



Análisis de ventas del Superstore para el período 2014-2017

D'Ercole Maria Victoria

TABLA DE CONTENIDOS:

- 1. Introducción3
- 2. Descripción de la temática de los datos4
- 3. Herramientas utilizadas5
- 4. Hipótesis6
- 5. Objetivo y alcance del proyecto8
- 6. Usuario final y nivel de aplicación del análisis10
- 7. Diagrama Entidad – Relación10
- 8. Listado de tablas12
- 9. Listado de las columnas de cada tabla17
- 10. Conexión de Base de Datos y Tablas22
- 11. Medidas Calculadas29
- 12. Diseño del tablero33
- 13. Dashboard Power BI.....34
- 14. Diagrama Entidad - Relación de Power BI (Final)47
- 15. Conclusiones Finales49
- 16. Futuras Líneas50
- 17. Referencia Bibliográfica50

1.Introducción

El análisis de datos es fundamental para comprender el funcionamiento de cualquier empresa. En este contexto, el conjunto de datos proporciona una oportunidad emocionante para explorar y comprender el rendimiento de una empresa minorista, con información detallada sobre ventas, clientes, productos y regiones. Este conjunto de datos, disponible en línea, ofrece una visión completa de las operaciones comerciales a lo largo del tiempo. Desde ventas hasta envíos y categorías de productos, cada aspecto está meticulosamente registrado y estructurado para análisis detallados. Consta de múltiples tablas interrelacionadas que permiten una exploración profunda de las relaciones entre diferentes aspectos del negocio.

Para aprovechar al máximo este recurso, se utilizarán herramientas como SQL y Power BI para extraer información significativa, visualizar patrones y tendencias, y extraer conclusiones que impulsen la toma de decisiones informadas. A través de consultas SQL, se profundizará en los datos para extraer información específica, mientras que Power BI permitirá la creación de visualizaciones dinámicas e interactivas para una comprensión más clara de los datos.

Este análisis no solo brindará insights sobre el rendimiento pasado del negocio, sino que también ofrecerá una base sólida para la formulación de estrategias futuras. Al comprender mejor las preferencias del cliente, las tendencias del mercado y los patrones de ventas, la empresa estará mejor equipada para optimizar sus operaciones, mejorar la satisfacción del cliente y maximizar su rentabilidad.

2.Descripción de la temática de los datos

El dataset aborda la temática de las operaciones comerciales de un negocio minorista, proporcionando una visión detallada de diversos aspectos relacionados con las ventas, los clientes y los productos.

El dataset presenta variables relevantes cómo ventas y ganancias, los cuales permiten analizar el rendimiento financiero del negocio y evaluar su rentabilidad a lo largo del tiempo; sobre los productos y categorías para comprender qué productos son los más populares entre los clientes y qué categorías tienen un mejor desempeño en términos de ventas y ganancias; clientes y segmentación que permite realizar análisis de la base de clientes, identificar patrones de compra y segmentar a los clientes según diferentes criterios demográficos o de comportamiento; envíos y logística para monitorear la eficiencia del proceso logístico y mejorar la experiencia del cliente en términos de tiempos de entrega y servicio postventa; región y geografía que permite realizar análisis de la distribución geográfica de la demanda y adaptar las estrategias comerciales a las características específicas de cada región.

Al analizar este conjunto de datos es posible identificar tendencias significativas en las operaciones del negocio minorista. Desde patrones estacionales en las ventas hasta cambios en las preferencias del cliente, el análisis detallado de los datos permite descubrir insights valiosos que pueden informar decisiones estratégicas. Al examinar las fluctuaciones en las ventas, las tendencias emergentes en la demanda de productos y los comportamientos de compra de los clientes, se pueden identificar oportunidades de crecimiento, anticipar

cambios en el mercado y ajustar las estrategias comerciales para maximizar el rendimiento del negocio. El objetivo principal de este conjunto de datos es proporcionar a las partes interesadas una visión completa y detallada de las operaciones comerciales de la empresa minorista. Al analizar los datos disponibles, se busca entender el rendimiento financiero del negocio, identificar áreas de mejora y oportunidades de crecimiento, y tomar decisiones estratégicas informadas que impulsen el éxito empresarial. Mediante el uso de técnicas de análisis de datos y herramientas como Power BI, se pretende transformar los datos en insights visuales claros y comprensibles, facilitando así la toma de decisiones informadas y la implementación de estrategias efectivas para el crecimiento y la rentabilidad del negocio.

3. Herramientas utilizadas

- Excel: Limpieza y lectura del dataset.
- SQL Server: Creación del diagrama Entidad – Relación.
- Power BI Desktop: Creación del tablero de control.

4.Hipótesis

En el análisis del Superstore, existen muchas hipótesis que se pueden formular e investigar. Las hipótesis son suposiciones iniciales que se pueden investigar con datos y sobre las cuales se puede afirmar si esa hipótesis es verdadera o falsa. A continuación, se muestran ejemplos de hipótesis que pueden ser relevantes en este caso:

- 1. *Variación estacional:*** La facturación del Superstore es cíclica y varía por cambio de estación. Por ejemplo, se podría asumir que la facturación aumenta significativamente durante las temporadas altas, como Navidad o después del Black Friday.
- 2. *Producto estrella:*** Existe un grupo de productos que aportan de manera desproporcionada al volumen de negocios total del Superstore. Por lo tanto, se podría averiguar si hay productos que representan la mayoría de los ingresos en comparación con otros.
- 3. *Segmentación de clientes:*** Diferentes segmentos de clientes pueden tener un impacto significativo en la facturación. Se podría investigar si ciertos grupos de clientes contribuyen más que otros a los ingresos del Superstore.
- 4. *Eficacia de promociones:*** Se podría hipotetizar que la implementación de promociones específicas, como descuentos en productos seleccionados, puede aumentar la facturación total al atraer a más clientes o alentar compras adicionales.

Estas son solo algunas ideas de hipótesis que podrían ser relevantes para un análisis de las operaciones del Superstore. Es importante recordar que las hipótesis deben ser respaldadas por datos y sometidas a pruebas estadísticas para determinar su validez.

5. Objetivo y Alcance del Proyecto

El objetivo de este proyecto es analizar los resultados de las estrategias comerciales y la rentabilidad de la tienda utilizando el Superstore Dataset, con el fin de proporcionar información valiosa para la toma de decisiones en futuras inversiones y estrategias de comercialización, impulsando el crecimiento y posicionamiento de la empresa en el mercado.

Este proyecto abarca la recopilación y limpieza de datos, el diseño de un modelo de datos en Power BI, y la creación de visualizaciones interactivas que muestren patrones de ventas y rentabilidad. Además, incluye la identificación de tendencias, el análisis de segmentación de clientes, y la implementación de métricas financieras clave. Finalmente, se desarrollarán informes y dashboards accesibles, y se ofrecerá capacitación y documentación para los usuarios finales.

El alcance del proyecto de Power BI centrado en el análisis del conjunto de datos del Superstore puede abarcar diversas etapas y objetivos claves para obtener una comprensión completa y valiosa de sus aspectos comerciales. A continuación, se describe un alcance típico para un proyecto de análisis de este tipo:

- 1. Definición de objetivos:** Establecer objetivos claros para el proyecto, como comprender las tendencias de ventas, analizar la rentabilidad de los productos y segmentar la base de clientes para estrategias de marketing más efectivas.
- 2. Recopilación de datos:** Obtener y preparar los datos del conjunto del Superstore, que incluirán información sobre ventas, productos, clientes, envíos y geografía.
- 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.
- 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 usuarios finales, recopilando comentarios y realizando ajustes si es necesario.

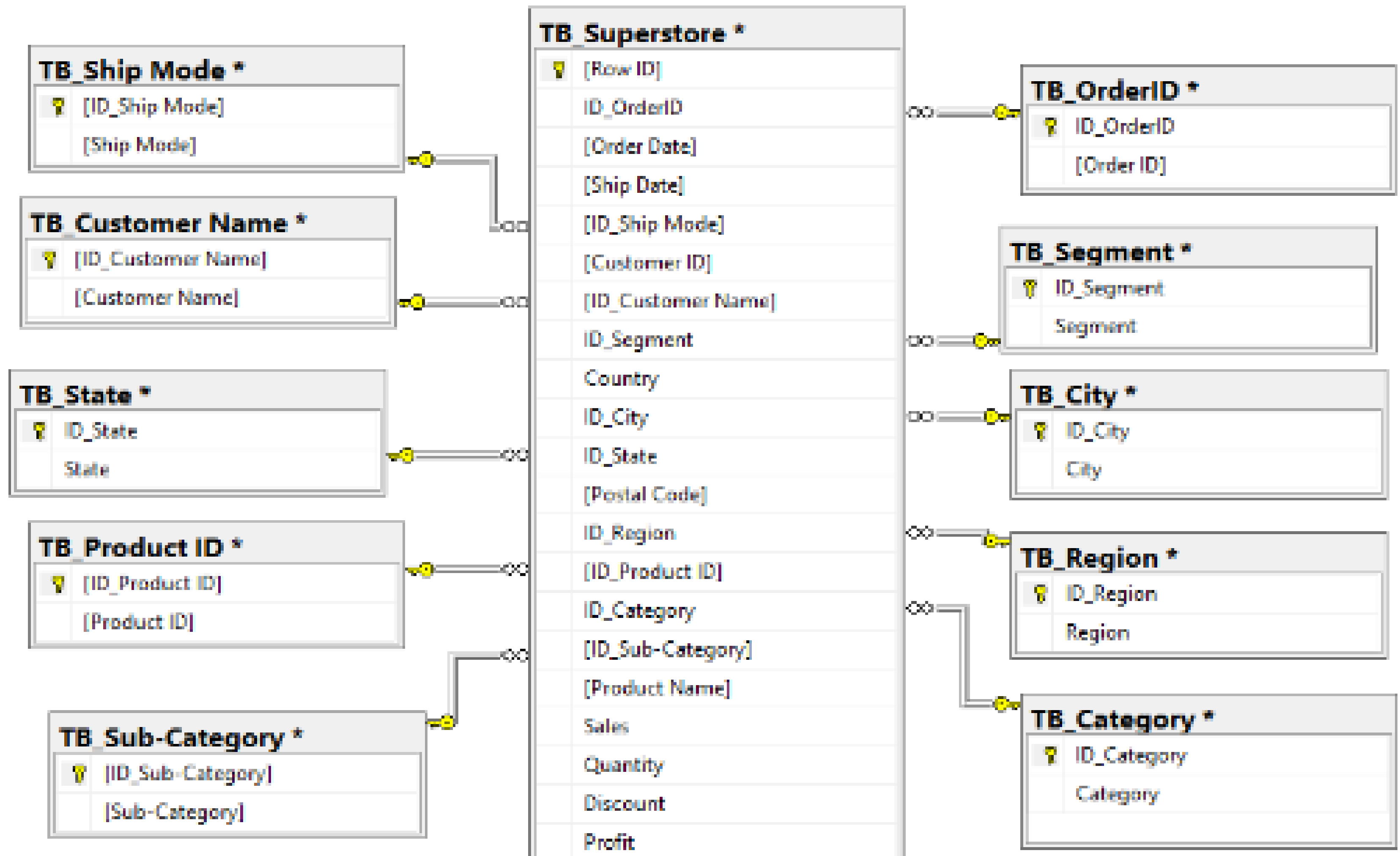
11. Mantenimiento continuo: Actualizar regularmente el tablero con nuevos datos para mantener la relevancia y precisión de las visualizaciones. El alcance específico del proyecto puede variar según las necesidades y metas de la organización.

6. Usuario Final y Nivel de Aplicación del Análisis

El usuario final de este análisis es el equipo de gestión de la empresa, incluyendo ejecutivos, gerentes de ventas, y analistas financieros. A nivel de aplicación, el análisis proporcionará insights detallados y prácticos que facilitarán la toma de decisiones estratégicas, tales como la optimización de estrategias de precios, la identificación de productos de alto rendimiento, y la mejora de la segmentación de clientes. Este enfoque permitirá a la empresa maximizar su rentabilidad y fortalecer su posición en el mercado.

7. Diagrama Entidad - Relación

A continuación, se detalla el diagrama Entidad – Relación creado:



8. Listado de tablas:

En el presente segmento se brinda una descripción de cada una de las tablas disponibles en la base de datos a trabajar, con una breve descripción de sus campos e identificando sus claves primarias y secundarias.

Tabla: TB_Superstore

Contiene las características principales de cada venta, identificando para cada número de operación, determinados caracteres de mercado, cómo fecha de venta, fecha de envío, tipo de envío, número de cliente, nombre de cliente, país de residencia del cliente, categoría, subcategoría, nombre de producto, etc.

Esta tabla almacena información detallada sobre cada venta realizada en el negocio minorista, proporcionando una visión completa de las transacciones y sus características asociadas para un análisis con datos relevantes sobre el desempeño del negocio

| | | |
|----|-----------------|--|
| PK | RowID | Identifica numéricamente a cada fila de la tabla superstore |
| FK | ID_Order ID | Identifica numéricamente a cada orden de cada cliente |
| FK | ID_Segment | El segmento al que pertenece el Cliente |
| FK | ID_City | Identifica numéricamente la ciudad de residencia del cliente |
| FK | ID_State | Identifica numéricamente el Estado de residencia del cliente |
| FK | ID_Región | Identifica numéricamente la región a la que pertenece cada cliente |
| FK | ID_Product ID | Identifica a cada producto |
| FK | ID_Category | Identifica numéricamente la categoría del producto solicitado. |
| FK | ID_Sub-category | Identifica numéricamente la categoría del producto solicitado. |
| FK | ID_Ship Mode | Modo de envío especificado por el cliente |
| | Order Date | Fecha de venta del producto |
| | Ship Date | Fecha de envío del producto |
| | Customer ID | Identifica únicamente a cada cliente |
| | Country | País de residencia del Cliente. |
| | Postal Code | Código postal de cada cliente |
| | Product Name | Identifica el nombre del producto |
| | Sales | Identifica numéricamente las ventas del producto |
| | Quantity | Identifica numéricamente las ventas del producto |
| | Discount | Identifica numéricamente los descuentos del producto |
| | Profit | Identifica numéricamente las ganancias o pérdidas del producto. |

Tabla: TB_State

Contiene los nombres de los Estados en los que viven los clientes y los identifica con un número.

| | | |
|----|----------|--|
| PK | ID_State | Identifica numéricamente a cada uno de los estados |
| | State | Nombre de los estados |

Tabla: TB_Region

Contiene el nombre de las regiones en las que viven cada uno de los clientes y los identifica con un número.

| | | |
|----|-----------|---------------------------------------|
| PK | ID Región | Identifica numéricamente las regiones |
| | Región | Nombre de las regiones |

Tabla: TB_Category

Identifica numéricamente a cada categoría del producto y muestra sus nombres

| | | |
|----|-------------|--|
| PK | ID_Category | Identifica numéricamente cada categoría del producto |
| | Category | Nombres de las categorías |

Tabla: TB_Subcategory

Identifica cada subcategoría numéricamente y contiene el nombre de cada uno.

| | | |
|----|-----------------|---|
| PK | ID_Sub-category | Identifica numéricamente cada subcategoría del producto |
| | Sub-Category | Nombres de las subcategorías |

Tabla: TB_City

Identifica numéricamente cada ciudad en la que viven los clientes y en qué ciudades se realizaron las ventas.

| | | |
|----|---------|--|
| PK | ID_City | Identifica numéricamente cada ciudad de los clientes |
| | City | Nombre de las ciudades de cada cliente |

Tabla: TB_Segment

Identifica numéricamente a qué segmento pertenece cada cliente.

| | | |
|----|------------|--|
| PK | ID_Segment | Identifica numéricamente el nombre de cada uno de los clientes |
| | Segment | Nombre de los clientes |

Tabla: TB_Ship Mode

Identifica numéricamente el modo de envío para cada cliente, elegido por el mismo.

| | | |
|----|--------------|---|
| PK | ID_Ship Mode | Identifica numéricamente el modo de envío para cada cliente |
| | Ship Mode | Modo de envío |

Tabla: TB_Order ID

Contiene el número de orden de cada pedido de cada cliente.

| | | |
|----|------------|---|
| PK | ID_OrderID | Identifica numéricamente a cada orden de cada cliente |
| | Order ID | Número de orden |

Tabla: TB_Product ID

Contiene todos los códigos de cada producto y los identifica con un número.

| | | |
|----|---------------|---|
| PK | ID_Product ID | Identifica numéricamente el código de cada producto |
| | Product ID | Código de cada producto |

9. Listado de las columnas de cada tabla:

Se detalla a continuación cada tabla con sus campos, tipo de campo, claves primarias y secundarias.

Tabla: TB_Superstore:

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE | RELACIÓN | |
|------------------|---------------|---------------|------------|------------------------------------|
| RowID | int | PRIMARY KEY | | |
| ID_Order ID | int | FOREIGN KEY | REFERENCES | TB_OrderID(ID_OrderID) |
| ID_Customer Name | int | FOREIGN KEY | REFERENCES | TB_Customer Name(ID_Customer Name) |
| ID_Segment | int | FOREIGN KEY | REFERENCES | TB_Segment(ID_Segment) |
| ID_City | int | FOREIGN KEY | REFERENCES | TB_City(ID_City) |
| ID_State | int | FOREIGN KEY | REFERENCES | TB_State(ID_State) |
| ID_Region | int | FOREIGN KEY | REFERENCES | TB_Region(ID_Region) |
| ID_Product ID | int | FOREIGN KEY | REFERENCES | TB_Product ID(ID_Product ID) |
| ID_Category | int | FOREIGN KEY | REFERENCES | TB_Category(ID_Category) |
| ID_Sub-Category | int | FOREIGN KEY | REFERENCES | TB_Sub-Category(ID_Sub-Category) |
| ID_Ship Mode | int | FOREIGN KEY | REFERENCES | TB_Ship Mode(ID_Ship Mode) |
| Order Date | datetime | | | |
| Ship Date | datetime | | | |
| Customer ID | varchar(50) | | | |
| Country | text | | | |
| Postal Code | int | | | |
| Product Name | varchar(80) | | | |
| Sales | decimal(18,2) | | | |
| Quantity | int | | | |
| Discount | decimal(18,2) | | | |
| Profit | decimal(18,2) | | | |

Tabla: TB_ProductID

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|--------------|---------------|---------------|
| ID_ProductID | int | PRIMARY KEY |
| ProductID | varchar(50) | |

Tabla: TB_OrderID

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|------------|---------------|---------------|
| ID_OrderID | int | PRIMARY KEY |
| OrderID | varchar(50) | |

Tabla: TB_State

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|----------|---------------|---------------|
| ID_State | int | PRIMARY KEY |
| State | text | |

Tabla: TB_Region

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|-----------|---------------|---------------|
| ID_Region | int | PRIMARY KEY |
| Region | text | |

Tabla: TB_Category

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|-------------|---------------|---------------|
| ID_Category | int | PRIMARY KEY |
| Category | text | |

Tabla: TB_Subcategory

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|-----------------|---------------|---------------|
| ID_Sub-Category | int | PRIMARY KEY |
| Sub-Category | text | |

Tabla: TB_City

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|---------|---------------|---------------|
| ID_City | int | PRIMARY KEY |
| City | text | |

Tabla: TB_Segment

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|------------|---------------|---------------|
| ID_Segment | int | PRIMARY KEY |
| Segment | text | |

Tabla: TB_Ship Mode

| CAMPO | TIPO DE CAMPO | TIPO DE CLAVE |
|--------------|---------------|---------------|
| ID_Ship Mode | int | PRIMARY KEY |
| Ship Mode | text | |

10.Conexión de Base de Datos y Tablas

Se comenzó realizando la conexión a la base de datos por medio del archivo Excel (que previamente fue normalizado desde excel).

Tabla TB_Superstore:

Se procedió a estandarizar y configurar correctamente los datos. Eliminar duplicados. Agregar, reordenar y renombrar columnas.

TB_Superstore

Opciones de presentación ▾

```
1 let
2   Origen = Excel.Workbook(File.Contents("C:\Users\ victo\OneDrive\Escritorio\super.xlsx"), null, true),
3   TB_Superstore_Sheet = Origen[Item="TB_Superstore",Kind="Sheet"] [Data],
4   #"Encabezados promovidos" = Table.PromoteHeaders(TB_Superstore_Sheet, [PromoteAllScalars=true]),
5   #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"Row ID", Int64.Type}, {"ID_OrderID", Int64.Type}, {"Order Date", type date}, {"Ship Date",
   type date}, {"ID_Ship Mode", Int64.Type}, {"Customer ID", type text}, {"Customer Name", type text}, {"ID_Segment", Int64.Type}, {"Country", type text}, {"ID_City",
   Int64.Type}, {"ID_State", Int64.Type}, {"Postal Code", Int64.Type}, {"ID_Region", Int64.Type}, {"ID_Product ID", Int64.Type}, {"ID_Category", Int64.Type},
   {"ID_Sub-Category", Int64.Type}, {"Product Name", type text}, {"Sales", type any}, {"Quantity", Int64.Type}, {"Discount", Percentage.Type}, {"Profit", type any}}),
6   #"Agregar columna personalizad ""sales""a" = Table.AddColumn(#"Tipo cambiado", "Personalizado", each let splitSales = Splitter.SplitTextByDelimiter(".", QuoteStyle.None)
   (Text.From([Sales], "es-MX")) in Text.Combine({Text.Start(splitSales{0}?, 2), Text.Reverse(Text.Middle(Text.Reverse(splitSales{0}?), 2))}), type text),
7   #"Columnas reordenadas ""sales"" nuevo" = Table.ReorderColumns(#"Agregar columna personalizad ""sales""a",{ "Row ID", "ID_OrderID", "Order Date", "Ship Date", "ID_Ship
   Mode", "Customer ID", "Customer Name", "ID_Segment", "Country", "ID_City", "ID_State", "Postal Code", "ID_Region", "ID_Product ID", "ID_Category", "ID_Sub-Category",
   "Product Name", "Sales", "Personalizado", "Quantity", "Discount", "Profit"}),
8   #"Columnas quitadas ""sales"" viejo" = Table.RemoveColumns(#"Columnas reordenadas ""sales"" nuevo",{ "Sales"}),
9   #"Columnas con nombre cambiado ""sales"" nuevo" = Table.RenameColumns(#"Columnas quitadas ""sales"" viejo",{{"Personalizado", "Sales"}}),
10  #"Tipo cambiado1 ""sales"" a ""entero"" = Table.TransformColumnTypes(#"Columnas con nombre cambiado ""sales"" nuevo",{{"Sales", Int64.Type}}),
11  #"Agregar columna personalizada" = Table.AddColumn(#"Tipo cambiado1 ""sales"" a ""entero""", "Personalizado", each let splitProfit = Splitter.SplitTextByDelimiter("-",
   QuoteStyle.None)(Text.From([Profit], "es-MX")) in Text.Combine(List.Transform(splitProfit, each Text.Start(_, 1)), "-"), type text),
12  #"Tipo cambiado1 ""int"" a ""%"" = Table.TransformColumnTypes(#"Agregar columna personalizada",{{"Personalizado", Percentage.Type}}),
13  #"Columnas quitadas ""profit"" viejo" = Table.RemoveColumns(#"Tipo cambiado1 ""int"" a ""%""",{ "Profit"}),
14  #"Columnas con nombre cambiado ""profit"" el duplicado" = Table.RenameColumns(#"Columnas quitadas ""profit"" viejo",{{"Personalizado", "Profit"}}),
15  #"Valor reemplazado" = Table.ReplaceValue(#"Columnas con nombre cambiado ""profit"" el duplicado",#date(2018, 1, 1),#date(2014, 1, 1),Replacer.ReplaceValue,{"Order Date"}
   )
16 in
17  #"Valor reemplazado"
```

Tabla TB_Sub-Category:

TB_Sub-Category

```
1 let
2     Origen = Excel.Workbook(File.Contents("C:\Users\Victo\OneDrive\Escritorio\super.xlsx"), null, true),
3     #"TB_Sub-Category_Sheet" = Origen[[Item="TB_Sub-Category",Kind="Sheet"]][Data],
4     #"Encabezados promovidos" = Table.PromoteHeaders(#"TB_Sub-Category_Sheet", [PromoteAllScalars=true]),
5     #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",({"ID_Sub-Category", Int64.Type}, {"Sub-Category", type text}))
6 in
7     #"Tipo cambiado"
```

Tabla TB_State:

TB_State

```
1 let
2     Origen = Excel.Workbook(File.Contents("C:\Users\Victo\OneDrive\Escritorio\super.xlsx"), null, true),
3     TB_State_Sheet = Origen[[Item="TB_State",Kind="Sheet"]][Data],
4     #"Encabezados promovidos" = Table.PromoteHeaders(TB_State_Sheet, [PromoteAllScalars=true]),
5     #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",({"ID_State", Int64.Type}, {"State", type text}))
6 in
7     #"Tipo cambiado"
```


Tabla TB_Ship Mode:

TB_Ship Mode

```
1 let
2   Origen = Excel.Workbook(File.Contents("C:\Users\victo\OneDrive\Escritorio\super.xlsx"), null, true),
3   #"TB_Ship Mode_Sheet" = Origen{[Item="TB_Ship Mode",Kind="Sheet"]}[Data],
4   #"Encabezados promovidos" = Table.PromoteHeaders(#"TB_Ship Mode_Sheet", [PromoteAllScalars=true]),
5   #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Ship Mode", Int64.Type}, {"Ship Mode", type text}})
6 in
7   #"Tipo cambiado"
```

Tabla TB_Segment:

TB_Segment

```
1 let
2   Origen = Excel.Workbook(File.Contents("C:\Users\victo\OneDrive\Escritorio\super.xlsx"), null, true),
3   TB_Segment_Sheet = Origen{[Item="TB_Segment",Kind="Sheet"]}[Data],
4   #"Encabezados promovidos" = Table.PromoteHeaders(TB_Segment_Sheet, [PromoteAllScalars=true]),
5   #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Segment", Int64.Type}, {"Segment", type text}})
6 in
7   #"Tipo cambiado"
```

Tabla TB_Region:

TB_Region

```
1 let
2     Origen = Excel.Workbook(File.Contents("C:\Users\victo\OneDrive\Escritorio\super.xlsx"), null, true),
3     TB_Region_Sheet = Origen[[Item="TB_Region",Kind="Sheet"]][Data],
4     #"Encabezados promovidos" = Table.PromoteHeaders(TB_Region_Sheet, [PromoteAllScalars=true]),
5     #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Region", Int64.Type}, {"Region", type text}})
6 in
7     #"Tipo cambiado"
```

Tabla TB_Product ID:

TB_Product ID

```
1 let
2     Origen = Excel.Workbook(File.Contents("C:\Users\victo\OneDrive\Escritorio\super.xlsx"), null, true),
3     #"TB_Product ID_Sheet" = Origen[[Item="TB_Product ID",Kind="Sheet"]][Data],
4     #"Encabezados promovidos" = Table.PromoteHeaders(#"TB_Product ID_Sheet", [PromoteAllScalars=true]),
5     #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Product ID", Int64.Type}, {"Product ID", type text}})
6 in
7     #"Tipo cambiado"
```

Tabla TB_Order ID:

TB_Order ID

```
1 let
2     Origen = Excel.Workbook(File.Contents("C:\Users\victo\OneDrive\Escritorio\super.xlsx"), null, true),
3     #"TB_Order ID_Sheet" = Origen[[Item="TB_Order ID",Kind="Sheet"]][Data],
4     #"Encabezados promovidos" = Table.PromoteHeaders(#"TB_Order ID_Sheet", [PromoteAllScalars=true]),
5     #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_OrderID", Int64.Type}, {"Order ID", type text}})
6 in
7     #"Tipo cambiado"
```

Tabla TB_City

TB_City

```
1 let
2     Origen = Excel.Workbook(File.Contents("C:\Users\victo\OneDrive\Escritorio\super.xlsx"), null, true),
3     TB_City_Sheet = Origen[[Item="TB_City",Kind="Sheet"]][Data],
4     #"Encabezados promovidos" = Table.PromoteHeaders(TB_City_Sheet, [PromoteAllScalars=true]),
5     #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_City", Int64.Type}, {"City", type text}})
6 in
7     #"Tipo cambiado"
```

Tabla TB_Category:

TB_Category

```
1 let
2   Origen = Excel.Workbook(File.Contents("C:\Users\victo\OneDrive\Escritorio\super.xlsx"), null, true),
3   TB_Category_Sheet = Origen{[Item="TB_Category",Kind="Sheet"]}[Data],
4   #"Encabezados promovidos" = Table.PromoteHeaders(TB_Category_Sheet, [PromoteAllScalars=true]),
5   #"Tipo cambiado" = Table.TransformColumnTypes(#"Encabezados promovidos",{{"ID_Category", Int64.Type}, {"Category", type text}})
6 in
7   #"Tipo cambiado"
```

Solo se realizaron cambios en la tabla TB_Superstore, ya que las demás tablas fueron corregidas en la normalización que se realizó mediante Excel.

Tabla Calendario:

La tabla calendario se realizó mediante las siguientes funciones:

Tabla Calendario = CALENDAR(MIN(TB_Superstore[Order Date]), MAX(TB_Superstore[Ship Date]))

Año = YEAR('Tabla Calendario'[Date])

Mes Corto = FORMAT('Tabla Calendario'[Date], "mmm")

Nombre del Mes = FORMAT('Tabla Calendario'[Date], "mmmm")

Número Mes = MONTH('Tabla Calendario'[Date])

1 Tabla Calendario = CALENDAR(MIN(TB_Superstore[Order Date]), MAX(TB_Superstore[Ship Date]))

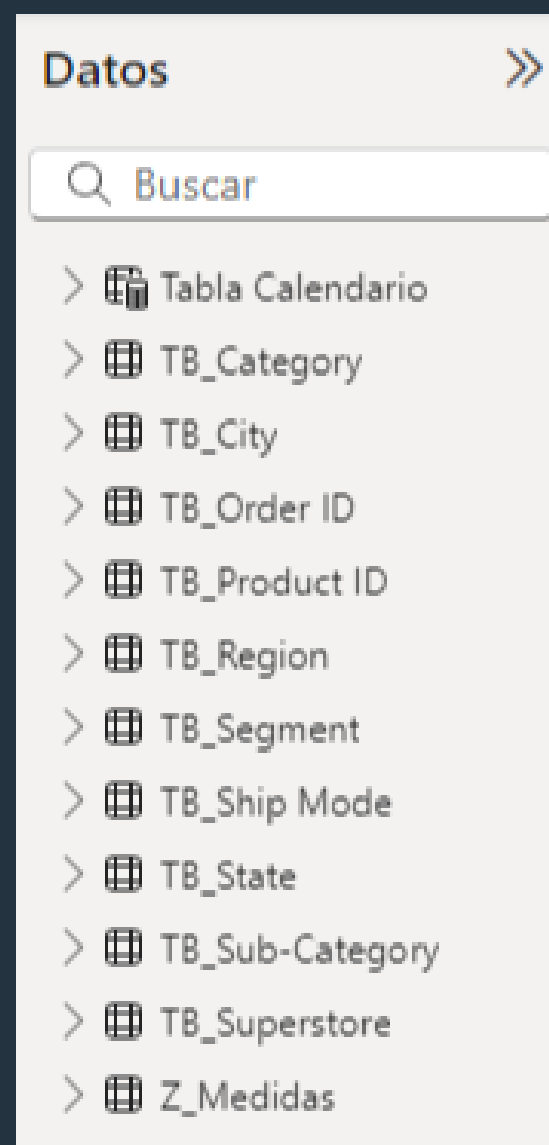
| Date | Año | Numero Mes | Nombre del Mes | Mes Corto |
|---------------------------|------|------------|----------------|-----------|
| 01/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 02/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 03/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 04/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 05/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 06/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 07/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 08/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 09/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 10/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 11/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 12/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 13/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 14/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 15/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 16/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 17/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 18/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 19/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 20/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |
| 21/01/2014 12:00:00 a. m. | 2014 | 1 | enero | ene |

Tabla Calendario (1,480 filas)

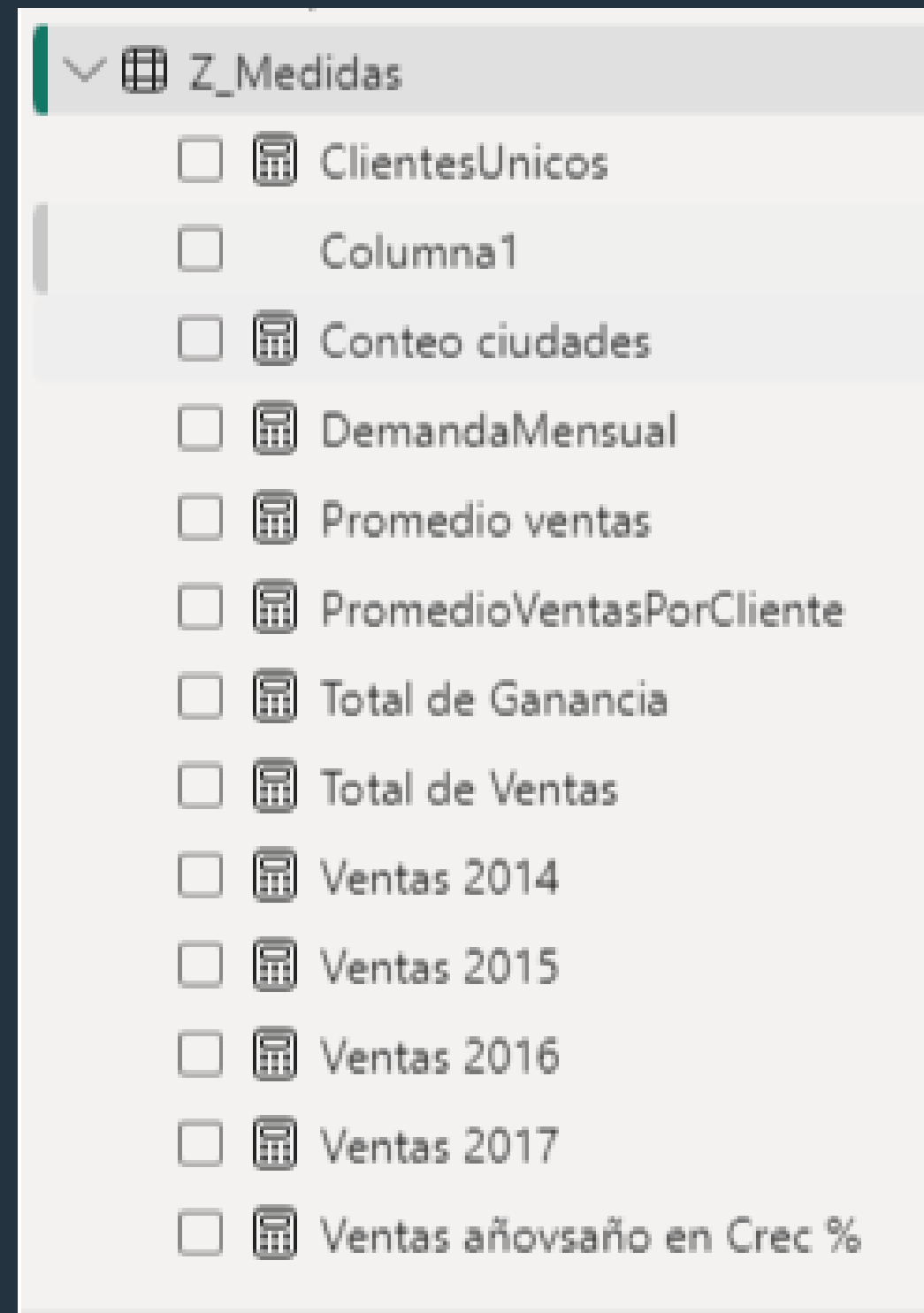
11. Medidas Calculadas

Luego de contar con todas las tablas de información del proyecto, configuradas y listas para el análisis de dichos datos, se procedió a crear las Medidas Calculadas con el fin de tener indicadores que ayuden al correcto análisis y conclusiones del proyecto.

Los datos que se utilizaron fueron los siguientes:



Se creó una tabla Medidas, donde se usaron funciones DAX para determinar las ventas respectivas de cada año, las ganancias, los promedios, entre otras cosas:



Las funciones DAX fueron las siguientes:

- ClientesUnicos = DISTINCTCOUNT('TB_Superstore'[Customer Name])

Clientes Únicos = Me da la cantidad exacta de clientes sin repetirlos

- Conteo ciudades = DISTINCTCOUNT(TB_City[City])

Conteo ciudades = Me da la cantidad exacta de las ciudades sin repetirlos

- DemandaMensual = CALCULATE(
SUM('TB_Superstore'[Quantity]),
DATESINPERIOD('TB_Superstore'[Order Date], MAX('TB_Superstore'[Order Date]), -12, MONTH))

Me devuelve la suma de la cantidad de productos vendidos en un período de 12 meses anteriores a la fecha máxima de la fecha de orden.

- Promedio ventas = AVERAGE(TB_Superstore[Sales])

Promedio ventas = Muestra el promedio de ventas de todos los artículos

- Total de Ganancia = SUM(TB_Superstore[Profit])

Me indica la ganancia que dejan los artículos.

- Total de Ventas = SUM(TB_Superstore[Sales])

Me calcula el total de las ventas.

- Ventas 2014 = CALCULATE(SUM(TB_Superstore[Sales]), 'Tabla Calendario'[Año] = 2014)

Me devuelve las ventas que hubo en el año 2014.

- Ventas 2015 = CALCULATE(SUM(TB_Superstore[Sales]), 'Tabla Calendario'[Año] = 2015)

Me devuelve las ventas que hubo en el año 2015.

- Ventas 2016 = CALCULATE(SUM(TB_Superstore[Sales]), 'Tabla Calendario'[Año] = 2016)

Me devuelve las ventas que hubo en el año 2016.

- Ventas 2017 = CALCULATE(SUM(TB_Superstore[Sales]), 'Tabla Calendario'[Año] = 2017)

Me devuelve las ventas que hubo en el año 2017.

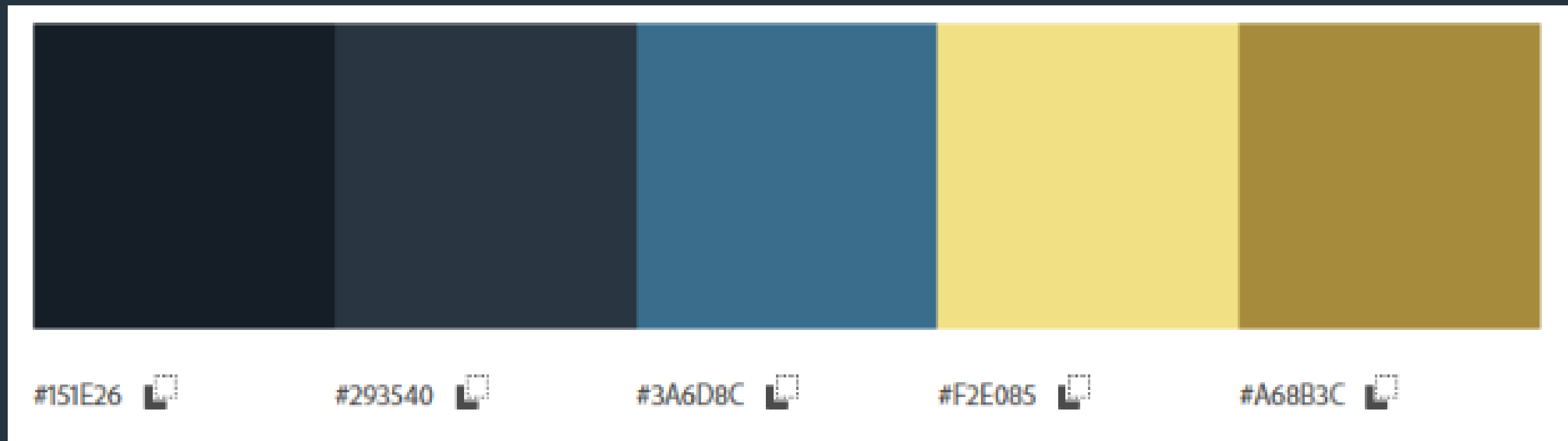
- Ventas año vs año en Crec % =

$$\text{DIVIDE}([\text{Total de Ventas}] - \text{CALCULATE}([\text{Total de Ventas}], \text{PARALLELPERIOD}('Tabla Calendario'[\text{Date}], -12, \text{MONTH})), \text{CALCULATE}([\text{Total de Ventas}], \text{PARALLELPERIOD}('Tabla Calendario'[\text{Date}], -12, \text{MONTH})))$$

Me calcula el porcentaje de crecimiento o decrecimiento de las ventas de un mes comparado con el mismo mes del año anterior.

12. Diseño del tablero

La paleta de colores utilizada es personalizada y se creó en la aplicación Colors:



13. Dashboard Power BI

Portada:



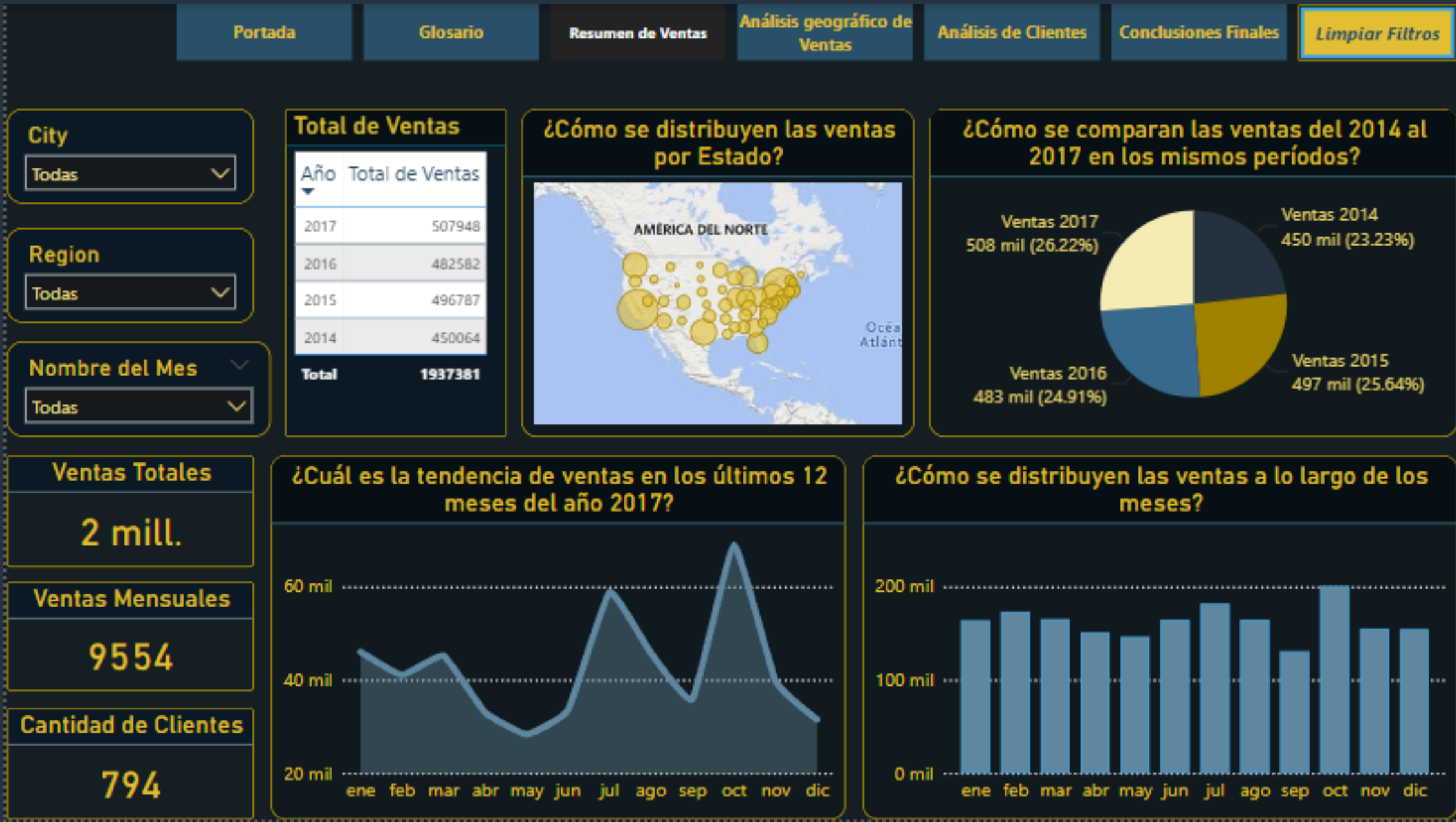
Se realizó una solapa de “Portada” con el fin de mejorar la navegabilidad, facilitando la comprensión del informe y permitiendo a los usuarios interactuar de manera más efectiva con los datos de la tienda “Superstore”.

Glosario:

| | | | | | | |
|--------------|---|----------|-------------------|-------------------------------|----------------------|----------------------|
| | Portada | Glosario | Resumen de Ventas | Análisis geográfico de Ventas | Análisis de Clientes | Conclusiones Finales |
| Glosario | | | | | | |
| RowID | Identifica numéricamente a cada fila de la tabla Superstore | | | | | |
| Order | ID Identifica numéricamente a cada orden de cada cliente | | | | | |
| Segment | El segmento al que pertenece el Cliente | | | | | |
| City | Identifica la ciudad de residencia del cliente | | | | | |
| State | Identifica el Estado de residencia del cliente | | | | | |
| Region | Identifica la región a la que pertenece cada cliente | | | | | |
| Product ID | Identifica a cada producto | | | | | |
| Category | Identifica la categoría del producto solicitado. | | | | | |
| Sub-category | Identifica la categoría del producto solicitado. | | | | | |
| Ship Mode | Modo de envío especificado por el cliente | | | | | |
| Order Date | Fecha de venta del producto | | | | | |
| Ship Date | Fecha de envío del producto | | | | | |
| Customer ID | Identifica únicamente a cada cliente | | | | | |
| Country | País de residencia del Cliente. | | | | | |
| Postal Code | Código postal de cada cliente | | | | | |
| Product Name | Identifica el nombre del producto | | | | | |
| Sales | Identifica numéricamente las ventas del producto | | | | | |
| Quantity | Identifica numéricamente las ventas del producto | | | | | |
| Discount | Identifica numéricamente los descuentos del producto | | | | | |
| Profit | Identifica numéricamente las ganancias o pérdidas del producto. | | | | | |

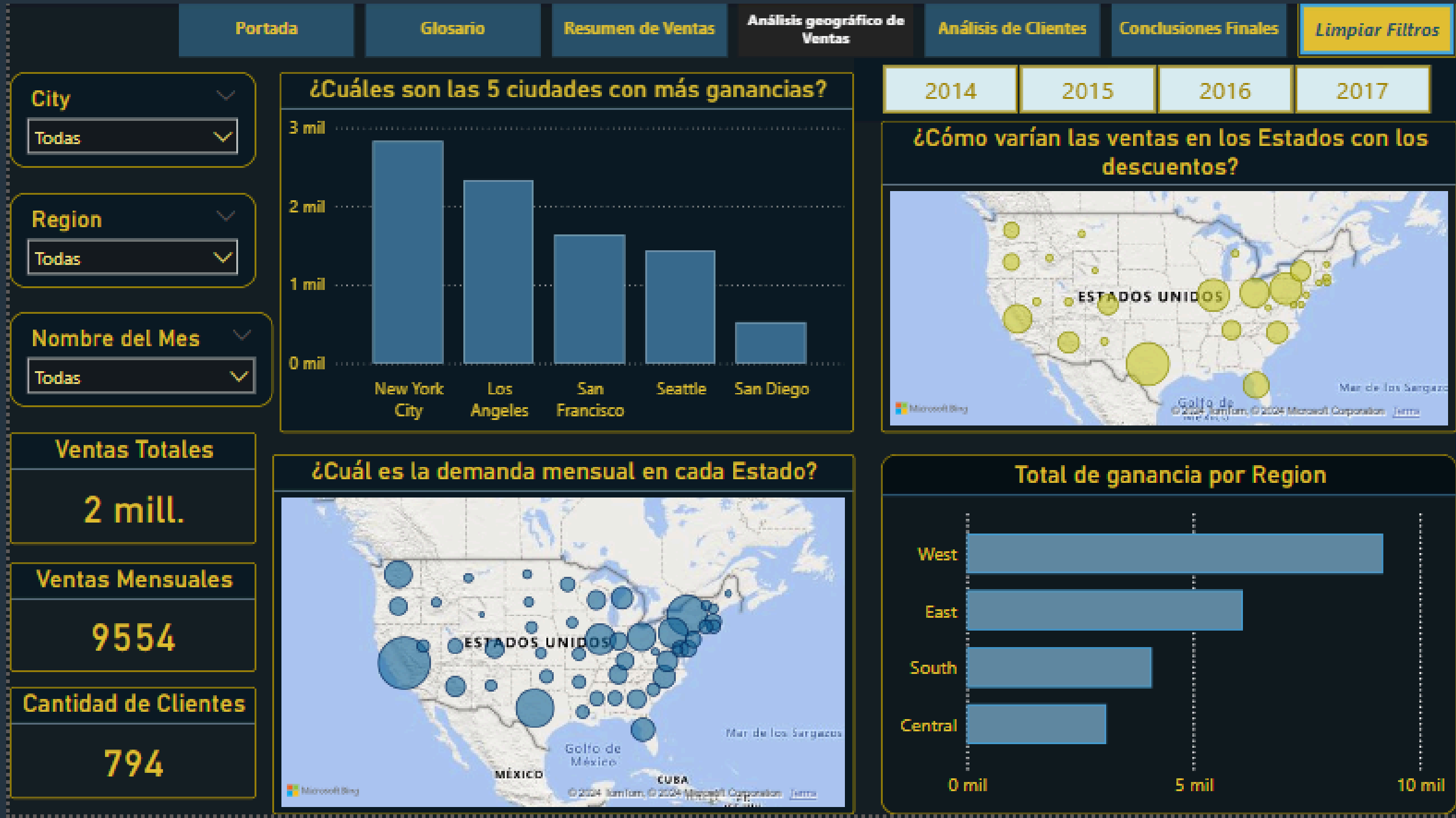
La solapa de “Glosario” explica que se va a encontrar en cada una de las columnas.

Resumen de Ventas:



La solapa de “Ventas” ofrece una visualización interactiva y detallada de las ventas a lo largo del tiempo y en diferentes regiones. Permite filtrar datos por ciudad, región y mes, lo que facilita un análisis personalizado. Se muestran las ventas totales por año (2014-2017), un mapa interactivo de distribución por estado, una comparación de ventas anuales y gráficos que presentan tendencias y distribuciones mensuales. También se destacan las ventas totales, el promedio mensual de ventas y la cantidad de clientes.

Análisis Geográfico de Ventas:

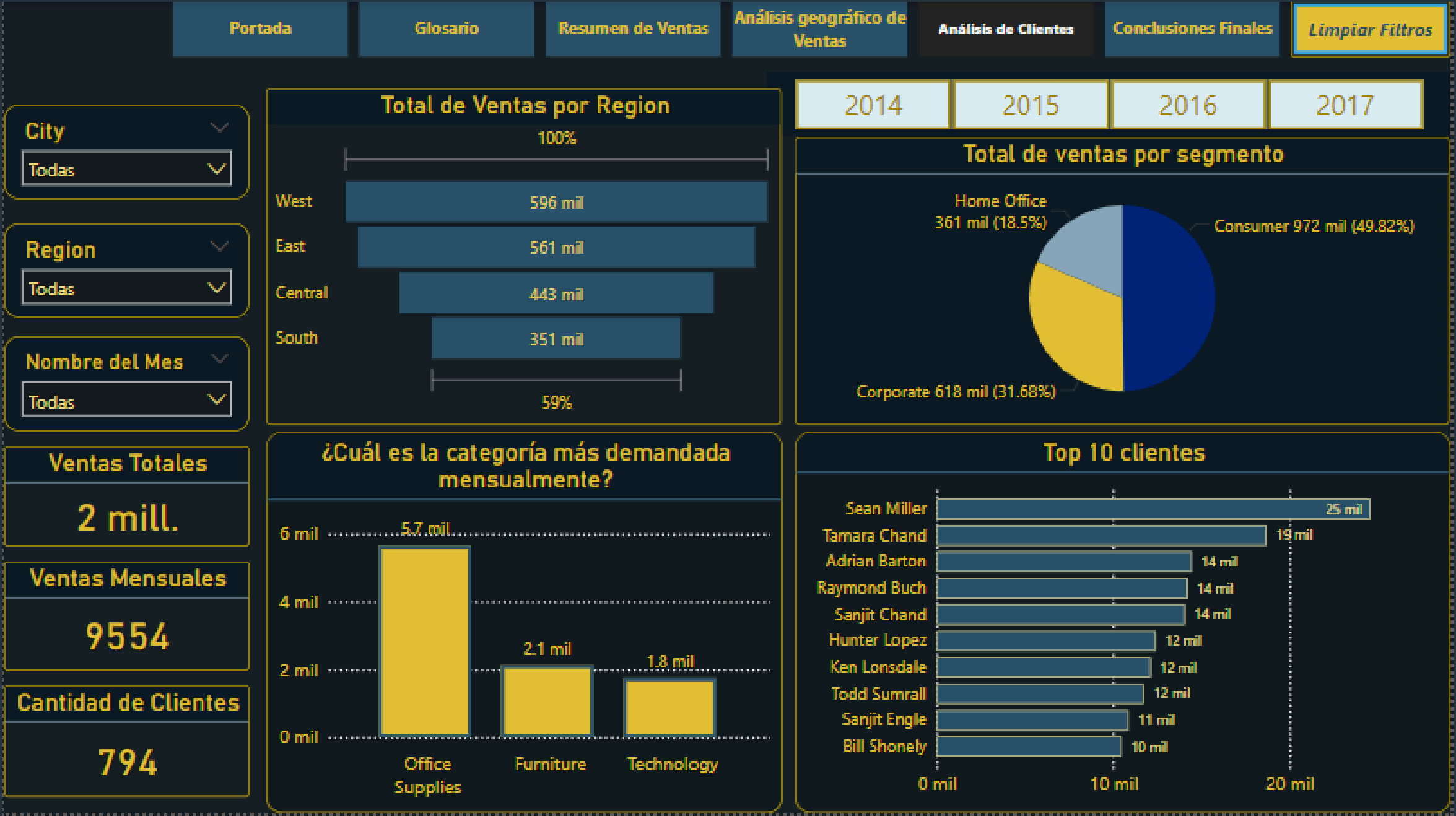


La solapa de “Análisis Geográfico de Ventas” proporciona una visión interactiva de las ganancias y la demanda mensual en diferentes regiones y ciudades. Permite filtrar los datos por ciudad, región y mes para obtener un análisis personalizado.

Se destacan las cinco ciudades con más ganancias, mostrando que New York City, Los Angeles, San Francisco, Seattle

y San Diego son las principales. Un mapa interactivo muestra cómo varían las ventas en los estados con los descuentos, y otro mapa indica la demanda mensual en cada estado, visualizando el tamaño de la demanda con diferentes tamaños de burbujas. Además, un gráfico de barras presenta el total de ganancias por región, revelando que la región Oeste es la más lucrativa.

Análisis de Clientes:



La solapa “Análisis de Clientes” proporciona una visión interactiva y detallada de las ventas por región, segmento, categoría y clientes. Permite filtrar los datos por ciudad, región y mes, facilitando un análisis personalizado.

Se muestra el total de ventas por región, destacando que la región Oeste tiene las mayores ventas con 596 mil unidades, seguida por la región Este, Central y Sur. Un gráfico de tortas desglosa el total de ventas por segmento, indicando que el segmento de consumidores representa el 49.82% de las ventas, seguido por el segmento corporativo con 31.68% y el segmento de oficinas en casa con 18.5%.

Además, un gráfico de barras muestra las categorías más demandadas mensualmente, donde "Office Supplies" es la categoría más popular con 5.7 mil unidades vendidas, seguida por muebles y tecnología. También se presentan los 10 principales clientes, con Sean Mile siendo el mayor cliente con 23 mil unidades adquiridas.

Conclusiones Finales:

Portada

Glosario

Resumen de Ventas

Análisis geográfico de Ventas

Análisis de Clientes

Conclusiones Finales

Según las Hipótesis previstas al iniciar el Proyecto, a través de todo lo analizado se pudo llegar a las siguientes Conclusiones:

1.Variación Estacional:

Varía mucho por año , pero se destacan picos de ventas por los meses de Julio y de Octubre

2.Producto Estrella:

El producto estrella varía mucho por zona año y mes, pero entre todos los productos que se ofrecen, se destaca una alta demanda por los "Chairs", "Phones", "Machines" , "Copiers" y "Storage" siendo los que más ventas tienen todos los años.

3.Segmentación de Clientes:

Por año se encuentran clientes que se destacan notoriamente por sus compras, encontrando clientes que superan las 22.3 compras por año.

4.Eficacia de Promociones:

Se encuentran zonas con más ganancias que otras, destacándose la región Oeste (West).

En la solapa “Conclusiones Finales” se detallan las conclusiones sacadas.

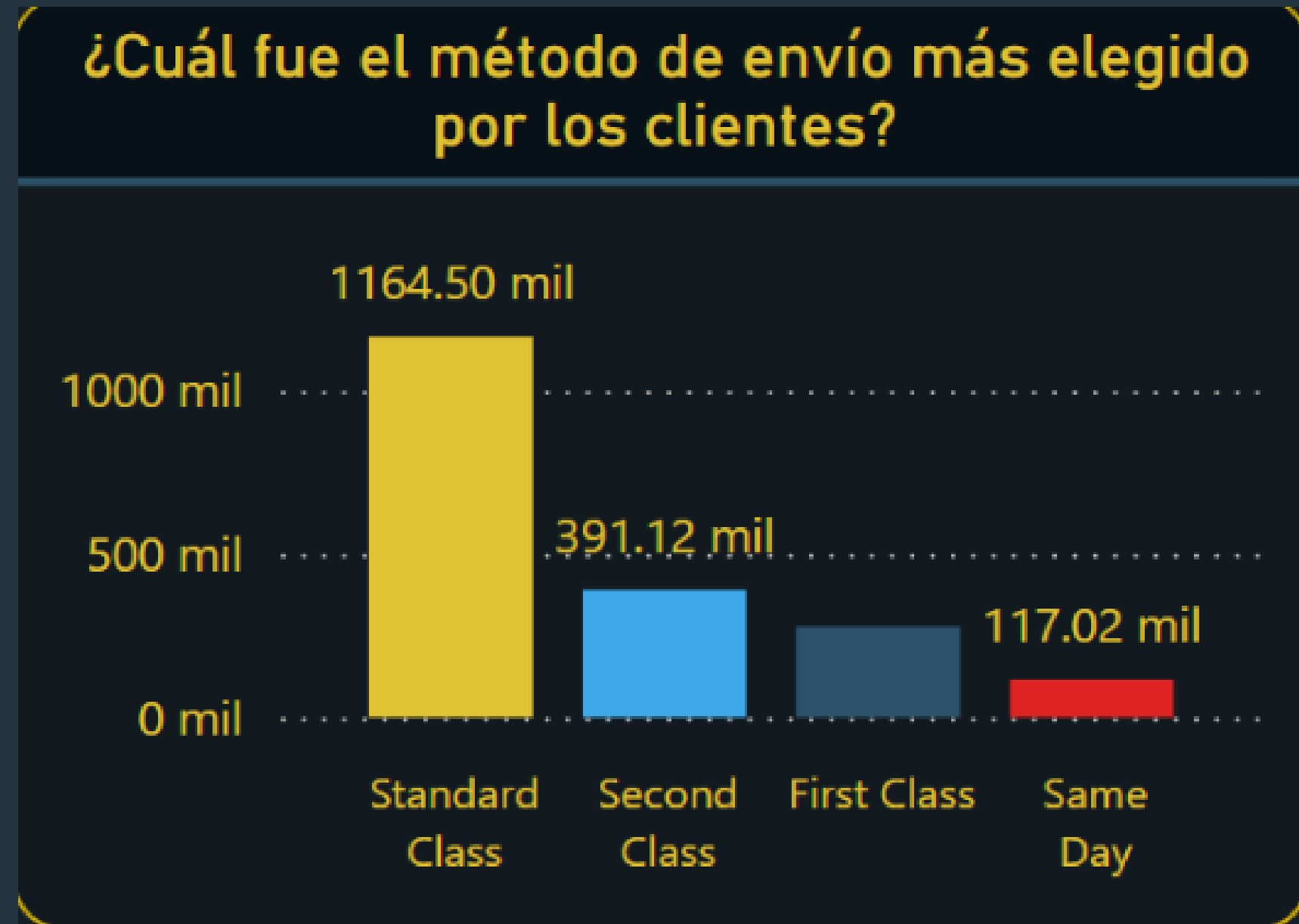
ToolTip 1 ventas:



El uso de ToolTips en la solapa de "ToolTip 1 ventas" permite una interacción adicional y un análisis más detallado. Específicamente, se implementó un ToolTip que muestra la categoría más vendida en cada estado. Cuando los usuarios pasan el cursor sobre un estado en el informe, se despliega una ventana emergente que revela la categoría dominante en términos de ventas.

En este caso, el gráfico circular muestra que la categoría "Technology" representa el 48.53% de las ventas, "Furniture" el 37.67%, y "Office Supplies" el 13.80%. Esta información brinda una perspectiva rápida y fácil de las preferencias de los consumidores en cada estado, lo que permite a los usuarios identificar tendencias de demanda y optimizar las estrategias de ventas según la categoría de producto en diferentes regiones.

ToolTip 2 ventas:



El uso de ToolTips en la solapa de "ToolTip 2 ventas" permite una interacción adicional y un análisis más detallado. Específicamente, se implementó un ToolTip que muestra el método de envío más elegido por los clientes. Cuando los usuarios pasan el cursor sobre un método de envío en el informe, se despliega una ventana emergente que revela la popularidad de cada opción de envío.

En este caso, el gráfico de barras muestra que "Standard Class" es el método de envío más elegido con 1,164.50 mil unidades, seguido por "Second Class" con 391.12 mil unidades, "First Class" con 117.02 mil unidades y "Same Day"

con 117.02 mil unidades. Esta información brinda una perspectiva rápida y fácil sobre las preferencias de los clientes en cuanto a los métodos de envío, permitiendo a los usuarios identificar tendencias y optimizar las opciones de envío ofrecidas según las preferencias del cliente.

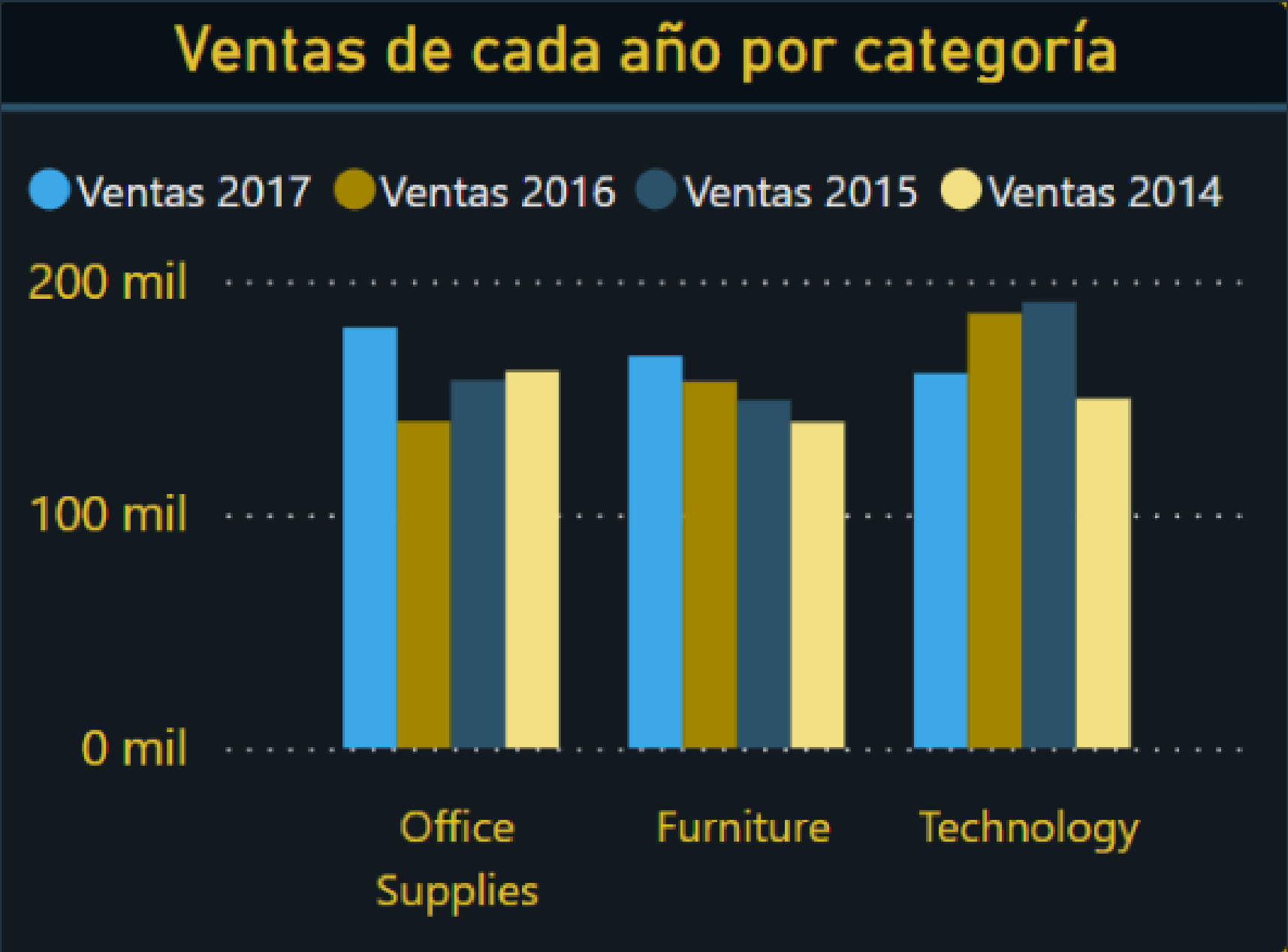
ToolTip 3 ventas:



El uso de ToolTips en la solapa “ToolTip 3 ventas” presenta un gráfico de barras invertido que muestra las subcategorías de productos más vendidas por mes. La visualización permite analizar y comparar fácilmente las ventas mensuales de diferentes sub-categorías, ordenadas de mayor a menor. La barra superior representa la subcategoría con mayores ventas, mientras que la barra inferior muestra la subcategoría con menores ventas. Este tipo de gráfico es útil para identificar cuáles son los productos más populares y cuáles tienen menor demanda.

Además, el gráfico proporciona valores específicos de ventas en miles de unidades para algunas sub-categorías clave, como "Tables" con 163 mil unidades vendidas y "Binders" con 156 mil unidades vendidas. Esta información detallada permite a los usuarios hacer un análisis más profundo sobre el rendimiento de cada subcategoría, facilitando la toma de decisiones estratégicas en términos de inventario o marketing.

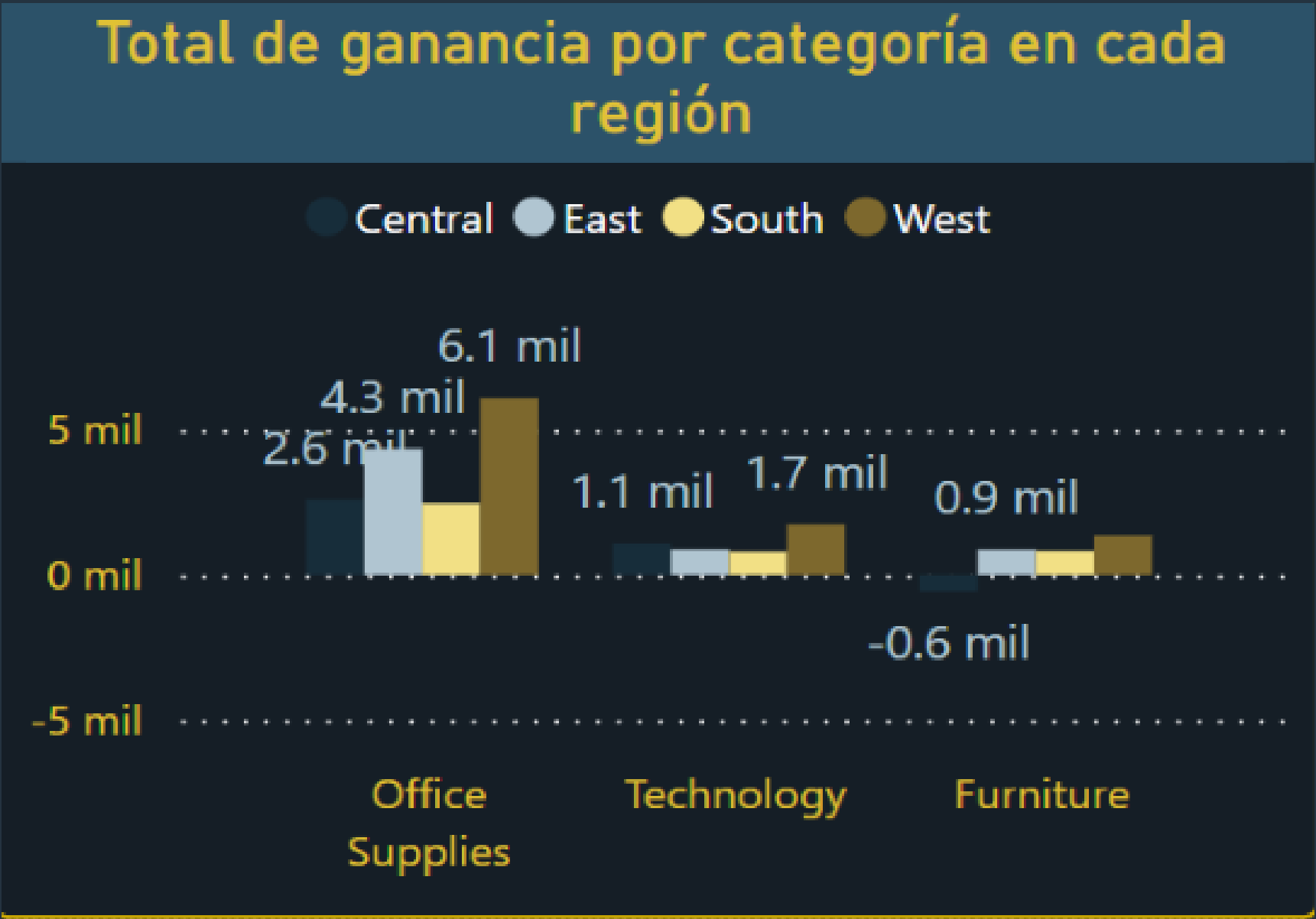
ToolTip 4 ventas:



El uso de ToolTips en la solapa “ToolTip 4 ventas” muestra las ventas anuales divididas en tres categorías: Office Supplies, Furniture y Technology. Los datos se representan a lo largo de cuatro años, identificados por colores: azul para 2017, amarillo para 2016, verde para 2015 y beige para 2014. Este gráfico de barras permite comparar fácilmente el rendimiento de las ventas de cada categoría a lo largo de estos años.

La visualización de estos datos proporciona una herramienta útil para el análisis de tendencias y patrones en las ventas. Por ejemplo, se puede observar si hay un crecimiento o una disminución en las ventas de cada categoría con el tiempo, lo que puede ayudar a identificar áreas que requieren más atención o recursos. Además, al tener los datos desglosados por año, los usuarios pueden identificar temporadas o periodos específicos que han sido más exitosos y correlacionarlos con estrategias de ventas o campañas de marketing específicas. Esta información es crucial para tomar decisiones informadas sobre la planificación y estrategia de ventas futuras.

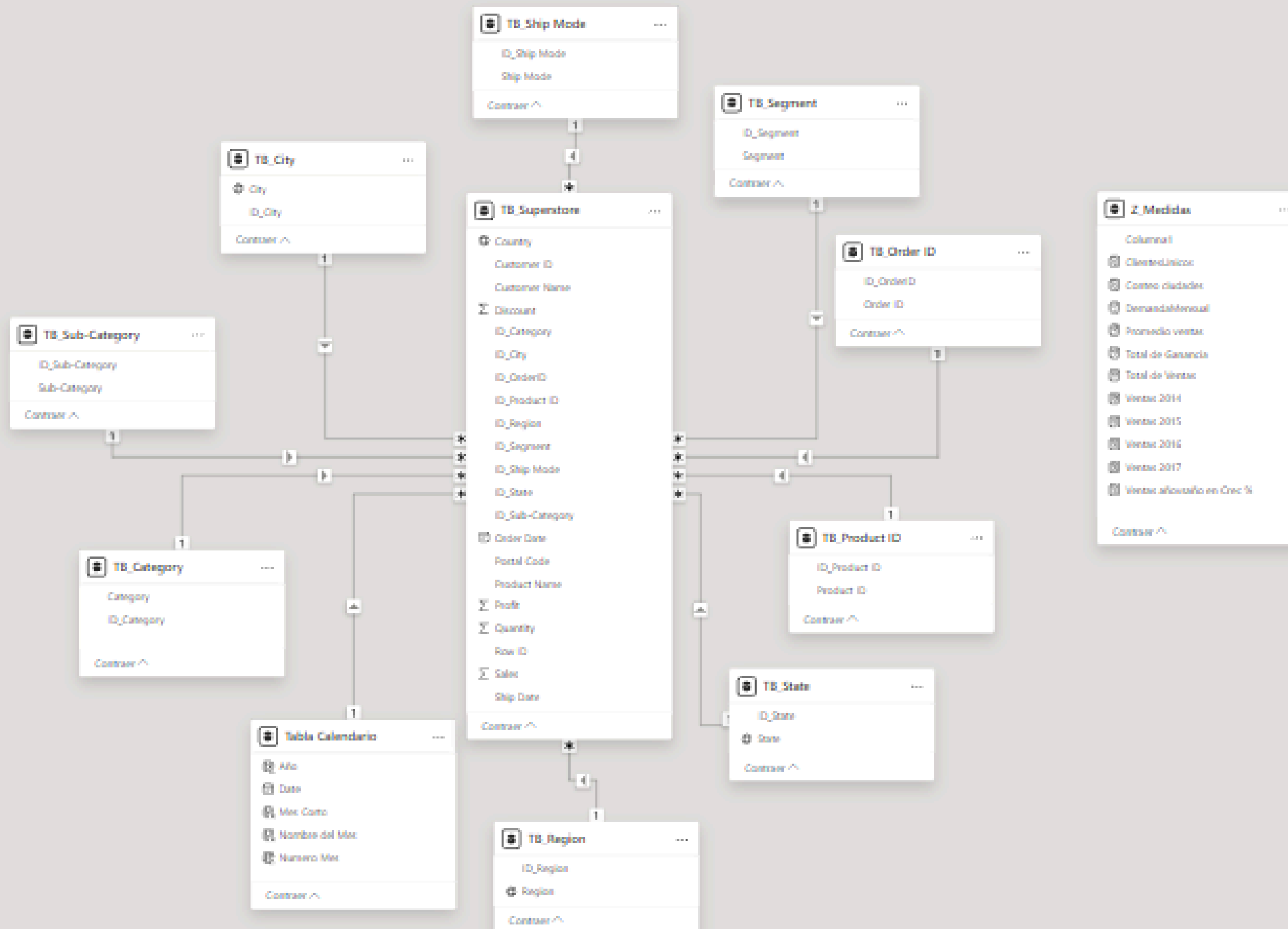
ToolTip 5 ventas:



El uso de ToolTips en la solapa de "ToolTip 5 ventas" permite una interacción adicional y un análisis más detallado. Específicamente, se implementó un ToolTip que muestra la ganancia total por categoría en cada región. Cuando los usuarios pasan el cursor sobre una

En este caso, el gráfico de barras muestra que en la región Oeste, la categoría "Suministros de Oficina" tiene la mayor ganancia con 6,1 mil, seguida por la región Este con 4,3 mil, y la región Sur con 2,6 mil en la misma categoría. En la categoría "Tecnología", la región Este lidera con 1,7 mil, mientras que en "Muebles", la región Central tiene una pérdida de -0,6 mil. Esta información es crucial para comprender el comportamiento del mercado en diferentes regiones y ajustar las estrategias de venta y marketing en consecuencia. Por ejemplo, se pueden dirigir más recursos y esfuerzos promocionales a la región Oeste, que muestra altos rendimientos en "Office Supplies", mientras que se podrían investigar las razones detrás de la pérdida en la región Central para "Furniture" y tomar medidas correctivas. La capacidad de ver estas cifras en un formato visual intuitivo también ayuda a identificar rápidamente patrones y tendencias en el rendimiento financiero por región y categoría, facilitando una toma de decisiones más informada y estratégica.

14. Diagrama Entidad - Relación de Power BI (Final)



15. Conclusiones Finales

Según las Hipótesis previstas al iniciar el Proyecto, a través de todo lo analizado se pudo llegar a las siguientes Conclusiones:

1. *Variación Estacional:* Varía mucho por año, pero se destacan picos de ventas por los meses de Julio y de octubre

2. *Producto Estrella:* El producto estrella varía mucho por zona año y mes, pero entre todos los productos que se ofrecen, se destaca una alta demanda por los "Chairs", "Phones", "Machines", "Copiers" y "Storage" siendo los que más ventas tienen todos los años.

3. *Segmentación de Clientes:* Por año se encuentran clientes que se destacan notoriamente por sus compras, encontrando clientes que superan las 22.3 compras por año.

4. *Eficacia de Promociones:* Se encuentran zonas con más ganancias que otras, destacándose la región Oeste (West).

16. Futuras Líneas

En este apartado, se consideraron ciertos puntos que no se explayaron en el trabajo.

Se consideraría oportuno:

1. Continuando con las hipótesis planteadas, se podría profundizar en el Análisis de algunos atributos particulares de los clientes.
2. Ver como son las ventas respecto a la influencia geográfica, Se podría investigar si la ubicación geográfica de las tiendas tiene algún efecto en los patrones de facturación, como si las ventas varían según la región o el país.
3. Averiguar cómo actúan las ventas respecto a respuestas a eventos externos. Podría plantearse la hipótesis de que eventos externos, como cambios en la economía o tendencias del mercado, tienen un impacto en la facturación del Superstore.

17. Referencia Bibliográfica

<https://www.kaggle.com/datasets/vivek468/superstore-dataset-final>