

# PROYECTO FINAL

**CONJUNTO DE DATOS DE TENDENCIAS  
DE COMPRA DE CLIENTES  
(PERIODO 2019 - 2022).**

# MANGO

**Profesor:** Leonardo Oscar Martin Claro.

**Tutor:** Ari Maldonado

(arimaldo19+tutor@gmail.com )

**Dataset:** Conjunto de datos de tendencias de compra de clientes.

**SEPTIEMBRE 2024**

**Presentado por:**

**Giuliana Vera.**

# MANGO

## ÍNDICE

### CONTENIDO DEL INFORME

PORADA	1
INDICE	2
DESCRIPCÓN/OBJETIVO/HIPOTESIS	3
HERRAMIENTAS TECNOLÓGICAS IMPLEMENTADAS	4
OBJETIVO	4
ALCANCE/NIVEL APLICACION	5
ERD/ENTIDADES/ATRIBUTOS	6
DIAGRAMA ENTIDAD-RELACION	7
LISTADO DE COLUMNAS POR TABLAS	8/9
MODELO RELACIONAL POWER BI.	10
DIAGRAMA ENTIDAD-RELACION PBI	11
VISUALIZACION DE LOS DATOS	12/13/14/15/16
CONCLUSION	17
FUTURAS LINEAS	18
BIBLIOGRAFIA	19

# DESCRIPCIÓN

## 3A. Temática:

### Descripción del Conjunto de Datos "tienda EDA"

Este dataset contiene información sobre las ventas de un supermercado. Incluye detalles de los pedidos realizados por los clientes, como la fecha del pedido, la fecha de envío, el método de envío, información del cliente (ID, nombre, segmento, ubicación) y detalles del producto (ID, categoría, subcategoría, nombre del producto) junto con las ventas realizadas por cada producto.

## 3B. Objetivo general:

El objetivo general será analizar los patrones de compra de los clientes del supermercado para comprender mejor su comportamiento y optimizar las estrategias de marketing y ventas.

## 3C. Hipótesis del programa "Supermercado EDA":

- Los clientes de diferentes segmentos (por ejemplo, región) tendrán patrones de compra distintos.
- Ciertos factores como el método de envío o la región geográfica pueden influir en las ventas.
- El 60% de los clientes son hombres.
- El 80% de las compras de Colorado son realizadas por corporativos.

# HERRAMIENTAS TECNOLÓGICAS IMPLEMENTADAS

Para el presente trabajo se utilizaron los siguientes programas:

- **Excel** para la lectura y limpieza de los datasets.
- **PowerPoint y Canva** para la creación del diseño del mockup.
- **Miro** para la creación del diagrama entidad-relación.
- **Power BI Desktop** para la creación del tablero de control.

## DIAGRAMA ENTIDAD-RELACION

### Objetivo del proyecto.

- El objetivo general de este dataset abarca varias características relacionadas con las preferencias de compra de los clientes y recopila información esencial para las empresas que buscan mejorar su comprensión de su base de clientes. Las características incluyen edad, sexo, monto de la compra, métodos de pago preferidos, frecuencia de compras y calificaciones de comentarios del cliente. Además, se incluyen datos sobre el tipo de artículos comprados, frecuencia de compra, temporadas de compra preferidas e interacciones con ofertas promocionales. Con una colección de 3900 registros, este conjunto de datos sirve como base para las empresas que buscan aplicar conocimientos basados en datos para una mejor toma de decisiones y estrategias centradas en el cliente.

## Alcance del Dataset

Este dataset abarca información detallada sobre clientes, sus compras, métodos de pago utilizados, envíos realizados, y productos comprados:

- Clientes: Datos demográficos y de contacto.
- Compras\_Clientes: Detalles de las compras realizadas por los clientes.
- Historial\_Compras: Historial de compras y comportamiento de compra.
- Método\_Pago: Información sobre los métodos de pago utilizados.
- Envíos: Información sobre los envíos realizados.
- Productos: Detalles de los productos comprados.

## Usuario final y nivel de aplicación del análisis (táctico, operativo, estratégico)

**Usuario Final:** Equipo de marketing, ventas, logística, finanzas y alta dirección.

Nivel de Aplicación:

- **Operativo:** Mejora de eficiencia diaria (logística y ventas).
- **Táctico:** Estrategias a corto y mediano plazo (marketing y finanzas).

Estratégico: Decisiones a largo plazo (alta dirección y gestión).

## Entidad-relación (ERD).

En la siguiente página se observará las relaciones entre las tablas trabajadas, los atributos principales y el tipo de relación.

- **Cliente a Compra\_Cliente:** Un cliente puede tener muchas compras, pero cada compra está asociada a un solo cliente.
- **Compras\_Clientes a Historial\_Compras:** Una compra puede tener un historial, y cada historial está asociado a una sola compra.
- **Compras\_Clientes a Método\_Pago:** Una compra puede tener múltiples métodos de pago, pero cada método de pago está asociado a una sola compra.
- **Clientes a Envíos:** Un cliente puede tener muchos envíos, pero cada envío está asociado a un solo cliente.
- **Compras\_Clientes a Envíos:** Una compra puede tener múltiples envíos, pero cada envío está asociado a una sola compra.
- **Compra\_Cliente a Productos:** Una compra puede tener múltiples productos.

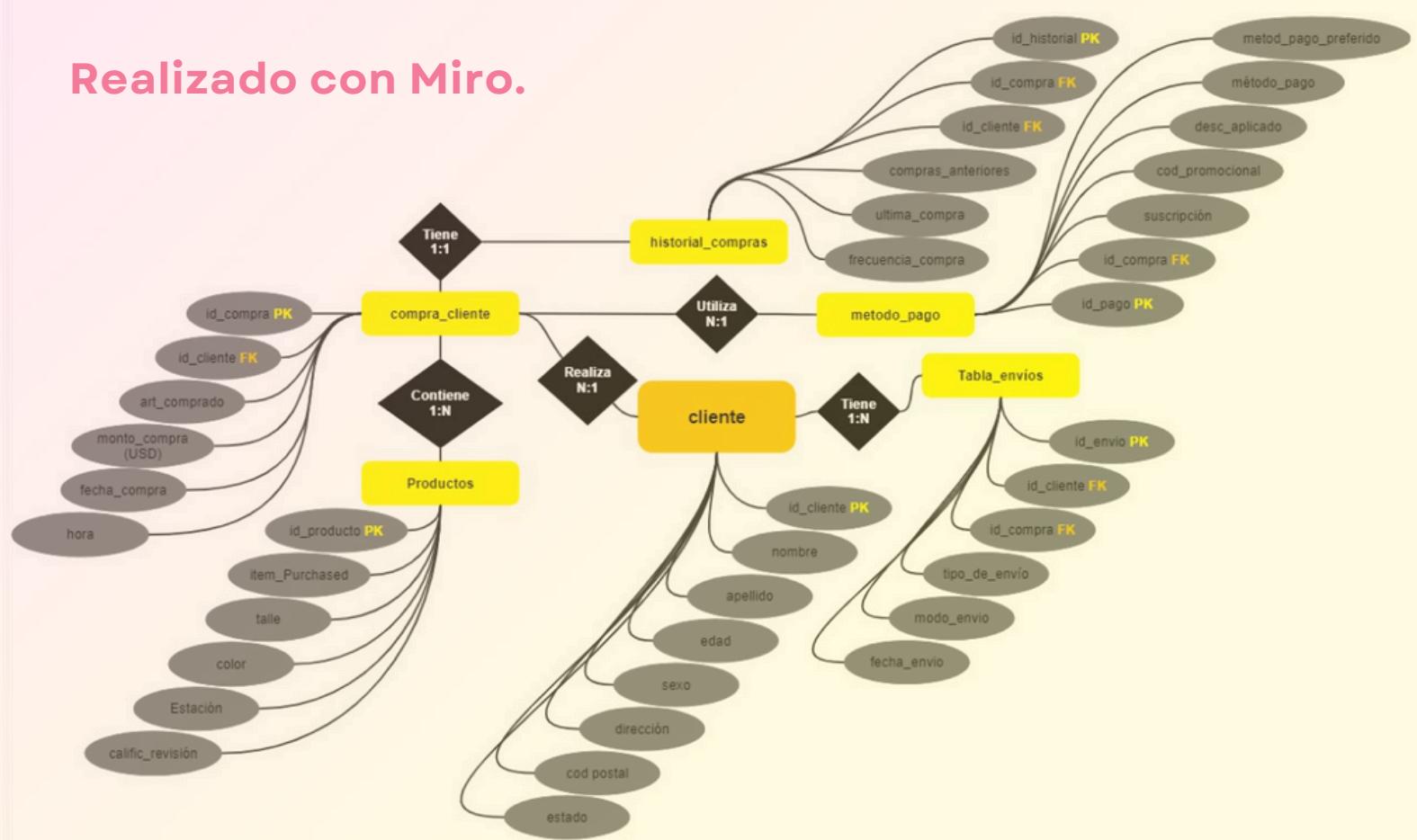
### Entidades y atributos:

<b>01</b>	<b>02</b>	<b>03</b>
<b>Cliente</b>	<b>Compra_Cliente</b>	<b>Historial_Compras</b>
<ul style="list-style-type: none"> <li>• id_cliente (PK)</li> <li>• nombre</li> <li>• apellido</li> <li>• edad</li> <li>• sexo</li> <li>• dirección</li> <li>• cod_postal</li> <li>• Estado</li> </ul>	<ul style="list-style-type: none"> <li>• id_compra (PK)</li> <li>• id_cliente (FK)</li> <li>• art_comprado</li> <li>• monto_compra (USD)</li> <li>• fecha_compra</li> <li>• hora</li> </ul>	<ul style="list-style-type: none"> <li>• id_historial (PK)</li> <li>• id_compra (FK)</li> <li>• id_cliente (FK)</li> <li>• compras_anteriores</li> <li>• ultima_compra</li> <li>• frecuencia_compra</li> </ul>
<b>04</b>	<b>05</b>	<b>06</b>
<b>Método_Pago</b>	<b>Envíos</b>	<b>Productos</b>
<ul style="list-style-type: none"> <li>• id_pago (PK)</li> <li>• id_compra (FK)</li> <li>• suscripción</li> <li>• cod_promocional</li> <li>• desc_aplicado</li> <li>• método_pago</li> <li>• metod_pago_preferido</li> </ul>	<ul style="list-style-type: none"> <li>• id_envio (PK)</li> <li>• id_cliente (FK)</li> <li>• id_compra (FK)</li> <li>• tipo_de_envío</li> <li>• modo_envio</li> <li>• fecha_envio</li> </ul>	<ul style="list-style-type: none"> <li>• id_producto (PK)</li> <li>• item_Purchased</li> <li>• talle</li> <li>• color</li> <li>• Estación</li> <li>• calific_revisión</li> </ul>

# DIAGRAMA ENTIDAD-RELACION

Aquí se muestra el DER del proyecto final, creado al principio del proyecto. El mismo busca describir las relaciones entre las distintas entidades, donde se muestra cada atributo principal del mismo.

Realizado con Miro.



\*BASE DE DATOS ORGANIZADA Y ESTRUCTURADA (FORMATO EXCEL)  
ADJUNTADA EN LA CARPETA DEL PROYECTO FINAL

# LISTADO DE COLUMNAS POR TABLAS

TABLA 1: TABLA\_CLIENTES

Tipo de clave	Campo	Tipo de datos
PK	id_cliente	INT
-	nombre	VARCHAR(30)
-	apellido	VARCHAR(30)
-	edad	INT
-	sexo	VARCHAR(10)
-	dirección	VARCHAR(100)
-	cod postal	INT
-	Estado	VARCHAR(50)

TABLA 2: COMPRA\_CLIENTE

Tipo de clave	Campo	Tipo de datos
PK	id_compra	VARCHAR
FK	id_cliente	INT
-	art_comprado	VARCHAR
-	monto_compra (USD)	MONEY
-	fecha_compra	DATE
-	hora	TIME

TABLA 3: HISTORIAL\_COMPRA

Tipo de clave	Campo	Tipo de datos
PK	id_historial	VARCHAR
FK	id_compra	VARCHAR
FK	id_cliente	INT
-	compras_anteriores	INT
-	ultima_compra	DATE
-	frecuencia_compra	VARCHAR

# LISTADO DE COLUMNAS POR TABLAS

**TABLA 4: METODO\_PAGO**

Tipo de clave	Campo	Tipo de datos
PK	id_pago	VARCHAR
FK	id_compra	VARCHAR
-	suscripción	BIT
-	cod_promocional	BIT
-	desc_aplicado	BIT
-	método_pago	VARCHAR
-	metod_pago_preferido	VARCHAR

**TABLA 5: TABLA\_ENVIOS**

Tipo de clave	Campo	Tipo de datos
PK	id_envio	VARCHAR
FK	id_cliente	INT
FK	id_compra	VARCHAR
-	tipo_de_envío	VARCHAR
-	modo_envio	VARCHAR
-	fecha_envio	DATE

**TABLA 6: PRODUCTOS**

Tipo de clave	Campo	Tipo de campo
PK	id_producto	VARCHAR
-	item_Purchased	VARCHAR
-	talle	VARCHAR
-	color	VARCHAR
-	Estación	VARCHAR
-	calific_revisión	DECIMAL(2,1)

# MODELO RELACIONAL POWER BI.

Una vez que los archivos planos fueron subidos a Power BI, para el correcto análisis realizamos las siguientes transformaciones:

- Se incorpora la columna “**id\_producto**” en “**historial\_compras**” con el fin de concatenar la tabla “**Productos**”.
- Se creó una tabla puente para las relaciones de muchos a muchos.

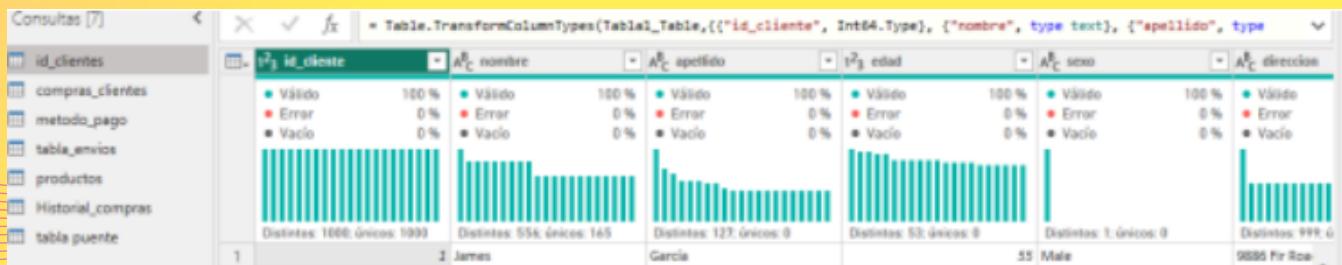
Lo primero que se realizó fue identificar las tablas y los campos involucrados. En este caso se trabajará

= **Table.Combine({compras\_clientes, tabla\_envios})**

Luego, creamos la tabla puente desde Power Query > Anexar consulta > Anexar consulta nueva.

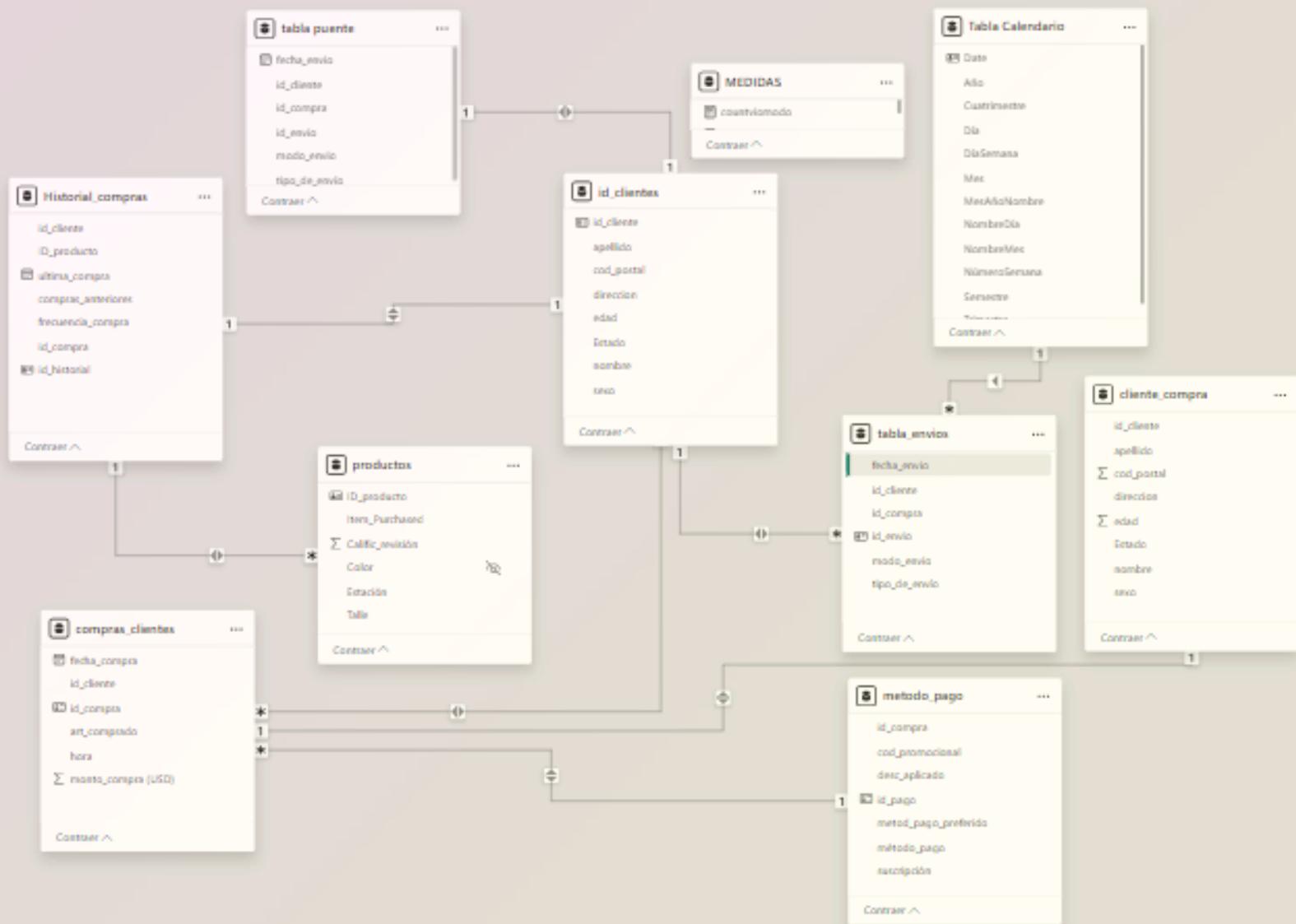
- Se modificó aquellos ID que contenían texto para quedar como campo numérico.
- Se quitó como medidas a aquellos campos que no pertenecen. por ejemplo los ID. Pasamos los campos de medida a dimensión para mantener una correcta estructura de datos del proyecto.
- Se revisó “**tipo de datos**” de cada uno de los campos de cada tabla, al importarlas a PBI. Se comprobó que

todos los datos estén correctos.



# DIAGRAMA ENTIDAD-RELACION PBI

El diagrama entidad-relación quedó graficado en Power BI como se puede observar a continuación:



# VISUALIZACION DE LOS DATOS

## Tabla de calendario.

Se creó una nueva tabla que abarca desde el primer día del año de la fecha de compra más temprana

registrada en la tabla de "**fecha\_envios**" hasta el último día del año más reciente en esa misma tabla,

generando un rango de fechas que cubre todo el período desde el primer hasta el último año en que se registraron compras.

Date
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
6/2/2019
7/2/2019
8/2/2019
9/2/2019

```
Tabla Calendario =
ADDCOLUMNS(
CALENDAR(
DATE(YEAR(MIN(tabla_envios[fecha_envio])), 1, 1),
DATE(YEAR(MAX(tabla_envios[fecha_envio])), 12, 31)),
"Año", YEAR([Date]),
"Mes", MONTH([Date]),
"Día", DAY([Date]),
"DíaSemana", WEEKDAY([Date],2),
"NúmeroSemana", WEEKNUM([Date],2),
"NombreDía", FORMAT([Date], "dddd"),
"NombreMes", FORMAT([Date], "MMMM"),
"MesAñoNombre", FORMAT([Date], "MMMM YYYY"),
"Trimestre", ROUNDUP(MONTH([Date]) / 3, 0),
"Cuatrimestre", ROUNDUP(MONTH([Date]) / 4, 0),
"Semestre", ROUNDUP(MONTH([Date]) / 6, 0))
```

Se anexa fragmento de la misma:

Date	Año	Mes	Día	DíaSemana	NúmeroSemana	NombreDía	NombreMes	MesAñoNombre	Trimestre	Cuatrimestre	Semestre
1/5/2015 00:00:00	2015	5	1	5	18	viernes	mayo	mayo 2015	2	2	1
2/5/2015 00:00:00	2015	5	2	6	18	sábado	mayo	mayo 2015	2	2	1
3/5/2015 00:00:00	2015	5	3	7	18	domingo	mayo	mayo 2015	2	2	1
4/5/2015 00:00:00	2015	5	4	1	19	lunes	mayo	mayo 2015	2	2	1
5/5/2015 00:00:00	2015	5	5	2	19	martes	mayo	mayo 2015	2	2	1
6/5/2015 00:00:00	2015	5	6	3	19	miércoles	mayo	mayo 2015	2	2	1
7/5/2015 00:00:00	2015	5	7	4	19	jueves	mayo	mayo 2015	2	2	1
8/5/2015 00:00:00	2015	5	8	5	19	viernes	Mayo	mayo 2015	2	2	1

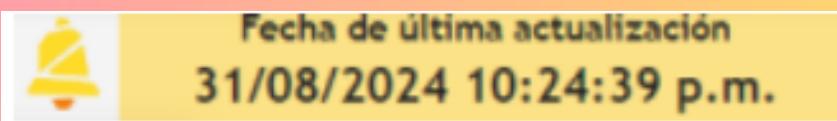
```
Calendario = CALENDAR(MIN(compras_clientes[fecha_compra]), MAX(compras_clientes[fecha_compra]))
```

# VISUALIZACION DE LOS DATOS

## Actualización.

Se agrego fecha de última actualización a la portada:

```
= DateTime.LocalNow()
```

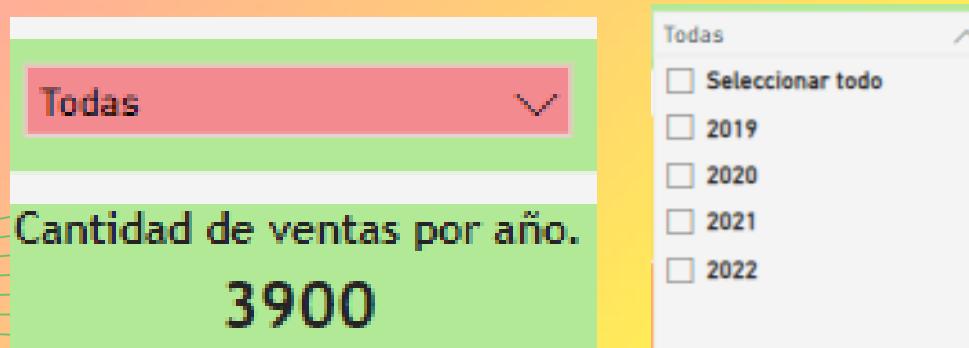


## Tabla Medidas.

Medidas calculadas creadas:

**CLIENTES Y VENTAS.** Análisis del total de ventas por año.

```
totalventas = SUM(compras_clientes[monto_compra (USD)])
```



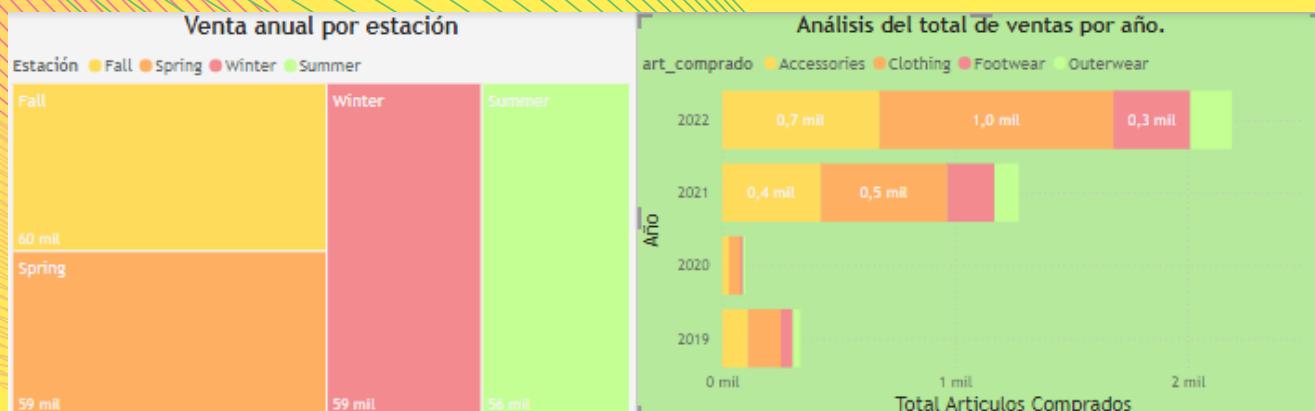
## Ventas por año, tipo de Producto y estacion.

Objetivo: Observar los productos más vendidos en un año específico.

Uso: Entender qué productos son más populares y cuándo.

Análisis de Compras por Estación. Se muestran las ventas de productos agrupadas por estación.

```
Total Articulos Comprados = COUNT(compras_clientes[art_comprado])
```



# VISUALIZACION DE LOS DATOS

## ENVÍOS.

### Análisis de Envíos por Modo de Envío

Descripción: Analiza los envíos (**id\_envio**) por tipo y modo de envío (**modo\_envio**). Se filtraran los envíos por modos preferidos por los clientes para optimizar el proceso logístico según la preferencia de los clientes.

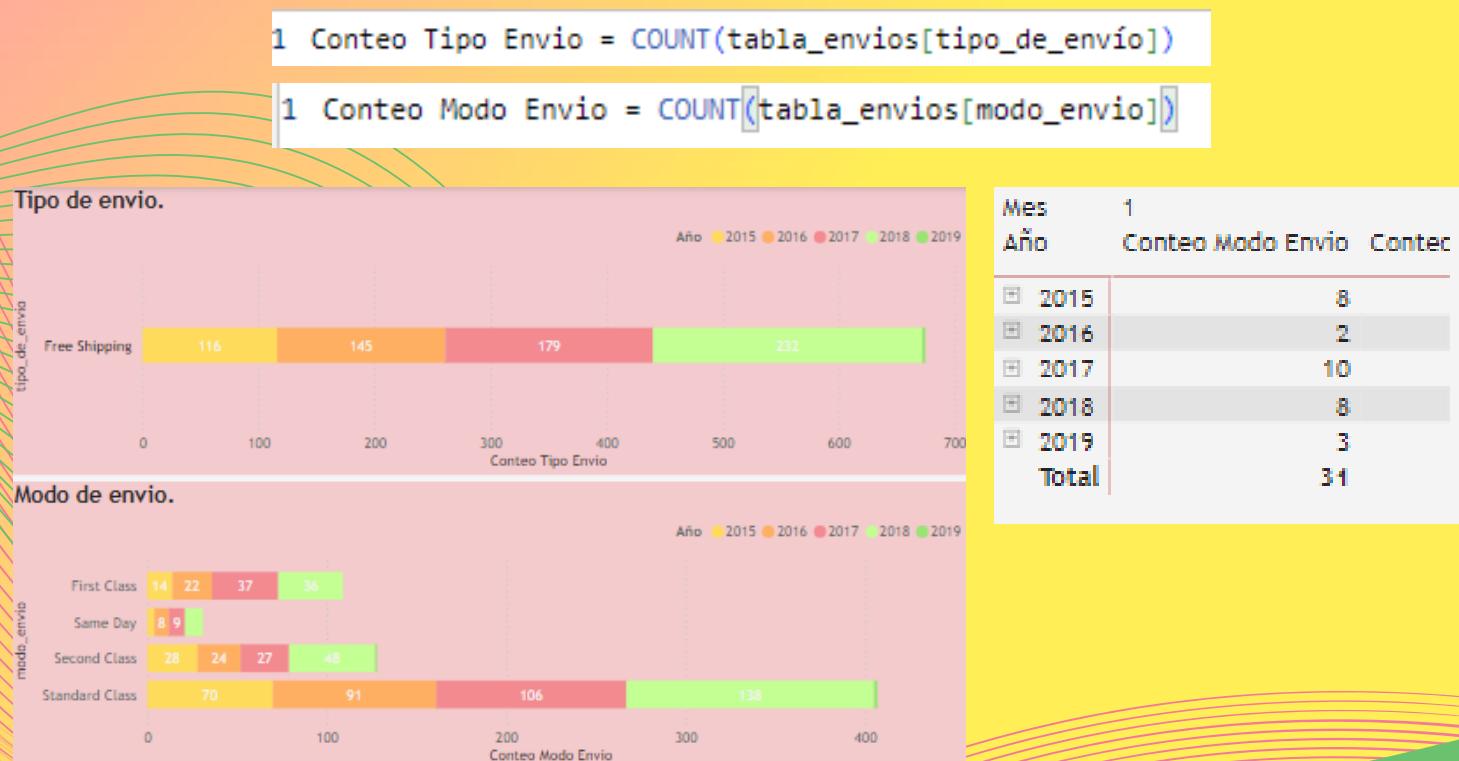
```
promenvios = DISTINCTCOUNT((tabla_envios[modo_envio]))  
counttipo = COUNT(tabla_envios[tipo_de_envio])  
countviomodo = COUNT(tabla_envios[modo_envio])
```

**Envios.**

Modo de envío mas utilizado

Standard Class

Ademas, se agrego un panel con filtros de mes año y modo de envio. Se agrego una matriz donde se puede observar detalladamente la cantidad de modo y tipo de envio. Realizado con medidas DAX ambos.

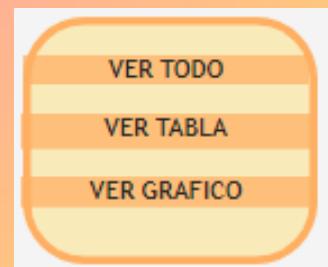


Para trabajar de forma mas eficiente se realizaron los siguientes filtros:

Tipo de envío							Mes	Año
Seleccionar todo	2-Day Shipping	Express	Free Shipping	Next Day Air	Standard	Store Pickup	<input type="checkbox"/> Seleccionar todo	<input checked="" type="checkbox"/> 2015
							<input type="checkbox"/> enero	<input checked="" type="checkbox"/> 2016
							<input type="checkbox"/> febrero	<input checked="" type="checkbox"/> 2017

## TOOLTIPS

Se agregan 3 tooltips para optimizar la pagina “envios”.



## MÉTODOS DE PAGO Y PROMOCIONES.

### Promociones Aplicadas

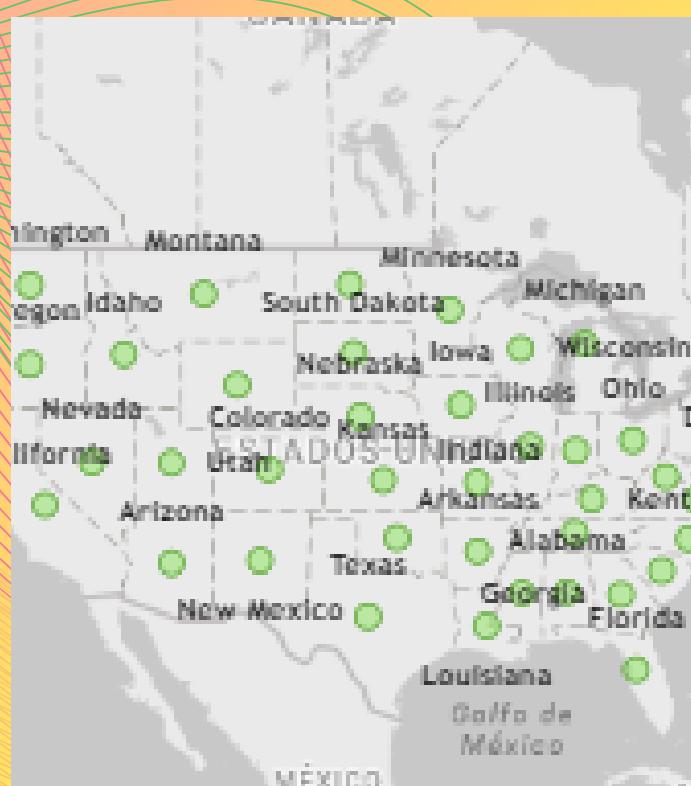
Se busca visualizar el uso de códigos promocionales (**cod\_promocional**) y descuentos aplicados (**desc\_aplicado**) para evaluar la efectividad de las promociones.

```
PorcentajeVentas = DIVIDE(SUM(compras_clientes[monto_compra (USD)]), SUMX(Historial_compras, [totalventas]), 0)
```

Además, de esta forma podremos observar aquellos métodos de Pago Preferidos. se filtra los métodos de pago más utilizados y los menos utilizados utilizando medidas DAX.

Uso: Identificar preferencias de pago y ajustar ofertas.

(Se agrega un mapa para poder localizar los estados y sus métodos de pago preferidos y el uso de suscripciones y descuentos). Ademas, agregamos un recuento de cantidad de compras (ventas) por Estado la cual tambien podemos filtrar por el metodo de pago utilizado.

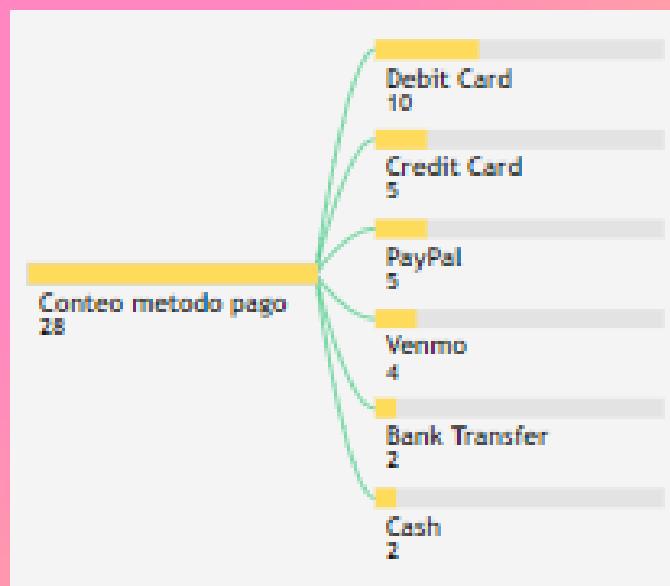


Por ejemplo, el Florida se realizaron un total de 37 compras. Al filtrar por “metodo de pago” podemos observar que solo 6 fueron realizadas con Paypal y de esas comprar solo un total de 3 utilizaron descuentos en el monto.

Estado	Metodo de pago	Desc. aplicado
Alabama	Bank Transfer	No
Alaska	Cash	No
Arizona	Credit Card	No
Arkansas	Debit Card	No
California	PayPal	No
Colorado	Venmo	No
Connecticut		
Delaware		
Florida		Yes

Se agregan medidas DAX para los métodos de pago:

```
Conteo metodo pago = COUNT(metodo_pago[método_pago])
```



## STOCK

```
Stock Actual =
CALCULATE(
    SUM(productos[ID_producto]),
    FILTER(
        productos,
        productos[ Item_Purchased ] = MAX(productos[ Item_Purchased ])
    )
)
```

# CONCLUSION

Durante el análisis de nuestro dataset, pudimos validar varias hipótesis relacionadas con el comportamiento de compra de los clientes en nuestra tienda de ropa. A continuación, se presentan las conclusiones principales obtenidas:

- **Preferencia de género en la compra de ropa y accesorios:** Nuestros datos confirman que las mujeres tienden a comprar más ropa que los hombres y muestran una tendencia a gastar más en accesorios. Esta tendencia sugiere una oportunidad para ajustar nuestras estrategias de marketing y oferta de productos según las preferencias específicas de género.
- **Patrones estacionales en las ventas de ropa:** Observamos patrones claros en las ventas de ropa de temporada, como la demanda de abrigos en invierno y trajes de baño en verano. Estos patrones predecibles nos permiten optimizar la gestión de inventario y la planificación de campañas promocionales para maximizar las ventas durante períodos específicos del año.
- **Incremento de ventas de ropa deportiva en verano:** Las ventas de ropa deportiva aumentan significativamente durante los meses de verano. Este hallazgo respalda la necesidad de enfocar nuestras estrategias de marketing y producto hacia artículos deportivos durante esta temporada alta.
- **Métodos de pago predominantes:** Finalmente, aproximadamente el 80% de nuestros clientes prefieren pagar con tarjeta de crédito, lo cual subraya la importancia de mantener opciones de pago flexibles y seguras para satisfacer las necesidades de nuestra clientela.

Estas conclusiones no solo reflejan los patrones observados en nuestros datos, sino que también proporcionan una base sólida para futuras estrategias comerciales y decisiones operativas dentro de nuestra tienda de ropa.

# FUTURAS LINEAS

En este apartado, se consideraron ciertos puntos que no fueron incluidos en el presente trabajo pero que se podrían añadir a futuro:

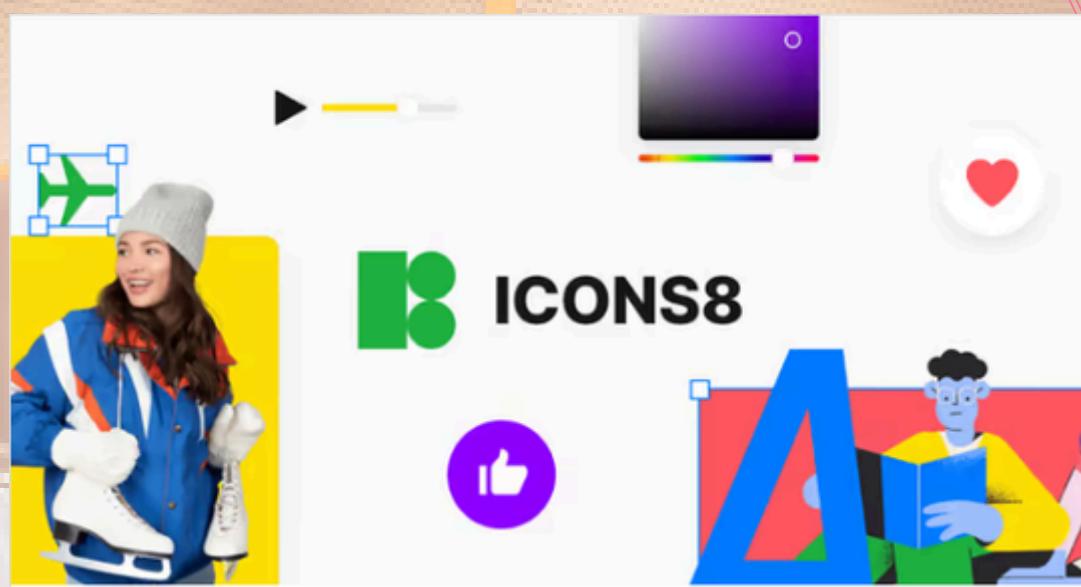
- **Optimización de inventario:** Con los patrones predecibles de ventas de ropa sería necesario expandir en dataset agregando un **stock por tipo de elemento y color**. Esto ayudaría a no revisar constantemente los datos de ventas anteriores y predecir las futuras necesidades de stock para evitar tanto faltantes como excesos de inventario.
- Incluir mas atributos como poder comparar cambios significativos en el tiempo, género y edad para conocer mas en detalle a los clientes.
- Expandir el análisis a nivel mundial ya que el análisis solo abarca las tiendas de los estados de ESTADOS UNIDOS.
- Incluir los costos de envíos y logística para obtener un detalle de los gastos y el nivel de satisfacción del cliente con el modo y tipo de envío seleccionado.



# BIBLIOGRAFIA



**Customer Shopping Trends Dataset**  
Journey into Consumer Insights and Retail Evolution with Synthetic Data  
[kaggle.com](https://kaggle.com)



**ICONS8**

Símbolos e iconos de Negocios en estilo Pieces, PNG, SVG

Encuentra símbolos, iconos y cliparts de Negocios para Android, iOS, Twitter, Discord. Iconos en PNG, SVG estilo Pieces. Edita los logos y símbolos de Negocios online.

 Icons8

**Vacunación contra el en la COVID-19 Ciudad Autónoma de Buenos Aires Administración de las vacunas y cuantificación del posible ahorro (Lic. Matias Tello )**



**SECRETOS DE POWER BI**  
datdata

Secretos de Power BI para que destakes en tu trabajo  
Únete a la comunidad más grande de entusiastas de Power BI. Tendrás acceso a cursos desde cero, eventos exclusivos, concursos y más.  
[datdata.com](https://datdata.com)