

O que são dados sujos?

Anteriormente, discutimos que **dados sujos** são dados incompletos, incorretos ou irrelevantes para o problema que você está tentando resolver. Esta leitura resume:

- Tipos de dados sujos que você pode encontrar
- O que pode ter causado a sujeira dos dados
- Como os dados sujos são prejudiciais às empresas

TIPOS DE DADOS SUJOS

DADOS DUPLICADOS

Descrição	Causas possíveis	Potenciais danos às empresas
Qualquer registro de dados que apareça mais de uma vez	Entrada manual de dados, importação de dados em lote ou migração de dados	Métricas ou análises distorcidas, contagens ou previsões inflacionadas ou imprecisas, ou confusão durante a recuperação de dados

DADOS DESATUALIZADOS

Descrição	Causas possíveis	Potenciais danos às empresas
Qualquer dado antigo que deva ser substituído por informações mais recentes e mais precisas	Pessoas mudando de papéis ou empresas, ou software e sistemas se tornando obsoletos	Percepções imprecisas, tomada de decisões e análises

DADOS INCOMPLETOS

Descrição	Causas possíveis	Potenciais danos às empresas
Qualquer dado que esteja faltando campos importantes	Coleta de dados incorreta ou entrada de dados incorreta	Diminuição da produtividade, percepções imprecisas ou incapacidade de completar serviços essenciais

DADOS INCORRETOS/INCISOS

Descrição	Causas possíveis	Potenciais danos às empresas
Qualquer dado completo, mas impreciso	Erro humano inserido durante a entrada de dados, informação falsa ou dados falsificados	Percepções imprecisas ou tomada de decisões baseadas em informações erradas resultando em perda de receita

DADOS INCONSISTENTES

Descrição	Causas possíveis	Potenciais danos às empresas
Qualquer dado que utilize formatos diferentes para representar a mesma coisa	Dados armazenados incorretamente ou erros inseridos durante a transferência de dados	Pontos de dados contraditórios que levam a confusão ou incapacidade de classificar ou segmentar os clientes

IMPACTO COMERCIAL DE DADOS SUJOS

Para uma leitura mais detalhada sobre o impacto comercial de dados sujos, insira o termo "dados sujos" na barra de busca de seu navegador preferido para trazer à tona numerosos artigos sobre o tópico. Aqui estão alguns impactos citados para certas indústrias a partir de uma pesquisa anterior:

- **Banking:** As imprecisões custam às empresas entre 15% e 25% da receita ([fonte](#)).
- **Comércio digital:** Até 25% dos contatos do banco de dados B2B contêm imprecisões ([fonte](#)).
- **Marketing e vendas:** 8 em cada 10 empresas disseram que dados sujos dificultam as campanhas de vendas ([fonte](#)).
- **Assistência médica:** Os registros duplicados podem ser de 10% e até 20% dos registros eletrônicos de saúde de um hospital ([fonte](#)).