

DA0-001^{Q&As}

CompTIA Data+

Pass CompTIA DA0-001 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.lead4pass.com/da0-001.html>

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by CompTIA
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers



QUESTION 1

Which of the following best describes the process of examining data for statistics and information about the data?

- A. Cleansing
- B. search
- C. Profiling
- D. Governance

Correct Answer: C

Explanation: Data profiling is the process of examining data for statistics and information about the data, such as the structure, format, quality, and content of the data. Data profiling can help to understand the characteristics, patterns, relationships, and anomalies of the data, as well as to identify and resolve any errors, inconsistencies, or missing values in the data. Data profiling can be done using various tools and methods, such as spreadsheets, databases, or programming languages¹².

QUESTION 2

Which of the following techniques is used to quantify data?

- A. Decoding
- B. Enumeration
- C. Coding
- D. Structure

Correct Answer: C

Answer: C. Coding Coding is a technique that is used to quantify data, especially qualitative data that are not expressed numerically. Coding involves assigning codes, such as numbers, letters, symbols, or colors, to different categories or themes that emerge from the data. For example, if you have a set of survey responses that ask about the satisfaction level of customers, you can code them as follows: Very satisfied = 5 Satisfied = 4 Neutral = 3 Dissatisfied = 2 Very dissatisfied = 1 By coding the data, you can convert them into quantitative data that can be analyzed using statistical methods, such as calculating the mean, median, mode, frequency, or percentage of each category¹². Option A is incorrect, as decoding is not a technique that is used to quantify data, but rather a process of interpreting or translating data from one form to another. For example, decoding can involve converting binary codes into text or images, or decrypting ciphertext into plaintext³. Option B is incorrect, as enumeration is not a technique that is used to quantify data, but rather a process of listing or naming data in a specific order. For example, enumeration can involve listing the names of the states in alphabetical order, or naming the planets in order of their distance from the sun⁴. Option D is incorrect, as structure is not a technique that is used to quantify data, but rather a property or characteristic of data that describes how they are organized or arranged. For example, structure can refer to the format, type, or schema of data, such as structured, semi-structured, or unstructured data.

QUESTION 3

Which of the following differentiates a flat text file from other data types?

- A. Data is separated by a delimiter.
- B. Data is stored in defined rows.
- C. Data is defined with key-value pairs.
- D. Data is housed in a markup language.

Correct Answer: A

Explanation: A flat text file is a type of data file that contains only plain text without any formatting or markup. Data in a flat text file is usually separated by a delimiter, which is a character that marks the boundary between different fields or values. For example, a comma-separated values (CSV) file is a flat text file that uses commas as delimiters. Other common delimiters are tabs, spaces, semicolons, and pipes. Therefore, the correct answer is A. References: Plain text - Wikipedia, Comparison of document markup languages - Wikipedia

QUESTION 4

A financial analyst is creating a daily billing report for a company. One night, the company's data warehouse did not update the data, which caused the data to be reported incorrectly the next day. Which of the following documentation elements should the analyst add to catch this error?

- A. Version number
- B. Data refresh
- C. Frequently asked questions tab
- D. Summary

Correct Answer: B

A data refresh is a documentation element that indicates when the data was last updated or refreshed from the source. A data refresh can help the analyst to catch the error of the data warehouse not updating the data, as it will show a discrepancy between the expected and actual date of the data update. A data refresh can also help the users of the report to verify the timeliness and accuracy of the data, and to avoid making decisions based on outdated or incorrect data

QUESTION 5

A data analyst is designing a dashboard that will provide a story of sales and determine which site is providing the highest sales volume per customer. The analyst must choose an appropriate chart to include in the dashboard. The following data is available: Which of the following types of charts should be considered?

Site	Customers	Sales volume	Average sales per customer
A1	2236	\$3,415,372.00	\$1,527.45
A2	885	\$1,405,437.00	\$1,588.06
A3	333	\$952,723.00	\$2,861.03
B1	483	\$4,871,380.00	\$10,085.67
B2	2969	\$780,381.00	\$262.84
B4	2357	\$4,917,436.00	\$2,086.31
C1	1524	\$1,135,204.00	\$744.88
C2	878	\$614,964.00	\$700.41
C3	1925	\$4,035,100.00	\$2,096.16

- A. Include a line chart using the site and average sales per customer.
- B. Include a pie chart using the site and sales to average sales per customer.
- C. Include a scatter chart using sales volume and average sales per customer.
- D. Include a column chart using the site and sales to average sales per customer.

Correct Answer: D

The best type of chart to display the data is D. Include a column chart using the site and sales to average sales per customer. A column chart is a good choice for comparing categorical data with numerical data, such as the site and sales to average sales per customer. A column chart can show the relative differences between the sites and highlight the site with the highest sales volume per customer. A column chart can also be easily labeled and formatted to make the data clear and understandable. A line chart is not suitable for this data, because it is used to show trends or changes over time, which is not relevant for the site and sales to average sales per customer data. A line chart would also be confusing and misleading, as it would imply a connection or correlation between the sites that does not exist. A pie chart is also not a good choice for this data, because it is used to show the proportion of a whole, not the comparison of different categories. A pie chart would also be difficult to read and interpret, as it would require labels or legends to identify the sites and their sales to average sales per customer. A pie chart would also not be able to show the exact values of the sales to average sales per customer, only their relative sizes. A scatter chart is another inappropriate option for this data, because it is used to show the relationship or correlation between two numerical variables, not between a categorical and a numerical variable. A scatter chart would also be cluttered and unclear, as it would plot each site as a point on a coordinate plane, without any labels or axes. A scatter chart would also not be able to show the differences or rankings between the sites and their sales to average sales per customer.

QUESTION 6

Which one of the following values will appear first if they are sorted in descending order?

- A. Aaron.
- B. Molly.

C. Xavier.

D. Adam.

Correct Answer: C

Explanation: The value that will appear first if they are sorted in descending order is Xavier. Descending order means arranging values from the largest to the smallest, or from the last to the first in alphabetical order. In this case, Xavier is the last name in alphabetical order, so it will appear first when sorted in descending order. The other names will appear in the following order: Molly, Adam, Aaron. Reference: Sorting Data - W3Schools

QUESTION 7

Angela is aggregating data from CRM system with data from an employee system.

While performing an initial quality check, she realizes that her employee ID is not associated with her identifier in the CRM system.

What kind of issues is Angela facing?

Choose the best answer.

A. ETL process.

B. Record linkage.

C. ELT process.

D. System integration.

Correct Answer: B

While this scenario describes a system integration challenge that can be solved with ETL or ELT, Angela is facing a Record linkage issue.

QUESTION 8

Samantha needs to share a list of her organization's top 50 customers with the VP of sales.

She would like to include the name of the customer, the business they represent, their contact information, and their total sales over the past year. The VP does not have any specialized analytics skills or software but would like to make some personal notes on the dataset.

What would be the best tool for Samantha to use to share this information?

A. Power BI.

B. Microsoft Excel.

C. Minitab.

D. SAS.

Correct Answer: B

Microsoft Excel.

This scenario presents a very simple use case where the business leader needs a dataset in an easy-to-access form and will not be performing any detailed analysis. A simple spreadsheet, such as Microsoft Excel, would be the best tool for

this job. There is no need to use a statistical analysis package, such as SAS or Minitab, as this would likely confuse the VP without adding any value. The same is true of an integrated analytics suite, such as Power BI.

QUESTION 9

Given the following graph: Which of the following summary statements upholds integrity in data reporting?



- A. Sales are approximately equal for Product A and Product B across all strategies.
- B. Strategy 4 provides the best sales in comparison to other strategies.
- C. While Strategy 2 does not result in the highest sales of Product D, over all products it appears to be the most effective.
- D. Product D should be promoted more than the other products in all strategies.

Correct Answer: B

Explanation: Strategy 4 provides the best sales in comparison to other strategies. This is because the total sales for Strategy 4 are the highest among all the strategies, as shown by the black line. The other statements are not accurate or do not uphold integrity in data reporting. Here is why:

Statement A is false because sales are not approximately equal for Product A and Product B across all strategies. For example, in Strategy 1, Product A has more sales than Product B, while in Strategy 3, Product B has more sales than Product A. Statement C is misleading because it does not account for the difference in scale between the products. While Strategy 2 has the highest total sales among all products, it does not necessarily mean that it is the most effective for each product. For instance, Product D has very low sales in Strategy 2 compared to other strategies. Statement D is biased because it does not provide any evidence or justification for why Product D should be promoted more than the other products in all strategies. It also ignores the fact that Product D has the lowest sales among all products in most of

the strategies.

QUESTION 10

Kelly wants to get feedback on the final draft of a strategic report that has taken her six months to develop.

What can she do to get prevent confusion as see seeks feedback before publishing the report?

Choose the best answer.

- A. Distribute the report to the appropriate stakeholders via email.
- B. Use a watermark to identify the report as a draft.
- C. Show the report to her immediate supervisor.
- D. Publish the report on an internally facing website.

Correct Answer: B

The best answer is to use a watermark to identify the report as a draft. A watermark is a faint image or text that appears behind the content of a document, indicating its status or ownership. By using a watermark, Kelly can clearly communicate that the report is not final and still subject to changes or feedback. This can prevent confusion among the readers and avoid any misuse or misinterpretation of the report. The other options are not as effective as using a watermark, as they either do not indicate the status of the report or do not reach the appropriate stakeholders. Distributing the report via email or publishing it on an internally facing website may not make it clear that the report is a draft and may cause confusion or errors. Showing the report to her immediate supervisor may not get enough feedback from other relevant stakeholders who may have different perspectives or insights. Reference: How to Add a Watermark in Microsoft Word - Lifewire

QUESTION 11

A data set was recorded using multimedia technology. Which of the following is a necessary step on the way to interpretation?

- A. Structural equation modeling
- B. Transcription
- C. Sequential analysis
- D. Sampling

Correct Answer: B

The correct answer is B. Transcription.

Transcription is a necessary step on the way to interpretation when a data set was recorded using multimedia technology. Multimedia technology refers to the use of various forms of media, such as audio, video, images, and text, to capture

and present information¹ Transcription is the process of converting multimedia data into written or textual form, which can then be analyzed using various methods and tools² Transcription can help to make the data more accessible,

searchable, and manageable, as well as to preserve the data for future use.

Structural equation modeling is not correct, because it is a statistical technique that tests the causal relationships between multiple variables using observed and latent variables. Structural equation modeling is not a necessary step on the

way to interpretation, but rather an optional method that can be applied to certain types of data. Sequential analysis is not correct, because it is a method of analyzing the order and timing of events or behaviors in a data set. Sequential

analysis is not a necessary step on the way to interpretation, but rather an optional method that can be applied to certain types of data. Sampling is not correct, because it is the process of selecting a subset of data from a larger population for

analysis. Sampling is not a necessary step on the way to interpretation, but rather a preliminary step that can be done before collecting or analyzing the data.

QUESTION 12

Joseph is interpreting a left skewed distribution of test scores. Joe scored at the mean, Alfonso scored at the median, and Gaby scored at the end of the tail.

Who had the highest score?

- A. Joseph
- B. Joe
- C. Alfonso
- D. Gaby

Correct Answer: C

Alfonso had the highest score. A left skewed distribution is a distribution where the tail is longer on the left side than on the right side, meaning that most of the values are clustered on the right side and there are some outliers on the left side.

In a left skewed distribution, the mean is less than the median, which is less than the mode. Therefore, Joseph, who scored at the mean, had the lowest score, Gaby, who scored at the end of the tail, had the second lowest score, and

Alfonso, who scored at the median, had the highest score.

Reference: Skewness - Statistics How To

QUESTION 13

A data analyst is creating a report that will provide information about various regions, products, and time periods. Which of the following formats would be the most efficient way to deliver this report?

- A. A workbook with multiple tabs for each region
- B. A daily email with snapshots of regional summaries

- C. A static report with a different page for every filtered view
- D. A dashboard with filters at the top that the user can toggle

Correct Answer: D

Explanation: The best format to deliver this report is D. A dashboard with filters at the top that the user can toggle. A dashboard is a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information can be monitored at a glance¹ A dashboard with filters at the top that the user can toggle would allow the user to easily and quickly access the information they need about various regions, products, and time periods, without having to navigate through multiple tabs, pages, or emails. A dashboard with filters would also enable the user to compare and contrast different views of the data and see how they change over time. A dashboard with filters would also be more interactive and engaging than a static or email report² A workbook with multiple tabs for each region would not be an efficient way to deliver this report, because it would require the user to switch between different tabs to see the information they need. This would make it harder to compare and contrast different regions, products, and time periods, and also increase the risk of errors or confusion. A workbook with multiple tabs would also be less visually appealing and more cluttered than a dashboard³ A daily email with snapshots of regional summaries would not be an efficient way to deliver this report, because it would limit the user's ability to explore the data in depth and customize their view. A daily email would also be dependent on the frequency and timing of the email delivery, which might not match the user's needs or preferences. A daily email would also be more likely to be ignored or deleted than a dashboard that is always accessible. A static report with a different page for every filtered view would not be an efficient way to deliver this report, because it would create a very long and cumbersome report that would be difficult to read and understand. A static report would also not allow the user to change or update the filters as they wish, or see how the data changes over time. A static report would also be less interactive and engaging than a dashboard.

QUESTION 14

Which of the following are reasons to create and maintain a data dictionary? (Choose two.)

- A. To improve data acquisition
- B. To remember specifics about data fields
- C. To specify user groups for databases
- D. To provide continuity through personnel turnover
- E. To confine breaches of PHI data
- F. To reduce processing power requirements

Correct Answer: BD

A data dictionary is a collection of metadata that describes the data elements in a database or dataset. It can help improve data acquisition by providing information about the data sources, formats, quality, and usage. It can also help remember specifics about data fields, such as their names, definitions, types, sizes, and relationships. Therefore, options B and D are correct.

Option A is incorrect because it is not a reason to create and maintain a data dictionary, but a benefit of doing so.

Option C is incorrect because specifying user groups for databases is not a function of a data dictionary, but a function of a database management system or a security policy.

Option E is incorrect because confining breaches of PHI data is not a function of a data dictionary, but a function of a

data protection or encryption system. Option F is incorrect because reducing processing power requirements is not a function of a data dictionary, but a function of a data compression or optimization system.

QUESTION 15

A data analyst needs to create a weekly recurring report on sales performance and distribute it to all sales managers. Which of the following would be the BEST method to automate and ensure successful delivery for this task?

- A. Use scheduled report delivery.
- B. Implement subscription access delivery.
- C. Print out a copy.
- D. Upload the report to the server.

Correct Answer: A

Explanation: Scheduled report delivery is a feature that allows a data analyst to automate the generation and distribution of a report at a specified time and frequency. This would be the best method to ensure that the sales managers receive the weekly report on sales performance without manual intervention. Subscription access delivery is a feature that allows users to subscribe to a report and access it on demand, but it does not automate the delivery. Printing out a copy or uploading the report to the server are manual methods that require more time and effort from the data analyst.

Reference: CertMaster Practice for Data+ Exam Prep - CompTIA

[Latest DA0-001 Dumps](#)

[DA0-001 VCE Dumps](#)

[DA0-001 Study Guide](#)