CompTIA

Exam Questions DA0-002

CompTIA Data+ Exam (2025)

NEW QUESTION 1
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 theData Analysisdomain, 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 DA0-002 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+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.
==========

NEW QUESTION 2
A sales manager wants a dashboard that shows sales aggregated by region and identifies high-volume sales by salesperson per region. Which of the following communication techniques best displays this information?

A. Defined parameters
B. Filter options
C. Level of detail
D. User persona

Answer: B

Explanation:
This question pertains to theVisualization and Reportingdomain, focusing on dashboard features for displaying data. The dashboard needs to show aggregated sales by region and allow identification of high-volume sales by salesperson within each region.
? Defined parameters (Option A): Parameters set specific values (e.g., a date range), but they don??t directly enable interaction with aggregated data.
? Filter options (Option B): Filter options allow the user to select a region and then view salespersons within that region, enabling the identification of high-volume sales per region interactively.
? Level of detail (Option C): Level of detail determines the granularity of data shown but doesn??t facilitate interactive exploration.
? User persona (Option D): User personas guide dashboard design but aren??t a communication technique for displaying data.
The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and filter options best enable the interactive analysis required.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.
==========

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 theData Concepts and Environmentsdomain, 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 DA0-002 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+ DA0-002 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 theData Concepts and Environmentsdomain, 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
A data analyst needs to get an accurate idea of how data components are automated. Which of the following types of documentation should the analyst review first?

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

Answer: A

Explanation:
This question pertains to theData Concepts and Environmentsdomain, focusing on documentation for understanding data processes. The analyst needs to understand automation of data components, which involves data movement and processes.
? Data flow diagram (Option A): A data flow diagram (DFD) visualizes how data
moves through systems, including automated processes, making it the best starting point.
? Data explainability report (Option B): This is related to AI/ML model transparency,
not data automation.
? Data dictionary (Option C): A data dictionary defines data elements, not how they??re automated.
? Data lineage (Option D): Data lineage tracks data origin and transformations but doesn??t focus on automation processes.
The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," and a data flow diagram is key for visualizing automation. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.
==========

NEW QUESTION 6
Which of the following best represents a type of infrastructure that requires a company to purchase and maintain all of its own servers?

A. Private
B. Cloud
C. Hybrid
D. Public

Answer: A

Explanation:
This question pertains to theData Concepts and Environmentsdomain, focusing on types of server infrastructure. The task is to identify an infrastructure where a company owns and maintainsall servers.
? Private (Option A): A private infrastructure (often on-premises) means the
company owns and maintains its own servers, typically in a private data center, which matches the requirement.
? Cloud (Option B): Cloud infrastructure is managed by third-party providers, not
owned by the company.
? Hybrid (Option C): Hybrid combines on-premises and cloud, so not all servers are owned by the company.
? Public (Option D): Public infrastructure is a cloud model shared across multiple organizations, not owned by the company.
The DA0-002 Data Concepts and Environments domain includes understanding "data environments," and a private infrastructure requires the company to purchase and maintain its own servers.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.
==========

NEW QUESTION 7
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 theData Concepts and Environmentsdomain, 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 8
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 theData Analysisdomain, 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 >= 125000This 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 >= 125000This 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 >
125000This 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 >=
125000This 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 9
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 theData Governancedomain, 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 DA0-002 Data Governance domain includes "data quality control concepts," and checking for duplication is critical to ensure accurate calculations like averages. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.
==========

NEW QUESTION 10
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 theData Concepts and Environmentsdomain, 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 DA0-002 Data Concepts and Environments domain includes understanding "data environments," and containerization is a standard method for ensuring consistency across systems.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.
==========

NEW QUESTION 10
A data analyst needs to create and deliver a dashboard that displays the company's financial transactions as they are updated. Which of the following delivery methods should the analyst consider? (Select two).

A. Real-time
B. Snapshot
C. Dynamic
D. Static
E. Ad hoc
F. Time series

Answer: AC

Explanation:
This question is part of theVisualization and Reportingdomain, focusing on delivery methods for dashboards. The requirement for displaying financial transactions "as they are updated" implies a need for real-time updates and interactivity, which narrows down the options.
? Real-time (Option A): Real-time delivery ensures the dashboard reflects the latest
data as transactions are updated, meeting the requirement.
? Snapshot (Option B): A snapshot provides a static view at a specific point, not suitable for ongoing updates.
? Dynamic (Option C): A dynamic dashboard allows for interactivity and can be updated as data changes, complementing real-time delivery.
? Static (Option D): Static dashboards don??t update automatically, making this incorrect.
? Ad hoc (Option E): Ad hoc delivery is for one-time reports, not ongoing updates.
? Time series (Option F): Time series refers to a data type or visualization, not a delivery method.
The DA0-002 Visualization and Reporting domain includes understanding "the appropriate visualization in the form of a report or dashboard" with delivery methods Real-time and dynamic methods best support the need for updated financial transaction dashboards. Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization andReporting

NEW QUESTION 14
A data analyst is generating a custom report for a Chief Executive Officer's executive meeting. Later, the analyst learns that other custom reports will be required for future executive meetings. Which of the following delivery methods should the analyst use?

A. Ad hoc
B. Real-time
C. Recurring
D. Self-service

Answer: C

Explanation:
This question falls under theVisualization and Reportingdomain of DA0-002, which involves selecting appropriate delivery methods for reports. The scenario describes a need for custom reports for future executive meetings, implying a scheduled, repeated delivery.
? Ad hoc (Option A): Ad hoc reports are generated on-demand for one-time use, not
suitable for ongoing needs.
? Real-time (Option B): Real-time delivery provides live data updates, which isn??t necessary for scheduled executive meetings.
? Recurring (Option C): Recurring delivery involves scheduling reports to be generated and delivered at regular intervals (e.g., weekly or monthly), which fits the need for future executive meetings.
? Self-service (Option D): Self-service allows users to generate reports themselves, but the scenario implies the analyst will create the reports.
The DA0-002 Visualization and Reporting domain includes understanding "the appropriate visualization in the form of a report" with delivery methods , and recurring delivery aligns with scheduled reporting needs.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting
==========

NEW QUESTION 18
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 theData Analysisdomain, 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 21
An analyst must use the same data to create reports for police patrol supervisors, the city council, and a public website. Which of the following is the best way to differentiate the reports?

A. Persona type
B. Detail levels
C. Accessibility
D. Sensitivity

Answer: A

Explanation:
This question falls under theVisualization and Reportingdomain, focusing on tailoring reports for different audiences. The same data is used for three distinct groups (supervisors, city council, public), requiring differentiation.
? Persona type (Option A): Persona types define the audience??s needs and
preferences (e.g., supervisors need operational details, the city council needs summaries, the public needs simplified data), making this the best way to differentiate the reports.
? Detail levels (Option B): Detail levels are a result of persona types, not the method of differentiation.
? Accessibility (Option C): Accessibility ensures access for all users (e.g., screen readers), but it??s not the primary way to differentiate content.
? Sensitivity (Option D): Sensitivity determines data access (e.g., confidential vs. public), but the scenario implies all reports use the same data, just presented differently.
The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and persona types guide report differentiation for diverse audiences.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.
==========

NEW QUESTION 23
The sales department wants to include the composition of total sales amounts across all three sales channels in a report. Given the following sample sales table:
Sales channel
Month
Sales (million $)
Digital January 135
Store February 145
Online March 165
Store April 200
Store May 125
Online June 155
Digital July 120
Online August 145

Digital September 160
Which of the following visualizations is the most appropriate?

A. Pivot table
B. Pie chart
C. KPI card
D. Box plot

Answer: B

Explanation:
This question pertains to theVisualization and Reportingdomain, focusing on selecting the appropriate visualization for a specific requirement. The task is to show the composition of total sales across three channels, which involves showing proportions.
? Pivot table (Option A): A pivot table summarizes data but isn??t a visualization; it??s
more for data exploration.
? Pie chart (Option B): A pie chart shows the proportion of total sales for each channel (Digital, Store, Online), which is ideal for displaying composition.
? KPI card (Option C): A KPI card displays a single metric, not suitable for showing composition across multiple channels.
? Box plot (Option D): A box plot shows data distribution (e.g., quartiles), not proportions.
The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and a pie chart is best for showing the composition of totals.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.
==========

NEW QUESTION 28
A table contains several rows of cellular numbers with call timestamps, call durations, called numbers, and carriers of the called number. Which of the following allows a data analyst to sort the cellular numbers based on the carriers of the called numbers and include the total call durations?

A. SELECT cellular_number, called_number_carrier, SUM(call_duration) FROM calls GROUP BY cellular_number ORDER BY called_number_carrier;
B. SELECT cellular_number, SUM(call_duration) FROM calls GROUP BY call_durationORDERBY called_number_carrier;
C. SELECT cellular_number, called_number_carrier, SUM(call_duration) FROM calls GROUP BY cellular_number, called_number_carrier ORDER BY called_number_carrier;
D. SELECT cellular_number, called_number_carrier, SUM(call_duration) FROM calls GROUP BY call_duration ORDER BY called_number_carrier;

Answer: C

Explanation:
This question falls under theData Analysisdomain of CompTIA Data+ DA0-002, focusing on SQL queries for data analysis. The task requires sorting cellular numbers by the carrier of the called number (called_number_carrier) and calculating the total call durations (SUM(call_duration)).
? Option A: SELECT cellular_number, called_number_carrier, SUM(call_duration)
FROM calls GROUP BY cellular_number ORDER BY called_number_carrierThis query groups by cellular_number only, but called_number_carrier is in the SELECT clause without being in the GROUP BY, which is invalid in SQL (it would raise an error in most databases).
? Option B: SELECT cellular_number, SUM(call_duration) FROM calls GROUP BY
call_duration ORDER BY called_number_carrierThis query doesn??t include called_number_carrier in the SELECT clause, so it cannot be used in the ORDER BY clause, making it invalid. Grouping by call_duration also doesn??t align with the task.
? Option C: SELECT cellular_number, called_number_carrier, SUM(call_duration)
FROM calls GROUP BY cellular_number, called_number_carrier ORDER BY called_number_carrierThis query correctly groups by both cellular_number and called_number_carrier (since both are in the SELECT clause), calculates the total call duration with SUM(call_duration), and sorts by called_number_carrier as required.
? Option D: SELECT cellular_number, called_number_carrier, SUM(call_duration)
FROM calls GROUP BY call_duration ORDER BY called_number_carrierGrouping by call_duration is incorrect because cellular_number and called_number_carrier are in the SELECT clause but not in the GROUP BY, making this query invalid.
The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods using SQL queries," and Option C correctly aggregates and sorts the data as specified.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.
==========

NEW QUESTION 33
The following SQL code returns an error in the program console: SELECT firstName, lastName, SUM(income)
FROM companyRoster SORT BY lastName, income
Which of the following changes allows this SQL code to run?

A. SELECT firstName, lastName, SUM(income) FROM companyRoster HAVING SUM(income) > 10000000
B. SELECT firstName, lastName, SUM(income) FROM companyRoster GROUP BY firstName, lastName
C. SELECT firstName, lastName, SUM(income) FROM companyRoster ORDER BY firstName, income
D. SELECT firstName, lastName, SUM(income) FROM companyRoster

Answer: B

Explanation:
This question falls under theData Analysisdomain, focusing on SQL query correction. The query uses an aggregate function (SUM) but has two issues: it uses "SORT BY" (incorrect syntax) and lacks a GROUP BY clause for non-aggregated columns.
? The query selects firstName, lastName, and SUM(income), but firstName and
lastName are not aggregated, requiring a GROUP BY clause.
? "SORT BY" is incorrect; the correct syntax is "ORDER BY."
? Option A: SELECT firstName, lastName, SUM(income) FROM companyRoster HAVING SUM(income) > 10000000This adds a HAVING clause but doesn??t fix the GROUP BY issue, so it??s still invalid.
? Option B: SELECT firstName, lastName, SUM(income) FROM companyRoster GROUP BY firstName, lastNameThis adds the required GROUP BY clause for firstName and lastName, fixing the aggregation error. While it removes the ORDER BY, the query will run without it, addressing the primary error.
? Option C: SELECT firstName, lastName, SUM(income) FROM companyRoster
ORDER BY firstName, incomeThis fixes "SORT BY" to "ORDER BY" but doesn??t address the missing GROUP BY, so the query remains invalid.
? Option D: SELECT firstName, lastName, SUM(income) FROM

companyRosterThis removes the ORDER BY but still lacks the GROUP BY clause, making it invalid.
The DA0-002 Data Analysis domain includes "applying the appropriate descriptive statistical methods using SQL queries," and adding GROUP BY fixes the aggregation error, allowing the query to run.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 3.0 Data Analysis.
==========

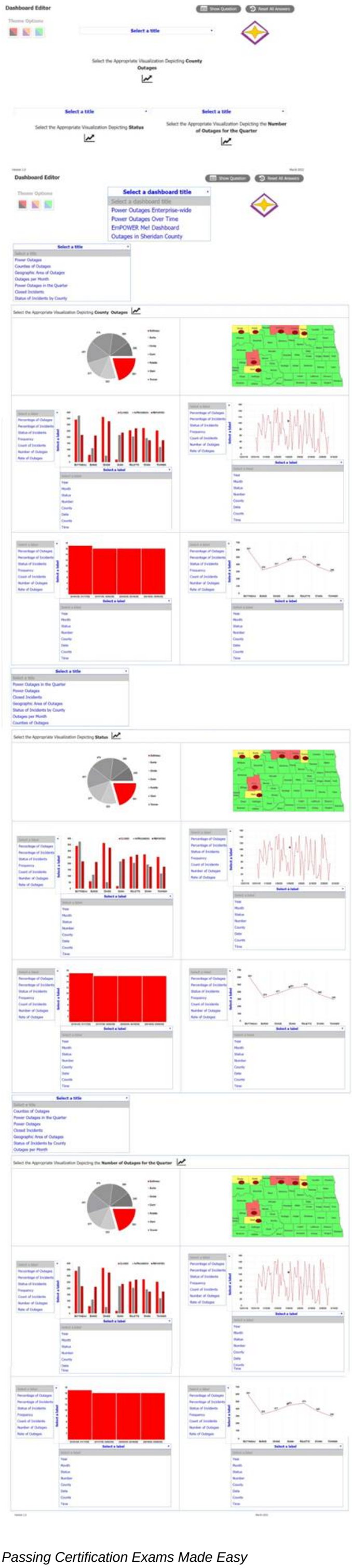
NEW QUESTION 34
Which of the following is the best tool for creating a dynamic dashboard?

A. Power BI
B. RStudio
C. Excel
D. SAS

Answer: A

Explanation:
The question asks for the best tool to create adynamic dashboard, which falls under the Visualization and Reportingdomain of CompTIA Data+ DA0-002. According to the DA0- 002 draft objectives, this domain includes understanding tools and techniques for creating effective visualizations, such as dashboards, that can be updated dynamically to reflect real-time or changing data. A dynamic dashboard typically allows for interactivity, real-time updates, and user-driven exploration of data, which is a key focus in this domain.
? Power BI (Option A): Power BI is a business intelligence tool by Microsoft
designed specifically for creating interactive and dynamic dashboards. It supports real-time data updates, user interactivity (e.g., filters, slicers), and integration with various data sources, making it ideal for dynamic dashboard creation.
? RStudio (Option B): RStudio is primarily an IDE for the R programming language,
used for statistical computing and data analysis. While it can create visualizations, it??s not optimized for dynamic dashboards without additional packages like Shiny, and even then, it requires more coding effort compared to Power BI.
? Excel (Option C): Excel is a spreadsheet tool that can create static charts and
basic dashboards, but it lacks the interactivity and real-time update capabilities of a true dynamic dashboard tool like Power BI.
? SAS (Option D): SAS is a statistical analysis software suite that excels in data
mining and analytics but is not primarily designed for creating dynamic, interactive dashboards.
The DA0-002 Visualization and Reporting domain emphasizes tools that facilitate "the appropriate visualization in the form of a report or dashboard with the proper design components," as noted in similar DA0-001 objectives (web ID: 1). Power BI aligns best with this requirement due to its focus on dynamic, user-friendly dashboard creation.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting
==========

NEW QUESTION 36
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.



Passing Certification Exams Made Easy *visit - https://www.surepassexam.com*

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 38
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 theData Governancedomain, 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 40
Which of the following AI types is the best option for time-series forecasting?

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

Answer: B

Explanation:
Foundational models are large AI models trained on vast amounts of data, often exhibiting strong generalization capabilities. While not specifically architected for time-series, their ability to learn complex patterns could potentially be leveraged for forecasting tasks through fine-tuning or specialized architectures built upon them.
In reality, the best AI types specifically designed for time-series forecasting include:
? Recurrent Neural Networks (RNNs), especially LSTMs and GRUs:These architectures are designed to handle sequential data and capture temporal dependencies.
? Transformer Networks:Originally developed for NLP, Transformers have shown remarkable success in time-series forecasting due to their ability to capture long-range dependencies.
? Traditional statistical models:ARIMA, Exponential Smoothing, and other statistical methods remain powerful and interpretable options for time-series analysis.
Therefore, while "foundational models" have some potential, it's important to understand that they aren't the primary or specifically designed AI type for time-series forecasting.

NEW QUESTION 41
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 theVisualization and Reportingdomain, 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 46
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 theData Analysisdomain, 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 48
Given the following tables:
Individual table ID
FirstName LastName 1
John Doe Output
ID
FullName
1
JohnDoe
Which of the following is the best option to display output from FirstName and LastName as FullName?

A. Concatenate
B. Filter
C. Join
D. Group

Answer: A

Explanation:
This question falls under theData Acquisition and Preparationdomain of CompTIA Data+ DA0-002, focusing on data manipulation techniques. The task is to combine FirstName and LastName into a single FullName field (e.g., "JohnDoe").
? Concatenate (Option A): Concatenation combines two or more strings into one (e.g., usingCONCAT in SQL or "+" in Python), which is the correct method to create FullName from FirstName and LastName.
? Filter (Option B): Filtering selects specific rows based on conditions, not suitable for combining fields.
? Join (Option C): Joining combines data from multiple tables, but the task involves manipulating data within a single table.
? Group (Option D): Grouping (e.g., GROUP BY in SQL) is for aggregation, not for combining fields into a new column.
The DA0-002 Data Acquisition and Preparation domain includes "executing data manipulation," and concatenation is the standard technique for combining fields like FirstName and LastName into FullName.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 2.0 Data Acquisition and Preparation.
==========

NEW QUESTION 51
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 theData Acquisition and Preparationdomain, 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 56
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 57
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 58
Which of the following file types separates data using a delimiter?

A. XML
B. HTML
C. JSON
D. CSV

Answer: D

Explanation:
This question falls under theData Concepts and Environmentsdomain, focusing on understanding file formats and their structures. The task is to identify a file type that uses delimiters to separate data.
? XML (Option A): XML uses tags to structure data, not delimiters.
? HTML (Option B): HTML is a markup language for web pages, not a data file format using delimiters.
? JSON (Option C): JSON uses key-value pairs and nested structures, not delimiters like commas.
? CSV (Option D): CSV (Comma-Separated Values) uses delimiters (typically commas) to separate data fields, making it the correct choice.
The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," such as file formats like CSV that use delimiters.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.
==========

NEW QUESTION 63
Which of the following best describes the reason an analyst would reference a data dictionary versus a source's metadata?

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

Answer: B

Explanation:
This question is part of theData Concepts and Environmentsdomain, focusing on the purpose of data documentation tools like data dictionaries and metadata. The question compares their uses.
? To gather information and resources about the data (Option A): This is too vague
and not specific to a data dictionary??s purpose.
? To find the content and specific attributes for a dataset (Option B): A data dictionary provides detailed definitions of data elements (e.g., field names, types, descriptions), which is more specific than metadata, which often includes broader information like creation date or source.
? To find a summary of basic information about the dataset (Option C): This better describes metadata, which provides high-level summaries, not detailed attributes.
? To gather information about the availability of the data (Option D): Neither a data dictionary nor metadata typically focuses on data availability.
The DA0-002 Data Concepts and Environments domain includes understanding "data schemas and dimensions," and a data dictionary is specifically used to find detailed attributes of a dataset.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.
==========

NEW QUESTION 66
A Chief Executive Officer requests a report that must:
• Summarize the company metrics in a simple way.
• Be clear and concise.
• Be easily understood by all company levels.
• Be accessible and updated without manual intervention.
Which of the following communication approaches best meets these requirements?

A. Executive summary
B. Slide presentation
C. Key performance indicator dashboard
D. Open data portal

Answer: C

Explanation:
This question pertains to theVisualization and Reportingdomain, focusing on selecting the appropriate communication method for a report. The requirements emphasize simplicity, clarity, accessibility, and automatic updates, which point to a specific approach.
? Executive summary (Option A): An executive summary is a written document that
summarizes metrics but isn??t typically updated automatically and may not be accessible toall levels without distribution.
? Slide presentation (Option B): A slide presentation can be clear but requires
manual updates and isn??t inherently accessible to all company levels.
? Key performance indicator dashboard (Option C): A KPI dashboard provides a simple, visual summary of metrics, is clear and concise, can be understood by all levels, and can be set up to update automatically, meeting all requirements.
? Open data portal (Option D): An open data portal provides raw data access, which may not be simple or easily understood by all levels.
The DA0-002 Visualization and Reporting domain emphasizes "translating business requirements to form the appropriate visualization," and a KPI dashboard is the best approach for meeting these requirements.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 4.0 Visualization and Reporting.
==========

NEW QUESTION 68
Which of the following data repositories stores unstructured and structured data?

A. Data store
B. Data silo
C. Data mart
D. Data lake

Answer: D

Explanation:
This question falls under theData Concepts and Environmentsdomain of CompTIA Data+ DA0-002, which involves understanding different types of data repositories and their characteristics. The task is to identify a repository that can store both unstructured and structured data.
? Data store (Option A): A data store is a general term for any data repository, but
it??s not specific enough to confirm it stores both unstructured and structured data.
? Data silo (Option B): A data silo is an isolated data repository, often structured, and not typically designed for unstructured data.
? Data mart (Option C): A data mart is a subset of a data warehouse, focused on structured data for specific business areas, not unstructured data.
? Data lake (Option D): A data lake is a centralized repository that stores raw data in its native format, including both structured (e.g., tables) and unstructured (e.g., text, images) data, 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 specifically designed to handle both unstructured and structured data.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 1.0 Data Concepts and Environments.
==========

NEW QUESTION 72
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 theData Analysisdomain, 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 77
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 theData Concepts and Environmentsdomain, 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 78
A data analyst receives a notification that a customized report is taking too long to load. After reviewing the system, the analyst does not find technical or operational issues. Which of the following should the analyst try next?

A. Check that the appropriate filters are applied.
B. Check data source connections.
C. Check for data structure changes in the report.
D. Check whether other peers have the same issue.

Answer: A

Explanation:
This question pertains to theData Governancedomain, focusing on data quality and report performance optimization. The report is slow despite no technical issues, suggesting a data-related inefficiency.
? Check that the appropriate filters are applied (Option A): Applying filters reduces
the dataset size by excluding irrelevant data, improving report performance. This is a logical next step after ruling out technical issues.
? Check data source connections (Option B): The analyst already reviewed the
system and found no operational issues, so connectivity is likely not the problem.
? Check for data structure changes in the report (Option C): While possible, this is a deeper investigation step and less likely to be the immediate cause of slowness.
? Check whether other peers have the same issue (Option D): This might confirm the issue??s scope but doesn??t directly address the performance problem.
The DA0-002 Data Governance domain emphasizes "data quality control concepts," including optimizing report performance through techniques like filtering.
Reference: CompTIA Data+ DA0-002 Draft Exam Objectives, Domain 5.0 Data Governance.
==========

NEW QUESTION 82
.....

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