

CompTIA

Exam Questions DA0-002

CompTIA Data+ Exam (2025)



NEW QUESTION 1

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:

This question falls under the Data Acquisition and Preparation domain of CompTIA Data+ DAO-002, which involves acquiring and combining data from different sources, such as a database and a web API, using tools like Python. The task requires joining the data, which in Python often involves using pandas DataFrames. ? Convert the data from the API and database to a varchar format and convert them to pandas DataFrames that are then merged together (Option A): VARCHAR is a database data type for strings, not a format for data exchange or merging in Python, making this incorrect.

? Convert the data from the API and database to a JSON format and convert them to pandas DataFrames that are then merged together (Option B): Web APIs commonly return data in JSON format, and databases can export data as JSON. In Python, JSON data can be easily converted to pandas DataFrames using `pandas.read_json()` or `pandas.DataFrame()`, and then merged using `pandas.merge()` on a common key, making this the best approach.

? Convert the data from the API and database to a TXT format and convert them to pandas DataFrames that are then merged together (Option C): TXT is a generic text format that lacks structure, making it less efficient for merging compared to JSON.

? Convert the data from the API and database to a string format and convert them to pandas DataFrames that are then merged together (Option D): Converting to a string format is vague and not a standard approach for structured data merging in Python.

The DAO-002 Data Acquisition and Preparation domain includes "executing data manipulation," such as combining data from APIs and databases, and JSON is a standard format for this purpose in Python.

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

=====

NEW QUESTION 2

A data analyst receives a flat file that includes dates. The analyst needs to calculate the number of days from the dates on the file to the current date. Which of the following is the best way to complete this task?

- A. Convert data to date format and use date functions.
- B. Validate the date format with logical functions and use date functions to analyze.
- C. Use date functions to analyze the data with no conversion.
- D. Transform data to a numerical value and use mathematical functions.

Answer: A

Explanation:

This question pertains to the Data Analysis domain, focusing on date calculations. The task is to calculate the difference between dates in a file and the current date, requiring proper date handling.

? Convert data to date format and use date functions (Option A): Flat files often store dates as strings (e.g., "2023-01-01"). Converting them to a date format (e.g., using Python's `datetime` or SQL's `TO_DATE`) allows the use of date functions (e.g., `DATEDIFF`) to calculate the difference to the current date, which is the best approach.

? Validate the date format with logical functions and use date functions to analyze (Option B): Validation is unnecessary if conversion handles format issues, making this overly complex.

? Use date functions to analyze the data with no conversion (Option C): Without converting to a date format, date functions may fail if the data is stored as strings.

? Transform data to a numerical value and use mathematical functions (Option D): This is inefficient and error-prone compared to using date functions.

The DAO-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods," and converting to date format followed by date functions is the standard method for such calculations.

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

=====

NEW QUESTION 3

Which of the following is a NoSQL database?

- A. PostgreSQL
- B. MySQL
- C. Oracle
- D. MongoDB

Answer: D

Explanation:

This question falls under the Data Concepts and Environments domain, focusing on types of databases. The task is to identify a NoSQL database among the options.

? PostgreSQL (Option A): PostgreSQL is a relational (SQL) database, not NoSQL.

? MySQL (Option B): MySQL is a relational (SQL) database, not NoSQL.

? Oracle (Option C): Oracle Database is a relational (SQL) database, not NoSQL.

? MongoDB (Option D): MongoDB is a NoSQL database that uses a document-based model, storing data in JSON-like structures, making it the correct choice.

The DAO-002 Data Concepts and Environments domain includes understanding "different types of databases and data repositories," and MongoDB is a well-known NoSQL database.

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

=====

NEW QUESTION 4

Which of the following best enables the retrieval and manipulation of data that is stored in a relational database?

- A. XML
- B. SQL
- C. Excel
- D. JavaScript

Answer: B

Explanation:

This question pertains to the Data Concepts and Environments domain, focusing on tools for interacting with relational databases. The task is to identify the best method for retrieving and manipulating data.

? XML (Option A): XML is a data format, not a language for retrieving or manipulating database data.

? SQL (Option B): SQL (Structured Query Language) is specifically designed for querying and manipulating data in relational databases (e.g., SELECT, UPDATE), making it the best choice.

? Excel (Option C): Excel can analyze data but isn't designed for direct database manipulation.

? JavaScript (Option D): JavaScript is a programming language for web development, not optimized for relational database operations.

The DA0-002 Data Concepts and Environments domain includes understanding "different types of databases," and SQL is the standard language for relational database operations. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

NEW QUESTION 5

Which of the following best describes an assessment a data analyst would use to validate that the number of records in a dataset matches the expected results?

- A. Source control
- B. Unit test
- C. Stress test
- D. Health check

Answer: B

Explanation:

This question pertains to the Data Governance domain, focusing on data quality validation techniques. The task is to validate that the number of records matches expectations, which requires a specific type of assessment.

? Source control (Option A): Source control (e.g., Git) manages code versions, not dataset validation.

? Unit test (Option B): A unit test checks a specific component of a process, such as verifying that the number of records in a dataset matches the expected count, making it the best fit.

? Stress test (Option C): Stress tests evaluate system performance under load, not record counts.

? Health check (Option D): A health check monitors system status but isn't specific to validating record counts.

The DA0-002 Data Governance domain includes "data quality control concepts," and unit tests are a standard method for validating specific data outcomes like record counts. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

NEW QUESTION 6

Which of the following is found in metadata?

- A. Transformations
- B. Data lineage
- C. Syntax
- D. Variable types

Answer: D

Explanation:

This question pertains to the Data Concepts and Environments domain, focusing on the content of metadata. Metadata describes data attributes, and the task is to identify what it typically includes.

? Transformations (Option A): Transformations (e.g., data cleaning steps) are part of data lineage, not metadata.

? Data lineage (Option B): Data lineage tracks data flow and transformations, which is related to metadata but not a direct component.

? Syntax (Option C): Syntax refers to code structure, not a metadata component.

? Variable types (Option D): Metadata includes information about data fields, such as variable types (e.g., integer, string), which is a standard component.

The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," and metadata typically contains details like variable types to describe the dataset.

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

NEW QUESTION 7

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+ DAO-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 DAO-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+ DAO-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW QUESTION 8

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:

This question pertains to the Data Governance domain, which in DAO-002 includes ensuring data quality and freshness, especially in dashboards. The issue is that the mean response time isn't updating, while traffic data is current, indicating a potential issue with the data refresh process for the response time metric.

? Refactoring the code base (Option A): Refactoring might improve long-term performance but doesn't directly address verifying data freshness.

? Testing for network connectivity issues (Option B): Network issues could cause delays, but since traffic data is updating, connectivity is likely not the issue.

? Checking the last time the calculation script ran (Option C): Mean response time is a calculated metric, likely derived from a script. Checking when the script last ran directly verifies if the data refresh process failed, making this the best approach.

? Determining the number of calls with no timestamps (Option D): Missing timestamps might indicate data quality issues, but it doesn't directly verify why the mean response time isn't updating.

The DAO-002 Data Governance domain focuses on "data quality control concepts," including ensuring data freshness in reporting. Checking the script's last run time aligns with this objective.

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

=====

NEW QUESTION 9

Which of the following best describes the function of a data type?

- A. To provide a generic identifier for files used in analysis
- B. To identify the program needed to open a file
- C. To differentiate the real value of the field in its context
- D. To make the addition of individual records simpler

Answer: C

Explanation:

This question falls under the Data Concepts and Environments domain, focusing on the purpose of data types in data management. Data types define how data is stored and interpreted.

? To provide a generic identifier for files used in analysis (Option A): Data types apply to fields within datasets, not files.

? To identify the program needed to open a file (Option B): File extensions (e.g., .csv) identify programs, not data types.

? To differentiate the real value of the field in its context (Option C): Data types (e.g., integer, string, date) define how a field's value is interpreted (e.g., "123" as a number vs. text), ensuring its real meaning in context, making this the correct answer.

? To make the addition of individual records simpler (Option D): Data types don't directly simplify record addition; they ensure proper data handling.

The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," and data types ensure fields are interpreted correctly in their context.

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

=====

NEW QUESTION 10

A manager wants to use the information in a recurring report on incomplete timesheets for the prior month to guide employee outreach. Which of the following report types is the best for this task?

- A. Summary
- B. Infographic
- C. Snapshot
- D. Ad hoc

Answer: A

Explanation:

This question is part of the Visualization and Reporting domain, focusing on selecting the appropriate report type for a specific purpose. The manager needs a recurring report on incomplete timesheets to guide outreach, which requires a concise, data-driven report.

? Summary (Option A): A summary report aggregates data (e.g., total incomplete timesheets per employee) and presents it concisely, making it ideal for recurring use to guide decisions like employee outreach.

? Infographic (Option B): Infographics are visual representations for broad audiences, not typically used for recurring, detailed employee outreach tasks.

? Snapshot (Option C): A snapshot report captures data at a specific point in time, but it's not ideal for recurring analysis of trends or aggregates.

? Ad hoc (Option D): Ad hoc reports are one-time, on-demand reports, not suitable for recurring needs.

The DA0-002 Visualization and Reporting domain includes "the appropriate visualization in the form of a report" , and a summary report best fits the need for recurring, actionable data.

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

=====

NEW QUESTION 10

A data analyst receives four files that need to be unified into a single spreadsheet for further analysis. All of the files have the same structure, number of columns, and field names, but each file contains different values. Which of the following methods will help the analyst convert the files into a single spreadsheet?

- A. Merging
- B. Appending
- C. Parsing
- D. Clustering

Answer: B

Explanation:

This question is part of the Data Acquisition and Preparation domain, which involves combining data from multiple sources. The files have the same structure but different

values, meaning they need to be stacked vertically into one dataset.

? Merging (Option A): Merging typically involves joining datasets on a common key (e.g., a customer ID), which isn't indicated here since the files only differ in values, not keys.

? Appending (Option B): Appending stacks datasets vertically, combining rows from files with the same structure into a single dataset, which matches the scenario.

? Parsing (Option C): Parsing involves breaking down data (e.g., splitting text), not combining files.

? Clustering (Option D): Clustering is a machine learning technique for grouping similar data points, not for combining files.

The DA0-002 Data Acquisition and Preparation domain includes "executing data manipulation," such as appending datasets with identical structures.

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

=====

NEW QUESTION 14

A data analyst needs to modify a dashboard that was created by another employee. Upon opening the dashboard, the analyst notices that the information is not loading properly. Which of the following should the analyst do to troubleshoot this error?

- A. Review the data layer and data source.
- B. Validate that the database is up-to-date.
- C. Check that the program is updated to the latest version.

D. Ensure the correct filters are displaying on the dashboard.

Answer: A

Explanation:

This question falls under the Data Governance domain of CompTIA Data+ DA0-002, focusing on troubleshooting data quality issues in dashboards. The dashboard isn't loading properly, indicating a potential issue with the data connection or configuration.

? Review the data layer and data source (Option A): The data layer (e.g., queries, connections) and data source (e.g., database) are the foundation of a dashboard. If the information isn't loading, the issue likely lies in the data connection or query configuration, making this the first step.

? Validate that the database is up-to-date (Option B): While this might be a subsequent step, it assumes the connection is working, which should be confirmed first.

? Check that the program is updated to the latest version (Option C): Software updates might fix bugs, but this isn't the most immediate cause of data not loading.

? Ensure the correct filters are displaying on the dashboard (Option D): Filters affect data display, but if the data isn't loading at all, the issue is more fundamental.

The DA0-002 Data Governance domain includes "data quality control concepts," and reviewing the data layer and source is the primary step in troubleshooting dashboard loading issues.

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

=====

NEW QUESTION 16

Which of the following supports capabilities such as automatic versioning, corruption checks, KPIs, and user authentication?

- A. Notebook
- B. REST API
- C. Pipeline
- D. Source control

Answer: D

Explanation:

This question falls under the Data Governance domain, focusing on tools that support data management and quality control features. The task is to identify a tool with capabilities like versioning, corruption checks, KPIs, and authentication.

? Notebook (Option A): Notebooks (e.g., Jupyter) are for data analysis and coding but don't inherently support versioning, corruption checks, or authentication.

? REST API (Option B): REST APIs enable data access but don't provide versioning or corruption checks as a primary function.

? Pipeline (Option C): Data pipelines automate data workflows but don't typically include versioning or authentication.

? Source control (Option D): Source control systems (e.g., Git) support automatic versioning (tracking changes), corruption checks (integrity verification), KPIs (e.g., commit frequency), and user authentication (access control), making this the best fit.

The DA0-002 Data Governance domain includes "data quality control concepts," and source control systems provide the listed capabilities to ensure data integrity and security. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

=====

NEW QUESTION 21

A company has a document that includes the names of key metrics and the standard for how those metrics are calculated company-wide. Which of the following describes this documentation?

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

Answer: A

Explanation:

This question falls under the Data Concepts and Environments domain, which involves understanding documentation types related to data management. The document describes key metrics and their calculation standards, which points to a specific type of metadata documentation.

? Data dictionary (Option A): A data dictionary defines data elements, including metrics, their meanings, and calculation methods, ensuring consistency across the organization. This matches the description.

? Data explainability report (Option B): This term is more associated with AI/ML, explaining model decisions, not metric definitions.

? Data lineage (Option C): Data lineage tracks the flow of data through systems, not metric definitions or calculations.

? Data flow diagram (Option D): A data flow diagram visualizes data processes, not metric standards.

The DA0-002 Data Concepts and Environments domain includes understanding "basic concepts of data schemas and dimensions", and a data dictionary is a foundational tool for defining metrics.

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

NEW QUESTION 25

A data analyst calculated the average score per student without making any changes to the following table:

Student Subject Score 123

Math 100

123

Biology 80

234

Math 96

123

Biology 80

345

Biology 88

234

Math 96

Which of the following exploration techniques should the analyst have considered before calculating the average?

- A. Duplication
- B. Redundancy
- C. Binning
- D. Grouping

Answer: A

Explanation:

This question pertains to the Data Governance domain, focusing on data quality issues that affect analysis. The table contains duplicate rows, which would skew the average score calculation if not addressed.

? Student 123: Math (100), Biology (80), Biology (80)– Duplicate Biology score.

? Student 234: Math (96), Math (96)– Duplicate Math score.

? Student 345: Biology (88)– No duplicates.

? Duplication (Option A): The table has duplicate rows (e.g., Student 123's Biology score of 80 appears twice), which would inflate the average if not removed.

The analyst should have checked for duplicates before calculating the average.

? Redundancy (Option B): Redundancy refers to unnecessary fields (e.g., storing the same data in multiple columns), not duplicate rows.

? Binning (Option C): Binning groups data into categories, not relevant for addressing duplicates in averaging.

? Grouping (Option D): Grouping (e.g., GROUP BY in SQL) might be part of the solution, but the issue to identify is duplication.

The DAO-002 Data Governance domain includes "data quality control concepts," and checking for duplication is critical to ensure accurate calculations like averages. Reference: CompTIA Data+ DAO-002 Draft Exam Objectives, Domain 5.0 Data Governance.

=====

NEW QUESTION 29

An analyst is building a reporting deck. The deck must include tracking and visualizing metrics and row-level security. Which of the following actions should the analyst take after meeting all of the requirements?

- A. Show a mock-up to the team.
- B. Explain the desired level of reporting detail.
- C. Present an analysis of the data.
- D. Find out the project due date.

Answer: A

Explanation:

This question pertains to the Visualization and Reporting domain, focusing on the process of building a reporting deck. After meeting requirements (tracking metrics, visualizing data, and row-level security), the next step involves validation with stakeholders.

? Show a mock-up to the team (Option A): Creating a mock-up and presenting it to the team ensures alignment on design and functionality before finalizing, which is a standard next step in report development.

? Explain the desired level of reporting detail (Option B): This should have been done earlier during requirements gathering, not after meeting them.

? Present an analysis of the data (Option C): Data analysis might inform the deck, but the task focuses on building the deck, not presenting analysis.

? Find out the project due date (Option D): The due date should have been established earlier in the project planning phase.

The DAO-002 Visualization and Reporting domain includes "translating business requirements to form the appropriate visualization," and showing a mock-up ensures the reporting deck meets stakeholder expectations.

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

=====

NEW QUESTION 33

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 DAO-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+ DAO-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW QUESTION 36

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

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

=====

NEW QUESTION 40

Which of the following best describes the method used to combine files, software, and libraries for use on various operating systems and environments?

- A. Package manager
- B. Code repository
- C. Virtual machine
- D. Containerization

Answer: D

Explanation:

This question pertains to the Data Concepts and Environments domain, focusing on methods for managing software and data environments. The task is to identify a method that combines files, software, and libraries for use across different systems.

? Package manager (Option A): Package managers (e.g., npm) manage software dependencies but don't combine files and libraries for cross-system use.

? Code repository (Option B): Code repositories (e.g., GitHub) store code but don't package it for deployment across environments.

? Virtual machine (Option C): Virtual machines emulate entire operating systems, which is heavier than needed for combining files and libraries.

? Containerization (Option D): Containerization (e.g., Docker) packages files, software, and libraries into a container that can run consistently across different operating systems and environments, making it the best choice.

The DAO-002 Data Concepts and Environments domain includes understanding "data environments," and containerization is a standard method for ensuring consistency across systems.

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

=====

NEW QUESTION 41

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 DAO-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+ DAO-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.

=====

NEW QUESTION 44

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 49

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 51

A data professional wants to identify all customers who made a purchase in January. Given the following table:

CustomerID	Month	Sales
------------	-------	-------

0001		
------	--	--

January	13000	
---------	-------	--

0002		
------	--	--

March	10000	
-------	-------	--

0003		
------	--	--

April	23000	
-------	-------	--

0004		
------	--	--

May	10000	
-----	-------	--

Which of the following types of functions should the professional use to flag the customers?

A. Statistical

B. Logical

C. Mathematical

D. Date

Answer: B

Explanation:

This question falls under the Data Analysis domain, focusing on selecting the appropriate function type to filter data in a query. The task is to flag customers who made a purchase in January, which involves a conditional check.

? Statistical (Option A): Statistical functions (e.g., AVG, STDEV) analyze data distributions, not suitable for flagging specific months.

? Logical (Option B): Logical functions (e.g., WHERE Month = 'January' in SQL) are used to apply conditions and flag rows based on criteria, which fits the task.

? Mathematical (Option C): Mathematical functions (e.g., SUM, ROUND) perform calculations, not conditional flagging.

? Date (Option D): Date functions (e.g., MONTH()) manipulate dates, but the Month column is already in text format, so a logical comparison is sufficient.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods using SQL queries," and logical functions are best for conditional flagging.

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

=====

NEW QUESTION 55

A data analyst learns that a report detailing employee sales is reflecting sales only for the current month. Which of the following is the most likely cause?

A. Lack of permissions

B. An error in SQL code

C. Report refresh failure

D. Connectivity issues

Answer: B

Explanation:

This question falls under the Data Analysis domain, focusing on troubleshooting issues in data reports. The report should show all employee sales but is limited to the current month, suggesting a data retrieval issue.

? Lack of permissions (Option A): Permissions issues would likely prevent access entirely, not limit data to the current month.

? An error in SQL code (Option B): The report likely uses an SQL query to retrieve data, and an error (e.g., a WHERE clause filtering for the current month) could restrict the data to the current month, making this the most likely cause.

? Report refresh failure (Option C): A refresh failure would result in outdated data, not specifically current-month data.

? Connectivity issues (Option D): Connectivity issues would likely prevent the report from running, not limit it to a specific time frame.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods using SQL queries," and errors in SQL code are a common cause of incorrect data retrieval in reports.

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

=====

NEW QUESTION 59

A report triggers an error that prevents information from being displayed. However, the report was functional before a database upgrade. Which of the following should a data analyst do first to troubleshoot the problem?

A. Ensure the system has permissions for the report service.

B. Change the report's refresh rate.

C. Verify the connection to the database.

D. Check whether the data structures were modified.

Answer: C

Explanation:

This question falls under the Data Concepts and Environments domain, which in DA0-002 involves understanding database environments, connections, and troubleshooting issues related to data access. The scenario describes a report failing after a database upgrade, indicating a potential issue with the database environment or connectivity.

? Ensure the system has permissions for the report service (Option A): While permissions issues can cause report failures, they are less likely to be the first issue after a database upgrade unless explicitly mentioned.

? Change the report's refresh rate (Option B): Refresh rate adjustments might address performance issues but won't resolve a fundamental error preventing data display.

? Verify the connection to the database (Option C): A database upgrade often involves changes to connection strings, drivers, or network configurations. Verifying the connection ensures the report can access the database, making this the most logical first step.

? Check whether the data structures were modified (Option D): While possible, checking data structures (e.g., schema changes) is a deeper troubleshooting step that should follow after confirming basic connectivity.

The DA0-002 Data Concepts and Environments domain includes understanding database connectivity, similar to DA0-001's focus on "data schemas and dimensions" and environments (web ID: 1). Verifying the connection is the first recommended step in troubleshooting post-upgrade issues.

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

=====

NEW QUESTION 60

A data analyst is evaluating all conditions in a query. Which of the following is the best logical function to accomplish this task?

A. OR

B. NOT

C. AND

D. IF

Answer: C

Explanation:

This question falls under the Data Analysis domain, focusing on SQL logical functions for query evaluation. The task is to evaluate "all conditions," implying multiple conditions must be true together.

? OR (Option A): OR returns true if any condition is true, not ensuring all conditions are met.

? NOT (Option B): NOT negates a condition, not suitable for combining multiple conditions.

? AND (Option C): AND requires all conditions to be true, which aligns with evaluating "all conditions" in a query.

? IF (Option D): IF is a conditional function for decision-making, not for evaluating multiple conditions together.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods using SQL queries," and AND is the best logical function for ensuring all conditions are met.

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

=====

NEW QUESTION 64

Before distributing a report, a marketing analyst notices that the total distinct promotional email messages is less than the combined total of emails sent. Which of the following is the most likely reason for this difference?

A. The aggregation did not include all emails.

B. Some emails were not delivered.

C. The report failed to run properly.

D. A recipient received duplicate emails.

Answer: D

Explanation:

This question falls under the Data Analysis domain, focusing on analyzing discrepancies in data reports. The total distinct messages are fewer than the total emails sent, indicating a specific issue.

? The aggregation did not include all emails (Option A): If the aggregation missed

emails, the total sent would be lower, not the distinct count.

? Some emails were not delivered (Option B): Undelivered emails would reduce the total sent, but the scenario implies the total sent is accurate.

? The report failed to run properly (Option C): A report failure would likely cause broader issues, not a specific discrepancy between distinct and total counts.

? A recipient received duplicate emails (Option D): If recipients received duplicates, the total emails sent would be higher than the distinct messages (unique email content), explaining the difference.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods," and identifying duplicates is a common analysis task to explain such discrepancies.

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

=====

NEW QUESTION 65

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

Explanation:

This question pertains to the Data Governance domain, focusing on data classification for security purposes. User information databases typically contain personal data, and security triggers (e.g., alerts for unauthorized access) require classifying data to determine protection levels.

? Public (Option A): Public data is openly accessible (e.g., company brochures), not suitable for user information requiring security triggers.

? Open (Option B): Open isn't a standard data classification; it's similar to public and not applicable here.

? Sensitive (Option C): Sensitive data includes information that, if exposed, could cause harm (e.g., user emails, roles), which fits user information and warrants security triggers.

? Non-Sensitive (Option D): Non-sensitive data doesn't require protection, so it wouldn't need security triggers.

? Private (Option E): Private data includes PII (e.g., names, addresses) in user information databases, requiring security measures like triggers to protect against breaches.

? Encrypted (Option F): Encrypted refers to a data state, not a classification; data can be classified as private or sensitive and then encrypted.

The DA0-002 Data Governance domain includes "data quality control concepts," such as

classifying data to apply appropriate security measures. Sensitive and private classifications are most relevant for user information.

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

=====

NEW QUESTION 69

Which of the following is the best reason for a company to use a CSV file to share data instead of an Excel file?

- A. CSV files can store different types of encoding.
- B. CSV files are not vendor-specific.
- C. CSV files are smaller in size.
- D. CSV files are easier to change in text editors.

Answer: B

Explanation:

This question pertains to the Data Concepts and Environments domain, focusing on file formats for data sharing. The task is to identify the best reason to choose CSV over Excel for sharing data.

? CSV files can store different types of encoding (Option A): While CSV files can use different encodings, this isn't the primary reason to choose them over Excel.

? CSV files are not vendor-specific (Option B): CSV is a plain-text format that can be opened by any software, unlike Excel files, which are tied to Microsoft Excel, making CSV more interoperable and the best reason for sharing.

? CSV files are smaller in size (Option C): CSV files are often smaller due to their simplicity, but this isn't always the primary reason for sharing.

? CSV files are easier to change in text editors (Option D): While true, this isn't the most compelling reason for sharing data across systems.

The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," and CSV's vendor-neutral nature makes it ideal for sharing data.

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

=====

NEW QUESTION 70

A manager wants a report that contains the days off for each direct report. The manager needs this report to always be up-to-date with the latest data. Which of the following describes the refresh frequency that the manager is requesting?

- A. Real-time
- B. Ad hoc
- C. Snapshot
- D. Dynamic

Answer: A

Explanation:

This question pertains to the Visualization and Reporting domain, focusing on report refresh frequencies. The manager needs the report to always be up-to-date, implying continuous data updates.

? Real-time (Option A): Real-time refresh frequency ensures the report reflects the latest data as soon as it changes, which matches the requirement to "always be up-to-date."

? Ad hoc (Option B): Ad hoc reports are generated on-demand, not continuously updated.

? Snapshot (Option C): A snapshot captures data at a specific point in time, not suitable for always being up-to-date.

? Dynamic (Option D): Dynamic reports allow interactivity, but the term doesn't specifically imply real-time updates.

The DA0-002 Visualization and Reporting domain includes "the appropriate visualization in the form of a report" with delivery methods, and real-time refresh frequency ensures the report is always current.

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

=====

NEW QUESTION 75

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: C

Explanation:

This question falls under the Data Analysis domain, focusing on analytical methods for dashboards. The analyst wants to identify patterns in historical data for further study, which points to a specific type of analytics.

? Predictive (Option A): Predictive analytics forecasts future outcomes, not focused on identifying patterns for further study.

? Prescriptive (Option B): Prescriptive analytics provides recommendations, which goes beyond identifying patterns.

? Diagnostic (Option C): Diagnostic analytics examines historical data to identify patterns, trends, and correlations, enabling further investigation, which fits the scenario.

? Descriptive (Option D): Descriptive analytics summarizes what happened but doesn't focus on identifying patterns for deeper study.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods," and diagnostic analytics is best for pattern identification in historical data.

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

=====

NEW QUESTION 76

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 79

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:

This question falls under the Data Analysis domain, focusing on calculating new values from existing data. The task is to determine the total cost per transaction, which involves multiplying Quantity by Item Price.

? Parsing (Option A): Parsing involves breaking down data (e.g., splitting a string), not calculating totals.

? Scaling (Option B): Scaling adjusts numerical values to a common range (e.g., normalization), not relevant for calculating totals.

? Compressing (Option C): Compressing reduces data size, not applicable to calculating costs.

? Deriving (Option D): Deriving involves creating new data fields by performing calculations on existing ones (e.g., Total Cost = Quantity ?? Item Price), which fits the task.

The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods," such as deriving new fields through calculations to analyze data. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.

=====

NEW QUESTION 81

A data analyst needs to create a combined report that includes information from the following two tables:

Managers table

ID

First_name Last_name Job_title 1001

John Doe Manager 1002

Jane Roe Director

Non-managers table ID

First_name Last_name Job_title 1003

Robert Roe

Business Analyst 1004

Jane Doe

Sales Representative 1005

John Roe

Operations Analyst

Which of the following query methods should the analyst use for this task?

- A. Group
- B. Join
- C. Union
- D. Nested

Answer: C

Explanation:

This question pertains to the Data Acquisition and Preparation domain, focusing on combining data from two tables. Both tables have the same structure (ID, First_name, Last_name, Job_title) and contain employee data, suggesting the task is to create a single list of all employees.

? Group (Option A): Grouping (e.g., GROUP BY in SQL) is for aggregation (e.g., counting employees by job title), not combining tables into a single report.

? Join (Option B): Joining tables (e.g., INNER JOIN) requires a common key and combines tables horizontally, but there??s no indication of a relationship between the tables (e.g., no shared key beyond ID, which isn??t linked).

? Union (Option C): UNION combines the rows of two tables with the same structure into a single result set, removing duplicates, which is ideal for creating a combined report of all employees from both tables.

? Nested (Option D): Nested queries (e.g., subqueries) are used for complex filtering, not for combining tables into a single list.

The DA0-002 Data Acquisition and Preparation domain includes "executing data manipulation," and UNION is the best method for combining two tables with identical structures into a single report.

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

NEW QUESTION 85

A data analyst is creating a new dataset that involves bringing together the following datasets:

Name ID
Date of birth
Frank 23525
3/19
Martha 11290
6/13
Ellen 12141
11/4
ID
Address City State 23525
1234 Harding Chicago
IL 11040
935 Terrace Hills Chino
CA 11290
2 Speedway Miami
FL

Which of the following would be the output if the data analyst does a FULL JOIN?

- A. Name IDDate of birth Address CityState Frank 235253/191234 Harding ChicagoILMartha 112906/13935 Terrace Hills ChinoCAEllen 1214111/42 Speedway MiamiFL
B. Name IDDate of birth Address CityState Frank 235253/191234 Harding ChicagoILMartha112906/13935 Terrace Hills ChinoCAEllen 1214111/42 Speedway MiamiFL
C. Name IDDate of birth Address CityState Frank 235253/191234 Harding ChicagoILMartha 112906/132 Speedway MiamiFLEllen 1214111/4935 Terrace Hills ChinoCA
D. Name IDDate of birth Address CityState Frank 235253/191234 Harding ChicagoILMartha 112906/132 Speedway MiamiFLEllen 1214111/411040935 Terrace Hills ChinoCA

Answer: D

Explanation:

This question falls under theData Concepts and Environmentsdomain, focusing on database operations like joins. A FULL JOIN combines all rows from both tables, including matches and non-matches, filling in NULLs where there??s no corresponding data.

? The first table has IDs: 23525 (Frank), 11290 (Martha), 12141 (Ellen).

? The second table has IDs: 23525, 11040, 11290.

? A FULL JOIN includes all IDs: 23525, 11290, 12141, 11040.

? Option A: Incorrect; it includes a row for Ellen with "2 Speedway," but Ellen??s ID (12141) doesn??t match any address, and 11040 is missing.

? Option B: Identical to Option A, so incorrect for the same reasons.

? Option C: Incorrect; it mismatches addresses (e.g., Ellen with 935 Terrace Hills, which belongs to 11040).

? Option D: Correct; it includes all IDs, with NULLs for non-matching rows (Ellen has no address, and 11040 has no name).

The DAO-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," such as performing joins in relational databases.

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

=====

NEW QUESTION 87

A data analyst receives the following sales data for a convenience store:

Item Quantity Price
Chocolate Bars 7
\$1.99
Vanilla Ice Bars
2
\$4.99
Chocolate Wafers 6
\$0.99
Peanut Butter 2
\$2.99
Cups 3
\$4.99
Strawberry Jam 3
\$4.99
Chocolate Cake 9
\$6.99
Milk Chocolate 2
\$2.99
Almonds 5
\$2.99

The analyst needs to provide information on the products that contain chocolate. Which of the following RegEx should the analyst use to filter the chocolate products?

- A. Chocolate!
B. Chocolate\$
C. %Chocolate&
D. #Chocolate#

Answer: B

Explanation:

This question falls under theData Acquisition and Preparationdomain, which includes techniques for manipulating and filtering data, such as using regular expressions (RegEx) to identify specific patterns in text data. The task is to filter items containing the word "Chocolate."

? Chocolate! (Option A): In RegEx, "!" is not a valid pattern for matching a word like

"Chocolate." It typically denotes negation in some contexts, but here it??s incorrect.

? Chocolate\$ (Option B): The "\$" in RegEx anchors the pattern to the end of the string, meaning it matches "Chocolate" at the end of an item name (e.g., "Milk

Chocolate"). This is the most appropriate pattern for identifying items ending with "Chocolate," which applies to the relevant items in the list.
? %Chocolate& (Option C): "%" and "&" are not standard RegEx anchors; they??re often used in SQL LIKE patterns, not RegEx, making this incorrect.
? #Chocolate#\$ (Option D): "#" is not a standard RegEx anchor, and this pattern would look for "Chocolate" surrounded by "#", which doesn??t match the data.
The DA0-002 Data Acquisition and Preparation domain includes "executing data manipulation" , and RegEx is a common technique for filtering text data. The pattern "Chocolate\$" correctly identifies items like "Chocolate Bars," "Chocolate Wafers," "Chocolate Cake," and "Milk Chocolate."
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 2.0 Data Acquisition and Preparation
=====

NEW QUESTION 89

A data breach occurs at a company. Which of the following actions should be taken?

- A. Make an announcement on social media so customers are aware as soon as possible.
- B. Tell the company management team and then tell regulatory agencies.
- C. Keep the incident a secret until the issue is resolved.
- D. Inform the entire IT sector, but ask for discretion.

Answer: B

Explanation:

This question falls under theData Governancedomain, focusing on data breach response protocols. A data breach requires a structured response to comply with legal and regulatory requirements.
? Make an announcement on social media so customers are aware as soon as possible (Option A): Public announcement without internal coordination or regulatory notificationcan lead to legal issues and loss of trust.
? Tell the company management team and then tell regulatory agencies (Option B): This follows best practices: inform internal leadership to coordinate a response, then notify regulatory agencies as required by laws (e.g., GDPR mandates notification within 72 hours).
? Keep the incident a secret until the issue is resolved (Option C): This violates regulations requiring timely breach notification.
? Inform the entire IT sector, but ask for discretion (Option D): Sharing with the IT sector is vague and risks leaks; regulatory agencies should be prioritized.
The DA0-002 Data Governance domain includes "data privacy concepts," such as proper breach response procedures, emphasizing internal and regulatory notification. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.
=====

NEW QUESTION 94

A project manager requests an unscheduled report that provides a list of clients. Which of the following frequencies is best for this report?

- A. Annual
- B. Daily
- C. Weekly
- D. Ad hoc

Answer: D

Explanation:

This question pertains to theVisualization and Reportingdomain, focusing on report delivery frequencies. The report is described as unscheduled, meaning it??s a one-time request.
? Annual (Option A): Annual frequency implies a scheduled report every year, not suitable for an unscheduled request.
? Daily (Option B): Daily frequency implies a scheduled report each day, not suitable.
? Weekly (Option C): Weekly frequency implies a scheduled report each week, not suitable.
? Ad hoc (Option D): Ad hoc reports are generated on-demand for one-time or unscheduledneeds, which matches the project manager??s request.
The DA0-002 Visualization and Reporting domain includes "the appropriate visualization in the form of a report" with delivery methods, and ad hoc is the best frequency for an unscheduled report.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.
=====

NEW QUESTION 98

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:

This question pertains to theData Governancedomain, focusing on ensuring data quality and reliability in dashboards over time. Daily troubleshooting indicates a recurring issue, and the task is to validate data moving forward.
? Inquiring about structure changes (Option A): This might identify past issues but doesn??t provide ongoing validation.
? Setting up monitoring alerts (Option B): Monitoring alerts can automatically notify the analyst of data issues (e.g., missing updates, errors), providing a proactive way to validate data continuously.
? Reaching out to users daily (Option C): This is inefficient and reactive, not a sustainable validation method.
? Rebuilding the dashboard (Option D): Rebuilding might fix current issues but doesn??t ensure future validation.
The DA0-002 Data Governance domain includes "data quality control concepts," such as implementing monitoring to ensure data reliability in dashboards.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.

NEW QUESTION 99

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 101

Which of the following best describes the semi-structured data that is gathered when web scraping?

- A. JSON
- B. CSV
- C. CSS
- D. HTML

Answer: A

Explanation:

This question pertains to the Data Acquisition and Preparation domain, which in DA0-002

includes understanding data acquisition concepts and the types of data gathered from various sources, such as web scraping. Web scraping involves extracting data from websites, and the data gathered is often semi-structured, meaning it has some organizational structure but isn't fully relational like a database table.

? JSON (Option A): JSON (JavaScript Object Notation) is a semi-structured data

format commonly used in web applications. Web scraping often retrieves data in JSON format via APIs or embedded scripts, as it's lightweight and structured with key-value pairs, making it ideal for semi-structured data.

? CSV (Option B): CSV (Comma-Separated Values) is a structured format typically

used for tabular data. It's not commonly the direct output of web scraping, though scraped data might be converted to CSV later.

? CSS (Option C): CSS (Cascading Style Sheets) is used for styling web pages and

isn't a data format, making it irrelevant for describing scraped data.

? HTML (Option D): HTML (HyperText Markup Language) is the structure of web pages and is often the raw format scraped during web scraping. While HTML is semi-structured due to its tag-based hierarchy, it's primarily a markup language, not a data format, and the actual data extracted is often parsed into formats like JSON.

The DA0-002 Data Acquisition and Preparation domain aligns with the DA0-001 focus on "data acquisition concepts" (web ID: 14), which includes identifying formats like JSON as semi-structured data commonly acquired through web scraping. JSON is the best fit here due to its prevalence in web data exchange.

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

=====

NEW QUESTION 104

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 108

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

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

=====

NEW QUESTION 113

A data analyst is creating a pivot table for a large dataset for an upcoming board meeting. Which of the following is the purpose of the pivot table?

- A. To visualize the data in a dashboard
- B. To retrieve and clean data from several sources
- C. To summarize and analyze the data
- D. To organize the data for reporting

Answer: C

Explanation:

This question pertains to the Data Analysis domain, focusing on the purpose of a pivot table. Pivot tables are a tool for summarizing and analyzing data, often used in preparation for reporting.

? To visualize the data in a dashboard (Option A): Pivot tables summarize data but aren't visualizations; charts in dashboards might be created from pivot tables.

? To retrieve and clean data from several sources (Option B): Retrieving and cleaning data is part of data preparation, not the purpose of a pivot table.

? To summarize and analyze the data (Option C): Pivot tables aggregate and summarize data (e.g., by calculating sums, averages) and allow for analysis (e.g., filtering, grouping), which is their primary purpose.

? To organize the data for reporting (Option D): While pivot tables can help organize data, their main purpose is summarization and analysis, not just organization.

The DAO-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods," and pivot tables are a key tool for summarizing and analyzing large datasets.

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

=====

NEW QUESTION 114

A grocery store wants to view the revenue from the previous year, highlighting individual departments. Which of the following is the most appropriate chart to communicate this information?

- A. Gantt
- B. Pie
- C. Area
- D. Line

Answer: B

Explanation:

This question is part of the Visualization and Reporting domain, focusing on selecting the appropriate visualization for a given dataset. The grocery store wants to view revenue by department, which requires a chart that shows proportions or comparisons across categories.

? Gantt (Option A): Gantt charts are used for project scheduling, not for comparing revenue across categories.

? Pie (Option B): Pie charts are ideal for showing proportions or percentages of a whole, such as revenue distribution across departments, making this the best choice.

? Area (Option C): Area charts are better for showing trends over time, not static categorical comparisons.

? Line (Option D): Line charts are used for trends over time, not for comparing discrete categories like departments.

The DAO-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization", and a pie chart is the most appropriate for showing departmental revenue proportions.

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

=====

NEW QUESTION 118

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 120

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 123

A sales manager wants to understand how sales are trending year over year. Which of the following chart types is the most appropriate to display the information?

A. Line

B. Donut

C. Bubble

D. Hierarchy

Answer: A

Explanation:

This question falls under the Visualization and Reporting domain, focusing on selecting the appropriate visualization for a specific data trend. The task is to show sales trends over time (year over year).

? Line (Option A): Line charts are ideal for displaying trends over time, such as year-over-year sales, as they clearly show changes and patterns across a continuous time axis.

? Donut (Option B): Donut charts show proportions or percentages of a whole, not suitable for time-based trends.

? Bubble (Option C): Bubble charts display three dimensions of data (e.g., size, x-axis, y-axis), not ideal for simple time trends.

? Hierarchy (Option D): Hierarchy charts (e.g., treemaps) show nested relationships, not time-based trends.

The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and a line chart is best for time-series trends.

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

=====

NEW QUESTION 127

.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questons and Answers in PDF Format

DA0-002 Practice Exam Features:

- * DA0-002 Questions and Answers Updated Frequently
- * DA0-002 Practice Questions Verified by Expert Senior Certified Staff
- * DA0-002 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * DA0-002 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year

100% Actual & Verified — Instant Download, Please Click
[Order The DA0-002 Practice Test Here](#)