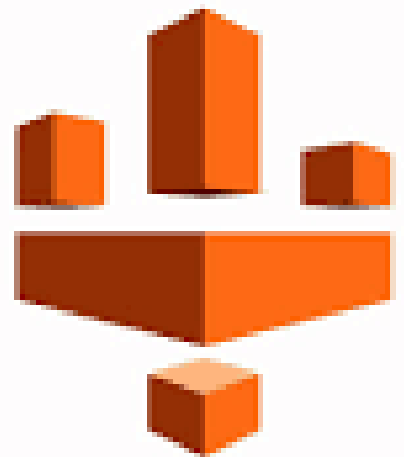


# 1 - Spark script editor

AWS Glue ETL Service



aws

Serviços

Search

[Alt+S]

Global

AdministratorAccess/marios.barros.pb@compasso.com.br

Amazon S3

Buckets

lb-01

lb-01

Informações

ObjetosPropriedadesPermissõesMétricasGerenciamentoPontos de acesso

Objetos (1)

Os objetos são as entidades fundamentais armazenadas no Amazon S3. Você pode usar o [Inventário do Amazon S3](#) para obter uma lista de todos os objetos em seu bucket. Para outras pessoas acessarem seus objetos, você precisará conceder permissões explicitamente a eles. [Saiba mais](#)

Copiar URI do S3

Copiar URL

Fazer download

Abrir

Excluir

Ações

Criar pasta

Carregar

Localizar objetos por prefixo

< 1 >

<input type="checkbox"/>	Nome	Tipo	Última modificação	Tamanho	Classe de armazenamento
<input type="checkbox"/>	lab-glue/	Pasta	-	-	-

aws

Serviços

Search

[Alt+S]

Global

AdministratorAccess/marios.barros.pb@compasso.com.br

Amazon S3

Amazon S3 > Buckets > lb-01 > lab-glue/ > input/

input/

ObjetosPropriedades

Objetos (1)

Os objetos são as entidades fundamentais armazenadas no Amazon S3. Você pode usar o [Inventário do Amazon S3](#) para obter uma lista de todos os objetos em seu bucket. Para outras pessoas acessarem seus objetos, você precisará conceder permissões explicitamente a eles. [Saiba mais](#)

Copiar URI do S3

Copiar URL

Fazer download

Abrir

Excluir

Ações

Criar pasta

Carregar

Localizar objetos por prefixo

< 1 >

<input type="checkbox"/>	Nome	Tipo	Última modificação	Tamanho	Classe de armazenamento
<input type="checkbox"/>	nomes.csv	csv	26 Sep 2023 06:45:36 PM -03	29.4 MB	Padrão

aws

Serviços

Search

[Alt+S]

Global

AdministratorAccess/marios.barros.pb@compasso.com.br

Identity and Access Management (IAM)

IAM > Usuários

Usuários (1) Informações

Um usuário do IAM é uma identidade com credenciais de longo prazo que é usada para interagir com a AWS em uma conta.

Pesquisar

< 1 >

<input type="checkbox"/>	Nome do usuário	Caminho	Grupo	Última atividade	MFA	Idade da senha	Último login no consó	ID da chave de acesso	Idade da chave ati	Última chave de a	ARN
<input type="checkbox"/>	marios-barros	/	0		+	5 horas	+	+	+		arn:aws:iam::369659558219:user/m

Consoles relacionados

Central de identidades do IAM

AWS Organizations

Services

Search

[Alt+S]

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

ETL jobs

Visual ETL

Notebooks

Job run monitoring

Interactive Sessions

Data classification tools

Sensitive data detection

Record Matching

AWS Glue

Getting started

Create IAM role

Step 1

Choose IAM users and roles for AWS Glue

Step 2

Grant Amazon S3 access

Step 3

Choose a default service role

Step 4

Review and confirm

Choose IAM users and roles for AWS Glue

How it works

Choose IAM identities (roles or users) in your account that need access to AWS Glue. AWS will attach the AWSGlueConsoleFullAccess IAM managed policy to these users and roles. This will grant these identities resources. [Learn more](#)

Selected roles (0)

Choose the roles you want to have access to AWS Glue.

Role name

Creation date (UTC)

No roles selected

Choose roles

Selected users (1)

Choose the users you want to have access to AWS Glue.

☒

User name

Console last activity

Creation date (UTC)

☒ marlos-barros

-

September 26, 2023 at 21:50

Services

Search

[Alt+S]

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

ETL jobs

Visual ETL

Notebooks

Job run monitoring

Interactive Sessions

Data classification tools

Sensitive data detection

Record Matching

Triggers

Workflows (orchestration)

Blueprints

Security configurations

Legacy pages

Connections (legacy)

Jobs (legacy)

Workflows (legacy)

AWS Glue

Getting started

Create IAM role

Step 1

Choose IAM users and roles for AWS Glue

Step 2

Grant Amazon S3 access

Step 3

Choose a default service role

Step 4

Review and confirm

Grant Amazon S3 access

How it works

Grant S3 access to the users and roles that you selected in Step 1. Glue will attach permissions to those identities based on the type of access you choose here. [Learn more](#)

Choose S3 locations

Targeted users and roles

1 users and 0 roles

Choose access to Amazon S3

☐ No additional access

Do not change permissions.

☐ Add access to specific Amazon S3 locations

Choose specific S3 paths that you want to grant access to.

☒ Grant full access to Amazon S3

Grant access to all S3 resources in your AWS account.

Data access permissions

Set the type of data access for the Glue users and roles.

Data access permissions

☒ Read only (recommended)

☐ Read and write

Cancel

Previous

Next

Services

Search

[Alt+S]

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

ETL jobs

Visual ETL

Notebooks

Job run monitoring

Interactive Sessions

Data classification tools

Sensitive data detection

Record Matching

Triggers

Workflows (orchestration)

Blueprints

Security configurations

Legacy pages

Connections (legacy)

Jobs (legacy)

Workflows (legacy)

AWS Glue

Getting started

Create IAM role

Step 1

Choose IAM users and roles for AWS Glue

Step 2

Grant Amazon S3 access

Step 3

Choose a default service role

Step 4

Review and confirm

Choose a default service role

How it works

AWS Glue uses an IAM service role to run jobs, access data, and run Data Quality tasks. We recommend that you start with the standard AWSGlueServiceRole as the default. [Learn more](#)

Choose a default AWS Glue service role

IAM role for AWS Glue

☒ Update the standard AWS Glue service role and set it as the default (recommended)

AWS will update the role with the IAM policies needed to run AWS Glue jobs, then set it as the default.

☐ Set an existing IAM role as the default

Select an IAM role that you've configured to use as an AWS Glue service role. Glue will set this role as the default, but won't add any permissions to it. [Learn more](#)

The following IAM role will be created and automatically configured for you:

- AWSGlueServiceRole

Cancel

Previous

Next

Serviços

Search

[Alt+S]

GlobalAdministratorAccess/marios.barros.pb@compasso.com.br

IAM > Funções > Criar função

Etapa 1

Selecionar entidade confiável

Etapa 2

Adicionar permissões

Etapa 3

Nomear, revisar e criar

Selecionar entidade confiável

Informações

Tipo de entidade confiável

☒ Serviço da AWS

Permitir que serviços da AWS, como o EC2, Lambda ou outros executem ações nessa conta.

☐ Conta da AWS

Permitir que entidades em outras contas da AWS pertencentes a você ou a terceiros executem ações nessa conta.

☐ Identidade Web

Permitir que os usuários federados pelo provedor de identidade da Web externo especificado assumam essa função para executar ações nessa conta.

☐ Federação SAML 2.0

Permitir que os usuários federados com o SAML 2.0 de um diretório corporativo executem ações nessa conta.

☐ Política de confiança personalizada

Crie uma política de confiança personalizada para permitir que outras pessoas executem ações nessa conta.

Caso de uso

Permitir que um serviço da AWS, como o EC2, o Lambda ou outros executem ações nessa conta.

Serviço ou caso de uso

Glue

Escolha um caso de uso para o serviço especificado.

Caso de uso

☒ Glue

Allows Glue to call AWS services on your behalf.

Cancelar

Próximo

Serviços

Search

[Alt+S]

GlobalAdministratorAccess/marios.barros.pb@compasso.com.br

IAM > Funções > Criar função

Etapa 1

Selecionar entidade confiável

Etapa 2

Adicionar permissões

Etapa 3

Nomear, revisar e criar

Nomear, revisar e criar

Detalhes da função

Nome da função

Insira um nome significativo para identificar esta função.

Máximo de 64 caracteres. Use caracteres alfanuméricos e "-,.,@,\_,:."

Descrição

Adicione uma breve explicação para esta função.

Allows Glue to call AWS services on your behalf.

Máximo de 1000 caracteres. Use caracteres alfanuméricos e "-,.,@,\_,:."

Etapa 1: Selecionar entidades confiáveis

Política de confiança

```
1 {
2   "Version": "2012-10-17",
3   "Statement": [
4     {
5       "Effect": "Allow",
6       "Principal": {
7         "Service": "glue.amazonaws.com"
8       },
9       "Action": "sts:AssumeRole"
10    }
11  ]
12 }
```

Editar

Etapa 2: Adicionar permissões

Resumo da política de permissões

Nome da política	Tipo	Anexado como
<a href="#">AmazonS3FullAccess</a>	Gerenciadas pela AWS	Política de permissões
<a href="#">AWSGlueConsoleFullAccess</a>	Gerenciadas pela AWS	Política de permissões
<a href="#">AWSLakeFormationDataAdmin</a>	Gerenciadas pela AWS	Política de permissões
<a href="#">CloudWatchFullAccess</a>	Gerenciadas pela AWS	Política de permissões

Etapa 3: Adicionar tags

Adicionar tags - opcional

Tags são pares de chave/valor que você pode adicionar aos recursos da AWS para ajudar a identificar, organizar ou pesquisar recursos.

Nenhuma tag associada ao recurso.

Adicione uma nova tag

Você pode adicionar até mais 50 tags.

Cancelar

Anterior

Criar função

Serviços

Search

[Alt+S]

Norte da VirgíniaAdministratorAccess/marios.barros.pb@compasso.com.br

AWS Lake Formation

Your database: glue- lab has been created successfully.

Introducing new Hybrid access mode

In hybrid access mode, your Lake Formation permissions and IAM policies for AWS Glue/S3 work together to provide access. [Learn more](#)

AWS Lake Formation > Databases

Databases (2)

Find databases

Name	Owner account ID	Lake Formation perm...	Default permissions f...	Shared resource	Shared resource owner	Shared resource owne...	Amazon S3 path	Description
<input type="radio"/> default	369659558219	No users	IAM/AWS Glue policies	-	-	-	-	-
<input type="radio"/> glue- lab	369659558219	No users	IAM/AWS Glue policies	-	-	-	-	-



Services

Search

[Alt+5]

Norte da Virginia

AdministratorAccess/marios.barras.pb@compasso.com.br

job\_aws\_glue\_lab\_4

Last modified on 26/09/2023, 22:50:47

ActionsSaveRun

ScriptJob detailsRunsData qualityNewSchedulesVersion Control

Basic properties

Name

job\_aws\_glue\_lab\_4

Description - optional

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.

AWSGlueServiceRole-Lab4

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 3.0 - Supports spark 3.1, 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.

2

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 jobs 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

Server-side encryption

☐ Select this option to enable S3-SSE encryption on both your data target and any data that is written to an S3 temporary directory. Ignored if a security configuration is specified [Learn More](#)

Use Glue data catalog as the Hive metastore

☒ To use Glue data catalog as the Hive metastore, the IAM role used for the job should have glue:CreateDatabase permissions. A database called "default" is created in the Glue data catalog if it does not exist.

Connections

Additional network connections

Info

Choose a VPC configuration to access Amazon S3 data sources located in your virtual private cloud (VPC). You can create and manage Network connections in AWS Glue.

Choose options

Current connections

These are the connections currently associated with the job.

Name	Type	VPC	Subnet	Security Groups
No connections				
No connections attached.				

Libraries

Info

Python library path

Dependent JARs path

Referenced files path

Job parameters

Info

Key	Value - optional
--S3_INPUT_PATH	s3://lb-01/lab-glue/input/
--S3_TARGET_PATH	s3://lb-01/lab-glue/output/

Add new parameter

You can add 48 more parameters.

Tags

No tags associated with the resource.

Add new tag

You can add 50 more tags.

aws

Services

Search

[Alt+5]

Norte da Virginia

AdministratorAccess/marios.barras.pb@compasso.com.br

job\_aws\_glue\_lab\_4

Last modified on 26/09/2023, 22:50:47

ActionsSaveRun

ScriptJob detailsRunsData qualityNewSchedulesVersion Control

aws

Services

Search

[Alt+5]

Norte da Virginia

AdministratorAccess/marios.barras.pb@compasso.com.br

job\_aws\_glue\_lab\_4

Last modified on 26/09/2023, 22:50:47

ActionsSaveRun

ScriptJob detailsRunsData qualityNewSchedulesVersion Control

Server-side encryption

☐ Select this option to enable S3-SSE encryption on both your data target and any data that is written to an S3 temporary directory. Ignored if a security configuration is specified [Learn More](#)

Use Glue data catalog as the Hive metastore

☒ To use Glue data catalog as the Hive metastore, the IAM role used for the job should have glue:CreateDatabase permissions. A database called "default" is created in the Glue data catalog if it does not exist.

Connections

Additional network connections

Info

Choose a VPC configuration to access Amazon S3 data sources located in your virtual private cloud (VPC). You can create and manage Network connections in AWS Glue.

Choose options

Current connections

These are the connections currently associated with the job.

Name	Type	VPC	Subnet	Security Groups
No connections				
No connections attached.				

Libraries

Info

Python library path

Dependent JARs path

Referenced files path

Job parameters

Info

Key	Value - optional
--S3_INPUT_PATH	s3://lb-01/lab-glue/input/
--S3_TARGET_PATH	s3://lb-01/lab-glue/output/

Add new parameter

You can add 48 more parameters.

Tags

No tags associated with the resource.

Add new tag

You can add 50 more tags.

EC2

S3

Athena

AWS Glue

AWS Lake Formation

Lambda

Search

[Alt+S]

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

ETL Jobs

Visual ETL

Notebooks

Job run monitoring

Interactive Sessions

Data classification tools

Sensitive data detection

Record Matching

Triggers

Workflows (orchestration)

Blueprints

Security configurations

Legacy pages

Connections (legacy)

Jobs (legacy)

Workflows (legacy)

Blueprints (legacy)

Schema registries (legacy)

Schemas (legacy)

What's New

Documentation

AWS Marketplace

job\_aws\_glue\_lab\_4

Successfully started job

Successfully started job job\_aws\_glue\_lab\_4. Navigate to [Run details](#) for more details.

ScriptJob detailsRunsData qualityNewSchedulesVersion Control

Script Info

```
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 awsglue import DynamicFrame
8 from pyspark.sql.functions import upper, count, desc
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
16 args = getResolvedOptions(sys.argv, ["JOB_NAME"])
17 sc = SparkContext()
18 glueContext = GlueContext(sc)
19 spark = glueContext.spark_session
20 job = Job(glueContext)
21 job.init(args["JOB_NAME"], args)
22
23 nomes_df = spark.read.csv("s3://laboratio-glue/lab-glue/input/nomes.csv", header=True)
24
25 nomes_df.printSchema()
26
27 nomes_df = nomes_df.withColumn("nome", upper(nomes_df["nome"]))
28
29 nomes_df.groupBy("ano", "sexo").agg(count("nome")).show()
30
31 nomes_df = nomes_df.orderBy(desc("ano"))
32
33 nome_feminino_mais_registrado = nomes_df.filter(nomes_df["sexo"] == "F").groupBy("nome", "ano").agg(count("nome").alias("contagem")).orderBy(desc("contagem")).first()
34
35 nome_masculino_mais_registrado = nomes_df.filter(nomes_df["sexo"] == "M").groupBy("nome", "ano").agg(count("nome").alias("contagem")).orderBy(desc("contagem")).first()
36
37 total_registros_por_ano = nomes_df.groupBy("ano").agg(count("").alias("total_registros")).orderBy("ano")
38 total_registros_por_ano.show(10)
39
40 nomes_df.write.partitionBy("sexo", "ano").json("s3://laboratio-glue/lab-glue/frequencia_registro_nomes_eua/")
41
```

PythonLn 3, Col 45Errors: 0Warnings: 0

# frequencia\_registro\_nomes\_eua/

Objetos

Propriedades

Objetos (2)

Os objetos são as entidades fundamentais armazenadas no Amazon S3. Você pode usar o [inventário do Amazon S3](#) para obter uma lista de todos os objetos em seu bucket. Para outras pessoas acessarem s

Copiar URI do S3

Copiar URL

Fazer download

Abrir

Excluir

Ações

Criar pasta

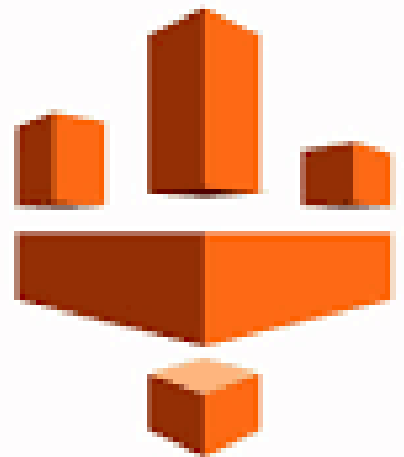
Carregar

Localizar objetos por prefixo

	Nome	Tipo	Última modificação	Tamanho	Class
	sexo=F/	Pasta	-	-	-
	sexo=M/	Pasta	-	-	-

# 2 - Visual with a source and target

AWS Glue ETL Service





Services

Search

[Alt+S]

Norte da Virginia

AdministratorAccess/marlos.barros.pb@compasso.com.br

AWS Glue

marlos-igor

Last modified on 27/09/2023, 00:29:14

Try new UI

Actions

Save

Run

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

ETL Jobs

Visual ETL

Notebooks

Job run monitoring

Interactive Sessions

Data classification tools

Sensitive data detection

Record Matching

Triggers

Workflows (orchestration)

Blueprints

Security configurations

Legacy pages

Connections (legacy)

Jobs (legacy)

Workflows (legacy)

Blueprints (legacy)

Schema registries (legacy)

Schemas (legacy)

What's New

Documentation

AWS Marketplace

Enable compact mode

Enable new navigation

Visual

Script

Job details

Runs

Data quality New

Schedules

Version Control

+

✕

🔍

🔍

🗑️

🔄

📄

Data source - Data Catalog

AWS Glue Data Catalog

Transform - SQL Query

SQL Query

Data target - S3 bucket

Amazon S3

Data source properties - Data Catalog

Output schema

Data preview

Name

AWS Glue Data Catalog

Database

Choose a database.

glue-lab

Use runtime parameters

Table

tb\_01

Use runtime parameters

Services

Search

[Alt+S]

Norte da Virginia

AdministratorAccess/marlos.barros.pb@compasso.com.br

AWS Glue

marlos-igor

Last modified on 27/09/2023, 00:29:14

Try new UI

Actions

Save

Run

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

ETL Jobs

Visual ETL

Notebooks

Job run monitoring

Interactive Sessions

Data classification tools

Sensitive data detection

Record Matching

Triggers

Workflows (orchestration)

Blueprints

Security configurations

Legacy pages

Connections (legacy)

Jobs (legacy)

Workflows (legacy)

Blueprints (legacy)

Schema registries (legacy)

Schemas (legacy)

What's New

Documentation

AWS Marketplace

Enable compact mode

Enable new navigation

Visual

Script

Job details

Runs

Data quality New

Schedules

Version Control

+

✕

🔍

🔍

🗑️

🔄

📄

Data source - Data Catalog

AWS Glue Data Catalog

Transform - SQL Query

SQL Query

Data target - S3 bucket

Amazon S3

Transform

Output schema

Data preview

Name

SQL Query

Node parents

Choose which nodes will provide inputs for this one.

Choose one or more parent node

AWS Glue Data Catalog

Catalog - DataSource

Associate an alias with each input source

Info

Edit the aliases used for the inputs to this node.

Input sources

AWS Glue Data Catalog

SQL aliases

myDataSource

SQL query

Enter a SQL statement to add to your job.

1 SELECT UPPER(none) AS none\_naluscilo, sexo, ano

2 FROM myDataSource

Services

Search

[Alt+S]

marlos-igor

Last modified on 27/09/2023, 00:29:14

Try new UI

Actions

Save

Run

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

ETL Jobs

Visual ETL

Notebooks

Job run monitoring

Interactive Sessions

Data classification tools

Sensitive data detection

Record Matching

Triggers

Workflows (orchestration)

Blueprints

Security configurations

Legacy pages

Connections (legacy)

Jobs (legacy)

Workflows (legacy)

Blueprints (legacy)

Schema registries (legacy)

Schemas (legacy)

What's New

Documentation

AWS Marketplace

Enable compact mode

Enable new navigation

Visual

Script

Job details

Runs

Data quality New

Schedules

Version Control

Data source - Data Catalog

AWS Glue Data Catalog

Transform - SQL Query

SQL Query

Data target - S3 bucket

Amazon S3

Data target properties - S3

Output schema

Data preview

Name

Amazon S3

Node parents

Choose which nodes will provide inputs for this one.

Choose one or more parent node

SQL Query

SQL Code - Transform

Format

JSON

Compression Type

None

S3 Target Location

Choose an S3 location in the format s3://bucket/prefix/object/ with a trailing slash (/).

s3://lb-01/lab-glue/frequencia\_registro\_nomes\_eua/

View

Browse S3

Data Catalog update options Info

Choose how you want to update the Data Catalog table's schema and partitions. These options will only apply if the Data Catalog table is an S3 backed source.

☒ Do not update the Data Catalog

☐ Create a table in the Data Catalog and on subsequent runs, update the schema and add new partitions

☐ Create a table in the Data Catalog and on subsequent runs, keep existing schema and add new partitions

Partition keys - optional

Add partition keys.

Partition (0)

sexo

Partition (1)

ano

Add a partition key

Services

Search

[Alt+S]

Amazon S3

Visualizar painel do Storage Lens

Buckets

Pontos de acesso

Pontos de acesso Lambda de objeto

Pontos de acesso de várias regiões

Operações em lotes

IAM Access Analyzer para S3

Configurações de bloqueio do acesso público desta conta

Storage Lens

Painéis

Configurações do AWS Organizations

Recurso em destaque

Amazon S3 > Buckets

Snapshot da conta

Última atualização: 24 Sep 2023 pelo Storage Lens. Métricas são geradas a cada 24 horas. Saiba mais

Visualizar painel do Storage Lens

Armazenamento total

29.4 MB

Quantidade de objetos

1

Tamanho médio do objeto

29.4 MB

Você pode habilitar métricas avançadas na configuração Configuração do "default-account-dashboard".

Buckets (2) Informações

Os buckets são contêineres para dados armazenados no S3. Saiba mais

Encontrar buckets por nome

	Nome	Região da AWS	Acesso	Data de criação
<input type="radio"/>	aws-glue-assets-369659558219-us-east-1	Leste dos EUA (Norte da Virgínia) us-east-1	Bucket e objetos não públicos	26 Sep 2023 10:50:48 PM -03
<input type="radio"/>	lb-01	Leste dos EUA (Norte da Virgínia) us-east-1	Bucket e objetos não públicos	26 Sep 2023 06:45:21 PM -03

Services

Search

[Alt+S]

Amazon S3

Copiar URI do S3

Buckets

Pontos de acesso

Pontos de acesso Lambda de objeto

Pontos de acesso de várias regiões

Operações em lotes

IAM Access Analyzer para S3

Configurações de bloqueio do acesso público desta conta

Storage Lens

Painéis

Configurações do AWS Organizations

Recurso em destaque

Amazon S3 > Buckets > lb-01 > lab-glue/

lab-glue/

Copiar URI do S3

Objetos

Propriedades

Objetos (2)

Os objetos são as entidades fundamentais armazenadas no Amazon S3. Você pode usar o [inventário do Amazon S3](#) para obter uma lista de todos os objetos em seu bucket. Para outras pessoas acessarem seus objetos, você precisará conceder permissões explicitamente a eles. Saiba mais

Copiar URI do S3

Copiar URL

Fazer download

Abrir

Excluir

Ações

Criar pasta

Carregar

Localizar objetos por prefixo

	Nome	Tipo	Última modificação	Tamanho	Classe de armazenamento
<input type="checkbox"/>	frequencia_registro_nomes_eua/	Pasta	-	-	-
<input type="checkbox"/>	input/	Pasta	-	-	-

Amazon S3

Buckets

Pontos de acesso

Pontos de acesso Lambda de objeto

Pontos de acesso de várias regiões

Operações em lotes

IAM Access Analyzer para S3

Configurações de bloqueio do acesso público desta conta

Storage Lens

Painéis

Configurações do AWS Organizations

Recurso em destaque

Amazon S3

Buckets

lb-01

lab-blue/

frequencia\_registro\_nomes\_eua/

Copiar URI do S3

Objetos

Propriedades

Objetos (2)

Os objetos são as entidades fundamentais armazenadas no Amazon S3. Você pode usar o [inventário do Amazon S3](#) para obter uma lista de todos os objetos em seu bucket. Para outras pessoas acessarem seus objetos, você precisará conceder permissões explicitamente a eles. [Saiba mais](#)

Copiar URI do S3

Copiar URL

Fazer download

Abriu

Excluir

Ações

Criar pasta

Carregar

Localizar objetos por prefixo

	Nome	Tipo	Última modificação	Tamanho	Classe de armazenamento
<input type="checkbox"/>	sexo=F/	Pasta	-	-	-
<input type="checkbox"/>	sexo=M/	Pasta	-	-	-

Amazon S3

Buckets

Pontos de acesso

Pontos de acesso Lambda de objeto

Pontos de acesso de várias regiões

Operações em lotes

IAM Access Analyzer para S3

Configurações de bloqueio do acesso público desta conta

Storage Lens

Painéis

Configurações do AWS Organizations

Recurso em destaque

Amazon S3

Buckets

lb-01

lab-blue/

frequencia\_registro\_nomes\_eua/

sexo=F/

Copiar URI do S3

Objetos

Propriedades

Objetos (135)

Os objetos são as entidades fundamentais armazenadas no Amazon S3. Você pode usar o [inventário do Amazon S3](#) para obter uma lista de todos os objetos em seu bucket. Para outras pessoas acessarem seus objetos, você precisará conceder permissões explicitamente a eles. [Saiba mais](#)

Copiar URI do S3

Copiar URL

Fazer download

Abriu

Excluir

Ações

Criar pasta

Carregar

Localizar objetos por prefixo

	Nome	Tipo	Última modificação	Tamanho	Classe de armazenamento
<input type="checkbox"/>	ano=1880/	Pasta	-	-	-
<input type="checkbox"/>	ano=1881/	Pasta	-	-	-
<input type="checkbox"/>	ano=1882/	Pasta	-	-	-
<input type="checkbox"/>	ano=1883/	Pasta	-	-	-
<input type="checkbox"/>	ano=1884/	Pasta	-	-	-
<input type="checkbox"/>	ano=1885/	Pasta	-	-	-
<input type="checkbox"/>	ano=1886/	Pasta	-	-	-
<input type="checkbox"/>	ano=1887/	Pasta	-	-	-
<input type="checkbox"/>	ano=1888/	Pasta	-	-	-
<input type="checkbox"/>	ano=1889/	Pasta	-	-	-
<input type="checkbox"/>	ano=1890/	Pasta	-	-	-
<input type="checkbox"/>	ano=1891/	Pasta	-	-	-
<input type="checkbox"/>	ano=1892/	Pasta	-	-	-
<input type="checkbox"/>	ano=1893/	Pasta	-	-	-
<input type="checkbox"/>	ano=1894/	Pasta	-	-	-
<input type="checkbox"/>	ano=1895/	Pasta	-	-	-
<input type="checkbox"/>	ano=1896/	Pasta	-	-	-
<input type="checkbox"/>	ano=1897/	Pasta	-	-	-
<input type="checkbox"/>	ano=1898/	Pasta	-	-	-

Amazon S3

Buckets

Pontos de acesso

Pontos de acesso Lambda de objeto

Pontos de acesso de várias regiões

Operações em lotes

IAM Access Analyzer para S3

Configurações de bloqueio do acesso público desta conta

Storage Lens

Painéis

Configurações do AWS Organizations

Recurso em destaque

Amazon S3

Buckets

lb-01

lab-blue/

frequencia\_registro\_nomes\_eua/

sexo=M/

Copiar URI do S3

Objetos

Propriedades

Objetos (135)

Os objetos são as entidades fundamentais armazenadas no Amazon S3. Você pode usar o [inventário do Amazon S3](#) para obter uma lista de todos os objetos em seu bucket. Para outras pessoas acessarem seus objetos, você precisará conceder permissões explicitamente a eles. [Saiba mais](#)

Copiar URI do S3

Copiar URL

Fazer download

Abriu

Excluir

Ações

Criar pasta

Carregar

Localizar objetos por prefixo

	Nome	Tipo	Última modificação	Tamanho	Classe de armazenamento
<input type="checkbox"/>	ano=1880/	Pasta	-	-	-
<input type="checkbox"/>	ano=1881/	Pasta	-	-	-
<input type="checkbox"/>	ano=1882/	Pasta	-	-	-
<input type="checkbox"/>	ano=1883/	Pasta	-	-	-
<input type="checkbox"/>	ano=1884/	Pasta	-	-	-
<input type="checkbox"/>	ano=1885/	Pasta	-	-	-
<input type="checkbox"/>	ano=1886/	Pasta	-	-	-
<input type="checkbox"/>	ano=1887/	Pasta	-	-	-
<input type="checkbox"/>	ano=1888/	Pasta	-	-	-
<input type="checkbox"/>	ano=1889/	Pasta	-	-	-
<input type="checkbox"/>	ano=1890/	Pasta	-	-	-
<input type="checkbox"/>	ano=1891/	Pasta	-	-	-
<input type="checkbox"/>	ano=1892/	Pasta	-	-	-
<input type="checkbox"/>	ano=1893/	Pasta	-	-	-
<input type="checkbox"/>	ano=1894/	Pasta	-	-	-
<input type="checkbox"/>	ano=1895/	Pasta	-	-	-
<input type="checkbox"/>	ano=1896/	Pasta	-	-	-
<input type="checkbox"/>	ano=1897/	Pasta	-	-	-
<input type="checkbox"/>	ano=1898/	Pasta	-	-	-

CloudShell

Comentários

© 2023, Amazon Web Services, Inc. ou suas afiliadas.

Privacidade

Termos

Preferências de cookies

Amazon Athena

Query editor

Editor

Recent queries

Saved queries

Settings

Workgroup primary

Data

Data source

AwsDataCatalog

Database

glue-lab

Tables and views

Create

Filter tables and views

Tables (2)

1

freuencia\_registro\_nomes\_eua

Partitioned

lb\_01

Partitioned

Views (0)

1

Query 14

Query 15

1 DESCRIBE lb\_01;

2

SQL

Ln 1, Col 16

Run again

Explain

Cancel

Clear

Create

Reuse query results up to 60 minutes ago

Query results

Query stats

Completed

Time in queue: 323 ms

Run time: 1.068 sec

Data scanned: -

nome string

sexo string

total bigint

ano bigint

partition\_0 string

partition\_1 string

# Partition Information

# col\_name data\_type comment

partition\_0 string

partition\_1 string

CloudShell

Comentários

© 2023, Amazon Web Services, Inc. ou suas afiliadas.

Privacidade

Termos

Preferências de cookies

Amazon Athena

Query editor

Editor

Recent queries

Saved queries

Settings

Workgroup primary

Data

Data source

AwsDataCatalog

Database

glue-lab

Tables and views

Create

Filter tables and views

Tables (2)

1

freuencia\_registro\_nomes\_eua

Partitioned

lb\_01

Partitioned

Views (0)

1

Query 14

1 -- Alterar a caixa dos valores da coluna nome para MAIUSCULO

2 SELECT UPPER(nome) AS nome\_maiusculo, sexo, total, ano FROM lb\_01;

3

SQL

Ln 3, Col 1

Run again

Explain

Cancel

Clear

Create

Reuse query results up to 60 minutes ago

Query results

Query stats

Completed

Time in queue: 234 ms

Run time: 3.618 sec

Data scanned: 29.43 MB

Results (1.825.433)

Copy

Download results

Search rows

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

# nome\_maiusculo sexo total ano

1 JENNIFER F 54336 1983

2 JESSICA F 45278 1983

3 AMANDA F 33752 1983

4 ASHLEY F 33292 1983

5 SARAH F 27228 1983

6 MELISSA F 23472 1983

7 NICOLE F 22392 1983

8 STEPHANIE F 22323 1983

9 HEATHER F 20749 1983

10 ELIZABETH F 19838 1983

11 CRYSTAL F 17904 1983

12 AMY F 17095 1983

13 MICHELLE F 16828 1983

14 TIFFANY F 15960 1983

15 KIMBERLY F 15374 1983

Services

Search

[Alt+S]

Norte da Virginia

AdministratorAccess/marlon.barnes.ph@compas.com.br

Amazon Athena

Query editor

Editor

Recent queries

Saved queries

Settings

Workgroup

primary

Data

Data source

AwsDataCatalog

Database

glue- lab

Tables and views

Create

Filter tables and views

Tables (0)

1

frequncia\_registro\_nomes\_cus

Partitioned

lb\_01

Partitioned

Views (0)

1

Query 14

Query 15

1 --:Imprimir a contagem de linhas presentes no dataframe

2 SELECT COUNT(\*) FROM lb\_01;

SQL

Ln 2, Col 28

Run again

Explain

Cancel

Clear

Create

Reuse query results

up to 60 minutes ago

Query results

Query stats

Completed

Time in queue: 156 ms

Run time: 1.126 sec

Data scanned: 29.43 MB

Results (1)

Copy

Download results

Search rows

#

▼

▼

\_col0

▼

1

1825453

Services

Search

[Alt+S]

Norte da Virginia

AdministratorAccess/marlon.barnes.ph@compas.com.br

Amazon Athena

Query editor

Editor

Recent queries

Saved queries

Settings

Workgroup

primary

Data

Data source

AwsDataCatalog

Database

glue- lab

Tables and views

Create

Filter tables and views

Tables (0)

1

frequncia\_registro\_nomes\_cus

Partitioned

lb\_01

Partitioned

Views (0)

1

Query 14

Query 15

1 --: Imprimir a contagem de nomes, agrupando os dados do dataframe pelas colunas ano e sexo

2 SELECT ano, sexo, COUNT(nome) FROM lb\_01 GROUP BY ano, sexo;

SQL

Ln 2, Col 61

Run again

Explain

Cancel

Clear

Create

Reuse query results

up to 60 minutes ago

Query results

Query stats

Completed

Time in queue: 188 ms

Run time: 2.179 sec

Data scanned: 29.43 MB

Results (270)

Copy

Download results

Search rows

#

▼

▼

ano

▼

▼

sexo

▼

▼

\_col2

▼

1

1983

M

7333

2

1984

M

7332

3

1987

M

8145

4

1988

F

13874

5

1988

M

8484

6

1989

M

9225

7

1992

F

15605

8

1994

M

10244

9

1996

M

10530

10

2000

M

12111

11

2001

M

12295

12

2006

F

20043

13

2009

F

20165

14

2013

F

19191

15

2013

M

14012

16

1881

F

938

CloudShell

Comentários

© 2023, Amazon Web Services, Inc. ou suas afiliadas.

Privacidade

Termos

Preferências de cookies

Services

Search

[Alt+S]

Norte da VirginiaAdministratorAccess/martins.barros.pb@compasso.com.br

Amazon AthenaQuery editor

EditorRecent queriesSaved queriesSettings

Workgroupprimary

Data

Data source

AwsDataCatalog

Database

glue- lab

Tables and views

Create

Filter tables and views

Tables (0)

1

freuencia\_registro\_nomes\_eua

Partitioned

lb\_01

Partitioned

Views (0)

1

Query 14Query 15

1 ... Ordenar os dados de modo que o ano mais recente apareça como primeiro registro do dataframe

2 SELECT \* FROM lb\_01 ORDER BY ano DESC;

SQLLn 1, Col 2

Run againExplainCancelClearCreate

Reuse query results up to 60 minutes ago

Query resultsQuery stats

Completed

Time in queue: 186 msRun time: 5.846 secData scanned: 29.43 MB

Results (1.825.433)

Search rows

CopyDownload results

116

#	nome	sexo	total	ano	partition_0	partition_1
1	Chaison	M	6	2014	lab-glue	input
2	Treyvon	M	80	2014	lab-glue	input
3	Chadd	M	6	2014	lab-glue	input
4	Tremaine	M	80	2014	lab-glue	input
5	Cesario	M	6	2014	lab-glue	input
6	Romello	M	80	2014	lab-glue	input
7	Ceazar	M	6	2014	lab-glue	input
8	Kentrell	M	80	2014	lab-glue	input
9	Caylin	M	6	2014	lab-glue	input
10	Kasyn	M	80	2014	lab-glue	input
11	Cayse	M	6	2014	lab-glue	input
12	Job	M	80	2014	lab-glue	input
13	Cayn	M	6	2014	lab-glue	input
14	Jimmie	M	80	2014	lab-glue	input
15	Caylin	M	6	2014	lab-glue	input
16	Gabe	M	80	2014	lab-glue	input

CloudShellComentários© 2023, Amazon Web Services, Inc. ou suas afiliadas.PrivacidadeTermosPreferências de cookies

Services

Search

[Alt+S]

Norte da VirginiaAdministratorAccess/martins.barros.pb@compasso.com.br

Amazon AthenaQuery editor

EditorRecent queriesSaved queriesSettings

Workgroupprimary

Data

Data source

AwsDataCatalog

Database

glue- lab

Tables and views

Create

Filter tables and views

Tables (0)

1

freuencia\_registro\_nomes\_eua

Partitioned

lb\_01

Partitioned

Views (0)

1

Query 14Query 15

1 ... Nome feminino com mais registros e em que ano ocorreu

2 SELECT nome, ano, COUNT(\*) as count FROM lb\_01 WHERE sexo = 'F' GROUP BY nome, ano ORDER BY count DESC LIMIT 1;

SQLLn 2, Col 112

Run againExplainCancelClearCreate

Reuse query results up to 60 minutes ago

Query resultsQuery stats

Completed

Time in queue: 298 msRun time: 2.46 secData scanned: 29.43 MB

Results (1)

Search rows

CopyDownload results

11

#	nome	ano	count
1	Jessica	1983	1

CloudShellComentários© 2023, Amazon Web Services, Inc. ou suas afiliadas.PrivacidadeTermosPreferências de cookies

Amazon Athena > Query editor

Editor Recent queries Saved queries Settings Workgroup: primary

Data source: AwsDataCatalog Database: glue-lab

Tables and views: Create Filter tables and views

Tables (0) < 1 > frequencia\_registro\_nomes\_eua Partitioned lb\_01 Partitioned Views (0) < 1 >

Query 14 Query 15

```
1 -- Nome masculino com mais registros e em que ano ocorreu
2 SELECT nome, ano, COUNT(*) AS count FROM lb_01 WHERE sexo = 'M' GROUP BY nome, ano ORDER BY count DESC LIMIT 1;
```

SQL Ln 2, Col 42

Run again Explain Cancel Clear Create Reuse query results up to 60 minutes ago

Query results Query stats

Completed Time in queue: 203 ms Run time: 2.129 sec Data scanned: 29.43 MB

Results (1) Copy Download results

Search rows

#	nome	ano	count
1	Chaites	1984	1

Amazon Athena > Query editor

Editor Recent queries Saved queries Settings Workgroup: primary

Data source: AwsDataCatalog Database: glue-lab

Tables and views: Create Filter tables and views

Tables (0) < 1 > frequencia\_registro\_nomes\_eua Partitioned lb\_01 Partitioned Views (0) < 1 >

Query 14 Query 15

```
1 -- Total de registros (masculinos e femininos) para cada ano presente no dataframe
2 SELECT ano, COUNT(*) FROM lb_01 GROUP BY ano;
```

SQL Ln 2, Col 27

Run again Explain Cancel Clear Create Reuse query results up to 60 minutes ago

Query results Query stats

Completed Time in queue: 597 ms Run time: 1.426 sec Data scanned: 29.43 MB

Results (155) Copy Download results

Search rows

#	ano	_col1
1	1988	22358
2	2005	51173
3	2004	32035
4	2005	32533
5	2006	34069
6	2009	34684
7	1880	2000
8	1888	2651
9	1895	3049
10	1898	3264
11	1901	3153
12	1903	3389
13	1909	4227
14	1911	4867
15	1919	10370
16	1922	10757

CloudShell Comentários

© 2023, Amazon Web Services, Inc. ou suas afiliadas. Privacidade Termos Preferências de cookies