



AWS  
re:Start  
LAB

# Working with Functions



**WEEK 7**





# Overview

---

Understanding and leveraging functions in database operations is vital for streamlined data handling. Functions like those used for summarizing data or extracting substrings play a pivotal role in simplifying complex tasks and improving data accuracy. By utilizing functions effectively within data queries and analyses, professionals can enhance the efficiency and accuracy of their data processing workflows.

Moreover, functions enable users to perform various data transformations and validations, such as checking string lengths or removing duplicates. This capability ensures data integrity and enhances the quality of analytical outputs. By incorporating functions seamlessly into data manipulation processes, professionals can achieve more precise results, optimize resource utilization, and make informed decisions based on reliable data insights.

## Topics covered

- Use aggregate functions `SUM( )`, `MIN( )`, `MAX( )`, and `AVG( )` to summarize data
- Use the `SUBSTRING_INDEX( )` function to split strings
- Use the `LENGTH( )` and `TRIM( )` functions to determine the length of a string
- Use the `DISTINCT( )` function to filter duplicate records
- Use functions in the `SELECT` statement and `WHERE` clause

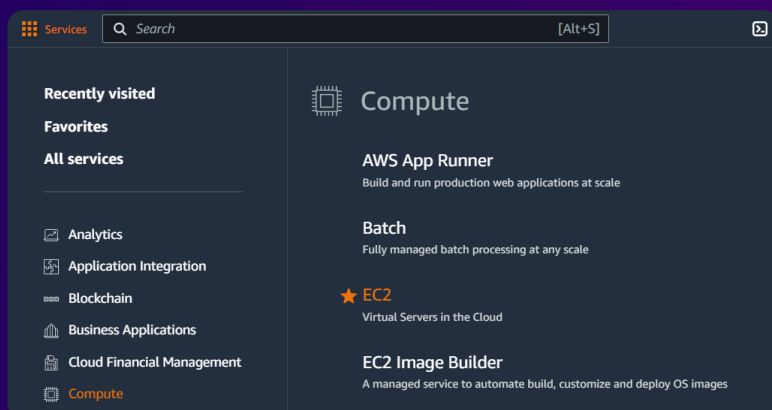


# Task 1

## Connect to the Command Host

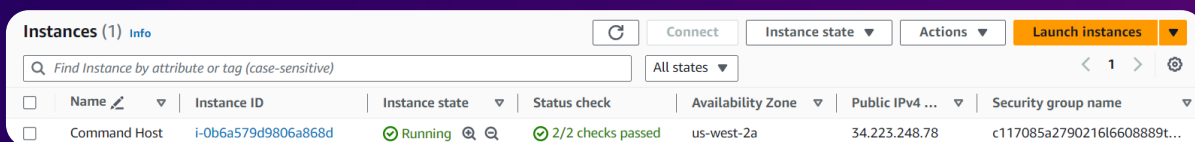
### Step 1: Access the EC2 Management Console

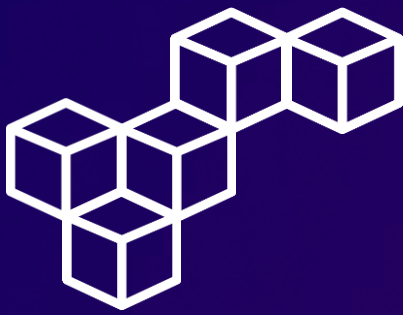
Open the AWS Management Console, and select EC2.



### Step 2: Review running instances

Navigate to the **Instances** section. The running **Command Host** instance is listed.



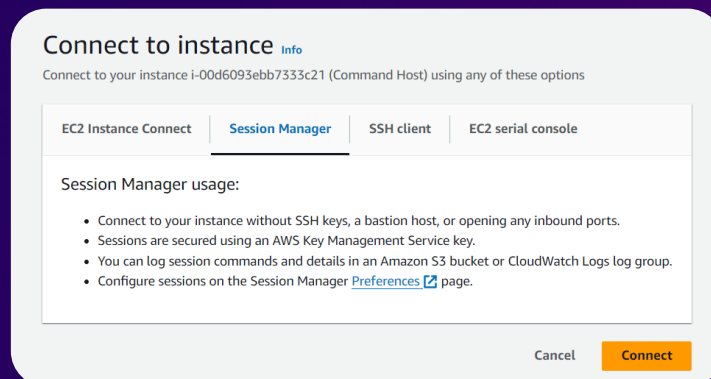


# Task 1

## Connect to the Command Host

### Step 3: Connect to the instance

Connect to the **Command Host** EC2 instance, which contains a database client, using [Session Manager](#).



### Step 4: Connect to the database server

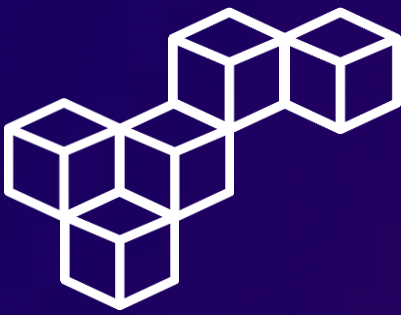
To connect to the database server, run the following commands in the terminal.

```
sh-4.2$ sudo su
[root@ip-10-1-11-196 bin]# cd /home/ec2-user/
[root@ip-10-1-11-196 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 14
Server version: 10.6.17-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)]>
```



# Task 2

## Query the world database

### Step 1: Show existing databases

Show the existing databases using the `SHOW DATABASES` query, and verify that a database named **world** is available.

```
MariaDB [(none)]> SHOW DATABASES;
+-----+
| Database |
+-----+
| information_schema |
| mysql |
| performance_schema |
| sys |
| world |
+-----+
5 rows in set (0.003 sec)
```

### Step 2: Review the table schema

Review the table schema, data, and number of rows in the **country** table, using the `SELECT * FROM` statement.

| MariaDB [(none)]> SELECT * FROM world.country; |                             |                |                           |             |           |            |                |           |           |                                  |   |         |       |
|--|-----------------------------|----------------|---------------------------|-------------|-----------|------------|----------------|-----------|-----------|----------------------------------|---|---------|-------|
| Code   | Name                        | Continent      | Region                    | SurfaceArea | IndepYear | Population | LifeExpectancy | GDP       | GDPGold   | LocalName                        | GovernmentForm                            | Capital | Code2 |
| AW   | Aruba                       | North America  | Caribbean                 | 193.00      | ND        | 100000     | 78.4           | 528.00    | 793.00    | Aruba                            | Metropolitan Territory of The Netherlands | 129     | AW    |
| AF   | Afghanistan                 | Asia           | Southern and Central Asia | 652090.00   | 1919      | 20720000   | 45.9           | 5994.00   | ND        | Afghanistan/Afghanistan          | Islamic Emirate                           | 1       | AF    |
| AO   | Angola                      | Africa         | Central Africa            | 1246700.00  | 1975      | 12870000   | 35.3           | 6649.00   | 7994.00   | Angola                           | Republic                                  | 56      | AO    |
| AI   | Anguilla                    | North America  | Caribbean                 | 94.00       | ND        | 8000       | 76.1           | 63.20     | ND        | Anguilla                         | Dependent Territory of the UK             | 62      | AI    |
| AD   | Andorra                     | Europe         | Southern Europe           | 2474.00     | 1912      | 261200     | 71.4           | 2220.00   | 2000.00   | Andorra                          | Republic                                  | 24      | AD    |
| AE   | United Arab Emirates        | North America  | Caribbean                 | 800.00      | ND        | 217000     | 74.7           | 1941.00   | ND        | Netherlands Antilles             | Metropolitan Territory of The Netherlands | 23      | AE    |
| AG   | Argentina                   | Asia           | Middle East               | 284000.00   | 1971      | 2442000    | 74.1           | 27946.00  | 26846.00  | al-Jazayer al-Ashjaq al-Mutahida | Islamic Emirate                           | 65      | AG    |
| AM   | Armenia                     | South America  | South America             | 2780400.00  | 1816      | 37022000   | 75.1           | 240228.00 | 222810.00 | Argentina                        | Federal Republic                          | 69      | AM    |
| AS   | American Samoa              | Asia           | Middle East               | 23800.00    | 1991      | 3520000    | 66.4           | 1812.00   | 1627.00   | Hajastan                         | Republic                                  | 126     | AS    |
| AT   | Austria                     | Europe         | Polynesia                 | 180.00      | ND        | 60000      | 75.1           | 234.00    | ND        | America Samoa                    | US Territory                              | 54      | AT    |
| AX   | Åland Islands               | Antarctica     | Antarctica                | 15120000.00 | ND        | 0          | ND             | 0.00      | ND        | -                                | Co-administrated                          | ND      | AX    |
| ATF  | French Southern territories | Antarctica     | Antarctica                | 7780.00     | ND        | 0          | ND             | 0.00      | ND        | Territoires australes françaises | Metropolitan Territory of France          | ND      | ATF   |
| AX   | Åland Islands               | North America  | Caribbean                 | 440.00      | 1941      | 1888000    | 70.5           | 612.00    | 384.00    | Åndipap and Baidip               | Constitutional Monarchy                   | 62      | AX    |
| AD   | Andorra                     | Europe         | Western Europe            | 462.00      | 1991      | 1888000    | 79.8           | 251152.00 | 292911.00 | Australia                        | Constitutional Monarchy, Federation       | 125     | AD    |
| AT   | Austria                     | Europe         | Western Europe            | 83850.00    | 1918      | 9051800    | 77.7           | 211460.00 | 206035.00 | Osterreich                       | Federal Republic                          | 1223    | AT    |
| AE   | Armenia                     | Asia           | Middle East               | 284000.00   | 1918      | 2774000    | 62.9           | 6127.00   | 6100.00   | Armenia                          | Constitutional Monarchy, Federation       | 144     | AE    |
| RU   | Russia                      | Africa         | Eastern Africa            | 27824.00    | 1962      | 6685000    | 46.2           | 503.00    | 582.00    | Ruundi/Urundi                    | Republic                                  | 552     | RU    |
| BE   | Belgium                     | Europe         | Western Europe            | 10512.00    | 1830      | 10259000   | 77.8           | 240704.00 | 240340.00 | België/Belgique                  | Constitutional Monarchy, Federation       | 179     | BE    |
| BN   | Brunei                      | Africa         | Western Africa            | 111420.00   | 1960      | 4097000    | 50.2           | 2292.00   | 2141.00   | Brunei                           | Republic                                  | 157     | BN    |
| BF   | Burkina Faso                | Asia           | Western Africa            | 274000.00   | 1960      | 11377000   | 46.7           | 2425.00   | 2201.00   | Burkina Faso                     | Republic                                  | 149     | BF    |
| BD   | Bangladesh                  | Europe         | Southern and Central Asia | 146960.00   | 1971      | 120110000  | 60.2           | 25262.00  | 21666.00  | Bangladesh                       | Republic                                  | 100     | BD    |
| BS   | Bahamas                     | Europe         | Eastern Europe            | 132984.00   | 1958      | 3105000    | 70.9           | 12178.00  | 10169.00  | Bahamas                          | Republic                                  | 239     | BS    |
| HR   | Croatia                     | Asia           | Middle East               | 564.00      | 1971      | 417000     | 72.0           | 4566.00   | 6097.00   | Al-Bahrayn                       | Monarchy (Deirats)                        | 149     | HR    |
| RU   | Russia                      | North America  | Caribbean                 | 13979.00    | 1978      | 897000     | 71.1           | 2527.00   | 2347.00   | The Bahamas                      | Constitutional Monarchy                   | 149     | RU    |
| RU   | Russia                      | Europe         | Southern Europe           | 51197.00    | 1992      | 3972000    | 71.5           | 2841.00   | ND        | Thomson & Desseguins             | Federal Republic                          | 201     | RU    |
| RU   | Russia                      | Europe         | Eastern Europe            | 207400.00   | 1991      | 10288000   | 68.0           | 17174.00  | ND        | Belarus                          | Republic                                  | 2320    | RU    |
| RU   | Russia                      | North America  | Central America           | 22498.00    | 1991      | 241800     | 62.9           | 776738.00 | 804109.00 | Beliz                            | Constitutional Monarchy                   | 183     | RU    |
| RU   | Russia                      | North America  | North America             | 53.00       | ND        | 40000      | 74.9           | 2228.00   | 2190.00   | Bermuda                          | Dependent Territory of the UK             | 191     | RU    |
| RU   | Russia                      | South America  | South America             | 1095881.00  | 1825      | 8329000    | 62.7           | 8571.00   | 7967.00   | Bolivia                          | Republic                                  | 194     | RU    |
| RU   | Russia                      | South America  | South America             | 8547400.00  | 1825      | 17011000   | 62.9           | 776738.00 | 804109.00 | Brazil                           | Federal Republic                          | 211     | RU    |
| RU   | Russia                      | North America  | Caribbean                 | 420.00      | 1966      | 270000     | 72.0           | 2222.00   | 2186.00   | Bahamas                          | Constitutional Monarchy                   | 174     | RU    |
| RU   | Russia                      | Asia           | Southeast Asia            | 3765.00     | 1984      | 225000     | 72.6           | 11705.00  | 12460.00  | Brunei Darussalam                | Monarchy (Sultanate)                      | 538     | RU    |
| RU   | Russia                      | Asia           | Southern and Central Asia | 47000.00    | 1910      | 2148000    | 52.4           | 372.00    | 353.00    | Indo-Nel                         | Monarchy                                  | 152     | RU    |
| RU   | Russia                      | Antarctica     | Antarctica                | 59.00       | ND        | 0          | ND             | 0.00      | ND        | South Georgia                    | Dependent Territory of Norway             | ND      | RU    |
| RU   | Russia                      | South America  | South America             | 501720.00   | 1966      | 1622000    | 35.3           | 4834.00   | 4835.00   | Botswana                         | Republic                                  | 204     | RU    |
| RU   | Russia                      | Central Africa | Central Africa            | 620948.00   | 1960      | 2612000    | 44.0           | 1054.00   | 993.00    | Central African Republic         | Republic                                  | 180     | RU    |
| RU   | Russia                      | North America  | North America             | 9970410.00  | 1867      | 21147000   | 79.4           | 588862.00 | 625626.00 | Canada                           | Constitutional Monarchy, Federation       | 1222    | RU    |



## Task 2

# Query the world database

### Step 3: Use aggregate functions

Use the aggregate functions `SUM()`, `AVG()`, `MAX()`, `MIN()`, and `COUNT()` to aggregate data from all records in the **country** table.

```
MariaDB [(none)]> SELECT SUM(Population), AVG(Population), MAX(Population), MIN(Population), COUNT(Population) FROM world.country;
```

| SUM(Population) | AVG(Population) | MAX(Population) | MIN(Population) | COUNT(Population) |
|-----------------|-----------------|-----------------|-----------------|-------------------|
| 6078749450      | 25434098.1172   | 1277558000      | 0               | 239               |

```
1 row in set (0.001 sec)
```

### Step 4: Split a string

Use the function `SUBSTRING_INDEX()` to split a string where a space occurs.

```
MariaDB [(none)]> SELECT Region, SUBSTRING_INDEX(Region, " ", 1) FROM world.country;
```

| Region                    | SUBSTRING_INDEX(Region, " ", 1) |
|---------------------------|---------------------------------|
| Caribbean                 | Caribbean                       |
| Southern and Central Asia | Southern                        |
| Central Africa            | Central                         |
| Caribbean                 | Caribbean                       |
| Southern Europe           | Southern                        |
| Southern Europe           | Southern                        |
| Caribbean                 | Caribbean                       |
| Middle East               | Middle                          |
| South America             | South                           |
| Middle East               | Middle                          |
| Polynesia                 | Polynesia                       |
| Antarctica                | Antarctica                      |
| Antarctica                | Antarctica                      |
| Caribbean                 | Caribbean                       |
| Australia and New Zealand | Australia                       |
| Western Europe            | Western                         |
| Middle East               | Middle                          |
| Eastern Africa            | Eastern                         |
| Western Europe            | Western                         |
| Western Africa            | Western                         |
| Western Africa            | Western                         |
| Southern and Central Asia | Southern                        |



# Task 2

## Query the world database

### Step 5: Search rows using a string fragment

Filter records that include "Southern" in the first part of the region name using the `SUBSTRING_INDEX()` function.

```
MariaDB [(none)]> SELECT Name, Region FROM world.country WHERE SUBSTRING_INDEX(Region, " ", 1) = "Southern";
```

| Name                   | Region                    |
|------------------------|---------------------------|
| Afghanistan            | Southern and Central Asia |
| Albania                | Southern Europe           |
| Andorra                | Southern Europe           |
| Bangladesh             | Southern and Central Asia |
| Bosnia and Herzegovina | Southern Europe           |
| Bhutan                 | Southern and Central Asia |
| Botswana               | Southern Africa           |
| Spain                  | Southern Europe           |
| Gibraltar              | Southern Europe           |
| Greece                 | Southern Europe           |
| Croatia                | Southern Europe           |
| India                  | Southern and Central Asia |
| Iran                   | Southern and Central Asia |
| Italy                  | Southern Europe           |
| Kazakhstan             | Southern and Central Asia |

### Step 6: Determine the length of a string

Return only regions that have fewer than 10 characters in their names using the `LENGTH()` function.

```
MariaDB [(none)]> SELECT Region FROM world.country WHERE LENGTH(TRIM(Region)) < 10;
```

| Region    |
|-----------|
| Caribbean |
| Caribbean |
| Caribbean |
| Polynesia |
| Caribbean |
| Caribbean |
| Caribbean |
| Polynesia |
| Caribbean |
| Caribbean |
| Caribbean |
| Caribbean |
| Melanesia |





## Task 2

# Query the world database

### Step 7: Filter duplicate records

You might have noticed duplicate records in the previous example. Use the `DISTINCT()` function to filter the duplicates.

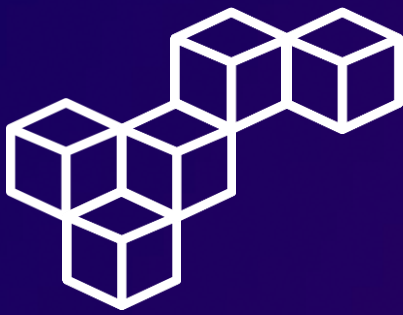
```
MariaDB [(none)]> SELECT DISTINCT(Region) FROM world.country WHERE LENGTH(TRIM(Region)) < 10;
+-----+
| Region |
+-----+
| Caribbean |
| Polynesia |
| Melanesia |
+-----+
3 rows in set (0.001 sec)
```

### Step 8: Challenge

Write a query to return rows that have Micronesia/Caribbean as the name in the region column. The output should split the region as Micronesia and Caribbean into two separate columns: one named Region Name 1 and one named Region Name 2.

```
MariaDB [(none)]> SELECT Name, SUBSTRING_INDEX(Region, "/", 1) AS "Region Name 1", SUBSTRING_INDEX(region, "/", -1)
AS "Region Name 2" FROM world.country WHERE Region = "Micronesia/Caribbean";
+-----+-----+-----+
| Name | Region Name 1 | Region Name 2 |
+-----+-----+-----+
| United States Minor Outlying Islands | Micronesia | Caribbean |
+-----+-----+-----+
1 row in set (0.000 sec)
```





# Conclusions

---

## **The SUM( ) and AVG( ) functions**

The SUM( ) and AVG( ) functions are essential for calculating total values and averages in numerical data sets, providing valuable insights into overall trends and patterns.

## **The MIN( ) and MAX( ) functions**

The MIN( ) and MAX( ) functions help identify the smallest and largest values within a dataset, facilitating data analysis and decision-making processes.

## **The SUBSTRING\_INDEX( ) function**

The SUBSTRING\_INDEX( ) function is instrumental in extracting substrings based on specified delimiters, enabling efficient string manipulation and data extraction tasks.

## **The LENGTH( ) and TRIM( ) functions**

The LENGTH( ) and TRIM( ) functions are useful for determining string lengths and removing leading or trailing spaces, ensuring data cleanliness and accuracy.

## **The DISTINCT( ) function**

The DISTINCT( ) function is crucial for filtering duplicate records, improving data quality, and ensuring accurate and relevant query results.



**Cristhian Becerra**



[cristhian-becerra-espinoza](#)



+51 951 634 354



[cristhianbecerra99@gmail.com](mailto:cristhianbecerra99@gmail.com)



Lima, Peru

