

O'Reilly – Databricks Fundamentals Bootcamp

by Mohit Batra

Lab 1 – Option 1: Setting up Azure Databricks Workspace & Cluster

Contents

Exercise 1 – Setup Azure Databricks Workspace	2
Exercise 2 – Launch Databricks Workspace & Create Cluster.....	2

Exercise 1 – Setup Azure Databricks Workspace

1. Go to Azure portal (portal.azure.com).
2. In the search bar, search for Azure Databricks. And select it.
3. Click on **Create**.
4. Fill up the properties to create account
 - a. [Basics Tab]
 - i. Subscription: Select your subscription
 - ii. Resource group: Click **Create new**, provide a name (example – O'Reilly), and click Ok
 - iii. Workspace name: Provide a unique name
 - iv. Region: Select region of your choice (example – East US 2)
 - v. Pricing tier: Premium
 - vi. Click **Review + Create**.

Home > Azure Databricks >

Create an Azure Databricks workspace

Basics Networking Encryption Security & compliance Tags Review + create

Project Details

Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * ⓘ MSDN Platforms

Resource group * ⓘ O'Reilly
[Create new](#)

Instance Details

Workspace name * MyDatabricksWorkspace ✓

Region * East US 2

Pricing Tier * ⓘ Premium (+ Role-based access controls)

Managed Resource Group name Enter name for managed resource group

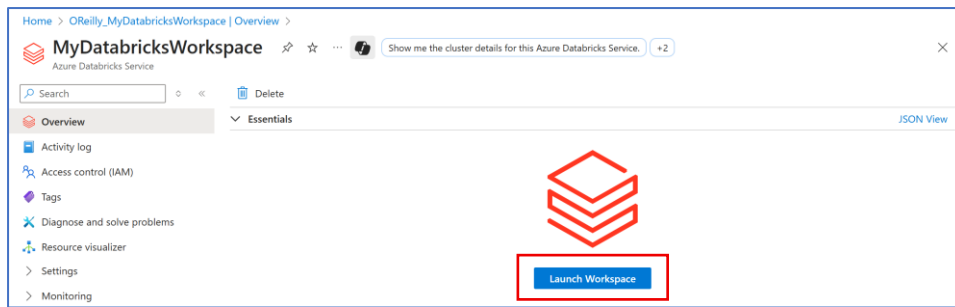
[Review + create](#) < Previous Next : Networking >

Notification: We selected the recommended pricing tier for your workspace. You can change the tier based on your needs.

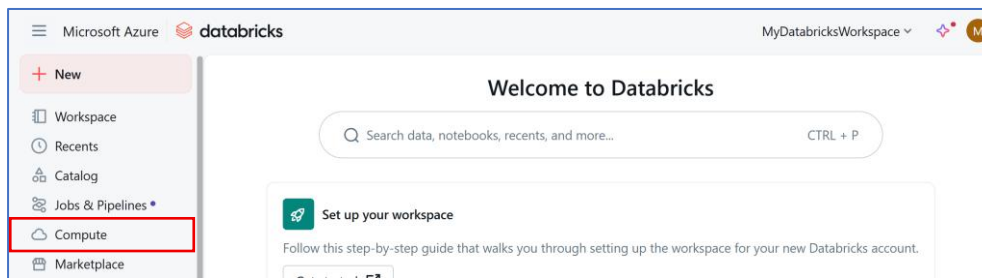
- b. Click **Create**.
 - This will take a few minutes to deploy.

Exercise 2 – Launch Databricks Workspace & Create Cluster

1. Open Azure Databricks instance created in the previous step.
2. Click on Launch workspace, to open Databricks UI.



3. In the workspace, from left pane, go to **Compute** tab.



4. Under **All-purpose compute**, Click on **Create Compute** to create a cluster.

5. Fill up cluster properties and click on **Create**. This will take few minutes to setup a single node cluster.
 - a. Compute name: Demo Cluster
 - b. Databricks Runtime: Select the latest runtime with LTS (long-term support).
 - c. Photon acceleration: Disable
 - d. Node type: Standard_DS3_v2 (if this type is not available, select any other)
 - e. Single node: Enable
 - f. Terminate minutes: 30 minutes

 This screenshot shows the 'Create new compute' form in the Databricks interface. The form is divided into several sections:

- General**: Contains 'Compute name' (set to 'Demo Cluster') and 'Policy' (set to 'Unrestricted').
- Performance**: Contains several settings:
 - 'Machine learning' is disabled.
 - 'Databricks runtime' is set to '16.4 LTS (Scala 2.13)'.
 - 'Photon acceleration' is disabled.
 - 'Node type' is set to 'Standard_DS3_v2'.
 - 'Single node' is checked.
 - 'Terminate after' is set to '30 minutes of inactivity'.
- Tags**: A section for adding tags, currently empty.

 At the bottom of the form, there are 'Create' and 'Cancel' buttons.

[Note]: If you want to setup multi-node cluster, deselect single node option from UI.

6. Once cluster is ready, it will show up as **running** in the list of clusters.

New

Workspace

Recents

Catalog

Jobs & Pipelines

Compute

Marketplace

Compute

All-purpose computeJob computeSQL warehousesPoolsPoliciesAppsOLTP Database

Filter compute you ...Created byOnly pinnedCreate with Personal ComputeCreate compute

State	Name	Policy	Runtime	Active ...	Active c...	Active ...	Source	Creator	Notebo...
	Demo Cluster	-	16.4	14 GB	4 cores	0.75	UI	Mohit Batra	-