SPRINT 04

Nivell 1 Exercici 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:

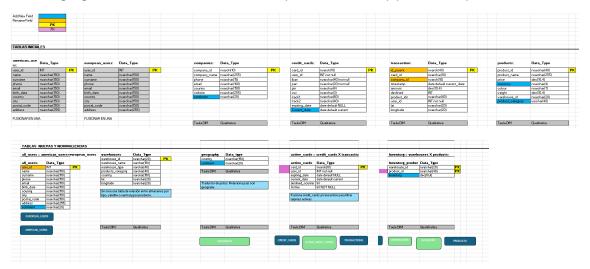
1.- Estudi del disseny de la base de dades.

Llistat de taules inicials de partida de l'exercici amb especificació de el tipus de dada en cada camp.

Llistat de taules noves addicionals, bé sigui per normalització o per expansió futur del model de dades.

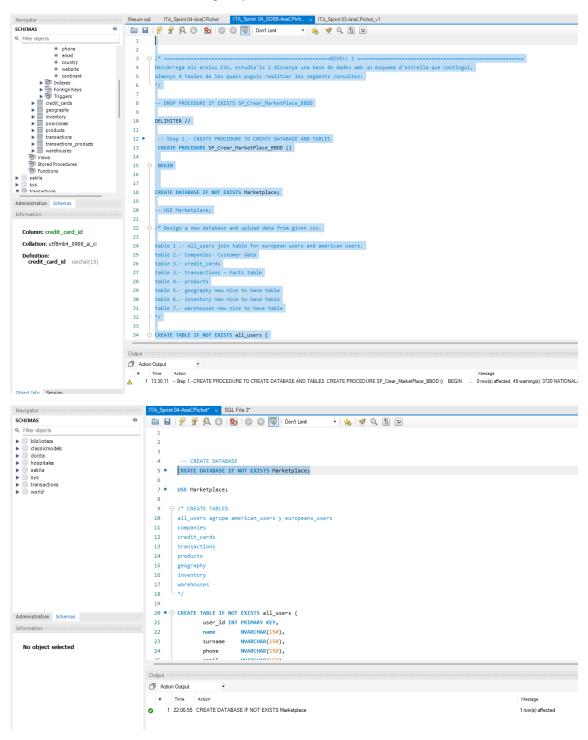
La premissa en aquest punt de l'exercici és que el model proposat sigui en estrella, per tant hi haurà una única taula de fets, transaccions i la resta seran de dimensions.

La taula geografies l'utilitzo com a taula auxiliar per traduir el camp país en camp continent.



Pas 1. Creació de la base de dades, taules i foreign keys.

Creo un procediment a la base de dades per poder tornar enrere en la creació de la base de dades. El procediment conté la descripció de les taules que existiran a la base de dades anomenada MarketPlace i la creació de les foreign keys.





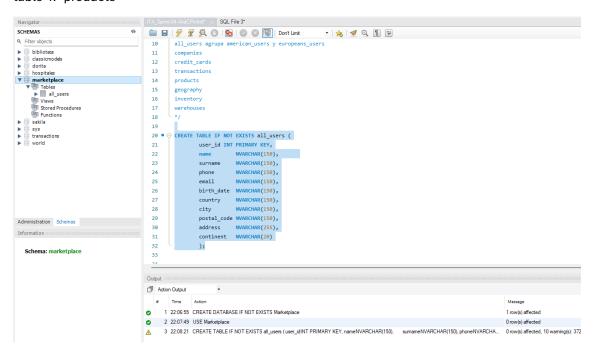
Pas 2.- Creació de les taules.

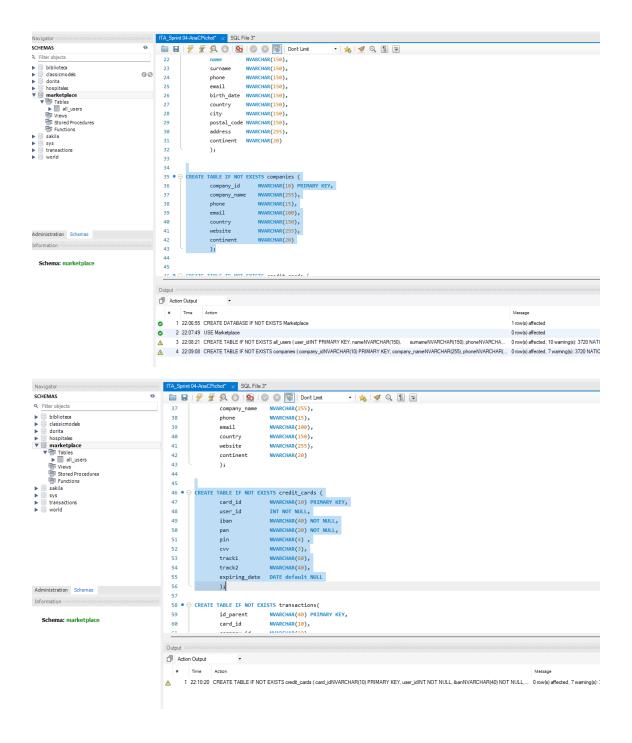
Creo una taula única pels usuaris. Aquesta taula és la suma de les taules usuaris Europa i usuaris América. Indexo el camp mail companyia com a únic per evitar que un mail estigui associat a més d'una persona.

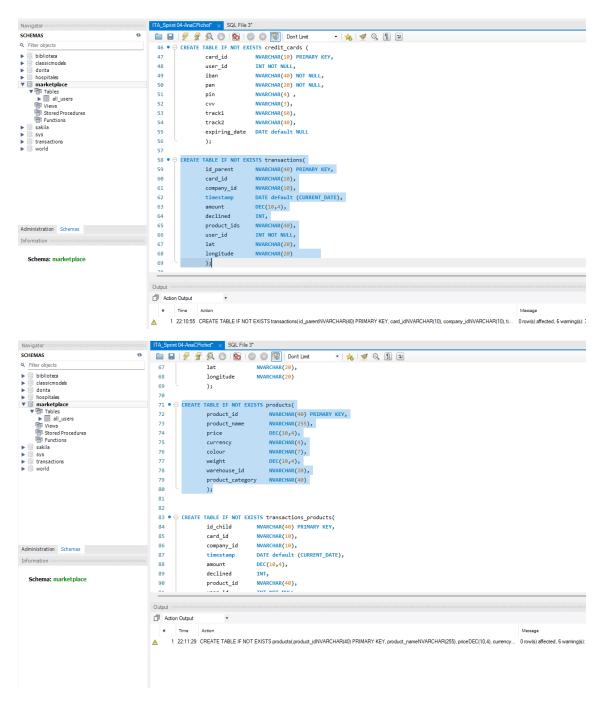
Creo la resta de taules.

- table 1 .- All_users join table for european users and american users.
- table 2.- Companies- Customer data
- table 3.- credit_cards
- table 3.- transactions Facts table

table 4.- products







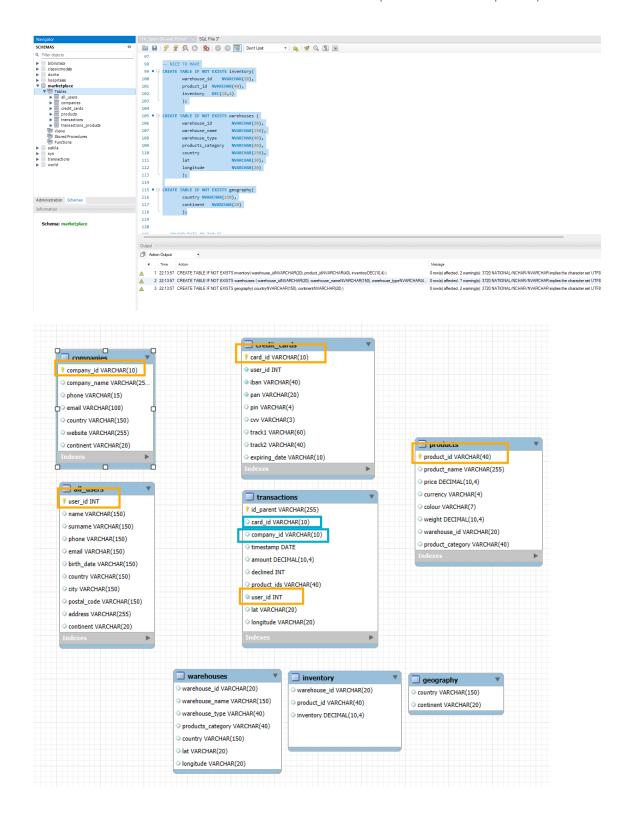
Pas 3.- Taules per un futur.

Afegeixo altres taules interessants per un futur però no les connecto perquè em demanen un model d'estrella i jo faria un model de floc de neu.

table 5.- geography new nice to have table

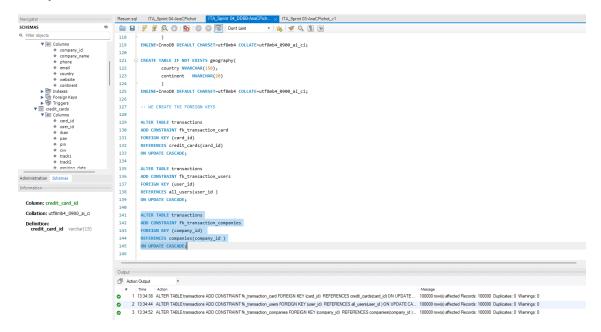
table 6.- inventory new nice to have table

table 7.- warehouses new nice to have table



Pas 4.- Creo les relacions entre les taules.

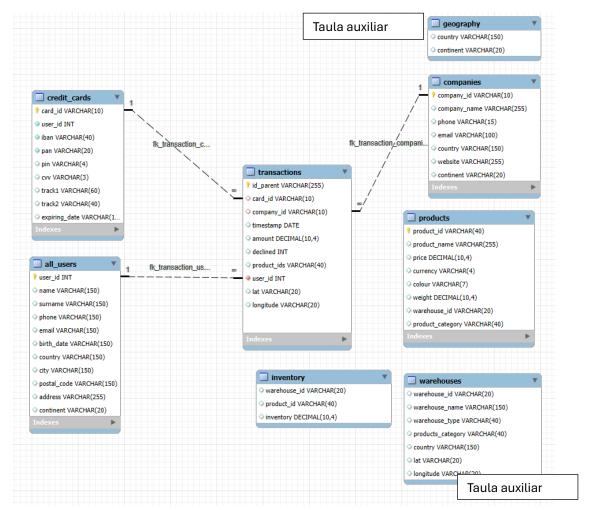
La taula productes la deixo sense connectar perquè en el següent exercici demana que es creí una nova taula transaccions_productes. Si la vulgues connectar en aquest punt del sprint la taula transaccions hauria de contenir un camp product_id únic és a dir una transacció per product_id. En aquest punt aquesta taula no conté aquest camp sinó un únic camp amb més d'un producte.



Pas 5.- Disseny final de la base de dades

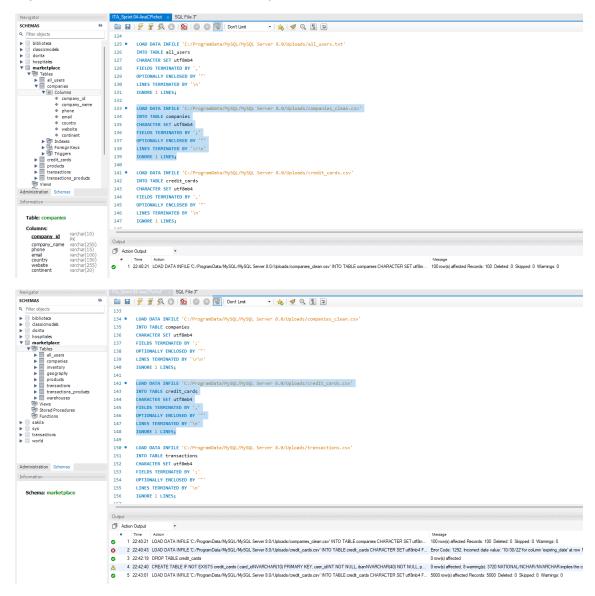
Dissenya una base de dades amb un esquema d'estrella que contingui, almenys 4 taules.

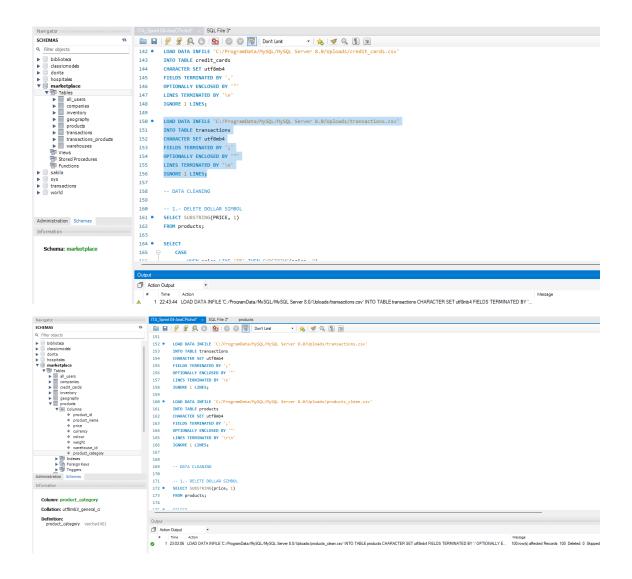
Les taules són : 1) Transactions, 2) crèdit_cards, 3) companyies, 4) all_users.



Pas 6.- Pujar les dades des de CSV.

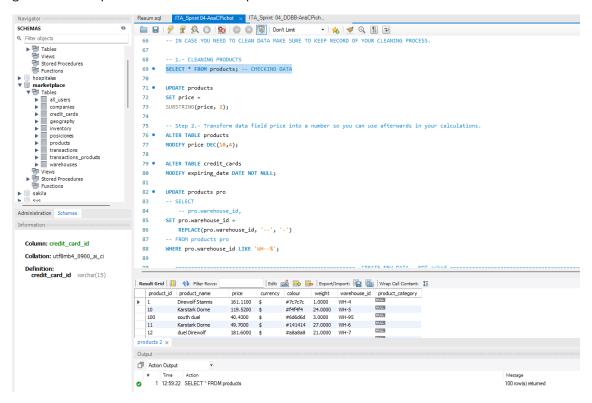
Pujo les dades a les taules utilitzant els csv que ens donen.





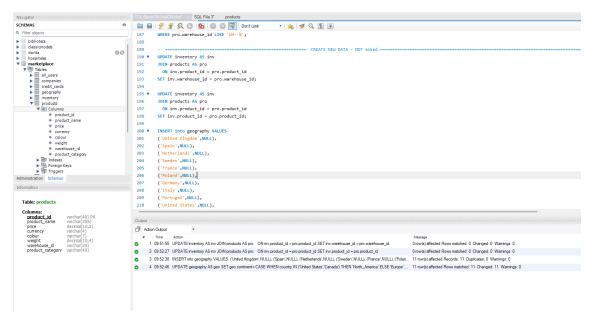
Pas 7.- Neteja de dades.

Netejo algunes dades com format preu per fer-lo decimal, netejar el camp warehouses on el guionet està duplicat o el format del camp dates.



Pas 8.- Creació d'un dataset.

Creo un set de dades per les taules que será bo afegir en un futur.

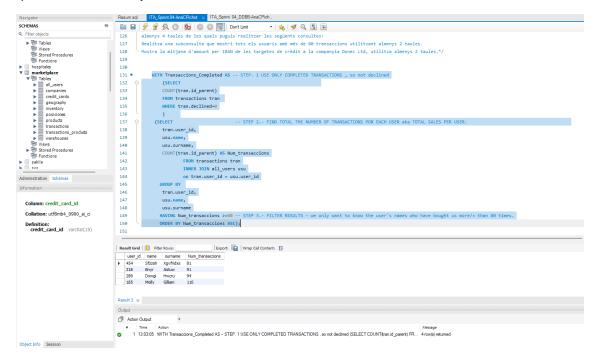


Nivell 1 Exercici 1.

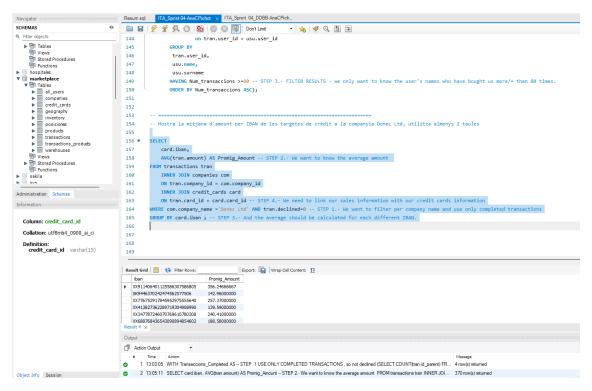
Descàrrega els arxius CSV, estudia'ls i dissenya una base de dades amb un esquema d'estrella que contingui, almenys 4 taules de les quals puguis realitzar les següents consultes:

1. Realitza una subconsulta que mostri tots els usuaris amb més de 80 transaccions utilitzant almenys 2 taules.

Utilitzo un primer select amb with per filtrar les transaccions completades i després faig un select que també conte dos joins.



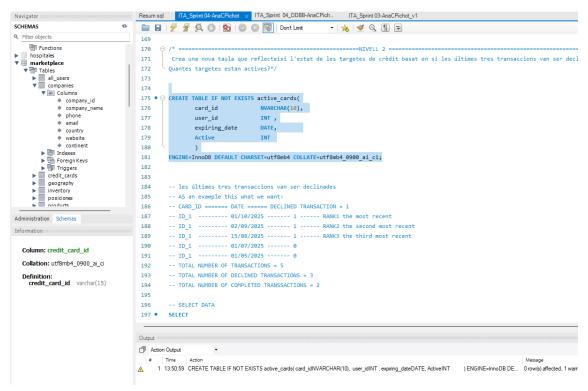
2. Mostra la mitjana d'amount per IBAN de les targetes de crèdit a la companyia Donec Ltd, utilitza almenys 2 taules.



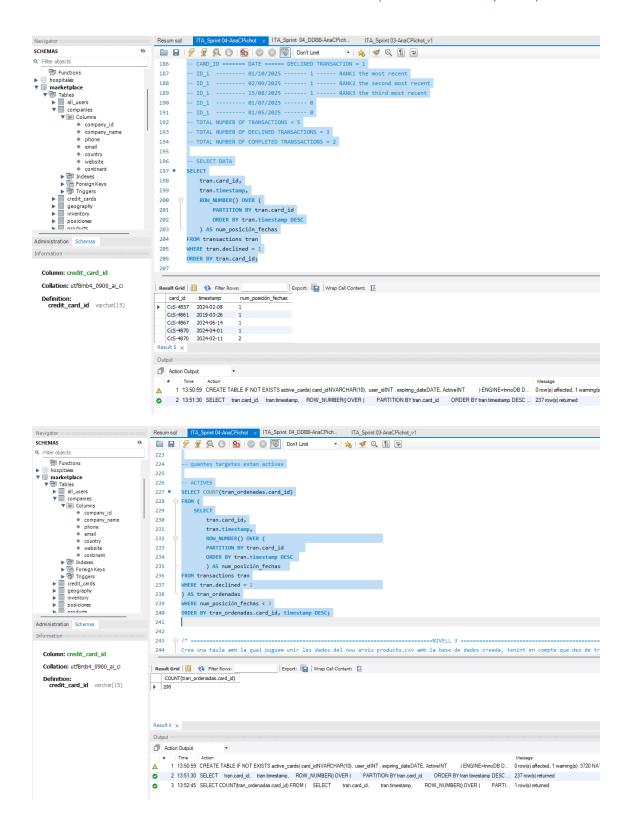
Nivell 2 - Exercici 1

Crea una nova taula que reflecteixi l'estat de les targetes de crèdit basat en si les tres últimes transaccions han estat declinades aleshores és inactiu, si almenys una no és rebutjada aleshores és actiu. Partint d'aquesta taula respon:

Quantes targetes estan actives?



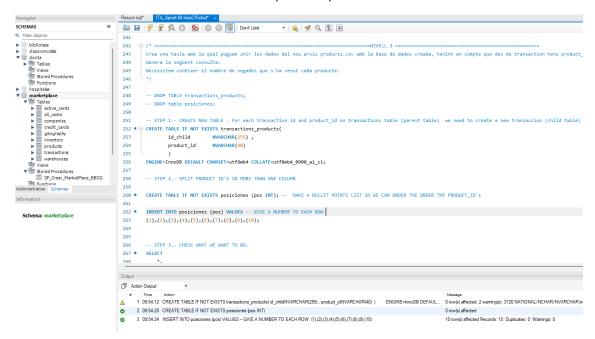




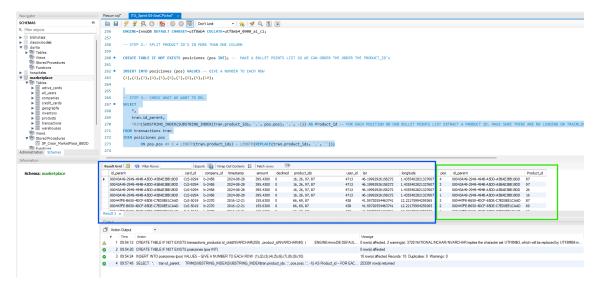
Nivell 3 Exercici 1

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:

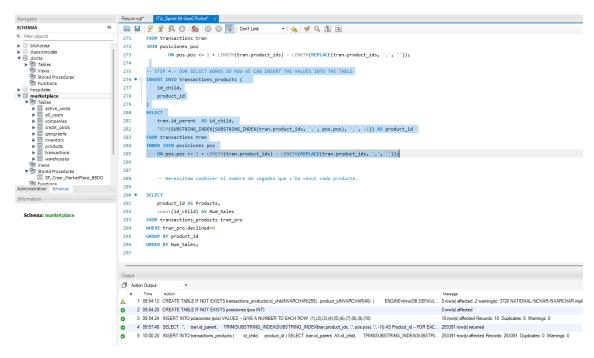
Creació de la nova taula per crear una única transacció per producte. Per similitud seria com tenir un document de comandes on hi ha una línia per cada producte.



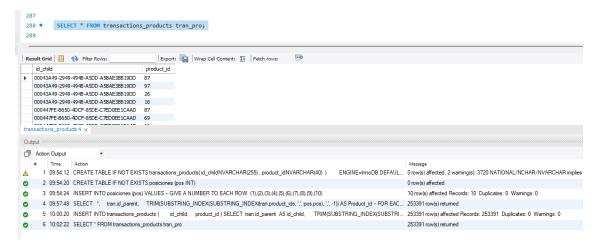
El resultat seria la taula verda. Una línia per producte i l'origen seria la taula blava una transacció per grup de productes.



Com el select anterior funciona bé, inserto els resultats.



Comprovo la nova taula



Exercici 1

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

