

## Tarea N° 1 – SQL OLAP

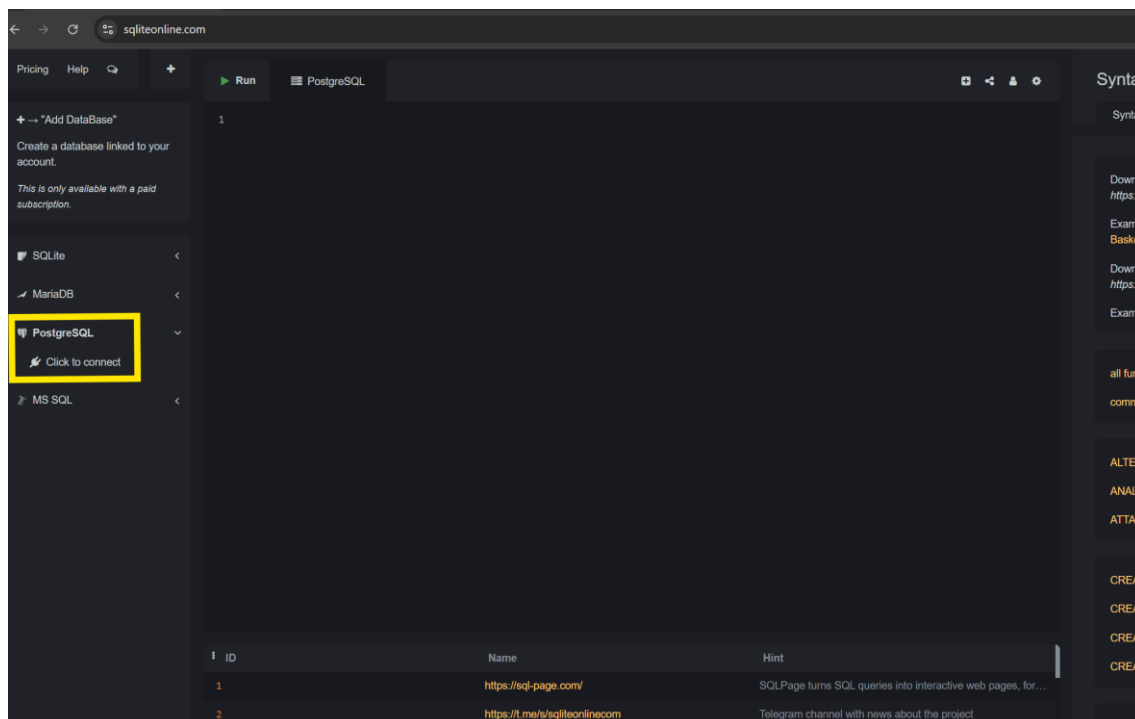
Se tiene la siguiente tabla con datos de viajes en transporte:

**VIAJES\_TRANSP\_ENE20** (dia, nombre\_empresa, línea, amba, tipo\_transporte, tipo\_jurisdiccion, provincia, municipio, cant\_viajes).

- dia: día de realización del viaje
- nombre\_empresa: nombre de la empresa donde se realizó el viaje
- línea: nombre de la línea donde se realizó el viaje
- amba: toma valores SI / NO, dependiendo si la línea es de amba o no
- tipo\_transporte: toma valores COLECTIVO, LANCHAS, SUBTE, TREN
- tipo\_jurisdiccion: indica si la línea es de jurisdicción Nacional, Provincial o Municipal.
- Provincia: nombre de la provincia de la línea (en caso de tratarse de jurisdicción provincial o municipal) o "JN" (en caso de tratarse de jurisdicción nacional)
- Municipio: nombre del municipio de la línea (en caso de tratarse de jurisdicción municipal) o "SD" (en caso de tratarse de jurisdicción nacional) o "SN" (en caso de tratarse de jurisdicción provincial)
- Cant\_viajes: cantidad de viajes realizados (en cada combinación de valores del resto de las variables)

✂ **IMPORTANTE:** los nombres de las líneas pueden repetirse en diferentes jurisdicciones. Para identificar unívocamente a una línea se necesita combinar Provincia, Municipio y Nombre de Línea.

- En el campus dejamos un script para la creación y población de esta tabla, de manera que puedan probar y verificar la resolución propuesta. Es para resolver sobre un motor PostgreSQL. En caso de que no lo hayan instalado en su equipo personal, pueden utilizar el sitio <https://sqliteonline.com/>, tiene una opción para ejecutar scripts PostgreSQL, no requiere tener usuario.



Se pide realizar las consultas SQL que devuelvan:

- a) Cantidad de viajes realizados agrupados por día, por día-tipo\_transporte, por día-tipo\_transporte-provincia.
- El resultado debe quedar en un solo dataset de respuesta, es decir, ejecutando una única consulta.
  - La salida esperada contiene los campos día, tipo\_transporte, provincia, total\_viajes. Los campos que no aplican para un nivel de agregación (por ej, tipo\_transporte y provincia no aplican para la agrupación sólo por día) pueden quedar en nulos.
  - Se lo quiere ordenado por día, dentro de día por tipo\_transporte, y dentro de día-tipo\_transporte ordenado por provincia, con una visibilidad jerárquica:
    - *Día 1*
    - *Día 1 - tipo\_transporte 1*
    - *Día 1 - tipo\_transporte 1 – provincia 1*
    - *Día 1 - tipo\_transporte 1 – provincia 2*
    - ...
    - *Día 1 - tipo\_transporte 1 – provincia n*
    - *Día 1 - tipo\_transporte 2*
    - *Día 1 - tipo\_transporte 2 – provincia 1*
    - *Día 1 - tipo\_transporte 2 – provincia 2*
    - ...
    - *Día 1 - tipo\_transporte 2 – provincia n*
    - ...
    - *Día 2*
    - *Día 2 - tipo\_transporte 1*
    - *Día 2 - tipo\_transporte 1 – provincia 1*
    - ...
    - ...
    - *Día m - tipo\_transporte k – provincia n*
- b) Cantidad de viajes realizados agrupados por día, por tipo de transporte, por línea (tres agrupaciones independientes, no se combinan entre sí).
- El resultado debe quedar en un solo dataset de respuesta, es decir, ejecutando una única consulta.
  - La salida esperada contiene los campos día, tipo\_transporte, línea, total\_viajes. Los campos que no aplican para un nivel de agregación (por ej, tipo\_transporte y línea no aplican para la agrupación sólo por día) pueden quedar en nulos.
  - Se lo quiere ordenado primero las agrupaciones por día, después las agrupaciones por tipo de transporte, y por ultimo las agrupaciones por línea.