DSA-C02 Q&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.pass4itsure.com/dsa-c02.html

100% Passing Guarantee
100% Money Back Assurance

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

QUESTION 1

Which object records data manipulation language (DML) changes made to tables, including inserts, updates, and deletes, as well as metadata about each change, so that actions can be taken using the changed data of Data Science Pipelines?

A. Task

B. Dynamic tables

C. Stream

D. Tags

E. Delta

F. OFFSET

Correct Answer: C

Explanation: A stream object records data manipulation language (DML) changes made to tables, including inserts, updates, and deletes, as well as metadata about each change,so that actions can be taken using the changed data. This process is referred to as change data capture (CDC). An individual table stream tracks the changes made to rows in a source table. A table stream (also referred to as simply a "stream") makes a "change table" available of what changed, at therow level, between two transactional points of time in a table. This allows querying and consuming a sequence of change records in a transactional fashion. Streams can be created to query change data on the following objects: Standard tables, including shared tables. Views, including secure views Directory tables Event tables

QUESTION 2

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 3

Secure Data Sharing do not let you share which of the following selected objects in a database in your account with other Snowflake accounts?

A. Sequences

B. Tables

C. External tables

D. Secure UDFs

Correct Answer: A

Explanation: Secure Data Sharing lets you share selected objects in a database in your account with other Snow-flake accounts. You can share the following Snowflake database objects: Tables External tables Secure views Secure materialized views Secure UDFs Snowflake enables the sharing of databases through shares, which are created by data providers and "imported" by data consumers.

QUESTION 4

Which command manually triggers a single run of a scheduled task (either a standalone task or the root task in a DAG) independent of the schedule defined for the task?

A. RUN TASK

B. CALL TASK

C. EXECUTE TASK

D. RUN ROOT TASK

Correct Answer: C

Explanation: The EXECUTE TASK command manually triggers a single run of a scheduled task (either a standalone task or the root task in a DAG) independent of the schedule defined for the task. A successful run of a roottask triggers a cascading run of child tasks in the DAG as their precedent task completes, as though the root task had run on its defined schedule. This SQL command is useful for testing new or modified standalone tasks and DAGs before you enable them to execute SQL code in production. Call this SQL command directly in scripts or in stored procedures. In addition, this command sup-ports integrating tasks in external data pipelines. Any third-party services that can authenticate into your Snowflake account and authorize SQL actions can execute the EXECUTE TASK command to run tasks.

QUESTION 5

In a simple linear regression model (One independent variable), If we change the input variable by 1 unit. How much output variable will change?

A. by 1

B. no change

C. by intercept

D. by its slope

Correct Answer: D

Explanation:

What is linear regression?

Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable\\'s value is

called the independent variable.

Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatoryvariable, and the other is considered to be a dependent variable. For

example, a modeler might want to relate the weights of individuals to their heights using a linear regression model. A linear regression line has an equation of the form Y = a + bX, where X is the explanatory variable and Y is the dependent

variable. The slope of the line is b, and a is the intercept (the value of y when x = 0).

For linear regression Y=a+bx+error.

If neglect error then Y=a+bx. If x increases by 1, then Y = a+b(x+1) which implies Y=a+bx+b. So Y increases by its slope.

For linear regression Y=a+bx+error. If neglect error then Y=a+bx. If x increases by 1, then Y = a+b(x+1) which implies Y=a+bx+b. So Y increases by its slope.