

PRACTICE QUESTIONS

SAS

A00-485

Modeling Using SAS Visual Statistics

SHARED BY [u/Plucky-Marmot-3573](#) on aiexamcerts

EDITION	Free Set 1
QUESTIONS	10
SHARED	October 15, 2026
EXAM	SAS A00-485, Modeling Using SAS Visual Statistics

Shared by [u/Plucky-Marmot-3573](#) on aiexamcerts

10 practice questions for the SAS A00-485 exam, Modeling Using SAS Visual Statistics, each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/sas/a00-485. Spotted a mistake? Email support@aiexamcerts.com with the exam code A00-485 and the question number.

1. Which principal component analysis step is most useful before modeling many correlated numeric inputs?
 - A Sort the inputs alphabetically by name and then keep only the first few columns for the model
 - B Create new uncorrelated components from the inputs to reduce dimensionality
 - C Replace each input with its frequency count
 - D Delete every input that has a missing value

Answer: B

Principal components combine correlated numeric inputs into a smaller set of uncorrelated variables. Replacing values with frequency counts is the tempting distractor, because it is a common encoding, but it is designed for categories rather than correlation structure. Sorting names and deleting inputs do not reduce correlated dimensions.

2. A numeric measurement such as age or income is called a _____ variable, while a category such as region is nominal.

Answer: continuous

Continuous measurements take many numeric values and support arithmetic such as means and regression slopes. Nominal variables only name categories and must be encoded before regression. Discrete is the tempting alternative, but it describes countable values with a small set of levels.

3. Which situation most clearly calls for a regularized logistic model?
 - A A single predictor with a clean linear trend
 - B A binary outcome with no missing values and no outliers
 - C Many correlated predictors relative to the number of events
 - D A response that is already a continuous measurement

Answer: C

Many correlated predictors against few events leads to unstable estimates, and a penalty such as the lasso or ridge stabilizes them. A single predictor with a clean trend gives no reason to shrink coefficients, which is the tempting distractor. The absence of missing values and outliers and a continuous response say nothing about the need for regularization in a logistic model.

4. Model version control and documentation help auditors trace which model produced a given score.
- A True
 - B False

Answer: A (True)

Versioned models and documentation let each score be traced to a specific model and its training data. Overwriting the model in place is the tempting shortcut, and it removes that traceability. Traceability is a common governance requirement for scored decisions.

5. Which two practices help keep a deployed scoring model reliable over time? (Choose two.)
- A Retrain on the scoring data's own labels at every run without review
 - B Track input distributions and prediction outcomes against the training baseline
 - C Delete the training data once the model is deployed
 - D Remove all version information from the model
 - E Re-evaluate and document the model when its performance drops

Answer: B, E

Monitoring inputs and outcomes against a baseline reveals drift early, and re-evaluation with documentation keeps the model accountable. Automatic retraining without review is the tempting distractor, because it sounds like maintenance, but it can silently accept bad labels. Deleting training data and removing version information both weaken the ability to diagnose and reproduce results.

6. Which metric does the area under the ROC curve summarize?
- A The model's ability to rank positive cases above negative cases across all cutoffs
 - B The total number of misclassified cases at a cutoff of 0.5
 - C The proportion of variance explained by the model
 - D The average residual of the training partition

Answer: A

AUC is the probability that a random positive case scores higher than a random negative case, which summarizes ranking across all cutoffs. The misclassification count at 0.5 is the tempting distractor, because it is a single-cutoff view, but AUC does not depend on one cutoff. Variance explained belongs to linear regression, and training residuals say nothing about ranking.

7. In gradient boosting, how is each new tree built?
- A It is fitted on a random sample of the target, ignoring the predictors
 - B It replaces the previous trees entirely
 - C It is fitted to the errors, or gradients, left by the trees before it
 - D It uses only the first split of the earlier trees

Answer: C

Gradient boosting adds trees in sequence, and each new tree models what the current ensemble still gets wrong. Fitting on a random sample of the target ignores the predictors, which is the tempting distractor. Replacing earlier trees would discard the sequential correction that boosting depends on.

8. Limiting a tree's _____ restricts the number of levels from the root to a leaf, which reduces its ability to fit small, noisy patterns.

Answer: depth (also accepted: tree depth, maximum depth)

Depth is the number of splits on the longest path from root to leaf, and limiting it is a standard way to control overfitting. Width is the tempting alternative, since the number of leaves also matters, but the question asks about path length. Depth and minimum leaf size are the two settings most often tuned together.

9. What is the main purpose of a residual plot in linear regression diagnostics?
- A To list the coefficients from largest to smallest with standard errors
 - B To look for patterns such as curvature or changing spread that signal violated assumptions
 - C To split the data into training and validation partitions
 - D To compute the p-value of the intercept alone

Answer: B

A residual plot shows whether residuals scatter randomly around zero, and systematic curves or funnels suggest a missing term or nonconstant variance. Listing coefficients is the tempting distractor because coefficient tables appear next to residual output, but it does not test assumptions. Splitting the data and computing one p-value are separate tasks.

10. A coefficient is statistically significant, but its sign is opposite to what domain knowledge predicts. Which explanation should be checked first?
- A The intercept is too large to interpret, which makes slopes flip sign
 - B Strong correlation with another predictor, or a coding error in the variable
 - C The model contains too many observations
 - D The p-value of the intercept is exactly one

Answer: B

Collinearity and miscoded variables can flip a coefficient's sign even when it is significant, so they are the first things to check. A large intercept is the tempting distractor, but it does not change the sign of a slope. The number of observations affects precision, not the direction of an effect, and an intercept p-value does not describe a slope.