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

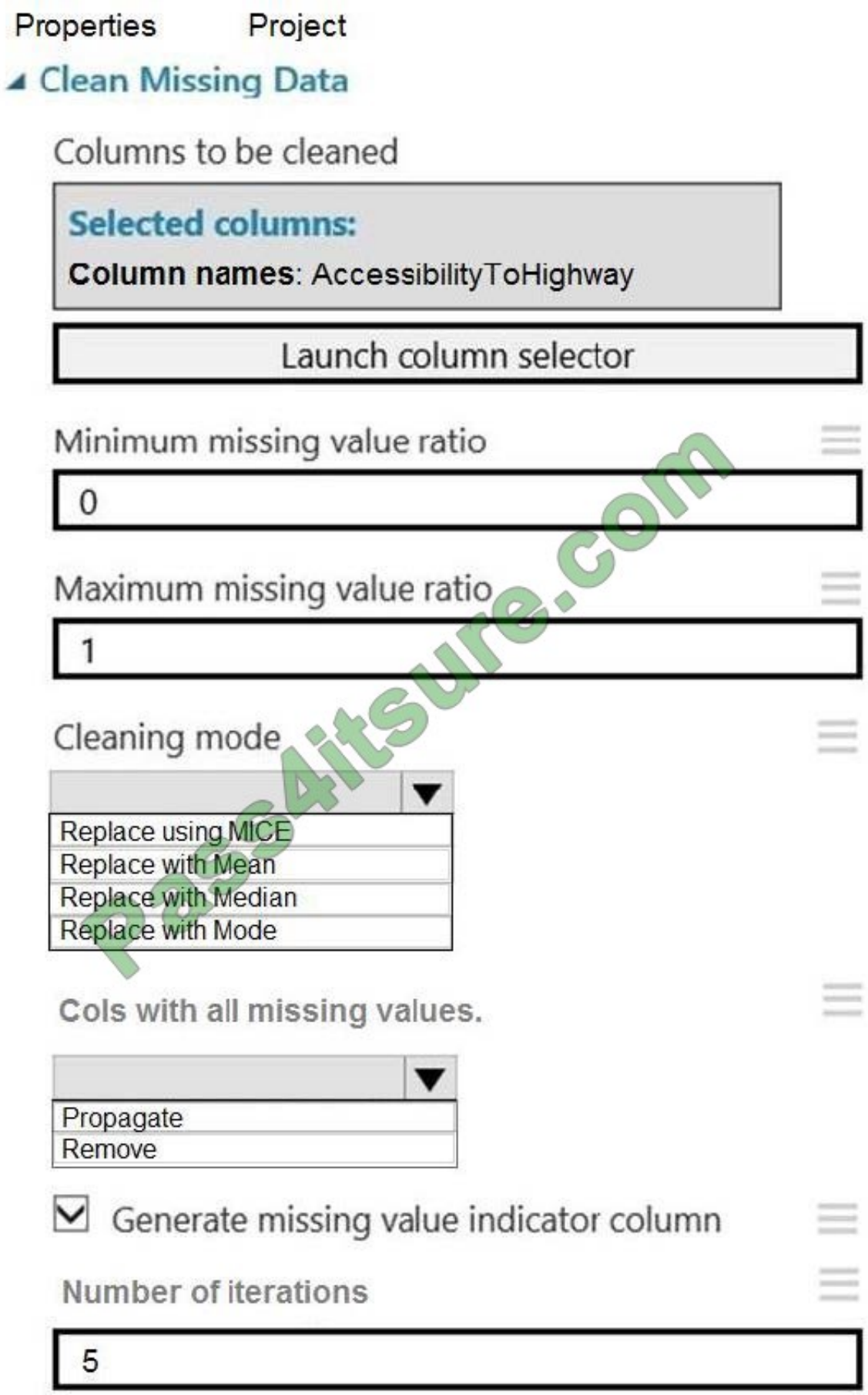
QUESTION 1

You need to replace the missing data in the AccessibilityToHighway columns.

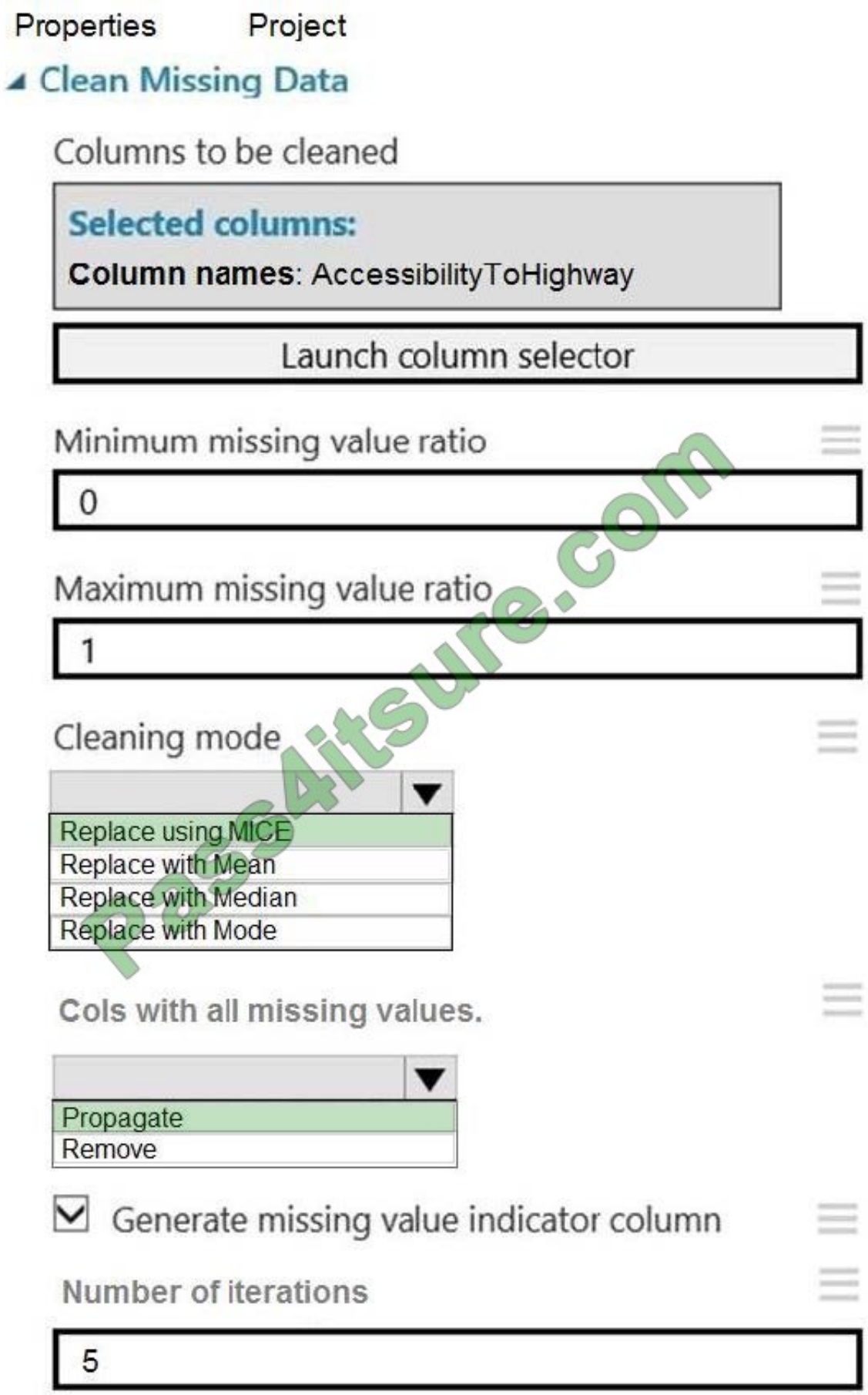
How should you configure the Clean Missing Data module? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:



Box 1: Replace using MICE

Replace using MICE: For each missing value, this option assigns a new value, which is calculated by using a method described in the statistical literature as "Multivariate Imputation using Chained Equations" or "Multiple Imputation by

Chained Equations". With a multiple imputation method, each variable with missing data is modeled conditionally using the other variables in the data before filling in the missing values.

Scenario: The AccessibilityToHighway column in both datasets contains missing values. The missing data must be replaced with new data so that it is modeled conditionally using the other variables in the data before filling in the missing

values.

Box 2: Propagate

Cols with all missing values indicate if columns of all missing values should be preserved in the output.

References:

https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data

---

QUESTION 2

You are creating a machine learning model. You have a dataset that contains null rows.

You need to use the Clean Missing Data module in Azure Machine Learning Studio to identify and resolve the null and missing data in the dataset.

Which parameter should you use?

A. Replace with mean

B. Remove entire column

C. Remove entire row

D. Hot Deck

E. Custom substitution value

F. Replace with mode

Correct Answer: C

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.

References: https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clean-missing-data

---

QUESTION 3

HOTSPOT

You plan to use Hyperdrive to optimize the hyperparameters selected when training a model. You create the following code to define options for the hyperparameter experiment:

```
import azureml.train.hyperdrive.parameter_expressions as pe
from azureml.train.hyperdrive import GridParameterSampling, HyperDriveConfig

param_sampling = GridParameterSampling({
  "max_depth" : pe.choice(6, 7, 8, 9),
  "learning_rate" : pe.choice(0.05, 0.1, 0.15)
  })
hyperdrive_run_config = HyperDriveConfig(
  estimator = estimator,
  hyperparameter_sampling = param_sampling,
  policy = None,
  primary_metric_name = "auc",
  primary_metruc_goal = PrimaryMetricGoal.MAXIMIZE,
  max_total_runs = 50,
  max_concurrent_runs = 4)
```

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

NOTE: Each correct selection is worth one point.

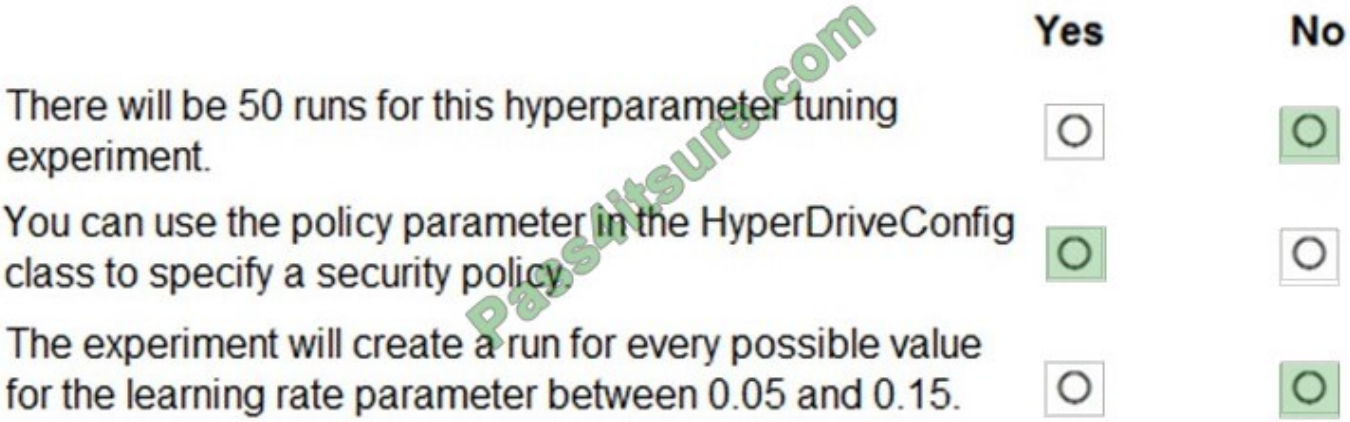
Hot Area:

Answer Area

	Yes	No
There will be 50 runs for this hyperparameter tuning experiment.	○	○
You can use the policy parameter in the HyperDriveConfig class to specify a security policy.	○	○
The experiment will create a run for every possible value for the learning rate parameter between 0.05 and 0.15.	○	○

Correct Answer:

Answer Area



Box 1: No

max_total_runs (50 here)

The maximum total number of runs to create. This is the upper bound; there may be fewer runs when the sample space is smaller than this value.

Box 2: Yes

Policy EarlyTerminationPolicy

The early termination policy to use. If None - the default, no early termination policy will be used.

Box 3: No

Discrete hyperparameters are specified as a choice among discrete values. choice can be:

1.

one or more comma-separated values

2.

a range object

3.

any arbitrary list object

Reference: https://docs.microsoft.com/en-us/python/api/azureml-train-core/azureml.train.hyperdrive.hyperdriveconfig
https://docs.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters

QUESTION 4

HOTSPOT

You are running a training experiment on remote compute in Azure Machine Learning.

The experiment is configured to use a conda environment that includes the mlflow and azureml-contrib-run packages.

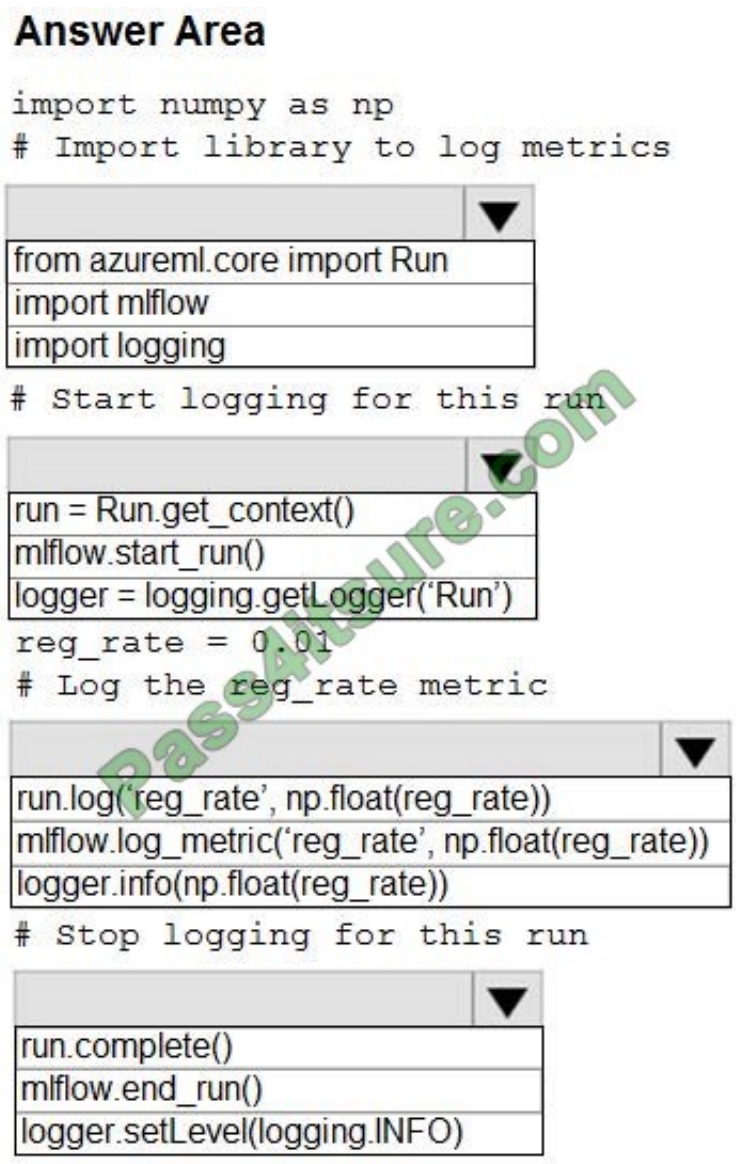
You must use MLflow as the logging package for tracking metrics generated in the experiment.

You need to complete the script for the experiment.

How should you complete the code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:

Answer Area

```
import numpy as np
# Import library to log metrics
```

[from azureml.core import Run | import mlflow | import logging]

```
# Start logging for this run
```

[run = Run.get_context() | mlflow.start_run() | logger = logging.getLogger('Run')]

```
reg_rate = 0.01
# Log the reg_rate metric
```

[run.log('reg_rate', np.float(reg_rate)) | mlflow.log_metric('reg_rate', np.float(reg_rate)) | logger.info(np.float(reg_rate))]

```
# Stop logging for this run
```

[run.complete() | mlflow.end_run() | logger.setLevel(logging.INFO)]

Box 1: import mlflow

Import the mlflow and Workspace classes to access MLflow\\'s tracking URI and configure your workspace.

Box 2: mlflow.start_run()

Set the MLflow experiment name with set_experiment() and start your training run with start_run().

Box 3: mlflow.log_metric(\\' ..\\')

Use log_metric() to activate the MLflow logging API and begin logging your training run metrics.

Box 4: mlflow.end_run()

Close the run:

run.endRun()

Reference:

https://docs.microsoft.com/en-us/azure/machine-learning/how-to-use-mlflow

QUESTION 5

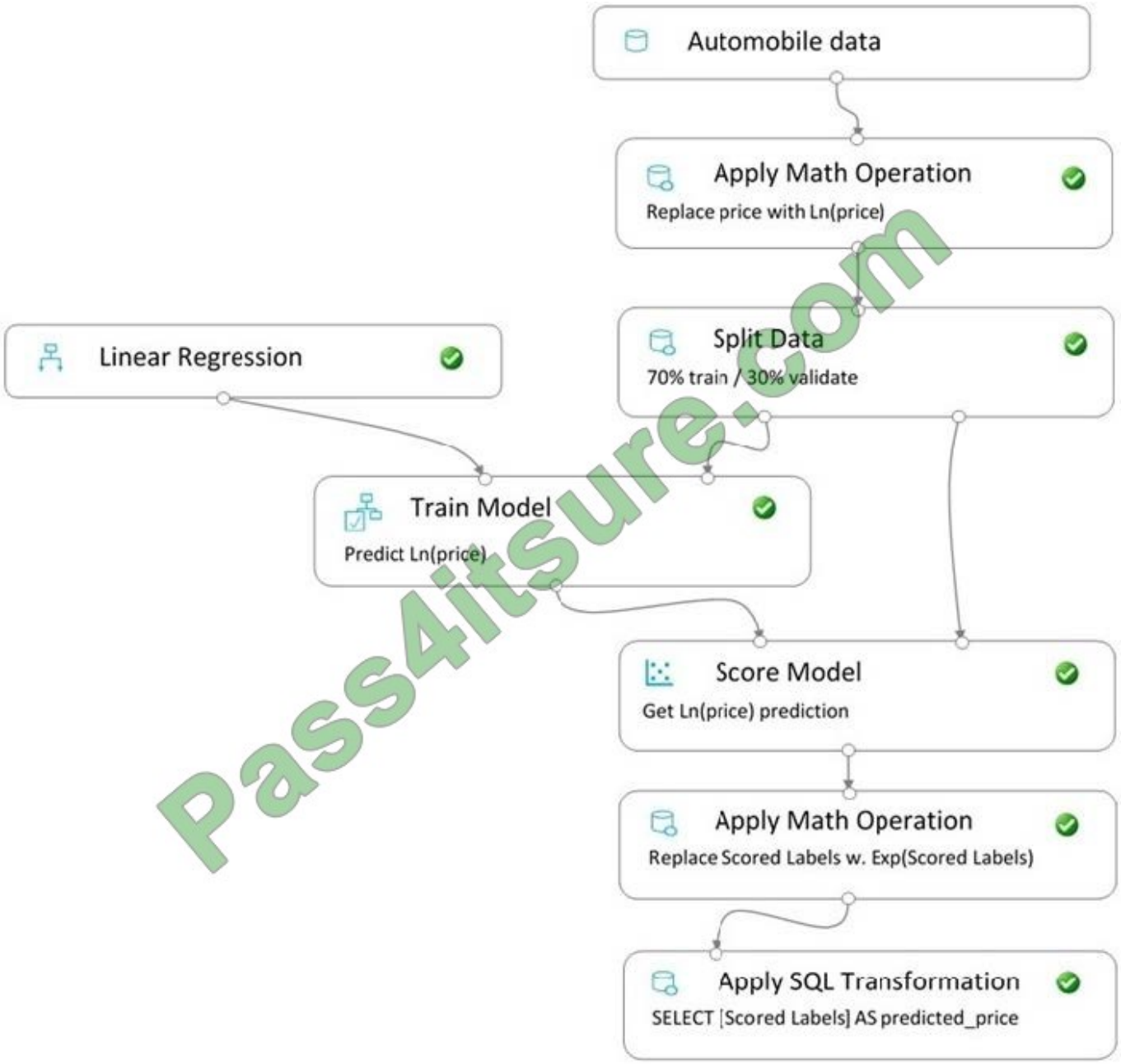
You create a pipeline in designer to train a model that predicts automobile prices.

Because of non-linear relationships in the data, the pipeline calculates the natural log (Ln) of the prices in the training data, trains a model to predict this natural log of price value, and then calculates the exponential of the scored label to get

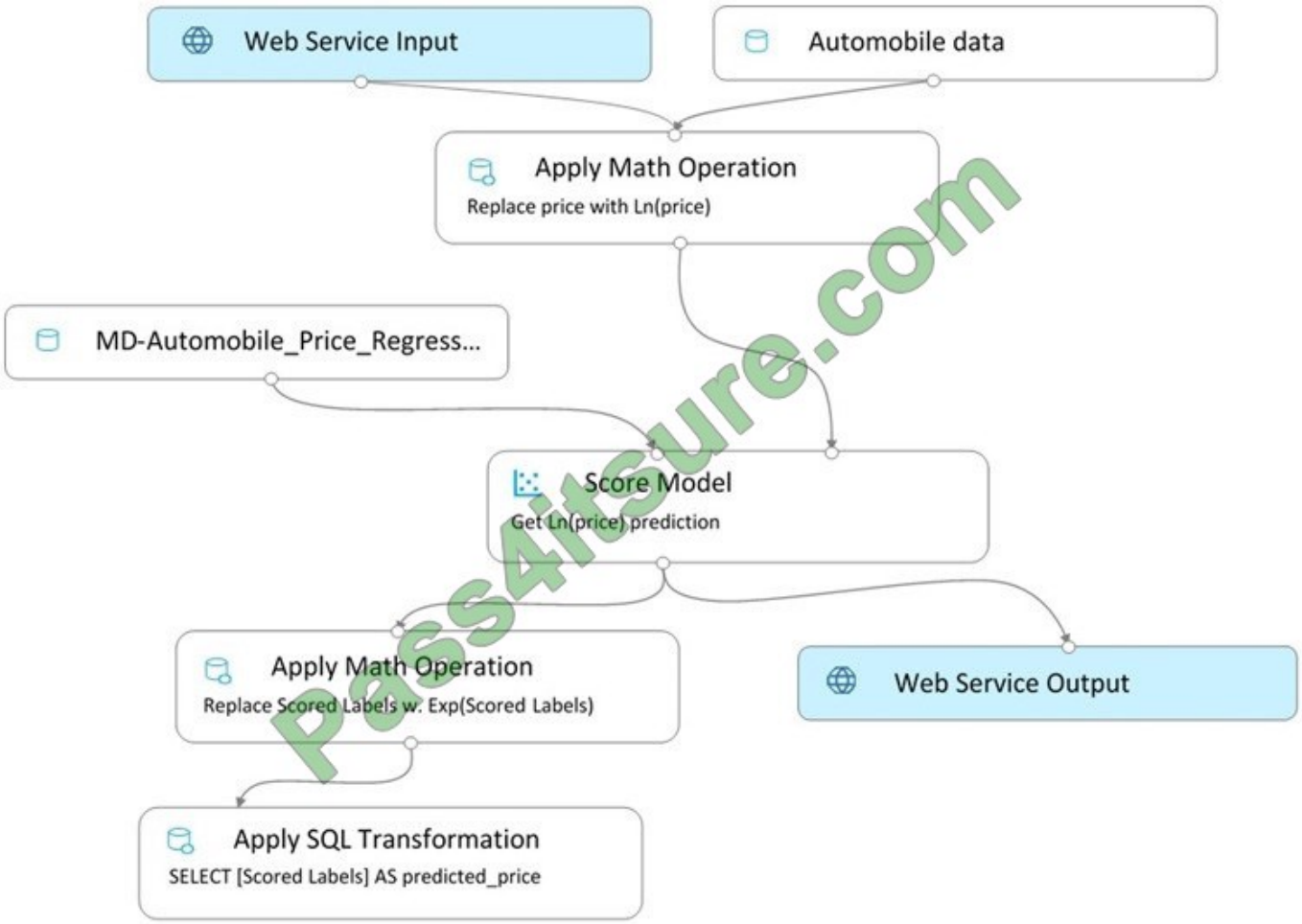
the predicted price.

The training pipeline is shown in the exhibit. (Click the Training pipeline tab.)

Training pipeline



You create a real-time inference pipeline from the training pipeline, as shown in the exhibit. (Click the Real-time pipeline tab.) Real-time pipeline



You need to modify the inference pipeline to ensure that the web service returns the exponential of the scored label as the predicted automobile price and that client applications are not required to include a price value in the input values.

Which three modifications must you make to the inference pipeline? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. Connect the output of the Apply SQL Transformation to the Web Service Output module.

B. Replace the Web Service Input module with a data input that does not include the price column.

C. Add a Select Columns module before the Score Model module to select all columns other than price.

D. Replace the training dataset module with a data input that does not include the price column.

E. Remove the Apply Math Operation module that replaces price with its natural log from the data flow.

F. Remove the Apply SQL Transformation module from the data flow.

Correct Answer: ACE