**Data Bricks On AWS**

[Data Bricks SignUp Link](https://accounts.cloud.databricks.com/login)

Keep the AWS account Open and finish the Sign Up Registration process.

Use Manual process to create a Databricks workspace

Click on Create workspace

A close-up of a computer screen

Description automatically generated

Use the following instructions to create a workspace with Manual.

1. In your account console, click the Workspaces tile.

2. Click the Create workspace dropdown and select Manual.

3. On the Let’s set up your workspace page, enter the following:

• A user-friendly Workspace name.

• The AWS region you want to host your Databricks workspace in (typically, choose the region closest to your primary work location). In this case choose **us-east-2**

A screenshot of a computer

Description automatically generated

There are four steps in creating the workspace,

* General
* Storage
* Credentials
* Summary

1. In General give the workspace name and select the region. Click on Next

A screenshot of a computer

Description automatically generated

2. In Storge step choose Create Storage Configuration give a storage configuration name.

A screenshot of a computer

Description automatically generated

You will be given access to an AWS account that's already set up with two IAM roles specifically designed for Databricks deployment and storage configurations. The roles are named **"Workspace-Deployment-Role"** and **"Workspace-Storage-Role"** respectively. Additionally, you'll be provided with an S3 bucket that begins with the **AWS account ID given to you**. This bucket will serve as the primary storage location for your data assets. These resources have been configured to facilitate your workflow seamlessly within the Databricks environment

Go to AWS console copy the S3 Bucket name and **Workspace-Storage-Role** IAM role

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Give the above copied bucket name, storage role ARN in Bucket and Role fields of storage settings.

A screenshot of a computer

Description automatically generated

Click on create storage configuration and continue.

3. For Databricks to launch clusters in your AWS account, you will need to create a cross-account IAM role that gives access to Databricks. You are responsible for the AWS cost of clusters you create. Select a credential you created previously, or a create a new one.

Choose Add new credentials and give a name to it.

In AWS IAM Console, Copy the **Workspace-Deployment-Role** ARN and give it in Role ARN field

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Click on Create credentials and continue

A close-up of a logo

Description automatically generated

Before moving on to 4th step, you are supposed to edit the Deployment, Storage IAM trust relationship policy and S3 bucket permissions policy.

Copy the account ID from your Databricks account.

A screenshot of a computer

Description automatically generated

Go to AWS IAM console, select and open **"Workspace-Deployment-Role".** Select the Trust relationships section and change the ExternalId with your Databricks ID

A screenshot of a computer

Description automatically generated

Similarly, do it for the **"Workspace-Storage-Role".** Select the Trust relationships section and change the ExternalId with your Databricks ID and update the Principal AWS section as well.

In Principal AWS section include the **"Workspace-Deployment-Role"** Arn as well

A screenshot of a computer

Description automatically generated

Go to AWS S3 service, select the Bucket. Choose the permissions section go to Bucket policy and update the "aws:PrincipalTag/DatabricksAccountId" field in the policy with your Databricks account ID.

A close up of a card

Description automatically generated

A screenshot of a computer

Description automatically generated

4. Go the summary section, preview the workspace name, region, storage and credentials configurations. Click on Create workspace.

A screenshot of a computer

Description automatically generated

In Data Bricks you will be seeing the WorkSpace getting created and it will be in running state in less than 2 minutes

A screenshot of a computer

Description automatically generated

A close-up of a computer screen

Description automatically generated

Open the URL to Access the workspace.

A screenshot of a computer

Description automatically generated

You will also be receiving the URL to your registered Email

A close-up of a computer screen

Description automatically generated

**In Case Roles got created and Policies are not attached to the role follow these steps.**

**Deployment-Role-Policies**

**Deployment-Policy-1**

1. Navigate to the IAM Console:

In the AWS Management Console, in the search bar, type "IAM" and select IAM.

A screenshot of a computer

Description automatically generated

1. In the IAM Dashboard, in the navigation pane on the left, select **Policies**.

 Click the **Create policy** button.

A screenshot of a computer

Description automatically generated

1. You will see two tabs: **Visual editor** and **JSON**. Choose JSON

A screenshot of a computer

Description automatically generated

You can directly input your policy in JSON format.

Copy the Json Policy here: [Deployment-Policy-1](https://handson-x-learn.s3.ap-south-1.amazonaws.com/DataBricks/CustomerManagedPolicies/Deployment-Policy-1.json) paste it in the Policy editor

1. Click Next: **Tags**. You can add tags if needed (optional).
2. Click Next: Review
3. Provide a Name “Deployment-Policy-1” and an optional Description for the policy

A screenshot of a computer

Description automatically generated

1. Review the policy summary and click **Create policy**.

A screenshot of a computer

Description automatically generated

**Deployment-Policy-2**

1. Create one more policy for Deployment role
2. Before creating the policy copy the Bucket ARN and Workspace-Deployment-Role ARN

Go to the S3 bucket Copy the Bucket ARN

A screenshot of a computer

Description automatically generated

Similarly go to the IAM and copy the Workspace-Deployment-Role ARN

A screenshot of a computer

Description automatically generated

1. Click the **Create policy** button.
2. You will see two tabs: **Visual editor** and **JSON**. Choose JSON
3. Copy the Json Policy here: [Deployment-Policy-2](https://handson-x-learn.s3.ap-south-1.amazonaws.com/DataBricks/CustomerManagedPolicies/Deployment-Policy-2.json) paste it in the Policy editor

In the JSON Resources field give the above copied ARN’s.

Bucket ARN Resources fields

A screen shot of a computer

Description automatically generated

Workspace-Deployment-Role ARN Resource field.

A screenshot of a computer

Description automatically generated

1. Click Next: **Tags**. You can add tags if needed (optional).
2. Click Next: Review
3. Provide a Name “Deployment-Policy-2” and an optional Description for the policy
4. Review the policy summary and click **Create policy**.

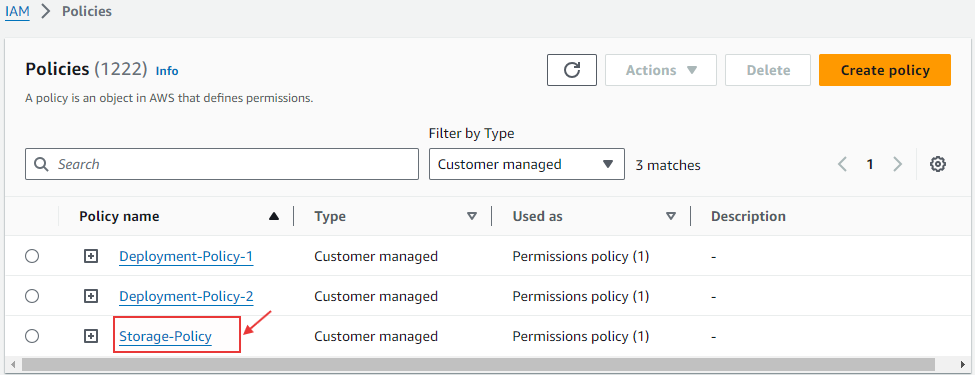
A screenshot of a computer

Description automatically generated

**Storage-Policy**

**Storage-Policy**

1. Similarly Create Policy for Storage Role
2. For Storage-Policy use the same policy created for ‘Deployment-Policy-2’, copy the JSON policy.
3. Click on Create policy, choose the JSON editor, paste the copied policy.
4. Click Next: **Tags**. You can add tags if needed (optional).
5. Click Next: Review
6. Provide a Name “Storage-Policy” and an optional Description for the policy
7. Review the policy summary and click **Create policy**.



**Attaching policies to roles**

1. Click the “Workspace-Deployment-Role” role to attach the policy.

**A screenshot of a computer

Description automatically generated**

1. In the **Permissions** tab, click the **Add permissions** button and select **Attach policies**.
2. In the list of policies, find and select the policy you created by checking the box next to its name.

A screenshot of a computer

Description automatically generated

1. Click **Add Permission**. Now the polices are attached to the “Workspace-Deployment-Role”
2. Similarly attach the policy ‘Storage-Policy’ to “Workspace-Storage-Role”