Tasca S5.01. Iniciació a l'anàlisi de dades amb Power BI i indicadors

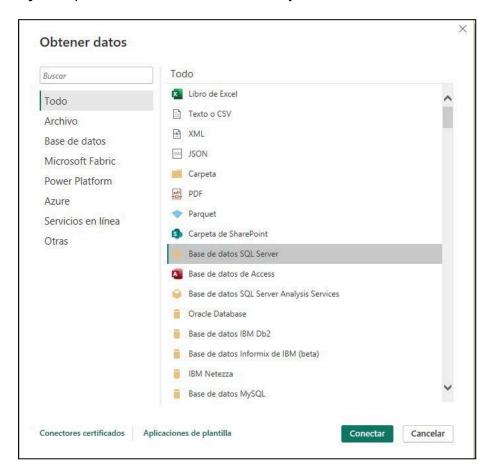
En aquest sprint, començaràs a aplicar els teus coneixements pràctics en Power BI utilitzant la base de dades prèviament utilitzada, que conté informació sobre una empresa dedicada a la venda de productes en línia. Durant els exercicis, és necessari que dediquis esforços a millorar la llegibilitat de les visualitzacions, assegurant-te de seleccionar les representacions visuals més adequades per a presentar la informació de manera clara i senzilla. No oblidis agregar títols descriptius als teus gràfics per a facilitar la comprensió de la informació visualitzada.

Nivell 1

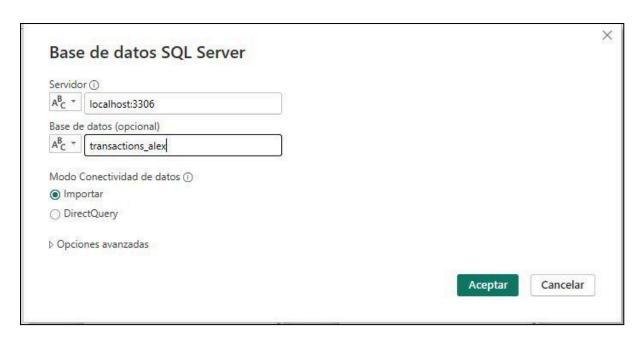
- Exercici 1

Importa les dades de la base de dades emprada prèviament. Després de carregar les dades, mostra el model de la base de dades en Power BI.

Importamos la base de datos del modelo mediante la opción **Importar base de datos MySQL**, previa instalación del conector MySQL Connector Net 9.1.0:



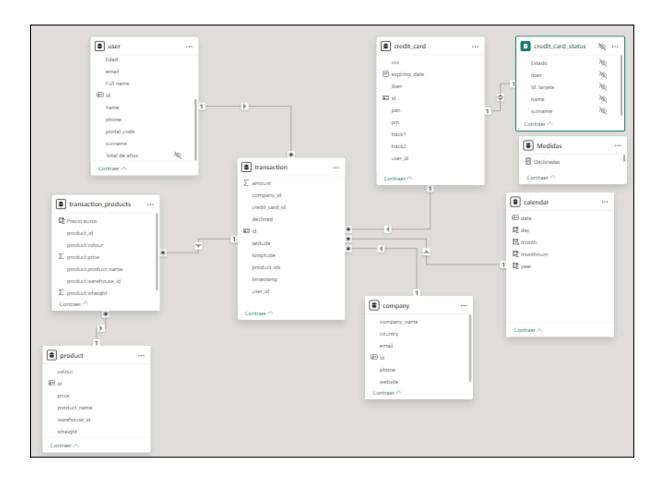
Introducimos los datos y las credenciales de la base de datos:



Y obtenemos el modelo de base de datos. Con Power Query, limpiamos los datos antes de incorporarlos. Hay que tener en cuenta:

- Los formatos de fecha: el campo **expired** de la tabla **credit_card** viene con la fecha en formato anglosajón mm/dd/yyyy; lo cambiamos a formato hispano dd/mm/yyyy
- Los precios en la tabla product vienen en dólares. Suponemos que el importe de las transacciones también están en dólares, que es consistente con el hecho que nuestros usuarios sean de USA, Canadá (dólares canadienses) y Reino Unido (libras esterlinas), que suponemos habrán pagado en dólares, conversión mediante. Transformamos este campo en decimal fijo.
- De paso, añadimos la columna age en la tabla user.
- Transformo el campo transaction.declined a booleano.

Comprobamos que las claves foráneas y la ordinalidad de las relaciones son correctas. Vemos que también se ha incorporado la tabla con la que identificamos las tarjetas activas o inactivas, que ocultaremos en el modelo. Simplificamos los nombres, eliminando el prefijo de la base de datos, y aprovechamos para crear una tabla de tiempos, con lo que nos quedaría así el modelo:



La teva empresa està interessada a avaluar la suma total del amount de les transaccions realitzades al llarg dels anys. Per a aconseguir això, s'ha sol·licitat la creació d'un indicador clau de rendiment (KPI). El KPI ha de proporcionar una visualització clara de l'objectiu empresarial d'aconseguir una suma total de 25.000 € per cada any.

Como estamos interesados en las ventas, vamos a crear una medida que discrimine por transacciones válidas:

Medida:

```
Ventas totales = CALCULATE(
    SUM('transaction'[amount]),
    'transaction'[declined]=false
)
```

Añadimos una medida constante para el objetivo de ventas anual:

```
Objetivo Ventas Anual = 25000
```

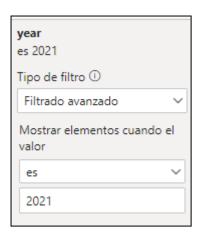
KPI año 2021:



KPI año 2022:



Para discriminar por año, hemos aplicado el siguiente filtro para el 2021 (y 2022):



La empresa batió el objetivo de ventas con creces en el año 2021, y este 2022, vista la tendencia (llevamos medio mes de marzo y los ingresos ya superan la mitad de los ingresos de marzo del 2021), la previsión es rebasar los ingresos del 2021.

- Exercici 3

Des de màrqueting et sol·liciten crear una nova mesura DAX que calculi la mitjana de suma de les transaccions realitzades durant l'any 2021. Visualitza aquesta mitjana en un mesurador que reflecteixi les vendes realitzades, recorda que l'empresa té un objectiu de 250.

Para calcular la media de las transacciones usaremos la siguiente medida en DAX para discriminar por las transacciones válidas:

```
Media Ventas = CALCULATE(
    AVERAGE('transaction'[amount]),
    'transaction'[declined]=false
)
```

CORRECCIÓN:

La empresa tiene un objetivo de 250 euros de media por compra, no 260 como indiqué erróneamente en la primera versión del ejercicio.



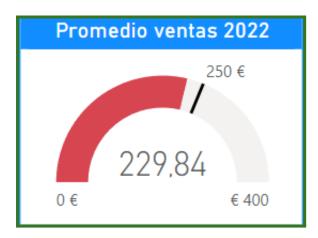
La empresa cumplió holgadamente el objetivo en el año 2021, aun cuando tenemos datos solo a partir de marzo.

Realitza el mateix procediment que vas realitzar en l'exercici 3 per a l'any 2022.

CORRECCIÓN:

Igual que en el ejercicio anterior, cambio el objetivo de 260 a 250 euros de media por compra.

Usamos el mismo objeto visual del ejercicio anterior, aplicando el filtro para el año 2022:



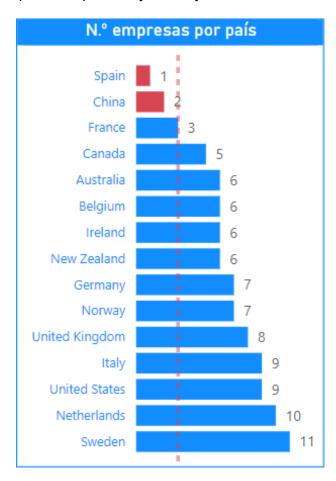
Aún estamos a mediados de marzo del 2022, con lo que se espera que el promedio se acerque al objetivo. Recomendaría, de cara a la campaña de primavera (meses de abril, mayo y junio, donde hay repunte importante de ventas en el 2021), ofrecer descuentos a nuestros mejores clientes para aumentar el número de unidades vendidas y mejorar este dato.

L'objectiu d'aquest exercici és crear una KPI que visualitzi la quantitat d'empreses per país que participen en les transaccions. La meta empresarial és garantir que hi hagi almenys 3 empreses participants per país. Per a aconseguir això, serà necessari utilitzar DAX per a calcular i representar aquesta informació de manera clara i concisa.

Creamos una medida DAX que nos permita hacer un DISTINCT COUNT de países de la tabla **company**:

```
Num empresas trans = DISTINCTCOUNTNOBLANK(
    'transaction'[company_id])
```

En el gráfico de barras segmentamos la medida DAX por país. Para el año 2021, la gráfica muestra los siguientes resultados. Marcamos el objetivo mínimo y discriminamos los países que están por debajo del objetivo:



CORRECCIÓN:

He invertido el eje y para resaltar los países con un número de empresas inferior al objetivo establecido, que, en este caso son España y China.

España apenas ha aportado 26 euros a los ingresos, como se puede comprobar al segmentar los datos por este país:



En el caso de China leemos un promedio de compra superior al objetivo, pero solo han realizado dos pedidos, uno de ellos declinado, con lo que se recomienda descartar este mercado.



Crea una nova KPI que permeti visualitzar la quantitat de transaccions declinades al llarg del temps. L'empresa va establir un objectiu de tenir menys de 10 transaccions declinades per mes.

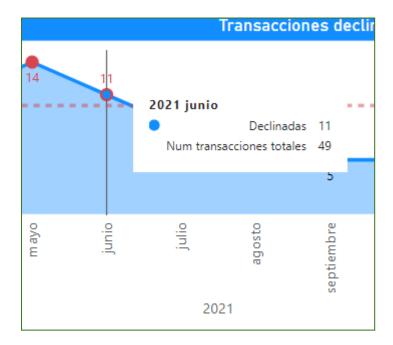
Creamos la medida DAX que calculará la suma de transacciones que cumplan con la condición de que el campo **declined** sea verdadero:

Dibujamos la evolución de transacciones declinadas por mes durante todo el periodo, indicando también el objetivo de transacciones máximas declinadas por mes y discriminando los meses que superan dicho valor:



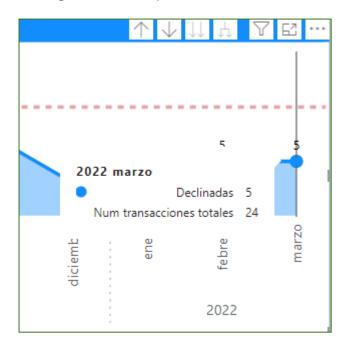
CORRECCIÓN:

Observamos que el número de transacciones rechazadas superó el máximo en los meses de abril, mayo y junio; siendo abril el tercer mes con mayor importe de ventas, mayo, el cuarto, y abril, el quinto, el impacto de esta tasa de transacciones fallidas ha afectado negativamente los ingresos de esos meses.



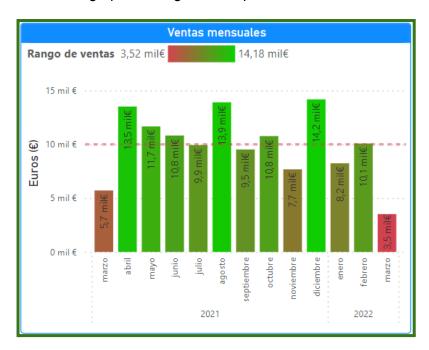
Vemos que se ha trabajado y se ha conseguido revertir esta tendencia a partir de junio del 2021, con un repunte en el mes de noviembre. Cabe señalar que la baja cantidad de transacciones fallidas en diciembre ha contribuido a conseguir que sea el mes con el valor máximo de ingresos por ventas.

Por otra parte, vemos que la tasa en marzo es demasiado alta en relación con el número de operaciones realizadas, revirtiendo la tendencia decreciente de finales del 2021. Hay que investigar las causas para volver a reducir esta tasa.



Crea un gràfic de columnes agrupades que reflecteixi la sumatòria de les vendes per mes. L'objectiu de l'empresa és tenir almenys 10.000 transaccions per mes.

Tomamos la medida DAX del ejercicio 2, **Ventas totales**, y dibujamos un gráfico de columnas agrupadas, segmentado por mes.

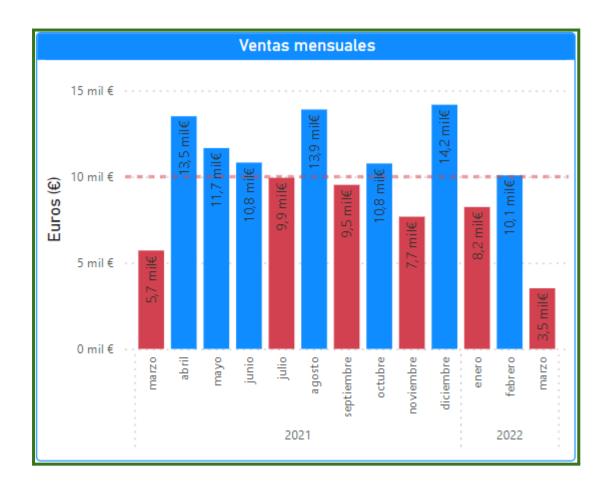


CORRECCIÓN:

Observamos tres picos de venta que coinciden con las vacaciones escolares de primavera (primera semana de abril), verano (sobre todo, agosto) y navidades (diciembre). En relación con el gráfico anterior, observamos el impacto negativo en los meses de mayo y junio, que coinciden con el final del curso escolar y que, dada la gama de productos que se comercializan (figuras de *Star Wars* y *Juego de tronos*, apreciadas entre escolares —y algunos sectores de adultos como el que analiza estos datos—), podrían ser superiores durante estos meses.



En la primera versión de este ejercicio había incorporado una graduación de colores para, mediante el color y el tono, valorar la tendencia de los resultados, pero resulta confuso. Dado que nos interesa tener claro qué meses han conseguido el objetivo y cuáles no, vamos a discriminar el color por estos dos casos:



En aquest exercici, es vol aprofundir en les transaccions realitzades per cada usuari/ària i presentar la informació de manera clara i comprensible. En una taula, presenta la següent informació:

- Nom i cognom dels usuaris/es (caldrà crear una nova columna que combini aquesta informació).
- Edat dels usuaris/es.
- o Mitjana de les transaccions en euros.
- o Mitjana de les transaccions en dòlars (conversió: 1 euro equival a 1,08 dòlars).

S'han de fer els canvis necessaris per a identificar als usuaris/es que van tenir una mitjana de 300 o més euros i 320 o més dòlars en les seves transaccions.

Creamos una nueva columna en la tabla **user** mediante Power Query, uniendo los campos **name** y **surname**. Cuando creamos la tabla, a esta columna le indicamos que la ordenamos por **surname**. Creamos la siguiente métrica DAX para calcular la media de las transacciones realizadas:

```
Promedio compra = CALCULATE(
         AVERAGE('transaction'[amount]),
         'transaction'[declined]=false
)
```

Creamos una tabla como objeto visual, en la que añadiremos el nombre completo, la edad, tal como habíamos calculado también en Power Query, la métrica **Promedio compra** y la métrica **Promedio compra dólares**, calculada simplemente multiplicando **Promedio compra** por el cambio:

Gasto medio por cliente						
Nombre	Edad	Dólares	Euros	% ▼	Desv.	% desv 🏽
🔕 Lynn Riddle	40	\$317,12 🛕	293,63 € 🕕	9 %	134,0	45,64 %
Hedwig Gilbert	33	\$275,16 🛕	254,78 € 🕕	7 %	161,2	63,25 %
🔕 Ocean Nelson	32	\$266,91 🛕	247,14 € 🕕	7 %	134,0	54,22 %
🔯 Brennan Wynn	36	\$285,80 🛕	264,63 € 🕕	5 %	137,4	51,91 %
⊗ Kenyon Hartman	42	\$255,72 📥	236,78 € ①	4 %	156,6	66,13 %
Slade Poole	23	\$242,34 🛕	224,39 € 🕕	4 %	131,3	58,50 %
Kim Mooney	23	\$220,93 🛕	204,56 € 🕕	3 %	137,8	67,36 %
Neil Powers	44	\$304,45 🛕	281,90 € 🕕	1 %	101,6	36,04 %
Lewis Melendez	34	\$496,57	459,79 € ⊘	1 %	34,4	7,48 %
Camilla Roach	31	\$481,12	445,48 € ⊘	1 %	17,5	3,93 %
Chloe Keith	25	\$469,39	434,62 € 🕢	1 %	40,4	9,30 %
Drake Walsh	40	\$469,30	434,54 € ⊘	1 %	50,5	11,62 %
Zephania Collins	30	\$451,55	418,11 € ⊘	1 %	2,8	0,68 %

Aprovechamos para calcular el porcentaje de negocio de cada cliente, que nos dará información importante sobre su impacto en los ingresos, así como la desviación estándar de la media, que nos indicará el rango de importes en que se mueve el cliente es sus operaciones.

El porcentaje se calcula así:

```
% compras cliente = DIVIDE(
   [Ventas totales],
   CALCULATE(
       [Ventas totales], ALL('transaction')
       )
   )
```

Y la desviación estándar:

```
Desviación standard = CALCULATE(
    STDEV.P('transaction'[amount]),
    'transaction'[declined]=false
)
```

Ordenamos por volumen de negocio y vemos que cinco de ellos operan para diversas empresas de varios países y, además, se mantienen activos a lo largo de todo el año, con lo que querremos fidelizarlos. Además, observamos que la media de compra de estos usuarios está por debajo del umbral especificado; así, pues, recomendaría ofrecerles ventajas y descuentos de cara a conseguir que aumenten las unidades y promocionar los artículos más caros, de cara a reducir la desviación estándar del promedio y aumentar este.

Vamos a presentar los Indicadores de los dos compradores más fieles: Lynn Riddle:



Y Hedwig Gilbert:



CORRECCIÓN:

Identifico con un icono verde, antepuesto al nombre, a los usuarios que cumplen con ambos requisitos: compras superiores a 300 euros **y también** superiores a 320 dólares, simplemente añadiendo la regla que el icono verde aparezca cuando **Promedio compra** >= 300 euros (que equivale a 324 \$, con lo que se cumplen ambas opciones). Aun así, en caso de que a alguien le pueda interesar saber qué usuarios están en el rango entre 320 y 324 \$ de media (que cumplirían la condición para esta moneda, pero no para la media en euros), mantengo la identificación en las columnas **Euros** y **Dólares**.

Redacta un paràgraf breu, de màxim 50 paraules, explicant el significat de les xifres presentades en les visualitzacions de Power BI. Pots interpretar les dades en general o centrar-te en algun país específic. Acompanya les interpretacions realitzades amb la captura de pantalla de les visualitzacions que analitzaràs.

Siguiendo la tendencia mensual de ventas del 2021, en abril de 2022 se superará el objetivo de ingresos.



Alemania lidera los ingresos con un 21 % (22.000 € sobre el total 107.000 €) con una baja tasa de transacciones rechazadas. En 2022 no ha realizado ninguna transacción; urge tomar acciones para recuperar este mercado.



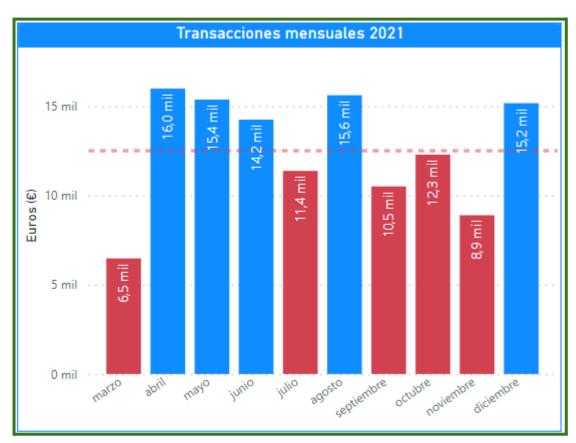
Nivell 2

Exercici 1

Des de l'àrea de màrqueting necessiten examinar la tendència mensual de les transaccions realitzades l'any 2021, específicament, volen conèixer la variació de les transaccions en funció del mes. Recorda visualitzar la meta empresarial d'aconseguir almenys 12.500 € en transaccions per mes. En aquest exercici, serà necessari que s'aconsegueixi identificar els mesos en què no es va aconseguir la meta establerta. De ser necessari pots realitzar dues visualitzacions.

CORRECCIÓN:

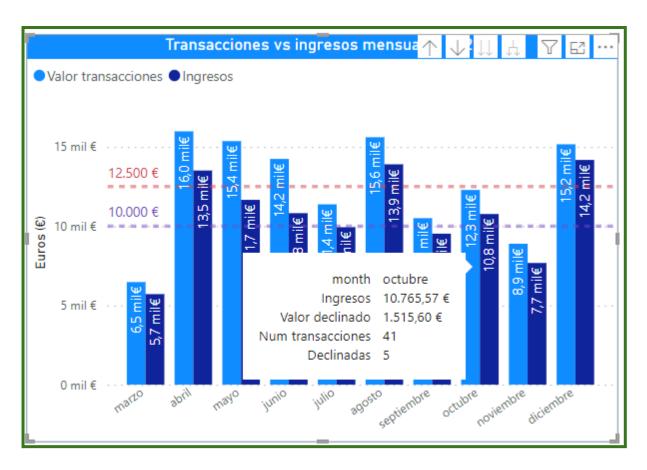
En este primer gráfico estamos interesados en el valor total de las transacciones, independientemente de si han sido aceptadas o no. Para esta primera gráfica usaremos el campo **amount** de la tabla **transaction**, y trazamos una gráfica de barras para ver la evolución mensual durante el 2021. Discriminamos por color los meses que alcanzan el objetivo de 12.500 €:



Para interpretarlo, vamos a apoyarnos en otra gráfica, donde comparamos el valor de las transacciones con el valor de los ingresos. Aquí marcaremos los respectivos objetivos con sendas líneas constantes. Situaremos una nueva medida:

```
Valor declinado = CALCULATE(SUM('transaction'[amount]) - [Ventas totales])
```

que mostrará la diferencia entre el valor de las transacciones y la cantidad ingresada, y que mostraremos como información de herramienta:



Se observan cinco meses en que se supera el objetivo del valor de las transacciones; por orden de cuantía: abril, agosto, mayo, diciembre y junio. Sin embargo, obtuvimos más ingresos reales en diciembre que en abril y agosto: en este mes se logró rentabilizar más las operaciones al disminuir las transacciones rechazadas. Por otra parte, mayo es el mes con mayor pérdida de ingresos, con 14 transacciones rechazadas y un monto no ingresado de 3.705 €.

Dicho esto, una vez identificados los meses con más demanda, vemos que los meses de primavera (abril, mayo y junio, desde las vacaciones de primavera al inicio de las de verano) conforman la campaña de ventas más potente y sostenida en el tiempo; aparte, tenemos los picos de transacciones de agosto y diciembre. Octubre se sitúa justo por debajo del objetivo de transacciones; septiembre, noviembre son los meses que más se alejan. No tenemos en cuenta los datos de marzo, ya que las primeras transacciones registradas son de fecha 17 de marzo.

Exercici 2

En el teu treball, es vol aprofundir en la comprensió de les transaccions realitzades a Alemanya. Per tant, et sol·liciten que desenvolupis mesures DAX per a crear visualitzacions que destaquin la mitjana de vendes a Alemanya. Tingues present que l'empresa té com a objectiu aconseguir una xifra de 250 euros anuals. Configura la visualització de manera que el valor mínim sigui 100 i el màxim 350, brindant així una representació més efectiva de la informació.

Calculamos la nueva medida DAX aprovechando la medida Media Ventas:

```
Media Ventas Alemania = CALCULATE(
    'Medidas'[Media Ventas],
    'company'[country] = "Germany"
)
```



CORRECCIÓN:

Vemos que el promedio de ventas para Alemania está cerca del objetivo marcado, aunque no llega a cumplirlo. Dado que, durante los primeros meses del 2022, no se han consolidado ventas para este país, se insiste en emprender acciones comerciales para estos clientes y analizar las causas de este hecho.

Aprovechamos para mirar el impacto de las compras de empresas alemanas en nuestro ejercicio, y vemos que una sola empresa, Nunc Interdum, ha realizado 104 operaciones y sus compras aportan el 20 % de nuestra facturación. El resto de empresas alemanas solo ha realizado una operación cada una.

Empresas Alemania					
Empresa	Ventas totales	Prom. venta	Uds.	N.º trans. ▼	% ventas
Nunc Interdum Incorporated	25.266,56 €	242,95 € ♦	261	104	20 %
Ac Fermentum Incorporated	293,57 €	293,57 €	1	1	0 %
Ac Industries	396,15 €	396,15 €	4	1	0 %
Aliquam PC	280,34 €	280,34 €	3	1	0 %
Auctor Mauris Corp.	308,99 €	308,99 €	1	1	0 %
Augue Foundation	15,05 €	15,05 € 🔷	1	1	0 %
Convallis In Incorporated	60,99 €	60,99 € 🔷	4	1	0 %
Rutrum Non Inc.	266,90 €	266,90 €	3	1	0 %
Total	26.888,55 €	242,24 €	278	111	21 %

Exercici 3

Escriu un breu paràgraf, màxim de 25 paraules, indica en quin mes no es va arribar a complir amb l'objectiu proposat de l'exercici 1.

Noviembre es el mes con menor cantidad y valor de transacciones, ya que las ventas se concentran en diciembre de cara a la Navidad.

Nivell 3

Exercici 1

La secció de màrqueting vol aprofundir en les transaccions realitzades pels usuaris i usuàries. En conseqüència, se't sol·licita l'elaboració de diverses visualitzacions que incloguin:

- Les mesures estadístiques claus de les variables que consideris rellevants per a comprendre les transaccions realitzades pels usuaris/es.
- o Quantitat de productes comprats per cada usuari/ària.
- Mitjana de compres realitzades per usuari/ària, visualitza quins usuaris/es tenen una mitjana de compres superior a 150 i quins no.
- o Mostra el preu del producte més car comprat per cada usuari/ària.
- o Visualitza la distribució geogràfica dels usuaris/es.

En aquesta activitat, serà necessari que realitzis els ajustos necessaris en cada gràfic per a millorar la llegibilitat i comprensió. En el compliment d'aquesta tasca, s'espera que avaluïs acuradament quines variables són rellevants per a transmetre la informació requerida de manera efectiva.

Para evitar problemas en la visualización, combinaremos en Power Query la tabla **transaction_products** con los datos de productos de la tabla **products**. Hemos de tener en cuenta que los precios de los productos están en dólares, y los valores de las transacciones están en euros. Crearemos una columna calculada que nos dé el valor de los productos en esta moneda.

```
Valor productos = CALCULATE(
        SUM('transaction_products'[Precio euros]),
        'transaction'[declined]=false
)
```

Donde [Precio euros] es una columna calculada en la tabla products:

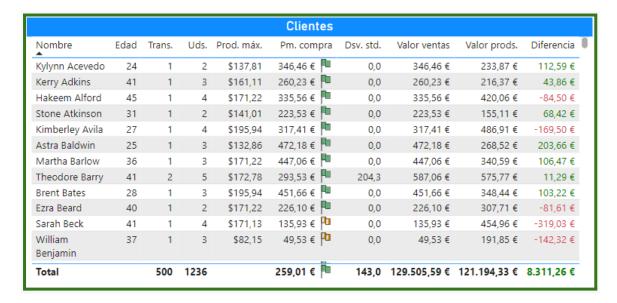
```
Precio euros = DIVIDE(transaction_products[product.price], 1.08, 0)
```

Y donde [Diferencia] es:

```
Diferencia = [Ventas totales] - [Valor productos]
```

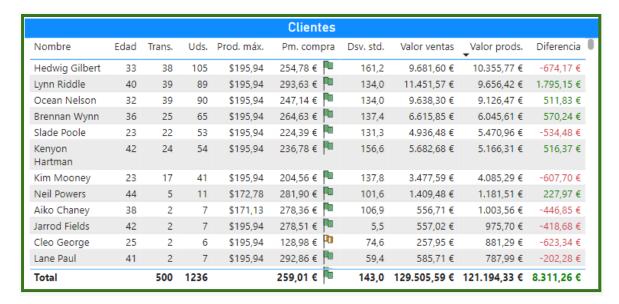
CORRECCIÓN:

Creamos una tabla de clientes, **ordenada alfabéticamente por el apellido**, que presente la siguiente información: edad, número de compras completadas, unidades totales compradas, valor promedio por venta realizada, **la desviación estándar para calibrar el rango en que suele situarse el importe de la compra del usuario** y el precio del producto más caro que ha adquirido, identificando con un icono verde aquellos compradores con una media de compra superior a 150, y con un icono amarillo los que no alcanzan esta media.



Esta información la queremos complementar con la ganancia (en color verde) o el descuento (en rojo) que hemos obtenido a través de las operaciones realizadas a través de nuestros agentes, y que se define por la diferencia entre el valor de los productos vendidos por la persona usuaria y el valor de las ventas. Queremos saber también quiénes son los clientes más fieles, que son los de mayor impacto en nuestros resultados, ya sea en ingresos o en unidades vendidas.

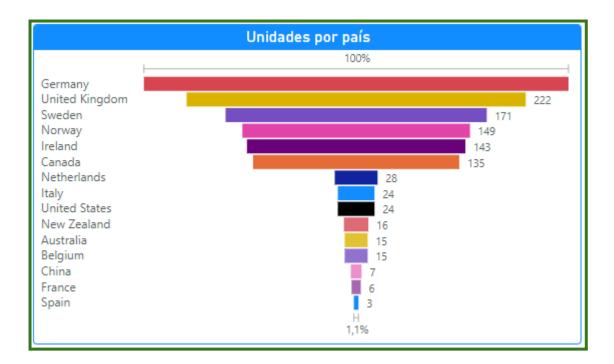
Si ordenamos la tabla por número de transacciones:



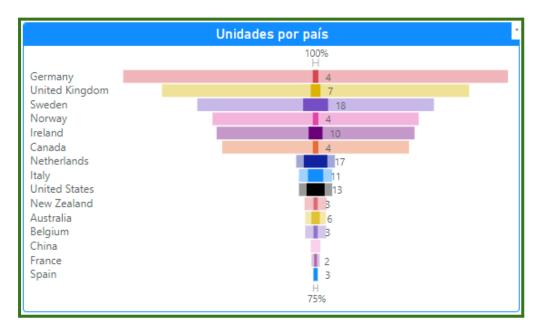
Comprobamos que la empresa, a lo largo de 2021 y 2022, obtiene unos beneficios netos totales de 8.311 €, pero también que, gracias a los descuentos, Hedwig Gilbert es el cliente que nos compra más productos. Por otra parte, Lynn Riddle es la que aporta mayores beneficios.

CORRECCIÓN:

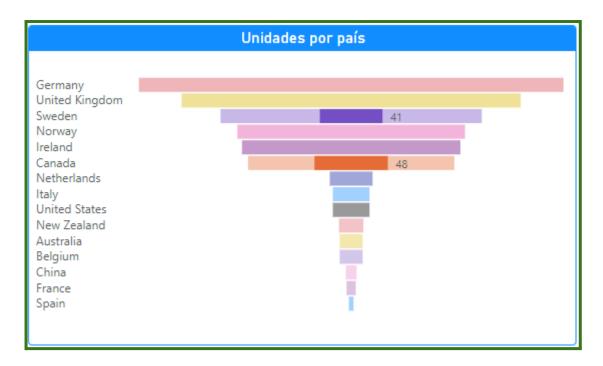
Como hay países que han realizado muy pocas compras y no se pueden graficar correctamente en una gráfico de áreas o de anillo, vamos a usar un gráfico de cascada para comparar el número de unidades compradas por país:



Vamos a fijarnos en la distribución por país de las compras por país de Hedwig Gilbert: lo seleccionamos en la tabla **Clientes** y nos fijamos en la distribución en la gráfica **Unidades por país**:



En el caso de Lynn Riddle, sus compras se centran en dos países, Suecia y Canadá:



Gilbert intenta abrir nuevos mercados para los productos mediante descuentos, pero el impacto de estos descuentos disperso en pocas ventas por cada país causa que la diferencia entre el valor de los productos vendidos y los ingresos obtenidos arroje un saldo negativo en su cuenta. Recomendaría identificar las empresas de los países que resultan más rentables y que Gilbert focalizarse en ellas las operaciones.

También nos interesará saber cuáles son los productos más vendidos, así como los que menos o los que no se han vendido, para ajustar los pedidos y stocks:

	Ventas ar	tículos			
id	Producto	Unids. ▼	Precio (\$)	Precio euros	ı
23	riverlands north	52	\$169,96	157,37 €	П
67	Winterfell	49	\$195,94	181,43 €	П
43	duel	45	\$59,80	55,37 €	П
79	Direwolf riverlands the	44	\$132,86	123,02 €	П
61	Winterfell Lannister	44	\$28,01	25,94 €	П
2	Tarly Stark	43	\$9,24	8,56 €	П
13	palpatine chewbacca	42	\$139,59	129,25 €	П
1	Direwolf Stannis	41	\$161,11	149,18 €	U
97	jinn Winterfell	41	\$65,25	60,42 €	
53	kingsblood Littlefinger the	40	\$137,81	127,60 €	
47	Tully	40	\$82,15	76,06 €	
17	skywalker ewok sith	39	\$91,89	85,08 €	
41	Lannister Barratheon Direwolf	37	\$141,01	130,56 €	
7	north of Casterly	37	\$63,33	58,64 €	
Total		997			

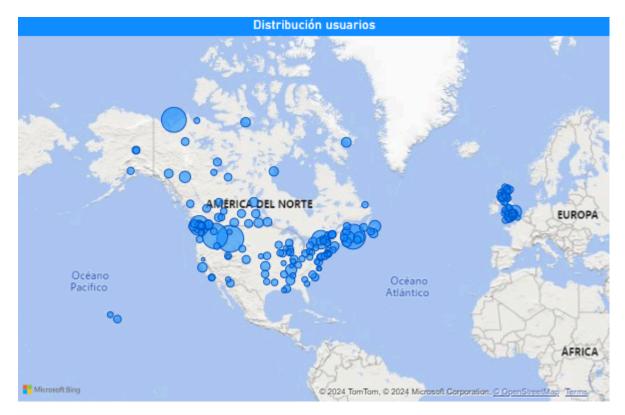
Y los artículos que no se han vendido, que serán aquellos cuyo **id** no esté en la tabla **transaction_products**. Para ello, crearemos una nueva columna en esta tabla, con la expresión en DAX:

Sin producto = INT(ISEMPTY('transaction_products'))

Y, en el filtro, indicaremos que buscamos los productos que tengan **Sin producto = 1**:

	Artículos sin v	ventas	
id	Producto	Color	warehouse
8	Winterfell	#383838	WH3
9	Winterfell	#b5b5b5	WH4
90	Winterfell	#757575	WH85
4	warden south duel	#111111	WH-1
20	warden Karstark	#b5b5b5	WH15
39	warden	#a0a0a0	WH34
55	Tully maester	#6d6d6d	WH50
35	the giantsblood maester	#8c8c8c	WH30
49	the giantsblood maester	#e2e2e2	WH44
16	the duel warden	#666666	WH11
62	the duel	#666666	WH57
81	the duel	#7f7f7f	WH76
91	the duel	#adadad	WH86

Por último, obtenemos la distribución geográfica de los usuarios:



Vemos que nuestros mejores agentes tienen su base de operaciones en Canadá y el noroeste de Estados Unidos.