

CODER HOUSE

Data Analytics

Iris Martinez

Comisión: 57440

Iris Martinez

[COMPANY NAME] [Company address]

Introducción

En el dinámico panorama empresarial actual, la capacidad de tomar decisiones informadas y estratégicas es esencial para el éxito de cualquier organización. El análisis de datos de ventas, como uno de los indicadores clave de rendimiento, desempeña un papel fundamental en la evaluación de la salud financiera y la dirección futura de una empresa. Con el advenimiento de herramientas tecnológicas avanzadas, como SQL y Power BI, las empresas tienen la oportunidad de transformar datos crudos en información valiosa y visualmente impactante.

Este documento se centra en el análisis de datos de ventas utilizando SQL y Power BI como herramientas principales. Exploraremos cómo SQL, un lenguaje de programación estándar para la gestión y manipulación de bases de datos, y Power BI, una plataforma de análisis de datos de Microsoft, pueden permitir a las organizaciones desglosar, comprender y aprovechar los datos de ventas de manera efectiva. Desde la consulta y extracción de datos con SQL hasta la creación de visualizaciones perspicaces y la identificación de tendencias y patrones ocultos con Power BI, el análisis de datos de ventas con estas herramientas ofrece a las empresas la capacidad de tomar decisiones más fundamentadas y estratégicas.

A lo largo de este documento, examinaremos las etapas clave para diseñar y desarrollar un tablero de análisis de datos de ventas efectivo utilizando SQL y Power BI. Además, exploraremos cómo la interactividad, la visualización de datos y las capacidades de generación de informes en tiempo real de Power BI pueden empoderar a los equipos de ventas y a los stakeholders para comprender mejor el rendimiento comercial de la empresa y formular estrategias basadas en evidencia.

En última instancia, este documento busca resaltar la importancia del análisis de datos de ventas en el contexto empresarial actual y cómo la combinación de SQL y Power BI se erige como una solución valiosa para transformar datos de ventas en conocimientos accionables.

Descripción de la temática de los datos.

El conjunto de datos utilizado en este análisis abarca información detallada sobre las ventas de una amplia variedad de productos, con un énfasis significativo en artículos deportivos. Este dataset comprende datos exhaustivos de cada transacción, incluyendo el comprador, el lugar y la fecha de la compra. Además, se detalla el tipo de producto vendido, junto con su categoría específica, género y una descripción completa del mismo. La información recopilada no solo permite identificar qué productos tienen mayor demanda, sino también analizar patrones de compra basados en variables demográficas y geográficas. Este nivel de detalle facilita un análisis profundo y multidimensional del comportamiento del consumidor, proporcionando a las empresas una valiosa oportunidad para optimizar sus estrategias de ventas y marketing. Con estos datos, es posible identificar tendencias emergentes, evaluar el rendimiento de diferentes líneas de productos y mejorar la toma de decisiones basada en evidencia concreta.

El propósito fundamental de este Dataset es proporcionar a los analistas financieros, gerentes y stakeholders de la empresa una visión detallada del rendimiento financiero, centrándose especialmente en los aspectos relacionados con la facturación. Mediante la aplicación de técnicas avanzadas de análisis de datos y el uso de herramientas como Power BI, este conjunto de datos puede ser transformado en información visualmente atractiva y fácilmente interpretable. Esta visualización de datos facilita la toma de decisiones estratégicas fundamentadas y contribuye al impulso del éxito empresarial al proporcionar insights relevantes y claros sobre el desempeño financiero de la organización.

Alcance

El alcance del proyecto de Power BI centrado en el análisis de la facturación de una empresa puede abarcar diversas etapas y objetivos clave para obtener una comprensión completa y valiosa de los aspectos financieros. A continuación, se describe un alcance típico para un proyecto de este tipo:

- 1. Definición de objetivos:** establecer objetivos claros para el proyecto, como identificar patrones de ventas, analizar la rentabilidad de los productos y comprender las tendencias de facturación a lo largo del tiempo.
- 2. Recopilación de datos:** obtener y preparar los datos de facturación de la empresa, que incluirán información sobre transacciones, productos, precios, clientes y fechas.
- 3. Limpieza y transformación de datos:** realizar limpieza de datos para manejar valores nulos, datos duplicados o incoherentes. Transformar los datos según sea necesario para crear una estructura adecuada para el análisis, como la creación de tablas de hechos y dimensiones.
- 4. Creación de un modelo de datos:** diseñar un modelo de datos en Power BI que refleje las relaciones entre las tablas y permita un análisis coherente y eficaz.
- 5. Desarrollo de visualizaciones:** crear visualizaciones interactivas, como gráficos de barras, líneas, tortas y mapas, para mostrar la facturación por período, productos, categorías, etc. Desarrollar tableros interactivos que permitan a los usuarios explorar los datos y obtener información relevante.
- 6. Análisis y descubrimiento:** identificar tendencias en los ingresos a lo largo del tiempo y realizar análisis comparativos entre productos, categorías o segmentos de clientes. Evaluar la influencia de los descuentos, promociones u otros factores en la facturación.
- 7. Implementación de métricas clave:** calcular y mostrar métricas financieras clave, como ingresos totales, promedio de ventas por cliente, margen de beneficio, etc.
- 8. Generación de informes y Dashboards:** diseñar informes ejecutivos y Dashboards que presenten los hallazgos y permitan a los usuarios acceder rápidamente a los datos relevantes.
- 9. Capacitación y documentación:** proporcionar capacitación a los usuarios finales sobre cómo interactuar con el tablero y realizar análisis personalizados. Documentar el proceso de preparación de datos, modelado y visualización para futuras referencias.
- 10. Entrega y seguimiento:** presentar el proyecto a los stakeholders y usuarios finales, recopilando comentarios y realizando ajustes si es necesario.
- 11. Mantenimiento continuo:** actualizar regularmente el tablero con nuevos datos para mantener la relevancia y precisión de las visualizaciones. Este alcance proporciona una estructura general para abordar un proyecto de Power BI relacionado con la facturación de la

empresa. Sin embargo, el alcance específico puede variar según las necesidades y metas de la organización.

Hipótesis

1. Variación por temporadas: La facturación de la empresa varía según las estaciones del año. Se hipotetiza que las ventas aumentan durante las temporadas festivas, como Navidad o el verano.
2. Producto estrella: Se supone que ciertos productos destacan significativamente en la facturación total de la empresa, lo que podría indicar la presencia de productos líderes en ventas que contribuyen en gran medida a los ingresos.
3. Impacto del día de la semana: La facturación puede ser diferente según el día de la semana. Se plantea la hipótesis de que los fines de semana generan un mayor volumen de ventas en comparación con los días laborables.
4. Segmentación de clientes: Se sugiere que diferentes segmentos de clientes tienen un impacto significativo en la facturación, y se investigará si ciertos grupos de clientes contribuyen más que otros a los ingresos de la empresa.
5. Efectividad de los descuentos: Se presume que la aplicación de descuentos en productos específicos puede aumentar la facturación total, al atraer a más clientes o estimular compras adicionales.
6. Tendencias históricas: Se plantea la hipótesis de que la facturación ha experimentado un crecimiento constante o una disminución a lo largo de un período de tiempo específico.
7. Correlación entre productos: Se investigará si existe una relación entre la venta de ciertos productos, sugiriendo que la compra de un artículo específico puede impulsar la venta de otro relacionado.
8. Impacto de los canales de venta: Se supone que ciertos canales de venta, como las tiendas físicas versus el comercio electrónico, tienen un impacto diferente en la facturación total de la empresa.
9. Respuesta a eventos externos: Se sugiere que eventos externos, como cambios en la economía o tendencias del mercado, pueden influir en la facturación de la empresa.
10. Influencia geográfica: Se investigará si la ubicación geográfica tiene algún efecto en los patrones de facturación, como si las ventas varían según la región o el país.

Herramientas de tecnología implementadas.

En el contexto de este proyecto, se seleccionan las siguientes aplicaciones y herramientas con el propósito de llevar a cabo las tareas de análisis requeridos:

- Microsoft Excel: para la lectura de archivos CSV que contienen los datos con los que trabajaremos.
- SQL Server Management Studio: para levantar y manipular los datos como una base de datos.
- Visual Paradigm: como herramienta para poder realizar los diagramas necesarios de la documentación.

Diagrama entidad relación.

A continuación, listamos cada tabla que compone la base de datos con la que trabajaremos:

Genero

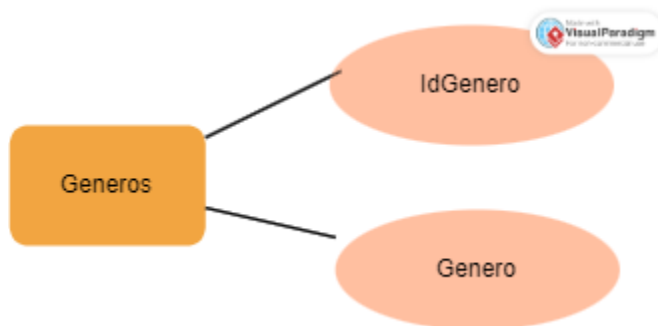


Ilustración 1: Contiene el género al que pertenece cada vendedor que realizó la venta.

Atributo	Tipo de dato	LLave
IdGenero	Int	PK
Genero	Varchar	

Categorías

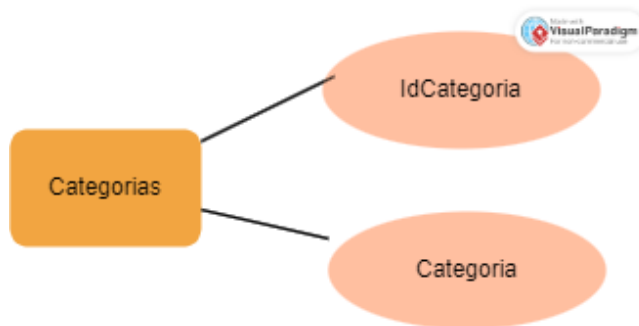


Ilustración 2: Contiene la categoría a la que pertenece la venta.

Atributo	Tipo de dato	LLave
IdCategoria	Int	PK
Categoria	Varchar	

Colores

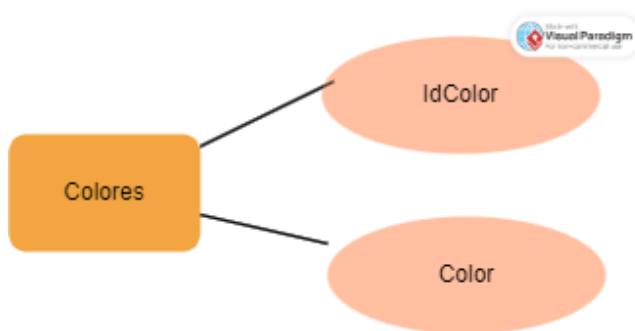


Ilustración 3: Contiene el color el cual pertenece al producto.

Atributo	Tipo de dato	LLave
IdColor	Int	PK
Color	Varchar	

Familias

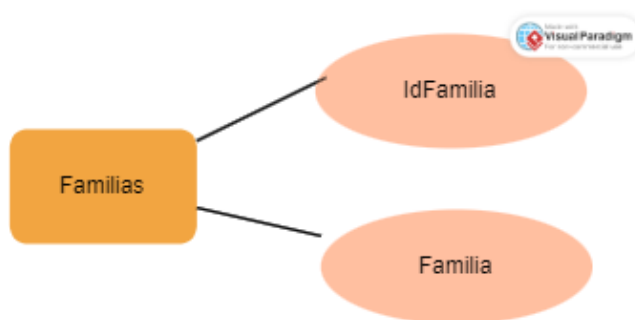


Ilustración 4: Contiene la familia al cual pertenece al producto.

Atributo	Tipo de dato	LLave
IdFamilia	Int	PK
Familia	Varchar	

Estados

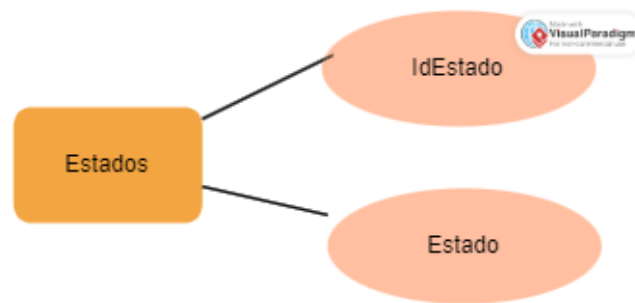


Ilustración 5: Contiene el nombre de los estados donde se realizan las ventas.

Atributo	Tipo de dato	LLave
IdEstado	Int	PK
Estado	Varchar	

Ciudades

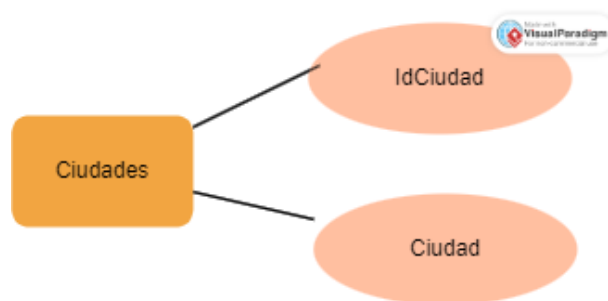


Ilustración 7: Contiene el nombre de los ciudades donde se realizan las ventas.

Atributo	Tipo de dato	LLave
IdCiudad	Int	PK
Ciudad	Varchar	

Tipo de producto

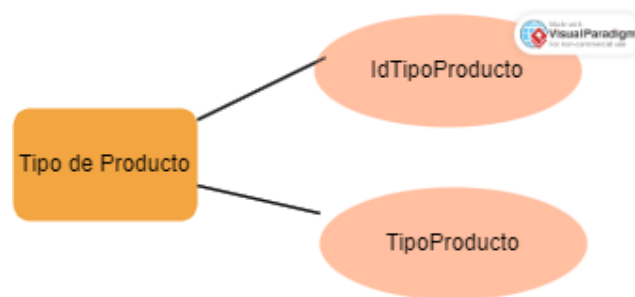


Ilustración 8: Contiene el tipo de producto al que pertenece el producto comprado.

Atributo	Tipo de dato	LLave
IdTipoProducto	Int	PK
TipoProducto	Varchar	

Tipo de venta

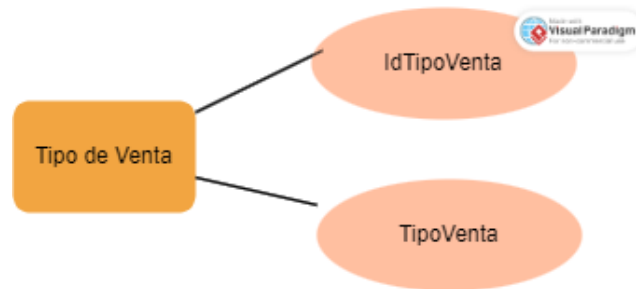


Ilustración 9: Contiene el tipo de venta al que pertenece la venta realizada.

Atributo	Tipo de dato	LLave
IdTipoVenta	Int	PK
TipoVenta	Varchar	

Vendedor

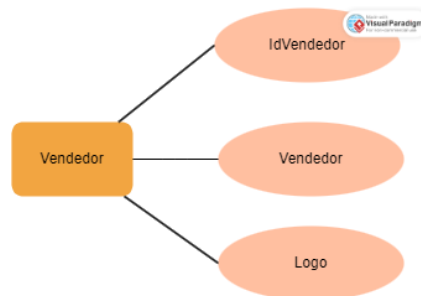


Ilustración 10: Contiene el vendedor que realizó la venta.

Atributo	Tipo de dato	LLave
IdVendedor	Int	PK
Vendedor	Varchar	
Logo	Varchar	

Región

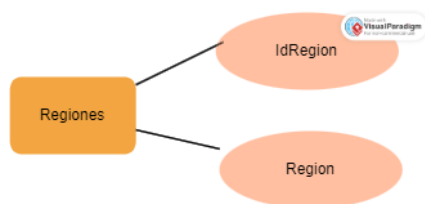


Ilustración 11: Contiene las regiones donde se realizó la venta.

Atributo	Tipo de dato	LLave
IdRegion	Int	PK
Region	Varchar	

Ventas

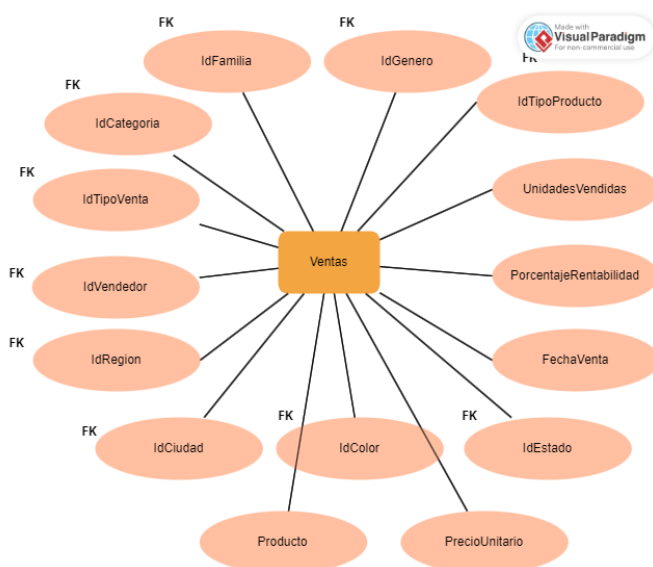


Ilustración 12: Contiene toda la información de las ventas realizadas.

Atributo	Tipo de dato	LLave
IdRegion	Int	FK
IdVendedor	Int	FK
IdTipoVenta	Int	FK
IdTipoProducto	Int	FK
IdCiudad	Int	FK
IdEstado	Int	FK
IdFamilia	Int	FK
IdColor	Int	FK
IdCategoria	Int	FK
IdGenero	Int	FK
Unidades vendidas	Int	
Porcentaje de Rentabilidad	Int	
Fecha de Venta	Date	

Precio Unitario	Int	
Producto	Varchar	

Generación de fondos para los tableros

La efectividad de estos tableros no solo depende de la precisión de los datos, sino también de la claridad y atractivo visual de su presentación. Es aquí donde la generación de fondos personalizados para tableros juega un papel crucial. Utilizando PowerPoint, hemos desarrollado una serie de fondos diseñados para mejorar la estética y funcionalidad de los tableros en PowerBI. Estos fondos no solo embellecen las presentaciones, sino que también facilitan la organización y la lectura de la información, permitiendo a los usuarios centrarse en los datos clave sin distracciones.

Aplicación de fondos en Power BI

Usar imágenes como fondo en Power BI es una técnica eficaz para mejorar la apariencia y funcionalidad de los informes. Sin embargo, es importante usarlas con moderación y cuidado para asegurar que la visualización de datos siga siendo clara y efectiva.

Hipótesis del Proyecto

En el documento de Power BI diferentes secciones la primera página del informe es fundamental ya que ofrece una descripción detallada del análisis del proyecto.

Esta sección no solo contextualiza la investigación actual, sino que también establece una base sólida para comprender los objetivos y resultados esperados. Se explora la hipótesis evaluada, lo que brinda claridad sobre el enfoque y propósito del proyecto. Además de la descripción del análisis y la hipótesis, se incluye un glosario que actúa como un recurso esencial para el lector.

Este elemento de diseño contribuye significativamente a la comodidad y accesibilidad general del informe, asegurando que los lectores puedan moverse fácilmente entre las diferentes secciones.

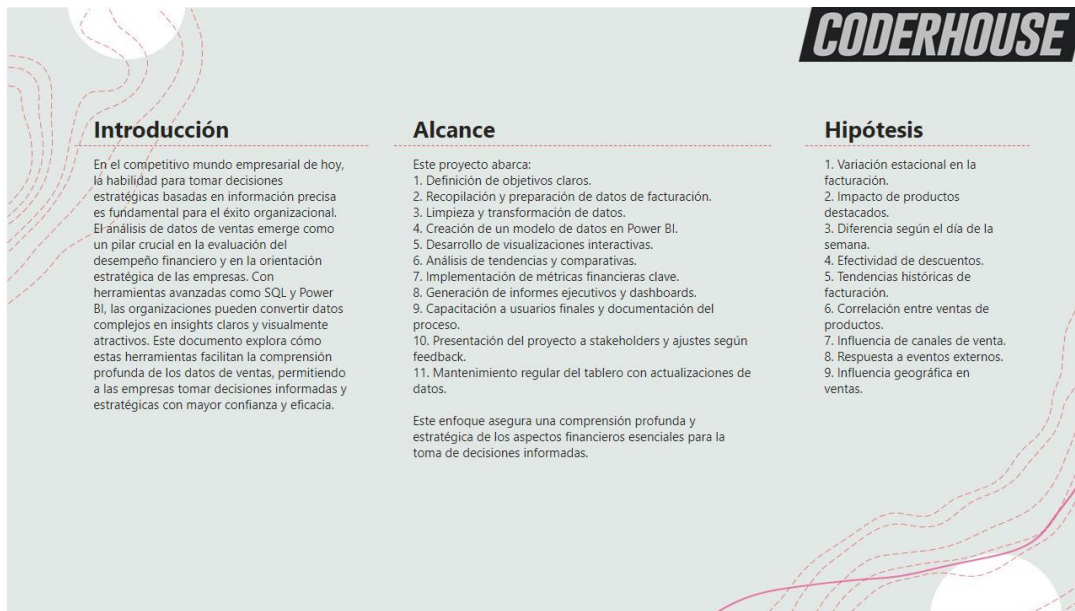


Tabla de fechas

Utilizando diversas funciones DAX, se crea una completa TABLA DE FECHAS que se basa en la fecha más antigua y la fecha más reciente registradas en la tabla VENTAS. Esta tabla ampliada se enriquece con la inclusión de columnas que representan diversos períodos temporales significativos. Estos períodos se configuran estratégicamente para facilitar la generación de visualizaciones impactantes, segmentaciones detalladas y cálculos avanzados relacionados con la inteligencia temporal.

A continuación, se presentan las fórmulas aplicadas en este proceso, diseñadas con precisión para optimizar el análisis de datos y mejorar la comprensión de las tendencias temporales:

Para la generación de la tabla:

CALENDARIO = CALENDAR(FIRSTDATE(VENTAS[FechaVenta]),
LASTDATE(VENTAS[FechaVenta]))

Para calcular los años: Año = YEAR(CALENDARIO[Fecha])

Para calcular los meses: MES = FORMAT(CALENDARIO[Fecha], "mmm")

Para calcular el número del mes: NUMERO MES = MONTH(CALENDARIO[Fecha])

Para calcular los trimestres: Trimestre = "T" & QUARTER(CALENDARIO[Fecha])

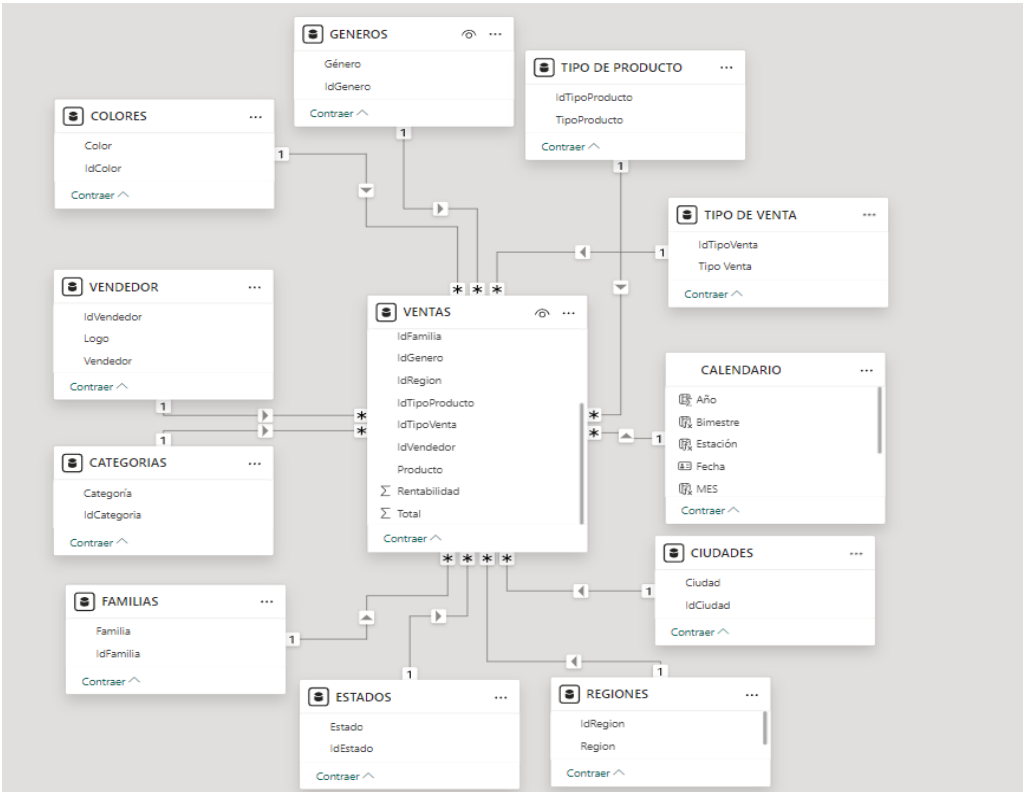
Para calcular los bimestres: Bimestre = SWITCH(TRUE(), CALENDARIO[Numero Mes]<=2, "B1",
CALENDARIO[Numero Mes]<=4, "B2", CALENDARIO[Numero Mes]<=6, "B3", CALENDARIO[Numero
Mes]<=8, "B4", CALENDARIO[Numero Mes]<=10, "B5", "B6")

Para calcular los semestres: Semestre = IF(CALENDARIO[NUMERO MES]<=6, "S1", "S2")

Relaciones entre tablas

Dentro del contexto del modelo de datos que hemos desarrollado, es esencial establecer una comprensión clara de cómo las diversas tablas interconectadas se relacionan entre sí. Estas relaciones son fundamentales para la integridad y la precisión de la información que se presenta y analiza en el modelo. Esta tabla de relaciones actúa como una hoja de ruta esencial para comprender cómo los datos fluyen y se vinculan entre las diferentes partes del modelo de datos. La claridad en las relaciones entre tablas es fundamental para garantizar que el análisis y la presentación de datos sean precisos y significativos. Al proporcionar esta tabla detallada de relaciones, se facilita a los usuarios la comprensión de cómo se estructura el modelo y cómo pueden navegar de manera efectiva a través de los datos para obtener insights valiosos.

Tabla	Tabla	Capo en Común
Regiones	Ventas	IdRegion
Vendedor	Ventas	IdVendedor
TipoVenta	Ventas	IdTipoVenta
TipodeProducto	Ventas	IdTipoProducto
Ciudad	Ventas	IdCiudad
Estado	Ventas	IdEstado
Familias	Ventas	IdFamilia
Color	Ventas	IdColor
Categoría	Ventas	IdCategoria
Genero	Ventas	IdGenero



Columnas Calculadas / Medidas/ Visualizaciones

En Power BI, las MEDIDAS son utilizadas para realizar cálculos específicos o para agregar datos en base a ciertas condiciones. Estas medidas son creadas utilizando el lenguaje de fórmulas de Power BI llamado DAX (Data Analysis Expressions).

Las medidas permiten realizar análisis más avanzados y obtener información específica que no está directamente disponible en los conjuntos de datos originales.

Aquí hay algunas situaciones comunes en las que las medidas son útiles:

- Cálculos personalizados: permiten crear medidas para realizar cálculos personalizados basados en fórmulas DAX.
- Indicadores clave de rendimiento: las medidas son fundamentales para definir y calcular KPI, que son indicadores clave que resumen el rendimiento de una organización en un área específica.
- Análisis de tendencias y variaciones: se pueden utilizar medidas para calcular tasas de crecimiento, variaciones porcentuales, o cualquier otra métrica que permitan analizar tendencias a lo largo del tiempo.
- Filtrado dinámico: al crear medidas, se pueden aplicar filtros dinámicos en los informes, permitiendo ajustar los cálculos según las selecciones realizadas por el usuario.
- Segmentación y agrupación: se pueden utilizar medidas para segmentar o agrupar datos de acuerdo con ciertos criterios, permitiendo un análisis más detallado y personalizado.
- Cálculos condicionales: se pueden crear medidas que realicen cálculos condicionales basados en ciertas condiciones, permitiendo obtener información específica en función de diferentes escenarios.
- Comparación con objetivos o benchmarks: las medidas son útiles para comparar el rendimiento real con objetivos predefinidos o benchmarks establecidos.

En resumen, las medidas en Power BI son herramientas poderosas que permiten realizar análisis más avanzados y personalizados, proporcionando información clave para la toma de decisiones empresariales. Estas medidas pueden ser utilizadas en tablas, gráficos, y otros elementos visuales en informes y paneles.

Para este proyecto, se generan las siguientes medidas:

- Total Facturado: esta medida calcula el total facturado por la empresa, sumando los valores contenidos en la columna TOTAL de la tabla VENTAS.
- TOTAL RENTABILIDAD: suma todos los valores del campo RENTABILIDAD
- FACTURACION PROMEDIO: promedia los valores del campo TOTAL de la tabla VENTAS.
- RENTABILIDAD PROMEDIO: promedia los valores del campo RENTABILIDAD de la tabla VENTAS.

- TASA: la tasa de crecimiento trimestral, tomando como referencia el total facturado por la empresa.
- DIFERENCIA: calcula la diferencia de facturación de un trimestre a otro.
- EVALUACION: califica la rentabilidad trimestral de la empresa, tomando en consideración las siguientes observaciones:
 - Si la rentabilidad total fue inferior a 1.000.000, mostrar una estrella.
 - Si la rentabilidad total se ubicó entre 1.000.000 y 2.500.000, mostrar 2 estrellas.
 - Si la rentabilidad total se ubicó entre 2.500.000 y 3.500.000, mostrar 3 estrellas.
 - Si la rentabilidad total se ubicó entre 3.500.000 y 5.000.000, mostrar 4 estrellas.
 - Si la rentabilidad total fue superior a 5.000.000, mostrar 5 estrellas.

Al finalizar pudimos generar la siguiente matriz de valores:

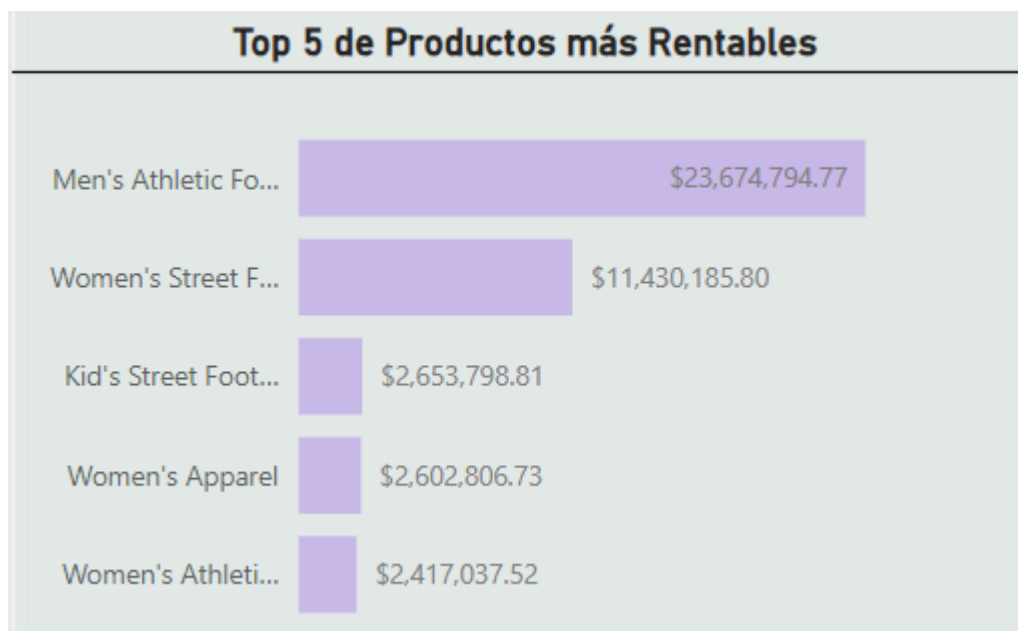
Factuación / Rentabilidad Trimestral por Año								
Año	Total Facturado	Total Rentabilidad	Ventas	Facturación Promedio	Rentabilidad Promedio	Tasa	Diferencia	Evaluación
2020	\$22,677,325	\$8,345,505.09	1,302	17,417.30	6,409.76			
T4	\$3,675,821	\$1,372,455.43	283	12,988.77	4,849.67	-45.88 %	(\$3,115,883)	★★
T2	\$5,962,039	\$2,170,623.96	285	20,919.44	7,616.22	-4.57 %	(\$285,722)	★★
T3	\$6,791,704	\$2,528,034.59	339	20,034.53	7,457.33	13.92 %	\$829,665	★★★
T1	\$6,247,761	\$2,274,391.12	395	15,817.12	5,757.95		\$6,247,761	★★
2021	\$95,264,450	\$37,922,771.15	8,346	11,414.38	4,543.83			
T4	\$25,231,709	\$10,136,371.40	2,085	12,101.54	4,861.57	-9.05 %	(\$2,510,816)	★★★★★
T3	\$27,742,525	\$11,259,987.29	2,112	13,135.67	5,331.43	17.11 %	\$4,053,152	★★★★★
T2	\$23,689,373	\$9,483,353.09	2,105	11,253.86	4,505.16	27.36 %	\$5,088,530	★★★★★
T1	\$18,600,843	\$7,043,059.37	2,044	9,100.22	3,445.72	406.03 %	\$14,925,022	★★★★★
Total	\$117,941,775	\$46,268,276.24	9,648	12,224.48	4,795.63			

Agregamos otros tres distintos gráficos:

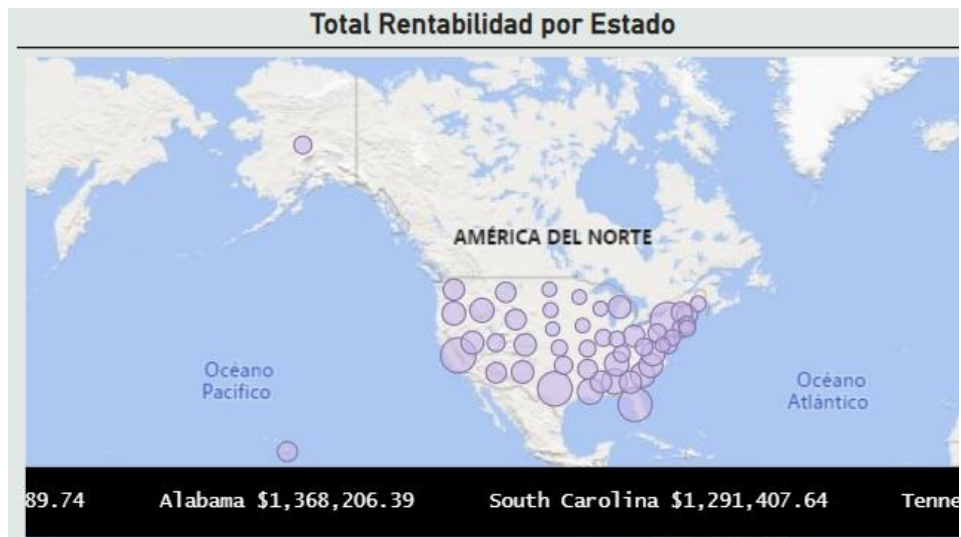
Un gráfico de columnas que muestre la rentabilidad generada por las ventas efectuadas por los distintos tipos de ventas (online, outlet y tienda).



Un gráfico de barras que muestre el TOP 5 de los productos más rentables.



Mapa que muestra la rentabilidad por estado.



Un medidor de rentabilidad:

