

Data+

1. A data analyst needs to calculate the mean for Q1 sales using the data set below:

Product	Q1 sales
Ground beef	\$2,667.60
Crab meet	\$1,768.41
Swiss cheese	\$3,182.40
Broccoli	\$1,509.60
Vegetable spread	\$3,202.87

Which of the following is the mean?

A. \$2,466.18

B. \$2,667.60

C. \$3,082.72

D. \$12,330.88

Answer(s): A

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

A. A workbook with multiple tabs for each region

B. A daily email with snapshots of regional summaries

C. A static report with a different page for every filtered view

D. A dashboard with filters at the top that the user can toggle

Answer(s): D

3. A customer list from a financial services company is shown below:

Name	Number of credit cards	Age	Income
Sean	0	27	\$60,000
Angela	4	31	\$50,000
Terry	3	40	\$170,000
Paula	1	25	\$70,000
Malcolm	3	28	\$150,000

A data analyst wants to create a likely-to-buy score on a scale from 0 to 100, based on an average of the three numerical variables: number of credit cards, age, and income. Which of the following should the analyst do to the variables to ensure they all have the same weight in the score calculation?

A. Recode the variables.

B. Calculate the percentiles of the variables.

C. Calculate the standard deviations of the variables.

D. Normalize the variables.

Answer(s): D

4. Which of the following actions should be taken when transmitting data to mitigate the chance of a data leak occurring? (Choose two.)

A. Data identification

B. Data processing

C. Data reporting

D. Data encryption

E. Data masking

F. Data removal

Answer(s): D E

5. A data analyst has been asked to organize the table below in the following ways:

By sales from high to low

By state in alphabetic order

First_name	Last_name	Address	City	State	Sales
Ed	Edens	2851 N. Southport	Chicago	IL	\$125,689
Pat	Mudd	710 Bridle Ridge Road	Eagan	MN	\$101,259
Katie	Hofstad	2851 S. Windwood Lane	Rosemount	NY	\$105,779
Edward	Frank	281 S. Northport	Chicago	IL	\$456,231
Rachel	Newman	305 Big Timber Trail	Wheaton	CO	\$99,876
Kaylyn	Korth	332 Richfield Drive	Lakeview	MN	\$166,874

Which of the following functions will allow the data analyst to organize the table in this manner?

A. Conditional formatting

B. Grouping

C. Filtering

D. Sorting

Answer(s): D

6. 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

Answer(s): D

7. Which of the following is a process that is used during data integration to collect, blend, and load data?

A. MDM

B. ETL

C. OLTP

D. BI

Answer(s): B

8. An analyst has received the requirements for an internal user dashboard. The analyst confirms the data sources and then creates a wireframe. Which of the following is the NEXT step the analyst should take in the dashboard creation process?

A. Optimize the dashboard.

B. Create subscriptions.

C. Get stakeholder approval.

D. Deploy to production.

Answer(s): C

9. A data analyst has been asked to derive a new variable labeled "Promotion_flag" based on the total quantity sold by each salesperson. Given the table below:

Store_ID	Item	Salesperson	Quantity_sold	Promotion_flag
104	Pax-2	James	1,000,300	
204	Pax-3	Paul	234,578	
304	Pax-1	Peter	2,000,432	
404	Pax-2	Esther	1,089,678	
204	Pax-3	May	126,578	
304	Pax-1	Park	200,432	
404	Pax-2	Mabel	1,089,000	

Which of the following functions would the analyst consider appropriate to flag “Yes” for every salesperson who has a number above 1,000,000 in the Quantity_sold column?

A. Date

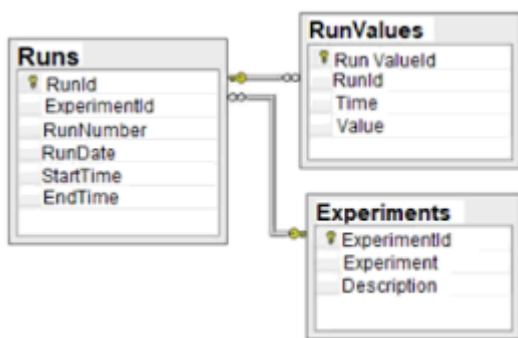
B. Mathematical

C. Logical

D. Aggregate

Answer(s): C

10. Given the diagram below:



Which of the following data schemas shown?

A. Key-value pairs

B. Online transactional processing

C. Data lake

D. Relational database

Answer(s): D

11. A company’s marketing department wants to do a promotional campaign next month. A data analyst on the team has been asked to perform customer segmentation, looking at how recently a customer bought product, at what frequency, and at what value. Which of the following types of analysis would this practice be considered?

A. Prescriptive

B. Trend

C. Gap

D. Custer

Answer(s): D

12. A publishing group has requested a dashboard to track submissions before publication. A key requirement is that all changes are tracked, as multiple users will be checking out documents and editing them before submissions are considered final. Which of the following is the BEST way to meet this stakeholder requirement?

A. Display the version number next to each submission on the dashboard.

B. Present a data refresh date at the top of the dashboard.

C. Confirm the dashboard is adhering to the corporate style guide.

D. Use permissions to ensure users only see certain versions of the submissions.

Answer(s): A

13. The number of phone calls that call center receives in a day is an example of:

A. continuous data.

B. categorical data.

C. ordinal data.

D. discrete data.

Answer(s): D

14. A data analyst is asked to create a sales report for the second-quarter 2020 board meeting, which will include a review of the business's performance through the second quarter. The board meeting will be held on July 15, 2020, after the numbers are finalized. Which of the following report types should the data analyst create?

A. Static

B. Real-time

C. Self-service

D. Dynamic

Answer(s): A

15. Which of the following would be considered non-personally identifiable information?

A. Cell phone device name

B. Customer's name

C. Government ID number

D. Telephone number

Answer(s): A

16. Which of the following is the correct data type for text?

A. Boolean

B. String

C. Integer

D. Float

Answer(s): B

17. Which of the following should be accomplished NEXT after understanding a business requirement for a data analysis report?

A. Rephrase the business requirement.

B. Determine the data necessary for the analysis.

C. Build a mock dashboard/presentation layout.

D. Perform exploratory data analysis.

Answer(s): B

18. 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

Answer(s): D

19. A data analyst needs to present the results of an online marketing campaign to the marketing manager. The manager wants to see the most important KPIs and measure the return on marketing investment. Which of the following should the data analyst use to BEST communicate this information to the manager?

A. A real-time monitor that allows the manager to view performance the day the campaign was launched

B. A self-service dashboard that allows the manager to look at the company's annual budget performance

C. A spreadsheet of the raw data from all marketing campaigns and channels

D. A summary with statistics, conclusions, and recommendations from the data analyst

Answer(s): D

20. A data analyst for a media company needs to determine the most popular movie genre. Given the table below:

MovieID	Name	Genre	Actors	Rating
01	Ghost Writer	Comedy, Actions	Joshua Wellington, Susana Summons	6.5
02	Life of Suffering	Drama, Foreign, Historical	Shelly May, Rita Moralle, Ethan Warner, Sean Houser	7.2

Which of the following must be done to the Genre column before this task can be completed?

A. Append

B. Merge

C. Concatenate

D. Delimit

Answer(s): D
