**Tasca S2.01. Nocions bàsiques SQL**

**Fernando Poblete**

**Objetivo:**

Repassar les nocions bàsiques per a l'ús de base de dades relacionals. En aquest sprint, iniciaràs la teva experiència pràctica amb una base de dades que conté informació d'una empresa dedicada a la venda de productes en línia. En aquesta activitat, t'enfocaràs en dades relacionades amb les transaccions efectuades i la informació corporativa de les empreses que van participar.

**Nivell 1**

**- Exercici 1**

A partir dels documents adjunts (estructura\_dades i dades\_introduir), importa les dues taules. Mostra les característiques principals de l'esquema creat i explica les diferents taules i variables que existeixen. Assegura't d'incloure un diagrama que il·lustri la relació entre les diferents taules i variables.

- Aunque solamente contamos con 2 tablas, podemos observar que el modelo de datos relacionales sigue el esquema de estrella. La tabla de ‘transacciones’ es la tabla de hechos, teniendo un gran volumen de datos, y llaves secundarias que pueden conectan a las llaves primarias de diversas tablas de dimensiones, como ‘credit\_card’, ‘user’ y ‘company’. En este ejemplo, solo contamos con la tabla de ‘company’

A diagram of a fact table

AI-generated content may be incorrect.

**Esquema Diagrama de estrella - Wikipedia**

A screenshot of a computer

AI-generated content may be incorrect.

**Diagrama entidad relacion**

- En el diagrama entidad-relacion, o modelo relacional, podemos observar que las tablas ‘company’ y ‘transaction’ tienen cada una, un primary key, que identifica los registros en cada tabla de manera única.

- Asi mismo, la tabla de transacciones se conecta a la de compañia mediante el ‘company\_id’, que en la tabla de ‘company’ simplemente se llama ‘id’.

- La relacion entre ‘company’ y ‘transaction’, es de 1:N, porque una compañia puede tener multiples transacciones, mientras que cada transacción, solo puede ser asignada a 1 compañia.

- A continuación, un listado de las variables de cada tabla con una breve descripción de cada variable:

|  |  |  |
| --- | --- | --- |
| **Variable** | **Tipo de dato** | **Descripción** |
| Id | VARCHAR(15) | PRIMARY KEY |
| company\_name | VARCHAR(255) | Nombre de la compañia |
| Phone | VARCHAR(15) | Telefono |
| email | VARCHAR(100) | email |
| country | VARCHAR(100) | pais |
| website | VARCHAR(255) | Sitio Web |

**Descripción Tabla ‘company’**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Tipo de dato** | **Descripción** |
| id | VARCHAR(255) | PRIMARY KEY |
| credit\_card\_id | VARCHAR(15) | REFERENCES credit\_card(id) |
| company\_id | VARCHAR(20) | FOREIGN KEY (company\_id) REFERENCES company(id) |
| user\_id | INT | REFERENCES user(id) |
| lat | FLOAT | Coordenadas latitud |
| longitude | FLOAT | Coordenadas longuitud |
| timestamp | TIMESTAMP | Tiempo de registro |
| amount | DECIMAL(10, 2) | Valor de la transaccion |
| declined | BOOLEAN | Valor 1= TRUE -> Transaccion declinada |

**Descripción Tabla ‘transaction’**

**- Exercici 2**

Utilitzant JOIN realitzaràs les següents consultes:

* Llistat dels països que estan fent compres.

- Hacemos un INNER JOIN entre la tabla de compañias y de transacciones, para obtener solamente los datos de compañias que tengan transacciones registradas. Luego obtenemos el listado de paises sin repeticiones utilizando DISTINCT en la selección de datos.

A screenshot of a computer

AI-generated content may be incorrect.

* Des de quants països es realitzen les compres.

- Para obtener la cantidad de paises que realizan compras, simplemente utilizamos la funcion de conteo, COUNT, en la selección de paises sin repetición.

A screenshot of a computer

AI-generated content may be incorrect.

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

- Para obtener la compañia con la media más grande de ventas, tenemos que extraer el promedio de los montos de las transacciones, con la funcion AVG sobre la variable ‘amount’

- Conservamos la estructura de los JOINS, y agrupamos por el ‘id’ de la compañia.

-Mostramos tanto el ‘id’ de la compañia como el ‘company\_name’, sabiendo que la relación de estas dos variables siempre es 1:1.

- Redondeamos el resultado a 2 decimales por claridad, ordenamos los resultados de forma descendente en el promedio calculado, y limitamos los resultados a 1 linea, para obtener la compañia con el mayor valor en esta nueva variable creada.

- La compañia con mayor media de ventas es: ‘Eget Ipsum Ltd’, con ‘id’ de compañia ‘b-2398’.

A screenshot of a computer

AI-generated content may be incorrect.

**(extra)**

- Finalmente verificamos los resultados solamente para las transacciones no declinadas. Vemos que aunque el valor de la media cambia, la compañia con la media más alta sigue siendo la misma.

A screenshot of a computer

AI-generated content may be incorrect.

**- Exercici 3**

Utilitzant només subconsultes (sense utilitzar JOIN):

* Mostra totes les transaccions realitzades per empreses d'Alemanya.

- Primero, mediante una subconsulta, creamos una tabla que contenga todos los ‘id’ de las compañias que se encuenten en Alemania (country = ‘Germany’).

- Después, seleccionamos toda la infomación de la tabla de transacciones, y agregamos la condición, con la clausula WHERE, que me entregue solamente resultados donde el ‘company\_id’ se encuentre en (IN) la tabla creada en el primera parte, que contiene solo identificadores de compañias en Alemania.

- Asi, sin utilizar JOINS, obtenemos todos los datos de transacciones exclusivamente para compañias en Alemania.

A screenshot of a computer

AI-generated content may be incorrect.

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

- En este ejercicio vamos a hacer consultas anidadas en 3 niveles.

- Primero, en una subconsulta en la tabla ‘transaction’, calculamos el valor promedio de todas las transacciones.

- Luego utilizamos este valor calculado, como parámetro en una clausula WHERE, para filtrar los registros de la tabla ‘transaction’, extrayendo unicamente los identificadores de empresa ‘company\_id´, sin repeticiones (las repeticiones se pueden eliminar tambien en el próximo nivel). Con repeticiones obtenemos 297 registros, y al agregar DISTINCT al SELECT, obtenemos solo 70 registros.

- Finalmente, para obtener los nombres de estas empresas y no sus códigos, utilizamos el resultado del paso anterior, de la misma manera que en el ejercicio 3.1, para extraer los nombres de las compañias, condicionando en la clausula WHERE, que sean solamente compañias cuyos ‘id’ se encuentren en el listado creado en el paso anterior.

A screenshot of a computer

AI-generated content may be incorrect.

**(extra)**

- De manera adicional, si queremos ser más exactos, podemos omitir las transacciones declinada, espeficicando mediante una clausula WHERE en el primero nivel, que se consideren solamente los montos de transacciones aceptadas (declined = FALSE).

Esto es importante para poder realizar decisiones de negocios, ya que las transacciones declinadas no representan ingresos para la compañia.

Al hacer esta correción, la cantidad de empresas que han realizado transacciones disminuye de 70 a 50 registros.

A screenshot of a computer

AI-generated content may be incorrect.

* Eliminaran del sistema les empreses que no tenen transaccions registrades, entrega el llistat d'aquestes empreses.
* Primero, extraemos un listado de ‘id\_company’ sin repetición de la tabla de transacciones.
* Luego utilizaremos el resultado anterior, como una subconsulta, para filtrar mediante una clausula WHERE, las empresas que no se encuentren en el listado, con el operador NOT IN.
* No obtenemos registros, lo que quiere decir, que todas las 100 compañias de la tabla ‘company’ tienen transacciones registradas en la tabla ‘transactions’

A screenshot of a computer

AI-generated content may be incorrect.

- Como era de esperarse, al ejecutar el código para eliminar las empresas sin transacciones, ningún registro se ve afectado.

- Incluimos para futuras referencias, un pantallazo de la opción que se debe cambiar para poder eliminar datos de una tabla (recordar reconectarse después de cambiar esta opción).

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.