

Proyecto Integrador - Módulo POWER BI

Base de Datos "Adventure Works cycles".



Nombre del autor:

Presa Germán Ariel

Email:

germanpresa92@gmail.com

Carrera:

Data Analytics

Cohorte:

DA FT10

Fecha de entrega:

09 de enero de 2025.

ÍNDICE:

ÍNDICE:	1
Introducción	2
Desarrollo del proyecto	2
AVANCE N° 1: Power Query - Conexión y limpieza de datos.	2
Conexión de datos.	2
Transformación de datos en Power Query.	3
AVANCE N° 2: Construcción del modelo relacional y mockup.	5
Modelo relacional:	6
Mockup:	6
AVANCE N° 3: DAX - Generación de medidas y columnas calculadas	9
AVANCE N° 4: Creación de tablero y documentación del proyecto BI	10
Optimización y sostenibilidad	11
Desafíos y soluciones	12
Reflexión personal	12

Introducción

El proyecto desarrollado para la empresa Adventure Works Cycles consistió en realizar el análisis sistematizado de sus ventas, ya que, anteriormente no contaba con indicadores que logren satisfacer las necesidades de información para una toma de decisiones eficiente y óptima.

Para ello se desarrollo un informe integral en Power BI que proporcione análisis detallados y visualizaciones interactivas sobre el rendimiento de ventas de la empresa Adventure Works Cycles.

Este informe permite a los usuarios finales comprender profundamente los factores que afectan a las ventas, los costos y la rentabilidad, facilitando la toma de decisiones estratégicas basadas en datos.

Desarrollo del proyecto

AVANCE N° 1: Power Query - Conexión y limpieza de datos.

En este primer avance, se utilizó Power Query para conectar y preparar los datos de la base "AdventureWorksDW2019" y la tabla adicional Customer. El objetivo principal es garantizar que los datos estén limpios y listos para su análisis en Power BI.

Esto implica descargar la base de datos, restaurarla en SQL Server, conectarla a Power BI y realizar diversas transformaciones en los datos para normalizarlos y combinarlos de manera adecuada. El enfoque principal está en preparar los datos para futuros análisis que desarrollaremos en los próximos avances

Conexión de datos.

- Se han conectado las siguientes tablas de la base de datos AdventureWorksDW2019 a Power BI:
 - DimProduct
 - DimProductCategory
 - DimProductSubcategory
 - DimDate
 - DimPromotion
 - DimSalesTerritory
 - DimGeography
 - FactInternetSales

- Fuente de datos Excel:
 - Se ha conectado la fuente de datos "DimCustomer" desde el archivo Excel proporcionado.

Transformación de datos en Power Query.

Transformación de Datos en Power Query

- Encabezados: Se verificaron y promovieron aquellos que no se generaron automáticamente.
- Limpieza de datos: Se eliminaron columnas por contener únicamente datos nulos o no relevantes,
 - Tabla FactInternetSales, columnas:
CarrierTrackingNumber
CustomerPONumber
UnitPriceDiscountPct
DiscountAmount
PromotionKey
SalesOrderLineNumber
RevisionNumber
OrderQuantity
 - DimProduct
FrenchProductName
FrenchDescription
ChineseDescription
ArabicDescription
HebrewDescription
ThaiDescription
GermanDescription
JapaneseDescription
TurkishDescription
 - DimProductCategory
FrenchProductCategoryName
ProductCategoryAlternateKey

-
- DimProductSubcategory
 - FrenchProductSubcategoryName
 - ProductSubcategoryAlternateKey
 - DimDate
 - FrenchDayNameOfWeek
 - DimPromotion
 - FrenchPromotionCategory
 - FrenchPromotionType
 - FrenchPromotionName
 - DimSalesTerritory
 - SalesTerritoryAlternateKey
 - Se elimino la ultima fila
 - DimGeography
 - FrenchCountryRegionName
 - DimCustomer
 - Se eliminaron columnas 18 y 31 sin datos
 - Name style
 - Sufix
 - Title
 - MiddleName
 - FrenchEducation
 - FrenchOccupation
 - AddressLine2
 - CustomerKey,
 - se eliminaron filas vacias
 - Se establecieron todas las relaciones entre las tablas.
 - Estructura de las tablas:
 - PK: Se identificaron y mantuvieron las claves primarias de cada tabla.
 - Descripción: Se incluyeron descripciones concisas de cada tabla para facilitar la comprensión.

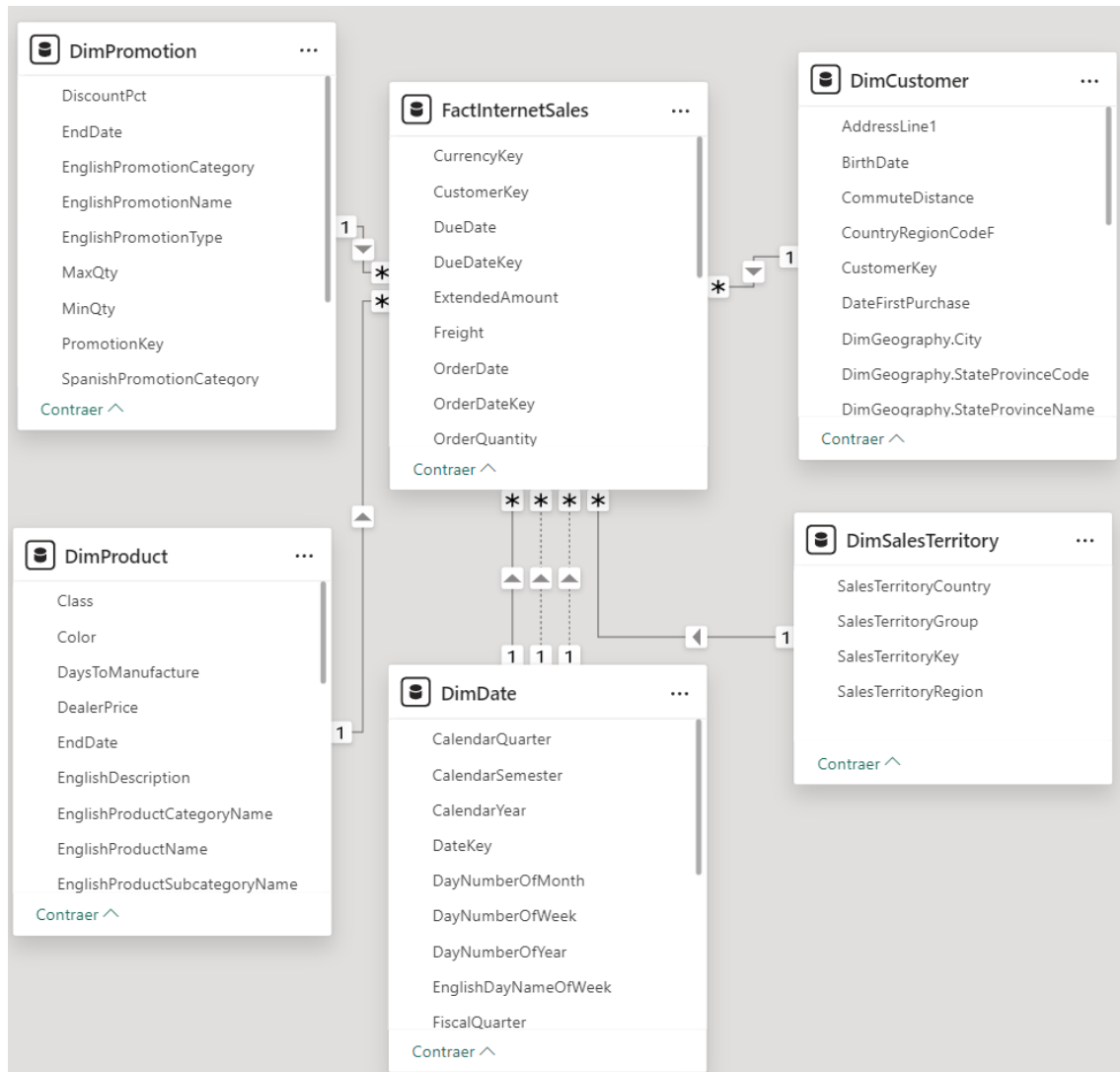
- FK: Se han identificado las claves foráneas de cada tabla y se han conectado las tablas de acuerdo con estas relaciones.
 - Columnas relevantes: Se han incluido columnas adicionales relevantes para el análisis, considerando el objetivo del proyecto.
- Cambio de nombres: Se han cambiado los nombres de algunas columnas para mejorar la legibilidad y facilitar el análisis.
- Normalización: Se han realizado las transformaciones necesarias para normalizar los datos, asegurando que cada hecho se almacene solo una vez y que las dimensiones estén debidamente estructuradas.
- Combinaciones:
 - Customer y Geography: Se ha realizado una combinación para incluir la información geográfica (ciudad, provincia, código) dentro de la tabla Customer.

Product, ProductCategory y ProductSubcategory: Se ha realizado una combinación para mostrar la categoría y subcategoría como clasificación dentro de la tabla Product.

AVANCE N° 2: Construcción del modelo relacional y mockup.

En este avance se acercó un primer diseño de un mockup y un modelo relacional en Power BI. El reporte responderá preguntas clave sobre ingresos, ventas, utilidades, costos y distribución geográfica de clientes y productos.

Se hizo foco en realizar un diseño eficiente del modelo de datos y del mockup, siguiendo las pautas para optimizar el rendimiento y facilitar el análisis. La estructura del informe prioriza el uso del patrón Z para ofrecer insights claros sobre el rendimiento de ventas.

Modelo relacional:**Mockup:**

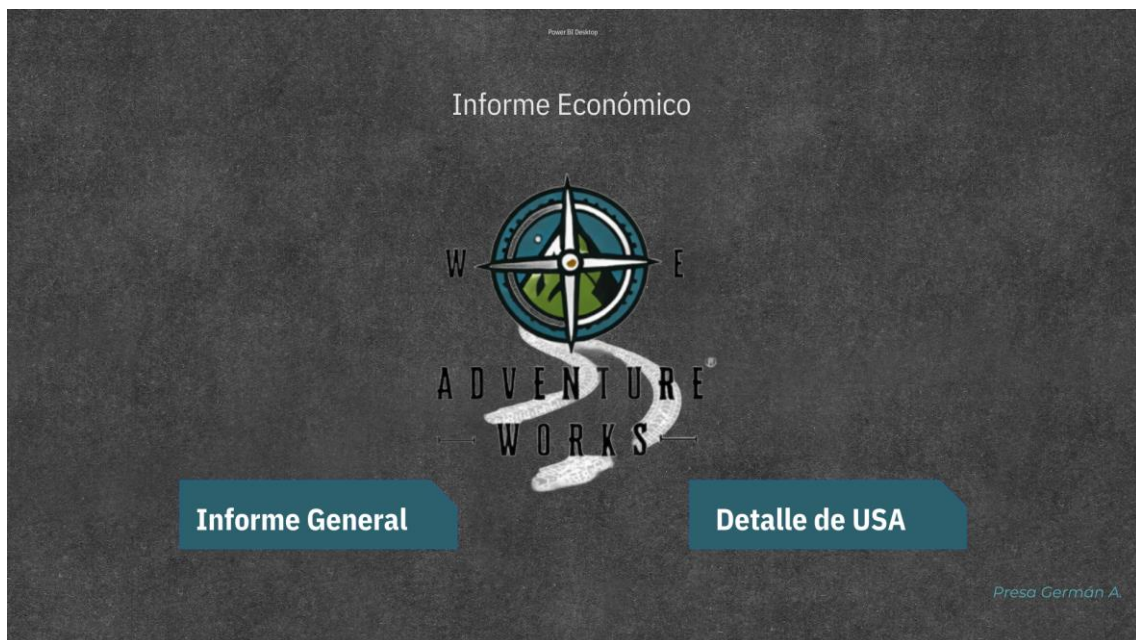
La intención del mockup es poder acercar al cliente un primer diseño de dashboard que se utilizará como base en el desarrollo del proyecto. Para con él, terminar de definir la información necesaria que se visualizará.

[Para abrir la presentación del Mockup click AQUI](#)

Para iniciar, se extrajo la paleta de colores del logo brindado por la empresa:

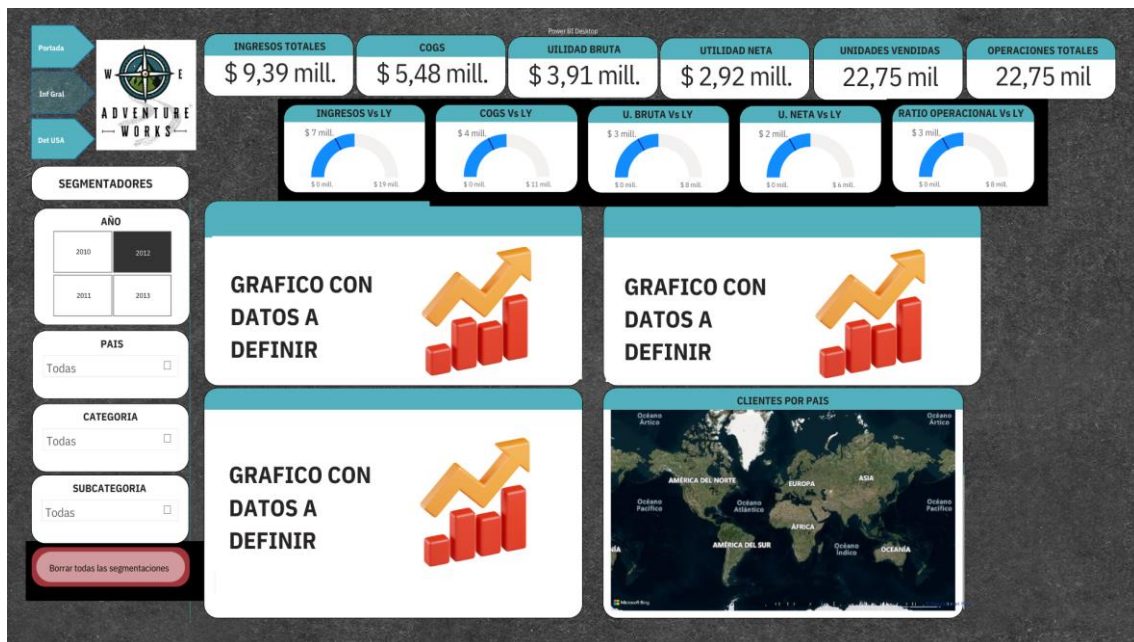


Luego se continuó con el desarrollo del mockup:



En la portada se colocaron dos botones que vinculan con las páginas del informe.

Módulo 3 - POWER BI



En la hoja del Informe General se colocarán filtros para facilitar la segmentación de los datos, como botones para una fluida navegación entre hojas del informe.

en la parte superior tarjetas y tacómetros que respondan a las necesidades planteadas por el usuario. debajo irán 3 gráficos a definir estilo y datos a visualizar y un gráfico de mapas.



Por último en la tercera hoja, la que se solicitó que sea de detalle de USA, se mantendrán filtros, botones, tarjetas y tacómetros, ya que se considera que dichos elementos son claves y de fácil lectura.

Se utilizarán dos gráficos a definir, se estima que el superior derecho será de líneas ya que se desea comparar ingresos acumulados del

periodo actual y el anterior, por esa razón a su derecha se colocará un KPI.

debajo se colocará una tabla y un mapa para mostrar detalles solicitados.

AVANCE N° 3: DAX - Generación de medidas y columnas calculadas

En este punto, previo a comenzar los trabajos con DAX, En Power Query, se agregó una columna personalizada llamada "Mes Abreviado", que muestra el nombre del mes en formato corto utilizando las primeras tres letras de cada mes.

Para optimizar el modelo de datos, se deshabilitó la carga de las tablas "ProductCategory", "ProductSubcategory" y "Geography".

Posteriormente, se marcó el campo "FullDateAlternateKey" como tabla de fechas.

Además, en la tabla "DimDate", se añadió una columna calculada llamada "Trimestre", que muestra el periodo en el formato "Trimestre XX" para facilitar la segmentación de los datos por trimestres.

Se generaron medidas que fueron agrupadas en carpetas para facilitar su comprensión. Las mismas son de vital importancia para poder responder los interrogantes del cliente.

Medidas (29)	
Fx-Bases	
Cantidades Operaciones	
Cantidades Vendidas	
COGS	
Costos de Envio	
Costos Totales	
Impuestos	
Ingresos Totales	
Total Clientes	
Utilidad Bruta	
Utilidad Neta	
Fx-Porcentuales	
% Margen Bruto	
% Margen Neto	
% Ratio Operacional	
Fx.Variaciones	
VARIACION COGS	
VARIACION COGS %	
VARIACION INGRESOS	
VARIACION INGRESOS %	
VARIACION UTILIDAD BRUTA	
VARIACION UTILIDAD BRUTA %	
VARIACION UTILIDAD NETA	
VARIACION UTILIDAD NETA %	
Medidas (29)	
Fx-Bases	
Fx-Porcentuales	
Fx.Variaciones	
Periodo actual	
Ingresos Acumulados	
Utilidad Bruta Periodo Actual	
Periodo Anterior	
% RATIO OPERACIONAL LY	
COGS LY	
Ingresos Acumulados LY	
INGRESOS LY	
Utilidad Bruta LY	
UTILIDAD NETA LY	

AVANCE N° 4: Creación de tablero y documentación del proyecto BI

El objetivo de esta etapa es desarrollar un informe final en Power BI que presente de manera clara y general los ingresos, costos, rentabilidad, y otros indicadores clave, además de un enfoque especial en el mercado de Estados Unidos. Interrogantes claves que fueron planteadas por la compañía.

Esto se realizó en base al mockup creado, las medidas y las columnas calculadas detalladas en el avance anterior.

Se configuró el lienzo de manera personalizada: Alto: 1080, Ancho: 1920.

Se crearon tres hojas: Portada, Informe General e Informe USA.

En la portada, se estableció un diseño con el logotipo de la empresa, y dos botones de navegación. Uno se dirige a la hoja Informe General y el otro a la hoja Informe detallado de USA.

En la segunda hoja, llamada Informe General, se colocó el logotipo de la empresa en la esquina superior izquierda. Justo a su izquierda, se incluyeron dos botones de acción que permiten navegar hacia la Portada y al Informe USA.

En la parte superior de la hoja, se dispusieron seis tarjetas que muestran los valores de los indicadores clave: Ingresos, COGS, Utilidad Bruta, Utilidad Neta, Transacciones y Cantidad Vendida.

Debajo de estas tarjetas, se añadieron cinco visualizaciones tipo tacómetro para comparar los siguientes indicadores: Ingresos vs Ingresos del año anterior, COGS vs COGS del año anterior, Utilidad Bruta vs Utilidad Bruta del año anterior, Utilidad Neta vs Utilidad Neta del año anterior, y el Ratio Operacional vs el del año anterior.

En el lado izquierdo de la hoja, se ubicaron cuatro segmentadores de datos que permiten filtrar por Año, País, Categoría y Subcategoría. Además, se añadió un botón que permite borrar todas las segmentaciones y restablecer los filtros aplicados.

En el centro, se colocó un gráfico de áreas apiladas y líneas que muestra los valores de Ingresos, COGS y Utilidad Bruta y neta distribuidos por mes y por día.

En la parte inferior izquierda, se incluyó un gráfico de barras agrupadas que presenta la Utilidad Bruta y la Utilidad Neta desglosadas por categorías y subcategorías de productos.

Finalmente, a la derecha del gráfico de mapa, que muestra la cantidad de clientes y parámetros claves que el cliente solicitó, se implementó un parámetro de campos denominado "Indicadores", que se emplea específicamente en esta visualización.

En la tercera hoja, "detalle de USA" se conservaron los filtros tarjetas y tacómetros.

Se colocó un gráfico de área y líneas, un gráfico de líneas para visualizar los ingresos acumulados del periodo actual versus los del periodo anterior. un mapa de EEUU y una tabla con los detalles de los indicadores mencionados ut supra.

Optimización y sostenibilidad

La base de datos fue optimizando paulatinamente a medida que se construyó, se estructuraron las tablas de forma tal para minimizar las redundancias y se puedan relacionar todos los datos de manera acorde. De este modo también se optimizó el acceso y la interpretación de la información.

Para garantizar la sostenibilidad a largo plazo, se introdujeron comentarios y procedimientos para cada actualización y manejo de tabla, facilitando así la comprensión total y su futura manipulación por personas externas.

Desafíos y soluciones

Enfrenté desafíos comunes al trabajar con Power BI, como ser:
Errores en DAX: No manejar correctamente las relaciones entre tablas, no usar CALCULATE cuando se necesita un cambio de contexto, usar SUM en lugar de SUMX en cálculos fila por fila, no manejar divisiones por cero con DIVIDE, no optimizar medidas o variables para grandes volúmenes de datos.

Errores en Power Query: No reducir la cantidad de datos importados, no cambiar los tipos de datos correctamente, aplicar pasos innecesarios en la transformación de datos, no usar claves únicas en relaciones entre tablas, cargar datos innecesarios en el modelo cuando no es requerido.

Reflexión personal

Durante este proyecto, he tenido la oportunidad de consolidar mis conocimientos básicos en análisis de datos y adquirir habilidades prácticas que me han permitido entender mejor el manejo de bases de datos.

Si tuviera que volver a empezar este proyecto, probablemente enfocaría más tiempo en la planificación y en la comprensión inicial de los datos. Me aseguraría de definir claramente los objetivos del análisis desde el principio y de explorar más herramientas y técnicas avanzadas para el análisis de datos.

Extra Credit