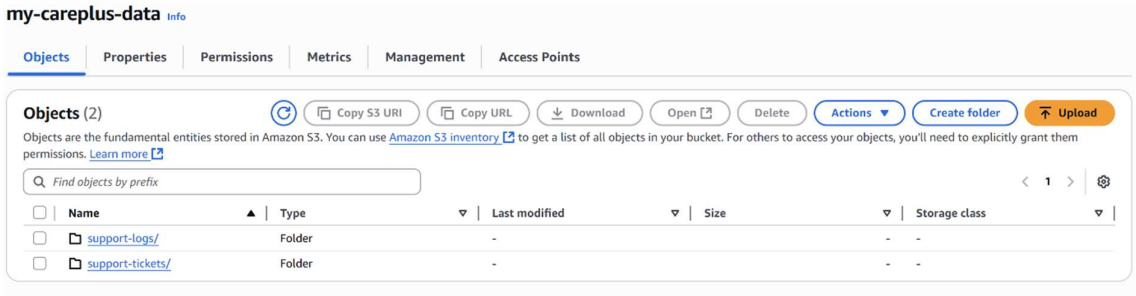
AWS ETL PIPELINE ARCHITECTURE VISUAL EXAMPLES

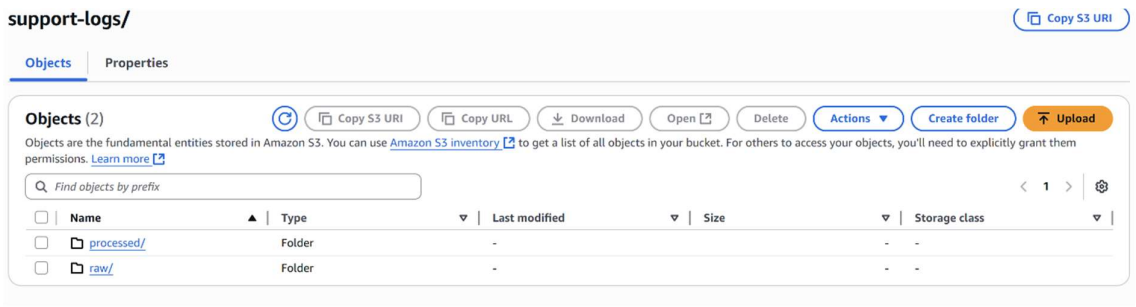
BY ABDUL RAYHAAN

1)AWS S3 WILL LOOK LIKE:

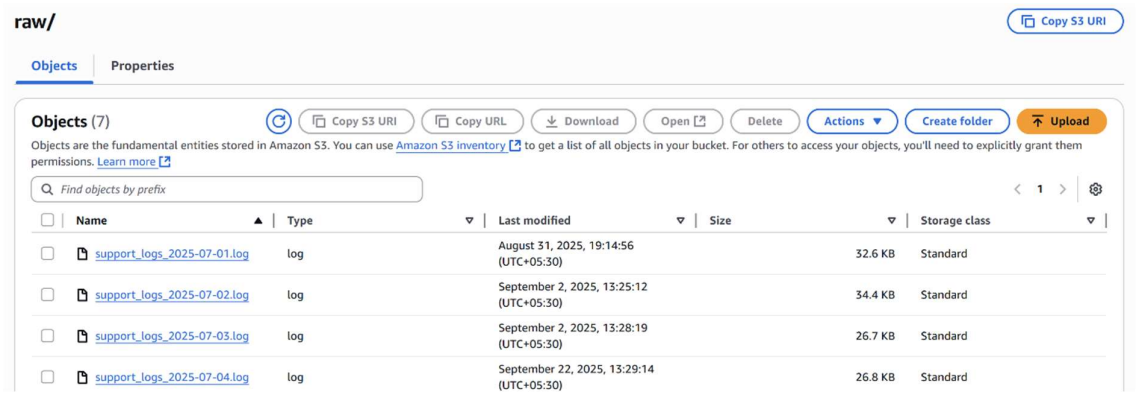
YOU WILL CREATE TWO FOLDERS FOR LOG DATA AND CSV DATA

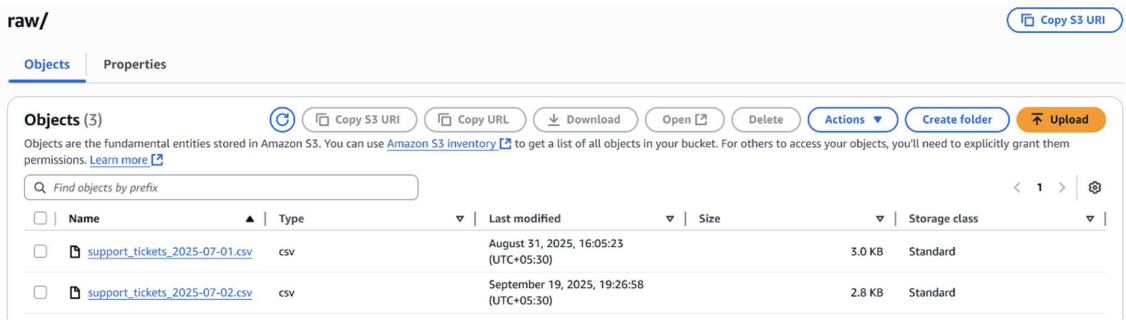


THEN, IN EACH FOLDER YOU WILL CREATE TWO MORE FOLDERS NAMED RAW AND PROCESSED

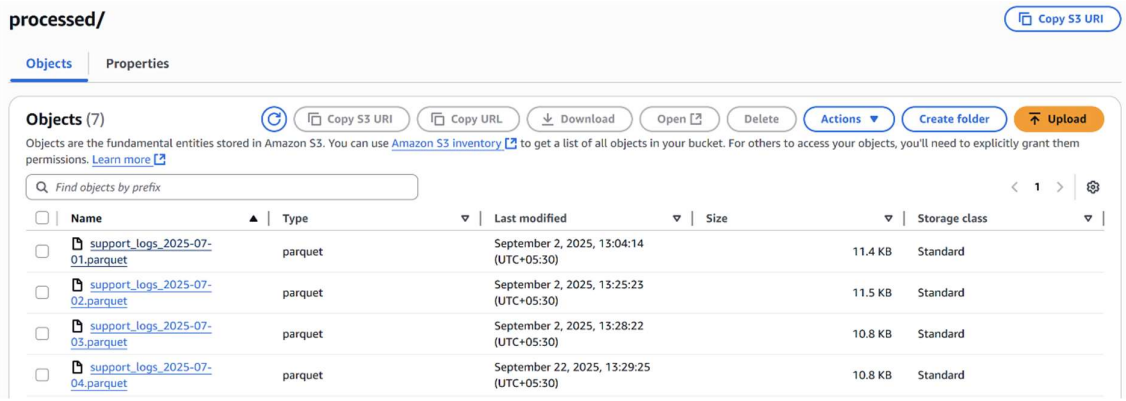


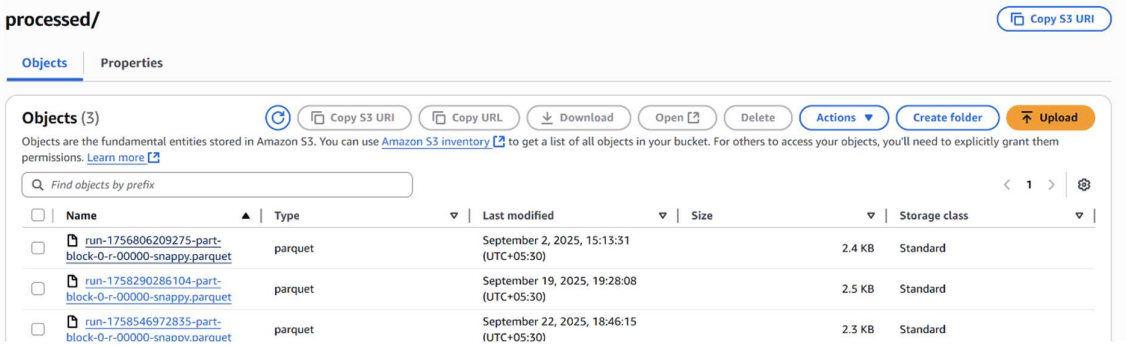
WHEN YOU DONE WITH INGESTION OF DATA, YOUR RAW FOLDERS WILL HAVE .LOG AND .CSV FILES AS SHOWN BELOW





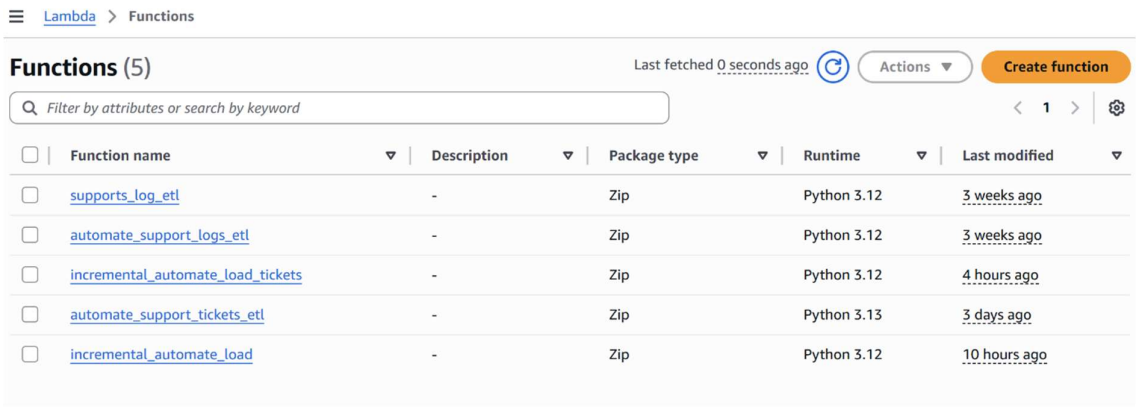
WHEN YOU ARE DONE WITH USING LAMBDA FUNCTION (automate_support_logs , automate_support_tickets) AND GLUE SERVICE , YOUR DATA WILL BE TRANSFORMED AND LOADED TO .processed folder AND WILL LOOK LIKE





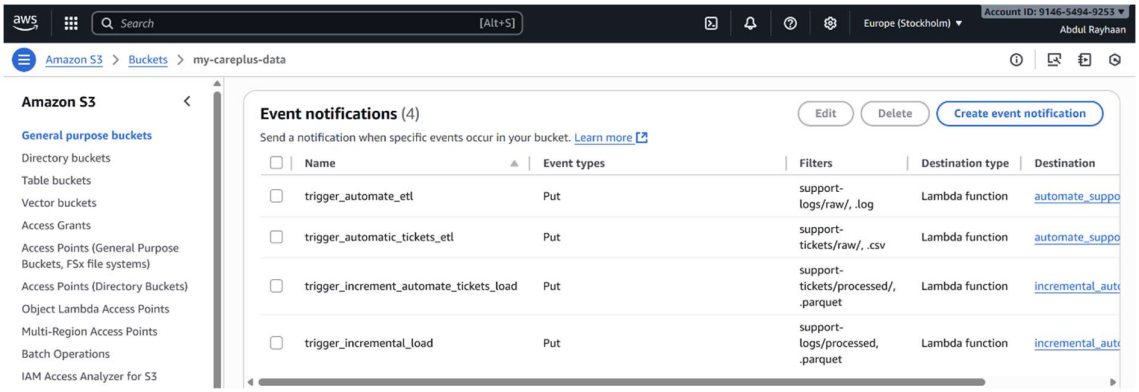
2)AWS LAMBDA FUNCTIONS:

YOUR AWS LAMBDA WILL LOOK LIKE THIS



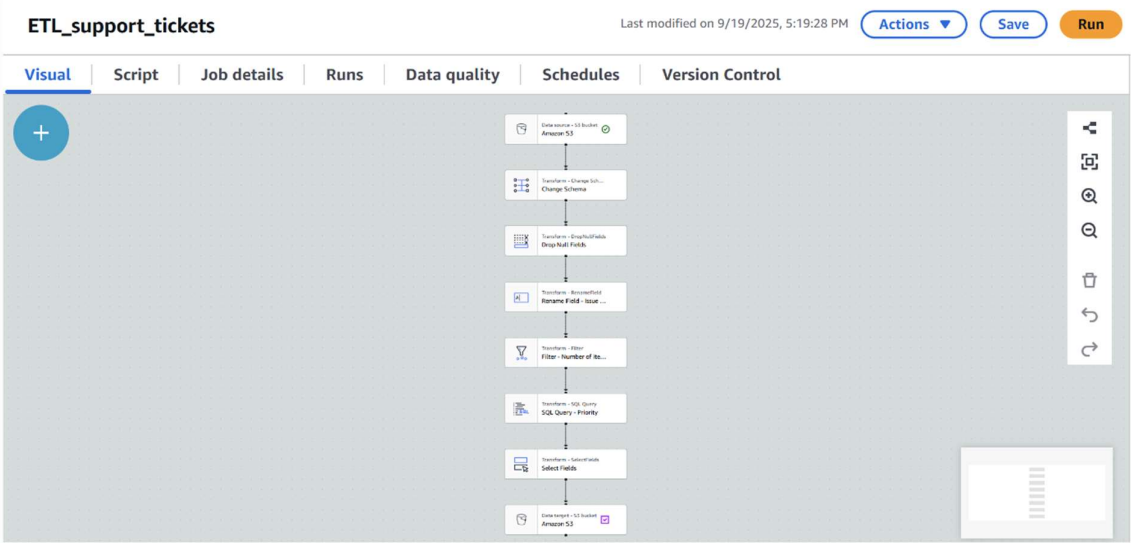
FOR EACH LAMBDA FUNCTION CREATE AN **IAM ROLE** WITH REQUIRED PERMISSIONS AND ADD REQUIRED LAYERS , ADJUST YOUR CONFIGURATIONS AS REQUIRED AND MAKE SURE TO CREATE EVENT NOTIFICATIONS FOR EACH FUNCTIONS SO THAT THEY CAN BE TRIGGERED

3)EVENT NOTIFICATIONS:

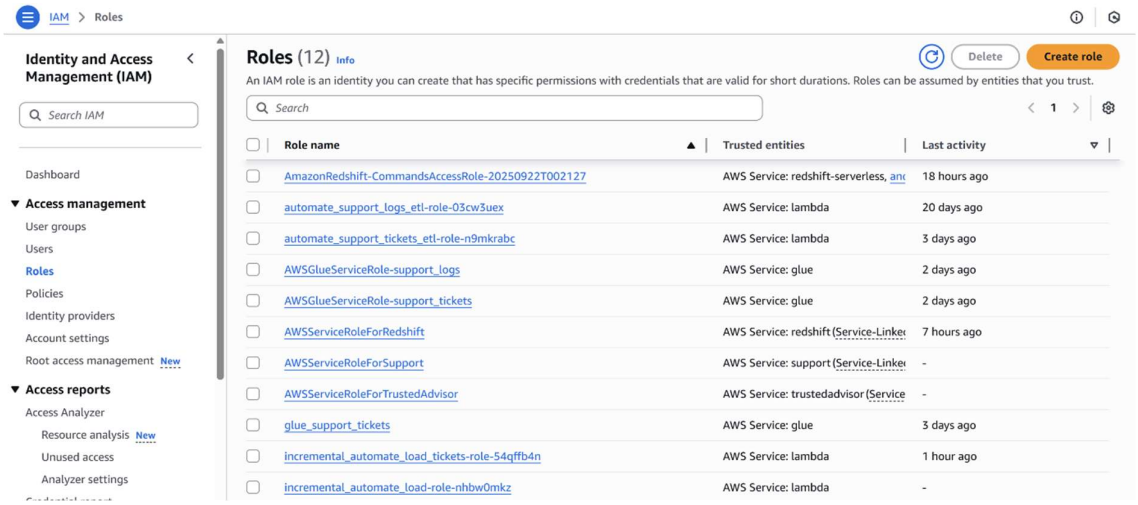


4)GLUE SERVICE:

YOU CAN USE VISUAL ETL INSTEAD OF SCRIPT AND PERFORM REQUIRED TRANSFORMATIONS, AT THE END THE TRANSFORMATION SEQUENCE WILL LOOK LIKE

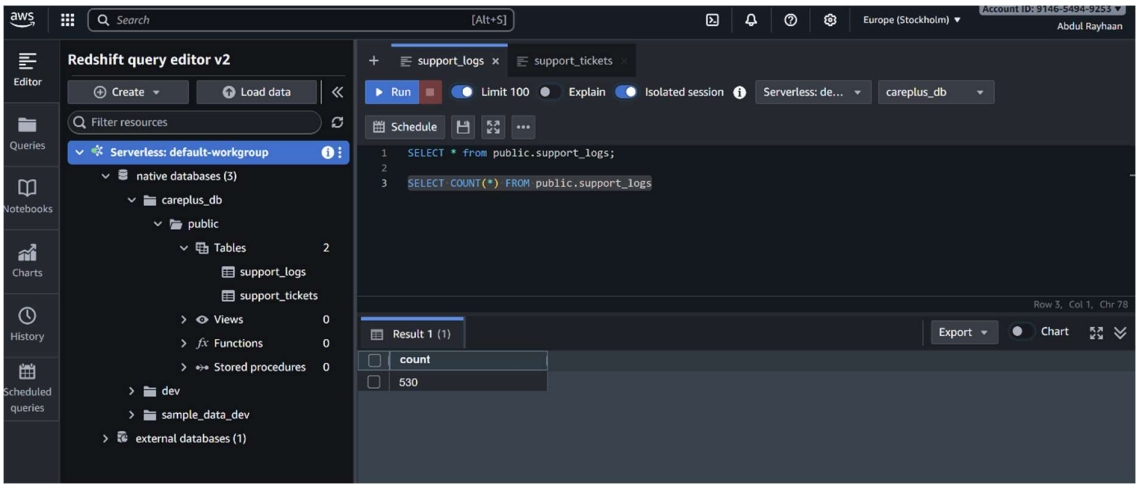


5)IAM ROLES:

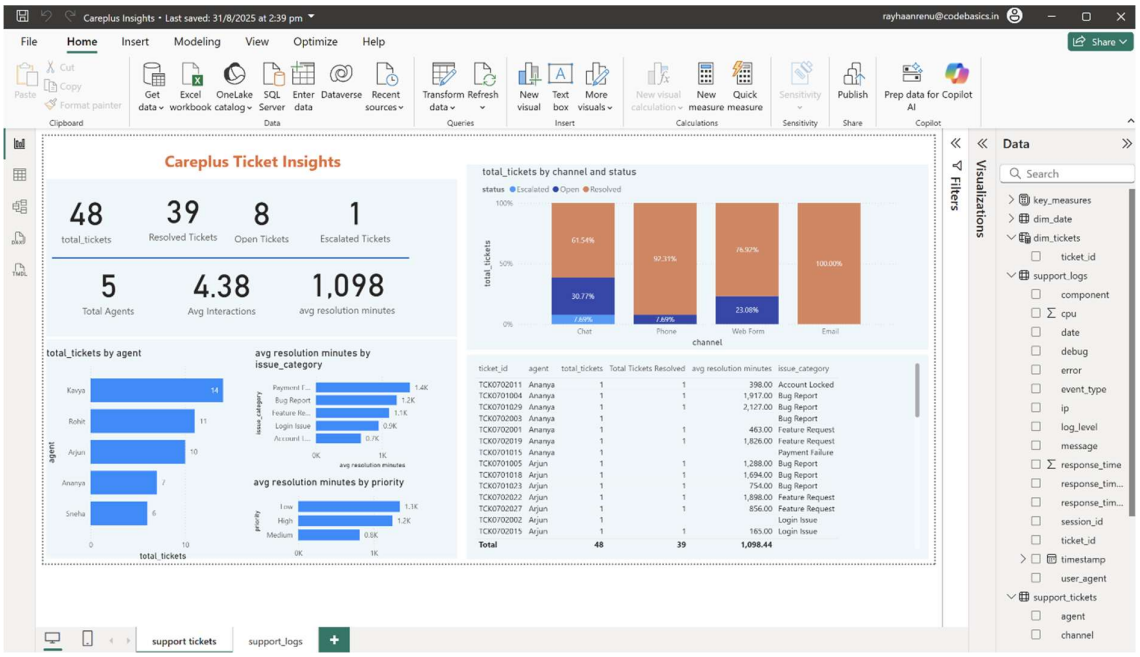


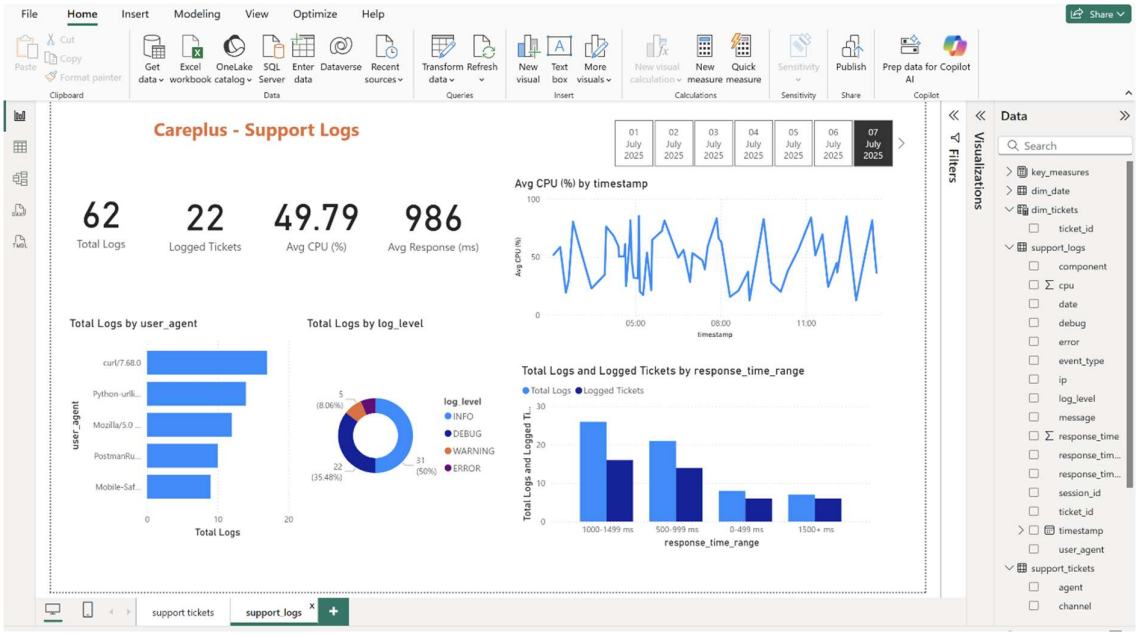
6)AWS REDSHIFT:

IN AWS REDSHIFT, AFTER YOU HAVE CREATED DATABASE AND LOADED TRANSFORMED DATA TO REDSHIFT, YOU WILL SEE SOMETHING LIKE THIS



7)POWERBI VISUALS




File
Home
Insert
Modeling
View
Optimize
Help
Share
Careplus - Support Logs
01 July 2025
02 July 2025
03 July 2025
04 July 2025
05 July 2025
06 July 2025
07 July 2025
62
Total Logs
22
Logged Tickets
49.79
Avg CPU (%)
986
Avg Response (ms)
Avg CPU (%) by timestamp
Total Logs by user_agent
curl/7.68.0
Python-urlli...
Mozilla/5.0 ...
PostmanRu...
Mobile-Saf...
Total Logs
Total Logs by log_level
log_level
INFO
DEBUG
WARNING
ERROR
Total Logs and Logged Tickets by response_time_range
Total Logs
Logged Tickets
1000-1499 ms
500-999 ms
0-499 ms
1500+ ms
response_time_range
Filters
Visualizations
Data
Search
key_measures
dim_date
dim_tickets
ticket_id
support_logs
component
cpu
date
debug
error
event_type
ip
log_level
message
response_time
session_id
ticket_id
timestamp
user_agent
support_tickets
agent
channel
support tickets
support_logs