PROYECTO ANÁLISIS DE DATOS

MINORISTA MUNDIAL DE PRODUCTOS ELECTRÓNICOS

Contenido

Método del análisis:	2
Acerca del dataset:	2
Usuario final:	2
Necesidades y Prioridades del Usuario:	2
Objetivos del análisis	3
Tablero de control en Power BI:	15
Conclusiones:	18

Método del análisis:

El análisis se va a desarrollar respondiendo a los objetivos mediante consultas en SQL y finalizará con la creación de un tablero de control en Power Bl.

Acerca del dataset:

Contiene datos de ventas de un minorista mundial de productos electrónicos ficticio, incluidas tablas que contienen información sobre transacciones, productos, clientes, tiendas y tipos de cambio de divisas. Los datos contenidos van desde 2016 hasta los primeros dos meses del 2021.

El conjunto de datos es de dominio público y se puede acceder a él mediante el siguiente link:

https://mavenanalytics.io/data-playground?order=date_added%2Cdesc&tags=Retail_el nombre es "Global Electronics Retailer"

Usuario final:

Nombre: Lucas Martín González

Rol: Director Global de Operaciones y Estrategia Comercial

Empresa: ElectroWorld

Descripción del Rol:

Lucas es el encargado de supervisar y optimizar las operaciones globales de ElectroWorld, un minorista líder en productos electrónicos. Con más de 10 años de experiencia en estrategia y análisis comercial, Lucas se enfoca en identificar oportunidades de crecimiento, mejorar la rentabilidad y garantizar una experiencia consistente para los clientes en todos los mercados. Su trabajo abarca desde la evaluación del desempeño de tiendas hasta la optimización de precios y análisis del comportamiento del cliente.

Necesidades y Prioridades del Usuario:

1. Maximizar la rentabilidad global:

Lucas necesita identificar qué mercados, productos y clientes generan el mayor valor para la empresa y cómo pueden mejorarse las operaciones menos rentables.

2. Evaluar eficiencia operativa:

Está interesado en analizar el rendimiento de las tiendas, especialmente en relación con su tamaño y antigüedad, para mejorar la eficiencia.

3. **Entender al cliente:**

Lucas quiere un análisis profundo del comportamiento del cliente para

segmentar su base de datos, identificar patrones y mejorar las campañas de marketing y retención.

Objetivos del análisis

1. Análisis de ventas globales:

- **Objetivo:** Evaluar el desempeño de las ventas por país, región y tienda.
 - Calcular los ingresos totales por país y continente.
 - Identificar las tiendas más y menos rentables basándose en los ingresos totales.

2. Análisis del comportamiento del cliente:

- **Objetivo:** Segmentar a los clientes para identificar patrones de compra y perfiles.
 - Agrupar clientes por género, país y estado para observar diferencias en los patrones de compra.
 - Analizar la edad promedio de los clientes y su relación con las categorías de productos más comprados.
 - Calcular la tasa de recompra y la cantidad promedio de compras por cliente.

3. Rendimiento de productos:

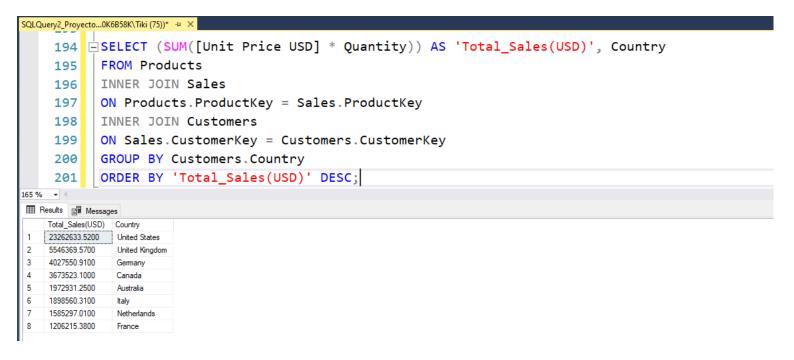
- **Objetivo:** Evaluar la popularidad y rentabilidad de los productos.
 - Determinar los productos más vendidos y las categorías/subcategorías más populares por región.
 - o Identificar los productos con mayores márgenes de ganancia.
 - Analizar qué marcas tienen un desempeño superior en ingresos totales y volumen de ventas.

4. Desempeño de las tiendas:

- Objetivo: Comparar el desempeño de las tiendas en función de su tamaño y antigüedad.
 - Relacionar las ventas totales de cada tienda con sus metros cuadrados para analizar eficiencia.
 - Evaluar cómo el tiempo de apertura de las tiendas impacta el volumen de ventas.
 - Identificar los países con mayor concentración de tiendas y su impacto en las ventas totales.

Análisis de ventas globales:

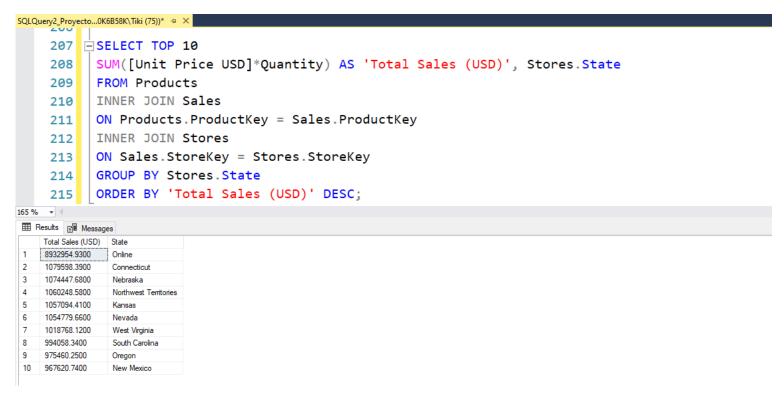
Comenzando por analizar las ventas globales cree una consulta uniendo 3 tablas mediante un JOIN para obtener las ventas totales por cada país:



Luego con una consulta similar obtuve las ventas totales agrupadas por continente:

```
SQLQuery2_Proyecto...0K6B58K\Tiki (75))* → ×
            SELECT (SUM([Unit Price USD] * Quantity)) AS 'Total_Sales(USD)', Continent
    204
    205
           FROM Products
           INNER JOIN Sales
    206
           ON Products.ProductKey = Sales.ProductKey
    207
    208
          INNER JOIN Customers
    209
           ON Sales.CustomerKey = Customers.CustomerKey
           GROUP BY Customers.Continent
    210
           ORDER BY 'Total_Sales(USD)' DESC;
    211
165 % + 4
Results Messages
    Total Sales(USD) Continent
   26936156.6200 América
    14263993.1800
             Europe
    1972931.2500
```

Para analizar las tiendas más rentables continué con la dinámica de las consultas anteriores y le sumé la sentencia TOP 10 ordenando descendentemente por el total de las ventas, dando como resultado las 10 tiendas con más ventas totales.

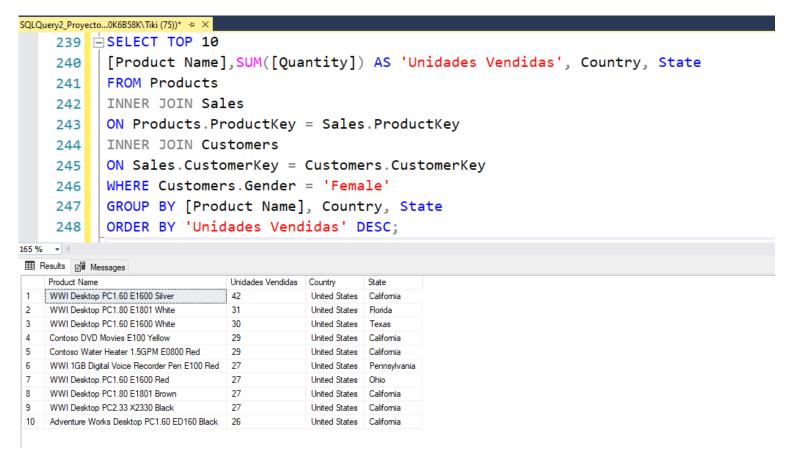


Para obtener las menos rentables utilicé la misma consulta, pero esta vez ordenando los resultados ascendentemente por las ventas totales.

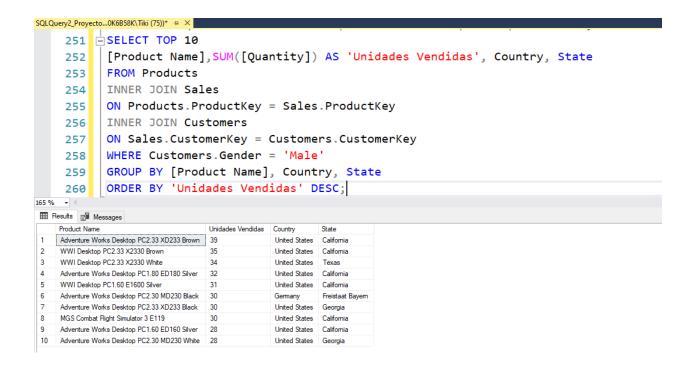
```
SQLQuery2_Proyecto...0K6B58K\Tiki (75))* → ×
     225 SELECT TOP 10
              SUM([Unit Price USD]*Quantity) AS 'Total Sales (USD)', Stores.State
     226
              FROM Products
     227
              INNER JOIN Sales
     228
              ON Products.ProductKey = Sales.ProductKey
     229
              INNER JOIN Stores
     230
              ON Sales.StoreKey = Stores.StoreKey
     231
              GROUP BY Stores.State
     232
     233
              ORDER BY 'Total Sales (USD)';
165 % +
Results Messages
    Total Sales (USD) State
    9380.0100
                 Northern Territory
    91653.0600
                 Franche-Comté
    122704.2000
    137518.5000
                 Martinique
    150018.6800
                 Limousin
6
    152756.2400
                 Basse-Normandie
    156696.2800
                 Caltanissetta
    167513.4700
                 La Reunión
8
9
    170238.7200
                 Mayotte
 10
    174302.2300
                 Australian Capital Territory
```

Análisis del comportamiento del cliente:

Avanzando con el análisis del comportamiento de los clientes, cree una consulta que me muestra los 10 productos más comprados filtrando por el género mujeres, indicando su país y Estado.



La misma consulta utilicé para descubrir los 10 productos más comprados por los hombres, cambiando el filtro para dicho género.



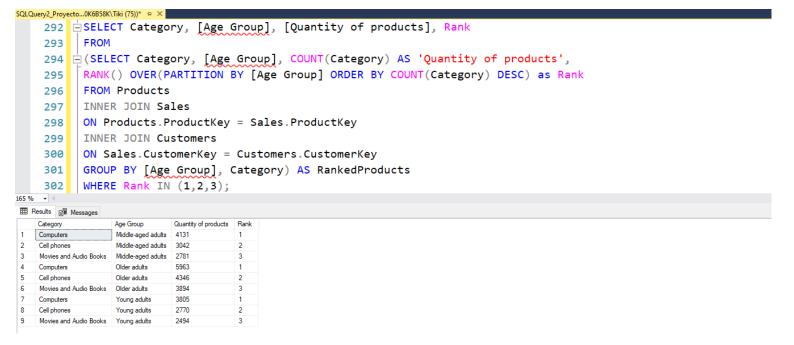
Para analizar la edad promedio de los clientes y su relación con las categorías de productos más comprados tuve que, primero crear una columna que calcule la edad de los clientes:

```
SQLQuery2_Proyecto...0K6B58K\Tiki (75))* → ×
          --Agrego la columna Edad para los clientes
    262
        ALTER TABLE Customers
    263
          ADD Age INT;
    264
    265
          --Realizo el cálculo de la edad para cada uno
    266
         UPDATE Customers
    267
          SET Age = DATEDIFF(YEAR, Birthday, GETDATE())
    268
    269
                       - CASE WHEN MONTH(Birthday) > MONTH(GETDATE())
                           OR (MONTH(Birthday) = MONTH(GETDATE()) AND DAY(Birthday) > DAY(GETDATE()))
    270
                           THEN 1
    271
                            ELSE 0
    272
                       END;
    273
```

Luego agregué una columna más para segmentar las edades en rangos etarios y asigné cada uno mediante una declaración CASE:

```
SQLQuery2_Proyecto...0K6B58K\Tiki (75))* → ×
          --Agrego otra columna más para segmentar las edades en rangos
    278
        ALTER TABLE Customers
    279
          ADD "Age Group" VARCHAR(50);
    280
    281
          --Calculo el grupo etario en base a rangos de edad
    282
        UPDATE Customers
    283
          SET [Age Group] = CASE
    284
                                WHEN Age >= 22 AND Age <=40 THEN 'Young adults'
    285
    286
                                WHEN Age >= 41 and Age <=60 THEN 'Middle-aged adults'
                                WHEN Age >60 THEN 'Older adults'
    287
    288
                                ELSE 'Uncategorized'
    289
                                END;
    290
```

Finalmente cree una consulta para obtener las 3 categorías de productos más comprados por cada rango etario, utilizando una función ventana RANK en una subconsulta que me devuelve un ranking con los productos comprados por cada rango etario agrupados por categorías y ordenados descendentemente, para luego filtrar los 3 más comprados:



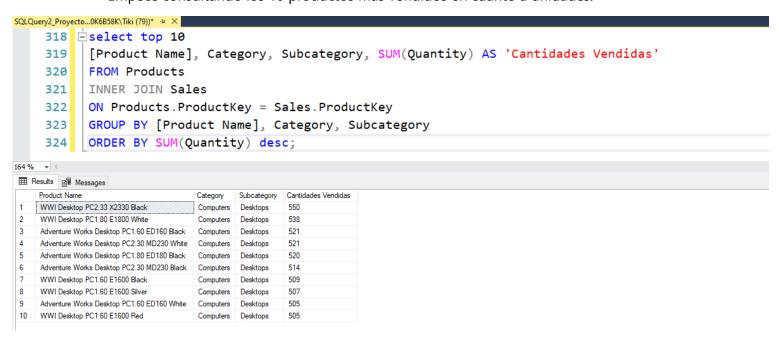
Siguiendo por calcular la **tasa de recompra** utilicé dos subconsultas dentro de la consulta principal, una para realizar el cálculo de la cantidad de clientes que tenemos y la otra para obtener la cantidad de clientes que han realizado más de una compra. Luego dividí el total de clientes con recompras por el total de clientes y multipliqué por 100 para obtener la tasa de recompra, que es del 86,75%.

```
SQLQuery2_Proyecto...0K6B58K\Tiki (79))* → ×
    305 SELECT ROUND((CAST(COUNT(*) AS FLOAT) /
    306
           (select count(distinct CustomerKey) from Sales)),4) * 100 AS [Tasa de recompra (%)]
    307
           select CustomerKey
    308
           from Sales
    309
           GROUP BY CustomerKey
    310
           having count(CustomerKey) > 1) AS compras_repetidas;
    311
164 %
Results 🗐 Messages
   Tasa de recompra (%)
   86.75
```

Finalmente obtuve la cantidad de compras promedio por cliente mediante una consulta sencilla:

Rendimiento de productos:

Empecé consultando los 10 productos más vendidos en cuanto a unidades.

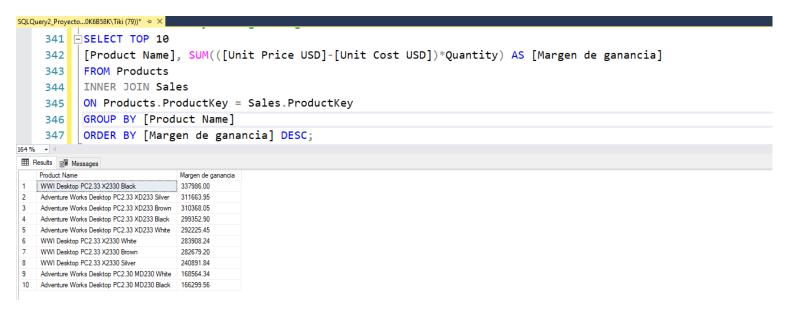


Continué creando una CTE para rankear las categorías más vendidas de cada país y luego mediante una consulta seleccionar las 3 primeras de cada uno dando como resultado una lista con las 3 categorías de productos más vendidos por cada país:

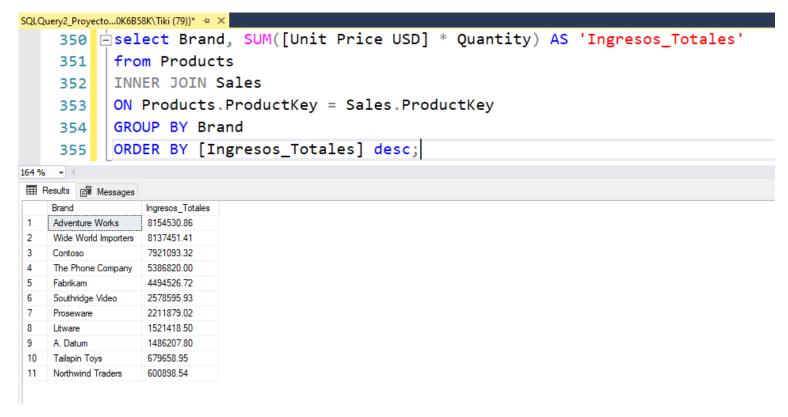
```
SQLQuery2_Proyecto...0K6B58K\Tiki (79))* → ×
    327 with ventas_productos as (
              select Country, Category, Subcategory, SUM(Quantity) AS [Cantidades Vendidas],
    328
              DENSE_RANK() OVER(PARTITION BY Country ORDER BY SUM(Quantity) DESC, Category) AS Rank
    329
    330
              FROM Sales
              INNER JOIN Products
    331
              ON Sales.ProductKey = Products.ProductKey
    332
              INNER JOIN Stores
    333
              ON Sales.StoreKey = Stores.StoreKey
    334
    335
              GROUP BY Country, Category, Subcategory)
          select country, Category, Subcategory, [Cantidades Vendidas], Rank
    336
    337
          FROM ventas_productos
    338
          WHERE Rank IN (1,2,3);
```

	country	Category	Subcategory	Cantidades Vendidas	Rank
1	Australia	Movies and Audio Books	Movie DVD	1061	1
2	Australia	Computers	Desktops	581	2
3	Australia	Audio	Bluetooth Headphones	511	3
4	Canada	Movies and Audio Books	Movie DVD	1843	1
5	Canada	Computers	Desktops	1281	2
6	Canada	Audio	Bluetooth Headphones	902	3
7	France	Movies and Audio Books	Movie DVD	648	1
8	France	Computers	Desktops	464	2
9	France	Games and Toys	Boxed Games	280	3
10	Germany	Movies and Audio Books	Movie DVD	2335	1
11	Germany	Computers	Desktops	1560	2
12	Germany	Audio	Bluetooth Headphones	1009	3
13	Italy	Movies and Audio Books	Movie DVD	969	1
14	Italy	Computers	Desktops	735	2
15	Italy	Audio	Bluetooth Headphones	465	3
16	Netherlands	Movies and Audio Books	Movie DVD	801	1
17	Netherlands	Computers	Desktops	578	2
18	Netherlands	Audio	Bluetooth Headphones	416	3
19	Online	Movies and Audio Books	Movie DVD	5932	1
20	Online	Computers	Desktops	4342	2
21	Online	Audio	Bluetooth Headphones	2835	3
22	United Kingdom	Movies and Audio Books	Movie DVD	3071	1
23	United Kingdom	Computers	Desktops	2245	2
24	United Kingdom	Audio	Bluetooth Headphones	1527	3
25	United States	Movies and Audio Books	Movie DVD	12142	1
26	United States	Computers	Desktops	8840	2
27	United States	Audio	Bluetooth Headphones	5731	3

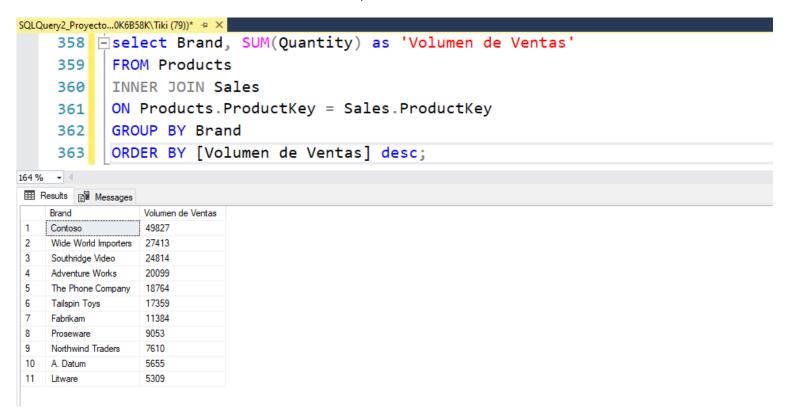
Creé una consulta que devuelve los 10 productos que mayor margen de ganancia generan:



Para finalizar con el rendimiento de los productos obtuve mediante una consulta simple los ingresos totales por cada marca ordenadas descendentemente:



Y los volúmenes de ventas de cada una, también ordenadas descendentemente:



Desempeño de las tiendas:

Comencé por analizar qué tiendas tienen el mayor ingreso por metro cuadrado, utilizando JOINS para unir las 3 tablas necesarias y la función NULLIF en el cálculo de la división, ya que en el caso de la tienda online los metros cuadrados son 0 y la división daría error:

```
SQLQuery2_Proyecto...0K6B58K\Tiki (63))* → ×
    371  Select StoreKey, State,
    372
          [Square Meters],
          (select SUM([Unit Price USD]*Quantity)/ NULLIF([Square Meters],0)
    373
    374
          FROM Products
    375
          INNER JOIN Sales
    376
          ON Products.ProductKey = Sales.ProductKey
          WHERE Stores.StoreKey = Sales.StoreKey) AS 'Ventas totales por metro cuadrado'
    377
    378
          ORDER BY 'Ventas totales por metro cuadrado' desc, [Square Meters], State;
    379
    380
```

Acá la lista con algunas de las tiendas con más ganancias por metro cuadrado:

	StoreKey	State	Square Meters	Ventas totales por metro cuadrado
1	66	Wyoming	840	1108.912595
2	47	Hawaii	1120	809.927526
3	43	Alaska	1190	780.872445
4	56	New Hampshire	1260	727.029888
5	51	Maine	1295	723.422409
6	64	Washington DC	1330	714.229548
7	9	Northwest Territories	1500	706.832386
8	53	Montana	1260	689.827079
9	10	Nunavut	1210	688.859082
10	29	Enna	1000	688.640340
11	26	Saarland	350	666.338914
12	57	New Mexico	1645	588.219294
13	65	West Virginia	1785	570.738442
14	40	Dungannon and South Tyrone	1300	563.464584
15	36	Amagh	1300	550.786684

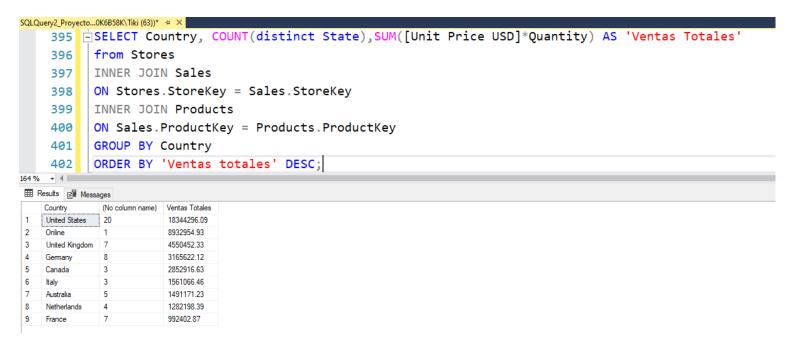
Continuando con el análisis de las tiendas cree una consulta que me permite obtener las ventas totales generadas por cada tienda y los días transcurridos desde su apertura para evaluar si el tiempo desde la apertura tiene un impacto en el volumen de ventas:

```
SQLQuery2_Proyecto...0K6B58K\Tiki (63))* → ×
    382 = select Stores.StoreKey,
    383
          State,
          DATEDIFF(DAY, [Open Date], GETDATE()) AS 'Días desde la apertura',
    384
          SUM([Unit Price USD]*Quantity) AS 'Ventas Totales'
    385
          FROM Stores
    386
          INNER JOIN Sales
    387
          ON Stores.StoreKey = Sales.StoreKey
    388
          INNER JOIN Products
    389
          ON Sales.ProductKey = Products.ProductKey
    390
          GROUP BY Stores.StoreKey, State, [Open Date]
    391
    392
          ORDER BY 'Ventas Totales' DESC, 'Días desde la apertura' DESC;
    393
```

Lista de las 15 tiendas que más ventas totales generaron y sus días desde la apertura:

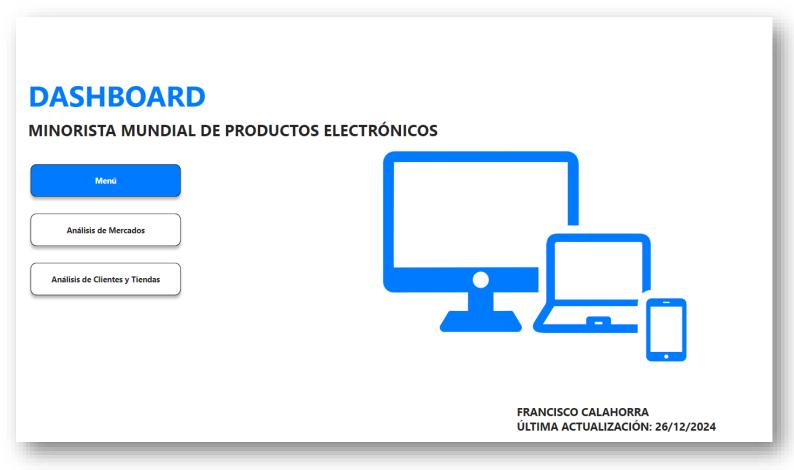
	StoreKey	State	Días desde la apertura	Ventas Totales
1	0	Online	5467	8932954.93
2	45	Connecticut	6375	1079598.39
3	54	Nebraska	4214	1074447.68
4	9	Northwest Territories	7231	1060248.58
5	50	Kansas	6133	1057094.41
6	55	Nevada	5484	1054779.66
7	65	West Virginia	4737	1018768.12
8	61	South Carolina	4388	994058.34
9	59	Oregon	4517	975460.25
10	57	New Mexico	5314	967620.74
11	8	Newfoundland and Labrador	3824	959148.56
12	64	Washington DC	5467	949925.30
13	44	Arkansas	5314	939762.31
14	51	Maine	5467	936832.02
15	66	Wyoming	4006	931486.58

Para finalizar esta parte del análisis cree una consulta que me proporciona la cantidad de tiendas por país y las ventas totales para cada uno de ellos, para analizar si la cantidad de tiendas en cada país incide en las ventas totales generadas por cada uno.



Tablero de control en Power BI:

El tablero presenta 3 solapas, la primera es el Menú con el título del proyecto y el menú de navegación para moverse entre las páginas:



La siguiente solapa se llama "Análisis de mercados" y permite visualizar cuáles son los países donde se obtienen mayores beneficios, las categorías con más ventas en usd, las subcategorías con sus ventas anuales y la variación porcentual de las ventas respecto al periodo anterior, y los beneficios totales de cada Estado. Todo esto pudiendo realizar una segmentación por categoría, año, país, Estado, género y grupo etario, permitiendo obtener un gran nivel de detalle:



La última solapa llamada "Análisis de clientes y Tiendas" proporciona información como la tasa de recompra, la cantidad de tiendas y clientes, el porcentaje de nacionalidades a la cual pertenecen nuestros clientes, también las ventas totales de las tiendas en los diferentes países y los productos más vendidos en usd, permitiendo al igual que la solapa anterior obtener un gran nivel de detalle mediante las segmentaciones:



Conclusiones:

En cuanto a rentabilidad, el mercado de las computadoras es ampliamente el más grande y el que mayor beneficio produce en general, seguido por los celulares y los electrodomésticos.

A su vez, el país en donde se obtienen mayores beneficios es Estados Unidos, seguido por Reino Unido y Alemania, no teniendo en cuenta la tienda online, ya que ocupa el segundo lugar en cuanto a beneficios generados.

Por el lado de los clientes tenemos una predominancia de adultos mayores (mayores de 60 años), en especial de mujeres, que son las que producen más ventas totales y, por ende, más beneficios para la empresa. Siendo sus artículos más comprados computadoras de escritorio.

En cuanto a volumen de ventas (unidades vendidas) los 10 productos con mayor cantidad también son computadoras de escritorio, siendo las de la marca "Adventure Works" y "Wide World Importers" las más compradas.

Siguiendo por el desempeño de las tiendas, y habiendo analizado factores como los metros cuadrados de cada tienda y su tiempo desde la apertura para ver si influyen en las ventas totales, puedo decir que estos son factores que no muestran incidencia en las ventas totales. Aquellas 3 que más ventas generaron pertenecen a Estados dentro de Estados Unidos, específicamente Connecticut, Nebraska y Northwest Territories en Canadá, dejando de lado la tienda online que es en realidad la que más generó por sí sola.

Por otro lado, tenemos que los productos con menos vendas en USD pertenecen a las categorías de juguetes, TV y video, y películas y audio libros. Creo que esto se debe principalmente al momento en el que vivimos en el que hay un cambio en los consumos, especialmente donde se van dejando de lado algunos productos como los juguetes para los niños y se van adentrando en la tecnología a una edad temprana. Otros productos como el DVD, juegos de mesa, MP4, simplemente quedaron obsoletos y poca gente lo consume.

Mi recomendación es priorizar aquellos productos de alta demanda, como las computadoras, laptops y sus accesorios, los celulares, y también los electrodomésticos. Orientando las campañas de marketing hacia ese mercado. Se podría ofrecer un servicio postventa de capacitación básica para aquellas personas que nop están tan relacionadas con la tecnología, ya que los clientes son en mayor medida adultos mayores, y esto podría ayudar a mejorar la retención de estos.

Aumentar la presencia de tiendas en lugares estratégicos de Europa también sería una buena idea, teniendo en cuanta que la cantidad es poca en comparación con Estados Unidos, lo que podría llevar a un descenso en los tiempos de entregas y aumento en las compras.