Tasca S2.01. Nocions bàsiques SQL Fernando Poblete

Objetivo:

Repassar les nocions bàsiques per a l'ús de base de dades relacionals. En aquest sprint, iniciaràs la teva experiència pràctica amb una base de dades que conté informació d'una empresa dedicada a la venda de productes en línia. En aquesta activitat, t'enfocaràs en dades relacionades amb les transaccions efectuades i la informació corporativa de les empreses que van participar.

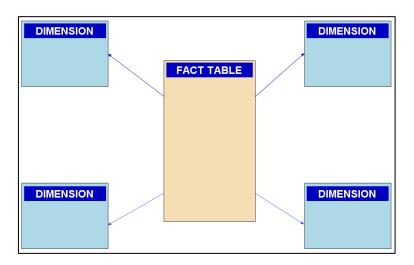
Nivell 1

- Exercici 1

A partir dels documents adjunts (estructura_dades i dades_introduir), importa les dues taules. Mostra les característiques principals de l'esquema creat i explica les diferents taules i variables que existeixen. Assegura't d'incloure un diagrama que il·lustri la relació entre les diferents taules i variables.

- Para crear la base de datos, primero ejecutamos el archivo 'estructura_dades.sql'. En este archivo se crea la base de datos llamada 'transactions', junto con sus dos tabla, 'company' y 'transaction'.
- Se específica para cada variable, el tipo de dato que contiene.
- Además se especifica si una variable será una Llave Primaria, o si será una Llave Secundaria, especificando a que Llave Primaria de otra tabla hace referencia.
- Después para cargar los datos ejecutamos el archivo 'dades_introduir.sql'. Alli se introducen datos en las tablas anteriormente creadas. En este caso tenemos 100 registros para la tabla 'company' y 587 registros para la tabla 'transaction'.

- Aunque solamente contamos con 2 tablas, podemos observar que el modelo de datos relacionales sigue el esquema de estrella. La tabla de 'transacciones' es la tabla de hechos, teniendo un gran volumen de datos, y llaves secundarias que pueden conectan a las llaves primarias de diversas tablas de dimensiones, como 'credit_card', 'user' y 'company'. En este ejemplo, solo contamos con la tabla de 'company'



Esquema Diagrama de estrella - Wikipedia

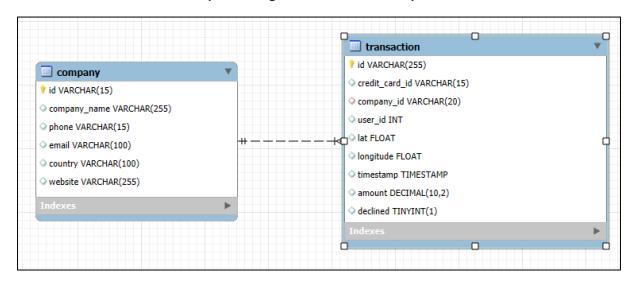


Diagrama entidad relacion

- En el diagrama entidad-relación, o modelo relacional, podemos observar que las tablas 'company' y 'transaction' tienen cada una, un primary key, que identifica los registros en cada tabla de manera única.
- Así mismo, la tabla de transacciones se conecta a la de compañia mediante el 'company_id', que en la tabla de 'company' simplemente se llama 'id'.

- La relación entre 'company' y 'transaction', es de 1:N, porque una compañía puede tener múltiples transacciones, mientras que cada transacción, solo puede ser asignada a 1 compañía.
- A continuación, un listado de las variables de cada tabla con una breve descripción de cada variable:

Variable	Tipo de dato	Descripción
Id	VARCHAR(15)	PRIMARY KEY
company_name	VARCHAR(255)	Nombre de la compañia
Phone	VARCHAR(15)	Telefono
email	VARCHAR(100)	email
country	VARCHAR(100)	pais
website	VARCHAR(255)	Sitio Web

Descripción Tabla 'company'

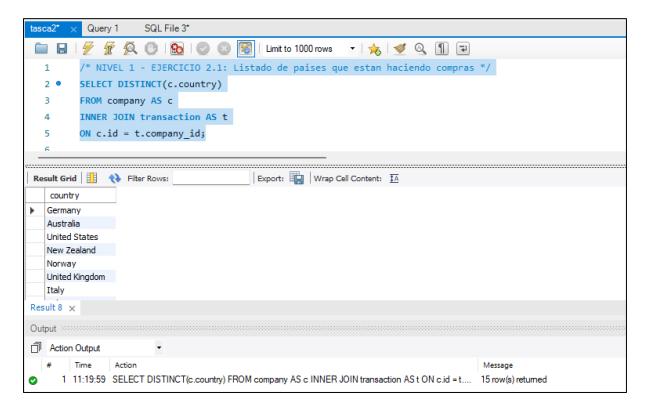
Variable	Tipo de dato	Descripción
id	VARCHAR(255)	PRIMARY KEY
credit_card_id	VARCHAR(15)	REFERENCES credit_card(id)
company_id	VARCHAR(20)	FOREIGN KEY (company_id) REFERENCES company(id)
user_id	INT	REFERENCES user(id)
lat	FLOAT	Coordenadas latitud
longitude	FLOAT	Coordenadas longitud
timestamp	TIMESTAMP	Tiempo de registro
amount	DECIMAL(10, 2)	Valor de la transacción
declined	BOOLEAN	Valor 1= TRUE -> Transacción declinada

Descripción Tabla 'transaction'

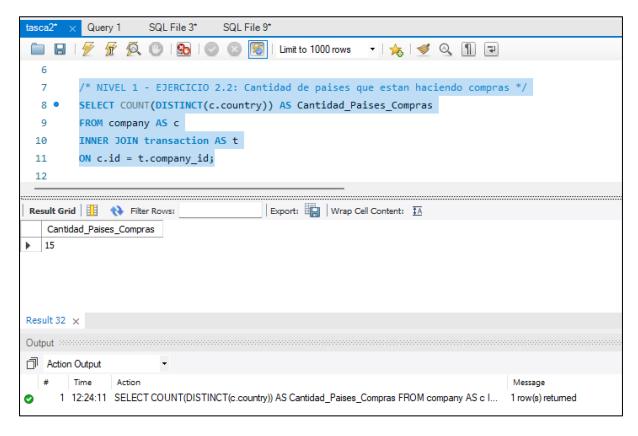
- Exercici 2

Utilitzant JOIN realitzaràs les següents consultes:

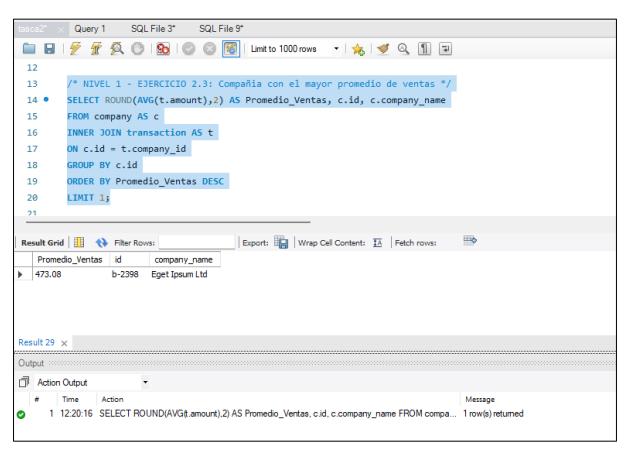
- Llistat dels països que estan fent compres.
- Hacemos un INNER JOIN entre la tabla de compañías y de transacciones, para obtener solamente los datos de compañías que tengan transacciones registradas. Luego obtenemos el listado de países sin repeticiones utilizando DISTINCT en la selección de datos.



- Des de quants països es realitzen les compres.
- Para obtener la cantidad de países que realizan compras, simplemente utilizamos la función de conteo, COUNT, en la selección de países sin repetición.

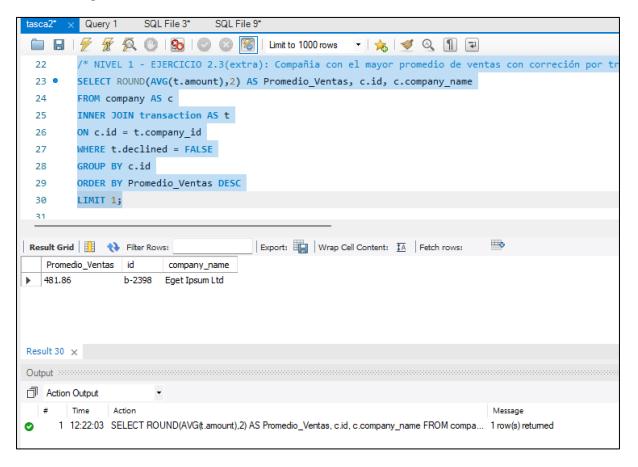


- Identifica la companyia amb la mitjana més gran de vendes.
- Para obtener la compañía con la media más grande de ventas, tenemos que extraer el promedio de los montos de las transacciones, con la función AVG sobre la variable 'amount'
- Conservamos la estructura de los JOINS, y agrupamos por el 'id' de la compañía.
- -Mostramos tanto el 'id' de la compañía como el 'company_name', sabiendo que la relación de estas dos variables siempre es 1:1.
- Redondeamos el resultado a 2 decimales por claridad, ordenamos los resultados de forma descendente en el promedio calculado, y limitamos los resultados a 1 línea, para obtener la compañía con el mayor valor en esta nueva variable creada.
- La compañía con mayor media de ventas es: 'Eget Ipsum Ltd', con 'id' de compañía 'b-2398'.



(extra)

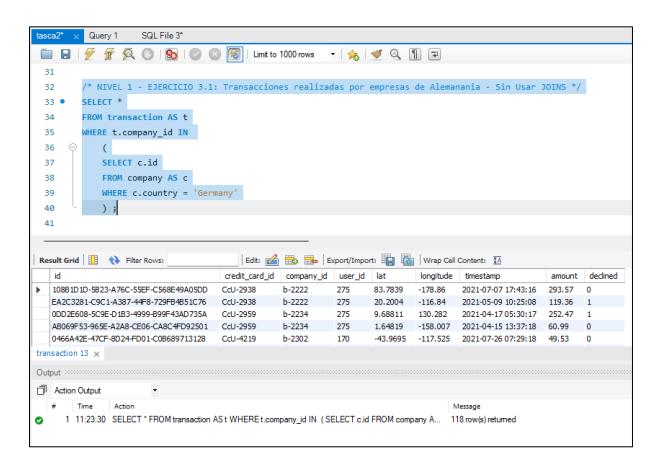
- Finalmente verificamos los resultados solamente para las transacciones no declinadas. Vemos que aunque el valor de la media cambia, la compañía con la media más alta sigue siendo la misma.



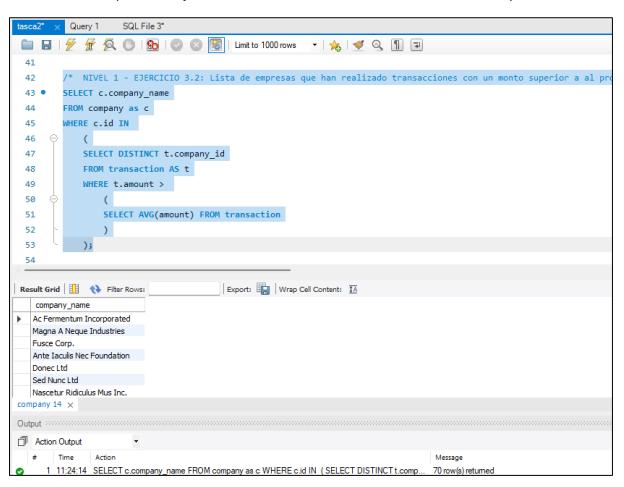
- Exercici 3

Utilitzant només subconsultes (sense utilitzar JOIN):

- Mostra totes les transaccions realitzades per empreses d'Alemanya.
- Primero, mediante una subconsulta, creamos una tabla que contenga todos los 'id' de las compañías que se encuentren en Alemania (country = 'Germany').
- Después, seleccionamos toda la información de la tabla de transacciones, y agregamos la condición, con la clausula WHERE, que me entregue solamente resultados donde el 'company_id' se encuentre en (IN) la tabla creada en el primera parte, que contiene solo identificadores de compañías en Alemania.
- Así, sin utilizar JOINS, obtenemos todos los datos de transacciones exclusivamente para compañías en Alemania.



- Llista les empreses que han realitzat transaccions per un amount superior a la mitjana de totes les transaccions.
- En este ejercicio vamos a hacer consultas anidadas en 3 niveles.
- Primero, en una subconsulta en la tabla 'transaction', calculamos el valor promedio de todas las transacciones.
- Luego utilizamos este valor calculado, como parámetro en una clausula WHERE, para filtrar los registros de la tabla 'transaction', extrayendo únicamente los identificadores de empresa 'company_id´, sin repeticiones (las repeticiones se pueden eliminar también en el próximo nivel). Con repeticiones obtenemos 297 registros, y al agregar DISTINCT al SELECT, obtenemos solo 70 registros.
- Finalmente, para obtener los nombres de estas empresas y no sus códigos, utilizamos el resultado del paso anterior, de la misma manera que en el ejercicio 3.1, para extraer los nombres de las compañías, condicionando en la cláusula WHERE, que sean solamente compañías cuyos 'id' se encuentren en el listado creado en el paso anterior.

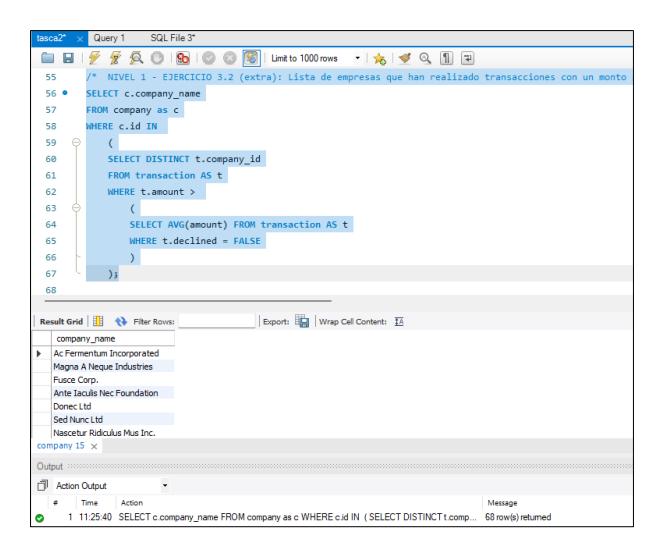


(extra)

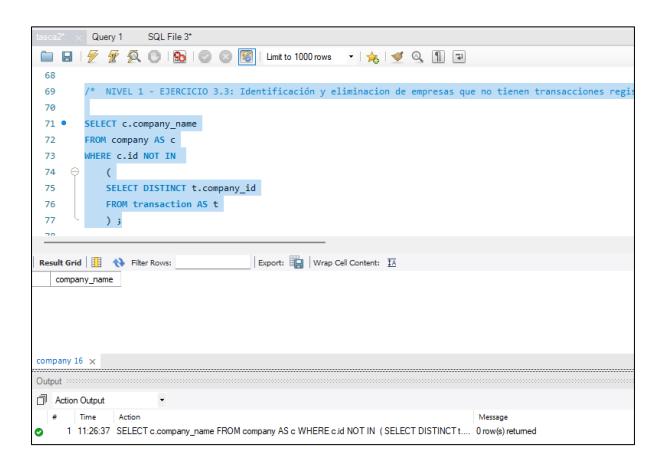
- De manera adicional, si queremos ser más exactos, podemos omitir las transacciones declinadas, especificando mediante una cláusula WHERE en el primero nivel, que se consideren solamente los montos de transacciones aceptadas (declined = FALSE).

Esto es importante para poder realizar decisiones de negocios, ya que las transacciones declinadas no representan ingresos para la compañía.

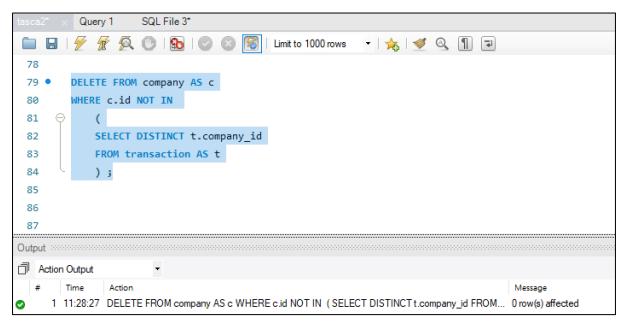
Al hacer esta corrección, la cantidad de empresas que han realizado transacciones disminuye de 70 a 68 registros.

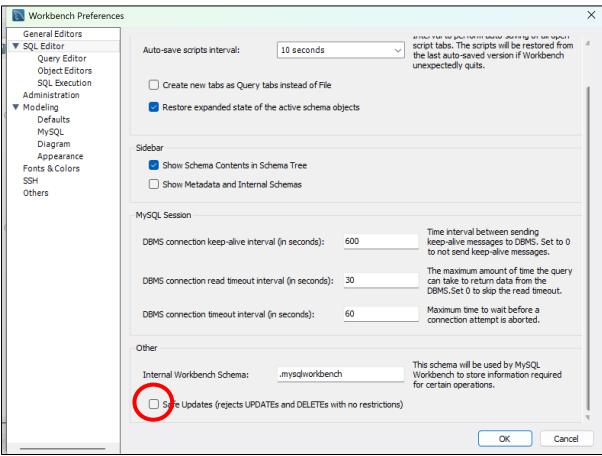


- Eliminaran del sistema les empreses que no tenen transaccions registrades, entrega el llistat d'aquestes empreses.
- Primero, extraemos un listado de 'id_company' sin repetición de la tabla de transacciones.
- Luego utilizaremos el resultado anterior, como una subconsulta, para filtrar mediante una cláusula WHERE, las empresas que no se encuentren en el listado, con el operador NOT IN.
- No obtenemos registros, lo que quiere decir, que todas las 100 compañías de la tabla 'company' tienen transacciones registradas en la tabla 'transactions'



- Como era de esperarse, al ejecutar el código para eliminar las empresas sin transacciones, ningún registro se ve afectado.
- Incluimos para futuras referencias, un pantallazo de la opción que se debe cambiar para poder eliminar datos de una tabla (recordar reconectarse después de cambiar esta opción).



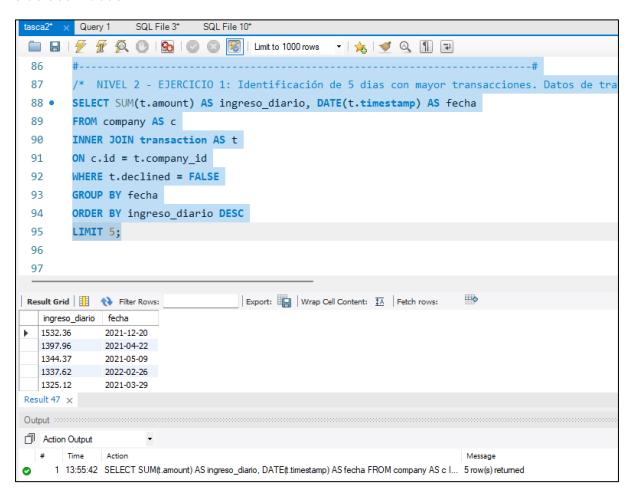


Nivell 2

Exercici 1

Identifica els cinc dies que es va generar la quantitat més gran d'ingressos a l'empresa per vendes. Mostra la data de cada transacció juntament amb el total de les vendes.

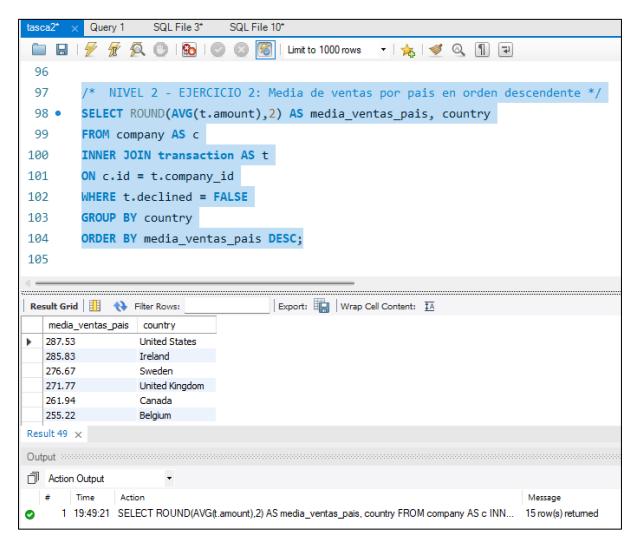
- Para obtener las ventas diarias, debemos extraer la agregación de valores de 'amount' que indican los ingresos de la compañía con la función SUM(), y agruparlos por día.
- Como no contamos con la variable de 'día' debemos crearla, a partir de la variable timestamp, que es una versión extensa de tiempo. Aplicamos la función DATE().
- Nuevamente consideramos en el análisis solamente las transacciones que no han sido declinadas.



Exercici 2

Quina és la mitjana de vendes per país? Presenta els resultats ordenats de major a menor mitjà.

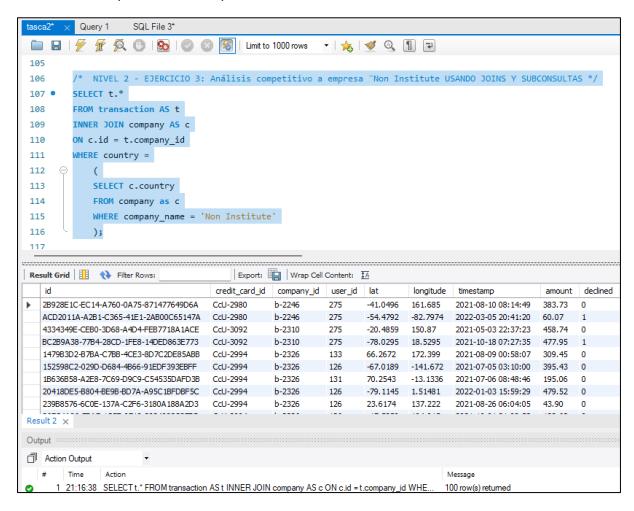
- Para obtener la media de ventas por país, calculamos el promedio de ventas con la función AVG sobre 'amount', y agrupamos los datos por país. Nuevamente excluimos las transacciones declinadas.



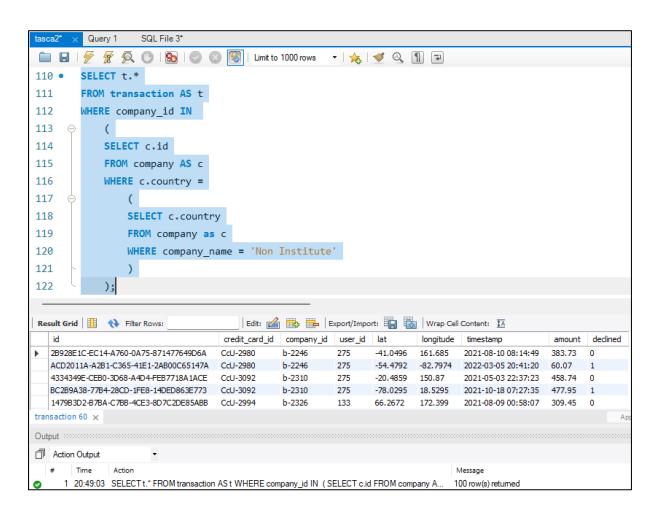
Exercici 3

En la teva empresa, es planteja un nou projecte per a llançar algunes campanyes publicitàries per a fer competència a la companyia "Non Institute". Per a això, et demanen la llista de totes les transaccions realitzades per empreses que estan situades en el mateix país que aquesta companyia.

- Mostra el llistat aplicant JOIN i subconsultes.
- Para ambos casos, decidimos incluir las transacciones declinadas, por si presentan valor para el análisis competitivo.
- Primero obtenemos de que país es la compañía 'Non Institute', directamente desde la base de datos, haciendo un SELECT de 'country' de la tabla 'company'.
- Luego de unir las tablas con un JOIN, extraemos las transacciones de la tabla 'transaction', especificando en su clausula WHERE, que el país coincida con lo obtenido en el paso anterior, que usamos como subconsulta.



- Mostra el llistat aplicant solament subconsultes.
- Primero obtenemos de que país es la compañía 'Non Institute', directamente desde la base de datos, haciendo un SELECT de 'country' de la tabla 'company'.
- Luego obtenemos un listado de los 'id' de todas las compañías que se encuentran en ese país. Para eso, extraemos los 'id' de la tabla 'company', indicando en una cláusula WHERE, que solamente queremos los 'id' de compañías cuyo 'country' sea igual al obtenido en la consulta anterior.
- Finalmente en la tabla de transacciones, extraemos el listado completo de columnas, condicionando en la clausula WHERE que se incluyan solamente aquellos registros que incluyen un 'company_id' que se incluya en el listado obtenido en el paso anterior.
- Así con las consultas anidadas, obtenemos las transacciones realizadas solamente en el Reino Unido, país donde se encuentra la compañía 'Non Institute'.
- Al comparar ambos métodos, el uso de JOINS parece mucho más eficiente, conectando la información a través de las diferentes tablas con mayor facilidad.

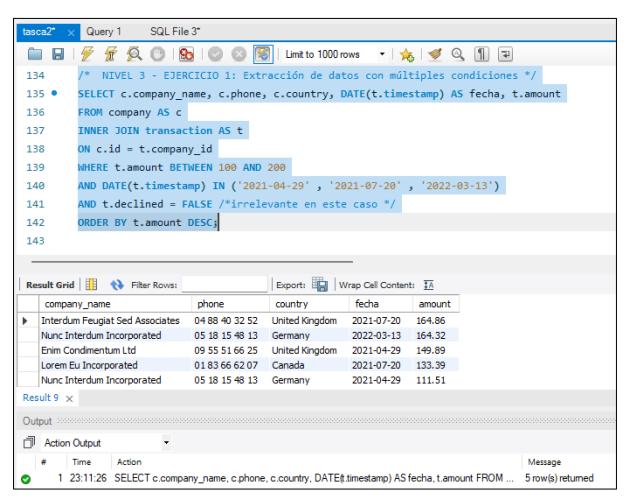


Nivell 3

Exercici 1

Presenta el nom, telèfon, país, data i amount, d'aquelles empreses que van realitzar transaccions amb un valor comprès entre 100 i 200 euros i en alguna d'aquestes dates: 29 d'abril del 2021, 20 de juliol del 2021 i 13 de març del 2022. Ordena els resultats de major a menor quantitat.

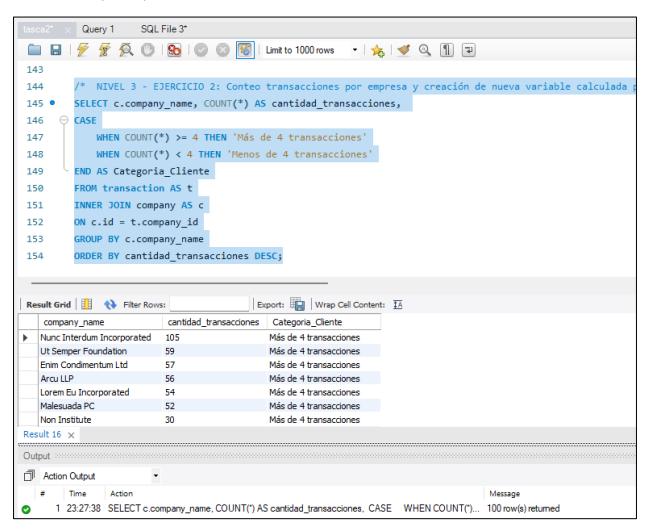
- Primero definimos todos los datos que queremos extraer de cada tabla, obteniendo la fecha con la función DATE() sobre la variable 'timestamp' de la tabla 'transaction'.
- Luego, después de definir el JOIN, ingresamos todas las condiciones solicitadas en la cláusula WHERE.
- Notamos que no podemos utilizar el alias 'fecha' en la cláusula WHERE, ya que, aunque el 'SELECT' se escribe antes del 'WHERE' en el código, se ejecutan de manera inversa, por lo que si el alias está definido en 'SELECT' al ejecutarse la cláusula 'WHERE' esta no tiene conocimiento del alias.



Exercici 2

Necessitem optimitzar l'assignació dels recursos i dependrà de la capacitat operativa que es requereixi, per la qual cosa et demanen la informació sobre la quantitat de transaccions que realitzen les empreses, però el departament de recursos humans és exigent i vol un llistat de les empreses on especifiquis si tenen més de 4 transaccions o menys.

- Para cumplir con lo solicitado, debemos crear una nueva variable, que nos diga en cuál de las dos categorías se encuentra la empresa listada.
- Aunque el caso de empresas con 4 transacciones no queda claro en la solicitud, en este caso no afecta los resultados.
- Unimos las tablas 'company' y 'transaction' con un JOIN, y le pedimos el nombre de la compañía, la cantidad de transacciones (haciendo uso de la función COUNT()), y finalmente la categoría en la que clasifica.
- La categoría la creamos usando CASE en la cláusula SELECT, para crear columnas calculadas, que dependen del conteo de transacciones.



- Finalmente, si queremos solo incluir las transacciones no declinadas, agregamos esta condición en la cláusula WHERE, obteniendo el siguiente resultado:

