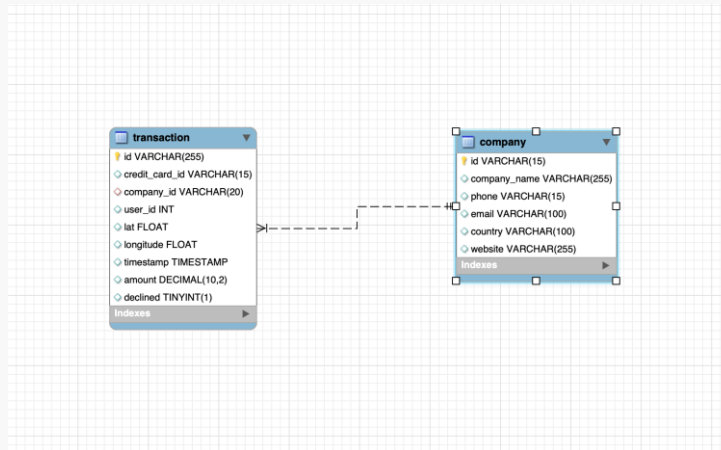


Sprint 2

Tarea S2.01. Nociones básicas SQL

- Ejercicio 1

-A partir de los documentos adjuntos (estructura_datos y datos_introducir), importa las dos tablas. Muestra las principales características del esquema creado y explica las diferentes tablas y variables que existen. Asegúrate de incluir un diagrama que ilustre la relación entre las diferentes tablas y variables



En este diagrama podemos encontrarnos con dos tablas, la tabla TRANSACTION que contiene datos cuantitativos, como, por ejemplo, el campo AMOUNT, cantidad por transacción y otros atributos como TIMESTAMP, LATITUD, LONGITUDE que darán contexto a cada transacción, y luego en la tabla COMPANY contiene datos descriptivos de cada empresa, como el nombre, número, información de contacto, país, entre otros.

La relación de las tablas es de muchas a una, es decir, muchas transacciones las realiza una compañía.

La llave foránea es COMPANY_ID de TRANSACTION que se relaciona con ID de COMPANY, y de esta manera las tablas pueden permanecer independientes entre sí a cambios independientes y no habrá conflictos entre sí.

La tabla de hechos es TRANSACTION porque tiene valores métricos, y la tabla de dimensión es COMPANY porque tiene datos descriptivos proporcionando información.

- Ejercicio 2

Utilizando JOIN realizarás las siguientes consultas:

1.

Llistat dels països que estan fent compres.

```
Select distinct company.country
from company
left Join transaction
on company.id = transaction.company_id;
```

En esta consulta, he realizado un **SELECT DISTINCT** sobre el campo country de la tabla **COMPANY** en el que debo hacer un listado de los países que hacen compras.

Utilizo **SELECT DISTINCT** para que me muestre valores únicos y no repetidos. (Porque hay varias transacciones por país).

Luego como las compras están en la tabla **TRANSACTION** y la información de las compañías y debo utilizar un **JOIN** para relacionarlas y la llave de relación entre ellas es **ID** en **COMPANY** y **COMPANY_ID** en la tabla **TRANSACTION**.

El resultado es:

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

76 12:57:51 Select distinct company.country from company left Join transaction on company.id = transaction.company_id 15 row(s) returned

2.

#Des de quants països es realitzen les compres.

```
Select count(distinct company.country)
from company
Join transaction
on company.id = transaction.company_id;
```

En esta consulta, me pide saber la cantidad de países que han realizado compras, siguiendo con la consulta anterior, en vez de saber que países, ahora quiero saber cuántos países y para ello voy a utilizar:

Select count(distinct company.country)

COUNT para realizar el conteo.

DISTINCT para que no se repitan porque recordemos tenemos varias transacciones por país.

Y volvemos hacer un **JOIN** para relacionar las tablas con su campo en común id y company_id

El resultado es:

count(distinct company.cou...
15

3.

#Identifica la companyia amb la mitjana més gran de vendes.

```
select company.id, company_name, avg(transaction.amount)
from transaction
join company
on transaction.company_id = company.id
group by transaction.company_id
order by avg(transaction.amount) desc
limit 1;
```

En esta consulta, pide realizar la media de la compañía con más ventas, entonces para ello busque donde podría conseguir los datos de ventas y lo encuentro en la tabla **TRANSACTION** en el campo de **AMOUNT**. Entonces una vez allí, debo sacar la media sobre el campo **AMOUNT**. Para ello utilizo la operación **AVG** que saca el promedio total de todas las empresas y en ese mismo tiempo ya se, que como hay muchas transacciones deberé agruparlo por las ID de empresa. Con **GROUP BY** company_id de la tabla **TRANSACTION** Como quiero tener la información completa hago una **JOIN** con **COMPANY** para que a la vez que me muestre el promedio me muestre el **ID** y el **NOMBRE DE LA EMPRESA**. Luego de hacer esta consulta me devolverá los promedios de cada compañía, pero como quiero saber cuál es la que más ventas tiene tengo que ordenarlo y limitarlo. Para ello utilizo el **ORDER BY** para ordenarlo descendientemente de mayor a menor Y el **LIMIT** para que solo me muestre 1 resultado.

El resultado es:

id	company_name	avg(transaction.amount)
b-2398	Eget Ipsum Ltd	473.075000

- Ejercicio 3

Utilizando sólo subconsultas (sin utilizar JOIN):

1.

```
#Mostra totes les transaccions realitzades per empreses d'Alemanya.
Select *
from transaction
) where company_id in (select id
                        from company
                        where country = 'germany');
```

En esta consulta lo primero que hice fue la subconsulta. Como el campo donde están los datos de los países es **COUNTRY** y una **SELECT** del **ID** de **COMPANY** donde en el campo **COUNTRY** sea igual a 'germany' (o a lo que escriba y este en ese campo), para asegurarme que desde la tabla **COMPANY**, solo veo las compañías ALEMANAS.

Luego que ya tengo la subconsulta, selecciono todo de la tabla **TRANSACTION** donde company_id sea igual al id filtrado ya para que aparezcan las empresas de **ALEMANIA**.

El resultado es:

id	credit_card...	company_id	user_id	lat	longitude	timestamp	amount	declined
108B1D1D-5B23-A76C-55EF-C568E49A05DD	CcU-2938	b-2222	275	83.7839	-178.86	2021-07-07 17:43:16	293.57	0
EA2C3281-C9C1-A387-44F8-729FB4B51C76	CcU-2938	b-2222	275	20.2004	-116.84	2021-05-09 10:25:08	119.36	1
0DD2E608-5C9E-D1B3-4999-B99F43AD735A	CcU-2959	b-2234	275	9.68811	130.282	2021-04-17 05:30:17	252.47	1
AB069F53-965E-A2A8-CE06-CA8C4FD92501	CcU-2959	b-2234	275	1.64819	-158.007	2021-04-15 13:37:18	60.99	0
0466A42E-47CF-8D24-FD01-C0B689713128	CcU-4219	b-2302	170	-43.9695	-117.525	2021-07-26 07:29:18	49.53	0
0A476ED9-0C13-1962-F87B-D3563924B539	CcU-4359	b-2302	221	-56.4901	114.801	2022-02-26 20:33:54	430.49	0
122DC333-E19F-D629-DCD8-9C54CF1EBB9A	CcU-4366	b-2302	221	29.6372	-166.173	2021-06-09 06:04:14	172.01	0
135267BA-2E7D-957C-C42C-6450A2B3ED54	CcU-4520	b-2302	210	20.6724	14.9732	2021-12-29 20:38:23	17.97	0
14CAE5B5-8FB1-3E4A-4C85-0EA4167534F4	CcU-4849	b-2302	189	-53.6202	93.0533	2021-12-31 00:29:42	388.04	0
158A3ACB-541C-DBCC-65BD-6373CC67BF1C	CcU-4849	b-2302	183	42.5424	-170.347	2022-03-08 05:02:19	240.29	0
162C7E78-2B6B-7971-A1E4-D2124E732451	CcU-4527	b-2302	210	-69.1381	58.0017	2021-04-11 05:59:18	231.26	0
1717FD6B-ADAD-7082-A748-9112BE892CCC	CcU-4219	b-2302	172	69.4892	-138.411	2021-12-29 16:18:54	249.91	0
1753A288-9FC1-52E6-5C39-A1FFB97B0D3A	CcU-4345	b-2302	222	57.9422	-114.729	2021-08-17 05:32:08	497.84	0
186F53DE-DE27-B1FE-B82F-15B61CEB7726	CcU-4310	b-2302	225	-72.7448	36.6211	2021-12-20 13:13:45	238.16	0
18C4E2DD-1E4C-F35E-2198-C660BD81DC25	CcU-4849	b-2302	177	68.0133	91.4839	2021-09-24 18:55:25	237.04	0
18C9RA7C-ARCI-1A13D-FAF3-4ARA9749936R	CcU-4219	b-2302	173	51.3881	-156.371	2021-06-21 03:21:34	58.16	0

79 13:56:58 Select * from transaction where company_id in (select id from company where country = 'germany') 118 row(s) returned

2.

#Llista les empreses que han realitzat transaccions per un amount superior a la mitjana de totes les transaccions.

```
select company_name
from company
where id in (select company_id
             from transaction
             where amount > (select avg(amount) from transaction));
```

En esta consulta, como debo mostrar un listado de las empresas con su nombre y operar en base a la tabla de TRANSACTION, primero debo hacer las subconsultas.

La primera es sacar el promedio del campo AMOUNT de la tabla TRANSACTION, y el listado de empresas que debo entregar debe ser todas las empresas que estén por encima de la media.

La segunda subconsulta es la que voy a usar como para conectar las tablas y además para dar la condición de que me enseñe las transacciones por encima de la media.

Y por último la consulta entera, conectar la tabla COMPANY y el campo COMPANY_NAME para enseñar la lista conectándolas tablas COMPANY y TRANSACTION a través del ID y el COMPANY_ID.

El resultado es:

company_name
Ac Fermentum Incorporated
Magna A Neque Industries
Fusce Corp.
Ante Iaculis Nec Foundation
Donec Ltd
Sed Nunc Ltd
Nascetur Ridiculus Mus Inc.
Vestibulum Lorem PC
Gravida Sagittis LLP
Mus Aenean Eget Foundation
Dis Parturient Institute
Sed LLC
Arcu LLP
Fringilla LLC
Elit Etiam Laoreet Associates
Nunc Interdum Incorporated
Aliquam Foundation

87 23:36:03 select company_name from company where id in (select company_id from transaction where amount > (select... 70 row(s) returned

3.

#Eliminaran del sistema les empreses que no tenen transaccions registrades, entrega el llistat d'aquestes empreses.

```
Select *
from company
where id not in (select company_id
                from transaction);
```

En esta consulta, sabemos que eliminarian del sistema las empresas que no tienen transacciones, entonces en un método fácil seria llamar a todo de la tabla COMPANY, usando una condición que me enseñe donde el ID, es decir, las empresas, donde NOT IN (no haya información) IN el campo COMPANY_ID de transacción.

Traducción: Muéstrame las empresas que no tengan transacciones.

El resultado es:

id	company_name	phone	email	country	website
NULL	NULL	NULL	NULL	NULL	NULL

En este caso es NULL.