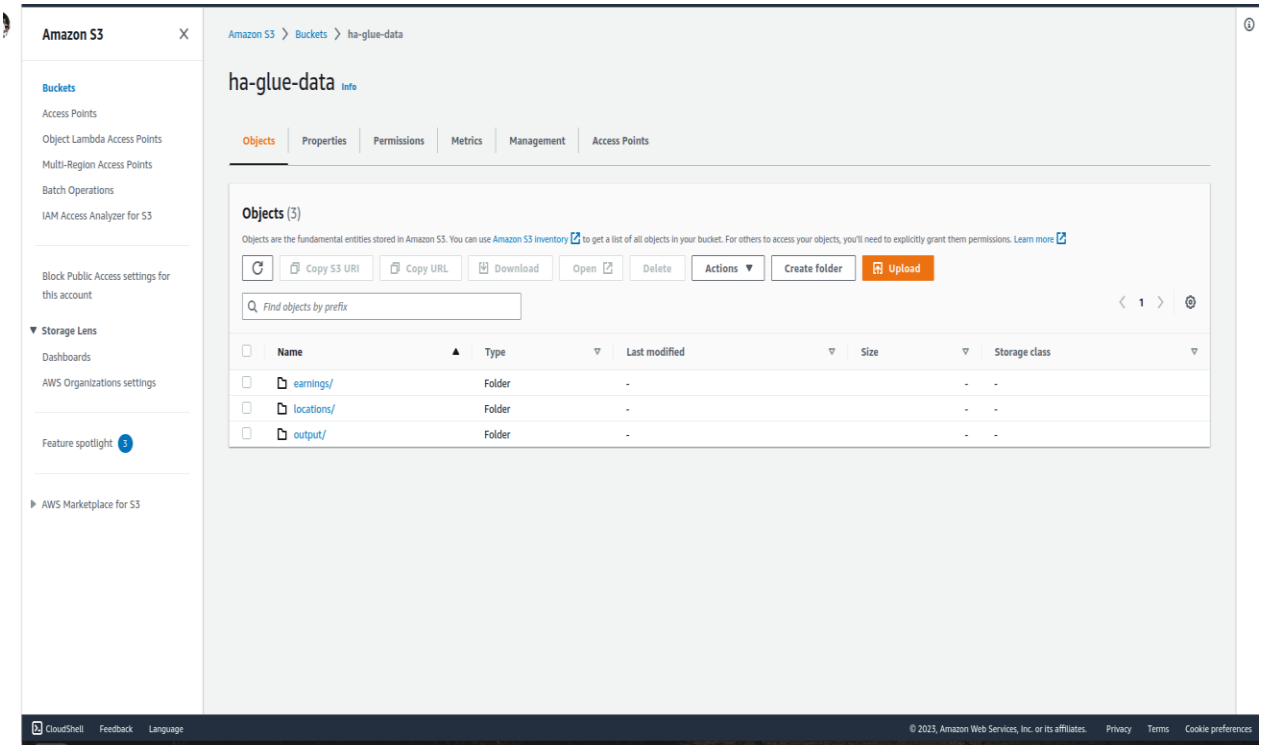
Members

Huzefa Anver (2303.KHI.DEG.002)

Syed Mohammad Anjil Hussain Rizvi (2303.KHI.DEG.031)

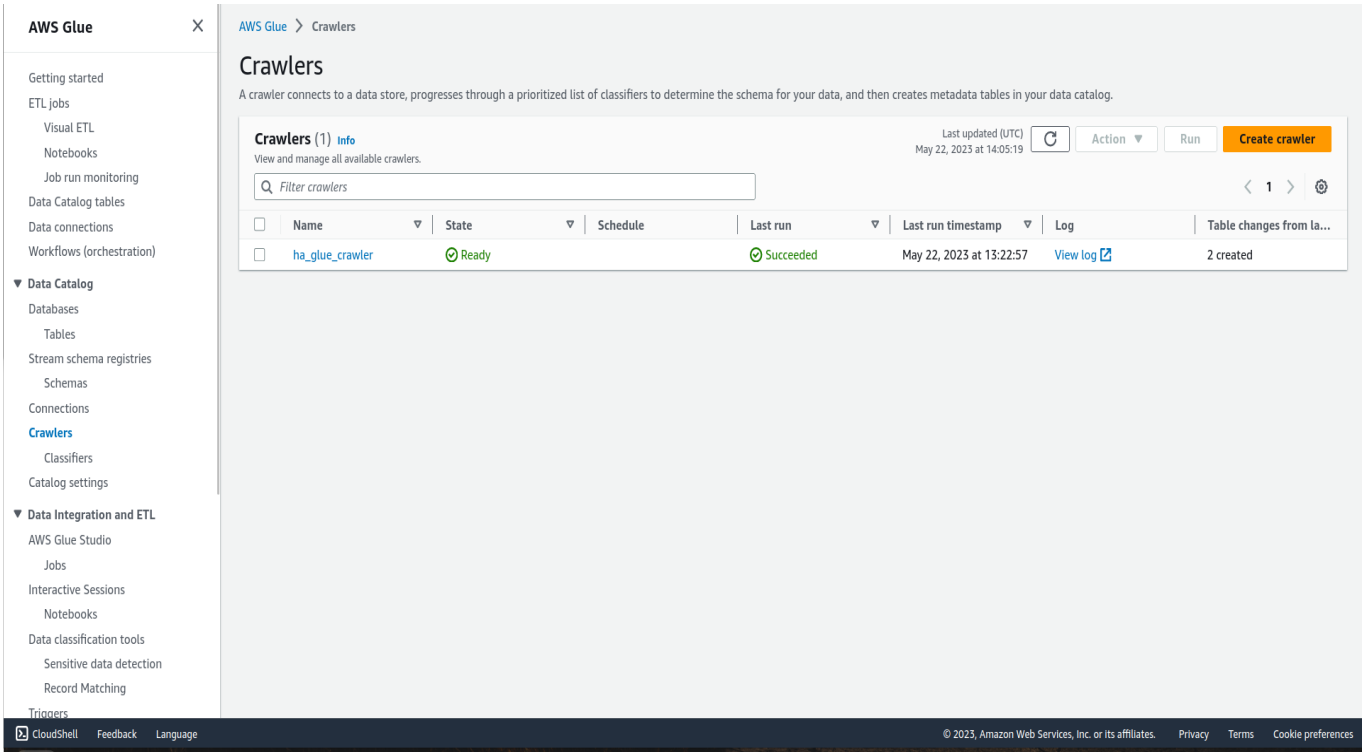
Ha-glue-data s3 bucket



This is our bucket containing 3 directories earnings and location containing our input data

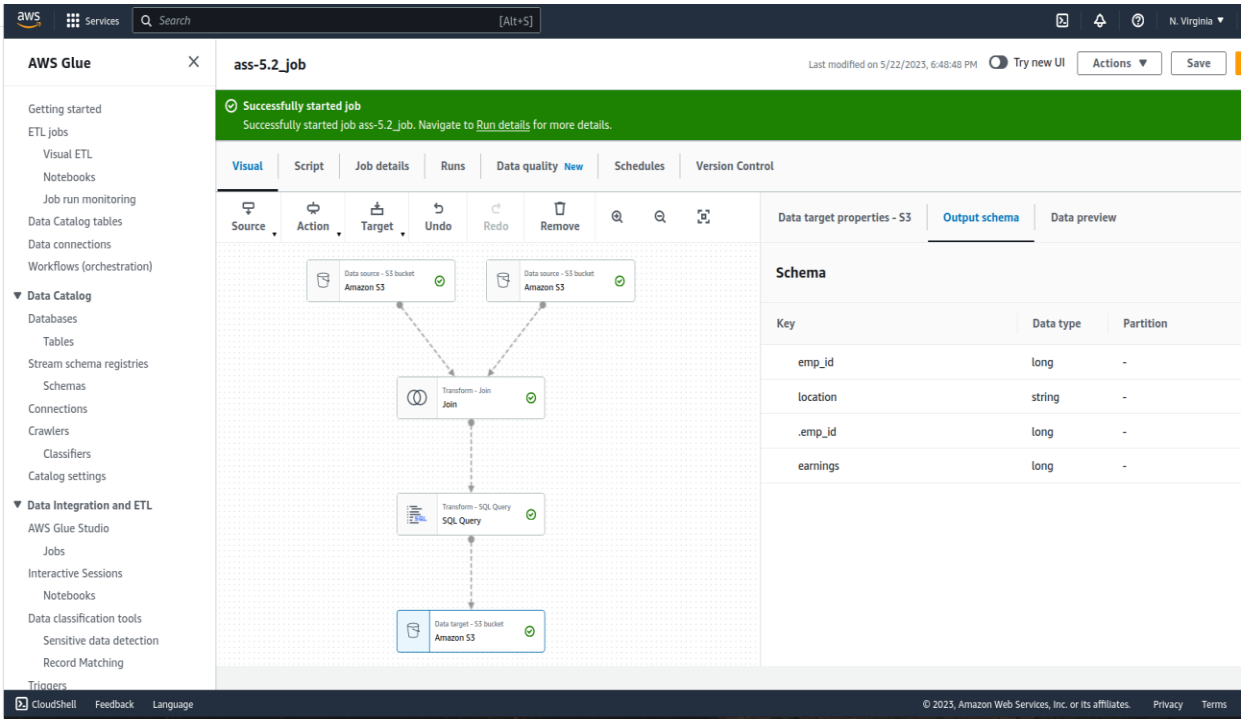
And output containing our output data after processing it using glue job crawlers.

AWS Glue Crawler



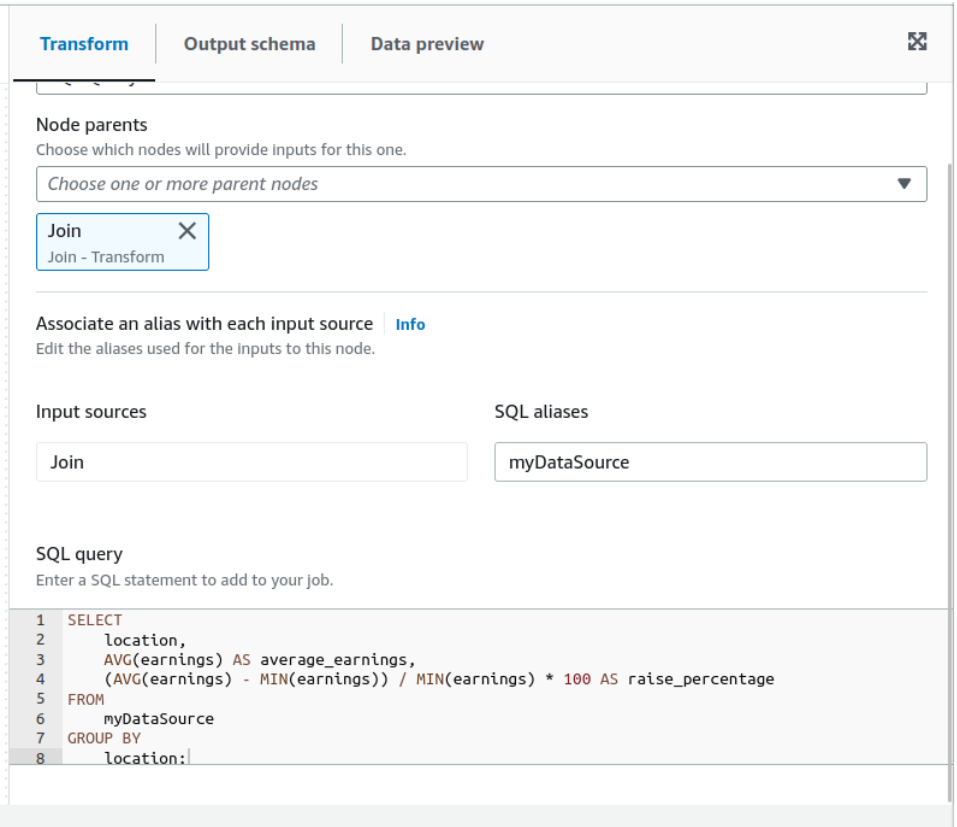
As we can see our crawler is up and running, which we will use to extract the data from s3 bucket from multiple directories earnings and locations, process it and load it into our output folder

Aws glue job



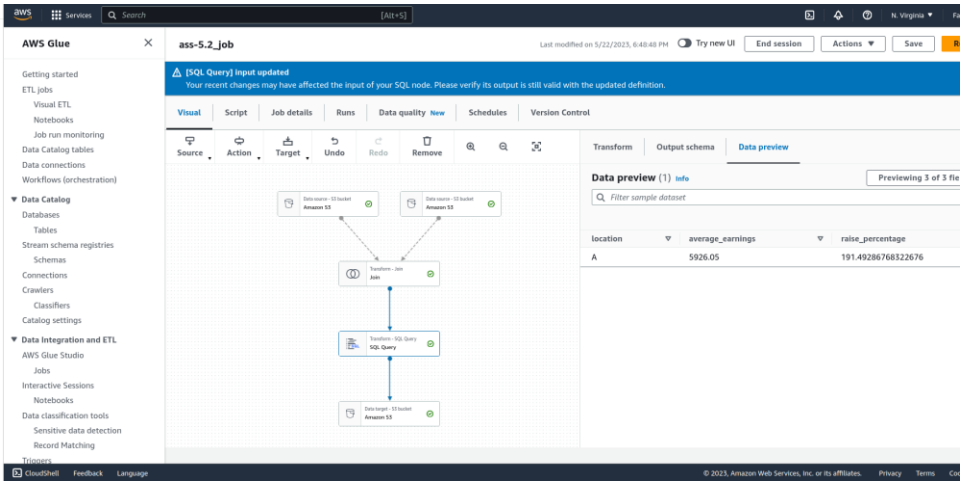
This is the overall glue job we created. As we can see the graph, we have 2 sources of our data, both directing towards the action we created join which will combine the data and then we will perform some queries from combine data and then the queried data will be saved in our output folder.

Performing query



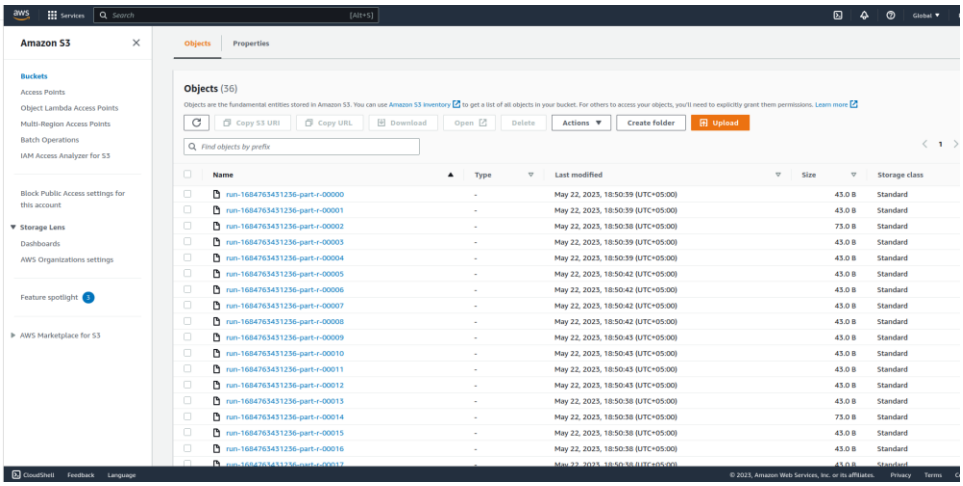
This is our query we performed after combining the data

Query preview



This is the preview we got after querying the data

Output directory



Finally, as we can see our data is loaded in our output folder which we created in our s3 bucket