Data Science and Data Gathering

MCQs and Short Questions

Q1. What is meant by data?

Raw facts and figures is called data, after that process data is called information.

Q2. What are the two broad categories of data?

The two categories of data are qualitative and quantitative.

Q3. What is the difference between qualitative and quantitative data?

Qualitative data	Quantitative data
Qualitative data describes qualities or characteristics (e.g., color, taste, opinion).	Quantitative data consists of numbers or measurements (e.g., height, weight, age).

Q4. What are the key characteristics of qualitative data?

Qualitative data is non-numeric and categorical.

Q5. What is meant by non-numeric data?

Non-numeric data is represented by words, labels, or symbols instead of numbers, e.g., names of students in class (Ali, Ahmed, Usman, Qasim etc).

Q6. What is categorical data?

Categorical data can be divided into groups or classes, e.g., type of fruits, job titles, hair colors.

Q7. What are the types of qualitative data?

The types of qualitative data are nominal data and ordinal data.

Q8. What is the difference between nominal and ordinal data?

Nominal data	Ordinal data
Nominal data categorizes items without implying any order, e.g., gender, types of fruits, or colors.	Ordinal data categorizes items with a meaningful order, e.g., education levels or shirt sizes.

Q9. What are the key characteristics of quantitative data?

Key characteristics of quantitative data are numerical, measurable, countable and arithmetical.

Q10. What is numerical data?

Numerical data is expressed in numbers, e.g., heights in centimeters, weights in kilograms.

Q11. What is measurable data?

Measurable data can be measured using instruments or tools, e.g., a ruler for length, a scale for weight.

Q12. What is countable data?

Countable data can be counted or enumerated, e.g., number of students in a class.

Q13. What is meant by arithmetical data?

Arithmetical data can be used in arithmetic operations like multiplication and addition.

Q14. What are the main types of quantitative data?

The main types of quantitative data are discrete data and continuous data.

Q15. What is the difference between discrete and continuous data?

Discrete data	Continuous data
Discrete data consists of distinct, separate values, often in whole numbers, e.g., number of oranges in a basket.	Continuous data is consists of measurable value within a given range, including fractions or decimals e.g., student heights (150.5 cm, 161.3 cm).

Q16. What operations can be performed on discrete data?

Logical, grouping, arithmetic (addition, subtraction), and statistical operations (average, range) can be performed on discrete data.

Q17: What operations can be performed on continuous data?

Logical, grouping, arithmetic, and division operations can be performed.

Q18. Give an example of continuous data and explain why it is considered continuous.

Height of students in school. Height can take any value within a range (for example, 150.5 cm, 150.75 cm), so it is continuous because it can be measured, not just counted.

Q19. What type of data is the number of students in your class?

It is quantitative discrete data because the number of students can only be whole numbers.

Imran Shaukat SST (IT)
Govt. Gulberg High School,
Lahore Cantt

MCQs and Short Questions

Q20. Why is it important to organize data into tables or charts before analyzing it?

Organizing data into tables or charts helps to reduce errors, saves time, present information clearly, and identify patterns to make simple for analyzing.

Q21. What are charts used for?

Graphs are used for visuals representation of complex data for easy to understand.

Q22. What are graphs used for?

Graphs are used to represent data and show relationships between different data points.

Q23. What is data collection?

Data collection is the process of gathering information to answer questions or make decisions.

Q24. What are the methods of data collection?

Data collection methods are surveys, questionnaires, interviews, observations and online data sources.

Q25. Which method would you use to collect opinions from a large group of people about a new school policy?

A survey or questionnaire would be used to collect opinions from a large group of people efficiently.

Q26. What is one advantage of using online tools like Google Forms for collecting survey data?

The advantage of Google Forms to collect and organize responses from many people quickly and easily. This data is automatically stored and can be easily analyzed.

Q27. What is data gathering and what methods are used?

Data gathering means collecting information from various places for research or school projects. Two methods are used online databases and online resources.

Q28. What is data extraction?

Data extraction is finding and saving the most relevant information from a large set of data.

Q29. What is meant b data integration?

The process of combining information from different sources to get a complete view.

Q30. Why might you need to integrate data from different sources when working on a project?

You might need to integrate data from different sources to get a complete and accurate picture of the topic, compare results, or make better decisions based on more information.

Q31. Describe a scenario where discrete data might be more useful than continuous data.

A scenario could be counting the number of students in each class. Discrete data is more useful because it deals with whole numbers that can be counted, not measured.

Q32. Describe a scenario where discrete data might be more useful than continuous data.

A scenario could be counting the number of students in each class. Discrete data is more useful because it deals with whole numbers that can be counted, not measured.

Q33. Differentiate structured data and unstructured data?

Structured data	Unstructured data
Structured data is organized and formatted so it can be easily searched and analyzed. e.g., spreadsheets, databases.	Unstructured data is free-form and does not fit into a specific format. e.g., Emails and social media posts.

Q34. What is meant by data visualization?

Data visualization is the process of turning numbers and information into pictures.

Q35. Explain why data visualization is important. How does it help in understanding complex information?

Data visualization is important because it presents data in a clear, visual form such as graphs or charts. It helps people quickly understand patterns, trends, and relationships in complex data.

Q36. Describe what a line graph is used for and provide an example of data that could be displayed using a line graph.

A line graph is used to show changes or trends over time. e.g., showing how the temperature changes during different hours of the day.

Imran Shaukat SST (IT)
Govt. Gulberg High School,
Lahore Cantt

MCQs and Short Questions

Q37. Explain the use of scatter plots in visualizing continuous data. Provide an example of a situation where a scatter plot would be useful.

Scatter plots are used to show the relationship between two continuous variables. e.g., comparing student weight with student height.

Q38. What is data pre-processing?

Data pre-processing is the process of preparing raw data for analysis by cleaning, organizing, and arranging it so that it becomes accurate and ready to use.

Q39. Why is data pre-processing important?

It is important because it removes errors and inconsistencies, making the data reliable and accurate for analysis.

Q40. What is meant by evaluating data quality?

Evaluating data quality means checking if the data is accurate, complete, consistent, and up-to-date before using it for analysis.

Q41. What are errors in data?

Errors are mistakes in the data, such as entering wrong or invalid values. e.g., if the maximum score is 100 but a student's score is entered as 105, it is an error.

Q42. What are outliers?

Outliers are unusual or extreme values that do not fit the general pattern of the data. e.g., if most students scored between 50 and 80, but one student scored 5 or 99, that score is an outlier.

Q43. What are biases in data?

Biases are systematic errors that cause certain results to be misrepresented in the data.

Q44. What is the use of cloud storage for data management?

Cloud storage provides facilities to store, access and share information for data management.

Q45. What did you know about Remote Access?

Remote access is the ability to connect and use computer or network from distant location. The following activities are performed by using remote access. Save your work, Access from Anywhere, Continue Working.

Q46. What is meant by data backup?

Data backup is another facility of cloud service to copies of important files store separately to protect against data loss.

Q47. What is collaborative authoring?

Cloud service provides collaborative authority for multiple people work together to create, edit, and improve a project in real time.

Q48. What are the benefits of collaborative tools?

Benefits of collaborative tools are enhanced productivity and version control.

Q49. What is data science?

Data science is the study of data to make better decisions and solve problems using computer skills and math.

Q50. Why is it important to know about data science?

It helps us make better decisions in daily life by analyzing information.

Q51. Why data science is called interdisciplinary?

Data science combines different fields to handle and analyze data effectively. These fields are Computer Science, Mathematics and statistics and Business Knowledge.

Q52. What is meant by Data Science Workflow?

Data Science Workflow is the systematic process used by data scientists to extract insights and knowledge from data.