Task: 05– by Bytewise Team

# **What is Historical Load?**

Historical load in data engineering is a way to store and organize data that happen in past, just like a history book. Let’s understand it better with help of a real-world example:

**E.g** Imagine you have a diary, where you write everything that happens to you daily. You write what ate for breakfast. What you played today and study and more. Now after a week, you can look back and read what you did on a specific day.

Just like the above stores store their daily sales data on the computer. Using that data they can check which products sell more and are more popular.

In data engineering, the historical load is important because it allows us to analyze and learn from data that happened in the past. Just like your diary helps you remember what you did last week, historical load helps companies remember what they did in the past, so they can make better decisions in the future.

# **What is a Full Load?**

In Full load, you will understand it by expanding on the previous examples. Just like above where you write everything about yourself. What if you have to write or store and organize data of a whole city? That's a lot of information, right?

In the same way, companies also need to keep track of a lot of information about their business. They need to keep track of things like how many products they sell, how much money they make, and how many customers they have. They all store it in their database obviously.

Now, sometimes they need to update this information, like when someone buys a product or returns something. When they update the information in the database, they need to make sure that they have all the latest information, so they do something called a **"full load"**.

So, in data engineering, a full load is when you take all the information you have and put it in the right place in the database. It's an important part of making sure the database has all the latest information, so the company can make good decisions based on that information.

# **What is Incremental Load?**

incremental load is a way of only updating the parts of the data that have changed, instead of looking at everything every time.

**E.g** Image you are a mechanic and you have a big list of all the tools in the toolbox and you want to keep track of which one you used each day. If you had to go through the whole list every day and mark off the tools you used, it would take a long time. But if you only looked at the new tools that you got that day and marked them off, it would be much faster. Well, that's kind of what incremental load means in data engineering.

In data engineering, incremental load means only updating the data that has changed since the last time you looked at it. For example, let's say you have a big spreadsheet with a list of all the tools in the toolbox and when you last used them. Instead of updating the whole spreadsheet every day, you can just add the new tools that you got that day and mark off the ones you used. This is much faster and saves a lot of time