

DP-100^{Q&As}

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.pass2lead.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

HOTSPOT

You create an Azure Machine Learning workspace and a dataset. The dataset includes age values for a large group of diabetes patients. You use the `dp.mean` function from the SmartNoise library to calculate the mean of the age value. You

store the value in a variable named `age.mean`.

You must output the value of the interval range of released mean values that will be returned 95 percent of the time.

You need to complete the code.

Which code values should you use? To answer, select the appropriate options in the answer area

NOTE: Each correct selection is worth one point.

Hot Area:

```
print(age_mean,  
```

A

- `compute_privacy_usage`
- placeholder

B

- 0.95
- 0.85
- 0.65

Correct Answer:

```
print(age_mean,  
```

A

- `compute_privacy_usage`
- placeholder

B

- 0.95
- 0.85
- 0.65

QUESTION 2

You are analyzing a raw dataset that requires cleaning.

You must perform transformations and manipulations by using Azure Machine Learning Studio.

You need to identify the correct modules to perform the transformations.

Which modules should you choose? To answer, drag the appropriate modules to the correct scenarios. Each module may be used once, more than once, or not at all.

You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Select and Place:

Answer Area

Methods

Clean Missing Data

SMOTE

Convert to Indicator Values

Remove Duplicate Rows

Threshold Filter

Scenario

Replace missing values by removing rows and columns.

Increase the number of low-incidence examples in the dataset.

Convert a categorical feature into a binary indicator.

Remove potential duplicates from a dataset.

Module

Correct Answer:

Answer Area

Methods	Scenario	Module
<input type="text"/>	Replace missing values by removing rows and columns.	Clean Missing Data
<input type="text"/>	Increase the number of low-incidence examples in the dataset.	SMOTE
<input type="text"/>	Convert a categorical feature into a binary indicator.	Convert to Indicator Values
<input type="text"/>	Remove potential duplicates from a dataset.	Remove Duplicate Rows
Threshold Filter		

Box 1: Clean Missing Data

Box 2: SMOTE Use the SMOTE module in Azure Machine Learning Studio to increase the number of underrepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

Box 3: Convert to Indicator Values 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.

Box 4: Remove Duplicate Rows

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

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

QUESTION 3

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 creating a new experiment in Azure Machine Learning Studio.

One class has a much smaller number of observations than the other classes in the training set.

You need to select an appropriate data sampling strategy to compensate for the class imbalance.

Solution: You use the Synthetic Minority Oversampling Technique (SMOTE) sampling mode.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: A

SMOTE is used to increase the number of underrepresented cases in a dataset used for machine learning. SMOTE is a better way of increasing the number of rare cases than simply duplicating existing cases.

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

QUESTION 4

HOTSPOT

You have a dataset that contains 2,000 rows. You are building a machine learning classification model by using Azure Learning Studio. You add a Partition and Sample module to the experiment.

You need to configure the module. You must meet the following requirements:

1.

Divide the data into subsets

2.

Assign the rows into folds using a round-robin method

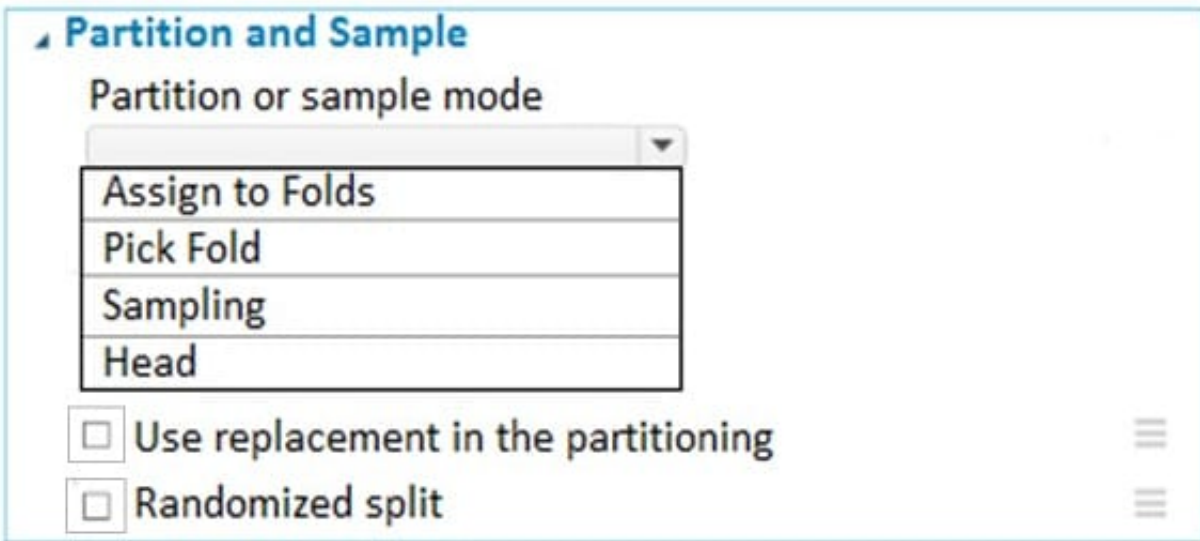
3.

Allow rows in the dataset to be reused

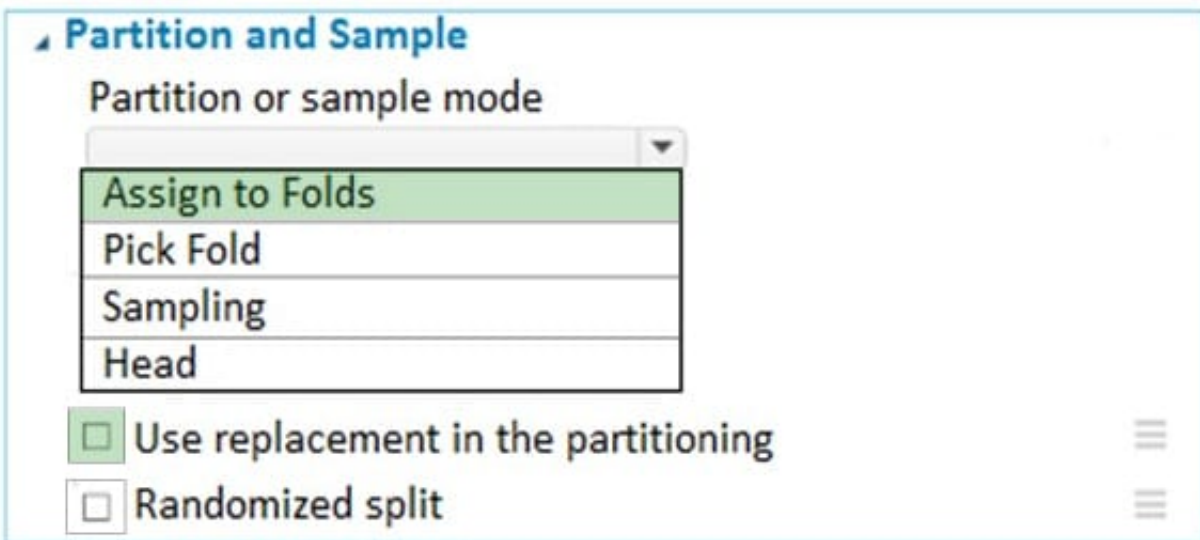
How should you configure the module? To answer, select the appropriate options in the dialog box in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:



Use the Split data into partitions option when you want to divide the dataset into subsets of the data. This option is also useful when you want to create a custom number of folds for cross-validation, or to split rows into several groups.

Add the Partition and Sample module to your experiment in Studio (classic), and connect the dataset.

For Partition or sample mode, select Assign to Folds.

Use replacement in the partitioning: Select this option if you want the sampled row to be put back into the pool of rows for potential reuse. As a result, the same row might be assigned to several folds.

If you do not use replacement (the default option), the sampled row is not put back into the pool of rows for potential reuse. As a result, each row can be assigned to only one fold.

Randomized split: Select this option if you want rows to be randomly assigned to folds.

If you do not select this option, rows are assigned to folds using the round-robin method.

References:

<https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/partition-and-sample>

QUESTION 5

You have recently concluded the construction of a binary classification machine learning model.

You are currently assessing the model. You want to make use of a visualization that allows for precision to be used as the measurement for the assessment.

Which of the following actions should you take?

- A. You should consider using Venn diagram visualization.
- B. You should consider using Receiver Operating Characteristic (ROC) curve visualization.
- C. You should consider using Box plot visualization.
- D. You should consider using the Binary classification confusion matrix visualization.

Correct Answer: D

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/how-to-understand-automated-ml#confusion-matrix>

[DP-100 PDF Dumps](#)

[DP-100 VCE Dumps](#)

[DP-100 Exam Questions](#)