**SYEDA FATIMA SHAH**

**16th March 2023**

***Task # 4:***

**- What is ETL? in detail.**

ETL stands for Extract, Transform, Load, which refers to a process of data integration and transfer. It involves extracting data from various sources, transforming the data to fit a common data model or structure, and then loading the transformed data into a destination database or data warehouse for analysis, reporting, and other purposes.

The ETL process is commonly used in business intelligence, data warehousing, and analytics. The extract step involves pulling data from various sources such as databases, spreadsheets, files, and web services. The transform step involves converting the data to a common format and structure, cleaning and filtering it, and possibly performing calculations or aggregations. The load step involves transferring the transformed data to a target system such as a database or data warehouse.

Effective ETL processes require careful planning, design, and testing to ensure data accuracy, consistency, and completeness. Various tools and platforms are available to facilitate ETL processes, including open-source software, commercial products, and cloud-based services.

**- What is ELT? in detail.**

ELT stands for Extract, Load, Transform, which is a data integration process similar to ETL, but with a different order of operations. While ETL first extracts data from source systems, then transforms it, and finally loads it into a target system, ELT reverses the order of the transform and load steps.

In an ELT process, data is first extracted from source systems and then loaded into a target system such as a data warehouse or a big data platform. After the data is loaded, it is transformed within the target system using tools such as SQL, MapReduce, or Spark.

The advantage of ELT over ETL is that it allows for greater flexibility and scalability in data processing. By loading data first and transforming it later, ELT avoids the need to transform all data before loading, which can be time-consuming and resource-intensive. ELT can also take advantage of the distributed processing power of big data platforms, allowing for faster processing of large datasets.

ELT is commonly used in modern data integration architectures that leverage cloud-based storage and computing resources, such as Amazon S3 and Redshift, Google Cloud Storage and BigQuery, or Microsoft Azure Data Lake and SQL Data Warehouse.

**- 3 Tier Architecture in DE**

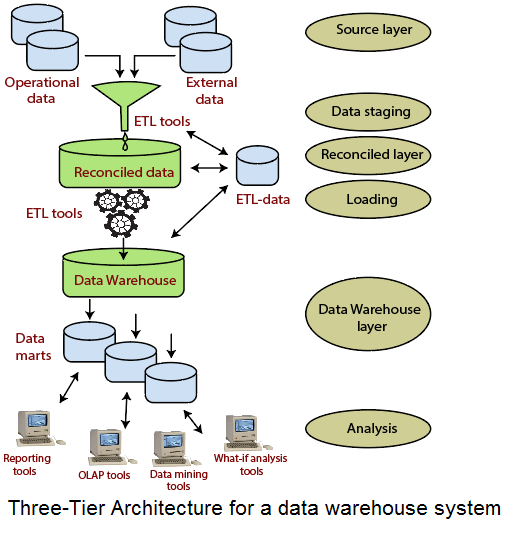
Three-tier architecture is a common approach to designing and implementing data engineering solutions. It is a client-server architecture that separates the presentation layer, application logic layer, and data storage layer into three distinct tiers.

The three tiers are as follows:

1. **Presentation tier:** This is the top layer of the architecture, which handles the user interface and user interaction. It includes applications such as web browsers, mobile apps, or desktop applications that interact with the middle tier to request and display data.
2. **Application logic tier:** This is the middle layer of the architecture, which processes user requests, performs business logic, and accesses data from the data storage layer. It includes web servers, application servers, and APIs that provide an interface between the presentation tier and the data storage tier.
3. **Data storage tier:** This is the bottom layer of the architecture, which stores data used by the application logic tier. It includes databases, data warehouses, and other data storage systems.

The advantages of the three-tier architecture in data engineering include modularity, scalability, and flexibility. By separating the presentation layer, application logic layer, and data storage layer, each tier can be designed and optimized independently, allowing for easier maintenance and upgrading of individual components. The architecture also enables scalability by allowing multiple instances of the application logic tier and data storage tier to be deployed and managed independently.

Overall, the three-tier architecture is a widely used and effective approach to designing and implementing data engineering solutions.



**- ETL Tools (any 3)**

Graphical user interface, text, application

Description automatically generated

***Task # 5:***

**- What is Historical Load**

Historical load refers to the past energy consumption patterns of a power grid or electrical system. It refers to the amount of electricity that has been consumed by a particular area or region over a specific period of time, usually measured in hours, days, weeks, or months.

Utility companies use historical load data to forecast and plan for future energy demand. By analyzing historical load data, they can identify trends, patterns, and peak usage periods that can help them anticipate future energy demand and adjust their supply accordingly.

Historical load data can also be used to evaluate the performance of energy conservation programs and to determine the effectiveness of energy management strategies. It is an important tool for energy planners, grid operators, and policymakers in managing the supply and demand of electricity in a reliable and efficient manner.

**- What is Full Load**

Full load refers to the maximum electrical load or power that a device, appliance, or electrical system can handle without exceeding its capacity. It is the point at which the power supply is providing the maximum amount of power to the device or system.

In the context of a power grid or electrical system, full load refers to the maximum power demand that can be sustained by the system without overloading or causing blackouts. It is the point at which the power generation capacity of the grid is fully utilized.

For example, a generator has a rated capacity, which is the maximum amount of power it can produce continuously without overheating or damaging the equipment. When the generator is operating at its rated capacity, it is said to be at full load.

In summary, full load is the maximum amount of power that a device, appliance, or electrical system can handle, and in the context of a power grid, it is the maximum power demand that can be sustained without overloading the system.

**- What is Incremental Load**

Incremental load refers to the additional load or demand that is placed on an electrical system when a new device, appliance, or piece of equipment is added to the system. It is the difference between the current load and the load that existed before the new device was added.

For example, if a household has an average power consumption of 5 kW, and a new air conditioning unit with a power rating of 2 kW is added, the incremental load would be 2 kW. The total load on the electrical system would now be 7 kW.

Incremental load is an important factor to consider when planning for new equipment or upgrades to an existing electrical system. It helps to determine the impact of the additional load on the system's capacity and whether it can handle the increased demand without overloading or causing other problems.

Electric utilities and grid operators often use incremental load data to forecast future energy demand and plan for new generation capacity or transmission infrastructure to meet the growing demand.