What is Historical Load?

A data warehouse contains historical data. Based on user requirements data in warehouse has to be retained for a particular duration of time. This duration could be anywhere from a single year to several decades.

When a data warehouse is created i.e., tables in them are created, it contains no records. This is since planning for creation of a warehouse could take several months or years. During this time there would be lot of data in OLTP systems which act as source systems for warehouse. Loading of this initial historical data into warehouse is initial historic load.

Sometimes it may so happen that when data is loaded regularly into the warehouse ETL process might break and fixing it would take several hours to several days. During this fix time there will be data in OLTP systems. Loading this data is also a historic load.

What is Incremental Load?

Incremental load is the periodic load of data into warehouse. This process loads the most recent data from OLTP systems. This process run periodically till the end of warehouse's life. Incremental loads could run daily, weekly, fortnightly, monthly, quarterly, yearly or at a scheduled time. For every incremental load there is a load window within which the ETL load process should start and finish loading into target warehouse. After end of load window, business users will usually start querying and analyzing data in warehouse.

What is Full load?

In a Full Data Load, the **complete dataset is emptied or loaded** and then entirely overwritten (i.e., deleted and replaced) with the newly updated dataset in the next data loading run. While comparing the Incremental Data Load vs Full Load, you also don’t need to maintain extra information such as timestamps to carry out a Full Data Load.

You can consider a simple example of a shopping mall that loads all the total daily sales via the ETL process into a Data Warehouse at the end of each day. Assume that there were 1000 sales done on Monday, thus, you would need to load data on Monday night with a dataset of 1000 records. Then, on Tuesday 700 more sales were done and need to be added. Similarly, on Tuesday night, 1000 Monday records, as well as 700 Tuesday records, will now be dumped in the Data Warehouse via the Full Load method.