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.pass4itsure.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

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 model to predict the price of a student\\'s artwork depending on the following variables:

the student\\'s length of education, degree type, and art form.

You start by creating a linear regression model.

You need to evaluate the linear regression model.

Solution: Use the following metrics: Accuracy, Precision, Recall, F1 score, and AUC.

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: B

Those are metrics for evaluating classification models, instead use: Mean Absolute Error, Root Mean Absolute Error, Relative Absolute Error, Relative Squared Error, and the Coefficient of Determination.

Reference: https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/evaluate-model

QUESTION 2

You are retrieving data from a large datastore by using Azure Machine Learning Studio.

You must create a subset of the data for testing purposes using a random sampling seed based on the system clock.

You add the Partition and Sample module to your experiment.

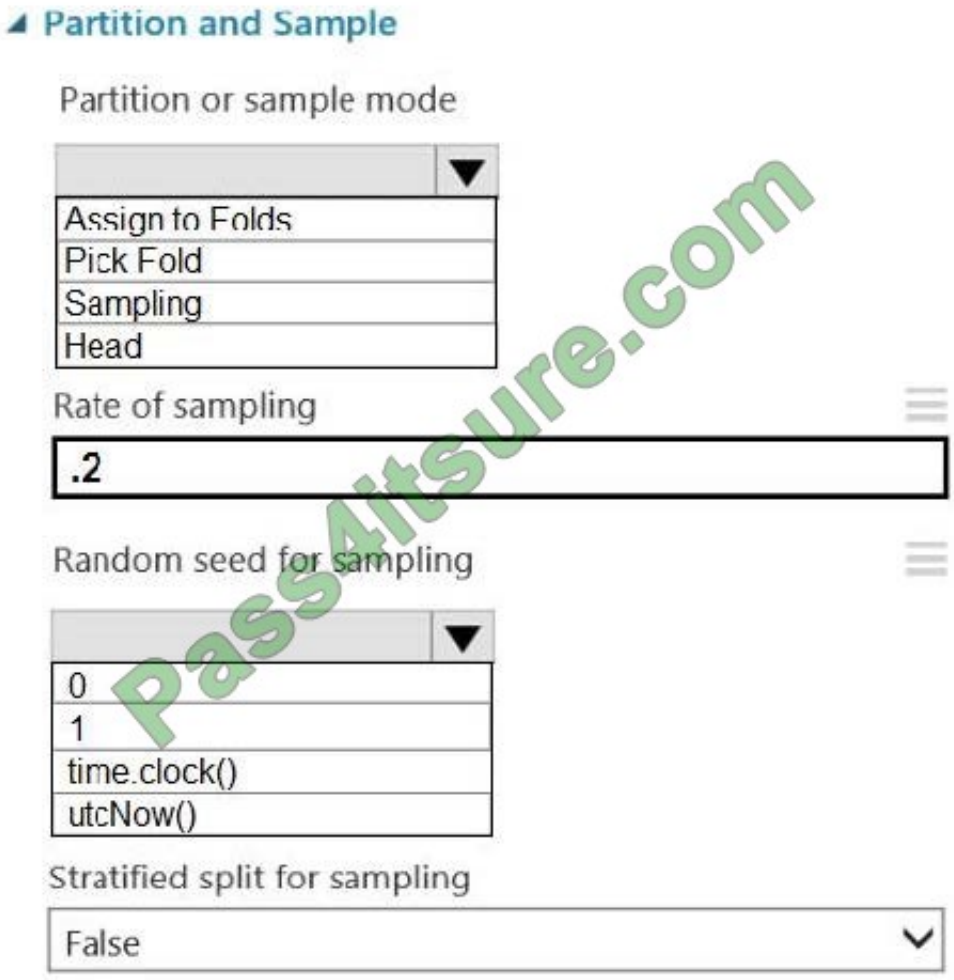
You need to select the properties for the module.

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

NOTE: Each correct selection is worth one point.

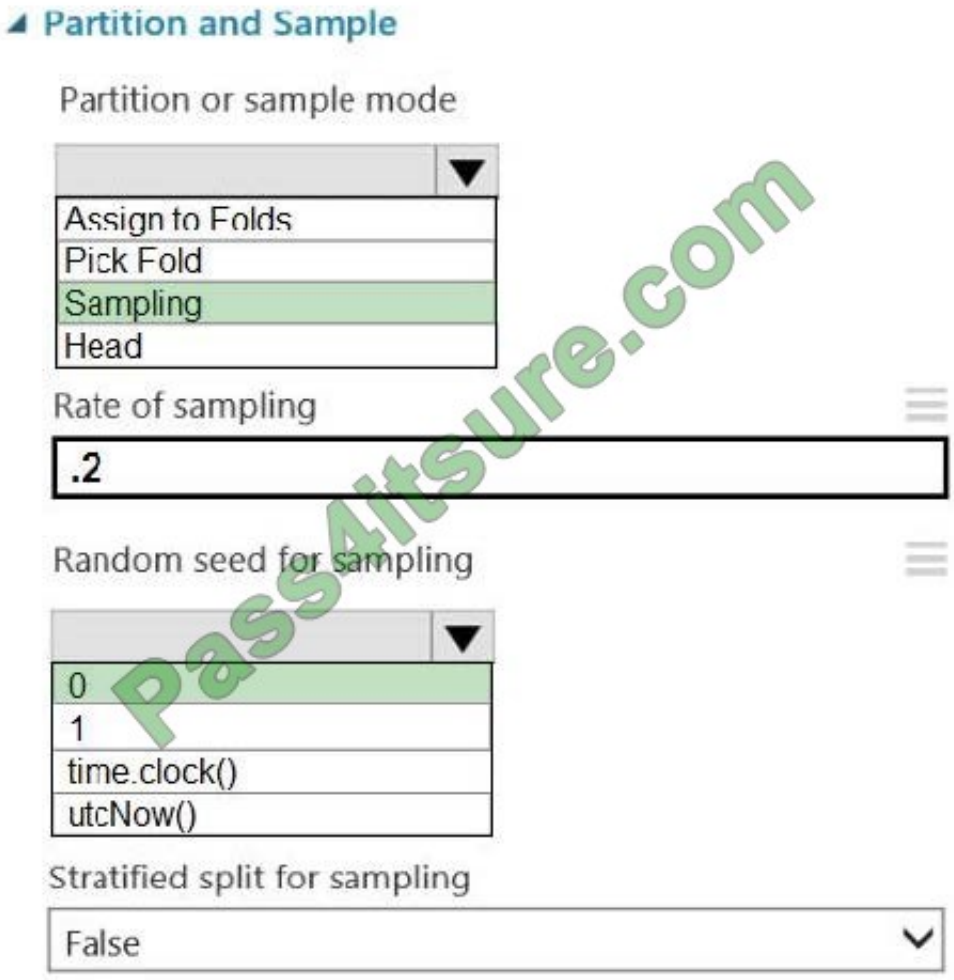
Hot Area:

Answer Area



Correct Answer:

Answer Area



Box 1: Sampling Create a sample of data This option supports simple random sampling or stratified random sampling. This is useful if you want to create a smaller representative sample dataset for testing.

1.

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

2.

Partition or sample mode: Set this to Sampling.

3.

Rate of sampling. See box 2 below.

Box 2: 0

3. Rate of sampling. Random seed for sampling: Optionally, type an integer to use as a seed value.

This option is important if you want the rows to be divided the same way every time. The default value is 0, meaning that a starting seed is generated based on the system clock. This can lead to slightly different results each time you run the experiment.

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

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 using Azure Machine Learning to run an experiment that trains a classification model.

You want to use Hyperdrive to find parameters that optimize the AUC metric for the model. You configure a HyperDriveConfig for the experiment by running the following code:

```
hyperdrive = HyperDriveConfig(estimator=your_estimator,
 hyperparameter_sampling=your_params,
 policy=policy,
 primary_metric_name='AUC',
 primary_metric_goal=PrimaryMetricGoal.MAXIMIZE,
 max_total_runs=6,
 max_concurrent_runs=4)
```

You plan to use this configuration to run a script that trains a random forest model and then tests it with validation data. The label values for the validation data are stored in a variable named y_test variable, and the predicted probabilities from

the model are stored in a variable named y_predicted.

You need to add logging to the script to allow Hyperdrive to optimize hyperparameters for the AUC metric.

Solution: Run the following code:

```
from sklearn.metrics import roc_auc_score
import logging
# code to train model omitted
auc = roc_auc_score(y_test, y_predicted)
logging.info("AUC: " + str(auc))
```

Does the solution meet the goal?

A. Yes

B. No

Correct Answer: A

Python printing/logging example: logging.info(message)

Destination: Driver logs, Azure Machine Learning designer

Reference: https://docs.microsoft.com/en-us/azure/machine-learning/how-to-debug-pipelines

QUESTION 4

HOTSPOT

You have an Azure blob container that contains a set of TSV files. The Azure blob container is registered as a datastore for an Azure Machine Learning service workspace. Each TSV file uses the same data schema.

You plan to aggregate data for all of the TSV files together and then register the aggregated data as a dataset in an Azure Machine Learning workspace by using the Azure Machine Learning SDK for Python.

You run the following code.

```
from azureml.core.workspace import Workspace
from azureml.core.datastore import Datastore
from azureml.core.dataset import Dataset
import pandas as pd
datastore_paths = (datastore, './data/*.tsv')
myDataset_1 = Dataset.File.from_files(path=datastore_paths)
myDataset_2 = Dataset.Tabular.from_delimited_files(path=datastore_paths, separator='\t')
```

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

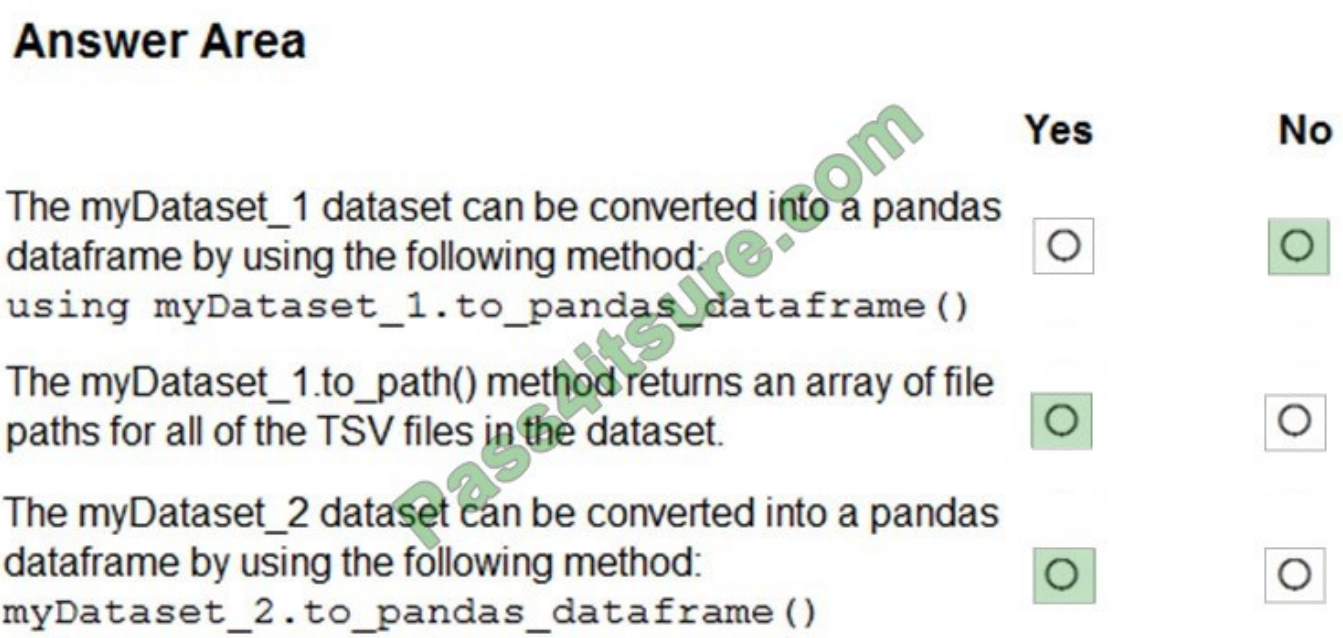
NOTE: Each correct selection is worth one point.

Hot Area:

Answer Area

	Yes	No
The myDataset_1 dataset can be converted into a pandas dataframe by using the following method: `using myDataset_1.to_pandas_dataframe()`	○	○
The myDataset_1.to_path() method returns an array of file paths for all of the TSV files in the dataset.	○	○
The myDataset_2 dataset can be converted into a pandas dataframe by using the following method: `myDataset_2.to_pandas_dataframe()`	○	○

Correct Answer:



Box 1: No

FileDataset references single or multiple files in datastores or from public URLs. The TSV files need to be parsed.

Box 2: Yes

to_path() gets a list of file paths for each file stream defined by the dataset.

Box 3: Yes

TabularDataset.to_pandas_dataframe loads all records from the dataset into a pandas DataFrame.

TabularDataset represents data in a tabular format created by parsing the provided file or list of files.

Note: TSV is a file extension for a tab-delimited file used with spreadsheet software. TSV stands for Tab Separated Values. TSV files are used for raw data and can be imported into and exported from spreadsheet software. TSV files are

essentially text files, and the raw data can be viewed by text editors, though they are often used when moving raw data between spreadsheets.

Reference:

https://docs.microsoft.com/en-us/python/api/azureml-core/azureml.data.tabulardataset

QUESTION 5

HOTSPOT

You deploy a model in Azure Container Instance.

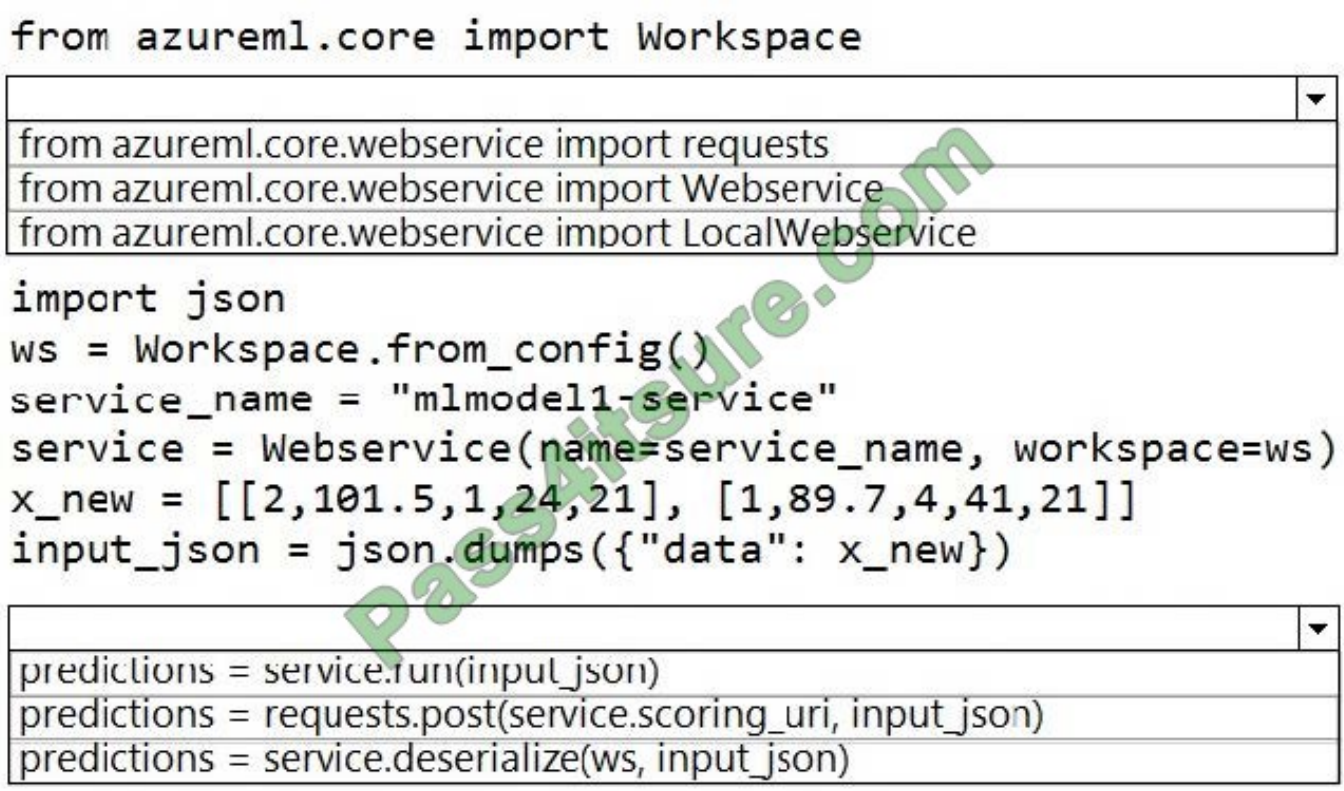
You must use the Azure Machine Learning SDK to call the model API.

You need to invoke the deployed model using native SDK classes and methods.

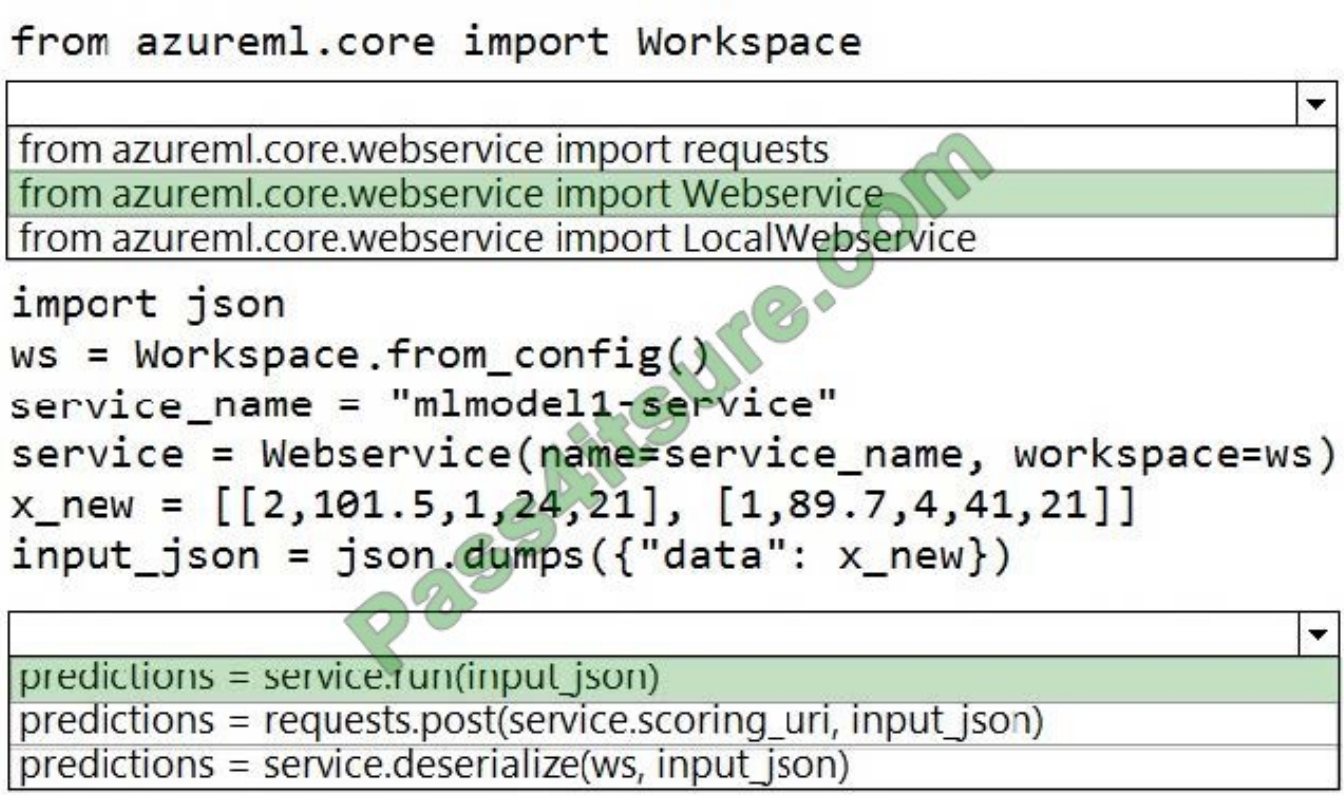
How should you complete the command? To answer, select the appropriate options in the answer areas.

NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:



Box 1: from azureml.core.webservice import Webservice

The following code shows how to use the SDK to update the model, environment, and entry script for a web service to Azure Container Instances:

from azureml.core import Environment

from azureml.core.webservice import Webservice

from azureml.core.model import Model, InferenceConfig

Box 2: predictions = service.run(input_json)

Example: The following code demonstrates sending data to the service:

import json

test_sample = json.dumps({\\'data\\': [[1, 2, 3, 4, 5, 6, 7, 8, 9, 10], [10, 9, 8, 7, 6, 5, 4, 3, 2, 1]

]})

test_sample = bytes(test_sample, encoding=\\'utf8\\')

prediction = service.run(input_data=test_sample) print(prediction)

Reference: https://docs.microsoft.com/bs-latn-ba/azure/machine-learning/how-to-deploy-azure-container-instance

https://docs.microsoft.com/en-us/azure/machine-learning/how-to-troubleshoot-deployment