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.geekcert.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 performing clustering by using the K-means algorithm.

You need to define the possible termination conditions.

Which three conditions can you use? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Centroids do not change between iterations.

B. The residual sum of squares (RSS) rises above a threshold.

C. The residual sum of squares (RSS) falls below a threshold.

D. A fixed number of iterations is executed.

E. The sum of distances between centroids reaches a maximum.

Correct Answer: ACD

AD: The algorithm terminates when the centroids stabilize or when a specified number of iterations are completed.

C: A measure of how well the centroids represent the members of their clusters is the residual sum of squares or RSS, the squared distance of each vector from its centroid summed over all vectors. RSS is the objective function and our goal is to minimize it.

Reference: https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/k-means-clustering

https://nlp.stanford.edu/IR-book/html/htmledition/k-means-1.html

QUESTION 2

HOTSPOT

You create an Azure Machine Learning workspace named workspace1. You assign a custom role to a user of workspace1.

The custom role has the following JSON definition:

```
{
 "Name": "MyRole",
 "IsCustom": true,
 "Description": "New custom role description.",
 "Actions": ["*"],
 "NotActions": [
   "Microsoft.MachineLearningServices/workspaces/write",
   "Microsoft.MachineLearningServices/workspaces/computes/*/write",
   "Microsoft.MachineLearningServices/workspaces/computes/*/delete",
   "Microsoft. Authorization/*/write"
 ],
 "AssignableScopes": [
   "/subscriptions/<subscription_id>/resourceGroups/resourcegroup1/providers/
   Microsoft.MachineLearningServices/workspaces/workspace1"
 ]
}
```

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

NOTE: Each correct selection is worth one point.

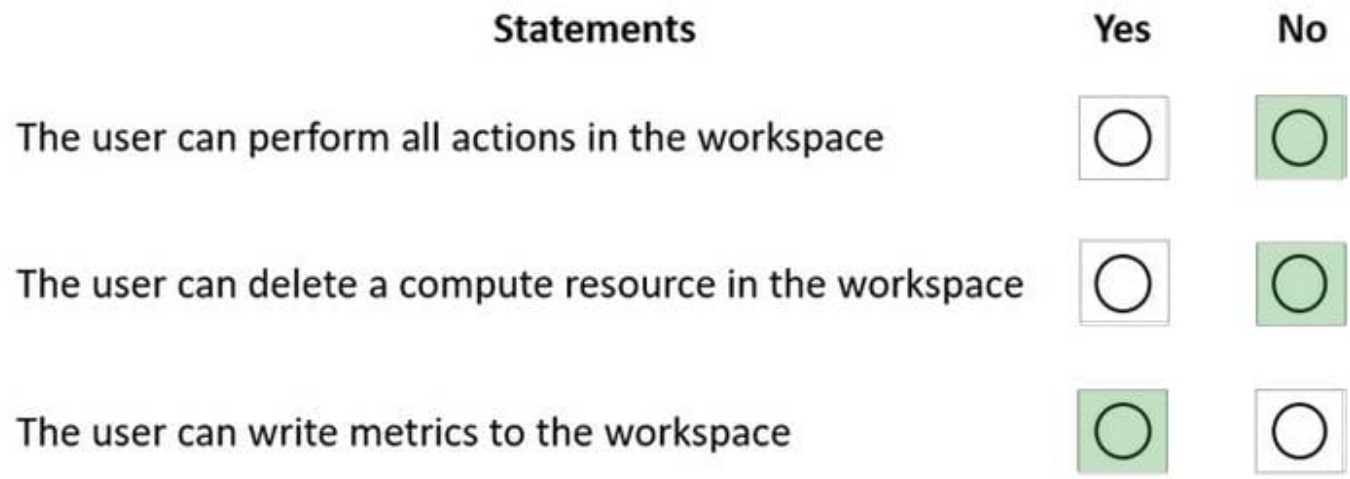
Hot Area:

Answer Area

Statements	Yes	No
The user can perform all actions in the workspace	○	○
The user can delete a compute resource in the workspace	○	○
The user can write metrics to the workspace	○	○

Correct Answer:

Answer Area



Box 1: No

The actions listed in NotActions are prohibited.

If the roles include Actions that have a wildcard (*), the effective permissions are computed by subtracting the NotActions from the allowed Actions.

Box 2: No

Deleting compute resources in the workspace is in the NotActions list.

Box 3: Yes

Writing metrics is not listed in NotActions.

Reference:

https://docs.microsoft.com/en-us/azure/role-based-access-control/overview#how-azure-rbac-determines-if-a-user-has-access-to-a-resource

QUESTION 3

You need to implement a feature engineering strategy for the crowd sentiment local models. What should you do?

A. Apply an analysis of variance (ANOVA).

B. Apply a Pearson correlation coefficient.

C. Apply a Spearman correlation coefficient.

D. Apply a linear discriminant analysis.

Correct Answer: D

The linear discriminant analysis method works only on continuous variables, not categorical or ordinal variables.

Linear discriminant analysis is similar to analysis of variance (ANOVA) in that it works by comparing the means of the variables.

Scenario:

Data scientists must build notebooks in a local environment using automatic feature engineering and model building in machine learning pipelines.

Experiments for local crowd sentiment models must combine local penalty detection data.

All shared features for local models are continuous variables.

Incorrect Answers:

B: The Pearson correlation coefficient, sometimes called Pearson\\'s R test, is a statistical value that measures the linear relationship between two variables. By examining the coefficient values, you can infer something about the strength of the relationship between the two variables, and whether they are positively correlated or negatively correlated.

C: Spearman\\'s correlation coefficient is designed for use with non-parametric and non-normally distributed data. Spearman\\'s coefficient is a nonparametric measure of statistical dependence between two variables, and is sometimes denoted by the Greek letter rho. The Spearman\\'s coefficient expresses the degree to which two variables are monotonically related. It is also called Spearman rank correlation, because it can be used with ordinal variables.

References: https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/fisher-linear-discriminant-analysis

https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/compute-linear-correlation

QUESTION 4

HOTSPOT

You train a machine learning model by using Aunt Machine Learning.

You use the following training script m Python to log an accuracy value.

```
from azureml.core.run import Run
run_logger = Run.get_context()
run_logger.log("accuracy", float(val_accuracy))
```

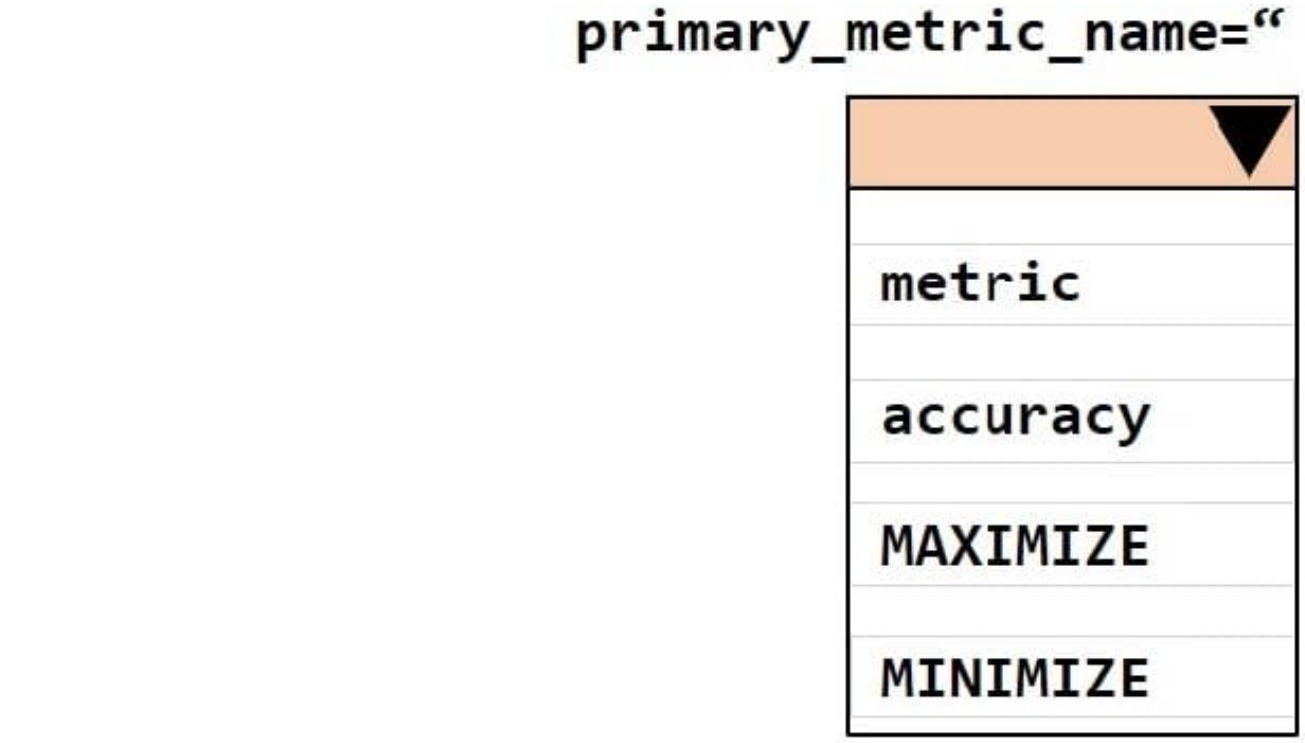
You must use a Python script to define a sweep job.

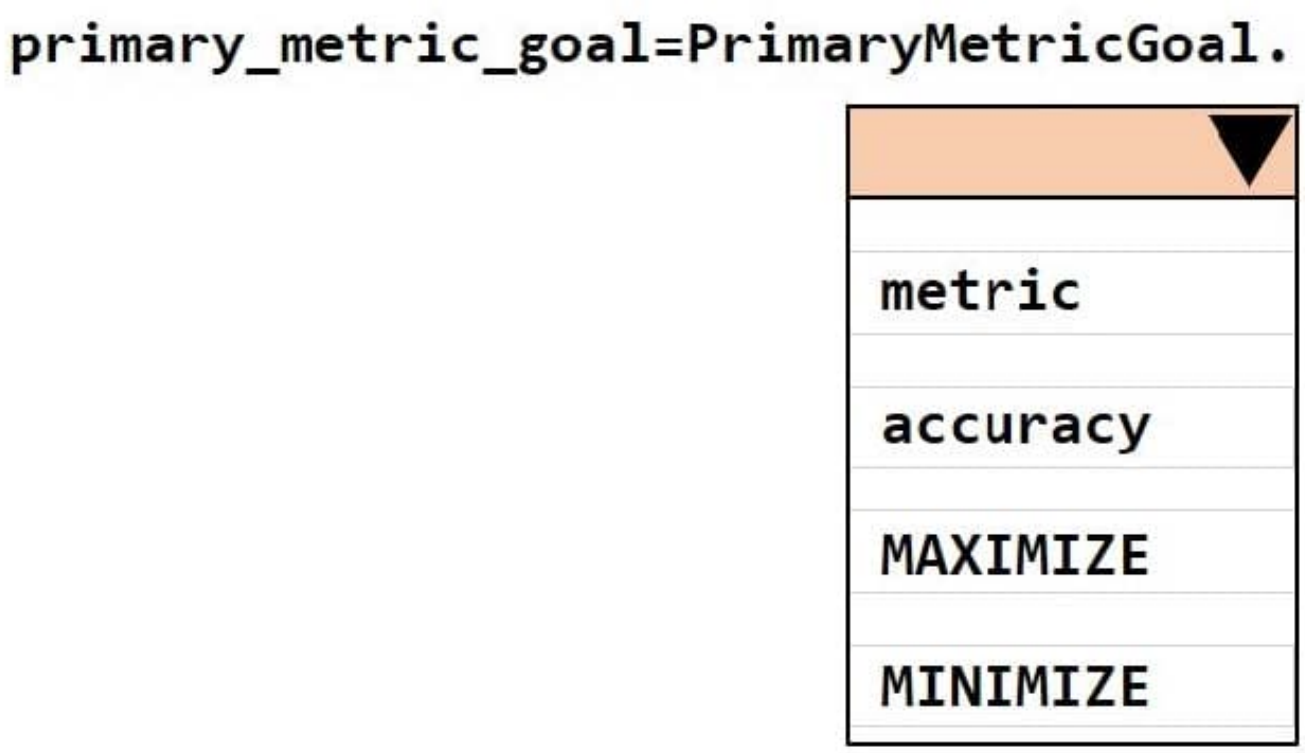
You need to provide the primary metric and goal you want hyper parameter tuning to optimize.

How should you complete the Python script? To answer select the appropriate options in the answer area

NOTE: Each correct selection is worth one point.

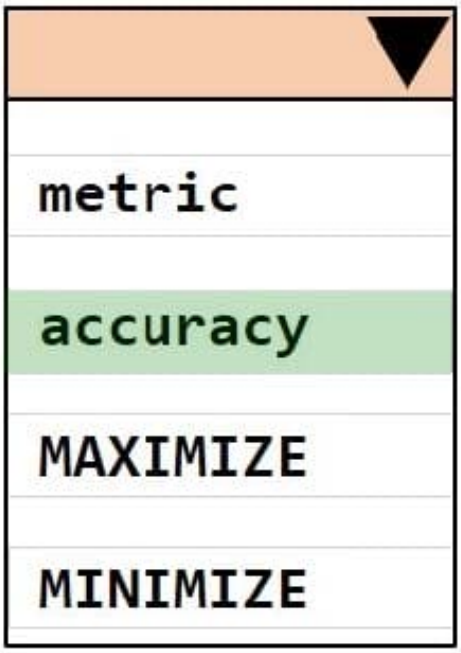
Hot Area:



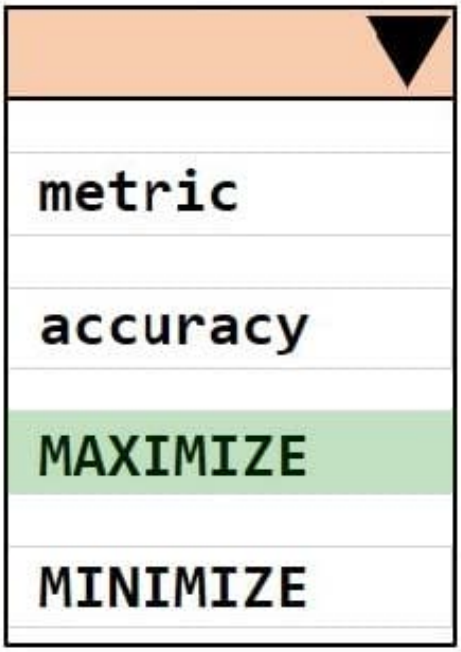


Correct Answer:

primary_metric_name="



primary_metric_goal=PrimaryMetricGoal.



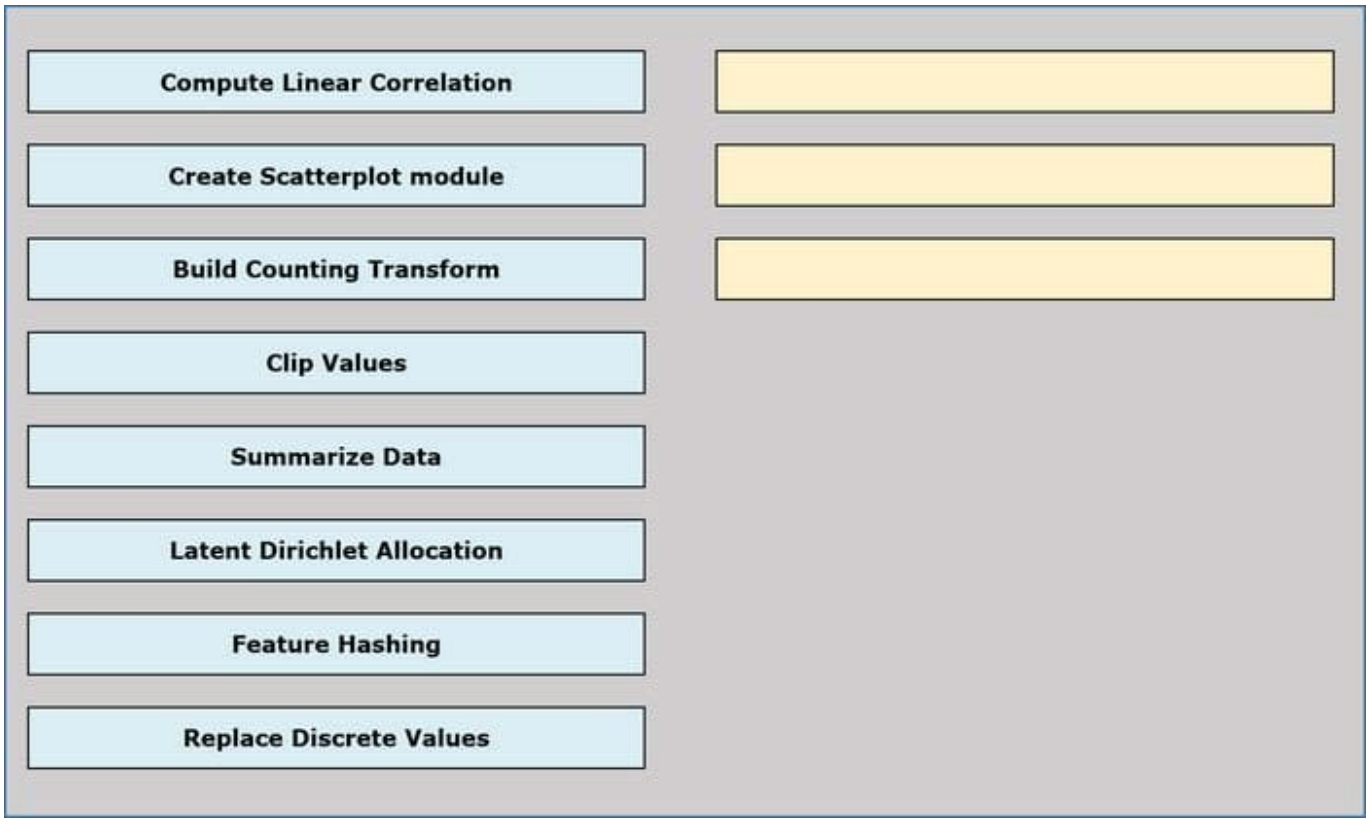
QUESTION 5

DRAG DROP

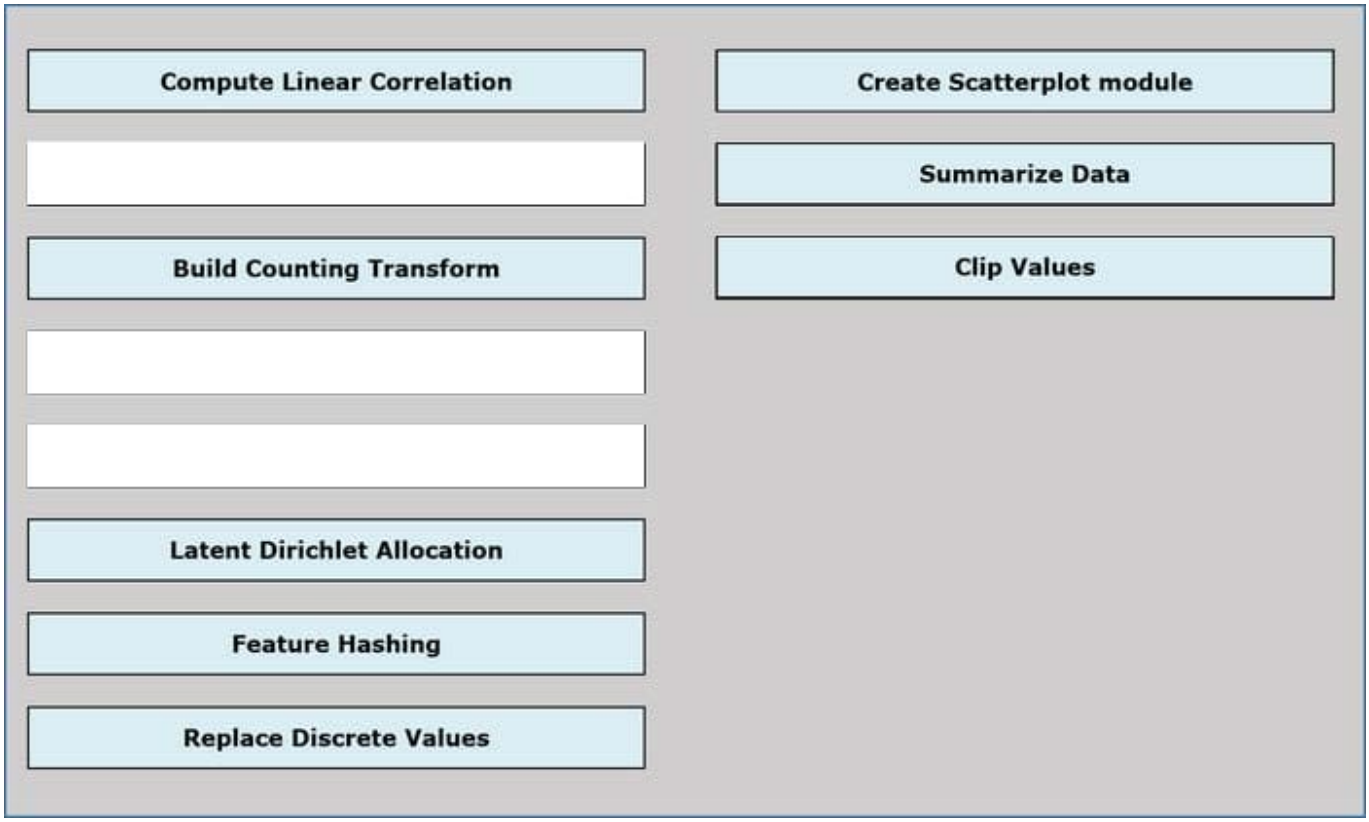
You need to visually identify whether outliers exist in the Age column and quantify the outliers before the outliers are removed.

Which three Azure Machine Learning Studio modules should you use in sequence? To answer, move the appropriate modules from the list of modules to the answer area and arrange them in the correct order.

Select and Place:



Correct Answer:



You can use the Clip Values module in Azure Machine Learning Studio, to identify and optionally replace data values that are above or below a specified threshold. This is useful when you want to remove outliers or replace them with a mean, a constant, or other substitute value.

References: https://blogs.msdn.microsoft.com/azuredev/2017/05/27/data-cleansing-tools-in-azuremachine-learning/ https://docs.microsoft.com/en-us/azure/machine-learning/studio-module-reference/clipvalues