Fariz khanzada

Task 5

Q .1

Historical load is the process of loading historical data into a data warehouse or data lake. It involves extracting data from various sources such as transactional databases, files, or APIs and transforming it into a format that can be stored and analyzed over time. Historical data can include data from several years or decades, and it can help organizations identify trends, patterns, and insights from the past to make informed decisions for the future. Historical load is typically done as a one-time process and is often followed by incremental loads to keep the data warehouse or data lake up to date.

Q.2

Full load is a type of data integration process where all data from a source system is extracted and loaded into a target system in a single batch. This process involves reading all the data from the source system, applying any necessary transformations, and loading the data into the target system. Full load is typically used when creating a data warehouse or data mart for the first time or when rebuilding it from scratch. This process can be time-consuming and resource-intensive, especially when dealing with large volumes of data, but it ensures that all the data is available in the target system. Full load can be contrasted with incremental load, which only extracts and loads new or changed data since the last load, and is typically faster and more efficient.

Q.3

Incremental load is a type of data integration process that involves extracting only the data that has changed or been added since the last data load and loading it into a target system. This process is typically used to keep a data warehouse or data lake up to date and to minimize the amount of time and resources required to update it. Incremental load can be implemented in different ways, such as using change data capture (CDC) technology to capture changes in the source system or using time-based filters to extract data that falls within a specific time window. Incremental load is faster and more efficient than full load, especially when dealing with large volumes of data, and it helps ensure that the data in the target system is always current.