SPRINT 4

Raquel Limpo Martínez

ÚLTIMO REPASO DE SQL EN EL QUE TRABAJAREMOS LA CREACIÓN DE BASES DE DATOS, LAS RELACIONES ENTRE TABLAS, TÉCNICAS DE MODELADO Y GRANULARIDAD DE LOS DATOS.

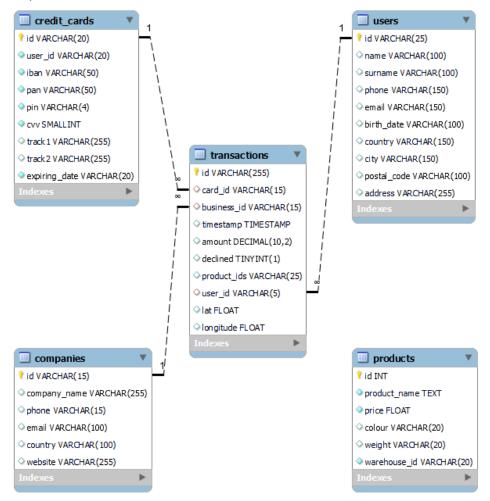
TABLA DE CONTENIDO

Nivel 1	. 2
Esquema en estrella	. 2
Creación de la base de datos e importación de los datos de los CSV	. 5
Ejercicio 1: Realiza una subconsulta que muestre a todos los usuarios con más de 30 transacciones utilizando al menos 2 tablas	11
Ejercicio 2 Muestra la media de amount por IBAN de las tarjetas de crédito a la compañía Donec Ltd., utiliza por lo menos 2 tablas.	
Nivel 21	12
Ejercicio 1: Crea una nueva tabla que refleje el estado de las tarjetas de crédito basado e si las últimas tres transacciones fueron declinadas y genera la siguiente consulta:	
¿Cuantas tarjetas están activas?1	
Nivel 3	13
Ejercicio 1: Necesitamos conocer el número de veces que se ha vendido cada producto. 1	14
Anexo: Los problemas de carga las tablas	15

Nivel 1

Descarga los archivos CSV, estúdialos y diseña una base de datos con un esquema en estrella que contengas al menos 4 tablas de las cuales puedas realizar las consultas de los ejercicios 1 y 2:

Esquema en estrella



En este esquema de estrella, la tabla factual o de hechos sería **Transaction**, no solo porque está conectada con todas las demás, sino que contiene los indicadores del negocio, cuanto se ha vendido, cuando a quien de parte de quien, cuanto dinero se ha ganado...

Esta unida a las tablas de dimensiones: **Companies, Credit_cards y Users.** No me deja conectarla a products, porque la foreign key es muchos a muchos.

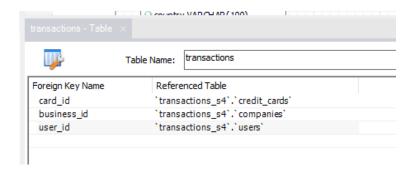


Tabla Transaction:

- id : El identificador de la transacción realizada. Es la clave primaria de esta tabla Como he comentado previamente, sería ideal que este campo fuera distintivo como *transaction_id*. Lo mismo aplica para el resto de tablas.
- card_id : El identificador de la tarjeta de crédito con la que se ha realizado la transacción, nos une a la tabla *Credit_card*
- business_id : El identificador de la compañía desde la que ha realizado la transacción, nos une a la tabla *Company*.
- timestamp : La hora a la que se ha realizado la transacción
- amount : La cantidad de dinero que se ha usado en la transacción
- declined : Campo booleano, nos indica si la transacción se ha realizado (0) o se ha cancelado (1)
- product_ids : los identificadores de los diferentes productos que se han intercambiado por dinero en la transacción. Cuando hay más de uno está separado por comas. Nos uniriamos a la tabla *Products* mediante una tabla intermedia que pondremos en el nivel 3..
- user_id : El identificador del usuario que ha realizado la transacción. Nos une por clave foránea a las tablas *Users*
- lat y longitude: Las coordenadas desde donde se ha hecho la transacción.

Tabla Company:

- id: El identificador para cada compañía. ES la clave primaria de esta tabla.
- company_name. El nombre de la compañía.
- phone : El número de contacto de cada compañía.
- email: el email de contacto de cada compañia
- country: el país donde esta la sede central de cada compañía. Este sería el único campo que no tendría por qué ser único de esta tabla.
- website: La página oficial de la compañia.

Tabla Credit_cards: En esta tabla hay varios datos NOT NULL ya que son importantes a la hora de hacer compra online

- id: Es el identificador de la tarjeta de crédito. Este campo es la clave primaria,
- iban: Es el código internacional de cuenta bancaria. Es un código alfanúmerico que identifica una cuenta bancaria en cualquier lugar del mundo.
- pan: es el acrónimo de personal account number y es el número que aparece en la tarjeta de crédito.
- pin: personal identification number es tu clave para entrar en tu cuenta.
- cvv: código valor de verificación. Es un número que está detrás de la tarjeta y que se suele pedir por seguridad en transacciones online.
- track1 y track2 : Son las pistas de banda magnética. Los lectores de losTPV siempre leen las pistas 1 o 2: la información mínima necesaria para realizar una transacción se encuentra en ambas pistas. La pista 1 tiene una mayor densidad de bits, es la única que puede contener caracteres alfanuméricos, y por tanto es la única que puede contener el nombre del portador de la tarjeta.
- Expiring_date La fecha de caducidad de la tarjeta. Al igual que el cvv se pide a menudo en transacciones online.

Tabla Product:

- id : El identificador del producto, es la primary key de la tabla.
- product_name : El nombre del producto.
- price El precio del producto, es un float y lleva símbolo de dólar, con lo cual puede dar problema al cargar la tabla.
- Colour: color en código hexadecimal del producto
- weight Peso del producto, sirve a la hora de eviar el producto y para tenerlo en cuenta al recibirlo.
- warehouse_id : Identificador del almacén, sirve para saber dónde está cada producto (qué almacén, qué zona del almacén, qué estante...).

Tablas Users de Canada, Reino Unido y Estados Unidos: todas tienen los mismos campos, así que las he unido todas en una sola tabla, para evitar repetición de código y simplificar las querys.

- id: Identificador del usuario, clave primaria.
- Name: nombre del usuario
- Surname: apellido del usuario
- Phone: número de contacto del usuario
- Email: Correo electronico de contactodel usuario.
- birth_date : Fecha de nacimiento del usuario
- country: país de residencia del usuario
- city: Ciudad de residencia del usuario
- postal code: Codigo postal del lugar de residencia del usuario
- address : dirección del lugar de residencia del usuario

Creación de la base de datos e importación de los datos de los CSV

Primero creamos la base de datos, para evitar problemas con bases de datos previas, le ponemos de nombre **transactions_s4**

Database y tabla companies

```
-- Creamos la base de datos

CREATE DATABASE IF NOT EXISTS transactions_s4;

USE transactions_s4;

-- Creamos las tablas

#companies

CREATE TABLE IF NOT EXISTS companies (
    id VARCHAR(15) PRIMARY KEY,
    company_name VARCHAR(255),
    phone VARCHAR(15),
    email VARCHAR(100),
    country VARCHAR(100),
    website VARCHAR(255)
```

Creamos la tabla **companies** y la cargamos. Para cargar los datos he buscado la ruta a través del código sql "SHOW VARIABLES LIKE 'secure_file_priv';", este te da la ruta exacta y te permite entrar en carpetas, aunque estén ocultas, pudiendo copiar ahí los archivos. Como estos dtos no tenian todos los campos de varchar entre comillas, he tenido que añadir "OPTIONALLY ENCLOSED BY ""

```
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/companies.csv'
INTO TABLE companies
FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 1 ROWS;
```

Luego comprobamos que se han cargado los datos.

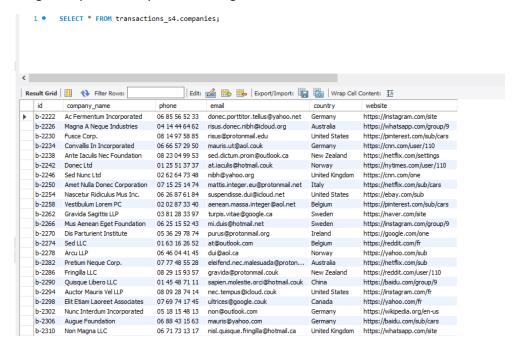


Tabla transactions

Repetimos el proceso con la tabla **transactions.** El reto con esta tabla fueron las foreign keys, ya que no deja hacer una relación de muchos a muchos con products_id.

En la parte de carga, además del problema con las comillas, se nos añadia que los campos no estaban separados por ',' sino por ';'.

```
# transactions

    ○ CREATE TABLE IF NOT EXISTS transactions (
      id VARCHAR(255) PRIMARY KEY,
      card_id VARCHAR(15) REFERENCES credit_card(id),
      business_id VARCHAR(15)REFERENCES companies(id),
      timestamp TIMESTAMP,
      amount DECIMAL(10,2),
      declined BOOLEAN,
      product_ids VARCHAR(25),
      user id VARCHAR(5) REFERENCES users(id),
      lat FLOAT,
      longitude FLOAT
      );
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/transactions.csv'
INTO TABLE transactions
FIELDS TERMINATED BY ';'
OPTIONALLY ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 1 ROWS;
 1 • SELECT * FROM transactions_s4.transactions;
```

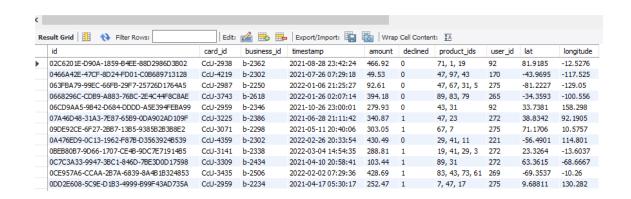
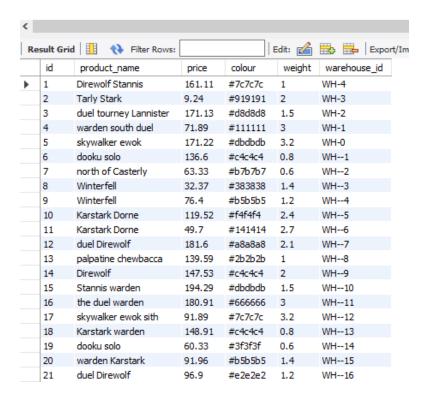


Tabla products

El precio estaba en dólares, así que no se cargaba correctamente, así que reemplacé el símbolo de dólar por un espacio.

```
#products
CREATE TABLE IF NOT EXISTS products (
    id VARCHAR(20) PRIMARY KEY,
    product_name TEXT NOT NULL ,
    price FLOAT NOT NULL,
    colour VARCHAR(20),
    weight VARCHAR(20),
    warehouse_id VARCHAR(20) NOT NULL
);
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/products.csv'
INTO TABLE products
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
IGNORE 1 ROWS
(id, product_name, @price_raw, colour, weight, warehouse_id)
SET price = REPLACE(@price_raw, '$', '');
        SELECT * FROM transactions_s4.products;
```



```
Tabla credit cards
  #credit cards

    CREATE TABLE credit_cards (
      id VARCHAR(20) PRIMARY KEY,
      user_id VARCHAR(20) NOT NULL,
      iban VARCHAR(50) NOT NULL,
      pan VARCHAR(50) NOT NULL,
      pin VARCHAR(4) NOT NULL,
      CVV SMALLINT NOT NULL,
      track1 VARCHAR(255),
      track2 VARCHAR(255),
      expiring_date VARCHAR(20) NOT NULL
 );
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/credit cards.csv'
INTO TABLE credit_cards
FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 1 ROWS;
🗎 📙 | 🐓 🙀 🧔 🕛 | 🗞 | 📀 🔕 🔞 | Don't Limit
                                         - | 🏡 | 🥩 🔍 👖 📦
      SELECT * FROM transactions s4.credit cards;
```

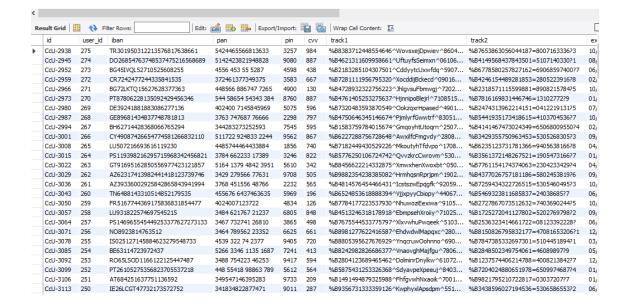


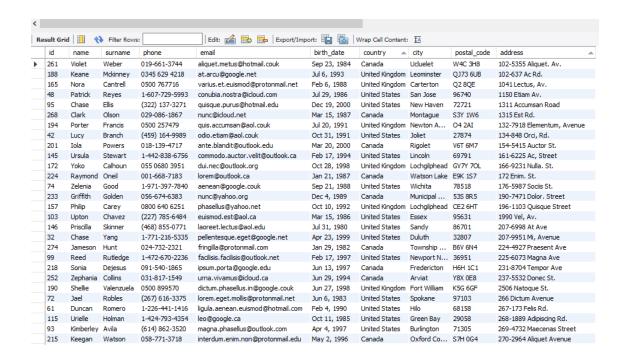
Tabla users:

He unido las tres tablas de users en una para evitar código repetido. En este caso el problema de carga era elsalto de línea. Según el manual depende de qué documentos y cómo hubieran sido creados hacia falta poner el /r

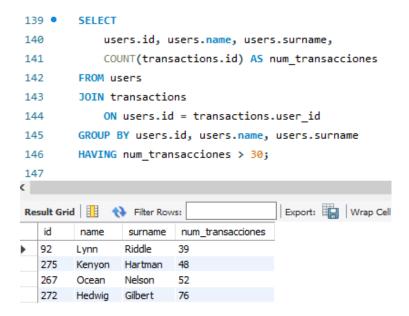
 $\frac{\text{http://download.nust.na/pub6/mysql/doc/refman/5.0/es/load-data.html#:}^{\text{ctext=Los}\%20 valores\%20 FIELDS\%20 TERMINATED\%20 BY,BY\%20'\%5 Cr\%5 Cn'\%20}{\text{ctext=Los}\%20 valores\%20 FIELDS\%20 TERMINATED\%20 BY,BY\%20'\%5 Cr\%5 Cn'\%20'\%5 Cr\%5 Cn'\%20'\%5 Cr\%5 Cn'\%20'\%5 Cr\%5 Cn'\%20'\%5 Cr\%5 Cn'\%20'\%5 Cr\%5 Cn'\%20'\%5 Cn'\%20'\%20'\%5 Cn'\%20'\%20'\%20'\%20'\%20'\%20'\$

"Nota: Si ha generado el fichero de texto en un sistema Windows, puede tener que usar lines terminated by '\r\n' para leer correctamente el fichero, ya que los programas de Windows típicamente usan dos caracteres como terminadores de línea. Algunos programas como wordPad, pueden usar \r como terminador de línea al escribir ficheros. Para leer tales ficheros, use lines terminated by '\r'."

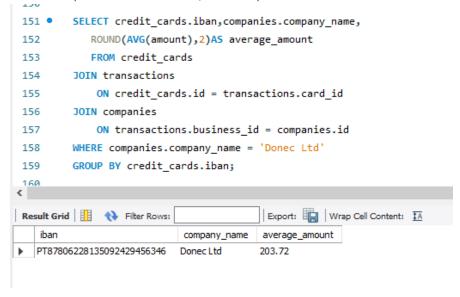
```
■ ○ CREATE TABLE IF NOT EXISTS users (
        id VARCHAR(20) PRIMARY KEY,
        name VARCHAR(100),
        surname VARCHAR(100),
        phone VARCHAR(150),
        email VARCHAR(150),
        birth_date VARCHAR(100),
        country VARCHAR(150),
        city VARCHAR(150),
        postal code VARCHAR(100),
         address VARCHAR(255)
        );
 LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/users_ca.csv'
 INTO TABLE users
FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY ""
LINES TERMINATED BY '\r\n'
 IGNORE 1 ROWS;
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/users_uk.csv'
INTO TABLE users
FIELDS TERMINATED BY ','
 OPTIONALLY ENCLOSED BY '"'
LINES TERMINATED BY '\r\n'
IGNORE 1 ROWS;
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/users_usa.csv'
 INTO TABLE users
FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY ""
LINES TERMINATED BY '\r\n'
IGNORE 1 ROWS;
```



Ejercicio 1: Realiza una subconsulta que muestre a todos los usuarios con más de 30 transacciones utilizando al menos 2 tablas.



Ejercicio 2 Muestra la media de amount por IBAN de las tarjetas de crédito a la compañía Donec Ltd., utiliza por lo menos 2 tablas.



Nivel 2

Ejercicio 1: Crea una nueva tabla que refleje el estado de las tarjetas de crédito basado en si las últimas tres transacciones fueron declinadas y genera la siguiente consulta: ¿Cuantas tarjetas están activas?

Crea una nueva tabla para el estado de las tarjetas.

```
■ ○ CREATE TABLE credit_card_status (
        card_id VARCHAR(15) PRIMARY KEY,
        status VARCHAR(15) NOT NULL
    );
   INSERT INTO credit_card_status (card_id, status)
    SELECT
        transactions.card_id,
       CASE
           WHEN SUM(CASE WHEN transactions.declined THEN 1 ELSE 0 END) = 3 THEN 'Inactiva'
            ELSE 'Activa'
        END AS status
 ⇒ FROM (
        SELECT card_id, declined
        FROM transactions
        ORDER BY timestamp DESC
        LIMIT 3
    ) transactions
    GROUP BY transactions.card_id;
   1 • SELECT * FROM transactions_s4.credit_card_status;
```



Nivel 3

Crea una tabla con la que podamos unir los datos del nuevo archivo products.csv con la base de datos creada, teniendo en cuenta que desde transaction tienes product_ids. Genera la siguiente consulta:

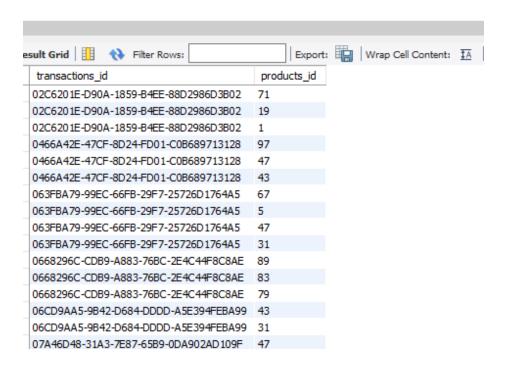
```
# *****Nivell 3*****
#Exercici 1: Necessitem conèixer el nombre de vegades que s'ha venut cada producte.

CREATE TABLE bridge_products (
    transactions_id VARCHAR(100),
    products_id VARCHAR(100),
    FOREIGN KEY (transactions_id) REFERENCES transactions(id),
    FOREIGN KEY (products_id) REFERENCES products(id)

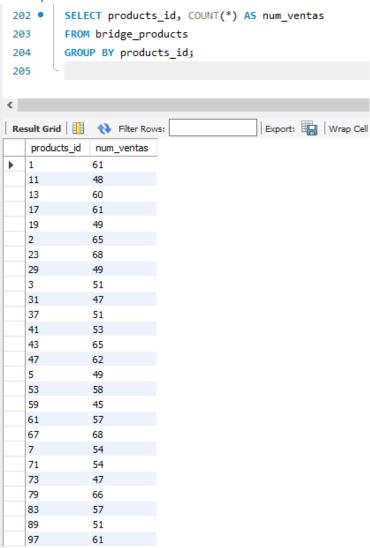
);

INSERT INTO bridge_products (transactions_id, products_id)
SELECT transactions.id AS transactions_id, products.id AS products_id
FROM transactions
JOIN products
ON FIND_IN_SET(products.id, REPLACE(transactions.product_ids, ' ', '')) > 0;
```

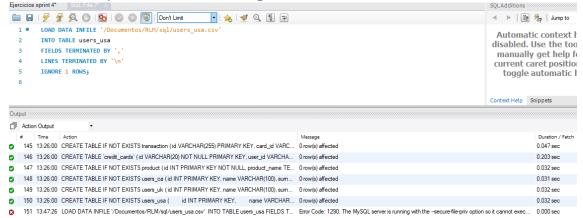
1 • SELECT * FROM transactions_s4.bridge_products;



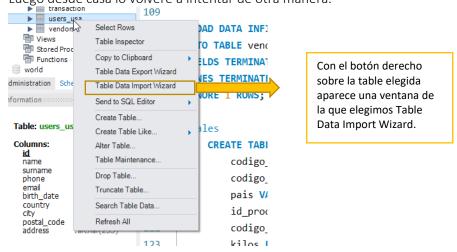
Ejercicio 1: Necesitamos conocer el número de veces que se ha vendido cada producto.



Anexo: Los problemas de carga las tablas



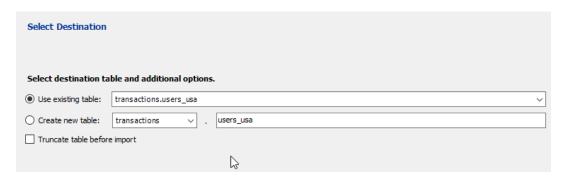
Originalmente iba a cargarlo así, però me aparecía el error "Error Code: 1290 - The MySQL server is running with the --secure-file-priv option so it cannot execute this statement." Y las maneras de corregirlo son difíciles en el ordenador del cibernàrium, ya que tengo que entrar en zonas seguras y no tengo los permisos, así que voy a ir avanzando, usando el import wizard. Luego desde casa lo volveré a intentar de otra manera.



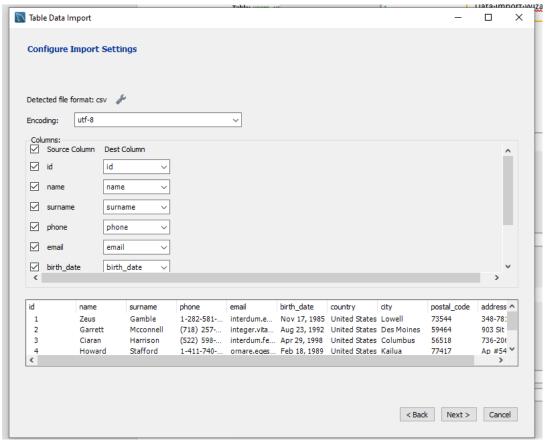
Nos abre el Wizard y le decimos de donde sacamos el archivo.

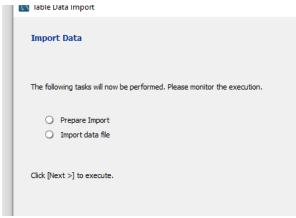


Le damos a Next y elegimos destino



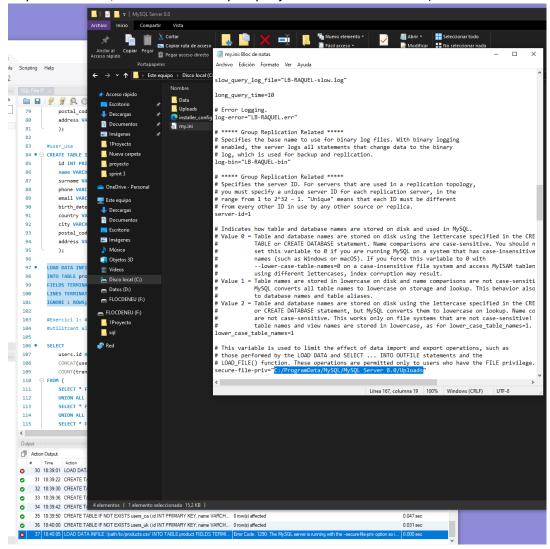
Le volvemos a dar a Next



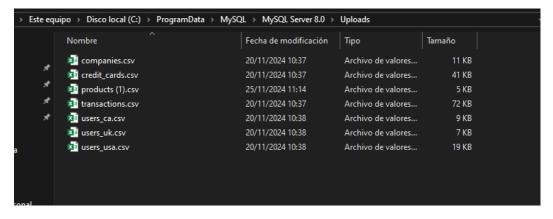


Como corregirlo desde casa:

Hay dos maneras, una abriendo el my.ini y dejando las comillas vacías (dónde esta resaltado).



La segunda, llevar los archivos csv a este directorio C:/ProgramData/MySQL/MySQL Server 8.0/Uploads , que es la que voy a hacer ahora.



Para cualquiera de ellas hay que volver a iniciar el mysql.

Bueno, a pesar de que ahora funciona, cada tabla da un error y hay que corregir cada una. Products, da problemas a causa del símbolo del dólar y hay que añadir SET price = REPLACE(@price_raw, '\$', ");

Companies , transactions y credit_cards, estan separadas por ; en vez de , y tienen problemas con la falta de comillas en algunos registros, mientras que users,