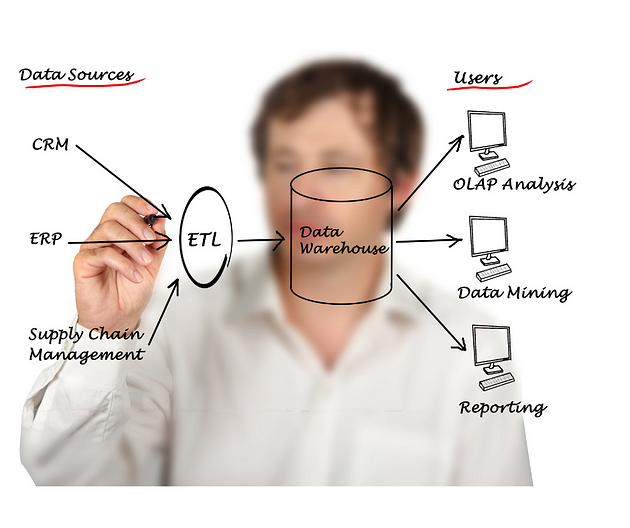
# **Data Warehousing- Historical Load, Full Load & Incremental Load**

The data warehouse is a large centralized repository that is designed to support business decision-making. It helps in storing, integrating, and analyzing data from multiple sources to provide a comprehensive view of the organization’s operations by identifying trends and patterns.

Data in the warehouse is often subject-oriented, meaning that it is organized around specific business topics, such as sales, customer behavior, or inventory, and are built using ETL.



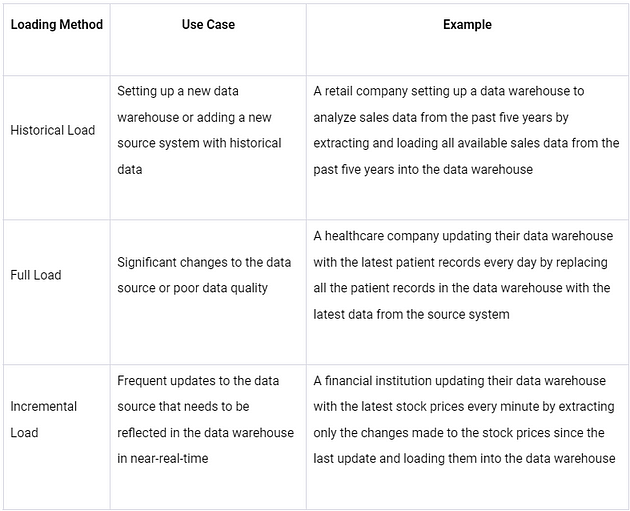
*ETL (detail) read: “*[*The Backbone of Data Management: A Deep Dive into the Responsibilities of a Data Engineer”*](https://mesum.medium.com/the-backbone-of-data-management-a-deep-dive-into-the-responsibilities-of-a-data-engineer-f1f49dfd2f24)

To populate or update a warehouse, different methods are used, including historical load, full load, and incremental load.

* Historical load refers to the process of populating a data warehouse with all available data from a certain point in the past up to the present. This is usually done when setting up a data warehouse for the first time or when adding a new source system that contains historical data. Historical load is a one-time process and can be time-consuming and resource-intensive.
* Full load, also known as a “rebuild,” is a process where all the data in a table or data warehouse is replaced or refreshed. This is typically done periodically, such as daily or weekly, to ensure that the data in the data warehouse is up-to-date. A full load is useful when there are significant changes to the data source or when the data quality is poor.
* Incremental load is a process where only the changes made since the last load is added to the data warehouse. This is usually done to update the data warehouse in real-time or near-real-time, and it is more efficient than a full load because it only processes the changes. The incremental load can be scheduled to run automatically at regular intervals or triggered by an event, such as a new transaction in the source system



Let’s understand with the help of examples and use case of each method for a clear image:



Note: The choice of loading method will depend on the specific requirements and constraints of each situation, and careful evaluation of the advantages and disadvantages of each method is necessary to choose the one that best meets the business needs.