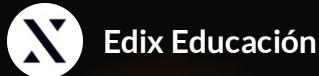


# Fastbook 03

## Medición de Eficacia Publicitaria

Fases de un proyecto de MMM y  
variables a modelizar



## 03. Fases de un proyecto de MMM y variables a modelizar

En este fastbook vamos a detallar las **fases en las que se divide habitualmente un proyecto de MMM** y lo importante que resulta no saltarse alguna de ellas. Además, vamos a revisar las **principales variables de negocio, tracking y web que se modelizan en este tipo de proyectos**.

Por tanto, el objetivo de este fastbook es continuar profundizando en el enfoque de MMM hablando, para ello, de las fases y de los datos.

*Autor: Carlos Real Ugena*

**Fases en un proyecto de MMM**

**Transformación de datos**

**Variables a modelizar**

**Variable de tracking**

**Métricas web**

Lesson 1 of 5

# Fases en un proyecto de MMM

 Edix Educación

---

Cada anunciante o compañía dedicada a la medición tendrá su propia metodología de trabajo. **Las fases de un proyecto de MMM tienen que garantizar que desarrollamos una medición fiable y que los resultados obtenidos se aproximan lo máximo a la realidad.**

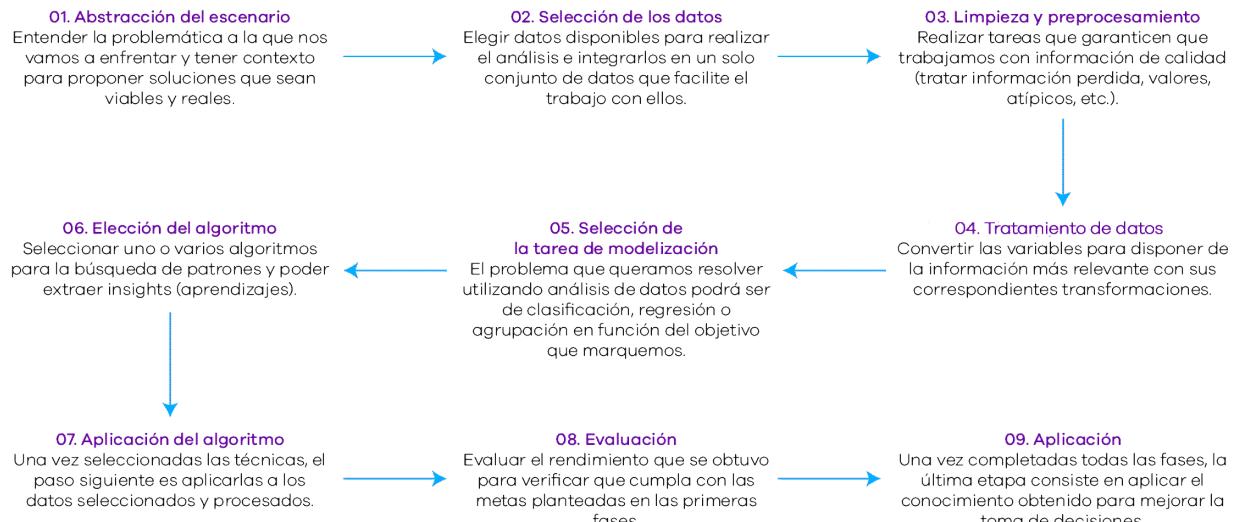
**¿Qué fases son recomendables abordar?** Esta tipología de análisis se engloba dentro de la categoría de proyectos de data mining, machine learning o inteligencia artificial. Por lo tanto, podemos utilizar un **framework (sistema)** que nos ayude a su planificación, desarrollo y mantenimiento.

---

KDD (*knowledge discovery in databases* o descubrimiento de conocimiento en bases de datos) es el proceso metodológico que se refiere al (amplio) concepto de extraer conocimiento de los datos.

A continuación, vemos en qué consiste este proceso metodológico y posteriormente nos centraremos en cómo aplicarlo a MMM.

## 9 pasos en un proyecto de KDD (*knowledge discovery in databases*)



**¿Cómo adaptamos los pasos anteriores a un proyecto de MMM?**

La forma más sencilla es verlo a través de un ejemplo.

Imaginemos que trabajamos en el equipo de data science de una compañía de perfumes de lujo y que vamos a tener que llevar a cabo un proyecto de medición del ROI a través de MMM.



Vamos a ver cómo se adaptan cada uno de los **pasos de KDD** a nuestro caso de modelización de perfumes.

## 1

### Abstracción del escenario

Lo primero que haremos es hablar con el equipo de marketing y con los expertos de negocio para **decidir la variable o variables dependientes que vamos a analizar**. Por ejemplo, la compañía podría estar interesada en analizar las ventas que se producen en un **establecimiento físico**, normalmente, denominado como Brick & Mortar en la industria de fragancias, y por el **canal online** (ecommerce).

2

## Selección de los datos

Ahora, tenemos que llevar a cabo una labor de investigación para tener en cuenta todas las fuentes de datos que vamos a emplear. Además de recolectar la información de ventas, necesitamos recopilar la información sobre los indicadores que pueden afectar a dichas ventas.

Tendremos que decidir la **información de publicidad (offline y online)** que vamos a incorporar al análisis, así como todos los indicadores que puedan afectar a las ventas.

---

**La distribución (número de tiendas y webs en las que se vende el producto), promociones, precios y variables macroeconómicas son algunos de los indicadores que debemos considerar.**

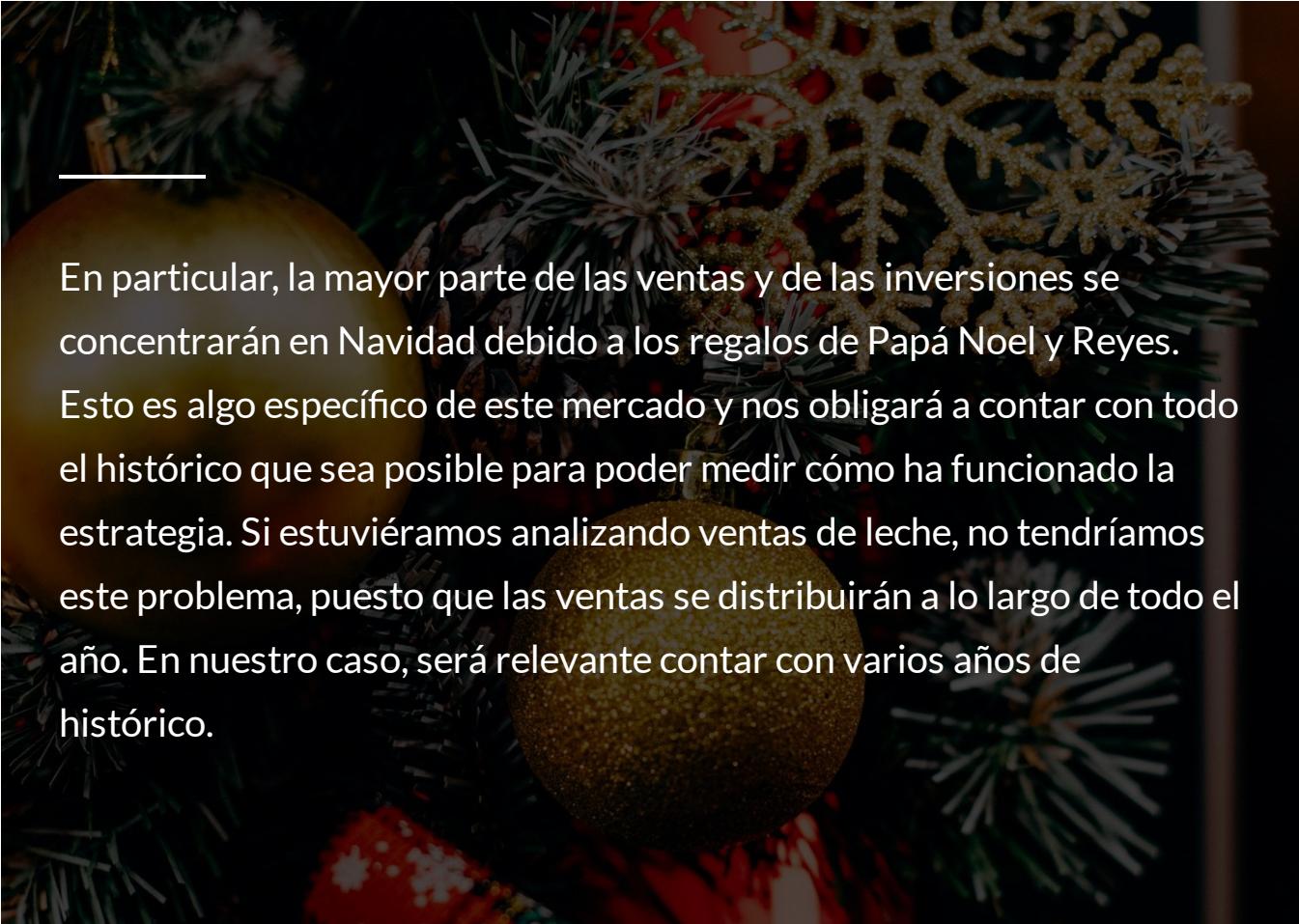
---

Además, incorporaremos fenómenos que puedan haber afectado al negocio de manera significativa como el covid-19, ¿piensas que las ventas de perfumes de lujo se han mantenido estables durante los meses de confinamiento?

3

## Limpieza y preprocesamiento

**Analizaremos la calidad de la información y decidiremos en base a esto, qué histórico de datos vamos a modelizar.** En el caso de perfumes de lujo observaremos que hay una fuerte estacionalidad positiva durante los periodos de Navidad, Día de la Madre si es fragancia femenina, Día del Padre si es masculina, San Valentín o Black Friday.



En particular, la mayor parte de las ventas y de las inversiones se concentrarán en Navidad debido a los regalos de Papá Noel y Reyes. Esto es algo específico de este mercado y nos obligará a contar con todo el histórico que sea posible para poder medir cómo ha funcionado la estrategia. Si estuviéramos analizando ventas de leche, no tendríamos este problema, puesto que las ventas se distribuirán a lo largo de todo el año. En nuestro caso, será relevante contar con varios años de histórico.

Una vez decidido el histórico en base a la información, tendremos que limpiar y preprocesar la información, tratando, por ejemplo, la información faltante y decidiendo qué hacer **con las observaciones atípicas**. Por ejemplo, podríamos encontrarnos ante dos situaciones distintas si no disponemos de ventas durante una semana:

#### Problema en la recogida de la información

No se almacenaron las ventas en una semana por fallo en el sistema tecnológico, sin embargo, sabemos que las ventas no fueron cero. En este caso, debemos estimar las posibles ventas de la semana o incluso eliminar la observación de nuestro conjunto de datos.

## No hay ventas

Debido a un problema de ventas por algún factor meteorológico como una gran nevada. En este caso, debemos imputar a cero las ventas y crear una variable que nos indique el periodo de duración de ese fenómeno. Otra opción sería eliminar esa semana para la modelización si consideramos que es un evento muy atípico que no se volverá a dar hasta dentro de varios años.

4

## Transformación de datos y análisis descriptivo

Una vez definidas las variables dependientes (ventas por canal), las variables independientes (publicidad, precio, promociones, etc.) y haber tratado la información faltante o atípica, debemos transformar la información y construir nuestra tabla de modelización.

Además, aprovecharemos esta fase para realizar el análisis descriptivo de la información. Es decir, realizaremos un análisis exploratorio de la información. A través del **análisis gráfico y de correlaciones**, entenderemos patrones de comportamiento de las series a analizar.

En resumen, esta es **una de las fases más importante de un proyecto de MMM** porque nos servirá para identificar comportamientos que habrá que tener en cuenta en la modelización. Una buena representación gráfica de los datos será fundamental.

5

## Selección de la tarea de modelización

Aquí no hay duda, **emplearemos modelos de regresión**. Esto nos permitirá analizar de qué depende el negocio a lo largo del tiempo. En nuestro caso de perfumes, aplicaremos dos modelos de regresión, uno para explicar las ventas del canal físico y otro para explicar las ventas del canal online.

## 6

## Elección del algoritmo

Existen múltiples modelos de regresión que se pueden emplear para medir el impacto de las **variables independientes**. Algunos ejemplos: regresión múltiple, Loglineal, TVP, Prophet, BSTS o Robyn. En nuestro caso de modelización de ventas de perfumes, nos haremos la siguiente pregunta: ¿qué conocimiento técnico tiene nuestro interlocutor?

- Si la respuesta es **conocimiento bajo**, optaremos por un enfoque de regresión múltiple que es más sencillo de entender.
- Si la respuesta es **medio o alto**, valoraremos el resto de los enfoques.

## 7

## Aplicación del algoritmo

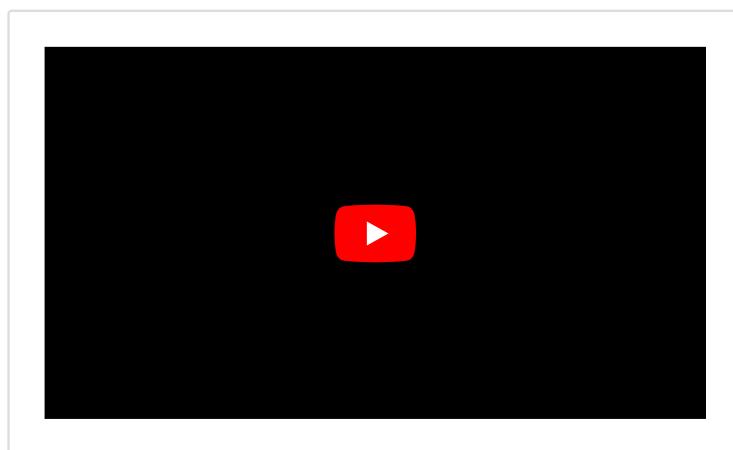
Una vez decidido el modelo de regresión que vamos a emplear, trabajaremos en un **proceso iterativo de ajuste del modelo y creación de nuevos indicadores**. Muy importante en nuestro caso capturar de forma adecuada la estacionalidad. Tendremos que aislar muy bien el efecto de la Navidad o 'días especiales'. Al igual que si estuviéramos modelizando helados, debemos dar relevancia al estudio de estacionalidad para no llegar a resultados erróneos.

## 8

## Evaluación

En nuestro caso, emplearemos las métricas habituales de **evaluación del error** de un modelo de regresión con series temporales (R<sup>2</sup>, RMSE, MAPE, etc.).

A continuación, un vídeo que resume cómo calcular el RMSE (*root-mean-square error*) sobre series temporales:



## 9

## Aplicación

Una vez desarrollado el modelo, nos toca aplicarlo. El objetivo principal que nos marcaremos en esta fase será la **optimización de la estrategia de marketing**. En nuestro caso de perfumes, queremos responder a este tipo de preguntas:

- ¿Cuánto debo invertir en cada una de las fragancias analizadas?
- ¿Es adecuado el mix de medios, es decir, inversión destinada a cada medio publicitario (TV, online, exterior, etc.)?
- ¿Cuál debe ser la estrategia de inversión en periodos estacionales positivos (Navidad, Día de la Madre, Día del Padre, Día de San Valentín o Black Friday)?
- ¿Es mejor comenzar a invertir dos semanas antes de los días especiales o únicamente en los días previos?
- ¿Cómo afectan los lanzamientos de nuevas referencias en las ventas y cómo se deben apoyar con publicidad?
- ¿Es el precio una variable muy relevante para la marca o afecta más la intensidad promocional?
- ¿Cómo afecta la publicidad que hacen los perfumes de la competencia?

Cada compañía o anunciante adaptará las fases a su metodología de trabajo. Por ejemplo, Nielsen presenta las fases de su modelo de trabajo como sigue:



Fuente: <https://marketingeffectiveness.nielsen.com>.

Lesson 2 of 5

# Transformación de datos

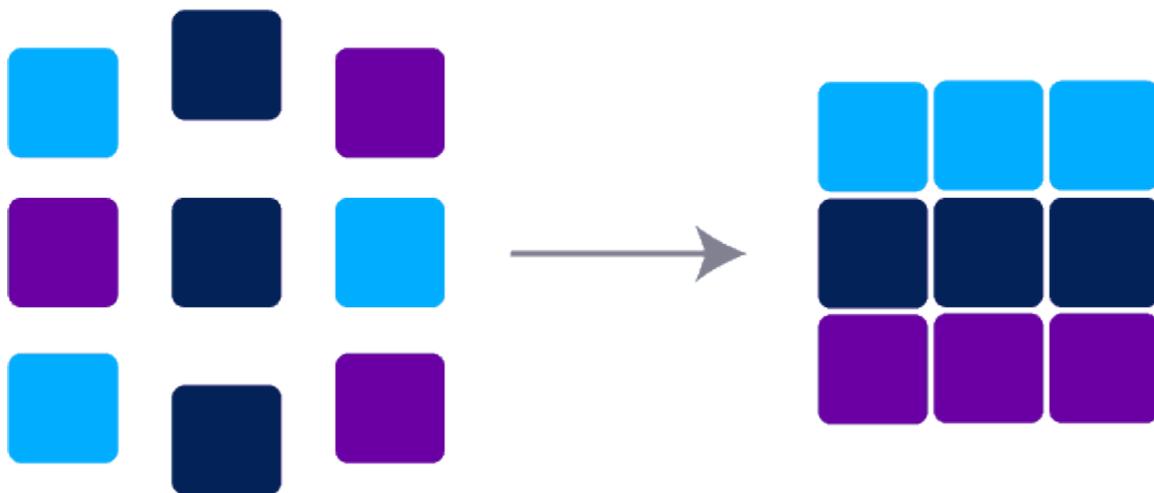
X Edix Educación

---

Una fase clave de nuestro trabajo será la transformación de datos y, en el caso del MMM, hay varias particularidades a tener en cuenta.

En esta fase de transformación de datos nos encontraremos distintos problemas con los datos que debemos resolver: cada fuente de datos puede tener una granularidad temporal distinta (datos diarios de visitas a la web, datos semanales de inversión en medios, datos mensuales de tracking, datos trimestrales macroeconómicos...), variables sin información para todo el histórico, etc. Además, necesitaremos crear nuevas variables a partir de las originales.

## Transformación de datos



---

**El fin de la transformación de datos es unificar toda la información disponible para la modelización en una sola tabla (base de datos).**

---

Nuestra base de datos, generalmente mensual, semanal o diaria, tendrá una fila por cada periodo y una columna por cada variable. Salvo la variable fecha, todas las columnas deberán tener **formato numérico**.

Como en cualquier otro caso de data mining, para poder unificar dos tablas, necesitamos que tengan, al menos, una variable en común. En nuestro caso, **la variable en común** que van a tener las distintas tablas a unificar es **la fecha**.

En la mayor parte de los casos, nos tocará cambiar la granularidad de la información. Si queremos construir un modelo de regresión de ventas semanales, necesitaremos que toda la información esté con esa periodicidad.

La granularidad de las distintas fuentes de datos que debemos integrar en la base de datos unificada podría ser diferente. Habrá desde **series diarias**, como las visitas a la web, hasta series trimestrales, como el PIB, pasando por **series semanales**, como la inversión publicitaria, o **series mensuales**, como la notoriedad de marca.

Dependiendo de la granularidad de origen del dato y de la de destino al que queramos llevarlo, el tratamiento es distinto. También el tratamiento depende de la naturaleza de la variable que queramos transformar (no se aplica el mismo procedimiento a un indicador que a un volumen).

## De menor a mayor granularidad: interpolar y repartir

Si tenemos que **pasar un dato de menor granularidad** (por ejemplo, mensual) **a uno de mayor granularidad** (por ejemplo, diario) aplicamos las dos siguientes transformaciones, dependiendo de la naturaleza de la variable a transformar:

### Interpolar



Si la variable es un indicador (por ejemplo, la tasa de paro), que habrá pasado por todos los puntos intermedios en las observaciones que no tenemos, la función de interpolación que utilizaremos normalmente es la lineal. Se desaconseja el uso de otras interpolaciones como la spline cúbica, ya que nuevos datos de la serie producen cambios en la estimación del pasado.

### Repartir



Si la variable es un volumen divisible (por ejemplo, inversión). Si no tenemos mejor información a priori, la transformación se hará de manera uniforme en todo el intervalo. Por ejemplo, si tenemos que pasar una inversión mensual a semanal, la repartimos uniformemente entre las semanas que formen el mes. Dicho de otro modo, si sabemos que la inversión en un mes ha sido de 100.000 euros, dividiremos entre los días que forman el mes para, posteriormente, sumar los días que forman cada semana.

## De mayor a menor granularidad: promediar y sumar

Si tenemos que **pasar un dato de mayor granularidad** (por ejemplo, diario) a **uno de menor granularidad** (por ejemplo, semanal) aplicaremos normalmente las dos siguientes transformaciones, dependiendo de la naturaleza de la variable a transformar:

### Promediar

Si la variable es un indicador (por ejemplo, precio), el promedio lo vamos a realizar si no tenemos mejor información *a priori* sin ponderar. Por ejemplo, si tenemos que pasar un precio medio diario a uno semanal, lo promediamos.

### Sumar

Si la variable es un volumen que se puede sumar (por ejemplo, visitas web).

---

**En definitiva: aquí lo que queremos construir son series temporales que tengan la misma granularidad. Así, podremos analizar la información (análisis descriptivo) y, posteriormente, realizar la medición (modelización).**

# Variables a modelizar

X Edix Educación

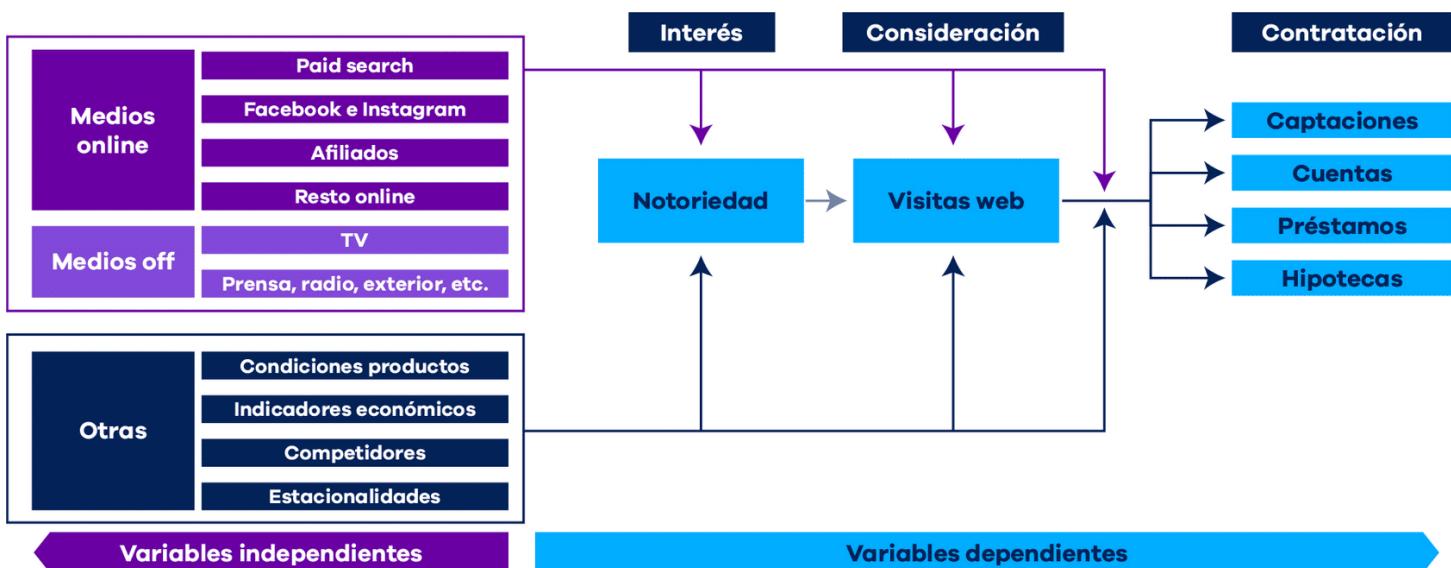
---

Cada anunciante deberá adaptar la medición a su negocio. Por tanto, **las variables a modelizar (variables dependientes) variarán en función del sector y del anunciante**. Podemos resumir qué tipo de variables de negocio son las más habituales a modelizar para algunos ejemplos de sectores:

- **Retail:** visitas tienda, visitas web, ventas en tienda y ventas web.
- **Hospital:** reservas y volumen de ingresos por región.
- **Banca:** captaciones y altas por producto financiero.
- **Seguros:** cotizaciones y pólizas.
- **Automoción:** visitas al concesionario, ventas y matriculaciones.
- **Gran consumo:** ventas en volumen.
- **Ecommerce:** descargas de la aplicación y altas de nuevos usuarios.

Los anunciantes están interesados en métricas que recojan el conocimiento que se tiene de la marca como son la notoriedad, número de búsquedas en internet o visitas web.

Para cada anunciante debemos diseñar cuál va a ser el esquema de modelización y cómo van a estar conectadas las variables dependientes e independientes. A continuación, se muestra un ejemplo de esquema de modelización para el sector financiero:



En resumen:

- La publicidad online y offline, junto con otros indicadores, afectan a la notoriedad, es decir, al número de personas que conocen la entidad financiera.
- Esta notoriedad hace incrementar las visitas a la web que se ven también influenciadas por las variables financieras.
- Nos encontramos con las variables relativas a la contratación (captaciones, altas de cuentas, altas de préstamos, altas de hipotecas, etc.) que, a su vez, están afectadas por las variables previas de visitas web y notoriedad (del tracking), así como por la publicidad y otros indicadores.

Las variables de negocio (ventas, captaciones,...) estarán, habitualmente, almacenadas en los sistemas de la compañía. Lo importante será garantizar que la información es de buena calidad, pero no debería suponer ningún problema adicional. En cambio, las variables de tracking proceden habitualmente de proveedores externos. De igual manera, las métricas web dependerán del sistema de analítica web disponible y será necesario chequear que son correctas.

Lesson 4 of 5

# Variable de tracking

X Edix Educación

---

En general, los anunciantes cuentan con estudios sobre la evolución de intangibles. Institutos de investigación como Millward Brown o GFK les ayudan a llevar el tracking (seguimiento) de su marca y compararla frente a los principales competidores. A través de encuestas, las compañías pueden saber cómo evolucionan en notoriedad e imagen de marca frente la competencia y analizar cómo afecta sobre el negocio.

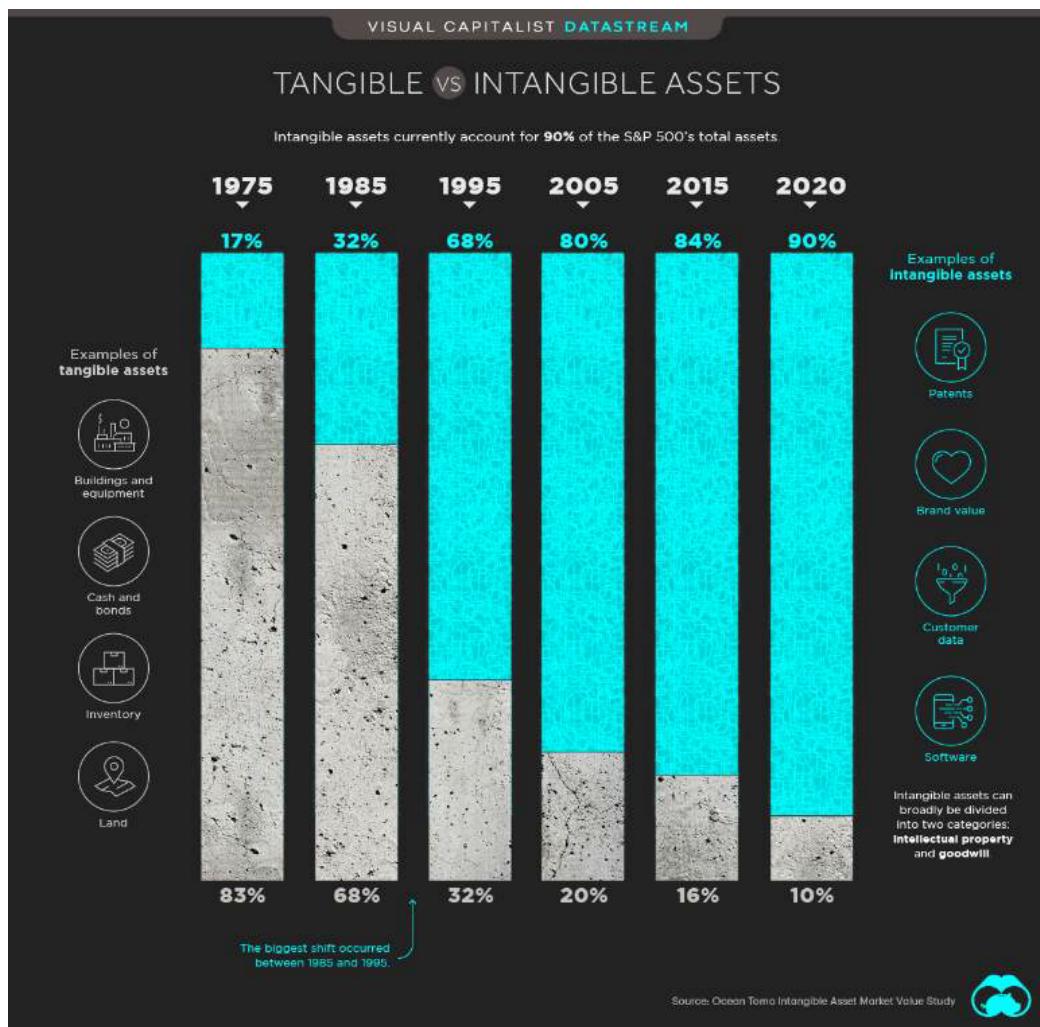


Accede a [este enlace](#) si quieras saber más sobre en qué consiste el tracking.

**¿Por qué es importante medir el valor intangible?**

Porque los intangibles representan cada vez más para el negocio de las compañías.

Según Ocean Tomo, banco comercial de propiedad intelectual que ofrece productos y servicios financieros, el valor de los activos intangibles para las compañías que forman el índice S&P 500 en 2020 ya supone el 90% de la importancia de su negocio.



Fuente: Ocean Tomo.

Por tanto, llevar a cabo esta medición será fundamental para analizar cómo mejora o empeora la salud de una marca a lo largo del tiempo. Podemos resaltar algunos indicadores clave del tracking que normalmente se modelizan en los modelos de MMM:

## Notoriedad de marca

Recoge el **conocimiento que se tiene sobre una marca**. Si pensamos en el sector financiero, la pregunta a hacer a los encuestados podría ser la siguiente: “De todos los bancos y cajas de ahorros que usted conoce, aunque solo sea de oídas, por favor, dígame el nombre del primer banco que se le venga a la mente”. Además, posteriormente, se podrá preguntar por los siguientes bancos o cajas de ahorro que vienen a su mente o sugerir las entidades para que confirme si conoce cada una de ellas.

## Recuerdo publicitario

Es similar a la notoriedad de marca, pero **centrada en la publicidad**. En el caso del sector financiero, podría ser la siguiente pregunta: “Ahora nos gustaría que pensase en publicidad. Por favor, indique en el recuadro que aparece a continuación, el primer banco o caja de ahorros del cuál recuerda haber visto, oído o leído publicidad recientemente”.

## Reputación

Recoge la opinión o estima que se tiene a una compañía, es **un concepto asociado al prestigio**. El Reputation Institute emplea las siguientes preguntas para medir la reputación: “Por favor, utilice una escala del 1 al 10, donde 1 significa ‘Completamente en desacuerdo’ y 10 significa ‘Completamente de acuerdo’: (1) tiene buena reputación, (2) es una empresa sobre la que tengo buena impresión, (3) es una empresa en la que confío y (4) es una empresa que admiro y respeto”.

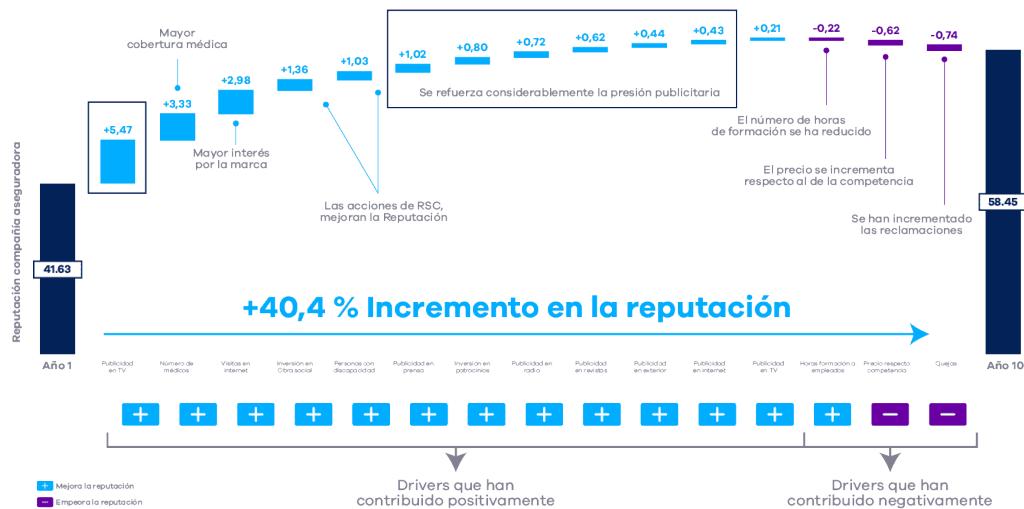
## Consideración (no clientes)

En este caso nos queremos centrar en la **intención de contratación o compra de no clientes**, es decir, **clientes potenciales**. Un ejemplo de pregunta que podríamos emplear en el sector financiero es este: “Sin considerar las entidades con las que usted trabaja actualmente, ¿en qué otras entidades le gustaría tener una cuenta u otro producto financiero? Diga todas las entidades que podría considerar para abrir una cuenta o contratar un producto financiero”.



Revisa la web y las encuestas del [Reputation Institute](#).

Analizar las variaciones de estos intangibles requerirá de un mayor histórico de datos porque no podemos esperar que la reputación o notoriedad de marca varíe en el corto plazo. En el ejemplo siguiente, se muestra un gráfico de '¿Por qué estamos cómo estamos?' Después de haber desarrollado un modelo de MMM para una serie de reputación de una compañía aseguradora utilizando diez años de histórico.



En resumen, la compañía aseguradora ha conseguido aumentar la reputación un 40,4% gracias a la estrategia seguida durante los últimos años: mayor publicidad en TV y en otro medios, mayor cobertura médica e inversión en obra social. Este gráfico se ha podido generar después de haber aplicado un modelo de regresión a la serie de reputación.

Como tenemos los coeficientes asociados a cada variable, podemos analizar la variación de las variables reales y multiplicarlas por dichos coeficientes para saber qué factores están ayudando a generar más reputación y qué factores están reduciéndola.

Lesson 5 of 5

# Métricas web

 Edix Educación

---

Todos estamos viviendo cómo el canal online va tomando cada vez más relevancia en nuestra vida y relación con las compañías. Esto ha provocado que los anunciantes no solo se centren en la modelización de variables de negocio o de tracking, sino **que sea muy relevante la modelización de métricas web**. Dependiendo del tipo de anunciante, será necesario modelizar unos KPIs web u otros.

---

**Interesa analizar las visitas web, visitas a formulario y conversiones a lo largo del histórico.**

---

Esto nos va a servir, por ejemplo, para medir la influencia que tiene la publicidad en TV y en tráfico web de la marca. Solo hace falta poner la TV y ver la cantidad de anuncios que hay de negocios puros digitales (Glovo, Idealista, Amazon,etc.), a pesar de ser compañías cuyo negocio procede exclusivamente del canal online.

## ¿Cuál es el problema más habitual al construir MMM sobre métricas web?

La medición de dichas métricas web.

A veces, las compañías cambian del proveedor de medición (Google Analytics, Adobe, etc.) o se producen cambios en la web que afectan a la navegación. Antes de modelizar alguna de estas variables, será indispensable conversar con los equipos especializados en analítica web para tratar la información de la mejor manera posible.

Adicionalmente, comienzan a surgir proveedores de métricas online que dan información para cualquier dominio. Un ejemplo es SimilarWeb, una herramienta de análisis de páginas web y aplicaciones que proporciona inteligencia de mercado de carácter global y multiplataforma para entender y monitorizar el rendimiento digital, así como analizar el comportamiento de la audiencia.



Accede a la [web](#) de SimilarWeb.

A continuación, el listado de variables que proporciona SimilarWeb para cualquier dominio web:

¿Qué datos proporciona SimilarWeb?			
<b>MÉTRICAS GLOBALES</b> A nivel total y desagregado por desktop y móvil. Dato diario, semanal, mensual. <small>*Algunos KPIs y análisis por sector solo dato mensual.</small> <ul style="list-style-type: none"> <li>Visitas web.</li> <li>Visitantes Únicos.</li> <li>Duración media de la visita.</li> <li>Páginas por visita.</li> <li>Tasa de rebote.</li> </ul>	<b>TRÁFICO REFERAL</b> A nivel total y desagregado por desktop y móvil. Dato mensual o total periodo.	<b>ANÁLISIS DE CONVERSIÓN</b> Dato mensual o total periodo. <small>*Dato para número de dominios limitado (no incluye aseguradoras por el momento)</small> <ul style="list-style-type: none"> <li>Dominio.</li> <li>Categoría.</li> <li>Cuota por dominio.</li> </ul>	<b>ANÁLISIS POR SEGMENTOS</b> Dato diario, semanal, mensual. <small>*Algunos KPIs y análisis por sector solo dato mensual</small> <ul style="list-style-type: none"> <li>Visitas web.</li> <li>Visitantes únicos.</li> <li>Duración media de la visita.</li> <li>Páginas por visita.</li> <li>Tasa de rebote.</li> </ul>
<b>VISITAS POR CANAL</b> Dato diario, semanal, mensual. <small>*Algunos KPIs y análisis por sector solo dato mensual.</small> <ul style="list-style-type: none"> <li>Visitas.</li> <li>Duración.</li> <li>Páginas.</li> <li>Tasa de rebote.</li> <li>Fuente (solo mensual).</li> </ul>	<b>TRÁFICO DISPLAY</b> A nivel total y desagregado por desktop y móvil. Dato mensual o total periodo.	<b>PERFIL SOCIODEMOGRÁFICO</b> A nivel agregado.	<ul style="list-style-type: none"> <li>Edad.</li> <li>Género.</li> <li>País.</li> </ul>
<b>ANÁLISIS DE KEYWORDS</b> Dato semanal y mensual.	<b>TRÁFICO REDES SOCIALES</b> A nivel total y desagregado por desktop y móvil. Dato mensual o total periodo.	<ul style="list-style-type: none"> <li>Término.</li> <li>Total de visitas search.</li> <li>Traffic orgánico vs pagado.</li> <li>Popularidad.</li> <li>CPC.</li> </ul>	

Al disponer de información sobre cualquier dominio web, podemos enriquecer el MMM, modelizando también las métricas web de la competencia.

En resumen:

- Las variables de negocio (ventas, contrataciones, etc.), las variables de tracking (notoriedad, reputación, etc.) y las variables web (visitas web, conversiones, etc.) serán buenas candidatas para ser modelizadas.
- A través de MMM, mediremos cuáles son las variables que tienen más influencia sobre ellas y obtendremos insights relevantes para la compañía.

Y para terminar este fastbook, un consejo centrado en las fases y variables de un proyecto de MMM:

“Disponer de información de calidad será un punto clave al desarrollar modelos de medición del ROI. En algunos casos, será necesario esperar un tiempo y recopilar la información que nutrirá los modelos. Sin duda, los datos son como el petróleo para muchos productos cotidianos, es decir, solo son útiles cuando se han extraído, refinado y puesto a disposición del equipo que vaya a modelizar. Seguir un plan ordenado de trabajo será el otro requisito que permitirá lograr el éxito en proyectos de MMM.”

**¡Enhорabuena! Fastbook superado**

edix

Creamos Digital Workers