

# Sprint2 BBDD relacionales e introducción a SQL

## Tarea S2.01. Nociones básicas SQL

### - Nivel 1

#### - Ejercicio 1

A partir de los documentos adjuntos (*estructura\_dades* y *dades\_introduir*), importa las dos tablas. Muestra las características principales del esquema creado y explica las diferentes tablas y variables que existen. Asegurate de incluir un diagrama que ilustre la relación entre las diferentes tablas y variables.

Primero, se crea la DATABASE apartir de las instrucciones facilitadas (*Imagen1*)

Segundo, se crea la tabla company con el código facilitado.

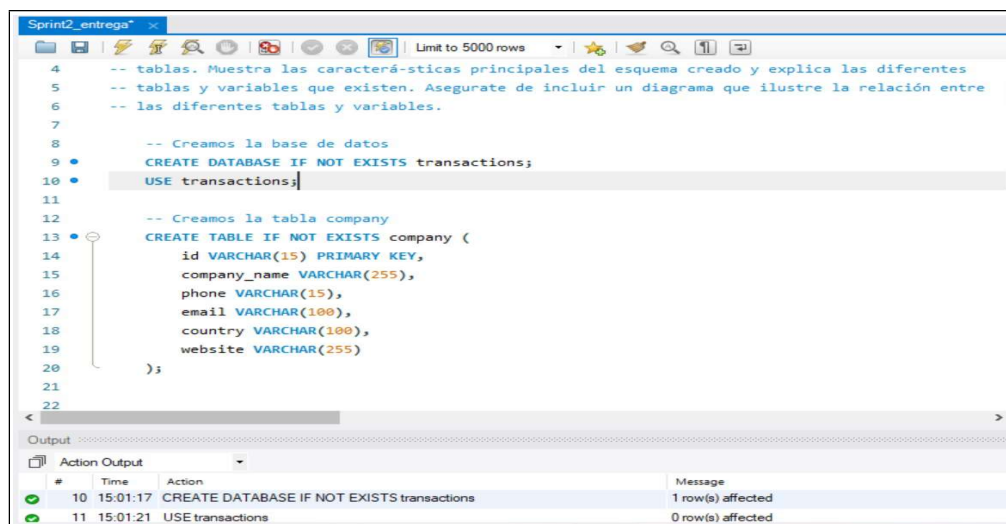


Imagen 1. Captura de pantalla de la creación de la base de datos transactions.

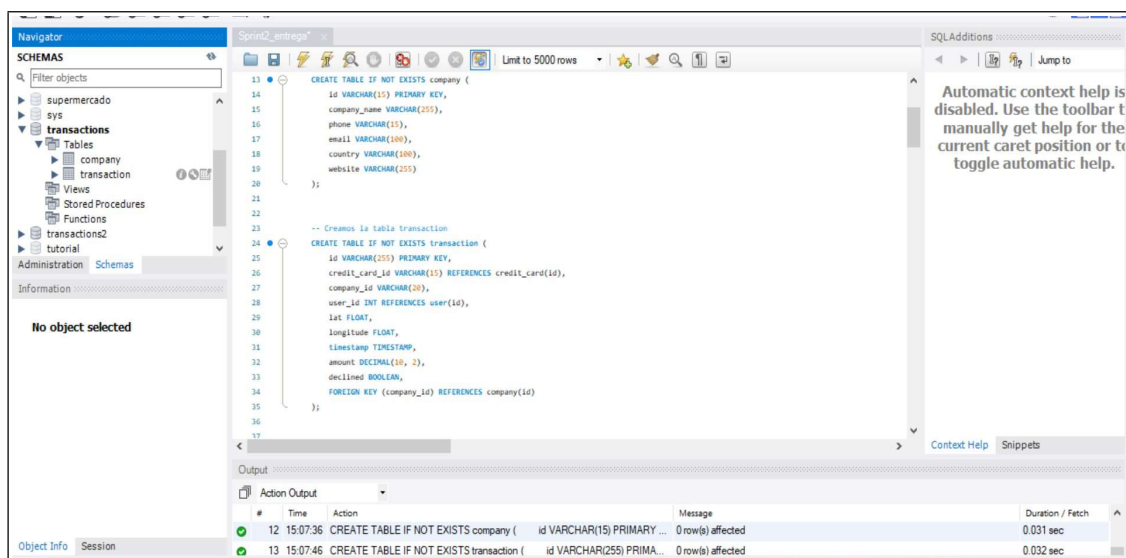


Imagen 2. Captura de pantalla de la creación de las tablas company y transaction.

Nuestra base de datos está formada por las tablas: *company* y *transaction*.

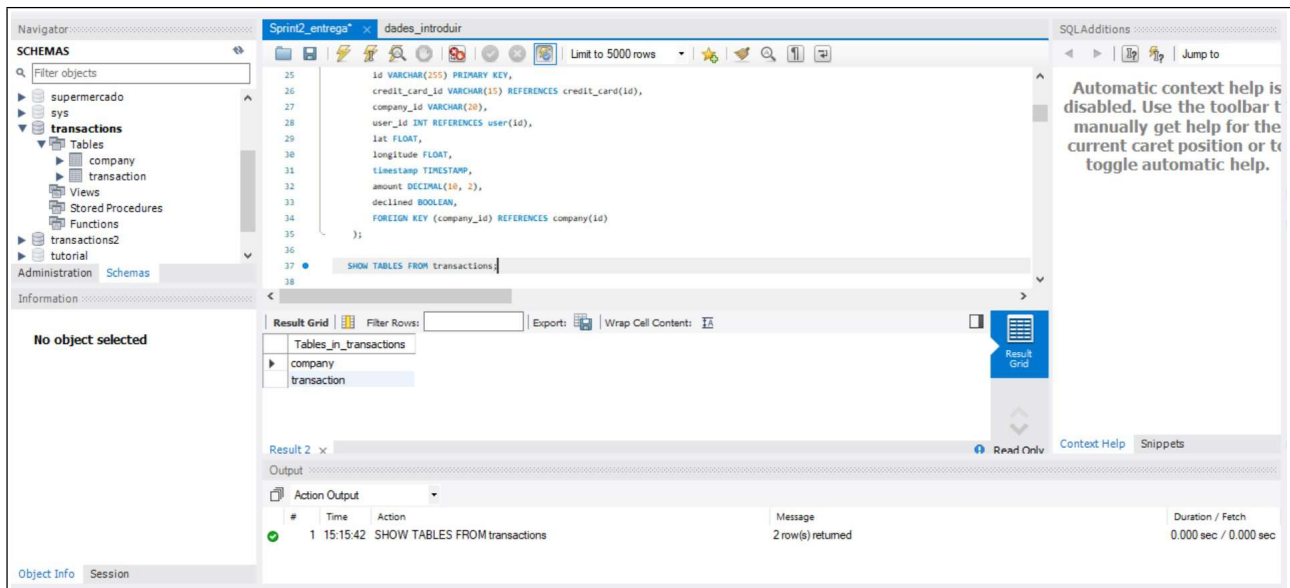


Imagen 3. Imagen de las tablas contenidas en mi DDBB *transactions*.

La tabla *company* contiene los siguientes campos:

- la clave primaria (o ID)
- nombre de la compañía
- teléfono, email
- país
- dirección web de la empresa

En este caso, todos los campos son variables de tipo texto.

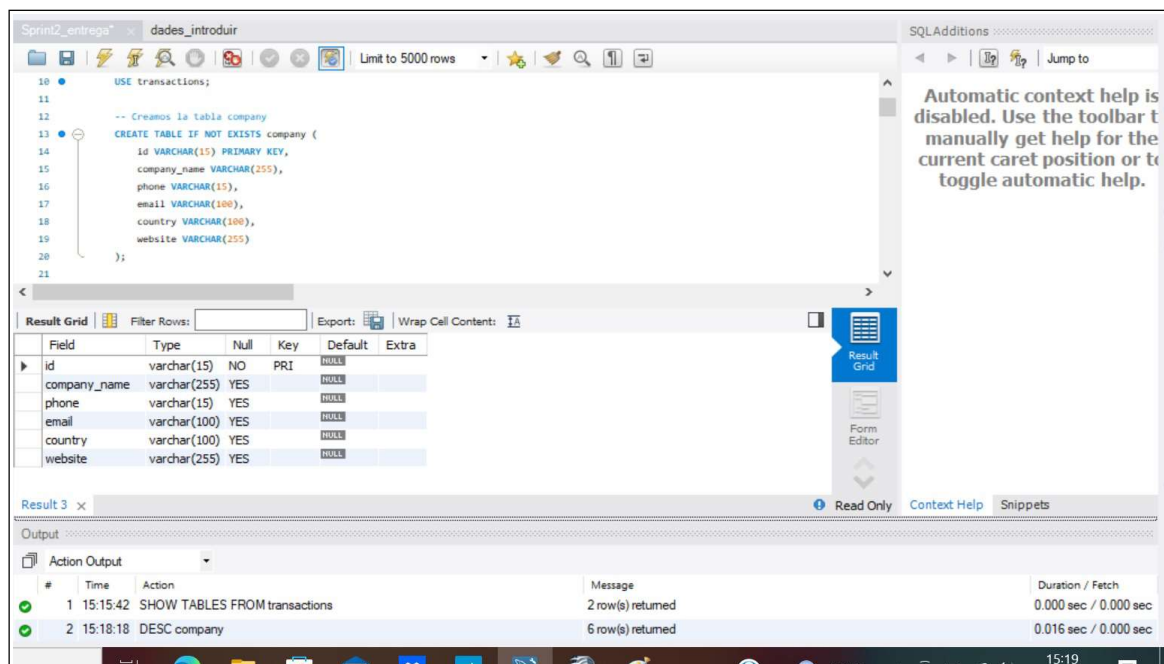


Imagen 2. Imagen de los campos que contenidos en la tabla *company*.

La tabla transaction contiene los siguientes campos:

- **id:** variable de tipo texto
- **tarjeta de crédito:** variable de tipo texto
- **id company** (corresponde a la clave foranea): variable de tipo texto
- **código de usuario:** variable de tipo integer
- **Latitud:** variable de tipo float
- **Longitud:** variable de tipo float
- **día y hora de la transacción:** variable de tipo timestamp, toma el día y la hora del momento.
- **cantidad:** variable de tipo decimal con redondeo a dos dígitos después de la coma.
- **declined:** variable de tipo tinyint. Es un integer con un peso muy pequeño.

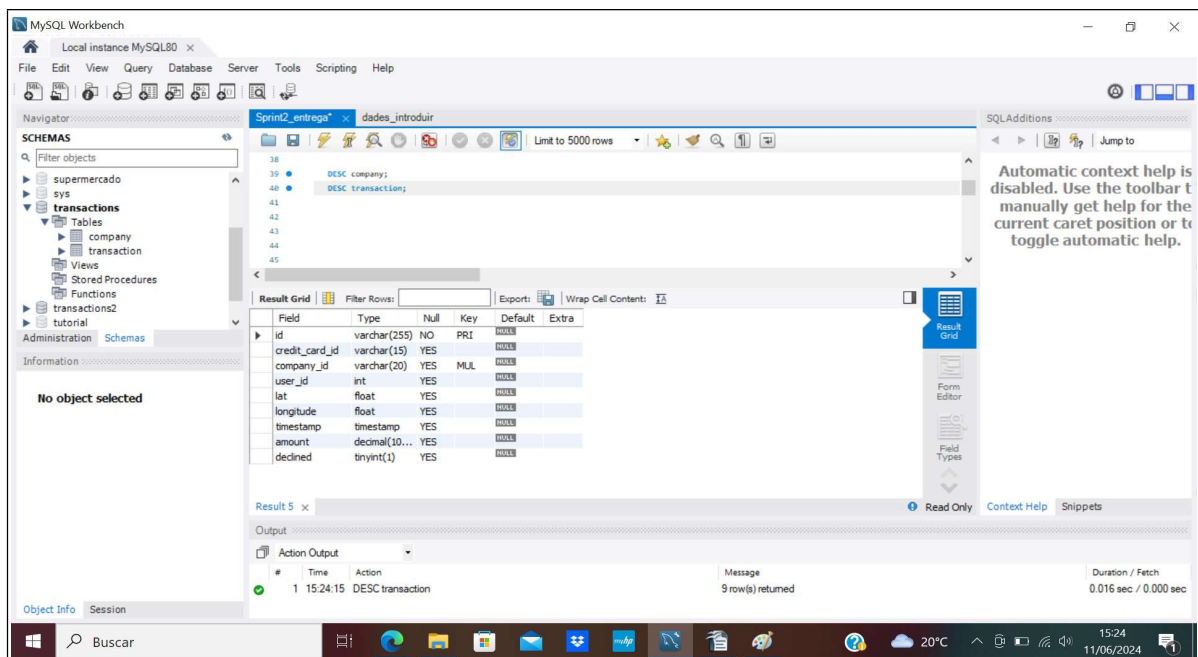


Imagen 3. Imagen de los campos que contenidos en la tabla **transaction**.

La relación entre las tablas de nuestro modelo es de 1:N. Una company tiene N registros en la tabla transacciones.

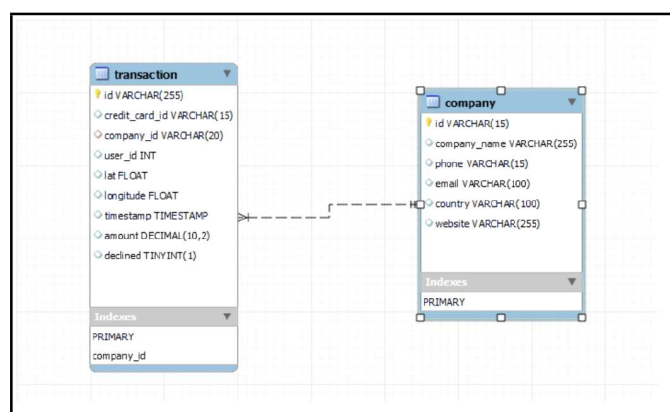


Imagen 4. Imagen de Diagrama Entidad Relación.

- **Ejercicio 2.** Utilizando JOIN realizarás las siguientes consultas.

### 2.1. Listado de los países que están haciendo compras.

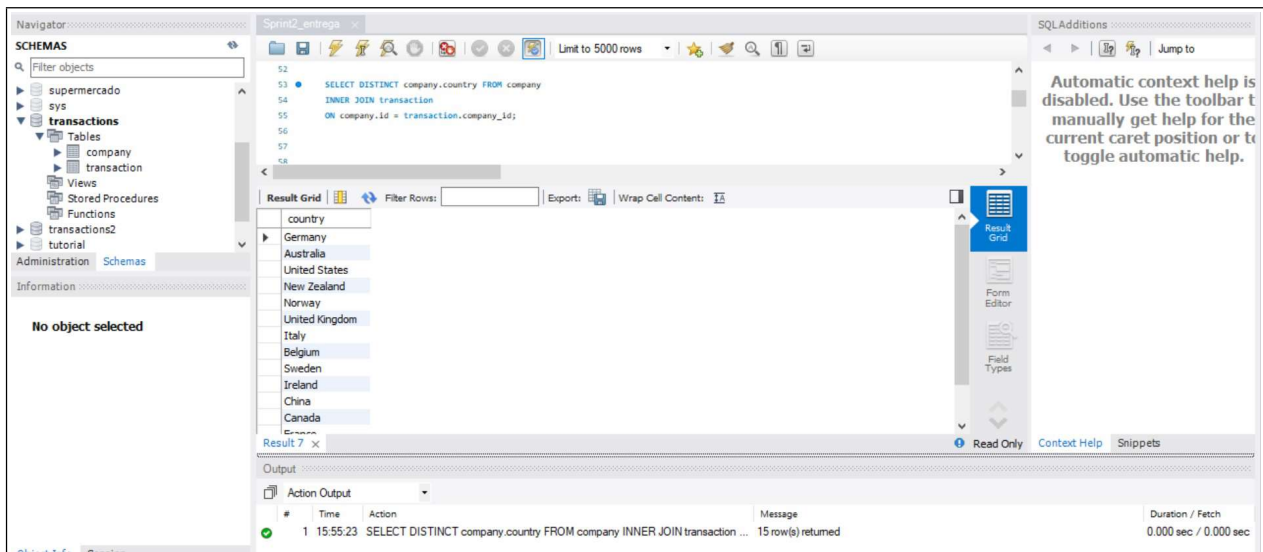


Imagen 5. Captura de pantalla de la ejecución de la consulta. El resultado son 15 países.

Primero se selecciona los country con una función DISTINCT que nos elimina los valores repetidos. Realizamos un JOIN para buscar los distintos países de la tabla **company** en la tabla **transaction**. En este caso, no tengo en cuenta el declined, porque es una consulta sobre los países que realizan compras.

El resultado de la consulta es el siguiente:

Australia  
Belgium  
Canada  
China  
France  
Germany  
Ireland  
Italy  
Netherlands  
New Zealand  
Norway  
Spain  
Sweden  
United Kingdom  
United States

## 2.2. Desde cuantos paises se realizan las compras.

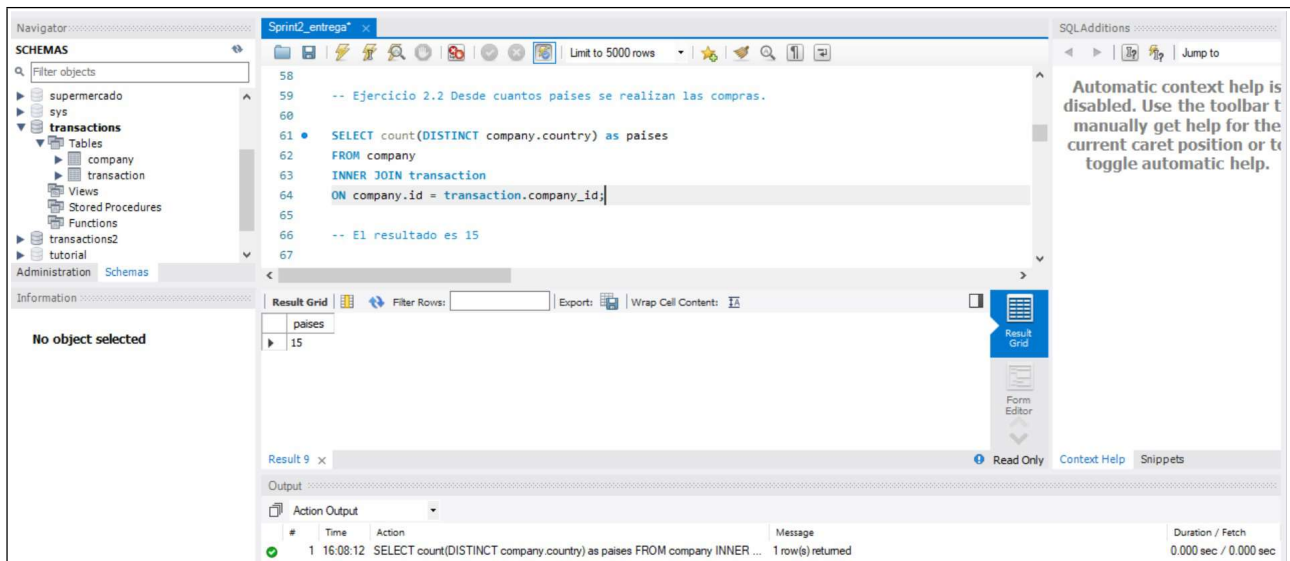


Imagen 5. Captura de pantalla de la ejecución de la consulta. El resultado del count son 15 paises.

Resultado: Se realizan las compras desde 15 paises diferentes

En este ejercicio, se realiza la repite la consulta del ejercicio anterior. Y se añade la función count para contar el número total de paises de la consulta anterior.

## 2.3. Identifica la compañía con la media más gran de ventas.

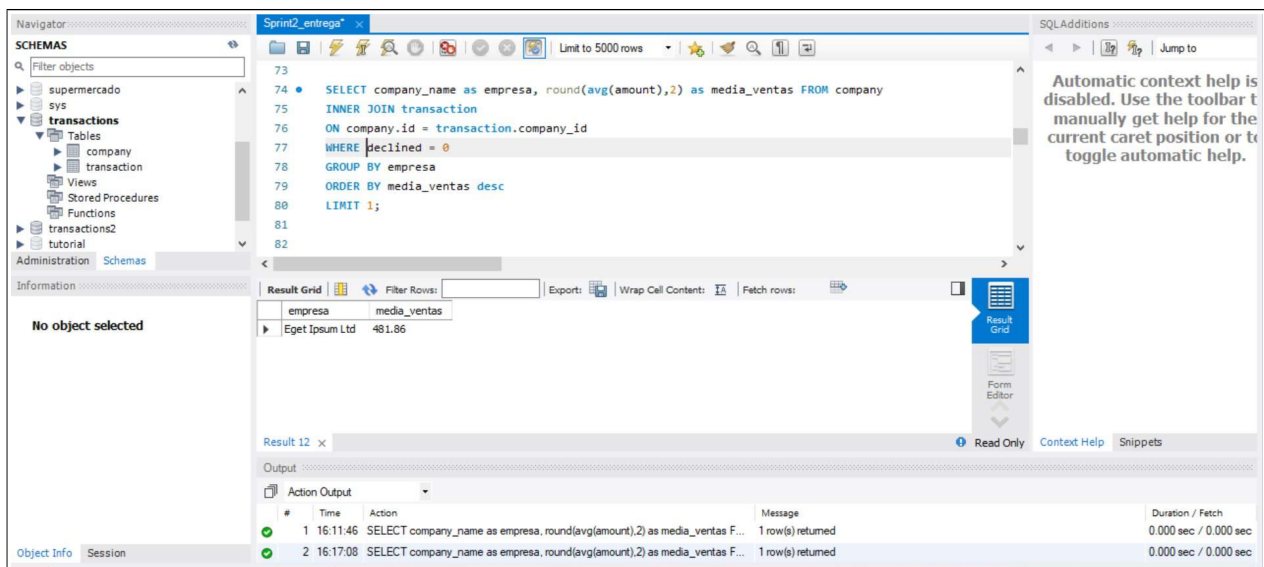


Imagen 6. Captura de pantalla de la ejecución de la consulta.

La compañía con la media más grande de ventas es: Eget Ipsum Ltd = 481.86

En esta consulta, Se busca listas los campos de nombre de la compañía y el average del campo amount. Este ultimo, como se trata de un campo de moneda se redodondea a dos decimales con la función Round.

Hacemos una JOIN para relacionar las dos tablas que contienen la información a relacionar, y filtramos con un WHERE con la condición que el **declined sea igual a cero**. Esto quiere decir que las transacciones realizadas se hayan hecho con éxito. Se utiliza el filtro del declined, ya que una venta se puede considerar venta si la transacción se ha realizado con éxito, para mí.

Al contener un campo de agrupación en la SELECT se debe agrupar por el campo descriptivo company\_name.

Por último, ordenamos (ORDER BY) por la media\_ventas en orden descendente y seleccionamos el primer registro con un LIMIT 1.



### - Ejercicio 3. Utilizando sólo subconsultas (sin utilizar JOIN)

#### 3.1. Muestra todas las transacciones realizadas por empresas de Alemania.

The screenshot shows a database IDE with a SQL query editor and a results grid. The query is as follows:

```
-- Ejercicio 3.1 Muestra todas las transacciones realizadas por empresas de Alemania.
SELECT *
FROM transaction
WHERE company_id IN
  (SELECT id FROM company
   WHERE country = 'Germany');
```

The results grid displays the following data:

id	credit_card_id	company_id	user_id	lat	longitude	timestamp
10881D1D-5B23-A76C-55EF-C568E49A05DD	CcU-2938	b-2222	275	83.7839	-178.86	2021-07-07 17:43:16
EA2C3281-C9C1-A387-4F8-729FB4B51C76	CcU-2938	b-2222	275	20.2004	-116.84	2021-05-09 10:25:08
0DD2E608-5C9E-D1B3-4999-B99F43AD735A	CcU-2959	b-2234	275	9.68811	130.282	2021-04-17 05:30:17
AB069F53-965E-A2A8-CE06-CA8C4FD92501	CcU-2959	b-2234	275	1.64819	-158.007	2021-04-15 13:37:18
0466A42E-47CF-8D24-FD01-C0B689713128	CcU-4219	b-2302	170	-43.9695	-117.525	2021-07-26 07:29:18

The output pane shows a message: "1 16:19:38 SELECT \* FROM transaction WHERE company\_id IN (SELECT id FROM comp... 118 row(s) returned".

Imagen 7. Captura de pantalla de la ejecución de la consulta.

El resultado es de 118 transacciones.

En esta consulta, se realiza una subquery donde se seleccionan los id de la tabla company, donde se cumpla la condición que el country sea igual a Germany. Esta lista se utiliza para filtrar a través de un IN, todos los campos de cada registro de la tabla transaction.

#### 3.2. Lista las empresas que han realizado transacciones por un amount superior a la media de todas las transacciones.

The screenshot shows a database IDE with a SQL query editor and a results grid. The query is as follows:

```
-- Ejercicio 3.2 lista las empresas que han realizado transacciones por un amount
-- superior a la media de todas las transacciones.
SELECT company_id, max(amount) as maximo
FROM transaction
GROUP BY company_id
HAVING maximo > (SELECT avg(amount) as mediaTotal
                  FROM transaction);
```

The results grid displays the following data:

company_id	maximo
b-2222	293.57
b-2226	480.13
b-2230	460.38
b-2238	401.53
b-2242	364.61
b-2246	383.73

The output pane shows a message: "1 16:31:15 SELECT company\_id, max(amount) as maximo FROM transaction GROUP BY c... 70 row(s) returned".

Imagen 8. Captura de pantalla de la ejecución de la consulta.

El resultado es una lista formada por 70 países.

En este caso, se seleccionan los campos `company_id`, y se busca el máximo del campo `amount` para cada `company`, asignándole un alias. Esto quiere decir que se agrupa por `company_id`, al tener en la `SELECT` una función de agregación, es necesario añadir el `GROUP BY`.

Para obtener las empresas que tengan alguna transacción mayor a la media, se compara con un 'mayor que' ( $>$ ), el valor máximo del `amount` por empresa con el average de todas las transacciones que hay en la tabla `transaction`.

Por último, en este caso, se calcula el máximo del `amount` por empresa, ya que si existe un valor que sea mayor a la media del total, ya cumpliría con la condición solicitada.

### 3.3. Eliminar del sistema las empresas que no tienen transacciones registradas, entrega el llistat d'aquestes empreses.

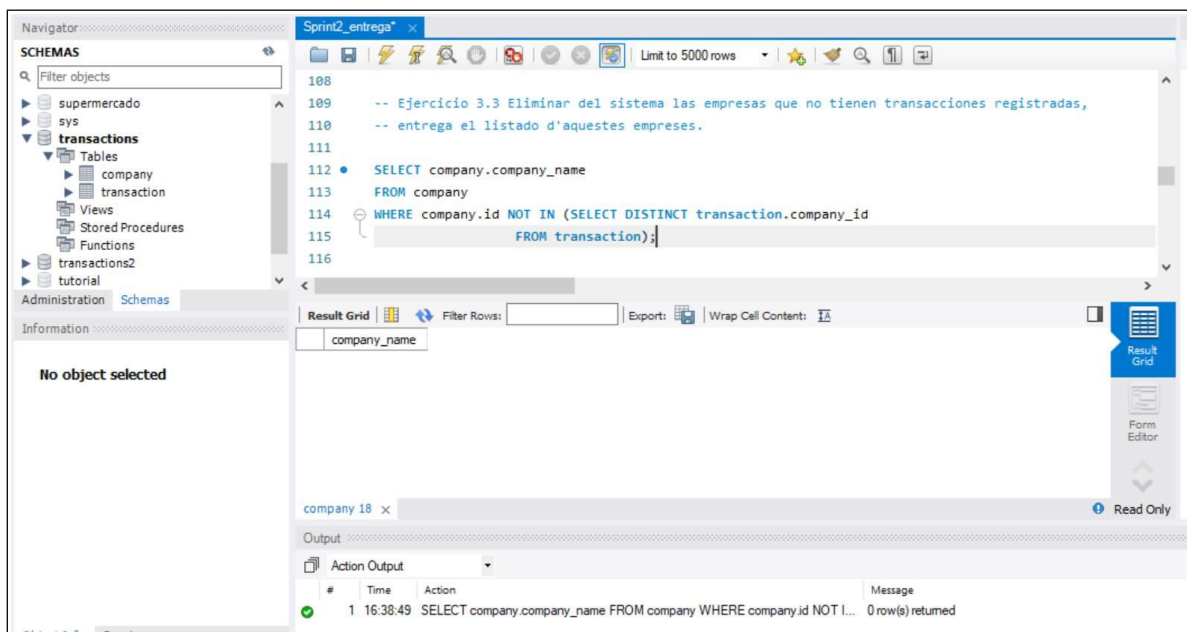


Imagen 9. Captura de pantalla de la ejecución de la consulta.

En esta consulta, se listarán los nombres de las compañías (`company_name`) de la tabla `company` que cumplen la condición (`WHERE`) que el campo `id` (de la tabla `company`) no se encuentre dentro de la siguiente subconsulta, con la función `NOT IN`. Esta es una lista con los distintos `company_id` de la tabla `transaction`. El resultado es una tabla vacía.

La conclusión es que todas las empresas de la tabla `company` tienen transacciones en la tabla `transcription`. NO es necesario eliminar ningún registro.

## - 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.

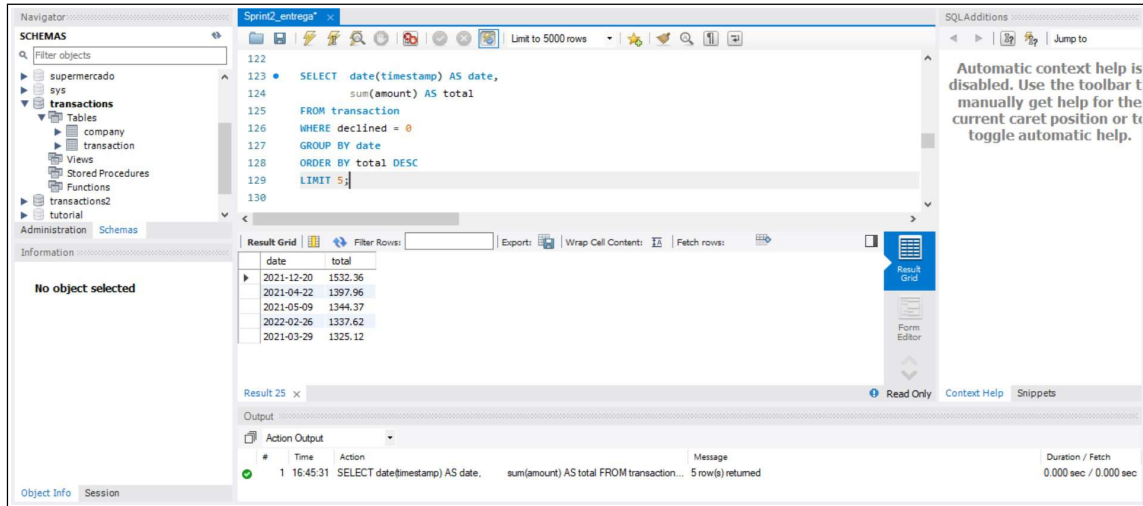


Imagen 10. Captura de pantalla de la ejecución de la consulta.

En esta consulta, se busca listar la fecha de la transacción (transaction.timestamp). A esta fecha, se le aplica una función date, para cambiar el formato de la fecha, y que muestre sólo el día. Y listar la suma (función de agregación) del amount agrupada por esa misma fecha.

No es necesario hacer un JOIN en esta consulta, ya que todos los campos que utilizo están en la tabla transactions.

En este caso, al tratarse de ventas, se tiene en cuenta el declined, considerando solo los valores cero, que son los que tienen transacciones efectivas. Considero que son Ventas, aquellas que tienen la transacción aceptada. Eso corresponde al declined = 0.

Para acabar, se ordena de manera descendiente por la suma del amount, y se limita a 5 los resultados. Así se seleccionan los 5 valores máximos.

### Ejercicio 2

¿Cuál es el promedio de ventas por país? Presenta los resultados ordenados de mayor a menor.

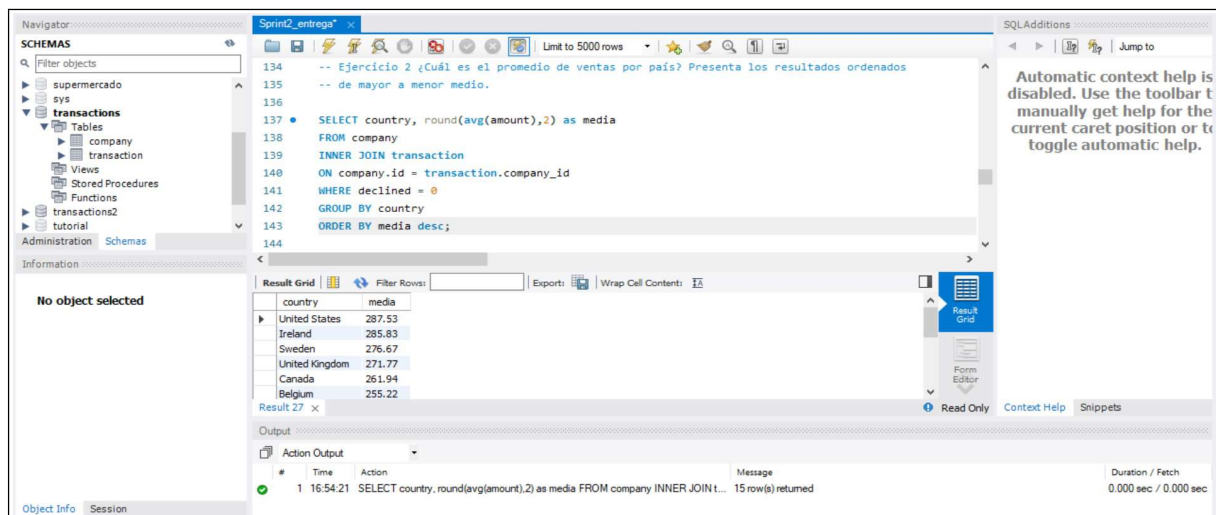


Imagen 11. Captura de pantalla de la ejecución de la consulta.



Para el cálculo de la media, se utiliza la función de agregación average sobre el campo amount. A ese valor se le aplica la función Round para limitar el valor a dos decimales, ya que es una moneda. Se lista el país y la media del amount que se agrupa por País. Ya que los datos requeridos se encuentran en tablas diferentes, se aplica una JOIN.

Por otro lado, como se está hablando de ventas aplicaremos el filtro para que cumplan la condición de declined igual a cero que se ha explicado anteriormente.

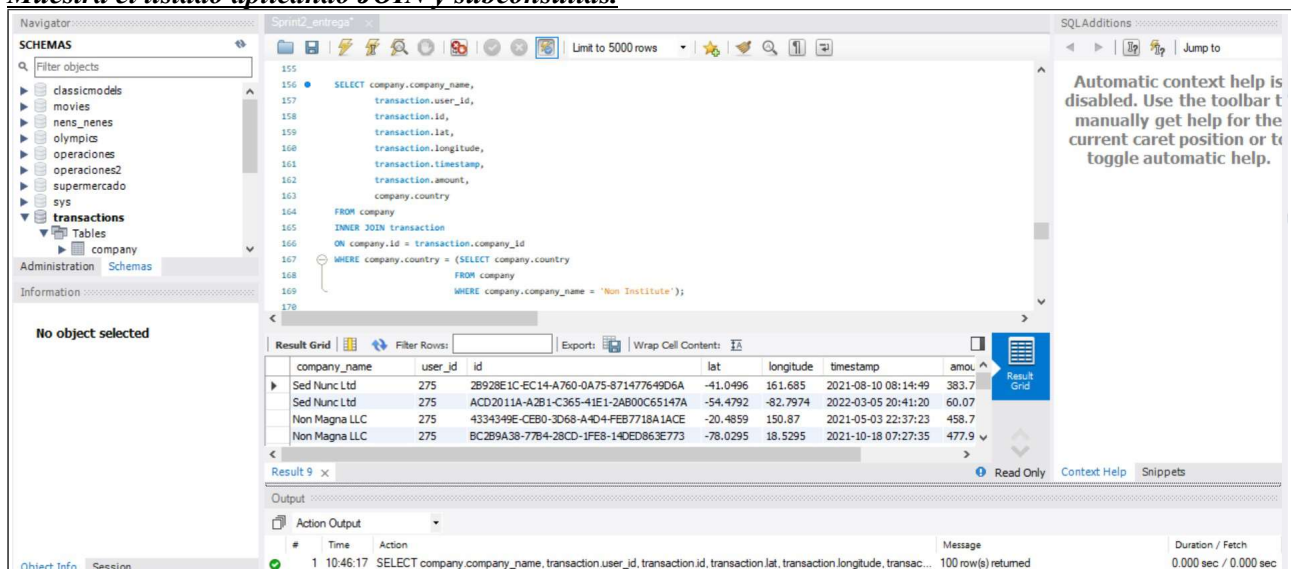
En el caso de las ventas, se puede tener encuesta como venta si las transacciones se han realizado con éxito, y no han sido devueltas. Yo entiendo que si haces una venta, y no te pagan, ésta no se debe tener encuesta cuando se habla de ventas.

Por último, se ordenan por la media del amount, de manera descendiente (ORDER BY 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 situadas en el mismo país que esta compañía.

**Muestra el listado aplicando JOIN y subconsultas.**



```
SELECT company.company_name,
transaction.user_id,
transaction.id,
transaction.lat,
transaction.longitude,
transaction.timestamp,
transaction.amount,
company.country
FROM company
INNER JOIN transaction
ON company.id = transaction.company_id
WHERE company.country = (SELECT company.country
FROM company
WHERE company.company_name = 'Non Institute');
```

company_name	user_id	id	lat	longitude	timestamp	amount
Sed Nunc Ltd	275	2B928E1C-EC14-A760-0A75-871477649D6A	-41.0496	161.685	2021-08-10 08:14:49	383.7
Sed Nunc Ltd	275	ACD2011A-A2B1-C365-41E1-2AB00C65147A	-54.4792	-82.7974	2022-03-05 20:41:20	60.07
Non Magna LLC	275	4334349E-CEB0-3D68-A4D4-FEB7718A1ACE	-20.4859	150.87	2021-05-03 22:37:23	458.7
Non Magna LLC	275	BC2B9A38-77B4-28CD-1FE8-14DED863E773	-78.0295	18.5295	2021-10-18 07:27:35	477.9

Imagen 12. Captura de pantalla de la ejecución de la consulta.

En la SELECT se detallan los campos: nombre de la compañía, id de la transacción, fecha de la transacción y el valor de la venta. Como los datos a buscar están en tablas distintas, se aplica una JOIN.

Se listan los datos que cumplan la condición que el país sea el mismo que el de Non Institute que se obtiene de la subquery.

Tienes razón, no leí bien. Daba por hecho que mi empresa era Non Institute. Se debería incluir en el estudio. Aunque en el enunciado pide las transacciones de todas las empresas que están en el mismo país que Non Institute.

En relación a los datos a incluir en la consulta. En este ejercicio, sólo tenemos dos tablas. Los datos son limitados. Estoy de acuerdo que para un análisis de datos para un estudio de Marketing necesitaríamos más datos.

Incluiría los siguientes campos:

company.company\_name  
transaction.user\_id  
transaction.id,

transaction.lat  
transaction.longitude  
transaction.timestamp  
transaction.amount  
company.country

*Muestra el listado aplicando solamente subconsultas.*

The screenshot shows a SQL IDE interface with a query editor, a schema navigator, and a result grid. The query is as follows:

```
176 SELECT company_name,  
177 user_id,  
178 transaction_id,  
179 lat,  
180 longitude,  
181 timestamp,  
182 amount,  
183 country  
184 FROM company, transaction  
185 WHERE country = (SELECT country  
186 FROM company  
187 WHERE company_name = 'Non Institute')  
188 and company_id = transaction.company_id  
189  
190 -- El resultado es 70 transacciones  
191
```

The result grid displays the following data:

company_name	user_id	id	lat	longitude	timestamp	amount
Sed Nunc Ltd	275	2B928E1C-EC14-A760-0A75-871477649D6A	-41.0496	161.685	2021-08-10 08:14:49	383.7
Sed Nunc Ltd	275	ACD2011A-A2B1-C365-41E1-2AB00C65147A	-54.4792	-82.7974	2022-03-05 20:41:20	60.07
Non Magna LLC	275	4334349E-CEB0-3D68-A4D4-FEB7718A1ACE	-20.4859	150.87	2021-05-03 22:37:23	458.7
Non Magna LLC	275	BC2B9A38-77B4-28CD-1FE8-14DED863E773	-78.0295	18.5295	2021-10-18 07:27:35	477.9

The Action Output pane shows the execution details:

#	Time	Action	Message	Duration / Fetch
1	10:46:17	SELECT company.company_name, transaction.user_id, transaction.id, transaction.lat, transaction.longitude, transaction.timestamp, transaction.amount, transaction.country	100 row(s) returned	0.000 sec / 0.000 sec

Imagen 13. Captura de pantalla de la ejecución de la consulta.

En el caso de sólo utilizar subconsultas, la diferencia con la anterior es que no se utiliza una JOIN, y se añade en el WHERE una condición más que es: **company.id = transaction.company\_id**  
Al utilizar AND, se deben cumplir las tres condiciones a la vez.

## Nivel 3

### Ejercicio 1

Presenta el nombre, teléfono, país, fecha y amount, de aquellas empresas que realizaron transacciones con un valor comprendido entre 100 y 200 euros y en alguna de estas fechas: 29 de abril de 2021, 20 de julio de 2021 y 13 de marzo de 2022. Ordena los resultados de mayor a menor cantidad.

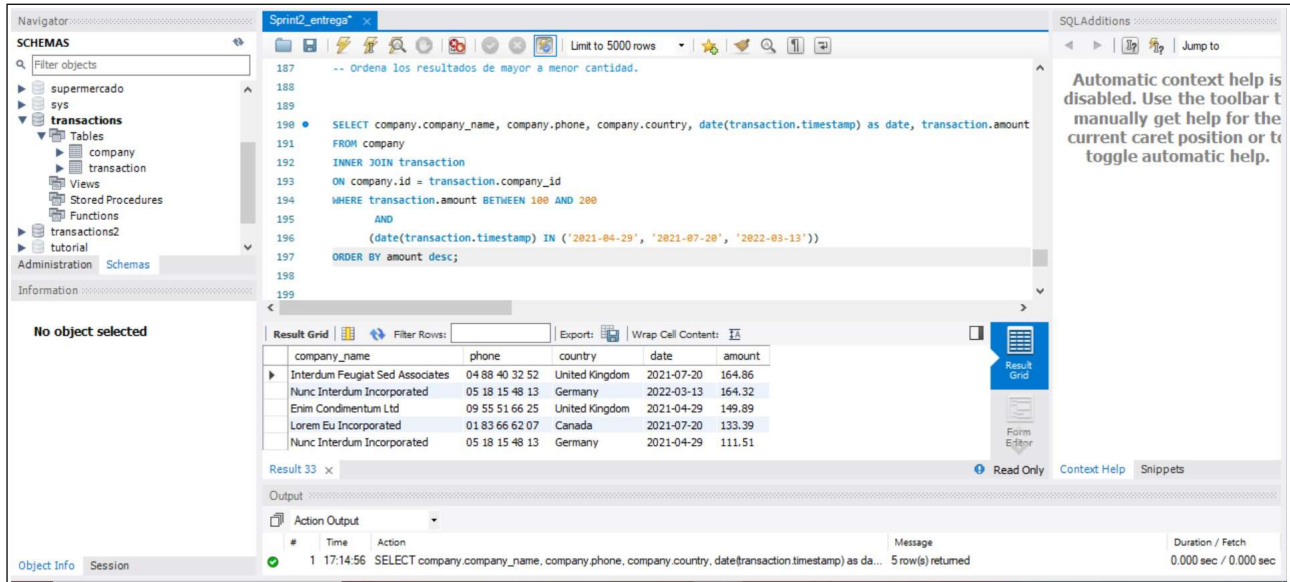


Imagen 14. Captura de pantalla de la ejecución de la consulta.

En esta consulta, se seleccionan los datos requeridos. Se aplica la función date para transformar el formato en solo fecha. Como los datos a listar se encuentran en diferentes tablas, se aplica una JOIN.

Se aplica un WHERE (un filtro) con dos condiciones que se deben cumplir ambas. La primera es un BETWEEN del campo amount definiendo un rango, y la segunda es un IN donde la fecha debe coincidir con una de esas tres fechas.

Para acabar, ordenamos por el campo cuantitativo amount de manera descendiente.

### 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 transacciones o menos.

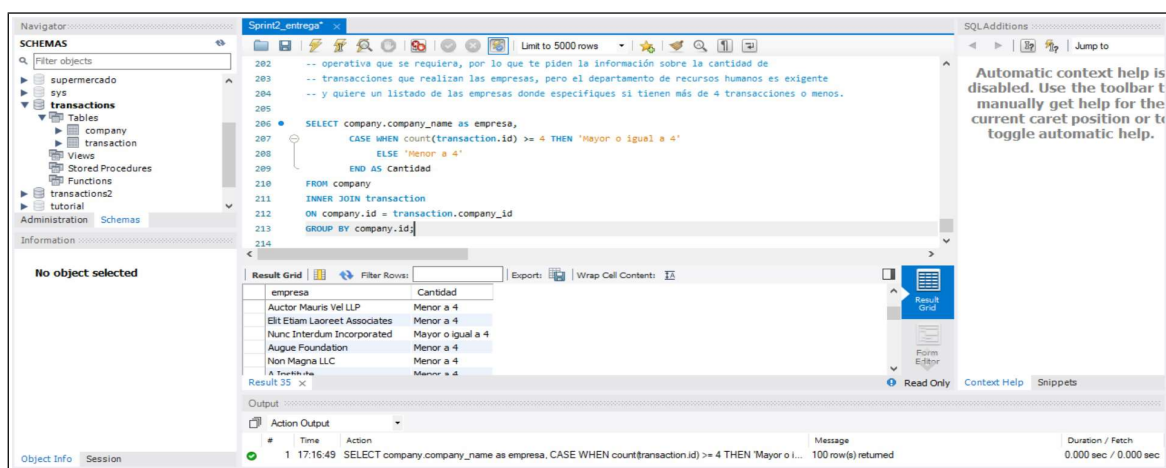


Imagen 15. Captura de pantalla de la ejecución de la consulta.

En esta última consulta, se listan el nombre de la empresa, y una nueva columna que recibe el nombre de Cantidad. Esta columna se define con la función CASE WHEN. Ésta compara el valor del conteo del número de la transacciones con un valor determinado. Si cumple la primera condición escribe, en este caso: 'Mayor o igual a 4', sino se escribe: 'Menor a 4'.