1		Des	crip	ció	2
2		Nive	ell 1.		2
	2.	1	Diss	seny de la base de dades amb almenys 4 taules	2
		2.1.1	1	Creació de la nova base de dades	2
		2.1.2	2	Creació de l'estructura de la taula companies	4
		2.1.3	3	Creació de l'estructura de la taula users	5
		2.1.4	4	Creació de la taula credit_card	6
		2.1.5	5	Creació de l'estructura de la taula transactions	8
		2.1.6 user		Automatitzar en Excel per generar codi SQL per inserir les dades a la taula 9	
		2.1.7 com		Automatitzar en Excel per generar codi SQL per inserir les dades a la taula ies	13
		2.′	1.7.1	Variant de la solució anterior utilitzant instrucció LOAD DATA (millor opció)	15
		2.1.8 cred		Automatitzar en Excel per generar codi SQL per inserir les dades a la taula ard	17
		2.1.9 tran		Automatitzar en Excel per generar codi SQL per inserir les dades a la taula ions	19
		2.1.1	10	Diagrama de la base de dades a MySQL Workbench	23
	2.	2	Exe	rcici 1	25
		2.2.1	I	Codi sql de la solució	25
	2.	3	Exe	rcici 2	26
		2.3.1	ı	Codi sql de la solució	26
3		Nive	ell 2.		27
	3.	1	Crea	ació taula Estat targetes de crèdit	27
	3.	2	Exe	rcici 1	29
		3.2.1	1	Codi sql de la solució	29
4		Nive	ell 3.		30
	4.	1	Exe	rcici 1	30
		4.1.1	ı	Creació estructura taula product_ids	30
		4.1.2 proc		Automatitzar en Excel per generar codi sql per inserir les dades a la taula _ids	31
		4.1.3	3	Construir taula intermèdia que relacioni transaccions i productes venuts	33
		4.1.4	4	Relacions taula intermèdia trans_x_product amb base de dades bbdd_tasca_	4
		4.1.5	5	Diagrama de la base de dades al afegir taula intermèdia trans_x_product	37
		110	•	Codi sal que déna respecta a la progunta de l'eversici	20

1 Descripció

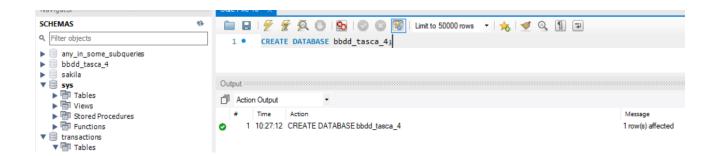
Partint d'alguns arxius CSV dissenyaràs i crearàs la teva base de dades.

2 Nivell 1

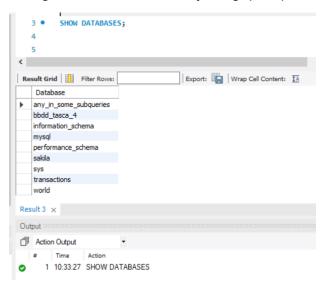
Descàrrega els arxius CSV, estudia'ls i dissenya una base de dades amb un esquema d'estrella que contingui, almenys 4 taules de les quals puguis realitzar les següents consultes:

2.1 Disseny de la base de dades amb almenys 4 taules

2.1.1 Creació de la nova base de dades



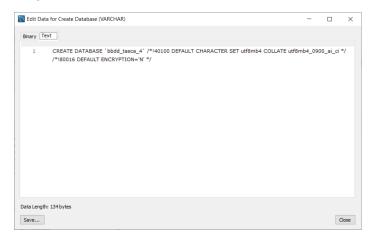
Si faig SHOW DATABASES; ja veig que apareix bbdd_tasca_4;

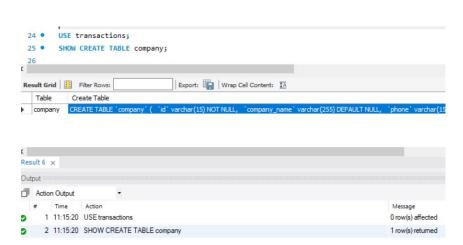






Veig automàticament que té les propietats que m'interessen



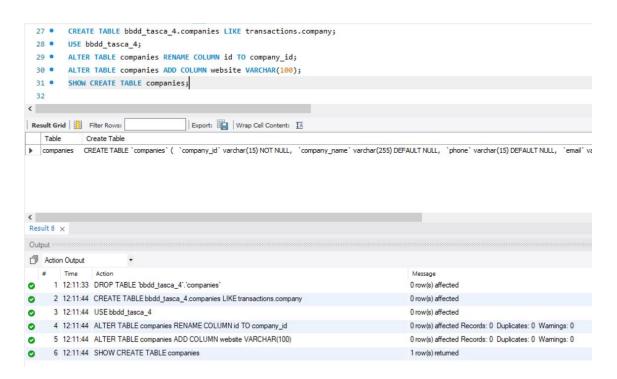




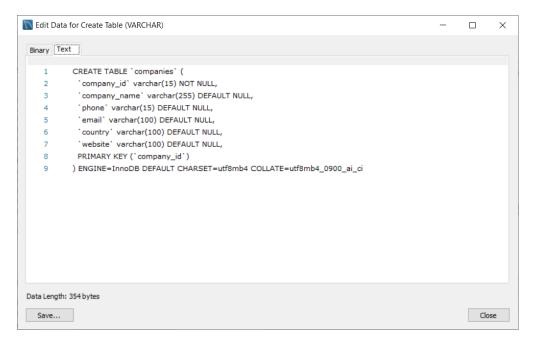
Així que l'estructura em serveix.

2.1.2 Creació de l'estructura de la taula companies

Com que les taules estan a diferents bases de dades el CREATE TABLE LIKE s'ha de fer com segueix, i faig ALTER per modificar el nom de la primera columna, i afegir una variable més de website.

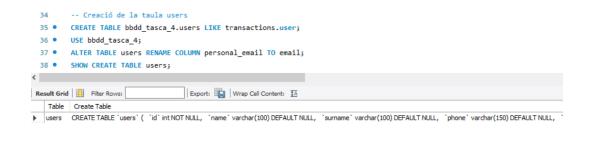


Així doncs aquesta estructura em serveix per omplir les dades;



2.1.3 Creació de l'estructura de la taula users

Faig el mateix per crear la taula users:





I l'estructura permetrà complimentar els valors:

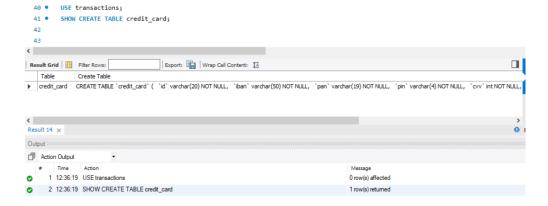


Però com que tinc varis *csv* de *users* de diferents països, crearé les diferents taules de *users* per país i després les fusionaré.

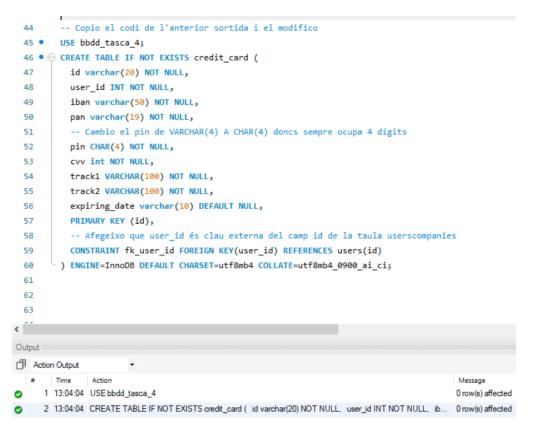


2.1.4 Creació de la taula credit_card

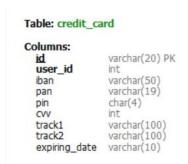
Com que veig que les columnes de credit_card.csv no tenen el mateix ordre que els Sprints anteriors, el que faré és un SHOW CREATE TABLE, copiaré el codi, i l'inseriré al Workbench i faré les modificacions pertinents.



```
Edit Data for Create Table (VARCHAR)
                                                                                                            X
 Binary Text
            CREATE TABLE 'credit_card' (
             'id' varchar(20) NOT NULL,
             `iban` varchar(50) NOT NULL,
            'pan' varchar(19) NOT NULL,
     5
            `pin` varchar(4) NOT NULL,
     6
             'cvv' int NOT NULL,
             `expiring_date` varchar(10) DEFAULT NULL,
            `fecha_actual` date DEFAULT NULL,
     8
            PRIMARY KEY ('id')
   10
         ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
 Data Length: 337 bytes
 Save...
                                                                                                             Close
```

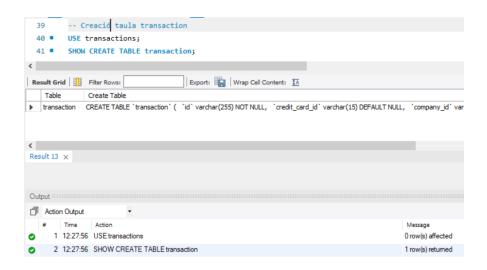


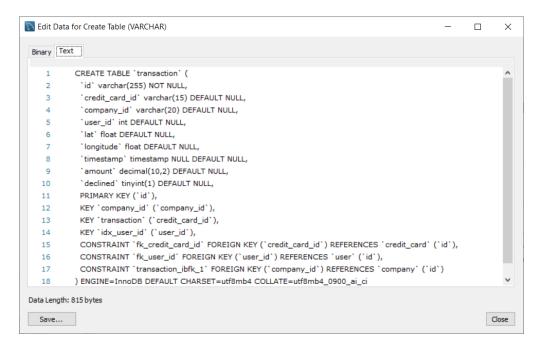
Refrescant l'esquema també veig com es la nova taula i també veig que user id és clau externa.



2.1.5 Creació de l'estructura de la taula transactions

Com que l'ordre de les columnes és diferent en aquest cas, es més fàcil copiar l'estructura del SHOW CREATE TABLE al Workbench i després canviar l'ordre i els noms tal com l'arxiu csv.





Com que l'ordre i els noms de les columnes em canvien respecte al arxiu *transactions.csv* com abans m'és més còmode copiar el codi al Workbench i modificar-lo.

Així doncs:

```
67 • USE bbdd_tasca_4;
 68 • ⊖ CREATE TABLE IF NOT EXISTS transactions (
         id varchar(255) NOT NULL,
          card_id varchar(15) DEFAULT NULL,
 71
         business_id varchar(20) DEFAULT NULL,
 72
         timestamp timestamp NULL DEFAULT NULL,
          amount decimal(10,2) DEFAULT NULL,
 73
          declined tinyint(1) DEFAULT NULL,
 74
 75
          -- Afegeixo aquest campo
           product_ids VARCHAR(20) DEFAULT NULL,
 76
          user_id int DEFAULT NULL,
          lat float DEFAULT NULL,
 79
          longitude float DEFAULT NULL,
         PRIMARY KEY (id),
 80
           -- Deixo els índexs que m'aniran bé per fer les consultes més ràpides
 81
 82
          KEY company_id (business_id),
          KEY transactions (card_id),
          KEY idx_user_id (user_id),
         CONSTRAINT fk_card_id FOREIGN KEY (card_id) REFERENCES credit_card (id),
          CONSTRAINT fk2 user id FOREIGN KEY (user id) REFERENCES users (id),
           CONSTRAINT fk business id FOREIGN KEY (business id) REFERENCES companies (company id)
 87
      ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4 0900 ai ci;
 88
                                                                          Message
    3 13:41:16 USE bbdd tasca 4
                                                                          0 row(s) affected
🛕 4 13:41:16 CREATE TABLE IF NOT EXISTS transactions ( id varchar(255) NOT NULL, card_id varchar(15) DE... 0 row(s) affected, 1 warning(s): 1681 Integer display width is deprecated and will be removed in a future...
```

He mirat el Warning i és un tema de versió, però no sembla que hagi de donar problemes.

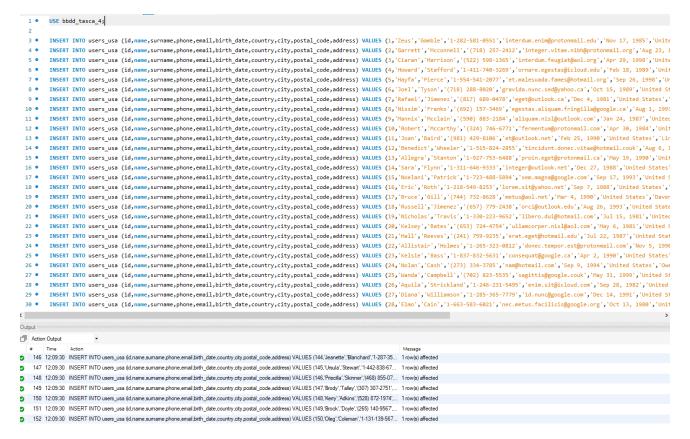
2.1.6 Automatitzar en Excel per generar codi SQL per inserir les dades a la taula users

De la taula user_usa.csv, fem separació de columnes i obtenim un arxiu user_usa.xlsx. A la última columna de la fila 2 inserim el següent codi (cel.la K2):

Arrastrem per totes les files de l'arxiu, i finalment peguem el codi al Workbench i ho executem en un nou arxiu denominat: .

Comprovo el codi sql, i veig que per culpa del cognom de *O'donnell* ens fallaria a partir del *id*=45. Així el que faig es escriure 'O\'donnell', per tal que l'apòstrof del cognom no l'agafi com un operador.

```
postal_code,address) VALUES (36, 'Raymond', 'Thornton', '1-581-648-7426', 'enim.nisl@google.edu', 'Dec 6, 1982', 'United States', 'Kansas City', '26231',
38 •
39 •
       postal_code,address) VALUES (37,'Graiden','Glover','(888) 957-3868','non.hendrerit@hotmail.ca','Nov 11, 1987','United States','Portland','35397',
       postal_code,address) VALUES (38,'Abra','Doyle','(367) 861-9621','augue.eu@outlook.edu','Aug 22, 1986','United States','Chicago','54137','409-9169
       postal_code,address) VALUES (39,'Nyssa','Shaffer','1-899-494-4941','malesuada.fames@google.ca','Jul 2, 1991','United States','Columbus','27828','
41 •
       postal_code,address) VALUES (40,'Astra','Alexander','1-877-546-5066','lectus.nullam.suscipit@google.org','Mar 12, 1983','United States','Columbus
42 •
43
       postal_code,address) VALUES (41, 'Uriel', 'Hebert', '1-265-846-2455', 'donec@outlook.com', 'Nov 30, 1981', 'United States', 'Juneau', '33549', '670-5867 E
       postal_code,address) VALUES (42,'Lucy','Branch','(459) 164-9989','odio.etiam@aol.couk','Oct 31, 1991','United States','Joliet','27874','134-848 0
44 •
45 •
       postal_code,address) VALUES (43,'Jayme','Chavez','(542) 568-7326','fusce.feugiat@outlook.com','Jul 29, 1984','United States','Aurora','28283','86
       postal_code,address) VALUES (44,'Aquila','Haley','1-951-243-1892','tempus.risus@outlook.edu','Jan 21, 1996','United States','Newark','51734','Ap
46
       postal_code,address) VALUES (45, 'Richard', 'O'dennell', '1-275-844-6782', 'ac.nulla@outlook.org', 'Oct 4, 1986', 'United States', 'Frankfort', '72520', '
47 🖸 🖂
       postal_code,address) VALUES (46,'Alika','Valdez','(781) 178-0838','dapibus.gravida@yahoo.net','Aug 8, 1990','United States','Salem','55729','P.O.
       postal_code,address) VALUES (47, 'Herrod', 'Wright', '(575) 645-3218', 'odio.auctor.vitae@yahoo.edu', 'Nov 11, 1999', 'United States', 'Fort Wayne', '728
49
       postal_code,address) VALUES (48, 'Patrick', 'Reyes', '1-607-729-5993', 'conubia.nostra@icloud.com', 'Jul 29, 1986', 'United States', 'San Jose', '96740',
50
       postal_code,address) VALUES (49, 'Lev', 'Roth', '1-685-331-7392', 'auctor.vitae.aliquet@yahoo.net', 'Dec 3, 1981', 'United States', 'Richmond', '52748',
       postal_code,address) VALUES (50,'Aretha','Chang','1-634-250-1977','suspendisse.aliquet.molestie@hotmail.edu','May 11, 1998','United States','Jone
52
       postal_code,address) VALUES (51, Lionel', 'Griffith', '1-816-243-8788', 'ultricies.sem.magna@yahoo.org', 'Apr 3, 1993', 'United States', 'Wichita', '176
```

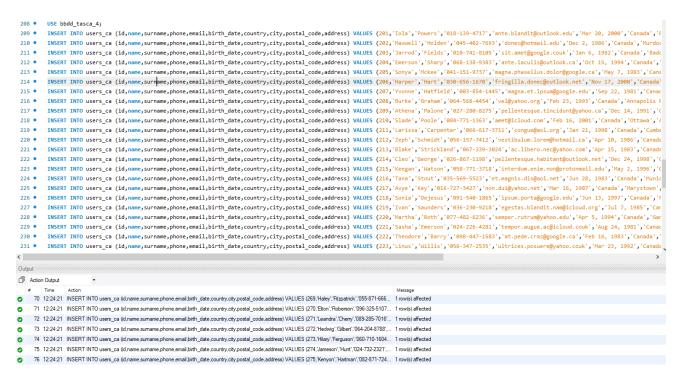


Fem el mateix per la taula *users_uk*. També ens apareix un error al id=181 al cognom que corregim com per tal que puguem executar sense errors. Després tornarem a posar els noms correctes com teníem a les dades del *csv* amb un UPDATE.

```
155
               -- Insertem les dades a la taula users uk
156 •
             USE bbdd tasca 4:
157 •
              INSERT INTO users uk (id.name.surname.phone.email.birth date.country.city.postal code.address) VALUES (151, 'Meghan', 'Hayden
              INSERT INTO users uk (id,name,surname,phone,email,birth date,country,city,postal code,address) VALUES (152, 'Hakeem', 'Alford
158 •
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (153,'Keegan','Pugh',
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (154,'Cooper','Bullocl
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (155,'Joshua','Russel
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (156,'Remedios','Case
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (157,'Philip','Carey'
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (158,'Fatima','Dyer',
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (159,'Kylynn','Aceved
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (160,'Lael','Moody','G
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (161,'Nora','Reeves',
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (162,'Francesca','Saw
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (163,'Denton','Blackbu
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (164,'Preston','Hood'
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (165,'Nora','Cantrell
171 •
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (166,'Matthew','Woodal
172 •
173 •
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (167,'Sheila','Dicker:
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (168, 'Donna', 'Rivers'
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (169,'Linda','Gaines'
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (170,'William','Benjam','Benjam','Benjam', 'Benjam', '
             INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (171,'Gary','Robbins'
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (172,'Yoko','Calhoun'
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (173,'Bertha','Sloan'
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (174,'Amal','Kennedy'
              INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (175,'Brent','Bates',
Action Output
                        Action
                                                                                                                                                                                       Message
      45 12:19:20 INSERT INTO users_uk (id,name,sumame,phone,email,birth_date,country,city,postal_code,address) VALUES (194, 'Porter', 'Francis', '0500 257479','...
     46 12:19:20 INSERT INTO users_uk (id,name,sumame,phone,email,birth_date,country,city,postal_code,address) VALUES (195,'Rosalyn','Blake','(015905) 2845... 1 row(s) affected
       47 12:19:20 INSERT INTO users_uk (id.name.sumame.phone.email.birth_date.country.city.postal_code.address) VALUES (196,'Blaze','Duke','0800 060 8337',q... 1 row(s) affected
     48 12:19:20 INSERT INTO users_uk (id,name,surname,phone,email,birth_date,country,city,postal_code,address) VALUES (197,'Carty','Mathews','(024) 5216 76... 1 row(s) affected
       49 12:19:20 INSERT INTO users_uk (id.name.sumame.phone.email.birth_date.country.city.postal_code.address) VALUES (198, 'Keely', 'Fox', '0800 1111', 'id.mollis... 1 row(s) affected
   50 12:19:20 INSERT INTO users_uk (id_name,sumame,phone,email_birth_date,country,city_postal_code,address) VALUES (199,'Lewis','Melendez','(016977) 575... 1 row(s) affected
       51 12:19:20 INSERT INTO users_uk (id_name_sumame_phone_email_birth_date_country_city_postal_code_address) VALUES (200,"Nero", "Mills", "056 4507 5712", "adi... 1 row(s) affected
```

Veiem que s'executa correctament.

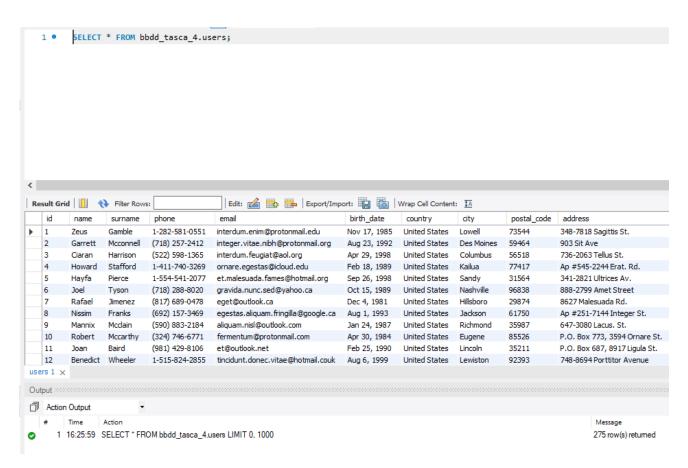
Fem el mateix per la taula users ca:



I ara fusionem les tres taules per inserir-les a la taula u*sers*, que sabem que tenen id diferents amb un UNION.



Si faig un SELECT de la taula users clarament té 275 files, i es verifica que és la suma de les files de les taules *users_usa*, *users_uk* i *users_ca*.



Podria eliminar les taules *users_usa*, *users_uk*, i *users_ca* però de moment les deixo a la base de dades ja que no molesten.

2.1.7 Automatitzar en Excel per generar codi SQL per inserir les dades a la taula companies

Fem una fórmula en Excel a la cel.la G2 de l'arxiu companies.xlsx tal com l'apartat anterior per generar el codi sql.

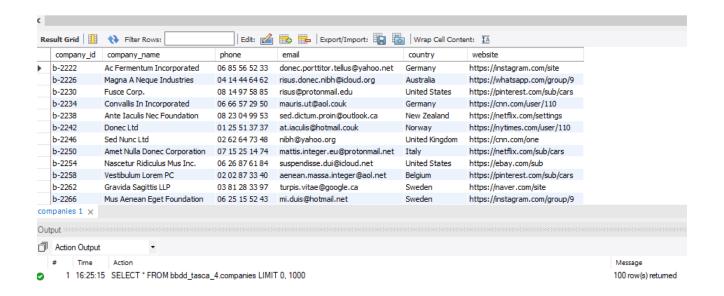
```
=CONCATENAR("INSERT INTO companies (";$A$1;",";$B$1;",";$C$1;",";$D$1;",";$E$1;",";$F$1; ") VALUES (";""";A2;"";",";""";B2;""";",";""";C2;"";",";"";D2;"";",";"";E2;""";",";"";F2;"";");")
```

Ho arrosseguem per totes les files i copiem el codi al Workbench. Executem i veiem que no hi ha errors.

```
USE bbdd tasca 4;
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2222','Ac Fermentum Incorporated','06 85 56 52 33','donec.portitor.tellus@yahoo.net','Germany','https:
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2226', 'Magna A Neque Industries', '04 14 44 64 62', 'risus.donec.nibh@icloud.org', 'Australia', 'https://wh
           INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2230','Fusce Corp.','08 14 97 58 85','risus@protonmail.edu','United States','https://pinterest.com/sub/c
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2234','Convallis In Incorporated','06 66 57 29 50', 'mauris.ut@aol.couk','Germany', 'https://cnn.com/user/
          INSERT INTO companies (company id, company name, phone, email, country, website) VALUES ('b-2238', 'Ante Iaculis Nec Foundation', '08 23 04 99 53', 'sed.dictum.proin@outlook.ca', 'New Zealand', 'https://doi.org/10.1006/
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2242', 'Donec Ltd','01 25 51 37 37', 'at.iaculis@hotmail.couk','Norway', 'https://nytimes.com/user/110');
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2246','Sed Nunc Ltd','02 62 64 73 48','nibh@yahoo.org','United Kingdom','https://cnn.com/one');
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2250','Amet Nulla Donec Corporation','67 15 25 14 74','mattis.integer.eu@protonmail.net','Italy','
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2254', 'Nascetur Ridiculus Mus Inc.','06 26 87 61 84', 'suspendisse.dui@icloud.net', 'United States', 'http
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2258','Vestibulum Lorem PC','02 02 87 33 40','aenean.massa.integer@aol.net','Belgium','https://pinterest
          INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2262','Gravida Sagittis LLP','03 81 28 33 97','turpis.vitae@google.ca','Sweden','https://naver.com/site
           INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES ('b-2266','Mus Aenean Eget Foundation','06 25 15 52 43','mi.duis@hotmail.net','Sweden','https://instagram.cc
Output :::
Action Output
     77 13:31:39 INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES (b-2522', Viverra Donec Foundation', '03 33 12 32 7...
                                                                                                                                            1 row(s) affected
Ø
    78 13:31:39 INSERT INTO companies (company_id.company_name_phone.email.country.website) VALUES (b-2526', 'Amet Luctus Vulputate Foundation', '03 18 ... 1 row(s) affected
      79 13:31:39 INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES (b-2530', Tincidunt Associates', 109 82 63 76 61', inc.
      80 13:31:39 INSERT INTO companies (company id, company name phone email country, website) VALUES (b-2534', Mattis Foundation', '06 74 03 95 74', 'sed@... 1 row(s) affected
      81 13:31:39 INSERT INTO companies (company_id,company_name.phone.email.country.website) VALUES (b-2538','At Associates','09 56 61 10 65',tristique.n...
      82 13:31:39 INSERT INTO companies (company_id_company_name_phone_email.country,website) VALUES (b-2542',Nunc In Foundation','03 79 81 18 92','ac.f.... 1 row(s) affected
      83 13:31:39 INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES (b-2546', Lorem Ipsum Dolor Corp.', '09 03 14 76 02',... 1 row(s) affected
      84 13:31:39 INSERT INTO companies (company_id,company_name_phone.email.country,website) VALUES (b-2550','Auctor Mauris Corp.','05 62 87 14 41','eget... 1 row(s) affected
      85 13:31:39 INSERT INTO companies (company_id,company_name.phone.email.country,website) VALUES (b-2554",Cras Vehicula Aliquet Industries",03 37 86 .
      86 13:31:39 INSERT INTO companies (company_id.company_name.phone.email.country.website) VALUES (b-2558', Nulla Integer Vulputate Corp.', '04 57 50 84... 1 row(s) affected
      87 13:31:39 INSERT INTO companies (company_id,company_name.phone.email.country,website) VALUES (b-2562',Dictum Eu Corp.','03 04 73 67 31',donec...
      88 13:31:39 INSERT INTO companies (company_id,company_name.phone.email,country,website) VALUES (b-2566','Aliquam PC','01 45 73 52 16','scelerisque....
      89 13:31:39 INSERT INTO companies (company_id,company_name.phone.email.country,website) VALUES (b-2570', Aliquet Diam Limited', '02 76 61 47 46', 'eu....
      90 13:31:39 INSERT INTO companies (company_id,company_name,phone,email.country,webste) VALUES (b-2574',Orci Adipiscing Limited','03 18 00 77 81','d... 1 row(s) affected
      91 13:31:39 INSERT INTO companies (company_id,company_name,phone,email,country,website) VALUES (b-2578','Dui Quis Institute',''06 93 28 72 81',luctus...
      92 13:31:39 INSERT INTO companies (company_id_company_name_phone_email_country,website) VALUES (b-2582',Nibh Phasellus Corporation','08 48 67 57 4... 1 row(s) affected
      93 13:31:39 INSERT INTO companies (company_id,company_name_phone.email.country,website) VALUES (b-2586',Tempor Diam Institute', 05 01 94 80 87',c...
      94 13:31:39 INSERT INTO companies (company_id_company_name_phone.email_country,website) VALUES (b-2590', Euismod Mauris Institute', '02 13 69 54 85', .... 1 row(s) affected
      95 13:31:39 INSERT INTO companies (company_id,company_name.phone.email.country.website) VALUES (b-2594', Et Magnis Ltd', '03 53 76 11 76', 'dapibus.li...
      96 13:31:39 INSERT INTO companies (company_id.company_name.phone.email.country.website) VALUES (b-2598','Aliquam laculis Lacus Corp.','04 43 07 91 ... 1 row(s) affected
      97 13:31:39 INSERT INTO companies (company id.company name.phone.email.country,website) VALUES (b-2602', Placerat LLP', '05 43 67 24 41', tellus.susp...
                                                                                                                                            1 row(s) affected
     98 13:31:39 INSERT INTO companies (company_id_company_name_phone,email_country_website) VALUES (b-2606',Sed Est Corp.\'04 58 02 37 91',tempor.bib... 1 row(s) affected
      99 13:31:39 INSERT INTO companies (company_id.company_name.phone.email.country.website) VALUES (b-2610', Egestas Nunc Sed Limited', '06 01 02 70 4... 1 row(s) affected
    100 13:31:39 INSERT INTO companies (company id,company name,phone,email,country,website) VALUES (b-2614',Rutrum Non Inc.','02 66 31 61 09',heque...
```

I fent un SELECT comprovo que companyies té 100 files tal com el seu csv.

```
1 • SELECT * FROM bbdd_tasca_4.companies;
```



2.1.7.1 Variant de la solució anterior utilitzant instrucció LOAD DATA (millor opció)

He investigat i una millor opció per introduir dades massives seria fent un LOAD ja que si tinguéssim més d'un milió de files al csv, ja no podríem passar per l'Excel per generar el codi INSERT. Ho provaré creant una taula companies 2 només per aprendre més.

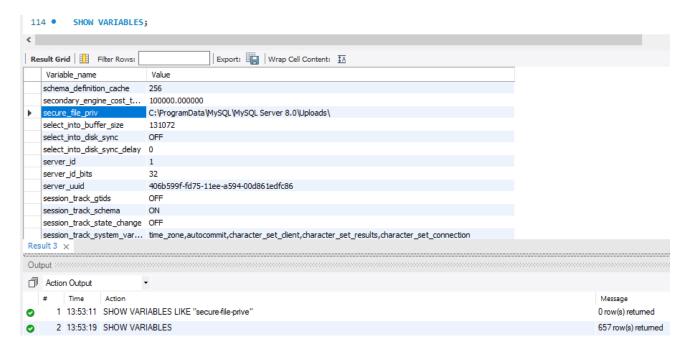
Executo el següent codi i...



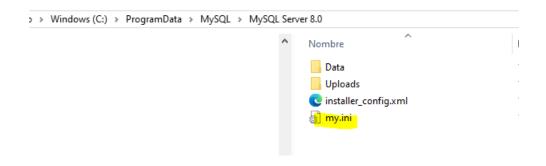
I em surt un error, que al principi em va costar una mica analitzar-ho

13:48:12 LOAD DATA INFILE 'D:/Classificació nova/36. Des pers/9. Form/2. Formació Data Science and Business Analytics\Curs Data Analyst (IT Academy 03-04-24)/4. Sprint 4. Modelat SQL/7. Tasca 4/companies.csv' INTO TABLE companies_2 FIELDS TERMINATED BY ',' ENCLOSED BY ''' IGNORE 1 ROWS Error Code: 1290. The MySQL server is running with the --secure-file-priv option so it cannot execute this statement 0.000 sec

Faig aleshores el següent per si puc veure aquesta variable secure-file-priv al Workbench:



Efectivament, veig que existeix i veig que té un valor que sembla com si fos una ruta on accepta pujar als arxius al MySQL Server. Potser és per un tema de seguretat. Investigo on es troba aquesta configuració, i és a l'arxiu marcat en groc.



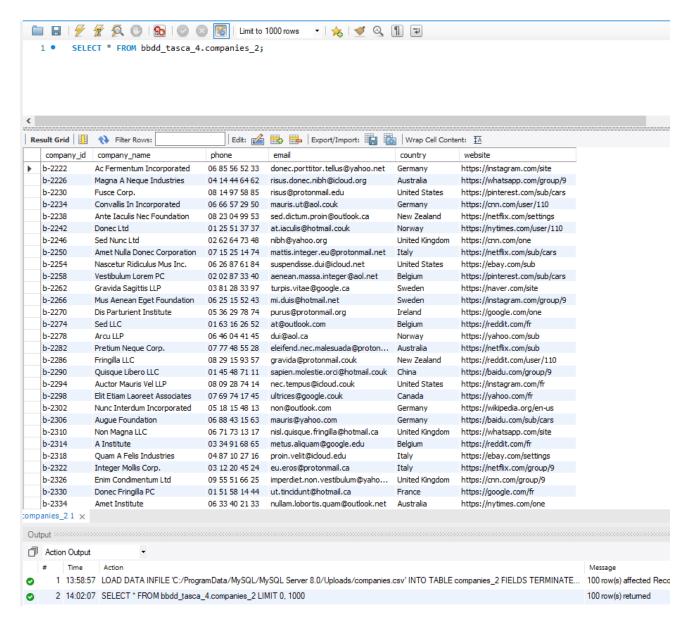
```
# This variable is used to limit the effect of data import and export operations, such as # those performed by the LOAD DATA and SELECT ... INTO OUTFILE statements and the # LOAD_FILE() function. These operations are permitted only to users who have the FILE privilege. secure-file-priv="C:/ProgramData/MySQL/MySQL Server 8.0/Uploads"
```

I efectivament m'explica el que em suposava. Aleshores mouré l'arxiu *companies.csv* a la carpeta *Uploads*.

Executo i ja funciona:

```
116 •
         LOAD DATA
117
          INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/companies.csv'
 118
          INTO TABLE companies_2
 119
         FIELDS TERMINATED BY ','
         ENCLOSED BY '"'
 120
         IGNORE 1 ROWS;
121
 122
Output ::::
Action Output
1 13:58:57 LOAD DATA INFILE 'C:/ProgramData:/MySQL/MySQL Server 8.0/Uploads/companies.csv' INTO TABLE companies_2 FIELDS TERMINATED... 100 row(s) affected Records: 100 Deleted: 0 Skipped: 0 Warnings: 0
```

Comprovo que companies 2 té 100 files com abans.



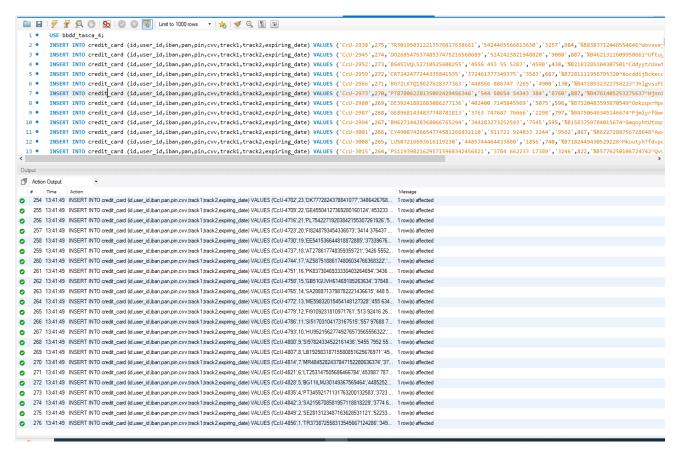
Posteriorment elimino la taula companies 2 que només era un exemple.

2.1.8 Automatitzar en Excel per generar codi SQL per inserir les dades a la taula credit_card

També genero el codi en Excel a la fila J2 a l'arxiu credit card.xlsx

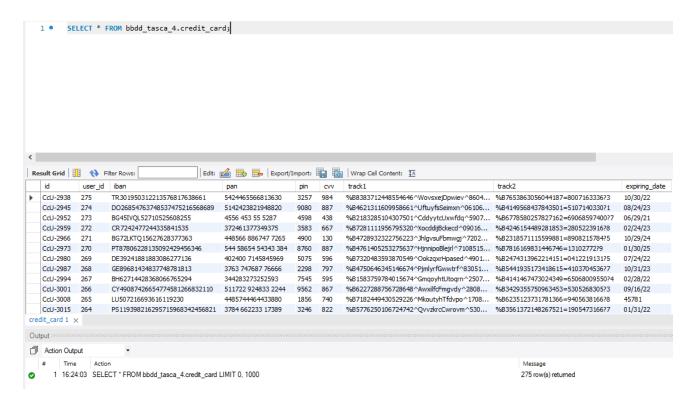
```
=CONCATENAR("INSERT INTO credit_card (";$A$1;",";$B$1;",";$C$1;",";$D$1;",";$E$1;",";$F$1;",";$G$1;",";$H$1;",";$I$1;") VALUES (";""";A2;"";",";B2;",";"";C2;"";",";"";D2;"";",";"";E2;"";",";F2;",";"";G2;"";",";"";H2;"";H2;"";",";"";D2;"";");")
```

Arrossego per totes les files i ja tinc tot el codi que pegaré al Workbench.



Si haguéssim inserit aquestes dades a la taula *credit_card* abans d'haver inserit dades de la taula *users*, no ens ho hauria deixat fer, doncs *user_id* és clau externa en aquesta taula de la taula *users*. També podem veure que qualsevol fila amb un valor *user id* està a la taula *users*.

I fent un SELECT veig que credit cart té 275 files tal com el seu csv.



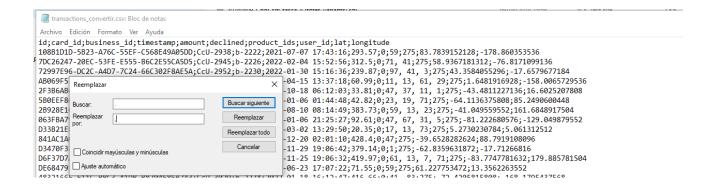
2.1.9 Automatitzar en Excel per generar codi SQL per inserir les dades a la taula transactions

Per la taula de *transactions* fem el mateix, convertim l'arxiu a transactions.xlsx i posem una fórmula al final de les columnes i de la fila 2 i arrastrem fins al final. Això sí, com que hi ha varis valors en format decimal amb punt, abans d'obrir l'Excel he de canviar la configuració regional per tal que el punt en l'Excel sigui el decimal. Sinó quan l'obri perdria els punts. He canviat en format regional el decimal per punt, però quan obro l'Excel perdo els decimals del camp Longitud i Latitud. Busco una alternativa. Torno a posar el decimal com el format regional.

El que farem és obrir l'arxiu *transactions.csv* en bloc de notes i canviarem tots els punts per decimals amb un reemplaçar massiu.

Originalment tenim:

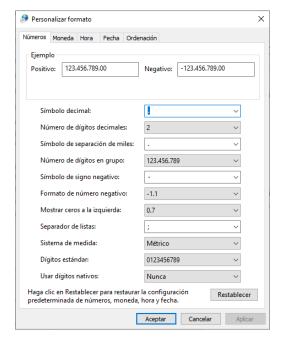
```
| d; card_id; business_id; timestamp; amount; declined; product_ids; user_id; lat; longitude
| 108B1D1D-5B23-A76C-55EF-C568E49A05DD; CcU-2938; b-2222; 2021-07-07 17:43:16; 293.57; 0; 59; 275; 83.7839152128; -178.860353536 |
| 7DC26247-20EC-53FE-E555-B6C2E55CA5D5; CcU-2945; b-2226; 2022-02-04 15:52:56; 312.5; 0; 71, 41; 275; 58.9367181312; -76.8171099136 |
| 72997E96-DC2C-A4D7-7C24-66C302F8AE5A; CcU-2952; b-2230; 2022-01-30 15:16:36; 239.87; 0; 97, 41, 3; 275; 43.3584055296; -17.6579677184 |
| AB069F53-965E-A2A8-CE06-CA8C4FD92501; CcU-2959; b-2234; 2021-04-15 13:37:18; 60.99; 0; 11, 13, 61, 29; 275; 1.6481916928; -158.0065729536 |
| 2F3B6AB6-147D-EB08-FE8D-9A4E2EA9DBD5; CcU-2966; b-2238; 2021-10-18 06:12:03; 33.81; 0; 47, 37, 11, 1; 275; -43.4811227136; 16.6025207808 |
| 5B0EEF86-B8A1-EFAA-5EE1-27E7DC8F54A4; CcU-2973; b-2242; 2022-01-06 01:44:48; 42.82; 0; 23, 19, 71; 275; -64.1136375808; 85.2490600448 |
| 2B928E1C-EC14-A760-0A75-87147764906A; CcU-2986; b-2246; 2021-08-10 08:14:49; 383.73; 0; 59, 13, 23; 275; -41.049559552; 161.6848917504 |
| 603FBA79-99EC-66FB-29F7-25726D1764A5; CcU-2987; b-2250; 2022-01-06 21:25:27; 92.61; 0; 47, 67, 31, 5; 275; -81.222680576; -129.049879552
```



L'obro amb Excel i ja amb comes els camps amount i lat i longitude:

id	card_id	business_id	timestamp	amount	declined	product_ids	user_id	lat	longitude
108B1D1D-56	CcU-2938	b-2222	07/07/2021 17:43	293,57	0	59	275	83,7839152	-178,860354
7DC26247-20	CcU-2945	b-2226	04/02/2022 15:52	312,5	0	71, 41	275	58,9367181	-76,8171099
72997E96-DC	CcU-2952	b-2230	30/01/2022 15:16	239,87	0	97, 41, 3	275	43,3584055	-17,6579677
AB069F53-96	CcU-2959	b-2234	15/04/2021 13:37	60,99	0	11, 13, 61, 29	275	1,64819169	-158,006573
2F3B6AB6-14	CcU-2966	b-2238	18/10/2021 6:12	33,81	0	47, 37, 11, 1	275	-43,4811227	16,6025208
5B0EEF86-B8	CcU-2973	b-2242	06/01/2022 1:44	42,82	0	23, 19, 71	275	-64,1136376	85,24906
2B928E1C-EC	CcU-2980	b-2246	10/08/2021 8:14	383,73	0	59, 13, 23	275	-41,0495596	161,684892
063FBA79-99	CcU-2987	b-2250	06/01/2022 21:25	92,61	0	47, 67, 31, 5	275	-81,2226806	-129,04988
D33B21EB-A6	CcU-2994	b-2254	02/03/2022 13:29	20,35	0	17, 13, 73	275	5,27302308	5,06131251
841AC1A0-90	CcU-3001	b-2258	20/12/2021 2:01	428,4	0	47	275	-39,6528283	88,7919108
D3470F3E-96	CcU-3008	b-2262	29/11/2021 19:06	379,14	0	1	275	-62,8359632	-17,7126682
D6F37D7A-A	CcU-3015	b-2266	25/11/2021 19:06	419,97	0	61, 13, 7, 71	275	-83,7747782	179,885782
DE684792-5D	CcU-3022	b-2270	23/06/2021 17:07	71,55	0	59	275	61,2277535	13,3562264

Ara torno al format regional a posar els decimals com a punts:



Passo l'arxiu csv a Excel i ja tinc el format dels números tal com vol el MySQL.

1 A	R	С	ט	E	F	G	Н	I	J
id	card_id	business_id	timestamp	amount	declined	product_ids	user_id	lat	longitude
108B1D1D-5	ECcU-2938	b-2222	07/07/2021 17:43	293.57	0	59	275	83.7839152	-178.860354
7DC26247-20	CcU-2945	b-2226	04/02/2022 15:52	312.5	0	71, 41	275	58.9367181	-76.8171099
72997E96-D0	CcU-2952	b-2230	30/01/2022 15:16	239.87	0	97, 41, 3	275	43.3584055	-17.6579677
AB069F53-9	6 CcU-2959	b-2234	15/04/2021 13:37	60.99	0	11, 13, 61, 29	275	1.64819169	-158.006573
2F3B6AB6-1	4 CcU-2966	b-2238	18/10/2021 6:12	33.81	0	47, 37, 11, 1	275	-43.4811227	16.6025208
5B0EEF86-B8	3 CcU-2973	b-2242	06/01/2022 1:44	42.82	0	23, 19, 71	275	-64.1136376	85.24906
2B928E1C-E	CcU-2980	b-2246	10/08/2021 8:14	383.73	0	59, 13, 23	275	-41.0495596	161.684892
063FBA79-99	9 CcU-2987	b-2250	06/01/2022 21:25	92.61	0	47, 67, 31, 5	275	-81.2226806	-129.04988
D33B21EB-A	CcU-2994	b-2254	02/03/2022 13:29	20.35	0	17, 13, 73	275	5.27302308	5.06131251
841AC1A0-9	(CcU-3001	b-2258	20/12/2021 2:01	428.4	0	47	275	-39.6528283	88.7919108
D3470F3E-96	5 CcU-3008	b-2262	29/11/2021 19:06	379.14	0	1	275	-62.8359632	-17.7126682
D6F37D7A-A	CcU-3015	b-2266	25/11/2021 19:06	419.97	0	61, 13, 7, 71	275	-83.7747782	179.885782
DE684792-5	CcU-3022	b-2270	23/06/2021 17:07	71.55	0	59	275	61.2277535	13.3562264
4832166E-E1	CcU-3029	b-2274	18/01/2022 16:12	416.66	0	41, 83	275	-72.4205816	-168.179544
1788C881-E	CcU-3036	b-2278	16/01/2022 8:19	486.83	0	7, 2, 79	275	55.9134551	110.554949
4B3FF0B5-15	CcU-3043	b-2282	04/12/2021 23:00	63.12	0	83	275	48.9972804	-50.5497548

A continuació al final de la columna J i a la fila 2 genero el codi que em servirà per inserir les dades.

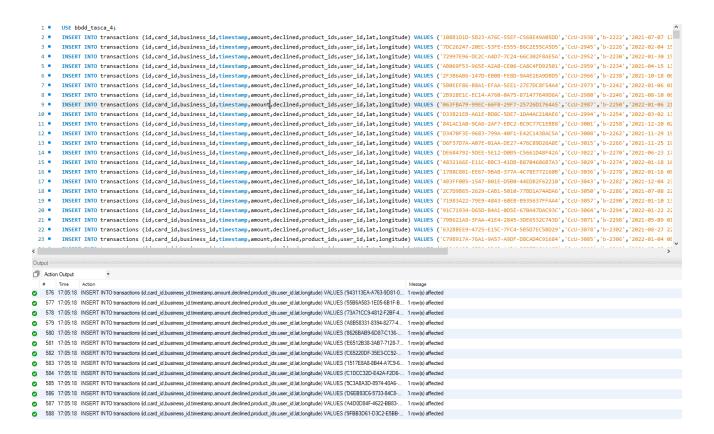
```
=CONCATENAR("INSERT INTO transactions (";$A$1;",";$B$1;",";$C$1;",";$D$1;",";$E$1;",";$F$1;",";$G$1;",";$H$1;",";$I$1;",";$J$1;") VALUES (";"";A2;"";",";"";B2;"";",";"";C2;"";",";"";K2;"";",";E2;",";F2;",";"";G2;"";",";H2;",";H2;",";J2;");")
```

Es important notar, que el valor K2 correspon a la columna *timestamp* (es troba a D2) però modificada. Com que *timestamp* (columna D) ens ve en format 07/07/2021 17:43, a l'Excel fem una altra columna auxiliar a K2 fent:

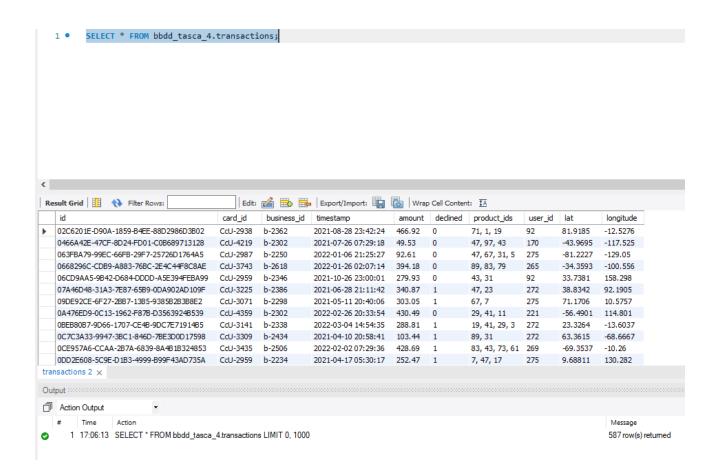
=TEXTO(D2;"AAAA-MM-DD HH:MM:SS"). D'aquesta manera el format de *timestamp* sí ens l'interpretarà al MySQL amb el format adequat i no ens ha suposat molta feina fer-ho a l' Excel.

Com sempre arrastrem al final per generar automàticament totes les fórmules i després pego el codi al Workbench.

id	card_id	business_id	timestamp	amount	declined	product	_ids	user_id	lat	longitude	timestamp_aux		
108B1D1D-5	CcU-2938	b-2222	07/07/2021 17:43	293.57	0		59	275	83.7839152	-178.860354	2021-07-07 17:43:16	INSERT INTO 1	transaction
7DC26247-20	CcU-2945	b-2226	04/02/2022 15:52	312.5	0	71, 41		275	58.9367181	-76.8171099	2022-02-04 15:52:56	INSERT INTO t	transaction
72997E96-DC	CcU-2952	b-2230	30/01/2022 15:16	239.87	0	97, 41, 3		275	43.3584055	-17.6579677	2022-01-30 15:16:36	INSERT INTO 1	transaction
AB069F53-96	CcU-2959	b-2234	15/04/2021 13:37	60.99	0	11, 13, 6	1, 29	275	1.64819169	-158.006573	2021-04-15 13:37:18	INSERT INTO 1	transaction
2F3B6AB6-14	CcU-2966	b-2238	18/10/2021 6:12	33.81	0	47, 37, 1	1, 1	275	-43.4811227	16.6025208	2021-10-18 06:12:03	INSERT INTO 1	transaction
SDUCCEOU DO	Cell 2072	h 2242	06/01/2022 1-44	42 Q2	n	22 10 7	1	275	6/ 1106076	05 2/IONE	2022 01 06 01-44-49	INICEDT INITO	transaction

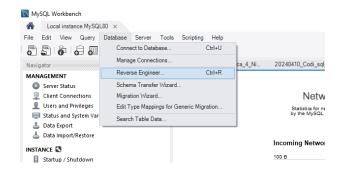


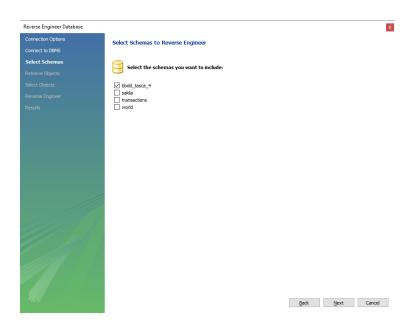
Finalment fem un SELECT, i veiem que *transactions* té 587 files com l'arxiu *csv* i les dades en el format que tocava.

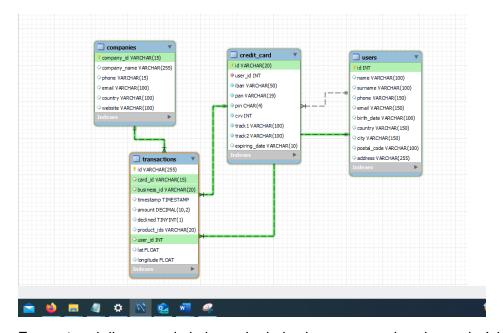


Com que hem pogut introduir totes les dades, significa que les clau externes que havíem creat inicialment són correctes.

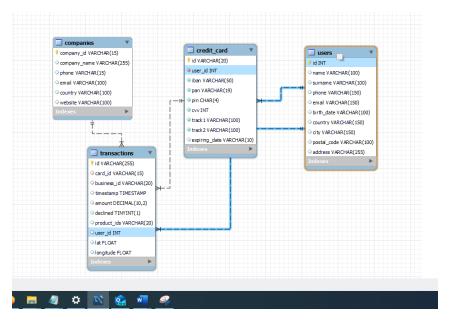
2.1.10 Diagrama de la base de dades a MySQL Workbench







Es mostra el diagrama de la base de dades i es marquen les claus primàries i les claus externes. company_id és clau primària a companies i externa a transactions. Es la relació de 1 a molts, ja que una mateixa companyia pot fer moltes transaccions.id a credit_cards, és una clau primària i té una relació amb transactions de 1 a molts. Es així doncs per una mateixa targeta de crèdit es poden fer moltes operacions.



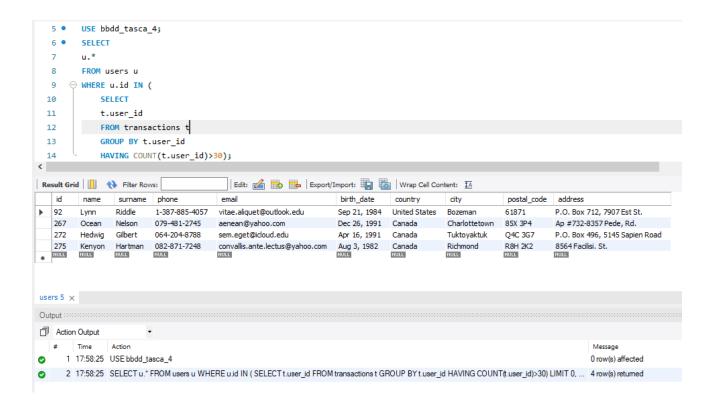
user_id apareix com una clau externa a credit_card i té una relació de 1 a molts ja que un usuari podria tenir més d'una targeta de crèdit. També user id apareix com una clau externa a credit card.

2.2 Exercici 1

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

2.2.1 Codi sql de la solució

La consulta més interna faig una agrupació de *user_id* i miro quants tenen més de 30 transaccions amb el HAVING COUNT. Gràcies al filtre del HAVING COUNT, puc mostrar com a sortida de la subconsulta t.user_id. Aleshores la consulta principal, mitjançant un WHERE em permet mostrar tots els camps de la taula d'usuaris amb més de 30 transaccions, que al final són 4 usuaris.

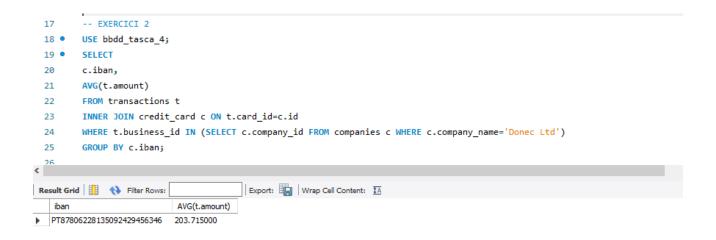


2.3 Exercici 2

Mostra la mitjana d'amount per IBAN de les targetes de crèdit a la companyia Donec Ltd, utilitza almenys 2 taules.

2.3.1 Codi sql de la solució

Sabem que *credit_card* i *transactions* estan relacionades per la variable id i *card_id*. Aleshores amb un INNER JOIN podem relacionar aquestes dues taules. Només volem fer l'agrupació del *iban* i la mitjana del *iban* de les companyies *t.business_id* de la taula *transactions* que compleixin la condició de la subconsulta més interna, és a dir que *t.business_id* sigui el de la companyia 'Donec Ltd'





3 Nivell 2

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.

3.1 Creació taula Estat targetes de crèdit

En aquest cas crearem una vista en comptes de fer un CREATE TABLE. El motiu és el següent. L'objectiu d'una vista és no contaminar la base de dades bbdd_tasca_4. del esquema principal amb altres i considerem més apropiat fer una vista

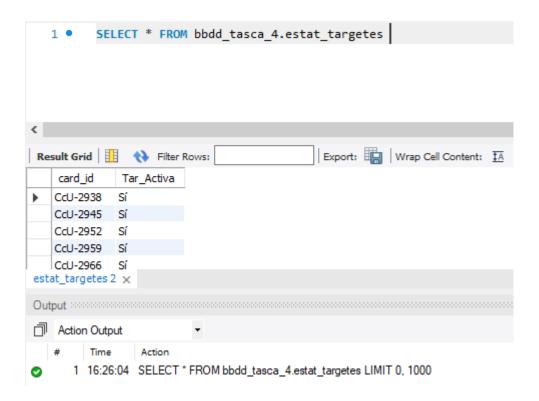
Al fina la solució més senzilla és utilitzant un ROW NUMBER().

La subconsulta més interna el que fem bàsicament és enumerar tots els registres que tinguin el mateix *t.card_id* (identificador de targetes de crèdit), amb numeració diferent segons el camp *t.timestamp* de forma descendent, de manera que els registres amb la columna *row_num*, que valguin 1, 2, i 3 seran les últimes tres transaccions per la mateixa *t.card_id*, és a dir les últimes tres dates de *t.timestamp*. Amb el WHERE següent ens quedem amb aquests tres registres per *t.card_id*.

Posteriorment la subconsulta de més alt nivell que l'anterior fem una agrupació de t_row_num.card_id i sumem el camp de declinades (t._row_num.declined). Finalment a la consulta de més alt nivell (la principal) és on identifiquem si una targeta estarà activa o no tal com passa a la vida real. Es a dir si falles amb el pin 3 vegades o es declina tres vegades la operació consecutivament, la targeta passarà a inactiva.

```
31 •
             USE bbdd_tasca_4;
 32
         -- Creem una vista que s'anomena estat_targetes
 33 •
            CREATE VIEW estat_targetes AS
 34
             SELECT
 35
             t_num_dec.card_id,
             IF (t_num_dec.num_declinades=3, 'No','Si') Tar_Activa
 36
 37
             FROM (SELECT
 38
                 t_row_num.card_id,
 39
                  SUM(t_row_num.declined) num_declinades
 40
                  FROM (SELECT
 41
                     t1.card_id,
                      t1.id,
 42
 43
                      t1.timestamp,
 44
                      t1.declined,
 45
                      ROW_NUMBER() OVER (PARTITION BY t1.card_id ORDER BY t1.timestamp DESC) row_num
                      FROM transactions t1) t_row_num
 46
 47
                      WHERE t_row_num.row_num<=3
                      GROUP BY t_row_num.card_id) t_num_dec;
 48
<
Output
Action Output
       Time
                Action
     1 16:24:52 USE bbdd_tasca_4
                                                                                                                          0 row(s) affected
     2 16:24:52 CREATE VIEW estat_targetes AS SELECT t_num_dec.card_id, IF (t_num_dec.num_declinades=3, "No", "Si") Tar_Activa FROM (SEL... 0 row(s) affected
```

A continuació mostrem el resultat d'alguns registres de la vista estat_targetes.

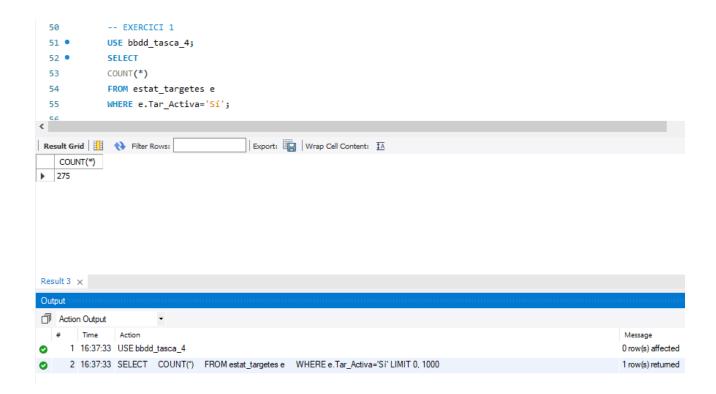


3.2 Exercici 1

Quantes targetes estan actives?

3.2.1 Codi sql de la solució

Veiem que totes les targetes estan Actives, doncs la vista estat_targetes té 275 files.



4 Nivell 3

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:

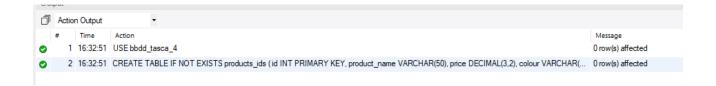
4.1 Exercici 1

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

4.1.1 Creació estructura taula product_ids

```
264
        -- NIVELL 3 -----
265
        -- EXERCICI 1
        USE bbdd_tasca_4;
266
267 • ⊖ CREATE TABLE IF NOT EXISTS products_ids (
        id INT PRIMARY KEY,
269
        product_name VARCHAR(50),
        price FLOAT,
270
        colour VARCHAR(20),
271
272
        weight DECIMAL(3,1),
273
        warehouse_id VARCHAR(20)
274
        ) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci;
```

Veient les dades de *products.csv* creem l'estructura de la taula *products_ids*.



4.1.2 Automatitzar en Excel per generar codi sql per inserir les dades a la taula product_ids

Aquestes són les fórmules Excel que introduïm per cada fila de l'arxiu products.xlsx i ho arrastrem per totes les files i ja tenim el codi per introduir totes les dades a la taula *products ids*

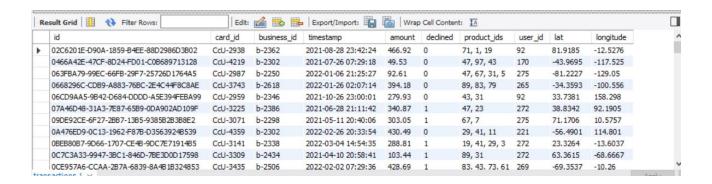
```
=CONCATENAR("INSERT INTO products_ids
(";$A$1;",";$B$1;",";$C$1;",";$D$1;",";$E$1;",";$F$1;") VALUES
(";A2;",";""";B2;"";",";EXTRAE(C2;2;LARGO(C2));",";""";D2;""";",";E2;",";""";F2;""";");")
```

Peguem el codi al Workbench.

```
USE bbdd tasca 4;
   3 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (1, 'Direwolf Stannis',161.11, '#7c7c7c',1, 'WH-4');
          INSERT INTO products ids (id, product name, price, colour, weight, warehouse id) VALUES (2, 'Tarly Stark', 9.24, '#919191', 2, 'WH-3');
          INSERT INTO products ids (id,product name,price,colour,weight,warehouse id) VALUES (3, 'duel tourney Lannister',171.13, '#d8d8d8',1.5, 'WH-2');
          INSERT INTO products ids (id.product name.price.colour.weight.warehouse id) VALUES (4, 'warden south duel',71.89, '#111111',3, 'WH-1');
          INSERT INTO products ids (id, product name, price, colour, weight, warehouse id) VALUES (5, 'skywalker ewok', 171.22, '#dbdbdb', 3.2, 'WH-0');
          INSERT INTO products ids (id, product name, price, colour, weight, warehouse id) VALUES (6, 'dooku solo', 136.60, '#c4c4c4', 0.8, 'WH--1');
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (7,'north of Casterly',63.33,'#b7b7b7',0.6,'WH--2');
  10 •
          INSERT INTO products ids (id, product name, price, colour, weight, warehouse id) VALUES (8, 'Winterfell', 32.37, '#383838', 1.4, 'WH--3');
  11 •
          INSERT INTO products ids (id, product name, price, colour, weight, warehouse id) VALUES (9, 'Winterfell', 76.40, '#b5b5b5', 1.2, 'WH--4');
  12 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (10, 'Karstark Dorne',119.52, '#f4f4f4',2.4, 'WH--5');
  13 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (11, 'Karstark Dorne',49.70, '#141414',2.7, 'WH--6');
  14 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (12,'duel Direwolf',181.60,'#a8a8a8',2.1,'WH--7');
  15 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (13,'palpatine chewbacca',139.59,'#2b2b2b',1,'WH--8');
  16 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (14,'Direwolf',147.53,'#c4c4c4',2,'WH--9');
  17 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (15,'Stannis warden',194.29,'#dbdbdb',1.5,'WH--10');
  18 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (16, 'the duel warden',180.91, '#666666',3, 'WH--11');
  19 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (17,'skywalker ewok sith',91.89,'#7c7c7c',3.2,'WH--12');
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (18, 'Karstark warden',148.91, '#c4c4c4',0.8, 'WH--13');
  20 •
  21 •
         INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (19,'dooku solo',60.33,'#3f3f3f',0.6,'WH--14');
  22 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (20,'warden Karstark',91.96,'#b5b5b5',1.4,'WH--15');
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (21, 'duel Direwolf',96.90, '#e2e2e2',1.2, 'WH--16');
  23 •
         INSERT INTO products_ids (id,product_name,price;colour,weight,warehouse_id) VALUES (22, 'chewbacca mustafar',150.02, '#fcfcfc',2.4, 'WH--17');
  24 •
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (23,'riverlands north',169.96,'#545454',2.7,'WH--18');
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (24,'south duel tourney',48.99,'#aaaaaa',2.1,'WH--19');
         INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (25,'skywalker_ewok',157.53,'#2b2b2b',1,'bM+-20');
          INSERT INTO products_ids (id,product_name,price,colour,weight,warehouse_id) VALUES (26,'Stark Karstark',53.01,'#898989',2,'WH--21');
         INSERT INTO products ids (id, product name, price, colour, weight, warehouse id) VALUES (27, 'Stannis riverlands', 172.93, '#7a7a7a', 1.5, 'WH--22');
Output :
Action Output
                                                                                                                             Message
94 16:39:00 INSERT INTO products_ids (id.product_name.price.colour.weight.warehouse_id) VALUES (93, Littlefinger Tarly Stark*, 192.60, "#efefef*, 1.2, "W....
     95 16:39:00 INSERT INTO products_ids (id.product_name.price.colour.weight,warehouse_id) VALUES (94.\in south Winterfell",51.94,\;\;\;\;43a3a3a',2.4.\;\;\;\;WH=8... 1 row(s) affected
96 16:39:00 INSERT INTO products_ids (id.product_name.price.colour.weight.warehouse_id) VALUES (95, riverlands north Direwolf'.6.90, #c1c1c1'.2.7,...
                                                                                                                            1 row(s) affected
     97 16:39:00 INSERT INTO products_ids (id.product_name.price.colour.weight.warehouse_id) VALUES (96,'dooku solo',20.92,'#282828',2.1,'WH-91')
   98 16:39:00 INSERT INTO products_ids (id.product_name.price.colour.weight,warehouse_id) VALUES (97.jinn Winterfell :65:25;#bababa'.1,"WH-92')
      99 16:39:00 INSERT INTO products ids (id.product name.price.colour,weight,warehouse id) VALUES (98, 'Direwolf Littlefinger', 38.33, '#bababa', 2, 'WH-93')
100 16:39:00 INSERT INTO products_ids (id_product_name.price.colour,weight.warehouse_id) VALUES (99.the duel',151.78;#212121',1.5,WH-94')
                                                                                                                            1 row(s) affected
    101 16:39:00 INSERT INTO products_ids (id_product_name_price_colour,weight,warehouse_id) VALUES (100,'south duel',40.43,'#6d6d6d',3,'WH-95')
                                                                                                                            1 row(s) affected
```

Es generen 100 files tal com l'arxiu products.csv

4.1.3 Construir taula intermèdia que relacioni transaccions i productes venuts



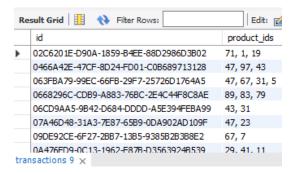
El camp *product_ids* de la taula *transactions* no ens anirà bé per poder extreure la informació, així doncs fem la taula intermèdia encara que haurem de generar codi amb sql. Aquesta taula intermèdia tindrà dos camps, el camp *id* i el camp *product_ids* nou.

Hem optimitzat el codi per tal d'evitar de crear taules auxiliar a la base de dades i després haver de eliminar-les.

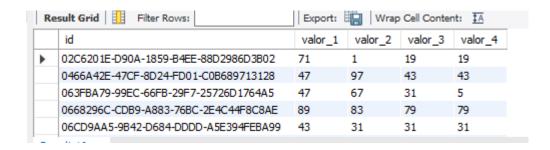
```
83 • USE bbdd_tasca_4;
        -- CREATE TABLE trans_x_product_aux
 84
        CREATE TABLE trans_x_product
 86
            WITH trans_x_product_aux_1 AS
 87
            (SELECT
 89
            t.product_ids
            FROM transactions t),
 90
 91
           el camp products_ids, que com a molt hem vist que només pot contenir 4 productes per transacció, el separem en diferents columnes
 92
            trans_x_product_aux_2 AS
 93
            (SELECT
 94
            t.id,
            TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 1), ',', -1)) AS valor_1,
 95
 96
            TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 2), ',', -1)) AS valor_2,
 97
            TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 3), ',', -1)) AS valor_3,
            TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 4), ',', -1)) AS valor_4
 98
            FROM trans_x_product_aux_1 t),
100
        -- Hem de convertir la consulta anterior en una taula de dues columnes
            trans_x_product_aux_3 AS
101
102
            (SELECT
103
            id,
104
            t.valor 1
            FROM trans_x_product_aux_2 t
106
            UNTON
            SELECT
107
109
            t.valor 2
            FROM trans_x_product_aux_2 t
110
            LINTON
111
            SELECT.
112
113
            t.id,
114
            t.valor_3
            FROM trans_x_product_aux_2 t
115
116
117
            SELECT
118
            t.id,
119
120
            FROM trans_x_product_aux_2 t)
 121
           SELECT
 122
           t3.id.
 123
           t3.valor 1 product ids
 124
           FROM trans_x_product_aux_3 t3
 125
           ORDER BY t3.id;
```

Anem mostrant resultats intermedis del codi anterior, tot i que tal com s'ha fet donarà directament la taula que necessitem *trans_x_product*.

El codi sql el que fem bàsicament és crear consultes temporals, amb la clàusula WITH. La primera subconsulta *trans_x_product_aux_1*, obté els camps *id* i *product_ids* de la taula *transactions*.

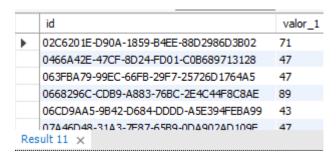


La següent subconsulta, que anomenem trans_x_product_aux_2 ens dóna el següent:



Bàsicament separem el camp *product_ids*, que com a molt sempre tindrà 4 productes en diferents columnes¹

Finalment, la última consulta la *trans_x_product_aux_3*, ens agafa les combinacions *id* amb els diferents columnes de valors, per obtenir la taula que busquem. Es interessant que fem UNION i no UNION ALL, ja que sinó hi haurien files repetides si s'analitza bé com funciona la consulta .



Posteriorment després d'una clàusula WITH niada, al final a d'acabar amb una consulta i per això posem

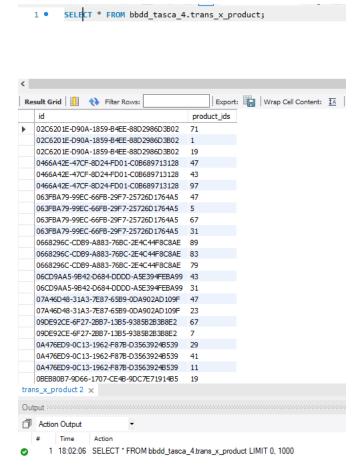
```
121 SELECT
122 t3.id,
123 t3.valor_1 product_ids
124 FROM trans_x_product_aux_3 t3
125 ORDER BY t3.id;
```

Com que tot la solució és temporal, encapsulem les tres subconsultes amb un CREATE TABLE al principi de tot per obtenir la taula *trans_x_product* que desitjàvem, on es pot veure el resultat de l'execució directa:

¹ Si el número de productes que es poguessin assignar a la columna *product_ids* fos indeterminat. El codi de trans_x_product_aux_2 no ens serviria. Sabem que hauríem de posar un codi on utilitza les avantatges dels arxius *json*. No ho fem tant general, doncs fem la solució òptima que ens serveix pel nostre exercici.



Mostrem el resultat de la taula *trans_x_product* intermèdia que necessitàvem, tot i que encara no hem acabat, ja que ens falten les relacions.



4.1.4 Relacions taula intermèdia trans_x_product amb base de dades bbdd_tasca_4

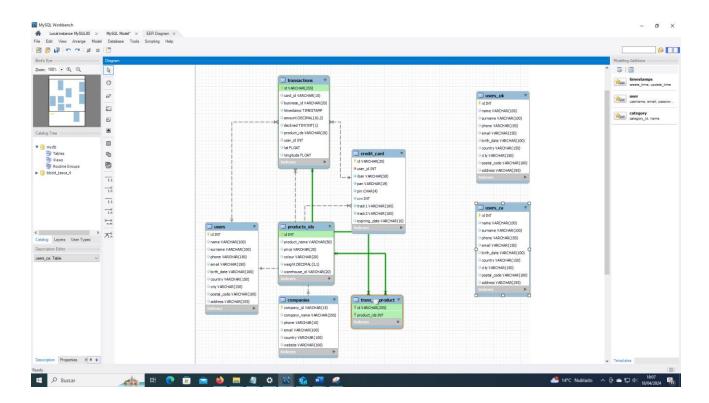
```
324 • SHOW CREATE TABLE trans_x_product;
325     -- ALTER TABLE trans_x_product DROP FOREIGN KEY trans_x_product_ibfk_2;
326 • ALTER TABLE trans_x_product MODIFY product_ids INT;
327 • ALTER TABLE trans_x_product ADD PRIMARY KEY (id,product_ids);
328     -- Afegeixo les claus externes, encara que el nom de la clau externa me'l triarà ell
329 • ALTER TABLE trans_x_product ADD FOREIGN KEY trans_x_product(id) REFERENCES transactions(id);
330 • ALTER TABLE trans_x_product ADD FOREIGN KEY trans_x_product(product_ids) REFERENCES products_ids(id);
331 • SHOW CREATE TABLE trans_x_product;
```



Veiem que hem creat una clau primària composta, formada per *id* i *product_ids* i també les relacions de cada un dels camps amb les altres taules de la base de dades *bbdd_tasca_4*.

4.1.5 Diagrama de la base de dades al afegir taula intermèdia trans_x_product

Aquí podem veure com ha canviat la base de dades al afegir trans_x_product.



4.1.6 Codi sql que dóna resposta a la pregunta de l'exercici

```
-- EXERCICI 1 -----
333
        -- Númera de ventas per cada producte
334
        USE bbdd tasca 4;
335 •
        SELECT
336
        t.product_ids,
337
        COUNT(t.product ids)
338
        FROM trans x product t
339
        GROUP BY 1
340
        ORDER BY 1;
341
```

I aquesta és la taula que conté el número de vendes per cada un dels productes, que veiem que com que només té 26 files, només s'han venut 26 productes dels 100 que té la taula *products_ids*.

	product_ids	COUNT(t.product_ids)
•	1	61
	2	65
	3	51
	5	49
	7	54
	11	48
	13	60
	17	61
	19	49
	23	68
	29	49
	31	47
	37	51
	41	53
	43	65
	47	62
	53	58
	59	45
	61	57
	67	68
	71	54
	73	47

