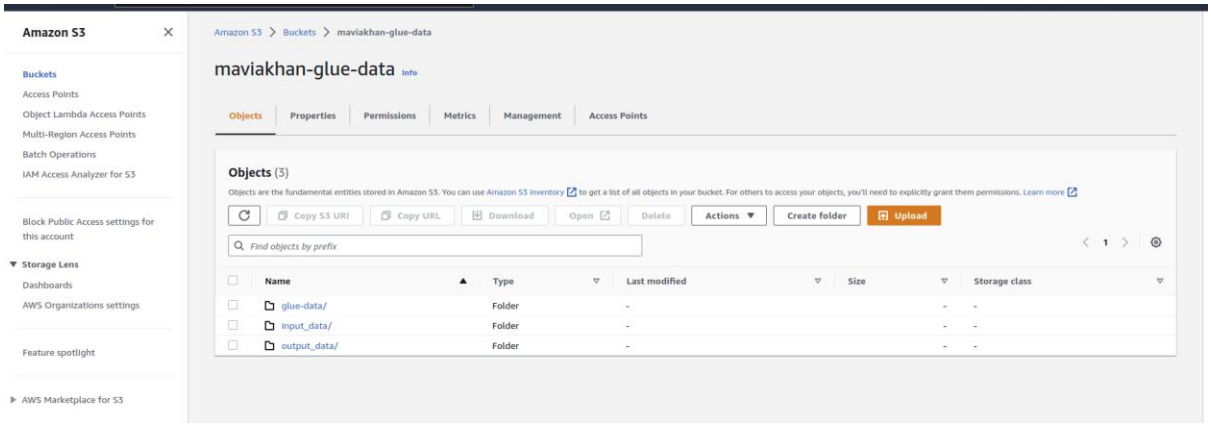
NAME: Mohammad Hussam (2303.KHI.DEG.020)
PARING WITH : MAVIA ALAM KHAN(2302.KHI.DEG.017)

ASSIGNMENT NO :5.2

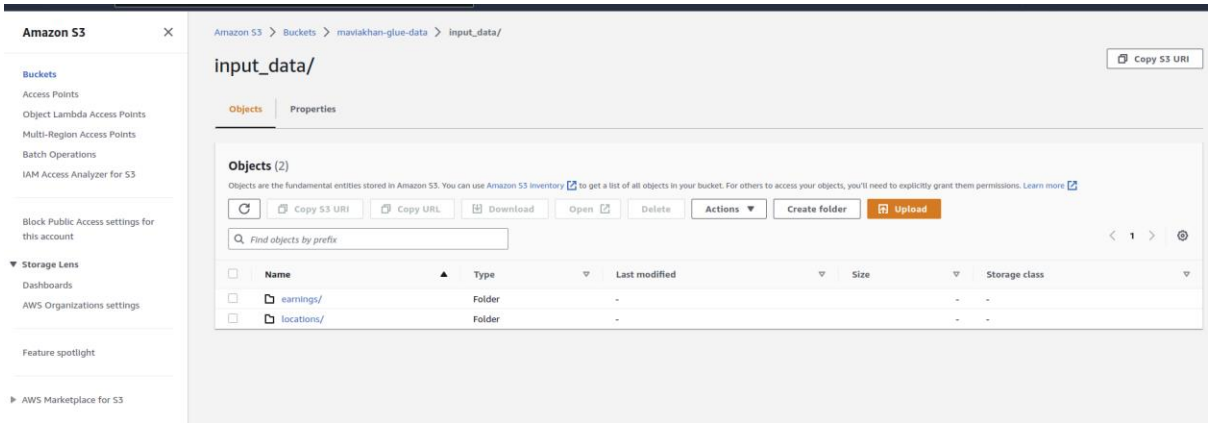
Using the salary CSV as a base, prepare a new data file with employees' office locations. Make sure there are 5-6 distinct locations that are shared between employees.
Create a Glue job that aggregates the data based on the office location to calculate
average salaries and raise percentages for these locations.

SOLUTION

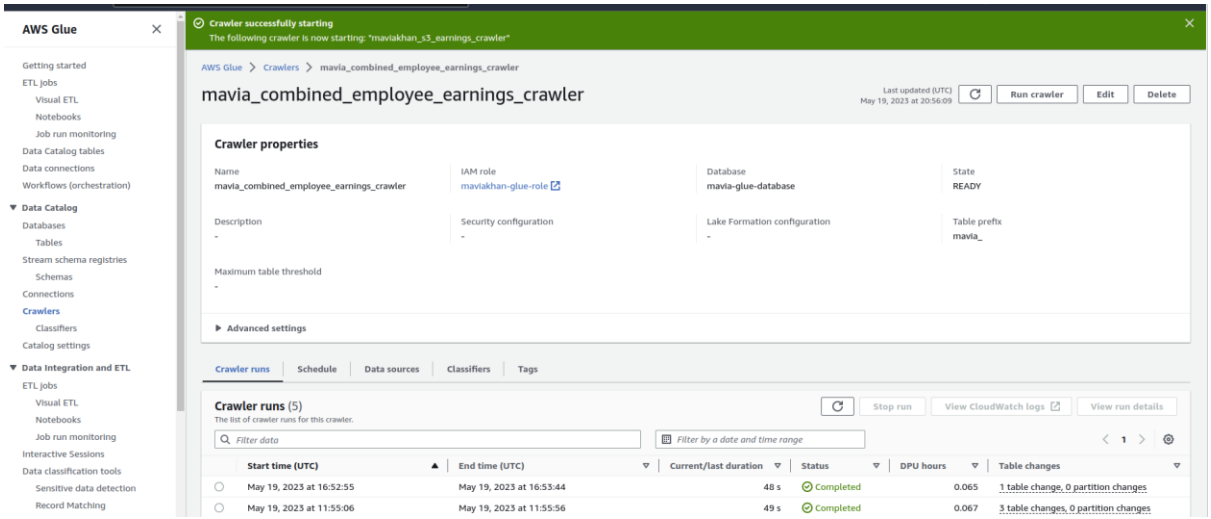
First create a folder in s3 bucket name (input and output)

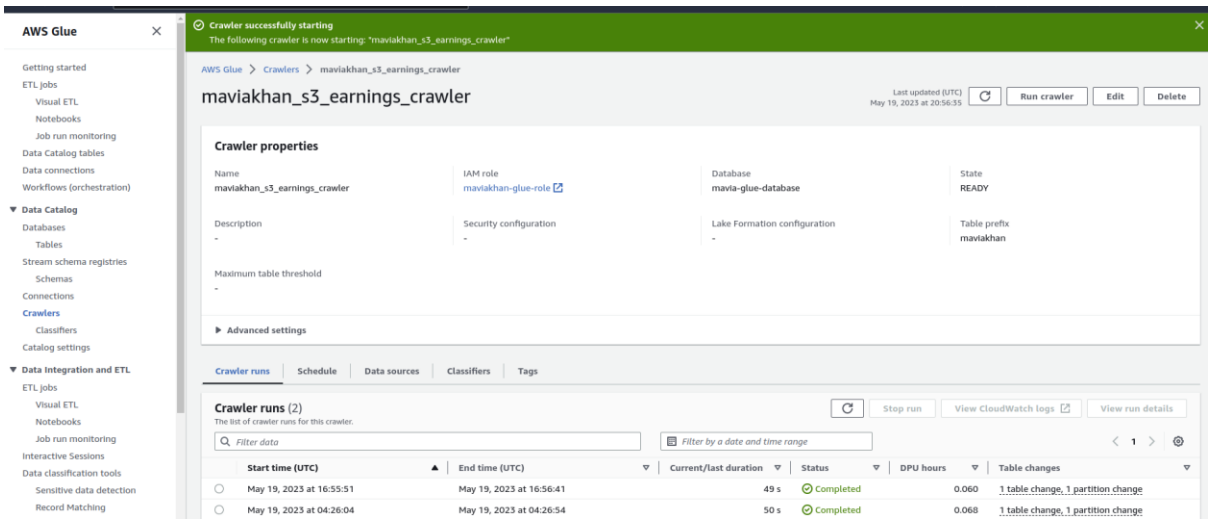


In input folder we store a two dataset earning.csv and location.csv

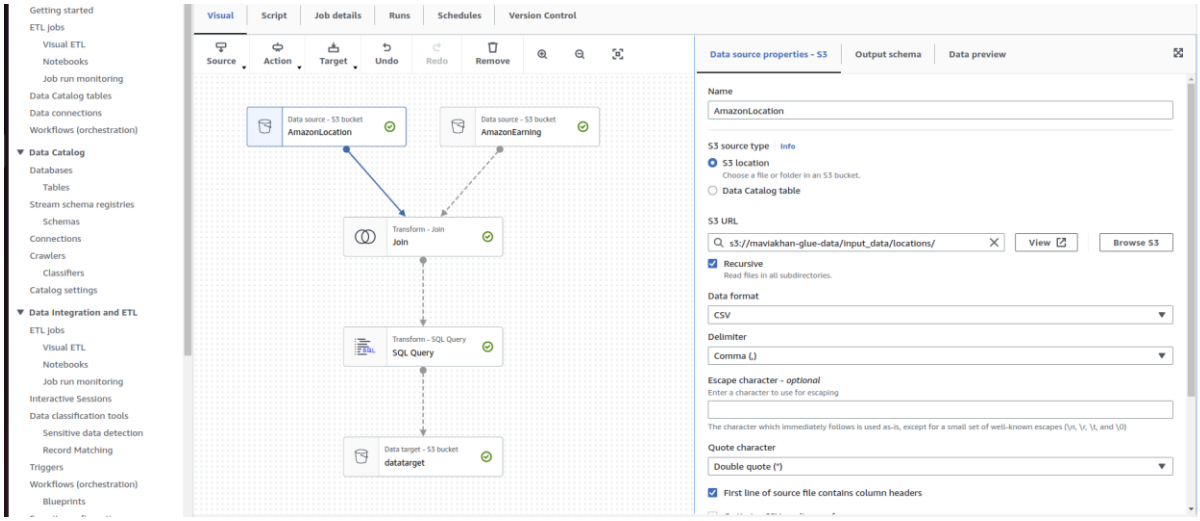


Now we create a crawler and extract the meta data

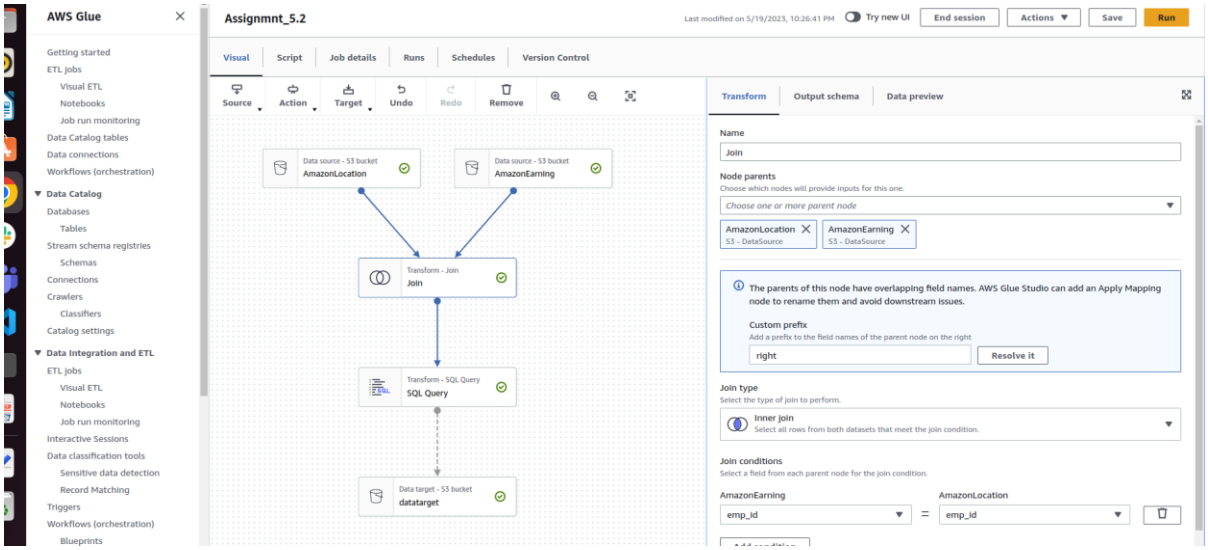


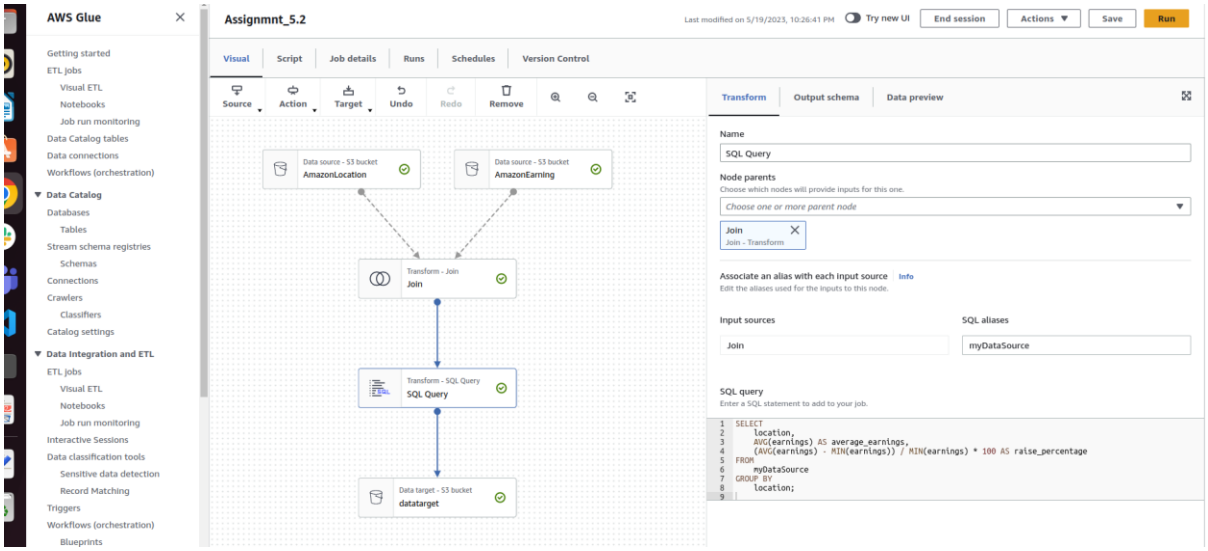


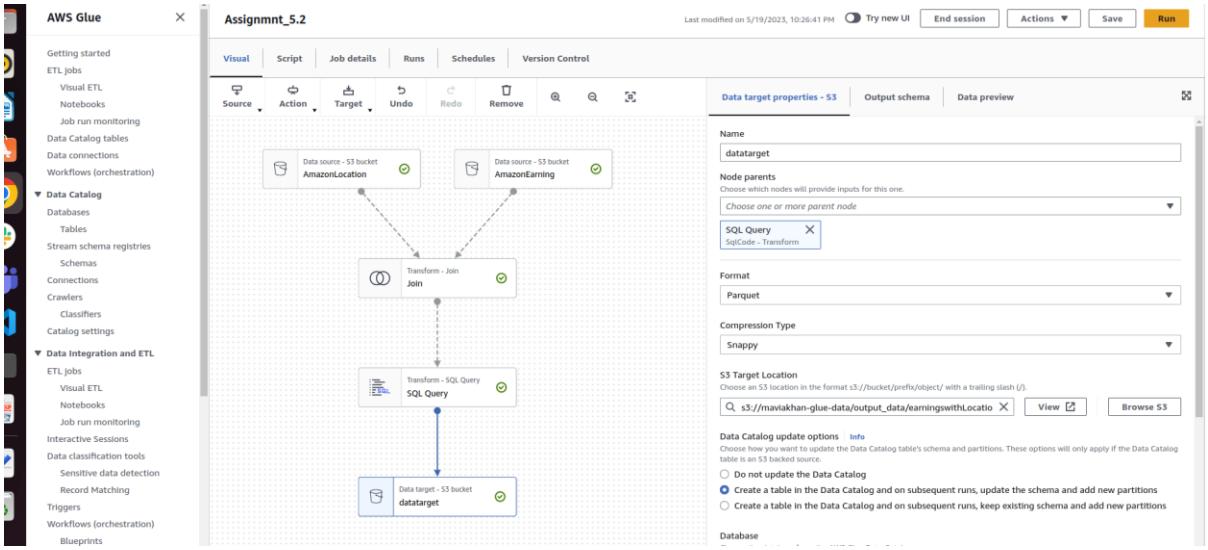
And after create the glue crawler we start creating the job



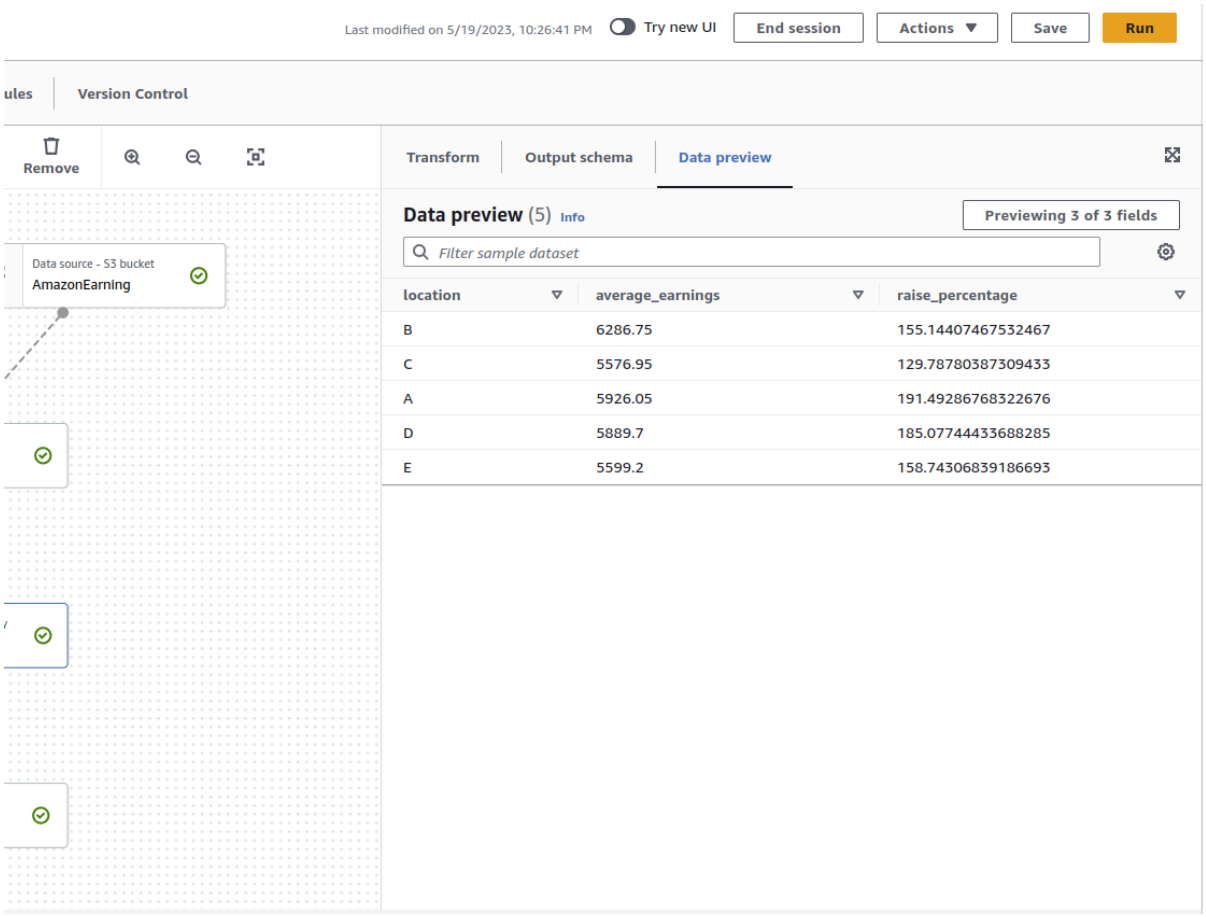
Now we create two s3 source one is employee earning data and location,we perform inner join on both data on emp_id and after that prepare for querying







Now here the queries based on the salaries and percentage of these locations.



Now finally we load the data and show in the output folder