

DATA CLEANING WITH MYSQL

Erroneous, missing, or duplicate data in a dataset can be fixed or eliminated through the process of data cleaning. This crucial phase ensures that data analysis and decision-making are accurate and reliable.

**Procedure for data cleansing that will aid in data preparation**:

* Eliminate redundant data
* Convert data types
* Ensure overall consistency
* Handle missing values
* Fix errors
* Handle errors
* Maintain data in a unified form

Why SQL?

Programming languages like SQL (Structured Query Language) are used to administer and work with relational databases. In addition to its effectiveness in retrieving, filtering, updating, and deleting data, it also has applications for data cleansing jobs.

With SQL, you can use a declarative approach, meaning you can tell the database engine what data you want and leave the rest to it, including how to obtain or edit it.

SQL is hence an effective tool for data purification procedures.

Most data pipelines need the use of SQL. When building a database on the cloud, SQL is considered far more effective at cleaning data than scripting languages.

**PROBLEM STATEMENT : Corporate Layoffs**

Layoffs across a variety of global companies are covered in detail in the dataset layoffs.csv. Analysis of these layoffs offers important insights about the status of various industries, firm stages, and geographic areas, as an increasing number of businesses experience financial strain and restructuring. Details about the following are included in the dataset:

* **Company Name**
* **Location**
* **Industry**
* **Number of Employees Laid Off**
* **Percentage of Workforce Affected**
* **Company Stage (e.g., Post-IPO, Series B)**
* **Funds Raised by the Company**
* **Date of Layoff Announcement**