

TEMARIO DEL CURSO DE SQL PARA ANÁLISIS DE DATOS

POR: Laureano Rojas

Aquí tienes un listado completo de todo lo que debes saber de SQL para análisis de datos:

1. Fundamentos de SQL

- **Sintaxis básica:** SELECT, FROM, WHERE, ORDER BY.
- **Operadores básicos:** =, <>, >, <, >=, <=.
- **Funciones de agregación:** COUNT, SUM, AVG, MIN, MAX.
- **Cláusulas básicas:** LIMIT, OFFSET.

2. Filtrado y agrupación de datos

- **WHERE:** Filtrado básico.
- **HAVING:** Filtrar después de aplicar funciones de agregación.
- **GROUP BY:** Agrupar datos por una o más columnas.
- **DISTINCT:** Eliminar duplicados.

3. Joins y combinaciones de tablas

- **INNER JOIN:** Filtra solo las coincidencias.
- **LEFT JOIN:** Incluye todos los registros de la izquierda.
- **RIGHT JOIN:** Incluye todos los registros de la derecha.
- **FULL OUTER JOIN:** Incluye todos los registros de ambas tablas.
- **CROSS JOIN:** Multiplicación cartesiana.
- **SELF JOIN:** Unir una tabla consigo misma.

4. Subconsultas

- **Subconsultas en SELECT:** Consultas dentro de una consulta.
- **Subconsultas en WHERE:** Usar resultados de una consulta como condición.
- **Subconsultas en FROM:** Tratar una subconsulta como una tabla temporal.

5. Funciones avanzadas

- **Funciones de ventana:** ROW_NUMBER(), RANK(), DENSE_RANK().
- **Funciones de agregación de ventana:** SUM() OVER(), AVG() OVER().
- **Funciones de cadena:** CONCAT(), LENGTH(), UPPER(), LOWER().
- **Funciones de fecha y hora:** NOW(), DATE_ADD(), DATE_SUB().
- **CAST() y CONVERT():** Cambiar tipos de datos.

6. Operadores avanzados

- **IN:** Coincidencia con un conjunto de valores.
- **BETWEEN:** Filtrar un rango de valores.
- **LIKE:** Búsqueda de patrones.
- **IS NULL:** Filtrar valores nulos.
- **EXISTS:** Verificar la existencia de registros.

7. Manejo de datos

- **INSERT INTO:** Insertar registros en una tabla.
- **UPDATE:** Actualizar registros existentes.
- **DELETE:** Eliminar registros.
- **TRUNCATE:** Eliminar todos los registros de una tabla sin eliminar la estructura.

8. Índices y optimización

- **Índices:** Mejorar el rendimiento de las consultas.
- **EXPLAIN:** Analizar y entender el plan de ejecución de una consulta.
- **JOINS eficientes:** Selección de índices adecuados.
- **Normalización y desnormalización:** Diseño de bases de datos eficientes.

9. Vistas y procedimientos almacenados

- **Vistas:** Crear y gestionar vistas como tablas virtuales.
- **Procedimientos almacenados:** Scripts que contienen lógica SQL reutilizable.
- **Funciones almacenadas:** Crear funciones personalizadas.
- **Triggers:** Automatizar acciones en la base de datos.

10. Transacciones y control de concurrencia

- **BEGIN TRANSACTION, COMMIT, ROLLBACK:** Control de transacciones.
- **ACID:** Propiedades de las transacciones (Atomicidad, Consistencia, Aislamiento, Durabilidad).
- **LOCKS:** Control de acceso concurrente a los datos.

11. Modelos de bases de datos

- **Modelo relacional:** Tablas, claves primarias, claves foráneas.
- **Modelo dimensional:** Estrella, copo de nieve, esquemas de dimensiones y hechos.

12. Análisis de datos en SQL

- **Pivot Tables:** Resumir datos en formato de tabla dinámica.

- **Análisis temporal:** Trabajar con series de tiempo.
- **Estadísticas:** Calcular medidas estadísticas (media, mediana, desviación estándar).
- **Segmentación de datos:** Filtrar y agrupar por categorías.

13. Prácticas avanzadas

- **Ctes (Common Table Expressions):** Consultas temporales.
- **UNION y UNION ALL:** Combinar resultados de múltiples consultas.
- **Funciones analíticas:** LEAD(), LAG(), FIRST_VALUE(), LAST_VALUE().
- **JSON en SQL:** Consultas sobre datos JSON.
- **Full-Text Search:** Búsquedas complejas en texto.

14. Trabajo con Data Warehouses

- **ETL (Extract, Transform, Load):** Extracción, transformación y carga de datos.
- **OLAP:** Procesamiento analítico en línea.
- **Data Marts:** Subconjuntos de datos dentro de un Data Warehouse.
- **Conformidad de dimensiones:** Dimensiones consistentes en todo el sistema.

15. SQL para Big Data

- **SQL en Hadoop (Hive, Impala):** Consultas en sistemas distribuidos.
- **SQL en NoSQL (Cassandra, MongoDB):** Consultas en bases de datos NoSQL.

Este es un panorama completo de SQL en el análisis de datos. Si necesitas más detalles sobre algún punto, ¡avísame!