



AWS CLI.

Como resultado de la práctica has de entregar un archivo en formato “pdf” con breves explicaciones y capturas de pantalla de los pasos principales que has dado para realizar los ejercicios.

CONTENIDO

APARTADO A

INTRODUCCIÓN

Investiga en Internet que comandos de AWS CLI hemos de utilizar para manejar *buckets* S3. Por ejemplo, en el siguiente enlace hay información de AWS sobre ello:

<https://docs.aws.amazon.com/cli/latest/userguide/cli-services-s3-commands.html>

Utilizando únicamente comando AWS CLI:

1. Crear tres *buckets* en la región por defecto (con nombres similares a: **primero**, **copia**, **sincro**).
2. Listar todos los cubos.
3. Ponte en una carpeta donde tengas varios archivos, mejor CSV's o JSON, Sube al menos dos archivos de esa carpeta al *bucket* **primero**.
4. Copia los archivos del *bucket* **primero** a **copia** de uno en uno.
5. Mover todos los objetos de **primero** a **copia** dentro de la carpeta **datos** de forma recursiva.
6. Listar los objetos de un **copia/datos**.
7. Listar recursivamente todos los objetos el cubo **copia**.
8. Lista recursivamente los datos del cubo en formato *legible por humanos*. ¿Qué significa eso?
9. Mostrar la información resumida, incluido el número de objetos y el tamaño total del cubo **copia**.
10. Muévete a otra carpeta en local y descarga todos los objetos de **copia** a ella. Borra en local los archivos descargados.
11. Sincronizar el contenido del un cubo **copia** con el cubo **sincro**.
12. Sincronizar el contenido del un cubo **copia** con la carpeta local.
13. Haz al menos tres consultas SQL cualesquiera sobre algunos de los archivos CSV o JSON que hayas subido.
14. Elimina los objetos de **primero** de uno en uno.
15. Elimina todos los objetos de **copia** de forma recursiva.
16. Elimina los cubos **primero** y **copia**.
17. Fuerza la eliminación del cubo **sincro** sin vaciarlo previamente.

CONTENIDO

APARTADO B



INTRODUCCIÓN

Para poder trasladar la información de HDFS a S3 haremos uso de la herramienta [S3DistCp \(s3-dist-cp\)](#). Analiza su documentación.

De HDFS a S3

- 1.- Crea un nuevo *bucket* en S3 y utilizando el comando anterior, copia el archivo `ventas.csv` que tenemos en HDFS y obtuviste en la práctica 7.1 a él.
- 2.- (INVESTIGA) ¿Cómo podrás desde HIVE lanzar las mismas consultas que hicimos en los apartados D y E de la práctica 7.1, pero en esta ocasión sobre ficheros en S3?