**TASK 3 QUESTIONS 15-03-23**

**Q1- Can a data base be used as DWH?**

Answer- No, database cannot be used as DWH because database stores real-time information about one particular part of the business. Its main job is to process the daily transactions that your company makes. Databases handle a massive volume of simple queries very quickly.

While a data warehouse is a system that pulls together data from many different sources within an organization for reporting and analysis. The reports created from complex queries within a data warehouse are used to make business decisions. A data warehouse stores historical data about your business so that you can analyze and extract insights from it. It does not store current information, nor is it updated in real-time.

**Q2- Major differences between structured and Un-Structured data?**

Answer- Here are the five differences between Structured and Un-structured data:

1. Structured data is standardized, clearly defined, and searchable data, while unstructured data is usually stored in its native format.
2. Structured data is quantitative, while unstructured data is qualitative.
3. Structured data is often stored in data warehouses, while unstructured data is stored in data lakes.
4. Structured data is easy to search and analyze, while unstructured data requires more work to process and understand.
5. Structured data exists in predefined formats, while unstructured data is in a variety of formats.

**Q3- What are the duties of a Data Engineer?**

Answer- A data engineer is responsible for collecting, managing, and converting raw data into information that can be interpreted by data scientists and business analysts. Data accessibility is their ultimate goal, which is to enable organizations to utilize data for performance evaluation and optimization.

These are some common tasks data engineer might perform when working with data:

* Acquire datasets that align with business needs
* Develop algorithms to transform data into useful, actionable information
* Build, test, and maintain database pipeline architectures
* Collaborate with management to understand company objectives
* Create new data validation methods and data analysis tools
* Ensure compliance with data governance and security policies