

Evidencia de IA

Julián Andrés Guisao Fernández

Escuela de ingenierías, Universidad Pontificia Bolivariana

Tópicos Avanzados en Bases de datos

Juan Rodas

3 de abril de 2025

estoy analizando una bd, debo realizar consultas para verificar la integridad de los datos, te voy a enviar el esquema de la bd y las consultas en lenguaje natural para que me ayudes a estructurarlas en lenguaje pgsq.

estructura de la bd:

```
"Table "corregido"."regiones" {  
  "id" SERIAL [pk, increment]  
  "nombre" VARCHAR(20) [unique, not null]  
}
```

```
Table "corregido"."fuentes_hidricas" {  
  "id" SERIAL [pk, increment]  
  "id_region" INT  
  "nombre" VARCHAR(4) [not null]  
}
```

```
Table "corregido"."embalses" {  
  "id" SERIAL [pk, increment]  
  "nombre" VARCHAR(4) [not null]  
  "id_fuente" INT  
}
```

```
Table "corregido"."aportes_hidricos" {  
  "id" SERIAL [pk, increment]  
  "fecha" DATE  
  "id_embalse" INT  
  "aporte_hidrico" FLOAT  
}
```

```
Ref:"corregido"."regiones"."id" <  
"corregido"."fuentes_hidricas"."id_region"
```

```
Ref:"corregido"."fuentes_hidricas"."id" <  
"corregido"."embalses"."id_fuente"
```

```
Ref:"corregido"."embalses"."id" <  
"corregido"."aportes_hidricos"."id_embalse"
```

* EN LOS SIGUIENTES MENSAJES TE VOY A ENVIAR UNA NNUEVA CONSULTA

¡Entendido! Envíame las consultas en lenguaje natural y te ayudaré a estructurarlas en PostgreSQL (PL/pgSQL).

Asuma que el año 2023 tiene 365 días y el año 2024 tiene 366 días. Si el archivo registra el aporte hídrico diario de cada uno de los embalses, es posible que algunos de ellos tengan datos ausentes.

Se requiere construir una consulta que indique para cada año y cada embalse, cual fue el porcentaje de completitud obtenido en las mediciones del aporte. Un indicador del 100% representa que todos los días del año se encontraron datos para dicho embalse.

A esta consulta, genere el plan de ejecución donde se evidencie el árbol de las acciones que realizó la base de datos para resolver la petición.

El plan de ejecución y los datos resultantes se deben colocar en el archivo en Word referenciado en los entregables.

AGRUPA POR REGIONES Y EMBALSES, YA QUE EL MISMO EMBALSE SE ENCUENTRA EN DIFERENTES REGIONES

Aquí tienes la consulta en PostgreSQL para calcular el porcentaje de completitud de los registros de aporte hídrico por embalse y año, agrupando también por región:

Consulta SQL

```
sql Copy Edit

WITH dias_por_anio AS (
    SELECT 2023 AS anio, 365 AS total_dias UNION ALL
    SELECT 2024 AS anio, 366 AS total_dias
),
conteo_aportes AS (
    SELECT
        e.id AS id_embalse,
        fh.id_region,
        EXTRACT(YEAR FROM ah.fecha) AS anio,
        COUNT(ah.id) AS dias_con_aporte
    FROM corregido.aportes_hidricos ah
    JOIN corregido.embalses e ON ah.id_embalse = e.id
    JOIN corregido.fuentes_hidricas fh ON e.id_fuente = fh.id
    GROUP BY e.id, fh.id_region, anio
)
SELECT
    r.nombre AS region,
    e.nombre AS embalse,
    c.anio,
    c.dias_con_aporte,
    d.total_dias,
    ROUND((c.dias_con_aporte::NUMERIC / d.total_dias) * 100, 2) AS porcentaje_completitud
FROM conteo_aportes c
JOIN dias_por_anio d ON c.anio = d.anio
JOIN corregido.embalses e ON c.id_embalse = e.id
JOIN corregido.fuentes_hidricas fh ON e.id_fuente = fh.id
JOIN corregido.regiones r ON fh.id_region = r.id
ORDER BY r.nombre, e.nombre, c.anio;
```

Generación del Plan de Ejecución

Para obtener el plan de ejecución de la consulta en PostgreSQL, puedes usar el siguiente comando:

```
sql Copy Edit

EXPLAIN ANALYZE
WITH dias_por_anio AS (
    SELECT 2023 AS anio, 365 AS total_dias UNION ALL
    SELECT 2024 AS anio, 366 AS total_dias
),
conteo_aportes AS (
    SELECT
        e.id AS id_embalse,
        fh.id_region,
        EXTRACT(YEAR FROM ah.fecha) AS anio,
        COUNT(ah.id) AS dias_con_aporte
    FROM corregido.aportes_hidricos ah
    JOIN corregido.embalses e ON ah.id_embalse = e.id
    JOIN corregido.fuentes_hidricas fh ON e.id_fuente = fh.id
    GROUP BY e.id, fh.id_region, anio
)
SELECT
    r.nombre AS region,
    e.nombre AS embalse,
    c.anio,
    c.dias_con_aporte,
    d.total_dias,
    ROUND((c.dias_con_aporte::NUMERIC / d.total_dias) * 100, 2) AS porcentaje_completitud
FROM conteo_aportes c
JOIN dias_por_anio d ON c.anio = d.anio
JOIN corregido.embalses e ON c.id_embalse = e.id
JOIN corregido.fuentes_hidricas fh ON e.id_fuente = fh.id
JOIN corregido.regiones r ON fh.id_region = r.id
ORDER BY r.nombre, e.nombre, c.anio;
```

Este comando generará un plan de ejecución detallado en la consola de PostgreSQL, mostrando el árbol de ejecución y los costos de cada operación. Copia ese resultado y colócalo en el archivo de Word solicitado.

Si necesitas ajustar la consulta o agregar más detalles, dime y te ayudo.

En el año 2024 se evidenció el fenómeno atmosférico del niño, caracterizado por reducción de lluvias en las zonas y cuencas hidrográficas donde se encuentran los embalses del SNI. Se desea conocer cuanto fue el delta porcentual entre el valor máximo y valor mínimo de cada embalse.

Utilizando CTE y Windows Functions, construya una consulta que permita conocer para el año 2024, para cada embalse, cual fue su valor mínimo y máximo de aporte hídrico y cuanto fue el porcentaje de reducción asumiendo que un porcentaje de 0% es que el embalse no se redujo nada y un porcentaje del 100% es que el embalse se redujo en su totalidad.

A esta consulta, genere el plan de ejecución donde se evidencie el árbol de las acciones que realizó la base de datos para resolver la petición.

El plan de ejecución y los datos resultantes se deben colocar en el archivo en Word referenciado en los entregables.

Aquí tienes la consulta en PostgreSQL utilizando CTE y funciones de ventana para calcular el delta porcentual de reducción del aporte hídrico para cada embalse en el año 2024.

Consulta SQL

```
sql Copy Edit

WITH aportes_2024 AS (
    SELECT
        ah.id_embalse,
        ah.aporte_hidrico
    FROM corregido.aportes_hidricos ah
    WHERE EXTRACT(YEAR FROM ah.fecha) = 2024
),
valores_extremos AS (
    SELECT
        id_embalse,
        MIN(aporte_hidrico) AS valor_minimo,
        MAX(aporte_hidrico) AS valor_maximo
    FROM aportes_2024
    GROUP BY id_embalse
)
SELECT
    e.nombre AS embalse,
    v.valor_maximo,
    v.valor_minimo,
    CASE
        WHEN v.valor_maximo = 0 THEN NULL -- Para evitar división por cero
        ELSE ROUND(((v.valor_maximo - v.valor_minimo) / v.valor_maximo) * 100, 2)
    END AS porcentaje_reduccion
FROM valores_extremos v
JOIN corregido.embalses e ON v.id_embalse = e.id
ORDER BY porcentaje_reduccion DESC;
```

Explicación de la Consulta

1. CTE `aportes_2024`: Filtra los datos de la tabla `aportes_hidricos` para incluir solo el año 2024.
2. CTE `valores_extremos`: Calcula el mínimo y máximo aporte hídrico por embalse.
3. Consulta Final:

- Une los resultados con la tabla `embalses` para obtener los nombres.
- Calcula el delta porcentual con la fórmula:

$$\frac{\text{valor máximo} - \text{valor mínimo}}{\text{valor máximo}} \times 100$$

- Maneja la posible división por cero.

Generación del Plan de Ejecución

Para obtener el plan de ejecución, ejecuta el siguiente comando en PostgreSQL:

```
sql Copy Edit

EXPLAIN ANALYZE
WITH aportes_2024 AS (
  SELECT
    ah.id_embalse,
    ah.aporte_hidrico
  FROM corregido.aportes_hidricos ah
  WHERE EXTRACT(YEAR FROM ah.fecha) = 2024
),
valores_extremos AS (
  SELECT
    id_embalse,
    MIN(aporte_hidrico) AS valor_minimo,
    MAX(aporte_hidrico) AS valor_maximo
  FROM aportes_2024
  GROUP BY id_embalse
)
SELECT
  e.nombre AS embalse,
  v.valor_maximo,
  v.valor_minimo,
  CASE
    WHEN v.valor_maximo = 0 THEN NULL
    ELSE ROUND(((v.valor_maximo - v.valor_minimo) / v.valor_maximo) * 100, 2)
  END AS porcentaje_reduccion
FROM valores_extremos v
JOIN corregido.embalses e ON v.id_embalse = e.id
ORDER BY porcentaje_reduccion DESC;
```

Este comando generará un plan de ejecución que muestra el árbol de acciones realizadas por la base de datos. Copia el resultado y colócalo en el archivo de Word solicitado.