Chapter 4 – Data Science (SLO Type Questions & Answers)

Prepared for Easy Understanding and Exam Practice

1. What is Data Science?

Data Science is the field that deals with collecting, cleaning, analyzing, and interpreting data to find useful information. It helps in understanding large amounts of data and supports decision-making in business, education, health, and other fields. It combines knowledge from computer science, statistics, and machine learning.

2. What is the scope of Data Science?

The scope of Data Science is very wide. It is used in almost every field like business, healthcare, finance, and education. As more data is created daily, more experts are needed to study it and help companies make smart, data-driven decisions.

3. How is Data Science used in Business and Marketing?

In business, Data Science helps companies understand customer behavior and improve sales. It is used to study market trends, segment customers into groups, and plan better marketing strategies to attract more buyers.

4. How does Data Science help in Healthcare?

In healthcare, Data Science is used to predict diseases and create personalized treatment plans for patients. It also helps doctors manage patient data more efficiently and make faster, more accurate medical decisions.

5. What is Artificial Intelligence (AI)?

Artificial Intelligence means making machines smart so they can think, learn, and act like humans. AI helps computers understand human language, make decisions, and even recognize faces or voices. Examples include chatbots, self-driving cars, and smart assistants like Siri or Alexa.

6. Who introduced the concept of the Turing Test and why?

The Turing Test was introduced by the British mathematician Alan Turing in 1950. It was designed to test if a machine could act so intelligently that people would not be able to tell whether they are talking to a human or a computer.

7. What are some areas where AI is used today?

AI is used in many areas like decision-making systems, healthcare diagnosis, robotics, smart devices, natural language processing, and computer vision. It helps make daily life easier through automation and smart technology.

8. What is Machine Learning (ML)?

Machine Learning is a branch of Artificial Intelligence that allows computers to learn from data and improve over time. Instead of being programmed for every task, machines learn patterns from data and use them to make predictions or decisions automatically.

9. What is the role of Machine Learning in real life?

Machine Learning is used in things we use every day — like YouTube recommending videos, Netflix suggesting movies, or Google predicting search results. It is also used in banking for fraud detection and in healthcare for early diagnosis.

10. What is Supervised Learning?

Supervised learning is a type of machine learning where the computer learns from data that already has correct answers or labels. For example, when we train a model with pictures of cats and dogs labeled correctly, the computer learns to identify them in new images.

11. What is Unsupervised Learning?

Unsupervised learning means the computer gets data without labels and tries to find patterns or groups by itself. It is used for tasks like grouping customers based on their buying habits or organizing similar news articles together.

12. What is Reinforcement Learning?

Reinforcement learning is when a computer learns by doing actions and receiving feedback in the form of rewards or penalties. Over time, it learns which actions are good. For example, a robot learns to walk by getting rewards for correct steps and penalties for falling.

13. What are the main differences between AI, ML, and Data Science?

AI focuses on making machines act like humans. ML helps machines learn automatically from data. Data Science deals with collecting and analyzing data to find insights. Together, they form the base of modern technology and intelligent systems.

14. What are the main skills needed for Data Science, AI, and ML?

You need skills like programming (Python or R), knowledge of machine learning algorithms, understanding of deep learning and neural networks, data cleaning, and the ability to analyze and visualize data effectively.

15. What is Data Visualization?

Data Visualization means showing data through graphs, charts, and maps instead of numbers. It helps people understand complex data easily and see

patterns or trends quickly. It makes information clear for both experts and non-technical people.

16. What are common methods of Data Visualization?

Common methods include bar charts, line charts, pie charts, scatter plots, heatmaps, and bubble charts. Each type of chart helps present data differently — for example, line charts show trends over time, and pie charts show parts of a whole.

17. What are the types of Data Visualization?

There are several types: Quantitative Visualization (for numbers), Categorical Visualization (for groups), Temporal Visualization (for time-based data), Spatial Visualization (for location data), Interactive Visualization, and Statistical Visualization.

18. Why is Data Visualization important?

It helps turn complex data into visuals that are easy to read and understand. Businesses, teachers, doctors, and scientists use it to make faster and better decisions by clearly seeing results and trends.

19. What are the advantages of Data Visualization?

It saves time, improves understanding, and helps in making accurate decisions. It also supports teamwork, as everyone can easily see the information through charts instead of long reports.

20. What is the relationship between Database and Machine Learning?

Databases store clean and organized data in tables. Machine Learning uses that data to learn patterns and make predictions. Without proper databases, machine learning models cannot work accurately.

21. What is a Database?

A database is an organized collection of data stored in tables with rows and columns. It helps store, manage, and retrieve information easily using tools like SQL, MySQL, or Oracle.

22. What is the Data Science Life Cycle?

The Data Science Life Cycle is a process that shows the steps used to solve problems using data. It starts with defining the problem and ends with monitoring the model after it is used in real life.

23. What happens in the Problem Definition stage?

In this stage, we clearly understand what problem we need to solve. We also set goals, called Key Performance Indicators (KPIs), to measure success and decide the scope of the work.

24. What is done during Data Collection?

We gather data from different sources such as surveys, experiments, databases, or online forms. The data must be correct, relevant, and sufficient to solve the problem.

25. What is Data Cleaning (Preprocessing)?

Data Cleaning means removing errors, duplicates, or missing values from the data. It ensures that the data is accurate before analysis because wrong data gives wrong results — 'Garbage in, garbage out.'

26. What happens in Data Analysis?

In this stage, we study the data using statistics and visual tools like graphs or charts to find patterns, relationships, or trends. It helps us test ideas and get insights from the data.

27. What is Data Modeling?

Data Modeling means organizing the data into structures or models that make it easy to understand and predict. It helps build systems that can analyze and give useful results from the data.

28. What is Model Evaluation?

Model Evaluation checks how well the model is working. We test its accuracy and reliability. If the model doesn't perform well, we adjust or improve it to get better results.

29. What is Model Deployment?

Model Deployment means putting the model into real use — for example, adding it to an app or website where it works with real data and gives real-time predictions or insights.

30. What is Maintenance and Monitoring in Data Science?

After deployment, the model needs regular checking to make sure it still works correctly. We update it when new data comes or user needs change. This helps keep the model accurate over time.

31. Why is Maintenance important in Data Science projects?

Because data and trends change with time. If the model is not updated, it may start giving wrong results. Continuous maintenance ensures accuracy and long-term success.

THE END