

S3 buckets | S3 | Global x +

s3.console.aws.amazon.com/s3/buckets?region=us-east-1&bucketType=general®ion=us-east-1

aWS Services Search [Option+S] Global voctabs/user3033498=9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92...

Amazon S3 > Buckets

Account snapshot

Last updated: Mar 15, 2024 by Storage Lens. Metrics are generated every 24 hours. Metrics don't include directory buckets. [Learn more](#)

Total storage Object count Average object size You can enable advanced metrics in the "default-account-dashboard" configuration.

33.3 MB 164 207.9 KB

[View Storage Lens dashboard](#)

General purpose buckets (4) [Info](#)

Buckets are containers for data stored in S3.

< 1 > ⚙️

[Create bucket](#)

Name	AWS Region	Access	Creation date
aws-glue-assets-924104828781-us-east-1	US East (N. Virginia) us-east-1	Bucket and objects not public	February 21, 2024, 19:36:26 (UTC-06:00)
project-stedi-lake-house-rishi	US East (N. Virginia) us-east-1	Bucket and objects not public	March 16, 2024, 19:45:09 (UTC-05:00)
rishitabucket	US East (N. Virginia) us-east-1	Objects can be public	January 29, 2024, 18:12:44 (UTC-06:00)
stedi-lake-house-rishi	US East (N. Virginia) us-east-1	Bucket and objects not public	February 21, 2024, 14:20:19 (UTC-06:00)

CloudShell Feedback © 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Screenshot of the AWS S3 console showing the 'Objects' tab for the bucket 'project-stedi-lake-house-rishi'. The 'Create folder' button is highlighted with a green box.

The browser address bar shows: s3.console.aws.amazon.com/s3/buckets/project-stedi-lake-house-rishi?region=us-east-1&bucketType=general&tab=objects

The AWS navigation bar includes: Services, Search, [Option+S], Global, and a user profile.

The breadcrumb navigation shows: Amazon S3 > Buckets > project-stedi-lake-house-rishi

The page title is: project-stedi-lake-house-rishi [Info](#)

The top navigation bar for objects includes: Objects, Properties, Permissions, Metrics, Management, Access Points, Actions, Create folder, and Upload.

The main content area displays the following table:

<input checked="" type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input checked="" type="checkbox"/>	accelerometer/	Folder	-	-	-
<input checked="" type="checkbox"/>	customer/	Folder	-	-	-
<input checked="" type="checkbox"/>	step_trainer/	Folder	-	-	-

A green box highlights the 'Create folder' button in the top right of the object list area.

Screenshot of the AWS S3 console showing the 'customer/' folder contents of the 'project-stedi-lake-house-rishi' bucket.

The browser address bar shows: `s3.console.aws.amazon.com/s3/buckets/project-stedi-lake-house-rishi?region=us-east-1&bucketType=general&prefix=customer/&showversions=false`

The AWS S3 navigation path is: `Amazon S3 > Buckets > project-stedi-lake-house-rishi > customer/`

The main interface displays the following details:

- Actions Bar:** Includes 'Objects (3) Info', 'Copy S3 URI', 'Copy URL', 'Download', 'Open', 'Delete', 'Actions ▾', 'Create folder' (highlighted with a green circle), and 'Upload'.
- Search Bar:** 'Find objects by prefix'.
- Table Headers:** Name, Type, Last modified, Size, Storage class.
- Table Data:** Three items under the 'customer/' folder:
 - curated/ (Folder)
 - landing/ (Folder)
 - trusted/ (Folder)

At the bottom of the page, there are links for CloudShell, Feedback, © 2024 Amazon Web Services, Inc. or its affiliates, Privacy, Terms, and Cookie preferences.

Screenshot of the AWS S3 console showing the 'landing/' folder contents in the 'customer/landing/' bucket.

The URL in the browser is: s3.console.aws.amazon.com/s3/buckets/project-stedi-lake-house-rishi?region=us-east-1&bucketType=general&prefix=customer/landing/&showversions=false

The 'Upload' button in the toolbar is circled in green.

Name	Type	Last modified	Size	Storage class
customer-1691348231425.json	json	March 16, 2024, 19:58:49 (UTC-05:00)	286.9 KB	Standard

AWS Services Search [Option+S] Global v vocabs/user3033498=9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92...

Amazon S3 > Buckets > project-stedi-lake-house-rishi > accelerometer/

accelerometer/

Copy S3 URI

Objects Properties

Objects (2) Info C 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 checked="" type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input checked="" type="checkbox"/>	landing/	Folder	-	-	-
<input checked="" type="checkbox"/>	trusted/	Folder	-	-	-

Screenshot of the AWS S3 console showing the 'landing/' folder in the 'accelerometer' bucket of the 'project-stedi-lake-house-rishi' bucket. The 'Upload' button in the top right of the object list is highlighted with a green circle.

Objects (9) [Info](#)

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](#)

<input checked="" type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input checked="" type="checkbox"/>	accelerometer-1691348231445.json	json	March 16, 2024, 19:59:27 (UTC-05:00)	825.2 KB	Standard
<input checked="" type="checkbox"/>	accelerometer-1691348231495.json	json	March 16, 2024, 19:59:28 (UTC-05:00)	825.4 KB	Standard
<input checked="" type="checkbox"/>	accelerometer-1691348231576.json	json	March 16, 2024, 19:59:28 (UTC-05:00)	826.5 KB	Standard
<input checked="" type="checkbox"/>	accelerometer-1691348231724.json	json	March 16, 2024, 19:59:29 (UTC-05:00)	825.4 KB	Standard
<input checked="" type="checkbox"/>	accelerometer-1691348231810.json	json	March 16, 2024, 19:59:30 (UTC-05:00)	826.0 KB	Standard
<input checked="" type="checkbox"/>	accelerometer-1691348231881.json	json	March 16, 2024, 19:59:30 (UTC-05:00)	826.3 KB	Standard
<input checked="" type="checkbox"/>	accelerometer-	json	March 16, 2024, 19:59:31 (UTC-05:00)	824.9 KB	Standard

[CloudShell](#) [Feedback](#) © 2024, Amazon Web Services, Inc. or its affiliates. [Privacy](#) [Terms](#) [Cookie preferences](#)

The screenshot shows the AWS S3 console interface. The URL in the browser is s3.console.aws.amazon.com/s3/buckets/project-stedi-lake-house-rishi?region=us-east-1&bucketType=general&prefix=step_trainer/&showversions=false. The page displays the contents of the 'step_trainer' folder within the 'project-stedi-lake-house-rishi' bucket. There are three objects listed: 'curated/' (Folder), 'landing/' (Folder), and 'trusted/' (Folder). The 'Actions' menu at the top right includes options like 'Create folder' (which is highlighted with a green box), 'Upload', 'Copy S3 URI', 'Copy URL', 'Download', 'Open', 'Delete', and 'Create folder'. A search bar at the bottom left allows filtering by prefix.

Name	Type	Last modified	Size	Storage class
curated/	Folder	-	-	-
landing/	Folder	-	-	-
trusted/	Folder	-	-	-

Screenshot of the AWS S3 console showing the contents of the 'landing/' folder in the 'step_trainer' bucket.

The URL in the browser is: https://s3.console.aws.amazon.com/s3/buckets/project-stedi-lake-house-rishi?region=us-east-1&bucketType=general&prefix=step_trainer/landing/&showversions=false

The page shows three objects in the 'Objects' list:

Name	Type	Last modified	Size	Storage class
step_trainer-1691348232038.json	json	March 16, 2024, 20:00:04 (UTC-05:00)	1.1 MB	Standard
step_trainer-1691348232085.json	json	March 16, 2024, 20:00:04 (UTC-05:00)	1.1 MB	Standard
step_trainer-1691348232132.json	json	March 16, 2024, 20:00:05 (UTC-05:00)	974.8 KB	Standard

The 'Upload' button in the top right of the object list is highlighted with a green box.

Databases - AWS Glue Console

us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/databases

AWS Services Search [Option+S] N. Virginia voclabs/user3033498=9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92...

AWS Glue

Getting started ETL jobs Visual ETL Notebooks Job run monitoring Data Catalog tables Data connections Workflows (orchestration)

Data Catalog

Databases

- Tables
- Stream schema registries
- Schemas
- Connections
- Crawlers
- Classifiers
- Catalog settings

▶ Data Integration and ETL

▶ Legacy pages

What's New Documentation AWS Marketplace

Enable compact mode

CloudShell Feedback

Databases (1/4) Last updated (UTC) March 17, 2024 at 01:06:57

A database is a set of associated table definitions, organized into a logical group.

Name	Description	Location URI	Created on (UTC)
default	Default Hive database	file:/user/hive/warehouse	February 26, 2024 at 19:18:04
stedi	-	-	February 26, 2024 at 18:35:29
stedi_project	-	-	March 17, 2024 at 00:57:11
stedi2	-	-	February 26, 2024 at 22:26:00

Add database

© 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preference

Tables - AWS Glue Console

us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/tables/add/step/1

AWS Services Search [Option+S] N. Virginia v ocilabs/user3033498=9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92...

AWS Glue

- Getting started
- ETL jobs
 - Visual ETL
 - Notebooks
 - Job run monitoring
- Data Catalog tables
- Data connections
- Workflows (orchestration)

Data Catalog

Databases

- Tables
- Stream schema registries
- Schemas
- Connections
- Crawlers
 - Classifiers
- Catalog settings

Data Integration and ETL

Legacy pages

What's New Documentation AWS Marketplace

Add table

Set table properties

Table details

Name: customer_landing

If you plan to access the table from Amazon Athena, then the name should be under 256 characters and contain only lowercase letters (a-z), numbers (0-9), and underscore (_). For more information, see [Athena names](#).

Database: stedi_project

Create database

Description - optional

Enter a description

Descriptions can be up to 2048 characters long.

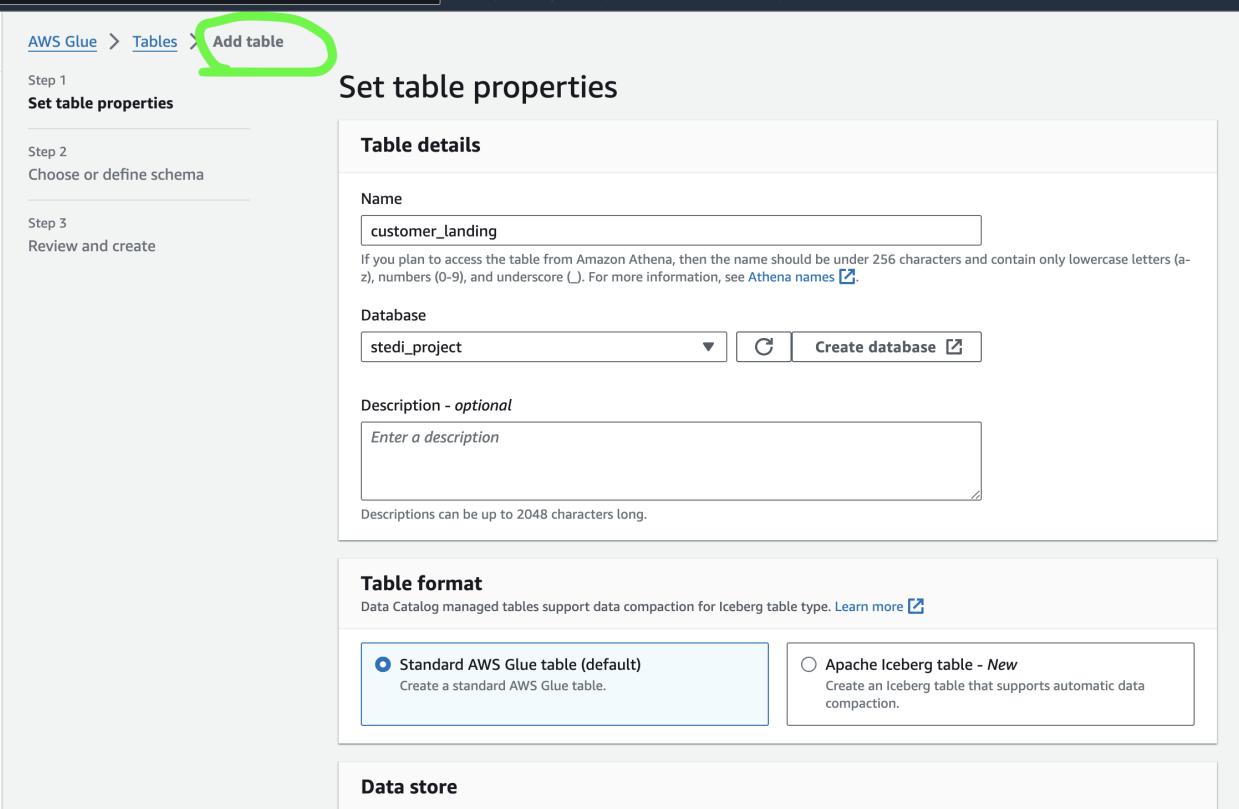
Table format

Data Catalog managed tables support data compaction for Iceberg table type. [Learn more](#)

Standard AWS Glue table (default)
Create a standard AWS Glue table.

Apache Iceberg table - New
Create an Iceberg table that supports automatic data compaction.

Data store



Tables - AWS Glue Console x +

us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/tables/add/step/1

aws Services Search [Option+S] N. Virginia vocompaction: voclabs/user3033498=9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92... ▾

AWS Glue x

Getting started
ETL jobs
 Visual ETL
 Notebooks
 Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)

▼ Data Catalog

Databases
 Tables
 Stream schema registries
 Schemas
Connections
Crawlers
 Classifiers
Catalog settings

► Data Integration and ETL

► Legacy pages

What's New Documentation AWS Marketplace

Enable compact mode

Data store

Select the type of source
 S3
 Kinesis
 Kafka

Data location is specified in
 my account
 another account

Include path
 li-lake-house-rishi/customer/landing/ X View Browse S3

Path must be in the form s3://bucket/prefix/. It must end with a slash (/) and not include any files.

Data format

Classification
Choose the format of the data in your table.
 Avro
 CSV
 JSON
 XML
 Parquet
 ORC

Cancel Next



Tables - AWS Glue Console

Table Detail - AWS Glue Console

us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/tables/add/step/2

Services Search [Option+S]

N. Virginia voclabs/user3033498=9ea2df2-322b-11ee-a55d-535da5610840 @ 92...

Step 1 Set table properties

Step 2 Choose or define schema

Step 3 Review and create

Choose or define schema

Schema

Define or upload schema Manually define schema

Choose from Glue Schema Registry Select existing schema from your Glue Schema Registry.

Schema (1/10)
View and manage the table schema.

#	Column name	Data type	Partition key	Comment
1	customername	string	-	-
2	email	string	-	-
3	phone	string	-	-
4	birthday	string	-	-
5	serialnumber	string	-	-
6	registrationdate	bigint	-	-
7	lastupdatedate	bigint	-	-
8	sharewithresearchasofdate	bigint	-	-
9	sharewithpublicasofdate	bigint	-	-
10	sharewithfriendsasofdate	bigint	-	-

Partition indexes - optional (0)
AWS Glue partition indexes are an important configuration to reduce overall data transfers and processing, and reduce query processing time. A maximum of 3 partition indexes can be created on a given table.

Index name	Index keys	Index status
No indexes Provide indexes Add indexes		

Cancel Previous **Next**

CloudShell Feedback © 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

us-east-1.console.aws.amazon.com/glue/home?region=us-east-1#/v2/data-catalog/tables/add/step/3

AWS Services Search [Option+S] N. Virginia v oclabs/user3033498=9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92...

Review and create

Step 1: Setup table properties

Table properties

Table name customer_landing	Database stedi_project	Description -
--------------------------------	---------------------------	------------------

Data store and format

Type of source S3	Include path s3://project-stedi-lake-house-rishi/customer/landing/	Classification JSON
----------------------	---	------------------------

Step 2: Choose or define schema

Schema (10)
View and manage the table schema.

#	Column name	Data type	Partition key	Comment
1	customername	string	-	-
2	email	string	-	-
3	phone	string	-	-
4	birthday	string	-	-
5	serialnumber	string	-	-
6	registrationdate	bigint	-	-
7	lastupdatedate	bigint	-	-
8	sharewithresearchersofdate	bigint	-	-
9	sharewithpublicasofdate	bigint	-	-
10	sharewithfriendsasofdate	bigint	-	-

Partition indexes - optional (0)
AWS Glue partition indexes are an important configuration to reduce overall data transfers and processing, and reduce query processing time. A maximum of 3 partition indexes can be created on a given table.

Index name	Index keys	Index status
	No indexes Provide indexes	Add indexes

Cancel Previous Create

AWS Services Search [Option+S] N. Virginia vocabs/user3033498-9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92... ▾

AWS Glue > Tables > Edit table

Edit table

Table details

Name: customer_landing

Input format: org.apache.hadoop.mapred.TextInputFormat

Output format: org.apache.hadoop.hive.io.HiveIgnoreKeyTextOutputFormat

Serde name:

Serialization lib: org.openx.data.jsonserde.JsonSerDe

Description:

Serde parameters

Key: serialization.format Value: 1 Remove Add

<https://repost.aws/knowledge-center/athena-failed-nullpointerexception-error>

Table properties

Key: classification Value: json Remove Add

Cancel Save

AWS Services Search [Option+S] N. Virginia vocabs/user3033498=9ea2dfd2-322b-11ee-a55d-535da5610840 @ 92...

Amazon Athena > Query editor

Editor Recent queries Saved queries Settings Workgroup primary

Data

Data source: AwsDataCatalog Database: stedi

Tables and views: Create Filter tables and views

Tables (5): accelerometer_landing, accelerometer_trusted, customer_curated, customer_landing, customer_trusted

Views (0)

Query 15: SHOW CREATE TABLE `customer_landing`;

SQL Ln 1, Col 38

Run again Explain Cancel Clear Create

Reuse query results up to 60 minutes ago

Completed Time in queue: 194 ms Run time: 882 ms Data scanned: -

```
CREATE EXTERNAL TABLE `customer_landing` (
  `customername` string COMMENT 'from deserializer',
  `email` string COMMENT 'from deserializer',
  `phone` string COMMENT 'from deserializer',
  `birthday` string COMMENT 'from deserializer',
  `serialnumber` string COMMENT 'from deserializer',
  `registrationdate` bigint COMMENT 'from deserializer',
  `lastupdate` bigint COMMENT 'from deserializer',
  `sharewithresearcherdate` bigint COMMENT 'from deserializer',
  `sharewithpublicasofdate` bigint COMMENT 'from deserializer',
  `sharewithfriendsasofdate` bigint COMMENT 'from deserializer')
ROW FORMAT SERDE
  'org.openx.data.jsonserde.JsonSerDe'
STORED AS INPUTFORMAT
  'org.apache.hadoop.mapred.TextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION
  's3://stedi-lake-house-rishi/customer/landing/'
TBLPROPERTIES (
  'classification'='json')
```

Step 1
Set table properties

Step 2
Choose or define schema

Step 3
Review and create

Set table properties

Table details

Name

If you plan to access the table from Amazon Athena, then the name should be under 256 characters and contain only lowercase letters (a-z), numbers (0-9), and underscore (_). For more information, see [Athena names](#).

Database

[Create database](#)

Description - optional

Enter a description

Descriptions can be up to 2048 characters long.

Table format

Data Catalog managed tables support data compaction for Iceberg table type. Learn more [\[?\]](#)

Standard AWS Glue table (default)

Create a standard AWS Glue table.

Apache Iceberg table - New

Create an Iceberg table that supports automatic data compaction.

Data store

Select the type of source

- S3
- Kinesis
- Kafka

Data location is specified in

- my account
- another account

Include path

Path must be in the form s3://bucket/prefix/. It must end with a slash (/) and not include any files.

[View](#)

[Browse S3](#)

Data format

Classification

Choose the format of the data in your table.

- Avro
- CSV
- JSON
- XML
- Parquet
- ORC

[Cancel](#)

Next

[Option+S] N. Virginia v vocabs/user3033498=9ea2df2-322b-11ee-a55d-535da5610840@92...

AWS Glue > Tables > Add table

Step 1 Set table properties

Step 2 Choose or define schema

Step 3 Review and create

Choose or define schema

Schema

Define or upload schema Manually define schema

Choose from Glue Schema Registry Select existing schema from your Glue Schema Registry.

Schema (1/5)
View and manage the table schema.

#	Column name	Data type	Partition key	Comment
1	user	string	-	-
2	timeStamp	bigint	-	-
3	x	float	-	-
4	y	float	-	-
5	z	float	-	-

Partition indexes - optional (0)
AWS Glue partition indexes are an important configuration to reduce overall data transfers and processing, and reduce query processing time. A maximum of 3 partition indexes can be created on a given table.

Index name	Index keys	Index status
	No indexes Provide indexes	Add indexes

Cancel Previous Next



Step 1
[Set table properties](#)
 Step 2
[Choose or define schema](#)
 Step 3
Review and create

Review and create

Step 1: Setup table properties

[Edit table properties](#)

Table properties

Table name	Database	Description
accelerometer_landing	stedi_project	-

Data store and format

Type of source	Include path	Classification
S3	s3://project-stedi-lake-house-rishi/accelerometer/landing/	JSON

Step 2: Choose or define schema

[Edit schema](#)

Schema (5)

View and manage the table schema.

<input type="checkbox"/>	#	Column name	Data type	Partition key	Comment
<input type="checkbox"/>	1	user	string	-	-
<input type="checkbox"/>	2	timeStamp	bigint	-	-
<input type="checkbox"/>	3	x	float	-	-
<input type="checkbox"/>	4	y	float	-	-
<input type="checkbox"/>	5	z	float	-	-

Partition indexes - optional (0)

AWS Glue partition indexes are an important configuration to reduce overall data transfers and processing, and reduce query processing time. A maximum of 3 partition indexes can be created on a given table.

Index name	Index keys	Index status
No indexes Provide indexes Add indexes		

[Cancel](#) [Previous](#) [Create](#)


[Optional]



N. Virginia

volcabs/user/S033498-9ea7d02-522b-11ee-a55d-555da5610840@52...

One table successfully created

The following table is now created: "accelerometer_landing (db:stedi_project)"

AWS Glue > Tables > Edit table

Edit table

Table details

Name

accelerometer_landing

Input format

org.apache.hadoop.mapred.TextInputFormat

Output format

org.apache.hadoop.hive.io.HiveIgnoreKeyTextOutputFormat

Serde name

Serialization lib

org.openx.data.jsonserde.JsonSerDe

Description

Serde parameters

Key

serialization.format

Value

1

Remove

Add

Table properties

Key

classification

Value

json

Remove

Add

Cancel

Save

AWS Services Search [Option+5] N. Virginia v vociabu/userS03549B=9ea2d8d2-322b-11ee-a55d-535da5610840 @ 92...

One table successfully created
The following table is now created: "accelerometer_landing (db:stedi_project)".

AWS Glue > Tables > Add table

Step 1 Set table properties

Step 2 Choose or define schema

Step 3 Review and create

Table details

Name: step_trainer_landing
If you plan to access the table from Amazon Athena, then the name should be under 256 characters and contain only lowercase letters (a-z), numbers (0-9), and underscore (_). For more information, see [Athena names](#).

Database: stedi_project [Create database](#)

Description - optional: Enter a description
Descriptions can be up to 2048 characters long.

Table format
Data Catalog managed tables support data compaction for iceberg table type. Learn more

Standard AWS Glue table (default)
Create a standard AWS Glue table.

Apache Iceberg table - New
Create an iceberg table that supports automatic data compaction.

Data store

Select the type of source:
 S3
 Kinesis
 Kafka

Data location is specified in:
 my account
 another account

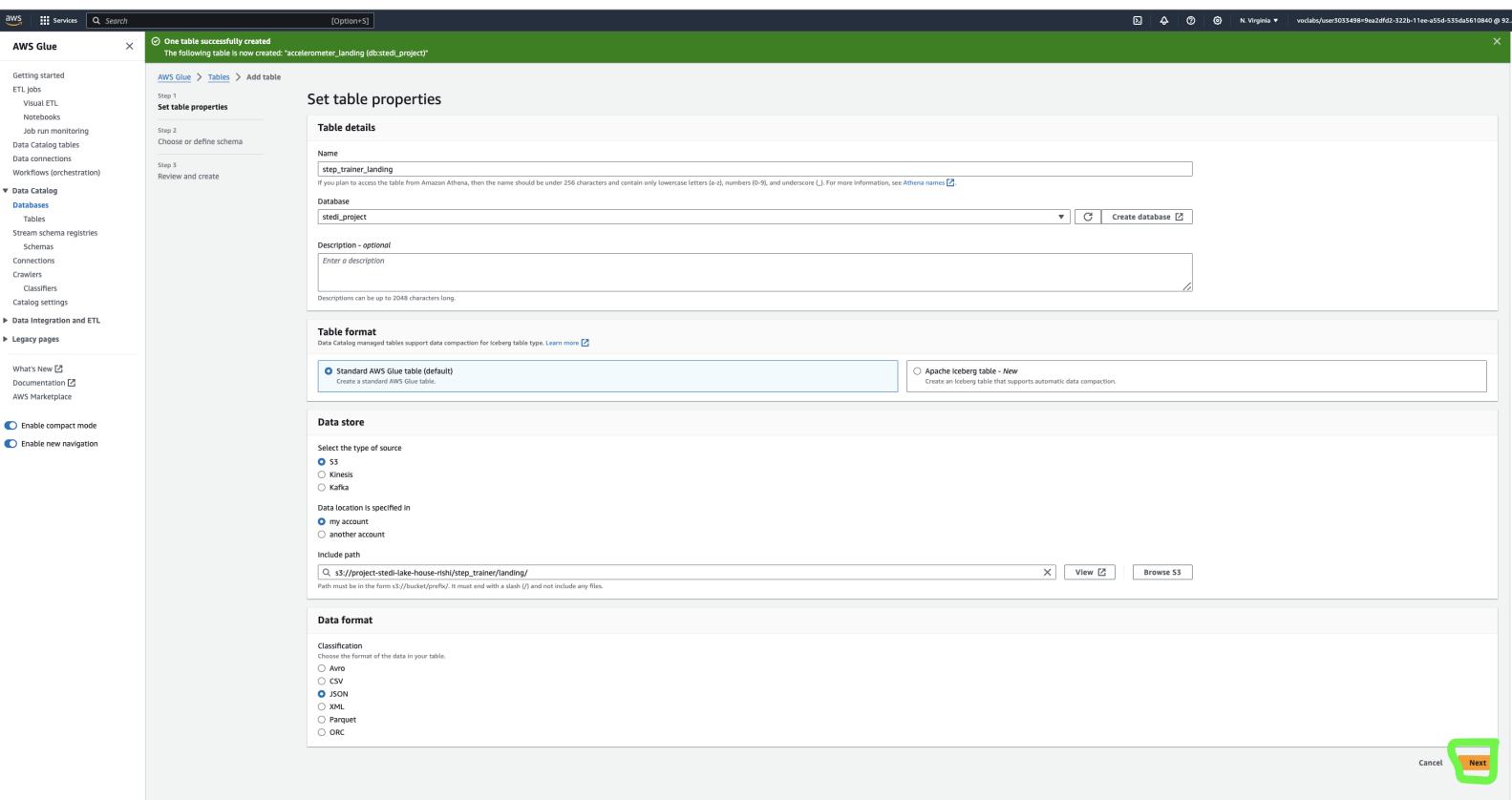
Include path:
 s3://project-stedi-lake-house-rishi/step_trainer/landing/ [View](#) [Browse S3](#)

Data format

Classification
Choose the format of the data in your table.

Avro
 CSV
 JSON
 XML
 Parquet
 ORC

Cancel **Next**



One table successfully created
The following table is now created: "accelerometer_landing (dbstedi_project)"

AWS Glue > Tables > Add table

Step 1 Set table properties

Step 2 Choose or define schema

Step 3 Review and create

Choose or define schema

Schema

Define or upload schema Manually define schema

Choose from Glue Schema Registry Select existing schemas from your Glue Schema Registry.

Schema (1/3)
View and manage the table schema.

#	Column name	Data type	Partition key	Comment
1	sensorReadingTime	bigint	-	-
2	serialNumber	string	-	-
3	distanceFromObject	int	-	-

Partition indexes - optional (0)
AWS Glue partition indexes are an important configuration to reduce overall data transfers and processing, and reduce query processing time. A maximum of 3 partition indexes can be created on a given table.

Index name	Index keys	Index status
	No Indexes Provide indexes	Add Indexes

Cancel Previous **Next**

One table successfully created
The following table is now created: "accelerometer_landing (db:stedi_project)"

AWS Glue > Tables > Add table

Step 1 Set table properties

Step 2 Choose or define schema

Step 3 Review and create

Review and create

Step 1: Setup table properties

Table properties

Table name	step_trainer_landing	Database	stedi_project	Description
------------	----------------------	----------	---------------	-------------

Data store and format

Type of source	S3	Include path	s3://project-stedi-lake-house-rishi/step_trainer/landing/	Classification	JSON
----------------	----	--------------	---	----------------	------

Step 2: Choose or define schema

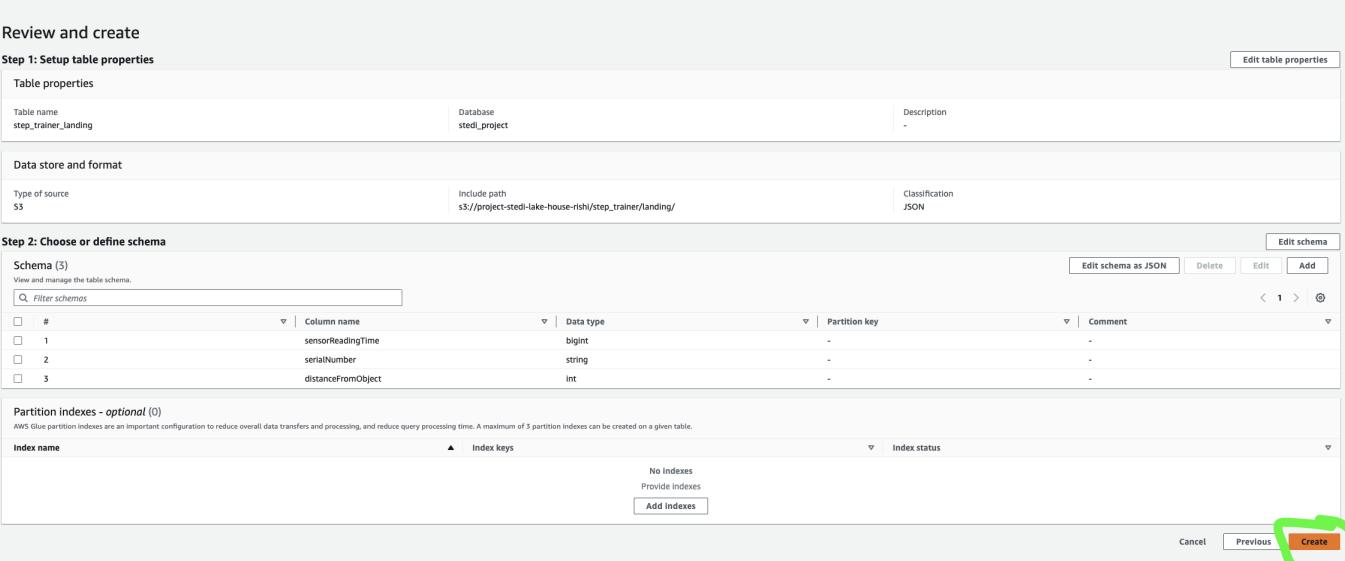
Schema (3)
View and manage the table schema.

#	Column name	Data type	Partition key	Comment
1	sensorReadingTime	bigint	-	-
2	serialNumber	string	-	-
3	distanceFromObject	int	-	-

Partition indexes - optional (0)
AWS Glue partition indexes are an important configuration to reduce overall data transfers and processing, and reduce query processing time. A maximum of 3 partition indexes can be created on a given table.

Index name	Index keys	Index status
	No Indexes Provide Indexes	Add Indexes

Cancel Previous Create



Some tables successfully created
The following tables are now created: "step_trainer_landing (db:stedl_project)", "accelerometer_landing (db:stedl_project)"

AWS Glue > Tables > Edit table

Edit table

Table details

Name: step_trainer_landing

Input format: org.apache.hadoop.mapred.TextInputFormat

Output format: org.apache.hadoop.hive.io.HiveIgnoreKeyTextOutputFormat

Serde name:

Serialization lib: org.openx.data.jsonserde.JsonSerDe

Description:

Serde parameters

Key	Value	Remove
serialization.format	1	<input type="button" value="Remove"/>

Add

Table properties

Key	Value	Remove
classification	json	<input type="button" value="Remove"/>

Cancel **Save**

Customer_Landing_to_Trusted

Visual Script Job details Runs Data quality - updated Schedules Version Control

Last modified on 5/18/2024, 4:51:58 PM Actions Save Run

Data source properties - S3

Name: Customer_Landing_Zone_Node

S3 source type: Info

S3 location: Choose a file or folder in an S3 bucket.

Data Catalog table

S3 URL: View Browse S3

Recursive: Read files in all subdirectories.

Data format: JSON

JsonPath - optional: Identify records with a JsonPath expression.

Multiline: Indicates if JSON records can span multiple lines.

Additional options

Data preview

No data to display. Try increasing the number of rows being sampled.

Filter sample dataset

Stack Trace

```
graph TD; A[Data source - S3 bucket Customer_Landing_Zone_Node] --> B[Transform - Filter Research_Filter]; B --> C[Data target - S3 bucket Customer_Trusted_Zone_Node]
```

AWS Glue Data Transformation Job: Customer_Landing_to_Trusted

The screenshot shows a Glue Data Transformation job named "Customer_Landing_to_Trusted". The visual interface displays a flow graph with three main components:

- Data source - S3 bucket:** Customer_Landing_Z...
- Transform - Filter:** Research_Filter
- Data target - S3 bucket:** Customer_Trusted_Z...

The flow starts with the data source, goes through the filter transform, and ends at the data target.

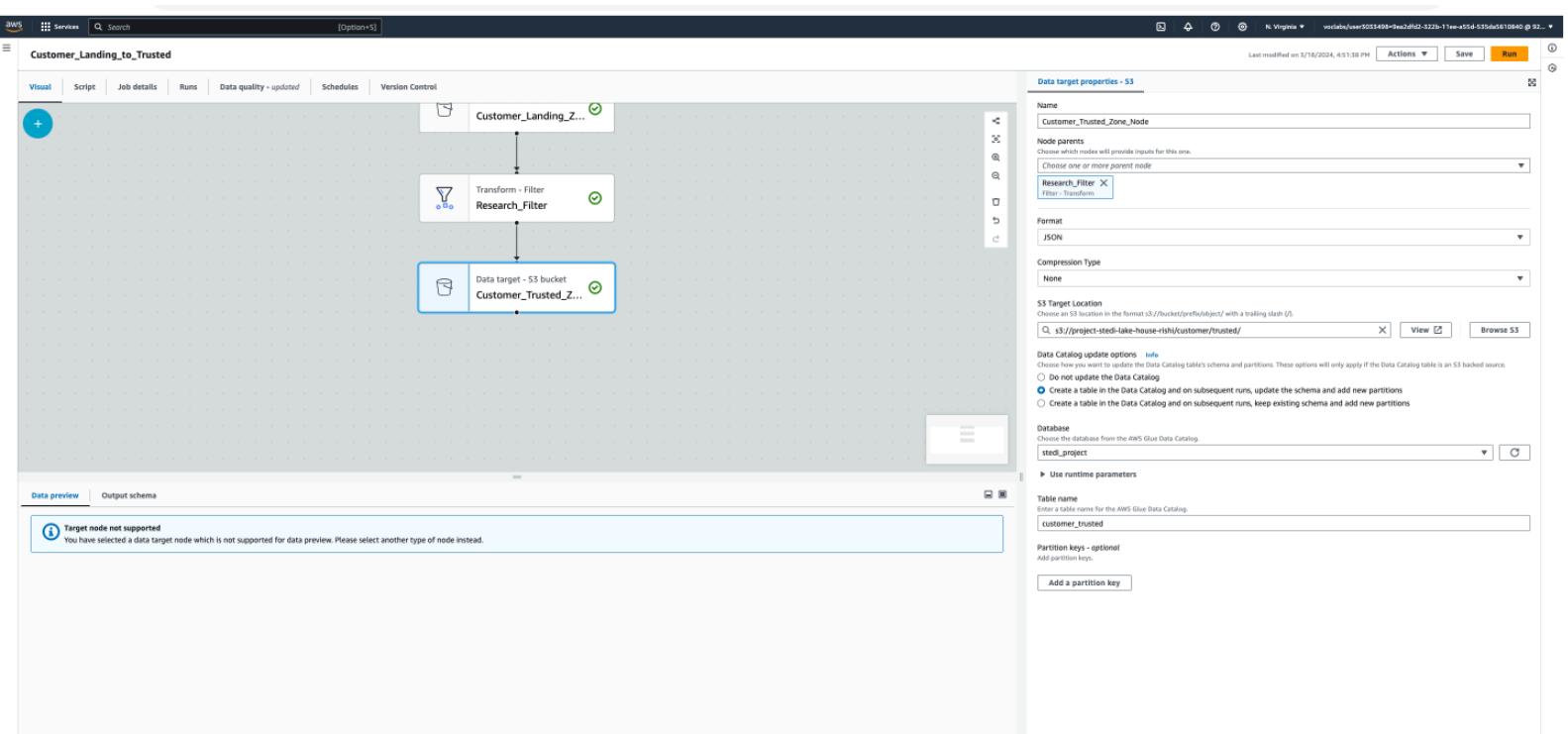
Transform Configuration:

- Name:** Research_Filter
- Node parents:** Choose one or more parent node. Selected: Customer_Landing_Zone_Node
- Filter Info:** Builds a new output by selecting records from the input data that satisfy a specified predicate function.
 - Global AND: All filter conditions will be applied as a global "AND".
 - Global OR: All filter conditions will be applied as a global "OR".
- Filter condition info:** Specify your filter condition by choosing the key operator, and entering a value.

Key	shareWithResearchAsOf...	Operation	=	Value	0
-----	--------------------------	-----------	---	-------	---

Data preview: No data to display. Try increasing the number of rows being sampled.

Output schema: Previewing 0 of 0 fields.



●	● Failed	0	2024/03/18 21:58:54	2024/03/18 22:01:12	2 m 7 s	10 DPU\$	G.1X	4.0
○	● Failed	0	2024/03/18 21:53:51	2024/03/18 21:55:47	1 m 46 s	10 DPU\$	G.1X	4.0

Run details	Input arguments (10)	Continuous logs	Run insights	Metrics	Spark UI		
● Error Category: RESOURCE_NOT_FOUND_ERROR; An error occurred while calling o117.pyWriteDynamicFrame. No such file or directory 's3://project-stedi-lake-house-rishi/customer/landing/customer-1691348231425.json'							
Job name Customer_Landing_to_Trusted	Start time (UTC) 2024/03/18 21:58:54	Glue version 4.0	Last modified on (UTC) 2024/03/18 22:01:12				
ID jr_b4e58d85952b0a459cae67c342cdaaaaab8e358aea042da28410d5cd2b17d3ad	End time (UTC) 2024/03/18 22:01:12	Worker type G.1X	Log group name /aws-glue/jobs				
Run status ● Failed	Start-up time 11 seconds	Max capacity 10 DPU\$	Number of workers 10				
Retry attempt number Initial run	Execution time 2 minutes 7 seconds	Execution class Standard	Timeout 2880 minutes				
Trigger name -	Security configuration -	Cloudwatch logs • All logs <input checked="" type="checkbox"/> • Output logs <input checked="" type="checkbox"/> • Error logs <input checked="" type="checkbox"/>					

SQS Services Search [Option+S]

IAM > Roles > my-glue-service-role > Edit policy

Step 1 Modify permissions in S3Access

Step 2 Review and save

Modify permissions in S3Access

Add permissions by selecting services, actions, resources, and conditions. Build permission statements using the JSON editor.

Policy editor

```
14 "Sid": "AllObjectActions",
15 "Effect": "Allow",
16 "Action": "s3:GetObject",
17 "Resource": [
18     "arn:aws:s3:::stedi-lake-house-rishi/*"
19 ],
20 },
21 ],
22 [
23     {
24         "Sid": "ListObjectsInBucketProj",
25         "Effect": "Allow",
26         "Action": [
27             "s3:ListBucket"
28         ],
29         "Resource": [
30             "arn:aws:s3:::project-stedi-lake-house-rishi"
31         ],
32     },
33     {
34         "Sid": "AllObjectActionsProj",
35         "Effect": "Allow",
36         "Action": "s3:GetObject",
37         "Resource": [
38             "arn:aws:s3:::project-stedi-lake-house-rishi/*"
39         ],
40     }
41 }
```

+ Add new statement

JSON Ln 22, Col 0

Security: 0 Errors: 0 Warnings: 0 Suggestions: 0

You need permissions
User: arn:aws:sts::924104828781:assumed-role/voclabs/userS033498=9ea2dfd2-322b-11ee-a55d-535da5610840 is not authorized to perform: access-analyzer:ValidatePolicy on resource: arn:aws:access-analyzer:us-east-1:924104828781.*

8399 of 8949 characters remaining

Check for new access

Cancel Next

Successfully started job Customer_Landing_to_Trusted. Navigate to Run details for more details.

Customer_Landing_to_Trusted

Last modified on 5/18/2024, 4:53:46 PM Actions Save Run

Run status	Retries	Start time (UTC)	End time (UTC)	Duration	Capacity (DPUs)	Worker type	Glue version
Success Succeeded	0	2024/03/18 22:14:06	2024/03/18 22:16:40	2 m 20 s	10 DPUs	G.1X	4.0
Failure Failed	0	2024/03/18 21:58:54	2024/03/18 22:01:12	2 m 7 s	10 DPUs	G.1X	4.0
Failure Failed	0	2024/03/18 21:53:51	2024/03/18 21:55:47	1 m 46 s	10 DPUs	G.1X	4.0

Run details Input arguments (10) Continuous logs Run insights Metrics Spark UI

Job name: Customer_Landing_to_Trusted Start time (UTC): 2024/03/18 22:14:06 Glue version: 4.0 Last modified on (UTC): 2024/03/18 22:16:40

Id: j_r1b7479013973188cab71c55275da43896b198d03679e1a96f5ebb691cf0b968 Log group name: /aws-glue/jobs

Run status: Success Succeeded End time (UTC): 2024/03/18 22:16:40 Worker type: G.1X

Retry attempt number: Initial run Start-up time: 14 seconds Max capacity: 10 DPUs Number of workers: 10

Execution time: 2 minutes 20 seconds Execution class: Standard Timeout: 2890 minutes

Trigger name: Security configuration Cloudwatch logs: All logs, Output logs, Error logs

Accelerometer_Landing_to_Trusted

Visual Script Job details Runs Data quality - updated Schedules Version Control

Actions Save Run

Data source properties - S3

Name Accelerometer_Landing_Zone_Node

S3 source type Info

S3 location Choose a file or folder in an S3 bucket.

Data Catalog table

S3 URL View Browse

Recursive Read files in all subdirectories.

Data format JSON

JsonPath - optional Identify records with a JsonPath expression.

Multiline Indicates if JSON records can span multiple lines.

Infer schema

Additional options

Data preview Output schema

Data preview (200) Info READY

Filter sample dataset

user	timestamp	x	y	z
Santosh.Clayton@test.com	165556444103	1	-1	-1
Santosh.Clayton@test.com	1655564440170	0	1	-1
Santosh.Clayton@test.com	1655564440170	-1	0	-1
Santosh.Clayton@test.co	^	^	^

Accelerometer_Landing_to_Trusted

Visual | Script | Job details | Runs | Data quality - updated | Schedules | Version Control

Data source properties - S3

Name: Customer_Trusted_Zone_Node

S3 source type: Info

S3 location: Choose a file or folder in an S3 bucket.

Data Catalog table

S3 URL: s3://project-stedi-lake-house-rishi/customer/trusted/

Recursive: Read files in all subdirectories.

Data format: JSON

JsonPath - optional: Identify records with a JsonPath expression.

Multiline: Indicates if JSON records can span multiple lines.

Infer schema

Additional options

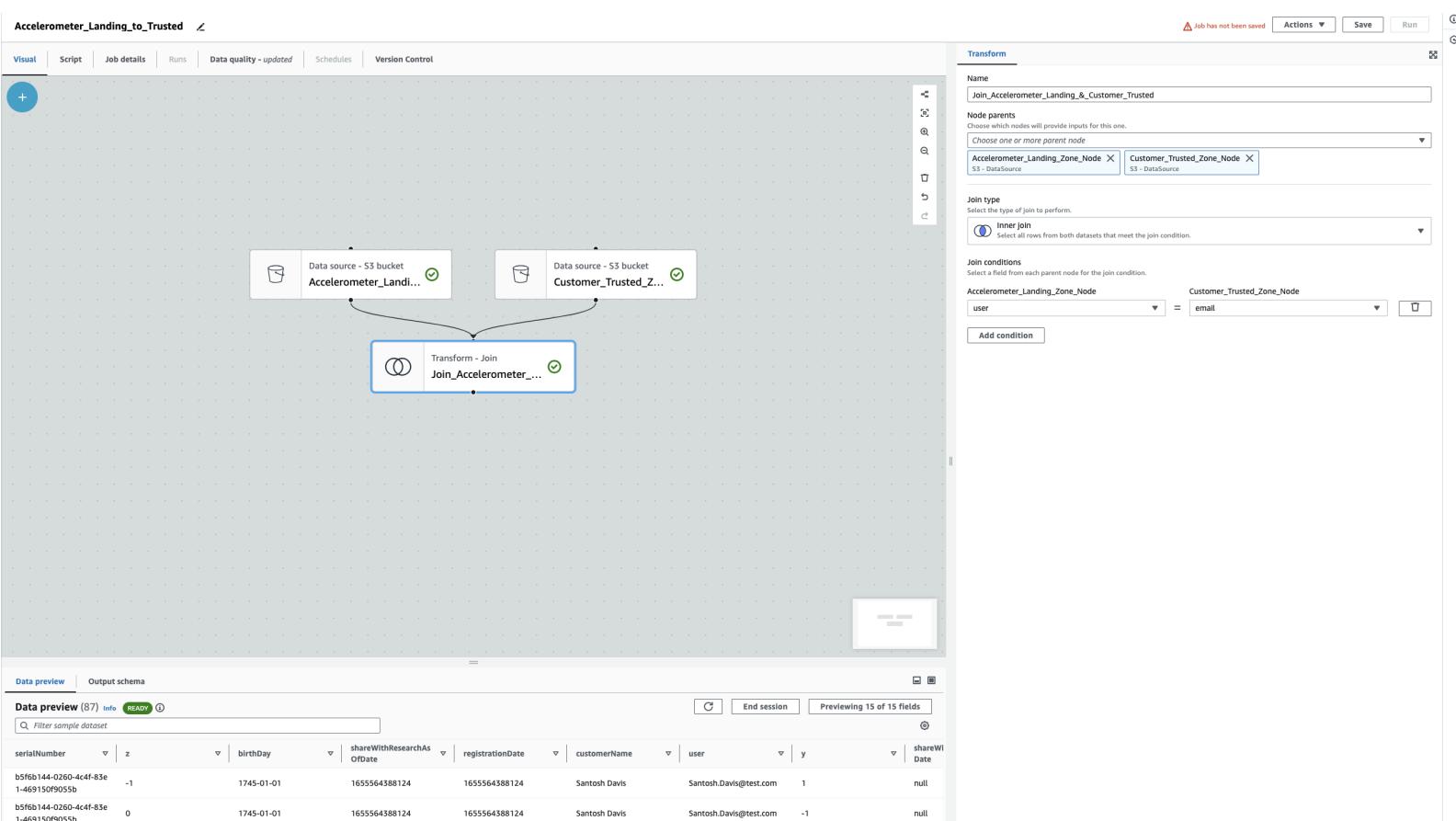
Data preview

Output schema

Data preview (200) Info READY

Filter sample dataset

serialnumber	birthday	registrationdate	sharewithresearchassociates	customername	sharewithfriendsasofdate	email	lastupdatedate	phone
50f7b4f3-7af5-4bd7-a42 1-7b902e8d2b7c	1900-01-01	1655564376361	1655564376361	Santosh Clayton	1655564376361	Santosh.Clayton@test.co m	1655564376361	801555
20400202-c1db-4328-8e 94-5bac7d61eb	1399-01-01	1655564414415	1655564414415	Ben Khatib	1655564414415	Ben.Khatib@test.com	1655564414415	801555
2b6fe805-9059-4e1b-b7 ed-720f5b4bf10c	1141-01-01	1655564433612	1655564433612	Jane Gonzalez	1655564433612	Jane.Gonzalez@test.com	1655564433612	801555



Screenshot of the AWS Glue Data Catalog interface showing a job flow named "Accelerometer_Landing_to_Trusted".

Job Status: Successfully started job Accelerometer_Landing_to_Trusted. Navigate to [Run details](#) for more details.

Job Details:

- Visual
- Script
- Job details
- Runs
- Data quality - updated
- Schedules
- Version Control

Transform:

```

graph TD
    A["Data source - S3 bucket  
Accelerometer_Landing_..."] --> C["Transform - Join  
Join_Accelerometer_..."]
    B["Data source - S3 bucket  
Customer_Trusted_Z..."] --> C
    C --> D["Transform - DropFields  
Drop_Fields_of_Cust..."]
    D --> E["Data target - S3 bucket  
Accelerometer_Trust..."]
  
```

DropFields:

Field	Data type
serialnumber	string
birthday	string
registrationdate	bigint
x	double
y	double
z	double
shareWithResearchersofDate	bigint
customerName	string
shareWithFriendsofDate	bigint
email	string
lastupdatedate	bigint
phone	string
shareWithPublicasofDate	bigint

Data preview: Previewing 13 of 13 fields

serialNumber	z	birthDay	shareWithResearchersofDate	registrationDate	customerName	user	y	shareWithPublicasofDate
b5f6b144-0260-4c4f-83e 1-4691509055b	-1	1745-01-01	1655564388124	1655564388124	Santosh Davis	Santosh.Davis@test.com	1	null
b5f6b144-0260-4c4f-83e 1-4691509055b	0	1745-01-01	1655564388124	1655564388124	Santosh Davis	Santosh.Davis@test.com	-1	null
b5f6b144-0260-4c4f-83e 1-4691509055b	-1	1745-01-01	1655564388124	1655564388124	Santosh Davis	Santosh.Davis@test.com	-1	null

AWS Glue job visualization and configuration interface.

Job Status: Successfully started job Accelerometer_Landing_to_Trusted. Navigate to [Run details](#) for more details.

Job Overview: Accelerometer_Landing_to_Trusted

Visual Editor: Shows the data flow graph:

```

graph TD
    A[Data source - S3 bucket Accelerometer_Landing] --> B[Transform - Join Join_Accelerometer...]
    C[Data source - S3 bucket Customer_Trusted_Z...] --> B
    B --> D[Transform - DropFields Drop_Fields_of_Cust...]
    D --> E[Data target - S3 bucket Accelerometer_Trust...]
  
```

Data Target Properties:

- Name: Accelerometer_Trusted_Zone_Node
- Node parents: Drop_Fields_of_Customer_Trusted
- Format: JSON
- Compression Type: None
- S3 Target Location: s3://project-stedi-lake-house-rish/accelerometer/trusted/
- Data Catalog update options: Do not update the Data Catalog
- Table name: accelerometer_trusted
- Partition keys - optional: Add partition keys

Job Details:

- Data preview: Target node not supported. You have selected a data target node which is not supported for data preview. Please select another type of node instead.
- Output schema: None

Bottom Right:

- Unsaved job found. We found an unsaved job, do you wish to restore it?
- © 2014 Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

Customer_Trusted_to_Curated

Actions ▾ Save ▾ Run

Visual Script Job details Runs Data quality - updated Schedules Version Control

Data source properties - S3

Name Customer_Trusted_Zone_Node

S3 source type Info

- S3 location Choose a file or folder in an S3 bucket.
- Data Catalog table

S3 URI s3://project-stedi-lake-house-rishi/customer/trusted/ View ▾ Browse S3

Recursive Read files in all subdirectories.

Data format JSON

JsonPath - optional Identify records with a JsonPath expression.

Multiline Indicates if JSON records can span multiple lines.

Infer schema

Additional options

Data preview Output schema

Data preview (200) Info **READY** End session Previewing 10 of 10 fields

Filter sample dataset

serialnumber	birthday	registrationdate	sharewithresearcherashowdate	customername	sharewithfriendsasofdate	email	lastupdatedate	phone	size
507f4f5f-7d5-4b07-a421-7b902c8d20fc	1900-01-01	1655564376361	1655564376361	Santosh Clayton	1655564376361	Santosh.Clayton@test.com	1655564376361	8015551212	1
20402022-c1da-4328-8e94-5bac7d61ecb	1399-01-01	1655564414415	1655564414415	Ben Khatib	1655564414415	Ben.Khatib@test.com	1655564414415	8015551212	n
2b6fe805-9059-4e1b-b7ed-720f5b4bf10	1141-01-01	1655564433612	1655564433612	Jane Gonzalez	1655564433612	Jane.Gonzalez@test.com	1655564433612	8015551212	n

AWS Services Search [Option+S] N. Virginia vocabs/layer033498+9ea2df2d-322b-11ee-a55d-535da5610840 @ 92... ▾

Customer_Trusted_to_Curated ↗

Visual Script Job details Runs Data quality - updated Schedules Version Control

Data source properties - S3

Name Accelerometer_Trusted_Zone_Node

S3 source type S3 location Choose a file or folder in an S3 bucket.
 Data Catalog table

S3 URL View

Recursive Read files in all subdirectories.

Data format

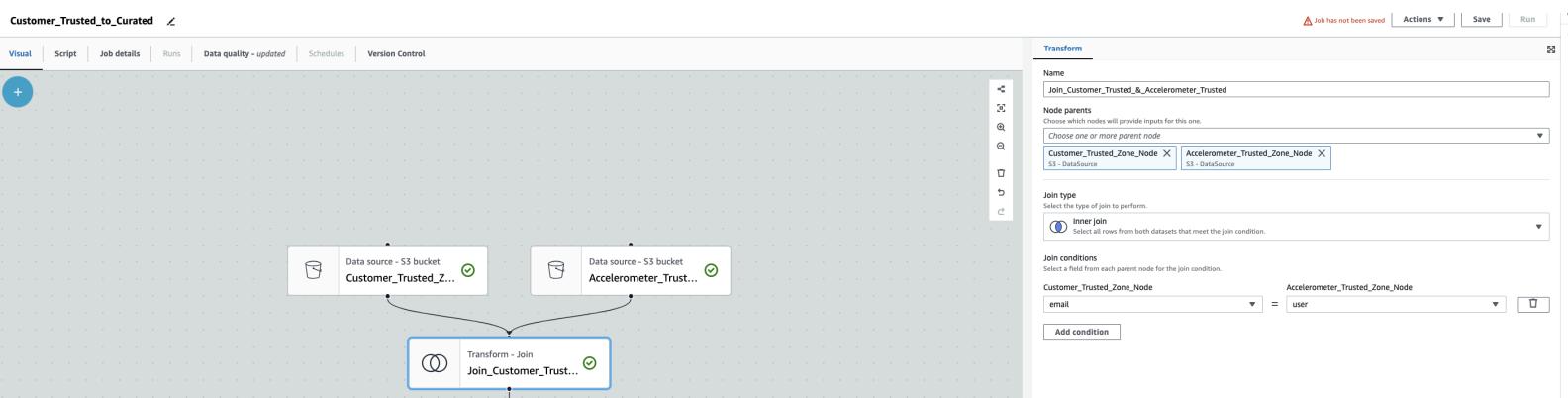
JSONPath - optional Identify records with a JSONPath expression.
 Multiline Indicates if JSON records can span multiple lines.
 Infer schema

▶ Additional options

Data preview Output schema

Data preview (200) **READY** Filter sample dataset

z	v	user	v	y	v	x	v	timestamp
-1		Lyn.Sanchez@test.com		-1		1		1655564469668
0		Lyn.Sanchez@test.com		1		0		1655564466935
0		Lyn.Sanchez@test.com		1		1		1655564458736
-1		Lyn.Sanchez@test.com		1		-1		1655564458736
-1		Lyn.Sanchez@test.com		1		-1		1655564456003



Customer_Trusted_to_Curated

Visual | Script | Job details | Runs | Data quality - updated | Schedules | Version Control

Actions ▾ Save Run

```

graph TD
    A[Data source - S3 bucket Customer_Trusted_Z...] --> C[Join_Customer_Trust...]
    B[Data source - S3 bucket Accelerometer_Trust...] --> C
    C --> D[Drop_Fields_of_Accelerometer_Trust...]
  
```

Transform

Name: `Drop_Fields_of_Accelerometer_Trust`

Node parents: Choose which nodes will provide inputs for this one. `Join_Customer_Trusted_&_Accelerometer_Trusted`

DropFields

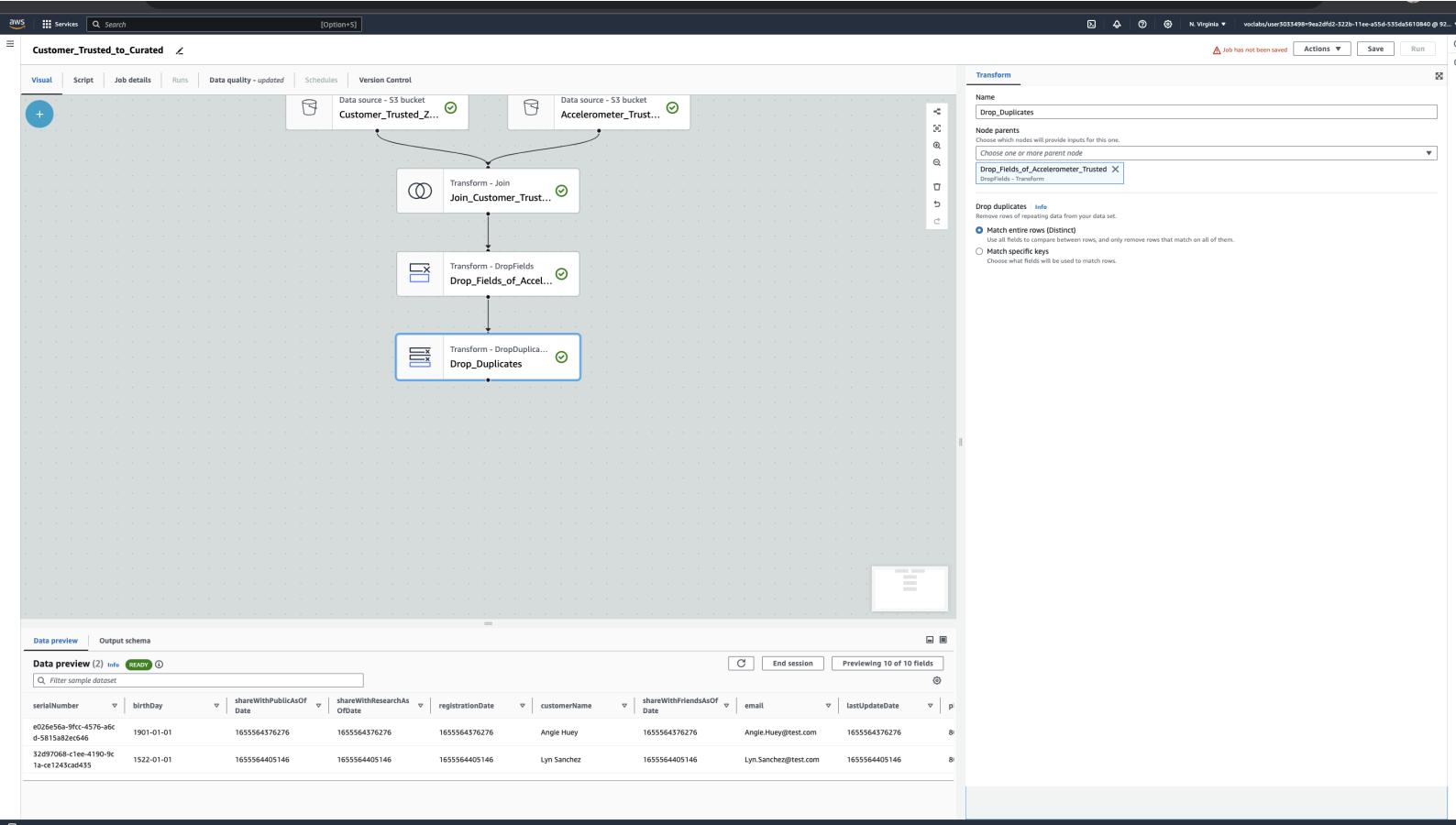
Field	Data type
serialNumber	string
<input checked="" type="checkbox"/> z	double
<input type="checkbox"/> birthDay	string
<input type="checkbox"/> shareWithPublicAsOfDate	long
<input type="checkbox"/> shareWithResearchAsOfDate	long
<input type="checkbox"/> registrationDate	long
<input type="checkbox"/> customerName	string
<input checked="" type="checkbox"/> user	string
<input type="checkbox"/> shareWithFriendsAsOfDate	long
<input checked="" type="checkbox"/> y	double
<input checked="" type="checkbox"/> x	double
<input checked="" type="checkbox"/> timestamp	long
<input type="checkbox"/> email	string
<input type="checkbox"/> lastUpdateDate	long
<input type="checkbox"/> phone	string

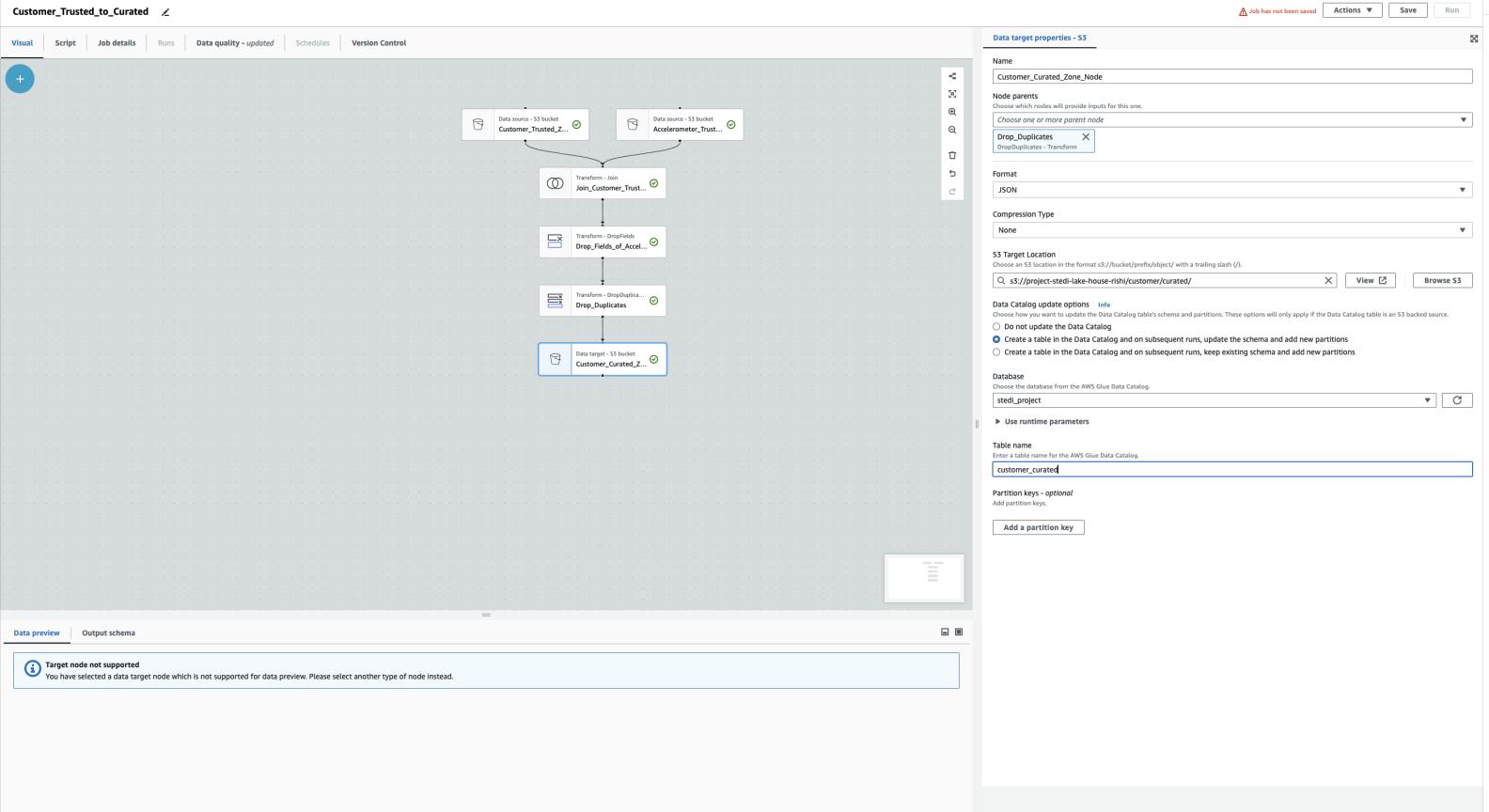
Data preview

Data preview (170) Info READY ⚡

Filter sample dataset

serialNumber	birthDay	shareWithPublicAsOfDate	shareWithResearchAsOfDate	registrationDate	customerName	shareWithFriendsAsOfDate	email	lastUpdateDate
52d97068-1ee-4190-9c1a-c1243cad435	1522-01-01	1655564405146	1655564405146	1655564405146	Lyn Sanchez	1655564405146	Lyn.Sanchez@test.com	1655564405146
52d97068-1ee-4190-9c1a-c1243cad435	1522-01-01	1655564405146	1655564405146	1655564405146	Lyn Sanchez	1655564405146	Lyn.Sanchez@test.com	1655564405146
52d97068-1ee-4190-9c1a-c1243cad435	1522-01-01	1655564405146	1655564405146	1655564405146	Lyn Sanchez	1655564405146	Lyn.Sanchez@test.com	1655564405146





Step_Trainer_Landing_to_Trusted

Last modified on 3/19/2024, 8:41:40 AM Actions Save Run

Visual Services Runs Data quality - updated Schedules Version Control

Data source properties - S3

Name Step_Trainer_Landing_Zone_Node

S3 source type Info
 S3 location Choose a file or folder in an S3 bucket.
 Data Catalog table

S3 URL s3://project-stedi-take-house-rishi/step_trainer/landing/ View Browse S3

Recursive Read files in all subdirectories.

Data format JSON
 JSONPath - optional Identify records with a JSONPath expression.

Multiline Indicates if JSON records can span multiple lines.
 Infer schema

Additional options

```

graph TD
    A[Data source - S3 bucket Step_Trainer_Landing...] --> C[Transform - SQL Query SQL_Query_Join_...]
    B[Data source - S3 bucket Customer_Curated_Z... ] --> C
    C --> D[Data target - S3 bucket Step_Trainer_Trusted...]
  
```

Data preview (200) Info Ready

Filter sample dataset

sensorreadingtime	serialnumber	distancefromobject
1655564444103	50f7bf13-7af5-4b07-a42	218
1-7b902c8d2b7c		
16555644440170	50f7bf13-7af5-4b07-a42	230
1-7b902c8d2b7c		
1655564416572	50f7bf13-7af5-4b07-a42	268
1-7b902c8d2b7c		

Step_Trainer_Landing_to_Trusted

Last modified on 3/19/2024, 8:41:40 AM Actions ▾ Save Run

Visual Script Job details Runs Data quality - updated Schedules Version Control

```

graph TD
    A[Data source - S3 bucket Step_Trainer_Landing...] --> C[Transform - SQL Query SQL_Query_Join_&...]
    B[Data source - S3 bucket Customer_Curated_Z... ] --> C
    C --> D[Data target - S3 bucket Step_Trainer_Trusted...]
  
```

Data source properties - S3

Name: Customer_Curated_Zone_Node

S3 source type: Info
 S3 location
 Subfolders in a folder in an S3 bucket.
 Data Catalog table

S3 URL: View Browse S3

Recursive: Read files in all subdirectories.

Data format: JSON
 JsonPath - optional
 Identify records with a jsonPath expression.

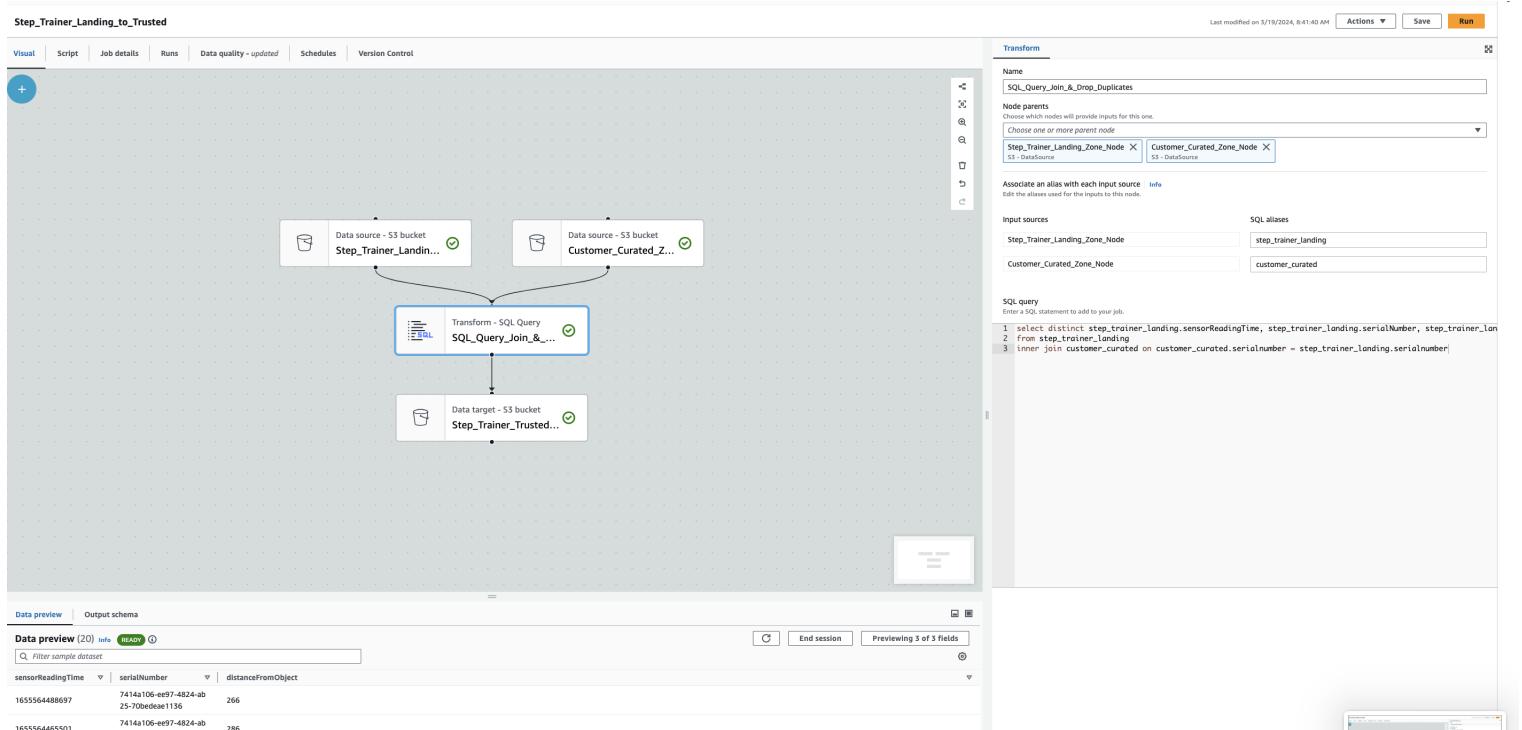
Multiline: Indicates if JSON records can span multiple lines.
 Infer schema

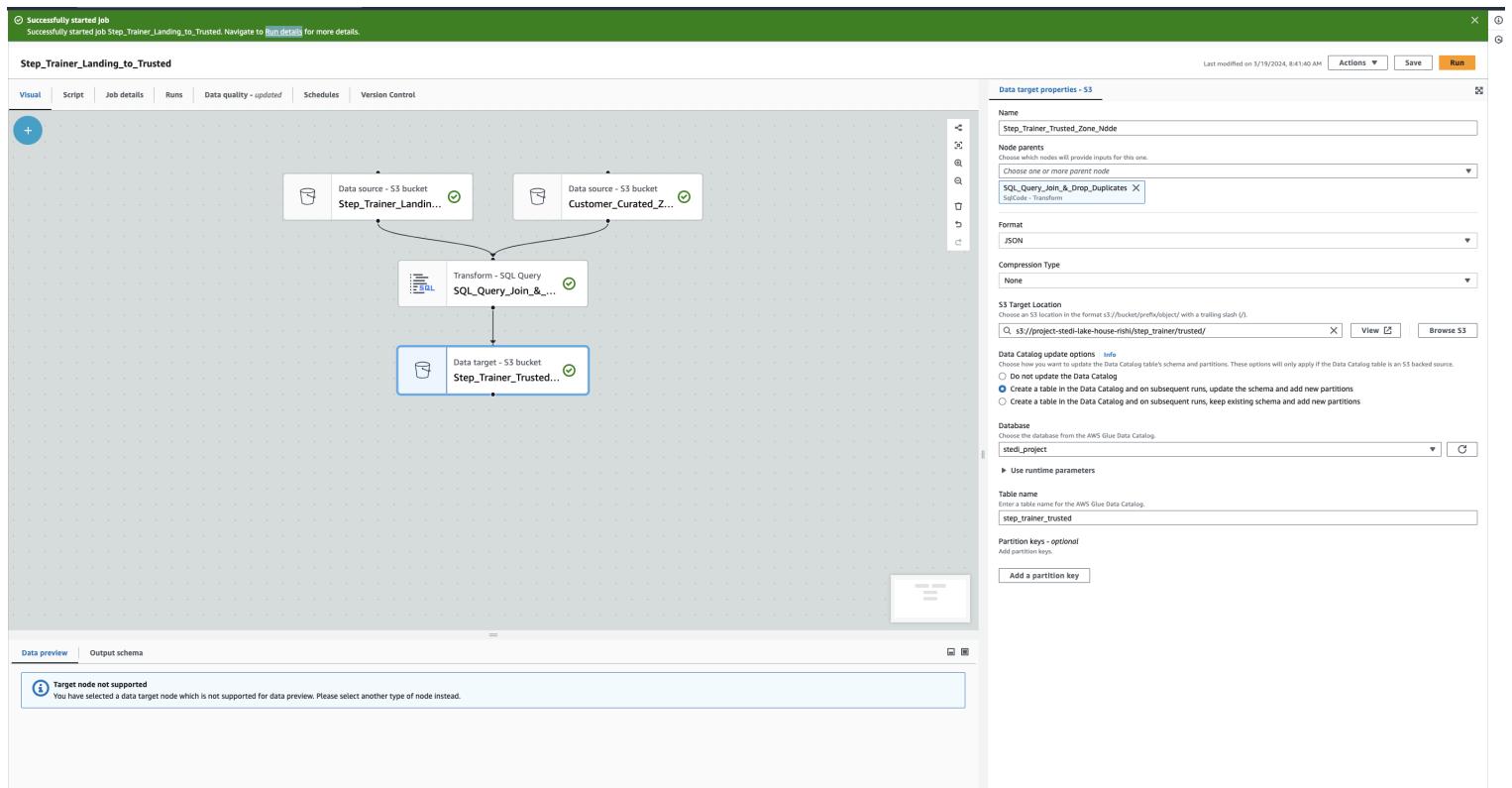
Additional options

Data preview | Output schema

Data preview (138) Info **READY** End session Previewing 10 of 10 fields

serialnumber	birthday	sharewithpublicate	sharewithresearcherdate	registrationdate	customername	email	lastupdatedate	phone
10a463c9-e723-454e-b5f9-25d0627487a6	1145-01-01	1655564433324	1655564433324	1655564433324	Dan Spencer	Dan.Spencer@test.com	1655564433324	8015551212
b97cfdd5-368d-42e2-b66f-72b77bf34e72	1195-01-01	null	1655564429652	1655564429652	Jaya Doshi	Jaya.Doshi@test.com	1655564429652	8015551212





aws Services Search [Option+S]

Machine_Learning_Curated

Visual Script Job details Runs Data quality - updated Schedules Version Control

Basic properties Info

Name
Machine_Learning_Curated

Description - optional
An aggregated table that has each of the Step Trainer Readings (step_trainer_trusted), and the associated accelerometer reading data for the same timestamp, but only for customers who have agreed to share their data (accelerometer_trusted).

Descriptions can be up to 2048 characters long.

IAM Role
Role assumed by the job with permission to access your data stores. Ensure that this role has permission to your Amazon S3 sources, targets, temporary directory, scripts, and any libraries used by the job.

my-glue-service-role

Type
The type of ETL job. This is set automatically based on the types of data sources you have selected.

Spark

Glue version Info
Glue 4.0 - Supports spark 3.3, Scala 2, Python 3

Language
Python 3

Worker type
Set the type of predefined worker that is allowed when a job runs.

G 1X
(4vCPU and 16GB RAM)

Automatically scale the number of workers
 AWS Glue will optimize costs and resource usage by dynamically scaling the number of workers up and down throughout the job run. Requires Glue 3.0 or later.

Requested number of workers
The number of workers you want AWS Glue to allocate to this job.
10

Generate job insights
 AWS Glue will analyze your job runs and provide insights on how to optimize your jobs and the reasons for job failures.

Job bookmark Info
Specifies how AWS Glue processes job bookmark when the job runs. It can remember previously processed data (Enable), update state information (Pause), or ignore state information (Disable).

Disable

Flex execution Info
 Reduce costs by running this job on spare capacity. Ideal for non-urgent workloads that don't require fast job start times or consistent execution times. See recommendations, limitations and pricing in the help panel by clicking on the Info link above.

Number of retries
0

Screenshot of the AWS Glue Data Catalog interface showing the creation of a new data source.

Machine_Learning_Curated Job

Data source properties - S3

Name: Step_Trainer_Trusted_Zone_Node

S3 source type: S3 location
Choose a file or folder in an S3 bucket.
 Data Catalog Table

S3 URL: `s3://project-steed-lake-house-rishi/step_trainer/trusted/`

Recursive: Read files in all subdirectories.

Data format: JSON

JsonPath - optional: Identify records with a JsonPath expression.

Multiline
Indicates if JSON records can span multiple lines.

Infer schema

Additional options

Data preview (200)

serialnumber	distancefromobject
1655564142276	0382c891-e0a2-4703-b7 a6-2ee8917fe4ba
1655564138310	0382c891-e0a2-4703-b7 a6-2ee8917fe4ba
1655564134344	0382c891-e0a2-4703-b7 a6-2ee8917fe4ba
.....

Screenshot of the AWS Glue Data Catalog interface showing the creation of a new data source named "Accelerometer_Trusted_Zone_Node".

The "Data source properties - S3" section includes:

- Name:** Accelerometer_Trusted_Zone_Node
- S3 source type:** S3 location (selected)
- S3 URL:** s3://project-steed-lake-house-rish/accelerometer/trusted/
- Data format:** JSON
- JSONPath - optional:** (empty)
- Multiline:** (unchecked)
- Infer schema:** (checked)

The "Data preview" section shows a sample of 200 rows from the dataset:

z	user	y	x	timestamp
-1	Lyn.Sanchez@test.com	-1	1	1655564469668
0	Lyn.Sanchez@test.com	1	0	1655564466935
0	Lyn.Sanchez@test.com	1	1	1655564458736
-1	Lyn.Sanchez@test.com	1	-1	1655564458736
-1	Lyn.Sanchez@test.com	1	-1	1655564456003
-1	Lyn.Sanchez@test.com	1	-1	1655564445071
0	Lyn.Sanchez@test.com	1	0	1655564423207

At the bottom, there are links for CloudShell, Feedback, and copyright information: © 2024, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences.

Screenshot of the AWS Glue Data Catalog interface showing a job named "Machine_Learning_Curated".

Job Overview:

- Name: Machine_Learning_Curated
- Status: Not yet saved
- Last modified: N. Virginia (3 hours ago)
- Actions: Save, Run

Job Details:

- VISUAL: Selected
- Script: Not selected
- Job details: Not selected
- RUNS: Not selected
- Data quality - updated: Not selected
- Schedules: Not selected
- Version Control: Not selected

Transform:

SQL_Query_Join_B_Drop_Duplicates

Node parents:
Choose which nodes will provide inputs for this one.
Choose one or more parent node
Step_Trainer_Trusted_Zone_Node X
Accelerometer_Trusted_Zone_Node X
S3 Dataviewer

Associate an alias with each input source
Edit the aliases used for the inputs to this node.

Input sources:

Step_Trainer_Trusted_Zone_Node	SQL aliases: step_trainer_trusted
Accelerometer_Trusted_Zone_Node	accelerometer_trusted

SQL query:
Enter a SQL statement to add to your job.

```

1 select distinct * from step_trainer_trusted inner join accelerometer_trusted
2 on accelerometer_trusted.timestamp = step_trainer_trusted.sensorReadingTime
    
```

Data preview:

Output schema

Data preview (0) info READY

Filter sample dataset

sensorReadingTime | serialNumber | distanceFromObject | z | user | y | x | timestamp

No data to display
Try increasing the number of rows being sampled.

