***Task -1 2023/03/14***

1. ***Big Data***: Big Data refers to large and complex datasets that cannot be effectively managed, processed or analyzed by traditional data processing applications. It typically involves data from various sources, including social media, internet searches, and sensor data. The three key characteristics of Big Data are volume, velocity, and variety. Big Data technologies such as Hadoop, Spark, and NoSQL databases have emerged to help process and analyze this data. Some key points to remember about Big Data include:

* Big Data is defined by its volume, velocity, and variety.
* Big Data requires specialized tools and technologies to manage, process and analyze.
* Hadoop, Spark, and NoSQL databases are commonly used Big Data technologies.
* Big Data is used in various industries including healthcare, finance, and e-commerce.

1. ***Data Lake:*** A Data Lake is a centralized repository that allows you to store all your structured and unstructured data in its native format. Unlike traditional data warehouses, a data lake doesn't require you to pre-define the schema of your data. Instead, you can store your data as-is and then apply the necessary transformations when you need to use it. Some key points to remember about Data Lake include:

* Data Lakes store all data types, including structured, semi-structured, and unstructured data.
* Data Lakes are designed for storing and processing large volumes of data in their raw form.
* Data Lakes support a variety of data processing and analytics technologies such as Hadoop and Spark.
* Data Lakes allow for greater data agility and flexibility than traditional data warehouses.

1. ***Database:*** A Database is a structured collection of data that is stored and organized in a specific way so that it can be easily accessed, managed, and updated. Databases are commonly used in businesses and organizations to store customer information, sales data, and other types of information. Some key points to remember about databases include:

* Databases come in various types such as Relational, NoSQL, and Graph databases.
* Databases support data management and transaction processing.
* Databases typically require you to pre-define the schema before storing data.
* Databases are widely used across many industries and applications.

1. ***Data Warehouse:*** A Data Warehouse is a centralized repository that allows you to store and manage large amounts of data from multiple sources. Unlike a database, a data warehouse is designed for analytical processing and decision-making. It typically uses a schema-on-write approach, which means that the data is pre-structured before it's loaded into the warehouse. Some key points to remember about Data Warehouse include:

* Data Warehouses are optimized for analytical processing, reporting, and decision-making.
* Data Warehouses use ETL (Extract, Transform, Load) processes to move data from source systems into the warehouse.
* Data Warehouses require you to pre-define the schema before storing data.
* Data Warehouses are used to support business intelligence and analytics applications.