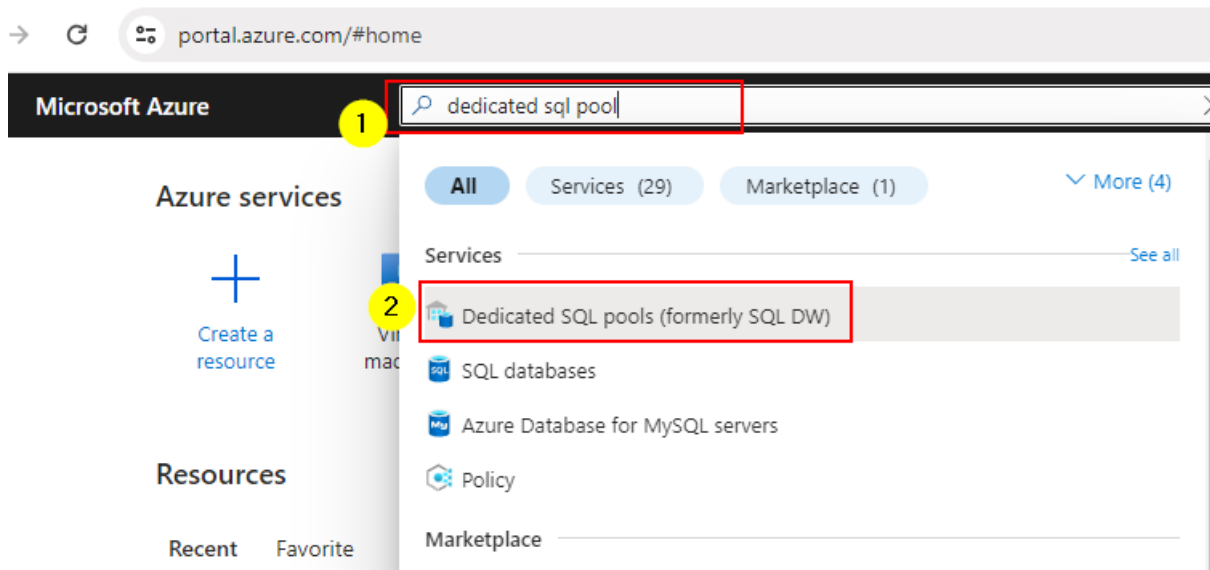


Setup Dedicated SQL Pool in Azure

1. Open Azure Portal (<https://portal.azure.com/#home>)
2. Search dedicated SQL pool in search bar



3. On dedicated SQL pool page click create
4. Create a new resource group: etl-rg
5. SQL pool name: dwh34x
6. Create server
 - a. Server Name: dwhserverxxxx (replace xxxx with random number)
 - b. Location: Central India
 - c. Authentication method: Use SQL authentication
 - i. Server admin login: dwhadmin
 - ii. Password: demo!pass123
 - iii. Confirm password: demo!pass123

[Home](#) > [Dedicated SQL pools \(formerly SQL DW\)](#) > [Create dedicated SQL pool \(formerly SQL DW\)](#) >

Create SQL Database Server

Microsoft

Server details

Enter required settings for this server, including providing a name and location. This server will be created in the same subscription and resource group as your database.

Server name *

dwhserver34x

.database.windows.net

Location *

(Asia Pacific) Central India

Authentication

Authentication method

☐ Use Microsoft Entra-only authentication
☐ Use both SQL and Microsoft Entra authentication
☒ Use SQL authentication

Server admin login *

dwhadmin

Password *

.....

Confirm password *

.....

d. Click Ok

7. Performance level: DW100c

Create dedicated SQL pool (formerly SQL DW)

Microsoft
manage all your resources.

Subscription * ⓘ MSDN Platforms Subscription

Resource group * ⓘ 1 (New) etl-rg
[Create new](#)

SQL pool details

Enter required settings for this SQL pool, including picking a logical server and configuration.

SQL pool name * 2 dwh34x

Server ⓘ 3 (new) dwhsrver34x (Central India)
[Create new](#)

Performance level * ⓘ 4 Gen2
DW1000c
[Select performance level](#)

Configure performance
Configure your performance level that best fits your needs.

Gen2
Offers the highest performance and storage scalability options for intensive workloads.
Starting at 143.59 INR / hour

5 Learn more about [performance level](#) ⓘ

Scale your system ⓘ

DW100c

Price of your system
143.59 INR / hour
100 cDWU

8. Click Next: Networking

9. In Networking SQL Tab:

- a. Connectivity method: Public endpoint
- b. Firewall rule:
 - i. Allow Azure services and resources to access this server: Yes
 - ii. Add current client IP address: Yes

Network connectivity

Choose an option for configuring connectivity to your server via public endpoint or private endpoint. You can create with defaults and you can configure connection method after server creation. [Learn more](#)

Connectivity method * ⓘ

1

☐

No access

☒

Public endpoint

☐

Private endpoint

Firewall rules

Setting 'Allow Azure services and resources to access this server' to Yes allows communication from the Azure boundary, that may or may not be part of your subscription. [Learn more](#)

Setting 'Add current client IP address' to Yes will add an entry for your client IP address to the firewall rules.

2

Allow Azure services and resources to access this server *

No

Yes

Add current client IP address *

No

Yes

[Review + create](#)

[< Previous](#)

[Next : Additional settings >](#)

10. Click Next: Additional Setting

11. On Additional setting tab page

a. Use existing data: Sample

Create dedicated SQL pool (formerly SQL DW) ...

Microsoft

[* Basics](#) [* Networking](#) [* Additional settings](#) [Tags](#) [Review + create](#)

Customize additional configuration parameters including collation & data source.

Data source

Start with a blank SQL pool, restore from a backup or select sample data to populate your new database.

Use existing data *

None

Backup

Sample

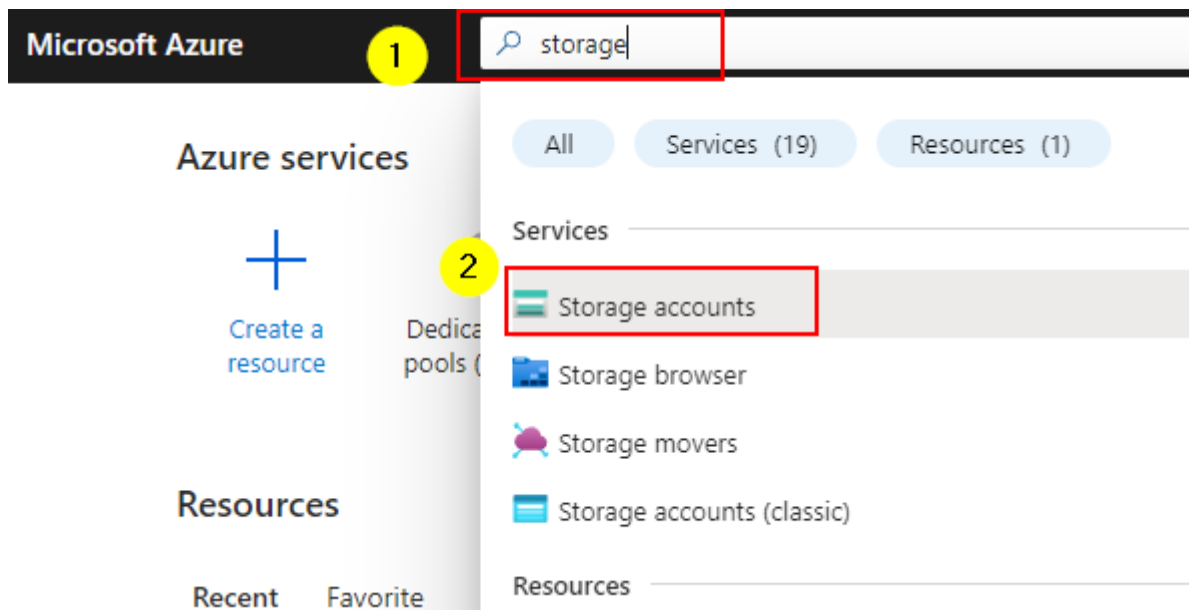
AdventureWorksDW will be created as the sample database.

12. Click Review + Create and then Click Create

13. Wait for the deployment to complete.

Setup Data Lake in Azure

1. Go to Azure home page (<https://portal.azure.com/#home>)
2. In the search bar search storage and click Storage accounts



3. On storage account page click **+ Create**
4. On **Create a storage account** page in the **Basics** tab, provide the following settings:
 - Under the project details, specify the following settings:
 - **Subscription:** the name of the subscription you are using in this lab
 - **Resource group:** etl-rg.
 - Under the Instance details, specify the following settings:
 - **Storage account name:** datalakexx, where xx are your initials.
 - **Region:** the name of the Azure region which is closest to the lab location. (Central India)
 - **Performance:** Standard.
 - **Redundancy:** Locally-redundant storage (LRS)

Microsoft Azure

Search resources, services, and docs (G+/)

[Home](#) > [Create a resource](#) > [Storage account](#) >

Create a storage account

Basics

Advanced

Networking

Data protection

Encryption

Tags

Review + create

Project details

Select the subscription in which to create the new storage account. Choose a new or existing resource group to organize and manage your storage account together with other resources.

Subscription *

Azure Pass - Sponsorship

Resource group *

kpmgadfpc

[Create new](#)

Instance details

If you need to create a legacy storage account type, please click [here](#).

Storage account name ⓘ *

datastoragepc1

Region ⓘ *

(US) East US

Performance ⓘ *

☒ Standard: Recommended for most scenarios (general-purpose v2 account)

Review + create

< Previous

Next : Advanced >

- Select **Advanced** tab in create storage account
 - Under Data Lake Storage Gen2 details
 - Enable **hierarchical namespace**: mark is Select

Microsoft Azure

Search resources, services, and docs (G+/)

[Home](#) > [Create a resource](#) > [Storage account](#) >

Create a storage account

Basics

Advanced

Networking

Data protection

Encryption

Tags

Review + create

Minimum TLS version ⓘ

Version 1.2

Data Lake Storage Gen2

The Data Lake Storage Gen2 hierarchical namespace accelerates big data analytics workloads and enables file-level access control lists (ACLs). [Learn more](#)

Enable hierarchical namespace ☒


SSH File Transfer Protocol (SFTP)

5. In the **Create storage account** screen, click **Review + create**.



6. After the validation of the **Create storage account*** screen, click **Create**.

Create a storage account ...

 Validation passed

Basics Advanced Networking Data protection Encryption Tags Review + create

Basics

Subscription	Azure Pass - Sponsorship
Resource Group	kpmgadfpc
Location	eastus
Storage account name	datastoragepc1
Deployment model	Resource manager
Performance	Standard
Replication	Locally-redundant storage (LRS)

Advanced

Secure transfer	Enabled
Allow storage account key access	Enabled
Allow cross-tenant replication	Disabled
Default to Azure Active Directory authorization in the Azure portal	Disabled
Blob public access	Enabled

Create < Previous Next > [Download a template for automation](#)

7. **Note:** The creation of the storage account will take approximately 90 seconds while it provisions the disks and the configuration of the disks as per the settings you have defined.

8. Once the storage account created create click on **Go to resource**, under **Data storage** blade click on **Containers** then click on **+ Container** and give name **data** then click on Create button.

The screenshot displays the Azure portal interface for a storage account named 'datastoragepc1'. The left-hand navigation pane shows the 'Data storage' section expanded, with 'Containers' selected. The main pane shows the 'Containers' view for 'datastoragepc1', featuring a '+ Container' button. A 'New container' dialog box is open on the right, with the name 'data' entered in the 'Name *' field and 'Private (no anonymous access)' selected for the 'Public access level'. The 'Create' button at the bottom of the dialog is highlighted.

Storage accounts

Default Directory

+ Create Manage view

Filter for any field...

Name ↑

- datablobpc
- datastoragepc1**
- sqlvmqjzombfvha5w

datastoragepc1 | Containers

Storage account

+ Container Change access level Restore containers Refresh Delete

Search containers by prefix

Name	Last modified	Public access level
<input type="checkbox"/> \$logs	3/1/2022, 1:28:34 PM	Private

New container

Name *
data

Public access level
Private (no anonymous access)

Advanced

Create Discard

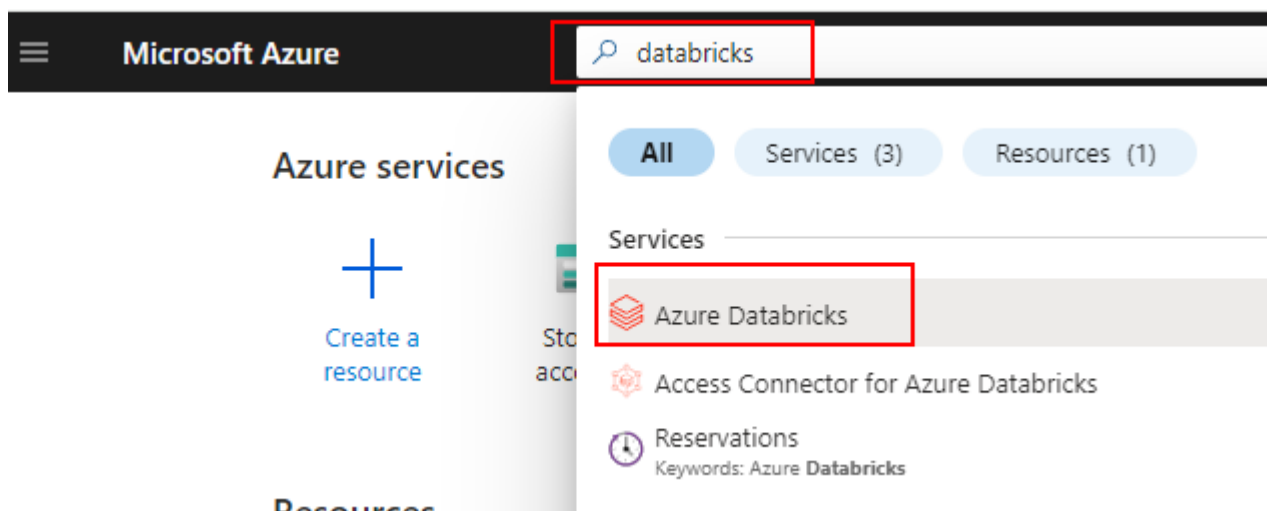
Setup Azure Databricks

Tasks:

1. Create and Configure Azure Databricks Instance.
2. Create a Spark Cluster in Azure Databricks.

Task 1: Create and configure an Azure Databricks instance.

1. In the Azure portal, at the top left search for databricks, click on Azure Databricks.



2. In the **Azure Databricks** page, click **+ Create**.
3. In the **Create Azure Databricks workspace** page, Under Basic tab provide the following settings:
 - **Subscription:** the name of the subscription you are using in this lab
 - **Resource group:** **adbwk-rg**, where **xx** are your initials.
 - **Workspace name:** **adbwkxx**, where **xx** are your initials.
 - **Region:** the name of the Azure region which is closest to you. (Central India)
 - **Pricing Tier:** **Standard (Apache Spark, Secure with Microsoft Entra ID)**.

Microsoft Azure

Search resources, services, and docs (G+)

Home > Azure Databricks >

Create an Azure Databricks workspace

Basics Networking Advanced 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 * ⓘ Azure Pass - Sponsorship

Resource group * ⓘ kpmgadfpc
[Create new](#)

Instance Details

Workspace name * adbwkpc

Region * West US

Pricing Tier * ⓘ Standard (Apache Spark, Secure with Azure AD)

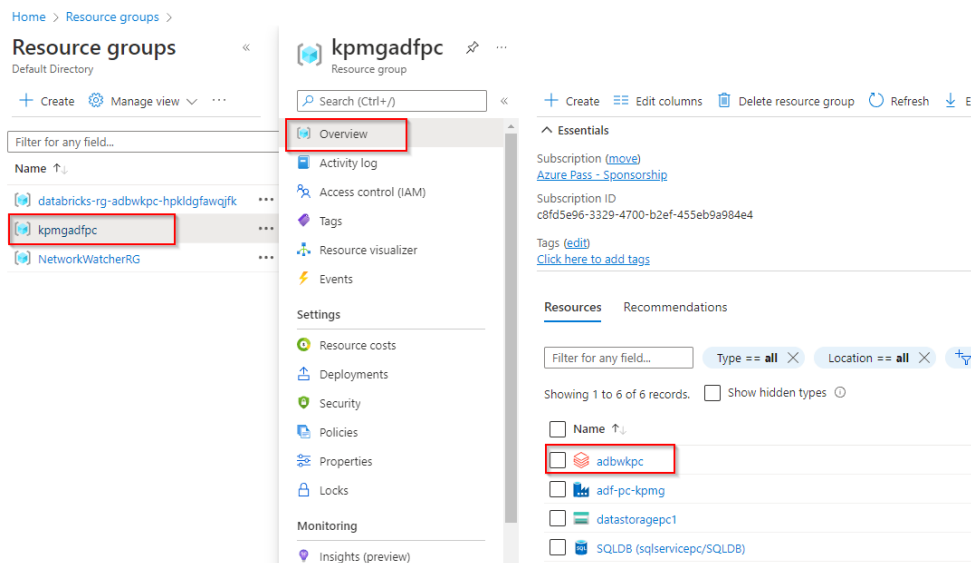
Review + create < Previous Next : Networking >

4. In the **Create Azure Databricks workspace** page, click **Review +create**. Then click **Create**

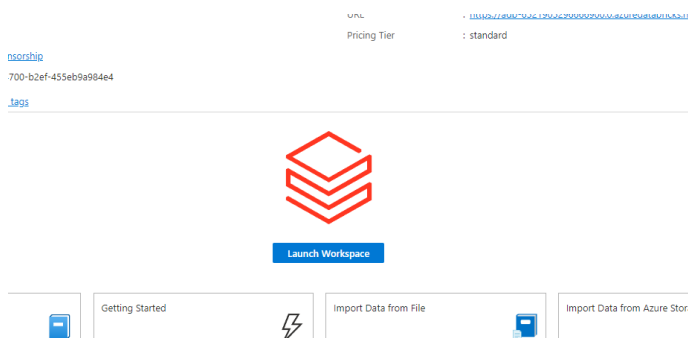
Note: The provision will take approximately 3 minutes. The Databricks Runtime is built on top of Apache Spark and is natively built for the Azure cloud. Azure Databricks completely abstracts out the infrastructure complexity and the need for specialized expertise to set up and configure your data infrastructure. For data engineers, who care about the performance of production jobs, Azure Databricks provides a Spark engine that is faster and performant through various optimizations at the I/O layer and processing layer (Databricks I/O).

5. Confirm that the Azure Databricks service has been created.
6. In the Azure portal, navigate to the **Resource group** screen.
7. In the Resource groups screen, click on the **adbwk-rg** resource group, where **xx** are your initials.

8. In the **adbwk-rg** screen, click **adbwkxx**, where **xx** are your initials to open Azure Databricks. This will open your Azure Databricks service.



9. Click on **adbwkxx** and click on **Launch Workspace**



Task 2: Create a Spark Cluster in Azure Databricks.

1. Once you Launch the Workspace.
2. Under **Compute Tab**, click **Create Cluster**.
3. In the **Create Cluster** screen, under New Cluster, create a Databricks Cluster with the following settings, and then click on **Create Cluster**:
 - **Cluster name:** My Cluster
 - **Cluster Mode:** Single Node
 - **Databricks Runtime Version:** Runtime: 11.3 LTS (Scala 2.12, Spark 3.1.2)
 - Make sure you select and set the **Terminate after 30** minutes of inactivity check box. If the cluster isn't being used, provide a duration (in minutes) to terminate the cluster.
 - Leave all the remaining options to their current settings.

Microsoft Azure | databricks | Search

Clusters / New Compute | UI Preview | Provide feedback

Pankaj Choudhary's Cluster ✎

☐ Multi node ☒ Single node

Access mode ⓘ Single user access ⓘ

Single user | ▾ Pankaj Choudhary (pankajacks@...) ▾

Performance

Databricks runtime version ⓘ

Runtime: 11.3 LTS (Scala 2.12, Spark 3.3.0) ▾

☐ Use Photon Acceleration ⓘ

Node type ⓘ

Standard_DS3_v2 14 GB Memory, 4 Cores ▾ ⓘ

☒ Terminate after 30 minutes of inactivity ⓘ

Tags ⓘ

4. In the **Create Cluster** screen, click on **Create Cluster** and leave the browser screen open.

Note: The creation of the Azure Databricks Spark cluster instance will take approximately 5-8 minutes.