

“TechXport” : Entrega de Proyecto Final - Data Analytics - Comisión 39880

DIEGO MURO

MARZO 2023

Versión Documento No. 1.0

Creación del Documento: 03/03/2023

Contenido

Objetivo del proyecto.....	2
Alcance del proyecto	2
Usuarios de la información.....	2
Descripción de la temática del proyecto (Análisis bajo Metodología SMART)	3
Hipótesis del proyecto (Análisis y situaciones problemáticas)	4
DER	5
Listado de tablas	6
Detalle de columnas por tabla.....	7
Datasets.....	11
Herramientas y tecnologías utilizadas	11
Modelo relacional en Power Bi.....	11
Medidas calculadas	13
Subcarpeta Clientes:.....	13
Subcarpeta Datos Estadísticos:.....	13
Subcarpeta KPI Retornos:.....	13
Subcarpeta Ordenes (Ventas – Retornos):.....	13
Subcarpeta Rentabilidad:.....	14
Subcarpeta Retornos:	14
Subcarpeta Ventas:.....	14
Subcarpeta Ventas comparadas:	14
Visualización del reporte:	15
Sección Portada:.....	15
Sección Resumen de Ventas:	16
Sección Resumen de retornos:.....	16
Sección Análisis de Clientes:.....	17
Sección Análisis de Rentabilidad:.....	17
Sección Datos Estadísticos:	18

Objetivo del proyecto

Analizar una base de datos de ventas de un negocio de exportación de productos de las categorías de muebles, suministros de oficina y tecnología para identificar patrones de ventas por ciudad, vendedor y producto, y generar informes que permitan tomar decisiones acertadas sobre inventario, oferta de productos, marketing y ventas.

A alcance del proyecto

El proyecto incluirá la identificación y selección de la base de datos de ventas a analizar, la extracción y limpieza de los datos, el análisis de los datos mediante técnicas de data analytics y la generación de informes con los resultados obtenidos (Power Bi). El proyecto se realizará dentro de un plazo definido y se establecerán entregables para monitorear y controlar el progreso del mismo.

Usuarios de la información

Con respecto a los usuarios que consumirán la información generada, se incluirán los siguientes puntos:

- Identificación de los usuarios clave (Key Users): se identificarán los usuarios que tendrán acceso a la información generada a partir del análisis de la base de datos de ventas, considerando sus necesidades y roles dentro de la organización.
- Diseño de informes y dashboards: se diseñarán informes y dashboards en Power Bi que permitan a los usuarios visualizar de manera clara y concisa la información relevante para su área de trabajo.
- Validación de los informes: se realizarán pruebas y validaciones de los informes y dashboards generados para asegurar su calidad y exactitud.
- Capacitación de usuarios: se capacitará a los usuarios en el uso de los informes y dashboards generados, con el fin de que puedan interpretar correctamente la información y tomar decisiones informadas.

Proyecto final – Diego Muro

- Actualización de informes: se establecerá un plan de actualización de los informes y dashboards generados para garantizar que la información se mantenga actualizada y relevante para los usuarios.

Los usuarios clave que tendrán acceso a la información generada a partir del análisis de la base de datos de ventas incluyen:

- ✓ Gerente General: Para ver informes globales de ventas y análisis de tendencias para tomar decisiones **estratégicas** sobre el negocio.
- ✓ Gerente de Ventas: Necesita información sobre las ventas por ciudad, vendedor y producto para tomar decisiones informadas sobre la estrategia de ventas y el rendimiento de los vendedores.
- ✓ Gerente de Marketing: Necesita información sobre las zonas geográficas y productos con mayor demanda para enfocar los esfuerzos de marketing y ventas.
- ✓ Analista de Datos: Necesita acceso a la base de datos de ventas y los informes generados para realizar análisis más detallados y extraer información adicional.
- ✓ Vendedores: Necesitan acceso a información sobre sus propias ventas y desempeño para evaluar su rendimiento y establecer objetivos de ventas.

En general, los usuarios clave dependerán de las necesidades y roles específicos dentro de la organización, y podrían incluir desde ejecutivos de alto nivel (Nivel Estratégico) hasta personal de ventas y marketing (Nivel Operativo y táctico). Es importante identificar a estos usuarios y diseñar informes y dashboards que satisfagan sus necesidades específicas.

Descripción de la temática del proyecto (Análisis bajo Metodología SMART)

Basado en la metodología SMART podemos describir el proyecto de la siguiente manera:

Proyecto final – Diego Muro

- 1- **Específico:** El proyecto tiene un objetivo claro y definido: Analizar una base de datos de ventas de un negocio de exportación de productos de las categorías de muebles, suministros de oficina y tecnología y que además me permita analizar las ventas por ciudad, vendedor, producto, etc., y mostrar las cantidades y los resultados de las ventas, así como las devoluciones y sus causas.
- 2- **Medible:** El éxito del proyecto se puede medir mediante el cumplimiento de los siguientes indicadores de desempeño:
 - La cantidad de ventas registradas en la base de datos.
 - La exactitud y coherencia de los datos de la base de datos.
 - La calidad y cantidad de los informes generados por la base de datos.
 - La satisfacción de los usuarios con la base de datos.
- 3- **Alcanzable:** El proyecto es alcanzable porque cuenta con los recursos necesarios, tales como el analista de datos que realizara el análisis, el software de gestión de bases de datos, el hardware y la infraestructura necesaria.
- 4- **Relevante:** El proyecto es relevante para el negocio porque le permitirá:
 - Conocer y analizar su desempeño de ventas.
 - Tomar decisiones informadas sobre su inventario y oferta de productos.
 - Identificar y resolver problemas de calidad o de servicio que afectan la satisfacción del cliente.
 - Enfocar sus esfuerzos de marketing y ventas en las zonas geográficas y productos con mayor demanda.
- 5- **A tiempo:** El proyecto tiene un plazo definido para su realización (cursada), el cual se debe cumplir para lograr los objetivos del negocio en el tiempo previsto. Además, se establecerán entregables para monitorear y controlar el progreso del proyecto (entregables) y asegurar su cumplimiento en tiempo y forma.

Hipótesis del proyecto (Análisis y situaciones problemáticas)

El alcance en el cual enfocare el análisis se basará en Ciudades, Vendedores, Productos, ventas y profit, Devoluciones y Periodos:

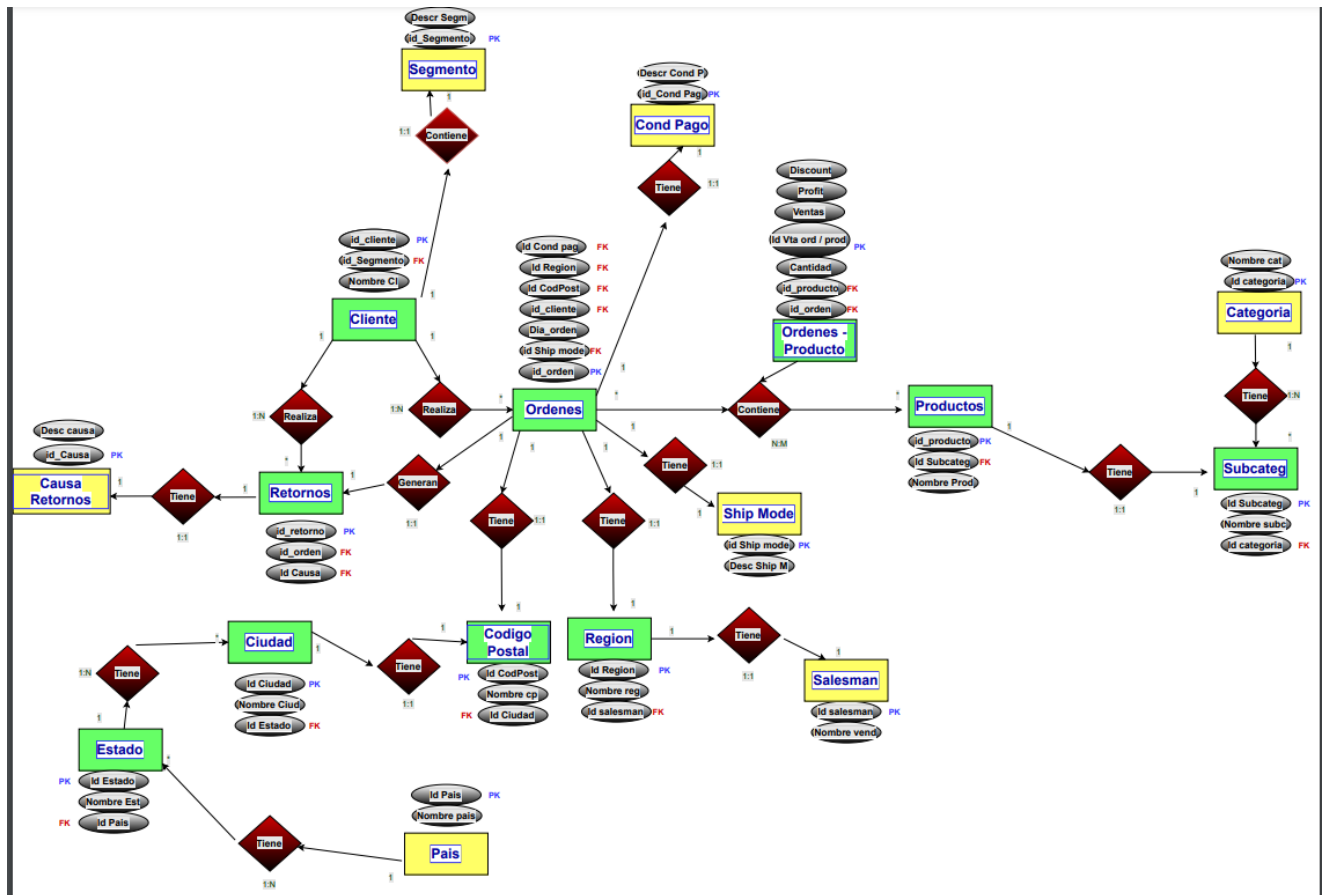
Proyecto final – Diego Muro

- El análisis de las **ventas por ciudad** permitirá identificar las zonas geográficas con mayor demanda de los productos, lo que permitirá al negocio enfocar sus esfuerzos de marketing y ventas en esas áreas.
- El análisis de las **ventas por vendedor** permitirá identificar a los vendedores más efectivos y los que necesitan más capacitación, lo que permitirá al negocio mejorar su desempeño de ventas.
- El análisis de las **ventas por producto** permitirá identificar los productos más vendidos y los menos vendidos, lo que permitirá al negocio tomar decisiones sobre el inventario y la oferta de productos.
- El análisis de las **devoluciones y sus causas** permitirá identificar los problemas de calidad o de servicio que afectan la satisfacción del cliente, lo que permitirá al negocio mejorar su calidad y su servicio al cliente.
- El análisis de las **ventas y profit** permitirá identificar adonde enfocar los equipos de marketing y de administración de ventas, respecto a promociones y acciones puntuales.
- El análisis de las **ventas por periodo** nos permitirá ver la evolución de los puntos anteriormente mencionados (performance de Ciudad, vendedores, productos, disminución de devoluciones)

DER

El Diagrama de entidad relación del proyecto tambien se encuentra en la carpeta así nombrada en el repositorio.

Proyecto final – Diego Muro



Listado de tablas

Se describen a continuación las tablas que contiene la BBDD del proyecto (Archivo completo en el repositorio, carpeta BBDD del proyecto final)

Proyecto final – Diego Muro

Comercial Global

		PK							
1	SEGMENTO (TABLA DIMENS)	id_Segm	descrip_seg						
		PK							
2	CONDICION DE PAGO (TABLA DIMENS)	id_payment	descrip_pay						
		PK							
3	CAUSA DE RETORNOS (TABLA DIMENS)	id_cause	descrip_causeret						
		PK							
4	CATEGORIA (TABLA DIMENS)	id_cat	name_cat						
		PK							
5	PAIS (TABLA DIMENS)	id_country	name_count						
		PK							
6	SHIP MODE (TABLA DIMENS)	id_ship	descrip_ship						
		PK							
7	VENDEDOR (TABLA DIMENS)	id_salesman	name_salesman						
		PK		FK					
8	CLIENTE (TABLA HECHO)	id_cust	name_cust	id_Segm					
		PK		FK					
9	PRODUCTOS (TABLA HECHO)	id_product	name_prod	id_subcateg					
		PK		FK	FK	FK	FK	FK	
10	ORDENES (TABLA HECHO)	id_order	order_date	id_ship	id_cust	id_pc	id_region	id_payment	
		PK	FK	FK					
11	ORDENES - PRODUCTO (TABLA HECHO)	id_vta_order_prod	id_order	id_product	sales	qty	discount	profit	
		PK	FK	FK					
12	RETORNOS (TABLA HECHO)	id_retorno	id_order	id_cause					
		PK		FK					
13	ESTADO (TABLA HECHO)	id_state	name_state	id_country					
		PK		FK					
14	CIUDAD (TABLA HECHO)	id_city	name_city	id_state					
		PK		FK					
15	SUB-CATEGORIA (TABLA HECHO)	id_subcateg	name_subcat	id_cat					
		PK		FK					
16	REGION (TABLA HECHO)	id_region	name_region	id_salesman					
		PK		FK					
17	CODIGO POSTAL (TABLA HECHO)	id_pc	name_pc	id_city					

Detalle de columnas por tabla

A continuación el detalle de columnas por tabla, también se encuentra en el repositorio (carpeta BBDD del proyecto final)

Proyecto final – Diego Muro

TABLA 1				
Nombre: SEGMENTO (TABLA DIMENSIONAL)				
Descripción: Datos de los segmentos de venta (Mayorista / Minorista, etc)				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_Segm	Id de Segmento	Int	Identificado de segmento del cliente
	descrip_seg	Descripción de segmento	Varchar(100)	Segmento del cliente
TABLA 4				
Nombre: CATEGORIA (TABLA DIMENSIONAL)				
Descripción: Datos de categoría de cada producto				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_categ	Id de Categoría	Int	Código de categoría
	name_cat	Nombre de categoría	Varchar(100)	Descripción de categoría
TABLA 7				
Nombre: VENDEDOR (TABLA DIMENSIONAL)				
Descripción: Datos de los vendedores asignados a cada región				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_salesman	Id de vendedor	Int	Código de vendedor
	name_salesman	Nombre de vendedor	Varchar(100)	Nombre del vendedor
TABLA 2				
Nombre: CONDICION DE PAGO (TABLA DIMENSIONAL)				
Descripción: Datos de las condiciones de venta (Contado, transferencia, cheque)				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_payment	Id de condición pago	Int	Identificado de condición pago
	descrip_pay	Descripción condición pago	Varchar(100)	Detalle de condición de pago
TABLA 5				
Nombre: PAIS (TABLA DIMENSIONAL)				
Descripción: Datos de Países adonde se envían productos				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_country	Id del País	Int	Código de país
	name_count	Nombre del país	Varchar(100)	Descripción de País
TABLA 8				
Nombre: CLIENTE (TABLA DE HECHO)				
Descripción: Datos de todos los clientes				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_cust	Id de cliente	Varchar(100)	Identificador del cliente
	name_cust	Nombre de cliente	Varchar(100)	Nombre del cliente
FK	id_Segm	Id de Segmento	Int	Identificado de segmento del cliente

Proyecto final – Diego Muro

TABLA 3				
Nombre: CAUSA DE RETORNOS (TABLA DIMENSIONAL)				
Descripcion: Datos de los causas de rechazo				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_cause	Causa de retorno	Int	Identificador de causa de retorno
	descrip_causeret	Descripcion de causa retorno	Varchar(100)	Detalle de causa del retorno
TABLA 6				
Nombre: SHIP MODE (TABLA DIMENSIONAL)				
Descripcion: Datos de modo de entrega de pedidos				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_ship	Id de Ship Mode	Int	Identificador de Ship Mode
	descrip_ship	Descripcion de ship mode	Varchar(100)	Detalle de modo de entrega
TABLA 9				
Nombre: PRODUCTOS (TABLA DE HECHO)				
Descripcion: Datos de productos				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_product	Id de producto	Varchar(100)	Identificador de producto
	name_prod	Nombre de producto	Varchar(100)	Descripcion del producto
FK	id_subcateg	Id de Subcategoria	Int	Codigo de subcategoria
TABLA 10				
Nombre: ORDENES (TABLA DE HECHO)				
Descripcion: Datos de todas las ventas				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_order	Id de orden	Varchar(100)	Identificador de Orden
	order_date	Dia de la orden	Date	Dia del pedido
FK	id_ship	Id de Ship Mode	Int	Identificador de Ship Mode
FK	id_cust	Id de cliente	Varchar(100)	Identificador del cliente
FK	id_pc	Id de Codigo postal	Int	Id de Codigo Postal
FK	id_region	Id de Region	Int	Codigo de region
FK	id_payment	Id de condicion pago	Int	Identificado de condicion pago
TABLA 13				
Nombre: ESTADO (TABLA DE HECHO)				
Descripcion: Datos de Estados adonde se envian productos				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_state	Id del Estado	Int	Codigo de Estado
	name_state	Nombre del Estado	Varchar(100)	Descripcion del Estado
FK	id_country	Id del Pais	Int	Codigo de pais
TABLA 16				
Nombre: REGION (TABLA DE HECHO)				
Descripcion: Datos de region donde se vende el producto				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_region	Id de Region	Int	Codigo de region
	name_region	Nombre de la region	Varchar(100)	Descripcion de region
FK	id_salesman	Id de vendedor	Int	Codigo de vendedor

Proyecto final – Diego Muro

Comercial Global

TABLA 11				
Nombre: ORDENES - PRODUCTO (TABLA DE HECHO)				
Descripción: Tabla intermedia de muchos a muchos de Ordenes / Productos				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_vta_order_prod	Venta productos en ordenes	Int	Identificador de Orden y producto
FK	id_order	Id de orden	Varchar(100)	Identificador de Orden
FK	id_product	Id de producto	Varchar(100)	Identificador de producto
	sales	Ventas en USD	decimal	Importe por producto
	qty	Cantidad vendida	Int	Cantidad vendida por producto
	discount	Descuento	decimal	Descuento por producto
	profit	Rentabilidad	decimal	Rentabilidad por producto
TABLA 14				
Nombre: CIUDAD (TABLA DE HECHO)				
Descripción: Datos de ciudades adonde se envian productos				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_city	Id de Ciudad	Int	Codigo de Ciudad
	name_city	Nombre de Ciudad	Varchar(100)	Descripcion de ciudad
FK	id_state	Id del Estado	Int	Codigo de Estado
TABLA 17				
Nombre: CODIGO POSTAL (TABLA DE HECHO)				
Descripción: Datos de los codigos postales de cada ciudad				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_pc	Id de Codigo postal	Int	Id de Codigo Postal
	name_pc	Descripcion de Codigo Postal	Varchar(100)	Descripcion de Codigo Postal
FK	id_city	Id de Ciudad	Int	Codigo de Ciudad
TABLA 12				
Nombre: RETORNOS (TABLA DE HECHO)				
Descripción: Datos de los rechazos de pedidos				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_retorno	Id de retorno	Int	Identificador de retorno
FK	id_order	Id de orden	Varchar(100)	Identificador de Orden
FK	id_cause	Causa de retorno	Int	Identificador de causa de retorno
TABLA 15				
Nombre: SUB-CATEGORIA (TABLA DE HECHO)				
Descripción: Datos de subcategoria de cada producto				
Tipo Clave	Campos abreviados	Nombre campo	Tipo de dato	Detalle
PK	id_subcateg	Id de Subcategoria	Int	Codigo de subcategoria
	name_subcat	Nombre de subcategoria	Varchar(100)	Descripcion de subcategoria
FK	id_categ	Id de Categoria	Int	Codigo de categoria

Datasets

Se utilizará un dataset en formato Excel que contiene 24 columnas y 9000 registros. El mismo debe ser normalizado para poder transformarlo en tablas que permitan armar la BBDD relacional.

Adjunto en el repositorio de Github donde se subirá la documentación de las entregas.

Herramientas y tecnologías utilizadas

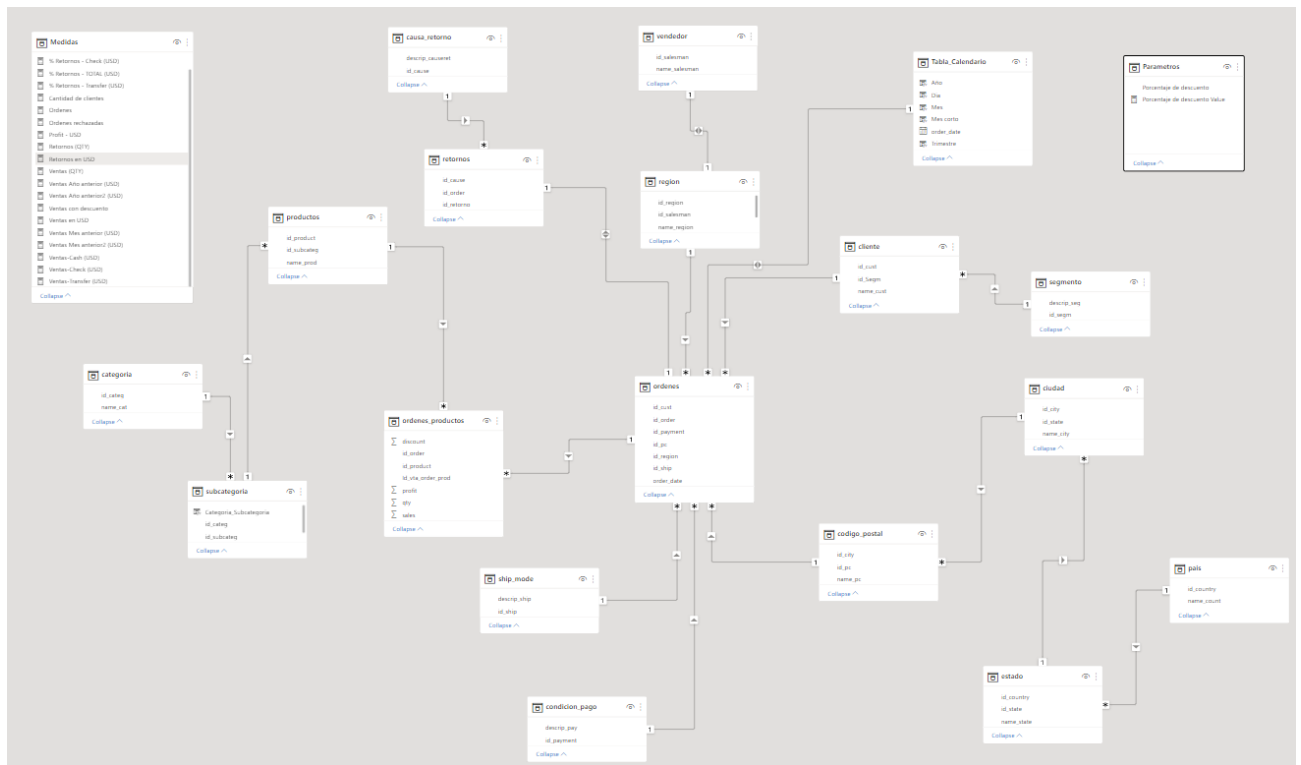
Para la realización del trabajo se utilizarán las siguientes tecnologías

- Microsoft SQL Server Management
- Drawio para el DER
- Microsoft Word
- Microsoft Excel
- Adobe PDF
- Git
- Github
- Google Drive

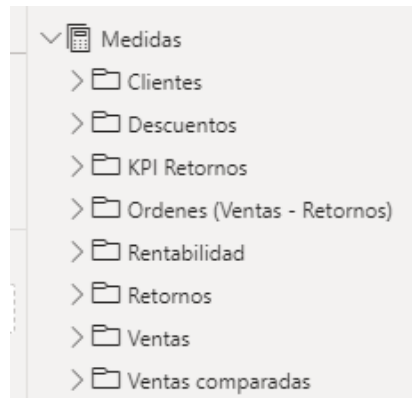
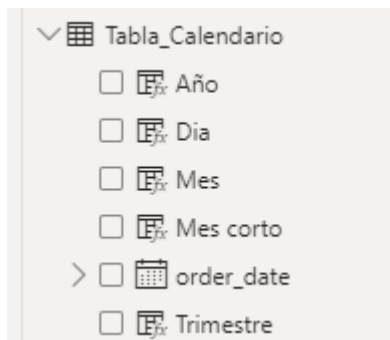
Modelo relacional en Power Bi

Para la subir la información a power bi, anteriormente se cargaron todas las tablas a SQL, como así tambien los datos.

Luego se exporto el xls desde SQL el cual se subió a Power bi y tambien quedo armada el Modelo relacional (Ver archivo BBDD1 en la carpeta BBDD Proyecto final / Subir a Power BI)



Se cargo la **TABLA CALENDARIO** y se generó la **CARPETA DE MEDIDAS**



Medidas calculadas

Se armaron las siguientes medidas calculadas las cuales quedaron agrupadas en carpeta por temática:

Subcarpeta Clientes:

- **Cantidad de clientes** = COUNT(cliente[id_cust])

Subcarpeta Datos Estadísticos:

- **Promedio de Profit en USD** = AVERAGE(ordenes_productos[profit])
- **Promedio de retornos en USD** = AVERAGEX(FILTER(ordenes_productos, RELATED(retornos[id_retorno])), [sales])
- **Promedio de ventas QTY** = AVERAGE(ordenes_productos[qty])
- **Promedio de ventas USD** = AVERAGE(ordenes_productos[sales])

Subcarpeta KPI Retornos:

- **% Retornos - Cash (USD)** = CALCULATE('Medidas'[Retornos en USD]/sum(ordenes_productos[sales]), condicion_pago[descrip_pay] = "cash")
- **% Retornos - Check (USD)** = CALCULATE('Medidas'[Retornos en USD]/sum(ordenes_productos[sales]), condicion_pago[descrip_pay] = "check")
- **% Retornos - Transfer (USD)** = CALCULATE('Medidas'[Retornos en USD]/sum(ordenes_productos[sales]), condicion_pago[descrip_pay] = "transfer")
- **% Retornos - TOTAL (USD)** = 'Medidas'[Retornos en USD]/sum(ordenes_productos[sales])

Subcarpeta Ordenes (Ventas – Retornos):

- **Ordenes** = CALCULATE (DISTINCTCOUNT (ordenes_productos [id_order]), categoria [id_categ])

Proyecto final – Diego Muro

- **Ordenes rechazadas** = CALCULATE (DISTINCTCOUNT (ordenes_productos [id_order]), retornos[id_retorno])

Subcarpeta Rentabilidad:

- **Profit - USD** = sum(ordenes_productos[profit])

Subcarpeta Retornos:

- **Retornos (QTY)** = CALCULATE(SUM(ordenes_productos[qty]),retornos[id_retorno])
- **Retornos en USD** = CALCULATE(SUM(ordenes_productos[sales]),retornos[id_retorno])

Subcarpeta Ventas:

- **Ventas (QTY)** = sum(ordenes_productos[qty])
- **Ventas en USD** = SUM(ordenes_productos[sales])
- **Ventas-Cash (USD)** = CALCULATE(sum(ordenes_productos[sales]), condicion_pago[descrip_pay] = "cash")
- **Ventas-Check (USD)** = CALCULATE(sum(ordenes_productos[sales]), condicion_pago[descrip_pay] = "check")
- **Ventas-Transfer (USD)** = CALCULATE(sum(ordenes_productos[sales]), condicion_pago[descrip_pay] = "Transfer")

Subcarpeta Ventas comparadas:

- **Ventas Año anterior2 (USD)** = CALCULATE(SUM(ordenes_productos[sales]), DATEADD(Tabla_Calendario[order_date], -1, YEAR))
- **Ventas Mes anterior2 (USD)** = CALCULATE(SUM(ordenes_productos[sales]), DATEADD(Tabla_Calendario[order_date], -1, MONTH))

Visualización del reporte:

Sección Portada:



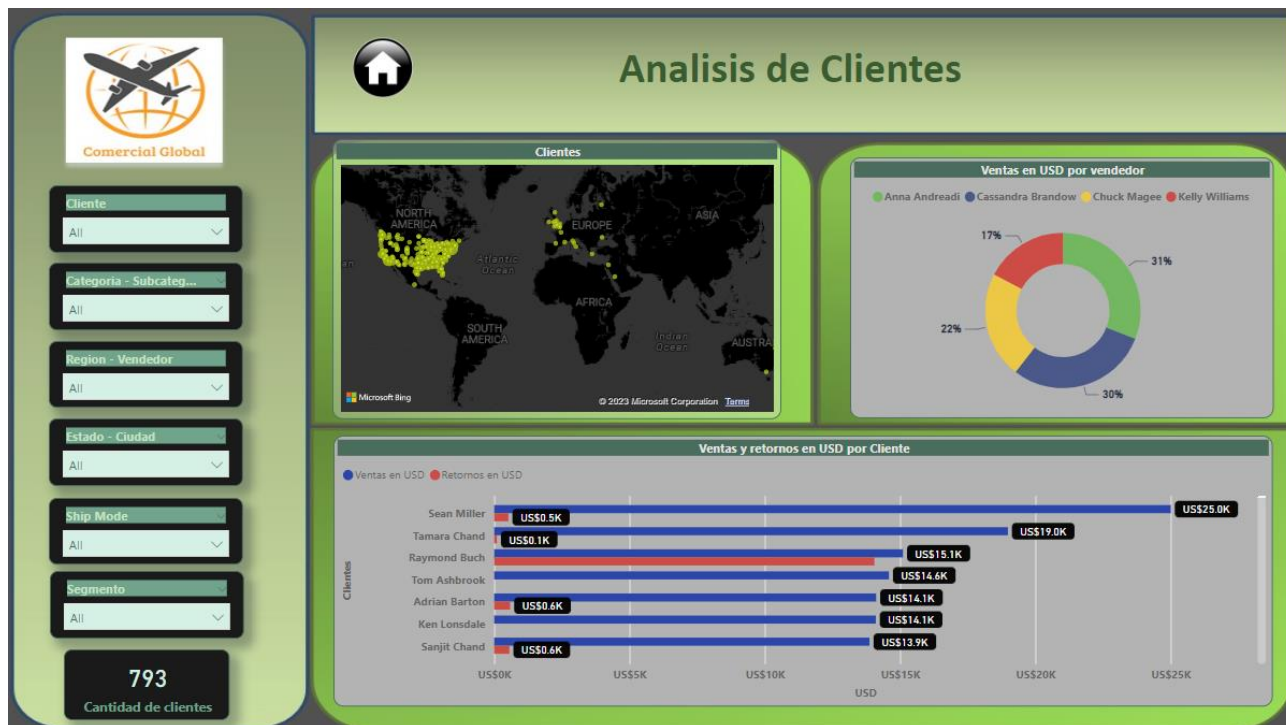
Sección Resumen de Ventas:



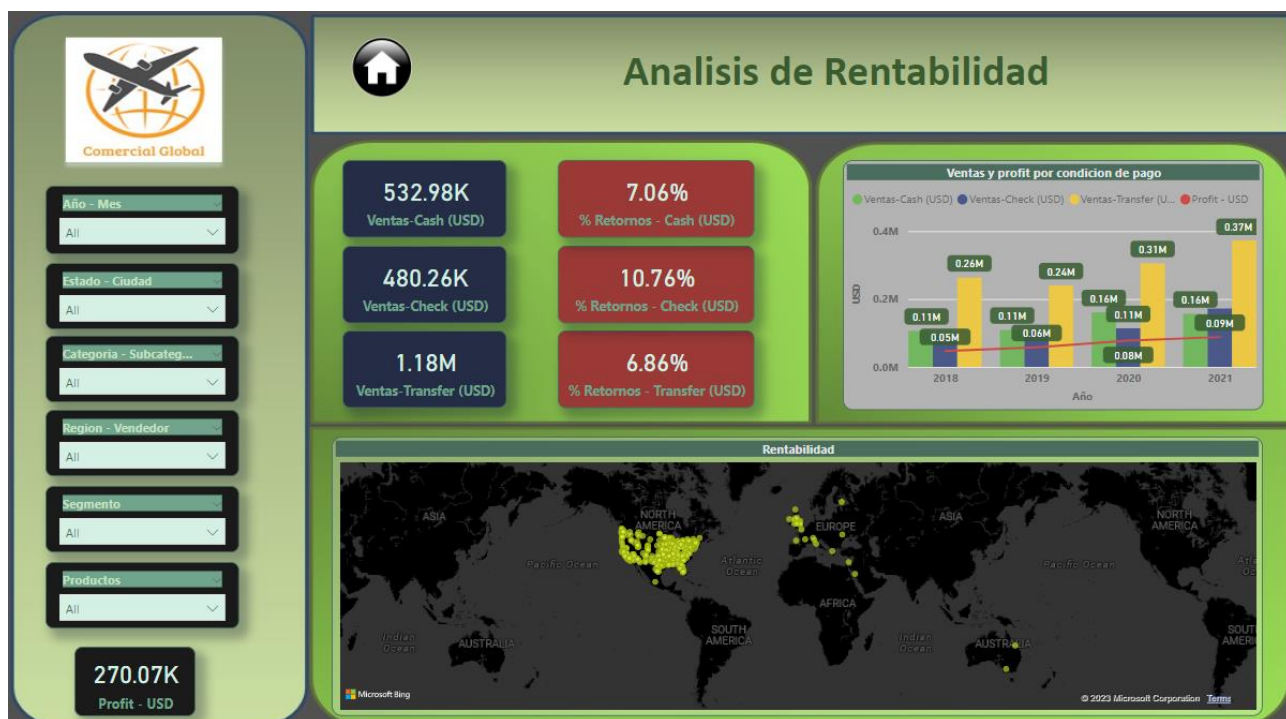
Sección Resumen de retornos:



Sección Análisis de Clientes:



Sección Análisis de Rentabilidad:



Sección Datos Estadísticos:

