

SPRINT S5.01 – INICIACIÓN AL ANÁLISIS DE DATOS CON POWER BI E IDICADORES

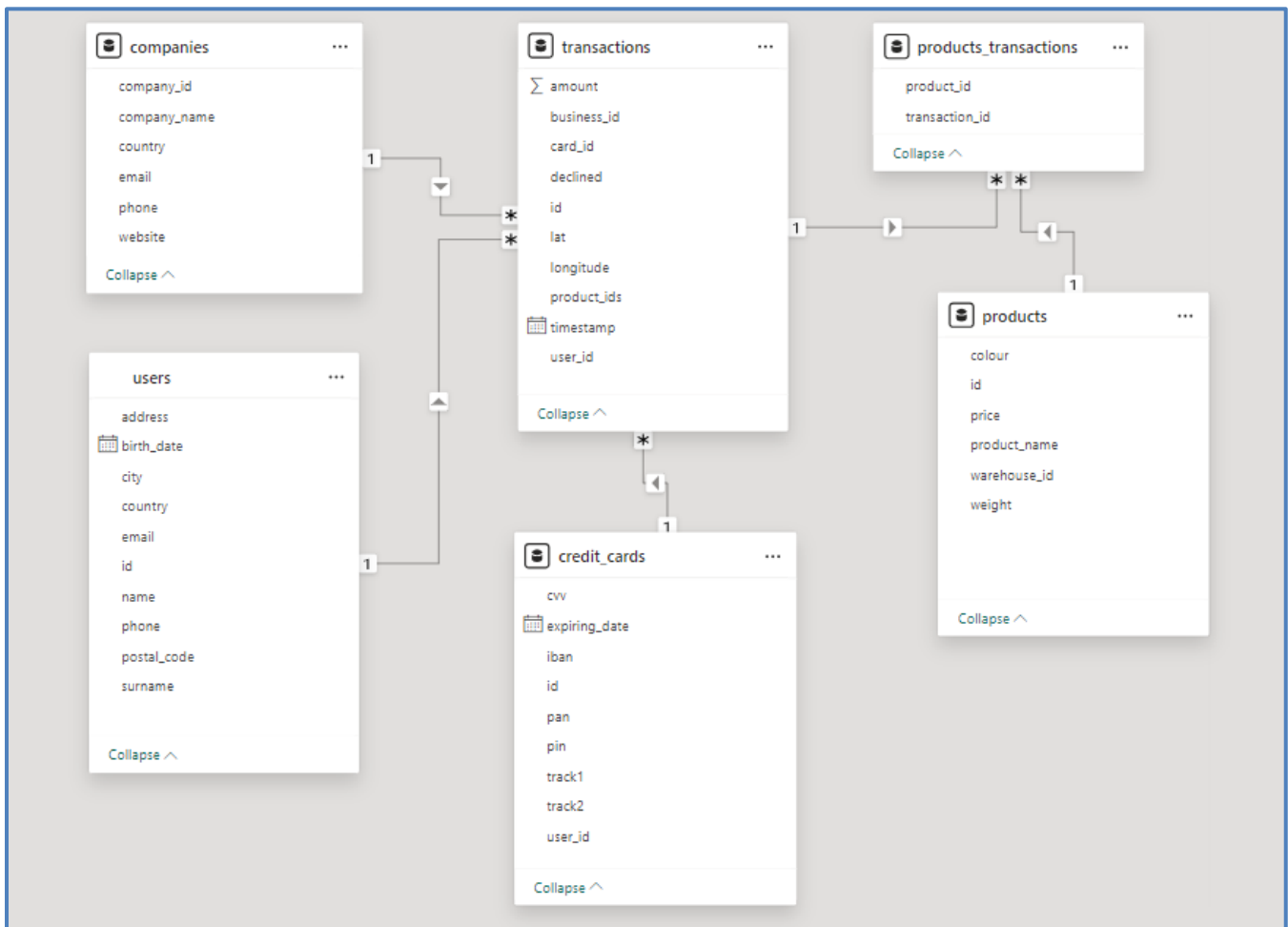
NIVEL 1

Ejercicio 1.1

Importar 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.

Solución:

El modelo resultante sería el siguiente:



Como se aprecia, es un modelo de estrella donde la tabla **transactions** vendría a ser la **tabla de hechos**, mientras que las otras tablas serían las **tablas de dimensiones**. Cabe destacar la **tabla intermedia** denominada **products_transactions**, que fue creada para solventar la relación de muchos a muchos entre **products** y **transactions**.

Para llegar a tal resultado se ha realizado el proceso de limpieza y transformación de datos los cuales se detallan en las siguientes líneas.

❖ **Tabla de hechos - Transactions:**

Como tabla de hechos es la que contiene las foreign key que relacionan con el resto de tablas de dimensiones. Se describen sus campos a continuación:

Amount : contiene el monto de las transacciones.

business_id : representa el identificador de las compañías. Es foreign key y se relaciona con el campo **company_id** de la tabla **companies**.

card_id: identificador de las tarjetas de crédito. Es foreign key y se relaciona con el campo **id** de la tabla **credit_cards**.

declined: Muestra si la relación ha sido declinada (valor = 1) o aceptada (valor = 0).

id: identifica la transacción. Vendría a ser la primary key y al mismo tiempo relaciona la tabla **transactions** con la tabla intermedia **products_transactions**.

lat y longitude: estas columnas conjuntas muestran las coordenadas o localización de la transacción.

products_ids: muestra los productos comprados en la transacción.

timestamp: muestra el momento preciso en que la transacción se ejecutó.

user_id: identifica al usuario. Es foreign key y se relaciona con el campo **id** de la tabla **users**.

❖ **Tablas de dimensiones: companies, users, credit_cards, products**

Companies : La **primary key** de esta tabla sería el campo **company_id** y que a su vez se relaciona con la **foreign key** denominada **business_id** de la tabla **transactions** con una cardinalidad de uno a muchos. El campo **company_name** describe el nombre de la compañía. **Country** indica el país de procedencia de la compañía. Los campos **email**, **phone** y **website** muestran las vías a través de las cuales se pueden contactar a la empresa, como email, teléfono y página web.

Users : Esta tabla comprende la union de los registros de los archivos **users_us**, **users_uk** y **users_ca**, como también se había realizado en el sprint anterior. La **primary key** vendría a ser el campo **id**, que a su vez se relaciona de uno a muchos con la **foreign key** denominada **user_id** de la tabla **transactions**. Los campos **name** y **surname** muestran el nombre y apellido, respectivamente. **Birth_date** indica la fecha de nacimiento del usuario. Los campos **address**, **postal_code**, **city** y **country** indican conjuntamente la localización del usuario. **Email** e **phone** muestran las vías de contactar al cliente.

Credit_cards : La **primary key** de esta tabla sería el campo **id**, que a su vez se relaciona de uno a muchos con la **foreign key** denominada **card_id** de la tabla **transactions**. **Iban** indica el número de cuenta asociado a la tarjeta. **Expiring_date** indica la fecha de expiración de la tarjeta. El código **CVV** muestra el código necesario para concretar operaciones cuando se efectúan en línea.

Products : La **primary key** de esta tabla sería el campo **id**, que a su vez se relacionaría de muchos a muchos con el campo **products_ids** de la table **transactions**. Sin embargo, se ha creado una tabla intermedia para romper la cardinalidad de muchos a muchos y así tener una estructura más eficiente y evitar complicaciones en la consulta de datos. El campo **price** indica el precio cada producto. **Product_name** indica el nombre del producto, mientras que **colour** y **weight** muestran los atributos de cada producto.

Además de estas tablas, como se mencionó anteriormente, fue necesaria la creación de la tabla intermedia **products_transactions**. Esta se relaciona con la tabla **transactions** a través del campo **transaction_id** y se la relaciona con la tabla **products** a través del campo **product_id**.

Ejercicio 1.2

La Empresa está interesada en evaluar la suma total de la cantidad de 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.

Solución:

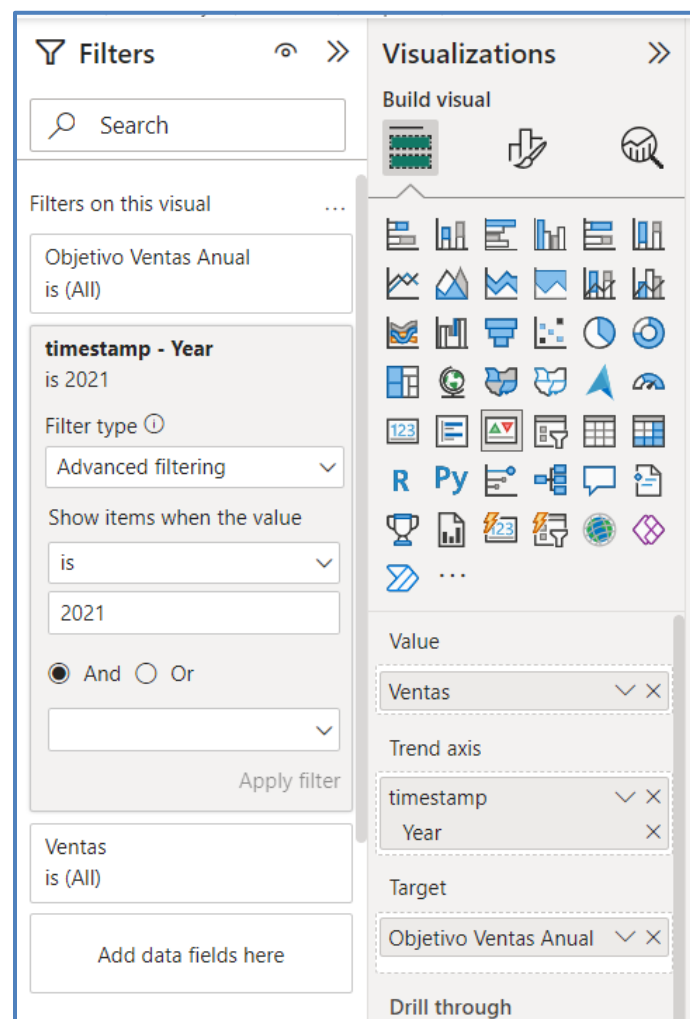
Para la generación del siguiente KPI se han creado **medidas**. La primera es **Ventas** , que viene a representar la suma del monto de las ventas de la tabla **transactions**. La segunda medida es **Objetivo de Ventas Anual** con un valor de 25.000 € , que es el objetivo empresarial que se ha trazado. Las medidas se crearon a través de las siguientes fórmulas, respectivamente:

Ventas = SUM(transactions[amount])



Objetivo Ventas Anual = 25000

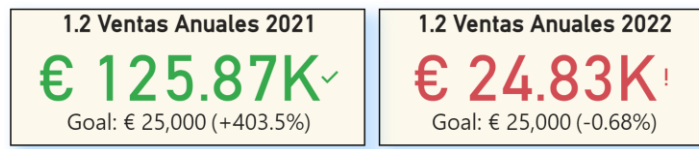
Luego de ello, para la creación del **KPI** en sí, se utilizó la medida **Ventas** como **value** (valor); el **timestamp** de la tabla **transactions** como **eje de tendencia** (trend axis) y la medida **Objetivo Ventas Anual** como **Objetivo** (Target). Además, en el apartado de filtros se filtró el resultado solo para el año 2021 y luego solo para el año 2022.



Interpretación:

Para el año 2021 se tiene que la suma de ventas ha superado con creces el objetivo planteado. En términos porcentuales, las ventas del 2021 han superado en un 403.5% el objetivo.

Para el año 2022 aún no se ha alcanzado el objetivo (datos hasta marzo), no obstante si extrapolamos valores es altamente probable que a finales del 2022 se supere el objetivo establecido al igual que en el año 2021.



Ejercicio 1.3

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.

Solucion:

Para el siguiente KPI se han creado las siguientes medidas:



```
Media Ventas 2021 = CALCULATE(AVERAGE('transactions'[amount]), YEAR('transactions'[timestamp])=2021)
```

Representa el promedio del monto de las ventas considerando únicamente las efectuadas en el año 2021.



```
Máximo Venta Media = 400
```

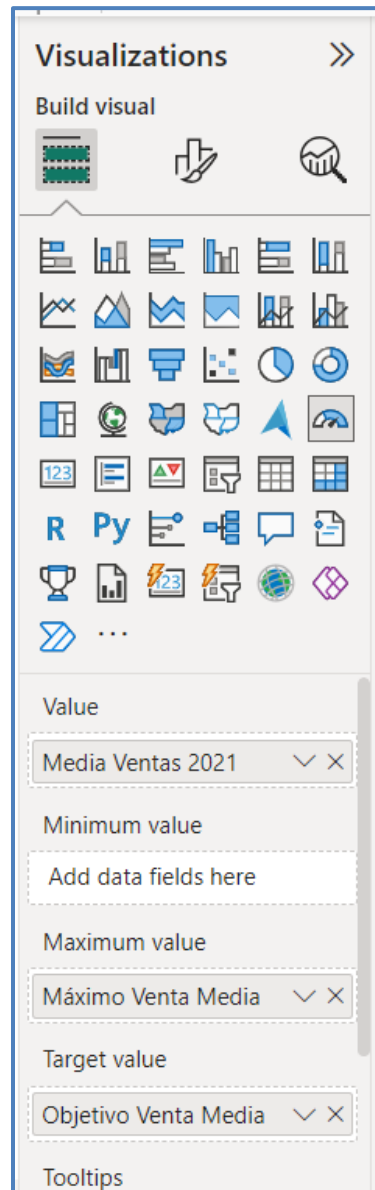
El valor máximo que mostrará el medidor. Es decir el medidor tiene un rango que va desde 0 € hasta 400 € para una mejor visualización.

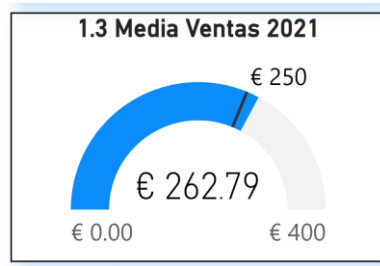


```
Objetivo Venta Media = 250
```

Se crea una medida como monto objetivo de media de ventas. Este valor de 250 € será el **target value** dentro del KPI.

Luego de ello, para la creación del **KPI** se utilizó la medida **Media Ventas 2021** como **valor** (value); la medida **Máximo Venta Media** como **Máximo Valor** (Maximum Value) y la medida **Objetivo Venta Media** como **Valor Objetivo** (Target Value).



Interpretación:

Para el año 2021 podemos apreciar que el valor promedio de ventas fue de 262.79 €; es decir se superó ligeramente el valor objetivo de 250 €. Lo que es lo mismo, podemos decir en términos porcentuales que se superó el objetivo en alrededor del 5% ($262.7 \text{ €} / 250 \text{ €}$).

Ejercicio 1.4

Realiza el mismo procedimiento realizado en el ejercicio 3, esta vez para el año 2022.

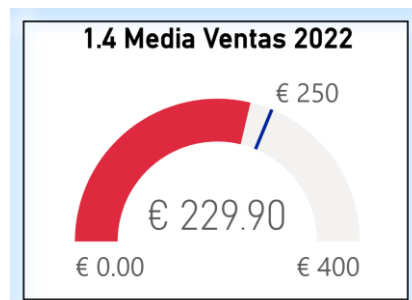
Solución:

Para el siguiente **KPI** se han utilizado las mismas medidas creadas anteriormente: Máximo Venta Media y Objetivo Venta Media. Solo se ha creado una nueva medida:



```
Media Ventas 2022 = CALCULATE(AVERAGE('transactions'[amount]), YEAR('transactions'[timestamp])=2022)
```

Representa el promedio del monto de las ventas considerando únicamente las efectuadas en el año 2022 (data hasta marzo).

Interpretación:

Para la data con la que se cuenta del 2022 podemos apreciar que el valor promedio de ventas fue de 229.90€; es decir, no se alcanzó el valor medio objetivo de ventas, estipulado en 250 €. En términos porcentuale podemos afirmar que el promedio de ventas para el año 2022 está por debajo del objetivo en 8% $[(229.9 - 250) / 250]$

Dado que se ha condicionado el color con el que se colorea el medidor, en este caso se muestra en rojo puesto que no se ha alcanzado el objetivo. A diferencia del caso anterior, en donde se coloreó de azul al superar el objetivo de 250 €.

Ejercicio 1.5

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.

Solución:

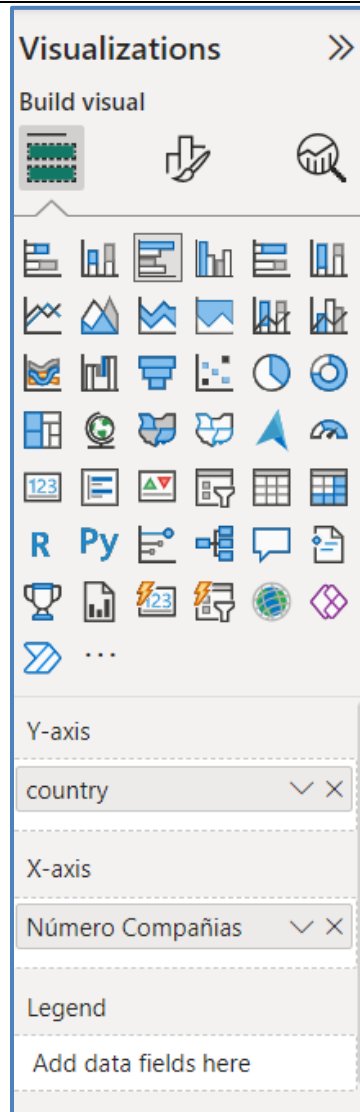
Hemos creado la siguiente medida:

```
Número Compañías = DISTINCTCOUNT('transactions'[business_id])
```

La cual vendría a contar el número de compañías que tiene cada país y que realizan transacciones.

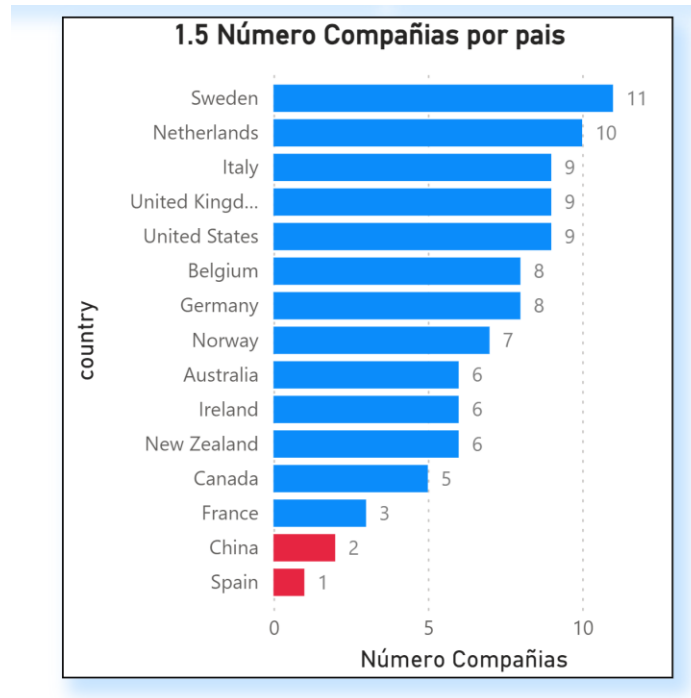
A partir de aquella medida junto con el campo de país de procedencia de la empresa se genera el siguiente KPI.

Se utilizó un gráfico de barras para representar el **KPI** en donde el **eje de y** (Y-axis) sería el **country** mientras que el **eje de x** (X-axis) vendría a ser la medida **Número Compañías**.



Interpretación:

Este KPI muestra el listado de países con su respectiva cantidad de empresas que realizan transacciones. Hemos condicionado el color de las barras para que si el valor es menos a la meta empresarial de 3 compañías por país entonces esta se muestre en rojo. En caso contrario se muestra en azul. A partir de esta tabla podemos sugerir que se tomen las estrategias y medidas necesarias para en un futuro cumplir con la meta empresarial para todos los países de procedencia de las empresas.



Ejercicio 1.6

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.

Solución:

Se han creado las siguientes medidas:

```
1 Número de Transacciones Declinadas =  
2 CALCULATE(  
3     COUNT(transactions[declined]),  
4     transactions[declined] = 1)
```

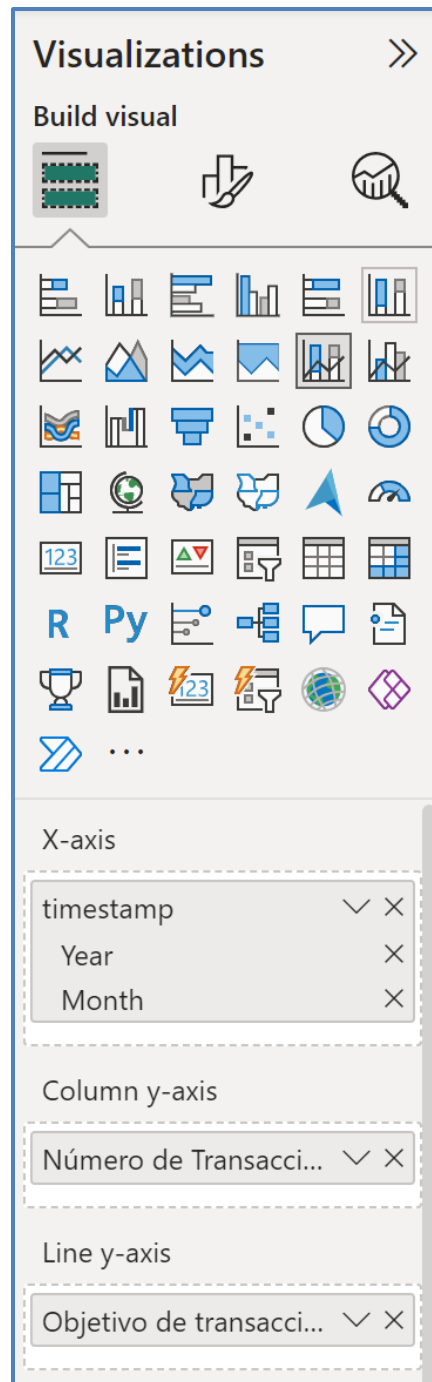
Esta medida realiza un conteo del número de transacciones declinadas, es decir, cuando el booleano **declined** sea igual a 1.



Objetivo de transacciones declinadas = 10

Esta medida formará parte del KPI estableciendo como objetivo de número de transacciones declinadas por mes menor a 10.

Se utilizó un gráfico de líneas y columnas apiladas (Line and stacked column chart) para elaborar el **KPI** en donde el **eje de x** (X-axis) sería el **timestamp**, considerando el mes y año; mientras que el **eje de y** (Y-axis) vendría a ser la medida **Número transacciones declinadas**. La **línea en el eje de y** (line Y-axis) sería la medida **Objetivo de transacciones declinadas**.



Por otro lado, al elaborar el **KPI** se ha condicionado el color de las barras que muestran el número de transacciones declinadas de cada mes. De esa manera si el valor fuere menor a 10 la barra será verde y si fuere mayor o igual a 10, sería roja.

Color - Categories

Format style

Rules

What field should we base this on?

Número de Transacciones Declinadas

Rules

Reverse color order

+ New rule

If value

>=

0

Number

and

<

10

Number

then

Green

If value

>=

10

Number

and

<=

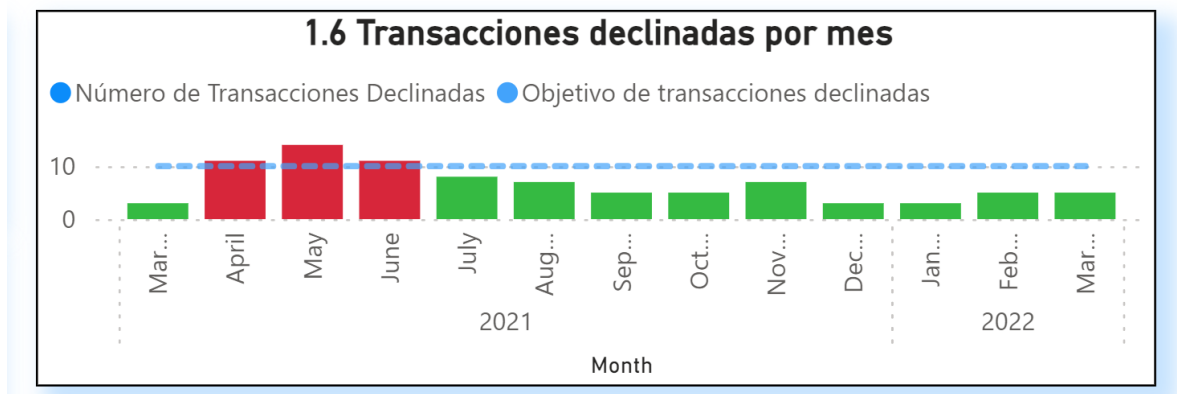
100

Number

then

Red

Interpretación:



Al organizar por mes el número de transacciones declinadas se ha obtenido el KPI de la figura mostrada arriba. Queda claro que en los meses de abril, mayo y junio el número de transacciones declinadas fue superior al objetivo. Este análisis es meramente descriptivo, por lo que queda pendiente indagar el porqué esto tuvo lugar en meses consecutivos.

Ejercicio 1.7

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 € en transacciones por mes.

Solución:

Se han creado las siguientes medidas:

Suma de Ventas € = `SUM(transactions[amount])`

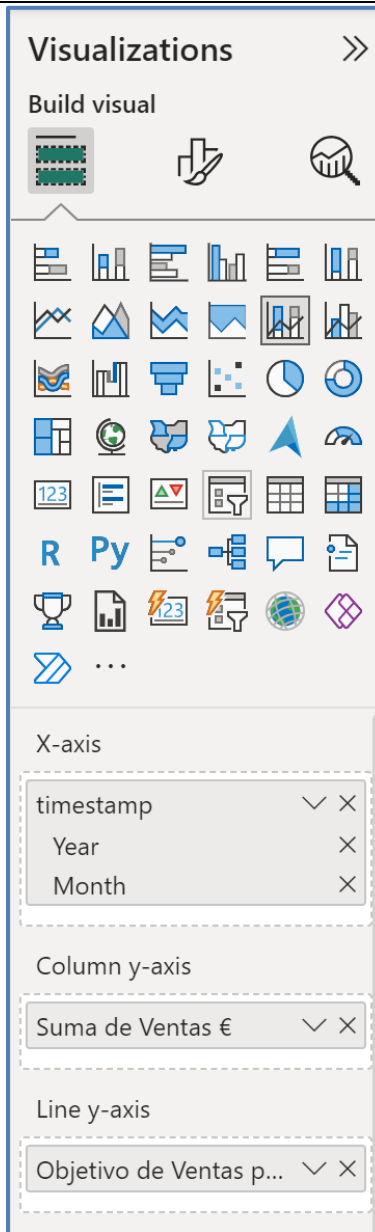
Esta medida calcula la suma del monto de las transacciones, que luego se agrupará por mes para la elaboración del KPI.



Objetivo de Ventas por mes = `10000`

Esta medida será la que establezca el umbral por encima del cual deben estar los monto de las ventas por mes para considerar que se ha cumplido el objetivo, según el KPI.

Se utilizó un gráfico de líneas y columnas apiladas (Line and stacked column chart) para elaborar el **KPI** en donde el **eje de x** (X-axis) sería el **timestamp**, considerando el mes y año; mientras que el **eje de y** (Y-axis) vendría a ser la medida **Suma de Ventas €**. La **línea en el eje de y** (line Y-axis) sería la medida **Objetivo de Ventas por mes**.



Por otro lado, al elaborar el **KPI** se ha condicionado el color de las barras que muestran la suma de ventas de cada mes. De esa manera si el valor fuere menor a 10.000 € la barra será roja y si fuere mayor o igual a 10.000 €, sería verde.

Color - Categories

Format style

Rules

What field should we base this on?

Suma de Ventas €

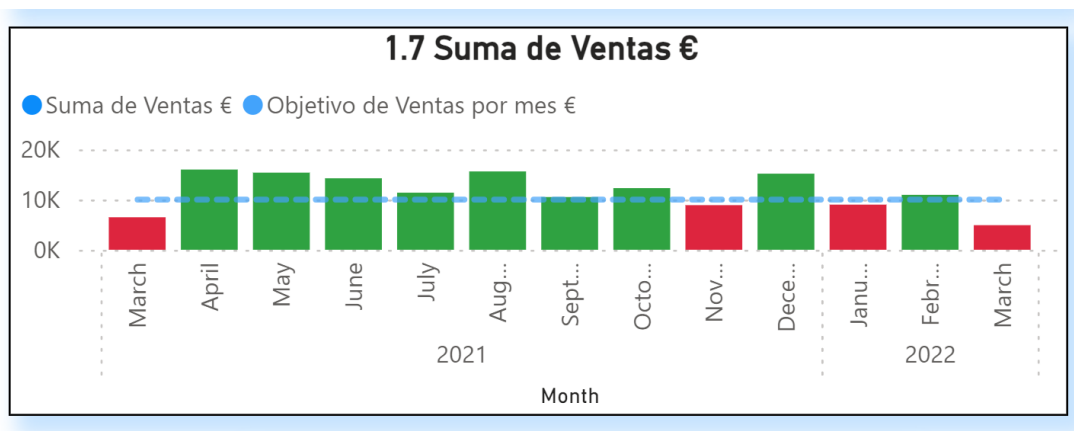
Rules

Reverse color order

New rule

If value	>=	0	Number	and	<	10000	Number	then		↑ ↓ ×
If value	>=	10000	Number	and	<	10000	Number	then		↑ ↓ ×

Interpretación:



La representación gráfica del KPI muestra que no en todos los meses se alcanzó el objetivo mensual de transacciones, establecido en 10.000 €. Concretamente, los meses de marzo y noviembre del 2021 y enero y marzo del 2022 evidencian valores de ventas por debajo del objetivo trazado. Dado que marzo se repite para ambos años, es muy probable que se deba a un factor estacional. Se recomienda realizar campañas agresivas de marketing orientadas a incrementar el monto de transacciones en tales meses flojos.

Ejercicio 1.8

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:

- Nombre y apellido de los usuarios/as (habrá que crear una nueva columna que combine esta información)
- Edad de los usuarios/as
- Media de las transacciones en euros
- Media de las transacciones en dólares (conversión: 1 euro equivale a 1,08 dólares)

Se deben realizar los cambios necesarios para identificar a los usuarios/as que tuvieron una media de 300 o más euros y 320 o más dólares en sus transacciones.

Solución:

Se han creado las siguientes medidas:

```
Edad = int(YEARFRAC(users[birth_date],TODAY()))
```

Esta medida calcula la edad de cada usuario a partir de su fecha de nacimiento y la fecha de hoy.



```
Promedio Ventas Euros = AVERAGE('transactions'[amount])
```

Esta medida nos da como resultado el promedio de ventas en euros, que es la moneda usada en la base de datos.



```
Promedio Ventas Dolares = [Promedio Ventas Euros] * 1.08
```

Esta medida muestra el promedio de ventas en dólares, para lo cual hemos considerado el valor en euros y multiplicado por 1.08, que vendría a ser la tasa de conversión de euros a dólares.



Previamente se generó una nueva columna denominada Nombre y Apellido, la cual viene a ser la concatenación de la columna name y surname. Se utilizó el siguiente código en DAX:


```
Nombre y Apellido = users[name]&" "&users[surname]
```



Si utilizamos el ratio de 1.08 para convertir 300 € en dólares el resultado sería 324 \$ y no 320 \$ como nos lo indica el enunciado como valor equivalente. Es por ello que hay clientes que aún habiendo comprado por un valor mayor o igual a 320 \$, no alcanzan a los 300 €, según la tabla.

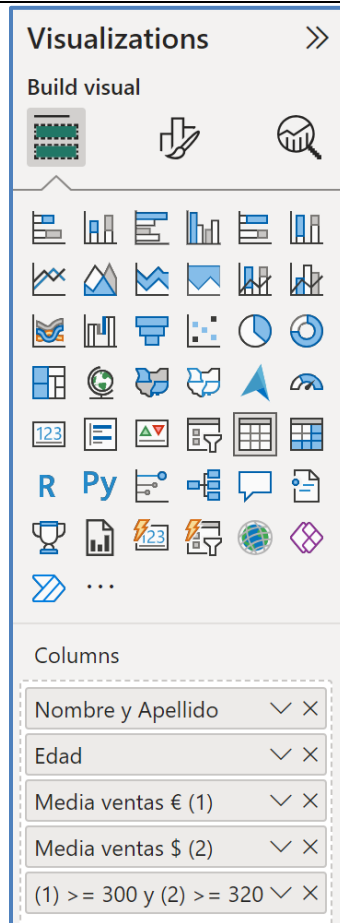
He interpretado que se deben cumplir las dos condiciones. Es decir:


- a) que el cliente haya comprado por un valor mayor o igual a 300 €. (**Promedio Ventas Euros >= 300€**)
- b) que el cliente haya comprado por un valor mayor o igual a 320 \$. (**Promedio Ventas Dólares >= 320\$**)

Es por ello que se ha creado una nueva medida denominada **IndicadorCompra** con un resultado booleano, en donde el valor será 1 sí y solo sí se cumplen a) y b). En caso contrario el valor resultante será 0.

```
IndicadorCompra =  
IF(  
    AND([Promedio Ventas Euros] >= 300, [Promedio Ventas Dólares] >= 320),  
    1,  
    0  
)
```

Se utilizó una tabla (table) para elaborar el **KPI** en donde las columnas estarían conformadas por los campos **Nombre y Apellido**, **Edad** y las medidas creadas: **Media ventas €** (Promedio Ventas Euros), **Media ventas \$** (Promedio Ventas Dólares), **(1) >= 300 y (2) >= 320** (IndicadorCompra).



Adicionalmente, a través de Visualizaciones > elementos de celda (cell elements) > icons hemos indicado que se señale a través de un icono de check verde  al lado izquierdo del campo **Nombre y Apellidos** a aquellos usuarios cuyo **IndicadorCompra** tiene el valor de 1, es decir, cumplen las condiciones (a) y (b).

Visualizations >>

Format visual

Search

Visual General ...

> Specific column

▼ Cell elements

Apply settings to

Series

Nombre y Apellido ▼

Background color ☐ Off

Font color ☐ Off

Icons ☒ On

Web URL ☐ Off

Icons - Icons ×

Format style ▼

Apply to ▼

What field should we base this on?

▼



Icon layout ▼

Icon alignment ▼










Style ▼

Rules

If value ▼ ▼ then ▼

De manera independiente se ha indicado que los campos **Media Ventas €** y **Media ventas \$** esten marcados con un check  al lado derecho cuando cumplan las condiciones (a) y (b), respectivamente. Pero lo que nos interesa es el  al inicio de la fila.

Interpretación:

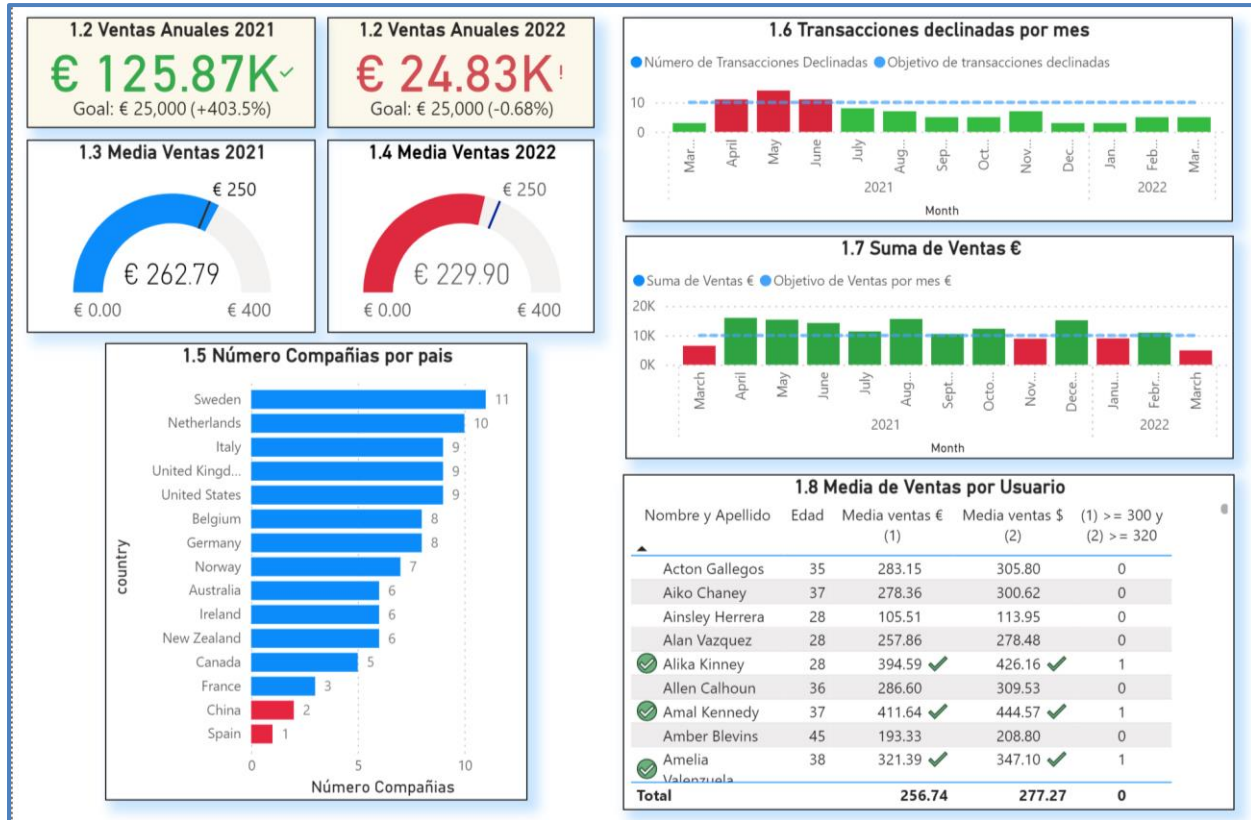
1.8 Media de Ventas por Usuario				
Nombre y Apellido	Edad	Media ventas € (1)	Media ventas \$ (2)	(1) >= 300 y (2) >= 320
Acton Gallegos	35	283.15	305.80	0
Aiko Chaney	37	278.36	300.62	0
Ainsley Herrera	28	105.51	113.95	0
Alan Vazquez	28	257.86	278.48	0
 Alike Kinney	28	394.59 	426.16 	1
Allen Calhoun	36	286.60	309.53	0
 Amal Kennedy	37	411.64 	444.57 	1
Amber Blevins	45	193.33	208.80	0
 Amelia Valenzuela	38	321.39 	347.10 	1
Total		256.74	277.27	0

Según el **KPI** generado que se muestra arriba no todos los usuarios superan la media de ventas de 300 € y 320 \$. Entonces se podría interpretar a los 300 € (o 320 \$) como un valor para el cohorte con el cual se puede segmentar a los clientes y enviarles campañas de marketing diferentes pues a primera vista se trata de clientes con distintos poderes adquisitivos y hábitos de consumo. Para ahondar en ese análisis es válido considerar la variable edad y agrupar a los clientes por rangos de edad.

Ejercicio 1.9

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

Se obtienen muchos insights interesantes a partir de las tablas y KPI's mostradas, por lo que se comentará brevemente líneas abajo.



Interpretación:

Se puede señalar sin lugar a dudas que la empresa está atravesando una fase de fuerte expansión, pues las ventas anuales han sido superadas con creces, tanto para el 2021 así como se espera para el 2022 (extrapolando los valores). Por otro lado los KPI's restantes detallan información valiosa para poder incrementar aún más las ventas, como por ejemplo captar más compañías de España y China, mercados con gran potencial pues poseen una gran demanda interna. Al mismo tiempo se podrían lanzar campañas de marketing más agresivas para los meses de bajas ventas, como enero, marzo y noviembre. Para finalizar, también se ha obtenido una primera segmentación de los clientes, de esa forma se podrían enfocar y personalizar las campañas de marketing.