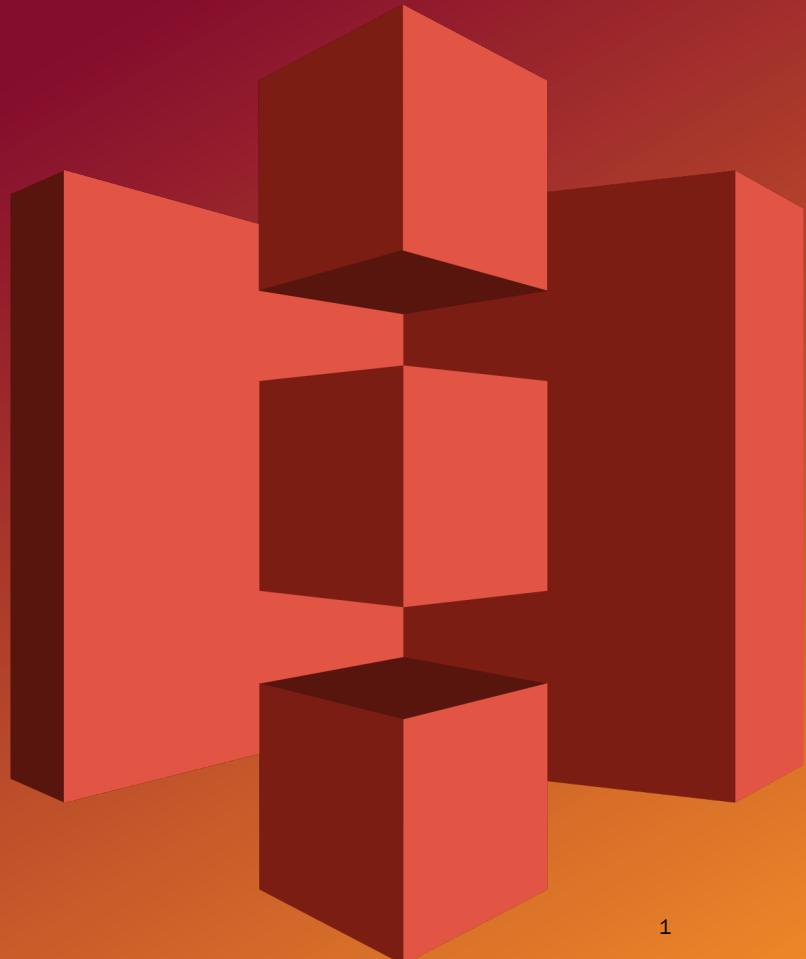


CSE 4510/5310

Big Data

Instructor: Fitzroy Nembhard, Ph.D.

**Setting up AWS S3,
Loading Data into a Bucket
and Performing Operations
on the Data**



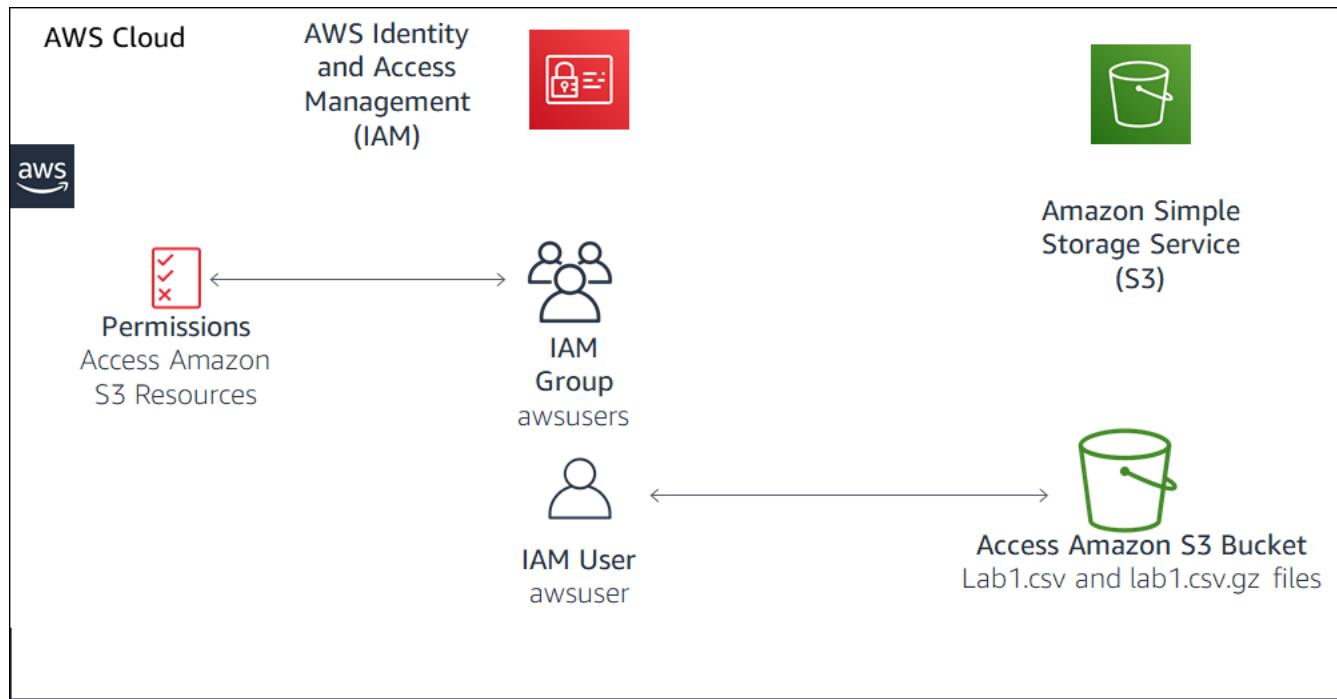
Goals

- To briefly explore AWS S3
- To upload data to S3 and perform basic SQL on the data

AWS S3

- Objectives
- Access Amazon S3 in the AWS Management Console
- Secure an S3 bucket with IAM
- Create a bucket with Amazon S3
- Load data into an S3 bucket
- Query an S3 bucket

Amazon Simple Storage Service (S3)



Hello,
I am Mr. Biggs! I specialize in Big Data. As you can see, I have my data toolkit, and I will be helping your set up and test AWS S3. Since Mr. DB specializes in databases, he may join us for DB-related topics.



Accept the invitation to join the AWS classroom

Join the class on AWS using the invitation email
that should have come to your Florida Tech inbox

AA

AWS Academy

Course Invitation

To: drfitz@fit.edu

Reply-To: reply+24e27c837fad421c-15691~7162367-1642520817@notifications.canvaslms.com

Inbox - Google

You've been invited to participate in a class at AWS Academy . The class is called AWS Academy Learner Lab - Foundation Services [12381]. Course role: Student

Name: **Fitz as Student**

Email: **drfitz@fit.edu**

Username: **none**

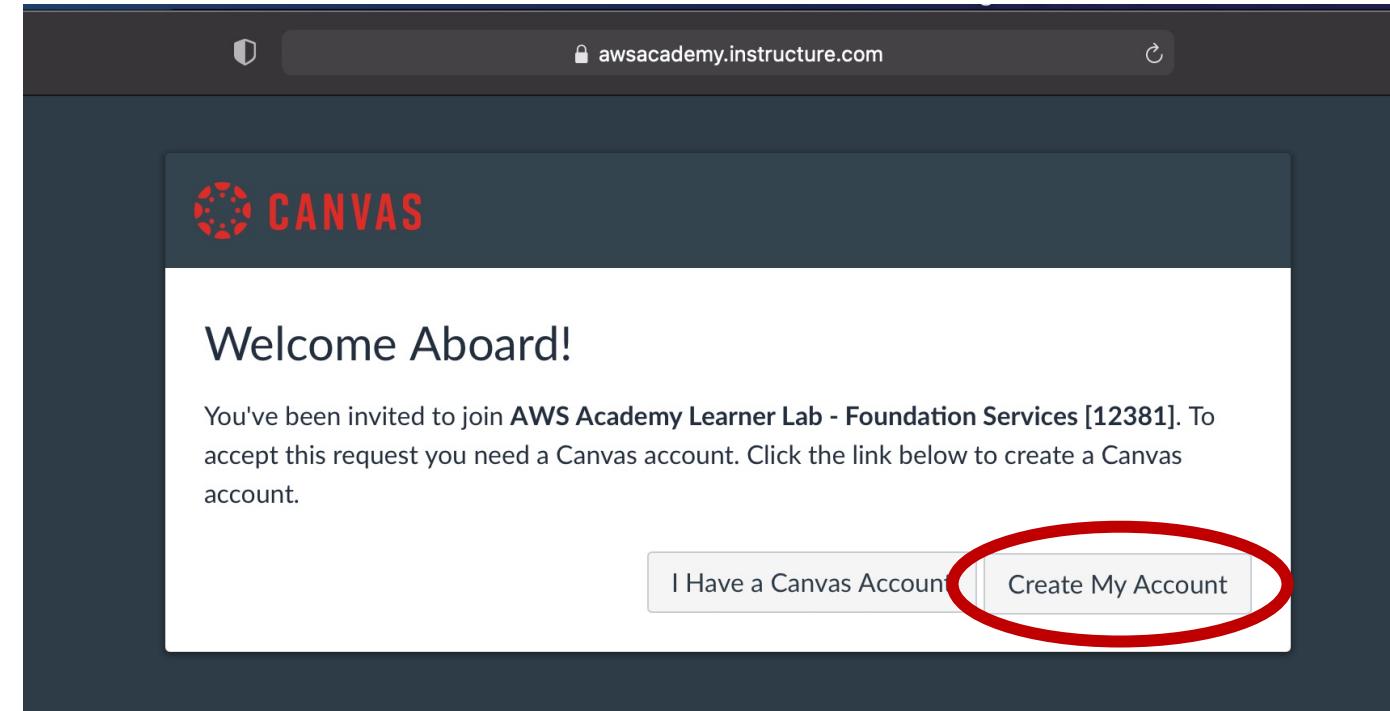
You'll need to register with Canvas before you can participate in the class.

Get Started



Log in Using Canvas

Your AWS Instructure account is different from your Florida Tech Canvas Account. If you have never used AWS Academy before, you will need to create an account by clicking the “Create My Account” button.



Creating a Canvas Account

Enter your credentials to register.

The screenshot shows a web browser window with the URL awsacademy.instructure.com in the address bar. The page title is "CANVAS". The main heading is "Welcome Aboard!". Below it, a message says: "In order to finish signing you up for the course AWS Academy Learner Lab - Foundation Services [12381], we'll need a little more information." There are three input fields: "Login" with the value "drfitz@fit.edu", "Password" with masked input, and "Time Zone" set to "Eastern Time (US & Canada) (-1)". Below these are two checked checkboxes: one for AWS to share AWS news via email, and another agreeing to the Canvas Instructure Acceptable Use Policy and AWS Learner Terms and Conditions. A large orange "Register" button is at the bottom.

Welcome Aboard!

In order to finish signing you up for the course **AWS Academy Learner Lab - Foundation Services [12381]**, we'll need a little more information.

Login: drfitz@fit.edu

Password:

Time Zone: Eastern Time (US & Canada) (-1)

Yes, I'd like Canvas to provide my contact information to [Amazon Web Services\(AWS\)](#) so AWS can share the latest news about AWS services and related offerings with me by email, post or telephone.

You may unsubscribe from receiving AWS news and offers from at any time by following the instructions in the communications received. AWS handles your information as described in the [AWS Privacy Notice](#). Providing Canvas with your information may involve transferring it to another country. For questions about how Canvas will handle your information, please contact Canvas directly or refer to its privacy policy.

I agree to the Canvas [Instructure Acceptable Use Policy](#) and to the [AWS Learner Terms and Conditions](#) The information you provide will be handled by AWS as described in the [AWS Privacy Notice](#).

Register

Welcome Screen

After logging in to Canvas, you should see a screen that resembles the following.

This text displayed is dependent on the course selected. The Big Data Containers course will be a learner lab whereas the general Big Data course will be named Data Analytics.

aws ALLFv1-12381

Home Modules Discussions Courses Calendar Inbox History Help

AWS Academy Learner Lab - Foundation Services [12381]



AWS Academy Learner Lab - Foundation Services provides a long-running sandbox environment for ad hoc exploration of AWS services. Within this class, students will have access to **a restricted set of AWS services**. Not all AWS documentation walk-through or sample labs that operate in an AWS Production account will work in the sandbox environment. You will retain access to the AWS resources set up in this environment for the duration of this course. We limit your budget (\$100), so you should exercise caution to prevent charges that will deplete your budget too quickly. If you exceed your budget, you will lose access to your environment and lose all of your work.

Each session lasts for 4 hours by default, although you can extend a session to run longer by pressing the start button to reset your session timer. At the end of each session, any resources you created will persist. However, we automatically shut EC2 instances down. Other resources, such as RDS instances, keep running. Keep in mind that we do not stop some AWS features, so they can still incur charges between sessions. For example, an Elastic Load Balancer or a NAT. You may wish to delete those types

[View Course Stream](#)
[View Course Calendar](#)
[View Course Notifications](#)

To Do
Nothing for now

Recent Feedback
Nothing for now

Returning to AWS

To return to AWS, go to

https://www.awsacademy.com/LMS_Login



The screenshot shows the AWS Academy login page for instructors. At the top right is the AWS Academy logo. Below it is a "Email" input field containing "drfitz@fit.edu". Below the email field is a "Password" input field, which is currently empty. To the right of the password field is a dropdown menu icon. Below the password field are two checkboxes: one for staying signed in and one for forgot password. To the right of these checkboxes is a "Log In" button. At the bottom of the page, there are links for Help, Privacy Policy, Acceptable Use Policy, Facebook, and Twitter. On the far right, there is a "INSTRUCTURE" logo.

Returning to AWS

The general Big Data Course has guided labs, which were created by AWS. Use this to explore various Big Data Platforms in a guided manner.

The Big Data Containers is for creating clusters and exploring various AWS Big Data platforms without guidance.

If you would like to work on your term project, I recommend you use the Big Data Containers.

The screenshot shows the AWS Dashboard interface. On the left is a sidebar with icons for Account, Dashboard, Courses, Calendar, Inbox, History, Help, and a back arrow. The main area displays two course cards. The first card, 'CSE 4510/5310: Big Data with Dr. Fitz', is highlighted with a red border. It contains the title 'BIG DATA', the subtitle 'AWS Academy Data Analytics [25... ADAv1EN-25002]', and a message icon. The second card, 'CSE 4510/5310: Big Data Containers', contains the title 'BIG DATA Containers', the subtitle 'AWS Academy Learner Lab [27361] ALLv1-27361', and a message icon. To the right of the dashboard are sections for 'To Do' (Nothing for now), 'Recent Feedback' (Nothing for now), and a 'View Grades' button.

Returning to AWS

Click Modules to access course modules.

aws ADAv1EN-25002

Home Modules Discussions Grades

AWS Academy Data Analytics [25002]

View Course Stream View Course Calendar View Course Notifications

To Do Nothing for now

Recent Feedback Nothing for now

AWS Academy Data Analytics is a foundational course for AWS Academy participants who plan to pursue careers in data analytics. The course helps learners develop skills with AWS services that are critical for conducting analysis of big data problems. The course consists of a series of labs that you can integrate with your existing courses on data mining, data analysis, or data science.

Get Started

Select [Modules](#) to start the course. Use [Discussions](#) to connect with peers. Visit [Course Support](#) for help.

Returning to AWS

This screen shows the example of a guided lab that introduces students to Amazon S3.

The screenshot displays the AWS Learning Lab interface. On the left is a dark sidebar with navigation links: Account, Dashboard, Courses (highlighted in blue), Grades, Calendar, Inbox, History, and Help. The main content area shows the path "ADAv1EN-25002 > Modules". Under the "Modules" section, there are three items: "Introduction", "Lab 1", and "Lab 2". "Lab 1" is expanded, showing a task titled "Store data in Amazon S3" with a yellow background and a red border around its icon and text. A "Complete All Items" button with a green checkmark is located to the right of the task. The "Introduction" and "Lab 2" sections have similar "Complete All Items" buttons.

IAM Users

On the AWS Management Console, on the **Services** menu, choose **Services**.

From the list of services, choose **IAM**.

In the navigation pane, choose **Groups**.

Choose the **awsusers** group.

Choose the **Permissions** tab.

Notice that the **AmazonS3FullAccess** policy is attached to the group.

Choose **Show Policy**

The screenshot shows the AWS IAM Groups page. In the left sidebar, under 'Access management', 'User groups' is selected. The main content area shows the 'awsusers' group details. The 'Summary' section includes the user group name (awsusers), creation time (September 10, 2022, 11:12 (UTC-04:00)), and ARN (arn:aws:iam::001495016464:group/awsusers). Below this, the 'Permissions' tab is selected, showing two attached policies: 'AmazonS3FullAccess' (AWS managed, provides full access to all buckets) and 'LimitS3Policy' (Customer inline).

Policy name	Type	Description
AmazonS3FullAccess	AWS managed	Provides full access to all buckets
LimitS3Policy	Customer inline	

IAM Users

The policy document is in JavaScript Object Notation (JSON) format. This policy states that users in that group are allowed to take all actions for Amazon S3 on all resources.

The screenshot shows the AWS Identity and Access Management (IAM) console. The left sidebar has 'User groups' selected under 'Access management'. The main area shows a group named 'awsusers' created on September 10, 2022. It lists one permission policy, 'AmazonS3FullAccess', which provides full access to all buckets via the AWS Management Console. The policy document is displayed in JSON format:

```
1 - [
2 -   "Version": "2012-10-17",
3 -   "Statement": [
4 -     {
5 -       "Effect": "Allow",
6 -       "Action": [
7 -         "s3:*",
8 -         "s3-object-lambda:*
```

S3 Limit Policy

The policy document is in JSON format. This policy states that users in the group are **not** allowed to perform the following specified actions on S3 objects:

ObjectLegalHold – A legal hold prevents an object version from being overwritten or deleted.

ObjectRetention – A retention period determines how long an object is retained.

BucketObjectLock – When an object is locked, it cannot be deleted or overwritten.

The screenshot shows the AWS Identity and Access Management (IAM) console. In the left sidebar, under 'Access management', 'User groups' is selected. The main area displays the 'awsusers' user group. The 'Permissions' tab is active, showing two policies attached to the group: 'AmazonS3FullAccess' (AWS managed) and 'LimitS3Policy' (Customer inline). The 'LimitS3Policy' is expanded, showing the following JSON code:

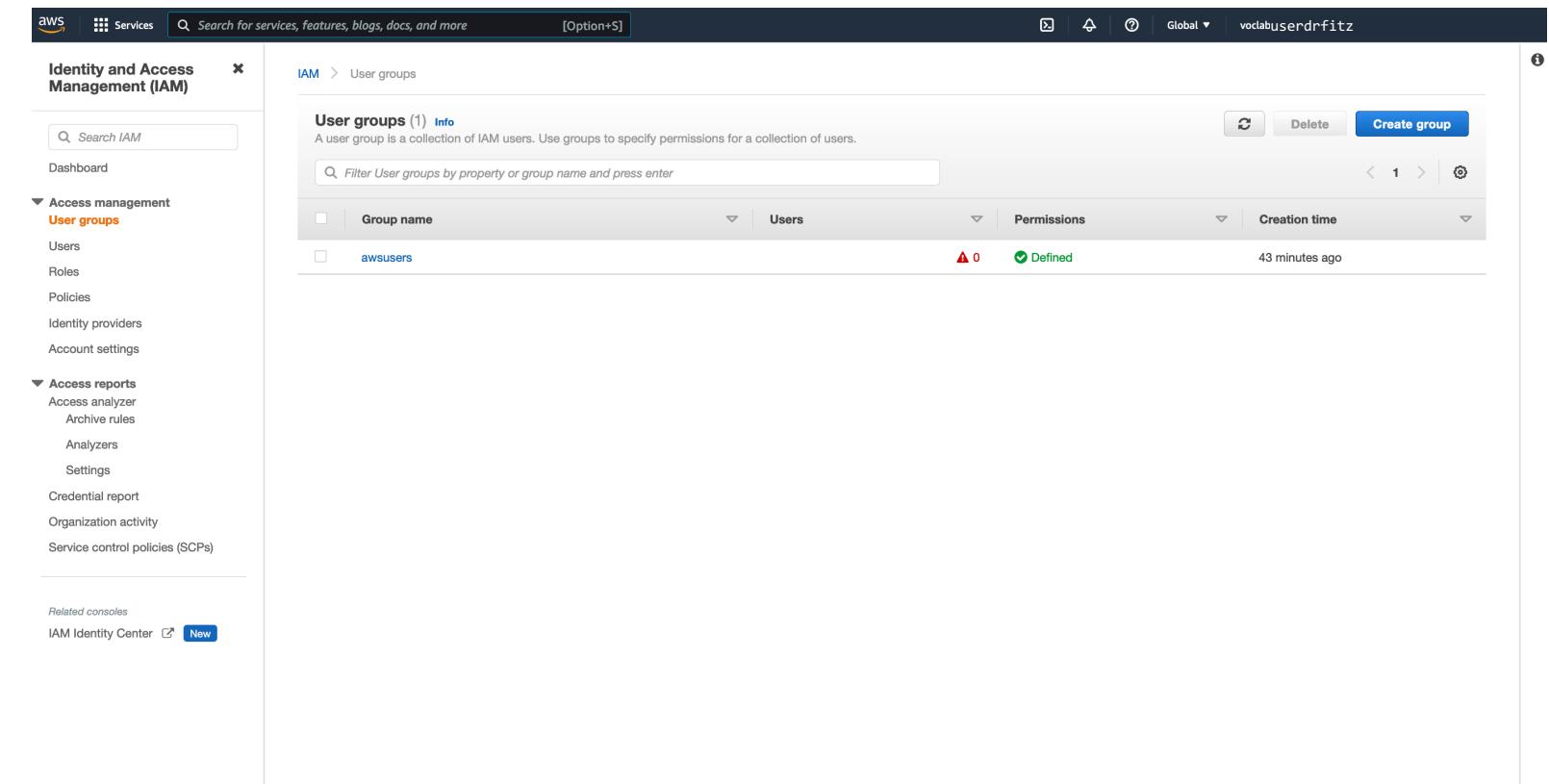
```
1 - {
2 -   "Version": "2012-10-17",
3 -   "Statement": [
4 -     {
5 -       "Action": [
6 -         "s3:*ObjectLegalHold",
7 -         "s3:*ObjectRetention",
8 -         "s3:*BucketObjectLock"
9 -       ],
10 -      "Resource": "*",
11 -      "Effect": "Deny"
12 -    }
13 -  ]
14 -}
```

Add a user

In this task, you will add the **awsuser** to the **awsusers** group. You will also log out of the console and log back in to the console with the **awsuser** account and password.

In the navigation pane, choose **Groups**.

Select the **awsusers** group.



The screenshot shows the AWS Identity and Access Management (IAM) service interface. In the top left, the 'Identity and Access Management (IAM)' logo is visible. The top bar includes the AWS logo, a 'Services' button, a search bar ('Search for services, features, blogs, docs, and more'), and a keyboard shortcut ('[Option+S]'). On the right of the top bar are icons for refresh, search, and help, followed by 'Global' and 'vocabuserdrfritz'. The main navigation menu on the left has 'Access management' expanded, with 'User groups' selected. Other options under 'Access management' include 'Users', 'Roles', 'Policies', 'Identity providers', and 'Account settings'. Below this is another expanded section for 'Access reports' with options like 'Access analyzer', 'Archive rules', 'Analyzers', 'Settings', 'Credential report', 'Organization activity', and 'Service control policies (SCPs)'. At the bottom of the left sidebar, there's a link to 'Related consoles' and 'IAM Identity Center' with a 'New' button. The main content area is titled 'User groups (1) Info' and contains a brief description: 'A user group is a collection of IAM users. Use groups to specify permissions for a collection of users.' A search bar at the top of the list table says 'Filter User groups by property or group name and press enter'. The table lists one item: 'awsusers'. The columns for the table are 'Group name', 'Users', 'Permissions', and 'Creation time'. The 'awsusers' row shows '0' users, 'Defined' status, and was created '43 minutes ago'. Action buttons for 'Delete' and 'Create group' are located at the top right of the table.

Add a user

From the Group Actions menu, choose **Add Users to Group**.

The screenshot shows the AWS Identity and Access Management (IAM) service in the AWS Management Console. The URL is https://us-east-1.console.aws.amazon.com/iamv2/home?region=us-east-1#/groups/details/awsusers)section=users. The left sidebar shows the navigation menu for IAM, with 'User groups' selected under 'Access management'. The main content area displays the 'awsusers' group summary. The 'Summary' section includes the user group name (awsusers), creation time (September 10, 2022, 11:12 (UTC-04:00)), and ARN (arn:aws:iam::001495016464:group/awsusers). Below this, there are tabs for 'Users', 'Permissions', and 'Access Advisor', with 'Users' selected. The 'Users in this group (0)' section contains a search bar and a table header with columns for 'User name', 'Groups', 'Last activity', and 'Creation time'. A red box highlights the 'Add users' button located at the top right of this section.

Add a user

Select the **awsuser** user.

Choose **Add Users**.

From the navigation header, open the list of account actions and copy the account ID

The screenshot shows the AWS Identity and Access Management (IAM) console. The left sidebar is titled "Identity and Access Management (IAM)" and includes sections for "Access management" (User groups, Users, Roles, Policies, Identity providers, Account settings), "Access reports" (Access analyzer, Archive rules, Analyzers, Settings, Credential report, Organization activity, Service control policies (SCPs)), and "Related consoles" (IAM Identity Center). The main content area is titled "Add users to awsusers" and shows a table of "Other users in this account". A single user, "awsuser", is listed with a checked checkbox next to its name. The "awsuser" row is highlighted with a red box. The table has columns for "User name", "Groups", "Last activity", and "Creation time". At the bottom right of the table are "Cancel" and "Add users" buttons.

User name	Groups	Last activity	Creation time
awsuser	0 None	45 minutes ago	

Locate User ID

From the navigation header, open the list of account actions and copy the account ID.

The screenshot shows the AWS Identity and Access Management (IAM) service interface. In the top left, the AWS logo and 'Services' button are visible. The search bar contains the placeholder 'Search for services, features, blogs, docs, and more'. The top right shows the user name 'userdrfritz' and a 'Global' dropdown. A blue info box at the top right says 'New feature to generate a policy based on CloudTrail events. AWS uses your CloudTrail events to identify the services and actions used and generate a least privileged policy that you can attach to this user.' Below this, the navigation path 'Users > awsuser' is shown. On the left, a sidebar menu includes 'Identity and Access Management (IAM)' (selected), 'Dashboard', 'Access management' (selected), 'User groups', 'Users' (selected), 'Roles', 'Policies', 'Identity providers', 'Account settings', 'Access reports' (selected), 'Access analyzer', 'Archive rules', 'Analyzers', 'Settings', 'Credential report', 'Organization activity', and 'Service control policies (SCPs)'. A search bar for 'Search IAM' is also present. The main content area is titled 'Summary' and displays the following information:

- User ARN: arn:aws:iam::001495016464:user/awsuser
- Path: /
- Creation time: 2022-09-10 11:13 EDT

A tab bar at the top of the summary section includes 'Permissions' (selected), 'Groups (1)', 'Tags (1)', 'Security credentials', and 'Access Advisor'. The 'Permissions' section shows 'Permissions policies (2 policies applied)'. It lists two policies:

Policy name	Policy type
AmazonS3FullAccess	AWS managed policy from group awsusers
LimitS3Policy	Inline policy from group awsusers

The 'Permissions boundary (not set)' section notes that a boundary can be set to control maximum permissions. A 'Set boundary' button is available, and a note states 'No permissions boundary is set for this user.' A note at the bottom of the summary section says 'This user can perform all actions that are allowed by the user's permission policies.'

Sign Out

In the list of account actions, choose **Sign Out**.

To sign back in with the awsuser credentials,
choose **Sign In to the Console**.

The screenshot shows the AWS Identity and Access Management (IAM) service. The left sidebar lists various IAM management options. The main content area is titled "Password policy" and describes the default password policy for the AWS account. It includes rules like minimum length of 8 characters and mixing character types. A "Change password policy" button is present. Below this is a section on "Security Token Service (STS)" with information about session tokens from regional and global endpoints. A warning message at the bottom states "You need permissions" because the user lacks the required permissions to perform the operation. The top right corner of the main content area is highlighted with a red box, showing the account ID (XXXXXX), federated user details (voclabs/userDrFitz), and navigation links for Account, Organization, Service Quotas, Billing Dashboard, and Settings. The "Sign out" button is also visible.

Identity and Access Management (IAM)

- Dashboard
- Access management
 - User groups
 - Users
 - Roles
 - Policies
 - Identity providers
 - Account settings**
- Access reports
 - Access analyzer
 - Archive rules
 - Analyzers
 - Settings
- Credential report
- Organization activity
- Service control policies (SCPs)

Search IAM

AWS account ID:
XXXXXXXXXX

>Password policy

This AWS account uses the following default password policy:

- Minimum password length is 8 characters
- Include a minimum of three of the following mix of character types: uppercase, lowercase, numbers, and ! @ # \$ % ^ & * () _ + - = [] { } | ' .
- Must not be identical to your AWS account name or email address

Change password policy

Security Token Service (STS)

Session Tokens from the STS endpoints

AWS recommends using regional STS endpoints to reduce latency. Session tokens from regional STS endpoints are valid in all AWS Regions. If you use regional STS endpoints, session tokens from the global STS endpoint (<https://sts.amazonaws.com>) are valid only in AWS Regions that are enabled by default. If you intend to enable a new Region for your account, you can use session tokens from regional STS endpoints or activate the global STS endpoint to issue session tokens that are valid in all AWS Regions. [Learn more](#)

Endpoints	Region compatibility of session tokens	Actions
Global endpoint	Valid only in AWS Regions enabled by default	Edit
Regional endpoints	Valid in all AWS Regions	

Endpoints

You can enable additional endpoints from which you can request temporary credentials. Activate only endpoints you intend to use. [Learn more](#)

You need permissions

You do not have the permission required to perform this operation. Ask your administrator to add permissions. [Learn more](#)

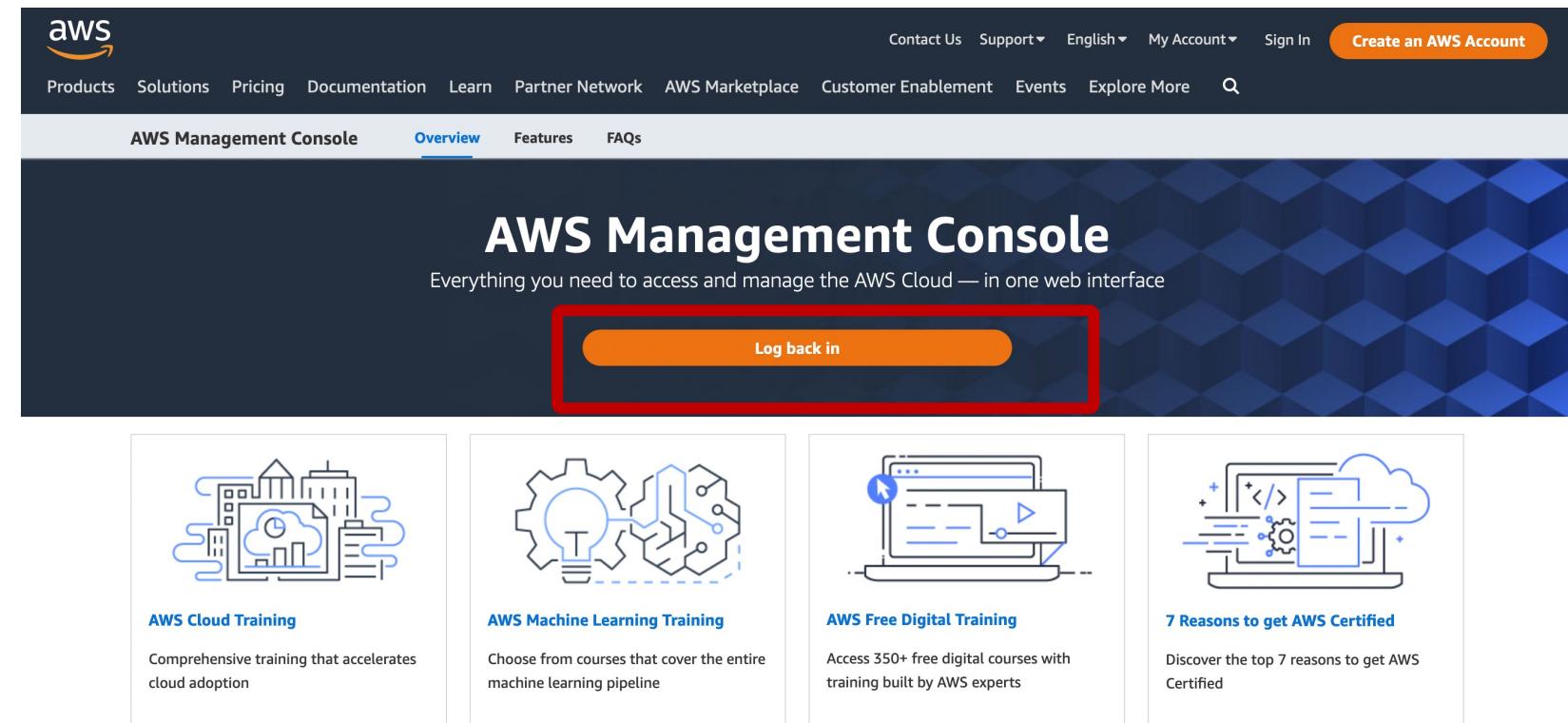
User: arn:aws:sts::XXXXXXXXXX userDrFitz =Fitzroy_Nembhard is not authorized to perform: iam:*

Region name	Endpoint	STS status	Actions

Log in as IAM User

In the list of account actions, choose [Sign Out](#).

To sign back in with the [awsuser](#) credentials, choose [Sign In to the Console](#).



The screenshot shows the AWS Management Console homepage. At the top, there is a navigation bar with links for Products, Solutions, Pricing, Documentation, Learn, Partner Network, AWS Marketplace, Customer Enablement, Events, Explore More, and a search icon. On the far right of the navigation bar are links for Contact Us, Support, English, My Account, Sign In, and a prominent orange button labeled "Create an AWS Account". Below the navigation bar, there are tabs for "AWS Management Console", "Overview", "Features", and "FAQs", with "Overview" being the active tab. The main title "AWS Management Console" is displayed prominently in large white text against a dark blue background. Below the title, a subtitle reads "Everything you need to access and manage the AWS Cloud — in one web interface". A large orange button with the text "Log back in" is centered on the page, surrounded by a red rectangular border. The page features four main call-to-action boxes: "AWS Cloud Training" (with an icon of a city skyline and clouds), "AWS Machine Learning Training" (with an icon of a brain and gears), "AWS Free Digital Training" (with an icon of a computer monitor displaying code), and "7 Reasons to get AWS Certified" (with an icon of a cloud and code). Each box contains a brief description and a link.

Log in as IAM User

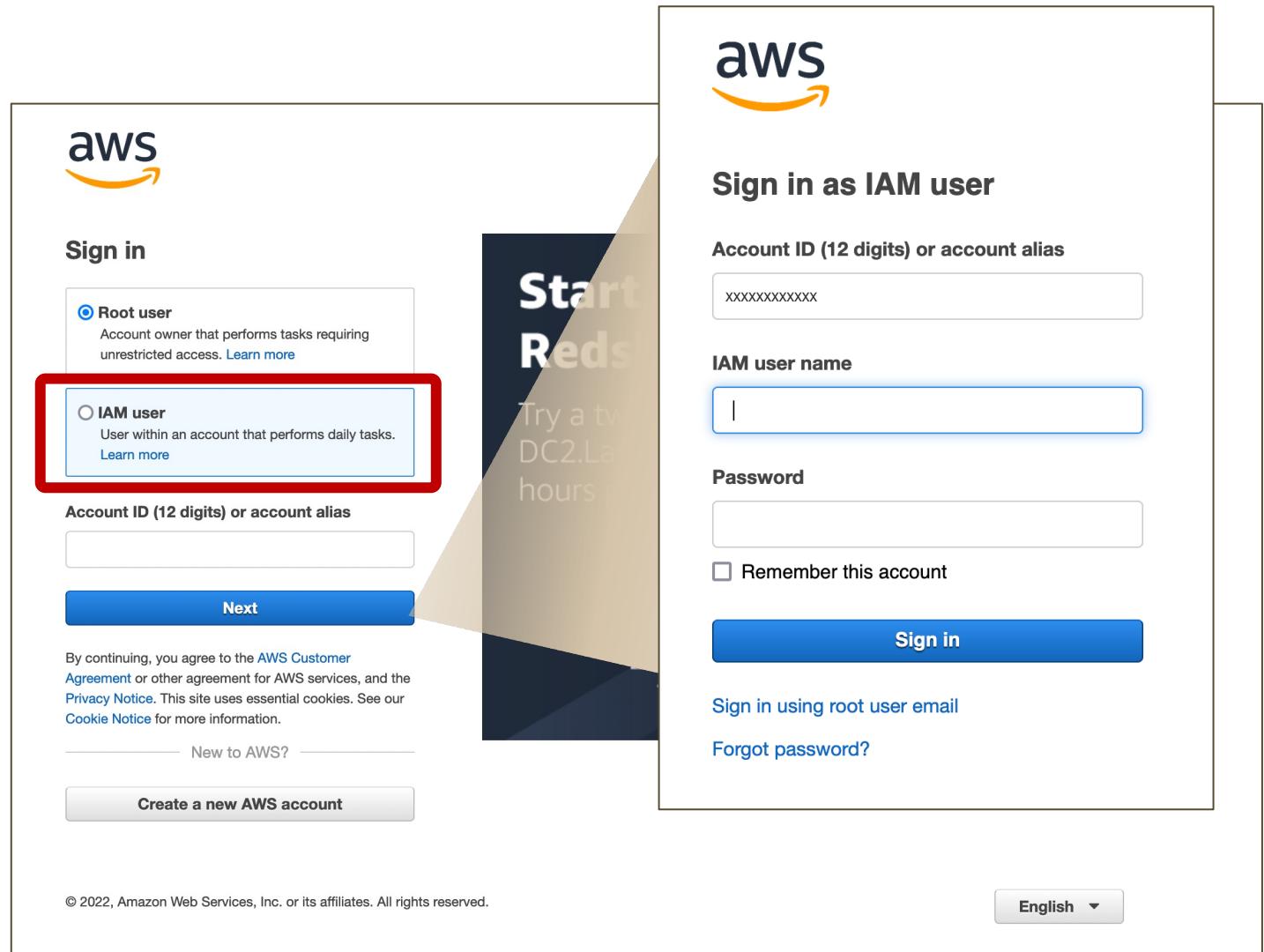
Select **IAM user** and then use the following information to sign in:

Note: Remove the dashes from the account number before you enter it.

Account: The account ID that you previously copied

IAM user name: awsuser

Password: myP@ssW0rd



The image shows two side-by-side AWS sign-in interfaces. On the left is the standard AWS Sign In page, and on the right is a simplified version. Both pages have identical fields: 'Account ID (12 digits) or account alias' (containing 'xxxxxxxxxx'), 'IAM user name' (containing 'I'), 'Password' (empty), and a 'Sign in' button. The left page includes additional options like 'Root user' and 'Create a new AWS account'. The right page has a large 'Start' button and a 'Try a free trial' banner.

AWS Sign In (Left):

- Account ID (12 digits) or account alias: xxxxxxxxxx
- Root user
- Account owner that performs tasks requiring unrestricted access. [Learn more](#)
- IAM user
- User within an account that performs daily tasks. [Learn more](#)
- Next
- By continuing, you agree to the [AWS Customer Agreement](#) or other agreement for AWS services, and the [Privacy Notice](#). This site uses essential cookies. See our [Cookie Notice](#) for more information.
- New to AWS? [Create a new AWS account](#)

AWS Sign in as IAM user (Right):

- Account ID (12 digits) or account alias: xxxxxxxxxx
- IAM user name: I
- Password: (empty)
- Sign in
- Remember this account:
- Sign in using root user email
- Forgot password?
- English ▾

Select S3

Buckets and objects are the basic building blocks for Amazon S3. You create buckets and add objects to the buckets. Objects in Amazon S3 can be up to 5 TB.

The screenshot shows the AWS Console Home page. On the left, there's a sidebar with 'Recently visited' (Console Home, S3, Athena, IAM), 'Favorites' (Analytics, Application Integration, AR & VR, AWS Cost Management, Blockchain, Business Applications, Compute, Containers, Customer Enablement, Database, Developer Tools, End User Computing, Front-end Web & Mobile, Game Development, Internet of Things, Machine Learning, Management & Governance, Media Services, Migration & Transfer, Networking & Content Delivery, Quantum Technologies, Robotics, Satellite), and 'All services'. The main area displays 'Recently visited' items: 'Console Home', 'S3 Scalable Storage in the Cloud' (which is highlighted with a red rectangle), 'Athena Query Data in S3 using SQL', and 'IAM Manage access to AWS resources'. To the right, there are 'Welcome to AWS' sections for 'Getting started with AWS', 'Training and certification', and 'What's new with AWS?'. At the bottom, it says 'No cost and usage'.

Create a Bucket

Choose [Create bucket](#).

Screenshot of the AWS S3 service page.

The page title is "Amazon S3" under the "Storage" category. The main heading is "Amazon S3" with the subtext "Store and retrieve any amount of data from anywhere". A description below states: "Amazon S3 is an object storage service that offers industry-leading scalability, data availability, security, and performance." To the right, there is a "Create a bucket" section with a large orange "Create bucket" button. Below this is a "Pricing" section stating "With S3, there are no minimum fees. You only pay for what you use. Prices are based on the location of your S3 bucket." It includes a link to the "Simple Monthly Calculator". Further down is a "Resources" section with links to "User guide", "API reference", and "FAQs". On the left side of the main content area, there is a video thumbnail titled "Introduction to Amazon S3" with the URL "aws.amazon.com/S3". The thumbnail features the AWS logo and a play button icon.

Create a Bucket

Enter a bucket name with three or more characters.
Uppercase characters are not allowed.

Choose **Create bucket**.

Note: S3 bucket names must be unique across all buckets in Amazon S3. If you get a conflict with another bucket, add a digit and try again.

Note: Write down the bucket name because it will be used in future steps

The screenshot shows the AWS S3 'Create bucket' wizard. The 'General configuration' tab is active, displaying fields for 'Bucket name' (containing 'cse4510bucket1') and 'AWS Region' (set to 'US East (N. Virginia) us-east-1'). A note about unique bucket names is present. The 'Object Ownership' tab is also visible, with 'ACLs disabled (recommended)' selected. The bottom of the screen shows the navigation bar with the Florida Tech logo.

Amazon S3 > Buckets > Create bucket

Create bucket Info

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

General configuration

Bucket name

cse4510bucket1

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

AWS Region

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

Copy settings from existing bucket - *optional*
Only the bucket settings in the following configuration are copied.

Choose bucket

Object Ownership Info

Control ownership of objects written to this bucket from other AWS accounts and the use of 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 enforced

Create a Bucket

You should see a message that the Bucket was successfully created.

The screenshot shows the AWS S3 Buckets page. At the top, a green banner displays the message "Successfully created bucket 'cse4510bucket1'". Below this, a blue bar offers guidance on "S3 Intelligent-Tiering". The main content area shows an "Account snapshot" and a table of buckets. The table has one item:

Name	AWS Region	Access	Creation date
cse4510bucket1	US East (N. Virginia) us-east-1	Bucket and objects not public	September 10, 2022, 12:47:53 (UTC-04:00)

Upload Data

In this task, you will upload an object to the S3 bucket that you created. First, download the lab1.csv file from Canvas to a local directory.

Choose the bucket that you created in the previous task.

In the Amazon S3 console, choose **Upload**.

The screenshot shows the Amazon S3 console interface. On the left, there is a sidebar with various options like Buckets, Storage Lens, and AWS Marketplace. The main area shows a bucket named 'cse4510bucket1'. At the top, there is a banner with a link to a tutorial. Below the banner, the 'Objects' tab is selected, showing a message that says 'Objects (0)' and 'No objects'. In the center, there is a large orange 'Upload' button, which is highlighted with a red rectangle. Above the 'Upload' button, there is a row of buttons for Copy S3 URI, Copy URL, Download, Open, Delete, Actions, and Create folder. There is also a search bar labeled 'Find objects by prefix' and a pagination indicator showing page 1 of 1.

Upload Data

Choose Add files.

The screenshot shows the AWS S3 'Upload' interface. At the top, there's a navigation bar with the AWS logo, a 'Services' dropdown, a search bar containing 'Search for services, features, blogs, docs, and more [Option+S]', and various global settings like 'Global' and a user profile. Below the navigation, the path is shown as 'Amazon S3 > Buckets > cse4510bucket1 > Upload'. The main title is 'Upload' with an 'Info' link. A sub-instruction says: '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' with a blue link icon. Below this is a dashed box with the text 'Drag and drop files and folders you want to upload here, or choose Add files, or Add folders.' A red box highlights the 'Add files' button in a row of buttons labeled 'Remove', 'Add files', and 'Add folder'. Below this is a table titled 'Files and folders (0)' with a single row: 'All files and folders in this table will be uploaded.' A search bar with 'Find by name' placeholder text is present. The table has columns: Name, Folder, Type, and Size. A message 'No files or folders' indicates no uploads have been made. In the 'Destination' section, it shows 'Destination s3://cse4510bucket1' and a 'Destination details' link with the sub-instruction: 'Bucket settings that impact new objects stored in the specified destination.'

Upload Data

Browse to the directory where you stored the lab1.csv file.

Choose the **lab1.csv** file.

Choose **Upload**.

The screenshot shows the AWS S3 console's 'Upload' interface. At the top, the navigation bar includes the AWS logo, 'Services' dropdown, search bar ('Search for services, features, blogs, docs, and more'), and global settings ('Global' and user 'userdrfitz'). The current path is 'Console Home > S3 > Buckets > cse4510bucket1 > Upload'. The main area is titled 'Upload' with an 'Info' link. A note says: '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'.

A large dashed blue box provides instructions: 'Drag and drop files and folders you want to upload here, or choose Add files, or Add folders.'

The 'Files and folders' section shows one item: '1 Total, 337.0 B'. It lists 'lab1.csv' as a 'text/csv' file. Buttons for 'Remove', 'Add files', and 'Add folder' are available. A search bar 'Find by name' is present.

The 'Destination' section shows the destination as 's3://cse4510bucket1'. It includes a 'Destination details' link and a note about bucket settings.

Expansion arrows lead to sections for 'Permissions' (Grant public access and access to other AWS accounts) and 'Properties' (Specify storage class, encryption settings, tags, and more).

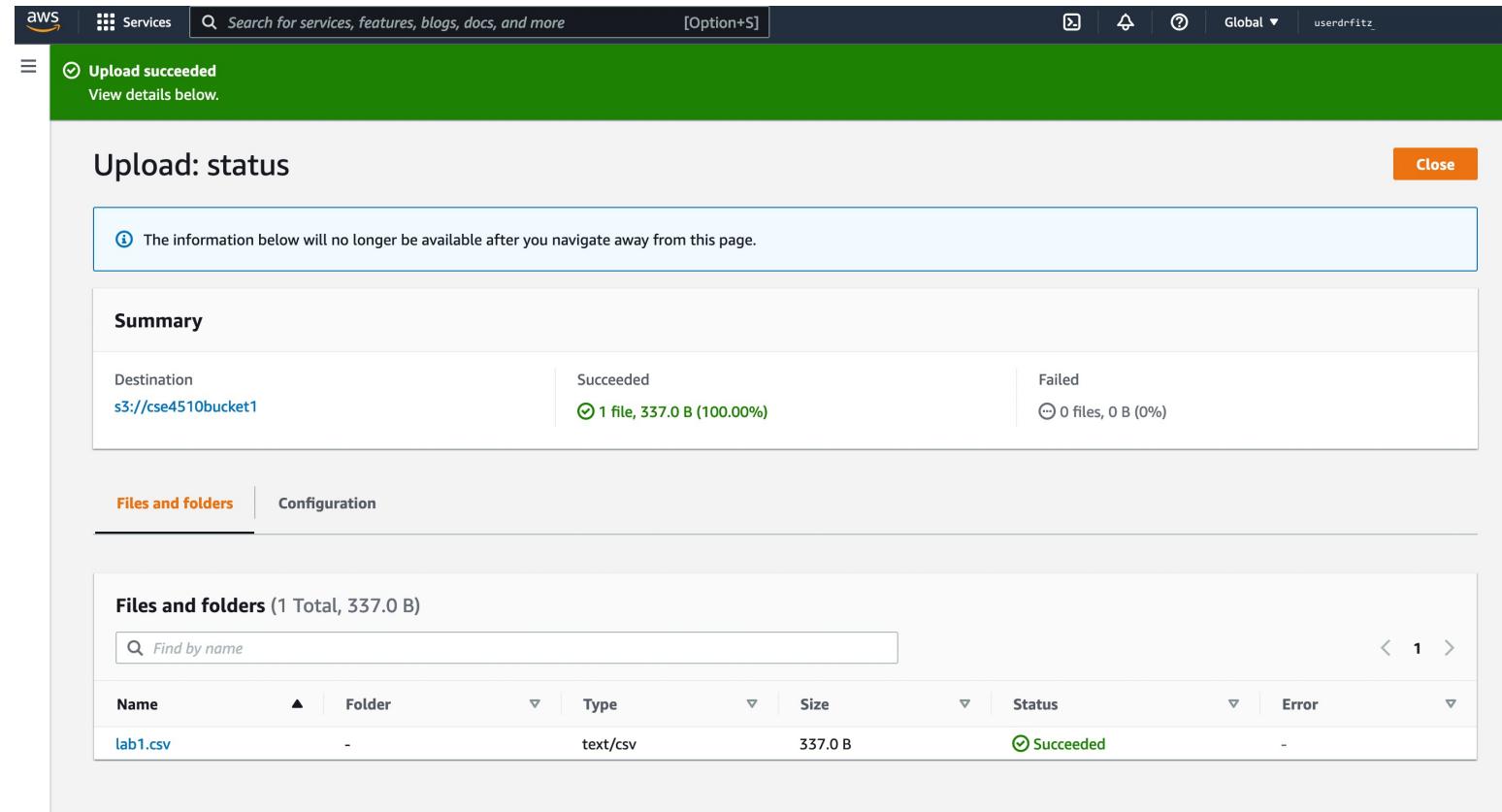
At the bottom right are 'Cancel' and a prominent orange 'Upload' button, which is highlighted with a red rectangle.

Upload Data

Browse to the directory where you stored the lab1.csv file.

Choose the **lab1.csv** file.

Choose **Upload**.



The screenshot shows the AWS S3 "Upload: status" page. At the top, a green banner indicates "Upload succeeded". Below it, a message states: "The information below will no longer be available after you navigate away from this page." The "Summary" section shows the destination as "s3://cse4510bucket1" with 1 file uploaded (100.00% success). The "Files and folders" tab is selected, showing a table with one item: "lab1.csv" (text/csv, 337.0 B, Succeeded).

Name	Folder	Type	Size	Status	Error
lab1.csv	-	text/csv	337.0 B	Succeeded	-

Review Uploaded File

In the Amazon S3 console, choose the **lab1.csv** file.

Review the file properties for the file that you uploaded.

Note: You should get a message stating that versioning is not enabled for the bucket. This behavior is expected.

The screenshot shows the Amazon S3 console interface. The left sidebar has 'Amazon S3' selected under 'Buckets'. The main area shows the file 'lab1.csv' in the bucket 'cse4510bucket1'. The 'Properties' tab is active. The 'Object overview' section displays the following details:

Attribute	Value
Owner	awslabsc0w4031887t1647838422
AWS Region	US East (N. Virginia) us-east-1
Last modified	September 10, 2022, 12:59:20 (UTC-04:00)
Size	337.0 B
Type	csv
Key	lab1.csv
S3 URI	s3://cse4510bucket1/lab1.csv
Amazon Resource Name (ARN)	arn:aws:s3:::cse4510bucket1/lab1.csv
Entity tag (Etag)	8cb3597be2d8c8774b3278562e27a9af
Object URL	https://cse4510bucket1.s3.amazonaws.com/lab1.csv

Query the object you uploaded

From the Object actions menu, choose **Query with S3 Select**.

The screenshot shows the AWS S3 console. On the left, the navigation pane is visible with sections like Buckets, Storage Lens, and Marketplace for S3. The main area displays the details for an object named 'lab1.csv' located in the 'cse4510bucket1' bucket. The 'Properties' tab is selected. In the top right, there is a 'Object actions' dropdown menu. The 'Query with S3 Select' option is highlighted with a red box.

Object overview

Owner	S3 URI
awslabsc0w4031887t1647838422	s3://cse4510bucket1/lab1.csv
AWS Region	Amazon Resource Name (ARN)
US East (N. Virginia) us-east-1	arn:aws:s3:::cse4510bucket1/lab1.csv
Last modified	Entity tag (Etag)
September 10, 2022, 12:59:20 (UTC-04:00)	8cb3597be2d8c8774b3278562e27a
Size	Object URL
337.0 B	https://cse4510bucket1.s3.amazonaws.com/lab1.csv
Type	
CSV	
Key	
lab1.csv	

Query the object you uploaded

From the **Object actions** menu, choose **Query with S3 Select**.

The screenshot shows the AWS S3 console. In the top navigation bar, the user is in the 'Services' section, with the search bar containing 'Search for services, features, blogs, docs, and more' and the keyboard shortcut '[Option+S]'. The user is signed in as 'userdrfitz_'. A banner at the top right encourages getting started with S3 Intelligent-Tiering.

The main content area shows the path: Amazon S3 > Buckets > cse4510bucket1 > lab1.csv > Query with S3 Select. The title 'Query with S3 Select' has an 'Info' link. Below it, a description explains using Amazon S3 Select to retrieve a subset of data from an object using standard SQL queries, mentioning pricing based on input size and data transferred, with links to 'Learn more' and 'Amazon S3 pricing'.

The 'Input settings' section contains the following configuration:

- Path:** s3://cse4510bucket1/lab1.csv
- Size:** 337.0 B
- Format:** CSV (radio button selected)
- CSV delimiter:** Comma (radio button selected)
- Exclude the first line of CSV data:** (checkbox is unchecked)
- Compression:** None (radio button selected)

Query the object you uploaded

Scroll down the page and choose Run SQL query.

The screenshot shows the AWS S3 console with a sidebar on the left containing navigation links like Buckets, Access Points, Object Lambda Access Points, Multi-Region Access Points, Batch Operations, and Access analyzer for S3. Below these are sections for Block Public Access settings for this account and Storage Lens. A Feature spotlight section is also present. The main content area has a blue banner at the top with the text: "Get hands-on guidance on how to get started with S3 Intelligent-Tiering and experience automatic storage cost savings." It includes a "View tutorial" button. On the right, there are settings for Format (CSV selected), CSV delimiter (Comma selected), and a SQL query editor. The SQL query editor contains two lines of code:

```
1 /* To create reference point for writing SQL queries, you can display the first 5 records of input data by running the following SQL command:  
2 : SELECT * FROM s3object s LIMIT 5 */  
2 SELECT * FROM s3object s LIMIT 5
```

A red box highlights the "Run SQL query" button.

Query the object you uploaded

You should see the first few records from the file.

The screenshot shows the AWS S3 console with the 'Select' feature open. The left sidebar lists 'Buckets', 'Access Points', 'Object Lambda Access Points', 'Multi-Region Access Points', 'Batch Operations', and 'Access analyzer for S3'. Under 'Storage Lens', there are 'Dashboards' and 'AWS Organizations settings'. A 'Feature spotlight' section is present. The main content area has a header: 'Get hands-on guidance on how to get started with S3 Intelligent-Tiering and experience automatic storage cost savings.' It includes a 'View tutorial' button and a 'Run SQL query' button. Below this is a code editor with the following SQL query:

```
1 /* To create reference point for writing SQL queries, you can display the first 5 records of input data by running the following SQL command
   : SELECT * FROM s3object s LIMIT 5 */
2 SELECT * FROM s3object s LIMIT 5
```

Under 'Query results', it says 'Successfully returned 5 records in 161 ms' and 'Bytes returned: 337 B'. There are 'Raw' and 'Formatted' tabs, with 'Formatted' selected. The output is a CSV file containing the following data:

CustomerID	First Name	Last Name	Join Date	Street Address	City	State	Phone
001	Alejandro	Rosalez	12/12/2013	123 Main St.	Baltimore	MD	765-234-2349
002	Jane	Doe	10/5/2014	456 State St.	Seattle	WA	415-889-4932
003	John	Stiles	9/20/20016	1980 8th St.	Brooklyn	NY	917-123-9308
004	Li	Juan	6/29/2011	1323 22nd Ave.	Albany	NY	917-332-3432

Add SQL Template

Choose Add SQL from templates.

The screenshot shows the AWS S3 console. On the left, there's a sidebar with options like Buckets, Access Points, Object Lambda Access Points, Multi-Region Access Points, Batch Operations, and Access analyzer for S3. Below that are sections for Block Public Access settings and Storage Lens. A Feature spotlight section is also present. The main area is titled "Amazon S3" and has a sub-header "Get hands-on guidance on how to get started with S3 Intelligent-Tiering and experience automatic storage cost savings." It includes settings for Format (CSV selected), CSV delimiter (Comma selected), and a SQL query editor. The SQL query editor contains two buttons: "Add SQL from templates" (which is highlighted with a red box) and "Run SQL query". Below the buttons is some sample SQL code:

```
1 /* To create reference point for writing SQL queries, you can display the first 5 records of input data by running the following SQL command
   : SELECT * FROM s3object s LIMIT 5 */
2 SELECT * FROM s3object s LIMIT 5
```

Add SQL Template

Choose **SELECT COUNT * FROM s3object s.**

Choose **Copy SQL.**

The screenshot shows the AWS Lambda console interface. On the left, there's a sidebar with links like 'Buckets', 'Access Points', 'Object Lambda Access Points', etc. In the center, a modal window titled 'SQL templates' is open. It contains a list of SQL query options, with the first one, 'SELECT * FROM s3object s LIMIT 5', selected. Below the list are buttons for 'Cancel', 'Copy SQL' (which is highlighted in orange), and 'Append SQL'. At the bottom of the modal, there's a 'Download results' button. To the right of the modal, there's some status information: 'Status' with a checkmark and the text 'Successfully returned 5 records in 161 ms', and 'Bytes returned: 337 B'. At the very bottom of the screen, there are 'Raw' and 'Formatted' tabs.

SQL templates

- SELECT * FROM s3object s LIMIT 5
- SELECT s._1, s._2 FROM s3object s
- SELECT * FROM s3object s WHERE s._N = 'xyz'
- SELECT * FROM s3object s WHERE s._N like "%xyz%"
- SELECT * FROM s3object s WHERE CAST(s._N as INTEGER) = 1234
- SELECT * FROM s3object s WHERE CAST(s._N as FLOAT) > 12.34
- SELECT count(*) FROM s3object s
- SELECT s.key1 FROM S3Object s
- SELECT s.array1[0].key1 FROM S3Object s

Cancel Copy SQL Append SQL

Download results

Status

Successfully returned 5 records in 161 ms

Bytes returned: 337 B

Raw Formatted

Add SQL Template

Replace the previous query by deleting it and then paste the query you copied.

Choose **Run SQL query**.

In the **Result** pane, you should get the total number of records, which is 5.

SQL query

Amazon S3 Select supports only the SELECT SQL command. Using the S3 console, you can extract up to 40 MB of records from an object that is up to 128 MB in size. To work with larger files or more records, use the AWS CLI, AWS SDK, or Amazon S3 REST API. For more complex SQL queries, use [Amazon Athena](#).

[Add SQL from templates](#) [Run SQL query](#)

```
1 /* To create reference point for writing SQL queries, you can display the first 5 records of input data by running the following SQL query
   : SELECT * FROM s3object s LIMIT 5 */
2 SELECT count(*) FROM s3object s
```

Query results

Query results are not available after you choose **Close** or navigate away. Choose **Download results** to download a copy of the following query results.

[Download results](#)

Status

Successfully returned 1 record in 547 ms

Bytes returned: 2 B

5

Change the encryption properties

You can set individual object properties—such as encryption at rest and storage class type—in the Amazon S3 console. Amazon S3 supports two kinds of encryption: Advanced Encryption Standard (AES)-256, and AWS Key Management Service (AWS KMS).

In the Amazon S3 console, choose the **lab1.csv** file. From the **Object actions** menu, choose **Edit server-side encryption**. Choose **Enable** and **Save changes**. To return to the object overview page, choose **Exit**.

From the **Object actions** menu, choose **Edit storage class**. Select **Intelligent-Tiering** and **Save changes**.

The screenshot shows the AWS S3 console interface. At the top, there's a navigation bar with the AWS logo, a search bar, and various global settings. Below that, the breadcrumb navigation shows: Amazon S3 > Buckets > cse4510bucket1 > lab1.csv. The main content area is titled "lab1.csv" with an "Info" link. There are three tabs: "Properties" (which is selected), "Permissions", and "Versions". On the right side, there's a "Object actions" dropdown menu with several options: "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" (this option is highlighted with a red box), "Edit metadata", "Edit tags", and "Make public using ACL". The "Properties" section contains details about the object: Owner (awslabsc0w4031887t1647838422), AWS Region (US East (N. Virginia) us-east-1), Last modified (September 10, 2022, 12:59:20 (UTC-04:00)), Size (337.0 B), Type (csv), and Key (lab1.csv). To the right of these details, there are corresponding values: S3 URI (s3://cse4510bucket1/lab1.csv), Amazon Resource Name (ARN) (arn:aws:s3:::cse4510bucket1/lab1.csv), Entity tag (Etag) (8cb3597be2d8c8774b3278562e27a9af), and Object URL (<https://cse4510bucket1.s3.amazonaws.com/lab1.csv>).

Change the encryption properties

You can set individual object properties—such as encryption at rest and storage class type—in the Amazon S3 console. Amazon S3 supports two kinds of encryption: Advanced Encryption Standard (AES)-256, and AWS Key Management Service (AWS KMS).

In the Amazon S3 console, choose the **lab1.csv** file. From the **Object actions** menu, choose **Edit server-side encryption**. Choose **Enable** and **Save changes**. To return to the object overview page, choose **Exit**.

The screenshot shows the 'Edit server-side encryption' page in the Amazon S3 console. At the top, there's a warning message about creating a copy of the object with updated settings. Below this, the 'Server-side encryption' section is visible, containing a radio button for 'Enable' which is selected and highlighted with a red box. Underneath, there are three options for 'Encryption key type': 'Amazon S3-managed keys (SSE-S3)' (selected), 'AWS Key Management Service key (SSE-KMS)', and 'AWS Lambda function ARN'.

Change the Storage Type properties

From the **Object actions** menu, choose **Edit storage class**.

Select **Intelligent-Tiering** and **Save changes**.

The screenshot shows the 'Edit storage class' page in the AWS S3 console. At the top, there is a warning message about creating a copy of the object with updated settings. Below this, a table lists various storage classes based on their designed用途 (Designed for) and availability zones.

Storage class	Designed for	Availability Zones	Min storage duration
Standard	Frequently accessed data (more than once a month)	> 3	-
Intelligent-Tiering	Data with changing or unknown access patterns	≥ 3	-
Standard-IA	Infrequently accessed data (once a month) with milliseconds access	≥ 3	30 days
One Zone-IA	Recreatable, infrequently accessed data (once a month) stored in a single Availability Zone with milliseconds access	1	30 days
Glacier Instant Retrieval	Long-lived archive data accessed once a quarter with instant retrieval in milliseconds	≥ 3	90 days
Glacier Flexible Retrieval	Long-lived archive data accessed once a year with retrieval of minutes to hours	≥ 3	90 days

Storage Pricing

From the **Object actions** menu, choose **Edit storage class**.

Select **Intelligent-Tiering** and **Save changes**.

Region: US East (Ohio) ▾		Storage pricing
S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data		
First 50 TB / Month		\$0.023 per GB
Next 450 TB / Month		\$0.022 per GB
Over 500 TB / Month		\$0.021 per GB
S3 Intelligent - Tiering* - Automatic cost savings for data with unknown or changing access patterns		
Monitoring and Automation, All Storage / Month (Objects > 128 KB)		\$0.0025 per 1,000 objects
Frequent Access Tier, First 50 TB / Month		\$0.023 per GB
Frequent Access Tier, Next 450 TB / Month		\$0.022 per GB
Frequent Access Tier, Over 500 TB / Month		\$0.021 per GB
Infrequent Access Tier, All Storage / Month		\$0.0125 per GB
Archive Instant Access Tier, All Storage / Month		\$0.004 per GB
S3 Intelligent - Tiering* - Optional asynchronous Archive Access tiers		
Archive Access Tier, All Storage / Month		\$0.0036 per GB
Deep Archive Access Tier, All Storage / Month		\$0.00099 per GB
S3 Standard - Infrequent Access** - For long lived but infrequently accessed data that needs millisecond access		
All Storage / Month		\$0.0125 per GB
S3 One Zone - Infrequent Access** - For re-createable infrequently accessed data that needs millisecond access		
All Storage / Month		\$0.01 per GB
S3 Glacier Instant Retrieval*** - For long-lived archive data accessed once a quarter with instant retrieval in milliseconds		
All Storage / Month		\$0.004 per GB
S3 Glacier Flexible Retrieval (Formerly S3 Glacier)*** - For long-term backups and archives with retrieval option from 1 minute to 12 hours		
All Storage / Month		\$0.0036 per GB
S3 Glacier Deep Archive*** - For long-term data archiving that is accessed once or twice in a year and can be restored within 12 hours		
All Storage / Month		\$0.00099 per GB

Upload Compressed Files

For big data scenarios, you should generally store compressed files in Amazon S3. Amazon S3 supports the .gzip and .bzip2 compression formats. Uploading a compressed file is essentially the same as uploading a file that is notebook compressed.

In this task, you will upload a file that is compressed as a .gzip file. First, download the compressed file from Canvas and save it to a local directory.

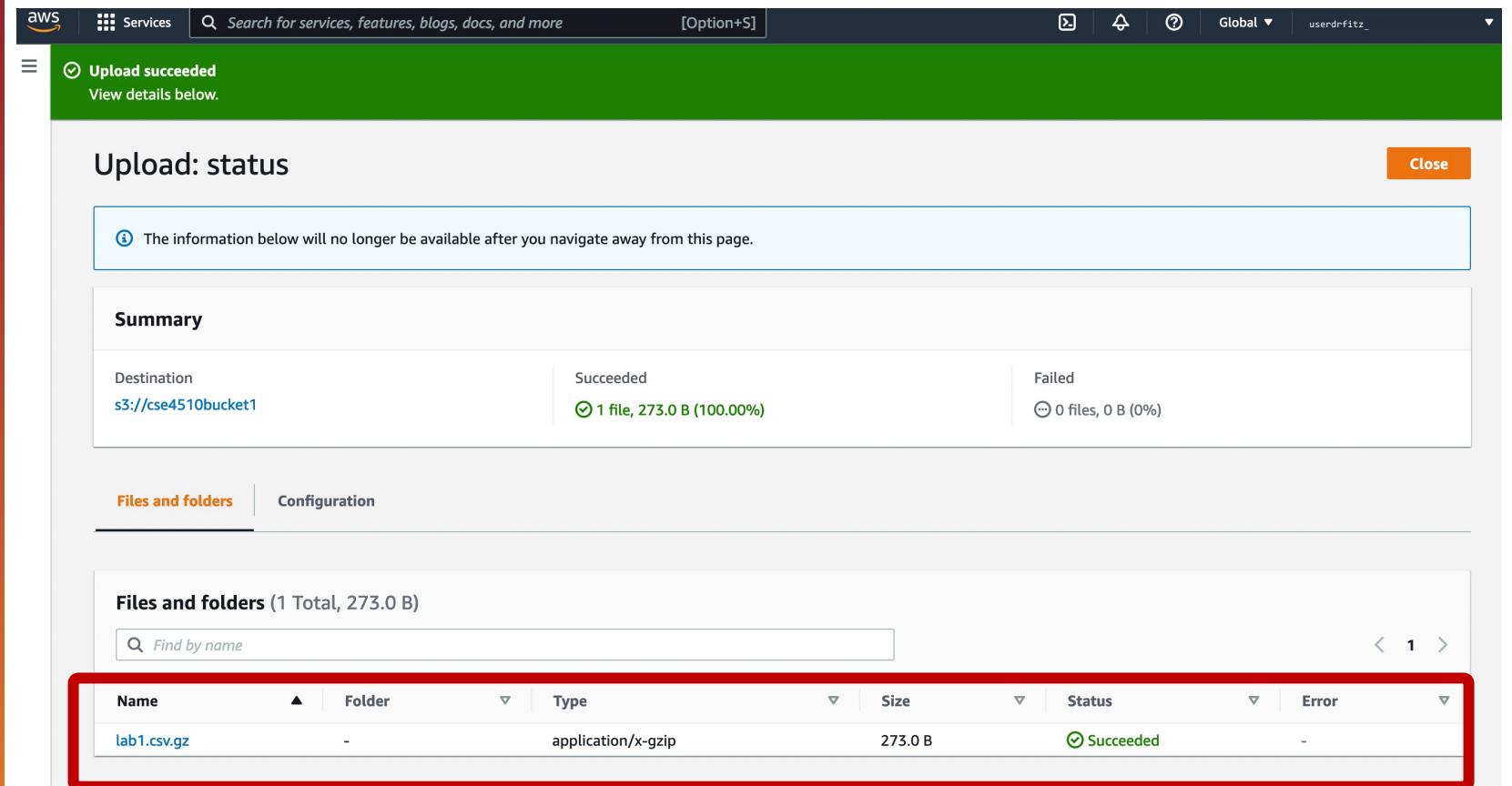
In the Amazon S3 console, choose your bucket from the breadcrumbs again.

Choose **Upload**.

Choose **Add files**, and choose the **lab1.csv.gz** file that you downloaded previously.

Choose **Upload**.

Select the **lab1.csv.gz** file.



The screenshot shows the AWS S3 console with the following details:

Upload succeeded
View details below.

Upload: status

Summary

Destination	Succeeded	Failed
s3://cse4510bucket1	1 file, 273.0 B (100.00%)	0 files, 0 B (0%)

Files and folders (1 Total, 273.0 B)

Name	Folder	Type	Size	Status	Error
lab1.csv.gz	-	application/x-gzip	273.0 B	Succeeded	-

Querying Compressed Data in S3

From the **Object actions** menu, choose **Query with S3 Select**.

Scroll down the page and choose **Run SQL query**.

You should get results that demonstrate that you can query the compressed file in the same way as a non-compressed file

The screenshot shows two sequential steps in the AWS S3 console:

Step 1: Input settings

The left sidebar shows the navigation path: Amazon S3 > Buckets > cse4510bucket1 > lab1.csv.gz > Query with S3 Select. The main area displays "Input settings" for the file "lab1.csv.gz". The "Path" is set to "s3://cse4510bucket1/lab1.csv.gz" and the "Size" is "273.0 B".

Step 2: SQL query and results

The left sidebar shows the navigation path: Amazon S3 > Buckets > cse4510bucket1 > lab1.csv.gz. The main area displays the "SQL query" section, which includes a note about the supported SELECT command and a template SQL query. A red box highlights the "Run SQL query" button. Below it, the "Query results" section shows a status message: "Successfully returned 5 records in 476 ms" and the raw query results:

```
CustomerID,First Name,Last Name,Join Date,Street Address,City,State,Phone
001,Alejandro,Rosalez,12/12/2013,123 Main St.,Baltimore,MD,765-234-2349
002,Jane,Doe,10/5/2014,456 State St.,Seattle,WA,415-889-4932
003,John,Stiles,9/20/20016,1980 8th St.,Brooklyn,NY,917-123-9308
004,Li,Juan,6/29/2011,1323 22nd Ave.,Albany,NY,917-332-3432
```

Working with S3 via Code

THIS SECTION IS TARGETED TO THOSE USING THE BIG DATA CONTAINERS COURSE WITHOUT FOLLOWING AWS ACADEMY'S GUIDED LABS.

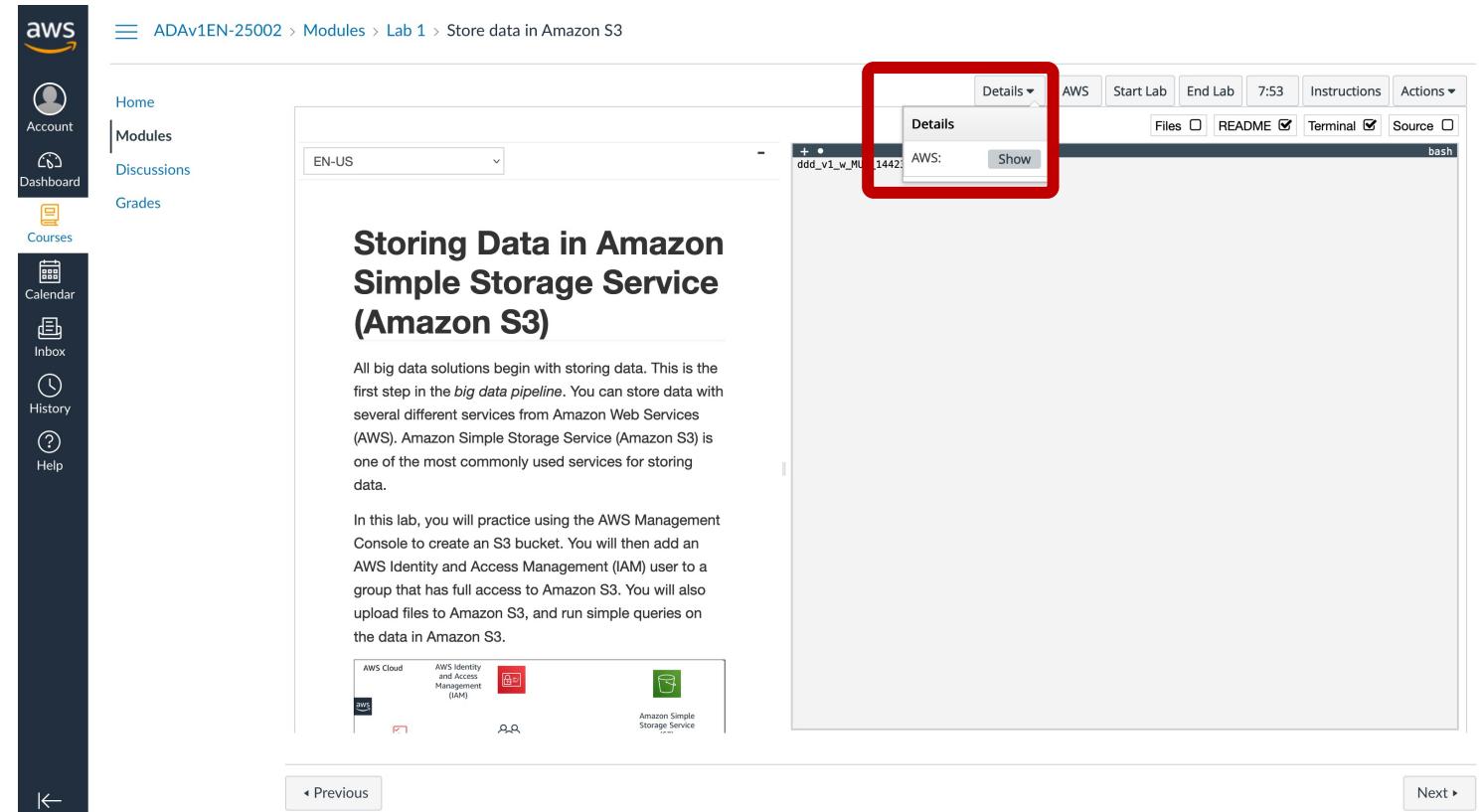
IF YOU HAVE TIME REMAINING ON YOUR GUIDED LAB, YOU MAY ALSO ACCESS THE PLATFORM VIA CODE.

Accessing Credentials & Starting a Session

First click start lab to start a lab.

Then click Details→Show to access your credentials.

If you are working with the Big Data Containers course, your screen will resemble the one on the next slide.



ADAv1EN-25002 > Modules > Lab 1 > Store data in Amazon S3

Home
Modules
Discussions
Grades
Calendar
Inbox
History
Help

EN-US

Storing Data in Amazon Simple Storage Service (Amazon S3)

All big data solutions begin with storing data. This is the first step in the *big data pipeline*. You can store data with several different services from Amazon Web Services (AWS). Amazon Simple Storage Service (Amazon S3) is one of the most commonly used services for storing data.

In this lab, you will practice using the AWS Management Console to create an S3 bucket. You will then add an AWS Identity and Access Management (IAM) user to a group that has full access to Amazon S3. You will also upload files to Amazon S3, and run simple queries on the data in Amazon S3.

AWS Cloud AWS Identity and Access Management (IAM) Amazon Simple Storage Service

◀ Previous Next ▶

Access AWS Details

Click **AWS details** to access AWS information such as keys for accessing resources using SSH, etc.

Click Show to see your credentials for using the AWS CLI.

The screenshot shows the AWS Learner Lab interface. At the top, there's a navigation bar with the AWS logo, the path 'ALLFv1-1... > Modules > Learner La... > Learner Lab - Foundational Services', and various status indicators like 'Used \$0 of \$100, Jan, 2022' and '05:54'. Below the navigation is a sidebar with links: Home, Modules (which is selected), Discussions, Courses, Calendar, Inbox, History, and Help. The main content area has a terminal window showing a session ID: 'ddd_v1_w_s4s_1015462@runweb46979:~\$'. To the right of the terminal is a 'Cloud Access' section. This section includes a 'AWS CLI' button with a 'Show' option, a 'Cloud Labs' section with session details (remaining time 05:55:35, started at 2022-01-18T08:23:42-0800, ended at 2022-01-18T14:23:42-0800), and an 'Accumulated lab time: 00:04:25 (5 minutes)'. It also shows 'No running instance'. Below these are buttons for 'SSH key' (Show, Download PEM, Download PPK) and 'AWS SSO' (Download URL). A table at the bottom lists 'AWSAccountid' (573703604721) and 'Region' (us-east-1). A red circle highlights the 'AWS Details' button in the top right corner of the main content area.



Set up Credentials

Create a file named **credentials** inside the folder
`~/.aws`

Run the following from your Terminal/cmd to
create the file

`cd ~ or cd %USERPROFILE% for Windows`

`mkdir .aws`

`cd .aws`

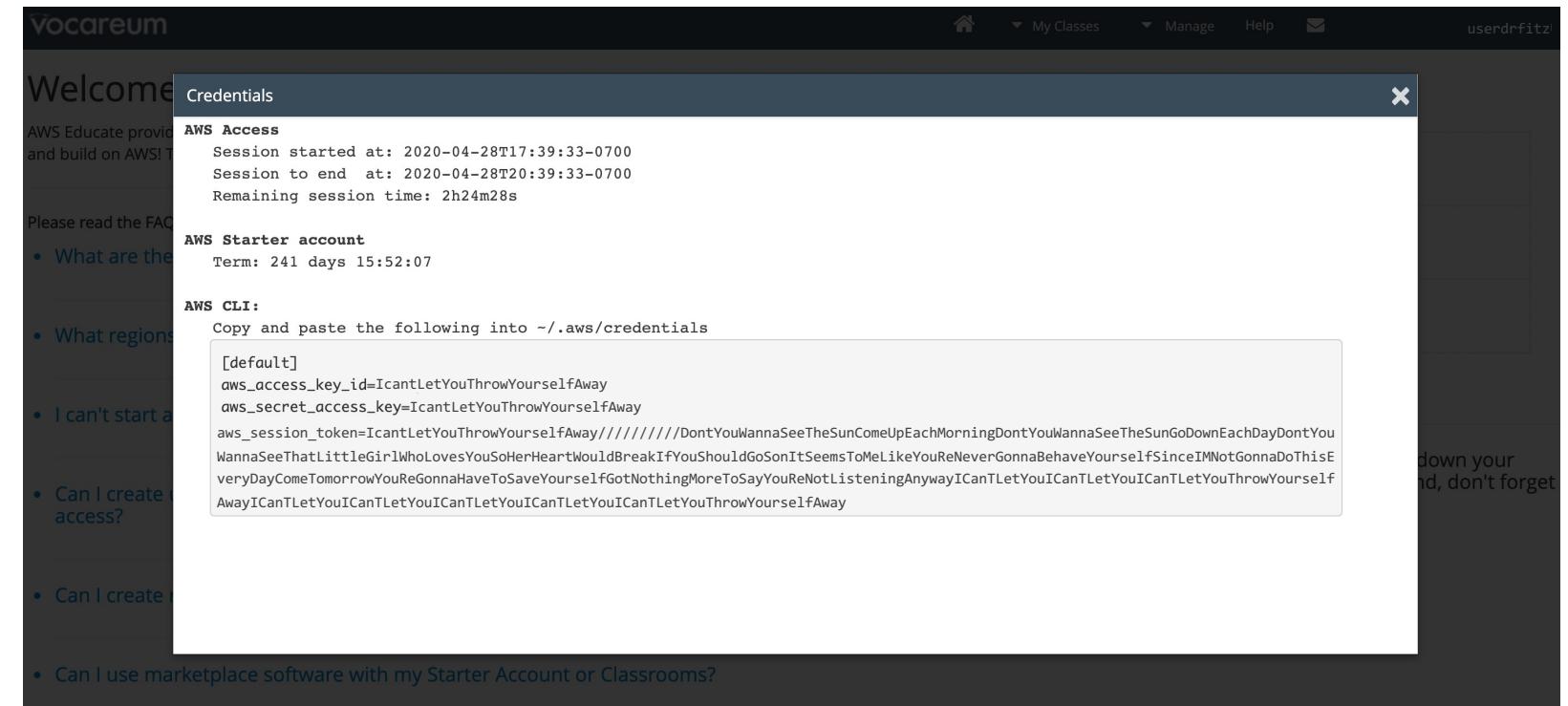
`sudo nano credentials`

Windows users:

Do not use sudo

To get nano, Install the Chocolatey
(<https://chocolatey.org/>) package manager and
enter

`choco install nano`



The screenshot shows a "Credentials" dialog box on the Vocareum platform. It displays session information: "Session started at: 2020-04-28T17:39:33-0700", "Session to end at: 2020-04-28T20:39:33-0700", and "Remaining session time: 2h24m28s". Below this, it shows an "AWS Starter account" with a term of "241 days 15:52:07". The "AWS CLI:" section contains a configuration snippet for the "[default]" profile, which includes the access key ID and secret access key, both set to a string of random characters. The secret access key is explicitly noted as being very long and random. At the bottom of the dialog, there is a link: "Can I use marketplace software with my Starter Account or Classrooms?".

Keep track of the session time and
Refresh temporary credentials five minutes
before their expiration.

Set up Credentials & Starting a Session

You may also store your credentials in a file or hard-code them (not recommended).

You may also set up multiple sets of credentials (profiles) in the same file and specify the profile name when connecting to s3.

Approach 2 – Hard-code or read from a text file

```
import boto3 session = boto3.Session( 's3',
aws_access_key_id=ACCESS_KEY,
aws_secret_access_key=SECRET_KEY,
aws_session_token=SESSION_TOKEN )
```

Approach 3 – Specify multiple accounts in a Shared Credentials File

[default]

```
aws_access_key_id=foo
aws_secret_access_key=bar
```

[dev]

```
aws_access_key_id=foo2
aws_secret_access_key=bar2
```

[prod]

```
aws_access_key_id=foo3
aws_secret_access_key=bar3
```

import boto3

```
session = boto3.Session(profile_name='dev')
dev_s3_client = session.client('s3')
```

```
dynamodb = session.resource('s3', region_name='us-east-1' )
```

Working with S3 Buckets via Code

Use the following examples to access your S3 Bucket via Python code.

```
import boto3
# Creating the low level functional client
client = boto3.client(
    's3',
    aws_access_key_id='AKIA46SFIWN5AMWMDQVB',
    aws_secret_access_key='yuHNxlcbEx7b9Vs6QEo2KWiaAPxj/k6RdEY4DfeS',
    region_name='ap-south-1'
)
```

```
# Creating the high level object oriented interface
```

```
resource = boto3.resource(
    's3',
    aws_access_key_id='AKIA46SFIWN5AMWMDQVB',
    aws_secret_access_key='yuHNxlcbEx7b9Vs6QEo2KWiaAPxj/k6RdEY4DfeS',
    region_name='ap-south-1'
)
```

<https://boto3.amazonaws.com/v1/documentation/api/latest/guide/s3-examples.html>

Working with S3 Buckets via Code

Use the following examples to access your S3 Bucket via Python code.

```
# Create the S3 object
obj = client.get_object(
    Bucket='sql-server-shack-demo-1',
    Key='sql-shack-demo.csv'
)

# Read data from the S3 object
data = pandas.read_csv(obj['Body'])

# Print the data frame
print('Printing the data frame...')
print(data)
```

<https://boto3.amazonaws.com/v1/documentation/api/latest/guide/s3-examples.html>

References

The following references were used to create this tutorial.

- AWS Academy
- <https://boto3.amazonaws.com/v1/documentation/api/latest/guide/s3-examples.html>