

PRACTICE QUESTIONS

DASCA

Senior Data Scientist

SHARED BY [u/Lucky-Compass-7853](#) on aiexamcerts

EDITION Free Set 2

QUESTIONS 10

SHARED October 15, 2026

EXAM DASCA Senior Data Scientist

Shared by [u/Lucky-Compass-7853](#) on aiexamcerts

10 practice questions for the DASCA Senior Data Scientist exam, each followed by its answer and a short explanation. More free sets and the full practice exam: www.aiexamcerts.com/dasca/senior-data-scientist. Spotted a mistake? Email support@aiexamcerts.com with the exam name and the question number.

1. A logistic regression coefficient has a 95 percent confidence interval that includes zero. Which interpretation is correct?
 - A The predictor has a large effect on the outcome.
 - B The data do not provide strong evidence that the coefficient differs from zero at the 5 percent level.
 - C The model is misspecified and must be discarded.
 - D The predictor's effect is exactly zero in the population, so the coefficient can safely be dropped from the model.

Answer: B

An interval that contains zero means the test at the 5 percent level does not reject a zero effect, which is weak evidence rather than proof of no effect. Concluding that the effect is exactly zero is the tempting overreach, because failing to reject is not the same as showing the effect is absent. A wide interval also does not indicate a large effect, and a single interval does not show misspecification.

2. The Central Limit Theorem requires the population distribution to be normal.
 - A True
 - B False

Answer: B (False)

The theorem requires only independent observations with finite variance, and the sampling distribution of the mean approaches normality for large samples regardless of the population shape. Assuming normal data is a common misreading. The statement is false.

3. Which practice best prevents hyperparameters from being tuned against the holdout test data?
 - A Reporting the test score after every tuning iteration and keeping the best configuration.
 - B Running grid search directly on the test split.
 - C Nested cross-validation with an inner tuning loop and an outer evaluation loop.
 - D Averaging the test predictions across several models.

Answer: C

Nested cross-validation keeps tuning inside the inner loop, so the outer folds remain unseen by the selection process. Reporting the test score after every tuning iteration is the tempting habit, since the same test score then guides tuning. Averaging predictions does not address selection bias.

4. Adding small random perturbations to training inputs as a regularizer is called input _____.

Answer: noise injection (also accepted: noise-injection)

Noise injection perturbs inputs during training so the model learns smoother functions that are less sensitive to small changes. Data augmentation is a broader term that also includes transformations such as flips or crops. Dropout is a different regularizer that disables units rather than perturbing inputs.

5. Which two statements about file sizing and partitioning are correct? (Choose two.)

- A Many tiny files slow query planning and reads because of per-file overhead.
- B Partitioning on a high-cardinality column such as user ID always improves performance.
- C Compacting small files is never useful once a table has been written.
- D Partition keys should match the most common filter predicates.
- E Compressing files prevents any query from using column statistics.

Answer: A, D

Each file adds listing and open overhead, so many tiny files slow planning and reads, and compaction fixes this. Partition keys pay off when queries filter on them, and high-cardinality partitions create many small partitions that hurt performance. Compression does not remove the statistics stored in Parquet footers.

6. What does a scaled dot-product attention head compute?

- A Gradient descent updates to the query, key and value weights for each token in the batch.
- B A weighted sum of value vectors, with softmax weights from scaled query-key scores.
- C The element-wise maximum of two embeddings.
- D A bag-of-words count of the input.

Answer: B

The head scores each query against every key, scales the scores, applies softmax, and uses the resulting weights to combine the value vectors. Gradient descent is the training procedure rather than the head's computation, which makes it a tempting confusion. Element-wise maxima and word counts are not part of attention.

7. Which statement about pretraining and fine-tuning is correct?
- A Fine-tuning is always done from random weights.
 - B Pretraining requires a human label for every token.
 - C Fine-tuning replaces the architecture and discards all of the pretrained weights before it starts.
 - D Pretraining learns general representations, and fine-tuning adapts them to a narrower task.

Answer: D

Pretraining uses large unlabeled or weakly labeled corpora to learn general language or visual representations, and fine-tuning reuses those weights for a specific task. Training from random weights is the opposite of fine-tuning, which makes it the tempting error. Fine-tuning keeps the architecture and adjusts its weights.

8. A watermark in stream processing lets the system wait indefinitely for late events, so no event is ever dropped.
- A True
 - B False

Answer: B (False)

A watermark is an estimate of event-time progress that bounds how long the system waits for late events before finalizing a window, and events arriving after that bound can be dropped or routed separately. The statement is false.

9. Which infrastructure approach gives teams reproducible training and serving environments?
- A Installing packages manually on each laptop.
 - B Pinning dependencies in a lockfile and building a versioned container image.
 - C Using the latest version of every library on each run to stay current with upstream.
 - D Running jobs without recording their configuration.

Answer: B

A lockfile and a versioned image define one environment that can be rebuilt identically across teams and time. Installing packages manually on each laptop is the tempting convenience, but machines drift apart. Floating versions and unrecorded configurations make results impossible to reproduce.

10. Which cost-control approach is most appropriate for an online endpoint with spiky traffic?
- A Keeping the maximum number of replicas running at all times, day and night, regardless of traffic.
 - B Disabling autoscaling altogether.
 - C Serving every request from a single CPU core.
 - D Autoscaling replicas on concurrency or queue length, with a nonzero minimum.

Answer: D

Scaling on concurrency or queue depth adds capacity when spikes arrive and removes it when they pass, and a non-zero minimum avoids cold-start delays. Keeping the maximum running is the tempting safe choice, but it pays for idle capacity. Disabling autoscaling or starving the endpoint of capacity harms either cost or latency.