

Create a Databricks workspace

Task 1: Provision Azure Databricks

Azure Databricks is an Apache Spark-based analytics platform optimized for Azure. It will be used in this lab to build and train a machine learning model used to predict flight delays.

1. In the Azure Portal (<https://portal.azure.com>), select + Create a resource, then type "Azure Databricks" into the search bar. Select Azure Databricks from the results.
2. Select Create on the bottom of the blade that follows.
3. Set the following configuration on the Azure Databricks Service creation form:
 - a. Name: Enter a unique name as indicated by a green checkmark.
 - b. Subscription: Select the subscription you are using for this hands-on lab.
 - c. Resource Group: Select the same resource group you created at the beginning of this lab.
 - d. Location: Select a region close to you. (If you are using an Azure Pass, select South Central US.)
 - e. Pricing: Select Premium.
4. Select Create to finish and submit.

Task 2: Create a cluster

1. In the left-hand menu of your Databricks workspace, select Clusters.
2. Select Create Cluster to add a new cluster.
3. Enter a name for your cluster. Use your name or initials to easily differentiate your cluster from your coworkers.
4. Select the Databricks RuntimeVersion. We recommend the latest runtime
5. Specify your cluster configuration.
 - a. For clusters that are created on a Community Edition, the default values are sufficient for the remaining fields.
 - b. For all other environments, refer to your company's policy on creating and using clusters.
6. Select Create Cluster.