

# Prepare with updated Snowflake DSA-C03 dumps - Get up to one year of free updates



BTW, DOWNLOAD part of Exams-boost DSA-C03 dumps from Cloud Storage: <https://drive.google.com/open?id=1hUfVkeCnFTM8QMIB1n-1eIJC-2N3iXbz>

We provide you with the latest prep material which is according to the content of Snowflake DSA-C03 certification exam and enhances your knowledge to crack the test. Exams-boost practice material is made by keeping in focus all the sections of the current syllabus. Our primary objective is to provide you with SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) actual questions to complete preparation for the test in few days. Our product includes SnowPro Advanced: Data Scientist Certification Exam real questions, desktop practice test software, and web-based practice exam. Keep reading to find out what are the specifications of these formats.

Snowflake DSA-C03 Exam Questions, applicants may study for and pass their desired certification exam. You may use Exams-boost's top DSA-C03 study resources to prepare for the SnowPro Advanced: Data Scientist Certification Exam exam. The Snowflake DSA-C03 Exam Questions offered by Exams-boost are dependable and trustworthy sources of preparation. Exams-boost provides valid exam questions and answers for customers, and free updates for 365 days.

>> DSA-C03 Exam Pattern <<

## DSA-C03 Test Vce Free | DSA-C03 Popular Exams

No doubt the SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) certification is one of the most challenging certification exams in the market. This SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) certification exam gives always a tough time to SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) exam candidates. The Exams-boost understands this hurdle and offers recommended and real Snowflake DSA-C03 Exam Practice questions in three different formats. These formats hold high demand in the market and offer a great solution for quick and complete SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) exam preparation.

## Snowflake SnowPro Advanced: Data Scientist Certification Exam Sample Questions (Q122-Q127):

### NEW QUESTION # 122

A healthcare provider has a Snowflake table 'MEDICAL RECORDS' containing patient notes stored as unstructured text in a column called 'NOTE TEXT'. They want to identify different patient groups based on the topics discussed in these notes. They aim to use a combination of unsupervised and supervised learning. Which of the following represents a robust workflow to achieve this goal?

- A. MultiOutputClassifier wrapped around a Logistic Regression model) within Snowflake (using Snowpark), using the original 'NOTE TEXT' as input features (TF-IDF or word embeddings) and the manually assigned topic labels as target variables. Use the trained model to classify the remaining patient notes into relevant patient groups.
- B. Perform topic modeling on a sample of the 'NOTE TEXT' data using a Snowflake Python UDF. Manually review the top documents for each identified topic, and assign labels describing the patient group represented by each topic. Train a supervised multi-label classification model (e.g., using scikit-learn's
- C. Use a Snowflake external function to call a pre-trained topic modeling model (e.g., BERTopic) hosted on Google Cloud AI Platform. Assign topic probabilities to each patient note. Then, perform K-Means clustering on the topic probabilities to identify patient segments. No manual labeling is performed.
- D. Perform topic modeling (e.g., LDA) directly on the 'NOTE\_TEXT' column using a Python UDF in Snowflake. Manually label a subset of the resulting topics. Then, train a supervised classifier (e.g., Naive Bayes) to predict the identified topics for new patient notes.
- E. Export all 'NOTE TEXT' data to an external system, use an existing NLP pipeline for topic modeling and manual labeling, then create a Snowflake UDF that replicates this entire pipeline internally.

**Answer: B**

Explanation:

Option D is the most comprehensive and practical. First, it uses unsupervised topic modeling to discover potential patient groups. Second, it uses manual labeling to create a supervised training dataset. Third, it trains a supervised multi-label classification model within Snowflake (using Snowpark), allowing for automated patient group assignment based on the text of their notes, leveraging TF-IDF or word embeddings for feature representation. This balances the efficiency of unsupervised learning with the accuracy of supervised learning. It also highlights Snowflake's ability to directly train and deploy models using Snowpark.

### NEW QUESTION # 123

You have a structured dataset in Snowflake containing customer information and purchase history. You aim to build a multi-class classification model to predict customer churn, categorizing customers into 'Low Risk', 'Medium Risk', and 'High Risk' of churning. After training the model, you want to evaluate its performance. Which of the following metrics and evaluation techniques, when used together, provide the MOST comprehensive understanding of the model's performance across all churn risk categories, especially when dealing with potential class imbalance?

- A. Only Overall Accuracy and a confusion Matrix.
- B. Area Under the ROC Curve (AUC-ROC) for each class (one-vs-rest approach), Precision-Recall Curve for each class, and Cumulative Accuracy Profile (CAP) curve.
- C. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (Coefficient of Determination).
- D. Log Loss (Cross-Entropy Loss), Gini Coefficient, and Kolmogorov-Smirnov (KS) statistic.
- E. Overall Accuracy, Precision, Recall, F I-Score for each class, and Confusion Matrix.

**Answer: E**

Explanation:

Option A offers the most comprehensive evaluation. Overall accuracy provides a general sense of performance, but can be misleading with imbalanced classes. Precision, recall, and F1-score, calculated for each class, give a detailed view of the model's performance on each churn risk category. The confusion matrix provides a visual representation of the model's classification errors, allowing you to identify patterns of misclassification between the different risk levels. Option B, ROC AUC and Precision-Recall curve are also relevant but is better for binary classification (with one-vs-rest extended for multiclass). CAP curves are less common. Option C (Log Loss, Gini, KS) is more suitable for binary classification or ranking problems. Option D (RMSE, MAE, R-squared) are regression metrics, not suitable for classification.

### NEW QUESTION # 124

You are tasked with deploying a time series forecasting model within Snowflake using Snowpark Python. The model requires significant pre-processing and feature engineering steps that are computationally intensive. These steps include calculating rolling statistics, handling missing values with imputation, and applying various transformations. You aim to optimize the execution time of

these pre-processing steps within the Snowpark environment. Which of the following techniques can significantly improve the performance of your data preparation pipeline?

- A. Force single-threaded execution by setting to avoid overhead associated with parallel processing.
- B. Convert the Snowpark DataFrame to a Pandas DataFrame using and perform all pre-processing operations using Pandas functions before loading the processed data back to Snowflake.
- C. Write the feature engineering logic directly in SQL and create a view. Use the Snowpark DataFrame API to query the view, avoiding Python code execution within Snowpark.
- D. Utilize Snowpark's vectorized UDFs and DataFrame operations to leverage Snowflake's distributed computing capabilities.
- E. Ensure that all data used is small enough to fit within the memory of the client machine running the Snowpark Python script, thus removing the need for distributed computing.

**Answer: C,D**

Explanation:

Vectorized UDFs and SQL Views are the key to optimizing data pre-processing. Options B and E are correct. B - Utilize Snowpark's vectorized UDFs and DataFrame operations: Snowpark is designed to push computation down to Snowflake's distributed compute engine. Vectorized UDFs allow you to execute Python code in a parallel and efficient manner directly within Snowflake. E - SQL View: Snowpark DataFrame API can query the view from SQL directly. Writing the data preparation logic in SQL leverages the snowflake's engine more effectively than Pandas or Python on a client machine. Options A, C, and D are generally incorrect: Option A is incorrect as it defeats the purpose of using Snowpark. Parallel execution is generally much faster. Option C is incorrect as moving data outside of snowflake is costly. Option D is incorrect. Snowpark is designed to manage a large scale of data.

#### NEW QUESTION # 125

You are building a predictive model for customer churn using linear regression in Snowflake. You have identified several features, including 'CUSTOMER AGE', 'MONTHLY SPEND', and 'NUM CALLS'. After performing an initial linear regression, you suspect that the relationship between 'CUSTOMER AGE' and churn is not linear and that older customers might churn at a different rate than younger customers. You want to introduce a polynomial feature of "CUSTOMER AGE (specifically, 'CUSTOMER AGE SQUARED') to your regression model within Snowflake SQL before further analysis with python and Snowpark. How can you BEST create this new feature in a robust and maintainable way directly within Snowflake?

- ALTER TABLE CUSTOMER\_DATA ADD COLUMN CUSTOMER\_AGE\_SQUARED FLOAT; UPDATE CUSTOMER\_DATA SET CUSTOMER\_AGE\_SQUARED = CUSTOMER\_AGE \* CUSTOMER\_AGE;
- CREATE OR REPLACE TEMPORARY TABLE CUSTOMER\_DATA\_WITH\_FEATURES AS SELECT \*, CUSTOMER\_AGE \* CUSTOMER\_AGE AS CUSTOMER\_AGE\_SQUARED FROM CUSTOMER\_DATA;
- CREATE OR REPLACE VIEW CUSTOMER\_DATA\_WITH\_FEATURES AS SELECT \*, CUSTOMER\_AGE \* CUSTOMER\_AGE AS CUSTOMER\_AGE\_SQUARED FROM CUSTOMER\_DATA;
- CREATE OR REPLACE TABLE CUSTOMER\_DATA AS SELECT \*, CUSTOMER\_AGE \* CUSTOMER\_AGE AS CUSTOMER\_AGE\_SQUARED FROM CUSTOMER\_DATA;
- CREATE OR REPLACE TABLE CUSTOMER\_DATA WITH FEATURES AS SELECT \*, POWER(CUSTOMER\_AGE, 2) AS CUSTOMER\_AGE\_SQUARED FROM CUSTOMER\_DATA;

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

**Answer: C**

Explanation:

Creating a VIEW (option C) is the BEST approach for several reasons. It doesn't modify the underlying data, which is crucial for data governance and prevents unintended side effects. The feature is calculated on-the-fly whenever the view is queried, ensuring that the feature is always up-to-date if the underlying changes. Options A, D, and E permanently alter the table, potentially leading to data redundancy and requiring manual updates if the column changes. Option B creates a temporary table, which is suitable for short-lived experiments but not ideal for a feature that will be used repeatedly. Using 2) is equivalent to CUSTOMER\_AGE \* CUSTOMER\_AGE. Views are efficient because Snowflake's query optimizer can often push down computations into the underlying table. Option C also avoids needing to manage the lifecycle of updated calculated columns.

#### NEW QUESTION # 126

You have a Snowpark DataFrame named 'product\_reviews' containing customer reviews for different products. The DataFrame includes columns like 'product\_id', 'review\_text', and 'rating'. You want to perform sentiment analysis on the 'review\_text' to

identify the overall sentiment towards each product. You decide to use Snowpark for Python to create a user-defined function (UDF) that utilizes a pre-trained sentiment analysis model hosted externally. You need to ensure secure access to this model and efficient execution. Which of the following represents the BEST approach, considering security and performance?

- A. Create an external function in Snowflake that calls a serverless function (e.g., AWS Lambda, Azure Function) that performs the sentiment analysis. Use Snowflake's network policies to restrict access to the serverless function and secrets management to handle API keys.
- B. Create a Snowpark Pandas UDF that calls the external sentiment analysis API. Use Snowflake secrets management to store the API key and retrieve it within the UDF.
- C. Create an inline Python UDF that directly calls the external sentiment analysis API with hardcoded API keys within the UDF code.
- D. Create a Java UDF that utilizes a library to call the sentiment analysis API. Pass the API key as a parameter to the UDF each time it is called.
- **E. Create an external function in Snowflake that calls a serverless function. Configure the API gateway in front of the serverless function to enforce authentication via Mutual TLS (mTLS) using Snowflake-managed certificates.**

**Answer: E**

Explanation:

Option E provides the BEST combination of security and performance. Using an external function that calls a serverless function allows for leveraging scalable compute resources. Configuring the API gateway with Mutual TLS (mTLS) provides a strong layer of authentication, ensuring that only Snowflake can access the serverless function. Snowflake's network policies further restrict access. Storing the API key using Snowflake secrets management within the serverless function provides additional security. Option A is insecure due to hardcoded API keys. Option B is better but can be less performant than external functions. Option C requires managing Java dependencies and might not be as scalable as serverless functions. Option D is good but mTLS gives the best protection available.

## NEW QUESTION # 127

.....

Firstly, we can give you 100% pass rate guarantee on the DSA-C03 exam. Our DSA-C03 practice quiz is equipped with a simulated examination system with timing function, allowing you to examine your learning results at any time, keep checking for defects, and improve your strength. Secondly, during the period of using DSA-C03 learning guide, we also provide you with 24 hours of free online services, which help to solve any problem for you on the DSA-C03 exam questions at any time and sometimes mean a lot to our customers.

**DSA-C03 Test Vce Free:** <https://www.exams-boost.com/DSA-C03-valid-materials.html>

Various version of DSA-C03 test dumps--- PDF & Software & APP version, The web-based Snowflake DSA-C03 practice exam is compatible with all operating systems and browsers, Snowflake DSA-C03 Exam Pattern Money Back Guarantee and 24/7 Customer Care, Our staff will create a unique study plan for you based on the choice of the right version of the DSA-C03 exam questions, So we are always very helpful in arranging our DSA-C03 Test Vce Free - SnowPro Advanced: Data Scientist Certification Exam exam questions with both high quality and reasonable price.

Offers expert guidance on improving retrospectives by integrating DSA-C03 Test Vce Free experiments and hypotheses, The third bit represents the location' of a packet in a series of fragmented packets.

Various version of DSA-C03 Test Dumps--- PDF & Software & APP version, The web-based Snowflake DSA-C03 practice exam is compatible with all operating systems and browsers.

## Free PDF DSA-C03 - SnowPro Advanced: Data Scientist Certification Exam –Efficient Exam Pattern

Money Back Guarantee and 24/7 Customer Care, Our staff will create a unique study plan for you based on the choice of the right version of the DSA-C03 exam questions.

So we are always very helpful in arranging DSA-C03 our SnowPro Advanced: Data Scientist Certification Exam exam questions with both high quality and reasonable price.

- Exam DSA-C03 Training ✓ DSA-C03 Training Courses ☐ DSA-C03 Test Pdf ☐ Search for ( DSA-C03 ) on ➡ [www.practicevce.com](http://www.practicevce.com) ☐ ☐ immediately to obtain a free download ☐ Valid DSA-C03 Exam Question

- 2026 Latest Exams-boost DSA-C03 PDF Dumps and DSA-C03 Exam Engine Free Share: <https://drive.google.com/open?id=1hUfVkeCnFTM8QMIB1n-1eIJC-2N3iXbz>

2026 Latest Exams-boost DSA-C03 PDF Dumps and DSA-C03 Exam Engine Free Share: <https://drive.google.com/open?id=1hUfVkeCnFTM8QMIB1n-1eIJC-2N3iXbz>