

ata source

AwsDataCatalog

atalog

None

atabase

stedi\_db

ables and views

Create

Filter tables and views

Tables (8)

< 1 >

accelerometer\_landing

:

accelerometer\_trusted

:

customer\_curated

:

customer\_landing

:

customer\_trusted

:

machine\_learning\_curated

:

1

select \* from machine\_learning\_curated;

SQL

Ln 1, Col 39

Run again

Explain

Cancel

Clear

Create

Reuse query results

up to 60 minutes ago

Query results

Query stats

Completed

Time in queue: 114 ms

Run time: 927 ms

Data scanned: 7.15 MB

Results (43,681)

Copy

Download results CSV

## Machine\_Learning\_Curated

Last modified on 4/22/2025, 1:42:21 PM

Actions Save Run

Visual Script Job details Runs Data quality Schedules Version Control

### Script (Locked) Info

Download script Edit script

```
11 for alias, frame in mapping.items():
12     frame.toDF().createOrReplaceTempView(alias)
13     result = spark.sql(query)
14     return DynamicFrame.fromDF(result, glueContext, transformation_ctx)
15 args = getResolvedOptions(sys.argv, ['JOB_NAME'])
16 sc = SparkContext()
17 glueContext = GlueContext(sc)
18 spark = glueContext.spark_session
19 job = Job(glueContext)
20 job.init(args['JOB_NAME'], args)
21
22 # Default ruleset used by all target nodes with data quality enabled
23 DEFAULT_DATA_QUALITY_RULESET = """
24 Rules = [
25     ColumnCount > 0
26 ]
27 """
28
```

Results (81,273)

Copy

Download results CSV

Search rows

< 1 ... > ⚙

#	user	timestamp	x	y	z
1	Jane.Olson@test.com	1655564439144	1.0	0.0	-1.0
2	Jane.Olson@test.com	1655564439144	1.0	-1.0	-1.0
3	Jane.Olson@test.com	1655564430945	-1.0	1.0	-1.0
4	Jane.Olson@test.com	1655564430945	-1.0	-1.0	-1.0
5	Jane.Olson@test.com	1655564417280	-1.0	-1.0	0.0
6	Jane.Olson@test.com	1655564414547	-1.0	0.0	0.0
7	Jane.Olson@test.com	1655564414547	0.0	-1.0	-1.0
8	Jane.Olson@test.com	1655564411814	0.0	-1.0	0.0
9	Jane.Olson@test.com	1655564411814	-1.0	-1.0	0.0
10	Danny.Hansen@test.com	1655564133234	1.0	-1.0	0.0

Query results

Query stats

✓ Completed

Time in queue: 107 ms

Run time: 724 ms

Data scanned: 286.89 KB

Results (956)

Copy

Download results CSV

Search rows

< 1 ... > ⚙

rationdate	lastupdatedate	sharewithresearchasofdate	sharewithpublicasofdate	sharewithfriendsasofdate
564376361	1655564376361	1655564376361	1655564376361	1655564376361
564414415	1655564414415	1655564414415		1655564414415
564131452	1655564131452			1655564131452
564435763	1655564435763			
564387419	1655564387419		1655564387419	1655564387419
564397161	1655564397161			1655564397161
564433612	1655564433612	1655564433612		1655564433612
564436700	1655564436700	1655564436700		

AWS Glue Studio [Info](#)

Create job [Info](#)



Author in a visual interface focused on data flow.

Visual ETL



Author using an interactive code notebook.

Notebook



Author code with a script editor.

Script editor

► Example jobs [Info](#)

Create example job

Your jobs (3) [Info](#)



Actions ▼

Run job

Q Filter jobs by property

< 1 > ⚙

<input type="checkbox"/>	Job name ▼	Type	Created by	Last modified ▼	AWS Glue version ▼
<input type="checkbox"/>	<a href="#">Customer_curated</a>	Glue ETL	Visual	4/22/2025, 11:27:24 AM	5.0
<input type="checkbox"/>	<a href="#">Accelerometer_trusted</a>	Glue ETL	Visual	4/22/2025, 11:15:35 AM	5.0
<input type="checkbox"/>	<a href="#">Customer_landing_to_trusted</a>	Glue ETL	Visual	4/22/2025, 10:54:57 AM	5.0

Results (482)

Copy

Download results CSV

Q Search rows

< 1 ... > ⚙

# ▼	customername ▼	email ▼	phone ▼	birthday ▼	serialnumber
1	Santosh Clayton	Santosh.Clayton@test.com	8015551212	1900-01-01	50f7b4f3-7af5-4b07-a421-7b902c8d2b
2	Ben Khatib	Ben.Khatib@test.com	8015551212	1399-01-01	20400202-c1da-4328-8e94-5bacc7d61e
3	Jane Gonzalez	Jane.Gonzalez@test.com	8015551212	1141-01-01	2b6fe805-9059-4e1b-b7ed-720f5b4bf1
4	Jaya Jackson	Jaya.Jackson@test.com	8015551212	1098-01-01	220437fd-77f5-463a-8d93-f59c944e60
5	Jerry Jefferson	Jerry.Jefferson@test.com	8015551212	1964-01-01	f1ad9449-f80b-4dcb-8fab-66352e9293
6	Senthil Smith	Senthil.Smith@test.com	8015551212	1587-01-01	dbfde3b7-69bf-4c56-a9e8-dfd48af6df8
7	Santosh Davis	Santosh.Davis@test.com	8015551212	1745-01-01	b5f6b144-0260-4c4f-83e1-469150f905
8	Sarah Doshi	Sarah.Doshi@test.com	8015551212	1126-01-01	466bb1b9-5dc3-4607-8f76-e0bcdcf6b9

Results (28,680)

Copy Download results CSV

#	sensorreadingtime	serialNumber	distancefromobject
1	1655564407947	81624e88-1469-4f54-977f-3dcb161ef79e	281
2	1655564365047	81624e88-1469-4f54-977f-3dcb161ef79e	272
3	1655564361147	81624e88-1469-4f54-977f-3dcb161ef79e	257
4	1655564353347	81624e88-1469-4f54-977f-3dcb161ef79e	260
5	1655564494769	058af4ef-68ed-4af7-b51d-b6223e5300ff	293
6	1655564488103	058af4ef-68ed-4af7-b51d-b6223e5300ff	278
7	1655564478104	058af4ef-68ed-4af7-b51d-b6223e5300ff	266
8	1655564471438	058af4ef-68ed-4af7-b51d-b6223e5300ff	269
9	1655564451440	058af4ef-68ed-4af7-b51d-b6223e5300ff	240
10	1655564441111	058af4ef-68ed-4af7-b51d-b6223e5300ff	222

Step\_trainer\_curated

Last modified on 4/22/2025, 1:27:13 PM Actions Save Run

Visual Script Job details Runs Data quality Schedules Version Control

Script (Locked) Info

Download script Edit script


```
1 import sys
2 from awsglue.transforms import *
3 from awsglue.utils import getResolvedOptions
4 from pyspark.context import SparkContext
5 from awsglue.context import GlueContext
6 from awsglue.job import Job
7 from awsgluedq.transforms import EvaluateDataQuality
8 from awsglue import DynamicFrame
9
10 def sparkSqlQuery(glueContext, query, mapping, transformation_ctx) -> DynamicFrame:
11     for alias, frame in mapping.items():
12         frame.toDF().createOrReplaceTempView(alias)
13     result = spark.sql(query)
14     return DynamicFrame.fromDF(result, glueContext, transformation_ctx)
15 args = getResolvedOptions(sys.argv, ['JOB_NAME'])
16 sc = SparkContext()
17 glueContext = GlueContext(sc)
18 spark = glueContext.spark_session
```

1

select \* from customer\_curated;

SQLLn 1, Col 31

Run again

Explain 


Cancel

Clear

Create ▼

Query results

Query stats

 Completed

Time in queue: 111 ms

Run

Results (482)

```
1 select * from accelerometer_trusted;
```

SQL Ln 1, Col 36

Run again

Explain [↗](#)

Cancel

Clear

Create

Query results

Query stats

✓ Completed

Time in

**Results** (40,981)

```
1 select * from step_trainer_landing;
```

SQL Ln 1, Col 35

Run again

Explain [↗](#)

Cancel

Clear

Query results

Query stats

✔ Completed

**Results** (28,680)