

Unit no 09: Data Science and Data Gathering

Long Question Answers:

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

Data can be divided into two broad categories, namely qualitative and quantitative.

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, behaviors, 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 colors 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' favorite extracurricular activities.

Surveys: 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'

COMPUTER SCIENCE

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.

Best Practices in Survey Design and Administration

- Be clear and specific
- Keep it short and simple
- Use multiple choice and rating scales
- Ensure anonymity
- Test your survey before sending it out

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 in measuring student performance or tracking attendance.

Continuous Data: Continuous data consists of values that can take any number within a given range, including fractions or decimals. Examples include student heights (150.5 cm, 160.2 cm), fruit weights (1.5 kg, 2.2 kg), and room temperatures (22.5°C, 23.7°C).

For example

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.

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

For example, counting oranges in a basket might show there are 16, answering "How many?" similarly, tracking the number of oranges falling from a tree each week, with counts like 5 on Monday, 7 on Wednesday, and 4 on Friday, helps answer "How often?"

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

Google forms are often used for creating surveys since they allow you to arrange events ask questions of your staff or clients and collect a variety of information in an easy effective manner. Google forms support a variety of questions format including short responses, paragraph, multiple selection, verification boxes, pull down, linear scale and grid with numerous possibilities among others.

Benefits:

1. Advanced users can utilize regular expressions to customize the types of data that can

be entered into a field this further customize the form.

2. Google forms allow us to preview how the survey will appear before sending it to the recipients.
3. We can send the form via email integrated into our website or share the link via social media or other channels.

Challenges:

1. Google forms has basic customization possibilities but it may not have advanced tools for more complex survey designs user seeking highly customizable layouts or intricate question branching may find the platform too restricted.
2. While Google forms offer a variety of question kinds such as multiple choice short answer and drop down it may not cover the entire range of question types required for specific research or survey needs user seeking sophisticated question kinds or experimental designs may need to look into alternate survey platform
5. **Imagine you are tasked with organizing a school event and need to collect data on students' preferences for activities and refreshments.**
Collecting information about student's preferences for activities and refreshments is critical when organizing a school event. Input can be gathered using a digital tool such as Google forms or printed service. The surveys should include areas for activities (e.g., sports, arts, games) and refreshments (check boxes for popular items such as pizza and beverages). Once the data has been spread and collected evaluate it to detect trends. For example if the majority of students prefer outside games, emphasize them, and base the menu on popular beverages sharing the final plan which students demonstrates that their input was respected resulting in a tailored and interesting event.
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.**
Tables and charts play a vital role in data analysis by helping to organize and present information in a clear and understandable way.
Tables: are useful for displaying raw data in a structured format, allowing for easy comparison of exact values across different categories. They are especially helpful when accuracy and detailed analysis are required.
Charts: on the other hand, provide a visual representation of data, making it easier to identify patterns, trends, and outliers at a glance. They are particularly effective when presenting data to a broader audience, as visuals are often more engaging and easier to interpret than numbers alone.
For example, to analyze students' grades in different subjects, a table can be used to list each student's scores in Math, Science, English, and History. This makes it easy to

compare individual performance across subjects. Alternatively, a bar chart can visually represent the same data by showing grouped bars for each student, with each bar representing a subject. This helps quickly identify which subjects students are performing well in and where they may need improvement. Using such tools in data analysis enhances both understanding and communication of the data's story.

Student Name	Math	Science	English	History
Ayesha	85	78	90	88
Bilal	92	84	75	80
Zara	76	88	82	91

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.**

Non-numeric data, also known as qualitative data, is essential in situations where understanding **opinions, feelings, behaviors, or descriptions** is more important than numbers. One common example is **analyzing customer feedback or student course evaluations**, where the goal is to understand people's experiences, satisfaction levels, or suggestions for improvement.

To collect such data, you might use methods like **interviews, open-ended survey questions, focus groups, or observation notes**. For instance, in a school setting, if you want to know how students feel about online learning, you could distribute a survey with open-ended questions asking about their experiences, challenges, and suggestions. Once collected, non-numeric data is typically stored in **documents, spreadsheets, or qualitative data analysis software** like **NVivo, Atlas.ti, or even Excel** for simpler cases. To analyze this data, techniques such as **thematic analysis or coding** are used. This involves reading through the responses, identifying common themes or patterns, and assigning codes or labels to different segments of the text.

To ensure accurate interpretation, it is important to use techniques like **triangulation** (comparing multiple data sources), **peer debriefing** (discussing interpretations with others), and **member checking** (verifying findings with participants). These strategies help reduce bias and improve the reliability of the analysis. Additionally, using word cloud generators or text mining tools can provide visual summaries of frequently mentioned words or ideas, making the insights easier to communicate.

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.**

COMPUTER SCIENCE

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.

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.

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

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.

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.

Nominal, ordinal, discrete, and continuous data are different types of data used in statistics and data analysis, each with unique characteristics and suitable visualization techniques. Here's a breakdown:

- ◆ 1. Nominal Data
 - Definition: Categorical data with no specific order.
 - Example: Types of fruits (apple, banana, orange).
 - Visualization Technique: Bar chart or Pie chart.
 - Use Case: A bar chart showing the number of students who prefer different fruits in a class.
- ◆ 2. Ordinal Data
 - Definition: Categorical data with a meaningful order, but not measurable differences between values.
 - Example: Customer satisfaction levels (Very Unsatisfied, Unsatisfied, Neutral, Satisfied, Very Satisfied).
 - Visualization Technique: Ordered bar chart or Stacked bar chart.
 - Use Case: An ordered bar chart displaying survey results on customer satisfaction, where responses are ranked from least to most satisfied.
- ◆ 3. Discrete Data
 - Definition: Numeric data that takes only specific whole values (countable).
 - Example: Number of students in a class.
 - Visualization Technique: Bar chart or Dot plot.
 - Use Case: A bar chart showing how many students are enrolled in each course (e.g., Math: 30, English: 25).
- ◆ 4. Continuous Data
 - Definition: Numeric data that can take any value within a range (measurable).
 - Example: Height of students in centimeters.
 - Visualization Technique: Histogram or Line graph.
 - Use Case: A histogram showing the distribution of students' heights in a school to analyze which height range is most common.