**Task 2**

**Data Marts**

A data mart is a structured data repository purpose-built to support the analytical needs of a particular department, line of business, or geographic region within an enterprise. Data marts are typically created as partitioned segments of an enterprise data warehouse, with each being relevant to a specific subject or department in your organization such as finance or sales. Data marts help you perform analysis faster given that you’re working with a smaller, more applicable dataset.

**Data Lakehouse**

A data lakehouse is a data management architecture that combines key capabilities of data lakes and data warehouses. It brings the benefits of a data lake, such as low storage cost and broad data access, and the benefits of a data warehouse, such as data structures and management features.

**Data Mesh**

**Data mesh** defines a platform architecture based on a decentralized network. The data mesh distributes data ownership and allows domain-specific teams to manage data independently.

The data mesh architecture treats data as a product. It distributes ownership to different end-users without requiring permission from a centralized system to access or manage it. This way, the linear data pipeline is eliminated while the central authority can easily monitor the system without approving every data request.

**- OLTP vs OLAP**

| **Basis of Differentiation** | **OLTP** | **OLAP** |
| --- | --- | --- |
| **Process** | Oltp refers to an online transactional system. These systems are used for the management of database modifications. | Olap systems are useful for online analysis and data retrieval processes. |
| **Characteristic** | Comprises of numerous short online transactions. | Comprises of large quantities of data. |
| **Functionality** | The oltp system is used for online database modifications. | Olap systems are designed to function as online database query management systems. |
| **Method** | Oltp makes use of traditional DBMS. | Olap makes use of a data warehouse to perform. |
| **Query** | Queries information to update, insert and delete data from a given database. | Mostly queries the select operations. |
| **Tables** | Oltp databases contain normalized databases. | The tables contained by olap databases are not normalized. |
| **Source** | The data sources in oltp systems comprise of OLTP and their transactions. | The sources of data in olap systems comprise of different OLTP databases. |
| **Data Integrity** | OLTP databases have to maintain the data integrity constraints. | As OLAP databases do not get modified frequently, data integrity does not serve to be an issue in them. |
| **Response time** | The response time of oltp systems is in milliseconds. | The response time of olap systems is in seconds to minutes |
| **Data quality** | The data available in OLTP databases/ processes is always organized and detailed. | The data available in olap processes may not be organized at all times. |
| **Utility** | These systems are useful for controlling and running different types of fundamental business tasks. | Olap systems are very helpful with planning, decision support and problem-solving. |
| **Operations** | Permits both read/write operations. | Olap systems are used for read operations and rarely write. |
| **Target audience** | Oltp caters via market orientated processes. | Olap takes care of customer orientated processes. |
| **Type of query** | The queries in oltp systems are simple and standardized. | The queries pertaining to olap systems involve aggregations that are generally complex in nature. |
| **Back-up** | An oltp system provides complete backup of all data in combination with incremental backups. | Olap systems require timely backup only. The need for back-up in their case is not as important as in the case of oltp systems. |
| **Database design** | DB design is mostly application oriented. For instance, the design changes with different industry verticals like banking, airlines, retail, etc. | DB design is generally subject oriented. For instances, database design changes comprising of subjects like purchasing, sales, marketing, etc. |
| **User type** | Oltp is useful for data critical users of the likes of clerks, Data Base and DBA professionals. | Useful for most data knowledge users of the kinds of workers, managers, CEOs, etc. |
| **Objective** | Oltps are designed for taking care of real time business operations. | Olap systems are designed for the analysis of business measures in terms of category and attributes. |
| **Performance metric** | In oltp, transaction throughput serves to be the key performance metric. | Query throughput serves as the main performance metric in olap systems. |
| **Number of users** | Thousands of users can access oltp databases. | Only hundreds of users are capable of using olap databases. |
| **Productivity** | Helps in increasing the productivity and self-service capabilities of users. | Increases the efficiency and productivity of business analysts. |
| **Challenges** | Historically, data warehouses have been built as costly, development projects. | OLAP cubes are not the same as open SQL server data warehouses. Given this, the knowledge and experience of trained professionals is required for the management of OLAP servers. |
| **Process** | Oltp systems offer fast results for the sake of daily used data. | Olap systems ensure that the response provided to queries is always quicker. |
| **Basic trait** | Oltps are easy to build and maintain. | Allows users to create views with the aid of spreadsheets. |
| **Style** | An oltp is designed for low data redundancy, fast response time, and is normalized. | As a data warehouse is built uniquely it is capable of integrating different data sources for the cause of building consolidated databases. |

**- DWH vs Data Lake**

| **Basis of Differentiation** | **Data Lake** | **Data Warehouse** |
| --- | --- | --- |
| Type of data | Unstructured and structured data from various company data sources | Historical data that has been structured to fit a relational database schema |
| Purpose | Cost-effective big data storage | Analytics for business decisions |
| Users | Data scientists and engineers | Data analysts and business analysts |
| Tasks | Storing data and big data analytics, like deep learning and real-time analytics. | Typically read-only queries for aggregating and summarizing data |
| Size | Stores all data that might be used—can take up petabytes! | Only stores data relevant to analysis |
|  | | |