Documentación del Laboratorio 0: AWS-EMR

Nombre: Liz Oriana Rodrigues Cruz
Fecha: 28 de Mayo

Prerrequisitos:

1. Clonar el Cluster del Lab3-0

		Nombre	ID	Estado
☐ ▶ ○		Lab3-0-Cluster	j-1ESO8DSDT4FTE	Comenzando
☐ ▶		Lab3-0-Cluster	j-3IR367QG51LBJ	Terminado Solicitud del usuario

Filter: Todos los clústeres ▼ Filtrar clústeres... Clústeres: 6 (todos cargados) ↻

2. Revisamos los datos del S3

```
[hadoop@ip-172-31-2-253 ~]$ hdfs dfs -ls /user/hadoop/datasets/gutenberg-small/
Found 1 items
-rw-r--r--   1 hadoop hdfsadmingroup    3419810 2022-05-28 02:18 /user/hadoop/datasets/gutenberg-small/gutenberg-small-e
s.zip
[hadoop@ip-172-31-2-253~]$
```

Creamos las carpetas y los datos del EMR

⌂ Inicio / user / hadoop / datasets / onu / **export-data.csv**

```
country,export
Afghanistan,15.48717104
Albania,29.77230976
Algeria,30.83040547
American Samoa,
Andorra,
Angola,56.83588283
Antigua and Barbuda,44.08266774
Argentina,21.70646878
Armenia,20.5836109
Aruba,
Australia,19.78024202
Austria,53.97135556
Azerbaijan,55.12575386
"Bahamas, The",42.02174624
Bahrain,
Bangladesh,18.40616841
Barbados,47.34395912
Belarus,54.6236092
Belgium,80.00874773
Belize,
Benin,14.12394984
```

Datos ERM en consola

```
[hadoop@ip-172-31-2-253 ~]$ hdfs dfs -ls /user/hadoop/datasets/
Found 2 items
drwxr-xr-x   - hadoop hdfsadmingroup          0 2022/05/28 2:18  /user/hadoop/datasets/gutenberg-smal
drwxr-xr-x   - hadoop hdfsadmingroup          0 2022/05/28 2:23  /user/hadoop/datasets/onu
[hadoop@ip-172-31-2-253 ~]$
[hadoop@ip-172-31-2-253 ~]$
```