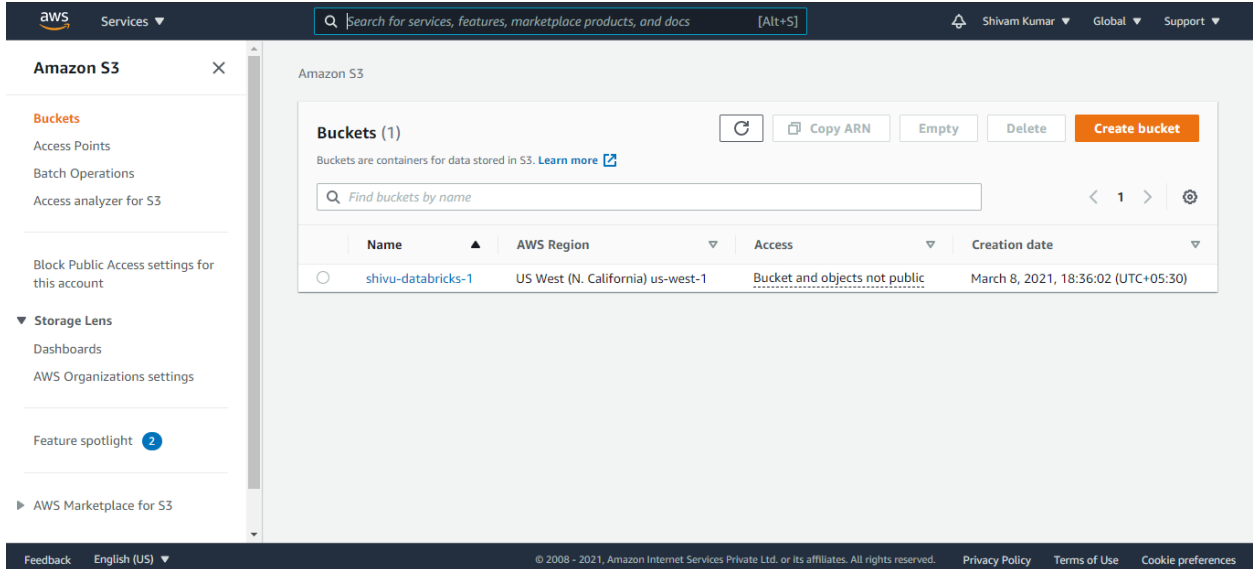


## AWS S3 Documentation:

### Step 1: Create a new Bucket - “shivu-databricks-1”

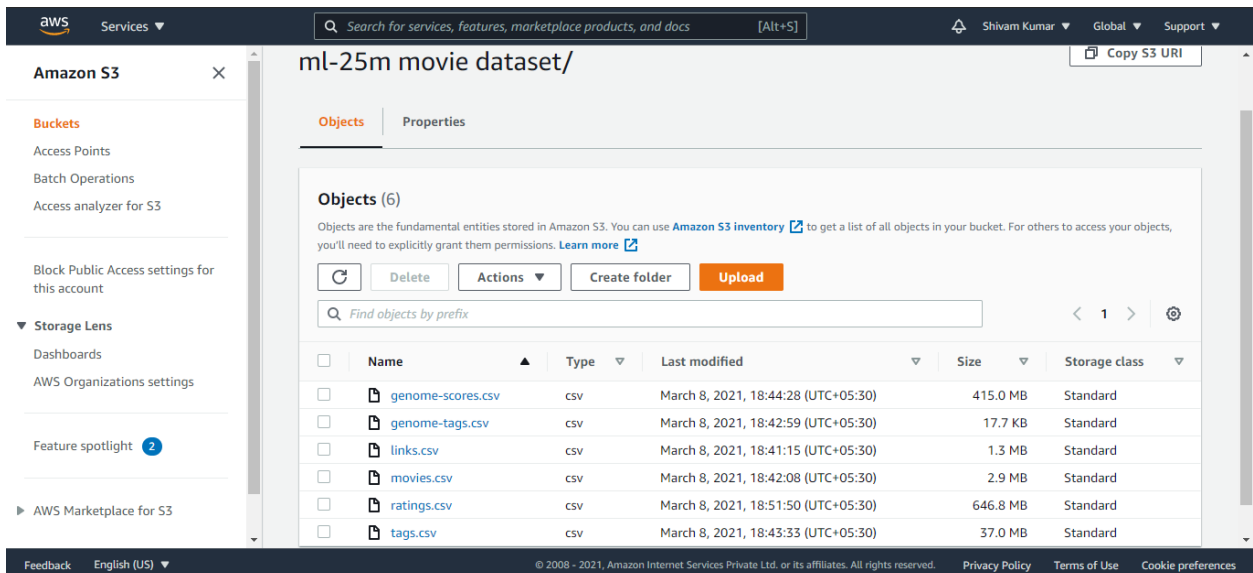


The screenshot shows the Amazon S3 console interface. On the left, the 'Amazon S3' sidebar is visible with options like Buckets, Access Points, Batch Operations, and Access analyzer for S3. The main content area displays 'Buckets (1)' with a table listing the bucket 'shivu-databricks-1'. The table columns are Name, AWS Region, Access, and Creation date. The bucket is located in the US West (N. California) us-west-1 region and has the access level 'Bucket and objects not public'. It was created on March 8, 2021, at 18:36:02 (UTC+05:30).

| Name               | AWS Region                        | Access                        | Creation date                       |
|--------------------|-----------------------------------|-------------------------------|-------------------------------------|
| shivu-databricks-1 | US West (N. California) us-west-1 | Bucket and objects not public | March 8, 2021, 18:36:02 (UTC+05:30) |

### Step 2: Create a Folder in Bucket and upload files as Objects.

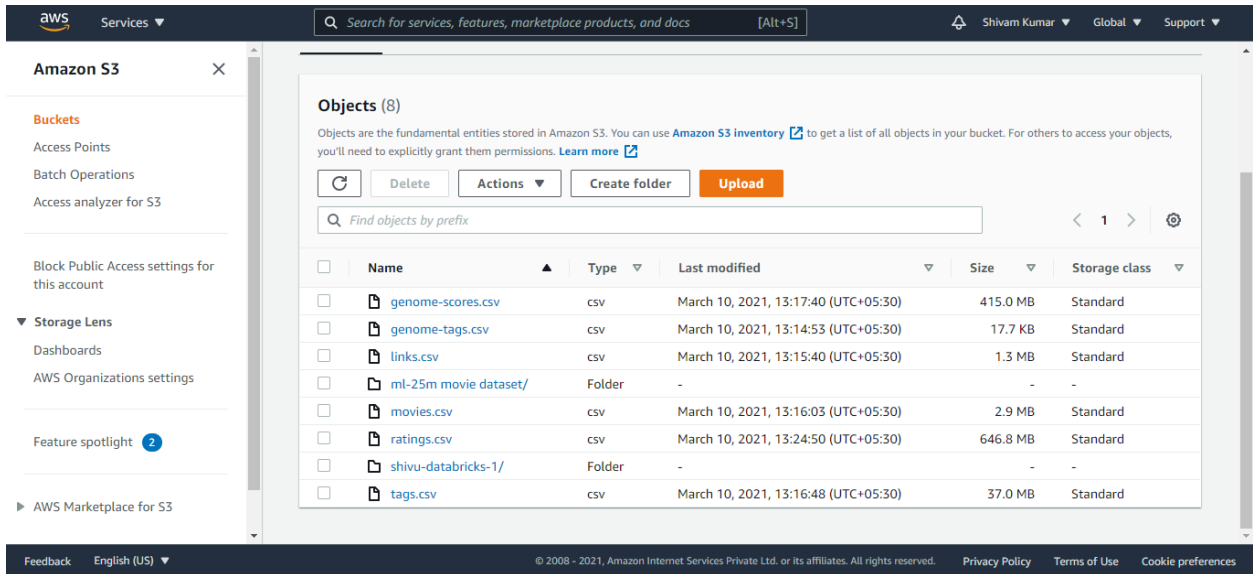
#### ml-25m movie dataset/



The screenshot shows the Amazon S3 console interface for the 'ml-25m movie dataset/' folder. The 'Objects (6)' page is displayed, showing a table of six CSV files. The table columns are Name, Type, Last modified, Size, and Storage class. The files are: genome-scores.csv (415.0 MB), genome-tags.csv (17.7 KB), links.csv (1.3 MB), movies.csv (2.9 MB), ratings.csv (646.8 MB), and tags.csv (37.0 MB). All files are stored in the Standard storage class.

| Name              | Type | Last modified                       | Size     | Storage class |
|-------------------|------|-------------------------------------|----------|---------------|
| genome-scores.csv | csv  | March 8, 2021, 18:44:28 (UTC+05:30) | 415.0 MB | Standard      |
| genome-tags.csv   | csv  | March 8, 2021, 18:42:59 (UTC+05:30) | 17.7 KB  | Standard      |
| links.csv         | csv  | March 8, 2021, 18:41:15 (UTC+05:30) | 1.3 MB   | Standard      |
| movies.csv        | csv  | March 8, 2021, 18:42:08 (UTC+05:30) | 2.9 MB   | Standard      |
| ratings.csv       | csv  | March 8, 2021, 18:51:50 (UTC+05:30) | 646.8 MB | Standard      |
| tags.csv          | csv  | March 8, 2021, 18:43:33 (UTC+05:30) | 37.0 MB  | Standard      |

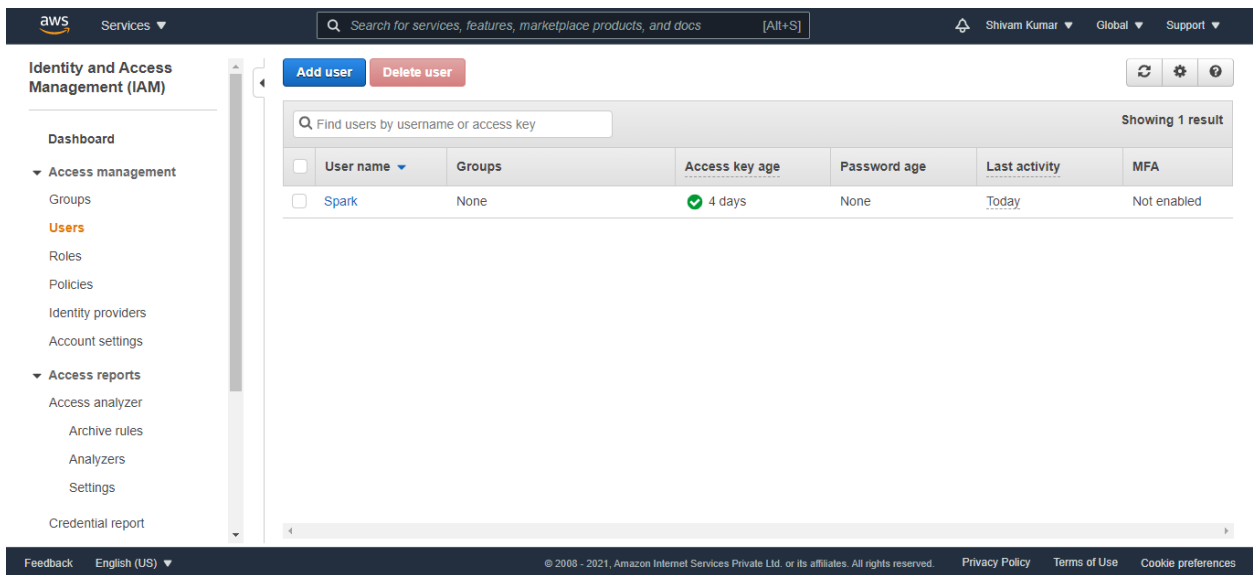
### Step 3: Upload some more CSV Files in same Bucket with public read access.



The screenshot shows the Amazon S3 console interface. On the left, the 'Amazon S3' sidebar is visible with options like Buckets, Access Points, Batch Operations, and Access analyzer for S3. The main content area displays 'Objects (8)' in a bucket. A search bar is present with the text 'Find objects by prefix'. Below the search bar, there is a table listing the objects. The table has columns for Name, Type, Last modified, Size, and Storage class. The objects listed are genome-scores.csv, genome-tags.csv, links.csv, ml-25m movie dataset/, movies.csv, ratings.csv, shivu-databricks-1/, and tags.csv. The 'Upload' button is highlighted in orange.

| <input type="checkbox"/> | Name                  | Type   | Last modified                        | Size     | Storage class |
|--------------------------|-----------------------|--------|--------------------------------------|----------|---------------|
| <input type="checkbox"/> | genome-scores.csv     | csv    | March 10, 2021, 13:17:40 (UTC+05:30) | 415.0 MB | Standard      |
| <input type="checkbox"/> | genome-tags.csv       | csv    | March 10, 2021, 13:14:53 (UTC+05:30) | 17.7 KB  | Standard      |
| <input type="checkbox"/> | links.csv             | csv    | March 10, 2021, 13:15:40 (UTC+05:30) | 1.3 MB   | Standard      |
| <input type="checkbox"/> | ml-25m movie dataset/ | Folder | -                                    | -        | -             |
| <input type="checkbox"/> | movies.csv            | csv    | March 10, 2021, 13:16:03 (UTC+05:30) | 2.9 MB   | Standard      |
| <input type="checkbox"/> | ratings.csv           | csv    | March 10, 2021, 13:24:50 (UTC+05:30) | 646.8 MB | Standard      |
| <input type="checkbox"/> | shivu-databricks-1/   | Folder | -                                    | -        | -             |
| <input type="checkbox"/> | tags.csv              | csv    | March 10, 2021, 13:16:48 (UTC+05:30) | 37.0 MB  | Standard      |

### Step4: Create a new user in Identity & Access Management IAM – “Spark” and provide full read access for AWS S3.



The screenshot shows the AWS Identity and Access Management (IAM) console. On the left, the 'Identity and Access Management (IAM)' sidebar is visible with options like Dashboard, Access management, Groups, Users, Roles, Policies, Identity providers, Account settings, Access reports, Access analyzer, Archive rules, Analyzers, Settings, and Credential report. The main content area displays a table of users. The table has columns for User name, Groups, Access key age, Password age, Last activity, and MFA. A single user named 'Spark' is listed with a green checkmark for the Access key age (4 days) and 'Not enabled' for MFA. The 'Add user' button is highlighted in blue.

| <input type="checkbox"/> | User name | Groups | Access key age | Password age | Last activity | MFA         |
|--------------------------|-----------|--------|----------------|--------------|---------------|-------------|
| <input type="checkbox"/> | Spark     | None   | ✓ 4 days       | None         | Today         | Not enabled |

aws Services Search for services, features, marketplace products, and docs [Alt+S] Shivam Kumar Global Support

Identity and Access Management (IAM)

Dashboard

Access management

Groups

Users

Roles

Policies

Identity providers

Account settings

Access reports

Access analyzer

Archive rules

Analyzers

Settings

Credential report

Users > Spark

### Summary

Delete user ?

User ARN: `arn:aws:iam::371801530033:user/Spark`

Path: `/`

Creation time: 2021-03-10 12:47 UTC+0530

Permissions Groups Tags Security credentials Access Advisor

Permissions policies (1 policy applied)

Add permissions Add inline policy

| Policy name        | Policy type        |
|--------------------|--------------------|
| Attached directly  |                    |
| AmazonS3FullAccess | AWS managed policy |

Permissions boundary (not set)

Feedback English (US) © 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences

## Step 5: Using access key and secret key for mounting S3 to Databricks Platform.

aws Services Search for services, features, marketplace products, and docs [Alt+S] Shivam Kumar Global Support

Identity and Access Management (IAM)

Dashboard

Access management

Groups

Users

Roles

Policies

Identity providers

Account settings

Access reports

Access analyzer

Archive rules

Analyzers

Settings

Credential report

Permissions Groups Tags Security credentials Access Advisor

### Sign-in credentials

Summary

- User does not have console management access

Console password: Disabled | Manage

Assigned MFA device: Not assigned | Manage

Signing certificates: None

### Access keys

Use access keys to make programmatic calls to AWS from the AWS CLI, Tools for PowerShell, AWS SDKs, or direct AWS API calls. You can have a maximum of two access keys (active or inactive) at a time.

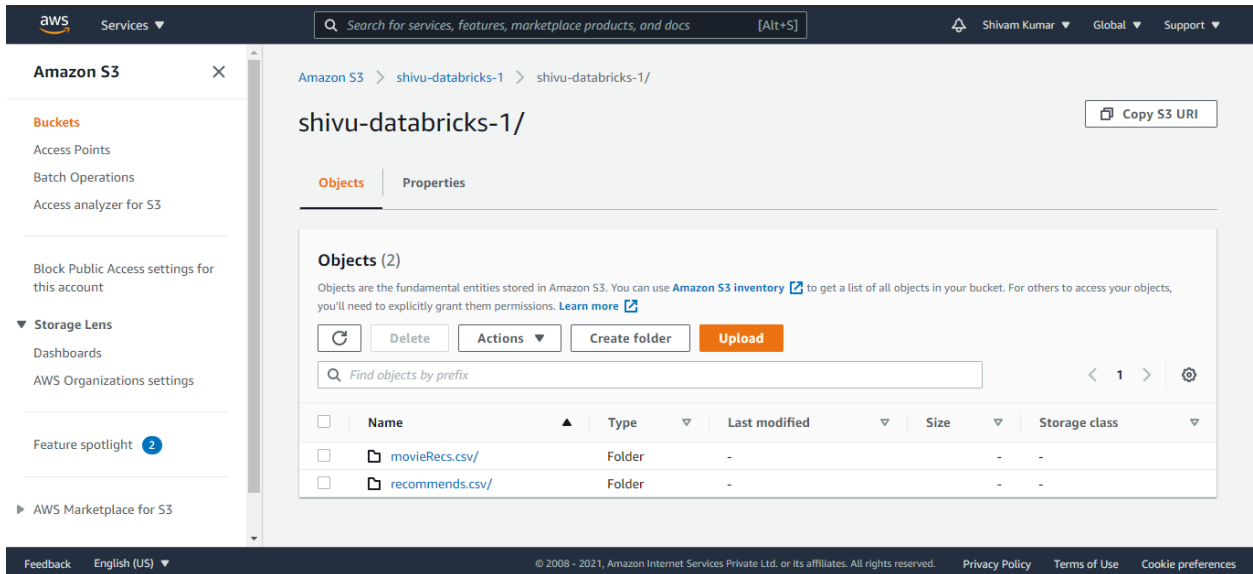
For your protection, you should never share your secret keys with anyone. As a best practice, we recommend frequent key rotation. If you lose or forget your secret key, you cannot retrieve it. Instead, create a new access key and make the old key inactive. Learn more

Create access key

| Access key ID        | Created                   | Last used                                | Status                 |
|----------------------|---------------------------|--|------------------------|
| AKIAVNEIZD2YR3P224EI | 2021-03-10 12:47 UTC+0530 | 2021-03-14 02:01 UTC+0530 with s3 in N/A | Active   Make inactive |

Feedback English (US) © 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences

## Step 6: Saving Final Recommendations Result in a Folder contains CSV Files in our S3 Bucket.



The screenshot shows the Amazon S3 console interface. The left sidebar contains navigation options: Buckets, Access Points, Batch Operations, Access analyzer for S3, Storage Lens, Dashboards, AWS Organizations settings, Feature spotlight (2), and AWS Marketplace for S3. The main content area displays the bucket 'shivu-databricks-1/'. The 'Objects' tab is active, showing a list of objects. There are two folders: 'movieRecs.csv/' and 'recommends.csv/'. The 'Properties' tab is also visible.

Amazon S3 > shivu-databricks-1 > shivu-databricks-1/

shivu-databricks-1/

Objects | Properties

Objects (2)

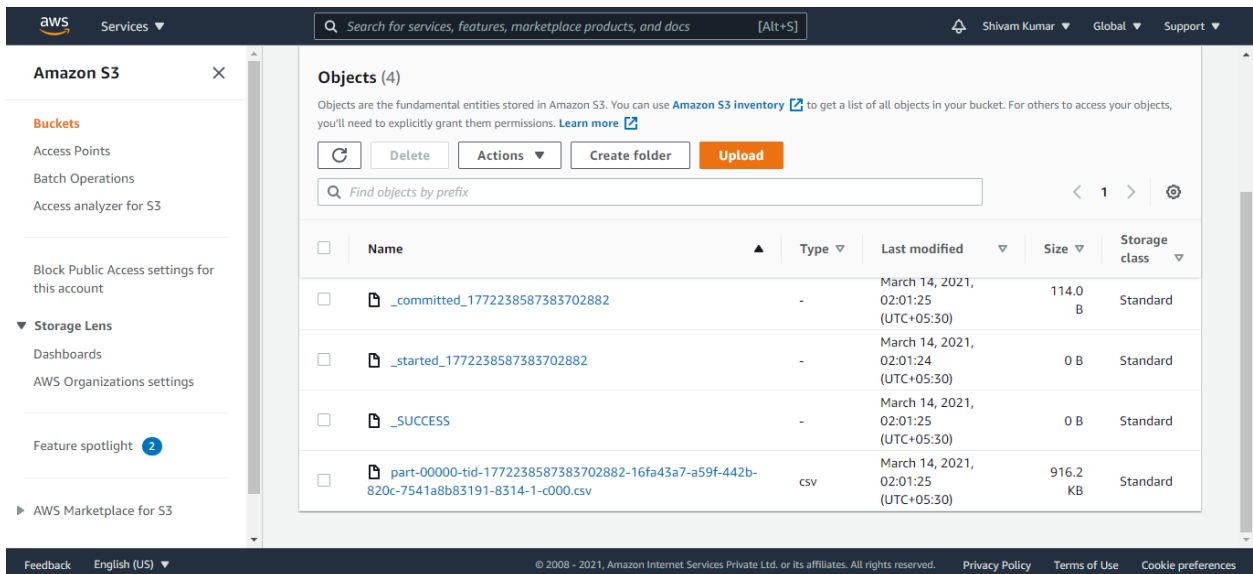
Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Refresh Delete Actions Create folder Upload

Find objects by prefix

| <input type="checkbox"/> | Name            | Type   | Last modified | Size | Storage class |
|--------------------------|-----------------|--------|---------------|------|---------------|
| <input type="checkbox"/> | movieRecs.csv/  | Folder | -             | -    | -             |
| <input type="checkbox"/> | recommends.csv/ | Folder | -             | -    | -             |

Feedback English (US) © 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences



The screenshot shows the Amazon S3 console interface. The left sidebar contains navigation options: Buckets, Access Points, Batch Operations, Access analyzer for S3, Storage Lens, Dashboards, AWS Organizations settings, Feature spotlight (2), and AWS Marketplace for S3. The main content area displays the bucket 'shivu-databricks-1/'. The 'Objects' tab is active, showing a list of objects. There are four objects: '\_committed\_1772238587383702882', '\_started\_1772238587383702882', '\_SUCCESS', and 'part-00000-tid-1772238587383702882-16fa43a7-a59f-442b-820c-7541a8b83191-8314-1-c000.csv'. The 'Properties' tab is also visible.

Amazon S3 > shivu-databricks-1 > shivu-databricks-1/

shivu-databricks-1/

Objects | Properties

Objects (4)

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](#)

Refresh Delete Actions Create folder Upload

Find objects by prefix

| <input type="checkbox"/> | Name  | Type | Last modified                        | Size     | Storage class |
|--------------------------|---|------|--------------------------------------|----------|---------------|
| <input type="checkbox"/> | _committed_1772238587383702882  | -    | March 14, 2021, 02:01:25 (UTC+05:30) | 114.0 B  | Standard      |
| <input type="checkbox"/> | _started_1772238587383702882  | -    | March 14, 2021, 02:01:24 (UTC+05:30) | 0 B      | Standard      |
| <input type="checkbox"/> | _SUCCESS  | -    | March 14, 2021, 02:01:25 (UTC+05:30) | 0 B      | Standard      |
| <input type="checkbox"/> | part-00000-tid-1772238587383702882-16fa43a7-a59f-442b-820c-7541a8b83191-8314-1-c000.csv | csv  | March 14, 2021, 02:01:25 (UTC+05:30) | 916.2 KB | Standard      |

Feedback English (US) © 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences

aws

Services

Search for services, features, marketplace products, and docs

[Alt+S]

Shivam Kumar

Global

Support

Amazon S3

Buckets

Access Points

Batch Operations

Access analyzer for S3

Block Public Access settings for this account

Storage Lens

Dashboards

AWS Organizations settings

Feature spotlight 2

AWS Marketplace for S3

Objects (4)

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](#)

Delete





Actions

Create folder

Upload

Find objects by prefix

< 1 > ⚙

| <input type="checkbox"/> | Name  | Type | Last modified                        | Size     | Storage class |
|--------------------------|---|------|--------------------------------------|----------|---------------|
| <input type="checkbox"/> |  <a href="#">_committed_1715591325667234136</a>  | -    | March 14, 2021, 01:51:44 (UTC+05:30) | 114.0 B  | Standard      |
| <input type="checkbox"/> |  <a href="#">_started_1715591325667234136</a>  | -    | March 14, 2021, 01:51:43 (UTC+05:30) | 0 B      | Standard      |
| <input type="checkbox"/> |  <a href="#">_SUCCESS</a>  | -    | March 14, 2021, 01:51:44 (UTC+05:30) | 0 B      | Standard      |
| <input type="checkbox"/> |  <a href="#">part-00000-tid-1715591325667234136-1f1bfd2a-6ae6-46ef-ad68-49b185977d89-8212-1-c000.csv</a> | csv  | March 14, 2021, 01:51:44 (UTC+05:30) | 515.3 KB | Standard      |

Feedback

English (US)

© 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved.

Privacy Policy

Terms of Use

Cookie preferences