

273-[DF]-Lab - Organización de datos

### **Datos Generales:**

Nombre: Tomás Alfredo Villaseca Constantinescu

País: Chile

Fecha: 08/10/2023

Contacto: tomas.villaseca.c@gmail.com

Este laboratorio demuestra cómo usar funciones de base de datos comunes con las cláusulas **GROUP BY** y **OVER**.

Después de completar este laboratorio, podrá realizar lo siguiente:

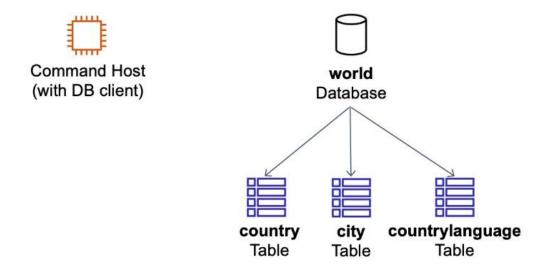
- Usar la cláusula GROUP BY con la función SUM()
- Usar la cláusula OVER con la función de ventana RANK()
- Usar la cláusula OVER con la función SUM() y la función de ventana RANK()

### Situación

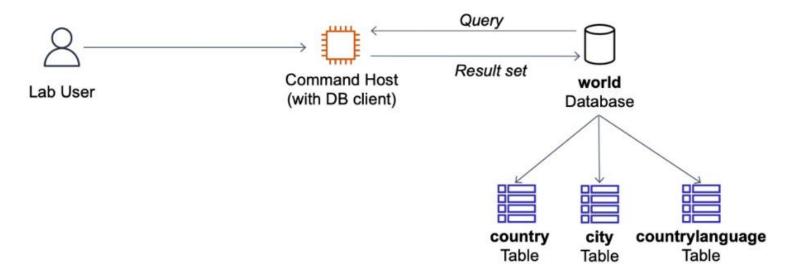
El equipo de operaciones de base de datos creó una base de datos relacional llamada **world** que contiene tres tablas: **city**, **country** y **countrylanguage**. Ayudará a escribir algunas consultas a los registros de grupo para su análisis usando ambas cláusulas **GROUP BY** y **OVER**.

### Entorno de laboratorio:

- Una instancia EC2 (Command Host) con un cliente de base de datos instalado.
- Una base de datos world que contiene 3 tablas: city, country, y countrylanguage.



Al final de este laboratorio habrá usado ambas cláusulas **GROUP BY** y **OVER** con algunos operadores de base de datos comunes:

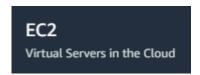


## Tarea 1: Conectarse a Command Host

En esta tarea, se conectará a una instancia de EC2 configurada con un cliente de base de datos (Command Host).

El cliente se usa para ejecutar las queries de SQL contra una base de datos relacional.

Paso 1: AWS Management Console → Services → Compute → EC2



### Paso 2: Instances → Command Host

| ✓        | Name         | Instance ID         | Instance state | ▽  | Instance type | $\triangledown$ | Status check      |
|----------|--------------|---------------------|----------------|----|---------------|-----------------|-------------------|
| <b>✓</b> | Command Host | i-0524ea72e85aedde0 |                | ⊕Q | t3.micro      |                 | 2/2 checks passed |

### Paso 3: Command Host → Connect → Session Manager

# Connect to instance Info Connect to your instance i-0524ea72e85aedde0 (Command Host) using any of these options EC2 Instance Connect Session Manager SSH client EC2 serial console Session Manager usage: • Connect to your instance without SSH keys, a bastion host, or opening any inbound ports. • Sessions are secured using an AWS Key Management Service key. • You can log session commands and details in an Amazon S3 bucket or CloudWatch Logs log group. • Configure sessions on the Session Manager Preferences ☑ page.

Paso 4: Configurar la terminal para acceder a todas las herramientas y recursos necesarios.

```
sudo su
cd /home/ec2-user/
```

- sudo = ejecutar comandos con privilegios.
- su = cambiar a otro usuario.
- cd = cambiar de directorio.

Session ID: user2741130=Tom\_sVillaseca-0ade39e12dc076495

```
sh-4.2$ sudo su
[root@ip-10-1-11-97 bin] # cd /home/ec2-user/
[root@ip-10-1-11-97 ec2-user] # pwd
/home/ec2-user
[root@ip-10-1-11-97 ec2-user]#
```

Paso 5: Conectarse a la base de datos relacional.

```
mysql -u root --password='re:St@rt!9'
```

- mysql = shell SQL para interactuar con los motores de bases de datos.
- -u = nombre de usuario de MySQL
- -p = contraseña de MySQL

Session ID: user2741130=Tom\_sVillaseca-0ade39e12dc076495 Instance ID: i-0524ea72e85aedde0

```
[root@ip-10-1-11-97 ec2-user] # mysql -u root --password='re:St@rt!9'
Welcome to the MariaDB monitor. Commands end with ; or \q.
Your MariaDB connection id is 13
Server version: 10.6.15-MariaDB MariaDB Server
Copyright (c) 2000, 2018, Oracle, MariaDB Corporation Ab and others.
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
MariaDB [(none)]>
```

# Tarea 2: Consulte la base de datos world

En esta tarea, consultará la base de datos **world** usando varias **SELECT** y funciones de la base de datos.

Paso 1: Mostrar las bases de datos existentes.

SHOW DATABASES;

Paso 2: Ver el esquema, los datos y la cantidad de filas de la tabla country.

SHOW COLUMNS FROM world.country; SELECT \* FROM world.country;

```
MariaDB [(none)] > SHOW COLUMNS FROM world.country;
 Field
                                                                                                               | Null | Key | Default | Extra
 Code
                                                                                                                 NO
                   enum('Asia','Europe','North America','Africa','Oceania','Antarctica','South America')
 Continent
                                                                                                                 NO
                                                                                                                               Asia
                   char (26)
decimal (10,2)
 Region
                                                                                                                 NO
                                                                                                                               0.00
 SurfaceArea
                                                                                                                 NO
                   smallint(6)
                                                                                                                 YES
 IndepYear
                                                                                                                               NULL
 Population
                   int (11)
                                                                                                                 NO
 LifeExpectancy
                   decimal(3,1)
                                                                                                                 YES
                                                                                                                               NULL
                   decimal(10,2)
                                                                                                                 YES
                                                                                                                               NULL
 GNPOld
                   decimal(10,2)
                                                                                                                 YES
                                                                                                                               NULL
 LocalName
                   char (45)
                                                                                                                 NO
 GovernmentForm
                   char (45)
                                                                                                                 NO
                                                                                                                               NULL
 Capital
                   int (11)
                                                                                                                 YES
 Code2
                   char(2)
                                                                                                                 NO
   rows in set (0.001 sec)
```

# **Paso 3:** Realizar una consulta en donde **Region** es "Australia" y "New Zealand" y se encuentre ordenada en base a la columna **Population** en orden descendente.

SELECT Region, Name, Population FROM world.country WHERE Region = 'Australia and New Zealand' ORDER By Population desc;

**Paso 4:** Realizar una consulta en donde obtenga el número total de población de las regiones de Australia y New Zealand.

- Region es "Australia" y "New Zealand"
- Agrupa ambas regiones en una utilizando GROUP BY
- Aplicar la función SUM() a los resultados para obtener el número total de población.

SELECT Region, SUM(Population) FROM world.country WHERE Region = 'Australia and New Zealand' GROUP By Region ORDER By SUM(Population) desc;

Paso 5: Realizar una consulta que utilice la función OVER() para agrupar los registros por Region y que aplique la función SUM() para sumar los registros.

```
SELECT Region, Name, Population, SUM(Population) OVER(partition by Region ORDER BY Population) as 'Running Total' FROM world.country WHERE Region = 'Australia and New Zealand';
```

El resultado muestra la población de un pais junto con el total continuo de la región.

| Region                    | Name                    | Population | Running Total |
|---------------------------|-------------------------|------------|---------------|
| Australia and New Zealand | Cocos (Keeling) Islands | 600        | 600           |
| Australia and New Zealand | Norfolk Island          | 2000       | 2600          |
| Australia and New Zealand | Christmas Island        | 2500       | 5100          |
| Australia and New Zealand | New Zealand             | 3862000    | 3867100       |
| Australia and New Zealand | Australia               | 18886000   | 22753100      |

**Paso 6:** Realizar una consulta que agrupe los registros por **Region** y los ordene por **Population** utilizando la función **OVER()**. Además utilizar la función **RANK()** para generar un número de rango que indica la posición de cada registro en el conjunto de resultados.

SELECT Region, Name, Population, SUM(Population) OVER(partition by Region ORDER BY Population) as 'Running Total', RANK() over(partition by region ORDER BY population) as 'Ranked' FROM world.country WHERE region = 'Australia and New Zealand';

| MariaDB [(none)]> SELECT Region try WHERE region = 'Austra | 2         | (Population) ( | OVER (partition by | y Region OR |
|--|-----------|----------------|--------------------|-------------|
| Region   | Name      | Population     | Running Total      | Ranked      |
| Australia and New Zealand                                  | . 3,      | •              | •                  |             |
| Australia and New Zealand  <br>  Australia and New Zealand |           | 2000<br>  2500 |                    |             |
| Australia and New Zealand                                  |           | 3862000        |                    |             |
| Australia and New Zealand                                  | Australia | 18886000       | 22753100           | 5           |
| 5 rows in set (0.001 sec)                                  | +         | +              | +                  | ++          |

### Desafió

Describa una consulta para calificar los países en cada región por su población de mayor a menor.

SELECT Region, Name, Population, RANK() OVER(partition by Region ORDER BY Population desc) as 'Ranked' FROM world.country order by Region, Ranked;

|                           | ion, Name, Population, RANK() OVER(partition by |            |        |
|---------------------------|---|------------|--------|
| Region                    | Name  | Population | Ranked |
| Antarctica                | French Southern territories                     | J 0        | 1      |
| Antarctica                | Bouvet Island                                   | 0          | 1      |
| Antarctica                | South Georgia and the South Sandwich Islands    | 0          | 1      |
| Antarctica                | Antarctica                                      | 0          | 1      |
| Antarctica                | Heard Island and McDonald Islands               | 0          | 1      |
| Australia and New Zealand | Australia                                       | 18886000   | 1      |
| Australia and New Zealand | New Zealand                                     | 3862000    | 2      |
| Australia and New Zealand | Christmas Island                                | 2500       | ] 3    |
| Australia and New Zealand | Norfolk Island                                  | 2000       | 4      |
| Australia and New Zealand | Cocos (Keeling) Islands                         | 600        | 5      |
| Baltic Countries          | Lithuania                                       | 3698500    | 1      |
| Baltic Countries          | Latvia  | 2424200    | 2      |
| Baltic Countries          | Estonia   | 1439200    | ] 3    |
| British Islands           | United Kingdom                                  | 59623400   | 1      |
| British Islands           | Ireland   | 3775100    | 2      |
| Caribbean                 | Cuba  | 11201000   | 1      |
| Caribbean                 | Dominican Republic                              | 8495000    | 2      |
| Caribbean                 | Haiti   | 8222000    | ] 3    |
| Caribbean                 | Puerto Rico                                     | 3869000    | 4      |
| Caribbean                 | Jamaica   | 2583000    | ] 5    |
| Caribbean                 | Trinidad and Tobago                             | 1295000    | 6      |
| Caribbean                 | Guadeloupe                                      | 456000     | 1 7    |
| Caribbean                 | Martinique                                      | 395000     | J 8    |
| Caribbean                 | Bahamas   | 307000     | 9      |
| Caribbean                 | Barbados  | 270000     | 10     |



