

SADAF SULTAN

Bytewise fellowship

Task #2

Data Mart

It is a data storage system that contain information of a specific organization business unit. It's a small and selected part of data that company store in a large storage system. In simple words we can say it's a data warehouse focused on a single subject.

Data Lakehouse

Combination of structured and unstructured data. Data lakehouse is a data repository which contain all type of data.

Data Mesh

Data mesh emphasizes data as a product and promotes the use of standards and APIs to enable easy discovery and consumption of data across the organization.

Data mesh emphasizes the decentralization of data ownership and management across an organization.

Data mesh delegates datasets ownership from centralized to the domain, enabling , agility and scalability.

DWH vs Data Lake

Data warehouse is repository for structured and filtered data that has already been processed for a specific purpose. While on the other hand data lake is a

repository of all type of data we can combination of structured and unstructured data is known as data lake . everyone can pick any kind of junk of data for their business can pick and processed on that data.

OLTP vs OLAP

Online transaction processing(OLTP) capture, stores and processes data from transaction in real time. It's a type of data processing that executes transaction-focused task. It involves inserting, deleting, or updating small quantities of database data.

Online analytical processing(OLAP) uses complex queries to analyze aggregated historical data from OLTP system.

Task # 3

Can a database be used as DWH?

yes database is data repository of any special organization. Like we can say that database of army or byte wise is special data that can't share it with others. And database store the current data require power an application.

While on the other hand data warehouse is a type of database which store a lot of data and it store current and historical data for one are more systems. Like statebank have the data of all the banks but all the banks don't have the data of others bank.

Major differences between structured and unstructured data.

Structured data is stored into warehouses, while unstructured data stored into data lake (data lake contain both type of data).

Structured data is qualitative ,while unstructured data is quantitative.

Structured data is easy to search and analyze. While unstructured data requires more work and understand.

Structured data is clearly define and searchable type while unstructured data is usually stored in its native format.

What are the duties of data engineer?

- Handling ETL jobs
- Apply transformation based on the business requirements.
- Analyze and organize raw data.
- Build data systems and pipelines.
- Evaluate business needs and objectives.
- Interpret trends and patterns.
- Conduct complex data analysis and report on results.
- Prepare data for prescriptive and predictive modeling.
- Build algorithms and prototypes.