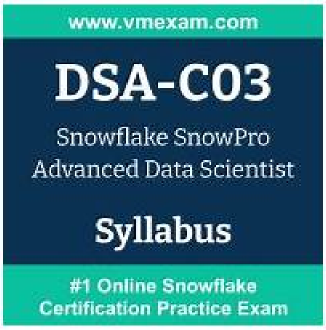
Snowflake DSA-C03 Test Answers & DSA-C03 Valid Test Testking



DOWNLOAD the newest Prep4King DSA-C03 PDF dumps from Cloud Storage for free: https://drive.google.com/open?id=1YMC1VkHcTRnLPHRjQ44QJNArb45kf0jw

You can land your ideal job and advance your career with the Snowflake DSA-C03 certification. Success in the Snowflake DSA-C03 exam verifies your talent to perform crucial technical tasks. Preparation for this Snowflake DSA-C03 exam is a tricky task. Make sure you choose the top-notch Snowflake DSA-C03 Study Materials to get ready for this exam. For your smooth DSA-C03 test preparation, Prep4King provides updated DSA-C03 practice material with a success guarantee.

If you choose our DSA-C03 study materials and use our products well, we can promise that you can pass the DSA-C03 exam and get the DSA-C03 certification. Then you will find you have so many chances to advance in stages to a great level of social influence and success. Our DSA-C03 Guide Torrent can also provide all candidates with our free demo, in order to exclude your concerts that you can check our DSA-C03 exam questions. We believe that you will be fond of our DSA-C03 learning guide.

>> Snowflake DSA-C03 Test Answers <<

DSA-C03 study materials: SnowPro Advanced: Data Scientist Certification Exam & DSA-C03 exam torrent & DSA-C03 actual exam

You will never know what kind of people you will be and what kind of future is waiting for you if you don't try your best to pursue. And our DSA-C03 learning prep can be one of your challege. Also your potential will be fully realized with the guidance of our DSA-C03 Exam Questions. It is a good chance for you to improve yourself. We are looking forward that you can choose our DSA-C03 study materials. It is up to you. Time and tides wait for no man. Come to purchase our DSA-C03 practice braindumps.

Snowflake SnowPro Advanced: Data Scientist Certification Exam Sample Questions (Q66-Q71):

NEW QUESTION # 66
You are troubleshooting an external function in Snowflake that calls a model hosted on Google Cloud A1 Platform. The external

function consistently returns 'SQL compilation error: External function error: HTTP 400 Bad Request'. You have verified the API integration is correctly configured, and the Google Cloud project has the necessary permissions. Which of the following is the most likely cause of this error, and how would you best diagnose it?

- A. There is a mismatch between the request headers sent by Snowflake and what the Google Cloud AI Platform endpoint expects, specifically the 'Content-Type'. Diagnose by examining the headers being sent by Snowflake and ensuring they match the expected format.
- B. The request payload being sent by Snowflake exceeds the maximum size limit allowed by Google Cloud AI Platform. Diagnose by reducing the size of the input data and testing again.
- C. The API integration in Snowflake is missing the necessary authentication credentials for Google Cloud. Diagnose by re-creating the API integration and ensuring the correct service account and scopes are configured.
- D. The Google Cloud AI Platform model is unavailable or experiencing issues. Diagnose by checking the Google Cloud status dashboard for AI Platform outages.
- E. The issue is most likely due to incorrect data types being passed from Snowflake to the Google Cloud A1 Platform model. Diagnose by examining the input data being sent to the function and comparing it to the model's expected input schema.

Answer: E

Explanation:
A 400 Bad Request error typically indicates that the server (Google Cloud A1 Platform in this case) received a request that it could not understand. This often means the data being sent is in an incorrect format or does not conform to the expected schema. While the other options could potentially cause issues, a 400 error is most directly linked to data type mismatches or schema violations. Diagnosing this involves carefully inspecting the data being sent by Snowflake and comparing it to the model's input requirements. Google Cloud logging or network tracing could be necessary in complex situations to identify discrepancies. The use of REQUEST and RESPONSE translators can mitigate these issues.

NEW QUESTION # 67
You've built a model in Snowflake to predict house prices based on features like location, square footage, and number of bedrooms. After deploying the model, you want to ensure that the incoming data used for prediction is similar to the data the model was trained on. You decide to implement a data distribution comparison strategy. Consider these options and select all that apply:

- A. Create a binary classification model in Snowflake that attempts to predict whether a given row of data comes from the training dataset or the incoming dataset. If the model achieves high accuracy, it indicates a significant difference in data distributions.
- B. Only focus on monitoring the target variable (house price) and assume that if the distribution of house prices remains stable, the input data distribution is also stable.
- C. Generate histograms for each numerical feature in both the training and incoming datasets using a Python UDF that leverages libraries like Pandas and Matplotlib. Visually compare the histograms to identify potential distribution shifts.
- D. Calculate the mean and standard deviation for each numerical feature in both the training and incoming datasets using Snowflake SQL. Create a Snowflake Alert that triggers if the difference in means or standard deviations exceeds a predefined threshold for any feature.
- E. Use Snowflake's built-in statistics functions to compute quantiles (e.g., 25th, 50th, 75th percentiles) for each numerical feature. Compare these quantiles between the training and incoming datasets and set up alerts for significant deviations.

Answer: A,D,E

Explanation:
Options A, C, and E are all valid and effective strategies for comparing data distributions. Option A provides a simple and easily implementable approach using basic descriptive statistics. Option C is more robust than just mean and standard deviation, providing a more detailed comparison of the distributions. Option E utilizes a machine learning approach to quantify the difference in distributions, which can be very sensitive to subtle changes. Option B is helpful for visualizations but lacks an automated, quantifiable alert mechanism. Option D is flawed because the target variable's distribution can remain stable even if the input data distribution changes due to compensatory effects or other factors.

NEW QUESTION # 68
You are tasked with identifying Personally Identifiable Information (PII) within a Snowflake table named 'customer data'. This table contains various columns, some of which may contain sensitive information like email addresses and phone numbers. You want to use Snowflake's data governance features to tag these columns appropriately. Which of the following approaches is the MOST effective and secure way to automatically identify and tag potential PII columns with the 'PII CLASSIFIED tag in your Snowflake

environment, ensuring minimal manual intervention and optimal accuracy?

- A. Write a SQL script to query the 'INFORMATION SCHEMA.COLUMNS' view, identify columns with names containing keywords like 'email' or 'phone', and then apply the 'PII_CLASSIFIED tag to those columns.
- B. Manually inspect each column in the 'customer_data' table and apply the 'PII_CLASSIFIED' tag to columns that appear to contain PII based on their names and a small sample of data.
- C. Use Snowflake's built-in classification feature with a pre-defined sensitivity category to identify potential PII columns. Associate a masking policy that redacts the data, and apply a tag 'PII_CLASSIFIED' via automated tagging to the columns identified as containing PII.
- D. Export the 'customer_data' to a staging area in cloud storage, use a third-party data discovery tool to scan for PII, and then manually apply the "PII_CLASSIFIED' tag to the corresponding columns in Snowflake based on the tool's findings.
- E. Create a custom Snowpark for Python UDF that uses regular expressions to analyze the data in each column and apply the 'PII_CLASSIFIED tag if a match is found. Schedule this UDF to run periodically using Snowflake Tasks.

Answer: C

Explanation:
Snowflake's built-in classification feature is the most effective because it uses machine learning models to automatically identify sensitive data with a high degree of accuracy. Associating masking policies with the identified columns provides additional data protection. Automated tagging further streamlines the governance process. Option A, while viable, requires custom code and maintenance. Option C is manual and error-prone. Option D is based solely on column names and can lead to false positives and negatives. Option E introduces unnecessary complexity and security risks by exporting data.

NEW QUESTION # 69
You've trained a sentiment analysis model in Snowflake using Snowpark Python and deployed it as a UDF. After several weeks, you notice the model's performance has degraded significantly. You suspect concept drift. Which of the following actions represent the MOST effective and comprehensive approach to address this situation, considering the entire Machine Learning Lifecycle, including monitoring, retraining, and model versioning? Assume you have monitoring in place that alerted you to the drift.

- A. Immediately replace the current UDF with a newly trained model using the latest data, ignoring model versioning and assuming the latest data will solve the drift issue.
- B. Retrain the model on a sample of the most recent data, overwriting the original model files in your Snowflake stage and updating the UDF definition. Keep no record of the old model.
- C. Adjust the existing model's parameters manually to compensate for the observed performance degradation without retraining or versioning.
- D. Disable the model and revert to a rule-based system, abandoning the machine learning approach altogether.
- E. Analyze the recent data to understand the nature of the concept drift, retrain the model with a combination of historical and recent data, version the new model, and perform AIB testing against the existing model before fully deploying the new version. Log both model version predictions during AIB testing.

Answer: E

Explanation:
Addressing concept drift requires careful analysis, retraining with appropriate data (historical + recent), and controlled deployment using A/B testing. Model versioning ensures that you can rollback if the new model performs poorly. Logging the predictions of both models assists in further performance analysis. Directly replacing (option A) or manually adjusting (option C) are risky without proper evaluation. Abandoning the ML approach (option D) is a last resort. Option E lacks model versioning and thus risks complete loss of the older model which is a common practice violation in ML engineering.

NEW QUESTION # 70
A data scientist is tasked with predicting customer churn for a telecommunications company using Snowflake. The dataset contains call detail records (CDRs), customer demographic information, and service usage data'. Initial analysis reveals a high degree of multicollinearity between several features, specifically 'total_day_minutes', 'total_eve_minutes', and 'total_night_minutes'. Additionally, the 'state' feature has a large number of distinct values. Which of the following feature engineering techniques would be MOST effective in addressing these issues to improve model performance, considering efficient execution within Snowflake?

- A. Apply Principal Component Analysis (PCA) to reduce the dimensionality of the CDR features ('total_day_minutes', 'total_eve_minutes', 'total_night_minutes') and use one-hot encoding for the 'state' feature.
- B. Create interaction features by multiplying 'total_day_minutes' with 'customer_service_calls' and applying a target encoding

to the 'state' feature.

- C. Apply min-max scaling to the CDR features to normalize them and use label encoding for the 'state' feature. Train a decision tree model, as it is robust to multicollinearity.
- D. Calculate the Variance Inflation Factor (VIF) for each CDR feature and drop the feature with the highest VIE Apply frequency encoding to the 'state' feature.
- E. Use a variance threshold to remove highly correlated CDR features and create a feature representing the geographical region (e.g., 'Northeast', 'Southwest') based on the 'state' feature using a custom UDF.

Answer: E

Explanation:
Option C is the most effective. Using a variance threshold directly addresses multicollinearity by removing redundant features. Creating a geographical region feature from 'state' reduces dimensionality and is more manageable than one-hot encoding for high cardinality features. A custom UDF can be used for efficient regional mapping. While PCA can reduce dimensionality, it can also make the features less interpretable. Target encoding (B) can introduce target leakage if not handled carefully. VIF calculation (D) is useful but doesn't directly address the high cardinality of 'state'. Label encoding (E) is not appropriate for nominal categorical features like 'state' as it introduces ordinality.

NEW QUESTION # 71
......

If you want to check the quality and validity of our DSA-C03 exam questions, then you can click on the free demos on the website. The free demo has three versions. We only send you the PDF version of the DSA-C03 study questions. We have shown the rest two versions on our website. All in all, you will have a comprehensive understanding of various DSA-C03 practice materials. Then after deliberate considerations, you can directly purchase the most suitable one for yourself.

DSA-C03 Valid Test Testking: https://www.prep4king.com/DSA-C03-exam-prep-material.html

After validating the product's license, you won't need an active internet connection to use the desktop SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) practice test software, Even the DSA-C03 test syllabus is changing every year, Are you still worried about the actuality and the accuracy of the DSA-C03 exam cram, Snowflake DSA-C03 Test Answers Our system is strictly protect the clients' privacy and sets strict interception procedures to forestall the disclosure of the clients' private important information, Additionally, you can set limit time to practice your DSA-C03 dumps pdf.

But we think it's useful to have southern DSA-C03 Test Engine states included, Tips and notes to help you do even more, After validating the product's license, you won't need an active internet connection to use the desktop SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) practice test software.

DSA-C03 Test Answers | Professional SnowPro Advanced: Data Scientist Certification Exam 100% Free Valid Test Testking

Even the DSA-C03 test syllabus is changing every year, Are you still worried about the actuality and the accuracy of the DSA-C03 exam cram, Our system is strictly protect the clients' privacy and sets strict DSA-C03 interception procedures to forestall the disclosure of the clients' private important information.

Additionally, you can set limit time to practice your DSA-C03 dumps pdf.

- DSA-C03 Exam Torrent - DSA-C03 Test Collection - DSA-C03 Top Quiz Search for ➡ DSA-C03 and obtain a free download on 【 www.vceengine.com 】 DSA-C03 Latest Exam Discount
- New Exam DSA-C03 Materials Valid Braindumps DSA-C03 Ppt DSA-C03 Valid Exam Cram Search for 「 DSA-C03 」 and download exam materials for free through www.pdfvce.com Reliable DSA-C03 Study Guide
- DSA-C03 – 100% Free Test Answers | Authoritative SnowPro Advanced: Data Scientist Certification Exam Valid Test Testking Search on ▷ www.itcerttest.com ◁ for （ DSA-C03 ） to obtain exam materials for free download Latest DSA-C03 Exam Camp
- DSA-C03 Reliable Exam Registration Free DSA-C03 Study Material DSA-C03 Latest Exam Discount Easily obtain DSA-C03 for free download through ➠ www.pdfvce.com DSA-C03 Reliable Torrent
- DSA-C03 – 100% Free Test Answers | Authoritative SnowPro Advanced: Data Scientist Certification Exam Valid Test Testking Search for ➤ DSA-C03 and obtain a free download on 【 www.real4dumps.com 】 Reliable DSA-C03 Exam Test
-