**ETL and 3-Tier in DE**

**Submitted by:** Afraz Ahmed

* **ETL:**

ETL (Extract, Transform, Load) is a process used to extract data from one or more sources, transform it to fit the desired data model or format, and then load it into a destination data storage system such as a data warehouse, database or a data lake. The ETL process is used to move data from transactional systems to analytical systems where it can be analyzed and used to make data-driven decisions.

Extract: This is the process of extracting data from various sources such as databases, flat files, APIs, and other sources. In this stage, the data engineer identifies the source system(s) that will provide the data for the ETL process. The data is then extracted from the source systems using various techniques such as API calls, database queries, or file transfers.

Transform: Once the data has been extracted, it is transformed into a format that is suitable for the target system. The transformation process involves cleaning, filtering, aggregating, joining, and enriching the data.

Load: After the data has been transformed, it is loaded into the destination system such as a data warehouse, database, or a data lake. This stage involves selecting the appropriate data storage technology and designing an optimized data loading process that is scalable and efficient.

* **ELT:**

"ELT" stands for Extract, Load, Transform, and it refers to a data integration process that involves three distinct steps:

Extract: In the first step of ELT, data is extracted from various sources such as databases, spreadsheets, or even web pages. The extracted data can be in any format such as CSV, XML, JSON, or other formats.

Load: The next step is to load the extracted data into a target system such as a data warehouse or a big data platform. This step involves moving the data from the source to the target system, where it can be stored and used for further analysis.

Transform: In the final step, the extracted data is transformed into a format that can be easily analyzed and queried. This transformation may involve cleaning the data, removing duplicates, enriching the data with additional information, or aggregating the data.

* **3 tools for ETL:**

1. Apache NiFi
2. Apache Spark
3. Talend