BDS-C00^{Q&As}

AWS Certified Big Data - Speciality (BDS-C00)

Pass Amazon BDS-C00 Exam with 100% Guarantee

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

https://www.leads4pass.com/aws-certified-big-data-specialty.html

100% Passing Guarantee
100% Money Back Assurance

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

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

QUESTION 1

A new algorithm has been written in Python to identify SPAM e-mails. The algorithm analyzes the free text

contained within a sample set of 1 million e-mails stored on Amazon S3. The algorithm must be scaled

across a production dataset of 5 PB, which also resides in Amazon S3 storage.

Which AWS service strategy is best for this use case?

A. Copy the data into Amazon ElastiCache to perform text analysis on the in-memory data and export the results of the model into Amazon Machine Learning.

B. Use Amazon EMR to parallelize the text analysis tasks across the cluster using a streaming program step.

C. Use Amazon Elasticsearch Service to store the text and then use the Python Elasticsearch Client to run analysis against the text index.

D. Initiate a Python job from AWS Data Pipeline to run directly against the Amazon S3 text files.

Correct Answer: C

Reference: https://aws.amazon.com/blogs/database/indexing-metadata-in-amazon-elasticsearch-serviceusing-aws-lambda-and-python/

QUESTION 2

An administrator needs to design a distribution strategy for a star schema in a Redshift cluster. The administrator needs to determine the optimal distribution style for the tables in the Redshift schema.

In which three circumstances would choosing Key-based distribution be most appropriate? (Select three.)

A. When the administrator needs to optimize a large, slowly changing dimension table.

B. When the administrator needs to reduce cross-node traffic.

C. When the administrator needs to optimize the fact table for parity with the number of slices.

D. When the administrator needs to balance data distribution and collocation data.

E. When the administrator needs to take advantage of data locality on a local node for joins and aggregates.

Correct Answer: ACD

QUESTION 3

The project you are working on currently uses a single AWS CloudFormation template to deploy its AWS infrastructure, which supports a multi-tier web application. You have been tasked with organizing the AWS CloudFormation resources so that they can be maintained in the future, and so that different departments such as Networking and Security can review the architecture before it goes to Production.

How should you do this in a way that accommodates each department, using their existing workflows?

A. Organize the AWS CloudFormation template so that related resources are next to each other in the template, such as VPC subnets and routing rules for Networking and Security groups and IAM information for Security

B. Separate the AWS CloudFormation template into a nested structure that has individual templates for the resources that are to be governed by different departments, and use the outputs from the networking and security stacks for the application template that you control

C. Organize the AWS CloudFormation template so that related resources are next to each other in the template for each department\\'s use, leverage your existing continuous integration tool to constantly deploy changes from all parties to the Production environment, and then run tests for validation

D. Use a custom application and the AWS SDK to replicate the resources defined in the current AWS CloudFormation template, and use the existing code review system to allow other departments to approve changes before altering the application for future deployments

Correct Answer: B

QUESTION 4

An organization is soliciting public feedback through a web portal that has been deployed to track the number of requests and other important data. As part of reporting and visualization, AmazonQuickSight connects to an Amazon RDS database to virtualize data. Management wants to understand some important metrics about feedback and how the feedback has changed over the last four weeks in a visual representation.

What would be the MOST effective way to represent multiple iterations of an analysis in Amazon QuickSight that would show how the data has changed over the last four weeks?

A. Use the analysis option for data captured in each week and view the data by a date range.

B. Use a pivot table as a visual option to display measured values and weekly aggregate data as a row dimension.

C. Use a dashboard option to create an analysis of the data for each week and apply filters to visualize the data change.

D. Use a story option to preserve multiple iterations of an analysis and play the iterations sequentially.

Correct Answer: D

QUESTION 5

What is the name of licensing model in which I can use your existing Oracle Database licenses to run Oracle deployments on Amazon RDS?

A. Bring Your Own License

B. Role Bases License

C. Enterprise License

D. License Included

Correct Answer: A