In the first week of Fellowship, we learned about new concept of data engineering. I will discuss it one by one.

Big Data: -

Big data refers to the large volume of structured and unstructured data that is generated from various sources such as social media, sensors, and other digital devices.

Data lake: -

A data lake is a centralized repository that allows organizations to store, manage, and analyze vast amounts of structured and unstructured data.

Data Warehouse: -

A data warehouse is a centralized repository of integrated data from one or more sources, designed to support business intelligence, analytics, and reporting activities. Unlike a data lake, a data warehouse is designed to store structured data, such as data from transactional systems or enterprise applications, in a predefined schema or structure.

Database: -

A database is an organized collection of data that is stored and managed in a way that allows for efficient retrieval, manipulation, and updating of the data. Databases are used in a wide range of applications, from small personal applications to large enterprise systems.

Data Mart: -

A data mart is a subset of a data warehouse that is designed to serve a particular business unit or department within an organization. Unlike a data warehouse, which stores data from across an organization, a data mart typically focuses on a specific area of business, such as sales, marketing, or finance.

Data lake house: -

A data lake house is a new data architecture that combines the advantages of both data lakes and data warehouses. It's a hybrid approach that allows organizations to store and analyze both structured and unstructured data in a single platform.

Data mesh: -

Data Mesh is a relatively new data architecture approach that aims to address some of the challenges associated with traditional centralized data architectures. It proposes a more decentralized and self-organizing approach to managing data in complex organizations.

ETL: -

ETL stands for Extract, Transform, and Load. It is a process that involves collecting data from various sources, transforming it into a format suitable for analysis or storage, and loading it into

a data warehouse or other storage system. ETL is a critical process in data integration and is used by organizations to centralize data from different sources for analysis and business intelligence purposes.

ELT: -

"Extract, Load, Transform." It's a data integration process where data is first extracted from various sources, then loaded into a target data store or warehouse, and finally transformed into a format suitable for analysis and reporting.

These are main topic cover by fellowship in this week.