

Structured Query Language

# SQL





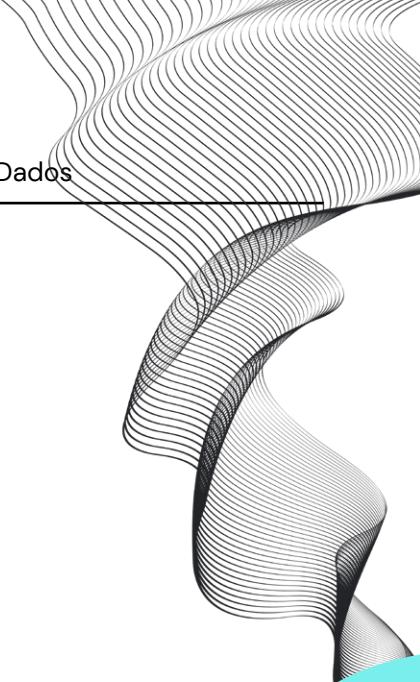
# Jornada do curso

01 Básico

02 Intermediário

03 Indo além

04 Processo de ETL com SQL

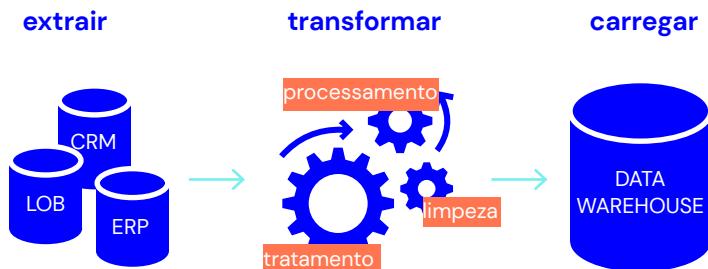




**E**XTRACT  
(extrair)

**T**RANSFORM  
(transformar)

**L**OAD  
(carregar)



## Processo de ETL

A sigla ETL corresponde à sequência de processos realizados com os dados – extração, transformação e carregamento.

Essa sequência de processos tem por objetivo **unir**, em um único banco, **dados de diferentes fontes**, podendo proporcionar uma visão geral dos dados da empresa. Isso é fundamental para a estratégia de análise e o tratamento de dados.



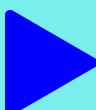
## 55% dos líderes

não confiam nas respostas dos dados que recebem.

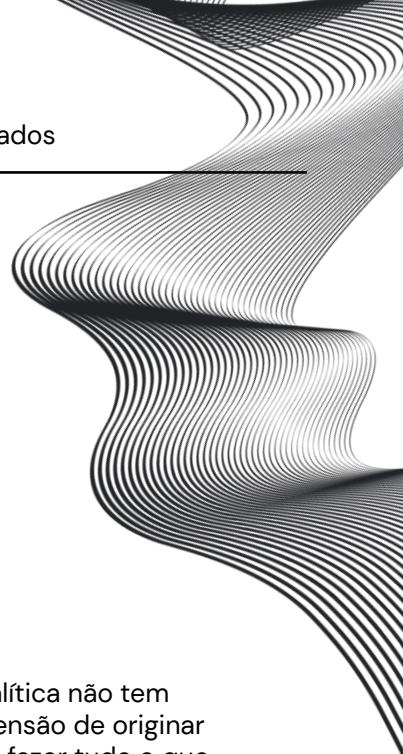
*Fonte: Experian, 2021.*

O ETL guia o tratamento de dados e, por isso, possibilita a criação de **dados confiáveis**.

# Dado **bom** é dado **bem tratado**.



O melhor dado que você tem é aquele que você pode **usar, confiar e basear** as tomadas de decisão e as ações da sua empresa. Para isso, ele precisa ser bem tratado.



## Missão da aula

Aprender a fazer ETL em suas bases para aumentar a confiabilidade da sua análise.



A máquina analítica não tem nenhuma pretensão de originar nada. Ela pode fazer tudo o que soubermos para que ela execute.”

[Ada Lovelace](#)

## Etapas

- › Importação dos dados e enriquecimento de informações.
- › Validação dos dados.
- › Reconstrução e padronização dos dados.
- › Remoção de duplicados e entrega diferenciada.



# Pulo do gato

Não faça alterações diretamente na sua fonte.

A melhor prática quando você está trabalhando com dados é não modificar o arquivo fonte nem as tabelas que você carregou. Faça uma cópia ou uma tabela temporária para poder realizar as transformações que precisa. Assim, se houver necessidade de conferir algo, você terá o dado original intacto.

---

## Comandos

### SET UPDATE

muda o que está marcado na coluna que escolher.

### DELETE

deleta os dados escolhidos.

Importante: usar **WHERE** em conjunto com ambos os comandos; assim, o comando só modifica ou deleta o que você de fato precisa, e não tudo.

### CREATE TABLE

cria uma nova tabela no banco de dados.

### /\* \*/

cria um bloco de anotações do próprio código, possibilitando organizar e categorizar cada bloco de código.



# Pulo do gato

Faça um **SELECT** antes de fazer **SET UPDATE** ou **DELETE**.

Para garantir que não haja alterações indevidas, o ideal é primeiro fazer um **SELECT** de algumas linhas ou um **SELECT COUNT(\*)**, para poder confirmar antes de executar o **SET UPDATE** ou o **DELETE**.

Isso simplifica o comando e valida se o código que você está escrevendo é realmente o que você precisa.

---

# Mão na massa

Quais são os 1.000 melhores clientes para oferecer empréstimo na ação da próxima semana?





# Tão importante quanto encontrar, é mostrar o que encontrou.

Não adianta fazer um ETL incrível e não mostrar os resultados do processo. Não só a base, mas, principalmente, um resumo dos seus dados ou até mesmo um relatório. Vale também pensar em levar esses dados em SQL para o Power BI e criar uma visualização que seja interessante. Recomendamos o curso de [Análise de Dados e Power BI](#) da **Conquer Plus** que aparece no “Quero mais” :)

## Desafio Conquer

Encontre uma base no site *Kaggle* que tenha mais de uma planilha e realize todas as etapas que aprendemos até aqui.



## Quero mais:



**Conquer Plus:**  
Análise de dados e Power BI



**Conquer Plus:**  
Lógica da programação



## Anotações

