



Power BI – Dashboard Ventas Infusiones USA

Joaquín Gabriel Manteca

Data Analytics 2024

26 de mayo de 2024

Última actualización: 16.58 pm

PROYECTO FINAL

Power BI

Dashboard Ventas Infusiones USA



Tabla de contenido

1.	Introducción de la temática.....	4
1.1	Descripción de la temática	4
1.2	Objetivo	4
1.3	Hipótesis.....	4
1.4	Nivel de aplicación.....	5
1.5	Tecnología a utilizar	5
1.5	Glosario de palabras claves	5
1.6	Selección de paleta de colores	5
1.7	Tabla de versionado.....	6
2.	Base de datos	7
2.1	Modelo Entidad-Relación	7
2.2	Listado de las tablas e información de sus campos.....	7
3.	Tablero en Power BI.....	10
3.1	Transformaciones realizadas	10
3.2	Tabla calendario	10
3.3	Medidas calculadas generadas y sus fórmulas.....	11
3.4	Presentación del tablero	12
3.5	Diagrama Entidad-Relación en Power BI.....	16
4.	Futuras líneas de investigación.....	17

1. Introducción de la temática

1.1 Descripción de la temática

Esta fuente de datos tiene información de las ventas de diferentes tipos de café y té en distintos estados de, valga la redundancia, Estados Unidos. La misma se obtuvo de una lista con bases de datos entregado por la profesora Laura.

Al mismo tiempo, cada estado tiene un código de área, dependiendo la zona en las que fueron realizadas las ventas dentro del estado. Por otro lado, las ventas están divididas dentro de 4 categorías correspondientes con la zona del país en la que se encuentra el estado. Las mismas son: Este, Oeste, Sur y Central. Además, cada estado tiene una calificación en cuanto al tamaño del mercado que se divide en mercado pequeño y mercado grande.

Por otro lado, la fuente de datos tiene información de las Ganancias, Margen, Ventas, Costos de producción, Total de las expensas, Marketing e Inventario por producto y por estado. Las siguientes ecuaciones muestran la manera de calcular el margen y las ganancias:

- 1) Margen: $\text{Ventas} - \text{Costo de producción}$
- 2) Ganancia: $\text{Margen} - \text{Expensas totales}$

En última instancia, se puede ver una fecha en la cual se obtuvieron estos datos que va desde el 1/1/2010 al 12/1/2010 y del 1/1/2011 al 12/1/2011.

1.2 Objetivo

El objetivo del proyecto es realizar una comparación de las ventas de Té y Café entre los primeros días del año 2010 y el año 2011, diferenciándolas por estado y tipo de producto. Con estos datos la idea es concluir si hubo un incremento o una disminución en las ganancias y, por lo tanto, evaluar si es necesario realizar un cambio en las ventas o si se debe seguir de la misma manera.




1.3 Hipótesis

Debido a la recesión económica de 2008 en los Estados Unidos, que para algunos expertos se extendió hasta el 2011, es probable que las ganancias de las ventas de café y té hayan disminuido entre 2010 y 2011. Igualmente, también es probable que el comportamiento no sea el mismo para todos los estados ya que todos tienen cierta independencia y, por lo tanto, los resultados de las ventas no se van a comportar de la misma manera.

1.4 Nivel de aplicación

Este análisis está dirigido a los gerentes de área (líderes) de la empresa para que tomen decisiones a mediano plazo, por lo tanto, se trata de un nivel de aplicación Táctico.

1.5 Tecnología a utilizar

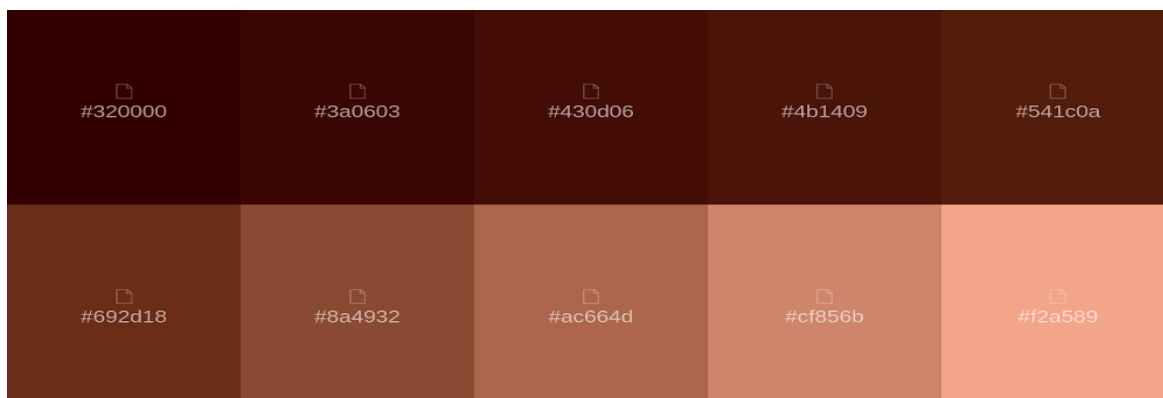
Función	Programa	Logo
Como base de datos	Excel	
Para el desarrollo del tablero	Power BI	
Para el diseño de los Mock-ups	Google slides	

1.5 Glosario de palabras claves

- a. Profit: Ganancias
- b. Margin: Margen
- c. Sales: Ventas
- d. COGS: Costos de producción
- e. Total expenses: Total de las expensas
- f. Inventory: Inventario
- g. Small market: mercado pequeño
- h. Major market: mercado grande

1.6 Selección de paleta de colores

Basado en la temática de la fuente de datos, es decir, la venta de infusiones como el café y el té y, tomando como principal a la primera de estas infusiones, se optó por la siguiente paleta de colores:



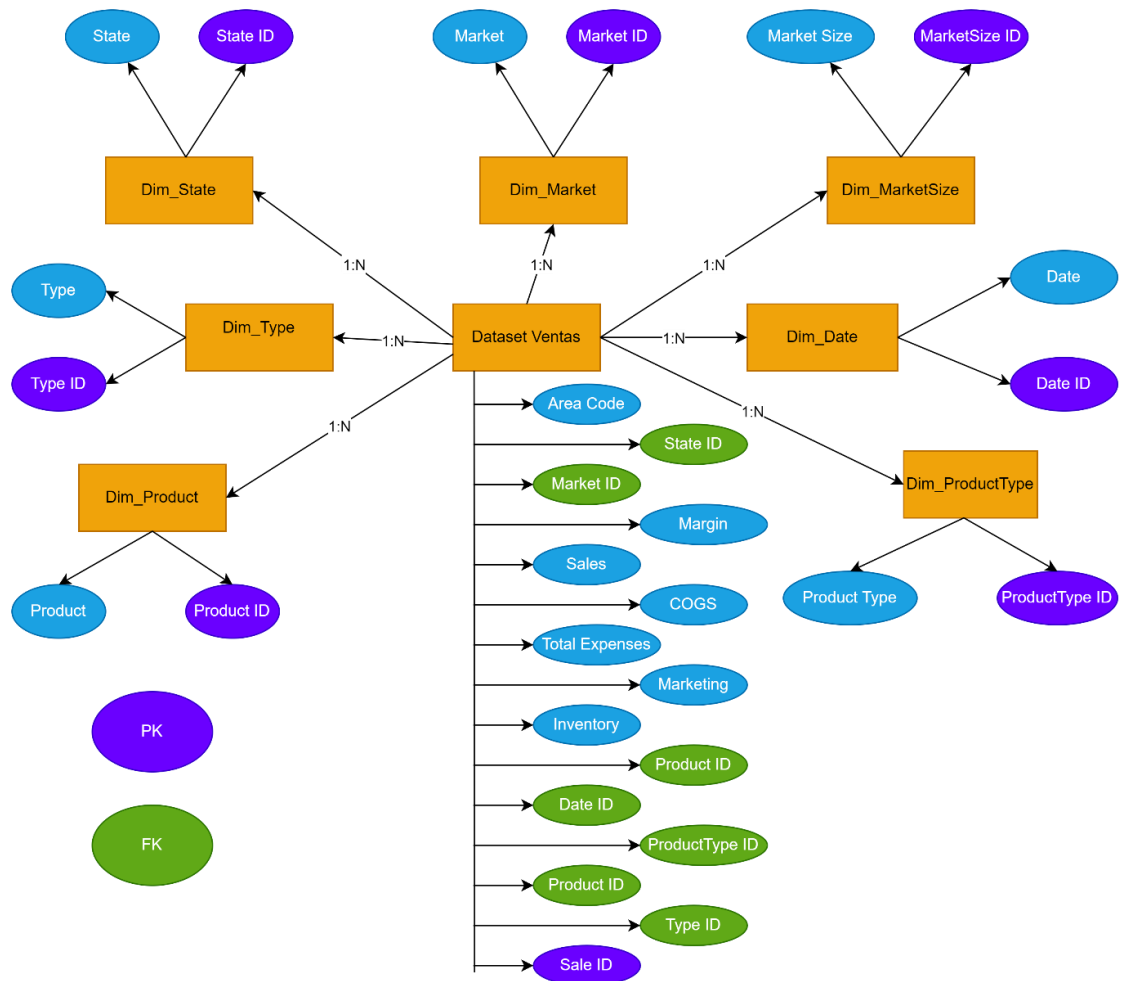
1.7 Tabla de versionado

A continuación, desarrollamos las tablas de versiones, en relación a las etapas que abordamos con las entregas intermedias y desafíos del curso.

VERSIÓN	ACCIONES
V1	Selección y carga de la base de datos Descripción de la temática Descripción de la hipótesis
V2	Normalización de la base de datos Realización del Diagrama Entidad-Relación Definición de alcance, usuario final y nivel de aplicación del análisis Definición y descripción de las tablas
V3	Desarrollo primer prototipo de tablero en Power BI Desarrollo de logo y paleta de colores Creación de la portada del tablero Diseño de los primeros mock ups Transformación de datos del tablero
V4	Realización de la Tabla calendario Realización de la tabla de medidas Descripción de las transformaciones realizadas Descripción de las medidas calculadas generadas y sus fórmulas
Versión Final	Corrección de las medidas calculadas Creación de botones de navegación entre solapas Creación de nuevas medidas calculadas Creación de gráficos, filtros y tarjetas informativas Creación de marcador de limpieza de filtros Creación de glosario Normalización y corrección de los parámetros visuales para su entrega.

2. Base de datos

2.1 Modelo Entidad-Relación



2.2 Listado de las tablas e información de sus campos

Dataset Ventas

Es la tabla principal y contiene información de las Ganancias, Margen, Ventas, Costos totales de producción, Gastos totales, Marketing e Inventario por cada Estado, Tamaño de mercado, Fecha, Producto, Tipo de producto y Sub-tipo de producto.

DataSet Ventas		
Campos	Tipo de dato	Clave
Sale ID	int	PK
Area Code	int	
State ID	int	FK
Market ID	int	FK
MarketSize ID	int	FK
Profit	int	
Margin	int	
Sales	int	
COGS	int	
Total Expenses	int	
Marketing	int	
Inventory	int	
Date ID	int	FK
ProductType ID	int	FK
Product ID	int	FK
Type ID	int	FK

Dim State

Es la tabla de dimensiones que define a cada estado con su número de ID.

Dim State		
Campos	Tipo de dato	Clave
State	varchar	
State ID	int	PK

Dim Market

Es la tabla de dimensiones que define a cada mercado con su número de ID.

Dim Market		
Campos	Tipo de dato	Clave
Market	varchar	
Market ID	int	PK

Dim Date

Es la tabla de dimensiones que define a cada fecha con su número de ID.

Dim Date		
Campos	Tipo de dato	Clave
Date	varchar	
Date ID	int	PK

Dim ProductType

Es la tabla de dimensiones que define a cada tipo de producto con su número de ID.

Dim Product Type		
Campos	Tipo de dato	Clave
Product Type	varchar	
ProductType ID	int	PK

Dim Type

Es la tabla de dimensiones que define a cada sub-tipo de producto con su número de ID.

Dim Type		
Campos	Tipo de dato	Clave
Type	varchar	
Type ID	int	PK

Dim Product

Es la tabla de dimensiones que define a cada producto con su número de ID.

Dim Product		
Campos	Tipo de dato	Clave
Product	varchar	
Product ID	int	PK

Dim MarketSize

Es la tabla de dimensiones que define los distintos tamaños de mercado con su número de ID.

Dim MarketSize		
Campos	Tipo de dato	Clave
Market Size	varchar	
Market Size ID	int	PK

3. Tablero en Power BI

3.1 Transformaciones realizadas

Al cargar la base de datos en Power BI es normal que se produzcan ciertos errores. A continuación, se detalla la corrección de los mismos.

En la tabla Dataset Ventas:

- Se eliminó la primer tabla producto ID (se encontraba duplicada)
- Se cambio el nombre de la tabla "ProductID_1" a "Product ID"
- Se modificaron los campos que en resumen figuraban como "Recuento" a "No resumir". Estos campos son Area Code, Date ID, Sale ID, Product ID.
- Se eliminaron las columnas 18, 19 y 20 que se encontraban vacias
- Se eliminaron las últimas dos filas que se encontraban vacias

En la tabla Dim_ProductType se modificó el nombre de "ProductType ID" a "ProductType" y el de "ProductType ID_1" a "Product Type ID". Además, se agregó una columna de nombre "Tipo" que define entre Café y Té para cada uno de los 4 tipos de producto. Por otro lado, también se cambió el campo ProductType ID que figuraban como "Recuento" a "No resumir".

3.2 Tabla calendario

Por otro lado, se creó la Tabla Calendario obteniendo de la fecha el año, el mes con número, el mes en letras, el mes en letras abreviado con 3 letras, y la semana.

La fórmula de la tabla calendario utilizada fue la siguiente:

let

Source = Dim_Date,

StartDate = Date.From(List.Min(Dim_Date[Date])), // Fecha de inicio

EndDate = Date.From(List.Max(Dim_Date[Date])), // Fecha de fin

// StartDate = #date(2023, 1, 1), // Fecha de inicio

// EndDate = #date(2023, 12, 31), // Fecha de fin

NumberOfDays = Duration.From(EndDate - StartDate) / #duration(1, 0, 0, 0),

DateList = List.Dates(StartDate, NumberOfDays, #duration(1, 0, 0, 0)),

// DateList = List.Dates(StartDate, NumberOfDays+1, #duration(1, 0, 0, 0)),

CalendarTable = Table.FromList(DateList, Splitter.SplitByNothing()),

RenameColumns = Table.RenameColumns(CalendarTable, {"Column1", "Fecha"}),

```

AddYear = Table.AddColumn(RenameColumns, "Año", each Date.Year([Fecha]),
Int64.Type),

AddMonthNumber = Table.AddColumn(AddYear, "Mes (Número)", each
Date.Month([Fecha]), Int64.Type),

AddMonthName = Table.AddColumn(AddMonthNumber, "Mes (Letra)", each
Date.ToText([Fecha], "MMMM")),

AddMonthAbbreviation = Table.AddColumn(AddMonthName, "Mes (Abreviado)",
each Date.ToText([Fecha], "MMM")),

AddWeek = Table.AddColumn(AddMonthAbbreviation, "Semana", each
Date.WeekOfYear([Fecha]))
in

AddWeek

```

3.3 Medidas calculadas generadas y sus fórmulas

En la tabla de medidas se realiza la comprobación de si los cálculos de margen y ganancia a través de medidas calculadas con variables son iguales a los campos de margen y ganancia.

Para la ganancia calculada se utiliza la siguiente fórmula:

```
Ganancia = [Margen] - sum('DataSet Ventas'[Total Expenses])
```

Por otro lado, se realiza la suma de todas las ganancias (campo que ya existía en la tabla) con la siguiente fórmula:

```
Ganancia_total = sum('DataSet Ventas'[Profit])
```

Para el margen calculado se utiliza la siguiente fórmula:

```
Margen = sum('DataSet Ventas'[Sales]) - sum('DataSet Ventas'[COGS])
```

Por último, se realiza la suma de todos los márgenes (campo que ya existía en la tabla) con la siguiente fórmula:

```
Margen_total = sum('DataSet Ventas'[Margin])
```

Como conclusión, había una diferencia en algunas ventas en particular en las que la ganancia calculada era diferente de la suma de las ganancias y el margen calculado era diferente a la suma de los márgenes. Al no poder encontrar de donde provenía esta

diferencia se procedió a modificar la base de datos (en Excel) para que efectivamente se cumplan las fórmulas anteriormente descritas:

- 1) Margen: Ventas – Costo de producción
- 2) Ganancia: Margen – Expensas totales

Por otro lado, se realiza un promedio de las ventas de Café con la siguiente fórmula:

```
Prom_Ventas_Café =  
calculate(average('DataSet Ventas'[Sales]), Dim_ProductType[Tipo ] = "Café")
```

Por último, se realiza un promedio de las ventas de Té con la siguiente fórmula:

```
Prom_Ventas_Té =  
calculate(average('DataSet Ventas'[Sales]),Dim_ProductType[Tipo ] = "Té")
```

3.4 Presentación del tablero

Solapa Portada:

Esta solapa es la portada del tablero que contiene el título del mismo, con su logo, el nombre del autor, la última actualización y los botones de navegación dentro del tablero.



Solapa Glosario

Esta solapa contiene una descripción del objetivo del tablero y se definen ciertos indicadores utilizados. Además, se describen los tipos y sub-tipos de infusiones en los que se basa el análisis.



GLOSARIO

El siguiente tablero tiene como principal objetivo analizar las ventas, margen y ganancia de Café y Té por tipo y sub-tipo en el mercado de Estados Unidos. También se analizan los datos en base a la ubicación de las ventas, diferenciando entre estados y mercados.

Algunas definiciones son:

Margen = Ventas - Costo de Producción

Ganancia = Margen - Expensas totales

Los tipos de café son: Espresso, Coffee.

Con sus sub-tipos: Decaf Espresso, Columbian, Caffé Mocha, Decaf Irish Cream, Caffé Late, Amaretto, Regular Espresso.

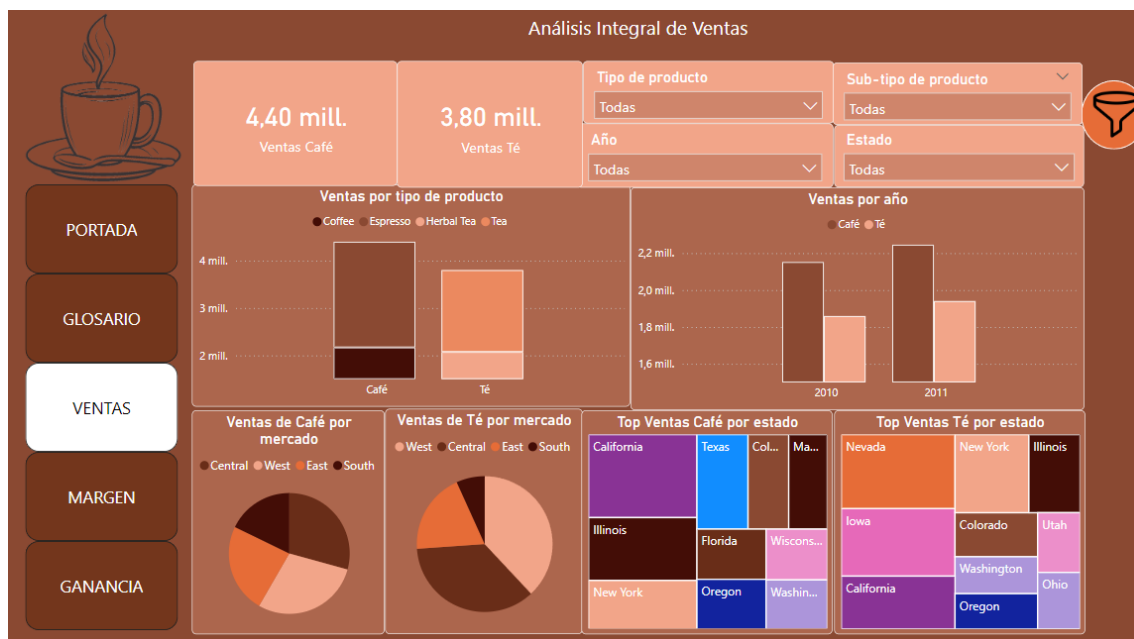
Los tipos de Té son: Herbal Tea, Tea.

Con sus sub-tipos: Lemon, Mint, Darjeeling, Green Tea, Chamomile, Earl Grey.

Solapa Ventas

En esta solapa se desarrolla un análisis integral de las ventas. La misma contiene tarjetas con la suma de las ventas para el Café y el Té; un gráfico de barras con la comparación de ventas por tipo y sub-tipo de Café y Té; un gráfico de barras comparando las ventas de Café y Té entre 2010 y 2011; dos gráficos circulares que muestran las ventas de Café y Té por mercado; dos gráficos Treemap con el top 10 de estados con mayor venta para cada infusión.

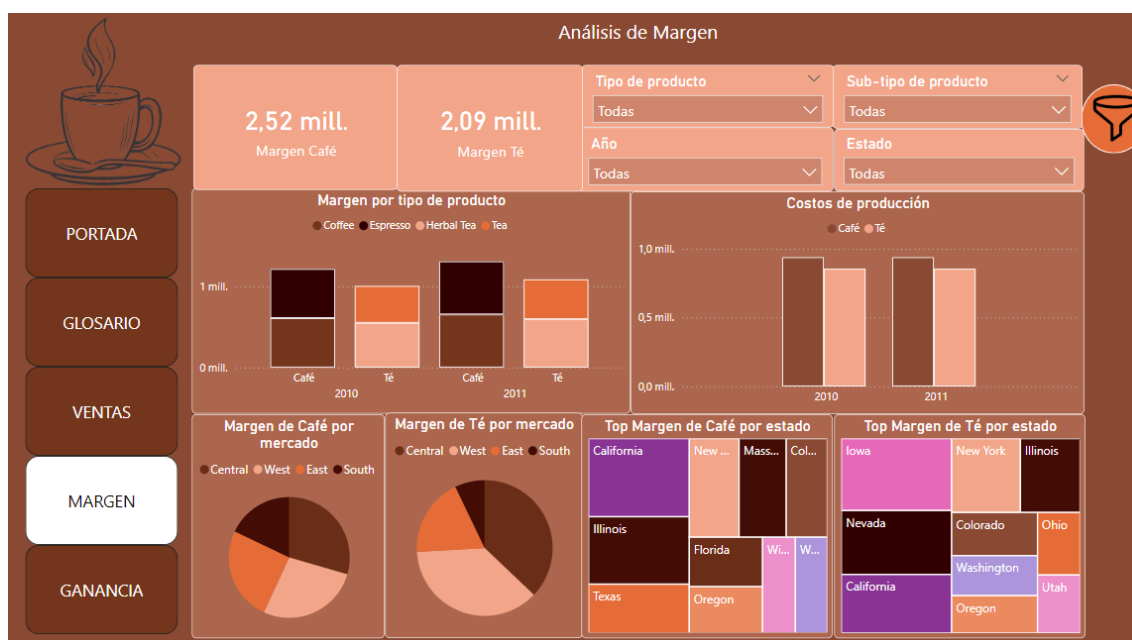
Por otro lado, el tablero permite al usuario filtrar la información por año, tipo de producto, sub-tipo de producto y estado de modo que los gráficos y tarjetas entreguen información específica para sacar conclusiones de las ventas por cada uno de estos segmentos. Además, posee un marcador para eliminar los filtros que facilita la interacción del usuario con el tablero.



Solapa Margen

En esta solapa se desarrolla un análisis integral de los márgenes. La misma contiene tarjetas con la suma de los márgenes para el Café y el Té; un gráfico de barras con la comparación de márgenes por año y entre tipo y sub-tipo de Café y Té; un gráfico de barras comparando los costos de producción de Café y Té entre 2010 y 2011; dos gráficos circulares que muestran los márgenes de Café y Té por mercado; dos gráficos Treemap con el top 10 de estados con mayor margen para cada infusión.

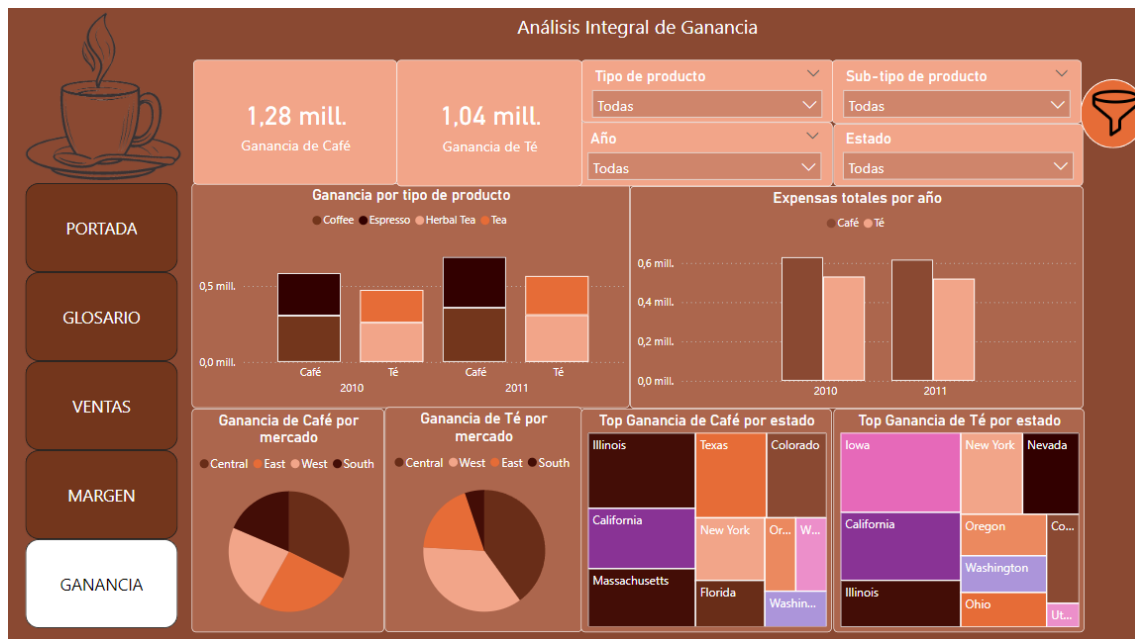
Por otro lado, el tablero permite al usuario filtrar la información por año, tipo de producto, sub-tipo de producto y estado de modo que los gráficos y tarjetas entreguen información específica para sacar conclusiones de los márgenes por cada uno de estos segmentos. Además, posee un marcador para eliminar los filtros que facilita la interacción del usuario con el tablero.



Solapa de Ganancia

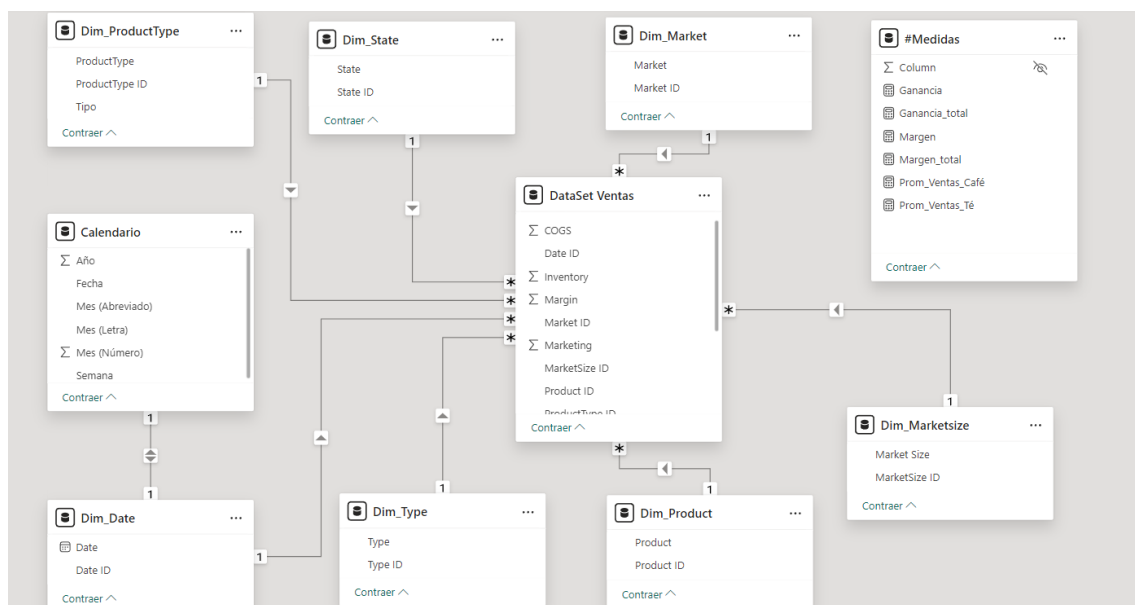
En esta solapa se desarrolla un análisis integral de las ganancias. La misma contiene tarjetas con la suma de las ganancias para el Café y el Té; un gráfico de barras con la comparación de ganancias por año y entre tipo y sub-tipo de Café y Té; un gráfico de barras comparando las expensas totales de Café y Té entre 2010 y 2011; dos gráficos circulares que muestran las ganancias de Café y Té por mercado; dos gráficos Treemap con el top 10 de estados con mayor ganancia para cada infusión.

Por otro lado, el tablero permite al usuario filtrar la información por año, tipo de producto, sub-tipo de producto y estado de modo que los gráficos y tarjetas entreguen información específica para sacar conclusiones de las ganancias por cada uno de estos segmentos. Además, posee un marcador para eliminar los filtros que facilita la interacción del usuario con el tablero.



3.5 Diagrama Entidad-Relación en Power BI

El programa Power BI realiza un diagrama de Entidad-Relación automático al cargar la base de datos en donde se pueden visualizar las diferentes tablas, con sus respectivas relaciones, lo que permite comprobar que no haya errores, para un correcto funcionamiento del tablero.



4. Futuras líneas de investigación

En primer lugar, se podría cargar información de la venta, margen y ganancia de los siguientes años y, de este modo, sacar conclusiones mas robustas del comportamiento de la empresa en el mercado estadounidense.

Por otro lado, se podría evaluar el inventario de mercadería por sub-tipo de producto para corroborar de que se posee lo suficiente para no llegar al punto de falta de stock. Sobre todo, para los productos que poseen mayor nivel de ventas.

Además, se podría evaluar modificaciones en las ventas de los productos que arrojaron un saldo negativo de ganancia. Los motivos podrían llegar a ser complejos, pero en base a los datos que poseemos, el balance positivo podría alcanzarse con un incremento en la campaña de marketing que conlleve mayores ventas o una modificación de alguno de los factores que hacen que las expensas totales sean mayores a los márgenes.