

Sprint 5 - Iniciación al análisis de datos con Power BI e indicadores

En este sprint, empezarás a aplicar tus conocimientos prácticos en Power BI utilizando la base de datos previamente utilizada, que contiene información sobre una empresa dedicada a la venta de productos online. Durante los ejercicios, es necesario que dediques esfuerzos a mejorar la legibilidad de las visualizaciones, asegurándote de seleccionar las representaciones visuales más adecuadas para presentar la información de forma clara y sencilla. No olvides agregar títulos descriptivos a tus gráficos para facilitar la comprensión de la información visualizada.

Nivel 1

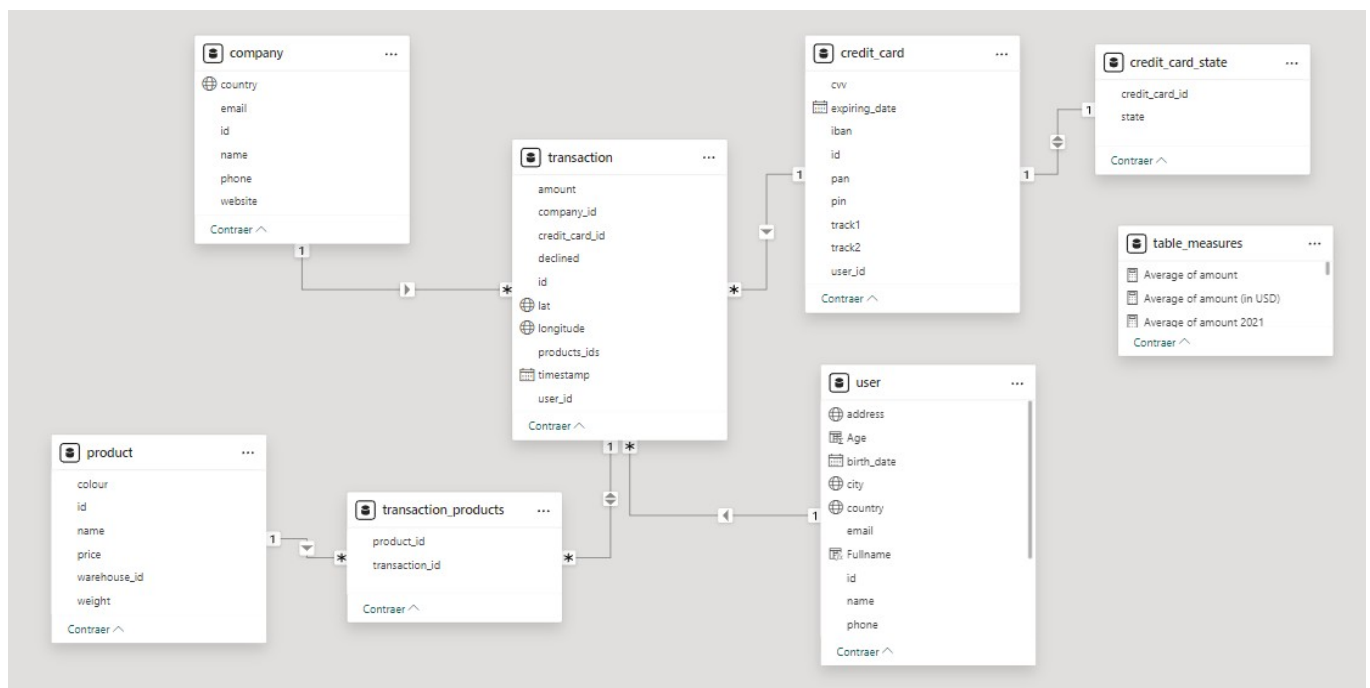
- Ejercicio 1:

Importa los datos de la base de datos utilizada previamente. Después de cargar los datos, muestra el modelo de la base de datos en Power BI.

Lo primero que hice fue extraer la base de datos desde **Power BI**, usando **Obtener datos** y escogiendo la opción **Base de datos MySQL**. Ahí ingresé el **Hostname**, el **Port**, y el nombre de mi base de datos.

Luego transformé los datos una vez tenía las tablas cargadas, empezando por cambiar los **data types**, luego removiendo los duplicados, después eliminando las columnas que hacen referencias a las relaciones porque después las crearé yo mismo, y por último renombrando las tablas para que no tengan el nombre de la base de datos como prefijo.

Lo siguiente fue crear las relaciones entre las distintas tablas que tengo.



- Ejercicio 2:

Tu empresa está interesada en evaluar la suma total del amount de las transacciones realizadas a lo largo de los años. Para ello, se ha solicitado la creación de un indicador clave de rendimiento (KPI). El KPI debe proporcionar una visualización clara del objetivo empresarial de conseguir una suma total de 25.000€ por cada año.

Lo primero que hice fue seleccionar la visualización **KPI** y, dentro de ella, colocar en el apartado **Valor** la medida **Sum of amount**, que creé para poder utilizarla libremente más tarde. En el **Eje de tendencia**, ingresé los valores **Year** de **timestamp** para poder mostrar los resultados filtrados por años. Además, en el apartado **Destino**, ingresé una medida creada que contenía la cifra objetivo que la empresa buscaba alcanzar. Finalmente, para filtrar por años, agregué una **Segmentación de datos** como visualización, que contenía los valores de **Year** en **timestamp**.

Sum of amount =

SUM('transaction'[amount])



Como podemos ver en las visualizaciones, el objetivo del año 2021 fue cumplido con creces, al superar en más de cuatro veces el objetivo empresarial. Por otro lado, en 2022 el objetivo no se cumplió por un margen menor al 1%, y, teniendo en cuenta que los datos disponibles llegan hasta el primer cuatrimestre del año, podríamos intuir que es cuestión de tiempo alcanzar dicho objetivo. Esto nos trae a colación la idea de si el objetivo de 25.000 euros debería ser aumentado o cambiado de anual a trimestral.

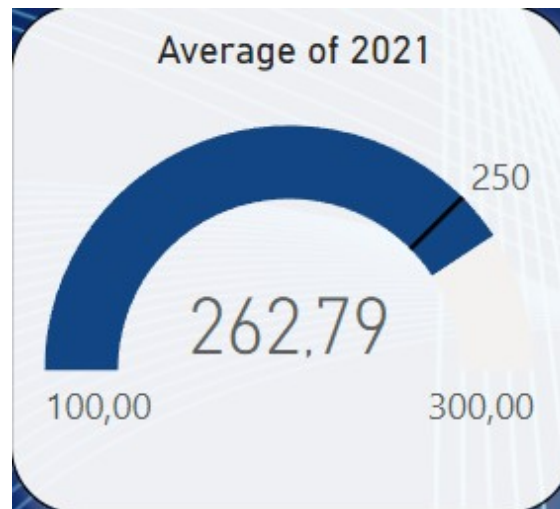
- **Ejercicio 3:**

Desde marketing te solicitan crear una nueva medida DAX que calcule la media de suma de las transacciones realizadas durante el año 2021. Visualiza esta media en un medidor que refleje las ventas realizadas, recuerda que la empresa tiene un objetivo de 250.

Para este ejercicio, seleccioné como visualización un **medidor**, donde ingresé en el apartado **Valor** una medida DAX creada que utiliza la función **AVERAGEX()** para calcular el valor de la columna **amount**, filtrándolo por el valor **Year = 2021** en la columna **timestamp**. Además, creé otra medida DAX que contiene el objetivo de 250, que es lo que la empresa aspira a alcanzar. Esta medida la ingresé en el apartado **Valor de destino**.

Average of amount 2021 =

```
AVERAGEX(  
    FILTER('transaction','transaction'[timestamp].[Year]=2021),  
    'transaction'[amount])
```



Como podemos apreciar en la visualización, el objetivo de un promedio de 250 fue alcanzado y superado exitosamente en el año 2021. Esto se refleja no solo en los valores, sino también en la intensidad del color azul de la visualización, que tiene un formato condicional en el que el valor mínimo se muestra en rojo y el objetivo en azul.

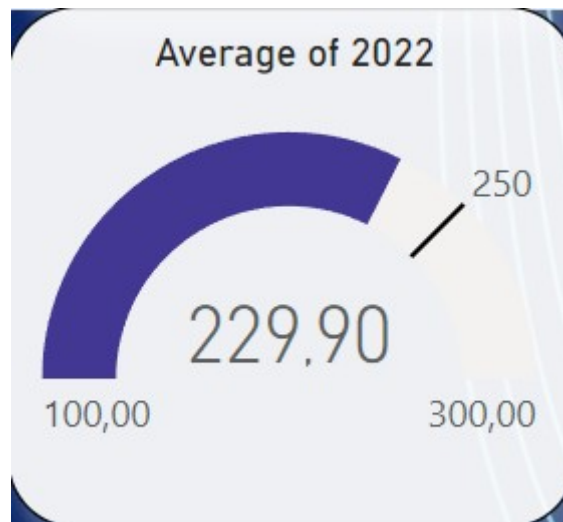
- **Ejercicio 4:**

Realiza el mismo procedimiento que realizaste en el ejercicio 3 para el año 2022.

Como este enunciado me pide hacer lo mismo que en el ejercicio anterior, para este ejercicio copié la misma fórmula DAX, pero cambiando el filtro de 2021 a 2022, y mantuve la medida DAX que contiene el objetivo de 250.

Average of amount 2022 =

```
AVERAGEX(  
    FILTER('transaction','transaction'[timestamp].[Year]=2022),  
    'transaction'[amount])
```



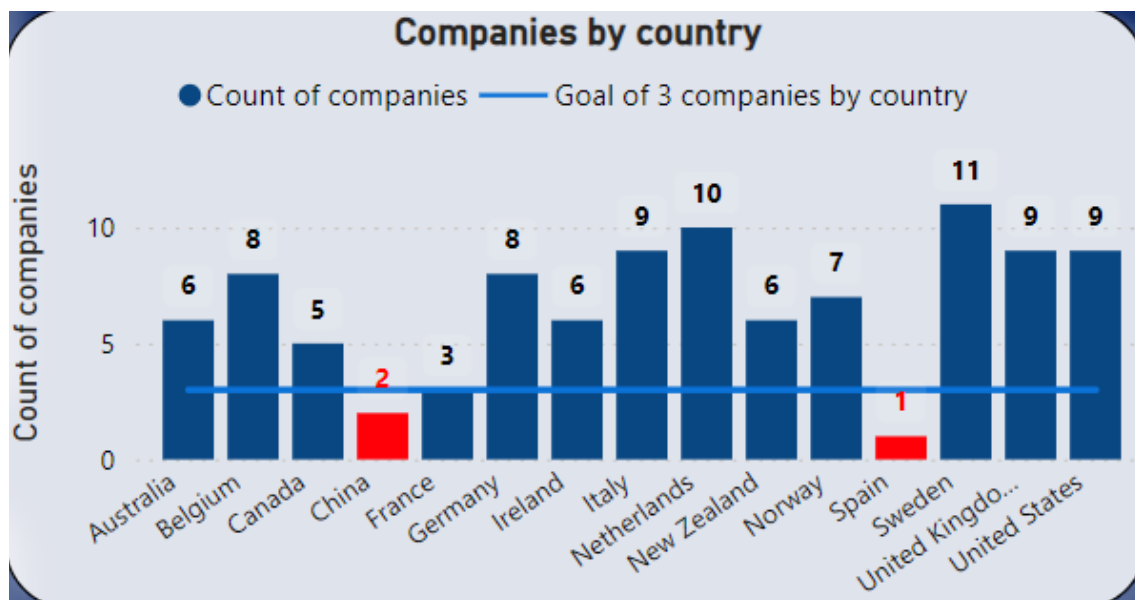
En esta visualización podemos observar que el objetivo de una media de 250 para el año 2022 no se está cumpliendo. Para alcanzar el objetivo de la media esperada, deberíamos implementar medidas que fomenten un mayor volumen de compras por parte de nuestros clientes durante los próximos tres trimestres de 2022.

- Ejercicio 5:

El objetivo de este ejercicio es crear una KPI que visualice la cantidad de empresas por país que participan en las transacciones. La meta empresarial es garantizar que haya al menos 3 empresas participantes por país. Para ello, será necesario utilizar DAX para calcular y representar esta información de manera clara y concisa.

Para esta visualización, seleccioné un **gráfico de columnas agrupadas y de líneas** con el objetivo de representar, por un lado, el conteo de empresas agrupadas por su país de origen, y por otro, mostrar gráficamente el objetivo empresarial de al menos 3 compañías por país. En el **Eje X**, coloqué la columna **country** de la tabla **company**; en el **Eje Y de columnas**, agregué una medida DAX creada que realiza un **DISTINCTCOUNT** de la columna **company_id** de la tabla **transaction**; y en el **Eje Y de líneas**, utilicé una fórmula DAX que contiene el objetivo empresarial de 3 compañías por país.

Count of distinct companies IDs from transaction =
DISTINCTCOUNT('transaction'[company_id])



Esta visualización no está filtrada por año, ya que no lo consideré necesario, dado que el objetivo es conocer la cantidad de compañías presentes en cada país. En la visualización, podemos observar claramente que hay dos países que no alcanzan el objetivo empresarial propuesto: España y China. También podríamos considerar a Francia, que cumple apenas con el objetivo.

Para fomentar la creación de empresas en España y China, sería recomendable implementar políticas que faciliten el acceso a financiación para emprendedores, así como crear programas de apoyo a pequeñas y medianas empresas (PYMEs). Además, sería necesario investigar si los trámites burocráticos y las cargas fiscales representan cuellos de botella que dificulten el aumento de la tasa de emprendimiento en estos países.

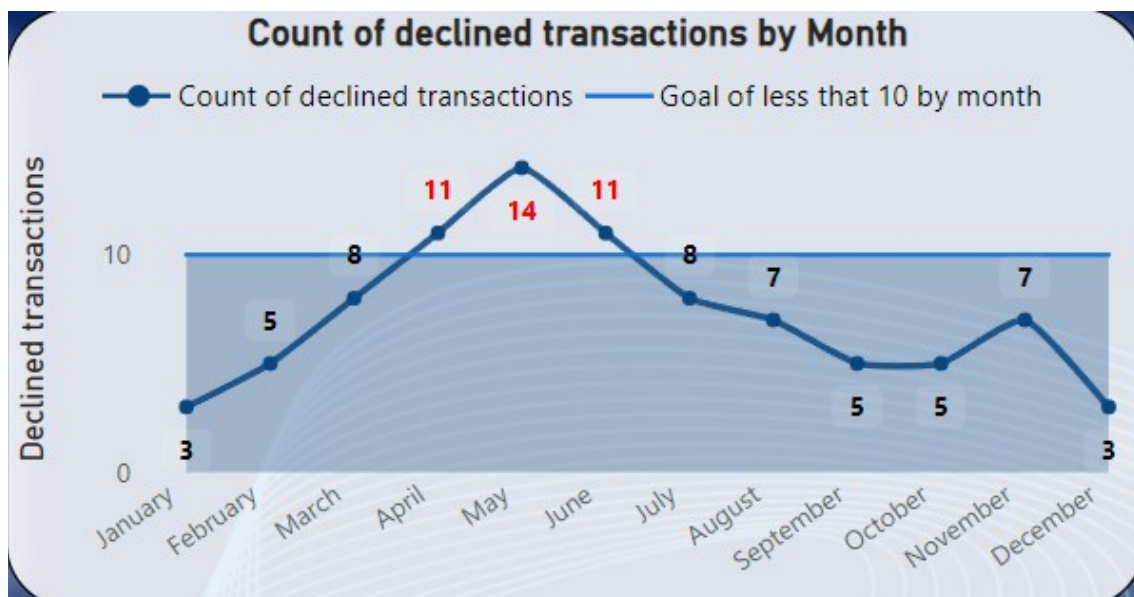
- Ejercicio 6:

Crea una nueva KPI que permita visualizar la cantidad de transacciones declinadas a lo largo del tiempo. La empresa estableció un objetivo de tener menos de 10 transacciones declinadas por mes.

Esta visualización será un **gráfico de líneas** que me permitirá observar el conteo de transacciones declinadas, agrupadas por mes, y su evolución a lo largo del tiempo. Para representar la visualización, he colocado en el **Eje X** los valores correspondientes a los meses de la columna **timestamp**. En el **Eje Y**, agregué una medida DAX que cuenta las filas filtradas por la condición de que el valor de la columna **declined** sea **TRUE()**. En el **Eje Y secundario**, utilicé una fórmula DAX que representa el objetivo empresarial de no más de 10 transacciones declinadas por mes.

Count of declined transactions =

```
COUNTROWS(  
    FILTER('transaction','transaction'[declined]=TRUE())  
)
```



Este gráfico se ve afectado por el filtro de años aplicado en el dashboard, ya que eso nos permite observar de manera precisa la evolución de las transacciones declinadas a lo largo de los meses en sus respectivos años. A futuro, sería importante tener en cuenta el total de transacciones registradas, ya que esto nos proporcionaría un mayor contexto para evaluar si se están haciendo las cosas correctamente. No es lo mismo tener 10 transacciones declinadas de un total de 50 que tener las mismas 10 de un total de 1.000.

Por otro lado, en el gráfico podemos observar una tendencia a la baja. Esto puede deberse a que no tenemos suficientes registros en nuestra base de datos de los últimos tres trimestres de 2022, lo que da como resultado un gráfico alentador pero sesgado. Otra posible explicación es que las medidas implementadas en la empresa han tenido un impacto positivo, reduciendo considerablemente el número de transacciones declinadas, especialmente después de los meses de abril, mayo y junio, que fueron los que presentaron un mayor conteo de transacciones declinadas y que superaron negativamente el objetivo empresarial.

- Ejercicio 7:

Crea un gráfico de columnas agrupadas que refleje la sumatoria de las ventas por mes. El objetivo de la empresa es tener por lo menos 10.000 transacciones por mes.

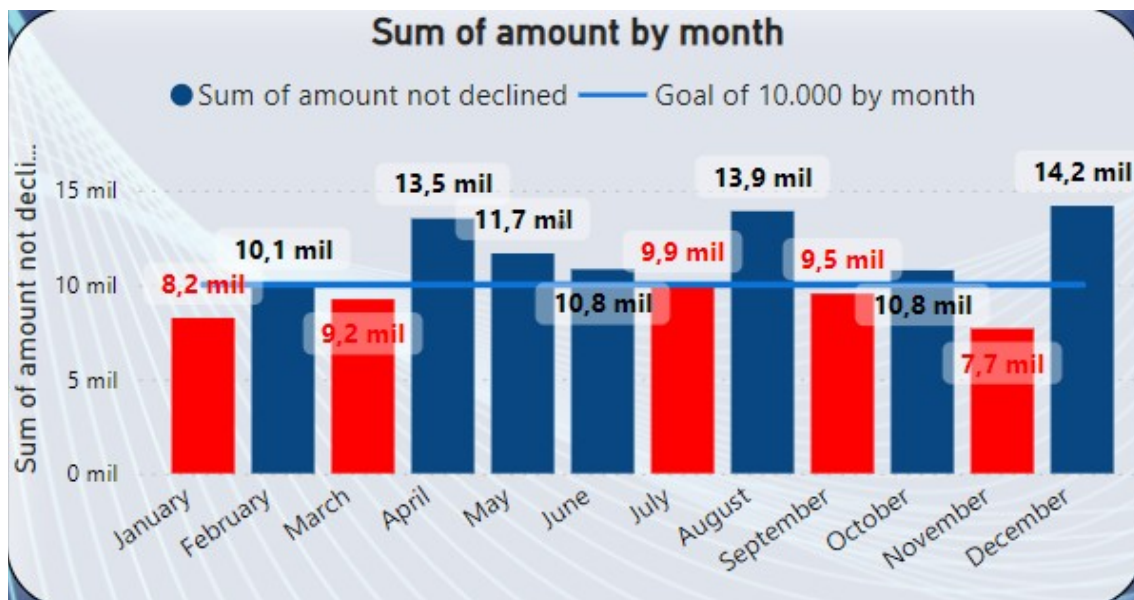
En esta ocasión, tal como sugiere el enunciado, utilicé un **gráfico de columnas agrupadas y líneas**, empleando la línea como representación gráfica del objetivo empresarial.

En el **Eje X** se encuentran los valores correspondientes a los meses de la columna **timestamp**.

En el **Eje Y de columnas**, agregué una fórmula DAX que calcula la sumatoria de la columna **amount**, filtrándola por la condición de que el valor de **declined** sea igual a **FALSE()**. En el **Eje Y de línea**, se encuentra la medida DAX que representa el objetivo empresarial.

Sum of amount not declined =

```
CALCULATE(  
    SUM('transaction'[amount]),  
    'transaction'[declined] = FALSE()  
)
```



Este gráfico se ve afectado por el filtro de años aplicado en el dashboard. Además, es importante destacar que, al realizar la sumatoria, solo se han contabilizado las transacciones no declinadas.

Como podemos observar, en cuatro meses no se ha logrado cumplir con el objetivo empresarial con total certeza: julio, septiembre, noviembre y enero. También debemos mencionar que en marzo, tanto en 2021 como en 2022, no se alcanzó el objetivo. En 2021, esto se debe a que los registros comienzan a partir del 17 de marzo, mientras que en 2022 los registros solo se extienden hasta el 16 de marzo.

Por otro lado, sería útil profundizar en los factores que explican los resultados de los meses de julio, septiembre, noviembre y enero, ya que en los meses posteriores a estos se observa un repunte en las ventas alcanzadas.

- Ejercicio 8:

En este ejercicio, se quiere profundizar en las transacciones realizadas por cada usuario y presentar la información de forma clara y comprensible.

En una tabla, presenta la siguiente información...

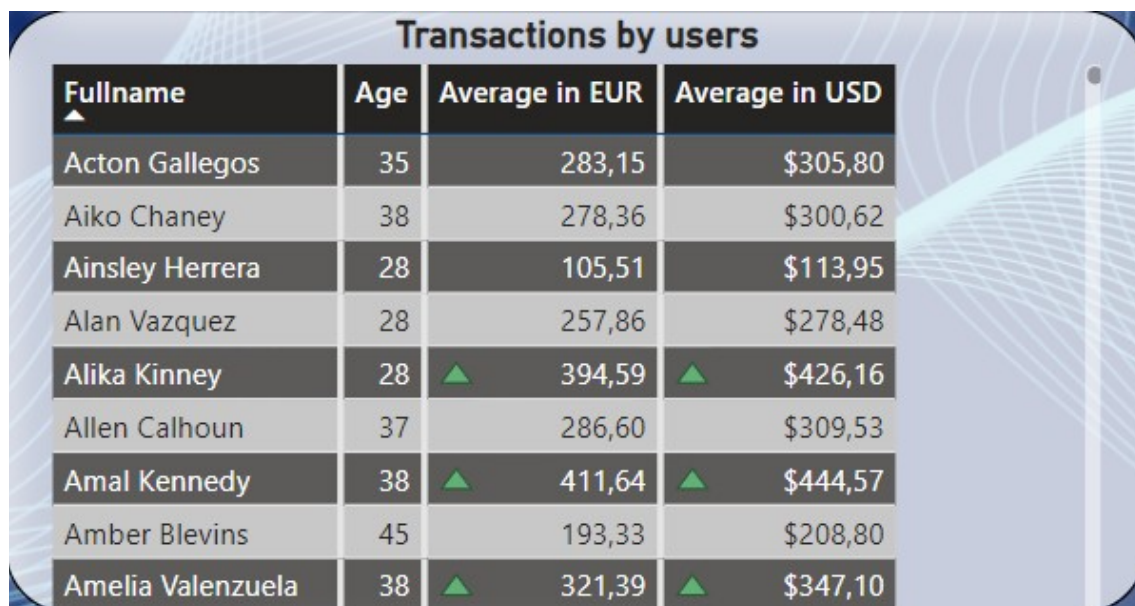
Para mostrar los nombres completos de los usuarios en esta **tabla**, creé una columna calculada llamada **Fullname**, que concatena los datos de las columnas **name** y **surname**, asegurándome de dejar un espacio entre ambos valores.

Luego, añadí la columna **Age**, que utiliza la función **DATEDIFF()** para calcular la diferencia entre las fechas de la columna **birth_date** y la fecha actual, proporcionada por la función **TODAY()**. El resultado se muestra en años, ya que se utiliza **YEAR** como unidad de medida.

La columna **Average in EUR** se obtiene mediante una medida DAX simple que calcula el promedio de los valores en la columna **amount** de la tabla **transaction**.

La columna **Average in USD** se basa en la medida anterior. Para esta nueva medida, multiplicamos el resultado de la columna **Average in EUR** por 1.08, considerando la tasa de conversión.

Finalmente, para identificar a los usuarios con promedios superiores a 300 euros o 320 dólares, formateé las celdas correspondientes, añadiendo un ícono de flecha verde hacia arriba a los valores que superan esos umbrales.



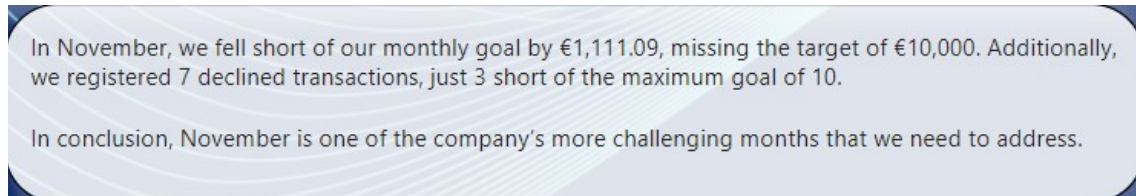
Fullname	Age	Average in EUR	Average in USD
Acton Gallegos	35	283,15	\$305,80
Aiko Chaney	38	278,36	\$300,62
Ainsley Herrera	28	105,51	\$113,95
Alan Vazquez	28	257,86	\$278,48
Alike Kinney	28	▲ 394,59	▲ \$426,16
Allen Calhoun	37	286,60	\$309,53
Amal Kennedy	38	▲ 411,64	▲ \$444,57
Amber Blevins	45	193,33	\$208,80
Amelia Valenzuela	38	▲ 321,39	▲ \$347,10

La tabla creada proporciona una visión detallada de los usuarios, facilitando un acceso rápido a información relevante que permita segmentarlos de manera más eficiente y personalizada. Al tener la información de los usuarios organizada por edad y sus promedios de compra, podemos identificar patrones de comportamiento y agruparlos en segmentos. Este análisis puede ser crucial para diseñar estrategias de marketing más efectivas, orientadas a grupos de usuarios con comportamientos de compra similares.

- Ejercicio 9:

Redacta un párrafo breve, máximo 50 palabras, explicando el significado de las cifras presentadas en las visualizaciones de Power BI. Puedes interpretar los datos por lo general o centrarte en algún país específico. Acompaña las interpretaciones realizadas con la captura de pantalla de las visualizaciones que vas a analizar.

Para esta ocasión utilicé una ventana de narrativa para escribir un pequeño análisis del mes de noviembre.



Como podemos ver en la visualización el mes de noviembre es un mes al que necesitamos prestarle más atención ya que en distintas métricas no se están alcanzando los objetivos propuestos.

Nivel 2

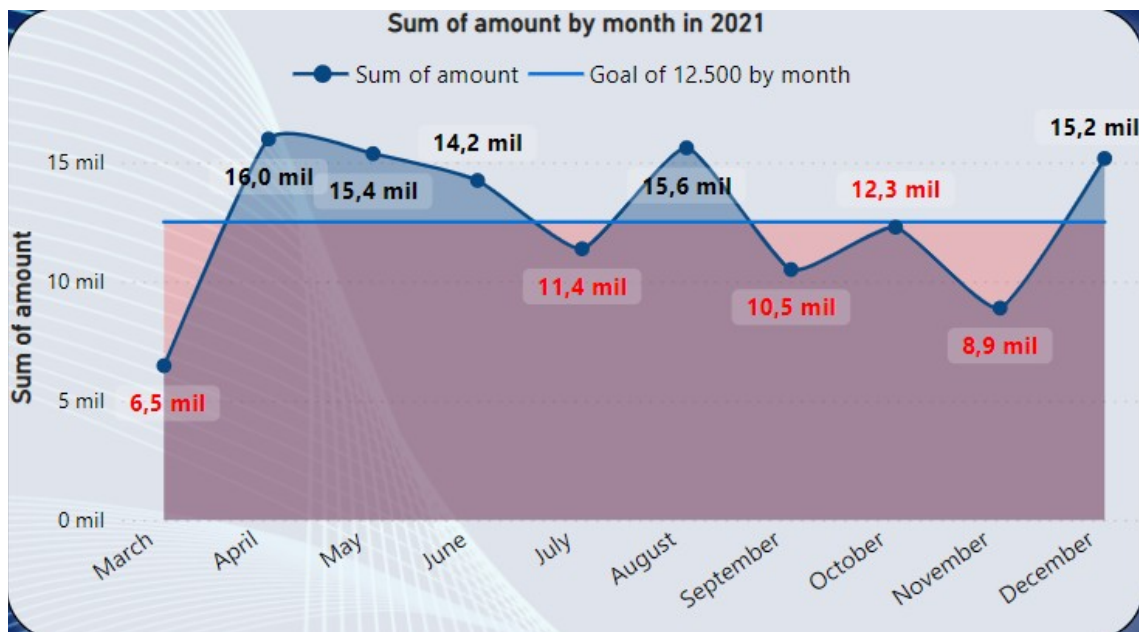
- Ejercicio 1:

Desde el área de marketing necesitan examinar la tendencia mensual de las transacciones realizadas en 2021, específicamente, quieren conocer la variación de las transacciones en función del mes. Recuerda visualizar la meta empresarial de alcanzar al menos 12.500€ en transacciones por mes. En este ejercicio, será necesario que se consiga identificar los meses en los que no se logró la meta establecida. Si es necesario puedes realizar dos visualizaciones.

En esta visualización utilicé nuevamente un **gráfico de líneas**, donde en el **Eje X** añadí los valores correspondientes a los meses de la columna **timestamp**. En el **Eje Y**, incluí una métrica DAX que representa la sumatoria de la columna **amount**, filtrada por el año 2021. Además, en el **Eje Y secundario**, agregué una métrica DAX que muestra el objetivo empresarial propuesto.

Sum of amount 2021 =

```
CALCULATE(  
    SUM('transaction'[amount]),  
    YEAR('transaction'[timestamp]) = 2021)
```



En esta visualización podemos observar que el objetivo empresarial no se ha alcanzado en cinco meses a lo largo del año, específicamente en marzo, julio, septiembre, octubre y noviembre. Esto resulta preocupante, ya que, aunque en los meses en los que se supera el objetivo se alcanza una cifra positiva, el margen no es lo suficientemente grande como para compensar los meses con déficit. Es necesario profundizar en el análisis de los dos últimos trimestres de 2021 para entender las razones detrás de la gran diferencia con respecto al objetivo empresarial.

- Ejercicio 2:

En tu trabajo, se quiere profundizar en la comprensión de las transacciones realizadas en Alemania. Por tanto, te solicitan que desarrolles medidas DAX para crear visualizaciones que destaquen la media de ventas en Alemania. Ten presente que la empresa tiene como objetivo conseguir una cifra de 250 euros anuales. Configura la visualización de forma que el valor mínimo sea 100 y el máximo 350, brindando así una representación más efectiva de la información.

Para este ejercicio, reutilicé un **medidor** como visualización, en el cual añadí, en el apartado **Valor**, una fórmula DAX que calcula el promedio de la columna **amount**, filtrada por el país Alemania y el año 2021. Además, reutilicé la misma métrica del objetivo empresarial utilizada en los medidores anteriores, ya que el objetivo es el mismo.

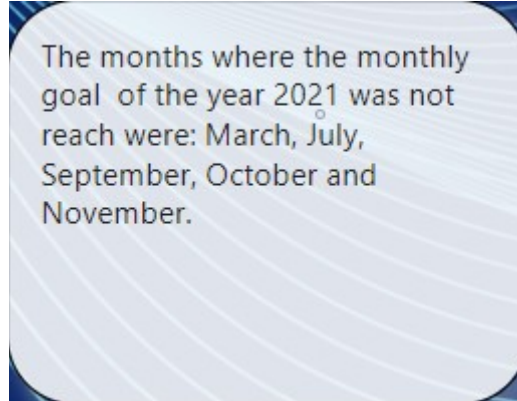


Como podemos observar en la visualización, el objetivo empresarial se está cumpliendo, pero con un margen bastante reducido. Por un lado, es reconfortante saber que se ha alcanzado el objetivo propuesto; sin embargo, esto también indica que hay trabajo por hacer en futuras campañas en Alemania, ya que estuvimos muy cerca de no cumplir con dicho objetivo empresarial.

- **Ejercicio 3:**

Escribe un breve párrafo, máximo de 25 palabras, indica en qué mes no se llegó a cumplir con el objetivo propuesto del ejercicio 1.

Nuevamente utilicé una visualización de tipo narrativa para este ejercicio.



Como se menciona en la primera visualización, los meses en los que no se cumplió con el objetivo empresarial fueron marzo, julio, septiembre y noviembre. Si consideramos que en 2021 no se registraron transacciones para los meses de enero y febrero, el hecho de que no se haya alcanzado el objetivo en 5 de los 10 meses restantes del año cobra mayor relevancia. Esto nos alerta de la necesidad de implementar medidas para evitar que esta situación se repita en los próximos años.

Nivel 3

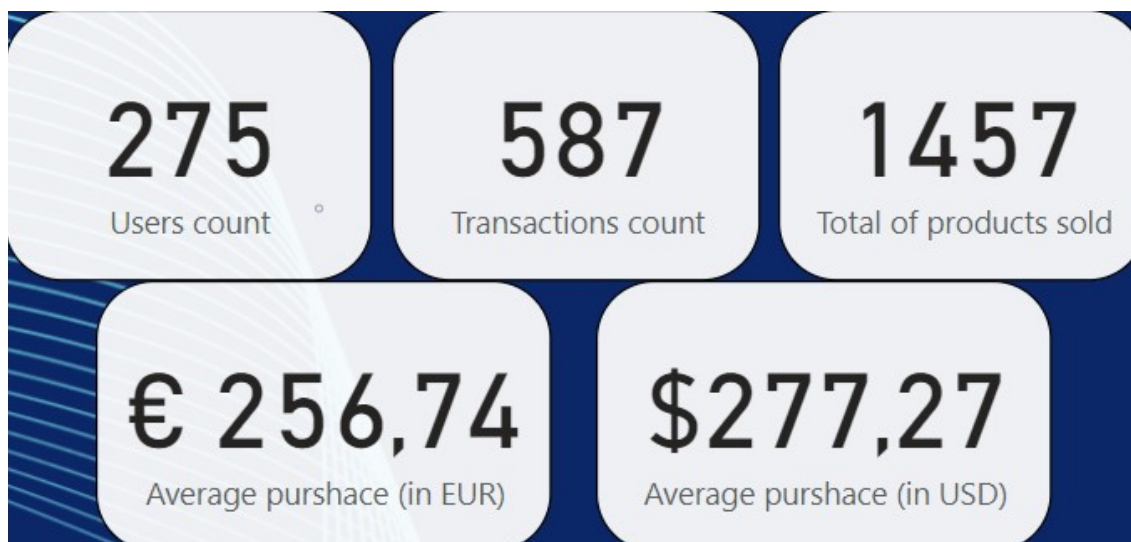
- Ejercicio 1:

La sección de marketing quiere profundizar en las transacciones realizadas por los usuarios y usuarias. En consecuencia, se te solicita la elaboración de diversas visualizaciones que incluyan:

- Las medidas estadísticas clave de las variables que consideres relevantes para comprender las transacciones realizadas por los usuarios/as.
- Cantidad de productos comprados por cada usuario/a.
- Media de compras realizadas por usuario/a, visualiza qué usuarios/as tienen una media de compras superior a 150 y cuáles no.
- Muestra el precio del producto más caro comprado por cada usuario/a.
- Visualiza la distribución geográfica de los usuarios/as.

-Las medidas estadísticas clave:

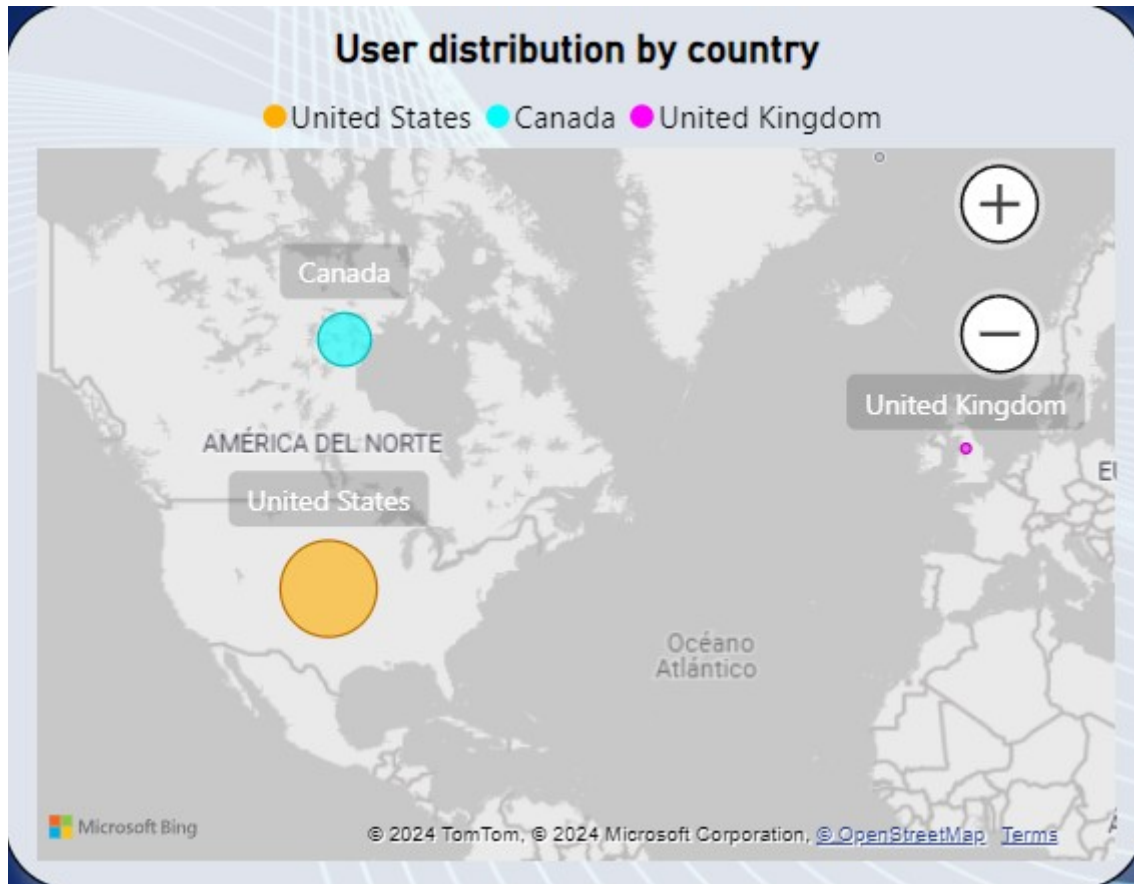
En esa ocasión, lo primero que hice fue crear algunas tarjetas gráficas que presentaran información relevante sobre las transacciones realizadas por los usuarios.



Estas tarjetas gráficas son interactivas, lo que permite filtrar sus valores generales por los tres países del mapa con la distribución geográfica, para obtener un perfil más específico de los usuarios en cada país.

-Visualización de la distribución geográfica de los usuarios:

Para mostrar el mapa con la distribución de los usuarios, lo primero que hice fue crear una jerarquía utilizando las columnas **country** y **city**. Añadí esta jerarquía al campo **Ubicación**, y en el campo **Leyenda** coloqué únicamente la parte correspondiente a **country**, para agrupar los resultados por país. Finalmente, en el campo **Tamaño de burbuja**, agregué el recuento de la columna **id** de la tabla **user**.



El mapa muestra de manera clara que el país con el mayor número de usuarios es Estados Unidos, seguido de Canadá, mientras que el que tiene menos usuarios es Reino Unido. Al interactuar con el mapa, podremos ver la distribución de los usuarios dentro de cada país. Por otro lado, al interactuar con la visualización observamos que el país con la media de compras más alta es Canadá, lo que sugiere que podríamos diseñar campañas de marketing dirigidas específicamente a este mercado, que tiene un mayor poder adquisitivo. Esta campaña también podría extrapolarse a Estados Unidos, ya que, además de contar con un poder adquisitivo similar cuenta a su vez con el mayor volumen de usuarios, y tienen una idiosincrasia similar a la de su vecino Canadá.

- Cantidad de productos comprados por cada usuario.
- Media de compras realizadas por usuario, visualizando quienes tienen una media de compras superior a 150 y cuáles no.
- Muestra el precio del producto más caro comprado por cada usuario.

Primero, creé la visualización de una tabla en la que agregué la columna **id** de la tabla **user**. A continuación, incluí las demás columnas solicitadas en el ejercicio, como la cantidad de productos comprados por cada usuario, el promedio de compras, el precio del producto más caro adquirido por cada usuario, y finalmente, el país y la ciudad de procedencia de cada uno.

Para darnos el precio del producto más caro comprado por cada usuario la fórmula hace lo siguiente:

1ro: Crea una variable **MaxPrice** donde guardará el precio del producto con el valor máximo que cada usuario haya comprado.

2do: Usa la función **CALCULATE()** para realizar ese cálculo.

3ro: Va a la tabla **transaction_products** y con la función **FILTER()** hace una filtración de los productos que corresponden a cada usuario, gracias a la función **RELATED()** que primero por un lado relaciona los **user_id** con sus respectivas transacciones, y por otro lado la parte correspondiente a **MAX('transaction'[user_id])** se asegura que los resultado se filtran únicamente por el usuario que estamos analizando actualmente.

4to: la función **MAXX()** busca el producto con el valor máximo comprado por cada usuario, y para realizar eso utiliza la función **RELATED()** que le permite comparar lo precios de los productos que se encuentran en la columna **price** de la tabla **product**.

5to: Finalmente **RETURN** nos devuelve el precio del producto guardado en la variable **MaxPrice** que el cálculo encontró.

Most expensive product purchased by user =

VAR MaxPrice =

```

CALCULATE(
    MAXX(
        FILTER(
            transaction_products,
            RELATED('transaction'[user_id]) = MAX('transaction'[user_id])),
        RELATED('product'[price]))

```

RETURN MaxPrice

Users ID	Products count	Purchase average	Highest price	Country	City
60	4 ▲	€ 156,23	€ 195,94	United States	Aurora
61	3 ▲	€ 460,82	€ 114,09	United States	Hilo
62	1 ▬	€ 133,39	€ 167,20	United States	Newark
63	1 ▬	€ 82,43	€ 26,51	United States	Pittsburgh
64	3 ▲	€ 471,47	€ 167,20	United States	Norman
65	1 ▲	€ 268,48	€ 63,33	United States	Salem
66	4 ▲	€ 381,17	€ 167,20	United States	Colchester
67	2 ▬	€ 114,58	€ 91,89	United States	Saint Louis
68	1 ▲	€ 157,20	€ 167,20	United States	Bloomington
69	2 ▲	€ 223,53	€ 141,01	United States	Columbus
70	1 ▲	€ 389,70	€ 49,70	United States	San Jose
71	2 ▲	€ 159,14	€ 141,01	United States	Mesa
72	4 ▲	€ 330,90	€ 195,94	United States	Spokane
73	1 ▲	€ 317,04	€ 139,59	United States	Boise
74	1 ▲	€ 481,75	€ 59,80	United States	Wichita
75	2 ▬	€ 21,97	€ 141,01	United States	Lewiston
76	3 ▲	€ 248,11	€ 139,59	United States	Newark
77	2 ▬	€ 33,10	€ 172,78	United States	Norfolk
78	3 ▲	€ 424,18	€ 171,22	United States	Aurora
79	3 ▲	€ 165,38	€ 195,94	United States	Savannah
80	2 ▬	€ 20,44	€ 132,86	United States	Paradise
81	3 ▲	€ 283,15	€ 161,11	United States	Lexington
82	1 ▲	€ 378,86	€ 171,22	United States	South Portland
83	4 ▲	€ 330,07	€ 169,96	United States	Hillsboro
84	2 ▲	€ 229,65	€ 132,86	United States	Dover
85	3 ▲	€ 306,94	€ 91,89	United States	Fairbanks
86	3 ▲	€ 390,67	€ 172,78	United States	Anchorage

Con esta tabla, obtendremos una visión detallada de las transacciones realizadas por cada usuario, lo que nos permitirá realizar un análisis exhaustivo de su comportamiento de compra. Esto facilitará la segmentación y el análisis de patrones, proporcionando información valiosa para identificar a los usuarios con mayor poder adquisitivo, entender sus preferencias y diseñar estrategias personalizadas, como campañas de marketing dirigidas o ajustes en los inventarios, con el fin de maximizar las ventas y mejorar la experiencia del cliente.