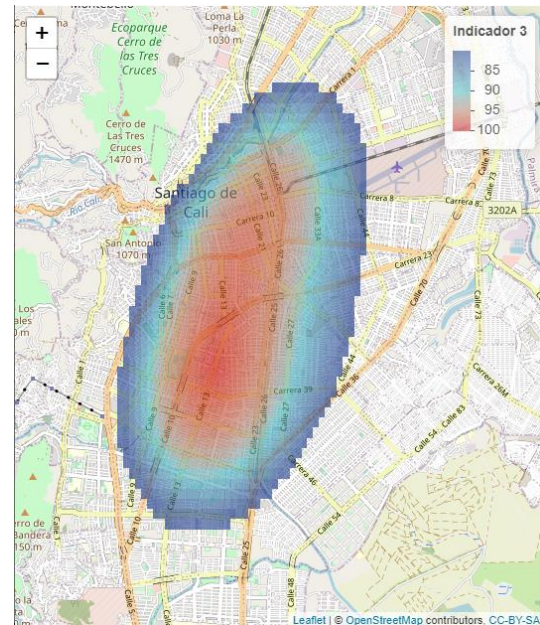
En las ilustraciones 20 y 21, se puede apreciar que la zona óptima demarcada se ubica entre la calle 13 y la carrera 32. Dentro de esta área en particular, el punto de servicio más cercano se encuentra en el barrio El Cedro (Of. 750).

Ilustración 20. Imagen estática indicador 3



Fuente: Elaboración propia

Ilustración 21. Imagen dinámica indicador 3



Fuente: Elaboración propia

6.4.4. Indicador 4 – Análisis de componentes principales

En el indicador 4, se emplea la metodología de Análisis de Componentes Principales (ACP) utilizando la librería "ade4". Para llevar a cabo este análisis, primero se examina la correlación entre las variables utilizando la función "cor". Esto permite comprender las relaciones existentes entre las variables y evaluar su grado de dependencia lineal. En la tabla 10, se observa una alta correlación entre la variable "Cantidad de transacciones" y "Ubicación de los clientes" con un 89%, y entre "Vías principales-secundarias" y "uso del suelo" con un 86%. Es crucial entender esta correlación entre variables, ya que el objetivo del análisis de componentes principales (ACP) es descomponer dicha matriz en valores y vectores propios para determinar la importancia de cada variable y la importancia u ponderación asociados a ellas.

Tabla 10. Matriz de Correlaciones de Variables

Características	Distancia de la Oficina de Servicios Especializados	Ubicación de los clientes	No de transacciones	Uso del Suelo	Vías principales y secundarias
Distante de la Oficina de Servicios Especializados	1	-0.76	-0.72	-0.33	-0.43
Ubicación de los clientes	-0.76	1	0.89	0.05	0.39
No de transacciones	-0.72	0.89	1	-0.14	0.07
Uso del Suelo	-0.33	0.05	-0.14	1	0.86
Vías principales y secundarias	-0.43	0.39	0.074	0.86	1

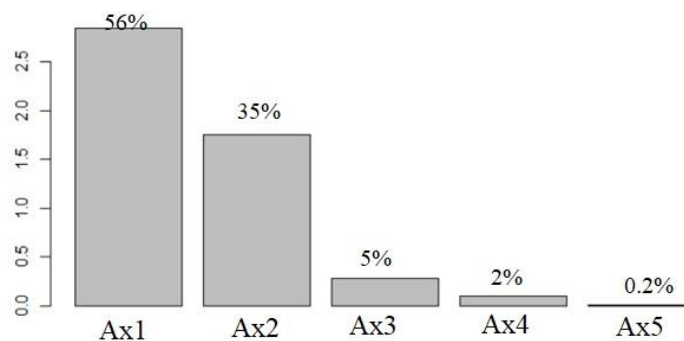
Fuente: Elaboración propia

A partir de la matriz de correlación, el ACP realiza una transformación lineal de los datos, generando nuevas variables llamadas componentes principales.

Las componentes principales se construyen de manera que la primera componente captura la mayor variabilidad posible en los datos, la segunda componente captura la segunda mayor variabilidad, y así sucesivamente. Cada componente principal es una combinación lineal de las variables originales, y se obtienen de forma que sean ortogonales entre sí. Estas componentes se ordenan por importancia, de manera que la primera componente principal explica la mayor parte de la variabilidad total en los datos.

En la ilustración 22 se observa que el primer componente Ax1 recoge el 56%, por encima de los otros componentes, Ax2 recoge el 35%, el Ax3 el 5%, el Ax4 el 2% y el Ax5 el 0.2% de la información. Para conocer el porcentaje de cada componente se utilizó la función “dudi.pca”.

Ilustración 22. Distribución gráfica de componentes principales – Indicador 4



Fuente: Elaboración propia

El resultado del análisis de componentes principales (ACP) consiste en valores numéricos llamados valores propios, los cuales están asociados a cada componente principal. Estos valores propios indican la cantidad de variabilidad explicada por cada componente principal. Además, es posible calcular el porcentaje de variabilidad explicada para cada componente, lo cual refleja la proporción de la variabilidad total de los datos capturada por cada componente. Para generar la descomposición espectral de los vectores propios, se utilizó la función "inertia.dudi".

En la tabla 11, se observa el cálculo de la inercia que es una medida de variabilidad total de los datos en un análisis de componentes principales y que se utiliza para evaluar la calidad del modelo ACP. En general, se busca retener la mayor cantidad de inercia posible con el menor número de componentes principales, para obtener una representación más compacta y eficiente de los datos.

Tabla 11. Componentes principales con la mayor parte de la variación en los datos

Característica	Inercia
Ax1	2.84
Ax2	1.75
Ax3	0.28
Ax4	0.10
Ax5	0.01

Fuente: Elaboración propia

El componente más representativo en el análisis es el Axis1, el cual tiene una inercia de 2.84. Este componente resume de manera significativa el análisis realizado. A partir de los vectores propios asociados a este componente, se procede a crear el indicador 4. La técnica sugiere ciertas ponderaciones para calcular dicho indicador, las cuales se detallan en la tabla 12.

Tabla 12. Porcentajes de prioridad componentes principales – Indicador 4

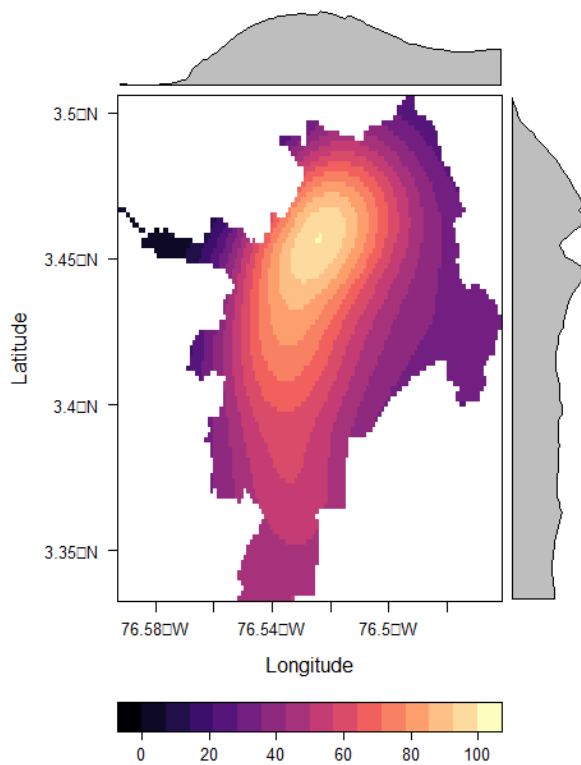
Características	Axis 1 (%)
Distante de la Oficina de Servicios Especializados	28.92
Ubicación de los clientes	28.71
No de transacciones	21.83
Uso del Suelo	6.31

Características	Axis 1 (%)
Vías principales y secundarias	14.21

Fuente: Elaboración propia

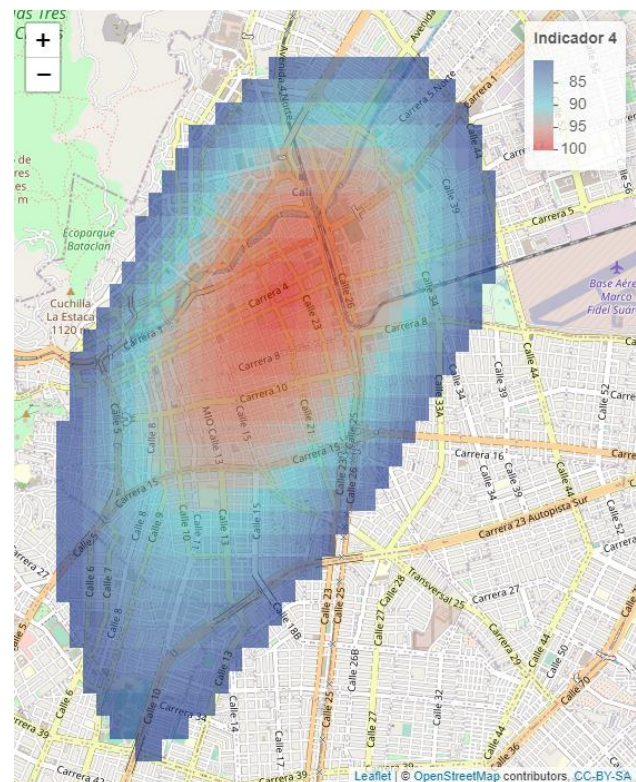
El resultado del indicador 4 indica que la zona delimitada como óptima se encuentra ubicada entre carrera 4 con calle 23 en el barrio San Nicolas. En esta zona específica el punto de servicio ubicado es el de “San Nicolas Cali” (Of. 62), que como se mencionó previamente en el resultado del indicador 1, es un área que se destaca por albergar sectores industriales de la ciudad de Santiago de Cali.

Ilustración 23. Imagen estática indicador 4 - ACP



Fuente: Elaboración propia

Ilustración 24. Imagen dinámica indicador 4 - ACP



Fuente: Elaboración propia

Luego de implementar los cuatro indicadores en cada uno de los escenarios, se llevó a cabo una comparación para determinar cuál de ellos era el mejor para delimitar la zona óptima. La evaluación de los indicadores se realizó mediante análisis y consideraciones basadas en los resultados obtenidos en cada escenario. En el capítulo 7 se mostrará el comparativo de acuerdo a similitud de las zonas obtenidas con cada uno de los indicadores.

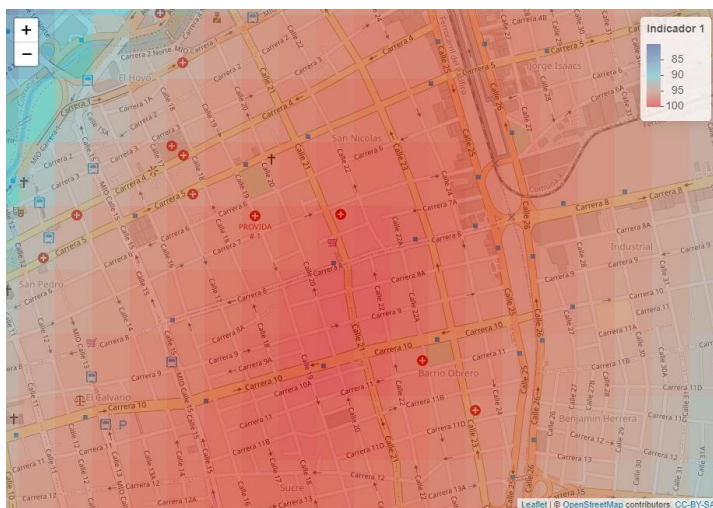
7. ANÁLISIS COMPARATIVO DE INDICADORES Y SELECCIÓN DE LA MEJOR OPCIÓN PARA LA DELIMITACIÓN DE LA ZONA ÓPTIMA

Según la generación de indicadores analizada en el capítulo anterior, se descubrió que los indicadores 1 y 4 delimitan como zona óptima el barrio San Nicolás, donde se encuentra ubicado el punto de servicio de San Nicolás Cali (Of. 62). Esto se debe a que las ponderaciones asignadas no muestran cambios significativos entre ellos. Sin embargo, al observar los indicadores 2 y 3, se identificó que la zona priorizada difiere considerablemente, por lo que no se profundizará en detalle sobre estas áreas.

7.1. INDICADOR 1 y 4 – Punto de servicio San Nicolas Cali (Of. 62)

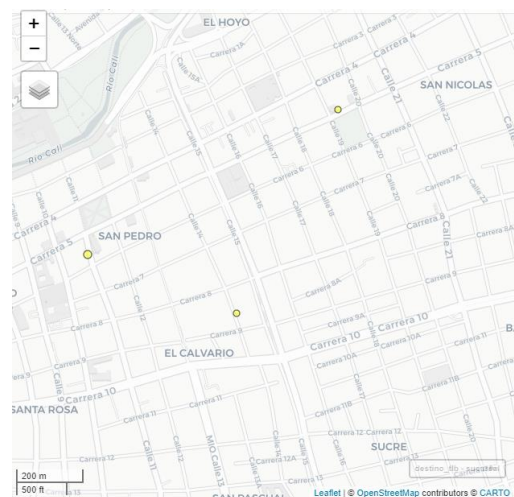
Al comparar el indicador 1, donde se asignó igual ponderación a las variables, con el indicador 4 generado mediante el Análisis de Componentes Principales, se identificó una zona común de alta densidad en el barrio San Nicolás. Al hacer un acercamiento a estas áreas y ubicar los puntos de servicio existentes, se observó que en la zona sombreada de mayor impacto del indicador como se mencionó antes, se encuentran tres puntos de servicio: la oficina 62, 60 y 803, correspondientes a San Nicolás Cali, Cali Principal y Calle Catorce, respectivamente. Esta información se puede apreciar en la ilustración 27.

Ilustración 25. Zona óptima Indicador 1



Fuente: Elaboración propia

Ilustración 26. Puntos de atención zona óptima



Fuente: Elaboración propia

El punto de servicio que se encuentra más cercano a la zona óptima señalada es el ubicado en la Cl. 20 #5-

49, COMUNA, en el barrio San Nicolás. Este punto se destaca por ser clave en la red de atención a clientes ya que se encuentra en un sector céntrico de la ciudad, rodeado de sectores industriales, comerciales y tradicionales de la historia caleña y se demuestra su importancia en las operaciones de la entidad financiera al observar una densidad significativa en el volumen de transacciones de los clientes del segmento empresarial. Este punto presenta un horario partido entre 8 y 11.30 de la mañana y entre 2 y 4:30 de la tarde para una atención de caja de divisas, caja general, asesoría y fila empresarial para Pymes.

Por otro lado, se tiene el punto de servicio llamado Cali Principal que queda ubicado cerca de la carrera 4 y la Plaza de Caicedo en el centro de la ciudad. Este punto tiene un horario en jornada continua de 8am a 4:30pm y los sábados de 8am hasta el mediodía y presta los mismos servicios del punto de San Nicolás.

El tercer punto de servicio que se encuentra dentro de la zona crítica óptima tiene el nombre de Oficina Calle Catorce (Of. 803), este se encuentra situado de igual forma en el centro de la ciudad, sin embargo, la ubicación es un poco más hacia el sur que los otros dos puntos, cercano a zonas críticas de comercio como el Pasaje Cali y la Estación de Transporte Petecuy. De manera similar a la oficina de Cali Principal, se ofrece atención al público en jornada continua hasta las 4.30 de la tarde y los sábados hasta el mediodía, no obstante, a diferencia de los otros dos puntos, en esta oficina no se presta el servicio de caja de divisas.

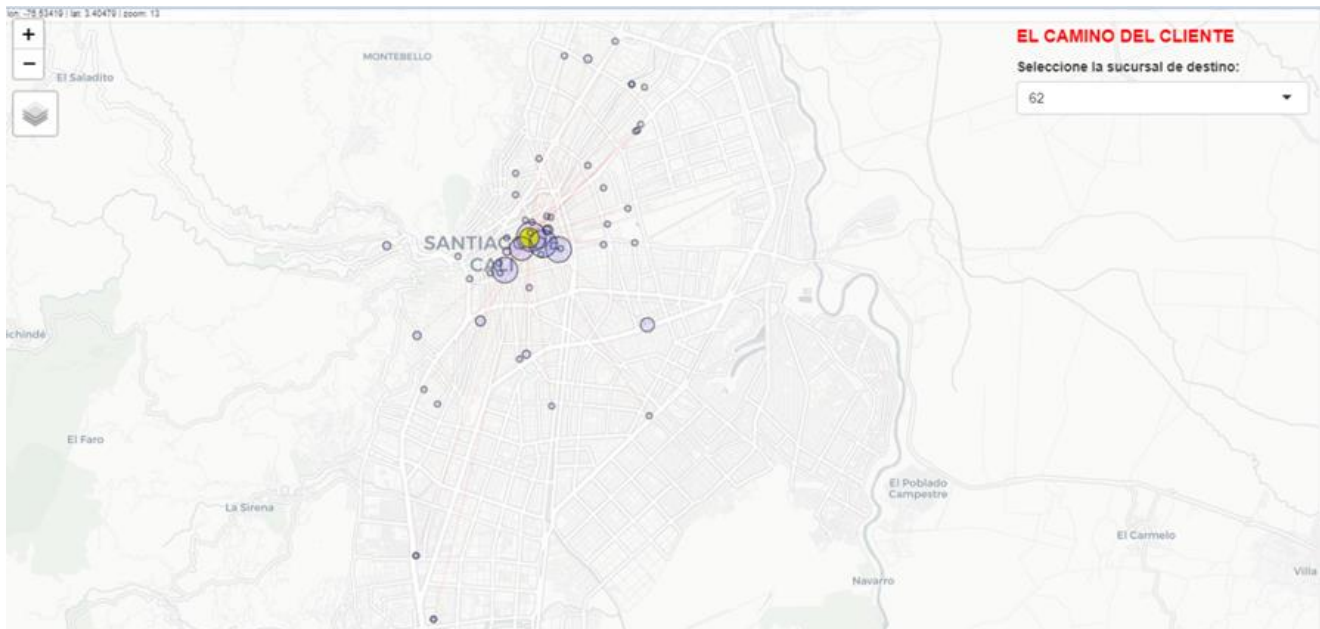
7.2. UBICACIÓN DEL POSIBLE PUNTO DE ATENCIÓN

Después de realizar el análisis previo, se puede concluir que el punto de atención potencial se encuentra en el punto de servicio No. 62, que corresponde a la Oficina de San Nicolás Cali ubicada en Cl. 20 #5-49, COMUNA 3, Cali, Valle del Cauca. Por medio de la librería “shiny” se desarrolló una aplicación que permitió identificar los lugares desde donde están llegando los clientes a esta oficina y la proporcionalidad en monto de transaccionalidad realizadas. En la ilustración 27, se representa el punto de servicio seleccionado el color amarillo, mientras que los círculos azules representan a los clientes. Además, se observa que estos clientes provienen de diferentes lugares de la ciudad, como el norte, sur, oriente y occidente.

Adicionalmente en la ilustración 27, el tamaño de los círculos refleja el monto transado por cada cliente y

es notable evidenciar que los clientes con los montos más altos se encuentran muy cerca del punto de servicio seleccionado, lo que indica una concentración de transacciones de gran valor en esa área específica.

Ilustración 27. Punto de servicio – San Nicolas (Of. 62)



Fuente: Elaboración propia

Por otro lado, el barrio San Nicolás de la ciudad de Cali también se destaca por su vibrante actividad comercial e industrial. En su entorno, se encuentran diversas zonas dedicadas al comercio, con una amplia variedad de tiendas, mercados y establecimientos que satisfacen las necesidades de los residentes y visitantes. Asimismo, el barrio alberga importantes áreas industriales, donde se desarrollan diversas actividades productivas y manufactureras. Estas zonas industriales son un motor económico para la ciudad, generando empleo y contribuyendo al crecimiento y desarrollo de la comunidad. San Nicolás se consolida como un barrio dinámico, donde la vida comercial e industrial se entrelaza en armonía con su identidad y tradiciones locales.

8. CONCLUSIÓN Y TRABAJOS FUTUROS

8.1. CONCLUSIONES

- ❖ El análisis y entendimiento de los datos de las transacciones realizadas por los clientes empresariales en el primer periodo de 2022 en la Ciudad de Cali, permitió identificar y definir variables de interés que brindaron información valiosa sobre la ubicación y comportamiento de los clientes y su relación con las transacciones comerciales y los puntos de servicio donde estos están acudiendo. Esta información, fue de gran utilidad para realizar el análisis que permitió obtener las zonas optimas con las que se tomaría la decisión de selección del punto de servicio.
- ❖ Los resultados del análisis exploratorio revelaron valiosa información acerca de la distribución geográfica de los clientes y los puntos de servicio en Cali. Se identificaron posibles áreas con una alta concentración de clientes o una densidad significativa de puntos de servicio, lo que fue de gran relevancia para el análisis realizado y para la toma de decisiones que esto podría representar para la entidad financiera en términos de planificación estratégica
- ❖ La generación de imágenes de densidad o mapas de calor de las variables seleccionadas, así como su combinación para crear diferentes escenarios e indicadores, brindó una metodología efectiva para delimitar las zonas óptimas para la posible habilitación del punto de servicio a clientes empresariales. Estos hallazgos proporcionan una base sólida para la toma de decisiones estratégicas en cuanto a la ubicación del punto de servicio, ofreciendo la posibilidad de que la empresa se adapte de manera eficiente a las necesidades de sus clientes empresariales. Además, este enfoque puede ser aplicado en otros contextos empresariales y de planificación, brindando una herramienta útil para la optimización espacial de servicios.
- ❖ Es uso de la metodología de Análisis de Componentes Principales - ACP, es una herramienta significativa para asignar priorización de variables. En este proyecto permitió priorizar e identificar posibles puntos de atención para una entidad financiera, transformando un conjunto de variables correlacionadas en un nuevo conjunto de variables no correlacionadas.

8.2. TRABAJOS FUTUROS

- ❖ Se puede investigar y desarrollar técnicas avanzadas de análisis de datos para mejorar la precisión y la calidad de la información obtenida. Esto puede incluir el uso de algoritmos de aprendizaje automático más sofisticados, como redes neuronales o algoritmos de agrupamiento más avanzados, que permitan descubrir patrones más complejos en los datos de transacciones.
- ❖ En lugar de realizar el análisis retrospectivo de los datos del primer periodo de 2022, se podría trabajar en el desarrollo de sistemas de análisis en tiempo real. Esto permitiría monitorear y analizar continuamente los datos de transacciones a medida que se generan, lo que proporcionaría información actualizada y permitiría tomar decisiones más ágiles y oportunas.
- ❖ Basándose en los datos históricos de transacciones y en la información obtenida sobre los puntos de servicio y la distribución geográfica de los clientes, se puede desarrollar un modelo de análisis predictivo. Esto permitiría predecir futuros patrones de comportamiento de los clientes y su relación con los puntos de servicio, lo que ayudaría a la entidad financiera a anticiparse a las necesidades y preferencias de los clientes y tomar decisiones estratégicas más informadas.
- ❖ Aunque el análisis actual se centró en variables específicas relacionadas con la ubicación y el comportamiento de los clientes, es importante considerar otras variables relevantes que puedan afectar las transacciones y la elección del punto de servicio. Una variable como "hurtos" puede proporcionar información valiosa sobre la seguridad en determinadas áreas y su impacto en la selección de puntos de servicio. Un trabajo futuro podría consistir en recopilar datos sobre hurtos y realizar un análisis adicional para evaluar cómo esta variable se relaciona con la distribución geográfica de los clientes y los puntos de servicio. Esto permitiría una comprensión más completa del entorno y ayudaría a tomar decisiones estratégicas más informadas sobre la ubicación del punto de servicio.
- ❖ Aunque es un proyecto elaborado a partir de las transacciones realizadas en las sucursales de una entidad financiera en la ciudad de Santiago de Cali y pensando en un adecuado servicio al cliente tiene como base fundamental el comportamiento y cumplimiento por parte de la organización frente

a las exigencias de quien recibe el servicio, es importante resaltar que este proyecto se puede aplicar en otras ciudades del país no solo para Oficina de Servicios Especializados, sino además la ubicación de cajeros automáticos, nuevas sucursales y otros.

- ❖ El enfoque utilizado en este proyecto puede ser aplicado a otros contextos empresariales y de planificación. Se puede investigar y adaptar la metodología y las técnicas utilizadas para brindar una herramienta útil para la optimización espacial de servicios en diferentes industrias y sectores.

BIBLIOGRAFÍA

- [1] Bolsa de Valores de Colombia, «Bolsa de Valores de Colombia,» 2014. [En línea]. Available: http://www.bvc.com.co/pps/tibco/portalbvc/Home/Empresas/Guia_Mercado_Valores?com.tibco.ps.pagesvc.action=updateRenderState&rp.currentDocumentID=-7ca0c036_147b6b20b27_5e970a0a600b&rp.revisionNumber=1&rp.attachmentPropertyName=Attachment&com.tibco.ps.pagesv.
- [2] M. A. Gutierrez Plata, 20 Noviembre 2014. [En línea]. Available: <https://repository.unimilitar.edu.co/handle/10654/13148>.
- [3] C. I. Value, 2020. [En línea]. Available: <https://www.civ.com.co/sites/default/files/EI%20ranking%20de%20la%20satisfacci%C3%B3n%20de%20los%20clientes%20de%20la%20banca%20en%20Colombia%202019.pdf>.
- [4] F. J. Rodríguez-Cortés, *¿Que tiene en común la estadística y los mapas?*, Medellín, Colombia: Escuela de Estadística, Univeridad Nacional de Colombia, , 2018.
- [5] F. J. Rodríguez-Cortés, *Análisis exploratorio de datos espaciales*, Medellín, Colombia: Escuela de Estadística, Univeridad Nacional de Colombia, 2018.
- [6] F. J. Rodríguez-Cortés, *¿Qué es la intensidad de un patrón de puntos y por qué usarla?*, Medellín, Colombia: Escuela de Estadística, Univeridad Nacional de Colombia, 2018.
- [7] J. M. C. M. D. Santos Preciado. [En línea]. Available: <https://canal.uned.es/uploads/materials/resources/pdf/3/6/1259093906763.pdf> .
- [8] D. H. L. Q. M. P. J. A. Dianda. [En línea]. Available: https://rehip.unr.edu.ar/bitstream/handle/2133/7560/Dianda%20y%20otros_%20indicadores%20multivariados.pdf?sequence=3&isAllowed=y.
- [9] G. d. G. S. d. Rosario. [En línea]. Available: https://www.fceia.unr.edu.ar/gps/cursos/Georreferenciacion_2019.pdf .
- [10] OSSO, 2003. [En línea]. Available: <https://www.osso.org.co/docu/tesis/2003/evaluacion/analisis.pdf>.
- [11] S. B. M. DÍAZ, «Unimilitar,» 2013. [En línea]. Available: <https://repository.unimilitar.edu.co/handle/10654/10914>.
- [12] D. C. D. M. J. P. M. S. R. R. R. a. S. T. Catherine Cabrera, «MINITIC,» 2020. [En línea]. Available: https://www.mintic.gov.co/micrositios/cienciadedatos/747/articles-151167_recurso_1.pdf.
- [13] H. T. M. M. P. a. A. K. A. Amoozad Mahdiraji, 01 Junio 2021. [En línea]. Available: <https://www.emerald.com/insight/content/doi/10.1108/BIJ-03-2021-0127/full/html#:~:text=The%20authors%20propose%20a%20novel%20hybrid%20multi-attribute%20data,a%20large-scale%20problem%20in%20the%20financial%20services%20industry..>