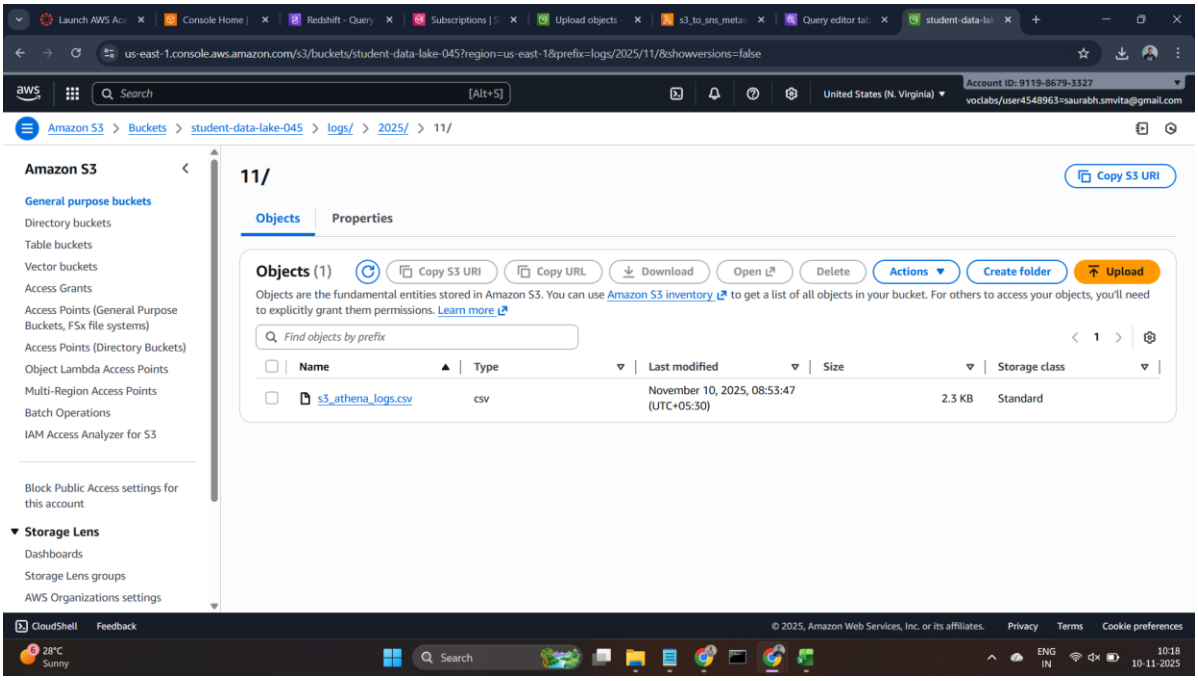
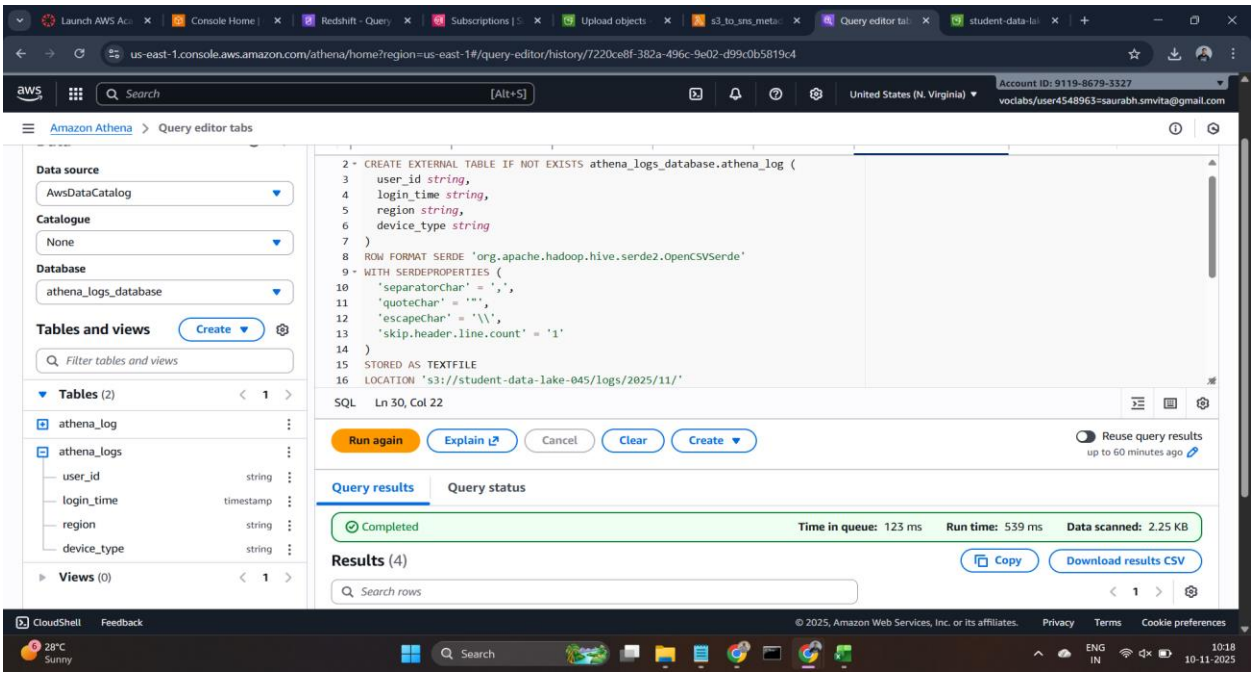
Name : Saurabh ShivPrakash Tripathi

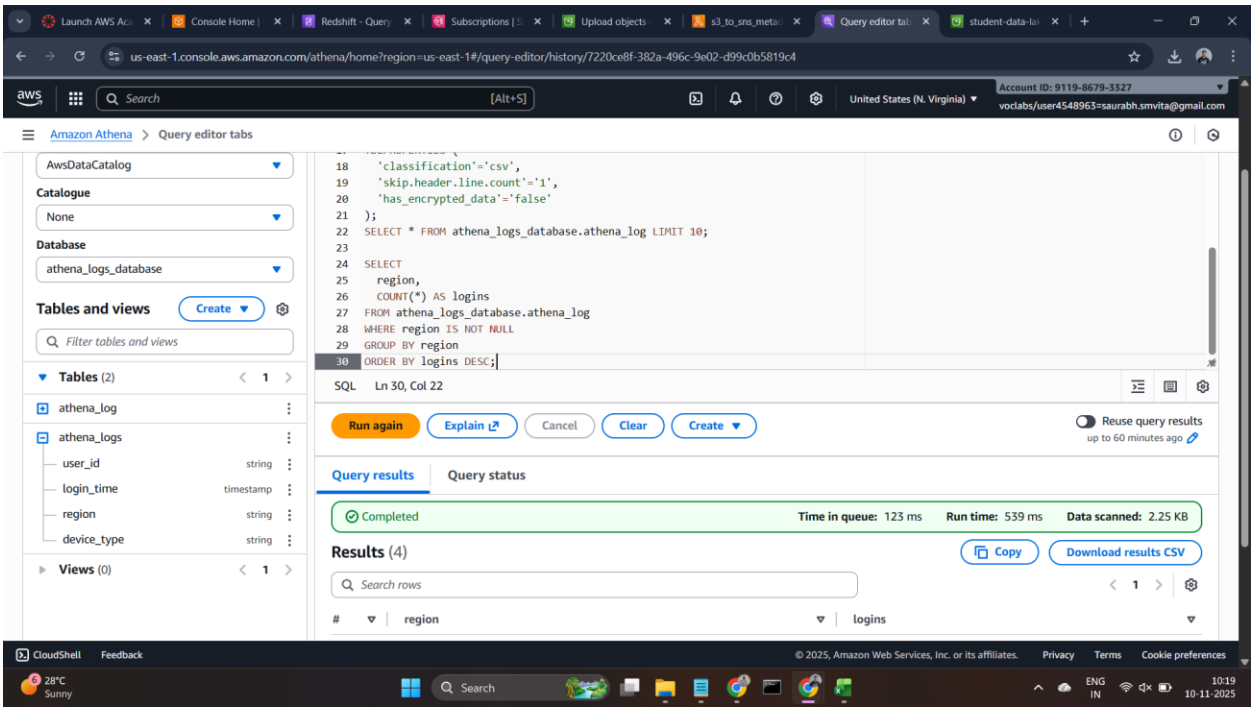
Prn No. : 045

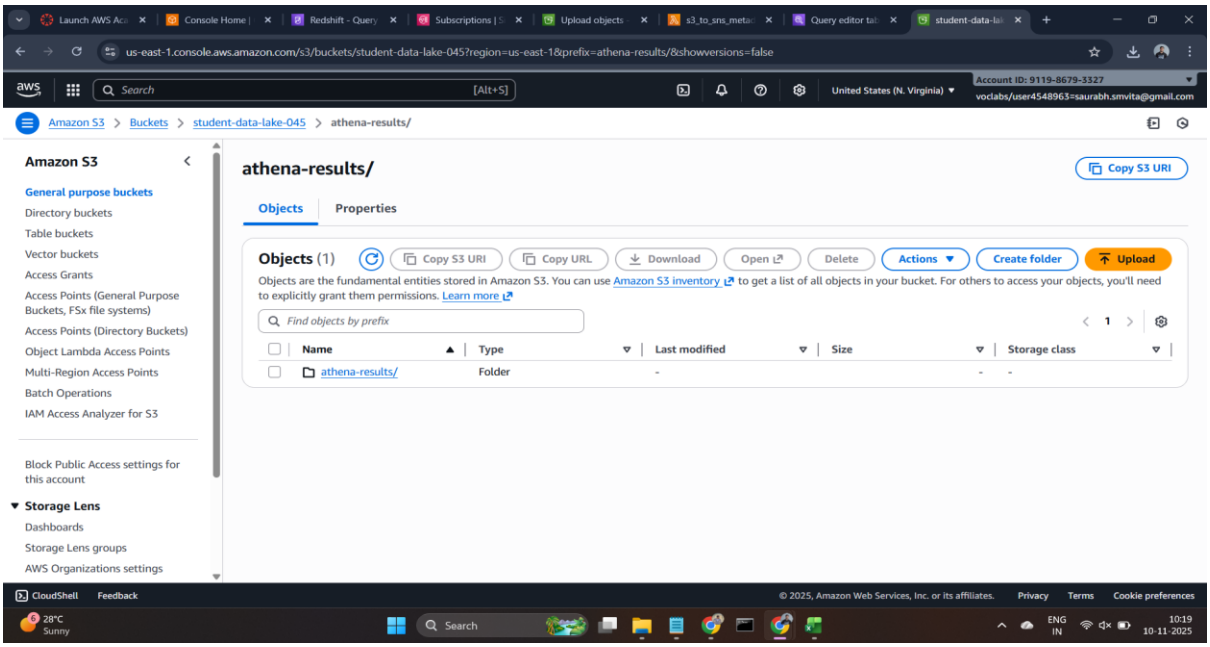
Q.1

Q.1) You have an S3 bucket student-data-lake containing multiple CSV files under the prefix logs/2025/11/. Each file has columns: user_id, login_time, region, device_type. Task: 1. Create an Athena external table to query CSV files directly from S3. 2. Write a query to return number of logins per region. 3. Save the query results to S3 folder athena-results


us-east-1.console.aws.amazon.com/s3/buckets/student-data-lake-045?region=us-east-1&prefix=logs/2025/11/&showversions=false
Amazon S3 > Buckets > student-data-lake-045 > logs/ > 2025/ > 11/
Amazon S3
General purpose buckets
Directory buckets
Table buckets
Vector buckets
Access Grants
Access Points (General Purpose Buckets, FSx file systems)
Access Points (Directory Buckets)
Object Lambda Access Points
Multi-Region Access Points
Batch Operations
IAM Access Analyzer for S3
Block Public Access settings for this account
Storage Lens
Dashboards
Storage Lens groups
AWS Organizations settings
11/
Copy S3 URI
Objects
Properties
Objects (1)
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Find objects by prefix
Name
Type
Last modified
Size
Storage class
s3_athena_logs.csv
csv
November 10, 2025, 08:53:47 (UTC+05:30)
2.3 KB
Standard
CloudShell
Feedback
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

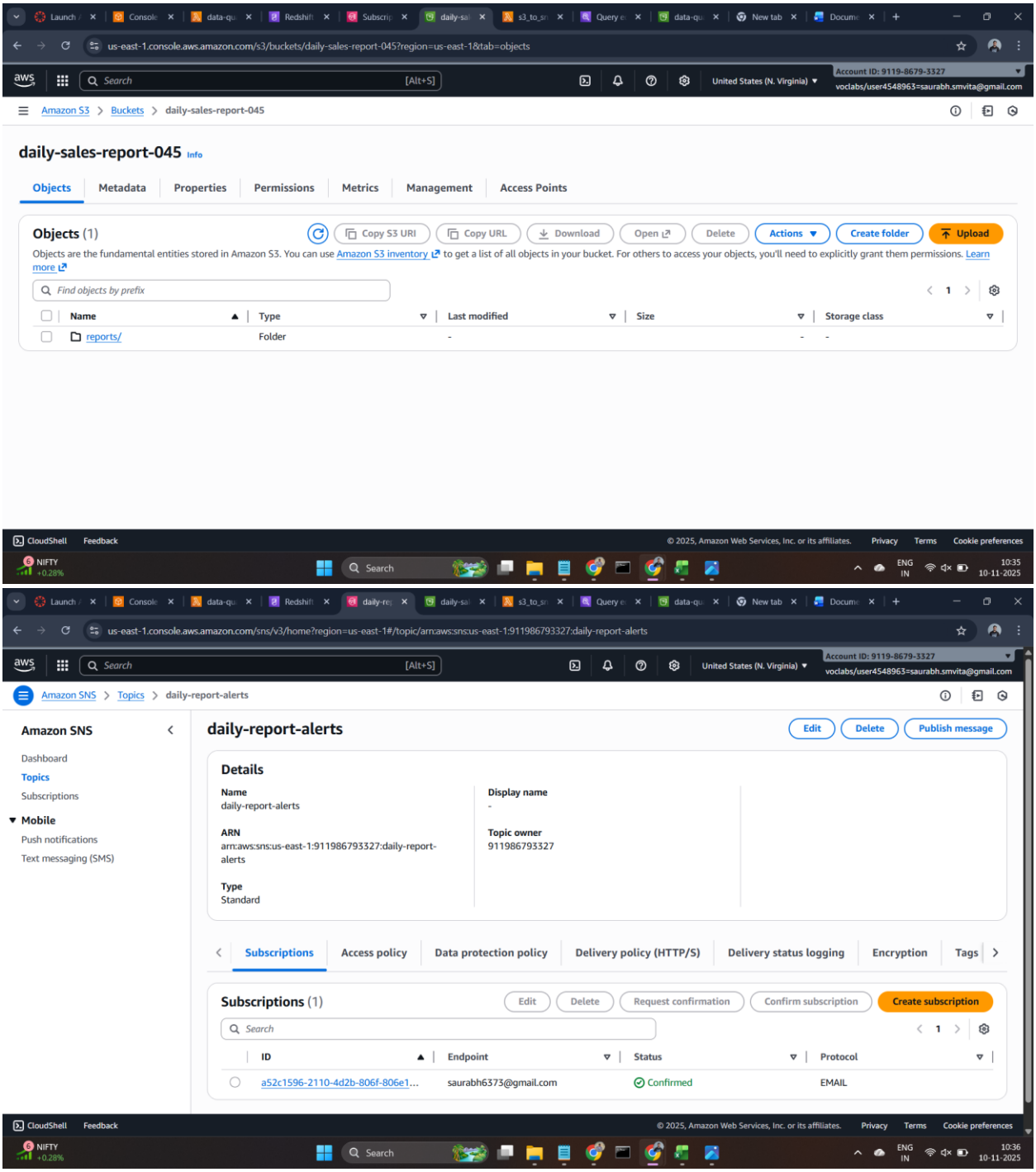

us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/7220ce8f-382a-496c-9e02-d99c0b5819c4
Amazon Athena > Query editor tabs
Data source
AwsDataCatalog
Catalogue
None
Database
athena_logs_database
Tables and views
Create
Filter tables and views
Tables (2)
athena_log
athena_logs
user_id string
login_time timestamp
region string
device_type string
Views (0)
CREATE EXTERNAL TABLE IF NOT EXISTS athena_logs_database.athena_log (
user_id string,
login_time string,
region string,
device_type string
)
ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.OpenCSVSerde'
WITH SERDEPROPERTIES (
'separatorChar' = ',',
'quoteChar' = '"',
'escapeChar' = '\\',
'skip.header.line.count' = '1'
)
STORED AS TEXTFILE
LOCATION 's3://student-data-lake-045/logs/2025/11/'
SQL Ln 30, Col 22
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query status
Completed
Time in queue: 123 ms
Run time: 539 ms
Data scanned: 2.25 KB
Results (4)
Copy
Download results CSV
Search rows
CloudShell
Feedback
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

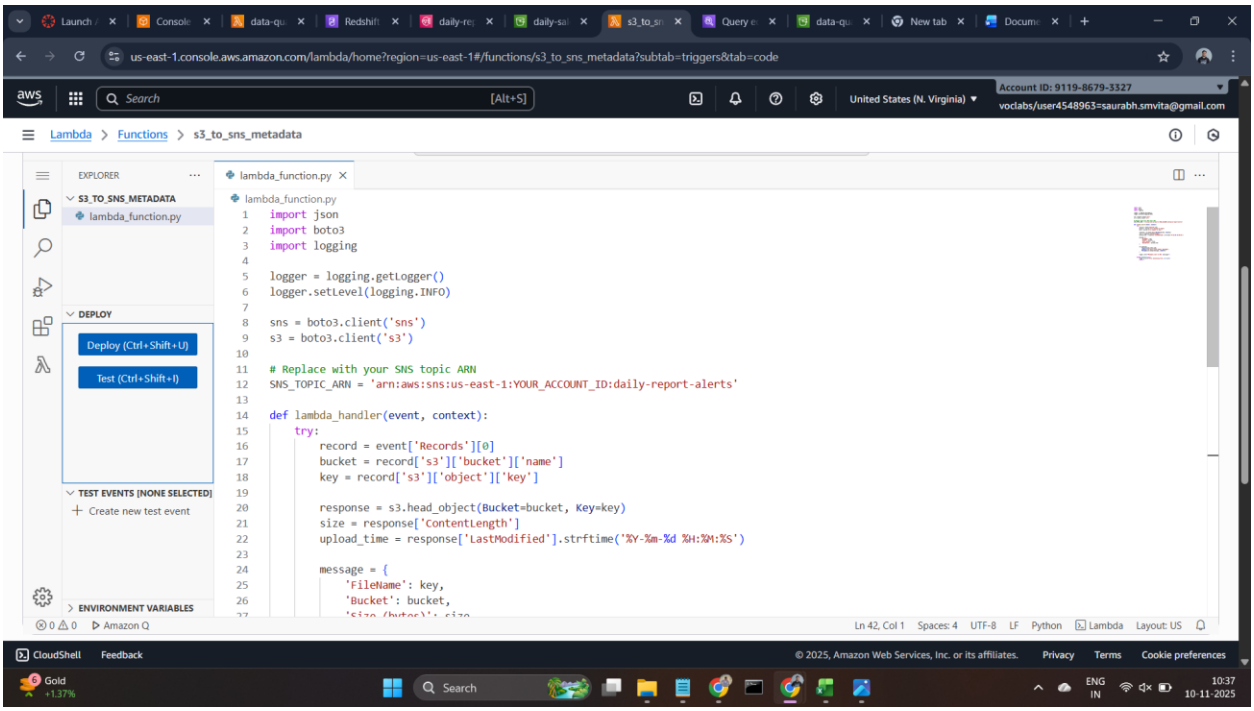


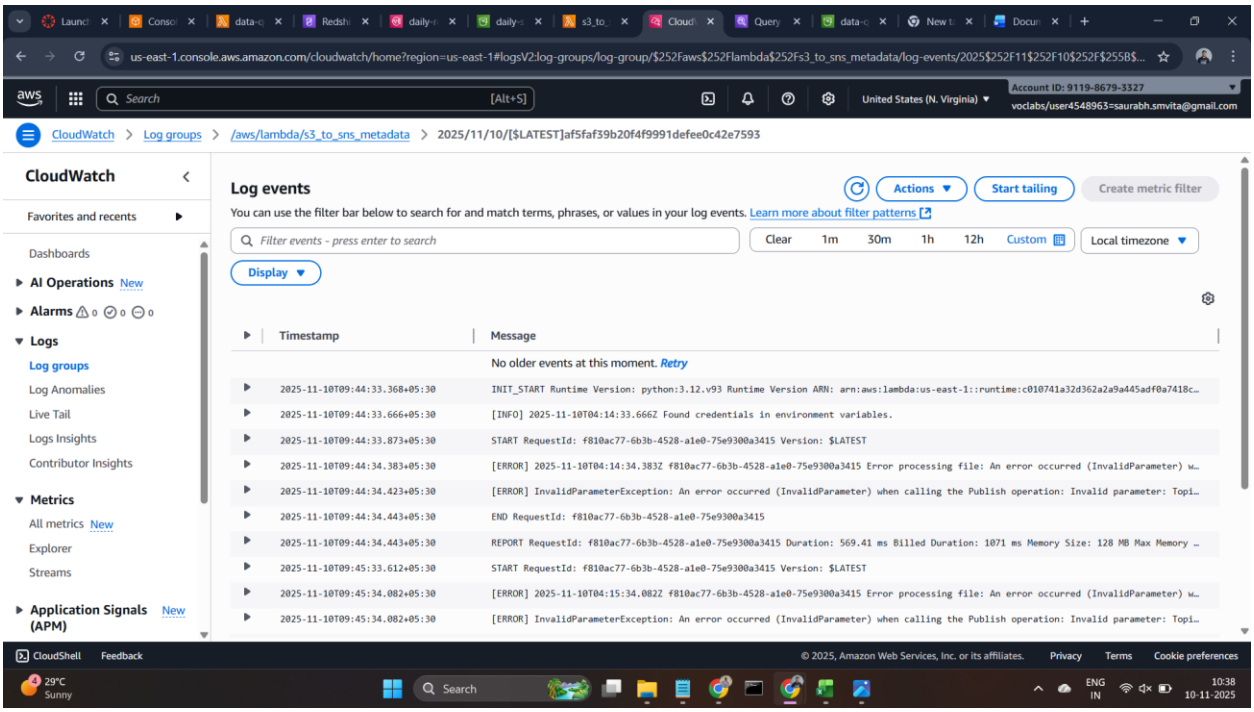


Save the query results to S3 folder athena-results

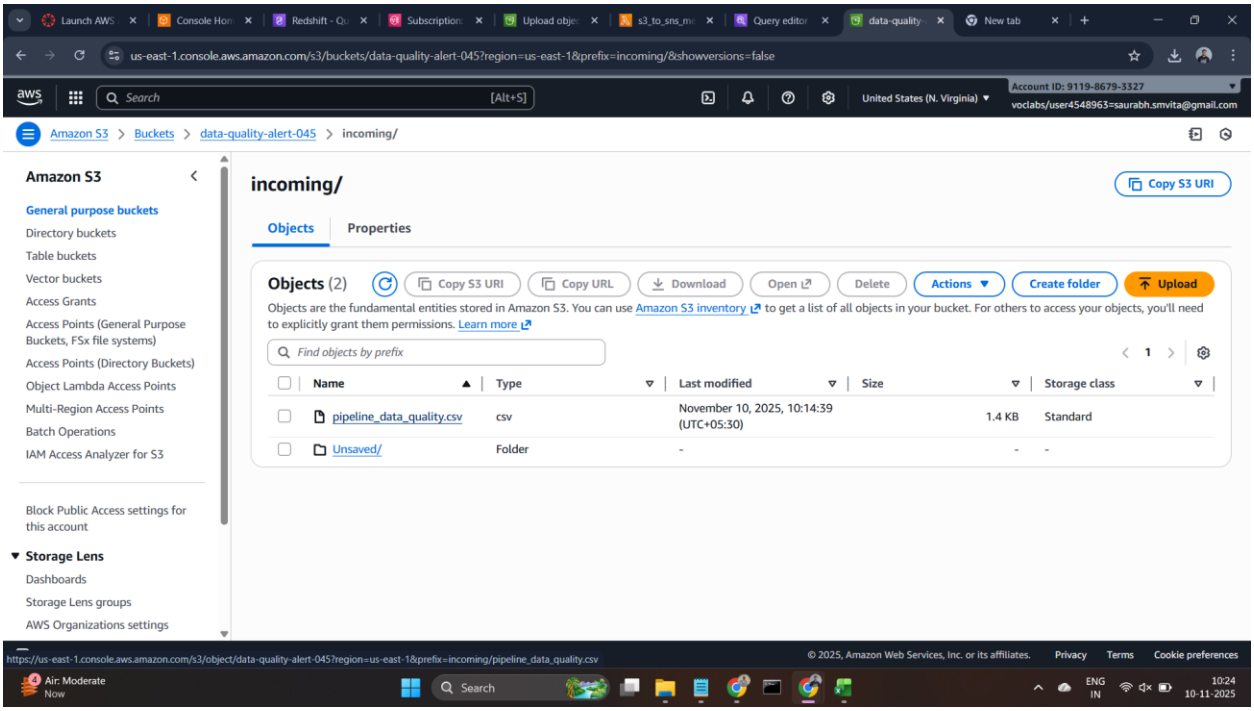
Q.2

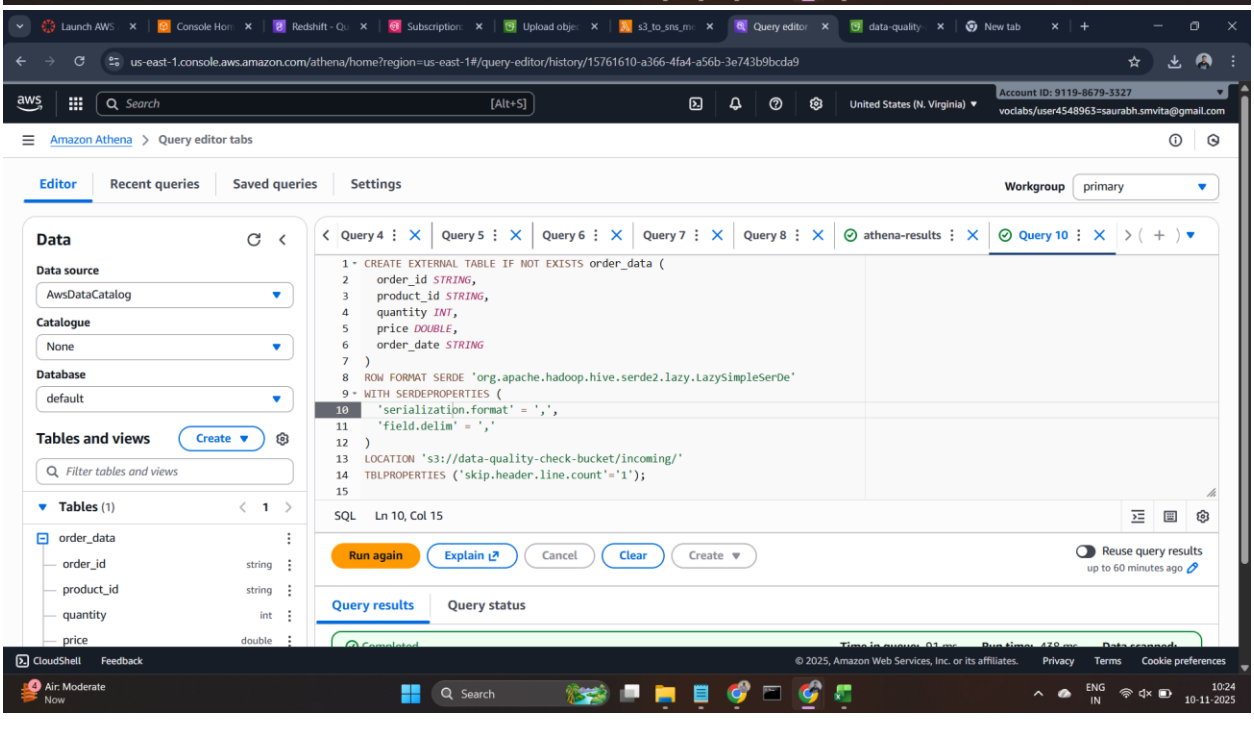

us-east-1.console.aws.amazon.com/s3/buckets/daily-sales-report-045?region=us-east-1&tab=objects
Amazon S3 > Buckets > daily-sales-report-045
daily-sales-report-045 Info
Objects
Metadata
Properties
Permissions
Metrics
Management
Access Points
Objects (1)
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Find objects by prefix
Name
Type
Last modified
Size
Storage class
reports/
Folder
us-east-1.console.aws.amazon.com/sns/v3/home?region=us-east-1#/topic/arn:aws:sns:us-east-1:911986793327:daily-report-alerts
Amazon SNS > Topics > daily-report-alerts
Amazon SNS
Dashboard
Topics
Subscriptions
Mobile
Push notifications
Text messaging (SMS)
daily-report-alerts
Edit
Delete
Publish message
Details
Name
daily-report-alerts
Display name
-
ARN
arn:aws:sns:us-east-1:911986793327:daily-report-alerts
Topic owner
911986793327
Type
Standard
Subscriptions
Access policy
Data protection policy
Delivery policy (HTTP/S)
Delivery status logging
Encryption
Tags
Subscriptions (1)
Edit
Delete
Request confirmation
Confirm subscription
Create subscription
Search
ID
Endpoint
Status
Protocol
a52c1596-2110-4d2b-806f-806e1...
saurabh6373@gmail.com
Confirmed
EMAIL
CloudShell
Feedback
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

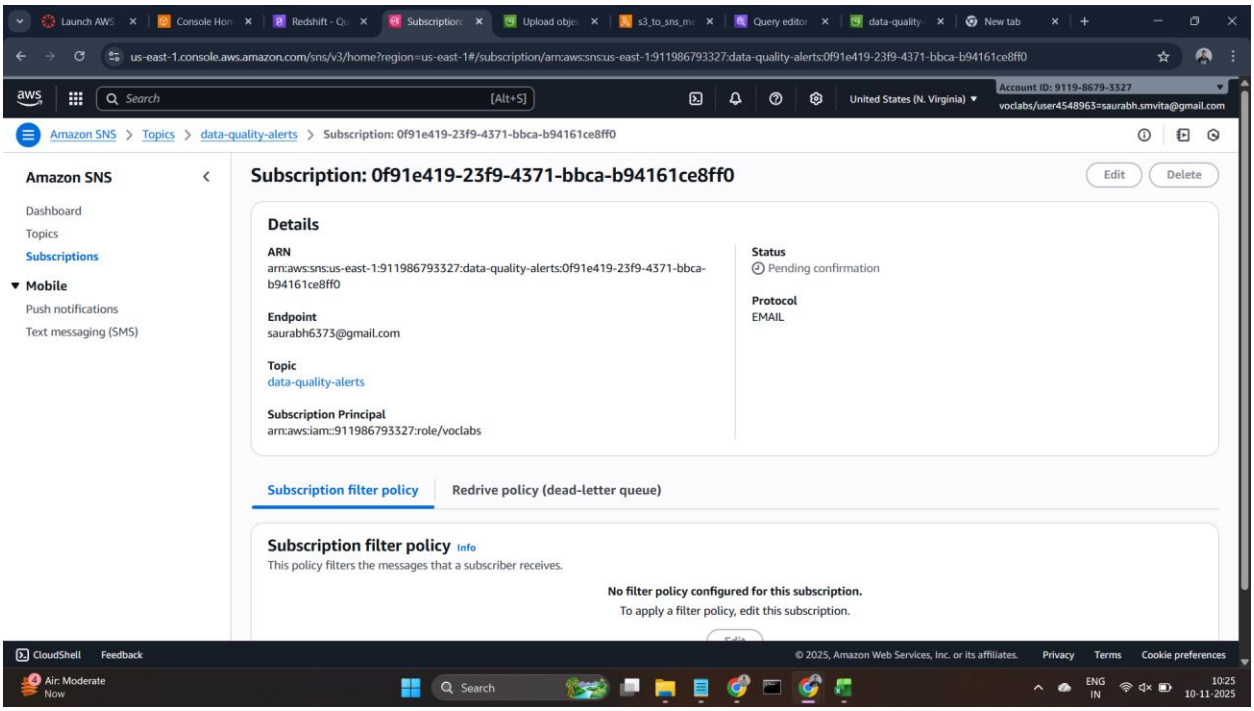

us-east-1.console.aws.amazon.com/lambda/home?region=us-east-1#/functions/s3_to_sns_metadata?subtab=triggers&tab=code
Lambda > Functions > s3_to_sns_metadata
lambda_function.py
import json
import boto3
import logging
logger = logging.getLogger()
logger.setLevel(logging.INFO)
sns = boto3.client('sns')
s3 = boto3.client('s3')
# Replace with your SNS topic ARN
SNS_TOPIC_ARN = 'arn:aws:sns:us-east-1:YOUR_ACCOUNT_ID:daily-report-alerts'
def lambda_handler(event, context):
try:
record = event['Records'][0]
bucket = record['s3']['bucket']['name']
key = record['s3']['object']['key']
response = s3.head_object(Bucket=bucket, Key=key)
size = response['ContentLength']
upload_time = response['LastModified'].strftime('%Y-%m-%d %H:%M:%S')
message = {
'FileName': key,
'Bucket': bucket,
Deploy (Ctrl+Shift+U)
Test (Ctrl+Shift+I)
Create new test event

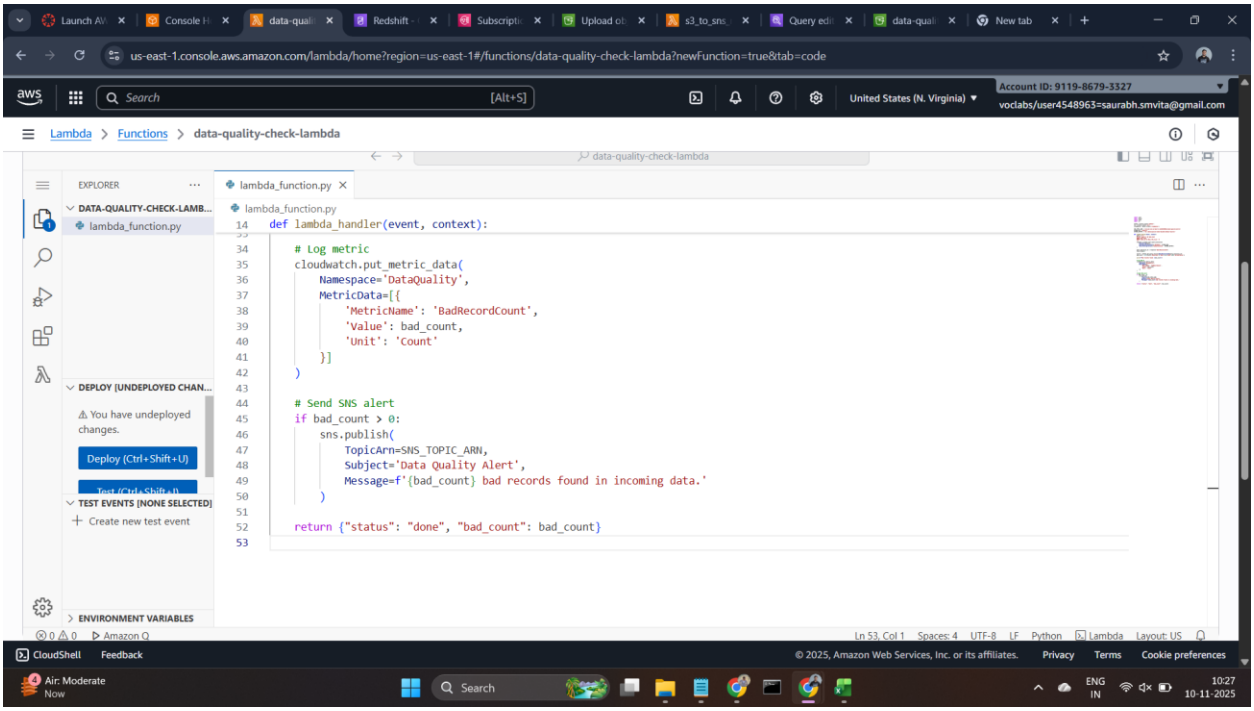

us-east-1.console.aws.amazon.com/cloudwatch/home?region=us-east-1#logsV2:log-groups/log-group/$252Faws$252Flambda$252Fs3_to_sns_metadata/log-events/2025$252F11$252F10$252F$255B$...
CloudWatch > Log groups > /aws/lambda/s3_to_sns_metadata > 2025/11/10/[$LATEST]af5faf39b20f4f9991defee0c42e7593
CloudWatch
Log events
Actions
Start tailing
Create metric filter
You can use the filter bar below to search for and match terms, phrases, or values in your log events. Learn more about filter patterns
Filter events - press enter to search
Clear
1m
30m
1h
12h
Custom
Local timezone
Display
Timestamp
Message
No older events at this moment. Retry
2025-11-10T09:44:33.368+05:30 INIT_START Runtime Version: python:3.12.v93 Runtime Version ARN: arn:aws:lambda:us-east-1::runtime:c010741a32d362a2a9a445adf0a7418c...
2025-11-10T09:44:33.666+05:30 [INFO] 2025-11-10T04:14:33.666Z Found credentials in environment variables.
2025-11-10T09:44:33.873+05:30 START RequestId: f810ac77-6b3b-4528-a1e0-75e9300a3415 Version: $LATEST
2025-11-10T09:44:34.383+05:30 [ERROR] 2025-11-10T04:14:34.383Z f810ac77-6b3b-4528-a1e0-75e9300a3415 Error processing file: An error occurred (InvalidParameter) w...
2025-11-10T09:44:34.423+05:30 [ERROR] InvalidParameterException: An error occurred (InvalidParameter) when calling the Publish operation: Invalid parameter: Topi...
2025-11-10T09:44:34.443+05:30 END RequestId: f810ac77-6b3b-4528-a1e0-75e9300a3415
2025-11-10T09:44:34.443+05:30 REPORT RequestId: f810ac77-6b3b-4528-a1e0-75e9300a3415 Duration: 569.41 ms Billed Duration: 1071 ms Memory Size: 128 MB Max Memory ...
2025-11-10T09:45:33.612+05:30 START RequestId: f810ac77-6b3b-4528-a1e0-75e9300a3415 Version: $LATEST
2025-11-10T09:45:34.082+05:30 [ERROR] 2025-11-10T04:15:34.082Z f810ac77-6b3b-4528-a1e0-75e9300a3415 Error processing file: An error occurred (InvalidParameter) w...
2025-11-10T09:45:34.082+05:30 [ERROR] InvalidParameterException: An error occurred (InvalidParameter) when calling the Publish operation: Invalid parameter: Topi...

Q.3

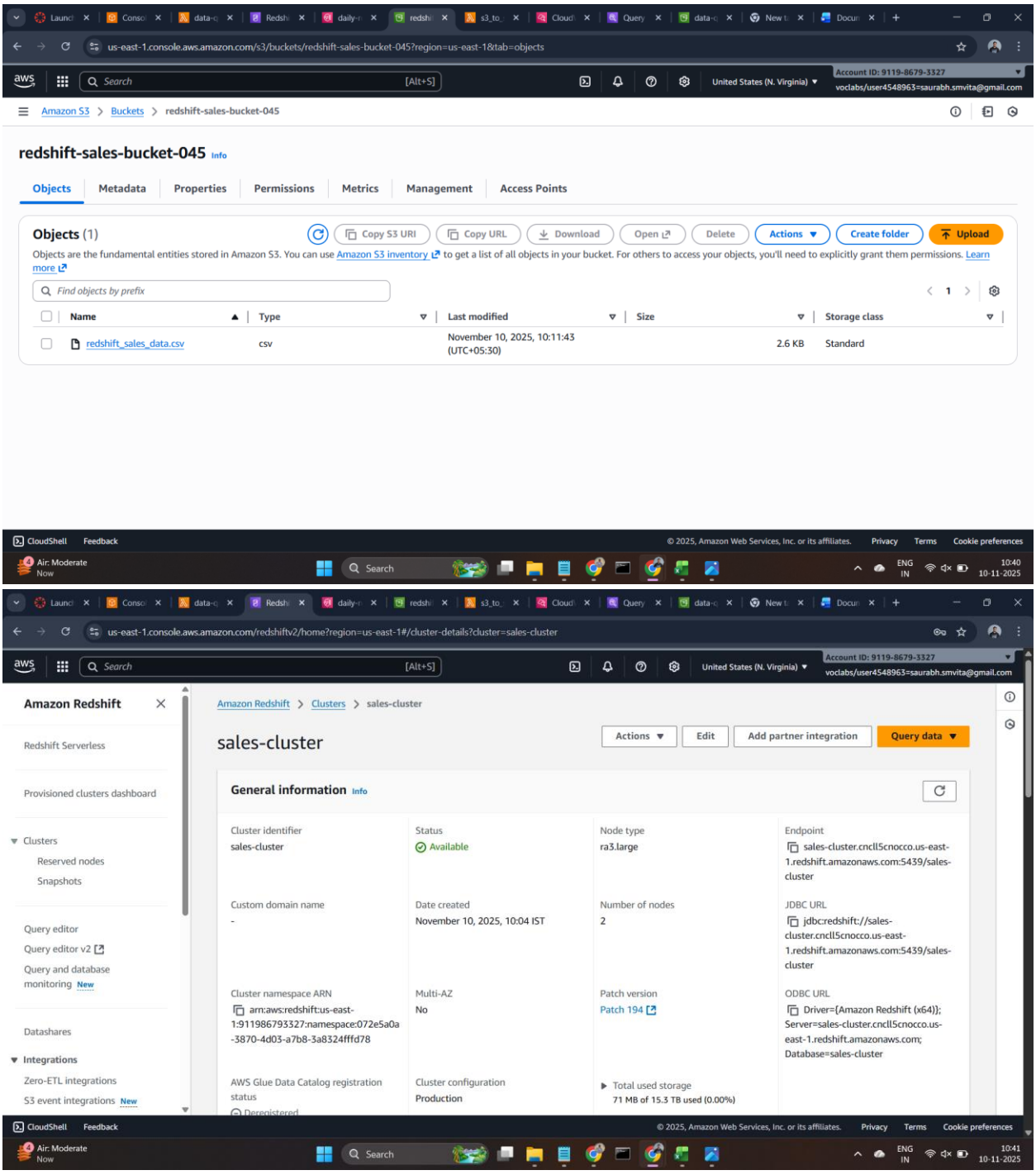

us-east-1.console.aws.amazon.com/s3/buckets/data-quality-alert-045?region=us-east-1&prefix=incoming/&showversions=false
Amazon S3 > Buckets > data-quality-alert-045 > incoming/
incoming/
Copy S3 URI
Objects
Properties
Objects (2)
Copy S3 URI
Copy URL
Download
Open
Delete
Actions
Create folder
Upload
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Find objects by prefix
Name
Type
Last modified
Size
Storage class
pipeline_data_quality.csv
csv
November 10, 2025, 10:14:39 (UTC+05:30)
1.4 KB
Standard
Unsaved/
Folder

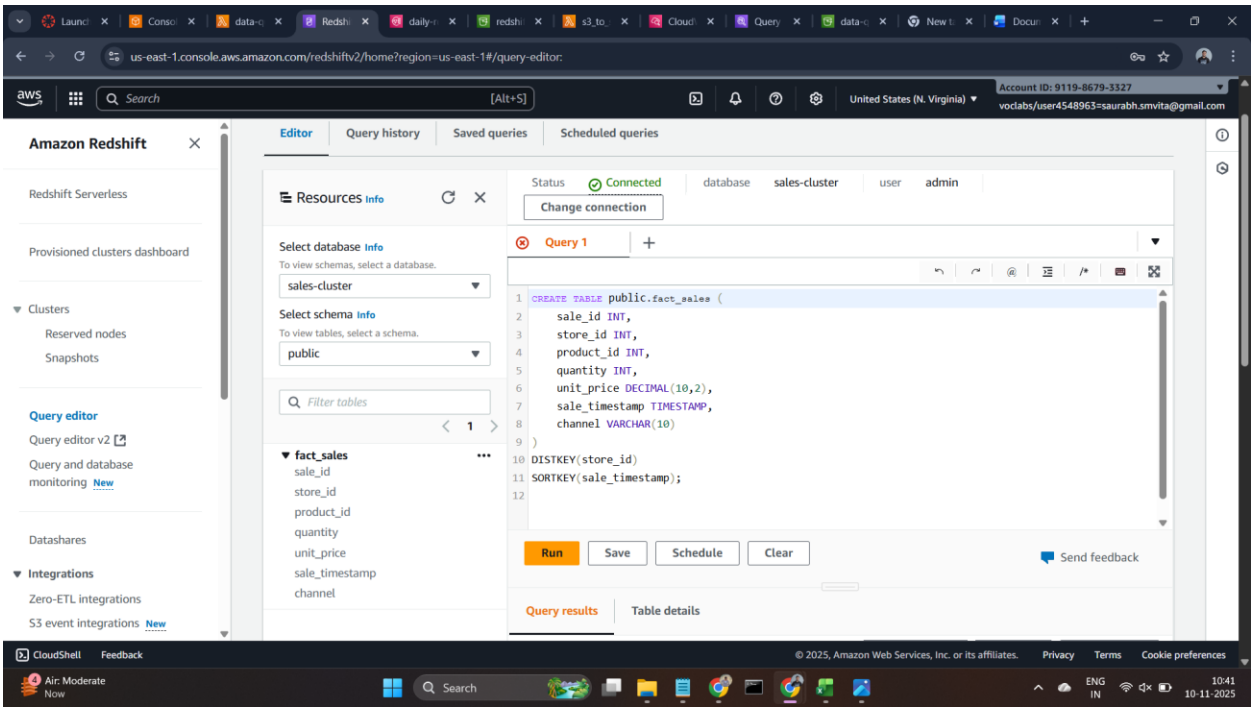

us-east-1.console.aws.amazon.com/athena/home?region=us-east-1#/query-editor/history/15761610-a366-4fa4-a56b-3e743b9bcda9
Amazon Athena > Query editor tabs
Editor
Recent queries
Saved queries
Settings
Workgroup
primary
Data
Data source
AwsDataCatalog
Catalogue
None
Database
default
Tables and views
Create
Filter tables and views
Tables (1)
order_data
order_id string
product_id string
quantity int
price double
Query 4
Query 5
Query 6
Query 7
Query 8
athena-results
Query 10
CREATE EXTERNAL TABLE IF NOT EXISTS order_data (
order_id STRING,
product_id STRING,
quantity INT,
price DOUBLE,
order_date STRING
)
ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'
WITH SERDEPROPERTIES (
'serialization.format' = ',',
'field.delim' = ','
)
LOCATION 's3://data-quality-check-bucket/incoming/'
TBLPROPERTIES ('skip.header.line.count'='1');
SQL
Ln 10, Col 15
Run again
Explain
Cancel
Clear
Create
Reuse query results
up to 60 minutes ago
Query results
Query status


us-east-1.console.aws.amazon.com/sns/v3/home?region=us-east-1#/subscription/arn:aws:sns:us-east-1:911986793327:data-quality-alerts:0f91e419-23f9-4371-bbca-b94161ce8ff0
Search
[Alt+S]
United States (N. Virginia)
Account ID: 9119-8679-3327
voclabs/user4548963=saurabh.smvita@gmail.com
Amazon SNS > Topics > data-quality-alerts > Subscription: 0f91e419-23f9-4371-bbca-b94161ce8ff0
Amazon SNS
Dashboard
Topics
Subscriptions
Mobile
Push notifications
Text messaging (SMS)
Subscription: 0f91e419-23f9-4371-bbca-b94161ce8ff0
Edit
Delete
Details
ARN
arn:aws:sns:us-east-1:911986793327:data-quality-alerts:0f91e419-23f9-4371-bbca-b94161ce8ff0
Endpoint
saurabh6373@gmail.com
Topic
data-quality-alerts
Subscription Principal
arn:aws:iam::911986793327:role/voclabs
Status
Pending confirmation
Protocol
EMAIL
Subscription filter policy
Redrive policy (dead-letter queue)
Subscription filter policy Info
This policy filters the messages that a subscriber receives.
No filter policy configured for this subscription.
To apply a filter policy, edit this subscription.
CloudShell
Feedback
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Air: Moderate
Now
Search
ENG
IN
10:25
10-11-2025


us-east-1.console.aws.amazon.com/lambda/home?region=us-east-1#/functions/data-quality-check-lambda?newFunction=true&tab=code
Search
[Alt+S]
United States (N. Virginia)
Account ID: 9119-8679-3327
voclabs/user4548963=saurabh.smvita@gmail.com
Lambda > Functions > data-quality-check-lambda
data-quality-check-lambda
EXPLORER
DATA-QUALITY-CHECK-LAMB...
lambda_function.py
DEPLOY (UNDEPLOYED CHAN...
You have undeployed changes.
Deploy (Ctrl+Shift+U)
TEST EVENTS [NONE SELECTED]
Create new test event
ENVIRONMENT VARIABLES
lambda_function.py
lambda_function.py
14 def lambda_handler(event, context):
34 # Log metric
35 cloudwatch.put_metric_data(
36 Namespace='DataQuality',
37 MetricData=[{
38 'MetricName': 'BadRecordCount',
39 'Value': bad_count,
40 'Unit': 'Count'
41 }]
42)
43
44 # Send SNS alert
45 if bad_count > 0:
46 sns.publish(
47 TopicArn=SNS_TOPIC_ARN,
48 Subject='Data Quality Alert',
49 Message=f'{bad_count} bad records found in incoming data.'
50)
51
52 return {"status": "done", "bad_count": bad_count}
53
Ln 53, Col 1
Spaces: 4
UTF-8
LF
Python
Lambda
Layout: US
CloudShell
Feedback
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Air: Moderate
Now
Search
ENG
IN
10:27
10-11-2025

Q.4


us-east-1.console.aws.amazon.com/s3/buckets/redshift-sales-bucket-045?region=us-east-1&tab=objects
Search [Alt+S]
United States (N. Virginia)
Account ID: 9119-8679-3327
voclabs/user4548963=saurabh.smvita@gmail.com
Amazon S3 > Buckets > redshift-sales-bucket-045
redshift-sales-bucket-045 Info
Objects Metadata Properties Permissions Metrics Management Access Points
Objects (1)
Copy S3 URI Copy URL Download Open Delete Actions Create folder Upload
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more
Find objects by prefix
Name Type Last modified Size Storage class
redshift_sales_data.csv csv November 10, 2025, 10:11:43 (UTC+05:30) 2.6 KB Standard
CloudShell Feedback
© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences
Air: Moderate Now
Search
ENG IN
10:40 10-11-2025
us-east-1.console.aws.amazon.com/redshiftv2/home?region=us-east-1#/cluster-details?cluster=sales-cluster
Amazon Redshift
Redshift Serverless
Provisioned clusters dashboard
Clusters
Reserved nodes
Snapshots
Query editor
Query editor v2
Query and database monitoring New
Datashares
Integrations
Zero-ETL integrations
S3 event integrations New
Amazon Redshift > Clusters > sales-cluster
sales-cluster
Actions Edit Add partner integration Query data
General information Info
Cluster identifier
sales-cluster
Status
Available
Node type
ra3.large
Endpoint
sales-cluster.cncll5cnocco.us-east-1.redshift.amazonaws.com:5439/sales-cluster
Custom domain name
-
Date created
November 10, 2025, 10:04 IST
Number of nodes
2
JDBC URL
jdbc:redshift://sales-cluster.cncll5cnocco.us-east-1.redshift.amazonaws.com:5439/sales-cluster
Cluster namespace ARN
arn:aws:redshift:us-east-1:911986793327:namespace:072e5a0a-3870-4d03-a7b8-3a8324fffd78
Multi-AZ
No
Patch version
Patch 194
ODBC URL
Driver={Amazon Redshift (x64)}; Server=sales-cluster.cncll5cnocco.us-east-1.redshift.amazonaws.com; Database=sales-cluster
AWS Glue Data Catalog registration status
Deregistered
Cluster configuration
Production
Total used storage
71 MB of 15.3 TB used (0.00%)
10:41 10-11-2025


us-east-1.console.aws.amazon.com/redshiftv2/home?region=us-east-1#/query-editor:
Search
[Alt+S]
United States (N. Virginia)
Account ID: 9119-8679-3327
voclabs/user4548963=saurabh.smvita@gmail.com
Amazon Redshift
Redshift Serverless
Provisioned clusters dashboard
Clusters
Reserved nodes
Snapshots
Query editor
Query editor v2
Query and database monitoring New
Datashares
Integrations
Zero-ETL integrations
S3 event integrations New
Editor
Query history
Saved queries
Scheduled queries
Resources Info
Status Connected
database sales-cluster
user admin
Change connection
Select database Info
To view schemas, select a database.
sales-cluster
Select schema Info
To view tables, select a schema.
public
Filter tables
fact_sales
sale_id
store_id
product_id
quantity
unit_price
sale_timestamp
channel
Query 1
CREATE TABLE public.fact_sales (
sale_id INT,
store_id INT,
product_id INT,
quantity INT,
unit_price DECIMAL(10,2),
sale_timestamp TIMESTAMP,
channel VARCHAR(10)
)
DISTKEY(store_id)
SORTKEY(sale_timestamp);
Run
Save
Schedule
Clear
Send feedback
Query results
Table details
CloudShell
Feedback
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences
Air: Moderate
Now
Search
ENG
IN
10:41
10-11-2025