

Proyecto final data analitycs 1

Análisis de ventas SuperShop en Power BI



23 de octubre de 2024

Contenido

[**Visualización** 2](#_Toc180857242)

[TABLA VERSIONADA 2](#_Toc180857243)

[OBJETIVO 2](#_Toc180857244)

[ALCANCE 2](#_Toc180857245)

[USUARIO FINAL Y APLICACIÓN DEL ANALISIS 2](#_Toc180857246)

[SOLAPAS 3](#_Toc180857247)

[DIAGRAMA E-R. 5](#_Toc180857248)

[LISTADO DE TABLAS. 6](#_Toc180857249)

[**Base de datos** 8](#_Toc180857250)

[DESCRIPCION DE LA TEMATICA DE LOS DATOS 8](#_Toc180857251)

[INDICADORES Y SEGMENTADORES UTILIZADOS 8](#_Toc180857252)

[DIAGRAMA DEL MODELADO DE POWER BI 8](#_Toc180857253)

[TRANSFORMACIONES REALIZADAS 9](#_Toc180857254)

[FUNCIONES EN MEDIDAS Y COLUMNAS 10](#_Toc180857255)

[CONCLUSION DE CADA SOLAPA 11](#_Toc180857256)

[FUTURAS LINEAS 12](#_Toc180857257)

# **Visualización**

## TABLA VERSIONADA

|  |  |  |
| --- | --- | --- |
|  | Instancia |  |
| 1 | 13/08/24 | Primera entrega |
| 2 | 27/08/24 | Segunda entrega |
| 3 | 03/10/24 | Tercera entrega |
| 4 | 22/10/24 | Entrega final |

## OBJETIVO

En este análisis podremos descubrir la eficiencia de esta Supertienda estadounidense y sus posibilidades de mejora. Para esto se analizarán las tendencias de ventas mediante las columnas fecha de pedidos y ventas buscando los periodos de alta y baja demanda, el rendimiento de los productos mediante las columnas de productos y ventas buscando los más vendidos, la distribución de las ventas en diferentes regiones, estados, ciudades buscando identificar las ubicaciones con las ventas más altas y más bajas, También podríamos analizar con las columnas de pedidos, ventas y segmento el tipo de cliente, identificando así clientes de alto valor.

## ALCANCE

El alcance de este proyecto sería interno, ya que con el análisis que estamos realizando buscamos mejorar la eficiencia de las ventas, por lo cual es una información que se va a manejar de manera interna dentro de la compañía/empresa.

## USUARIO FINAL Y APLICACIÓN DEL ANALISIS

Nivel operativo: Dentro del nivel operativo los usuarios finales serían los equipos de operaciones, gerentes y atención al cliente. La forma de aplicación seria a través de operaciones diarias y decisiones inmediatas.

Nivel táctico: Dentro del nivel táctico los usuarios finales serían las áreas de marketing, logística y analistas de ventas. La forma de aplicación seria optimizando las ventas y logística, evaluando las campañas de marketing y la segmentación de clientes.

Nivel estratégico: Dentro del nivel estratégico los usuarios finales serían los altos mandos. La forma de aplicación estaría dado a través de planificaciones a largo plazo y crecimiento del negocio.

## SOLAPAS



1. Portada: En esta solapa inicial tenemos la portada con su respectivo título, nombre del autor del proyecto, fecha de última actualización del proyecto y los botones para navegar entre solapas.

Interfaz de usuario gráfica

Descripción generada automáticamente

1. País: En esta segunda solapa contamos con información sobre las ventas por ciudad, región y trimestres lo que nos permite analizar de una manera mas eficiente las zonas con más o menos cantidad de ventas y total de ingresos.

Interfaz de usuario gráfica

Descripción generada automáticamente

1. Productos: En esta tercera solapa tenemos toda la información referida a los productos ya sean sus categorías, subcategorías, máximo de ventas por productos entre otros, con estos datos podremos analizar de una manera mas sencilla cuales son nuestros mejores y peores productos en cuanto a venta.

Interfaz de usuario gráfica

Descripción generada automáticamente

1. Clientes: En esta cuarta y ultima solapa contamos con toda la información sobre nuestros clientes, cantidad de clientes, genero, edad, segmento entre otros, contando con esta información se nos facilita encontrar cuales son nuestros mejores clientes en cuanto a cantidad de ventas y a montos gastados en cada compra.

## DIAGRAMA E-R.

Diagrama

Descripción generada automáticamente

## LISTADO DE TABLAS.

* Ventas:

Esta tabla contiene los valores, las ordenes, los clientes y los productos de cada venta efectuada.

* PK: Ventas\_id
* FK: Pedidos\_id
* FK: Clientes\_id
* FK: Producto\_id
* Pedidos:

Esta tabla contiene los números de orden, fecha de orden, fecha de envió, métodos de envió y datos de clientes.

* PK: Pedidos\_id
* Clientes:

Esta tabla contiene el id, el nombre, el segmento y el país de cada cliente.

* PK: Clientes\_id
* FK: Estado
* Producto:

Esta tabla contiene el id, la categoría, la subcategoría y el nombre de cada producto vendido.

* PK: Producto\_id
* País:

Esta tabla contiene el país, la ciudad, el estado, el código postal y la región de cada cliente.

* PK: Estado

|  |  |  |
| --- | --- | --- |
| Sales | | |
| Campo | Tipo de campo | Tipo de clave |
| Ventas\_id | int | PK |
| Pedidos\_id | int | FK |
| Clientes\_id | Varchar (15) | FK |
| Ventas | int | - |
| Producto\_id | Varchar (25) | FK |

|  |  |  |
| --- | --- | --- |
| Order | | |
| Campo | Tipo de campo | Tipo de clave |
| Pedidos\_id | Varchar (20) | PK |
| Fecha\_pedidos | Datetime | - |
| Fecha\_envios | Datetime | - |
| Modo\_envio | Varchar (15) | - |

|  |  |  |
| --- | --- | --- |
| Customer | | |
| Campo | Tipo de campo | Tipo de clave |
| Clientes\_id | Varchar(15) | PK |
| Nombre\_cliente | Varchar (30) | - |
| Segmento | Varchar (20) | - |

|  |  |  |
| --- | --- | --- |
| Product | | |
| Campo | Tipo de campo | Tipo de clave |
| Producto\_id | Varchar(25) | PK |
| Categoría | Varchar (20) | - |
| Sub\_categoría | Varchar (15) | - |
| Nombre\_producto | Varchar (110) | - |

|  |  |  |
| --- | --- | --- |
| Country | | |
| Campo | Tipo de campo | Tipo de clave |
| Ciudad | Varchar (25) | - |
| Estado | Varchar (20) | PK |
| Codigo\_postal | int | - |
| Región | Varchar(15) | - |
| País | Varchar (20) | - |

# **Base de datos**

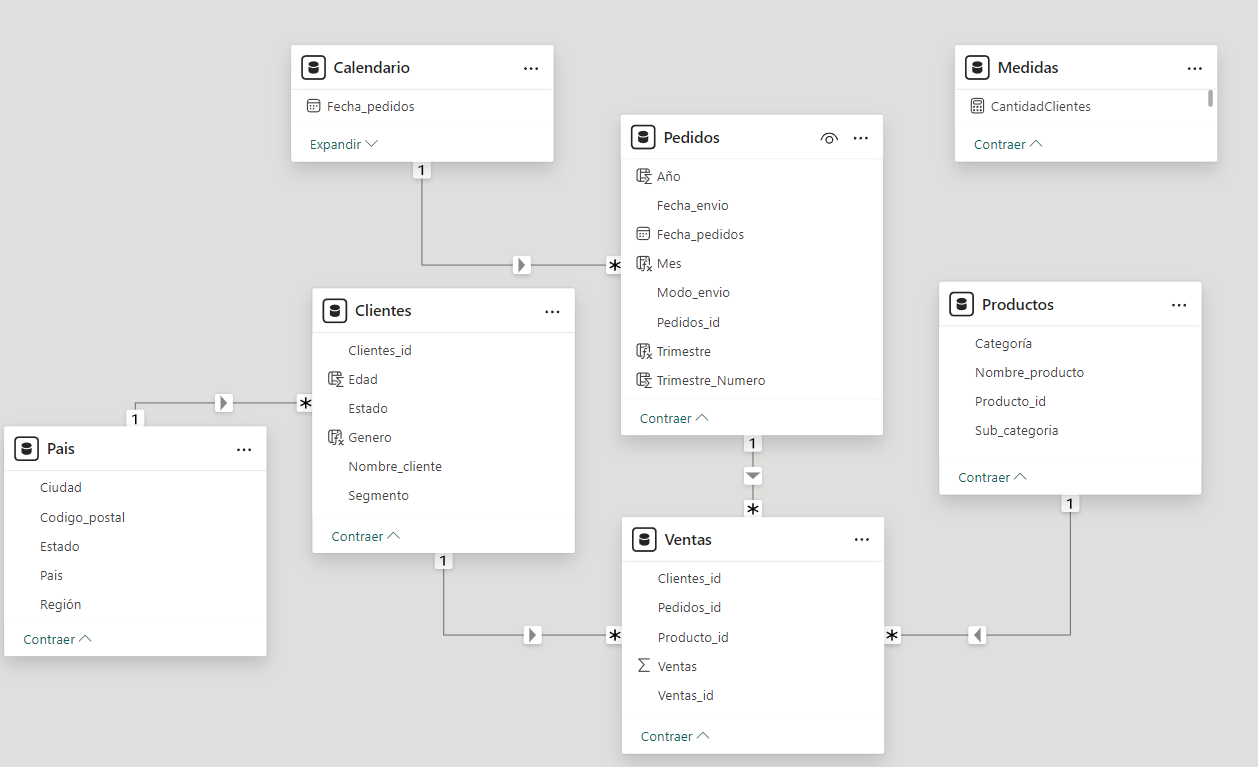
## DESCRIPCION DE LA TEMATICA DE LOS DATOS

Para este proyecto, voy a tomar información de las ventas online de una supertienda de estados unidos. Además, tendremos en cuenta otros datos particulares de cada venta como fechas de ventas, ventas, fechas de envíos, modos de envíos, datos de clientes, datos de productos. Teniendo la primera venta registrada en nuestro dataset el día tres de enero del año dos mil quince, luego el análisis de las ventas se extiende hasta el día treinta de diciembre del año dos mil dieciocho.

## INDICADORES Y SEGMENTADORES UTILIZADOS

Para este análisis vamos a utilizar algunos indicadores de tarjeta como cantidad de ventas y monto total para analizar por ciudad, estado, mes, año o región cuando vayamos cambiando con los segmentadores o datos introducidos en los gráficos. También se utilizaron los mismos tres segmentadores en las 3 solapas de análisis, los cuales serán año, mes y región para poder analizar con una mayor precisión en el tiempo y por región los datos que tenemos ya sean de clientes productos o ventas.

## DIAGRAMA DEL MODELADO DE POWER BI



## TRANSFORMACIONES REALIZADAS

* Se obtuvieron datos de un libro de Excel
* Con el transformador de datos:
* Se duplico la table madre 4 veces para poder hacer las otras 4 tablas (Productos, Clientes, País, Pedidos).
* Se eliminaron las columnas correspondientes de cada tabla y quitaron duplicados correspondientes.
* Se cambiaron los formatos de los números de ventas y se quitaron decimales.
* Se completo un valor NULL en código postal.
* Se agrego una columna índice en la tabla ventas.
* El dataset estaba en inglés, se cambió todos los datos a español, los títulos cambiándolos y los datos con remplazar valor.
* Se agrego una tabla calendario duplicando la tabla que contenía las fechas, quitando el resto de las columnas y quitando duplicados de las fechas.
* Se hicieron las uniones correspondientes entre tablas, con sus respectivos PK y FK.

* Se uso introducir datos para generar una tabla vacía, la cual se llamó medidas y fui agrupando en esa tabla todas las medidas que fui calculando
* En la tabla pedidos se agregaron 4 columnas nuevas una que muestra el año en números, una que muestra el mes en letras, una que muestra el trimestre en letras y una que muestra el trimestre en números para que me quede bien ordenado el grafico de trimestres.
* En la tabla Clientes se agregaron 2 columnas una de edad en números y otra de género masculino y femenino.
* Luego se generaron todas las medidas necesarias para poder trabajarlas en el dashboard con los gráficos.
* Luego inicié por la portada poniéndole nombre del autor, título del dashboard, botones de navegación (con símbolos), fecha de elaboración y el lienzo de la hoja es el logo que figuraba en la página de kaggle donde extraje mi dataset. Se configuro toda la portada ya sea colores, tamaño y tipo de letras, vínculos de botones y transparencia del lienzo para mejor visualización.
* Se eligió el tema en colores azules y celestes ya que son los colores del logo que figuraba en kaggle.
* Se hicieron tres hojas aparte de la portada las cuales fueron País, Producto y Clientes.
* En las 3 hojas se aplicaron los mismos filtros (año, mes, región) y los botones de navegación a las distintas hojas y portada con distintos símbolos.
* Cada hoja tiene su enfoque:

* País: está enfocada en mostrar datos de ventas y clientes en base a las regiones ciudades y estados, como cantidad de ventas por estado, monto total por estado, top 10 ventas por ciudad, entre otras. Se aplicaron dos marcadores uno para los datos generales y otro para mostrar unos top 10.
* Producto: está enfocada en mostrar datos en base a los productos, categorías y subcategorías, como venta máxima por categoría, ventas totales por categoría y subcategoría, entre otros. Se aplicaron 3 marcadores uno que muestra datos generales, otros dos que muestran un top 10 positivo y top 10 negativo.
* Clientes: está enfocado en mostrar datos de clientes en base segmento, genero, edad, como cantidad de clientes, cantidad de clientes hombres y mujeres, promedio de ventas por segmento, entre otros. Se aplicaron 2 marcadores uno general con todos los datos y otro que muestra un top 6 positivo de cantidad de clientes por ciudad.

## FUNCIONES EN MEDIDAS Y COLUMNAS

* Columnas:
  + - Año **“YEAR ('Pedidos'[Fecha pedidos])”**
    - Mes **“FORMAT ('Pedidos'[Fecha pedidos],"mmmm")”**
    - Trimestre **“SWITCH( TRUE(), 'Pedidos'[Mes]= "Enero"|| 'Pedidos'[Mes]= "Febrero" ||'Pedidos'[Mes]= "Marzo" , "Primer trimestre", 'Pedidos'[Mes]= "Abril"||'Pedidos'[Mes]= "Mayo"|| 'Pedidos'[Mes]= "Junio" , "Segundo trimestre", 'Pedidos'[Mes]= "Julio" ||'Pedidos'[Mes]= "Agosto"||'Pedidos'[Mes]= "Septiembre" , "Tercer trimestre", 'Pedidos'[Mes]= "Octubre" ||'Pedidos'[Mes]= "Noviembre" ||'Pedidos'[Mes]= "Diciembre" , "Cuarto trimestre")”**
    - Trimestre\_numero **“QUARTER (Pedidos [Fecha pedidos])”.**
    - Edad **“RANDBETWEEN (18,65)”**
* Medidas:
  + - VentasTotales **SUM (Ventas [Ventas])**
    - VentaMaxima **MAX (Ventas [Ventas])**
    - CantidadVentas **COUNT (Ventas [Ventas\_id])**
    - PromedioVentas **AVERAGE (Ventas [Ventas])**
    - PromedioEdad **AVERAGE (Clientes [Edad])**
    - CantidadClientes

**DISTINCTCOUNT(Clientes [Nombre\_cliente])**

* + - Hombres
    - **COUNTAX(FILTER(Clientes,Clientes[Genero]="Masculino"), Clientes [Genero])**
    - Mujeres

**COUNTAX(FILTER(Clientes,Clientes[Genero]="Femenino"), Clientes [Genero])**

## CONCLUSION DE CADA SOLAPA

País:

En la primer solapa gracias a los gráficos y los segmentadores podemos notar que todos los años aumentaron las cantidad de ventas y el dinero que ingreso por dichas ventas, también podemos notar cual es la región con mas ventas, en este caso es la región Oeste aunque no es la que mejor promedio tiene, si nos referimos al promedio de venta en cuanto a cantidad de ventas y monto ingresado seria la región Central, también podemos ver el estado con mas ventas a lo largo del tiempo que fue Los Ángeles menos en 2016 que fue Fort Worth. Por último, podemos ver que el tercer trimestre fue por lo general el trimestre en el que más se vendió, aunque si queremos verlo por región o año especifico puede variar.

Producto:

En la segunda solapa se puede apreciar distintos datos sobre los productos los cuales nos ayudan a analizar las categorías y subcategorías más vendidas en cuanto a cantidad y monto total de cada producto las cuales no coinciden, ya que la categoría con mas cantidad de ventas es la de materiales de oficina y la subcategoría es carpetas, por otro lado la categoría con mas monto total ingresado es Muebles y la subcategoría es Librerías, también podemos apreciar los productos mas vendidos y menos vendidos con el fin de mejorar los números de los menos vendidos y aprovechar los más vendidos.

Clientes:

En esta última solapa podemos notar todos los datos acerca de los clientes los cuales nos indican que tenemos una cantidad total actualmente de 793 clientes, que el promedio de edad de nuestros clientes es de 42 años, que tenemos más clientes masculinos (554) que femeninos (239), que el segmento que mas compra es Consumidor y el que menos Oficina en casa, que el método de envió más utilizado es el de Clase estándar mientras que el menos utilizado es Mismo día, también que a pesar de tener mas de el doble de clientes masculinos que femeninos en cuento a promedio de ventas está por encima el genero femenino. También podemos notar que nuestras tres regiones mas fuertes son las de California, New York y Texas mientras que las dos peores regiones en las cuales se debería trabajar mas son Illinois y Washington. Por último, tenemos un ranking de clientes en cuanto a cantidad de ventas y monto total, el cual no coinciden el top 1 demostrando que quien mas compras hizo no es quien mas dinero gasto, teniendo en primer lugar en cantidad de ventas a William Brown con 35 ventas y Anna Gayman en monto total con 25 millones, estos datos nos sirven para poder saber quienes son nuestros clientes mas valiosos y en quienes hay que trabajar e indagar un poco más.

## FUTURAS LINEAS

En mi opinión, creo que se podrían obtener mas datos para hacer un mejor análisis como costo de cada producto para poder sacar el profit de cada producto, categoría y subcategoría. También un dato que podría ser relevante es los métodos de pagos utilizados por los clientes con el cual podríamos analizar si hay un método de pago que sea mas utilizado o muy poco utilizado, por ultimo en cuanto a la parte de productos podríamos hacer un stock con sus costos, cantidades y ventas, para poder analizar las ganancias o hasta quizás si en algún trayecto en la línea temporal hubo perdidas para la empresa/compañía.

Lautaro Falco