

AWS SQL Querying Data Project

Creating Buckets: Used Standard Bucket (for using current data usage)

Bucket name:

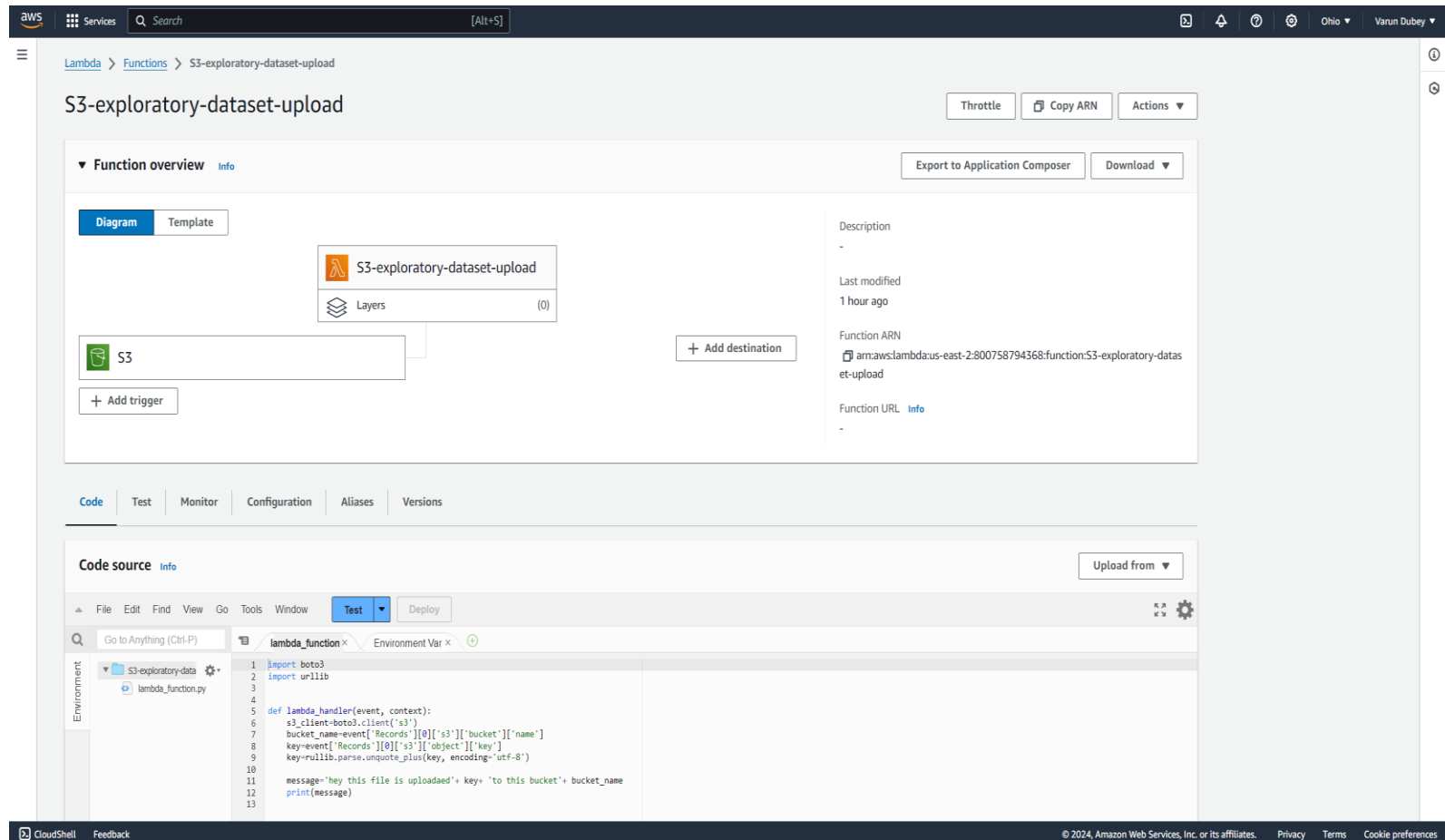
1.)exploratory-dataset-bucket (empty source bucket)

2.) result-exploratory-dataset-bucket (empty destination bucket)

The screenshot shows the AWS Management Console interface for the 'Buckets' page under 'Amazon S3'. The top navigation bar includes the AWS logo, 'Services', a search bar, and user information 'Varun Dubey'. The main content area features an 'Account snapshot' section with a 'View Storage Lens dashboard' button. Below this, there are tabs for 'General purpose buckets' (selected) and 'Directory buckets'. The 'General purpose buckets (2)' section includes a search bar, a 'Find buckets by name' input, and a table listing two buckets. The table has columns for Name, AWS Region, Access, and Creation date. Both buckets are in the 'US East (Ohio) us-east-2' region and have 'Bucket and objects not public' access. The page also includes a 'Create bucket' button and pagination controls showing 1 page.

Name	AWS Region	Access	Creation date
exploratory-dataset-bucket	US East (Ohio) us-east-2	Bucket and objects not public	January 27, 2024, 13:40:23 (UTC-05:00)
result-exploratory-dataset-bucket	US East (Ohio) us-east-2	Bucket and objects not public	January 27, 2024, 13:41:21 (UTC-05:00)

Defined Lambda Function: It triggers, and logs when we upload csv file in the source table.
i.e exploratory-dataset-bucket.



AWS cloud Watch: This records log files when the file is being uploaded in exploratory-dataset-bucket .

The screenshot displays the AWS CloudWatch console interface. On the left is a navigation sidebar with sections like 'CloudWatch', 'Alarms', 'Logs', 'Metrics', 'X-Ray traces', 'Events', 'Application Signals', 'Network monitoring', and 'Insights'. The main area shows the breadcrumb path: 'CloudWatch > Log groups > /aws/lambda/S3-exploratory-dataset-upload > 2024/01/27/[\${LATEST}]d79825ece6c54fa795794e48429e3c0c'.

Log events

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](#)

Buttons: [Refresh] [Actions ▼] [Start tailing] [Create metric filter]

Filter bar: [Filter events] [Clear] [1m] [30m] [1h] [12h] [Custom] [Local timezone ▼] [Display ▼] [Settings ⚙️]

▶	Timestamp	Message
No more records within selected time range Retry		
▶	2024-01-27T14:17:12.212-05:00	INIT_START Runtime Version: python:3.9.v44 Runtime Version ARN: arn:aws:lambda:us-east-2::runtime:3fee1447de2c6c6b67487c8862ea51f5b376c221844a0e7711..
▶	2024-01-27T14:17:12.514-05:00	START RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f Version: \$LATEST
▶	2024-01-27T14:17:14.558-05:00	[ERROR] NameError: name 'rullib' is not defined Traceback (most recent call last): File "/var/task/lambda_function.py", line 9, in lambda_handler ..
▶	2024-01-27T14:17:14.579-05:00	END RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f
▶	2024-01-27T14:17:14.579-05:00	REPORT RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f Duration: 2064.91 ms Billed Duration: 2065 ms Memory Size: 128 MB Max Memory Used: 74 MB Init D..
▶	2024-01-27T14:18:20.744-05:00	START RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f Version: \$LATEST
▶	2024-01-27T14:18:20.823-05:00	[ERROR] NameError: name 'rullib' is not defined Traceback (most recent call last): File "/var/task/lambda_function.py", line 9, in lambda_handler ..
▶	2024-01-27T14:18:20.842-05:00	END RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f
▶	2024-01-27T14:18:20.842-05:00	REPORT RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f Duration: 97.34 ms Billed Duration: 98 ms Memory Size: 128 MB Max Memory Used: 75 MB
▶	2024-01-27T14:20:15.986-05:00	START RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f Version: \$LATEST
▶	2024-01-27T14:20:16.061-05:00	[ERROR] NameError: name 'rullib' is not defined Traceback (most recent call last): File "/var/task/lambda_function.py", line 9, in lambda_handler ..
▶	2024-01-27T14:20:16.062-05:00	END RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f
▶	2024-01-27T14:20:16.063-05:00	REPORT RequestId: b04e10df-5bc5-49f5-9b74-469808ff041f Duration: 76.04 ms Billed Duration: 77 ms Memory Size: 128 MB Max Memory Used: 75 MB
No more records within selected time range Auto retry paused. Resume		

AWS Source Bucket: Source Bucket after loading csv data file (California-housing.csv).

aws

Services

Search

[Alt+S]

Global

Varun Dubey

Amazon S3

Buckets

exploratory-dataset-bucket

exploratory-dataset-bucket

Info

Objects

Properties

Permissions

Metrics

Management

Access Points

Objects (1)

Info

Refresh

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

< 1 >

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	 California-housing.csv	csv	January 27, 2024, 14:17:11 (UTC-05:00)	1.1 MB	Standard

CloudShell

Feedback

© 2024, Amazon Web Services, Inc. or its affiliates.

Privacy

Terms

Cookie preferences

AWS Athena: SQL query executed on California-housing dataset stored in exploratory-dataset-bucket.

aws

Services

Q s3

Ohio

Varun Dubey

Amazon Athena > Query editor tabs

EditorRecent queriesSaved queriesSettings

Workgroupresult-exploratory-...

> Query 3 :

1

2

3

SELECT * FROM "default"."exploratory_dataset_bucket"
where population>500

SQLLn 1, Col 1

Run again

Explain

Cancel

Clear

Create

Reuse query results

up to 60 minutes ago

Query resultsQuery status

Completed

Time in queue: 137 msRun time: 735 msData scanned: 1.12 MB

Results (18,660)

Copy

Download results

Search rows

< 1 ... >

housing_median_ag

total_rooms

total_bedrooms

population

household

median_incom

median_house

AWS Destination Bucket: Destination Bucket(result-exploratory-dataset-bucket) after running SQL query result of AWS Athena saved (processed_california-_housing .csv).

aws

Services

Search

[Alt+S]

Global

Varun Dubey

Amazon S3

Buckets

result-exploratory-dataset-bucket

California-result-dataset/

California-result-dataset/

Copy S3 URI

ObjectsProperties

Objects (2) Info

Refresh

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

< 1 >

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	 c64681ab-eede-4e61-b7c7-0236b6a8c49a.csv.metadata	metadata	January 27, 2024, 15:40:23 (UTC-05:00)	566.0 B	Standard
<input type="checkbox"/>	 processed_california-_housing .csv	csv	January 27, 2024, 15:40:52 (UTC-05:00)	1.4 MB	Standard

CloudShell

Feedback

© 2024, Amazon Web Services, Inc. or its affiliates.

Privacy

Terms

Cookie preferences