

## DP-100<sup>Q&As</sup>

Designing and Implementing a Data Science Solution on Azure

### Pass Microsoft DP-100 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.leads4pass.com/dp-100.html>

100% Passing Guarantee  
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft  
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers



## QUESTION 1

You plan to provision an Azure Machine Learning Basic edition workspace for a data science project.

You need to identify the tasks you will be able to perform in the workspace.

Which three tasks will you be able to perform? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. Create a Compute Instance and use it to run code in Jupyter notebooks.
- B. Create an Azure Kubernetes Service (AKS) inference cluster.
- C. Use the designer to train a model by dragging and dropping pre-defined modules.
- D. Create a tabular dataset that supports versioning.
- E. Use the Automated Machine Learning user interface to train a model.

Correct Answer: ABD

Incorrect Answers:

C, E: The UI is included the Enterprise edition only.

Reference:

<https://azure.microsoft.com/en-us/pricing/details/machine-learning/>

---

## QUESTION 2

You create a binary classification model.

You need to evaluate the model performance.

Which two metrics can you use? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

- A. relative absolute error
- B. precision
- C. accuracy
- D. mean absolute error
- E. coefficient of determination

Correct Answer: BC

The evaluation metrics available for binary classification models are: Accuracy, Precision, Recall, F1 Score, and AUC.

Note: A very natural question is: 'Out of the individuals whom the model, how many were classified correctly (TP)?'

This question can be answered by looking at the Precision of the model, which is the proportion of positives that are classified correctly.

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio/evaluate-model-performance>

---

### QUESTION 3

You are a data scientist working for a hotel booking website company. You use the Azure Machine Learning service to train a model that identifies fraudulent transactions.

You must deploy the model as an Azure Machine Learning real-time web service using the Model.deploy method in the Azure Machine Learning SDK. The deployed web service must return real-time predictions of fraud based on transaction

data input.

You need to create the script that is specified as the entry\_script parameter for the InferenceConfig class used to deploy the model.

What should the entry script do?

- A. Register the model with appropriate tags and properties.
- B. Create a Conda environment for the web service compute and install the necessary Python packages.
- C. Load the model and use it to predict labels from input data.
- D. Start a node on the inference cluster where the web service is deployed.
- E. Specify the number of cores and the amount of memory required for the inference compute.

Correct Answer: C

The entry script receives data submitted to a deployed web service and passes it to the model. It then takes the response returned by the model and returns that to the client. The script is specific to your model. It must understand the data

that the model expects and returns.

The two things you need to accomplish in your entry script are:

Loading your model (using a function called init())

Running your model on input data (using a function called run())

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-deploy-and-where>

---

**QUESTION 4**

DRAG DROP

You have been tasked with moving data into Azure Blob Storage for the purpose of supporting Azure Machine Learning.

Which of the following can be used to complete your task? Answer by dragging the correct options from the list to the answer area.

Select and Place:

# Options

# Answer

AzCopy

Bulk Copy Program  
(BCP)

SSIS

Bulk Insert SQL Query

Azure Storage  
Explorer

Correct Answer:

# Options

# Answer

Bulk Copy Program (BCP)
Bulk Insert SQL Query

AzCopy
SSIS
Azure Storage Explorer

You can move data to and from Azure Blob storage using different technologies: Azure Storage-Explorer AzCopy Python SSIS

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/team-data-science-process/move-azure-blob>

## QUESTION 5

You are analyzing a dataset by using Azure Machine Learning Studio.

You need to generate a statistical summary that contains the p-value and the unique count for each feature column.

Which two modules can you use? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Computer Linear Correlation

- B. Export Count Table
- C. Execute Python Script
- D. Convert to Indicator Values
- E. Summarize Data

Correct Answer: BE

The Export Count Table module is provided for backward compatibility with experiments that use the Build Count Table (deprecated) and Count Featurizer (deprecated) modules.

E: Summarize Data statistics are useful when you want to understand the characteristics of the complete dataset. For example, you might need to know:

How many missing values are there in each column?

How many unique values are there in a feature column?

What is the mean and standard deviation for each column? The module calculates the important scores for each column, and returns a row of summary statistics for each variable (data column) provided as input.

Incorrect Answers:

A: The Compute Linear Correlation module in Azure Machine Learning Studio is used to compute a set of Pearson correlation coefficients for each possible pair of variables in the input dataset.

C: With Python, you can perform tasks that aren't currently supported by existing Studio modules such as:

Visualizing data using matplotlib

Using Python libraries to enumerate datasets and models in your workspace Reading, loading, and manipulating data from sources not supported by the Import Data module

D: The purpose of the Convert to Indicator Values module is to convert columns that contain categorical values into a series of binary indicator columns that can more easily be used as features in a machine learning model.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/export-count-table>

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/summarize-data>

---

## QUESTION 6

You need to implement a new cost factor scenario for the ad response models as illustrated in the performance curve exhibit. Which technique should you use?

- A. Set the threshold to 0.5 and retrain if weighted Kappa deviates +/- 5% from 0.45.
- B. Set the threshold to 0.05 and retrain if weighted Kappa deviates +/- 5% from 0.5.
- C. Set the threshold to 0.2 and retrain if weighted Kappa deviates +/- 5% from 0.6.
- D. Set the threshold to 0.75 and retrain if weighted Kappa deviates +/- 5% from 0.15.

Correct Answer: A

Scenario:

Performance curves of current and proposed cost factor scenarios are shown in the following diagram:



The ad propensity model uses a cut threshold is 0.45 and retrains occur if weighted Kappa deviated from 0.1 +/- 5%.

## QUESTION 7

You have been tasked with employing a machine learning model, which makes use of a PostgreSQL database and needs GPU processing, to forecast prices.

You are preparing to create a virtual machine that has the necessary tools built into it.

You need to make use of the correct virtual machine type.

Recommendation: You make use of a Geo AI Data Science Virtual Machine (Geo-DSVM) Windows edition.

Will the requirements be satisfied?

A. Yes

B. No

Correct Answer: B

The Azure Geo AI Data Science VM (Geo-DSVM) delivers geospatial analytics capabilities from Microsoft's Data Science VM. Specifically, this VM extends the AI and data science toolkits in the Data Science VM by adding ESRI's market-leading ArcGIS Pro Geographic Information System.

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/data-science-virtual-machine/overview>

## QUESTION 8

You are in the process of creating a machine learning model. Your dataset includes rows with null and missing values.

You plan to make use of the Clean Missing Data module in Azure Machine Learning Studio to detect and fix the null and missing values in the dataset.

Recommendation: You make use of the Remove entire row option.

Will the requirements be satisfied?

A. Yes

B. No

Correct Answer: A

Remove entire row: Completely removes any row in the dataset that has one or more missing values. This is useful if the missing value can be considered randomly missing. Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data>

---

## QUESTION 9

You have the following Azure subscriptions and Azure Machine Learning service workspaces:

Subscription	Workspace	Comment
385bdf5e-4cef-4ad4-b977-3f86d92727c9	ml-default	This is the default subscription.
5a5891d1-557a-4234-9b83-2e90412b1068	ml-project	The information required to uniquely identify this workspace is stored in the file config.json in the same folder as the Python script.

You need to obtain a reference to the ml-project workspace. Solution: Run the following Python code:

```
from azureml.core import Workspace  
ws = Workspace(workspace_name="ml-project")
```

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

## QUESTION 10

You have an Azure Machine Learning (ML) model deployed to an online endpoint.

You need to review container logs from the endpoint by using Azure ML Python SDK v2. The logs must include the console log from the inference server, with print/log statements from the model's scoring script.

What should you do first?



- A. Connect by using SSH to the inference server.
- B. Create an instance of the MLClient class.
- C. Connect by using Docker tools to the inference server.
- D. Create an instance of the OnlineDeploymentOperations class.

Correct Answer: B

Get container logs

To see log output from container, use the get\_logs method as follows:

Python

```
ml_client.online_deployments.get_logs(  
name="", endpoint_name="", lines=100  
)
```

You can also get logs from the storage initializer container by adding container\_type="storage-initializer" option.

Python

```
ml_client.online_deployments.get_logs(  
name="", endpoint_name="", lines=100, container_type="storage-initializer" )
```

Note (B, not D): OnlineDeploymentOperations Class You should not instantiate this class directly. Instead, you should create an MLClient instance that instantiates it for you and attaches it as an attribute.

Reference: <https://learn.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-online-endpoints>  
<https://learn.microsoft.com/en-us/python/api/azure-ai-ml/azure.ai.ml.operations.onlinedeploymentoperations>

---

## QUESTION 11

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while

others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You are using Azure Machine Learning Studio to perform feature engineering on a dataset.

You need to normalize values to produce a feature column grouped into bins.

Solution: Apply an Entropy Minimum Description Length (MDL) binning mode.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: A

Entropy MDL binning mode: This method requires that you select the column you want to predict and the column or columns that you want to group into bins. It then makes a pass over the data and attempts to determine the number of bins that minimizes the entropy. In other words, it chooses a number of bins that allows the data column to best predict the target column. It then returns the bin number associated with each row of your data in a column named quantized.

References: <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/group-data-into-bins>

---


## QUESTION 12

HOTSPOT

Complete the sentence by selecting the correct option in the answer area.

Hot Area:

### Answer Area

To move a large dataset from Azure Machine Learning Studio to a Weka environment, the data must be converted to the  format.

	▼
CSV	
DOCX	
ARFF	
TXT	

Correct Answer:

## Answer Area

To move a large dataset from Azure Machine Learning Studio to a Weka environment, the data must be converted to the



A dropdown menu with a downward arrow icon. The menu is open, showing four options: CSV, DOCX, ARFF (highlighted in green), and TXT.

format.

Use the Convert to ARFF module in Azure Machine Learning Studio, to convert datasets and results in Azure Machine Learning to the attribute-relation file format used by the Weka toolset. This format is known as ARFF.

The ARFF data specification for Weka supports multiple machine learning tasks, including data preprocessing, classification, and feature selection. In this format, data is organized by entities and their attributes, and is contained in a single text file. Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/convert-to-arff>

### QUESTION 13

You use Azure Machine Learning designer to load the following datasets into an experiment: Dataset1

Age	Length	Width
3	22	13
7	11	96
18	32	85

Dataset2

Age	Length	Width
11	101	65
6	98	23
33	22	54
17	52	12

You need to create a dataset that has the same columns and header row as the input datasets and contains all rows from both input datasets.

Solution: Use the Add Rows module.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: A

Add Rows component

Use this component to concatenate two datasets. In concatenation, the rows of the second dataset are added to the end of the first dataset.

Concatenation of rows is useful in scenarios such as these:

You have generated a series of evaluation statistics, and you want to combine them into one table for easier reporting.

You have been working with different datasets, and you want to combine the datasets to create a final dataset.

Reference: <https://learn.microsoft.com/en-us/azure/machine-learning/component-reference/add-rows>

---

## QUESTION 14

DRAG DROP

You are producing a multiple linear regression model in Azure Machine Learning Studio.

Several independent variables are highly correlated.

You need to select appropriate methods for conducting effective feature engineering on all the data.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:

Action	Answer area
Evaluate the probability function	
Remove duplicate rows	
Use the Filter Based Feature Selection module	
Test the hypothesis using t-Test	
Compute linear correlation	
Build a counting transform	

Correct Answer:

Action	Answer area
Evaluate the probability function	Use the Filter Based Feature Selection module
Remove duplicate rows	Build a counting transform
	Test the hypothesis using t-Test
Compute linear correlation	

Step 1: Use the Filter Based Feature Selection module

Filter Based Feature Selection identifies the features in a dataset with the greatest predictive power.

The module outputs a dataset that contains the best feature columns, as ranked by predictive power. It also outputs the names of the features and their scores from the selected metric.

Step 2: Build a counting transform

A counting transform creates a transformation that turns count tables into features, so that you can apply the transformation to multiple datasets.

Step 3: Test the hypothesis using t-Test

References:

<https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/studio-module-reference/filter-based-feature-selection>

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/build-counting-transform>

---

## QUESTION 15

You retrain an existing model.

You need to register the new version of a model while keeping the current version of the model in the registry.

What should you do?

- A. Register a model with a different name from the existing model and a custom property named version with the value 2.
- B. Register the model with the same name as the existing model.
- C. Save the new model in the default datastore with the same name as the existing model. Do not register the new model.
- D. Delete the existing model and register the new one with the same name.

Correct Answer: B

Model version: A version of a registered model. When a new model is added to the Model Registry, it is added as Version 1. Each model registered to the same model name increments the version number.

Reference: <https://docs.microsoft.com/en-us/azure/databricks/applications/mlflow/model-registry>

[Latest DP-100 Dumps](#)

[DP-100 PDF Dumps](#)

[DP-100 Exam Questions](#)