Day 1

Name: Krushnakumar Patle

Email: krishnapatle128@gmail.com

Data Engineering- Batch 1

❖ Explain data warehousing concept.

- Data warehousing is a concept and practice in the field of information technology that involves the collection, storage, and management of data from various sources within an organization.
- The goal of a data warehouse is to provide a centralized and unified repository of data for analysis and reporting. This helps organizations make informed decisions based on a comprehensive view of their data.

❖ Features of **Data Warehouse**: -

1. Subject-Oriented:

   - Focus on Business Areas: A data warehouse is subject-oriented, meaning it is designed to provide information about a specific subject or business area rather than the ongoing operations of an organization.
   - Tailored for Analysis: Data is organized and presented in a way that is most relevant to the needs of decision-makers, allowing them to analyze and gain insights into specific business topics.

2. Integrated:

   - Combination of Data Sources: Integration involves bringing together data from different sources, which may include various departments, operational systems, and external entities.
   - Consistent Format: The data is transformed and standardized during the ETL process to ensure a consistent format, allowing for seamless analysis and reporting.

3. Time-Variant:

   - Historical Perspective: Data warehousing systems store historical data, allowing users to analyse changes and trends over time.
   - Time-Stamping: Historical data is often time-stamped to track when it was collected, providing a temporal dimension for analysis.

4. Non-volatile:

   - Data is Read-Only: Once data is loaded into the data warehouse, it becomes non-volatile, meaning it is primarily for querying and reporting purposes.

- Minimal Data Modification: Unlike operational databases, where data is constantly updated, inserted, or deleted, data warehouses are optimized for read-intensive operations.