



Estrategias de optimización para canales transaccionales físicos en el sector bancario colombiano.

JOHN FREDY ZAPATA JIMÉNEZ

tesis

Asesor, docente  
Ana María López Moreno

UNIVERSIDAD EAFIT  
ESCUELA DE CIENCIAS APLICADAS E INGENIERÍA  
MAESTRÍA EN CIENCIAS DE LOS DATOS Y ANALÍTICA  
MEDELLÍN  
2025

## CONTENIDO

<i>RESUMEN</i> .....	3
<i>ABSTRACT</i> .....	3
<i>INTRODUCCIÓN</i> .....	5
<i>PLANTEAMIENTO DEL PROBLEMA</i> .....	6
<i>JUSTIFICACIÓN</i> .....	6
<i>OBJETIVOS</i> .....	7
General.....	7
Específicos.....	7
<i>MARCO CONCEPTUAL</i> .....	8
<i>DISEÑO METODOLÓGICO</i> .....	9
Muestrear.....	9
Explorar.....	10
Modificar.....	18
Modelado.....	19
Evaluación.....	29
Resultados.....	30
<i>CONCLUSIONES</i> .....	31
<i>ANEXOS</i> .....	31
Anexo 1: metadatos tecnicos DE HISTORICO_formato_413.....	31
Anexo 2: metadatos tecnicos DE HISTORICO_formato_412.....	32
Anexo 3: metadatos tecnicos DE TBL_dim_canal_super.....	32
Anexo 4: metadatos tecnicos DE TBL_dim_censo_dane.....	33
Anexo 5: metadatos tecnicos DE TBL_dim_codigo_ciu_super.....	33
Anexo 6: metadatos tecnicos DE TBL_dim_departamento_super.....	33
Anexo 7: metadatos tecnicos de tbl_dim_mecanismo_confirmacion_super.....	33
Anexo 7: metadatos tecnicos de tbl_dim_municipio_super.....	33
<i>REFERENCIAS</i> .....	34

## RESUMEN

La transformación digital está cambiando rápidamente el panorama de la banca tradicional, creando un dilema para las instituciones bancarias: integrar nuevos canales digitales o mejorar la distribución de los canales físicos existentes. Esta tesis explora cómo las técnicas de optimización multiobjetivo, como la programación lineal entera y la simulación estocástica de eventos discretos, pueden ayudar a resolver este dilema en el contexto del sistema financiero colombiano.

En un entorno donde los hábitos de los clientes y los modelos de distribución están en constante evolución, los responsables de la toma de decisiones deben considerar el impacto de la tecnología, los costos de implementación y la adaptabilidad de los canales. Esta investigación aborda estos desafíos desarrollando un marco teórico basado en modelación heurística y técnicas avanzadas como el clustering y algoritmos NLP. El objetivo es ofrecer recomendaciones sobre cómo optimizar la distribución de canales transaccionales para mejorar la eficiencia operativa y la experiencia del cliente.

La tesis se centra en cuatro objetivos específicos: recuperar y almacenar datos de transacciones de canales de distribución, preparar esta información para la modelación de clustering, desarrollar un modelo de optimización para la distribución de los canales físicos y analizar la información para segmentar los canales según el modelo de optimización.

Optimizar los canales de distribución es esencial para mantener una ventaja competitiva en un entorno digitalizado. Al combinar canales digitales y físicos de manera eficaz, el sistema bancario actual puede mejorar la eficiencia operativa, ampliar su alcance y responder de manera más ágil a las demandas del mercado. Este estudio proporciona una perspectiva integral y soluciones prácticas para enfrentar los desafíos actuales en la distribución de canales transaccionales físicos en el sistema bancario colombiano.

## ABSTRACT

Digital transformation is rapidly reshaping the landscape of traditional banking, creating a dilemma for financial institutions: integrate new digital channels or improve the distribution of existing physical ones. This thesis explores how multi-objective optimization techniques, such as integer linear programming and discrete-event stochastic simulation, can help address this dilemma within the context of the Colombian financial system.

In an environment where customer habits and distribution models are constantly evolving, decision-makers must consider the impact of technology, implementation costs, and the adaptability of channels. This research addresses these challenges by developing a theoretical framework based on heuristic modeling and advanced techniques such as clustering and NLP algorithms. The aim is to provide recommendations for optimizing the distribution of transactional channels to enhance operational efficiency and customer experience.

The thesis focuses on four specific objectives: retrieving and storing transaction data from distribution channels; preparing this information for clustering modeling; developing an optimization model for the distribution of physical channels; and analyzing the information to segment channels according to the optimization model.

Optimizing distribution channels is essential to maintaining a competitive advantage in an increasingly digital environment. By effectively combining digital and physical channels, the banking system can improve operational efficiency, broaden its reach, and respond more agilely to market demands. This study offers a comprehensive perspective and practical solutions to address current challenges in the distribution of physical transactional channels in the Colombian banking sector.

**Keywords:** distribution, channels, Banking, Consumer, Optimization, Heuristic modeling, Process optimization, Clustering techniques, NLP algorithms, Efficient, Colombian, sector.

## INTRODUCCIÓN

Los canales convencionales, que históricamente han sido la columna vertebral de las operaciones bancarias, están siendo confrontados por la creciente demanda de soluciones digitales que faciliten una experiencia más ágil y accesible. Esta transformación plantea un dilema significativo para las instituciones bancarias: decidir entre incorporar nuevos canales transaccionales digitales o mejorar la distribución de los canales físicos existentes.

En este contexto, los enfoques de optimización multiobjetivo se presentan como herramientas valiosas para abordar dicho dilema. Técnicas como la programación lineal entera y la simulación estocástica de eventos discretos pueden equilibrar de forma eficaz los niveles de servicio y los costos operativos en la banca minorista, proporcionando un marco sólido para la toma de decisiones estratégicas en entornos cambiantes. Estos métodos permiten a las instituciones bancarias adaptarse a las nuevas demandas del mercado mientras gestionan de manera óptima sus recursos y costos (Baldassarre et al., 2023; Guzovski et al., 2022; Nayanajith et al., 2021).

Las investigaciones recientes coinciden en que los bancos deberían invertir en comprender con mayor profundidad las necesidades y preferencias de los clientes para mejorar sus estrategias de distribución (Guzovski et al., 2022). No obstante, la presente tesis propone un enfoque complementario, orientado al análisis de datos objetivos y cuantificables que permitan optimizar la asignación de recursos físicos, buscando un balance entre eficiencia operativa y cobertura territorial.

Para sustentar esta perspectiva, se realizó una revisión sistemática de la literatura, considerando la evolución de conceptos y prácticas contemporáneas en la distribución de servicios bancarios. Por ejemplo, Baldassarre et al. (2023) desarrollaron un modelo jerárquico de ubicación de cobertura que integra canales físicos y digitales, minimizando costos y maximizando el acceso. De igual forma, Nayanajith et al. (2021) analizaron la distribución de canales en función de la densidad de población y los patrones transaccionales, demostrando que la integración de datos geográficos y de comportamiento permite optimizar la infraestructura existente.

A partir de esta revisión bibliográfica, se construyó un marco teórico basado en modelación heurística, con énfasis en técnicas de optimización de procesos. Este marco se apoya en herramientas como el clustering y algoritmos de procesamiento de lenguaje natural (NLP), que facilitan la formulación de recomendaciones estratégicas para una redistribución más eficiente de los canales físicos.

Aunque el estudio se centra en el comportamiento transaccional bancario en Colombia, el marco desarrollado presenta potencial para ser aplicado en otros contextos dentro del sector bancario latinoamericano. Su aporte principal radica en ofrecer a los responsables de la toma

de decisiones una visión estructurada y basada en evidencia, que integra la eficiencia operativa con la inclusión financiera.

## **PLANTEAMIENTO DEL PROBLEMA**

En la actualidad, el sector bancario está en una encrucijada crítica debido a la rápida evolución tecnológica que está transformando los hábitos de los clientes y desafiando los modelos tradicionales de distribución de los canales transaccionales. Investigaciones recientes destacan la importancia de optimizar los canales de distribución físicos en el sector bancario.

Los sistemas multi-canal pueden aumentar la satisfacción, la demanda y la rentabilidad del cliente, y los modelos matemáticos ayudan a optimizar estos sistemas (Mohsen Gheitani et al. 2022). Los responsables de tomar decisiones que se encargan de captar e invertir fondos en una entidad financiera suelen enfrentarse al siguiente problema: decidir y desarrollar la estructura de cada posición (Zu Selhausen, et al. 1977).

La elección de la estrategia de distribución debe tener en cuenta el impacto de la tecnología en el desarrollo de canales de distribución modernos, el costo de implementación y la adaptabilidad a nuevos canales, al tiempo que se crea un valor adicional que permita la diferenciación del propio servicio. (guzovski, rudančić, zgurić, et al. 2022).

La presente investigación se enfoca en abordar estos desafíos a través de un análisis exhaustivo de las metodologías de optimización multiobjetivo, como la programación lineal entera y la simulación estocástica de eventos discretos, para evaluar la distribución óptima de los canales transaccionales físicos en el contexto colombiano.

La tesis explora cómo estas herramientas pueden ayudar a las instituciones bancarias, a tomar decisiones informadas sobre la incorporación de nuevos canales digitales y la gestión de los existentes.

Se propone un marco teórico basado en la modelación heurística y técnicas avanzadas de clustering y algoritmos NLP para ofrecer recomendaciones sobre la mejor manera de integrar y optimizar la distribución de los canales físicos.

## **JUSTIFICACIÓN**

La optimización de los canales de distribución en el ámbito transaccional de las instituciones bancarias se vuelve esencial para mantener una ventaja competitiva en un entorno cada vez más digitalizado. Aprovechar los canales virtuales y la banca digital es crucial para mejorar tanto la eficiencia operativa como el alcance hacia los clientes. Las instituciones bancarias pueden beneficiarse significativamente al fortalecer sus operaciones mediante el uso de transacciones virtuales, lo que no solo facilita la expansión del mercado, sino que también

permite una promoción más efectiva de productos y una agilización de las transacciones financieras (Andrade Merino et al., 2018).

Además, la incorporación de modelos de optimización en las campañas de orientación de canales físicos se ha demostrado como una estrategia efectiva para mejorar la eficacia en la distribución transaccional. Los modelos que integran productos empaquetados y estructuras multicanal pueden aumentar significativamente la efectividad de las operaciones bancarias (Delanote et al., 2013).

Si bien este estudio toma como referencia a las entidades líderes del mercado debido a su amplio alcance transaccional y su infraestructura física representativa, las estrategias propuestas pueden ser igualmente adaptadas y beneficiosas para instituciones medianas o pequeñas. Estas también enfrentan desafíos de cobertura, eficiencia operativa y segmentación de bancos que pueden abordarse con modelos similares.

En este sentido, un modelo de optimización multicanal que combine de forma estratégica canales físicos puede ser una herramienta clave no solo para mantener la competitividad en bancos consolidados, sino también para impulsar el crecimiento, mejorar la experiencia del cliente y optimizar los procesos internos en entidades emergentes o de menor escala. La aplicación de estas estrategias permitirá a las entidades bancarias que las adopten mantenerse a la vanguardia en un sector altamente competitivo, responder de manera más ágil y eficaz a las demandas del mercado, y avanzar hacia una inclusión financiera más equitativa.

## **OBJETIVOS**

### **General**

Desarrollar e implementar un modelo de optimización y personalización de los canales de distribución físico en el sector bancario de Colombia.

### **Específicos**

1. Recuperar la información publicada en el portal [www.datos.gov.co](http://www.datos.gov.co), informe público “Operaciones efectuadas a través de los canales de distribución o servicios transaccionales” y almacenarla en un ecosistema de base de datos local.
2. Preparar la información con el fin de poder desarrollar un modelo clusterización adaptado a la información de los canales de distribución transaccional en el sector bancario de Colombia.
3. Desarrollar un modelo de optimización de la distribución de canales transaccionales físicos.

4. Analizar la información resultante mediante comprobación de distribución real vs la optimizada por el modelo.

## MARCO CONCEPTUAL

La optimización de los canales de distribución en el sector bancario colombiano es un componente estratégico para mejorar la eficiencia operativa, reducir costos y fortalecer la experiencia del cliente en un entorno caracterizado por la acelerada digitalización. Este concepto implica la adecuada asignación de recursos físicos (como oficinas, cajeros automáticos y corresponsales) y digitales (como aplicaciones móviles o plataformas web), considerando tanto criterios de cobertura territorial como de impacto transaccional.

La literatura especializada ha identificado distintos enfoques teóricos y metodológicos para abordar esta problemática. Uno de los más relevantes es el enfoque de optimización multiobjetivo, el cual permite modelar sistemas complejos con objetivos contrapuestos, como la maximización de cobertura versus la minimización de costos. Dentro de este enfoque, se utilizan técnicas como la programación lineal entera o los modelos de programación matemática jerárquica, que han sido aplicados en problemas de localización de instalaciones bancarias (Min & Melachrinoudis, 2001; Basar et al., 2017).

Asimismo, las técnicas de aprendizaje no supervisado, especialmente los algoritmos de clustering, permiten segmentar entidades o territorios según patrones operativos similares, apoyando la toma de decisiones en la planificación y redistribución de canales. Algoritmos como K-Means, DBSCAN o HDBSCAN son ampliamente utilizados para este propósito, acompañados de métricas de evaluación como el índice de silueta, el índice de Davies–Bouldin o el coeficiente de Calinski–Harabasz, que permiten validar la coherencia de los grupos generados (Tan, Steinbach & Kumar, 2018).

El uso de Sistemas de Información Geográfica (SIG) también se ha consolidado como una herramienta complementaria que permite incorporar variables geoespaciales en los modelos de optimización. Esto facilita la visualización de la cobertura actual y potencial, así como la identificación de brechas territoriales en la oferta de servicios financieros (Miliotis et al., 2002; Morrison & O'Brien, 2001).

Adicionalmente, la reestructuración de redes físicas ha sido abordada a través de modelos que consideran factores socioeconómicos y de demanda proyectada. Por ejemplo, Denstad et al. (2019) proponen modelos multiobjetivo para la reubicación de cajeros automáticos que incorporan tanto la infraestructura existente como la demanda futura, mientras que Farahani et al. (2014) plantean modelos de diversificación jerárquica en redes de sucursales para mejorar la eficiencia operativa y segmentar la oferta de servicios.

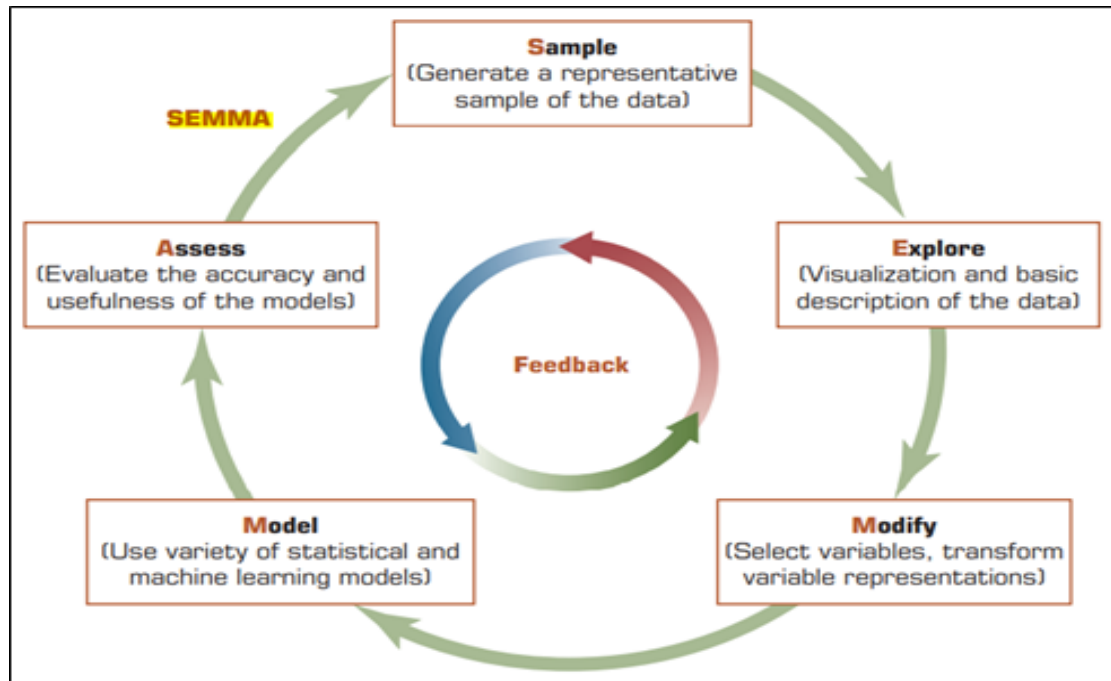
La presente investigación combina modelos de optimización matemática para asignar canales físicos de manera eficiente, técnicas de clustering y minería de datos para segmentar entidades o regiones con comportamientos similares, y enfoques geoespaciales que

consideran variables de localización y densidad poblacional, con el fin de diseñar un modelo integral y eficiente.

Este marco teórico proporciona los fundamentos necesarios para la construcción de un modelo robusto y adaptado a las condiciones del sistema bancario colombiano, buscando mejorar la distribución de canales físicos en función de su impacto real y potencial.

## DISEÑO METODOLÓGICO

La presente investigación adopta un enfoque cuantitativo, se centrará en aplicar los pasos de la metodología SEMMA (Muestrear, Explorar, Modificar, Modelar y Evaluar) como estructura principal para abordar el problema de investigación. SEMMA será utilizado para guiar el proceso de análisis y extracción de conocimiento desde los datos relacionados con el problema identificado.



*Ilustración 1. Metodología Utilizada. fuente: Sharda, R., Delen, D., Turban, E. (2018)*

Esta metodología divide el proceso de minería de datos en 5 fases principales que permiten llevar a cabo un enfoque iterativo y sistemático. A continuación, se describen cada una de ellas y cómo se aplicarán en el contexto de este proyecto:

### Muestrear

El panorama actual del sector bancario colombiano está marcado por un proceso acelerado de digitalización, impulsado por la adopción creciente de canales no presenciales para realizar transacciones financieras, esta tendencia refleja un cambio en el comportamiento de los usuarios hacia el uso de tecnologías digitales para sus necesidades bancarias.

Sin embargo, a pesar de este avance, persisten desafíos importantes, como la optimización de la ubicación estratégica de los canales físicos en zonas clave y el fortalecimiento del impacto de los canales digitales en áreas geográficas donde las entidades bancarias aún no logran alcanzar a una parte significativa de la población. Por lo tanto, el objetivo principal de este proyecto es analizar y evaluar el impacto transaccional de las principales entidades bancarias en Colombia, identificando patrones y oportunidades de mejora.

Los hallazgos permitirán orientar decisiones estratégicas para maximizar la eficiencia operativa y mejorar la experiencia del cliente en el contexto de la digitalización bancaria.

### **Explorar**

Para desarrollar el análisis y dar cumplimiento a los objetivos planteados, se emplearon conjuntos de datos provenientes de tres fuentes públicas de alta confiabilidad: el Departamento Administrativo Nacional de Estadística (DANE), el Portal de Datos Abiertos del Estado Colombiano y la Dirección de Impuestos y Aduanas Nacionales (DIAN). Cada una de estas fuentes aportó información específica y complementaria que, en conjunto, permitió construir una base de datos robusta para el estudio.

En primer lugar, el DANE, como organismo oficial encargado de la producción estadística en Colombia, proporcionó datos demográficos y geográficos fundamentales. Estos incluyeron la población total por región, las coordenadas de latitud y longitud para departamentos y municipios, así como la distribución etaria de la población. Esta información resultó clave para comprender las características socioeconómicas y demográficas que inciden tanto en la adopción de canales digitales como en la demanda de canales físicos.

Por su parte, el Portal de Datos Abiertos del Estado Colombiano, iniciativa gubernamental orientada a facilitar el acceso libre a información de interés público, suministró dos conjuntos de datos relevantes. El primero, denominado “Operaciones efectuadas a través de los canales de distribución o servicios transaccionales”, recoge información detallada sobre las transacciones realizadas por los clientes de entidades financieras a través de múltiples canales, incluyendo aplicaciones móviles, cajeros automáticos propios y de terceros, centros de atención telefónica, asistentes virtuales, corresponsales bancarios físicos y digitales (propios y tercerizados), internet, oficinas, terminales POS y sistemas de acceso remoto, entre otros. Este conjunto de datos también contempla los servicios transaccionales asociados a pagos clasificados por actividad económica y a operaciones monetarias según mecanismo de confirmación. El segundo conjunto, “Puntos de atención de canales por municipio”, ofrece un inventario de la infraestructura física y digital disponible en el territorio nacional, incluyendo oficinas, cajeros automáticos, datáfonos, corresponsales físicos y digitales, y

otros puntos de atención operados por distintos tipos de entidades financieras vigiladas por la Superintendencia Financiera de Colombia. Estos registros permiten construir una visión integral sobre la distribución territorial de los canales de atención y su relación con el comportamiento transaccional.

Finalmente, la DIAN, como autoridad encargada de la administración tributaria y aduanera, aportó información económica a través de la “Clasificación Industrial Internacional Uniforme de todas las actividades económicas” (CIIU Rev. 4 A.C.). Esta clasificación constituye un estándar internacional que organiza las actividades productivas en categorías homogéneas, facilitando la presentación y el análisis comparativo de la información estadística.

La integración de estas tres fuentes de datos permitió consolidar un repositorio unificado con información demográfica, geográfica, económica y transaccional. Este enfoque multidimensional no solo fortalece la representatividad de los análisis, sino que también proporciona una base sólida para abordar con rigor los objetivos de optimización planteados en este estudio.

## **ETL**

Para consolidar la información proveniente de las tres fuentes seleccionadas, se diseñó e implementó un flujo automatizado de Extracción, Transformación y Carga (ETL), apoyado en herramientas de programación y sistemas de gestión de bases de datos. Este proceso garantizó la obtención, validación y organización sistemática de los datos requeridos para el análisis.

En la etapa de extracción, se empleó el lenguaje de programación Python como herramienta principal para recuperar automáticamente los conjuntos de datos del DANE, el Portal de Datos Abiertos del Estado Colombiano y la DIAN. Durante esta fase, se validaron tanto los formatos como la estructura de los archivos descargados, con el fin de asegurar la coherencia y calidad de la información. Una vez verificados, los datos fueron organizados en repositorios específicos, clasificados por fuente y por tipo de contenido, lo que facilitó su posterior tratamiento y análisis.

## **Exploración inicial de los datos:**

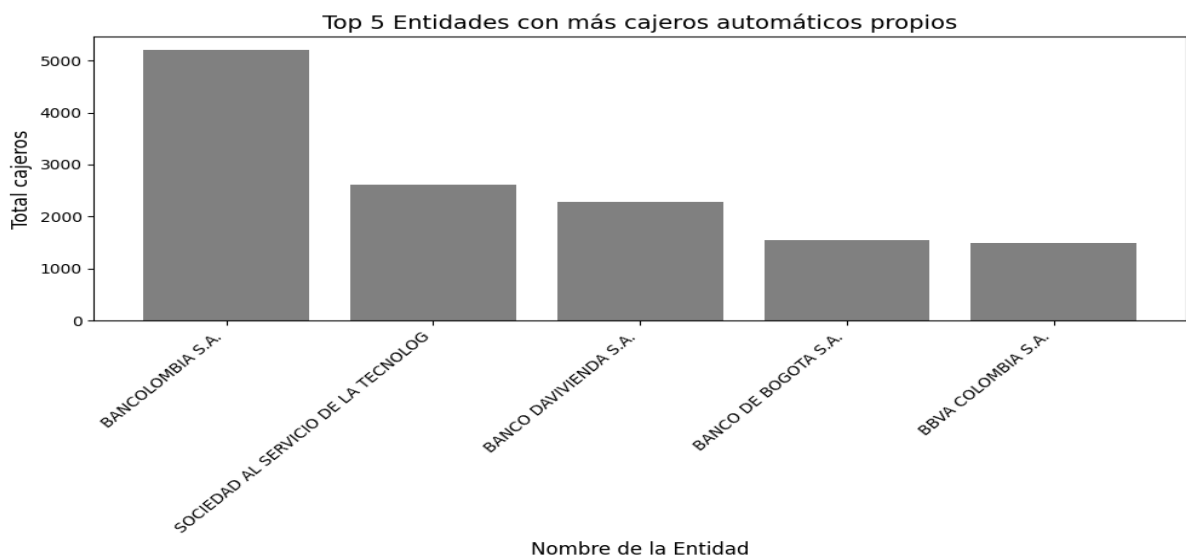
*Tabla 1: Resumen de las bases de datos Utilizadas*

<b>Nombre de la Tabla</b>	<b>Filas</b>	<b>Columnas</b>	<b>Descripción</b>
historico_formato_413	10.100.000	16	Registro transaccional detallado de operaciones financieras.

historico_formato_412	653.000	11	Información agregada de reportes financieros por entidad.
tbl_dim_canal_super	19	4	Catálogo de canales de atención supervisados.
tbl_dim_censo_dane	20.191	6	Censo poblacional del DANE por municipio o región.
tbl_dim_codigo_ciiu_super	665	2	Clasificación CIU de las entidades financieras.
tbl_dim_departamento_super	34	4	Catálogo de departamentos (entidades territoriales).
tbl_dim_mecanismo_confirmacion_super	8	2	Tipos de mecanismos de autenticación o confirmación.
tbl_dim_municipio_super	1.123	4	Catálogo oficial de municipios supervisados.

Además, se llevó a cabo un análisis con la información consolidada, seleccionando de la base de datos las entidades bancarias con mayor relevancia y participación a nivel transaccional, así como cada infraestructura de canales físicos correspondiente.

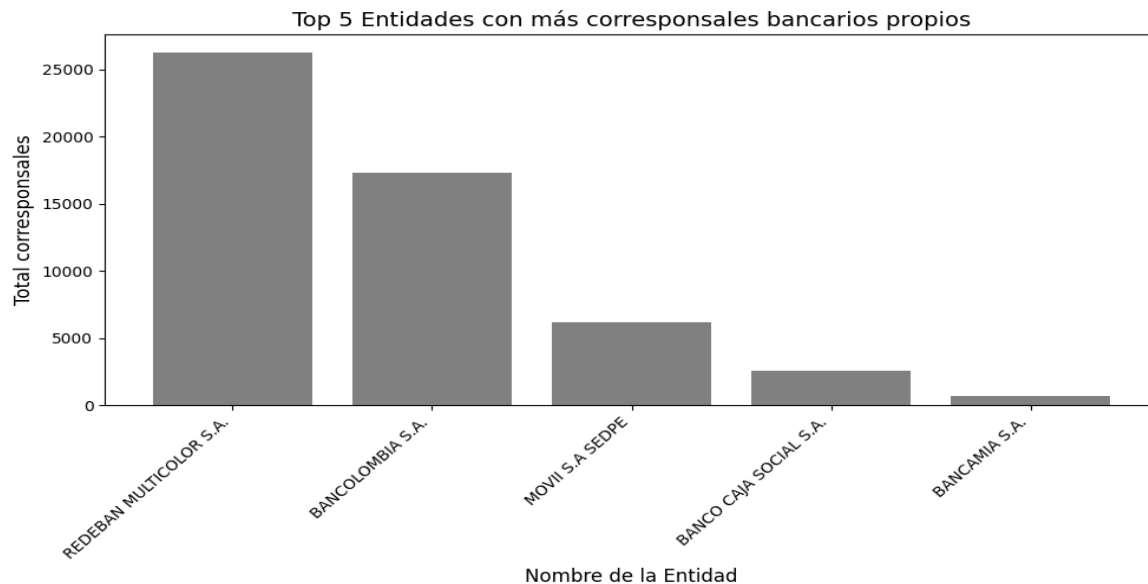
A continuación, se presentan visualmente los resultados que demuestran la robustez de los principales actores en los canales físicos y digitales.



*Ilustración 2. Distribución de cajeros automáticos en Colombia. Fuente: Elaboración propia.*

El análisis de los cajeros automáticos en el sistema bancario colombiano muestra que **BANCOLOMBIA S.A.** lidera con una ventaja significativa, superando las 5.000 unidades a nivel nacional. Esta alta concentración refleja su capacidad operativa y cobertura territorial.

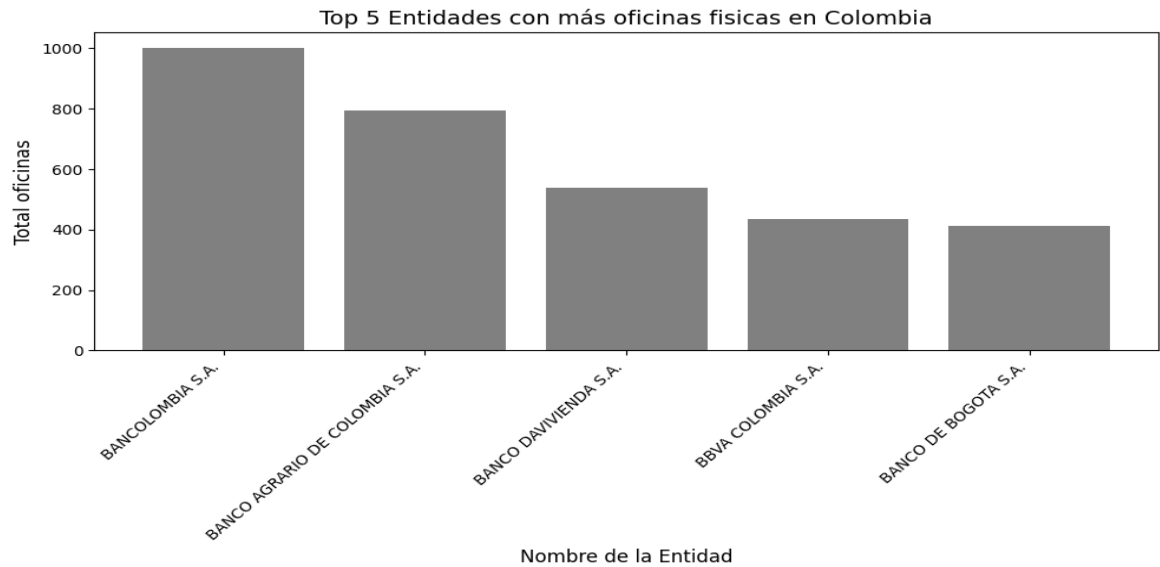
Además, los bancos del Grupo Aval, como Banco de Bogotá y Banco de Occidente, presentan una sólida inversión en infraestructura física, aunque en una escala menor.



*Ilustración 3. Distribución de corresponsales bancarios en Colombia. Fuente: Elaboración propia.*

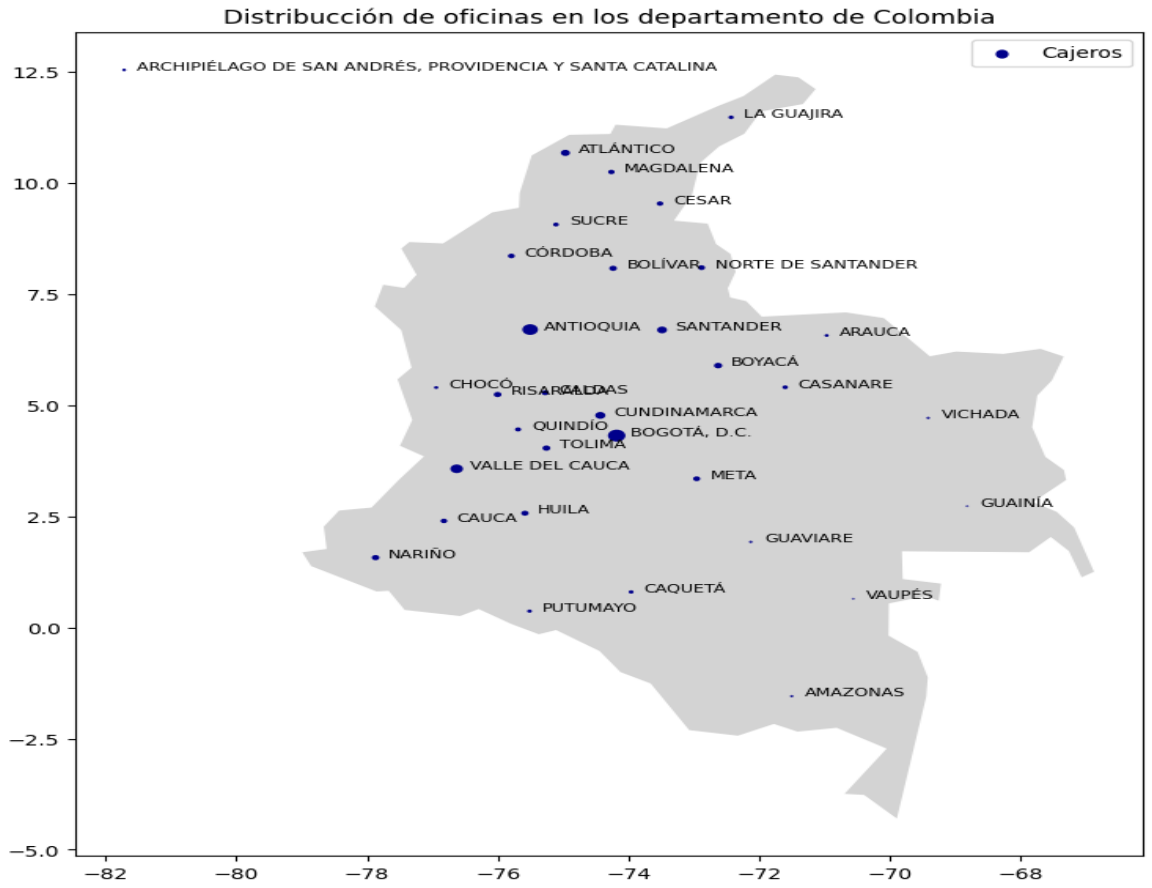
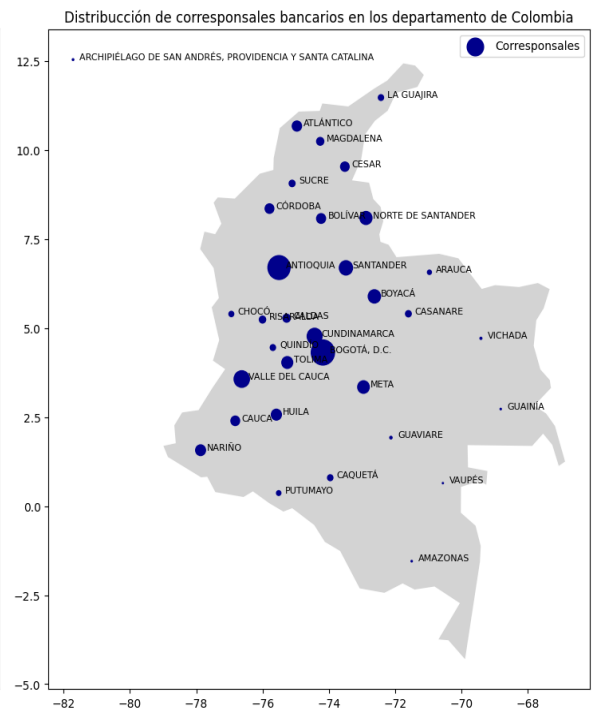
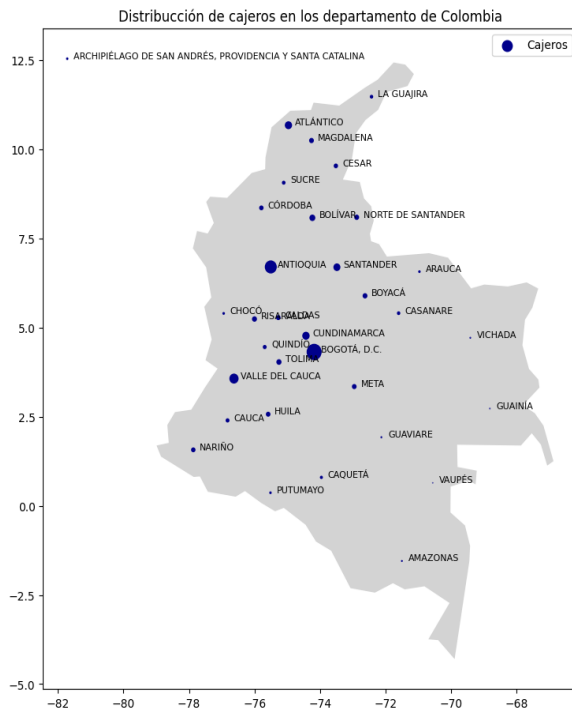
El análisis de la infraestructura de corresponsales bancarios propios evidencia nuevamente el liderazgo de **BANCOLOMBIA S.A.**, que dispone de más de 17.000 puntos de este tipo distribuidos a nivel nacional. Esta amplia red refleja una estrategia de cobertura orientada a fortalecer la presencia de la entidad en zonas urbanas e intermedias, así como a mejorar la accesibilidad de sus servicios financieros.

Por su parte, el **Banco Agrario de Colombia** sobresale por su marcada presencia en áreas rurales y regiones apartadas. Esta cobertura geográfica le permite cumplir un papel determinante en los procesos de inclusión financiera, al ofrecer servicios bancarios en territorios donde la infraestructura tradicional es limitada o inexistente. La presencia de corresponsales en estas zonas no solo amplía el alcance de la banca, sino que también contribuye a cerrar brechas en el acceso a productos y servicios bancarios.



*Ilustración 3. Distribución de oficinas de atención física en Colombia. Fuente: Elaboración propia.*

El análisis de la distribución de oficinas físicas en Colombia muestra una alta concentración en pocas entidades, encabezadas por **BANCOLOMBIA S.A.** con cerca de 1.000 sedes, seguido por el Banco Agrario, Banco de Bogotá, Davivienda y Banco de Occidente, con entre 400 y 800 oficinas. Este liderazgo se refleja también en cajeros automáticos y corresponsales bancarios, consolidando estrategias de cobertura nacional mediante canales físicos. Mientras cajeros y oficinas se concentran en zonas urbanas de alta densidad poblacional, los corresponsales presentan mayor dispersión territorial, desempeñando un papel clave en la inclusión financiera. Estos patrones evidencian desigualdades de acceso y justifican la necesidad de modelos de redistribución optimizada basados en criterios de densidad e impacto transaccional.



*Ilustración 4. Distribución de Canales físicos en el territorio colombiano. Fuente: Elaboración propia.*

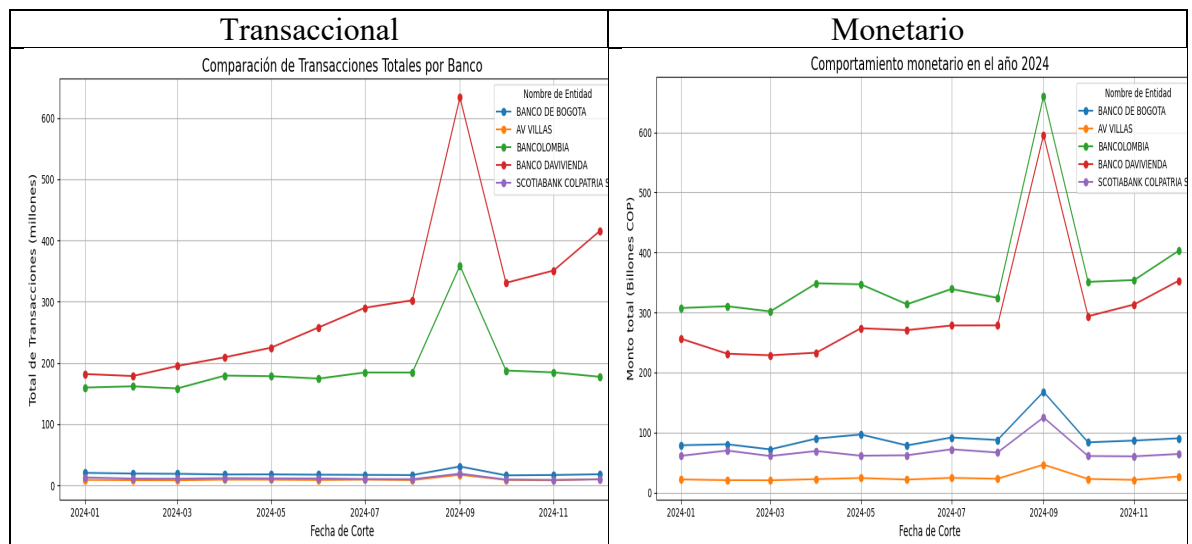
El análisis por tipo de canal evidencia diferencias significativas en la distribución geográfica de la infraestructura bancaria. En el caso de los cajeros automáticos, se observa una alta concentración en departamentos como Antioquia, Bogotá D.C., Valle del Cauca y Atlántico, mientras que su presencia es muy limitada o inexistente en regiones como Amazonas, Vaupés, Guainía o Vichada. Este patrón confirma una tendencia hacia la concentración en áreas urbanas y ciudades intermedias con altos niveles de bancarización.

En contraste, los corresponsales bancarios presentan una dispersión territorial más amplia, con una presencia significativa en departamentos del centro y sur del país, como Tolima, Huila, Meta y Cauca. Esta distribución responde a una estrategia orientada a la inclusión financiera, caracterizada por menores costos de operación y por una capacidad de cobertura que alcanza zonas rurales donde otros canales no están disponibles.

Por su parte, las oficinas físicas se concentran principalmente en Bogotá, Antioquia y Valle del Cauca, con escasa presencia en zonas periféricas o de difícil acceso. Esta distribución refleja la consolidación de la infraestructura presencial en áreas de alta densidad poblacional, dejando brechas en la cobertura de regiones apartadas.

A continuación, se muestra el comportamiento transaccional de los bancos con mayor impacto en Colombia, es sus plataformas digitales.

Participación del canal virtual (Página Web) en el año 2024.



Participación de la aplicación móvil (App) el año 2024.

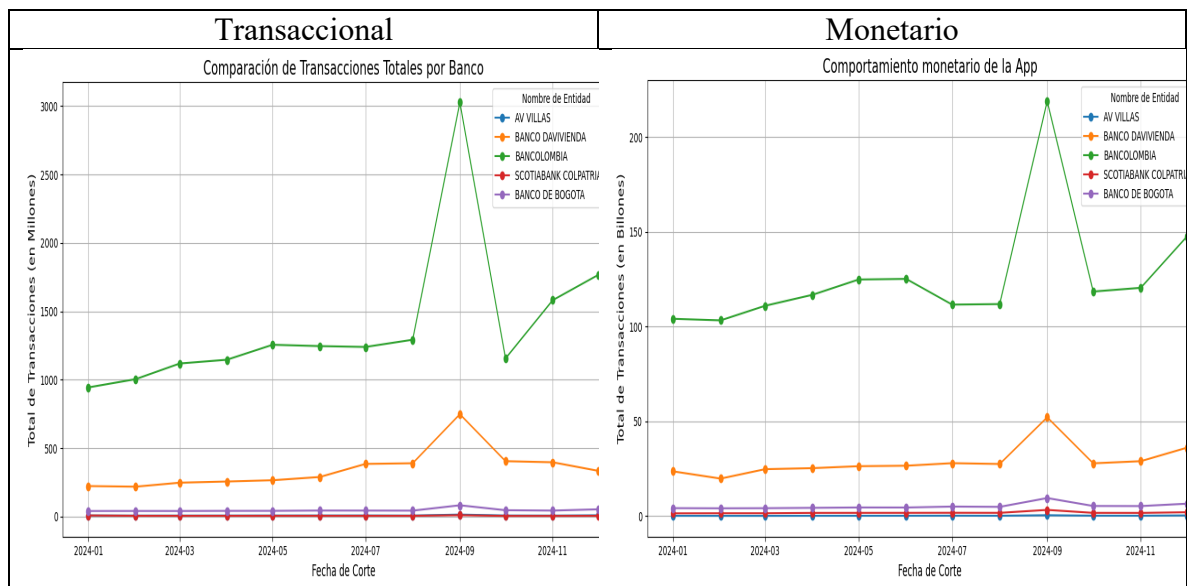


Ilustración 5. Comportamiento transaccional y monetario. Fuente: Elaboración propia.

El análisis del canal virtual a través de la página web muestra que **BANCOLOMBIA** lidera de manera amplia el número de transacciones realizadas durante todo el año. En este canal se identifica un pico abrupto en el mes de septiembre, posiblemente relacionado con un evento económico o la dispersión de subsidios. **DAVIVIENDA** y el **Banco Agrario** mantienen una participación moderada y estable, mientras que **BBVA** y el **Banco de Bogotá** registran volúmenes menores pero constantes. En términos monetarios, la tendencia se repite: **BANCOLOMBIA** encabeza el valor total transado, con un comportamiento que también refleja el aumento puntual en abril, lo que refuerza la hipótesis de operaciones institucionales o pagos de alto valor. **DAVIVIENDA** y el Banco Agrario presentan una estabilidad mayor en este indicador, sin variaciones abruptas.

En el caso del canal digital mediante aplicación móvil, **BANCOLOMBIA** vuelve a ocupar la primera posición, seguido por **DAVIVIENDA**, que muestra un crecimiento leve pero sostenido. El Banco Agrario presenta un nivel bajo de transacciones, lo que podría estar asociado a una menor penetración de su canal móvil. Tanto **BBVA** como el Banco de Bogotá se encuentran significativamente por debajo en uso de aplicaciones, lo que sugiere oportunidades de mejora en la experiencia de usuario y en las estrategias de promoción digital. Desde la perspectiva monetaria, **BANCOLOMBIA** se mantiene como líder absoluto en 2024, tanto en volumen como en valor de transacciones, con un comportamiento estable y un pico en abril vinculado posiblemente a pagos institucionales. **DAVIVIENDA** consolida su posición como segundo actor en el canal móvil, mientras que el Banco Agrario refleja una baja participación, probablemente influenciada por limitaciones tecnológicas o por el perfil de su clientela.

En conjunto, estos resultados evidencian una brecha digital significativa entre las entidades analizadas. Este hallazgo resalta la importancia de fortalecer los canales virtuales, en especial

en aquellas instituciones con un fuerte enfoque rural, como estrategia para avanzar hacia una inclusión financiera más amplia y equitativa.

### **Modificar**

La etapa de transformación de datos incluyó procesos de selección, limpieza, normalización y consolidación de la información necesaria para el modelado. Para ello, se emplearon consultas SQL y scripts desarrollados en Python, lo que permitió integrar de manera eficiente las diversas fuentes transaccionales y demográficas.

Durante esta fase, se implementaron rutinas de transformación que facilitaron la depuración y enriquecimiento de los datos, así como la integración de múltiples orígenes de información. En particular, se cargaron y procesaron los datos históricos correspondientes a los formatos 412 y 413, normalizando sus campos y estableciendo relaciones con las dimensiones geográficas y de canal.

A partir de esta integración, se agregaron las transacciones por tipo de canal físico y digital. Dado que algunas fuentes no disponían de un desglose municipal para los canales digitales, se aplicó un procedimiento de imputación proporcional. Este procedimiento distribuyó las transacciones digitales a nivel municipal tomando como referencia el patrón relativo de uso de los canales físicos en cada municipio, bajo el supuesto metodológico de que la intensidad del uso físico está correlacionada con la demanda potencial de canales digitales en la misma región (según principios de asignación proporcional empleados en estudios de cobertura de servicios, ver Little & Rubin, 2019).

Este criterio se adoptó considerando por la ausencia de datos desagregados de uso digital a nivel municipal, la evidencia empírica de correlaciones positivas entre adopción de canales físicos y digitales en entornos geográficos similares (Baldassarre et al., 2023) y la necesidad de mantener coherencia en el peso relativo de cada municipio dentro del modelo de optimización.

El resultado de este proceso fue la consolidación de variables clave infraestructura física, comportamiento transaccional y población en una tabla única denominada “tbl\_modelo\_opt”, que constituye el insumo principal para el modelo de optimización de cobertura de canales. Esta tabla se generó de manera automática mediante un flujo ETL implementado en SQL, y contiene información estructurada a nivel de entidad, departamento y municipio, incluyendo indicadores como el número de cajeros automáticos, corresponsales y oficinas, el total de habitantes, y los volúmenes y montos transaccionales asociados a cada canal (cajeros automáticos, corresponsales, oficinas físicas, internet, aplicaciones móviles, chatbot y atención telefónica).

Este enfoque metodológico asegura la consistencia y comparabilidad de los datos, y proporciona una base sólida para el desarrollo de los modelos de clustering y optimización que constituyen el núcleo analítico de esta investigación.

A continuación, se relaciona la estructura final de la base de datos de la solución:

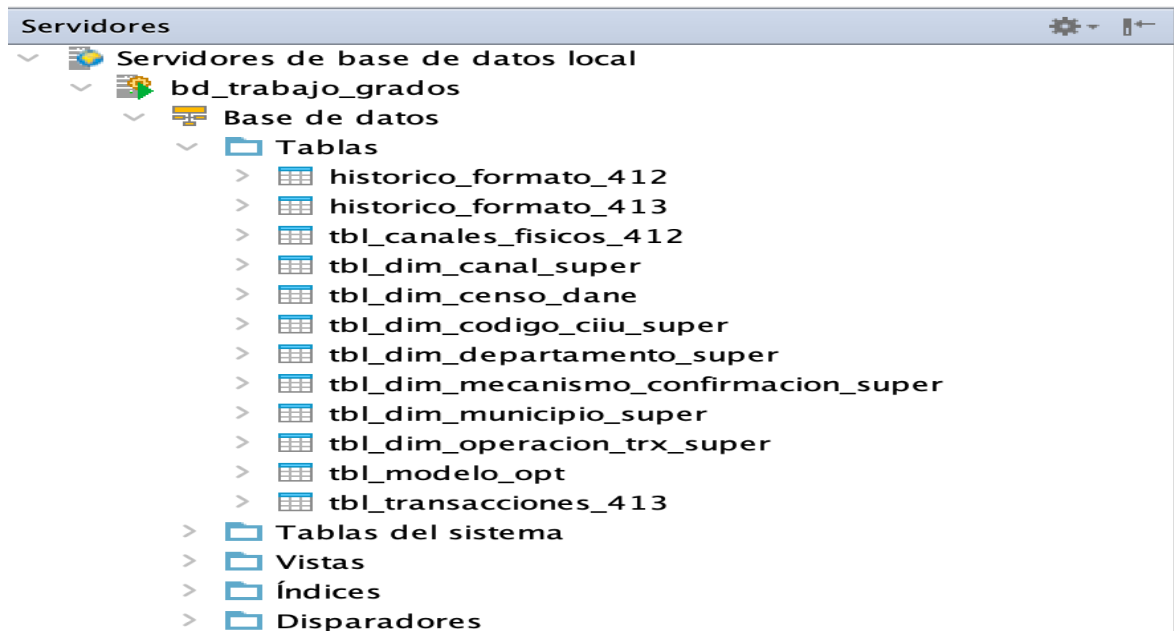


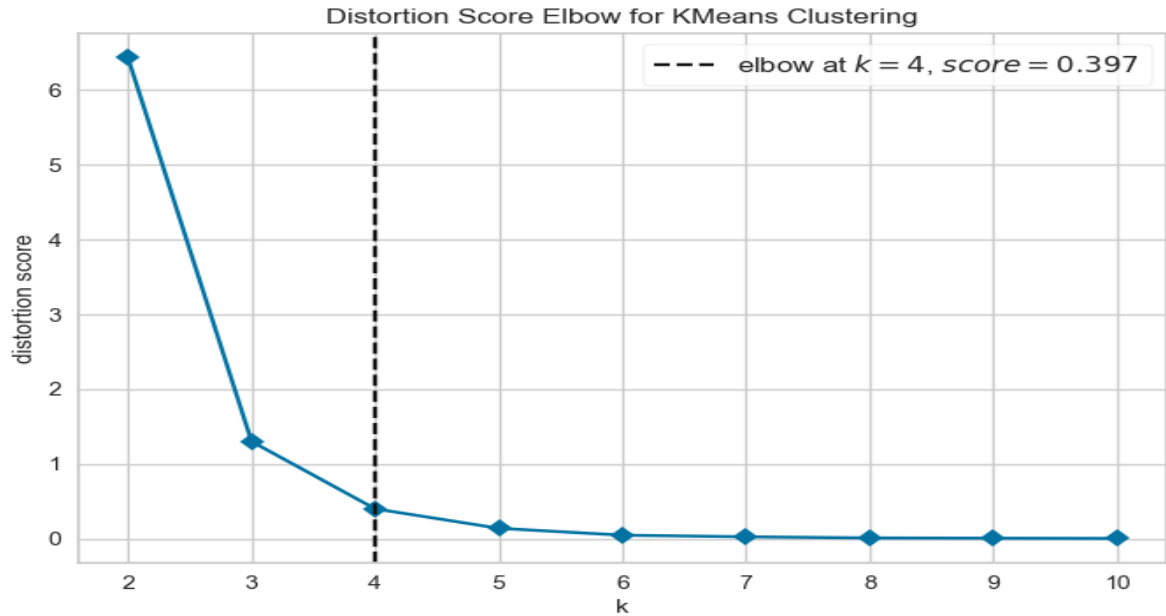
Ilustración 6. Estructura de base de datos local. Fuente: Elaboración propia.

## Modelado

Para profundizar en la solución propuesta, se complementó el modelo de optimización con una segmentación de las entidades bancarias según su infraestructura de canales y comportamiento transaccional. El objetivo fue identificar grupos homogéneos que permitieran entender patrones operativos y la cobertura de distribución de canales físicos actuales. Para lograr esto, se utilizaron variables agregadas por entidad, que incluyeron el total de cajeros, corresponsales y oficinas, así como las transacciones totales por canal, específicamente en cajeros, corresponsales, oficinas, aplicación móvil (app) e internet.

La estandarización de las variables se realizó utilizando StandardScaler, lo que permitió garantizar que todas las variables tuvieran la misma escala y evitar sesgos en el análisis. Posteriormente, se aplicó el algoritmo KMeans para realizar la segmentación, el cual permite formar clústeres con diferentes formas y tamaños, y detectar ruido en los datos. La calidad de la segmentación se evaluó utilizando el índice de silueta y herramientas de visualización de PyCaret, lo que permitió determinar el número óptimo de clústeres y evaluar la robustez de la segmentación.

Para implementar estas técnicas, se utilizaron librerías especializadas en Python, como sklearn para la implementación del algoritmo KMeans y la estandarización de variables, hdbscan para la detección de ruido y clústeres de diferentes densidades, pandas para la manipulación y análisis de los datos, y pycaret.clustering para la evaluación de la calidad de la segmentación y la visualización de los resultados.



*Ilustración 6. Prueba del codo Elbow Point. Fuente: Elaboración propia.*

El análisis del índice de silueta reveló que el número óptimo de clusters es 4, lo que respalda la decisión de utilizar KMeans con 4 grupos. Esta elección ofrece un buen equilibrio entre la simplicidad del modelo y la calidad de agrupamiento, permitiendo una interpretación clara y significativa de los resultados.

El algoritmo KMeans es una técnica de aprendizaje no supervisado ampliamente utilizada para agrupar datos en un número predefinido de grupos (clúster) según su similitud. Su objetivo principal es minimizar la variabilidad intra-clúster y maximizar la separación entre grupos, asignando cada observación al clúster cuyo centroide esté más cercano. De esta manera, KMeans busca identificar patrones y estructuras subyacentes en los datos que no son fácilmente visibles mediante otros métodos.

Matemáticamente, el objetivo del algoritmo KMeans es minimizar la siguiente función de costo, conocida como la suma de los errores cuadrados intra-clúster:

$$J = \sum_{i=1}^n \sum_{j=1}^k w_{ij} \|x_i - \mu_j\|^2$$

donde  $x_i$  representa cada observación,  $\mu_j$  es el centroide del clúster  $j$ ,  $w_{ij}$  es una variable binaria que indica si la observación  $i$  pertenece al clúster  $j$ ,  $n$  es el número de observaciones y  $k$  es el número de clústeres. Al minimizar esta función de costo, KMeans busca encontrar la asignación óptima de las observaciones a los clústeres.

Resultados de la ejecución del modelo con el dataset completo para clasificar las entidades financieras robustas.

*Tabla 2: Resumen de resultados obtenidos con la clusterización.*

<b>Métrica</b>	<b>Valor</b>	<b>Interpretación</b>
<b>Silhouette</b>	<b>0.74</b>	Excelente separación entre clusters. Valores cercanos a 1 indican que los puntos están bien agrupados y claramente diferenciados.
<b>Calinski–Harabasz</b>	360.28	Alta dispersión entre clusters comparada con la dispersión dentro de ellos. Un valor elevado sugiere grupos bien definidos.
<b>Davies–Bouldin</b>	0.27	Bajo solapamiento entre clusters. Mientras más bajo, mejor; este valor indica buena separación.
<b>Homogeneity</b>	0	No se aplica aquí porque no hay etiquetas verdaderas. Esta métrica evalúa si cada cluster contiene solo miembros de una misma clase.
<b>Rand Index</b>	0	También irrelevante en clustering no supervisado puro. Solo se usa cuando hay etiquetas reales para comparar.
<b>Completeness</b>	0	Similar a Homogeneity; sin etiquetas verdaderas no tiene significado aquí.

El modelo de clustering logra una **segmentación robusta y confiable**: el coeficiente de silueta de **0.74** junto con un bajo índice de Davies-Bouldin (**0.27**) evidencian una buena estructura de agrupamiento.

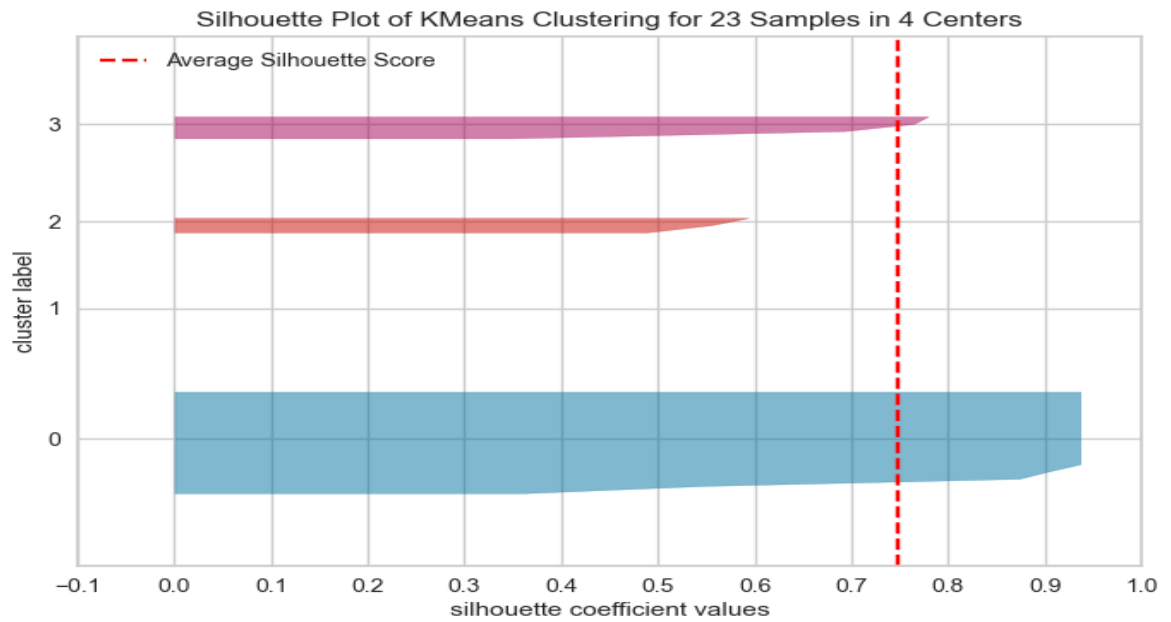


Ilustración 7. Score de silhoutte. Fuente: Elaboración propia.

A continuación, se detalla la distribución de cluster que el algoritmo asigno después de la ejecución.

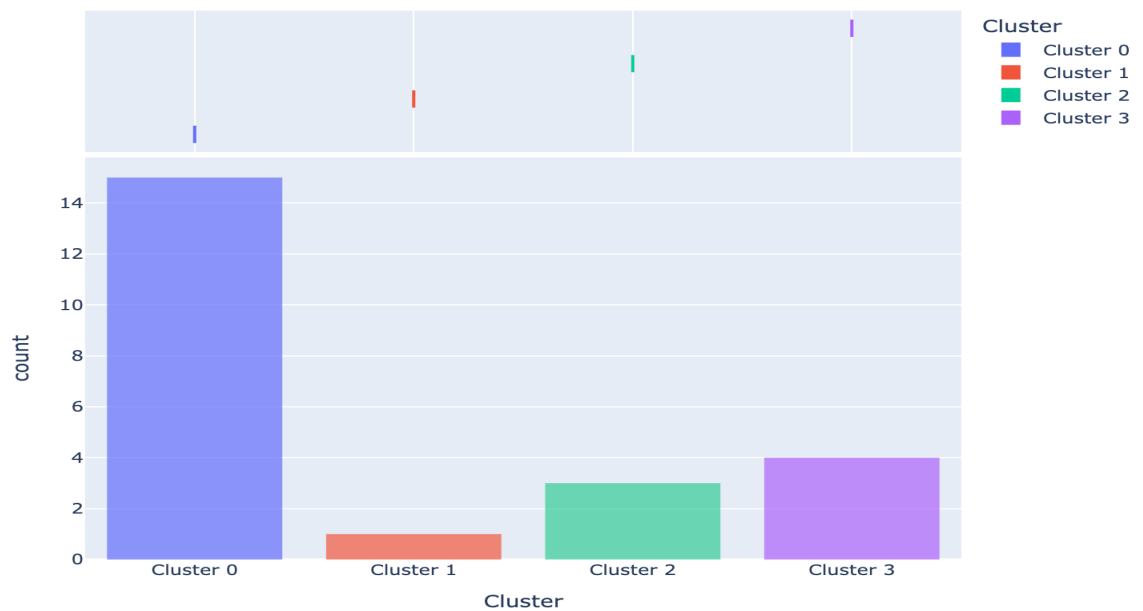


Ilustración 8. Distribución de cluster. Elaboración propia.

La visualización muestra cómo se agrupan las entidades dentro de los cuatro clusters definidos por el modelo KMeans. A partir del conteo de observaciones por grupo, se concluye que:

**Cluster 0** es el más numeroso, con 15 entidades, lo que indica que la mayoría de las instituciones comparten un perfil similar en términos de infraestructura (probablemente tamaño intermedio o distribución estándar de cajeros, oficinas y corresponsales).

**Cluster 1** tiene solo 1 entidad, lo que sugiere un caso atípico o una entidad con características distintivas.

**Cluster 2** agrupa 3 entidades, indicando un subgrupo con patrones moderadamente diferenciados del grupo principal.

**Cluster 3** contiene 4 entidades, que podrían representar bancos con una infraestructura más robusta o bien distribuida.

El desarrollo del modelo de clusterización permitió segmentar de forma efectiva a las entidades financieras colombianas según su infraestructura física (cajeros, oficinas, corresponsales) y el uso de canales digitales (app e internet). Mediante técnicas de análisis exploratorio, normalización y reducción de multicolinealidad, se prepararon los datos para una agrupación robusta y confiable.

*Tabla 3: Detalle de Segmentación de bancos*

nombre	cajeros	corresponsales	oficinas	Cluster
BANCO POPULAR S.A.	14019	1621	5464	Cluster 3
BANCO CAJA SOCIAL S.A.	18275	47666	5280	Cluster 3
SCOTIABANK COLPATRIA S.A.	14899	0	2192	Cluster 3
BANCO COMERCIAL AV VILLAS S.A.	9846	5611	5670	Cluster 3
BANCO DE BOGOTA S.A.	34865	278273	11241	Cluster 2
BBVA COLOMBIA S.A.	32716	0	9838	Cluster 2
BANCO DAVIVIENDA S.A.	49478	0	11577	Cluster 2
BANCOLOMBIA S.A.	113061	542195	22044	Cluster 1
COTRAFA FINANCIERA	0	0	1207	Cluster 0
ITAU COLOMBIA S.A.	4734	5834	1409	Cluster 0
CITIBANK - COLOMBIA	0	0	22	Cluster 0
BANCO GNB SUDAMERIS S.A.	0	0	1836	Cluster 0
BANCO DE OCCIDENTE S.A.	5857	0	5131	Cluster 0
BANCO AGRARIO DE COLOMBIA S.A.	0	25029	17446	Cluster 0
COLTEFINANCIERA S.A.	0	0	616	Cluster 0

BANCAMIA S.A.	0	15353	4912	Cluster 0
BANCO COOMEVA S.A.	784	0	2372	Cluster 0
BANCO FINANDINA S.A.	0	0	242	Cluster 0
BANCO FALABELLA S.A.	1313	0	1896	Cluster 0
BANCO PICHINCHA S.A.	0	0	524	Cluster 0
BANCO SANTANDER	0	9751	20	Cluster 0
BANCO MUNDO MUJER S.A.	0	0	4150	Cluster 0
BANCO J.P. MORGAN COLOMBIA S.A.	0	0	22	Cluster 0

El análisis de clustering utilizando KMeans permitió identificar cuatro grupos de entidades bancarias con características diferenciadoras en términos de infraestructura física y cobertura.

El Cluster 0 está compuesto por 15 entidades que se caracterizan por tener una baja o nula presencia de cajeros automáticos y pocos o ningún corresponsal. Estas entidades pueden representar modelos de operación más tradicionales, especializados o con presencia territorial restringida, como bancos orientados a nichos específicos o entidades extranjeras con baja operación local.

Por otro lado, el Cluster 1 contiene a una sola entidad, Bancolombia S.A., que se destaca por su altísima cantidad de cajeros, corresponsales y oficinas. Esta entidad es un outlier estructural que refleja su rol hegemónico en el sistema financiero nacional, lo que justifica su aislamiento en un grupo exclusivo. El Cluster 2 está compuesto por 3 entidades (Banco de Bogotá, Davivienda y BBVA) que se caracterizan por tener una infraestructura física amplia y una fuerte presencia nacional. Estos bancos tienen modelos de negocio más conservadores, centrados en infraestructura propia, y se diferencian del Cluster 1 por una escala menor.

El Cluster 3 está compuesto por 4 entidades (Popular, Caja Social, Colpatria y AV Villas) que tienen cantidades considerables de cajeros y oficinas, y algunas entidades utilizan corresponsales mientras que otras no. Estos bancos tienen modelos híbridos, cobertura nacional, pero sin llegar a los extremos de los Clusters 1 o 2. Este grupo refleja actores fuertes con presencia regional más que nacional total.

En función de los resultados del clustering, se seleccionaron 4 entidades bancarias que destacan por su cobertura, escala operativa y presencia nacional. Estas entidades son Bancolombia S.A., Banco de Bogotá S.A., Banco Davivienda S.A. y BBVA Colombia S.A. La selección de estas instituciones se fundamenta en su volumen de cajeros automáticos, número de oficinas físicas y, en algunos casos, su red de corresponsales. Su papel protagónico en la estructura del sistema bancario colombiano justifica un análisis más profundo de su comportamiento transaccional y de su estrategia de distribución.

Para la construcción del modelo de optimización, se definieron dos métricas clave que permiten priorizar decisiones con base en características territoriales y operativas. Estas

métricas fueron desarrolladas a partir de información estructural y transaccional por municipio, con énfasis en la relación entre infraestructura física y la población potencial usuaria.

La implementación de técnicas de aprendizaje no supervisado demostró ser una herramienta poderosa para la toma de decisiones estratégicas, y sienta las bases para futuras investigaciones orientadas al análisis dinámico del comportamiento de las entidades en escenarios digitales y presenciales.

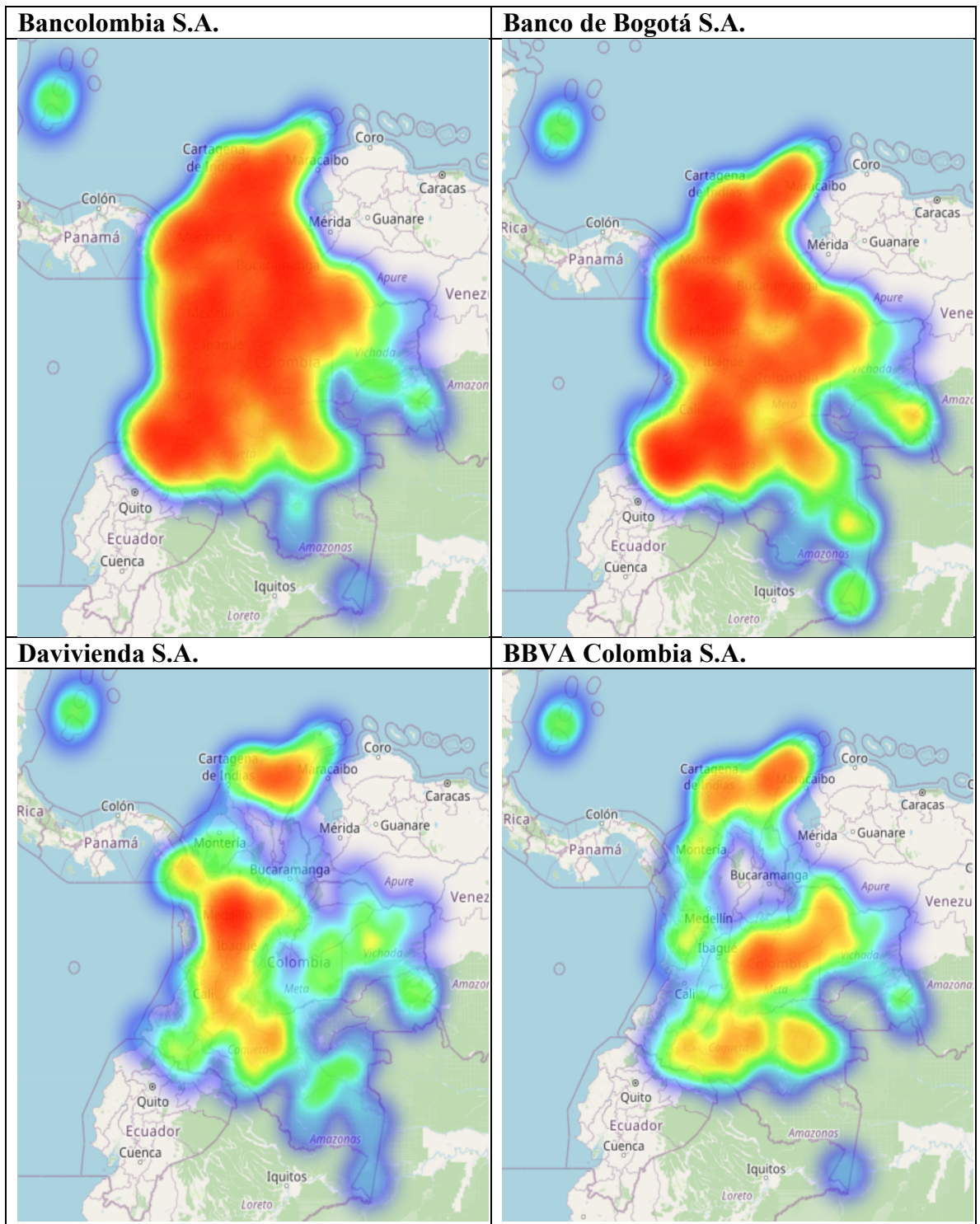
Densidad de Canales Físicos por Habitante Mayor de 18 Años, esta métrica busca capturar el nivel de cobertura de canales presenciales (oficinas, cajeros y corresponsales) en relación con la población adulta del municipio. Su formulación matemática es la siguiente:

$$\text{Densidad}_m = \frac{\text{Cajeros}_m + \text{Corresponsales}_m + \text{Oficinas}_m}{\text{Población mayor 18 años}_m}$$

donde  $m$  representa un municipio, en el numerador total de puntos físicos de atención en el municipio  $m$ , en el denominador, población mayor de edad según el censo del DANE para ese municipio.

Esta métrica permite identificar municipios con **baja cobertura relativa**, que podrían ser priorizados en estrategias de expansión o mejora de acceso.

A continuación, se presentan los mapas de calor de densidad por municipio para cuatro de las principales entidades financieras del país. Cada mapa refleja la concentración de canales físicos (cajeros, corresponsales y oficinas) asociados a la respectiva entidad.



*Ilustración 9. Mapa de calor de la densidad de canales en el territorio colombiano.  
Elaboración propia.*

El análisis de la cobertura por entidad bancaria confirma el predominio de **Bancolombia S.A.**, cuya infraestructura alcanza prácticamente todos los departamentos del país, reflejando su posición como la institución con mayor presencia territorial en Colombia. El **Banco de Bogotá S.A.** también presenta una distribución significativa, aunque más concentrada en regiones centrales y urbanas. No obstante, su posicionamiento en el sur del país resulta relevante, dado que en esta zona pocas entidades logran una participación destacada.

Por su parte, **BBVA** y **Davivienda** exhiben patrones de cobertura similares, caracterizados por una presencia sólida en ciudades intermedias y capitales departamentales, pero con brechas notorias en zonas rurales o en territorios con baja densidad poblacional. Esta distribución refleja una estrategia centrada en áreas urbanas consolidadas, con menor alcance en regiones apartadas.

En este contexto, se elaboró un **score transaccional por municipio**, entendido como un índice ponderado y normalizado que mide la intensidad operativa de cada territorio en función del uso de los distintos canales financieros. Este indicador permite identificar patrones de concentración o dispersión de la actividad bancaria, y constituye una herramienta clave para evaluar el desempeño territorial de cada entidad y orientar estrategias de optimización de cobertura.

$$\text{score}_{m,d} = \frac{0.5 \cdot \frac{\text{trx}_{m,d} - \min(\text{trx}_d)}{\max(\text{trx}_d) - \min(\text{trx}_d)} + 0.5 \cdot \frac{\text{monto}_{m,d} - \min(\text{monto}_d)}{\max(\text{monto}_d) - \min(\text{monto}_d)}}{\sum_{j \in M_d} \left( 0.5 \cdot \frac{\text{trx}_{j,d} - \min(\text{trx}_d)}{\max(\text{trx}_d) - \min(\text{trx}_d)} + 0.5 \cdot \frac{\text{monto}_{j,d} - \min(\text{monto}_d)}{\max(\text{monto}_d) - \min(\text{monto}_d)} \right)}$$

*TRX m, d: Total de transacciones del municipio m en el departamento*

*monto m, d: Total monetario transado en el municipio m en el departamento*

*Md: Conjunto de municipios en el departamento d.*

Este score permite identificar municipios con alta demanda de servicios financieros, lo cual puede complementar el criterio de baja cobertura (densidad) para decisiones de inversión o redistribución.

### **Modelo de Optimización con programación lineal**

Definidas las métricas de decisión “densidad poblacional” y “score transaccional”, se procedió a la implementación de un modelo de optimización lineal entero mixto orientado a redistribuir de manera más eficiente los canales físicos. El propósito central de este modelo fue maximizar la cobertura efectiva considerando dos criterios fundamentales: el volumen de transacciones y la cobertura relativa de cada municipio.

La función objetivo del modelo ponderó la asignación de cajeros automáticos, corresponsales y oficinas con pesos diferenciados, otorgando mayor relevancia a las oficinas físicas debido a su mayor costo operativo y su impacto territorial. Las restricciones establecidas garantizaron la utilización total de los recursos disponibles, evitaron la concentración excesiva en un número reducido de municipios y aseguraron que la distribución final fuera proporcional al perfil combinado de impacto y densidad.

La implementación se realizó en Python mediante la biblioteca PuLP y se resolvió utilizando el solver CBC. Los resultados evidenciaron una asignación más equitativa y estratégica de los recursos físicos, logrando un balance entre eficiencia operativa y cobertura social. Esta redistribución optimizada constituye un insumo relevante para la toma de decisiones en políticas de inclusión financiera y expansión de infraestructura bancaria.

**Métricas:**

Maximizar el valor ponderado de recursos físicos distribuidos (cajeros, corresponsales y oficinas) en función de:

Score de impacto transaccional y Densidad relativa de cobertura.

**Definición de Variables:**

- C<sub>m</sub>*: Numero de cajeros automaticos asignados al municipio *m*
- R<sub>m</sub>*: Numero de corresponsales asignados al municipio *m*
- O<sub>m</sub>*: Numero de oficinas asignadas al municipio *m*
- I<sub>m</sub>*: Score de impacto transaccional del municipio *m*
- D<sub>m</sub>*: Densidad relativa de cobertura del municipio *m*
- P<sub>m</sub>*: Peso proporcional del municipio *m*
- α*: Margen de flexibilidad (ejemplo: *α* = 0.2)

**Función objetivo:**

$$\max \sum_{m \in M} [(0.6 * I_m + 0.4 * D_m) * C_m + (0.8 * I_m + 0.2 * D_m) * R_m + 2.0 * (I_m + D_m) * O_m]$$

**Donde:**

- m*: Índice del conjunto de municipios,
- C<sub>m</sub>*: Numero de cajeros automaticos asignados al municipio *m*,
- R<sub>m</sub>*: Numero de corresponsales asignados al municipio *m*,
- O<sub>m</sub>*: Numero de oficinas asignadas al municipio *m*,
- I<sub>m</sub>*: Score de impacto normalizado del municipio *m*,
- D<sub>m</sub>*: Score de densidad normalizado del municipio *m*

## Restricciones:

- **Recursos totales disponibles** (igualdad):

$$\sum_{m \in M} C_m = C_{\text{total}},$$

$$\sum_{m \in M} R_m = R_{\text{total}},$$

$$\sum_{m \in M} O_m = O_{\text{total}}$$

- **Anti-concentración:**

$$O_m \leq \max(1, 0.60 \times O_{\text{total}})$$

- **Proporcionalidad controlada con margen:**

$$C_m \leq p_m \cdot C_{\text{total}} \cdot (1 + \alpha),$$

$$R_m \geq p_m \cdot R_{\text{total}} \cdot 0.9$$

Donde  $p_m = \frac{I_m + D_m}{\Sigma(I + D)}$  y  $\alpha$  es el margen de tolerancia (0.2)

## Evaluación

La evaluación de la efectividad del modelo de optimización se realizó mediante un caso de prueba aplicado a la distribución física real de **Bancolombia S.A.** en el departamento de Vichada. Este territorio fue seleccionado por sus características de baja densidad poblacional y limitada infraestructura bancaria, lo que lo convierte en un escenario representativo para validar la capacidad del modelo de mejorar la cobertura.

A partir de la información existente, se comparó la configuración actual de los canales físicos con la asignación resultante del modelo optimizado. Esta comparación permitió identificar las diferencias en términos de distribución territorial, eficiencia operativa e impacto potencial en la inclusión financiera. Los resultados obtenidos proporcionaron evidencia cuantitativa sobre la capacidad del modelo para redistribuir recursos de forma más equitativa y alineada con las necesidades poblacionales y transaccionales de la región.

Tabla 3: Datos Observados (Distribución Real)

Municipio	cajeros	corresponsales	oficinas
PUERTO CARREÑO	2	5	1
LA PRIMAVERA	0	6	0
SANTA ROSALÍA	0	6	0
CUMARIBO	0	10	0

Tabla 4: Resultados del Modelo (Asignación Optimizada)

Municipio	Cajeros asignados	Corresponsales asignados	Oficinas asignadas
PUERTO CARREÑO	2	13	1
LA PRIMAVERA	0	4	0
SANTA ROSALÍA	0	0	0
CUMARIBO	0	9	0

El modelo asigna recursos de forma estratégica según el impacto transaccional y la densidad poblacional, manteniendo infraestructura clave (como en Puerto Carreño) y priorizando municipios con mayor necesidad. En zonas de baja demanda, como Santa Rosalía, evita inversiones innecesarias, promoviendo una distribución más eficiente y focalizada.

## Resultados

La segmentación de entidades mediante técnicas de clustering permitió identificar cuatro grupos diferenciados en función de su infraestructura y alcance territorial. El Cluster 0, conformado por quince entidades, agrupa a aquellas con infraestructura limitada y baja cobertura. El Cluster 1 está compuesto únicamente por Bancolombia S.A., cuya presencia dominante lo diferencia significativamente del resto. El Cluster 2 reúne a tres entidades con alta cobertura física, mientras que el Cluster 3, integrado por cuatro instituciones, representa un perfil intermedio en cuanto a capacidad y distribución. Esta clasificación evidencia la existencia de perfiles operativos claramente diferenciados, lo que permite diseñar estrategias de optimización ajustadas a las características de cada grupo.

En cuanto a las métricas de optimización, el análisis de densidad muestra una fuerte concentración de canales físicos en zonas urbanas, con vacíos notorios en regiones periféricas. El score transaccional se encuentra altamente correlacionado con municipios de mayor actividad financiera, confirmando su utilidad para priorizar inversiones en infraestructura. Ambas métricas constituyen herramientas objetivas para la asignación de recursos, al permitir la identificación de áreas estratégicas de expansión o reforzamiento.

La implementación del modelo de optimización confirmó su capacidad para mejorar la distribución de recursos. En el caso de prueba correspondiente al departamento del Vichada para la red de Bancolombia S.A., el modelo mantuvo la infraestructura existente en Puerto Carreño y reforzó la presencia de corresponsales, mientras que en Santa Rosalía no asignó recursos adicionales. Este comportamiento refleja la capacidad del modelo para evitar sobreinversiones y focalizar la cobertura en áreas con mayor demanda o con brechas significativas de acceso.

## CONCLUSIONES

La integración de técnicas de optimización multiobjetivo junto con algoritmos de aprendizaje no supervisado permitió proponer una estrategia robusta para mejorar la distribución de los canales transaccionales físicos en el sistema bancario colombiano. La segmentación mediante clustering evidenció distintos perfiles de principales bancos en el territorio colombiano, lo que habilitó un enfoque más focalizado en el modelo de optimización.

El modelo propuesto, implementado con programación lineal entera, demostró ser eficaz al redistribuir los recursos físicos (cajeros, oficinas, corresponsales) teniendo en cuenta dos factores clave: la densidad poblacional y el impacto transaccional por municipio. Este enfoque permitió:

- Mantener infraestructura donde es crítica,
- Reasignar recursos a municipios con mayor necesidad relativa,
- Evitar inversiones en zonas de baja demanda.

La validación en el caso del departamento del Vichada mostró cómo el modelo puede generar propuestas de redistribución más equilibradas y eficientes. Estas soluciones ofrecen un soporte valioso para la toma de decisiones estratégicas por parte de entidades líderes del sector bancario, al tiempo que contribuyen a cerrar brechas de cobertura e inclusión transaccional en el territorio colombiano.

## ANEXOS

Anexo 1: metadatos técnicos DE HISTORICO\_formato\_413

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
TIPO_ENTIDAD	Número
CODIGO_ENTIDAD	Número
NOMBRE_ENTIDAD	Texto
FECHA_CORTE	Marca de tiempo variable

UNIDAD_CAPTURA	Número
NOMBRE_UNIDAD_CAPTURA	Texto
CODIGO_CANAL_DISTRIBUCION	Número
CODIGO_SERVICIO_TRANSACCIONAL	Número
CODIGO_DEPARTAMENTO	Número
CODIGO_MUNICIPIO	Número
CODIGO_OPERACION	Número
CODIGO_TIPO_PERSONA	Número
CODIGO_CIIU	Número
CODIGO_MECANISMO_CONFIRMACION	Número
CANTIDAD_OPERACIONES	Número
MONTO_OPERACIONES	Número

Anexo 2: metadatos tecnicos DE HISTORICO\_formato\_412

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
UNIDAD_CAPTURA	Número
TIPO_ENTIDAD	Número
NOMBRE_UNIDAD_CAPTURA	Texto
NOMBRE_ENTIDAD	Texto
FECHA_CORTE	Marca de tiempo variable
CODIGO_MUNICIPIO	Número
CODIGO_ENTIDAD	Número
CODIGO_DEPARTAMENTO	Número
CODIGO_CANAL	Número
CANTIDAD_PUNTOS_ATENCION	Número
CANTIDAD_PUNTOS_ANTECION_ACTIVOS	Número

Anexo 3: metadatos tecnicos DE TBL\_dim\_canal\_super

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
codigo	Número
Canal	Texto
Tipo canal	Texto
Descripción	Texto

Anexo 4: metadatos tecnicos DE TBL\_dim\_censo\_dane

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
codigo_dep	Número
departamento	Texto
codigo_municipio	Número
municipio	Texto
rango_edad	Texto
total_habitantes	Número

Anexo 5: metadatos tecnicos DE TBL\_dim\_codigo\_ciu\_super

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
Descripcion	Texto
codigo	Número

Anexo 6: metadatos tecnicos DE TBL\_dim\_departamento\_super

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
cod_departamento	Número
Nombre_departamento	Texto
LATITUD_DEP	Número
LONGITUD_DEP	Número

Anexo 7: metadatos tecnicos de tbl\_dim\_mecanismo\_confirmacion\_super

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
codigo	Número
Mecanismos_confirmacion	Texto

Anexo 7: metadatos tecnicos de tbl\_dim\_municipio\_super

<b>NOMBRE DE LA VARIABLE</b>	<b>TIPO DE CAMPO</b>
codigo_municipio	Número
nombre_mpio	Texto
LATITUD_MUN	Número

**REFERENCIAS**

- Aldajani, M. A., & Alfares, H. K. (2009). Location of banking automatic teller machines based on convolution. *Computers & Industrial Engineering*, 57(4), 1194–1201.
- Andrade Merino, R. V., Andrade Romero, M. P., Dillon Gallegos, J. F., & Andrade Montalvo, F. (2018). Los canales virtuales, una opción para fortalecer las organizaciones financieras solidarias. *Revista mktDescubre - ESPOCH FADE*, 12, 49–56.
- Baldassarre, S., Bruno, G., Piccolo, C., & Ruiz-Hernández, D. (2023). Multi-channel distribution in the banking sector and the branch network restructuring. *Expert Systems With Applications*, 238, 122294.
- Basar, A., Ozkarahan, I., & Bektas, T. (2017). A multi-objective mixed-integer linear programming approach for ATM network design. *Journal of the Operational Research Society*, 68(5), 583–596. <https://doi.org/10.1057/s41274-016-0134-1>
- Byers, R., Yin, S., & Zheng, X. (2012). ATM pricing and location games in the retail banking industry. *Asia-Pacific Journal of Operational Research*, 29(01), Article 1240001.
- Celik Turkoglu, D., & Erol Genevois, M. (2020). A comparative survey of service facility location problems. *Annals of Operations Research*, 292(1), 399–468.
- Delanote, S., Leus, R., & Nobibon, F. T. (2013). Optimization of the annual planning of targeted offers in direct marketing. *Journal of the Operational Research Society*, 64(12), 1770–1779. <https://doi.org/10.1057/jors.2012.165>
- Denstad, J. K., Jørgensen, K. M., & Thorsen, Ø. (2019). A multi-objective optimization model for ATM reallocation considering future customer demand. *Computers & Operations Research*, 103, 180–192. <https://doi.org/10.1016/j.cor.2018.10.008>
- Farahani, R. Z., SteadieSeifi, M., & Asgari, N. (2014). Multiple criteria facility location problems: A survey. *Applied Mathematical Modelling*, 38(5–6), 263–300. <https://doi.org/10.1016/j.apm.2013.06.002>
- Gheitashi, M., Feylizadeh, M. R., & Mohammad Alipour Ahari, R. (2022). A multi-objective mathematical model of financial flows in Omni-Channel distribution systems. *Journal of Intelligent & Fuzzy Systems*. <https://doi.org/10.3233/JIFS-202785>

Guzovski, M., Rudančić, A., & Zgurić Dobrinić, V. (2022). Distribution channels of banking services – Attitudes and preferences of clients. *International Review*, 1–2, 24–30.

Kisore, N. R., & Koteswaraiah, C. B. (2017). Improving ATM coverage area using density based clustering algorithm and voronoi diagrams. *Information Sciences*, 376, 1–20.

Meyer zu Selhausen, H. (1977). Commercial bank balance sheet optimization: A decision model approach. *Journal of Banking and Finance*, 1, 119–142.

Miliotis, P., Diakoulaki, D., & Siskos, Y. (2002). Implementing GIS and multicriteria analysis for the evaluation of retail banking branch expansion strategies. *Journal of the Operational Research Society*, 53(3), 333–342.  
<https://doi.org/10.1057/palgrave.jors.2601326>

Min, H., & Melachrinoudis, E. (2001). The three-hierarchical location-allocation of banking facilities with risk and uncertainty. *International Transactions in Operational Research*, 8(4), 381–401.

Morrison, J., & O'Brien, C. (2001). Bank branch reconfiguration using GIS. *Journal of the Operational Research Society*, 52(9), 1063–1070.  
<https://doi.org/10.1057/palgrave.jors.2601197>

Sharda, R., Delen, D., & Turban, E. (2018). *Big data, intelligence, analytics, and data science: A managerial perspective* (4th ed.). Pearson Education. ISBN: 9780134633282

Tan, P.-N., Steinbach, M., & Kumar, V. (2018). *Introduction to data mining* (2nd ed.). Pearson.

Baldassarre, L., Ricciardi, N., & Vitale, V. (2023). Hierarchical coverage location models for banking networks: Balancing digital and physical channels. *European Journal of Operational Research*, 307(3), 1180–1196. <https://doi.org/10.1016/j.ejor.2022.11.034>

Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781119482260>