What is a Data Mart?

Data marts are specialized, subset of data warehouses focused on a particular line of business or department. They make it easy for that particular department to query data and make critical insights without having to search and query the entire data warehouse looking for relevant data.
While a data warehouse tasks data from many different sources and aggregates it into one repository, a data mart stores user/department specific data from that DWH. A data mart is optional in a business

What is a Data Lakehouse?

Data Lakehouses are a relatively new type of data management architecture that combine the flexibility, scale and cost-efficiency of data lakes with the data management and control of data warehouses. A data lakehouse, like a data lake, can store both structured and unstructured data allowing users to query only one site for all types of data.
Data lakehouses make use of intelligent metadata layers that can effectively “structure” unstructured data. For example, part of this metadata extraction might include using computer vision or NLP to read the contents of an image or text file and categorize them accordingly.

What is a Data Mesh?

Similar to how modern data applications have adopted the microservice architecture as opposed to the monolithic architecture, A data mesh is like the microservice approach in terms of data. A data mesh leverages a domain-driven, self-serve data infrastructure.
In summary, a data mesh architecture:

- Gives individual teams control over certain datasets (domain-driven data ownership, not relying on data product owners)
- Focuses on data as a product
- Emphasizes a self-service infrastructure
- Ensures data governance and security

Data Warehouse VS Data Lake

Data lakes and data warehouses are both widely used for storing big data, but they are not interchangeable terms.
Here are 4 key differences betweent the two:

Property	Data Lake	Data Warehouse
Data Structure	Raw	Processed
Purpose of Data	Not yet determined	Currently in use
Users	Data Scientists/ ML Engineers	Business Professionals/Analytics
Accessibility	Highly accessible and quick to update	Not updated frequently

OLTP Vs OLAP:

The main difference between the two is the type of processing that each of them are optimized for.

OLAP is optimized for conducting complex data analysis for smarter decision-making. OLAP systems are designed for use by data scientists and business analysts, and they support business intelligence (BI), data mining and other decision support applications.

OLTP, on the other hand, is optimized for processing a massive number of transactions. OLTP systems are designed for use by frontline workers (e.g., cashiers, bank tellers, hotel desk clerks) or for customer self-service applications (e.g., online banking, e-commerce, travel reservations).

Can a database be used as a Data Warehouse?

Yes they can be used as a data warehouse

What is the major difference between structured and unstructured data?

The main difference between structured and unstructured data is that the former adheres to a predefined format and stored in such a way while the later doesnt follow any format and is a combination of different types of formats depending on where the data originated from.

What are the duties of a Data Engineer?

Some of the main responsibilities of a data engineer are:

1. Build and maintain data systems and pipelines
2. Analyze and organize raw data
3. Prepare data for predictive modeling
4. Evaluate business needs