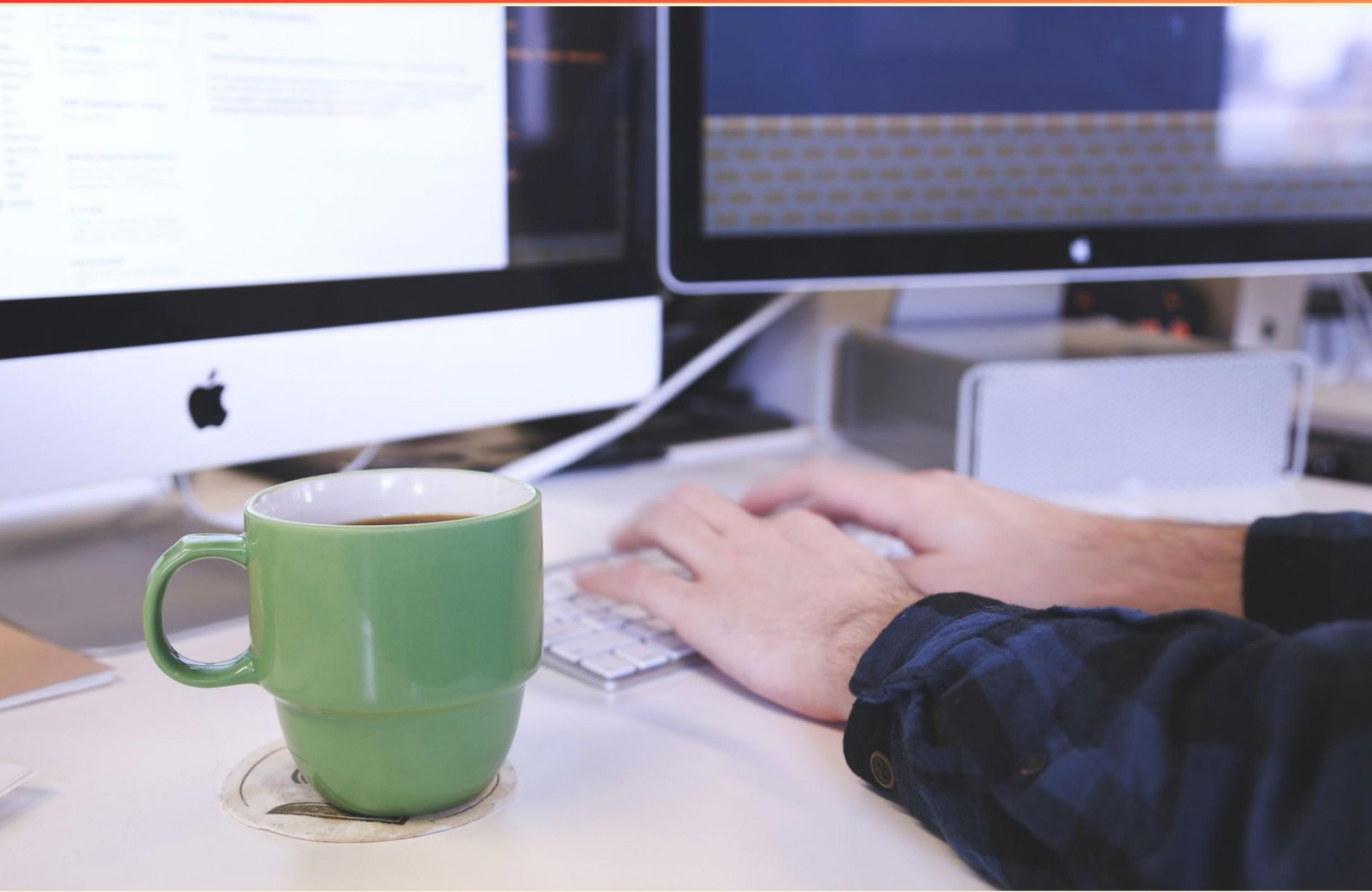


SNOWFLAKE DATA ENGINEER PRACTICE EXAM (SAMPLE)

STUDY GUIDE



SAMPLE STUDY GUIDE WITH LIMITED QUESTIONS

**VISIT US ONLINE FOR
THE FULL VERSION STUDY GUIDE**

<https://snowflakedataengineer.examzify.com>



Copyright © 2026 by Examzify - A Kaluba Technologies Inc. product.

ALL RIGHTS RESERVED.

No part of this book may be reproduced or transferred in any form or by any means, graphic, electronic, or mechanical, including photocopying, recording, web distribution, taping, or by any information storage retrieval system, without the written permission of the author.

Notice: Examzify makes every reasonable effort to obtain accurate, complete, and timely information about this product from reliable sources.

SAMPLE

Table of Contents

Copyright	1
Table of Contents	2
Introduction	3
How to Use This Guide	4
Questions	5
Answers	8
Explanations	10
Next Steps	15

SAMPLE

Introduction

Preparing for a certification exam can feel overwhelming, but with the right tools, it becomes an opportunity to build confidence, sharpen your skills, and move one step closer to your goals. At Examzify, we believe that effective exam preparation isn't just about memorization, it's about understanding the material, identifying knowledge gaps, and building the test-taking strategies that lead to success.

This guide was designed to help you do exactly that.

Whether you're preparing for a licensing exam, professional certification, or entry-level qualification, this book offers structured practice to reinforce key concepts. You'll find a wide range of multiple-choice questions, each followed by clear explanations to help you understand not just the right answer, but why it's correct.

The content in this guide is based on real-world exam objectives and aligned with the types of questions and topics commonly found on official tests. It's ideal for learners who want to:

- Practice answering questions under realistic conditions,
- Improve accuracy and speed,
- Review explanations to strengthen weak areas, and
- Approach the exam with greater confidence.

We recommend using this book not as a stand-alone study tool, but alongside other resources like flashcards, textbooks, or hands-on training. For best results, we recommend working through each question, reflecting on the explanation provided, and revisiting the topics that challenge you most.

Remember: successful test preparation isn't about getting every question right the first time, it's about learning from your mistakes and improving over time. Stay focused, trust the process, and know that every page you turn brings you closer to success.

Let's begin.

How to Use This Guide

This guide is designed to help you study more effectively and approach your exam with confidence. Whether you're reviewing for the first time or doing a final refresh, here's how to get the most out of your Examzify study guide:

1. Start with a Diagnostic Review

Skim through the questions to get a sense of what you know and what you need to focus on. Your goal is to identify knowledge gaps early.

2. Study in Short, Focused Sessions

Break your study time into manageable blocks (e.g. 30 – 45 minutes). Review a handful of questions, reflect on the explanations.

3. Learn from the Explanations

After answering a question, always read the explanation, even if you got it right. It reinforces key points, corrects misunderstandings, and teaches subtle distinctions between similar answers.

4. Track Your Progress

Use bookmarks or notes (if reading digitally) to mark difficult questions. Revisit these regularly and track improvements over time.

5. Simulate the Real Exam

Once you're comfortable, try taking a full set of questions without pausing. Set a timer and simulate test-day conditions to build confidence and time management skills.

6. Repeat and Review

Don't just study once, repetition builds retention. Re-attempt questions after a few days and revisit explanations to reinforce learning. Pair this guide with other Examzify tools like flashcards, and digital practice tests to strengthen your preparation across formats.

There's no single right way to study, but consistent, thoughtful effort always wins. Use this guide flexibly, adapt the tips above to fit your pace and learning style. You've got this!

Questions

SAMPLE

- 1. If a table has a data retention of 10 days and is loaded with 256 MB of data daily, what will be the total storage after 5 days of ingestion?**
 - A. 1280 MB
 - B. 256 MB
 - C. 2560 MB
 - D. 0 MB

- 2. Which use case is valid for SnowCD?**
 - A. Integrating SnowCD into automated deployment scripts
 - B. Using SnowCD to load data from Kafka
 - C. Visualizing query performance with SnowCD
 - D. Scheduling data replication with SnowCD

- 3. What is the typical micro-partition size before compression?**
 - A. 50 to 500 MB
 - B. 1 to 2 GB
 - C. 5 to 10 MB
 - D. 2 to 10 GB

- 4. Is it best practice to avoid binding data using Python's string formatting functions due to SQL injection risk?**
 - A. False
 - B. True
 - C. It depends on the context
 - D. Not related

- 5. Which technique speeds up queries on a table with city demography data for both range and equality searches?**
 - A. Create a materialized view.
 - B. Turn on search optimization for the table.
 - C. Create a secure view for each search pattern.
 - D. Cluster the table.

6. In Databricks, Snowflake's query pushdown is enabled by default.
- A. It is enabled only in some regions
 - B. It is not enabled at all
 - C. True
 - D. Only in premium plans
7. Streams have no Fail-safe period or Time Travel retention period. The metadata in these objects cannot be recovered if a stream is dropped.
- A. True
 - B. False
 - C. Only fail-safe
 - D. Only Time Travel
8. What does setting VALIDATION_MODE in COPY INTO TABLE do?
- A. Validate the data files without loading them.
 - B. Load the data first, then validate
 - C. Ignore errors during validation
 - D. Validate a random subset of rows only
9. Which error indicates that the storage integration cannot be found for a stage after recreation?
- A. 000139
 - B. 003139=SQL compilation error: Integration "{0}" associated with the stage "{1}" cannot be found.
 - C. 001139
 - D. 999999
10. If both FORMAT_NAME and TYPE are specified in a COPY command, what is the likely outcome?
- A. No effect
 - B. Might result in unexpected behavior
 - C. Guaranteed error
 - D. It will ignore TYPE

Answers

SAMPLE

1. A
2. A
3. A
4. B
5. B
6. C
7. A
8. A
9. B
10. B

SAMPLE

Explanations

SAMPLE

1. If a table has a data retention of 10 days and is loaded with 256 MB of data daily, what will be the total storage after 5 days of ingestion?

A. 1280 MB

B. 256 MB

C. 2560 MB

D. 0 MB

Understanding how data retention interacts with daily ingestion helps you predict storage growth. If you load 256 MB each day and keep data for 10 days, the storage increases by 256 MB per day until the retention window is full; no data is purged yet. After 5 days, all five days of data are still within the 10-day retention, so you simply have 256 MB times 5 days. That equals 1280 MB of total storage. So 1280 MB is the correct total after 5 days. The other amounts would imply holding only one day of data, holding data for the full 10 days, or having no data at all, which isn't the case here.

2. Which use case is valid for SnowCD?

A. Integrating SnowCD into automated deployment scripts

B. Using SnowCD to load data from Kafka

C. Visualizing query performance with SnowCD

D. Scheduling data replication with SnowCD

SnowCD is built to automate and orchestrate Snowflake deployments as code, so its strongest use case is integrating it into automated deployment scripts. By treating schema changes, grants, and other environment updates as versioned, codified steps, SnowCD enables repeatable, auditable deployments that fit neatly into CI/CD pipelines. It's about applying changes in a controlled, automated way across environments, not about data ingestion, visualization, or scheduling tasks. Therefore loading data from Kafka isn't a SnowCD function, visualizing query performance isn't its purpose, and scheduling data replication falls outside its deployment-focused role.

3. What is the typical micro-partition size before compression?

A. 50 to 500 MB

B. 1 to 2 GB

C. 5 to 10 MB

D. 2 to 10 GB

Micro-partitions are the basic storage unit Snowflake uses to organize data. Before any compression, a typical micro-partition holds about 50 to 500 MB of data. This size is chosen to balance efficient pruning during queries—Snowflake can skip partitions that don't match filters—while keeping the partition small enough to load, cache, and scan quickly. Data is stored columnar within the partition and then compressed, so the on-disk size after compression is smaller, but the uncompressed size remains in that 50–500 MB range. Sizes much smaller (like 5–10 MB) would increase metadata overhead and reduce pruning efficiency, while sizes in the gigabyte range (1–2 GB or more) would hurt pruning and increase scan costs.

4. Is it best practice to avoid binding data using Python's string formatting functions due to SQL injection risk?

- A. False
- B. True**
- C. It depends on the context
- D. Not related

Binding data with parameterized queries is the safer, standard approach for including values in SQL. Building SQL by inserting values directly with Python's string formatting (or f-strings) mixes code and data, and any user-supplied input can alter the SQL itself, leading to SQL injection risks. Parameter binding sends the SQL statement with placeholders first, and the values are supplied separately by the driver. The database then treats those values strictly as data, applying proper escaping and typing, so the input cannot change the intended logic of the query. This not only prevents injection but also often improves performance through statement plan reuse when the same query runs with different values. In Snowflake with the Python connector, you typically provide the query with placeholders and pass the parameters separately, which enforces safe separation of code and data. So yes, it's best practice to avoid Python string formatting for binding data due to SQL injection risk.

5. Which technique speeds up queries on a table with city demography data for both range and equality searches?

- A. Create a materialized view.
- B. Turn on search optimization for the table.**
- C. Create a secure view for each search pattern.
- D. Cluster the table.

The technique being tested focuses on accelerating predicate-based searches using a specialized indexing feature. Turning on search optimization for a table makes Snowflake maintain a lightweight search index that speeds up predicates, especially for both range queries (for example, filtering by a range of population or year) and exact-match lookups (such as filtering by city name or code). This optimization helps the query engine prune irrelevant micro-partitions more effectively, so queries can locate matching rows quickly without scanning the entire table. It's particularly advantageous for large city demography tables where you run a mix of range and equality filters. While clustering and materialized views have their uses, they address different patterns or require more manual tuning and maintenance. Secure views don't inherently improve performance, and clustering requires careful design of cluster keys and ongoing maintenance. Search optimization provides a targeted, often easier-to-manage speedup for the combined range and exact-match searches described.

6. In Databricks, Snowflake's query pushdown is enabled by default.

- A. It is enabled only in some regions
- B. It is not enabled at all
- C. True**
- D. Only in premium plans

Query pushdown means the Databricks-Snowflake connector translates eligible Spark operations—like filters and projections—into Snowflake SQL so that Snowflake runs the work directly, only returning the necessary data. This offloads work to Snowflake's engine, reducing data movement and usually speeding up queries. The default behavior of the connector is to enable this pushdown, so the statement is correct: pushdown happens by default without extra configuration. The other choices aren't accurate because pushdown isn't gated by region, isn't disabled entirely, and isn't limited to premium plans. Keep in mind that not every operation can be pushed down; some complex expressions or unsupported constructs may still be evaluated after data is fetched, but the default remains enabled.

7. Streams have no Fail-safe period or Time Travel retention period. The metadata in these objects cannot be recovered if a stream is dropped.

- A. True**
- B. False
- C. Only fail-safe
- D. Only Time Travel

Streams are change-tracking metadata, not data with its own retention. Time Travel and Fail-safe apply to data stored in tables, not to the stream's metadata. Because a stream doesn't hold data and its state isn't preserved by those retention mechanisms, dropping a stream means its internal offsets and change-tracking history cannot be recovered. To resume change tracking, you'd recreate a stream on the base table, and you can use Time Travel on the base table to inspect past data, but the dropped stream's previous state isn't recoverable.

8. What does setting VALIDATION_MODE in COPY INTO TABLE do?

- A. Validate the data files without loading them.**
- B. Load the data first, then validate
- C. Ignore errors during validation
- D. Validate a random subset of rows only

Setting VALIDATION_MODE on COPY INTO TABLE makes the operation a dry run: Snowflake reads the input data files and checks that they conform to the target table's schema and the chosen file format, but it does not insert any rows. This lets you catch data-type, length, nullability, and format errors before loading, and it returns a report of any issues found so you can fix the data. It does not perform the load, so it won't actually write data, and it isn't about skipping errors or validating only a random subset.

9. Which error indicates that the storage integration cannot be found for a stage after recreation?

- A. 000139
- B. 003139=SQL compilation error: Integration "{0}" associated with the stage "{1}" cannot be found.**
- C. 001139
- D. 999999

In Snowflake, a stage can reference a storage integration to access external storage. The stage stores the name of that integration, and during SQL compilation Snowflake verifies that the named integration actually exists. If the integration has been dropped or hasn't been recreated with that name, the compiler can't resolve the reference and returns an error stating that the integration associated with the stage cannot be found. The exact message flags both the missing integration and the affected stage, making it the clear indicator that the problem lies with the storage integration not being present for the stage after recreation. This type of error is distinct from syntax or general object-not-found issues because it specifically points to the unresolved storage integration binding on a stage.

10. If both FORMAT_NAME and TYPE are specified in a COPY command, what is the likely outcome?

- A. No effect
- B. Might result in unexpected behavior**
- C. Guaranteed error
- D. It will ignore TYPE

Mixing a named file format with inline format options creates conflicting instructions for how to parse the input files. The system isn't guaranteed to pick one source of truth over the other, so the behavior can vary in unpredictable ways. As a result, you might see differences in delimiter handling, header treatment, or other parsing details across runs or across files, depending on which setting takes effect. This is not typically a guaranteed error, nor is it a no-op or a guaranteed ignore of one option. The safer approach is to choose one method (either rely on the named format or specify inline options consistently) to avoid ambiguous behavior.

SAMPLE

Next Steps

Congratulations on reaching the final section of this guide. You've taken a meaningful step toward passing your certification exam and advancing your career.

As you continue preparing, remember that consistent practice, review, and self-reflection are key to success. Make time to revisit difficult topics, simulate exam conditions, and track your progress along the way.

If you need help, have suggestions, or want to share feedback, we'd love to hear from you. Reach out to our team at hello@examzify.com.

Or visit your dedicated course page for more study tools and resources:

<https://snowflakedataengineer.examzify.com>

We wish you the very best on your exam journey. You've got this!

SAMPLE