Microsoft Certified: Azure Data Scientist Associate – Skills Measured

This document contains the skills measured on the exams associated with this certification. It does not include any upcoming or recent changes that have been made to those skills. For more information about upcoming or recent changes, see the associated exam details page(s).

Set up an Azure Machine Learning workspace

Create an Azure Machine Learning workspace

May include but is not limited to:

- create an Azure Machine Learning workspace
- configure workspace settings
- manage a workspace by using Azure Machine Learning Studio

Manage data objects in an Azure Machine Learning workspace

May include but is not limited to:

- register and maintain data stores
- create and manage datasets

Manage experiment compute contexts

May include but is not limited to:

- create a compute instance
- determine appropriate compute specifications for a training workload
- create compute targets for experiments and training

Run experiments and train models

Create models by using Azure Machine Learning Designer

May include but is not limited to:

- create a training pipeline by using Designer
- ingest data in a Designer pipeline
- use Designer modules to define a pipeline data flow
- use custom code modules in Designer

Run training scripts in an Azure Machine Learning workspace

May include but is not limited to:

- create and run an experiment by using the Azure Machine Learning SDK
- consume data from a data store in an experiment by using the Azure Machine Learning SDK
- consume data from a dataset in an experiment by using the Azure Machine Learning SDK
- choose an estimator

Generate metrics from an experiment run

May include but is not limited to:

- log metrics from an experiment run
- retrieve and view experiment outputs
- use logs to troubleshoot experiment run errors

Automate the model training process

May include but is not limited to:

- create a pipeline by using the SDK
- pass data between steps in a pipeline
- run a pipeline
- monitor pipeline runs

Optimize and manage models

Use Automated ML to create optimal models

May include but is not limited to:

- use the Automated ML interface in Studio
- use Automated ML from the Azure ML SDK
- select scaling functions and pre-processing options
- determine algorithms to be searched
- define a primary metric
- get data for an Automated ML run
- retrieve the best model

Use Hyperdrive to rune hyperparameters

May include but is not limited to:

- select a sampling method
- define the search space
- define the primary metric
- define early termination options
- find the model that has optimal hyperparameter values

Use model explainers to interpret models

May include but is not limited to:

- select a model interpreter
- generate feature importance data

Manage models

May include but is not limited to:

- register a trained model
- monitor model history
- monitor data drift

Deploy and consume models

Create production compute targets

May include but is not limited to:

- consider security for deployed services
- evaluate compute options for deployment

Deploy a model as a service

May include but is not limited to:

- configure deployment settings
- consume a deployed service
- troubleshoot deployment container issues

Create a pipeline for batch inferencing

May include but is not limited to:

- publish a batch inferencing pipeline
- run a batch inferencing pipeline and obtain outputs

Publish a Designer pipeline as a web service

May include but is not limited to:

- create a target compute resource
- configure an Inference pipeline
- consume a deployed endpoint