

PRACTICE QUESTIONS

CompTIA DataAI

SHARED BY [u/Mellow-Tiger-4451](#) on aiexamcerts

EDITION	Free Set 1
QUESTIONS	10
SHARED	October 15, 2026
EXAM	CompTIA DataAI

Shared by [u/Mellow-Tiger-4451](#) on aiexamcerts

10 practice questions for the CompTIA DataAI exam, each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/comptia/comptia-dataai. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. Which statement about the central limit theorem is correct?
 - A Individual observations become normally distributed once the dataset has more than 30 rows.
 - B The sample median converges to zero as the number of observations grows.
 - C The distribution of sample means tends toward a normal distribution as sample size grows, whatever the population shape, provided its variance is finite.
 - D The sample standard deviation always exceeds the population standard deviation.

Answer: C

The central limit theorem concerns the distribution of sample means, which approaches a normal shape as the sample size increases, even when the population is skewed. Believing that individual observations become normal once the sample is large enough is the most tempting mistake, because it confuses raw data with the sampling distribution of the mean. The statement about sample medians and standard deviations is not a property the theorem provides.

2. When a sorted dataset is divided into four equal parts, the three boundary values that separate those parts are called _____.

Answer: quartiles

Quartiles are the values at the 25th, 50th and 75th percentiles, which split ordered data into four groups of equal count. The interquartile range is built from the first and third quartiles. Deciles and percentiles divide the data into ten or one hundred parts, so they are not the correct term here.

3. Which technique estimates model performance across several train and test partitions by rotating which subset is held out?
 - A Principal component analysis
 - B K-fold cross-validation
 - C One-hot encoding
 - D Min-max scaling

Answer: B

K-fold cross-validation splits the data into k folds and uses each fold once as the held-out set, which gives a more stable performance estimate than a single split. Principal component analysis reduces dimensionality and does not estimate performance. One-hot encoding and min-max scaling are preprocessing steps and say nothing about evaluation.

4. L2 regularization penalizes the sum of squared coefficient values, which shrinks coefficients toward zero but usually does not set them exactly to zero.

- A True
- B False

Answer: A (True)

The L2 penalty adds the squared magnitudes of the coefficients to the loss, which pushes them smaller but rarely makes them exactly zero. L1 regularization, by contrast, can drive some coefficients to exactly zero and performs feature selection. The statement is true.

5. Which two statements about reducing overfitting in neural networks are correct? (Choose two.)

- A Increasing model complexity always reduces overfitting.
- B Dropout randomly disables units during training to reduce co-adaptation between them.
- C Removing the validation set makes overfitting easier to detect.
- D Data augmentation increases the effective variety of training examples.
- E Training for more epochs without monitoring is a standard regularizer.

Answer: B, D

Dropout disables a random subset of units on each training step, which discourages units from relying on one another, and augmentation creates varied versions of the same examples. Adding complexity usually increases overfitting, and removing validation data hides it. Training longer without monitoring is a path to overfitting rather than a remedy.

6. Which activation function is most commonly used in hidden layers of modern deep networks because it is cheap to compute and avoids vanishing gradients for positive inputs?

- A Sigmoid
- B Step function
- C ReLU
- D Linear identity

Answer: C

ReLU outputs zero for negative inputs and the input itself for positive ones, so its gradient is constant on the positive side and cheap to compute. Sigmoid saturates at both ends, which causes the vanishing gradient problem the question describes. A step function has zero gradient almost everywhere, and a linear identity stacks into a single linear map.

7. Which storage pattern keeps raw data in its original format in low-cost storage and applies a schema when the data is read?
- A Data lake
 - B Operational OLTP database
 - C Star-schema data mart
 - D In-memory key-value cache

Answer: A

A data lake stores raw files cheaply in open formats and leaves interpretation to schema-on-read at query time. An operational transactional database enforces structure on write and serves application transactions. Star-schema marts and in-memory caches are curated or transient stores, so they do not keep raw data in its original form.

8. Monitoring a deployed model for changes in the distribution of its input data is commonly called detecting data _____.

Answer: drift

Data drift means the inputs arriving in production differ from those seen in training, which can degrade accuracy even when the model code is unchanged. Concept drift is a related change in the relationship between inputs and targets. Drift is detected by comparing distributions over time rather than by checking the model file.

9. What is the primary goal of object detection?
- A Producing a single label for the whole image.
 - B Compressing the image to a smaller file size.
 - C Estimating the camera's GPS position.
 - D Locating and classifying multiple objects in an image with bounding boxes.

Answer: D

Object detection outputs a class and a bounding box for each object it finds, so it answers both what and where. Whole-image labeling is image classification, which is the most tempting confusion. Compression and camera localization are separate tasks.

10. A time series has an upward trend and a repeating weekly cycle. What is a common first step before fitting an ARIMA model?
- A Difference or otherwise transform the series so that it is stationary.
 - B Remove the time index from the data.
 - C Shuffle the observations to remove the pattern.
 - D Convert the numeric values into categorical labels such as low, medium and high.

Answer: A

ARIMA assumes a stationary series, so differencing removes trend and, with seasonal differencing, a weekly cycle before the model is fitted. Removing the time index or shuffling the observations destroys the temporal structure the model needs. Converting to categories discards the numeric information that forecasting depends on.