

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

ESTRUCTURACIÓN DE ESTRATEGIAS DE MARKETING PARA LA TOMA DE
DECISIONES APLICANDO LA MINERÍA DE DATOS EN BANCOLOMBIA

Claudia Sorany Cuervo Hernández

Marilin Muñoz Ospina

Adriana Marcela Zapata Cano

Administración Tecnológica

Julián Alberto Uribe Gómez
(Magister en Gestión Tecnológica)

INSTITUTO TECNOLÓGICO METROPOLITANO

22 de octubre 2020

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

RESUMEN

En la sociedad actual, el mercado evoluciona constantemente, y las personas pueden cambiar de percepciones tan rápido como el mercado, por lo cual, es importante que las empresas estén en todo momento al tanto de estos cambios, conociendo y monitoreando los comportamientos, patrones y relaciones de las compras y transacciones que realizan sus clientes, esto con el fin de actuar de una manera preventiva y predictiva, anticipándose a hechos que puedan afectar a la organización en sus ingresos.

El Big data se ha convertido en una herramienta esencial que ayuda a procesar grandes cantidades de información, y extraer de allí la información más relevante aplicando técnicas estadísticas y reglas asociativas que apoyen la toma de decisiones y determinar, entre otras cosas estrategias de marketing acordes al comportamiento de los consumidores afianzando aún más la relación con sus clientes al interior de las organizaciones.

Por tal motivo, en el sector bancario el Data Mining y por lo tanto el Big Data ha servido para crear modelos y determinar patrones de diferentes medios de pago que permiten segmentar los clientes, y medir la efectividad de las campañas de marketing desarrolladas por dichas entidades. Desde estas premisas, se partió de una base de datos en Excel suministrada por la entidad financiera Bancolombia con 504.914 registros, en el cual se aplicó el modelo CRISP-DM.

Finalmente, se encontraron datos representativos para la organización, por ejemplo, que se realizan mayor número de transacciones con tarjeta débito, lo que sugiere que los clientes están cada vez más familiarizados con medios de pago electrónicos; con dicha información se podrán establecer estrategias de marketing más dirigidas al cliente de Bancolombia.

Palabras clave: Minería de datos, Big Data, Toma de decisiones, Estrategias de marketing, Clientes

	<p style="text-align: center;">INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm</p>	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

RECONOCIMIENTOS

Inicialmente agradecemos a nuestro tutor Julian Alberto Uribe Gómez, quien nos compartió su conocimiento y nos guio para realizar el presente trabajo, también a Dios por permitirnos darnos la oportunidad de cumplir una meta más en nuestras vidas, a nuestra alma Mater, quien nos aprovisiona de conocimiento y de crecimiento personal para orientar nuestro camino, a nuestras familias por brindarnos su apoyo incondicional fomentando en nosotras el deseo de superación, humildad y sacrificio para ser cada días mejores y tener una mejor calidad de vida, continuar adelante y ser un ejemplo para las personas que nos rodean, también a Bancolombia por permitirnos desarrollar nuestro proyecto de grado en la entidad, a cada una de las personas que intervinieron en nuestra formación académica y humana, y a nosotras por todos los esfuerzos y sacrificios que hicimos para salir adelante y poder hacer realidad este sueño.

 Institución Universitaria	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

ACRÓNIMOS

KDD: Knowledge Discovery in Databases

H1N1: virus A, hemaglutinina y neuramidasa

SQL : Structured Query Language

OLAP: On-Line Analytical Processing

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

CONTENIDO E ÍNDICES

1. JUSTIFICACIÓN	8
2. PLANTEAMIENTO DEL PROBLEMA O SITUACIÓN	9
3. OBJETIVOS	10
3.1. Objetivo General	10
3.2. Objetivos Específicos.....	10
3.3 Alcance	10
4. MARCO REFERENCIAL Y DOCUMENTACIÓN CONSULTADA	11
4.1. Bases de datos	11
4.2 Big Data	11
4.3 Data mining o Minería de datos.....	14
4.3.1 ¿Para qué sirve la minería de datos?	14
4.4 Análisis de datos con Microsoft Excel.....	19
4.5 Segmentación de clientes:	20
4.6 Marketing.....	20
4.7 Estrategias de marketing	21
4.7.1 Estrategias de marketing Mix	23
4.7.1.2 Decisiones sobre Distribución.....	23
4.7.1.3 Decisiones sobre Organización	24
4.7.1.5 Decisiones sobre Precio	24
4.7.1.6 Decisiones sobre Servicio.....	24
4.8 Fidelización de clientes	25
4.9 Toma de decisiones.....	26
5. Toma de decisiones estratégicas	26
5. METODOLOGÍA	27
6. RESULTADOS.....	31
7. CONCLUSIONES Y RECOMENDACIONES	40

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Árbol de clasificación Datamining.	17
Ilustración 2: Fases de Metodología Crisp DM.	30
Ilustración 3. Porcentaje medios de pago.....	31
Ilustración 4. Cantidad de transacciones por género.....	32
Ilustración 5. Porcentaje de compras por categorías.....	32
Ilustración 6. Porcentajes compra por meses	33
Ilustración 7. Porcentaje de compras por días.	33
Ilustración 8. Porcentaje de compras por rangos de edades.....	34
Ilustración 9. Porcentaje de compras por días de la semana.	35
Ilustración 10. Compras por días del mes	35
Ilustración 11. Porcentaje de compras por estratos socioeconómicos.	36
Ilustración 12. Porcentaje de compras por días en el estrato 4.	37
Ilustración 13. Porcentaje de compras por categoría en el estrato 4.	38
Ilustración 14. Porcentaje de usuarios en rangos de edades.....	38

ÍNDICE DE TABLAS

Tabla 1. Definición de variables de la base de datos Bancolombia	28
Tabla 2. Ficha técnica de investigación.	30
Tabla 3. Coeficientes de regresión.....	39
Tabla 4. Análisis de varianza	39
Tabla 5. Coeficientes.....	39

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

INTRODUCCIÓN

Actualmente, las organizaciones realizan estudios de mercado que les ayuda a conocer mejor a sus clientes actuales y potenciales, saber cuáles son sus gustos, preferencias, ubicación, estrato, educación y ocupación, lo cual les permite generar estrategias más personalizadas a cada segmento de clientes, con el fin de aumentar sus ventas y garantizar la satisfacción del usuario; y así poder describir los grandes volúmenes de datos estructurados y no estructurados que abundan en las organizaciones.

Toda la información generada como consecuencia del uso del internet, redes sociales y uso de tarjetas débito y crédito, es una fuente potencial para que las empresas puedan conocer más a sus clientes, identificando sus gustos y preferencias, asimismo al interior de las organizaciones existen áreas dedicadas exclusivamente a la consolidación de Bases de datos; esta investigación se dirige a la forma en cómo recopilar y analizar información fundamental para crear y dirigir dichas estrategias hacia sus clientes, complementando las estrategias con un buen manejo de la cantidad y calidad de los datos.

El objetivo de este trabajo es describir cómo la implementación de metodologías de investigación aplicadas a los datos de los clientes, como Data Mining, Big data, Análisis exploratorio, Predicción, entre otras, posibilita mejorar la eficiencia en los procesos de marketing dentro de una entidad financiera.

En la actualidad, este tema se encuentra en auge a nivel mundial dado que se han venido buscando diferentes tipos de estrategias para el manejo de la información, donde se tienen en cuenta todos los movimientos y transacciones que los clientes generan día a día, y así poder crear nuevos métodos que permitan mejorar la calidad y velocidad de los procesos de mercadeo, encaminado a que el cliente consuma lo que le gusta, creando un impacto de innovación con beneficios tanto para la organización como para los clientes generando así un valor agregado en los productos y servicios que se están ofreciendo.

El presente trabajo se inicia con la justificación del tema a desarrollar, donde luego se abordan referentes teóricos y metodológicos sobre la minería de datos y su uso en las organizaciones, siguiendo con el planteamiento del problema y se establecen los objetivos que se buscan alcanzar con este proyecto. Así mismo, se define el alcance y la metodología que se va a utilizar para desarrollar el trabajo.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

1. JUSTIFICACIÓN

En la actualidad, las empresas obtienen una gran cantidad de información de sus clientes con la cual pueden conocer el comportamiento de ellos, gustos, preferencias, actividades, tendencias, edades, géneros y frecuencias; debido a esto se busca que las entidades financieras hagan uso de dicha información, con el fin de enfocar las estrategias de mercadeo de forma más personalizada basada en el conocimiento del cliente.

Esto permitirá que dichas entidades obtengan fidelidad de sus clientes, puesto que al identificar las necesidades de cada uno de ellos podrán encaminar sus estrategias al cumplimiento de estas y así generar satisfacción, anticipándose a los requerimientos de ellos; adicional, esto les permitirá a las organizaciones optimizar los costos en los que incurre por publicidad, ya que pasará de ser general a particular, siendo más dirigida, oportuna y eficiente.

Este proyecto tendrá utilidad con el pasar del tiempo, gracias a que las metodologías para análisis de datos permitirán a las entidades financieras, adaptar sus estrategias de mercadeo a los cambios del entorno, y así generar una ventaja competitiva a dichas entidades, dado que podrán anticiparse a situaciones, dándoles la posibilidad de enfrentarlas con mayor asertividad y éxito para la organización.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

2. PLANTEAMIENTO DEL PROBLEMA O SITUACIÓN

En los últimos años, ha crecido en gran proporción la cantidad de información que las personas suministran al entorno, por medio de las redes sociales, internet, uso de tarjetas débito, crédito, entre otras, y toda esta información puede ser usada por las empresas para conocer a sus clientes a través de gustos, frecuencias, precio y lugares de las compras.

Examinando las herramientas que existen en la actualidad para el análisis de datos, se determina que su implementación permitiría a las organizaciones financieras recopilar la información para reconocer patrones y relaciones de los datos, con el fin de enfocar sus estrategias de marketing, y de esta manera fidelizar a los clientes satisfaciendo sus necesidades, y así mismo, al conocer las tendencias se podría atraer a nuevos clientes.

El desconocimiento de las necesidades de cada cliente conlleva a que se ofrezcan productos y servicios que no son atractivos para muchos, cumpliendo las expectativas de solo un pequeño segmento del mercado y adicionalmente, se incurre en grandes costos de publicidad debido a que no le está llegando al público idóneo, lo que se traduce en pérdidas para las compañías; teniendo datos más específicos se pueden tomar decisiones más acertadas acerca de los recursos que se invierten en cada campaña.

De ahí surge que el análisis de datos sea una gran oportunidad para las entidades financieras del país, puesto que su enfoque principal no es realizar este tipo de análisis, así mismo cabe resaltar que actualmente estos tipos de estudios solo están enfocados a un número limitado de áreas más especializadas; también es importante debido a que permite una orientación a satisfacer y generar valor a cada uno de sus consumidores de una forma más personalizada, entendiendo a cada cliente para que les permita adecuar los servicios que ofrecen.

El valor real del Data mining se da de acuerdo con la información que pueda extraerse y que ayude a la toma de decisiones y a mejorar las estrategias, empleando la analítica como herramienta fundamental en todo el proceso para optimizar eficiencia, reducir costos, incrementar las ganancias y de esta manera aumentar la satisfacción del cliente De acuerdo a lo anterior, se busca responder a la siguiente pregunta:

¿Cómo ayuda la implementación de metodologías de análisis de datos a Bancolombia, al desarrollo de estrategias de marketing de manera personalizada, con el fin de que permita el conocimiento de los gustos y preferencias de los clientes?

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

3. OBJETIVOS

3.1. Objetivo General

Estructurar las estrategias de marketing dirigidas a los clientes de Bancolombia a través de la definición de patrones y relaciones de las transacciones de estos.

3.2. Objetivos Específicos

1. Comprender los diferentes conceptos asociados a la minería de datos, a través de una búsqueda en bases de datos indexadas que permita conocer la aplicación del Big Data.
2. Aplicar herramientas de análisis de datos, a través de la definición de variables que permita realizar un análisis descriptivo de la información suministrada por la entidad financiera.

3.3 Alcance

Dentro del alcance del proyecto se busca obtener información valiosa y relevante de una base de datos suministrada por una entidad financiera, con el fin de buscar diferentes estrategias comerciales adaptadas específicamente a todos los segmentos de clientes; utilizando Data Mining como herramienta de negocio en la cual se pueden generar nuevas oportunidades para diversificar el portafolio de las empresas y optimizar el proceso de encontrar información predecible que normalmente requiere de un extenso análisis manual.

Es un proyecto que se puede aplicar a nivel Nacional, dado que la información contenida en la base de datos es de clientes de la entidad financiera Bancolombia de todo el país, y cuya información transaccional se ha generado por el uso de tarjetas débito y crédito principalmente, de esta manera el objetivo es obtener toda la información posible que pueda indicar el comportamiento y hábito de consumo de los clientes, y así segmentarlos en diferentes tipos de estrategias comerciales orientadas a mejorar las ventas a corto y a largo plazo, según sea la necesidad de la organización.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

4. MARCO REFERENCIAL Y DOCUMENTACIÓN CONSULTADA

A continuación, se realizará un recorrido por algunos conceptos teóricos que servirán de guía para el entendimiento del objetivo de la presente investigación:

4.1. Bases de datos

Según Trentin (1992), “una base de datos es un conjunto de datos estructurados y permanentes agrupados por su homogeneidad y relacionados entre ellos, organizados con la mínima redundancia para ser usados en aplicaciones diversas, de modo controlado” (p. 81).

Otra definición de base datos que sirve para explicarla es: "todo aquel material que sirve para certificar los resultados de la investigación que se realiza, que se han registrado durante ella y que se ha reconocido por la comunidad científica", es decir, se trata de cualquier información reunida, utilizada o generada en la experimentación, observación, medición, simulación, cálculo, análisis, interpretación, estudio o cualquier otro proceso de indagación que sustente y justifique las aportaciones científicas que se difunden en las publicaciones de investigación. Se presentan en cualquier formato y soporte, por ejemplo:

- **Archivos numéricos**, hojas de cálculo, tablas, etc.
- **Documentos de texto** en distintas versiones.
- **Imágenes**, gráficos, archivos de audio, vídeo, etc.
- **Código** o registros de software, bases de datos, etc.
- **Datos geoespaciales**, información georreferenciada.

(Sistema de bibliotecas de la Universidad de Extremadura, 2020)

4.2 Big Data

El término big data, terminología anglosajona ampliamente utilizada que se suele traducir por datos masivos, apareció a principios del siglo XXI en el entorno de las ciencias, en particular, de la astronomía y de la genética, debido a que ambos campos experimentaron una gran explosión en la disponibilidad de datos.

En estos últimos años la explosión de datos se ha generalizado en muchos de los campos que rodean nuestra vida cotidiana. Entre otros, el incremento del número de dispositivos con conexión a internet, el

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

auge de las redes sociales y el internet de las cosas (IoT).² han provocado una explosión en el volumen de datos disponibles. (Casas, Nin, & Julbe, 2019, p.23-24)

Según Castañeda, Benítez, Zúñiga y Vásquez (2013), definen el Big Data como: “Un conjunto de datos cuyo tamaño, se encuentra por encima de la capacidad de las herramientas de base de datos típicas utilizadas para capturar, almacenar, administrar y analizar información”. (p. 1)

Por su parte, González (2018) lo define como una herramienta que permite recolectar, almacenar y analizar gran cantidad de datos de diferentes tipos. Cada contacto con la tecnología deja una huella que permite recolectar toda la información y correlacionar con otros datos, por ejemplo, las páginas web que se visitan, los clics en las páginas, las búsquedas en internet, las interacciones en redes sociales, las compras en línea y la ubicación de las personas. De esta manera en las grandes empresas le dan uso comercial a esta información detallada y así poder direccionar mejor sus decisiones y acciones estratégicas (p. 211).

Teniendo en cuenta que los datos son la materia prima que se debe procesar y transformar para generar valor, se debe realizar una gestión de los datos los cuales se pueden englobar en tres: explotación, almacenamiento y producción (Chen y otros, 2014). El proceso de almacenamiento se debe acompañar de procesos propios de la gestión de la información como son la evaluación y el mantenimiento para garantizar la calidad de los datos almacenados para su recuperación. En la misma línea, Padgavankar y Gupta (2014) incluyen también el tratamiento de información.

En cuanto al proceso de producción, este debe producir información y conocimiento para la acción, base de la creación de valor para la organización. Esta cadena de producción se ha concretado en la cadena de valor de los datos masivos, que se compone de cuatro fases: generación, adquisición, almacenamiento y análisis de datos (García, M. 2017, p.18).

Los datos masivos también se han definido teniendo en cuenta otras características, además de su volumen, en la definición se han considerado aspectos como la variedad de datos y la velocidad de procesamiento (Mayor-Ríos et al. 2019). Sucesivamente otros autores, cuando profundizan en su estudio han ido añadiendo más V. Por lo tanto, a las tres V iniciales (volumen, variedad y velocidad) se añadieron otras características vinculadas con calidad de los datos, tratamiento y gestión, o prestaciones de tecnología y software. Estas otras características son: veracidad, valor, visualización, verificación, variabilidad y viabilidad.

A pesar de que la descripción de las características de los datos masivos a partir de las V es la más conocida, existen muchas otras. Todas ellas ejemplifican proliferación de definiciones propias de una

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

disciplina en formación (instrumentos, investigación, formación del corpus teórico y comunidad científica).(García, M. 2017, p.11)

El Big data tiene unas dimensiones que son sus características más importantes, Volumen, Variedad y Velocidad (Instituto de Ingeniería del conocimiento, 2016):

- **Volumen:** Cada día las empresas tienen un aumento significativo de sus datos, el volumen se refiere a la cantidad de datos que son generados cada segundo, minuto y días en nuestro entorno. Es la característica más asociada al Big Data, ya que hace referencia a las cantidades masivas de datos que se almacenan con la finalidad de procesar dicha información, transformando los datos en acciones.
- **Variedad:** La variedad se refiere a las formas, tipos y fuentes en las que se registran los datos. Estos datos pueden ser datos estructurados y fáciles de gestionar como son las bases de datos, o datos no estructurados, entre los que se incluyen documentos de texto, correos electrónicos, datos de sensores, audios, vídeos o imágenes que tenemos en nuestro dispositivo móvil, hasta publicaciones en nuestros perfiles de redes sociales, artículos que leemos en blogs, las secuencias de clic que se hacen en una misma página, formularios de registro e infinidad de acciones más que realizamos desde nuestro Smartphone, Tablet y ordenador.
- **Velocidad:** La velocidad se refiere a los datos en movimiento por las constantes interconexiones que se realizan, es decir, a la rapidez en la que son creados, almacenados y procesados en tiempo real.

Para los procesos en los que el tiempo resulta fundamental, tales como la detección de fraude en una transacción bancaria o la monitorización de un evento en redes sociales, estos tipos de datos deben estudiarse en tiempo real para que resulten útiles para el negocio y se consigan conclusiones efectivas.

Estas habían sido las 3 principales características, pero con el avance tecnológico que se ha venido presentando se han adicionado otras 4 características más (Instituto de Ingeniería del conocimiento, 2016):

- **Veracidad de los datos:** Se refiere a que la información obtenida debe ser fiable, por esta razón es importante invertir tiempo a través de métodos que permitan conseguir datos de calidad.
- **Viabilidad:** Es la capacidad de las organizaciones para realizar un uso eficaz de toda la información que guardan en sus bases de datos, mediante la selección, análisis y monitoreo, con el fin de conocer a todos los involucrados tanto internos como externos que permitan diseñar estrategias apropiadas.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

- **Visualización de los datos:** Es la forma en que se presentan los datos, para que sean entendibles y de fácil acceso, y así permita de una manera más sencilla encontrar patrones y relaciones.
- **Valor de los datos:** El valor se obtiene del proceso de transformarlos en información; los cuales permiten adquirir conocimiento importante, y esto le facilita a los altos directivos la toma de decisiones (Instituto de Ingeniería del conocimiento, 2016).

4.3 Data mining o Minería de datos

La minería de datos es un paso en el proceso KDD que consiste en la aplicación de análisis de datos y algoritmos de descubrimiento que, bajo limitaciones de eficiencia computacional aceptables, produce una enumeración particular de patrones sobre los datos. Tenga en cuenta que el espacio de los patrones es a menudo infinito, y la enumeración de patrones implica alguna forma de búsqueda en este espacio. Existen bases de datos en redes sociales, bancos, tiendas, hospitales y más. Para las empresas u organizaciones los datos son materia prima para poder encontrar patrones que favorezcan a interpretar fenómenos o sucesos, por ejemplo, un usuario desea saber si puede acceder a un préstamo en el banco, qué producto se vende más según temporadas o cuáles son las causas de una enfermedad. (Flores, Cadena, Quinatoa, Villa., 2019, p. 961).

El Data mining es una etapa dentro del proceso completo del descubrimiento del conocimiento, este intenta obtener patrones o modelos a través de los datos recopilados. (Valcárcel, 2004, p. 84).

La minería de datos puede apreciarse como una disciplina que combina técnicas de Inteligencia Artificial, Aprendizaje Computacional, Probabilidad, Estadística, y Bases de Datos para extraer información y conocimientos útiles desde grandes cantidades de datos. (Velarde, 2003).

4.3.1 ¿Para qué sirve la minería de datos?

Es necesario conocer las utilidades de la minería de datos, con el fin de encaminar el presente proyecto adecuadamente, por lo tanto, a continuación, se mencionan algunas de sus funcionalidades.

A nivel global se crean y almacenan una incontable cantidad de datos y continúa aumentando, lo que aumenta su potencial de obtener una visión clave de los negocios y cómo se pueden analizar y aplicar, lo importante es saber cómo utilizar la información que se analiza y qué soluciones puede aportar dicha información.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Cuando se combina el Big Data con la analítica se pueden aplicar en campos relacionados con negocios como:

- Determinar la causa de un problema
- Generar estrategias a partir del hábito de compra de los clientes
- Calcular riesgos
- Detectar fraudes que afecten a la organización. (SAS, s.f.)

Una empresa en posesión de una base de datos de calidad y tamaño suficiente puede emplear el Datamining para generar nuevas oportunidades de negocio, dada su capacidad para proporcionar (Garrido, 2001, p.500):

- **Predicción automática de comportamientos:** Generalmente se trata de problemas de clasificación, como por ejemplo se puede citar el marketing dirigido. Data Mining usa los resultados de campañas de marketing realizadas anteriormente para identificar el perfil de los clientes que son más propensos a comprar el producto y de este modo permitirá sustituir el correo masivo por el correo dirigido.
- **Predicción automática de tendencias:** Partiendo de base de datos históricas, Data mining creará un modelo para predecir las tendencias. Como ejemplo se puede citar la predicción de ventas en el futuro.
- **Descubrimiento automático de comportamientos desconocidos anteriormente:** Las herramientas Datamining de visualización y clustering, permiten “ver” los datos desde una perspectiva distinta y por ello nuevas relaciones entre ellos. (Garrido, 2001, p.500)

Dentro de los usos más conocidos de la minería de datos, se encuentran ejemplos como (Casas, Nin, & Julbe, 2019, p.37):

Redes sociales. Su uso que cada vez está más extendido provoca la tendencia a que sus usuarios incorporen cada vez más una gran parte de su actividad y la de sus conocidos. Las empresas utilizan toda esta información con muchas finalidades. Por ejemplo, para realizar estudios de marketing, evaluar su reputación o incluso cruzar los datos de los candidatos a un puesto de trabajo determinado.

Consumo. Amazon es líder en ventas cruzadas. Gran parte de su éxito se basa en el análisis masivo de datos de patrones de compra de un usuario cruzados con los datos de compra de otros, creando así anuncios personalizados y boletines electrónicos que incluyen justo aquello que el usuario quiere en ese instante.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Salud y medicina. En 2009, el mundo experimentó una pandemia de gripe A, también conocida como gripe porcina o H1N1. El website Google Flu Trends 12 fue capaz de predecir gracias a los resultados de las búsquedas de palabras clave en su buscador. Flu Trends usó los datos de las búsquedas de los usuarios que contienen influenza-like illness symptoms (‘Síntomas parecidos a la enfermedad de la gripe’) y los agregó según ubicación y fecha, siendo capaz de predecir la actividad de la gripe hasta con dos semanas más de antelación respecto a los sistemas tradicionales.

Política. Barak Obama fue el primer candidato a la presidencia de Estados Unidos en basar toda su campaña electoral en los análisis realizados por su equipo de big data. Este análisis ayudó a la victoria de Barak Obama frente al otro candidato republicano Mitt Romney con el 51,06 % de los votos, siendo esta una de las elecciones presidenciales más disputadas.

Telefonía. Las compañías de telefonía utilizan la información generada por los teléfonos móviles (posición GPS y los CDR)¹⁴ para estudios demográficos, planificación urbana, etc.

Finanzas. Los grandes bancos disponen de sistemas de trading algorítmico que analizan una gran cantidad de datos de todo tipo para decidir qué operaciones en bolsa son las más rentables en cada momento.

El Data Mining para el conocimiento de datos emplea herramientas. Se clasifican en 2 grandes grupos (Valcárcel, 2004, p. 84-85):

- **Técnicas de verificación:** en las que el sistema se limita a comprobar hipótesis suministradas por el usuario.
- **Métodos de descubrimiento:** en los que se han de encontrar patrones potencialmente interesantes de forma automática. Y los resultados de este pueden ser de carácter descriptivo o predictivo. Las predicciones sirven para prever el comportamiento futuro de algún tipo de entidad mientras que una descripción puede ayudar a su comprensión.

La aplicación de técnicas en Data Mining persiguen los siguientes resultados (Valcárcel, 2004, p. 84-85), tal como se muestra en la Ilustración 1:

- Clasificación
- Regresión
- Agrupamiento (Clustering)
- Resumen
- Modelado de dependencias
- Análisis de secuencias

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Taxonomía de técnicas de Minería de Datos

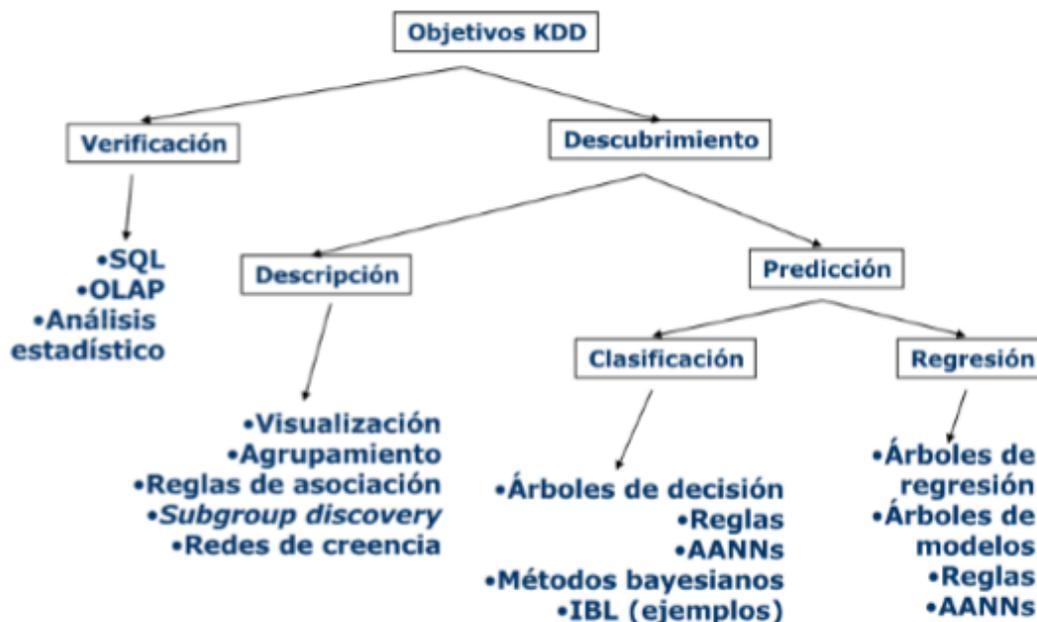


Ilustración 1: Árbol de clasificación Datamining. Fuente: (Escobar, Alcibar y Puris, 2016, p. 507)

Según la ilustración 1, las técnicas de minería de datos se dividen en dos: Verificación donde las herramientas más utilizadas son: SQL, OLAP y Análisis estadístico, y la segunda técnica es Descubrimiento, esta a su vez se divide en técnicas de Descripción (herramientas de visualización, Agrupamiento y reglas de asociación); y de Predicción (Árboles de decisión , Reglas, Árboles de regresión y Árboles de modelos) (Escobar, Alcibar y Puris, 2016, p. 507).

En el proceso de abordar un problema se sugiere tener en cuenta la realización de las etapas (Medina y Gómez, 2014, P. 32-33):

1. **Selección del modelo:** Con la estructuración del problema basado en el tipo de datos y objetivo que se quiere obtener, los datos se deben preparar y explorar de manera que se pueda llegar a la generación del modelo, con el cual se logra dar explicación al comportamiento de los datos, para posteriormente realizar su validación, implementación y actualización.
2. **Clasificación de los datos:** Agrupa las reglas de asociación, decisión, métodos, funciones, núcleos, vectores para descubrir las relaciones entre los atributos de un

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

conjunto de datos de acuerdo con un porcentaje de error permitido (umbral), descubriendo relaciones de asociación, segmentación, dependencias funcionales y no funcionales entre los diferentes atributos. A partir de la predefinición o búsqueda de las clases categóricas o de reglas, toma un dato y lo ubica dentro de un rango, haciendo para ello el uso de algoritmos.

3. **Predicción de valores en función de datos:** Basado en funciones matemáticas o estadísticas como:
 - a. La interpolación, la cual estima un valor dado dentro de los límites de una función, construida a partir de unos valores conocidos.
 - b. La extrapolación, la cual estima un valor dado por fuera de los límites de una función, construida a partir de unos valores conocidos.
 - c. Regresión: técnica estadística para crear relación de variable, dependiente o independiente, para que a partir de un atributo de entrada o un valor de predicción obtener el valor estimado o de salida de acuerdo a un valor de error permitido.
4. **Modelamiento de las dependencias para resolver el problema:** Usando los datos disponibles en el proceso de aprendizaje inductivo, deductivo, transducción y estimación, se predicen las dependencias desconocidas, con que se crea el modelo para el entrenamiento de los datos que satisfaga el porcentaje de error permitido y que genere el modelo nuevo.
5. **Descubrimiento de nuevas reglas:** Con el entrenamiento de datos, las dependencias o las asociaciones de los atributos de entrada con los de salida y sus predicciones, acompañadas del porcentaje de error permitido en las muestras de los datos, de acuerdo con el algoritmo seleccionado se obtienen nuevas reglas y condiciones que pueden hacer variar las entradas o descubrir nuevas dependencias o asociaciones entre los datos de la muestra.
6. **Visualización de resultados:** La información obtenida de la minería de datos se muestra para permitir sus análisis e interpretación, medir el alcance del objetivo, la comprensibilidad de los patrones extraídos y la facilidad del modelo. (Medina y Gómez, 2014, P. 32-33).

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Luego de abordar diferentes conceptos sobre el Big data, es importante comprender algunos términos básicos que permitan orientar la aplicación de estos, hacia el establecimiento de estrategias que lleven a las empresas a adquirir más clientes y/o a fidelizar los que ya se tienen.

4.4 Análisis de datos con Microsoft Excel

Actualmente se requiere de herramientas potentes que permitan desarrollo de investigaciones y el manejo de gran cantidad de datos y así mismo agilizar la obtención de resultados teniendo una mayor precisión en la información; una herramienta que puede permitir el manejo de dicha información es el paquete de Office de Windows y dentro de este el programa Microsoft Excel, el cual tiene una gama amplia de utilidades.

Se tiende a menospreciar infundadamente a Microsoft Excel como una herramienta para la investigación. Este software podría ser muy útil para este fin. Excel es básicamente una hoja de cálculo electrónica, pero es mucho más, es una opción muy buena para resolver la mayoría de las necesidades estadísticas de las investigaciones.

Excel presenta diferentes opciones que permiten el análisis y presentación de los datos de una investigación, a través de tablas dinámicas y análisis estadístico descriptivo e inferencial paramétrico de los datos, además se pueden construir gráficos a partir de los datos presentados en una hoja de cálculo o gráficos dinámicos, de gran utilidad para un mejor análisis y resumen de los datos.

También este programa cuenta con la posibilidad de desarrollar un enorme número de funciones matemáticas, económicas y estadísticas. (Perez, L, 2006, P. 68 - 70)

La hoja de cálculo Excel de Microsoft es una aplicación integrada en el entorno Windows que permite la realización de cálculos, así como la representación de estos valores de forma gráfica. A estas capacidades se suma la posibilidad de utilizarla como base de datos.

Ofrece una gran ayuda, siendo relativamente fácil la comprensión de las operaciones básicas, y contando, además, con funciones estadísticas que facilitan los distintos cálculos. Dispone, además de un abundante repertorio de figuras y proporciona una hoja de trabajo estándar para registro de datos que permite una rápida transferencia a otros programas estadísticos con una dificultad mínima (Bausela, E, 2005, P. 2).

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

4.5 Segmentación de clientes:

La segmentación es la práctica de clasificar a los clientes en distintos grupos, con base en su información multidimensional (características socio-demográficas, patrones de compra y uso, preferencias, necesidades y actitudes). Permite identificar los factores que mejor describen el comportamiento de cada grupo o segmento, y que los diferencian de los demás.

Un ejercicio de segmentación robusto comúnmente se apoya en modelos de analítica predictiva, que permiten descubrir patrones de similitud y diferencia entre los clientes. El uso de dichos modelos asegura que la segmentación es objetiva, permite sintetizar una gran cantidad de datos, y permite encontrar patrones que no se distinguen a simple vista. (Shimada, s.f)

4.6 Marketing

La actual publicidad digital es parte de un complejo proceso de comunicación donde usuarios, públicos, audiencias y consumidores conviven en una incansable demanda y oferta de novedosas experiencias relacionadas con multitud de marcas y productos. Como los procesos de gestión, tanto de marketing como de comunicación comercial, una constante transformación que los hace más tecnológicos, directos y participativos. (Martínez, E. y Nicolás, M. 2016, p. 10).

Principales características del marketing bancario:

Dado que los bancos hacen cientos de campañas de publicidad al año, los empleados de marketing tienen que maximizar y optimizar sus resultados en cada etapa del proceso. El tipo de marketing sistemático y beneficioso tiene cuatro características básicas:

- **Planear** las campañas de marketing más efectivas y las estrategias a seguir.
- **Fijar** un objetivo en las actividades de la campaña dirigido a los segmentos definidos del mercado con propensión a comprar nuestros servicios.
- **Actuar** de acuerdo con los planes, con herramientas de gestión automatizadas de marketing, como módulos creadores de listas, generación de e-mails, etc.
- **Aprendizaje** a partir de la experiencia de la campaña, midiendo los resultados y dando automáticamente una visión sobre el sistema para mejorar en el futuro.

En general, se puede afirmar que el modo más utilizado por las entidades bancarias es el marketing directo, que es personalizado e individualizado. En un sector tan competitivo como es el bancario o

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

financiero, el valor de la información es doblemente importante. Con un entorno en continuo desarrollo y cambio, con los clientes con mayores niveles de exigencia y más volátiles, las decisiones de los gestores de las entidades bancarias han de ser certeras al milímetro para el éxito de sus empresas. Y esto solo es posible a través de un análisis minucioso de la mayor cantidad posible de información relevante.

Por ello, el intercambio de información entre entidades financieras se vuelve cada vez más necesario en el sector. Estos ficheros de información interna permiten desarrollar bases de datos relacionales, que dotan a las entidades de un mejor conocimiento del mercado y de sus clientes, lo que es positivo para la segmentación. (Rivera, J. y Garcillán, M, 2014, p. 24-25).

4.7 Estrategias de marketing

Las estrategias de marketing, también conocidas como estrategias de mercadotecnia, estrategias de mercadeo o estrategias comerciales, son acciones meditadas que se llevan a cabo para alcanzar objetivos de marketing, tales como aumentar las ventas o lograr una mayor participación de mercado.

Como todo elemento del marketing, las estrategias de marketing comprenden los cuatro elementos que conforman la mezcla de marketing, por lo que se dividen o clasifican en estrategias para el producto, estrategias para el precio, estrategias para la distribución, y estrategias para la promoción.

Al permitir alcanzar los objetivos de marketing y comprender las decisiones y acciones relacionadas con el producto, el precio, la distribución y la promoción, las estrategias de marketing suelen ser las estrategias más utilizadas en una empresa, y las más determinantes a la hora de generar las ventas y las utilidades.

Estas suelen ser formuladas anualmente o siempre que se consideren necesarias por los directivos o gerentes de marketing, o los encargados del marketing (mercadólogos), ser parte de la planeación del marketing, y estar especificadas en el plan de marketing de la empresa. (Kamiya, A, 2019).

Cuando una empresa tiene un cliente activo, debe realizar un proceso de incrementar ventas, además realizar un proceso de retención de clientes.

Las empresas difieren en sus estrategias de marketing. Sin embargo, es posible identificar tres tipos de estrategias de marketing (Farías, P, 2014, p. 10):

- **Estrategia pasiva:** Las empresas que siguen una estrategia pasiva realizan las tácticas comúnmente usadas en marketing y esperan que la mayor cantidad de su mercado

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

objetivo responda a esos esfuerzos de marketing, los esfuerzos de marketing más rentables son mantenidos por la empresa, y los menos rentables son abandonados (enfoque prueba y error).

- **Estrategia activa:** Las empresas que siguen una estrategia activa, por medio de la experiencia, juicio y conocimiento de los administradores, identifican a sus mejores clientes y los satisfacen de acuerdo con sus características. Por ejemplo, para adquirir clientes identifican y perfilan los mejores prospectos y diseñan una oferta para atraerlos. Para incrementar ventas, el administrador con base en su experiencia identifica a los potenciales clientes y los potenciales productos a ofrecerles. Para retener clientes, la empresa identifica a los potenciales desertores, y luego realiza los esfuerzos de marketing necesarios para retenerlos.
- **Estrategia científica:** Las empresas siguen una estrategia científica cuando usan modelos matemáticos para seleccionar al cliente. Por ejemplo, para adquirir clientes examinan los prospectos y proyectan sus patrones de comportamiento. Para incrementar ventas, mediante modelos matemáticos seleccionan a los clientes más receptivos a los esfuerzos para incrementar ventas. Para la retención de clientes, la empresa identifica potenciales desertores a través de modelos matemáticos.

Las estrategias de marketing son acciones que se ejecutan para alcanzar los objetivos de comercialización. Dichas estrategias se clasifican en estrategias para el producto, estrategias para el precio, estrategias para la distribución, y estrategias para la promoción y se redefinen anualmente según sean los objetivos de la organización.

Dentro de las estrategias de marketing se pueden diferenciar tres tipos principales: como la estrategia pasiva, que busca que el mercado objetivo responda a las campañas de marketing y así mantener la estrategia más rentable para la empresa. En segundo lugar se encuentra la estrategia activa, que se basa en identificar sus clientes potenciales y diseñar ofertas para atraerlos y retener a los posibles desertores. Por último está la estrategia científica, donde se aplican modelos matemáticos para seleccionar a los clientes, por ejemplo, analizando sus hábitos de comportamiento para incrementar sus ventas (Farías, P, 2014, p. 10).

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

4.7.1 Estrategias de marketing Mix

El objetivo de los empresarios en las operaciones de marketing es encontrar una oferta de productos y servicios que incremente el nivel de satisfacción del comprador y genere utilidades para la empresa. Esta oferta está conformada, desde el punto de vista mercadotécnico, por un conjunto de variables controlables por la empresa, denominadas comúnmente mezcla de mercadotecnia o, en su expresión en inglés, marketing mix, a saber: producto (product), distribución (distribution), comunicación (communication) y precio (price). En este mix se incorpora el servicio (service) como quinto elemento, por considerar que el paquete de valor ofrecido está compuesto por beneficios donde el servicio es un componente de valor agregado que puede constituir una ventaja competitiva, toda vez que los clientes actúan estimulados por la búsqueda de valor (Peñaloza, 2005, p. 74).

4.7.1.1 Decisiones sobre el Producto/Servicio

En términos generales, producto es todo lo que pueda ser mercadeado. Está conformado por un conjunto de elementos tangibles e intangibles, como son el producto básico, la marca, el empaque, la etiqueta, la calidad y los servicios de apoyo que en un todo armónico contribuyen a satisfacer las necesidades y las expectativas del comprador (Ibarra-Morales et al. 2020).

Al decidir sobre los componentes que se adicionan al producto básico, como la marca, el envase, el logo y los colores, la empresa está creando medios para posicionarse en la mente del comprador y, al mismo tiempo, modos para facilitar y minimizar los esfuerzos que realiza el individuo en sus procesos de búsqueda de información, de evaluación de alternativas y de la compra propiamente dicha. (Peñaloza, 2005, p. 75).

4.7.1.2 Decisiones sobre Distribución

La distribución, entendida como el conjunto de actividades, organizaciones y personas que hacen posible que el bien llegue al consumidor final o al usuario organizacional, representa una decisión estratégica de trascendencia, pues compromete a la empresa a largo plazo.

En tal sentido, la gestión del elemento distribución ha adquirido enorme importancia, pues patentiza el servicio al cliente, al proveer beneficios sustanciales y medibles que el comprador valora en la transacción. (Peñaloza, 2005, p. 76).

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

4.7.1.3 Decisiones sobre Organización

Esta dimensión incluye decisiones relativas al diseño del canal, uso de los intermediarios, amplitud del canal, selección del canal, control sobre la cadena de distribución, características del mercado, del producto y de la empresa; la estrategia comercial, el análisis de la competencia, las características del entorno, así como los propios objetivos que la empresa pretende alcanzar con la distribución (Peñaloza, 2005, p. 76).

4.7.1.4 Decisiones sobre Comunicación

Por comunicación se entiende el proceso mediante el cual un individuo transmite estímulos para modificar la conducta de otros. Uno de sus objetivos básicos es informar al mercado-objetivo sobre la oferta de la empresa, lo que se traduce en facilidades para el comprador, al conocer la propuesta de la empresa, los lugares de venta, los precios de la misma, etc. (Peñaloza, 2005, p. 77).

4.7.1.5 Decisiones sobre Precio

Este componente de la mezcla de mercadotecnia es vital para la permanencia y éxito de la empresa, por cuanto es el elemento del mix que por excelencia, a través de las cantidades vendidas proporciona los ingresos necesarios para el funcionamiento y la consecución de los fines de rentabilidad de la empresa (Peñaloza, 2005, p. 78).

4.7.1.6 Decisiones sobre Servicio

Se entiende al servicio como un conjunto de prestaciones realizadas por la organización que, en comunión con los demás componentes de la mixtura, se realizan antes, en el momento y después de la venta del bien, agregando valor e incrementando el nivel de satisfacción del cliente.

De manera que, el elemento servicio dentro de la mezcla, es un componente integrador, pues está presente en todo el proceso mercadotécnico y constituye, junto al producto, el precio, la comunicación y la distribución, los medios que la empresa, sobre todo la orientada hacia el mercado, utiliza para proporcionar satisfacción y bienestar a los consumidores (Peñaloza, 2005, p. 79).

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

4.8 Fidelización de clientes

La fidelización de clientes pretende que los compradores o usuarios de los servicios de la empresa mantengan relaciones comerciales estables y continuas, o de largo plazo con ésta. La fidelidad se produce cuando existe una correspondencia favorable entre la actitud del individuo frente a la organización y su comportamiento de compra de los productos y servicios de la misma. Constituye la situación ideal tanto para la empresa como para el cliente. El cliente es fiel, «amigo de la empresa» y, muy a menudo, actúa como «prescriptor» de la compañía (Apaolaza, Forcada, y Hartmann, 2002).

Un cliente fiel es aquel que:

- Regularmente compra el producto o utiliza el servicio,
- Le gusta realmente la organización y piensa muy bien acerca de ella, y
- Nunca ha considerado usar otro proveedor para ese servicio.

Este concepto se relaciona con la habitualidad del cliente para realizar una compra o usar un servicio, lo cual tiene una relación directa con su nivel de satisfacción, ya que un alto grado de satisfacción convierte la repetición en fidelización; toda herramienta o práctica que mejore el nivel de satisfacción facilita la consecución de la lealtad del cliente y, por ende, su fidelización.

En la fidelización se persigue acaparar la atención del cliente y desplazar a cualquier competidor por medio de la diferenciación del producto o servicio de acuerdo con las necesidades del cliente, el valor agregado que perciba el cliente, las relaciones públicas o cualquier otra técnica de fidelización (Figuroa, V, 2011, p. 30).

El objetivo de la fidelización de clientes es mantener una relación a largo plazo con el consumidor y a su vez que éste recomiende el producto o servicio a otras personas; la fidelización del cliente está directamente relacionada con el grado de satisfacción acerca de cierto producto y es determinante para desplazar a los demás competidores y seguir conservando la relación con el cliente.

Según Cabrera (2013), en el marketing actual, se entiende que la fidelización implica establecer sólidos vínculos y el mantenimiento a largo plazo de las relaciones con los clientes. Lo que trajo aparejado un desplazamiento de un marketing centrado en el corto plazo, a un marketing con un enfoque estratégico (Cabrera, S, 2013, p. 155).

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

4.9 Toma de decisiones

Una decisión no es un hecho simple unitario, sino el producto de un proceso social complejo generalmente extendido sobre un considerable periodo de tiempo. En este sentido la toma de decisiones incluye procesos de atención, dirección y de inteligencia que determinan las oportunidades de la decisión. Procesos para descubrir y diseñar posibles cursos de acción, y procesos para evaluar alternativas y seleccionar entre ellas. (Rubio, G. y Fierro, A, 2014, p. 84-85).

Su propósito está asociado a disminuir los riesgos organizacionales, solucionar problemas y aprovechar oportunidades (Rodríguez, Y, 2015, p. 150; Londoño-Patiño, 2020).

El proceso de toma de decisiones consistente está formado por un conjunto de fases que las empresas deben seguir para incrementar la probabilidad de que sus elecciones sean lógicas y óptimas. Este proceso puede tomar más de una vertiente, pero usualmente se realiza en siete pasos:(Franklin, E, 2011, p. 116):

- Diagnosticar y definir el problema.
- Establecer metas.
- Buscar soluciones alternativas.
- Comparar y evaluar las soluciones alternativas.
- Elegir entre soluciones alternativas.
- Implementar la solución seleccionada.
- Dar seguimiento y controlar resultados.

Actualmente el gran flujo de información ha permitido que los sistemas de información sirvan de soporte para la toma de decisiones, constituyen un conjunto de componentes interrelacionados para reunir, procesar, almacenar y distribuir información para apoyar la toma de decisiones, la coordinación, el control, el análisis y la visualización de una organización. (Rodríguez, Y, 2015, p. 155).

5. Toma de decisiones estratégicas

Según Quinn, la decisión estratégica es un proceso fragmentado que se da dentro de las distintas unidades organizativas a medida que se adaptan de forma reactiva a un entorno cambiante. (Navas, J, Guerras, L y Montero, A, 2010, p.192).

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Las decisiones estratégicas son generalmente sobre los problemas más externos de la organización y no de los internos, potenciando los productos de la organización en el mercado. (Rodríguez, Y, 2015, p. 152).

Lo estratégico es lo global, integrado por decisiones que afectan a la empresa como totalidad y está relacionado con respuestas integrales frente a las realidades del medio externo; lanzar una nueva línea de producto, o propiciar el desarrollo de una diferenciación por servicio al cliente, son decisiones estratégicas. (Mendoza, J, 2011, p. 12).

5. METODOLOGÍA

En primer lugar, se realiza una investigación para sentar las bases teóricas del trabajo, para ello se investigó y se recolectó información relacionada con el tema por medio de: fuentes secundarias, como los sitios web de las entidades especializadas en el tema, libros y artículos de Revistas. Y como fuentes primarias se gestiona permisos con la Entidad Bancaria BANCOLOMBIA para acceder a la base de datos de la sección Analítica Personas. Se define el tipo de Investigación como: Investigación cuantitativa correlacional y descriptiva.

Hoy en día con la nueva revolución tecnológica se han implementado muchos métodos de innovación, optimización del tiempo y recursos tecnológicos, encontrando que hay más tendencia a hacer desarrollos de implementación de sistemas de información. Por tal motivo, las organizaciones deben incorporar a sus áreas nuevas metodologías y herramientas a desarrollar, una de ellas es el Big Data, la Minería de datos entre otros, que actualmente permiten abordar la información de los datos de las personas con una mayor agilidad, sin olvidar que también trae asociado la actualización de datos, sumando a ello, una forma de obtener información que ayuda a las entidades a descubrir cuáles son sus preferencias, gustos y lugares más visitados, para así enviar información constante de interés y mantener una relación activa entre ambas partes.

Con este trabajo se quiere llevar a cabo un sistema más directo de la información que se le envía al cliente y personalizar sus preferencias. El enfoque de la investigación será descriptivo y correlacional, ya que a través de variables se pretende describir una tendencia o patrón de una población, y adicional se permitirá realizar predicciones por medio de la relación de dichas variables y de esta manera presentar resultados que puedan aportar información relevante para la toma de decisiones.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

La información se obtendrá de la entidad financiera Bancolombia, dada la cantidad de datos almacenados se busca extraer la información más útil con la cual se podrá realizar los respectivos análisis de datos, y para la construcción del modelo se utilizará como insumo el Excel, el cual requiere una aplicación ordenada y eficiente de las técnicas para que puedan ser adaptadas al desarrollo del proyecto.

La base de datos pertenece al área de Analítica de Personas de Bancolombia, en la cual se encuentran registradas las transacciones realizadas por los clientes con tarjeta débito y crédito entre las fechas 01/01/2020 y 30/06/2020, con dicha información se busca identificar las preferencias y frecuencia de compra de los clientes y cómo se da en diferentes edades y estratos socioeconómicos, con base en los resultados del análisis de la base de datos, se pretende orientar las estrategias de marketing generadas actualmente por el banco con el fin de desarrollar estrategias más personalizadas hacia los clientes, que permitan identificar sus necesidades y de esta manera llegarles de una forma más acertada y eficaz, generando mayores ingresos para el banco y así mismo fidelizar al cliente.

Dicha base contiene 504.914 registros, una dimensión de 4.544.235 registros, y como se evidencia en la tabla 1, abarca las siguientes columnas:

Tabla 1. Definición de variables de la base de datos Bancolombia

Variables	Descripción	Tipo de variable
Id mask	Es el número de identificación del cliente, el cual se encuentra enmascarado por seguridad de la información	Numérica
Género	Género del cliente F (Femenino), M (Masculino), N (Ninguno)	Alfabética
Edad	Tiempo transcurrido desde el nacimiento	Numérica
Segmento	Es grupo de clientes con características específicas	Alfanumérica
Estrato	Característica de lugares similares dentro una jerarquía social	Numérica
Día, Mes , año	Indica el día, mes y año en que	Fecha

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

	se realizó una transacción	
Día de compra	Día de la semana en que se realiza la compra	Alfabética
Valor pesos	Valor de las transacciones realizadas por cada cliente	Numérica
Descripción	Se explica el tipo de transacción y lugar donde se realizó la compra	Alfanumérica
Descripción Código Categoría	Tipo de compra realizada por el cliente	Alfanumérica
Descripción Categoría	Tipo de categoría a la que pertenece la compra realizada por el cliente	Alfabética
Código tipo TRX	Código con el cual se identifica la transacción Débito (E) o crédito (T)	Binaria

Se trabajará con 2 bases de datos debido a que la columna Estrato no cuenta con toda la información, la primer base contiene todos los registros 504.914 pero la columna Estrato tiene campos vacíos, con esta se tendrá mayor precisión en los resultados al contener un mayor número de información y de variables, la segunda base consta de 8.188 registros donde dicha columna tiene la información completa, cuenta con menos información pero se pueden establecer relaciones y dependencias importantes para el desarrollo del trabajo.

Se utilizará el modelo metodológico CRISP-DM que permitirá abordar los problemas que se presenten al momento de realizar el análisis de la información, dado que este se puede trabajar con minería de datos y es un modelamiento por fases que se adapta a las necesidades del presente trabajo.

La metodología propuesta para el desarrollo de esta investigación es la presentada en la ilustración 2.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

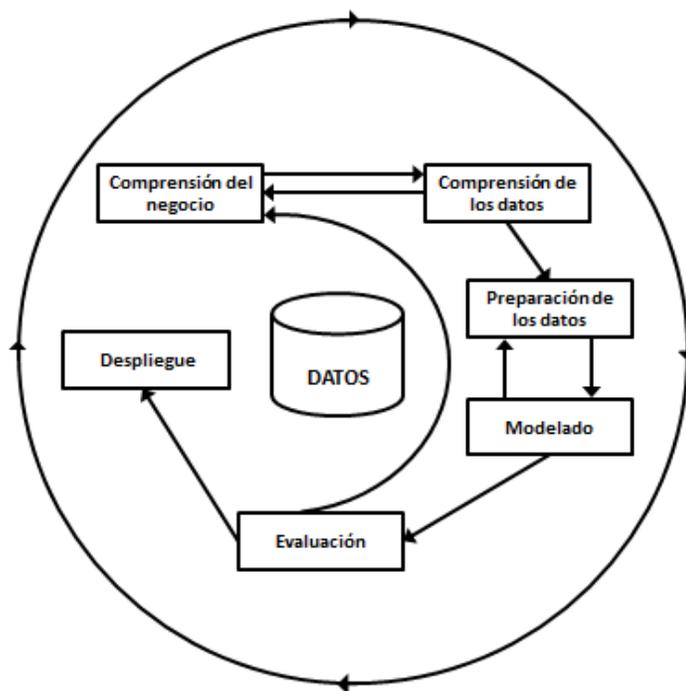


Ilustración 2: Fases de Metodología Crisp DM. Fuente: (Peralta, F. 2014, p 278)

La ficha técnica de la investigación es la presentada a continuación en la tabla 2:

Tabla 2. Ficha técnica de investigación. **Fuente:** Elaboración propia.

Universo	Base de datos de personas con tarjeta débito o crédito de Bancolombia.
Ámbito geográfico	Nacional
Lugar de aplicación	Entidad Financiera Bancolombia
Periodo de recogida de la información	6 últimos meses
Técnicas de análisis	Cuantitativo, Correlacional y descriptivo
Herramienta utilizada	Microsoft Excel versión 2016

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

6. RESULTADOS

Al realizar el análisis de la base de datos completa que contiene 504.914 registros suministrados por Bancolombia, se halla en la ilustración 3 que el 92% de las transacciones se realizan con tarjeta débito:

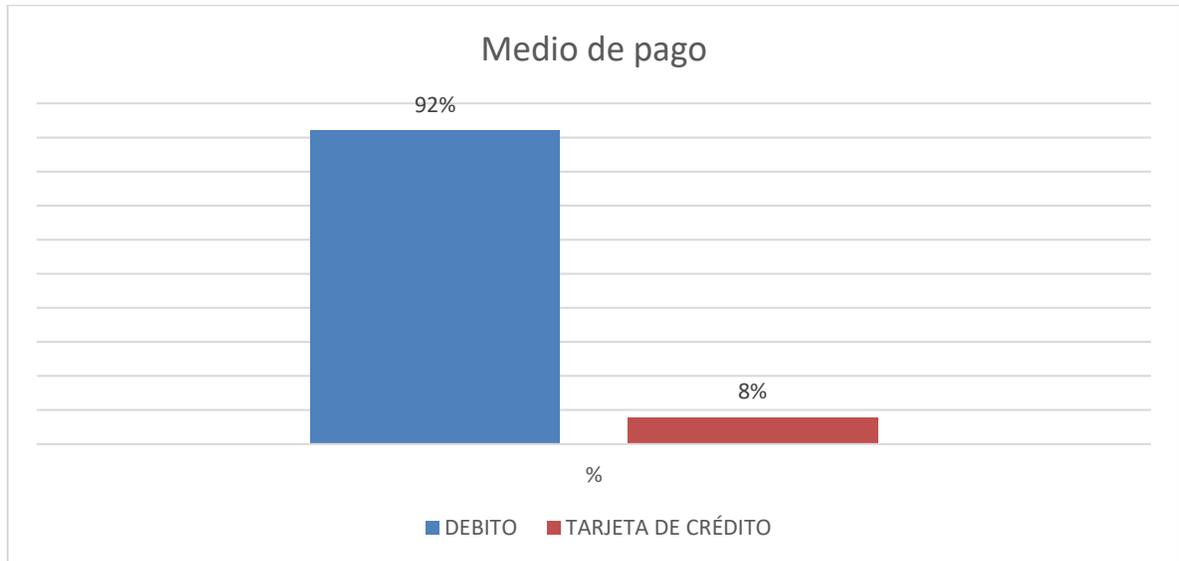


Ilustración 3. Porcentaje medios de pago

También se encuentra que la mayor parte de las transacciones realizadas son por parte de las mujeres, sin embargo, los hombres tienen un comportamiento muy similar en términos de cantidad, ya que al analizar las transacciones por montos este comportamiento cambia, siendo los hombres quienes generan compras por mayor valor, como se puede ver en la Ilustración 4.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

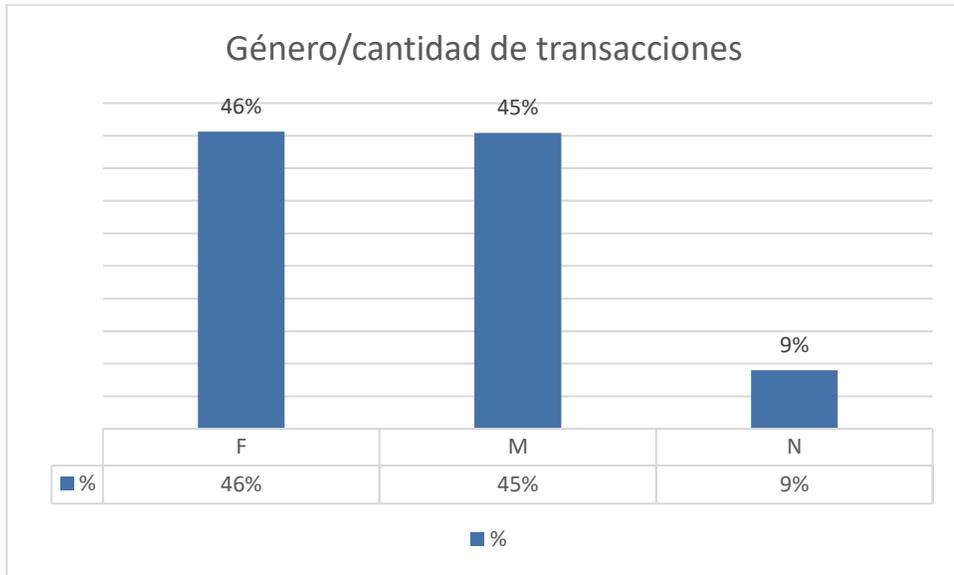


Ilustración 4. Cantidad de transacciones por género.

Según la categoría de compra tal como se muestra en la Ilustración 5, la mayoría de estas pertenecen a “Otros”, que representa el 59% de las transacciones, y es donde se agrupan los sectores de Vestuario, Restaurante, Belleza y Diversión.

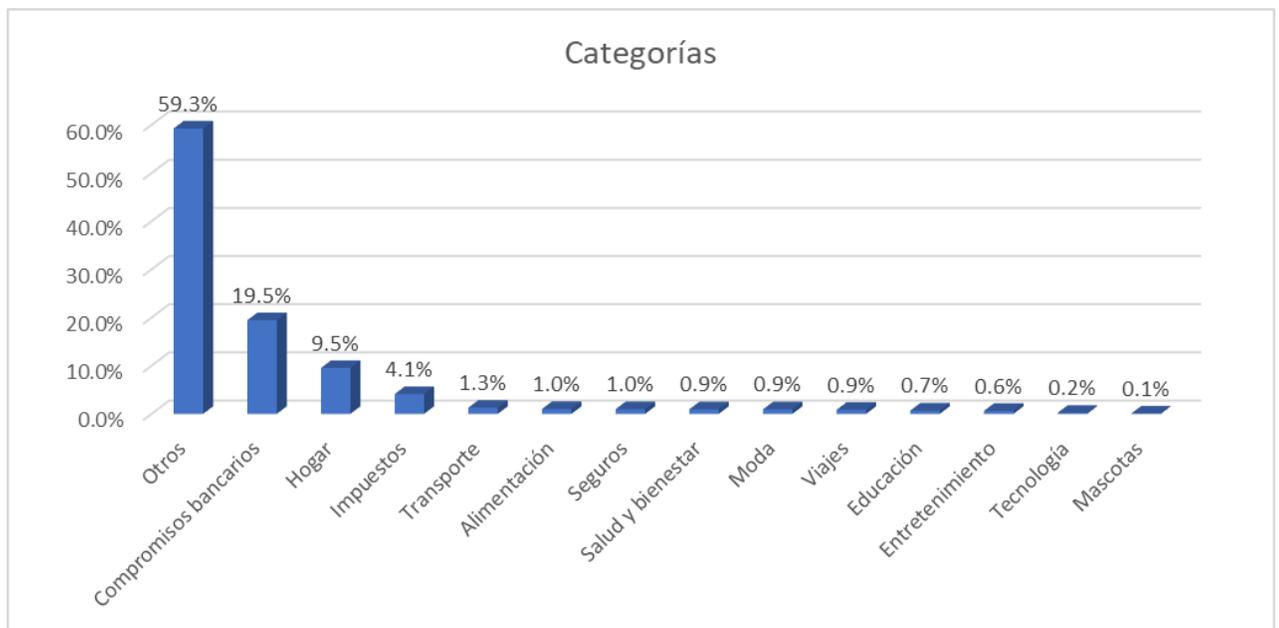


Ilustración 5. Porcentaje de compras por categorías.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Además, como se observa en la Ilustración 6, el mes en el cual se realizaron más transacciones fue en el mes de enero con un 19%, seguido de febrero con un 18%, teniendo que dicha base es del primer semestre del año 2020.

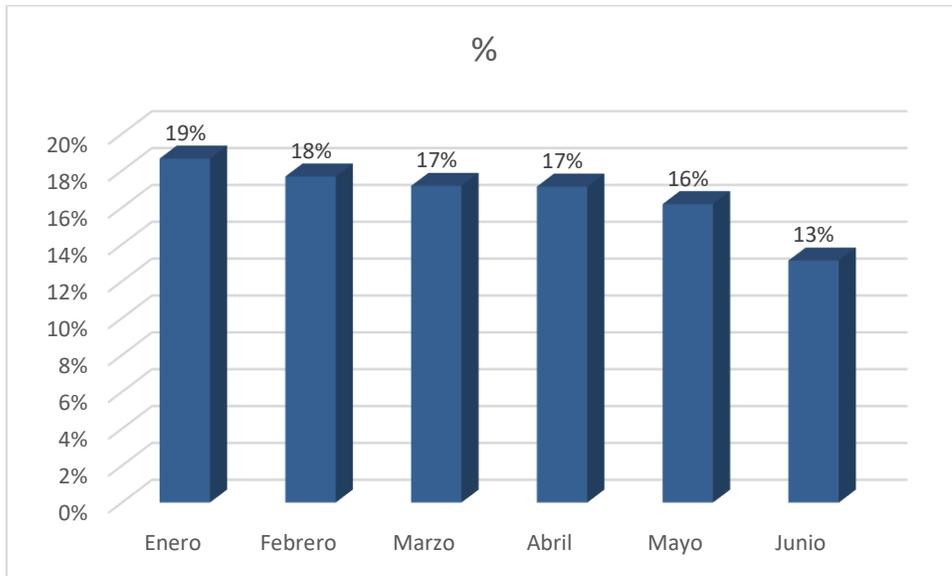


Ilustración 6. Porcentajes compra por meses

Se analizan los días de mayor frecuencia de compra, y se evidencia en la ilustración 7, que el martes es cuando más transacciones se generan representando un 19% del total de las compras semanales.

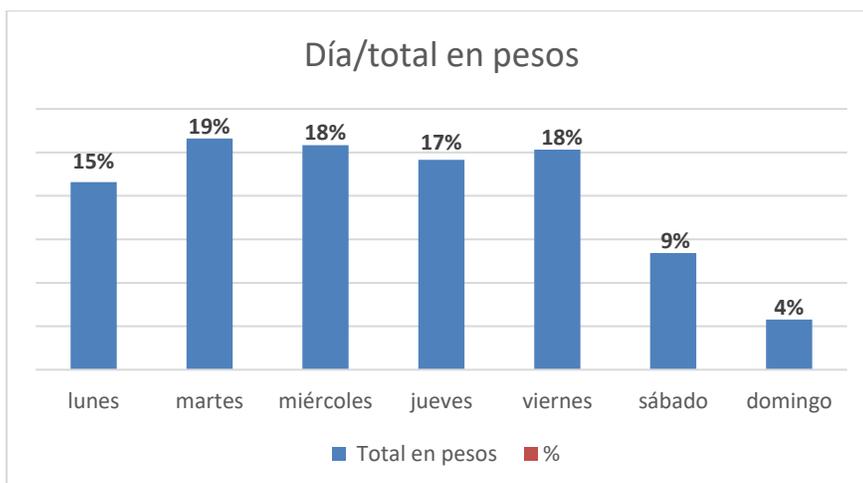


Ilustración 7. Porcentaje de compras por días.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Luego, se tomaron los rangos de edades de la base de datos, y se halla que las personas con edades entre 27 a 37 años conforman la mayor parte de clientes del banco representando un 37,5%, seguido por las personas con edades entre 38 a 48 años representados por un 24,3% del total de la base, y asimismo son quienes realizan compras por mayor valor, como se ve en la Ilustración 8:

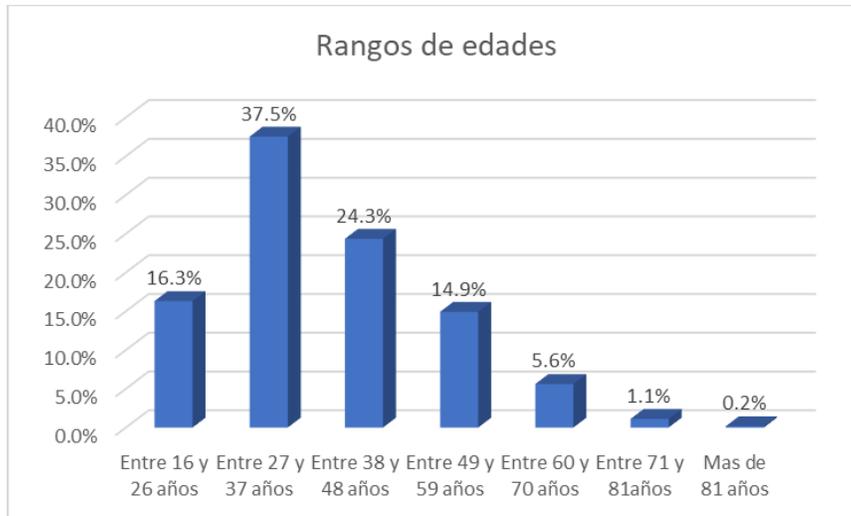


Ilustración 8. Porcentaje de compras por rangos de edades.

Así mismo se valida por rangos de edad, y según la Ilustración 9 se halla que las personas entre 16 a 37 años realizan la mayor parte de transacciones los días miércoles, mientras que los que están en un rango mayor a 37 años generan sus compras los días martes.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

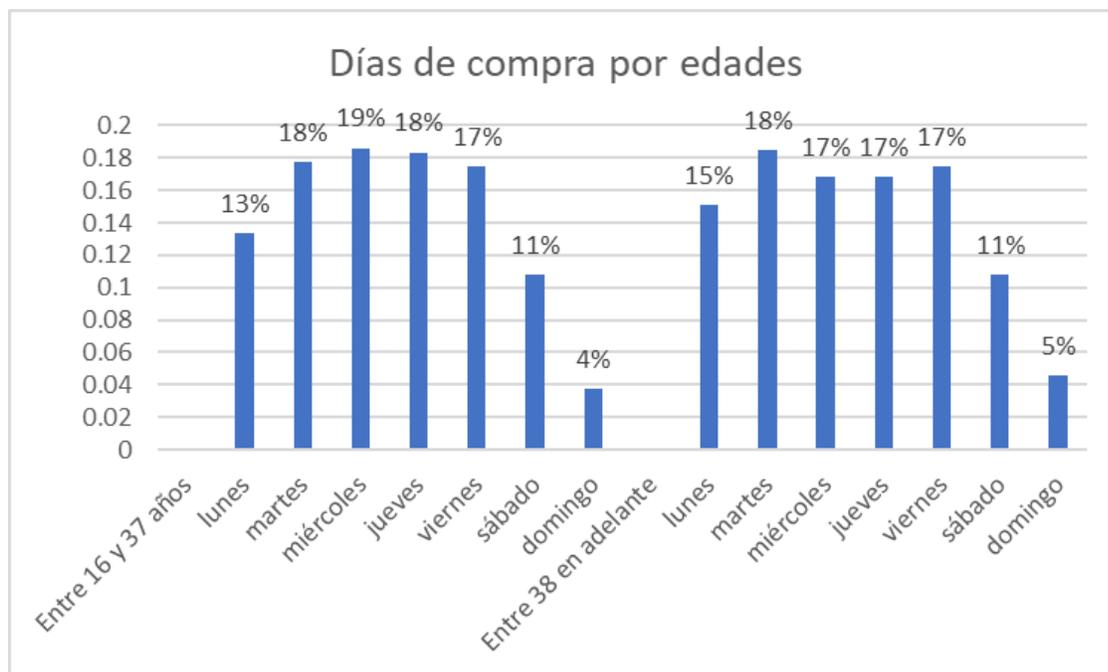


Ilustración 9. Porcentaje de compras por días de la semana.

Se analiza los días de compra y se evidencia según la Ilustración 10, que las fechas de mayor compra son los 2 y 15 de cada mes, donde es muy claro que la tendencia de compra son los días posteriores a los pagos de las quincenas.



Ilustración 10. Compras por días del mes

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Luego de estudiar la base de datos que contiene los campos de la columna estrato completos, y que cuenta con 8.188 registros se encuentra que el estrato 4 es donde realizan compras por mayores montos que representan un 38,2% del total de la base de datos, como se puede observar en la Ilustración 10.

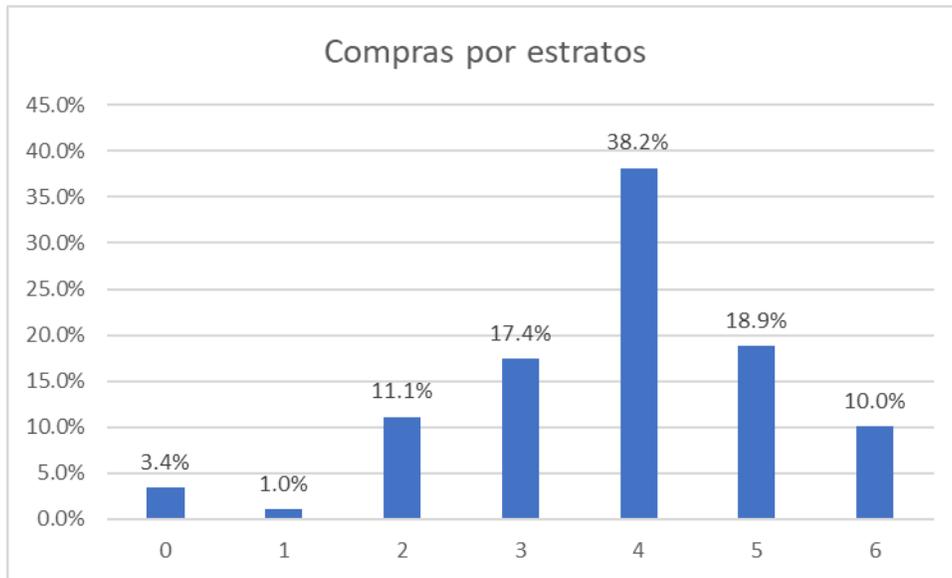


Ilustración 11. Porcentaje de compras por estratos socioeconómicos.

Dado que el estrato 4 es donde más transacciones se realizan, el trabajo se enfocará en dicho estrato, dirigiendo las estrategias de marketing para aumentar su consumo según las necesidades identificadas, los demás estratos también son importantes, pero por cuestiones de tiempo y de lo extenso que sería el trabajo, serán abarcados más adelante por el banco.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

Basados en la Ilustración 11, se puede determinar que en el estrato 4 se genera mayor cantidad de transacciones los días miércoles, representados por un 25%.

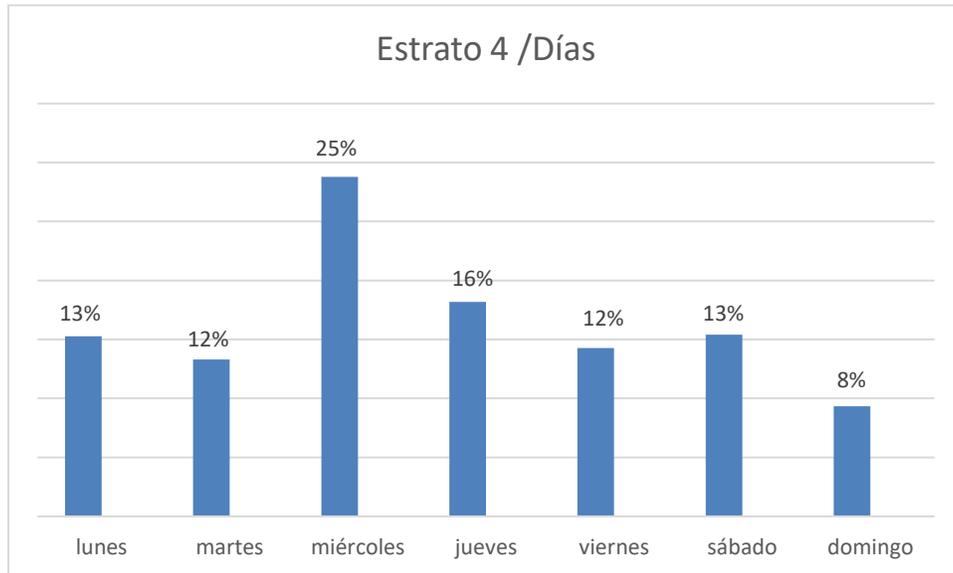


Ilustración 12. Porcentaje de compras por días en el estrato 4.

Y se evidencia que también las personas que pertenecen a este estrato generan la mayor parte de transacciones en la categoría de “Otros”, como se observa en la Ilustración 12.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

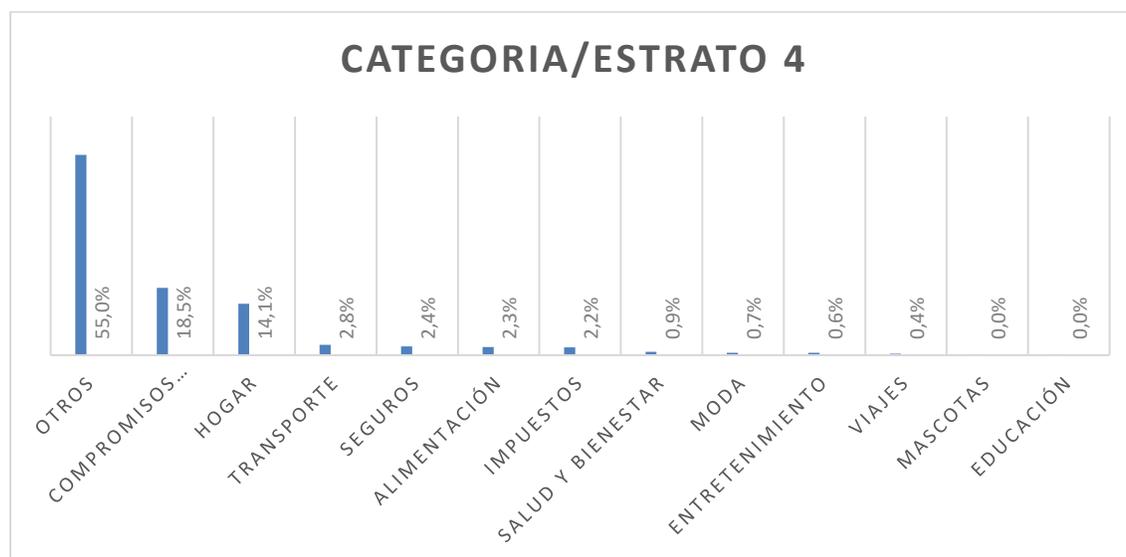


Ilustración 13. Porcentaje de compras por categoría en el estrato 4.

Al revisar los rangos de edades del estrato 4, se evidencia que la mayoría de los usuarios están en edades entre 38 y 59 años. Ver Ilustración 13.

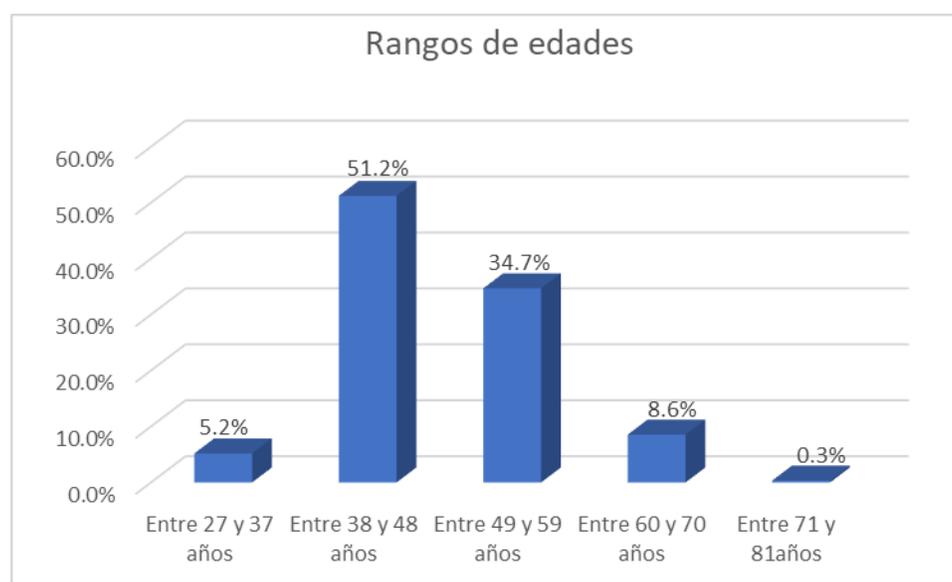


Ilustración 14. Porcentaje de usuarios en rangos de edades.

Al realizar un modelo de regresión lineal para las variables Edad y Valor pesos, con el fin de encontrar si existe una relación entre estas, se pudo determinar que no existe una relación lineal, tal como se muestra

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

en la tabla 3, por lo tanto, la ecuación resultante $y = mx + b$ no es confiable predecir la relación según la edad, dado que los coeficientes de relación tienden a 0, esto se debe a que al efectuar el estudio de la base de datos se pudo evidenciar que las personas con edades inferiores a 25 años y mayores a 60 realizan menos compras, por tal motivo se en vez de generar una tendencia lineal su tendencia es similar a una parábola.

Se llevaron a cabo tres procesos de relaciones lineales entre variables: Edades vs Valor compra, Estratos vs Valor compra y Edades vs Estratos. En la tabla 3 de acuerdo con los resultados de los coeficientes de correlación se puede evidenciar que no existen relaciones lineales entre las variables y se explican poco las variables.

Tabla 3. Coeficientes de regresión

<i>Estadísticas de la regresión</i>	EDADES/VALOR	ESTRATO/VALOR	EDADES/ESTRATO
Coefficiente de correlación	0.007587709	0.010359171	0.050625927
Coefficiente de determinación R ²	5.75733E-05	0.000107312	0.002562984

De acuerdo a la tabla 4, se presentan las pruebas de significancia de la regresión, se puede observar que, para grupo, al ser el indicador F mayor que el valor crítico de F, se puede concluir que las regresiones son significativas.

Tabla 4. Análisis de varianza

ANÁLISIS DE VARIANZA		
	F	Valor crítico de F
EDADES/VALOR	29.07119834	6.97977E-08
ESTRATO/VALOR	0.87844654	0.348655167
EDADES ESTRATO	21.03193249	4.58403E-06

Finalmente, la tabla 5, presenta los coeficientes interceptos y pendientes para cada de las relaciones.

Tabla 5. Coeficientes

Coeficientes	EDADES/VALOR	ESTRATO/VALOR	EDADES ESTRATO
Intercepción	38.57835469	3.76570614	40.12057981
Variable X 1	3.34845E-08	2.6793E-08	0.400518861

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

A partir de los resultados hallados del análisis de datos para el estrato 4, se puede percibir que algunas estrategias que se puede utilizar son las siguientes: Marketing diferenciado realizando una segmentación la cual puede ser geográfica, demográfica, conductual o psicográfica, y también se puede usar el Marketing Mix abordando cada variable con el fin de promover el uso de las tarjetas del banco.

7. CONCLUSIONES Y RECOMENDACIONES

En conclusión, Bancolombia debe establecer estrategias de marketing de segmentación diferenciada, a través de alianzas estratégicas con sectores de vestuario, restaurantes, belleza y diversión, y adicional enfocarse en las personas que pertenezcan a estrato 4, generando campañas especiales los días martes y miércoles que son los días de mayor índice de compras, ofreciendo descuentos al hacer uso de las tarjetas del banco, con esto se buscará aumentar el número y volumen de las compras de este segmento, siendo un beneficio tanto para el aliado como para Bancolombia; aunque actualmente ya se ofrecen este tipo de descuentos, no se encuentra focalizado por segmentos ni por sectores, sino que van dirigidos al público en general ofreciendo tal vez descuentos en artículos que no son de su interés, lo que causa que no sea atractivo y el cliente efectúe ninguna compra; adicional se debe orientar dichas campañas a las personas que se encuentren en edades entre 27 y 48 años. Esto permitirá llegar directamente al grupo de clientes que más compras realizan y así atraer su atención para que generen nuevas compras en los sectores y establecimientos que son de su gusto.

Por ende, se pudo determinar los patrones de compra y definir el tipo de características demográficas y socioeconómicas que actualmente son más frecuentes en los clientes de Bancolombia, y con esto se pueda establecer estrategias de marketing diferenciadas y dirigidas al nicho de mercado que genera mayores ingresos a la entidad, satisfaciendo sus necesidades particulares.

Paralelamente, se puede llevar a cabo el marketing mix, dado que se deben cubrir las variables de producto, precio, distribución y comunicación para que la segmentación que se va a realizar sea efectiva, a través del Producto, que en este caso serían las tarjetas débito y crédito de Bancolombia, que son sinónimo de facilidad de pago y cuentan con el respaldo de la compañía; en el Precio se tendrán en cuenta los descuentos con los aliados estratégicos con el fin de atraer más clientes; en la Distribución se cuenta con diferentes canales para realizar los pagos sea vía Online, Datafonos, Transferencias por Sucursal virtual o App, y pagos con código QR; y en la variable Comunicación se debe diseñar publicidad

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

llamativa que pueda enviarse por diferentes medios como mensajes de texto, correos electrónicos, mensajes por Whatsapp, etc...donde se le comuniqué al cliente los beneficios que le está ofreciendo la entidad.

La implementación de un análisis de la información de una forma periódica le permite al banco identificar las necesidades más importantes de sus clientes y asimismo las oportunidades para la organización, con el fin de generar información que pueda ser útil para la toma de decisiones a nivel de marketing y servicios.

A través del Big data, las empresas pueden explorar y hacer uso de todos sus beneficios aplicando sus bondades en el área de marketing y en muchas otras según las necesidades de la organización, permitiendo conocer a sus clientes de una manera más detallada generando mayor valor a las estrategias de marketing mencionadas anteriormente.

Una de las dificultades que se presentó durante el desarrollo de la investigación, fue la calidad de los datos debido a que no se contaba con los registros completos de cada variable, esto se debe a que en la entidad financiera no es obligatorio diligenciar cierta información personal de los clientes, por lo cual, la base de datos tenía campos vacíos y esto afecta el resultado final del estudio realizado, como consecuencia se tuvo que separar la información en 2 bases de datos, para tener un mejor manejo de la información.

Se recomienda que las entidades garanticen la calidad de la información al momento de obtener la información de sus clientes, ya que esto agiliza y permite obtener una mayor fiabilidad de los resultados. También se presentó un inconveniente con el manejo de la información suministrada por el banco, esto debido a que es información sensible de los clientes, por lo tanto no se pudo realizar el trabajo en un software especializado ben Big data, ya que esto conllevaba a enviar la base de datos a un equipo que tuviera dicha licencia, y por seguridad Bancolombia no lo permitía, por tal motivo se decidió trabajar con la herramienta Excel; para futuros estudios se sugiere tener en cuenta el tema de confidencialidad desde el inicio del proyecto para llegar a acuerdos, o por medio de contratos de confidencialidad la entidad permita hacer uso de la información en equipos de cómputo diferentes a los de la organización.

	INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm	Código	FDE 169
		Versión	02
		Fecha	30-07-2019

REFERENCIAS

- Bausela, E. (2005). Utilidad de la hoja de cálculo Excel en el análisis de datos cuantitativos. *Revista de Informática Educativa y Medios Audiovisuales*, 2(6), 2. <http://laboratorios.fi.uba.ar/lie/Revista/Articulos/020206/A1dic2005.pdf?>
- Casas, J., Guerrero, J., & Julbe, F. (2019). *BIG DATA* (UOC, Issue c).
- Escobar, H., Alcivar, M., & Puris, A. (2016). Aplicaciones de Minería de Datos en Marketing. *Revista Publicando*, 3(8), 503–512.
- Flores, G., Cadena, J., Quinatoa, E., & Villa, M. (2019). Minería de datos como herramienta estratégica. *Recimundo*, 3(1), 955–970. [https://doi.org/10.26820/recimundo/3.\(1\).enero.2019.955-970](https://doi.org/10.26820/recimundo/3.(1).enero.2019.955-970)
- Gonzalez, L. (2018). Control de nuestros datos personales en la era del big data: El caso del rastreo web de terceros. *Estudios Socio-Jurídicos*, 21(1), 209–244. <https://doi.org/10.12804/revistas.urosario.edu.co/sociojuridicos/a.6941>
- Gonzalez, L. D., Valcárcel Asencios, V., Cap, E. U. O. C., De, C., Medina Rojas, F., Gomez, C., & Gual Solé, J. (2014). Estrategias empresariales en “infocom”: el nuevo sector de telecomunicaciones e Internet. *Ingeniería y Región*, 7(2), 31. <https://doi.org/10.15381/idata.v7i2.6140>
- Ibarra-Morales, LE, Woolfolk-Gallego, LE y Meza-López, BI. (2020). Evaluación de la calidad en el servicio: una aplicación práctica en un establecimiento de Café. *Revista CEA* 6 (11), 89-107
- Medina, F., & Gomez, C. (2014). Funcionalidades de la minería de datos. *Ingeniería y Región*, 12(2), 31. <https://doi.org/10.25054/22161325.728>
- Mayor-Ríos, J. A., Pacheco-Ortiz, D. M., Patiño-Vanegas, J. C. & Ramos-y-Yovera, S. E. (2019). Análisis de la integración del Big Data en los programas de contaduría pública en universidades acreditadas en Colombia. *Revista CEA*, 5(9), 53-76. <https://doi.org/10.22430/24223182.1257>
- Londoño-Patiño, J.A. (2020). Toma de decisiones basada en la productividad en Pymes manufactureras: aproximación desde la Lógica Difusa. *Revista CEA*, 6 (12), 181-207. <https://doi.org/10.22430/24223182.1507>
- Peñaloza, M. (2005). El Mix de Marketing: Una herramienta para servir al cliente. *Actualidad Contable Faces*, 8, 71–81.
- Peralta, F. (2014). Proceso de Conceptualización del Entendimiento del Negocio para Proyectos de Explotación de Información. *Revista Latinoamericana de Ingeniería de Software*, 2(5), 273. <https://doi.org/10.18294/relais.2014.273-306>
- Perez, L. (2006). Microsoft Excel: una herramienta para la investigación. *Revista Electrónica de Las Ciencias Médicas En Cienfuegos*, 23(3), 68–70.
- T., S. (n.d.). Segmentación de clientes: ¿Disparar con rifle o Escopeta? *Sintec*, 1–2.
- Trentin, G. (1992). Estructura y organización de una base de datos. *Comunicación, Lenguaje y Educación*, 4(13), 81–88. <https://doi.org/10.1080/02147033.1992.10821002>
- Velarde, A. (2003). Minería de Datos una Introducción. *Red de Revistas Científicas de América Latina, El*

	<p style="text-align: center;">INFORME FINAL DE TRABAJO DE GRADO - FCEyAdm</p>	Código	FDE 169
		Versión	02
		Fecha	30-07- 2019

Caribe, España y Portugal, 23. <https://doi.org/1405->

Zuñiga, C., Castañeda, L., Benítez, J., & Vásquez, K. (2013). Big Data ¿ Tecnología o una prioridad empresarial para la gestión de datos digitales ? *Universidad Nacional Autónoma De Honduras*, 10.