



## **SPRINT 2**

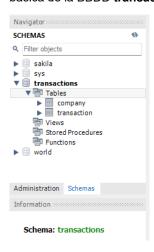
### Nivell 1

#### - 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.

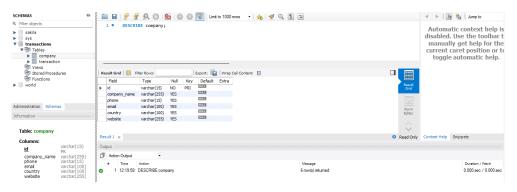
El primer pas es executar primer el doc. estructura\_de\_dades.sql i posteriorment executar el doc. dades\_introduir.sql

Després de crear la base de dades i de importar les dues taules podem veure a primer cop d'ull l'estructura bàsica de la BBDD **transactions**:



Per veure l'estructura de les taules utilitzem la query DESCRIBE, posteriorment exportem les taules per a detallar el contingut dels camps.

## company



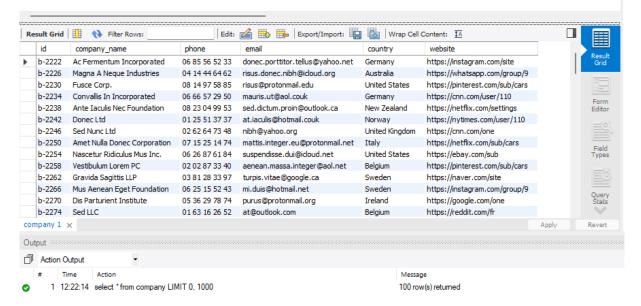
company							
Camp	Tipus	Null	Clau	Valor per defecte	Extra		
id	varchar(15)	NO	PRIMARY KEY		És la Primary Key Conté el id de company		
company_name	varchar(255)	YES			Conté el nom de la company		
phone	varchar(15)	YES			Conté el telèfon format per 8 xifres encapçalades per 0		
email	varchar(100)	YES			Conté la adreça de correu electrònic		
country	varchar(100)	YES			Conté el país de la company		
Website	varchar(255)	YES			Conté la adreça web de la company		

Oriac Gimeno Lozano Reskilling Data Analytics

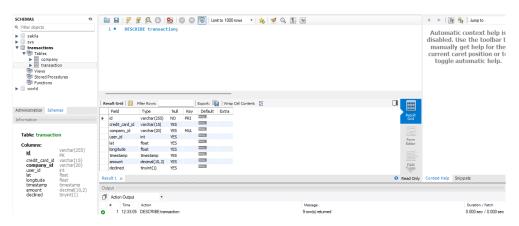




1 • select \* from company



#### transaction

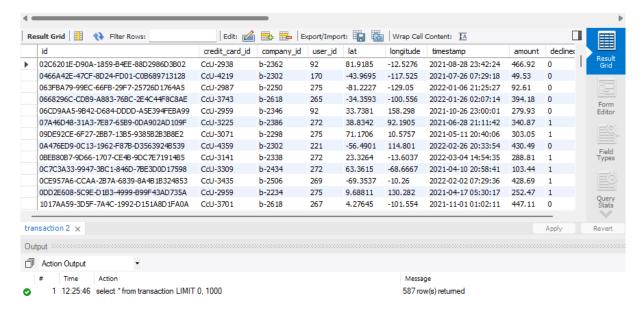


transaction								
Camp	Tipus	Null	Clau	Valor per de- fecte	Extra			
Id	varchar(255)	NO	PRIMARY KEY		Conté el id de la transacció			
credit_card_id	varchar(15)	YES			Conté el id de la targeta de crèdit			
company_id	varchar(20)	YES	FOREING KEY		Conté el id de company			
user_id	vint	YES			Conté el id de usuari			
Lat	float	YES			Conté la latitud de la transacció			
Longitude	float	YES			Conté la longitud de la transacció			
Timestamp	timestamp	YES			Conté el timestamp de la transacció			
Amount	decimal(10,2)	YES			Conté l'import de la transacció amb 2 decimals			
Declined	tinyint(1)	YES			Utilitzat com booleà. 1 = TRUE = Rebutjada 0 = FALSE = Acceptada			





1 • select \* from transaction



Fem el diagrama EER de la base de dades de la següent manera:

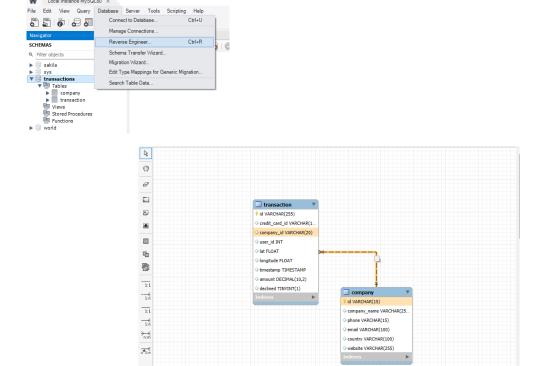


Diagrama EER de la base de dades

Hi ha dues taules amb les seves respectives claus primàries:

- transaction
  - id (VARCHAR(255)) → Clau primària (identifica cada transacció)
- company
  - id (VARCHAR(15)) → Clau primària (identifica cada companyia)

Oriac Gimeno Lozano Reskilling Data Analytics





Relació entre taules:

transaction.company\_id (VARCHAR(20)) està relacionat amb company.id (VARCHAR(15))

És una **relació 1:N** (una empresa pot tenir múltiples transaccions, però cada transacció pertany a una sola empresa).

L'enllaç es fa amb una **clau forana** (**company\_id**) en la taula **transaction**, apuntant a la clau primària **id** de **company**.

#### **Convencions:**

- Per als quèries tractarem les taules amb els àlies per la seva inicial  ${\bf c}$ . per a  ${\bf company}$  i  ${\bf t}$ . per a  ${\bf transaction}$ 
  - u.m. → unitats monetàries
- Exercici 2

Utilitzant JOIN realitzaràs les següents consultes:

2.1.Llistat dels països que estan fent compres.

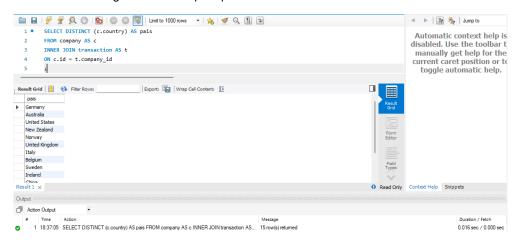
Podem veure el llistat amb els noms dels països que apareixen (15 rows)

**SELECT DISTINCT (c.country)** → Selecciona els països únics que apareixen a la taula.

**FROM company AS c** → Treballem amb la taula company, que conté la informació de les empreses.

**INNER JOIN transaction AS t ON c.id = t.company\_id** → Uneix la taula company amb transaction per obtenir només les empreses que tenen compres registrades.

**DISTINCT** → Ens assegurem de no repetir països a la llista.



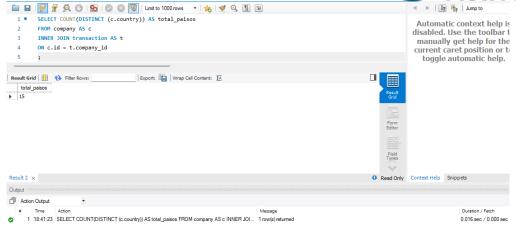
2.2.Des de quants països es realitzen les compres.

**COUNT(DISTINCT (c.country))** → Compta el nombre de països diferents que han realitzat compres.

**INNER JOIN transaction AS t \rightarrow \text{Filtra}** només les empreses que tenen transaccions registrades.







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

**SELECT company\_name**, **ROUND(AVG(amount),2)** → Seleccionem el nom de l'empresa i la mitjana de les seves vendes, arrodonida a 2 decimals.

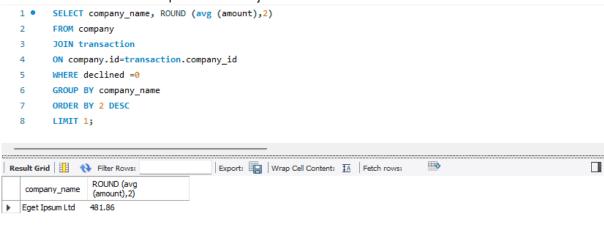
**INNER JOIN transaction** → Connectem les empreses amb les seves transaccions.

WHERE declined = 0 → Ens assegurem que només es comptabilitzen les transaccions aprovades.

**GROUP BY company\_name** → Agrupem les dades per cada empresa per calcular la mitjana individual.

**ORDER BY 2 DESC** → Ordena les empreses per la seva mitjana de vendes de major a menor (la columna 2 correspon a ROUND(AVG(amount),2)).

**LIMIT 1** → Retorna només l'empresa amb la mitjana de vendes més alta.









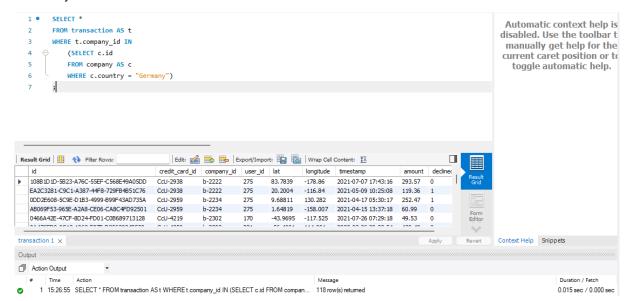
#### - Exercici 3

Utilitzant només subconsultes (sense utilitzar JOIN):

3.1. Mostra totes les transaccions realitzades per empreses d'Alemanya.

WHERE t.company\_id IN (...) → Filtra només les transaccions de les empreses situades a Alemanya.

**SELECT c.id FROM company WHERE c.country = "Germany"** → Retorna els ID de les empreses d'Alemanya.



3.2.Llista les empreses que han realitzat transaccions per un **amount** superior a la mitjana de totes les transaccions. **– a treballar** 

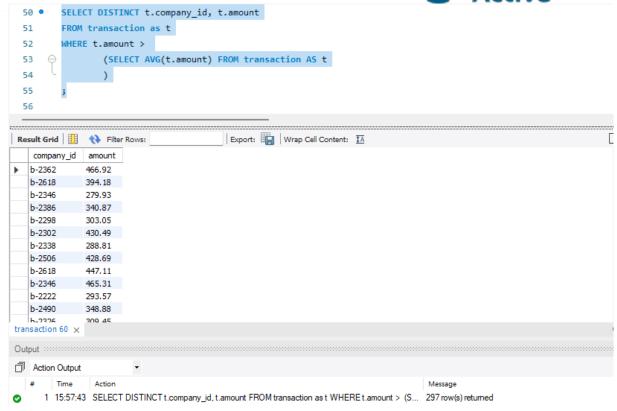
 $AVG(t.amount) \rightarrow Calcula la mitjana de totes les transaccions.$ 

 $\textbf{WHERE t.amount > (SELECT AVG(...))} \rightarrow \textbf{Filtra nom\'es les transaccions que superen aquesta mitjana}.$ 

**DISTINCT** → Evita duplicats a la llista.

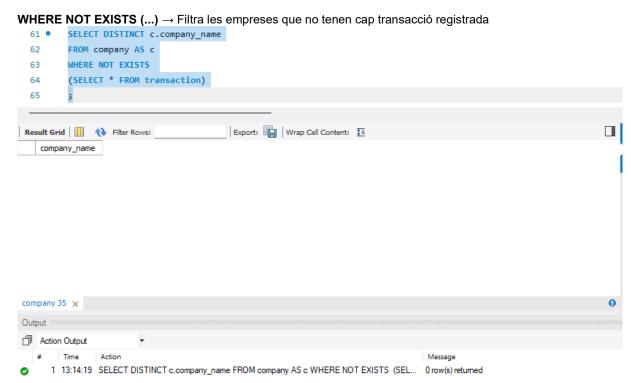






3.3. Eliminaran del sistema les empreses que no tenen transaccions registrades, entrega el llistat d'aquestes empreses.

## No hi ha cap empresa que no tingui ninguna transacció registrada





#### 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.

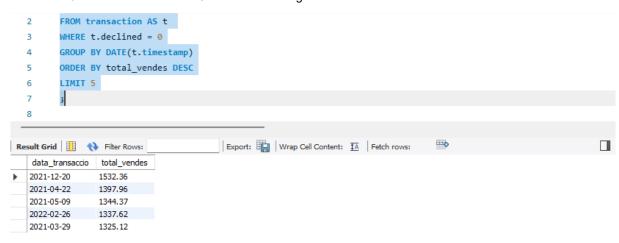
**DATE(t.timestamp)** → Extreu només la data de la transacció.

**SUM(t.amount)** → Calcula el total d'ingressos per cada dia.

**GROUP BY DATE(t.timestamp)** → Agrupa els ingressos per dia.

**ORDER BY total\_vendes DESC**  $\rightarrow$  Ordena de major a menor.

**LIMIT 5** → Retorna només els 5 dies amb més ingressos.









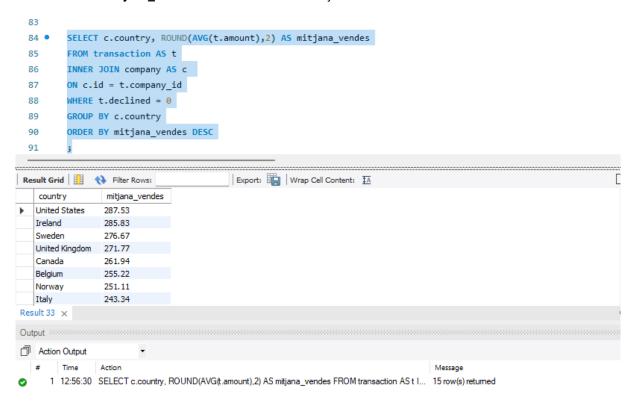
#### Exercici 2

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

**AVG(t.amount)** → Calcula la mitjana de vendes per país.

**GROUP BY c.country** → Agrupa les dades per país.

**ORDER BY mitjana\_vendes DESC** → Ordena de major a menor.



## 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.

3.1. Mostra el llistat aplicant JOIN i subconsultes.

**SELECT \* FROM transaction** → Seleccionem totes les dades de la taula transaction (totes les transaccions).

**INNER JOIN company AS c ON transaction.company\_id = c.id** → Relacionem cada transacció amb la seva empresa mitjançant company id.

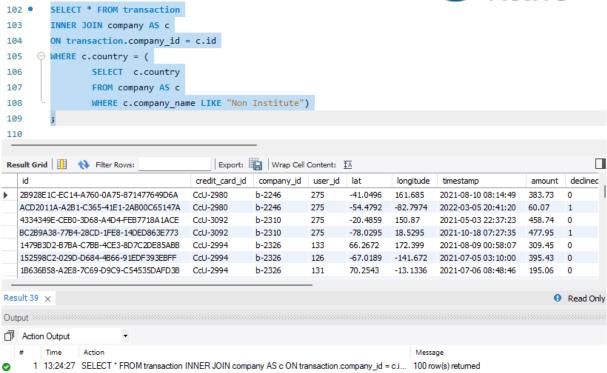
**WHERE c.country = (...)** → Ens quedem només amb les transaccions d'empreses situades en el mateix país que "Non Institute".

(SELECT c.country FROM company AS c WHERE c.company\_name LIKE "Non Institute")  $\rightarrow$  Aquesta subconsulta busca el país de l'empresa "Non Institute"

**LIKE** → Utilitzem LIKE perquè es una cadena de text.

## **IT ACADEMY**





3.2. Mostra el llistat aplicant solament subconsultes.

**SELECT \* FROM transaction AS t**  $\rightarrow$  Selectionem totes les transactions.

WHERE t.company\_id IN (...) → Filtra només les transaccions d'empreses que compleixen una condició.

Subconsulta 1:

SELECT c.id FROM company AS c

WHERE c.country = (...) → Retorna els ID de totes les empreses del país de "Non Institute".

Subconsulta 2:

(SELECT c.country FROM company AS c WHERE c.company\_name LIKE "Non Institute") → Retorna el país de "Non Institute".

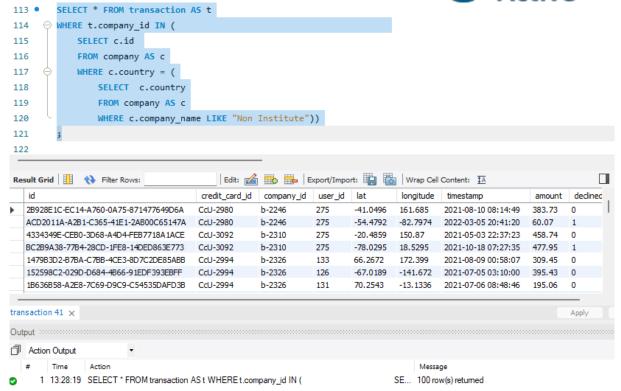
La subconsulta interna busca quin país pertany "Non Institute"

La segona subconsulta selecciona els id de totes les empreses situades en Alemanya.

La consulta principal selecciona totes les transaccions on company\_id coincideix amb algun d'aquests ID.

# **IT ACADEMY**





#### Nivell 3

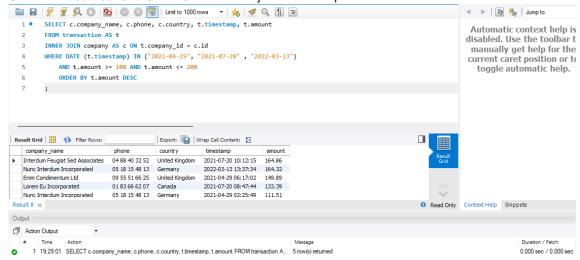
### Exercici 1

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.

WHERE DATE(t.timestamp) IN  $(...) \rightarrow$  Filtra per dates específiques.

AND t.amount >= 100 AND t.amount <= 200 → Filtra només les transaccions dins del rang de 100 a 200 μ m

**ORDER BY t.amount DESC** → Ordena de major a menor import.







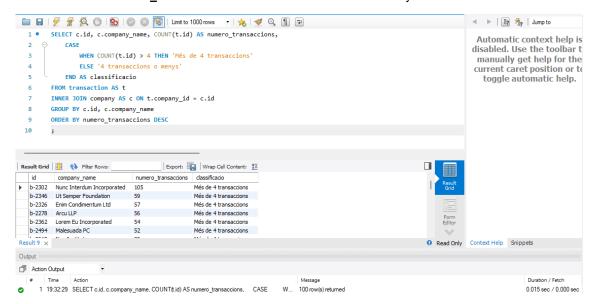
#### Exercici 2

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.

**COUNT(t.id)** → Compta el nombre de transaccions per empresa.

**CASE WHEN** ... **THEN** ... **ELSE** ... **END** → Classifica segons el nombre de transaccions.

**ORDER BY numero\_transaccions DESC** → Ordena de més a menys transaccions.



#### Reptes, dificultats afrontades i aspectes a millorar

- El primer i principal repte ha estat familiaritzar-me amb l'entorn de treball de Workbench i el funcionament del programa en general, ja que la meva experiència prèvia era amb phpMyAdmin i PostgreSQL i he trobat diferències amb aquests dues en les seves funcionalitats així com per exemple per a elaborar el diagrama EER.
- Utilització adequada de INNER JOIN + IN per a obtenir els resultats desitjats
- Jerarquitzar adequadament els conceptes a buscar per a poder fer les seleccions pas a pas incloent escaladament els criteris de cerca.
- Tenir cura en utilitzar els àlies quan correspon a la consulta principal i quan correspon a les subconsultes.
- A les comparacions utilitzar adequadament IN als WHERE i a DATE

Ex: A l'exercici N3 - 3.1 WHERE DATE (t.timestamp) IN ("2021-04-29", "2021-07-20", "2022-03-13")

- Dificultat en l'ús de CASE, WHEN, ELSE i END i en utilitzar-lo adequadament

Ex: A l'exercici Nivell 3 Exercici 2

SELECT c.id, c.company\_name, COUNT(t.id) AS numero\_transaccions,

## CASE

WHEN COUNT(t.id) > 4 THEN 'Més de 4 transaccions'

ELSE '4 transaccions o menys'

**END AS classificacio** 

FROM transaction AS t

INNER JOIN company AS c ON t.company\_id = c.id

GROUP BY c.id, c.company\_name

ORDER BY numero\_transaccions DESC;

Oriac Gimeno Lozano Reskilling Data Analytics