Task 2

15 March 2023

Q: Can a database be used as DWH?

Yes, a database can be used as DWH but the real purpose for which database is built is not the same as DWH. A database is used for transactional processing and handle day to day business operations such as inventory, customer, orders etc. On the other hand, DWH is a centralised repository where the data is stored from different sources and is prepared/optimised for querying, analysing and reporting large amount of data after series of ETL process.

While a database can be used as DWH, but they are not built to handle large amount of data efficiently and provide analytical capabilities as provided by DWH.

Q: Structured Data vs Un Structured Data

* Structured data is data which has already defined and organised structure.
* It is usually stored in databases or in pre defined format/schema.
* It is easier to analyze, query and process it using RDBMS.
* Examples include data like customer info, inventory etc.
* Un structured data is data which has no pre defined structure.
* It is usually stored in text, audio, video and pictures format.
* It is difficult to analyze, query and process it because of its unorganised nature.
* Examples include social media posts, youtube videos etc.

Q: What are duties of Data Engineer?

* Design data pipelines(ETL pipelines).
* Build data infrastructure.
* Building and maintaining data storage systems.
* Integrate data from different sources to create a unified view of data.
* Ensure data quality.
* Collaborate with data scientists/analysts.