



**Universidad**  
Zaragoza

## Trabajo Fin de Grado

# Nuevas tendencias en el uso de los datos para el marketing: Big Data.

**Autor**

José Daniel Torrubia Sánchez

**Directora**

Pilar Olave Rubio

Facultad de Economía y Empresa



Facultad de  
Economía y Empresa  
**Universidad** Zaragoza

2016

**Autor:** José Daniel Torrubia Sánchez

**Directora:** Pilar Olave Rubio

**Título del trabajo:** Nuevas tendencias en el uso de los datos para el marketing: Big Data.

**Titulación:** Grado en Marketing e Investigación de Mercados

**Modalidad:** Análisis de datos

## RESUMEN

---

Este trabajo pretende poner en valor la importancia que supone para los departamentos de marketing, analizar toda la información que obtienen a través de diversas fuentes de datos. Analizar el comportamiento del cliente para predecir qué es lo que van a necesitar mañana y ofrecérselo es el gran reto que las empresas tienen hoy en día. El objetivo del proyecto es demostrar el gran interés que para las empresas tiene el tratamiento de esta información para generar conocimiento y valor, en particular en el campo del marketing y la investigación de mercados. Este objetivo lo llevaremos a cabo mediante una introducción al concepto de Big Data, seguido de un análisis de la información como activo estratégico y finalmente con la realización de un caso práctico de data mining. Éste tendrá como objetivo, analizar los perfiles de los clientes de IKEA FAMILY y calcular la probabilidad de compra de un determinado producto a través de un modelo de regresión binario. Así pues, este proyecto consta de 4 partes: Introducción al Big Data, la información como activo estratégico, caso práctico y conclusiones, además de la correspondiente bibliografía consultada.

## ABSTRACT

---

This essay tries to point out the value of analyzing the information that marketing departments have available from several data sources. Currently, the companies main goal is to analyze and examine customer's behavior in order to predict what they will need in the future. The aim of this project is to show the large interest that companies have in the treatment of information in order to generate knowledge and value, in the marketing field and the market research particularly. This aim will be achieved through the Big Data, continued by a review of the information as a strategic asset and finally creating an example of data mining exercise. This case will analyze the customer's profiles obtained from the IKEA FAMILY card and to calculate the probability of one particular product purchase through a binary logistic regression model. Therefore, this essay consists of 4 parts: Introduction to Big Data, the information as a strategic asset, a practical exercise and a conclusion. Moreover, its related bibliography will be attached.

## INDICE

1. INTRODUCCIÓN AL BIG DATA .....	6
2. LA INFORMACIÓN COMO ACTIVO ESTRATÉGICO .....	17
2.1. Transformación digital. ....	21
2.2. La digitalización en el proceso de compra. ....	23
2.3. Caso de éxito IKEA FAMILY .....	24
3. CASO PRACTICO: DATA MINING .....	27
3.1 Descripción de las variables .....	29
3.2 Análisis exploratorio.....	31
3.3 Análisis de correspondencias .....	34
3.6 Análisis de regresión .....	39
4. CONCLUSIONES .....	45
5. BIBLIOGRAFIA.....	46

## INDICE DE GRÁFICOS

Gráfico 1. Distribución de la muestra por sexo. ....	31
Gráfico 2. Distribución de la muestra por edad.....	31
Gráfico 3. Distribución de la muestra por tipo de casa .....	32
Gráfico 4. Distribución de la muestra por tipo de estudios .....	32
Gráfico 5. Distribución de la muestra por renta media.....	32
Gráfico 6. Distribución de la muestra por visitas media. ....	33
Gráfico 7. Distribución de la muestra por gasto media. ....	33
Gráfico 8. Distribución de la muestra por consumo IKEA FOOD .....	33
Gráfico 9. Distribución de la muestra por NPS. ....	33
Gráfico 10. Perfil Fila. Según Renta .....	34
Gráfico 11. Perfil Columna. Según Visitas. ....	35
Gráfico 12. Perfil Fila. Según el número de visitas.....	36
Gráfico 13. Perfil Columna. Según la edad .....	36
Gráfico 14. Perfil Fila. Según el Gasto Medio .....	37
Gráfico 15. Perfil Columna. Según la renta.....	37
Gráfico 16. Perfil Fila. Consumo IKEA FOOD según renta.....	38

## INDICE DE TABLAS

Tabla 1. Prueba ómnibus sobre los coeficientes del modelo.....	39
Tabla 2. Prueba de Hosmer - Lemeshow.....	39
Tabla 3. Resultado del modelo .....	40
Tabla 4. Tabla de clasificación .....	41
Tabla 5. Variables en la ecuación.....	42

## 1. INTRODUCCIÓN AL BIG DATA

Los datos llevan mucho tiempo presentes en el marketing de cualquier negocio, generándose de forma automática o manual a lo largo de su actividad diaria. Desde aspectos generales como el número de clientes, ventas relacionadas de productos o conocimiento de marca a través de encuestas, hasta aspectos más concretos como el análisis de las compras de un determinado producto, en un punto de venta a una hora fija, la tasa de respuesta a una campaña de correo directo o la tasa de cupones de descuento canjeados en un código postal.

Todos estos datos procesados son información para la empresa, porque tienen un significado; relevancia, propósito y contexto. Por lo tanto, son de utilidad para quién debe tomar decisiones, al disminuir su incertidumbre. Los datos se pueden transformar en **conocimiento** añadiéndoles valor:

- Contextualizando: se sabe en qué contexto y para qué propósito se generaron.
- Categorizando: se conocen las unidades de medida que ayudan a interpretarlos.
- Calculando: los datos pueden haber sido procesados estadísticamente.
- Corrigiendo: se han eliminado errores e inconsistencias de los datos.
- Condensando: los datos se han podido resumir de forma más concisa.

El conocimiento se deriva de la información, así como la información se deriva de los datos. Para que la información se convierta en conocimiento es necesario realizar acciones como: comparación con otros elementos, predicción de consecuencias, búsqueda de conexiones o conversación con otros portadores de conocimiento.

El reto que las empresas tienen que superar para poder seguir generando conocimiento a través de sus datos es ampliar de manera exponencial sus canales de toma de datos, hoy en día el dato proviene de una conexión IP, del comportamiento de un cliente en una página web, un comentario de un consumidor potencial en una red social, la geolocalización del consumo en base al uso de tarjetas de crédito. Por primera vez los gustos y recomendaciones se pueden medir y modelizar en un algoritmo que prediga el siguiente producto que va a comprar nuestro consumidor, en base a dónde está, qué perfil tiene, qué le gusta o a quién sigue. La información se ha convertido en un bien cada vez más importante y en uno que las compañías deben emplear de forma incuestionable para sobrevivir a los cambios que imponen los tiempos.

La información como tal es un valor al alza y sin duda cada vez más puede marcar la diferencia entre el éxito o el fracaso. Esta realidad ha llevado a las empresas a obsesionarse con la información. Todo el mundo quiere saber qué debe hacer para tener más y más datos sobre sus consumidores y todo el mundo quiere acumular bases de datos inmensas y muy completas con el potencial para saber todo sobre sus compradores, sobre el mercado y sobre los movimientos de la competencia. Cuantos más datos se posean, mejor preparada estará la compañía. Y por ello las marcas han entrado en una suerte de guerra sin cuartel por el dominio de la información. Todo el mundo quiere ser quien se posicione de la manera más notable en esta carrera.

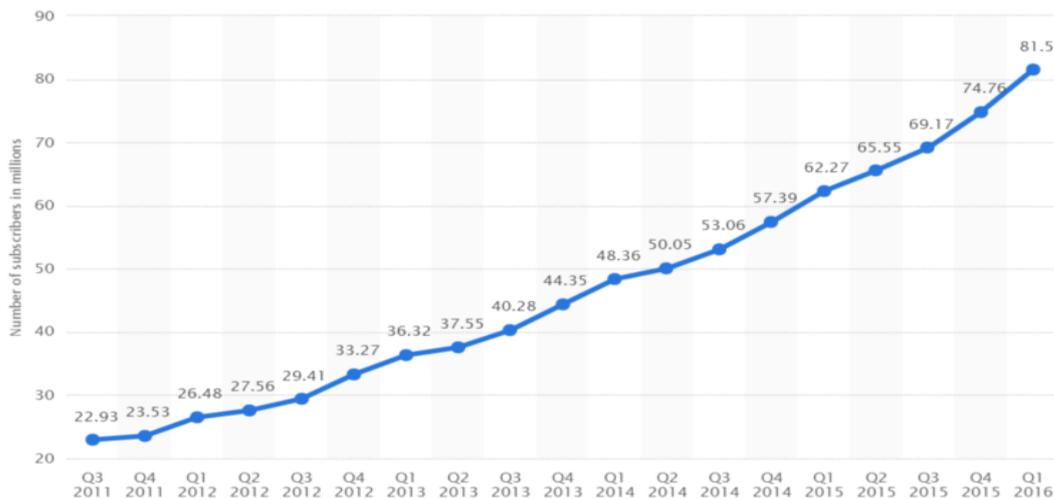
Pero la obsesión por la información está haciendo que los responsables de las empresas pierdan de vista una cosa. No todo es información y no todo son datos. También hay que tener en cuenta otras cuestiones. *Los datos por sí solos no sirven para nada.* Es necesario saber hacerles las preguntas correctas para que éstos se conviertan en valor para las empresas. En los algoritmos es donde se encuentra el valor real que define la acción y forma el corazón de las nuevas interacciones con los consumidores. **Este proyecto pretende poner en valor todas estas ideas y destacar la importancia de la digitalización de las empresas.**

A continuación, se detallan dos ejemplos en los que el éxito está muy relacionado no sólo con la información que poseen sino con lo que hacen con ella. En primer lugar, tenemos como ejemplo la empresa Amazon la cual ha patentado un nuevo método de trabajo llamado «*Method and System for Anticipatory Package Shipping*», con el que el gigante de las compras online pretende minimizar el tiempo transcurrido entre el momento en que se realiza una compra y llega a casa. Con el método de envíos anticipados de Amazon será posible enviar un producto antes de que el cliente lo compre, permitiendo reducir el tiempo de entrega a tan sólo unas horas.

Este método de envíos anticipados basado en modelos predictivos de comportamiento del consumidor, constituirá todo un desafío estratégico que será aplicable principalmente a productos de consumo masivo. La clave del sistema está en la información almacenada en las bases de datos de Amazon y en el uso de la tecnología de Big Data, para establecer modelos teóricos de comportamientos futuros basándose en el historial de compra de cada cliente y sus hábitos de consumo. Estos modelos teóricos serán utilizados por el aparato logístico de la compañía para la predicción de preferencias y conocer de antemano qué

cantidad de cada producto se servirá en un futuro próximo a un determinado código postal o barrio de una ciudad.

En segundo lugar, tenemos a la compañía norteamericana Netflix que actualmente tiene más de 80 millones de clientes en todo el mundo.



Fuente: Statista 2016. Global Insight Netflix

Gracias al análisis de enormes cantidades de datos, Netflix, toma decisiones que contentan a sus suscriptores. Por ejemplo, Netflix analiza cuántos espectadores han visto una serie completa, qué día y a qué hora ven un episodio, desde qué dispositivos, cuándo paran o aceleran la reproducción, y hasta qué hacen cuando llegan los títulos de crédito, para ver si el espectador quiere ver otro episodio cuando termina o cierra la aplicación. Por otro lado, Netflix pide a sus clientes que valoren su interés en diferentes géneros o películas que han visto. El algoritmo de recomendación les sugiere después títulos adaptados a sus gustos. Y parece que acierta bastante, ya que el 75% de lo que ven los usuarios procede de las recomendaciones.

Gracias al análisis de todos los datos que la empresa acumula sobre sus usuarios ha conseguido crear productos específicos para su audiencia en base a sus gustos. Es el ejemplo de la serie House of Cards, la primera que produce la plataforma con un coste de 100 millones de dólares y que ha alcanzado el éxito internacional, con un nivel de calidad que le ha valido un Emmy y dos Globos de Oro. La serie se ha creado en base a los gustos e intereses de los espectadores, información obtenida con un proyecto de Big Data, que permite saber qué quieren ver, cómo, cuándo y así cualquier aspecto que permite crear la serie perfecta para el público idóneo. Antes de dar luz verde al proyecto, Netflix sabía

que la temática (ficción política), el director, los actores y la versión inglesa de House of Cards tendrían buena acogida entre sus suscriptores. No sólo eso: prepararon 10 tráiler diferentes para promocionar la serie, adaptados a los diferentes tipos de espectador.

Como hemos visto, la utilización de **análisis de Big Data en los departamentos de marketing tiene mucha importancia** ya que los beneficios pueden resultar muy interesantes. El motivo principal es que al conocer a tus clientes adquieres una ventaja frente a tu competencia muy importante. En este caso, al saber los gustos, las necesidades que tienen o su comportamiento de cara a tu marca o a un determinado estímulo, se pueden generar estrategias y acciones específicas con un resultado muy positivo.

La utilización de una forma eficaz del Big Data es complicada, ya que los sistemas analíticos no siguen los procesos y decisiones que le gustaría al departamento de marketing. Por ello es muy importante saber qué datos hay que recopilar, qué soluciones analíticas se deben usar y conocer cómo convertir los datos en conocimiento que genere un impacto positivo en tus estrategias. Se debe utilizar el Big Data para obtener información relevante y útil para tu empresa. Pero a continuación se debe confiar en soluciones analíticas que permitan tratar y analizar esos datos con óptica de negocio, conociendo de antemano qué información te van a facilitar y cómo aplicar esa información en la actividad comercial de la empresa. De esta manera se pueden desarrollar acciones y estrategias concretas que te impulsen a alcanzar los objetivos previamente marcados.

Actualmente, por abuso del lenguaje se **denomina Big Data** a la gestión y análisis de enormes volúmenes de datos que no pueden ser tratados de manera convencional ya que superan los límites y capacidades de las herramientas de software habitualmente utilizadas para la captura, gestión y procesamiento de datos, aunque hace unos años el concepto de Big Data hacía referencia exclusivamente al volumen de datos y su tratamiento era un tratamiento estadístico estándar. Así pues, las herramientas clásicas de análisis de datos conocidas como minería de datos, son las idóneas si el procesamiento de la información al dato depurado ha sido la adecuada. Este concepto engloba infraestructuras, tecnologías y servicios que han sido creados para dar solución al procesamiento de enormes conjuntos de datos estructurados, no estructurados (mensajes en redes sociales, señales de móvil, archivos de audio, sensores, imágenes digitales, emails, logs, entre otros).

Así pues, el trabajo lo vamos a desarrollar en dos grandes bloques, el primero de ellos es una introducción al Big Data y la información como valor estratégico y el segundo bloque consistirá en la utilización de técnicas de análisis multivariante para el procesamiento de la información en un departamento de marketing.

Las tecnologías Big Data no sólo ayudan a recopilar grandes cantidades de datos, sino que además permiten su almacenamiento, organización y recuperación para aprovechar todo su valor. Y con el objetivo puesto en que su uso permita optimizar la toma de decisiones. El Big Data es al mismo tiempo un reto y una oportunidad para las empresas porque les permite mejorar su competitividad y adaptarse al nuevo escenario de la economía global y digital, en el que nuevos agentes están revolucionando las diferentes industrias con nuevos modelos de negocio y propuestas de valor Según el informe “*Open Data in Europe*”, realizado por la Fundación DemosEuropa, **el Big Data generará 4,4 millones de empleos** en todo el mundo en los próximos cinco años. El comercio, la industria, la salud, la información, las comunicaciones, la banca, los seguros son los sectores donde el aumento de la inversión será más relevante.

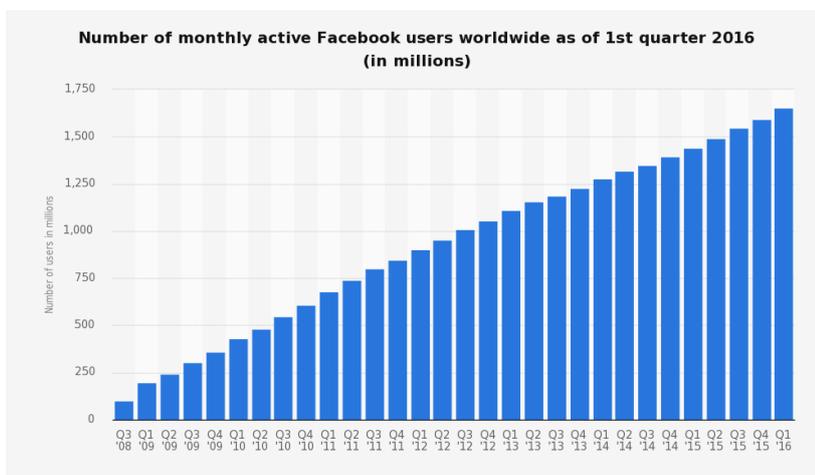
El volumen de los datos generados –junto a la velocidad, la variedad de su naturaleza y la veracidad– es una de las señas de identidad del fenómeno Big Data, señas que también identifican los principales retos y problemas que ha de afrontar en su desarrollo. Desde la llegada de Internet, los científicos de computación han trabajado para aumentar considerablemente el poder de procesamiento de las máquinas. Durante las tres últimas décadas, cada diez años la velocidad de procesamiento de los ordenadores se ha multiplicado por mil. En 30 años, se ha multiplicado por mil millones. Esto ha supuesto que el supercomputador más rápido del mundo hace 12 años quepa hoy día en un chip.



También hay que destacar cómo la capacidad de almacenamiento se ha reducido drásticamente, una tarjeta de memoria MicroSD del año 2010 almacenaba solo un 0,1% de lo que almacena una del mismo tamaño en el año 2015. (1GB = 1000MB)

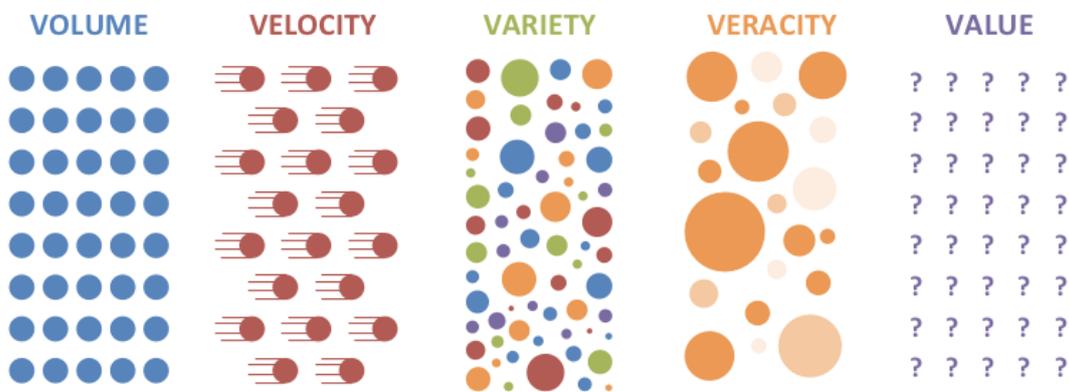
Junto a otros factores, a lo que se enfrentan hoy las plataformas de Big Data es a las consecuencias de ese aumento exponencial de los órdenes de magnitud. Un ejemplo del mundo científico: el acelerador de partículas LHC (Large Hadron Collider) genera 1 PetaByte (1 millón de GigaByte) de datos por segundo. Esta gigantesca cantidad de datos producidos en un solo segundo es similar al volumen de información que ocupan 10.000 millones de fotografías o 13 años de televisión de alta definición. Además, el Colisionador de Hadrones, que forma parte del Centro de Datos del CERN (Organización Europea para la Investigación Nuclear), comparte la información con 170 centros colaboradores de 36 países en todo el mundo que están conectados con el CERN. Esta red de centros pone en funcionamiento centenares de miles de ordenadores que proporcionan los recursos necesarios para almacenar, distribuir y procesar toda la información generada. El poder combinado de esta red en un sólo día es el equivalente al trabajo continuado de un ordenador durante más de 600 años.

Si analizamos todo lo que ocurre en internet las proporciones de escalabilidad se disparan, por ejemplo, sólo en la plataforma de vídeos Youtube cada minuto se ven 138.889 horas de vídeo (aproximadamente 16 años), si analizamos la red social de fotografías Instagram cuenta mensualmente con 400 millones de usuarios activos que suben diariamente 80 millones de fotos. En el caso de la red social Facebook en el primer cuarto del año 2016 alcanzó una media de más de 1.5 billones de usuarios diarios que compartieron 4,2 millones de publicaciones.



Fuente: Facebook @Statista 2016

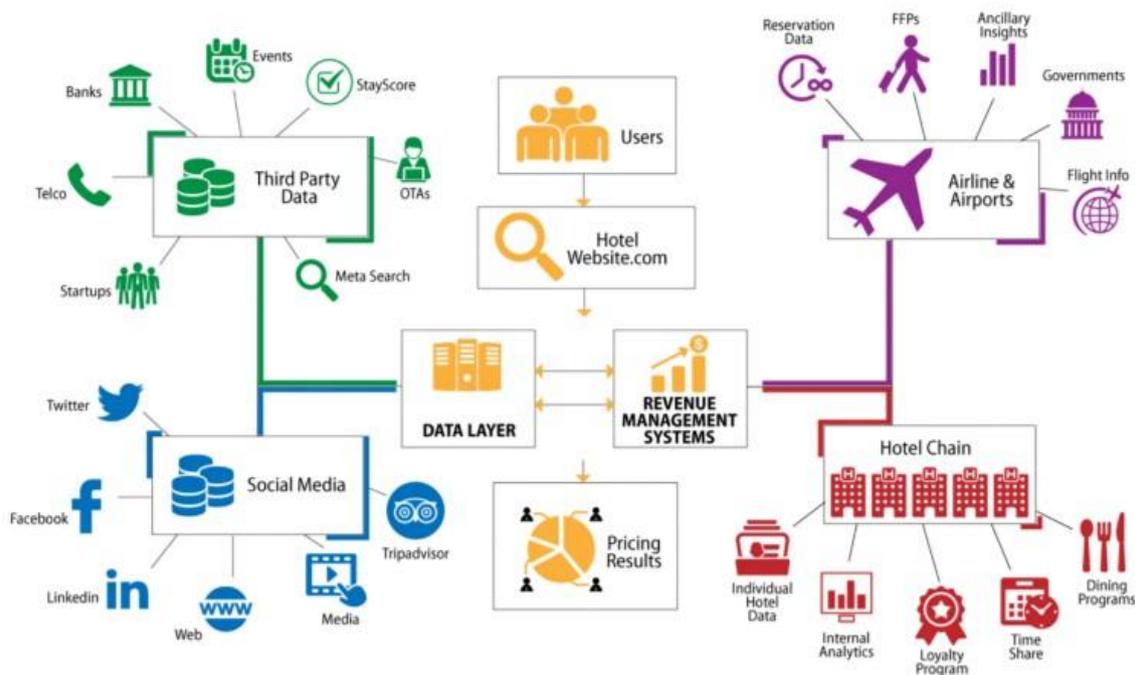
Puesto en contexto, según Eric Schmidt (ex CEO de Google) “Entre el nacimiento del mundo y el año 2005 se crearon 5 Exabytes (1.000 Millones de Gigas) de información. Actualmente se crean 5 Exabytes cada dos días. También se puede afirmar que el 90% de toda la información creada por el ser humano se ha creado en los dos últimos años. Se calcula que para el año 2020 se acumulen 50 veces más datos que los que había en 2011. Este aumento se debe en gran medida al “Internet de las cosas”; millones de sensores por todo el mundo creando datos cada segundo. Por ejemplo, para almacenar los datos que genera un avión de última generación durante un año, necesitaríamos un millón de ordenadores de sobremesa domésticos (1M de Terabytes).



Por todo lo comentado, el **volumen** parece ser la cualidad principal del Big Data pero hay otras, a continuación, dos características principales; la variedad y la velocidad. Hasta el 2005 todo el dato que se generaba era estructurado, todo se podía almacenar en columnas y filas. Esta información se almacenaba en bases de datos que se unificaba en datawarehouse y provenía principalmente de los sistemas informacionales de las empresas como, por ejemplo, ERP, CRM o Sistemas contables. Pero la realidad es que hoy ese tipo de información sólo corresponde al 10% de la información que pueden obtener las empresas, el otro 90% se encuentra en formato no estructurado, esto se debe a que los datos tienen múltiples formatos porque se originan en diferentes servicios y canales como máquinas de procesamiento. Parece lógico que analizando todos los mails intercambiados a lo largo de una semana en un departamento de atención al cliente de una empresa multinacional podamos analizar, sentimientos de consumidores, rapidez de la respuesta, principales problemas ocasionados...pero ¿cómo se estructura un correo electrónico para analizarlo? O, por ejemplo, qué relación tiene esos correos del departamento de atención al cliente, con lo que nuestros consumidores dicen en las redes

sociales, de qué manera debe actuar la empresa para mantener o mejorar el compromiso con sus consumidores.

Como hemos visto la **variedad** de los datos conforman una realidad a tener en cuenta en un proyecto de Big Data y esta variedad no sólo afecta al almacenamiento en forma no estructurada o en formato clave-valor, sino que para cada tipo de dato tenemos que amoldar el tipo de análisis y tratamiento que vamos a realizar



Fuente: 2016 markrs.co

La tercera cualidad principal del Big Data es la **velocidad**, las cantidades de datos que se generan por minuto son inmensas y la variedad de los datos obliga a las empresas a actuar con mucha rapidez porque seguramente si corriges un mes después lo que has analizado que la gente percibe de tu producto a través de todos los datos recogidos de las redes sociales, será demasiado tarde. El Big Data necesita tiempo real y procesamientos iterativos. Cada vez que un usuario se conecta con su ordenador o con su dispositivo móvil a internet, está generando información constantemente se envían señales a los servidores donde están las páginas web alojadas. Volviendo al ejemplo anterior, Cada búsqueda en el portal google.com define el perfil del usuario asociado a una dirección IP y automáticamente analiza millones de páginas web para dar la mejor respuesta a la búsqueda. Para una empresa como Amazon, es vital saber todo lo que pueda de quien está conectado a su web en todo el mundo para alimentar su algoritmo de predicción y poder

recomendar productos a sus clientes en relación a sus búsquedas e incluso antes de que sea consciente de que lo va a necesitar. Amazon facturó \$12M cada hora del 2015 (total *facturación 2015; \$100.000M*)

*Después del Volumen*, la Variedad y la Velocidad, la comunidad científica entorno al Big Data se ha encargado de seguir analizando más en profundidad de qué manera se define el Big Data. *Valor, Veracidad, Variabilidad y Visualización.*

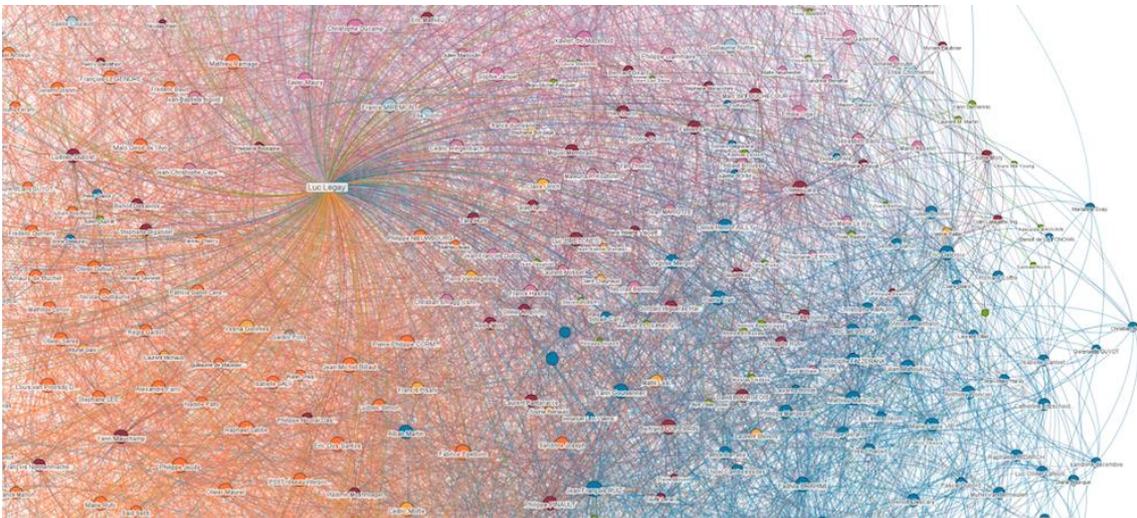
Lo primero de todo, como se ha comentado al principio de la introducción de los datos se obtiene información y de la información se obtiene **valor**. Es muy fácil caer en la acumulación de datos por parte de las empresas sin un claro objetivo detrás, simplemente como están ahí y se pueden almacenar las empresas gastan parte de su presupuesto en hacerlo, aunque luego no obtengan nada de ellos. El valor del Big Data reside en ayudar a las compañías y organismos a entender mejor lo que sus consumidores quieren, permite optimizar procesos, anticipar demanda; amoldar de manera dinámica los precios para maximizar el beneficio en función de las posibilidades de compra. Si la empresa no obtiene valor de un proyecto de Big Data es porque no se está llevando a cabo bien.

El siguiente punto es la **variabilidad**. Todos los análisis de datos son variables porque, en función de los datos obtenidos o de las variables estudiadas, el resultado puede variar, por eso los algoritmos que se usan en Big Data son interactivos en base a “Machine Learning”. Todos los nuevos casos son estudiados para mejorar el modelo e ir variándolo en función de cómo se comporten los inputs. Si un proyecto de Big Data no incorpora la variabilidad a sus procesos, se quedará obsoleto en un mes.

Tener mucho volumen de datos de muchas fuentes de información supone un problema para la empresa, no sólo por su almacenamiento o análisis. Saber en todo momento que los datos que corren por nuestros sistemas de Big Data reflejan la realidad de los procesos de una empresa es fundamental. Para garantizar la **veracidad** en todo momento de la información es necesario un análisis exhaustivo que se ha denominado “gobierno del dato” (Data Governance). La información es extraída, compartida, analizada, muchas veces retroalimentada y está distribuida a lo largo de los sistemas de procesamiento, de los que hablaremos más adelante, por todo ello es necesario que la información que analizamos en el punto final de nuestro proceso de datos es la misma que ha entrado. En los sistemas informacionales de datos estructurados controlar esto a través de metadatos

era relativamente sencillo, hoy en día con los nuevos sistemas de particionado de archivos o arquitecturas de sistemas distribuidos es mucho más complejo.

Como última cualidad para definir el Big Data encontramos la **visualización**. Todos los esfuerzos en este sentido están enfocados hacia la usabilidad de la información. Hacer accesible la información en tiempo real para los usuarios finales es el objetivo principal de los procesos de visualización. A través de complejos sistemas de información, el usuario final antes de tomar una decisión puede ver y consultar de manera sencilla los inputs que necesita. Mapas de calor en función de las ventas, geolocalización de consumidores, análisis de sentimientos o métricas de impacto de campañas en forma de Key Performance Indicator (KPI).



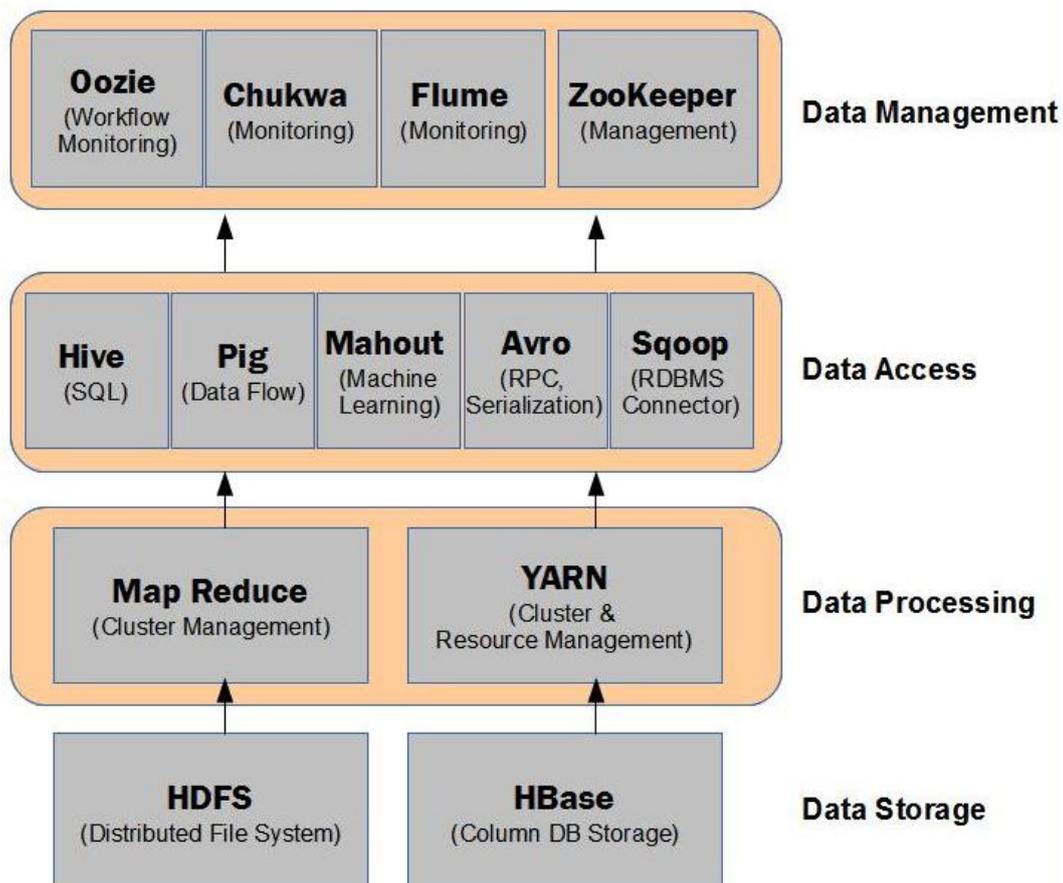
La tecnología que ha permitido esta revolución se llama Hadoop. Hasta la aparición de esta tecnología, las empresas almacenaban su información en servidores locales que tenían en sus oficinas con capacidad limitada a la cantidad de memoria del servidor. En diciembre del año 2004 Google publica un informe científico titulado MapReduce: Simplified Data Processing on Large Clusters donde explican los avances que habían descubierto en torno a la computación horizontal distribuida en clusters. Esta publicación es el origen del Big Data como hoy en día lo conocemos, hablar de Big Data es hablar de Hadoop. Aunque esta tecnología existe y se distribuye como una plataforma de servicios, son muy pocas (<10%) las empresas que incorporan esta tecnología.

Hadoop se creó en el año 2006 a raíz de la necesidad de sistemas nuevos para gestionar la explosión de datos de la web. De descarga gratuita, y software libre para potenciarlo y mejorarlo, Hadoop es un método de código abierto para almacenar y procesar los datos que permite el procesamiento en paralelo distribuido de enormes cantidades de datos en

servidores estándar del sector, económicos, que almacenan y procesan los datos, y que pueden escalarse sin límite.

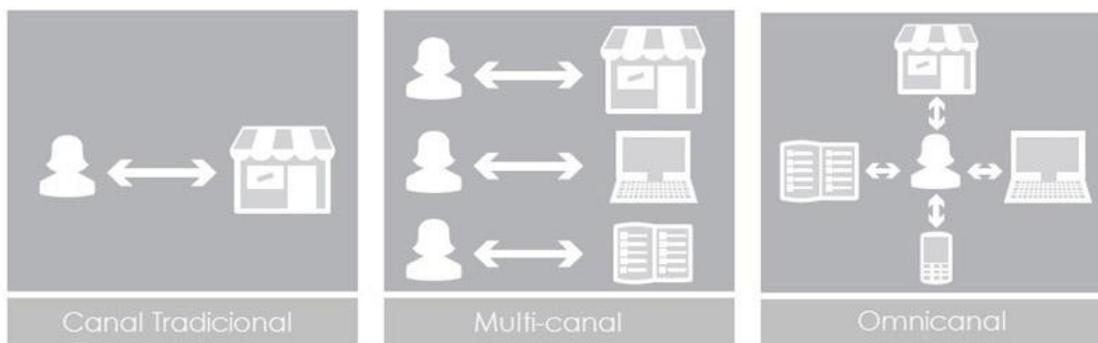
Hadoop se compone de dos partes principales:

- Hadoop Distributed File System (HDFS): Se trata de un sistema de ficheros responsable de almacenar los datos en el clúster dividiéndolos en bloques y distribuyéndolos a través de los nodos del clúster.
- MapReduce: Es un modelo de programación encargado de distribuir una tarea por diversos nodos del clúster y procesar los datos que se encuentran en esos nodos. Consta de dos fases: Map y Reduce.



## 2. LA INFORMACIÓN COMO ACTIVO ESTRATÉGICO

La disrupción digital ha obligado a las empresas a pasar de una estrategia de producto monocal a una omnicanal. Ofrecer a los consumidores una experiencia fluida sea cual sea el punto de contacto entre la marca y el cliente, unificando el proceso de compra. Este es el gran reto de una estrategia omnicanal.



Los consumidores están cada vez más conectados. Las barreras entre el comercio físico y el digital tienden a difuminarse progresivamente, una realidad que las empresas deben integrar en sus estrategias. El concepto de omnicanal implica que el cliente puede ir a la tienda a recoger el pedido que ha realizado por Internet. Y del mismo modo, se trata de hacer accesible online el inventario de productos que están disponibles en una tienda en tiempo real y viceversa. El 70% de los visitantes de una tienda navegan en su plataforma digital antes de comprar. De hecho, no menos del 43% de los visitantes de una página web de una marca compran en esta misma marca en los próximos tres meses. También hay que destacar, la tasa de conversión de un visitante multicanal es un 20% más alto que el de un cliente monocal.

Lo primero de todo es colocar el conocimiento sobre el cliente en el centro de la estrategia de marketing de una compañía. Hay que aprovechar los datos recogidos de los consumidores a través de su experiencia en la página web y también en la tienda física. El cliente debe ser conocido y reconocido, y ser recibido adecuadamente. Paralelamente, acompañando esta transformación, la tienda física debe evolucionar. El recorrido de compra en tienda debe facilitarse siguiendo una lógica omnicanal y los vendedores deben estar informados de lo que están comercializando en tiempo real, con el objetivo de responder a las expectativas de un consumidor sobre informado.

Esta nueva experiencia de cliente obliga a las empresas a reformular los procesos y canales de venta, la forma de comunicarse con sus clientes y de atraerles. Las marcas de hoy en día tienen que centrarse en fidelizar a sus clientes a través del valor real que aportan. **Esta estrategia basada en poner en el centro al cliente y priorizar el valor que aporta la compañía, tiene que incluirse en una cultura de innovación abierta, derribar los departamentos funcionales que están establecidos en las empresas, en definitiva, colaborar y trabajar de una manera diferente.**

Si el nuevo consumidor es omnicanal y digital, las empresas tienen que introducir esos valores y principios de la cultura digital en sus procesos internos y externos. De esa manera pueden surgir nuevos modelos de negocio capaces de responder de una manera más eficiente a las necesidades del consumidor.

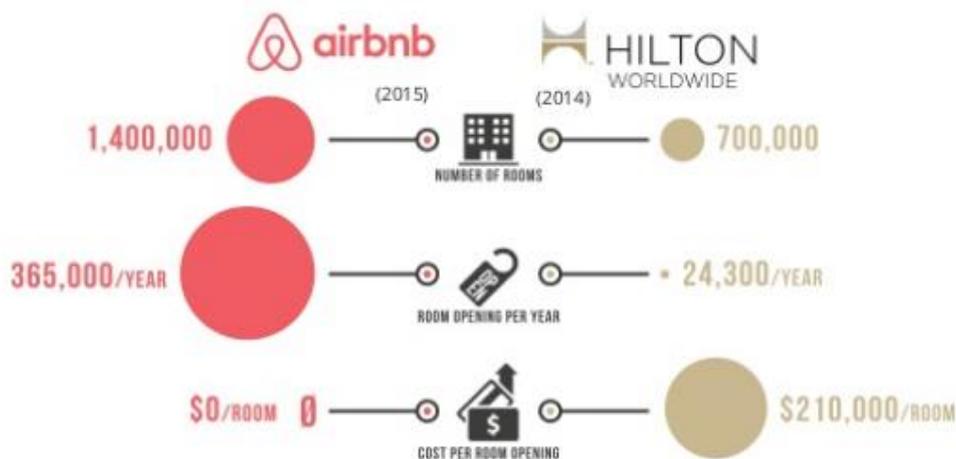
Las principales empresas centradas en la estrategia digital y con el objetivo claro de diferenciarse aportando valor real a sus usuarios son empresas que ya han nacido con este nuevo paradigma empresarial. A gran escala, se podría decir que los gigantes Google, Apple, Facebook y Amazon son los que, en estos últimos años, han dado un vuelco a la economía tradicional obligando a todas las empresas a cambiar o amoldar sus estrategias en las direcciones que marcan estos. Sin duda estas cuatro empresas han establecido las nuevas reglas de la estrategia empresarial. Apple y Google son hoy las marcas más valoradas del mundo, por encima de Coca-Cola, que supera por poco Amazon y Facebook. Google es, además, la mayor intermediaria de medios del mundo, con ingresos por publicidad que superan los de todas las multinacionales de publicidad y medios juntas.

- Estas empresas han redefinido el concepto de cliente y han potenciado la información como activo estratégico. En primer lugar, no hacen ninguna diferencia entre un cliente que paga y uno que no paga porque están dispuestos a hacerse indispensables a tantas personas como sea posible.
- Se enfocan en ahorrar tiempo y esfuerzos en la vida diaria de los clientes y en ofrecer a todo el mundo la mejor experiencia posible, para ganar su atención. Su cadena de valor incluye y trata de la misma forma a los visitantes (alguien que demuestra interés), a los amigos (aquellas personas que proporcionan datos o cualquier tipo de información), y a los clientes que realmente hacen una compra.

- Por otro lado, han redefinido la creación de valor. Primero se piensa en términos de compromiso con el cliente en lugar de valores financieros. La entrega de valor al cliente sostenible prevalece sobre la rentabilidad a corto plazo. Internamente estas cuatro empresas han redefinido la gestión del talento. Han creado un ambiente y entornos de trabajo propicios a la innovación para impulsar el potencial de rendimiento y adelantarse al futuro.

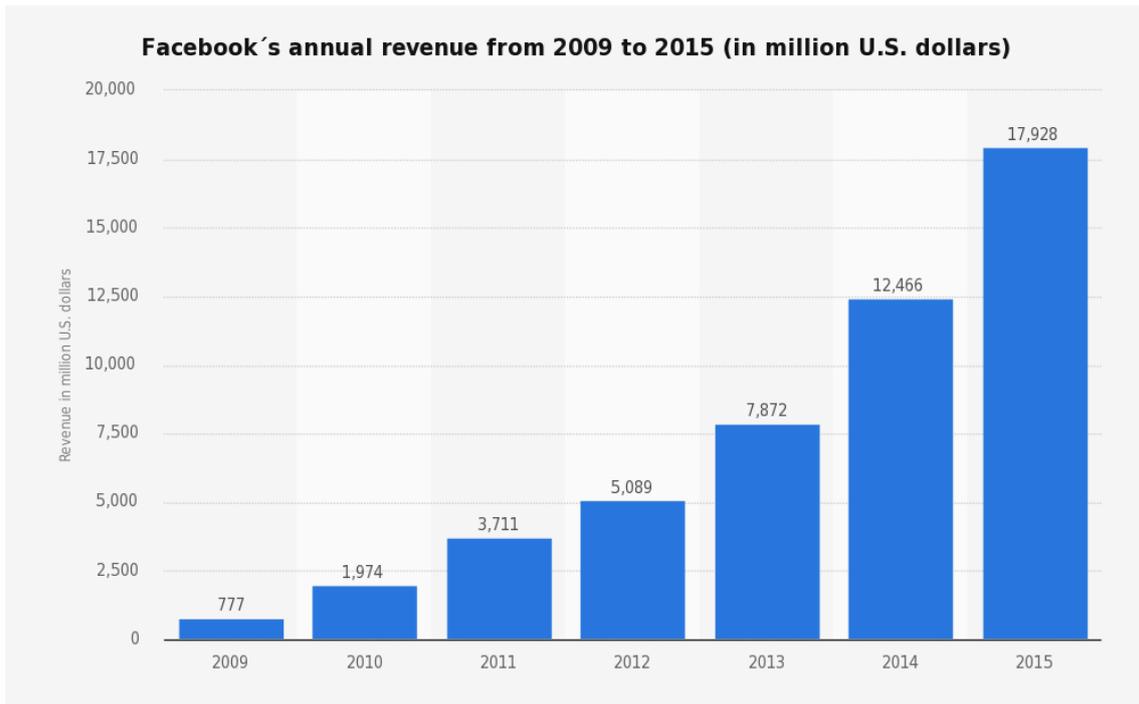
Estos nuevos planteamientos, junto a las funcionales que ofrecen estas empresas, ha permitido a otros gigantes como Uber, Netflix, Tesla o AirBnb construir sus modelos en la misma línea, con el plus de que se apoyan en sus plataformas para operar. Todas ellas se basan en obtener el máximo partido de las economías en red y en la información como su activo estratégico máspreciado.

El principal cambio de estas empresas es entender el poder de la economía en red. Las empresas deben estar integradas por unidades pequeñas que, unidas, suman un gran potencial de generación de valor. Como el iphone, un producto que se convierte en único para el usuario al poder alojar innumerables apps que le sean útiles, y que son producto de la cocreación de millones de desarrolladores. O el modelo de AirBnb, que identifica “stock” disponible y lo pone en el mercado a escala global, y con coste cero, como muestra el gráfico inferior.



Fuente: Quartz, Wall Street Journal, FABERNOVEL (estimates)

Por otro lado, el objetivo de rentabilidad se basa en la escalabilidad del modelo, de modo que los desarrollos de software y servicios alcancen el coste cero por cliente (100% de rentabilidad por cada uno). Un ejemplo sería el de Waze, la app que integra en tiempo real datos de tráfico procedentes de usuarios de todo el mundo en un modelo de cocreación, a coste zero y utilizando las redes sociales para viralizar la información. O el de Facebook, que con un incremento constante de usuarios desde 2002, alcanzando en el año 2015 casi 18.000M\$ en ingresos de publicidad.



Fuente: Facebook @Statista 2016

Junto a la escalabilidad, estas empresas tratan de obtener y aplicar aprendizajes en tiempo real, como hace Amazon al recoger y analizar a gran velocidad fuentes masivas de datos que le permiten optimizar la oferta de producto, precio y unidades a lo largo de todo su ciclo de vida. Por ultimo en lo más profundo de la estrategia de estas empresas está el optimizar la experiencia de cliente a gran escala y a la vez de forma individualizada. Netflix o Spotify serían los mejores ejemplos de este tipo de enfoque, en el que se utilizan algoritmos para ofrecer recomendaciones de consumo de contenidos en base a la acumulación de insights obtenidos sobre intereses declarados, recomendaciones de otros usuarios y hábitos consolidados. Este es el reto de este trabajo, la optimización de la información y su tratamiento para que los datos estén accesibles para definir los perfiles de los clientes.

## 2.1. Transformación digital.

La principal razón para que las empresas actuales incorporen en su cultura estrategias digitales tiene que ver con su continuidad en el mercado a largo plazo, todas las que no lo hagan sencillamente quedarán excluidas. Actualmente las empresas digitalizadas venden más y son más rentables. A continuación, se analizan las tres partes del retorno de la inversión de la transformación digital de las empresas donde el uso de los datos y su analítica adquiere un peso importante.

La transformación digital del departamento de marketing y ventas, nos lleva a conocer mucho más a nuestro cliente. A través de esta información las ofertas que envían las empresas pueden estar más personalizadas que nunca. Las empresas que siguen estrategias sociales calculan cuál es su valor de adquisición, cual es el valor que un cliente va a aportar para la empresa y centrarse en retener a los que más valor potencial tengan. Esto se conoce como fidelización 2.0 las empresas gastan sus recursos en fidelizar y retener a los clientes que realmente les interesa.

Los nuevos canales digitales de venta, como las tiendas online, aplicaciones comerciales o el nuevo “retail digital” interacción directa con los productos a través de soporte digital en los puntos de venta físicos. Generan nuevos escenarios de venta donde el consumidor una vez más nos sigue dando mucha información sobre qué es lo que realmente le interesa de un producto, que es lo que necesita y cuál es el recorrido que ha seguido desde la información y el conocimiento hasta la compra final.



Toda esta información y conocimiento de los usuarios amplían los productos y servicios digitales que las empresas y las marcas pueden ofrecer. Generando más oportunidades de venta e interacción con los consumidores o incluso nuevos modelos de negocio diferentes a la estrategia principal de la compañía.

En primer lugar, la digitalización permite a las empresas tener acceso a conocimiento de expertos a nivel externo y generar nuevos espacios para la gestión del conocimiento interno de la empresa por lo que es posible reducir costes de innovación y desarrollo de producto.

Dentro de la cadena de valor de las empresas, la cadena de logística y distribución también se ha visto drásticamente afectada por el uso de los datos y la digitalización, permitiendo la automatización de los procesos, la mejora continua de las previsiones y las gestiones de inventarios, reduciendo directamente los costes.

Por último dentro del departamento de marketing la reducción de costes que originan las estrategias digitales, se basa en la capacidad analítica y de medición que permiten optimizar al máximo la compra de medios o los procesos de captación de nuevos clientes.

Los activos digitales que las nuevas empresas están desarrollando tienen que ver con las comunidades, los canales sociales y los blogs de contenidos. Estos nuevos activos que van más allá de los activos intangibles de las marcas, se basan en la vinculación de los usuarios con la marca, pero de manera indirecta y prácticamente sin intervención de la propia marca porque son los propios usuarios los que consumen y crean el contenido que se genera. Las marcas con alta reputación corporativa digital, generan mucha más notoriedad que las que no siguen este tipo de estrategias digitales. Por último, una cultura digital realmente innovadora, asegura la atracción y la retención de talento y aumenta el activo más importante de las empresas que es el conocimiento y capacidades de sus trabajadores.



## 2.2. La digitalización en el proceso de compra.

Estos datos pretenden poner en contexto el cambio que supone la economía digital para las empresas a través de datos de los clientes.

- De media un utilizamos internet 150 veces al día.
- Más de la mitad de los españoles entre 16 y 74 lo hacen diariamente.
- 18,6 millones de personas lo hacen principalmente a través del móvil.
- EL 50% de los menores de 34 años echa de menos poder comprar sus marcas favoritas a través de Facebook.
- Un 50% se muestra proclive a compartir sus productos favoritos es Facebook
- El 84 %de las personas que compran en una tienda física, utilizan su móvil para obtener información sobre el producto. Buscando el producto y sus características en google o preguntando alguien que le parece por Whatsapp.

La unión de la experiencia de compra digital y física es clave para atraer a este tipo de consumidor digital, acostumbrado a visitar la tienda y posteriormente realizar una búsqueda online de la mejor oferta para ese producto. O por el contrario primero el cliente realiza el proceso de búsqueda de información online y luego procede a la compra del producto en una tienda física. Por ello las estrategias de marketing y posicionamiento tienen que estar integradas por ejemplo digitalizando el punto de venta físico para crear una experiencia de compra única. Motivar las interacciones con la marca desde el dispositivo móvil en el interior del punto de venta físico, poder realizar pedidos online desde la propia tienda o fomentar compartir en el acto experiencias del punto de venta físico en redes sociales.



### 2.3. Caso de éxito IKEA FAMILY

Son muchas las empresas que están empezando a moldarse a las nuevas tecnologías y optimizando sus procesos a través del análisis de los datos que disponen. En el sector retail y de gran consumo esta tendencia se incrementa. Para reflejar este enfoque al cliente 360° hemos elegido trabajar sobre el caso de éxito en la gestión interna de la información de clientes. El problema que planteamos es como IKEA a través de su tarjeta de IKEA FAMILY puede predecir qué clientes tienen mayor probabilidad de comprar un tipo de colchón específico.

IKEA es una compañía multinacional de origen sueco. Líder en el mercado de la distribución de mobiliario y objetos de decoración para el hogar. Ofrece un amplio abanico de productos funcionales, de calidad y diseño, a unos precios asequibles para la mayoría de los consumidores. IKEA tiene pasión por sus clientes, y para todos aquellos que además de clientes son incondicionales de la marca, decidió crear un club especial denominado. IKEA FAMILY En el club se consiguen, aparte de otros beneficios exclusivos, descuentos especiales en productos del surtido normal de IKEA, y también en el surtido único y exclusivo de productos IKEA FAMILY.



El objetivo de IKEA España es dotar de valor al club IKEA FAMILY. La compañía ha identificado que sabiendo cómo son sus clientes y cómo se comportan en el contexto de cada tienda, puede amoldar sus estrategias de comunicación y optimizar la gestión de stocks de almacén en función de los productos que más demanda futura espere tener. Para cumplir este objetivo el primer paso es dotar de los recursos técnicos necesarios, empezando por construir una base de datos operacional e informacional donde poder integrar y consolidar la información que se origina de las distintas fuentes como las tiendas (caja), tiendas (monolitos), web, correo postal y teléfono.

Los clientes pueden darse de alta en IKEA FAMILY desde cualquiera de estos 5 soportes, aunque el más importante es el monolito expendedor instalados en todas las tiendas. Toda la información de los clientes tiene que ser almacenada para luego vincularla mediante el ID de cada usuario a sus compras. De esta manera gracias a los históricos de las compras, el departamento de analítica sabe que productos ha comprado ya y en función de esos estimar cuales son los productos que tiene más probabilidad de comprar en su próxima visita a la tienda. La forma de conseguir que



los clientes se identifiquen con la tarjeta de IKEA FAMILY a la hora de pasar por caja es ofreciendo descuentos exclusivos en determinados productos, un descuento directo el día de tu cumpleaños, participar en seminarios talleres y eventos de la marca y lo más valorado café gratis todos los días.

El café gratis para todos los socios es uno de los productos que sirven para realizar estimaciones sobre cuanta gente ha pasado por la tienda. Un indicador muy útil para analizar si una persona en concreto ha pasado por la tienda y no ha comprado nada. En ese momento se activa una alerta en un sistema programado que tiene que enviarle un correo electrónico que le recuerde que tiene un descuento en unos artículos seleccionados.

En su afán por conocer a sus clientes, IKEA analiza las relaciones de las compras de los clientes para poder segmentar sus comunicaciones y predecir de antemano la mayor tasa de éxito para el menor coste operativo.

Como ya hemos introducido nuestro estudio en este trabajo se centra en optimizar el envío de publicidad en concreto para la activación de una promoción sobre un tipo de colchón específico. Para ello vamos a analizar qué perfil de cliente tiene una alta probabilidad de compra de este tipo de colchón **utilizando un modelo logit binario**.

- IKEA FAMILY en cifras de 2015 según los datos de la consultora Pentasoft.



En el año 2015 IKEA FAMILY sobre paso los 5 millones de personas. En tan solo dos años ha aumentado casi en 1 millón de personas. En el año 2013 eran 4,5 millones y en el año 2015 suman un total de 5,3. IKEA FAMILY realiza más de 9 millones de envíos de correos electrónicos en forma de newsletters totalmente segmentados y personalizados. Según los datos oficiales presentados en el informe anual de 2015, ser socio de IKEA FAMILY incrementa un 45% la frecuencia de visita y un 13% la probabilidad de compra en cada visita, gracias a los productos reclamo que están estratégicamente colocados para que los compres. Respecto a la página web los usuarios de IKEA FAMILY entran un 28% más que no lo son.

<b>Datos Oficiales IKEA España, 2015</b>	2013	2014	2015
Facturación (mill. €)	1.162	1.165	1.281
Facturación IKEA Food (mill. €)	52	52,8	57,3
Inversión nuevas instalaciones y mejoras (mill. €)	116,4	64,5	48,57
Reducción de precios en productos (%)	2,3	1	2,5
Gastos en salarios (mill. €)	124,8	141,9	149,6
Proveedores de productos y servicios indirectos (nº)	1.834	1.796	1.765
Proveedores españoles de productos y servicios indirectos (nº)	1.753	1.720	1.687
Gasto en proveedores de productos y servicios indirectos (mill. €)	257	219	168
Gasto en proveedores españoles de productos y servicios indirectos (mill. €)	86,8	94,8	106
Visitantes tiendas IKEA (mill.)	38,2	37,5	39,4
Miembros IKEA Family (mill.)	4,5	4,9	5,3
Visitantes web IKEA.es (mill.)	58,2	70,2	83,7
Clientes IKEA Food (mill.)	10,5	10	11
Facebook (fans)	450.000	650.000	953.000
Twitter (fans)	42.000	72.000	104.000
Índice de satisfacción de clientes (0-100%)	76	74	77
Reclamaciones de clientes (ud.)	61.032	74.523	102.816
Número de tiendas (ud.)	15	16	17
Centros de distribución (ud.)	2	2	2
Tasa de llenado de camiones del centro de distribución a las tiendas (%)	59	62,3	64,92
Catálogos distribuidos (mill.)	8	8	10

### 3. CASO PRACTICO: DATA MINING

Para la parte práctica de este estudio como se ha señalado vamos a utilizar una fuente interna del departamento de marketing de la empresa IKEA España. Concretamente se trata la información que IKEA recoge de los clientes que tienen la tarjeta IKEA FAMILY. Este tipo de información está completamente anonimizada y cumple la Ley de Protección de datos.

La base utilizada corresponde a la ola del año 2015 (la proporción de clientes que permanecen en el panel del año anterior la desconocemos) y tiene una muestra de 1000 clientes de los que se recogen un total de 15 variables con las que el departamento de marketing realiza estudios para optimizar el envío de comunicaciones o predecir qué tipo de cliente tiene una mayor probabilidad de compra según características de su perfil, por lo tanto primero hablaremos de perfiles de clientes para amoldar los modelos LOGIT a ese tipo de perfiles.

La compañía está en un proceso interno de digitalización, como los que se han descrito en el punto 2 de este trabajo, para posicionarse a la vanguardia del sector, por ello cuenta con un departamento de analistas de datos que conforman y enriquecen las variables que son importantes para su negocio. En sus análisis incluyen nuevas variables que surgen del comportamiento del cliente y que pueden ajustarse mejor a la realidad que las variables introducidas por los propios clientes a la hora de darse de alta en IKEA FAMILY.

Un ejemplo que ilustra esta idea, resulta más fiable detectar a través del consumo de los clientes si tienen hijos y de qué edad aproximada que basarse en la información sobre si tiene hijos que el cliente relleno cuando se dio de alta. Por un lado, puede ser simplemente que dejó sin rellenar la pregunta y por otro, que en el momento en que se hizo la tarjeta, hace 2 años cuando abrió el centro, no tuviera hijos y ahora sí. Esta información se va a tener en cuenta a través del estudio de las correlaciones entre variables, imputar el número de hijos que tiene en el momento, en función de la compra de productos infantiles que ha realizado.

Este tipo de información también se pone en contexto por ejemplo con los tickets recogidos en IKEA FOOD asociados a una tarjeta de IKEA FAMILY para detectar comportamientos como que la gente que dijo que no tenía hijos realice algún tipo de consumo infantil (Menú Infantil disponible solo para menores de 14 años). Otro indicador que se mide de una forma parecida, buscando correlaciones es la renta media, que no se

pregunta directamente en ningún formulario pero que, sí se extrapola del nivel de estudios, zona de residencia y tipo de vivienda.

En definitiva, el departamento cuenta con distintos tipos de origen de datos. Dependiendo del origen, los datos hay que analizarlos a nivel individual o agregado en forma de indicadores. Los datos, pueden ser estructurados o no estructurados.

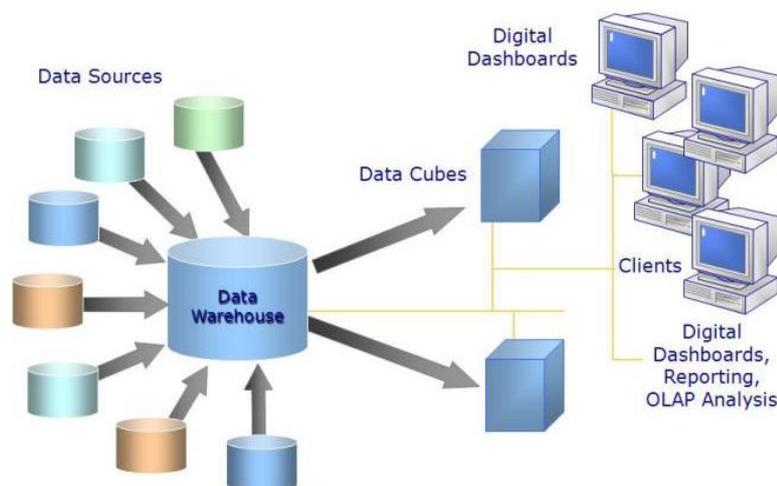
1. Individuales

- Formulario de inscripción
- Tickets de compra asociados a la tarjeta
- Comportamiento online – aplicación (No estructurados)

2. Agregados

- Dispositivos de tránsito y movimiento (No estructurados)
- Social Media (No estructurados)
- Tiempo atmosférico
- Tipos de decoración de la tienda.

Toda la información que recoge IKEA le permite analizar en tiempo real el comportamiento de sus tiendas y lo que sus clientes están demandando. Esto les aporta una ventaja competitiva que les permite ahorrar en costes por ejemplo en logística y comunicación comercial. En este segundo aspecto, la comunicación comercial, es en el que se va a centrar el análisis que vamos a realizar. Evidentemente en un trabajo de estas características, las variables no estructuradas en dispositivos de entrada complejos no van a formar parte de momento de nuestro análisis, pero sería interesante que, en trabajos posteriores, se pudieran incluir.



### 3.1 Descripción de las variables

- Variables que hacen referencia a las características demográficas del cliente IKEA FAMILY

V1) **Sexo**; Recogida en el formulario inicial. Dicotómica (Hombres o Mujeres)

V2) **Edad**; Recogida en el formulario inicial. Dividida en intervalos (de 18 a 30 años, de 31 a 40, de 41 a 55 y mayores de 55)

V3) **Nivel de Estudios**; Recogida en el formulario inicial. (Sin Estudios, Estudios Básicos, Estudios Medios y Estudios Superiores)

V4) **Estado Civil**; Recogida en el formulario inicial (En pareja o Soltero)

V5) **Tipo de casa**; Recogida en el formulario inicial. Inicialmente esta variable recoge multitud de combinaciones, pero queda reducida a tres categorías que recogen todas las opciones (Apartamento Pequeño <100m<sup>2</sup>, Apartamento Grande >100m<sup>2</sup> y Viviendas Unifamiliares >1 planta)

V6) **Hijos**; Variable obtenida a través de correlacionar historial de compra de productos infantiles de decoración o mobiliario y de la relación con los pedidos infantiles en IKEA FOOD. (Sí o No)

V7) **Renta Media**; Variable obtenida a través de la suma de tres indicadores: localización del domicilio (C.P.), tipo de vivienda y nivel de estudios. (Renta Baja, Renta Media-Baja, Renta Media, Renta Media-Alta, Renta Alta)

- Variables que hacen referencia al comportamiento del cliente IKEA FAMILY

V8) **Visitas 2015**; Número de visitas que ha realizado el cliente a lo largo del año 2015 (de 1 a 5, de 5 a 10 y >10)

V9) **Gasto Medio**; Este indicador se obtiene de realizar la media del total del ticket dividido por el número de visitas que ha realizado el cliente de IKEA FAMILY. Posteriormente se agrupa categóricamente de la siguiente manera (<50€, de 50 a 200€, de 200 a 500€ y >500€)

V10) **IKEAFOOD**; Variable dicotómica (Si, No) que se obtiene de analizar si en la mayoría de las ocasiones que el cliente visita la tienda, compra en IKEAFOOD. Analizando los clientes que consumen productos de IKEAFOOD, se contabilizan visitas del cliente a la tienda, aunque no haya comprado ningún otro producto de IKEA.

V11) **Promo Especial**; Variable dicotómica (Si, No) que analiza si el cliente compra frecuentemente artículos recogidos en las promociones especiales que son exclusivas para socios de IKEA FAMILY

V12) **Promo Email**; Variable dicotómica (Si, No) que analiza si el cliente interactúa frecuentemente con las comunicaciones que se realizan a los socios que así lo especifican por correo electrónico. Por ejemplo, encuestas de satisfacción (NPS) o presentaciones de catálogos de temporada.

V13) **Indicador NPS**; Indicador que permite medir la lealtad a la marca de los clientes basándose en las recomendaciones es una métrica fácil de obtener y de medir que refleja como los clientes perciben a la marca, calidad o atención al cliente. Este indicador se mide por correo electrónico. Y se pide que valore del 1 al 10 si recomendaría a IKEA. (Detractores, Pasivos y Promotores)

V14) **Eventos**; Indicador que muestra si el cliente ha asistido alguna vez a algún evento organizado por IKEA en alguna de sus tiendas. (Si, NO)

V15) **TRG Colchón**; Variable target para entrenar un modelo para predecir la probabilidad de que un determinado cliente compre un colchón de IKEA. Responde a la pregunta; ¿Ha comprado un colchón en IKEA? (Si, NO)

### 3.2 Análisis exploratorio

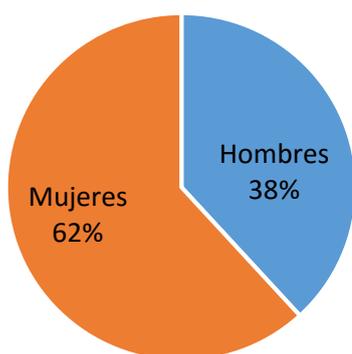
Para comenzar el análisis estadístico y antes de realizar el modelo logit binario hemos se ha realizado un análisis exploratorio de las variables más significativas para comprobar cuál es su distribución dentro de la muestra.

Las variables escogidas, nos aportan una visión sobre como es el cliente de IKEA FAMILY. En primer lugar, vemos en la variable sexo la primera diferencia significativa y es que un 62% de los clientes que tienen la tarjeta son mujeres. La mujer toma una especial relevancia para marcas como IKEA, porque en la mayoría de los casos es ella la que toma las decisiones sobre decoración o menaje en un hogar. Por eso no es significativo que la tarjeta de fidelización de IKEA FAMILY esté a nombre de mujeres.

Otra variable demográfica que la empresa tiene muy en cuenta es la edad de sus clientes, aunque IKEA vende sus productos en España desde hace 20 años, su crecimiento como marca se ha producido en los últimos 10 años llegando a tener hoy en día un total de 17 puntos de venta. Debido a ser una marca enfocada en combinar precios bajos y diseño funcional sigue las tendencias sociológicas actuales de cada país en el que vende sus productos, en España en concreto se centra en amoldar sus muebles a las casas de reducido tamaño y sacar el máximo partido al espacio disponible en ellas.

Gráfico 1. Distribución de la muestra por sexo.

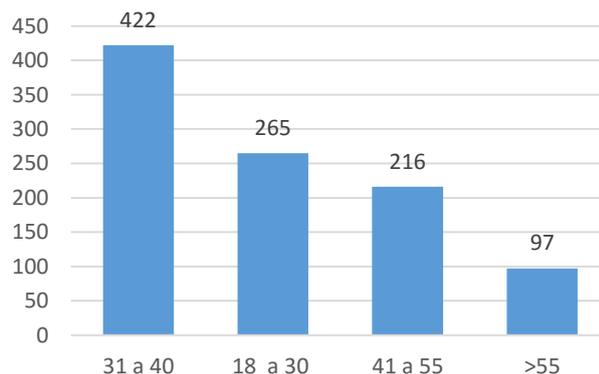
#### DISTRIBUCIÓN SEXO



Fuente: Elaboración propia

Gráfico 2. Distribución de la muestra por edad.

#### DISTRIBUCION EDAD

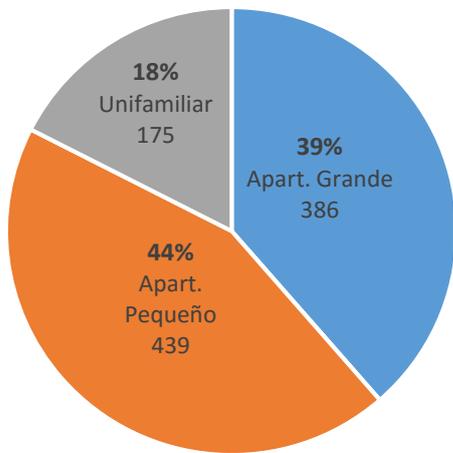


Fuente: Elaboración propia

Como se observa en el gráfico 2, la mayoría de los clientes tienen entre 31 y 40 años seguidos de la franja de edad de los 18 a los 30. En total los menores de 40 años representan casi el 70% de la muestra. Tan solo un 10% de los clientes tiene más de 55 años.

Gráfico 3. Distribución de la muestra por tipo de casa

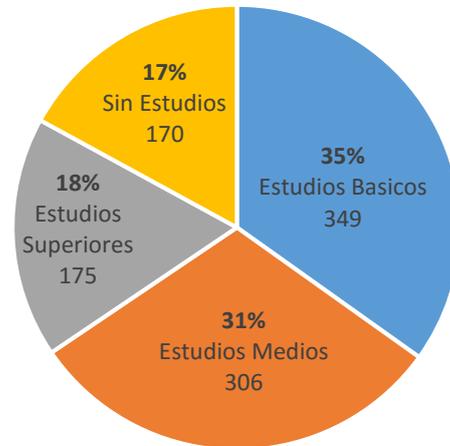
DISTRIBUCIÓN TIPO DE CASA



Fuente: Elaboración propia

Gráfico 4. Distribución de la muestra por tipo de estudios

DISTRIBUCIÓN TIPO DE ESTUDIOS



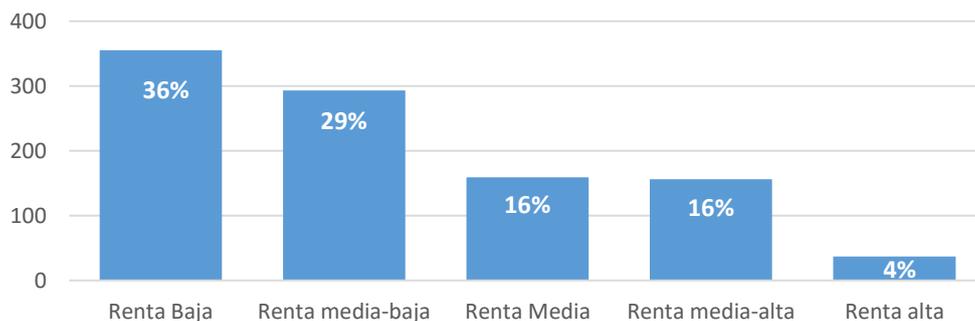
Fuente: Elaboración propia

Según el tipo de casa y el nivel de estudios que los clientes que poseen la tarjeta de IKEA FAMILY observamos que un 44% afirman vivir en apartamentos más pequeños de 100 m<sup>2</sup> y para más del 50% su nivel de estudios es básico o medios. Esta información está alineada con la estrategia de la empresa de precios baratos y decoración moderna para hogares con espacios reducidos y que responden a una nueva organización familiar.

Hemos comprobado, como se muestra en el gráfico 5, que la información obtenida concuerda con el nivel de renta media de la muestra. Solo un 4% de la muestra se encuentra en una situación de renta alta. Un 65% pertenece a perfil socio económico inferior que se sitúa entre rentas media-baja y baja. Por todo ello podemos afirmar que el perfil del comprador fidelizado de IKEA responde a una persona joven, que vive en un apartamento pequeño y que tiene una renta media-baja. En el apartado siguiente, trataremos de ver si los perfiles, se corresponden con este análisis exploratorio básico.

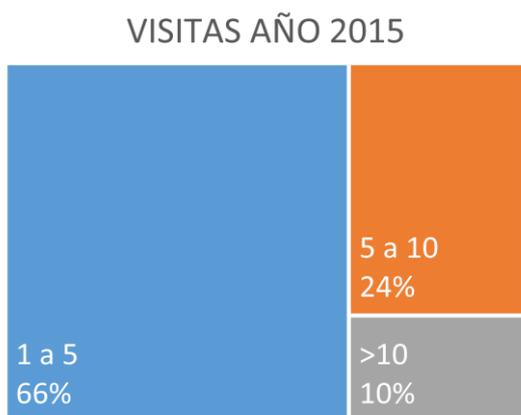
Gráfico 5. Distribución de la muestra por renta media.

DISTRIBUCIÓN RENTA MEDIA



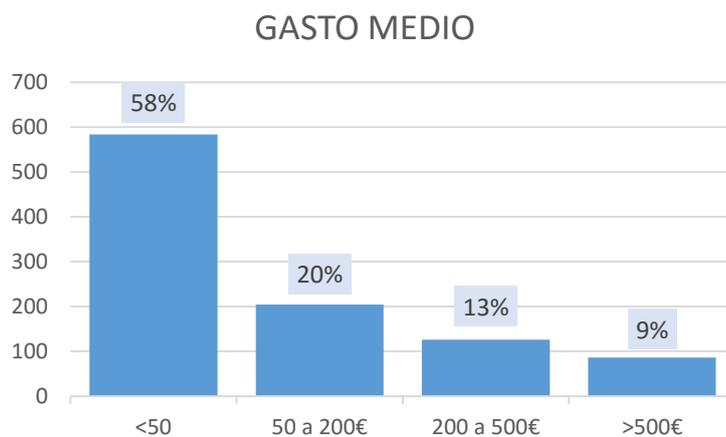
Fuente: Elaboración propia

Gráfico 6. Distribución de la muestra por visitas medias.



Fuente: Elaboración propia

Gráfico 7. Distribución de la muestra por gasto medio.

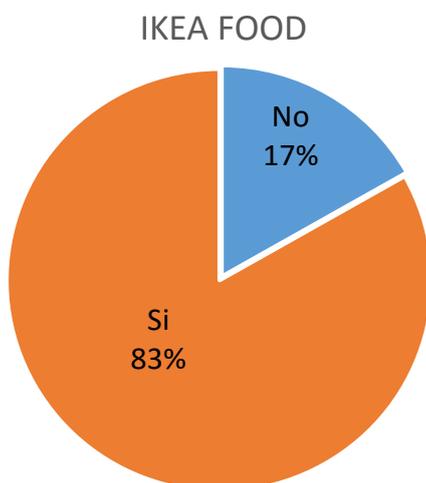


Fuente: Elaboración propia

Como se muestra en los gráficos 6 y 7 las variables del número de visitas al año y el gasto medio de ellas presentan muchas diferencias entre unos valores y otros. Por un lado, el 66% de los socios de IKEA FAMILY compran entre 1 y 5 veces al año en una tienda IKEA, lo que supone una tasa de recurrencia a priori relativamente baja. Por otro lado, el gasto medio de los clientes en un 58% de los casos se sitúa por debajo de los 50€.

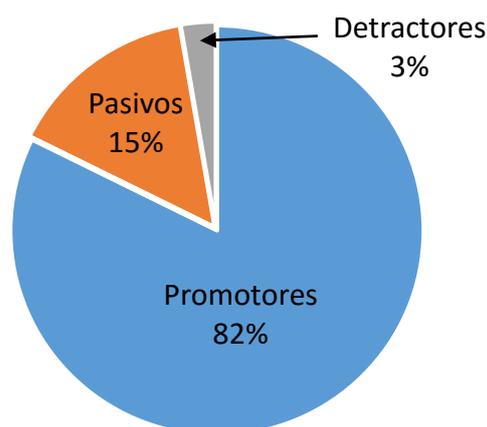
De todo ello deducimos que el tipo de compra es esporádico motivado por una necesidad concreta que lleva a visitar la tienda por un tipo de producto determinado.

Gráfico 8. Distribución de la muestra por consumo IKEA FOOD



Fuente: Elaboración propia

Gráfico 9. Distribución de la muestra por NPS NETWORK PROMOTORE SCORE



Fuente: Elaboración propia

Una de las claves del valor que aporta IKEA para sus clientes, es la experiencia de compra de sus tiendas. Motivada en cierta parte por su área de restauración IKEA FOOD, con productos típicos suecos. Un 83% de los clientes que tienen la tarjeta IKEA FAMILY compra de manera frecuente productos de IKEA FOOD. Para la marca conseguir que los

clientes comprenden estos productos y se sientan identificados con ellos, aporta a IKEA un valor diferencial a través de lo que ocurre entorno a sus productos.

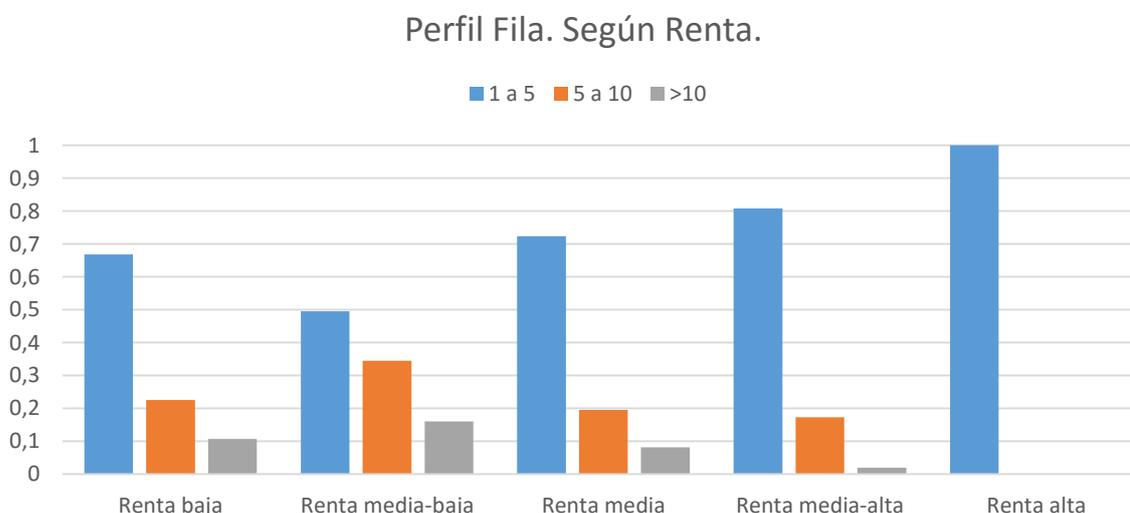
Todo ello sumado a la relación calidad precio, la variedad, diseño y a la atención del servicio post-venta, hacen que un 82% de los encuestados, recomiende IKEA a amigos y familiares. Cabe destacar que la encuesta de NPS se realiza por email y solo hay respuestas de un 15% de la muestra, por lo tanto, en un trabajo posterior, sería de interés hacer un modelo de regresión en el que imputásemos la falta de respuesta de NPS a través de otras variables con tasa de respuesta alta. Como ya hemos dicho, esto es un análisis exploratorio básico, vamos a ver si en el análisis de perfiles coincide con las expectativas del cliente que posee la tarjeta IKEA FAMILY.

### 3.3 Análisis de correspondencias

Una vez llevado a cabo el análisis descriptivo de las variables, debemos estudiar las posibles correspondencias entre las variables que hemos identificado más importantes. Si no identificáramos las relaciones entre las variables, difícilmente el departamento de marketing podría llevar a cabo una estrategia de optimización de recursos como la que plantea este trabajo.

Para detectar el perfil del cliente de IKEA FAMILY vamos a realizar un análisis de correspondencias. En primer lugar, vamos a analizar la variable que recoge el número de visitas que realiza el cliente a lo largo de un año en función de su nivel de renta.

Gráfico 10. Perfil Fila. Según Renta



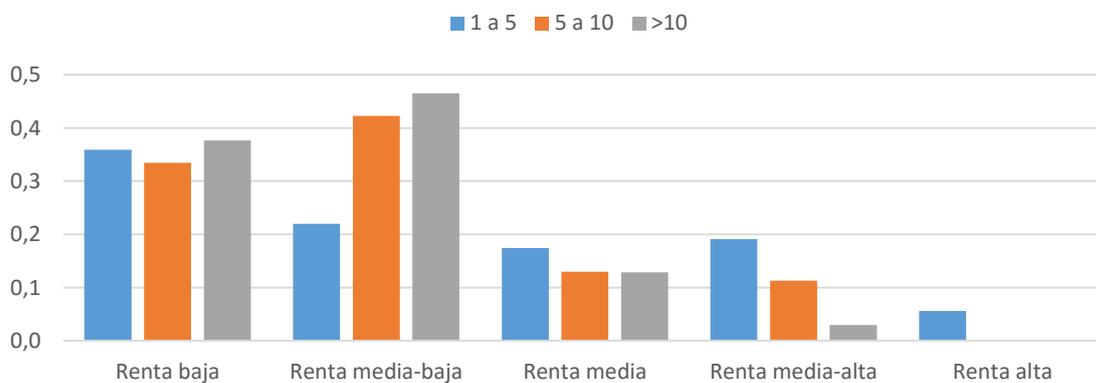
Fuente: Elaboración propia

Como se ve en el gráfico del perfil fila, las personas con una renta media-alta o alta, realizan pocas visitas a la tienda (entre 1 y 5), por lo que sus visitas no son recurrentes.

Por lo tanto, IKEA, debería plantear dos estrategias la fidelización de los clientes con rentas intermedias por un lado y captación de clientes en rentas altas por otro. Una estrategia que acercaría los perfiles de rentas altas más a la marca IKEA podría ser hacer que en sus catálogos se mostrarán casas con un diseño diferente al habitual, es decir, un formato en el que vea una casa de alto standing perfectamente amueblada con la decoración típica de IKEA. Esto supondría un cambio de paradigma en la cultura de cómo se amueblan los hogares en España.

Gráfico 11. Perfil Columna. Según Visitas

### Perfil Columna. Según Visitas.



Fuente: Elaboración propia

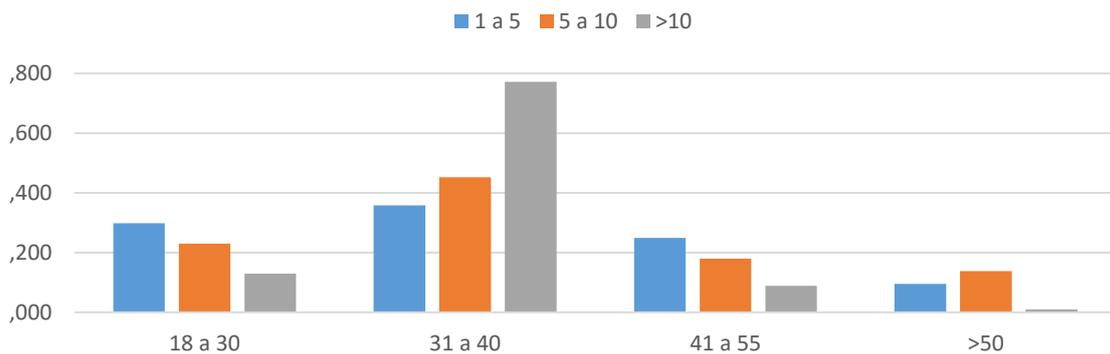
Si observamos el gráfico que recoge el perfil columna, los perfiles que visitan más de 10 veces al año son clientes con menor renta. Analizando este patrón de comportamiento deducimos que corresponde a que el nivel de renta influye mucho a la hora de comprar en IKEA, los perfiles con rentas superiores tienen otras alternativas para comprar productos sustitutivos porque para ellos el precio no es un factor decisivo.

Por otro lado, también analizamos que las rentas con menos ingresos, encuentran en IKEA un lugar de ocio, donde acudir a ver las novedades de decoración, disfrutar de la restauración que ofrece a precios económicos e incluso gracias a la guardería para los socios de IKEA FAMILY un lugar donde acudir en familia.

Para buscar más indicios sobre las variables que influyen en la frecuencia de visita a las tiendas de IKEA hemos analizado la edad. Como se observa en el gráfico 12 que muestra el perfil fila, el 75% de la gente que visita la tienda con una frecuencia superior a 10 veces al año tiene entre 31 y 40 años. Lo que guarda relación con lo expuesto anteriormente, faltaría por comprobar que en esa franja es donde más familias con niños se encuentra.

Gráfico 12. Perfil Fila. Según el número de visitas

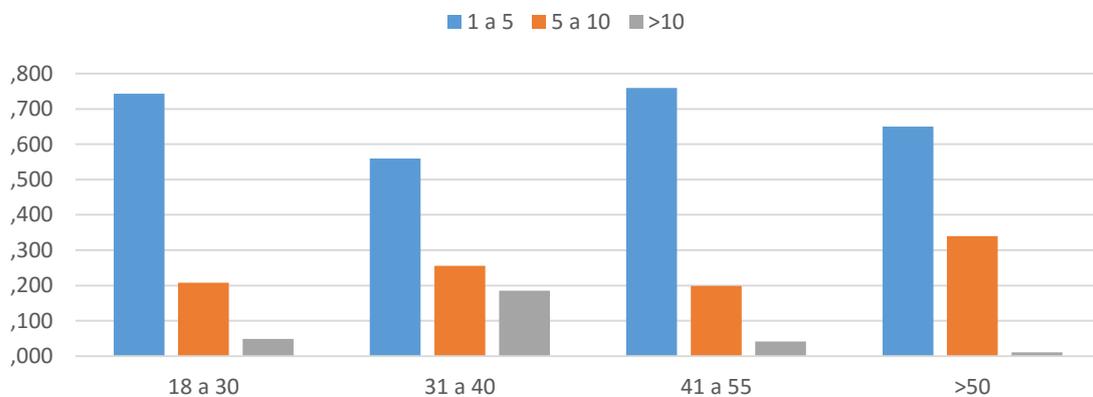
Perfil Fila. Según el numero de visitas



Fuente: Elaboración propia

Gráfico 13. Perfil Columna. Según la edad

Perfil Columna. Según la Edad



Fuente: Elaboración propia

Si observamos el gráfico 13 que muestra el perfil columna, deducimos que independientemente de la edad, lo que predomina es una tasa de recurrencia menor a 5 visitas por año, a pesar de los esfuerzos de la marca por atraer a los clientes a su tienda. Esto es debido al tipo de producto que oferta IKEA siendo un producto estacional y con consumidores muy informados con mucha oferta donde poder elegir.

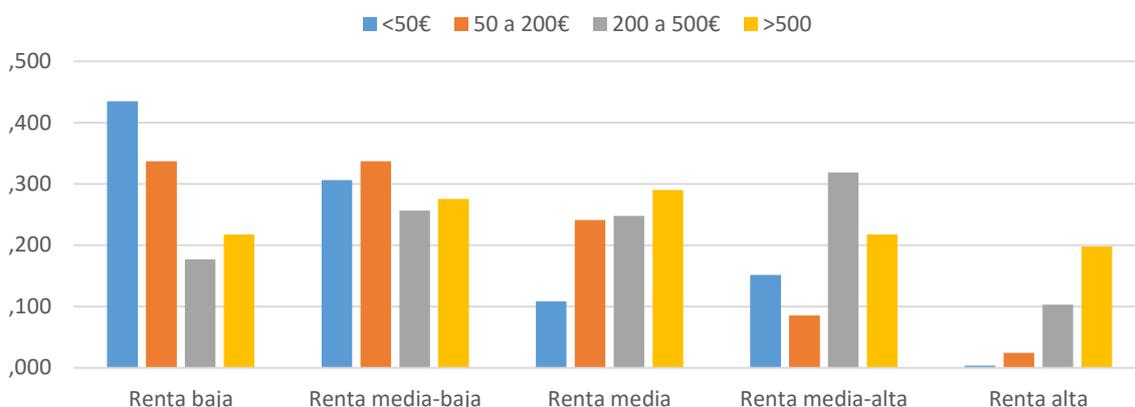
A continuación, hemos analizado cuál es el gasto medio en función del nivel de renta. En el gráfico 14 que muestra el perfil fila observamos que el 74% de los clientes que se gastan

menos de 50€ de media en IKEA tienen renta baja o media-baja. Por el contrario, el gasto superior a 500€ se encuentra repartido por los distintos niveles de renta, esto es debido a la gama extensa y distribuida de productos que posee IKEA.

Los perfiles de rentas bajas, encuentran en IKEA electrodomésticos o mobiliario asequible a su nivel, esto supone que haya compras de alto valor económico, pero realmente sean productos de gama “low cost”.

Gráfico 14. Perfil Fila. Según el gasto medio

Perfil Fila. Según el Gasto medio

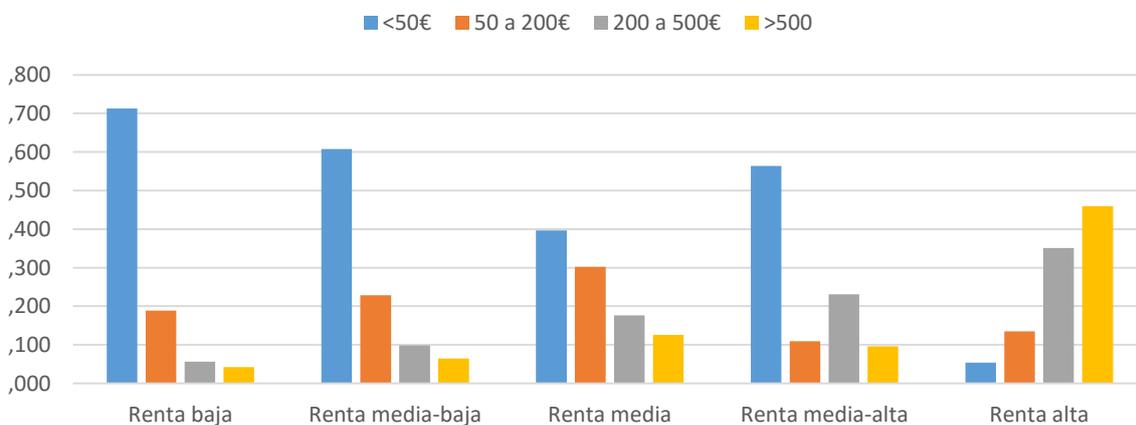


Fuente: Elaboración propia

En el gráfico 15 que corresponde al perfil columna observamos cómo el 80% de los clientes con mayor poder adquisitivo realizan compras superiores a los 200€. En el otro extremo, dentro de los perfiles con menores rentas el 70% tiene un consumo medio menor a 50€. De estos datos y los anteriores, se deduce que los perfiles de rentas altas realizan menos visitas a las tiendas de IKEA, pero cuando lo hacen su consumo es mayor, por el

Gráfico 15. Perfil Columna. Según la renta

Perfil Columna. Según Renta

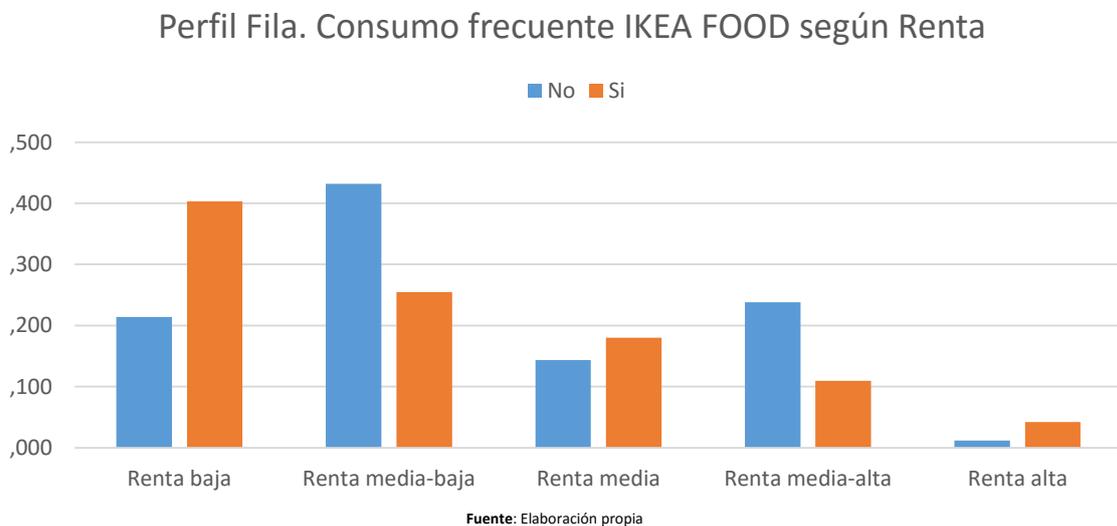


Fuente: Elaboración propia

contrario, los perfiles con rentas más bajas visitan con más frecuencia las tiendas, pero su consumo en ellas es menor.

Siguiendo con el análisis sobre el comportamiento de los clientes de IKEA hemos detectado un hecho que concuerda con el análisis realizado hasta ahora, se trata del consumo de los clientes en las áreas de restauración IKEA FOOD. En el grafico X que corresponde al perfil fila, observamos como la proporción de consumir frecuentemente productos de IKEA FOOD, disminuye conforme aumenta la renta de las clientes. Esto es debido a que conforme mayor es el nivel adquisitivo disminuye la intención de consumir comida “rápida” en el establecimiento de compra.

Gráfico 16. Perfil Fila. Consumo IKEA FOOD según renta.



Al analizar el grafico del perfil columna, se observa como independientemente de la renta, el consumo frecuente de IKEA FOOD al realizar una visita al establecimiento es muy superior a la gente que no consume. Esto se debe a los precios atractivos y la situación estratégica de los establecimientos de restauración. IKEA ha conseguido que la experiencia del consumidor que pasa por sus tiendas esté relacionada con el consumo de sus productos.

### 3.6 Análisis de regresión

Para finalizar con la parte práctica de este proyecto vamos a llevar a cabo un análisis de regresión con variable dependiente categórica. El modelo que buscamos a priori con este tipo de análisis, es un modelo que explique de que variables son más influyentes y de qué forma aumenta o disminuye la compra de un tipo de colchón específico de IKEA. Por lo tanto, la variable dependiente que vamos a estudiar va a ser si el cliente ha comprado o no ha comprado dicho colchón, y las variables explicativas son aquellas que están relacionadas con el perfil del comprador:

- **EDAD**; Variable categórica definida de 18 a 30 años, de 31 a 40, de 41 a 55 y mayores de 55. Esta variable se recoge en el formulario inicial.
- **ESTUDIOS**; Nivel de estudios de la persona titular de la tarjeta, dividida categóricamente en los rangos; Sin Estudios, Estudios Básicos, Estudios Medios y Estudios Superiores.
- **HIJOS**; Variable obtenida a través de correlacionar el historial de compra de productos infantiles de decoración o mobiliario y de la relación con los pedidos infantiles en IKEA FOOD. Dividida de forma dicotómica, sí o no.
- **SEXO**

Las variables que hemos seleccionado son todas demográficas, porque de esta manera nos permiten ajustar el modelo con variables primarias, que podemos recoger en el formulario inicial, y que no se conforman con a lo largo del tiempo. En el caso de la variable HIJOS es una correlación que incluye los datos del formulario inicial con el historial de compra de productos infantiles que puedan mostrar que el cliente tiene hijos y no lo había especificado en el formulario inicial.

El método utilizado es el habitual en estos modelos como en los de regresión con variable dependiente cuantitativa, es decir, regresión por pasos hacia delante. Esto es en cada paso va introduciendo aquellas variables que en ese paso son significativas. En el paso uno, incluye la variable más importante del modelo (en nuestro caso es la variable EDAD). Y en los pasos sucesivos continúa incluyendo aquellas variables significativas, mientras que las que no lo son las deja fuera del modelo.

Para determinar la calidad del modelo, tenemos que centrar nuestra atención en los siguientes estadísticos:

- Prueba ómnibus: indica la significatividad del modelo según los diferentes pasos en los que se incluyen nuevas variables al modelo. Tenemos que fijarnos en concreto en la significatividad de los distintos pasos para aceptar el modelo. Como nuestra regresión logística es por pasos, nos va a indicar los valores de Chi-cuadrado en los distintos pasos.

Tabla 1. Prueba ómnibus sobre los coeficientes del modelo

		Chi cuadrado	gl	Sig.
Paso 1	Paso	401,203	3	,000
	Bloque	401,203	3	,000
	Modelo	401,203	3	,000
Paso 2	Paso	181,196	3	,000
	Bloque	582,399	6	,000
	Modelo	582,399	6	,000
Paso 3	Paso	57,455	1	,000
	Bloque	639,854	7	,000
	Modelo	639,854	7	,000
Paso 4	Paso	59,685	1	,000
	Bloque	699,539	8	,000
	Modelo	699,539	8	,000

Fuente: Elaboración propia

En el primer paso observamos que el Chi-cuadrado del modelo es significativo con  $p < 0,001$ , es decir, las variables independientes describen la variable dependiente de forma significativa. En los pasos sucesivos observamos que hay una reducción del estadístico Chi-cuadrado pero el modelo sigue siendo significativo ( $p < 0,001$ ).

- Estadístico de Hosmer-Lemeshow: en el caso de que el p-valor para este estadístico fuera menor de 0,05 nos encontraríamos con un mal ajuste del modelo y no nos serviría. En este ejemplo, el estadístico tiene un valor de 0,287, es decir, un buen ajuste en el paso 2.

Tabla 2. Prueba de Hosmer - Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,000	2	–
2	8,548	7	,287

Fuente: Elaboración propia

- Resumen del modelo: nos indica el valor del estadístico  $\ll -2 \log$  de la verosimilitud  $\gg$  útil para hacer comparaciones, así como dos valores de R cuadrado. El  $r^2$  de Cox y Snell no varía entre 0 y 1 por lo que se ha modificado en lo que se llama el  $r^2$  de Nagelkerke que sí varía entre 0 y 1. En nuestro caso, en el paso 4, el modelo explica un 73,5% de la variabilidad de los datos.

Tabla 3. Resultado del modelo

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	753,196 <sup>a</sup>	,330	,483
2	571,999	,441	,645
3	514,544	,473	,690
4	454,86	,503	,735

a. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: Elaboración propia

Estos tres últimos estadísticos los podemos utilizar para comparar en el modelo las diferentes versiones que podemos obtener de él (ya que en cada paso se introduce una nueva variable). Modelos con la R cuadrado más alta o menor estimación en la función de verosimilitud será el que mejor modelo. En nuestro caso, nos encontramos con un R cuadrado muy alto, por lo que es un modelo que explica muy bien la variabilidad de los datos.

Para seguir con el análisis del modelo hemos analizado la tabla de clasificación la cual nos indica el porcentaje de clasificaciones correctas de nuestro modelo. En dicha tabla observamos que en el paso 4 el modelo predice correctamente el 97,6% de las personas que no han comprado el colchón. Sin embargo, lo importante es que es capaz de predecir correctamente el 78,4% de las personas que si lo han comprado. Por lo tanto, es un modelo acertado para predecir cuando una persona potencialmente puede comprar este tipo de colchón.

Tabla 4. Tabla de clasificación

Observado			Pronosticado		
			¿Ha comprado un COLCHON IKEA?		Porcentaje correcto
			No	Si	
Paso 1	¿Ha comprado un COLCHON IKEA?	No	635	101	86,3
		Si	52	212	80,3
	Porcentaje global				84,7
Paso 2	¿Ha comprado un COLCHON IKEA?	No	681	55	92,5
		Si	91	173	65,5
	Porcentaje global				85,4
Paso 3	¿Ha comprado un COLCHON IKEA?	No	656	80	89,1
		Si	50	214	81,1
	Porcentaje global				87,0
Paso 4	¿Ha comprado un COLCHON IKEA?	No	718	18	97,6
		Si	57	207	78,4
	Porcentaje global				92,5

a. El valor de corte es ,500

Fuente: Elaboración propia

Por último, hemos analizado la tabla de las variables de la ecuación seleccionadas en cada paso. Un dato de especial importancia que encontramos en esta tabla es el ODDS ratio (columna “Exp (B)”). Esta ratio indica el coeficiente entre la probabilidad de que ocurra el suceso que indica la variable dependiente (comprar el colchón), nuestra variable de referencia en nuestro modelo, frente a la probabilidad de que no ocurra en presencia o ausencia de las variables explicativas.

El valor 67,67 que encontramos en la variable “Nivel de estudios (3) = Estudios Medios”, nos indica que es 67,67 veces más grande el cociente entre la probabilidad de comprar el colchón que de no comprarlo en aquellas personas que tienen estudios medios. De esta misma forma, el valor 10,42 correspondiente a la variable “Hijos = Tiene hijos”, indica que el cociente entre la probabilidad de comprar el colchón que de no comprarlo es 10,42 veces más grande en aquellas personas que tienen hijos. Todo ello respecto a la variable de referencia que calcula por defecto el programa estadístico SPSS.

Por último, encontramos el valor 8,54 en la variable “Sexo (1) = Mujer” lo que nos indica que es 8,54 veces más grande el cociente entre la probabilidad de comprar el colchón que de no hacerlo si el cliente es mujer. Como comentábamos al principio del caso son las mujeres las que mayoritariamente toman las decisiones de compra en IKEA por lo que este indicador guarda concordancia.

Tabla 5. Variables en la ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)
Paso 1 A.	EDAD			291,223	3	,000	
	EDAD (1)	-4,106	,376	119,437	1	,000	,016
	EDAD (2)	-3,314	,285	134,772	1	,000	,036
	EDAD (3)	-,447	,272	2,692	1	,101	,640
	Constante	1,058	,232	20,763	1	,000	2,880
Paso 2 B.	EDAD			83,813	3	,000	
	EDAD (1)	-7,537	,831	82,204	1	,000	,001
	EDAD (2)	-6,439	,784	67,454	1	,000	,002
	EDAD (3)	-,403	,305	1,752	1	,186	,668
	ESTUDIOS			46,690	3	,000	
	ESTUDIOS (1)	-1,402	,474	8,753	1	,003	,246
	ESTUDIOS (2)	-1,521	,465	10,699	1	,001	,218
	ESTUDIOS (3)	3,623	,752	23,198	1	,000	37,448
	Constante	1,904	,487	15,279	1	,000	6,711
Paso 3 C.	EDAD			63,171	3	,000	
	EDAD (1)	-6,607	,861	58,859	1	,000	,001
	EDAD (2)	-6,014	,806	55,726	1	,000	,002
	EDAD (3)	-,306	,328	,870	1	,351	,736
	HIJOS	1,730	,236	53,730	1	,000	5,643
	ESTUDIOS			43,466	3	,000	
	ESTUDIOS (1)	-1,194	,498	5,755	1	,016	,303
	ESTUDIOS (2)	-1,270	,486	6,837	1	,009	,281
	ESTUDIOS (3)	3,889	,756	26,486	1	,000	48,873
	Constante	,440	,547	,647	1	,421	1,553
Paso 4 D.	EDAD			68,115	3	,000	
	EDAD (1)	-6,637	,859	59,684	1	,000	,001
	EDAD (2)	-6,048	,804	56,603	1	,000	,002
	EDAD (3)	-,056	,359	,025	1	,875	,945
	HIJOS	2,345	,278	71,100	1	,000	10,428
	SEXO (1)	2,145	,308	48,407	1	,000	8,543
	ESTUDIOS			47,641	3	,000	
	ESTUDIOS (1)	-1,397	,531	6,914	1	,009	,247
	ESTUDIOS (2)	-,756	,518	2,132	1	,144	,470
	ESTUDIOS (3)	4,215	,795	28,133	1	,000	67,675
	Constante	-1,682	,624	7,258	1	,007	,186

A. Variable(s) introducida(s) en el paso 1: EDAD.

B. Variable(s) introducida(s) en el paso 2: ESTUDIOS.

C. Variable(s) introducida(s) en el paso 3: HIJOS.

D. Variable(s) introducida(s) en el paso 4: SEXO.

Fuente: Elaboración propia

Con los coeficientes de las ecuaciones que hemos obtenido en el paso 4, vamos a simular la probabilidad de compra de dos casos posibles como los que se recogen en nuestra muestra y luego analizaremos el perfil ideal que maximiza la probabilidad.

- Mujer de 50 años con hijos y sin estudios. Los coeficientes utilizados son -0,56 por Edad (3), 2,345 por tener hijo, 2,145 por ser mujer y -1.397 por no tener estudios. El resultado de la fórmula es que, en este caso, tiene un 80% de probabilidad de comprar el colchón por lo que lo incluiríamos en el envío de publicidad comercial.

$$P1(\text{comprar colchon}) = \frac{1}{1 + e^{-(-1,682-0,56+2,345+2,145-1,397)}} = 0.795$$

- Mujer de 28 años con hijos y estudios básicos. Los coeficientes utilizados son -6,04 por Edad (1), 2,345 por tener hijo, 2,145 por ser mujer y 4.2147 por tener estudios básicos. El resultado nos dice que en este caso tiene un 72,6% de comprar el colchón.

$$P2(\text{comprar colchon}) = \frac{1}{1 + e^{-(-1,682-6,04+2,345+2,145+4,214)}} = 0.726$$

- Mujer de 42 años con hijos y estudios medios. Los coeficientes utilizados son -0,05 por Edad (1), 2,345 por tener hijo, 2,145 por ser mujer y 4.2147 por tener estudios medios. Según el modelo calculado, el perfil que maximiza la probabilidad de compra es Mujer entre 41 a 55 años con hijos y estudios medios. En este caso, la probabilidad de comprar un colcho de IKEA es del 98%

$$Pmax(\text{comprar colchon}) = \frac{1}{1 + e^{-(-1,682-0,05+2,345+2,145+4,214)}} = 0.989$$

#### 4. CONCLUSIONES

Con este trabajo se ha pretendido poner en valor el gran interés que para las empresas tiene el tratamiento de la información para generar conocimiento, en particular, en el campo del marketing y la investigación de mercados. Así pues, hemos analizado en el caso práctico al disponer de un sistema de información completo y tomar decisiones en base a esos datos, lo que supone maximizar el beneficio de determinadas acciones estratégicas. Gracias al modelo de regresión que hemos planteado podríamos dar respuesta a este problema de negocio; A qué clientes de IKEA FAMILY de los 5,3 millones de socios, tiene interés enviarles un nuevo folleto con este tipo de colchón propuesto en nuestro hipotético caso empírico ya que tiene un coste para la empresa de un 1€ por folleto. Si el departamento de marketing tiene asignados 100.000€ de presupuesto para realizar esta acción: ¿Cómo seleccionaríamos esos 100.000 clientes?

La respuesta nos la da precisamente el modelo de regresión binario planteado, con un porcentaje de acierto de más de un 90% y con un ODD ratio muy superiores al 50% para los perfiles de clientes según edad y estudios que en el modelo hemos visto. Concretamente uno de los perfiles que maximiza la probabilidad de compra (98%) es una mujer de 42 años con al menos 2 hijos y estudios medios.

Por lo tanto, plantear modelos de minería de datos dan respuesta a esta pregunta y otras similares, lo que supone una gran ventaja competitiva para la empresa ya que nos permite conocer mejor el entorno competitivo y a nuestros clientes. Una cultura basada en la transformación digital que pone en el centro de todo al cliente y en la gestión de los datos, ayudará en el futuro próximo de cualquier empresa a desarrollarse y crecer. Para conseguirlo, es necesario un cambio de mentalidad por parte de las personas que lideran las compañías y la absorción de herramientas y perfiles tecnológicos, para explotar y sacar el máximo rendimiento de las bases de datos propias de la empresa y de esta forma optimizar los recursos de la empresa.

Este trabajo pone de manifiesto la necesidad de intensificar los esfuerzos en materias estadísticas, de extracción y modelización de datos y de comprensión de las decisiones que estratégicamente se pueden tomar con datos, en estudios de Marketing e Investigación de mercados. En los próximos años, el perfil de analista de datos de marketing también conocido como **Marketing Intelligence**, aportará mucha más relevancia a los departamentos de marketing dentro de las organizaciones.

## 5. BIBLIOGRAFIA

- SIEGEL, E. (2013) *Predictive analytics: the power to predict who will click, buy, lie, or die*. Wiley. Hoboken.
- GUISANDE, C. *et al*; (2011) *Tratamiento de datos con R, STATISTICA y SPSS*. Díaz de Santos. Madrid.
- GEORGE, J. y DESPINA, A. (2000) *The impact of internet use on Business-to-Business Marketing*. Elsevier Science Inc. New York
- URQUIZU, P. *Ejemplo DSS: Pañales y Cervezas* (WEB) <http://www.businessintelligence.info/dss/ejemplo-data-mining-panales-y-cerveza.html> , 29/10/2009 [17/02/2016]
- BARTON, D. y COURT, D. (2012) Making Advanced Analytics Work for You. *Harvard Business Review*. [www.hbr.org/2012/10/making-advanced-analytics-work-for-you/ar/4](http://www.hbr.org/2012/10/making-advanced-analytics-work-for-you/ar/4) [26/01/2016]
- PRESS, G. *A Very Short History Of Big Data* (WEB) <http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/#5e6e993455da> 9/05/2013 [18/03/2016]
- MIKE, O. (2010) *HADOOP: Scalable, flexible data storage and analysis* (WEB) [http://www.cloudera.com/content/dam/cloudera/Resources/PDF/Olson\\_IOT\\_Quarterly\\_Spring\\_2010.pdf](http://www.cloudera.com/content/dam/cloudera/Resources/PDF/Olson_IOT_Quarterly_Spring_2010.pdf)
- POLO, F. y MAGALHAES, V. (2014) *Transformación digital: reinventando organizaciones ante la nueva revolución industrial* (WEB) Territorio Creativo <https://www.territoriocreativo.es/etc/2014/11/transformacion-digital-whitepaper.html>
- MAÑE, S. (2016) *Qué es el Big Data y cómo usarlo en el marketing* (WEB) Comunidad IEB SCHOOL [http://comunidad.iebschool.com/iebs/marketing-digital/como-aplicar-el-big-data-en-el-marketing/?utm\\_source=face&utm\\_medium=feed](http://comunidad.iebschool.com/iebs/marketing-digital/como-aplicar-el-big-data-en-el-marketing/?utm_source=face&utm_medium=feed) [16/05/2016]
- PULIDO, E. (2014). Big data: solución o problema. Lección inaugural curso 2014-2015 en la Universidad Autónoma de Madrid. <http://arantxa.ii.uam.es/~epulido/bigdata.pdf>
- PENTASOFT, *Club de fidelización IKEA FAMILY* (WEB) <http://www.pentasoftware.com/cliente/ikea/club-de-fidelizacion-ikea-family> [19/05/2016]
- DIOSDADO, S. (2014) *El cliente ya no es multicanal, es omnicanal* (WEB) <http://www.samuediosdado.com/02/el-cliente-ya-es-multicanal-es-omnicanal/> [19/05/2016]
- GARCIA-ABRIL, C. y BLANCO, C. (2016) *Lecciones de éxito de Google, Apple, Facebook y Amazon para la digitalización* (WEB) <https://www.territoriocreativo.es/etc/2016/05/lecciones-de-exito-de-google-apple-facebook-y-amazon-para-la-digitalizacion.html> [19/05/2016]
- PEREZ, M. (2016) *Amazon crea un método para enviar pedidos antes de que se compren basado en Big Data* (WEB) <http://blogthinkbig.com/metodo-de-envios-anticipados-de-amazon/> [19/05/2016]