



Snowflake DSA-C02 Practice Questions

Shared by Adams on 19-08-2026

For More Free Questions and Preparation Resources

Check the Links on Last Page



Question 1

Question Type: MultipleChoice

Which of the following is a Python-based web application framework for visualizing data and analyzing results in a more efficient and flexible way?

Options:

- A- StreamBI
- B- Streamlit
- C- Streamsets
- D- Rapter

Answer:

B

Explanation:

Streamlit is a Python-based web application framework for visualizing data and analyzing results in a more efficient and flexible way. It is an open source library that assists data scientists and academics to develop Machine Learning (ML) visualization dashboards in a short period of time. We can build and deploy powerful data applications with just a few lines of code.

Why Streamlit?

Currently, real-world applications are in high demand and developers are developing new libraries and frameworks to make on-the-go dashboards easier to build and deploy. Streamlit is a library that reduces your dashboard development time from days to hours. Following are some reasons to choose the Streamlit:

It is a free and open-source library.

Installing Streamlit is as simple as installing any other python package

It is easy to learn because you won't need any web development experience, only a basic understanding of Python is enough to build a data application.

It is compatible with almost all machine learning frameworks, including Tensorflow and Pytorch, Scikit-learn, and visualization libraries such as Seaborn, Altair, Plotly, and many others.

Question 2

Question Type: MultipleChoice

Which one is incorrect understanding about Providers of Direct share?

Options:

- A- A data provider is any Snowflake account that creates shares and makes them available to other Snowflake accounts to consume.
- B- As a data provider, you share a database with one or more Snowflake accounts.
- C- You can create as many shares as you want, and add as many accounts to a share as you want.
- D- If you want to provide a share to many accounts, you can do the same via Direct Share.

Answer:

D

Explanation:

If you want to provide a share to many accounts, you might want to use a listing or a data exchange.

Question 3

Question Type: MultipleChoice

A Data Scientist as data providers require to allow consumers to access all databases and database objects in a share by granting a single privilege on shared databases. Which one is incorrect SnowSQL command used by her while doing this task?

Assuming:

A database named product_db exists with a schema named product_agg and a table named Item_agg.

The database, schema, and table will be shared with two accounts named xy12345 and yz23456.

1. USE ROLE accountadmin;
2. CREATE DIRECT SHARE product_s;

3. GRANT USAGE ON DATABASE product_db TO SHARE product_s;
4. GRANT USAGE ON SCHEMA product_db. product_agg TO SHARE product_s;
5. GRANT SELECT ON TABLE sales_db. product_agg.Item_agg TO SHARE product_s;
6. SHOW GRANTS TO SHARE product_s;
7. ALTER SHARE product_s ADD ACCOUNTS=xy12345, yz23456;
8. SHOW GRANTS OF SHARE product_s;

Options:

- A- GRANT USAGE ON DATABASE product_db TO SHARE product_s;
- B- CREATE DIRECT SHARE product_s;
- C- GRANT SELECT ON TABLE sales_db. product_agg.Item_agg TO SHARE product_s;
- D- ALTER SHARE product_s ADD ACCOUNTS=xy12345, yz23456;

Answer:

C

Explanation:

CREATE SHARE product_s is the correct Snowsql command to create Share object.

Rest are correct ones.

<https://docs.snowflake.com/en/user-guide/data-sharing-provider#creating-a-share-using-sql>

Question 4

Question Type: MultipleChoice

All Snowpark ML modeling and preprocessing classes are in the _____ namespace?

Options:

- A- snowpark.ml.modeling
- B- snowflake.sklearn.modeling
- C- snowflake.scikit.modeling

D- snowflake.ml.modeling

Answer:

D

Explanation:

All Snowpark ML modeling and preprocessing classes are in the snowflake.ml.modeling namespace. The Snowpark ML modules have the same name as the corresponding module from the sklearn namespace. For example, the Snowpark ML module corresponding to sklearn.calibration is snowflake.ml.modeling.calibration.

The xgboost and lightgbm modules correspond to snowflake.ml.modeling.xgboost and snowflake.ml.modeling.lightgbm, respectively.

Not all of the classes from scikit-learn are supported in Snowpark ML.

Question 5

Question Type: MultipleChoice

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

Options:

- A- Denodo
- B- DvSUM
- C- DiYotta
- D- HEX

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 6

Question Type: MultipleChoice

Mark the Incorrect statements regarding MIN / MAX Functions?

Options:

- A- NULL values are skipped unless all the records are NULL
- B- NULL values are ignored unless all the records are NULL, in which case a NULL value is returned
- C- The data type of the returned value is the same as the data type of the input values
- D- For compatibility with other systems, the DISTINCT keyword can be specified as an argument for MIN or MAX, but it does not have any effect

Answer:

B

Explanation:

NULL values are ignored unless all the records are NULL, in which case a NULL value is returned

Question 7

Question Type: MultipleChoice

What is the formula for measuring skewness in a dataset?

Options:

- A- MEAN - MEDIAN

- B- MODE - MEDIAN
- C- $(3(\text{MEAN} - \text{MEDIAN}))/ \text{STANDARD DEVIATION}$
- D- $(\text{MEAN} - \text{MODE})/ \text{STANDARD DEVIATION}$

Answer:

C

Explanation:

Since the normal curve is symmetric about its mean, its skewness is zero. This is a theoretical explanation for mathematical proofs, you can refer to books or websites that speak on the same in detail.

Question 8

Question Type: MultipleChoice

Which of the following is a useful tool for gaining insights into the relationship between features and predictions?

Options:

- A- numpy plots
- B- sklearn plots
- C- Partial dependence plots(PDP)
- D- FULL dependence plots (FDP)

Answer:

C

Explanation:

Partial dependence plots (PDP) is a useful tool for gaining insights into the relationship between features and predictions. It helps us understand how different values of a particular feature impact model's predictions.

Question 9

Question Type: MultipleChoice

Which of the following metrics are used to evaluate classification models?

Options:

- A- Area under the ROC curve
- B- F1 score
- C- Confusion matrix
- D- Compare disparity between selection rates and performance metrics across ethnicities.

Answer:

D

Explanation:

Evaluation metrics are tied to machine learning tasks. There are different metrics for the tasks of classification and regression. Some metrics, like precision-recall, are useful for multiple tasks. Classification and regression are examples of supervised learning, which constitutes a majority of machine learning applications. Using different metrics for performance evaluation, we should be able to improve our model's overall predictive power before we roll it out for production on unseen data. Without doing a proper evaluation of the Machine Learning model by using different evaluation metrics, and only depending on accuracy, can lead to a problem when the respective model is deployed on unseen data and may end in poor predictions.

Classification metrics are evaluation measures used to assess the performance of a classification model. Common metrics include accuracy (proportion of correct predictions), precision (true positives over total predicted positives), recall (true positives over total actual positives), F1 score (harmonic mean of precision and recall), and area under the receiver operating characteristic curve (AUC-ROC).

Confusion Matrix

Confusion Matrix is a performance measurement for the machine learning classification problems where the output can be two or more classes. It is a table with combinations of predicted and actual values.

It is extremely useful for measuring the Recall, Precision, Accuracy, and AUC-ROC curves.

The four commonly used metrics for evaluating classifier performance are:

1. Accuracy: The proportion of correct predictions out of the total predictions.
2. Precision: The proportion of true positive predictions out of the total positive predictions

(precision = true positives / (true positives + false positives)).

3. Recall (Sensitivity or True Positive Rate): The proportion of true positive predictions out of the total actual positive instances (recall = true positives / (true positives + false negatives)).

4. F1 Score: The harmonic mean of precision and recall, providing a balance between the two metrics (F1 score = $2 * ((\text{precision} * \text{recall}) / (\text{precision} + \text{recall}))$).

These metrics help assess the classifier's effectiveness in correctly classifying instances of different classes.

Understanding how well a machine learning model will perform on unseen data is the main purpose behind working with these evaluation metrics. Metrics like accuracy, precision, recall are good ways to evaluate classification models for balanced datasets, but if the data is imbalanced then other methods like ROC/AUC perform better in evaluating the model performance.

ROC curve isn't just a single number but it's a whole curve that provides nuanced details about the behavior of the classifier. It is also hard to quickly compare many ROC curves to each other.

Question 10

Question Type: MultipleChoice

Consider a data frame df with 10 rows and index ['r1', 'r2', 'r3', 'row4', 'row5', 'row6', 'r7', 'r8', 'r9', 'row10']. What does the expression `g = df.groupby(df.index.str.len())` do?

Options:

- A- Groups df based on index values
- B- Groups df based on length of each index value
- C- Groups df based on index strings
- D- Data frames cannot be grouped by index values. Hence it results in Error.

Answer:

D

Explanation:

Data frames cannot be grouped by index values. Hence it results in Error.

To Get Premium Files for DSA-C02 Visit

<https://www.p2pexams.com/products/dsa-c02>

For More Free Questions Visit

<https://www.p2pexams.com/snowflake/pdf/dsa-c02>

20%
DISCOUNT

P2P
exams