**Introducción a Consultas a bases de Datos**

**Sobre la importancia de las consultas (o Queries) en una base de datos relacional.**

Las consultas son la parte fundamental de una base de datos. De qué te sirve hacer una consulta o un Query efectivo?

Pues bien, una consulta bien hecha podría, incluso, *salvar un negocio o una empresa.* Las empresas tienen una gran necesidad por información correcta, exacta, oportuna; dicha información, se puede extraer generalmente de las bases de datos que ya tenemos (de aquí la importancia de las *consultas*).

En las bases de datos tenemos muchas veces *clientes, productos, proveedores* y diferentes tipos de entidades que, por separado, no tendrían mucho sentido; pero, cuando se comienzan a unir, a través de Queries, puedes presentar: un informe, o una tendencia, etc...; de tal forma que, las personas que están encargadas de *maniobrar* la empresa para un lado o para el otro,... tengan la información que necesitan para tomar las decisiones correctas.

Entonces, alrededor de toda esa cultura de “*extraer y transformar información”,* se han creado varias especialidades en el esquema de datos que tiene que ver con, por ejemplo, ***ETL*** (transformación de datos), ***Business Intelligent*** (extracción & utilización de datos para la toma correcta de decisiones) e, incluso, que tiene que ver con: ***Machine leearning*** (toma los datos existentes de una base de datos y los presenta de tal manera que haga sentido, que sea coherente y, sobre todo, que sea útil).

*En resumen, los Queries son una parte fundamental de una base de datos;* pues, nos permiten conocer las estructuras de los **Query** y resolver cualquier duda que tenga cualquier miembro de la empresa que pueda ser trabajada a través de una base de datos.

En resumen, con los ***Queries*** es que se brinda valor a una organización al final del día; pues, es lo que nos trae la información, *el diario vivir dentro de una organización*; con lo cual, es que se podrán tomar decisiones de gestión e inversión diariamente. Entonces, para traer toda esa información de nuestro interés, de una base de datos, necesitas hacer ***Queries.*** Los ***Queries*** son la forma en la que estructuramos las preguntas que le vamos a hacer a la base de datos; es decir, al final del día, nuestro camino a seguir para encontrar todas las respuestas.

Ahora, lo complicado es que hay muchas formas de hacer un ***Queri,*** desde lo más básico a lo más complejo.

***Vamos de las Queries básicas a las complejas: Select***

Empecemos a practicar la habilidad de convertir nuestras preguntas en ***Queries.*** El ***Queri*** tiene básicamente dos partes: la sentencia ***Select*** y la sentencia ***From*** y, en varias ocasiones, nos encontramos con la sentencia ***Where.***



***Select*** nos permite traer los datos que queremos mostrar, todos ***(\*)*** o de alguna ***columna*** en especifica; en este caso puntual, queremos traer los datos de la columna o atributo ***city*** y una cuenta ***count(\*)*** la cual se proyecta como ***total*** (en adelante explicaremos más a fondo los dos últimos elementos señalados en negrilla).

*Un dato:* En caso que usted no sólo desee ver los datos de una única columna, de una tabla, puede agregar más. Simplemente, dentro de la misma sentencia ***Select,*** debe llamar al otro u otros campos en cuestión por medio de separaciones de coma; por ejemplo: ***SELECT city, name*** (Ahora veríamos, ademas de ***city,*** la columna ***name).***

En ***From*** indicamos de dónde vamos a seleccionar esos datos que queremos mostrar con el ***select;*** es decir, más precisamente, de que ***tabla***; en contexto, entonces, estamos indicando que queremos los datos de la columna o atributo ***city*** de la tabla ***people.***

Ahora, en ***Where,*** el filtro de la consulta y recolección de datos se vuelve más exigente; pues, con ***where,*** lo que se establece es una condición; en consecuencia, dicha condición debe cumplirse para poder seleccionar los datos deseados de una columna para una tabla en especifica. En contexto, cuando definimos un where con ***active = true*** le estás diciendo a nuestra base de datos de trabajo actual que desea solamente los datos de la columna ***city*** de la tabla ***people*** que cumplan con la condición ***active = true.*** De hecho, ***active*** vendría siendo también otra columna o atributo de la misma tabla ***people*** y ***true*** es uno de los posibles valores que puede tener, dentro de dicha tabla, la columna ***active.***

***Group by*** lo que nos permite es agrupar los datos, justamente, según un criterio pasado; en este caso, el criterio pasado fue: ***city;*** es decir, vamos a agrupar los datos de la tabla ***people***, que cumplen con la condición ***active = true*** y son de la la columna o atributo ***city,*** dentro de la misma columna ***city*** (nos mostrará la info de ***city***).

***Order by*** define de qué manera se organizarán los datos recolectados; en este caso, nos ayuda a definir de qué manera ordenar los datos seleccionados de la columna ***city***, de la tabla ***people***, que cumplen con la condición ***active = true*** (la totalidad de ellos); por ejemplo, definir el orden de los datos seleccionados de tal manera que: se *enlisten de mayor a menor* las ciudades ***(city)*** según la cantidad de personas que tenga cada una.

Por último, en ***Having,*** sedefine otro parámetro de limitación, va a aplicar otro filtro o condición mucho más especifico; en este caso, va a filtrar (elegir) el ***total*** de los datos seleccionados de la columna ***city***, de la tabla ***people*** y que tengan un ***active = true;*** pero, además, *cuando haya 2 o más de 2 personas en cada ciudad* ***(total >= 2).*** Tengo la sensación que ***total*** es otra columna o atributo de la tabla ***people.***

***un dato:*** La sentencia ***Select \**** significa *selecciona todos los datos;* sin embargo, qué es *todo?* SQL interpreta esta solicitud, por sí sola, como muy ambigua y tiene razón (por lo cual te lanza un error). Es decir, a ***Select \**** se le debe acompañar en todos los casos con otra sentencia. Por ejemplo, evaluemos nuestro primer caso de ***Query:***



En efecto, con esto estamos diciendole a la base de datos que nos traiga absolutamente todo ***(SELECT \*)***; pero, recuerde, falta algo más,... Queremos todo, sí, pero de la tabla ***posts (FROM posts),*** que nos muestre absolutamente todo de ahí (sus campos, columnas o atributos).

De hecho, generalmente, te encuentras las ***Queries*** con una tercera sentencia declarada que es, como te contabamos, ***Where;*** que nos permite filtrar cuáles de esos datos, de la tabla ***posts***, queremos ver.



En este caso, sólo hemos filtrado los datos de la tabla ***posts*** que tienen como fecha de publicación una fecha igual o superior al año ***2024.*** Es decir, la base de datos nos arrojará todas las noticias, ***posts,*** que sean del ***2024*** para ***adelante*** en su ***fecha\_publicacion.***

***Nota:*** *para SQL “>” es igual a “>=”, por lo que en el caso anterior se incluye también al año 2024. Sin embargo, “<” no es lo mismo que “<=”.* **En resumen, con *Queries* y *Select,* filtramos información.**

En adelante, probaremos ***Queries*** más complejas según el nivel de complejidad que tenga, propiamente, la consulta de los datos que se desean ver; que, obviamente, deberían de ser en la medida unos datos con unas caracteristicas más puntuales o con mayor especificaciones. Sin embargo, para hacer consultas, se requiere tener primeramente datos; dicho eso, empecemos por agregar datos de prueba a nuestras tablas de modelación.

***Propiedad AS***

Ahora, existe una propiedad que no es propiamente de ***Select,*** pero que sirve de asistencia a las consultas o ***Queries*** que hagamos con ***Select;*** esa propiedad es: ***AS.***

***AS*** nos permite pasarle un *alias,* o nombre temporal *que no sustituye de manera permanente al nombre original,* a una o varias columnas; esto, si se acompaña con la sentencia ***Select***; es decir, ***AS*** le pasaría un nombre temporal, cambiando temporalmente el nombre original, a una o varias columnas seleccionadas a la hora de querer hacer una consulta de ellas con ***Select***. Es decir, ***AS*** nos permite sustituir temporalmente, por lo menos sólo en el momento de hacer propiamente la consulta, el nombre original de una o de varias columna; de tal manera que se le asigne un nombre temporal a la columna (o columnas) de nuestro interés.

El proceso es más o menos así: Por medio de ***Select,*** se llama a la *o a las columnas* de nuestro interés con su nombre original (no se olvide hacer separaciones por coma por cada columna llamada)**;**posteriormente, se le debe pasar a cada nombre por separado, e inmediatamente después del nombre de cada columna llamada, la sentencia ***AS;*** entonces, luego de declarar ***AS,*** se le pasaría un nuevo nombre (que sería el *Alias*), el de nuestro interés, a cada columna llamada en cuestión,... tal que así, por ejemplo:



Advertir que lo único que cambia, del *dataset,* y de forma *temporal,* es el nombre original al hacer una consulta con ***Select;*** del resto, todo sigue igual, los mismos datos para las mismas columnas: sólo que ahora está viendo que esas columnas afectadas están siendo llamadas con otro nombre; los cuales, fueron pasados con la propiedad ***AS.***

**Ahora, algo interesante.** En adelante, de hecho, usted podrá llamar cualquiera de las dos formas a sus columnas de interés para manipular sobre ellas, ya sea que las llame con su nombre original o por medio del *Alias* recientemente pasado. Por ejemplo, en futuras ***Queries*** más complejas puedes referirte a los campos en cuestión, ya no con su nombre original; sino, también, con su *Alias* asignado (si así lo deseas).

***Propiedad Count***

***Select,*** también, nos permite hacer datos que se crean en el momento y que nos pueden agregar información; un buen ejemplo de eso sería combinar el ***Select*** con el uso de la sentencia ***Count(),*** ¿y qué hace ***Count()***?

La sentencia ***Count()*** se encarga de contar, ¿contar qué? Contar ***Registros.*** Es decir, a partir de unos parametros pasados para ***Count(),*** éste últimose encargará de contar la totalidad de los registros que hay dentro de una tabla, si y sólo si, se cumplen a cabalidad unas condiciones establecidas según dichos parametros pasados en cuestión, ¿qué parametros? Pues, en principio, los que establecemos dentro de los parentesis de ***Count().***

El ejemplo más básico de uso de la sentencia ***Count(),*** en connivencia con la sentencia ***Select,...*** es que en ***Count()*** se le pase, como parametro, que cuente absolutamente todos (sin filtro alguno) los registros existentes para una tabla seleccionada con la sentencia ***From***; entonces, en este caso, ese “nuevo dato que se crearía en el momento y que nos podría agregar información” sería el dato númerico, que se nos arroja, equivalente a la totalidad de registros que se alcanzarón a contar, en su totalidad, para una tabla seleccionada en cuestión.

Pongamos, por ejemplo, el siguiente caso: 

Esto nos arrojaría un nuevo dato, un simple número, que reflejaría el conteo total de los *registros* que tenemos en nuestra tabla *posts;* es decir, prácticamente, la cantidad de *posts* que tenemos publicados. Tal que así:

***, 22*** sería el número de *registros,* o posts, que tenemos en nuestra tabla *posts* parcialmente.

Ahora, si se da cuenta, ese nuevo dato que se ha creado, el dato númerico, está sustentado por un nuevo atributo, campo o columna que también ha sido creado; esa nueva columna si se percata se llama: ***COUNT(\*)***

Luego, se dará cuenta que, a la nueva columna resultante ***COUNT(\*)*** se le puede pasar un *Alias* tambiéncon la sentencia ***AS;*** esto para recordar o identificar, a dicha columna, de una manera más facil e *identificativa*; pues, es *engorroso* tener que llamar a una columna con una palabra clave como ***Count(\*****).* Probemos:



Esto quiere decir que, nuestra columna inicialmente llamada ***COUNT(\*),*** ahora puede ser llamada también como: ***numero\_posts***; lo cual, tendría mucho más sentido pues ***numero\_posts***, como nueva columna, nos estaría mostrando justamente el número de postsque tiene la tabla ***posts***,... Quedando así:



Por supuesto, hay que decir que realmente este dato no está almacenado de forma permanente en la base de datos. Con todo y eso, lo podemos aún así consultar por medio de ***Select.***

***From + Join***

Como usted ya sabe, por cuestiones de *normalización,* siempre se busca que cada entidad se estudie por separado; esto, al asignarles tablas independientes, aunque relacionadas, a cada una de ellas. Sin embargo, a la hora de presentar un informe, a la hora de traer información que sea valiosa, conviene juntar dichas tablas independientes; pues, nunca dejan de relacionarse y nos ayudaría, justamente, a *desarrollar información.*

En consecuencia, para lograr ese *junte,* como parte de la sentencia ***From,*** tenemos un compañero inseparable de ella***,...*** es la sentencia ***Join.*** Entonces, con ***From*** no sólo se selecciona de qué ***tabla*** queremos visualizar unos datos; sino que, asistiendose de ***join,*** también podríamos unir varias tablas entre sí a través de las relaciones que les asignamos en un principio; es decir, más concretamente, uniendo llaves foráneas con llaves primarias en caso de existir dicho vinculo. Eso es lo que veremos ahora.

Los ***Join*** tratan o parten de una base matematica relacionada a la *teoría de conjuntos;* esto lo veremos, de una forma gráfica, con algo que se llama: ***Diagramas de Venn. Los Diagramas de Venn*** son simplemente circulos que se tocan, en algún punto, para ver dónde está la intersección de conjuntos y, en últimas, ver dónde hay un traslape de valores de un lado y del otro. Veamos esto.

***Diferentes tipos de Join***

**1. Diferencia:** Cuando hablamos de“Diferencia” nos referimos a lo que tenemos en un lado, ignorando lo que está en el otro (se seleccionan los datos que están en una tabla ***A***, por ejemplo, y se excluyen los datos que están en una tabla ***B***); es decir, hay un ***Join*** que se encarga de mostrarnos solamente la información que tiene cada tabla por aparte (el ***Join*** de la ***Diferencia***). Ahora bien, si quisieramos ser más exigentes en nuestra toma de datos con el tipo de ***Join “Diferencia”***; de nosotros depende si queremos excluir, también, los datos de la tabla ***A***, continuando con el mismo ejemplo, que se relacionen o se repitan en la tabla ***B***.

O sea, quiero que quede claro, uno podría hacer una ***diferencia*** al tomar la totalidad de los datos de una tabla ***A***, sin importar si uno de sus datos se relacione o se repita en tabla ***B***; o, bien, siendo más rigurosos, podríamos tomar la totalidad de los datos de una tabla ***A***; pero, excluyendo aquellos datos que se repitan en tabla ***B;*** es decir, ignorando la intersección entre los conjuntos,***...*** En cualquiera de los dos casos, sepa que se estaría hablando de ***Left Join*** (si se supone a priori que la tabla ***A*** está situada en la parte izquierda y la tabla ***B*** está situada en la parte derecha).

Tal que así:



Ahora, suponiendo que la tabla ***A*** estuviera situada en la parte derecha y la tabla ***B*** en la parte izquierda,... estaríamos hablando de ***Diferencia*** de ***Right Join.*** Por suerte conservamos cada tabla en su mismo lado:



**2. Intersección:** Cuando hablamos de “Intersección”, nos referimos a los datos que tenemos en común dentro de cada tabla, los datos que se repiten, que están están en ambas tablas; es decir, más puntualmente, estamos hablando de seleccionar los datos de la intersección entre los conjuntos. Si retomamos el ejemplo anterior estaríamos hablando de seleccionar los datos que están tanto en la tabla ***A*** como en la tabla **B**. Por lo general, este es el tipo de ***Join*** que más se ve y el nombre correcto que recibe es: ***Inner Join*** o en español ***Join interno.***



**3. Unión:** Cuando hablamos de “Unión”, estamos definitivamente hablando de unir ambas tablas, tanto ***A*** como ***B,*** por ejemplo; es decir, acá se seleccionarían y se unirían la totalidad de los datos, sin excepción alguna, de ambas tablas. *Unión* trae absolutamente todos los datos de ambas tablas.



***4. Diferencia simétrica:*** Acá también se tomarían los datos de ambas tablas,... menos los datos que se repitan en cada una de ellas; es decir, los datos que tengan en común ambas tablas, tanto la tabla ***A*** como la tabla ***B,*** por ejemplo. En resumen, se tomarían todos los datos de ambas tablas; pero se omitirían, solamente, los datos de la intersección entre los conjuntos. Se ve así:



***Ahora sí, veamos cada uno de los 4 ejemplos por medio de código SQL en MySQL Workbench,... Le daremos sentido a todo esto.***

***Utilizando la sentencia FROM en MySQL***

***Poniendo en práctica Left Join & Right Join: bajo este tipo de Join uniremos “usuarios” con “posts”.***

Primero que todo nos interesa efectuar la práctica de ***FROM*** en la tabla ***Usuarios***; pero, antes que nada, visualizemos toda la tabla ***usuarios,*** en su conjunto, corriendo las siguientes sentencias que ya conoce:



Se nos *muestra*, como resultado, lo siguiente:



***Left Join.***

***Si hacemos un simil con los Diagramas de Venn, “usuarios” será “tabla A” & “posts” será “tabla B”;*** entonces, en ese orden de ideas, la tabla que usted seleccione de primero será instantaneamente la de la *parte* *izquierda,...* adicionalmente, debajo de ***FROM usuarios,*** se ha declarado la sentencia ***LEFT JOIN***,... con eso estamos afirmando que la prioridad la tendrá la tabla ***usuarios***; es decir, la ***tabla izquierda,*** la ***tabla A*** (Recuerde, la 1ra tabla que se pase siempre será la tabla del *lado izquierdo* o *tabla A*)***.*** Entonces, como la ***tabla A*** tendría la prioridad, de la unión resultante con la ***tabla B***, ***los datos de la tabla A se verían primero; y, de hecho, se verían en su totalidad,... es decir, se verían todos los usuarios independientemente de si los usuarios hayan escrito o no posts. Esto último es lo más importante. Y acuerdese siempre de la relación que tengan ambas entidades; pues, serán fundamentales para las conclusiones que podríamos tener para las Queries que hagamos.***

Sin embargo, retomemos, este conjunto de sentencias está incompleta; recuerde que su intención es, precisamente, hacer un tipo de ***Join*** que es de ***Diferencia,*** al tratar de unir una ***tabla A*** con una ***tabla B***; entonces, para señalar cúal pasaría a ser esa ***tabla B***,... se le pasa el nombre de la segunda tabla de su interés, en este caso ***Posts,*** justo al lado de la sentencia ***LEFT JOIN.*** Tal que así:



Ahora, ya dejamos claro qué tablas vamos a unir y hemos decidido, entre las dos, cual será ***tabla A*** y cual será ***tabla B,*** y que la unión será de tipo ***LEFT JOIN*** entre ambas; sin embargo, hace falta especificar la manera en cómo lograremos que ambas tablas, en definitiva, se unan; y, eso se logra, ***conectando la llave primaria de la tabla A***, ***usuarios,*** que es la tabla de nuestra prioridad por ser la tabla de ***la parte*** ***izquierda*** y por hacer la unión desde un ***left join,...*** ***con la llave foranea de la tabla B*** que, justamente, se relaciona o viene importada de la llave primaria de ***usuarios.*** Lo anterior se logra al *igualar* ambas columnas, la de la llave primaria de ***usuarios*** con la de la llave foranea de ***posts*** (relacionada a usuarios), por medio de la palabra reservada **ON.** Se *igualan* porque se supone que, en efecto, sus datos son iguales entre sí. Veamos esto en código:



En ***usuarios.id, “usuarios”*** hace referencia a la tabla como tal, luego hay una separación de palabras por medio de un punto (.), e ***“id”*** hace referencia, ahora sí, a un campo de la tabla ***usuarios;*** que, como ya debe saber, es el campo correspondiente a la llave primaria de dicha tabla ***usuarios.***

Ahora, en ***posts.usuario\_id, “posts”*** hace referencia a la tabla como tal, luego hay una separación de palabras por medio de un punto (.), y ***“usuario\_id”*** hace referencia, ahora sí, a un campo de la tabla ***posts;*** que, como ya debe saber, es el campo correspondiente a una llave foránea (que se relaciona con la llave primaria de la tabla ***usuarios)*** de dicha tabla ***posts.***

En resumen, las tres sentencias conjuntas, al ser ejecutadas, nos dice lo siguiente:

***Selecciona todo, de la tabla “usuarios”, y unela con la tabla “posts”; reflejando primero la tabla “usuarios”, por ser la de la izquierda, y luego la tabla “posts”, por ser la de la derecha. Dicha unión se logra por medio de las dos llaves que las relaciona; concretamente, para esto, “usuarios.id” debe ser igual a “posts.usuario\_id” (como se supone).***

Con esto tenemos los datos, unidos, de ambas tablas: ***usuarios & posts.***

Ahora, si queremos saber puntualmente que usuarios NO han escrito ningún post; recuerdese de la sentencia ***WHERE*** que nos ayuda a, precisamente, buscar un dato o unos datos en especifico, de una tabla pasada (o, en este caso, de la combinación de dos tablas por medio de sus llaves en común, una propia y otra foránea), según un filtro declarado; en esta ocasión el filtro es: ***querer saber qué posts están vacios y a que usuario le corresponde;*** más concretamente, en qué ***Registros*** las celdas de la tabla ***posts*** estén vacias o ***NULL*** para x o y ***usuario.***

Como en este caso puntual estamos haciendo busquedas especificas con la sentencia ***WHERE;*** pero, **sobre dos tablas unidas**; entonces, para respetar y mantener dicha unión entre ambas tablas, a ***WHERE,*** hay que pasarle la llave foranéa de la tabla ***posts****;* no cualquiera, sino, la que se vincula con la tabla ***usuarios*** justamente***.*** Ahora, uno habla de ***posts*** porque el filtro se aplica es sobre dicha tabla, no sobre otra,... se pretende buscar sus campos vacios o ***NULL***; y, principalmente, por lo que ya dijimos, porque es ***posts.usuario\_id*** la llave que relaciona a ambas tablas al fin; pues, ya están unidas y se debe respetar dicha unión,... realmente cualquier ***NULL*** que encuentre, entre una tabla u otra, sobre un ***registro***, lo arrojará (mostrará). Como en este caso la prioridad la tiene la tabla ***usuarios,*** se mostrarán todos los datos de los ***usuarios*** registrados; por lo que, entonces, se espera ver los ***NULL*** en los ***registros*** de la tabla ***posts.*** La declaración de la sentencia ***WHERE*** quedaría así:



y ahora todo junto se vería así:





***En el caso contrario, Right Join...***

La ***tabla B*** tendría la prioridad de la unión resultante con la ***tabla A***; es decir, ahora, ***los datos de la tabla B se verían primero que los de la tabla A; y, de hecho, ahora se verían en su totalidad los datos de la tabla B; esto, en otras palabras, significa que se verían todos los posts,... independientemente de si haya registro o no del usuario que lo o los escribió.***

Como usted ya sabe, la 1ra tabla pasada siempre será la ***tabla A*** o tabla de la ***izquierda***; es decir, eso no va a variar ahora. Lo único que cambiará ahora es la declaración de la sentencia ***LEFT JOIN*** por la de ***RIGHT JOIN,...*** para darle la prioridad a la ***tabla B,*** a la tabla de la ***derecha,*** en este tipo de ***Join (y ya sabe lo que significa eso)***





Por lo general las tablas se unen para hacer *filtros* entre ellas según la relación, de cardinalidad, que ellas tengan. Recuerdese que la cardinalidad entre ***usuarios*** y ***tablas*** era de ***1:N*** y la ***relación*** era ***usuarios escriben posts.*** Base sus ***joins*** de tablas, y los posteriores *filtros* que haga para ambas, al estar unidas, según la relación que conserven (como lo hicimos, justamente, con ***usuarios y posts***).

***Poniendo en práctica Inner (Intersección): bajo este tipo de Join uniremos, ahora, “usuarios” con “posts”.***

***Inner***



La ***intersección*** tiene que ver con que, recordemos, los datos de cada registro de ambas tablas (a & b) se correspondan entre sí y, en consecuencia, no se encuentren celdas vacías al unirse las tablas en el ***query;*** es decir, acá al hacer la consulta, no se estaría llamando todas celdas con información de una única tabla A, por ejemplo, sin importar si cada una de sus celdas se encuentre correspondida, en términos de información (no estén vacias tampoco), con las celdas de la tabla B; o, por el contrario, no se estarían llamando tampoco todas las celdas con información de una única tabla B sin importar que sus celdas estén correspondidas, en términos de información, con las celdas de la tabla A, NO. Lo que estamos pidiendo en la consulta, ***query,*** es que veamos todas las celdas, tanto de la tabla A como de la tabla B, cuando ninguna de sus celdas estén vacias en cada uno de los ***registros.*** Entonces, esta es la condición: las celdas, en cada uno de los registros para ambas tablas, no pueden estar vacías para que el ***query*** las pueda mostrar; de lo contrario, el ***join inner*** no las mostrará. En contexto, ***join inner*** sólo nos mostrará los ***usuarios*** con necesariamente registro en ***posts*** y los ***posts*** con necesariamente registro en ***usuarios.***



Acá se respeta el mismo orden en el que son nombradas las tablas. La ***tabla A*** o ***de la izquierda,*** manteniendo el simil con el diagrama de Venn, será la primera tabla llamada; es decir, ***FROM*** ***usuarios.*** Luego, la tabla B o de la derecha será la segunda llamada.

La palabra ***INNER JOIN*** quiere decir: *“sólo traeme lo que está internamente ligado, traeme solamente los que tienen dependencia de ambos lados”* (donde no haya celdas vacías en tabla A ni en tabla B).



Este es uno de los ***Joins, Inner Joins,*** de los que más se van a utilizar y es, generalmente, el que más valor brinda; pero, hay muchas veces que es necesario utilizar ***Left*** o ***Right*** para que no te importe uno de los lados y siempre traigas todos los datos de un lado de la ecuación.

***Poniendo en práctica Unión: bajo este tipo de Join uniremos, ahora, “usuarios” con “posts”.***

***Unión***



Este ***Join*** es muy claro. Como se dijo antes, cuando hablamos de *Unión,* estamos hablando de unir definitivamente ambas tablas, tanto ***A*** como ***B,*** sin importar si en los *registros* de cada una de las tablas hay celdas *vacías.*

Este tipo de ***join,*** en algunos manejadores de bases de datos, no tiene una forma propia de sentencia para ser ejecutado; por lo que el *estandar* es que sea ejecutado al *unir* las sentencias ***left join*** (que muestra la totalidad de los registros completados de *la tabla izquierda* o tabla A con el registro correspondiente en tabla B; sin importar si las celdas, de cada registro de la tabla B, estén vacías o no) ***& right join*** (que muestra la totalidad de los registros completados de *la tabla derecha* o tabla B con el registro correspondiente en tabla A; sin importar si las celdas, de cada registro de la tabla A, estén vacías o no); evidensiandose así, a lo sumo, por medio de dicha unión, la totalidad de las celdas de cada tabla, estén vacías o no. De hecho, si tú unes los diagrams de Venn del ***Left join*** con el de ***Right join*** te dará ***Unión.*** Es decir, en contexto, ***se verían todos los usuarios independientemente de si los usuarios hayan escrito o no posts; pero, también, se verían todos los posts independientemente de si haya registro o no del usuario que lo o los escribió.*** En código se vería así:





***Por último, pongamos en práctica Diferencia simétrica: bajo este tipo de Join uniremos, ahora, “usuarios” con “posts”.***

***Diferencia simétrica***



Sin extendernos mucho, básicamente consiste en: ***ver lo que existe en A pero no en B y ver lo que existe en B pero no en A* (es justamente todo lo contrario a *Inner join*).** Tal que así:





***Where***

***Where*** es ahora la sentencia que nos ayudará a filtrar, no columnas; sino, filas o ***registros.*** Esta vez nos va a permitir filtrar qué datos en especificos, según unas caracteristicas puntuales pasadas, queremos mostrar; no en el sentido generalizado de *qué campos visualizar en su conjunto*, no; más bien, qué registros puntuales ver según una condición de *filtro* pasada que deban cumplir. Vamos al grano con ejemplos.

En concreto, la sentencia ***Where*** lo que nos va a permitir hacer es filtrar y decir cuáles ***registros,*** de una o varias tablas combinadas con ***join,*** nos van a servir en nuestro reporte; esto según un parámetro o condición especificada.

**Ejemplo 1.** Queremos ver los ***registros*** de la tabla ***posts*** donde el campo ***id sea igual o menor*** que ***50,*** tal así:





*Tenga en cuenta que los ID para los posts comienzan desde el 43*

**Ejemplo 2.** Queremos ver los ***registros*** de la tabla ***posts*** donde el campo ***estatus No sea igual*** a ***“activo”,*** tal así:





Ahora, hay un ejemplo interesante con los tipos de dato ***string*** (cadenas de texto) que vamos a conocer. El tipo de dato ***string*** dispone de una propiedad llamada ***LIKE,*** que es muy utilizada en ***Queries,*** en los casos en los que se desee visualizar cadenas de texto, dentro de todos los ***registros*** para un ***campo,*** que lleven consigo alguna palabra en especifica (que las cadenas de texto tengan); es decir, ***like***hace la busqueda de todas las cadenas de texto que tengan consigo una palabra en concreto; dicha ***palabra*** es pasada por medio de la propiedad ***like***, luego se abre comillas *“”*,y se asiste de simbolos de porcentajes antes y después de la escritura, dentro de las propias comillas ***(“% ... %”)***, de la palabra en cuestión. La *palabra* vendría siendo realmente el filtro para poder visualizar o no las cadenas de texto; en concreto, con la palabra pasada es que se buscan las cadenas de texto en sí. Conclusión: la cadena de texto que tenga dicha palabra, pasada con ***LIKE,*** será arrojada como resultado dentro del ***query.*** Veamos esto mejor con un ejemplo.

**Ejemplo 3.** Queremos ver los registros de la tabla ***posts*** donde el campo ***titulo,*** dentro de sus cadenas de texto, tenga la palabra ***“escandalo”***(o una extremadamente parecida a ella), tal así:





Ahora, si por ejemplo la palabra ***“escandalo”*** queremos que sea necesariamente la primera escrita, dentro de todas las cadenas de textos del campo ***titulo*** de la tabla ***posts,*** y queremos visualizar cuáles ***registros*** cumplen con dicha condición... ***El primer simbolo de porcentaje (%) se omite,*** lo demás se deja igual**;** tal que así:





Si quisieramos que la palabra ***“escandalo”*** esté al final... ***Es el segundo simbolo de “%” el que se omitiría.***



Como ***ninguna cadena de texto*** del campo ***titulo*** (de la tabla ***posts***) finaliza con la palabra ***“escandalo”,*** entonces, no se logra visualizar ***ningún registro*** en el ***Query.***



**Ejemplo 4.** Queremos ver los ***registros*** de la tabla ***posts*** donde el campo ***fecha\_publicacion*** sean ***superior*** a la fecha ***“2025-01-01”,*** tal así:





**Una acotación:** *Los tipos de datos “fecha”* (date, datetime, time, timestamp, year) se escriben entre comillas.

Ahora, si quisieramos ver los ***registros*** de la tabla ***posts*** que respeten un intervalo de tiempo; es decir, que estén entre una fecha y otra, nos debemos asistir de la propiedad ***Between*** dentro de la sentencia ***Where;*** luego de declarar la propiedad ***Between,*** se pasan las fechas en cuestión.

Puntualmente, **ejemplo 5,** queremos ver todos los ***posts*** que hayan sido escritos entre el ***1 de enero del 2023 y el 31 de diciembre del 2025***, tal que así:





La propiedad ***Between*** no sólo funciona sobre tipos de datos ***Data;***sino, también, sobre ***cualquier tipo de dato númerico***; por ejemplo, veamos ahora los ***registros*** de la tabla ***posts*** que vayan desde el ***id 50*** hasta el ***id 60***, tal así:





**Formas más exoticas de proyectar los datos:** Con ***Where,*** en ocasiones, no hace falta ser siempre genericos a la hora de hacer una ***consulta*** sobre un ***campo*** de cualquier ***tabla***; por ejemplo, en el caso del ***atributo*** o campo ***fecha\_publicacion***, no hace falta pasar toda la *fecha* en concreto para que el ***query*** haga la consulta sobre dicha columna; podríamos, incluso, basar (y hacer más detallada) nuestra intención de busqueda según un parámetro propio, más especifico, extraído del campo en cuestión; en este caso, de ***fecha\_publicacion.* Veamos varios ejemplos.**

  
Acá se especificó algo más puntual que toda la fecha de publicación, el año ***(year);*** por lo tanto, al ***Query*** sólo le interesa que se cumpla la condición de que los ***posts*** hayan sido publicados ***entre*** los años ***2023*** & ***2024*** (no le interesa ni en qué mes ni que día); del resto le importa poco; incluso, repito, no le importa los meses ni los días; sólo que se cumpla la condición de que hayan sido escritos, los posts, entre el año 2023 y el año 2024.

***Un ejemplo más exótico todavía sería el siguiente:***

***,.. lo saca?* Acá sólo importa que el mes sea “04” (sin importar año ni fecha).**

***Where para Nulo o no Nulos***

Hasta ahora no hemos dejado muy claro cómo usar la sentencia ***Where*** para hacer consultas en aquellas celdas o valores de tipo ***null (or not null)***. Cuando nos referimos a celdas tipo ***null,*** o también ***default,*** hablamos de celdas *vacías,...* celdas que no han recibido, o pasado, ningún tipo de valor (no han sido definidas). Este tipo de celdas ***null*** reciben un tratamiento particular, digamos, para hacer consultas de ellas; se asisten de la sentencia: ***NULL (IS NULL).*** Vamos directo al *grano* con un ejemplo.

***Ejemplo 1.*** Queremos consultar, dentro de la tabla ***posts,*** aquellos ***registros*** donde los ***posts*** no tengan un ***usuario asociado.***

Recuerdese, como estamos viendo sólo la información de una tabla, la tabla ***posts;*** entonces, para saber puntualmente qué ***columna,*** ***campo*** o ***atributo*** de la tabla ***posts*** tiene la información de los ***usuarios,*** teniendo claro que ***usuarios*** vendría siendo realmente otra entidad o tabla,... A qué recurso, de la misma tabla ***posts,*** tendría que recurrir para ver los ***usuarios*** ***asociados*** a ***los*** ***posts? Pues bien, a la llave foránea de usuarios dentro de la tabla posts;*** es decir, al campo o columna: ***“usuario\_id”.*** Lo dicho, la tarea consiste en consultar qué registros, de la tabla ***posts,*** tienen un ***usuario\_id*** equivalente a ***NULL***;es decir, donde ***usuarios\_id sea null.*** Tal así:





*La base de datos entiende que le estás preguntando donde el usuario es “nulo”; es decir, trae todos los registros donde el usuario es nulo (is null).*

Sin embargo, hay que precisar que la mayoria de las veces no nos interesa traer los ***nulos***; sino, mas bien, los ***no nulos***; siguiendo el mismo ejemplo, estaríamos hablando entonces de los posts que sí tienen un usuario asociado. Cuando queremos consultar registros, de cualquier campo, que ***NO*** tengan celdas ***null*** nos debemos asistir de la sentencia ***IS NOT NULL.*** Tal así:





***AND***

Ahora, se pueden combinar varias condiciones en***WHERE*** asistiendose de ***AND;*** es decir, se pueden unir varios criterios a nuestra busqueda o ***query.***

Por ejemplo, qué tal si además de ver todos los posts asociados a un usuario se cumpla también la condición de que el post esté activo; es decir, un ***registro*** para ser mostrado por el ***query*** en mención debe cumplir con ambas condiciones: *que el* *post tenga un usuario asociado y que, dicho post, se encuentre también activo,* tal así:





Entonces, aquí se están filtrando los registros por la validez, no sólo de una condición; sino, de dos condiciones (pueden ser más incluso): que el post necesariamente tenga un ***usuario asociado*** y que, además, esté ***“activo”.***

Ahora un ***query*** más exigente (donde se cumplan, ya no dos; sino, tres condiciones al tiempo). A las dos criterios de busqueda pasados, anteriormente, agreguemos que: el ***ID sea menor que 50.*** Tal así:





Agreguemos otra condición más, que la ***categoria\_id sea 2.*** Tal así:





Con ***Where,*** si lo nota, podemos filtrar tanto como queramos!

Si, por ejemplo, además de las otras condiciones,... le agregamos una última que nos arroje en el ***query*** solo los ***posts*** que hayan sido publicados en el año ***2025,*** sería así:





Nos arroja sólo un ***registro*** porque, precisamente, es ese el único registro que cumple, en su conjunto, con todas las condiciones pasadas.

***Group by***

Esta nueva sentencia, si bien no es tan común, también es muy usada. Tiene que ver, como ya se imagina, con ***agrupar.*** La sentencia ***Group*** ***by*** lo que hará es indicarle a la base de datos qué criterios o valores (datos) nos va a arrojar, de un campo o columna seleccionada, para agrupar cierto número de registros en ellos; esto a partir de seleccionar desde qué campo o columna se verán los registros, ahí agrupados, según un criterio pasado en ***Select*** (que, entre otras cosas, suele coincidir con la columna o campo seleccionado en ***group by***). Por lo general, ***Group by*** trabaja de la mano de la sentencia ***COUNT();*** pues, muchas veces la intención de *agrupar* consiste en saber una totalidad numérica de registros (los cuenta) para cierta columna pasada en ***Select***, según *x* o *y* parámetro, criterio. El criterio o parametro suelen ser los distintos valores o datos con los que ha sido definido el campo o columna seleccionada con ***Select;*** es decir, **se me hará un conteo de los registros que hay para un determinado campo o columna seleccionada, sí; pero, haciendo el conteo de registros, por separado, para cada uno de los posibles valores que asume el campo o la columna en cuestión.**

Lo siguiente: cuando se trabaja con ***COUNT(),*** lo que se crea es una nueva columna, con ese mismo nombre ***“COUNT()”;*** que, justamente, nos arrojará el resultado de un conteo; en este caso, el conteo de registros para un campo o columna seleccionada con ***Select***. Ahora, cuando el conteo es sobre el **total** de los registros, de dicho campo seleccionado, hablamos de ***COUNT(\*);*** el criterio sería: ***selecciona y cuenta todos los registros del campo en cuestión.***

Ahora, si quizás a usted no le interesa que el campo creado con ***COUNT(\*)*** se llame “***COUNT(\*)”***; entonces, usted podría sustituir el nombre del campo, antes de crearlo, asistiendose de la palabra reservada ***AS*** y pasandole un nuevo nombre identificativo al campo “***COUNT(\*)”*** en cuestión***.*** Veamos todo esto, en su conjunto, con un ejemplo.

**Ejemplo 1.** Queremos ***contar,*** con count(\*), la totalidad de registros (posts publicados) que hay para cada uno de los valores, ***active*** e ***inactive,*** de la columna seleccionada ***estatus;*** todo esto por medio de ***agrupar,*** con ***group by***, dicha información en la misma columna existente ***estatus.***

Cambie el nombre de la columna ***“count(\*)”***, posiblemente resultante, por el nombre ***“post\_quantity”.*** Tal así:





Hemos contado la cantidad de registros que hay en cada uno de los valores que asume la columna **estatus.**

**Veamos otros ejemplos. Ejemplo 2.** Queremos ***contar,*** con count(\*), la totalidad de posts publicados que hay para cada año por separado; los cuales, fueron registrados en la totalidad de los datos o valores que tiene la columna ***fecha\_publicacion***; es decir, toca ***seleccionar*** dicha columns *(fecha\_publicacion)* para evaluar cuántos posts fueron publicados por cada año en cuestión; en consecuencia, adicionalmente, toca ***agrupar*** con ***group*** ***by***, la información de la misma columna ***fecha\_publicación***. Ésta última, para efectos prácticos, la identificaremos ahora con un nuevo campo dinámico llamado ***post\_year*** y ahí, en ***post\_year***,sólo nos interesa ocupar la información relacionada a los años, no las fechas completas de la columna fecha\_publicacion: esto para decir puntualmente la cantidad de posts por año, OJO, no la cantidad de posts publicados en cada una de las fechas registradas en fecha\_publicacion; si ese fuera el caso, no funcionaría nuestro filtro por años (YEAR); por lo cual, entonces, nos tocaría expresar la información de fecha\_publicacion, que está con un tipo de dato de ***TIMESTAMP***, en términos del tipo de dato ***YEAR.*** Veamos:





Agrupemos ahora, la cantidad de posts publicados, pero en términos de los 12 meses del año ***(MONTHNAME);*** es decir, la cantidad de posts por mes sin importar el año. Toca entonces ahora expresar la información de fecha\_publicacion, que está con un tipo de dato ***YEAR***, en términos del tipo de dato ***MONTHNAME.*** Tal así:





***Varios criterios en Group by***

Usted recuerda lo anterior: *“La sentencia* ***Group******by*** *lo que hará es indicarle a la base de datos qué criterios o valores (datos) nos va a arrojar, de un campo o columna seleccionada, para agrupar cierto número de registros en ellos; esto a partir de seleccionar desde qué campo o columna se verán los registros, ahí agrupados, según un criterio pasado en* ***Select*** *(que, entre otras cosas, suele coincidir con la columna o campo seleccionado en* ***group by****)*”... Pues bien, *pueden ser más de un criterio los pasados* tanto para el ***Select*** como para el ***Group By;*** es decir, para que se organice o se agrupe la información según los datos o valores de más de una columna; como estamos trabajando con count(\*),... para que se cuenten los registros; más precisamente, **haciendo el conteo de los registros por cada valor, separadamente, de cada una de las columnas seleccionadas**; en este caso, el conteo de la cantidad de posts publicados para cada uno de los valores que contemplan los campos o columnas pasadas como criterios en el ***Select*** y en el ***Group By***.

Entonces, cada criterio pasado en el ***Select*** y en el ***Group*** ***By,*** cada nueva columna, es separada por comas en cada una de las líneas de código del ***Select*** y del ***Group By***. Esta vez los criterios pasados, para hacer la agrupación, serán las columnas del ***estatus*** y la nueva columna dinámica ***post\_month***, tal que así:





Si se da cuenta, me hizo el conteo de la totalidad del posts publicados (registros), no sólo para cada mes por separado; sino, para cada mes por separado y también para cada *estatus* dentro de cada uno de esos meses, ya sea *activo* o *inactivo;* es decir, el filtro ahora está más elaborado.

***Order by***

***Order by*** es una sentencia que tiene que ver con el ordenamiento de los datos; ***order by*** quiere decir, justamente, ***ordenar por***; es decir, le vas a decir a la base de datos *que te ordene los datos por x o y criterio y/o campo*. Es decir, con ***Order by*** no se hacen filtros; no, sino que se pretende mostrar una misma cantidad de información según un tipo de orden que le pasemos. Veamos esto con un ejemplo.

***Ejemplo 1.*** Si quisiera ver todos los registros de la tabla ***posts*** (sin hacer ningún tipo de filtro en ella, aclaremos); de tal manera que, se me ordene u organice la lista de todos los posts publicados según su fecha de publicación, podríamos hacerlo con ***Order by;*** de hecho, podríamos establecer si queremos que el ordenamiento se haga desde la fecha de publicación más reciente a la más antigua (con la propiedad ***DESC***) o desde la fecha de publicación más antigua a la más reciente (con la propiedad ***ASC***); sin embargo, si no se le pasa ninguna de las dos propiedades, el ordenamiento por defecto se hará desde la fecha más antigua a la más reciente; es decir, como si fuera ***ASC. Probemos ambas propiedades de Order By: ASC & DESC.***

**Ordenamiento de registros por tipo de dato “*data”*** o fecha; desde la fecha más antigua a la más reciente o, bien, de forma ascendente: propiedad ***ASC***.





*Ahora,* **Ordenamiento de registros por tipo de dato “*data”*** o fecha; desde la fecha más reciente a la más antigua o, bien, de forma descendente: propiedad ***DESC***.





*Ahora,***nos interesaría probar un order by por tipo de dato *“string” o*** cadena de texto; acá el ordenamiento también se asiste de las palabras reservadas: ***ASC & DESC.*** Trabaja de la siguiente manera: El order by, en cadenas de texto, ordena u organiza los registros de forma alfabetica; de tal manera que, la propiedad ASC organiza los registros desde la A a la Z; y, por el contrario, la propiedad DESC organiza los registros desde la Z a la A. En ambas se parte o se toma como referencia, como criterio de ordemiento, la inicial de cada registro. Tal así:

**Ordenamiento de registros por tipo de dato “*string”*** o cadenas de texto; desde los registros que empiezan con la vocal A hasta los que empiezan con la letra Z o, bien, de forma ascendente: propiedad ***ASC***. Veamos esto en la columna o atributo ***titulo*** (que me ordene de forma *ascedente* los registros de los posts según sus títulos).





**Ahora, revisemos el caso contrario,** ordenemos los registros desde las cadenas de texto que empiecen con la letra Z hasta las que empiecen con la letra A o, bien, de forma descendente: propiedad ***DESC***. Veamos esto en la misma columna o atributo ***titulo*** (que me ordene de forma *descedente* los registros de los posts según sus títulos).





*También* existe el **Ordenamiento de registros por tipo de dato “*number”*** o numérico. Vamos al grano salvajemente: usted decidirá si quiere organizar los registros de una columna, que es de tipo numérica, de menor a mayor (con la propiedad ***ASC***) o de mayor a menor (con la propiedad ***DESC***). Probemos esto con el atributo o columna ***usuario\_id.***

*Ordenamiento de registros, de menor a mayor número. según la columna* ***usuario\_id****.*





*Ordenamiento de registros, de mayor a menor número. según la columna* ***usuario\_id****.*





Ahora, hay una sentencia de complemento a la sentencia ***Order by*** que se llama ***Limit;*** en ***Limit*** se determina la cantidad de resultados que queremos que arroje el manejador de bases de datos; esto, claro, luego de determinar en qué orden, y con qué columna de referencia, serán presentados los datos al asistirse con ***Order by.*** Veamos.

Por ejemplo, si queremos ver ***los primeros 5 posts que se publicaron*** (es decir, los primeros 5 registros)tomando como referencia la columna de ***fecha\_publicacion***, sería así:





Como estamos hablando de ***las 5 primeras fechas (LIMIT 5)***, entonces, tendríamos que asistirnos también de la sentencia auxiliar ***ASC;*** pues, los registros los queremos ver de forma ascendente según, o tomando como referencia, la columna ***fecha\_publicacion;*** es decir, que vayan desde la fecha más antigua ***(los primeros)*** a la fecha más reciente.

Con la herramienta ***Limit*** podemos hacer ***Top10,*** o ***Top100,*** de cualquier cosa que usted se imagine y sea medible; es decir, de cualquier tipo de dato en últimas: ya sea fecha, número o cadena de texto.

***Having***

***Having*** tiene una gran similitud con ***Where;*** pero, justamente, de lo que se trata acá es de ver cuáles son sus posibles diferencias. Usted se recuerda de esto:

***... Where*** es ahora la sentencia que nos ayudará a filtrar, no columnas; sino, filas o ***registros.*** Esta vez nos va a permitir filtrar qué datos en especificos, según unas caracteristicas puntuales pasadas, queremos mostrar; no en el sentido generalizado de *qué campos visualizar en su conjunto*, no; más bien, qué registros puntuales ver según una condición de *filtro* pasada que deban cumplir. Vamos al grano con ejemplos.

En concreto, la sentencia ***Where*** lo que nos va a permitir hacer es filtrar y decir cuáles ***registros,*** de una o varias tablas combinadas con ***join,*** nos van a servir en nuestro reporte; esto según un parámetro o condición especificada.

**Ejemplo 1.** Queremos ver los ***registros*** de la tabla ***posts*** donde el campo ***id sea igual o menor*** que ***50,*** tal así:





*Tenga en cuenta que los ID para los posts comienzan desde el 43*

... Pues bien, la sentencia ***Having*** hace exactamente lo mismo; sólo que, en vez de ***Where,*** la sentencia ***Having*** es la que se usa en los casos puntuales donde se pretende hacer filtrossobre **registros agrupados y/o de columnas dinámicas.**

Es decir, ***Where*** nos ayuda hacer filtros sobre los registros, sí; pero los filtra antes, o sin la interferencia, de la sentencia ***group by***; dicho eso, ***Where*** no puede trabajar de la mano con ***Group*** ***by,*** porque where *desconoce* dicha sentencia y, en consecuencia, no puede hacer una seleccion o filtro de filas o registros agrupados ni de columnas dinámicas; en cambio, ***Having*** sí. La sentencia ***Having*** se coloca justo después de la sentencia ***Group by.***

***Veamos el siguiente ejemplo de uso de Having:*** Se pretende ***agrupar,*** y ***contar,*** la cantidad de posts publicados, según cada uno de los valores que asumen las columnas de referencia ***fecha\_publicacion*** & ***estatus,*** con la condición de que se muestren solamente los registros donde se publicó ***más de un posts*** (para cada uno de los valores de las columnas ***fecha\_publicacion*** & ***estatus*** donde se coincida, conjuntamente y en un mismo ***registro,*** con más de un post publicado). Además, se desea tomar como referencia la columna ***fecha\_publicacion*** para ordenar los datos seleccionados. Tal que así:





*La idea es que primero el manejador de datos agrupe y cuente, posteriormente que filtre (sobre los datos contados); y, luego, que se ordenen dichos datos según unos criterios deseados.*