CHAPTER 9:

Data Science and Data Gathering

SOLVED EXERCISE

Tick (✓) the correct answer.

1. **What is data?**
a) Processed information
b) Raw facts gathered about things
c) A collection of numbers only.
d) A list of observed events

2. **Which of the following is an example of qualitative data?**
a) Temperature readings in degrees Celsius
b) Number of students in a class
c) Favourite ice cream flavours
d) Test scores out of 100

3. **What type of data involves distinct, separate values that are countable?**
a) Nominal Data
b) Ordinal Data
c) Discrete Data
d) Continuous Data

4. **What is an example of continuous data?**
a) Number of cars in a parking lot
b) Height of students in centimetres
c) Types of fruits
d) Shirt sizes (small, medium, large)

5. **What type of data is used to categorize items without implying any order?**
a) Ordinal Data
b) Discrete Data
c) Nominal Data
d) Continuous Data

6. **How can you organize data to make it easier to analyze?**
a) By writing it in long paragraphs
b) By creating tables, charts, and graphs
c) By storing it in random files
d) By keeping it in a messy notebook

7. **Which tool can be used to create surveys online?**
a) Microsoft Word
b) Google Forms
c) Excel Spreadsheets
d) Adobe Photoshop

8. **What is the main purpose of data collection?**
a) To create random numbers
b) To gather information to answer questions or make decisions
c) To delete old data
d) To format text documents

9. **What is the primary purpose of data visualization?**
a) To generate random numbers
b) To convert text into data
c) To make data easier to understand by turning it into pictures
d) To hide complex data.

10. Which tool is specifically designed for creating detailed and interactive visualizations?
a) Microsoft Excel
b) Google Sheets
c) Tableau
d) PowerPoint

11. What is the first step in the data science process?
a) Data Cleaning
b) Data Analysis
c) Data Collection
d) Understanding the problem

12. What does the 'Volume' characteristic of Big Data refer to?
a) The speed at which data is generated
b) The different forms data can take
c) The sheer amount of data being collected
d) The way data is processed

13. What is an outlier in a dataset?
a) The most frequent value
b) The average of all values
c) An unusual or extreme value that doesn't fit the pattern
d) The middle value when all values are arranged in order

14. What does data encryption do?
a) It converts data into a code to prevent unauthorized access.
b) It makes data available to everyone online.
c) It automatically deletes old data.
d) It speeds up internet connection.

1	2	3	4	5	6	7	8	9	10
b	c	c	b	c	b	b	b	c	c

11	12	13	14
d	c	c	a

Short Answer Questions

1. **What is the difference between qualitative and quantitative data?**
Ans: **Qualitative data:** Qualitative data refer to categories or labels used to describe the qualities or characteristics of something rather than its quantity. This type of data offer a way to categorize and provide insights into opinions, behaviours, and experiences through descriptions rather than numbers.

Quantitative Data: Quantitative data consists of numbers used to measure the quantity or amount of something. These data types answer questions like "How much?" or "How long?" and can be useful for mathematical calculations and statistical analyses.

2. **Give an example of continuous data and explain why it is considered continuous.**
Ans: Continuous data are used when measuring items and can help answer questions about the extent or duration of something. For example, measuring

the height of a plant might reveal that it is 32.5 centimetres tall. If you measure how long it takes for the plant to grow from 20 to 30 centimetres, might show that it took 15 days. These measurements provide detailed information about the amount (height) or duration (growth time), helping us understand how much or how long something occurs.

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

Ans: I will use survey method because by using a survey method, I can collect valuable opinions and feedback from a large group of people about the new school policy. This method is effective for gathering diverse viewpoints.

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

Ans: The number of students in my class is **quantitative data**, specifically **discrete data**. Quantitative data refers to numerical data that can be measured or counted. It can be expressed in terms of numbers, quantities, or amounts.

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

Ans: •Proper organization saves time.

• When data is organized, it's easier to find and analyze.

Organized data improves clarity.

• By organizing data into tables, charts, and graphs, you can quickly grasp what the data is saying, making it simpler to draw conclusions and make decisions.

• To make data easy to understand, it can be organized into tables, charts, and graphs.

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

Ans: One significant advantage of using online tools like Google Forms for collecting survey data is that they automatically collect and organize the data in a spreadsheet, making it easier to analyze and interpret the results. With Google Forms, responses are instantly recorded in a linked Google Sheets document.

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

Ans: Integrating data from different sources is often necessary when working on a project to gain a comprehensive and accurate understanding of the topic. It increases the efficiency and also helps in better decision making.

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

Ans: A scenario where discrete data is more useful than continuous data is when dealing with attendance of students in a classroom. Discrete data makes it easier to analyze and compare attendance patterns across different students, classes, or schools.

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

Ans: Data visualization is important because it makes understanding data much easier. Imagine trying to understand a list of hundreds of numbers. It's tough, right? But if those numbers are displayed as a bar chart, it's much easier to see what's going on. Visualizations provide quick insights, allowing you to grasp what the data is saying without having to read through all the details. This is particularly useful for making better and faster decisions. For example, a business can quickly see which products are selling well and which are not.

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

Ans: A line graph is used to display data trends over time or to show the relationship between two continuous variables. It is particularly effective for visualizing changes, patterns, or progresses. For example a line graph could be used to display the daily temperatures over a week. This would clearly show how the temperature rises, falls, or remains steady across the days, making it easy to observe trends.

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

Ans: A scatter plot is a type of data visualization that displays the relationship between two continuous variables. It is a useful tool for exploring and understanding the correlation between two variables. For example a scatter plot would be useful to analyze the relationship between hours of study and exam scores for a group of students.

In this example, the scatter plot is useful because it:

- Helps to identify the relationship between study time and exam score.
- Reveals patterns and outliers in the data.
- Provides a visual representation of the data, making it easier to understand and communicate the findings.

Long Questions

1. Explain the differences between qualitative and quantitative data. Provide examples of each type.

Ans: Qualitative Data:

Qualitative data refer to categories or labels used to describe the qualities or characteristics of something rather than its quantity. This type of data offers a way to categorize and provide insights into opinions, behaviours, and experiences through descriptions rather than numbers. Key characteristics of qualitative data include non-numeric, descriptive, and categorical attributes.

• Non-Numeric: Qualitative data is represented by words, labels, or symbols instead of numbers. It describes attributes rather than quantities. Examples

include the names of students in a class (e.g., Ali, Badar, Qasim) and the colours of cars in a parking lot (e.g., red, blue, green), where "name" and "colour" are attributes.

• Categorical: Qualitative data can be into categories or classes based on their characteristics. Examples include types of fruit (e.g., apple, banana, orange), job titles (e.g., manager, engineer, accountant), and book genres (e.g., fiction, non-fiction, mystery)

Quantitative Data:

Quantitative data consists of numbers used to measure the quantity or amount of something. These data types answer questions like "How much?" or "How long?" and can be useful for mathematical calculations and statistical analyses. Key characteristics of quantitative data include being numerical, measurable, countable, and arithmetical.

• Numerical: Quantitative data is expressed in numbers, representing a measurable quantity. Examples include heights in centimetres, weights in kilograms, and test scores.

• Measurable: Quantitative data can be measured using instruments or tools. Examples include using a ruler for length, a scale for weight, and a thermometer for temperature.

• Countable: Quantitative data can be counted or enumerated, particularly for discrete data. Examples include the number of students in a class or the number of cars in a parking lot.

• Arithmetical: Quantitative data can be used in arithmetic operations. For instance, multiplying the unit price of a fruit by its weight to find the total price, or calculating annual school fees by multiplying monthly fees by the number of months.

2. Describe the process of conducting a survey to gather data about students'favourite extracurricular activities.

Ans: To gather data about students' favorite extracurricular activities. We will follow the following process:

Step 1: Define the Objectives:

It is obvious that the purpose of our survey is to collect data about students' favorite extracurricular activities. First of all we will identify the target population, and then we will decide a sample size, and the time frame for the survey.

Step 2: Develop the Survey Tool:

Develop a survey questionnaire that includes relevant questions to collect the required data. The questionnaire should include:

- Questions about age, grade, gender etc.
- Questions about extracurricular activities (e.g., sports, clubs, music programs)
- Questions about students' preferences and interests.
- Open-ended questions for additional comments or suggestions.

Step 3: Pilot Testing:

Conduct a pilot test with a small group of students to:

- Ensure the survey questions are clear and understandable
- Identify any biases or ambiguities in the questions
- Refine the survey instrument as needed

Step 4: Conduct the Survey:

Distribute the survey to the target population, either:

- Online (e.g., Google Forms, SurveyMonkey)
- Face to face (e.g., interview)
- Through email or social media

Step 5: Data Organization:

Collect the survey responses and arrange the data by:

- Checking for completeness and accuracy
- Handling missing or incomplete data
- Coding and categorizing responses as needed

Step 6: Data Analysis:

- Analyze the data using statistical methods and data visualization techniques to:
- Identify trends and patterns in students' preferences
- Compare preferences across different demographic groups
- Draw conclusions about the most popular extracurricular activities

Step 7: Findings:

Interpret the results in the context of the research question and objectives. Report the findings in a clear and concise manner, using visual aids and summaries to facilitate understanding.

Step 8: Evaluation:

Evaluate the survey process and refine it for future surveys. Identify areas for improvement, such as survey design, administration, or analysis.

3. Compare and contrast continuous and discrete data. Use examples to show how each type of data might be used in a school setting, such as measuring student performance or tracking attendance.

Ans:
Comparison between Continuous and Discrete Data:

Feature	Continuous Data	Discrete Data
Definition	Continuous data consists of values that can take any number within a given range,	Discrete data consist of distinct, separate values that are countable, often in whole numbers. Discrete data answer questions like "How many?" or "How often?"
Purpose	Measures the extent or duration of something	Counts specific items or occurrences.
Representation	Numbers, including fractions and decimals	Numbers, whole or discrete values.
Characteristics	Measurable, can include fractions or decimals.	Countable, often whole numbers.
Examples in School	Test scores (e.g., 75.5%), Height of students (e.g., 156.3 cm)	Number of students in a class Count of books in a library.
Operations	Equality, grouping, arithmetic operations including division	Equality, grouping, counting, arithmetic operations without division
Statistical Analysis	Mean, median, range, standard deviation.	Mean, median, range, standard deviation.
Usage	Measuring precise amounts, durations, or extents.	Counting specific occurrences or items
Measurement Type	Measured quantities (e.g., time, temperature).	Counted quantities (e.g., attendance, grades as whole numbers).
Visualization Tools	Line graphs, histograms.	Bar graphs, pie charts.

School Setting Examples:
1. Measuring Student Performance
Continuous Data:

...king test scores over time. A student might score 83.5% on one ...ther, showing a little change in performance.

Con...
Example: Trac...
test and 87.5% o...

Use: To analyze trends or compute averages with high precision.

Visualization: A line graph could illustrate how scores improve or decline across multiple tests.

Discrete Data:

Example: Counting the number of students who scored above a required grade, such as "A" or "B."

Use: To understand how many students fall into each grade category.

Visualization: A bar graph or pie chart showing the distribution of grades.

2. Tracking Attendance

Continuous Data:

Example: Measuring the total time each student spends in school (e.g., 6 hours per day).

Use: To evaluate trends in punctuality or late arrivals.

Visualization: A histogram showing the distribution of hours attended.

Discrete Data:

Example: Counting the number of students present each day (e.g., 45 on Monday, 40 on Tuesday).

Use: To calculate average daily attendance or detect patterns in absenteeism.

Visualization: A bar graph showing attendance by day of the week.

Both types of data are valuable in a school setting, but their usage depends on the context:

• Continuous data is ideal for measuring and analyzing precise, variable factors like performance or time.

• Discrete data is better for counting and categorizing items such as student numbers or event occurrences.

4. Analyze the benefits and challenges of using digital tools like Google Forms for data collection.

Ans: Benefits of Using Digital Tools like Google Forms for Data Collection

1. Efficiency and Convenience:

Digital tools like Google Forms allow respondents to complete surveys at their own pace, anytime, and from any location with an internet connection. Surveys can be shared via email reaching a large audience directly. Responses are recorded automatically and can be accessed immediately for analysis.

2. Cost-Effectiveness:

Digital surveys eliminate printing and distribution costs, making them a cost-effective option. No expenses for physical storage as data are stored digitally.

3. Accessibility:

Can be accessed from any device with an internet connection, increasing reach to participants.

Accommodates diverse users with features like multiple languages or accessibility modes.

4. Data Organization and Analysis:
Automatically compiles responses into spreadsheets for easy organization.
Provides built-in analytics, such as response summaries, charts, and graphs.
Export capabilities allow integration with other analysis tools like Excel or statistical software.

5. Environment Friendly:
Reduces paper waste and minimizes the carbon footprint associated with traditional paper-based surveys.

Challenges of Using Digital Tools like Google Forms

1. Limited Accessibility:
Some respondents may not have access to the internet, computers, or mobile devices, limiting the survey's reach.

2. Technical Issues:
Technical Issues: Respondents may experience technical difficulties, such as connectivity problems or browser incompatibility.

3. Data Quality:
Respondents may provide incomplete, inaccurate, or fake data, compromising the survey's validity.

4. Privacy and Security Concerns:
Storing data online raises concerns about unauthorized access or data breaches.

5. Lack of Personal Interaction:
Digital surveys may lack the personal interaction and nuance of in-person or phone interviews.

6. Survey Fatigue:
Respondents may experience survey fatigue, leading to decreased response rates or lower-quality data.

Conclusion:
Digital tools like Google Forms are highly efficient and accessible for most data collection scenarios, especially when large-scale participation is required.

5. Imagine you are tasked with organizing a school event and need to collect data on students' preferences for activities and refreshments.

Ans: Steps to Collect Data on Students' Preferences for Activities and Refreshments

1. Define Objectives:
Determine the purpose of the data collection:
- Identify the most popular activities to include in the event.
- Understand preferences for refreshments (e.g., snacks, drinks).

2. Design the Survey:
Create a simple and engaging survey that captures key information. Use a mix of:

Closed-ended questions for quick analysis:

> "Which type of activities do you prefer? (Select one): Sports, Music, Art, Games, Other."

> "What kind of refreshments do you prefer? (Select one): Chips, Sandwiches, Pizza, Fruits, Other."

Open-ended questions to allow creativity:

> "What particular activity would you like to see at the event?"

> "Do you have any eating restrictions or preferences?"

3. Select the Survey Tool:

- Use tools like **Google Forms** or **Microsoft Forms** for quick and easy digital collection.
- Offer a **paper-based option** for students without internet access.
- Announce the survey through multiple channels (school website, social media, class announcements).

4. Data Collection:

- Set a deadline to ensure timely responses.
- Encourage participation by highlighting the importance of their input and ensuring secrecy.

5. Data Analysis:

- **Quantitative Analysis:** Use built-in analytics or spreadsheets to calculate the popularity of activities and refreshments.
- **Qualitative Analysis:** Group and categorize open-ended responses to identify repeated themes or unique suggestions.

6. Interpret and Plan:

Based on the data, finalize the event plan:

- Select the most popular activities and refreshments.
- Accommodate dietary restrictions to ensure inclusivity.
- Consider creative ideas from open-ended responses.

7. Feedback:

Share a summary of the planned event with students, showing how their feedback shaped the decisions. This transparency encourages engagement and builds excitement.

8. Future Recommendations:

During the event, gather informal feedback to see if the activities and refreshments meet expectations. Use this information to improve future events. By following these steps, the data collected will ensure the event is tailored to students' preferences, creating a more enjoyable and inclusive experience.

6. Explain the role of tables and charts in data analysis. Provide an example of how you could use a table or chart to present data about students' grades in different subjects.

Ans: **Tables:** Organizing data systematically is very important for clear analysis and interpretation. When data is well-organized, it helps reduce errors. For

example, imagine you have a list of students and their test scores, a disordered list might lead to accidentally recording a score under the wrong student's name. Organizing the data neatly in a table reduces such mistakes. Tables help present data clearly, making it easy to compare and analyse the scores of different students across various subjects.

Charts: Charts are visual representation of data designed to make complex information easier to understand. Charts help identify patterns, trends and outliers in datasets. Common types of charts include:

• Bar charts
• Line charts
• Pie charts

Example: Data about students' grades Using a Table:

A table is useful for organizing and presenting detailed data about students' grades in different subjects.

Name of Student	Physics	Chemistry	Computer	Math	Average	Grade
Ali Jan	75	60	68	70	68.25	C
Muhammad Aayan	78	75	80	82	78.75	B
Atif Mehmood	85	80	92	90	86.75	A
Nazim Ali	90	85	95	85	88.75	A

This table:

•   shows individual performance of the students in each subject.
•   calculates the average grade to give a comprehensive view of performance.

Example: Making Chart:

A **bar chart** provides a visual summary of data.

•   **Y-axis:** Subjects (Physics, Chemistry, Computer, Math).
•   **X-axis:** Grades (0-100).

Each student has a separate bar for each subject, making it easy to compare their performance across subjects. **It helps:**

•   Rapid identification of strong and weak subjects for each student.
•   Shows overall trends in class performance for specific subjects (e.g., high scores in Math but lower scores in Computer).
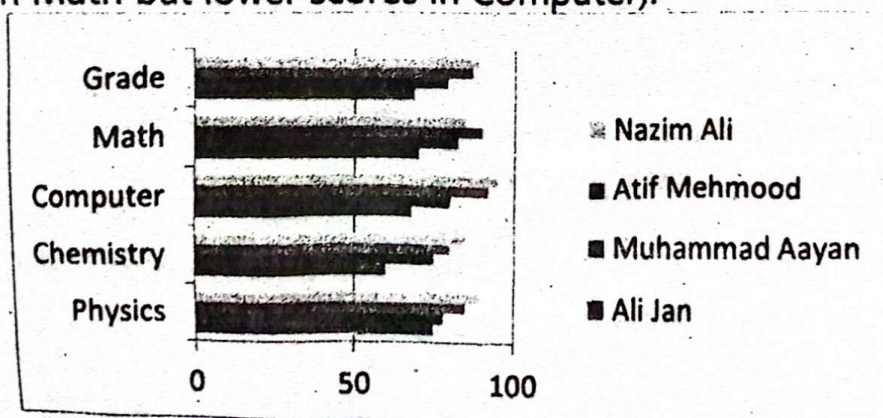
7. Describe a situation where non-numeric data is essential. How would you collect, store, and analyze this type of data? Discuss the tools and techniques you would use to ensure the data is accurately interpreted.

Ans: Non-Numeric is represented by words, labels, or symbols instead of numbers. It describes attributes rather than quantities. Examples include the names of students in a class (e.g., Ali, Badar, Qasim) and the colours of cars in a parking lot (e.g., red, blue, green), where "name" and "colour"are attributes.

Situation Where Non-Numeric Data is Essential:

A school principal wants to understand students' opinions about a proposed change to the school dress code. This requires collecting **qualitative data**, which is non-numeric, such as comments, opinions, and suggestions.

Tools and techniques to ensure the accurate interpretation of data:

1. Data Gathering:

Surveys: Use open-ended questions like, "What do you think about the proposed dress code change?"

Interviews: Conduct small group discussions or one-on-one interviews to gain deeper insights.

Observation: Record non-verbal feedback (e.g., reactions during meetings).

Digital Tools: Use platforms like Google Forms, Microsoft Forms, or Typeform to collect responses.

2. Data Storage:

Save text responses in spreadsheets, databases, or qualitative data analysis software like SPSS etc.

3. Data Extraction:

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

• Identify Relevant Data: Look for specific details about how change of dress code impacts students. Search for keywords like "change dress code" and find sections in articles with important statistics or quotes.

• Copy and Save: Highlight the key information you find. For example, if an article says "70% of students feel stressed because of change of dress code," copy this information and paste it into a document. Keep your data organized by topic or source.

• Use Tools: Some websites have tools to help you gather and organize data. Look for options like "Export" or "Download" to get articles in formats like PDFs.

4. Data Analysis:

• Identify recurring themes or patterns in the responses.

• Remove irrelevant or duplicate responses.

• Clarify ambiguous responses by following up with participants.

- **Word Cloud or Frequency Analysis:** Use tools online tools to visualize commonly mentioned words.
- Classify responses as positive, neutral, or negative to gauge overall sentiment.
- **Use different tools and techniques to ensure accurate Interpretation.**
- Involve multiple analysts to review the data, ensuring objectivity and consistency.

5. Data Integration:

Data integration involves combining information from different sources to get a complete perspective.

- Organize Data: Put all your saved information into one place, such as a document or spreadsheet. For example, create a Google Sheets file titled "Change Dress Code Impact Research" where you list all key statistics and quotes.
- Compare and Merge: Review data from different sources and find common themes. For instance, if many sources say that change dress code increases anxiety among students, note this. Combine similar findings into one summary.
- Create Summaries: Write a brief overview of what you found. Highlight the main points, such as how many students report feeling anxious due to change of Dress Code. Make charts or graphs to show trends, such as bar graph of anxiety levels among students.

6. Data Protection:

Use password-protected files or cloud services like Google Drive to ensure privacy and compliance with data protection laws.

8. Explain the concept of data visualization. How does it help in understanding complex data? Provide examples of different types of visualizations and their applications in real-life scenarios.

Ans: Data Visualization:

Data visualization is the process of turning numbers and information into pictures. These pictures make it easier for us to understand what the data is telling us. When we look at data in the form of charts or graphs, it becomes simpler to see patterns, trends, and relationships.

Understanding Complex Data:

Data visualization plays a crucial role in understanding complex data by:

1. Simplifying complex information: Visualization helps to break down intricate data into easily digestible bits.
2. Revealing patterns and trends: Visualizations can expose underlying patterns, correlations, and anomalies that might be difficult to discern from raw data.
3. Facilitating exploration and discovery: Interactive visualizations enable users to explore data from different angles, fostering a deeper understanding of the data.

Visualizing Different Data Types:

Different types of data can be visualized in different ways. Here are some common types of data and how we can visualize them:

Nominal Data: Nominal data represents categories without any specific order. Bar charts and pie charts are great for showing nominal data.

Ordinal Data: Ordinal data represents categories with a specific order but without a uniform scale. Bar charts and stacked bar charts are effective for visualization for ordinal data.

Discrete Data: Discrete data consists of distinct, separate values. Visualization technique of histograms and dot plots are useful.

Continuous Data: Continuous data can take any value within a range. 193 Visualization techniques of line graphs scatter plot and box plot are commonly used for continuous data.

9. Discuss the importance and benefits of data visualization. Why is it essential for businesses and decision-makers to use charts, graphs, and dashboards?

Ans: Importance and Benefits of Data Visualization:

Data visualization is important because it makes understanding data much easier. Imagine trying to understand a list of hundreds of numbers. It's tough, right? But if those numbers are displayed as a bar chart, it's much easier to see what's going on. Visualizations provide quick insights, allowing you to grasp what the data is saying without having to read through all the details. This is particularly useful for making better and faster decisions. For example, a business can quickly see which products are selling well and which are not. There are several techniques and tools for visualizing data including charts, graphs and dashboards. Some popular tools for visualizing data include Microsoft Excel, Google Sheets and Tableau.

• **Microsoft Excel:** A spreadsheet tool that can create various charts and graphs.

• **Google Sheets:** Similar to Excel, it allows you to create and share visualizations online.

• **Tableau:** A powerful tool specifically designed for creating detailed and interactive visualizations.

• **Microsoft Power BI:** Power BI allows users to create a wide variety of visualizations, including charts, graphs, maps, and more, to represent data in an intuitive and understandable format.

Why Businesses and Decision-Makers Need Data Visualization:

• **Data Overload:** The increasing amount of data available makes it challenging to analyze and make decisions without visualization tools.

• Complexity of Data: Modern data sets often involve complex relationships and patterns, which are difficult to understand without visualization.

• Need for Speed: Businesses require quick insights to stay competitive, and data visualization enables fast and accurate decision-making.

• Communication: Charts and graphs make it easier to present data to a diverse audience, including non-technical stakeholders. They engage viewers with visuals that are more digestible and impactful than raw data or lengthy reports.

• Strategic Decision-Making: Data visualization informs strategic decisions by providing a clear understanding of market trends, customer behavior, and operational performance. Visuals provide clarity, enabling faster and more accurate decisions. It also helps to identify key areas for improvement or growth.

Benefits of Using Charts, Graphs, and Dashboards

1. Charts:

• **Bar Charts:** Compare categories or groups, such as monthly sales by region.

• **Pie Charts:** Show proportions, like market share among competitors.

• **Histograms:** Display frequency distributions, such as customer age ranges.

2. Graphs

• **Line Graphs:** Track changes over time, like revenue trends over quarters.

• **Scatter Plots:** Analyze relationships between variables, such as advertising spend vs. customer acquisition.

3. Dashboards:

• Integrate multiple visualizations into a single interface.

• Provide real-time monitoring of KPIs (e.g., sales, customer satisfaction, and inventory levels).

• Enable interactive exploration, such as filtering data by date or region.

10. Differentiate between nominal, ordinal, discrete, and continuous data. For each type, describe a suitable visualization technique and provide a specific example of how this technique can be used to represent that type of data effectively.

Ans: Visualizing Different Data Types:

Different types of data can be visualized in different ways. Here are some common types of data and how we can visualize them:

Nominal Data:

• Definition: Nominal data represents categories without any specific order.

• Examples: A school wants to analyze the distribution of favorite extracurricular activities among students (e.g., Sports, Music, and Drama).

• **Visualization Technique:** Bar charts and pie charts are great for showing nominal data.

• **Visualization:** Use a bar chart with each category (activity type) on the x-axis and the number of students on the y-axis, or a pie chart showing the percentage of students in each category.

Ordinal Data:

• Definition: Ordinal data represents categories with a specific order but without a uniform scale.

• **Example:** Collect data on students' satisfaction with the school cafeteria (options: Very Unsatisfied, Unsatisfied, Neutral, Satisfied, Very Satisfied).

• **Visualization Technique** Bar charts and stacked bar charts are effective for visualization for ordinal data.

• **Visualization:** Use a bar chart where categories are arranged in order on the x-axis, and the frequency of each response is plotted on the y-axis. A stacked bar chart can show proportions within subgroups (e.g., by grade).

Discrete Data:

Definition: Discrete data consist of distinct, separate values that are countable, often in whole numbers. Discrete data answer questions like "How many?" or "How often?"

• Examples: Track the number of books read by students in a month.

• Visualization technique of histograms and dot plots are useful.

• **Visualization:** A histogram shows data distribution, the count of students (y-axis) for each number of books read (x-axis). A dot plot can also visually represent the frequency of each value.

Continuous Data:

• Definition: Continuous data consists of values that can take any number within a given range, including fractions or decimals..

• **Example:** Monitor students' heights in a PE class.

• Visualization techniques of line graphs scatter plot and box plot are commonly used for continuous data.

• **Visualization:** A histogram can show the distribution of students' heights. Alternatively, a line graph can be used to track the average height changes across grade levels.

Additional Multiple Choice Questions (MCQs)

1. What does data consist of?
a) Only numbers   b) Only words
c) information   d) Raw facts and figures

2. Which type of data is used to describe qualities or characteristics?
a) Quantitative   b) Qualitative   c) Discrete   d) Continuous

3. What is an example of discrete data?
a) Height of students   b) Weight of fruits
c) Number of students   d) Temperature

4. Which characteristic of quantitative data refers to its ability to be expressed in numbers?
a) Measurable
b) Countable
c) Numerical
d) Arithmetical

5. Which characteristic of quantitative data refers to its ability to be used in arithmetic operations?
a) Measurable   b) Countable   c) Numerical   d) Arithmetical

6. What is the characteristic of qualitative data?
a) Numeric   b) Descriptive c) Measurable   d) Countable

7. Which of the following is an example of nominal data?
a) Weigt of fruits b) Height of students c) Gender   d) Temperature

8. Which of the following is an example of continuous data?
a) Number of students in a class   b) Height of students
c) Types of fruits   d) Satisfaction ratings

9. Which operation can be performed on continuous data but not on discrete data?
a) Addition   b) Division   c) Multiplication   d) Mode calculation

10. Which of the following is an example of discrete data?
a) Weight of a fruit   b) Temperature
c) Number of cars   d) Height of a student

11 Which operation can be performed on nominal data?
a) Median   b) Ranking   c) Counting   d) Division

12. What type of data is used to measure the extent or duration of something?
a) Qualitative data   b) Ordinal data
c) Discrete data   d) Continuous data

13. Which of the following is an example of ordinal data?
a) Types of fruits   b) Education level
c) Gender   d) Satisfaction ratings

14. Which of the following is a benefit of organizing data?
a) Increases errors   b) Wastes time
c) Improves clarity   d) Makes data collection harder

15. Which tool is a popular choice for creating surveys?
a) Google Docs   b) Google Forms
c) Microsoft Excel   d) SurveyMonkey

16. Which of the following is a visual representation of data designed to identify patterns and trends?
a) Table   b) Survey   c) Chart   d) Text document

17. What is an example of a method used for collecting data?
a) Organizing data   b) Creating a chart
c) Conducting a survey   d) Analyzing data trends

18. Which of the following is a common tool for creating surveys?
a) Google Forms
b) Microsoft Office Word
c) Excel
d) Notepad

19. Which of the following data types is represented by a pie chart?
a) Nominal data b) Ordinal data c) Quantitative data d) Discrete data

20. Which of the following is an example of continuous data?
a) The number of students
b) The height of students i
c) The types of fruit
d) Satisfaction ratings

21. What operation can be performed on a data table to make it easier to analyse?
a) Grouping and sorting the data
b) Printing the data
c) Writing data manually
d) Storing the data in a physical folder

22. What is the primary purpose of data collection?
a) To make decisions
b) To answer questions
c) improve understanding
d) All of these

23. Which of the following is a method of data collection?
a) Data analysis
b) Data visualization
c) Questionnaires
d) Data storage

24. What is the main characteristic of structured data?
a) It is free-form and unorganised
b) It is organised and formatted
c) It is only used for data analysis
d) It is only used for data visualization

25. A questionnaire used for collecting information through:
a) One-on-one conversations
b) Written questions
c) Observing behaviour
d) Conducting experiments

26. Which of the following is an example of unstructured data?
a) A table of student grades
b) A list of customer names
c) A social media post
d) A sales report in Excel

27. Which method involves watching and noting what happens in a situation?
a) Interview   b) Survey   c) Observation   d) Questionnaire

28. Finding and saving the most relevant information from a larger set of data is called:
What is data extraction?
a) Data extraction
b) Data collection
c) Data integration
d) Data analysis

29. What is an example of structured data?
a) A list of customer reviews
b) A photo album
c) A student record in a spreadsheet   d) A handwritten letter

30. What type of data is harder to organize and requires special tools for processing?
a) Structured data
b) Unstructured data
c) Qualitative data
d) Quantitative data

31. Which of the following is an example of structured data?
a) A video file
b) A blog post
c) A table of product prices
d) An email message

32. Which of the following is an example of online data sources?
a) A personal notebook
b) A printed newspaper
c) Google Scholar
d) A handwritten questionnaire

33. What is the first step in gathering data from online sources?
a) Organizing data
b) Data extraction
c) Data gathering
d) Writing a report

34. Which data storage technology is used for personal budgeting and simple data analysis?
a) Databases b) Spreadsheets   c) Data Warehouses d) NoSQL

35. Which data storage technology is designed for storing and analysing large amounts of data?
a) Databases b) Spreadsheets   c Data Warehouses d) NoSQL

36. Which of the following is an example of a NoSQL database?
a) MySQL   b) MongoDB   c) Microsoft Access d) PostgreSQL

37. What type of data can NoSQL databases handle?
a) Structured data b) Unstructured data c) Both a & B d) Textual data

38. Which of the following is NOT a type of data storage method mentioned in the text?
a) Spreadsheets
b) Data warehouses
c) NoSQL
d) Cloud storage

39. Which database is designed to handle large amounts of data across many servers?
a) MongoDB   b) Cassandra c) Microsoft Excel   d) Google Sheets

40. What is the main purpose of data visualization?
a) To analyse data
b) To collect data
c) To make data easier
d) To make data complicated

41. Which tool is specifically designed for interactive visualizations?
a) Microsoft Excel
b) Google Sheets
c) Tableau
d) Microsoft Power BI

42. Which of the following is NOT a tool for data visualization?
a) Microsoft Excel b) Google Sheets c) Microsoft Word d) Tableau

43. What type of data does a pie chart typically visualize?
a) Nominal data
b) Ordinal data
c) Continuous data
d) Discrete data

44. Which chart is most suitable for showing the distribution of discrete data?
a) Line graph b) Dot plot   c) Scatter plot   d) Bar chart

45. Which of the following is used to visualize continuous data?
a) Pie chart   b) Stacked bar chart c) Line graph   d) Histogram

46. Which type of data has a specific order but no uniform scale?
a) Nominal data b) Ordinal data   c) Discrete data   d) Continuous data

47. Which of the following is a common visualization technique for ordinal data?
a) Pie chart   b) Stacked bar chart c) Scatter plot   d) Box plot

48. What does continuous data represent?
a) Categories with no order   b) Values within a range
c) Separate distinct values   d) Ordered categories

49. What is the purpose of data pre-processing?
a) To analyse data   b) To visualize data
c) To clean and organize data   d) To collect data

50. Which statistical measure is used to describe the central point or typical value in a set of data?
a) Mean   b) Median   c) Mode   d) All of these

51. Which of the following is an example of qualitative data?
a) Test scores   b) Text from interviews
c) Sales numbers   d) Temperatures

52. A value that deviates significantly from other values
a) Outlier   b) Variance   c) Missing value   d) Incorrect value

53. Which measure of central tendency is best when dealing with outliers?
a) Mean   b) Median   c) Mode   d) Range

54. A measure of how much values deviate from the mean is known as:
a) Mean   b) Median   c) Mode   d) Variance

55. Which of the following is a technique for analyzing qualitative data?
a) Statistical analysis   b) Content analysis
c) Regression analysis   d) Correlation analysis

56. What is the first step in data pre-processing?
a) Analyzing data   b) Visualizing data
c) Collecting data   d) Cleaning and organizing data

57. It checks for completeness and accuracy of data:
a) Validation b) Visualization   c) Interpretation   d) Integration

58. Which of the following is an example of a measure of spread?
a) Median   b) Mode   c) Range   d) Mean

59. In which type of data analysis would you count the frequency of specific words in text?
a) Quantitative analysis   b) Qualitative analysis
c) Statistical analysis   d) Predictive analysis

60. What action is typically taken when encountering an error in data during the cleaning process?
a) Ignore the error
b) Highlight the error
c) Correct or delete the error
d) Include the error in analysis

61. What is the primary purpose of cloud storage?
a) To analyse data
b) To visualize data
c) To collect data
d) To store information online

62. Which collaborative tool allows multiple users to edit a document in real-time?
a) Google Drive
b) Google Docs
c) Google Slides
d) Microsoft Excel

63. Which of the following is an example of cloud storage?
a) External hard drive
b) USB flash drive
c) Google Drive
d) CD-ROM

64. What is the main benefit of remote access?
a) Faster internet connection
b) Use a computer from a distant location
c) Increased physical storage
d) Enhanced graphics processing

65. Why are data backups important?
a) To increase file size
b) To allow collaboration
c) To protect against data loss
d) To improve internet speed

66. Which tool is commonly used for collaborative authoring?
a) Photoshop
b) Google Slides
c) VLC Media Player
d) Microsoft Solitaire

67. What is the primary goal of data science?
a) To collect data
b) To analyse data
c) To solve problems using data
d) To visualize data

68. Which of the following is NOT a field combined in data science?
a) Computer science
b) Mathematics and statistics
c) Business knowledge
d) History

69. Extracting insights and solving problems using data is the main purpose of:
a) Data science
b) Data visualization
c) Data integration
d) Data gathering

70. What is the second step in the Data Science Workflow?
a) Data collection
b) Data analysis
c) Data cleaning
d) Data visualization

71. What does data cleaning resemble?
a) Writing a story
b) Drawing a map
c) Cooking a meal
d) Organizing a messy room

72. Which step involves understanding the meaning of analysed data?
a) Data cleaning
b) Data visualization
c) Data interpretation
d) Data collection

73. What is the primary challenge with Big Data?
a) Managing its size
b) Analyzing its complexity
c) Ensuring its security
d) All of these

74. Which data science tool is used for managing and querying databases?
a) Excel   b) Python   c) R   d) SQL

75. Which tool is best for statistical analysis and data visualization?
a) Excel   b) Python   c) R   d) SQL

76. What is an example of Big Data in transportation?
a) Predicting health trends
b) Detecting fraudulent transactions
c) Monitoring traffic to optimize routes
d) Analyzing retail sales patterns

77. Which technique helps visualize relationships between data points?
a) Graph analytics
b) Predictive modelling
c) Data cleaning
d) Data extraction

78. Which tool would you use to create charts from survey data?
a) SQL   b) Java   c) R   d) Python

79. How do banks use Big Data?
a) To improve public transport routes
b) To detect fraudulent activities
c) To analyse sports performance
d) To predict disease outbreaks

Answers:

1	2	3	4	5	6	7	8	9	10	11	12
D	B	C	C	D	B	C	B	B	C	C	D
13	14	15	16	17	18	19	20	21	22	23	24
B	C	B	C	C	A	A	B	A	D	C	B
25	26	27	28	29	30	31	32	33	34	35	36
B	C	C	A	C	B	C	C	C	B	C	B
37	38	39	40	41	42	43	44	45	46	47	48
C	D	B	C	C	C	A	B	B	B	B	B
49	50	51	52	53	54	55	56	57	58	59	60
C	D	B	A	B	D	B	C	A	C	B	C
61	62	63	64	65	66	67	68	69	70	71	72
D	B	C	B	C	B	C	D	B	A	C	C
73	74	75	76	77	78	79					
D	D	C	C	A	D	B					

Topic Wise Additional Short Questions and Answers

9.1. DATA

1. What is data?

Ans: Data consists of raw facts collected about things around us that we can process to generate useful information. It can take many forms, such as numbers, words, measurements, observations, or even images and sounds, and may originate from various sources.

2. What are some examples of data?

Ans: Weather data, sales data, survey responses, website data, and social media data.

3. Why is understanding data important?

Ans: Understanding data is essential in today's world, as it allows us to comprehend situations, make informed decisions, solve problems, and drive innovation

4. How is weather data used?

Ans: : Information on temperature, humidity, and wind speed collected from weather stations.

5. What is the impact of weather data?

Ans: Weather data helps meteorologists predict storms and extreme weather conditions. By analyzing weather data, scientists can issue early warnings for hurricanes, potentially saving lives.

6 What is sales data?

Ans: Sales data includes records of product sales, quantities and prices at a store. For example, selling 50 units of a product at 100 Rupees each in a day.

9.2. Data Types

7. What are the two main types of data?

Ans: Data can be divided into two broad categories: qualitative and quantitative.

8. What is qualitative data?

Ans: Qualitative data refer to categories or labels used to describe the qualities or characteristics of something. This type of data provides insights into opinions, behaviours and experiences through descriptions.

9. Give two examples of non-numeric data.

Ans: Qualitative data is represented by words, labels, or symbols instead of numbers. It describes attributes rather than quantities. Examples include the names of students in a class (e.g., Ali, Badar, Qasim) and the colours of cars in a parking lot (e.g., red, blue, green), where "name" and "colour "are attributes.

10. Give two examples of categorical data.

Ans: Qualitative data can be into categories or classes based on their characteristics. Examples include types of fruit (e.g., apple, banana, orange), job titles (e.g., manager, engineer, accountant), and book genres (e.g., fiction, non-fiction, mystery).

11. What are the two types of qualitative data?

Ans: The two types of qualitative data are Nominal and Ordinal data.

12. What is the main difference between nominal and ordinal data?

Ans: *Nominal data* labels or categorizes items without implying any order. Examples include gender (male, female), types of fruits (apple, banana, orange), and colours (red, blue, green).

Ordinal data represents categories with a meaningful order, though the differences between categories are not uniform. Examples include education levels (high school, bachelor's, master's), and shirt sizes (small, medium, large, extra-large).

13. What is quantitative data?

Ans: Quantitative data consists of numbers used to measure the quantity or amount of something. It is useful for mathematical calculations and statistical analyses.

14. What are key characteristics of quantitative data?

Ans: Key characteristics of quantitative data include being numerical, measurable, countable and arithmetical.

15. What are the two types of quantitative data?

Ans: Discrete data and continuous data.

Discrete data consists of distinct, separate values that are countable, often in whole numbers.

Continuous data consists of values that can take any number within a given range, including fractions or decimals.

16. What are some operations that can be performed on discrete data?

Ans: All operations that can be performed on nominal and ordinal data can also be performed on discrete data. In addition to logical and grouping operations, arithmetic operations (addition, subtraction) and statistical operations (average, range) can be performed on discrete data.

17. What are some operations that can be performed on continuous data?

Ans: All operations that can be performed on discrete data can also be performed on continuous data. In addition to logical, grouping and arithmetic operations division can also be performed on continuous

9.3. Organising and Analysing Data

18. What does organizing data mean? Give an example.

Ans: Organizing data provide clear analysis and interpretation. When data is well-organised, it helps reduce errors. For example, imagine you have a list of

students and their test scores, a messy, list might lead to accidentally recording a score under the wrong student's name. Organizing the data neatly in a table reduces such mistakes.

19. Why is organizing data important?
Ans: • Proper organisation saves time. When data is organised, it's easier to find and analyse.
• Organised data improves clarity. When data is presented clearly, it's easier to understand and interpret.
• By organising data into tables, charts, and graphs, we can quickly grasp what the data is saying, making it simpler to draw conclusions and make decisions.

20. How can you organize data to improve clarity?
Ans: By presenting data in a clear and understandable format, such as tables, charts, and graphs.

21. What are some common ways to organize data?
Ans: Data can be organized into tables, charts, and graphs. These formats help in clearly presenting and analyzing data.

22. What is an example of data organized into a table?
Ans: An example of data organized into a table is student test scores in different subjects, like Math, Science, and English, where each student's score is listed across these subjects.

23. What is the role of charts in organizing data?
Ans: Charts are visual representation of data designed to make complex information easier to understand. Charts help identify patterns, trends and outliers in datasets.

24. What are some types of charts used to organize data?
Ans: Common types of charts include bar charts, line charts, and pie charts.

25. What is the role of graphs in organizing data?
Ans: Graphs visually represent data and show relationships between different data points. Examples include line graphs, bar graphs, scatter plots, and histograms.

26. What are some types of graphs used to organize data?
Ans: Common types of graphs include line graphs, bar graphs, scatter plots, and histograms.

27. What is data collection?
Ans: Data collection is the process of gathering information to answer questions, make decisions, or understand something better.

28. What are some common methods of data collection?
Ans: Questionnaires (Surveys), observations, interviews, experiments and online data sources.

29. What are surveys in data collection?

Ans: Surveys collect information from people by asking them questions. This can be done on paper, over the phone, or online. For example, to find out your classmates' favourite ice cream flavours, you might create a survey with questions like "What is your favourite ice cream flavour?" and give it to your classmates to fill out.

30. What are some popular tools for creating surveys?

Ans: Google Forms, Microsoft Forms, and SurveyMonkey.

31. What are best practices in Survey Design and Administration?

Ans: • Be clear and specific
• Use multiple choice and rating scales
• Test your survey before sending it out
• Keep it short and simple
• Ensure anonymity
• Analyse the results

32. What is the purpose of data visualization?

Ans: To present complex data in a clear and understandable format, making it easier to identify patterns, trends, and outliers.

33. Define Questionnaire. OR What is Questionnaire?

Ans: Questionnaires are similar to surveys but are often written forms that people fill out. They usually
have a set of questions that need to be answered.

34. How interviews are conducted? OR How is an interview different from a survey?

Ans: Talking to individuals one-on-one to gather detailed information. Example: Interviewing a school teacher to understand their experience and challenges. While questionnaires generally involve written forms that collect answers from multiple people.

35. What is meant by observation in data collection?

Ans: Watching and noting what happens in a particular situation. Example: Observing how students behave during a group project to understand how they work in a team.

36. How online data sources help in data collection?

Ans: These include websites, databases, and digital tools where you can find and collect information. For example, if you're researching the most popular pets, you might use the internet to find statistics or articles about pet ownership.

37. What is data gathering?

Ans: Data gathering means collecting information from various places to help with your research or school projects.

38. What is the purpose of data extraction?

Ans: Data extraction is about finding and saving the most relevant information from a larger set of data.

39. What is data integration?

Ans: Data integration involves combining information from different sources to get a complete perspective

9.4. Data Types:

40. What are the two types of data with respect to storage and processing?

Ans: Structured and unstructured data.

41. What is structured data?

Ans: Structured data is organised and formatted to be easily searchable and analysable. Examples
include data in spread sheets and traditional databases.

42. What is unstructured data?

Ans: Structured data is organised and formatted to be easily searchable and analysable. Examples include data in spread sheets and traditional databases.

43. What is the difference between structured and unstructured data?

Ans: Structured data, like spreadsheets, is neatly organised into rows and columns, making it easy to search and analyse. Unstructured data, like emails or social media posts, is more chaotic and requires special tools to process.

9.5 Data Storage Techniques:

44. What are the four important data storage technologies?

Ans: Spreadsheets, Databases, Data Warehouses, and NoSQL.

45. What are spreadsheets used for?

Ans: Spreadsheets are tools that help us organise data in rows and columns. We can perform calculations, create charts, and sort data easily. These are commonly used for personal budgeting, simple data analysis and small business inventory management.

46. How do databases store data?

Ans: Databases store data in tables, where each row represents a record and each column represents the attributes of that record. They are designed to handle large amounts of structured data efficiently.

47. What is the purpose of a data warehouse?

Ans: A data warehouse is designed for storing and analyzing large amounts of data collected from various sources. It helps organizations to make decisions and understand business performance. Prominent examples of data warehouse include Amazon Redshift, a data warehouse service offered by Amazon Web Services, and Google BigQuery, a fast and scalable data warehouse offered by Google Cloud.

48. What does NoSQL stand for?

Ans: NoSQL stands for "Not Only SQL" and refers to database technologies designed to handle unstructured data, offering more flexibility than traditional SQL databases.

49. What is an example of a NoSQL database?

Ans: MongoDB is an example of a NoSQL database that stores data in JSON-like documents. Another example is Cassandra, which handles large amounts of data across many servers.

9.6 Data Visulization:

50. What is data visualization?

Ans: Data visualization is the process of turning numbers and information into pictures. These pictures make it easier for us to understand what the data is telling us. When we look at data in the form of charts or graphs, it becomes simpler to see patterns, trends, and relationships.

51. Why is data visualization important?

Ans: It makes understanding data easier and provides quick insights. This is particularly useful for making better and faster decisions. For example, a business can quickly see which products are selling well and which are not.

52. What are some popular tools for visualizing data?

Ans: Some popular tools for visualizing data are:
• Microsoft Excel: A spreadsheet tool that can create various charts and graphs.
• Google Sheets: Similar to Excel, it allows you to create and share visualizations online.
• Tableau: A powerful tool specifically designed for creating detailed and interactive visualizations.
• Microsoft Power BI: Power BI allows users to create a wide variety of visualizations, including charts, graphs, maps, and more, to represent data in an intuitive and understandable format.

53. How can nominal data be visualized?

Ans: Using bar charts and pie charts.

54. What type of visualization is suitable for continuous data?

Ans: Line graphs, scatter plots, and box plots.

55. What type of data is best visualized using a bar chart or pie chart?

Ans: Nominal data, which represents categories without a specific order.

56. What type of data is represented by a line graph?

Ans: Continuous data, which can take any value within a range.

57. What is the main difference between ordinal and nominal data?

Ans: Ordinal data has a specific order, while nominal data consists of categories without any particular order.

58. Which visualization technique is commonly used for discrete data?

Ans: Histograms and dot plots.

59. What tool is specifically designed for creating detailed and interactive visualizations?

Ans: Tableau.

9.7 Data Pre-Processing and Analysis:

60. What is data pre-processing?

Ans: Data pre-processing is the first and most important step in working with data. It involves getting the data ready for analysis by cleaning and organizing it.

61. Why is evaluating data quality important in data pre-processing?

Ans: Before using data, we need to check its quality, ensuring it is accurate, complete, and reliable. We ask questions like: Is any data missing? Are there errors? Is the data consistent and up-to-date?

62. What are outliers in data?

Ans: Outliers are unusual or extreme values that don't fit the pattern of the rest of the data.

Example: In a list of test scores, if most students scored between 50 and 80, but one student scored 5, the score of 5 is an outlier.

63. What are biases? How they affect the data?

Ans: Biases are distortions that affect the accuracy of the data.

Example If a survey only includes answers from students in one school, it may not represent the opinions of all students in the city. Hence if you want to analyse opinion of entire city students then this data would be biased due to the limited sample.

64. What is the difference between data validation and data cleaning?

Ans: Data Validation involves checking data completeness and accuracy., while data cleaning involves error removing, handling missing data and dealing with outliers. Errors are either corrected or the incorrect data is deleted.

65. How can missing data be handled?

Ans: Missing data can be handled by filling the gaps with estimated values, such as using the average value of a dataset, or by deleting the incomplete record.

66. What is quantitative analysis?

Ans: Quantitative analysis deals with numbers and measurable data. It helps us understand patterns, relationships, and trends in numeric data.

67. What is the purpose of statistical analysis?

Ans: Statistical analysis uses mathematical techniques to analyse data, helping to summarize, describe, and infer patterns and trends.

68. What is the mean of a dataset?

Ans: The sum of all values divided by the number of values, representing the average.

69. What is the mode of a dataset? OR What is the mode in statistical analysis?

Ans: The value that appears most frequently. There can be more than one mode if multiple values appear with the same frequency.

70. What is the median of a dataset?

Ans: The middle value when all values are arranged in order. If there is an even number of values, the median is the average of the two middle values.

71. What does standard deviation measure?

Ans: Standard deviation measures the amount of variation or spread in a dataset, indicating how much individual data points deviate from the mean.

72. What is measure of spread? OR What is measure of dispersion?

Ans: Measures of spread (also known as measures of dispersion) are statistical tools used to describe the amount of variation or diversity in a dataset. They provide degree to which data points differ from the average value (mean) or median.

73. What is range of data? How it is found?

Ans: The range is the simplest measure of spread. It is the difference between the highest and lowest values. For example, for the test scores 60, 75, 83, 91, and 95, the range is 95 - 60 = 35.

74. What is variance? Write its formula.

Ans: Variance measures how spread out the values are from the mean. It gives a sense of how much the values in a dataset vary from the mean. Formula for the variance is:

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

75. What is standard deviation? Write its formula.

Ans: Standard Deviation: This measures the amount of variation or spread in a set of data. A low standard deviation means that the data points are close to the mean, while a high standard deviation means that the data points are spread out over a wide range. Formula for standard deviation is:

$$S^2 = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

76. What is Qualitative analysis?

Ans: Qualitative analysis deals with non-numeric data such as text, images, and sounds. It helps us understand meanings, concepts, and experiences.

77. What is content analysis?

Ans: One common method is Content Analysis, which involves counting the number of times specific words or themes appear in a text. For example, researchers might analyse a set of interviews to count how many times students mention "homework stress." This method helps to quantify the presence of specific themes or ideas within qualitative data.

78. What is thematic analysis?

Ans: Thematic Analysis, which involves identifying and interpreting themes or patterns within qualitative data. For instance, by reading through student

essays, researchers can find common themes like "difficulty with math" or "enjoyment of science." This method provides a deeper understanding of the underlying themes and patterns in the data.

9.8 Collaborative Tools and Cloud Storage

79. What is use of collaborative tools?

Ans: In today's world, working together and sharing information online is essential. Collaborative tools make this possible, allowing people to work together, share files, and access their work from anywhere.

80. What is cloud storage used for in data management?

Ans: It allows us to save files on the internet for access from any device, make backups to protect our data, and work on documents with others in real-time.

81. What is remote access?

Answer: Remote access refers to the ability to connect to and use a computer or network from a distant location. This means you can work on files, use software, or access resources on a computer or network that is not physically near you. For example, Google Drive.

82. What is the activity flow for remote access?

Ans: Following is the activity flow for remote access.

• **Save Your Work:** Open Google Drive and upload your project file.

• **Access from Anywhere:** Log into Google Drive from any device (computer, tablet, phone) and open your project.

• **Continue Working:** Make changes, add new information, or review your work from wherever you are.

83. Why are data backups important?

Ans: Data Backups are copies of important data or files stored separately from the original to protect against data loss. Backups are essential to ensure that you can recover your data if something goes wrong, such as accidental deletion, hardware failure, or a computer virus.

84. What is the activity flow for data backup?

Ans: Following is the activity flow for data backup on a cloud service.

• **Automatic Backups:** Set up your device to automatically back up important files to a cloud service like OneDrive.

• **Restore Files:** If you lose a file, log into your cloud account and download the backup version.

• **Resume Work:** Continue working on your restored file without worrying about losing data.

85. What is collaborative authoring?

Ans: The process of multiple people working together to create, edit, and improve a document or project in real-time. People use online tools that allow them to contribute and see each other's changes.

86. What is the activity flow for collaborative authoring?

Ans: Following is the activity flow for collaborative authoring.

• *Create a Shared Presentation:* In Google Slides, create a presentation called "Climate Change Presentation" and share it with your friends.

• *Upload Files:* Each group member work on their slides and notes.

• *Edit Together:* Open the files and make edits together, ensuring everyone's input is included.

87. What are the benefits of collaborative tools?

Ans: Enhanced productivity, version control, and real-time collaboration.

• *Enhanced Productivity:* When multiple students work on different sections of a project simultaneously the project gets done faster. This saves time and allows everyone to contribute their best work.

• *Version Control:* Google Docs automatically saves every change, so you can go back to previous versions if needed. You don't lose work, and you can see who made which changes in the document.

88: Give an example of a collaborative tool for creating presentations.

Ans: Google Slides.

89: How do collaborative tools help with version control?

Ans: Collaborative tools like Google Docs save every change and allow users to revert to previous versions and see who made changes.

90. What is global collaboration?

Ans: Global Collaboration: Cloud-based tools have enabled people to work together on projects from different parts of the world. For instance, a student in Pakistan can collaborate on a science project with peers in the USA and Australia simultaneously.

9.9 Introduction to Data Science

91. What is data science?

Ans: Data science is the process of solving problems using data by analyzing patterns and extracting insights.

92. Why is data science important?

Ans: It helps us make better decisions and improve outcomes in various fields such as education, business, sports, and healthcare. By learning data science, we can solve problems more effectively and make smarter choices in everyday life.

93. Name three fields combined in data science. OR What fields does data science combine?

Ans: Computer science, mathematics and statistics, and business knowledge.

94. What is the data science workflow?

Ans: A systematic process involving problem identification, data collection, data cleaning, data analysis, data interpretation, and data visualization.

95. What is the first step in the Data Science Workflow?

Ans: Problem identification; Understanding and clearly defining the problem you are trying to solve.

96. What does data cleaning involve?

Ans: After collecting data, it's important to clean it by removing errors and organizing it. Data cleaning involves fixing mistakes and making the data useful.

97. How can data visualization help in data science?

Ans: It makes data easier to understand by using charts and graphs to represent findings clearly.

98. Provide an example of data science in sports.

Ans: Sports teams use data to create better strategies and enhance performance.

99. What is the role of business knowledge in data science?

Ans: It applies insights from data to solve real-life problems and make informed decisions.

100. What is difference in data analysis and data interpretation in data science?

Ans: • **_Data Analysis:_** This step involves looking closely at the data to find patterns or answers. For example, if you have weather data from the past month, you can analyse it to predict if it will rain tomorrow. It's like solving a puzzle by examining the pieces.

• **_Data Interpretation:_** Once you have analysed the data, you need to understand it and draw conclusions. This is like reading a story and understanding its meaning. For example, after analysing survey results, you might conclude that most students prefer outdoor sports over indoor ones.

9.10 Big Data and its Applications:

101. What is Big Data?

Ans: Big Data refers to extremely large and complex sets of data that are difficult to process using traditional methods.

102. What are the three main characteristics of Big Data? Or What are the three Vs of Big Data?

Ans: Volume (amount of data), Velocity (speed of data generation and processing), and Variety (different forms of data).

103. Provide an example of Big Data's "Volume" characteristic.

Ans: Social media platforms generate a massive number of posts, likes, and comments daily, demonstrating the "Volume" aspect.

104. How is Big Data used in the retail industry?

Ans: Retailers use Big Data to analyse customer behaviour, improve product offerings, and enhance the shopping experience.

105. What is the use of Python in data science?

Ans: Python is a popular programming language used in data science for analysing data and building data models. It comes with powerful libraries like Pandas, for data manipulation and Matplotlib for creating graphs.

106. What is the role of R language in data science?

Ans: R is another programming language designed for statistical analysis and data visualization. It is particularly good at handling complex data and presenting it in a clear way.

107. Which programming language is widely used for data science, with libraries like Pandas and Matplotlib?

Ans: Python.

108. Name one tool used for querying databases in data science.

Ans: SQL (Structured Query Language).

109. What is the use of SQL in data science?

Ans: SQL (Structured Query Language) is used to manage and query databases. SQL allows to extract specific information quickly. These tools are important for data scientists as they help in make sense of large amounts of data, provide valuable insights, and support informed decisions.

110. What is predictive modelling?

Ans: Predictive modelling uses historical data to forecast future events, such as identifying students who may need extra academic support.

111. What are graph analytics in data science?

Ans: Graph analytics is a method for analysing relationships between different data points. It helps to visualize and understand connections and interactions.

112. How does data science impact various industries?

Ans: Data science is used in retail, healthcare, finance, sports, and other industries to solve problems, make better decisions, and improve outcomes.

113. What role does AI play in the future of data management?

Ans: AI will automate tasks, enhance analysis speed, improve visualization, and enable more effective decision-making. AI helps everyday gadgets like virtual assistants (Siri and Alexa) understand our questions and perform tasks, making technology feel more intuitive.

114. What is the role of data science in healthcare?

Ans: Hospitals use data science techniques to monitor patient health and predict trends. For example, doctors use predictive modelling to analyse patients' medical histories and identify their diseases, such as diabetes or heart disease. This helps doctors take preventative measures and offer early treatment to save lives.

115. What is the role of data science in sports?

Ans: In sports, data science techniques are used to enhance player performance and team strategies. For example, coaches analyse player statistics and game data to identify strengths and weaknesses. They can develop strategies to improve team performance and make decisions during games.