

COMPORTAMIENTO DE COMPRA DE CLIENTES

EMPRESA DE VENTAS ELECTRÓNICAS

AUTOR: AGOSTINA ALTINA

COMISIÓN: 65180

INSTITUCIÓN: CODERHOUSE

FECHA ENTREGA FINAL: 06/02/2025

ÍNDICE

Descripción de la temática de los datos.	2
Descripción de la hipótesis	3
<u>Dataset.</u>	3
Objetivo	4
Alcance	4
Usuario final y Nivel de Aplicación	4
Diagrama Entidad-Relación	5
Listado de tablas	5
Listado de las columnas de las tablas	7
Documentación del Proyecto – Transformaciones Tablas	8
Documentación del Proyecto – Medidas Generadas	12
Documentación del Proyecto – Últimas modificaciones	14
Conclusiones	18
Futuras Líneas	19

Specific (Específico):

El análisis de datos se enfoca en comprender el comportamiento de compra de los clientes de una tienda de ventas electrónicas.

Esto incluye objetivos como:

- 🖶 Analizar el perfil demográfico y de fidelidad de los clientes.
- Las categorías de productos más populares y su rendimiento.
- Los métodos de pago y envíos preferidos por los clientes.
- Los patrones en la adquisición de complementos.

Measurable (Medible):

Evaluaremos métricas claves como:

- Cantidad de productos vendidos por tipo.
- Análisis de tendencias en métodos de pago y opciones de envío.
- ♣ Porcentaje de pedidos completados vs. cancelados.
- Frecuencia de clientes leales.
- Índices como la calificación promedio por producto o grupo demográfico.

Estas métricas servirán para analizar las preferencias de los clientes y el rendimiento de los productos.

Achievable (Alcanzable):

Los datos necesarios están disponibles en el dataset, que incluye información detallada sobre las transacciones de clientes, calificaciones, y métodos de compra. La calidad y estructura del dataset permiten realizar análisis integrales para obtener buenos resultados.

Relevant (Relevante):

Este análisis es esencial para identificar patrones de compra y áreas de mejora en las estrategias de ventas. Los resultados pueden servir como base para optimizar las ofertas personalizadas; Incrementar la retención de clientes leales; Mejorar los productos y servicios mediante el análisis de calificaciones y tendencias.

Time-bound (Limitado en el tiempo):

El proyecto se llevará a cabo en tres entregas parciales, con plazos definidos para cada etapa, y concluirá con un proyecto final. La primera entrega abarcará la presentación de la temática, la formulación de la hipótesis y la selección del dataset. Las siguientes fases se centrarán en un análisis más detallado de los datos y la creación de visualizaciones, culminando con la

presentación de un tablero en Power BI y la elaboración de la documentación completa del proyecto.

DESCRIPCIÓN DE LA HIPÓTESIS

Hipótesis

Los clientes con características específicas, como membresía de lealtad, edad y género muestran patrones de compra recurrente y de mayor valor y la personalización de ofertas basada en sus preferencias mejora la frecuencia de compra y el valor total generado, impactando positivamente en los ingresos de la empresa.

Se parte de la premisa de que los clientes con mayor fidelidad tienden a generar un impacto positivo en los ingresos a través de compras recurrentes y de mayor valor. Se analiza la hipótesis de que el gasto promedio de Miembros es mayor al gasto promedio de los No Miembros.

A través del estudio del dataset, se buscará confirmar si existe una correlación entre las características de los clientes (edad, género, membresía de lealtad) y sus patrones de compra y cómo estas pueden ser utilizadas para diseñar estrategias de marketing más efectivas. Planteamos hipótesis como:

- La mayoría de los miembros son jóvenes adultos de entre 25-40 años.
- El método de pago preferido es Transferencia Bancaria
- > El Tipo de envío preferido es el "Normal"
- > El producto más vendido es "Smartphone"
- > La mayoría de las compras contienen además complementos extra adquiridos.

DATASET

El Dataset cuenta con una tabla de 10000 registros con 16 columnas.

Columnas y descripción de las mismas:

- IdCliente: Identificador único para cada cliente.
- Edad: Edad del cliente.
- Sexo: Sexo del cliente (Masculino o Femenino).
- EsMiembro: (Sí/No).
- *TipoProducto:* Tipo de producto electrónico vendido.

- IdProducto: un código único para cada producto.
- Rating: Valoración del producto por parte del cliente (de 1 a 5 estrellas).
- EstadoOrden: Estado del pedido (Completado, Cancelado).
- MetodoPago: Método utilizado para el pago (por ejemplo, Efectivo, Tarjeta de crédito, Paypal).
- PrecioTotal: Precio total de la transacción.
- PrecioUnidad: Precio por unidad del producto.
- Cantidad: Número de unidades compradas.
- FechaCompra: Fecha de la compra (formato: DD-MM-YYYY).
- TipoEnvio: Tipo de envío elegido (por ejemplo, Estándar, Al día siguiente, Exprés).
- Complementos: Lista de los artículos adicionales comprados (por ejemplo, Accesorios, Garantía ampliada).
- TotalComplementos: Precio total de los complementos adquiridos.
- *idPedido*: valor único para cada compra

El dataset será utilizado para extraer información relevante que permita evaluar la hipótesis planteada. Se analizarán los comportamientos de compra de los clientes, enfocándose en las relaciones entre las características demográficas (edad, género, membresía de lealtad) y las decisiones de compra (productos principales y complementos adquiridos). Además, se considerarán las preferencias de métodos de pago y tipos de envío, con el fin de entender cómo estos factores influyen en el gasto total y en la fidelidad del cliente, buscando así identificar patrones que permitan tomar una mejor decisión que optimice las ventas y la experiencia del cliente.

OBJETIVO

Identificar mediante el análisis detallado del comportamiento de compra de los clientes, incluyendo su perfil demográfico, patrones de consumo y preferencias de envío y pago, estrategias efectivas para aumentar la retención de clientes y maximizar los ingresos de la tienda electrónica.

ALCANCE

El proyecto abarcará:

- ♣ Datos del período de tiempo correspondiente a septiembre del año 2023 a septiembre del año 2024.
- Limpieza y normalización del dataset para establecer posteriormente relaciones entre las tablas y datos de calidad que permitan seguir con el análisis
- → Desarrollo de visualizaciones que permitan mostrar aspectos evaluados como cantidad de productos vendidos por tipo, frecuencia de miembros de la tienda, tendencias en métodos de pago y tipo de envío y calificación promedio de productos.
- ≠ Descripción en documentación de los hallazgos encontrados y de las tablas del dataset.

USUARIO FINAL Y NIVEL DE APLICACIÓN DE ANÁLISIS

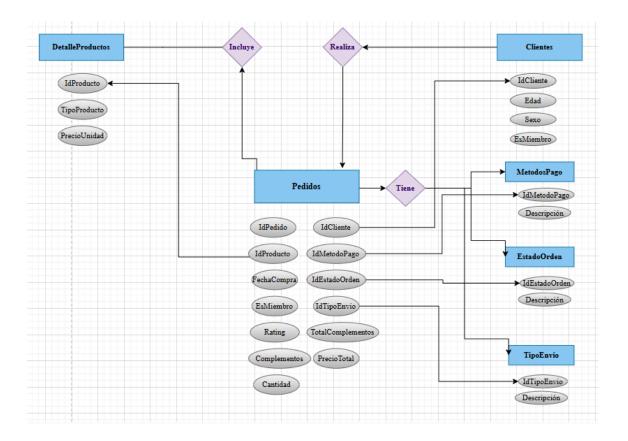
Usuario final

El tablero está diseñado para ser utilizado por gerentes de la tienda de electrónica, así como también personal de marketing y contadores o responsables de la planificación de recursos de la tienda.

Nivel de aplicación de análisis

El tablero mostrará principalmente un nivel de aplicación Táctico y Estratégico ya que proporcionara métricas que permitirán a los usuarios finales tomar decisiones que ajusten sus estrategias tanto a mediano como a largo plazo. También habrá métricas que permitan tomar decisiones a corto plazo (como el producto más vendido para asegurarse que haya stock) dando lugar al análisis Operativo en menor medida.

DIAGRAMA DE ENTIDAD-RELACIÓN



LISTADO DE LAS TABLAS

Tabla Pedidos

• En la tabla pedidos se encuentran cada uno de los pedidos realizados. Hay un pedido por producto comprado y un pedido puede tener muchos complementos agregados (pero solo un producto principal que puede comprarse en cantidad. Ej.: 6 Laptop).

Nombre de Campo	Tipo de clave	Relación
IdPedido	PK	Con la tabla Complementos por medio de IdPedido
IdProducto	FK	Con la tabla DetalleProductos por medio de IdProducto
IdCliente	FK	Con la tabla Clientes por medio de IdCliente
IdMetodoPago	FK	Con la tabla MetodosPago por medio de IdMetodoPago
IdEstadoOrden	FK	Con la tabla EstadoOrden mediante IdEstadoOrden
IdTipoEnvio	FK	Con la tabla TipoEnvio mediante IdTipoEnvio
FechaCompra		
EsMiembro		
Complementos		
TotalComplementos		
PrecioTotal		

Tabla DetalleProductos

• Esta tabla contiene el detalle de los tipos de productos y su precio unitario.

Nombre de Campo	Tipo de clave	Relación
IdProducto	PK	Con la tabla Pedidos por medio de IdProducto
TipoProducto		
PrecioUnidad		

Tabla Clientes

• Esta tabla contiene los datos exclusivamente del cliente.

Nombre de Campo	Tipo de clave	Relación
IdCliente	PK	Con la tabla Pedidos mediante IdCliente
Edad		
Sexo		
EsMiembro		

Tabla Metodos Pago

• Contiene el ID del pago y la descripción del tipo de método de pago que se usó según ese ID.

Nombre de Campo	Tipo de clave	Relación
IdMetodoPago	PK	Con la tabla Pedidos mediante IdMetodoPago
descripcion		

Tabla EstadoOrden

• Contiene la descripción asociada al ID del estado del pedido.

Nombre de Campo	Tipo de clave	e Relación	
IdEstadoOrden	PK	Con la tabla Pedidos mediante IdEstadoOrden	
descripcion			

Tabla TipoEnvio

• Contiene la descripción asociada al ID del tipo de envío que se eligió.

Nombre de Campo	Tipo de clave	Relación
IdTipoEnvio	PK	Con la tabla Pedidos mediante IdTipoEnvio
descripcion		

LISTADO DE LAS COLUMNAS DE LAS TABLAS

Tabla Pedidos

Nombre de Campo	Descripción	Tipo de dato	Pk o FK	Contiene NULL
IdPedido	Identificacion de c/pedido	VARCHAR	PK	NO
IdProducto	Identificacion de producto	VARCHAR	FK	NO
IdCliente	Identificacion de cliente	INT	FK	NO
IdMetodoPago	Identificacion del metodo de pago	INT	FK	NO
IdEstadoOrden	Identificacion del estado del pedido	INT	FK	NO
IdTipoEnvio	Identificacion del tipo de envio elegido	INT	FK	NO
FechaCompra	Fecha de cuando se realizo el pedido	DATE (dd/mm/yy)		NO
EsMiembro	Si/No es miembro de la tienda	VARCHAR		NO
Complementos	Contiene el nombre de los complementos adquiridos	VARCHAR		SI
TotalComplementos	Precio total de complementos agregados	FLOAT		NO
PrecioTotal	Precio total de los productos pedidos(sin complementos)	FLOAT		NO

Tabla DetalleProductos

Nombre de Campo	Descripción	Tipo de dato	Pk o FK	Contiene NULL
IdProducto	Identificacion de producto	INT	PK	NO
TipoProducto	Descripcion del producto asociado al ID	VARCHAR		NO
PrecioUnidad	Precio por unidad del producto	FLOAT		NO

Tabla Clientes

Nombre de Campo	Descripción	Tipo de dato	Pk o FK	Contiene NULL
IdCliente	Identificacion de cliente	INT	PK	NO
Edad	Edad del cliente registrada	INT		NO
Sexo	Sexo del cliente	VARCHAR		NO
EsMiembro	Último estado del cliente para saber si es miembro o no	INT		NO

Tabla Metodos Pago

Nombre de Campo	Descripción	Tipo de dato	Pk o FK	Contiene NULL
IdMetodoPago	Identificacion del metodo de pago	INT	PK	NO
Descripcion	Descripcion del método de pago asociado al ID	VARCHAR		NO

Tabla EstadoOrden

Nombre de Campo	Descripción	Tipo de dato	Pk o FK	Contiene NULL
IdEstadoOrden	Identificacion del estado del pedido	INT	PK	NO
Descripcion	Descripcion del estado del pedido asociado al ID	VARCHAR		NO

Tabla TipoEnvio

Nombre de Campo	Descripción	Tipo de dato	Pk o FK	Contiene NULL
IdTipoEnvio	Identificacion del tipo de envio elegido	INT	PK	NO
Descripcion	Descripcion del tipo de envio asociado al ID	VARCHAR		NO

DOCUMENTACIÓN DEL PROYECTO - TRANSFORMACIONES EN TABLAS

Se realizaron algunas transformaciones para las tablas de DetalleProductos, Pedidos y Clientes. Se añadió además una nueva tabla calculada necesaria para las medidas DAX posteriores llamada ClientesMembresia.

Tabla DetalleProductos

• Se cambió el tipo de dato de entero a decimal de la columna PrecioUnidad).

- = Table.TransformColumnTypes(#"Encabezados promovidos",{{"IdProducto", type text}, {"TipoProducto", type text}, {"PrecioUnidad", type number}})
- Se conservaron únicamente los id únicos, se eliminaron duplicados.
 - = Table.Distinct(#"Tipo cambiado", {"IdProducto"})
- Se convirtio la columna PrecioUnidad a Medida (estaba como Dimensión).

<pre>fx = Table.Distinct(#"Tipo cambiado", {"IdProducto"})</pre>							
A ^B C IdProducto	▼ A ^B _C TipoProducto	-	1.2 PrecioUnidad	v			
Válido 100 9	6 • Válido	100 %	 Válido 	100 %			
• Error 0 9	6 Error	0 %	• Error	0 %			
• Vacío 0 9	% ● Vacío	0 %	Vacío	0 %			
Distintos: 5; únicos: 5	Distintos: 2; únicos: 1		Distintos: 3; únicos:	2			
SMP234	Smartphone			79119			
TBL345	Tablet			24703			
LTP123	Laptop			46396			
SWT567	Smartwatch			84483			
HDP456	Headphones			36118			

Tabla Pedidos

- Se cambio el tipo de dato de PrecioTotal de entero a decimal.
 - = Table.TransformColumnTypes(#"Encabezados promovidos",{{"IdPedido", type text}, {"IdCliente", Int64.Type}, {"EsMiembro", type text}, {"IdProducto", type text}, {"Rating", Int64.Type}, {"IdEstadoOrden", Int64.Type}, {"IdMetodoPago", Int64.Type}, {"PrecioTotal", type number}, {"Cantidad", Int64.Type}, {"FechaCompra", type date}, {"IdTipoEnvio", Int64.Type}, {"Complementos", type text}, {"TotalComplementos", Int64.Type}})
- Se cambio de entero a decimal el TotalComplementos.
 - = Table.TransformColumnTypes(#"Tipo cambiado",{{"TotalComplementos", type number}})
- Se cambió el formato de las fechas a (DD/MM/YYYY).

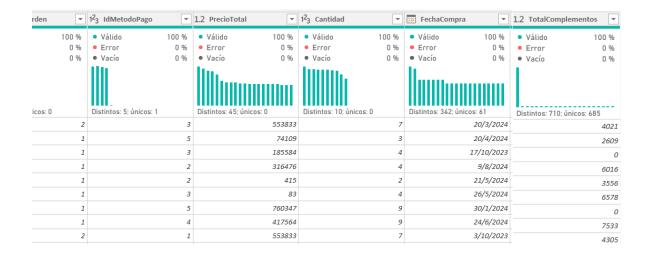


Tabla Clientes

- Se conservaron solo los valores únicos de ID de clientes, se borraron duplicados.
 - = Table.Distinct(#"Tipo cambiado", {"IdCliente"})
- Se agregó una columna calculada para saber el estado final de un cliente en cuanto a su membresía. Este atributo se agregó a la documentación de las páginas anteriores y al diagrama.

```
EsMiembro =
IF(
    CALCULATE(
        MAX(Pedidos[EsMiembro]),
        FILTER(
            Pedidos,
            Pedidos[IdCliente] = EARLIER(Clientes[IdCliente]) &&
            Pedidos[FechaCompra] =
                CALCULATE(MAX(Pedidos[FechaCompra]),
                           FILTER(Pedidos, Pedidos[IdCliente] =
EARLIER(Clientes[IdCliente]))
        )
    ) = "SI",
    1,
    0
)
```

• Se agregó una columna calculada llamada FechaPrimeraCompra para registrar posteriormente la cantidad de clientes que compraron en determinada fecha sin que se repita el mismo cliente.

FechaPrimeraCompra =

```
CALCULATE(
    MIN(Pedidos[FechaCompra]),
    FILTER(Pedidos, Pedidos[IdCliente] = Clientes[IdCliente])
)
```

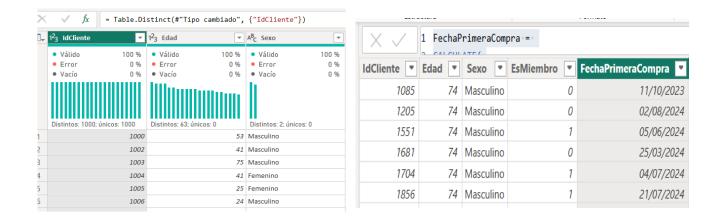


Tabla EstadoMiembro

• Se creó la tabla para guardar la descripción de los posibles estados de un cliente ya que los mismos se guardan en la tabla clientes con valores enteros de cero y uno. "Código": identificación en INTEGER del estado.

"Estado": descripción en STRING del estado.



Tabla Calendario

• Se creó una Tabla Calendario a partir del campo FechasCompra de la tabla Pedidos.

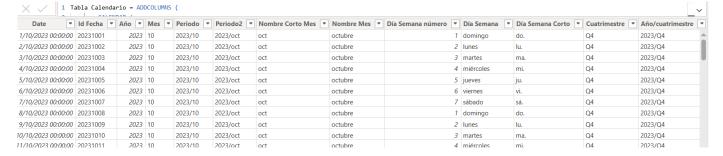
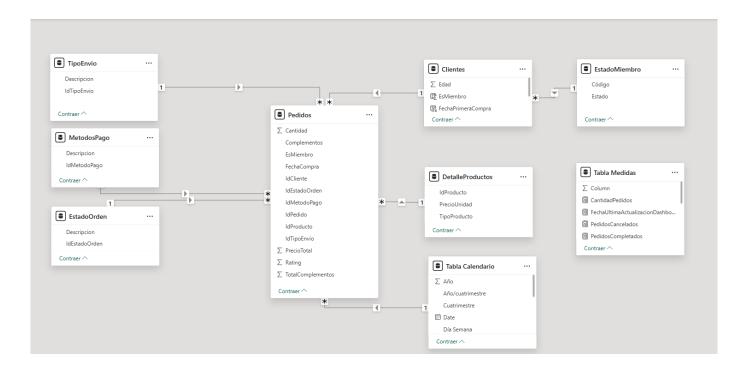


Tabla Medidas

• Se creo esta tabla para almacenar todas las medidas calculadas en ella.

Modelo Final



DOCUMENTACIÓN DEL PROYECTO - MEDIDAS CALCULADAS GENERADAS

El proyecto consta de 3 solapas además de la Portada: KPI Pedidos, KPI Clientes, Tendencias de Preferencia.

Se generaron las siguientes medidas calculadas aplicadas el día 22/01:

Cantidad de pedidos: guarda el total de pedidos creado como variable para ser reutilizado.

CantidadPedidos = VAR CantidadPedidos = COUNT(Pedidos[IdPedido]) RETURN

CantidadPedidos

Ventas por Producto: para tener un seguimiento de la cantidad de ventas por producto. VentasPorProducto = COUNT(Pedidos[IdProducto])

Rating Promedio por Producto: para saber la calificación promedio que está recibiendo cada producto.

RatingPromedioPorProducto = AVERAGEX(

```
SUMMARIZE(
Pedidos,
Pedidos[IdProducto],
"RatingPromedio", AVERAGE(Pedidos[Rating])
),
[RatingPromedio]
```

Pedidos con Complementos: para saber qué cantidad de pedidos comparado el total de pedidos, adquirieron complementos además de su compra principal.

PedidosConComplementos = CALCULATE(COUNTROWS(Pedidos),NOT ISBLANK(Pedidos[Complementos]))

Pedidos Completados: para saber la cantidad de pedidos que fueron completados exitosamente.

PedidosCompletados = CALCULATE(COUNT(Pedidos[IdPedido]),Pedidos[IdEstadoOrden]=1)

Pedidos Cancelados: para saber la cantidad de pedidos que fueron cancelados

PedidosCancelados = CALCULATE(COUNT(Pedidos[IdPedido]),Pedidos[IdEstadoOrden]=2)

Total Clientes: fue creada como variable para poder reutilizarla.

TotalClientes = VAR TotalClientes = COUNT(Clientes[IdCliente]) RETURN TotalClientes

Total Clientes Mujeres:

TotalMujeres = VAR TotalMujeres = CALCULATE(COUNT(Clientes[IdCliente]), Clientes[Sexo] = "Femenino") RETURN TotalMujeres

Total Clientes Hombres:

TotalHombres = VAR TotalHombres = CALCULATE(COUNT(Clientes[IdCliente]), Clientes[Sexo] = "Masculino") RETURN TotalHombres

Pedidos por Método de Pago: para saber la cantidad de pedidos por cada Método de Pago.

PedidosPorMetodoPago = CALCULATE(COUNTROWS(Pedidos), ALLSELECTED(Pedidos[IdMetodoPago]))

Pedidos por Tipo de Envío: para saber la cantidad de pedidos por Tipo de Envío.

PedidosPorTipoEnvio = CALCULATE(COUNTROWS(Pedidos), ALLSELECTED(Pedidos[IdTipoEnvio]))

Fecha última Actualización Dashboard: Guarda la fecha del día actual en la Portada del Dashboard.

FechaUltimaActualizacionDashboard = FORMAT(TODAY(), "dd/mm/yyyy")

DOCUMENTACIÓN DEL PROYECTO - ÚLTIMAS MODIFICACIONES

Se realizó en la fecha del 04/02/2025 modificaciones en torno a las medidas y se diseñaron las visualizaciones correspondientes.

Se diseñaron 3 solapas además de las solapas correspondientes a Portada, Glosario y Tooltip:

Portada



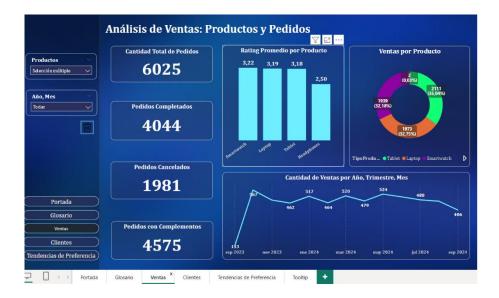
 La Portada quedó diseñada con un título, navegador de páginas, información del proyecto y la fecha de la última actualización mencionada previamente mediante el método "TODAY"

Glosario



• Se diseño esta solapa para tener una breve introducción de lo que se verá en las siguientes solapas. Cuenta también con un navegador de página y dado que no se utilizaron términos especiales no se aclaro esto mismo.

Ventas



- Esta solapa cuenta con información general importante entorno a los pedidos. Cuenta con un navegador de páginas, borrador de filtros y con filtros para las Fechas y Tipo de Productos.
- Aquí se usaron las medidas DAX de CantidadPedidos, PedidosCompletados, PedidosCancelados, PedidosConComplementos, RatingPromedioPorProducto, VentasPorProducto.
- También cuenta con una visualización asociada al gráfico de Cantidad de Ventas por Año, Trimestre y Mes mediante un tooltip que muestra el precio promedio de compra en esa fecha.

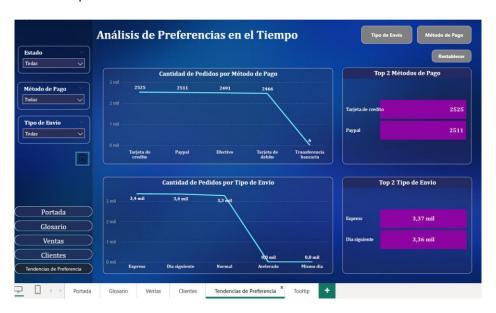
Clientes



- Esta solapa cuenta con información de los clientes así como con un navegador de páginas, borrador de filtros y filtros para el Estado del cliente (Miembro/ No miembro) y Tipo de Productos.
- Se utilizaron aquí las medidas DAX de TotalClientes, TotalHombres, TotalMujeres, FechaPrimeraCompra. Lo demás fue calculado con los atributos de las tablas.
- Se creó la medida DAX FechaPrimeraCompra para guardar la primera vez que compró un cliente pudiendo así luego en conjunto con la Fecha calcular la cantidad de clientes que hubo en el tiempo.
 FechaPrimeraCompra = CALCULATE(MIN(Pedidos[FechaCompra]), FILTER(Pedidos, Pedidos[IdCliente] = Clientes[IdCliente]))

Tendencias de Preferencia

- Esta solapa cuenta con información más detallada de los pedidos que nos muestran las preferencias de los clientes en cuanto al Método de Pago y Tipo de Envío. También cuenta con un navegador de páginas, borrador de filtros y filtros de Estado (Miembro/ No miembro), Método de Pago y Tipo de Envío. Además cuenta con 3 botones para poder visualizar únicamente Método de Pago o Tipo de Envío y para Restablecer y visualizar ambos.
- Se utilizaron aquí las medidas DAX de PedidosPorMétodoDePago y PedidosPorTipoEnvio.



Tooltip

• Finalmente hay una solapa Tooltip que se aplica a la solapa de Ventas en el gráfico de Cantidad de Ventas en el tiempo para visualizar además de la cantidad, el precio promedio de las compras de esa fecha.



CONCLUSIONES

En base a las hipótesis generadas se obtuvo las siguientes conclusiones:

El gasto promedio y la cantidad de compras de Miembros es mayor al gasto promedio de los No Miembros.

No se encontró correlación entre que un cliente sea miembro y gaste más o haya comprado más. En términos generales se encontró que los No miembro tienen un promedio de precio de compra mayor a los Miembros y se han vendido más unidades de todos los tipos a No miembros que a Miembros.

Esto puede indicar que los Miembros no tienen ninguna incentivación para marcar una diferencia con los No miembros en cuanto al valor y cantidad de sus compras.

La mayoría de los miembros son jóvenes adultos de entre 25-40 años.

La edad promedio para Miembros ronda los 49 años, similar a la de los No Miembros por lo que esta hipótesis no refleja fielmente la edad dominante. Sin embargo, habría que analizar cuanto varían las edades de la edad promedio para tener información más clara.

El método de pago preferido es Transferencia Bancaria

Incorrecto. La tarjeta de crédito es el método más usado por todos los clientes. Luego los Miembros utilizan más la Tarjeta de Débito y los No miembros la Tarjeta de Crédito.

El Tipo de envío preferido es el "Normal"

Incorrecto. El tipo de Envío más usado es el Express. Los miembros si utilizan principalmente el Normal y los No miembros Día Siguiente.

El producto más vendido es "Smartphone"

Correcto. Los Smartphones representan un 39,74% de las ventas seguidos por las Tablets con el 21,11% de las ventas.

La mayoría de las compras contienen además complementos extra adquiridos.

Correcto. Más de la mitad de las ventas contienen también elementos extra comprados, lo que sugiere mantenerlos en la tienda.

En términos generales podemos ver que no hay una correlación directa entre características específicas del cliente como membresía, edad o sexo con un patrón de adquisición de compras en particular. Podría realizarse un análisis más minucioso con nuevos datos para ver si se encuentra correlación en otros factores como ubicación, facilidad de adquisición de productos, entre otros.

FUTURAS LÍNEAS

Para complementar el análisis efectuado, se pueden abordar diferentes líneas para obtener análisis más valiosos:

- Analizar los complementos adquiridos para tener un mayor conocimiento de las preferencias entre estos y tener un stock para los más vendidos siempre.
- Realizar campañas de promociones con nuevos productos para analizar si sería conveniente invertir en nuevos productos y así ampliar los consumidores.
- Realizar promociones y/o beneficios a miembros para aumentar la lealtad entre clientes y la adquisición de una membresía, llevando también a posibles compras de mayor valor.
- Analizar la localización de las compras para analizar la conveniencia de apertura de tiendas físicas en los lugares más solicitados o ajustar las promociones a cada región para incentivar compras con publicidades más personalizadas.