



Entrega Final

Proyecto Ventas

Estados Unidos

Curso: Data Analytics

Comisión: 42310

Autor: Leandro Montes Catari

Tutor: Lucas Bernal

Índice

Contenido

1.	Tabla de Versionado.....	1
2.	Glosario de Términos.	1
3.	Definición de la fuente de información.	2
4.	Análisis de datos.....	3
5.	Descripción de la temática.	4
6.	Hipótesis.	4
7.	Datasets.	4
8.	Objetivo de proyecto.....	4
9.	Herramientas Tecnológicas Implementadas.	5
10.	Alcance.....	5
11.	Usuario Final y nivel de Aplicación del Análisis (Operativo, Táctico, Estratégico).....	5
12.	Diagrama Entidad – Relación.	6
13.	Listado de Tablas.	6
14.	Listado de Columnas de cada tabla.	7
15.	Mockups, diseño de Dashboard y solapas.....	9
16.	Conexión de Base de Datos.....	12
17.	Modelo de Relacional de Datos	13
18.	Medidas y columnas calculadas.....	14
19.	Medidas Calculadas Avanzadas.	16
20.	Futuras Líneas.....	18

1. Tabla de Versionado.

Versión	Fecha	Actividad
V1	27/06/2023	Modelado de datos definicion de columnas en tablas.
V2	17/07/2023	Diseño de prototipo.
V3	21/08/2023	Diseño de Graficos y Formato de Dashboard.
V4	06/09/2023	Correcciones recomendadas, incorporacion de glosario y tabla de versiones.

2. Glosario de Términos.

- **Cliente:** Persona o entidad que compra productos o servicios de una empresa.
- **Empresa:** Entidad comercial que vende productos o servicios.
- **Gross margin / Margen de Ganancia:** Margen bruto, es decir, la diferencia entre los ingresos totales y los costos de los bienes vendidos.
- **Order Number / Orden de Venta:** Documento que registra la venta de un producto o servicio a un cliente.
- **Producto:** Bien o servicio que se vende a los clientes.
- **Región:** Área geográfica en la que se realizan ventas.
- **Revenue:** Ingresos totales generados por una empresa.
- **Sales channel:** Canal de ventas, es decir, el medio a través del cual se realizan las ventas (por ejemplo, para este proyecto, In store, Online, Distributor, Wholesale).
- **Sales team:** Equipo de ventas, es decir, el grupo de personas que se encarga de vender productos o servicios de una empresa.
- **Tienda:** Establecimiento físico o virtual donde se venden productos o servicios.
- **Vendedor:** Persona que se encarga de vender productos o servicios de una empresa a los clientes.

3. Definición de la fuente de información.

- Ventas en Estados Unidos.
- Peleas y Peleadores de la UFC.
- Accidentes de tránsito en Chile.

1. **Ventas en Estados Unidos:** Tome como principal temática las ventas en Estados Unidos y aprovechando que me fue facilitado este Dataset que esta OK y aparte de eso es una temática que me llama mucho la atención ya que últimamente me he estado adentrando en el mundo del consumo masivo de un producto. En esta temática podremos realizar diferentes análisis. Productos más vendidos, por zona, por canal de venta, por tienda. Tiendas con más ventas realizadas, top 10 clientes que más han comprado. Son algunos de los análisis que podre realizar con este Dataset.
2. **Peleas y Peleadores de UFC:** En esta temática podremos analizar a los peleadores y las peleas más comerciales que se han dado en la franquicia a través del tiempo. Podremos determinar cuáles fueron las peleas con más recaudación, peleadores mejores pagados, en que países o provincia se han realizado más combates, récords por peleadores, ranking de peleadores por divisiones, la nacionalidad que más predomina en peleadores, arte marcial o especialidad más usada por peleadores.
3. **Accidentes de tránsito en Chile:** En la temática podremos analizar un periodo de tiempo determinado los accidentes de tránsito que ocurrieron en las diferentes regiones del país, donde podremos medir las causas que generaron el accidente, resultado del accidente en las personas, pérdidas materiales ocasionadas, etc.

4. Análisis de datos.

Para esta actividad ocupare la empresa donde actualmente trabajo, a esta la llamaremos **Empresa LOS POLLOS, SA** para no usar su nombre real, esta empresa se dedica a la comercialización de productos cárnicos (pollo) en sus diferentes presentaciones. Aunque no trabajo directamente en el área de comercialización podre aplicar los cuatro tipos de Análisis a un objetivo común para esta área que es maximizar sus ventas / ingresos, lo que se traduce a mantener la mayor cantidad de consumidores posibles.

Iniciativa: Maximizar Ventas / Ingresos, aumentar cartera de Clientes.

Iniciativa: Maximizar Ventas / Ingresos, aumentar cartera de Clientes.

Análisis descriptivo: Recopilamos los datos de las ventas diarias de los productos, los precios, los canales de distribución, ubicación geográfica de las tiendas, entre otros. En base a esto podemos generar informes y visualizaciones que describan las tendencias de venta, los productos más populares, las ubicaciones con mayor demanda, los patrones estacionales, entre otros. Esto proporciona una comprensión clara de la situación actual de la empresa y ayuda a identificar áreas de oportunidad y patrones clave

Análisis diagnóstico: Al realizar este análisis podremos comprender los factores que influyen en las ventas y la satisfacción del cliente. Por ejemplo, podemos investigar las preferencias de los clientes en términos de características del producto, presentaciones, precios y calidad. También podemos analizar la eficacia de las estrategias de marketing, propaganda, descuentos y promociones utilizadas anteriormente. El análisis diagnóstico nos permitirá identificar los factores clave que influyen en las ventas y el comportamiento del cliente.

Análisis predictivo: Utilizando los datos históricos de ventas, los datos demográficos de los clientes y otros factores relevantes, podremos predecir las tendencias de venta futuras y anticipar la demanda de productos. Por ejemplo, podemos utilizar modelos de series de tiempo para prever la demanda en temporadas específicas o utilizar técnicas de segmentación de clientes para identificar grupos con mayor potencial de compra. Esto nos permite planificar la producción, la distribución y las estrategias de precios para satisfacer la demanda y aumentar las ventas.

Análisis prescriptivo: Una vez que hemos obtenido las predicciones, podremos realizar un análisis prescriptivo para determinar las mejores acciones a tomar para aumentar las ventas y expandir nuestra cartera de clientes. Por ejemplo, podemos analizar la competencia y los patrones de compra para identificar oportunidades de diferenciación y mejorar la propuesta de valor. El análisis prescriptivo nos permite tomar decisiones informadas y recomendar acciones específicas para alcanzar los objetivos comerciales.

5. Descripción de la temática.

Cómo se mencionó anteriormente la temática elegida fue de Ventas en Estados Unidos para el periodo del 31-05-2018 al 30-12-2020. En este datasets contamos con los datos de todas las ventas realizadas. Cada venta tiene asignado su número de orden, también tiene como campo el ID cliente, ID de tienda, ID de producto, ID de equipo de venta, entre otros. Cabe mencionar que en el dataset cuento con las tablas de cliente, tienda, equipo de ventas, producto y región, las cuales facilitaran la construcción de las Bases de datos para realizar distintos análisis.

6. Hipótesis.

Percibiremos por medio de los datos cuales fueron los canales de ventas más efectivos, los productos más comerciales o con mayor salida o al contrario podremos detectar los canales de venta más débiles, las tiendas con menor facturación, en base a estos hallazgos podremos aplicar estrategias para recuperarnos y posicionar mejor el producto en el mercado.

7. Datasets.

El datasets se encuentra en el Drive compartido. Esta nombrado de la siguiente manera "DATASETS_VENTAS_ESTADOS_UNIDOS.xlsx".

8. Objetivo de proyecto.

El objetivo del proyecto es poder crear desde 0 una estructura de datos mediante el modelo de entidad de relación y el modelo relacional, donde podremos detallar todas las tablas que usaremos, con sus debidos atributos, identificando cada Primary y Foreign Key. Posterior a esto podremos analizar todos los datos de las ventas realizadas en estados unidos en un periodo aproximado de 3 años. En detalle podremos analizar por medio de consultas todos los datos del dataset como, por ejemplo: Productos más vendidos en cierto periodo de

tiempo, equipo de venta que más ventas realizo, tienda que más ventas online realizo, canal de ventas por donde más se realizaron ventas y muchas otras más.

9. Herramientas Tecnológicas Implementadas.

Para la elaboración de este proyecto ocupamos **Microsoft Excel** ya que la base de datos la tenemos en formato .xls, para la transformación de datos ocupamos el editor de **Power Query**, para la elaboración de Mockup ocupo **Microsoft Power Point**, para la elaboración de este documento se ocupo **Microsoft Word** y para la visualización de Dashboard de ocupo **Power BI**.

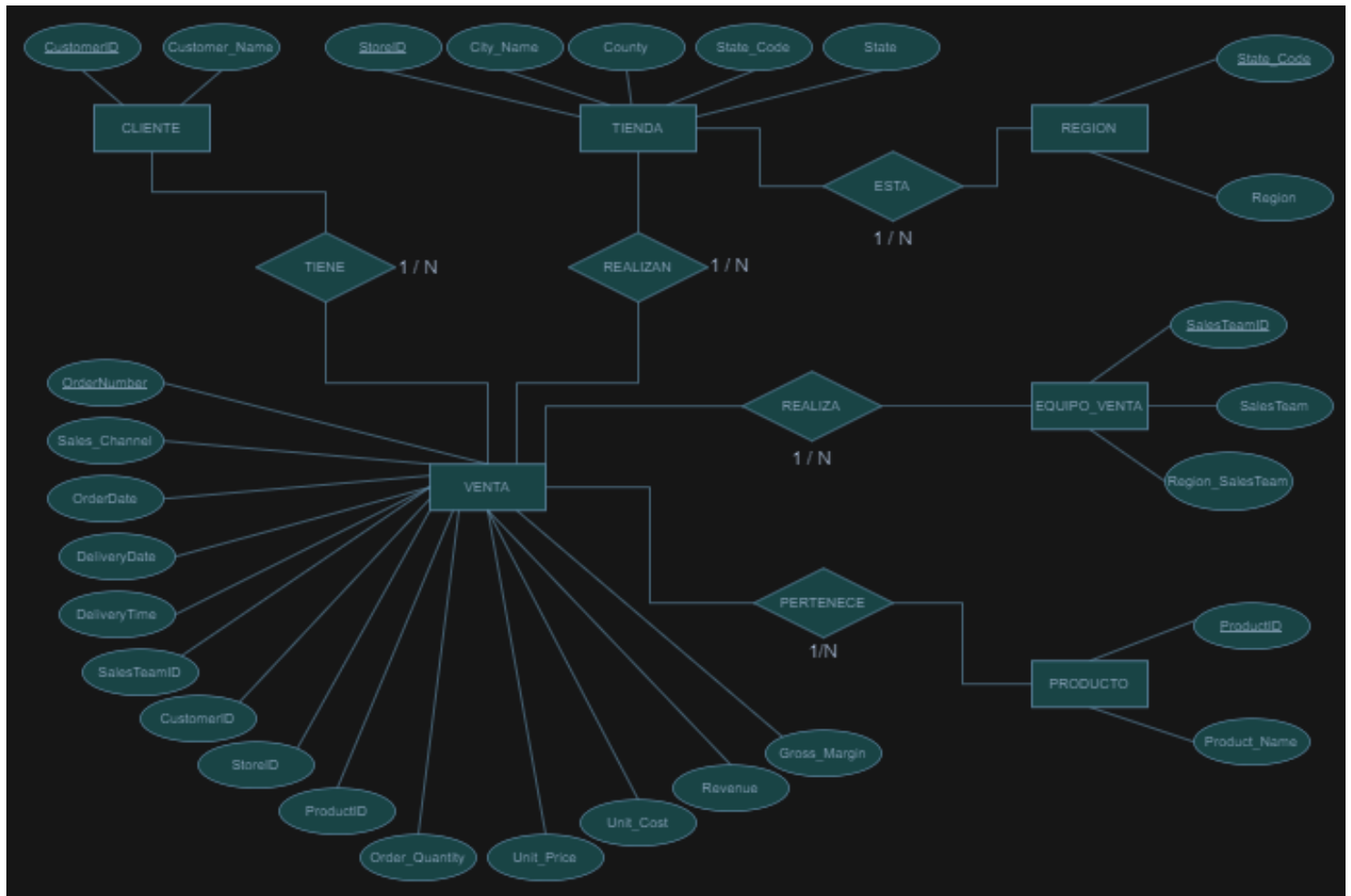
10. Alcance.

Por medio de los datos que tenemos se realizaran diversos análisis lo cual nos permitirá tener una visión panorámica sobre las ventas realizadas en aproximadamente 3 años, esto nos permitirá tomar decisiones y determinar acciones que a corto o mediano plazo nos darán resultados satisfactorios o no. Cabe mencionar que para obtener buenos resultados tenemos que estar en constante análisis y revisión de los datos.

11. Usuario Final y nivel de Aplicación del Análisis (Operativo, Táctico, Estratégico).

Como se comentó anteriormente por medio de los análisis de datos que se realicen la Gerencia se encargara de tomar decisiones y transmitir a las jefaturas las estrategias y acciones que se deben realizar con el fin de mejorar el negocio, ahorrar en costos, obtener mayores ganancias, posicionarse en más sitios. Por lo tanto, para este análisis lo aplicaríamos a nivel estratégico ya que sin una estrategia difícilmente podríamos obtener buenos resultados en las ventas.

12. Diagrama Entidad – Relación.



13. Listado de Tablas.

- a) VENTA
- b) CLIENTE
- c) TIENDA
- d) EQUIPO_VENTA
- e) PRODUCTO
- f) REGION

14. Listado de Columnas de cada tabla.

Tabla Venta: En esta tabla encontraremos los datos que se generan al realizar la venta de un producto, tenemos el número de Orden como PK, para clasificar medio por donde se realizó la venta tendríamos el canal de venta, tenemos fecha de la orden y fecha de entrega de producto, en el campo DeliveryTime registramos el tiempo que se tardó en entregar el pedido en minutos. Tenemos cantidad de productos vendidos, precio unitario, costo unitario, revenue vendría siendo el total y margen de ganancia. Como FK tenemos los id's de todas las tablas con las que se relaciona la tabla Ventas.

Nota: Cabe mencionar que, para este dataset cada venta tiene solo un producto.

TIPO DE CLAVE	CAMPO	TIPO DE DATO
PK	OrderNumber	VARCHAR(100)
-	Sales_Channel	VARCHAR(100)
-	OrderDate	DATE
-	DeliveryDate	DATE
-	DeliveryTime	TIME
-	Order_Quantity	INT
-	Unit_Price	DECIMAL(10,2)
-	Unit_Cost	DECIMAL(10,2)
-	Revenue	DECIMAL(10,2)
-	Gross_Margin	DECIMAL(10,2)
FK	SalesTeamID	INT
FK	CustomerID	INT
FK	StoreID	INT
FK	ProductID	INT

Tabla Cliente: En esta tabla podemos encontrar el ID del cliente que se va incrementando correlativamente, y su respectivo nombre.

TIPO DE CLAVE	CAMPO	TIPO DE DATO
PK	CustomerID	INT
-	Customer_Name	VARCHAR(100)

Tabla Tienda: En esta tabla podremos encontrar el ID de la tienda, nombre de la ciudad donde está ubicada, condado, estado y en código del estado que vendría siendo una FK de la tabla Región.

TIPO DE CLAVE	CAMPO	TIPO DE DATO
PK	StoreID	INT
-	City_Name	VARCHAR(100)
-	County	VARCHAR(100)
-	State	VARCHAR(100)
FK	State_Code	VARCHAR(100)

Tabla Equipo_Venta: Tenemos el ID del equipo de venta que se autoincrementa, el nombre del vendedor y la región a la cual pertenece.

TIPO DE CLAVE	CAMPO	TIPO DE DATO
PK	SalesTeamID	INT
-	SalesTeam	VARCHAR(100)
-	Region_SalesTeam	VARCHAR(100)

Tabla Producto: En esta tabla podemos encontrar el ID del producto que se va incrementando correlativamente, y su respectivo nombre.

TIPO DE CLAVE	CAMPO	TIPO DE DATO
PK	ProductID	INT
-	Product_Name	VARCHAR(100)

Tabla Región: En esta tabla podremos encontrar el código del estado y la región a la cual pertenece dicho estado.

TIPO DE CLAVE	CAMPO	TIPO DE DATO
PK	State_Code	VARCHAR(100)
-	Region	VARCHAR(100)

15. Mockups, diseño de Dashboard y solapas.

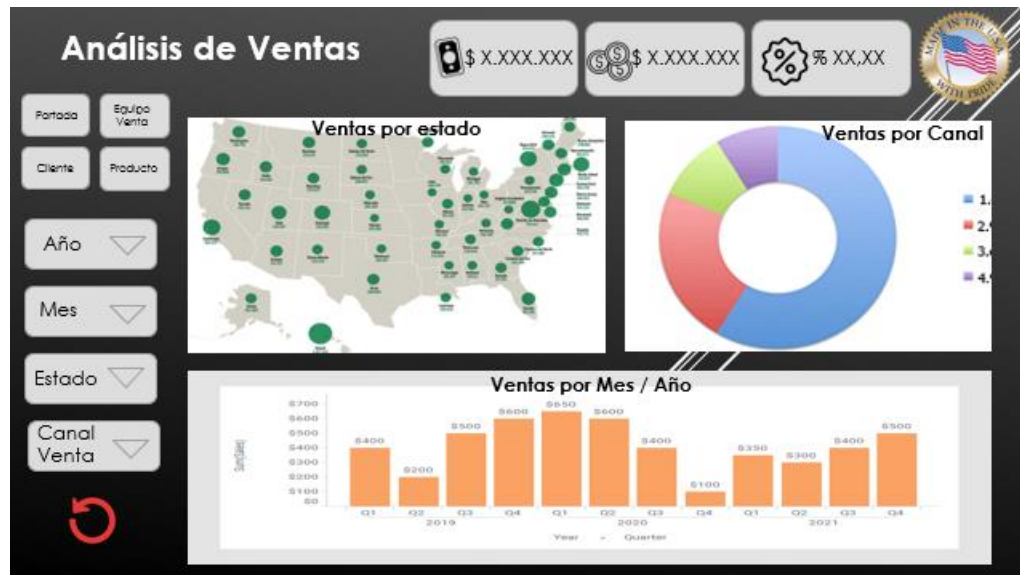
En el diseño inicialmente se incluirán 6 solapas, donde la primera solapa será la portada, donde se incluirá el nombre del proyecto, algunos datos generales y los botones que nos ayuden en la navegación del tablero.

- **Solapa Portada:**



- **Solapa Glosario:** En esta solapa se incluirán todos los conceptos y términos técnicos utilizados en el desarrollo de este proyecto. Esto con el fin de un rápido y fácil entendimiento.

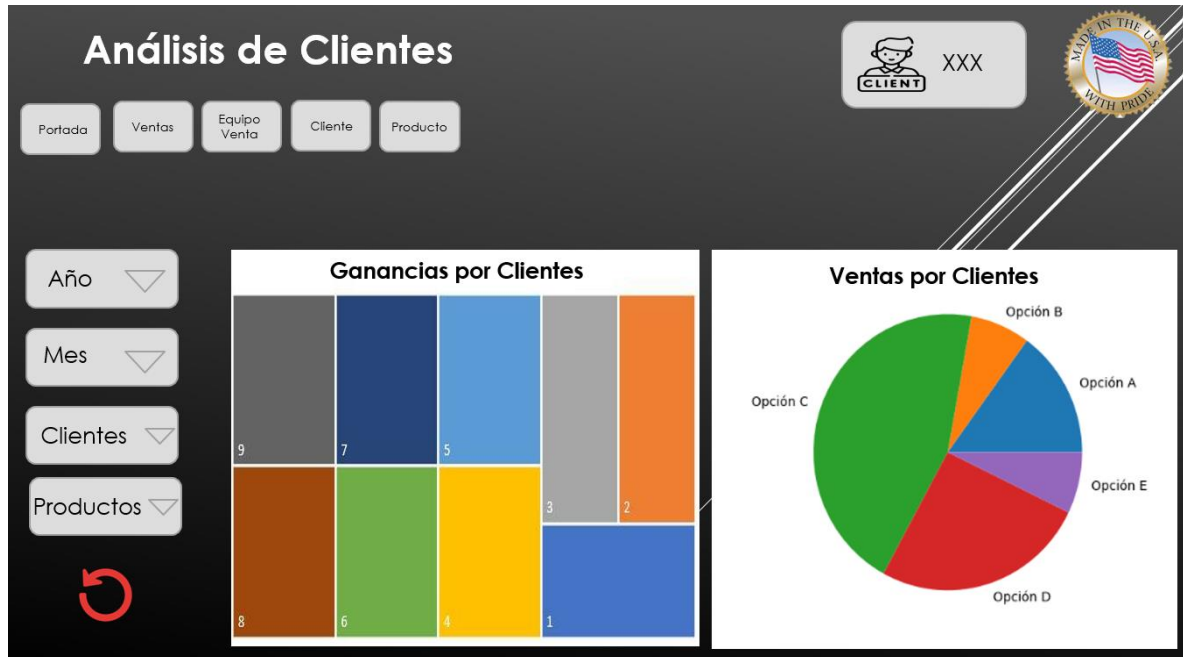
- **Solapa Ventas:** En esta solapa nos enfocaremos en analizar las ventas realizadas en el periodo de tiempo que tenemos en nuestro dataset, de igual manera podremos analizar ventas por estado o región, ventas por canal o canales determinados. Para todas estas visualizaciones tendremos su segmentadores (Año, Mes, Canal de venta, estado o cualquier otro que necesitemos). También mostraremos algunos datos generales en las tarjetas.



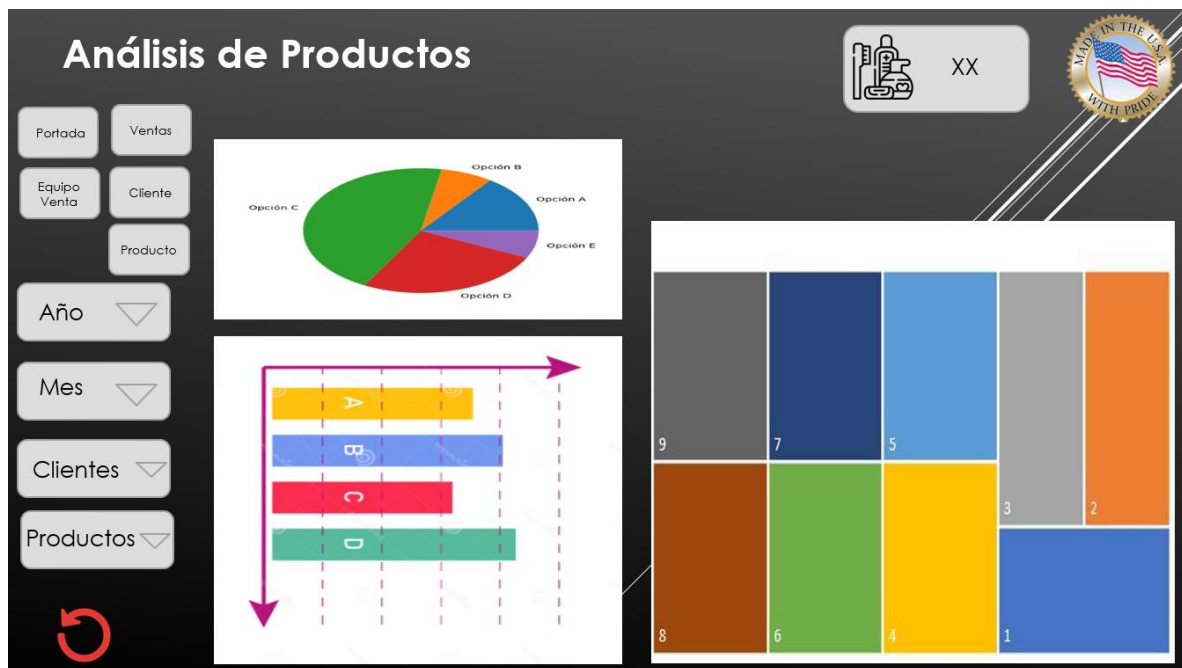
- **Solapa Equipo de Ventas:** En esta solapa analizaremos todas las ventas realizadas por cada vendedor, también podremos analizar las ganancias por vendedor, también podremos visualizar algunos KPI como la cantidad de vendedores y Ventas Máximas o mínimas. Y podremos destacar el vendedor con más ventas o más ganancias.



- **Solapa Clientes:** En esta solapa tendremos un top de ventas por clientes, donde podremos apreciar en grafico los clientes que más gastaron o compraron, podremos ver los productos que compraron ciertos clientes, tendremos una tarjeta donde veremos la cantidad de clientes registrados hasta el momento y tendremos los filtros generales para segmentar los datos (año, mes, cliente, producto).

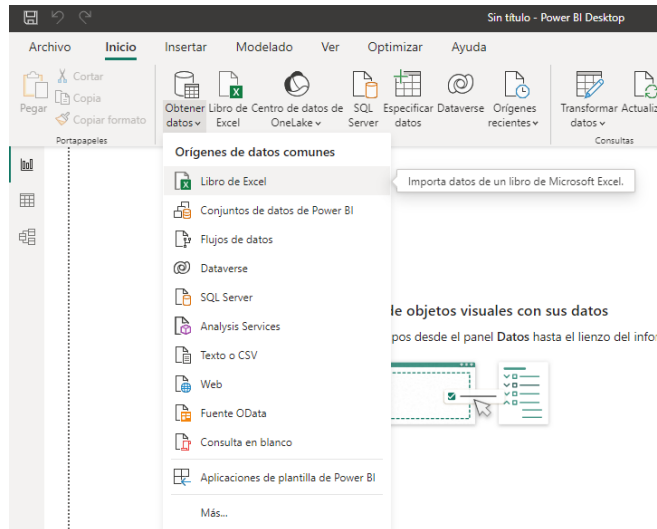


- **Solapa Productos:** Al igual que en la solapa de clientes tendremos algún top que muestre los productos más vendidos, tendremos las ganancias conseguidas por producto y las ventas por producto.

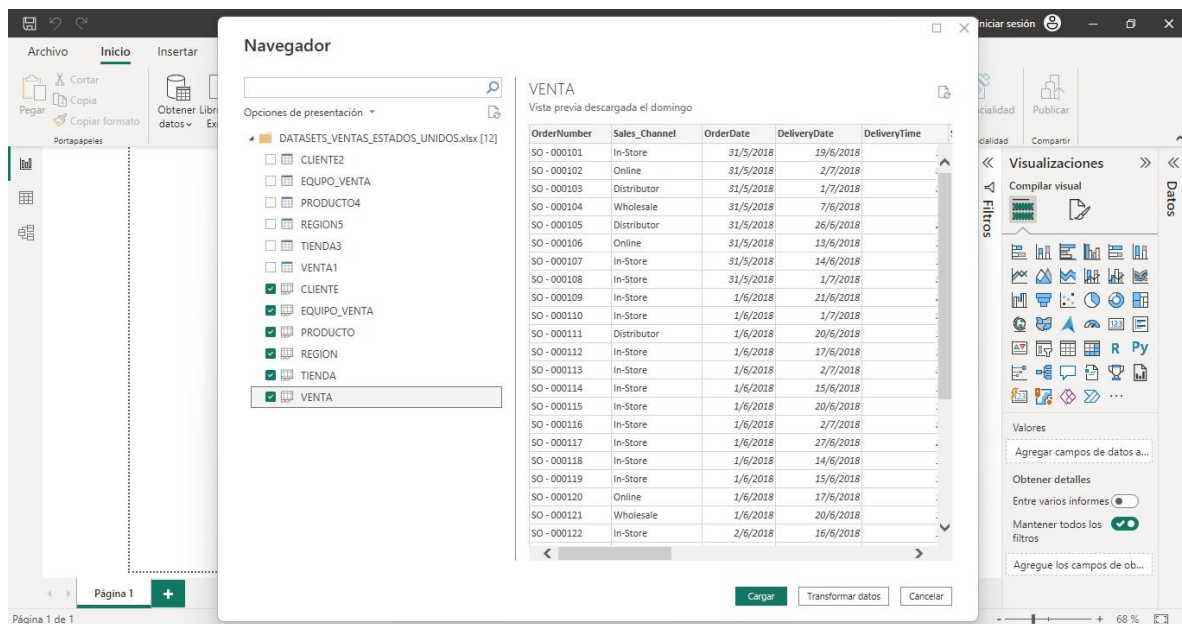


16. Conexión de Base de Datos

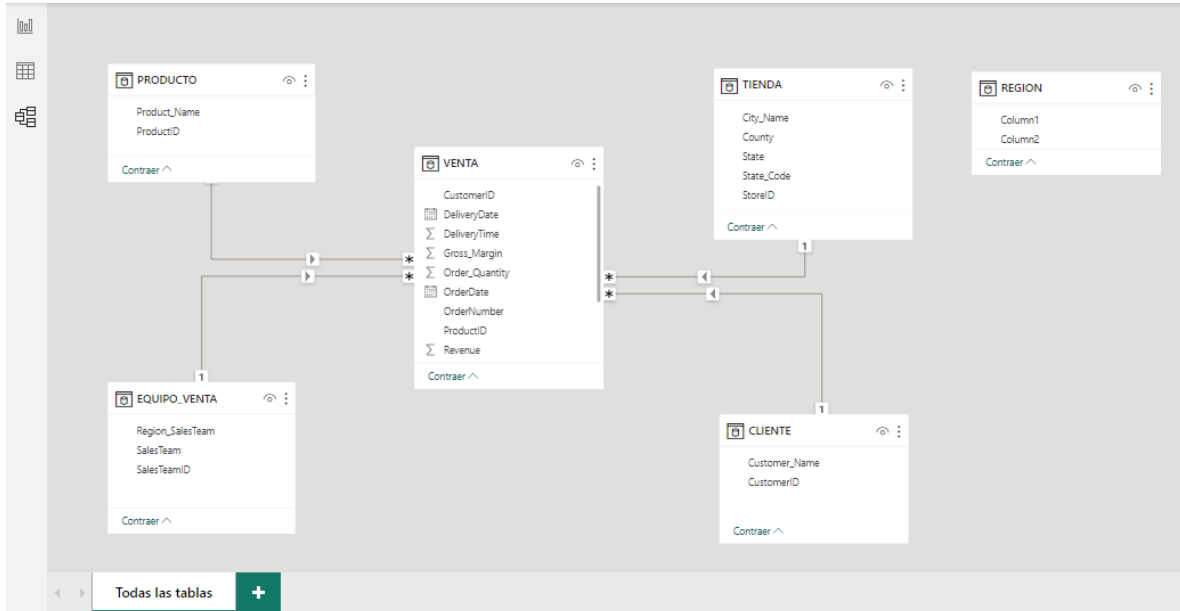
Se realiza la conexión de la base de datos elegida para el desarrollo del proyecto final, la conexión del archivo se realiza a través de un libro Excel como se aprecia en la imagen “DATASETS_VENTAS_ESTADOS_UNIDOS.xlsx”.



Conexión realizada con archivo Excel. Por algún motivo aparece una lista de tablas que no tengo definidas en el dataset. Pero aclarar que se seleccionaron solo las tablas que se definieron en la segunda entrega.

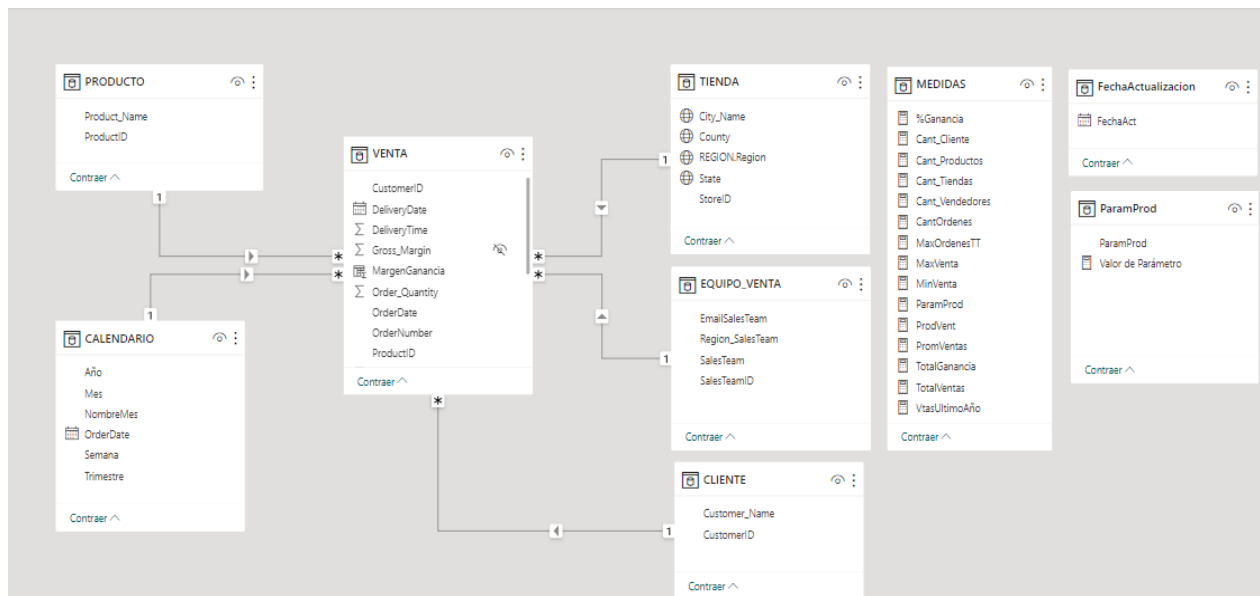


Luego de cargar las tablas Power Bi genero automáticamente el siguiente modelo relacional, el cual será modificado de ser necesario en la siguiente actividad.



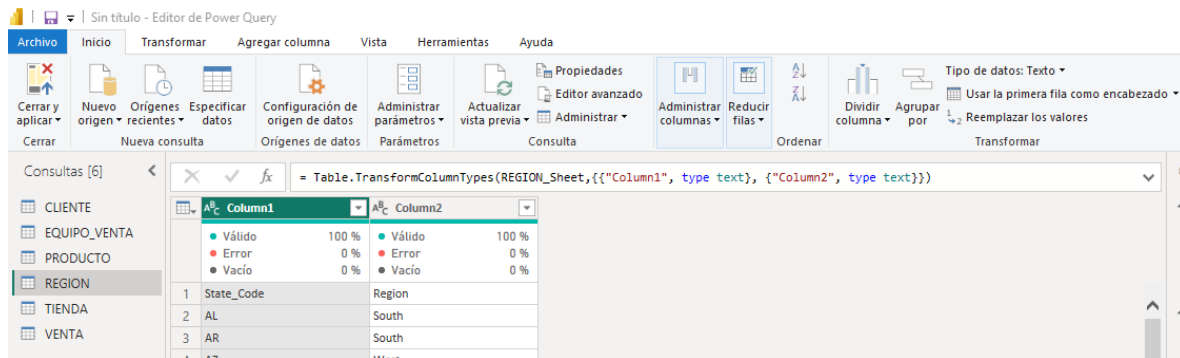
17. Modelo de Relacional de Datos

El modelo relacional quedaría de la siguiente manera ya que por el dataset elegido no sería necesario agregar tablas puentes. (tabla Región se deshabilita, se deja en stand by)

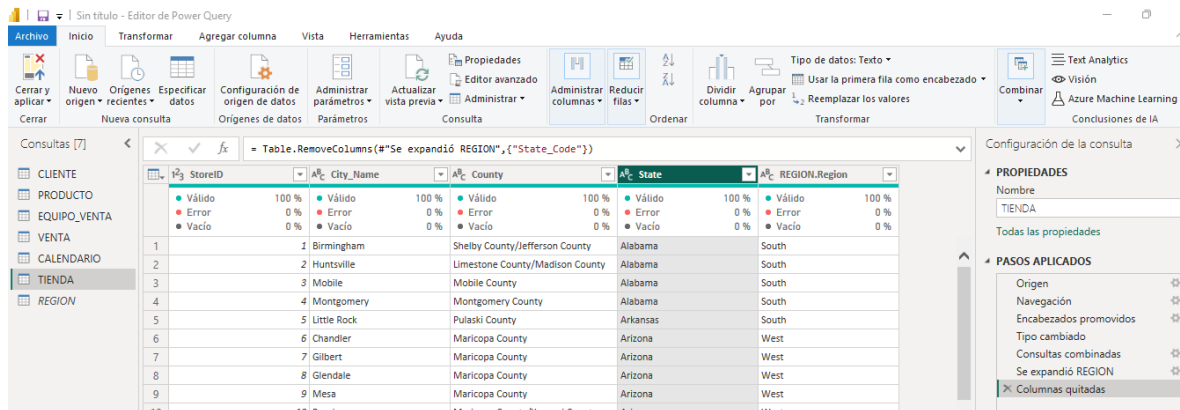


- **Transformación de datos.**

Luego de realizar la conexión de la base de datos con Power BI procedí a corregir un error en el encabezado de las columnas de la tabla **REGION**. Este error lo corregí clickeando en el botón de usar fila como encabezado.



En la tabla **TIENDA** decidí realizar un Join (Combinar Consultas) para traerme el nombre de la región de la tabla **REGION** y así dejar esta inhabilitada o en stand by mientras no la necesite. Al realizar el Join ya no sería necesario tener en la tabla **TIENDA** el código de la región por lo que procedí a eliminar la columna StateCode.



18. Medidas y columnas calculadas.

- **Medidas Calculadas**

En la tabla **PRODUCTO** se generó una medida llamada Cant_Productos, como su nombre lo indica cuenta la cantidad de productos existentes en la tabla.

Cant_Productos = COUNT(PRODUCTO[ProductID])

En la tabla **EQUIPO_VENTA** se generó una medida llamada Cant_Vendedores y como su nombre lo indica cuenta la cantidad de vendedores en la tabla. Esta vez con una variable.

```
Cant_Vendedores =
VAR CantVendedores =
    COUNT(EQUIPO_VENTA[SalesTeamID])
return
    CantVendedores
```

En la tabla **VENTA** se generó una medida llamada TotalVentas que se encarga de sumar todas las ventas realizadas en el periodo de tiempo que tenemos en la tabla.

```
TotalVentas = SUM(VENTA[TotalOrder])
```

También en tabla **VENTA** se generó una medida llamada TotalGanancia que suma todas las ganancias de todas las ventas.

```
TotalGanancia = SUM(VENTA[MargenGanancia])
```

- **Columnas Calculadas**

Si bien la tabla **VENTA** contiene las columnas Revenue que se refiere al total de la venta, la columna Gross_Margin que se refiere al margen de ganancia y DeliveryDate que sería el tiempo en minutos que se tardó la entrega, decidí asegurarme de que el cálculo estuviese correcto, por lo que procedí a generar estas columnas calculadas y trabajar con estas.

Columna TotalOrder, donde se calcula el total de la orden de venta.

```
TotalOrder = VENTA[Order_Quantity] * VENTA[Unit_Price]
```

Columna MargeGanancia, donde se calcula la ganancia total de la orden de venta.

```
MargenGanancia = VENTA[TotalOrder]-(VENTA[Order_Quantity]*VENTA[Unit_Cost])
```

A parte de esto también agregue una columna personalizada llamada TipoVenta, donde definí que si la cantidad de la orden (Order_Quantity) es menor o igual a 5 es una venta “Al Detal”, si la cantidad es mayor a 5 entonces sería una venta “Al Mayor”.

Agregar una columna condicional

Agregue una columna condicional que se calcula a partir de las otras columnas o valores.

Nuevo nombre de columna:

	Nombre de columna	Operador	Valor	Salida
Si	Order_Quantity	es menor o igual...	5	Enton... Al Detal
O si	Order_Quantity	es mayor que	5	Enton... Al Mayor

De lo contrario:

Para agregar la tabla **CALENDARIO** fue necesario duplicar la tabla **VENTA**, luego de esto elimine todos los campos que no necesitaba, dejando solamente la fecha de la orden declarada como OrderDate. Elimino las filas duplicadas y agrego a la tabla **CALENDARIO** los campos año, nombre del mes, semana y trimestre.

	OrderDate	Año	Nombre del mes	Semana del año	Trimestre
1	31/5/2018	2018	mayo	22	2
2	31/5/2018	2018	mayo	22	2
3	31/5/2018	2018	mayo	22	2
4	31/5/2018	2018	mayo	22	2
5	31/5/2018	2018	mayo	22	2
6	31/5/2018	2018	mayo	22	2
7	31/5/2018	2018	mayo	22	2
8	31/5/2018	2018	mayo	22	2
9	1/6/2018	2018	junio	22	2
10	1/6/2018	2018	junio	22	2
11	1/6/2018	2018	junio	22	2

19. Medidas Calculadas Avanzadas.

- Medidas Calculadas con Variables

Medida con una variable: cree dos medidas con una variable cada una con el fin de mostrarlas en una tarjeta respectivamente en su visualización, estas son las siguientes:

Cant_Cliente, se encarga de contar todos los clientes que han comprado y están registrados en la BD. El cálculo lo realiza a partir del campo CustomerID de la tabla **CLIENTE**.

```
Cant_Cliente =
VAR CantidadClientes =
    COUNT(CLIENTE[CustomerID])
return
    CantidadClientes
```

Cant_Vendedores, se encarga de contar todos los vendedores registrados en la BD. El cálculo lo realiza a partir del campo SalesTeamID de la tabla **EQUIPO_VENTA**.

```
Cant_Vendedores =
VAR CantVendedores =
    COUNT(EQUIPO_VENTA[SalesTeamID])
return
    CantVendedores
```

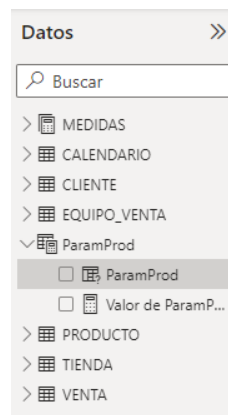
Medida con dos variables: Se crea una medida calculada llamada % Ganancia con el fin de mostrar en una tarjeta la ganancia en porcentaje, se declara una variable TotalVenta, donde se suman todas las ventas del campo TotalOrder de la tabla **VENTAS**, la segunda variable llamada TotalGanancia donde se suman todas las ganancias del campo MargenGanancia de la misma tabla **VENTAS** por último, se declara la variable Ganancia donde se realiza la operación con la función DIVIDE, se generó de la siguiente manera:

```
%Ganancia =
VAR TotalVenta =
    SUM(VENTA[TotalOrder])
VAR TotalGanancia =
    SUM(VENTA[MargenGanancia])
VAR Ganancia =
    DIVIDE (TotalGanancia, TotalVenta,0)
return
    Ganancia
```

Medida con función de inteligencia de tiempo: Se agrega una medida calculada donde usamos una función de inteligencia de tiempo DATE, la cual usamos para mostrar todas las ventas del último año del dataset o por así decirlo año en curso. En esta se usan los campos TotalOrder y OrderDate de la tabla **VENTAS**.

```
VtasUltimoAño =
VAR Vta2020 =
    CALCULATE(SUM(VENTA[TotalOrder]), VENTA[OrderDate] >
        DATE(2020, 01,01) && DATE(2020,12,31))
return
    Vta2020
```

Parámetro: Se creo una medida calculada con parámetro, para controlar el grafico directamente con el filtro de parámetro generado.



La medida para enlazar el parámetro con el grafico es ParamProd, en esta podremos visualizar las Ganancias Por Productos vendidos.

```
ParamProd =  
    VAR VentParam =  
        [TotalGanancia] * (1 + ParamProd[Valor de ParamProd])  
    return  
    VentParam
```

20. Futuras Líneas.

Para este proyecto en un futuro, lo que realizaré será incorporar más datos de las distintas tablas e incorporarlos al dataset. Esto lo hago con la intención principal de seguir aprendiendo, pero también tendría la posibilidad de llevar a cabo análisis más complejos.

Con la ampliación de datos, podré explorar nuevas tendencias y patrones que puedan surgir, lo que enriquecerá aún más la calidad de los análisis y las conclusiones que se puedan extraer. Además, considero que este proyecto tiene un gran potencial para brindar información valiosa que pueda ser de utilidad en futuras investigaciones o decisiones estratégicas.

En general, la incorporación de más datos al dataset me permitirá hacer un mejor uso del modelo y generar resultados más beneficiosos.

Nota: Como información adicional se agregaron correcciones o recomendaciones transmitidas por el tutor en la 3era entrega, se reemplaza grafico donde no se apreciaban bien los datos debido a la cantidad de información manejada. De igual manera se agrega un ToolTips en cada solapa para complementar el análisis.