DSA-C02Q&As

SnowPro Advanced: Data Scientist Certification

Pass Snowflake DSA-C02 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.geekcert.com/dsa-c02.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Snowflake Official Exam Center

- **Instant Download** After Purchase
- **100% Money Back** Guarantee
- **365 Days** Free Update
- **800,000+** Satisfied Customers

QUESTION 1

Which one is not the types of Feature Engineering Transformation?

A. Scaling

B. Encoding

C. Aggregation

D. Normalization

Correct Answer: C

Explanation:

What is Feature Engineering?

Feature engineering is the process of transforming raw data into features that are suitable for ma-chine learning models. In other words, it is the process of selecting, extracting, and transforming the most relevant features from the available

data to build more accurate and efficient machine learning models.

The success of machine learning models heavily depends on the quality of the features used to train them. Feature engineering involves a set of techniques that enable us to create new features by combining or transforming the existing

ones. These techniques help to highlight the most important pat-terns and relationships in the data, which in turn helps the machine learning model to learn from the data more effectively.

What is a Feature?

In the context of machine learning, a feature (also known as a variable or attribute) is an individual measurable property or characteristic of a data point that is used as input for a machine learning al-gorithm. Features can be numerical,

categorical, or text-based, and they represent different aspects of the data that are relevant to the problem at hand. For example, in a dataset of housing prices, features could include the number of bedrooms, the square footage, the location,

and the age of the property. In a dataset of customer demographics, features could include age, gender, income level, and occupation. The choice and quality of features are critical in machine learning, as they can greatly impact the accuracy and performance of the model.

Why do we Engineer Features?

We engineer features to improve the performance of machine learning models by providing them with relevant and informative input data. Raw data may contain noise, irrelevant information, or missing values, which can lead to inaccurate or

biased model predictions. By engineering features, we can extract meaningful information from the raw data, create new variables that capture important patterns and relationships, and transform the data into a more suitable format for

machine learning algorithms. Feature engineering can also help in addressing issues such as overfitting, underfitting, and high di-mensionality. For example, by reducing the number of features, we can prevent the model from be-coming too

complex or overfitting to the training data. By selecting the most relevant features, we can improve the model\\'s accuracy and interpretability. In addition, feature engineering is a crucial step in preparing data for analysis and decision- making

in various fields, such as finance, healthcare, marketing, and social sciences. It can help uncover hidden insights, identify trends and patterns, and support data-driven decision-making.

We engineer features for various reasons, and some of the main reasons include:

Improve User Experience: The primary reason we engineer features is to enhance the user experience of a product or service. By adding new features, we can make the product more intuitive, efficient, and user-friendly, which can increase

user satisfaction and engagement. Competitive Advantage: Another reason we engineer features is to gain a competitive advantage in the marketplace. By offering unique and innovative features, we can differentiate our product from

competitors and attract more customers. Meet Customer Needs: We engineer features to meet the evolving needs of customers. By analyzing user feedback, market trends, and customer behavior, we can identify areas where new features

could enhance the product\\'s value and meet customer needs. Increase Revenue: Features can also be engineered to generate more revenue. For example, a new feature that streamlines the checkout process can increase sales, or a feature

that provides additional functionality could lead to more upsells or cross-sells. Future-Proofing: Engineering features can also be done to future-proof a product or service. By an-ticipating future trends and potential customer needs, we can

develop features that ensure the product remains relevant and useful in the long term.

Processes Involved in Feature Engineering

Feature engineering in Machine learning consists of mainly 5 processes: Feature Creation, Feature Transformation, Feature Extraction, Feature Selection, and Feature Scaling. It is an iterative process that requires experimentation and

testing to find the best combination of features for a given problem. The success of a machine learning model largely depends on the quality of the features used in the model.

Feature Transformation

Feature Transformation is the process of transforming the featuresinto a more suitable representation for the machine learning model. This is done to ensure that the model can effectively learn from the data.

Types of Feature Transformation:

Normalization: Rescaling the features to have a similar range, such as between 0 and 1, to prevent some features from dominating others.

Scaling: Rescaling the features to have a similar scale, such as having a standard deviation of 1, to make sure the model considers all features equally. Encoding: Transforming categorical features into a numerical representation. Examples

are one-hot encoding and label encoding.

Transformation: Transforming the features using mathematical operations to change the distribution or scale of the features. Examples are logarithmic, square root, and reciprocal transformations.

QUESTION 2

Select the Data Science Tools which are known to provide native connectivity to Snowflake?

A. Denodo

B. DvSUM

C. DiYotta

D. HEX

Correct Answer: D

Explanation:

Hex -- collaborative data science and analytics platform Denodo -- data virtualization and federation platform DvSum -- data catalog and data intelligence platform Diyotta -- data integration and migration

QUESTION 3

Which one is the incorrect option to share data in Snowflake?

A. a Listing, in which you offer a share and additional metadata as a data product to one or more accounts.

B. a Direct Marketplace, in which you directly share specific database objects (a share) to another account in your region using Snowflake Marketplace.

C. a Direct Share, in which you directly share specific database objects (a share) to anoth- er account in your region.

D. a Data Exchange, in which you set up and manage a group of accounts and offer a share to that group.

Correct Answer: B

Explanation:

Options for Sharing in Snowflake

You can share data in Snowflake using one of the following options:

a Listing, in which you offer a share and additional metadata as a data product to one or more ac-counts,

a Direct Share, in which you directly share specific database objects (a share) to another account in your region,

a Data Exchange, in which you set up and manage a group of accounts and offer a share to that group.

QUESTION 4

Which of the following Functions do Support Windowing?

A. HASH_AGG

B. ENCRYPT

C. EXTRACT

D. LISTAGG

Correct Answer: D

Explanation:

What is a Window?

A window is a group of related rows. For example, a window might be defined based on timestamps, with all rows in the same month grouped in the same window. Or a window might be defined based on location, with all rows from a

particular city grouped in the same window.

A window can consist of zero, one, or multiple rows. For simplicity, Snowflake documentation usually says that a window contains multiple rows.

What is a Window Function?

A window function is any function that operates over a window of rows. A window function is generally passed two parameters:

A row. More precisely, a window function is passed 0 or more expressions. In almost all cases, at least one of those expressions references a column in that row. (Most window functions require at least one column or expression, but a few

window functions, such as some rank-related functions, do not required an explicit column or expression.) A window of related rows that includes that row. The window can be the entire table, or a subset of the rows in the table.

For non-window functions, all arguments are usually passed explicitly to the function, for example:

MY_FUNCTION(argument1, argument2, ...)

Window functions behave differently; although the current row is passed as an argument the normal way, the window is passed through a separate clause, called an OVER clause. The syntax of the OVER clause is documented later.

LISTAGG

Returns the concatenated input values, separated by the delimiter string.

Window function

1.LISTAGG([DISTINCT] [,]) 2.[WITHIN GROUP ( )]

3.OVER ([PARTITION BY])

HASH_AGG

Returns an aggregate signed 64-bit hash value over the (unordered) set of input rows. HASH_AGG never returns NULL, even if no input is provided. Empty input "hashes" to 0.

Window function

HASH_AGG([DISTINCT] [, ...]) OVER ([PARTITION BY]) HASH_AGG(*) OVER ([PARTITION BY])

QUESTION 5

Mark the incorrect statement regarding usage of Snowflake Stream and Tasks?

A. Snowflake automatically resizes and scales the compute resources for serverless tasks.

B. Snowflake ensures only one instance of a task with a schedule (i.e. a standalone task or the root task in a DAG) is executed at a given time. If a task is still running when the next scheduled execution time occurs, then that scheduled time is skipped.

C. Streams support repeatable read isolation.

D. An standard-only stream tracks row inserts only.

Correct Answer: D

Explanation:

All are correct except a standard-only stream tracks row inserts only. A standard (i.e. delta) stream tracks all DML changes to the source object, including inserts, up-dates, and deletes (including table truncates).