



## DP-100T01 Designing and Implementing a Data Science Solution on Azure

### Summary

Length: 24 hours Level: Experienced

Gain the necessary knowledge about how to use Azure services to develop, train, and deploy, machine learning solutions. The course starts with an overview of Azure services that support data science. From there, it focuses on using Azure's premier data science service, Azure Machine Learning service, to automate the data science pipeline. This course is focused on Azure and does not teach the student how to do data science. It is assumed students already know that. This course may earn a Credly Badge.

### Learning Objectives

This role is responsible for solving the business problem that initiated the project. While the Data Engineer will prepare the data to be used for the models, the Data Scientist determines what data is needed for model training, creates model features from the data, determines what machine learning model to use, trains and evaluates the model, and often has involvement in model deployment. Often the data scientist needs to evaluate multiple models to determine which performs the best.

### Course Outline

1. **Doing Data Science on Azure**

- Introduce the Data Science Process
- Overview of Azure Data Science Options
- Introduce Azure Notebooks

2. **Doing Data Science with Azure Machine Learning service**

- Introduce Azure Machine Learning (AML) service
- Register and deploy ML models with AML service

3. **Automate Machine Learning with Azure Machine Learning service**

- Automate Machine Learning Model Selection
- Automate Hyperparameter Tuning with HyperDrive

4. **Manage and Monitor Machine Learning Models with the Azure Machine Learning service**

- Manage and Monitor Machine Learning Models

### Audience

This course is aimed at data scientists and those with significant responsibilities in training and deploying machine learning models.

### Prerequisites

Before attending this course, students must have: Azure Fundamentals Understanding of data science including how to prepare data, train models, and evaluate competing models to select the best one. How to program in the Python programming language and use the Python libraries: pandas, scikit-learn, matplotlib, and seaborn.

