



DIPLOMADO “CIENCIA DE DATOS” Módulo 3 ANÁLISIS Y MANIPULACIÓN DE BASE DE DATOS

1. Introducción
2. Exploración de datos
3. Restricciones y ordenamiento de datos
4. Funciones
5. **Agrupamiento de datos**

¿Qué es una función de grupo?

Una función de grupo en SQL es una categoría especial de funciones que se utilizan con la cláusula GROUP BY en una consulta SELECT. Estas funciones realizan cálculos sobre grupos de filas relacionadas en lugar de operar en filas individuales. Los grupos se forman agrupando los registros que comparten el mismo valor o conjunto de valores en una o varias columnas específicas.

Las funciones de grupo se utilizan comúnmente en combinación con la cláusula GROUP BY para resumir y obtener resultados agregados de una base de datos. Estas funciones toman como entrada un conjunto de valores dentro de cada grupo y devuelven un solo valor para cada grupo.

5.1 Funciones de grupo

Algunas de las funciones de grupo más utilizadas en SQL son:

- **SUM:** Calcula la suma de los valores de una columna numérica.
- **AVG:** Calcula el promedio de los valores de una columna numérica.
- **COUNT:** Cuenta el número de registros en una tabla o el número de valores no nulos en una columna.
- **MIN:** Encuentra el valor mínimo en una columna numérica o de texto.
- **MAX:** Encuentra el valor máximo en una columna numérica o de texto.

Funciones que agrupan todos los datos de la tabla.

Ejemplo 1.

Esta consulta muestra el salario máximo, el salario mínimo y la suma de todos los salarios de todos los profesores

```
SELECT MAX(salario), MIN(salario), SUM(salario)
FROM profesores;
```

Función COUNT

La función COUNT en SQL se utiliza para contar el número de registros o filas en una tabla o el número de valores no nulos en una columna específica. Es una función de





agregación que puede aplicarse en combinación con la cláusula GROUP BY para obtener el recuento de registros en grupos específicos o sin la cláusula GROUP BY para obtener el recuento total en toda la tabla.

Ejemplo 1.

Esta consulta muestra la cantidad de alumnos

```
SELECT COUNT() AS cantidad
FROM alumnos;
```

Función AVG

La función AVG en SQL es una función de agregación que se utiliza para calcular el promedio de los valores numéricos en una columna específica. Es decir, devuelve el valor promedio de los datos numéricos que se encuentran en la columna seleccionada en una tabla.

La sintaxis básica de la función AVG() es la siguiente:

```
SELECT AVG(columna)
FROM tabla;
```

Ejemplo 1.

Esta consulta muestra el promedio de los salarios de los profesores

```
SELECT AVG(salario) AS promedio
FROM profesores;
```

Ejemplo 2.

Mostrar el precio promedio de los productos vendidos

```
SELECT AVG(precio_unitario) AS precio_promedio
FROM ventas;
```

Clausula GROUP BY

La cláusula GROUP BY en SQL se utiliza junto con funciones de agregación para agrupar los resultados de una consulta según uno o varios valores de una o varias columnas. Con GROUP BY, puedes dividir los datos en grupos basados en los valores comunes en una o varias columnas y luego aplicar funciones de agregación como SUM(), AVG(), COUNT(), MAX(), MIN(), entre otras, para obtener resultados resumidos dentro de cada grupo.





```
SELECT  función (columna), ...  
FROM    tabla  
[WHERE  condición]  
[GROUP BY columna];
```

Ejemplo 1.

Esta consulta despliega el promedio de los salarios de los profesores por departamento

```
SELECT departamento, AVG(salario) AS promedio  
FROM  profesores  
GROUP BY departamento;  
FROM  profesores;
```

departamento	promedio
Mecanica	3100
Sistemas	14500
Logistica	12800

Ejemplo 2.

Esta consulta despliega la cantidad de empleados en cada departamento

```
SELECT departamento, COUNT(*) AS total_empleados  
FROM empleados  
GROUP BY departamento;
```

departamento	total_empleados
Mecanica	34
Sistemas	26
Logistica	52





5.2 Restringir agrupamientos

Clausula HAVING

La cláusula HAVING en SQL se utiliza junto con la cláusula GROUP BY para filtrar los resultados de una consulta basándose en los resultados de funciones de agregación. Mientras que la cláusula WHERE se utiliza para filtrar filas individuales antes de aplicar funciones de agregación, la cláusula HAVING se utiliza para filtrar grupos de resultados después de aplicar funciones de agregación.

```
SELECT  función (columna), ...  
FROM    tabla  
[WHERE  condición]  
[[GROUP BY columna]  
[HAVING condición]];
```

Ejemplo 1.

Si queremos obtener el total de ventas y solo mostrar los resultados para los productos que han tenido más de 5 unidades vendidas, podemos usar HAVING con SUM:

```
SELECT producto, SUM(cantidad) AS total_ventas  
FROM    ventas  
GROUP BY producto  
HAVING SUM(cantidad) > 5;
```

