Source code -- AWS Big Data Project

Github link : https://github.com/manishsharma1402/AwsBigDataProject

For s3

Click on s3 -> create new bucket 🡪 give unique name and all the required permissions 🡪 create

For object in s3

Click on bucket name 🡪 create instance 🡪 add file 🡪 upload the file 🡪 uploaded successfully

For crawler

Go to glue 🡪 select crawler 🡪 create new 🡪 give unique name 🡪 give IAM permission and souce as s3 bucket 🡪 create db (database) 🡪 click on create

For metadata catalog

Click on crawler name 🡪 click run 🡪 click on logs

For glue job

Click on job 🡪 click on create 🡪 in source give s3 instance 🡪 in target give parquet file location 🡪 give unique name 🡪 save 🡪 run 🡪 check files in s3 bucket

For query

Go to athena 🡪 select db 🡪 select table 🡪 write query 🡪 click on run 🡪 get result