

Data Science Fundamentals

Data Science · Practice Test · 15 Questions

1. What is a common way to represent the relationship between two numerical variables in data science?

- A) Bar Chart
- B) Line Graph
- C) Scatter Plot
- D) Pie Chart

2. Which of these is an example of 'categorical' data?

- A) Height in centimeters
- B) Temperature in Celsius
- C) Color of a car
- D) Number of steps taken

3. In data science, what does 'cleaning' or 'wrangling' data typically involve?

- A) Creating new datasets from scratch
- B) Removing irrelevant or incorrect data points
- C) Developing complex algorithms
- D) Designing user interfaces

4. What is the primary purpose of a 'histogram' in data visualization?

- A) To show parts of a whole
- B) To display trends over time
- C) To illustrate the distribution of a single numerical variable
- D) To compare values across different categories

5. Which programming language is widely used in data science for its libraries like Pandas and NumPy?

- A) Java
- B) C++
- C) Python
- D) Ruby

6. What does 'machine learning' enable computers to do with data?

- A) Only perform basic calculations
- B) Learn from data without being explicitly programmed
- C) Store data permanently
- D) Connect to the internet

7. What is the 'mean' of a dataset?

- A) The most frequent value
- B) The middle value when data is ordered
- C) The sum of all values divided by the number of values
- D) The range between the highest and lowest values

8. Which of the following is a type of 'structured' data?

- A) An email message
- B) A social media post
- C) A customer database table
- D) A handwritten note

9. What is the term for analyzing data to find patterns and insights that can help make decisions?

- A) Data Entry
- B) Data Storage
- C) Data Analysis
- D) Data Deletion

10. A 'correlation' between two variables indicates:

- A) One variable directly causes the other
- B) The variables tend to change together
- C) There is no relationship between them
- D) One variable is always greater than the other

11. What is the main goal of 'data visualization'?

- A) To make data harder to understand
- B) To present data in a clear and effective graphical format
- C) To delete unnecessary data
- D) To speed up data entry

12. Which type of data is represented by numbers that can be measured or counted?

- A) Categorical Data
- B) Textual Data
- C) Numerical Data
- D) Boolean Data

13. What is a 'dataset' in the context of data science?

- A) A single number
- B) A collection of related data points
- C) A computer program
- D) A mathematical formula

14. Which of these is an example of 'ordinal' data?

- A) City name
- B) Customer satisfaction rating (e.g., 'Poor', 'Fair', 'Good')
- C) Number of items purchased
- D) Yes/No response

15. What is a 'feature' in a dataset?

- A) A unique identifier for a row
- B) A single data point
- C) A measurable characteristic or attribute of an observation
- D) The entire dataset