Real DSA-C03 Questions | DSA-C03 Reliable Exam Sample

P.S. Free & New DSA-C03 dumps are available on Google Drive shared by PassCollection: https://drive.google.com/open?id=1tqEPUk9zTuu8EhLH309uKk2Erx_2CosI

The emerging Snowflake field creates a space for SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) certification exam holders to accelerate their careers. Many unfortunate candidates don't get the SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) certification because they prepare for its SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) exam questions from an Snowflake DSA-C03 exam that dumps outdated material. It results in a waste of time and money. You can develop your skills and join the list of experts by earning this SnowPro Advanced: Data Scientist Certification Exam (DSA-C03) certification exam.

PassCollection provides you with actual Snowflake DSA-C03 dumps in PDF format, Desktop-Based Practice tests, and Web-based Practice exams. These 3 formats of SnowPro Advanced: Data Scientist Certification Exam exam preparation are easy to use. This is a printable Snowflake DSA-C03 PDF dumps file. The Snowflake DSA-C03 Pdf Dumps enables you to study without any device, as it is a portable and easily shareable format, thus you can study Snowflake DSA-C03 dumps on your preferred smart device such as your smartphone or in hard copy format.

>> Real DSA-C03 Questions <<

Fantastic Snowflake - Real DSA-C03 Questions

As an old saying goes: Practice makes perfect. Facts prove that learning through practice is more beneficial for you to learn and test at the same time as well as find self-ability shortage in DSA-C03 test prep. The PC test engine of our DSA-C03 exam torrent is designed for such kind of condition, when the system of the DSA-C03 Exam Torrent has renovation of production techniques by actually simulating the test environment. Until then, you will have more practical experience and get improvement rapidly through our DSA-C03 quiz guide.

Snowflake SnowPro Advanced: Data Scientist Certification Exam Sample Questions (Q200-Q205):

NEW QUESTION # 200
You have deployed a machine learning model in Snowflake to predict customer churn. The model was trained on data from the past year. After six months of deployment, you notice the model's recall for identifying churned customers has dropped significantly. You suspect model decay. Which of the following Snowflake tasks and monitoring strategies would be MOST appropriate to diagnose and address this model decay?

- A. Use Snowflake's data sharing feature to share the model's predictions with a separate analytics team. Let them monitor the overall customer churn rate and notify you if it changes significantly.
- B. Establish a Snowflake pipe to continuously ingest feedback data (actual churn status) into a feedback table. Write a stored procedure to calculate performance metrics (e.g., recall, precision) on a sliding window of recent data. Create a Snowflake Alert that triggers when recall falls below a defined threshold.
- C. Back up the original training data to secure storage. Ingest all new data as it comes in. Retrain a new model and compare its performance with the backed-up training data.
- D. Create a Snowflake Task that automatically retrains the model weekly with the most recent six months of data. Monitor the model's performance metrics using Snowflake's query history to track the accuracy of the predictions.
- E. Implement a Shadow Deployment strategy in Snowflake. Route a small percentage of incoming data to both the existing model and a newly trained model. Compare the predictions from both models using a UDF that calculates the difference in predicted probabilities. Trigger an alert if the differences exceed a certain threshold.

Answer: B,E

Explanation:
Option B is the most comprehensive. It establishes a system for continuous monitoring of model performance using real-world feedback, and alerts you when performance degrades. Option E is also strong because it allows for direct comparison of a new model against the existing model in a production setting, identifying model decay before it significantly impacts performance. Options A and D are insufficient for monitoring as they lack real-world feedback loops for continuous assessment. Simply retraininig frequently does not guarantee model improvements, and option C relies on manual intervention and lacks granular monitoring of the model's specific performance. Shadow Deployment is costly but more robust.

NEW QUESTION # 201
You are preparing a dataset in Snowflake for a K-means clustering algorithm. The dataset includes features like 'age', 'income' (in USD), and 'number of_transactions'. 'Income' has significantly larger values than 'age' and 'number of_transactions'. To ensure that all features contribute equally to the distance calculations in K-means, which of the following scaling approaches should you consider, and why? Select all that apply:

- A. Apply StandardScaler to all three features ('age', 'income', 'number_of_transactions') to center the data around zero and scale it to unit variance.
- B. Do not scale the data, as K-means is robust to differences in feature scales.
- C. Apply PowerTransformer to transform income and StandardScaler to other features to handle skewness.
- D. Apply MinMaxScaler to all three features to scale them to a range between O and 1 .
- E. Apply RobustScaler to handle outliers and then StandardScaler or MinMaxScaler to further scale the features.

Answer: A,D,E

Explanation:
K-means clustering is sensitive to the scale of the features because it relies on distance calculations. Features with larger values will have a disproportionate influence on the clustering results. StandardScaler centers the data around zero and scales it to unit variance, which ensures that all features have a similar range and variance. MinMaxScaler scales the features to a range between O and 1, which also addresses the issue of different scales. RobustScaler handles outliers which will then use the other two scaling techniques. Therefore A, B and D are the appropriate scaling techniques. C is not correct as K-means relies on distance calculations and not scaling the data could give some feature a larger weight which isn't the desired outcome. Option E: Using PowerTransformer on 'income' to reduce skewness and StandardScaler on the other features can be a valid approach, but it depends on the distribution of 'income' and the presence of outliers. If 'income' is highly skewed and/or contains outliers, this combination might be more effective than using StandardScaler or MinMaxScaler alone.

NEW QUESTION # 202
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. 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.
- C. Create interaction features by multiplying 'total_day_minutes' with 'customer_service_calls' and applying a target encoding to the 'state' feature.
- D. 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.
- 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 # 203
A data scientist is tasked with building a predictive maintenance model for industrial equipment. The data is collected from IoT sensors and stored in Snowflake. The raw sensor data is voluminous and contains noise, outliers, and missing values. Which of the following code snippets, executed within a Snowflake environment, demonstrates the MOST efficient and robust approach to cleaning and transforming this sensor data during the data collection phase, specifically addressing outlier removal and missing value imputation using robust statistics? Assume necessary libraries like numpy and pandas are available via Snowpark.

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

Answer: B

Explanation:
Option E is the MOST robust and efficient. It uses the interquartile range (IQR) method, which is less sensitive to extreme outliers than the z-score method in Option A. It also utilizes 'approx_quantile' and is therefore more optimized for Snowflake large datasets. The median is also a more robust measure of central tendency for imputation than the mean when dealing with outliers. Option C uses a hard-coded threshold for outlier removal and imputes with 0, which is not adaptive or robust. Option D skips data cleaning altogether.Option A uses z-score which may work however, since IoT has continuous streaming data quantile based outlier removal is better. It is more optimised for large dataset and better at handling streaming datasets.

NEW QUESTION # 204
You are tasked with creating a new feature in a machine learning model for predicting customer lifetime value. You have access to a table called 'CUSTOMER ORDERS which contains order history for each customer. This table contains the following columns: 'CUSTOMER ID', 'ORDER DATE, and 'ORDER AMOUNT. To improve model performance and reduce the impact of outliers, you plan to bin the 'ORDER AMOUNT' column using quantiles. You decide to create 5 bins, effectively creating quintiles. You also want to create a derived feature indicating if the customer's latest order amount falls in the top quintile. Which of the following approaches, or combination of approaches, is most appropriate and efficient for achieving this in Snowflake? (Choose all that apply)

- A. Create a temporary table storing quintile information, then join this table to original table to find the top quintile order amount.
- B. Use 'WIDTH_BUCKET function, after finding the boundaries of quantile using 'APPROX_PERCENTILE' or 'PERCENTILE_CONT. Using MAX(ORDER to determine recent amount is in top quantile.

- C. Calculate the 20th, 40th, 60th, and 80th percentiles of the 'ORDER AMOUNT' using 'APPROX PERCENTILE or 'PERCENTILE CONT and then use a 'CASE statement to assign each order to a quantile bin. Calculate and see if on that particular date is in top quintile.
- D. Use the window function to create quintiles for 'ORDER AMOUNT and then, in a separate query, check if the latest 'ORDER AMOUNT for each customer falls within the NTILE that represents the top quintile.
- E. Use a Snowflake UDF (User-Defined Function) written in Python or Java to calculate the quantiles and assign each 'ORDER AMOUNT to a bin. Later you can use other statement to check the top quintile amount from result set.

Answer: B,C,D

Explanation:
Options A, B, and E are valid and efficient approaches. Option A using 'NTILE' is a direct and efficient way to create quantile bins within Snowflake SQL, and can find the most recent order date for customer with a case statement. Option B calculates the percentiles directly and then uses a CASE statement to assign bins. This is also efficient for explicit boundaries. Option E finds the boundaries of the quantile using 'APPROX_PERCENTILE or 'PERCENTILE_CONT , after that you can use 'WIDTH_BUCKET to categorize into quantile bins based on ranges. Option C is possible but generally less efficient due to the overhead of UDF execution and data transfer between Snowflake and the UDF environment. Option D is valid, but creating a temporary table adds complexity and potentially reduces performance compared to window functions or direct quantile calculation within the query.

NEW QUESTION # 205
......

From the time you purchase, use, and pass the DSA-C03 exam, we will be with you all the time. You can seek our help anytime, anywhere. If you have experienced a very urgent problem while using DSA-C03 exam simulating, you can immediately contact online customer service, you'd praise the staff of DSA-C03 study engine, because they can solve any problems you have encountered while using DSA-C03 exam simulating. All we do is just want you to concentrate on DSA-C03 exam learning, Do not hesitate anymore. You will never regret buying DSA-C03 study engine!

Snowflake Real DSA-C03 Questions When I chose the IT industry I have proven to God my strength, Snowflake Real DSA-C03 Questions With such a high hit rate, it becomes much easier to pass the exam, We provide excellent services for passing DSA-C03 exam, Snowflake Real DSA-C03 Questions In case of any trouble relating o your purchase or downloading, our online support chat service is available all the time, Using our DSA-C03 study materials can also save your time in the exam preparation for the content is all the keypoints covered.

Even the earphones and their little wired remote are Real DSA-C03 Questions subject to destruction, and not merely by vandals with scissors, For iPhone users, for example, you can track your monthly talk time, how many text DSA-C03 messages you've sent and received, and track your wireless data usage with a few taps on the screen.

DSA-C03 Quiz Materials - DSA-C03 Exam Guide & DSA-C03 Exam Collection

When I chose the IT industry I have proven to God my strength, With such a high hit rate, it becomes much easier to pass the exam, We provide excellent services for passing DSA-C03 exam.

In case of any trouble relating o your purchase DSA-C03 Exam Pattern or downloading, our online support chat service is available all the time, Using our DSA-C03 study materials can also save your time in the exam preparation for the content is all the keypoints covered.

- DSA-C03 Latest Test Question Vce DSA-C03 Test Simulator Latest DSA-C03 Exam Questions Vce Search for ✔ DSA-C03 ✔ and obtain a free download on 《 www.exam4labs.com 》 Reliable DSA-C03 Exam Questions
- 100% Pass DSA-C03 - Perfect Real SnowPro Advanced: Data Scientist Certification Exam Questions Download 《 DSA-C03 》 for free by simply entering { www.pdfvce.com } website Certification DSA-C03 Training
- DSA-C03 Latest Study Questions Top DSA-C03 Exam Dumps DSA-C03 Latest Test Cost Open ➡ www.dumpsquestion.com enter " DSA-C03 " and obtain a free download Latest DSA-C03 Exam Materials
- Pass Guaranteed Quiz Snowflake - Accurate Real DSA-C03 Questions Simply search for ✔ DSA-C03 ✔ for free download on ⇒ www.pdfvce.com ⇐ New DSA-C03 Test Tutorial
- Real DSA-C03 Questions: 2026 Realistic Snowflake SnowPro Advanced: Data Scientist Certification Exam Reliable Exam Sample Pass Guaranteed The page for free download of DSA-C03 on www.vceengine.com will open immediately DSA-C03 Latest Test Question
- Latest DSA-C03 Exam Questions Vce Latest DSA-C03 Exam Materials DSA-C03 Latest Test Question Open website www.pdfvce.com and search for { DSA-C03 } for free download New DSA-C03 Test Tutorial
- Snowflake - Fantastic Real DSA-C03 Questions The page for free download of （ DSA-C03 ） on " www.prep4sures.top " will open immediately Latest DSA-C03 Exam Materials
- www.stes.tyc.edu.tw, 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, 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, allprotrainings.com, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, essarag.org, www.stes.tyc.edu.tw, Disposable vapes