1.	What is data literacy?
Ans:	Data literacy is the ability to read, interpret, create, and communicate data. It involves understanding data in various forms (e.g., charts, graphs, tables) and being able to use data for decision-making, problem-solving, and analyzing trends.
2.	What is the difference between qualitative and quantitative data?
Ans:	▰ **Qualitative data** is non-numeric and describes qualities or characteristics (e.g., colors, names, or categories). ▰ **Quantitative data** is numeric and can be measured or counted (e.g., age, height, or number of students).
3.	Why is data visualization important?
Ans:	Data visualization helps to present data in a graphical format such as charts, graphs, and maps. It allows us to understand complex data easily, identify trends, make decisions, and communicate information effectively to others.
4.	What are the basic types of data?
Ans:	The basic types of data are: ● **Qualitative data** (categorical data): Data that describes characteristics. ● **Quantitative data** (numerical data): Data that can be measured or counted.
5.	What is the role of data in decision-making?
Ans:	Data plays a critical role in decision-making by providing factual information and evidence to support or refute ideas. It helps to make informed choices based on trends, patterns, and insights derived from the data.
6.	How is data collection useful in the AI project cycle?
Ans:	Data collection is crucial in the AI project cycle as it provides the raw material needed to train AI models. The quality and quantity of the data collected directly impact the performance of the AI system.

7.	What is the importance of data privacy?
Ans:	Data privacy ensures that personal and sensitive information is protected from unauthorized access, use, or disclosure. It is important to maintain trust, comply with laws, and protect individuals' rights to privacy.
8.	What is the significance of cleaning data?
Ans:	Data cleaning is important because it helps remove errors, inconsistencies, and irrelevant information from the dataset. Clean data ensures accurate analysis, better predictions, and more reliable results from AI models.
	4 Marks
1.	Explain the different types of data.

Ans:	Data can be classified into two main types:
	
	● **Qualitative Data**: Also known as categorical data, this type of data describes characteristics or qualities. It is non-numeric and is used to categorize or label attributes. Examples include the color of a car, the type of music, or gender.
	
	● **Quantitative Data**: This type of data is numeric and can be measured or counted. It represents amounts or quantities and can be subjected to mathematical operations. Quantitative data is further divided into:
	
	   o **Discrete Data**: This refers to countable data, which can only take certain values (e.g., number of students in a class).
	
	   o **Continuous Data**: This refers to measurable data that can take any value within a range (e.g., height, weight, or temperature).

2.	What are the steps involved in data analysis?
Ans:	The process of data analysis involves several steps: 1. **Data Collection**: Gathering raw data from various sources. 2. **Data Cleaning**: Removing errors, duplicates, or irrelevant data to ensure the quality of the dataset. 3. **Data Exploration**: Analyzing the data to find patterns, trends, or insights.

	4. **Data Visualization**: Representing data visually through graphs, charts, and maps to make it easier to understand. 5. **Data Interpretation**: Drawing conclusions from the visualized data to inform decisions. 6. **Reporting**: Presenting the findings to stakeholders in an understandable and actionable format.
3.	What is the importance of data visualization in understanding data?
Ans:	Data visualization plays a crucial role in making complex data easier to understand and interpret. It involves the graphical representation of data through charts, graphs, and maps. Key benefits of data visualization include: ● **Simplifies complex information**: Visuals make it easier to understand large volumes of data at a glance. ● **Identifies trends and patterns**: Visualization helps in recognizing patterns, correlations, and trends that may not be evident in raw data. ● **Improves decision-making**: By presenting data clearly, it supports informed decision-making. ● **Engages the audience**: Graphs and visuals are more engaging than textual data, making it easier to communicate findings effectively.
4.	What are the different data privacy practices that should be followed?
Ans:	Data privacy involves protecting personal and sensitive information from unauthorized access or use. Some key practices for ensuring data privacy include: ● **Data Encryption**: Encrypting data ensures that it is unreadable to unauthorized users, even if it is intercepted. ● **Access Control**: Limiting access to data based on roles and responsibilities to ensure only authorized personnel can view or use sensitive information. ● **Anonymization**: Removing personally identifiable information (PII) from datasets to protect individuals' privacy.

	● **Regular Audits**: Performing regular security audits to identify and address vulnerabilities in data handling systems. ● **Compliance with Legal Regulations**: Adhering to data privacy laws such as GDPR or CCPA to ensure legal protection of personal data.
5.	Describe the concept of data cleaning and its importance.
Ans:	Data cleaning is the process of identifying and correcting errors or inconsistencies in a dataset to improve its quality. This step is crucial in ensuring that the data used for analysis or training AI models is accurate and reliable. The importance of data cleaning includes: ● **Improved Accuracy**: Clean data leads to more accurate analysis and insights. ● **Better Decision Making**: Reliable data supports better business or research decisions. ● **Efficiency**: Clean data reduces the time and effort spent analyzing and interpreting erroneous or incomplete datasets. ● **Enhanced Model Performance**: In AI projects, data cleaning ensures that the model is trained on high-quality data, which improves its performance and predictions.