

**UNIVERSIDADE FEDERAL DE ITAJUBÁ**

**OSWALDO HENRIQUE DE OLIVEIRA MACHADO**

**MELHORIA DA QUALIDADE DA INFORMAÇÃO DO CADASTRO DOS  
CLIENTES EM UMA EMPRESA NO RAMO DE BEBIDAS**

**ITAJUBÁ**

**2021**

OSWALDO HENRIQUE DE OLIVEIRA MACHADO

MELHORIA DA QUALIDADE DA INFORMAÇÃO DO CADASTRO DOS CLIENTES  
EM UMA EMPRESA NO RAMO DE BEBIDAS

Dissertação apresentada como requisito para a  
obtenção do título de Mestre em  
Administração no Programa de Pós-Graduação  
de Mestrado Profissional em Administração.

Orientador: Prof. Dr. Fabio Favaretto

ITAJUBÁ

2021

## **AGRADECIMENTOS**

À minha família, Jerusa minha esposa e Alice minha filha, meus amores, pelo carinho compreensão e paciência nesta jornada.

Aos meus pais, familiares e amigos, pelo apoio em todos os momentos.

Ao meu orientador, Fabio Favaretto, por acreditar em minha capacidade e orientar-me com paciência e dedicação ao longo desse processo.

Aos meus colegas de trabalho que direta ou indiretamente foram fundamentais para a concretização desta pesquisa em especial Tiago Rossi, Ricardo Andrade e Júlio Cesar.

## RESUMO

A qualidade da informação é um importante pilar para que o trabalho entre empresas e clientes sejam assertivos. O cadastro correto dos clientes é fundamental para que problemas processuais de vendas e logísticos sejam minimizados. Um erro cadastral pode ocasionar desde a não venda de produtos a até problemas a níveis logísticos e fiscais. O estudo tem o objetivo de melhorar a qualidade da informação do processo de alteração de cadastro dos clientes em uma empresa no ramo de bebidas. Tendo em vista a necessidade de realizar uma pesquisa aplicada fez uso de um método focado na melhoria da qualidade da informação (TDQM - Gestão da Qualidade Total dos Dados). A pesquisa propiciará a revisão e mudança dos processos existentes apontando oportunidade de melhorias ao longo do trabalho focada na qualidade da informação. Para maior interação com o tema e conhecer mais a respeito, foram feitas revisão da literatura sobre qualidade da informação e verificados casos de aplicação prática do tema. Como resultado do trabalho, houve a criação de novos processos de coleta e gestão dos dados, criação de métricas e identificação das dimensões da qualidade informação que mais se aderiram ao trabalho (temporalidade, relevância, reputação e qualidade dos dados). O trabalho também apresenta sugestões para melhoria e as aplicou na empresa alvo do estudo. O estudo demonstra que a qualidade a informação da empresa melhora em 59,1% comparando o índice de qualidade da informação antes e após a aplicação do TDQM. Destaca-se também que não existe uma única forma de utilizar as dimensões da informação e que o contexto e a necessidade do momento são decisivos para escolher uma ou outra dimensão.

**Palavras-Chaves:** Qualidade da Informação. Dimensões da qualidade da informação. TDQM.

## ABSTRACT

The information quality is an important pillar for the work between companies and customers to be assertive. The correct registration of customers is essential for procedural sales and logistics problems to be minimized. A registration error can cause non-sale of products to problems at logistical and fiscal levels. The study has the objective to improve the information quality in the process of changing the registration of customers in a beverage industry company. In view of the need to carry out an applied research, made use of a method focused on improving the quality of information (TDQM - Total Data Quality Management). The research will provide the review and change of existing processes pointing to the opportunity for improvements throughout the work focused on the quality of information. For greater interaction with the theme and to know more about it, a literature review on information quality and cases of practical application of the theme were verified. As a result of the work, there was the creation of new processes of data collection and management, creation of metrics and identification of the dimensions of quality information that most adhered to the work (temporality, relevance, reputation and data quality). The work also presents suggestions for improvement and applied them in the target company of the study. The study demonstrates that the quality of the company's information improves by 59.1% comparing the information quality index before and after the application of the TDQM. It is also noteworthy that there is no single way to use the dimensions of information and that the context and the needs of the moment are decisive for choosing one dimension or another.

**Keywords:** Information Quality. Quality Information Dimensions. TDQM.

## LISTA DE FIGURAS

<b>Figura 1</b> - Esquema ilustrativo da metodologia TDQM.....	31
<b>Figura 2</b> – Processo de solicitação de modificação de cadastro.....	41
<b>Figura 3</b> – Processo de solicitação após pesquisa .....	42
<b>Figura 4</b> - Sistema de produção de informação .....	48
<b>Figura 5</b> - Processo de coleta de dados, armazenamento e tratativa das solicitações .....	49
<b>Figura 6</b> - Processo para avaliação e melhoria da qualidade da informação.....	59
<b>Figura 7</b> - Processo para avaliar a etapa definir .....	68
<b>Figura 8</b> - Processos do fluxo de dados desde a coleta até a tratativa e exposição da informação em um painel ( <i>dashboard</i> ). .....	73

## LISTA DE QUADROS

<b>Quadro 1</b> - Produtos versus Informação.....	16
<b>Quadro 2</b> - Categorias e dimensões da qualidade da informação .....	18
<b>Quadro 3</b> - Foco do estudo a respeito da qualidade da informação por autor.....	19
<b>Quadro 4</b> - Característica inerentes da qualidade dos dados ISO/25012 .....	22
<b>Quadro 5</b> - Artigos selecionados .....	25
<b>Quadro 6</b> - Quantidade de dimensões por artigo e dimensões listadas .....	27
<b>Quadro 7</b> - Classificação da pesquisa.....	30
<b>Quadro 8</b> - Etapas do procedimento metodológico adotado .....	32
<b>Quadro 9</b> – Resumo e descrição dos entrevistados .....	44
<b>Quadro 10</b> - Resumo das dimensões selecionadas por campo e pergunta .....	47
<b>Quadro 11</b> - Esquema ilustrativo dos dados coletados e controlados .....	47
<b>Quadro 12</b> - Métricas por campo do formulário e dimensão da informação .....	56
<b>Quadro 13</b> - Tamanho da amostra por campo e dados pontuados .....	60
<b>Quadro 14</b> - Avaliação de percentual da qualidade da informação por campo.....	61
<b>Quadro 15</b> - Percentual de qualidade da informação por dimensão.....	63
<b>Quadro 16</b> - Campos associados a dimensão quantidade de dados apropriados.....	64
<b>Quadro 17</b> – Avaliação qualidade da informação por campo da etapa definir .....	68
<b>Quadro 18</b> – Avaliação das dimensões da qualidade da informação .....	69
<b>Quadro 19</b> – Comparação da qualidade da informação. ....	69
<b>Quadro 20</b> - Comparação antes e depois do TDQM.....	71
<b>Quadro 21</b> – Comparação entre TDQM e etapa definir.....	71
<b>Quadro 22</b> - Resumo das etapas do projeto.....	75

## LISTA DE TABELAS

<b>Tabela 1</b> – Distribuição de trabalho por ano de trabalhos entre 2015 e 2020 .....	24
---	----

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>11</b>
<b>2 REFERENCIAL TEÓRICO .....</b>	<b>14</b>
2.1 QUALIDADE DA INFORMAÇÃO E O USUÁRIO DA INFORMAÇÃO .....	14
2.2 DIMENSÕES DA QUALIDADE DE DADOS .....	17
2.3 APLICAÇÃO DA AVALIAÇÃO DA QUALIDADE DA INFORMAÇÃO .....	23
<b>3 METODOLOGIA.....</b>	<b>30</b>
3.1 GESTÃO DA QUALIDADE TOTAL DOS DADOS (TDQM) .....	30
3.2 PROCEDIMENTO METODOLÓGICO ADOTADO .....	32
3.3 ETAPAS DO PROCEDIMENTO METODOLÓGICO .....	33
<b>3.3.1 Etapa Definir .....</b>	<b>33</b>
<b>3.3.2 Etapa Medir .....</b>	<b>36</b>
<b>3.3.3 Etapa Analisar .....</b>	<b>37</b>
<b>3.3.4 Etapa Melhorar .....</b>	<b>38</b>
<b>4 DESENVOLVIMENTO.....</b>	<b>39</b>
4.1 ETAPA DEFINIR.....	39
4.1.1 Identificação da necessidade de melhoria de um processo.....	40
4.1.2 Mudanças na maneira como os dados eram coletados .....	42
4.1.3 Análise e proposta de mudança frente ao problema.....	43
4.1.4 Definição das dimensões da qualidade da informação .....	44
4.2 ETAPA MEDIR.....	49
4.3 ETAPA ANÁLISE DOS DADOS.....	59
4.3.1 Avaliação dos dados coletado .....	59
4.3.2 Avaliação percentual.....	61
4.3.3 Consolidação das avaliações.....	62
4.3.4 Avaliação geral da qualidade .....	63
4.3.5 Seleção das dimensões com percentual abaixo .....	63
4.4 ETAPA MELHORAR.....	65
<b>5 RESULTADOS .....</b>	<b>67</b>
5.1 COMPARATIVO DA QUALIDADE DA INFORMAÇÃO .....	67
5.2 MITIGAÇÃO DE PROBLEMAS DA QUALIDADE DA INFORMAÇÃO .....	72
5.3 IMPLEMENTAÇÃO DE PROCESSO PARA VISUALIZAÇÃO DA QUALIDADE DA INFORMAÇÃO.....	73
5.4 PROGRAMAÇÃO DE TREINAMENTOS .....	74

5.5 LIÇÕES APRENDIDAS E RESUMO DAS ETAPAS DO TDQM .....	74
<b>6. CONCLUSÕES.....</b>	<b>76</b>
<b>REFERÊNCIAS .....</b>	<b>79</b>
<b>ANEXOS .....</b>	<b>84</b>
ANEXO A - FORMULÁRIO PARA SOLICITAR ALTERAÇÃO DE CADASTRO DE CLIENTES CRIADO NA ETAPA MEDIR.....	84

## 1 INTRODUÇÃO

Para a comercialização de bens de consumo e serviços entre empresas ocorrer é necessário o compartilhamento de dados. Com o aumento da quantidade de dados trocados e o número crescente de transações os sistemas integrados de gestão passaram a ser utilizados para facilitar a integração das várias áreas presentes dentro das empresas e entre as empresas. Gradualmente a qualidade dos dados adquiriu mais e mais importância dentro das empresas sendo que dados que possuem alto grau de qualidade passaram a ser determinantes para o desempenho ideal de análises e eficácia na tomada de decisões. A falta de qualidade de dados, por outro lado, implica em consequências ruins aos clientes internos e externos à organização gerando descrédito e graves consequências negativas nos processos decisórios (WAND; WANG, 1996; MADHIKERMI *et al.*, 2016; FERREIRA; FERREIRA, 2017; HAZEN *et al.*, 2017).

Os vários parâmetros que podem ser identificados em uma operação tornam a aquisição de dados e a análise críticas dos dados elementos desafiadores e essenciais para tomada de decisões do negócio embasadas e com uma boa qualidade de informação (GÜNTHER *et al.*, 2019). O cuidado na coleta e gestão dos dados é essencial por ser tratar de um item crítico e de grande impacto na tomada de decisões principalmente quando se trata de casos individuais tal como o exemplo a seguir: o produto deve ir do ponto A para o ponto B, porém por um erro de cadastro vai para o ponto C. Esse erro de informação além de ocasionar a entrega no local errado poderá impactar negativamente no custo do transporte, tempo de resposta ao pedido, cancelamento do pedido, redução do faturamento dentre outras consequências tais como perda de confiabilidade e participação no mercado. Em uma escala macro um erro poderá virar meramente uma estatística considerando clientes menos importantes para o negócio, porém para clientes chave, esse erro pode causar grandes transtornos.

Constatou-se em uma empresa do ramo de bebidas, que será usada neste estudo, que os processos de modificação dos dados dos clientes não estavam satisfatórios. O atual processo estava moroso, sem controle, sem processos formalizados e não era transparente para o solicitante da alteração dos dados do cliente. As principais dificuldades elencadas eram:

- Tempos de espera para modificação dos dados de clientes estava levando semanas para ser analisado sem dar visibilidade dos motivos dos tempos de resolução dos pedidos;

- Não havia registros formalizados dos pedidos de modificação dos dados o que ocasionava em perda de informação e desordem no processo;
- Durante as tratativas para modificar os dados dos clientes era necessário pedir mais de uma vez informações ou documentos.
- Os solicitantes das modificações de cadastro do cliente não tinham visibilidade dos processos.

As dificuldades citadas pelos colaboradores da empresa seguem um processo que se inicia com a abordagem de um cliente por algum membro da equipe comercial, que em geral é um vendedor. O vendedor faz o cadastro por meio de um sistema de cadastro de clientes via celular. Este cadastro entra em um fluxo de aprovação dentro do sistema da empresa. Caso o cadastro esteja correto o cliente pode comprar.

Caso o cadastro do cliente necessite de qualquer tipo de tratativa como, por exemplo, mudança de endereço, razão social ou qualquer outro tipo modificação de dado que ocorra após o primeiro cadastro é necessário solicitar que o dado seja corrigido.

Quando qualquer um destes exemplos ocorre se fazem necessárias alterações nos dados do cliente. O processo simplificado para solicitação de modificação dos dados do cliente era feito da seguinte forma: o solicitante precisava contatar um analista da equipe de suporte para analisar e tratar os casos que estavam impossibilitando as vendas.

Nem sempre as informações repassadas aos analistas estavam corretas ou faltavam informações para concretização da alteração dos dados dos clientes. A falta de informações, processos bem definidos e comunicação ineficiente tinham como consequência um processo pouco fluído e com baixa qualidade da informação.

Vislumbrando este contexto, foi proposto o seguinte problema de pesquisa: como seria possível reestruturar o processo de modificação dos dados (cadastro) dos clientes garantindo a qualidade da informação?

Baseado nesta questão, este trabalho tem como objetivo geral melhorar a qualidade da informação no processo de solicitação de alteração dos cadastros de clientes na empresa analisada.

Para atingir este objetivo o trabalho é subdividido nos seguintes objetivos específicos:

- Mapear os processos de coleta de dados e fluxo dos dados;
- Padronizar a maneira como os dados são colhidos em campo;
- Reduzir as interações informais;

- Medir a qualidade da informação coletada em campo;
- Utilizar a metodologia TDQM (*Total Data Quality Management*).

Este trabalho utilizará um ambiente prático para a avaliação da qualidade da informação aplicada na alteração de cadastros de clientes. Serão utilizados conceitos teóricos na verificação dos dados coletados em campo bem como será adotado um procedimento metodológico focado na melhoria dos processos que envolvem a informação.

É importante enfatizar que uma coleta de dados bem-feita pode mitigar muitos problemas e representa benefícios consideráveis para a empresa, porém este processo não é garantido devido a erros de digitação e interpretação. Criar um processo que aponte a informação boa e a informação ruim, que previna a entrada de informação ruim no processo, bem como métricas de controle são necessárias para mitigar possíveis entradas de informação com baixa qualidade no sistema. Sendo assim, identificou a necessidade deste trabalho neste ambiente selecionado para esta pesquisa.

Este trabalho divide-se em 5 outras partes: o referencial teórico demonstrará o embasamento sobre os temas quando será caracterizada a qualidade da informação e o usuário da informação, dimensões teóricas da qualidade da informação e casos práticos da avaliação da qualidade da informação. A metodologia da pesquisa descreverá os procedimentos metodológicos. O desenvolvimento demonstrará as etapas que o pesquisador percorreu para propor melhorias dentro do contexto do estudo. Os resultados apresentarão a aplicação das melhorias propostas. A conclusão passará o desfecho da pesquisa e indicará oportunidades de pesquisas futuras.

## 2 REFERENCIAL TEÓRICO

Este tópico tem como objetivo embasar o tema discutido acercando-se dos conceitos pertinentes utilizados para o desenvolvimento do projeto. De uma forma geral, serão abordados os seguintes temas: a qualidade da informação e o usuário da informação, as dimensões da qualidade dos dados e casos de aplicação da avaliação da qualidade da informação.

### 2.1 QUALIDADE DA INFORMAÇÃO E O USUÁRIO DA INFORMAÇÃO

Pesquisas recentes demonstram que a qualidade da informação tem um impacto significativo no conhecimento das organizações e conseqüentemente é um dos pontos cruciais para a tomada de decisão dos gestores (SETIA *et al.*, 2013; KWON; LEE; SHIN, 2014; HEINRICH; KLIER, 2015). Apesar da importância deste tema, a qualidade da informação apresenta definições que podem depender o ponto de vista de quem a utiliza, não possuindo assim, uma definição única. Para Ferreira e Ferreira (2017) a qualidade da informação possui sentidos distintos de acordo com o contexto que está inserido na linha de pesquisa ou indústria. A qualidade da informação, de acordo com Burgess, Gray e Fiddian (2004), não possui somente uma definição, o consumidor da informação é capaz de criar definições personalizadas de acordo com a necessidade.

Alguns autores buscam fazer definições ligadas a necessidades que podem surgir de acordo com um contexto direcionado a satisfação do usuário ou necessidade de uma organização. Isso pode ser verificado quando Shahbazi *et al.* (2019) citam que umas das maiores preocupações das pessoas que desenham sistemas é o usuário final sendo a qualidade da informação um dos maiores desafios suprir a necessidade dos usuários na íntegra, sendo necessário estar alinhado com a necessidade dos indivíduos que utilizam a informação. Para Carretero *et al.* (2017) o tema da qualidade da informação não deve ser tratado de forma isolado, ela deve se adequar ao contexto da organização e desta forma a estratégia direcionada aos dados terá mais chances de sucesso ao suportar as estratégias da organização.

Um dos principais estudos a respeito do tema feito por Wang e Strong (1996) busca definir que a qualidade da informação deve ser adequada ao uso do usuário, sendo o usuário definido como consumidor da informação. Seguindo a abordagem empírica destes autores, eles também buscam elencar efeitos que a qualidade da informação pode causar em uma organização podendo ser benéficos ou não. Segundo Strong, Lee e Wang (1997) uma organização que possui uma baixa qualidade de informações tem um alto potencial de ter problemas nos mais variados âmbitos associados a impactos sociais e econômicos. Segundo

Côrte-Rea, Ruivo e Oliveira (2020) um bom nível de qualidade de dados pode potencializar a eficiência de grandes massas de dados e extrair valor no negócio em sua totalidade se estiverem de acordo com as necessidades do negócio, representando um diferencial competitivo. Caso contrário, podem prejudicar a organização.

Favaretto e Vieira (2007) comentam que a qualidade da informação é valiosa e deve ser tratada como produto ou serviço. Este tema deve ser mensurado através de parâmetros que viabilizem a medição do quanto a informação consegue suprir as necessidades das pessoas que irão fazer uso desta informação. Vidgen, Shaw e Grant (2017) citam que as empresas necessitam ter a capacidade de lidar com grandes quantidades de dados e, por isso, a qualidade dos dados se torna peça fundamental para um processo de tomada de decisão. Logo a qualidade de dados pode ser considerada como um facilitador chave e potencialmente uma barreira para a criação de valor.

Favaretto (2007) cita que qualidade da informação e qualidade de dados podem ser tratados como a mesma coisa. Logo, a qualidade da informação será utilizada também para a qualidade dos dados. Por isso, a qualidade da informação e dos dados da empresa são de extrema importância para todos os processos sendo eles comerciais, produtivos ou administrativos. Ela deve ser gerida como um bem da empresa uma vez que representa um diferencial competitivo. Para Machado (2013), a informação constitui um insumo básico para as atividades laborais, sendo assim, o ponto central da avaliação da qualidade da informação. Oleto (2006) acrescenta que, a informação do ponto de vista de produto é abordada como “coisa” onde se atribuem algumas perspectivas de produtos transformados para um conceito de informação: o quanto ela é abrangente, acessível, atualizada, confiável, objetiva, precisa e válida.

Desta forma, para facilitar as analogias feitas entre produto físico e produto de informação, Wang (1998) introduz o conceito de produto de informação (*Information Product – IP*) quando é feita uma analogia entre um sistema de manufatura e informação onde ele utiliza um processo de entrada, processamento e saída. No âmbito da manufatura, as entradas são as matérias primas, no da informação, os dados sem tratativas. No processamento da matéria prima são os processos que o produto passa até sair como um produto acabado. No caso da informação, os sistemas de informação tomam a frente do processamento e a saída é a informação tratada, ou seja, o produto final, conforme Quadro 1.

**Quadro 1- Produtos versus Informação**

Etapas	Produção de produto	Produção de informação
Entradas	Matéria prima	Dados brutos
Processo	Linha de montagem	Sistema de informação
Saída	Produtos físicos	Produtos da informação

Fonte: Adaptado de Wang (1998, p 59).

A visão da qualidade de dados vista como um produto ou um ativo, determina que as empresas devem ter atenção especial a este tema visto que quanto melhor os dados, melhores serão os benefícios que eles poderão prover para as organizações (FLECKENSTEIN; FELLOWS, 2018; MAHANTI, 2019). Ferreira, Ruivo e Reis (2021), citam que a qualidade de dados boa tem um efeito positivo no desempenho financeiro, organizacional e possui relação direta com o grau de maturidade da empresa. Para Côrte-Real, Ruivo e Oliveira (2020), a qualidade dos dados traz vantagens competitivas para as organizações ao passo que orientam os gestores a tomada de decisão com informações relevante de boa procedência.

Favaretto (2007) menciona que a qualidade é um fator dependente da satisfação do usuário, ou seja, ele deve ter condições de mensurar se a informação está chegando até ele com qualidade ou não. Mattioda e Favaretto (2009) citam que além do usuário a qualidade da informação deve atender também as necessidades da organização. Segundo Kakar (2013), a satisfação do usuário é um indicador chave para medir se o sistema de informação desenvolvido obteve sucesso ou não. Valencia-Parra *et al.* (2021) resumizam que as definições mais utilizadas na literatura estão cunhadas na noção de atender aos requisitos do cliente ou atender a uma adequação ao uso. No entendimento desses autores, a primeira definição relaciona-se ao quão bem os dados estão construídos (exemplo: o número de vezes que os dados estão bons ou ruins). A segunda o quão utilizáveis os dados são, ou seja, relacionam-se ao risco que a organização quer correr quanto a confiabilidade dos dados (exemplo: 90% dos dados estão acurados, isso pode ser (ou não) considerado insuficiente acurado para a empresa “A” não tomar uma decisão).

Janssen, Van Der Voort e Wahyudi (2017) explicam que a qualidade dos dados, o processamento dos dados e como a transferência dos dados é tratada influencia a qualidade da tomada de decisão. Visto este cenário, se faz necessário o uso de critérios para avaliar a qualidade da informação que chega aos usuários. Estes critérios são denominados, dimensões da qualidade da informação e serão explorados no tópico a seguir.

## 2.2 DIMENSÕES DA QUALIDADE DE DADOS

Para avaliar a qualidade dos dados é necessário ter critérios a respeito do tema. De acordo com Shahbazi *et al.* (2019) quando se tem a intenção de lidar com a qualidade de dados é necessário considerar pontos que vão desde erros ortográficos a dados inválidos. Segundo Gürdür, El-Khoury e Nyberg (2019) a qualidade de dados deve vir seguida de sistemas que possibilitem o monitoramento para avaliar a qualidade dos dados, políticas e protocolos que possibilitem verificar os dados bons e ruins. Ainda, segundo o mesmo autor, ela possui várias dimensões. Gualo *et al.* (2021) citam que a avaliação da qualidade dos dados ajuda a garantir sustentabilidade da organização no longo prazo, melhorar e conhecer os aspectos internos do negócio e as formas de trabalho da organização.

Wand e Wang (1996) explicam que não existe um acordo geral em uma definição exata a respeito das dimensões da qualidade sendo um tema com múltiplos conceitos. Mattioda e Favaretto (2006) citam também que não existe rigor relativo as definições das dimensões da qualidade de dados sendo a escolha de determinada dimensões associadas a experiência, revisão da literatura ou percepção sobre o assunto. Oliveira e Favaretto (2021) procuram tratar o tema das dimensões da qualidade da informação como uma dimensão sendo uma divisão (especificidade) para se analisar um todo (mensurar e gerir dados). Dimensões são critérios criados para que os dados sejam avaliados dentro de determinados parâmetros. A respeito das classificações das dimensões pode-se citar que:

A literatura sobre qualidade de dados fornece uma classificação completa das dimensões da qualidade de dados; no entanto, há uma série de discrepâncias na definição da maioria das dimensões devido à natureza contextual da qualidade. As seis classificações mais importantes das dimensões da qualidade são fornecidas por Wand e Wang [1996]; Wang e Strong [1996]; Redman [1996]; Jarke *et al.* [1995]; Bovee *et al.* [2001]; e Naumann [2002]. Ao analisar essas classificações, é possível definir um conjunto básico de dimensões da qualidade dos dados, incluindo precisão, completude, consistência e oportunidade, que constituem o foco da maioria dos autores. No entanto, não existe um acordo geral sobre qual conjunto de dimensões define a qualidade dos dados, ou sobre o significado exato de cada dimensão (BATINI *et al.*, 2009, p.6).

Wand e Wang (1996) fazem referência a quatro dimensões intrínsecas à qualidade de dados sendo que eles devem ser: completos, sem ambiguidade, significativo e correto. No artigo de Wang e Strong (1996) eles enfatizam a necessidade da qualidade de dados possuir uma estrutura compreensível na perspectiva do consumidor da informação. Desta forma, eles formulam quatro aspectos que a qualidade de dados deve possuir: ser acessível (o consumidor sabe de onde capturar os dados), interpretável (é possível de ser lida), relevante (deve ser oportuna para o uso direcionado ao propósito do usuário) e deve ser precisa (vinda de uma

fonte confiável). Por fim, Wang e Strong propõem uma estrutura representada na Quadro 2 com quatro categorias e 15 dimensões resumidas:

**Quadro 2** - Categorias e dimensões da qualidade da informação

<b>Categorias da qualidade</b>	<b>Dimensões da qualidade</b>
Intrínseca	Acurácia, objetividade, reputação e credibilidade.
Acessível	Acessível e segura.
Contextual	Relevante, valor agregado, temporalidade, completude, qualidade apropriada de dados.
Representativa	Interpretável, fácil de entender, representatividade consistente e representatividade concisa.

Fonte: Adaptado de Wang e Strong (1996, p 20).

O modelo conceitual criado por Wang e Strong (1996) foi feito para organizar as dimensões da qualidade dos dados. As dimensões foram definidas por meio de um estudo que considerou as que eram mais importantes para o consumidor da informação e o contexto da atividade segundo pesquisa destes autores. As categorias representam uma definição geral que resumem as dimensões que estão em cada categoria. Desta forma o modelo desenvolvido visou atender inicialmente indústria e áreas governamentais (WANG; STRONG, 1996). Cabe ressaltar que antes de chegar neste modelo conceitual, Wang e Strong (1996) verificaram 179 atributos, chegaram um modelo com 20 dimensões e por fim entenderam que “vinte dimensões seriam muito para uma proposta de avaliação prática” (WANG; STRONG, 1996, p. 16).

Segundo Redman (1998), os principais problemas relacionados à qualidade da informação estão concentrados em quatro dimensões: modo como o mundo capta a informação, valores dos dados, apresentação dos dados e privacidade (segurança). Ele cita que a baixa qualidade dos dados nestas dimensões causa insatisfação dos clientes, aumenta custos, proporciona decisões ineficientes de mercado e reduz a satisfação dos funcionários.

Na visão de Jeusfeld, Quix e Jarke (1998) as dimensões da qualidade devem ser direcionadas ao objetivo das partes interessadas (*stakeholder*) de acordo com a prioridade de cada envolvido. Eles não definem as dimensões da qualidade em grupos, como Wang e Strong (1996), citam que elas existem (temporalidade, qualidade, acessibilidade, consistência, dentre outras) e que cabe a cada *stakeholder* definir as suas dimensões prioritárias.

Naumann e Rolker (2000), sugerem três classes para a verificação da qualidade da informação. São sugeridas classes subjetivas (ponto de vista do usuário e suas experiências, tem relação ao quanto os dados são compreensíveis), classes objetivas (avaliação feito por um

analista da informação relacionado ao quanto completo os dados estão) e classes processuais (determinada pelo processo de consulta relacionado ao tempo). Eles propõem um modelo para fazer esta avaliação e descrevem além dos três critérios mais 22 pontos e o método para fazer as avaliações direcionada ao detentor de cada processo.

Bovee, Srivastava e Mak (2003) propõem um modelo para avaliação da qualidade da informação flexível que possa ser aplicado em várias situações que vão desde áreas de auditoria que necessitam de um documento para comprovação da existência da informação a até a percepção do usuário. Eles propõem um modelo que seja aplicável em qualquer situação e é explicado da seguinte forma:

(...) consiste em três atributos (ou afirmações) que podem ser vistos como extrínsecos às informações: 'Acessibilidade', 'Interpretabilidade' e 'Relevância', e um que pode ser visto como intrínseco: "Integridade." Os atributos extrínsecos são determinados pelas percepções do usuário sobre os atributos de qualidade e o atributo intrínseco. "Integridade", determinada por aspectos inerentes à informação e com critérios que dependem das necessidades do usuário, consiste em quatro subatributos (ou subafirmações): 'Exatidão', 'Completeness', 'Consistência' e 'Existência' (BOVEE, SRIVASTAVA E MAK, 2003, p.13).

O Quadro 03 sintetiza o foco dos autores citados por Batini *et al.* (2009). Ele demonstra o número ou a forma como cada um deles trata as dimensões da informação e cita o número de categorias que cada autor adotou no estudo. Alguns autores tratam categorias com o nome de atributos ou classes. Esta análise tratará todos estes nomes como categorias visto que em todos os casos representam nomenclaturas para agrupar um conjunto de dimensões.

**Quadro 3 - Foco do estudo a respeito da qualidade da informação por autor**

<b>Autores</b>	<b>Ano</b>	<b>Foco do estudo</b>	<b>Número de categoria</b>	<b>Número de dimensões</b>
Wand e Wang; Strong e Wang	1996	Organizar as dimensões da qualidade dos dados.	4	15
Redman	1998	Problemas relacionados a baixa qualidade dos dados.	4	4
Jeusfeld, Quix e Jarke	1998	O objetivo das partes interessadas define as dimensões.	De acordo com necessidade	Partes interessadas definem
Naumann e Rolker	2000	Verificar a qualidade da informação em diferentes perspectivas (classes).	3	22
Bovee, Srivastava e Mak	2003	Descreve um Modelo flexível aplicado a qualquer área do conhecimento	4	4 fixas mais as que o usuário achar necessária

Fonte: Elaborado pelo autor.

As dimensões da qualidade variam de autor para autor, muitas vezes tem semelhanças entre as dimensões e são aplicadas de acordo com a necessidade e contexto que as pesquisas foram feitas. Entretanto, elas possuem em todos os casos, algum tipo de convergência no momento que são avaliadas individualmente cada critério das dimensões da qualidade da informação.

Günther *et al.* (2019) exemplificam que as dimensões mais comuns nas abordagens da avaliação da qualidade são: precisão, integridade, consistência, temporalidade e relevância. Para Scannapieco e Catarci (2002) as dimensões mais recorrentes da maioria dos autores são: acuracidade, integridade, consistência e temporalidade.

Ge e Lewoniewski (2020) comentam que uma dimensão da qualidade é importante em um contexto, mas em outro contexto pode ser ignorado. Gharib e Giorgini (2019) concluem em seus trabalhos que os requerimentos de qualidade não podem ser definidos adequadamente sem considerar o contexto das organizações e dos sistemas existentes. A sensibilidade da avaliação da qualidade dos dados depende dos recursos e das fontes de dados e da dimensão específica da qualidade dos dados: diferentes dimensões podem ser sensíveis de uma maneira diferente (ARDAGNA; 2018). Nenhum autor restringe seu modelo sem termos tais como acessibilidade, integridade ou temporalidade dentre outros que ou são sinônimos ou tem relação próximas entre os mais comuns visto anteriormente. Em uma visão prática, os modelos devem ser adaptados e utilizados de forma a atender a necessidade dos consumidores em sintonia com as necessidades das partes envolvidas.

Valencia-Parra *et al.* (2021) destaca dois modelos genéricos: um proposto por Wang (1998) e outro proposto pela ISO/25012 e 25024. Ainda segundo Valencia-Parra *et al.* (2021) o modelo de Wang (1998) vem sendo o mais utilizado por ser a referência mais confiável na área e auxilia na identificação dos requisitos utilizados em um determinado contexto. As ISO não devem ser consideradas como um modelo alternativo a Wang (1998), elas fornecem indicações importantes para a definição de medidas e métodos de medição para as dimensões ou características da qualidade dos dados.

Wang (1998) propôs uma metodologia intitulada TQDM (*Total Data Quality Management*). “O objetivo desta metodologia TQDM é fornecer produtos de informação (IP) de alta qualidade para consumidores de informação” (WANG, 1998, p.58). O TDQM é um ciclo de melhoria contínua focada em melhorar a qualidade da informação. Este ciclo é constituído de quatro etapas: definir, medir, analisar e melhorar.

Na etapa de definição, Wang (1998) procura entender o contexto geral dos dados que estão analisados, compreender as características do produto de informação os requerimentos e

especificações dos fornecedores, gestores, mantenedores e consumidores da informação. Nesta etapa devem ser definidas as dimensões da qualidade da informação de acordo com a necessidade dos consumidores da informação. O ponto de vista deles deve ser abordado pois será a partir desta visão que será possível compreender o as dimensões que necessitam ser capturadas e reconciliadas. As 15 dimensões definidas por Wang e Strong (1996) são utilizadas como referências no TDQM para avaliação do ambiente estudos e, de acordo com o resultado desta avaliação preliminar, será focada a melhoria nas dimensões que obtiveram pior avaliação e poderá envolver a criação de um sistema no sistema e processos de negócios.

Na etapa de medir, Wang (1998) propõe o desenvolvimento das métricas que serão a bases para avaliar a qualidade da informação. Ao definir as métricas, Wang (1998) sugere que as métricas sejam vinculadas a alguma dimensão da qualidade da informação. A medição constitui uma fase crítica, pois caso não ela seja feita de forma correta não será possível gerenciar a qualidade dos dados de forma eficaz e significativa (GE; LEWONIEWSKI, 2020).

Logo após a definição, Wang (1998) sugere que sejam feitas as análises dos dados coletados, sendo estes avaliados de acordo com o critério das métricas. Estas análises deverão ser os direcionadores para identificar áreas que serão necessárias tomadas de ações para melhoria da qualidade da informação.

A etapa de melhoria deve ser iniciada somente após a conclusão das análises (WANG, 1998). Os executores do TDQM precisarão intensificar as áreas chaves para execução das melhorias. As melhorias podem conter as mais variadas ações. Wang (1998) exemplifica que estas melhorias podem ser direcionadas a uma revisão do fluxo da informação ou realinhar as características da qualidade da informação da organização alvo.

A respeito da avaliação de qualidade da ISO/25012 e 25024, para que ela ocorra, é necessário um ambiente adequado, segundo Caballero *et al.* (2018) são necessário três elementos: (i) um ambiente preparado para a aplicação das ISO, (ii) um processo de avaliação baseada na ISO e (iii) um ambiente semiautomático de avaliação de dados. As duas ISO fornecem definições para medir e indica métodos de medição para as dimensões da qualidade dos dados (VALENCIA-PARRA *et al.*, 2021). A ISO/25012 define 5 características inerentes à qualidade de dados: acurácia, completude, consistência, credibilidade e atualizada. O Quadro 04 resume as características inerentes a qualidade dos dados. Note que o termo característica utilizado na ISO e o termo dimensão utilizado em outros trabalhos são muito próximos. Este trabalho adotará os termos como sinônimos para simplificar a exposição da nomenclatura adotada na ISO.

**Quadro 4 - Características inerentes da qualidade dos dados ISO/25012**

<b>Características da qualidade dos dados</b>	<b>Definições</b>
Acurácia	O grau em que os dados têm atributos que representam corretamente o verdadeiro valor do atributo pretendido de um conceito ou evento em um contexto de uso específico.
Completude	O grau em que os dados do sujeito associados a uma entidade têm valores para todos os atributos esperados e instâncias de entidade relacionadas em um contexto de uso específico.
Consistência	O grau em que os dados têm atributos livres de contradição e são coerentes com outros dados em um contexto de uso específico.
Credibilidade	O grau em que os dados têm atributos considerados verdadeiros e verossímeis pelos usuários em um contexto de uso específico.
Atualizada	O grau em que os dados têm atributos da idade certa em um contexto de uso específico.

Fonte: Adaptado de Valencia-Parra *et al.* (2021, p 2).

A ISO 25024 não explica especificamente a forma como as medidas da qualidade de dados devem ser agrupadas para calcular a qualidade para cada tipo de qualidade de dados. Sobre a ISO é explicado que:

A ISO/IEC 25024 não aborda especificamente como as medidas correspondentes às propriedades de qualidade de dados devem ser agregadas para calcular o nível geral de qualidade para cada característica de qualidade de dados. É uma questão para a organização que realiza a avaliação da qualidade dos dados especificar adequadamente como agregar as medidas para as propriedades de qualidade de dados, a fim de obter o valor do nível de qualidade para as características de qualidade de dados, bem como para estabelecer o processo de avaliação de qualidade de dados específico (GUALO *et al.*, 2021, p. 3).

Os métodos de medição da qualidade da informação serão baseados na identificação das características da qualidade da informação que mais se adequem a necessidade da organização. Apesar disso os laboratórios certificadores da ISO possuem um processo rigoroso e sistemático para avaliação da qualidade, são 5 processos: (i) estabelecer os requerimentos para avaliar a qualidade dos dados, (ii) especificar a avaliação da qualidade dos dados, (iii) desenhar o escopo da avaliação dos dados, (iv) executar a avaliação e (v) concluir a avaliação (MERINO; 2017, GUALO *et al.*;2021).

Um bom exemplo da adaptabilidade da aplicação das dimensões nos mais diversos setores pode ser verificado no trabalho de Oliveira e Favaretto (2021) quando eles selecionam as dimensões reputação, acessibilidade e valor agregado após entrevistar vários participantes de um estudo focado na área médica. O objetivo foi melhorar a qualidade da informação dos prontuários eletrônicos dos pacientes em um hospital. É importante citar que neste estudo foi utilizado um processo adaptado, os *Ten Steps to Quality Data and Trusted Information*

(MCGILVRAY, 2008) para avaliar quais as dimensões teriam mais aderência com o ambiente que a pesquisa se enquadrava.

No tópico a seguir serão abordados mais estudos que têm um viés prático de aplicação da avaliação da qualidade de dados (*data assessment*) e/ou avaliação da qualidade da informação (*quality assessment*) com o objetivo de exemplificar casos de aplicação real.

### 2.3 APLICAÇÃO DA AVALIAÇÃO DA QUALIDADE DA INFORMAÇÃO

Nesta sessão foram selecionados artigos que apresentaram viés de aplicação prática das dimensões da qualidade da informação objetivando demonstrar como outros trabalhos utilizaram as dimensões da qualidade da informação para o desenvolvimento da avaliação (*assessment*). Mattioda *et al.* (2015) elencam três questões para identificar artigos que tenham relevância e possam servir de suporte para revisões literárias. (i) Quais as pesquisas recentes da área? (ii): Quais são as referências científicas para o estudo? (iii): Existe referências científicas nos estudos relevantes? Este raciocínio metodológico será adaptado (simplificado) neste trabalho para criar um funil com o objetivo de selecionar artigos relevantes para a este trabalho.

A pesquisa bibliográfica foi feita por meio do portal Periódicos CAPES na coleção principal da base de dados *Web of Science* (WOS) e *ScienceDirect*. Os termos usados para a seleção dos estudos foram: *Data\_Qualit*, *Dimension* e *Assessment*, no campo título, para obtenção de uma busca mais objetiva direcionada ao problema. A linha temporal considerada foi entre 2015 e 2020. Foram utilizados filtros com as áreas que apresentam maior possibilidade de similaridade com trabalho em questão. As áreas selecionadas para o estudo foram: informática aplicada a sistemas de informação, automação de sistemas, engenharia de produção, ciência da computação, métodos de engenharia multidisciplinar, ciência da computação interdisciplinar, inteligência artificial, negócios, informática cibernética, aplicações de matemática interdisciplinares, engenharia industrial, de telecomunicações, informática engenharia de software, transporte, tecnologia da ciência do transporte, ciência da informação biblioteca, finanças de negócios, ciência da arquitetura computacional e estatística e probabilidade de hardware.

Além dos filtros selecionados, os trabalhos devem apresentar as seguintes características para integrarem a seleção:

- devem identificar algumas dimensões da informação citados na literatura;
- ser qualitativos;
- demonstraram aplicação prática;

- tratar empiricamente as dimensões da qualidade da informação.

Foram identificados 95 trabalhos de pesquisa e lidos os resumos de todos os trabalhos. Dos 95, 46 apresentaram relação muito baixa com o problema de pesquisa. Desta forma foram lidos na íntegra 49 artigos. Dos 49 artigos lidos na íntegra foram selecionados 8 e excluídos 41. A Tabela 1 demonstra o um resumo da distribuição dos artigos por ano dentro do período estipulado neste estudo.

**Tabela 1** – Distribuição de trabalho por ano de trabalhos entre 2015 e 2020

Ano	Total	Percentual	Não Selecionados	Selecionados
2020	14	15%	12	2
2019	20	21%	16	4
2018	18	19%	17	1
2017	16	17%	16	0
2016	12	13%	11	1
2015	15	16%	15	0
<b>Total</b>	<b>95</b>	<b>100%</b>	<b>87</b>	<b>8</b>

Fonte: Elaborado pelo autor.

A pesquisa demonstrou uma concentração de trabalhos maior no ano de 2019 de acordo com os parâmetros avaliados, 21% do número total seguido por 2018 com 19% e 17% para 2017. Dentre os 8 trabalhos selecionados 4 estão no ano de 2019 e 2 no 2020 sendo estes os anos com maior concentração de artigos.

Os principais motivos das exclusões de 41 artigos foram: (i) 19 tratavam de ensaios teóricos com proposição de modelos não aplicados de fato. Os trabalhos enfatizavam questões relativas a como avaliar a qualidade e quais dimensões da qualidade seriam mais adequadas a determinadas áreas, tais como, indústria de manufatura, área da saúde e sistema de informação. Não havia uma aplicação de fato em um ambiente prático. (ii) 9 não faziam referências às dimensões da qualidade da informação e/ou eram pobres em referencial. Estes trabalhos tinham um rigor acadêmico pobre ao referenciar erroneamente trabalhos de um autor em vez de outro ou não referenciavam o que eram as dimensões e partiam direto para partes práticas com baixo embasamento teórico. (iii) 13 não apresentavam um processo definido para a construção do modelo de avaliação da qualidade da informação. Estes trabalhos aplicaram em um ambiente prático as dimensões, mas não descreveram o caminho para chegar a um modelo, ou seja, os trabalhos não demonstraram claramente as etapas que forma necessárias desenvolver para criação do modelo aplicado em um ambiente prático. O Quadro 05 ilustra os selecionados ao final da análise.

**Quadro 5 - Artigos selecionados**

Referência	Área de pesquisa	Número de citações
REIMER, Andrew P.; MILINOVICH, Alex; MADIGAN, Elizabeth A. Data quality assessment framework to assess electronic medical record data for use in research. <i>International journal of medical informatics</i> , v. 90, p. 40-47, 2016.	Computer Science; Health Care Sciences & Services; Medical Informatics.	20
ARDAGNA, Danilo <i>et al.</i> Context-aware data quality assessment for big data. <i>Future Generation Computer Systems</i> , v. 89, p. 548-562, 2018.	Computer Science.	34
GÜNTHER, Lisa C. <i>et al.</i> Data quality assessment for improved decision-making: a methodology for small and medium-sized enterprises. <i>Procedia Manufacturing</i> , v. 29, p. 583-591, 2019.	Engineering; Metallurgy & Metallurgical Engineering.	36
GÜRDÜR, Didem; EL-KHOURY, Jad; NYBERG, Mattias. Methodology for linked enterprise data quality assessment through information visualizations. <i>Journal of Industrial Information Integration</i> , v. 15, p. 191-200, 2019.	Computer Science; Engineering.	28
SÁNCHEZ, Roberto Álvarez <i>et al.</i> TAQIH, a tool for tabular data quality assessment and improvement in the context of health data. <i>Computer methods and programs in biomedicine</i> , v. 181, p. 104824, 2019.	Computer Science; Engineering; Medical Informatics.	27
SHAHBAZI, Mehri <i>et al.</i> Development of a scale for data quality assessment in automated library systems. <i>Library &amp; Information Science Research</i> , v. 41, n. 1, p. 78-84, 2019.	Information Science & Library Science.	48
ANDREWS, Robert <i>et al.</i> Quality-informed semi-automated event log generation for process mining. <i>Decision Support Systems</i> , v. 132, p. 113265, 2020.	Computer Science; Operations Research & Management Science.	64
CÔRTE-REAL, Nadine; RUIVO, Pedro; OLIVEIRA, Tiago. Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value?. <i>Information &amp; Management</i> , v. 57, n. 1, p. 103141, 2020.	Computer Science; Information Science & Library Science; Business & Economics.	114

Fonte: Elaborado pelo autor.

Todos os estudos analisados descritos no Quadro 2 possuem aplicação de alguma dimensão na qualidade da informação no trabalho de pesquisa. Os trabalhos tinham como proposta construir um modelo (*framework*) e que já apresentavam algum tipo de citação de dimensões teóricas que seriam aplicadas nos modelos práticos. A seguir serão apresentados uma síntese sobre o contexto geral de cada estudo, quais dimensões da qualidade da informação cada trabalho utilizou na aplicação prática da avaliação da qualidade da informação e os autores que foram citados para referenciar as dimensões.

Reimer, Milinovich e Madigan (2016) criaram um modelo para avaliação da qualidade da informação de registros médicos eletrônicos baseado em 6 etapas: análise preliminar, documentação, amplitude (número e tipo dados por paciente), presença de elementos de dados (presença de variáveis), densidade (frequência de registros para cada paciente ao longo do tempo) e predição para avaliar a qualidade dos dados. Este modelo foi

aplicado em uma base de registros hospitalar. No que abrange as dimensões da qualidade da informação os autores optaram por utilizar 5: integridade, exatidão, concordância, plausibilidade e temporalidade (WENG; WEISKOPF, 2013).

Ardagna *et al.* (2018) fizeram um estudo prático a partir de dados coletados proveniente de uma cidade inteligente. Foi proposto um modelo adaptativo para gerenciar a avaliação da qualidade dos dados a ajudar o usuário a escolher melhores parâmetros para executar avaliações de qualidade de acordo com os objetivos principais do modelo: minimização de tempo, orçamento e melhorar a confiança. Eles dividiram o estudo em 3 etapas: coleta da amostra dos dados, modelo de regressão e exploração de modelo orientada para o custo. O modelo considera 7 dimensões avaliar: acuracidade, completude, consistência, distinção, precisão, temporalidade e volume (WANG; STRONG, 1996; BOVEE; SRIVASTAV; MAK, 2003; BATINI *et al.*, 2016).

Sánchez *et al.* (2019) propuseram um método denominado TAQIH para avaliação e melhoria da qualidade da informação e apoiar usuários não técnicos da informação para explorar os dados com foco na área da saúde. São consideradas 4 dimensões: completude, precisão, redundância e legibilidade (SCANNAPIECO; BATINI, 2016)

Shahbazi *et al.* (2019) fizeram seus estudos baseados em sistemas automatizados de bibliotecas (ASL). Eles desenvolveram uma escala para avaliar a qualidade dos dados sistemas automatizados de livrarias com 62 itens de qualidade de dados e 4 fatores: qualidade do conteúdo de dados, qualidade organizacional de dados, qualidade de apresentação de dados e qualidade de uso de dados baseada em dimensões da qualidade. Os quatro fatores foram sintetizados em 4 dimensões: utilidade, consistência, precisão e acessibilidade (STRONG; LEE; WANG, 1997).

Günther *et al.* (2019) fizeram um estudo em uma fábrica do ramo de chapas de metal. Eles propuseram uma metodologia para simplificar a avaliação da qualidade da informação e entender melhor os resultados com foco em indicadores de desempenho logístico. O trabalho utilizou 6 dimensões: completude, relevância, precisão, temporalidade, consistência e plausibilidade (WANG; STRONG, 1996; BATINI *et al.*, 2009; PRICE; SHANKS, 2016).

Gürdür, El-Khoury e Nyberg (2019) desenvolveram um modelo para avaliação e gestão da qualidade dos dados no ramo automotivo. O modelo possui em 6 etapas: (planejamento, extração, identificação, consulta, visualização e melhoria). Como resultado do estudo foi desenvolvido um painel que possibilitou a empresa gerir de forma visual a qualidade dos dados. No total forma 5 dimensões: completude, usabilidade, rastreabilidade, consistência e validada (NETO *et al.* 2016; ZAVERI *et al.* 2016).

Andrews *et al.* (2020), desenvolveram uma abordagem semiautomática para avaliar a qualidade dos dados por meio da geração de *logs* de eventos de um processo de extração e dados (*data mining*) em uma empresa de eletrônicos. Neste trabalho foram consideradas 3 dimensões relevantes para o estudo: precisão, completude e singular (FOX; LEVITIN; REDMAN, 1994; WANG; STRONG, 1996; BATINI *et al.* 2009).

Côrte-Real, Ruivo e Oliveira (2020), desenvolveram um modelo que avaliou a qualidade da informação em mais de 600 empresas. O modelo explica a importância da qualidade dos dados para criar alicerce para que o *Big Data* e a *Internet of Things* se tornem diferenciais competitivos. Foram 4 dimensões utilizadas para avaliar o modelo: completude, acuracidade, formato e moeda (STRONG; LEE; WANG, 1997; CAI, L; ZHU, 2015). O Quadro 6 faz uma síntese de cada autor e elenca a quantidade e as dimensões que cada um utilizou.

**Quadro 6** - Quantidade de dimensões por artigo e dimensões listadas

<b>Autor(es)</b>	<b>Quantidade</b>	<b>Dimensões</b>
Reimer, Milinovich e Madigan (2016)	5	Integridade, exatidão, concordância, plausibilidade e temporalidade.
Ardagna <i>et al.</i> (2018)	7	Acuracidade, completude, consistência, distinção, precisão, temporalidade e volume.
Sánchez <i>et al.</i> (2019)	4	Completude, precisão, redundância e legibilidade.
Shahbazi <i>et al.</i> (2019)	4	Utilidade, consistência, precisão e acessibilidade.
Günther <i>et al.</i> (2019)	6	Completude, relevância, precisão, temporalidade, consistência e plausibilidade.
Gürdür, El-Khoury e Nyberg (2019)	5	Completude, usabilidade, rastreabilidade, consistência e validade.
Andrews <i>et al.</i> (2020)	3	Precisão, completude e singular.
Côrte-Real, Ruivo e Oliveira (2020)	4	Completude, acuracidade, formato e moeda.

Fonte: Elaborado pelo autor.

Dos 8 trabalhos pesquisados o número máximo de dimensões utilizadas foi 7 e mínimo de 3 sendo que a média que dimensões por trabalho foi de 4,75 dimensões e a mediana 4. As dimensões que obtiveram maior frequência de citações foram completude (6 trabalhos), precisão (5 trabalhos e consistência (4 trabalhos). Cabe ressaltar que as dimensões que cada autor elenca foram traduzidas do inglês para o português, desta forma podem haver desvios quanto ao significado da palavra e ao conceito central de cada dimensões, podendo haver semelhanças entre dimensões com nomes distintos.

Embora não se tenha um padrão ou norma para utilizar as dimensões da informação (BATINI *et al.* 2009) coube a cada autor utilizar a que mais se adequou ao trabalho. Reimer,

Milinovich e Madigan (2016) justificaram o uso das dimensões por serem frequentemente utilizadas na área da saúde. Ardagna *et al.* (2018) selecionaram 7 dimensões contidas na literatura definidas previamente. O modelo desenvolvido no trabalho detecta o tipo de dados e avalia o dado de acordo com a dimensão que mais se adequa. Sánchez *et al.* (2019) justifica o uso por serem as dimensões mais utilizadas na área de EHR (armazenagem de dados eletrônicos de saúde).

Shahbazi *et al.* (2019) colheram opiniões de especialista da área de ALSs (sistemas automatizados de livraria) sobre as dimensões que eles consideravam mais importante, depois identificaram as mais importantes de acordo com os usuários finais e por fim selecionaram as dimensões com o auxílio de professores da área. Günther *et al.* (2019) estudaram as dimensões mais frequentes na literatura e fizeram uma avaliação preliminar dos dados coletados na pesquisa e definiram quais dimensões da qualidade da informação mais se adequaram a cada parcela de dados avaliados. Gürdür, El-Khoury e Nyberg (2019) justificaram a escolha das dimensões por meio do alinhamento com as partes interessadas em avaliar a qualidade da informação na empresa.

Este alinhamento consistiu em explicar os conceitos e importância da qualidade da informação incluído detalhes técnicos e acordos para entrega do modelo final abrangendo as dimensões da qualidade da informação selecionadas. Andrews *et al.* (2020), utilizaram as dimensões mais importantes para o processo de mineração de dados segundo a necessidade do projeto. A avaliação foi feita pelos próprios autores do trabalho baseadas na literatura que pesquisada. Côte-Real, Ruivo e Oliveira (2020), de acordo com a pesquisa teórica que os autores fizeram sobre dimensões da qualidade da informação, eles definiram as mais podem influenciar a *Big data* e *Internte of Things*.

É possível identificar também que as referências mais citadas pelos autores para referenciar e justificar o uso das dimensões estão concentradas em nos trabalhos de Wang e Strong e Batini. Os primeiros são citados em 5 trabalhos, o que representa 62,5% dos trabalhos (Andrews *et al.*, 2020; Ardagna *et al.*, 2018; CÔRTE-REAL; RUIVO; OLIVEIRA, 2020, GÜNTHER *et al.*, 2019; SHAHBAZI *et al.*, 2019). O segundo aparece em 4 trabalho, ou seja 50% dos trabalhos analisados (ANDREWS *et al.*, 2020, ARDAGNA *et al.* 2018; GÜNTHER *et al.*, 2019; SÁNCHEZ *et al.*, 2019).

Estes estudos práticos serviram de auxílio para compreender como outros autores selecionaram as dimensões, quais foram os processos utilizados para criar os modelos para avaliar a qualidade da informação e para identificar as literaturas mais recorrente. Foi possível também compreender que não existe uma única forma de utilizar as dimensões da informação

e que o contexto e a necessidade do momento são decisivos para escolher uma ou outra dimensão. Outra contribuição que esta pesquisa trouxe ao trabalho refere-se aos múltiplos caminhos que foram explorados pelos autores para definir as dimensões e criar modelos que apesar de distintos possuem um fim em comum: avaliar a qualidade da informação de uma determinada base dados. A falta de uniformidade justifica-se uma vez cada autor trabalhou em áreas distintas na maioria dos casos e adequou a quantidade de dimensões, os processos e o modelo de avaliação de acordo com a necessidade do momento.

### 3 METODOLOGIA

A pesquisa em questão classifica-se como pesquisa qualitativa, de caráter exploratória, de natureza aplicada. De acordo com Deslauriers (1991) a pesquisa qualitativa objetiva produzir novo conteúdo aprofundado sobre um determinado assunto não importando o tamanho da amostra e busca proporcionar ao pesquisador condições de preparar declarações válidas.

Segundo Gil *et al.* (2007) a pesquisa exploratória busca oferecer maior intimidade com o problema objetivando torná-lo mais claro. Este trabalho classifica-se como pesquisa aplicada, ao passo que “procura produzir conhecimentos para aplicação prática dirigidos à solução de problemas específicos” (PRADANOV; FREITAS, 2013, p. 126). O Quadro 7 sintetiza os pontos elencados anteriormente a respeito da metodologia.

**Quadro 7 - Classificação da pesquisa**

Ponto de vista	Caracterização
Abordagem do problema	Qualitativa
Objetivos	Exploratória
Natureza	Pesquisa Aplicada

Fonte: adaptado SILVA e MENEZES (2001).

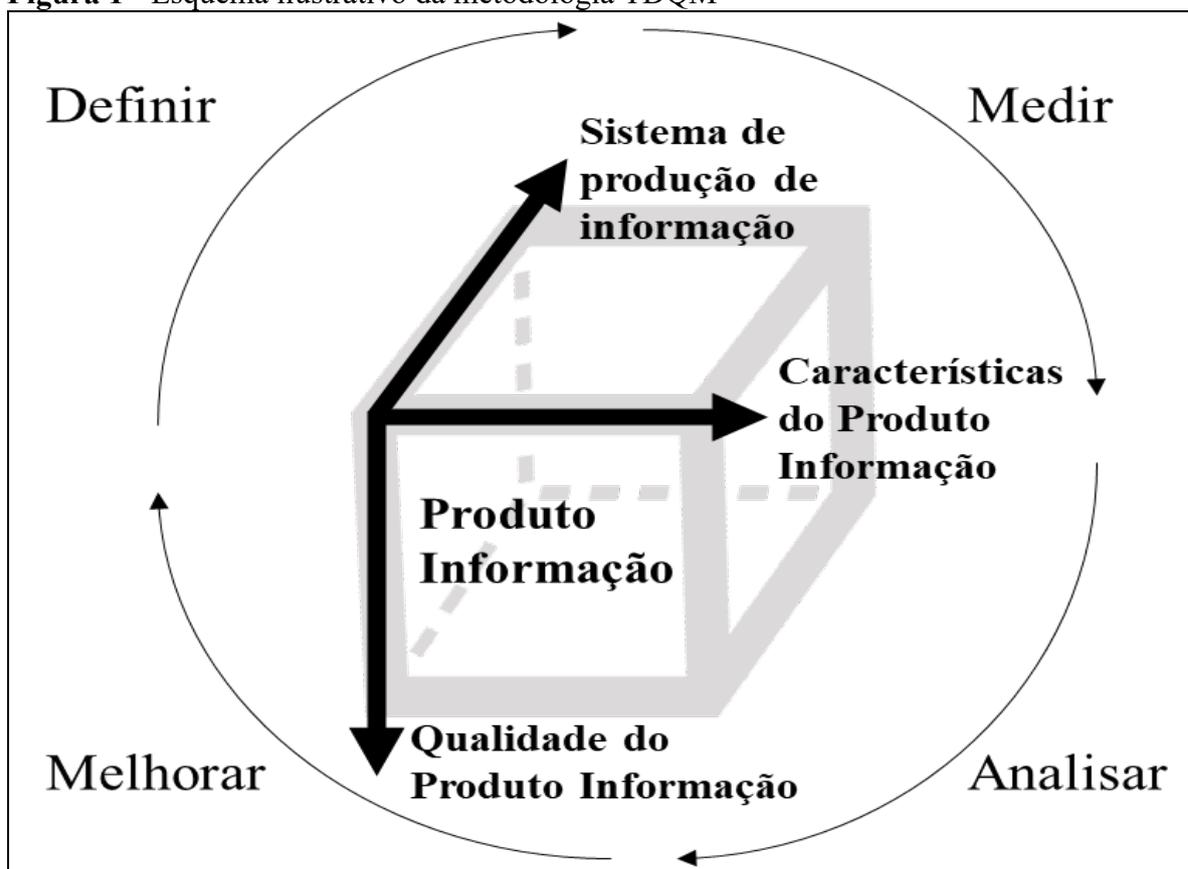
Como procedimento técnico a pesquisa adotará o TDQM (*Total Data Quality Managment*). Esta metodologia tem foco na abordagem de problemas relacionados a qualidade da informação por meio de 4 etapas (definir, medir, analisar e melhorar). Segundo Wang (1998), “a metodologia proposta TDQM é baseada em pesquisas acumuladas e extensas experiências práticas”.

No Tópico 3.1 serão descritas as metodologias que foram utilizadas neste trabalho e o Tópico 3.2 destinará a expor o procedimento metodológico que demonstrará claramente as sequências e etapas da condução da pesquisa.

#### 3.1 GESTÃO DA QUALIDADE TOTAL DOS DADOS (TDQM)

O TDQM (*Total Data Quality Management*) proposto por Wang (1998) é um método que trata a informação como produto tendo como objetivo proporcionar produtos de qualidade aos usuários e melhorar o produto de informação. A informação deve ser adequada para o uso do usuário da informação, ou consumidor da informação (*information consumer*). Esta metodologia é dividida em quatro etapas, conforme Figura 1.

**Figura 1** - Esquema ilustrativo da metodologia TDQM



Fonte: Adaptado de Wang (1998, p 60).

A Figura 1 proporciona uma visão resumida das etapas do desenvolvimento do TDQM. Cada um dos passos (definir, medir, analisar e melhorar) possui um detalhamento para que o ciclo não seja entendido de maneira superficial. A seguir serão detalhadas as etapas do ciclo segundo Wang (1998).

**Definição:** identifica as dimensões da qualidade da informação e os requerimentos correspondentes ao que se procura melhorar. Em um nível mais amplo deve-se compreender quais as principais funcionalidades necessárias para os consumidores da informação (a concepção do que é importante deve ser aderente a necessidade do consumidor da informação) relacionadas ao produto de informação. Em um nível mais restrito quais são as informações básicas que o usuário necessita para utilizar o produto de informação. Após a compreensão das características do produto de informação, é necessário compreender como ocorre a produção desta informação e as interações entre os envolvidos no processo. Wang (1998) cita que são quatro envolvidos no processo do produto de informação: fornecedores (coletores da informação), fabricantes (desenham o sistema que os dados serão armazenados), consumidores (são os usuários dos produtos informação) e gerentes (gerenciam os processos, produção e ciclo de vida do produto de informação). Wang (1998, p. 63) explica que de uma

forma geral a etapa de definição deve entregar: “(...) (1) um modelo de relacionamento de entidade de qualidade que define o IP (produto de informação) e seus requisitos de QI, e (2) um sistema de manufatura de informações que descreve como o IP é produzido e as interações entre os fornecedores de informação (fornecedores), fabricantes, consumidores e gerentes de IP.”

**Medição:** métricas correspondentes aos requerimentos da qualidade da informação. Estas métricas podem ser as dimensões mais importantes para o usuário do produto de informação e podem envolver regras de negócios específicas da organização. Elas devem ser implementadas nas rotinas de um sistema de informação.

**Análise:** identifica a causa raiz do problema e evidencia os impactos do baixo nível da qualidade da informação. Esta etapa tem como base os resultados das métricas da etapa anterior para que a equipe do produto de informação tenha bases para investigar possíveis causas de problemas da qualidade da informação. Segundo Wang (1998) esta etapa pode usar métodos de análise simples ou complexos abrangendo desde reconhecimento de padrões, pareto e controle estatístico de processos.

**Melhoria:** indica caminhos e técnicas para prover melhorias no processo analisado. Ela deve ser realizada após a etapa de análise. Nela devem ser indicadas as áreas que e os processos que devem ser feitas melhorias.

### 3.2 PROCEDIMENTO METODOLÓGICO ADOTADO

Este tópico destina demonstrar a sequência das etapas adotadas para condução da pesquisa. Este procedimento será baseado no TDQM. O Quadro 8 ilustra a sequência das etapas do procedimento adotado:

**Quadro 8** - Etapas do procedimento metodológico adotado

TDQM			
Definir	Medir	Analisar	Melhorar

Fonte: Elaborado pelo autor.

As etapas do TDQM serão utilizadas para identificar as funções do produto de informação. O projeto de estudo busca melhorar a qualidade da informação no processo de solicitação de alteração dos cadastros de clientes na empresa analisada. Este processo inicia-se a partir da identificação de um erro cadastral. Este erro é reportado e deve ser tratado para que o cadastro seja corrigido e a empresa tenha possibilidade de executar transações comerciais

com clientes de uma maneira adequada. A seguir serão detalhadas as etapas do procedimento metodológico de acordo com o método TDQM proposto por Wang (1998).

### 3.3 ETAPAS DO PROCEDIMENTO METODOLÓGICO

Os tópicos a seguir irão demonstrar os passos que a pesquisa seguirá para o desenvolvimento do trabalho. Serão explicadas as etapas do procedimento metodológico utilizado no trabalho.

#### 3.3.1 Etapa Definir

Wang (1998) define que nesta etapa devem ser abordadas as características, requerimentos e o sistema de manufatura da informação. Este projeto tem como principal objetivo melhorar a qualidade da informação do processo de solicitação de alteração de cadastro de clientes. Nesta etapa serão feitas as seguintes atividades:

- Identificação da necessidade de melhoria de um processo;
- Mudanças na maneira como os dados serão coletados;
- Análise e proposta de mudança frente ao problema;
- Definição das dimensões da qualidade da informação.

A partir dos problemas relatados pela equipe que solicita a modificação dos dados dos clientes, a equipe de suporte ao time comercial fará a identificação dos problemas do processo que causam a baixa qualidade da informação.

Após esta etapa serão identificados os dados que necessitam ser coletados. Será construída uma tabela onde cada coluna terá como cabeçalho a necessidade dos dados e cada linha será o conjunto dos dados coletados. Os solicitantes das mudanças de cadastro dos clientes serão os membros da equipe comercial que é constituída por vendedores e gestores comerciais (supervisores e gerentes).

Os gestores dos dados serão o time de suporte comercial (assistentes, analistas e especialistas). Eles serão responsáveis por inserir na planilha os dados repassados pelo time comercial e dar prosseguimento as tratativas de modificação do cadastro dos clientes. Eles serão responsáveis também por indicar ao time comercial a necessidade de documentos extras, tais como comprovantes de endereço, cadastros fiscais e fotos do local. À medida que os dados forem coletados a equipe de suporte comercial se encarregará de analisar e apontar oportunidades de melhoria para o processo de coleta de dados e melhoria da qualidade da informação.

Serão feitas as seguintes atividades para auxiliar a execução deste trabalho: (i) listar os motivos mais frequentes de solicitação de mudança dos dados dos clientes, (ii) identificar outros dados que são repassados com frequência pelo time comercial e que o formulário não contempla, (iii) remover campos que tem uma frequência baixa de solicitação, (iv) padronizar os motivos de solicitação, v) propor melhorias para o processo de coleta e acompanhamento das solicitações e (vi) planejar próximas ações. A coleta de dados e resultados obtidos nesta etapa serão utilizados para dar suporte para iniciar a próxima etapa da pesquisa e servirá de base para comparar os processos antes e depois do TDQM. Esta etapa auxiliará na definição das características do produto da informação tal como o consumidor da informação necessitará consumi-la em termos de entender a funcionalidade para o usuário, quais serão as unidades básicas de informação necessárias para o funcionamento adequado do banco de dados e quais serão as maiores necessidades do usuário da informação. O sistema deve ser funcional para o usuário.

As principais ideias para aplicação do processo se concentrarão em um sistema onde seja possível que a equipe comercial (coletores dos dados) solicite por meio de um sistema on-line a alteração do cadastro dos clientes. Além disso, o processo deverá proporcionar um feedback para o solicitante, registro da abertura da solicitação sem necessidade de intervenção humana para inserção de data da solicitação e atenda às necessidades de informação do usuário da informação coletada.

Será necessário padronizar os tipos de solicitações possíveis de serem feitas por meio do sistema, mapear as regras de negócios e inserir no sistema as evidências necessárias para evitar retrabalho de solicitar novamente informações a equipe comercial.

Wang (1998) sugere definir os requerimentos da qualidade da informação de acordo com as perspectivas do produto de informação dos fornecedores de informação, dos fabricantes de informação, dos consumidores e dos gestores do produto de informação. Serão avaliadas e selecionadas dimensões da qualidade da informação de acordo com a necessidade do consumidor da informação e do negócio.

De acordo com a metodologia TDQM, são 15 dimensões: acuracidade, objetividade, credibilidade, reputação, acessibilidade, segurança, relevante, valor agregado, temporal, completa, quantidade de dados, interpretável, fácil de entender, representação concisa e representação consistente (WANG, 1998). Para selecionar as que mais se adequarão será utilizada a necessidade do usuário como parâmetro para escolha, as que mais auxiliarão na resolução do problema de pesquisa e as necessidades do negócio da empresa alvo.

O pesquisador, juntamente com o usuário da informação e os gestores da empresa serão responsáveis por fazer a leitura deste ambiente. O pesquisador fará entrevistas por meio de um roteiro, que será detalhado no Tópico 4, com alguns membros da equipe comercial a acompanhará o dia a dia deles. Após coletar as informações, o pesquisador determinará a quantidade a quais dimensões se adequarão melhor ao contexto da pesquisa baseada nas respostas dos entrevistados (WANG, 1998). Ao longo do desenvolvimento deste trabalho serão selecionadas e justificadas o porquê do uso de determinada dimensão que mais se adequaram ao trabalho por meio do uso da metodologia sugerida no TDQM, esta etapa será melhor detalhada no Tópico 4.

Os dados coletados por meio do formulário que será mais bem abordado no Tópico 4.2 serão utilizados para abordar a avaliação da qualidade da informação bem como o novo processo. De acordo com o processo de desenvolvimento do TDQM de Wang (1998), será feito um modelo de relacionamento de entidade de qualidade para definir o produto de informação e seus requisitos (dimensões) da qualidade da informação. Será proposto um sistema de manufatura de informações que descreveu o sistema produtivo do produto de informação e as interações dos envolvidos no produto de informação que, no caso deste estudo foram: a equipe comercial e a equipe de suporte comercial. A partir dos resultados obtidos será desenvolvido um sistema para armazenamento, tratativa e gestão destes dados.

Será alvo do projeto um centro de distribuição de uma empresa do ramo de bebidas, localizado na cidade de Poços de Caldas, Minas Gerais. Os principais pontos para a seleção desta unidade foram: identificação problemas de qualidade da informação no processo de abertura e solicitação de modificação e dados dos clientes, baixa complexidade para fazer um projeto piloto de melhoria da qualidade da informação e unidade autorizada fazer o projeto exploratório

A respeito da técnica utilizada para coleta de dados, será feita coleta de dados documental sendo que serão utilizados os dados coletados nesta etapa (definir) para estudar o registro das solicitações coletadas em planilhas. De acordo com as conclusões retiradas do estudo dos registros feitos nas planilhas será construído um formulário que resultará em um banco de dados à medida que o time comercial (coletores dos dados) preencherem os formulários com solicitações de alteração de cadastros dos clientes.

A situação em que se encontra o processo de solicitação de alteração de cadastro de clientes está desorganizado, gera queixas da equipe comercial, problemas de comunicação e impossibilita fazer um acompanhamento das solicitações de alteração do cadastro dos clientes feitas pela equipe comercial. Diante deste cenário, será feito um estudo para entender melhor

o problema, coletar dados, compreender melhor o problema e gerar ações para melhorar a qualidade da informação. Para ajudar nesta avaliação o processo de cíclico do TDQM é essencial para garantir alta qualidade do produto de informação.

Visando melhorar a qualidade da informação e segregar por categoria a informação avaliada, serão selecionadas dimensões da qualidade da informação de acordo com a necessidade do usuário da informação. Assim como nos trabalhos de Reimer, Milinovich e Madigan (2016), Ardagna *et al.* (2018), Sánchez *et al.* (2019), Shahbazi *et al.* (2019), Günther *et al.* (2019) e Gürdür, El-Khoury e Nyberg (2019) serão selecionadas as dimensões que se adequarão mais a necessidade e contexto que a informação está inserida com o propósito de melhorar a qualidade da informação de uma situação específica. Serão selecionadas dentre as 15 dimensões citadas por Wang (1998) as que mais se adequarem ao contexto e a necessidade deste trabalho. O uso será justificado adequadamente à escolha sempre tendo como foco suprir a necessidade do usuário da informação.

Para facilitar a compreensão do produto de informação, Wang (1998) sugere que seja feito um desenho (esquema) demonstrando o sistema de manufatura de informação deixar clara as etapas de produção do produto de informação.

### **3.3.2 Etapa Medir**

Os dados serão coletados pelo time comercial por meio de um formulário, Anexo A, que será construído para coletar os dados dos clientes que precisarão alterar os dados cadastrais no sistema da empresa. Ele será construído para facilitar a interação entre a equipe comercial e a equipe de apoio ao time comercial, garantir que todos os dados necessários para o usuário da informação sejam coletados, formalizar o processo de coleta de dados e construir uma base para avaliar a qualidade da informação dos dados coletados.

A construção do formulário ocorrerá da seguinte forma: serão analisados os dados coletados na etapa de definição. Será feita uma verificação da necessidade da informação que o usuário necessita coletar. Serão consideradas as premissas para alteração cadastral do negócio. Será feita uma análise para entender os motivos mais frequentes de necessidade de alteração da modificação do cadastro dos clientes para padronizar os tipos de solicitação que os coletores de dados poderão selecionar. O formulário deverá ser on-line para que a equipe comercial tenha condições de solicitar mudanças com o mínimo de apoio e para que a tabulação dos dados seja automática.

Neste processo também serão criadas as métricas para a avaliação da qualidade da informação. Cada um dos dados coletados nos formulários serão avaliados de acordo com as

métricas construídas ao longo do desenvolvimento do trabalho. De acordo com Wang (1998) as métricas devem estar alinhadas aos requerimentos da qualidade da informação, possuir vínculos de avaliação relacionadas as dimensões mais importantes e é ideal que sejam implementadas nas rotinas do sistema de informação.

Serão criadas métricas e seus respectivos vínculos com as dimensões da qualidade da informação e relacionadas a cada campo do formulário. Elas serão criadas tendo como base a necessidade de verificação dos dados inseridos no formulário (necessidade do negócio) e como a informação deverá estar em cada em cada campo do formulário.

É importante frisar que as métricas e as dimensões deste trabalho serão definidas ao longo da pesquisa. Elas serão implementadas para subdividir o todo em mais partes, para classificar e facilitar as análises que serão feitas (OLIVEIRA e FEVARETTO, 2021; WANG,1998).

### **3.3.3 Etapa Analisar**

A análise dos dados coletados será feita pelo pesquisador que trabalhava na empresa e fazia parte da equipe de suporte comercial, ele representará o time suporte comercial e terá apoio do dos gestores comerciais. Segundo Coughlan e Coughlan (2002), o processo de análise dos dados é aconselhável que seja feito de maneira colaborativa entre os clientes e o pesquisador uma vez que se entende que eles serão os principais afetados pelas ações decorrentes. Esta análise colaborativa se justifica uma vez que os gestores serão os responsáveis por auxiliar no processo de mudança e implementação do plano de ação.

Para apoiar esta análise serão avaliados os resultados métricas da qualidade da informação. De acordo com Wang (1998), é necessário compreender a causa raiz dos problemas de qualidade da informação independentemente do método sendo ele simples ou complexo. Segundo a metodologia do TDQM a etapa de analisar deve vir após a etapa de medir.

Para avaliação das métricas será utilizada uma classificação binária sendo 0 para não atende e 1 para atende aos critérios da qualidade da informação de acordo com a métrica definida para determinado campo. Caso todos os dados estejam dentro dos padrões, uma determinada métrica receberá uma nota igual a 100% caso todos os dados estejam fora dos padrões receberá uma avaliação igual a 0%. Esta métrica estará vinculada a uma dimensão da qualidade da informação que terá nota igual da métrica avaliada. O peso de todas as métricas será igual a um (1) pois todas são igualmente importantes no caso estudado. Exemplo: foram

avaliados 10 dados vinculados a dimensão representatividade. 9 estavam dentro da métrica e 1 fora. Logo a avaliação desta métrica e da dimensão é igual a 90%.

Ao final destas análises serão separadas as métricas e dimensões que tiveram pior desempenho. Para Wang (1998), com base nos resultados esperados (*targeted payoffs*), é possível separar fatores críticos e os dados que deverão ser o foco da análise. Além de avaliar as métricas será proposto, também, um valor mínimo para dados bons, médios e ruins. Como todas as métricas possuirão o mesmo peso de importância, os casos classificados como ruins serão o alvo de implementação de melhorias. A qualidade dos dados coletados será avaliada de acordo com as dimensões da qualidade da informação definidas ao longo da pesquisa e avaliadas para verificar a aderência do que se esperava na coleta dos dados versus a coleta real dos dados. A análise terá como base as métricas criadas correspondentes às dimensões que o produto de informação foi medido.

Baseado nas análises que serão feitas, será identificada a causa raiz dos possíveis problemas de baixo nível de qualidade da informação e sugerido um plano de ação que será abordado no próximo Etapa 3.3.4.

### **3.3.4 Etapa Melhorar**

A construção do plano de ação ocorrerá com o alinhamento entre as equipes que participaram do processo de produção de informação com foco na melhoria da qualidade da informação para o usuário da informação e necessidade do negócio. A execução ocorrerá com o apoio dos gestores que ajudarão na implantação do plano de ação junto a equipe comercial. Wang (1998) sugere que nesta etapa deverão ser indicados os processos ou área que as melhorias deverão ser aplicadas partindo dos resultados obtidos na etapa da análise.

Para evidenciar se houve (ou não) melhorias no processo, será apresentada como resultado uma comparação entre a qualidade da informação da Etapa 3.3.1 definir e da pesquisa utilizando o TDQM com o objetivo de demonstrar se houve melhoria da qualidade da informação no comparativo entre as etapas do procedimento metodológico, ou seja, demonstrar um antes e um depois.

Por fim, será feito fechamento do trabalho descrevendo resumidamente o caminho percorrido desde o início até o fechamento do trabalho elencando dificuldades e lições aprendidas.

## 4 DESENVOLVIMENTO

Nesta etapa do trabalho será exposto o desenvolvimento das etapas do procedimento metodológico de acordo com o TDQM (*Total Data Quality Management*). Serão demonstradas a evolução do trabalho e as ações que serão tomadas de acordo com resultados das análises dos dados coletados.

### 4.1 ETAPA DEFINIR

Este tópico irá expor como foram executadas as etapas para definir as dimensões da qualidade da informação e os requisitos do que se procura melhorar. Além disso será apresentado um contexto do local (empresa) que será alvo do processo de melhoria e as áreas envolvidas no processo de produção de informação. Nesta etapa serão apresentados o modelo geral do produto qualidade da informação e as interações entre os envolvidos no sistema.

A pesquisa foi feita devido a necessidade da empresa e do usuário da informação aprimorarem o processo e a qualidade da informação coletada. O local escolhido para o trabalho foi um centro de distribuição de uma empresa do ramo de bebidas. Possui cerca de 50 funcionários e está localizado na cidade de Poços de Caldas, Minas Gerais. Globalmente a empresa alvo possui cerca de 80 mil funcionários, atua em mais de 80 países, possui um portfólio de mais de 300 marcas e mais de 150 anos de história.

Os principais motivos iniciais que levaram a escolha da unidade para fazer a pesquisa utilizando esta metodologia foram:

- Identificação de um problema de qualidade de informação no processo de modificação dos dados (cadastro) dos clientes.
- Empresa de referência na área,
- Unidade com menos pessoas se comprado outras unidades, sendo considerada pela empresa um bom local (laboratório) para testes de projetos futuros em outras localidades de acordo com os resultados alcançados.

De uma forma geral, o usuário da informação necessitou formalizar as solicitações de alteração do cadastro dos clientes que não seguiam um padrão e era feita de maneira informal. O projeto de melhoria da qualidade da informação baseado no TQDM teve início a partir da necessidade de melhorar a qualidade da informação das solicitações de alteração dos dados dos clientes vindas da equipe comercial. A seguir serão descritas as etapas utilizadas para definir o produto de informação.

#### 4.1.1 Identificação da necessidade de melhoria de um processo

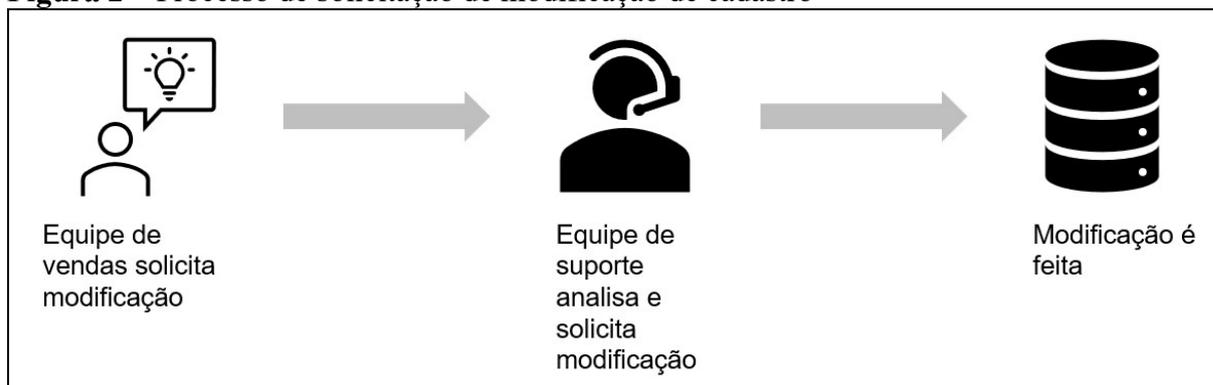
O cliente que compra os produtos da empresa necessita que os seus dados estejam corretos no banco de dados. Caso algum dado esteja incorreto pode impedir ou dificultar transações entre a empresa e os clientes. Exemplos: alteração de endereço, razão social, tabela de preços canal de vendas do cliente dentre outros. Para facilitar a classificação e identificação dos atores deste processo foi utilizada a abordagem Wang (1998) para caracterizar o papel de cada um dentro do processo.

De acordo com a abordagem de Wang (1998) que conceitua um sistema de manufatura como um sistema de informação que produz produtos de informação, são elencados quatro participantes: os fornecedores de informação (coletores e/ou criadores de informação), os fabricantes de informação (projetistas, desenvolvedores ou mantenedores da infraestrutura de dados e sistemas para o produto de informação), os consumidores (usuário do produto de informação no trabalho) e os gestores do produto de informação (gerenciam os processos e ciclos de vida referente ao produto de informação). Baseado nestes conceitos foram identificados e definidos no trabalho os seguintes envolvidos neste processo;

- Equipe comercial (vendedores, supervisores ou gerentes). Esta equipe é responsável por sinalizar os clientes que necessitam que o cadastro seja modificado. São os fornecedores de informação ou coletores de dados;

- Equipe de suporte comercial (especialistas, analistas e assistentes comerciais). Estes configuram os outros participantes do produto de informação. São ao mesmo tempo fabricantes de informação, os consumidores e os gestores do produto de informação. São eles que projetam como o sistema será construído, as informações necessárias que o banco de dados necessitar ter, consomem a informação e gerenciam o ciclo de vida do produto de informação.

Devidamente classificados o papel de cada área no processo de solicitação de alteração do cadastro dos clientes, a Figura 2 proporciona uma visão geral do processo primordial.

**Figura 2** – Processo de solicitação de modificação de cadastro

Fonte: Elaborado pelo autor.

O processo não possui controles em nenhuma etapa impossibilitando qualquer tipo análise a respeito da qualidade da informação. Todas as necessidades de alteração do cadastro dos clientes eram feitas pela equipe comercial em contato direto com a equipe de suporte comercial. Os pedidos de mudança de cadastro eram feitos por telefone, e-mail, mensagem de texto, redes sociais, recados em papel ou em conversas informais dentro do escritório.

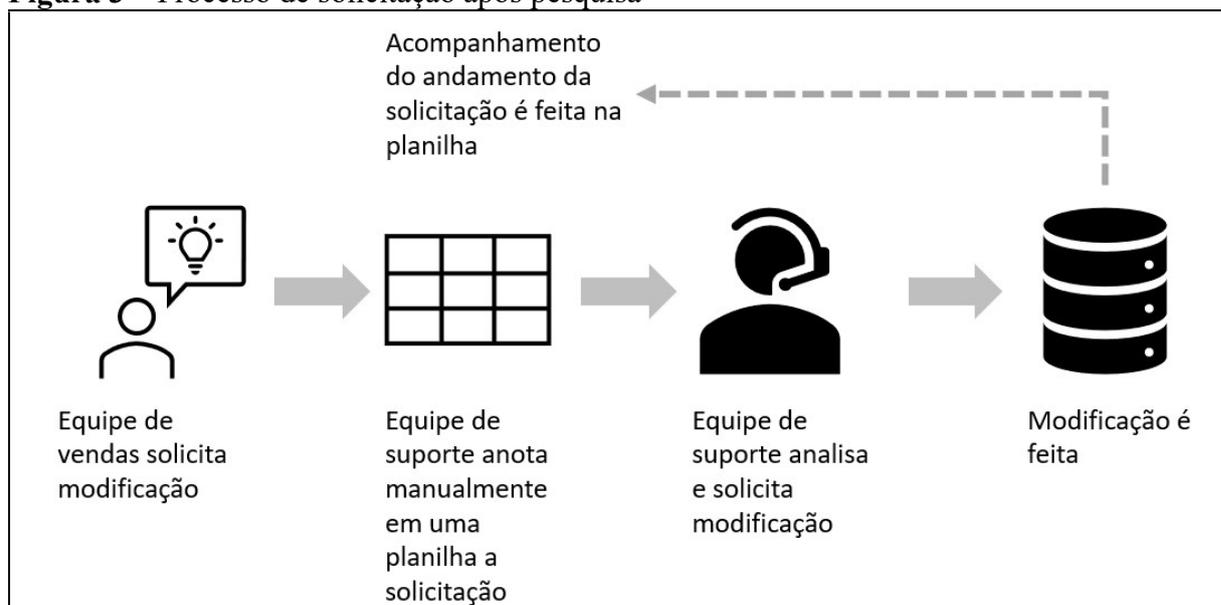
As solicitações de alteração do cadastro dos clientes eram pouco estruturadas, o nível de controle das informações era baixo e de difícil acesso quando havia a necessidade de rever as solicitações. As consequências deste fluxo eram:

- Falta de controle: tanto a equipe comercial quanto a equipe de suporte não tinham nenhuma forma de comprovar sistemicamente as solicitações de alteração de dados. Consequentemente as solicitações se perdiam devida a falta de controle e a falta de processos.
- Regras de negócios pouco claras: para que as solicitações fossem executadas pela equipe de suporte comercial são necessários alguns tipos de documentação de acordo com a necessidade de alteração dos dados. Apesar destas regras existirem e haver um documento formalizado, não foi dado treinamento a respeito dos tipos de documento necessários em cada caso de necessidade de alteração do cadastro do cliente para a equipe comercial. Além disso a equipe de suporte comercial era nova, o que tinha como consequência uma curva de aprendizado.
- Comunicação ineficiente: solicitações de modificação de dados de clientes muitas vezes se tornavam morosas devido à necessidade da equipe de suporte ter que voltar a falar com a equipe comercial solicitando documentos que não foram apresentados no ato da solicitação de modificação dos dados do cliente.

#### 4.1.2 Mudanças na maneira como os dados eram coletados

Como resultado das constatações feitas no Tópico 4.1.1, ficou visível que a forma como as tratativas das solicitações vinham sendo feitas estava ruim. Por isso, a equipe de suporte comercial passou a registrar as solicitações em uma planilha para que o processo ficasse menos desorganizado. A planilha construída possuía os seguintes campos que os coletores dos dados deveriam repassar para os usuários da informação: solicitante (nome), responsável (pessoa responsável pela tratativa da solicitação), setor (número do setor de vendas), matrícula (número), código do cliente, nome do cliente, data de nascimento do cliente, e-mail, cep, endereço, CNPJ ou CPF, canal, telefone, motivo, abertura (sistema que foi aberta a solicitação), data da solicitação, data da inserção (no sistema) e data final (data de resolução). A Figura 03 ilustra a mudança que foi feita no fluxo de coleta dos dados e controle das solicitações.

**Figura 3** – Processo de solicitação após pesquisa



Fonte: Elaborado pelo autor.

Os usuários da informação deveriam controlar o andamento das solicitações de acordo com os seguintes pontos: número da solicitação, e-mail de autorização (para modificação do cadastro), *status* (da solicitação), detalhe da alteração, abertura (data) e finalização (data).

Nesta pesquisa foram coletados 1.156 dados entre os meses de novembro de 2018 e agosto de 2019. Toda e qualquer solicitação passou a ser registrada na nesta planilha.

### 4.1.3 Análise e proposta de mudança frente ao problema

A partir da coleta de dados feita no Tópico 4.1.2 foi possível mapear as solicitações mais frequentes e padronizá-las reduzindo o número de solicitações distintas e ajudou a compreender se faltava algum tipo de informação necessária para que o usuário da informação trabalhasse com todos os dados que ele necessitava ou se havia dados desnecessários que estavam sendo coletados.

Foi possível identificar que existiam 85 descrições distintas no campo “motivo” solicitando modificações no cadastro dos clientes. Foi feita uma análise que identificou e agrupou estes motivos em 7 categorias de solicitações:

- desbloqueio ou reativação;
- correção de cadastro de clientes;
- expansão de clientes;
- alteração de tabela de preços;
- alteração de canal;
- produtos sem preço;
- impossibilidade de digitar pedido.

Essa segmentação ajudou a compreender melhor as necessidades dos coletores de dados (equipe comercial) e facilitou a análise dos usuários da informação no momento de alterar o cadastro dos clientes. Verificou-se também que alguns dados solicitados pelos coletores de dados não necessitavam ser preenchidos pois não necessitavam ser utilizados. Os campos eram: matrícula, CEP, endereço, canal, abertura, data da inserção, abertura e finalização.

Apesar dos dados dos clientes terem ficado mais organizados a qualidade dos dados ainda estava baixa, na avaliação (percepção) do usuário (consumidor da informação), que neste caso era o time de suporte comercial (assistentes, analistas e especialista).

Foram analisados os dados do campo “CNPJ ou CPF” e “E-mail” para terem bases mínimas para justificar ou comprovar a sensação (percepção) da baixa qualidade da informação e assim evidenciar a necessidade de um estudo focado na melhoria da qualidade da informação no processo de solicitação de alteração do cadastro do cliente. O campo do formulário “CNPJ ou CPF” é um campo que deve ser preenchido com números e estes números devem ter uma quantidade determinada de caracteres (no caso entre 11 e 18 para ser considerado um CNPJ ou CFP). Foram identificados 789 casos preenchidos inadequadamente

dos 1.156 casos analisados. 74,2% dos casos estavam com dados preenchidos corretamente. O campo “E-mail” deve ser preenchido por todos os clientes. Ele deve ter sempre “@” e “com”. Foram identificados 731 dados com essa informação preenchida de maneira incorreta dos 1.156 casos existentes. 31,7% dos dados estavam preenchidos corretamente.

Por esses motivos, percepção de baixa qualidade da informação e algumas evidências numéricas de um conjunto de dados, viu-se necessária uma pesquisa mais aprofundada sobre o tema com atuação ativa dos usuários da informação.

O TDMQ se enquadra como uma boa opção para a pesquisa aplicada na empresa alvo pois ele oferece um alicerce teórico relacionado ao tema qualidade da informação para auxiliar na resolução do problema, foi aprovado pelos gestores da empresa e foi considerado uma metodologia de fácil entendimento e aplicação. Foram utilizados também trabalhos com aplicação práticas citados no Tópico 2.3 “Aplicação da avaliação da qualidade da informação” que auxiliaram no entendimento de como as questões de avaliação da qualidade da informação foram aplicadas em um ambiente prático.

#### 4.1.4 Definição das dimensões da qualidade da informação

Wang (1998) cita que as características do produto de informação devem ser definidas pelo consumidor da informação. Para identificar as dimensões da qualidade da informação que mais se adequam a este trabalho foram feitas entrevistas com 4 gestores da área comercial (2 gerentes e 2 supervisores) e com 5 pessoas da equipe de suporte da área comercial (1 especialista, 3 analistas e 1 assistentes). O Quadro 9 resume o número de entrevistados, nível hierárquico de cada um dentro da empresa, descrição do cargo e como utiliza a informação no dia a dia.

**Quadro 9** – Resumo e descrição dos entrevistados

Hierarquia	Número de pessoas	Descrição do cargo	Uso da informação
Gerente	2	Gere toda a equipe comercial, toma decisões e aprova mudanças.	Utiliza a informação macro sem entrar em detalhes de cadastro.
Supervisor	2	Gere vendedores, executa decisões e faz a gestão das vendas.	Utiliza informação micro a nível de cadastro do cliente, sem interação com o sistema.
Especialista	1	Analisa e aponta oportunidades de melhoria	Utiliza informação micro a nível de cadastro do cliente, sem interação com o sistema.
Analista	3	Faz a gestão sistêmica dos clientes e cria ferramentas de acompanhamento	Utiliza informação micro a nível de cadastro do cliente, interação com o sistema.
Assistente	1	Auxilia na gestão das solicitações do chamado	Utiliza informação micro a nível de cadastro do cliente, interação com o sistema.

Fonte: Elaborado pelo autor.

As entrevistas foram realizadas com o objetivo de compreender o que este grupo considerava importante avaliar nos dados que serão coletados. O roteiro da entrevista foi estruturado com 3 perguntas abertas:

- A primeira com enfoque em dados preenchidos pelo time comercial, que não dependem de informação do cliente;
- A segunda com enfoque em informações que os clientes passam para o time comercial;
- A terceira focada em informações de controle das solicitações.

O pesquisador fez as entrevistas por meio de telefone com os envolvidos nas primeiras duas semanas do mês de agosto de 2019. Antes de cada entrevista o pesquisador passava uma rápida explicação sobre o problema de pesquisa, os objetivos do projeto e proporcionou uma noção ampla de qualidade da informação e dimensões da qualidade a informação para os entrevistados. Após a execução das entrevistas, elas foram analisadas. Estas análises identificaram as dimensões da qualidade da informação que mais tinham relação com as necessidades do negócio. Em seguida foi apresentado um panorama geral das análises das entrevistas para os envolvidos no trabalho. Desta forma, das 15 dimensões da qualidade da informação que Wang (1998) sugere, foram definidas 4 que apresentaram maior relação com as entrevistas e necessidade do negócio: temporalidade, relevância, reputação e quantidade de dados. Estas dimensões foram determinadas de acordo com a interpretação que o pesquisador fez das respostas dos entrevistados comparando com as definições das medidas propostas na literatura com enfoque nos estudos de Wang (1998), Wang e Strong (1996) e Wand e Wang (1996).

Cabe ressaltar que as demais dimensões elencadas no TDQM (credibilidade, acuracidade, objetividade, valor agregado, completa, interpretável, fácil de entender, representatividade consistente, representatividade concisa, acessível e acesso seguro) apresentaram menor relação com o projeto e necessidades do negócio de acordo com as análises das respostas das entrevistas, interpretação do pesquisador e necessidade do negócio. Por este motivo não serão utilizadas no trabalho.

A definição de cada dimensão não é universal e estão sujeitas a interpretação (FERREIRA; FERREIRA, 2017; BATINI *et al.* 2009; MATTIODA; FAVARETTO, 2006; BURGESS; GRAY; FIDDIAN, 2004; WAND; WANG, 1996). É possível que a seleção e quantidades de dimensões selecionadas varie de acordo com os participantes da pesquisa e a necessidade do negócio. Como o trabalho baseia-se na metodologia do TDQM optou-se por fazer uso, apenas das 15 dimensões elencadas por Wang (1998) não estendendo-se a outros

trabalhos que abrangem outros tipos de dimensão. A seguir serão detalhados os motivos da seleção das 4 dimensões selecionadas tendo como base as respostas das entrevistas.

Dos 9 entrevistados, 9 citaram (100%) ter necessidade de identificar o momento exato em que uma solicitação foi feita. Sem essa medida em todas as solicitações e com um alto grau de acuracidade não é possível saber com certeza quando a solicitação foi feita. Esta informação é muito importante para monitorar o momento da solicitação da modificação de cadastro e o tempo que as solicitações estão levando para serem concluídas. Wand e Wang (1996) definem que a temporalidade pode ser relacionada com a velocidade com que um sistema é atualizado após mudanças no mundo real e o tempo que os dados são de fato usados. Visto a definição da dimensão temporalidade elencada por estes autores e as respostas das entrevistas, esta necessidade do grupo foi identificada como a dimensão da temporalidade.

Dos 9 entrevistados, 8 citaram (89%) que saber a pessoa que fez a solicitação, o número do setor de vendas, o motivo da solicitação, o código e nome do cliente são de extrema importância para acompanhar as solicitações. Sem estas informações não é possível saber o responsável pela abertura da solicitação (o coletor do dado), o motivo e nem a qual cliente a solicitação faz referência. Sem estes dados não seria possível, por exemplo, contactar o solicitante da alteração de cadastro do cliente caso seja necessário confirmar algum dado do cliente, tão pouco identificar ao certo qual cliente é necessário fazer a alteração dos dados. Assim, dados errados tornam-se inúteis para o usuário da informação. Essa necessidade dos entrevistados se aproxima da dimensão relevância quando ela é definida da seguinte forma: “até que ponto os dados são aplicáveis e úteis para a tarefa em questão” (WANG; STRONG, 1996, p. 31).

Dos 9 entrevistados, 8 citaram (89%) que os dados vindos dos clientes (CPF ou CNPJ, e-mail data de nascimento e número do telefone) devem ter um alto nível de credibilidade ao passo que, se o cliente passar alguma informação errada não será possível que nenhuma ação de cadastro ou modificação no cadastro sejam feitas. Logo, para que uma venda ocorra é necessário que os dados inseridos dos clientes estejam corretos. São dados que necessitam de algum tipo de validação do cliente caso tenham mudado. Segundo Wang e Strong (1996) a reputação dos dados é uma dimensão que tem relação a confiabilidade dos dados em relação a fonte ou conteúdo. Visto que as fontes dos dados são externas, vindas dos clientes, a dimensão selecionada, tendo como embasamento a visão dos entrevistados, foi a reputação.

Dos 9 entrevistados, 7 (78%) mencionaram que os campos de adição de imagem e observação são de extrema importância em casos que são necessários passar mais detalhes

sobre um determinado dado já informado. Por exemplo, quando um cliente precisa mudar o canal de venda, passar de bar para restaurante, é necessário detalhar o que deve ser feito no campo observação e adicionar 3 fotos do local de venda. Caso as informações não estejam lá, a solicitação pode ser negada. De acordo com Wang e Strong (1996) a dimensão quantidade de dados apropriada tem relação com o volume ou quantidade de dados apropriado para a atividade. Visto o contexto das entrevistas e a definição da dimensão quantidade apropriada de dados foi selecionada para fazer parte da avaliação.

A Quadro 10 ilustra um resumo do processo que envolveu a definição das dimensões da qualidade da informação neste trabalho:

**Quadro 10** - Resumo das dimensões selecionadas por campo e pergunta

Foco da pergunta	Dados necessários coletados	Dimensão selecionada
Dados internos da empresa	Nome do Vendedor, Setor do vendedor, Motivo do Chamado, Cod_Cliente, Nome Cliente	Relevância
Informação do cliente	CPF/CNPJ, e-mail, data de nascimento do cliente, telefone	Reputação
Dados de controles das solicitações	Data da solicitação, Data de finalização, Aprovação (e-mail do gestor), Número da solicitação, Ponto de ação (campo livre), responsável pela tratativa	Temporalidade, Relevância e quantidade de dados apropriados
Dados adicionais	Observação e Adição de Imagens	Quantidade de dados apropriados

Fonte: Dados da pesquisa

É importante que seja construído um esquema para representar as características do produto de informação para facilitar a análise para pessoas que não são familiarizadas com o sistema e consigam interagir com ele facilmente (WANG, 1998). O Quadro 11 demonstra as informações que foram coletadas pelo time comercial (coletores de dados) e as informações que os usuários da informação inseriram de acordo com o andamento do processo de solicitação de modificação de cadastro dos clientes. Ao lado de cada campo está explícito qual dado deve ser inserido em cada campo da coleta e controle dos dados. Este esquema auxiliou na construção do formulário usado para coleta de dados na próxima etapa da pesquisa.

**Quadro 11** - Esquema ilustrativo dos dados coletados e controlados

Equipe responsável	Função	Campos	Dimensão
Comercial	Coletar dados	Nome do Vendedor, Setor do vendedor, Motivo do Chamado, Cod_Cliente, Nome Cliente	Relevância

Continua

Conclusão

Comercial	Coletar dados	CPF/CNPJ, e-mail, data de nascimento do cliente, telefone	Reputação
Equipe de suporte comercial	Controlar e avaliar	Data da solicitação, Data de finalização	Temporalidade
Equipe de suporte comercial	Controlar e avaliar	Aprovação (e-mail do gestor), Número da solicitação, responsável pela tratativa	Relevância
Equipe de suporte comercial	Controlar e avaliar	Ponto de atenção (campo livre)	Quantidade de dados apropriados
Comercial	Coletar dados	Observação e Adição de Imagens	Quantidade de dados apropriados

Fonte: Dados da pesquisa, formato adaptado de Wang (1998).

Para complementar o esquema representado do Quadro 11 foi construído um sistema de produção de informação para proporcionar uma visão geral da produção e informação (produto de informação) representado na Figura 4 que ilustra o sistema de produção de informação utilizado na pesquisa. Este fluxo é uma analogia entre um sistema de informação e um sistema de produção de produtos físicos com entradas, processos e saídas.

**Figura 4 - Sistema de produção de informação**



Fonte: Elaborado pelo autor.

Definidos critérios do produto de informação, os requisitos de qualidade da informação, as interações entre participantes do processo e como o produto de informação é produzido, a organização possui duas opções segundo Wang (1998): desenvolver um novo sistema de informação ou desenvolver mecanismos para remediar a deficiência do sistema existente. Estas duas opções foram compartilhadas com os gestores da empresa e optou-se por desenvolver um novo sistema.

Partindo desta premissa, definiu-se mudar o processo de coleta, armazenagem e gestão dos dados. Ele passará a ser feito por meio de um formulário eletrônico compartilhado com a equipe comercial (coletores de dados) que preencherá este formulário sempre que houver a necessidade de solicitar a mudança dos dados dos clientes. Ao preencher o formulário, os dados já serão tabulados automaticamente e documentados (armazenados) em

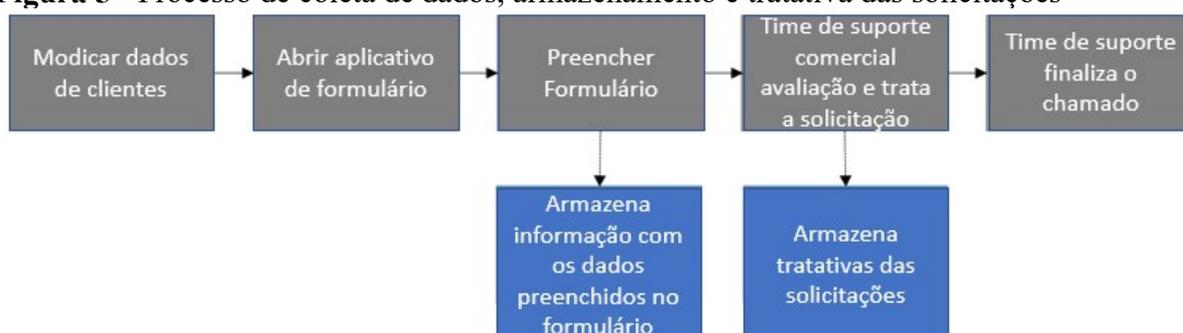
um banco de dados. Após a coleta dos dados, o time de suporte comercial fará a tratativa e o acompanhamento das solicitações pelos usuários da informação. A construção e o processo destes itens serão abordados em detalhes no próximo tópico.

#### 4.2 ETAPA MEDIR

Os dados desta pesquisa foram coletados pela equipe comercial por meio de um formulário eletrônico (ANEXO A). O formulário foi confeccionado para atender a necessidade do usuário relativo as informações que ele necessitava para executar os processos de alteração dos dados dos clientes.

O formulário possui um total de 15 campos para preenchimento sendo 11 perguntas obrigatórias (3 perguntas de múltiplas escolhas com apenas uma opção de escolha, 9 questões abertas) e 4 optativas (uma para observações e 3 campos para inserção de imagem). Após a entrada dos dados no formulário o time de suporte comercial faz o controle e a tratativa das solicitações individualmente. O controle dos dados possui 7 campos sendo um campo preenchido automaticamente (data e hora) assim que a solicitação é inserida pelo solicitante, e 2 campos de preenchimento obrigatório e 4 campos de preenchimento opcional. A Figura 5 ilustra como foi o processo da coleta e tratativa dos dados. As etapas em cinza sinalizam as etapas que possuem interação humana no processo. As em azul representam as etapas de armazenamento de dados sem a intervenção de pessoas.

**Figura 5** - Processo de coleta de dados, armazenamento e tratativa das solicitações



Fonte: Elaborado pelo autor.

O processo inicia-se com a necessidade de alteração dos dados de algum cliente. O aplicativo para solicitar a modificação é aberto. É necessário preencher o formulário campo a campo. Após o preenchimento dos campos o solicitante do chamado confirma as informações. As informações são enviadas para um banco de dados. Estas informações chegam até o time de suporte comercial. A equipe de suporte fará as tratativas necessárias referentes a

solicitação. As tratativas serão armazenadas em um banco de dados. Após finalizar o processo da tratativa o time de suporte comercial finaliza o processo.

A coleta de dados deste formulário teve início no dia 16/08/2019 e foi finalizada no dia 29/12/2020 sendo coletados um total de 425 solicitações de alteração de dados e contou com a participação de 26 coletores de dados (equipe comercial).

O mecanismo para medir como a qualidade da informação é a criação de métricas (WANG, 1996). Neste trabalho as métricas foram construídas de acordo com cada dado que foi incluso no formulário e de acordo com as necessidades do negócio. No total foram criadas 19 métricas sendo 7 relacionadas as tratativas do time de suporte e 12 as solicitações do time comercial.

Os dados que serão criadas métricas relativas às tratativas do time de suporte são: Data de abertura (Data/hora), E-mail de aprovação, Número do Chamado, Status do Chamado, Pontos de atenção, Aberto por e Data de abertura.

Os dados que serão criadas métricas relativas as solicitações são: Vendedor, Setor, Motivo do Chamado, Código do Cliente, Nome do Cliente, CPF/CNPJ, E-mail, Data de nascimento do Cliente, Telefone, Atenção para a alteração, Observação e Adição de Imagens.

De acordo com Wang (1998), independentemente da natureza da métrica de qualidade da informação, ela deve ser implementada na rotina do sistema. Para cada métrica criada será necessário implementar um procedimento que seja capaz de fazer a verificação desta métrica. Adicionalmente para cada campo do formulário foi atribuída uma das 4 dimensões selecionadas na etapa de definir: temporalidade, relevância, reputação e quantidade de dados. Cada uma destas dimensões será associada a um dado coletado. Por exemplo: os dados de “hora e data” da solicitação são associados a dimensão temporalidade. Os dados relacionados ao “Setor” de vendas associam-se a dimensão relevância. Cada uma dessas associações foi definida na etapa de planejamento. Logo, para cada dado inserido no formulário haverá uma métrica e uma dimensão associada.

Serão apresentados mais detalhes relativos às 19 métricas criadas para cada campo do formulário. Além disso será indicada a dimensão da informação que aquele campo pertence e uma explicação do que o campo deve conter de informação. Este detalhamento será dividido em duas partes: métricas de controle e métricas de coleta de dados. As métricas de controle de solicitações são relativas aos controles que são feitos a respeito dos dados coletados pela equipe do time comercial, são 7 campos no total.

As Métricas de coleta de dados, são os dados que a equipe comercial insere no formulário. São 15 dados no total.

### Métricas de controle de solicitações

Data de abertura (Data/hora): este campo corresponde a data e hora que um dado foi inserido no formulário. A geração deste campo ocorre de forma automática, sem a necessidade de interação manual para a criação deste dado. A métrica associada a ele é o preenchimento ou não do campo. Exemplo: caso o campo esteja com dados a informação tem qualidade, caso o campo não esteja com dados (campo em branco) a informação é de má qualidade. Cabe ressaltar que este campo é gerado automaticamente, logo, para que um dado não exista, é necessário que tenha havido uma falha no sistema ao registrar o chamado. A dimensão associada ao campo é a temporalidade.

E-mail de aprovação: este campo refere-se ao e-mail de aprovação do gestor da área frente à necessidade de alteração dos dados do cliente. O preenchimento deste campo é manual, feito por algum membro da equipe de suporte comercial. A métrica associada é o preenchimento ou não do campo. Exemplo: um vendedor solicitou a alteração do cliente “A”. O gestor é comunicado por e-mail e posteriormente deve dar um “ok” para esse e-mail. Caso o gestor tenha aprovado é inserido um “sim” no campo e-mail de aprovação, caso não seja necessária aprovação do gestor é inserido um “não”. Caso o campo não esteja preenchido (campo em branco), a informação é de má qualidade. O não preenchimento deste campo não impede que a equipe de suporte comercial dê prosseguimento a solicitação, o não preenchimento deste campo é considerado como falha processual e pode ser inserida a informação a qualquer momento. A dimensão associada ao campo é a relevância.

Status do Chamado: campo destinado a verificar o andamento da solicitação. Este campo pode possuir 4 tipos de Status: “Novo”, “Finalizado”, “Cancelado” ou “Em Branco”. Qualquer solicitação que entra no sistema com o campo “em branco” no status do chamado. Quando um analista da área de suporte comercial inicia as tratativas o status vai para “novo”. Quando as tratativas são finalizadas o status vai para “finalizado”. Se faltou algum tipo de informação ou autorização que impeça a modificação do cadastro do cliente, o status será classificado como “cancelado”. A métrica associada é o preenchimento do campo “Status do Chamado” com "Novo", "Cancelado", "Finalizado" ou "Em branco". Caso o campo “Número do Chamado” esteja preenchido o status deve estar com qualquer tipo de status exceto “Em branco”. Exemplo: se uma pessoa do suporte comercial iniciar a análise de uma solicitação, ela deve passar o status do processo de “em branco” para “novo”. Caso esse processo já tenha um “Número de Chamado” e estiver ainda com um status de “em branco”, ele estará fora da métrica. Esta informação é preenchida manualmente e pode ser alterada a qualquer momento. A dimensão associada ao campo é a relevância.

**Número do Chamado:** este campo relaciona-se a um número de controle das solicitações feitas pela equipe comercial. O número do chamado é gerado automaticamente após a formalização da solicitação ser feita no sistema ERP da empresa. Ele é gerado automaticamente, porém o preenchimento desta informação na área de controle das solicitações é feito manualmente. A métrica associada é o preenchimento ou não do campo (exceção à regra – solicitação cancelada não precisa ter número de chamado). Exemplo: o analista da equipe de suporte comercial recebeu uma solicitação de alteração de endereço de um cliente. Esta alteração será feita no sistema ERP da empresa. Assim que o analista fizer esta alteração nos campos do ERP, o sistema gerará um número de controle. Esse número é o “número do chamado”. Após esta geração o analista deve copiar este número e inserir no campo “Número do chamado”. Caso o campo não esteja preenchido (campo em branco), a informação é de má qualidade. A exceção a esta regra é quando a solicitação foi cancelada, não havendo a necessidade de um número de controle. O não preenchimento deste campo não impede que a equipe de suporte comercial prossiga com a solicitação. A dimensão associada ao campo é a relevância.

**Pontos de atenção:** esse campo é destinado ao detalhamento das solicitações quando necessário. É um campo livre para escrita. Ele deve ser preenchido obrigatoriamente quando o campo “Status do Chamado” estiver preenchido como seja "Cancelado". Logo a métrica associada ao campo é campo preenchido obrigatório caso o chamado seja "Cancelado", caso não seja “Cancelado”, o campo não necessita estar preenchido. Exemplo: uma solicitação foi analisada e finalizada com sucesso. O analista colocou uma observação “vendedor forneceu o comprovante de endereço por e-mail”. Essa observação não era obrigatória, porém não tem problema quanto à avaliação da qualidade da informação caso haja adição de texto. Cabe ressaltar que apenas os casos que se enquadrarem na regra de preenchimento obrigatório que serão avaliados a qualidade da informação. A dimensão associada ao campo é a quantidade de dados apropriados.

**Aberto por:** este dado indica quem é o responsável pela tratativa da solicitação, deve ser preenchido com o nome do analista responsável. A métrica associada a este campo é se o campo foi preenchido ou não. A dimensão associada ao campo é relevância.

**Data da finalização:** é um campo preenchido quando todas as tratativas relacionadas a solicitação foram feitas e ela estiver com status de “finalizado” no campo “status do chamado”. A métrica associada é se o campo deve estar preenchido com uma data se o status da solicitação estiver como “finalizado”. Para qualquer outro status o campo não precisa estar preenchido. Cabe ressaltar que apenas os casos que esse enquadrarem à regra desta métrica

serão avaliadas a qualidade da informação. A dimensão associada ao campo é a temporalidade.

#### Métricas de coleta de dados

**Vendedor:** refere-se ao nome do coletor de dados. O time comercial preenche esse campo com o nome da pessoa que está solicitando a modificação do dado do cliente. É um campo livre para escrita. A métrica associada é: preenchimento com letras. Caso o campo não esteja preenchido com letras ele será considerado sem qualidade de informação. A dimensão associada ao campo é a relevância.

**Sector:** o setor corresponde ao grupo de clientes associadas a uma região de vendas (setor de vendas) e a um vendedor. Esse setor deve ser constituído sempre por 5 números. Para o preenchimento deste campo não são permitidas inserções de letras (regra estabelecida previamente no formulário). Exemplo: o setor 15900 possui 150 clientes e é de responsabilidade do vendedor João da Silva. A métrica associada é: campo preenchido com exatamente 5 números. Se forem mais ou menos que 5 números o valor não é considerado como um dado de qualidade. A dimensão associada ao campo é a relevância.

**Motivo do Chamado:** o motivo do chamado é a razão (justificativa) que motivou a solicitação de mudança do cadastro do cliente. Ele é constituído por 7 motivos distintos que foram mapeados na Etapa 4.1 (Definir) e padronizados posteriormente. Os motivos são:

- **Desbloqueio/reactivação:** clientes que o sistema não permite a venda por motivos diversos e é necessário fazer algum tipo de tratativa que permita que o cliente volte a ter a permissão para comprar produtos;
- **Expansão de clientes:** cliente não está associado ao local de venda correto. Exemplo: o cliente 001 deve pertencer ao local de vendas L1, porém no sistema ele está no local L2. Logo, não será possível fazer vendas para ele até que seja alterado do local L2 para L1;
- **Alteração de canal (detalhar no campo observação a mudança):** o estabelecimento comercial está com cadastro divergente do que ele é. Exemplo: o cliente é um bar, porém está cadastrado no sistema como supermercado, o canal está divergente e deve ser alterado;
- **Alteração de tabela de preços:** o cliente está com uma tabela de preço divergente da que deveria ter. Exemplo: o cliente é um bar e está com uma tabela de preços de supermercado;
- **Produtos sem preço:** o produto não está com preço no sistema. Exemplo: quando o vendedor consulta o preço de determinado produto, o produto está com o preço zerado;

- Impossibilidade de digitar pedido: o cliente aparece com apto para compra, porém não é possível executar o processo de venda de produtos;
- Correção de cadastro de clientes (detalhar no campo observação): campo destinado a modificação e endereço, razão social ou qualquer tipo de cadastro relacionado a localização ou questão fiscais do cliente. Exemplo: o cliente mudou de endereço e de dono. O vendedor deverá solicitar esta modificação no sistema para que a entrega não ocorra no local errado e os documentos fiscais sejam destinados à pessoa correta.

Não é permitida a seleção de mais de um motivo por solicitação. Esse campo é de preenchimento obrigatório sendo impossível enviar um formulário com este campo sem preenchimento a menos que haja um erro no sistema. A métrica associada é: campo preenchido para que esta informação tenha qualidade. A dimensão associada ao campo é a relevância.

Código do cliente: campo destinado ao preenchimento do código do cliente. Todo cliente possui um tipo de código que pode variar entre 6 e 8 caracteres e deve conter apenas números. Para o preenchimento deste campo não são permitidas inserções de letras (regra estabelecida previamente no formulário). Métrica associada: campo preenchido de 6 a 8 caracteres. Caso haja mais ou menos números que o estabelecido, a informação não possui qualidade. A dimensão associada ao campo é a relevância.

Nome do Cliente: local onde deverá ser inserido o nome do cliente. Pode ser nome popular ou nome de registro fiscal (CNPJ/CPF). A métrica associada é: campo preenchido. Caso o campo esteja em branco a qualidade da informação é ruim. A dimensão associada ao campo é a relevância.

CPF/CNPJ: campo deve ser preenchido com os números de registro fiscal ou de identificação de pessoa física. A métrica associada a estes campos é o preenchimento entre 11 e 19 caracteres. Caso o campo não tenha o número de caracteres dentro deste intervalo a informação possui baixa qualidade. Exemplo: o vendedor precisa preencher o CPF do José. 000.000.000-95 (14 caracteres) ou 00000000095 (11 caracteres) ou CNJP 00.000.000/0001-00 (18 caracteres) ou 00000000000100 (14 caracteres). A dimensão associada ao campo é a reputação.

E-mail: endereço eletrônico necessário para o cliente receber documento enviados após a compra de produtos. Caso este campo esteja incorreto a informação não chega ao cliente. A métrica associada é: existência de “@” e “com” no texto digitado. Caso não tenha os dois caracteres no texto a informação será considerada de baixa qualidade. A dimensão associada ao campo é a reputação.

Data de nascimento do cliente: deve ser preenchida a data de nascimento do cliente. Métrica associada é: campo preenchido com uma data. Caso não contenha uma data a informação será considerada de baixa qualidade. A dimensão associada ao campo é a reputação.

Telefone: número referente ao contato telefônico do cliente. Este campo deve ter entre 8 e 13 caracteres sendo possível apenas a inserção de números. A métrica associada é: campo preenchido entre 8 e 13 números. A dimensão associada ao campo é a reputação.

Atenção para a alteração: campo de deve ser preenchido com sim ou não. A métrica segue a seguinte regra: será feita a seguinte verificação, caso “Cod\_Cliente” inicie com 1 e o motivo da solicitação seja "alteração de canal (detalhar no campo observação a mudança)" ou "correção de cadastro de clientes (detalhar no campo observação)" esse campo deve ser necessariamente respondido como “sim” para ser considerada uma informação com qualidade. Caso contrário, a informação será considerada de baixa qualidade. Cabe ressaltar que apenas os casos que se enquadrarem as estas regras serão avaliadas a qualidade da informação. Logo, para esta métrica, a amostra avaliada sempre será menor que o total dos dados coletados. A dimensão associada ao campo é a qualidade de dados adequados.

Observação: campo de livre escrita. Pode ser preenchido independentemente do tipo de solicitação sendo obrigatório caso o motivo do chamado seja “alteração de canal (detalhar no campo observação a mudança)". Exemplo: solicitação de alteração de canal. Mudar de bar para restaurante. A métrica para avaliar esse campo é: solicitação “alteração de canal” é obrigatório o preenchimento do campo “observação”, caso não esteja preenchido é considerado informação com baixa qualidade. Cabe ressaltar que apenas os casos que se enquadrarem a esta regra serão avaliadas a qualidade da informação dentro dos critérios desta métrica. A dimensão associada ao campo é a qualidade de dados adequados.

Adição de Imagens: campo dedicado a adição de imagens. As imagens podem ser imagens do local, documento que comprovem o endereço ou qualquer outra necessidade que o solicitante tenha de evidenciar por imagem. A imagem é obrigatória em casos que o número do cliente se inicia com o número um em solicitações quando o motivo for ou “alteração de canal (detalhar no campo observação a mudança)” ou "correção de cadastro de clientes (detalhar no campo observação)". Exemplo: o cliente que se inicia com o número 1 tem necessidade de alteração de canal de vendas bar para restaurante. Logo, o vendedor que está solicitando a mudança deve adicionar uma imagem do local que será avaliada posteriormente pelo time de suporte comercial se a solicitação pode ser executada ou não. Cabe ressaltar que apenas os casos que se enquadrarem a estas regras serão avaliados a qualidade da informação

considerando os critérios de avaliação desta métrica. A dimensão associada ao campo é a qualidade de dados adequados. Observações: clientes que iniciam com o número 1 são cliente cadastrados em um CPF, por não terem dados em um sistema que vincule o CPF a uma atividade comercial a conferência, a regra de negócios da empresa exige uma comprovação por foto. Existem mais dois campos no formulário para inserção de imagem, qualquer um dos campos que a imagem for adicionada já está suficiente para que este ponto seja avaliado não sendo, portanto, necessária a avaliação de cada campo em separado.

CDA: campo correspondente ao centro de distribuição que o membro do time comercial pertence. Como é um campo que analisa apenas uma localidade ele não será avaliado.

O Quadro 12 apresenta um resumo de todas as considerações feitas anteriormente a respeito de cada campo do formulário.

**Quadro 12 - Métricas por campo do formulário e dimensão da informação**

<b>Pergunta (texto igual ao formulário)</b>	<b>Dimensão da qualidade a informação avaliada</b>	<b>Métrica para avaliação</b>
Data de Abertura	Temporalidade	Campo preenchido
E-mail de aprovação	Relevância	Campo preenchido
Número do chamado	Relevância	Campo preenchido (se for cancelado não necessita estar preenchido)
Status do Chamado	Relevância	Campo preenchido com "Novo", "Cancelado", "Finalizado" ou "Em branco".
Pontos de atenção	Quantidade de dados apropriados	Campo preenchido obrigatório caso o chamado seja "Cancelado"
Aberto por	Relevância	Campo preenchido
Data de finalização	Temporalidade	Campo preenchido caso o chamado seja "Finalizado"
Vendedor	Relevância	Campo preenchido com letras
Setor	Relevância	Campo preenchido com 5 números
Motivo do Chamado	Relevância	Campo preenchido
Código do cliente	Relevância	Campo preenchido entre 6 e 8 números
Nome do Cliente	Relevância	Campo preenchido com letras
CPF/CNPJ	Reputação	Campo preenchido com números entre 11 e 18
E-mail	Reputação	Campo preenchido que contenha @ e .com
Data de nascimento do cliente	Reputação	Campo preenchido que contenha uma data
Telefone	Reputação	Campo preenchido com números entre 8 e 13

Continua

## Conclusão

Atenção para a alteração	Quantidade de dados apropriados	Campo preenchido com "sim" caso o Cod_Cliente inicie com 1 e o motivo da solicitação seja "ALTERAÇÃO DE CANAL (DETALHAR NO CAMPO OBSERVAÇÃO A MUDANÇA)" ou "CORREÇÃO DE CADASTRO DE CLIENTES (DETALHAR NO CAMPO OBSERVAÇÃO)"
Observação	Quantidade de dados apropriados	Campo preenchido obrigatoriamente caso o motivo da solicitação seja "motivo do chamado seja "ALTERAÇÃO DE CANAL (DETALHAR NO CAMPO OBSERVAÇÃO A MUDANÇA)"
Adição de Imagens Obs.: 3 campos no formulário para inserção de imagens	Quantidade de dados apropriados	Campo preenchido caso o Cod_Cliente inicie com 1 e o motivo da solicitação seja "ALTERAÇÃO DE CANAL (DETALHAR NO CAMPO OBSERVAÇÃO A MUDANÇA)" ou "CORREÇÃO DE CADASTRO DE CLIENTES (DETALHAR NO CAMPO OBSERVAÇÃO)"

Fonte: Elaborado pelo autor.

Para entender o grau de importância de cada métrica foi questionado aos gestores da área comercial e a equipe de suporte comercial qual métrica era mais ou menos importante. Como todas as informações eram importantes e para simplificar a forma de analisar optou-se por considerar todas com o mesmo peso. Para este questionamento não foi aplicado nenhum questionário. Também foi compartilhada qual seria a forma mais adequada de avaliar cada métrica relativa à qualidade do dado inserido em cada campo ou seria uma escala de intervalos numéricos ou escala nominal (atente ou não atente).

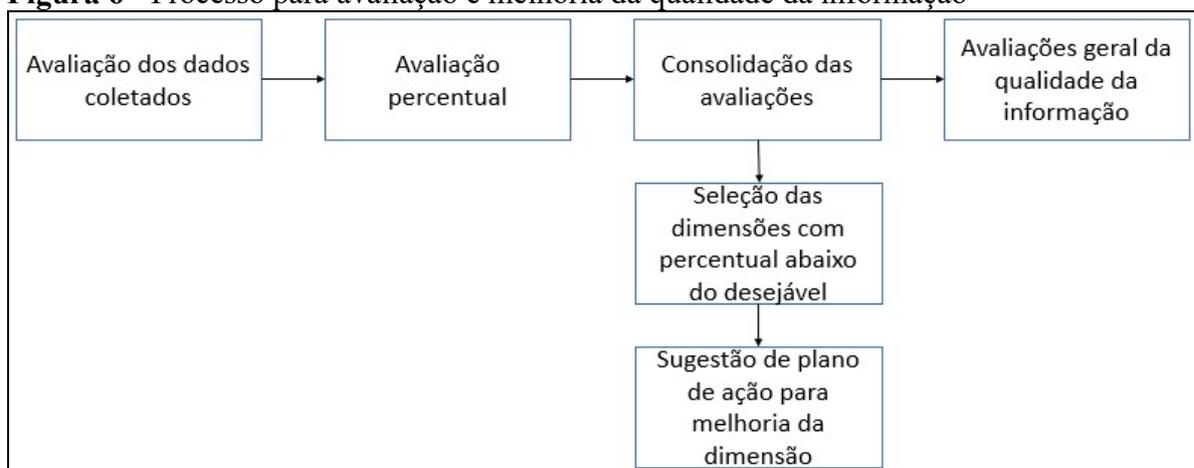
Foi escolhida pelo grupo de trabalho uma escala nominal (um ou zero - atende ou não atende) para calcular as métricas. Para o dado que atende seria atribuída uma nota igual a um e para a que não atende zero. Os principais argumentos utilizados pelos gestores e pelo time de suporte comercial para esta escolha foram: os dados precisam ter informações assertivas não sendo aceitável dados “meio certos”, não são campos extensos para preenchimento, simplificação do cálculo (facilidade de explicar e replicar para outras áreas da companhia) e facilidade para automatizar. Foi considerado que todas as dimensões têm o mesmo grau de importância, não sendo priorizada nenhuma dimensão sendo mais ou menos importante para o trabalho. Desta forma, a escala seguiu padrão binário – 0 ou 1 sendo zero não tem qualidade e 1 tem qualidade, o peso de todas os campos é o mesmo.

Para demonstrar o resultado consolidado por dado coletado será feita uma representação percentual de 0% a 100%. Foi definido junto a gestão que a informação com mais de 95% considerada boa, entre 90% e 95% média e abaixo de 90% ruim. Para as métricas com baixa avaliação serão feitas análise detalhadas para entendimento da causa raiz da baixa qualidade da informação e serão traçados planos para que esse resultado melhore em

conjunto com os gestores da área comercial. O numerador da equação será a somatória dos dados que atendem a métrica e o denominador será a somatório dos dados coletados na maior parte dos casos. Exemplo: a métrica “vendedor” teve 100 dados coletados. Dos 100, 98 estavam aderentes aos requisitos da qualidade da informação da métrica. Resultado percentual é:  $98/100 = 98\%$ . Logo o campo, “Vendedor” foi preenchido 98% das vezes de maneira correta de acordo com a métrica proposta, este conjunto de dado possui 98% de qualidade da informação.

Para as métricas “Atenção para a alteração”, “Observação”, “Adição de Imagens”, “Pontos de atenção” e “Data de Finalização”, como já citado na definição das métricas, não serão utilizados todos os dados coletados para avaliação. Serão utilizados apenas quando os dados coletados se enquadrarem a regra da métrica. Exemplo: foram coletados no total 100 dados. A métrica para avaliar “Data de Finalização” consiste em ter data de finalização da solicitação somente se ela estiver com status de finalizado. Das 100 solicitações, 40 possuem o status de “cancelada” e 60 “finalizado”. Dos 60 finalizados, 5 não possuem data de finalização e 55 tinham. Logo, o numerador desta equação seriam os 55 com data e o denominador seria 60, resultando em 91,7% de qualidade de informação para este caso.

A avaliação das dimensões será feita por meio da somatória dos pontuados dividido pelo total de dados coletados respeitando regras estabelecidas por métrica. Exemplo: dimensão temporalidade tem 2 campos que compõem a avaliação desta dimensão (Data de Abertura e Data de Finalização). Dos 100 formulários respondidos foram considerados 100 para o campo Data de Abertura sendo que 100 estavam de acordo com os critérios da métrica da métrica. Para o campo Data de Finalização foram considerados 60 formulários por motivo da mecânica da métrica e 55 estavam dentro dos parâmetros da métrica. Logo o cálculo para a dimensão ficaria 155 dados com qualidade de informação em 160 considerados na análise. Para a dimensão da temporalidade o percentual de qualidade da informação atingido foi de 96,9%. A Figura 6 demonstra resumidamente o processo de avaliação desde a inserção do dado por campo do formulário até a consolidação das avaliações.

**Figura 6 - Processo para avaliação e melhoria da qualidade da informação**

Fonte: Elaborado pelo autor.

O uso das métricas auxiliará na verificação da qualidade da informação inserida pelos coletores dos dados nos controles feitos pela equipe de suporte comercial e a quantificar o quanto a informação está boa ou ruim. As métricas criadas e os processos descritos nesta etapa do trabalho darão o suporte necessário para que as análises dos dados sejam executadas.

#### 4.3 ETAPA ANÁLISE DOS DADOS

A partir dos dados coletados e das métricas traçadas será possível determinar a qualidade da informação. Foram selecionadas as dimensões e demonstrado como serão executados os cálculos para mensurar numericamente a qualidade da informação. Serão demonstrados nesta análise os resultados da qualidade da informação vinculadas as dimensões que permeiam cada um dado coletado.

Além da análise e vinculações das análises com a teoria, Wang (1998) sugere que, a partir da mensuração da qualidade da informação seja investigada a causa raiz dos problemas de qualidade da informação. Considerando os critérios de avaliação estabelecidos, a dimensão que apresentar resultados abaixo de 90% será foco da análise para melhorar a qualidade da informação.

A seguir serão executados os passos descritos na Figura 6 para avaliar a qualidade da informação. Em seguida será selecionada a dimensão com pior avaliação. Para esta dimensão será feito um plano de ação para melhorar a qualidade da informação.

##### 4.3.1 Avaliação dos dados coletado

Os dados coletados receberam, individualmente uma avaliação de 0 ou 1 de acordo com a métrica associada ao dado coletado sendo 1 para um dado com qualidade e 0 para dado

sem qualidade. Para facilitar a compreensão desta análise, os dados com qualidade serão chamados de dados pontuados, pois recebem uma nota igual a 1. Para os dados com baixa qualidade serão chamados de não pontuados, pois receberam uma nota igual a 0.

No total foram coletados e preenchidos 425 formulários. Os dados que considerarão o total de dados coletados para avaliação (425) foram: data de abertura, vendedor, setor, motivo do chamado, código do cliente, nome do cliente, CPF/CNPJ, e-mail, data de nascimento do cliente, telefone, e-mail de aprovação, número do chamado e status do chamado e aberto por.

Os dados que considerarão parte da amostra devido a premissas das métricas associadas aos dados foram: Atenção para alteração (15 dados considerados), Adição de Imagens (15 dados considerados), Observação (25 dados considerados), Pontos de atenção (133) e Data de Finalização (288).

O Quadro 8 demonstra o número de dados pontuados por campo de dados coletado. Este resultado é proveniente da aplicação das métricas linha a linha de acordo com os parâmetros descritos do Tópico “4.2.2 - Coleta de dados” no conjunto dos dados.

**Quadro 13 - Tamanho da amostra por campo e dados pontuados**

<b>Nome do campo</b>	<b>Tamanho da amostra</b>	<b>Pontuado</b>
Data de Abertura	425	425
Vendedor	425	425
Setor	425	423
Motivo do Chamado	425	425
Código do cliente	425	418
Nome do Cliente	425	425
CPF/CNPJ	425	404
E-mail	425	402
Data de nascimento do cliente	425	420
Telefone	425	419
Atenção para a alteração	15	6
Observação	25	16
Adição de Imagens	15	2
E-mail de aprovação	425	364
Número do chamado	425	425
Status do Chamado	425	425
Pontos de atenção	133	125
Aberto por	425	419
Data de Finalização	288	284

Fonte: Elaborado pelo autor.

Cada um dos campos calculados de acordo com a premissa da métrica gerou uma pontuação. Cada pontuação é referente ao número de pontos que apresentaram qualidade da informação. Por exemplo: todos os 425 casos coletados relativos ao campo “Data de Abertura” estavam dentro do parâmetro estabelecido pela meta. Logo foram pontuados para este campo 425 itens de 425 itens considerados. O campo Adição de Imagem tem 15 dados que devem ser analisados devido as restrições da métrica para considerar ou não um dado para avaliação desta métrica. De 15 dados 2 foram pontuados, o seja, 2 dados têm qualidade ao passo que 13 não possuem qualidade.

#### 4.3.2 Avaliação percentual

Considerando o total de dados por campo e o total de dados coletados, é possível avaliar percentualmente o nível de acuracidade dos dados de cada campo analisado. Por exemplo: o campo “Data de Abertura” teve 100% dos dados com qualidade pois o total de dados pontuados sobre o total de dados coletados foi o mesmo ( $425/425 = 100\%$ ). O campo Adição de Imagem obteve uma pontuação igual 13,3% pois dos 15 dados considerados 2 pontuaram ( $2/15=13,3\%$ ).

Sabendo-se o percentual (%) alcançado por campo é possível avaliar a qualidade da informação que ele se enquadra. Essa avaliação foi feita de acordo com parâmetros estabelecidos pelos gestores da empresa em conjunto com o pesquisador (já explicada na Etapa “4.2.2 -Coleta de dados”). Exemplo: os dados coletados no campo Data de abertura estão bons pois obtiveram uma avaliação percentual acima de 95%. Já os dados coletados para o campo Adição de imagem estão ruins pois obtiveram uma pontuação abaixo de 90%.

O Quadro 14 demonstra quanto cada campo avaliado obteve de pontuação percentual de acordo com a lógica da métrica associada à avaliação dos dados. A informação boa (maior de 95%) está em verde, a média (entre 90% e 95%) está em amarelo e a ruim (menor que 90%) está em vermelho.

**Quadro 14 - Avaliação de percentual da qualidade da informação por campo**

Nome do campo	%
Data de Abertura	100,0%
Vendedor	100,0%
Setor	99,5%
Motivo do Chamado	100,0%
Código do cliente	98,4%
CPF/CNPJ	95,1%

Continua

	Conclusão
E-mail	94,6%
Data de nascimento do cliente	98,8%
Telefone	98,6%
Atenção para a alteração	40,0%
Observação	64,0%
Adição de Imagens	13,3%
E-mail de aprovação	85,6%
Número do chamado	100,0%
Status do Chamado	100,0%
Pontos de atenção	94,0%
Aberto por	98,6%
Data de Finalização	98,6%
>95%	Boa
Entre 90% e 95%	Média
<90%	Ruim

Fonte: Elaborado pelo autor.

Cada um dos campos calculados de acordo com a premissa da métrica gerou um percentual de aproveitamento. A avaliação demonstra que dos 19 campos analisados 13 foram classificados como boa a qualidade da informação por apresentar um percentual acima de 95%. Dois campos (E-mail e Pontos de atenção) ficaram com uma avaliação média pois apresentaram um resultado entre 90% e 95%. Quatro campos (Atenção para a alteração, Observação, Adição de Imagens e E-mail de aprovação) foram classificados como ruim a qualidade da informação. Cabe ressaltar que “Atenção para a alteração”, “Observação” e “Adição de Imagens” ficaram com um percentual abaixo de 65% e E-mail de aprovação ficou com um percentual de 85,6%.

#### 4.3.3 Consolidação das avaliações

Cada um dos campos avaliados possui uma relação com uma dimensão da qualidade da informação. Esta dimensão foi definida junto aos gestores por meio da análise das entrevistas feitas junto aos gestores explicada com mais detalhes na estrutura conceitual teórica.

Para quantificar as quatro dimensões selecionadas (temporalidade, relevância, reputação e quantidade de apropriados) será utilizada a somatória da pontuação de cada campo relacionada a aquela dimensão dividida pelo somatório do tamanho das amostras de cada campo avaliados daquela dimensão. Exemplo: a dimensão temporalidade possui dois campos avaliados (Data de abertura e Data de finalização). O tamanho da amostra para esta

dimensão é 425 mais 288 totalizando 713. A positividade destes dois campos somados foi de 709. O percentual de qualidade a informação para a dimensão foi de 99,4% (709/713). A Tabela 15 demonstra o resultado que cada dimensão atingiu.

**Quadro 15** - Percentual de qualidade da informação por dimensão

Dimensão	%	Tamanho da amostra
Temporalidade	99,4%	713
Relevância	98,0%	3.825
Reputação	96,8%	1.700
Quantidade de dados apropriados	79,3%	188

Legenda: qualidade da informação

>95%	Boa
Entre 90% e 95%	Média
<90%	Ruim

Fonte: Elaborado pelo autor.

Três dimensões (temporalidade, relevância e reputação) possuíram uma avaliação muito satisfatória quanto ao desempenho da qualidade da informação. Elas foram classificadas como boas dentro da classificação alinhada entre pesquisador e a empresa alvo do estudo. A melhor avaliação foi da dimensão temporalidade com uma aderência as regras das métricas de 99,4% seguidas de relevância com 98,0% e reputação com 96,8%. Uma dimensão (quantidade de dados apropriados) ficou com uma avaliação considerada ruim com um aproveitamento de 79,3%.

#### 4.3.4 Avaliação geral da qualidade

A avaliação geral da qualidade da informação objetiva dar uma visão geral para empresa relacionada a qualidade dos dados para o processo avaliado de uma maneira geral. Deve somar o total pontuado por dimensão e subtraído pelo tamanho da amostra.

O total de pontos obtidos nas dimensões foi de 6.252 e a somatória das amostras foi de 6.426. O percentual alcançado foi de 97,3%. A avaliação geral da qualidade da informação para processo de solicitação de modificação de cadastro do cliente foi considerada bom pois foi mais do que 95%.

#### 4.3.5 Seleção das dimensões com percentual abaixo

Apesar do indicador geral da qualidade da informação sinalizar que ela está boa, é importante verificar que ainda existem oportunidades de melhoria. Esta evidência pode ser

coletada ao analisar que a dimensão “quantidade de dados apropriados”. Ela está com uma qualidade considerada ruim com um percentual de 79,3%. O Quadro 16 explicita os campos associados a dimensão qualidade de dados apropriado. As dimensões são: atenção para a alteração, observação, adição de imagens e pontos de atenção.

**Quadro 16 - Campos associados a dimensão quantidade de dados apropriados**

<b>Campo</b>	<b>%</b>	<b>Tamanho da amostra</b>	<b>Dimensão</b>	<b>Pontuado</b>
Atenção para a alteração	40,0%	15	Quantidade de dados apropriados	6
Observação	64,0%	25	Quantidade de dados apropriados	16
Adição de Imagens	13,3%	15	Quantidade de dados apropriados	2
Pontos de atenção	94,0%	133	Quantidade de dados apropriados	125

Fonte: Elaborado pelo autor.

Dos quatro campos avaliados, três (atenção para a alteração, observação e adição de imagens) possuíram uma qualidade da informação considera ruim com um grau de aproveitamento que variou entre 64,0% e 13,3%. Um campo (ponto de atenção) foi pontuado como médio alcançando um aproveitamento de 94,0%. Considerando que dos quatro campos um está com um aproveitamento muito próximo de bom ele será descartado para esta rodada de melhoria da qualidade da informação.

Os campos selecionados para entender a causa raiz e determinar ações de melhoria serão os que foram classificados com uma qualidade da informação ruim (abaixo de 90%): atenção para a alteração, observação e adição de imagens. Ao analisar os três campos que apresentaram qualidade da informação ruim: (i) Atenção para a alteração com 40,0%; (ii) Observação com 64,0% e (iii) Adição de Imagens com 13,3% para a dimensão “quantidade de dados apropriados” eles foram os principais ofensores pela baixa qualidade da informação nesta. As métricas associadas a avaliação dos dados destes campos possuem regras mais restritas no seu uso e são informações coletadas em campo.

Estes campos têm uma avaliação menos frequente visto o número de restrições que as métricas possuem para serem consideradas para avaliar. Isso fica evidente ao analisar o tamanho da amostra destes 3 campos: 15 para “Atenção para a alteração”, 25 para “Observação” e 15 para “Adição de Imagens”. A frequência de dados avaliados é baixa frente ao número total de formulários analisados que foi de 425 formulários. Apesar da baixa relevância no contexto total da qualidade da informação do processo de solicitação do cadastro do cliente, estes campos são os maiores ofensores na dimensão “quantidade de dados

apropriados”. De 188 dados considerados para avaliar o indicador, estes três campos juntos correspondem a 55 dados, ou seja, 29,3%.

#### 4.4 ETAPA MELHORAR

Está etapa irá propor ações para melhorar a qualidade da informação de dimensões que foram classificadas com um baixo nível de qualidade. As ações, análises e direcionamentos frente ao problema da baixa qualidade da informação foram feitas com a participação da equipe de gestão, suporte comercial e o pesquisador. Após a conclusão das análises dos dados marcou-se uma reunião com os gestores e vendedores da área. Foram demonstrados os principais pontos do estudo, compartilhadas ideias e percepções qualitativas a fim compreender as causas da baixa qualidade de dados para a dimensão “quantidade de dados apropriados”.

Os motivos dos solicitantes de modificação do cadastro dos clientes, os coletores de dados, não preencherem os campos com pior avaliação (atenção para a alteração, observação e adição de imagens) de forma correta foram:

- São campos com pouca frequência de uso;
- São campos com regras muito específicas;
- O formulário permite que os dados sejam preenchidos desta forma.

Além dos pontos elencados anteriormente, eles citaram também que nem sempre os vendedores sabem a relação de causa e efeito do preenchimento errôneo de uma informação. O processo não é claro para todos os envolvidos. Os gestores elencaram também, que não possuem um acompanhamento da qualidade da informação. Sem este acompanhamento eles não tem visibilidade de como está a qualidade da informação nem tem embasamentos numéricos para fazer cobranças embasadas em números. Esta falta de controle também foi uma das causas elencadas para que a dimensões “quantidade de dados apropriada” estivesse ruim. Segundo os gestores, um acompanhamento seria crucial para a melhoria da qualidade da informação e desenvolvimento da equipe que faz as solicitações por meio do formulário. A partir da análise dos dados e reunião com a equipe comercial surgiram as seguintes ações:

- Criar regras no formulário para impedir o envio da informação de acordo com as métricas de qualidade da informação estabelecidas. O formulário deve reconhecer as regras das métricas para impedir que dados fora daquele conjunto de regras sejam inseridos no sistema;

- Construir um acompanhamento e controle da qualidade da informação. Fazer um painel (*dashboard*) para que os gestores tenham controle do status da qualidade da informação à medida que a informação é coletada e possibilite compartilhar com o time comercial oportunidades de melhoria por meio da análise deste painel;
- Treinar ao time comercial (coletores dos dados). Para os itens medidos com baixa, será feito um treinamento focado na melhoria destes itens junto a equipe que faz a coleta de dados.

## 5 RESULTADOS

Este tópico destina-se a apresentar os resultados da pesquisa. Ele apresentará a comparação entre os resultados obtidos no ciclo de melhoria do TDQM em comparação com os dados colhidos na etapa definir, ou seja, será representado um antes e depois. Serão descritos também os resultados da aplicação das ações propostas no Tópico “4.4 – Melhorar”. Serão elencadas as regras criadas no formulário para mitigar problemas de qualidade da informação antes dessa informação serem armazenadas no banco de dados. Será demonstrado o novo processo de fluxo que coleta de dados que incluirá a avaliação automática da qualidade da informação que dará suporte aos gestores acompanharem e analisar a qualidade da informação das solicitações de mudança de cadastro em tempo real. Por fim, serão descritas as atividades de treinamento focada na melhoria da qualidade da informação.

### 5.1 COMPARATIVO DA QUALIDADE DA INFORMAÇÃO

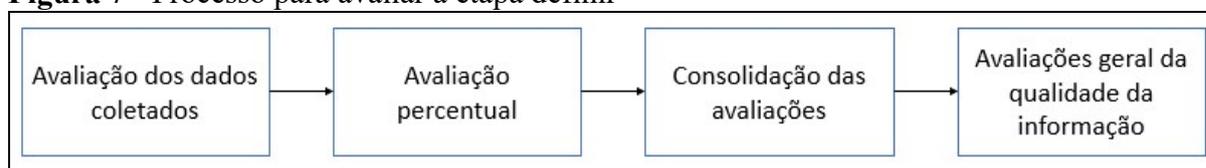
Este tópico irá demonstrar as melhorias obtidas na qualidade da informação após a aplicação do TDQM. A qualidade da informação da Etapa 4.1 (definir) foi medida utilizando as mesmas métricas e dimensões utilizadas para medir o novo processo. Dentre os 19 campos avaliados no novo processo, 14 estão presentes no processo antigo: data de abertura, vendedor, setor, motivo do chamado, código do cliente, CPF/CNPJ, e-mail, data de nascimento do cliente, telefone, e-mail de aprovação, número do chamado, status do chamado e aberto por. Visando os campos que serão avaliados, as dimensões Temporalidade, Relevância a Reputação também serão avaliados consequentemente.

Cinco campos não serão feitas comparações, pois eles não existiam no processo antes da aplicação do TDQM. Os 5 campos são: atenção para a alteração, observação, adição de imagens, e-mail de aprovação, pontos de atenção e data de finalização. Como a dimensão quantidade de dados apropriados está ligada aos campos atenção para a alteração, observação, adição de imagens e pontos de atenção, estas dimensões não poderão ser verificadas se houve ou não melhoria da qualidade da informação. O campo data de finalização está relacionada a dimensão temporalidade. Desta forma, para esta comparação da dimensão temporalidade, será utilizada somente a data de abertura como critério de comparação.

Para avaliação da qualidade da informação da Etapa 4.1 (definir) foram utilizadas as mesmas etapas utilizadas para avaliação da qualidade da informação no TDQM, porém sem a necessidade de construção de um plano de ação para a melhoria da qualidade da informação uma vez que esta avaliação visa comprovar a qualidade da informação preliminar versus os

resultados alcançados com a aplicação TDQM. A Figura 7 resume os processos de avaliação da qualidade da informação da Etapa 4.1 (definir).

**Figura 7 -** Processo para avaliar a etapa definir



Fonte: Elaborado pelo autor.

No total foram coletados 1.156 dados para avaliação da qualidade a informação. Os 14 campos avaliados utilizaram o total da amostra para a avaliação da qualidade da informação. O Quadro 17 representa o tamanho da amostra associada a cada campo avaliado e o valor pontuado representa o número de campos com informação consideradas boas de acordo com o total avaliado. Cada um dos campos gerou um percentual relativo ao número de pontos obtidos por cada campo dividido pelo tamanho da amostra. Assim o critério de avaliação percentual será o mesmo que foi utilizado no TDQM para manter os mesmos padrões de avaliação.

**Quadro 17 –** Avaliação qualidade da informação por campo da etapa definir

% de informação com qualidade	Tamanho da amostra	Pontuado	Aproveitamento % etapa definir
Data de Abertura	1.156	1.124	97,2%
Vendedor	1.156	1.061	91,8%
Setor	1.156	1.117	96,6%
Motivo do Chamado	1.156	1.153	99,7%
Código do cliente	1.156	1.149	99,4%
Nome do Cliente	1.156	858	74,2%
CPF/CNPJ	1.156	367	31,7%
E-mail	1.156	330	28,5%
Data de nascimento do cliente	1.156	3	0,3%
Telefone	1.156	435	37,6%
E-mail de aprovação	1.156	557	48,2%
Número do chamado	1.156	1.147	99,2%
Status do Chamado	1.156	16	1,4%
Aberto por	1.156	631	54,6%

>95%	Boa
Entre 90% e 95%	Média
<90%	Ruim

Fonte: Elaborado pelo autor.

Os dados demonstram que a qualidade da informação representada em vermelho (nome do cliente, CPF/CNPJ, e-mail, data de nascimento do cliente, telefone, e-mail de aprovação, status do chamado e aberto por), apresentaram uma avaliação muito ruim. Já os campos data de abertura, vendedor, setor, motivo do chamado, código do cliente e número do chamado apresentaram uma avaliação acima de 90%. Estes dados representam que o processo da Etapa 4.1 (definir) apresentavam muitas oportunidades de melhoria na qualidade da informação. Por meio da avaliação destes campos foi possível compreender como estava a avaliação por dimensão da qualidade da informação representada no Quadro 18.

**Quadro 18** – Avaliação das dimensões da qualidade da informação

Dimensão	Etapa definir	Tamanho da amostra
Temporalidade	97,2%	1.156
Relevância	73,9%	10.404
Reputação	24,5%	4.624

Legenda: qualidade da informação

>95%	Boa
Entre 90% e 95%	Média
<90%	Ruim

Fonte: Elaborado pelo autor.

As dimensões relevância e reputação obtiveram uma avaliação ruim, obtendo pontuações inferiores a 90%. A dimensão Reputação obteve uma avaliação boa de acordo com os critérios que consideram a qualidade da informação boa ou ruim no estudo. A qualidade da informação dos dados coletados na Etapa 4.1 (definir) utilizando as métricas construídas durante o processo de melhoria do TDQM foi de 61,5%. Esse baixo número se deve ao fato de que durante a Etapa 4.1 (definir) não foram estipuladas métricas para avaliação da qualidade da informação em nenhum momento. Logo, esta avaliação trouxe um parâmetro para verificar se a qualidade da informação melhorou ou não considerando um antes e depois da aplicação do método do TDQM para melhoria da qualidade da informação.

O Quadro 19 representa a comparação percentual da qualidade da informação entre a Etapa 4.1 (definir), antes, e o resultado após o ciclo de melhoria do TDQM, depois.

**Quadro 19** – Comparação da qualidade da informação.

% de informação com qualidade	Depois TDQM	Antes Etapa Definir	Diferença percentual (pesquisa TDQM x Etapa Definir)
Data de Abertura	100,0%	97,2%	2,8%

Continua

			Conclusão
Vendedor	100,0%	91,8%	8,2%
Setor	99,5%	96,6%	2,9%
Motivo do Chamado	100,0%	99,7%	0,3%
Código do cliente	98,4%	99,4%	-1,0%
Nome do Cliente	100,0%	74,2%	25,8%
CPF/CNPJ	95,1%	31,7%	63,3%
E-mail	94,6%	28,5%	66,0%
Data de nascimento do cliente	98,8%	0,3%	98,6%
Telefone	98,6%	37,6%	61,0%
Atenção para a alteração	40,0%	"Não avaliado"	Não Avaliado
Observação	64,0%	"Não avaliado"	Não Avaliado
Adição de Imagens	13,3%	"Não avaliado"	Não Avaliado
E-mail de aprovação	85,6%	48,2%	37,5%
Número do chamado	100,0%	99,2%	0,8%
Status do Chamado	100,0%	1,4%	98,6%
Pontos de atenção	94,0%	"Não avaliado"	Não Avaliado
Aberto por	98,6%	54,6%	44,0%
Data de Finalização	98,6%	"Não avaliado"	Não Avaliado

Legenda: qualidade da informação

>95%	Boa
Entre 90% e 95%	Média
<90%	Ruim

Fonte: Elaborado pelo autor.

Os resultados demonstram que para a grande maioria dos casos a qualidade da informação teve melhorias. Apenas para o campo, “código do cliente” houve uma pequena redução da qualidade, porém ainda é considerada boa de acordo o percentual alcançado de acurácia maior que 95% considerada uma boa qualidade a informação. Os Tópicos 5.2 e 5.3 irão expor as ações tomadas para melhorar a qualidade da informação.

Por meio da Etapa 4.1 (definir) foi possível fazer esta comparação entre a qualidade da informação antes e depois da aplicação do ciclo completo do TDQM. Como alguns campos eram similares, estes foram comparados por meio dos critérios criados para avaliação da qualidade da informação ao longo do processo indicado pelo TDQM. Para comparar as dimensões foi utilizado o mesmo critério de considerar na contagem de dados coletados e dados pontuados somente os campos semelhantes em cada um dos estudos para que a avaliação e critérios fossem os mesmos dentro da comparação feita neste estudo. O Quadro 20 demonstra esta comparação considerando as premissas elencadas anteriormente.

**Quadro 20** - Comparação antes e depois do TDQM

Dimensão	Depois TDQM	Antes Etapa definir	Diferença percentual TDQM versus Etapa definir
Temporalidade	100,0%	97,2%	2,8%
Relevância	98,0%	73,9%	24,1%
Reputação	96,8%	24,5%	72,2%

Fonte: Elaborado pelo autor.

É possível perceber uma evidente melhoria na qualidade da informação considerando os campos e conseqüentemente as dimensões em comum entre os modelos. Destaca-se a expressiva melhora da qualidade da informação nas dimensões relevância e reputação apresentando evoluções de 24,1% e 72,2% respectivamente.

De uma forma geral, a pesquisa, utilizando as métricas criadas na metodologia do TDQM em comum a Etapa 4.1 (definir), apresentou um resultado geral de 97,8% frente a 61,5% da Etapa 4.1 (definir). É possível afirmar que o processo de melhoria da qualidade de informação cresceu 36,3% pontos percentuais e uma evolução de 59,1% comparando o processo o estado inicial e o final. Logo, as mudanças feitas nos processos trouxeram melhoria significativas para a qualidade da informação e maturidade do processo visto.

O Quadro 21 faz uma comparação entre o TDQM e a Etapa 4.1 (definir) no que se refere a melhoria do processo qualitativo e de aprendizagem e demonstra o resultado geral das duas pesquisas. O “sim” significar que o existiu a etapa indicada no “processo qualitativo e de aprendizagem”, o “não” significa que a etapa não existiu. A porcentagem representa o percentual de qualidade da informação alcançado.

**Quadro 21** – Comparação entre TDQM e etapa definir

Processo qualitativo e de aprendizagem	TDQM	Etapa definir
Processos mapeados	Sim	Sim
Segue uma metodologia científica consagrada	Sim	Não
Coleta de dados por formulário	Sim	Não
Tabulação de dados automática	Sim	Não
Métricas definidas para cálculo da qualidade da informação	Sim	Não
Dimensões da qualidade da informação	Sim	Não
Avaliação da qualidade da informação	Sim	Sim
Analísado com objetivo de mudar para melhorar	Sim	Sim
Proposição e implementação de plano de ação	Sim	Sim
Trouxe melhorias e novos conhecimentos para a empresa	Sim	Sim
Resultado da avaliação da qualidade utilizando somente os campos em comum entre etapa definir e TDQM	97,8%	61,5%

Fonte: Elaborado pelo autor.

Uso do TDQM agregou novos processos para a atividade de avaliação da qualidade da informação e agregaram conhecimento aos participantes/empresa. Um ponto muito importante nesta avaliação é que a etapa de definir serviu como motivador para que fosse feita uma pesquisa mais aprofundada sobre qualidade da informação. O TDQM teve papel fundamental para o desenvolvimento do projeto.

## 5.2 MITIGAÇÃO DE PROBLEMAS DA QUALIDADE DA INFORMAÇÃO

Outro resultado relevante para a pesquisa foi a utilização de regras diretamente aplicadas no formulário. Estas regras visam garantir a qualidade pois ela não espera o processo acontecer para identificar o erro antes da entrega da informação no banco de dados. Para os campos que apresentaram pior avaliação foram criadas regras específicas no formulário para evitar a entrada de informação com baixa qualidade ou o campo foi excluído. A seguir serão detalhadas as ações aplicadas diretamente no formulário que os coletores de dados preenchem.

Nome do campo: “Atenção para a alteração de Canal/Tabela/Correção de Cadastro é necessário o envio de três fotos do local (balcão para fora, por dentro e por fora) e incluir no chamado. É o seu caso?”. Ação: este campo foi excluído do formulário pois, após um alinhamento com os gestores e usuários do formulário chegou-se à conclusão que ele é um campo condicional manual. Se o indivíduo que preencher o formulário com a intenção de alterar “Canal”, “Tabela” ou corrigir cadastro deve seguir regras estabelecidas por cada requisição. Logo colocou-se como obrigatoriedade a inclusão de fotos para mudança de “Canal”, “Tabela” ou “Cadastro” a inclusão de imagens e texto na observação detalhando o que deve ser feito. A não inclusão não permite o envio do formulário.

Nome do campo: “Observação”. Ação: foi criada uma regra condicional baseada na métrica deste campo (“Campo preenchido obrigatoriamente caso o motivo da solicitação seja "motivo do chamado seja "Alteração de canal (detalhar no campo observação a mudança)”) Caso a regra não seja obedecida, o formulário não poderá ser enviado.

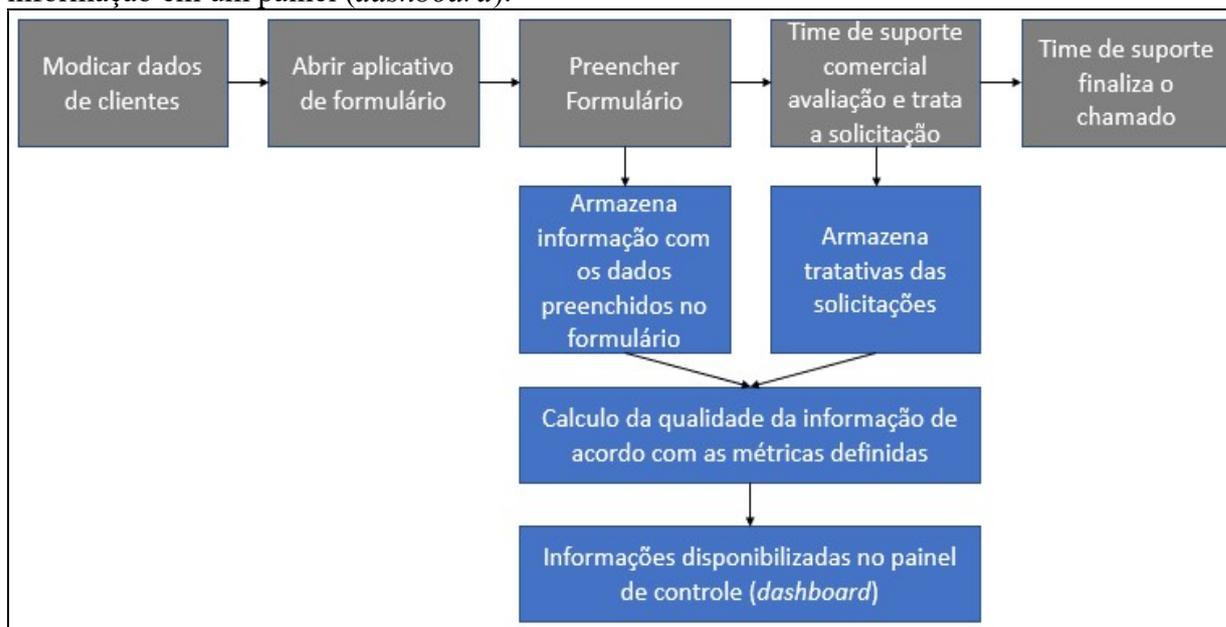
Nome do Campo: “Adição de Imagens”. Ação: foi criada uma regra condicional baseada na métrica deste campo (“Campo preenchido caso o Cod\_Cliente inicie com 1 e o motivo da solicitação seja "Alteração de canal (detalhar no campo observação a mudança)" ou "correção de cadastro de clientes (detalhar no campo observação)”). Caso a regra não seja obedecida, o formulário não poderá ser enviado.

A expectativa é que estas regras melhorem a qualidade da informação dos campos “Observação” e “Adição de Imagem”.

### 5.3 IMPLEMENTAÇÃO DE PROCESSO PARA VISUALIZAÇÃO DA QUALIDADE DA INFORMAÇÃO

Relativo à ação de: “construir um acompanhamento e controle da qualidade da informação. Fazer um painel (dashboard) para que os gestores tenham controle do status da qualidade da informação à medida que a informação é coletada e possibilite compartilhar com o time comercial oportunidades de melhoria por meio da análise deste painel”. Foram incluídos processos de cálculos baseados nas métricas estabelecidas para que os campos fossem avaliados automaticamente a partir dos dados inseridos nos bancos de dados. Estes cálculos foram necessários, pois sem eles não seria possível construir um painel que fosse atualizado em tempo real para os gestores. A Figura 8 representa o processo completo de fluxo de dados após a ação incluindo o processo de ida dos dados para o painel (*dashbaod*).

**Figura 8** - Processos do fluxo de dados desde a coleta até a tratativa e exposição da informação em um painel (*dashboard*).



Fonte: Elaborado pelo autor.

O novo fluxo criado proporcionou a possibilidade de cálculos automáticos das métricas para cada campo da qualidade da informação e proporcionou a exposição destes dados em um painel. A expectativa é que os gestores tenham maior visibilidade da qualidade da informação e possam tomar decisões mais assertivas quanto a ações para melhorar continuamente a qualidade da informação.

#### 5.4 PROGRAMAÇÃO DE TREINAMENTOS

Finalizando os resultados do plano de ação proposto na etapa melhorar do TDQM, não é possível afirmar que os treinamentos trarão um resultado efetivo para a melhoria da qualidade da informação, porém segundo os gestores, os treinamentos trouxeram ganhos qualitativos para a equipe. Os mais relevantes serão elencados a seguir:

- Compreensão dos processos de coleta de dados: o time comercial passou a entender melhor o fluxo da informação. Para onde vão as solicitações de mudança de cadastro, quem são os responsáveis, os tempos entre os processos e os envolvidos;
- Qualidade da informação é vital: os treinamentos também trouxeram uma visão de causa e consequência, ou seja, se o cadastro estiver errado, o cliente não recebe, a venda não é feita. O responsável por isso são os coletores de dados, a informação com qualidade é vital para que eles tenham sucesso na jornada de vendas;
- Mudança de mentalidade: como consequência de um entendimento melhor dos processos e do impacto da qualidade da informação nas vendas, o time comercial passou a ser mais criterioso na coleta dos dados. Passou a prover feedbacks mais frequentes aos gestores sobre a necessidade de ter um base de cliente com informações mais assertivas.

#### 5.5 LIÇÕES APRENDIDAS E RESUMO DAS ETAPAS DO TDQM

O trabalho executado na empresa alvo foi concluído com êxito. De acordo com os resultados obtidos foi possível demonstrar que houve melhorias tanto qualitativas quanto quantitativas. As melhorias qualitativas mais perceptíveis estão relacionadas ao fluxo, coleta, armazenamento, gestão dos dados, criação de métricas para avaliação da qualidade da informação e sistemas preventivos de coleta de dados com informação ruim.

A melhorias quantitativas foram perceptíveis ao comparar os resultados da qualidade da informação obtidos na Etapa 4.1 (definir) (61,5%) versus a pesquisa utilizando o TDQM (97,8%).

As principais lições aprendidas ao longo do projeto estão relacionadas às necessidades de criar estruturas e ambientes propícios para avaliar a qualidade da informação. Neste processo, isto iniciou-se com a conscientização dos gestores do negócio e disseminação para o restante da organização.

A abordagem inicial na Etapa 4.1 (definir) foi importante para que obter resultados rápidos para a organização do trabalho e fez parte do processo de amadurecimento do conhecimento para instigar os envolvidos no processo a procurar maneiras de melhorar ainda mais os processos relativos à qualidade da informação. A utilização de uma metodologia

focada em qualidade da informação e ao mesmo tempo em melhorar processos, no caso o TDQM, foi importante para mudar o ponto de vista da organização frente a avaliação, processo de coleta de dados, gestão e armazenamento dos dados.

As principais dificuldades encontradas foram na adaptação dos coletores da dados para entender o funcionamento do formulário construído e compreender a necessidade de passar as informações de uma forma padronizada. Para que esse ponto não fosse impeditivo para a execução do projeto, a influência dos gestores foi essencial para o sucesso da implantação do formulário. O Quadro 22 resume o que foi feito ao longo processo de melhoria da qualidade da informação.

**Quadro 22 - Resumo das etapas do projeto**

<b>Etapas</b>	<b>Período</b>	<b>Objetivo principal</b>	<b>Critério de Avaliação</b>	<b>Atividades executadas</b>	<b>Avaliação</b>	<b>Melhoria e aprendizagem</b>
Definir	Nov 2018 a ago 2019	Organizar processo de coleta de dados e verificar as causas mais frequentes de necessidade de alteração do cadastro dos clientes e definir dimensões da qualidade da informação.	Definição de principais causas de solicitações de modificação de cadastro de clientes, dimensões da qualidade da informação definidas e partes interessadas no projeto de acordo com as definições.	Definição do escopo inicial do projeto da empresa, coleta e registro de dados em planilha, identificação dos envolvidos no processo, mapeamento de processos, entrevista, criação de estrutura preliminar para segunda onda de coleta de dados e definição das dimensões da qualidade da informação.	Processo de registro e coleta de informação melhorou; processo ficou melhor definido, qualidade da informação ainda pode melhorar, produto de informação necessita ser mais bem estruturado.	Estudar metodologias que avaliem a qualidade da informação e que conduzam as etapas do trabalho de forma mais organizada. Melhorar a coleta e avaliação dos dados. Aplicar padronizações e ideias de melhoria para outros projetos.
Medir, análise e melhorar	Ago 2018 a Dez 2020	Melhorar a qualidade da informação no processo de solicitação de mudança de cadastro dos clientes.	Comparar melhoria da qualidade da informação com processo anterior, definir métricas para avaliar a qualidade da informação, analisar os dados e traçar ações.	Definição de necessidades do consumidor da informação; Construção de processo para coleta de dados e formulário; Definição de métricas; Análise dos dados; Construção de plano de Ação; Comparação do antes e depois.	A qualidade da informação melhorou em comparação com o processo anterior considerando as métricas e dimensões criadas a partir da necessidade do usuário. O estudo identificou novos pontos que não estavam sendo avaliados que propiciaram uma visão mais clara a respeito dos dados coletados, criação de métricas e uso das dimensões da qualidade da informação.	O estudo apontou pontos de melhoria do formulário, treinamentos necessários para a equipe coletora e a necessidade de acompanhar de perto a qualidade da informação por meio de um painel. O trabalho ajudou a avaliar a qualidade da informação de uma forma científica e com um procedimento metodológico adequado a realidade da organização. Cabe ressaltar que a etapa 4.1 (definir) serviu de gatilho para demonstrar a empresa como a qualidade da informação estava. Sendo assim, uma etapa primordial para instigar a equipe envolvida no projeto a melhorar os processos relacionados a qualidade da informação aprofundar no tema do trabalho.

Fonte: Elaborado pelo autor.

## 6. CONCLUSÕES

O trabalho executado com foco na melhoria da qualidade da informação aplicado a alteração do cadastro de clientes, demonstrou a importância da criação de ambientes e processos que favoreçam a qualidade da informação. O trabalho trouxe uma perspectiva de aplicação prática e acadêmica que evidenciou um processo de amadurecimento processual e conceitual relacionados a melhoria da qualidade informação aplicada em um ambiente prático.

No que se refere ao objetivo geral do trabalho, houve melhoria na qualidade da informação visto que o índice criado para medir a qualidade da informação passou de 61,5% para 97,8%. Essa melhoria foi perceptível pois foram criadas métricas e definidas dimensões que possibilitaram a comparação dos resultados antes e depois do ciclo do TDQM.

Quanto aos objetivos específicos, todos foram alcançados de acordo com os tópicos abaixo:

- Mapear os processos de coleta de dados e fluxo dos dados: os processos foram mapeados em todas as etapas proporcionando visibilidade do trabalho ao longo do ciclo do TDQM;
- Padronizar a maneira como os dados são colhidos em campo: os dados passaram a ser coletados e padronizados por meio de um sistema de coleta de dados on-line em um formulário;
- Reduzir as interações informais: as interações informais foram reduzidas pois os coletores de dados receberam treinamento sobre o processo de coleta de dados e passaram utilizar o formulário de maneira autônoma;
- Medir a qualidade da informação coletada em campo: a criação das métricas em conjunto com a padronização de um formulário e processos de coleta dos dados trouxe a possibilidade de medir adequadamente a qualidade da informação dos dados coletados em campo;
- Utilizar a metodologia TDQM (*Total Data Quality Management*): TDQM foi utilizada de forma satisfatória proporcionando contribuições práticas e contribuições científicas.

A metodologia TDQM atendeu de maneira satisfatória a necessidade de pesquisa e da empresa sendo uma ela de fácil compreensão para leigos, na opinião dos envolvidos na pesquisa. Ela ajudou a melhorar e criar processos e métricas que impactaram positivamente na melhoria da qualidade da informação e ajudaram a direcionar os envolvidos a criarem planos de ação direcionados a partir dos resultados extraídos das análises dos dados coletados. O uso

do TDQM ajudou a criar métricas para avaliar quantitativamente a qualidade da informação e auxiliou a categorizar cada uma das métricas dentro de dimensões da qualidade da informação.

A etapa de definir foi essencial para repensar a forma como a empresa estava coletando e avaliando a qualidade da informação. Ela incentivou os envolvidos no processo a buscarem mais conhecimento a respeito da qualidade da informação, o que facilitou muito a aplicação completa do ciclo TDQM. Ela foi importante também para que fosse identificada uma necessidade real de aplicar uma metodologia de pesquisa em um ambiente prático e que esta pesquisa formal trouxe uma visão de avaliação da qualidade da informação não abordada anteriormente.

Os processos que envolveram o TDQM foram essenciais para repensar a forma como a empresa estava coletando, avaliando a qualidade da informação e gerindo os processos. Isso foi percebido ao comparar a qualidade da informação utilizando um processo formalizado, o TDQM. Houve evolução de 59,1% da qualidade da informação comparando os dados coletados na etapa de definir em comparação com o resultado obtido após a aplicação do ciclo completo do TDQM. É importante frisar que o processo do TDQM foi o responsável pela criação de métricas, sem elas, não haveria a possibilidade de fazer a comparação de um antes e depois. Esta abordagem foi interessante pois pode ser utilizada, como um incentivador para que mais projetos práticos passem a ter suporte científico e contribuir de uma forma mais ampla na sociedade.

O trabalho trouxe como contribuição científica uma visão de como outros autores trabalharam na seleção das dimensões expondo os processos utilizados para criar os modelos para avaliar a qualidade da informação e as dimensões mais recorrentes. Este propiciou também compreender que os contextos são cruciais para definição de critérios e medidas para avaliar a qualidade da informação justificando-se os múltiplos caminhos fazer avaliações. Além disso, proporcionou a criação de um fluxograma para análise de pontos críticos para avaliação da qualidade da informação mesmo ela estando de uma maneira geral boa frente aos padrões da organização avaliada.

Para trabalhos futuros, sugere-se identificar problemas de qualidade da informação que não só os consumidores da informação estejam expostos, mas também, que seja um problema que incomode mais pessoas ou áreas envolvidas em um processo que esta informação seja importante e não esteja boa. Desta forma, mais pessoas poderão estar comprometidas com a melhoria da qualidade da informação. Outros pontos relevantes para trabalhos futuros é expor para o público leigo as vantagens de quantificar a qualidade da

informação, por meio da criação métricas, vínculo com dimensões da qualidade da informação para criação de indicadores e quantificar o quanto a organização alvo do estudo pode deixar de perder ou ganhar monetariamente tendo indicadores saldáveis de qualidade da informação.

## REFERÊNCIAS

- ANDREWS, Robert *et al.* *Quality-informed semi-automated event log generation for process mining.* **Decision Support Systems**, v. 132, p. 113265, 2020.
- ARDAGNA, D. *et al.* *Context-aware data quality assessment for big data.* **Future Generation Computer Systems**, v. 89, p. 548-562, 2018.
- BATINI, C. *et al.* *Methodologies for data quality assessment and improvement.* **ACM computing surveys (CSUR)**, v. 41, n. 3, p. 1-52, 2009.
- BATINI, C. *et al.* *Data and information quality.* **Cham, Switzerland: Springer International Publishing**, 2016.
- BATINI C., SCANNAPIECO M., *Data quality dimensions*, in: *Data Inf. Qual.*, Cham, Springer, 2016, pp. 21–51
- BOVEE, M.; SRIVASTAVA, R. P.; MAK, Brenda. *A conceptual framework and belief-function approach to assessing overall information quality.* **International journal of intelligent systems**, v. 18, n. 1, p. 51-74, 2003.
- BURGESS, M. S. E.; GRAY, W. A.; FIDDIAN, N. J. *Quality Measures and the Information Consumer.* *Proceedings International conference on information quality*, 9, 2004.
- CABALLERO, I.; MONJE, M. R.; SÁNCHEZ, C. M. F. *Calidad de datos digitales certificada.* **AENOR: Revista de la Normalización y la Certificación**, n. 331, p. 36-39, 2017
- CAI, L; ZHU, Y. *The challenges of data quality and data quality assessment in the big data era.* **Data Science journal**, v. 14, 2015.
- CARRETERO, A. G. *et al.* *MAMD 2.0: Environment for data quality processes implantation based on ISO 8000-6X and ISO/IEC 33000.* **Computer Standards & Interfaces**, v. 54, p. 139-151, 2017.
- CÔRTE-REAL, N.; RUIVO, P.; OLIVEIRA, T. *Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value?.* **Information & Management**, v. 57, n. 1, p. 103141, 2020.
- COUGHLAN, P.; COGHLAN, D. *Action research for operations management.* **International Journal of Operations & Production Management**, 2002.
- DESLAURIERS, J. P. *Recherche Qualitative- Guide pratique.* Montreal: McGraw-Hill, 1991.
- FAVARETTO, F. **Experimento Para Análise da Implantação da Medição da Qualidade da Informação.** Produção. Vol. 17, n. 1, p. 151-161, 2007.
- FAVARETTO, F.; VIEIRA, G. E. *Estudo Descritivo da Qualidade da Informação no Planejamento da Produção.* **Revista Gestão Industrial**, v.03 n. 02 p. 17-27, 2007.

FERREIRA, E.; FERREIRA, De. *Towards altruistic data quality assessment for mobile sensing. In: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers. 2017.* p. 464-469.

FERREIRA, Humberto; RUIVO, Pedro; REIS, Carolina. *How do data scientists and managers influence machine learning value creation?. Procedia Computer Science*, v. 181, p. 757-764, 2021.

FLECKENSTEIN, M.; FELLOWS, L. *Modern Data Strategy*. Springer, 2018.

FOX, C; LEVITIN, A.; REDMAN, T. *The notion of data and its quality dimensions. Information Processing & Management*, v. 30, n. 1, p. 9-19, 1994.

GE, M.; LEWONIEWSKI, W. *Developing the Quality Model for Collaborative Open Data. Procedia Computer Science*, v. 176, p. 1883-1892, 2020.

GHARIB, M.; GIORGINI, P. *Information quality requirements engineering with STS-IQ. Information and Software Technology*, v. 107, p. 83-100, 2019.

GIL, A. C. *et al. Como Elaborar Projetos de Pesquisa*. São Paulo: Atlas, 2002.

GUALO, Fernando *et al. Data quality certification using ISO/IEC 25012: Industrial experiences. Journal of Systems and Software*, v. 176, p. 110938, 2021.

GÜNTHER, Lisa C. *et al. Data quality assessment for improved decision-making: a methodology for small and medium-sized enterprises. Procedia Manufacturing*, v. 29, p. 583-591, 2019.

GÜRDÜR, D.; EL-KHOURY, J.; NYBERG, M. *Methodology for linked enterprise data quality assessment through information visualizations. Journal of Industrial Information Integration*, v. 15, p. 191-200, 2019.

HAMMER, M. *The process audit. Harvard Business Review*, v. 85, n. 4, p. 111, 2007.

HAZEN, B. T. *et al. Toward understanding outcomes associated with data quality improvement. International Journal of Production Economics*, v. 193, p. 737-747, 2017.

HEINRICH, B.; KLIER, M. *Metric-based data quality assessment—Developing and evaluating a probability-based currency metric. Decision Support Systems*, v. 72, p. 82-96, 2015.

ISO/IEC, 2015. **ISO/IEC 25012**. *Software Engineering — Software Product Quality Requirements and Evaluation (SQuaRE) — Data Quality Model*. ISO/IEC, International Standard.

ISO/IEC, 2015. **ISO/IEC 25024:2015** *Systems and Software Engineering – Systems and Software Quality Requirements and Evaluation (SQuaRE) – Measurement of Data Quality*. ISO/IEC, International Standard.

JANSSEN, M.; VAN DER VOORT, H.; WAHYUDI, A. *Factors influencing big data decision-making quality. Journal of Business Research*, v. 70, p. 338-345, 2017.

JEUSFELD, M. A.; QUIX, C.; JARKE, M. *Design and analysis of quality information for data warehouses. In: International Conference on Conceptual Modeling*. Springer, Berlin, Heidelberg, 1998. p. 349-362.

KAKAR, A. K. *Feature Selection for Evolutionary Commercial-Off-The-Shelf*. Tuscaloosa, AL: University of Alabama, 2013.

KWON, O.; LEE, N.; SHIN, B. *Data quality management, data usage experience and acquisition intention of big data analytics. International Journal of Information Management*, v. 34, n. 3, p. 387-394, 2014.

MACHADO, O. A. **Qualidade da Informação: Uma Abordagem Orientada para o Contexto**. São Paulo, SP: Universidade de São Paulo, 2013.

MAHANTI, Rupa. *Data Quality: Dimensions, Measurement, Strategy, Management, and Governance*. Quality Press, 2019.

MATTIODA, R. A. *et al. Determining the principal references of the social life cycle assessment of products. The International Journal of Life Cycle Assessment*, v. 20, n. 8, p. 1155-1165, 2015.

MATTIODA, R. A.; FAVARETTO, F. *Qualidade da Informação em Duas Empresas que Utilizam Data Warehouse na Perspectiva do Consumidor de Informação: um estudo de caso. Gestão & Produção*, v. 16, n. 4, p. 645-666, 2009.

MATTIODA, R. A.; FAVARETTO, F. *Qualidade de dados e produtos—uma analogia. 3º Congresso Nacional de Excelência em Gestão*, 2006.

MERINO G. J. *Environment for The Evaluation and Certification of Data Products Quality*, 2017.

MADHIKERMI, M. *et al. Data quality assessment of maintenance reporting procedures. Expert Systems with Applications*, v. 63, p. 145-164, 2016.

MCGILVRAY, D. *Executing Data Quality Projects: Ten Steps to Quality Data And Trusted Information™*. Elsevier, 2008.

NETO, C. B. *et al. Assessing quantity and quality of links between link data datasets. In: LDOW@ WWW*. 2016.

NAUMANN, F.; ROLKER, C. *Assessment Methods for Information Quality Criteria.*, 2000.

OLIVEIRA, S. A.; FAVARETTO, F. *Qualidade da Informação do Prontuário Eletrônico do Paciente no Processo de Apoio à Decisão Clínica. Journal of Health Informatics*, v. 13, n. 1, 2021.

- OLETO, R. R. **Percepção da qualidade da informação**. Ciência da informação, v. 35, 2006.
- PRADANOV, C. C.; FREITAS, E. C. **Metodologia do Trabalho Científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. 2ª Ed. – Novo Hamburgo: Feevale, 2013.
- PRICE, R.; SHANKS, G. *A semiotic information quality framework: development and comparative analysis*. In: **Enacting Research Methods in Information Systems**. Palgrave Macmillan, Cham, 2016. p. 219-250.
- PRICE, R.; SHANKS, G. *Empirical Refinement of a Semiotic Information Quality Framework*. **Proceedings of 38th Hawaii International Conference on System Sciences**, Hawaii, p. 1-10, 2005.
- REDMAN, T. C. **Data Quality for The Information Age**. Norwood, Mass.: Artech House. 1996.
- REDMAN, T. C. *The impact of poor data quality on the typical enterprise*. **Communications of the ACM**, v. 41, n. 2, p. 79-82, 1998.
- REIMER, A. P.; MILINOVICH, A.; MADIGAN, E. A. *Data quality assessment framework to assess electronic medical record data for use in research*. **International Journal of Medical Informatics**, v. 90, p. 40-47, 2016.
- SETIA, P. *et al. Leveraging digital technologies: How information quality leads to localized capabilities and customer service performance*. **MIS quarterly**, p. 565-590, 2013.
- SHAHBAZI, M. *et al. Development of a scale for data quality assessment in automated library systems*. **Library & Information Science Research**, v. 41, n. 1, p. 78-84, 2019.
- SHI, N. S.; SILVIUS, G. (Ed.). *Enterprise IT Governance, Business Value and Performance Measurement*. **IGI Global**, 2010.
- SILVA, E. L.; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração de Dissertação**. 2001.
- SCANNAPIECO, M.; CATARCI, T. *Data quality under a computer science perspective*. **Archivi & Computer**, v. 2, p. 1-15, 2002.
- STRONG, D. M.; LEE, Y. W.; WANG, R. Y. **10 Potholes in the Road to Information Quality**. **Computer**, v. 30, n. 8, p. 38-46, Aug. 1997
- VALENCIA-PARRA, A. *et al. DMN4DQ: when data quality meets DMN*. **Decision Support Systems**, v. 141, p. 113450, 2021.
- VIDGEN, R.; SHAW, S.; GRANT, D. B. *Management challenges in creating value from business analytics*. **European Journal of Operational Research**, v. 261, n. 2, p. 626-639, 2017.

WANG, Y.; WANG, R. *Anchoring Data Quality Dimensions in Ontological Foundations. Communications of the ACM*, v. 39, n. 11, 1996.

WANG, R. Y. *A Product Perspective on Total Data Quality Management. Communications of the ACM*, v. 41, n. 2, p. 58-65, 1998.

WANG, R. Y.; STRONG, D. M. Beyond accuracy: *What data quality means to data consumers. Journal of Management Information Systems*, v. 12, n. 4, p. 5-33, 1996.

WANG, R.; ZIAD, M.; LEE, Y. W. *Data Quality*. Kluwer Academic Publishers, 2000.

WEISKOPF, N. G.; WENG, C. *Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. Journal of the American Medical Informatics Association*, v. 20, n. 1, p. 144-151, 2013.

YIN, R. K. *Estudo de Caso: Planejamento e Métodos*. Porto Alegre, RS. 2005.

ZAVERI, A. *et al. Quality assessment for linked data: A survey. Semantic Web*, v. 7, n. 1, p. 63-93, 2016.

## ANEXOS

ANEXO A - FORMULÁRIO PARA SOLICITAR ALTERAÇÃO DE CADASTRO DE CLIENTES CRIADO NA ETAPA MEDIR.

### Formulário para solicitar alteração de cadastro de clientes criado na etapa medir.

Formulário aplicado para coleta de dados destinado a solicitações de modificação de cadastro de clientes.

Solicitação de alteração de dados dos clientes

Este formulário é para abertura de chamados referente a alteração de canal, tabela de preço, expandir clientes e desbloquear/reactivar clientes (ATENÇÃO AO CAMPOS OBRIGATÓRIOS DE PREENCHIMENTO)

CDA

POÇOS DE CALDAS

Vendedor

Setor

Motivo do Chamado

DESBLOQUEIO/REATIVAÇÃO

EXPANSÃO DE CLIENTES

ALTERAÇÃO DE CANAL (DETALHAR NO CAMPO OBSERVAÇÃO A MUDANÇA)

ALTERAÇÃO DE TAB DE PREÇOS

PRODUTOS SEM PREÇO

IMPOSSIBILIDADE DE DIGITAR PEDIDO

CORREÇÃO DE CADASTRO DE CLIENTES (DETALHAR NO CAMPO OBSERVAÇÃO)

Cod\_Cliente

Nome Cliente

CPF/CNPJ

E-mail

DATA DE NASCIMENTO DO CLIENTE

---

TELEFONE

---

Atenção para a alteração de Canal/Tabela/Correção de Cadastro é necessário o envio de três fotos do local (balcão para fora, por dentro e por fora) e incluir no chamado. É o seu caso?

- Sim
- Não

Observações

---

---

---

Add imagem  
(Upload de imagem)

Add imagem 2  
(Upload de imagem)

Add imagem 3  
(Upload de imagem)