

Evaluación: Módulos 2 y 3 AWS Data Engineering

- Identificar las principales características y funciones de Amazon S3 que permiten el acceso y análisis de datos.
- Describir cómo se utilizan los servicios de Amazon S3 para almacenar y recuperar grandes volúmenes de datos.
- Configurar un bucket de Amazon S3 para almacenar datos de manera eficiente. • Comparar las ventajas de usar Amazon S3 frente a otras soluciones de almacenamiento de datos en términos de coste y rendimiento.
- Evaluar las implicaciones de seguridad y privacidad al utilizar Amazon S3 en proyectos de análisis de datos
- Enumerar los cinco Vs de los datos: volumen, velocidad, variedad, veracidad y valor. • Explicar cómo cada uno de los cinco Vs afecta la captura, el procesamiento y la visualización de datos.
- Utilizar un ejemplo práctico para demostrar cómo planificar y configurar una pipeline de datos considerando los cinco Vs
- Analizar los diferentes tipos de variabilidad de los datos (velocidad, volumen, etc.) influyen en las decisiones tecnológicas para proyectos de big data. • datos haya impactado negativamente en los resultados del análisis de datos. • Desarrollar un plan estratégico para implementar una pipeline de datos efectiva que maneje adecuadamente los cinco Vs y maximice el valor de los datos para una organización específica.

Instrucciones de entrega:

- Crea un documento en markdown que incluya una breve explicación de las partes que creas más importantes de tus respuestas junto con las capturas de pantalla o texto correspondiente del laboratorio: Lab: Accessing and Analyzing Data by Using Amazon S3
- Para el módulo 3 presenta la verificación del Knowledge Check
- Subir el documento final a tu repositorio personal hasta el día 5 de mayo (23:59).

Identificar las principales características y funciones de Amazon S3 que permiten el acceso y análisis de datos:

- Amazon S3 es un servicio de almacenamiento de objetos escalable y duradero, que permite almacenar y recuperar cualquier cantidad de datos, desde cualquier lugar de la web.
- Algunas características clave de S3 que facilitan el acceso y análisis de datos son:
 - Escalabilidad ilimitada para almacenar grandes volúmenes de datos.
 - Múltiples clases de almacenamiento para optimizar el costo y el acceso.
 - Integración con otros servicios AWS como Athena, Redshift y EMR para análisis de datos.
 - Controles de acceso y seguridad flexibles para gestionar permisos.

Describir cómo se utilizan los servicios de Amazon S3 para almacenar y recuperar grandes volúmenes de datos:

- Los datos se pueden cargar en S3 a través de diversos mecanismos como la consola web, la CLI o SDK.
- Los objetos almacenados en S3 se organizan en "buckets" (cubos) que actúan como contenedores lógicos.
- S3 escala automáticamente para almacenar y procesar grandes cantidades de datos sin necesidad de aprovisionar.
- Los datos se pueden recuperar a través de consultas y descargas, o integrándolos con otros servicios analíticos de AWS.

Configurar un bucket de Amazon S3 para almacenar datos de manera eficiente:

- Definir una estructura de carpetas y nomenclatura para organizar los datos.
- Elegir la clase de almacenamiento adecuada según el patrón de acceso esperado.
- Habilitar características como versionamiento, ciclo de vida y cifrado para proteger los datos.
- Configurar políticas de acceso y permisos para controlar quién puede acceder a los datos.

Comparar las ventajas de usar Amazon S3 frente a otras soluciones de almacenamiento de datos en términos de costo y rendimiento:

- S3 ofrece un almacenamiento escalable y de bajo costo, especialmente para grandes volúmenes de datos.
- Proporciona un alto rendimiento y disponibilidad, adecuado para cargas de trabajo de análisis.
- Permite pagar solo por lo que se usa, sin necesidad de aprovisionar capacidad.
- Otras soluciones pueden tener costos más altos o requerir más mantenimiento y gestión por parte del usuario.

Evaluar las implicaciones de seguridad y privacidad al utilizar Amazon S3 en proyectos de análisis de datos:

- S3 proporciona controles de acceso flexibles y cifrado de datos en reposo y en tránsito.
- Es necesario configurar adecuadamente los permisos y políticas de acceso para evitar accesos no autorizados.
- Hay que considerar la ubicación geográfica de los buckets y cumplir con las regulaciones de privacidad de datos.
- Implementar prácticas de seguridad como registro de auditoría, monitoreo y

protección contra pérdida de datos.

Enumerar los cinco Vs de los datos: volumen, velocidad, variedad, veracidad y valor.

Explicar cómo cada uno de los cinco Vs afecta la captura, el procesamiento y la visualización de datos:

- **Volumen:** Impacta el diseño de la infraestructura de almacenamiento y procesamiento para manejar grandes cantidades de datos.
- **Velocidad:** Afecta la elección de tecnologías y arquitecturas para ingerir, procesar y analizar datos en tiempo real o por lotes.
- **Variedad:** Exige la capacidad de manejar diferentes formatos y estructuras de datos, y realizar transformaciones para integrarlos.
- **Veracidad:** Requiere procesos de limpieza, validación y gobernanza de datos para garantizar su integridad y fiabilidad.
- **Valor:** Determina qué datos son relevantes y valiosos para los objetivos de negocio, guiando las decisiones de captura y análisis.

Utilizar un ejemplo práctico para demostrar cómo planificar y configurar una pipeline de datos considerando los cinco Vs:

Por ejemplo:

Una empresa de comercio electrónico que quiere analizar el comportamiento de los clientes en su sitio web.

Volumen: Capturar y procesar registros de clics y eventos de navegación a gran escala.

Velocidad: Ingerir y procesar los datos de clics en tiempo real para realizar recomendaciones personalizadas.

Variedad: Integrar datos estructurados (transacciones) y no estructurados (text, imágenes) de diferentes fuentes.

Veracidad: Implementar procesos de limpieza y validación de datos para garantizar su integridad.

Valor: Identificar los datos más relevantes para analizar el comportamiento del cliente y mejorar la experiencia.

Analizar cómo los diferentes tipos de variabilidad de los datos (velocidad, volumen, etc.) influyen en las decisiones tecnológicas para proyectos de big data:

- El alto volumen de datos requerirá soluciones de almacenamiento escalables como Amazon S3.
- La necesidad de procesar datos en tiempo real implicará el uso de tecnologías de streaming como Amazon Kinesis.
- La variedad de datos estructurados y no estructurados sugerirá la adopción de plataformas flexibles como Amazon EMR.
- La veracidad de los datos impulsará la implementación de procesos de limpieza, validación y gobernanza.

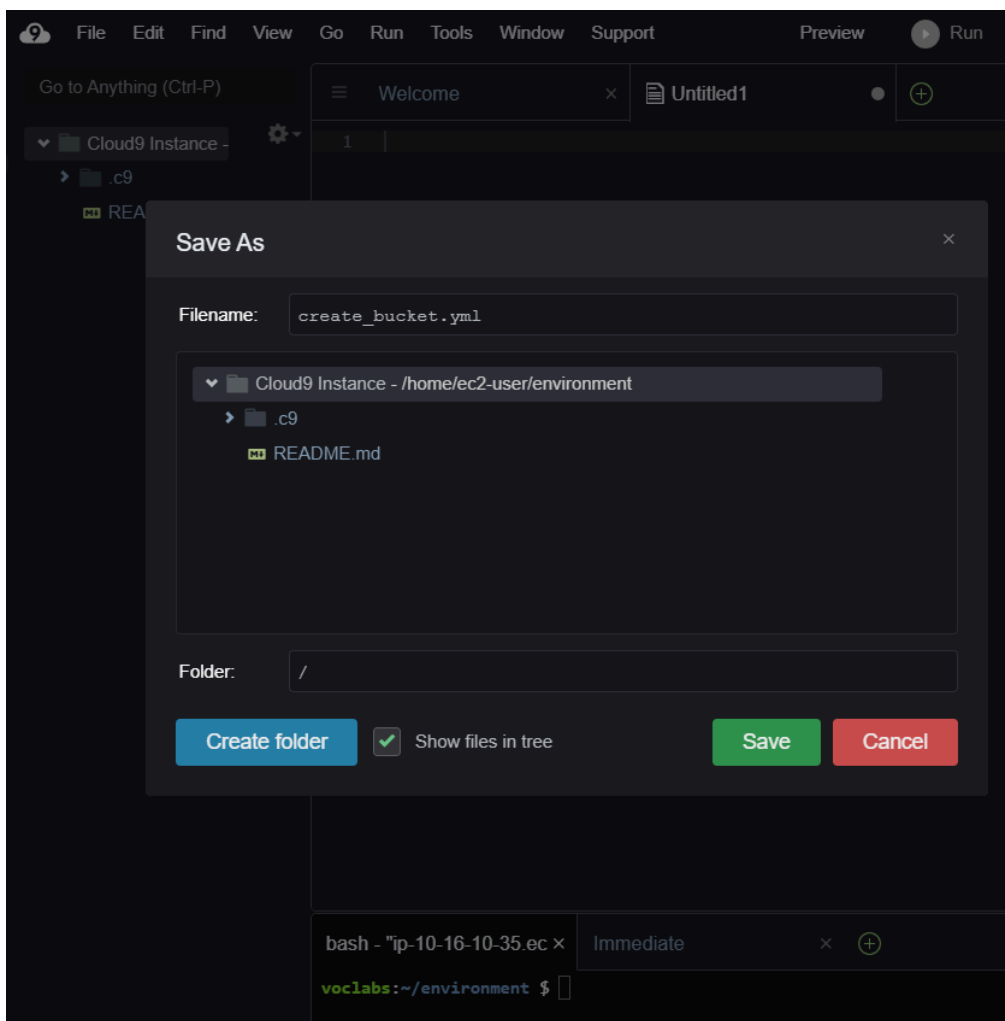
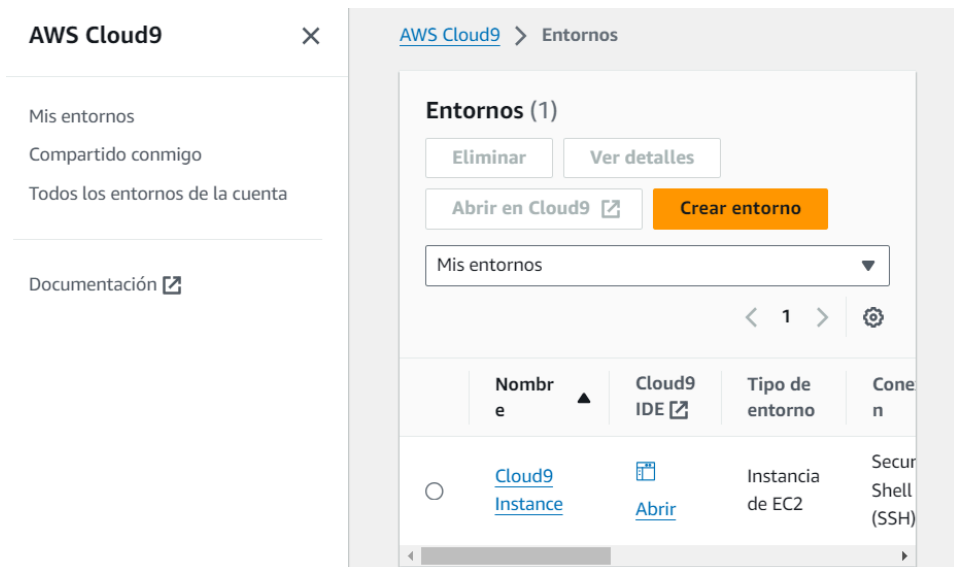
Desarrollar un plan estratégico para implementar una pipeline de datos efectiva que maneje adecuadamente los cinco Vs y maximice el valor de los datos para una organización específica:

- Evaluar las necesidades de la organización en cuanto a los cinco Vs de los datos.
- Diseñar una arquitectura de pipeline de datos que incluya servicios como S3, Kinesis, Glue, Athena, etc.
- Implementar procesos de gobernanza de datos, limpieza, transformación e integración.
- Establecer mecanismos de visualización y análisis para extraer insights valiosos.

- Monitorizar y optimizar continuamente la pipeline para adaptarse a los cambios en los requisitos de datos.

Lab: Accessing and Analyzing Data by Using Amazon S3

Tarea 1: Crear una plantilla y una pila de CloudFormation



```
1 AWSTemplateFormatVersion: "2010-09-09"
2 Description: "TEST for S3 infrastructure"
3 Resources:
4   S3Bucket:
5     Type: "AWS::S3::Bucket"
6     Properties:
7       BucketName: !Join
8         - "-"
9         - "ade-my-bucket"
10        - !Select
11          - 0
12          - !Split
13            - "-"
14            - !Select
15              - 2
16              - !Split
17                - "/"
18                - !Ref "AWS::StackId"
19        PublicAccessBlockConfiguration:
20          BlockPublicAcls: True
21          BlockPublicPolicy: True
22          IgnorePublicAcls: True
23          RestrictPublicBuckets: True
24        VersioningConfiguration:
25          Status: Enabled
26        BucketEncryption:
27          ServerSideEncryptionConfiguration:
28            - ServerSideEncryptionByDefault:
29              SSEAlgorithm: 'aws:kms'
30              KMSMasterKeyID: KMS-KEY-ARN
```

Buckets de uso general (1) Información Todas las regiones de AWS

Buscar buckets por nombre

Copiar ARN

V

Los buckets son contenedores de datos almacenados en S3.

	Nombre ▲	Región de AWS ▼	Analizador de acceso de IAM	Fecha de creación
<div></div>	ade-s3lab-bucket--b20d4860	EE. UU. Este (Norte de Virginia) us-east-1	Ver analizador para us-east-1	5 May 2024 5:56:30 PM -05

```
bash - "ip-10-16-10-35.ec × Immediate × +
voclabs:~/environment $ aws cloudformation validate-template --template-body file:///create_bucket.yml
{
  "Parameters": [],
  "Description": "TEST for S3 infrastructure"
}
voclabs:~/environment $
```

```
voclabs:~/environment $ aws cloudformation create-stack --stack-name ade-my-bucket --template-body file:///create_bucket.yml
{
  "StackId": "arn:aws:cloudformation:us-east-1:050670574482:stack/ade-my-bucket/11c28e30-0b35-11ef-b54a-0affcfa1a77b"
}
voclabs:~/environment $
```

Buckets de uso general (2)

Información

Todas las regiones de AWS

Copiar ARN

Vaci

Los buckets son contenedores de datos almacenados en S3.

Q

Buscar buckets por nombre

	Nombre ▲	Región de AWS ▼	Analizador de acceso de IAM	Fecha de creación
<input type="radio"/>	ade-my-bucket-11c28e30	EE. UU. Este (Norte de Virginia) us-east-1	Ver analizador para us-east-1	5 May 2024 6:13:29 PM -05
<input type="radio"/>	ade-s3lab-bucket--b20d4860	EE. UU. Este (Norte de Virginia) us-east-1	Ver analizador para us-east-1	5 May 2024 5:56:30 PM -05

```
voclabs:~/environment $ aws s3api list-buckets
{
  "Buckets": [
    {
      "Name": "ade-my-bucket-11c28e30",
      "CreationDate": "2024-05-05T23:13:29+00:00"
    },
    {
      "Name": "ade-s3lab-bucket--b20d4860",
      "CreationDate": "2024-05-05T22:56:30+00:00"
    }
  ],
  "Owner": {
    "DisplayName": "awslabsc0w4896291t1669396780",
    "ID": "76b269964e7013683bd5009a765ed8a5cd6194992e05f6be668e388001afce7c"
  }
}
voclabs:~/environment $
```

elimina

```
voclabs:~/environment $ aws cloudformation delete-stack --stack-name ade-my-bucket
voclabs:~/environment $
```

Buckets de uso general (1) Información

Todas las regiones de AWS

Copiar ARN

Vaciar

Los buckets son contenedores de datos almacenados en S3.

Q

Buscar buckets por nombre

	Nombre	Región de AWS	Analizador de acceso de IAM	Fecha de creación
	ade-s3lab-bucket--b20d4860	EE. UU. Este (Norte de Virginia) us-east-1	Ver analizador para us-east-1	5 May 2024 5:56:30 PM -05

```
{
  "StackSummaries": [
    {
      "StackId": "arn:aws:cloudformation:us-east-1:050670574482:stack/ade-my-bucket/11c28e30-0b35-11ef-b54a-0affcf1a77b",
      "StackName": "ade-my-bucket",
      "TemplateDescription": "TEST for S3 Infrastructure",
      "CreationTime": "2024-05-05T23:13:23.153000+00:00",
      "DeletionTime": "2024-05-05T23:20:09.134000+00:00",
      "StackStatus": "DELETE_COMPLETE",
      "DriftInformation": {
        "StackDriftStatus": "NOT_CHECKED"
      }
    },
    {
      "StackId": "arn:aws:cloudformation:us-east-1:050670574482:stack/aws-cloud9-Cloud9-Instance-77f0847c0e144a68b1b3e86af7b78a91/c100cc20-0b32-11ef-8f14-0e0b789071eb",
      "StackName": "aws-cloud9-Cloud9-Instance-77f0847c0e144a68b1b3e86af7b78a91",
      "CreationTime": "2024-05-05T22:56:48.664000+00:00",
      "StackStatus": "CREATE_COMPLETE",
    }
  ]
}
```

```
voclabs:~/environment $ aws s3api list-buckets
{
  "Buckets": [
    {
      "Name": "ade-s3lab-bucket--b20d4860",
      "CreationDate": "2024-05-05T22:56:30+00:00"
    }
  ],
  "Owner": {
    "DisplayName": "awslabsc0w4896291t1669396780",
    "ID": "76b269964e7013683bd5009a765ed8a5cd6194992e05f6be668e388001afce7c"
  }
}
```

Tarea 2: cargar datos de muestra en un depósito de S3

```
voclabs:~/environment $ wget https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-200-ACDSCI-1-DEV/lab-01-s3/code.zip -P /home/ec2-user/environment
--2024-05-05 23:27:19-- https://aws-tc-largeobjects.s3.us-west-2.amazonaws.com/CUR-TF-200-ACDSCI-1-DEV/lab-01-s3/code.ziplargeobjects.s3.us-west-2.amazonaws.com/CUR-TF-200-ACDSCI-1-DEV/1
Resolving aws-tc-largeobjects.s3.us-west-2.amazonaws.com (aws-tc-largeobjects.s3.us-west-2.amazonaws.com)... 3.5.87.117, 52.92.130.162, 52.92.205.226, ...
Connecting to aws-tc-largeobjects.s3.us-west-2.amazonaws.com (aws-tc-largeobjects.s3.us-west-2.amazonaws.com)[3.5.87.117]:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 751 [application/zip]
Saving to: '/home/ec2-user/environment/code.zip'

100%[----->] 751 --.-K/s in 0s

2024-05-05 23:27:20 (13.8 MB/s) - '/home/ec2-user/environment/code.zip' saved [751/751]

voclabs:~/environment $ ll
total 12
-rw-rw-r-- 1 ec2-user ec2-user 751 Sep  3 2022 code.zip
-rw-r--r-- 1 ec2-user ec2-user 836 May  5 23:04 create_bucket.yml
-rw-r--r-- 1 ec2-user ec2-user 569 Apr 30 02:48 README.md
voclabs:~/environment $
```

```
voclabs:~/environment $ unzip code.zip
Archive:  code.zip
  inflating: lab1.csv
replace code.zip? [y]es, [n]o, [A]ll, [N]one, [r]ename: n
voclabs:~/environment $ ll
total 16
-rw-rw-r-- 1 ec2-user ec2-user 751 Sep  3  2022 code.zip
-rw-r--r-- 1 ec2-user ec2-user 836 May  5 23:04 create_bucket.yml
-rwxr-xr-x 1 ec2-user ec2-user 328 Sep  2  2022 lab1.csv
-rw-r--r-- 1 ec2-user ec2-user 569 Apr 30 02:48 README.md
voclabs:~/environment $
```

```
voclabs:~/environment $ cat lab1.csv
CustomerID,First Name,Last Name,Join Date,Street Address,City,State,Phone
1,Alejandro,Rosalez,12/12/2013,123 Main St.,Any Town,MD,301-555-0158
2,Jane,Doe,10/5/2014,456 State St.,Anywhere,WA,360-555-0163
3,John,Stiles,9/20/2016,1980 8th St.,Nowhere,NY,914-555-0122
4,Li,Juan,6/29/2011,1323 22nd Ave.,Anytown,NY,914-555-0149
voclabs:~/environment $
```

```
voclabs:~/environment $ aws s3api list-buckets
{
  "Buckets": [
    {
      "Name": "ade-s3lab-bucket--b20d4860",
      "CreationDate": "2024-05-05T22:56:30+00:00"
    }
  ],
  "Owner": {
    "DisplayName": "awslabsc0w4896291t1669396780",
    "ID": "76b269964e7013683bd5009a765ed8a5cd6194992e05f6be668e388001afce7c"
  }
}
```

```
voclabs:~/environment $ aws s3 cp lab1.csv s3://ade-s3lab-bucket--b20d4860
upload: ./lab1.csv to s3://ade-s3lab-bucket--b20d4860/lab1.csv
voclabs:~/environment $
```

ade-s3lab-bucket--b20d4860 [Información](#)

[Objetos](#)
[Propiedades](#)
[Permisos](#)
[Métricas](#)
[Administración](#)
[Puntos de acceso](#)

Objetos (1) [Información](#)

Acciones ▾

Cargar

Los objetos son las entidades fundamentales que se almacenan en Amazon S3. Puede utilizar el [inventario de Amazon S3](#) para obtener una lista de todos los objetos de su bucket. Para que otras personas obtengan acceso a sus objetos, tendrá que concederles permisos de forma explícita. [Más información](#)

< 1 > ⚙️

<input type="checkbox"/>	Nombre	Tipo	Última modificación	Tamaño	Clase de almacenamiento
<input type="checkbox"/>	lab1.csv	csv	5 May 2024 6:35:38 PM -05	328.0 B	Estándar


```
voclabs:~/environment $ aws s3 ls s3://ade-s3lab-bucket--b20d4860
2024-05-05 23:35:38          328 lab1.csv
voclabs:~/environment $
```

Tarea 3: Consultar los datos

Configuración de entrada

Ruta

s3://ade-s3lab-bucket--b20d4860/lab1.csv

Tamaño

328.0 B

Formato

☒ CSV

☐ JSON

☐ Apache Parquet

CSV delimitador

☒ Coma

☐ Tabulador

☐ Personalizado

☒ Excluir la primera línea de CSV datos

Habilite esta configuración si CSV contiene una fila de encabezado.

Compresión

☒ Ninguno

☐ GZIP

☐ BZIP2

1

/* Para crear un punto de referencia para escribir consultas SQL, puede mostrar los primeros 5 registros de datos de entrada ejecutando la siguiente consulta SQL: SELECT * FROM s3object s LIMIT 5 */

2

SELECT * FROM s3object s LIMIT 5

SQLLn 2, Col 330 Errores: 00 Advertencias: 0

Resultados de la consulta

Descargar resultados

Los resultados de las consultas no están disponibles después de elegir **Cerrar** o salir de la página. Elija **Download results** (Descargar resultados) para descargar una copia de los siguientes resultados de la consulta.

Estado

Se han devuelto correctamente 4 registros en 1337 ms

Bytes devueltos: 253 B

Sin procesar

Formateado

1

1,Alejandro,Rosalez,12/12/2013,123 Main St.,Any Town,MD,301-555-0158

2

2,Jane,Doe,10/5/2014,456 State St.,Anywhere,WA,360-555-0163

3

3,John,Stiles,9/20/2016,1980 8th St.,Nowhere,NY,914-555-0122

4

4,Li,Juan,6/29/2011,1323 22nd Ave.,Anytown,NY,914-555-0149

5

Estado
✔ Se han devuelto correctamente 4 registros en 1337 ms
Bytes devueltos: 253 B

Sin procesar | **Formateado**

< 1 >

1	Alejandro	Rosalez	12/12/2013	123 Main St.	Any Town	MD	301-555-0158
2	Jane	Doe	10/5/2014	456 State St.	Anywhere	WA	360-555-0163
3	John	Stiles	9/20/2016	1980 8th St.	Nowhere	NY	914-555-0122
4	Li	Juan	6/29/2011	1323 22nd Ave.	Anytown	NY	914-555-0149

1

/ Para crear un punto de referencia para escribir consultas SQL, ejecutando la siguiente consulta SQL: SELECT * FROM s3object*

2

SELECT "First Name" FROM s3object s LIMIT 3

SQL Ln 2, Col 44 Errores: 0 Advertencias: 0

Resultados de la consulta

Los resultados de las consultas no están disponibles después de elegir **Close** (Cerrar) o salir de la página resultados de la consulta.

Estado
✔ Se han devuelto correctamente 3 registros en 1973 ms
Bytes devueltos: 20 B

Sin procesar | Formateado

1

Alejandro

2

Jane

3

John

4

Sin procesar | **Formateado**

Alejandro

Jane

John

Tarea 4: Modificar las propiedades de cifrado y el tipo de almacenamiento de un objeto

Objetos (1) Información

Copiar URI de S3

Copiar URL

Descargar

Abrir

Eliminar

Acciones ▲

Crear carpeta

Descargar como

Compartir con una URL prefirmada

Calcular el tamaño total

Copiar

Trasladar

Iniciar restauración

Consultar con S3 Select

Editar acciones

Cambiar el nombre del objeto

Editar clase de almacenamiento

Editar cifrado del lado del servidor

Editar metadatos

Editar etiquetas

Hacer público mediante ACL

Los objetos son las entidades fundamentales que se almacenan en Amazon S3. Puede utilizar el [Inventario de Amazon S3](#) para obtener una lista de todos los objetos de su tienda que concederles permisos de forma explícita. [Más información](#)

Buscar objetos por prefijo

Nombre

▲

Tipo

▼

Última modificación

▼

Tamaño

lab1.csv

csv

5 May 2024 6:35:38 PM -05

Clase de almacenamiento

Amazon S3 ofrece una gama de tipos de almacenamiento diseñados para diferentes casos de uso. [Obtenga más información](#) o consulte los [precios de Amazon S3](#)

	Clase de almacenamiento	Diseñado para	Zonas de disponibilidad	Duración de almacenamiento mínima	Tiempo de recuperación
<input type="radio"/>	Estándar	Datos a los que se accede con frecuencia (más de una vez al mes) con acceso en milisegundos	≥ 3	-	-
<input checked="" type="radio"/>	Agrupación por niveles inteligente	Datos con patrones de acceso desconocidos o cambiantes	≥ 3	-	-
<input type="radio"/>	Estándar - Acceso poco frecuente	Datos a los que se accede con poca frecuencia (una vez al mes) con acceso en milisegundos	≥ 3	30 días	1
<input type="radio"/>	Única zona - Acceso poco frecuente	Datos recreables y a los que se accede con poca frecuencia (una vez al mes) almacenados en una única zona de	1	30 días	1

Se editó correctamente la clase de almacenamiento

Consulte los detalles a continuación.

Editar clase de almacenamiento: estado Información Cerrar

La información que aparece a continuación ya no estará disponible una vez que abandone la página.

Resumen

Origen

s3://ade-s3lab-bucket-b20d4860

Editado correctamente

1 objeto, 328.0 B

Error al editar

0 objetos

Error al editar (0)

Buscar objetos por nombre

Nombre

▲

Carpeta

▼

Tipo

▼

Última modificación

▼

Tamaño

▼

Error

▼

No hay ediciones con error.

Tarea 5: comprimir y consultar el conjunto de datos

```
voclabs:~/environment $ zip lab lab1.csv
adding: lab1.csv (deflated 29%)
voclabs:~/environment $ ll
total 20
-rw-rw-r-- 1 ec2-user ec2-user 751 Sep  3  2022 code.zip
-rw-r--r-- 1 ec2-user ec2-user 836 May  5 23:04 create_bucket.yml
-rwxr-xr-x 1 ec2-user ec2-user 328 Sep  2  2022 lab1.csv
-rw-rw-r-- 1 ec2-user ec2-user 398 May  5 23:59 lab.zip
-rw-r--r-- 1 ec2-user ec2-user 569 Apr 30 02:48 README.md
voclabs:~/environment $
```

```
voclabs:~/environment $ gzip -v lab1.csv
lab1.csv:      29.3% -- replaced with lab1.csv.gz
voclabs:~/environment $ ls -la
total 24
drwxr-xr-x  3 ec2-user ec2-user  109 May  6 00:00 .
drwx----- 13 ec2-user ec2-user 4096 May  5 23:19 ..
drwxr-xr-x  5 ec2-user ec2-user  189 May  5 23:04 .c9
-rw-rw-r--  1 ec2-user ec2-user  751 Sep  3  2022 code.zip
-rw-r--r--  1 ec2-user ec2-user  836 May  5 23:04 create_bucket.yml
-rwxr-xr-x  1 ec2-user ec2-user  259 Sep  2  2022 lab1.csv.gz
-rw-rw-r--  1 ec2-user ec2-user  398 May  5 23:59 lab.zip
-rw-r--r--  1 ec2-user ec2-user  569 Apr 30 02:48 README.md
voclabs:~/environment $
```

```
voclabs:~/environment $ aws s3 cp lab1.csv.gz s3://ade-s3lab-bucket--b20d4860 --cache-control max-age=60
upload: ./lab1.csv.gz to s3://ade-s3lab-bucket--b20d4860/lab1.csv.gz
voclabs:~/environment $
```

ade-s3lab-bucket--b20d4860 Información

Objetos

Propiedades

Permisos

Métricas

Administración

Puntos de acceso

Objetos (2) Información

Copiar URI de S3

Copiar URL

Descargar

Abrir

Eliminar

Acciones

Crear carpeta

Cargar

Los objetos son las entidades fundamentales que se almacenan en Amazon S3. Puede utilizar el [Inventario de Amazon S3](#) para obtener una lista de todos los objetos de su bucket. Para que otras personas obtengan acceso a sus objetos, tendrá que concederles permisos de forma explícita. [Más información](#)

< 1 >

<input type="checkbox"/>	Nombre	Tipo	Última modificación	Tamaño	Clase de almacenamiento
<input type="checkbox"/>	lab1.csv	csv	5 May 2024 6:57:33 PM -05	328.0 B	Agrupación por niveles inteligente
<input type="checkbox"/>	lab1.csv.gz	gz	5 May 2024 7:04:24 PM -05	259.0 B	Estándar

Objetos (2) Información

Copiar URI de S3

Copiar URL

Descargar

Abrir

Eliminar

Acciones

Crear carpeta

Los objetos son las entidades fundamentales que se almacenan en Amazon S3. Puede utilizar el [Inventario de Amazon S3](#) para obtener una lista de todos los objetos de su bucket. Para que otras personas obtengan acceso a sus objetos, tendrá que concederles permisos de forma explícita. [Más información](#)

< 1 >

<input type="checkbox"/>	Nombre	Tipo	Última modificación	Tamaño
<input type="checkbox"/>	lab1.csv	csv	5 May 2024 6:57:33 PM -05	
<input checked="" type="checkbox"/>	lab1.csv.gz	gz	5 May 2024 7:04:24 PM -05	

Descargar como

Compartir con una URL prefirmada

Calcular el tamaño total

Copiar

Trasladar

Iniciar restauración

Consultar con S3 Select

Editar acciones

Cambiar el nombre del objeto

Editar clase de almacenamiento

Editar cifrado del lado del servidor

Editar metadatos

Configuración de entrada

Ruta
s3://ade-s3lab-bucket--b20d4860/lab1.csv.gz

Tamaño
259.0 B

Formato
☒ CSV
☐ JSON
☐ Apache Parquet

CSV delimitador
☒ Coma
☐ Tabulador
☐ Personalizado

☒ Excluir la primera línea de CSV datos
Habilite esta configuración si CSV contiene una fila de encabezado.

Compresión
☐ Ninguno
☒ GZIP
☐ BZIP2

```
1  /* Para crear un punto de referencia para escribir consultas SQL, puede
    SELECT * FROM s3object s LIMIT 5 */
2  SELECT * FROM s3object s limit 3
```

SQL Ln 2, Col 33 Errores: 0 Advertencias: 0

Resultados de la consulta

Los resultados de las consultas no están disponibles después de elegir **Close** (Cerrar) o salir de la página. Elija l

Estado
 Se han devuelto correctamente 3 registros en 852 ms
Bytes devueltos: 193 B

Sin procesar | Formateado

1	1,Alejandro,Rosalez,12/12/2013,123 Main St.,Any Town,MD,301-555-0158
2	2,Jane,Doe,10/5/2014,456 State St.,Anywhere,WA,360-555-0163
3	3,John,Stiles,9/20/2016,1980 8th St.,Nowhere,NY,914-555-0122
4	

✔ Se han devuelto correctamente 3 registros en 852 ms
Bytes devueltos: 193 B

Sin procesar | **Formateado**

								< 1 >	
1	Alejandro	Rosalez	12/12/2013	123 Main St.	Any Town	MD	301-555-0158		
2	Jane	Doe	10/5/2014	456 State St.	Anywhere	WA	360-555-0163		
3	John	Stiles	9/20/2016	1980 8th St.	Nowhere	NY	914-555-0122		

Tarea 6: Administrar y probar el acceso restringido para un miembro del equipo

Visitados recientemente

Favoritos

Todos los servicios

Administración financiera en la nube

Administración y gobierno

Almacenamiento

Análisis

Aplicaciones empresariales

Visitados recientemente

☆ IAM
Administrar el acceso a los recursos de AWS

S3
Almacenamiento escalable en la nube

Cloud9
Un IDE en la nube para escribir, ejecutar y depurar código

Página de inicio de la consola
Ver información sobre recursos, accesos directos a servicios y actualizaciones de características

Identity and Access Management (IAM)

Buscar en IAM

Panel

Administración del acceso

Grupos de usuarios

Usuarios

Roles

Políticas

Proveedores de identidad

Configuración de cuenta

Informes de acceso

Access Analyzer

Acceso externo

Acceso no utilizado

Configuración del analizador

Informe de credenciales

IAM > Panel

Panel de IAM

Recursos de IAM
Recursos de esta cuenta de AWS

Grupos de usuarios	Usuarios	Roles	Políticas
1	1	15	2

Proveedores de identidad

0

Novedades [Ver todo](#)
Actualizaciones de características en IAM

- Presentamos las API de AWS IAM Identity Center para visualizar el acceso de la fuerza laboral a AWS.
Hace 6 meses
- Nuevas claves de condición de IAM para toda la

IAM > Usuarios

Usuarios (1) Información

Eliminar

Crear usuario

Un usuario de IAM es una identidad con credenciales válidas a largo plazo que se utiliza para interactuar con AWS en una cuenta.

Buscar

< 1 > ⚙

☐

Nombre de usuario

▲

Ruta

▼

Grupo:

▼

Última actividad

▼

MFA

▼

Antigüedad de

▼

Último inicio de se

▼

☐

paulo

/

1

-

-

🟢 1 hora

-

paulo Información

Eliminar

Resumen

ARN
arn:aws:iam::050670574482:user/paulo

Creado
May 05, 2024, 17:56 (UTC-05:00)

Acceso a la consola
⚠️ Habilitado sin MFA

Último inicio de sesión en la consola
🕒 Nunca

Clave de acceso 1
AKIAQXTBTTOJDHHKUXVI - Active
🕒 Nunca usado. Creado hoy.

Clave de acceso 2
[Crear clave de acceso](#)

Permisos

Grupos (1)

Etiquetas (1)

Credenciales de seguridad

Access Advisor

Políticas de permisos (1)

Eliminar

Agregar permisos ▼

Los permisos se definen mediante políticas asociadas al usuario directamente o a través de grupos.

Buscar

Filtrar por Tipo
Todos los tipos ▼

< 1 > ⚙

☐

Nombre de la política

▲

Tipo

▼

Adjuntado a través de

🔗

☐

[Policy-For-Data-Scien...](#)

Administrada ...

Grupo [DataScienceGroup](#)

paulo Información

Eliminar

Resumen

ARN
arn:aws:iam::050670574482:user/paulo

Creado
May 05, 2024, 17:56 (UTC-05:00)

Acceso a la consola
⚠️ Habilitado sin MFA

Último inicio de sesión en la consola
🕒 Nunca

Clave de acceso 1
AKIAQXTBTTOJDHHKUXVI - Active
🕒 Nunca usado. Creado hoy.

Clave de acceso 2
[Crear clave de acceso](#)

Permisos

Grupos (1)

Etiquetas (1)

Credenciales de seguridad

Access Advisor

Pertenencia a grupos de usuarios (1)

Eliminar

Agregar el usuario a los grupos

Un grupo de usuarios es un conjunto de usuarios de IAM. Los grupos se utilizan para especificar permisos que se aplican a un conjunto de usuarios. Un usuario puede ser miembro de hasta 10 grupos a la vez.

☐

Nombre del grupo

▲

Políticas adjuntas

🔗

☐

[DataScienceGroup](#)

[Policy-For-Data-Scientists](#)

Nombre de la política

Tipo

Adjuntado a través de

Policy-For-Data-Scien...

Administrada ...

Grupo DataScienceGroup

Policy-For-Data-Scientists

Copiar JSON

1-

{

2-

"Version": "2012-10-17",

3-

"Statement": [

4-

{

5-

"Action": [

6-

"s3:ListAllMyBuckets"

7-

],

8-

"Resource": "arn:aws:s3:::",

9-

"Effect": "Allow"

10-

},

11-

{

12-

"Action": [

13-

"s3:*"

14-

],

15-

"Resource": [

16-

"arn:aws:s3:::ade-s3lab-bucket--b20d4860",

17-

"arn:aws:s3:::ade-s3lab-bucket--b20d4860/*"

18-

],

19-

"Effect": "Allow"

20-

},

CloudFormation

Pilas

StackSets

Exportaciones

Application Composer

Generador de laC

Registro

CloudFormation

Pilas

Pilas (2)

Eliminar

Actualizar

Acciones de pila

Crear pila

Filtrar por nombre de pila

Filtrar estado

Activo

Vista anidada

< 1 >

Nombre de la pila	Estado	Hora de creación	Descripción
aws-cloud9-Cloud9-Instance-77f0847c0e144a68b1b3e86af7b78a91	CREATE_COMPLETE	2024-05-05 17:56:48 UTC-0500	-
c116055a2759811f6616091t1w050670574482	CREATE_COMPLETE	2024-05-05 17:56:23 UTC-0500	ADE_1.0

Pilas (2)

Filtrar estado

Filtrar por nombre de pila

Activo

Vista anidada

< 1 >

Pilas

aws-cloud9-Cloud9-Instance-77f0847c0e144a68b1b3e86af7b78a91	2024-05-05 17:56:48 UTC-0500	CREATE_COMPLETE
c116055a2759811f6616091t1w050670574482	2024-05-05 17:56:23 UTC-0500	CREATE_COMPLETE

Eliminar

Actualizar

Acciones de pila

Crear pila

< Pila

Eventos

Recursos

Salidas

>

Salidas (3)

Buscar resultados

< 1 >

Clave	Valor
BucketName	ade-s3lab-bucket--b20d4860
PaulosAccessKey	AKIAQXTBTT0JDHHKUXVI
PaulosSecretAccessKey	+5ScGprup6Zd4DX6mxaDEw6x76a5Uhp8NJt+uSfW

```
voclabs:~/environment $ AK=AKIAQXTBTT0JDHHKUXVI
voclabs:~/environment $ SAK=+5ScGprup6Zd4DX6mxaDEw6x76a5Uhp8NJt+uSfW
voclabs:~/environment $
```

```
voclabs:~/environment $ AWS_ACCESS_KEY_ID=$AK AWS_SECRET_ACCESS_KEY=$SAK aws e
c2 describe-instances --region us-east-1

An error occurred (UnauthorizedOperation) when calling the DescribeInstances o
peration: You are not authorized to perform this operation. User: arn:aws:iam:
:050670574482:user/paulo is not authorized to perform: ec2:DescribeInstances b
ecause no identity-based policy allows the ec2:DescribeInstances action
voclabs:~/environment $
```



```
voclabs:~/environment $ cat pulled_lab.csv
CustomerID,First Name,Last Name,Join Date,Street Address,City,State,Phone
1,Alejandro,Rosalez,12/12/2013,123 Main St.,Any Town,MD,301-555-0158
2,Jane,Doe,10/5/2014,456 State St.,Anywhere,WA,360-555-0163
3,John,Stiles,9/20/2016,1980 8th St.,Nowhere,NY,914-555-0122
4,Li,Juan,6/29/2011,1323 22nd Ave.,Anytown,NY,914-555-0149
voclabs:~/environment $
```

AWS

00:22

▶ Start Lab

■ End Lab

i AWS Details

i Details

✕

Submit

Submission Report

Grades

es
sions

1. To record your progress, choose **Submit** at the top of these instructions.

2. When prompted, choose **Yes**.

After a couple of minutes, the grades panel appears and shows you how many points you earned for each task. If the results don't display after a couple of minutes, choose **Grades** at the top of these instructions.

⚠ Important: Some of the checks made by the submission process in this lab will only give you credit if it has been at least 5 minutes since you completed the action. If you do not receive credit the first time you submit, you may need to wait a couple minutes and the submit again to receive credit for these items.

💡 Tip: You can submit your work multiple times. After you change your work, choose **Submit** again. Your last submission is recorded for this lab.

3. To find detailed feedback about your work, choose **Submission Report**.

Total score

35/35

[Task 1a] CloudFormation stack was crea

[Task 1b] S3 bucket was created

[Task 1c] CloudFormation stack was dele

[Task 1d] S3 bucket was deleted

[Task 2] CSV file copied to bucket

[Task 4] S3 encrypted object found

[Task 5] GZIP file copied to bucket

Module 2

Module 2 Knowledge Check

Due

No Due Date

Points

100

Submitting

an external tool

Knowledge check results

Your score:

100%

(100 points)

Required score:

70%

(70 points)

Result:

Congratulations! You have completed this module.

To continue, choose **Next** in the lower-right corner.

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Module 3

Module 3 Knowledge Check

Due

No Due Date

Points

100

Submitting

an external tool

Knowledge check results

Your score:

100%

(100 points)

Required score:

70%

(70 points)

Result:

Congratulations! You have completed this module.

To continue, choose **Next** in the lower-right corner.

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.