Data-related Termilogies:-

1. Data Lake: A centralized repository that allows you to store all your structured and unstructured data at any scale. It provides a flexible and cost-effective solution for big data processing and analytics.

2. Data Warehouse: A large, centralized repository of data that is used to support business intelligence and reporting activities. Data warehouses are typically used to consolidate and analyzedata from different sources.

3. Data Wrangling: The process of cleaning, structuring, and enriching raw data into a desired format for better decision-making in less time. It often involves tasks like cleaning, transforming, and aggregating data.

4. ETL (Extract, Transform, Load): A process of extracting data from source systems, transforming it into a more usable format, and loading it into a target system, such as a data warehouse or database.

5. Data Mining: The process of discovering patterns and information from large datasets. It involves techniques from statistics, machine learning, and database systems.

6. Data Governance: The management of data availability, usability, integrity, and security in an organization. It involves defining processes, roles, policies, and standards for handling data.

7. Data Quality: The accuracy, completeness, and reliability of data. Data quality management involves processes and technologies to ensure that data meets the required standards.

8. Master Data Management (MDM): The process of creating and managing a single, consistent, accurate, and complete version of master data across an organization. Master data includes entitiessuch as customers, products, and employees.

9. Metadata: Information that describes the characteristics of data, such as its structure, format, source, and usage. Metadata is crucial for data management and understanding the context of thedata.

10. Big Data: A term that refers to extremely large and complex datasets that traditional data processing applications are inadequate to deal with. It involves processing and analyzing data withinnovative technologies to extract valuable insights.

11. Data Mining: The process of discovering patterns and knowledge from large datasets using various techniques, including machine learning, statistical analysis, and artificial intelligence.

12. Data Integration: The process of combining data from different sources to provide a unified view. It includes tasks like data cleaning, transformation, and loading (ETL).

13. Data Modeling: The process of creating a visual representation of data structures and relationships, often using diagrams, to design databases and systems.

14. Data Mart: A smaller, focused subset of a data warehouse that is designed for the needs of a specific business unit or team.

15. Data Scientist: A professional who uses scientific methods, processes, algorithms, and systems to extract insights and knowledge from structured and unstructured data.

16. Data Security: The practice of protecting data from unauthorized access, disclosure, alteration, and destruction. It involves implementing measures such as encryption, access controls, and monitoring.

17. Data Privacy: The protection of personal and sensitive information, ensuring that data is handled in compliance with privacy regulations and ethical standards.

18. Data Catalog: A centralized repository that stores metadata and information about the available data assets within an organization, making it easier to discover and manage data.

19. Data Governance: The overall management of the availability, usability, integrity, and security of the data used in an enterprise.

20. Data Pipelines: A set of processes and tools for ingesting, processing, and moving data from one place to another, often used in data integration and ETL workflows.