DA0-001 Q&As

CompTIA Data+

Pass CompTIA DA0-001 Exam with 100% Guarantee

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

https://www.pass4itsure.com/da0-001.html

100% Passing Guarantee
100% Money Back Assurance

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

QUESTION 1

A Chief Executive Officer (CEO) is requesting more up-to-date sales data for improved visibility prior to month-end. An analyst must determine the frequency of a sales report that was previously distributed on an as-needed basis. Which of the following would be the most appropriate frequency for this report?

A. Monthly

B. Quarterly

C. Weekly

D. Every other month

Correct Answer: C

Explanation: The most appropriate frequency for the sales report is weekly, as this will provide the CEO with more up-to-date sales data for improved visibility prior to month-end. A weekly sales report can show the sales performance, trends, and issues of the sales team on a regular basis, and help the CEO to monitor and evaluate the progress and results of the sales activities. A weekly sales report can also help the CEO to identify and address any problems or opportunities that may arise during the month, and to make timely and informed decisions.

QUESTION 2

Randy scored 76 on a math test, Katie scored 86 on a science test, Ralph scored 80 on a history test, and Jean scored 80 on an English test. The table below contains the mean and standard deviation of the scores for each of the courses:

Course	Mean	Standard deviation
Math	70	2
Science	80	3
History	75	2
English	90	1

Using this information, which of the following students had the BEST score?

A. Randy

B. Katie

C. Ralph

D. Jean

Correct Answer: B

Explanation: To compare the students\\' scores, we need to standardize them by using the z-score formula, which is:

z = (x -) /

where x is the raw score, is the mean, and is the standard deviation. The z-score tells us how many standard deviations a score is above or below the mean. A higher z-score means a better score relative to the average.

Using the table, we can calculate the z-scores for each student as follows:

Randy: z = (76 - 70) / 2 = 3 Katie: z = (86 - 80) / 3 = 2 Ralph: z = (80 - 75) / 2 = 2.5 Jean: z = (80 - 90) / 1 = -10

The student with the highest z-score is Randy, with a z-score of 3. This means that Randy scored 3 standard deviations above the mean in math, which is the best performance among the four students. Therefore, the correct answer is A.

References: Comparing with z-scores (video) | Z-scores | Khan Academy, 17 Important Data Visualization Techniques | HBS Online

QUESTION 3

Which of the following BEST describes the issue in which character values are mixed with integer values in a data set column?

A. Duplicate data

B. Missing data

C. Data outliers

D. Invalid data type

Correct Answer: D

Explanation: The invalid data type is the best description for the issue in which character values are mixed with integer values in a data set column. Invalid data type means that the data does not match the expected or required format or structure for a given variable or attribute. For example, if a column is supposed to store numerical values, but some rows contain text values, then those rows have an invalid data type. References: CompTIA Data+ Certification Exam Objectives, page 10

QUESTION 4

Which of the following is a common data analytics tool that is also used as an interpreted, high-level, general-purpose programming language?

A. SAS

B. Microsoft Power BI

C. IBM SPSS

D. Python

Correct Answer: D

Explanation: Python is a common data analytics tool that is also used as an interpreted, high-level, general-purpose programming language. Python has a simple and expressive syntax that makes it easy to read and write code. Python also has a rich set of libraries and frameworks that support various tasks and applications in data analytics, such as data manipulation, visualization, machine learning, natural language processing, web scraping, and more. Some

examples of popular Python libraries for data analytics are pandas, numpy, matplotlib, seaborn, scikit-learn, nltk, and beautifulsoup. Python is different from other data analytics tools that are not programming languages but rather software applications or platforms that provide graphical user interfaces (GUIs) for data analysis and visualization. Some examples of these tools are SAS, Microsoft Power BI, IBM SPSS. Therefore, the correct answer is D. References: [What is Python? | Definition and Examples], [Python Libraries for Data Science]

QUESTION 5

Which of the following is a characteristic of a relational database?

A. It utilizes key-value pairs.

B. It has undefined fields.

C. It is structured in nature.

D. It uses minimal memory.

Correct Answer: C

Explanation: It is structured in nature. This is because a relational database is a type of database that organizes data into tables, which consist of rows and columns. A relational database is structured in nature, which means that the data has a predefined schema or format, and follows certain rules and constraints, such as primary keys, foreign keys, or referential integrity. A relational database can be used to store, query, and manipulate data using a structured query language (SQL). The other characteristics are not true for a relational database. Here is why:

It utilizes key-value pairs. This is not true for a relational database, because key-value pairs are a way of storing data that associates each value with a unique key, such as an identifier or a name. Key-value pairs are typically used in non-relational databases, such as NoSQL databases, which do not have tables, rows, or columns, but rather store data in various formats, such as documents, graphs, or columns. It has undefined fields. This is not true for a relational database, because fields are another name for columns in a table, which define the attributes or properties of each row or record in the table. Fields have defined names, types, and lengths in a relational database, which specify the format and size of the data that can be stored in each field. It uses minimal memory. This is not true for a relational database, because memory is the amount of space or storage that is used by a database to store and process data. Memory usage depends on various factors, such as the size, complexity, and number of tables and queries in a relational database. A relational database can use a lot of memory if it has many tables with many rows and columns, or if it performs complex or frequent queries on the data.

QUESTION 6

What role in a data governance is typically responsible for day-to-day oversight of data use?

A. Data processors.

B. Data custodians

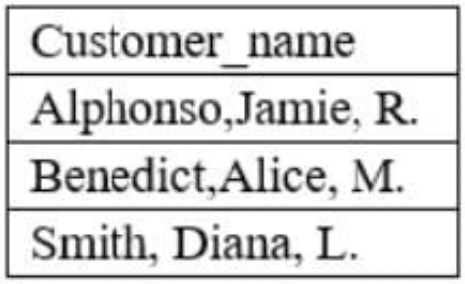
C. Data owners.

D. Data stewards.

Correct Answer: D

QUESTION 7

A data analyst must separate the column shown below into multiple columns for each component of the name: Which of the following data manipulation techniques should the analyst perform?



Customer_name
Alphonso,Jamie, R.
Benedict,Alice, M.
Smith, Diana, L.

A. Imputing

B. Transposing

C. Parsing

D. Concatenating

Correct Answer: C

Explanation: Parsing is the data manipulation technique that should be used to separate the column into multiple columns for each component of the name. Parsing is the process of breaking down a string of text into smaller units, such as words, symbols, or numbers. Parsing can be used to extract specific information from a text column, such as names, addresses, phone numbers, etc. Parsing can also be used to split a text column into multiple columns based on a delimiter, such as a comma, space, or dash1. In this case, the analyst can use parsing to split the column by the comma delimiter and create three new columns: one for the last name, one for the first name, and one for the middle initial. This will make the data more organized and easier to analyze.

QUESTION 8

A data analyst needs to create a dashboard using the company\\'s yearly revenue data sets. Which of the following would be the best way to plot the information to show the top- performing region?

A. A line chart

B. A waterfall chart

C. A heat map

D. A stacked bar chart

Correct Answer: D

QUESTION 9

Which of the following best describes the law of large numbers?

A. As a sample size decreases, its standard deviation gets closer to the average of the whole population.

B. As a sample size grows, its mean gets closer to the average of the whole population

C. As a sample size decreases, its mean gets closer to the average of the whole population.

D. When a sample size doubles. the sample is indicative of the whole population.

Correct Answer: B

The best answer is B. As a sample size grows, its mean gets closer to the average of the whole population. The law of large numbers, in probability and statistics, states that as a sample size grows, its mean gets closer to the average of the whole population. This is due to the sample being more representative of the population as it increases in size. The law of large numbers guarantees stable long-term results for the averages of some random events1 A. As a sample size decreases, its standard deviation gets closer to the average of the whole population is not correct, because it confuses the concepts of standard deviation and mean. Standard deviation is a measure of how much the values in a data set vary from the mean, not how close the mean is to the population average. Also, as a sample size decreases, its standard deviation tends to increase, not decrease, because the sample becomes less representative of the population.

C. As a sample size decreases, its mean gets closer to the average of the whole population is not correct, because it contradicts the law of large numbers. As a sample size decreases, its mean tends to deviate from the average of the whole population, because the sample becomes less representative of the population. D. When a sample size doubles, the sample is indicative of the whole population is not correct, because it does not specify how close the sample mean is to the population average. Doubling the sample size does not necessarily make the sample indicative of the whole population, unless the sample size is large enough to begin with. The law of large numbers does not state a specific number or proportion of samples that are indicative of the whole population, but rather describes how the sample mean approaches the population average as the sample size increases indefinitely.

QUESTION 10

Which of the following can be used to translate data into another form so it can only be read by a user who has a key or a password?

A. Data encryption.

B. Data transmission.

C. Data protection.

D. Data masking.

Correct Answer: A

Data encryption can be used to translate data into another form so it can only be read by a user who has a key or a password. Data encryption is a process of transforming data using an algorithm or a cipher to make it unreadable to anyone except those who have the key or the password to decrypt it. Data encryption is a common method of protecting data from unauthorized access, modification, or theft. Reference: Guide to CompTIA Data+ and Practice Questions - Pass Your Cert

QUESTION 11

Which of the following describes the method of sampling in which elements of data are selected randomly from each of the small subgroups within a population?

A. Simple random

B. Cluster

C. Systematic

D. Stratified

Correct Answer: D

Explanation: This is because stratified is a type of sampling in which elements of data are selected randomly from each of the small subgroups within a population, such as age groups, gender groups, or income groups. Stratified sampling can be used to ensure that the sample is representative and proportional of the population, as well as reduce the sampling error or bias. For example, stratified sampling can be used to select a sample of voters from different political parties based on their proportion in the population. The other types of sampling are not the types of sampling in which elements of data are selected randomly from each of the small subgroups within a population. Here is why:

Simple random is a type of sampling in which elements of data are selected randomly from the entire population, without dividing it into any subgroups. Simple random sampling can be used to ensure that every element in the population has an equal chance of being selected, as well as avoid any systematic error or bias. For example, simple random sampling can be used to select a sample of students from a school by using a lottery or a computer-generated number. Cluster is a type of sampling in which elements of data are selected randomly from a few large subgroups within a population, such as regions, districts, or schools. Cluster sampling can be used to reduce the cost and complexity of sampling, as well as increase the feasibility and convenience of sampling. For example, cluster sampling can be used to select a sample of households from a few neighborhoods by using a map or a list. Systematic is a type of sampling in which elements of data are selected at regular intervals from an ordered list or sequence within a population, such as every nth element or every kth element. Systematic sampling can be used to simplify and speed up the sampling process, as well as ensure that the sample covers the entire range or scope of the population. For example, systematic sampling can be used to select a sample of books from a library by using an alphabetical order or a numerical order.

QUESTION 12

A junior web developer is developing a new application where users can upload short videos. The first task is to create a homepage that shows the headline "Upload Your Short Videos" and a clickable button that says "upload now".

Which of the following HTML commands would help the developer to complete the task successfully?

A. Upload Your Short Videosupload now

B. Upload Your Short Videosupload now

C. Upload Your Short Videosupload now

D. Upload Your Short Videosupload now

Correct Answer: C

The HTML commands that would help the developer to complete the task successfully are

Upload Your Short Videos

and upload now. The

tag defines a heading level 1, which is the largest and most

important heading on a webpage. The tag defines a clickable button that can perform some action when clicked. The other options are not suitable for the task, as they either use the wrong tags or do not create a clickable button. The tag defines a section of text with no specific meaning or formatting. The

tag defines a paragraph of text. The tag does not exist in HTML. Reference: HTML Tags - W3Schools

---

QUESTION 13

When analyzing the values of two variables, you decide to convert both variables so they are on a scale of 0 to 1.

What term describes this action?

A. Filtering.

B. Normalization.

C. Transposition.

D. Aggregation.

Correct Answer: B

Normalization is the process of reorganizing data in a database so that it meets two basic requirements: There is no redundancy of data, all data is stored in only one place. Data

dependencies are logical, all related data items are stored together. Put simply, data normalization ensures that your data looks, reads, and can be utilized the same way across all of the records in your customer database. This is done by standardizing the formats of specific fields and records within your customer database.

QUESTION 14

Jenny wants to study the academic performance of undergraduate sophomores and wants to determine the average grade point average at different points during an academic year.

What best describes the data set she needs?

A. Sample.

B. Observation.

C. Variable.

D. Population.

Correct Answer: A

Correct answer A. Sample.

Jenny does not have data for the entire population of all undergraduate sophomores. While a specific grade point

average is an observation of variable, jenny needs sample data.

QUESTION 15

A development company is constructing a new Init in its apartment complex. The complex has the following floor plans:

Unit name	Sq. Ft.	Price	$/Sq. Ft.
Jasmine	1,000	$345,000	$345
Orchid	1,100	$425,000	$386
Azalea	1,300	$460,000	$354
Tulip	1,640	$525,000	$320
Rose	2,000		

Using the average cost per square foot of the original floor plans. which of the following should be the price of the Rose Init?

A. $640,900

B. $690,000

C. $705,200

D. $702,500

Correct Answer: D

Explanation: The correct answer is D. $702,500. To find the price of the Rose unit, we need to use the average cost per square foot of the original floor plans. The average cost per square foot is calculated by dividing the price by the square footage of each unit type. Using the data from the table, we can do the following: Jasmine: $345,000 / 1,000 = $345 per square foot Orchid: $525,000 / 2,000 = $262.5 per square foot Azalea: $375,000 / 1,500 = $250 per square foot Tulip: $450,000 / 1,800 = $250 per square foot The average cost per square foot of the original floor plans is the mean of these four values, which is ($345 + $262.5 + $250 + $250) / 4 = $276.875 per square foot. To find the price of the Rose unit, we need to multiply the average cost per square foot by the square footage of the Rose unit. The Rose unit has a square footage of 2,535, according to the table. Therefore, the price of the Rose unit is $276.875 x 2,535 = $702,421.875.

Rounding to the nearest whole number, we get $702,500 as the price of the Rose unit.