# SPRINT 4

**Edoardo Brega** 



 Descàrrega els arxius CSV, estudia'ls i dissenya una base de dades amb un esquema d'estrella que contingui, almenys 4 taules de les quals puguis realitzar les següents consultes:

Realitza una subconsulta que mostri tots els usuaris amb més de 30 transaccions utilitzant almenys 2 taules.

#### creo database y tablas

```
create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

create database sprint 4 visual; use sprint 4 visual;

index product data varchar (100), amount double, declined int,

product data varchar (100), index id, key (card_id),

create TABLE companies (company_id) varchar (100), pr. money_id) varchar (100), pr. money_id), pr. money_id)
```

## \* excecuto comprobando que no haya errores

```
1 • create database sprint_4_visual; use sprint_4_visual;
   2 • \odot CREATE TABLE transactions (id varchar (200),card_id varchar (100) not null,business_id varchar (100),timestamp varchar (100),amount double,declined int,
          product_ids varchar (100),user_id varchar (100),lat double,longitude double,index company_key (business_id),index card_key (card_id),
          index product_ids_key (product_ids),index user_key(user_id),PRIMARY KEY id (id));
   5 • CREATE TABLE companies (company_id varchar (200) not null,company_name varchar (100),phone varchar (100),email varchar (100),country varchar (100),
         website varchar (100), PRIMARY KEY company_id (company_id), INDEX id_key (company_id), FOREIGN KEY (company_id) REFERENCES transactions(business_id));
   7 • ⊝ CREATE TABLE credit cards (id varchar (100) not null, user id varchar (100), jban varchar (100), pan varchar (100), pin int, cvv varchar (100), track1 varchar (100),
          track2 varchar (100),expiring_date varchar (100),PRIMARY KEY (id),INDEX id_key (id),FOREIGN KEY (id) REFERENCES transactions(card_id));
   9 • 🔆 CREATE TABLE products (id varchar (100) not null, product_name varchar (100), price varchar (100), colour varchar (100), weight double, warehouse_id varchar (100),
        PRIMARY KEY (id), index id_key(id));
Output :::
Action Output
     1 23:18:38 create database sprint_4_visual
                                                                                                        1 row(s) affected
                                                                                                                                                                                             0.000 sec
                                                                                                                                                                                             0.000 sec
2 23:18:38 use sprint_4_visual
                                                                                                       0 row(s) affected
      3 23:18:38 CREATE TABLE transactions (id varchar (200),card_id varchar (100) not null,business_id varchar (100),times... 0 row(s) affected
                                                                                                                                                                                             0.031 sec
4 23:18:38 CREATE TABLE companies (company_id varchar (200) not null.company_name varchar (100).phone varch... 0 row(s) affected
                                                                                                                                                                                             0.031 sec
      5 23:18:38 CREATE TABLE credit_cards (id varchar (100) not null,user_id varchar (100),iban varchar (100),pan varchar... 0 row(s) affected
                                                                                                                                                                                             0.031 sec
6 23:18:38 CREATE TABLE products (id varchar (100) not null product_name varchar (100) price varchar (100), colour v... 0 row(s) affected
                                                                                                                                                                                             0.032 sec
      7 23:18:39 CREATE TABLE user ca (id varchar (100) not null name varchar (100) sumame varchar (100) phone varcha... 0 row(s) affected
                                                                                                                                                                                             0.031 sec
8 23:18:39 CREATE TABLE user_usa (id varchar (100) not null name varchar (100),sumame varchar (100),phone varch... 0 row(s) affected
                                                                                                                                                                                             0.031 sec
      9 23:18:39 CREATE TABLE user_uk (id varchar (100) not null name varchar (100), sumame varchar (100), phone varcha... 0 row(s) affected
                                                                                                                                                                                             0.031 sec
```

#### código:

create database sprint\_4\_visual; use sprint\_4\_visual;

CREATE TABLE transactions (id varchar (200),card\_id varchar (100) not null,business\_id varchar (100),timestamp varchar (100),amount double,declined int, product\_ids varchar (100),user\_id varchar (100),lat double,longitude double,index company\_key (business\_id),index card\_key (card\_id), index product\_ids\_key (product\_ids),index user\_key(user\_id),PRIMARY KEY id (id));

CREATE TABLE companies (company\_id varchar (200) not null,company\_name varchar (100),phone varchar (100),email varchar (100),country varchar (100), website varchar (100),PRIMARY KEY company\_id (company\_id),INDEX id\_key (company\_id),FOREIGN KEY (company\_id) REFERENCES transactions(business\_id));

CREATE TABLE credit\_cards (id varchar (100) not null,user\_id varchar (100),iban varchar (100),pan varchar (100),pin int,cvv varchar (100),track1 varchar (100), track2 varchar (100),expiring\_date varchar (100),PRIMARY KEY (id),INDEX id\_key (id),FOREIGN KEY (id) REFERENCES transactions(card\_id));

CREATE TABLE products (id varchar (100) not null,product\_name varchar (100),price varchar (100),colour varchar (100),weight double,warehouse\_id varchar (100), PRIMARY KEY (id),index id\_key(id));

CREATE TABLE user\_ca (id varchar (100) not null,name varchar (100),surname varchar (100),phone varchar (100),email varchar (100),birth\_date varchar (100), country varchar (100),city varchar (100),postal\_code varchar(100),address varchar (100),PRIMARY KEY (id),index id key(id));

CREATE TABLE user\_usa (id varchar (100) not null,name varchar (100),surname varchar (100),phone varchar (100),email varchar (100),birth\_date varchar (100), country varchar (100),city varchar (100),postal\_code varchar(100),address varchar (100),PRIMARY KEY (id).index id key(id));

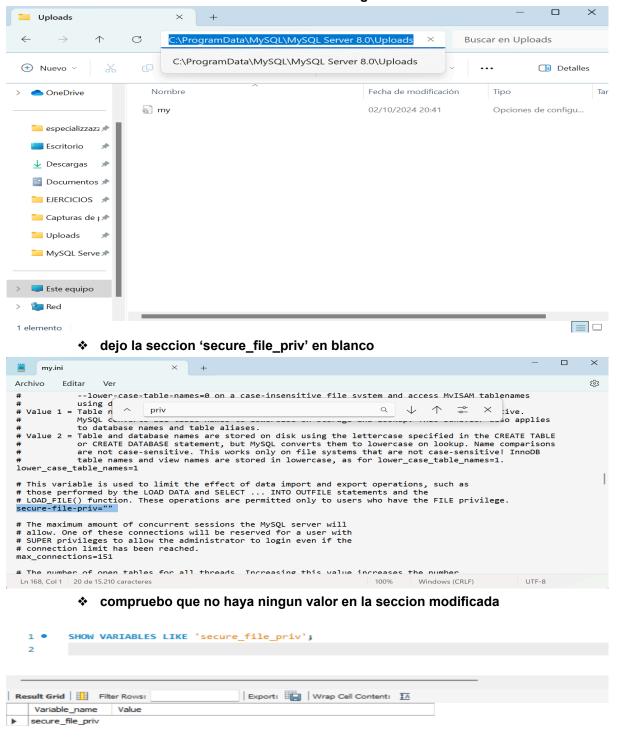
CREATE TABLE user\_uk (id varchar (100) not null,name varchar (100),surname varchar (100),phone varchar (100),email varchar (100),birth\_date varchar (100), country varchar (100),city varchar (100),postal\_code varchar(100),address varchar (100),PRIMARY KEY (id),index id key(id));

## > explicacion:

Primero creo el database y luego creo las tablas, empezando por la tabla de hechos que será 'transactions'.

No añado la 'foreign key' en las tablas 'products, user\_ca, user\_usa, user\_uk' porque los valores en los archivos darían error al intentar de insertarlos manualmente.

busco el file my.ini para desactivar la seccion 'secure\_file\_priv' para poder insertar los dados directamente con codigo



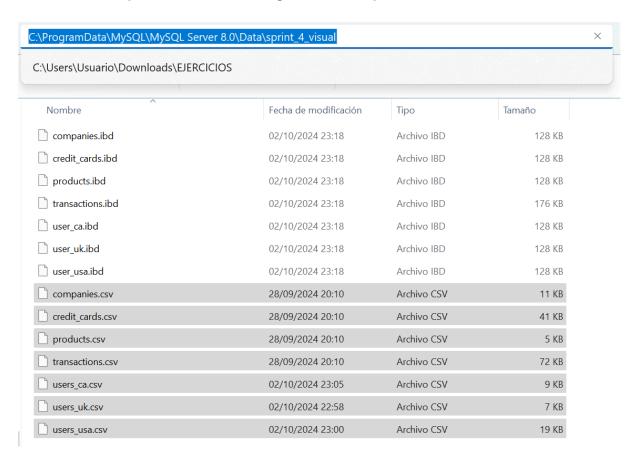
1 row(s) returned

Result 3 ×

Action Output

1 21:00:13 SHOW VARIABLES LIKE secure file privi

## copio los archivos descargados en la carpeta del database creada



#### inserto los valores manualmente



#### > código:

LOAD DATA INFILE 'transactions.csv'INTO TABLE transactions FIELDS TERMINATED BY ',' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA INFILE 'companies.csv' INTO TABLE companies FIELDS TERMINATED BY ',' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA INFILE 'credit\_cards.csv' INTO TABLE credit\_cards FIELDS TERMINATED BY ',' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA INFILE 'products.csv' INTO TABLE products FIELDS TERMINATED BY ',' LINES TERMINATED BY '\n' IGNORE 1 LINES;

LOAD DATA INFILE 'users\_ca.csv' INTO TABLE user\_ca FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY ""
LINES TERMINATED BY '\r' IGNORE 1 LINES;

LOAD DATA INFILE 'users\_usa.csv' INTO TABLE user\_usa FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY ""
LINES TERMINATED BY '\r' IGNORE 1 LINES;

LOAD DATA INFILE 'users\_uk.csv' INTO TABLE user\_uk FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY ""
LINES TERMINATED BY '\r' IGNORE 1 LINES;

## > explicacion:

#### LOAD DATA INFILE '...' INTO TABLE ..

Es para seleccionar el file con los valores y la tabla donde subirlos

#### FIELDS TERMINATED BY '...'

Indico por que caracter estan separados los campos del file seleccionado

## LINES TERMINATED BY '...'

Indico por que caracter estan separadas las lineas del file seleccionado

# IGNORE ... LINES

Indico cuantas lineas no se tienen que subir a la tabla

#### OPTIONALLY ENCLOSED BY '...'

Es para no tomar los caracteres dentro del caracter seleccionado como si fueran delimitadores de campo.

Para los 3 archivos de utentes ha sido necesario tambien eliminar manualmente la ultima linea blanca de estos files.

## \* añado las foreign keys que me faltaban

```
• SET FOREIGN_KEY_CHECKS=0;

2

3 • ALTER TABLE products ADD CONSTRAINT fkproducts FOREIGN KEY (id)REFERENCES transactions(product_ids);

4

5 • ALTER TABLE user_ca ADD CONSTRAINT fk_user_ca FOREIGN KEY (id)REFERENCES transactions(user_id);

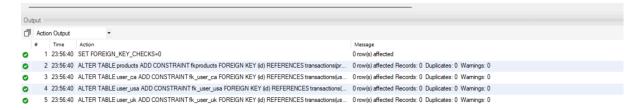
6

7 • ALTER TABLE user_usa ADD CONSTRAINT fk_user_usa FOREIGN KEY (id)REFERENCES transactions(user_id);

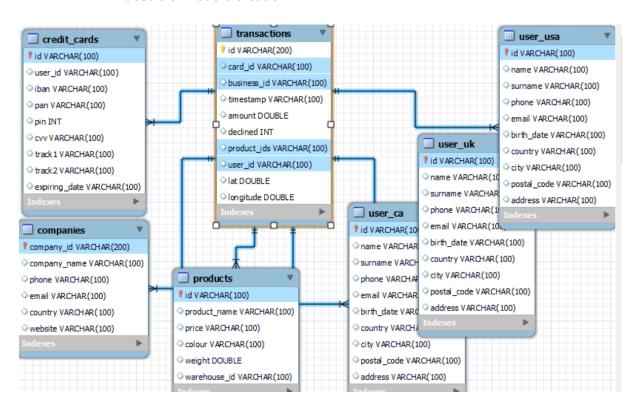
8

9 • ALTER TABLE user_uk ADD CONSTRAINT fk_user_uk FOREIGN KEY (id)REFERENCES transactions(user_id);

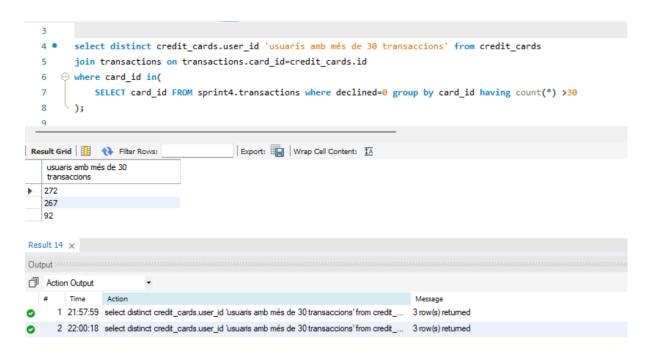
10
```



#### \* muestro el modelo creado



#### selecciono los usuarios con mas de 30 transacciones



## > código:

select distinct credit\_cards.user\_id 'usuaris amb més de 30 transaccions' from credit\_cards join transactions on transactions.card\_id=credit\_cards.id where card\_id in(

SELECT card\_id FROM transactions where declined=0 group by card\_id having

SELECT card\_id FROM transactions where declined=0 group by card\_id having count(\*) >30 );

## > explicación:

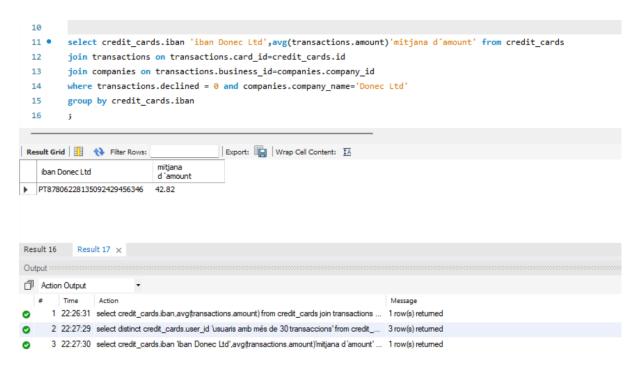
En la subconsulta:

SELECT card\_id FROM transactions where declined=0 group by card\_id having count(\*) >30

Filtro los id de las tarjetas de credito que tienen mas de 30 transacciones no declinadas.

En la consulta principal relaciono los id de los usuarios con los id de las tarjetas filtradas.

- Mostra la mitjana d'amount per IBAN de les targetes de crèdit a la companyia Donec Ltd, utilitza almenys 2 taules.
  - \* selecciono la media por iban de la compañia Donec Ltd



#### > código:

select credit\_cards.iban 'iban Donec Ltd',avg(transactions.amount)'mitjana d'amount' from credit\_cards

join transactions on transactions.card\_id=credit\_cards.id join companies on transactions.business\_id=companies.company\_id where transactions.declined = 0 and companies.company\_name='Donec Ltd' group by credit\_cards.iban;

## > explicación:

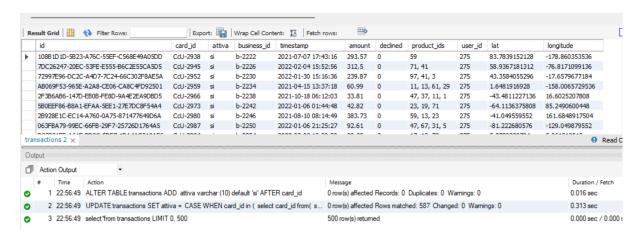
Selecciono los iban desde la tabla 'credit\_cards', la media de los importes de las transacciones no declinadas desde la tabla 'transactions'.

Relaciono los resultados con la tabla 'companies' para filtrar los valores de la compañia 'Donec Ltd'.



- → Crea una nova taula que reflecteixi l'estat de les targetes de crèdit basat en si les últimes tres transaccions van ser declinades i genera la següent consulta: Quantes targetes estan actives?
  - creo la tabla que refleccione si las tarjetas estan activas

```
1 • ALTER TABLE transactions ADD attiva varchar (10) default 'si' AFTER card_id;
2 ● ⊖ UPDATE transactions SET attiva = CASE
   select card_id from(
   6
      select card_id,timestamp,declined,ROW_NUMBER() OVER (PARTITION BY card_id ORDER BY timestamp desc) AS rn FROM transactions
      where rn<4
9
     ) y
10
      group by card_id having sum(declined)>2
      ) THEN REPLACE(attiva, 'si', 'no')
11
12
      else 'si'
13
      END;
14 • select*from transactions;
```



> codigo:

```
ALTER TABLE transactions ADD attiva varchar (10) default 'si' AFTER card_id;
UPDATE transactions SET attiva = CASE
WHEN card_id in (
select card_id from(
select * from(
select card_id,timestamp,declined,ROW_NUMBER() OVER (PARTITION BY card_id
ORDER BY timestamp desc) AS rn FROM transactions
) x
where rn<4
) y
group by card_id having sum(declined)>2
) THEN REPLACE(attiva, 'si', 'no')
else 'si'
END:
```

#### > explicación:

ALTER TABLE transactions ADD attiva varchar (10) default 'si' AFTER card id;

-Creo la nueva columna donde aparecerà el estado de la tarjeta.

UPDATE transactions SET attiva = CASE WHEN card id in (...

-Modifico la columna creada segun la condiciones requeridas.

select card\_id,timestamp,declined,ROW\_NUMBER() OVER (PARTITION BY card\_id ORDER -BY timestamp desc) AS rn FROM transactions

Relleno una nueva columna temporal con 'Row Number' que asigna valores 1,2,3 a las ultimas tres transacciones agrupadas por 'card\_id'

```
where rn<4) y
group by card_id having sum(declined)>2) THEN REPLACE(attiva, 'si', 'no')
else 'si'
```

-En caso que la suma de los declined de las ultimas tres transacciones (rn<4) sea mas de dos remplazo el valor 'si' con 'no'

## cuento las tarjetas activas



## > codigo:

select count(\*) 'targetes actives' from (select distinct card\_id from transactions where attiva='si')x;

### > explicación:

Cuento el número de tarjetas distintas que tengan valor 'si' a la columna que define si están activas.



 Crea una taula amb la qual puguem unir les dades del nou arxiu products.csv amb la base de dades creada, tenint en compte que des de transaction tens product\_ids.
 Genera la següent consulta:

Necessitem conèixer el nombre de vegades que s'ha venut cada producte.

\* muestro el resultado

```
1 • \ominus WITH RECURSIVE cte_count (n) AS (
            SELECT 1 UNION ALL SELECT n + 1
            FROM cte_count
  3
            WHERE n < 1000)
  5

⇒ select idprod 'producte', count(*) 'nombre de vegades que s'ha venut' from(
  6
            SELECT TRIM( BOTH FROM SUBSTRING_INDEX( SUBSTRING_INDEX(product_ids, ',', n) )AS idprod
             FROM transactions
  q
            JOIN cte_count cnt WHERE cnt.n <= LENGTH(product_ids) -LENGTH(REPLACE(product_ids,',','')) +1) x
 10
        group by idprod
 11
        order by idprod;
Result Grid Filter Rows:
                                     Export: Wrap Cell Content: IA
   producte nombre de vegades que s'ha venut
  11
          48
   13
            60
  17 61
Result 17 ×
Output:
Action Output
    1 10:58:08 WITH RECURSIVE cte_count (n) AS ( SELECT 1 UNION ALL SELECT n + 1 FR... 26 row(s) returned
         codigo:
```

## > explicación:

- -WITH RECURSIVE cte\_count (n) AS ( SELECT 1,3 UNION ALL SELECT n + 1,n\*n FROM cte\_count WHERE n < 1000)
- -En esta parte del codigo creo una tabla (Common Table Expression) con numeros sequenciales que necesitaré para relacionarla con la consulta siguiente. Suponiendo que ninguna singola compra pase los 1000 productos.
- -SELECT idprod 'producte', count(\*) 'nombre de vegades que s'ha venut':
- -Selecciono el ID del producto y cuenta cuántas veces se ha vendido de los valores que voy a extraer que al final agruparé y ordenaré por el ID.
- -TRIM(BOTH FROM..)
- -Elimina los espacios en blanco alrededor del valor extraído.
- -SUBSTRING\_INDEX(product\_ids, ',', n)
- -Esta función toma la cadena product\_ids y devuelve la subcadena hasta el n-ésimo separador.

Por una compra de 3 productos "1,25,74" y n=2 SUBSTRING\_INDEX(product\_ids, ',', 2) devolverá '1,25'.

- -SUBSTRING INDEX( SUBSTRING\_INDEX(product\_ids, ',', n) ,',',-1)
- -Esta función toma la subcadena obtenida en el 'Substring\_index' anterior y devuelve el último elemento después del último separador.

Continuando con el ejemplo anterior, SUBSTRING INDEX('1,25', ',', -1) devolverá '25'.

- -JOIN cte\_count cnt
- -Une la tabla transactions con la CTE cte\_count para contar todos los ID en la columna 'product ids'
- -LENGTH(product ids)
- -Es la longitud de caracteres de toda la cadena
- -REPLACE(product ids,',',")
- -Elimina las comas
- -LENGTH(REPLACE(product ids,',',"))
- -Es la longitud de la cadena sin las comas
- -WHERE cnt.n <= LENGTH(product\_ids) LENGTH(REPLACE(product\_ids, ',', ")) + 1
- la resta entre la longitud de todos los caracteres menos la longitud de caracteres sin sin las comas nos da como resultado el numero de comas -1.

Con el +1 al final obtengo el numero de comas (que es igual al numero de productos-1) y me aseguro que 'n' no exceda el numero de productos en 'poduct\_ids'.