**Proyecto Integrador Módulo 3: “Visualizando el rendimiento de AWC con Power BI”**

**Nombre del autor:** Fernando Gómez

**Email:** fernandomgomez8@gmail.com

**Cohorte:** DA-PT04

**Fecha de entrega:** 14/11/2024

**Institución:** Adventure Works Cycles



# 

# Introducción

# Este proyecto, denominado “Visualizando el rendimiento de AWC con Power BI", tiene como propósito realizar un análisis estructurado de las ventas de la empresa. Actualmente, Adventure Works Cycles no cuenta con indicadores que proporcionen la información necesaria para tomar decisiones de forma efectiva y eficiente.

# Adventure Works Cycles es una reconocida compañía multinacional que fabrica y distribuye bicicletas, componentes y accesorios en mercados de Norteamérica, Europa y Asia.

# Para cubrir las necesidades de información sobre el rendimiento de la empresa, se diseñó un informe completo en Power BI que ofrece un análisis detallado y visualizaciones interactivas del desempeño de ventas de AWC.

# Este proyecto permite a los usuarios finales comprender en profundidad los factores que afectan las ventas, los costos y la rentabilidad, facilitando decisiones estratégicas informadas, basadas en datos y cumpliendo así con los requerimientos informativos de la empresa.

# Desarrollo del proyecto

# Para comenzar el proyecto se llevó a cabo la restauración de la base de datos *AdventureWorksDW2019* en SQL Server, para luego conectar las siguientes tablas a Power BI:

# DimProduct

# DimProductCategory

# DimProductSubcategory

# DimPromotion

# DimDate

# DimSalesTerritory

# DimGeography

# FactInternetSales (tabla de hechos)

# En las bases de datos, una tabla de hechos es la pieza central dentro de un esquema dimensional; contiene los valores de las métricas empresariales o, en otras palabras, los indicadores de negocio.

# El modelo seguirá un esquema en estrella, en el cual las tablas de dimensiones permiten filtrar y agrupar, mientras que la tabla de hechos actúa como un resumen. Para garantizar la solidez de este esquema, se revisaron y configuraron manualmente sus relaciones.

# Posteriormente, se conectó a Power BI una fuente de datos adicional, la tabla DimCustomer, desde un archivo de Excel. A continuación, se realizaron las transformaciones de datos en Power Query (limpieza, formatos adecuados, etc.).

# Un paso crucial de estas transformaciones fue asegurar que las claves primarias tuvieran valores únicos y no nulos para mantener la integridad del modelo. Una vez verificado esto, se eliminaron las columnas que no aportaban valor al modelo, ya sea porque estaban vacías o contenían muchos valores atípicos. Una columna con muchos valores atípicos, puede sesgar el análisis, por lo cual se decidió eliminarlas.

# Las tablas DimCustomer y DimGeography se combinaron para extraer las columnas de ciudad, provincia y código de provincia del cliente, y se almacenaron en DimCustomer, mientras que DimGeography quedó oculta por no ser necesaria. Luego, se aplicó el mismo procedimiento con las tablas DimProduct, DimProductCategory y DimProductSubcategory, extrayendo de estas dos últimas el nombre en inglés de cada categoría y subcategoría. Con esta información almacenada en DimProduct, se ocultaron DimProductCategory y DimProductSubcategory.

# Se creó una tabla de calendario llamada "Calendar" para optimizar el modelo, usando la función CALENDAR(). Se decidió hacer esto porque, en comparación con la tabla de fechas existente “DimDate”, la tabla Calendario creada de esta manera proporciona mayor claridad para el análisis, funcionalidad más óptima y mayor facilidad para personalización de la tabla.

# Esta tabla contiene todas las fechas desde la primera hasta la última orden de compra registrada y cuenta con columnas adicionales que ofrecen distintos formatos de fecha, como nombre abreviado del mes, número de mes y trimestre, entre otros. La tabla “Calendar” se marcó como la tabla de fechas en uso y se vinculó al esquema en estrella activando su relación “uno a varios” con la tabla de hechos, la relación se hizo por medio de la columna “Date” de la tabla “Calendar” hacia la columna “OrderDate” de la tabla de hechos. Las relaciones de la tabla “DimDate” con la tabla de hechos se dejaron inactivas. A continuación, se detalla su sintaxis:

Calendar =

VAR MinFecha = MIN (FactInternetSales[OrderDate])

VAR MaxFecha = MAX (FactInternetSales[OrderDate])

RETURN

ADDCOLUMNS (

    CALENDAR(MinFecha, MaxFecha),

    "Year", YEAR([Date]),

    "Month", MONTH([Date]),

    "Year Quarter", YEAR([Date]) & "-"  & QUARTER([Date]),

    "Year Month", YEAR([Date]) & "-"  & FORMAT([Date], "MM"),

    "Quarter", QUARTER([Date]),

    "Month Name", FORMAT([Date], "MMMM"),

    "Short Month Name", FORMAT([Date], "MMM"))

# El siguiente paso del proyecto fue comenzar a crear las medidas necesarias para responder a las consultas solicitadas por el cliente:

* ¿Cuál es el total de ingresos del período actual y del período anterior? ¿Qué porcentaje representa dicha variación?

IngresosTotales = SUM(FactInternetSales[SalesAmount])

IngresosTotalesLY = CALCULATE([IngresosTotales], SAMEPERIODLASTYEAR('Calendar'[Date]))

%VariacionIngresosTotales = ([IngresosTotales]-[IngresosTotalesLY])/[IngresosTotales]

* ¿Cuál es la cantidad vendida?

CantidadVendida = SUM(FactInternetSales[OrderQuantity])

* ¿Cuál es la utilidad bruta del período actual y del período anterior? ¿Y la utilidad neta? ¿Cuál es el porcentaje de variación de ambas utilidades?

UtilidadBruta = ([IngresosTotales]-[COGS])

UtilidadBrutaLY = CALCULATE([UtilidadBruta], SAMEPERIODLASTYEAR('Calendar'[Date]))

UtilidadNeta = ([IngresosTotales]-[COGS]-[CostoEnvio]-[Impuestos])

UtilidadNetaLY = CALCULATE([UtilidadNeta], SAMEPERIODLASTYEAR('Calendar'[Date]))

%VariacionUtilidadBruta = ([UtilidadBruta]-[UtilidadBrutaLY])/[UtilidadBruta]

%VariacionUtilidadNeta = ([UtilidadNeta]-[UtilidadNetaLY])/[UtilidadNeta]

* ¿Cuál es el costo de los bienes vendidos (COGS) del período actual y del período anterior? ¿En qué porcentaje varía?

COGS = SUM(FactInternetSales[TotalProductCost])

COGSLY = CALCULATE([COGS], SAMEPERIODLASTYEAR('Calendar'[Date]))

%VariacionCOGS = ([COGS]-[COGSLY])/[COGS]

* ¿Cuántos clientes hay en cada país?

TotalClientes = DISTINCTCOUNT(FactInternetSales[CustomerKey])

* ¿Cómo se distribuyen los ingresos, el COGS y la utilidad bruta mensualmente?

Esta pregunta se respondió creando gráficos utilizando las medidas mencionadas en las respuestas anteriores.

* ¿Qué utilidad (bruta y neta) tuvo cada segmento (categoría) y subcategoría de producto?

Esta pregunta se respondió creando gráficos utilizando las medidas mencionadas en las respuestas anteriores.

* Los usuarios desean ver además el Ratio Costo operacional versus LY (COGS + freight / Ingresos), el porcentaje de margen de utilidad bruta y utilidad neta y el porcentaje de COGS mostrado de manera eficiente en medidores (o tacómetros).

%CostoOp/Ingresos = DIVIDE(([CostoEnvio]+[COGS]), [IngresosTotales])

%CostoOp/IngresosLY = CALCULATE([%CostoOp/Ingresos], SAMEPERIODLASTYEAR('Calendar'[Date]))

%MargenBruto = DIVIDE([UtilidadBruta], [IngresosTotales])

%MargenNeto = DIVIDE([UtilidadNeta],[IngresosTotales])

* Como adicional, el usuario solicita ver de manera detallada indicadores del negocio de Estados Unidos donde se muestre por cada provincia y ciudad el segmento de producto (categoría), los ingresos, utilidades, COGS, márgenes (bruto y neto), y el costo de envío. Todo lo anterior desean ver resumido en una tabla. Por otro lado se solicitó un gráfico que muestre el COGS y el % de margen bruto (utilidad bruta) por ciudad y otro comparativo que muestre los ingresos acumulados del período actual versus los del período anterior.

Para responder a esta consulta en forma más óptima se creó un Grupo de Cálculo llamado “VariaciónTiempo” y se creó el Parámetro de Campo llamado “Indicadores”.

El Grupo de Cálculo “VariacionTiempo” se creó mediante la siguiente sintaxis:

PeriodoActual = SELECTEDMEASURE()

LY = CALCULATE(SELECTEDMEASURE(), SAMEPERIODLASTYEAR('Calendar'[Date]))

YOY = SELECTEDMEASURE()-CALCULATE(SELECTEDMEASURE(), SAMEPERIODLASTYEAR('Calendar'[Date]))

YOY% = DIVIDE(CALCULATE(SELECTEDMEASURE(), 'GCInteligenciaTemporal'[VariacionTiempo]="YOY"), CALCULATE(SELECTEDMEASURE(), 'GCInteligenciaTemporal'[VariacionTiempo]="LY"))

El Parámetro de Campo llamado “Indicadores” se creó mediante la siguiente sintaxis:

Indicadores = {

    ("IngresosTotales", NAMEOF('Medidas'[IngresosTotales]), 0),

    ("UtilidadBruta", NAMEOF('Medidas'[UtilidadBruta]), 1),

    ("UtilidadNeta", NAMEOF('Medidas'[UtilidadNeta]), 2),

    ("COGS", NAMEOF('Medidas'[COGS]), 3),

    ("%MargenNeto", NAMEOF('Medidas'[%MargenNeto]), 4),

    ("%MargenBruto", NAMEOF('Medidas'[%MargenBruto]), 5),

    ("CostoEnvio", NAMEOF('Medidas'[CostoEnvio]), 6)}

Finalmente se organizaron todas las medidas en diferentes carpetas para mejorar la organización del reporte.

**Análisis general del tablero**

Se configuró el diseño del informe (lienzo) de forma personalizada con un tamaño de 1080 de alto y 1920 de ancho. El informe se divide en tres páginas principales:

* **PORTADA**: cumple el rol de Menú Principal del reporte e incluye el logo de la empresa, una breve descripción del contenido y objetivos del reporte, y dos enlaces, uno al Reporte Financiero (Dashboard) General y otro al Reporte Financiero (Dashboard) específico de EEUU.
* **TABLERO REPORTE FINANCIERO GENERAL**: ofrece una visión de los indicadores financieros clave de la empresa y se divide en **3 secciones diferentes, “Ingresos y Ventas”, “Costos” y “Clientes”, cada una de estas secciones es una vista (marcador)** **diferente del mismo tablero** y puede accederse a cada una de ellas mediante sus correspondientes botones accionables.

En este dashboard (en sus 3 vistas) se destacan las métricas principales como ingresos totales, costos operativos, utilidad neta, etc. presentadas en tarjetas o medidores comparativos (tacómetros) de KPI’s y gráficos comparativos para una rápida referencia. También se incluyen gráficos de tendencias de ventas y costos a lo largo del tiempo, gráficos de ingresos y costos por categoría de producto, un mapa que muestra la distribución geográfica de los clientes y un conjunto de segmentadores intuitivos que permiten filtrar la información por año, mes y categoría de producto (más un segmentador para selección de indicadores en la vista de “Clientes”). Con esta configuración, el tablero permite a los usuarios obtener una visión clara y completa de la situación financiera general de la empresa, ayudando a identificar patrones de rendimiento y áreas de oportunidad.

* **TABLERO REPORTE FINANCIERO EEUU:** ofrece una visión detallada del rendimiento de la empresa enfocado en la zona de Estados Unidos.

Al igual que el Tablero General, en su parte superior muestra KPI’s como ingresos totales y márgenes de ganancia o producto más vendido en forma de tarjetas. Proporciona una exploración bastante completa de los Indicadores de negocio (ingresos totales, utilidad bruta, utilidad neta, etc.) para distintos períodos temporales y su variación entre sí, mediante una tabla matriz creada a partir del Parámetro de Campo llamado “Indicadores” que incluye las Medidas ya mencionadas (ingresos totales, utilidad bruta, etc.) y de un Grupo de Cálculo llamado “VariacionTiempo”. La mencionada tabla incluye las jerarquías de Estado y Ciudad y responde a segmentadores de “Categoría”, “Año” e “Indicadores” (que permite visualizar uno a uno o en simultáneo los diferentes valores de los indicadores de negocio.

Además, el tablero contiene un gráfico de columnas agrupadas que pemite visualizar los COGS y la Utilidad Bruta por Ciudad (con líneas de promedio para cada serie de datos) y un gráfico de líneas que compara los ingresos acumulados del período actual frente a los del período anterior (con líneas de promedio para cada serie de datos).

**Cabe destacar también** que ambos tableros (Reporte Financiero General y EEUU) cuentan con botones interactivos e intuitivos por ejemplo para ir a la página siguiente o a la anterior, para volver al menú principal, para recorrer las vistas o marcadores de una misma página, para obtener información sobre uso de un segmentador o para borrar todas las segmentaciones. También se editaron las interacciones de los filtros o segmentadores por ejemplo para que un segmentador por año no afecte a un gráfico de líneas de ingresos por año. Además se agregaron tooltips por ejemplo para obtener información extra sobre indicadores de negocios en los distintos países al posicionar el cursor sobre el mapa de clientes.

# 

# Resultados principales y líneas futuras de análisis

# A continuación, se presentan las conclusiones sobre los principales resultados e insights obtenidos para cada segmento del análisis de rendimiento de la emresa, así como las recomendaciones para futuros estudios o profundización en alguno de los insights encontrados. Cabe aclarar aquí que si bien el Reporte

* **Ingresos y Ventas:** 
  + La mayor cantidad de ventas se produce en el mes de diciembre (época de fiestas) tanto a nivel global como en el mercado de EEUU. Esto es un insight para tomar la decisión por ejemplo de mayores inversiones en promociones y campañas publicitarias navideñas.

En este punto se podría profundizar en un estudio sobre las ventas en la temporada vacacional de mitad de año para intentar igualar las ventas que se logran a fin de año.

* + Los márgenes de ganancia bruto y neto, se mantienen bastante estables, aunque hay que destacar que en el año 2013 (último año para el que se tienen datos completos) tanto el margen de ganancia bruto como el neto estuvieron aproximadamente un 0,2% por debajo respecto al año anterior. Este es un dato importante a considerar por la dirección de la empresa.
* **Costos:**
  + Los COGS y los costos de envío siguen líneas de tendencia muy similares tanto anualmente como mensualmente.
  + El ratio de COGS y costos operacionales sobre los Ingresos se mantienen bastante estables en los períodos interanuales. Este al igual que el punto anterior son datos que aportan tranquilidad al verse una estabilidad en las finanzas de la empresa.
* **Clientes:**
  + A primer golpe de vista del mapa de clientes se observa que el principal mercado de la empresa es Estados Unidos. Este hecho justifica que se haya hecho un reporte especialmente enfocado en este mercado y se recomienda seguir profundizando en el análisis del mismo. Cabe también destacar a Oceanía que se encuentra en segundo lugar con una importante cantidad de Clientes.
* **Reporte Estados Unidos:**
  + Los márgenes de ganancia (bruta y neta) en el mercado de Estados Unidos son aproximadamente un 1% mayores que a nivel global de todo el territorio abarcado por la empresa.
  + La ciudad que generó la mayor utilidad bruta fue la ciudad de Bellflower, del Estado de California, siendo también California el Estado que generó mayores ingresos en todo el período temporal estudiado. Basados en este insight recomendamos inversiones para mantener y potenciar el negocio en este estado.
  + En el gráfico de ingresos acumulados puede observarse una caída de los ingresos en el período que va del año 2011 al 2012, y luego un notable crecimiento de los mismos entre el 2012 y el 2013 (último año para el que se tienen datos completos). Profundizar en el estudio de las condiciones del mercado durante el año en que se produjo esa caída de los ingresos puede ser una oportunidad para identificar nuevas claves de negocio.

# Reflexión personal

# Durante la realización de este proyecto pude notar el enorme potencial que tiene Power BI para el análisis de datos. En este potencial radica también lo que veo como el mayor desafío en su uso que es la gran cantidad de herramientas que posee (que además se actualizan constantemente) y la posibilidad de llegar a un mismo resultado por múltiples caminos diferentes, lo que implica tener la capacidad de decidir cuál es el más óptimo en cada caso.

# El cursado del módulo y la realización del proyecto me han provisto de habilidades para enfocarme tanto en la calidad de los datos como en su comprensión, manipulación y presentación de manera clara y visualmente atractiva (sin dejar de lado la sobriedad y formalidad).

# Si tuviera que recomenzar este trabajo, desde el principio le daría aún más importancia a la limpieza de los datos, a los insights a los que preveo arribar y a las relaciones generadas en el modelo de datos. Además usaría la función CALENDAR() desde el comienzo para crear una tabla de fechas personalizada que simplifique el modelo desde el principio."

# 