

Exam **DAO-001**

Title **CompTIA Data+ Exam**

Version **3.0**

**Product
Type** **90 Q&A with explanations**

QUESTION 1

Your CEO wants to be able to explore how retail sales have been changing over time. What is the best way to present the information to the CEO? (Choose the best answer.)

- A. One-time report
- B. Recurring report
- C. Interactive dashboard
- D. Historical report

Answer: C

Explanation

D is the correct answer. Since your CEO wants to explore retail sales data, that rules out any type of static report.

QUESTION 2

You work for an organization that uses the Cognos analytics suite. What component of the suite could you use to best create scorecards for business leaders?

- A. Metric Studio
- B. Query Studio
- C. Report Studio
- D. Analysis Studio

Answer: A

Explanation

D is the correct answer. It is feasible that you might be able to create a scorecard using any of these tools. However, Metric Studio is the best tool for the job because it is designed to create scorecards for business leaders seeking to analyze key metrics from across their organizations.

QUESTION 3

What is the mean of the following numbers: 1, 1, 2, 3, 3, 4, 5, 6, 7, 8

- A. 1
- B. 4
- C. 3.5
- D. 3

Answer: B

Explanation

D is the correct answer. To get the mean, you add all the numbers of the dataset together and divide by the number of observations:

1 is the minimum, 1 and 3 are the mode, and 3.5 is the median.

QUESTION 4

Your colleague is looking at a dataset about movies. One of the attributes is the genre a movie belongs to. For a specific science fiction movie, the genre is listed as "æstar." What is this an example of?

- A. Redundant data
- B. Duplicate data
- C. Missing data
- D. Invalid data

Answer: D

Explanation

C is the correct answer. Since the data is for a single movie and comes from a single dataset, it is neither redundant nor duplicate. Since the data exists, it is not missing. While star is a valid word, it isn't a valid movie genre; as such, this is an example of invalid data.

QUESTION 5

To glean insights from the data, many analysts and data scientists rely on _____.

- A. Data mining
- B. All the mentioned above
- C. Data warehouse
- D. Data visualization

Answer: D

Explanation

B is the correct answer. Many analysts and data scientists use data visualization, or the graphical depiction of data, to assist individuals visually explore and find patterns and outliers in the data in order to get insights. Data visualization features are included in a good data analytics system, making data exploration easier and faster.

QUESTION 6

Your friend's company just acquired a competitor, and he is responsible for integrating the Human Resources systems. One of the things he needs to account for are employees who work at his current company who used to work at the acquired company, and vice versa. What is this an example of?

- A. Duplicate data
- B. Redundant data
- C. Data type validation
- D. Data outlier

Answer: A

Explanation

C is the correct answer. The choice here is between redundant and duplicate data. In this case, your friend is sourcing data from multiple systems, and redundant data comes from a single system, so duplicate is the correct answer

QUESTION 7

You work for an organization that relies heavily on the machine learning and analytics tools offered by Amazon Web Services. You would like to select a dash boarding tool that will easily integrate with those other tools. What product would best meet your needs?

- A. QuickSight
- B. Power BI

- C. Cognos
- D. MicroStrategy

Answer: B

Explanation

A is the correct answer. It is likely that any of these tools could meet your needs. However, the key requirement mentioned here is the tight integration with AWS services. AWS QuickSight is an AWS service itself, so it would be the best way to meet your business requirements.

QUESTION 8

Your friend is debating between a snowflake and a star schema design for a data warehouse. Which of the following is not a factor in his selection?

- A. Number of records in the fact table
- B. Degree of normalization
- C. Query complexity
- D. Storage space

Answer: A

Explanation

C is the correct answer. A snowflake approach is more normalized, has more complicated queries, and requires less space than a star approach. The number of records in a fact table has no bearing on whether to use a snowflake or star approach.

QUESTION 9

Logistic regression is used to find the probability of event = Success and event = _____.

- A. None of the mentioned above
- B. Success
- C. Failure
- D. Both A and B

Answer: C

Explanation

A is the correct answer. The likelihood of event=Success and event=Failure is calculated using logistic regression. When the dependent variable is in nature, we should utilize logistic regression. For classification difficulties, logistic regression is commonly employed. There is no requirement for a linear relationship between the dependent and independent variables in logistic regression. Because it uses a non-linear log transformation on the anticipated odds ratio, it can handle a wide range of relationships.

QUESTION 10

Which of the following is the correct advantage of the Update-Driven Approach?

- A. The data can be copied, processed, integrated, annotated, summarized and restructured in the semantic data store in advance.
- B. This approach provides high performance.
- C. None of the above

D. Both A and B

Answer: D

Explanation

C is the correct answer. The statements given in both A and B are the advantage of the Update-Driven Approach in Data Warehousing.

QUESTION 11

Your friend has a new job with the Coca-Cola Company. As part of his role, he has access to a list of logistics suppliers and pricing information. What is the most appropriate classification for this data?

- A. Highly Sensitive
- B. Public
- C. Sensitive
- D. Internal

Answer: D

Explanation

B is the correct answer. Internal applies to data for use within an organization. Public applies to data intended for public consumption, Sensitive applies to data that is for limited use within an organization, and Highly Sensitive data is reserved for restricted use.

QUESTION 12

Clustering belongs to _____ data analysis.

- A. None of the mentioned above
- B. Both A and B
- C. Supervised
- D. Unsupervised

Answer: D

Explanation

B is the correct answer. Unsupervised data analysis includes clustering. Without any prior knowledge, the data's hidden structure is discovered and emphasized. Popular clustering techniques include K-means, K-nearest neighbors, and hierarchical clustering.

QUESTION 13

Your colleague is creating an inventory report for the manager of a local convenience store. On average, replacement items are delivered within 3 days after the manager places an order. What is the most appropriate frequency for this report? (Choose the best answer.)

- A. Daily
- B. Weekly
- C. Hourly
- D. Real-time

Answer: A