



CompTIA

Exam Questions DA0-002

CompTIA Data+ Exam (2025)

About ExamBible

Your Partner of IT Exam

Found in 1998

ExamBible is a company specialized on providing high quality IT exam practice study materials, especially Cisco CCNA, CCDA, CCNP, CCIE, Checkpoint CCSE, CompTIA A+, Network+ certification practice exams and so on. We guarantee that the candidates will not only pass any IT exam at the first attempt but also get profound understanding about the certificates they have got. There are so many alike companies in this industry, however, ExamBible has its unique advantages that other companies could not achieve.

Our Advances

* 99.9% Uptime

All examinations will be up to date.

* 24/7 Quality Support

We will provide service round the clock.

* 100% Pass Rate

Our guarantee that you will pass the exam.

* Unique Gurantee

If you do not pass the exam at the first time, we will not only arrange FULL REFUND for you, but also provide you another exam of your claim, ABSOLUTELY FREE!

NEW QUESTION 1

Given the following table:

ID	Value
1	1.5
2	24.456
3	113

Which of the following data types should an analyst use for the numeric values in the Value column?

- A. Double
- B. Float
- C. Boolean
- D. Integer

Answer: B

Explanation:

This question falls under the Data Concepts and Environments domain of CompTIA Data+ DA0-002, focusing on selecting appropriate data types for a given dataset. The Value column contains decimal numbers (1.5, 24.456, 113), requiring a data type that supports such values.

? Double (Option A): Double is a floating-point data type that supports decimals with higher precision than Float, but it's often overkill for typical datasets unless very high precision is needed, which isn't indicated here.

? Float (Option B): Float is a floating-point data type that supports decimal numbers (e.g., 1.5, 24.456) and is commonly used for such values in databases, making it the best choice.

? Boolean (Option C): Boolean is for true/false values, not numeric data.

? Integer (Option D): Integer is for whole numbers, but the values (e.g., 1.5, 24.456) have decimals, so Integer is not suitable.

The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," such as selecting data types like Float for decimal numeric values.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW 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:

This question falls under the Data Governancedomain, which in DA0-002 includes understanding data privacy and compliance with regulations like GDPR. The question asks which piece of information, if made public, constitutes a privacy violation, meaning it must be personally identifiable information (PII).

? Gender (Option A): Gender is not typically considered PII on its own, as it's not uniquely identifiable.

? Driver's license (Option B): A driver's license number is PII because it uniquely identifies an individual and can be linked to other personal information, such as name and address. Making it public violates privacy regulations.

? Age (Option C): Age alone isn't PII, as it's not uniquely identifiable.

? Employment status (Option D): Employment status (e.g., employed, unemployed) isn't PII, as it doesn't uniquely identify an individual.

The DA0-002 Data Governance domain includes "identifying PII and data privacy concepts," and a driver's license is a clear example of PII that, if exposed, results in a privacy violation.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

=====

NEW QUESTION 3

The human resources department wants to know the number of employees who earn \$125,000 or more. However, the department is concerned about duplicates in the dataset. Given the following table:

Employee_ID Level

Salary

001

1

10000

002

2

20000

003

2

256000

004

2

125000

001

1

10000

002

2

20000

Which of the following SQL statements resolves this issue?

- A. SELECT DISTINCT Employee_ID FROM Employee WHERE Salary >= 125000
- B. SELECT COUNT(DISTINCT Employee_ID) FROM Employee WHERE Salary >= 125000
- C. SELECT DISTINCT Employee_ID FROM Employee WHERE Salary > 125000
- D. SELECT COUNT(Employee_ID) FROM Employee WHERE Salary >= 125000

Answer: B

Explanation:

This question falls under the Data Analysis domain, focusing on SQL queries to handle duplicates while counting employees. The task is to count unique employees with a salary of \$125,000 or more, addressing duplicates in the dataset.

? Option A: SELECT DISTINCT Employee_ID FROM Employee WHERE Salary >= 125000 This lists unique Employee_IDs but doesn't provide a count, which the department needs.

? Option B: SELECT COUNT(DISTINCT Employee_ID) FROM Employee WHERE

Salary >= 125000 This counts unique Employee_IDs (using DISTINCT) with a salary of \$125,000 or more, correctly addressing duplicates and providing the required count (2 employees: 003 and 004).

? Option C: SELECT DISTINCT Employee_ID FROM Employee WHERE Salary >

125000 This lists unique Employee_IDs with a salary strictly greater than \$125,000 (missing 004), and doesn't provide a count.

? Option D: SELECT COUNT(Employee_ID) FROM Employee WHERE Salary >=

125000 This counts all rows without addressing duplicates, resulting in an incorrect count (2 rows, but only 2 unique employees).

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods using SQL queries," and COUNT(DISTINCT) is the correct method to count unique employees while handling duplicates.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.

=====

NEW QUESTION 4

An administrator needs to design a table that will include foreign words. Which of the following is the best option for storing non-native language characters?

- A. VARCHAR
- B. NVARCHAR
- C. CLOB
- D. CHAR

Answer: B

Explanation:

This question falls under the Data Concepts and Environments domain, focusing on selecting appropriate data types for storing specific kinds of data. The task requires storing foreign words, which often include non-native characters (e.g., accents, Unicode characters).

? VARCHAR (Option A): VARCHAR stores variable-length strings but typically uses

ASCII or single-byte encoding, which may not support all foreign characters.

? NVARCHAR (Option B): NVARCHAR (National VARCHAR) stores variable-length strings in Unicode, supporting a wide range of non-native characters, making it the best choice.

? CLOB (Option C): CLOB (Character Large Object) is for large text data, but it's overkill for most foreign words and not specifically designed for Unicode.

? CHAR (Option D): CHAR stores fixed-length strings, but like VARCHAR, it often uses single-byte encoding, which may not support foreign characters.

The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," such as selecting data types like NVARCHAR for Unicode support.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW QUESTION 5

Which of the following explains the purpose of UAT?

- A. To begin the software application development process to enhance user experience
- B. To ensure all parts of the software application work together after each sprint
- C. To review software application crashes, create patches, and deploy to users
- D. To validate and verify that a software application meets the needs and requirements of users

Answer: D

Explanation:

This question is related to the Data Governance domain of DA0-002, which includes understanding processes like User Acceptance Testing (UAT) to ensure data-related applications meet governance and quality standards. UAT is a critical step in ensuring software aligns with user needs and organizational requirements.

? To begin the software application development process to enhance user experience (Option A): UAT occurs near the end of development, not at the beginning.

? To ensure all parts of the software application work together after each sprint

(Option B): This describes integration testing, not UAT, which focuses on user validation.

? To review software application crashes, create patches, and deploy to users (Option C): This refers to post-deployment maintenance, not UAT.

? To validate and verify that a software application meets the needs and

requirements of users (Option D): UAT is specifically designed to ensure the software meets user requirements and functions as intended in a real-world scenario, aligning with governance standards for quality.

The DA0-002 Data Governance domain emphasizes "data quality control concepts" (similar to DA0-001, web ID: 1), which include ensuring applications meet user needs through processes like UAT.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance

=====

NEW QUESTION 6

Which of the following data repositories stores unformatted data in its original, raw form?

- A. Data warehouse
- B. Data silo
- C. Data mart
- D. Data lake

Answer: D

Explanation:

This question pertains to the Data Concepts and Environments domain, focusing on data repositories. The task is to identify a repository that stores raw, unformatted data.

? Data warehouse (Option A): A data warehouse stores structured, processed data in a predefined schema, not raw data.

? Data silo (Option B): A data silo is an isolated repository, often structured, not designed for raw data storage.

? Data mart (Option C): A data mart is a subset of a data warehouse, also storing structured data.

? Data lake (Option D): A data lake stores raw, unformatted data in its original format (structured, semi-structured, or unstructured), making it the correct choice.

The DA0-002 Data Concepts and Environments domain includes understanding "different types of databases and data repositories," and a data lake is designed for raw data storage.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW QUESTION 7

An analyst needs to create a collection of dashboards for multiple teams within their organization. Which of the following should the analyst do first before starting the project?

- A. Evaluate the user persona type for the dashboards.
- B. Determine the number of team members who need to access the dashboards.
- C. Determine the delivery method of the dashboards.
- D. Evaluate the KPIs for the dashboards.

Answer: A

Explanation:

This question pertains to the Visualization and Reporting domain, focusing on the initial steps in dashboard creation. The analyst is starting a project for multiple teams, so understanding the audience is critical.

? Evaluate the user persona type for the dashboards (Option A): User personas define the needs, roles, and preferences of the audience (e.g., executives vs. analysts), which is the first step to ensure the dashboards meet user requirements.

? Determine the number of team members who need to access the dashboards

(Option B): This is important for access control but not the first step in design.

? Determine the delivery method of the dashboards (Option C): Delivery method (e.g., self-service) is decided later, after understanding user needs.

? Evaluate the KPIs for the dashboards (Option D): KPIs are defined after understanding the user personas and their goals.

The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and evaluating user personas is the first step in this process.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.

=====

NEW QUESTION 8

A product goes viral on social media, creating high demand. Distribution channels are facing supply chain issues because the testing and training models that are used for sales forecasting have not encountered similar demand. Which of the following best describes this situation?

- A. Model bias
- B. Data drift
- C. Incorrect sizing
- D. Skewing

Answer: B

Explanation:

This question pertains to the Data Analysis domain, focusing on issues with forecasting models. The scenario describes a sudden change in demand (viral product) that the model couldn't predict because it hasn't seen similar patterns before.

? Model bias (Option A): Model bias occurs when a model systematically favors

certain outcomes due to flawed training data, but this scenario is about a change in data patterns, not bias.

? Data drift (Option B): Data drift occurs when the statistical properties of the data

change over time (e.g., sudden high demand due to virality), causing the model to perform poorly because it was trained on different patterns, which fits the scenario.

? Incorrect sizing (Option C): This term is vague and not a standard concept in data analysis for this context.

? Skewing (Option D): Skewing refers to data distribution asymmetry, not a change in data patterns affecting model performance.

The DA0-002 Data Analysis domain includes understanding "applying the appropriate descriptive statistical methods," and data drift is a key concept in forecasting when data patterns change unexpectedly.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.

NEW QUESTION 9

A data company needs a visualization that shows the availability zones from the last ten years and any future availability zones that the company will be using in the next five years. Which of the following is the most appropriate visualization to display this information?

- A. Bar chart
- B. Mosaic plot
- C. Map
- D. Pie chart

Answer: C

Explanation:

This question falls under the Visualization and Reporting domain of CompTIA Data+ DA0-002, focusing on selecting the appropriate visualization for a specific dataset. The task is to show availability zones over a 15-year period (past and future), which involves a geographical element since availability zones are typically location-based.

? Bar chart (Option A): Bar charts are good for comparing categorical data but don't effectively show geographical locations or time-based trends across zones.

? Mosaic plot (Option B): Mosaic plots display relationships between categorical variables, not suitable for geographical or time-series data.

? Map (Option C): A map can display availability zones geographically, with annotations or layers to show changes over time (past 10 years and future 5 years), making it the most appropriate visualization.

? Pie chart (Option D): Pie charts show proportions of a whole, not suitable for geographical or time-based data.

The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and a map is the best choice for displaying geographical availability zones over time.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.

=====

NEW QUESTION 10

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

Explanation:

This question pertains to the Visualization and Reporting domain, focusing on dashboard design for specific audiences. Internal technical audiences (e.g., data analysts, IT staff) need actionable, data-driven insights.

? Methodology section (Option A): Methodology is important for research reports, not dashboards, especially for technical audiences who prioritize data.

? Dynamic features (Option B): Dynamic features (e.g., interactivity) are useful but not the most critical element for technical audiences.

? Key performance indicators (Option C): KPIs provide critical metrics (e.g., system uptime, error rates) that technical audiences need to monitor and act on, making this the most important element.

? Company branding (Option D): Branding is more relevant for external audiences, not internal technical ones.

The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and KPIs are essential for technical dashboards.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.

=====

NEW QUESTION 10

SIMULATION

The director of operations at a power company needs data to help identify where company resources should be allocated in order to monitor activity for outages and restoration of power in the entire state. Specifically, the director wants to see the following:

- * County outages
- * Status
- * Overall trend of outages

INSTRUCTIONS:

Please, select each visualization to fit the appropriate space on the dashboard and choose an appropriate color scheme. Once you have selected all visualizations, please, select the appropriate titles and labels, if applicable. Titles and labels may be used more than once.

If at any time you would like to bring back the initial state of the simulation, please click the Reset All button.

1000000

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

This is a simulation question that requires you to create a dashboard with visualizations that meet the director's needs. Here are the steps to complete the task:

• Drag and drop the visualization that shows the county outages on the top left

space of the dashboard. This visualization is a map of the state with different colors indicating the number of outages in each county. You can choose any color scheme that suits your preference, but make sure that the colors are consistent and clear. For example, you can use a gradient of red to show the counties with more outages and green to show the counties with less outages.

• Drag and drop the visualization that shows the status of the outages on the top

right space of the dashboard. This visualization is a pie chart that shows the percentage of outages that are active, restored, or pending. You can choose any color scheme that suits your preference, but make sure that the colors are distinct and easy to identify. For example, you can use red for active, green for restored, and yellow for pending.

• Drag and drop the visualization that shows the overall trend of outages on the

bottom space of the dashboard. This visualization is a line graph that shows the number of outages over time. You can choose any color scheme that suits your preference, but make sure that the color is visible and contrasted with the background. For example, you can use blue for the line and white for the background.

• Select appropriate titles and labels for each visualization. Titles and labels may be

used more than once. For example, you can use "County Outages" as the title for the map, "Status" as the title for the pie chart, and "Trend" as the title for the line graph. You can also use "County", "Number of Outages", "Active", "Restored", "Pending", "Time", and "Number of Outages" as labels for the axes and legends of the visualizations.

NEW QUESTION 13

The human resources department wants to understand the relationship between the ages and incomes of all employees. Which of the following graphics is the most appropriate to present the analysis?

- A. Scatter plot
- B. Area plot
- C. Bar chart
- D. Pie chart

Answer: A

Explanation:

This question pertains to the Visualization and Reporting domain, focusing on selecting the appropriate visualization to show a relationship between two continuous variables (ages and incomes).

• Scatter plot (Option A): A scatter plot displays individual data points on two axes

(age vs. income), making it ideal for showing the relationship and potential correlation between two continuous variables.

• Area plot (Option B): Area plots are used for showing trends over time, not relationships between two variables.

• Bar chart (Option C): Bar charts are better for categorical data comparisons, not continuous variable relationships.

• Pie chart (Option D): Pie charts show proportions of a whole, not suitable for showing relationships between variables.

The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and a scatter plot is best for showing the relationship between age and income.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.

=====

NEW QUESTION 16

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:

This question pertains to the Data Analysis domain, focusing on statistical measures. The middle 50% of a dataset refers to a specific statistical concept related to data distribution.

• Interquartile range (Option A): The interquartile range (IQR) is the range between

the first quartile (Q1, 25th percentile) and the third quartile (Q3, 75th percentile), representing the middle 50% of the data, which matches the description.

• The difference between mode and median (Option B): This measures the spread between two central tendency metrics but doesn't represent the middle 50% of the data.

• Mean variance (Option C): Variance measures data dispersion around the mean, not the middle 50%.

• Skewness from the slope (Option D): Skewness measures data asymmetry, and "slope" is irrelevant here.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods," and the IQR is the standard measure for the middle 50% of a dataset. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.

=====

NEW QUESTION 20

A manager needs a report to be sent by email every Monday for the next six months. Which of the following is the best way to accomplish this task?

- A. Building self-service access
- B. Creating a data snapshot

- C. Developing a recurring process
- D. Waiting for the request each week

Answer: C

Explanation:

This question falls under the Visualization and Reporting domain, focusing on report delivery methods. The task requires a report to be emailed every Monday for six months, indicating a scheduled, repeating process.

? Building self-service access (Option A): Self-service allows users to generate reports on-demand, but the manager wants automatic delivery.

? Creating a data snapshot (Option B): A snapshot captures data at a specific point, not suitable for recurring delivery over six months.

? Developing a recurring process (Option C): A recurring process schedules the report to be generated and emailed every Monday, meeting the requirement for automated delivery over six months.

? Waiting for the request each week (Option D): This is manual and inefficient, not suitable for a six-month schedule.

The DA0-002 Visualization and Reporting domain includes "the appropriate visualization in the form of a report" with delivery methods, and a recurring process is ideal for scheduled email delivery.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.

=====

NEW QUESTION 24

Which of the following tables holds relational keys and numeric values?

- A. Fact
- B. Graph
- C. Dimensional
- D. Transactional

Answer: A

Explanation:

This question falls under the Data Concepts and Environments domain, focusing on understanding table types in data warehousing. The task is to identify a table that holds relational keys and numeric values, typically used in a star schema.

? Fact (Option A): Fact tables in a star schema store quantitative data (numeric values, e.g., sales amounts) and foreign keys (relational keys) linking to dimension tables, making this the correct choice.

? Graph (Option B): Graph tables are used in graph databases for relationships (e.g., nodes, edges), not typically for relational keys and numeric values in a traditional sense.

? Dimensional (Option C): Dimension tables store descriptive attributes (e.g., product names) and primary keys, not typically numeric measures.

? Transactional (Option D): Transactional tables are used in OLTP systems and may contain numeric values, but they're not specifically designed for relational keys in a data warehousing context.

The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," and fact tables are designed to hold relational keys and numeric values in a data warehouse.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW QUESTION 27

A company gives users adequate data access permissions to allow them to fulfill their duties but nothing more. Which of the following concepts best describes this practice?

- A. Active Directory
- B. Hierarchical access
- C. Zero Trust
- D. Least privilege

Answer: D

Explanation:

This question pertains to the Data Governance domain, focusing on data security and access control principles. The company restricts access to the minimum needed for duties, which aligns with a specific security concept.

? Active Directory (Option A): Active Directory is a tool for managing users and permissions, not a concept.

? Hierarchical access (Option B): Hierarchical access implies access based on roles in a hierarchy, but it doesn't specifically focus on minimal access.

? Zero Trust (Option C): Zero Trust requires continuous verification for all access, which is broader than just minimal permissions.

? Least privilege (Option D): Least privilege ensures users have only the permissions necessary for their duties, which matches the scenario.

The DA0-002 Data Governance domain includes "data privacy concepts," and least privilege is a fundamental principle for secure access control.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

=====

NEW QUESTION 32

Due to new reporting requirements, a data analyst must add new classification codes to historical data. Which of the following is the best technique for this task?

- A. Append
- B. Binning
- C. Parsing
- D. Union

Answer: A

Explanation:

This question falls under the Data Acquisition and Preparation domain, focusing on modifying historical data. The task is to add new classification codes to existing data, which involves adding new rows or columns.

? Append (Option A): Appending adds new rows to a dataset, which is suitable if the

classification codes are new records (e.g., a new table of codes to combine with historical data). If the codes are a new column, a join or update might be used, but append fits the context of adding new data.

? Binning (Option B): Binning groups data into categories, not suitable for adding classification codes.

? Parsing (Option C): Parsing breaks down data (e.g., splitting strings), not relevant for adding codes.

? Union (Option D): Union stacks tables with identical structures, but the task involves adding new data (codes) to historical data, not combining identical tables.

The DA0-002 Data Acquisition and Preparation domain includes "executing data manipulation," and appending is a common technique for adding new data to historical datasets.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 2.0 Data Acquisition and Preparation.

=====

NEW QUESTION 34

A company reports on seven years of data in a sales dashboard. The dashboard pulls from a sales database that has 30 years of data. The dashboard performance is slow. Which of the following is the best way to improve the dashboard's performance?

- A. Performing a code review
- B. Checking network connectivity
- C. Filtering to include only relevant data
- D. Adding more RAM and rerunning

Answer: C

Explanation:

This question falls under the Data Governance domain, focusing on optimizing data quality and performance in dashboards. The dashboard is slow because it pulls from a large database (30 years) but only needs seven years of data.

? Performing a code review (Option A): A code review might identify inefficiencies, but it's not the most direct solution for this scenario.

? Checking network connectivity (Option B): Network issues might cause delays, but the primary issue is the data volume, not connectivity.

? Filtering to include only relevant data (Option C): Filtering the data to include only the last seven years reduces the dataset size, directly improving performance by minimizing the data processed.

? Adding more RAM and rerunning (Option D): Adding RAM might help, but it's a hardware solution that doesn't address the root cause of excessive data.

The DA0-002 Data Governance domain includes "data quality control concepts," such as optimizing performance by filtering data to improve efficiency.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

=====

NEW QUESTION 36

A data analyst receives a new data source that contains employee IDs, job titles, dates of birth, addresses, years of service, and employees' birth months. Which of the following inconsistencies should the analyst identify?

- A. Redundancy
- B. Equivalence
- C. Parallel
- D. Duplication

Answer: A

Explanation:

This question falls under the Data Governance domain, focusing on identifying data quality issues. The dataset includes dates of birth and birth months, which suggests a potential inconsistency.

? Redundancy (Option A): The dataset includes both dates of birth (e.g., 1990-05-15) and birth months (e.g., May), which is redundant because the birth month can be derived from the date of birth, indicating a data quality issue.

? Equivalence (Option B): Equivalence isn't a standard data quality term in this context; it might refer to data matching, which isn't the issue here.

? Parallel (Option C): Parallel isn't a recognized data quality term; it might relate to processing, not data inconsistencies.

? Duplication (Option D): Duplication refers to identical records, but the issue here is redundant fields, not duplicate rows.

The DA0-002 Data Governance domain includes "data quality control concepts," and redundancy is a key inconsistency when the same information is stored in multiple forms unnecessarily.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

=====

NEW QUESTION 40

A data analyst needs to remove all duplicate values between two tables, "Employees" and "Managers," using SQL SELECT statements. Which of the following should the analyst use for this task?

- A. SELECT * FROM Employees UNION ALL SELECT * FROM Managers
- B. SELECT * FROM Employees UNION SELECT * FROM Managers
- C. SELECT * FROM Employees JOIN SELECT * FROM Managers
- D. SELECT * FROM Employees CROSS JOIN SELECT * FROM Managers

Answer: B

Explanation:

This question pertains to the Data Acquisition and Preparation domain, focusing on combining and deduplicating data using SQL. The task is to remove duplicates between two tables, meaning the analyst needs a unique set of records from both.

? SELECT * FROM Employees UNION ALL SELECT * FROM Managers (Option A):

UNION ALL combines all rows from both tables, including duplicates, which doesn't meet the requirement.

? SELECT * FROM Employees UNION SELECT * FROM Managers (Option B):

UNION combines rows from both tables and automatically removes duplicates, providing a unique set of records, which fits the task.

? SELECT * FROM Employees JOIN SELECT * FROM Managers (Option C): This

syntax is incorrect; a JOIN requires an ON clause, and it wouldn't remove duplicates.

? SELECT * FROM Employees CROSS JOIN SELECT * FROM Managers (Option

D): A CROSS JOIN creates a Cartesian product, resulting in all possible combinations, not removing duplicates.

The DA0-002 Data Acquisition and Preparation domain includes "executing data manipulation," and UNION is the correct SQL operation for combining tables while removing duplicates.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 2.0 Data Acquisition and Preparation.

=====

NEW QUESTION 43

Software end users are happy with the quality of product support provided. However, they frequently raise concerns about the long wait time for resolutions. An IT manager wants to improve the current support process. Which of the following should the manager use for this review?

- A. Infographic
- B. KPI
- C. Survey
- D. UAT

Answer: C

Explanation:

This question falls under the Data Analysis domain, focusing on methods to gather data for process improvement. The IT manager needs to review user concerns about wait times, which requires collecting feedback.

? Infographic (Option A): An infographic visualizes data but isn't a method for gathering feedback.

? KPI (Option B): KPIs (e.g., average resolution time) measure performance but don't directly gather user feedback.

? Survey (Option C): A survey collects detailed feedback from users about their experiences, such as wait times, making it the best method for this review.

? UAT (Option D): User Acceptance Testing validates software functionality, not support processes.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods," and surveys are a standard method for collecting user feedback to analyze and improve processes.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.

=====

NEW QUESTION 48

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:

This question is part of the Data Concepts and Environments domain, focusing on identifying data quality issues. The table shows Type and TypeID columns, where TypeID seems to repeat information from Type with an additional identifier.

? Redundancy (Option A): The TypeID column (e.g., "Full Time 1") redundantly includes the Type value ("Full Time") with an extra identifier, which is unnecessary and could be simplified by using a numeric ID instead.

? Outlier (Option B): Outliers are data points that deviate significantly, which isn't applicable here.

? Missing data (Option C): There are no missing values in the table.

? Duplication (Option D): Duplication refers to identical rows, but the rows here are unique; the issue is with the column content.

The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," and redundancy is a common data quality issue in schema design.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW QUESTION 52

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:

This question pertains to the Data Governancedomain, focusing on troubleshooting data freshness issues in reports. The report isn't showing the latest data despite a recent source update, indicating a potential refresh or connectivity issue.

? Event log (Option A): Event logs might provide insight into errors, but they're not

the first step for checking data freshness.

? User privileges (Option B): Privileges might affect access, but the user can see the report, so this isn't the issue.

? Database connection (Option C): If the database connection failed or isn't refreshing properly, the report won't reflect the latest data, making this the first thing to check.

? Report corruption (Option D): Corruption might cause errors, but it's less likely than a connectivity issue for this scenario.

The DA0-002 Data Governance domain includes "data quality control concepts," such as ensuring data freshness by verifying database connections.

Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

=====

NEW QUESTION 54

.....

Relate Links

100% Pass Your DA0-002 Exam with Exam Bible Prep Materials

<https://www.exambible.com/DA0-002-exam/>

Contact us

We are proud of our high-quality customer service, which serves you around the clock 24/7.

Viste - <https://www.exambible.com/>