Create and Explore an Azure Machine Learning Workspace

In this exercise, you will create and explore an Azure Machine Learning workspace.

Create an Azure Machine Learning workspace

As its name suggests, a workspace is a centralized place to manage all of the Azure ML assets you need to work on a machine learning project.

1. In the Azure portal, create a new **Machine Learning** resource, specifying the following settings:

   - **Subscription**: *Your Azure subscription*
   - **Resource group**: *Create or select a resource group*
   - **Workspace name**: *Enter a unique name for your workspace*
   - **Region**: *Select the geographical region closest to you*
   - **Storage account**: *Note the default new storage account that will be created for your workspace*
   - **Key vault**: *Note the default new key vault that will be created for your workspace*
   - **Application insights**: *Note the default new application insights resource that will be created for your workspace*
   - **Container registry**: None (*one will be created automatically the first time you deploy a model to a container*)

   > **Note**: When you create an Azure Machine Learning workspace, you can use some advanced options to restrict access through a *private endpoint* and specify custom keys for data encryption. We won't use these options in this exercise - but you should be aware of them!

2. When the workspace and its associated resources have been created, view the workspace in the portal.

Explore Azure Machine Learning studio

You can manage some workspace assets in the Azure portal, but for data scientists, this tool contains lots of irrelevant information and links that relate to managing general Azure resources. *Azure Machine Learning studio* provides a dedicated web portal for working with your workspace.

1. In the Azure portal blade for your Azure Machine Learning workspace, click the link to launch studio; or alternatively, in a new browser tab, open https://ml.azure.com. If prompted, sign in using the Microsoft account you used in the previous task and select your Azure subscription and workspace.

   > **Tip** If you have multiple Azure subscriptions, you need to choose the Azure *directory* in which the subscription is defined; then choose the subscription, and finally choose the workspace.

2. View the Azure Machine Learning studio interface for your workspace - you can manage all of the assets in your workspace from here.

3. In Azure Machine Learning studio, toggle the ☰ icon at the top left to show and hide the various pages in the interface. You can use these pages to manage the resources in your workspace.

Create a compute instance

One of the benefits of Azure Machine Learning is the ability to create cloud-based compute on which you can run experiments and training scripts at scale.

1. In Azure Machine Learning studio, view the **Compute** page. This is where you'll manage compute resources for your data science activities. There are four kinds of compute resource you can create:

   - **Compute instances**: Development workstations that data scientists can use to work with data and models.
   - **Compute clusters**: Scalable clusters of virtual machines for on-demand processing of experiment code.
   - **Inference clusters**: Deployment targets for predictive services that use your trained models.
   - **Attached compute**: Links to other Azure compute resources, such as Virtual Machines or Azure Databricks clusters.

   For this exercise, you'll create a compute instance so you can run some code in your workspace.

2. On the **Compute instances** tab, add a new compute instance with the following settings. You'll use this as a workstation to run code in notebooks.

   - **Region**: *The same region as your workspace*
   - **Virtual machine type**: CPU
   - **Virtual machine size**: Standard_DS11_v2
   - **Compute name**: *enter a unique name*
   - **Enable SSH access**: Unselected (you can use this to enable direct access to the virtual machine using an SSH client)
   - **Show advanced settings**: Note the following settings, but do not select them:
     - **Enable virtual network**: Unselected (you would typically use this in an enterprise environment to enhance network security)
     - **Assign to another user**: Unselected (you can use this to assign a compute instance to a data scientist)

3. Wait for the compute instance to start and its status to change to **Running**.

Clone and run a notebook

A lot of data science and machine learning experimentation is performed by running code in *notebooks*. Your compute instance includes fully featured Python notebook environments (*Jupyter* and *JuypyterLab*) that you can use for extensive work; but for basic notebook editing, you can use the built-in **Notebooks** page in Azure Machine learning studio.

1. In Azure Machine Learning studio, view the **Notebooks** page.

2. Open a **Terminal**, and ensure its **Compute** is set to your compute instance.

3. Enter the following commands to clone a Git repository containing notebooks, data, and other files to your workspace:

```
cd Users
git clone https://github.com/MicrosoftLearning/mslearn-dp100
```

4. When the command has completed, in the **My files** pane, click ↻ to refresh the view and verify that a new **Users/mslearn-dp100** folder has been created. This folder contains multiple **.ipynb** notebook files.

5. Close the terminal pane, terminating the session.

6. In the **Users/mslearn-dp100** folder, open the **Get Started with Notebooks** notebook. Then read the notes and follow the instructions it contains.

> **Tip**: To run a code cell, select the cell you want to run and then use the ▷ button to run it.

Stop your compute instance

If you've finished exploring Azure Machine Learning for now, you should shut down your compute instance to avoid incurring unnecessary charges in your Azure subscription.

1. In Azure Machine Learning studio, on the **Compute** page, select your compute instance.
2. Click **Stop** to stop your compute instance. When it has shut down, its status will change to **Stopped**.