

# **Proyecto Integrador**

## **Módulo 5: Data Visualization**

### **Estrategias de Marketing: Análisis de Campañas y Respuestas del Cliente**

#### **Sector: Business Intelligence y Data Analytics**

##### **Integrantes:**

GARCÍA PÉREZ GIANNY RONALDO  
HEREDIA VILLAMAR KIMBERLY ELIZABETH  
MURILLO LOOR RAUL ABRAHAM  
VELIZ RIVERA FELIX JAIR  
VELIZ CEVALLOS LUISA FERNANDA

##### **Curso:**

Paralelo - 04

**Coding Bootcamps - MINTEL**

**Programa Data-Driven-Decision Specialist**

##### **Profesor:**

MSc. Martha Tomalá

**24 de enero del 2026**

1. Introducción y objetivo del análisis.....	3
Introducción.....	3
Objetivo del análisis.....	3
2. Descripción del dataset y alcance del análisis.....	4
3. Metodología de preparación de datos.....	4
3.1 Limpieza y validaciones aplicadas.....	4
3.2 Tratamiento de inconsistencias.....	5
3.3 Reconstrucción y creación de variables.....	5
4. Análisis exploratorio: hallazgos clave.....	6
4.1 Hallazgos principales.....	6
5. Preguntas de negocio.....	6
6. Definición de KPIs y métricas.....	7
Para responder las preguntas planteadas, se definieron los siguientes indicadores clave de desempeño (KPIs):.....	7
6.1 Tasa de aceptación de campañas.....	7
6.2 Número de clientes que aceptan campañas.....	7
6.3 Recurrencia promedio de aceptación de campañas.....	8
6.4 Tasa de aceptación por canal de compra.....	8
6.5 Ingresos totales por campaña.....	8
7. Visualizaciones y justificación de selección.....	8
Insights clave.....	9
Recomendaciones.....	9

# 1. Introducción y objetivo del análisis.

## Introducción

En el contexto actual del marketing basado en datos, resulta fundamental comprender no solo el comportamiento de compra de los clientes, sino también su respuesta ante distintas campañas de marketing. Un análisis inadecuado o basado en datos inconsistentes puede conducir a decisiones erróneas en la asignación de recursos, selección de canales y segmentación de clientes.

El presente estudio se apoya en un proceso riguroso de preparación y validación de datos, con el fin de garantizar que los resultados obtenidos sean confiables y útiles para la toma de decisiones estratégicas. A partir de un dataset de campañas de marketing y comportamiento del cliente, se construye una base sólida que permite evaluar la efectividad de las campañas, identificar perfiles con mayor tasa de aceptación y determinar qué acciones generan mayor impacto económico para la empresa.

## Objetivo del análisis

El objetivo del análisis es evaluar la efectividad de las campañas de marketing y el comportamiento de respuesta de los clientes, mediante un dataset depurado y validado, con el propósito de:

- Identificar el porcentaje de clientes que aceptan campañas de marketing.
- Determinar los perfiles de clientes con mayor tasa y recurrencia de aceptación.
- Analizar los canales de compra más efectivos.
- Identificar las campañas que generan mayor ingreso total para la empresa.

Los resultados del análisis están orientados a apoyar la toma de decisiones estratégicas en la planificación, segmentación y optimización de futuras campañas de marketing.

## 2. Descripción del dataset y alcance del análisis.

El análisis se realizó sobre el dataset `marketing_raw.csv`, el cual contiene información relacionada con características demográficas de los clientes, comportamiento de compra, interacción con campañas de marketing y montos de gasto por categoría de producto.

En su estado original, el dataset estaba compuesto por **39 columnas y 2205 registros**, incluyendo variables numéricas, categóricas y binarias. Tras el proceso de limpieza y validación, se obtuvo un dataset final de **2021 registros únicos**, representando clientes individuales sin duplicados.

El alcance del análisis se limita a la evaluación descriptiva y analítica del comportamiento de los clientes frente a campañas de marketing. No se incluyen modelos predictivos, sino que se prioriza la obtención de **insights accionables** mediante indicadores clave de desempeño (KPIs) y visualizaciones en Power BI.

## 3. Metodología de preparación de datos

### 3.1 Limpieza y validaciones aplicadas

Con base en la exploración inicial del dataset, se identificaron diversos problemas de calidad que podrían afectar la confiabilidad del análisis. Para corregirlos, se aplicaron las siguientes acciones:

- Eliminación de **184 registros duplicados**, reduciendo el dataset a **2021 clientes únicos**.
- Conversión de columnas monetarias que contenían símbolos de moneda a formato numérico.
- Validación de rangos lógicos en la variable edad, definiendo un intervalo entre **18 y 100 años**.
- Identificación de valores negativos en columnas de gasto, los cuales fueron reemplazados por **0**, bajo el criterio de que representan ausencia de consumo.

Estas acciones permitieron garantizar la consistencia básica del dataset y evitar sesgos en los análisis posteriores.

### 3.2 Tratamiento de inconsistencias

Durante el proceso de validación, se detectó una inconsistencia significativa entre la variable **MntTotal** (gasto total declarado) y la suma de los gastos por categoría de producto. Se observó que **1902 registros** presentaban diferencias mayores a una unidad monetaria entre ambos valores.

Dado que no se cuenta con información suficiente sobre el origen o la metodología de cálculo de la variable **MntTotal**, Dado que no se dispone de información suficiente sobre el origen o la metodología de cálculo de la variable **Mnt Total**, se tomó la decisión de no corregir automáticamente estos valores, con el fin de preservar la integridad de los datos originales. Como criterio analítico, los cálculos relacionados con ingresos se basan en los montos por categoría de producto, los cuales presentan coherencia interna y consistencia en todo el dataset.

Como criterio analítico, los cálculos relacionados con ingresos y comportamiento de gasto se basan en los **montos por categoría de producto**, ya que estos presentan coherencia interna y permiten un análisis más confiable.

### 3.3 Reconstrucción y creación de variables

Con el objetivo de facilitar el análisis y reducir la complejidad del dataset, se realizó la reconstrucción de variables categóricas originalmente representadas mediante múltiples columnas binarias. En particular:

- Se consolidaron las variables relacionadas con **estado civil** en la columna **Marital\_Status**.
- Se consolidaron las variables relacionadas con **nivel educativo** en la columna **Education**.

Adicionalmente, se crearon variables derivadas que enriquecen el análisis, tales como:

- **Total\_Children**: número total de hijos en el hogar. (calculado como **Kidhome** + **Teenhome**).
- **Income\_Category**: segmentación de clientes por nivel de ingreso.
- **Spending\_Category**: clasificación de clientes según nivel de gasto.
- **Total\_Purchases**: total de compras realizadas en todos los canales.
- **Deals\_Percentage**: porcentaje de compras realizadas con descuento.

- **MntCalculado**: suma de gastos por categoría de producto.
- **Diferencia\_Mnt**: diferencia entre **MntTotal** y **MntCalculado**.

Estas variables permiten realizar segmentaciones más claras y análisis orientados al negocio.

## 4. Análisis exploratorio: hallazgos clave.

El análisis exploratorio se realizó sobre el dataset final depurado, con el objetivo de identificar patrones relevantes en el comportamiento de los clientes y su respuesta ante las campañas de marketing. A diferencia de una exploración descriptiva básica, este análisis se enfoca en extraer **hallazgos accionables** que permitan responder las preguntas de negocio planteadas.

### 4.1 Hallazgos principales

- **Baja tasa general de aceptación**: Se observa que una proporción reducida del total de clientes acepta campañas (evidenciado por valores bajos en **AcceptedCmpOverall**), lo que evidencia que la respuesta positiva no es generalizada y depende de características específicas del cliente.
- **Diferencias por perfil demográfico**: Existen diferencias claras en la tasa de aceptación según el perfil del cliente, particularmente al segmentar por:
  - **Nivel educativo (Education)**: Clientes con **PhD** y **Master** muestran mayor propensión.
  - **Estado civil (Marital\_Status)**: Diferencias entre **Married**, **Single** y **Together**.
  - **Nivel de ingreso (Income\_Category)**: Clientes de ingresos **High** responden más frecuentemente.
- **Relación entre fidelización y respuesta**: Los clientes con mayor volumen de compras históricas (**Total\_Purchases**) y mayor gasto acumulado (**MntTotal**, **Spending\_Category = "High"**) tienden a mostrar una mayor probabilidad de aceptar campañas, lo que sugiere una relación directa entre fidelización y efectividad de las campañas.
- **Importancia de los descuentos**: Las compras realizadas con descuento representan una proporción relevante dentro del total de transacciones (**Deals\_Percentage**), lo que sugiere que este factor puede estar asociado a la respuesta positiva en determinados segmentos de clientes.
- **Efectividad diferenciada por canal**: No todos los canales de compra presentan el mismo nivel de efectividad; algunos concentran una mayor tasa de aceptación:
  - **NumCatalogPurchases** muestra correlación positiva con aceptación de campañas.

- **NumWebVisitsMonth** puede indicar engagement previo.

Estos hallazgos justifican la necesidad de analizar el comportamiento del cliente de manera segmentada y no asumir una respuesta homogénea ante las campañas.

## 5. Preguntas de negocio

Con base en los objetivos del análisis, se plantearon las siguientes preguntas de negocio:

1. ¿Cuál es el porcentaje de clientes que aceptan al menos una campaña de marketing?
2. ¿Qué perfil de cliente presenta mayor tasa de respuesta positiva a las campañas?
3. ¿Qué segmentos de clientes muestran mayor recurrencia en la aceptación de campañas?
4. ¿Qué canal de compra está asociado con una mayor respuesta positiva a la última campaña?
5. ¿Qué campaña genera el mayor ingreso total para la empresa?

Estas preguntas orientan el análisis hacia la evaluación de la efectividad de las campañas y la identificación de oportunidades de mejora en la estrategia de marketing.

## 6. Definición de KPIs y métricas

Para responder las preguntas planteadas, se definieron los siguientes indicadores clave de desempeño (KPIs):

### 6.1 Tasa de aceptación de campañas

**Definición:**

Porcentaje de clientes que aceptaron al menos una campaña de marketing respecto al total de clientes.

**Cálculo:**

Clientes con aceptación  $\geq 1$  / Total de clientes

**Uso:**

Medir la efectividad general de las campañas y evaluar el alcance real de las acciones de marketing.

### 6.2 Número de clientes que aceptan campañas

**Definición:**

Cantidad absoluta de clientes que respondieron positivamente a al menos una campaña.

**Uso:**

Complementa la tasa de aceptación y permite dimensionar el impacto real de las campañas en términos de volumen de clientes.

### 6.3 Recurrencia promedio de aceptación de campañas

**Definición:**

Promedio de campañas aceptadas por cliente. Según los datos disponibles, este valor es de 0.46, lo que indica una baja recurrencia general.

**Uso:**

Identificar clientes recurrentes y evaluar el nivel de fidelización generado por las campañas. Este indicador se analiza de forma segmentada por categoría de ingreso, permitiendo identificar patrones de fidelización asociados al comportamiento de aceptación de campañas. Los datos confirman que los clientes de mayor ingreso (High) muestran una mayor recurrencia de aceptación.

### 6.4 Tasa de aceptación por canal de compra

**Definición:**

Porcentaje de aceptación de campañas segmentado por canal de compra (web, tienda, catálogo, etc.). Los datos muestran una tasa del 19.33% para Catálogo, 15.63% para Store (tienda física) y 0.2% para Web.

**Uso:**

Determinar qué canales presentan mayor efectividad (Catálogo) e identificar problemas críticos de desempeño (canal Web con 0.2%) para optimizar la asignación de recursos, priorizar los canales efectivos y corregir las deficiencias en futuras campañas.

### 6.5 Ingresos totales por campaña

**Definición:**

Monto total de ingresos (en millones) asociados a cada campaña de marketing, clasificadas de mayor a menor impacto económico. Los datos disponibles muestran los ingresos totales generados por cada campaña específica (Cmp 1 a Cmp 6).

**Uso:**

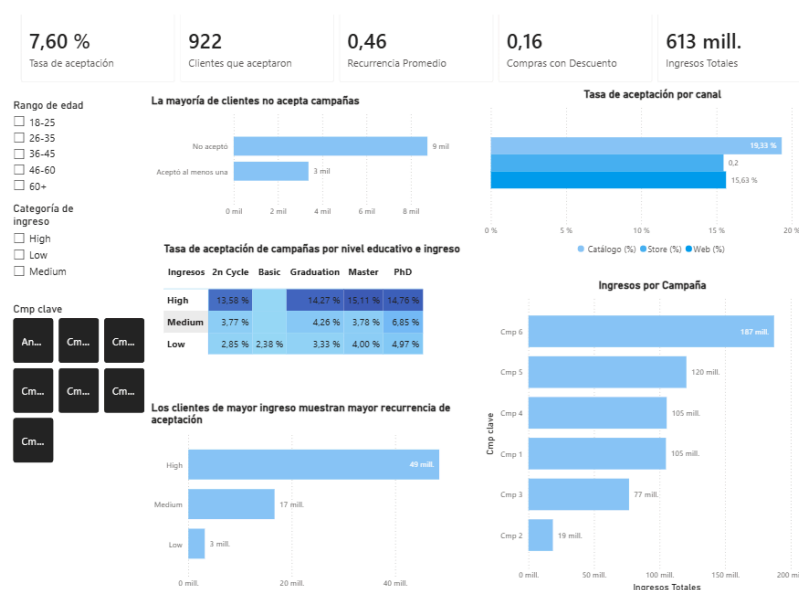
Identificar las campañas con mayor impacto económico (Cmp 6 con 187 millones) y evaluar el retorno de la inversión en marketing. Permite priorizar recursos en las campañas más rentables y analizar el desempeño financiero de cada iniciativa de marketing.



## 7. Visualizaciones y justificación de selección

Las visualizaciones fueron diseñadas con el objetivo de facilitar la interpretación de los KPIs y comunicar los resultados de forma clara y efectiva.

- **Gráficos de barras:** Utilizados para comparar visualmente los ingresos totales generados por cada campaña (Cmp 1 a Cmp 6). Este formato permite identificar de un vistazo la campaña más rentable (Cmp 6 con 187 millones) y comparar el impacto económico relativo de cada iniciativa, apoyando la priorización de recursos en futuras estrategias.
- **Tablas y matrices de datos:** Empleadas para presentar de manera estructurada y detallada la Tasa de aceptación de campañas por nivel educativo e ingreso. Esta visualización permite un análisis cruzado preciso, identificando que los clientes con nivel educativo Graduation y PhD en el segmento de ingresos High presentan las tasas de aceptación más altas (15.11% y 14.76% respectivamente), lo cual es fundamental para la segmentación de campañas.
- **Tarjetas de KPIs (Key Performance Indicators):** Utilizadas para resumir y destacar las métricas principales en un formato claro y conciso, como la Tasa de aceptación general (7.60%), el Número de clientes que aceptaron (922), la Recurrencia promedio (0.46), las Compras con descuento (0.16) y los Ingresos totales (613 mill.). Esto permite una comprensión inmediata del desempeño global.
- **Segmentaciones (Slicers / Filtros interactivos):** Implementadas en la sección de Rango de edad y Categoría de ingreso, permiten analizar dinámicamente los resultados por perfil del cliente. Esto facilita la exploración de datos, apoyando un análisis segmentado y la toma de decisiones más granular, como focalizar esfuerzos en el segmento de High income que muestra mayor recurrencia.



## 8. Insights y recomendaciones

### Insights clave

- **Baja efectividad general con alta dependencia del segmento:** Las campañas de marketing no generan una respuesta homogénea. La tasa de aceptación general es baja (7.60%), evidenciando que la respuesta positiva se concentra en segmentos específicos, particularmente en clientes de ingreso alto (High).
- **El nivel de ingreso es el principal predictor de éxito:** La tasa de aceptación varía drásticamente según el perfil del cliente. Los datos muestran que los clientes de ingreso alto (High) tienen tasas de aceptación consistentemente superiores (entre 13.58% y 15.11%), independientemente de su nivel educativo. El nivel educativo más alto (PhD) solo muestra un impacto positivo significativo dentro del segmento de ingreso bajo (Low), donde alcanza un 4.97%.
- **Los clientes de alto ingreso son los más fidelizables:** Los clientes de la categoría High no solo aceptan más, sino que muestran una mayor recurrencia, confirmando que las campañas funcionan principalmente como una herramienta de fidelización para este segmento rentable, y no como un mecanismo efectivo de captación masiva.
- **El canal web es críticamente deficiente:** Existe una oportunidad clara y un problema grave en la efectividad por canal. Mientras el canal Catálogo logra una alta aceptación (19.33%), el canal Web tiene un desempeño mínimo (0.2%), indicando una falla crítica en la estrategia digital o en la experiencia del usuario online.
- **Desempeño económico desigual entre campañas:** No todas las campañas generan el mismo impacto. El análisis de ingresos revela una gran disparidad: Cmp 6 genera 187 millones (la más rentable), mientras Cmp 2 solo genera 19 millones. Esto destaca la necesidad de evaluar el ROI de cada iniciativa de manera individual.

### Recomendaciones estratégicas

- **Segmentación prioritaria:** Enfocar los esfuerzos y la inversión en marketing en el segmento de clientes High income, ya que demuestran ser los más receptivos y recurrentes. Diseñar campañas premium exclusivas para este grupo.

- **Revisión y rediseño del canal digital:** Investigar urgentemente las causas del fracaso del canal Web (0.2%). Se debe auditar la usabilidad, el diseño de las ofertas y la integración del proceso de aceptación de campañas en la plataforma online.
- **Optimización del mix de campañas:** Asignar más recursos a las campañas de alto rendimiento (Cmp 6, Cmp 5, Cmp 4) y reevaluar o rediseñar las de bajo impacto (Cmp 2, Cmp 3), analizando qué elementos de las exitosas pueden replicarse.
- **Potenciar el canal más efectivo:** Aumentar la inversión y la frecuencia en el \*\*canal Catálogo\*\*, que presenta la tasa de aceptación más alta (19.33%), posiblemente integrando ofertas personalizadas basadas en el historial de compras.
- **Estrategia diferenciada para ingresos bajos (Low):** Para el segmento Low, desarrollar campañas específicas que resalten el valor y el ahorro inmediato, utilizando el nivel educativo (PhD) como un indicador de mayor probabilidad de aceptación dentro de este grupo.