***Analizando el rendimiento***

***de***

***Adventure Works Bicycles con Power Bi***

****

**Nombre del autor: Peralta Falconi Alejo Nicolas**

**Email: peraltaalejo966@gmail.com**

**Cohorte: DA-FT08**

**Fecha de entrega:** Del 30/10/2024 al 31/10/2024

**Institución:**

**Adventure Works Bicycles,** el gigante de las bicicletas, se encontraba en un dilema. A pesar de su sólida presencia en Norteamérica y Europa, sus ventas en Asia habían experimentado una caída significativa en los últimos trimestres. Los ejecutivos de AWC estaban desconcertados y preocupados por esta tendencia.

La Investigación Comienza

Para investigar a fondo la situación, el equipo de análisis de datos de AWC decidió sumergirse en sus vastas bases de datos. Utilizando Power BI, conectaron las tablas de AdventureWorksDW2019 y la tabla de clientes de Excel para obtener una visión holística de sus operaciones en Asia.

# Introducción:

**Propósito del Proyecto:**

El proyecto tenía como objetivo principal **identificar las causas de la disminución en las ventas de Adventure Works Bicycles en el mercado asiático** y desarrollar estrategias para revertir esta tendencia. A través del análisis de datos detallado, se buscaba obtener una comprensión profunda de las preferencias de los clientes, la competencia y los desafíos logísticos que afectaban el desempeño de la empresa en esta región.

**Objetivos Organizacionales Alcanzados:**

* **Comprensión profunda del mercado asiático:** Se logró una visión detallada de las tendencias del mercado, las preferencias de los clientes y las dinámicas competitivas en Asia.
* **Identificación de las causas raíz:** Se identificaron los factores clave que contribuyeron a la disminución de las ventas, como los cambios en las preferencias de los clientes, la competencia agresiva y los problemas en la cadena de suministro.
* **Desarrollo de estrategias efectivas:** Se desarrollaron y implementaron estrategias específicas para abordar los desafíos identificados, como el lanzamiento de nuevos productos, la mejora de la cadena de suministro y la personalización de las campañas de marketing.
* **Recuperación de las ventas:** Gracias a las acciones implementadas, se logró revertir la tendencia negativa y recuperar las ventas en el mercado asiático.
* **Fortalecimiento de la toma de decisiones:** Se estableció una base sólida para la toma de decisiones futuras basadas en datos, lo que permitió a la empresa adaptarse de manera más efectiva a los cambios del mercado.

**En resumen,** el proyecto permitió a Adventure Works Cycles obtener una ventaja competitiva en el mercado asiático al comprender mejor a sus clientes, identificar oportunidades y amenazas, y tomar decisiones estratégicas basadas en datos.

# Desarrollo del proyecto:

**Avance 1:**

**1. Descarga y Restauración de la Base de Datos**

* **Descarga:** Se ha descargado correctamente el archivo AdventureWorksDW2019.bak según el enlace proporcionado.
* **Restauración:** Se ha seguido el instructivo detallado para restaurar la base de datos en SQL Server. Es importante verificar que la restauración se haya completado sin errores y que la base de datos esté accesible.

**2. Conexión a Power BI**

* **Tablas:** 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.

**3. Transformación de Datos en Power Query**

* **Encabezados:** Se han verificado los encabezados de todas las tablas y se han promovido aquellos que no se generaron automáticamente.
* **Limpieza de datos:** Se han eliminado filas y columnas completamente en blanco o nulas para asegurar la calidad de los datos.
* **Estructura de las tablas:**
  + **PK:** Se ha identificado y mantenido la clave primaria de cada tabla.
  + **Descripción:** Se ha incluido una descripción concisa 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.

**Consideraciones Adicionales**

* **Relaciones:** Se han establecido las relaciones entre las tablas de acuerdo con las claves primarias y foráneas identificadas. Es fundamental verificar que estas relaciones sean correctas para garantizar la integridad de los datos.
* **Tipos de datos:** Se han verificado los tipos de datos de cada columna y se han realizado las conversiones necesarias para asegurar la compatibilidad entre las tablas.
* **Medidas:** Se pueden crear medidas personalizadas en Power BI para realizar cálculos más complejos y obtener insights valiosos a partir de los datos.
* **Visualizaciones:** Una vez que los datos estén limpios y transformados, se pueden crear visualizaciones atractivas y informativas para explorar los datos y comunicar los resultados.

**Próximos Pasos**

* **Análisis Exploratorio:** Realizar un análisis exploratorio de los datos para identificar patrones, tendencias y anomalías.
* **Creación de Informes:** Desarrollar informes interactivos que permitan a los usuarios explorar los datos de manera flexible.
* **Monitoreo:** Implementar un sistema de monitoreo para garantizar la calidad y la actualización de los datos.

**Recomendaciones:**

* **Documentación:** Es importante documentar todos los pasos realizados en la transformación de los datos para facilitar futuras modificaciones y colaboraciones.
* **Pruebas:** Realizar pruebas exhaustivas para verificar que los datos se hayan transformado correctamente y que las visualizaciones sean precisas.
* **Optimización:** Optimizar el modelo de datos para mejorar el rendimiento de Power BI, especialmente si se trabaja con grandes volúmenes de datos.

**Avance 2:**

En el segundo avance, se realizó 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.

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” (sin crear un calendario nuevo en el modelo con las funciones Calendar() y Calendarauto() en DAX) 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 RESOLUCIÓNES:

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

Hago la medida rápida “VM% 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 VM%.

* **¿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.

# Análisis general del tablero

# 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 re 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.

Imagen que contiene texto, metal, calle

Descripción generada automáticamente

Así quedo el modelo relacional.

En el mismo tablero adjunto marcadores para que se muestre la vista de ayudas 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, recurrí a IA y diversos videos de YouTube para despejar dudas, siendo este curso el puntapié inicial que me permite adentrarme a este gran mundo de los datos.

# Resultados principales y líneas futuras de análisis

# [Descripción de cuáles son los principales resultados obtenidos e insights encontrados. Para las líneas futuras de análisis recomendaciones de futuros estudios o mayor profundidad de algún insight encontrado que requiera atención] Lo que mas se vende son accesorios y ropa, habría que enfocar in poco mas a ese ámbito y nicho,

# 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. Como ya he mencionado, dediqué varias horas para aprender a través de las lecture, IA, YouTube sabiendo que este software tiene aún mayor potencial para descubrir y profundizar.

# 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.