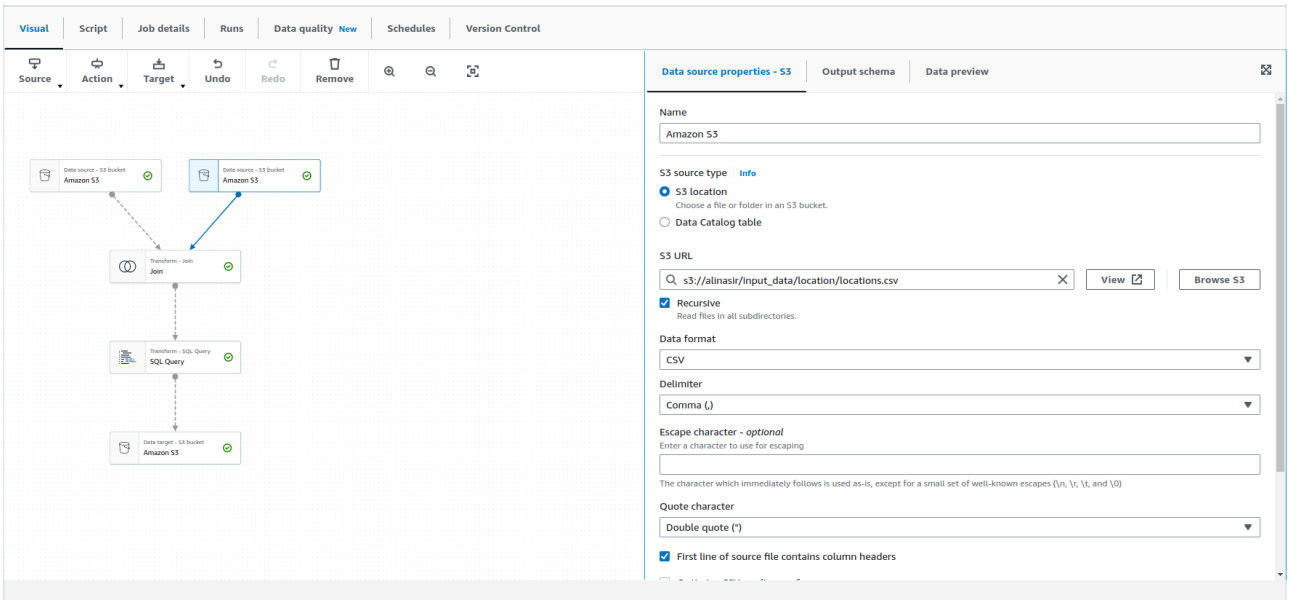
Assignment 5.2 Submitted by Ali Nasir(2023-KHI-DEG-032) & Saif ur Rehman(2023-KHI-DEG-007):

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.

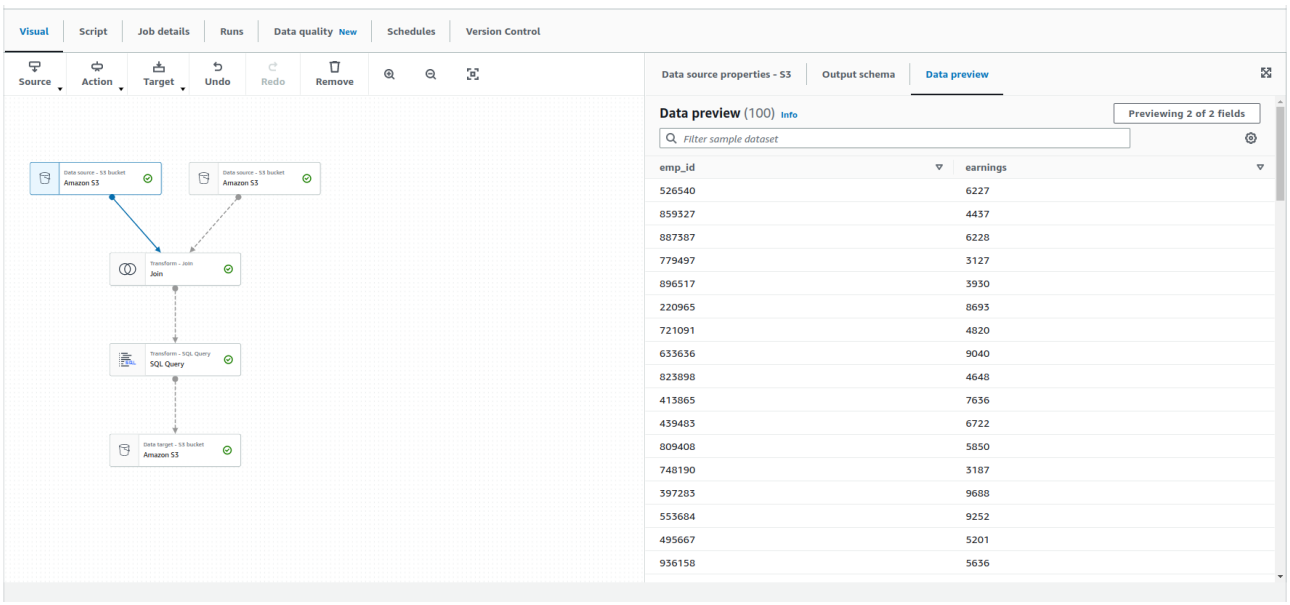
Solution:

Daily task:

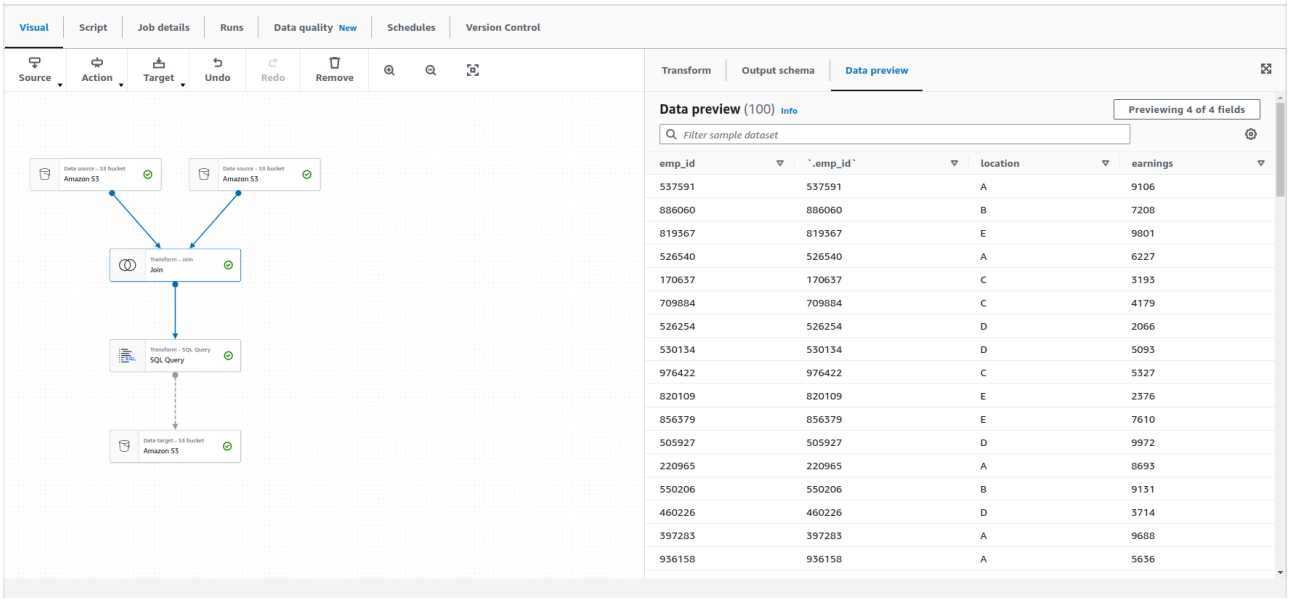
- We have used S3 bucket input/location to extract the csv



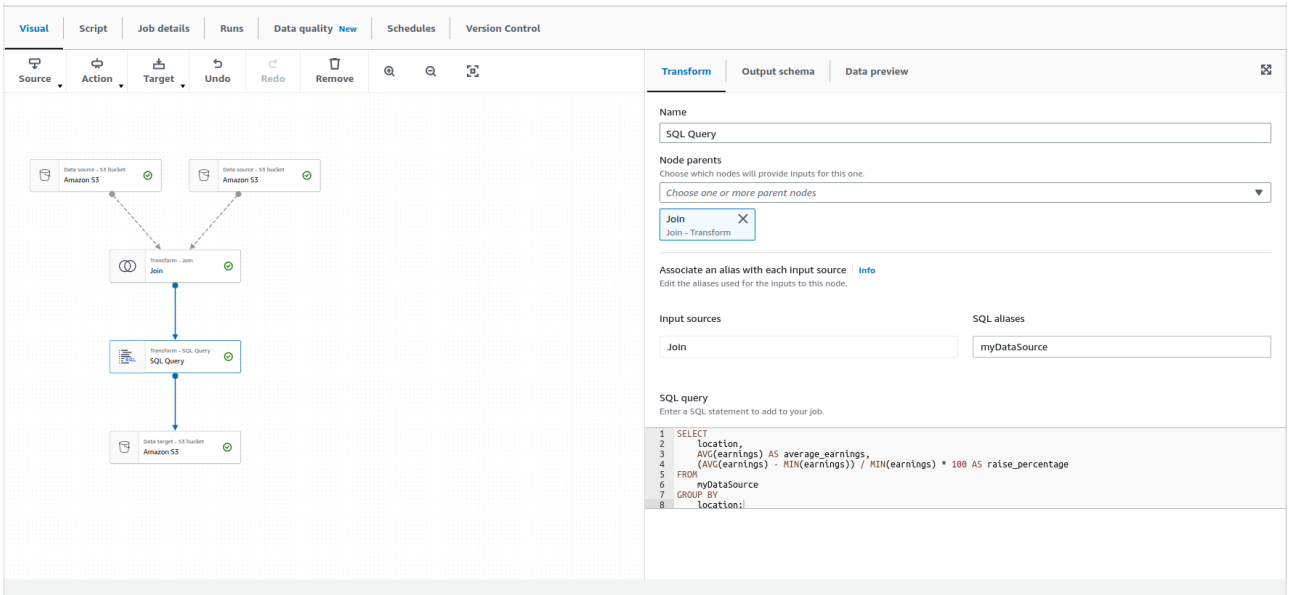
- We have used S3 bucket input/earning to extract the csv

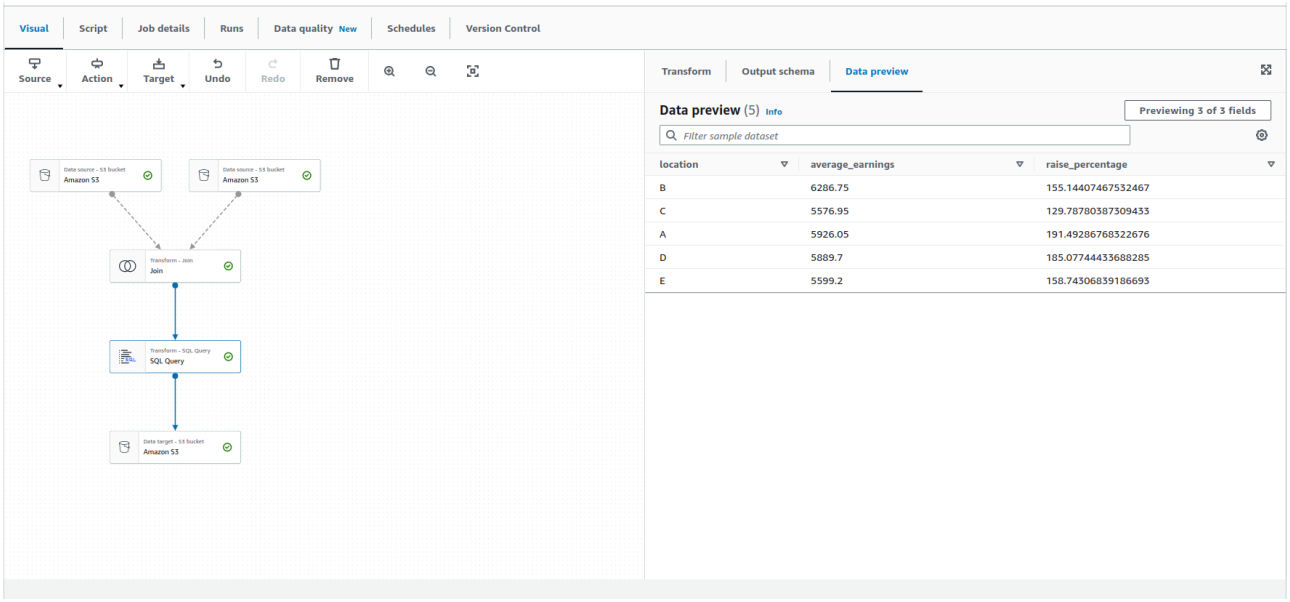


- Here we have implemented join on the both dataset



- Here is the query that we have implemented to find the aggregate data.





- Here is the output of the data from our crawler.