**Task 02**

**Abdul Razaque**

* **Data Mart:**

It is a data store that is designed for a particular department of an organization, or Data Mart is a subset of a Data Warehouse that is usually oriented to a specific purpose.

Data Mart is small slices of the data warehouse.

These are the reasons we use Data Mart.

* Easy access to frequent data.
* Improve end-user response time
* Easy to create Data Mart.
* Less Cost
* Faster query performance
* Better data quality

Different Type Of D.M:

1. Dependent D.M
2. Independent D.M

* **Dependent D.M**:

In this data mart, is built by drawing data from a control data warehouse that already exists.

* **Independent D.M:**

In this data mart, is built by drawing from the operation or external source of data or both.

* **Data Lakehouse:**

A data lakehouse is a data management architecture that combines the benefits of a traditional [data warehouse](https://www.techtarget.com/searchdatamanagement/definition/data-warehouse) and a [data lake](https://www.techtarget.com/searchaws/definition/data-lake). Like a data lake, it can handle a wide variety of data formats and types, making it more flexible than a traditional data warehouse.

A data lakehouse is a centralized repository of data that can handle both structured and unstructured data, as well as real-time and batch data processing.

key features:

* Data integration capabilities.
* data quality
* data governance
* **Low-cost object store**
* **Declarative DataFrame API support**
* **Data Mesh:**

A data mesh is a decentralized data architecture that organizes data by a specific business domain, for example, marketing, sales, customer service, and more.

 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 the adoption of modern data architectures.

The benefits of Data Mesh

* **Total clarity into data’s value**
* **10X faster innovation cycles**
* **More than 70% reduction in data engineering**
* improved data agility
* faster time-to-market
* better data quality
* **Different between DWH vs Data Lake:**
* **DWH:**
* . A data warehouse is a repository of highly structured historical data which has been processed for a defined purpose.
* Data is structured and processed
* Processing schema-on-write
* Storage is expensive for large data volumes
* Security is mature
* Users business professionals
* Data warehouses typically have strict data governance and quality controls
* **Data Lake:**
* A data lake is a massive repository of structured and unstructured data, and the purpose of this data has not been defined
* Data is structured, semi-structured, and unstructured.
* Processing schema-on-read
* is designed by low-cost Storage
* Data lakes have looser governance and rely on data consumers to manage their own data quality and governance.
* Security is maturing
* User data scientists
* **Different between OLTP vs OLAP:**
* **OLTP:**
* OLTP (Online Transaction Processing)
* OLTP is designed for day-to-day transactional processing
* OLTP databases have a normalized data structure,
* OLTP databases are optimized for fast read and write operations
* OLTP databases typically handle large volumes of small transactions
* **OLAP:**
* OLAP (Online Analytical Processing)
* OLAP is designed for data analysis and decision-making
* OLAP databases have a denormalized data structure.
* OLAP databases are optimized for fast querying and reporting.
* OLAP databases typically handle large volumes of data for analysis and reporting.