✓ 100 XP ▶

Use datastores

5 minutes

To add a datastore to your workspace, you can register it using the graphical interface in Azure Machine Learning studio, or you can use the Azure Machine Learning SDK. For example, the following code registers an Azure Storage blob container as a datastore named **blob_data**.

```python
from azureml.core import Workspace, Datastore

ws = Workspace.from_config()

# Register a new datastore
blob_ds = Datastore.register_azure_blob_container(workspace=ws,
                                                  datastore_name='blob_data',
                                                  container_name='data_container',
                                                  account_name='az_store_acct',
                                                  account_key='123456abcde789…')
```

Managing datastores

You can view and manage datastores in Azure Machine Learning Studio, or you can use the Azure Machine Learning SDK. For example, the following code lists the names of each datastore in the workspace.

```python
for ds_name in ws.datastores:
    print(ds_name)
```

You can get a reference to any datastore by using the **Datastore.get()** method as shown here:

```python
blob_store = Datastore.get(ws, datastore_name='blob_data')
```

The workspace always includes a *default* datastore (initially, this is the built-in **workspaceblobstore** datastore), which you can retrieve by using the **get_default_datastore()** method of a **Workspace** object, like this:

```
Python                                                                    Copy
```

```
default_store = ws.get_default_datastore()
```

Considerations for datastores

When planning for datastores, consider the following guidelines:

- When using Azure blob storage, *premium* level storage may provide improved I/O performance for large datasets. However, this option will increase cost and may limit replication options for data redundancy.
- When working with data files, although CSV format is very common, Parquet format generally results in better performance.
- You can access any datastore by name, but you may want to consider changing the default datastore (which is initially the built-in **workspaceblobstore** datastore).

To change the default datastore, use the **set_default_datastore()** method:

```
Python                                                                    Copy
```

```
ws.set_default_datastore('blob_data')
```

Next unit: Introduction to datasets

Continue >