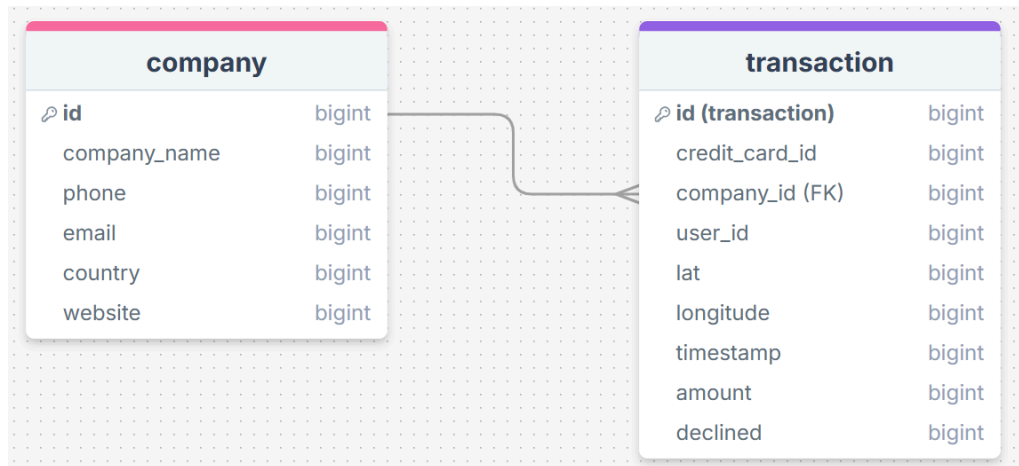


# Tasca S2.01. Nocions bàsiques SQL

## - 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.



**TAULA COMPANY:** es tracta de la taula de dimensions. La seva clau primària (id) es relaciona amb la taula de fets (transaction). A més aquesta relació és d'1 a molts en direcció a la taula transaction el que ens assegura que company és una taula de dimensions on totes les variables són de tipus varchar.

- id: és la clau primària de la taula. És una dimensió tipus varchar, és a dir, una cadena de text variable que pot contenir lletres, números i caràcters especials. Id fa referència al codi d'identificació únic que té cada empresa (no dues empreses amb el mateix codi id).
- company\_name: és una dimensió que fa referència al nom de cada empresa.
- phone: és una dimensió que fa referència al número de telèfon de contacte associat a cada empresa.
- email: és una dimensió que fa referència a l'adreça de correu associat a cada empresa.
- country: és una dimensió que fa referència al país en el que es troba situada cada empresa.
- website: és una dimensió que fa referència a la URL web associada a cada empresa.

# Tasca S2.01. Nocions bàsiques SQL

**TAULA TRANSACTION:** es tracta de la taula de fets o fact table. Aquesta taula en principi hauria d'anar afegint dades per cada transacció que fa cada empresa. Com s'ha comentat abans, la cardinalitat és de molts a 1 ó n a 1 amb la taula company a través de la foreign key o clau forània (company\_id), confirmant aquesta taula com a taula de fets.

- id: és la clau primària de la taula. En aquest cas es tracta d'una dimensió que fa referència al codi associat a cada transacció.
- credit\_card\_id: fa referència al codi de targeta de crèdit utilitzat per realitzar cada transacció, aquests codis poden repetir-se al llarg de la taula, perquè una mateixa targeta pot utilitzar-se per fer diverses transaccions.
- company\_id: es la clau forània de la taula. És una dimensió s'associa amb la clau primària de la taula de dimensions *company* (id) de molts a 1. És a dir, pot haver més d'una transacció vinculada a una mateixa empresa.
- user\_id: és un tipus de variable INT (número enter). És una dimensió que fa referència al codi vinculat a cada usuari que fa una transacció. Aquesta variable, al igual que credit\_card\_id es pot repetir al llarg de la taula.
- lat: és un tipus de dimensió FLOAT (número decimal aproximat). Fa referència a la latitud geogràfica en la que es troba l'empresa.
- longitude: és la mateixa dimensió que abans, en aquest cas fent referència a la longitud geogràfica.
- timestamp: és una dimensió tipus TIMESTAMP que fa referència el moment del dia, mes i any en el que s'ha fet la transacció.
- amount: és una mètrica tipus DECIMAL (número decimal exacte) que fa referència a la quantitat de diners per transacció.
- declined: és una dimensió que es podria entendre com BOOLEAN. Perquè diu si la transacció s'ha declinat o no mitjançant 1 i 0 (que es podria traduir com a True i False, respectivament).

## - Exercici 2

Utilitzant JOIN realitzaràs les següents consultes:

- Llistat dels països que estan fent compres

The screenshot shows a SQL IDE interface. The query editor contains the following SQL code:

```
2 # Exercici 2, Llistat dels països que estan fent compres.
3 • SELECT DISTINCT country
4 FROM company
5 JOIN transaction
6 ON transaction.company_id = company.id;
7
```

Below the query editor, the 'Result Grid' shows the results of the query. The first column is 'country'. The results are:

country
Germany
Australia
United States
New Zealand
Norway
United Kingdom
Italy

The status bar at the bottom indicates that the query was executed successfully at 11:56:24, returning 15 rows in 0.000 seconds.

# Tasca S2.01. Nocions bàsiques SQL

Mitjançant el `SELECT DISTINCT` ens assegurem que estem seleccionant únicament els països sense repetir, que han participat en alguna transacció. Després fem la join amb la taula `transactions` mitjançant la columna `id` (foreign key a `transactions` com `company_id`) per obtenir la informació de les transaccions dels països. Observem que hi ha 15 països diferents que fan compres.

- Des de quants països es realitzen les compres.

```
8 # Des de quants països es realitzen les compres.
9 • SELECT count(distinct country) as païses_distintos
10 FROM company
11 JOIN transaction
12 ON transaction.company_id = company.id;
```

Result Grid

païses_distintos
15

Result 1 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	12:08:38	SELECT count(distinct country) as païses_distintos FROM company JOIN transaction...	1 row(s) returned	0.015 sec / 0.000 sec

En aquest cas, en el `SELECT`, mitjançant la funció d'agregació `count(distinct country)` comptem quants països diferents hi ha entre les empreses que han fet alguna transacció. Necessitem tornar a fer la `JOIN` per obtenir les dades del país al que pertanyen les empreses que han fet transaccions.

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

```
17 # Identifica la companyia amb la mitjana més gran de vendes.
18 • SELECT company.company_name, avg(transaction.amount) as media_ventas
19 FROM company
20 JOIN transaction
21 ON transaction.company_id = company.id
22 WHERE declined = 0
23 group by 1
24 order by media_ventas desc
25 limit 1;
```

Result Grid

company_name	media_ventas
Eget Ipsum Ltd	481.860000

Result 2 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	12:08:38	SELECT count(distinct country) as païses_distintos FROM company JOIN transaction...	1 row(s) returned	0.015 sec / 0.000 sec
2	12:10:04	SELECT company.company_name, avg(transaction.amount) as media_ventas FROM...	1 row(s) returned	0.016 sec / 0.000 sec

## Tasca S2.01. Nocions bàsiques SQL

En aquest cas, demanem el nom de les empreses i la funció d'agregació `avg(amount)` treu el promig de les transaccions de cada empresa. Un fet important es que en el filtre `where` estem considerant **només** les transaccions efectives, és a dir aquelles que apareixen a la taula com a `declined = 0` (considerant que el 0 vindria a ser un valor booleà FALS). Ordenem pel nombre de ventes (àlies que li hem donat al `average del amount`) de major a menor i limitem el primer resultat, és a dir el top 1 o la empresa que té una mitjana de valor de transacció més gran.

### - Exercici 3

Utilitzant només subconsultes (sense utilitzar JOIN)

- Mostra totes les transaccions realitzades per empreses d'Alemanya

```
28 # EXERCICI 3
29 # Mostra totes les transaccions realitzades per empreses d'Alemanya.
30 • SELECT *
31 FROM transaction
32 WHERE declined = 0 and
33     company_id IN (
34     SELECT id
35     FROM company
36     WHERE country = "Germany");
```

id	credit_card_id	company_id	user_id	lat	longitude	timestamp	amount	declined
10881D1D-5B23-A76C-55EF-C568E49A05DD	CcU-2938	b-2222	275	83.7839	-178.86	2021-07-07 17:43:16	293.57	0
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

transaction 1 x

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	21:04:10	SELECT * FROM transaction WHERE declined = 0 and company_id IN (SELECT...	111 row(s) returned	0.000 sec / 0.000 sec

El `SELECT *` from `transaction` selecciona totes les columnes de cada transacció que es realitza. Tornem a filtrar per aquelles transaccions que no s'han declinat i filtrem el codi `company_id` per la instrucció de la subquery. La subquery troba els id de totes les empreses de la taula `company` que són d'Alemanya.

En aquest cas, potser no caldria afegir el filtre `WHERE declined = 0` perquè l'enunciat no especifica si la transacció ha estat efectiva o no, però per mantenir una continuïtat amb els exercicis, s'ha mantingut. Hi ha 111 transaccions realitzades per empreses d'Alemanya.

## Tasca S2.01. Nocions bàsiques SQL

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

```
39 # Llista les empreses que han realitzat transaccions per un amount superior a la mitjana de totes les transaccions.
40 • SELECT company_name
41 FROM company
42 WHERE id IN (
43     SELECT company_id
44     FROM transaction
45     WHERE amount > (
46         SELECT AVG(amount)
47         FROM transaction
48     );
```

Result Grid

company_name
Ac Fermentum Incorporated
Magna A Neque Industries
Fusce Corp.
Ante Iaculis Nec Foundation
Donec Ltd

company 4 x Read Only

Output

Action Output

#	Time	Action	Message	Duration / Fetch
3	12:12:56	SELECT * FROM transaction WHERE company_id IN ( SELECT id FROM com...	118 row(s) returned	0.000 sec / 0.000 sec
4	12:22:25	SELECT company_name FROM company WHERE id IN ( SELECT company_id ...	70 row(s) returned	0.000 sec / 0.000 sec

En aquest cas seleccionem el nom de la empresa de la taula companyia i filtrem per una subquery anidada, en la que la subquery interior selecciona els valors promig del valor de les transaccions. La subquery de fora, filtra aquelles empreses, les quals els valors de les transaccions són majors que la mitjana del amount. Hi ha 70 empreses que han realitzat transaccions per valors superiors a la mitjana de totes les transaccions.

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

```
51 # Eliminaren del sistema les empreses que no tenen transaccions registrades, entrega el llistat d'aquestes empreses.
52 • SELECT company_name
53 FROM company
54 WHERE id NOT IN (
55     SELECT company_id
56     FROM transaction);
57
58
```

Result Grid

company_name
--------------

company 6 x Read Only

Output

Action Output

#	Time	Action	Message	Duration / Fetch
5	12:24:55	SELECT company_name FROM company WHERE id NOT IN ( SELECT company...	0 row(s) returned	0.000 sec / 0.000 sec
6	12:25:04	SELECT company_name FROM company WHERE id NOT IN ( SELECT company...	0 row(s) returned	0.000 sec / 0.000 sec

# Tasca S2.01. Nocions bàsiques SQL

En aquest cas seleccionem el nom de les empreses que NO es troben dins la taula transacció. És a dir amb el filtre WHERE NOT IN ens assegurem d'estar agafant els valors de les empreses que es troben a la taula company però no a la transaction. És a dir que existeixen però no han fet cap transacció. Com observem al resultat, no hi ha cap empresa registrada que no hagi fet almenys una transacció.

## 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.

```
59 # Nivell 2
60 # 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 ju
61 • SELECT DATE(timestamp) as dia, sum(amount) as ingresos
62 FROM transaction
63 WHERE declined = 0
64 GROUP BY dia
65 ORDER BY ingresos desc
66 LIMIT 5;
67
```

dia	ingresos
2021-12-20	1532.36
2021-04-22	1397.96
2021-05-09	1344.37
2022-02-26	1337.62
2021-03-29	1325.12

Result 7

Output

Action Output

#	Time	Action	Message	Duration / Fetch
6	12:25:04	SELECT company_name FROM company WHERE id NOT IN ( SELECT company...	0 row(s) returned	0.000 sec / 0.000 sec
7	12:30:39	SELECT DATE(timestamp) as dia, sum(amount) as ingresos FROM transaction WH...	5 row(s) returned	0.000 sec / 0.000 sec

En aquest exercici, s'introdueix una nova fórmula al SELECT. Aquesta fórmula es DATE i ens serveix per donar-li un format YYYY-MM-DD al camp timestamp. D'aquesta manera SQL pot comparar i filtrar dates sense problema i li diem "dia". A continuació, fem una funció d'agregació SUM de la quantitat i li donem un àlies com "ingresos". Agrupem pel dia i ordenem de major a menor, mostrant només els 5 primers resultats.

# Tasca S2.01. Nocions bàsiques SQL

## Exercici 2

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

```
68 # Quina és la mitjana de vendes per país? Presenta els resultats ordenats de major a menor mitjà.
69 • SELECT company.country, avg(amount) as ventas
70 FROM transaction
71 JOIN company on transaction.company_id = company.id
72 WHERE declined = 0
73 GROUP BY company.country
74 ORDER BY ventas desc;
75
```

country	ventas
United States	287.531111
Ireland	285.825357
Sweden	276.668382
United Kingdom	271.767527
Canada	261.941930

Result 1 x

Output

#	Time	Action	Message	Duration / Fetch
1	10:27:01	SELECT company.country, avg(amount) as ventas FROM transaction JOIN company...	15 row(s) returned	0.016 sec / 0.000 sec

En aquest exercici, seleccionem el país i la quantitat mitjana del amount de les transaccions. En aquest cas, haurem de tornar a fer una JOIN, ja que aquests dos camps es troben en taules diferents. Agruparem llavors per país i ordenarem de major a menor per la quantitat de les vendes. Observem que el resultat ens dona 15 països, tants com hi ha de diferents a la taula company.

# Tasca S2.01. Nocions bàsiques SQL

## 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.

```
77 # En la teva empresa, es planteja un nou projecte per a llançar algunes campanyes publicitàries per a fer competència a la companyia "Non I
78 # fet amb joins i subqueries
79 • SELECT *
80 FROM transaction
81 JOIN company on
82     transaction.company_id = company.id
83 WHERE declined = 0 and
84     company.country in (
85     SELECT country
86     FROM company
87     WHERE company_name = "Non Institute");
```

id	credit_card_id	company_id	user_id	lat	longitude	timestamp	amount	declined	id	company_name
2B928E1C-EC14-A760-0A75-871477649D6A	CcU-2980	b-2246	275	-41.0496	161.685	2021-08-10 08:14:49	383.73	0	b-2246	Sed Nunc Ltd
4334349E-CEB0-3D68-A4D4-FEB7718A1ACE	CcU-3092	b-2310	275	-20.4859	150.87	2021-05-03 22:37:23	458.74	0	b-2310	Non Magna LLC
1479B3D2-87BA-C7BB-4CE3-8D7C2DE85ABB	CcU-2994	b-2326	133	66.2672	172.399	2021-08-09 00:58:07	309.45	0	b-2326	Enim Condimentum Ltd
152598C2-029D-D684-4B66-91EDF393EBFF	CcU-2994	b-2326	126	-67.0189	-141.672	2021-07-05 03:10:00	395.43	0	b-2326	Enim Condimentum Ltd

Result 9 x

Output

#	Time	Action	Message	Duration / Fetch
8	12:32:23	select company.country, avg(amount) as ventas from transaction join company on tr...	15 row(s) returned	0.000 sec / 0.000 sec
9	12:35:05	SELECT * FROM transaction JOIN company on transaction.company_id = compa...	93 row(s) returned	0.000 sec / 0.000 sec

En aquest exercici, seleccionem tots els camps de la taula transaction i realitzem una JOIN amb la taula company per accedir als valors de country i company\_name que ens demana l'enunciat. A la subquery filtrem pel país que correspon a la empresa "Non Institute". D'aquesta manera, quan el resultat de la subquery ens filtra per "United Kingdom", quan s'executa la query, agafarà només els valors de les transaccions del país que ha filtrat la nostra subquery.



## Tasca S2.01. Nocions bàsiques SQL

- Mostra el llistat aplicant solament subconsultes.

```
90 • SELECT *
91 FROM transaction
92 WHERE declined = 0 AND
93 company_id IN (
94     SELECT id
95     FROM company
96     WHERE country = (
97         SELECT country
98         FROM company
99         WHERE company_name = "Non Institute")
100 )
```

id	credit_card_id	company_id	user_id	lat	longitude	timestamp	amount	declined
2B928E1C-EC14-A760-0A75-871477649D6A	CcU-2980	b-2246	275	-41.0496	161.685	2021-08-10 08:14:49	383.73	0
4334349E-CEB0-3D68-A4D4-FEB7718A1ACE	CcU-3092	b-2310	275	-20.4859	150.87	2021-05-03 22:37:23	458.74	0
1479B3D2-B7BA-C7BB-4CE3-8D7C2DE85ABB	CcU-2994	b-2326	133	66.2672	172.399	2021-08-09 00:58:07	309.45	0
152598C2-029D-D684-4B66-91EDF393EBFF	CcU-2994	b-2326	126	-67.0189	-141.672	2021-07-05 03:10:00	395.43	0
1B636B58-A2E8-7C69-D9C9-C54535DAFD3B	CcU-2994	b-2326	131	70.2543	-13.1336	2021-07-06 08:48:46	195.06	0

transaction 10 x

Output

#	Time	Action	Message	Duration / Fetch
9	12:35:05	SELECT * FROM transaction JOIN company on transaction.company_id = compa...	93 row(s) returned	0.000 sec / 0.000 sec
10	12:38:12	SELECT * FROM transaction WHERE declined = 0 AND company_id IN ( SELECT ...	93 row(s) returned	0.000 sec / 0.000 sec

Ara, el mateix exercici només aplicant subqueries té un funcionament molt similar, només afegint una subquery anidada. La subquery de més a dins segueix seleccionant el país que correspon a la empresa "Non Institute". Ara, la subquery de fora, fa la funció que feia la join en l'apartat anterior. És a dir, filtra les transaccions fetes les empreses les quals el company\_id es troba inclòs en la llista de països filtrat per la subquery. Per últim a la query es seleccionen tots els camps de la taula transaction, igual que abans.

# Tasca S2.01. Nocions bàsiques SQL

## 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.

```
102 # Nivell 3
103 # Presenta el nom, telèfon, país, data i amount, d'aquelles empreses que van realitzar transaccions amb un valor comprès entre 100 i 200 eu
104 • SELECT c.company_name, c.phone, c.country, date(t.timestamp) as fecha, t.amount
105 FROM company c
106 JOIN transaction t on c.id=t.company_id
107 WHERE amount between 100 and 200
108 and declined = 0
109 and date(t.timestamp) in ("2021-04-29", "2021-07-20", "2022-03-13")
110 ORDER BY amount desc;
```

company_name	phone	country	fecha	amount
Interdum Feugiat Sed Associates	04 88 40 32 52	United Kingdom	2021-07-20	164.86
Nunc Interdum Incorporated	05 18 15 48 13	Germany	2022-03-13	164.32
Enim Condimentum Ltd	09 55 51 66 25	United Kingdom	2021-04-29	149.89
Lorem Eu Incorporated	01 83 66 62 07	Canada	2021-07-20	133.39
Nunc Interdum Incorporated	05 18 15 48 13	Germany	2021-04-29	111.51

Result 1 x Read Only

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	11:05:06	SELECT c.company_name, c.phone, c.country, date(t.timestamp) as fecha, t.amount ...	5 row(s) returned	0.000 sec / 0.000 sec

En aquest exercici, seleccionem els camps que ens demana l'enunciat al SELECT. En aquest cas, per fer el codi més llegible, s'ha abreviat la taula company com "c" i la taula transaction com "t". D'aquesta manera quan ens referim als camps de les taules ho podem buscar com "t.camp" o "c.camp". Haurem de fer la mateixa join dels exercicis anteriors per poder accedir a els valors que ens demanen. Ara, per filtrar, ens demana transaccions entre dos valors. Per fer-ho fem servir la funció "BETWEEN". També ens demana que aquestes transaccions siguin entre tres dates. Per fer-ho fem servir la funció date com abans i IN per filtrar per aquells valors en concret. Ho hem ordenat per quantitat de transacció de major a menor.

# Tasca S2.01. Nocions bàsiques SQL

## 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.

```
122 # Necessitem optimitzar l'assignació dels recursos i dependrà de la capacitat operativa que es requereixi, per la qual cosa et demanen la i
123 • SELECT company_id, count(*) as num_transacciones,
124 CASE
125     WHEN count(*) >= 4 THEN "mayor/igual a 4"
126     ELSE "menor a 4"
127 END AS especificación_empresas
128 FROM transaction
129 WHERE declined = 0
130 GROUP BY company_id
131 ORDER BY num_transacciones desc;
```

company_id	num_transacciones	especificación_empresas
b-2302	104	mayor/igual a 4
b-2346	58	mayor/igual a 4
b-2326	56	mayor/igual a 4
b-2278	55	mayor/igual a 4
b-2362	53	mayor/igual a 4

Result 2 x

Output

#	Time	Action	Message	Duration / Fetch
1	11:05:06	SELECT c.company_name, c.phone, c.country, date(timestamp) as fecha, t.amount ...	5 row(s) returned	0.000 sec / 0.000 sec
2	11:08:05	SELECT company_id, count(*) as num_transacciones, CASE WHEN count(*) >= 4 T...	100 row(s) returned	0.016 sec / 0.000 sec

En aquest exercici, introduïm una nova funció anomenada CASE. Seleccionem l'identificador de l'empresa i fem un comptatge de les transaccions que realitzen (no necessari per l'exercici però va bé per comprovar que l'script funciona). La funció CASE agafa el nombre de transaccions i li posa una condició: en el cas que aquest nombre sigui major o igual a 4 es classificarà l'empresa en "mayor/igual a 4". En qualsevol altre cas, (inferiors a 4) es classificarà com "menor/igual a 4". Agrupem per l'id de l'empresa perquè la funció COUNT(\*) pugui comptar quantes transaccions fa cada empresa i ordenem pel nombre de transaccions de major a menor.