**[Título del proyecto]**

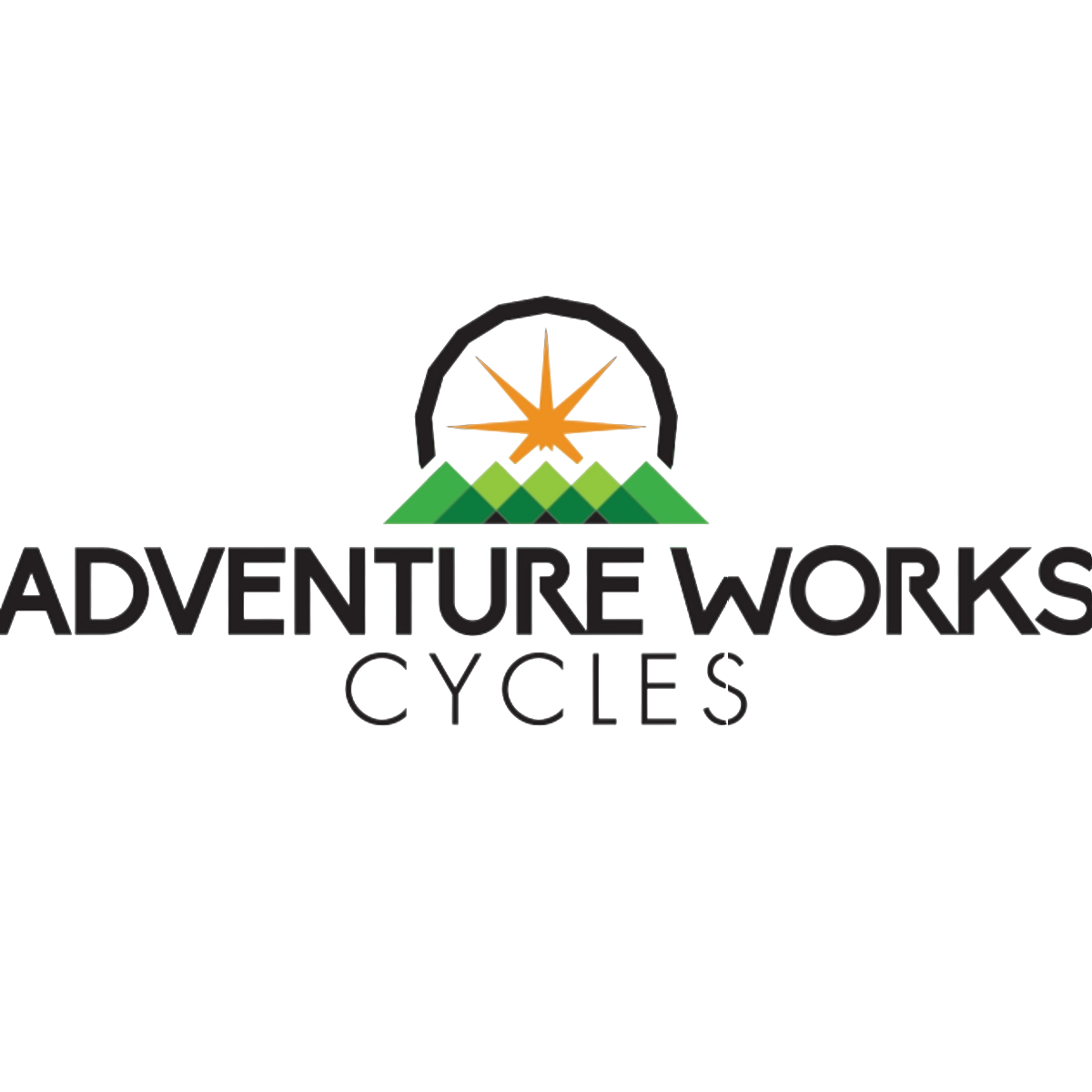
**Nombre del autor: Gerardo Morsucci**

**Email: geramorsucci@gmail.com**

**Cohorte: DA-PT01**

**Fecha de entrega:** 20/06/2024

**Institución:** Adventure Works Cycles

****

# Introducción

# El presente proyecto, titulado "Visualizando el Rendimiento de AWC en Power BI", tiene como finalidad llevar a cabo un análisis estructurado de las ventas de la empresa. Actualmente, la empresa carece de indicadores que cumplan con las necesidades informativas necesarias para una toma de decisiones eficiente y óptima.

# 

# Adventure Works Cycles es una destacada empresa multinacional de fabricación que produce y distribuye bicicletas, piezas y accesorios para mercados comerciales en Norteamérica, Europa y Asia.

# 

# Para alcanzar el objetivo de satisfacer las necesidades de información sobre el desempeño de la empresa, se desarrolló un informe integral en Power BI. Este informe ofrece un análisis detallado y visualizaciones interactivas sobre el rendimiento de ventas de AWC.

# 

# A través de este proyecto, los usuarios finales pueden obtener una comprensión profunda de los factores que influyen en las ventas, los costos y la rentabilidad, facilitando la toma de decisiones estratégicas basadas en datos y satisfaciendo así las necesidades informativas de la empresa.

# Desarrollo del proyecto

# Se realizó la restauración de la base de datos adventure works dw 2019.bank en sql server para luego conectar las siguientes tablas a power bi:

1. **Dim product**
2. **DimProductCategory**
3. **DimProductSubcategory**
4. **DimPromotion**
5. **DimDate**
6. **DimSalesTerritory**
7. **DimGeography**
8. **FactInternetSales ( Tabla de hechos )**

En las bases de datos, y más concretamente en un data warehouse, una tabla de hechos es la tabla central de un esquema dimensional y contiene los valores de las medidas de negocio o dicho de otra forma los indicadores de negocio.

**El modelo se basará en un esquema de estrella**, buscando que el mismo nos proporciona tablas para filtrar y agrupar (tablas de dimensiones) y también una tabla a modo resumen (tabla de hechos). Para que este esquema sea sólido, se le revisaron y configuraron apropiadamente sus relaciones.

Luego se conectó la fuente de datos de **DimCustomer** desde un archivo Excel a Power BI:

Se procedió a realizar las transformaciones de datos en power query que se detallan a continuación:

**Un paso importante fue verificar que las llaves primarias tengan valores únicos, y no sean nulas,** esto sirve para garantizar la integridad del modelo. Una vez hecho esto **se eliminaron columnas que no aportan información** al modelo (columnas vacías o con una gran predominancia de outliers) Si una columna tiene mucha cantidad de outliers corremos el riesgo de sesgar el análisis, por tal razón se tomó la decisión mencionada.

Se unificaron las tablas **DimCustome**r con **DimGeography** para extraer las columnas que contienen ciudad, provincia y código de provincia del cliente, las mismas quedarían almacenadas en DimCustomer y **la tabla DimGeography quedaría oculta**, sin relevancia. Siguiente a este cambió, se hizo algo similar entre **DimProduct**, DimProductCategory y DimProductSubCategory, esta vez se buscó extraer de estas últimas dos tablas el nombre en inglés de cada categoría y subcategoría. Quedando DimProduct con estas columnas mencionadas, **se ocultó DimProductCategory y DimProductSubCategory.**

Se creó una tabla Calendario llamada “Calendar” para generar una optimización en el modelo, esto se logró a través de la función **CALENDAR() y la columna OrderDate de la tabla FactInternetSales.** Dentro de esta tabla están las fechas desde la primera compra registrada hasta la última. En la misma se crearon varias columnas para darles distintos formatos a las fechas, como por ejemplo: Short month name, número de més, “Trimestre X”, por nombrar algunas.. **La tabla Calendar se marcó como tabla de fechas, luego de generar su vinculación al esquema estrella.** Adjunto su sintaxis a continuación:

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"))

**SE FILTRARON LOS AÑOS ENTRE 2012 Y 2013 YA QUE EL ÚLTIMO AÑO (2014) TENÍA DATOS EXTREMADAMENTE SESGADOS. EL AÑO 2011 NO SE UTILIZÓ PESE A TENER DATOS PORQUE SE PREFIRIÓ REALIZAR UNA CONCLUSIÓN EN BASE A LOS AÑOS 2012 Y 2013.**

Como siguiente proceso se empezaron a generar las medidas para responder las siguientes consultas solicitadas por el cliente:

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

TotalSales = SUM(FactInternetSales[SalesAmount])

TotalSalesLY = CALCULATE([TotalSales],SAMEPERIODLASTYEAR('Calendar'[Date]))

TotalSalesDIFF = ([TotalSales] - [TotalSalesLY])

% TotalSales = DIVIDE([TotalSalesDIFF], [TotalSalesLY])

* ¿Cuál es la cantidad vendida?

SoldProducts = 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?

GrossRevenue = ([TotalSales]-[TotalCOGS])

GrossRevenueLY = CALCULATE([GrossRevenue], SAMEPERIODLASTYEAR('Calendar'[Date]))

GrossRevenueDIFF = ([GrossRevenue] - [GrossRevenueLY])

% GrossRevenue = DIVIDE([GrossRevenueDIFF], [GrossRevenueLY])

NetRevenue = ([GrossRevenue]-[TotalFreight]-[TotalTaxes])

NetRevenueLY = CALCULATE([NetRevenue], SAMEPERIODLASTYEAR('Calendar'[Date]))

NetRevenueDIFF = ([NetRevenue] - [NetRevenueLY])

% NetRevenue = DIVIDE([NetRevenueDIFF], [NetRevenueLY])

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

TotalCOGS = SUM(FactInternetSales[TotalProductCost])

TotalCogsLY = CALCULATE([TotalCOGS], SAMEPERIODLASTYEAR('Calendar'[Date]))

TotalCogsDIFF = ([TotalCOGS] - [TotalCogsLY])

% TotalCOGS = DIVIDE([TotalCogsDIFF], [TotalCogsLY])

* ¿Cuántos clientes hay en cada país?

Customers = DISTINCTCOUNT(FactInternetSales[CustomerKey])

Acá se generó un parámetro de campos:

Indicators = {

("Sales", NAMEOF('Medidas'[TotalSales]), 0),

("UtilityNet", NAMEOF('Medidas'[NetRevenue]), 1),

("UtilityBrut", NAMEOF('Medidas'[GrossRevenue]), 2),

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

("VariationUNet", NAMEOF('Medidas'[% NetRevenue]), 4),

("VariationUBrut", NAMEOF('Medidas'[% GrossRevenue]), 5),

("Freight", NAMEOF('Medidas'[TotalFreight]), 6)

}

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

Se puede responder con las medidas ya creadas anteriormente

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

Se puede responder con las medidas ya creadas anteriormente

* 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).

TotalFreight = SUM(FactInternetSales[Freight])

TotalCogsFreight = ([TotalCOGS] + [TotalFreight])

OperationalCostRatio = DIVIDE([TotalCogsFreight], [TotalSales])

OperationalCostRatioLY = CALCULATE([OperationalCostRatio], SAMEPERIODLASTYEAR('Calendar'[Date]))

Difference Operational Cost Ratio vs LY = ([OperationalCostRatio] - [OperationalCostRatioLY])

% MargenGrossRevenue = DIVIDE([GrossRevenue], [TotalSales])

% MargenNetRevenue = DIVIDE([NetRevenue], [TotalSales])

* 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.

Se creó un grupo de cálculo a modo de optimización:

Total = SELECTEDMEASURE()

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

YOY = SELECTEDMEASURE() -

CALCULATE(

SELECTEDMEASURE(),

SAMEPERIODLASTYEAR('Calendar'[Date]))

YOY% = DIVIDE(CALCULATE(SELECTEDMEASURE(),Variacion\_Tiempo[Calculation group column]= "YOY"), CALCULATE(SELECTEDMEASURE(), Variacion\_Tiempo[Calculation group column]= "LY"))

Se optimizaron las medidas al organizarlas en carpetas.

# Análisis general del tablero

Se configuró el lienzo de manera personalizada:

Alto: 1080

Ancho: 1920

El tablero consta de 3 partes:

* **COVER:** Contiene 2 hipervínculos y un “clustered column chart” que nos representa los COGS, el total de ventas y la utilidad bruta.

Visualizar Tendencias: Puedes observar cómo han variado los ingresos, COGS y la utilidad bruta a lo largo del tiempo, identificando patrones estacionales o tendencias crecientes/decrecientes.

Comparar Componentes: Al ver los tres componentes juntos, es fácil comprobar cómo cada uno contribuye a la utilidad bruta. Esto nos ayuda a identificar meses en los que los costos fueron inusualmente altos o los ingresos fueron excepcionalmente bajos.

Los dos hipervínculos que contiene COVER nos llevan a las otras dos hojas (Dashboard y U.S.A).

* **Dashboard:** todas las páginas poseen 3 segmentadores para filtrados intuitivos (Categoría, Año, Mes)

Contiene un boton para volver a COVER, y **4 bookmarks:**

* **Sales:** Cuenta con 4 tarjetas de KPi’s: TotalSales, TotalSalesDIFF, UniqueProducts, SoldProducts. También podemos encontrar una matriz y un “clustered column chart” que nos ayudan a visualizar una comparación completa de periodos tanto de mes como de año para las ventas totales.. La matriz nos muestra datos en mayor crudeza, logrando complementar la buena diferenciación visual que nos brinda el gráfico logrado. El mismo posee un tooltip que nos brinda las variaciones en porcentaje de las ventas totales. Otro gráfico presente es el de columnas apiladas, que representa la cantidad de productos vendidos por periodo, además hay un gráfico de barras reflejando el total de ventas por categoría en cada periodo. Se enseña un tacómetro reflejando el porcentaje de ventas totales.
* **COGS**: Contiene un “clustered column chart” que no ayuda a visualizar una comparación completa de periodos tanto de mes como de año para los COGS, contiene un tooltip que nos muestra porcentajes de variación de COGS, además una matriz que respalda el gráfico.

Cuenta con 5 tarjetas KPi’s: TotalCOGS, TotalCogsLY, TotalCogsDiff

Estas tres tarjetas nos muestran como varían entre año los COGS y cual es su diferencia.

Las otras dos tarjetas son para el Operational Ratio Cost actual y del periodo anterior.Se enseña un tacómetro reflejando el porcentaje de TotalCOGS. Contiene un “line clustered column chart”, él mismo nos enseña la diferencia entre el operational ratio cost por periodo. Además hay un gráfico de barras reflejando el total COGS por categoría en cada periodo.

* **Revenues:** Contiene dos “line clustered column chart” que reflejan utilidad bruta, neta por periodo y a través de una línea se compara su diferencia. Ambos gráficos están respaldados por sus respectivas matrices. También hay un clustered column chart que nos compara utilidad bruta y neta en cada categoría y subcategoría de productos (también respaldado por una matriz). Hay 3 tarjetas de KPi’s. Las mismas reflejan el porcentaje de margen neto o bruto y la diferencia que hay entre ellos.
* **Customer:** Contiene un mapa mundial con un tooltip por país que refleja todos los aspectos solicitados y un segmento para filtrar por medida preferente. (ejemplo: Total COGS, Variación de utilidad neta, etc)
* **U.S.A:** Contiene el mismo parámetro de campo utilizado anteriormente, vinculado con un grupo de cálculo para ahorrar medidas y aumentar la performance del modelo. También contiene un tooltip`informativo.

Está compuesto por un clustered bar chart, que no es afectado por el parámetro de campo, que refleja TotalCOGS y utilidad bruta por período entre las ciudades de U.S.A. También contiene un line stacked column chart que refleja Total de ventas por período y contiene un segundo eje Y, graficado como una línea que refleja el porcentaje de variación por período.

Contiene 2 tarjetas KPi’s: Sold Products, UniqueProducts y un botón para volver al COVER

Lo más importante de esta página es la matriz vinculada al grupo de cálculo que nos arroja datos en bruto para las provincias y ciudades de U.S.A filtrando a gusto con el parámetro de campos. (También contiene 3 segmentadores al igual que las otras páginas del dashboard que reflejan Año,Mes y Categoría.

# 

# 

# 

# Resultados principales y líneas futuras de análisis

* **COVER:** Se aprecia como insight clave del modelo que la empresa creció en todo ámbito mas de un 250% esto es por ver una simple comparativa (Clustered column chart) entre periodos en los aspectos de TotalSales, Cogs y Utilidades .
* **Dashboard:** Destaco como insight clave que el país qué mas generó fue U.S.A y que la categoría de Bicicletas arrasó en comparativa con las otras categorías.
* **Sales:** Se vendieron 53mil productos en comparación al 2012 que se vendieron solo 3mil. Se amplió la variedad de productos poco más de un 45% en comparativa al año anterior.
* **COGS:** Se refleja un ratio operational cost con un valor clave para reflejar el grandioso crecimiento de la empresa **(61%).** Si el coeficiente de explotación de una empresa es 0,60, es decir, el 60%, significa que por cada dólar de ventas generado se gastan 0,60 en gastos de explotación.

Los 0,40 dólares restantes se destinan a gastos no operativos o se convierten en ingresos netos, que pueden conservarse como beneficios no distribuidos o repartirse como dividendos a los accionistas.

* **Revenues:** La clave estuvo en el margen de utilidades. El margen de utilidad bruta(41.42%) y neta (30.9%) (Diferencia: 10.5%) Lo más importante es que tenemos casi un 31% final de ganancia
* **Customer:** Nos arroja el insight de que pese a que estados unidos es el que más utilidad genera, australia lo sigue con una diferencia de 200mil dólares, casualmente Australia tiene la mitad de clientes que estados unidos, por lo tanto sería ideal generar una campaña para ampliarse en Australia, ya que en un futuro promete más que Estados Unidos.
* **U.S.A:** California lideró en utilidad generando por sí sola alrededor de 50%, se recomienda generar una campaña para invertir en ciudades aledañas. U.S.A compró un 33% de la cantidad de productos que vendió la empresa.

# Reflexión personal

# Siento que power bi tiene una escalabilidad de aprendizaje alta, que al principio cuesta debido a la cantidad de elementos que posee, a pesar de esto con el tiempo se vuelve muy interactivo y dinámico. Ahora poseo habilidades para enfocarme tanto en la calidad de los datos como en representarlos de una forma visualmente atractiva. Me siento muy cómodo con DAX, debuggie muchísimos aspectos en distintas funciones, rompí y arreglé muchas veces el modelo, lo cual me proporcionó una muy buena habilidad con la herramienta. Respecto a la pregunta… NO! Luego de haber transformado los datos utilizaría de entrada las herramientas de los parámetros de campo, grupos de cálculo y la función CALENDAR(), para simplificar y a la vez optimizar de entrada el modelo.