

**MÁSTER EN DATA SCIENCE,
BIG DATA & BUSINESS ANALYTICS
2023/2024**

**TAREA DE BASE DE
DATOS
RELACIONALES
(SQL)**

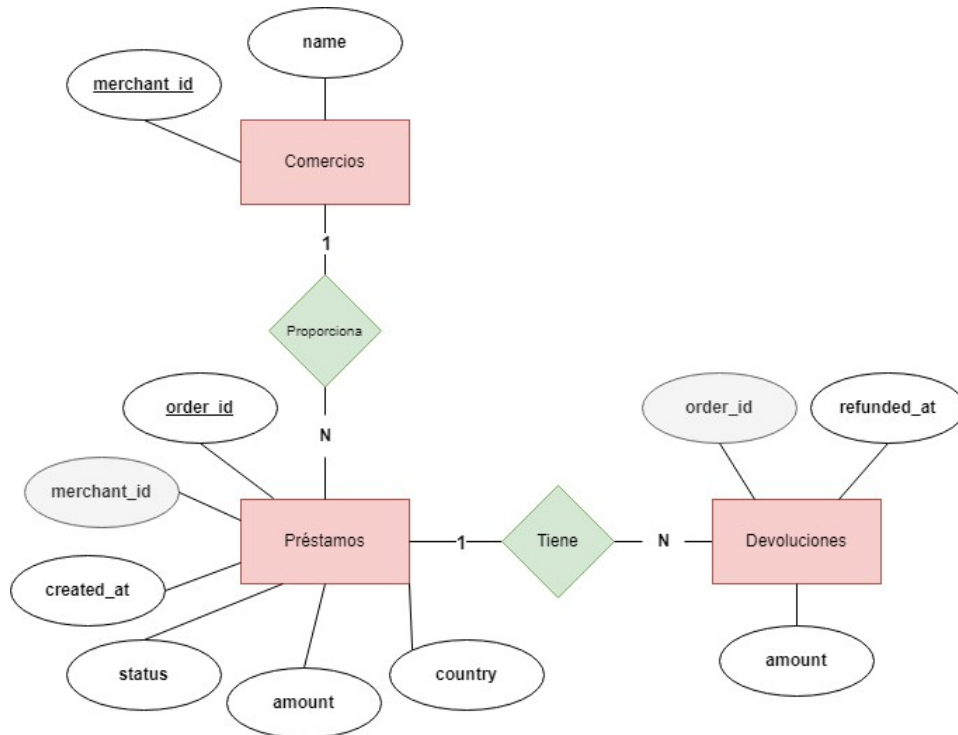
INSTITUCIÓN: UCM
ESTUDIANTE: Hao, Qi Xu

Índice

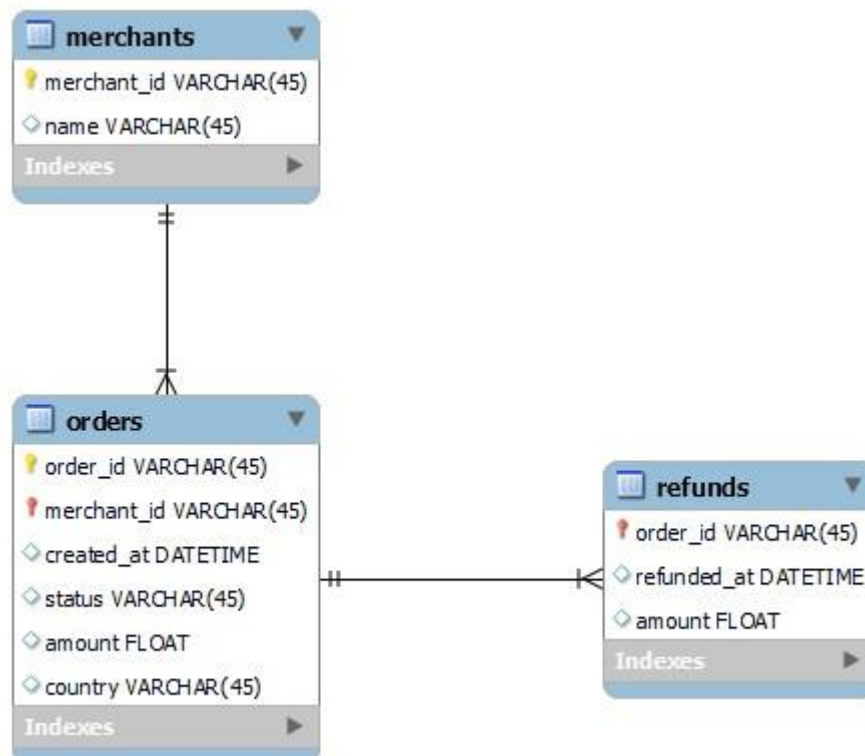
Ejercicio 1	1
Modelo Conceptual de Entidad-Relación.....	1
Esquema del base de datos operaciones_ucm	1
Consulta de la creación de la base de datos	2
Ejercicio 2	3
Consulta 2.1.....	3
Consulta 2.2.....	4
Ejercicio 3	5
Consulta 3.1.....	5
Consulta 3.2.....	7
Ejercicio 4	9
Funcionalidad desarrollada	9
Consulta de la funcionalidad y reflexión:	9

Ejercicio 1

Modelo Conceptual de Entidad-Relación:



Esquema del base de datos operaciones_ucm:



Consulta de la creación de la base de datos:

Cosas raras al importar las tablas:

- El nombre de la primera columna de cada tabla comienza con `id`;
- La primera fila de la tabla merchants se corresponde con los **nombres de las columnas** `id`, `merchant_id` y `name` cuando el primer registro del fichero merchants se había importado correctamente como nombres de las columnas.

DDL mediante Table Inspector después de importar los ficheros csv por import Wizard:

```
CREATE SCHEMA `tarea_ucm` ;
```

```
CREATE TABLE `merchants` (  
  `id` text,  
  `merchant_id` text,  
  `name` text  
)
```

```
CREATE TABLE `orders` (  
  `id` text,  
  `order_id` text,  
  `created_at` datetime DEFAULT NULL,  
  `status` text,  
  `amount` double DEFAULT NULL,  
  `merchant_id` text,  
  `country` text  
)
```

```
CREATE TABLE `refunds` (  
  `id` text,  
  `order_id` text,  
  `refunded_at` datetime DEFAULT NULL,  
  `amount` double DEFAULT NULL  
)
```

Ejercicio 2

1- Realizamos una consulta donde obtengamos por país y estado de operación, el total de operaciones y su importe promedio. La consulta debe cumplir las siguientes condiciones:

- a) Operaciones posteriores al 01-07-2015
- b) Operaciones realizadas en Francia, Portugal y España.
- c) Operaciones con un valor mayor de 100 € y menor de 1500€

Ordenamos los resultados por el promedio del importe de manera descendente.

Entiendo por total de operaciones el número total (conteo) de las operaciones en vez de la suma total.

Consulta 2.1:

```
SELECT
    country,
    status,
    count(amount) AS num_operaciones,
    avg(amount) AS importe_promedio
FROM orders
WHERE
    (date(created_at) > '2015-07-01') AND
    (country IN ('Francia', 'Portugal', 'España')) AND
    (amount BETWEEN 100 AND 1500)
GROUP BY country, status
ORDER BY importe_promedio DESC;
```

	country	status	num_operaciones	importe_promedio
►	Portugal	CLOSED	7	522.3814285714285
	España	DELINQUENT	21	423.4757142857144
	España	ACTIVE	172	410.67296511627904
	Portugal	ACTIVE	5	392.97999999999999
	Francia	CLOSED	44	391.22272727272724
	Francia	ACTIVE	93	347.2960215053766
	España	CLOSED	160	344.95725
	Francia	DELINQUENT	12	336.53416666666667
	Francia	CANCELLED	10	322.339
	España	CANCELLED	21	296.72238095238095

2- Realizamos una consulta donde obtengamos los 3 países con el mayor número de operaciones, el total de operaciones, la operación con un valor máximo y la operación con el valor mínimo para cada país. La consulta debe cumplir las siguientes condiciones:

- a) Excluimos aquellas operaciones con el estado “Delinquent” y “Cancelled”
- b) Operaciones con un valor mayor de 100 €

Consulta 2.2:

```
SELECT
    country,
    count(*) AS num_operaciones,
    sum(amount) AS total_operaciones,
    max(amount) AS max_operación,
    min(amount) AS min_operación
FROM orders
WHERE
    status NOT IN ('DELINQUENT', 'CANCELLED') AND
    amount > 100
GROUP BY country
ORDER BY num_operaciones DESC
LIMIT 3;
```

```
1 • SELECT
2     country,
3     count(*) AS num_operaciones,
4     sum(amount) AS total_operaciones,
5     max(amount) AS max_operación,
6     min(amount) AS min_operación
7 FROM orders
8 WHERE
9     status NOT IN ('DELINQUENT', 'CANCELLED') AND
10    amount > 100
11 GROUP BY country
12 ORDER BY num_operaciones DESC
13 LIMIT 3;
```

	country	num_operaciones	total_operaciones	max_operación	min_operación
►	España	342	145906.029999999994	2960.87	101
	Francia	141	56222.760000000001	1863.98	100.88
	Italia	75	23208.739999999998	1299	107.99

Ejercicio 3

A partir de las tablas incluidas en la base de datos tarea_ucm vamos a realizar las siguientes consultas:

1- . Realizamos una consulta donde obtengamos, por país y comercio, el total de operaciones, su valor promedio y el total de devoluciones. La consulta debe cumplir las siguientes condiciones:

- a) Se debe mostrar el nombre y el id del comercio.
- b) Comercios con más de 10 ventas.
- c) Comercios de Marruecos, Italia, España y Portugal.
- d) Creamos un campo que identifique si el comercio acepta o no devoluciones. Si no acepta (total de devoluciones es igual a cero) el campo debe contener el valor “No” y si sí lo acepta (total de devoluciones es mayor que cero) el campo debe contener el valor “Sí”. Llamaremos al campo “acepta_devoluciones”.

Ordenamos los resultados por el total de operaciones de manera ascendente.

Consulta 3.1:

SELECT

```
m.ï»¿merchant_id AS merchant_id,  
m.name AS nombre_comercio,  
o.country,  
count(*) AS num_operaciones,  
avg(o.amount) AS promedio_operaciones,  
count(r.refunded_at) AS num_devoluciones,  
CASE WHEN count(r.refunded_at) > 0 THEN 'Sí' ELSE 'No' END AS acepta_devoluciones
```

FROM orders AS o

INNER JOIN merchants AS m ON o.merchant_id = m.ï»¿merchant_id

LEFT JOIN refunds AS r ON o.ï»¿order_id = r.ï»¿order_id

GROUP BY nombre_comercio, o.country, merchant_id

HAVING

```
(num_operaciones > 10) AND
```

```
(country IN ('Marruecos', 'Italia', 'Españ±a', 'Portugal'))
```

ORDER BY num_operaciones ASC;

```

1 • SELECT
2     m.ï»¿merchant_id AS merchant_id,
3     m.name AS nombre_comercio,
4     o.country,
5     count(*) AS num_operaciones,
6     avg(o.amount) AS promedio_operaciones,
7     count(r.refunded_at) AS num_devoluciones,
8     CASE WHEN count(r.refunded_at) > 0 THEN 'Sí' ELSE 'No' END AS acepta_devoluciones
9 FROM orders AS o
10 INNER JOIN merchants AS m ON o.merchant_id = m.ï»¿merchant_id
11 LEFT JOIN refunds AS r ON o.ï»¿order_id = r.ï»¿order_id
12 GROUP BY nombre_comercio, o.country, merchant_id
13 HAVING
14     (num_operaciones > 10) AND
15     (country IN ('Marruecos', 'Italia', 'España', 'Portugal'))
16 ORDER BY num_operaciones ASC;

```

	merchant_id	nombre_comercio	country	num_operaciones	promedio_operaciones	num_devoluciones	acepta_devoluciones
►	pk_743f2fdecb876b75e975c005	Pepe Jeans	España	11	171.99363636363634	0	No
	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	Marruecos	13	365.356923076923	3	Sí
	pk_736c7094ea96eda38b098f56	Massimo Dutti	España	13	169.87692307692308	0	No
	pk_c15afcbd3a31b732f097ba7b	Havainas	España	16	323.02124999999995	0	No
	pk_07225590b8fea17e739aa451	Netflix	España	21	363.57190476190476	0	No
	pk_a3aa2fa07c5436f4c8ca1e03	fnac	España	22	531.8445454545455	0	No
	pk_c447a91e755425d163df6837	YouTube music	España	25	669.2836	1	Sí
	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	Italia	26	229.30153846153846	2	Sí
	pk_b9ee4936f19ba28d96f6001e	K-tuin	España	46	373.36478260869563	0	No
	pk_19d9ed34a670cbd04543ec35	Spotify	España	64	644.2579687500001	0	No
	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	España	137	314.75700729927	5	Sí

2- Realizamos una consulta donde vamos a traer todos los campos de las tablas operaciones y comercios. De la tabla devoluciones vamos a traer el conteo de devoluciones por operación y la suma del valor de las devoluciones. Una vez tengamos la consulta anterior, creamos una vista con el nombre orders_view dentro del esquema tarea_ucm con esta consulta.

Nota: La tabla refunds contiene más de una devolución por operación por lo que, para hacer el cruce, es muy importante que agrupemos las devoluciones.

Decidí agrupar la consulta por cada una de las columnas individuales de las tables ‘orders’ y ‘merchants’ con el fin de poder mostrarlos con las dos columnas agrupadas provenientes de la tabla ‘refunds’.

Consulta 3.2:

```
CREATE VIEW orders_view AS

SELECT

    o.ï»¿order_id AS order_id,

    date(o.created_at),

    o.status,

    o.amount,

    o.merchant_id,

    o.country,

    m.name AS nombre_comercio,

    sum(r.amount) AS monto_devoluciones,

    count(r.refunded_at) AS num_devoluciones

FROM orders AS o

INNER JOIN merchants AS m ON o.merchant_id = m.ï»¿merchant_id

LEFT JOIN refunds AS r ON o.ï»¿order_id = r.ï»¿order_id

GROUP BY 1, 2, 3, 4, 5, 6, 7

;

SELECT * FROM orders_view;
```

```

1 • CREATE VIEW orders_view AS
2 SELECT
3     o.ï»¿order_id AS order_id,
4     date(o.created_at),
5     o.status,
6     o.amount,
7     o.merchant_id,
8     o.country,
9     m.name AS nombre_comercio,
10    sum(r.amount) AS monto_devoluciones,
11    count(r.refunded_at) AS num_devoluciones
12 FROM orders AS o
13 INNER JOIN merchants AS m ON o.merchant_id = m.ï»¿merchant_id
14 LEFT JOIN refunds AS r ON o.ï»¿order_id = r.ï»¿order_id
15 GROUP BY 1, 2, 3, 4, 5, 6, 7
16 ;

```

	order_id	date(o.created_at)	status	amount	merchant_id	country	nombre_comercio	monto_devoluciones	num_devoluciones
►	5c3ef8170aee697c1ba8432a	2015-07-17	CLOSED	163.08	pk_317b4fc6fd80a5f8fb2ff216	Francia	Calcedonia	163.07999999999998	2
	5c3ef8170aee697c1ba8432b	2015-07-21	CLOSED	773.14	pk_317b4fc6fd80a5f8fb2ff216	Portugal	Calcedonia	773.14	1
	5c3ef8170aee697c1ba8432c	2015-07-23	CLOSED	191.05	pk_317b4fc6fd80a5f8fb2ff216	Alemania	Calcedonia	191.05	2
	5c3ef8170aee697c1ba8432d	2015-07-23	CLOSED	235.53	pk_317b4fc6fd80a5f8fb2ff216	Alemania	Calcedonia	235.53	1
	5c3ef8170aee697c1ba8432e	2015-07-24	CLOSED	302.06	pk_317b4fc6fd80a5f8fb2ff216	Alemania	Calcedonia	302.06	1
	5c3ef8170aee697c1ba8432f	2015-07-25	CLOSED	282.72	pk_317b4fc6fd80a5f8fb2ff216	Marruecos	Calcedonia	282.72	2
	5c3ef8170aee697c1ba84330	2015-07-25	CLOSED	440.63	pk_317b4fc6fd80a5f8fb2ff216	España	Calcedonia	440.63	1
	5c3ef8170aee697c1ba84331	2015-07-26	CLOSED	194.22	pk_317b4fc6fd80a5f8fb2ff216	Francia	Calcedonia	194.22	1
	5c3ef8170aee697c1ba84332	2015-07-26	CLOSED	214.24	pk_317b4fc6fd80a5f8fb2ff216	España	Calcedonia	214.24	1
	5c3ef8170aee697c1ba84333	2015-07-26	DELINQUENT	217.36	pk_317b4fc6fd80a5f8fb2ff216	España	Calcedonia	217.36	2
	5c3ef8170aee697c1ba84334	2015-07-26	CLOSED	261.28	pk_317b4fc6fd80a5f8fb2ff216	Italia	Calcedonia	261.28	1
	5c3ef8170aee697c1ba84335	2015-07-27	CLOSED	266.17	pk_317b4fc6fd80a5f8fb2ff216	Italia	Calcedonia	266.17	1
	5c3ef8170aee697c1ba84336	2015-07-28	CANCELLED	356.51	pk_317b4fc6fd80a5f8fb2ff216	España	Calcedonia	356.51	1
	5c3ef8170aee697c1ba84337	2015-07-28	CANCELLED	3.11	pk_317b4fc6fd80a5f8fb2ff216	Alemania	Calcedonia	3.11	1
	5c3ef8170aee697c1ba84338	2015-07-29	CLOSED	230.23	pk_317b4fc6fd80a5f8fb2ff216	Marruecos	Calcedonia	230.23	1
	5c3ef8170aee697c1ba84339	2015-07-29	CLOSED	221.62	pk_317b4fc6fd80a5f8fb2ff216	España	Calcedonia	NULL	0

Ejercicio 4

Funcionalidad desarrollada:

Se quiere llevar a cabo dos consultas con el propósito de analizar el volumen y la cantidad de préstamos efectuados en España en relación con otros países. Además, se quiere realizar una comparación de la cantidad total de préstamos/operaciones netas en España (descontando las devoluciones), desglosada por mes y comercio, con el objetivo de identificar los meses y comercios que han tenido un mayor impacto en la oferta total neta de préstamos proporcionados por los comercios ubicados en territorio español.

Consulta de la funcionalidad y reflexión:

-- Operaciones por países

```
SELECT
    country,
    count(amount) AS num_operación,
    round(sum(amount), 2) AS total_operación
FROM orders
GROUP BY country
ORDER BY 3 DESC;
```

	country	num_operación	total_operación
▶	España	420	164132.27
	Francia	175	66078.49
	Belgica	63	25954.16
	Italia	96	25800.56
	Alemania	63	18659.14
	Marruecos	42	16096.6
	Portugal	14	7479.62

Se observa un total de 420 operaciones realizadas en España en el 2015, el cual compone cerca del 48% de las operaciones totales, 873. Además, España cuenta también con el mayor monto de transacciones respecto a los demás países, 164.132,27.

-- Crear vista con las columnas necesarias (España) para consultas posteriores

```
CREATE VIEW españa_funcionalidad AS
```

```
SELECT
```

```
    o.country,
```

```
    m.name as nombre_comercio,
```

```
    o.created_at,
```

```
    o.amount as monto_operación,
```

```
    CASE WHEN r.amount IS NOT NULL THEN r.amount ELSE 0 END AS monto_devolución
```

```
FROM orders AS o
```

```
INNER JOIN merchants AS m ON o.merchant_id = m.í»¿merchant_id
```

```
LEFT JOIN refunds AS r ON o.í»¿order_id = r.í»¿order_id
```

```
WHERE country = 'España';
```

```
SELECT * FROM españa_funcionalidad;
```

```
1 • CREATE VIEW españa_funcionalidad AS
2   SELECT
3     o.country,
4     m.name as nombre_comercio,
5     o.created_at,
6     o.amount as monto_operación,
7     CASE
8       WHEN r.amount IS NOT NULL THEN r.amount ELSE 0
9     END AS monto_devolución
10  FROM orders AS o
11  INNER JOIN merchants AS m ON o.merchant_id = m.í»¿merchant_id
12  LEFT JOIN refunds AS r ON o.í»¿order_id = r.í»¿order_id
13  WHERE country = 'España'
14  ;
15 • SELECT * FROM españa_funcionalidad;
```

	country	nombre_comercio	created_at	monto_operación	monto_devolución
►	España	Calcedonia	2015-07-25 17:22:06	440.63	440.63
	España	Calcedonia	2015-07-26 15:59:23	214.24	214.24
	España	Calcedonia	2015-07-26 16:04:24	217.36	200
	España	Calcedonia	2015-07-26 16:04:24	217.36	17.36
	España	Calcedonia	2015-07-28 15:32:34	356.51	356.51
	España	Calcedonia	2015-07-29 19:08:20	221.62	0
	España	Calcedonia	2015-07-31 10:30:08	173.85	0
	España	Calcedonia	2015-08-01 11:37:34	159	0
	España	Calcedonia	2015-08-02 23:08:27	321.61	0
	España	Calcedonia	2015-08-03 16:30:43	403.1	0
	España	Kiko	2015-08-03 17:55:24	92.78	0

-- Conocer el sumatorio del total_operación_neto en España

```
SELECT ROUND(SUM(total_operación_neto),2) AS sumatorio_total_operación_neto
```

```
FROM (SELECT (SUM(monto_operación) - SUM(monto_devolución)) AS total_operación_neto FROM devolucion_mes) AS subconsulta;
```

```
1 • SELECT ROUND(SUM(total_operación_neto),2) AS sumatorio_total_operación_neto
2 FROM (SELECT (SUM(monto_operación) - SUM(monto_devolución)) AS total_operación_neto FROM devolucion_mes) AS subconsulta;
```

	sumatorio_total_operación_neto
▶	162394.14

Dicha cuantía, una vez descontado el monto de las devoluciones, resulta en un total de operaciones netas de 162.394,14. Esta cantidad será utilizado para consultadas posteriores para calcular la contribución en porcentaje por mes y comercio en España.

-- Operaciones desglosadas por mes en España

```
SELECT
```

```
    country,
```

```
    month(created_at) AS mes,
```

```
    count(monto_operación) AS num_operación,
```

```
    round(sum(monto_operación),2) AS total_operación,
```

```
    sum(monto_devolución) AS total_devolución,
```

```
    round(sum(monto_operación) - sum(monto_devolución),2) AS total_operación_neto,
```

```
    round((sum(monto_operación) - sum(monto_devolución)) /
```

```
        (SELECT ROUND(SUM(total_operación_neto),2)
```

```
        FROM (SELECT (SUM(monto_operación) - SUM(monto_devolución)) AS total_operación_neto FROM devolucion_mes) AS subconsulta)
```

```
    * 100,2) AS contribución_total_operación_neto
```

```
FROM españa_funcionalidad
```

```
GROUP BY mes;
```

```
1 • SELECT
2     country,
3     month(created_at) AS mes,
4     count(monto_operación) AS num_operación,
5     round(sum(monto_operación),2) AS total_operación,
6     sum(monto_devolución) AS total_devolución,
7     round(sum(monto_operación) - sum(monto_devolución),2) AS total_operación_neto,
8     round((sum(monto_operación) - sum(monto_devolución)) /
9         (SELECT ROUND(SUM(total_operación_neto),2)
10            FROM (SELECT (SUM(monto_operación) - SUM(monto_devolución)) AS total_operación_neto FROM devolucion_mes)
11                AS subconsulta) * 100,2) AS contribución_total_operación_neto
12 FROM españa_funcionalidad
13 GROUP BY mes;
14
```

	country	mes	num_operación	total_operación	total_devolución	total_operación_neto	contribución_total_operación_neto
▶	España	7	7	1841.57	1228.74	612.83	0.38
	España	8	28	8215.27	0	8215.27	5.06
	España	9	31	8896.7	0	8896.7	5.48
	España	10	37	11669.93	0	11669.93	7.19
	España	11	89	32160.57	0	32160.57	19.8
	España	12	229	101565.59	726.75	100838.84	62.1

Al descomponer el número total de operaciones por meses, se observa que en diciembre y noviembre se llevan a cabo más operaciones que el resto de los meses posiblemente debido a la temporada navideña, lo cual indica una cierta estacionalidad en el número de operaciones. En cuanto al cuantía total de préstamos netos de 162.394,14, los préstamos solamente en diciembre (62.1%) y noviembre (19.8%) juntos contribuyen casi el 81.9% de los préstamos netos totales ofrecidos por los comercios en España.

-- Operaciones desglosadas por comercio en España

```
SELECT
    country,
    nombre_comercio,
    count(monto_operación) AS num_operación,
    round(sum(monto_operación),2) AS total_operación,
    sum(monto_devolución) AS total_devolución,
    round(sum(monto_operación) - sum(monto_devolución),2) AS total_operación_netto,
    round((sum(monto_operación) - sum(monto_devolución)) /
        (SELECT ROUND(SUM(total_operación_netto),2)
        FROM (SELECT (SUM(monto_operación) - SUM(monto_devolución)) AS
        total_operación_netto FROM devolucion_mes) AS subconsulta)
        * 100,2) AS contribución_total_operación_netto
FROM españa_funcionalidad
GROUP BY nombre_comercio
ORDER BY contribución_total_operación_netto DESC;
```

```
1 • SELECT
2     country,
3     nombre_comercio,
4     count(monto_operación) AS num_operación,
5     round(sum(monto_operación),2) AS total_operación,
6     sum(monto_devolución) AS total_devolución,
7     round(sum(monto_operación) - sum(monto_devolución),2) AS total_operación_netto,
8     round((sum(monto_operación) - sum(monto_devolución)) /
9         (SELECT ROUND(SUM(total_operación_netto),2)
10        FROM (SELECT (SUM(monto_operación) - SUM(monto_devolución)) AS
11        total_operación_netto FROM devolucion_mes) AS subconsulta)
12        * 100,2) AS contribución_total_operación_netto
13 FROM españa_funcionalidad
14 GROUP BY nombre_comercio
15 ORDER BY contribución_total_operación_netto DESC;
```

	country	nombre_comercio	num_operación	total_operación	total_devolución	total_operación_netto	contribución_total_operación_netto
►	España	Calcedonia	137	43121.71	1228.74	41892.97	25.8
	España	Spotify	64	41232.51	0	41232.51	25.39
	España	K-tuin	46	17174.78	0	17174.78	10.58
	España	YouTube music	25	16732.09	726.75	16005.34	9.86
	España	fnac	22	11700.58	0	11700.58	7.21
	España	Netflix	21	7635.01	0	7635.01	4.7
	España	Havainas	16	5168.34	0	5168.34	3.18
	España	Speedo	7	3830.19	0	3830.19	2.36
	España	Massimo Dutti	13	2208.4	0	2208.4	1.36
	España	Kiko	8	2179.44	0	2179.44	1.34
	España	Pepe Jeans	11	1891.93	0	1891.93	1.17
	España	Loreal	2	1885.84	0	1885.84	1.16
	España	Sony	3	1566.9	0	1566.9	0.96

Por último, el análisis de las operaciones netas totales en España por comercio muestra que cerca del 83.54% de los préstamos netos totales provienen de los top 6 primeros comercios que ofrecen más préstamos: Calcedonia (25,8), Spotify (25,39), K-tuin (10,58), Youtube music (9,86), Fnac (7,21), y Netflix (4,7).

-- Recuento de los distintos comercios en España

Select count(distinct nombre_comercio) FROM devolucion_mes;

```
1 • Select count(distinct nombre_comercio) FROM españa_funcionalidad;
```

	count(distinct nombre_comercio)
▶	32

Por otro lado, los 26 comercios restantes en conjunto contribuyen solamente al restante 16,46% del préstamo neto total generado por las operaciones en España durante 2015.