**Data Engineering**

**By: Shaharyar Sabir**

**Week#01**

**Task # 2:**

Familiarize yourself with the following topics:

- **Data Marts** is a part of data warehouse related to a specific field like business and sales. It works on a single function. As warehouse consists of large amount of data from multiple different sources so data mart is made as a subset of data warehouse that belongs to individual departments.

- **Data Lakehouse**  as the name suggests, is a new data architecture that merges a data warehouse and a data lake into a single whole, with the purpose of addressing each one’s limitations. Allows low-cost storage of large volume of data in a raw format just like data lakes.

- **Data Mesh** collection of data from specific business domain like sales, marketing, customer services etc. so that it will be easily accessible for business users.

- **DWH vs Data Lake** data lake contains organizational data in a raw, unstructured manner and can store it indefinitely. While data ware house contains structured, clean and processed data.

- **OLTP vs OLAP:**

**OLTP:** handles large number of small transactions. Simple queries. Based on insert, update, delete commands. Control and run essential business operations in real time.

**OLAP:** Handles large volume of data with complex queries. Based on select command. Plan, solve problems, support decisions, discover hidden insights.

**Task # 3:**

**After you complete these topics, please answer the following questions in your document:**

**- Can a database be used as DWH?**

Data ware house is a type of database that may serve different purposes. OLTP are best used with database while OLAP are best with DWH. Database can handle thousands of users at a time while data ware house handles a small number of users.

**- Major differences between structured and Un-structured data.**

Structured data is easier to search and use as it is older so many analytic tools are available while unstructured data involves complex search and for use they need processing.

- **What are the duties of a data engineer? (high-level)**

Analyzing raw data, building datasets, pipelines , maintaining quality and improving efficiency.