

PRACTICE QUESTIONS

NVIDIA

NCA-ADS

NVIDIA-Certified Associate: Accelerated Data Science

SHARED BY [u/Grand-Switch-6290](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 18, 2026

EXAM NVIDIA NCA-ADS, NVIDIA-Certified Associate: Accelerated Data Science

Shared by u/Grand-Switch-6290 on aiexamcerts

10 practice questions for the NVIDIA NCA-ADS exam, NVIDIA-Certified Associate: Accelerated Data Science, each followed by its answer and a short explanation. More free sets and the full practice exam:

aiexamcerts.com/nvidia/nca-ads. Spotted a mistake? Email support@aiexamcerts.com with the exam code NCA-ADS and the question number.

1. Which plot best shows the distribution of a single numeric variable, including its spread and skew?
 - A Pie chart
 - B Line chart over time
 - C Scatter plot showing two variables
 - D Histogram

Answer: D

A histogram bins the values and shows the shape of the distribution, including spread and skew. A scatter plot is tempting because it looks similar, but it displays the relationship between two variables. A line chart shows change over time, and a pie chart shows parts of a whole.

2. A pipeline stores large analytical tables in a columnar, compressed format named _____.

Answer: Parquet (also accepted: Apache Parquet)

Parquet stores each column together, so queries that read a few columns avoid scanning the rest, and its compression reduces storage. CSV is a row-oriented text format that is often confused with it.

3. Which two algorithms available in RAPIDS cuML are unsupervised? (Choose two.)
 - A Linear regression
 - B K-means clustering
 - C Logistic regression
 - D Principal component analysis
 - E Random forest classifier

Answer: B, D

K-means groups unlabeled points into clusters, and principal component analysis finds directions of variance without using a target. Linear regression, logistic regression and random forest classification all learn from labeled targets.

4. Which tool is most suitable for running a multi-step data pipeline so that each task starts only after its upstream tasks succeed?
- A A workflow orchestrator such as Apache Airflow that defines task dependencies
 - B A text editor with saved notes and a manual checklist for each run
 - C A spreadsheet macro that someone runs by hand each morning after checking the inputs
 - D A GPU driver update

Answer: A

Orchestrators such as Apache Airflow model a pipeline as tasks with dependencies, retries and schedules. Hand-run spreadsheet macros and notes give no dependency control, and a driver update does not run any pipeline at all.

5. Which two statements about overfitting are correct? (Choose two.)
- A Adding more noise features is the standard fix for overfitting
 - B A model that scores far better on training data than on held-out data is likely overfitting
 - C Overfitting means the model fails on both training and test data
 - D Regularization, such as L2 penalties, can reduce overfitting
 - E Overfitting only happens in neural networks

Answer: B, D

A large gap between training and held-out performance signals overfitting, and regularization limits model complexity to reduce it. Adding noise features makes overfitting worse. Failing on both sets describes underfitting, which is the usual confusion.

6. Which two statistics are robust to a few extreme outliers? (Choose two.)
- A Mean
 - B Median
 - C Standard deviation
 - D Interquartile range
 - E Maximum

Answer: B, D

The median depends only on the middle rank, and the interquartile range depends only on the middle half of the data, so a few extreme values barely move either one. The mean is the most tempting choice, but a single large value pulls it upward. Standard deviation and maximum are both inflated by outliers.

7. A reproducible pipeline produces the same outputs from the same inputs and code, so random seeds should be fixed or recorded.
- A True
 - B False

Answer: A (True)

Fixing or recording seeds makes random steps repeatable, which is a core part of reproducibility alongside versioned data and code. Seeds do not by themselves guarantee identical results across different hardware or library versions.

8. In pandas, the method used to change a DatetimeIndex to a regular frequency such as daily is called _____.

Answer: resample

The resample method groups time-indexed rows into regular bins and then aggregates or interpolates them. Reindexing can also produce a regular index, but it does not aggregate rows by time bin.

9. A sensor table has timestamps at irregular intervals. Which operation best prepares it for a model that expects evenly spaced rows?
- A Sort it by the sensor name only
 - B Resample to a fixed frequency, then aggregate or interpolate the gaps
 - C Drop the timestamp column entirely and keep only the raw sensor readings for the model
 - D Convert each timestamp to a text label

Answer: B

Resampling to a fixed frequency produces evenly spaced rows, and aggregation or interpolation fills the gaps by a documented rule. Sorting by sensor name ignores time order. Dropping the timestamps removes the time structure the model needs.

10. Which metric equals true positives divided by the sum of true positives and false positives?
- A Recall
 - B Accuracy
 - C Precision
 - D Specificity

Answer: C

Precision measures how many predicted positives were actually positive. Recall divides true positives by the sum of true positives and false negatives, which is the most tempting confusion. Accuracy and specificity use the true negatives as well.