ByteWise Fellowship – Week 1

Task#4 – ETL/ELT/3tier Date: 16-March-2023

Question#1

**What is 3tier architecture?**

It involves separating the application into three distinct layers or tiers, each with its own set of responsibilities and functionalities. The three tiers are:

* **Presentation Tier**: This is the user-facing layer where the user interface is located. The presentation tier is responsible for presenting data to the user in a way that is easy to understand and interact with. It handles user input and sends requests to the application server for processing.
* **Application Tier**: This is the middle layer of the 3tier.The application tier is responsible for processing user requests, managing data, and communicating with the data tier. It acts as a bridge between the presentation tier and the data tier.
* **Data Tier**: This is the layer that contains the data storage and retrieval systems. The data tier is responsible for managing and storing data, providing data access to the application tier, and ensuring data integrity and security. It can include relational databases, NoSQL databases, DWH, and other data storage systems.

**Application:**

A real-life example of a working 3-tier architecture in data engineering is an **online shopping application.**

* Presentation Tier: In an online shopping application, the presentation tier would include the user interface where customers can browse products, add items to their cart, and check out. The presentation tier would also handle user authentication and authorization.
* Application Tier: This layer handle tasks such as calculating shipping costs, managing inventory, and processing payments. It would also communicate with the data tier to retrieve and store data.
* Data Tier: It would include a database that stores information about products, customer orders, and payment details.

Question#2

**What is ETL?**

The ETL process involves extracting data from various sources, transforming it into a usable format, and loading it into a data warehouse or another target system. The process typically involves several components and steps

* **Extraction**: The first step in the ETL process is to extract data from various sources. These sources could include relational databases, flat files, web APIs, or other systems. The extraction process involves identifying the data sources and selecting the data to be extracted. This could involve filtering the data based on specific criteria or selecting only the required columns.
* **Transformation**: Once the data is extracted, the next step is to transform it into a usable format. This involves applying business rules, data validation, data cleaning, and data enrichment techniques to ensure that the data is consistent, accurate, and complete. The transformation process could include merging data from multiple sources, converting data types, and creating new fields.
* **Loading**: Once the data is transformed, it is loaded into a target system such as a data warehouse, a data lake, or a BI system. This step involves identifying the target schema, creating tables, and defining the relationships between the tables. The data is loaded into the target system in batches or in real-time, depending on the requirements.

Question#3

**What is ELT?**

**ELT (Extract, Load, Transform)** is a data integration process that is similar to ETL but with the transformation step being performed after the data has been loaded into the target system. The ELT process involves extracting data from various sources, loading it into a target system, and then transforming it into a usable format. The process typically involves several components and steps.

The main difference is that in ELT, the data is available in raw for to the consumers who can transform it as per their need this option is not available in ETL as the consumer get the specific transformed data.

Due to this “Scalability” advantage of ELT, it has become more popular and more in use than ETL.

Question#4

**ETL Tools:**

* Apache Kafka
* Talend Open Studio
* Apache Airflow