



CBDA^{Q&As}

Certification in Business Data Analytics (IIBA - CBDA)

Pass IIBA CBDA Exam with 100% Guarantee

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

<https://www.pass4itsure.com/cbda.html>

100% Passing Guarantee
100% Money Back Assurance

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

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





QUESTION 1

A lab is conducting a study on protein interactions. They have used the data to create a graph visualization. In graph visualization, what would an edge represent?

- A. A single datapoint
- B. A link between two datapoints
- C. A collection of datapoints and links
- D. A dedicated algorithm that calculates the node positions

Correct Answer: B

A graph visualization is a type of visualization that shows the relationships among data points by using nodes (or vertices) to represent the data points and edges (or links) to represent the connections between them¹. A graph visualization can help reveal patterns, clusters, outliers, or hierarchies in the data². In a graph visualization, an edge represents a link between two data points, indicating that they have some kind of association, interaction, similarity, or dependency³. For example, in a study on protein interactions, an edge could represent a physical or functional interaction between two proteins, such as binding, signaling, or regulation⁴.

A single data point, a collection of data points and links, and a dedicated algorithm that calculates the node positions are not correct definitions of an edge in a graph visualization. A single data point is represented by a node, not an edge, in a graph visualization. A collection of data points and links is the whole graph, not an edge, in a graph visualization. A dedicated algorithm that calculates the node positions is a method of graph layout, not an edge, in a graph visualization. A graph layout is the way the nodes and edges are arranged in a graph visualization, which can affect the readability, aesthetics, and interpretation of the graph.

QUESTION 2

A manufacturing company, specializing in turf maintenance equipment, has recently seen a decline in their lawn mower sales. As a result, the analytics team is asked to review the latest customer satisfaction survey results. An analyst on this team creates a report for senior management with attractive visuals, supported by the KPI results. Upon reviewing the report, the analyst's manager mentions that the report is missing the narrative. What does this mean?

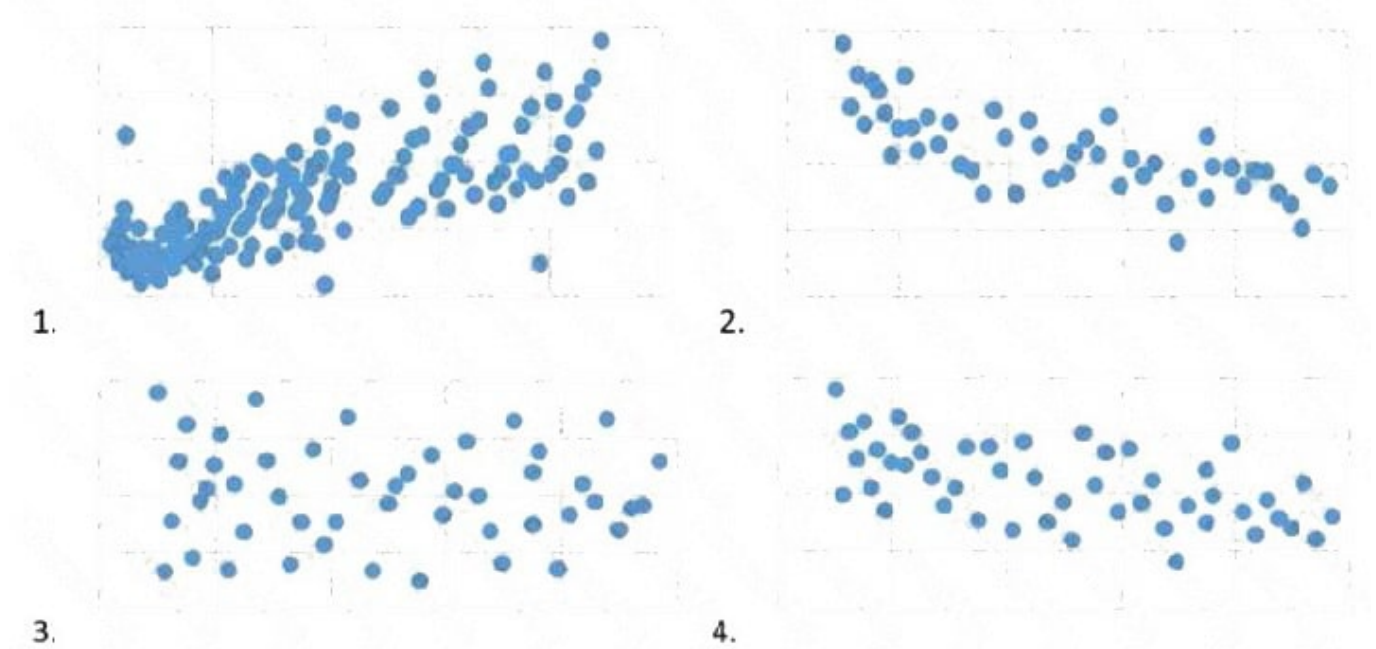
- A. The data tables that support the visuals and help answer questions
- B. A narrative that supports insights with additional context and draws correlations
- C. Notes on assumptions and unavailable data for analysis
- D. Commentary around why each graphic was selected to provide additional context

Correct Answer: B

A narrative is a written or spoken explanation of the data analysis results that tells a story with the data, provides additional context and background information, highlights the key insights and findings, and draws correlations and implications for the decision makers¹². The report is missing the narrative, meaning that it does not communicate the meaning and value of the data analysis effectively, and it leaves the interpretation and action to the senior management without any guidance or recommendation

**QUESTION 3**

A data scientist is analyzing a dataset to determine if there is a strong relationship between two variables. A measure of covariance is done. Which of the following graphs indicate Zero Covariance between variables?



A. 3

B. 1

C. 4

D. 2

Correct Answer: C

In the context of Business Data Analytics (IIBA - CBDA), zero covariance between two variables indicates that there is no linear relationship between those variables. When the covariance is zero, it means the variables are independent of each other. In the provided options, graph 4 shows a random scatter of data points without any apparent trend or pattern, indicating zero covariance.

QUESTION 4

A data scientist at a consumer goods company, has been asked to do a detailed analysis on customer profiles. The Data Scientist has identified an external data source that carries valuable additional information on their customers. The data scientist also identifies the address column as the most reliable column to join the internal data source with the external data source. Addresses may appear in different formats for example:

File A = "13 Smith St"

File B = "Unit 7, 13 Smith Street"

Which of the following techniques would be useful in this situation?



- A. Deterministic linkage
- B. Probabilistic linkage
- C. Genetic linkage
- D. Cuff linkage

Correct Answer: B

Probabilistic linkage is a technique that uses statistical methods to match records from different data sources based on the similarity of key variables, such as name, address, date of birth, etc¹. Probabilistic linkage can handle variations, errors, or missing values in the data, and assign a score or probability to each potential match². Probabilistic linkage would be useful in this situation, as the address column may have different formats, spellings, or abbreviations in the internal and external data sources, and a deterministic linkage (which requires exact matches) might miss some valid matches or create false matches.

Deterministic linkage is a technique that uses predefined rules or criteria to match records from different data sources based on the exact agreement of key variables, such as identifiers, codes, or hashes³. Deterministic linkage would not be useful in this situation, as the address column may not have consistent or unique values in the internal and external data sources, and a probabilistic linkage (which allows for some variation or uncertainty) might find more accurate matches or avoid false matches.

Genetic linkage is a term used in genetics to describe the tendency of genes or DNA sequences that are located close together on a chromosome to be inherited together⁴. Genetic linkage is not relevant to this situation, as it has nothing to do with matching records from different data sources based on the address column.

Cuff linkage is a term used in sewing to describe the process of attaching a cuff to a sleeve by stitching or fastening. Cuff linkage is not relevant to this situation, as it has nothing to do with matching records from different data sources based on the address column.

QUESTION 5

With the recent departure of two of its employees, an IT helpdesk team is now understaffed and finding it difficult to keep up with the current workload. The number of tickets being received has increased as well as the number of days to resolve the tickets. The IT manager has set up a meeting with the IT director to request funding for two new helpdesk agents. To prepare for the meeting, the manager is interested in showing the tickets processed against ticket volume over the past year. What type of chart should the manager use to effectively show the change in processing rate over time?

- A. A pie chart to compare the number of tickets coming in versus tickets being processed each month, over the past year
- B. A column chart to compare the number of tickets coming in versus tickets being processed each month, since June
- C. A line chart to show the widening gap between the number of tickets being processed against the number coming over the past year
- D. A waterfall chart to show the number of tickets coming in are a lot higher than those being processed as of year to date

Correct Answer: C

A line chart is the type of chart that the manager should use to effectively show the change in processing rate over time, because it is a technique that displays data as a series of points connected by straight lines. A line chart can help the



manager visualize the trends and patterns in the ticket volume and processing rate over the past year, and highlight the widening gap between them. A line chart can also show the seasonal variations and fluctuations in the data, and compare the performance of different categories or groups. Options A, B, and D are not suitable for showing the change in processing rate over time, because they are techniques that display data as proportions (A), comparisons (B), or accumulations (D) of different categories or groups at a single point in time or over a fixed period.

QUESTION 6

A government agency is conducting a study on the performance of 12th grade students in mathematics across the country. In particular, they want to understand if there is a relationship between intelligence and scores, as well as the difference in performance between various locations. Which combination of inferential statistics procedures should be used?

- A. Range, standard deviation
- B. Mean, median
- C. Correlation co-efficient, analysis of variance
- D. Frequency distribution, time-series

Correct Answer: C

A correlation co-efficient is a measure of the strength and direction of the linear relationship between two variables, such as intelligence and scores. A correlation co-efficient can range from -1 to 1, where -1 indicates a perfect negative relationship, 0 indicates no relationship, and 1 indicates a perfect positive relationship¹². An analysis of variance (ANOVA) is a procedure that tests whether the means of two or more groups are significantly different from each other, such as the performance of students across various locations. ANOVA can compare the variation within each group and the variation between groups to determine if there is a statistically significant difference among the group means³⁴.

QUESTION 7

The analytics team is struggling with which recommendation to make. Their challenge is that they have five good options and this indecision is stopping them from moving forward. To help the team finalize their recommendation, the BA professional on the team recommends they complete:

- A. Root cause analysis
- B. Business rules analysis
- C. Data flow diagrams
- D. Acceptance and evaluation criteria

Correct Answer: D

Acceptance and evaluation criteria are the techniques that the BA professional on the team should recommend they complete, because they are the standards or measures that are used to evaluate the suitability and value of each option. Acceptance and evaluation criteria can help the team compare the benefits, costs, risks, and impacts of each option, and determine which one best meets the needs and expectations of the stakeholders. Acceptance and evaluation criteria can also help the team communicate the rationale and evidence behind their recommendation, and ensure that the recommendation is aligned with the business goals and objectives.

**QUESTION 8**

A research marketer is interested in collecting information about the spending habits of families in North America. Concerned about the volume of data required to conduct the research, they choose to use sampling. The dataset is sourced using all credit card transactions from a leading North American credit card company for Quarter 1 of the prior year. The sample used is:

- A. Statistically representative
- B. Not relevant
- C. Too large to be helpful
- D. Biased

Correct Answer: D

The sample used in this case is biased, meaning that it is not representative of the population of interest. The population of interest is the families in North America, but the sample is drawn from only one source of data: the credit card transactions from a leading North American credit card company. This sample excludes the families who do not use credit cards, or who use other credit card companies, or who use other payment methods. Therefore, the sample is not random or fair, and it may introduce sampling bias into the research results

QUESTION 9

The team has completed their analysis on a vast amount of collected data and agree on their recommendations for action.

However, they are having difficulty in developing the appropriate messages to support their recommendations. The business analysis professional suggests which technique to assist the team?

- A. T-Testing
- B. Simulation
- C. Visioning
- D. Storyboarding

Correct Answer: D

Storyboarding is a technique that helps the team to develop the appropriate messages to support their recommendations by creating a visual sequence of the main points, evidence, and actions. Storyboarding helps the team to organize their thoughts, identify gaps, and communicate their findings in a clear and compelling way

QUESTION 10

The analytics team is identifying research questions to address a business problem. The business analysis professional reminds the team that the most important dimension to consider is the:

- A. Sources of data



- B. Quality of the data
- C. Timeframe of analysis
- D. Measurement scale

Correct Answer: B

The quality of the data is the most important dimension to consider when identifying research questions, as it affects the validity, reliability, and accuracy of the analysis and the results. Data quality refers to the degree to which the data meets the requirements and expectations of the stakeholders and the purpose of the analysis¹². Poor data quality can lead to erroneous conclusions, ineffective decisions, and wasted resources

QUESTION 11

An analytics team is sourcing data for a new analytics initiative and is deciding between two comparable data sources. One source being considered is a very large dataset and another consists of three smaller sources. What advantage will the larger dataset provide over the three smaller sources?

- A. More significant results
- B. Higher validity
- C. More reproducibility
- D. Higher reliability

Correct Answer: A

A larger dataset may provide more significant results than three smaller sources, as it may have more statistical power to detect differences or relationships among variables¹. Statistical power is the probability of finding a statistically significant result when there is a true effect in the population². A larger dataset may have more power because it may have more variability, less sampling error, and higher precision than smaller datasets³. More significant results may lead to more confident and valid conclusions and recommendations for the analytics initiative.

Higher validity, more reproducibility, and higher reliability are not necessarily advantages of a larger dataset over three smaller sources, as they depend on other factors besides the size of the data. Validity is the degree to which the data and the analysis measure what they are intended to measure⁴. Reproducibility is the degree to which the data and the analysis can be replicated by another analyst using the same methods and data sources. Reliability is the degree to which the data and the analysis produce consistent results under the same conditions. These qualities may be affected by the quality, accuracy, completeness, and relevance of the data, as well as the appropriateness, transparency, and rigor of the analysis methods. A larger dataset may not be valid, reproducible, or reliable if it has errors, biases, missing values, or irrelevant variables, or if the analysis methods are not suitable, documented, or verified.

QUESTION 12

The finance manager has reported that customers are taking much longer to remit payments this year than last. They would like help in finding a solution to address the situation. One suggestion was to offer a 10% discount to entice customers to pay their invoices in full within the first 30 days. Before offering the discount, the finance manager would like the analytics team to do some research to determine if there is value in addressing the accounts receivable problem. Which of the following is a valid question to ask in this situation?

- A. Have discounts been offered before?



- B. Are sales decreasing when accounts receivables are increasing?
- C. How does credit score impact the customer's ability to pay?
- D. Should the discount offered be set at 10% or 15%?

Correct Answer: A

According to the Guide to Business Data Analytics, one of the steps in conducting business data analytics is to identify the research questions that will guide the analysis and help answer the business problem or opportunity. The research questions should be relevant, specific, measurable, achievable, and testable. In this situation, the business problem is the delay in customer payments and the potential solution is to offer a discount. A valid question to ask in this situation is whether discounts have been offered before, and if so, what was the effect on customer behavior and profitability. This question is relevant because it can help assess the feasibility and effectiveness of the proposed solution. It is also specific, measurable, achievable, and testable, as it can be answered by collecting and analyzing historical data on customer payments and discounts..

QUESTION 13

What is the relationship between a Customer entity and an Order entity, where a customer entry will be present in the Customer entity regardless of whether an order was made?

- A. zero-to-one
- B. many-to-many
- C. zero-to-many
- D. one-to-one

Correct Answer: C

A zero-to-many relationship between two entities means that one instance of the first entity can be associated with zero or more instances of the second entity, and one instance of the second entity can be associated with only one instance of the first entity¹. In this case, a customer entry will be present in the Customer entity regardless of whether an order was made, which means that a customer can have zero or more orders, but an order can only belong to one customer. Therefore, the relationship between Customer and Order is zero-to-many.

QUESTION 14

A clinical research organization is using predictive analytics to improve patient safety and decrease costs on its clinical trials. To ensure that a standard set of tools/techniques is identified and best practices adhered to, teams are required to create scenarios to generate appropriate data for initial analysis. This practice is required because it is almost certain that data will be difficult to come by for most research. Which concern would lead the team to establish scenario development as a required technique?

- A. Data validity
- B. Data privacy
- C. Data reliability
- D. Data reproducibility



Correct Answer: A

Data validity refers to the extent to which data accurately represents the phenomenon or concept that it is intended to measure¹. Data validity is essential for predictive analytics, as it affects the quality and credibility of the analysis results and the subsequent decisions or actions based on them. If data is invalid, the predictions may be inaccurate, misleading, or irrelevant. However, data validity may be challenging to ensure in clinical research, as data may be scarce, incomplete, inconsistent, or subject to errors or biases². Therefore, the team may establish scenario development as a required technique to address this concern. Scenario development is a form of document analysis that involves creating hypothetical situations or stories based on assumptions, evidence, and logic to explore the possible outcomes or implications of a problem or opportunity³. Scenario development can help the team generate appropriate data for initial analysis by simulating different conditions, variables, or events that may affect the clinical trials, and by testing the validity of the data against the scenarios⁴.

QUESTION 15

The analytics team has established two equally strong potential recommendations which will deliver the desired outcomes with similar benefits to be derived from each one. On the surface there is no discernable difference in costs or schedule for either option. To help the analytics team reach a recommendation the business analysis professional recommends the team:

- A. Complete market research
- B. Assess risks for each option
- C. Vote to choose the recommendation
- D. Seek management guidance

Correct Answer: B

Assessing risks for each option is the recommendation that the business analysis professional should make to the analytics team, because it is a technique that involves identifying, analyzing, and evaluating the potential positive or negative impacts of each option on the project, the organization, or the stakeholders. Assessing risks can help the team compare the pros and cons of each option, and determine which one has the highest expected value or the lowest expected loss. Assessing risks can also help the team prepare contingency plans or mitigation strategies for the chosen option, and communicate the rationale and assumptions behind their recommendation

[Latest CBDA Dumps](#)

[CBDA Study Guide](#)

[CBDA Exam Questions](#)