

PRACTICE QUESTIONS

Microsoft

DP-800

Microsoft Certified: SQL AI Developer Associate

SHARED BY [u/Quick-Yak-7653](#) on aiexamcerts

EDITION	Free Set 1
QUESTIONS	10
SHARED	October 15, 2026
EXAM	Microsoft DP-800, Microsoft Certified: SQL AI Developer Associate

Shared by [u/Quick-Yak-7653](#) on aiexamcerts

10 practice questions for the Microsoft DP-800 exam, Microsoft Certified: SQL AI Developer Associate, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/microsoft/dp-800. Spotted a mistake? Email support@aiexamcerts.com with the exam code DP-800 and the question number.

1. Which artifact stores the endpoint and deployment that SQL uses to call an external embedding model?

- A A database trigger that runs an embedding call on each insert
- B A full-text catalog
- C An external model definition
- D A partition function

Answer: C

An external model definition holds the endpoint, deployment, and model metadata that AI functions use. A trigger runs on data changes, a full-text catalog supports keyword search, and a partition function divides table data. None of those describes a model endpoint.

2. The number of passages retrieved for each question is usually called the top-_____ value.

Answer: k

Top-k sets how many nearest passages are returned to the prompt builder. A larger value raises recall but adds tokens and possible noise, and a smaller value risks missing the answer. Teams tune it against evaluation questions.

3. Which metric helps evaluate whether the retrieved passages contain the answer to a question?

- A Average row size in bytes
- B Index fragmentation percent
- C Number of tokens per embedding
- D Recall at k

Answer: D

Recall at k measures how often the relevant passage appears among the top k retrieved results, which is the core retrieval question. Row size and fragmentation describe storage, and token counts describe cost. None of those shows whether the answer was retrieved.

4. Azure SQL runs embedding models inside the database engine, so an external model definition is never needed.
- A True
 - B False

Answer: B (False)

Embedding generation in Azure SQL calls a hosted endpoint that is registered as an external model, and functions such as `AI_GENERATE_EMBEDDINGS` send the text there. The engine does not host the model itself. The statement is false because the external model definition is what makes the call possible.

5. Which two design choices keep an embedding table maintainable when the embedding model changes? (Choose two.)
- A Record the model name and version in a column for each row
 - B Store a hash or version of the source text so stale embeddings can be detected
 - C Overwrite old embeddings without recording which model created them
 - D Use one global embedding for every document regardless of model
 - E Drop the source text after it has been embedded

Answer: A, B

Storing the model name and version lets you find rows that need regeneration, and a text hash or version shows when the source has changed. Overwriting without provenance, sharing one embedding across documents, and dropping source text all make later migrations unreliable.

6. Which setting ensures that calls from SQL to an Azure OpenAI resource avoid the public internet?
- A A private endpoint for the Azure OpenAI resource, reached through a private DNS zone
 - B A public IP on the SQL server that any client on the internet can reach
 - C A firewall rule that allows every IP address to connect to the server
 - D An HTTP-only connection string with transport encryption turned off

Answer: A

A private endpoint gives the Azure OpenAI resource a private address in your virtual network, and a private DNS zone makes the name resolve to it. A public IP or an open firewall exposes the path to the internet. An HTTP-only connection string removes transport encryption.

7. Which SQL feature captures query runtime statistics and plan history inside a user database for regression analysis?
- A The default trace only
 - B Query Store
 - C Transaction log backups
 - D Database Mail

Answer: B

Query Store keeps per-query runtime statistics and plan history in the database, which supports regression analysis and plan forcing. The default trace records events on the server, transaction log backups protect recovery, and Database Mail sends messages. Only Query Store captures plan history for queries.

8. A _____ key on the chunk table enforces that each chunk belongs to an existing document.

Answer: foreign

A foreign key from the chunk table to the document table stops orphan chunks from being inserted or left behind. It also lets citations join back to the source document reliably. Enforcing this in the database is more dependable than relying on application code alone.

9. A query must return the five most similar chunks. Which clause shape does that?
- A TOP (5) with ORDER BY VECTOR_DISTANCE('cosine', @query, Embedding) DESC
 - B LIMIT 5 with GROUP BY Embedding
 - C TOP (5) with ORDER BY VECTOR_DISTANCE('cosine', @query, Embedding) ASC
 - D HAVING VECTOR_DISTANCE('cosine', @query, Embedding) < 5 only

Answer: C

Ordering by ascending cosine distance puts the closest vectors first, and TOP limits the result to five rows. Descending order returns the least similar chunks, LIMIT is not T-SQL, and a HAVING filter does not rank or limit the results.

10. A column holds national identification numbers that analysts should not see in full. Which control hides them in query results while keeping the data stored?
- A Dynamic data masking on the column
 - B A clustered index on the column, which orders rows by its values
 - C A default constraint of zero on the column for rows that omit it
 - D A check constraint on the length of the column values

Answer: A

Dynamic data masking applies a masking function to the column in results for users without UNMASK, while the stored value is unchanged. An index changes access paths, a default fills missing values, and a check constraint validates input. None of those hides values.