Study Guide

Exam DP-100: Designing and Implementing a Data Science Solution on Azure

Purpose of this document

This study guide should help you understand what to expect on the exam and includes a summary of the topics the exam might cover and links to additional resources. The information and materials in this document should help you focus your studies as you prepare for the exam.

Useful links	Description
How to earn the certification	Some certifications only require one exam, while others require more. On the details page, you'll find information about what skills are measured and links to registration. Each exam also has its own details page covering exam specifics.
Certification renewal	Once you earn your certification, don't let it expire. When you have an active certification that's expiring within six months, you should renew it—at no cost—by passing a renewal assessment on Microsoft Learn. Remember to renew your certification annually if you want to retain it.
Your Microsoft Learn profile	Connecting your certification profile to Learn brings all your learning activities together. You'll be able to schedule and renew exams, share and print certificates, badges and transcripts, and review your learning statistics inside your Learn profile.
Passing score	All technical exam scores are reported on a scale of 1 to 1,000. A passing score is 700 or greater. As this is a scaled score, it may not equal 70% of the points. A passing score is based on the knowledge and skills needed to demonstrate competence as well as the difficulty of the questions.
Exam sandbox	Are you new to Microsoft certification exams? You can explore the exam environment by visiting our exam sandbox. We created the sandbox as an opportunity for you to experience an exam before you take it. In the sandbox,

Useful links	Description
	you can interact with different question types, such as build list, case studies, and others that you might encounter in the user interface when you take an exam. Additionally, it includes the introductory screens, instructions, and help topics related to the different types of questions that your exam might include. It also includes the non-disclosure agreement that you must accept before you can launch the exam.
Request accommodations	We're committed to ensuring all learners are set up for success. If you use assistive devices, require extra time, or need modification to any part of the exam experience, you can request an accommodation.
Take a practice test	Taking a practice test is a great way to know whether you're ready to take the exam or if you need to study a bit more. Subject-matter experts write the Microsoft Official Practice Tests, which are designed to assess all exam objectives.

Objective domain: skills the exam measures

The English language version of this exam was updated on October 18, 2022.

Some exams are localized into other languages, and those are updated approximately eight weeks after the English version is updated. Other available languages are listed in the **Schedule Exam** section of the **Exam Details** webpage. If the exam isn't available in your preferred language, you can request an additional 30 minutes to complete the exam.

> **Note**
> The bullets that follow each of the skills measured are intended to illustrate how we are assessing that skill. Related topics may be covered in the exam.

> **Note**
> Most questions cover features that are general availability (GA). The exam may contain questions on Preview features if those features are commonly used.

Skills measured

- Design and prepare a machine learning solution (20–25%)
- Explore data and train models (35–40%)
- Prepare a model for deployment (20–25%)
- Deploy and retrain a model (10–15%)

Functional groups

Design and prepare a machine learning solution (20–25%)

Design a machine learning solution

- Determine the appropriate compute specifications for a training workload
- Describe model deployment requirements
- Select which development approach to use to build or train a model

Manage an Azure Machine Learning workspace

- Create an Azure Machine Learning workspace
- Manage a workspace by using developer tools for workspace interaction
- Set up Git integration for source control

Manage data in an Azure Machine Learning workspace

- Select Azure Storage resources
- Register and maintain datastores
- Create and manage data assets

Manage compute for experiments in Azure Machine Learning

- Create compute targets for experiments and training
- Select an environment for a machine learning use case
- Configure attached compute resources, including Azure Databricks and Azure Synapse Analytics
- Monitor compute utilization

Explore data and train models (35–40%)

Explore data by using data assets and data stores

- Load and transform data
- Analyze data by using Azure Data Explorer
- Use differential privacy

Create models by using the Azure Machine Learning designer

- Create a training pipeline
- Consume data assets from the designer
- Use designer components to define a pipeline data flow
- Use custom code components in designer
- Evaluate the model, including responsible AI guidelines

Use automated machine learning to explore optimal models

- Use automated machine learning for tabular data
- Use automated machine learning for computer vision

- Use automated machine learning for natural language processing (NLP)
- Select and understand training options, including preprocessing and algorithms
- Evaluate an automated machine learning run, including responsible AI guidelines

Use notebooks for custom model training

- Develop code by using a compute instance
- Consume data in a notebook
- Track model training by using MLflow
- Evaluate a model
- Train a model by using Python SDK
- Use the terminal to configure a compute instance

Tune hyperparameters with Azure Machine Learning

- Select a sampling method
- Define the search space
- Define the primary metric
- Define early termination options

Prepare a model for deployment (20–25%)

Run model training scripts

- Configure job run settings for a script
- Configure compute for a job run
- Consume data from a data asset in a job
- Run a script as a job by using Azure Machine Learning
- Use MLflow to log metrics from a job run
- Use logs to troubleshoot job run errors
- Configure an environment for a job run
- Define parameters for a job

Implement training pipelines

- Create a pipeline
- Pass data between steps in a pipeline
- Run and schedule a pipeline
- Monitor pipeline runs
- Create custom components
- Use component-based pipelines

Manage models in Azure Machine Learning

- Describe MLflow model output
- Identify an appropriate framework to package a model

- Assess a model by using responsible AI guidelines

Deploy and retrain a model (10–15%)

Deploy a model

- Configure settings for real-time deployment
- Configure compute for a batch deployment
- Deploy a model to a real-time endpoint
- Deploy a model to a batch endpoint
- Test a real-time deployed service
- Invoke the batch endpoint to start a batch scoring job

Apply machine learning operations (MLOps) practices

- Trigger an Azure Machine Learning pipeline, including from Azure DevOps or GitHub
- Automate model retraining based on new data additions or data changes
- Define event-based retraining triggers

Study Resources

We recommend that you train and get hands-on experience before you take the exam. We offer self-study options and classroom training as well as links to documentation, community sites, and videos.

Study resources	Links to learning and documentation
Get trained	Choose from self-paced learning paths and modules or take an instructor led course
Find documentation	Azure Databricks Azure Machine Learning Azure Synapse Analytics MLflow and Azure Machine Learning
Ask a question	Microsoft Q&A | Microsoft Docs
Get community support	AI - Machine Learning - Microsoft Tech Community AI - Machine Learning Blog - Microsoft Tech Community
Follow Microsoft Learn	Microsoft Learn - Microsoft Tech Community
Find a video	Microsoft Learn Shows