Name: Ahsan Bilal

Data Marts

- It is a data store which is design for particular department f an organization or Data mart is a subset of data warehouse usually oriented to a specific purposes

- Reasons:
  - ➔ Easy access of frequent data
  - ➔ Improve end user response time
  - ➔ Less cost
  - ➔ Easy creation process

- Overall, data marts are a way to make it easier for businesses to access and analyze the data they need to make informed decisions.

Data Lakehouse

- A data lakehouse is a modern data architecture that combines the benefits of a data lake and a data warehouse. It is designed to store and manage large volumes of data, both structured and unstructured, from a variety of sources.

- In a data lakehouse, data is stored in a centralized repository where it can be easily accessed by various teams and applications. This data can be raw or structured and does not need to be transformed or processed before it is stored.

- The data lakehouse also includes features of a traditional data warehouse, such as data quality checks, data transformation, and organization into relevant schemas. This allows for easy querying, analysis, and reporting of data for business insights and decision-making.

- Overall, a data lakehouse offers a flexible, scalable, and cost-effective solution for managing large volumes of data and enabling data-driven decision-making in an organization.

Date: 23/3/2023

Name: Ahsan Bilal

Data Mesh

- Data Mesh is a relatively new approach to designing and managing data architecture within organizations. It is based on the idea that data should be treated as a product and that data ownership and responsibility should be distributed across teams instead of centralized within a single data team.

- In simple words, Data Mesh is an organizational structure that aims to enable more efficient and effective use of data by breaking down silos between different teams and empowering them to take ownership of their data domains. It emphasizes collaboration and decentralization of data management, allowing teams to work autonomously and making it easier to scale data projects across an organization.

OLTP VS OLAP

- OLTP (Online Transaction Processing) and OLAP (Online Analytical Processing) are two different types of database systems used for different purposes. Here are some of the key differences between the two:

- Purpose: OLTP systems are designed for transactional processing, where data is constantly being added, updated or deleted. OLAP systems are designed for analytical processing, where data is aggregated, sorted, and analyzed for reporting and decision-making purposes.

- Data structure: OLTP databases are typically normalized to minimize redundancy and optimize transaction processing. OLAP databases are usually denormalized to simplify queries and optimize analytical processing.

- Volume of data: OLTP systems deal with a high volume of real-time transactions and typically store smaller amounts of data. OLAP systems deal with a large volume of historical data and are optimized for queries and reporting.

- Complexity of queries: OLTP systems are optimized for simple queries that retrieve a small set of records. OLAP systems are optimized for complex queries that involve aggregations, grouping, and sorting of large datasets.

- User base: OLTP systems are typically used by operational staff, such as customer service representatives, while OLAP systems are used by analysts and decision-makers who need access to historical data for reporting and analysis.

Date: 23/3/2023

Name: Ahsan Bilal

Data Warehouse Vs Data Lake

Philosophy	Think first Load later	Load First Think later
Processing	Can only handle structured Data	Can handle Structered, Unstructured, Semi structured Data
Storage	Expensive for large data storage	Built for low cost storage
Ability	Less Agile in nature	Higly Agile in nature
Usage	Operational reporting	Use for advance data analytics
Security	Matured	Still maturing

Date: 23/3/2023