

Actividad:	Athena							
Ciclo:	IABD	Modulo:	SBD	Curso:		Agrupación:	1	
Alumno/a:							Grupo:	
CIPFP www.fpmislata.com								

Consultas sobre S3 con Athena

Supongamos que tienes un conjunto de datos en Amazon S3 que registra las ventas de una tienda en diferentes regiones. Los datos están almacenados en formato CSV y contienen las columnas: fecha, producto, cantidad vendida y región.

Pasos:

- **Acceso a Athena:**
 - Inicia sesión en la consola de AWS y accede a Amazon Athena.
- **Creación de Base de Datos y Tabla:**
 - Crea una nueva base de datos llamada **ventas_db**.
 - Define una tabla **ventas** con las siguientes columnas: **fecha (string)**, **producto (string)**, **cantidad (int)**, **region (string)**. Asegúrate de especificar el formato de datos adecuado.
- **Carga de Datos:**
 - Simula la carga de datos ejecutando una consulta para agregar los datos almacenados en tu bucket de S3 al esquema creado. Utiliza la ubicación correcta del archivo CSV.

```
CREATE EXTERNAL TABLE IF NOT EXISTS `empresa`.`ventas` (
  fecha string,
  producto string,
  cantidad int,
  region string
)
ROW FORMAT SERDE
'org.apache.hadoop.hive.serde2.OpenCSVSerde'
WITH SERDEPROPERTIES (
  'separatorChar' = ',',
  'quoteChar' = '"',
  'escapeChar' = '\\'
)
LOCATION 's3://tu-bucket';
```

- **Consulta Básica:**
 - Realiza una consulta para mostrar las primeras 10 filas de datos de ventas.

Actividad:	Athena							
Ciclo:	IABD	Modulo:	SBD	Curso:		Agrupación:	1	
Alumno/a:							Grupo:	
CIPFP www.fpmislata.com								

- **Consultas SQL:**

- Escribe consultas SQL para responder a las siguientes preguntas:

- ¿Cuál es el total de ventas por región?
- ¿Cuál es el producto más vendido en total?

- **Búsqueda Avanzada:**

- Realiza una consulta que muestre la cantidad vendida por producto y región, pero solo para productos que hayan vendido más de 100 unidades en total.