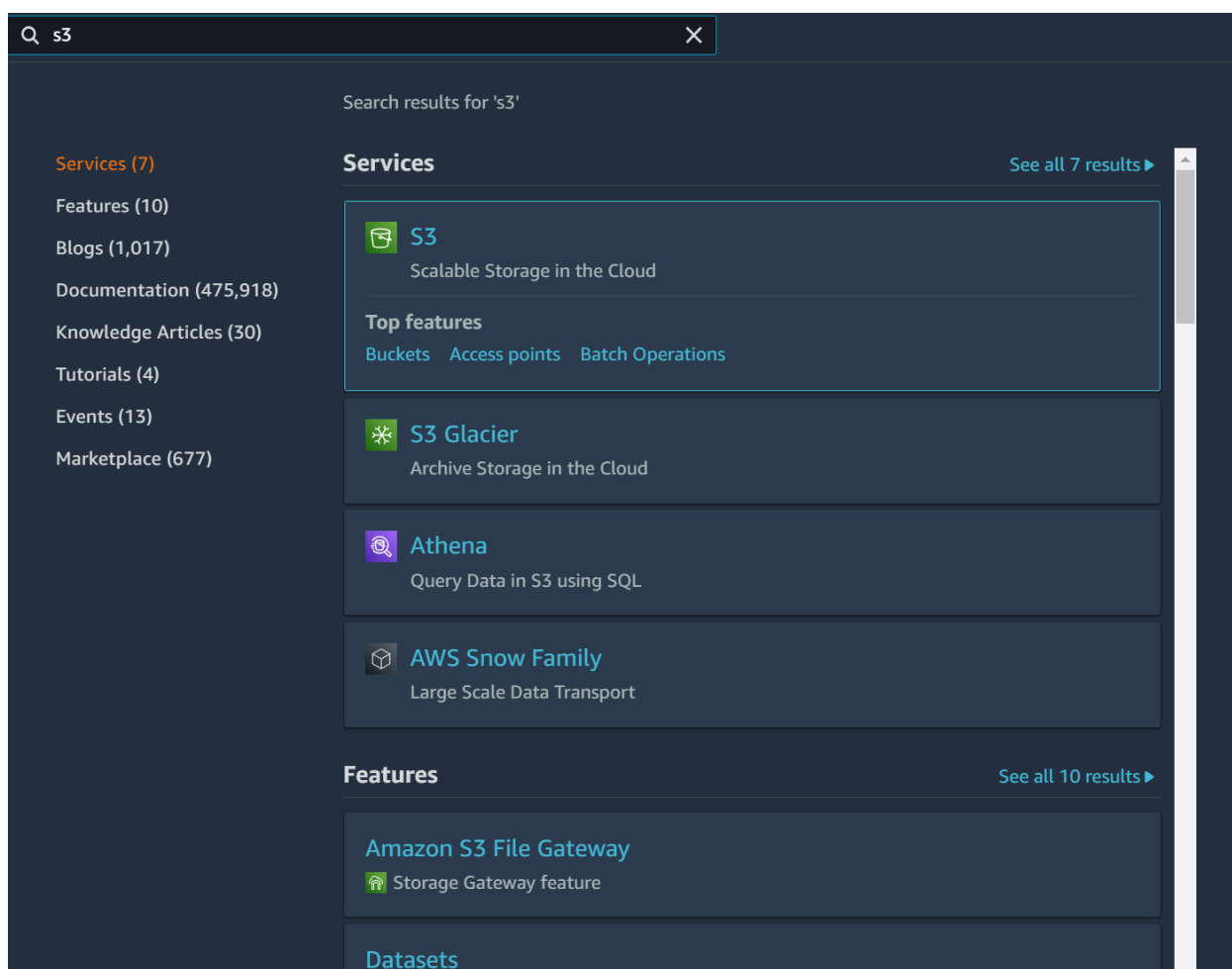


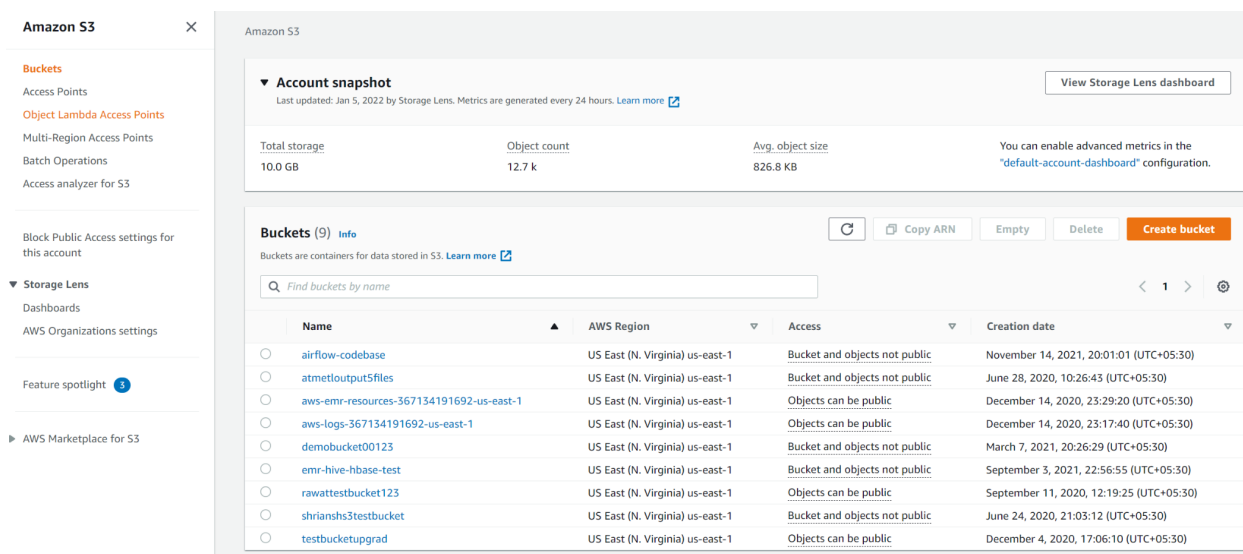
Creating an S3 Bucket

You will use Amazon S3 to read and write data while running your queries during the course. For this purpose, you will need to create buckets on Amazon S3. The detailed steps for the same are given below.

1. Choose the S3 option from your Amazon dashboard.



2. You will then be redirected to the S3 management console.



Amazon S3

Buckets

Access Points

Object Lambda Access Points

Multi-Region Access Points

Batch Operations

Access analyzer for S3

Block Public Access settings for this account

Storage Lens

Dashboards

AWS Organizations settings

Feature spotlight

AWS Marketplace for S3

Account snapshot

Last updated: Jan 5, 2022 by Storage Lens. Metrics are generated every 24 hours. [Learn more](#)

Total storage: 10.0 GB

Object count: 12.7 k

Avg. object size: 826.8 KB

You can enable advanced metrics in the "default-account-dashboard" configuration.

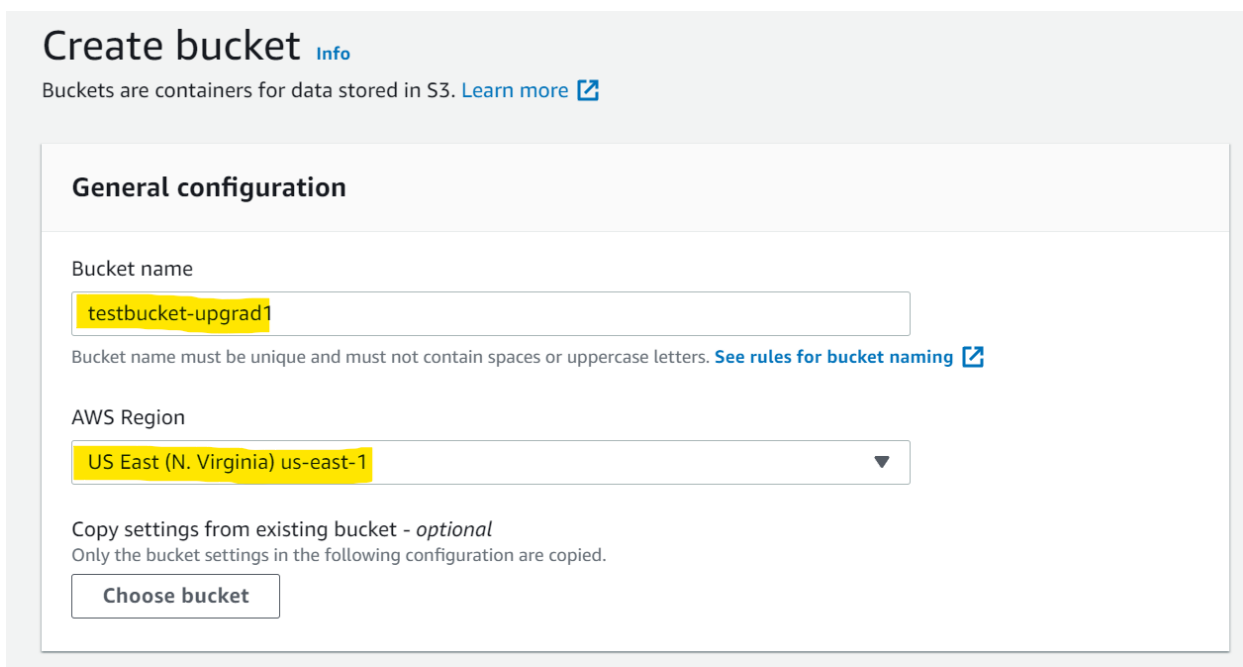
Buckets (9) [Info](#)

Buckets are containers for data stored in S3. [Learn more](#)

[Copy ARN](#) [Empty](#) [Delete](#) [Create bucket](#)

Name	AWS Region	Access	Creation date
airflow-codebase	US East (N. Virginia) us-east-1	Bucket and objects not public	November 14, 2021, 20:01:01 (UTC+05:30)
atmetioutput5files	US East (N. Virginia) us-east-1	Bucket and objects not public	June 28, 2020, 10:26:43 (UTC+05:30)
aws-emr-resources-367134191692-us-east-1	US East (N. Virginia) us-east-1	Objects can be public	December 14, 2020, 23:29:20 (UTC+05:30)
aws-logs-367134191692-us-east-1	US East (N. Virginia) us-east-1	Objects can be public	December 14, 2020, 23:17:40 (UTC+05:30)
demobucket00123	US East (N. Virginia) us-east-1	Bucket and objects not public	March 7, 2021, 20:26:29 (UTC+05:30)
emr-hive-hbase-test	US East (N. Virginia) us-east-1	Bucket and objects not public	September 3, 2021, 22:56:55 (UTC+05:30)
rawattestbucket123	US East (N. Virginia) us-east-1	Objects can be public	September 11, 2020, 12:19:25 (UTC+05:30)
shriansh3testbucket	US East (N. Virginia) us-east-1	Bucket and objects not public	June 24, 2020, 21:03:12 (UTC+05:30)
testbucketupgrad	US East (N. Virginia) us-east-1	Objects can be public	December 4, 2020, 17:06:10 (UTC+05:30)

3. Click on the **Create Bucket** button to get to the following screen.



Create bucket [Info](#)

Buckets are containers for data stored in S3. [Learn more](#)

General configuration

Bucket name

Bucket name must be unique and must not contain spaces or uppercase letters. [See rules for bucket naming](#)

AWS Region

Copy settings from existing bucket - *optional*

Only the bucket settings in the following configuration are copied.

[Choose bucket](#)

To create a new bucket, you will need to enter the following details under **General configuration**:

- **Bucket Name:** Enter the bucket's name that you want to make. **The bucket names are unique across the entire S3 system, and the region should be N. Virginia.**
- **AWS Region:** Please ensure you **select the same region as your instance. If the bucket is in a separate region from the instance, you will not be able to read/write the data to the bucket.** For example, if your instance is in US East (N. Virginia), select the same region for your S3 bucket.

You can skip the third option, “Copy settings from an existing bucket”. After this, scroll down to the Bucket settings for **Object Ownership**.

4. Under the **Object Ownership** configuration, you need to select and choose the **ACLs enabled** option. Without ACL enabled, you will not be able to make the files in your bucket **public**. After this, scroll down to the **Block Public Access settings for this bucket** configuration.

Object Ownership [Info](#)

Control ownership of objects written to this bucket from other AWS accounts and granted using access control lists (ACLs). Object ownership determines who can specify access to objects.

☐ **ACLs disabled (recommended)**
 All objects in this bucket are owned by this account. Access to this bucket and its objects is specified using only policies.

☒ **ACLs enabled**
 Objects in this bucket can be owned by other AWS accounts. Access to this bucket and its objects can be specified using ACLs.

Object Ownership

☒ **Bucket owner preferred**
 If new objects written to this bucket specify the bucket-owner-full-control canned ACL, they are owned by the bucket owner. Otherwise, they are owned by the object writer.

☐ **Object writer**
 The object writer remains the object owner.

i If you want to enforce object ownership for new objects only, your bucket policy must specify that the bucket-owner-full-control canned ACL is required for object uploads. [Learn more](#) [↗](#)

5. In this configuration, you need to **untick** the **Block all public access** setting and then **tick** on the acknowledgement for turning off block all public access as shown in the image below. You can then click on the “**Create Bucket**” button at the end of this page.

Block Public Access settings for this bucket

Public access is granted to buckets and objects through access control lists (ACLs), bucket policies, access point policies, or all. In order to ensure that public access to this bucket and its objects is blocked, turn on Block all public access. These settings apply only to this bucket and its access points. AWS recommends that you turn on Block all public access, but before applying any of these settings, ensure that your applications will work correctly without public access. If you require some level of public access to this bucket or objects within, you can customize the individual settings below to suit your specific storage use cases. [Learn more](#)

☐ Block all public access

Turning this setting on is the same as turning on all four settings below. Each of the following settings are independent of one another.

☐ Block public access to buckets and objects granted through *new* access control lists (ACLs)

S3 will block public access permissions applied to newly added buckets or objects, and prevent the creation of new public access ACLs for existing buckets and objects. This setting doesn't change any existing permissions that allow public access to S3 resources using ACLs.

☐ Block public access to buckets and objects granted through *any* access control lists (ACLs)

S3 will ignore all ACLs that grant public access to buckets and objects.

☐ Block public access to buckets and objects granted through *new* public bucket or access point policies

S3 will block new bucket and access point policies that grant public access to buckets and objects. This setting doesn't change any existing policies that allow public access to S3 resources.

☐ Block public and cross-account access to buckets and objects through *any* public bucket or access point policies

S3 will ignore public and cross-account access for buckets or access points with policies that grant public access to buckets and objects.



Turning off block all public access might result in this bucket and the objects within becoming public

AWS recommends that you turn on block all public access, unless public access is required for specific and verified use cases such as static website hosting.

☒ I acknowledge that the current settings might result in this bucket and the objects within becoming public.

6. You will now be able to see your bucket on the S3 management screen.

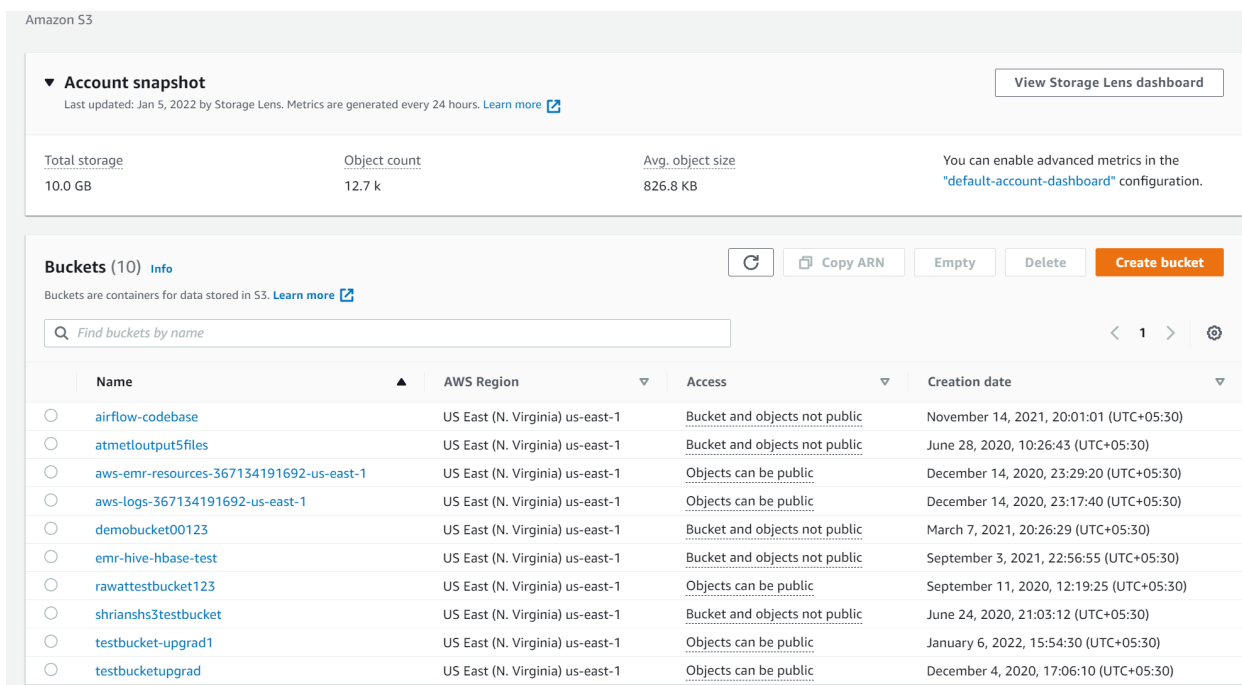
<input type="radio"/>	airflow-codebase	US East (N. Virginia) us-east-1	Bucket and objects not public	November 14, 2021, 20:01:01 (UTC+05:30)
<input type="radio"/>	atmetloutput5files	US East (N. Virginia) us-east-1	Bucket and objects not public	June 28, 2020, 10:26:43 (UTC+05:30)
<input type="radio"/>	aws-emr-resources-367134191692-us-east-1	US East (N. Virginia) us-east-1	Objects can be public	December 14, 2020, 23:29:20 (UTC+05:30)
<input type="radio"/>	aws-logs-367134191692-us-east-1	US East (N. Virginia) us-east-1	Objects can be public	December 14, 2020, 23:17:40 (UTC+05:30)
<input type="radio"/>	demobucket00123	US East (N. Virginia) us-east-1	Bucket and objects not public	March 7, 2021, 20:26:29 (UTC+05:30)
<input type="radio"/>	emr-hive-hbase-test	US East (N. Virginia) us-east-1	Bucket and objects not public	September 3, 2021, 22:56:55 (UTC+05:30)
<input type="radio"/>	rawattestbucket123	US East (N. Virginia) us-east-1	Objects can be public	September 11, 2020, 12:19:25 (UTC+05:30)
<input type="radio"/>	shrianshs3testbucket	US East (N. Virginia) us-east-1	Bucket and objects not public	June 24, 2020, 21:03:12 (UTC+05:30)
<input type="radio"/>	testbucket-upgrad1	US East (N. Virginia) us-east-1	Objects can be public	January 6, 2022, 15:54:30 (UTC+05:30)
<input type="radio"/>	testbucketupgrad	US East (N. Virginia) us-east-1	Objects can be public	December 4, 2020, 17:06:10 (UTC+05:30)

Note: Please delete your buckets once you have gone through the module. If you want to practice the concepts, you can create the buckets again later on.

Uploading files to S3

Once you have created a bucket on S3, you will also need to upload files to your newly created S3 bucket during this program. You can easily accomplish this task by following these steps.

1. Navigate to the S3 management console as you did in the steps above.



Amazon S3

▼ Account snapshot View Storage Lens dashboard

Last updated: Jan 5, 2022 by Storage Lens. Metrics are generated every 24 hours. [Learn more](#)

Total storage	Object count	Avg. object size	You can enable advanced metrics in the "default-account-dashboard" configuration.
10.0 GB	12.7 k	826.8 KB	

Buckets (10) [Info](#) Refresh Copy ARN Empty Delete Create bucket

Buckets are containers for data stored in S3. [Learn more](#)

Find buckets by name

Name	AWS Region	Access	Creation date
airflow-codebase	US East (N. Virginia) us-east-1	Bucket and objects not public	November 14, 2021, 20:01:01 (UTC+05:30)
atmetloutput5files	US East (N. Virginia) us-east-1	Bucket and objects not public	June 28, 2020, 10:26:43 (UTC+05:30)
aws-emr-resources-367134191692-us-east-1	US East (N. Virginia) us-east-1	Objects can be public	December 14, 2020, 23:29:20 (UTC+05:30)
aws-logs-367134191692-us-east-1	US East (N. Virginia) us-east-1	Objects can be public	December 14, 2020, 23:17:40 (UTC+05:30)
demobucket00123	US East (N. Virginia) us-east-1	Bucket and objects not public	March 7, 2021, 20:26:29 (UTC+05:30)
emr-hive-hbase-test	US East (N. Virginia) us-east-1	Bucket and objects not public	September 3, 2021, 22:56:55 (UTC+05:30)
rawattestbucket123	US East (N. Virginia) us-east-1	Objects can be public	September 11, 2020, 12:19:25 (UTC+05:30)
shrianshs3testbucket	US East (N. Virginia) us-east-1	Bucket and objects not public	June 24, 2020, 21:03:12 (UTC+05:30)
testbucket-upgrad1	US East (N. Virginia) us-east-1	Objects can be public	January 6, 2022, 15:54:30 (UTC+05:30)
testbucketupgrad	US East (N. Virginia) us-east-1	Objects can be public	December 4, 2020, 17:06:10 (UTC+05:30)

2. Now click on the bucket that you want to upload your file to. For this demo, we will be uploading a file to the same S3 bucket that we created in the previous section of this document. You will be navigated to the following screen.

testbucket-upgrad1 [Info](#)

[Objects](#)
[Properties](#)
[Permissions](#)
[Metrics](#)
[Management](#)
[Access Points](#)

Objects (0)
 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](#)

	Name	Type	Last modified	Size	Storage class
No objects					
You don't have any objects in this bucket.					
<input type="button" value="Upload"/>					

3. Click on the **Upload** button. You will then be navigated to the following screen.

Upload [Info](#)

Add the files and folders you want to upload to S3. To upload a file larger than 160GB, use the AWS CLI, AWS SDK or Amazon S3 REST API. [Learn more](#)

Drag and drop files and folders you want to upload here, or choose **Add files**, or **Add folders**.

Files and folders (0)

All files and folders in this table will be uploaded.

	Name	Folder	Type	Size
No files or folders				
You have not chosen any files or folders to upload.				

Destination

Destination
[s3://testbucket-upgrad1](#)

- Click on the **Add files** button and select the file/dataset you want to upload. After adding the file, you can scroll down and click on the **Upload button**.

Drag and drop files and folders you want to upload here, or choose **Add files**, or **Add folders**.

Files and folders (1 Total, 444.1 KB)

Remove

Add files

Add folder

All files and folders in this table will be uploaded.

Find by name

< 1 >

<input type="checkbox"/>	Name	Folder	Type	Size
<input type="checkbox"/>	clickstream-dataset.csv	-	application/vnd.ms-excel	444.1 KB

Destination

Destination

[s3://testbucket-upgrad1](#)

► **Destination details**

Bucket settings that impact new objects stored in the specified destination.

► **Permissions**

Grant public access and access to other AWS accounts.


► **Properties**

Specify storage class, encryption settings, tags, and more.

Cancel


Upload

- After clicking on the upload button, you should see the following screen where after some time, once the upload is done, you will see the upload succeeded status at the top. You can then go ahead and click on the **Close** button. After this, you will be navigated to your S3 bucket, where you can see your file uploaded. Click on the file that you have just uploaded.

 Upload succeeded
View details below.

Upload: status

Close

 The information below will no longer be available after you navigate away from this page.

Summary

Destination

s3://testbucket-upgrad1

Succeeded

1 file, 444.1 KB (100.00%)

Failed

0 files, 0 B (0%)

Files and folders

Configuration

Files and folders (1 Total, 444.1 KB)

Find by name

Name

Folder

Type

Size

Status


Error

clickstream-dataset.csv

-

application/vnd.ms-excel

444.1 KB

 Succeeded

-

- After clicking on the file, you will be navigated to the following screen. Next, try clicking on the link highlighted in the image under the Object URL.

Amazon S3

>

testbucket-upgrad1

>

clickstream-dataset.csv

clickstream-dataset.csv

Info

Copy S3 URI

Download

Open

Object actions

Properties

Permissions

Versions

Object overview

Owner

upgraduser820

AWS Region

US East (N. Virginia) us-east-1

Last modified

January 6, 2022, 16:16:35 (UTC+05:30)

Size

444.1 KB

Type

csv

Key

clickstream-dataset.csv

S3 URI

s3://testbucket-upgrad1/clickstream-dataset.csv

Amazon Resource Name (ARN)

arn:aws:s3:::testbucket-upgrad1/clickstream-dataset.csv

Entity tag (Etag)

17b0b4b82a5fb7cdf0f47429ee0257f4

Object URL

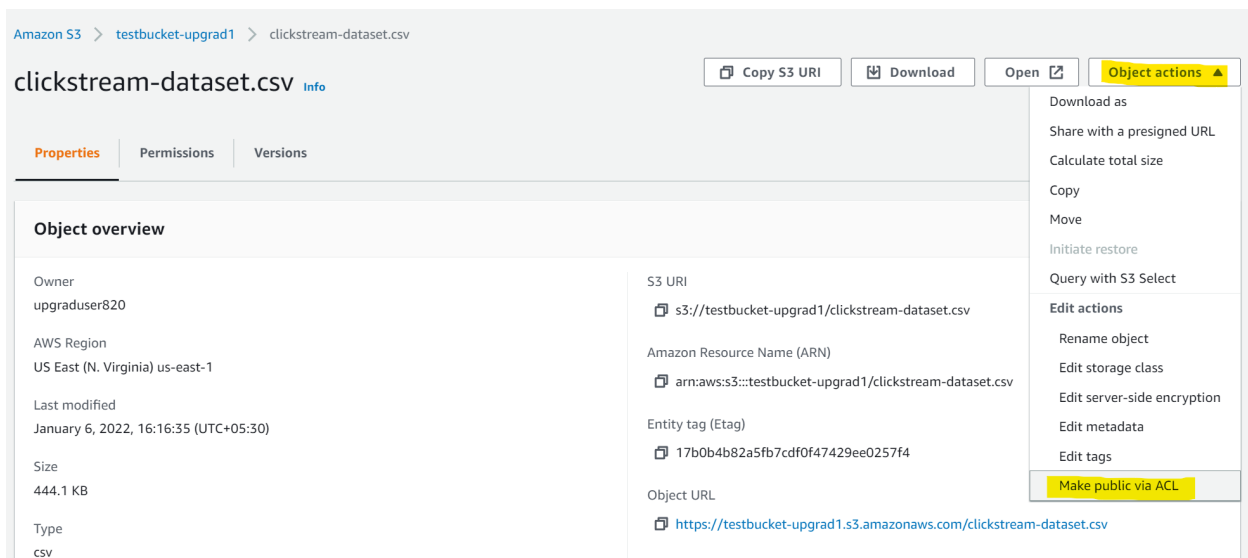
<https://testbucket-upgrad1.s3.amazonaws.com/clickstream-dataset.csv>

- You will be sent to the following page when you click on the URL, saying that access is denied. Now, go back to the S3 console by clicking on the back button on your web browser.

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?xml version="1.0" encoding="UTF-8" ?>
<Error>
  <Code>AccessDenied</Code>
  <Message>Access Denied</Message>
  <RequestId>R5879M1RQ8RE0PWS</RequestId>
  <HostId>OK14b1VnQMkKCX13ahvapCpIGehJXW2FFIdv97ZKjosTYgJ0RJ/+11Z0Mqk3U3so+7VAyJ6P4X0=</HostId>
</Error>
```

- Now, to solve the access issue, you need to make the object or file **public**. For this, you need to click on the **Object actions** button towards the top right of the screen and then in the drop-down menu, click on **Make public via ACL**.



Amazon S3 > testbucket-upgrad1 > clickstream-dataset.csv

clickstream-dataset.csv [Info](#)

[Copy S3 URI](#) [Download](#) [Open](#) [Object actions](#)

Properties | [Permissions](#) | [Versions](#)

Object overview

Owner upgraduser820 AWS Region US East (N. Virginia) us-east-1 Last modified January 6, 2022, 16:16:35 (UTC+05:30) Size 444.1 KB Type csv	S3 URI s3://testbucket-upgrad1/clickstream-dataset.csv Amazon Resource Name (ARN) arn:aws:s3:::testbucket-upgrad1/clickstream-dataset.csv Entity tag (Etag) 17b0b4b82a5fb7cdf0f47429ee0257f4 Object URL https://testbucket-upgrad1.s3.amazonaws.com/clickstream-dataset.csv
---	--


Object actions

- Download as
- Share with a presigned URL
- Calculate total size
- Copy
- Move
- Initiate restore
- Query with S3 Select
- Edit actions**
 - Rename object
 - Edit storage class
 - Edit server-side encryption
 - Edit metadata
 - Edit tags
 - Make public via ACL**

- On the next screen, simply click on the **Make public** button. You can then click on the **Close** button once the public access has been successfully edited.


Make public Info

The make public action enables public read access in the object access control list (ACL) settings. [Learn more](#)

 When public read access is enabled and not blocked by Block Public Access settings, anyone in the world can access the specified objects.

Specified objects

< 1 >

Name	Type	Last modified	Size
 clickstream-dataset.csv	csv	January 6, 2022, 16:16:35 (UTC+05:30)	444.1 KB

Cancel
Make public

- Once, again try to click on the **Object URL** for your file. This time, as soon as you click on the URL, a download will start for the file you uploaded on the S3 bucket signifying that the file has been provided public access.

Amazon S3

>

testbucket-upgrad1

>

clickstream-dataset.csv

clickstream-dataset.csv

Info

Copy S3 URI

Download

Properties

Permissions

Versions

Object overview

Owner

upgraduser820

AWS Region

US East (N. Virginia) us-east-1

Last modified

January 6, 2022, 16:16:35 (UTC+05:30)


Size

444.1 KB


Type

csv


Key

 clickstream-dataset.csv


S3 URI

 s3://testbucket-upgrad1/clickstream-dataset.csv


Amazon Resource Name (ARN)

 arn:aws:s3:::testbucket-upgrad1/clickstream-dataset.csv

Entity tag (Etag)

 17b0b4b82a5fb7cdf0f47429ee0257f4

Object URL

 <https://testbucket-upgrad1.s3.amazonaws.com/clickstream-dataset.csv>

© Copyright. upGrad Education Pvt. Ltd. All rights reserved