

## **DATA PREPARATION**

Data preparation is the process of preparing raw data so that it is suitable for further processing and analysis. Key steps include collecting, cleaning, and labeling raw data into a form suitable for machine learning (ML) algorithms and then exploring and visualizing the data. Data preparation, also sometimes called “pre-processing,” is the act of cleaning and consolidating raw data prior to using it for business analysis and machine learning. Good data preparation allows for efficient data analysis, limits errors and inaccuracies that can occur to data during processing, and makes all processed data more accessible to users. It's also gotten easier with new tools that enable any user to cleanse and qualify data on their own.