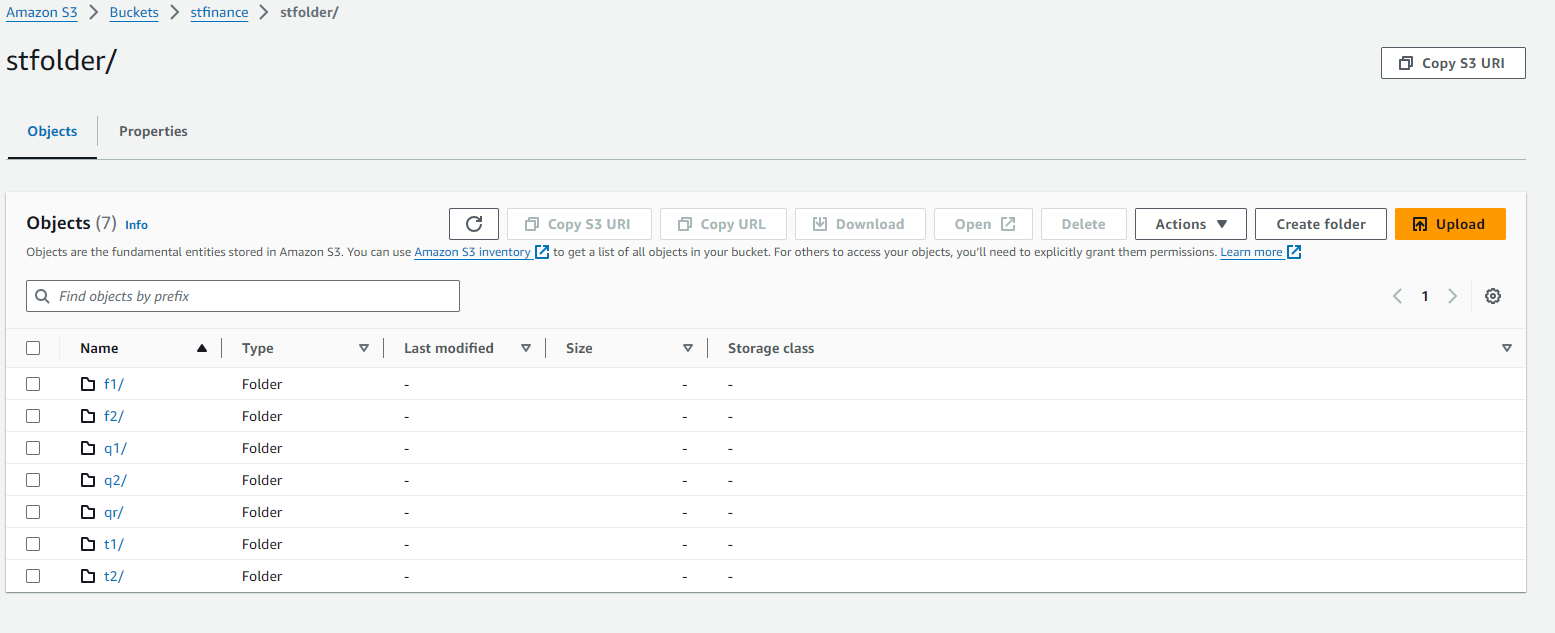
**Creating S3 Buket:**

* To begin, we establish a bucket called statefinance, a folder called stfolder, and two more folders called f1, f2, along with the csv files.

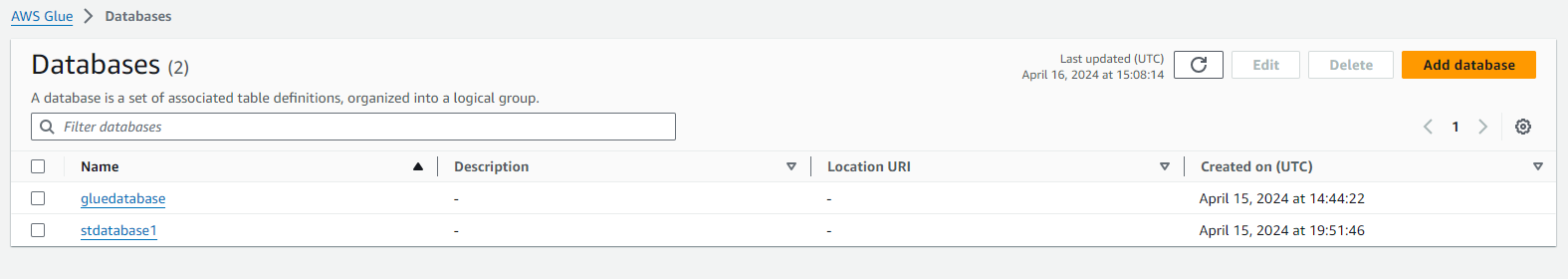


**IAM ROLE:**

* I have used the old IAM role

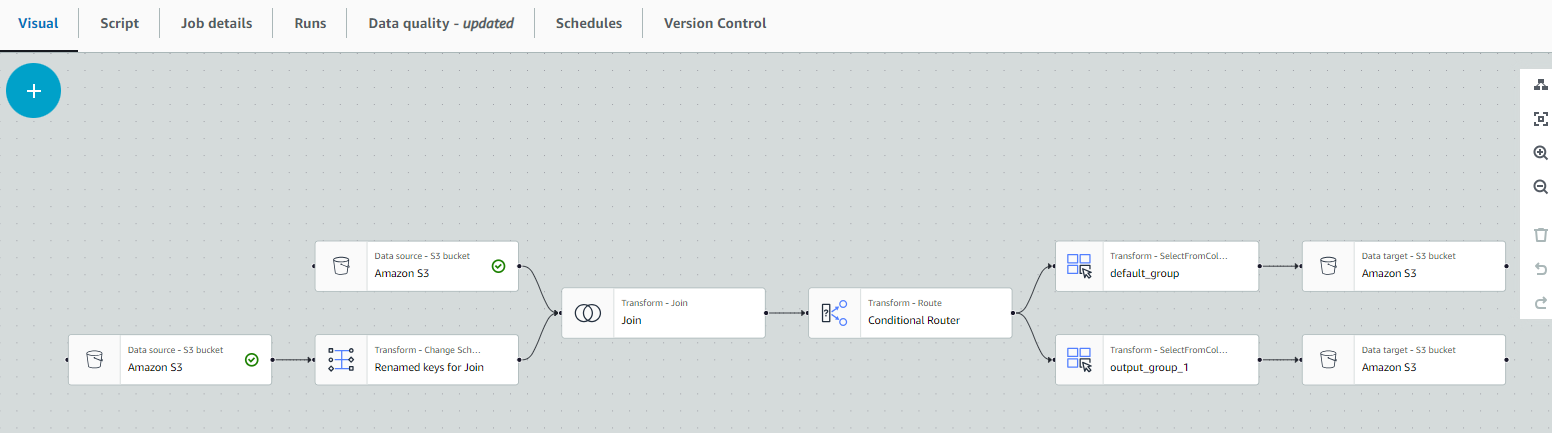
**AWS Glue**

* Since we have two input files, I have made two crawlers.
* After that, I made one database specifically for crawlers.

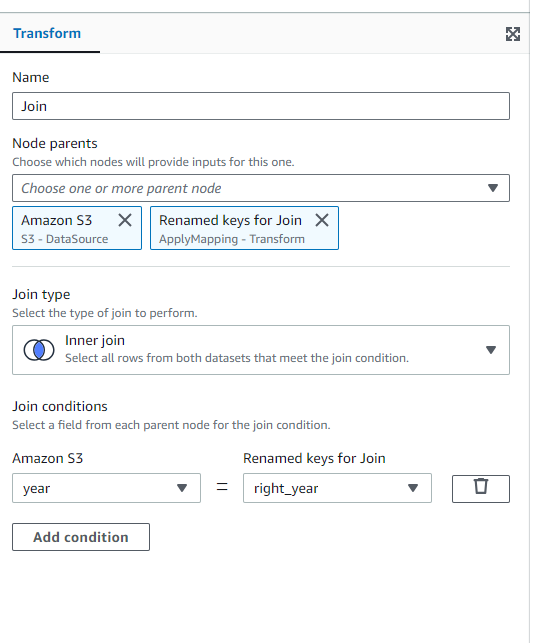


**ETL Visual:**

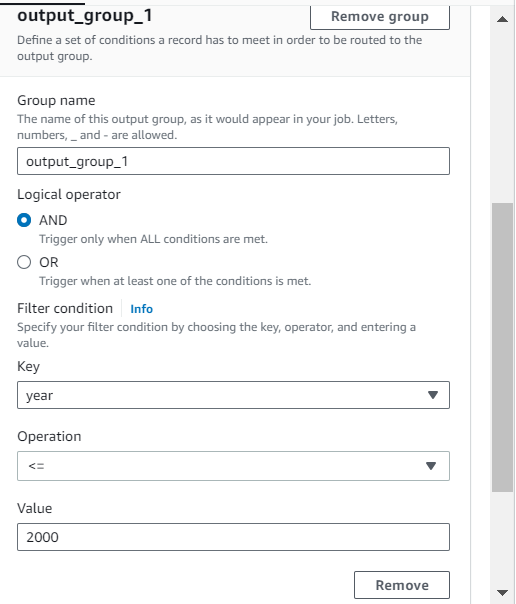
* The stages we are doing in ETL Visual are listed below.



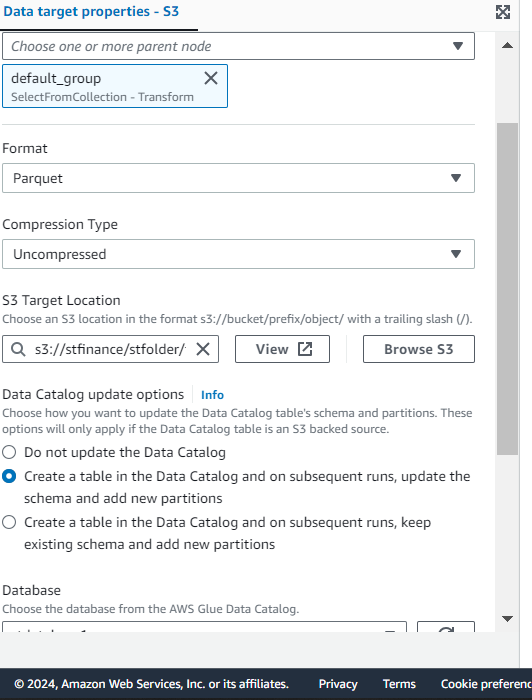
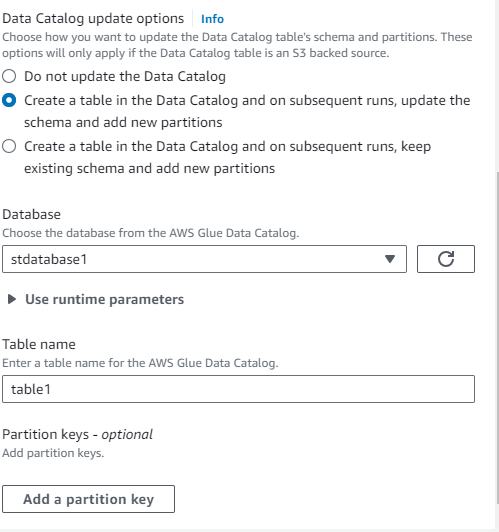
* As a starting point, we're adding two S3 buckets.
* To place the database and files that are already in the AWS Glue, I am utilizing the data catalog as the source in both AWS Buckets.
* After adding the join, we will notice an overlapping issue, which we will fix by adding a right join.



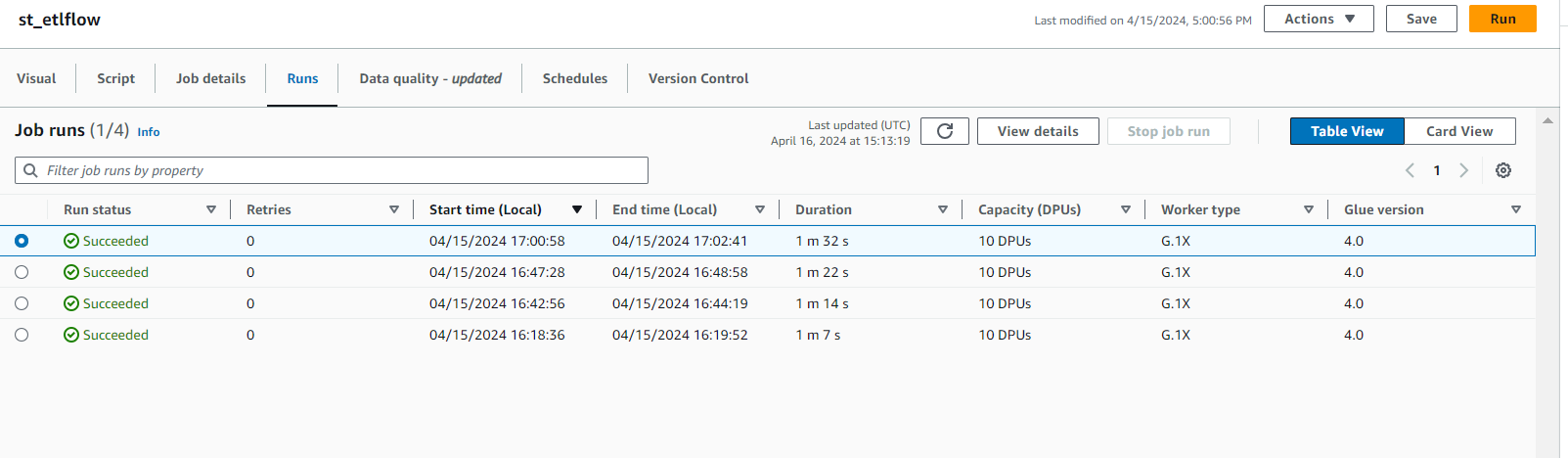
* I used an inner join in the joins, selecting year for Amazon S3 and right\_year for renamed keys. We'll have AWS Glue handle it through mapping.
* The AND operator and the condition year<=2000 are used in Conditional Routing.



* For our data goals, we selected two AWS buckets.
* For our data goals, we selected two AWS buckets.
* We choose the Uncompressed compression type and Parquet format, and we keep the table I need to import our data into.

* Save the job now, then execute it. If it is successful, go to Athena.



* We use Athena to query the data catalog table.   
    
  