



UNIVERSIDAD
COMPLUTENSE
MADRID



Bases de Datos con SQL: TAREA FINAL

Autor: Gabriela Rivero Di Leonardo

Profesor: Juan Fernando Sánchez Martínez

Índice

Ejercicios bases de datos	3
Ejercicio 1	3
Modelo de Entidad-Relación	3
Modelo Lógico	4
Modelo Físico	4
Ejercicio 2	6
Respuesta:	6
Respuesta:	7
Ejercicio 3	8
Respuesta:	8
Respuesta:	9
Ejercicio 4	11
Estudio del Ratio de Operaciones Problemáticas.....	11

Ejercicios bases de datos

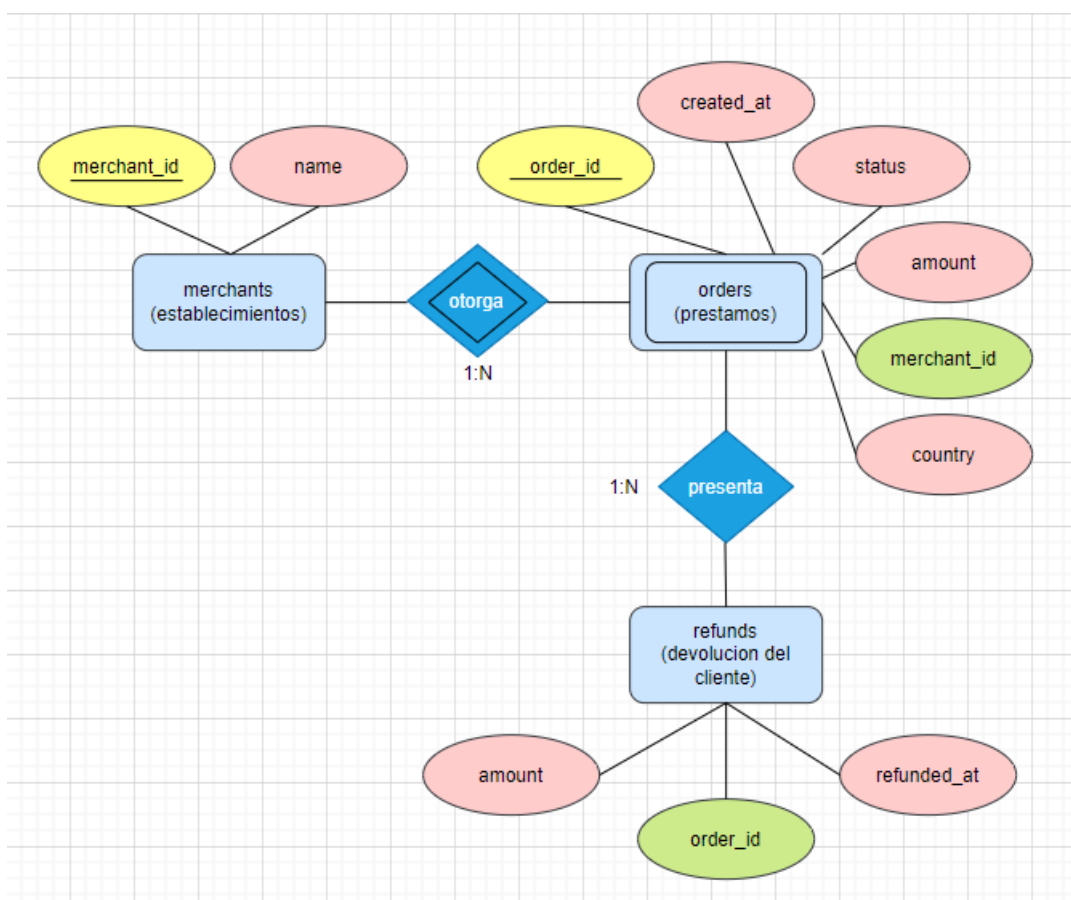
A continuación, se presentan los ejercicios a resolver para la evaluación del módulo de Bases de Datos SQL que desarrollaremos con las siguientes tablas:

- Orders
- Refunds
- Merchants

Ejercicio 1

Desarrollar un **Modelo de Entidad-Relación** indicando las entidades, atributos, relaciones y cardinalidad entre las tablas.

Modelo de Entidad-Relación



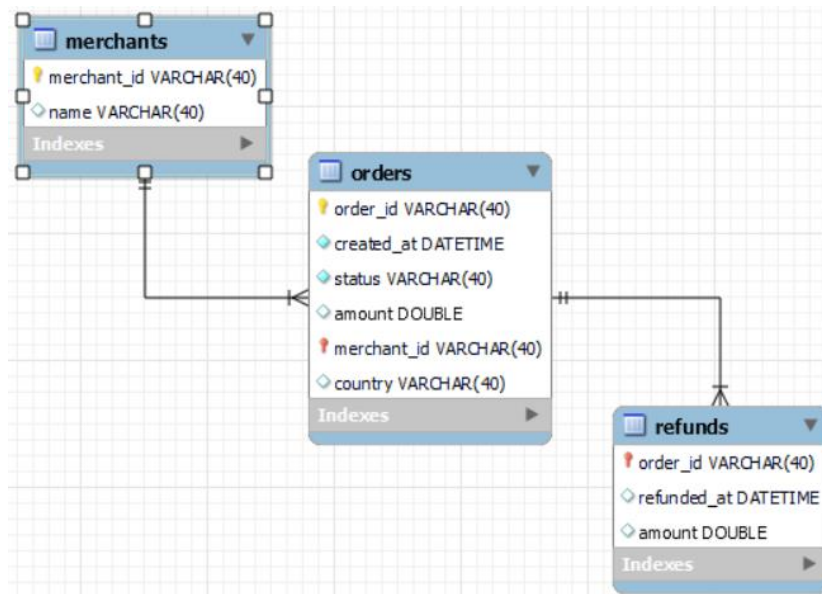
FK: Verde / PK: Amarillo / Entidades: Azul claro / Atributos: Rosa / Relación: Azul oscuro

Los préstamos dependen de un establecimiento para poderse emitir, por eso es una entidad débil.

Relación 1:N de merchant/orders: un establecimiento puede conceder muchos préstamos, pero los préstamos son concedidos solo por un establecimiento.

Relación 1:N de orders/refunds: Cada devolución está asociada a un order en específico, pero un order puede tener varios refunds. Si se trata de préstamos, el cliente, por ejemplo, podría realizar devoluciones parciales del préstamo.

Modelo Lógico



Modelo Físico

Tabla Merchants:

```
CREATE TABLE IF NOT EXISTS `tarea_prestamos`.`merchants` (  
  `merchant_id` VARCHAR(40) NOT NULL,  
  `name` VARCHAR(40) NULL,  
  PRIMARY KEY (`merchant_id`))  
ENGINE = InnoDB
```

Tabla Orders:

```
CREATE TABLE IF NOT EXISTS `tarea_prestamos`.`orders` (  
  `order_id` VARCHAR(40) NOT NULL,  
  `created_at` DATETIME NOT NULL,  
  `status` VARCHAR(40) NOT NULL,  
  `amount` DOUBLE NULL,  
  `merchant_id` VARCHAR(40) NOT NULL,  
  `country` VARCHAR(40) NULL,  
  PRIMARY KEY (`order_id`, `merchant_id`),
```

```
INDEX `merchant_id_idx` (`merchant_id` ASC) VISIBLE,  
CONSTRAINT `merchant_id`  
  FOREIGN KEY (`merchant_id`)  
    REFERENCES `tarea_prestamos`.`merchants` (`merchant_id`)  
  ON DELETE NO ACTION  
  ON UPDATE NO ACTION)  
ENGINE = InnoDB
```

Tabla Refunds :

```
CREATE TABLE IF NOT EXISTS `tarea_prestamos`.`refunds` (  
  `order_id` VARCHAR(40) NOT NULL,  
  `refunded_at` DATETIME NULL,  
  `amount` DOUBLE NULL,  
  INDEX `order_id_idx` (`order_id` ASC) VISIBLE,  
  PRIMARY KEY (`order_id`),  
  CONSTRAINT `order_id`  
    FOREIGN KEY (`order_id`)  
      REFERENCES `tarea_prestamos`.`orders` (`order_id`)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION)  
ENGINE = InnoDB
```

Ejercicio 2

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 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.

Respuesta:

```
SELECT `country` AS 'País', `status` AS Estado, COUNT(`order_id`) AS Total_operaciones,
      ROUND(AVG(`amount`),2) AS Importe_promedio
FROM `orders`
WHERE `created_at` > 01-07-2015
AND `country` IN ('Francia', 'Portugal', 'España')
AND `amount` > 100
AND `amount` < 1500
GROUP BY `country`, `status`
ORDER BY Importe_promedio DESC;
```

```
3 • SELECT `country` AS 'País', `status` AS Estado, COUNT(`order_id`) AS Total_operaciones, ROUND(AVG(`amount`),2) AS Importe_promedio
4 FROM `orders`
5 WHERE `created_at` > 01-07-2015
6     AND `country` IN ('Francia', 'Portugal', 'España')
7     AND `amount` > 100
8     AND `amount` < 1500
9 GROUP BY `country`, `status`
```

Result Grid				
Filter Rows:		Export:	Wrap Cell Content:	
País	Estado	Total_operaciones	Importe_promedio	
Portugal	CLOSED	7	522.38	
España	DELINQUENT	21	423.48	
España	ACTIVE	172	410.67	
Portugal	ACTIVE	5	392.98	
Francia	CLOSED	44	391.22	
Francia	ACTIVE	93	347.3	
España	CLOSED	160	344.96	
Francia	DELINQUENT	12	336.53	
Francia	CANCELLED	10	322.34	
España	CANCELLED	21	296.72	

Result 11 x

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

Respuesta:

```
SELECT `country` AS 'País', COUNT(`order_id`) AS Total_operaciones, MAX(`amount`) AS Valor_max, MIN(`amount`) AS Valor_min
FROM `orders`
WHERE NOT `status` IN ("Delinquent","Cancelled")
AND `amount` > 100
GROUP BY `country`
ORDER BY Total_operaciones DESC
LIMIT 3;
```

```
12 #Pregunta 2
13 • SELECT `country` AS 'País', COUNT(`order_id`) AS Total_operaciones, MAX(`amount`) AS Valor_max, MIN(`amount`) AS Valor_min
14 FROM `orders`
15 WHERE NOT `status` IN ("Delinquent","Cancelled")
16     AND `amount` > 100
17 GROUP BY `country`
18 ORDER BY Total_operaciones DESC
19 LIMIT 3;
20
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:	Fetch rows:
País	Total_operaciones	Valor_max	Valor_min	
España	342	2960.87	101	
Francia	141	1863.98	100.88	
Italia	75	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.

Respuesta:

```
SELECT
    m.`merchant_id` AS ID,
    `name` AS Nombre,
    `country` AS 'País',
    COUNT(o.`order_id`) AS Total_operaciones,
    ROUND(AVG(o.`amount`),2) AS Importe_promedio,
    COUNT(`refunded_at`) AS Total_devoluciones,
    CASE
        WHEN COUNT(`refunded_at`) > 0 THEN "Sí"
        ELSE "NO"
    END AS Acepta_devoluciones
FROM
    `merchants` AS m
INNER JOIN `orders` AS o ON m.merchant_id = o.merchant_id
LEFT JOIN `refunds` AS r ON o.order_id = r.order_id
WHERE `country` IN ('Marruecos','Italia','España','Portugal')
GROUP BY country, m.`merchant_id`, `name`
HAVING Total_operaciones>10
ORDER BY Total_operaciones ASC;
```



```

23 • SELECT
24     m.`merchant_id` AS ID,
25     `name` AS Nombre,
26     `country` AS 'País',
27     COUNT(o.`order_id`) AS Total_operaciones,
28     ROUND(AVG(o.`amount`),2) AS Importe_promedio,
29     COUNT(`refunded_at`) AS Total_devoluciones,
30     CASE
31         WHEN COUNT(`refunded_at`) > 0 THEN "Sí"
32         ELSE "NO"
33     END AS Acepta_devoluciones
34 FROM
35     `merchants` AS m
36     INNER JOIN `orders` AS o ON m.merchant_id = o.merchant_id
37     LEFT JOIN `refunds` AS r ON o.order_id = r.order_id
38 WHERE `country` IN ('Marruecos','Italia','España', 'Portugal')
39 GROUP BY country, m.`merchant_id`, `name`
40 HAVING Total_operaciones>10
41 ORDER BY Total_operaciones ASC;

```

ID	Nombre	País	Total_operaciones	Importe_promedio	Total_devoluciones	Acepta_devoluciones
pk_743...	Pepe Jeans	España	11	171.99	0	NO
pk_317...	Calcedonia	Marruecos	13	365.36	3	Sí
pk_736...	Massimo Dutti	España	13	169.88	0	NO
pk_c15...	Havainas	España	16	323.02	0	NO
pk_072...	Netflix	España	21	363.57	0	NO
pk_a3a...	fnac	España	22	531.84	0	NO
pk_c44...	YouTube music	España	25	669.28	1	Sí
pk_317...	Calcedonia	Italia	26	229.3	2	Sí
pk_b9e...	K-tuin	España	46	373.36	0	NO
pk_19d...	Spotify	España	64	644.26	0	NO
pk_317...	Calcedonia	España	137	314.76	5	Sí

- 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.

Respuesta:

```

CREATE VIEW orders_view AS

SELECT m.`merchant_id` AS Comercio_ID, `name` AS Comercio,
       o.`order_id`, `created_at`, `status`, o.`amount`, `country`,
       COUNT(r.`order_id`) AS Numero_devoluciones, SUM(r.`amount`) AS Suma_devoluciones
FROM
    `merchants` AS m

```

INNER JOIN `orders` AS o ON m.merchant_id = o.merchant_id

LEFT JOIN `refunds` AS r ON o.order_id = r.order_id

GROUP BY m.merchant_id, o.order_id, name, created_at, status, amount, country;

```
44 • CREATE VIEW orders_view AS
45 SELECT m.`merchant_id` AS Comercio_ID, `name` AS Comercio,
46        o.`order_id`, `created_at`, `status`, o.`amount`, `country`,
47        COUNT(r.`order_id`) AS Numero_devoluciones, SUM(r.`amount`) AS Suma_devoluciones
48 FROM
49     `merchants` AS m
50     INNER JOIN `orders` AS o ON m.merchant_id = o.merchant_id
51     LEFT JOIN `refunds` AS r ON o.order_id = r.order_id
52 GROUP BY m.merchant_id, o.order_id, name, created_at, status, amount, country;
```

The screenshot shows a database management interface. On the left, a 'SCHEMAS' panel displays a tree view of the database structure, including tables like 'merchants', 'orders', and 'refunds', and a view named 'orders_view'. The 'orders_view' view is selected, and its columns are listed: Comercio_ID, Comercio, order_id, created_at, status, amount, country, Numero_devoluciones, and Suma_devoluciones. The main area shows the SQL query for the view: 'SELECT * FROM tarea_prestamos.orders_view;'. Below the query, a 'Result Grid' displays the data. The grid has 10 columns corresponding to the view's fields. The data is filtered to show 10 rows. The first row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The second row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The third row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The fourth row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The fifth row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The sixth row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The seventh row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The eighth row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The ninth row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'. The tenth row shows a merchant named 'Calcedonia' with order_id '5c3ef8170a...' and status 'CLOSED'.

Comercio_ID	Comercio	order_id	created_at	status	amount	country	Numero_devoluciones	Suma_devoluciones
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-17 16:5...	CLOSED	163.08	Francia	2	163.07999999999998
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-21 11:2...	CLOSED	773.14	Portugal	1	773.14
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-23 16:5...	CLOSED	191.05	Alemania	2	191.05
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-23 17:3...	CLOSED	235.53	Alemania	1	235.53
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-24 14:0...	CLOSED	302.06	Alemania	1	302.06
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-25 08:4...	CLOSED	282.72	Marruecos	2	282.72
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-25 17:2...	CLOSED	440.63	España	1	440.63
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-26 15:0...	CLOSED	194.22	Francia	1	194.22
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-26 15:5...	CLOSED	214.24	España	1	214.24
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-26 16:0...	DELINQUENT	217.36	España	2	217.36
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-26 19:5...	CLOSED	261.28	Italia	1	261.28
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-27 09:4...	CLOSED	266.17	Italia	1	266.17
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-28 15:3...	CANCELLED	356.51	España	1	356.51
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-28 19:4...	CANCELLED	3.11	Alemania	1	3.11
pk_317b4fc6fd80a5f...	Calcedonia	5c3ef8170a...	2015-07-29 17:1...	CLOSED	230.23	Marruecos	1	230.23

Ejercicio 4

A partir de los datos disponibles diseñar una **funcionalidad** a tu elección que permita obtener un insight de interés sobre el caso de uso estudiado.

Para ello debes plantear primeramente en un breve texto el objetivo de tu funcionalidad, la queries desarrollada y una reflexión sobre el insight obtenido. Para ello puedes usar cualquier recurso estudiado en clase.

Algunos ejemplos de funcionalidad podría ser: segmentación de clientes en función del valor de las operaciones, sistema de alertas para operaciones delictivas, identificación de estacionalidad, etc.. Tienes libertad total para desarrollar tu funcionalidad, lo importante es que tenga tu sello personal

Estudio del Ratio de Operaciones Problemáticas

El propósito de esta funcionalidad es identificar países y, posteriormente, comercios, con una alta proporción de operaciones problemáticas, que serían las Delinquent o Cancelled. Así asignar una especie de “Protocolo de Seguridad” en función de su proporción, lo que permitiría priorizar la atención y medidas de seguridad en las áreas de mayor riesgo.

En una primera fase, identificaremos los países con una alta proporción de operaciones problemáticas, eligiendo los tres que mayor ratio poseen.

```
SELECT
    `country`,
    SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
        END) AS Total_Delinquent_Cancelled,
    COUNT(*) AS Total_Operaciones,
    (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
        END) / COUNT(*)) AS 'Proporción'
FROM `orders`
GROUP BY `country`
ORDER BY (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
    END) / COUNT(*)) DESC
LIMIT 3;
```

```
55 • SELECT
56     `country`,
57     SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
58         END) AS Total_Delinquent_Cancelled,
59     COUNT(*) AS Total_Operaciones,
60     (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
61         END) / COUNT(*)) AS 'Proporción'
62 FROM `orders`
63 GROUP BY `country`
64 ORDER BY (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
65     END) / COUNT(*)) DESC
66 LIMIT 3;
```

Result Grid			
Filter Rows: <input type="text"/>			
Export:			
Wrap Cell Content:			
Fetch rows:			
country	Total_Delinquent_Cancelled	Total_Operaciones	Proporción
Marruecos	10	42	0.2381
Francia	27	175	0.1543
España	57	420	0.1357

A continuación, nos enfocaremos en estos tres países para determinar el protocolo de seguridad a aplicar. Los comercios que obtengan una proporción mayor o igual a 20% se les deberá prestar mayor atención.

```

SELECT
    `country`, name,
    SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
        END) AS Total_Delinquent_Cancelled,
    COUNT(*) AS Total_Operaciones,
    (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
        END) / COUNT(*)) AS 'Proporción',
    (CASE
        WHEN (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
            END) / COUNT(*)) >= 0.20 THEN "ALTO"
        ELSE "BAJO"
        END) AS "Protocolo_seguridad"
FROM `orders` AS o INNER JOIN `merchants` AS m ON m.`merchant_id`=o.`merchant_id`
WHERE `country` IN ('Marruecos','Francia','España')
GROUP BY `country`, name
HAVING Total_Delinquent_Cancelled >= 3
ORDER BY (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
    END) / COUNT(*)) DESC;

```

```

69 • SELECT
70     `country`, name,
71     SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
72         END) AS Total_Delinquent_Cancelled,
73     COUNT(*) AS Total_Operaciones,
74     (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
75         END) / COUNT(*)) AS 'Proporción',
76     (CASE
77         WHEN (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
78             END) / COUNT(*)) >= 0.20 THEN "ALTO"
79         ELSE "BAJO"
80         END) AS "Protocolo_seguridad"
81 FROM `orders` AS o INNER JOIN `merchants` AS m ON m.`merchant_id`=o.`merchant_id`
82 WHERE `country` IN ('Marruecos','Francia','España')
83 GROUP BY `country`, name
84 HAVING Total_Delinquent_Cancelled >= 3
85 ORDER BY (SUM(CASE WHEN `status` = 'Delinquent' OR `status` = 'Cancelled' THEN 1 ELSE 0
86     END) / COUNT(*)) DESC;

```

Result Grid						
Filter Rows: <input type="text"/> Export: <input type="button" value=""/> Wrap Cell Content: <input type="checkbox"/>						
country	name	Total_Delinquent_Cancelled	Total_Operaciones	Proporción	Protocolo_seguridad	
Francia	Speedo	3	7	0.4286	ALTO	
Marruecos	K-tuin	4	10	0.4000	ALTO	
España	Havainas	6	16	0.3750	ALTO	
Marruecos	YouTube music	3	9	0.3333	ALTO	
España	Massimo Dutti	3	13	0.2308	ALTO	

Los resultados obtenidos muestran que, en Francia el comercio “Speedo” presenta la mayor proporción, lo que justifica la aplicación de un protocolo de seguridad de alto nivel. En Marruecos, los comercios “K-tuin” y “YouTube music” requieren una atención especial y un protocolo de seguridad alto debido a su alta proporción de actividades problemáticas. Asimismo, en España, los comercios “Havainas” y “Massimo Dutti” también tienen proporciones significativas de operaciones problemáticas, lo que destaca la necesidad de elevar los protocolos de seguridad.

Este insight es valioso ya que, con esta información, se tomarán decisiones en términos de seguridad y gestión de riesgos, y con ello se podrán priorizar recursos y así reducir pérdidas y mejorar la seguridad del entorno empresarial.