**Bytewise fellowship: Data Engineering**

**Task 2:**

**Data marts:**

A data mart is a subset of a data warehouse focused on a particular line of business, department, or subject area. Data marts make specific data available to a defined group of users, which allows those users to quickly access critical insights without wasting time searching through an entire data warehouse. For example, many companies may have a data mart that aligns with a specific department in the business, such as finance, sales, or marketing.

**Data Lakehouse:**

A data Lakehouse is a new, open data management architecture that combines the flexibility, cost-efficiency, and scale of data lakes with the data management and ACID transactions of data warehouses, enabling business intelligence (BI) and machine learning (ML) on all data.

**Data Mesh:**

Data mesh is a sociotechnical approach to building a decentralized data architecture by leveraging a domain-oriented, self-serve design, and borrows Eric Evans’ theory of domain-driven design and Manuel Pais’ and Matthew Skelton’s theory of team topologies.

data mesh is a new approach to thinking about data based on a distributed architecture for data management. The idea is to make data more accessible and available to business users by directly connecting data owners, data producers, and data consumers. Data mesh aims to improve business outcomes of data-centric solutions as well as drive adoption of modern data architectures.

**Data Warehouse vs Data Lake:**

A data lake contains all an organization's data in a raw, unstructured form, and can store the data indefinitely — for immediate or future use. while A data warehouse contains structured data that has been cleaned and processed, ready for strategic analysis based on predefined business needs.

Data from a data lake — with its large volume of unstructured data — is typically used by data scientists and engineers who prefer to study data in its raw form to gain new, unique business insights. while Data from a data warehouse is typically accessed by managers and business-end users looking to gain insights from business KPIs, as the data has already been structured to provide answers to pre-determined questions for analysis.

**OLTP Vs OLAP:**

* Online Analytical Processing (OLAP): Online Analytical Processing consists of a type of software tools that are used for data analysis for business decisions. OLAP provides an environment to get insights from the database retrieved from multiple database systems at one time. Examples – Any type of Data warehouse system is an OLAP system.
* Online transaction processing (OLTP): Online transaction processing provides transaction-oriented applications in a 3-tier architecture. OLTP administers the day-to-day transactions of an organization.