***Visualizando el rendimiento de AWC con Power BI* 🚀**

**Nombre del autor: Agustina Maria Luaces Perez**

**Email: agustinaluaces@yahoo.com**

**Cohorte: DAPT-02**

**Fecha de entrega: 08/08/2024**

**Institución: Adventure Works Cycle (AWC)**

****

# Introducción

# Se desarrolla un informe integral en Power BI que proporciona el análisis detallado y las distintas visualizaciones interactivas sobre el rendimiento de ventas de la empresa AWC. Este informe cumple con las consignas dadas a lo largo del módulo M3 y permite a los usuarios finales comprender profundamente los factores que afectan a los ingresos, los costos y la rentabilidad, facilitando así la toma de decisiones estratégicas basadas en datos.

# Los objetivos alcanzados a lo largo de este proyecto son:

* Conectar y preparar la base de datos de AWC, mejorando la calidad de la misma para su análisis en Power BI.
* Diseñar el mockup en Powerpoint y crear un modelo de datos relacional que refleje las necesidades del negocio y sirva para responder preguntas claves sobre ingresos, costos y rentabilidad.
* Utilizar DAX para calcular métricas claves a través de la creación de medidas y columnas calculadas.
* Crear el tablero final en Power BI, siendo visualmente atractivo, interactivo y coherente para los usuarios finales.

# El reporte financiero final consta de una portada y dos páginas, una general y una detallada sobre el mercado de Estados Unidos.

# Desarrollo del proyecto

En el primer avance, descargo la base de datos “AdventureWorksDW2019” y la restauro en SQL Server siguiendo los pasos otorgados en la consigna. Posteriormente, conecto esta BD restaurada a Power BI seleccionando las siguientes tablas:

* DimProduct
* DimProductCategory
* DimProductSubcategory
* DimDate
* DimPromotion
* DimSalesTerritory
* DimGeography
* FactInternetSales

Luego, conecto el Excel “DimCustomer” y comienzo la transformación en Power Query. Verifico que los encabezados de las tablas se generan automáticamente y con la calidad de columnas observó aquellas con alto porcentaje de “Vacíos”. A partir de esto y con el contexto del negocio que tengo, intento interpretar si tienen utilidad para el análisis de datos que voy a hacer. Borro columnas totalmente vacías y con valores válidos < 1%, principalmente en la tabla “DimCustomer” y notó la importancia de los pasos aplicados para retrotraer alguna acción.

Particularmente, no eliminó tantas columnas o filas nulas teniendo en cuenta que en este paso como analista de datos me gustaría tener reuniones previas con el negocio y las partes interesadas para volver a chequear sus objetivos y ver si de alguna forma se pueden subsanar estos nulos por si se necesitan en el futuro.

En cada tabla cargada, identificó los PK y los FK necesarios para realizar las relaciones. Combino la tabla “DimCustomer” con “DimGeography” con la opción Externa Izquierda que sería como un LEFT JOIN de SQL. De esta forma, traigo los campos que contienen la ciudad, la provincia y su código a la primera tabla.

Posteriormente, combino las tablas “DimProduct” con “DimProductCategory” y “DimProductSubcategory”, teniendo en cuenta las PK y FK ya existentes para que la relación entre las tablas sea exitosa y el modelo sea lo más estrellado posible. Eliminó las combinaciones entre tablas existentes que no voy a utilizar (especialmente en “DimProduct”) para que la visualización y el recorrido al editar las consultas sea más amigable y resulte más sencillo entender qué datos estoy manejando.

Elimino los vacíos en la columna “CustomerKey” de la tabla “DimCustomer” ya que se trata de una clave PK y de otra forma las relaciones entre tablas no se puede generar. Activo las relaciones que necesito y borro aquellas que son redundantes. Finalmente, decido usar “DimDate” como tabla calendario y no generar una nueva ya que está unida directamente a la tabla de ventas.

En el segundo avance, realizo el diseño del mockup teniendo en cuenta las preguntas a responder otorgadas en la consigna. En este paso analizo en profundidad qué tipo de gráficos me servirían para representar de forma eficiente la información solicitada. También tengo en cuenta el patrón Z para facilitar la comprensión y el análisis de los usuarios finales. Genero el logo de la empresa y decido utilizar Power Point para crear el mockup ya que comparte muchos items de diseño con Power BI al tratarse de dos herramientas propias de Microsoft (*ver anexo “DA\_Luaces\_Agustina\_MockupPIM3”*).

A lo largo del tercer avance, realizo en Power Query una columna personalizada con el nombre del mes en formato corto en la tabla “DimDate”, deshabilito la carga de las tablas “DimProductCategory”, “DimProductSubcategory” y “DimGeography, marco “DimDate” como tabla de fechas con la columna “FullDateAlternateKey” y creo una columna calculada en la tabla “DimDate” indicando el trimestre como “Trimestre XX”.

Además, genero las medidas necesarias para resolver las preguntas detalladas en el avance anterior y creo una tabla de Medidas y carpetas en ella para ordenarlas y facilitar la búsqueda al ir generando las distintas visualizaciones en el tablero.

PREGUNTAS y RESPUESTAS:

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

Hago la medida rápida “MoM% de SalesAmount” de cambio mes a mes con las fechas “FullDateAlternateKey” de la tabla “DimDate” y con el campo “SalesAmount” de la tabla “FactInternetSales”. Creó el gráfico de barras para la suma del campo SalesAmount (medida con DAX) mes a mes desde 2011 a 2013 y la medida rápida MoM%.

* **¿Cuál es la cantidad vendida?**

El campo Cantidad Vendida lo hago con una medida y DAX de la suma SalesAmount y lo muestro en una tarjeta (Ingresos Totales).

* **¿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?**

Utilidad Bruta = SalesAmount- (TotalProductCost + Freight)

El costo de flete (“Freight”) generalmente se incluye como parte de los costos directos asociados con la producción y venta de productos. Por lo tanto, el costo de flete se considera en el cálculo de la utilidad bruta. Esta fórmula se traduce en DAX para generar la medida, restando el costo total del producto y el costo de flete de los ingresos totales por ventas, sin incluir el monto de impuestos sobre las ventas.

Utilidad Neta = Utilidad Bruta - TaxAmt

Esta fórmula se traduce en DAX para calcular la utilidad neta en una medida, restando los impuestos a la medida de la Utilidad Bruta.

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

COGS = TotalProductCost

Esta fórmula tiene en cuenta solo el costo total del producto. La fórmula se traduce en DAX para crear la medida correspondiente.

* **¿Cuántos clientes hay en cada país? El usuario desea ver esta demografía representada en mapas.**

Utilizo el campo “SalesTerritoryCountry” de la tabla “DimSalesTerritory”, que está relacionada con la tabla “FactInternetSales” mediante “CustomerKey”. Finalmente, la cantidad de clientes la calculo con el “CustomerKey” que es un ID único para cada cliente.

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

Relacionada con las primeras preguntas. Se hicieron las medidas necesarias para visualizar cada una de ellas.

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

Se utilizan las medidas generadas anteriormente y se organizan por categoría y subcategoría, creando un parámetro para usar el mismo gráfico en el reporte.

* **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).**

Genero los márgenes dividiendo las métricas por los ingresos totales y el ratio mediante esa fórmula en DAX (DIVIDE()).

* **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.**

Con las medidas creadas, se seleccionan las necesarias y se filtran por el país United States.

* **El usuario quiere tener la posibilidad de segmentar la información en ambos casos (general y detalle USA) por año y categoría de producto.**

Se establecen los filtros correspondientes en el reporte *(ver anexo DA\_Luaces\_Agustina\_ReportePIM3)*

Una vez creadas las medidas necesarias con DAX, se crea el reporte final en Power BI en el avance cuatro. Este reporte se centra en los ingresos, costos y rentabilidad con detalle en el mercado de Estados Unidos. Genero las visualizaciones basandome en el mockup, incorporando los parámetros creados, los segmentadores y filtros correspondientes, títulos dinamicos para los gráficos, formato condicional, grupo de cálculo, navegadores, marcadores y tooltip. Estas herramientas aprendidas fueron de gran utilidad, para crear una narrativa coherente con lo solicitado por el negocio, visualmente atractivo y con gran nivel de interactividad.

# Análisis general del tablero

# En el mismo tablero adjunto marcadores para que se muestre la vista de ayuda mediante botones a los usuarios finales, describiendo los componentes y lo que implica cada visualización. Fue un proyecto al que le dediqué muchas horas para poder utilizar todas las herramientas aprendidas y lograr familiarizarme con ellas e incorporarlas en mi futuro labor como analista de datos. Además de las lecture vistas en el módulo, recurri a IA y diversos videos de YouTube para despejar dudas.

# Resultados principales y líneas futuras de análisis

# En una revisión rápida al tablero creado, cómo analista de datos puedo observar que el mercado de Estados Unidos es el más significativo pero para el mercado Australiano también se podría tener un informe detallado ya que tiene menos clientes pero con ingresos elevados. Por otro lado, noto que en el año 2014 tengo faltante de datos (hasta marzo 2014) pero puedo decir que desde la mitad del año 2013 en adelante hubo un incremento en las ventas. Esta falta de datos puede ser subsanada con la empresa a futuro si necesitan evaluar una mayor ventana de tiempo.

La categoría que más utilidades tiene es la de las bicicletas, permitiendo al negocio revisar si hay posibilidades de expansión en las demás categorías o si en cambio se deberá reforzar el stock de los productos más vendidos (subcategorías) para abastecer el mercado existente. Estos insights se pueden comunicar a la parte interesada, para que el negocio pueda establecer el camino a seguir mediante la toma de decisiones estratégicas basadas en datos.

# Reflexión personal

Durante este proyecto fui capaz de incorporar muchas herramientas nuevas y sentí que finalmente pude usar Power BI de manera profesional y responsable, creando un tablero que percibo como amigable, fácil de navegar, coherente y presentable.

Si tuviese que volver a empezar este proyecto, intentaría entender un poco más las tablas y cuales son los objetivos del cliente. Aún en esta fase del aprendizaje en el rol de analista de datos, creo que esto recién lo comprendo completamente cuando veo las distintas relaciones volcadas en gráficos y no tanto cuando hago el modelo de datos relacional. También considero que es cuestión de práctica a medida que me involucre en más proyectos de distinta complejidad.

**EXTRA CREDIT**

Realizo un gráfico de líneas de la medida “Ingreso Totales” vs tiempo, utilizando “DueDate” ya que necesito un campo calendario que tenga jerarquización de fechas (no uso la tabla “DimDate” establecida como tabla de fechas). En la parte de Analytics, selecciono la opción de Previsión y observo que automáticamente me da una línea recta desde el último punto tomado, sin que esto tenga mucho sentido para la previsión.

Empiezo a modificar los parámetros dados como predeterminados y defino reducir la duración a 6 puntos, omitiendo los últimos 4 puntos con 4 puntos de estacionalidad y con un intervalo de confianza más angosto (90% - menor diferencia entre los límites superior e inferior y menor precisión de que el dato real esté en ese rango). De esta forma, puedo tener una predicción distinta a una línea recta con los datos otorgados, pudiendo comparar los últimos 4 datos reales con los datos generados por la previsión y notando que los mismos se encuentran en el rango de confianza establecido.

Estos resultados sirven para un análisis predictivo rápido y pueden mostrarse a la parte interesada, teniendo en cuenta que los datos históricos utilizados para la predicción son relevantes y de calidad gracias al proceso de transformación y limpieza realizado al principio de este proyecto. El gráfico obtenido se muestra en el reporte de USA, utilizando marcadores y botones para visualizarlo.

# 