<https://www.simplilearn.com/data-profiling-tools-and-best-practices-article>

## **What Is Data Profiling (DF)?**

It is the process of examining source data and understanding structure, content, and interrelationships between data. The method uses a set of business rules and analytical algorithms to analyze data minutely for discrepancies. [Data Analysts](https://www.simplilearn.com/what-does-a-data-analyst-do-article) then use that information to interpret how those factors can align with business growth and objectives.

Tools : Informatica, Erwin Data Quality, Data Ladder,Microsoft Purview Data Quality

Data profiling is increasingly vital for businesses as it helps determine data accuracy and validity, risks, and overall trends. It can eliminate costly errors that usually occur in customer databases, like missing values, redundant values, values that do not follow expected patterns, etc. Companies can use the valuable insight gained from data profiling to make critical business decisions.

Most commonly, it  is used in combination with an ETL (Extract, Transform, and Load) process for data cleansing or [data scrubbing](https://www.simplilearn.com/what-is-data-scrubbing-article) and moving quality data from one system to another. An example can help you understand what is DF in ETL. Often ETL tools are used to move data to a data warehouse. Data profiling can come in handy to identify which data quality issues need to be fixed in the source and which issues can be fixed during the ETL process.

<https://www.simplilearn.com/what-is-data-scrubbing-article>

In [Data Analytics](https://www.simplilearn.com/data-analytics-bootcamp?source=GhPreviewCoursepages), [data cleaning](https://www.simplilearn.com/data-cleaning-why-and-how-to-get-started-article), also called data cleansing, is a less involved process of tidying up your data, mostly involving correcting or deleting obsolete, redundant, corrupt, poorly formatted, or inconsistent data.

 Data professionals do the actual cleaning, checking the database and making corrections and edits as needed, and practicing good data entry habits.

Consider data scrubbing as a subset of data cleaning. Data scrubbing employs actual tools to do a much “deeper clean” than just having a user pore over database spreadsheets and making corrections.