

# Bases de Datos con SQL: TAREA FINAL

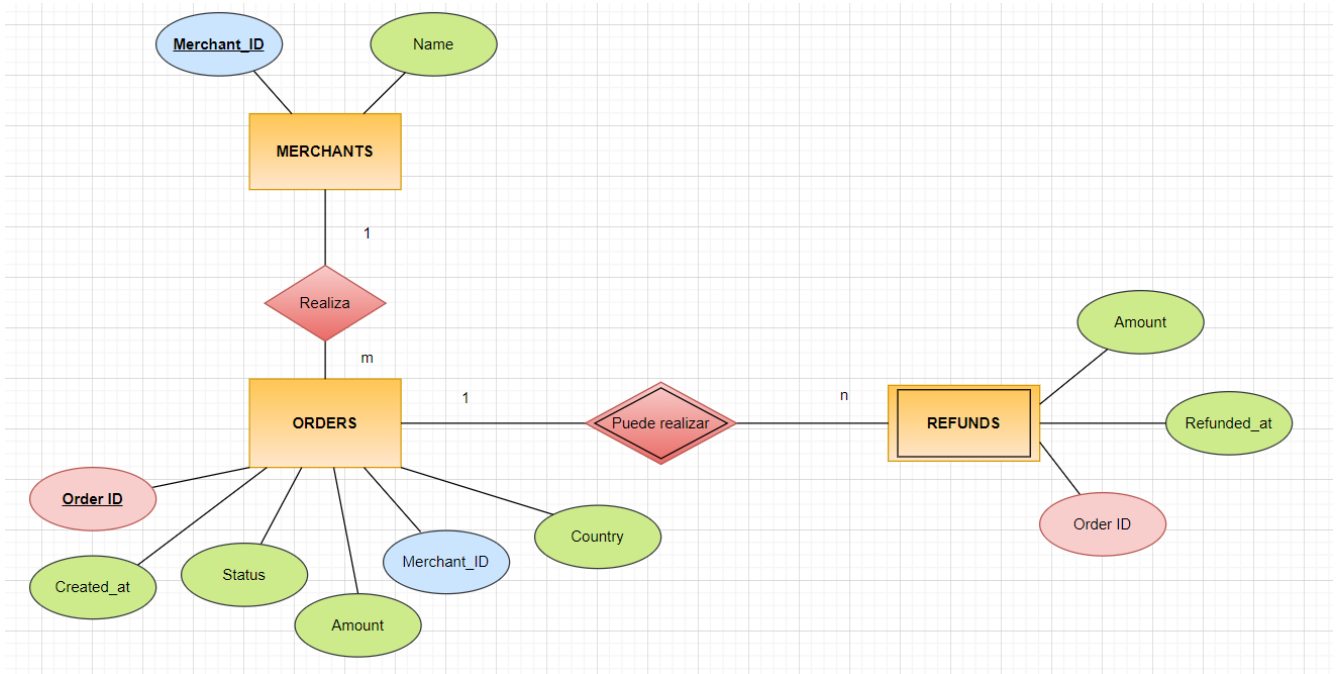
## Índice

<b>Ejercicio 1 .....</b>	<b>2</b>
1. Modelo Entidad-Relación. ....	2
2. Modelo Lógico. ....	3
3. Modelo Físico. ....	4
<b>Ejercicio 2 .....</b>	<b>6</b>
1. Primera consulta.....	6
2. Segunda consulta.....	7
<b>Ejercicio 3 .....</b>	<b>9</b>
1. Primera consulta.....	9
2. Segunda consulta.....	10
<b>Ejercicio 4 .....</b>	<b>14</b>
1. Objetivo de análisis.....	14
2. Análisis .....	14
3. Estrategias y acciones posteriores .....	17

## Ejercicio 1

Desarrollar un **Modelo de Entidad-Relación** indicando las entidades, atributos, relaciones y cardinalidad entre las tablas, el **Modelo Lógico** y el **Modelo Físico** mediante DDL.

### 1. Modelo Entidad-Relación.



#### Entidades:

- **Orders:** Representa los préstamos realizados por la empresa financiera.
- **Merchants:** Representa los establecimientos donde se pueden adquirir los productos.
- **Refunds:** Representa las devoluciones realizadas por los clientes.

#### Atributos:

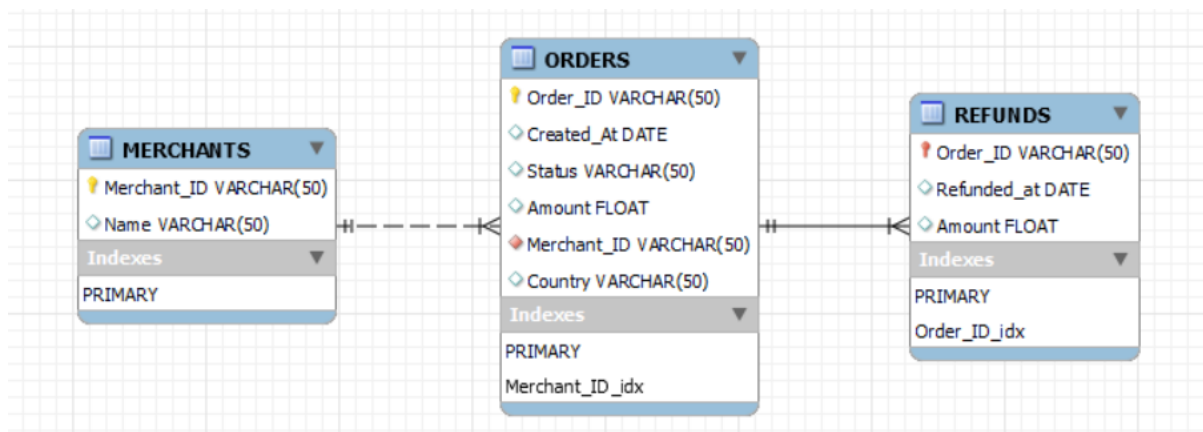
- **ORDERS:**
  - **Order\_ID:** Identificador único del pedido (*Primary Key*).
  - **Created\_at:** Fecha en la que se realizó el pedido.
  - **Status:** Estado del pedido (*Active, Closed, Delinquent, Cancelled*).
  - **Amount:** Importe del pedido.
  - **Merchant\_ID:** Identificador del establecimiento donde se realizó el pedido (*Foreign Key*).
  - **Country:** País donde se realizó el pedido.
- **MERCHANTS:**
  - **Merchant\_ID:** Identificador único del establecimiento (*Primary Key*).
  - **Name:** Nombre del establecimiento.
- **Refunds:**
  - **Order\_ID:** Identificador del pedido asociado a la devolución (*Foreign Key*).
  - **Refunded\_at:** Fecha en la que se realizó la devolución.
  - **Amount:** Importe de la devolución.

- No necesitaría clave primaria porque para identificar de forma única cada fila la combinación de **Order\_ID** y **Refunded\_at** sería suficiente, puesto que se trata de una combinación única.

### Relaciones y cardinalidad:

- Un cliente **realiza** un pedido en un comercio. Un comercio (*Merchants*) puede realizar muchas órdenes de compras (*Orders*), pero cada orden está asociada a un único comercio. La cardinalidad es de **uno a muchos (1:M)**: un comercio se relaciona con muchas órdenes, pero cada orden pertenece a un único comercio.
- Un cliente **puede realizar** una devolución de un pedido. Una orden (*Orders*) puede tener varias devoluciones (*Refunds*) si un cliente decide devolver el producto más de una vez o en múltiples pagos, pero una devolución está asociada a una única orden, ya que solo puede referirse a la devolución de un pedido específico. La cardinalidad es de **uno a varios (1:N)**: una orden puede tener varias devoluciones, pero cada devolución pertenece a una única orden. Existe una **relación identificativa**, ya que la entidad débil (*Refunds*) necesita que exista la entidad fuerte (*Orders*) para existir y se identifica mediante la entidad fuerte a la que está asociada y algún atributo de la entidad débil.

## 2. Modelo Lógico.



### Claves primarias:

- Order\_ID
- Merchant\_ID

### Relaciones:

- Cardinalidad de **uno a muchos (1:M)**: un comercio (*Merchants*) se relaciona con muchas órdenes (*Orders*), pero cada orden pertenece a un único comercio. La relación entre las dos tablas se establece mediante una clave foránea (*Merchant\_ID* en *Orders*) que referencia la clave primaria (*Merchant\_ID* en *Merchants*), indicando que cada orden está asociada a un comercio específico.
  - Se relacionan mediante una **línea discontinua** para indicar que una entidad puede existir independientemente de la otra, es decir, existe una relación opcional.

- Un pedido (*Orders*) solo puede tener una devolución (*Refunds*) y una devolución solo puede ser de un pedido (**1:1**). La relación entre las dos tablas se establece mediante una clave foránea *Order\_ID* en *Refunds*) que referencia la clave primaria (*Order\_ID* en *Orders*), indicando que cada devolución está asociada a una orden específica.
  - Se relacionan mediante una **línea continua** para indicar que los registros de una tabla dependen de los registros de la otra, es decir, existe una relación obligatoria ya que las devoluciones están directamente vinculadas a las órdenes correspondientes.

### 3. Modelo Físico.

```
CREATE SCHEMA prestamos_2015;
```

```
CREATE TABLE prestamos_2015.merchants(
    merchant_id VARCHAR(50),
    name VARCHAR(50)
);
```

```
CREATE TABLE prestamos_2015.orders(
    order_id VARCHAR(50),
    created_at DATE,
    status VARCHAR(50),
    amount FLOAT,
    merchant_id VARCHAR(50),
    country VARCHAR(50)
);
```

```
CREATE TABLE prestamos_2015.refunds(
    order_id VARCHAR(50),
    refunded_at DATE,
    amount FLOAT
);
```

```
4  -- Creamos un esquema de base de datos con el nombre prestamos_2015
5  • CREATE SCHEMA prestamos_2015;
6
7  -- Creamos las tres tablas correspondientes a los 3 archivos:
8  • CREATE TABLE prestamos_2015.merchants(
9      merchant_id varchar(50),
10     name varchar(50)
11 );
12
13 • CREATE TABLE prestamos_2015.orders(
14     order_id varchar(50),
15     created_at DATE,
16     status varchar(50),
17     amount FLOAT,
18     merchant_id varchar(50),
19     country varchar(50)
20 );
21
22 • CREATE TABLE prestamos_2015.refunds(
23     order_id varchar(50),
24     refunded_at DATE,
25     amount FLOAT
26 );
```

## Ejercicio 2

### 1. Primera consulta.

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:

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

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

```
SELECT o.country,
       o.status,
       COUNT(*) AS total_operaciones,
       ROUND(AVG(o.amount),2) AS importe_promedio
FROM orders AS o
WHERE o.created_at > '2015-07-01 23:59:59'
AND o.country IN ('Francia', 'Portugal', 'España')
AND o.amount BETWEEN 100 AND 1500
GROUP BY o.country, o.status
ORDER BY importe_promedio DESC;
```

```
4 • SELECT o.country,
5       o.status,
6       COUNT(*) AS total_operaciones,
7       ROUND(AVG(o.amount),2) AS importe_promedio
8 FROM orders AS o
9 WHERE
10    o.created_at > '2015-07-01 23:59:59'
11    AND o.country IN ('Francia', 'Portugal', 'España')
12    AND o.amount BETWEEN 100 AND 1500
13 GROUP BY o.country, o.status
14 ORDER BY importe_promedio DESC;
15
```

country	status	total_operaciones	importe_promedio
Portugal	CANCELLED	1	773.14
Portugal	CLOSED	6	480.59
Espana	DELINQUENT	20	433.78
Espana	ACTIVE	171	408.82
Portugal	ACTIVE	5	392.98
Espana	CANCELLED	5	391.1
Francia	CANCELLED	4	387.4
Francia	CLOSED	52	386.15
Francia	ACTIVE	91	341.83
Espana	CLOSED	178	339.4
Francia	DELINQUENT	12	336.53

### Conclusiones:

- Francia, a pesar de tener un alto volumen de operaciones, cuenta con el importe promedio más bajo en todos los estados, excepto en préstamos pagados ('Closed'), donde le gana España con un importe promedio de 339,40€.

- Portugal tiene la menor cantidad de operaciones en todos los estados y el importe promedio más alto en préstamos finalmente no realizados y/o reembolsados y préstamos pagados ('Cancelled' y 'Closed').
- España tiene el mayor número de operaciones en todos los estados (la mayoría activos), con un importe promedio que varía considerablemente, siendo el más alto el de los préstamos no pagados con plazo expirado ('Delinquent').
- En general, las operaciones 'Closed', muestran importes promedio más bajos, lo que podría indicar que los clientes que completan sus préstamos tienden a tener importes más bajos en promedio en comparación con aquellos en otros estados.

## 2. Segunda consulta.

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:

- Excluimos aquellas operaciones con el estado "Delinquent" y "Cancelled".
- Operaciones con un valor mayor de 100 €.

```
SELECT o.country,
COUNT(*) AS total_operaciones,
MAX(o.amount) AS importe_maximo,
MIN(o.amount) AS importe_minimo
FROM orders AS o
WHERE o.status NOT IN ('Delinquent', 'Cancelled')
AND o.amount > 100
GROUP BY o.country
ORDER BY total_operaciones DESC
LIMIT 3;
```

```
23 • SELECT o.country,
24     COUNT(*) AS total_operaciones,
25     MAX(o.amount) AS importe_maximo,
26     MIN(o.amount) AS importe_minimo
27 FROM
28     orders AS o
29 WHERE
30     o.status NOT IN ('Delinquent', 'Cancelled')
31     AND o.amount > 100
32 GROUP BY o.country
33 ORDER BY total_operaciones DESC
34 LIMIT 3;
```

Result Grid				
		Filter Rows:	Export:	Wrap Cell Content:
	country	total_operaciones	importe_maximo	importe_minimo
▶	Espana	359	2960.87	101
	Francia	147	1863.98	100.88
	Italia	77	1299	107.99



### **Conclusiones:**

- Aunque España lidera en términos de cantidad de operaciones (con un total de 359 operaciones), también tiene un rango amplio de importes, lo que sugiere una diversidad de clientes y tipos de transacciones en el mercado español. Cuenta con la operación con el importe máximo más alto entre los tres países (2.960,87€). Esto puede indicar que se realizaron transacciones significativas o préstamos de mayor valor en el mercado español.
- Francia tiene el segundo mayor número de operaciones (147) y el importe mínimo más bajo (100,88€). Esto sugiere que, aunque Francia tiene operaciones de elevado importe, también presenta una diversidad de transacciones que abarcan desde importes más bajos hasta importes relativamente altos.
- Italia tiene el menor número de operaciones (77), lo que indica una actividad financiera menor en comparación con España y Francia. También, cuenta con el tercer importe máximo más alto, lo que podría indicar que las transacciones de mayor valor son menos frecuentes en Italia (en comparación con los otros dos países) en el período considerado.

## Ejercicio 3

### 1. Primera consulta.

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.

```
SELECT o.country,
       m.merchant_id,
       m.name AS nombre_comercio,
       COUNT(o.merchant_id) AS total_operaciones
       ROUND(AVG(o.amount),2) AS valor_promedio
       COALESCE(SUM(r.conteo_devoluciones),0) AS conteo_devoluciones,
       CASE WHEN SUM(r.suma_devoluciones) > 0 THEN 'Sí' ELSE 'No' END AS
       acepta_devoluciones
FROM orders AS o
LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
LEFT JOIN (
    SELECT order_id,
           ROUND(SUM(amount),2) AS suma_devoluciones,
           COUNT(*) AS conteo_devoluciones
    FROM refunds
    GROUP BY order_id
) AS r ON o.order_id = r.order_id
WHERE o.country IN ('Marruecos', 'Italia', 'España', 'Portugal')
GROUP BY o.country, m.merchant_id, m.name
HAVING total_operaciones > 10
ORDER BY total_operaciones ASC;
```

```

49 • SELECT o.country,
50       m.merchant_id,
51       m.name AS nombre_comercio,
52       COUNT(o.merchant_id) AS total_operaciones,
53       ROUND(AVG(o.amount),2) AS valor_promedio,
54       COALESCE(SUM(r.conteo_devoluciones),0) AS conteo_devoluciones,
55       CASE WHEN SUM(r.suma_devoluciones) > 0 THEN 'Sí' ELSE 'No' END AS acepta_devoluciones
56 FROM orders AS o
57 LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
58 LEFT JOIN (
59     SELECT order_id,
60           ROUND(SUM(amount),2) AS suma_devoluciones,
61           COUNT(*) AS conteo_devoluciones
62 FROM refunds
63 GROUP BY order_id
64 ) AS r ON o.order_id = r.order_id
65 WHERE o.country IN ('Marruecos', 'Italia', 'España', 'Portugal')
66 GROUP BY o.country, m.merchant_id, m.name
67 HAVING total_operaciones > 10
68 ORDER BY total_operaciones ASC;

```

country	merchant_id	nombre_comercio	total_operaciones	valor_promedio	conteo_devoluciones	acepta_devoluciones
Espana	pk_743f2fdec876b75e975c005	Pepe Jeans	11	171.99	0	No
Marruecos	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	12	372.24	3	Sí
Espana	pk_736c7094ea96eda38b098f56	Massimo Dutti	13	169.88	0	No
Espana	pk_c15afcbd3a31b732f097ba7b	Havaianas	16	323.02	0	No
Espana	pk_07225590b8fea17e739aa451	Netflix	21	363.57	0	No
Espana	pk_a3aa2fa07c5436f4c8ca1e03	fnac	22	531.84	0	No
Espana	pk_c447a91e755425d163df6837	YouTube music	25	669.28	1	Sí
Italia	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	26	229.3	2	Sí
Espana	pk_b9ee4936f19ba28d96f6001e	K-tuin	46	373.36	0	No
Espana	pk_19d9ed34a670cbd04543ec35	Spotify	64	644.26	0	No
Espana	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	136	315.47	5	Sí

## Conclusiones:

- Calcedonia se muestra como líder en operaciones (174 operaciones totales), teniendo la mayor cantidad de operaciones totales en España (136). Además, es el único comercio con presencia en todos los países que aparecen en la consulta (Marruecos, Italia y España).
- España es el país que en el que se encuentra el comercio con menor número total de operaciones (Pepe Jeans) y mayor número total de operaciones (Calcedonia). Pepe Jeans también es el comercio con el menor importe promedio de la consulta (171,99€), mientras que YouTube music (también en España), cuenta con el mayor importe promedio (669,28€).
- YouTube music y Calcedonia son los únicos comercios que aceptan devoluciones, lo que puede influir positivamente en la satisfacción del cliente. Calcedonia en España tiene la mayor cantidad de devoluciones (5). Dentro de los comercios que no aceptan devoluciones, Spotify es el que mayor importe promedio cuenta con diferencia (644,26€).
- España es el país que aparece, con diferencia, más veces en esta consulta por lo que se puede identificar como un país clave.

## 2. Segunda consulta.

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 *prestamos\_2015* 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.

```
SELECT o.*,
       m.merchant_id AS merch_id,
       m.name AS nombre_comercio,
       COALESCE(r.conteo_devoluciones,0) AS conteo_devoluciones,
       COALESCE(r.suma_devoluciones,0) AS suma_devoluciones
FROM orders AS o
LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
LEFT JOIN (
    SELECT order_id,
           ROUND(SUM(amount),2) AS suma_devoluciones,
           COUNT(*) AS conteo_devoluciones
    FROM refunds
    GROUP BY order_id
) AS r ON o.order_id = r.order_id;
```

```
79 • SELECT o.*,
80     m.merchant_id AS merch_id,
81     m.name AS nombre_comercio,
82     COALESCE(r.conteo_devoluciones,0) AS conteo_devoluciones,
83     COALESCE(r.suma_devoluciones,0) AS suma_devoluciones
84 FROM orders AS o
85 LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
86 LEFT JOIN (
87     SELECT order_id,
88         ROUND(SUM(amount),2) AS suma_devoluciones,
89         COUNT(*) AS conteo_devoluciones
90     FROM refunds
91     GROUP BY order_id
92 ) AS r ON o.order_id = r.order_id;
93
```

order_id	created_at	status	amount	merchant_id	country	merch_id	nombre_comercio	conteo_devoluciones	suma_devoluciones
5c3ef8170aee697c1ba84361	2015-08-17 05:55:44	ACTIVE	178.79	pk_317b4fc6fd80a5f8fb2ff216	Espana	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3ef8170aee697c1ba84362	2015-08-17 09:16:55	ACTIVE	141.76	pk_317b4fc6fd80a5f8fb2ff216	Italia	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3ef8170aee697c1ba84363	2015-08-17 18:25:23	ACTIVE	263.57	pk_317b4fc6fd80a5f8fb2ff216	Espana	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3c56224490281cfd202df0	2015-12-04 22:22:23	ACTIVE	299	pk_07225590b8fea17e739aa451	Espana	pk_07225590b8fea17e739aa451	Netflix	0	0
5c347803c3d42e57ea5fe6cb	2015-12-04 23:17:45	ACTIVE	184.99	pk_a3aa2fa07c5436f4c8ca1e03	Alemania	pk_a3aa2fa07c5436f4c8ca1e03	fnac	0	0
5c3f38c30aee693b32b834c9	2015-12-05 08:40:37	ACTIVE	13.94	pk_d9b9215223a9d14515ae0b42	Espana	pk_d9b9215223a9d14515ae0b42	Ryanair	0	0
5c347803c3d42e57ea5fe6cc	2015-12-05 09:46:28	ACTIVE	165.98	pk_a3aa2fa07c5436f4c8ca1e03	Italia	pk_a3aa2fa07c5436f4c8ca1e03	fnac	0	0
5c3ef8170aee697c1ba8440b	2015-12-05 15:04:08	ACTIVE	778.85	pk_317b4fc6fd80a5f8fb2ff216	Espana	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c36ff3a3e78fa60e0ad6baf	2015-12-06 08:38:17	ACTIVE	159	pk_736c7094ea96eda38b098f56	Francia	pk_736c7094ea96eda38b098f56	Massimo Dutti	0	0
5c3ef8170aee697c1ba8440d	2015-12-06 11:46:56	ACTIVE	202.82	pk_317b4fc6fd80a5f8fb2ff216	Belgica	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3c56224490281cfd202df1	2015-12-06 15:41:05	ACTIVE	230	pk_07225590b8fea17e739aa451	Francia	pk_07225590b8fea17e739aa451	Netflix	0	0
5c3ef3f70aee697c1ba7e936	2015-12-06 16:15:07	ACTIVE	251.03	pk_c15afcbd3a31b732f097ba7b	Francia	pk_c15afcbd3a31b732f097ba7b	Havainas	0	0
5c347803c3d42e57ea5fe6cd	2015-12-06 16:51:03	ACTIVE	199.99	pk_a3aa2fa07c5436f4c8ca1e03	Italia	pk_a3aa2fa07c5436f4c8ca1e03	fnac	0	0

32 12:18:18 SELECT o.\*, m.merchant\_id AS merch\_id, m.name AS nombre\_comercio, COALESCE(r.conteo\_devoluciones,0) AS conteo\_devoluciones, ... 873 row(s) returned

CREATE VIEW orders\_view AS

```
SELECT o.*,
       m.merchant_id AS merch_id,
       m.name AS nombre_comercio,
       COALESCE(r.conteo_devoluciones,0) AS conteo_devoluciones,
       COALESCE(r.suma_devoluciones,0) AS suma_devoluciones
FROM orders AS o
LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
LEFT JOIN (
    SELECT order_id,
           ROUND(SUM(amount),2) AS suma_devoluciones,
           COUNT(*) AS conteo_devoluciones
    FROM refunds
    GROUP BY order_id
) AS r ON o.order_id = r.order_id;
```

SELECT \* FROM prestamos\_2015.orders\_view;

```
96 SELECT o.*,
97     m.merchant_id AS merch_id,
98     m.name AS nombre_comercio,
99     COALESCE(r.conteo_devoluciones,0) AS conteo_devoluciones,
100    COALESCE(r.suma_devoluciones,0) AS suma_devoluciones
101 FROM orders AS o
102 LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
103 LEFT JOIN (
104     SELECT order_id,
105            ROUND(SUM(amount),2) AS suma_devoluciones,
106            COUNT(*) AS conteo_devoluciones
107     FROM refunds
108     GROUP BY order_id
109 ) AS r ON o.order_id = r.order_id;
110
111 SELECT * FROM prestamos_2015.orders_view;
```

order_id	created_at	status	amount	merchant_id	country	merch_id	nombre_comercio	conteo_devoluciones	suma_devoluciones
5c3ef8170aee697c1ba84361	2015-08-17 05:55:44	ACTIVE	178.79	pk_317b4fc6fd80a5f8fb2ff216	Espana	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3ef8170aee697c1ba84362	2015-08-17 09:16:55	ACTIVE	141.76	pk_317b4fc6fd80a5f8fb2ff216	Italia	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3ef8170aee697c1ba84363	2015-08-17 18:25:23	ACTIVE	263.57	pk_317b4fc6fd80a5f8fb2ff216	Espana	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3c56224490281cfd202df0	2015-12-04 22:22:23	ACTIVE	299	pk_07225590b8fea17e739aa451	Espana	pk_07225590b8fea17e739aa451	Netflix	0	0
5c347803c3d42e57ea5fe6db	2015-12-04 23:17:45	ACTIVE	184.99	pk_a3aa2fa07c5436f4c8ca1e03	Alemania	pk_a3aa2fa07c5436f4c8ca1e03	fnac	0	0
5c3f38c30aee693b32b834c9	2015-12-05 08:40:37	ACTIVE	13.94	pk_d9b9215223a9d14515ae0b42	Espana	pk_d9b9215223a9d14515ae0b42	Ryanair	0	0
5c347803c3d42e57ea5fe6cc	2015-12-05 09:46:28	ACTIVE	165.98	pk_a3aa2fa07c5436f4c8ca1e03	Italia	pk_a3aa2fa07c5436f4c8ca1e03	fnac	0	0
5c3ef8170aee697c1ba8440b	2015-12-05 15:04:08	ACTIVE	778.85	pk_317b4fc6fd80a5f8fb2ff216	Espana	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c36ff3a3e78fa60eadbaf	2015-12-06 08:38:17	ACTIVE	159	pk_736c7094ea96eda38b098f56	Francia	pk_736c7094ea96eda38b098f56	Massimo Dutti	0	0
5c3ef8170aee697c1ba8440d	2015-12-06 11:46:56	ACTIVE	202.82	pk_317b4fc6fd80a5f8fb2ff216	Belgica	pk_317b4fc6fd80a5f8fb2ff216	Calcedonia	0	0
5c3c56224490281cfd202df1	2015-12-06 15:41:05	ACTIVE	230	pk_07225590b8fea17e739aa451	Francia	pk_07225590b8fea17e739aa451	Netflix	0	0
5c3ef3f70aee697c1ba7e936	2015-12-06 16:15:07	ACTIVE	251.03	pk_c15afcbd3a31b732f097ba7b	Francia	pk_c15afcbd3a31b732f097ba7b	Havainas	0	0
5c347803c3d42e57ea5fe6cd	2015-12-06 16:51:03	ACTIVE	199.99	pk_a3aa2fa07c5436f4c8ca1e03	Italia	pk_a3aa2fa07c5436f4c8ca1e03	fnac	0	0

## Conclusiones:

En la consulta, se observa un total de 25 devoluciones con un importe acumulado de 6.312,22€. Dentro de los países analizados:

- España tiene la mayor suma total de devoluciones (1.955,49€), a pesar de tener el mismo número de devoluciones que Francia (6). Esto indica que, en promedio, las devoluciones en España son de mayor valor que en otros países.
- Portugal solo registra una devolución, pero el importe total es significativo (773,14€). Esto sugiere que esa única devolución tuvo un impacto financiero importante.

- Marruecos tiene el menor importe total de devoluciones (650,32€), a pesar de contar con 4 devoluciones. Esto indica que las devoluciones en Marruecos tienen un valor relativamente bajo en comparación con otros países.
- Bélgica no registra devoluciones, lo que puede ser una señal positiva en términos de satisfacción del cliente o eficiencia en la gestión de devoluciones por parte de los comercios en ese país.

Con respecto a los comercios:

- Calzedonia se presenta como el líder en las devoluciones (4.131,33€), lo que sugiere que este comercio tiene una alta incidencia de devoluciones y/o que las devoluciones que realiza son de mayor valor en promedio.
- Kindle, en cambio, cuenta con tan solo una devolución en Italia (124,46€), siendo así el comercio con la menor suma de devoluciones.

## Ejercicio 4

Con los datos que tienes a tu disposición, tu tarea es diseñar un análisis o funcionalidad que te permita descubrir un aspecto interesante del caso de estudio.

Primero, deberás definir el objetivo de tu análisis en un texto breve. A continuación, desarrolla las consultas necesarias para llevar a cabo tu análisis. Finalmente, reflexiona sobre los resultados obtenidos.

Puedes utilizar cualquier herramienta o recurso que hayas aprendido en clase. Algunas ideas para tu funcionalidad podrían ser: segmentar a los clientes según el valor de sus operaciones, establecer un sistema de alertas para operaciones sospechosas, identificar patrones estacionales, etc. Eres libre de desarrollar la funcionalidad que prefieras, lo importante es que refleje tu enfoque personal y cumpla los objetivos de análisis.

**OBJETIVOS DE ANÁLISIS:** En la documentación encontrarás un ejemplo de funcionalidad que puede servirte de guía. Para la evaluación de este ejercicio, se valorará una propuesta bien fundamentada en los datos, un análisis exploratorio que respalde el objetivo del análisis elegido, el uso del lenguaje SQL para mostrar los recursos estudiados durante el módulo (pueden ser una o varias consultas, a tu elección) y, finalmente, una conclusión bien elaborada basada en los resultados obtenidos que aporte decisiones basadas en los resultados.

### 1. Objetivo de análisis

Detectar posibles casos de fraude en las operaciones financieras mediante la identificación de patrones.

Gracias a este análisis podremos:

- Proteger a los clientes del fraude.
- Reducir las pérdidas por fraude.
- Mejorar la seguridad del sistema de pagos y fortalecer la capacidad de las organizaciones para detectar y prevenir fraudes en sus operaciones financieras.

### 2. Análisis

Primero, analizaremos las transacciones con importes significativamente altos o bajos en comparación con el promedio de transacciones de cada país. Esto puede ayudar a detectar patrones de fraude específicos en cada país.

```
SELECT o.country,
       o.order_id,
       o.amount,
       o_avg.promedio_pais AS promedio_pais,
       CASE
         WHEN o.amount > o_avg.promedio_pais * 2 THEN 'Importe alto'
         WHEN o.amount < o_avg.promedio_pais / 2 THEN 'Importe bajo'
         ELSE 'Normal'
       END AS tipo_transaccion
FROM orders AS o
LEFT JOIN (
  SELECT country,
         ROUND(AVG(amount),2) AS promedio_pais
```



```

FROM orders
GROUP BY country
) AS o_avg ON o.country = o_avg.country
GROUP BY o.country, o.order_id, o.amount, o_avg.promedio_pais
HAVING tipo_transaccion != 'Normal';

```

```

121 • SELECT o.country,
122         o.order_id,
123         o.amount,
124         o_avg.promedio_pais AS promedio_pais,
125         CASE
126             WHEN o.amount > o_avg.promedio_pais * 2 THEN 'Importe alto'
127             WHEN o.amount < o_avg.promedio_pais / 2 THEN 'Importe bajo'
128             ELSE 'Normal'
129         END AS tipo_transaccion
130     FROM orders AS o
131     LEFT JOIN (
132         SELECT country,
133                ROUND(AVG(amount),2) AS promedio_pais
134         FROM orders
135         GROUP BY country
136     ) AS o_avg ON o.country = o_avg.country
137     GROUP BY o.country, o.order_id, o.amount, o_avg.promedio_pais
138     HAVING tipo_transaccion != 'Normal';

```

country	order_id	amount	promedio_pais	tipo_transaccion
Espana	5c3ef8170aee697c1ba84361	178.79	390.79	Importe bajo
Espana	5c3f38c30aee693b32b834c9	13.94	390.79	Importe bajo
Francia	5c36ff3a3e78fa60e0ad6baf	159	377.59	Importe bajo
Belgica	5c3ef8170aee697c1ba8440d	202.82	411.97	Importe bajo
Espana	5c3c4cc14490281cfd1fbaca	129.85	390.79	Importe bajo
Espana	5c370dfc3e78fa6aacf5cb27	1295.95	390.79	Importe alto
Espana	5c347803c3d42e57ea5fe6cf	810.99	390.79	Importe alto
Espana	5c3ef8170aee697c1ba84412	165.65	390.79	Importe bajo
Espana	5c36ff3a3e78fa60e0ad6bb0	160.2	390.79	Importe bajo
Marruecos	5c1a2e1a2dea1c29c61969f9	156.8	383.25	Importe bajo

35 22:40:47 SELECT o.country, o.order\_id, o.amount, o\_avg.promedio\_pais AS promedio\_pais, CASE WHE... 286 row(s) returned

Por otro lado, realizaremos la misma consulta, pero comparando los importes de las operaciones con el promedio de las transacciones de cada comercio. La presencia de transacciones consideradas de 'importe alto' o 'bajo' en comparación con el promedio del comercio también puede ser una señal de alerta para posibles fraudes.

```

SELECT m.name AS nombre_comercio,
       o.order_id,
       o.amount,
       o_avg.promedio_comercio AS promedio_comercio,
       CASE
           WHEN o.amount > o_avg.promedio_comercio * 2 THEN 'Importe alto'
           WHEN o.amount < o_avg.promedio_comercio / 2 THEN 'Importe bajo'
           ELSE 'Normal'
       END AS tipo_transaccion
FROM orders AS o
LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
LEFT JOIN (

```



```

SELECT merchant_id,
ROUND(AVG(amount),2) AS promedio_comercio
FROM orders
GROUP BY merchant_id
) AS o_avg ON o.merchant_id = o_avg.merchant_id
GROUP BY m.name, o.order_id, o.amount, o_avg.promedio_comercio
HAVING tipo_transaccion != 'Normal';

```

```

29 • SELECT m.name AS nombre_comercio,
30     o.country,
31     o.order_id,
32     o.amount,
33     o_avg.promedio_comercio AS promedio_comercio,
34     CASE
35         WHEN o.amount > o_avg.promedio_comercio * 2 THEN 'Importe alto'
36         WHEN o.amount < o_avg.promedio_comercio / 2 THEN 'Importe bajo'
37         ELSE 'Normal'
38     END AS tipo_transaccion
39 FROM orders AS o
40 LEFT JOIN merchants AS m ON o.merchant_id = m.merchant_id
41 LEFT JOIN (
42     SELECT merchant_id,
43         ROUND(AVG(amount),2) AS promedio_comercio
44     FROM orders
45     GROUP BY merchant_id
46 ) AS o_avg ON o.merchant_id = o_avg.merchant_id
47 GROUP BY m.name, o.country, o.order_id, o.amount, o_avg.promedio_comercio
48 HAVING tipo_transaccion != 'Normal';

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

	nombre_comercio	country	order_id	amount	promedio_comercio	tipo_transaccion
▶	Calcedonia	Italia	5c3ef8170aee697c1ba84362	141.76	298.92	Importe bajo
	fnac	Alemania	5c347803c3d42e57ea5fe6cb	184.99	582.02	Importe bajo
	fnac	Italia	5c347803c3d42e57ea5fe6cc	165.98	582.02	Importe bajo
	Calcedonia	Espana	5c3ef8170aee697c1ba8440b	778.85	298.92	Importe alto
	fnac	Italia	5c347803c3d42e57ea5fe6cd	199.99	582.02	Importe bajo
	Spotify	Espana	5c370dfc3e78fa6aacf5cb27	1295.95	615.86	Importe alto
	fnac	Alemania	5c347803c3d42e57ea5fe6ce	199.99	582.02	Importe bajo

40 23:01:24 SELECT m.name AS nombre\_comercio, o.order\_id, o.amount, o\_avg.promedio\_comercio AS promedio\_... 242 row(s) returned

## Conclusiones:

Análisis por países:

- Según la consulta realizada, existen 206 transacciones que se encuentran por debajo del promedio de sus respectivos países y 80 operaciones que se encuentran por encima. Sería conveniente enfocarnos, sobre todo, en aquellas transacciones muy por encima del promedio, ya que podrían ser intentos de fraude mediante compras de alto valor.
- España tiene la mayor cantidad de operaciones tanto de importe bajo (105 operaciones) como de importe alto (44 operaciones), en comparación con el promedio del país. Este país muestra una diversidad significativa de operaciones, lo que indica una amplia gama de comportamientos financieros.

- En cambio, Portugal presenta una distribución más equilibrada, con pocas operaciones tanto por debajo (3) como por encima (1) del promedio del país.
- Francia e Italia muestran un patrón similar entre sí, con más operaciones consideradas de importe bajo (prácticamente el triple) que de importe alto.

Con respecto a los comercios:

- Existen 173 transacciones que se encuentran por debajo del promedio de sus respectivos comercios, y 69 operaciones que se encuentran por encima del promedio.
- Calzedonia es el comercio con más cantidad de operaciones consideradas de ‘importe alto’, mientras que Spotify lidera (en términos de valor) tanto en operaciones de importe alto como bajo, con valores significativos en ambos extremos. Esto puede indicar actividades fraudulentas a través de compras masivas de artículos de gran valor o intentos de aprovechar promociones u ofertas de forma indebida.
- Carhart es el único comercio de la consulta que tan solo presenta una transacción considerada de ‘importe bajo’ y otra de ‘importe alto’, lo que muestra una variedad limitada de transacciones extremas. Esto nos puede indicar que dicho comercio no es tan susceptible al fraude como otros, aunque sería recomendable monitorear cualquier cambio en su patrón de transacciones.
- En general, las transacciones atípicas en términos de importe (alto o bajo en comparación con el promedio) se concentran en España, donde destaca Spotify como un caso notable tanto en operaciones de importe alto como bajo. Todo ello señala que este país y este comercio en particular podrían ser objetivos frecuentes de actividades fraudulentas.

Dada la diversidad de comportamientos financieros entre países y comercios, sería conveniente implementar estrategias específicas adaptadas a cada contexto, poniendo el foco principal en la prevención de fraudes y el análisis de riesgos financieros.

### 3. Estrategias y acciones posteriores

Basándonos en el análisis realizado sobre las transacciones financieras y la detección de posibles fraudes, es importante considerar algunas estrategias y acciones que podrían ser útiles para su prevención:

- **Controles de seguridad adicionales y sistemas de detección de fraudes avanzados**

Implementar controles de seguridad adicionales para validar estas transacciones, por ejemplo, que se requiera la verificación de identidad y autenticación para transacciones que se consideren de alto riesgo según los patrones identificados; y sistemas de detección de fraudes avanzados que integren inteligencia artificial (IA) y *machine learning* para identificar patrones y comportamientos fraudulentos de forma más precisa y en tiempo real.

- **Técnicas de análisis predictivo y alertas automáticas**

Relacionado con la IA y *machine learning*, utilizar técnicas de análisis predictivo para detectar posibles fraudes en etapas iniciales y así poder tomar medidas preventivas antes de que se produzcan pérdidas significativas. Asimismo, configurar alertas automáticas (basadas patrones predefinidos o comportamientos sospechosos) para reducir al máximo el tiempo de respuesta ante un posible fraude sobre transacciones destacables de importe alto o bajo en comparación con el promedio del país o del comercio.

- **Feedback y revisiones para una mejora continua**

Recopilar el *feedback* de las acciones llevadas a cabo, además de realizar auditorías internas y revisiones periódicas para identificar posibles puntos débiles o áreas de mejora de las estrategias implementadas, con el objetivo de realizar mejoras continuas en base a los resultados obtenidos; así como actualizar regularmente las políticas y estrategias de prevención de fraudes en función de la evolución del comportamiento fraudulento y las nuevas tecnologías disponibles. Por ejemplo, en el caso de identificar transacciones con importes inusualmente altos, como en España, se deberían revisar las políticas de devolución y reembolso para asegurarse de que sean robustas y que eviten posibles abusos o fraudes en este aspecto.

Esta revisión constante y adaptación de políticas asegura una respuesta efectiva y proactiva ante las tendencias emergentes en actividades fraudulentas.

- **Análisis geográfico y de tendencias**

Utilizar análisis geográficos para identificar determinadas áreas geográficas donde las actividades fraudulentas puedan ser más comunes. Por ejemplo, si España (según el análisis realizado) concentra un gran número de transacciones atípicas, se deberían llevar a cabo análisis más detallados en dicha región para detectar posibles patrones de fraude.

Por otro lado, como se comentó con Carhart, sería necesario analizar las tendencias (tanto de comercios como de países) para poder identificar cambios o picos inusuales en las transacciones consideradas de ‘importe alto’ o ‘bajo’, pues ello podría indicar actividades fraudulentas temporales.

- **Monitoreo continuo y especializado**

En relación con el punto anterior, sería importante establecer un sistema de monitoreo continuo y especializado para detectar actividades fraudulentas de manera efectiva. Esto se traduce en implementar, por ejemplo, algoritmos avanzados de detección de anomalías que sean capaces de identificar transacciones que se desvíen significativamente del comportamiento histórico o del promedio esperado. Este monitoreo también debe ser capaz de identificar cambios bruscos en los hábitos de transacción de los usuarios que puedan indicar actividades fraudulentas, como accesos desde ubicaciones inusuales o patrones de compra atípicos.

Asimismo, sería interesante implementar un monitoreo especializado para países. Por ejemplo, dada la diversidad significativa de operaciones tanto de importe bajo como alto en España, se recomendaría establecer un monitoreo más frecuente y detallado para este país, con el fin de detectar posibles actividades fraudulentas en tiempo real.

Estas estrategias y acciones combinadas pueden ayudar a combatir eficazmente el fraude en transacciones financieras, mediante la detección de patrones inusuales con la ayuda de algoritmos avanzados, el monitoreo continuo y especializado de transacciones, la revisión periódica para mejorar, y la actualización regular de políticas antifraude. Además, como hemos visto, el seguimiento detallado de transacciones atípicas por país y comercio es clave.

Al proporcionar un enfoque integral, las organizaciones pueden fortalecer sus defensas contra actividades fraudulentas en un entorno en constante evolución. Sin embargo, es esencial adaptar estas medidas a las necesidades específicas de cada empresa y mantener una revisión constante de las estrategias antifraude.

