

**Planes de expansión del comercio minorista: un acercamiento desde la
analítica, grandes conjuntos de datos e información de transacciones
financieras**

Carlos Antonio Gutiérrez Barrera
cgutie23@eafit.edu.co

Trabajo de grado presentado como requisito para optar al título de Magíster
en Administración Financiera

Asesor: PhD Juan David Ospina Arango

Universidad EAFIT
Escuela de Economía y Finanzas
Maestría en Administración Financiera
Medellín
2018

Resumen

La expansión física de las empresas de comercio minorista es un asunto que podría condicionar la sostenibilidad en el tiempo de estas empresas. Sin embargo, la carencia de información dificulta elegir la localización óptima de los nuevos puntos de venta. A través de la utilización de *big data* financiero se desarrolla una metodología para estimar la compatibilidad demográfica entre las empresas de comercio y los centros comerciales, que además permite medir el desempeño agregado de las empresas de comercio y de sus diferentes establecimientos. Esta metodología tiene potencial para suplir la creciente necesidad de información que tienen los directivos de las empresas para localizar sus nuevos puntos de venta.

Palabras clave: *Big data*, Comercio minorista, Georreferenciación, Planes expansión.

Abstract

The physical expansion of retail companies is a matter that could condition the sustainability over time of these companies. However, the lack of information makes it difficult to choose the optimal location of the new points of sale. Through the use of financial big data a methodology is developed to estimate the demographic compatibility between commercial companies and shopping centers, which also allows measuring the aggregate performance of commercial companies and their different establishments. This methodology has the potential to meet the growing need for information that business managers have to locate their new points of sale.

Key words: *Big data*, Retail, Geolocation, Expansion plans.

Contenido

Introducción.....	5
1 Marco teórico	7
1.1 Expansión del comercio minorista	7
1.2 Definición de <i>big data</i>	9
1.3 Características del <i>big data</i>	10
1.4 Tipos del <i>big data</i>	12
1.5 La expansión física de las empresas de comercio minorista	18
2 Método de solución	22
2.1 Transacciones	22
2.2 Demografía	23
2.3 Establecimientos	24
2.4 Modelo lógico de unión de las fuentes	25
2.5 Método a aplicar	27
3 Resultados	30
Caso 1 – Empresa A	30
Caso 2 – Empresa B.....	35
4 Conclusiones y recomendaciones.....	40
Referencias	41
Anexos	43

Lista de ilustraciones

Ilustración 1. Ejemplo de isócronas.....	8
Ilustración 2. Analítica: dificultad vs valor	16
Ilustración 3. Modelo lógico de datos	25
Ilustración 4. Desempeño empresa A	31
Ilustración 5. Comparación áreas de influencia caso 1	32
Ilustración 6. Desempeño empresa B	35
Ilustración 7. Comparación áreas de influencia caso 2	36

Lista de tablas

Tabla 1 - Resumen variables Caso 1 – Comercio calzado	33
Tabla 2 - KPI's Caso 1 – Comercio calzado.....	34
Tabla 3 - Resumen variables Caso 2 – Comercio vestuario	37
Tabla 4 - KPI Caso 2 – Comercio vestuario	38

Introducción

En algunos casos las decisiones de expansión de las compañías de comercio minorista no son tomadas con toda la información necesaria y, por lo tanto, estas decisiones pueden terminar encauzando a las empresas en un rumbo diferente al que originalmente se había planeado. Un ejemplo de esto han sido los planes fallidos de expansión de empresas de comercio minorista que hemos visto en el país en los últimos años, en casos como Ripley, la Polar, compañías que cerraron su operación en el país o ejemplos no tan drásticos como almacenes que han debido ser cerrados tras no cumplir con las expectativas de ventas inicialmente previstas. Es aquí donde el acceso a buenas fuentes de información toma relevancia, ya que esta permite disminuir la incertidumbre de los posibles resultados a los que puede llevar las decisiones que se toman.

A cierre de 2017 se estima que en Colombia existe algo más de 220 centros comerciales, tanto en ciudades principales como intermedias. Los actuales proyectos de construcción sumarán cerca de 20 centros comerciales en 2018 (La República, 2018). Las diferentes empresas del comercio minorista han visto el aumento de nuevas aperturas de centros comerciales como una oportunidad para desarrollar sus planes de expansión y, por lo tanto, han decidido ubicar sus tiendas en estos nuevos proyectos. Sin embargo, como se mencionó anteriormente, estas estrategias no siempre han resultado exitosas. Una de las posibles razones puede haber sido la falta de información para la toma de decisión del lugar de ubicación óptima de las nuevas tiendas, ya que actualmente los directivos de las empresas de comercio minorista no siempre cuentan con información que les permita determinar el perfil

demográfico de sus posibles compradores y si este será compatible con el perfil demográfico del centro comercial en donde buscan expandirse.

Actualmente las entidades financieras capturan y almacenan una gran cantidad de datos de sus clientes. Estos datos van desde la información demográfica que estos les comparten en la vinculación y su actualización al solicitar un nuevo producto, hasta, por ejemplo, todas las compras que realizan con sus tarjetas débito y crédito.

El objetivo de esta investigación es utilizar información demográfica y transaccional para desarrollar una metodología que permita determinar la compatibilidad entre los perfiles demográficos de una empresa de comercio minorista y los compradores del centro comercial donde planea realizar su expansión, además de medir el comportamiento agregado y de los diferentes establecimientos de esta empresa de comercio. El periodo de análisis comprende los años 2014 a 2017.

Al final de la investigación se analizan dos empresas de los sectores comercialización de vestuario y comercialización de calzado, en donde se estiman los perfiles demográficos de los compradores de la empresa, antes de que realicen una nueva apertura y se comparan con el perfil demográfico de los lugares donde planean expandirse.

Este documento se divide en las siguientes secciones: primero, se realiza un resumen de literatura, luego se da la especificación del método de solución y posteriormente se exponen los resultados.

1 Marco teórico

1.1 Expansión del comercio minorista

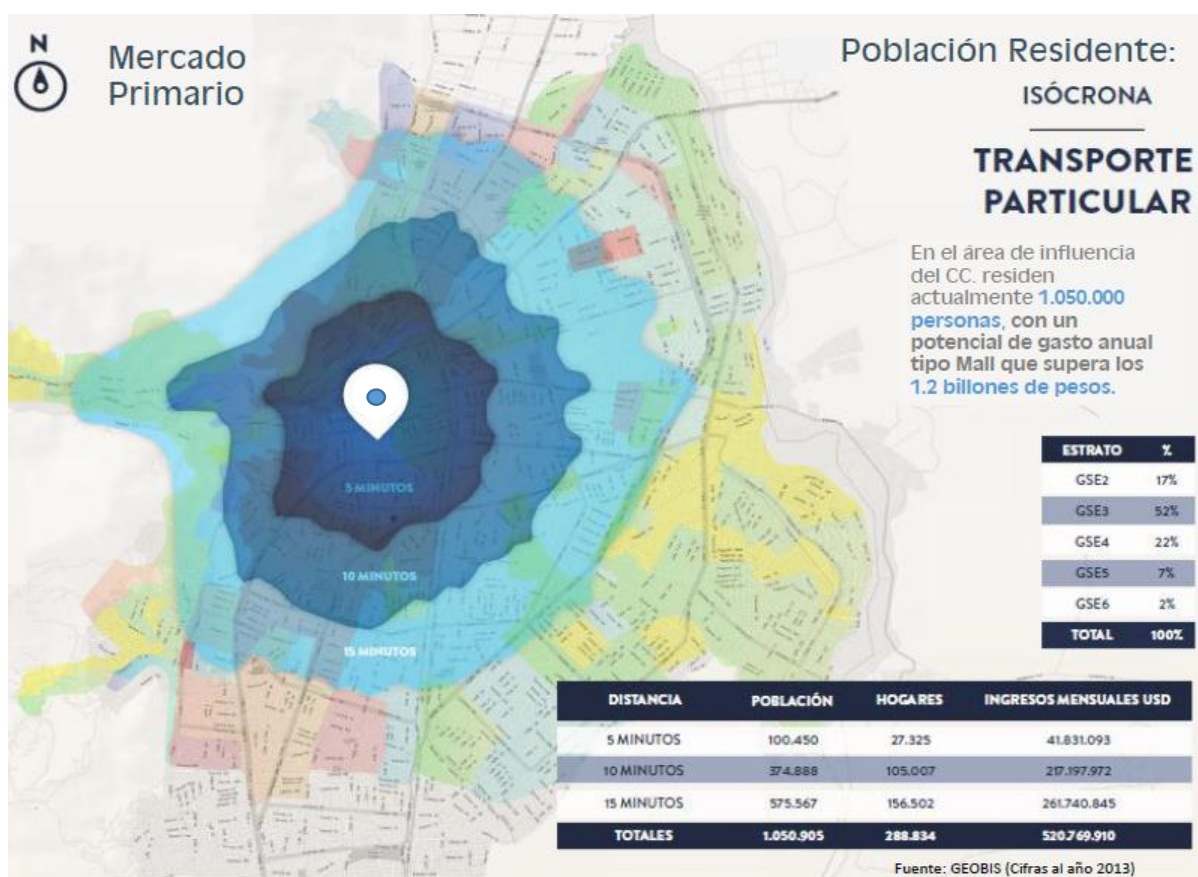
Las estrategias de expansión de las empresas son un tema de vital importancia para sus directivos, según Edith Penrose, la empresa nunca alcanzará la situación de equilibrio. Por tanto, el proceso de expansión es fundamental para la supervivencia de la empresa en el mercado (Penrose, 1959). Sin embargo, el proceso de toma de decisión de apertura de nuevos locales comerciales es complejo, principalmente debido a la incertidumbre que puede existir sobre los hábitos de consumo de los consumidores, ya que usualmente no se cuenta con esta información. Gracias a ello, esta investigación podía ser de gran utilidad para los directivos de las empresas de comercio minorista en Colombia.

En el proceso de toma de decisiones, una parte fundamental es la recolección y el análisis de información de manera que la decisión se pueda tomar con la mayor certidumbre posible. Esta situación aplica también para las empresas de comercio minorista, ya que para cualquiera de estas es de vital importancia tener información que le permita conocer los hábitos de consumo de su cliente, qué le gusta comprar, qué lugares frecuenta, cuánto compra, entre otros aspectos, con el fin de poder diseñar productos y/o servicios más acordes a sus necesidades y poder seleccionar la óptima ubicación de los almacenes.

En Colombia, los constructores de centros comerciales contratan estudios de movilidad, llamados Isócronas, alrededor del lugar donde planean desarrollar un proyecto, con el fin de analizar el mercado potencial y, por lo tanto, argumentar la viabilidad del proyecto. La Fuente: Geobis, es una extracción de un caso real de un proyecto de construcción de un centro comercial presentado a una entidad financiera para su financiación. Como se puede apreciar

en la Fuente: Geobis, , este estudio realiza áreas de influencia potencial del proyecto, en función de la cantidad de personas que viven en ciertos rangos de desplazamiento, en este caso, en transporte particular, de cinco, diez y quince minutos. Una vez tienen el número total de personas que residen en esta área de influencia, calculan las ventas potenciales del centro comercial como una proporción de los ingresos de la posible población.

Ilustración 1. Ejemplo de isócronas



Fuente: Geobis, 2013.

El potencial de mercado de estos estudios se fundamentan en tres supuestos; en primer lugar, las personas que frecuentan un centro comercial lo hacen porque este es cercano a su lugar de residencia; en segundo lugar, los ingresos de las personas están asociados y condicionados

al estrato socioeconómico en el que viven; por último, a todas las personas de un mismo estrato socioeconómico se les asocia el mismo ingreso.

Los directivos de las empresas de comercio minorista no siempre cuentan con información que les permita validar estos análisis, ya que estudios de hábitos de consumo y demográfica de sus compradores deben ser adquiridos a un tercero. Por lo tanto, en los casos donde los directivos de las empresas de comercio minorista no cuenten con la información de sus consumidores, se dificulta determinar si existe compatibilidad entre las características demográficas y comportamentales de los clientes de su empresa y los clientes potenciales que van a frecuentar el centro comercial.

1.2 Definición de *big data*

Big data es un concepto que se ha vuelto cada vez más popular en los últimos años, sin embargo, aún genera bastante confusión, ya que no existe una definición rigurosa del término. Ya en 1997, científicos de la NASA publicaron un documento en donde describieron los problemas de visualización que tenían: “los conjuntos de datos son generalmente bastante grandes, limitando las capacidades de la memoria principal, el disco local e incluso el disco remoto. Llamamos a esto el problema del *big data*. Cuando los conjuntos de datos no caben en la memoria principal (en el núcleo), o cuando no encajan ni siquiera en el disco local, la solución más común es adquirir más recursos” (Press, 2014, p. 1).

En 2008, varios científicos informáticos estadounidenses popularizaron el término, y aunque no dieron una definición exacta del *big data*, pronosticaron que “la computación de *big data*

transformará las actividades de las empresas, investigaciones científicas, prácticas médicas y las operaciones de inteligencia y defensa de la nación” (Press, 2014, p. 1).

Posteriormente, IBM habla del término así: “los avances de la tecnología han abierto las puertas hacia un nuevo enfoque de análisis y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos que tomaría demasiado tiempo y sería muy costoso tratarlos de manera tradicional para su análisis” (Barranco, 2012, p. 1).

En 2013 el *Oxford English Dictionary* (OED) incluyó el término *big data* dentro de sus publicaciones, lo que implicó que por primera vez se tuviera una definición oficial: “datos de un tamaño muy grande, generalmente en la medida en que su manipulación y gestión presentan importantes desafíos logísticos” (Press, 2013, p. 1).

La firma de consultoría Gartner define el concepto como: “*big data* son activos de información de gran volumen, alta velocidad y / o alta variedad que demandan formas innovadoras de procesamiento de información rentables que permiten un mejor conocimiento, toma de decisiones y automatización de procesos” (Gartner, 2014, p. 1), en donde introduce un concepto importante que son las tres V’s del *big data*, las cuales hacen referencia al volumen, la velocidad y la variedad de los datos. Sin embargo, algunas empresas han ampliado aún más la definición al adicionar otras dos características (V’s), el valor del dato y la veracidad (BBVA, 2017).

1.3 Características del *big data*

Volumen: el término *big data* hace referencia a grandes volúmenes de datos, los cuales no pueden ser procesados con métodos tradicionales. Esto se da, gracias a que actualmente los

seres humanos estamos constantemente generando grandes cantidades de información a través de las interacciones que tenemos diariamente con la tecnología, como por ejemplo las transacciones financieras, las compras que se realizan en diferentes comercios con medios de pago digitales, las redes sociales que usamos, entre otros.

Variación: hace referencia a que en la actualidad se almacenan datos de fuentes convencionales y no convencionales, como redes sociales, imágenes, videos, dispositivos electrónicos o sensores que se encuentran conectados, entre otros.

Velocidad: se refiere a la posibilidad que genera el *big data* de crear, almacenar y procesar datos en tiempo real, lo cual no era posible con los métodos tradicionales. Desde el punto de vista de una empresa, el *big data* le permite tomar decisiones con más información y mayor rapidez.

Veracidad: “cuando hablamos de veracidad nos referimos a la incertidumbre de los datos, es decir, al grado de fiabilidad de la información recibida. Es necesario invertir tiempo para conseguir datos de calidad, aplicando soluciones y métodos que puedan eliminar datos imprevisibles que puedan surgir, como datos económicos y el comportamiento de los consumidores que puedan influir en las decisiones de compra” (IC, 2016, p. 1).

Valor: este concepto no se refiere a que los datos por sí solos tengan valor, sino a cuando los datos se transforman en información; esta a su vez se convierte en conocimiento, y este en acciones o en decisiones que después le generan valor a las compañías que los poseen.

1.4 Tipos del *big data*

Como se mencionó anteriormente, el *big data* supone la generación de grandes cantidades de datos, los cuales pueden ser estructurados, como la información contable o financiera de una empresa, la cual puede ser procesada fácilmente con métodos convencionales de tratamiento de datos y no estructurados como correos electrónicos, videos, publicaciones en redes sociales, entre otros, los cuales representan cerca del 80% del total de datos que son generados constantemente. De acuerdo con IBM (Barranco, 2012), los datos pueden clasificarse de acuerdo a cinco grandes fuentes:

- Web y redes sociales: contenido web e información de redes sociales como *Facebook*, *Twitter*, *LikedIn*, entre otras, así como búsqueda en *Google*, clics en vínculos, etc.
- Entre máquinas (M2M): esto se refiere a la interconexión entre máquinas, las cuales a través de sensores capturan eventos, tales como velocidad, temperatura, variables meteorológicas, calidad del aire, entre otros, los cuales son transmitidos de forma alámbrica o inalámbrica para ser traducidos en datos.
- Transacciones: se refiere a las interacciones de las personas, las cuales pueden ser transacciones monetarias y no monetarias. Dentro de las monetarias están las compras realizadas con medios de pago digitales (tarjetas débito y crédito). Incluye transacciones monetarias, registros de facturación, en telecomunicaciones registros detallados de las llamadas (CDR), etc. Estos datos transaccionales están disponibles en formatos tanto semiestructurados como no estructurados.

- Biometría: información biométrica en la que se incluye huellas digitales, escaneo de la retina, reconocimiento facial, genética, etc.
- Generado por personas: las personas generan diversas cantidades de datos como la información de los *call center*, notas de voz, correos electrónicos, documentos electrónicos, estudios médicos, etc.

Uno de los primeros casos conocidos de utilización de *big data* en empresas es el de Amazon, empresa que pudo evolucionar estratégicamente gracias a que le dio uso y valor a los datos que almacenaba diariamente de sus clientes. Amazon comenzó como una empresa de comercio electrónico enfocada en la venta de libros, pero rápidamente entendió que podía obtener información de los clics que sus consumidores realizaban con el *mouse* mientras estos compraban libros en la página web. La empresa logró capturar y analizar el historial de navegación de sus clientes, incluidas las palabras que fueron buscadas, los libros comprados, los que no se compraron, los que se colocaron en las listas de deseos y la cantidad de tiempo que se vieron los artículos. Esto condujo a Amazon a tener una mayor y mejor selección de libros, de acuerdo a las preferencias de sus compradores, un mejor sistema de recomendación y a expandirse a otros segmentos de mercado. En la actualidad Amazon vende prácticamente cualquier producto en su sitio web, el cual incluye productos electrónicos, equipos deportivos, indumentaria e incluso materiales de construcción (Mazzei & Noble, 2017).

Bernard Marr, menciona en un artículo publicado en Forbes en el año 2015 que: “*big data* no es una moda pasajera. Estamos justo al comienzo de una revolución que tocará todos los negocios y todas las vidas en este planeta” (Marr 2015, p. 1). En el mismo artículo menciona que el término *big data* es un concepto que muchas personas están tratando como algo que

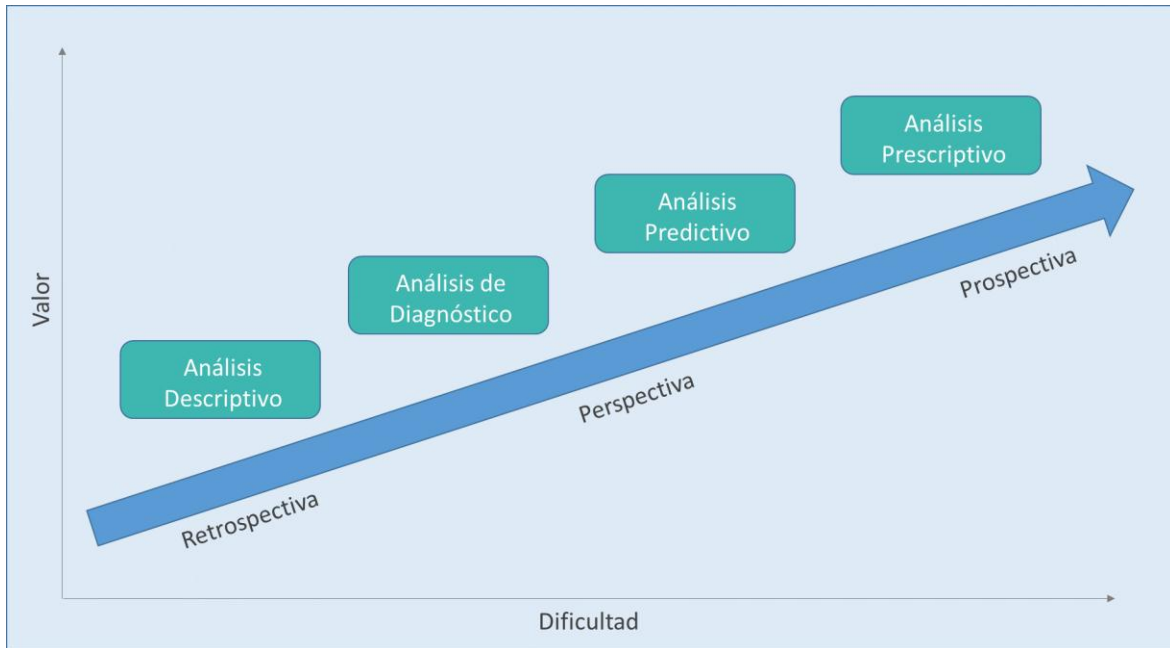
pueden elegir ignorar, cuando en realidad cambiará la forma en la que vivimos. Para ilustrar su punto, el autor menciona datos curiosos que ilustran la dimensión e importancia del término; a continuación algunos de ellos:

- Los volúmenes de datos están explotando, se han creado más datos en los últimos dos años que en toda la historia de la raza humana.
- Cada segundo creamos nuevos datos. Por ejemplo, realizamos 40.000 consultas de búsqueda por segundo (solo en Google), lo que lo convierte en 3,5 búsquedas por día y 1,2 billones de búsquedas por año.
- Los usuarios de Facebook envían un promedio de 31,25 millones de mensajes y ven 2,77 millones de videos cada minuto.
- Estamos viendo un crecimiento masivo en los datos de video y fotos, donde cada minuto hasta 300 horas de video se cargan solo en YouTube.
- Para 2020, tendremos más de 6.100 millones de usuarios de teléfonos inteligentes en todo el mundo (superando las suscripciones de telefonía fija).
- Dentro de cinco años habrá más de 50 mil millones de dispositivos inteligentes conectados en el mundo, todos desarrollados para recopilar, analizar y compartir datos, los cuales superarán ampliamente el número de personas en el planeta.
- Las empresas de comercio minorista que aprovechan toda el potencial de *big data* podrían aumentar sus márgenes operativos hasta en un 60%.
- De momento solo se analiza y utiliza el 0,5% de todos los datos que existen, por lo que debemos considerar el potencial que este mercado presenta.

La información se ha convertido en uno activo importante en la actualidad, gracias a que la tecnología ha permitido y llevado a que todos los agentes económicos (personas, empresas y gobierno) estén constantemente creando huellas digitales, debido a las interacciones que realizan entre ellos, como por ejemplo dispositivos electrónicos, búsquedas en internet, interacción en redes sociales, compras, pagos, recaudos, entre otros. *Big data* permite a las empresas que recopilan este tipo de información, poder analizarla con el fin de encontrar patrones de comportamiento, los cuales le ayudan a darle respuesta a preguntas que antes no habían podido ser contestadas, llevando a estas empresas a encontrar nuevas oportunidades de negocio, así como optimizar los procesos de toma de decisiones, lo que se traduce en una ventaja competitiva en la industria en la que se encuentre. “En general, para cualquier industria, el *big data* puede servir de apoyo (i) en los procesos complejos de toma de decisiones, (ii) en extraer “inteligencia” de la combinación de diversas fuentes y (iii) en el trabajo con grandes volúmenes de información” (Asobancaria, 2018, p. 2).

La firma de analítica *Datameer* (Datameer, 2017), menciona que la analítica de *big data* (*Big Data Analytics*) busca responder a una nueva serie de preguntas de diagnóstico sobre su negocio, mediante el uso de más datos y análisis sofisticados para entregar resultados procesables a sus equipos de negocios. La analítica de *big data* permite explorar preguntas más profundas, algunas de las cuales anteriormente era imposible responder, para revelar un nuevo nivel de conocimiento e identificar los pasos necesarios para mejorar el rendimiento de los negocios. *Big data* tiene el potencial de cambiar los procesos de toma de decisiones de las empresas, ya que les permite a estas dar pasos más allá del simple análisis descriptivo. Fuente:, inspirada en una gráfica elaborada por la firma Gartner resume esta situación.

Ilustración 2. Analítica: dificultad vs valor



Fuente: Gartner, 2013.

Análisis descriptivo: muchas empresas han pasado años generando análisis descriptivos, respondiendo lo que sucedió. Esta información es valiosa, pero solo proporciona una vista general del rendimiento del negocio.

Análisis de diagnóstico: la mayoría de las empresas comienzan a aplicar la analítica de *big data* para responder preguntas de cómo y por qué sucedió algo. El análisis de diagnóstico con *big data* ayuda de dos maneras: (i) los datos adicionales generados por la era digital eliminan los puntos ciegos analíticos, y (ii) el cómo y el por qué ofrecen información que identifica las acciones a realizar.

Análisis predictivo: responder las preguntas de cómo y por qué revelan patrones específicos para identificar cuándo las acciones o los resultados están a punto de ocurrir. Los análisis

predictivos se basan en el análisis de diagnóstico para buscar estos patrones y ver qué sucederá.

Análisis prescriptivo: sobre la base de los análisis previos, se proporcionan recomendaciones (prescriptivas) específicas. Estas responden a la pregunta: ¿qué debería hacer si pasa “x”?

En síntesis, analítica de *big data* permite pasar del análisis descriptivo tradicional, a realizar diagnósticos y predicciones, lo cual le genera valor agregado a las empresas, ya que les responde nuevas preguntas, lo que las llevará a producir más y mejores resultados.

Como menciona Marr, *big data* revolucionará todos los negocios e incluso la vida de todas las personas, sin embargo, hay algunos sectores económicos que por sus características son más propensos a adoptar estas tecnologías. Algunos sectores son los siguientes:

Sector financiero: desde el punto de vista del negocio, la aplicación de *big data* y la analítica es enorme y se focaliza en diversos procesos, como por ejemplo segmentación, fidelización, gestión de riesgos, mayor gestión de la omnicanalidad, cumplimiento, solvencia, gestión de fraude, atención al cliente, prevención de lavado de activos, entre otros (Asobancaria, 2018).

Con la competencia cada vez mayor, la banca podría utilizar la analítica de *big data* para identificar nuevos productos y servicios para llegar a nuevos clientes y crear nuevas fuentes de ingresos. Las entidades financieras capturan y almacenan una gran cantidad de información de sus clientes. En el caso de las personas naturales, estas entidades almacenan toda la información entregada por los usuarios, como lo es su información demográfica, sin embargo, las entidades financieras también capturan y almacenan el historial de información transaccional (compras con medios de pagos digitales), o en otras palabras, las huellas

digitales de las personas naturales. En el caso de los clientes empresariales, las entidades financieras cuentan con la información que estos le proveen, como por ejemplo los estados financieros de cierto periodo, o los establecimientos con los que cuenta y donde están ubicados, entre otros.

Salud: los gobiernos y grandes grupos de atención médica pueden utilizar *big data* para el análisis de salud de la población y con este poder predecir epidemias, curar enfermedades, mejorar la calidad de vida y evitar muertes prevenibles. Los servicios de atención médica tradicionales están utilizando la aplicación de datos de dispositivos médicos, notas médicas, resultados de laboratorio y más, para impulsar el análisis de *big data* a mejorar la atención del paciente y aumentar la eficiencia (Datameer, 2017).

Comercio: las empresas de comercio pueden aplicar *big data* para comprender por qué los consumidores están comprando y luego adaptar sus respuestas para influir en las ventas. Con ciclos de vida de los productos cada vez más rápidos y las operaciones cada vez más complejas, las empresas de comercio pueden utilizar *big data* para comprender las cadenas de suministro y la distribución de productos con el fin de reducir sus costos.

1.5 La expansión física de las empresas de comercio minorista

Una de las preguntas que las empresas de comercio han empezado a responder con el uso de *big data* es: ¿dónde ubicar mi próxima tienda?, ya que las tiendas físicas aún representan el 94% de las ventas minoristas y, por lo tanto, la selección de la ubicación de estas tiendas juega un papel descomunal en la determinación de si sus cadenas tienen éxito o fracasan (Thau, 2014). En un artículo, publicado por la revista Forbes, Simon Thompson, director

comercial de Esri, compartió cómo ArcGIS, su plataforma tecnológica para visualizar datos en forma de mapas, está ayudando a algunas empresas de comercio a ajustar la ubicación de sus tiendas con el objetivo de generar más tráfico y aumentar las ventas. Algunos ejemplos que el autor menciona en el artículo son los siguientes:

- **PETCO:** comenzaron a utilizar el software de Esri para determinar los riesgos potenciales asociados con la construcción de nuevas tiendas. Cuantas más tiendas abra Petco, mayor será el riesgo de que la compañía construya una tienda en una ubicación no deseada. Por ejemplo, una tienda en una comunidad que tiene una baja concentración de dueños de mascotas dañaría las ventas. Además, abrir una tienda cercana a la tienda de un competidor también impediría el buen desempeño de la tienda.
- **Starbucks:** en 2008, el CEO de Starbucks, Howard Schultz, se vio obligado a salir de su retiro para cerrar cientos de tiendas y reconsiderar el plan de crecimiento estratégico de la compañía. Esta vez, Starbucks adoptó un enfoque más disciplinado e impulsado por el software de Esri pudo analizar datos demográficos y de ubicación de sus consumidores, para determinar el mejor lugar para abrir nuevas tiendas de la cadena sin perjudicar las ventas en otras ubicaciones de las ya existentes. El software también está ayudando a determinar dónde deben colocarse las siguientes, más de 1.500 tiendas, no solo para ayudar a la compañía a expandirse, sino también para generar ingresos para los nuevos desarrollos de la compañía.

Esri, fundada en 1969, es una empresa dedicada a lo que ellos mismos han denominado “la ciencia del dónde (The Science of Where)”. Esta empresa desarrolló el software de

geolocalización ArcGIS, lo cual los ha posicionado como líderes mundiales en sistemas de información geográfica, sin embargo, quizá una de sus ventajas competitivas más importantes es que han complementado la información de geográfica con información demográfica de los consumidores (edad, género, ocupación, nivel de ingresos), así como características comportamentales de las personas (hábitos de consumo, patrones de desplazamiento, tenencia de mascotas, entre otros). Actualmente Esri ofrece soluciones para diferentes industrias, como sistema financiero, gobierno, educación, salud, recursos naturales y comercio, en donde para este último, una de sus principales propuesta es la de selección óptima de ubicación de tiendas físicas.

Para una selección óptima de ubicación de tienda físicas, Esri propone una serie de pasos o preguntas:

- ¿Quién es el cliente actual de la marca?: comprender a sus clientes y su comportamiento socioeconómico y de compras es esencial para tomar buenas decisiones comerciales.
- Posicione sus nuevas tiendas para atraer nuevos clientes: una ubicación comercial no óptima puede hacer o deshacer un negocio, sin importar cuán bueno sea el servicio o producto.
- Minimizar la canibalización entre canales y promover el crecimiento general de la empresa.
- Generar el mayor rendimiento de las inversiones en la apertura de nuevas tiendas.

- Establecer su marca en mercados con altas concentraciones de sus clientes objetivo, antes de que sus competidores lo hagan.

Dado lo anterior, existe un área de oportunidad para las entidades financieras en Colombia, ya que estas capturan y almacenan información con la cual se podrían replicar ejercicios como los que realiza Esri. Las entidades financieras capturan y almacenan todos los registros que dejan las personas naturales al realizar compras con sus medios de pago digitales, con lo cual podrían tener aproximaciones de los hábitos de consumo de sus clientes, ya que podrían saber qué están comprando cada uno de sus clientes, dónde lo están comprando, cuánto están gastando en cada transacción, entre otras cosas. Igualmente, las entidades financieras podrían agregar las compras que realizan sus clientes persona natural en los diferentes establecimientos que visitan y conociendo a quién pertenecen estos establecimientos, podrían ver de forma agregada todas las compras que realizan estos en sus clientes empresariales. En resumen, las entidades financieras podrían tener la información que los directivos de las empresas de comercio minorista necesitan y las características demográficas y comportamentales de sus clientes y, por lo tanto, con esto, podrían determinar los perfiles de los compradores de las diferentes empresas de comercio.

En esta investigación se plantea realizar una aproximación a la oportunidad anteriormente descrita, utilizando la información que hoy en día capturan y almacenan las entidades financieras, en tanto es posible determinar el perfil de los compradores de una empresa de comercio determinada y, por lo tanto, también se podrían evaluar si el lugar donde planea abrir su próxima tienda es compatible con ese perfil de su comprador promedio. Adicionalmente, también se podría evaluar y determinar el comportamiento de cada uno de

los establecimiento y el nivel agregado para la empresa de comercio en función de las ventas estimadas.

2 Método de solución

Para esta investigación, se trabaja con la información capturada y almacenada por una entidad financiera. Esta información es el resultado de la agregación de tres fuentes diferentes, las compras realizadas con tarjetas débito y crédito emitidas a los clientes de la entidad, las variables demográficas de estos clientes y la información de los establecimientos en donde estas personas realizan sus compras. A continuación, se describirán estas variables de estas tres fuentes:

2.1 Transacciones

En primer lugar, se extrae la información transaccional de los compradores, para esto se toman todas las transacciones que realizaron estos clientes entre enero de 2014 y diciembre de 2017. Esta información se revisa y estructura de manera que pueda ser utilizada. Una de las revisiones que se realiza es garantizar que todas las transacciones hayan sido compras efectivas, con el fin de no involucrar en el análisis otro tipo de transacciones como devoluciones de compras, transacciones rechazadas, entre otras. Una vez finalizadas estas validaciones, para cada compra (transacción), se cuenta con los siguientes campos:

- Fecha en la que se realizó la transacción.
- Valor de la transacción.

- Número de la tarjeta.
- Código de establecimiento donde se realizó la transacción.
- Número de ID del cliente.

2.2 Demografía

A cada número de las tarjetas (débito y crédito) se le asigna la información demográfica del propietario. A continuación, se relacionan las variables con las que se cuenta y las subsecuentes correcciones o transformaciones:

- Número de ID del cliente.
- Género.
- Dirección de residencia. Estas direcciones se estandarizan para posteriormente calcular la georreferenciación de este lugar.
- Ciudad donde está ubicada la residencia.
- Latitud: se calcula en función de la dirección de residencia de la persona.
- Longitud: se calcula en función de la dirección de residencia de la persona.
- Ingresos mensuales: estos se dividen por el valor de salario mínimo mensual legal vigente de cada año respectivo, con el fin de estandarizar la serie y facilitar el análisis.
- Edad de la persona: esta variable se calcula para cada año, teniendo en cuenta la fecha de nacimiento de la persona.

2.3 Establecimientos

Como se mencionó en la sección Transacciones, todos los establecimientos de comercio tienen un código que los identifica, el cual es único e irrepetible. A través de este código, a cada establecimiento se le asocian las siguientes variables:

- Nit de la empresa dueña del establecimiento.
- Nombre de la empresa.
- Código del establecimiento.
- Nombre del establecimiento.
- Código MCC¹.
- Ciudad de ubicación del establecimiento.
- Dirección del establecimiento: esta variable se estandariza para poder posteriormente calcular su georreferenciación. Sin embargo, algunos establecimientos ubicados dentro de centros comerciales no reportan una dirección que pudiese ser georreferenciada (e.g. Centro comercial xxxx local 34), a estos se les asigna la dirección del centro comercial².
- Latitud: se calcula en función de la dirección del establecimiento. En algunos casos se utiliza la latitud del centro comercial donde está ubicado el establecimiento.

¹ MCC = Merchant Category Code. Este código identifica la actividad comercial de cada establecimiento. Existen 204 códigos MCC distintos.

² Para poder realizar estas asociaciones, se realizó un trabajo previo de identificación de cada uno de los establecimientos que componen los principales centros comerciales de la ciudad de Medellín.

- Longitud: se calcula en función de la dirección del establecimiento. En algunos casos se utiliza la longitud del centro comercial donde está ubicado el establecimiento.
- Nombre del centro comercial donde está ubicado el establecimiento en los casos que aplique.

Esta investigación se acotó únicamente a las transacciones realizadas en establecimientos de comercio ubicados en Medellín.

2.4 Modelo lógico de unión de las fuentes

La Fuente: E representa la lógica de entidades, y cómo se relacionan las transacciones con la información demográfica de clientes y la información de los establecimientos.

Ilustración 3. Modelo lógico de datos



Fuente: Elaboración propia, 2018.

Una vez unificadas las tres fuentes de información, se procedió a construir una base de datos final con toda la información agregada y filtrada, que solo contuviera transacciones realizadas

en la ciudad de Medellín y su área metropolitana. Dentro de los cuatro años de estudio se contó con 145.651.712 transacciones, las cuales fueron realizadas por 1,94 millones de ID distintos, en 31.151 establecimientos.

A la base de datos final se le realizan las siguientes modificaciones:

- Se elimina el campo fecha y se crean tres nuevos campos, Año, Mes y un campo concatenado de los dos anteriores (Año&Mes).
- Se agrega un nuevo campo que contará el número de transacciones, ya que se encuentran casos de ID's que realizan múltiples transacciones durante un mes en un mismo establecimiento.
- El campo valor de la transacción se modifica para que contenga la suma de los valores de las transacciones, de acuerdo a los cambios anteriores. Este nuevo campo contiene el valor total de las transacciones que cada ID hizo durante un mes de un año en un establecimiento específico.
- Se agrega la variable Ticket Promedio (valor promedio de cada compra), la cual es la división entre la suma del valor total de las transacciones realizadas por un ID en una fecha específica en un establecimiento específico.
- Se agrega la variable Distancia km. Esta variable es la distancia en kilómetros entre la residencia de cada individuo y la ubicación del establecimiento donde efectuó una compra. Esta variable se calcula en función de la georreferenciación de los puntos anteriormente mencionados (ubicación residencia cliente y ubicación establecimiento) y para ello se utiliza la Fórmula del Haversine («Fórmula del

haversine», 2017), la cual es utilizada en trigonometría esférica para calcular la distancia sobre la tierra entre dos puntos georreferenciados (teniendo su latitud y longitud), asumiendo que el radio equivolumen de la tierra es de 6.371 km.

Todos los datos anteriormente mencionados están almacenados en un cluster de Hadoop y toda la minería de datos fue trabajada a través de Cloudera Impala utilizando SQL (Structured Query Language por sus siglas en ingles). Toda la modelación posterior fue realizada en R.

2.5 Método a aplicar

1. Seleccionar dos empresas de comercio minorista para analizar. Estas empresas deben cumplir los siguientes criterios:
 - Caso 1: una empresa que tenga por lo menos un establecimiento activo durante la mayor parte de los 48 meses de estudio. Además, que durante el periodo de análisis haya abierto y cerrado un nuevo establecimiento dentro de uno de los principales centros comerciales de la ciudad de Medellín. Este nuevo establecimiento debe haber estado en funcionamiento por lo menos 12 meses. El(los) establecimiento(s) inicial(es) y activo(s) debe tener un periodo de actividad previa a la apertura del nuevo establecimiento de por lo menos 12 meses, además este(os) sirve como parámetro de control para analizar el desempeño del nuevo establecimiento y el impacto agregado en las ventas de la empresa.
 - Caso 2: una empresa que tenga por lo menos un establecimiento activo durante la mayor parte de los 48 meses de estudio. Además, que durante el

periodo de análisis haya abierto un nuevo establecimiento dentro de uno de los principales centros comerciales de la ciudad de Medellín y este esté aún en funcionamiento al final del periodo de análisis. El(los) establecimiento(s) inicial(es) y activo(s) debe tener un periodo de actividad previa a la apertura del nuevo establecimiento de por lo menos 12 meses, además este(os) sirve como parámetro de control para analizar el desempeño del nuevo establecimiento y el impacto agregado en las ventas de la empresa.

2. Determinar el área de influencia de(los) establecimiento(s) inicial(es) de cada empresa seleccionada, al posicionar todos las personas que realizaron compras en un mapa alrededor del(os) establecimiento(s) inicial(es).
3. Determinar el perfil demográfico de los compradores de cada una de las empresas seleccionadas durante el periodo de tiempo previo a la apertura del nuevo establecimiento. El perfil demográfico se hace en función de las variables, edad, ingresos, distancia entre el lugar de residencia y el establecimiento y el ticket promedio.
4. Determinar el área de influencia de los compradores del centro comercial donde cada una de las empresas seleccionadas va a abrir su nuevo establecimiento, de acuerdo a la categoría MCC, al posicionar todas las personas que realizaron compras en un mapa alrededor de la ubicación del centro comercial.
5. Determinar el perfil demográfico del centro comercial donde cada una de las empresas seleccionadas va a abrir su nuevo establecimiento. Este análisis se hace para el mismo periodo de tiempo en que esté activo el(los) establecimiento(s) inicial(es)

de cada empresa. Además, se evalúan solo las compras realizadas en establecimientos de la misma categoría que las empresas seleccionadas. El perfil demográfico se hace en función de las variables, edad, ingresos, distancia entre el lugar de residencia y el establecimiento y el ticket promedio.

6. Comparar los mapas de áreas de influencia entre cada una de las empresas caso de estudio y el área de influencia del respectivo centro comercial donde planea abrir su nuevo establecimiento, esto con el fin de definir a priori si ambas áreas se traslapan y por lo tanto podría haber evidencia de posible canibalización entre el establecimiento inicial y el nuevo.
7. Una vez determinados los perfiles de cada empresa caso de estudio y del respectivo centro comercial donde se ubicará el nuevo establecimiento, se contrastan estos resultados con el fin de determinar la compatibilidad entre ambos perfiles demográficos.
8. Analizar el desempeño de cada uno de los establecimientos de las empresas seleccionadas a través de los siguientes KPI's (indicadores clave de desempeño por sus siglas en inglés):
 - Variación acumulada anual de las ventas estimadas de cada establecimiento de las empresas seleccionadas.
 - Variación acumulada anual de las ventas estimadas de las empresas seleccionadas.

- Participación de cada establecimiento en las ventas totales estimadas de las empresas seleccionadas.
- Ticket promedio de los compradores de cada establecimiento y del total de la empresa.
- Canibalización (%) entre establecimientos. Este se mide como el porcentaje de clientes que realizaron compras en el establecimiento inicial y después de la apertura del nuevo establecimiento, realizan compras en este. Este indicador se calcula para periodos anuales.
- Variación (%) del Ticket Promedio.

3 Resultados

Para el caso 1 se seleccionó una empresa de comercialización de calzado que llamaremos empresa A. Para el caso 2 se seleccionó una empresa de comercialización de vestuario que llamaremos empresa B³.

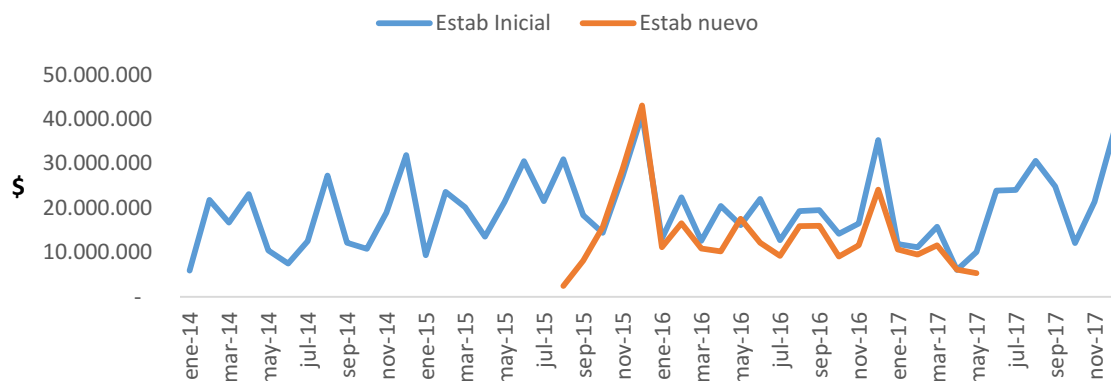
Caso 1 – Empresa A:

La empresa A cuenta con un punto de venta activo en la ciudad desde enero de 2014 y abrió un nuevo punto de venta en agosto de 2015. Este segundo establecimiento fue cerrado 22 meses después (mayo 2017).

³ Por temas de confidencialidad, se mantendrán anónimos los nombres de las empresas seleccionadas.

Desempeño general de las ventas estimadas de la empresa:

Ilustración 4. Desempeño empresa A



Fuente: Elaboración propia, 2018.

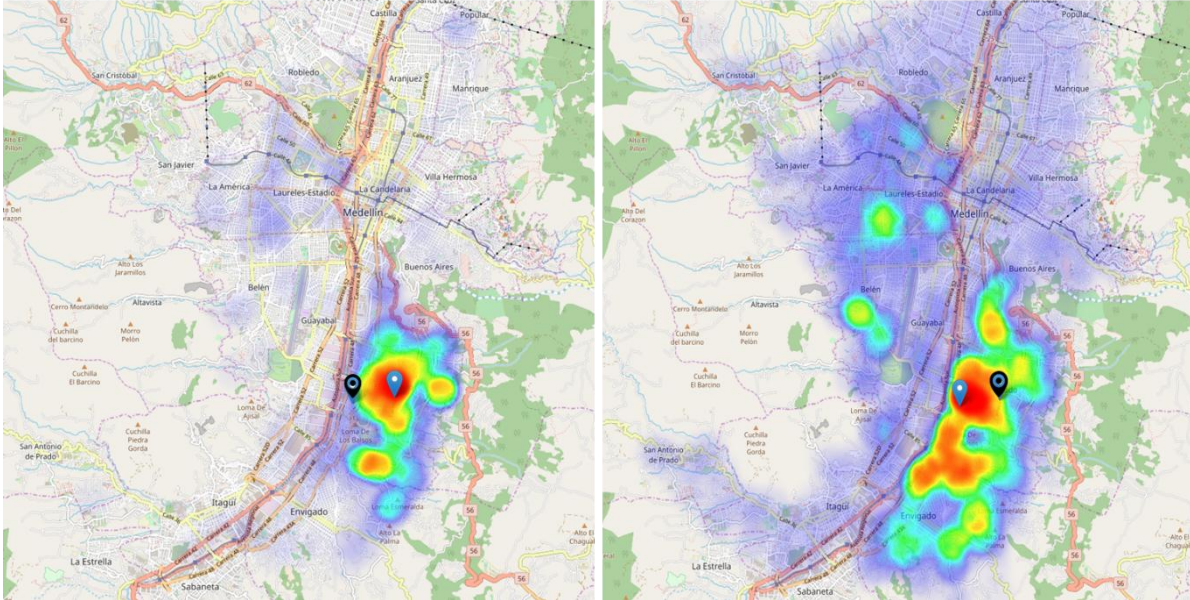
La Fuente: Elaboración propia, muestra el desempeño de las ventas estimadas de cada uno de los establecimientos, en donde se puede inferir que el nuevo establecimiento tuvo un desempeño inferior al del establecimiento inicial durante todo 2016 y 2017.

Comparación áreas de influencia antes de nueva apertura:

Ilustración 5. Comparación áreas de influencia caso 1

Empresa A (marcador azul)

Centro comercial – categoría calzado (Marcador Azul)



Fuente: Elaboración propia, 2018.

El mapa del lado izquierdo de la Fuente: Elaboración propia, 2018.

muestra la ubicación espacial de los clientes del establecimiento inicial de la empresa A (marcador rojo), donde se puede ver que existe una concentración importante en zonas cercanas a la ubicación del establecimiento, por lo que se podría inferir que este establecimiento hace gran parte de sus ventas a distancias cortas. El mapa del lado derecho muestra el área de influencia de los compradores de la categoría de calzado, donde se puede concluir que a este centro comercial van personas de áreas muchos más apartadas.

Comparación perfiles demográficos:

La Tabla 1 muestra la comparación entre el perfil demográfico de los clientes de la empresa A (parte superior) y el perfil demográfico de los clientes del centro comercial (parte inferior), donde abrió el nuevo punto de venta.

Comprador promedio empresa A entre enero de 2014 y julio de 2015: el cliente promedio de esta empresa de comercio minorista es una persona de 47,3 años, lo que lo ubica dentro de la generación X, gana en promedio \$11,4 millones de pesos mensuales, vive a 2,22 kilómetros de distancia de la ubicación del punto de venta y su gasto promedio en la marca es de \$235.000.

Comprador promedio centro comercial categoría calzado entre enero de 2014 y julio de 2015: el cliente promedio de este centro comercial es una persona de 42,4 años, lo que lo ubica dentro de la generación X, gana en promedio \$5,4 millones de pesos mensuales, vive a 4,2 kilómetros de distancia de la ubicación del punto de venta y su gasto promedio en establecimientos de calzado dentro del centro comercial es de \$147.000.

Tabla 1. Resumen variables caso 1

Resumen Variables - Comercio Calzado				
	Edad	Ingresos	Distancia (km)	Ticket promedio
Min.	14.00	1.18	0.04	16,960
1er Cuartil	38.00	6.31	0.85	123,438
Mediana	48.00	10.70	1.55	198,000
Media	47.35	14.64	2.22	235,813
3er Cuartil	56.00	21.14	2.50	295,000
Max.	97.00	49.19	6.41	310,000
Desviación Estándar	12.7	11.4	2.1	174,088
Resumen Variables - Centro Comercial (calzado)				
	Edad	Ingresos	Distancia (km)	Ticket promedio
Min.	4.00	-	0.05	100
1er Cuartil	32.00	3.39	2.00	74,900

Mediana	39.00	5.46	3.46	130,000
Media	42.38	6.84	4.21	147,691
3er Cuartil	51.00	9.14	5.67	185,000
Max.	98.00	1.51	8.22	519,300
Desviación Estándar	13	4.6	3.1	107,780

Fuente: Elaboración propia, 2018.

Desempeño de las ventas:

La siguiente tabla muestra los principales indicadores con los cuales se puede evaluar el desempeño de cada uno de los establecimientos y del agregado de las ventas de la empresa de calzado, antes y después de la apertura del segundo establecimiento.

Tabla 2. KPI's empresa A

Desempeño agregado Empresa A				
	KPI's	Establecimiento inicial	Establecimiento nuevo	Total
2015	Crec (%) acumulado	35.3%	NA	83.8%
	Part (%) ventas	48.8%	51.2%	100%
	Ticket Promedio	356,476	336,116	345,760
	% canibalización		2.73%	
2016	Crec (%) acumulado	-17.9%	70.4%	5.4%
	Part (%) ventas	60.5%	39.5%	100%
	Ticket Promedio	301,788	278,961	292,331
	Var (%) Ticket Promedio	-15.3%	-17.0%	-15.5%
	% canibalización		3.95%	
2017	Crec (%) acumulado	5.6%	-73.2%	-28.1%
	Part (%) ventas	66.5%	33.5%	100%
	Ticket Promedio	300,914	264,960	287,840
	Var (%) Ticket Promedio	-0.3%	-5.0%	-1.5%
	% canibalización		5.72%	

Fuente: Elaboración propia, 2018.

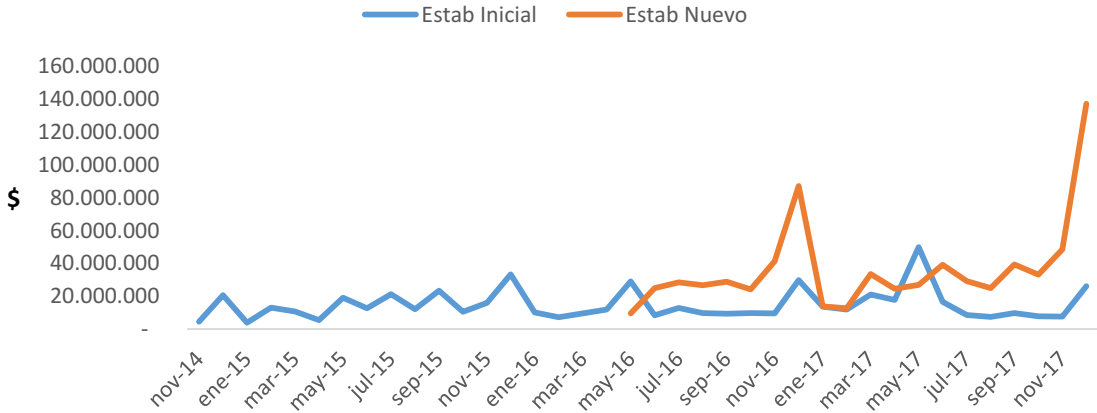
Como se puede ver en la Tabla 2, el nuevo establecimiento creció un 70% el primer año, sin embargo, en el segundo año presenta un decrecimiento del 73.2%, por lo que se puede concluir que este establecimiento empezó a deteriorar su desempeño desde el mismo 2016. Adicionalmente, se puede apreciar que el ticket promedio del nuevo establecimiento fue siempre menor al del establecimiento antiguo; adicionalmente, el establecimiento nuevo tuvo los mayores decrecimientos en el ticket promedio. Finalmente, la canibalización entre ambos establecimientos fue menor al 6% para todos los años.

Caso 2 – Empresa B:

La empresa cuenta con un punto de venta activo en la ciudad en noviembre de 2014 y abrió un nuevo punto de venta en mayo de 2016 en un centro comercial de la ciudad. Este segundo establecimiento continúa abierto hasta el momento en el que se realizaron los análisis.

Desempeño general de las ventas estimadas de la empresa:

Ilustración 6. Desempeño empresa B



Fuente: Elaboración propia, 2018.

La Fuente: Elaboración propia, 2018.

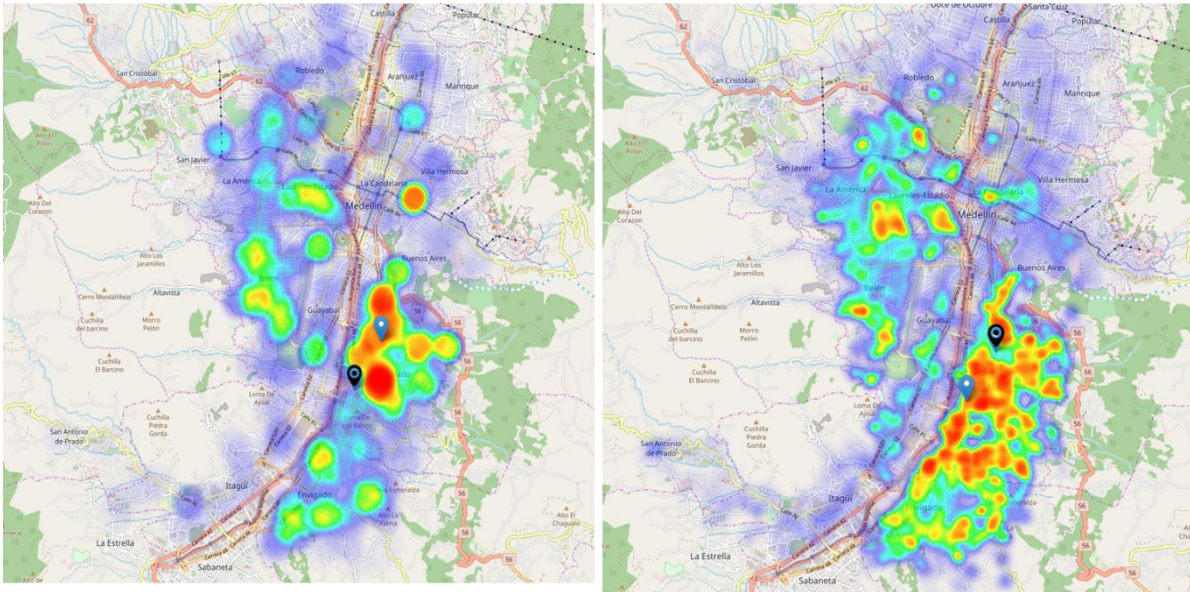
muestra el desempeño de los establecimientos (inicial y nuevo) de la empresa de vestuario. Se puede concluir que el establecimiento nuevo tuvo, desde su apertura, un comportamiento superior al establecimiento anterior.

Comparación áreas de influencia antes de nueva apertura:

Ilustración 7. Comparación áreas de influencia caso 2

Empresa B (marcador azul)

Centro comercial – categoría vestuario (marcador azul)



Fuente: Elaboración propia, 2018.

Al comparar ambos mapas de la Fuente: Elaboración propia, 2018.

, se puede concluir que ambas áreas de influencia abarcan zonas importantes de la ciudad, sin embargo, el área de influencia del centro comercial parecer agrupar un número mayor de clientes y tener mayor cubrimiento de la ciudad.

Comparación perfiles demográficos:

La Tabla 1Tabla 3 muestra la comparación entre el perfil demográfico de los clientes de la empresa B (parte superior) y el perfil demográfico de los clientes del centro comercial (parte inferior), donde abrió el nuevo punto de venta.

Comprador promedio de la empresa B: el cliente promedio de esta empresa de comercio minorista es una persona de 34 años, lo que lo ubica dentro de la generación *Millenials*, gana en promedio \$3,2 millones de pesos mensuales, vive a 4,4 kilómetros de distancia de la ubicación del punto de venta y su gasto promedio en la marca es de \$85.000.

Comprador promedio centro comercial categoría vestuario: el cliente promedio de este centro comercial es una persona de 39 años, lo que lo ubica dentro de la generación X, gana en promedio \$4,2 millones de pesos mensuales, vive a 4,9 kilómetros de distancia de la ubicación del punto de venta y su gasto promedio en establecimientos de calzado dentro del centro comercial es de \$132.000.

Tabla 3. Resumen variables caso 2

Resumen variables – empresa B				
	Edad	Ingresos	Distancia (km)	Ticket promedio
Min.	4	0.00	0.19	3,200
1er Cuartil	26	1.99	1.96	43,900
Mediana	30	3.20	4.09	68,658
Media	34	4.03	4.40	84,947
3er Cuartil	39	5.37	5.80	104,920
Max.	78	2.80	5.60	558,758
Desviación Estándar	11.15	2.65	2.99	61,748
Resumen Variables - Centro Comercial (vestuario)				

	Edad	Ingresos	Distancia (km)	Ticket promedio
Min.	2	-	0.05	1
1er Cuartil	29	2.43	2.36	67,600
Mediana	36	4.26	4.13	102,348
Media	39	5.35	4.91	132,278
3er Cuartil	48	7.26	6.65	158,700
Max.	118	7.96	8.22	590,000
Desviación Estándar	13.48	3.8	3.37	123,344

Fuente: Elaboración propia, 2018.

Desempeño de las ventas:

La siguiente tabla muestra los principales indicadores con los cuales se puede evaluar el desempeño de cada uno de los establecimientos y del agregado de las ventas de la empresa de vestuario, antes y después de la apertura del segundo establecimiento.

Tabla 4. KPI caso 2

Desempeño Empresa Vestuario				
KPI's	Establecimiento Inicial	Establecimiento nuevo	Total	
2015	Crec (%) acumulado	98.6%	NA	98.63%
	Part (%) ventas	100.0%	0.0%	100%
	Ticket Promedio	105,952	NA	105,952
2016	Crec (%) acumulado	-13.6%	NA	135.5%
	Part (%) ventas	25.6%	74.4%	100%
	Ticket Promedio	97,511	92,269	93,572
	Var (%) Ticket Promedio	-8.0%	NA	-11.7%
	% canibalización		8.82%	
2017	Crec(%) acumulado	26.3%	70.9%	54.6%
	Part (%) ventas	16.2%	83.8%	100%
	Ticket Promedio	104,014	93,975	95,454
	Var (%) Ticket Promedio	6.7%	1.8%	2.0%
	% canibalización		7.02%	

Fuente: Elaboración propia, 2018.

Como se puede ver en el anterior cuadro, el nuevo establecimiento, el cual fue abierto en mayo del 2016, creció un 70,9% el primer año acumulado completo (2017), sin embargo, al calcular el crecimiento anual entre mayo de 2017 y mayo de 2016 (mes de la apertura), el crecimiento es el 179,8%. Se puede apreciar que el ticket promedio del nuevo establecimiento ha sido siempre menor al del establecimiento antiguo. Finalmente, la canibalización entre ambos establecimientos fue menor al 8,82% para el primer año y del 7,02% para el segundo año.

Discusión de los resultados:

Al comparar ambos casos es posible concluir que la apertura del nuevo establecimiento de la empresa A no debió haberse hecho, ya que los perfiles demográficos de los compradores del establecimiento inicial y los compradores de la categoría de calzado en el centro comercial son bastante diferentes. Para la empresa B, los resultados obtenidos permiten observar el éxito que ha tenido la apertura de la nueva tienda, el cual puede deberse a las similitudes presentadas entre los perfiles demográficos del establecimiento inicial y el centro comercial. Adicional, los indicadores de desempeño de la tienda de vestuario (KPI's), corroboran el buen comportamiento del nuevo establecimiento durante el periodo analizado.

Al analizar las áreas de influencia de los puntos de venta inicial, y nuevos, para ambos casos, se evidencia que éstas se traslapan, lo que conducía a esperar un nivel alto de canibalización. Sin embargo, los resultados muestran una realidad diferente, ya que el nivel de canibalización

para ambos casos es inferior al 10%, por lo que se puede concluir que una gran porción de los consumidores de un centro comercial son afines, independiente de qué tan cercano sea a su residencia.

4 Conclusiones y recomendaciones

Luego de desarrollar la metodología, se estiman los perfiles demográficos de las dos empresas seleccionadas para los casos 1 y 2 y, de forma paralela, los perfiles demográficos de las ubicaciones de los nuevos puntos de venta. Adicionalmente, se logra analizar el comportamiento de las ventas agregadas estimadas de ambas empresas a través de KPI's, en donde se estima al nivel de canibalización, el cual es el porcentaje de clientes que migraron de comprar en el establecimiento inicial, al establecimiento nuevo.

Al inicio de la investigación se esperaba contar con un mayor número de variables demográficas, como, por ejemplo, la ocupación, nivel académico, estrato socioeconómico del lugar de residencia, entre otras; sin embargo, al incluir estas variables, se reducía considerablemente la muestra, ya que solo una porción inferior al 50% de la muestra contaba con esta información.

La información de transacciones financieras tiene sesgos, ya que esta no involucra a las personas que no están bancarizadas. Adicionalmente, no se contemplan las transacciones que se realizan con efectivo y, por lo tanto, personas que sean afines a este medio de pago no son tenidas en cuenta en este ejercicio.

Esta investigación se podría extender al determinar los perfiles demográficos a partir del lugar de residencia de las personas y no del lugar donde realizan sus compras, con lo cual se podría estimar un proxy del potencial de mercado de las diferentes zonas de la ciudad. Este ejercicio podría enriquecer los análisis de isócronas presentados en el marco teórico y sería de gran utilidad para varios agentes económicos, como directivos de empresas de comercio y desarrolladores de centros comerciales, entre otros.

Referencias

- Asobancaria (2018, febrero 12). Edición 1124 | *Big Data*: aprovechando los datos, el nuevo gran activo. Recuperado 24 de febrero de 2018, a partir de <http://www.asobancaria.com/2018/02/12/edicion-1124/>
- Barranco, R. (2012, junio 18). ¿Qué es *Big Data*? Recuperado 18 de febrero de 2018, a partir de <http://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/index.html>
- BBVA (2017, mayo 8). Las cinco uves del *big data* | BBVA. Recuperado 25 de febrero de 2018, a partir de <https://www.bbva.com/es/las-cinco-uves-del-big-data/>
- Datameer (2017, mayo 12). What Is *Big Data* Analytics? Recuperado 25 de febrero de 2018, a partir de <https://www.datameer.com/what-is-big-data-analytics/>
- Fórmula del haversine (2017, octubre 14). En *Wikipedia, la enciclopedia libre*. Recuperado el 25 de febrero de 2018 a partir de https://es.wikipedia.org/w/index.php?title=F%C3%B3rmula_del_haversine&oldid=102570421

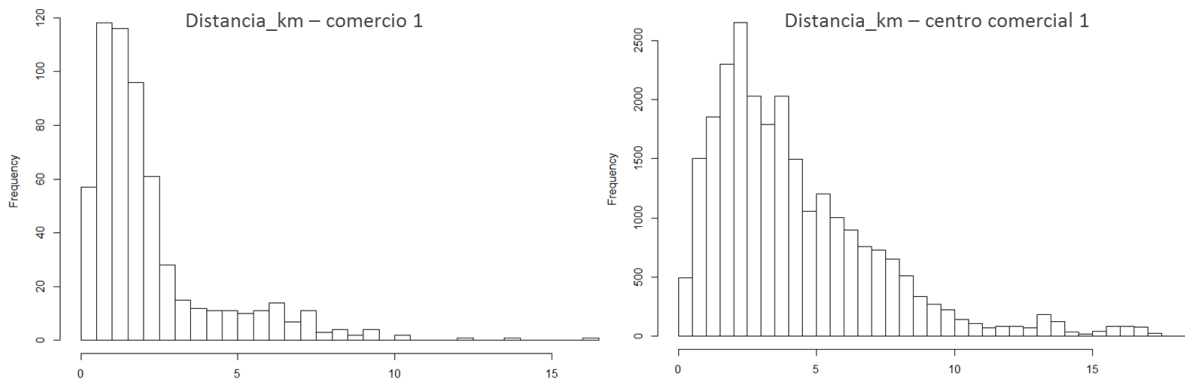
- Gartner (2014). What Is *Big Data*? - Gartner IT Glossary - *Big Data*. Recuperado 18 de febrero de 2018, a partir de <https://www.gartner.com/it-glossary/big-data>
- IIC (2016, junio 28). Las 7 V del *Big data*: Características más importantes - IIC. Recuperado 26 de febrero de 2018, a partir de <http://www.iic.uam.es/innovacion/big-data-caracteristicas-mas-importantes-7-v/>
- La República (2018, enero 15). Más de 20 centros comerciales abrirán sus puertas este año, duplicando cifra de 2017. Recuperado 9 de abril de 2018, a partir de <https://www.larepublica.co/empresas/mas-de-20-centros-comerciales-abriran-sus-puertas-este-ano-2588268>
- Marr, B. (2015). *Big Data*: 20 Mind-Boggling Facts Everyone Must Read. Recuperado 19 de marzo de 2018, a partir de <https://www.forbes.com/sites/bernardmarr/2015/09/30/big-data-20-mind-boggling-facts-everyone-must-read/>
- Mazzei, M. J., & Noble, D. (2017). *Big data* dreams: A framework for corporate strategy. *Business Horizons*, 60(3), 405-414. <https://doi.org/10.1016/j.bushor.2017.01.010>
- Meza, J. (2013). Evaluación financiera de proyectos: 10 casos prácticos resueltos en Excel. Bogotá, Colombia: Ecoe Ediciones.
- Penrose, E. T. (1959). *The Theory of the Growth of the Firm*. Oxford: Oxford University Press.
- Press, G. (2013). *Big Data* News: A Revolution Indeed. Recuperado 17 de febrero de 2018, a partir de <https://www.forbes.com/sites/gilpress/2013/06/18/big-data-news-a-revolution-indeed/>

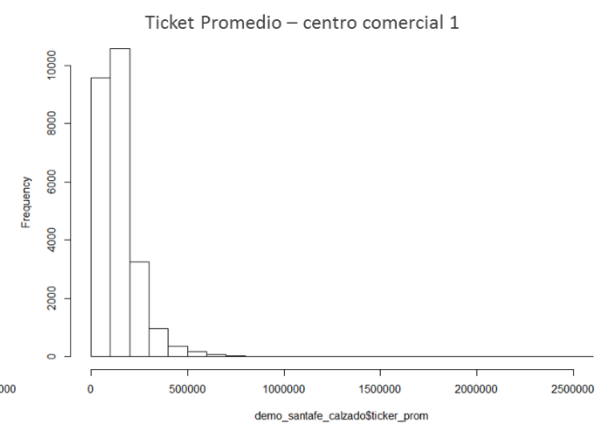
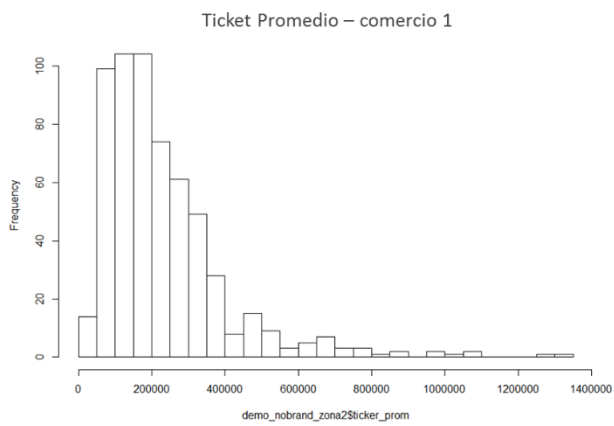
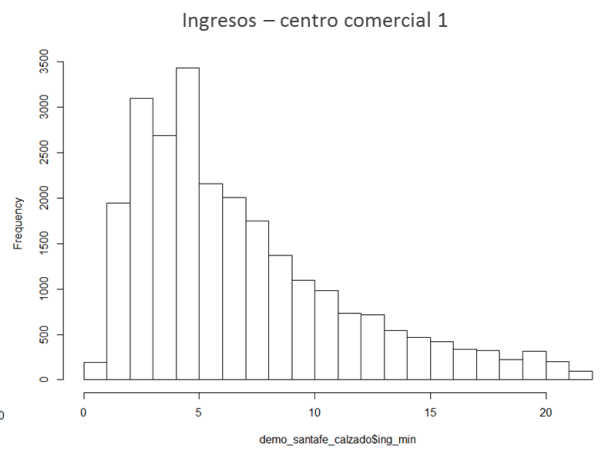
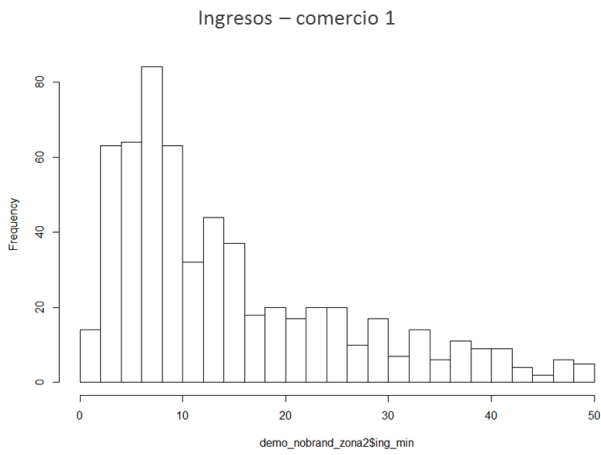
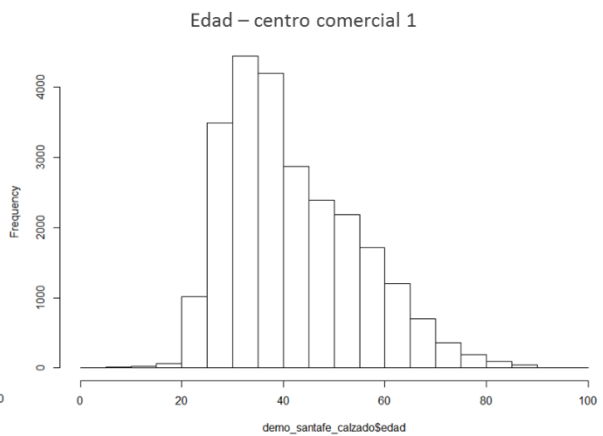
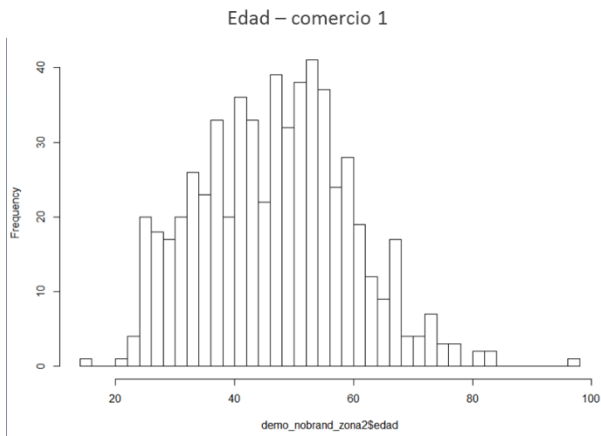
Press, G. (2014). 12 *Big Data* Definitions: What's Yours? Recuperado 17 de febrero de 2018, a partir de <https://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/>

Thau, B. (2014). How *Big Data* Helps Chains Like Starbucks Pick Store Locations -- An (Unsung) Key To Retail Success. Recuperado 20 de marzo de 2018, a partir de <https://www.forbes.com/sites/barbarathau/2014/04/24/how-big-data-helps-retailers-like-starbucks-pick-store-locations-an-unsung-key-to-retail-success/>

Anexos:

Demografía Empresa A vs centro comercial categoría calzado:





Demografía Empresa B vs centro comercial categoría vestuario:

