Number: CBDA
Passing Score: 800.0
Time Limit: 120.0
File Version: 1.0

Exam Code: CBDA

Exam Name: Certification in Business Data Analytics

Website: www.Vdumps.com

Exam A

QUESTION 1
An analyst is interested in providing a visual diagram to compare and contrast the characteristics of four different solution options. Each option should be represented by their cost, value, and risk level. What type of chart would accomplish this task?

A. Bubble
B. Waterfall
C. Pie
D. Bullet

Correct Answer: A
Section:
Explanation:
A bubble chart is a type of chart that displays three dimensions of data: the x-axis, the y-axis, and the size of the bubble. A bubble chart can be used to compare and contrast the characteristics of different solution options by plotting their cost, value, and risk level on the three axes. For example, a solution option with a high cost, high value, and low risk would be represented by a large bubble on the upper left corner of the chart, while a solution option with a low cost, low value, and high risk would be represented by a small bubble on the lower right corner of the chart. A bubble chart can help the analyst and the stakeholders to visualize the trade-offs and benefits of each solution option and to select the most optimal one based on the business objectives and constraints.
Reference:Guide to Business Data Analytics, page 77;Introduction to Business Data Analytics: A Practitioner View, page 16; [Business Data Analytics: A Practical Guide], page 121.

QUESTION 2
A movie production company wants to use analytics to decide which customers would choose to watch or not watch a particular movie after seeing a promotional teaser. The business analysis professional suggests they could make that prediction by identifying characteristics of the new movie and determining if the customer has watched other movies with similar characteristics. This is an example of using the following technique:

A. Logistic regression
B. Ouster analysis
C. Integer programming
D. Analysis of variance

Correct Answer: A
Section:
Explanation:
Explanation: Logistic regression is a technique that can be used to model the probability of a binary outcome, such as choosing to watch or not watch a movie, based on one or more predictor variables, such as the characteristics of the movie and the customer's viewing history. Logistic regression can help the business analysis professional to identify the factors that influence the customer's decision and to estimate the likelihood of each customer's preference. Logistic regression can also be used to test hypotheses and to evaluate the performance of the predictive model.
Reference: [Guide to Business Data Analytics], page 55; [Business Data Analytics: A Practical Guide], page 93; [Introduction to Business Data Analytics: A Practitioner View], page 14.

QUESTION 3
An analytics team has completed some initial data analysis but is considering revising their research question based on their analysis findings. The team was concerned the original question was too broad. What outcome would lead the team to have this concern?

A. Data once analyzed had significant data quality issues
B. Data the team had planned to use was not available
C. Difficult to identify the KPIs to measure
D. The source data sets could not be merged

Correct Answer: C
Section:
Explanation:
A research question is a clear and focused question that guides the data analytics process and defines the expected outcome or value of the analysis1. A research question that is too broad may lead to the concern of being difficult to identify the key performance indicators (KPIs) to measure, as KPIs are specific, quantifiable, and relevant metrics that indicate the progress and success of the analysis in relation to the research question23. A broad research question may also result in too much or too little data, unclear or conflicting objectives, or irrelevant or ambiguous results4.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 202: Guide to Business Data Analytics, IIBA, 2020, p. 233: Key Performance Indicators: Developing, Implementing, and Using Winning KPIs, David Parmenter, 2015, p. 34: How to Write a Good Research Question, ThoughtCo, 2021, 1.

QUESTION 4
While sourcing data, an analyst runs into a situation where different business units are using different names to refer to the same data element. This lack of standardization is resulting in confusion and additional time required to properly prepare data for analysis. Which practice, if implemented would address this situation and mature the organization's business analytics practice?

A. Data quality management
B. Database operations management
C. Data warehousing
D. Meta data management

Correct Answer: D
Section:
Explanation:
Meta data management is the practice that, if implemented, would address the situation and mature the organization's business analytics practice, because it is a technique that involves defining, documenting, and maintaining the information about the data elements, such as their names, definitions, formats, sources, and relationships. Meta data management can help the analyst resolve the inconsistencies and ambiguities in the data element names, and ensure that the data is standardized, consistent, and understandable across different business units. Meta data management can also help the analyst improve the data quality, accessibility, and usability for the analysis.
Reference:
* Business Analysis Certification in Data Analytics, CBDA | IIBA, CBDA Competencies, Domain 2: Source Data
* Guide to Business Data Analytics - Iiba - Google Books, page 14
* Business Data Analytics (IIBA-CBDA Exam preparation) | Udemy, Section 2: Source Data, Lecture 8: Meta Data Management

QUESTION 5
A dataset contains 10 measures of workplace sustainability. The analytics team is in need of producing a single score of sustainability. Which of the following techniques if used would achieve this objective?

A. Logistic regression
B. Linkage algorithms
C. Factor analysis
D. K means clustering

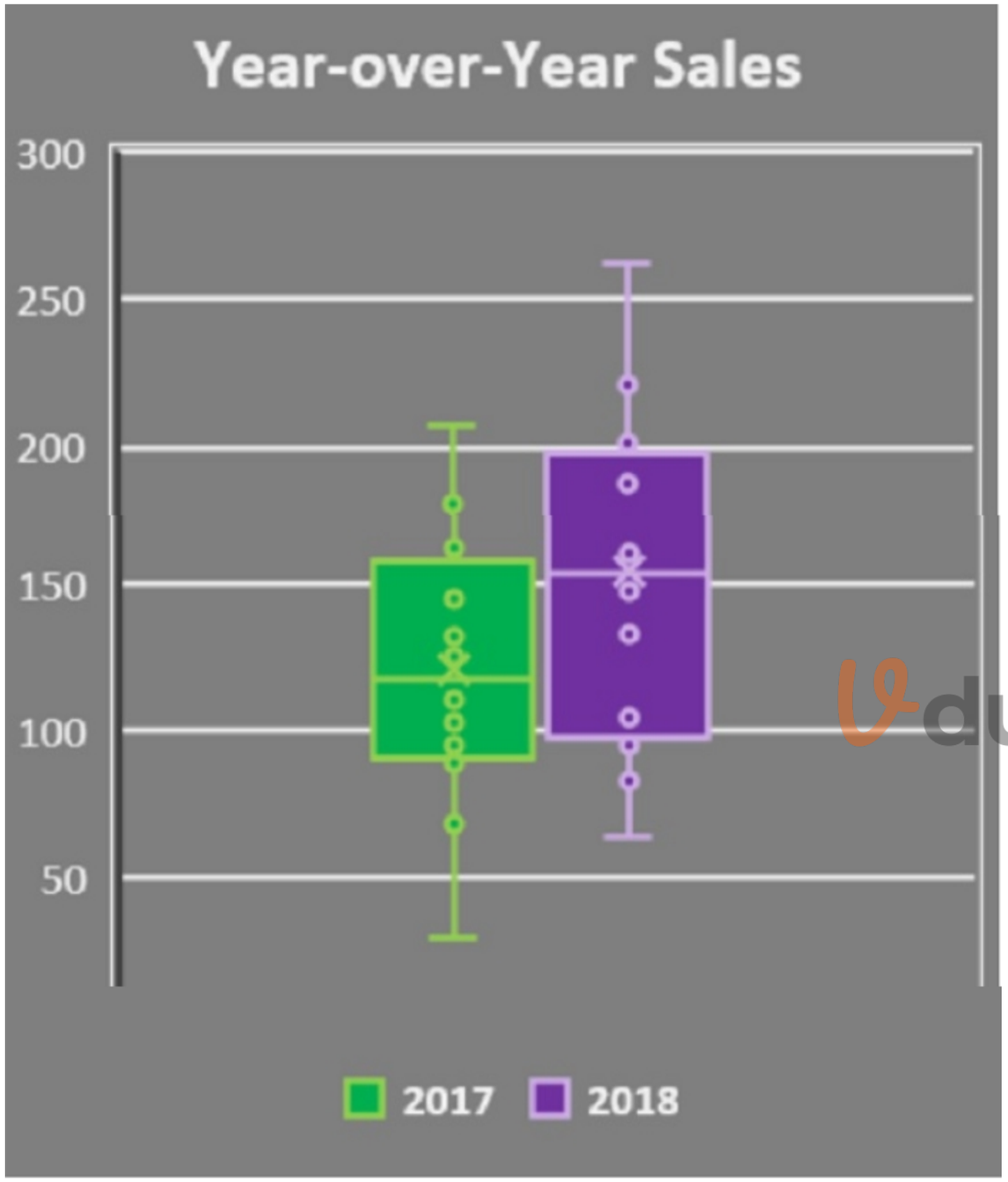
Correct Answer: C
Section:
Explanation:
Factor analysis is the technique that, if used, would achieve the objective of producing a single score of sustainability, because it is a technique that reduces the dimensionality of a data set by identifying the underlying factors or latent variables that explain the variation and correlation among the observed variables. Factor analysis can help the analytics team combine the 10 measures of workplace sustainability into a smaller number of factors, and then derive a composite score of sustainability based on the factor loadings and weights. Factor analysis can also help the analytics team simplify and interpret the data, and identify the key drivers of sustainability.
Reference:
* Business Analysis Certification in Data Analytics, CBDA | IIBA, CBDA Competencies, Domain 3: Analyze Data
* Understanding the Guide to Business Data Analytics, page 17
* Business Data Analytics (IIBA-CBDA Exam preparation) | Udemy, Section 3: Analyze Data, Lecture 15: Factor Analysis

QUESTION 6
An analyst is looking at a particular dataset that includes the scores across all 8th grade students, across three schools. The analyst is trying to determine which type of statistics average to use to best represent the results. On looking through the dataset, the analyst has identified a few extreme outliers. As a result, the analyst was led to use the following type of average:

A. Median
B. Range
C. Mean
D. Mode

Correct Answer: A
Section:
Explanation:
The median is the type of statistics average that the analyst should use to best represent the results, because it is a measure of central tendency that divides the data set into two equal halves. The median is the middle value of the data set when it is arranged in ascending or descending order. The median is not affected by extreme outliers, unlike the mean, which is the arithmetic average of the data set. The median can give a more accurate representation of the typical score of the 8th grade students across the three schools. Options B, C, and D are not types of statistics average, but types of statistics measures that describe other aspects of the data set. The range is a measure of dispersion that shows the difference between the highest and the lowest values of the data set. The mean is a measure of central tendency that shows the sum of the values of the data set divided by the number of values. The mode is a measure of central tendency that shows the most frequent value of the data set.
Reference:
* Business Analysis Certification in Data Analytics, CBDA | IIBA, CBDA Competencies, Domain 3: Analyze Data
* Understanding the Guide to Business Data Analytics, page 17
* Business Data Analytics (IIBA-CBDA Exam preparation) | Udemy, Section 3: Analyze Data, Lecture 13: Descriptive Statistics

QUESTION 7
A software company launched a new product in late 2016. The product manager is reviewing a Box and Whisker plot used to compare year-over-year sales, from 2017 to 2018. What is the conclusion he can make from this chart?



A. 2017 minimum and maximum sales are higher than 2018, and the 2017 median result is higher than the 2018 median result
B. 2017 minimum and maximum sales are higher than 2018, but the 2017 median result is lower than 2018 1st quartile result
C. 2018 minimum and maximum sales are higher than 2017, and the 2018 quartile results are higher than 2017 quartile results
D. 2018 minimum and maximum sales are higher than 2017, and the 2018 1st quartile is higher than 2017 median result

Correct Answer: D

Section:

QUESTION 8
The interplay between enterprise systems and data analytics can be envisioned at various layers. The layer that connects the business processes to data analytics is the:

A. information layer
B. physical layer
C. technical layer
D. infrastructure layer

Correct Answer: A
Section:
Explanation:
The information layer is the layer that connects the business processes to data analytics. It consists of the data models, data quality, data governance, and data security that enable the data to be accessed, analyzed, and transformed into insights. The information layer also supports the communication and collaboration among the stakeholders involved in the data analytics process. The other layers are the physical layer, which deals with the hardware and software components of the data infrastructure; the technical layer, which handles the data integration, data storage, data processing, and data analysis techniques; and the infrastructure layer, which provides the network, cloud, and security services for the data environment12
Reference: 1: Data and Analytics (D&A) - Gartner 2: Enterprise Data Analytics - SelectHub

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
Section:
Explanation:
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 way12
Reference: 1: Developing Key Messages for Effective Communication - MSKTC 2: 11 Ways Highly Successful Leaders Support Their Team - Redbooth

QUESTION 10
A large number of text messages are received by Twitter each year making Twitter one example of Big Data. What data characteristic represents this large number of text messages?

A. Veracity
B. Velocity
C. Value
D. Variety

Correct Answer: B
Section:
Explanation:
Velocity is one of the four V's of Big Data, along with Volume, Variety, and Veracity. Velocity refers to the speed at which data is generated, collected, and processed. A large number of text messages received by Twitter each year is an example of high-velocity data, as it requires real-time or near-real-time processing and analysis to extract insights and value from it. High-velocity data poses challenges and opportunities for business data analytics,

as it requires efficient and scalable data infrastructure, streaming analytics, and timely decision-making.

QUESTION 11
Allegra Consulting is planning on establishing an analytics system to track career progression of their consultants. Elicitation will be used to identify the required features. How would brainstorming be used to prepare for elicitation?

A. To identify sources of business information to consider
B. To identify the key metrics to be collected
C. Determine the value for establishing the analytics system
D. To choose the statistical methods required

Correct Answer: A
Section:
Explanation:
Explanation: According to the Guide to Business Data Analytics, one of the tasks under the domain of ''Identify the Research Questions'' is to identify sources of business information to consider. This task involves reviewing existing business information, such as documents, reports, databases, and systems, to determine what data is available, relevant, and reliable for answering the research questions. This task also involves identifying any gaps or limitations in the existing information and proposing ways to address them.

QUESTION 12
A job satisfaction study is being considered. Half of the employees of the company will be interviewed by senior managers and the other half of the employees will be interviewed by an external market research company, using the same set of questions. Which of the following might be a concern for using this approach to collect study data?

A. Reliability
B. Validity
C. Timeliness
D. Precision

Correct Answer: A
Section:
Explanation:
Explanation: Reliability is the degree to which a data collection method produces consistent results under the same conditions1. In this case, the reliability of the study data might be compromised by the different interviewers (senior managers vs. external market research company), who might have different biases, expectations, or rapport with the employees. This could affect how the employees respond to the same set of questions, and thus introduce variability in the dat a.Validity, timeliness, and precision are not directly affected by the choice of interviewers, as they depend more on the quality, relevance, and accuracy of the questions and the data analysis.
Reference:1: Guide to Business Data Analytics, IIBA, 2020, p. 26.

QUESTION 13
The analytics team has been asked to provide an estimate of the number of customers they expect to have in 12 months. They debated how accurate that figure needs to be and determined that based on the availability of good data, they could predict within + or - 10%. This is an example of a:

A. ROM estimate
B. Delphi estimate
C. Parametric estimate
D. Definitive estimate

Correct Answer: A
Section:
Explanation:

A ROM estimate is a rough order of magnitude estimate that provides a quick and approximate estimate of the cost, time, or effort required for a project or a task. A ROM estimate is based on expert opinion or experience from past projects, and it usually has a large range of variation, such as + or - 10%. A ROM estimate is useful when there is limited information or data available, or when a high-level estimate is needed for planning or budgeting purposes. However, a ROM estimate also has a high degree of uncertainty and variability, and it should be refined as more details become available12
Reference: 1: Project Estimation Techniques Business Analysts Should Know About 2: Estimation techniques for business analysts -- The Functional BA

QUESTION 14
To gain traction on online sales, a retailer initiated a marketing campaign using banner ads. The company has requested their analytics team to evaluate the performance of the campaign. During the presentation, the analyst confirmed that the campaign did bring in a large number of net new customers to the website and met the target sales conversion rate. They also noted that there was a high number of repeat visitors not completing a sale. What decision would help the retailer improve sales conversion rates for repeat visitors?

A. Increase investment in banner ads
B. Incentivize customers to subscribe to promotional notifications
C. Add additional new products to attract customers
D. Ensure the sales checkout process is streamlined

Correct Answer: D
Section:
Explanation:
According to the Business Data Analytics: A Decision-Making Paradigm1, one of the key steps in the analytics process is to communicate insights and recommendations to stakeholders. The analyst should present the findings in a clear and concise manner, and provide actionable suggestions to improve the business outcomes. In this case, the analyst has identified that repeat visitors are not completing a sale, which indicates a possible issue with the sales checkout process. Therefore, the analyst should recommend the retailer to streamline the sales checkout process, which could reduce friction, increase customer satisfaction, and boost sales conversion rates for repeat visitors.
Reference: Business Data Analytics: A Decision-Making Paradigm

QUESTION 15
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
Section:
Explanation:
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.
Reference:
* Business Analysis Certification in Data Analytics, CBDA | IIBA, CBDA Competencies, Domain 4: Interpret and Report Results
* Understanding the Guide to Business Data Analytics, page 18
* 16 Best Types of Charts and Graphs for Data Visualization [+ Guide]

QUESTION 16
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
Section:
Explanation:
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.
Reference:
* Business Analysis Certification in Data Analytics, CBDA | IIBA, CBDA Competencies, Domain 5: Use Results to Influence Business Decision Making
* Understanding the Guide to Business Data Analytics, page 9
* CERTIFICATION IN BUSINESS DATA ANALYTICS HANDBOOK - IIBA, page 8, CBDA Exam Sample Questions and Self-Assessment, Question 12

QUESTION 17
Based on the results of a recently completed analytics initiative, the Human Resource department for a major department store implemented a change to its hiring practice to address the attrition rates of its sales associates. The new policy stated that candidates applying for sales positions must possess at least 3 years of relevant sales experience to be considered. After implementing the change, attrition rates are 10% higher and management is frustrated. Which of the following could result in this outcome?

A. The results of analysis have been incorrectly interpreted
B. Sales experience is not a relevant skill
C. Analytics is not helpful given this situation
D. The change proposed is not aligned to company strategy

Correct Answer: D
Section:
Explanation:
The change proposed is not aligned to company strategy, because it may not address the root cause of the attrition problem, or it may conflict with other organizational goals or values. For example, the change may reduce the pool of qualified candidates, increase the hiring costs, or lower the diversity or customer satisfaction of the sales team. The change may also ignore other factors that influence the attrition rates, such as compensation, training, feedback, or recognition. Therefore, the change may not achieve the desired outcome of reducing attrition, and may even worsen it.
Reference:
* Business Analysis Certification in Data Analytics, CBDA | IIBA, CBDA Competencies, Domain 5: Use Results to Influence Business Decision Making
* Understanding the Guide to Business Data Analytics, page 9
* CERTIFICATION IN BUSINESS DATA ANALYTICS HANDBOOK - IIBA, page 8, CBDA Exam Sample Questions and Self-Assessment, Question 13

QUESTION 18
A company wants to gauge the thoughts of their employees towards a new company product. On the 25th of March the interviewer makes a list of all employees who were at work on that day and then chooses a subset of those employees to interview. Which term describes the list of all employees present on March 25th?

A. Population of interest
B. Survey sample
C. Sampling frame
D. Sample weights

Correct Answer: C
Section:
Explanation:
The sampling frame is the term that describes the list of all employees present on March 25th, because it is a technique that defines the set of elements from which a sample is drawn. The sampling frame should ideally match the population of interest, which is the group of elements that the researcher wants to study or make inferences about. In this case, the population of interest is the employees of the company, and the sampling frame is the subset of employees who were at work on a specific day. The survey sample is the technique that selects a portion of the sampling frame to participate in the survey. The sample weights are the technique that assigns different values or importance to each element in the sample, based on their representation in the population.
Reference:
* Business Analysis Certification in Data Analytics, CBDA | IIBA, CBDA Competencies, Domain 2: Source Data
* Understanding the Guide to Business Data Analytics, page 14
* CERTIFICATION IN BUSINESS DATA ANALYTICS HANDBOOK - IIBA, page 8, CBDA Exam Sample Questions and Self-Assessment, Question 14

QUESTION 19
Interested in building out the analytics capability based on the positive results obtained by past analytics efforts, the Chief Marketing Officer (CMO) pitches the idea of using analytics to guide future decision making across the enterprise. Before allocating budget to build up an enterprise analytics practice, the decision makers should:

A. Request that a small team be assembled to brainstorm a list of capabilities to develop with any approved monies
B. Identify the sponsor and a project manager who can collaborate on the development of the project charter
C. Oversee the completion of up-front analysis to determine how value can be achieved through an enterprise-wide analytics practice
D. Determine if the company has the sufficient resources to build up the analytics practice

Correct Answer: C
Section:
Explanation:
Before investing in an enterprise analytics practice, the decision makers should have a clear understanding of the expected value and benefits of such a practice. This requires conducting an up-front analysis that identifies the business problems or opportunities that can be addressed by analytics, the data sources and technologies that are needed, the analytical models and methods that are appropriate, and the metrics and indicators that will measure the impact and outcomes of the analytics solutions12. This analysis will help to define the scope, objectives, and requirements of the enterprise analytics practice, as well as the resources, roles, and governance structures that are necessary to support it34. An up-front analysis will also help to prioritize the analytics initiatives based on their feasibility, alignment with the business strategy, and potential value creation
Topic 2, Exam Pool B

QUESTION 20
A business analyst is conducting a series of interviews to understand the research questions that will be explored within a new analytics project. Which of the following is true about interviews?

A. Planned interviews are less effective than unplanned
B. Interviews must be structured to be effective
C. Goals for the interview should be clearly articulated
D. Interviews should only be conducted with one interviewee

Correct Answer: C
Section:
Explanation:
Interviews are a technique to elicit information from stakeholders and subject matter experts. Interviews can be planned or unplanned, structured or unstructured, depending on the context and purpose of the interview. However, regardless of the type of interview, it is important to have clear goals for the interview, such as what information is needed, what questions will be asked, and how the information will be used. Having clear goals for the interview helps the interviewer to prepare, conduct, and follow up the interview effectively, and also helps the interviewee to understand the expectations and provide relevant and accurate information.
Reference:Guide to Business Data Analytics, page 25;Certification in Business Data Analytics Handbook, page 9;How to Ace Your Next Business Analysis Job Interview

QUESTION 21
An insurance company would like to develop a range of insurance products for different types of customers. The analytics team is asked to conduct some research and share their insights with senior management. Which technique would be useful to divide the customer base into groups?

A. Linear regression
B. Survey sampling
C. Factor analysis
D. K-means clustering

Correct Answer: D
Section:
Explanation:
K-means clustering is a technique that partitions a set of data points into a predefined number of clusters, based on their similarity or distance. This technique can be useful to divide the customer base into groups that have similar characteristics, preferences, or behaviors, and then design insurance products that cater to each group's needs and expectations. K-means clustering can also help identify outliers or anomalies in the customer data that may require further investigation or attention.

QUESTION 22
A brainstorming session is conducted to identify the research questions to be explored within an analytics project. During the brainstorming activity which of the following should happen?

A. The number of questions generated should be limited to contain scope
B. Participants should make sure the questions are unique and realistic
C. Participants should add questions as they come to mind without restriction on time limit
D. Participants should avoid critiquing suggested questions raised by the group

Correct Answer: D
Section:
Explanation:
Explanation: According to the Guide to Business Data Analytics, brainstorming is a technique used to generate a large number of ideas or questions in a short period of time1. The purpose of brainstorming is to encourage creativity and divergent thinking, not to evaluate or judge the ideas or questions. Therefore, participants should avoid critiquing suggested questions raised by the group, as this could inhibit the flow of ideas and discourage participation.The other options are not consistent with the principles of brainstorming, as they could limit the quantity or quality of the questions generated.

QUESTION 23
An analyst is doing a clinical study on the value of analyte among a large population of healthy people. The analyst is going to use a Gaussian Distribution to share the results. Which of the following represents a Gaussian Distribution? (IMAGE TAKEN)

A. D
B. B
C. A
D. C

Correct Answer: C
Section:
Explanation:
Explanation: As explained in the previous question, a Gaussian Distribution, also known as a normal distribution, is represented by a symmetrical bell-shaped curve. The mean, median, and mode of the distribution are equal and are at the center of the distribution. This type of distribution is characterized by its mean and standard deviation. The curve is symmetrical around the mean. In the image, the curve labeled A is the only one that matches this description. The other curves are either skewed or irregular.

QUESTION 24
The analytics team has been asked to determine if the organization should launch their highest revenue generating product into the North American market. To date, this has only been available in Eastern Europe. To answer this, the team formulates several research questions, including:

A. What product launch related costs can we expect?
B. How much revenue does the product generate in Eastern Europe?
C. Why does management need to know this?
D. Do existing customers really like the product?

Correct Answer: D
Section:
Explanation:
One of the steps in identifying the research questions for business data analytics is to assess the feasibility and desirability of the proposed solution or change1. This involves understanding the needs, preferences, and satisfaction of the existing and potential customers. Therefore, asking whether the existing customers really like the product is a relevant research question for the analytics team.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 22.

QUESTION 25
An analyst has just completed building a data model that shows the table structures including table names, table relationships with primary and foreign keys and column names with respective data types. What type of data model has the analyst just built?

A. Physical
B. Hierarchical
C. Conceptual
D. Logical

Correct Answer: A
Section:
Explanation:
A physical data model is the most detailed and specific type of data model, which shows how the data is stored, accessed, and manipulated in the database. It includes the table structures, column names, data types, primary and foreign keys, constraints, indexes, and other physical attributes of the data12.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 542: Data Modeling Essentials, Graeme Simsion and Graham Witt, 2005, p. 15.

QUESTION 26
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
Section:
Explanation:
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 resources3.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 282: Data Quality Assessment, Arkady Maydanchik, 2007, p. 33: Data Quality: The Field Guide, Thomas C. Redman, 2001, p. 1.

QUESTION 27
An analyst at a supermarket chain has been asked to extract data from multiple data sources to complete a study on customer spending habits. The analyst is going to query data from various databases. Which statement is true about database querying?

A. Querying can be used to create predictive data models

B. Irrespective of the querying language used, data results retrieved are always in a tabular format
C. A querying language is independent of the type of database being used
D. Querying is a structured way of searching, manipulating and managing data

Correct Answer: D
Section:
Explanation:
Querying is a technique that allows analysts to access, filter, join, aggregate, and transform data from various databases using a specific syntax and logic1. Querying can be used for different purposes, such as data exploration, data preparation, data analysis, and data visualization2. Querying is not limited to creating predictive data models, nor does it always produce tabular results. Moreover, querying languages may vary depending on the type and structure of the database, such as relational, hierarchical, or document-based3.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 552: Data Analysis Using SQL and Excel, Gordon S. Linoff, 2016, p. 33: Database Systems: Design, Implementation, and Management, Carlos Coronel and Steven Morris, 2019, p. 17.

QUESTION 28
A lab is conducting a study on protein interactions. They have used the data to create a graph visualization. In graph visualization, what would a layout be?

A. A single data point
B. A link between two data points
C. A dedicated algorithm that calculates the node positions
D. A collection of data points and links

Correct Answer: C
Section:
Explanation:
A layout is a way of arranging the nodes and links of a graph visualization to convey meaningful information about the data. A layout is determined by a dedicated algorithm that calculates the node positions based on certain criteria, such as minimizing edge crossings, maximizing node spacing, or emphasizing clusters12. A layout can also be influenced by user interaction, such as zooming, panning, or dragging3.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 642: Graph Drawing: Algorithms for the Visualization of Graphs, Giuseppe Di Battista et al., 1999, p. 33: Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward et al., 2015, p. 227.

QUESTION 29
An analyst at a bank is trying to identify research questions for an analytical study on top customer issues across branches. During an interview with a branch manager, the analyst asks the manager what their top customer concerns are relating to this branch?
After the manager's reply, the analyst asks a follow up question on how their top customer concerns compare against the top customer concerns across all branches? Was the analyst's follow-up question valid?

A. No, there is no value comparing the results of a single branch with results across all branches
B. Yes, it builds on the previous question and allows the analyst to identify branch-specific concerns
C. No, the question is not valid in this particular scenario
D. Yes, only for the purpose of ensuring that the manager is aware of the company-wide reports

Correct Answer: B
Section:
Explanation:
The analyst's follow-up question is valid because it helps to refine the scope and context of the research questions for the analytical study. By comparing the top customer concerns across branches, the analyst can identify the common and unique issues that affect customer satisfaction and loyalty. This can also help to prioritize the most critical or urgent problems that need to be addressed by the bank12.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 212: Business Analysis for Practitioners: A Practice Guide, PMI, 2015, p. 43.

QUESTION 30
Interested in experimenting with analytics, a manufacturing company hires an analyst to see how the capability can be developed within its organization. The analyst is getting started and recognizes the need to show value

from the onset of their work to gain upper management's trust and future funding. What action will accomplish these objectives?

A. Solve the biggest problem the organization has first to quickly grab the support and attention of senior management
B. Develop a question that can be answered quickly regardless of alignment to strategy, just to get started
C. Develop a meaningful question that can be answered with data the company already has in its possession
D. Perform a market analysis to understand how competitors are using analytics and then launch a similar initiative

Correct Answer: C
Section:
Explanation:
The best action for the analyst to show value from the onset of their work is to develop a meaningful question that can be answered with data the company already has in its possession. This way, the analyst can demonstrate the potential of analytics to solve relevant business problems, without spending too much time or resources on data collection or market research. The question should also be aligned with the organization's strategy and goals, and provide actionable insights for decision making12.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 202: Data Science for Business, Foster Provost and Tom Fawcett, 2013, p. 14.

QUESTION 31
A large car manufacturer is interested in comparing the number of sales for a specific model of electric car across all 50 US states.
The data analytics team sourced and acquired the data, and the business analyst created the model to compare sales across states.
In a meeting to review the results, the feedback received included several complaints concerning an inability to distinguish the number of sales per state. What model would result in such confusion?

A. Bullet chart
B. Dual axis chart
C. Bar chart
D. Pie chart

Correct Answer: D
Section:
Explanation:
A pie chart is a circular chart that shows the proportion of each category in a whole by dividing the circle into slices. A pie chart would result in confusion when comparing the number of sales for a specific model of electric car across all 50 US states, because it is difficult to compare the angles and areas of the slices, especially when there are many categories with similar values. A pie chart also does not show the absolute values of each category, unless they are labeled or annotated12. A better alternative would be a bar chart, which can show the number of sales for each state along a common axis, making it easier to compare and rank the values3.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 652: Storytelling with Data, Cole Nussbaumer Knaflic, 2015, p. 673: The Visual Display of Quantitative Information, Edward R. Tufte, 2001, p. 178.

QUESTION 32
The definition of data elements is different across various data sources. The organization is looking to improve the usability of data across the organization. Which practice would help address this problem?

A. Data governance
B. Data quality
C. Data architecture
D. Data ethics

Correct Answer: A
Section:
Explanation:
Data governance is the practice of establishing and enforcing policies, standards, roles, and responsibilities for the management and use of data across the organization. Data governance helps to address the problem of inconsistent data definitions across various data sources by ensuring that data is properly defined, documented, classified, and aligned with the business objectives and requirements12.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 292: Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program, John Ladley, 2012, p. 3.

QUESTION 33
Insights based on the data collected indicate that a multi-national company could increase its sales of a mature product by reducing its price by 20% which would result in increased revenues of 2% over a 6-month period. The team recommends this as an appropriate goal for its organization. This is considered a good goal because:

A. It meets all the criteria for a well-defined objective
B. The organization can derive additional revenue from the product
C. It indicates that the company does not have to incur costs associated with retiring this product
D. Management will be pleased that the mature product can still contribute to revenue

Correct Answer: A
Section:
Explanation:
A well-defined objective is one that is specific, measurable, achievable, relevant, and time-bound (SMART)1. The goal of increasing sales of a mature product by reducing its price by 20% which would result in increased revenues of 2% over a 6-month period meets all these criteria, as it clearly states what the desired outcome is, how it will be measured, whether it is realistic and attainable, how it aligns with the organization's strategy, and when it will be achieved2.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 192: SMART Goals: How to Make Your Goals Achievable, MindTools, 2021, 1.

QUESTION 34
The marketing department for a major restaurant chain is interested in testing a Kids Eat Free campaign to determine if it will help to increase sales. They are interested in piloting the campaign to determine which day of the week will improve sales the most.
The campaign is launched across 7 cities with each city promoting a different day of the week. The sales data is collected and provided to a team for analysis. What concern might the analytics team have regarding data quality across cities?

A. Normality
B. Heteroskedacity
C. Linearity
D. Variation

Correct Answer: D
Section:
Explanation:
Variation is the degree to which the data values differ from each other or from a central tendency measure, such as the mean or median. Variation can affect the data quality across cities, as it can indicate the presence of outliers, errors, noise, or inconsistency in the data collection or processing methods. Variation can also influence the statistical analysis and interpretation of the results, as it can affect the significance, confidence, and validity of the findings12.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 302: Statistics for Business and Economics, David R. Anderson et al., 2014, p. 83.

QUESTION 35
A call center has requested to review their sales conversion data for the month. The analyst working on this request is trying to identify the chart that will effectively present the data, which includes: the number of leads, the number of calls made, the number of calls completed, the number of customers interested and the number of sales. What chart should the analyst use to show the values across each stage of the pipeline?

A. Pie chart
B. Funnel chart
C. Bar chart
D. Bullet chart

Correct Answer: B
Section:
Explanation:

A funnel chart is a type of chart that shows the values of different stages of a process, such as a sales pipeline, where each stage represents a subset of the previous one. A funnel chart is useful for showing the conversion rate, the drop-off rate, and the potential revenue or profit at each stage12. A funnel chart would be an effective way to present the data requested by the call center, as it would show the number of leads, calls, customers, and sales, as well as the percentage of change between each stage.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 662: Data Visualization: A Practical Introduction, Kieran Healy, 2018, p. 233.

QUESTION 36
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
Section:
Explanation:
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.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 582: Statistics for Business and Economics, David R. Anderson et al., 2014, p. 7133: Guide to Business Data Analytics, IIBA, 2020, p. 594: Statistics for Business and Economics, David R. Anderson et al., 2014, p. 849.

QUESTION 37
An organization's customers are categorized based on the amount of purchases completed over the last 12 months. The analytics team would like to ensure the accuracy of their survey results and decide to randomly select 500 customers to participate in a survey from this large pool of customers. This is an example of:

A. Stratified sampling
B. Quota sampling
C. Purposive sampling
D. Snowball sampling

Correct Answer: A
Section:
Explanation:
Stratified sampling is a technique that divides the population into homogeneous subgroups (strata) based on a relevant characteristic, such as the amount of purchases, and then randomly selects a proportional number of elements from each subgroup to form the sample. Stratified sampling ensures that the sample is representative of the population and reduces the sampling error and bias12.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 312: Statistics for Business and Economics, David R. Anderson et al., 2014, p. 262.

QUESTION 38
An insurance company has seen an upward trend in winter-related accidents over the past three years. The company has just completed an analytics study to better understand the primary reasons for these accidents and assess how many of the drivers were using winter tires. This analysis will help the company decide how to move forward with drivers not taking precautionary measures during winter. What type of analysis will help in determining the primary reasons and percentage of those drivers with winter tires?

A. Prescriptive
B. Descriptive and Predictive
C. Descriptive

D. Descriptive and Diagnostic

Correct Answer: D
Section:
Explanation:
Descriptive analytics is a type of analytics that summarizes and visualizes the data to provide an overview of what has happened or is happening, such as the trend of winter-related accidents over the past three years, or the percentage of drivers using winter tires12. Diagnostic analytics is a type of analytics that explores and analyzes the data to understand why something has happened or is happening, such as the primary reasons for these accidents, or the factors that influence the drivers' decisions13. To answer the question, both descriptive and diagnostic analytics would be needed to provide the relevant information and insights for the company.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 182: Business Analytics: Data Analysis & Decision Making, S. Christian Albright and Wayne L. Winston, 2015, p. 53: Data Science for Business, Foster Provost and Tom Fawcett, 2013, p. 13.

QUESTION 39
A Human Resource manager recently learned that their competitor reduced employee attrition rates by 20% after implementing personality tests as part of their screening process. Intrigued by the idea, the manager suggests collecting data on personality tests and attrition rates over the next year. The data from this year is then analyzed to explore possible relationships. What type of analytics has the team been asked to perform?

A. Predictive
B. Descriptive
C. Prescriptive
D. Diagnostic

Correct Answer: B
Section:
Explanation:
Descriptive analytics is a type of analytics that summarizes and visualizes the data to provide an overview of what has happened or is happening, such as the attrition rates and the personality test scores of the employees12. The team has been asked to perform descriptive analytics to explore possible relationships between the data variables, without making any predictions or prescriptions for the future.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 182: Business Analytics: Data Analysis & Decision Making, S. Christian Albright and Wayne L. Winston, 2015, p. 5.

QUESTION 40
A large telecommunications company wants to increase their Average Revenue Per User per month by 5%, by end of year, to increase revenue in a highly competitive market. From a SMART target perspective, what is missing?

A. T - The increase should be seen sooner
B. A - It is too easy of a target to attain
C. R - Since competition is high, focus should be on increasing customer base and not on ARPU
D. S - There is no mention of which product group/line the target pertains to

Correct Answer: D
Section:
Explanation:
A SMART target is one that is specific, measurable, achievable, relevant, and time-bound1. The target of increasing the Average Revenue Per User (ARPU) per month by 5%, by end of year, to increase revenue in a highly competitive market is missing the specificity criterion, as it does not mention which product group or line the target applies to. The target should be more specific and clear about the scope and context of the desired outcome, such as which segment, region, or service the target relates to23.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 192: SMART Goals: How to Make Your Goals Achievable, MindTools, 2021, 13: How to Set SMART Marketing Goals, CoSchedule, 2021, 2.

QUESTION 41
A database analyst is modelling a database for a large toy manufacturer. Which statement describes a logical database model?

A. The layer of views created to summarize data or provide another perspective of certain data

B. A model that depicts the actual design of the relational database
C. An abstraction of the conceptual data model that includes rules of normalization
D. Modelling that involves objects being defined at the schema level

Correct Answer: C
Section:
Explanation:
A logical database model is a data model of a specific problem domain expressed independently of a particular database management product or storage technology. It describes data using notation that corresponds to a data organization used by a database management system, such as relational tables and columns. It also includes rules of normalization, which are the process of converting complex data structures into simple, stable data structures12
Reference: 1: Logical schema - Wikipedia 2: What Is a Data Model? | Coursera

QUESTION 42
A professional association is funded by membership fees. The membership renewal occurs every 5 years. Although, they have a strong subscription rate each year, their renewal rate is low. They are working with an external firm specializing in Business Analytics to identify the groups of customers that have a high likelihood of cancelling their subscription after their first 5-year term ends. This type of study is called:

A. Untrained learning
B. Supervised learning
C. Trained learning
D. Unsupervised learning

Correct Answer: D
Section:
Explanation:
Unsupervised learning is a type of study that involves finding patterns or clusters in data without any predefined labels or outcomes. It is useful for exploring data and discovering hidden structures or groups of customers. For example, the professional association can use unsupervised learning to identify the characteristics of customers who are likely to cancel their subscription after their first 5-year term ends, and then design strategies to retain them12
Reference: 1: What is Unsupervised Learning? - IBM 2: Unsupervised Learning - IIBA BABOK Guide v3

QUESTION 43
Which attributes from the Order entity will need to be normalized to avoid redundancies?
. OrderId
. OrderDate
. ItemId
. ItemName
. Quantity
. ItemPrice

A. OrderDate ItemPrice
B. ItemName ItemPrice
C. OrderDate ItemName
D. Item Name Quantity

Correct Answer: B
Section:
Explanation:
The attributes ItemName and ItemPrice need to be normalized to avoid redundancies because they depend on the attribute ItemId, which is not part of the primary key of the Order entity. This is a case of partial dependency, which violates the second normal form (2NF) of database normalization. To achieve 2NF, the Order entity should be split into two entities: Order and Item, where Item contains the attributes ItemId, ItemName, and ItemPrice,

and Order contains the attributes OrderId, OrderDate, ItemId, and Quantity. This way, the ItemName and ItemPrice are stored only once for each ItemId, and the Order entity references them through a foreign key12
Reference: 1: Balancing Data Integrity and Performance: Normalization vs ... 2: Normalization Process in DBMS - GeeksforGeeks

QUESTION 44
The architecture team puts forth a solution architecture that integrates multiple data sources from within and outside the organization. The architecture provides the foundation to source a new analytics program. If one of the objectives of the analytics team was to provide 'one source of the truth', this objective would be referring to which of the following?

A. Identifying one key stakeholder, who can make final decisions about which sources to relate/merge
B. Evaluating the completeness, validity, and reliability of the data from source systems
C. Ensuring stakeholders always have clear insight into the final requirements at all times
D. Enforcing master data management principles and practices

Correct Answer: D
Section:
Explanation:
Providing 'one source of the truth' means ensuring that there is a single, consistent, and authoritative source of data that can be used for analytics and decision making across the organization. This objective can be achieved by enforcing master data management principles and practices, which involve defining, governing, and maintaining the quality and integrity of the core data entities that are shared by multiple systems and processes. Master data management helps to eliminate data silos, reduce data duplication and inconsistency, and improve data accuracy and reliability12
Reference: 1: What is Master Data Management (MDM)? - Informatica 2: Master Data Management - IIBA BABOK Guide v3

QUESTION 45
The results of the data analytics work led to some clear and strongly supported outcomes and the analytics team is very confident in their recommendations; particularly given that the payback on the required changes are a short 3 months. However, there is concern because the organization operates in a highly regulated environment and some new regulatory changes are being considered with announcements and implementation in the next 6 months. Under these conditions the team decides to:

A. Recommend no action be taken at this time and revisit in 6 months
B. Reassess their results to ensure their validity and then decide what to do
C. Identify and carefully document assumptions for their recommendation
D. Postpone recommendations for 6 months until the announcements are made

Correct Answer: C
Section:
Explanation:
The best option for the team under these conditions is to identify and carefully document the assumptions for their recommendation, such as the expected impact of the regulatory changes, the risks and benefits of implementing the changes before or after the announcements, and the sensitivity of the results to different scenarios. This way, the team can communicate their findings and recommendations clearly and transparently, while also acknowledging the uncertainty and limitations of their analysis. This can help the decision makers to evaluate the trade-offs and make informed choices12.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 242: Data-Driven Decision Making: A Primer for Beginners, Anand Rao, 2018, 1.

QUESTION 46
A colleague proposes measuring job satisfaction by asking the question 'What is your salary?'. What is the concerning factor about this question?

A. Validity
B. Clarity
C. Reproducibility
D. Subjectivity

Correct Answer: A
Section:

Explanation:
Validity is the extent to which a measure or a question accurately captures the intended concept or construct1. The question ''What is your salary?'' is not a valid measure of job satisfaction, as it does not reflect the various aspects of job satisfaction, such as work environment, recognition, autonomy, growth, etc. Salary is only one possible factor that may influence job satisfaction, but it is not a direct or comprehensive indicator of it23. Therefore, the question is not valid for measuring job satisfaction.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 302: Job Satisfaction: Application, Assessment, Causes, and Consequences, Paul E. Spector, 1997, p. 23: Job Satisfaction Survey, 1.

QUESTION 47
A marketing director has asked the question 'How many product purchases are expected this coming year given the current marketing campaign?'. What type of analytics would be performed to answer this question?

A. Descriptive
B. Predictive
C. Diagnostic
D. Prescriptive

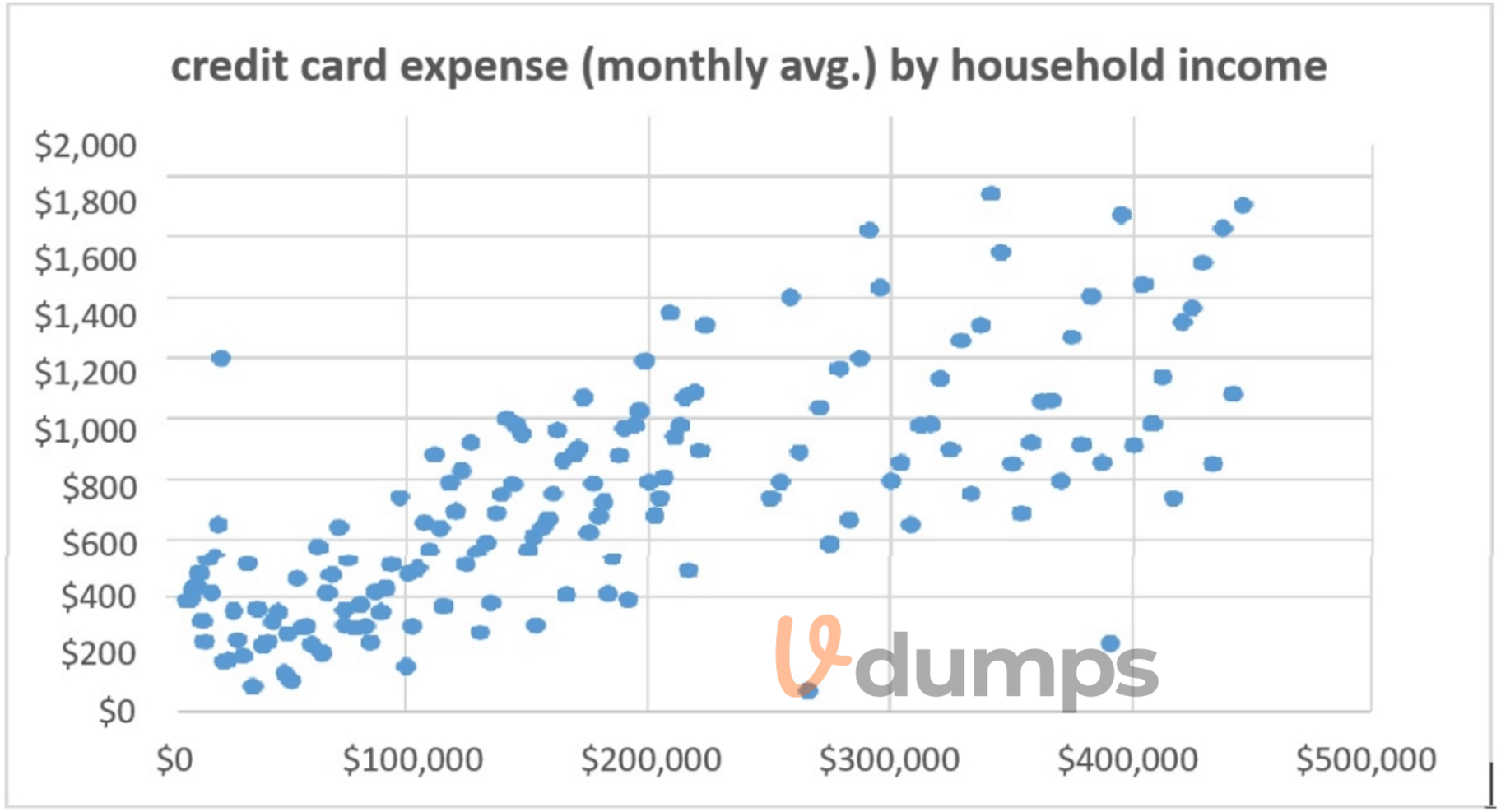
Correct Answer: B
Section:
Explanation:
Predictive analytics is a type of analytics that uses historical and current data, as well as statistical and machine learning techniques, to forecast future events or outcomes, such as product purchases, customer behavior, or market trends12. To answer the question 'How many product purchases are expected this coming year given the current marketing campaign?', predictive analytics would be performed to estimate the demand and sales based on the existing data and the marketing campaign variables.
Reference: 1: Guide to Business Data Analytics, IIBA, 2020, p. 182: Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die, Eric Siegel, 2016, p. 3.

QUESTION 48
As the team discusses how to utilize the results of their data analysis to put forth a business recommendation, an analyst on the team voices concern over the current organizational culture presenting a roadblock to their ability to influence business decision making. Which of the following would be a justifiable concern at this stage of the team's efforts?

A. Difficulty bringing business stakeholders to a shared understanding about value when sharing data assets across business domains
B. Changing the mindsets of business stakeholders to trust insights gleaned from data over experience and intuition
C. Applying a myopic view of data and establishing data silos which create roadblocks to exploring available data sources
D. Finding data that creates value creating difficulties, as not all data helps a business make better decisions

Correct Answer: B
Section:
Explanation:
A justifiable concern at this stage of the team's efforts is changing the mindsets of business stakeholders to trust insights gleaned from data over experience and intuition. This is because some stakeholders may have a strong attachment to their own opinions or beliefs, and may resist or ignore data that contradicts them. This can create a barrier to data-driven decision making, which requires a culture of curiosity, openness, and evidence-based reasoning. The team needs to communicate the value and validity of their data analysis, and persuade the stakeholders to adopt a data-driven mindset12
Reference: 1: Use Data to Accelerate Your Business Strategy 2: Data-Driven Decision Making: A Step-by-Step Guide

QUESTION 49
An analytics team employed at a leading credit card company is utilizing data analytics to identify unusual credit card purchases.
They have created the following visual. How many extreme outliers exists in this dataset?



A. 0
B. 5
C. 3
D. 2

Correct Answer: C
Section:
Explanation:
According to the Business Data Analytics (IIBA- CBDA) principles, extreme outliers in a dataset can be identified visually on a scatter plot as points that are distinctly separate from the bulk of the data. In this visual, there are three points that are significantly higher on the y-axis (credit card expense) relative to their position on the x-axis (household income), indicating unusual credit card purchases.
Reference: The identification and interpretation of outliers is a standard practice in data analytics and is covered under the Business Data Analytics (IIBA- CBDA) learning resources.

QUESTION 50
A professor at a university has received a few complaints of the exams being too difficult. The professor is looking at exam performance results over the past 5 years to understand the normal tendency and outliers. Which chart should the professor use?

A. Sunburst

B. Scatterplot
C. Pie chart
D. Line

Correct Answer: B
Section:
Explanation:
A scatterplot is a type of chart that shows the relationship between two variables by plotting data points on a two-dimensional plane. A scatterplot can help the professor to understand the normal tendency and outliers of exam performance results over the past 5 years by displaying the distribution, trend, and correlation of the data. For example, the professor can use the x-axis to represent the year and the y-axis to represent the exam score, and see how the scores vary over time and across different exams. Outliers can be identified as data points that are far away from the main cluster or the line of best fit12
Reference: 1: Scatter Plot - Statistics How To 2: Scatterplots - IIBA BABOK Guide v3