

Actividad:	Glue_2						
Ciclo:	IABD	Modulo:	SBD	Curso:		Agrupación:	1
Alumno/a:						Grupo:	
CIPFP www.fpmislata.com							

Tarea. AWS Glue Studio

La práctica consiste en hacer una ETL con AWS Glue utilizando un Job Visual.

- Los datos de partida son un fichero en formato csv con datos de ciclistas.
- Cargaremos los datos en un bucket S3 que tendrá dos carpetas, una los datos de entrada y otro para los datos procesados.
- Crearemos un job que elimine los duplicados y haga un filtro y seleccione solo las filas con el campo Severity = "Grave".
- Crea una función lambda que ejecute el job cada vez que se suba un fichero .csv a la carpeta de datos de entrada.

Revisaremos como se monitorizan los jobs y las funciones lambda.

Ayuda: <https://aws-dojo.com/workshoplists/workshoplist21/>