

# Informe de Análisis de Datos de Campaña de Marketing

## Índice de Contenidos

1. Resumen Ejecutivo
2. Introducción
3. Metodología
4. Resultados y Hallazgos
  - 4.1. Análisis Descriptivo de los Datos
  - 4.2. Distribuciones de Variables Clave
  - 4.3. Análisis de Correlación
  - 4.4. Análisis de Gasto por Nivel Educativo y Estado Civil
  - 4.5. Relación entre Ingresos, Edad y Gasto Total
5. Conclusiones
6. Recomendaciones
7. Apéndice Técnico

## 1. Resumen Ejecutivo

Este informe presenta un análisis exhaustivo de los datos de una campaña de marketing de un supermercado. El análisis incluye la exploración de las características demográficas de los clientes, sus patrones de compra y la relación entre diversas variables. Los resultados revelan información valiosa sobre los perfiles de clientes, sus hábitos de gasto y la efectividad de las campañas de marketing. Se identificaron patrones significativos en la distribución de edad, ingresos, gasto total y nivel educativo, así como correlaciones entre estas variables. Las conclusiones y recomendaciones derivadas de este análisis tienen como objetivo mejorar la segmentación de clientes y optimizar las estrategias de marketing para aumentar la rentabilidad.

## 2. Introducción

El presente informe tiene como objetivo analizar los datos de una campaña de marketing de un supermercado para comprender mejor el comportamiento de los clientes y optimizar las estrategias de marketing. La base de datos contiene información demográfica de los clientes, sus hábitos de compra y su respuesta a las campañas de marketing. El análisis exploratorio de datos (EDA) se realizó utilizando Python y las bibliotecas Pandas, Matplotlib y Seaborn. Este informe proporciona una visión detallada de los hallazgos y ofrece recomendaciones basadas en los resultados obtenidos.

## 3. Metodología

El análisis se llevó a cabo siguiendo los siguientes pasos:

1. Carga de Datos: Se cargó el conjunto de datos desde un archivo CSV utilizando la biblioteca Pandas.
2. Limpieza y Preprocesamiento: Se verificaron y manejaron los valores faltantes en la columna 'Income' imputándolos con la mediana. Se convirtió la columna 'Dt\_Customer' al formato de fecha y hora.
3. Ingeniería de Características: Se crearon nuevas variables como 'Age', 'Total\_Spending', 'Total\_Kids' y 'Total\_Purchases' para facilitar el análisis. Se eliminaron las columnas innecesarias 'Z\_CostContact' y 'Z\_Revenue'.
4. Análisis Exploratorio de Datos (EDA): Se realizaron visualizaciones para comprender la

distribución de las variables clave y las relaciones entre ellas. Se utilizaron histogramas, gráficos de barras, mapas de calor, diagramas de caja y diagramas de dispersión.

5. Interpretación de Resultados: Se analizaron los resultados de las visualizaciones y se identificaron patrones y tendencias significativas.

6. Conclusiones y Recomendaciones: Se derivaron conclusiones basadas en los hallazgos y se proporcionaron recomendaciones para mejorar las estrategias de marketing.

## 4. Resultados y Hallazgos

### 4.1. Análisis Descriptivo de los Datos

El conjunto de datos contiene información sobre 2240 clientes. Las variables incluyen datos demográficos como edad, nivel educativo, estado civil e ingresos, así como datos de comportamiento de compra como el gasto en diferentes categorías de productos y el número de compras realizadas a través de diferentes canales.

### 4.2. Distribuciones de Variables Clave

\* Edad: La distribución de la edad de los clientes muestra una concentración en el rango de 40 a 60 años, con una distribución relativamente normal.

Imagen no encontrada: ./imgs/age\_distribution.png

\* Ingresos: La distribución de los ingresos muestra una concentración en el rango de 50,000 a 80,000, con algunos valores atípicos de ingresos más altos.

Imagen no encontrada: ./imgs/income\_distribution.png

\* Gasto Total: La distribución del gasto total muestra una concentración en el rango de 0 a 1000, con algunos clientes que gastan significativamente más.

Imagen no encontrada: ./imgs/total\_spending\_distribution.png

\* Estado Civil: La mayoría de los clientes están casados o en pareja, seguidos por solteros.

Imagen no encontrada: ./imgs/marital\_status\_distribution.png

\* Nivel Educativo: La mayoría de los clientes tienen un nivel educativo de graduado o posgrado.

Imagen no encontrada: ./imgs/education\_level\_distribution.png

\* Total de Hijos: La mayoría de los clientes tienen 0 o 1 hijo.

Imagen no encontrada: ./imgs/total\_kids\_distribution.png

\* Total de Compras: La distribución del total de compras muestra una concentración en el rango de 5 a 15 compras.

Imagen no encontrada: ./imgs/total\_purchases\_distribution.png

### 4.3. Análisis de Correlación

El mapa de calor de la matriz de correlación muestra las siguientes correlaciones significativas:

\* Existe una correlación positiva fuerte entre el gasto total y el gasto en diferentes categorías de productos.

\* Existe una correlación positiva moderada entre el ingreso y el gasto total.

\* Existe una correlación negativa débil entre la edad y el número de compras realizadas.

Imagen no encontrada: ./imgs/correlation\_matrix.png

#### 4.4. Análisis de Gasto por Nivel Educativo y Estado Civil

\* Gasto por Nivel Educativo: Los clientes con nivel educativo de posgrado tienden a tener un gasto total ligeramente superior en comparación con otros niveles educativos.

Imagen no encontrada: ./imgs/spending\_by\_education.png

\* Gasto por Estado Civil: Los clientes casados o en pareja tienden a tener un gasto total ligeramente superior en comparación con los solteros.

Imagen no encontrada: ./imgs/spending\_by\_marital\_status.png

#### 4.5. Relación entre Ingresos, Edad y Gasto Total

\* Ingresos vs Gasto Total: Existe una correlación positiva entre los ingresos y el gasto total, lo que indica que los clientes con mayores ingresos tienden a gastar más.

Imagen no encontrada: ./imgs/income\_vs\_spending.png

\* Edad vs Gasto Total: No se observa una correlación fuerte entre la edad y el gasto total.

Imagen no encontrada: ./imgs/age\_vs\_spending.png

\* Pairplot: El pairplot muestra la relación entre las variables numéricas seleccionadas.

Imagen no encontrada: ./imgs/pairplot.png

### 5. Conclusiones

\* Los clientes del supermercado tienen una edad promedio entre 40 y 60 años, con ingresos que se concentran entre 50,000 y 80,000.

\* La mayoría de los clientes están casados o en pareja y tienen un nivel educativo de graduado o posgrado.

\* Existe una correlación positiva entre los ingresos y el gasto total, lo que sugiere que los clientes con mayores ingresos son un segmento clave para el supermercado.

\* El nivel educativo y el estado civil influyen ligeramente en el gasto total, con los clientes de posgrado y casados gastando un poco más.

\* No se observa una correlación fuerte entre la edad y el gasto total, lo que indica que el gasto no está directamente relacionado con la edad.

### 6. Recomendaciones

1. Segmentación de Clientes: Segmentar a los clientes en función de sus ingresos y nivel educativo para personalizar las campañas de marketing.

2. Campañas Dirigidas: Crear campañas de marketing dirigidas a clientes con mayores ingresos y nivel educativo de posgrado, ya que tienden a gastar más.

3. Programas de Lealtad: Implementar programas de lealtad para clientes casados o en pareja, ya que muestran un gasto ligeramente superior.

4. Análisis Continuo: Realizar un análisis continuo de los datos para identificar nuevas tendencias y patrones en el comportamiento de los clientes.

5. Optimización de Canales: Optimizar los canales de compra en función de las preferencias de los clientes, ya que se observan diferencias en el número de compras realizadas a través de diferentes canales.

### 7. Apéndice Técnico

\* Lenguaje de Programación: Python

\* Bibliotecas: Pandas, Matplotlib, Seaborn

\* Conjunto de Datos: marketing\_campaign.csv

- \* Manejo de Valores Faltantes: Imputación con la mediana en la columna 'Income'.
- \* Ingeniería de Características: Creación de nuevas variables como 'Age', 'Total\_Spending', 'Total\_Kids' y 'Total\_Purchases'.
- \* Visualizaciones: Histogramas, gráficos de barras, mapas de calor, diagramas de caja y diagramas de dispersión.