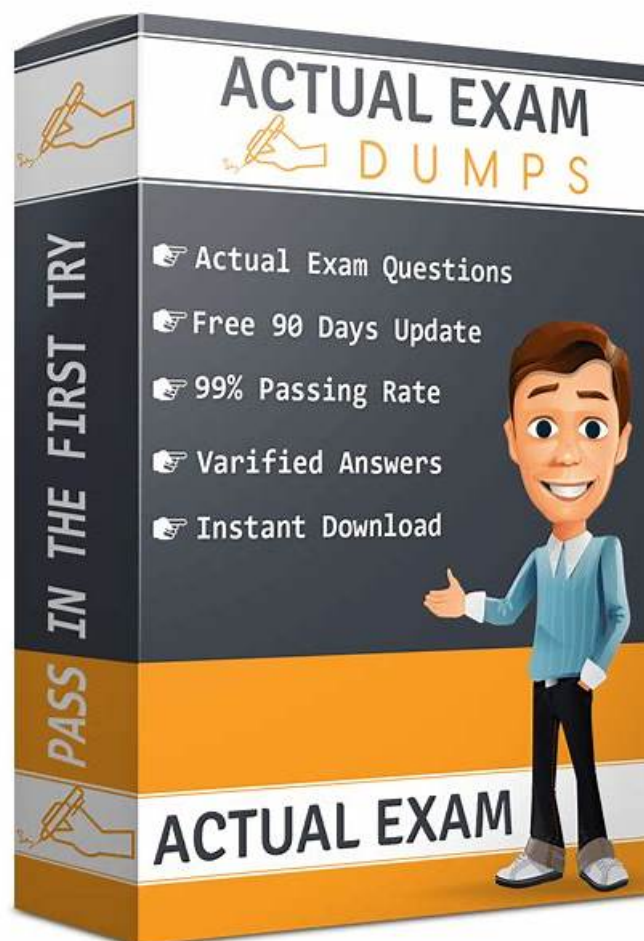


# DASCA SDS Exam PDF Dumps And Practice Test Software Is Ready For Download



The Senior Data Scientist SDS practice test is available in three compatible and user-friendly formats. These formats are SDS desktop practice test software, Senior Data Scientist SDS web-based practice exam, and DASCA SDS PDF dumps file. All three formats of SDS study material contain actual and verified Senior Data Scientist SDS Exam Dumps that will help you boost your exam preparation. The DASCA desktop practice test software and web-based SDS practice test both simulate the actual exam environment and identify your mistakes.

Your chances of passing the Senior Data Scientist (SDS) certification exam the first time around can be greatly improved if you attempt the ITCertMagic DASCA SDS practice exam. To help you succeed on your first try at the Senior Data Scientist (SDS) exam, ITCertMagic has created three formats of Senior Data Scientist (SDS) practice exam.

>> SDS Latest Dumps Questions <<

## Excellent SDS Latest Dumps Questions Help You to Get Acquainted with Real SDS Exam Simulation

We have brought in an experienced team of experts to develop our SDS study materials, which are close to the exam syllabus. With the help of our SDS study materials, you don't have to search all kinds of data, because our products are enough to meet your needs. You also don't have to spend all your energy to the exam because our SDS Study Materials are very efficient. Only should you spend a little time practicing them can you pass the exam successfully.

## DASCA Senior Data Scientist Sample Questions (Q25-Q30):

### NEW QUESTION # 25

Which classification steps are performed in inductive techniques?

- i. Training Step
- ii. Test Step
- iii. Validation Step
- iv. Application Step

- A. i, ii
- B. i, ii, iv
- C. ii, iii
- **D. i, ii, iii, iv**

**Answer: D**

Explanation:

Inductive learning techniques in machine learning (such as decision trees, neural networks, or SVMs) follow a systematic sequence of steps for classification:

Training Step (i): A model is built using training data, where the system learns relationships between features and target labels.

Test Step (ii): The trained model is evaluated on unseen test data to measure its performance and generalizability.

Validation Step (iii): Often, a validation set is used to fine-tune model parameters, avoid overfitting, and choose the best model configuration.

Application Step (iv): The final validated model is applied to classify new, real-world data.

Since all four steps (i, ii, iii, iv) are essential to inductive classification, the correct answer is Option D (i, ii, iii, iv).

Reference:

DASCA Data Scientist Knowledge Framework (DSKF) - Analytics & Machine Learning: Classification and Inductive Learning Techniques.

### NEW QUESTION # 26

IoT is built on:

- **A. Both A and B**
- B. Cloud Computing
- C. None of the above
- D. Networks of data gathering devices

**Answer: A**

Explanation:

The Internet of Things (IoT) is an ecosystem of interconnected devices that collect, transmit, and analyze data. IoT relies on two critical foundations:

Option A (Cloud Computing): IoT generates massive amounts of data, and cloud platforms provide scalable storage, analytics, and computing resources for real-time and batch processing.

Option B (Networks of data gathering devices): IoT relies on physical devices - sensors, smart appliances, industrial machines - that collect and transmit data through networks (Wi-Fi, Bluetooth, 5G, LPWAN).

Thus, IoT is fundamentally built on both cloud computing and networks of devices, making Option C correct.

Reference:

DASCA Data Scientist Knowledge Framework (DSKF) - Big Data & IoT Ecosystem Fundamentals.

### NEW QUESTION # 27

Example of amortized performance is:

- A. HDFS dictionaries
- **B. Python dictionaries**
- C. Hadoop dictionaries
- D. All of the above
- E. MapReduce dictionaries

**Answer: B**

Explanation:

Amortized performance refers to averaging the cost of operations over a sequence of actions, ensuring that while some operations may be costly, the overall average time per operation remains efficient.

Python Dictionaries (Option B): Implemented using hash tables. Insertions, deletions, and lookups typically run in  $O(1)$  average time, but occasionally require rehashing (costly). The high cost of rehashing is spread over many operations, giving amortized constant-time performance.

Option A (Hadoop dictionaries): Not standard terminology.

Option C (HDFS dictionaries): HDFS doesn't use dictionary structures in this sense.

Option D (MapReduce dictionaries): MapReduce uses key-value pairs, but amortized dictionary performance is not its focus.

Thus, the correct answer is Option B (Python dictionaries).

Reference:

DASCA Data Scientist Knowledge Framework (DSKF) - Programming for Data Science: Hash Tables & Amortized Analysis.

## NEW QUESTION # 28

Which of the following is TRUE about Avro?

- **A. Both A and B**
- B. Avro is based on Remote Procedure Call (RPC)
- C. Avro is a data serialization framework
- D. None of the above

**Answer: A**

Explanation:

Apache Avro is a widely used framework within the Hadoop ecosystem for data serialization and data exchange.

Option A (Correct): Avro is a compact, fast, binary data serialization format. It allows efficient storage and exchange of structured data.

Option B (Correct): Avro supports Remote Procedure Call (RPC). It provides a framework for RPC communication, making it easier for distributed applications to exchange data across systems.

Option C: Correct, since both statements are true.

Option D: Incorrect because Avro is indeed both a serialization framework and RPC-based.

In data engineering workflows, Avro is valuable because it is schema-based (defined using JSON), highly interoperable, and ensures compatibility across different programming languages. This makes it essential in big data pipelines, Kafka messaging, and Hadoop ecosystem tools.

Thus, the correct answer is Option C (Both A and B).

Reference:

DASCA Data Scientist Knowledge Framework (DSKF) - Big Data Ecosystem Tools & Data Serialization Techniques.

## NEW QUESTION # 29

Which of the following is correct for Markov chain?

- A. A Markov chain is a sequence of fixed variables  $X_1, X_2$
- B. A Markov chain is the state of a system at sequential points in time
- C. A Markov chain is a sequence of random variables  $X_1, X_2$
- D. Both A and B
- **E. Both B and C**

**Answer: E**

Explanation:

A Markov chain is a stochastic process describing a sequence of possible events, where the probability of each event depends only on the state attained in the previous step (the Markov property).

Option A: Incorrect. The variables are random, not fixed.

Option B: Correct. Markov chains represent the state of a system at sequential time points.

Option C: Correct. A Markov chain is indeed a sequence of random variables  $\{X_1, X_2, \dots\}$  that satisfy the Markov property.

Option D: Incorrect, since A is wrong.

Option E: Correct, because both B and C are valid.

Thus, the correct answer is Option E (Both B and C).

Reference:

- [illegible]