

GUIA PARA LA CONFIGURAR LA BASE DE DATOS AWS/RDS/MYSQL

La siguiente base de datos contiene aproximadamente 300,000 registros de empleados con 2.8 millones de ingresos salariales. Los datos de exportación son 167 MB, que no son enormes, pero lo suficientemente pesados como para no ser triviales para las pruebas.

Los datos fueron generados, y como tal hay inconsistencias y problemas sutiles. En lugar de eliminarlos, los autores de este conjunto de datos decidieron dejar los contenidos intactos y utilizar estos problemas como ejercicios de limpieza de datos.

Prerrequisitos:

Verifique que el usuario con el que va a ejecutar el procedimiento de instalación tenga los siguiente privilegios en la instancia AWS/RDS/MySQL:

```
SELECT, INSERT, UPDATE, DELETE,  
CREATE, DROP, RELOAD, REFERENCES,  
INDEX, ALTER, SHOW DATABASES,  
CREATE TEMPORARY TABLES,  
LOCK TABLES, EXECUTE, CREATE VIEW
```

Instalación:

1. Descargue el repositorio de la siguiente dirección:

https://github.com/datacharmer/test_db

2. Cargue el archivo comprimido en bucket s3 que ha seleccionado para almacenar los recursos del proyecto: `s3://resources/`
3. Descargue el archivo comprimido test-db-master.zip en la instancia EC2 desde la ruta s3 de recursos del proyecto, utilice el siguiente comando:

```
aws s3 cp s3://resources/test-db-master.zip ./test-db-master.zip
```

4. Descomprima el archivo test-db-master.zip y ubíquese en el directorio generado.
5. Instale la base de datos ejecutando el siguiente comando:

```
mysql -h <endpoint> -P 3306 -u <mymasteruser> -p < employees.sql
```

6. Verifique la instalación de la base de datos ejecutando el siguiente comando:

```
mysql -h <endpoint> -P 3306 -u <mymasteruser> -u admin -p <
test_employees_md5.sql
```

7. Debería obtener un resultado similar al siguiente:

```
mysql -h <endpoint> -P 3306 -u <mymasteruser> -u admin -p <
test_employees_md5.sql
```

```
+-----+
| INFO |
+-----+
| TESTING INSTALLATION |
+-----+

+-----+-----+-----+
| table_name | expected_records | expected_crc |
+-----+-----+-----+
| employees | 300024 | 4ec56ab5ba37218d187cf6ab09ce1aa1 |
| departments | 9 | d1af5e170d2d1591d776d5638d71fc5f |
| dept_manager | 24 | 8720e2f0853ac9096b689c14664f847e |
| dept_emp | 331603 | ccf6fe516f990bdaa49713fc478701b7 |
| titles | 443308 | bfa016c472df68e70a03facafa1bc0a8 |
| salaries | 2844047 | fd220654e95aea1b169624ffe3fca934 |
+-----+-----+-----+

+-----+-----+-----+
| table_name | found_records | found_crc |
+-----+-----+-----+
| employees | 300024 | 4ec56ab5ba37218d187cf6ab09ce1aa1 |
| departments | 9 | d1af5e170d2d1591d776d5638d71fc5f |
| dept_manager | 24 | 8720e2f0853ac9096b689c14664f847e |
| dept_emp | 331603 | ccf6fe516f990bdaa49713fc478701b7 |
| titles | 443308 | bfa016c472df68e70a03facafa1bc0a8 |
| salaries | 2844047 | fd220654e95aea1b169624ffe3fca934 |
+-----+-----+-----+

+-----+-----+-----+
| table_name | records_match | crc_match |
+-----+-----+-----+
| employees | OK | ok |
| departments | OK | ok |
| dept_manager | OK | ok |
| dept_emp | OK | ok |
| titles | OK | ok |
| salaries | OK | ok |
+-----+-----+-----+
```