

# **CODING BOOTCAMPS ESPOL:**

## **Data-Driven Decisions Specialist**

### **Data Visualization**

#### **Grupo 4**

#### **Proyecto: Campañas y Respuestas**

#### **Entregable Final**

#### **Integrantes:**

Forero Villota Katherine Sheila

Gonzales Ochoa Jean Pierre

Mora Verdezoto Luis Enrique

Morales Lucas José Daniel

<b>Estado Inicial del Dataset.....</b>	<b>3</b>
<b>Estado Actual del Dataset con Mejoras Aplicadas.....</b>	<b>3</b>
<b>Enfoque Analítico, Preguntas de Negocio y KPIs.....</b>	<b>4</b>
Preguntas de negocio.....	4
Perfil del Cliente.....	4
<b>Definición de KPIs.....</b>	<b>5</b>
<b>Descripción de KPIs.....</b>	<b>6</b>
<b>Filtros del Dashboard.....</b>	<b>7</b>
<b>Justificación de las Visualizaciones.....</b>	<b>7</b>
Tabla de detalle del dashboard.....	10
Resumen.....	10
<b>Respuesta a las Preguntas de Negocio: Insights Relevantes.....</b>	<b>11</b>
¿Cómo cambia el gasto total de los clientes en función de su nivel educativo?.....	11
Insights relevantes.....	11
Explicación.....	11
¿De qué manera influye la constitución familiar del cliente en sus hábitos de compra?.....	11
Insights relevantes.....	11
Explicación.....	11
¿Existe una relación entre el nivel de ingresos y el volumen de gasto total del cliente?.....	12
Insights relevantes.....	12
Explicación.....	12
¿Hay una correlación entre el tiempo que lleva un cliente y la cantidad de dinero que ha gastado?.....	12
Insights relevantes.....	12
Explicación.....	12
¿Qué grupos de clientes tienen la tasa más alta de aceptación de las campañas de marketing?.....	12
Insights relevantes.....	12
Explicación.....	13

## Estado Inicial del Dataset

El dataset presentaba varios problemas de calidad que afectan directamente su uso para análisis y modelado. Las columnas monetarias estaban almacenadas como texto debido a la presencia del símbolo “\$”, lo que impedía realizar cálculos numéricos correctamente. Además, existían variables categóricas representadas mediante múltiples columnas one-hot, lo que aumentaba innecesariamente la dimensionalidad del dataset y dificultaba su interpretación. Aunque el número de registros era consistente y no había valores nulos, se detectaron valores inválidos como montos negativos en columnas monetarias y una variable binaria (AcceptedCmpOverall) con valores fuera de su dominio esperado. También se encontraron registros duplicados y columnas que no aportan valor analítico, lo que indicaba que el dataset contenía ruido, redundancia y errores lógicos que podían distorsionar cualquier análisis posterior.

## Estado Actual del Dataset con Mejoras Aplicadas

Tras el preprocesamiento, el dataset fue transformado en una versión más limpia, coherente y lista para análisis. Las columnas monetarias fueron convertidas a formato numérico, eliminando el símbolo “\$” para permitir cálculos y agregaciones correctas. Se reconstruyeron las variables categóricas Marital\_Status y Education a partir de sus columnas one-hot, concentrando la información en una sola columna por variable y reduciendo la dimensionalidad y redundancia. Se identificaron y trataron valores inválidos como montos negativos (también se dio un ejemplo de cómo tratar una edad inválida dadas las instrucciones, pero realmente en el dataset no existían valores fuera de lo normal), proponiendo imputaciones basadas en estadísticas robustas como la mediana. Asimismo, se eliminan duplicados y columnas irrelevantes que no aportan valor analítico. En conjunto, estos cambios mejoran significativamente la calidad del dataset al aumentar su consistencia, reducir el ruido, corregir errores lógicos y hacer que los datos sean más interpretables y confiables para análisis estadísticos y modelos de datos.

## Enfoque Analítico, Preguntas de Negocio y KPIs

El enfoque analítico del panel de control se enfoca en definir el perfil del cliente, combinando variables demográficas y socioeconómicas con comportamientos para detectar patrones significativos que respalden la toma de decisiones estratégicas.

El análisis tiene como objetivo entender de qué modo estas particularidades afectan al gasto y a la reacción ante las campañas de marketing, lo que posibilita segmentar a los clientes con mayor eficacia y mejorar los esfuerzos comerciales.

### Preguntas de negocio

#### *Perfil del Cliente*

Con el objetivo de examinar la conducta y las particularidades del cliente desde un enfoque completo, se sugieren cinco interrogantes estratégicas:

1. **¿De qué manera cambia el gasto total de los clientes en función de su grado educativo?**
  - **Justificación:** Actúa como un indicador del perfil socioeconómico del cliente, permitiendo determinar si las variaciones significativas en el comportamiento de gasto están vinculadas con el nivel educativo.
2. **¿De qué manera influye la constitución familiar del cliente en sus hábitos de compra?**
  - **Justificación:** El análisis de la presencia de hijos o adolescentes contribuye a comprender la manera en que las responsabilidades del hogar afectan las decisiones relativas al consumo.
3. **¿Existe alguna relación entre el nivel de ingresos y el volumen de gasto total del cliente?**
  - **Justificación:** Analizar el vínculo entre el poder adquisitivo del cliente y el gasto promedio realizado en las tiendas, permite validar si el ingreso anual es el único motor del consumo o si existen otros factores que influyen en el comportamiento de compra.
4. **¿Hay una correlación entre el tiempo que lleva un cliente y la cantidad de dinero que ha gastado?**
  - **Justificación:** Determinar si los clientes que tienen una relación más prolongada con la compañía muestran un valor acumulado más alto posibilita medir la fidelización de manera indirecta.

## 5. ¿Qué grupos de clientes tienen la tasa más alta de aceptación de las campañas de marketing?

- **Justificación:** Determinar los perfiles con mayor receptividad a las campañas hace posible que la inversión en marketing se optimice y que la eficacia de las estrategias de promoción se incremente.

### Definición de KPIs

Se han establecido indicadores principales alineados con las preguntas de negocio formuladas, con el fin de cuantificar las características y el comportamiento del perfil del cliente.

Los KPIs del dashboard permiten:

- Analizar el comportamiento de gasto promedio de los clientes.
- Evaluar la capacidad adquisitiva a través del ingreso promedio.
- Identificar el perfil demográfico predominante mediante la edad promedio.
- Analizar la relación entre antigüedad y comportamiento de consumo.
- Contextualizar los análisis a partir del tamaño total de la base de clientes.

Cada indicador ha sido diseñado para ser coherente con la narrativa visual del dashboard, permitiendo una interpretación clara y comparaciones consistentes entre distintos segmentos.

KPI	Definición Técnica	Utilidad
<b>Total de Clientes</b>	Conteo de clientes únicos	Conocer el tamaño de la base de clientes analizada.
<b>Gasto Promedio General</b>	Promedio del monto total gastado por cliente (Promedio de <i>MntTotal</i> )	Analizar el nivel de consumo promedio de los clientes.
<b>Ingreso Promedio General</b>	Promedio del ingreso de los clientes (Promedio de <i>Income</i> )	Evaluar la capacidad adquisitiva promedio de la base de clientes.
<b>Edad Promedio</b>	Promedio de la edad de los clientes	Identificar el perfil etario predominante.
<b>Antigüedad Promedio</b>	Promedio de años de antigüedad del cliente	Analizar la permanencia y lealtad del cliente.

## Descripción de KPIs

Los indicadores clave de desempeño (KPIs) presentados en el dashboard han sido definidos con el objetivo de **caracterizar el perfil del cliente** y analizar su comportamiento de consumo a partir de variables demográficas y económicas.

Estos KPIs permiten obtener una visión general del cliente promedio, facilitando la comparación entre segmentos y apoyando la toma de decisiones estratégicas.

- El **Total de Clientes** indica el tamaño de la base analizada y permite contextualizar los resultados mostrados en el dashboard.
- El **Gasto Promedio General** refleja el monto promedio que gasta un cliente, siendo un indicador clave para identificar patrones de consumo.
- El **Ingreso Promedio General** representa la capacidad adquisitiva promedio de los clientes y permite analizar su relación con el nivel de gasto.
- La **Edad Promedio** permite identificar el perfil etario dominante dentro de la base de clientes.
- Finalmente, la **Antigüedad Promedio** muestra el tiempo promedio que los clientes han permanecido vinculados a la empresa, aportando información relevante sobre la lealtad del cliente.

## Filtros del Dashboard

El filtro por **rango de edad** permite segmentar a los clientes según su grupo etario, facilitando el análisis de patrones de gasto, ingresos y conversión en distintos momentos del ciclo de vida del cliente. Este filtro es clave para identificar qué segmentos presentan mayor valor o mayor respuesta a campañas, apoyando decisiones de segmentación y personalización.

El filtro por **rango de antigüedad** permite analizar cómo el tiempo de relación del cliente con la empresa impacta en su comportamiento de compra. Su inclusión es relevante para evaluar la fidelización del cliente y detectar diferencias entre clientes nuevos y antiguos.

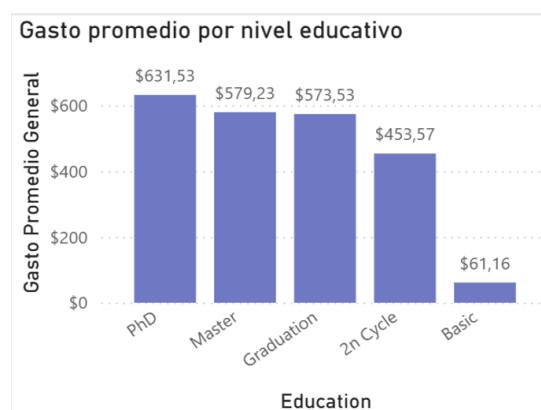
El filtro por **nivel educativo** facilita la segmentación de los clientes según su formación académica, permitiendo analizar diferencias en el gasto promedio y en el nivel de ingresos. Este filtro funciona como un proxy del nivel socioeconómico del cliente y apoya la definición de estrategias diferenciadas por perfil.

El filtro por **composición familiar (clientes con o sin hijos)** permite evaluar cómo la carga familiar influye en el consumo y en la respuesta a campañas. Su uso es relevante para diseñar ofertas y campañas ajustadas a las necesidades de cada segmento.

## Justificación de las Visualizaciones

- **Gasto Promedio General por Nivel Educativo (Gráfico de Columnas)**

Este gráfico permite analizar cómo varía el gasto promedio de los clientes según su nivel educativo. Su utilidad radica en identificar qué segmentos presentan mayor disposición al consumo, lo que apoya la segmentación de clientes y la definición de estrategias comerciales acordes a su perfil socioeconómico.



- **Gasto Promedio General según Composición Familiar (Gráfico de Anillos)**

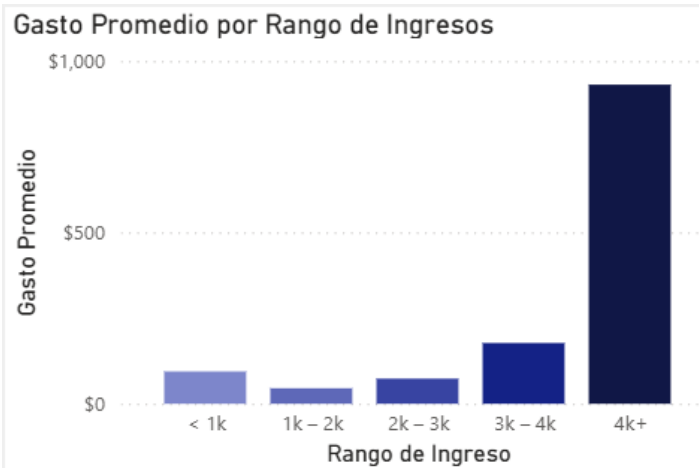
La visualización permite comparar el gasto promedio entre clientes con y sin hijos, mostrando el impacto de la carga familiar en las decisiones de consumo. Esta información es útil para adaptar la oferta de productos y servicios a las prioridades de cada segmento.

Gastos Promedio de Clientes con Hijos y sin Hijos



- **Gasto Promedio por Rango de Ingresos (Gráfico de Columnas)**

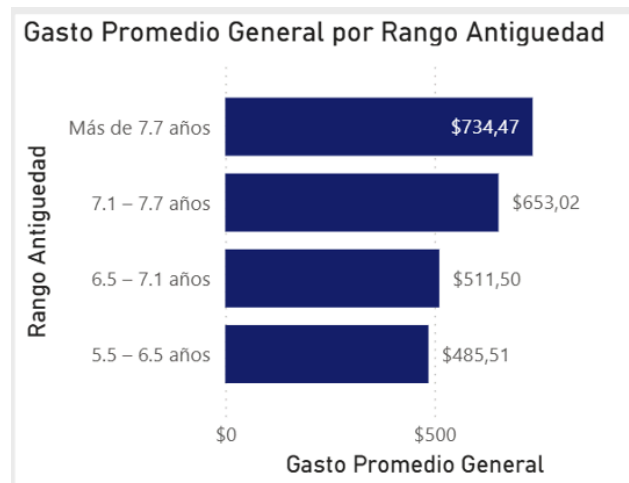
Este gráfico facilita la comparación del gasto promedio según rangos de ingresos, permitiendo analizar la relación entre **capacidad adquisitiva y nivel de consumo**.





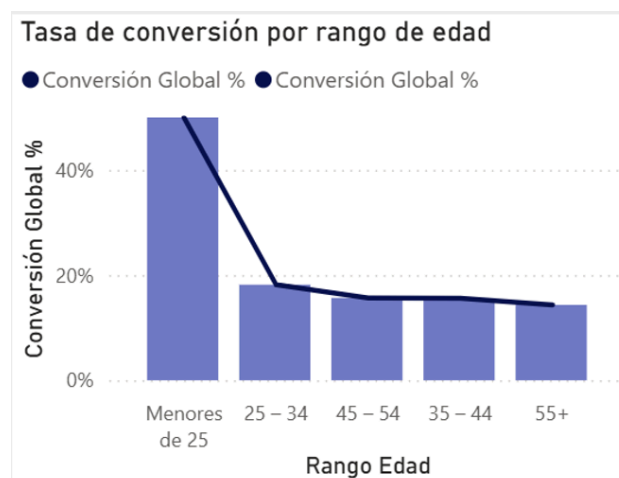
- **Gasto Promedio General por Rango de Antigüedad (Gráfico de Barras)**

Este gráfico permite analizar cómo el gasto promedio varía según el tiempo de relación del cliente con la empresa. Su análisis es clave para evaluar el impacto de la fidelización y para identificar oportunidades de retención y crecimiento del valor del cliente.



- **Conversión Global y Conversión Global (%) por Rango de Edad (Gráfico de Columnas y Línea)**

Esta visualización permite evaluar la efectividad de las campañas de marketing en distintos segmentos de edad. Su utilidad radica en identificar qué grupos presentan una mayor tasa de conversión, facilitando la identificación de perfiles que tengan respuesta a las campañas propuestas para futuras estrategias.



## Tabla de detalle del dashboard

Nivel de Educación	Estado Civil	Rango de Edad	Rango de Antigüedad	Gasto Total	Ingreso Total	Número de Compras Total
2n Cycle	Divorced	35 – 44	7.1 – 7.7 años	\$31	\$1,127.75	7
2n Cycle	Divorced	45 – 54	5.5 – 6.5 años	\$36	\$3,550.9167	9
2n Cycle	Divorced	45 – 54	6.5 – 7.1 años	\$1,070	\$5,475.5	27
2n Cycle	Divorced	45 – 54	7.1 – 7.7 años	\$978	\$7,821	23
2n Cycle	Divorced	45 – 54	Más de 7.7 años	\$1,110	\$4,093.1667	28
2n Cycle	Divorced	55+	5.5 – 6.5 años	\$1,665	\$13,730.25	56
2n Cycle	Divorced	55+	6.5 – 7.1 años	\$1,542	\$6,529.4167	24
2n Cycle	Divorced	55+	7.1 – 7.7 años	\$3,754	\$24,383.0833	94
2n Cycle	Divorced	55+	Más de 7.7 años	\$2,749	\$12,623.4167	50
2n Cycle	Married	25 – 34	5.5 – 6.5 años	\$345	\$8,332.5833	31
2n Cycle	Married	25 – 34	6.5 – 7.1 años	\$1,133	\$11,504.5833	40
2n Cycle	Married	25 – 34	7.1 – 7.7 años	\$1,005	\$9,184.6667	35
2n Cycle	Married	25 – 34	Más de 7.7 años	\$58	\$1,944.25	8
2n Cycle	Married	35 – 44	5.5 – 6.5 años	\$1,874	\$30,616.8333	83
2n Cycle	Married	35 – 44	6.5 – 7.1 años	\$1,777	\$31,069.8333	114
2n Cycle	Married	35 – 44	7.1 – 7.7 años	\$1,082	\$11,361.3333	53
2n Cycle	Married	35 – 44	Más de 7.7 años	\$35	\$1,233	6
<b>Total</b>				<b>\$1,139,418</b>	<b>\$8,704,995.8333</b>	<b>30108</b>

- La tabla complementa las visualizaciones al presentar el detalle completo de los datos, combinando múltiples variables como nivel educativo, estado civil, rango de edad y antigüedad.
- A diferencia de los gráficos, la tabla utiliza valores totales, ya que su objetivo es permitir un análisis más detallado y validar los resultados observados en las visualizaciones agregadas.

## Resumen

Preguntas de negocio	Nombre del Gráfico	Definición Técnica	Propósito del Negocio
¿De qué manera cambia el gasto de los clientes según su nivel educativo?	Gasto Promedio por Nivel de Educación	Promedio de <i>MntTotal</i> agrupado por nivel educativo	Identificar segmentos con mayor disposición al consumo según su perfil educativo.
¿Cómo influye la constitución familiar del cliente en sus hábitos de compra?	Gasto Promedio según Tenencia de Hijos	Promedio de <i>MntTotal</i> segmentado por clientes con hijos y sin hijos	Analizar el impacto de la carga familiar en el comportamiento de gasto.
¿Existe alguna relación entre el nivel de ingresos y el gasto del cliente?	Gasto Promedio por Rango de Ingresos	Promedio de <i>MntTotal</i> agrupado por rangos de ingresos	Evaluar si una mayor capacidad adquisitiva se traduce en un mayor nivel de consumo.
¿Existe una relación entre la antigüedad del cliente y su nivel de gasto?	Gasto Promedio por Rango de Antigüedad	Promedio de <i>MntTotal</i> agrupado por rangos de antigüedad	Analizar la lealtad del cliente y su impacto en el gasto promedio.
¿Qué grupos de clientes presentan mayor aceptación a las campañas de marketing?	Tasa de Conversión por Rango de Edad	(Aceptaciones / Total de clientes por rango de edad) × 100	Optimizar las campañas de marketing enfocándolas en los segmentos con mayor conversión.

## Respuesta a las Preguntas de Negocio: Insights Relevantes

### ¿Cómo cambia el gasto total de los clientes en función de su nivel educativo?

#### *Insights relevantes*

Los clientes con niveles educativos más altos, especialmente PhD y Master, presentan un gasto promedio superior en comparación con los demás niveles educativos. En contraste, los clientes con educación Básica registran el gasto promedio más bajo.

#### *Explicación*

El gráfico de Gasto Promedio por Nivel de Educación muestra una tendencia creciente del gasto conforme aumenta el nivel educativo. Esto sugiere que un mayor nivel de formación académica está asociado a una mayor capacidad adquisitiva y, por ende, a un mayor consumo. El nivel educativo actúa como un indicador indirecto del perfil socioeconómico del cliente.

### ¿De qué manera influye la constitución familiar del cliente en sus hábitos de compra?

#### *Insights relevantes*

Los clientes sin hijos presentan un mayor gasto promedio en comparación con los clientes que tienen hijos.

#### *Explicación*

El gráfico de anillos de Gasto Promedio de Clientes con Hijos y sin Hijos muestra que el segmento sin hijos concentra el mayor gasto promedio. Esto sugiere que estos clientes disponen de una mayor proporción de ingresos para consumo personal, al no tener gastos asociados a responsabilidades familiares, lo que se traduce en un mayor gasto individual.

## **¿Existe una relación entre el nivel de ingresos y el volumen de gasto total del cliente?**

### *Insights relevantes*

Existe una relación positiva entre el nivel de ingresos y el gasto promedio. Los clientes ubicados en el rango de ingresos más altos (4k+) presentan un gasto significativamente mayor que el resto de los segmentos.

### *Explicación*

El gráfico de Gasto Promedio por Rango de Ingresos muestra un incremento progresivo del gasto a medida que aumenta el ingreso, con un salto marcado en el segmento de mayores ingresos. Esto indica que la capacidad adquisitiva influye directamente en el nivel de consumo, aunque no de manera estrictamente proporcional en todos los rangos.

## **¿Hay una correlación entre el tiempo que lleva un cliente y la cantidad de dinero que ha gastado?**

### *Insights relevantes*

Los clientes con mayor antigüedad presentan un gasto promedio más alto que aquellos con menor tiempo de relación con la empresa.

### *Explicación*

El gráfico de Gasto Promedio por Rango de Antigüedad evidencia que los clientes con más de 7.7 años de antigüedad lideran el gasto promedio. Esto sugiere que la permanencia del cliente está asociada a un mayor valor acumulado, reflejando fidelidad y confianza con la empresa.

## **¿Qué grupos de clientes tienen la tasa más alta de aceptación de las campañas de marketing?**

### *Insights relevantes*

El grupo de clientes menores de 25 años presenta la mayor tasa de conversión, mientras que la conversión disminuye progresivamente en los grupos de mayor edad.

### *Explicación*

El gráfico combinado de Porcentaje de Conversión por Rango de Edad muestra un pico claro en el segmento más joven. Esto indica que los clientes jóvenes son más receptivos a las campañas de marketing, posiblemente debido a una mayor afinidad con canales digitales y promociones, aunque no necesariamente sean los de mayor gasto.