Vendor: CompTIA

Exam Code: DA0-002

Exam Name: CompTIA Data+ Certification Exam

Version: DEMO

QUESTION 1
A data analyst created a dashboard to illustrate the traffic volume and mean response time for a call center. The traffic data is current, but the mean response time has not updated for more than an hour. Which of the following is the best way to verify the data's freshness?

A. Refactoring the code base
B. Testing for network connectivity issues
C. Checking the last time the calculation script ran
D. Determining the number of calls with no timestamps

Answer: C
Explanation:
Checking when the calculation script last ran verifies whether the metric is being refreshed properly, which directly confirms the data's freshness.

QUESTION 2
Which of the following pieces of information, if made public, results in a data privacy violation?

A. Gender
B. Driver's license
C. Age
D. Employment status

Answer: B
Explanation:
A driver's license number is personally identifiable information that uniquely identifies an individual and can be used for identity theft or fraud, making its public disclosure a data privacy violation.

QUESTION 3
A data analyst deployed a report for public access. A user states that the report is not showing the latest information, even though the user updated the source an hour ago. Which of the following should the data analyst check first?

A. Event log
B. User privileges
C. Database connection
D. Report corruption

Answer: C
Explanation:
The first thing to verify is the database connection between the report and the data source. If the connection is not refreshing correctly or is pointing to outdated data, the report will not display recent updates even though the source data has changed.

QUESTION 4
A data analyst team needs to segment customers based on customer spending behavior. Given one million rows of data like the information in the following sales order table:

Customer_ID	Region	Amount Spent	Product Category	Quantity of Items
00123	East	20,000	Baby	4
00124	West	30,000	Home	6
00125	South	40,000	Garden	7
00126	North	50,000	Furniture	8
00127	East	60,000	Baby	10

Which of the following techniques should the team use for this task?

A. Standardization
B. Concatenate
C. Binning
D. Appending

Answer: C
Explanation:
Binning groups continuous numerical values such as customer spending into defined ranges, enabling segmentation of customers based on spending behavior.

QUESTION 5
A data analyst creates a report that identifies the middle 50% of the collected data. Which of the following best describes the analyst's findings?

A. Interquartile range
B. The difference between mode and median
C. Mean variance
D. Skewness from the slope

Answer: A
Explanation:
The interquartile range (IQR) represents the spread of the middle 50% of a dataset. It is calculated as the difference between the third quartile (75th percentile) and the first quartile (25th percentile), making it a useful measure of variability that is less affected by outliers.

QUESTION 6
Which of the following best describes the reason an analyst would reference a data dictionary versus a source's metadata?

A. To gather information and resources about the data
B. To find the content and specific attributes for a dataset
C. To find a summary of basic information about the dataset
D. To gather information about the availability of the data

Answer: B
Explanation:

A data dictionary provides detailed definitions of fields, including their meaning, structure, and attributes, which helps an analyst understand the specific contents of a data set.

QUESTION 7
A data analyst is joining two tables with different content and one common field. Which of the following should the analyst do to most efficiently meet this requirement?

A. Match the records of the related columns and merge the tables.
B. Create a cluster to facilitate data integration between the tables.
C. Explode both tables to identify unique values and reorder the fields in one table.
D. Append the values of the matching columns and concatenate the other data fields.

Answer: A
Explanation:
When two tables contain different data but share a common field, the most efficient approach is to match the records using the common field and merge the tables through a join operation. This combines related information from both tables while preserving the relationships between the records.

QUESTION 8
A data analyst pulls a table similar to the following one:

ID	Type	TypeID	Phone
1	Full Time	Full Time 1	Mobile
2	Part Time	Part Time 2	Work
3	Full Time	Full Time 3	Mobile

Which of the following best explains the data issue with TypeID?

A. Redundancy
B. Outlier
C. Missing data
D. Duplication

Answer: A
Explanation:
The TypeID column contains information that is already present in the Type column. Values such as "Full Time 1", "Part Time 2", and "Full Time 3" repeat the employment type while adding an identifier, resulting in unnecessary duplication of information. This is an example of redundancy because the same data is being stored in multiple places.

QUESTION 9
Which of the following AI types is the best option for time-series forecasting?

A. Generative AI

B. Foundational models
C. Natural language processing
D. Robotic process automation

Answer: B
Explanation:
Foundation models are the best option because they can be adapted for a wide range of predictive tasks, including time-series forecasting. These models are trained on large datasets and can identify patterns over time to generate forecasts for future values.

QUESTION 10
A data analyst is following up on a recent, company-wide data audit of customer invoice data. Which of the following is the best option for the analyst to use?

A. PCI DSS
B. GDPR
C. ISO
D. PII

Answer: B

QUESTION 11
A database administrator needs to implement security triggers for an organization's user information database. Which of the following data classifications is the administrator most likely using? (Select two).

A. Public
B. Open
C. Sensitive
D. Non-Sensitive
E. Private
F. Encrypted

Answer: CE

QUESTION 12
A data analyst is analyzing the following dataset:

Transaction Date	Quantity	Item	Item Price
12/12/12	11	USB Cords	$9.99
11/11/11	3	Charging Block	$8.89
10/10/10	5	Headphones	$50.15

Which of the following methods should the analyst use to determine the total cost for each transaction?

A. Parsing
B. Scaling
C. Compressing
D. Deriving

Answer: D
Explanation:
Deriving creates a new value from existing fields, such as calculating total cost by multiplying quantity by item price for each transaction.

QUESTION 13
A business intelligence analyst is creating an employee retention dashboard that looks at data from the last five years. The analyst is interested in identifying patterns that can be studied further. Which of the following is the best method to apply to the dashboard?

A. Predictive
B. Prescriptive
C. Diagnostic
D. Descriptive

Answer: D
Explanation:
Descriptive analytics summarizes historical data to identify trends, patterns, and relationships. Since the analyst is examining employee retention data from the last five years to discover patterns for further study, descriptive analytics is the most appropriate method.

QUESTION 14
Which of the following elements is the most important to include in a dashboard for internal technical audiences?

A. Methodology section
B. Dynamic features
C. Key performance indicators
D. Company branding

Answer: C

QUESTION 15
A data analyst creates a report, and some of the fields are empty. Which of the following conditions should the analyst add to a query to provide a list of all the records with empty fields?

A. WHERE
[ColumnName] = NULL
B. WHERE
[ColumnName] IS NULL
C. WHERE
[ColumnName] IS NOT NULL
D. WHERE
[ColumnName] = 'NULL'

Answer: B
Explanation:
In SQL, NULL represents a missing or unknown value and cannot be compared using the equals operator. The IS NULL condition is specifically designed to identify records where a column contains a NULL value, allowing the analyst to retrieve all records with empty fields.

QUESTION 16
A data analyst needs to get an accurate idea of how data components are automated. Which of the following types of documentation should the analyst review first?

A. Data flow diagram
B. Data explainability report
C. Data dictionary
D. Data lineage

Answer: A

QUESTION 17
A data analyst troubleshoots a dashboard every day for a week. Which of the following techniques best addresses how to validate the data moving forward?

A. Inquiring about structure changes
B. Setting up monitoring alerts
C. Reaching out to users daily
D. Rebuilding the dashboard

Answer: B
Explanation:
Setting up monitoring alerts enables continuous validation by automatically detecting anomalies or issues in the data, reducing the need for repeated manual troubleshooting.

QUESTION 18
A data analyst needs to join together a table data source and a web API data source using Python. Which of the following is the best way to accomplish this task?

A. Convert the data from the API and database to a varchar format and convert them to pandas DataFrames that are then merged together.
B. Convert the data from the API and database to a JSON format and convert them to pandas DataFrames that are then merged together.
C. Convert the data from the API and database to a TXT format and convert them to pandas DataFrames that are then merged together.
D. Convert the data from the API and database to a string format and convert them to pandas DataFrames that are then merged together.

Answer: B
Explanation:
API data is commonly returned in JSON format, which can be easily converted into pandas DataFrames and merged with database data, making it the most appropriate and efficient approach.

Thank You for Trying Our Product

Lead2pass Certification Exam Features:

★ More than **99,900** Satisfied Customers Worldwide.

★ Average **99.9%** Success Rate.

★ **Free Update** to match latest and real exam scenarios.

★ **Instant Download** Access! No Setup required.

★ Questions & Answers are downloadable in **PDF** format and **VCE** test engine format.

★ Multi-Platform capabilities - **Windows, Laptop, Mac, Android, iPhone, iPod, iPad**.

★ **100%** Guaranteed Success or **100%** Money Back Guarantee.

★ **Fast**, helpful support **24x7**.

View list of all certification exams: http://www.lead2pass.com/all-products.html

10% Discount Coupon Code: **ASTR14**