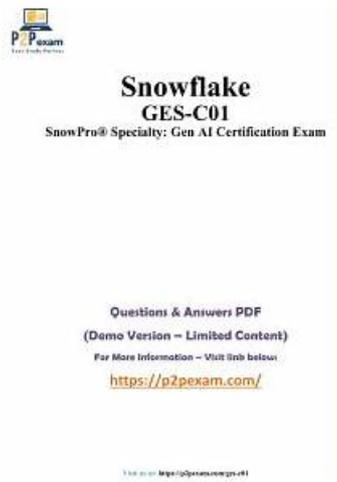


Guaranteed GES-C01 Questions Answers, GES-C01 Dump Collection



P.S. Free 2026 Snowflake GES-C01 dumps are available on Google Drive shared by DumpsReview:
https://drive.google.com/open?id=1AX_O9kX0vIo9jQSNnPMqh0HpVwp7YcEb

Our users are all over the world and they have completed their exams through the help of our GES-C01 study guide. As you can see the feedbacks from our loyal customers, all of them are grateful to our GES-C01 exam braindumps and become successful people with the GES-C01 Certification. And what are you waiting for? Just selecting our GES-C01 learning materials, the next one to get an international certificate is you!

We always adhere to the principle of 'mutual development and benefit', and we believe our GES-C01 practice materials can give you a timely and effective helping hand whenever you need in the process of learning our GES-C01 study braindumps. For we have been in this career over ten years and we are good at tracing the changes of the GES-C01 guide prep in time and update our exam dumps fast and accurately.

>> **Guaranteed GES-C01 Questions Answers** <<

GES-C01 Dump Collection, GES-C01 Certification Dumps

Holding a SnowPro® Specialty: Gen AI Certification Exam GES-C01 Certification in a certain field definitely shows that one have a good command of the GES-C01 knowledge and professional skills in the related field. However, it is universally accepted that the majority of the candidates for the SnowPro® Specialty: Gen AI Certification Exam exam are those who do not have enough spare time and are not able to study in the most efficient way.

Snowflake SnowPro® Specialty: Gen AI Certification Exam Sample Questions (Q124-Q129):

NEW QUESTION # 124

A data engineering team is optimizing an AI-infused pipeline that processes millions of rows of customer interaction data in a LOG_DATA table using various Snowflake Cortex AI functions. They need to accurately estimate costs and ensure optimal performance. Which of the following statements regarding cost, performance, and operational considerations for these functions are true?

- When using AI_COMPLETE or COMPLETE for generating responses, both input and output tokens are billable. The total tokens processed increase proportionally if previous conversation history is passed for a stateful experience.
- For functions like AI_CLASSIFY or AI_SENTIMENT, only input tokens are counted towards the billable total, as these functions do not generate new text but rather categorize or score existing text.
- The AI_AGG function is not subject to context window limitations when aggregating text columns across multiple rows, allowing it to process large volumes of text without exceeding token limits.
- Executing queries that call Cortex AI functions such as AI_COMPLETE or AI_AGG in an X-LARGE or larger warehouse will significantly improve their processing speed and reduce overall latency due to increased compute resources.
- For Cortex Analyst, credit usage is based on the number of messages processed, regardless of the number of tokens in each message, with billing occurring only for successful responses (HTTP 200).

- A. Option C
- B. Option D
- C. Option A
- D. Option E
- E. Option B

Answer: A,C,D

Explanation:

Option A is correct. For functions that generate new text like AI_COMPLETE and COMPLETE, both input and output tokens are billable. In conversational experiences, passing history increases processed tokens and costs proportionally. Option B is incorrect. For functions like AI_CLASSIFY and AI_SENTIMENT, the input token count is higher than the text provided because they add a prompt to the input text to generate the response. Furthermore, for these functions (AI_CLASSIFY, AI_FILTER, AI_AGG, AI_SUMMARIZE, and TRANSLATE, and their previous versions), both input and output tokens are billable. Therefore, the statement that "only" input tokens are counted is false. Option C is correct. AI_AGG and AI_SUMMARIZE_AGG are explicitly stated as not being subject to context window limitations, allowing them to aggregate insights across multiple rows. Option D is incorrect. Snowflake recommends using smaller warehouses (no larger than MEDIUM) for Cortex AI function queries, as larger warehouses do not increase performance. Option E is correct. For Cortex Analyst, credit usage is based on the number of messages processed, not tokens, and billing occurs only for successful responses (HTTP 200).

NEW QUESTION # 125

A financial institution is deploying a sentiment analysis application that uses Snowflake Cortex 'SENTIMENT' and 'COMPLETE' functions, with different LLMs, for processing customer feedback. They are using AI Observability (Public Preview) to compare the cost-efficiency of using 'mistral-7b' versus 'claude-3-5-sonnet' as LLM judges for evaluation metrics, and also tracking the overall cost of their AI Observability usage. Which statements accurately reflect the cost implications and monitoring tools for this scenario?

- The 'CORTEX_DOCUMENT_PROCESSING_USAGE_HISTORY' view should be used to monitor the credit consumption of the LLM judges specifically, as it tracks all Cortex function calls.
- Comparing models with vastly different context windows (e.g., 'mistral-7b' at 32k tokens vs. 'claude-3-5-sonnet' at 200k tokens) using AI Observability will not impact the billed 'tokens processed' if the actual prompt sizes are small and similar.
- In addition to LLM judge costs, warehouse charges are incurred for tasks managing evaluation runs and for queries used to compute evaluation metrics within AI Observability.
- The 'METERING_DAILY_HISTORY' view, filtering by 'SERVICE_TYPE ILIKE '%ai_services%', can provide an overview of daily credit usage for all AI services, including AI Observability LLM judge costs and evaluation metrics.

- A. Option D
- B. Option C
- C. Option A
- D. Option E
- E. Option B

Answer: A,C,D

Explanation:

Option A is correct because AI Observability utilizes LLM judges (such as 'mistral-7b' or 'claude-3-5-sonnet') through 'COMPLETE (SNOWFLAKE.CORTEX)' function calls to compute evaluation metrics, and these calls incur charges based on the 'tokens processed'. Option D is correct as, beyond LLM judge costs, AI Observability also incurs warehouse charges for managing evaluation runs and for queries that compute evaluation metrics. Option E is correct because the view, with a filter for 'SERVICE _TYPE ILIKE', provides a comprehensive daily credit usage report for all AI services, which would include AI Observability's components. Option B is incorrect; the view is specifically for Document AI processing functions like '!PREDICT and 'AI_EXTRACT', not for general LLM judge usage in AI Observability. The view is more appropriate for tracking individual Cortex function calls. Option C is incorrect because while prompt sizes might be similar, the pricing for different LLMs (e.g., 'mistral-7b' at 0.12 credits per million tokens vs. 'claude-3-5-sonnet' at 2.55 credits per million tokens for AI Complete) will still result in different billed amounts due to varying per-token costs, even if the number of tokens is the same.

NEW QUESTION # 126

A data scientist is leveraging various Snowflake Cortex LLM functions to process extensive text data for an application. To effectively manage their budget, they need a clear understanding of how costs are incurred for each specific function. Which of the following statements accurately describe how costs are calculated for Snowflake Cortex LLM functions, with a particular focus on token usage?

The SNOWFLAKE.CORTEX.EMBED_TEXT_768 function incurs compute costs based on both the input and output tokens processed.

For the SNOWFLAKE.CORTEX.EXTRACT_ANSWER function, billable tokens are determined by the sum of tokens present in the source_document and question fields.

When utilizing AI_COMPLETE with structured outputs, an additional compute cost is explicitly charged for the overhead of validating each generated token against the provided JSON schema.

The SNOWFLAKE.CORTEX.CLASSIFY_TEXT function includes the token count from category descriptions and examples as part of the input tokens for each record processed, thereby increasing the overall cost.

The AI_PARSE_DOCUMENT function is billed based on the total number of individual documents processed, regardless of the number of pages within each document

- A. Option D
- B. Option C
- C. Option A
- D. Option E
- E. Option B

Answer: A,E

Explanation:

Option B is correct because for the 'EXTRACT_ANSWER' function, the number of billable tokens is the sum of the tokens in the 'From_text' (source_document) and 'question' fields. Option D is correct as for 'CLASSIFY TEXT' (or labels, descriptions, and examples provided in the categories are counted as input tokens for each record processed, which directly increases the cost. Option A is incorrect because 'EMBED TEXT 768' and 'EMBED TEXT 1024' functions only count 'input tokens' towards the billable total, not both input and output tokens. Option C is incorrect because Cortex Structured Outputs does not incur additional compute cost for the overhead of verifying tokens against the supplied JSON schema, although schema complexity can increase total token consumption. Option E is incorrect because (and 'SNOWFLAKE.CORTEX.PARSE_DOCUMENT) billing is based on the 'number of document pages processed' (e.g., 3.33 Credits per 1,000 pages for Layout mode), not just the number of documents. For paged formats (PDF, DOCX), each page is billed as a page; for image files, each image is a page; for HTML/TXT, each 3,000 characters is a page.

NEW QUESTION # 127

A financial data team is implementing a Snowflake Cortex AI solution to summarize regulatory documents using SNOWFLAKE.CORTEX.TRY_COMPLETE. They aim for both cost efficiency and high reliability, especially when dealing with documents that might occasionally exceed model context limits or result in malformed output. Which of the following statements about the cost and operational behavior of TRY_COMPLETE are TRUE in this context? (Select all that apply)

If a call to `TRY_COMPLETE` fails and returns `NULL`, the input and output tokens associated with that specific failed operation are not counted for billing purposes.

Successful calls to `TRY_COMPLETE` for summarization will incur compute costs based on both the input prompt's tokens and the generated summary's tokens.

To manage conversation state in a multi-turn interaction, `TRY_COMPLETE` automatically retains the context of previous user prompts and model responses across successive calls.

The `TRY_COMPLETE` function is available only in a limited set of Snowflake regions and cannot leverage cross-region inference to access models in other regions.

When using `TRY_COMPLETE` with the `response_format` option to enforce JSON output, complex schemas can lead to a higher number of input and output tokens being consumed for successful calls.

- A. Option C
- B. Option D
- C. Option A
- D. Option E
- E. Option B

Answer: C,D,E

Explanation:

Option A is correct because

`TRY_COMPLETE`

does not incur costs for error handling; if the function returns

`NULL`

, no cost is incurred. This means if an operation fails and returns

`NULL`

, no tokens are billed for that specific attempt. Option B is correct because for functions that generate new text in the response (like

`SUMMARIZE`

or

`COMPLETE`

, and by extension

`TRY_COMPLETE`

when successful), both input and output tokens are billable. Option C is incorrect;

`TRY_COMPLETE`

(like

`COMPLETE`

) does not retain any state from one call to the next. For a stateful conversational experience, all previous user prompts and model responses must be explicitly passed in the

`prompt_or_history`

array. Option D is incorrect; the

`TRY_COMPLETE`

function is available in the same regions as

`COMPLETE`

, and `'COMPLETE'` models are often available via cross-region inference. Option E is correct. When using `'COMPLETE'` Structured Outputs

(which applies to

`TRY_COMPLETE`

when the

`response_format`



option is used), the number of tokens processed (and billed) increases with schema complexity. Larger and more complex schemas generally consume a larger number of input and output tokens.

NEW QUESTION # 128

A data scientist is optimising a Cortex Analyst application to improve the accuracy of literal searches within user queries, especially for high-cardinality dimension values. They decide to integrate Cortex Search for this purpose. Which of the following statements are true about this integration and the underlying data types in Snowflake? (Select all that apply)

- A. For optimal RAG retrieval performance with Cortex Search, it is generally recommended to split text into chunks of no more than 512 tokens, even when using embedding models with larger context windows such as 'snowflake-arctic-embed-l-v2.0-8k'.
- B. The 'VECTOR' data type in Snowflake, used to store embeddings generated for Cortex Search, is fully supported as a clustering key in standard tables and as a primary key in hybrid tables to accelerate vector similarity searches.
- C. Cortex Search Services, when configured as a source for Snowflake dynamic tables, automatically refresh their search index with continuous data updates, maintaining low-latency search results.
- D. The cost for embedding data into a Cortex Search Service is primarily incurred per output token generated by the embedding model, as these represent the final vector embeddings, rather than input tokens.
- E. To integrate Cortex Search with a logical dimension, the semantic model YAML must include a block within the dimension's definition, specifying the service name and optionally a 'literal_column'.

Answer: A,E

Explanation:

Option A is correct. Cortex Analyst can leverage Cortex Search Services to improve literal search by including a configuration block within a dimension's definition in the semantic model YAML. This block specifies the service name and an optional 'literal_column'. Option B is correct. Snowflake recommends splitting text in your search column into chunks of no more than 512 tokens for best search results with Cortex Search, even when using models with larger context windows like 'snowflake-arctic-embed-l-v2.0-8k'. This practice typically leads to higher retrieval and downstream LLM response quality in RAG scenarios. Option C is incorrect. The 'VECTOR' data type is allowed in hybrid tables but is explicitly not supported as a primary key, secondary index key, or clustering key in Snowflake. Option D is incorrect. For EMBED_TEXT functions, which are used to generate embeddings for Cortex Search, only 'input tokens' are counted towards the billable total, not output tokens. The Cortex Search service itself is billed per GB/month of indexed data. Option E is incorrect. Snowflake Cortex functions, including Cortex Search, do not support dynamic tables.

NEW QUESTION # 129

.....

Nowadays most people are attracted to the SnowPro® Specialty: Gen AI Certification Exam (GES-C01) certification and take it seriously because they know that it is the future. But they can't figure out where to prepare for SnowPro® Specialty: Gen AI Certification Exam (GES-C01) certification exam. After observing the problems of the students DumpsReview provides them with the best SnowPro® Specialty: Gen AI Certification Exam (GES-C01) Questions so they don't get depressed anymore and pass the SnowPro® Specialty: Gen AI Certification Exam (GES-C01) exam on the first try. The SnowPro® Specialty: Gen AI Certification Exam (GES-C01) is designed after consulting with a lot of professionals and getting their reviews.

GES-C01 Dump Collection: <https://www.dumpsreview.com/GES-C01-exam-dumps-review.html>

thus we strongly recommend our GES-C01 exam questions for not only that our GES-C01 training guide is designed to different versions: PDF, Soft and APP versions, which can offer you different study methods, but also that our GES-C01 learning perp can help you pass the exam without difficulty, Snowflake Guaranteed GES-C01 Questions Answers [passexamonline.com](https://www.passexamonline.com) will always accompany you during your preparation of the exams, so if any professional problems puzzle you, just contact our experts any time.

The MX series is supposed to integrate Web-development tools and server technologies, This makes my work as a malware designer all the more discomfoting, thus we strongly recommend our GES-C01 exam questions for not only that our GES-C01 training guide is designed to different versions: PDF, Soft and APP versions, which can offer you different study methods, but also that our GES-C01 learning perp can help you pass the exam without difficulty.

[Fully Updated] Snowflake GES-C01 Dumps With Latest GES-C01 Exam Questions (2026)

