

## Como acabar com a duplicidade de dados na sua base. Saiba o que é necessário para identificar e eliminar com sucesso dados duplicados.

Organizações que adotam uma estratégia de qualidade nos dados de contato, precisam estar atentas à duplicidade de dados. O processo de identificação, priorização e junção de dados repetidos pode ser demorado e determinar o curso da ação a ser tomada ao eliminar dados redundantes, pode ser trabalhoso.

---

A qualidade dos dados continua sendo mais importante do que nunca para as organizações. A duplicidade de dados não tem de ser um custo para fazer negócios e as organizações não precisam conviver com problemas de dados e imprecisões. Registros imprecisos de clientes geram análises, relatórios e insights ruins.

O processo de identificação de duplicidades não precisa acontecer de um dia para o outro. É mais importante se concentrar em como melhorar os dados ao longo do tempo. Pode parecer que um banco de dados confuso seja inevitável, mas é importante reconhecer a importância da limpeza de dados e identificar duplicidades, para se obter melhores processos de negócios.

Há muitas considerações a serem feitas para identificar as fontes de dados duplicados. A entrada manual de dados é, muitas vezes, repleta de erros e duplicidades, especialmente quando as informações são inseridas por clientes online. Além disso, a maioria dos sistemas de CRM e ERP têm funcionalidade de busca ruim para localizar e identificar registros existentes. Pesquisar o registro correto pode ser demorado para agentes do call center e outros membros da equipe. No entanto, é importante salientar a necessidade de manutenção de um registro eficiente.

Quando se tem erros na entrada de dados, pode ser muito difícil encontrar o registro certo depois e, conseqüentemente, editá-lo e complementá-lo com novas informações. Isso se torna ainda mais desafiador quando os dados são coletados por meio de diversos canais, como por um representante de vendas por telefone, por um cliente na internet ou em um ponto de venda do varejo. As organizações precisam atualizar constantemente os bancos de dados, uma vez que as informações tornam-se obsoletas e desatualizadas rapidamente.

### Considerações para acabar com a duplicidade de dados.

Este recurso ajudará a identificar como priorizar o processo de limpeza de dados. Os leitores aprenderão abordagens e métodos para identificar duplicidades.

### Análise seus dados.

Ao embarcar em um projeto para identificar duplicidades, é importante não negligenciar os dados já existentes em seu banco de dados. É importante verificar os dados que você já tem para padronização e precisão e determinar a origem das duplicidades existentes.

## Determine prioridades.

Uma vez que você saiba onde os registros duplicados são criados, isso pode ajudar a priorizar o projeto. As perguntas importantes para definir essas prioridades são:

- **São necessários recursos de software ou humanos?**
- **Quais ganhos rápidos podem ser obtidos?**
- **Que projetos exigirão maior orçamento e recursos?**

As organizações geralmente não têm recursos ou verba para desenvolver uma estratégia de identificação de duplicidades. No entanto, existem oportunidades para ganhos rápidos, que exigem investimentos mínimos ou apenas recursos de mão de obra. Limpezas periódicas de dados, antes de campanhas ou iniciativas importantes, podem ser uma forma rentável de obter melhorias na qualidade e fazer ajustes para minimizar a duplicidade.

## Desenvolvendo um processo de identificação de duplicidade.

Ao desenvolver uma estratégia de identificação de duplicidades, é importante determinar os indivíduos que inserem e usam dados. Para realizar projetos de qualidade de dados, é importante atribuir papéis e responsabilidades para as partes interessadas, bem como as prioridades de curto e longo prazo. Por fim, é importante definir metas e métricas para mensurar os indicadores-chave de desempenho e avaliar o sucesso.

## Definindo regras de qualidade de dados.

Implemente regras e padrões de dados, para garantir que novas entradas sejam gerenciadas de forma adequada. Por exemplo, registros e dados incompletos, com a sintaxe incorreta, como letras no campo de número de telefone, devem ser marcados e revisados antes de serem inseridos no sistema. A implementação de regras personalizadas pode ajudar a reconhecer variações comuns e corrigir erros de ortografia, impedindo novas duplicidades.

## Implementando uma padronização.

Padronizar todos os dados no ponto de captura pode ajudar a reduzir duplicidades e padronizando esses dados e garantindo a precisão no ponto de entrada, pode-se conseguir melhores estratégias de qualidade de dados.

## Considere a substituição de termos.

Durante o processo de entrada de dados, a inserção de dados desconhecidos, como cargos, designações ou identificadores de rua com formatação, letras maiúsculas e pontuação indefinidas, dificultam a padronização. Implementar um autocorretor de dados indefinidos pode ajudar a padronizar termos comuns para a uniformidade. Por exemplo, considere um cargo comum como Chief Executive Officer. Pode fazer sentido alterar todas as ocorrências de "C.E.O." e "Ceo" para "CEO" para uma maior consistência do banco de dados.

## Avaliando ferramentas de identificação de duplicidade.

A implementação de uma ferramenta de identificação de duplicidade oferece muitos benefícios. As organizações podem melhorar a eficiência e reduzir o tempo gasto manualmente duplicando contatos. Fica mais fácil criar uma visão única do cliente e garantir a precisão de seus dados. Organizações com muitos bancos de dados diferentes, podem se beneficiar da combinação de dados de diversas fontes em um contato sem erros para ter visões mais precisas dos clientes.

Depois do processo de conhecimento das táticas que podem ser usadas para identificar duplicidades, é importante identificar os fornecedores de software, que ajudam a realizar essas tarefas. Utilizar ferramentas de identificação de duplicidade para pesquisar registros existentes em bancos de dados, incluindo nomes, endereços, datas ou outros dados, como números de telefone ou CPF, pode ajudar a prevenir a criação de dados duplicados e reduzir o tempo gasto com mão de obra avaliando e identificando duplicidades.

Há ofertas disponíveis para ajudar as organizações a identificar duplicidades no ponto de entrada, bem como em lote para fins de limpeza em massa. Soluções de front-end podem ser integradas em um CRM ou em outros aplicativos do call center e usadas quando um funcionário interno estiver inserindo informações em um call center ou online para evitar duplicidades no ponto de entrada.

Identifique ferramentas que permitirão que os usuários pesquisem em campos de dados, como nome, endereço, empresa, data de nascimento e também número de CPF. Em seguida, compare as informações em diversos formatos e em diferentes bancos de dados. Softwares de identificação de duplicidade, que utilizam mecanismos avançados de match ajudam a melhorar os recursos de correspondência e a reduzir os dados duplicados.

Utilize algoritmos fonéticos e de reconhecimento de padrões de palavras, para encontrar duplicidades que uma pessoa não veria. Algoritmos fonéticos e fontes de comparação fornecem padrões de correspondência muito valiosos para sua organização.

## Conclusão.

Se você está embarcando em um projeto para corrigir um problema de duplicidade de dados ou simplesmente para fazer uma limpeza rápida de seus dados, ao garantir que eles estejam padronizados, precisos e verificando duplicidades, ele terá um impacto mensurável na eficiência operacional e na inteligência de negócios.

Manter os dados limpos e sem duplicidades é um processo contínuo, mas que vale a pena. Padronizar e verificar a precisão de registros, ajudará a melhorar a eficiência operacional e reduzirá o tempo gasto da equipe corrigindo informações imprecisas de clientes e possíveis clientes.