Task 2

15 March 2023

Data Marts:

1. Data mart is a subset of a data warehouse which is used by different departments within an organisation, such as sales,finance marketing etc.
2. They are created to address specific needs of each department so that they dont have to waste time on searching the required data in entire datawarehouse by allowing faster access to it.
3. They contain summarised data which stakeholders can use to quickly make informed decisions. For example a sales department could use sales related data mart to compile yearly sales report.

Data Lakehouse:

1. Data lakehouse is a mixture of both, data lake and data warehouse.
2. It stores structured and unstructured data like what data lake does but it also has this additional feature like structuring and transforming data like in data warehouses to make it ready and optimised for analysis.
3. It allows for faster data access and analysis since data teams can access it without need to access multiple systems.

Data Mesh:

1. Data mesh is a decentralised and domain oriented approach to manage data effectively.
2. Each domain in an organisation like sales,marketing etc has its own data who they are responsible for.
3. It breaks data silos where data is in isolation and there is no connection between different types of data.
4. It increases collaboration between different domains as they can share and access different domain’s data. They work together to maintain the data quality and make sure that it is available for anyone who needs it. This reduces time to find the desired data.

Datawarehouse vs Data lake:

1. In DWH, structured data is stored where as in data lakes structured, semi structured and non structured data are also stored.
2. DWH is designed to support business intelligence and reporting whereas data lake is used for analysis.
3. In DWH, before storing data, the schema/structure should be defined , whereas in data lake there is sometimes no need to define schema/structure of data.
4. After passing through ETL process, data is stored in DWH, whereas in data lakes , the data is stored in its raw/original form.

OLTP vs OLAP:

Online Transactional Processing and Online Analytical Processing, both are online processing frameworks used in organisations.

1. OLAP systems are used to analyze huge volume of data and handle compex queries which are aimed for decision making purposes.
2. It provides multi dimensional analytics capabilities to analyze large amount of data.
3. It is mostly built on top of DWH or data lake.
4. It is used by analytics and reporting applications such as Business Intelligence and Data Science tools.
5. OLTP systems are used to manage transactional data in real time. It focuses on processing individual transactions.
6. It is optimised for faster and efficient transactions processing.
7. The data stored is structured and normalized.
8. It is mostly built on top of relational database.
9. It is used by transactional applications such as Point of Sale, CRM, and ERP systems.