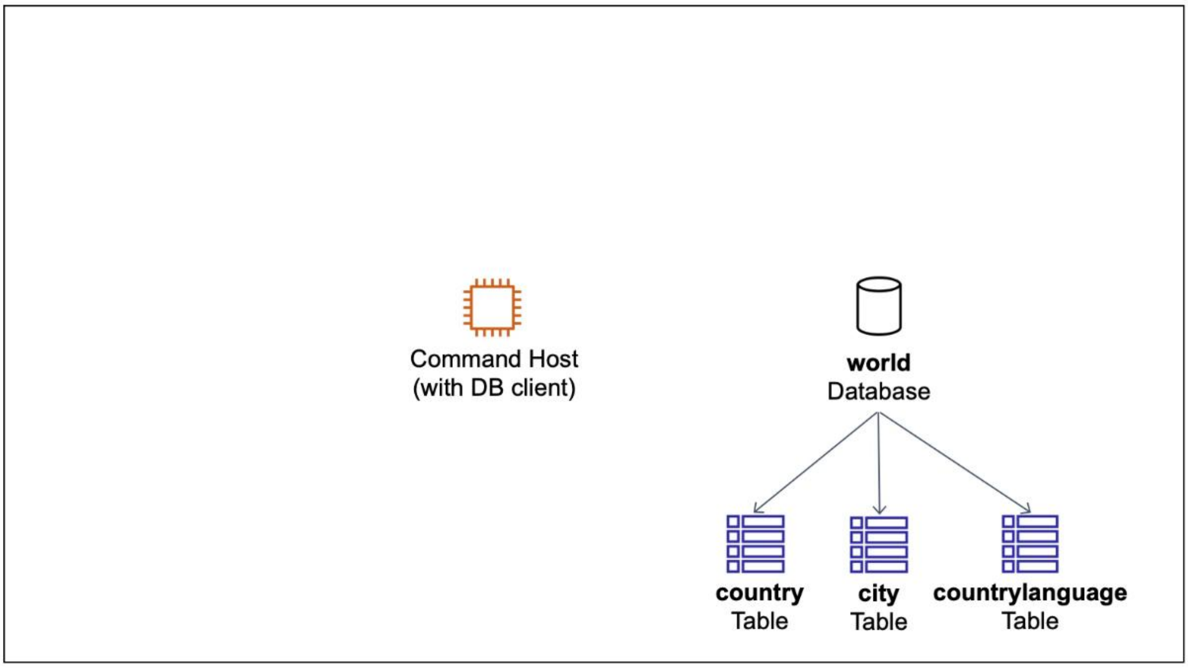
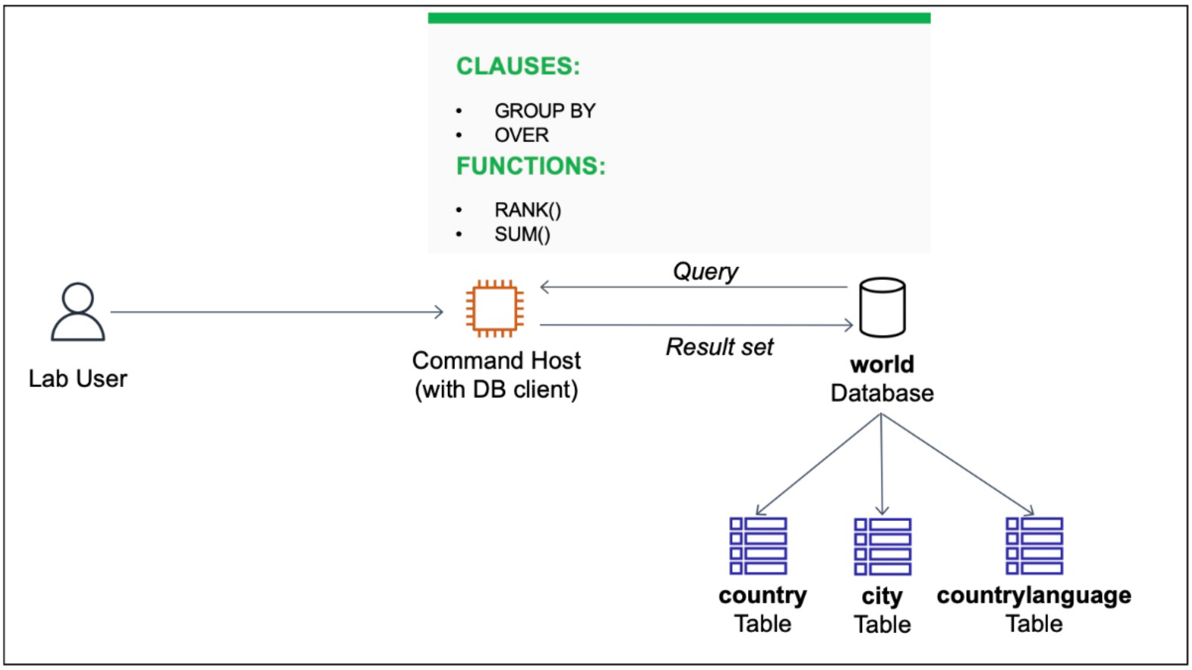
Organizing Data



A Command Host instance and world database containing three tables

At the end of this lab, you would have used both the **GROUP BY** and **OVER** clauses with some common database operators:

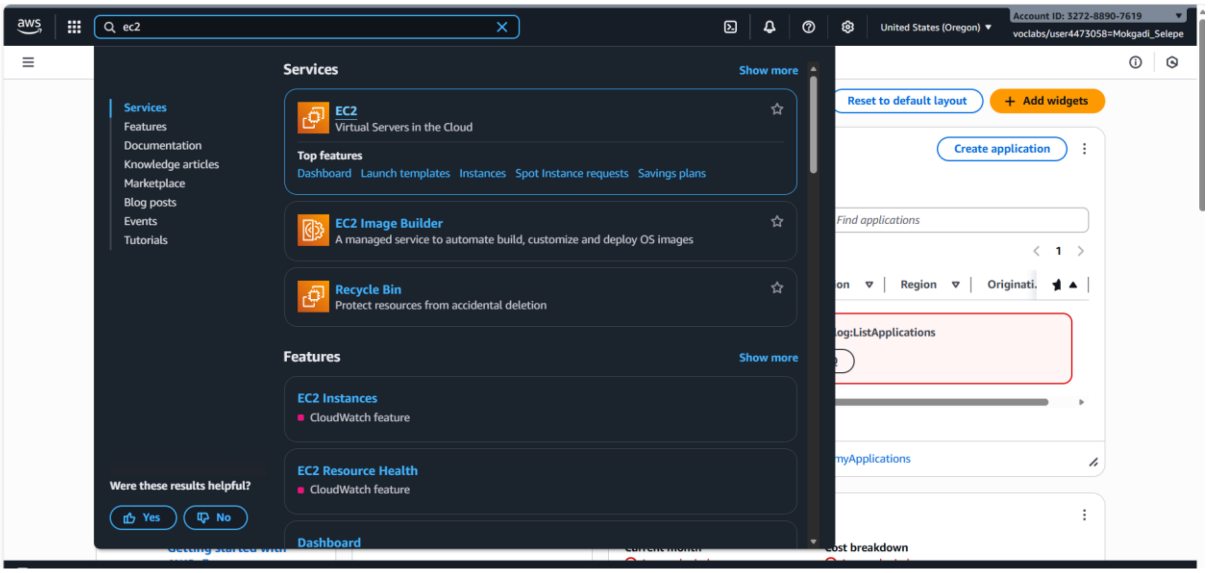


A lab user is connected to a database instance. It also displays some commonly used SQL clauses and database functions.

Sample data in this course is taken from Statistics Finland, general regional statistics, February 4, 2022.

MOKGADI SELEPE

1: Connect to the Command Host



Services

EC2
Virtual Servers in the Cloud

Top features
Dashboard Launch templates Instances Spot Instance requests Savings plans

EC2 Image Builder
A managed service to automate build, customize and deploy OS images

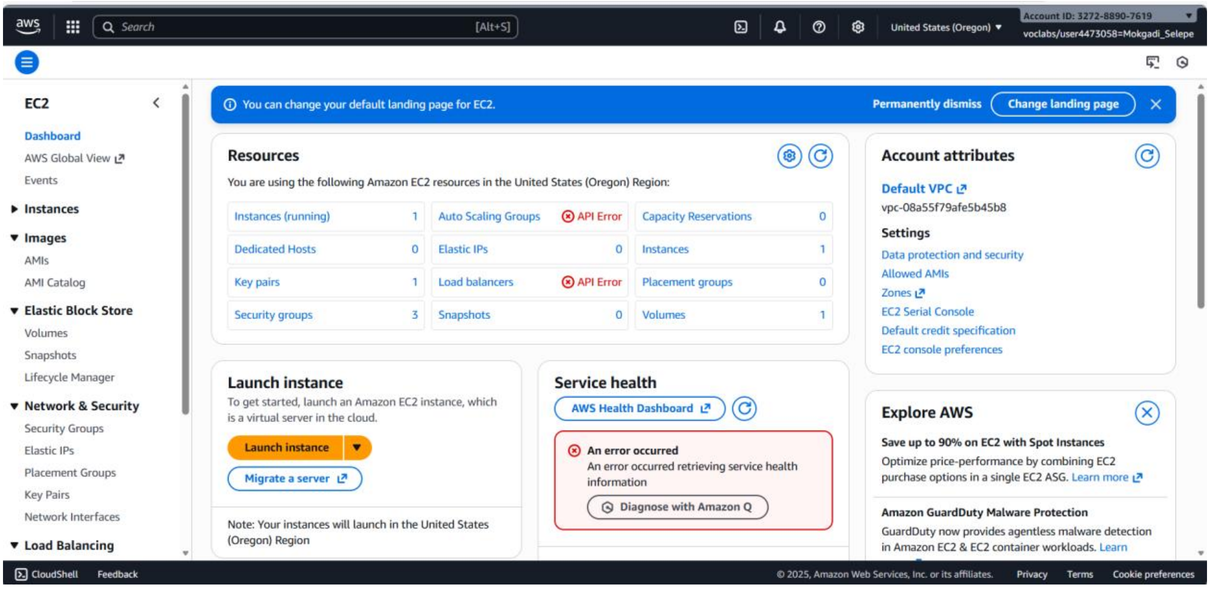
Recycle Bin
Protect resources from accidental deletion

Features

EC2 Instances
CloudWatch feature

EC2 Resource Health
CloudWatch feature

Dashboard



You can change your default landing page for EC2. Permanently dismiss Change landing page

Resources

You are using the following Amazon EC2 resources in the United States (Oregon) Region:

Instances (running)	1	Auto Scaling Groups	API Error	Capacity Reservations	0
Dedicated Hosts	0	Elastic IPs	0	Instances	1
Key pairs	1	Load balancers	API Error	Placement groups	0
Security groups	3	Snapshots	0	Volumes	1

Account attributes

Default VPC
vpc-08a55f79afe5b45b8

Settings
Data protection and security
Allowed AMIs
Zones
EC2 Serial Console
Default credit specification
EC2 console preferences

Launch instance

To get started, launch an Amazon EC2 instance, which is a virtual server in the cloud.

Launch instance
Migrate a server

Note: Your instances will launch in the United States (Oregon) Region

Service health

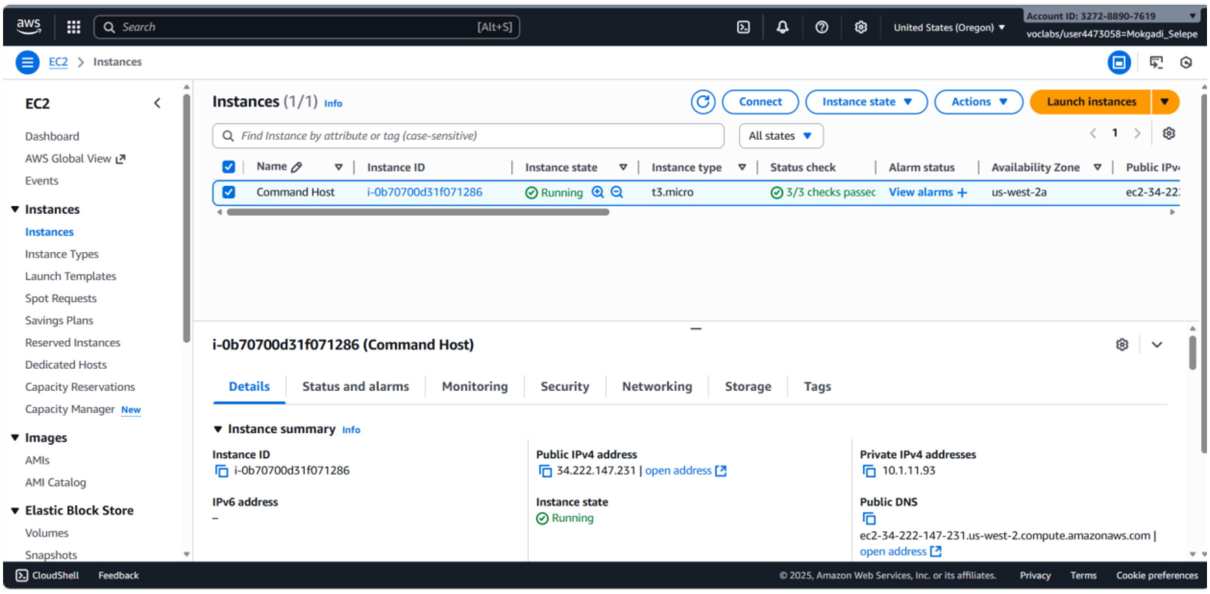
AWS Health Dashboard

An error occurred
An error occurred retrieving service health information
Diagnose with Amazon Q

Explore AWS

Save up to 90% on EC2 with Spot Instances
Optimize price-performance by combining EC2 purchase options in a single EC2 ASG. Learn more

Amazon GuardDuty Malware Protection
GuardDuty now provides agentless malware detection in Amazon EC2 & EC2 container workloads. Learn



Instances (1/1) Info
Connect Instance state Actions Launch instances

Find Instance by attribute or tag (case-sensitive) All states

☑	Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv
☑	Command Host	i-0b70700d31f071286	✓ Running	t3.micro	✓ 3/3 checks passec	View alarms +	us-west-2a	ec2-34-22

i-0b70700d31f071286 (Command Host)

Details Status and alarms Monitoring Security Networking Storage Tags

Instance summary Info

Instance ID
i-0b70700d31f071286

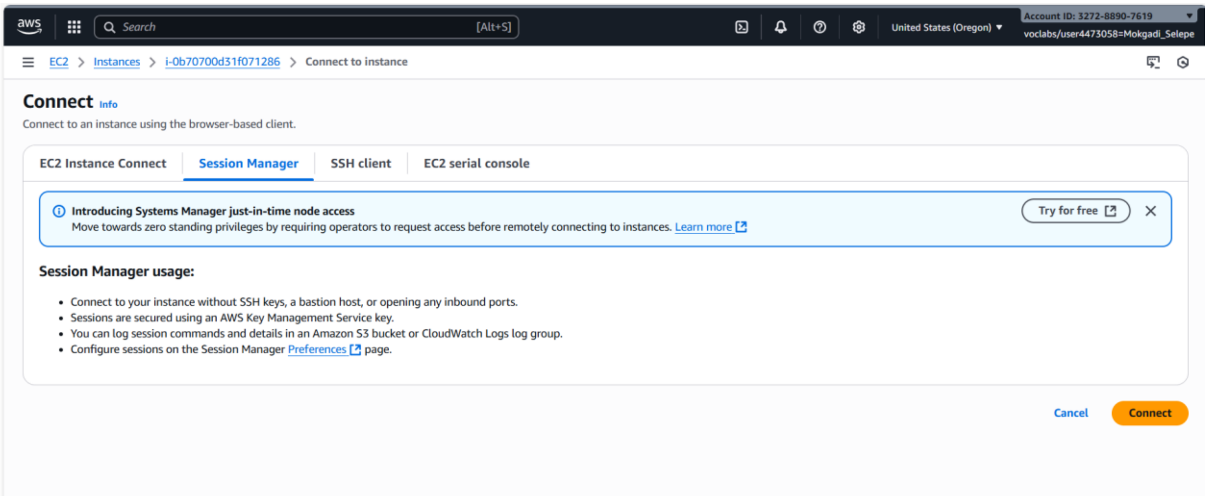
IPv6 address
–

Public IPv4 address
34.222.147.231 | open address

Instance state
✓ Running

Private IPv4 addresses
10.1.11.93

Public DNS
ec2-34-222-147-231.us-west-2.compute.amazonaws.com | open address



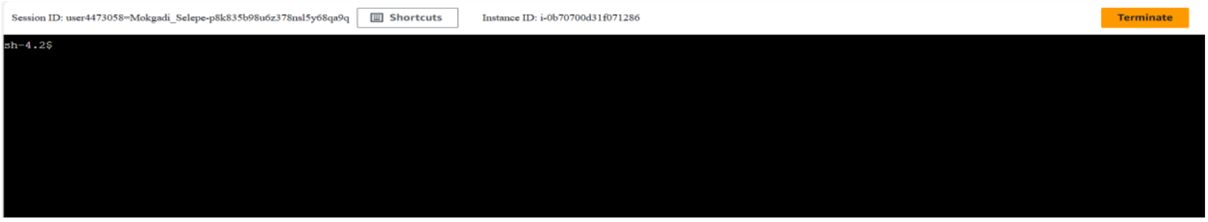
Connect Info

Connect to an instance using the browser-based client.

Introducing Systems Manager just-in-time node access
Move towards zero standing privileges by requiring operators to request access before remotely connecting to instances. Learn more

Session Manager usage:

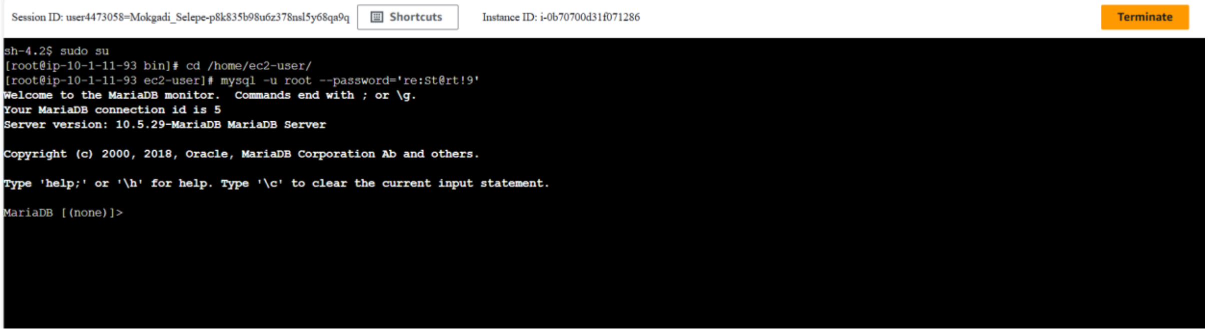
- Connect to your instance without SSH keys, a bastion host, or opening any inbound ports.
- Sessions are secured using an AWS Key Management Service key.
- You can log session commands and details in an Amazon S3 bucket or CloudWatch Logs log group.
- Configure sessions on the Session Manager Preferences page.



Session ID: user4473058=Mokgadi_Selepe-p8k835b98u6z378nsl5y68qa9q

Instance ID: i-0b70700d31f071286

```
sh-4.2$
```



Session ID: user4473058=Mokgadi_Selepe-p8k835b98u6z378nsl5y68qa9q

Instance ID: i-0b70700d31f071286

```
sh-4.2$ sudo su
[root@ip-10-1-11-93 bin]# cd /home/ec2-user/
[root@ip-10-1-11-93 ec2-user]# mysql -u root --password='re:St@rt!9'
Welcome to the MariaDB monitor.  Commands end with ; or \g.
Your MariaDB connection id is 5
Server version: 10.5.29-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)]>
```

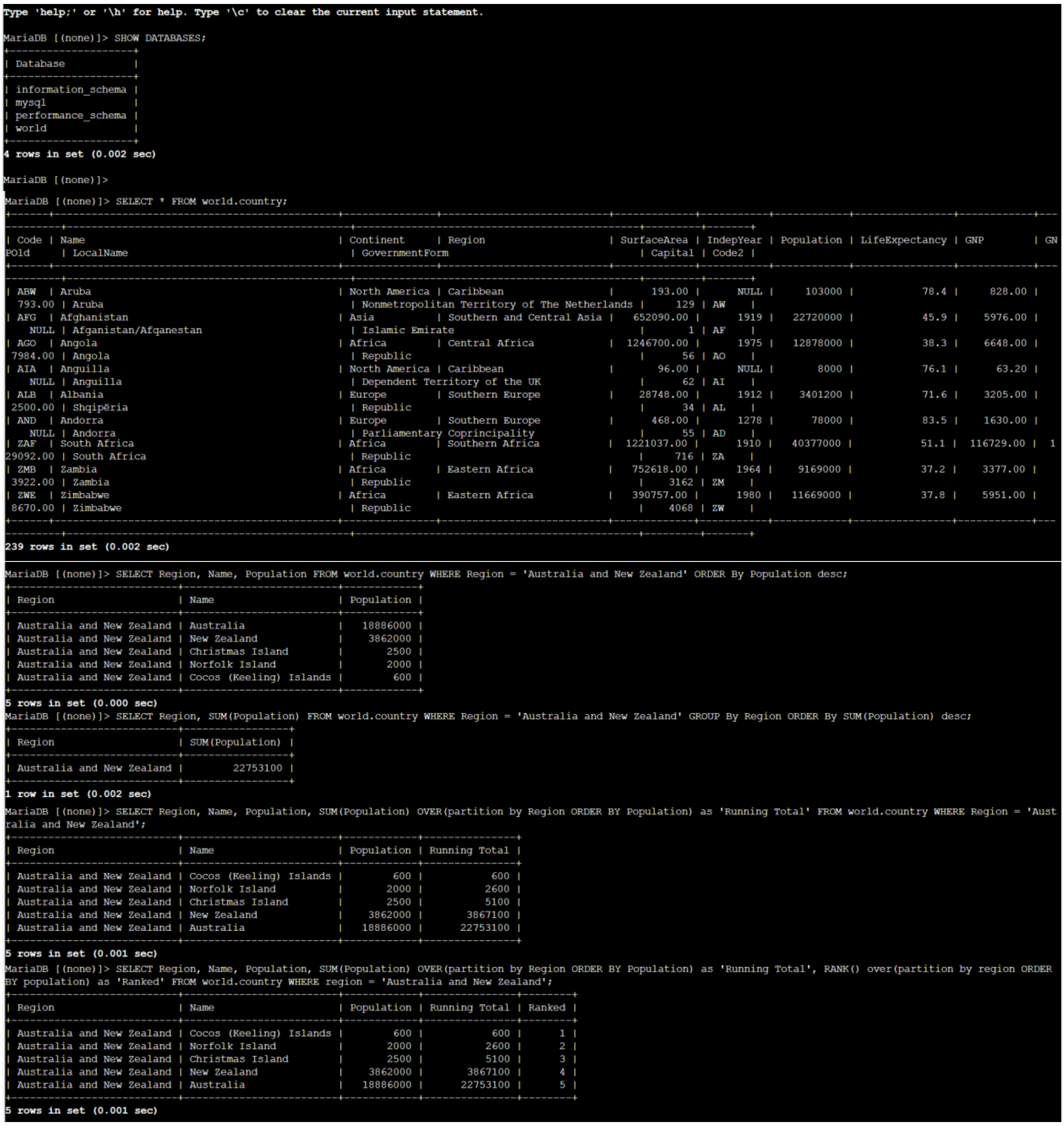
Here's what happened:

I connected to a special computer in the cloud that has a tool to talk to a database. Here's how I did it:

1. I went to the AWS website and found the special computer, called the Command Host.
2. I clicked some buttons to connect to it, and a new window opened up.
3. In the new window, I typed some commands to get everything set up.
4. Then, I typed another command to connect to the database, using a secret password.

Now I'm connected to the database and can start working with it!

Think of it like:
- Finding a special computer in the cloud
- Opening a door to that computer
- Getting everything ready to use
- Unlocking the database with a secret password



```
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

MariaDB [(none)]> SHOW DATABASES;
+--------------------+
| Database           |
+--------------------+
| information_schema |
| mysql              |
| performance_schema |
| world              |
+--------------------+
4 rows in set (0.002 sec)

MariaDB [(none)]>

MariaDB [(none)]> SELECT * FROM world.country;
+------+----------------------------------------------+---------------+---------------------------+-------------+-----------+------------+----------------+------------+---
---------+----------------------------------------------+----------------------------------------------+---------+-------+
| Code | Name                                         | Continent     | Region                    | SurfaceArea | IndepYear | Population | LifeExpectancy | GNP        | GN
POld     | LocalName                                    | GovernmentForm                               | Capital | Code2 |
+------+----------------------------------------------+---------------+---------------------------+-------------+-----------+------------+----------------+------------+---
---------+----------------------------------------------+----------------------------------------------+---------+-------+
| ABW  | Aruba                                        | North America | Caribbean                 |      193.00 |      NULL |     103000 |           78.4 |     828.00 |
  793.00 | Aruba                                        | Nonmetropolitan Territory of The Netherlands |     129 | AW    |
| AFG  | Afghanistan                                  | Asia          | Southern and Central Asia |   652090.00 |      1919 |   22720000 |           45.9 |    5976.00 |
    NULL | Afganistan/Afqanestan                        | Islamic Emirate                              |       1 | AF    |
| AGO  | Angola                                       | Africa        | Central Africa            |  1246700.00 |      1975 |   12878000 |           38.3 |    6648.00 |
 7984.00 | Angola                                       | Republic                                     |      56 | AO    |
| AIA  | Anguilla                                     | North America | Caribbean                 |       96.00 |      NULL |       8000 |           76.1 |      63.20 |
    NULL | Anguilla                                     | Dependent Territory of the UK                |      62 | AI    |
| ALB  | Albania                                      | Europe        | Southern Europe           |    28748.00 |      1912 |    3401200 |           71.6 |    3205.00 |
 2500.00 | Shqipëria                                    | Republic                                     |      34 | AL    |
| AND  | Andorra                                      | Europe        | Southern Europe           |      468.00 |      1278 |      78000 |           83.5 |    1630.00 |
    NULL | Andorra                                      | Parliamentary Coprincipality                 |      55 | AD    |
| ZAF  | South Africa                                 | Africa        | Southern Africa           |  1221037.00 |      1910 |   40377000 |           51.1 |  116729.00 | 1
29092.00 | South Africa                                 | Republic                                     |     716 | ZA    |
| ZMB  | Zambia                                       | Africa        | Eastern Africa            |   752618.00 |      1964 |    9169000 |           37.2 |    3377.00 |
 3922.00 | Zambia                                       | Republic                                     |    3162 | ZM    |
| ZWE  | Zimbabwe                                     | Africa        | Eastern Africa            |   390757.00 |      1980 |   11669000 |           37.8 |    5951.00 |
 8670.00 | Zimbabwe                                     | Republic                                     |    4068 | ZW    |
+------+----------------------------------------------+---------------+---------------------------+-------------+-----------+------------+----------------+------------+---
---------+----------------------------------------------+----------------------------------------------+---------+-------+
239 rows in set (0.002 sec)

MariaDB [(none)]> SELECT Region, Name, Population FROM world.country WHERE Region = 'Australia and New Zealand' ORDER By Population desc;
+---------------------------+-------------------------+------------+
| Region                    | Name                    | Population |
+---------------------------+-------------------------+------------+
| Australia and New Zealand | Australia               |   18886000 |
| Australia and New Zealand | New Zealand             |    3862000 |
| Australia and New Zealand | Christmas Island        |       2500 |
| Australia and New Zealand | Norfolk Island          |       2000 |
| Australia and New Zealand | Cocos (Keeling) Islands |        600 |
+---------------------------+-------------------------+------------+
5 rows in set (0.000 sec)
MariaDB [(none)]> SELECT Region, SUM(Population) FROM world.country WHERE Region = 'Australia and New Zealand' GROUP By Region ORDER By SUM(Population) desc;
+---------------------------+-----------------+
| Region                    | SUM(Population) |
+---------------------------+-----------------+
| Australia and New Zealand |        22753100 |
+---------------------------+-----------------+
1 row in set (0.002 sec)
MariaDB [(none)]> SELECT Region, Name, Population, SUM(Population) OVER(partition by Region ORDER BY Population) as 'Running Total' FROM world.country WHERE Region = 'Aust
ralia and New Zealand';
+---------------------------+-------------------------+------------+---------------+
| 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 |
+---------------------------+-------------------------+------------+---------------+
5 rows in set (0.001 sec)
MariaDB [(none)]> 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';
+---------------------------+-------------------------+------------+---------------+--------+
| Region                    | Name                    | Population | Running Total | Ranked |
+---------------------------+-------------------------+------------+---------------+--------+
| Australia and New Zealand | Cocos (Keeling) Islands |        600 |           600 |      1 |
| Australia and New Zealand | Norfolk Island          |       2000 |          2600 |      2 |
| Australia and New Zealand | Christmas Island        |       2500 |          5100 |      3 |
| Australia and New Zealand | New Zealand             |    3862000 |       3867100 |      4 |
| Australia and New Zealand | Australia               |   18886000 |      22753100 |      5 |
+---------------------------+-------------------------+------------+---------------+--------+
5 rows in set (0.001 sec)
```

Here's what happened:

I played around with a database called "world" and asked it questions using special commands.

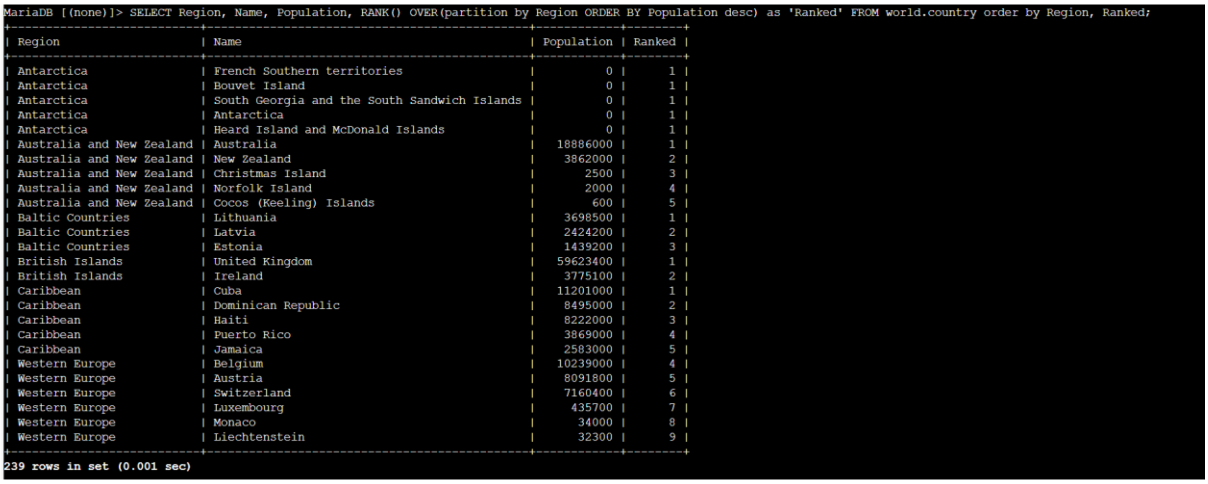
Here's what I did:

1. I checked what databases are available and found the "world" database.
2. I looked at a table called "country" in the "world" database and saw all the data in it.
3. I asked for countries in the "Australia and New Zealand" region, sorted by population from highest to lowest.
4. I grouped countries in the "Australia and New Zealand" region together and calculated their total population.
5. I calculated a running total of population for countries in the "Australia and New Zealand" region.

6. I ranked countries in the "Australia and New Zealand" region by population, showing their position in the list.

I used special tools like GROUP BY, SUM, OVER, and RANK to get the information I needed!

Challenge



```
MariaDB [(none)]> SELECT Region, Name, Population, RANK() OVER(partition by Region ORDER BY Population desc) as 'Ranked' FROM world.country order by Region, Ranked;
+---------------------------+----------------------------------------------+------------+--------+
| Region                    | Name                                         | Population | Ranked |
+---------------------------+----------------------------------------------+------------+--------+
| Antarctica                | French Southern territories                  |          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 |
| Western Europe            | Belgium                                      |   10239000 |      4 |
| Western Europe            | Austria                                      |    8091800 |      5 |
| Western Europe            | Switzerland                                  |    7160400 |      6 |
| Western Europe            | Luxembourg                                   |     435700 |      7 |
| Western Europe            | Monaco                                       |      34000 |      8 |
| Western Europe            | Liechtenstein                                |      32300 |      9 |
+---------------------------+----------------------------------------------+------------+--------+
239 rows in set (0.001 sec)
```

Here's what happened:

I wrote a special command to rank countries in each region by their population, from largest to smallest.

Here's what I did:

1. I chose to use the OVER clause to group countries by region, because I wanted to rank countries within each region.
2. I used the RANK() function to give each country a rank number based on its population.
3. I sorted the results by region and then by rank, so countries in each region are listed in order of population.

The command says: "For each region, rank countries by population from largest to smallest, and show me the region, country name, population, and rank."

So, I got a list of countries in each region, ranked by population!

Conclusion

Here's what happened:

I finished working with the database and learned lots of cool things!

Here's what I did:

1. I grouped data together using GROUP BY and calculated totals using SUM.
2. I used OVER to rank countries within each region using RANK.
3. I also used OVER to calculate running totals and ranks at the same time.

I used these tools to get the information I needed and now I'm done! I know how to work with databases and get insights from data.