Tasca S2.01. Nocions bàsiques SQL

Nivel 1

Ejercicio 1:

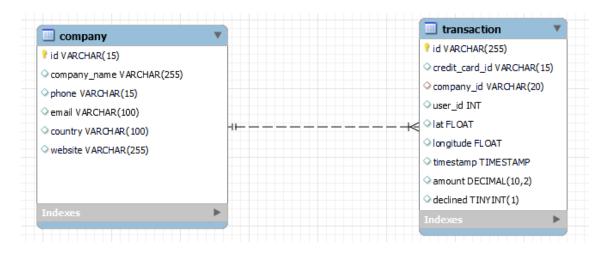
Características principales del esquema creado, explicación de las tablas y variables que existen, diagrama que ilustre la relación entre las diferentes tablas.

```
estructura_dades* × dades_introduir
Don't Limit
 1
       -- Creamos la base de datos
           CREATE DATABASE IF NOT EXISTS transactions;
 2 •
 3 •
           USE transactions;
 4
 5
           -- Creamos la tabla company
           CREATE TABLE IF NOT EXISTS company (
 7
              id VARCHAR(15) PRIMARY KEY,
 8
               company_name VARCHAR(255),
 9
               phone VARCHAR(15),
               email VARCHAR(100),
10
11
               country VARCHAR(100),
12
               website VARCHAR(255)
13
           );
15
16
           -- Creamos la tabla transaction
         CREATE TABLE IF NOT EXISTS transaction (
              id VARCHAR(255) PRIMARY KEY,
18
               credit_card_id VARCHAR(15) REFERENCES credit_card(id),
19
               company_id VARCHAR(20),
               user_id INT REFERENCES user(id),
21
               lat FLOAT,
22
              longitude FLOAT,
24
              timestamp TIMESTAMP,
25
               amount DECIMAL(10, 2),
26
               declined BOOLEAN,
               FOREIGN KEY (company_id) REFERENCES company(id)
27
```

Tenemos 2 tablas: company y transaction. En la tabla company, tenemos tipos de datos que se utilizan para almacenar cadenas de texto de tipo variable (VARCHAR), de ahí que usemos un paréntesis con la longitud que deseamos que tenga. Se utilizan para mejorar la eficiencia y evitar usar más espacio de lo necesario. Las variables incluyen una clave primaria (columna con valores únicos) "id", nombre de la compañía, teléfono, email, país y página web.

En la tabla transaction tenemos más variedad de tipos de datos. Incluyen varchar, variables numéricas ('int', 'float' y 'decimal'), referencias ('REFERENCES') que se utilizan para establecer relaciones entre tablas, variables Booleanas (True or False), variables que indican la

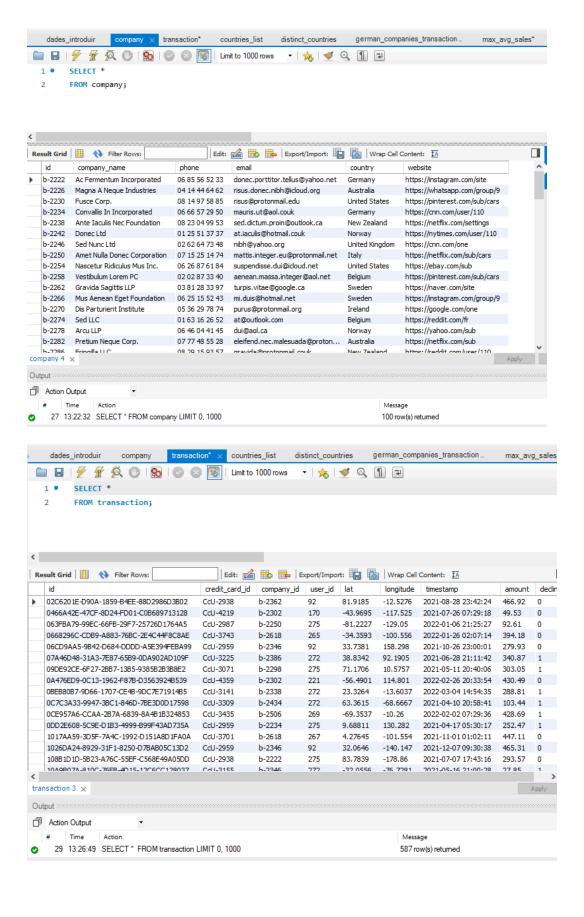
hora en que se realiza una transacción (Timestamp) y una clave foránea. Las variables son id de la tarjeta de crédito, id de la compañía, id del usuario, variables de localización geográfica (latitud y longitud), variable de fecha y hora ('timestamp'), importe de la transacción ('amount'), variable que indica si se ha declinado o no la transacción (declined), descrita con 0 (no se declina la transacción) y 1 (se declina) y la clave foránea id de la compañía.



Aquí podemos observar el diagrama de relación 1:N entre ambas tablas. Esta relación significa que 1 elemento de la tabla company está relacionado con varios elementos de la tabla transaction. Es decir, una compañía puede tener múltiples transacciones asociadas a ella. Esta relación se representa por la clave foránea "company_id" en la tabla "transaction" que referencia el campo o la columna "id" en la tabla "company".

```
dades_introduir × company transaction countries_list* distinct_countries* max_avg_sales*
1 •
       USE transactions;
       -- Insertamos datos de company
       INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2222', 'Ac Fermentum Incorp
 4 •
       INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2226', 'Magna A Neque Indus
                                                                                        'b-2230', 'Fusce Corp.', '08 1
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
 6 •
 7 • INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                        'b-2234', 'Convallis In Incorp
                                                                                         'b-2238', 'Ante Iaculis Nec Fo
 8 •
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
       INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2242', 'Donec Ltd', '01 25
 9 •
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2246', 'Sed Nunc Ltd', '02
10 •
11 • INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2250', 'Amet Nulla Donec Co
                                                                                         'b-2254', 'Nascetur Ridiculus
12 • INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
13 •
                                                                                         'b-2258', 'Vestibulum Lorem PC
14 •
       INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2262', 'Gravida Sagittis LL
15 •
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2266', 'Mus Aenean Eget Fou
16 • INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2270', 'Dis Parturient Inst
17 • INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2274', 'Sed LLC', '01 63 16
                                                                                         'b-2278', 'Arcu LLP', '06 46 0
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
18 •
       INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2282', 'Pretium Neque Corp.
20 •
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2286', 'Fringilla LLC', '08
21 • INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                        'b-2290', 'Quisque Libero LLC'
22 •
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2294', 'Auctor Mauris Vel L
      INSERT INTO company (id, company_name, phone, email, country, website) VALUES (
                                                                                         'b-2298', 'Elit Etiam Laoreet
```

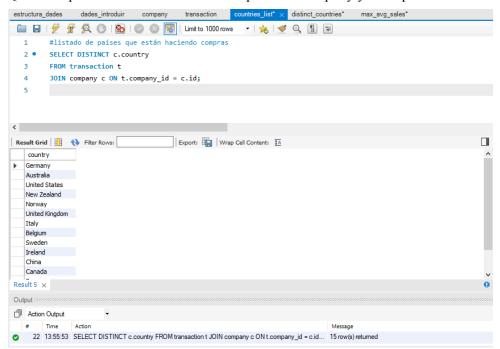
Aquí tenemos el archivo con los valores que se introducirán en las tablas creadas.



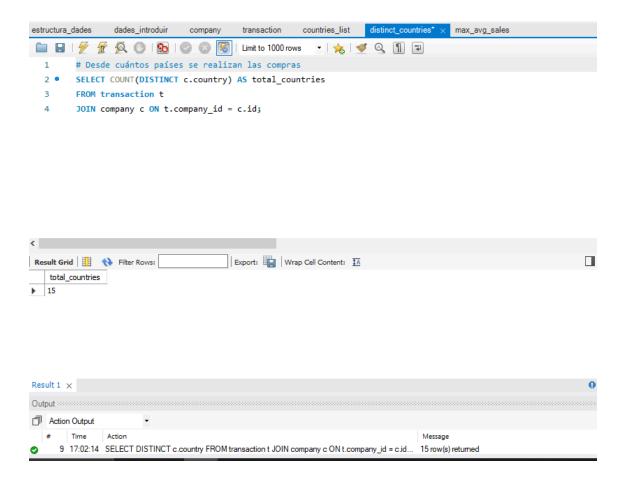
Aquí tenemos las dos tablas con sus columnas y los valores introducidos. Basta con hacer SELECT * (asterisco = seleccionar todos los valores) FROM "nombre de la tabla". Conviene fijarse en la última sección de Action Output para ver si la consulta se ejecuta o no correctamente y cuántas filas nos devuelve ésta para tener una idea del tamaño de cada tabla y posteriores consultas que realicemos. Nos ayuda un poco a comprobar si tiene sentido o no el output obtenido.

Ejercicio 2: Utilizando JOIN realizamos las siguientes consultas.

Listado de países que están haciendo compras. Para obtener dicho listado empezamos seleccionando con SELECT DISTINCT los valores únicos de la columna 'country' que se encuentra en la tabla company. Así eliminamos duplicados. Para agilizar el código de la query, podemos usar abreviaciones como alias e introducirlo seguido de "." y el nombre de la columna. Para company usaremos 'c' y en la tabla transaction usaremos 't'. Si nos fijamos, country está solamente en la tabla company, pero queremos unir ambas tablas para ver los países realizando compras o transacciones. La forma de relacionar o unir ambas tablas puede hacerse usando JOIN a través de la columna que tengan en común. En este caso, en la tabla transactions es company_id y en la tabla company es id. Veremos que la fila 'JOIN company c ON t.company_id = c.id' se repite en las 3 consultas del ejercicio 2. Usamos FROM transactions t para que la query acuda a la tabla de transacciones y busque a partir de company_id todas las filas con transacciones realizadas y el país asociado a cada transacción. Todas las consultas conviene finalizarlas con ';' en caso de tener que escribir en un mismo script varias. Así MySQL sabrá que son consultas distintas. Aquí vemos la query y el output obtenido:

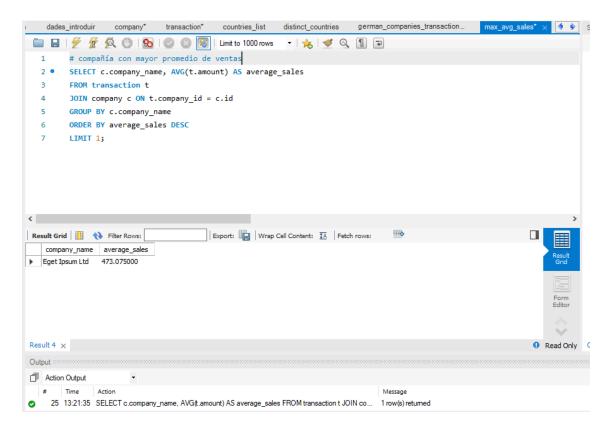


Desde cuántos países se realizan las compras. Para obtener el número de países que realizan las compras, sólo necesitamos incluir en la selección 'COUNT(DISTINCT c.country)'. La palabra clave COUNT nos hace el recuento total de la consulta realizada anteriormente. Podemos incluir un alias para que nos muestre la columna renombrada como 'total_countries' escribiendo 'AS total_countries'. A continuación mostramos la consulta en MySQL:



Identifica la compañía con la media más grande de ventas. Para realizar esta consulta primero seleccionamos las columnas que queremos visualizar en el output. Queremos el nombre de la empresa que está en la tabla company y la media de ventas, cuya columna 'amount' con el importe de éstas se encuentra en la tabla transaction. Para ello seleccionamos c.company_name y haremos el cálculo del promedio de ventas con AVG(t.amount), recordar que el importe se encuentra en la tabla de transaction, por eso empezamos con 't.' Podemos renombrar esta columna con un alias, así que pondremos 'AS average_sales' después del promedio calculado con AVG(t.amount). Las dos siguientes líneas de código son idénticas a las consultas anteriores. Para poder ver la empresa con el mayor promedio agrupamos por nombre de empresa usando GROUP BY c.company_name y ordenamos por valor del promedio usando ORDER BY y el alias que hemos creado, 'average_sales'. Para ver sólo el promedio de ventas más alto, usamos DESC (palabra clave que ordena de mayor a menor valor) y debajo usamos LIMIT 1 para ver la

primera fila con el nombre de la empresa y el mayor valor de promedio de ventas. A continuación mostramos el código SQL y el output.

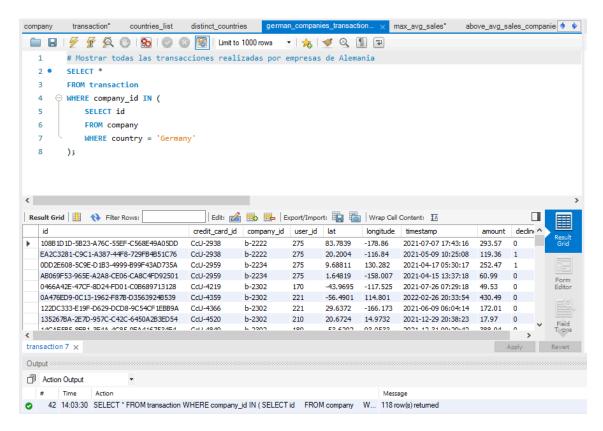


Como vemos en la consulta, la empresa Eget Ipsum Ltd tiene el mayor promedio de ventas.

Ejercicio 3

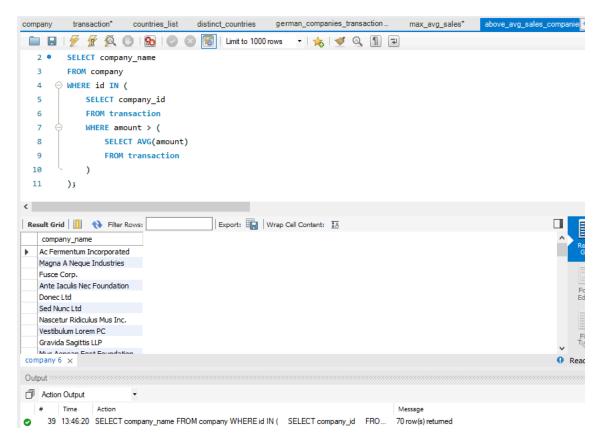
Crear subconsultas sin utilizar JOIN.

Mostrar todas las transacciones realizadas por empresas de Alemania. En esta consulta queremos ver todas las columnas de la tabla transaction cuyas empresas están localizadas en Alemania. Para ello seleccionamos todas las columnas con '*' FROM transaction e introducimos el criterio WHERE donde hacemos la relación entre company_id de la tabla transaction e id de la tabla company sin usar JOIN pero utilizando una subconsulta. El criterio de WHERE es company_id IN (subconsulta que seleccione todas las id de empresas ubicadas en Alemania, SELECT id FROM company WHERE country = 'Germany'). Vemos como queda en MySQL:



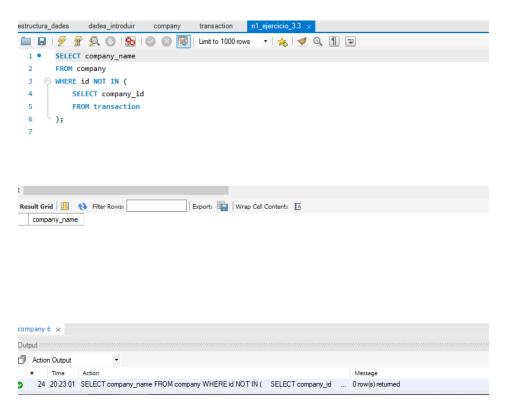
118 transacciones realizadas por empresas ubicadas en Alemania.

Mostrar todas las empresas cuyas transacciones realizadas tengan un importe superior al promedio de todas las transacciones. Para realizar esta consulta seleccionamos company_name de la tabla company, usamos una primera subconsulta con el criterio WHERE id IN para seleccionar company_id de la tabla transaction con otro criterio WHERE con la segunda subconsulta que nos coja la id de las empresas cuyo importe (amount) sea > al promedio de todas las transacciones en la tabla transaction. Basta con una selección de AVG(amount) de la tabla transaction. Según la teoría mostrada en este sprint, sería un ejemplo de consulta anidada o subconsulta dentro de otra subconsulta. Cabe destacar cómo se colocan los paréntesis finales, para que el usuario entienda que hay 2 subconsultas, se suele cerrar en una nueva fila y debajo del criterio WHERE correspondiente a cada subconsulta. El código se ejecuta correctamente si cerramos los paréntesis justo en la misma línea donde acabamos nuestro código pero supongo que si se diese el caso en que se debe hacer 3 o más subconsultas, es fácil olvidarse o confundirse al final cuántos paréntesis debemos introducir. Si se escribe el cierre de paréntesis debajo de cada WHERE, facilita la comprensión. Veamos como queda:



70 empresas realizan transacciones con importes por encima de la media de transacciones.

Eliminarán del sistema las empresas que carecen de transacciones registradas, mostrar el listado de estas empresas. Para realizar esta última consulta del ejercicio 3 en el nivel 1 seleccionamos el nombre de la compañía de la tabla company y usamos el criterio WHERE id NOT IN para identificar aquellas empresas que no tengan transacciones realizadas. Es decir, que no tengan ningún registro en la tabla transaction. Seleccionamos en la subconsulta company_id de la tabla transaction. Veamos como queda.

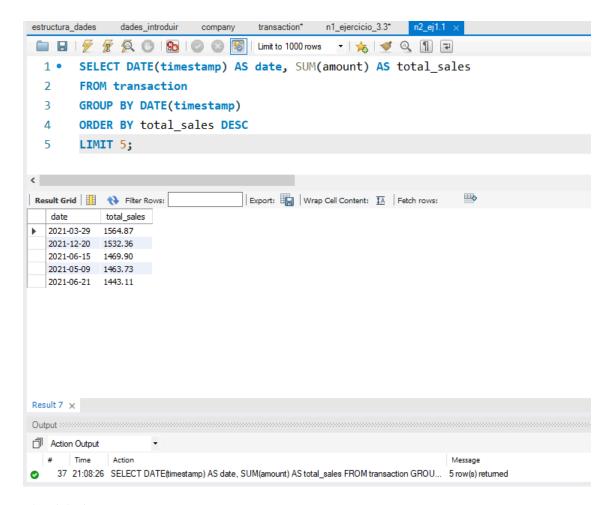


Si el listado nos sale vacío, todas las empresas tienen al menos un registro en la tabla transaction. No hay ninguna empresa sin realizar transacciones.

Nivel 2

Ejercicio 1

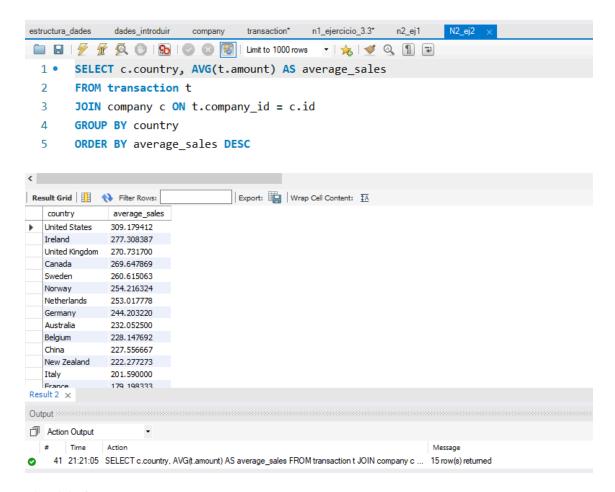
Identifica los cinco días que se generó la mayor cantidad de ingresos en la empresa por ventas. Muestra la fecha de cada transacción junto con el total de las ventas. Para hacer esta consulta basta con seleccionar la variable timestamp donde se indica la fecha y la hora, asegurándonos que cogemos solo la fecha con la cláusula DATE(timestamp). Para ver el total de ventas haremos SUM(amount). Podemos renombrar todo con alias 'AS date' y 'AS total_sales'. Agrupamos por fecha, ordenamos por importe de ventas totales y limitamos a 5 días con LIMIT 5.



Ejercicio 2

¿Cuál es la media de ventas por país? Presenta los resultados ordenados de mayor a menor medio. Para ello, seleccionamos country de la tabla company y AVG(t.amount) para coger los

importes de la tabla transaction y hacer el promedio, usamos 'average_sales' como alias. Agrupamos por país y ordenamos por average sales de mayor a menor valor usando DESC.

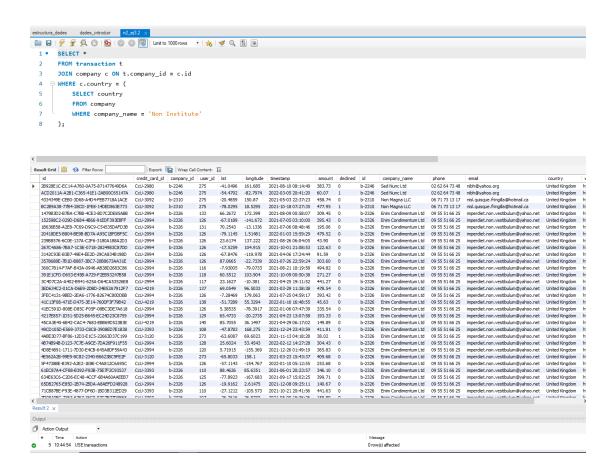


Ejercicio 3

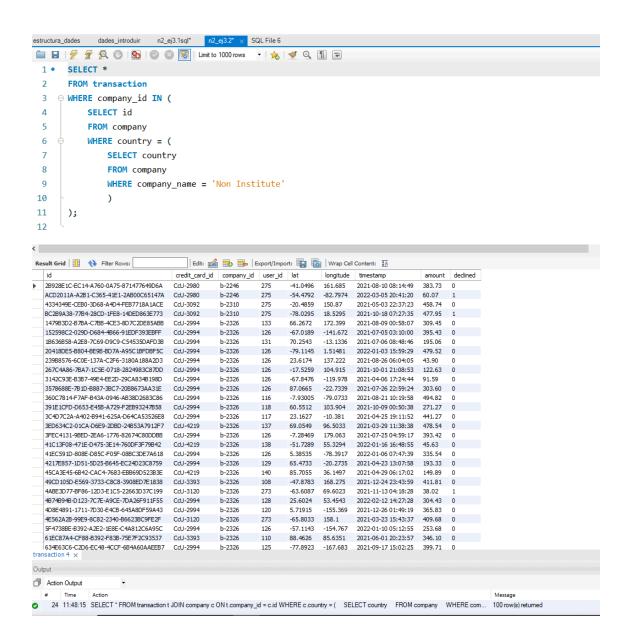
En tu empresa, se plantea un nuevo proyecto para lanzar algunas campañas publicitarias para hacer competencia a la compañía "Non Institute". Para ello, te piden la lista de todas las transacciones realizadas por empresas que están ubicadas en el mismo país que esta compañía.

Mostrar el listado aplicando JOIN y subconsultas:

Para esta consulta, seleccionamos todas las columnas de la tabla transaction, usamos el JOIN que hemos estado realizando en todas las consultas uniendo por t.company_id = c.id y en el criterio WHERE hacemos la subconsulta, queremos el país de la tabla company cuyo nombre de empresa sea 'Non Institute'.



Mostrar el listado con subconsultas solamente. En este caso haremos una misma selección de todas las columnas en la tabla transaction con SELECT * FROM transaction y hacemos coincidir ids de la compañía con WHERE company_id IN (subconsulta que encuentre ID de la tabla company donde el país coincida con el país de la empresa 'Non Institute'). Vemos como queda:

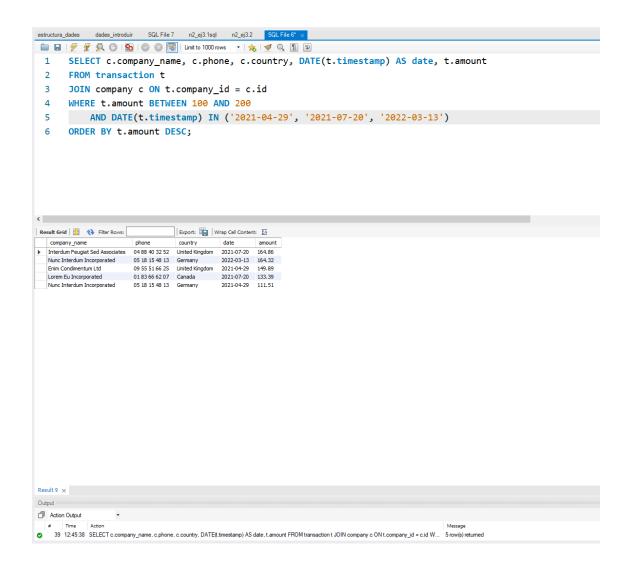


La diferencia entre ambas queries es que la primera muestra todas las columnas de ambas tablas y la segunda, usando solamente subconsultas, solo nos muestra las columnas de la tabla transaction. No obstante, el objetivo era mostrar transacciones de empresas que se encontrasen en el mismo país que Non Institute (UK).

Nivel 3

Ejercicio 1

Para ello, seleccionamos las variables pedidas, recordar coger la fecha con DATE(t.timestamp). Hacemos el JOIN que hemos usado durante toda la tarea y el criterio WHERE será t.amount entre 100 y 200 y que las fechas sean las indicadas con DATE(t.timestamp) IN (fechas en formato 'YY-MM-DD'). Ordenamos de mayor a menor valor con ORDER BY t.amount DESC.

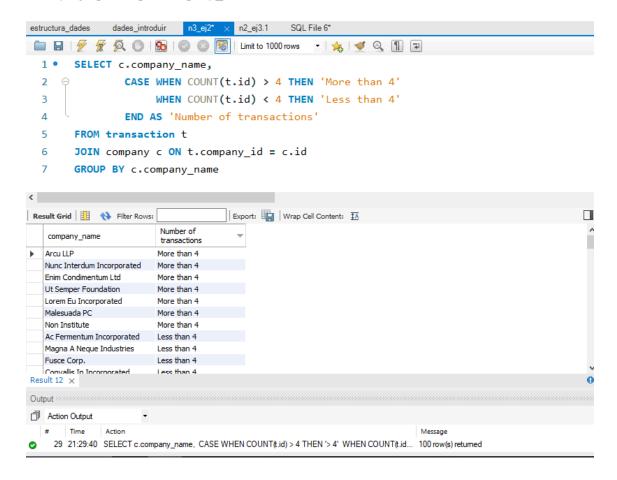


Ejercicio 2

Necesitamos optimizar la asignación de los recursos y dependerá de la capacidad operativa que se requiera, por lo que te piden la información sobre la cantidad de transacciones que realizan

las empresas, pero el departamento de recursos humanos es exigente y quiere un listado de las empresas donde especifiques si tienen más de 4 o menos transacciones.

Para realizar esta última consulta, haremos uso de CASE WHEN para hacer un listado de empresas con una columna que muestre si tienen más o menos de 4 transacciones. Usamos JOIN y agrupamos por company name.



Si hacemos click en el encabezado de la columna 'Number of transactions' se ordenan en 2 partes (más o menos de 4 transacciones).