Generate Dummy data and upload to S3

New arrival of the file will be notified via Event Bridge

Amazon EventBridge Triggers Step function

AWS Step Functions involkes lambda function

1. Read the s3 CSV file
2. parse the file
3. upsert the records in Redshift

Upsert logic will be implmented on Redshift

Incremental Data Processing Architecture

Project Assignment: Building an Incremental Data Processing Pipeline with AWS

1. s3 bucket - incrementaldataprocessing
2. Event bridge notification is enabled
3. Event bridge rule (incrementalDataProcessing) is defined that captures the events - create file/upoad file to s3 bucket
4. The event tirggers step function (incrementalDataProcecssing) which in turn invokes lambda function(incrementalDataProcessing)
5. Function interacts with redshift with the help of library pg8000 to perform upsert operation using merge.
6. python scrip (generate_dummy_data.py) generates the transactions file