

## DP-100: Microsoft Azure Data Scientist Associate Exam Study Guide

April 19, 2022 by [manish](#)

★★★★★ 5/5 - (2 votes)

Azure DP-100 is the Data Science Certification from Microsoft for all the people who are interested in Data Handling. The learning is self-paced and flexible. Once you get certified, you can work on Azure models, build and operate the models, track the experiments, build pipelines, and tune hyperparameters in the Azure way.

The topics covered in this blog are:

- [What is Azure Data Scientist Certification?](#)
- [Why You Should Learn Data Science?](#)
- [Who This Certification is For?](#)
- [Benefits of DP-100 Certification](#)

- [DP-100 Exam Detail](#)
- [DP-100 Exam Skills Measured](#)
- [How to Register for Azure Dp 100 Exam](#)
- [Prerequisites for DP-100 Certification](#)
- [DP 100 Study Guide](#)
- [DP-100 Exam Retake Policy](#)
- [DP 100 Exam Day Tips](#)
- [Conclusion](#)

## What is Azure Data Scientist Certification?

With the DP 100 Microsoft Azure Data Scientist Certification, you will be able to work on machine learning workloads using Azure Machine Learning Service and apply your knowledge of data science and machine learning.

In addition, it involves planning and creating a suitable working environment for data science workloads on Azure, executing data experiments, and training predictive machine learning models.



Azure Data Scientist Associate

Are you new to the Azure cloud? Do check out our blog post on the [Azure Certification Path 2022](#) and choose the best certification for you.

## Why You Should Learn Data Science?

In IT industries, a lot of raw data is generated daily. Therefore, an IT industry needs a dedicated team to work on this raw data, evaluate it and plot the data to make inferences. Moreover, the team is required to apply the Machine Learning algorithm to make predictions.

Also, there is a huge gap between demand and supply for Data Scientists and right now Data Scientist jobs are trending in the IT industry.

## Who This Certification is For?

The certificate is applicable to,

- Candidates who are interested in Machine Learning Like Scikit-Learn, PyTorch, and Tensorflow.
- Some knowledge of Data handling and Artificial Intelligence
- People good at statistics
- The Data Scientist who prepares data, trains models, and evaluates competing models, but hasn't used Azure to do so before.

**Check Out:** [Azure Data Factory Interview Questions](#)

## Benefits of DP-100 Certification

- Create different workspaces for running machine learning models.
- Using Machine Learning to implement and run Machine Learning on [Azure](#).
- Planning, running data experiments, training predictive, and creating a working environment for Data science workloads on Azure.
- AI/ML Data Scientist Jobs are trending in the IT industry.
- The IT industry is flooded with jobs for candidates who have basic knowledge about using Microsoft Azure Services. Candidates having

Microsoft certification has a good job opportunity.

- The certificate is proof that your skills have been validated.

## DP-100 Exam Details

<b>Exam Name</b> Exam DP-100: Designing and Implementing a Data Science Solution on Azure	<b>Exam Duration</b> 180 Minutes
<b>Exam Type</b> Multiple Choice Examination	<b>Number of Questions</b> 40 - 60 Questions
<b>Exam Fee</b> \$165	<b>Eligibility/Pre-Requirement</b> None
<b>Exam validity</b> 2 years	<b>Exam Languages</b> English, Japanese, Korean, and Simplified Chinese

## DP-100 Exam Skills Measured

<b>Manage Azure resources for machine learning</b>	<b>25–30%</b>
<b>Run experiments and train models</b>	<b>20–25%</b>
<b>Deploy and operationalize machine learning solutions</b>	<b>35-40%</b>
<b>Implement responsible machine learning</b>	<b>5-10%</b>

## How to Register for Azure Dp 100 Exam

You can register for the DP-100 Exam (Designing And Implementing A Data Science Solution On Azure) by going to the [Official Microsoft Page](#).

**Languages:** English, Japanese, Chinese (Simplified), Korean, German, Chinese (Traditional), French, Spanish, Portuguese (Brazil), Russian, Arabic (Saudi Arabia), Italian, Indonesian (Indonesia)

**Retirement date:** none

This exam measures your ability to accomplish the following technical tasks: manage Azure resources for machine learning; run experiments and train models; deploy and operationalize machine learning solutions; and implement responsible machine learning.

**\$165 USD\***

Price based on the country in which the exam is proctored.

[Schedule exam >](#)

Official practice test for Designing and Implementing a Data Science Solution on Azure  
All objectives of the exam are covered in depth so you'll be ready for any question on the exam.

## Prerequisites For DP 100 Exam

- An understanding of Machine Learning, Artificial Intelligence, and Data Science.
- Candidate having existing knowledge of Python and Machine Learning. Experience of working on it for at least 3-6 months makes it easy to prepare for the exam.
- Basic knowledge of Azure data fundamentals
- [DP-900 Certification](#) (not mandatory)

By taking the DP-900 certification exam, it would be easy to learn the Basic knowledge of Azure. Moreover, it will be easy to crack the DP-100 certification exam.

Knowledge of Machine learning helps to make sense of the codes, and answer Machine Learning questions during the exam.

## DP 100 Study Guide

### Manage Azure resources for machine learning (25-30%)

#### Create an Azure Machine Learning workspace

- Create an Azure Machine Learning workspace

- [Create and Manage Azure Machine Learning workspaces](#)
- Configure workspace settings
  - [az security workspace-setting](#)
- Manage a workspace by using Azure Machine Learning studio
  - [Manage a Machine Learning Studio \(classic\) workspace](#)

## **Manage data in an Azure Machine Learning workspace**

- Select Azure storage resources
  - [Connect to storage services on Azure with datastores](#)
- Register and maintain datastores
  - [Create and register datastores](#)
- Create and manage datasets
  - [Create Azure Machine Learning datasets](#)

## **Manage compute for experiments in Azure Machine Learning**

- Determine the appropriate compute specifications for a training workload
  - [Configure and submit training runs](#)
- Create compute targets for experiments and training
  - [What are compute targets in Azure Machine Learning?](#)
  - [Set up compute targets for model training and deployment](#)
- Configure Attached Compute resources including Azure Databricks
  - [Azure Databricks](#)
- Monitor compute utilization
  - [Monitor Azure Machine Learning](#)

## **Implement security and access control in Azure Machine Learning**

- Determine access requirements and map requirements to built-in roles
  - [Default roles](#)
  - [Manage workspace access](#)

- Create custom roles
  - [Create custom role](#)
- Manage role membership
  - [Manage access to an Azure Machine Learning workspace](#)
- Manage credentials by using Azure Key Vault
  - [Use authentication credential secrets in AML training runs](#)

## **Set up an Azure Machine Learning development environment**

- Create compute instances
  - [What is an Azure Machine Learning compute instance?](#)
  - [Create and manage an Azure Machine Learning compute instance](#)
- Share compute instances
  - [share compute instance with other user](#)
- Access Azure Machine Learning workspaces from other development environments
  - [Organize and set up Azure Machine Learning environments](#)
  - [Create & use software environments in Azure Machine Learning](#)

## **Set up an Azure Databricks workspace**

- Create an Azure Databricks workspace
  - [Create an Azure Databricks workspace by using an ARM template](#)
- Create an Azure Databricks cluster
  - [Create a cluster](#)
- Create and run notebooks in Azure Databricks
  - [Create a notebook](#)
  - [Run a Databricks notebook](#)
- Link and Azure Databricks workspace to an Azure Machine Learning workspace
  - [Develop with AML & Azure Databricks](#)

**Run experiments and train models (20-25%)**

## Create models by using the Azure Machine Learning designer

- Create a training pipeline by using Azure Machine Learning designer
  - [Use pipeline parameters to retrain models in the designer](#)
- Ingest data in a designer pipeline
  - [Data ingestion options for Azure Machine Learning workflows](#)
- Use designer modules to define a pipeline data flow
  - [Designer – train a no-code regression model](#)
- Use custom code modules in designer
  - [Define custom R modules for Machine Learning Studio \(classic\)](#)

## Run model training scripts

- Create and run an experiment by using the Azure Machine Learning SDK
  - [Azure Machine Learning in Jupyter Notebooks](#)
- Configure run settings for a script
  - [Configure and submit training runs](#)
  - [ScriptRunConfig Class](#)
- Consume data from a dataset in an experiment by using the Azure Machine Learning SDK
  - [Create and register datastores](#)
- Run a training script on Azure Databricks compute
  - [Develop with AutoML & Azure Databricks](#)
- Run code to train a model in an Azure Databricks notebook
  - [Example of building ML models on Azure Databricks](#)

## Generate metrics from an experiment run

- Log metrics from an experiment run
  - [Log & view metrics and log files](#)
- Retrieve and view experiment outputs
  - [Evaluate automated machine learning experiment results](#)



- [Set input & output directories](#)
- Use logs to troubleshoot experiment run errors
  - [Collect ML pipeline log files in Application Insights for alerts and debugging](#)
- Use MLflow to track experiments
  - [Track Azure Databricks ML experiments with MLflow and Azure ML](#)
  - [Track ML models with MLflow and Azure ML](#)
- Track experiments running in Azure Databricks
  - [Track Azure Databricks experiments in Azure ML](#)

## **Use Automated Machine Learning to create optimal models**

- Use the Automated ML interface in Azure Machine Learning studio
  - [Forecast demand with automated machine learning](#)
- Use Automated ML from the Azure Machine Learning SDK
  - [Set up AutoML training with Python](#)
- Select pre-processing options
  - [Preprocessing Data in Azure Machine Learning Studio](#)
  - [What is automated machine learning \(AutoML\)?](#)
- Select the algorithms to be searched
  - [How to select algorithms for Azure Machine Learning](#)
- Define a primary metric
  - [Primary metric](#)
- Get data for an Automated ML run
  - [Explore the results](#)
- Retrieve the best model
  - [Retrieve the best model](#)

## **Tune hyperparameters with Azure Machine Learning**

- Select a sampling method
  - [Sampling the hyperparameter space](#)

- Define the search space
  - [Hyperparameter tuning a model with Azure Machine Learning](#)
- Define the primary metric
  - [Primary metric](#)
  - [Specify primary metric](#)
  - [Choosing a primary metric to handle imbalanced data](#)
- Define early termination options
  - [Specify early termination policy](#)
    - [Bandit policy](#)
    - [Median stopping policy](#)
    - [Truncation selection policy](#)
    - [No termination policy \(default\)](#)
- Find the model that has optimal hyperparameter values
  - [Find the best model](#)

## Deploy and operationalize machine learning solutions (35-40%)

### Select compute for model deployment

- Consider security for deployed services
  - [Use TLS to secure a web service through Azure ML](#)
- Evaluate compute options for deployment
  - [Deploy machine learning models to Azure](#)

### Deploy a model as a service

- Configure deployment settings
  - [Deploy machine learning models to Azure](#)
- Deploy a registered model
  - [Deploy machine learning models](#)
- Deploy a model trained in Azure Databricks to an Azure Machine Learning endpoint
  - [Deploy and Serve Model from Azure Databricks onto Azure ML](#)

- Consume a deployed service
  - [Consume an Azure ML model deployed as a web service](#)
- Troubleshoot deployment container issues
  - [Troubleshooting remote model deployment](#)

## **Manage models in Azure Machine Learning**

- Register a trained model
  - [Register the model](#)
- Monitor model usage
  - [Monitor Azure Machine Learning](#)
  - [Monitoring data from Azure Machine Learning](#)
- Monitor data drift
  - [Detect data drift \(preview\) on datasets](#)
  - [Monitor data drift with Azure Machine Learning](#)

## **Create an Azure Machine Learning pipeline for batch inferencing**

- Configure a ParallelRunStep
  - [ParallelRunStep Class](#)
- Configure compute for a batch inferencing pipeline
  - [Create & attach the remote compute target](#)
- Publish a batch inferencing pipeline
  - [Run batch predictions using Azure ML designer](#)
- Run a batch inferencing pipeline and obtain outputs
  - [Build an Azure ML pipeline for batch scoring](#)
- Obtain outputs from a ParallelRunStep
  - [ParallelRunStep Class](#)

## **Publish an Azure Machine Learning designer pipeline as a web service**

- Create a target compute resource

- [Deploy machine learning models to Azure](#)
- Configure an inference pipeline
  - [Designer – deploy a machine learning model](#)
- Consume a deployed endpoint
  - [Deploy machine learning models to Azure](#)

## **Implement pipelines by using the Azure Machine Learning SDK**

- Create a pipeline
  - [Use automated machine learning to predict taxi fares](#)
- Pass data between steps in a pipeline
  - [Pass data between pipeline steps](#)
- Run a pipeline
  - [Running automated machine learning experiments](#)
- Monitor pipeline runs
  - [Monitor Azure Machine Learning](#)

## **Apply ML Ops practices**

- Trigger an Azure Machine Learning pipeline from Azure DevOps
  - [Trigger machine learning pipelines](#)
  - [Azure Machine Learning training pipeline using CI/CD with Azure DevOps](#)
  - [Enabling CI/CD for Machine Learning project with Azure Pipelines](#)
- Automate model retraining based on new data additions or data changes
  - [Retrain your model on new data](#)
  - [Retraining & updating AML models](#)
- Refactor notebooks into scripts
  - [Convert ML experiments to production Python code](#)
- Implement source control for scripts
  - [Git integration for Azure Machine Learning](#)
  - [Using Azure Machine Learning from GitHub Actions](#)

## Implement responsible machine learning (5-10%)

### Use model explainers to interpret models

- Select a model interpreter
  - [Model interpretability in Azure Machine Learning \(preview\)](#)
  - [Model explainability in automated ML \(preview\)](#)
  - [Interpret & explain ML models in Python](#)
- Generate feature importance data
  - [Generate feature importance value on your personal machine](#)
  - [Permutation Feature Importance](#)

### Describe fairness considerations for models

- Evaluate model fairness based on prediction disparity
  - [Machine learning fairness \(preview\)](#)
  - [Assess ML models' fairness in Python](#)
- Mitigate model unfairness
  - [Detect and mitigate unfairness in models with Azure Machine Learning](#)

### Describe privacy considerations for data

- Describe principles of differential privacy
  - [What is differential privacy in machine learning \(preview\)?](#)
- Specify acceptable levels of noise in data and the effects on privacy
  - [Use differential privacy in Azure Machine Learning](#)

## DP-100 Exam Retake Policy

The DP-100 exam retake policy is as follows:

1. If a candidate fails on the first attempt, they must wait for 24 hours

- before retaking the exam.
2. If a candidate again fails on the second attempt, then the candidate will have to wait for 14 days.
  3. A candidate will be given a maximum of five attempts to retake an exam in a year.

## DP-100 Exam Day Tips

- Schedule the exam a week or two in advance.
- Set up a certification profile (if you haven't already)
- Verify personal information and payment method
- It is recommended to take the exam online.
- The test-taking space should be quiet.
- Do not read the question aloud, this may lead to disqualification.
- Before taking photos, remove any paper, pencils, external keyboard, etc, from sight in compliance with test rules.
- You are not allowed to write, you can use a virtual notepad during the test.
- Be mindful of your eyes during your test. You will be recorded by your front-facing camera during your test, and a Pearson VUE proctor will monitor you throughout the session. The lack of attention to your computer screen during the test may lead to suspicions that you are cheating.
- If you stare off into the distance while thinking, it is best to resist that urge during your exam.

## Conclusion

There is a huge demand for data scientists, machine learning experts, Artificial intelligence, etc. and the future of Data science is defined by large firms such as [Microsoft](#), Amazon, Databricks, and google. Due to this reason, it makes sense to get certification from any of these big firms.

The Azure DP-100 certification is just the beginning of your learning journey toward Machine Learning, not the end. Once you get the certification it is on you to put that skills into good use whether it is a job or for your personal project.

## Related/References

- [DP-300 Exam Study Guide](#)
- [AZ-500 Exam Study Guide](#)
- [AZ-204 Exam Study Guide](#)
- [SC-900 Exam Study Guide](#)
- [AZ-900 Exam Study Guide](#)
- [AI-900 Exam Study Guide](#)



## Leave a Comment

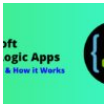
Save my name, email, and website in this browser for the next time I comment.

Post Comment

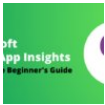
## Recent Posts



Azure Storage Explorer: Download, Install, and Setup Overview



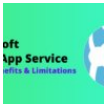
What are Azure Logic Apps: Components, Advantages and How it Works



Microsoft Azure Application Insights: A Complete Beginners Guide



Microsoft Azure Service Bus: A Complete Beginners Guide



Azure App Service: Types, Benefits and Limitations



