Free Questions for CBDA by actualtestdumps

Shared by Harris on 27-02-2024

For More Free Questions and Preparation Resources

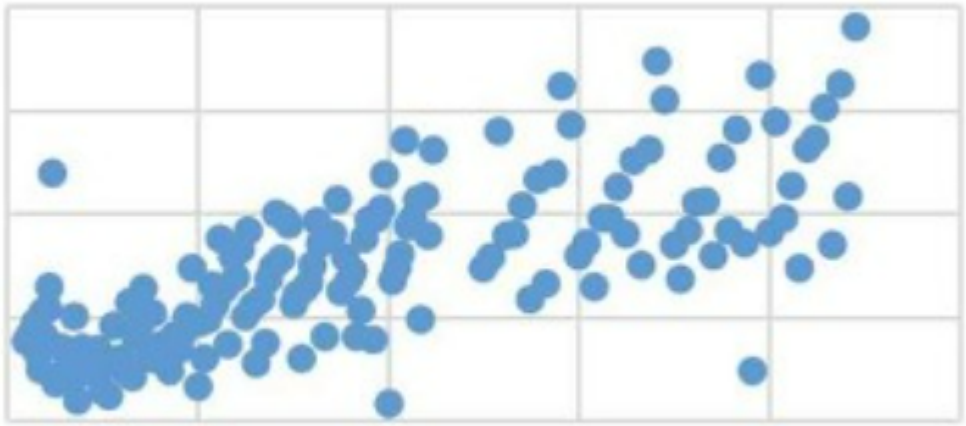
Check the Links on Last Page

Question 1

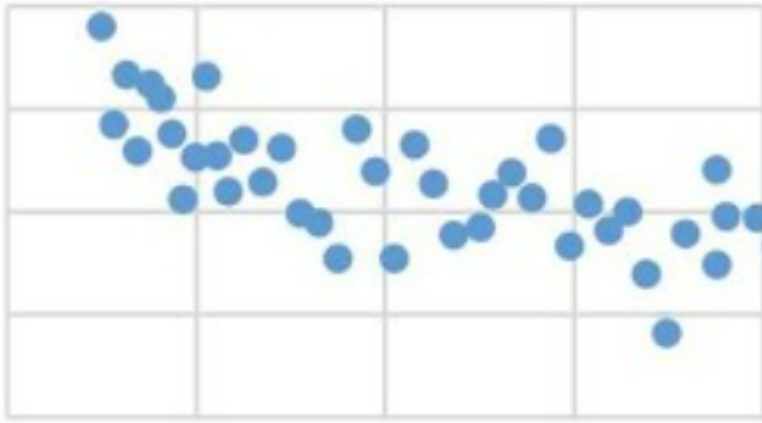
Question Type: MultipleChoice

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?

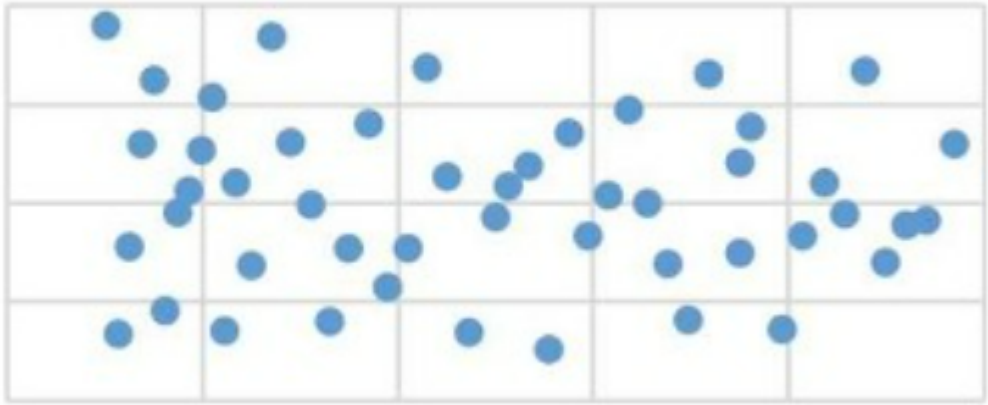
1.



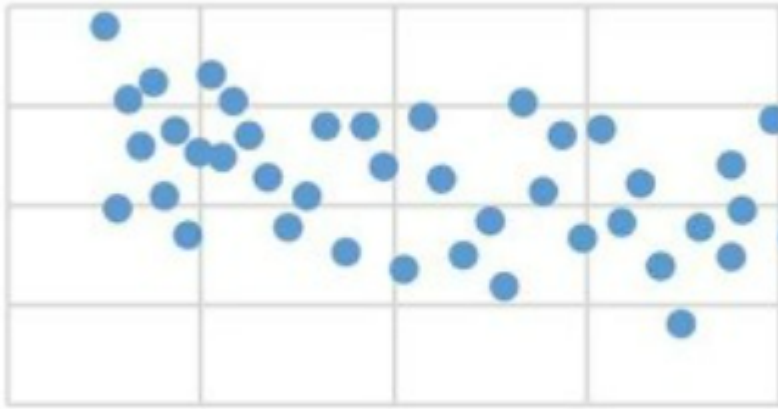
2.



3.



4.



Options:

A- 2

B- 1

C- 4

D- 3

Answer:

C

Explanation:

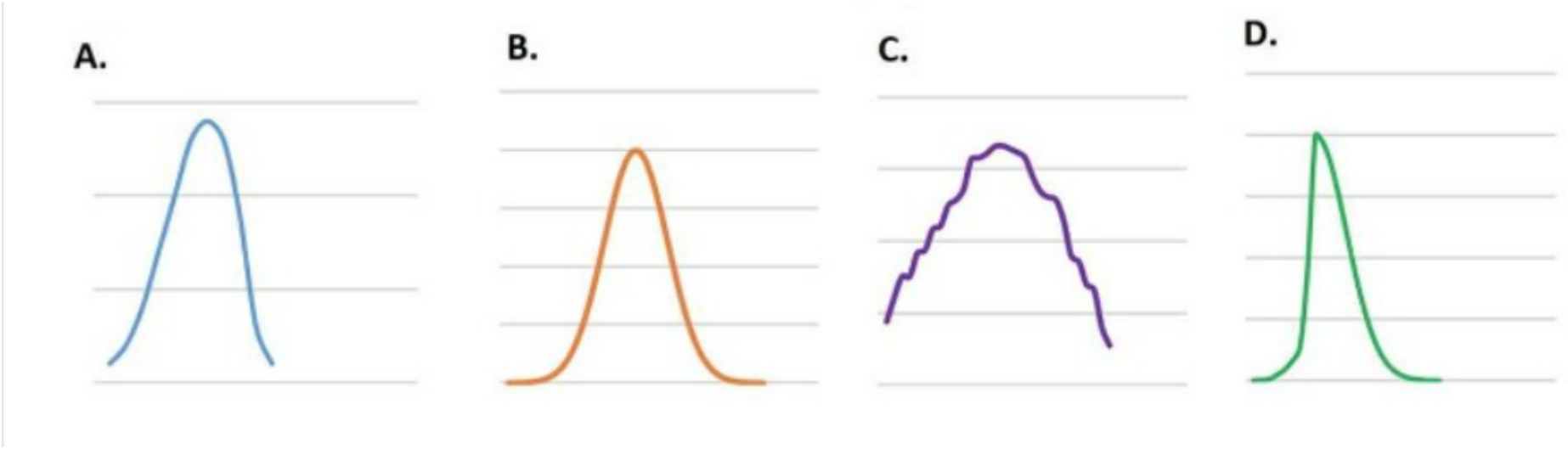
Covariance measures the directional relationship between the returns on two assets. A positive covariance means that asset returns move together while a negative covariance means they move inversely. Zero covariance indicates that the returns on the two assets move independently of each other. In the context of a scatter plot, zero covariance is represented by a plot where the points do not show any upward or downward trend but are rather scattered randomly on the graph with no discernible pattern.

Graph 4 displays such a pattern where there is no apparent relationship between the variables on the x and y axes, indicating that there is zero covariance between them.

Question 2

Question Type: MultipleChoice

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?



Options:

A- D

B- A

C- C

D- B

Answer:

B

Explanation:

The Gaussian distribution, also known as the normal distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph form, the Gaussian distribution will appear as a bell curve, which is the case with option A. It is characterized by its bell-shaped curve and is defined by the mean () and the standard deviation (). It is a common assumption for the distribution of independent, randomly generated variables.

Question 3

Question Type: MultipleChoice

An analyst is working through data on comparing performance scores in different schools across the state, for ranking purposes. Since there is a lot of data and some extreme outliers, the analyst is trying to determine which type of statistical average would best represent the results. Which of the following is a concern when relying too heavily on summary statistics during data analysis?

Options:

A- Contextualization

B- Data variation

C- Data properties

D- Frequency

Answer:

A

Explanation:

Summary statistics are numerical measures that describe certain characteristics of a data set, such as the mean, median, mode, standard deviation, range, or quartiles. Summary statistics can help simplify and communicate complex data, but they can also obscure or distort important information, such as the distribution, shape, outliers, or trends of the data. Contextualization is the process of providing relevant background information, assumptions, limitations, or explanations for the data analysis and its results. Contextualization can help avoid misinterpretation, confusion, or bias when using summary statistics. Contextualization can also help connect the data analysis to the business problem, objectives, and stakeholders.

Question 4

Question Type: MultipleChoice

After completing their data analysis, an analyst is drawing out the results, explaining the methods and processes used, and identifying any limitations or weaknesses in the data or methods applied. While performing these steps, which recommended practice would the analyst apply?

Options:

A- Use exploratory analysis to determine the best mathematical method to use

B- Understand the communication needs of stakeholders

C- Let the data drive the conclusions and the insights reached

D- Learn a variety of visualization techniques for effective communications

Explanation:

According to the IIBA Guide to Business Data Analytics, communication is a key skill for analysts, as it involves conveying the results, methods, and limitations of the data analysis to various stakeholders in a clear, concise, and meaningful way. To communicate effectively, analysts need to understand the communication needs of stakeholders, such as their level of interest, knowledge, and influence, their preferred format and frequency of communication, and their expectations and objectives. By understanding the communication needs of stakeholders, analysts can tailor their messages, choose the appropriate language and tone, and select the most suitable communication channels and medi

a. Therefore, the correct answer is B, as understanding the communication needs of stakeholders is a recommended practice for analysts while performing the steps of drawing out the results, explaining the methods and processes used, and identifying any limitations or weaknesses in the data or methods applied.

Answer:

B

Question 5

Question Type: MultipleChoice

An analyst is performing regression analysis and reviewing the results. They would like to rescale the variables in the model to more clearly reflect the relationship between the regression coefficients. Which technique could be used to rescale the variables?

Options:

A- Dimension Reduction

B- Mean Centering

C- Normalization

D- Clustering

Explanation:

Normalization is a technique that rescales the values of the variables in a data set to a common range, such as [0,1] or [-1,1]. Normalization can help reduce the effect of outliers, improve the performance of some algorithms, and make the interpretation of the regression coefficients easier and more consistent. Normalization can be done using different methods, such as min-max scaling, z-score scaling, or unit vector scaling.

Answer:

C

Question 6

Question Type: MultipleChoice

A real estate broker is tracking monthly sales between two of its teams. The results have been visualized using a Treemap chart. What is the advantage of using a Treemap chart, over a Sunburst chart to visualize the results?

Options:

A- With its colour scheme, it is easy to compare the variables within a Treemap

B- With its rectangles and straight lines, a Treemap is optimized to include more

C- A Treemap is meant to represent a hierarchical result set as opposed to a Sunburst chart

D- A Treemap shows all the hierarchical levels of data as opposed to a Sunburst chart

Explanation:

A Treemap chart is a type of chart that displays hierarchical data as a set of nested rectangles, where the size and color of each rectangle represent a quantitative value and a categorical variable, respectively1.A Sunburst chart is a type of chart that displays

hierarchical data as a set of concentric circles, where the size and color of each slice represent a quantitative value and a categorical variable, respectively2. Both charts are useful for visualizing hierarchical data structures, but they have different advantages and disadvantages.One advantage of using a Treemap chart over a Sunburst chart is that a Treemap chart is optimized to include more data points, as it uses a Cartesian coordinate system that fills the entire rectangular space of the chart area, whereas a Sunburst chart uses a polar coordinate system that leaves empty spaces in the corners of the chart area3. This means that a Treemap chart can display more levels of hierarchy, more categories, and more details than a Sunburst chart, without compromising readability or clarity.Therefore, the correct answer is B, as a Treemap chart is optimized to include more data than a Sunburst chart.

Answer:

B

Question 7

Question Type: MultipleChoice

A company wants to run a monthly promotion on batteries that cost 15 cents each and sells for 50 cents. At this price, they typically sell 1000 batteries and generate a profit of 35 cents per battery for a total profit of $350. The analytics team was asked to test two price points - 20% off (i.e. a sale price of 40 cents) and 40% off (i.e., a sale price of 30 cents). The survey data completed by 10000 participants was analyzed and showed that a 20% savings would result in sales of 1200 batteries and the 40% savings would result in 1800 batteries being sold. The team's initial recommendation was to recommend the 40% discount. Now that they are validating their recommendations, they decide to:

Options:

A- Question why management would only want them to test two price points

B- Change their recommendation realizing they have been victims of linear bias

C- Redo the survey looking for a larger sample size

D- Use their original recommendation given that the volume of sales is much higher

Explanation:

Linear bias is a type of cognitive bias that assumes a linear relationship between two variables, when in fact the relationship may be more complex or nonlinear. In this case, the analytics team assumed that the higher the discount, the higher the sales and profit, without considering other factors that may affect customer behavior, such as price elasticity, perceived quality, or competition. By changing their recommendation, the team can avoid making a suboptimal decision that may result in lower profit or customer satisfaction.

Answer:

B