

# Free PDF Snowflake - GES-C01 - SnowPro® Specialty: Gen AI Certification Exam High Hit-Rate Authentic Exam Questions



P.S. Free & New GES-C01 dumps are available on Google Drive shared by Pass4suresVCE: <https://drive.google.com/open?id=16JmeVMjlpY7gcS7qbOhZgNMewuMRgq3>

The web-based GES-C01 practice test is accessible via any browser. This GES-C01 mock exam simulates the actual SnowPro® Specialty: Gen AI Certification Exam (GES-C01) exam and does not require any software or plugins. Compatible with iOS, Mac, Android, and Windows operating systems, it provides all the features of the desktop-based GES-C01 Practice Exam software.

GES-C01 valid study test give you an in-depth understanding of the contents and help you to make out a detail study plan for GES-C01 preparation. All the questions are edited according to the analysis of data and summarized from the previous test, which can ensure the high hit rate. You just need take the spare time to study GES-C01 Training Material, the effects are obvious. You will get a high score with the help of Snowflake GES-C01 study pdf.

>> GES-C01 Authentic Exam Questions <<

## Best GES-C01 Vce & GES-C01 Examcollection Vce

In the present market you are hard to buy the valid GES-C01 study materials which are used to prepare the GES-C01 exam like our GES-C01 latest question. Both for the popularity in the domestic and the international market and for the quality itself, other kinds of study materials are incomparable with our GES-C01 Test Guide and far inferior to them. Our GES-C01 certification tool has their own fixed clients base in the domestic market and have an important share in the international market to attract more and more foreign clients.

## Snowflake SnowPro® Specialty: Gen AI Certification Exam Sample Questions (Q167-Q172):

### NEW QUESTION # 167

A software development team is building a conversational AI application within Snowflake, aiming to provide a dynamic and stateful chat experience for users. The application needs to handle follow-up questions while maintaining context, provide responses with a degree of creative variation, and actively filter out any potentially harmful content. The team utilizes the SNOWFLAKE.CORTEX.COMPLETE (or AI\_COMPLETE) function.

```
SELECT SNOWFLAKE.CORTEX.COMPLETE(  
  model => 'llama3.1-70b',  
  prompt => [{ 'role': 'user', 'content': 'First question.' }],  
  options => { 'temperature': 0.2, 'guardrails': TRUE, 'max_tokens': 50 }  
);
```

- A.);

```
SELECT SNOWFLAKE.CORTEX.TRY_COMPLETE(
  model => 'mistral-large2',
  prompt => [
    { 'role': 'user', 'content': 'What are the sales trends?' },
    { 'role': 'assistant', 'content': 'Sales are up 10% this quarter.' },
    { 'role': 'user', 'content': 'What about next quarter?' }
  ],
  options => { 'temperature': 0.8, 'guardrails': TRUE }
);
```

• B.

```
SELECT SNOWFLAKE.CORTEX.COMPLETE(
  model => 'snowflake-arctic',
  prompt => 'Analyze the latest market report.',
  options => { 'temperature': 0.5, 'guardrails': FALSE }
);
```

• C.

• D.

```
SELECT SNOWFLAKE.CORTEX.COMPLETE(
  model => 'gemma-7b',
  prompt => [
    { 'role': 'user', 'content': 'Summarize recent news.' },
    { 'role': 'user', 'content': 'Highlight key economic impacts.' }
  ],
  options => { 'temperature': 0.7, 'guardrails': TRUE }
);
```

```
SELECT SNOWFLAKE.CORTEX.COMPLETE(
  model => 'claude-3-5-sonnet',
  prompt => [
    { 'role': 'system', 'content': 'Respond as a financial advisor.' },
    { 'role': 'user', 'content': 'What is my portfolio performance?' }
  ],
  options => { 'temperature': 0.0, 'guardrails': TRUE }
);
```

• E.

**Answer: B**

Explanation:

The scenario requires statefulness (multi-turn conversation), creative variation in responses, and safety (filtering harmful content).

Option B correctly demonstrates a multi-turn conversation by passing a history of user and assistant messages in the prompt array. It uses a of which allows for creative variation in the output, and sets guardrails to TRUE to enable content filtering. Option A's prompt is temperature 0.8, single-turn, not multi-turn. Option C uses a single-turn prompt and disables guardrails. Option D uses a of which produces temperature 0.0, deterministic results and lacks creative variation. Option E has an incorrect conversation history structure with two consecutive 'user' roles without an 'assistant' response in between, which is not supported for stateful conversations.

### NEW QUESTION # 168

A development team is constructing a Gen AI application using Snowflake Cortex LLM functions, particularly for conversational and text generation tasks. They are concerned about potential high costs due to token consumption. Which of the following strategies would most effectively help minimize token usage and optimize costs when working with these Cortex LLM functions?

- A. Utilize the COUNT\_TOKENS (SNOWFLAKE.CORTEX) helper function to pre-validate the prompt length against the model's context window, thereby preventing truncation errors and subsequent re-runs.
- B. For multi-turn conversational experiences using SNOWFLAKE.CORTEX.COMPLETE, only send the most recent user prompt in each API call, as the model automatically retains previous context.
- C. To encourage more succinct LLM responses and reduce completion\_tokens, configure the temperature option to a higher value (e.g., 0.7) in COMPLETE function calls.
- D. When employing AI\_COMPLETE for structured output tasks, providing concise and highly descriptive explanations for

each field within the JSON schema will reduce the input tokens required for the LLM to understand and adhere to the schema accurately.

- E. In multi-turn conversations within Cortex Analyst, integrate a dedicated LLM summarization agent to rephrase follow-up questions, which reduces the total conversational history passed as context to the main LLM.

**Answer: A,D,E**

Explanation:

Option B is correct because while schema verification itself doesn't incur extra cost, a large or complex schema can increase token consumption. Providing precise and concise descriptions for schema fields helps the LLM understand and adhere to the desired format more efficiently, potentially reducing the overall tokens consumed for accurate responses. Option C is correct as the 'COUNT\_TOKENS' function allows developers to determine the token count of an input prompt for a specific model, enabling them to pre-emptively avoid exceeding the model's context window, thus preventing errors and wasted compute from re-runs. Option E is correct because for multi-turn conversations in Cortex Analyst, a summarization agent is specifically used to rephrase follow-up questions by incorporating previous context, without passing the entire, potentially long, conversation history. This significantly reduces the 'prompt\_tokens' sent to the main LLM for each turn and optimizes inference times. Option A is incorrect because 'COMPLETE' (and 'TRY\_COMPLETE') functions are stateless; to maintain conversational context, all previous user prompts and model responses must be included in the array, which increases token count proportionally. Simply sending the latest prompt would lose context. Option D is incorrect as setting a higher 'temperature' value (e.g., 0.7) increases the 'randomness and diversity' of the LLM's output, not necessarily its conciseness for cost optimization. For the most consistent (and often direct) results, a 'temperature' of 0 is recommended.

#### NEW QUESTION # 169

A data engineer has successfully experimented with a prompt and various model settings in the Snowflake Cortex Playground for a text classification task using the mistral-large2 model and Cortex Guard. They now want to operationalize this solution within their Snowflake environment. Which of the following statements correctly describe capabilities or considerations when moving from the Cortex Playground to a production pipeline?

- A. The exported SQL query, when used with dynamic tables, supports incremental refresh for efficient processing of new data without recomputing the entire table.
- B. The Playground allows exporting the exact SQL query with all defined model settings, including temperature and Cortex Guard enablement, for direct use in a Snowflake worksheet or task.
- C. If the mistral-large2 model is not natively available in the target production region, cross-region inference must be enabled by setting the CORTEX\_ENABLED\_CROSS\_REGION parameter.
- D. To filter unsafe LLM responses in production, the Cortex Guard option, which is built with Meta's Llama Guard 3, must be explicitly enabled in the COMPLETE function's options argument.
- E. For continuous processing of new data, the exported SQL query can be automated using

**Answer: B,C,D,E**

Explanation:

The Cortex Playground offers the capability to export the exact SQL query, including specified settings like 'temperature' and whether Cortex Guard is enabled. This exported SQL can be directly used in Snowflake worksheets, notebooks, or automated via streams and tasks for continuous execution. Cortex Guard is a feature of the (or 'COMPLETE') function, built with Meta's Llama Guard 3, and needs to be explicitly enabled using the 'guardrails' option to filter unsafe responses. Automation with streams and tasks is a valid approach for document processing pipelines using 'PREDICT' and 'COMPLETE' functions. If a model like 'mistral-large?' is not natively available in a specific Snowflake region, cross-region inference must be enabled for Cortex LLM functions. However, dynamic tables do not currently support incremental refresh when using the 'COMPLETE' function.

#### NEW QUESTION # 170

A data processing team is using Snowflake Document AI to extract data from incoming supplier invoices. They observe that many documents are failing to process, and successful extractions are taking longer than expected, leading to increased costs. Upon investigation, they find error messages such as

```
{ "__processingErrors": [ "Document has too many pages. Actual: 130. Maximum: 125." ] }
,
{ "__processingErrors": [ "File exceeds maximum size. Actual: 54096026 bytes. Maximum: 50000000 bytes." ] }
, and
{ "__processingErrors": [ "cannot identify image file <_io.BytesIO object at 0x...>" ] }
```

. Additionally, their 'X-LARGE' virtual warehouse is constantly active, contributing to higher-than-anticipated bills. Which two of the following actions are essential steps to troubleshoot and address the root causes of these processing errors and optimize their Document AI pipeline?

- A. Increase the 'max\_tokenS' parameter within the ' !PREDICT' function options to accommodate longer document responses from the model.
- **B. Implement a pre-processing step to split documents exceeding 125 pages or 50 MB into smaller, compliant files before loading to the stage.**
- C. Scale down the virtual warehouse to 'X-SMALL' or 'SMALL' size, as larger warehouses do not increase Document AI query processing speed and incur unnecessary costs.
- **D. Configure the internal stage used for storing invoices with 'ENCRYPTION = (TYPE = 'SNOWFLAKE\_SSE'Y.**
- E. Redefine extraction questions to be more generic and encompassing, reducing the number of distinct questions needed per document.

**Answer: B,D**

Explanation:

The error messages 'Document has too many pages. Actual: 130. Maximum: 125.' and 'File exceeds maximum size. Actual: 54096026 bytes. Maximum: 50000000 bytes.' directly indicate that the documents do not meet Document AI's input requirements, which specify a maximum of 125 pages and 50 MB file size. Therefore, implementing a pre-processing step to split or resize these documents is an essential solution (Option B). The error 'cannot identify image file <\_io.BytesIO object at 0x...>' is a known issue that occurs when an internal stage used for Document AI is not configured with 'SNOWFLAKE\_SSE' encryption. Correctly configuring the stage with this encryption type is crucial for resolving this processing error (Option D). Option A, while addressing cost optimization, is not a root cause of the 'processing errors' themselves, although it is a best practice for cost governance as larger warehouses do not increase Document AI query processing speed. Option C is incorrect; best practices for question optimization suggest being specific, not generic. Option E is incorrect as 'max\_tokenS' relates to the length of the model's output, not the input document's size or page limits.

## NEW QUESTION # 171

A data application developer is tasked with building a multi-turn conversational AI application using Streamlit in Snowflake (SiS) that leverages the COMPLETE (SNOWFLAKE.CORTEX) LLM function. To ensure the conversation flows naturally and the LLM maintains context from previous interactions, which of the following is the most appropriate method for handling and passing the conversation history?

The developer should store the entire conversation history in a temporary table in Snowflake and query it with each new turn, passing only the latest user message to the COMPLETE function.

- Snowflake automatically manages conversational context for COMPLETE within the session, so the developer only needs to pass the current user prompt as a string.
- The conversation history must be explicitly managed within the Streamlit application's state, typically by initializing `st.session_state.messages = []` and appending each user and assistant message as an object with 'role' and 'content' keys, then passing the full list to the `prompt_or_history` argument of COMPLETE.
- The developer should concatenate all previous user prompts and assistant responses into a single, long string, and pass this as the `<prompt>` argument to COMPLETE for each turn.
- The COMPLETE function has an optional 'conversation\_id' parameter that automatically retrieves and manages conversation history when provided.

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

**Answer: D**

Explanation:

To provide a stateful, conversational experience with the 'COMPLETE (SNOWFLAKE.CORTEX)' function (or its latest version, 'AI\_COMPLETE'), all previous user prompts and model responses must be explicitly passed as part of the argument. This argument expects an array of objects, where each object represents a turn and contains a 'role' ('system', 'user', or 'assistant') and a 'content' key, presented in chronological order. In Streamlit, 'st.session\_state' is the standard and recommended mechanism for storing and managing data across reruns of the application, making it ideal for maintaining chat history, by initializing `st.session_state.messages = []` and appending messages to it. Option A is incorrect because 'COMPLETE' does not inherently manage history from external tables. Option B is incorrect as 'COMPLETE' does not retain state between calls; history must be explicitly managed. Option D is a less effective form of prompt engineering compared to passing structured history, as it loses the semantic role distinction and can be less accurate for LLMs. Option E describes a non-existent parameter for the 'COMPLETE' function.



• • • • •

**Best GES-C01 Vce:** <https://www.pass4suresvce.com/GES-C01-pass4sure-vce-dumps.html>

With the basics covered, let's look at each type of positioning in turn. The value of a brand is that the GES-C01 exam questions are more than just exam preparation tool -- it should be part of our lives, into our daily lives.

You don't have to wait a long time to start your preparation for the GES-C01 Exam, As the quick development of the world economy and intense competition in the international, the world GES-C01 labor market presents many new trends: company's demand for the excellent people is growing.

If you choose GES-C01 study guide, you will find the test questions and test answers are certainly high-quality, which is the royal road to success.

- [illegible]

myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt,  
myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, blogfreely.net, www.stes.tyc.edu.tw,  
hhi.instructure.com, estar.jp, rupeebazar.com, gettr.com, Disposable vapes

BTW, DOWNLOAD part of Pass4suresVCE GES-C01 dumps from Cloud Storage: <https://drive.google.com/open?id=16JmeVMjlpY7gcS7qbOhZgNMewuMRgq3>