



# 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.lead4pass.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 are planning to register a trained model in an Azure Machine Learning workspace.

You must store additional metadata about the model in a key-value format. You must be able to add new metadata and modify or delete metadata after creation.

You need to register the model.

Which parameter should you use?

- A. description
- B. model\_framework
- C. tags
- D. properties

Correct Answer: D

`azureml.core.Model.properties`:

Dictionary of key value properties for the Model. These properties cannot be changed after registration, however new key value pairs can be added.

Reference:

<https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.model.model>

---

### QUESTION 2

You plan to use a Deep Learning Virtual Machine (DLVM) to train deep learning models using Compute Unified Device Architecture (CUDA) computations.

You need to configure the DLVM to support CUDA. What should you implement?

- A. Solid State Drives (SSD)
- B. Computer Processing Unit (CPU) speed increase by using overclocking
- C. Graphic Processing Unit (GPU)
- D. High Random Access Memory (RAM) configuration
- E. Intel Software Guard Extensions (Intel SGX) technology

Correct Answer: C

A Deep Learning Virtual Machine is a pre-configured environment for deep learning using GPU instances.

References: <https://azuremarketplace.microsoft.com/en-au/marketplace/apps/microsoft-ads.dsvm-deep-learning>

---



### QUESTION 3

You need to configure the Feature Based Feature Selection module based on the experiment requirements and datasets.

How should you configure the module properties? To answer, select the appropriate options in the dialog box in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

**Answer Area**

Filter Based Feature Selection

Feature scoring method

	▼
Fisher Score	
Chi-squared	
Mutual information	
Counts	

Operate on feature columns only

Target column

	▼
MedianValue	
AvgRooms/nHouse	

Launch column selector

Number of desired features

1
---

Correct Answer:



## Answer Area

### Filter Based Feature Selection

Feature scoring method

	▼
Fisher Score	
Chi-squared	
Mutual information	
Counts	

Operate on feature columns only

Target column

	▼
MedianValue	
AvgRooms/nHouse	

Launch column selector

Number of desired features

1

Box 1: Mutual Information.

The mutual information score is particularly useful in feature selection because it maximizes the mutual information between the joint distribution and target variables in datasets with many dimensions.

Box 2: MedianValue

MedianValue is the feature column, , it is the predictor of the dataset.

Scenario: The MedianValue and AvgRoomsInHouse columns both hold data in numeric format. You need to select a feature selection algorithm to analyze the relationship between the two columns in more detail.

References:

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

## QUESTION 4

You create a Python script that runs a training experiment in Azure Machine Learning. The script uses the Azure Machine Learning SDK for Python.

You must add a statement that retrieves the names of the logs and outputs generated by the script.

You need to reference a Python class object from the SDK for the statement.



Which class object should you use?

- A. Run
- B. ScriptRunConfig
- C. Workspace
- D. Experiment

Correct Answer: A

A run represents a single trial of an experiment. Runs are used to monitor the asynchronous execution of a trial, log metrics and store output of the trial, and to analyze results and access artifacts generated by the trial.

The run Class `get_all_logs` method downloads all logs for the run to a directory.

Incorrect Answers:

A: A run represents a single trial of an experiment. Runs are used to monitor the asynchronous execution of a trial, log metrics and store output of the trial, and to analyze results and access artifacts generated by the trial.

B: A `ScriptRunConfig` packages together the configuration information needed to submit a run in Azure ML, including the script, compute target, environment, and any distributed job-specific configs.

Reference: [https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.run\(class\)](https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.core.run(class))

---

## QUESTION 5

You are with a time series dataset in Azure Machine Learning Studio.

You need to split your dataset into training and testing subsets by using the Split Data module.

Which splitting mode should you use?

- A. Recommender Split
- B. Regular Expression Split
- C. Relative Expression Split
- D. Split Rows with the Randomized split parameter set to true

Correct Answer: D

Split Rows: Use this option if you just want to divide the data into two parts. You can specify the percentage of data to put in each split, but by default, the data is divided 50-50. Incorrect Answers:

B: Regular Expression Split: Choose this option when you want to divide your dataset by testing a single column for a value.

C: Relative Expression Split: Use this option whenever you want to apply a condition to a number column.

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

---



### QUESTION 6

You are evaluating a completed binary classification machine learning model.

You need to use the precision as the evaluation metric.

Which visualization should you use?

- A. Violin plot
- B. Gradient descent
- C. Box plot
- D. Binary classification confusion matrix

Correct Answer: D

Incorrect Answers:

A: A violin plot is a visual that traditionally combines a box plot and a kernel density plot.

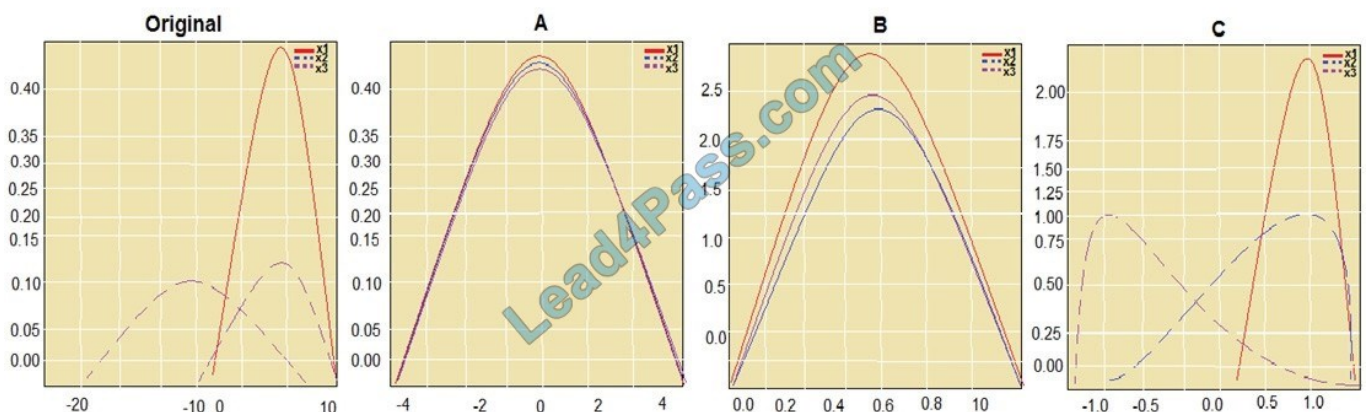
B: Gradient descent is a first-order iterative optimization algorithm for finding the minimum of a function. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or approximate gradient) of the function at the current point.

C: A box plot lets you see basic distribution information about your data, such as median, mean, range and quartiles but doesn't show you how your data looks throughout its range.

References: <https://machinelearningknowledge.ai/confusion-matrix-and-performance-metrics-machine-learning/>

### QUESTION 7

You are performing feature scaling by using the scikit-learn Python library for  $x_1$ ,  $x_2$ , and  $x_3$  features. Original and scaled data is shown in the following image.



Use the drop-down menus to select the answer choice that answers each question based on the information presented in the graphic. NOTE: Each correct selection is worth one point.



Hot Area:

## Answer Area

Question	Answer choice
Which scaler is used in graph A?	<input type="text"/> Standard Scaler Min Max Scale Normalizer
Which scaler is used in graph B?	<input type="text"/> Standard Scaler Min Max Scale Normalizer
Which scaler is used in graph C?	<input type="text"/> Standard Scaler Min Max Scale Normalizer

Correct Answer:





## Answer Area

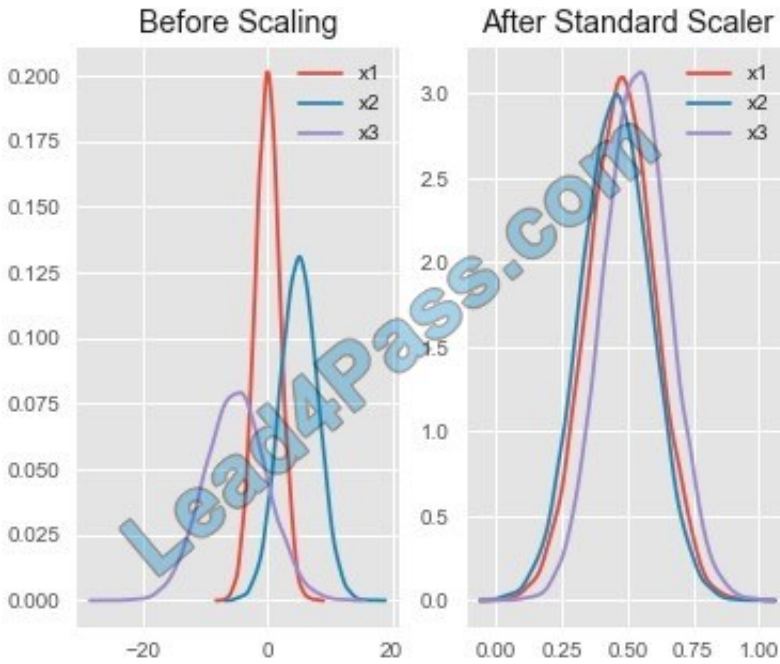
Question	Answer choice
Which scaler is used in graph A?	<input type="text"/> Standard Scaler Min Max Scale Normalizer
Which scaler is used in graph B?	<input type="text"/> Standard Scaler Min Max Scale Normalizer
Which scaler is used in graph C?	<input type="text"/> Standard Scaler Min Max Scale Normalizer

Box 1: StandardScaler

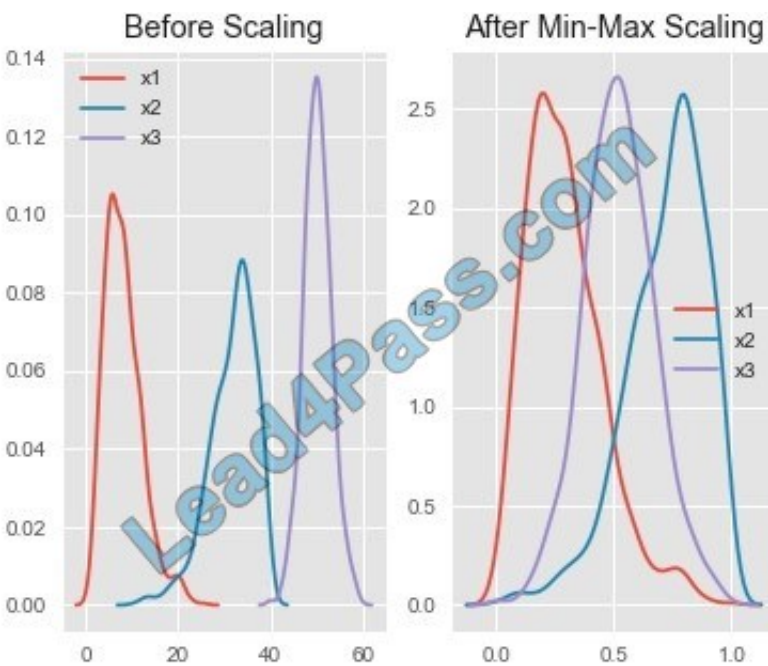
The StandardScaler assumes your data is normally distributed within each feature and will scale them such that the distribution is now centred around 0, with a standard deviation of 1.

Example:





All features are now on the same scale relative to one another. Box 2: Min Max Scaler



Notice that the skewness of the distribution is maintained but the 3 distributions are brought into the same scale so that they overlap.

Box 3: Normalizer

References:

<http://benalexkeen.com/feature-scaling-with-scikit-learn/>



### QUESTION 8

You use the Azure Machine Learning designer to create and run a training pipeline. You then create a real-time inference pipeline.

You must deploy the real-time inference pipeline as a web service.

What must you do before you deploy the real-time inference pipeline?

- A. Run the real-time inference pipeline.
- B. Create a batch inference pipeline.
- C. Clone the training pipeline.
- D. Create an Azure Machine Learning compute cluster.

Correct Answer: D

You need to create an inferencing cluster.

Deploy the real-time endpoint

After your AKS service has finished provisioning, return to the real-time inferencing pipeline to complete deployment.

1.

Select Deploy above the canvas.

2.

Select Deploy new real-time endpoint.

3.

Select the AKS cluster you created.

4.

Select Deploy.

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-deploy>

---

### QUESTION 9

You are building a machine learning model for translating English language textual content into French language textual content.

You need to build and train the machine learning model to learn the sequence of the textual content.

Which type of neural network should you use?

- A. Multilayer Perceptions (MLPs)
- B. Convolutional Neural Networks (CNNs)



C. Recurrent Neural Networks (RNNs)

D. Generative Adversarial Networks (GANs)

Correct Answer: C

To translate a corpus of English text to French, we need to build a recurrent neural network (RNN).

Note: RNNs are designed to take sequences of text as inputs or return sequences of text as outputs, or both. They're called recurrent because the network's hidden layers have a loop in which the output and cell state from each time step become inputs at the next time step. This recurrence serves as a form of memory. It allows contextual information to flow through the network so that relevant outputs from previous time steps can be applied to network operations at the current time step.

Reference: <https://towardsdatascience.com/language-translation-with-rnns-d84d43b40571>

---

### QUESTION 10

You need to consider the underlined segment to establish whether it is accurate.

To transform a categorical feature into a binary indicator, you should make use of the Clean Missing Data module.

Select "No adjustment required" if the underlined segment is accurate. If the underlined segment is inaccurate, select the accurate option.

A. No adjustment required.

B. Convert to Indicator Values

C. Apply SQL Transformation

D. Group Categorical Values

Correct Answer: B

Explanation:

Use the Convert to Indicator Values module in Azure Machine Learning Studio. The purpose of this 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.

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/convert-to-indicator-values>

---

### 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 train a classification model by using a logistic regression algorithm.

You must be able to explain the model's predictions by calculating the importance of each feature, both as an overall global relative importance value and as a measure of local importance for a specific set of predictions.

You need to create an explainer that you can use to retrieve the required global and local feature importance values.

Solution: Create a MimicExplainer.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

Instead use Permutation Feature Importance Explainer (PFI).

Note 1: Mimic explainer is based on the idea of training global surrogate models to mimic blackbox models. A global surrogate model is an intrinsically interpretable model that is trained to approximate the predictions of any black box model as accurately as possible. Data scientists can interpret the surrogate model to draw conclusions about the black box model.

Note 2: Permutation Feature Importance Explainer (PFI): Permutation Feature Importance is a technique used to explain classification and regression models. At a high level, the way it works is by randomly shuffling data one feature at a time for the entire dataset and calculating how much the performance metric of interest changes. The larger the change, the more important that feature is. PFI can explain the overall behavior of any underlying model but does not explain individual predictions.

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/how-to-machine-learning-interpretability>

---

## QUESTION 12

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

Explanation:

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 13**

**HOTSPOT**

You have a multi-class image classification deep learning model that uses a set of labeled photographs. You create the following code to select hyperparameter values when training the model.

```
from azureml.train.hyperdrive import BayesianParameterSampling
param_sampling = BayesianParametersSampling ({
    "learning_rate": uniform(0.01, 0.1),
    "batch_size": choice(16, 32, 64, 128)}
)
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No.

NOTE: Each correct selection is worth one point.

Hot Area:

	<b>Yes</b>	<b>No</b>
Hyperparameter combinations for the runs are selected based on how previous samples performed in the previous experiment run.	<input type="radio"/>	<input type="radio"/>
The learning rate value 0.09 might be used during model training.	<input type="radio"/>	<input type="radio"/>
You can define an early termination policy for this hyperparameter tuning run.	<input type="radio"/>	<input type="radio"/>

Correct Answer:

	<b>Yes</b>	<b>No</b>
Hyperparameter combinations for the runs are selected based on how previous samples performed in the previous experiment run.	<input checked="" type="radio"/>	<input type="radio"/>
The learning rate value 0.09 might be used during model training.	<input checked="" type="radio"/>	<input type="radio"/>
You can define an early termination policy for this hyperparameter tuning run.	<input type="radio"/>	<input checked="" type="radio"/>

Box 1: Yes

Hyperparameters are adjustable parameters you choose to train a model that govern the training process itself. Azure



Machine Learning allows you to automate hyperparameter exploration in an efficient manner, saving you significant time

and resources. You specify the range of hyperparameter values and a maximum number of training runs. The system then automatically launches multiple simultaneous runs with different parameter configurations and finds the configuration

that results in the best performance, measured by the metric you choose. Poorly performing training runs are automatically early terminated, reducing wastage of compute resources. These resources are instead used to explore other

hyperparameter configurations.

Box 2: Yes

uniform(low, high) - Returns a value uniformly distributed between low and high

Box 3: No

Bayesian sampling does not currently support any early termination policy.

Reference:

<https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters>

[Latest DP-100 Dumps](#)

[DP-100 PDF Dumps](#)

[DP-100 VCE Dumps](#)



To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

## Try our product !

100% Guaranteed Success

100% Money Back Guarantee

365 Days Free Update

Instant Download After Purchase

24x7 Customer Support

Average 99.9% Success Rate

More than 800,000 Satisfied Customers Worldwide

Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

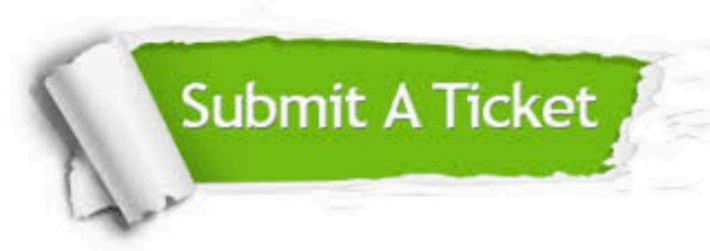
We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.lead4pass.com/allproducts>

## Need Help

Please provide as much detail as possible so we can best assist you.

To update a previously submitted ticket:



 <p><b>One Year Free Update</b> Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p><b>Money Back Guarantee</b> To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p><b>Security &amp; Privacy</b> We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.

All trademarks are the property of their respective owners.

Copyright © lead4pass, All Rights Reserved.