

Definition and Applications

1. How would you define data science, and what sets it apart from traditional data analysis?
2. What are some key real-world applications of data science across various industries?
3. How does data science contribute to improving business operations and efficiency?
4. Can you provide an example of how data science is used in fields like healthcare or finance?

Basic Concepts

5. What is the role of data in the data science pipeline, and how is it collected and processed?
6. Why is data quality critical for ensuring reliable insights in data science projects?
7. How does domain knowledge shape the interpretation and application of data science models?
8. How do mathematics and statistics support the algorithms and techniques used in data science?
9. How is probability theory applied to handle uncertainty and risk in data science models?
10. In what ways does business knowledge influence the development of data science solutions?
11. How do data, information, and knowledge interact in the data science process?
12. Why is an understanding of linear algebra crucial for developing machine learning models?

Data Science Process

13. What are the major steps involved in a typical data science project?
14. How does data preprocessing impact the outcome of the entire data science process?

15. What is the purpose of exploratory data analysis (EDA), and how does it guide further analysis?
16. Why is data validation an essential step in building reliable data science models?

Importance in Decision-Making Process

17. How does data science facilitate data-driven decision-making in organizations?
18. What are the key advantages of using data science for improving business strategies?
19. How do predictive analytics assist businesses in forecasting future trends and behaviors?

Data Science Methodologies

20. What are some commonly used methodologies in data science, and how are they applied in practice?
21. How does the CRISP-DM framework guide data science projects from start to finish?
22. How is the agile methodology adapted in data science, and why is it effective for iterative improvements?

Principal Tasks

23. What are the primary tasks a data scientist performs during the course of a project?
24. How does model evaluation and tuning fit into the overall data science workflow, and why is it important?

Tools and Technologies

25. What are some essential tools and technologies used by data scientists for data collection, analysis, and visualization?
26. How do platforms like Python, R, or SQL contribute to efficient data manipulation and modeling in data science?