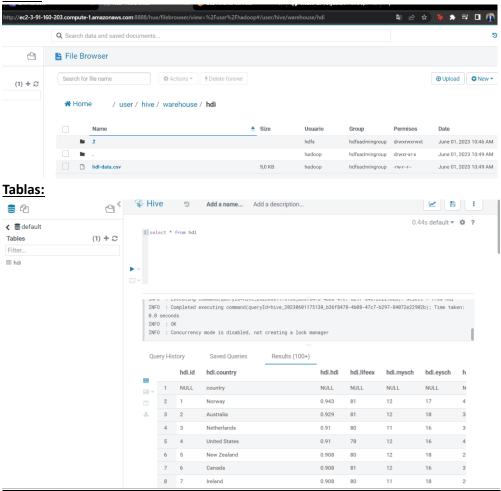
## 1. EMR HIVE / SPARK SQL

- a. Realice las actividades de hive con
  - i. tablas nativas

### Datos:

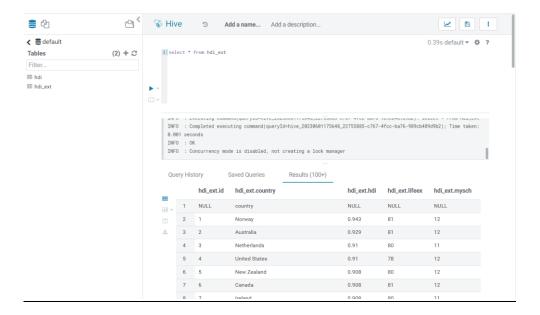


ii. tablas externas

#### Datos:

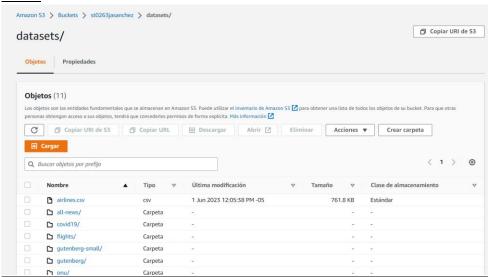


**Tablas:** 

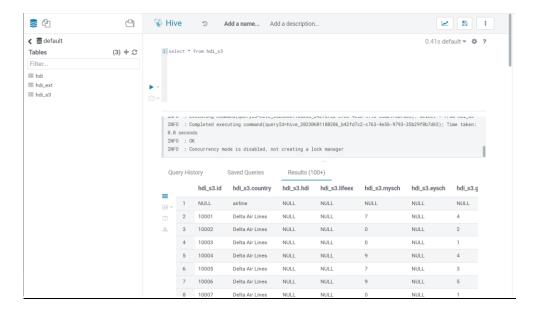


### iii. tablas desde S3

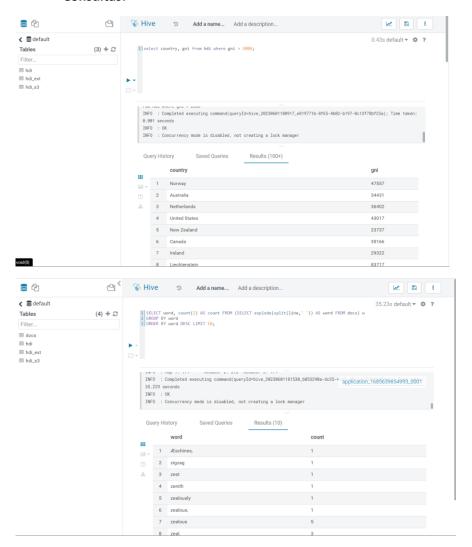
### Datos:

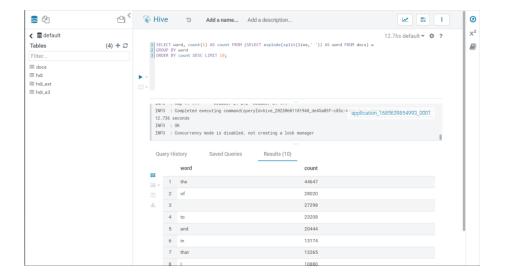


Tablas:

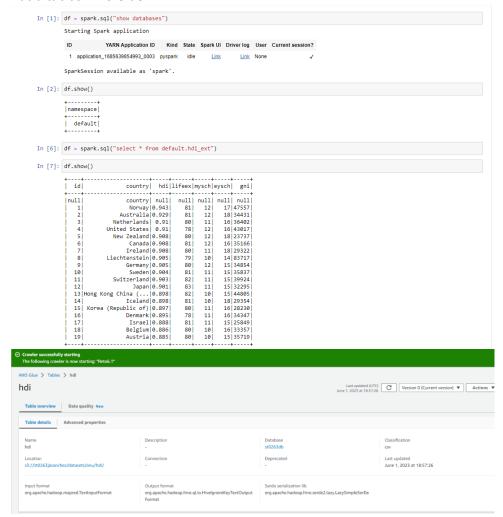


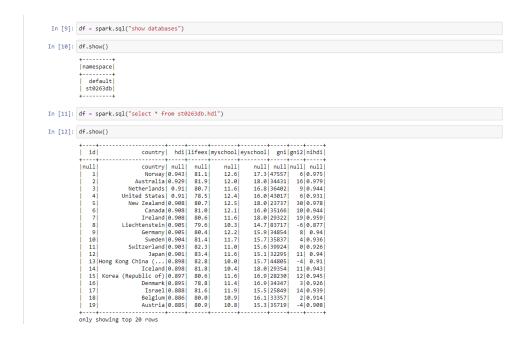
### Consultas:





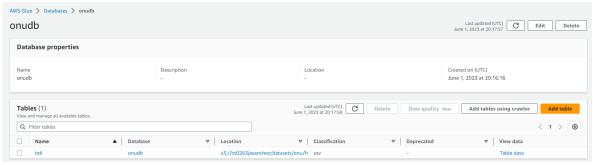
b. Realice la ejecución de consultas SQL desde Apache SparkSQL hacia tablas Hive y hacia tablas AWS Glue.





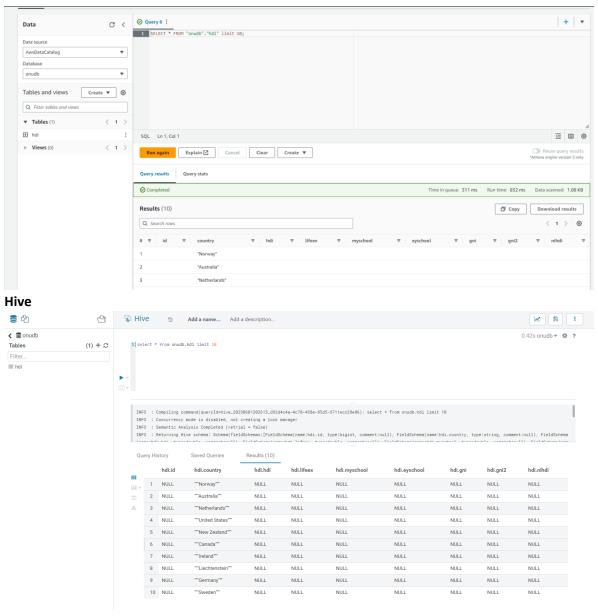
### 2. IMPLEMENTACIÓN DE UN DATA WAREHOUSE SENCILLO CON AWS ATHENA

## **Catalogando datos:**



# **Consultas:**

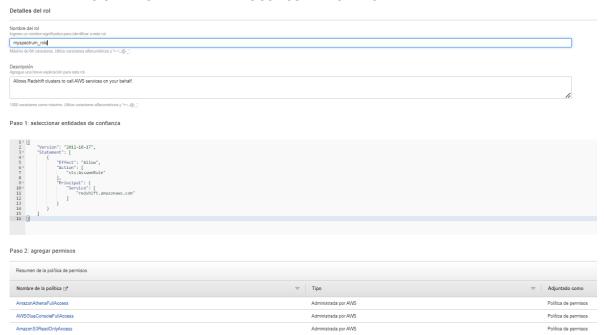
**Athena** 



**Spark** 

```
In [2]: df = spark.sql("show databases")
 In [3]: df.show()
           namespace
              default
 In [9]: df = spark.sql("select * from onudb.hdi")
In [10]: df.show()
                      | id|
           Inul11
                                  Norway 0.943
                                                     81.1
                                                                12.6
                                                                           17.3 | 47557 | 6 | 0.975 | 18.0 | 34431 | 16 | 0.979
                               Australia 0.929
                                                    81.9
                        Netherlands | 0.91 | 80.7 |
United States | 0.91 | 78.5 |
                                                                11.6
                                                                           16.8|36402|
16.0|43017|
                                                                                            9|0.944|
6|0.931|
                          New Zealand 0.908
Canada 0.908
                                                     80.7
81.0
                                                                12.5
                                                                           18.0 | 23737 | 30 | 0.978 |
16.0 | 35166 | 10 | 0.944 |
               6
                                                                11.6
                                 Ireland 0.908
                                                     80.6
                                                                           18.0 29322 19 0.959
                        Liechtenstein 0.905
                                                     79.6
                                                                           14.7 83717
                                                                                           -6 0.877
                             Germany 0.905
Sweden 0.904
                                                                12.2
11.7
                                                                           15.9 34854
15.7 35837
                                                                                            8 0.94
4 0.936
               9
                                                     80.4
                            Switzerland 0.903
              11
                                                     82.3
                                                                11.0
                                                                           15.6 39924
                                                                                            0 0.926
              12| Japan|0.901|
13|Hong Kong China (...|0.898|
14| Iceland |0.898|
15| Korea (Republic of)|0.897|
                                                                           15.1|32295| 11| 0.94|
15.7|44805| -4| 0.91|
                                                     82.8
                                                                10.0
                                                                10.4
                                                                           18.0 29354 11 0.943
                                                                11.6
                                                                           16.9 28230 12 0.945
                                                     80.6
                                                                11.4
                                                                           16.9|34347| 3|0.926|
15.5|25849| 14|0.939|
                                  Denmark 0.895
                                                     78.8
              17
                                   Israel 0.888 81.6
                                  Belgium|0.886| 80.0|
Austria|0.885| 80.9|
                                                                10.9
                                                                           16.1|33357| 2|0.914|
15.3|35719| -4|0.908|
              18
              19
           only showing top 20 rows
```

#### 3. IMPLEMENATCIÓN DE UN DATA WAREHOUSE CON AWS REDSHIFT



| ② No se pudo crear el rol myspectrum_role. User, am aws sts::726662720187.assumed-role/voclabs/user1891167=jasanchez@eafit.edu.co is not authorized to perform: iam.CreateRole on resource: am aws iam::725662720187.role/myspectrum_role because no identity-based policy allows the tam: CreateRole action |   |
|--|---|
| IAM > Roles > Crearrol   |   |
| Paso 1<br>Seleccionar entidad de<br>confianza  | Asignar nombre, revisar y crear   |
| Paso 2<br>Agregar permisos<br>Paso 3<br>Asignar nombre, revisar y<br>crear   | Detailes del rol  |
|  | Nombre del rol Ingrese un nombre significativo para identificar a este rol. |
|  | myspectrum_role   |
|  | Máximo de 64 caracteres. Utilice caracteres alfanuméricos y "+ =, @".       |
|  | Descripción Agregue una breve explicación para este rol.                    |
|  | Allows Redshift clusters to call AWS services on your behalf.               |