



UNISUL

UNIVERSIDADE DO SUL DE SANTA CATARINA

EDUARDO FARIAS

STORYTELLING DE DADOS:
CONTANDO HISTÓRIAS COM *DASHBOARDS*

Palhoça
2020

EDUARDO FARIAS

***STORYTELLING DE DADOS:
CONTANDO HISTÓRIAS COM DASHBOARDS***

Trabalho de Conclusão de Curso
apresentado ao Curso de Bacharelado em
Sistemas de Informação da Universidade
do Sul de Santa Catarina como requisito
parcial à obtenção do título de Bacharel em
Sistemas de Informação.

Orientador: Prof. Aran Bey Tcholakian Morales, Dr.

Coorientadora: Prof. Maria Inés Castiñeira, Dra.

Palhoça
2020

EDUARDO FARIAS

STORYTELLING DE DADOS:
CONTANDO HISTÓRIAS COM *DASHBOARDS*

Este Trabalho de Conclusão de Curso foi julgado adequado à obtenção do título de Bacharel em Sistemas de informação e aprovado em sua forma final pelo Curso de Sistemas de informação da Universidade do Sul de Santa Catarina.

Palhoça, 01 de dezembro de 2020.

Professor e orientador Prof. Aran Bey Tcholakian Morales, Dr.
Universidade do Sul de Santa Catarina

Prof. Maria Inés Castiñeira, Dra.
Universidade do Sul de Santa Catarina

Prof. Vera Rejane Niedersberg Schuhmacher, Dra.
Universidade do Sul de Santa Catarina

Dedico esse trabalho a todas as mulheres presentes em minha vida. Que em momento algum deixaram de honrar as suas maiores qualidades: amar e cuidar.

AGRADECIMENTOS

A Deus por ter me dado saúde e força para superar as dificuldades.

Aos meus pais, Adauto Rosa Farias (*in memoriam*) e Sarlene Stock Farias pelo amor e educação.

Um agradecimento especial para minha mãe que fez muitas marmitas com a melhor refeição do mundo. Com isso consegui mais tempo para me concentrar neste trabalho. Muito obrigado, mãe!

Ao meu irmão Bruno Farias pelo apoio e união nos momentos difíceis que passamos.

A minha esposa Thais Edir Dalló, que fez de tudo para me incentivar dando apoio em todos os momentos. Ajudando na correção deste trabalho e ainda lidou com meus vários momentos de mau humor.

Aos meus padrinhos Lorena Stock Chaves e Udo Stock pela educação e incentivo para finalizar essa jornada.

A esta universidade, seu corpo docente, direção e administração que me deram oportunidade para aprender e hoje trabalhar aplicando os conhecimentos adquiridos.

Aos professores Dr. Aran Bey Tcholakian Morales e a Dra. Maria Inés Castiñeira, responsáveis pela orientação deste trabalho. Obrigado pelo direcionamento, esclarecimento, incentivo e serem tão atenciosos e pacientes.

E a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

“Os números têm uma história importante para contar. Eles confiam em você para lhes dar uma voz clara e convincente”.
(Stephen Few)

RESUMO

O *Dashboard* é uma ferramenta de apresentação visual de dados com objetivos bem definidos e que incorporam uma solução de *Business Intelligence*. Uma solução de *Business Intelligence* deve melhorar a tomada de decisão provendo dados de qualidade, para isso o *Dashboard* deve revelar *insights* que apontem para uma decisão correta. Em muitas ocasiões os *Dashboards* não entregam seus objetivos. São desenvolvidos fora de contexto e não são criados pensando no seu público-alvo. Também são desenvolvidos usando recursos de visualização de dados inadequados ou mal elaborados. Em alguns casos são implementados sem uma estrutura ordenada para os critérios mais importantes e que exigem maior atenção. Este trabalho tem como objetivo desenvolver um *Dashboard* para uma solução de *Business Intelligence* aplicando técnicas e conceitos de *Storytelling* de Dados. Para isso, foi conceituado sobre *Business Intelligence*, *Dashboard* e *Storytelling* de Dados. Foi descrita a importância de se apresentar os dados em um contexto correto e de se conhecer bem o seu público-alvo, assim como a importância de empregar as técnicas corretas de visualização de dados e narrativa. Cada dado tem a sua melhor forma de ser apresentado, a escolha do gráfico correto é fundamental para que os objetivos sejam alcançados. Para aplicar e validar os conceitos foi aplicado um experimento através de uma solução de *Business Intelligence* para a empresa estudo de caso. Como primeira etapa foi feita uma análise de requisitos e levantamento de indicadores para a modelagem de dados dimensional e construção do *Data Warehouse* (DW). Seguindo para o processo de ETL onde os dados foram extraídos, transformados e carregados no DW. Com os dados prontos e os requisitos levantados foi possível desenvolver um *Dashboard* aplicando os conceitos de *Storytelling* de dados. Como resultado obtido temos uma solução de *BI* que atende as necessidades do diretor da empresa estudo de caso, com um *Dashboard* que cumpre seu objetivo proposto tornando a tomada de decisão mais rápida e precisa. O emprego do *Storytelling* de dados no *Dashboard* propiciou a visualização de forma clara e evidenciou o estado dos principais indicadores. Além disso, o *Dashboard* conseguiu orientar, de uma forma preliminar, o que pode ser feito para reverter o estado ruim revelado por um indicador. Desta forma verificou-se que a aplicação de uma narrativa de dados, condizente ao contexto da empresa e público-alvo, bem como utilização adequada de elementos visuais, conceitos base para um bom *Storytelling* de dados, trouxeram resultados expressivos no desenvolvimento do *Dashboard* para monitoramento de negócio.

Palavras-chave: *Data Warehouse*. *Business Intelligence*. *Dashboard*. *Storytelling*. Visualização de Dados. ETL. Modelagem Dimensional.

ABSTRACT

The Dashboard is a visual data presentation tool with well-defined objectives and incorporating a Business Intelligence solution. A Business Intelligence solution should improve decision making by providing quality data, for which the Dashboard should reveal insights that point to a correct decision. On many occasions Dashboards do not deliver their goals. They are developed out of context and are not created with your target audience in mind. They are also developed using inadequate or poorly designed data visualization capabilities. In some cases they are implemented without an orderly structure for the most important criteria and that require more attention. This work aims to develop a Dashboard for a Business Intelligence solution applying Data Storytelling techniques and concepts. For that, it was conceptualized on Business Intelligence, Dashboard and Data Storytelling. The importance of presenting the data in a correct context and of getting to know your target audience was described, as well as the importance of employing the correct data visualization and narrative techniques. Each data has its best way of being presented, the choice of the correct graph is fundamental for the objectives to be achieved. To apply and validate the concepts, an experiment was applied through a Business Intelligence solution for the case study company. As a first step, a requirements analysis and survey of indicators for the modeling of dimensional data and construction of the Data Warehouse (DW) was carried out. Moving on to the ETL process where the data was extracted, transformed and loaded into the DW. With the data ready and the requirements raised it was possible to develop a Dashboard applying the concepts of Data Storytelling. As a result we have a BI solution that meets the needs of the company's case study director, with a Dashboard that fulfills its proposed objective, making decision making faster and more accurate. The use of Data Storytelling on the Dashboard provided a clear visualization and evidenced the status of the main indicators. In addition, the Dashboard was able to guide, in a preliminary way, what can be done to revert the bad state revealed by an indicator. In this way it was found that the application of a data narrative, consistent with the context of the company and target audience, as well as the appropriate use of visual elements, basic concepts for a good data Storytelling, brought expressive results in the development of the Dashboard for monitoring business.

Keywords: Data Warehouse. Business Intelligence. Dashboard. Storytelling. Data Visualization. ETL. Dimensional Modeling.

LISTA DE FIGURAS

Figura 1 – Componentes de um sistema de <i>Business Intelligence</i>	22
Figura 2 – Exemplo visual de um fluxo de ETL	23
Figura 4 – Exemplo de tabelas em uma modelagem dimensional	25
Figura 5 – Exemplo de uma tabela de dimensão	26
Figura 6 – Exemplo de um meio entidade relacionamento de uma tabela fato	27
Figura 7 – Exemplo visual de um modelo estrela	28
Figura 8 – Exemplo de <i>Dashboard</i> de vendas	32
Figura 9 – Exemplo de <i>Dashboard</i> com mais visualizações de dados.....	33
Figura 10 – União de elementos: Narrativa, visuais e dados	35
Figura 11 – Exemplo de visualização de dados: Gráfico de barras.....	39
Figura 12 – Exemplo de visualização de dados: Gráfico de barras.....	40
Figura 13 – Exemplo de visualização de dados: Gráfico de barras.....	40
Figura 14 – Exemplo de visualização de dados: Gráfico de barras.....	41
Figura 15 – Exemplo de visualização de dados gráfico de linha	42
Figura 16 – Exemplo de visualização gráfico de pontos	42
Figura 17 – Exemplo de visualização de texto e número	43
Figura 18 – Exemplo de visualização de texto e número	43
Figura 19 – Exemplo de visualização de texto e número	44
Figura 20 – Exemplo de uma tabela como elemento de visualização de dados	44
Figura 21 – Exemplo de uma tabela como elemento de visualização de dados	45
Figura 22 – Exemplo de uma tabela como elemento de visualização de dados	46
Figura 23 – Exemplo de narrativa sequencial em <i>Dashboard</i>	49
Figura 24 – Modelagem dimensional – Esquema estrela.....	61
Figura 25 – Canvas para <i>Storytelling</i>	71
Figura 26 – Arquitetura da solução	73
Figura 27 – Staging st_cliente	75
Figura 28 – Data Warehouse dim_cliente	75
Figura 29 – Staging st_produto	76
Figura 30 – Data Warehouse dim_produto	76
Figura 31 – Staging st_vendedor	77
Figura 32 – Data Warehouse dim_vendedor.....	78
Figura 33 – Dimensão geografia	78

Figura 34 – Dimensão tempo	79
Figura 35 – Staging st_imposto_ipi_item_pedido.....	79
Figura 36 – Staging st_imposto_st_item_pedido.....	80
Figura 37 – Staging st_item_pedido.....	80
Figura 38 – Staging st_pedido.....	81
Figura 39 – Data Warehouse ft_venda.....	82
Figura 40 – Carga completa da <i>staging</i>	83
Figura 41 – Carga dimensões para DW	84
Figura 42 – Carga DW completa	84
Figura 43 – Data <i>Storytelling</i> Canvas.....	85
Figura 44 – Dashboard Estrutura Atendimento	87
Figura 45 – Dashboard Controle Faturamento	87

LISTA DE QUADROS

Quadro 1 – Aplicações de gráfico	46
Quadro 2 – Quando usar gráficos ou tabelas.....	47
Quadro 3 – Classificação do tipo de pesquisa	51
Quadro 4 – Levantamento das regras de negócio	56
Quadro 5 – Levantamento das regras de negócio	59
Quadro 6 – Os requisitos funcionais descrevem as regras de negócio.....	59
Quadro 7 – Requisitos não funcionais.....	60
Quadro 8 – Dimensão cliente	62
Quadro 9 – Dimensão geografia	63
Quadro 10 – Dimensão produto	64
Quadro 11 – Dimensão tempo	65
Quadro 12 – Dimensão vendedor	66
Quadro 13 – Fato venda	67
Quadro 14 – Experimento	72
Quadro 15 – Arquitetura da solução.....	73

LISTA DE ABREVIATURAS

BI	Business Intelligence
DM	Data Mart
DW	Data Warehouse
ETL	Extract, Transform and Load
OLAP	Online Analytical Processing
OLTP	Online Transaction Processing
SQL	Structure Query Language
ERP	Enterprise Resource Planning
DAX	Data Analysis Expressions
BD	Base de Dados
SSMS	SQL Server Management Studio
SGBD	Sistema de Gerenciamento de Banco de Dados

SUMÁRIO

1	INTRODUÇÃO.....	15
1.1	PROBLEMÁTICA	15
1.2	OBJETIVOS	16
1.2.1	Objetivo geral.....	16
1.2.2	Objetivos específicos.....	16
1.3	JUSTIFICATIVA	17
1.4	ESTRUTURA DA MONOGRAFIA	17
2	REVISÃO BIBLIOGRÁFICA	19
2.1	CONCEITOS DE BUSINESS INTELLIGENCE (BI).....	19
2.1.1	O que é Business Intelligence?.....	19
2.1.2	Importância de BI.....	20
2.1.3	Aplicação de BI.....	21
2.1.4	Arquitetura de BI.....	21
2.1.4.1	Back-end de BI.....	22
2.1.4.2	Front-end de BI.....	22
2.2	BACK-END DE UMA ARQUITETURA DE BI	23
2.2.1	ETL.....	23
2.2.2	Data Warehouse.....	24
2.2.3	Modelo dimensional	24
2.2.3.1	Tabela dimensão	26
2.2.3.2	Tabela fato.....	26
2.2.3.3	Métricas.....	27
2.2.3.4	Modelo estrela	27
2.3	FRONT-END DE UMA ARQUITETURA DE BI.....	28
2.3.1	OLAP.....	28
2.3.1.1	Cubos OLAP.....	29
2.3.1.1.1	Drill Down e Roll Up	29
2.3.1.1.2	Drill Across e Drill Through.....	29
2.3.1.1.3	Slice and Dice e Pivot.....	30
2.3.2	KPI.....	30
2.3.3	Dashboards	31
2.4	STORYTELLING DE DADOS.....	33
2.4.1	Storytelling de dados	34
2.4.2	Contexto e público-alvo	35
2.4.2.1	Contexto	35
2.4.2.2	Público-alvo	36
2.4.3	Visualização de dados	37
2.4.3.1	Gráficos de barra.....	38
2.4.3.2	Gráficos de Linha	41
2.4.3.3	Gráficos de ponto	42
2.4.3.4	Sparkline (mini gráficos).....	43
2.4.3.5	Bullet Charts.....	43
2.4.3.6	Textos e cartões	43
2.4.3.7	Tabelas.....	44
2.4.3.8	Mapas geográficos	45

2.4.3.9	<i>Ícones</i>	45
2.4.3.10	<i>Aplicações para cada tipo de gráfico</i>	46
2.4.3.11	<i>Quando usar gráficos ou tabelas</i>	47
2.4.4	Narrativas e Storyboard de dados	47
2.4.4.1	<i>Narrativas de dados</i>	48
2.4.4.2	<i>Storyboard de dados</i>	49
3	MÉTODO DE PESQUISA	51
3.1	CARACTERIZAÇÃO DO TIPO DE PESQUISA.....	51
3.2	ETAPAS METODOLÓGICAS	52
3.3	PROPOSTA DE SOLUÇÃO	52
3.4	DELIMITAÇÕES.....	53
4	MODELAGEM	54
4.1	CONTEXTO DO NEGÓCIO	54
4.2	REQUISITOS	54
4.2.1	Entrevista	55
4.2.2	Levantamento das regras de negócio	56
4.2.3	Requisitos funcionais	59
4.2.4	Requisitos não funcionais	59
4.3	MODELAGEM DIMENSIONAL.....	60
4.3.1	Dimensão cliente	61
4.3.2	Dimensão geografia	62
4.3.3	Dimensão produto	63
4.3.4	Dimensão tempo	64
4.3.5	Dimensão vendedor	65
4.3.6	Fato venda	66
5	DESENVOLVIMENTO	68
5.1	TECNOLOGIAS E FERRAMENTAS	68
5.1.1	Microsoft SQL Server	68
5.1.2	Microsoft SQL Server Management Studio	68
5.1.3	SQL Power Architect	69
5.1.4	Pentaho Data Integration	69
5.1.5	Microsoft Power BI	70
5.1.6	DAX	70
5.1.7	Canvas para Storytelling	70
5.2	EXPERIMENTO	72
5.2.1	Arquitetura da solução	72
5.2.2	Integração de dados	74
5.2.2.1	<i>Dimensão cliente</i>	74
5.2.2.1.1	Staging st_cliente	74
5.2.2.1.2	Data Warehouse dim_cliente	75
5.2.2.2	<i>Dimensão produto</i>	75
5.2.2.2.1	Staging st_produto	76
5.2.2.2.2	Data Warehouse dim_produto.....	76
5.2.2.3	<i>Dimensão vendedor</i>	77
5.2.2.3.1	Staging st_vendedor.....	77

5.2.2.3.2	Data Warehouse dim_vendedor	77
5.2.2.4	<i>Dimensão geografia</i>	78
5.2.2.5	<i>Dimensão tempo</i>	78
5.2.2.6	<i>Fato vendas</i>	79
5.2.2.6.1	Staging st_imposto_ipi_item_pedido	79
5.2.2.6.2	Staging st_imposto_st_item_pedido	80
5.2.2.6.3	Staging st_item_pedido	80
5.2.2.6.4	Staging st_pedido	81
5.2.2.6.5	Data Warehouse ft_venda	81
5.2.2.7	<i>Carga no DW</i>	82
5.2.2.7.1	Carga completa da <i>staging</i>	82
5.2.2.7.2	Carga dimensões para DW	83
5.2.2.7.3	Carga DW completa	84
5.2.3	Planejamento do <i>storytelling</i>	85
5.2.4	Implementação do <i>Dashboard</i>	85
5.3	CONSIDERAÇÕES	88
6	CONCLUSÕES E TRABALHOS FUTUROS	90
6.1	CONCLUSÕES	90
6.2	TRABALHOS FUTUROS	91
	REFERÊNCIAS	93

1 INTRODUÇÃO

Possuir acesso à informação e dados de qualidade se tornou algo extremamente estratégico nas organizações. A informação é algo que toda empresa necessita para tomar decisões nos negócios, sejam elas estratégicas, táticas e/ou operacionais. Obter esses dados e informações qualificados e de fácil entendimento se tornou cada vez mais trabalhoso devido à complexidade dos negócios e às novas necessidades dos clientes. Para ajudar na apresentação e organização desses dados e informações os sistemas de *Business Intelligence* (BI) estão cada vez mais presentes nas organizações.

Os sistemas de *Business Intelligence* são conjuntos de conceitos e ferramentas que ajudam a organizar e padronizar a extração e apresentação dos dados e informações. O desenvolvimento de um sistema de *Business Intelligence* deve estar focado diretamente no negócio da empresa e em seu público-alvo. Segundo Turban *et al.* (2009, p. 27):

Os principais objetivos do *Business Intelligence* são permitir o acesso interativo aos dados (às vezes, em tempo real), proporcionar a manipulação desses dados e fornecer aos gerentes e analistas de negócios a capacidade de realizar a análise adequada.

Um sistema de *Business Intelligence* possui recursos gráficos de interação com usuários, representados como: ferramentas de *front-end*. Dentre as ferramentas a que apresentam um acesso mais rápido e personalizado aos dados e informações são os *Dashboards*. Segundo Few (2006, p. 4), “Os *Dashboards* oferecem uma única poderosa solução para a necessidade de uma organização em obter informação”.

Os *Dashboards*, como ferramentas de visualização de dados, podem ser desenvolvidos com de técnicas de Storytelling de dados. De acordo com Dykes (2016, p. 1), “*Storytelling* de dados é uma abordagem estruturada para comunicar informações de dados e envolve uma combinação de três elementos principais: dados, recursos visuais e narrativa”.

1.1 PROBLEMÁTICA

Os profissionais de *Business Intelligence* têm um grande desafio nas organizações. Eles precisam abstrair os dados e informações, captados das mais

variadas fontes, em ferramentas visuais, como exemplo os *Dashboards*, para que estes satisfaçam as expectativas do seu público-alvo.

Para que os anseios dos usuários sejam alcançados os desenvolvedores precisam ter um conhecimento do contexto do problema que irão resolver. O conjunto dessas circunstâncias orientará o desenvolvedor a aplicar técnicas visuais eficientes.

Muitas vezes o processo de desenvolvimento não é orientado ao contexto correto e nem ao público-alvo, tampouco utiliza de técnicas visuais adaptadas a uma melhor interpretação humana. Isso pode tornar o *Dashboard* uma ferramenta deficiente em seu propósito de tomada de decisão.

Storytelling de dados contextualizam a fim de criar uma conexão entre os dados apresentados e o público-alvo. Isso faz com que o propósito do *Dashboard*, em ser uma ferramenta de comunicação, seja alcançado.

Diante disso, a questão que direciona esse trabalho é: Como utilizar *Storytelling* de dados em *Dashboards* a fim de criar uma melhor experiência ao público-alvo?

1.2 OBJETIVOS

Os objetivos desta monografia são divididos em objetivo geral e objetivos específicos. Conforme descritos a seguir.

1.2.1 Objetivo geral

Este trabalho propõe-se a apresentar, através de um exemplo, técnicas que visam o aprimoramento de visualização de dados em Dashboards a fim de orientar desenvolvedores de Business Intelligence a produzirem com foco em *Storytelling* de dados considerando a escolha de um visual eficaz bem como a compreensão de contexto e público-alvo e dessa forma alcançar a expectativa na qual foi proposto.

1.2.2 Objetivos específicos

- Conceituar *Business Intelligence* e apresentar suas ferramentas com ênfase nas ferramentas de *front-end*.

- Apresentar conceitos de visualização de dados que aprimoram as narrativas de dados.
- Contribuir para a comunidade que desenvolve *Dashboards*, tornando o trabalho uma ferramenta de consulta e referência em língua Portuguesa.
- Aprofundar conhecimentos de visualização de dados e *Storytelling*.
- Projetar como estudo de caso, um *Dashboard* comercial para uma empresa do segmento de distribuição de produtos alimentícios com os conceitos estudados.
- Apresentar um roteiro para a produção de um *Dashboard* orientado a *Storytelling* de Dados.

1.3 JUSTIFICATIVA

Devido ao desenvolvimento de *Dashboards* ineficientes em seu propósito, conforme Few (2007, p. 1), “A maioria os painéis usados nas empresas hoje falham. Na melhor das hipóteses fornecem apenas uma fração do *insight* necessário para monitorar o negócio”. Assim sendo esta pesquisa se justifica através da aplicação de técnicas de *Storytelling* de dados na criação de *Dashboards* pois segundo Davenport (2015, p. 1):

As histórias sempre foram ferramentas eficazes para transmitir a experiência humana. *Storytelling* é a maneira como simplificamos e compreendemos um mundo complexo. Ele fornece contexto, *insight*, interpretação - tudo o que torna os dados significativos e analíticos mais relevantes e interessantes.

Desta forma teremos o benefício de aperfeiçoar a comunicação, por conseguinte melhorar a tomada de decisão.

1.4 ESTRUTURA DA MONOGRAFIA

Esta monografia está dividida em cinco capítulos. O capítulo 1 explana o tema a ser estudado, os objetivos e a justificativa do trabalho.

No capítulo 2 são abordados conceitos de *Business Intelligence*, ferramentas *front-end*, conceitos de visualização de dados, *Dashboards* e *Storytelling* de dados.

O capítulo 3 apresenta a metodologia de como serão alcançados os objetivos propostos e o planejamento deste trabalho.

No capítulo 4 é descrita a modelagem e requisitos para desenvolvimento do *Data Warehouse*, base para a leitura dos dados utilizados nos *Dashboards*.

O capítulo 5 apresenta o desenvolvimento do *Dashboard* orientado a *Storytelling* de Dados.

No capítulo 6 consta as conclusões, resultados obtidos e trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados os conceitos que oferecem a base teórica que sustenta a obtenção dos objetivos deste trabalho. O capítulo é segmentado em três seções principais: Conceitos e arquitetura de *Business Intelligence* (BI), *Back-end* de uma arquitetura de BI e *Front-end* de uma arquitetura de BI.

2.1 CONCEITOS DE BUSINESS INTELLIGENCE (BI).

Esta seção aborda os principais conceitos de *Business Intelligence* (BI), a importância e aplicação do *Business Intelligence* nas organizações e a arquitetura de BI.

2.1.1 O que é Business Intelligence?

O termo *Business Intelligence* (BI) pode ter definições diferentes para vários autores. Algumas definições são apresentadas a seguir:

Turban, Sharda, Delen (2016, p. 16) esclarecem que:

Business Intelligence (BI) é um termo abrangente que combina arquiteturas, ferramentas, bancos de dados, ferramentas analíticas, aplicativos e metodologias. O principal objetivo do BI é permitir o acesso interativo aos dados, permitir a manipulação dos dados e fornecer aos gerentes de negócios e análises a capacidade de realizar análises apropriadas. Ao analisar dados históricos e atuais, situações e performances, os tomadores de decisão obtêm informações valiosas que lhes permitem tomar decisões mais informadas e melhores. O processo de BI é baseado na transformação de dados em informação, depois em decisões e, finalmente, em ações.

Stachowiak, Rayman e Greenwald (2007, p. 3) afirmam:

Business Intelligence pode ser definida como o acesso correto aos dados ou informações necessárias para tomar as decisões comerciais corretas no momento certo. Os dados podem estar brutos ou podem ter sido analisados de alguma maneira. O acesso a essas informações permite o gerenciamento dos negócios de fato, em vez de depender principalmente da intuição.

Luhn (1958, p. 314):

São técnicas baseadas em procedimentos estatísticos que podem ser executados nas máquinas de processamento de dados. Juntamente com os recursos adequados de comunicação e o equipamento de entrada e saída, um sistema abrangente pode ser montado para acomodar todos os

problemas de informação de uma organização. Chamamos isso de Sistema de *Business Intelligence*.

Para Davenport e Harris (2007, p. 3):

Business Intelligence é um termo genérico para um conjunto de sistemas, aplicativos e processos de governança que permitem sofisticada análises, permitindo dados, conteúdo e análises fluir para aqueles que precisam quando eles precisam.

Segundo Prat e Fruhlinger (2019, p. 1): “O termo *Business Intelligence* também se refere a uma variedade de ferramentas que fornecem acesso rápido, fácil e insights sobre o estado atual de uma organização, com base nos dados disponíveis”.

Foley e Guillemette (2010, p. 4) conceituam:

Business Intelligence (BI) é uma combinação de processos, políticas, cultura e tecnologias para reunir, manipular, armazenar e analisar dados coletados de fontes internas e fontes externas, a fim de comunicar informações, criar conhecimento e informar a tomada de decisão.

Defronte das várias definições percebe-se que o termo *Business Intelligence* é bastante abrangente. Nota-se um consenso entre os autores como sendo um processo em que se utiliza recursos computacionais como: banco de dados e ferramentas de análises e dados provenientes do negócio, com o propósito de propiciar aos gestores e analistas um ambiente que possibilite o acesso a informações com o intuito de auxiliar o processo de tomada de decisão nas organizações.

2.1.2 Importância de BI

O crescimento dos negócios e a competitividade entre as organizações tornam cada vez as decisões complexas e exigem sistemas próprios para esse apoio, como o BI.

Turban, Sharda e Delen (2016, p. 11) afirmam:

Muitas decisões envolvem cálculos. Os dados para esses podem ser armazenados em diferentes bancos de dados, em qualquer lugar da organização e até mesmo possivelmente fora da organização. Os dados podem incluir texto, som, gráficos e vídeo, e estes podem estar em diferentes idiomas. Muitas vezes é necessário transmitir dados rapidamente de locais distantes. Atualmente, os sistemas podem pesquisar e transmitir os dados necessários de forma rápida, econômica, segura e transparente.

O BI tem uma grande relevância nas organizações pois segundo Foley e Guillemette (2010, p. 4):

BI ajuda a relatar o desempenho dos negócios, descobrir novas oportunidades e melhorar decisões de negócios em relação aos concorrentes, fornecedores, clientes, questões financeiras, questões estratégicas e de produtos e serviços.

Loh (2014, p. 14), informa que o grande objetivo do BI é acabar com o “achismo” ou “empirismo”.

2.1.3 Aplicação de BI

Alguns exemplos de aplicações e benefícios ao utilizar BI:

Howston (2008 p.17 e 18):

Em marketing, ganhos de desempenho podem ser alcançados melhorando taxas de respostas para campanhas específicas, identificando características de clientes mais responsivos. Eliminar campanhas ineficazes economiza milhões de dólares a cada ano. Business Intelligence permite empresas a aumentar as receitas através da venda cruzada de produtos a clientes existentes. O pessoal da contabilidade pode usar BI para reduzir o envelhecimento das contas recebíveis pela identificação de clientes que pagam atrasados. Na manufatura, o BI facilita a análise de lacunas para entender por que certas plantas operam com mais eficiência do que outras.

As aplicações e benefícios de BI são inúmeras, Foley e Guillemette (2010, p. 12) ainda afirmam que “BI é reconhecido como gerador de benefícios intangíveis” pois podem gerar novas ideias a partir de análises e insights.

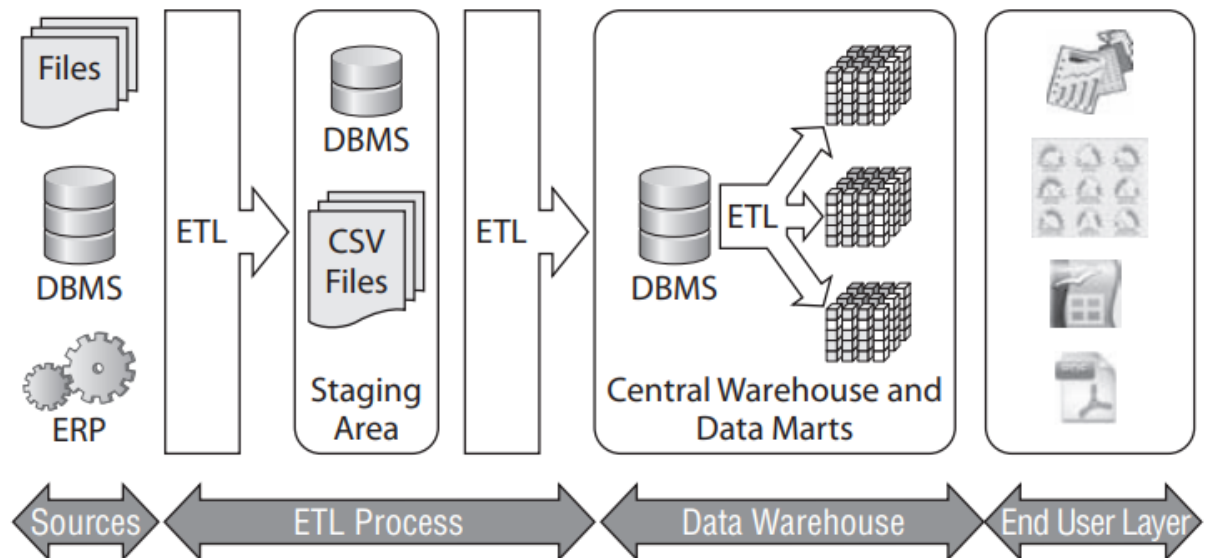
Loh (2014, p. 10) entende que a aplicação de BI “tem por objetivo encontrar causas ou explicações para eventos ou resultados”. Acrescenta que com BI deve-se procurar as práticas de sucesso. Identificando o vendedor que tem a melhor performance procuramos entender por que ele é o melhor. Com esses conhecimentos podemos replicar em boas práticas para os outros vendedores.

2.1.4 Arquitetura de BI

A arquitetura de BI pode ser segmentada em duas partes: *back-end* e *front-end*. A *back-end* é representada por aquilo que o usuário final não vê, mas é necessária para construção de um sistema de BI. A arquitetura de *front-end* de BI pode ser

descrita como o conjunto de ferramentas, visuais ou analíticas que o usuário final de um sistema de BI utiliza para tomada de decisão ou *insights* sobre o negócio.

Figura 1 – Componentes de um sistema de *Business Intelligence*



Fonte: Bouman e Dongen (2009, p. 8).

2.1.4.1 *Back-end de BI*

De forma simplificada conforme Davenport e Harris (2007, p. 3) podemos descrever o *back-end* de BI da seguinte forma:

1. Fontes de onde os dados serão extraídos.
2. Ferramentas e processos para extração, transformação e carga (ETL) para fim de armazenar em um banco de dados para posteriores análises.
3. Repositórios como *Data Warehouse* e *Data Marts* que utilizam Banco de Dados para armazenar os dados extraídos e já transformados.

2.1.4.2 *Front-end de BI*

Apresentada, conforme Davenport e Harris (2007, p. 3) pelas ferramentas:

1. Ferramentas para análises como Cubos OLAP e aplicações para *Data Mining*.
2. Aplicações que permitam visualização dos dados como *Dashboards*, relatórios e KPI.

2.2 BACK-END DE UMA ARQUITETURA DE BI

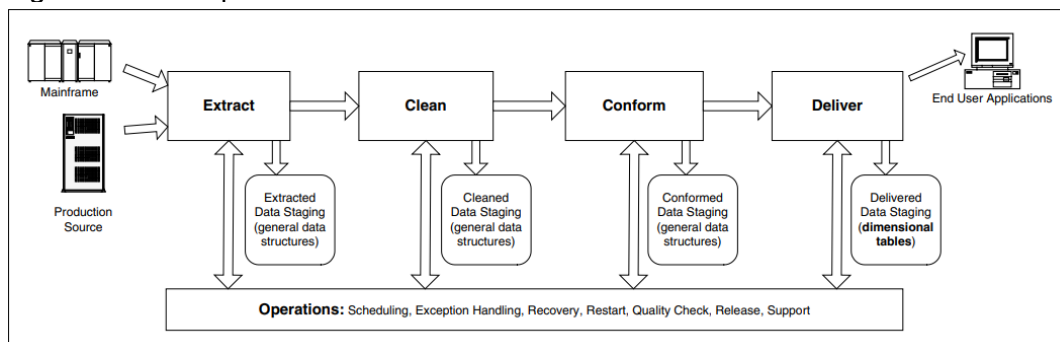
Esta seção aborda conceitos sobre ETL (*Extract, Transform and Load*), *Data Warehouse* e sobre modelagem dimensional.

2.2.1 ETL

Os sistemas de BI necessitam de dados para uma posterior análise. Os dados podem estar em diferentes fontes. Como: arquivos, planilha de Excel, banco de dados, webservices provedores de informações e entre outros.

O processo de ETL é descrito pelo acrônimo em inglês: (*Extract, Transform and Load*) ou Extração, Transformação e Carga que representa suas três etapas. Inmon (2005, p. 111) cita como o ETL sendo um processo de integração de dados, pois é desta forma que unifica os dados de vários locais.

Figura 2 – Exemplo visual de um fluxo de ETL



Fonte: Kimball (2004, p. 18).

O processo de Extração buscará em todas as fontes mapeadas os dados necessários para satisfazerem os requisitos dos usuários. O resultado dessa integração será armazenado em um banco de dados em uma área chamada de *staging*. Segundo Machado (2004, p. 58) o objetivo da área de *staging* é criar um ambiente intermediário de armazenamento e processamento dos dados oriundos de aplicações OLTP e outras fontes e com isso evitar problemas após a criação do DW e a concorrência com o ambiente transacional no consumo de recursos.

Após os dados serem extraídos, precisam ser limpos e transformados para atender as necessidades do projeto. Davenport e Harris (2007, p. 7) citam exemplos de limpezas e padronizações como: CEP americano deve conter nove dígitos ou uma

adequação de negócios onde um cliente pode ser uma empresa em um sistema e em outro pode ser um fornecedor.

Por último o processo de Carga é o processo que irá popular os dados em um *Data Warehouse* para serem utilizados em posteriores análises.

2.2.2 Data Warehouse

No sistema de BI o *Data Warehouse* (DW) será o repositório dos dados. O DW armazena os dados do resultado obtido pelo processo de ETL. Conforme descrito por Inmon (2000, p. 1) o *Data Warehouse* “armazena os dados necessários para um processamento analítico e informativo durante um período histórico”.

Um *Data Warehouse* consolida dados de muitas fontes e assuntos das várias áreas de negócios, dispostas de maneira a promover o acesso e interpretação dos dados pelos usuários corporativos (DAGAN, 2007, p. 23). Kimball e Caserta (2004, p. 22) destacam que a missão chave do DW é publicar os dados. Os dados segundo Inmon (2000, p. 1) são relacionados aos assuntos que permeiam os negócios da organização, tais como as operações da empresa.

Um DW pode ser formado por uma integração de vários repositórios relativos a cada departamento ou assuntos de uma organização. Esses repositórios segmentados são chamados de *Data Mart*. Davenport e Harris (2007, p. 8) descrevem que “um *Data Mart* pode se referir a um repositório separado ou a uma seção particionada de um *Data Warehouse* completo.

Os *Data Marts* são geralmente usados para oferecer suporte a uma única função”. Dagan (2007, p. 8) destaca que o termo *Data Mart* normalmente é uma versão menor de um DW e geralmente pertence a uma única área de negócios, enquanto o DW compreende várias áreas de uma empresa.

2.2.3 Modelo dimensional

Adamson (2010, p. 3) indica que existe dois tipos principais de sistemas de informação. O primeiro é o sistema que suporta a execução dos processos de negócio das organizações e o outro são os sistemas que suportam as análises dos processos de negócio. Os conceitos da modelagem dimensional evoluíram, a fim de atender aos requisitos dos sistemas analíticos.

Adamson (2010, p. 3) conceitua:

O núcleo de todo modelo dimensional é um conjunto de métricas de negócios que captura como um processo é avaliado e uma descrição do contexto de cada medição. Sistemas analíticos e sistemas operacionais servem a propósitos fundamentalmente diferentes. Um sistema operacional suporta a execução de um processo de negócios, enquanto um sistema analítico suporta a avaliação do processo. Seus objetivos distintos são refletidos em perfis de uso contrastantes, que por sua vez sugerem que princípios diferentes guiarão seu design.

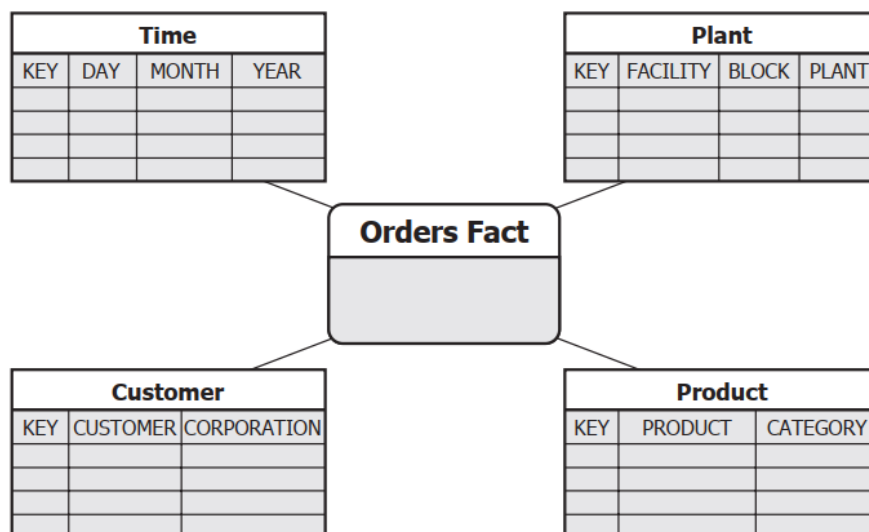
Kimball e Ross (2013, p. 7) apresentam a modelagem dimensional como sendo a técnica preferida para apresentar dados analíticos pois abrange dois pontos importantes:

- Apresenta os dados de forma simples para o entendimento dos usuários.
- Prove rapidez nas consultas aos dados.

Ainda Kimball e Ross (2013, p. 7) indicam que a simplicidade para uma modelagem dimensional é muito importante para os usuários. Eles devem entender com facilidade todas as informações necessárias. As nomenclaturas e informações devem ser comumente utilizadas nas áreas de negócio. A simplicidade ao criar também garante que os Softwares entreguem as informações de forma mais rápida para os usuários e com o menor uso de recurso computacional possível.

A modelagem dimensional é composta por três elementos importantes: tabelas de dimensão, tabelas fato e métricas.

Figura 3 – Exemplo de tabelas em uma modelagem dimensional



Fonte: Howston (2008, p. 31).

2.2.3.1 Tabela dimensão

A tabela de dimensão contém colunas que representam as dimensões/visões e atributos/características do negócio. Adamson (2010, p. 10) afirma que as tabelas de dimensões servem para fornecer o contexto necessário para o estudo dos fatos. Em consultas e relatórios, as dimensões serão usadas para especificar como os fatos serão agregados.

Adamson (2010, p. 29 a 30) complementa:

Um conjunto bem desenvolvido de tabelas de dimensões fornece recursos analíticos poderosos e diversos. As dimensões fornecem informações contextuais, sem as quais os relatórios não teriam sentido. O design de dimensão bem-sucedido depende do uso adequado de chaves, do desenvolvimento de um conjunto de colunas de dimensão ricamente detalhado e da rejeição da necessidade de economizar espaço.

Figura 4 – Exemplo de uma tabela de dimensão

Product Dimension
Product Key (PK)
SKU Number (Natural Key)
Product Description
Brand Name
Category Name
Department Name
Package Type
Package Size
Abrasive Indicator
Weight
Weight Unit of Measure
Storage Type
Shelf Life Type
Shelf Width
Shelf Height
Shelf Depth
...

Fonte: Kimball (2013, p. 13).

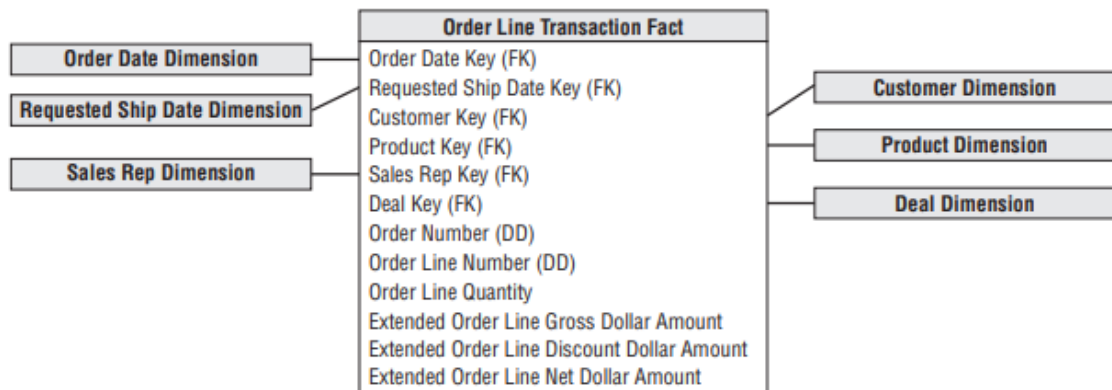
2.2.3.2 Tabela fato

A tabela de fatos, em um modelo dimensional, armazena as medições de desempenho resultantes dos eventos de processos de negócios de uma organização. Segundo Kimball e Ross (2013, p. 10) cada linha em uma tabela de fatos corresponde

a um evento de medição. Os dados em cada linha estão em um nível específico de detalhe, conhecido como granulação.

Adamson (2010, p. 286) afirma que um dos princípios básicos da modelagem dimensional é que todas as linhas de medição em uma tabela de fatos devem estar no mesmo grão, isso garante que as medições não sejam inadequadamente contadas duas vezes.

Figura 5 – Exemplo de um modelo entidade relacionamento de uma tabela fato



Fonte: Kimball (2013, p. 169).

Machado (2004, p. 141 e 142) elucida que:

Os elementos que participam de um fato, seja qual for ele, no mínimo são estes quatro:

- Onde aconteceu o fato.
- Quando aconteceu o fato.
- Quem executou o fato.
- O que é objeto do fato.

2.2.3.3 Métricas

As métricas ou medidas são as representações numéricas de uma transação presente na tabela fato. Elas podem estar prontas, obtidas no evento ou calculadas (MACHADO 2004, p. 127).

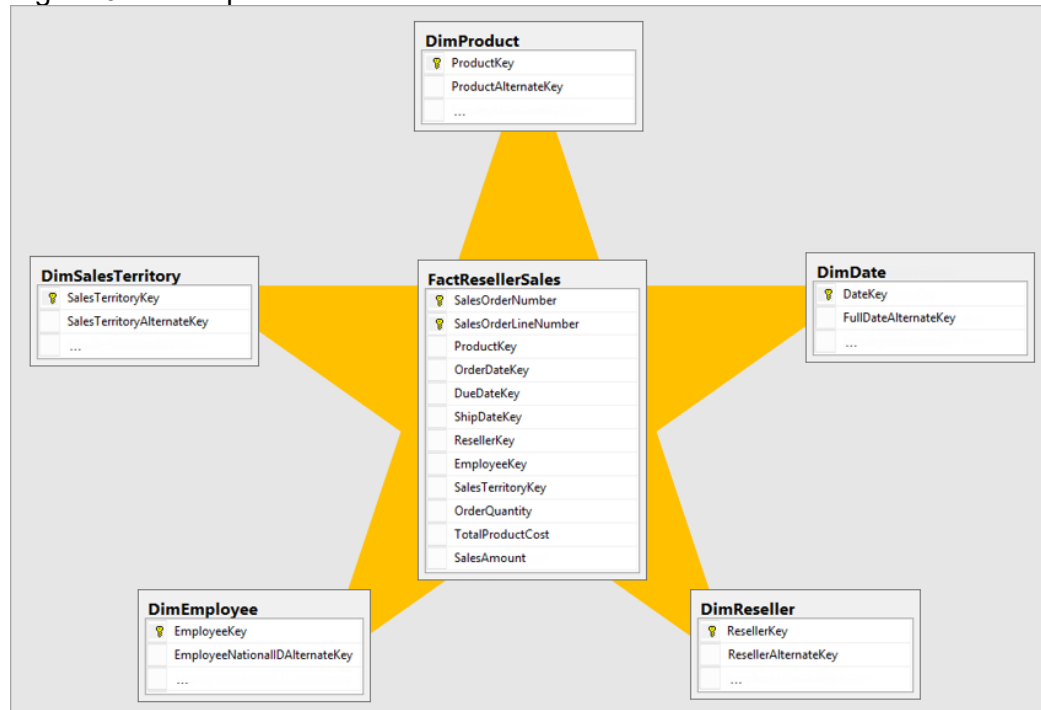
2.2.3.4 Modelo estrela

Segundo Machado (2004, p. 28 e 29) modelo estrela é a representação de processos de negócio e do relacionamento dos participantes desse processo ao longo do tempo.

Adamson (2010, p. 11) conceitua:

Em um esquema em estrela, cada tabela de dimensão recebe uma chave substituta. Esta coluna é um identificador exclusivo, criado exclusivamente para o armazém de dados. Chaves substitutas são atribuídas e mantidas como parte do processo que carrega o esquema em estrela. A chave substituta não tem significado intrínseco; normalmente é um número inteiro. A chave substituta é a chave primária da tabela de dimensões.

Figura 6 – Exemplo visual de um modelo estrela



Fonte: Microsoft (2019, p. 1).

2.3 FRONT-END DE UMA ARQUITETURA DE BI

Esta seção aborda conceitos de front-end de uma arquitetura de BI. Será exposto conceitos de Cubos OLAP (*Online Analytical Processing*) e suas operações. Definição sobre KPI (*Key Performance Indicators*) e sobre *Dashboards* (Painéis de Indicadores).

2.3.1 OLAP

OLAP (*Online Analytical Processing*) ou Processamento Analítico Online, compreende princípios da modelagem dimensional a fim de prover dados para sistemas analíticos com otimização para consultas. Conforme Adamson (2010, p. 4 e 5) explica que as análises podem acessar grandes volumes de dados, não apenas os dados individuais. Outra característica importante, citada por Adamson (2010, p. 4 e

5) é que deve compreender, também, a preservação de dados históricos mesmo que nos sistemas transacionais alterem ou excluam valores.

2.3.1.1 Cubos OLAP

Os cubos implementam os benefícios do OLAP em uma ferramenta para os analistas e gestores aplicarem suas análises e descobertas em dados.

Microsoft (2019, p. 1) entende que:

Um cubo OLAP é uma estrutura de dados que supera as limitações dos bancos de dados relacionais, proporcionando rápida análise de dados. Os cubos podem exibir e somar grandes quantidades de dados enquanto fornecem aos usuários acesso pesquisável a quaisquer pontos de dados. Dessa forma, os dados podem ser acumulados, segmentados e incluídos, conforme necessário, para lidar com a maior variedade de perguntas relevantes à área de interesse de um usuário.

Os cubos OLAP permitem que os usuários mudem as perspectivas e interajam com os dados (ADANSON, 2010, p. 53). Para executarem essas interações os cubos OLAP implementam algumas operações. Machado (2004, p. 112 e 113) entende que as operações podem ser: *Drill Down*, *Roll Up*, *Drill Across*, *Drill Through*, *Slice and Dice* e *Pivot*.

2.3.1.1.1 Drill Down e Roll Up

Para Machado (2004, p. 112 e 113) são operações para movimentar a visão dos dados ao longo das duas hierarquias. *Drill Down* ocorre quando se aumenta o nível de detalhe, diminuindo a granularidade. *Roll up* ou *Drill Up* ocorre quando se aumenta o nível de granularidade, diminuindo o nível de detalhamento.

2.3.1.1.2 Drill Across e Drill Through

De acordo com Machado (2004, p. 115 e 116) *Drill Across* ocorre quando o usuário passa para um outro nível da hierarquia na mesma dimensão.

Drill Through no entendimento de Machado (2004, p. 115 e 116) compreende quando o usuário passa de uma informação contida de uma dimensão para outra.

2.3.1.1.3 Slice and Dice e Pivot

Para Machado (2004, p. 116 e 117) *Slice and Dice* é definido como “[...] a redução do escopo dos dados em análise, além de mudar a ordem das dimensões, mudando desta forma a orientação segundo a qual os dados são visualizados”. Machado (2004, p. 116 e 117) define que *Slice* é a operação que corta o cubo e mantém a mesma perspectiva da visualização. *Dice* é a mudança da perspectiva da visão.

Machado (2004, p. 116 e 117) apresenta *Pivot* como uma troca da posição entre as linhas e as colunas em uma tabela.

2.3.2 KPI

KPI acrônimo de (*Key Performance Indicators*) ou Indicadores Chave de Desempenho. São instrumentos usados, em sua maioria, por gerentes, a fim de visualizarem se seus negócios estão no caminho certo (MARR, 2012, p. 25).

Rasmussen, Bansal e Chen (2009, p. 23) indicam que KPI são métricas, mas uma métrica não será necessariamente um KPI.

Rasmussen, Bansal e Chen (2009, p. 24) acrescentam:

Normalmente, os KPIs são representados como uma proporção (porcentagem) de um valor real comparado a um valor de destino predefinido. Eles geralmente são exibidos como, ou acompanhados por um símbolo gráfico, como um semáforo, para facilitar aos usuários identificarem instantaneamente se estão no alvo, acima ou abaixo.

Aspin (2015, p. 3) apresenta *scorecards* como uma coleção de KPIs. Podendo estar relacionados a uma área de negócios ou uma visão de várias áreas. Acrescenta que os *Scorecards* podem fazer parte de um *Dashboard*.

Aspin (2015, p. 19) destaca alguns pontos que um KPI deve conter:

- Um valor: valor de uma medida/métrica.
- Uma meta: um valor que deveria ser atingido.
- Um status: um símbolo que indica se o KPI está alinhado com a meta.
- Uma tendência: apresenta uma tendência das medidas desse KPI ao longo de um período.

2.3.3 Dashboards

Few (2007, p. 34) tem a definição para *Dashboard* como sendo:

Uma exibição visual das informações mais importantes necessárias para alcançar um ou mais objetivos, consolidados e organizados em uma única tela para que as informações possam ser monitoradas praticamente em um piscar de olhos.

Gartner (2020, p. 1) destaca que os *Dashboards* agregam indicadores de desempenho (KPIs), tornando possível serem utilizados rapidamente por todos os usuários antes de uma eventual exploração adicional por meio de ferramentas de análise.

Brath e Peters (2004, p. 1) conceituam:

Dashboards e visualização são ferramentas cognitivas que melhoram sua “extensão de controle” sobre muitos dados corporativos. Essas ferramentas ajudam as pessoas a identificar visualmente tendências, padrões e anomalias, raciocinar sobre o que veem e ajudar a orientá-las em direção a decisões efetivas. Como tal, essas ferramentas precisam aproveitar as capacidades visuais das pessoas. Com a prevalência de scorecards, painéis e outras ferramentas de visualização agora amplamente disponíveis para usuários corporativos revisarem seus dados, a questão do design de informações visuais é mais importante do que nunca.

Os *Dashboards* apresentam suas informações através de combinação de textos e gráficos em uma tela. Os gráficos nas maiorias das vezes têm o maior destaque pois eles comunicam de maneira mais eficiente e abrangente que textos (FEW, 2007, p. 35). Few (2012, p. 61) acrescenta “a visão, de todos os sentidos, é o nosso canal mais poderoso e eficiente para receber informações ao nosso redor. Aproximadamente 70% dos receptores sensoriais em nosso corpo são dedicados à visão”.

Few (2007, p. 34) considera que os *Dashboards* exibem informações para que seja possível alcançar objetivos específicos e que para sua elaboração é necessário coleta de informações provenientes de diversas fontes e relacionadas a várias atividades do negócio. As informações em *Dashboards* não são apenas para executivos ou gerentes. São para qualquer um que possui objetivos para cumprir. As informações necessárias podem ser um conjunto de KPIs, mas também podem conter outros elementos como Aspin (2015, p. 3) indica: Scorecards, textos, gráficos, tabelas e mapas.

Na visão de Few (2007, p. 34) um *Dashboard* deve ser apresentado em uma única tela. As informações devem ser simples e disponíveis aos olhos do usuário, para que possam ser vistas de uma só vez. Few (2007, p. 34) acredita que os *Dashboards* são usados para monitorar as informações rapidamente. O *Dashboard* tem que ser apto para indicar que algo requer atenção e pode exigir ação. “O painel executa seu trabalho principal se indicar apenas com um olhar que você deve agir”. Few (2007, p. 36) define que “*Dashboards* têm mecanismos de exibição pequenos, concisos, claros e intuitivos”.

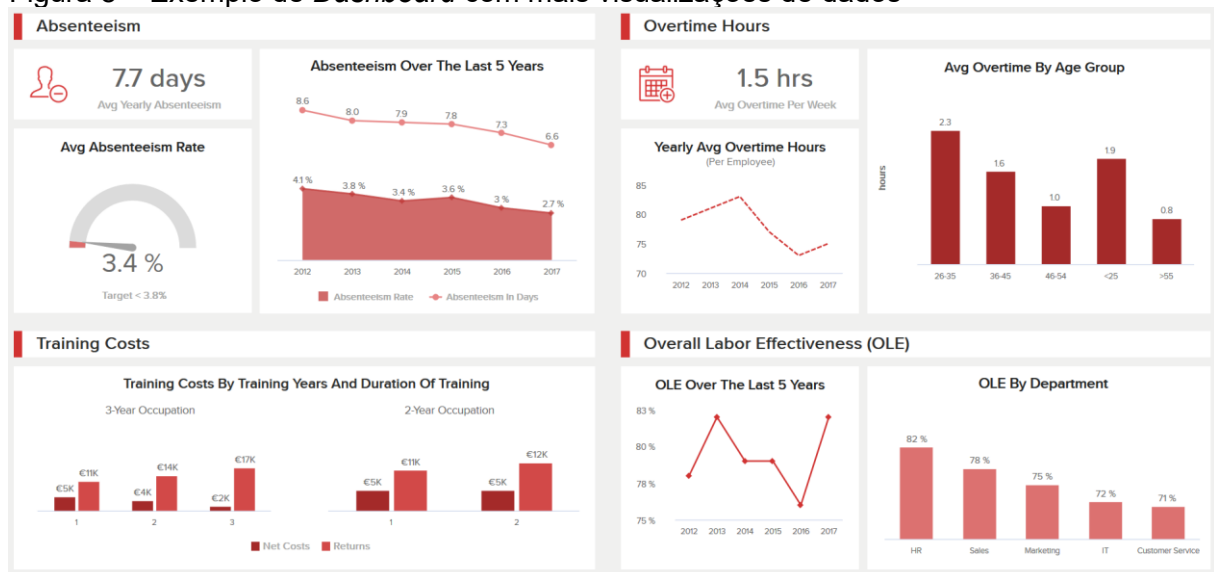
Figura 7 – Exemplo de *Dashboard* de vendas



Fonte: Few (2006, p. 177).

Few (2007, p. 36) entende que os *Dashboards* são personalizáveis. As informações apresentadas para atender aos requisitos de uma determinada pessoa, grupo ou função. Few (2007, p. 36) acrescenta que “Um *Dashboard* é um tipo de exibição, uma forma de apresentação, não um tipo específico de informação ou tecnologia”. Gartner (2020, p. 1) sustenta que: “Os *Dashboards* ajudam a melhorar a tomada de decisões, revelando e comunicando informações contextuais sobre o desempenho dos negócios”.

Figura 8 – Exemplo de *Dashboard* com mais visualizações de dados



Fonte: Datapine (2020, p. 1).

Wexler, Shaffer e Cotgreave (2017, p. 14) cita alguns exemplos práticos de aplicação de *Dashboards*:

- Uma exibição interativa que permite que as pessoas explorem reivindicações de remuneração de trabalhadores por região, setor.
- Um PDF mostrando as principais medidas enviado por e-mail a um executivo toda segunda-feira de manhã.
- Uma tela grande na parede que mostra as estatísticas da central de suporte em tempo real.
- Um aplicativo móvel que permite aos gerentes de vendas revisar o desempenho em diferentes regiões e comparar as vendas do ano atual com o anterior ano.

2.4 STORYTELLING DE DADOS

Esta seção aborda conceito de *Storytelling* e de *Storytelling* em dados. Conceitos e técnicas como: Contexto, visualização de dados e narrativa de dados e a importância em reduzir a carga cognitiva e realçar o que realmente importa nos *Dashboards*.

2.4.1 Storytelling de dados

Xavier (2015, p. 1 e 2) apresenta uma definição pragmática para Storytelling: “é a técnica arte de elaborar e encadear cenas, dando-lhes um sentido envolvente que capte a atenção das pessoas e enseje a assimilação de uma ideia central”. Antônio Núñez (*apud* XAVIER, 2015, p. 2) define *storytelling* como “uma ferramenta de comunicação estruturada em uma sequência de acontecimentos que apelam a nossos sentidos e emoções”. Para Berengueres, Sandell e Fenwick (2019, p. 14) “*Storytelling* é um conto de eventos”.

Duarte (2019, p. 9) afirma que “Contar histórias faz o cérebro acender de uma maneira que nenhuma outra forma de comunicação”. Knafllic (2017, p. 166) acrescenta que “A história é uma estrutura testada e aprovada; os seres humanos têm se comunicado com histórias ao longo das eras. Podemos usar essa poderosa ferramenta para nossas comunicações comerciais”.

Ryan (2018, p. 85) complementa afirmando que:

Há até evidências dos efeitos cognitivos de *storytelling* em nossa neurologia. É uma maneira central de aprender, lembrar e comunicar informações - o que tem implicações importantes quando o objetivo de uma visualização ou de uma história de dados visuais é preparar os tomadores de decisão de negócios para deixar uma apresentação de dados com uma história na cabeça que os ajude a lembrar sua mensagem e aja.

Ao aplicar storytelling para apresentar dados Dykes (2020, p. 15) acredita que é possível ampliar o poder que os números têm. Ao criar uma narrativa, aplicando *storytelling*, o público-alvo fica mais envolvido e tende a agir por mudanças, pois conecta os pontos envolvidos com mais facilidade. Acrescenta que “Quando você apresenta suas ideias como *storytelling* de dados, é mais provável que você influencie decisões e conduza ações que levam à criação de valor”. Davenport (2015, p. 1) afirma que *Storytelling* “fornece contexto, insight, interpretação - tudo o que torna os dados significativos e analíticos mais relevantes e interessantes”.

Dykes (2020, p. 16) considera que apenas fatos e números não influenciarão as decisões e nem levarão as pessoas a agir, apenas com *Storytelling* de dados bem elaborados. Em concordância Daniel Kahneman (*apud* DYKES, 2020, p. 298) “Ninguém nunca tomou uma decisão por causa de um número. Eles precisam de uma história”. Davenport (2015, p. 1) lembra que o objetivo geralmente é mudar a maneira de como alguém toma uma decisão. Com *Storytelling* tentamos “convencer, inspirar

confiança e liderar mudanças”. Conforme apresentado por Stephen Few (*apud* DYKES, 2020, p. 15) “Os números têm uma história importante para contar. Eles confiam em você para lhes dar uma voz clara e convincente”.

Dykes (2020, p. 16) compreende que *Storytelling* de dados combina três elementos: dados, narrativa e visual e em conjunto eles ajudam a explicar, esclarecer e envolver o público. Davenport (2015, p. 1) destaca que precisamos de simplicidade nas representações, não podemos apresentar todos os dados pois causaria confusão para o espectador ou desfrutador dos dados. Steele e Illinsky (2010, p. 16) explica que aplicar *Storytelling* de dados é associado ao processo de colocar a visualização de dados em um contexto.

Figura 9 – União de elementos: Narrativa, visuais e dados



Fonte: Dykes (2020, p. 32).

2.4.2 Contexto e público-alvo

Esta seção apresenta conceitos referente ao contexto em *storytelling* de dados e importância do público-alvo.

2.4.2.1 Contexto

Ao desenvolver um *Dashboard*, utilizando de conceitos de *Storytelling* de dados, deve-se entender o contexto em que será aplicado. Segundo Thomas e Millar (*apud* BERENGUERES; SANDELL; FENWICK, 2019, p. 150):

O contexto é um dos fatores mais fortes de influência no julgamento e na tomada de decisão. Isso ocorre porque nosso cérebro usa todos os sentidos para coletar informações sobre o ambiente e criar uma história coerente do que está acontecendo.

Steele e Illinsky (2010, p. 16) explica que as histórias de visualização iniciam com uma pergunta que direciona o espectador ao assunto e contexto em que os dados são mais importantes. Steele e Illinsky (2010, p. 16) ressalta que “Muitas das partes principais de uma história estão relacionadas como parte do processo de colocar a visualização em um contexto”. Steele e Illinsky (2010, p. 16) afirma que o contexto apresenta informações que elucidam questões como:

- Quais dados estamos vendo?
- Em que período esses dados existem?
- Que eventos ou variáveis notáveis influenciaram os dados?

2.4.2.2 Público-alvo

Aplicamos *Storytelling* de Dados com eficiência quando temos conhecimento do nosso público-alvo. Knafllic (2017, p. 20) destaca que “[...] é importante entender bem quem é seu público e como ele o interpreta. Isso pode ajudá-lo a identificar pontos em comum que o ajudarão a garantir que sua mensagem seja ouvida”.

Ao respeito Few (2009, p. 5) afirma:

Conheça o seu público. Não é segredo que devemos ter um senso das pessoas com quem estamos falando e personalizar a apresentação para trabalhar para elas. Não falo com médicos e enfermeiros da mesma forma que falo com pessoas que trabalham em um departamento de marketing. Posso estar contando a mesma história para os dois grupos, mas escolho minhas palavras e imagens de maneira diferente. Por quê? Porque eu quero expressar a história em termos familiares.

Knafllic (2017, p. 21) entende que quanto mais específico for um público-alvo mais fácil será de se comunicar pois iremos atingir nosso objetivo de uma maneira mais ampla. Quando o público-alvo é distinto, pessoas diferentes e com necessidades distintas, acabamos nos colocando em uma posição na qual não conseguimos nos comunicar com todas de uma maneira eficiente.

Knafllic (2017, p. 21) explica:

Às vezes, isso significa criar diferentes comunicações para diferentes públicos. Identificar o tomador de decisão é uma maneira de restringir seu público. Quanto mais você souber sobre seu público, mais bem posicionado

estará para saber como ser entendido por ele e para estabelecer uma comunicação que satisfaça a necessidade dele e a sua.

2.4.3 Visualização de dados

Após entender o contexto e conhecer o público-alvo, Knafllic (2017, p. 188) afirma que conheceremos o que nosso público-alvo necessita saber ou fazer. Knafllic (2017, p. 26) acrescenta que agora podemos considerar os dados e identificar: “Quais dados estão disponíveis e que ajudarão a apresentar minha ideia?”. Knafllic (2017, p. 26) acredita que: “Dados se tornam a evidência que corrobora a história que você vai construir e contar”. Yau (2013, p. 261) considera que: “A visualização geralmente é enquadrada como um meio de contar histórias. Os números são o material de origem e os gráficos são como você descreve a fonte”.

Knafllic (2017, p. 8) esclarece:

Uma visualização de dados eficaz pode significar a diferença entre o sucesso e o fracasso na hora de comunicar as constatações de seu estudo, levantar dinheiro para sua organização sem fins lucrativos, apresentar informações para seus diretores ou simplesmente mostrar sua ideia para seu público.

No entender de Few (2014, p. 1) visualização de dados é “a exibição gráfica de informações abstratas para dois propósitos: análise de dados e comunicação”. Complementa ainda que: “Histórias importantes que vivem em nossos dados e a visualização de dados é um meio poderoso para descobrir e entender essas histórias e, em seguida, apresentá-las aos outros”. Few (2014, p. 1) comenta que “visualização de dados é eficaz porque altera o equilíbrio entre percepção e cognição para aproveitar ao máximo as habilidades do cérebro”.

Few (2009, p.12) define visualização de dados como:

Um termo genérico que abrange todos os tipos de representações visuais que dão suporte à exploração, exame e comunicação de dados. Qualquer que seja a representação, desde que seja visual, e o que ela representa, desde que seja informação.

Os elementos de visualização de dados nos *Dashboards* podem ser: tabelas, números, palavras ou textos, símbolos, ícones, figuras, mapas e gráficos. Entre esses elementos, os gráficos são os mais utilizados pois são os que melhor resumem os dados para serem apresentados. Tufte (2013, p. 11) considera que: “Os gráficos de dados exibem visualmente as quantidades medidas por meio do uso combinado de

pontos, linhas, um sistema de coordenadas, números, símbolos, palavras, sombreamento e cor”.

Tufte (2013, p. 11) acrescenta que:

Os gráficos de dados modernos podem fazer muito mais do que simplesmente substituir pequenas tabelas estatísticas. Na melhor das hipóteses, os gráficos são instrumentos para raciocinar sobre informações quantitativas. Frequentemente, o mais eficaz é descrever, explorar e resumir um conjunto de números, mesmo um conjunto muito grande, é ver fotos desses números.

Tufte (2013, p. 14) entende que os gráficos devem comunicar com clareza, precisão e eficiência todos os dados que eles revelam. Tufte (2013, p. 14) ainda afirma que: “De fato, os gráficos podem ser mais precisos e reveladores do que os cálculos estatísticos convencionais”.

Tufte (2013, p. 51) define princípios de para uma excelência na apresentação de dados com gráficos:

- Bem projetados, aplicando bom design e apresentando boa informação.
- Comunicar ideias complexas com precisão e clareza.
- Oferecer ideias para o público-alvo no menor tempo possível e com menos informação visual possível.
- Ser possível de analisar múltiplas variáveis ao mesmo tempo.
- Sempre apresentar a verdade sobre os dados.

Knafllic (2017, p. 43) esclarece que existem muitos tipos de gráfico, mas que conseguimos atingir nossos objetivos de comunicar com clareza com poucos tipos. Acrescenta que os tipos de gráficos que são mais frequentemente usados são das seguintes categorias: pontos, linhas, barras.

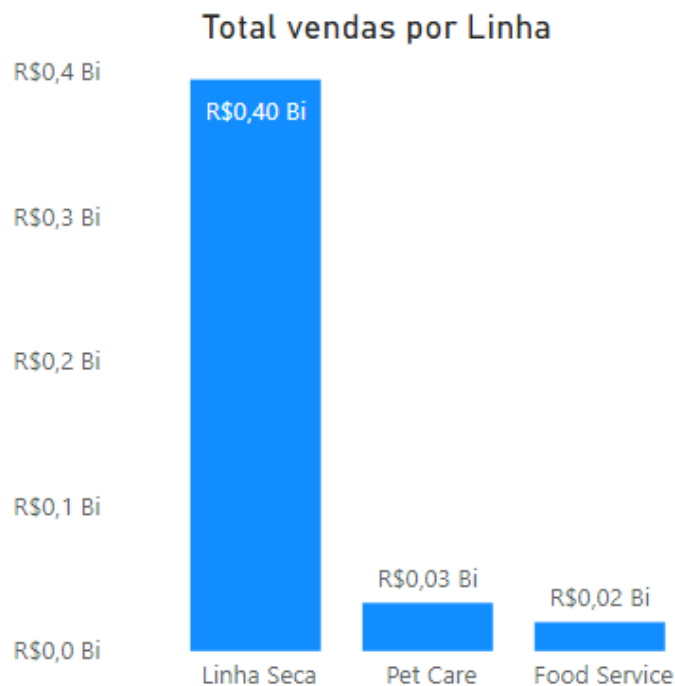
2.4.3.1 *Gráficos de barra*

Os gráficos de barras, podendo ser verticais ou horizontais, segundo Few (2006, p. 130) “são projetados para exibir várias instâncias, em vez de uma única instância, de uma ou mais medidas principais”. Complementa que os gráficos de barras são ótimos para “exibir medidas associadas a itens em uma categoria, como regiões ou departamentos”.

Few (2006, p. 131) destaca que os gráficos de barras possuem três tipos de escalas categóricas, que são: escalas nominais, escalas ordinais e escalas de intervalo.

Escalas nominais: são gráficos que apresentam dados de uma categoria em comum, porém sem relação, exemplo: regiões, departamentos e linhas de produtos. Esses itens não possuem uma ordenação específica e não representam valores quantitativos.

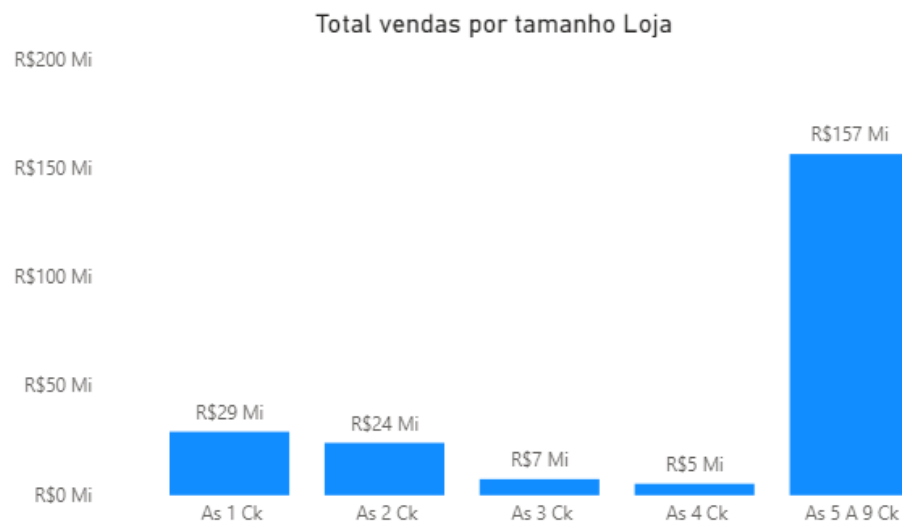
Figura 10 – Exemplo de visualização de dados: Gráfico de barras



Fonte: elaborado pelo autor.

Escalas ordinais: são semelhantes às nominais, porém possuem uma ordenação intrínseca, como exemplo escala de grandeza pequeno, médio, grande e classificações como: péssimo, ruim, regular, bom e ótimo. Da mesma forma que as escalas nominais eles ainda não correspondem a valores quantitativos.

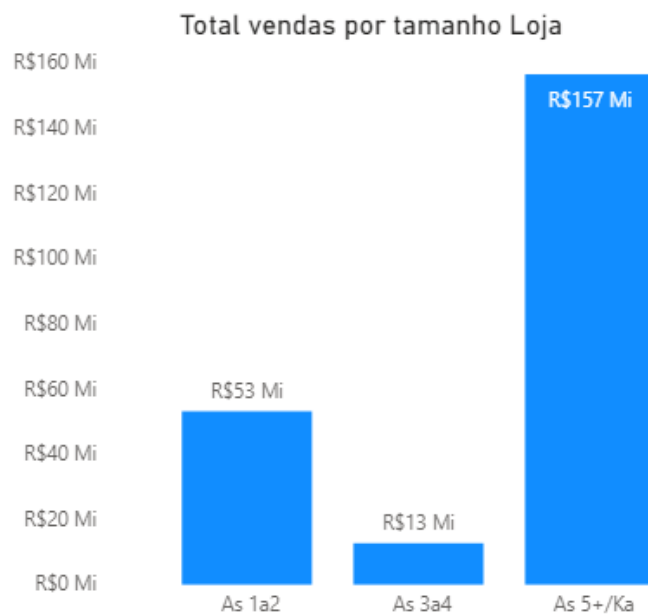
Figura 11 – Exemplo de visualização de dados: Gráfico de barras



Fonte: elaborado pelo autor.

Escalas de intervalo: semelhante as escalas ordinais, porém representam valores quantitativos. Essa escala geralmente se inicia em uma escala quantitativa e na sequência é agrupada convertendo em uma escala categórica com pequenos intervalos exemplo: intervalo valores de pedidos de R\$ 0,00 à R\$ 100,00 correspondem a uma categoria é de R\$ 100 a R\$ 1.000,00 correspondem a outra.

Figura 12 – Exemplo de visualização de dados: Gráfico de barras



Fonte: elaborado pelo autor.

Gráfico de barra empilhado é uma variação de gráficos de barra que permite empilhar múltiplas variáveis ou categorias de dimensões diferentes. Segundo Few

(2006, p. 135) ele é recomendado para mostrar as subcategorias de uma parte do todo.

Figura 13 – Exemplo de visualização de dados: Gráfico de barras

Total vendas por categoria de cliente

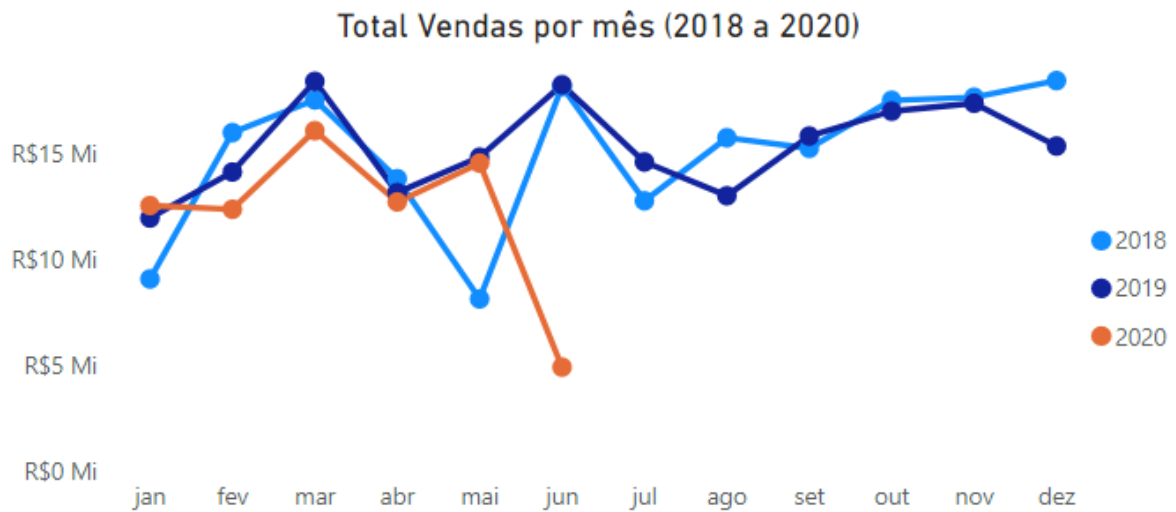


Fonte: elaborado pelo autor.

2.4.3.2 Gráficos de Linha

Conforme Few (2006, p. 139) “os gráficos de linha fazem um trabalho excepcional ao revelar a forma do movimento dos dados para cima e para baixo de um valor para o próximo, especialmente à medida que muda ao longo do tempo”. Acrescenta que esse gráfico é ideal para “ênfatizar padrões nos dados, como tendências, flutuações, ciclos, taxas de mudança e como dois conjuntos de dados variam em relação um ao outro”. Few (2006, p.139) elucida que os gráficos de linha costumam ser a melhor forma de apresentar uma visão geral rápida de uma série temporal em *Dashboards*.

Figura 14 – Exemplo de visualização de dados gráfico de linha



Fonte: elaborado pelo autor.

2.4.3.3 Gráficos de ponto

De acordo com Few (2012, p. 88) esses gráficos são utilizados para apresentar dados quantitativos. Utiliza-se de pontos para plotar cada valor no gráfico.

Figura 15 – Exemplo de visualização gráfico de pontos

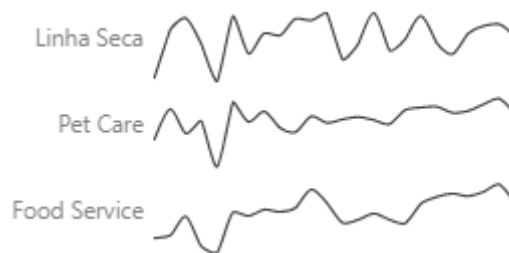


Fonte: elaborado pelo autor.

2.4.3.4 Sparkline (mini gráficos).

Para Few (2006, p. 140) são gráficos que fornecem um contexto rápido de tendência de uma determinada série temporal. Esses gráficos são menores e sem detalhes visuais, apenas para dar uma noção de tendências, pico, média ou vale de uma determinada série de dados temporal.

Figura 16 – Exemplo de visualização de texto e número

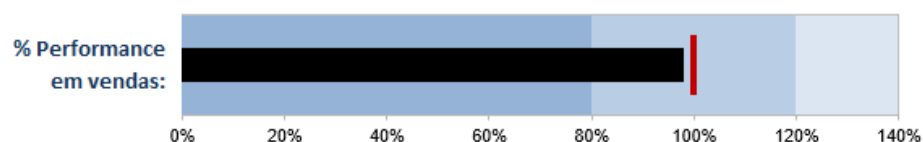


Fonte: elaborado pelo autor.

2.4.3.5 Bullet Charts

Few (2006, p. 125) explica que os gráficos *Bullet Charts* foram criados especificamente para *Dashboards* a fim de preservar espaço e mesmo assim conseguir entregar de forma plena a informação proposta. A sua concepção veio para substituir os visuais do tipo gauge que tentam simular velocímetros ou termômetros, mas acabam não fornecendo um bom aproveitamento de espaço e seu aspecto visual destoa dos demais gráficos. O modelo consiste em duas barras, uma para apontar uma tendência ou estágio desejado, outra barra ao centro demonstrando o valor alcançado e por último um indicador de alvo ou objetivo.

Figura 17 – Exemplo de visualização de texto e número



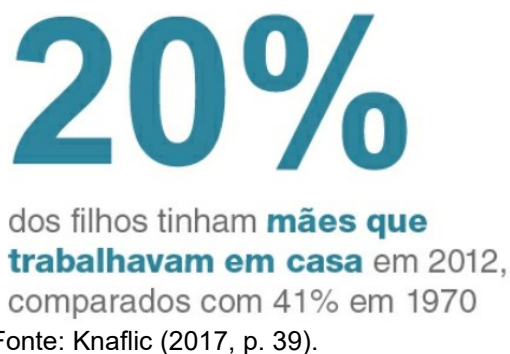
Fonte: elaborado pelo autor.

2.4.3.6 Textos e cartões

Texto simples e números também são visualizações de dados importantes. Knaflic (2017, p. 38) indica que “Quando você tem apenas um número ou dois para

compartilhar, um simples texto pode ser uma ótima maneira de comunicar”. Destaca ainda que é ideal usar textos simples de forma destacada, fora de tabelas e gráficos, pois desta forma perderia o destaque.

Figura 18 – Exemplo de visualização de texto e número



2.4.3.7 Tabelas

As tabelas complementam as nossas visualizações trazendo informações mais genéricas e com isso atingindo anseios de um público maior. Knaflit (2017, p. 40) afirma que “As tabelas interagem com nosso sistema verbal, o que significa que nós as lemos”. Ainda afirma que as tabelas são para comunicar para um público misto. Knaflit (2017, p. 40) sustenta que “Se você precisa comunicar várias unidades de medida diferentes, normalmente isso também é mais fácil com uma tabela do que com um gráfico”.

Figura 19 – Exemplo de uma tabela como elemento de visualização de dados

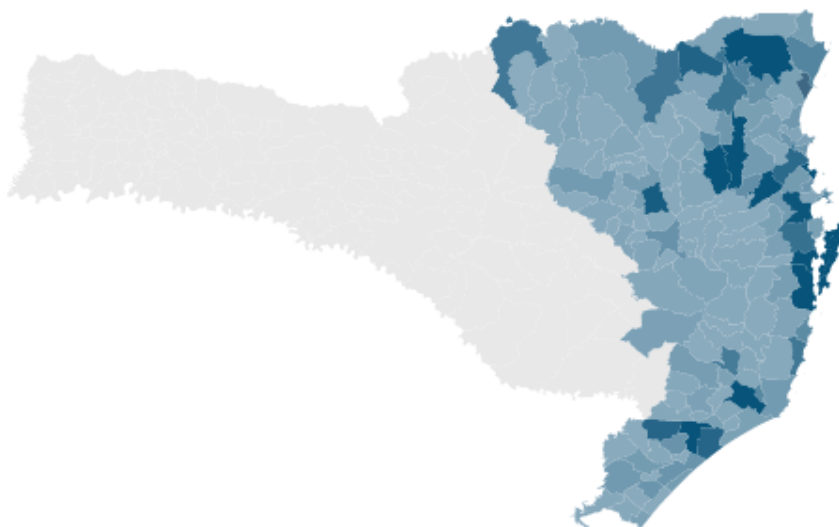
Mês	Food Service	Linha Seca	Pet Care	Total
junho	R\$1.463.318,46	R\$32.534.395,04	R\$2.541.818,04	R\$36.539.531,54
março	R\$1.340.142,89	R\$32.704.968,57	R\$2.028.602,06	R\$36.073.713,52
novembro	R\$1.840.153,36	R\$30.842.404,89	R\$2.480.682,08	R\$35.163.240,33
outubro	R\$1.609.775,05	R\$30.801.880,01	R\$2.230.423,41	R\$34.642.078,47
dezembro	R\$1.551.076,60	R\$30.149.223,59	R\$2.226.031,37	R\$33.926.331,56
setembro	R\$1.535.779,41	R\$27.459.622,43	R\$2.201.533,08	R\$31.196.934,92
fevereiro	R\$1.113.753,89	R\$26.773.762,96	R\$2.350.070,32	R\$30.237.587,17
agosto	R\$1.578.812,73	R\$24.921.303,39	R\$2.374.166,70	R\$28.874.282,82
julho	R\$1.493.002,45	R\$23.673.237,79	R\$2.321.933,82	R\$27.488.174,06
abril	R\$1.014.397,86	R\$23.948.971,60	R\$2.124.923,57	R\$27.088.293,03
maio	R\$930.897,15	R\$20.391.265,54	R\$1.757.545,37	R\$23.079.708,06
janeiro	R\$1.062.491,51	R\$18.107.162,63	R\$1.969.166,22	R\$21.138.820,36
Total	R\$16.533.601,36	R\$322.308.198,44	R\$26.606.896,04	R\$365.448.695,84

Fonte: elaborado pelo autor.

2.4.3.8 Mapas geográficos

Na concepção de Few (2006, p. 157) mapas geográficos apresentam os dados em uma associação de categoria e quantidade em um espaço físico. Como exemplo: Podemos ter um mapa onde sua área pode ter destaque das vendas realizadas. Auxilia a dar um contexto para algumas análises.

Figura 20 – Exemplo de uma tabela como elemento de visualização de dados



Fonte: elaborado pelo autor.

2.4.3.9 Ícones

De acordo com Few (2013, p. 153) os *Dashboards* podem incluir outras formas de visuais além de gráficos, tabelas e números a fim de enriquecer e melhorar a comunicação. Conforme descrito por Few (2013, p. 177) os ícones são imagens simplistas que expõe informações de forma clara e compreensível e que podemos incluir nos *Dashboards* para indicar três significados.

- Alerta: atenção.
- Direção: para cima e para baixo.
- Estado: ligado e desligado.

Few (2013, p. 177) acrescenta:

Muitas vezes é útil chamar a atenção para informações específicas em um painel. Isso é especialmente verdadeiro quando algo está errado e requer atenção. Um ícone que funciona como um alerta grita para o visualizador: "Ei, olhe aqui!" Para um ícone desempenhar bem essa função, ele precisa ser excepcionalmente simples e perceptível.

Figura 21 – Exemplo de uma tabela como elemento de visualização de dados

▲	Ano	Total Vendas	Total Vendas LM	% MoM	Total Vendas LY	% YoY	Total Vendas YTD
☒	2018	R\$180.822.860,57	162.312.336,68	11,40%			180.822.860,57
☒	2019	R\$184.625.835,27	187.720.551,49	-1,65%	180.822.860,57	2,10%	184.625.835,27
☒	janeiro	R\$12.009.914,93	18.510.523,89	✗ -35,12%	9.128.905,43	▲ 31,56%	12.009.914,93
☒	fevereiro	R\$14.187.327,19	12.009.914,93	✓ 18,13%	16.050.259,98	▼ -11,61%	26.197.242,12
☒	março	R\$18.475.284,65	14.187.327,19	✓ 30,22%	17.598.428,87	▲ 4,98%	44.672.526,77
☒	abril	R\$13.215.557,10	18.475.284,65	✗ -28,47%	13.872.735,93	▼ -4,74%	57.888.083,87
☒	maio	R\$14.891.860,90	13.215.557,10	✓ 12,68%	8.187.847,16	▲ 81,88%	72.779.944,77
☒	junho	R\$18.313.679,32	14.891.860,90	✓ 22,98%	18.225.852,22	▲ 0,48%	91.093.624,09
☒	julho	R\$14.661.557,12	18.313.679,32	✗ -19,94%	12.826.616,94	▲ 14,31%	105.755.181,21
☒	agosto	R\$13.064.835,65	14.661.557,12	✗ -10,89%	15.809.447,17	▼ -17,36%	118.820.016,86
☒	setembro	R\$15.888.447,69	13.064.835,65	✓ 21,61%	15.308.487,23	▲ 3,79%	134.708.464,55
☒	outubro	R\$17.063.730,78	15.888.447,69	✓ 7,40%	17.578.347,69	▼ -2,93%	151.772.195,33
☒	novembro	R\$17.437.832,27	17.063.730,78	✓ 2,19%	17.725.408,06	▼ -1,62%	169.210.027,60
☒	dezembro	R\$15.415.807,67	17.437.832,27	✗ -11,60%	18.510.523,89	▼ -16,72%	184.625.835,27
☒	2020	R\$73.516.274,56	83.969.996,30	-12,45%	91.093.624,09	-19,30%	73.516.274,56
	Total	R\$438.964.970,40	434.002.884,47	1,14%	271.916.484,66	61,43%	73.516.274,56

Fonte: elaborado pelo autor.

2.4.3.10 Aplicações para cada tipo de gráfico

Few (2012, p. 310) indica um resumo com os melhores tipos de gráfico para cada informação que deseja apresentar.

Quadro 1 – Aplicações de gráfico

	Pontos	Linhas	Barras
Comparação nominal: Uma comparação simples de valores para um conjunto de itens não ordenados (por exemplo, produtos ou regiões).	Utilizar quando não conseguir utilizar barras quando a escala não começar no zero	Evitar	Utilizar
Série de tempo: Os valores mostram como algo mudou através tempo (anual, mensal etc.).	Apenas quando os valores não foram coletados em intervalos de tempo consistentes	Utilizar	Utilizar
Ranqueamento: Os valores são ordenados por tamanho (descendente ou ascendente).	Utilizar quando não conseguir utilizar barras quando a escala não começar no zero	Evitar	Utilizar
Parte do todo: Os valores representam partes (proporções) de um todo (por exemplo, porções regionais de vendas totais).	Evitar	Utilizar para mostrar como as partes mudam com o passar do tempo	Utilizar

Desvio: A diferença entre dois conjuntos de valores (por exemplo, a variação entre despesas reais e orçadas).	Utilizar quando a escala quantitativa não começar no zero	Utilizar combinado com série de tempo	Utilizar, vertical quando combinado com série de tempo
Distribuição simples: Contagens de valores por intervalo do menor ao maior (por exemplo, contagens de pessoas por intervalos de idade de 10 anos cada).	Utilizar	Utilizar	Utilizar
Distribuição múltipla	Utilizar	Utilizar	Evitar
Correlação: Comparação de dois conjuntos de valores emparelhados (por exemplo, as alturas e pesos de várias pessoas) para determinar se há uma relação entre eles.	Utilizar	Evitar	Utilizar
Geoespacial: Os valores são exibidos em um mapa para mostrar sua localização.	Utilizar, a variação do tamanho dos pontos indica os valores	Utilizar para marcar rotas	Evitar

Fonte: elaborado pelo autor de acordo com Few (2012, p. 310).

2.4.3.11 Quando usar gráficos ou tabelas

Few (2012, p. 51) esclarece, com a tabela abaixo, quando devemos usar tabelas ou gráficos para apresentar os dados.

Quadro 2 – Quando usar gráficos ou tabelas

Use tabelas quando:	Use gráficos quando:
<ul style="list-style-type: none"> • O visual é usado para demonstrar valores individuais. • Será necessário comparar valores individuais. • Um detalhamento dos valores é requerido. • Os valores quantitativos incluem mais que uma unidade de medida. • Valores detalhados e sumarizados são incluídos. 	<ul style="list-style-type: none"> • A mensagem está contida na forma dos valores. Exemplo: padrões, tendências e exceções. • O visual será usado para revelar relacionamentos entre conjuntos inteiros de valores.

Fonte: elaborado pelo autor de acordo com Few (2012, p. 51).

2.4.4 Narrativas e *Storyboard* de dados

Nesta seção é apresentado conceitos de narrativa de dados em *Dashboard* e a aplicação de *Storyboard* como esboço inicial para desenvolvimento de *Dashboard*.

2.4.4.1 Narrativas de dados

A narrativa, conforme dicionário Oxford (*apud* SEGEL; HERR, 2010, p. 1) é definida como: “um relato de uma série de eventos, fatos etc., dados em ordem e com o estabelecimento de conexões entre eles”. Segel e Herr (2010, p. 1) acrescentam que a ideia principal dessa definição:

É a noção de uma cadeia de eventos causalmente relacionados. Histórias dessa Forma geralmente têm começo, meio e fim: uma introdução à situação, uma série de eventos que geralmente envolvem tensão ou conflito e uma resolução.

O uso de narrativas para explicar e apresentar ideias é usado a muito tempo. Segel e Herr (2020, p. 1) afirmam que:

Desde os tempos antigos, as pessoas tentam entender e formalizar os elementos da narrativa. Escritores desenvolveram tipologias de situações dramáticas e identificaram tramas comuns a muitas narrativas, como a jornada do herói.

Berengueres, Sandell e Fenwick (2019, p. 26) descrevem narrativa como “um conjunto de crenças, valores ou cosmovisão. Portanto, a narrativa escolhida interpreta a história (e consequentemente os dados / realidade subjacentes)”.

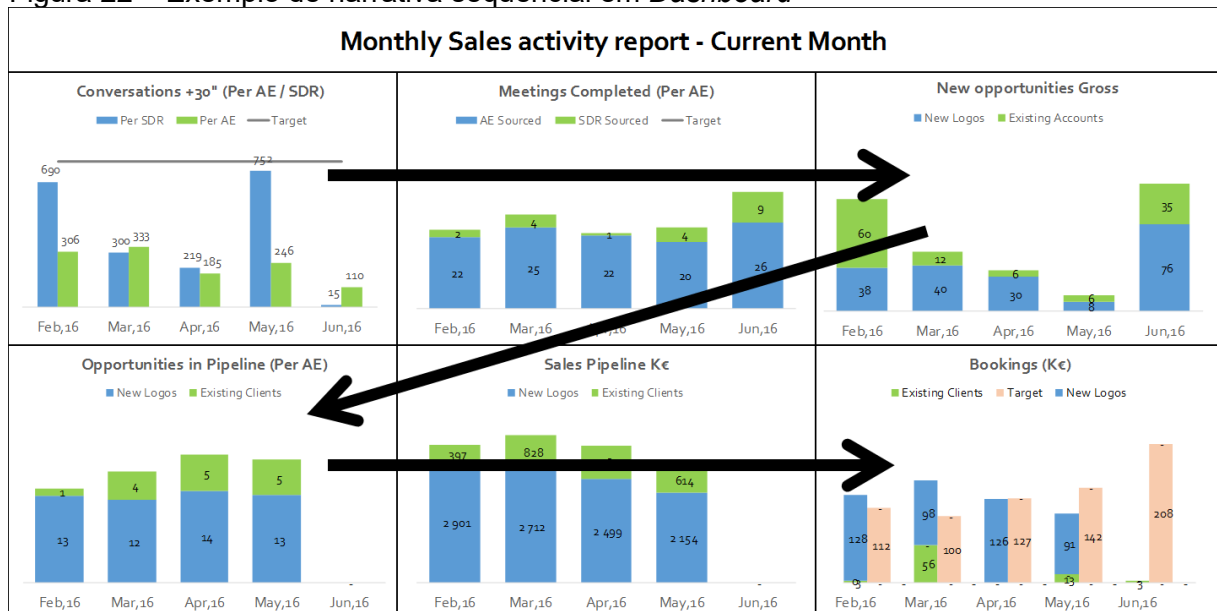
Few (2009, p. 8) elucida que:

A narrativa se desenrola de maneira serial, com começo, meio e fim. As histórias geralmente são contadas sequencialmente, revelando fatos apenas no tempo adequado. Podemos sugerir o que está por vir para criar antecipação, mas ordenamos as partes de uma história de uma maneira que corresponda à cronologia dos eventos, construa conceitos de componentes simples a conjuntos mais complexos, ou guie o público através de uma série de ideias relacionadas, afirmando o validar com evidência e conectá-la aos valores da audiência, um ponto de cada vez em ordem lógica.

Quando abordamos a narrativa na construção de *Dashboards* temos que pensar na ordem que o nosso público-alvo quer ver os dados. Entender o público-alvo é muito importante para a narrativa. Podemos ter um público que não se importa com o processo, quer apenas uma resposta e podemos ter um público na qual precisamos convencê-los de algo. Nesse caso a narrativa deve ser elaborada para que se consiga prender a concentração e facilitar o entendimento a fim de atingir o objetivo proposto (KNAFLIC 2017, p. 175). O mesmo autor Knafllic (2017, p. 175 e 176) acrescenta: “Uma coleção de números e palavras sobre determinado assunto, sem estrutura para organizá-los e dar significado, é inútil”.

Quando utilizamos meios de apresentar dados e informações como Dashboards, que não possuem uma sequência de conteúdos explícitos como slides, páginas ou capítulos temos que identificar maneiras de tornar a sequência evidente. Como teremos que contar histórias quantitativas em uma única tela temos que escolher métodos que tornem a sequência mais transparente para o público (FEW 2009, p. 10). O mesmo autor fundamenta que um dos métodos pode ser o mesmo usado nas histórias em quadrinhos que “sequenciam a narrativa em painéis organizados da esquerda para a direita, de cima para baixo. Outro método envolve numerar seções da tela sequencialmente ou sugerir caminhos visuais, como o uso de setas” (FEW, 2009, p.10).

Figura 22 – Exemplo de narrativa sequencial em *Dashboard*



Fonte: Prats (2017, p. 1).

2.4.4.2 Storyboard de dados

Antes de iniciarmos a implementação de um *Dashboard* podemos criar um esboço, um pequeno protótipo da narrativa e visualização de dados que serão aplicados em nosso *Dashboard* (KNAFLIC, 2015, p. 1). Para executar essa etapa utilizaremos uma técnica chamada *Storyboard*. David Becker (*apud* DYKES, 2020, p. 170) diz que: “O principal valor do *storyboard* é que ele o força a ter uma razão e uma abordagem consistente para tudo o que faz”. Dykes (2020, p. 170) acrescenta que Walt Disney Studio introduziu essa técnica nos anos 30 a fim de planejar a sequência de cenas das animações. Hoje essa técnica é adotada para várias outras situações.

Quando estruturamos a narrativa do nosso *Dashboard* com *storyboard* temos uma organização maior dos fluxos, etapas e uma visualização do que é mais importante a ser apresentado. Se pularmos essa etapa e não visualizarmos a narrativa perde-se uma oportunidade de construir uma estrutura de narrativa melhor, desenvolver um *Dashboard* com mais rapidez e que atenda melhor a necessidade do nosso público (DYKES, 2020, p. 171).

Dykes (2020, p. 170) elucida que:

O processo de *storyboard* ajuda você a criar uma narrativa mais forte e impactante, mas também pode economizar muito tempo. Em vez de desperdiçar uma quantidade excessiva de tempo gerando conteúdo que pode ou não ser incluído na sua história de dados, você pode identificar antecipadamente o que exatamente é necessário. Embora as histórias de dados mais simples possam não exigir muito *storyboard*, você não pode ignorar esse importante método de visualizar sua narrativa quando ela é complexa com vários pontos da narrativa.

Dykes (2020, p. 171 a 178) propõe 4 etapas para desenvolver um *Storyboard*:

- 1- Identificar a informação mais importante que você quer apresentar para o público.
- 2- Determinar o início do seu *Storytelling* de dados. Qual contexto conecta o público com seu *Dashboard* ou apresentação.
- 3- Desenvolva sua narrativa e os dados. Determine o que será aplicado para criar os insights.
- 4- Conclua apresentando uma proposta de solução.

3 MÉTODO DE PESQUISA

Nesta seção será abordado conceitos sobre a metodologia de pesquisa. Os temas abordados são: caracterização do tipo de pesquisa, etapas realizadas na pesquisa, proposta da solução para o problema definido e as delimitações desta pesquisa.

Silva e Menezes (2005, p. 10) explanam que a “metodologia tem como função mostrar a você como andar no “caminho das pedras” da pesquisa, ajudá-lo a refletir e instigar um novo olhar sobre o mundo: um olhar curioso, indagador e criativo”. Sampieri, Collado e Lucio (2013, p. 30) descrevem pesquisa como: “um conjunto de processos sistemáticos, críticos e empíricos aplicados no estudo de um fenômeno”.

3.1 CARACTERIZAÇÃO DO TIPO DE PESQUISA

Pesquisas podem abordar várias atividades, desde a realização de pesquisas eleitorais até a pesquisa científica que procuram ampliar o conhecimento humano. As pesquisas científicas podem ser classificadas em diferentes critérios (WAZLAWICK 2014, p. 21). Ainda segundo Wazlawick (2014, p. 21) acrescenta que:

É possível diferenciar tipos de pesquisa de acordo com sua natureza, objetivos ou procedimentos técnicos. Nem sempre um trabalho de pesquisa limita-se a um único tipo. Além disso, alguns tipos de pesquisa podem ser a base para outros.

Quadro 3 – Classificação do tipo de pesquisa

Critério	Abordagem adotada	Descrição
Quanto à natureza	Aplicada	Silva e Menezes (2005, p. 20) “objetiva gerar conhecimentos para aplicação prática e dirigidos à solução de problemas específicos. Envolve verdades e interesses locais”.
Quanto à abordagem	Qualitativa	Silva e Menezes (2005, p. 20) “A interpretação dos fenômenos e a atribuição de significados são básicas no processo de pesquisa qualitativa. Não requer o uso de métodos e técnicas estatísticas. O ambiente natural é a fonte direta para coleta de dados e o pesquisador é o instrumento-chave. É descritiva. Os pesquisadores tendem a analisar seus dados indutivamente. O processo e seu significado são os focos principais de abordagem”.
Quanto aos objetivos	Exploratória	Gil (2002, p. 41) “[...] estas pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições. Seu planejamento é, portanto, bastante flexível, de modo que possibilite a consideração dos mais variados aspectos relativos ao fato estudado”.

Quanto aos procedimentos técnicos	Bibliográfica	Gil (2002, p. 44) “A pesquisa bibliográfica é desenvolvida com base em material já elaborado, constituído principalmente de livros e artigos científicos”.
---	---------------	--

Fonte: elaborado pelo autor.

3.2 ETAPAS METODOLÓGICAS

Este trabalho é dividido em 7 etapas, todas descritas abaixo:

1. Definição do tema a ser pesquisado, que neste trabalho é aplicação de *Storytelling* de Dados em *Dashboards*.
2. O levantamento de bibliografia para entender conceitos e funcionamento de sistemas de *Business Intelligence* e conceitos de *Storytelling* de Dados e princípios de design aplicados a *Dashboards*.
3. Escolha das ferramentas de BI que serão utilizadas para o desenvolvimento da solução.
4. Entendimento do negócio da organização e levantamento de requisitos, contexto e público-alvo a fim de desenvolver a solução.
5. Utilização da modelagem dimensional, ETL e carregamento de um *Data Warehouse* para servir de base para desenvolvimento do *Dashboard*.
6. Conclusão apresentando um *Dashboard* com os conceitos de *Storytelling* de Dados.
7. Apresentar as possibilidades para trabalhos futuros.

3.3 PROPOSTA DE SOLUÇÃO

As etapas serão de acordo com o descrito abaixo:

- a) Análise de requisitos: levantamento de requisitos com a área de negócio a fim de identificar as métricas, KPIs e as dimensões que desejam analisar.
- b) Modelagem dimensional: definir um modelo dimensional de acordo com os requisitos.
- c) Aplicar ETL e carregar o DW: extrair, transformar e carregar os dados da modelagem em um *Data Warehouse*.
- d) Desenhar protótipo do Dashboard: definir um modelo de Dashboard aplicando princípios de design e *Storytelling*.

- e) Desenvolvimento: desenvolver os itens acima e uma versão final de um Dashboard aplicando conceitos de *Storytelling*.

3.4 DELIMITAÇÕES

Este trabalho tem como objetivo desenvolver um Dashboard, amparado com as técnicas e conceitos de BI, com dados de uma empresa fictícia e usando conceitos de *Storytelling* de dados com as seguintes limitações:

- a) Não serão feitas atualizações automáticas dos dados.
- b) Os dados utilizados serão fictícios a fim de preservar qualquer analogia a alguma empresa.
- c) Não será implementada uma solução de BI completa com documentação e metadados.
- d) Serão desenvolvidos apenas os requisitos necessários para validar o objetivo proposto.

4 MODELAGEM

Neste capítulo é apresentado os requisitos da área de negócio para desenvolvimento da modelagem dimensional. Essa etapa é fundamental para o projeto do *Data Warehouse*, onde também é mostrado os requisitos necessários para implementação dos *Dashboards* para monitoramento do negócio.

4.1 CONTEXTO DO NEGÓCIO

A aplicação prática desenvolvida neste trabalho será para a empresa estudo de caso (não será falado o nome por questões de confidência). A empresa atua no ramo de distribuição de produtos alimentícios, humano e pet. Os clientes são de vários ramos, exemplo: mercearias, supermercados, restaurantes, padarias, farmácias, atacados, conveniências, entre outros. Possui equipes de vendas externa que visitam os clientes para oferecer os produtos e realizar a venda. As entregas dos pedidos são realizadas por motoristas em caminhões.

Todo o processo de venda é feito pelo sistema ERP da empresa. Que armazena as informações no banco de dados SQL Server. O ERP já possui a maioria dos relatórios operacionais. Alguns são fornecidos por Excel através de planilhas eletrônicas. Os relatórios operacionais e analíticos não são adequados para apresentar os dados para a diretoria. Diante desse cenário o diretor da empresa apontou uma necessidade, ter acesso a informações para tomada de decisão de forma mais independente, ágil e segura.

4.2 REQUISITOS

Segundo Sbrocco (2012, p. 44) os requisitos indicam o que a solução deve fazer ou ter a fim de resolver um ou mais problemas dos usuários. Do ponto de vista dos desenvolvedores os requisitos são divididos em funcionais e não funcionais e devem passar por um processo a fim de se estabelecer o entendimento das necessidades dos usuários e/ou patrocinadores do projeto, sempre objetivando a busca de soluções.

Para realizar a análise de requisitos algumas etapas são importantes, tais como:

- Identificar quem são os patrocinadores.
- Acordar sobre a definição do que será resolvido.
- Entender os problemas ou necessidades.
- Estabelecer as delimitações da solução.
- Captar as restrições da solução.

4.2.1 Entrevista

Sbrocco (2012, p. 44) explica sobre a metodologia ciclo PDCA (em inglês, *Plan, Do, Check e Action*) muito empregada para melhoria de processo, criado por Walter Shewhart. Sbrocco (2012) explica que esse ciclo possui um método conhecido como 5W2H, que simboliza as palavras em inglês: *What* (O quê), *Who* (Quem), *Why* (Por quê), *When* (Quando), *Where* (Onde), *How* (Como) e *How Much* (Quanto custa). São perguntas básicas e importantes para o levantamento de requisitos. Para entender a necessidade do negócio e público-alvo, baseando-se no 5W2H, serão utilizadas as seguintes perguntas:

- a. “O que aconteceu?”
- b. “Quando aconteceu o fato?”
- c. “Onde aconteceu?”
- d. “Quem executou o fato?”
- e. “Quem demandou o fato?”
- f. “Qual é o objeto do fato?”
- g. “Quantas vezes ocorreu o fato?”
- h. “Por quê aconteceu o fato?”
- i. “Como aconteceu?”

Ao finalizar a entrevista e ter o entendimento do negócio é necessário conhecer o público-alvo. Neste trabalho será atendida uma necessidade do diretor da empresa sendo que ela já possui relatórios operacionais e de gestão, porém o diretor carece de informações que facilitem o entendimento do andamento das vendas. Especificamente ele necessita de um *Dashboard* que obtenha informações consolidadas e resumidas para ter uma reunião de alinhamento com o gerente de vendas.

Diante deste cenário temos algumas perguntas estratégicas, a fim de captar os pontos mais importantes que o diretor precisa para tomar suas decisões com rapidez e confiança.

- a. Quais são as métricas mais importantes?
- b. Qual o período para análise?
- c. Quais são os indicadores mais importantes?
- d. Existe alguma meta para os indicadores?
- e. Quais dados possuem relação para análises?

Após a definição das perguntas estratégicas, é possível definir as regras de negócio, modelagem e o *Dashboard*, conforme apresentado nas seções seguintes.

4.2.2 Levantamento das regras de negócio

Vazquez e Simões (2016, p. 103) definem regras de negócio como leis que orientam o mesmo e determinam como os processos devem funcionar. É a maneira que a conduta do negócio é orientada, seguindo alguns pontos como: legislação, regulamentações e padrões de mercado.

O quadro 01, apresenta as regras de negócio levantadas para desenvolvimento

Quadro 4 – Levantamento das regras de negócio

Regra de Negócio	Descrição
RN01 – Dimensão dim_tempo	<p>O DW deve conter uma tabela de tempo com granularidade de dias e com os campos auxiliares e complementares, como feriados conforme descritos a seguir:</p> <ol style="list-style-type: none"> 1. sk_data 2. data 3. nr_ano_mes 4. nr_ano 5. nr_mes 6. nr_dia 7. nr_dia_semana 8. nr_dia_do_ano 9. nm_dia_semana 10. nm_dia_semana_abreviado 11. nm_mes 12. nm_mes_abreviado 13. nr_mes_ano_txt_curto 14. nr_quinzena_mes 15. nr_trimestre

	16. nr_ano_trimestre 17. nr_semestre 18. nr_ano_semestre 19. nr_semana_mes 20. nr_semana_ano
RN02 – Dimensão dim_cliente	<p>O DW deve conter os dados dos clientes sendo que existe uma ampla classificação de tipos de canais ou categorias, correspondentes ao tipo do estabelecimento do cliente. Os dados devem ser conforme a lista abaixo:</p> 1. sk_cliente 2. cd_cliente 3. nm_cliente 4. cd_uf 5. nm_uf_sigla 6. cd_cidade_ibge 7. nm_cidade 8. nm_bairro 9. cd_categoria_cliente 10. nm_categoria_cliente 11. cd_unid_negocio_cliente 12. nm_unid_negocio_cliente 13. cd_segmento_cliente 14. nm_segmento_cliente 15. flag_status_cliente 16. dt_cadastro_cliente
RN03 – Dimensão dim_vendedor	<p>O DW deve conter uma tabela com a equipe de vendas, referenciada pelo vendedor que realizou o fato da venda. Os campos devem ser conforme a lista abaixo:</p> 1. sk_vendedor 2. cd_vendedor 3. nm_vendedor 4. cd_supervisor 5. nm_supervisor 6. cd_equipe_vendedor 7. nm_equipe_vendedor 8. flag_status_vendedor
RN04 – Dimensão dim_geografia	<p>O DW deve possuir uma tabela de geografia com a hierarquia de cidade e estado, contendo todos os municípios que abrange a área de atuação do negócio, são as cidades dos clientes. Foi solicitado também a inclusão da quantidade de habitantes por município. Os dados que compõem a tabela são os listados a seguir:</p> 1. sk_cidade 2. cd_cidade 3. nm_cidade 4. cd_uf 5. nm_uf_sigla 6. nm_uf 7. qtd_habitantes

RN05 – Dimensão dim_produto	<p>O DW deve incluir o objeto do fato, que nesse caso são os produtos comercializados. Eles possuem uma hierarquia de linha, grupo, família e produto. Além de informações sobre peso e quantidade de produtos por caixa e pallet. Os dados são os listados a seguir:</p> <ol style="list-style-type: none"> 1. sk_produto 2. cd_produto 3. nm_produto 4. cd_familia_produto 5. nm_familia_produto 6. nm_grupo_produto 7. cd_grupo_produto 8. cd_linha_produto 9. nm_linha_produto 10. qtd_und_por_caixa 11. vl_peso_bruto_und_kg 12. vl_peso_liquido_und_kg 13. flag_status_produto
RN06 – Fato ft_venda	<p>O DW deve conter uma tabela que armazene os fatos. Essa tabela deve ser associada com as dimensões: dim_tempo, dim_cliente, dim_produto, dim_vendedor e dim_geografia. Outros atributos ligados a transação/fato serão armazenados na própria tabela fato como dimensões degeneradas. Os dados são listados a seguir:</p> <ol style="list-style-type: none"> 1. sk_data 2. sk_cliente 3. sk_produto 4. sk_vendedor 5. sk_cidade 6. cd_pedido 7. nr_nota_fiscal 8. vl_venda 9. vl_imposto_st 10. vl_imposto_ipi 11. vl_total_imposto 12. vl_venda_liquida 13. vl_desconto 14. vl_peso_bruto 15. vl_peso_liquido 16. qtd_unidade_venda
RN07 – Extrair dados	O sistema deverá extrair os dados disponibilizados pela empresa através da base de dados do ERP e de tabelas auxiliares.
RN08 – Carregar dados <i>Staging</i>	O sistema deverá carregar apenas os dados necessários para a solução e armazená-los em um banco de dados chamado stage.
RN09 – Transformar dados	O sistema deverá transformar os dados para que se ajustem as necessidades da empresa e tenha uma correta estruturação. Não armazenando dados nulos nas <i>surrogate key</i> .
RN10 – Carregar dados DW	O sistema deverá carregar os dados já transformados em um banco de dados modelado no modelo estrela a fim de permitir conexão para exploração por <i>softwares</i> específicos.

Fonte: elaborado pelo autor.

As tabelas e colunas do DW possuem prefixo para reduzir o tamanho dos nomes no banco de dados. A seguir segue os prefixos utilizados e seu significado.

Quadro 5 – Levantamento das regras de negócio

Prefixo	Descrição
dim	Dimensão
ft	Fato
vl	Valor
qtd	Quantidade
nm	Nome (utilizado para descrição ou nome)
nr	Número
cd	Código
sk	<i>Surrogate Key</i> (Chave das dimensões)
flag	Indicador, usado para armazenar a situação
dt	Data

Fonte: elaborado pelo autor.

4.2.3 Requisitos funcionais

No entendimento de Sbrocco (2012, p. 50) os requisitos funcionais são a relação das funcionalidades que o sistema deve prover para o usuário, orientadas ao âmbito do problema.

Quadro 6 – Os requisitos funcionais descrevem as regras de negócio

Requisitos Funcionais	Descrição
RF01 – Disponibilizar <i>Dashboard</i>	Disponibilizar o <i>Dashboard</i> para visualização dos dados ao diretor da empresa para facilitar a tomada de decisão.
RF02 – Dimensões atualizadas automaticamente	O ETL deverá ter atualização automática para manter as dimensões atualizadas, uma carga diária a fim de manter exatamente com as informações que o ERP possui. Manter histórico da localização dos clientes.
RF03 – Atualização incremental da Fato automática	O ETL deverá ter atualização automática para manter os fatos atualizados com uma carga diária. A carga deverá ser incremental a fim de demandar o menor recurso computacional possível.

Fonte: elaborado pelo autor.

4.2.4 Requisitos não funcionais

Sbrocco (2012, p. 51) descreve os requisitos não funcionais como características ligadas ao domínio da solução e demandadas pelo cliente, fatores

externos, organizacionais ou tecnológicos. Os requisitos não funcionais podem ser restrições que podem estar presentes nos contextos da solução.

Quadro 7 – Requisitos não funcionais

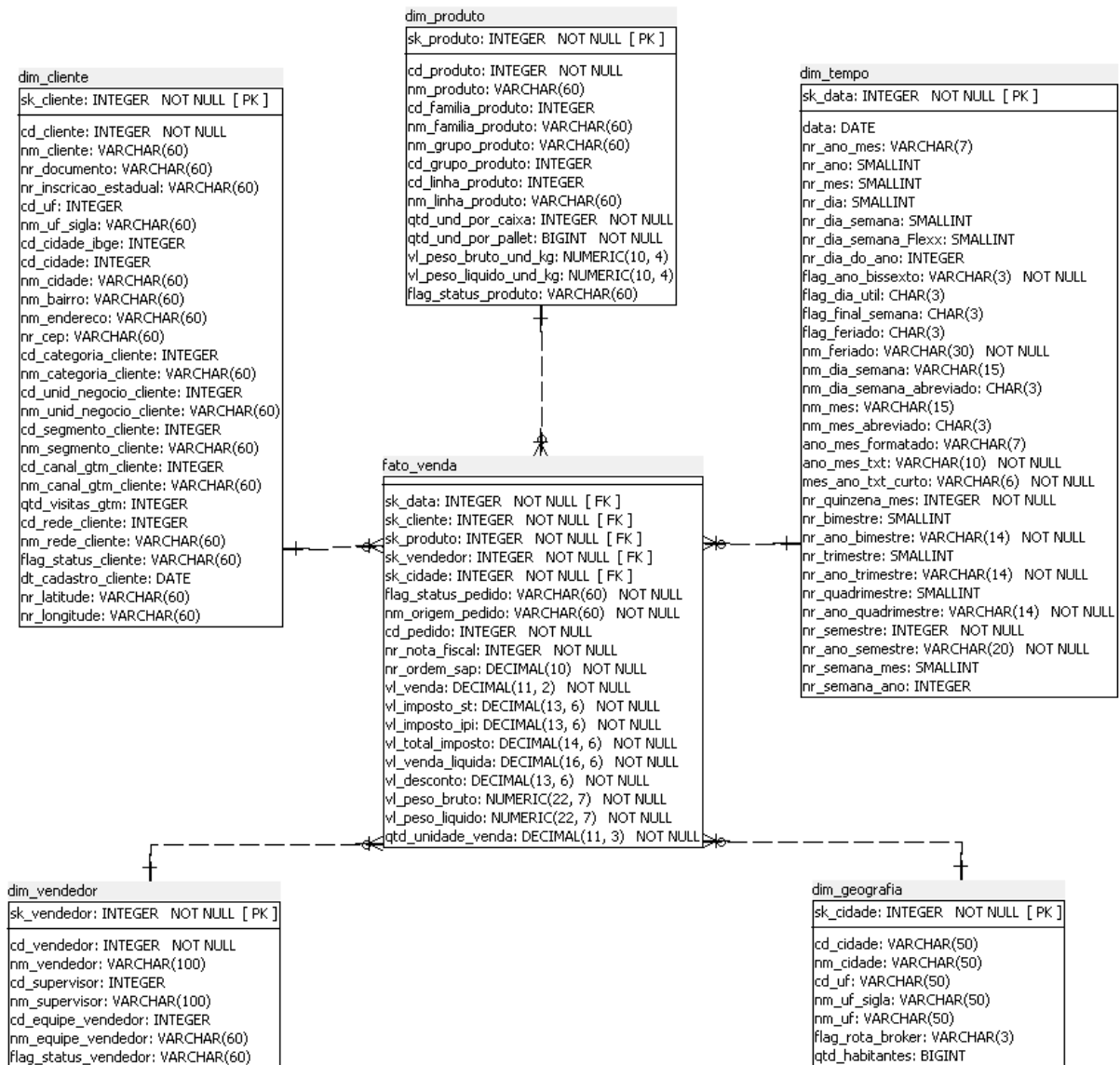
Requisitos Não Funcionais	Descrição
RNF01 – Banco de Dados SQL Server 2019	O DW deverá utilizar o banco de dados SQL Server 2019 Developer.
RNF02 – Disponibilizar <i>Dashboard</i> em <i>Power BI</i>	O <i>Dashboard</i> deverá ser desenvolvido no sistema Microsoft Power BI Desktop.
RNF03 – Modelagem Entidade Relacionamento	A modelagem lógica deve ser feita no <i>Power Architect Community Edition</i> 1.0.8 do modelo entidade relacionamento (ER).
RNF04 – ETL	O sistema deverá utilizar o <i>Pentaho Data Integration Community Edition</i> 9.0 para etapas de integração de dados (ETL).

Fonte: elaborado pelo autor.

4.3 MODELAGEM DIMENSIONAL

Neste projeto de DW foi aplicado o modelo de esquema estrela por ser uma estrutura com melhor performance no processamento de dados e pelo fato do tamanho em disco não ser um problema neste cenário. O modelo lógico possui uma tabela fato: *ft_venda* e cinco tabelas dimensões: *dim_tempo*, *dim_geografia*, *dim_cliente*, *dim_produto* e *dim_vendedor*, destacada na Figura 24.

Figura 23 – Modelagem dimensional – Esquema estrela




Fonte: elaborado pelo autor.

4.3.1 Dimensão cliente

A dimensão de cliente, modelada como o nome da tabela “dim_cliente” é responsável por fornecer informações referentes a quem participa do fato. Que nesse caso foi o cliente que realizou a compra de produtos. Essa dimensão contém o nome do cliente, o código que é obtido do ERP da empresa estudo de caso. Além de informações com: endereço, classificação quanto ao tipo de estabelecimento e a “dt_cadastro_cliente” que foi a data que o cliente foi incluído no sistema ERP.

Quadro 8 – Dimensão cliente


dim_cliente	
	sk_cliente
	cd_cliente
	nm_cliente
	cd_uf
	nm_uf_sigla
	cd_cidade_ibge
	nm_cidade
	nm_bairro
	cd_categoria_cliente
	nm_categoria_cliente
	cd_unid_negocio_cliente
	nm_unid_negocio_cliente
	cd_segmento_cliente
	nm_segmento_cliente
	flag_status_cliente
	dt_cadastro_cliente
	etl_versao
	etl_dt_inicio
	etl_dt_fim

Fonte: elaborado pelo autor.

4.3.2 Dimensão geografia

A dimensão geografia, modelada na tabela “dim_geografia” é criada para armazenar informações referentes a localidade, atende a questão referente a: “onde o fato ocorreu”. Na modelagem ela é usada para determinar a localidade física do cliente, cidade, estado além de conter a informação de quantidade de habitantes para cálculos específicos como média de valor de compra por habitante de uma determinada cidade.

Quadro 9 – Dimensão geografia


dim_geografia	
	sk_cidade
	cd_cidade
	nm_cidade
	cd_uf
	nm_uf_sigla
	nm_uf
	qtd_habitantes
	etl_versao
	etl_dt_inicio
	etl_dt_fim

Fonte: elaborado pelo autor.

4.3.3 Dimensão produto

A dimensão de produto atende o requisito de “qual é o objeto do fato”. Ela contém informações dos produtos comprados pelos clientes. A empresa também classifica os produtos quanto a linha, grupo e família, similar a categoria e subcategoria.

Quadro 10 – Dimensão produto


dim_produto	
	sk_produto
	cd_produto
	nm_produto
	cd_grupo_produto
	nm_grupo_produto
	cd_familia_produto
	nm_familia_produto
	cd_linha_produto
	nm_linha_produto
	qtd_und_por_caixa
	vl_peso_bruto_und_kg
	vl_peso_liquido_und_kg
	flag_status_produto
	etl_versao
	etl_dt_inicio
	etl_dt_fim

Fonte: elaborado pelo autor.

4.3.4 Dimensão tempo

A dimensão de tempo atende o a pergunta do requisito “quando o fato ocorreu”. Contempla a data que ocorreu o fato e mais as suas segmentações como: dias, meses, anos, trimestres. Ela apresenta os meses na forma de número e extenso, de forma completa e abreviada. Exemplo: mês: 1, janeiro, jan.

Quadro 11 – Dimensão tempo


dim_tempo	
	sk_data
	data
	nr_ano_mes
	nr_ano
	nr_mes
	nr_dia
	nr_dia_semana
	nr_dia_do_ano
	nm_dia_semana
	nm_dia_semana_abreviado
	nm_mes
	nm_mes_abreviado
	nm_mes_ano_txt_curto
	nr_quinzena_mes
	nr_trimestre
	nr_ano_trimestre
	nr_semestre
	nr_ano_semestre
	nr_semana_mes
	nr_semana_ano
	nm_data_completa
	primeiro_dia_mes
	ultimo_dia_mes
	etl_versao
	etl_dt_inicio
	etl_dt_fim

Fonte: elaborado pelo autor.

4.3.5 Dimensão vendedor

A dimensão de vendedor, representada pela tabela “dim_vendedor”, atende ao requisito de “quem gerou o fato”. Indica o vendedor que realizou a venda para o cliente. Ela armazena informações referentes ao seus respectivos supervisor e gerente.

Quadro 12 – Dimensão vendedor

dim_vendedor	
	sk_vendedor
	cd_vendedor
	nm_vendedor
	cd_supervisor
	nm_supervisor
	cd_equipe_vendedor
	nm_equipe_vendedor
	etl_versao
	etl_dt_inicio
	etl_dt_fim
	cd_tipo_equipe
	nm_tipo_equipe

Fonte: elaborado pelo autor.

4.3.6 Fato venda

As vendas, que são os fatos, representada pela tabela “ft_venda”, armazena informações referentes ao valor vendido, peso dos produtos, quantidade de produtos vendida e as chaves *Surrogate Key* (sk) que fazem a associação com todas as tabelas de dimensão.

Quadro 13 – Fato venda

ft_venda	
sk_data	
sk_dt_retorno	
sk_cliente	
sk_produto	
sk_vendedor	
sk_motorista	
sk_veiculo	
sk_cidade	
cd_pedido	
nr_nota_fiscal	
vl_venda	
vl_imposto_st	
vl_imposto_ipi	
vl_total_imposto	
vl_venda_liquida	
vl_desconto	
vl_peso_bruto	
vl_peso_liquido	
qtd_unidade_venda	

Fonte: elaborado pelo autor.

5 DESENVOLVIMENTO

Este capítulo objetiva explicar e ilustrar a solução proposta. Apresentando as ferramentas e tecnologias escolhidas para implementar a solução. Será apresentada as etapas executadas e ao final a solução proposta com as considerações relatadas.

5.1 TECNOLOGIAS E FERRAMENTAS

Está seção apresenta as tecnologias e ferramentas utilizadas no processo de desenvolvimento da solução proposta. Tecnologias de armazenamento de dados, ferramenta de ETL, ferramenta de modelagem de dados, Canvas para Planejamento do *Storytelling* de dados, Documento para *Storyboard* e ferramenta para implementação do *Dashboard*.

5.1.1 Microsoft SQL Server

O Microsoft SQL Server é um produto da Microsoft que é composto por vários softwares. O principal sistema, e o que utilizaremos no desenvolvimento deste trabalho, é o Sistema Gerenciamento de Banco de Dados (SGBD). Com ele iremos executar os comandos *Transact-SQL*, um dialeto da linguagem SQL desenvolvida pela Microsoft.

SQL (*Structure Query Language*) é uma linguagem de programação usada para definir e consultar bancos de dados. *View* é um objeto do SQL semelhante a uma tabela virtual em que o conjunto de resultados é extraído de uma query (consulta em SQL). Utilizaremos as *views* para abstrair a complexidade das consultas realizadas no BD do ERP da empresa estudo de caso.

No banco de dados SQL Server armazenaremos a área de *Staging* e o *Data Warehouse* e extrairemos os dados através de *Views* do ERP da na empresa estudo de caso.

5.1.2 Microsoft SQL Server Management Studio

O SSMS (*SQL Server Management Studio*) é um software usado para gerenciar a infraestrutura do SQL Server. O SSMS possui ferramentas para

configurar, monitorar e administrar instâncias do SQL Server e bancos de dados e permite também criar e executar consultas e scripts SQL (MICROSOFT, 2020, p. 1).

5.1.3 SQL Power Architect

SQL *Power Architect* é um *software* especializado com recursos e ferramentas de modelagem de dados. Criado por designers de *Data Warehouse* com função que permite que os usuários façam engenharia reversa em bancos de dados, apliquem a modelagem física em banco de dados e gerem metadados ETL automaticamente. Essas ferramentas de modelagem facilitam e simplificam os esforços de modelagem de dados, maximizando o uso de seus recursos (BEST OF BI, 2020, p. 1).

5.1.4 Pentaho Data Integration

Pentaho Data Integration (PDI, também chamado de *Kettle*), é utilizado para integração de dados (processos de ETL), exportação de dados de bancos de dados para arquivos planos, limpeza de dados e muito mais. PDI é uma ferramenta multiplataforma, o que significa que você poderá instalar a ferramenta independentemente do seu sistema operacional (PULVIRENTI; ROLDAN, 2011, p. 1).

Ferramentas e utilitários *Pentaho Data Integration* compreende o seguinte conjunto de ferramentas e utilitários:

- *Spoon* – Um IDE de integração de dados gráficos para criar transformações e Jobs.
- *Kitchen* – Uma ferramenta de linha de comando para executar Jobs.
- *Pan* – Uma ferramenta de linha de comando para executar transformações.

O PDI possui algumas estruturas, entre elas podemos citar: Repositórios, Transformações e Jobs.

Repositórios são os locais onde armazenaremos as transformações, conexões e Jobs do ETL.

Transformações são etapas que realizam operações nos dados. Existe uma variedade muito grande de transformações no PDI. Extração de dados de arquivos texto, Excel e Banco de Dados. Mudanças de tipo de dados, de textos para número.

Operações específicas para textos, como passar todas as letras para maiúsculo. São alguns exemplos.

Jobs são compostos de uma ou mais transformações e até outros Jobs. Os Jobs operam os status de saída resultante da execução da tarefa (BOUMAN; DONGEN, 2009, p. 25 a 42).

5.1.5 Microsoft Power BI

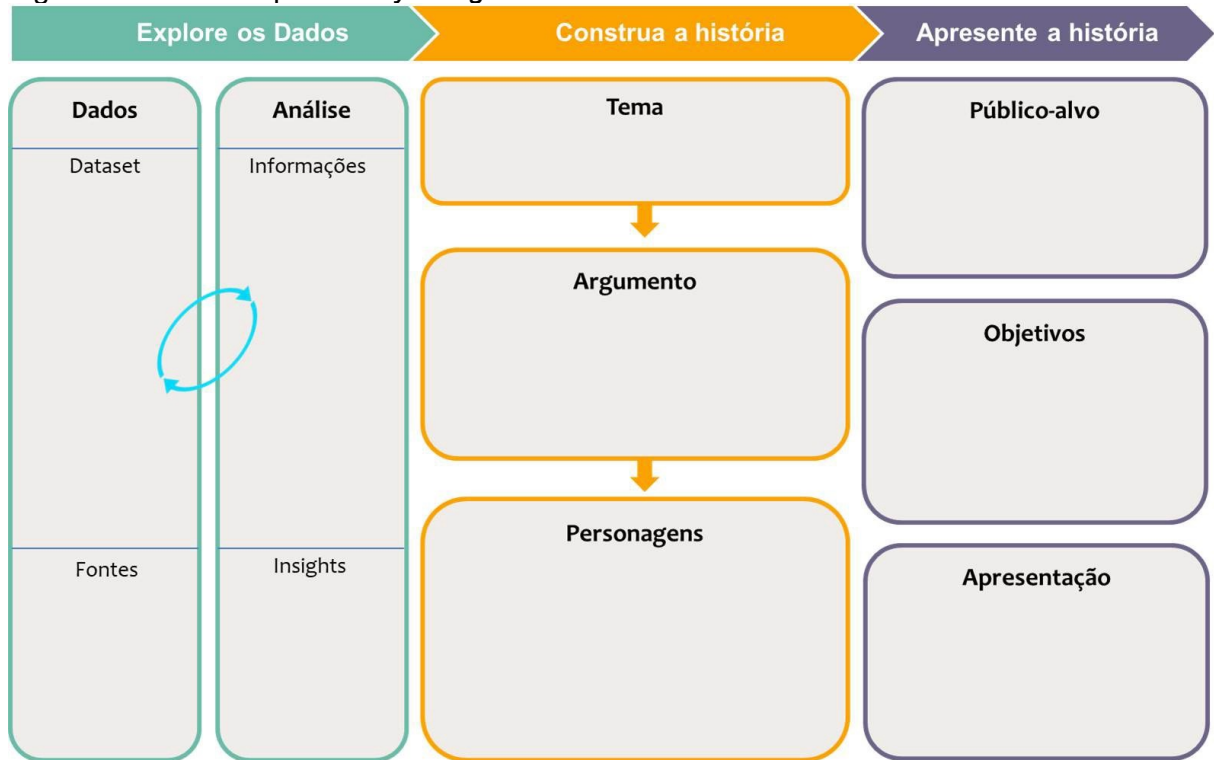
O Microsoft Power BI é um conjunto de softwares, serviços e conectores que trabalham em sinergia fornecendo um ambiente para que possamos transformar os dados em informações coerentes, visualmente envolventes e interativas com *Dashboards*. Os dados podem estar em uma planilha do Excel, arquivos textos, aplicações web ou *data warehouses* híbridos locais ou baseados na nuvem. Power BI, permite conectar facilmente nas fontes de dados, criar visualizações e *Dashboards* e compartilhar com o público-alvo permitindo tomarem as melhores decisões para os seus negócios (MICROSOFT, 2020, p. 1).

5.1.6 DAX

DAX, significa *Data Analysis Expressions*, é a linguagem de programação do Microsoft Power BI, Microsoft *Analysis Services* e Microsoft *Power Pivot* para Excel. Essa linguagem foi projetada para aplicar fórmulas e negócio em um modelo de dados, facilitando a construção de métricas e KPIs (RUSSO; FERRARI, 2020, p. 1).

5.1.7 Canvas para *Storytelling*

Canvas para *Storytelling* de Dados é um quadro que auxilia no processo e análise de dados. Criado por Stéfano Carnevalli e Raphael Peruccini para auxiliar nos trabalhos da startup Simples Data fundada por eles. Por uma dificuldade em apresentar análises e dados para o público-alvo a utilização do Canvas facilita e direciona para um entendimento mais amplo das necessidades em dados. O desenvolvimento do Data *Storytelling* Canvas se concebeu reunindo conhecimento de negócio e de apresentações usando técnicas de *storytelling* (CARNEVALLI, 2020, p. 1).

Figura 24 – Canvas para *Storytelling***Data Storytelling Canvas**

Elaborado por Stéfano Carnevalli e Raphael Peruccini. Simples Data.
Agradecemos a todos que de alguma forma apoiaram essa iniciativa.

Fonte: Carnevalli (2020, p. 1).

Carnevalli (2020, p. 1) apresenta a Data *Storytelling* Canvas dividida em 3 áreas:

1. Explore os dados
 - a. Dados (Quais dados disponibilizados? Quais outros dados serão necessários? Quais fontes? Onde estão armazenadas?).
 - b. Análise (Que interpretações têm dos dados? O que surpreenderia o público-alvo).
2. Construa a história
 - a. Tema (Qual tema central da apresentação? Será utilizada alguma metáfora?).
 - b. Argumento (Palavras chaves do roteiro da apresentação. Passos do roteiro).
 - c. Personagens (Que tipos gráficos utilizar? Imagens? Textos?).
3. Apresente a história
 - a. Público-Alvo (Para quem vou apresentar? Qual perfil?).
 - b. Objetivos (Qual a pergunta que temos que responder? Para quê? Qual finalidade?).

- c. Apresentação (Como será? Quais recursos disponíveis? Quanto tempo? Quantos participantes?).

5.2 EXPERIMENTO

O objetivo deste trabalho, detalhado no capítulo 1, é desenvolver um *Dashboard* aplicando conceitos e fundamentos que elevem o aprimoramento no desenvolvimento do *Dashboard* a fim de entregar ao público-alvo o propósito de informar e facilitar as decisões. Para tanto, será desenvolvido uma solução de BI atendendo os requisitos descritos no capítulo 4.

As etapas para realização do desenvolvimento deste trabalho são de acordo com os passos da tabela abaixo:

Quadro 14 – Experimento

Passo	Descrição
1	É realizado a extração dos dados necessários da base do sistema ERP da empresa estudo de caso.
2	Dados são armazenados no banco de dados “ <i>stage</i> ”.
3	Os dados passam por tratamento.
4	Dados são carregados para o DW no banco de dados “ <i>datawarehouse</i> ” e estão prontos para serem consumidos.
5	Os dados são carregados no Power BI para desenvolvimento do Dashboard.
6	Definir o <i>storytelling</i> de dados: contexto, público-alvo e objetivo com o Canvas.
7	Esboçar o <i>Storyboard</i> para definir a ordem que os visuais serão apresentados.
8	O Dashboard é desenvolvido e entregue para a diretoria da empresa estudo de caso.

Fonte: elaborado pelo autor.

Nas seções a seguir serão descritos: arquitetura da solução, integração dos dados, planejamento do *Storytelling* de dados, *Storyboard* e a implementação do *Dashboard*.

5.2.1 Arquitetura da solução

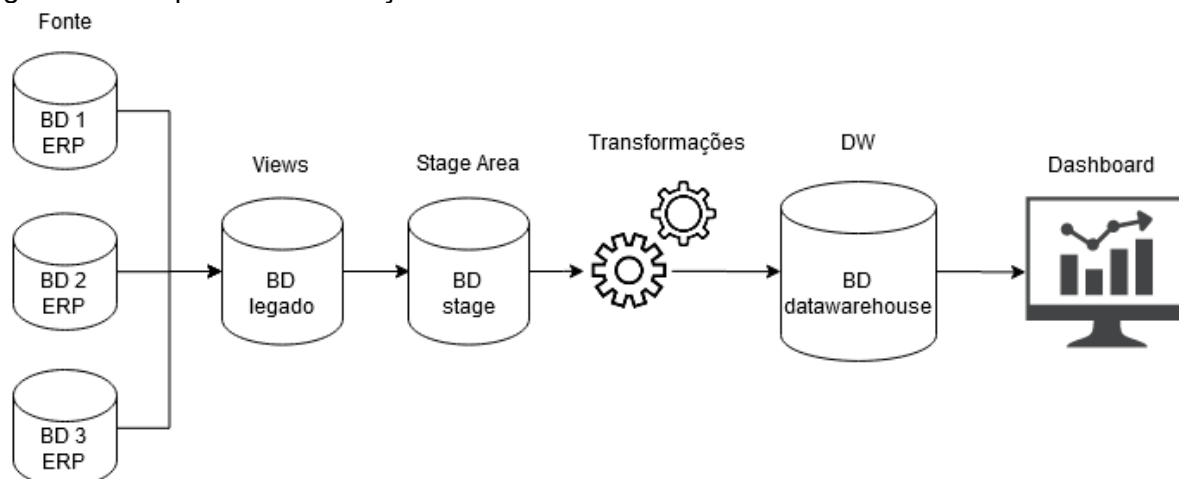
A solução proposta inicia-se na elaboração da arquitetura, ela sintetiza como será o fluxo das informações e tecnologias para embasar o desenvolvimento.

Inicialmente, o Administrador do SQL Server da empresa estudo de caso, propicia as bases de dados físicas e os acessos conforme descrito a seguir:

1. Cria 3 bases de dados no SQL Server: (“legado”, “stage” e “datawarehouse”).
2. Cria usuário (“datawarehouse”) no banco de dados com permissão de leitura na base “legado” e leitura/escrita no banco de dados “stage” e “datawarehouse”.
3. Cria as *views* (no banco de dados “legado”) que permitem a extração das informações na fonte (Banco de Dados do Sistema ERP) necessárias para carregar cada dimensão e fato.

A imagem a seguir demonstra a arquitetura implementada para a solução.

Figura 25 – Arquitetura da solução



Fonte: elaborado pelo autor.

A seguir, a tabela contém as fontes necessárias para a extração dos dados e elaboração das tabelas de dimensão e fato descritas no capítulo 4.

Quadro 15 – Arquitetura da solução

Nome	Formato	Fonte	Descrição
cliente	View	Banco “legado”	Possui informações dos clientes. Atende as necessidades apresentadas no capítulo 4, RN02.
imposto_ipi_item_pedido	View	Banco “legado”	Possui informações referente ao imposto IPI de cada produto e transação da tabela fato. Atende RN06.
imposto_st_item_pedido	View	Banco “legado”	Possui informações referente ao imposto ST (substituição tributária) de cada produto e transação da tabela fato. Atende RN06.
item_pedido	View	Banco “legado”	Possui informações referente ao valor da venda de cada produto e transação da tabela fato. Atende RN06.

pedido	View	Banco “legado”	Possui informações referente ao pedido da transação da tabela fato. Atende RN06.
produto	View	Banco “legado”	Possui informações dos produtos. Atende a necessidade apresentadas no capítulo 4, RN05.
vendedor	View	Banco “legado”	Possui informações dos vendedores e equipe de vendas. Atende a necessidade apresentadas no capítulo 4, RN03.
geografia	Tabela	Banco “legado”	Possui informações da geografia. Atende a necessidade apresentadas no capítulo 4, RN04.
tempo	Tabela	Banco “legado”	Possui informações do tempo, data com o grão de dias. Atende a necessidade apresentadas no capítulo 4, RN01.

Fonte: elaborado pelo autor.

Com os dados apresentados acima e a arquitetura definida daremos início da integração de dados para carregamento do DW.

5.2.2 Integração de dados

Conforme apresentado na seção 2.2.1 deste trabalho, o processo de ETL irá integrar os dados, permitindo a extração dos dados do Banco de Dados do ERP da empresa estudo de caso, *Staging* e por último para o *Data Warehouse*. Com o PDI iremos automatizar e padronizar as etapas necessárias. Nesta seção será apresentada todos os passos para conclusão desse processo.

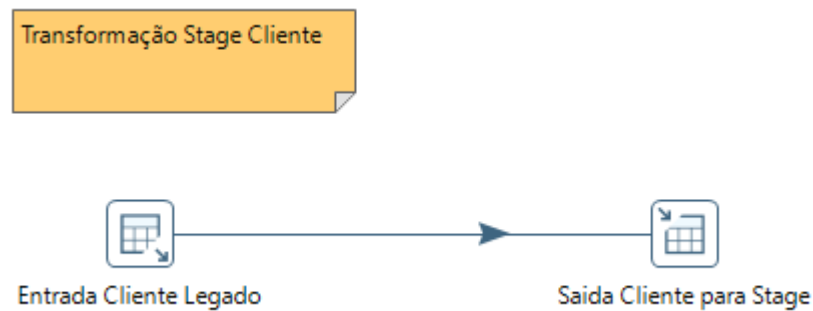
5.2.2.1 Dimensão cliente

Esta seção apresenta os passos para transformação da dimensão cliente. Primeiramente será realizada a transformação da *Staging* com os dados dos clientes e após será carregada a tabela dimensão no *Data Warehouse*.

5.2.2.1.1 Staging st_cliente

Este processo extrai dados clientes do ERP através da *View* cliente no BD legado, utilizando-se a transformação *Table input* e armazena na tabela st_cliente do BD stage através da transformação *Table output*.

Figura 26 – Staging st_cliente

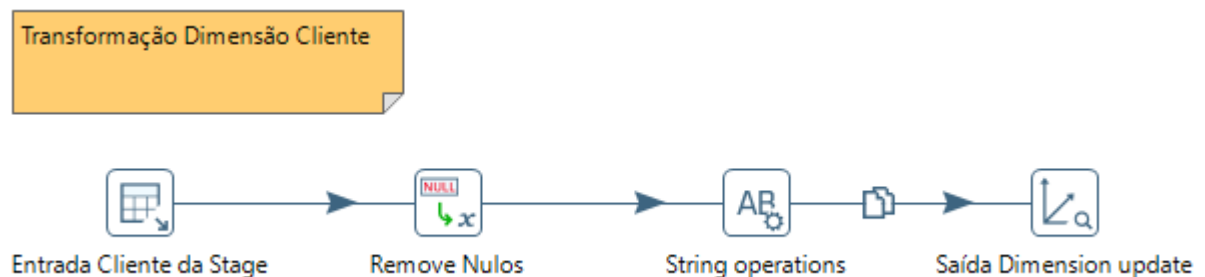


Fonte: elaborado pelo autor.

5.2.2.1.2 Data Warehouse dim_cliente

Extraí dados da tabela st_cliente no BD *stage*. Remove os valores nulos substituindo os códigos por -3 e nomes por *** Não informado ***. Ajusta os nomes para o padrão da primeira letra da palavra em maiúscula e as demais minúsculas e remove espaços iniciais e finais em branco. Armazena os dados na tabela dimensão dim_cliente do BD *datawarehouse*. Mantendo os dados atualizados quando houver mudança no ERP.

Figura 27 – Data Warehouse dim_cliente



Fonte: elaborado pelo autor.

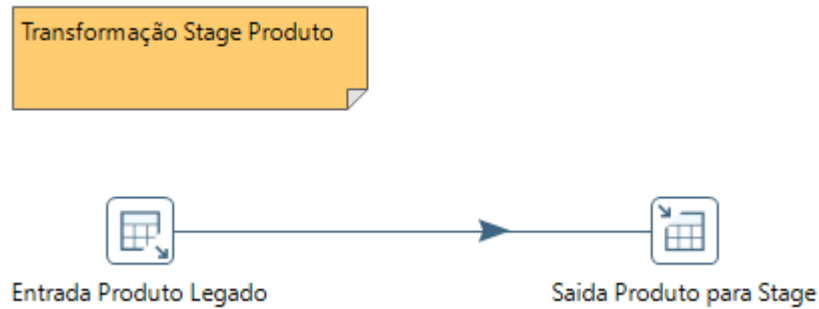
5.2.2.2 Dimensão produto

Esta seção apresenta os passos para transformação da dimensão produto. Primeiramente será realizada a transformação da *Staging* com os dados dos produtos e após será carregada a tabela dimensão no *Data Warehouse*.

5.2.2.2.1 Staging st_produto

Este processo extrai dados clientes do ERP através da *View* produto no BD legado, utilizando-se a transformação *Table input* e armazena na tabela st_produto do BD stage através da transformação *Table output*.

Figura 28 – Staging st_produto

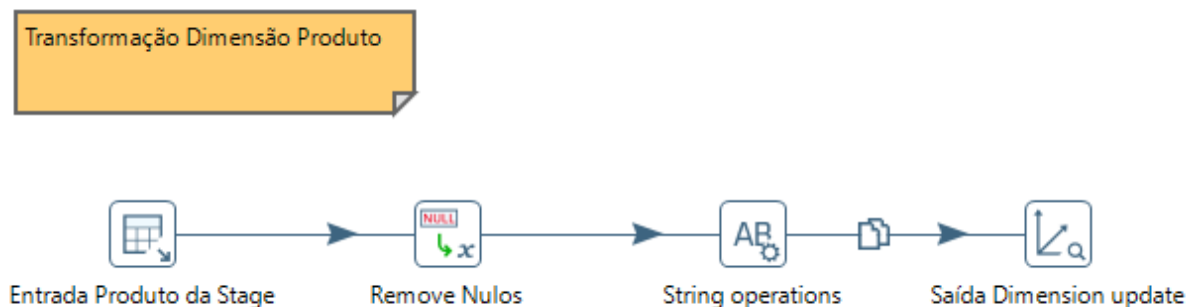


Fonte: elaborado pelo autor.

5.2.2.2.2 Data Warehouse dim_produto

Extrai dados da tabela st_produto no BD *stage*. Remove os valores nulos substituindo os códigos por -3 e nomes por *** Não informado ***. Ajusta os nomes para o padrão da primeira letra da palavra em maiúscula e as demais minúsculas e remove espaços iniciais e finais em branco. Armazena os dados na tabela dimensão dim_produto do BD *datawarehouse*. Mantendo os dados atualizados quando houver mudança no ERP.

Figura 29 – Data Warehouse dim_produto



Fonte: elaborado pelo autor.

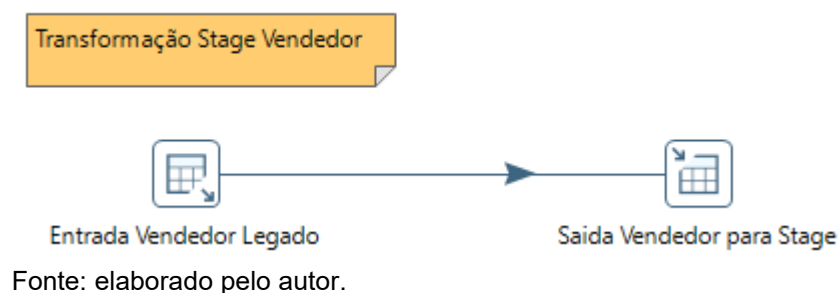
5.2.2.3 Dimensão vendedor

Esta seção apresenta os passos para transformação da dimensão vendedor. Primeiramente será realizada a transformação da *Staging* com os dados dos vendedores e após será carregada a tabela dimensão no *Data Warehouse*.

5.2.2.3.1 Staging st_vendedor

Este processo extrai dados vendedores do ERP através da *View* vendedor no BD legado, utilizando-se a transformação *Table input* e armazena na tabela st_vendedor do BD stage através da transformação *Table output*.

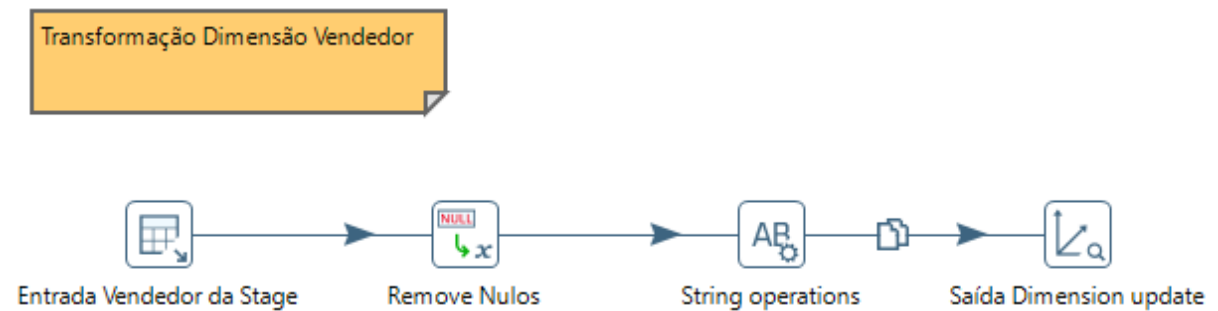
Figura 30 – Staging st_vendedor



5.2.2.3.2 Data Warehouse dim_vendedor

Extrai dados da tabela st_vendedor no BD *stage*. Remove os valores nulos substituindo os códigos por -3 e nomes por *** Não informado ***. Ajusta os nomes para o padrão da primeira letra da palavra em maiúscula e as demais minúsculas e remove espaços iniciais e finais em branco. Armazena os dados na tabela dimensão dim_vendedor do BD *datawarehouse*. Mantendo os dados atualizados quando houver mudança no ERP.

Figura 31 – Data Warehouse dim_vendedor

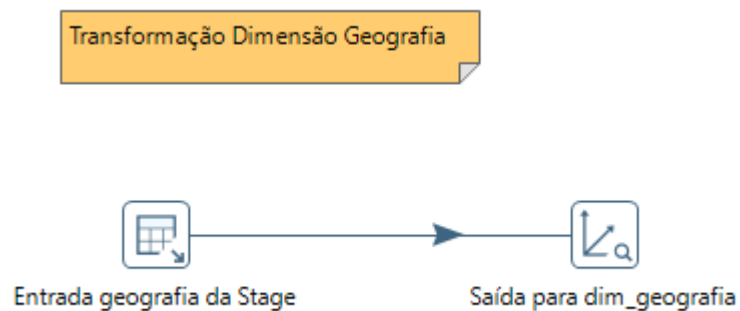


Fonte: elaborado pelo autor.

5.2.2.4 Dimensão geografia

Este processo extrai dados de geografia do ERP através da tabela geografia no BD legado, utilizando-se a transformação *Table input* e armazena na tabela dim_geografia do BD *stage* através da transformação *Table output*.

Figura 32 – Dimensão geografia

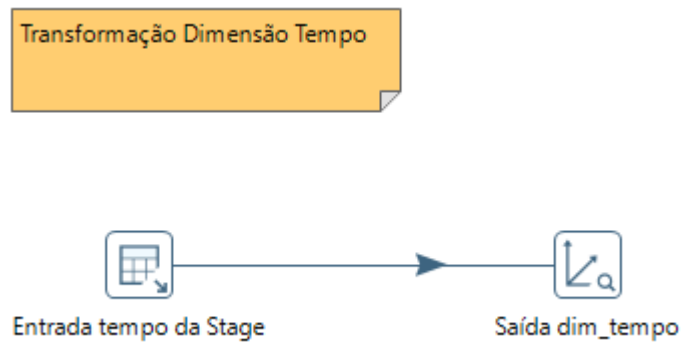


Fonte: elaborado pelo autor.

5.2.2.5 Dimensão tempo

Este processo extrai dados de tempo do ERP através da tabela tempo no BD legado, utilizando-se a transformação *Table input* e armazena na tabela dim_tempo do BD *stage* através da transformação *Table output*.

Figura 33 – Dimensão tempo



Fonte: elaborado pelo autor.

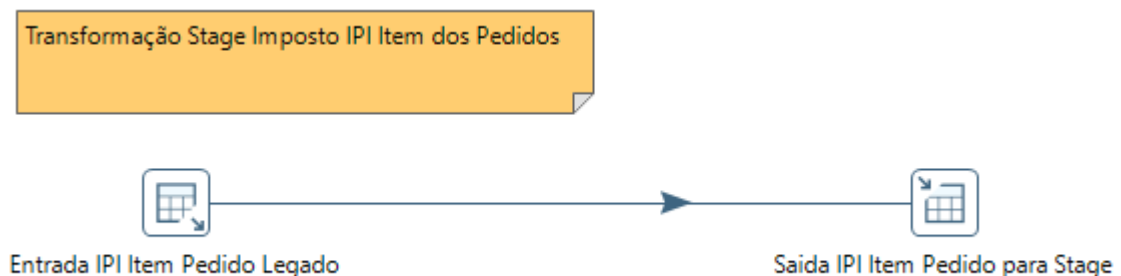
5.2.2.6 Fato vendas

Esta seção apresenta os passos para transformação da tabela fato venda. Primeiramente será realizada a transformação da *Staging*, carregando os dados das transações de venda do legado e na sequência será carregada a tabela ft_venda no *Data Warehouse*. A ft_venda é carregada com dados de 4 tabelas diferentes. Será demonstrado nas próximas subseções cada tabela necessária.

5.2.2.6.1 Staging st_imposto_ipi_item_pedido

Este processo extrai dados dos itens dos pedidos de venda que possuem imposto de IPI do ERP através da *View* imposto_ipi_item_pedido no BD legado, utilizando-se a transformação *Table input* e armazena na tabela st_imposto_ipi_item_pedido do BD *stage* através da transformação *Table output*.

Figura 34 – Staging st_imposto_ipi_item_pedido

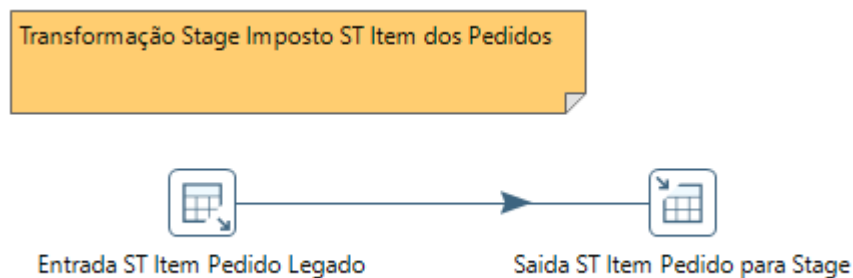


Fonte: elaborado pelo autor.

5.2.2.6.2 Staging st_imposto_st_item_pedido

Este processo extrai dados dos itens dos pedidos de venda que possuem imposto de Substituição Tributária (ST) do ERP através da *View* imposto_st_item_pedido no BD legado, utilizando-se a transformação *Table input* e armazena na tabela st_imposto_st_item_pedido do BD *stage* através da transformação *Table output*.

Figura 35 – Staging st_imposto_st_item_pedido

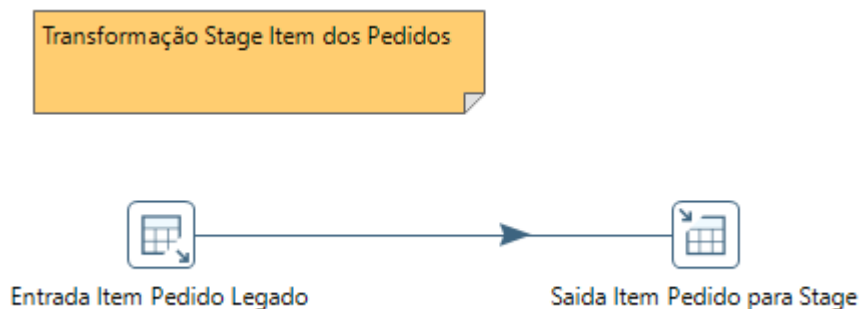


Fonte: elaborado pelo autor.

5.2.2.6.3 Staging st_item_pedido

Este processo extrai dados dos itens dos pedidos de venda do ERP através da *View* item_pedido no BD legado, utilizando-se a transformação *Table input* e armazena na tabela st_item_pedido do BD *stage* através da transformação *Table output*.

Figura 36 – Staging st_item_pedido

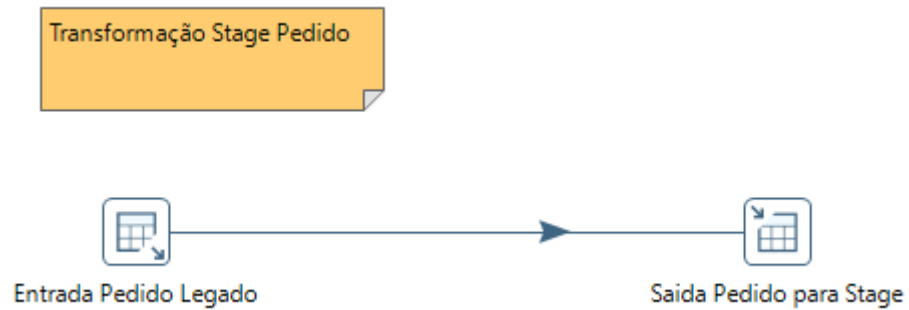


Fonte: elaborado pelo autor.

Este processo extrai dados referente aos pedidos de venda do ERP através da *View* pedido no BD legado, utilizando-se a transformação *Table input* e armazena na tabela st_pedido do BD *stage* através da transformação *Table output*.

5.2.2.6.4 Staging st_pedido

Figura 37 – Staging st_pedido

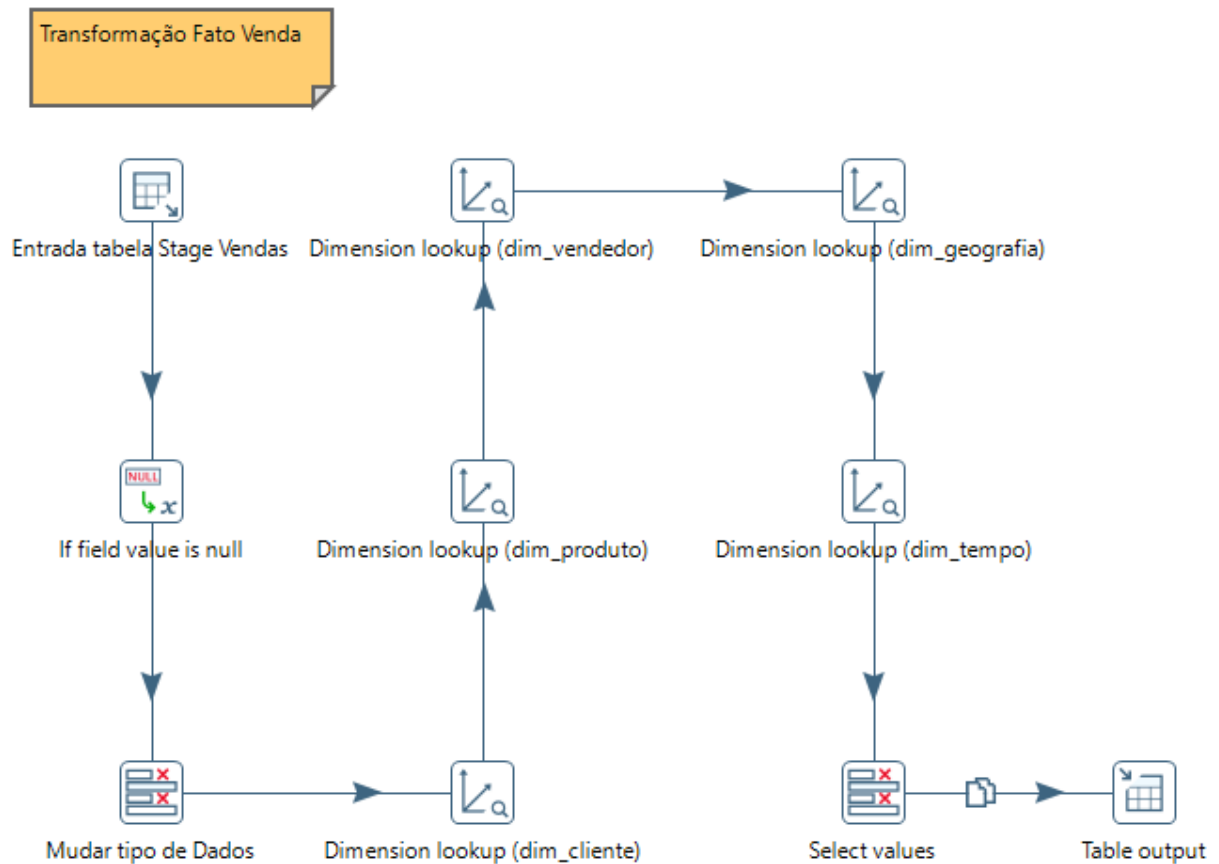


Fonte: elaborado pelo autor.

5.2.2.6.5 Data Warehouse ft_venda

Extrai dados da *View* st_vendas do BD *stage*. Remove os campos nulos substituindo códigos e chaves por -3, valores por 0 e nomes por *** Não informado ***. Procura as chaves substitutas (surrogate key / SK) em todas as dimensões. Carrega tabela ft_venda no BD *datawarehouse*.

Figura 38 – Data Warehouse ft_venda



Fonte: elaborado pelo autor.

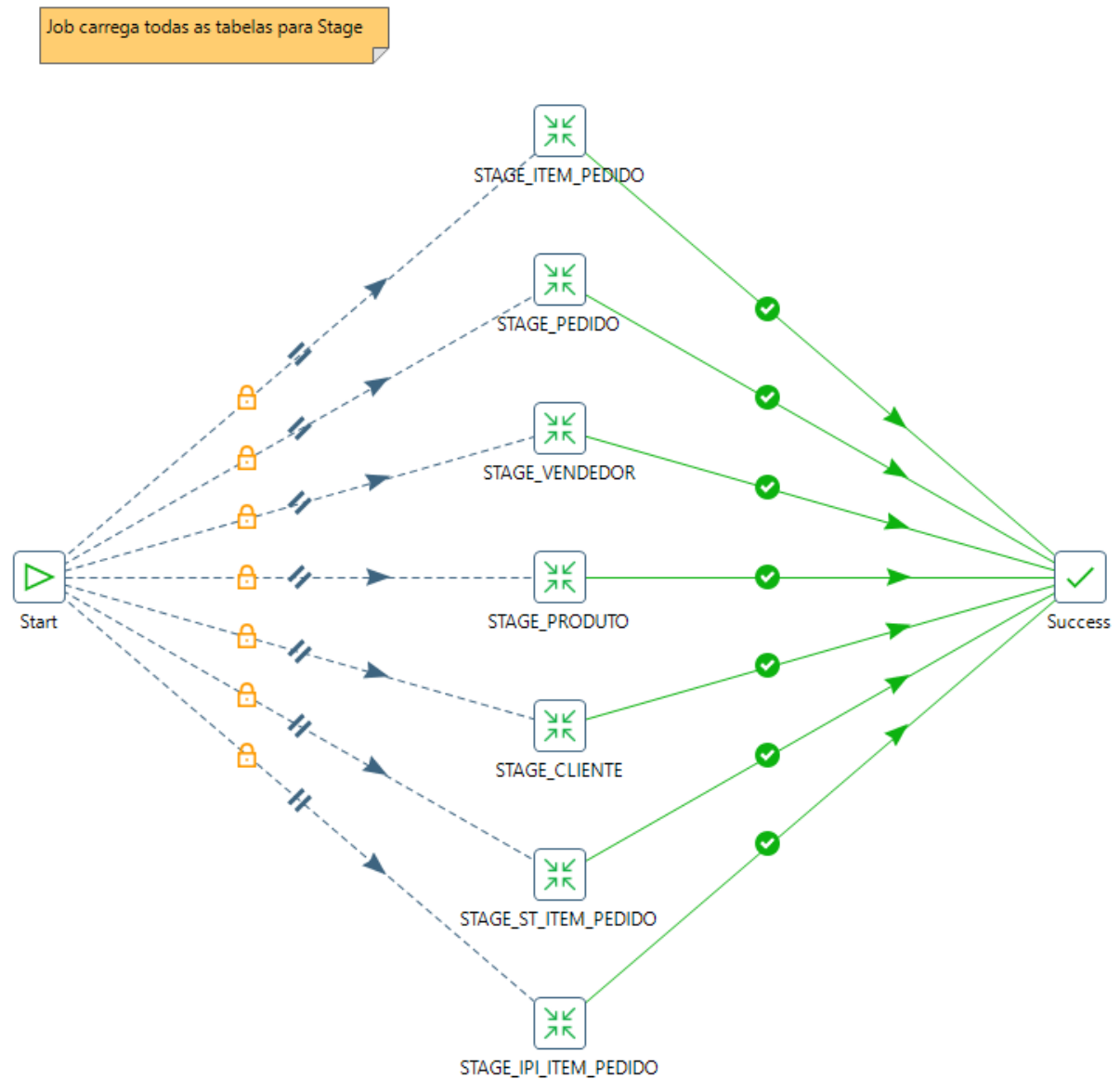
5.2.2.7 Carga no DW

Essa etapa é composta pela união das transformações em processos conhecidos como *Job* no PDI. Automatizaremos o fluxo criando uma linha contínua de trabalho para todo o processo de ETL. Será composto por duas etapas intermediárias, Carga completa da *Staging*. Carga completa das dimensões, e uma carga completa que implementa as duas etapas intermediárias citadas acrescentando a carga da tabela fato no *Data Warehouse*.

5.2.2.7.1 Carga completa da *staging*

Esse *Job* agrega todas as transformações que carregam os dados para a *Staging* em um único processo. Conforme imagem:

Figura 39 – Carga completa da *staging*

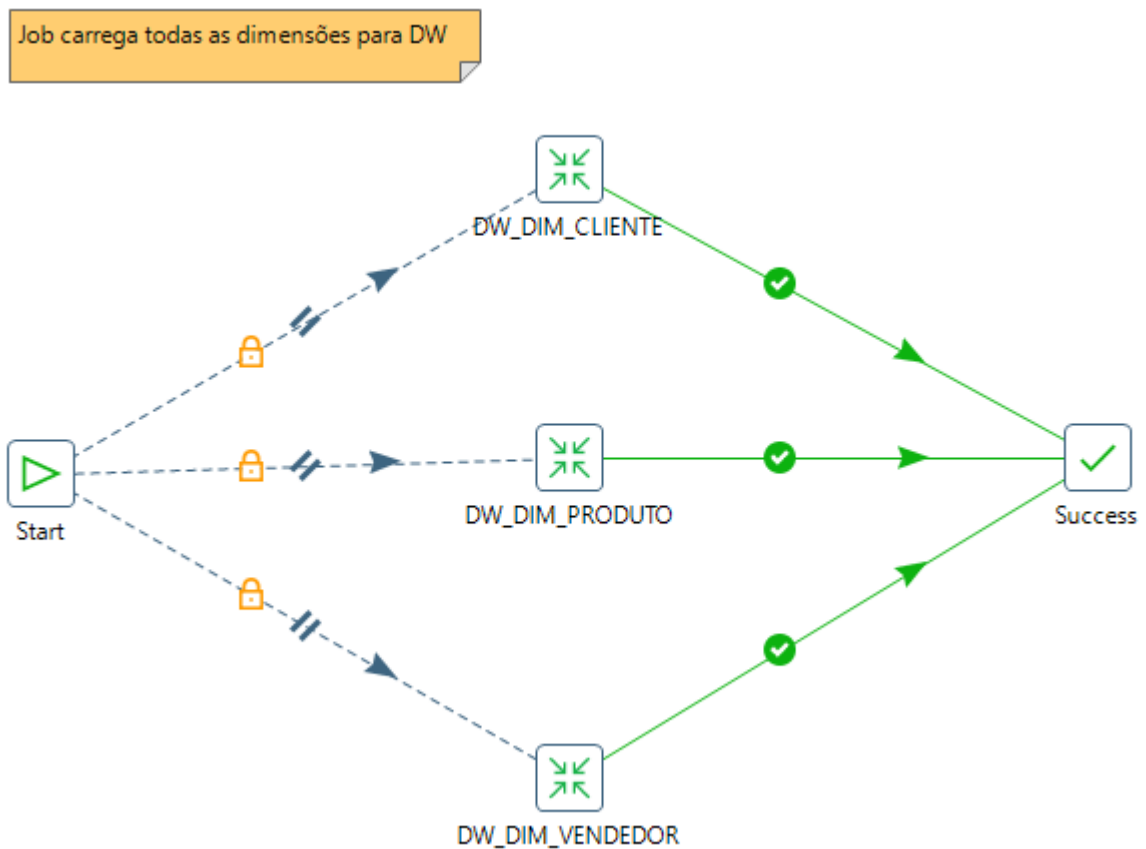


Fonte: elaborado pelo autor.

5.2.2.7.2 Carga dimensões para DW

Esse *Job* agrega todas as transformações que carregam os dados das dimensões para o *Data Warehouse* em um único processo. Conforme imagem:

Figura 40 – Carga dimensões para DW

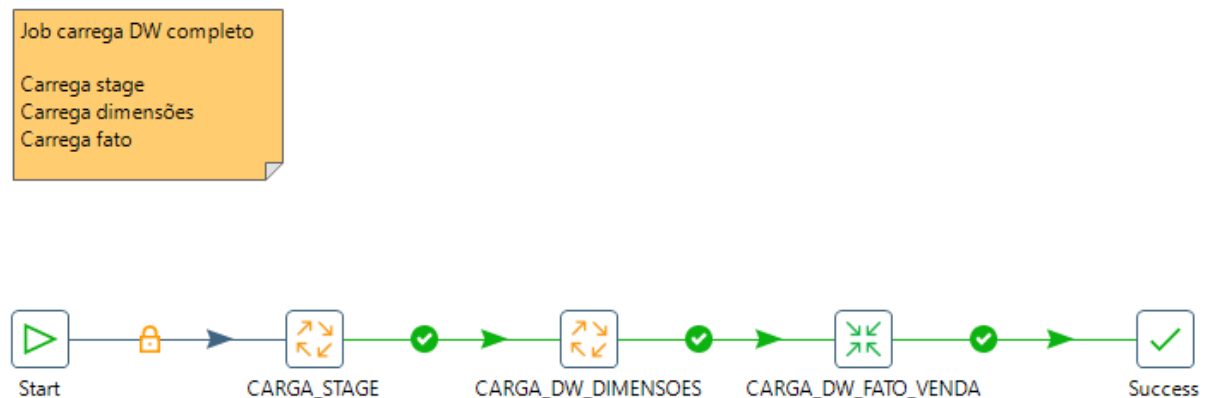


Fonte: elaborado pelo autor.

5.2.2.7.3 Carga DW completa

Este *Job* agrega os processos completos de carga da *Staging*, carga das dimensões para o DW e carga da Fato Venda para o DW. Conforme imagem:

Figura 41 – Carga DW completa



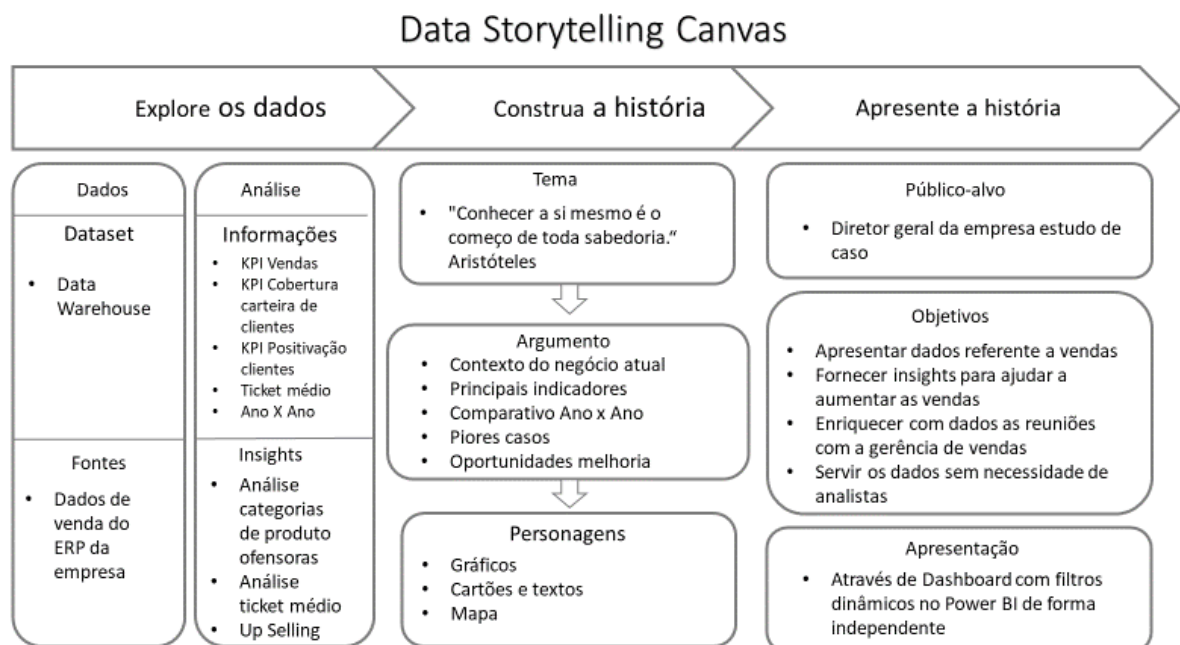
Fonte: elaborado pelo autor.

5.2.3 Planejamento do *storytelling*

O planejamento do *Storytelling* de dados para o desenvolvimento do *Dashboard* será elaborado com o auxílio do Canvas para *Storytelling* de dados. Um bom planejamento desta etapa nos auxiliará no desenvolvimento do *Dashboard* e trará todos os benefícios da aplicação dos conceitos de *Storytelling* de dados para o público-alvo conforme apresentado na seção 2.4 deste trabalho.

A figura 43 apresenta o planejamento do *Storytelling* de dados para o desenvolvimento do *Dashboard*.

Figura 42 – Data *Storytelling* Canvas



Fonte: elaborado pelo autor.

5.2.4 Implementação do *Dashboard*

Conforme apresentado na subseção 5.1.5 a ferramenta para desenvolvimento do protótipo do *Dashboard* escolhida foi o Microsoft Power BI. O *Dashboard* irá apresentar métricas e KPIs levantados na análise de requisitos na subseção 4.2 deste trabalho.

As métricas são apresentadas a seguir:

1. Quantidade de clientes ativos
2. Quantidade de vendedores ativos
3. Quantidade de cidades atendidas

4. Valor faturamento
5. Valor peso transportado em Kg
6. Quantidade de clientes cobertos na carteira
7. Quantidade de positividade de vendas positivadas
8. Ticket médio por cliente
9. Quantidade de habitantes
10. Ticket médio por habitante
11. Carteira de clientes ativa

Os indicadores são apresentados a seguir:

1. Vendas mês realizada pela meta
2. Projeção de vendas mês pela meta
3. Vendas mês anterior pela meta
4. Vendas acumulada do ano pela meta
5. Projeção de vendas do ano pela meta

O período de análise disponível é de janeiro de 2018 até 15 de junho de 2020.

Apresentação da evolução das vendas pelo tempo.

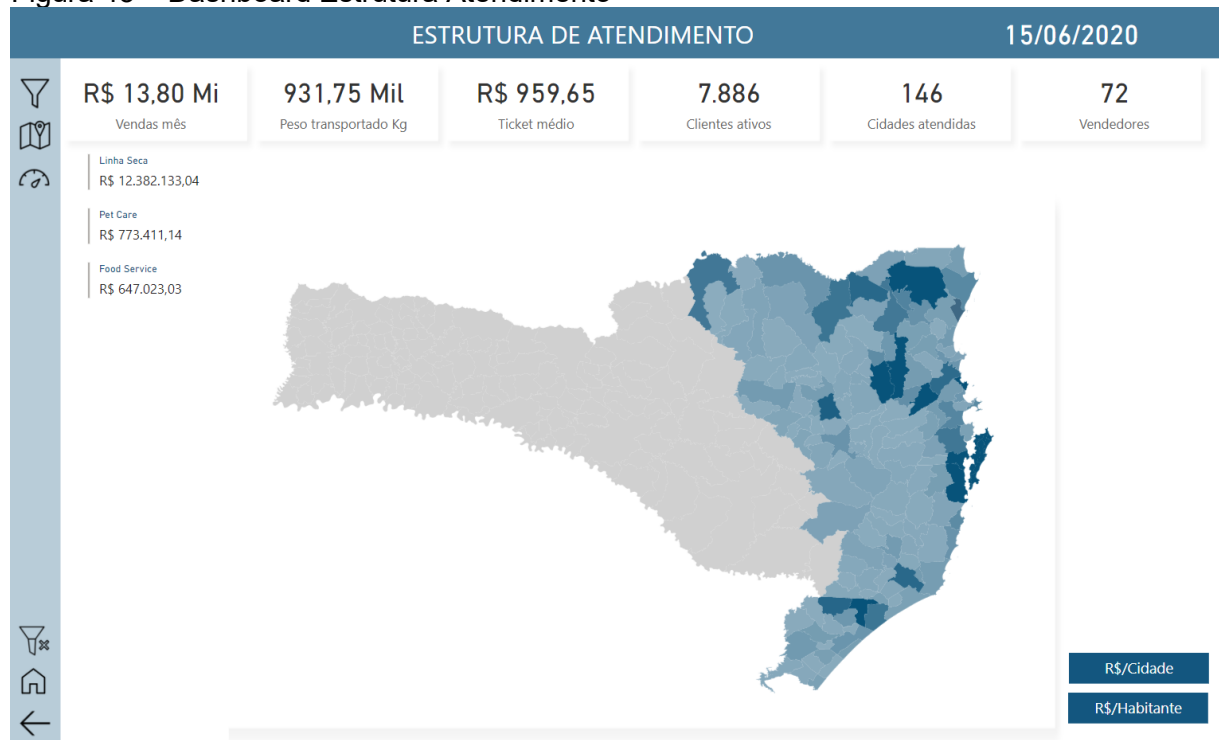
Será feita análise do ano atual comparada com o ano anterior.

Análise referente ao ticket médio, buscando realizar *Up Selling* (aumento das vendas) nos clientes com *ticket* médio abaixo do canal.

Os filtros disponíveis serão:

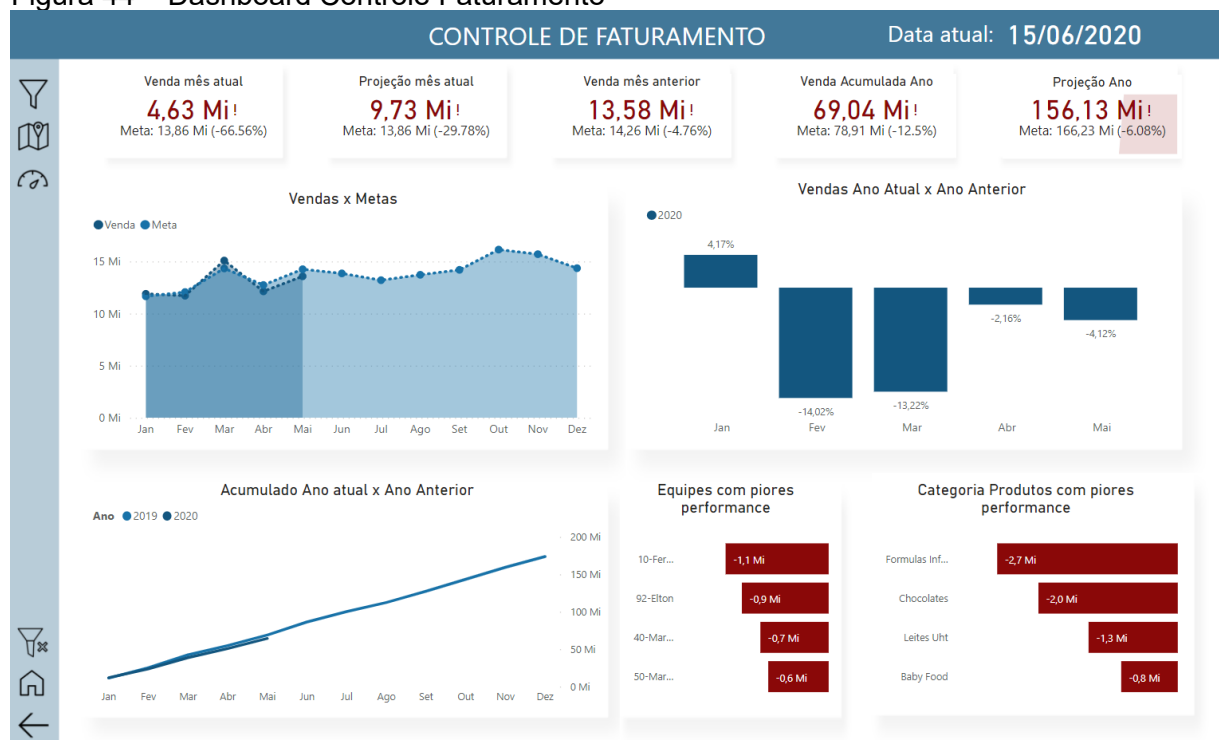
1. Ano/Mês
2. Data
3. Linha de produto
4. Grupo de produto
5. Categoria de produto
6. Categoria de cliente
7. Canal de vendas/equipe.

Figura 43 – Dashboard Estrutura Atendimento



Fonte: elaborado pelo autor.

Figura 44 – Dashboard Controle Faturamento



Fonte: elaborado pelo autor.

5.3 CONSIDERAÇÕES

No processo de desenvolvimento da solução foi utilizada a teoria apresentada no capítulo 2 e foi possível constatar que com o entendimento da aplicação de *Storytelling* de dados e a utilização dos visuais corretos nos *Dashboards*, teve-se uma melhor orientação facilitando o desenvolvimento e a entrega de um *Dashboard* muito mais perto do seu propósito de informar.

Tendo em vista o desafio de colocar em prática a teoria aprendida durante o desenvolvimento, foi identificado alguns pontos com maior complexidade. Para aplicar os conceitos de BI temos que utilizar algumas ferramentas. Embora a curva de aprendizado das ferramentas utilizadas fosse pequena, foi um grande desafio utilizá-las para aplicar os conceitos.

O PDI exigiu inicialmente, algumas configurações para fazer a interface com o SQL Server. Um ponto importante aqui é utilizar a versão do drive de conexão indicado tanto para a versão do Java (necessário para executar o PDI) quanto da versão do Banco de Dados Microsoft SQL Server. Usar o repositório de dados do PDI em um banco de dados, notou-se ser uma ótima prática.

Ainda na etapa de integração de dados, notou-se uma demora muito grande ao realizar a carga da tabela fato no DW. Após uma leitura da documentação oficial da aplicação, observou-se que ao realizar o processo de *lookup* para encontrar as SK não pode existir valores nulos. Para isso foi criado dois registros em todas as dimensões, um como: SK e NK como 0 (zero) nome “Não Encontrado” e outro registro como SK e NK como -3 com nome “Não se aplica”. Também foi substituído todos os códigos chave por -3. Isso garantiu que não tivéssemos valores nulos durante o *lookup*.

Na utilização do Microsoft Power BI Desktop, primeiramente notou-se que seria inviável trabalhar com muitos registros na tabela fato pois causa muita lentidão na aplicação, não possibilitando uma utilização fluída. Como o trabalho proposto foi apenas um protótipo, foi decidido trabalhar com apenas uma amostra dos dados. Apenas para registro, para contornar de forma mais correta, no dia a dia da empresa, é interessante trabalhar com a função de atualização incremental do Power BI.

Inicialmente foi exigido um estudo da ferramenta Power BI, a fim de propiciar o desenvolvimento adequado dos *Dashboards*. A linguagem DAX, notou-se fundamental para garantir uma boa construção das métricas e indicadores. Um estudo

dessa linguagem para um bom entendimento foi requerido e percebido um desafio grande.

O emprego do *Storyboard* e do Canvas para *Storytelling* de dados foi algo que exigiu um tempo considerável, porém todo tempo empregado nessa etapa foi recompensado durante o desenvolvimento do *Dashboard*.

Apesar das dificuldades encontradas a solução proposta atende os requisitos solicitados, notou-se que a solução trouxe facilidade para encontrar os indicadores mais importantes para o negócio e com isso facilita a tomada de decisão.

6 CONCLUSÕES E TRABALHOS FUTUROS

Este capítulo apresenta as conclusões obtidas com o desenvolvimento da solução proposta e sugestões de trabalhos futuros, usando os conceitos apresentados, que poderão ser desenvolvidos para aprofundar, complementar e até inspirar outras áreas de estudo.

6.1 CONCLUSÕES

Neste trabalho, foi desenvolvido uma solução de BI utilizando dados e indicadores da empresa estudo de caso. A solução se aprofundou em técnicas e conceitos para desenvolvimento de *Dashboard* utilizando *Storytelling* de dados. A ideia inicial desse projeto nasceu por uma dificuldade do autor em iniciar o desenvolvimento de um *Dashboard*, além de ter observado visuais que não entregavam a seu propósito de informar e ajudar nas decisões.

Toda solução de BI deve iniciar com uma boa modelagem e levantamento dos requisitos. Respeitar esse processo garante que todo o esforço para a integração de dados e desenvolvimento dos *Dashboards* seja justificado com as melhores escolhas e decisões por parte dos gestores.

Entender a necessidade do público-alvo e o contexto é fundamental para desenvolver melhores visualizações e análises. Desta forma, utilizar de perguntas estratégicas foi fundamental para que as necessidades fossem entendidas. O uso do Canvas para *Storytelling* de Dados foi muito importante e primordial para ter uma visão geral de como aplicar o *Storytelling* de dados na construção do *Dashboard*.

As visualizações de dados são recursos cruciais para o desenvolvimento do *Dashboard*. São elas que serão nossos “personagens” que contaram as histórias dos dados ao público-alvo. Entender a melhor forma de apresentar cada dado e informação foi muito importante na etapa de desenvolvimento do *Dashboard*.

A narrativa de dados, esclarece melhor a tomada de decisão pois cria uma fluidez na leitura e captação de insights com os visuais presentes no *Dashboard*. Novamente, as perguntas estratégicas para conhecer os indicadores mais importantes para o negócio foi fundamental nesta etapa. O *Storyboard* colaborou para criar uma narrativa prévia à implementação do *Dashboard*. Com isso pode-se ganhar tempo tendo menos ajustes e mudanças durante o desenvolvimento.

O maior desafio e complexidade deste trabalho, destacado como a dificuldade do autor em desenvolver *Dashboards* que entregam maior valor, foi resolvido com estudos apresentados no capítulo 2, como *Storytelling* de dados e visualizações de dados. Esses conceitos e técnicas ajudaram a construir *Dashboards* de forma mais orientada e precisa.

Com os requisitos definidos, as etapas de integração de dados finalizada e o planejamento do *Storytelling* de dados feito. Foi desenvolvido um *Dashboard* utilizando a ferramenta Power BI da Microsoft. O *Dashboard* apresentou uma solução para a Diretoria da empresa estudo de caso que pode visualizar os indicadores de forma rápida e sem necessidade de analistas para apresentar relatórios. Com isso conseguiu ter insights mais rápidos e decisões mais assertivas.

O desenvolvimento da solução foi feito utilizando metodologias e conceitos descritos no trabalho e concluindo esta etapa, conseguiu-se realizar todos os passos planejados de acordo com os requisitos. Em razão disto, é possível afirmar que este trabalho cumpriu com a proposta e que a utilização de conceitos de *Storytelling* de Dados em *Dashboards* para aplicações reais funcionam e agregam muito valor pois ajudam a informar melhor e com isso aperfeiçoam a tomada de decisão.

6.2 TRABALHOS FUTUROS

A partir dos estudos apresentados neste trabalho e no desenvolvimento da solução proposta é possível aprimorar, aprofundar e tomar novas frentes de estudo. Visto que esse tema pode ser utilizado em várias áreas, tornando essencial o domínio desses conceitos e técnicas, pois cada vez mais os dados estarão presentes nos mais variados campos de estudo. Utilizar dados de qualidade e apresentar de forma mais clara e que conecte o público-alvo é fundamental para o sucesso de qualquer projeto. Desta forma, temos como perspectivas futura para continuidade deste trabalho, os seguintes itens:

- Uso de *Storytelling* de dados na construção de infográficos em mídias de informação.
- Aprofundar os estudos em visualizações de dados. É um tema bastante abrangente e que pode ser mais bem explorado.

- Design de *Dashboard* foi um tema que poderia ser abordado neste trabalho, mas por razões de tempo não foi abordado. É um tema interessante para se aprofundar.
- Para a arquitetura da solução poderia ser desenvolvida 100% em nuvem (*cloud computing*). Demonstrando os passos e desafios de uma solução de BI totalmente na nuvem.
- Uso de *Storytelling* de dados para apresentação de dados em *Data Science (Big Data)* usando a Jornada do Herói para narrativa.
- Desenvolver um trabalho que faça avaliação final, através de pesquisas com os usuários do *Dashboard* a fim de identificar se a construção do *Dashboard* com *Storytelling* de dados melhorou a experiência do público-alvo.
- Apresentar na forma de um guia diretrizes que visam o aprimoramento de visualização de dados em Dashboards a fim de orientar desenvolvedores de Business Intelligence a produzirem com foco em *Storytelling* de dados considerando a escolha de um visual eficaz bem como a compreensão de contexto e público-alvo.

REFERÊNCIAS

ADAMSON, C. **Star schema**: The complete reference. Sebastopol: The McGraw-Hill Companies, 2010.

ASPIN, A. **Business intelligence with SQL server reporting services**. Apress. 2015.

BERENGUERES, J.; SANDELL, M.; FENWICK, A. **Introduction to data visualization & storytelling**: A guide for the data scientist. Stokes-Hamilton, 2019.

BEST OF BI. **SQL power architect**. Data modeling & profiling tool. 2020. Disponível em: <http://www.bestofbi.com/page/architect>. Acesso em: 04 nov. 2020.

BOUMAN, R.; VAN DONGEN, J. **Pentaho solutions**: Business intelligence e data warehousing com pentaho e MySQL. Wiley: Edição do Kindle, ago. 2009.

BRATH, R.; PETERS, M. Dashboard design: Why design is important. **DM Review**, v. 85, p. 1011285-1011289, 15 out. 2004. Disponível em: http://cs.furman.edu/~pbatchelor/csc105/articles/TUN_DM_ONLINE.pdf. Acesso em: 04 nov. 2020.

CARNEVALLI, S. **Data Storytelling Canvas**: Conte a história dos seus dados. 16 abr. 2020. Disponível em: <http://datastorytelling.com.br/data-storytelling-canvas-conte-a-historia-do-seus-dados/>. Acesso em: 05 maio 2020.

DAGAN, B. Business intelligence simply stated. **Natural Gas & Electricity**, v. 23, n. 10, p. 23-27, maio 2007. Disponível em: <https://vdocuments.net/business-intelligence-simply-stated.html>. Acesso em: 18 mar. 2020.

DATAPINE. **Absenteeism, Overtime hours, Training costs, Overall labor effectiveness (OLE)**. Disponível em: <https://public.datapine.com/#board/XpSRIGq2yshO0Uqg24N2b>. Acesso em: 10 nov. 2020.

DAVENPORT, T. H. **Why data storytelling is so importante - and why we're so bad at it**. 22 jan. 2015. Disponível em: <https://www2.deloitte.com/us/en/insights/topics/analytics/data-driven-storytelling.html>. Acesso em: 26 abr. 2020.

DAVENPORT, T. H.; HARRIS, J. G. **The architecture of business intelligence**. Mar. 2007. Disponível em: https://whitepapers.em360tech.com/wp-content/files_mf/white_paper/Accenture_BI_analytics_white_paper.pdf. Acesso em: 28 abr. 2020.

DUARTE, N. **DataStory**: Explain data and inspire action through story. Publicação Ideapress: Edição Kindle, 2019.

DYKES, B. **Data storytelling**: The essential data science skill everyone needs. 31 mar. 2016. Disponível em: <https://www.forbes.com/sites/brentdykes/2016/03/31/data->

storytelling-the-essential-data-science-skill-everyone-needs/#2c7e392552ad. Acesso em: 26 abr. 2020.

DYKES, B. **Effective data storytelling**: How to drive change with data, narrative, and visuals. Wiley: Edição Kindle, 2020.

ELIAS, D. **Entendendo o processo de ETL**. Canal tech, 11 jun. 2014. Disponível em: <https://canaltech.com.br/business-intelligence/entendendo-o-processo-de-etl-22850/>. Acesso em: 10 nov. 2020.

FEW, S. **Data visualization for human perception**. The encyclopedia of human-computer interaction. 2. ed. 2014. Disponível em: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception>. Acesso em: jun. 2020.

FEW, S. **Information dashboard design**: The effective visual communication of data. Sebastopol, CA: O'Reilly Media, Inc, 2006.

FEW, S. **Now you see it**: Simple visualization techniques for quantitative analysis. Oakland, California, USA: Analytics Press, 2009.

FEW, S. **Show me the numbers**: Design tables and graphs to enlighten. Burlingame, USA: Analytics Press, 2012.

FEW, S. **Statistical Narrative**: Telling compelling stories with numbers. Visual Business Intelligence Newsletter, jul./ago. 2009. Disponível em: https://www.perceptualedge.com/articles/visual_business_intelligence/statistical_narrative.pdf. Acesso em: 15 jun. 2020.

FEW, S. **Why Most Dashboards Fail**. 2007. Disponível em: <http://www.perceptualedge.com/articles/misc/WhyMostDashboardsFail.pdf>. Acesso em: 26 abr. 2020.

FOLEY, E.; GUILLEMETTE, M. G. **What is business intelligence?** Sherbrooke, Canadá: International Journal of Business Intelligence Research, 2010.

GARTNER. Glossary. **Dashboards**. Disponível em: <https://www.gartner.com/en/information-technology/glossary/dashboard> Acesso em: 12 jun. 2020.

GARTNER. Glossary. **Information delivery**. Disponível em: <https://www.gartner.com/en/information-technology/glossary/information-delivery> Acesso em: 12 jun. 2020.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

HOWSTON, C. **Successful business intelligence**: Secrets to make bi a killer app. New York, New York: The McGraw-Hill Companies, 2008.

INMON, W. H. **Building the data warehouse**. 4. ed. Indianapolis: Wiley Publishing, Inc. 2005.

INMON, W. H. **What is data warehouse?** 2000. Disponível em: <http://repository.binus.ac.id/2009-2/content/M0584/M058459913.pdf>. Acesso em: 17 mai. 2020.

KIMBALL, R. *et al.* **The data warehouse lifecycle toolkit**. Indianapolis: Wiley Publishing Inc, 1998.

KIMBALL, R.; CASERTA, J. **The data warehouse ETL toolkit**: Practical techniques for extracting, cleaning, confirming and delivering data. Indianapolis: Wiley Publishing Inc, 2004.

KIMBALL, R.; ROSS, M. **The data warehouse toolkit**: The definitive guide to dimensional modeling. Indianapolis: Wiley Publishing Inc, 2013.

KNAFLIC, C. N. **How I storyboard**. 25 ago. 2015. Disponível em: <https://www.storytellingwithdata.com/blog/2015/8/24/how-i-storyboard>. Acesso em: 30 jun. 2020.

KNAFLIC, C. N. **Storytelling com dados**: Um guia sobre visualização de dados para profissionais de negócios. Alta Books: Edição Kindle, 2017.

LOH, S. **BI na era do big data para cientistas de dados**: Indo além de cubos e dashboards na busca pelos porquês, explicações e padrões. Porto Alegre: Edição Kindle, 2014.

LUHN, H. P. A business intelligence system. **IBM Journal of Research and Development**, v. 2, n. 4, p. 314-319, 1958. Disponível em: <https://ieeexplore.ieee.org/abstract/document/5392644>. Acesso em: 30 jun. 2020.

MACHADO, F. N. R. **Tecnologia e projeto de data warehouse**: Uma visão Multidimensional. 6. ed. São Paulo: Saraiva, Edição Kindle, 2004.

MARR, B. **Key performance indicadores**: The 75 measures every manager needs to know. Great Britain: Person Education, 2012.

MICROSOFT. **Baixar o SQL Server Management Studio (SSMS)**. 2020. Disponível em: <https://docs.microsoft.com/pt-br/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-ver15>, Acesso em: 04 nov. 2020.

MICROSOFT. **Entenda o esquema em estrela e a importância para o Power BI**. 09 set. 2019. Disponível em: <https://docs.microsoft.com/pt-br/power-bi/guidance/star-schema>. Acesso em: 10 nov. 2020.

MICROSOFT. **Visão geral dos cubos OLAP do service manager para análise avançada**, 2019. Disponível em: <https://docs.microsoft.com/pt-br/system-center/scsm/olap-cubes-overview?view=sc-sm-2019> Acesso em: 16 jul. 2018

PRATS, N. **Tell a story through your dashboards**. 2017. Disponível em: <https://www.nathanprats.com/dashboards-tell-story/>. Acesso em: 10 nov. 2020.

PRATT, M. K.; FRUHLINGER, J. **What is business intelligence? Transforming data into business insights**. 16 out. 2019. Disponível em: <https://www.cio.com/article/2439504/business-intelligence-definition-and-solutions.html>. Acesso em: 28 abr. 2020.

PULVIRENTI, A. S.; ROLDAN, M. C. **Pentaho data integration 4 Cookbook**. Packt Publishing: Edição Kindle, 2011.

RASMUSSEN, N. H.; BANSAL, M.; CHEN, C. Y. **Business dashboards: A visual catalog for design and deployment** (Locais do Kindle 437-442). Wiley: Edição Kindle, 2009.

RUSSO, M.; FERRARI, A. **The definitive guide to DAX: Business intelligence with Microsoft Power BI, SQL Server Analysis Services, and Excel**. Pearson Education, Inc., 2020.

RYAN, L. **Visual data storytelling with tableau**. Addison-Wesley, 2018.

SAMPIERI, R. H.; COLLADO, C. F.; LUCIO, M. D. P. B. **Metodologia de pesquisa**. Porto Alegre: Peso, 2013.

SBROCCO, J. H. T. C. **Metodologias ágeis: Engenharia de software sob medida**. São Paulo: Érica, 2012.

SEGEL, E.; HERR, J. Narrative visualization: Telling stories with data. **IEEE Transactions on Visualization and Computer Graphics**, v. 16, n. 6, p. 1139-1148, nov. 2010. Disponível em: <https://ieeexplore.ieee.org/abstract/document/5613452>. Acesso em: 28 abr. 2020.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. Florianópolis: UFSC, 2005.

SPITZER, R. D. **Transforming performance measurement: Rethinking the way we measure and drive organizational success**. USA: Amazon Books, 2007.

STACHOWIAK, R.; RAYMAN, J.; GREENWALD, R. **Oracle datawarehousing and business intelligence solutions**. Indianapolis, Indiana: Wiley Publishing, 2007.

STEELE, J.; ILLINSKY, N. **Beautiful Visualization: Looking at data through the eyes of experts**. O'Reilly Media, Inc, 2010.

TUFTE, E. R. **The visual display of quantitative information**. 2. ed. Cheshire, Connecticut, 2013.

TURBAN, E. *et al.* **Business intelligence: Um enfoque Gerencial**. Porto Alegre: Bookman, 2009.

VAZQUEZ, C. E.; SIMÕES, G. S. **Engenharia de requisitos: Software orientado ao negócio**. Rio de Janeiro: Brasport, 2016.

WAZLAWICK, R. S. **Metodologia de Pesquisa Para Ciência da Computação**. Rio de Janeiro: Elsevier. 2014.

WEXLER, S.; SHAFFER, J.; COTGREAVE, A. **The big book of dashboards**: Visualizing your data using real-world business scenarios. New Jersey, USA: Wiley, 2017.

XAVIER, A. **Storytelling**: Histórias que deixam marcas. Rio de Janeiro: Best Business, 2015.

YAU, N. **Data points**: Visualization that means something. Wiley: Edição do Kindle, 2013.