Universidade de Brasília - UnB
Faculdade UnB Gama - FGA
Engenharia de Software

Implementação de *Business Intelligence* nas corporações: estudo de caso.

Autor: Vinicius da Silva Carvalho
Orientador: Prof. Dr. Vandor Roberto Vilardi Rissoli

Brasília, DF
2019

Vinicius da Silva Carvalho

Implementação de *Business Intelligence* nas corporações: estudo de caso.

Monografia submetida ao curso de graduação em Engenharia de Software da Universidade de Brasília, como requisito parcial para obtenção do título de Bacharel em Engenharia de Software.

Universidade de Brasília - UnB

Faculdade UnB Gama - FGA

Orientador: Prof. Dr. Vandor Roberto Vilardi Rissoli

Brasília, DF

2019

Vinicius da Silva Carvalho
Implementação de *Business Intelligence* nas corporações: estudo de caso./ Vinicius da Silva Carvalho. – Brasília, DF, 2019-
73 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Vandor Roberto Vilardi Rissoli

Trabalho de Conclusão de Curso – Universidade de Brasília - UnB
Faculdade UnB Gama - FGA , 2019.

1. *Business Inteligence*. 2. Análise de Dados. 3. Informação Corporativa. I. Prof. Dr. Vandor Roberto Vilardi Rissoli. II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Implementação de *Business Intelligence* nas corporações: estudo de caso.

CDU 02:141:005.6

Vinicius da Silva Carvalho

Implementação de *Business Intelligence* nas corporações: estudo de caso.

Monografia submetida ao curso de graduação em Engenharia de Software da Universidade de Brasília, como requisito parcial para obtenção do título de Bacharel em Engenharia de Software.

Trabalho aprovado. Brasília, DF, 10 de dezembro de 2019:

Dr. Vandor Roberto Vilardi Rissoli
Universidade de Brasília (Orientador)

Dr. Wander Cleber Pereira da Silva
Universidade de Brasília

MSc. Ricardo Ajax Dias Kosloski
Universidade de Brasília

Brasília, DF
2019

Agradecimentos

Agradeço, primeiramente, à minha família por todo o apoio e compreensão durante minha caminhada. Em especial, ao meu pai, Antônio Morais Carvalho Neto, por toda sua dedicação e esforço pela educação de seus filhos e pelo exemplo de paternidade que obtive em casa, sempre preocupado com nosso futuro e nos ajudando a sermos o melhor de nós mesmos.

Agradeço ao meu orientador, prof^{o}. Vandor Roberto Rissoli, por todo o apoio durante as disciplinas, projetos e este Trabalho de Conclusão de Curso. Sua orientação foi fundamental durante toda a graduação e sua confiança em meu trabalho me fez crescer como profissional. E meus sinceros agradecimentos aos professores Wander Cléber Pereira e Ricardo Ajax, por me apoiarem e contribuírem imensamente na construção deste trabalho.

Agradeço em especial à Orc'estra Gamificação, empresa júnior de engenharia de software, que me ofereceu grandes oportunidades de aprendizado, de crescimento pessoal e profissional além de grandes amigos que me apoiaram durante essa jornada. A minha formação se tornou completa pelas experiências vividas na empresa júnior, sendo elas de sucessos e de fracassos, me tornando um líder capaz de alcançar o sucesso.

Resumo

A otimização do espaço físico é uma das principais tendências de estudo nas empresas, pois um ambiente organizado motiva e estimula a maior produtividade dos colaboradores. Ao considerar a importância do espaço físico para uma área logística, em que é preciso armazenar produtos em estoque, diversos outros fatores contribuem além de produtividade, como maior capacidade de produção e de armazenamento, todos estes convergindo para o aumento de operações. Todos os processos de armazenamento nas empresas geram dados, e estes, armazenados e integrados, podem ser estudados para promover melhorias no processo. Este estudo de caso é realizado por meio de uma pesquisa-ação com objetivo de analisar os dados provenientes da armazenagem de produtos de uma empresa industrial do ramo alimentício, com a utilização de técnicas de análise de inteligência de negócios. Seguiu-se as etapas de construção e levantamento de requisitos de sistemas e indicadores significativos para o setor de estoque da empresa em estudo, as fases de desenvolvimento de sistemas de *Business Intelligence*, focadas em coletar dados, integrá-los e oferecer um indicador conciso. Os resultados apresentados foram por experiências subjetivas dos usuários e os painéis de indicadores-chave para o departamento logístico.

Palavras-chaves: Inteligência de Negócios, Análise de Dados, Informação Corporativa.

Abstract

The optimization of the physical space is one of the main trends of study in the companies, because an organized environment, motivates and stimulates the greater productivity of the collaborators. When considering the importance of physical space for a logistics area, where it is necessary to store products in stock, several other factors contribute in addition to productivity, such as greater production capacity and storage, all of which converge to increase operations. All the storage processes in the companies generate data and these, stored and integrated, can be studied to promote improvements in the process. This case study is carried out through an action research with the objective of analyzing the data from the storage of products of an industrial food company, using business intelligence analysis techniques. This was followed by the stages of building and gathering system requirements and significant indicators for the company's inventory sector under study, the Business Intelligence system development phases, focused on collecting data, integrating it and offering it. a concise indicator. Results were presented by subjective user experiences and key indicator panels for the logistics department.

Key-words: Business Intelligence, Data Analysis, Corporate Information.

Lista de Figuras

Figura 1 – Estrutura típica de um sistema de BI. 22
Figura 2 – Arquitetura da modelagem *Star Schema* em banco de dados. 26
Figura 3 – Arquitetura da modelagem *Snowflake* em banco de dados. 27

Figura 4 – Fases da Pesquisa-Ação. 34
Figura 5 – Atividades e fluxo na fase de Diagnóstico da pesquisa-ação. 35
Figura 6 – Atividades da fase de Validação da pesquisa-ação. 38
Figura 7 – Visão geral do desenvolvimento de um sistema de BI. 38
Figura 8 – Fluxo de atividades da Primeira Fase. 39
Figura 9 – Fluxo de atividades da Segunda Fase. 42
Figura 10 – Ferramentas de apoio ao desenvolvimento de escrita. 44
Figura 11 – Ferramentas de apoio ao desenvolvimento de BI. 49

Figura 12 – Produção de dados pelo sistema WMS em 2019. 53
Figura 13 – Fluxo de Movimentações realizadas em 2019. 54
Figura 14 – Área de estoque monitorada pelo sistema WMS até o momento. 55
Figura 15 – Modelagem do BI. 57
Figura 16 – Exemplo do *workflow* do processo ETL que popula a tabela dimensão. 58
Figura 17 – Exemplo de JOB que executa as transformações em sequência. 59
Figura 18 – Painel apresentando indicadores de gestão de estoque. 61
Figura 19 – Painel com indicadores sobre produtos em estoque. 62

Lista de tabelas

Tabela 1 – Quantidade de tuplas de dados no Data Mart. 59

Lista de Quadros

2.1 Definição de Dado, Informação e Conhecimento. 24

3.2 Classificação da metodologia. 32
3.3 Cronograma de atividade da primeira fase do trabalho. 41
3.4 Cronograma de atividades da segunda fase do trabalho. 43

4.5 Matriz Dimensão X Indicador. 56

Lista de abreviaturas e siglas

BI	Business Intelligence
BPMN	Business Process Modeling Notation
DW	Data Warehouse
ETL	Extract, Transform, Load
FGA	Faculdade do Gama - Universidade de Brasília
GPS	Global Positioning System
IC	Inteligência Competitiva
MDX	Multidimensional Expression
OLAP	Online Analytical Processing
PDF	Portable Document Format
SCM	Supply Chain Management
SGBD	Sistema Gerenciador de Banco de Dados
TCC	Trabalho de Conclusão de Curso

Sumário

Sumário 11

1 INTRODUÇÃO 13
1.1 Contextualização 14
1.2 Justificativa 15
1.3 Problema de Pesquisa 16
1.4 Objetivos 17
1.4.1 Objetivo Geral 17
1.4.2 Objetivos Específicos 17
1.5 Metodologia 17
1.6 Organização do Trabalho 17

2 REFERENCIAL TEÓRICO 19
2.1 *Business Intelligence* 20
2.1.1 Benefícios e Dificuldades 21
2.1.2 Componentes do BI 21
2.1.3 Base de Dados Corporativa 23
2.1.3.1 Dado, Informação e Conhecimento 23
2.1.3.2 Modelagem em BI 24
2.1.3.2.1 Esquema Estrela 25
2.1.3.2.2 Floco de Neve 26
2.2 Inteligência Competitiva 27
2.2.1 Armazenagem logística 28
2.2.2 *Suply Chain Management* 29

3 PROPOSTA 31
3.1 Metodologia 31
3.1.1 Classificação de Pesquisa 31
3.1.1.1 Abordagem da Pesquisa 31
3.1.1.2 Natureza da Pesquisa 32
3.1.1.3 Procedimentos 32
3.1.1.3.1 Pesquisa Bibliográfica 33
3.1.1.3.2 Pesquisa-ação 33
3.2 Processo Metodológico 39
3.2.1 Primeira fase 39
3.2.2 Segunda fase 41

3.3 Suporte Tecnológico . . . 43
3.3.1 Desenvolvimento teórico . . . 43
3.3.2 Desenvolvimento prático . . . 44
3.3.2.1 SQL Power Architect . . . 44
3.3.2.2 MySQL . . . 45
3.3.2.3 HeidiSQL . . . 45
3.3.2.4 Pentaho Data Integration . . . 46
3.3.2.5 Pentaho Server . . . 46
3.3.2.6 Pentaho Workbench . . . 47
3.3.2.7 Microsoft Power BI . . . 47

4 DESENVOLVIMENTO . . . 50
4.1 Entendimento de Negócio . . . 50
4.2 Entendendo os dados . . . 50
4.3 Pré-processamento . . . 51
4.3.1 Entendendo as variáveis do Banco de Dados . . . 51
4.3.2 Análise dos Dados . . . 52
4.4 Solução de *Business Intelligence* . . . 55
4.4.1 Fontes externas . . . 55
4.4.2 Modelagem de BI . . . 56
4.4.3 ETL e Data Mart . . . 58
4.4.4 OLAP . . . 59
4.4.5 Visualização . . . 60

5 CONCLUSAO . . . 63
5.1 Lições aprendidas pelo pesquisador . . . 63
5.2 Trabalhos futuros . . . 64

REFERÊNCIAS BIBLIOGRÁFICAS . . . 66

APÊNDICES 69

APÊNDICE A – *SCRIPT* DE CRIAÇÃO DE BANCO MYSQL . . . 70

1 Introdução

Desde o surgimento dos sistemas computacionais, um dos principais objetivos das organizações tem sido o de armazenar um grande volume de dados. Essa tendência ficou ainda mais visível com a grande queda nos custos de aquisição de hardware, devido a implantação de novas tecnologias, tornando possível armazenar quantidades cada vez maiores de informações. Novas e mais complexas estruturas de armazenamento foram desenvolvidas, tais como: depósitos de dados virtuais, bibliotecas virtuais e armazenamento em nuvem, pois além de armazenar estes dados é necessário buscá-los com facilidade (CAMILO; SILVA, 2009).

Atividades rotineiras como enviar e-mails, solicitar serviços por aplicativos, utilizar o GPS (*Global Positioning System*) do smartfone e tirar uma foto armazenando-a em nuvem são todas atividades que viraram comuns e acessíveis podendo ocorrer várias vezes ao longo de um dia. Cada uma dessas atividades podem gerar milhares de metadados, que unidos podem ser interpretados e gerar a informação útil ao usuário. Essa captura de informações é realizada de forma automática e o usuário pode nem ter observado a coleta de todas essas informações.

No meio corporativo não é diferente, todas as transações realizadas por um software geram metadados que são armazenados em bancos de dados virtuais, que isolados podem não transmitir uma informação coesa diretamente, porém ao se analisar o conteúdo, o conceito e o histórico daquele dado, pode-se obter uma informação útil do processo. Esta informação pode ser suficiente para a compreensão de situações-chave podendo contribuir na otimização deste, seja na diminuição de recursos no processo, redução do número de atividades ou custos, tempo de espera, entre outros.

O armazenamento de grande volume de dados não é uma proposta recente e remonta do início da maior utilização de sistemas de informação pelas empresas, na década de 1950 (CHEN; CHIANG; STOREY, 2012). Um dos destaques é o crescente aumento de fontes de informações externas, advindas de novas tecnologias como as redes sociais, da interoperabilidade de sistemas corporativos e acesso a dados públicos abertos à comunidade.

Desde a década de 1980, uma ampla variedade de softwares tem sido criada para auxiliar na análise de dados qualitativos e quantitativos. Com o volume de dados armazenados crescendo diariamente, responder uma questão tornou-se crucial: O que fazer com todos esses dados armazenados? (CAMILO; SILVA, 2009).

Conforme Demchenko et al. (2013), a quantidade de dados cresce exponencialmente, assim como a sua possibilidade de manipulação e utilização. Neste sentido, um

dos maiores desafios está no manuseio e na análise desses dados de forma global, a fim de que sejam geradas informações úteis aos diferentes níveis envolvidos e para promover conhecimento que poderá aprimorar os negócios, como na maior compreensão dos tipos e estilos de clientes ao seu negócio, realizar análises mercadológicas mais acuradas, monitorar produtos e serviços de concorrentes, gerar histórico de vendas de produtos e promover impacto através de tecnologias.

1.1 Contextualização

O mercado atual tem grande foco em otimizar processos e trabalhar diversos fatores que contribuam e agreguem valor ao cliente final. Dessa forma, de acordo com Ballou (1993), a logística ganhou destaque no cenário empresarial buscando formas de melhorar e prover a rentabilidade nos serviços de entrega aos clientes e consumidores. É possível à logística agregar valor ao produto através do planejamento, organização e controle das atividades de movimentação e armazenagem visando a facilitação do fluxo de compra e venda.

A gestão logística é uma área em constante crescimento nas organizações, gerando maior rentabilidade e eficácia nos negócios, e está estabelecendo novos patamares competitivos nos mercados. Em termos práticos, a área de logística não gera receita à empresa, então estudos de redução de custos nesse processo são realizados e estimulados por grandes empresas.

Segundo Viana (2000), a administração de estoques é importante para qualquer organização, pois, de maneira geral, ela visa controlar e buscar por melhores resultados em termos de lucratividade e produtividade. As empresas comerciais podem ter um grande gasto com armazenamento de produtos e/ou insumos, por isso deve-se administrar com eficiência e exatidão a movimentação de entrada e saída de materiais necessários à organização, considerando que em todos os passos há gastos envolvidos.

A execução de todos esses processos de logística que objetivam o aprimoramento de armazenagem, apoiadas em ferramentas de gestão ou planilhas, são ineficientes no processo de melhoria se não considerada a etapa de análise dos dados. A partir das informações coletadas é possível identificar pontos de atenção, de retenção e gargalos no processo, em que recebendo o tratamento adequado, podem aumentar a competitividade de inovação na empresa.

Dentro de um setor com tantas operações, a correlação de diversos dados de uma empresa não é trivial, pois necessita obter a quantidade de informações suficientes para gerar uma análise mais específica e particular, que apresente padrões a serem otimizados. A série de dados provenientes de movimentações físicas, compras de insumos, quebras e avarias devido a locomoção dos produtos são armazenados em diversas fontes e por

diferentes setores, gerando interdependências entre diversas áreas para obter uma visão mais ampla e geral do processo.

O tratamento e a manipulação dos dados são os pontos fortes das principais ferramentas de apoio a análise de dados qualitativos, trazendo grande vantagem na forma como os dados podem ser acessados, encontrados e visualizados. Segundo Dolan e Ayland (2001), esses softwares oferecem excepcionais facilidades para o armazenamento, o gerenciamento e a manipulação dos dados, permitindo trabalhar com volumes proibitivos no processo manual.

A existência (ou não) de viés no processo analítico é provavelmente o aspecto mais debatido em meios acadêmicos. Os problemas relatados muitas vezes não estão relacionados com os softwares de análise em si, mas referem-se à forma como o sistema ou software é utilizado. Mesmo os usuários mais entusiastas, algumas vezes, mostram-se indecisos quanto a esse aspecto, especialmente porque existem situações nas quais as facilidades desses softwares podem levar ao viés nos processos analíticos (LAGE; GODOY, 2008).

Os autores Dembkowski e Hanmer-Lloyd (1995) argumentam que a facilidade e a velocidade obtidas com o uso de uma ferramenta analítica podem levar o usuário a gerar um volume de textos codificados muito maior que o necessário, dificultando a tarefa de análise e saindo do foco das questões iniciais de pesquisa. Na mesma linha de argumento de Dolan e Ayland (2001) e Lage e Godoy (2008), a grande facilidade para gerar nós e árvores de nós (agrupamentos hierárquicos de conceitos relacionados a partes de dados) estimula o usuário do sistema a elaborar centenas de estruturas, dificultando os processos analíticos, podendo causar uma superficialidade de análise e o comprometimento do resultado final da pesquisa.

Diante desse cenário, é perceptível que apesar da existência de softwares de análise robustos que permitem grande capacidade de armazenamento e com muitas funcionalidades ao usuário, há a dificuldade em alcançar o estado ótimo com a ferramenta de apoio, devido a vários fatores que influenciam, como desconhecimento sobre a utilização do sistema e seus limites, muita codificação que pode deixar a análise mais complexa, a falta, na geração de indicadores, de alerta no âmbito gerencial ao invés de várias métricas isoladas, entre outros.

1.2 Justificativa

A análise dos dados relevantes para a tomada de decisões gerada pelos sistemas de *Business Intelligence* (BI) tornou-se uma grande aliada para a gestão e planejamento de indicadores estratégicos associados a muitos processos. Esse tipo de sistema objetiva viabilizar a análise de informação por diversos níveis da organização, buscando, principalmente, aumento da produtividade, minimização dos riscos, assertividade das decisões,

velocidade das respostas, visualização de tendências e diminuição de custos.

A principal diferença entre as aplicações do conceito de BI atuais e de tempos anteriores é a tecnologia. Essa foi bastante evoluída apresentando mais recursos e maior facilidade de utilização, além da eficiência em seus resultados. Segundo Gangadharan e Swami (2004), a principal utilização do BI nos dias atuais é agregar, em uma mesma ferramenta, variadas informações produzidas e disponibilizadas pelos mais diversos setores de uma organização. Assim, a alta administração tem acesso facilitado e sistematizado a um conjunto de informações passíveis de variadas análises, que subsidiarão as suas tomadas de decisões e geração de conhecimento.

O conhecimento e a utilização de técnicas de análises de dados mais adequadas a um contexto da empresa podem garantir a melhor apresentação de informações aos seus gestores, de forma a agregar valor ao trabalho e aumentar o monitoramento de indicadores do setor ou da empresa em geral.

1.3 Problema de Pesquisa

Segundo Pozo (2000), a gestão logística é indispensável para o sucesso de uma organização pois contempla uma visão empresarial que direciona o melhor desempenho nas empresas, tendo como meta reduzir o *lead-time* (tempo de processamento do pedido) entre a solicitação de compra, a produção e a entrega de modo que o cliente receba seus bens ou serviços no momento oportuno, com o preço desejado. Desta forma, buscar táticas de administração logística visando aperfeiçoar os resultados agregando valor aos clientes finais e oferecendo um atendimento de qualidade é um fator competitivo muito importante no mercado atual.

A logística consiste em dispor dos materiais necessários no momento apropriado e no lugar certo ao menor custo global para o estabelecimento. Não é somente uma questão de técnica de armazenagem e de movimentação de embalagens e transportes; é também um método de direção e gestão que determina o grau de utilização das instalações fabris, o volume de estoque necessário, o relacionamento e a disposição de fornecedores, entre outros (POZO, 2000).

De acordo com a problemática apresentada, a questão de pesquisa norteadora deste trabalho é *"Como potencializar a armazenagem logística com base em dados históricos do processo de deslocamento de produtos em estoque?"*.

1.4 Objetivos

1.4.1 Objetivo Geral

O objetivo geral deste trabalho é apresentar um modelo de análise de dados direcionado ao processo de armazenagem logística, permitindo a visualização de indicadores-chave no processo.

1.4.2 Objetivos Específicos

Como parte do desdobramento do objetivo geral de pesquisa, descrito na seção 1.4.1, foram propostos os seguintes objetivos específicos:

- Analisar o potencial do BI na área em estudo de forma a garantir a entrega de valor aos usuários da empresa;
- Estruturar um projeto de análise e sintetização de dados provenientes do processo de armazenagem que utilize métodos de inteligência de negócios;
- Permitir a aplicação do projeto de aperfeiçoamento de análise de dados na empresa participante do estudo.

1.5 Metodologia

A metodologia adotada para o desenvolvimento deste trabalho consiste em compreender os processos logísticos de armazenagem e projetar a conciliação de dados por meio de métodos de Inteligência de Negócios (*Business Intelligence*), apoiando-se nos princípios mercadológicos, políticas de estoque e armazenagem, geração de indicadores e análise dos dados como a efetivação prática desse trabalho.

A pesquisa realizada é um estudo de caso de natureza qualitativa e quantitativa, realizado em uma instituição localizada no Distrito Federal. De forma procedimental, é utilizado a pesquisa-ação como norteadora do trabalho prático, baseada em diagnóstico da empresa, planejamento da ação, tomada de ação e avaliação do que foi executado. No capítulo 4 deste documento a forma de aplicação é melhor descrita.

1.6 Organização do Trabalho

Este Trabalho de Conclusão de Curso (TCC) está organizado e estruturado em cinco capítulos, sendo eles:

1. Introdução: refere-se à contextualização, levantamento da questão de pesquisa, justificativa e definição dos objetivos do trabalho;

2. Referencial teórico: objetiva fornecer ao leitor o conhecimento necessário para compreender a pesquisa realizada, por meio de revisão de literatura;

3. Proposta: a apresentação da proposta do trabalho, definindo a metodologia a ser utilizada e as ferramentas de suporte tecnológico estabelecidas para este trabalho. Este capítulo busca apresentar as técnicas utilizadas para a realização da pesquisa, definindo as atividades a serem desempenhadas para conclusão do trabalho;

4. Desenvolvimento: a visualização resumida do processo de desenvolvimento prático da proposta, apresentando como as ferramentas foram utilizadas e em qual situação, além de permitir visualizar o resultado de cada fase do projeto de implementação;

5. Conclusão: apresenta o status do trabalho, as conclusões possíveis com sua implementação e possíveis ganhos do projeto.

2 Referencial Teórico

Independente do ramo de atividade, é importante tratar os dados de uma empresa como parte de seu patrimônio, ou seja, deve ser considerado como um capital da empresa, devendo este ser bem administrado permanentemente. A análise desses dados é uma das formas mais eficazes de nutrir os direcionamentos, planejamentos e iniciativas de um negócio. Dados corporativos explorados significa ganhos em visualização rápida e útil de informação interna e externa como indicativos e sinais para as tomadas de decisão, preponderando a capacidade de melhorias de processos.

Os dados de um departamento de uma empresa podem revelar muitas informações e conhecimentos a cerca dela mesma, como seu comportamento, suas estratégias e até suas fraquezas, que são oportunidades de melhorias. Pode-se imaginar, por exemplo, algumas aplicações no contexto empresarial, como:

- Vendas: informações sobre clientes, produtos e compras;
- Recursos Humanos: informações completas sobre funcionários, históricos, salários, encargos sociais e benefícios;
- Contabilidade: pagamentos, recebimentos, saldo de contas;
- Instituições de Ensino: informações pessoais dos alunos, matrículas em cursos e notas.

Como apresentado, o banco de dados constitui uma parte essencial em toda empresa, que é de armazenar informações de contexto geral e informações específicas e destinadas para o setor da empresa ou área. A utilização futura deste dado armazenado gera a necessidade de garantir a sua fidedignidade, a integridade dessa informação.

Um dado é persistente se sobrevive ao processo que o criou, dados em armazenamento não volátil. A gestão de dados envolve tanto a definição de estruturas para o armazenamento de informações quanto os mecanismos que preveem a manipulação da informação.

As informações que fluem, e se relacionam dentro de um sistema de gestão, são gerenciadas no Sistema Gerenciador de Bancos de Dados (SGBD), os quais consistem em uma coleção de tabelas relacionadas, que são geralmente integradas, vinculadas ou referenciadas uma a outra. O SGBD tem a funcionalidade de criar, editar, atualizar, armazenar e recuperar dados em tabelas de banco de dados (SIPPERT; SEIFERT; PORCIUNCULA, 2018). Os SGBDs são projetados para gerir grandes massas de informação e

surgiram em resposta aos métodos iniciais de gestão informatizados de dados comerciais (SILBERSCHATZ; SUNDARSHAN; KORTH, 2016). Este sistema permite colocar dados à disposição de usuários para uma consulta assegurando-se dos direitos atribuídos aos mesmos, sendo o banco local armazenado em um dispositivo por um usuário, ou compartilhado, podendo ser acessado por vários usuários, sendo esta a sua grande vantagem.

Para o armazenamento dos dados, pode-se pensar em vários meios de gravação, como arquivos de texto, planilhas, papéis, entre outros. Ao aglutinar os dados de forma lógica, organizada, específica e não volátil, gera-se um banco de dados. Banco de Dados é uma coleção de dados persistentes, usada pelos sistemas de aplicação para trabalhar um tema específico. Esses dados só podem ser removidos dessa coleção posteriormente, por uma requisição explícita ao Sistema de Gerenciamento deste Banco de Dados, dessa forma, ele se torna um dado persistente.

2.1 *Business Intelligence*

Visando apoiar a Inteligência Estratégica das organizações, o *Business Intelligence* (BI) surge como uma tecnologia que permite transformar dados em informação significativa. Dados são fontes de grande riqueza, pois é possível realizar estudos extraindo comportamentos a partir de diferentes cenários. O objetivo do *Business Intelligence* é auxiliar na tomada de decisão apoiando-se em uma visão global do negócio, oferecendo informações consistentes e atualizadas. Visando o acesso rápido às informações, os recursos de BI proporcionam a disseminação do conhecimento fazendo com que os usuários estejam alinhados na estratégia da organização (LORENZZETTI, 2016).

O *Business Intelligence*, ou Inteligência de Negócios, define um conjunto de regras e técnicas objetivando organizar adequadamente um grande volume de dados, visando transformá-los em depósitos estruturados de informações. BI é uma evolução da modelagem de dados onde é possível promover a sua estruturação gerando informações valiosas, como os históricos por meio de ferramentas analíticas (PALESTINO, 2001).

> *Business Intelligence* – conceito que engloba um vasto conjunto de aplicações de apoio à tomada de decisão que possibilitam um acesso rápido, partilhado e interativo das informações, bem como a sua análise e manipulação; através destas ferramentas, os utilizadores podem descobrir relações e tendências e transformar grandes quantidades de informação em conhecimento útil (PALESTINO, 2001, p.10).

2.1.1 Benefícios e Dificuldades

Os benefícios organizacionais do uso das informações processados pelos sistemas de BI podem ser percebidos por meio da geração da receita, redução de custos e efetividade gerencial. Esses benefícios estratégicos estão associados ao desempenho da organização (WRIGHT, 1993).

A inteligência dá a empresa uma vantagem competitiva por meio da geração da receita. A melhora no desempenho é esperada através do sucesso da introdução de um novo produto/serviço ou desenvolvimento de um novo mercado (TEO; CHOO, 2001). As empresas que possuem sistemas avançados que monitoram as atividades de seus concorrentes exibem maior lucratividade. Sistemas de BI permitem que as empresas aproveitem a rentabilidade da racionalização da cadeia de suprimentos para aumentar a competitividade. Além disso, a possibilidade de fechar novos negócios é aumentada quando um sistema integrado facilita as transações.

Percebe-se que a informação traz mais benefícios que os dados para a camada tomadora de decisão de uma organização. Segundo Primak (2008), a informação faz parte da base da construção do conhecimento. Infelizmente muitas das iniciativas de *Business Intelligence* (BI) não alcançam o sucesso. As empresas buscam no BI formas para melhorar o desempenho organizacional e o diferencial competitivo. Apesar das vantagens e benefícios encontrados nesta solução, algumas tentativas de adoção não vão adiante devido a barreiras encontradas pelo percurso.

As dificuldades mais comuns são encontradas dentro da própria organização (ambiente interno). Entre elas estão a barreira cultural, falta de visão estratégica, inexistência ou má qualidade dos dados, centralização da solução e escolha incorreta da ferramenta de BI para os usuários.

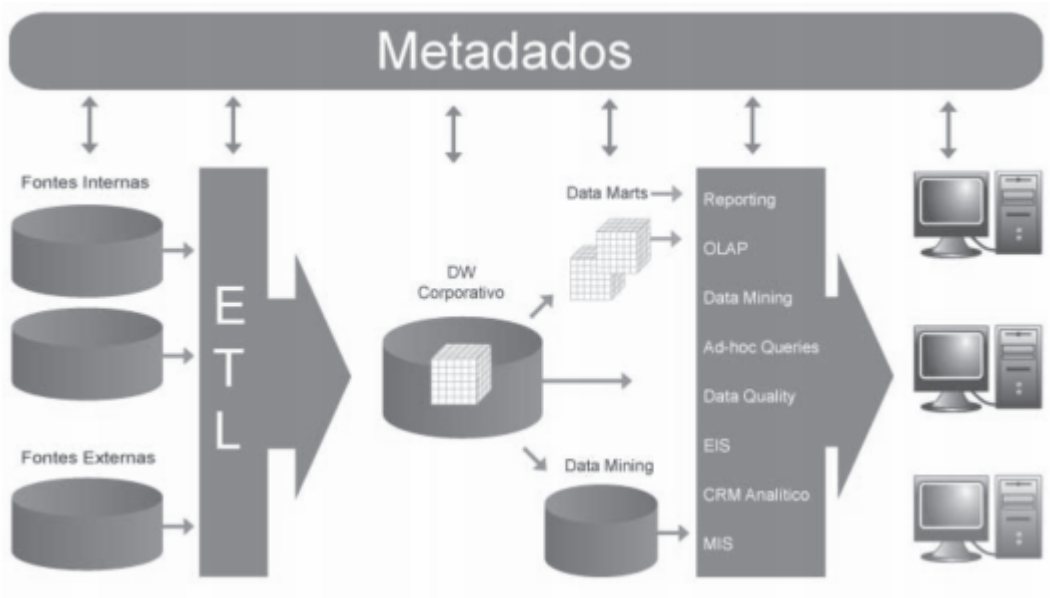
O BI interfere na forma que a organização pensa e age no dia a dia. E, como toda mudança, há empecilhos a serem enfrentados. A barreira cultural na empresa é um grande problema a ser encarado. Muitas pessoas não querem mudar sua forma de trabalho pelo enraizamento do processo atual e pela comodidade de continuar como está. Também, muitas vezes, é devido a insegurança de terem seus trabalhos ameaçados pela nova tecnologia. Seus benefícios não serão visíveis se a empresa não dispuser de uma estrutura organizacional analítica séria e responsável a disposição dos gestores.

2.1.2 Componentes do BI

Um sistema de *Business Intelligence*, encarado na perspectiva da sua vertente tecnológica, deve ser enquadrado na infraestrutura global dos sistemas de informação da organização. Por um lado, deve-se ter sempre em mente que um sistema de BI não subsiste por si próprio – está ligado, de forma umbilical, às fontes de dados subjacentes, sejam

elas os sistemas transacionais ou as planilhas de suporte; enfim, tudo o que se possa considerar um repositório primário de informação resultante dos processos de negócio da organização. Por outro lado, é necessário perceber a interação entre o conhecimento produzido e os seus destinatários (usuários finais), que, através das várias interfaces e ferramentas de visualização, tiram partido do que foi produzido, filtrado e sintetizado (SEZÕES; OLIVEIRA; BAPTISTA, 2006).

Figura 1 – Estrutura típica de um sistema de BI.



Fonte: (SEZÕES; OLIVEIRA; BAPTISTA, 2006).

Os sistemas de BI podem ser analisados a partir de duas perspectivas: técnica e de negócios. Do ponto de vista técnico, são um conjunto integrado de ferramentas, produtos, tecnologias e softwares utilizados para coletar dados de fontes heterogêneas dispersas e posteriormente consolidá-los num único banco de dados (OLSZAK; ZIEMBA, 2012). Essas etapas são apresentadas na figura 1 e descritas a seguir.

- Ferramentas para extrair, transformar e carregar dados (*ETL - Extract, Transform and Load*) - principais responsáveis pela transferência de dados dos sistemas de transação e Internet para o *Data Warehouse* (DW). É o processo mais crítico e demorado na construção de um DW, pois consiste na extração dos dados de bases heterogêneas, na transformação e limpeza destes dados e na carga dos dados na base do DW;

- *Data Warehouse* (DW) - banco de dados para armazenamento de dados consolidados e analisados. O DW é a espinha dorsal da infraestrutura que suporta vários processamentos analíticos e aplicações visuais;

- Ferramentas analíticas OLAP (On-Line Analytical Processing) - permitem que os usuários acessem, analisem e compartilhem as informações que estão armazenadas no *Data Warehouse*;
- Ferramentas de *Data Mining* - permitem a descoberta de padrões diferentes, generalizações, regularidades e regras em dados;
- Ferramentas para relatórios e *ad hoc* - permitem a criação e a utilização de diferentes relatórios sintéticos;
- Camada de apresentação - aplicações, incluindo interfaces gráficas e multimídia cuja tarefa é fornecer aos usuários informações de uma forma confortável e acessível.

Do ponto de vista organizacional, os sistemas de BI possuem uma metodologia e filosofia específica para trabalhar com a informação e com o conhecimento, propiciando uma comunicação aberta e compartilhada, juntamente com uma abordagem holística e analítica dos processos de negócios nas organizações (OLSZAK; ZIEMBA, 2012).

2.1.3 Base de Dados Corporativa

Os bancos de dados corporativos centralizam uma coleção de dados comuns aos vários sistemas de informação de toda a empresa, tais como: dados de pessoas físicas e jurídicas, estrutura organizacional do estabelecimento, dos departamentos, das áreas e das equipes. Uma base corporativa permite a melhor integridade dos dados, evitando as duplicidades e as inconsistências resultantes de uma gestão descentralizada.

Como apresentado, o banco de dados constitui uma parte essencial em qualquer empresa, armazenando informações de contexto geral a todas as corporações e informações específicas e destinadas para o segmento. A finalidade deste dado armazenado gera a necessidade de rastrear sua confiabilidade, garantindo a integridade da base de dados. Visto que a informação é tão importante na maioria das organizações, cientistas da computação têm desenvolvido um vasto conjunto de conceitos e técnicas de gerenciamento de dados devido ao aumento de complexidade em armazenamento e a chegada de tecnologias que facilitam o acesso a esses dados.

2.1.3.1 Dado, Informação e Conhecimento

Segundo DAVENPORT (1998), dados representam um conjunto de fatos objetivos e estruturados sobre um evento e não contêm nenhum elemento de julgamento ou interpretação. A informação representa um significado e um propósito atribuídos aos dados de um evento, influenciados pela análise do seu criador no contexto da realidade na qual está inserido. Nesse sentido, a informação gerada está sujeita à interpretação de seu criador, exigindo a mediação humana e sendo transmitida através de uma mensagem para um ou

mais receptores. O termo "Conhecimento"é uma combinação de informação contextualizada, valores e experiências acumuladas, o que resulta em sua natureza intuitiva e é difícil de traduzir em termos lógicos. Davenport apresenta definições destes conceitos conforme a tabela 2.1.

Dado	Informação	Conhecimento
Simples observação sobre o estado do mundo	Dados dotados de relevância e propósito	Informação valiosa da mente humana
Facilmente estruturado Facilmente obtido por máquinas Frequentemente quantificado Facilmente transferível	Requer unidade de análise Exige consenso em relação ao significado Exige necessariamente a mediação humana	Difícil estruturação Difícil captura em máquinas Frequentemente tácito Difícil transferência

Fonte: (DAVENPORT, 1998).

Quadro 2.1 – Definição de Dado, Informação e Conhecimento.

Dados são registros de símbolos quantificados e quantificáveis de acordo com uma sintaxe pré-definida, enquanto informação é uma abstração informal e matéria-prima para o conhecimento, isto é, dados têm uma natureza sintática enquanto informação tem natureza semântica (SETZER, 1999).

Para Miranda (1999), a definição de dados se dá por "um conjunto de registros qualitativos ou quantitativos conhecido e que organizado, agrupado, categorizado e padronizado adequadamente, transforma-se em informação". Dessa forma, identifica-se que os dados precisam ser estruturados e organizados para que possam se transformar em informação e serem úteis para a tomada de decisão. A informação é matéria-prima fundamental para o entendimento de uma empresa, seu nicho de atuação e, consequentemente, para a formulação ou adequação ágil de estratégias assertivas e inovadoras (PINTO, 2012).

2.1.3.2 Modelagem em BI

A modelagem é uma etapa de concepção e visualização de um modelo de armazenamento de dados. Para a modelagem dos dados armazenados em um *Data warehouse* e em um *Data Mart* é utilizada a modelagem dimensional. Para Kimball e Ross (2011), a modelagem dimensional é uma técnica que possibilita a criação de um modelo de dados dimensional.

O modelo dimensional é constituído por um conjunto de medidas que descrevem aspectos de negócios. Esta modelagem permite sumarizar e estruturar os dados para dar suporte à análise de dados. Três elementos formam esse modelo, são eles: Fatos, Dimensões e Medidas.

De acordo com Kimball e Ross (2011), a tabela de fatos é a principal tabela do modelo. Os fatos são coleções de itens de dados, compostas de dados de medida e de contexto. Essas coleções são compostas pelas medições numéricas que representam a evolução dos negócios de uma organização. A tabela fato registra o dado que será analisado e é composto pela chave primária e um conjunto único de valores de chaves de dimensões.

As tabelas de dimensão contêm as descrições de negócio, são os elementos das tabelas fato do negócio. Cada dimensão pode ter vários níveis hierárquicos para proporcionar um melhor entendimento e uma melhor visualização dos indicadores. Os atributos das dimensões são os principais atributos usados para obter vistas do processo de negócio, tais como filtro nas consultas, agrupamentos e relatórios (DÍAZ, 2012). As medidas ou variáveis são os atributos numéricos que representam os indicadores que mostram a evolução do negócio da empresa.

Existem principalmente duas abordagens dentro da modelagem multidimensional, o modelo *Star Schema* e o modelo *Snow Flake* (DÍAZ, 2012). O *Star Schema* ou Esquema Estrela consiste em estruturar a informação em processos, vistas e medidas em forma de estrela. Tem como característica básica a presença de dados redundantes para proporcionar um melhor desempenho. Em termos de desenho, esse esquema é composto por uma tabela de fatos no centro para o fato objeto de análise, e uma ou várias tabelas auxiliares chamadas tabelas de dimensões para cada ponto de vista da análise que participa da descrição do fato (DÍAZ, 2012). Nesse esquema, a consulta ocorre inicialmente nas tabelas de dimensões e depois na tabela de fatos, garantindo a precisão dos dados através de uma estrutura completa.

O segundo modelo, o *Snow Flake*, é um esquema derivado do modelo Estrela, o qual as tabelas de dimensões se normalizam em diversas tabelas. Com isso, a tabela de fatos deixa de ser a única tabela que se relaciona com as outras, e assim surgem novas uniões (DÍAZ, 2012).

2.1.3.2.1 Esquema Estrela

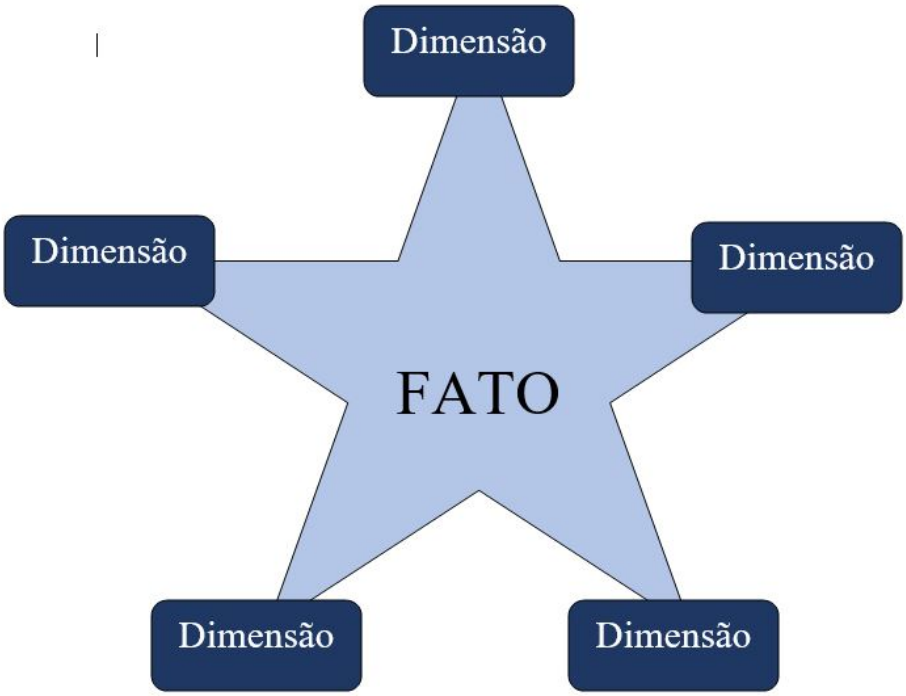
O *Star Schema* (esquema em estrela) é talvez o modelo mais utilizado na concepção de um *data warehouse*. O conceito pressupõe a criação de tabelas dimensionais que ficam ligadas entre si através de uma tabela de fato, conforme apresentado na figura 2. A sua interligação baseia-se num esquema lógico e simples: as tabelas dimensionais contêm as definições das características dos eventos, enquanto as tabelas de fatos, por sua vez, armazenam os fatos decorridos e as chaves estrangeiras para as características respectivas que se encontram nas tabelas dimensionais.

Sua composição típica possui uma grande entidade central denominada fato (*fact table*) e um conjunto de entidades menores denominadas dimensões (*dimension tables*),

arranjadas ao redor dessa entidade central, formando uma estrela.

Já Palestino (2001), define o modelo estrela como sendo um conjunto de tabelas Fatos (nas quais se concentram os dados de interesse, passíveis de manipulação numérica e estatística), e tabelas Dimensão (tabelas satélites que possuem as chaves de entrada do modelo, além das informações descritivas de cada dimensão).

Figura 2 – Arquitetura da modelagem *Star Schema* em banco de dados.



Fonte: Autor.

Contudo, neste modelo, as dimensões não são normalizadas, apresentando uma possível desvantagem. Caso uma dimensão possua mais de um atributo e esse atributo seja um subatributo de outra, uma coluna pode aparecer repetidas vezes no modelo de dados. Desta forma, é necessário que o menor nível de detalhamento de uma tabela Fato esteja em uma tabela Dimensão, ou seja, é necessária a aplicação de granularidade no modelo de dados respeitando o modelo de negócios.

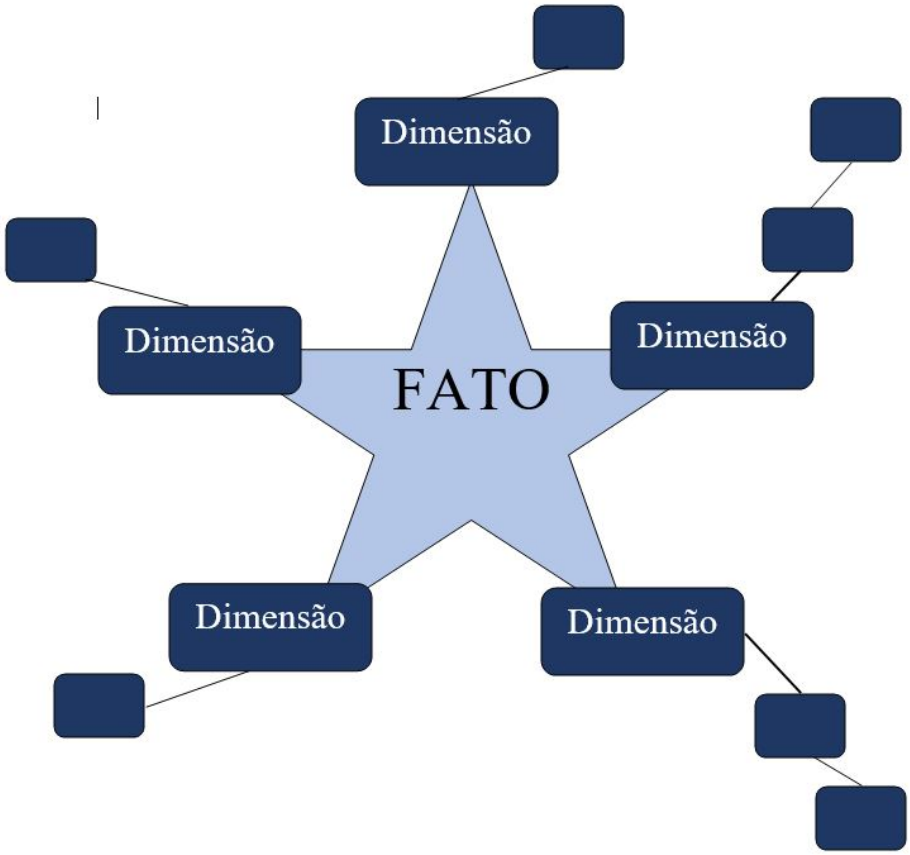
2.1.3.2.2 Floco de Neve

Na modelização dos dados, o esquema *Snowflake* é uma evolução mais complexa do *Star Schema*, sendo o resultado da decomposição de uma ou mais dimensões que possuem hierarquia entre seus membros. O seu objetivo é normalizar as dimensões, agrupando os dados de cada uma em várias tabelas (e não apenas numa única grande tabela como no esquema em estrela), evitando assim redundâncias.

O modelo de dados multidimensional é usualmente similar a uma estrela em sua arquitetura, obtendo um fato no centro da estrela e as dimensões ao seu redor, porém permitindo a decomposição das dimensões, como visualizado na figura 3.

É uma modelagem de fácil entendimento pelos desenvolvedores de sistemas OLTP, pois é resultado da terceira forma normal como um projeto relacional, permitindo preservar a utilização de meios de armazenamento.

Figura 3 – Arquitetura da modelagem *Snowflake* em banco de dados.



Fonte: Autor.

2.2 Inteligência Competitiva

A Inteligência Competitiva (IC) é um subconjunto de *Business Intelligence*, que foca no monitoramento do ambiente externo e usa recursos públicos para localizar e desenvolver informações sobre os clientes e os concorrentes. Desta forma, proporciona uma imagem balanceada do ambiente externo para os tomadores de decisão (OLSZAK; ZIEMBA, 2012).

O conceito de IC tem fortes fundamentos da ciência militar e tem uma rica história que remonta a mais de 5.000 anos (TAO; PRESCOTT, 2000). Nos últimos anos, esse

tema retornou com popularidade, especialmente no campo da gestão. O conceito tem sido estudado com diferentes títulos. A IC é considerada um sistema de *scanning* ambiental que integra o conhecimento de todos na empresa (AHITUV; ZIF; MACHLIN, 1998). Refere-se às informações sobre o ambiente externo de negócios que possam afetar a posição competitiva de uma empresa.

Para Tao e Prescott (2000), a Inteligência Competitiva é o processo de previsão de recurso em relação à dinâmica competitiva e fatores diversos que podem ser usados para aumentar a vantagem competitiva, tais como regulamentações governamentais, tarifas e cultura do país. A Inteligência Competitiva engloba os efeitos potenciais (ameaças e oportunidades) criados por todos os elementos externos ao ambiente de negócios que têm impacto sobre a competitividade atual e futura de uma empresa.

A Inteligência Competitiva é usada no acompanhamento da atividade dos concorrentes diretos e indiretos, incluindo suas atividades de negócios em geral, táticas e estratégias relativas a várias questões como a penetração de mercado, desenvolvimento de produtos, registros de patentes entre outros.

2.2.1 Armazenagem logística

A gestão de estoque, em interação com o transporte e a armazenagem, é uma função fundamental da logística integrada. Uma gestão efetiva de estoques é aquela que garante o nível de serviço desejado com o mínimo custo logístico total.

A grande maioria dos custos de armazenagem – aluguel, mão de obra, depreciação de instalações e equipamentos de movimentação - são fixos e indiretos. Estas duas características dificultam respectivamente o gerenciamento da operação e a alocação de custos.

A elevada parcela de custos fixos na atividade de armazenagem faz com que os custos sejam proporcionais à capacidade instalada. Desta maneira, pouco importa se o armazém está quase vazio ou se está movimentando menos produtos do que o planejado. Ainda assim, a maior parte dos custos de armazenagem continuarão ocorrendo, pois, na sua grande maioria, estão associados ao espaço físico, aos equipamentos de movimentação, ao pessoal e aos investimentos em tecnologia.

Para tornar a situação ainda mais complexa, é importante lembrar que a demanda pela atividade de armazenagem não é constante nem ao longo dos meses, nem ao longo dos dias do mês ou da semana. Um exemplo claro disso é a concentração da expedição nos últimos dias do mês. Isso tende a levar ao super dimensionamento da capacidade para atender os dias de pico, ou, ao contrário, faz com que o armazém opere acima da sua capacidade, prejudicando o nível de serviço – através de falhas, avarias e atrasos. Assim, medidas que venham amortecer essas variações na demanda serão sempre positivas do

ponto de vista da expedição. Principalmente quando as razões para os picos de demanda são induzidas por políticas internas como, por exemplo, cotas mensais de vendas, prazo para faturamento, falta de uma política de ressuprimento contínuo com os clientes etc.

O fato de os custos de armazenagem serem indiretos dificulta a sua alocação aos produtos e clientes, pois a alocação, neste caso, é realizada através de rateios, deixando-os sujeitos a distorções. Para minimizar as distorções é importante que:

- Os itens de custos sejam contabilizados de acordo com a sua função (movimentação, acondicionamento, administração) e não por contas naturais (depreciação, mão-de-obra);
- A alocação seja condizente com o real consumo de recursos na operação.

2.2.2 *Suply Chain Management*

Suply Chain Management (SCM) ou Gerenciamento da Cadeia de Suprimentos é o sistema cujo recurso permite a integração de uma empresa com as demais organizações envolvidas no processo produtivo (clientes e fornecedores), buscando otimizar o funcionamento como um todo, com reduções de custos e ganhos de produtividade e qualidade.

Segundo a ISO 9000 (2015), a qualidade dos produtos e serviços de uma organização é determinada pela capacidade de satisfazer os clientes e pelo impacto pretendido e não pretendido nas partes interessadas pertinentes, inclui não apenas sua função e desempenho pretendidos, mas também seu valor percebido e o benefício para o cliente. Ou seja, a qualidade é o nível de perfeição de um processo, serviço ou produto entregue pela empresa, de maneira que atenda às exigências definidas pelas normas internas e, é claro, pelos seus clientes.

O setor de *Suply Chain* nas empresas é o principal alinhador entre os custos e a qualidade em relações logísticas, e possuem grandes necessidades tecnológicas para apoio a tomadas de decisões. De acordo com Chopra e Meindl (2003), os sistemas de tecnologia da informação (TI) são importantes em todo estágio da cadeia de suprimentos, pois permitem que as empresas reúnam e analisem as informações que são necessárias para a tomada de decisão. Eles podem ser segmentados de acordo com os estágios da cadeia de suprimentos e possuem níveis diferentes de funcionalidade que podem receber e apresentar informações diferenciadas e analisá-las para solucionar problemas a curto ou a longo prazo, podem ser utilizados para tomar decisões de estratégia, planejamento ou operação.

A maioria das empresas não têm definido em seus processos claramente a logística reversa, pois representa um percentual muito pequeno em relação aos valores movimentados pela logística direta. É por isso que não existem no mercado sistemas de informação específicos para seu controle. Isso representa uma deficiência para a empresa, pois se não

houver sistemas de informação que operacionalizem o gerenciamento de resíduos, estes muitas vezes ficam na empresa ocupando um espaço inadequado, podendo gerar penalidades legais e obstruir o setor produtivo da empresa.

Uma das principais atividades de *Suply Chain* é monitorar os indicadores de armazenagem no estoque, como ocupação do estoque, segmentação de produtos por ocupação, giro de estoque, produtos de maior venda, entre outros.

3 Proposta

A metodologia adotada para o desenvolvimento deste trabalho consiste em compreender os processos logísticos de armazenagem de produtos alimentícios e projetar a conciliação de dados por meio de métodos de Inteligência de Negócios (BI), apoiando-se nos princípios mercadológicos, políticas internas de armazenagem, geração de métricas e indicadores e análise dos dados como a efetivação prática deste trabalho.

A proposta deste trabalho é apresentar um painel de indicadores-chave da área de gestão de estoque que possa auxiliar no gerenciamento e controle do estoque em tempo presente e oferecer insumos confiáveis para justificar tomadas de decisão pela liderança.

3.1 Metodologia

De acordo com Prodanov e Freitas (2013), a metodologia de pesquisa adotada neste trabalho pode ser classificada nas seguintes categorias: classificação quanto à natureza da pesquisa, quanto à escolha do objeto de estudo, quanto à técnica de coleta de dados e quanto à técnica de análise de dados. A classificação escolhida em cada categoria busca garantir conhecimento amplo sobre diferentes técnicas de solução de BI, possibilitando a adaptação de algumas para cada contexto corporativo. Nas seções seguintes, são apresentadas as categorias e suas respectivas classificações de pesquisa adotadas para este trabalho.

3.1.1 Classificação de Pesquisa

Segundo a classificação apresentada por Oliveira (2011), a pesquisa realizada possui abordagem qualitativa e quantitativa, apoiando-se em técnicas de coleta de dados, conforme apresentado resumidamente na tabela 3.2. A pesquisa é realizada em uma empresa alimentícia na região do Distrito Federal, responsável pela fabricação, comercialização e distribuição de bebidas para três estados próximos, envolvendo em sua amostra o setor relacionado a processos logísticos, mais especificamente, as atividades de armazenamento de estoque de acabados, ou seja, de produtos prontos para a venda e consumo.

Nos próximos tópicos serão tratados, detalhadamente, cada uma destas categorias.

3.1.1.1 Abordagem da Pesquisa

A classificação deste trabalho em relação à abordagem da pesquisa é definida como qualitativa e quantitativa. Segundo Oliveira (2011), abordagens de cunho qualitativo buscam o significado dos dados obtidos, capturando, além da aparência do fenômeno

Natureza da Pesquisa	Escolha do Objeto de Estudo	Procedimentos
Qualitativa e Quantitativa	Estudo de caso único	Pesquisa-Ação

Fonte: Autor.

Quadro 3.2 – Classificação da metodologia.

estudado, suas essências. Dessa forma, a pesquisa qualitativa envolve a obtenção e a análise dos dados descritivos por meio do contato direto entre o pesquisador e o fenômeno estudado. Já a pesquisa quantitativa é caracterizada pelo emprego da quantificação, tanto nas modalidades de coleta de informações quanto no tratamento delas por meio de técnicas estatísticas.

A abordagem qualitativa não busca definir resultados de maneira quantificável, procurando aprender de forma subjetiva o evento estudado. Neste trabalho, os dados serão coletados em ambiente controlado, buscando caracterizar os resultados como, por exemplo, satisfatórios ou não, caracterizando-a, assim, uma pesquisa qualitativa. Sua utilização, durante este trabalho, teve como objetivo entender, detalhadamente, o funcionamento de processos logísticos, viabilizando a adaptação em análise de dados para um contexto limitado.

3.1.1.2 Natureza da Pesquisa

A pesquisa científica visa conhecer um ou mais aspectos de determinado assunto de forma técnica. Para tanto, deve ser sistemática, metódica e crítica. Segundo Prodanov e Freitas (2013), o produto da pesquisa científica deve contribuir para o avanço do conhecimento humano. Na vida acadêmica, a pesquisa é um exercício que permite despertar o espírito de investigação diante dos trabalhos e problemas sugeridos ou propostos pelos professores e orientadores.

No decorrer deste trabalho será realizada uma pesquisa básica relacionada ao tema, a fim de responder a questão de pesquisa e contextualizar a possível solução. Não foi possível realizar uma pesquisa muito aprofundada sobre o conceito de BI, uma vez que este não é foco do trabalho, e ser apenas uma pesquisa bibliográfica, em que aspectos históricos não seriam fundamentais. No entanto, a pesquisa foi realizada de forma a esclarecer conceitualmente os processos e tecnologias utilizadas para atender às características de solução do problema prático que será explorado como estudo de caso neste TCC.

3.1.1.3 Procedimentos

Quanto aos procedimentos, o trabalho foi dividido em duas etapas. Durante a primeira etapa do procedimento seguido está a pesquisa bibliográfica a respeito do assunto e

levantamento dos objetivos em uso da tecnologia. Nesse caso, segundo Prodanov e Freitas (2013), esse procedimento é feito a partir do levantamento de referências teóricas publicadas por meio de livros, artigos, páginas da web, entre outros recursos interessantes que possam agregar aos principais assuntos do trabalho proposto. Através deste procedimento foram apresentadas as definições em uso, suas teorias e referências para este trabalho, formulando uma base conceitual para prosseguimento nas tarefas.

Já na segunda etapa do trabalho, o procedimento seguido será o de pesquisa-ação. Como Prodanov e Freitas (2013) afirmam, a pesquisa-ação pressupõe uma participação planejada do pesquisador na situação-problema a ser investigada. Além da participação do pesquisador, o objeto de estudo da pesquisa-ação não pode ser representado por um conjunto de variáveis que poderiam ser analisadas separadamente. A utilização da pesquisa-ação durante este trabalho se concretizou pela necessidade da presença do pesquisador na empresa gerando a participação ativa na situação estudada, o que possibilitou ampla análise de todo o contexto.

3.1.1.3.1 Pesquisa Bibliográfica

A pesquisa bibliográfica teve o objetivo de estabelecer uma base teórica e definir os conceitos que são utilizados no trabalho. A definição de *Business Intelligence*, por exemplo, é modificada por diferentes autores, alguns classificando-a como ferramenta, uma série de métodos e até mesmo como apenas um conceito. Dessa forma, tornou-se necessário o levantamento das bases e conceitos entrelaçados ao BI, compreendendo uma porção de tempo dedicada para leitura e interpretação sobre o assunto em bases científicas.

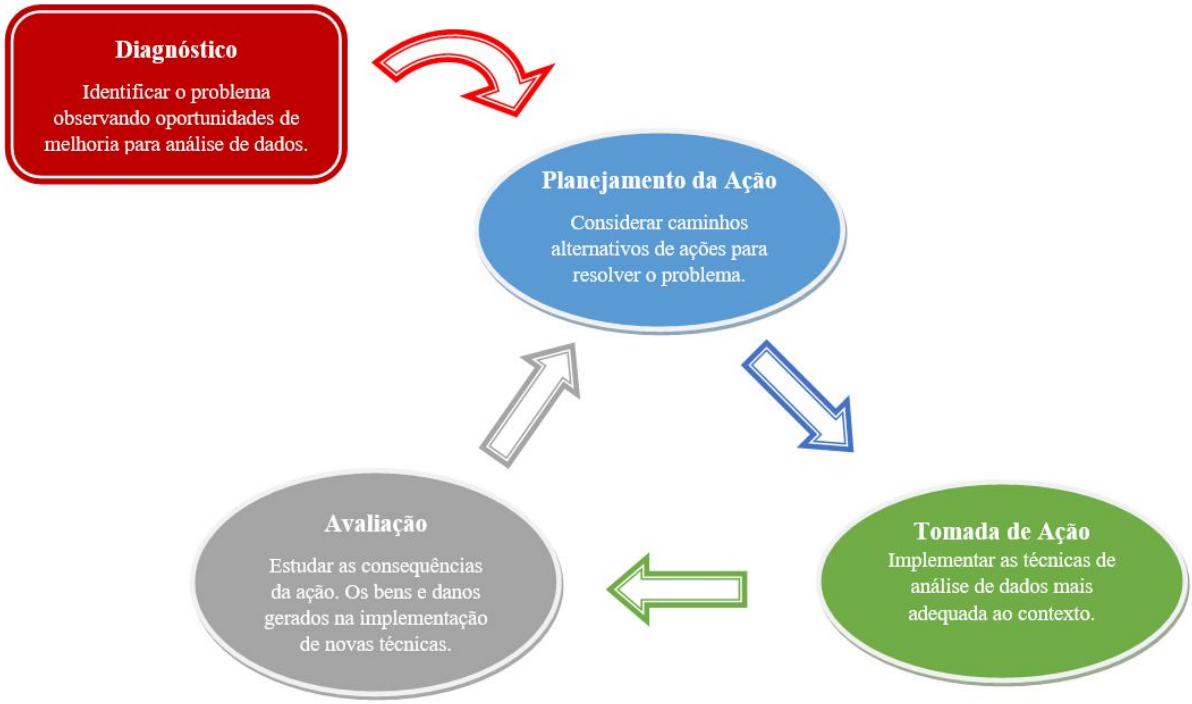
3.1.1.3.2 Pesquisa-ação

Para atender ao objetivo traçado em tempo hábil, esta será fundamentada em quatro fases, sendo elas de Diagnóstico, Planejamento, Ação e Avaliação, como mostra a figura 4. Essas fases não seguem um processo linear de desenvolvimento, em que após a fase de diagnóstico, são realizadas iterações evolutivas, muito útil para gerar novas versões cada vez mais complexas do sistema.

Observado esse processo, pode-se descrever os principais objetivos das fases e sua forma de funcionamento, conforme apresentado a seguir:

- **Diagnóstico:** É a primeira etapa do processo, realizada apenas uma vez. Esta fase constitui-se em analisar o ambiente corporativo em que o processo de análise será inserido, verificar o processo diário dos colaboradores, o fluxo de informações necessárias e quais dados são mais relevantes no cotidiano e em análises gerenciais.

Figura 4 – Fases da Pesquisa-Ação.



Fonte: Autor.

- **Planejamento:** Nesta fase espera-se identificar e propor as ações a serem realizadas, avaliando as variáveis de ambiente, tempo e técnicas que solucionem o problema identificado anteriormente. Espera-se que no final dessa fase sejam apresentadas quais técnicas de análises de dados serão utilizadas.

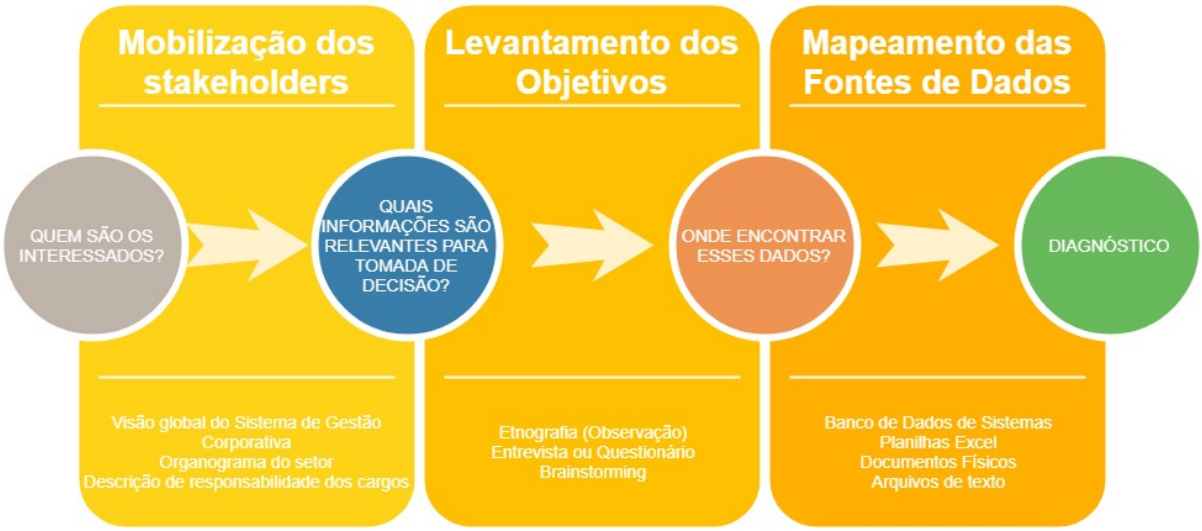
- **Ação:** É a completa execução do planejamento realizado. Implementação das técnicas de análise propostas.

- **Avaliação:** Etapa final do ciclo em que são levantados os resultados das ações que foram tomadas para resolução do problema. O método de avaliação é baseado em métricas de software que possam avaliar a performance do sistema, coerência de indicadores, periodicidade de consulta aos dados, usabilidade, entre outros.

Neste trabalho foi realizado apenas um ciclo da metodologia pesquisa-ação apresentada, sendo executadas as etapas de Diagnóstico, de Planejamento e de Ação. Este fluxo de perguntas como inicializadoras de tarefas dentro das fases auxiliam a tornar as atividades objetivas e mais claras. Este modelo pode ser visualizado na figura 5.

Conforme apresentado por Silva, Silva e Gomes (2016), o processo de BI possui atividades que necessitam ser bem compreendidas para que a implementação da solução tenha efetividade, ou seja, é considerado importante a fase de diagnóstico do negócio do cliente, o entendimento da estrutura organizacional, processos e modelos de gestão. Com

Figura 5 – Atividades e fluxo na fase de Diagnóstico da pesquisa-ação.



Fonte: Autor.

posse da informação, foi definido seguir um fluxo de atividades de desenvolvimento destacando a opinião do principal *stakeholder*, o usuário final do sistema. Dessa forma, será trabalhado e descrito no decorrer deste trabalho as atividades de levantamento de objetivos do sistema de BI, utilizando algumas técnicas de Levantamento de Requisitos para o melhor entendimento das atividades do setor, dos indicadores operacionais e estratégicos. Alguma dessas atividades e sua sequência é determinada pela figura 5.

As etapas de diagnóstico consistem em:

- **Mobilização dos *stakeholders*:** É necessária a identificação dos stakeholders, ou seja, das partes interessadas do processo de *Business Intelligence*. Precisa ser explicitado quem são os principais beneficiados da solução além de adquirir o apoio total da alta gestão. Deve-se mostrar os benefícios que o projeto trará para a organização, além de promover discussões sobre os diversos assuntos relacionados à gestão utilizando os indicadores mais apropriados.

  Na empresa participante do estudo, foi realizado um treinamento sobre BI com pessoas-chave de alguns setores da empresa, estes desenvolvem todos os modelos de painel utilizados nos departamentos, porém só conseguem trabalhar com o BI nos tempos livres, pois esta tarefa não faz parte da sua função. Estas pessoas, além da alto gerência, são apontadas como partes interessadas neste projeto.

- **Levantamento dos objetivos:** Antes mesmo de mapear os dados, é preciso listar as necessidades informacionais que os gestores possuem. Para isso, é necessário fazer rodadas iterativas com os usuários da solução e, neste momento, não há preocupação com a viabilidade das informações nessa atividade. Nas entrevistas serão levantados

os principais anseios dos gestores, que deverão partir do questionamento: "Quais informações necessito para tomar decisões mais assertivas?".

- **Mapeamento das fontes dos dados:** Só aqui começa a preocupação com a existência ou não dos dados necessários para gerar as informações solicitadas na etapa anterior. É analisado todo o tipo de fontes de dados existentes na empresa, tais como CDs, pendrives, rede corporativa, arquivos de texto, planilhas, banco de dados, etc. Se os dados existirem, é feito o mapeamento estruturado de como o dado entrará na etapa de desenvolvimento da solução de BI. No caso de inexistência de tais dados, os participantes do levantamento serão informados, para que seja provocado na organização o questionamento sobre a necessidade desses dados ou de serem criadas alternativas de fontes para sua busca.

Na fase de planejamento da ação é utilizada a teoria estudada na pesquisa bibliográfica para programar os próximos passos do trabalho. Ao chegar nessa etapa espera-se que já se tenha avaliado o contexto, documentado às possibilidades de melhoria do processo atual, à indicação dos dados do setor que são relevantes para a tomada de decisão gerencial, entre outros aspectos sobre o ambiente e fluxo de atividades.

Essa etapa exige muito conhecimento sobre as ferramentas que serão utilizadas, os processos envolvidos da empresa, o potencial de análise de dados e a visão ampla de todo o projeto. Pode ser considerada a fase que exige maior conhecimento do pesquisador, em que sugerem melhorias a um sistema, e estas são avaliadas e verificadas junto ao professor orientador deste projeto, a fim de validar a viabilidade e a correta execução.

A etapa de Ação é onde ocorre o desenvolvimento ou melhoria da aplicação de BI. Por questões de acessos e proteção da propriedade intelectual, o que já foi desenvolvido pelos funcionários da empresa até o momento não será considerado no desenvolvimento deste trabalho, dessa forma, o primeiro ciclo de desenvolvimento de pesquisa-ação contemplará o início do projeto de desenvolvimento, e as iterações seguintes abrangerão novas versões desse projeto, aumentando sua complexidade. A fase de execução contemplará de forma mais detalhada as atividades propostas para a segunda etapa do trabalho, sendo que no fim desta fase espera-se ter uma versão entregável aos usuários finais da empresa utilizada para o estudo.

A parte de Ação, neste trabalho, representa a atividade:

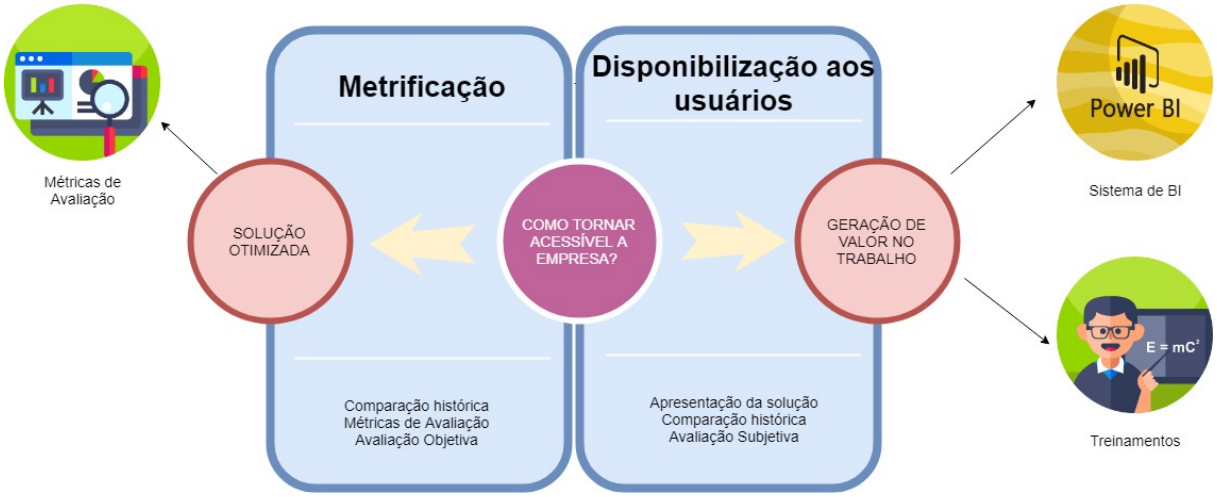
- **Construção da solução BI:** Nessa etapa é dado o início à construção tecnológica da solução de *Business Intelligence*. É feita a modelagem, o processo de ETL, carga do banco de dados no *Data Warehouse*, criação de *Data Marts*, processo de OLAP e criação de relatórios.

O *Data Warehouse* (DW), ou Armazém de Dados, é uma coleção de dados integrados, orientados por assunto, variáveis com o tempo e não voláteis, usados para dar suporte ao processo gerencial de tomada de decisão. Trata-se de um conjunto de dados derivados de bancos de dados operacionais heterogêneos, com o objetivo de dar suporte à tomada de decisão, apresentando-os de forma analítica, detalhados ou resumidos. Não é uma base de dados transacional, os acessos são somente para carga, através de aplicativos extratores, no processo chamado de ETL (*Extration, Transformation and Load*) e consultas por parte da equipe que gerenciará a informação (COSTA et al., 2019).

A fase de avaliação levantará quais foram os resultados obtidos através da nova versão do sistema de BI, além de analisar e realizar a consolidação dos dados. Com os dados analisados e consolidados, a última atividade desta fase pode ser realizada, que é a conclusão do projeto e apresentação dos dados e considerações para aos gestores da empresa em estudo e a banca avaliadora deste trabalho.

Nessa fase serão realizados dois tipos de testes de validação, um deles será mais técnico comparando o sistema de BI desenvolvido no trabalho e o modo atual de análise. Essa análise será feita com apoio de métricas relacionadas a consultas ao banco de dados, número de indicadores disponíveis, performance e desempenho do sistema. Já o segundo teste é um *feedback* direto do gestor que acompanha os devidos indicadores, é um teste mais objetivo que serve para verificar a satisfação do usuário. Os produtos finais do testes de validação são as métricas de avaliação e seus resultados, comparando se o sistema de BI apresentou melhora no que foi trabalhado, e demais informações com um viés técnico relacionado à Engenharia de Software; e o outro possui como produto final o sistema de BI para utilização direta pela equipe, além de treinamentos para uso e para manutenção do sistema, conforme ilustrado na figura 6.

Figura 6 – Atividades da fase de Validação da pesquisa-ação.



Fonte: Autor.

- **Disponibilização aos usuários:** Enfim, é disponibilizado o uso da solução para os usuários (gestores). Mas não é somente liberar o acesso à ferramenta e ao sistema de BI, isso também envolve um grande trabalho de capacitação e envolvimento dos usuários, conscientizando-os sobre o uso correto da informação e definição dos níveis de permissão de cada participante, pois nem todos poderão ter acesso irrestrito às informações do BI.

Pelo processo macro, tem-se a execução de passos conforme a figura 7. Os autores Silva, Silva e Gomes (2016) apresentam e recomendam que a solução de BI atenda de forma abrangente toda a corporação, porém, para este trabalho serão realizadas atividades para a obtenção de uma solução de *Business Intelligence* apenas para o setor específico da empresa em que está sendo realizado o estudo.

Figura 7 – Visão geral do desenvolvimento de um sistema de BI.

Fonte: Autor.

3.2 Processo Metodológico

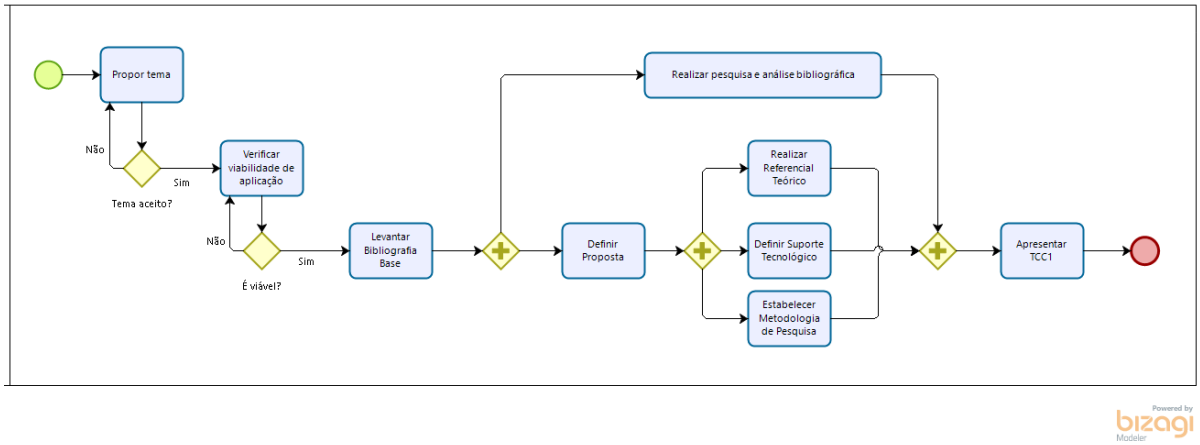
O presente Trabalho de Conclusão de Curso (TCC) corresponde a um projeto realizado em duas fases: uma fase de planejamento e viabilidade do projeto, e outra de elaboração de uma proposta de solução para o problema apresentado inicialmente. Durante a primeira fase, o foco é estabelecer os pilares teóricos para embasamento do projeto total por meio de pesquisas bibliográficas e maior detalhamento da proposta de trabalho, além de sua metodologia. Já na segunda fase, o trabalho possui foco em atingir o objetivo geral de "aprimorar a gestão de armazenagem de produtos no estoque de uma indústria alimentícia, potencializando sua eficácia em ocupação e direcionamento no processo de armazenagem logística", com base nas fundamentações teóricas e na metodologia proposta, ambos realizados e definidos ao longo da fase de planejamento e viabilidade.

3.2.1 Primeira fase

Nesta primeira fase do projeto, pode-se subdividir o processo em oito atividades: selecionar tema, verificar a viabilidade de implantação do projeto em uma empresa externa à universidade, realizar pesquisa bibliográfica, definir proposta, escrever referencial teórico, estabelecer suporte tecnológico, evoluir metodologia e apresentar a proposta à comissão avaliadora deste trabalho.

Com o objetivo de definir e acompanhar o processo de desenvolvimento da primeira etapa do trabalho de conclusão de curso - a pesquisa -, foi modelado um processo metodológico, o qual se encontra na Figura 8.

Figura 8 – Fluxo de atividades da Primeira Fase.



Fonte: Autor.

A seguir será descrita o que cada atividade do processo representa:

- **Propor Tema:** A atividade de propor tema engloba desde a escolha do contexto em que se deseja trabalhar, área de estudo, até a definição do professor orientador do trabalho. Após a escolha do contexto e do orientador, buscou-se definir um escopo que será abordado durante o trabalho. Os orientadores devem validar o tema escolhido para concluir a atividade.

- **Verificar viabilidade de aplicação:** Dado o contexto do trabalho, é necessária a permissão de acesso e a solicitação de colaboração de uma empresa na área do Distrito Federal para a análise e aplicação do projeto. Esta fase consiste em verificar com a empresa a possibilidade de trabalho, além de solicitar as ferramentas utilizadas referentes a BI.

- **Levantar Bibliografia Base:** Esta atividade refere-se à definição de pilares para o estudo proposto, ou seja, estabelecer os principais pontos que serão tratados e validar se são importantes para a obtenção do resultado esperado. Esse levantamento garante o entendimento do contexto trabalhado e as possibilidades de atuação, especificando mais adequadamente o escopo.

- **Definir Proposta:** Documentar a proposta de pesquisa para este trabalho. A proposta inclui, principalmente, uma introdução com a contextualização do tema, os objetivos do trabalho, qual problema foi identificado e deseja resolver, o cronograma macro destas atividades e uma metodologia de pesquisa.

- **Realizar pesquisa e análise bibliográfica:** A pesquisa bibliográfica foi realizada a partir da definição de tema, e consiste na obtenção de conhecimentos mais específicos sobre o que está sendo estudado academicamente e relacionado ao problema exposto, sendo necessária a pesquisa por autores e publicações pertinentes aos tópicos de interesse do tema.

- **Realizar Referencial Teórico:** O referencial teórico tem o completo objetivo de definir quais as fontes conceituais que irão embasar e dar suporte teórico ao trabalho. É apresentado no capítulo 2, em que são explanados os principais conceitos utilizados ao longo deste trabalho. Como insumos para esta atividade, encontram-se todas as pesquisas bibliográficas obtidas durante a atividade *Realizar pesquisa e análise bibliográfica.*

- **Definir Suporte Tecnológico:** Esta atividade realiza o levantamento das principais tecnologias, ferramentas de apoio ao desenvolvimento da aplicação, a gestão de atividades e pesquisa e escrita do trabalho com o objetivo de definir os tipos de suporte necessário de acordo com a disponibilidade e fácil acesso deste e por preferências da equipe de TI da empresa em estudo.

- **Estabelecer Metodologia de Pesquisa:** Esta atividade é responsável por definir toda a metodologia de trabalho para orientar e tornar a pesquisa válida, além de um desenvolvimento mais detalhado e com foco em resultados. A partir da definição da metodologia é possível compreender qual processo pode atender às necessidades.

- **Apresentar Trabalho parcial:** Esta atividade, fim do processo de TCC1, consiste na apresentação de todo o trabalho aos membros da banca examinadora, relacionando ao problema exposto e à proposta de solução apresentada, de forma a obter a validação parcial das atividades e possíveis tratativas metodológicas.

Etapas do TCC - 1ª fase	**CRONOGRAMA - Ano 2019**				
	Meses				
	Março	Abril	Maio	Junho	Julho
Escolha do tema	X				
Justificativa	X				
Hipóteses ou questões norteadoras		X			
Objetivos		X			
Fundamentação teórica	X	X			
Metodologia, pesquisa e dados.		X			
Previsão dos recursos			X		
Encontros com o orientador	X	X	X		X
Revisão Bibliográfica	X	X	X	X	X
Criação dos instrumentos para a coleta de dados					
Coleta e análise dos dados.			X		
Análise ferramental para a solução de BI			X		
Elaboração das Considerações Finais				X	
Redação final da 1ª fase do trabalho				X	
Encaminhamento à correção linguística - orientador				X	
Entrega da 1ª fase				X	
Apresentação à Banca Examinadora				X	
Validação da proposta pela Banca Examinadora					X

Fonte: Autor.

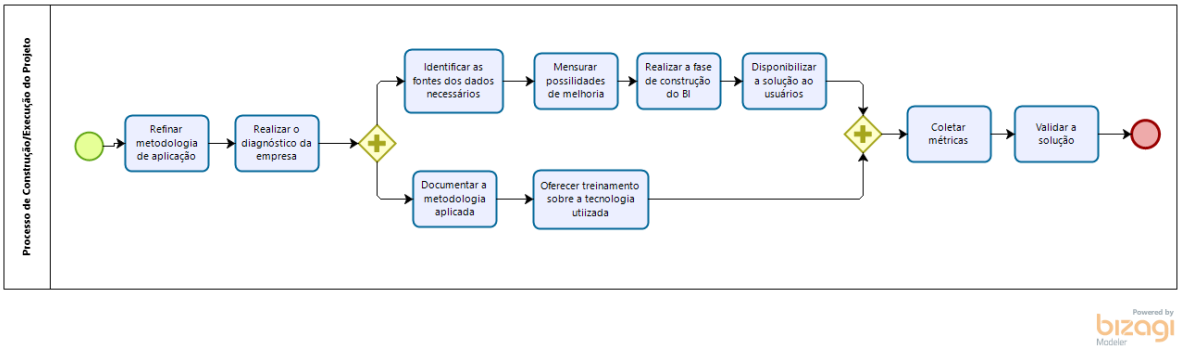
Quadro 3.3 – Cronograma de atividade da primeira fase do trabalho.

3.2.2 Segunda fase

No desenvolvimento prático deste projeto será seguido o fluxo de atividades do procedimento de pesquisa-ação já determinado e especificado neste documento na seção 3.1.1.4.2 e de forma resumida pela figura 7.

Para facilitar o acompanhamento do processo de desenvolvimento da segunda etapa do projeto - desenvolvimento prático -, foi modelado um processo metodológico, o qual se encontra na Figura 9.

Figura 9 – Fluxo de atividades da Segunda Fase.



Fonte: Autor.

A seguir, algumas atividades serão descritas no que representam ao processo. Algumas dessas atividades já foram melhor descritas no capítulo 4 deste documento, pois retratam etapas de implementação do BI.

- **Refinar metodologia de aplicação:** Essa atividade consiste em revisar e atualizar a metodologia proposta e sugerida na primeira fase do trabalho. Essa verificação é realizada após o resultado da banca examinadora, validando a proposta de TCC e oferecendo sugestões sobre o trabalho apresentado. Refinar também é necessário devido ao início da fase de execução, ou seja, a etapas de implementação do BI, em que pode haver mudanças devido às características e solicitações da empresa em estudo.

- **Mensurar possibilidades de melhoria:** Neste momento, após a análise do sistema de BI vigente na empresa e identificação das técnicas de análise de dados utilizadas, será levantada possibilidade de melhorias sistêmicas no aspecto de análise de dados, de forma a oferecer os melhores resultados e visualização de indicadores à alta gestão da empresa.

- **Validar a solução com o cliente/usuário final:** Após o processo de melhoria, é esperado que ocorra a atividade de validação direta com os usuários do sistema, que são colaboradores da empresa contribuinte com este trabalho. Um procedimento subjetivo de análise será utilizado, como um questionário, por exemplo, em que poderão ser listados pontos a favor da solução, contra e pontos a melhorar.

- **Apresentar resultados:** Essa é a apresentação dos resultados finais aos *stakeholders* do projeto, podendo ser apresentadas opções futuras de desenvolvimento.

Etapas do TCC - 2ª fase	**CRONOGRAMA - Ano 2019**				
	Meses - Semanas				
	Agosto	Setembro	Outubro	Novembro	Dezembro
Correção de apontamentos da Banca Examinadora	X				
Submissão da 1ª fase do trabalho corrigido	X	X			
Instalação e reparação de ferramentas de BI	X	X	X	X	
Descrição das fontes de dados	X	X			
Modelagem do BD		X		X	
Tratamento de Dados para incluir na base	X	X			
Processo de ETL		X	X		
Manipulação de Data Marts		X	X		
Criação de OLAP				X	
Criação de Visualização dos indicadores				X	
Escrever sobre o desenvolvimento do trabalho				X	
Elaboração das Conclusão e Trabalhos futuros				X	
Redação final da 2ª fase do trabalho				X	
Encaminhamento à correção linguística - orientador				X	
Entrega do TCC					X
Apresentação à Banca Examinadora					X
Validação da proposta pela Banca Examinadora					X

Fonte: Autor.

Quadro 3.4 – Cronograma de atividades da segunda fase do trabalho.

3.3 Suporte Tecnológico

Nesta seção, serão apresentadas ferramentas e tecnologias utilizadas para auxiliar o desenvolvimento deste projeto, desde a organização e definição da metodologia de pesquisa até o desenvolvimento e construção do BI durante o trabalho.

3.3.1 Desenvolvimento teórico

Com o objetivo de contribuir com o desenvolvimento da escrita, gerenciamento e armazenamento de artigos, foram utilizadas algumas ferramentas para o desenvolvimento do trabalho, entre elas: de edição de texto em plataforma de escrita colaborativa, como o Overleaf Latex, utilizando o LaTeXcomo mecanismo de formatação; para armazenamento de arquivos em nuvens, além da criação, edição e armazenamento de textos, apresentações de slides, planilhas, desenhos e formulários online, permitindo a visualização e o compartilhamento total ou parcial dos arquivos, como o Google Drive; e uma ferramenta para gerenciamento de processos, adotada para a modelagem de fluxogramas, mapas mentais e diagramas em geral, utilizando o Modelo e Notação de Processos de Negócio (BPMN) como o Bizagi Modeler.

Figura 10 – Ferramentas de apoio ao desenvolvimento de escrita.

Fonte: Autor.

3.3.2 Desenvolvimento prático

No desenvolvimento prático deste trabalho, foram utilizadas ferramentas de apoio e de execução. Estas foram escolhidas por levantamento prévio de necessidades e requisitos que precisavam ser atendidos pelo setor de TI, por arquitetura de rede e sistemas, sistemas operacionais utilizados, capacidade de gerenciamento e armazenamento de dados e conhecimento da equipe sobre as ferramentas. Foram utilizadas as seguintes ferramentas:

3.3.2.1 SQL Power Architect

A ferramenta de modelagem de dados SQL Power Architect foi criada por designers de *data warehouse* e possui muitos recursos exclusivos voltados especificamente para o arquiteto de armazém de dados. Ela permite que os usuários façam engenharia reversa de bancos de dados existentes, realizem perfis de dados nos bancos de dados de origem e gerem automaticamente metadados de ETL. Algumas de suas características são:

- Acesso a bancos de dados através de código via JDBC;
- Conecta-se a vários bancos de dados ao mesmo tempo;
- Compara modelos de dados e estruturas de bancos de dados;
- Engenharia reversa para bancos como PostgreSQL, Oracle, MS SQL Server, entre outros;
- Todos os dados do projeto são armazenados em um formato *eXtensible Markup Language* (XML) de fácil análise;
- Existe uma versão livre.

Arquitetos de dados, analistas e designers contam com ferramentas de modelagem de dados para facilitar e simplificar seus esforços de modelagem de dados e, ao mesmo

tempo, maximizar o uso de seus recursos. Com o SQL Power Architect também é possível visualizar informações sobre o tamanho dos dados em cada coluna, os valores máximos e mínimos, a distribuição de frequência dos valores de uma coluna e outras diversas informações.

3.3.2.2 MySQL

O MySQL é um sistema de gerenciamento de banco de dados (SGBD), que utiliza a linguagem SQL (Linguagem de Consulta Estruturada, do inglês Structured Query Language) como interface. É atualmente um dos sistemas de gerenciamento de bancos de dados mais populares da Oracle Corporation, com mais de 10 milhões de instalações pelo mundo.

Para utilizar o MySQL é necessário instalar um servidor e uma aplicação cliente. O servidor é o responsável por armazenar os dados, responder às requisições, controlar a consistência dos dados, bem como a execução de transações concomitantes, entre outras. O cliente se comunica com o servidor através da SQL. A versão gratuita do MySQL é chamada de *Community Edition* e possui o servidor e uma interface gráfica cliente.

O servidor deve ser instalado e configurado para receber conexões dos clientes. No MySQL, o principal cliente é a interface gráfica cliente fornecida pela Oracle, denominada MySQL Workbench. Através do MySQL Workbench, pode-se executar consultas SQL, administrar o sistema e modelar, criar e manter a base de dados através de um ambiente integrado. O MySQL Workbench está disponível para Windows, Linux e Mac OS.

3.3.2.3 HeidiSQL

O HeidiSQL é um software livre e tem como objetivo ser fácil de aprender. O software permite ver e editar dados e estruturas de computadores executando um destes sistemas de banco de dados: MariaDB, MySQL, Microsoft Server ou PostgreSQL.

Ao utilizá-lo com o banco de dados de dados MySQL, o SGBD atua como servidor, enquanto o HeidiSQL como cliente, sendo uma interface de gerenciamento. O conceito de cliente e servidor é definido da seguinte forma por Takai, Italiano e Ferreira (2005): "o servidor é o elemento computacional responsável por prover serviços, por consequência, o cliente é o elemento computacional responsável por requisitar e utilizar tais serviços".

Dessa forma, o MySQL (servidor) e o HeidiSQL (cliente) são elementos computacionais em forma de software livre e fornecem as funcionalidades necessárias para executar as operações computacionais deste estudo, a alocação do banco de dados. Por proporcionarem um ambiente genérico, os procedimentos podem ser operados por outros tipos de SGBD, ou até mesmo por outras linguagens de programação ou programas de computadores. Neste estudo, o SGBD MySQL e o cliente HeidiSQL serão aplicados devido

à simplicidade de operação, resposta em tempo ágil devido ao grande volume de dados utilizados e com foco na integração desta modelagem computacional a outras modelagens computacionais comumente utilizadas no mercado por empresas que não possuem o desenvolvimento de software como atividade fim.

3.3.2.4 Pentaho Data Integration

O Pentaho Data Integration (PDI) fornece os recursos de Extração, Transformação e Carga (ETL) que facilitam o processo de captura, limpeza e armazenamento de dados usando um formato uniforme e consistente, acessível e relevante para os usuários finais e as tecnologias de IoT. Trata-se da ferramenta mais popular e madura da suíte Pentaho, com mais de 15 anos de existência.

Os usos comuns do Pentaho Data Integration incluem:

- Migração de dados entre diferentes bancos de dados e aplicativos;
- Carregar enormes conjuntos de dados em bancos de dados, aproveitando ao máximo os ambientes de processamento em nuvem, cluster e paralelamente;
- Limpeza de dados com etapas que variam de transformações muito simples a muito complexas, integração de dados, incluindo a capacidade de alavancar ETL em tempo real como fonte de dados para o Pentaho Reporting;
- Popular um *data warehouse* com suporte embutido para alterar lentamente as dimensões e criar substitutos de chave.

O PDI, também conhecido como Kettle, dispõe de poderosas ferramentas de ETL, integradas em um ambiente gráfico rico e atrativo, com ampla conectividade a qualquer tipo de dados e que proporciona escalabilidade. É composto por quatro ferramentas, nomeadamente a Spoon, Pan, Chef e Kitchen. O PDI permite tratar os dados para posterior análise (Analysis), visualização (Dashboards) e representação (Relatórios).

Neste projeto será utilizada a ferramenta Spoon, sendo esta que permite desenhar de forma gráfica a transformação ETL. O PDI Client (Spoon) é um aplicativo de desktop que pode ser instalado na estação de trabalho, permitindo criar transformações e tarefas ou agendar quando as tarefas devem ser executadas.

3.3.2.5 Pentaho Server

O Pentaho BI Server é responsável por toda a camada servidora da plataforma do Pentaho, fornece a arquitetura e infraestrutura necessária para a construção de soluções para business intelligence (BI).

3.3.2.6 Pentaho Workbench

O Mondrian Schema Workbench é uma interface de design que permite criar e testar os esquemas de cubos do Mondrian OLAP visualmente. O mecanismo Mondrian processa solicitações MDX com os esquemas ROLAP (Relational OLAP). Esses arquivos de esquema são modelos de metadados XML criados em uma estrutura específica usada pelo mecanismo Mondrian. Os modelos XML podem ser considerados estruturas do tipo cubo que utilizam as tabelas FACT e DIMENSION existentes encontradas em seu RDBMS e não requerem que um cubo físico real seja construído ou mantido, somente que o modelo de metadados seja criado.

Com um modelo físico de dados multidimensionais, pode ser criado um modelo lógico que o mapeie. Um esquema Mondrian é, essencialmente, um arquivo XML que executa esse mapeamento, definindo assim uma estrutura de banco de dados multidimensional.

Em um cenário muito básico, um esquema Mondrian poderá ser feito como um cubo que consiste em uma única tabela de fatos e em algumas dimensões, cada uma com uma única hierarquia que consiste em vários níveis. Esquemas mais complexos podem envolver vários cubos virtuais e, em vez de mapear diretamente para a única tabela de fatos no centro de um esquema em estrela, eles podem mapear para visualizações ou tabelas embutidas. Os esquemas em Mondrian são definidos como documentos XML e o Pentaho Schema Workbench (PSW) oferece uma interface gráfica de usuário para editar esses esquemas. É possível editar os esquemas manualmente usando um editor XML, IDE ou mesmo um editor de texto sem formatação.

3.3.2.7 Microsoft Power BI

O Power BI é um serviço de *Business Analytics* baseado em nuvem que oferece uma visão única dos dados comerciais mais críticos. É possível monitorar a saúde do negócio da empresa usando um dashboard interativo e em tempo real, criar relatórios interativos com o Power BI Desktop e acessar os dados com aplicativos nativos do Power BI Mobile. É uma aplicação gratuita e adotada pela empresa como sua ferramenta de BI.

Algumas vantagens do Power BI relativas as outras ferramentas de visualização são: capacidade de carregar visualizações personalizadas através do AppSource, acesso à informação em tempo real, acesso de informação em qualquer aplicativo móvel, uso empresarial, atualização mensal com novas funções e facilidade de uso.

O Power BI pode ser usado de forma simples para pequenos trabalhos no Excel e banco de dados local, ou até mesmo para grandes e robustos trabalhos a nível empresarial.Com esta ferramenta é possível gerar uma visão de 360 graus com suas métricas mais importantes em um só lugar, atualizadas em tempo real.

O Power BI é uma solução de análise de negócios que permite a visualização de dados e o compartilhamento de *insights* em toda a organização, conectando centenas de fontes de dados e possibilitando dar vida aos dados corporativos com dashboards e relatórios. O software possui uma capacidade de compactação impressionante para arquivos do Microsoft Excel e extensões text/csv. Isso permite editar, analisar e visualizar enormes quantidades de dados que não podem ser abertas no Excel. Logo, grandes conjuntos de dados não precisam estar agregados e reduzidos em tamanho para mostrar análises mais abrangentes.

Dentro do contexto do negócio, a ferramenta permite realizar as seguintes funções:

- Possibilidade de transformações e edições, incluindo a alteração de formatos, exclusão de linhas, adição de colunas e transposição de dados;
- Adicionar relações entre tabelas diferentes, o que é útil se você tiver um esquema estrela ou floco de neve.
- Adicionar novos conjuntos de dados facilmente à medida que dados úteis são descobertos. Novos dados podem ser integrados no Power BI sem a necessidade de recombinar. Isso proporciona uma enorme flexibilidade na vinculação de conjuntos de dados.
- Com o Power BI, a repetição dolorosa das etapas de preparação de dados é reduzida ou totalmente eliminada, pois o software se lembrará dessas etapas e as repetirá automaticamente durante as atualizações.

Segundo a própria Microsoft, o Power BI é uma solução voltada para a análise de negócios possibilitando realizar a análise de informações e o compartilhamento de ideias. Os painéis do Power BI fornecem uma visão ampla e integrada para os usuários corporativos, de modo que eles possam consultar as suas métricas mais importantes. As informações podem ser atualizadas em tempo real e podem ficar disponíveis em diversos dispositivos.

Em questões ferramentais, são muitas as combinações de sistemas para chegar a entrega de um BI ao cliente, portanto, nesta implementação serão utilizados os softwares mencionados como padronizados pela empresa em estudo e seu uso é condicionado a uma sequência lógica de implementação, como apresentado na figura 11.

Figura 11 – Ferramentas de apoio ao desenvolvimento de BI.

Fonte: Autor

4 Desenvolvimento

No decorrer da análise da solução, foi realizada uma exploração aprofundada e detalhada do estudo de caso, com a proposição de uma arquitetura de software, que se apoia no processo de identificações de eventos anormais, com possíveis indicações de otimização de espaço de estocagem e aperfeiçoamentos logísticos.

4.1 Entendimento de Negócio

O Entendimento de Negócio estabelece uma visão geral sobre as funções, processos e estrutura de logística em armazenagem de produtos alimentícios prontos para venda, focando no objeto estudado e subsidiando o entendimento das atividades e estrutura do modelo de compra e produção de produtos.

As atividades envolvidas neste trabalho dizem respeito à disciplina de gestão de estoques, que podem ser decompostas em quatro aspectos básicos. Sendo eles: as políticas e modelos quantitativos utilizados, as questões organizacionais envolvidas, o tipo de tecnologia utilizada e, finalmente, o monitoramento do desempenho do processo.

Um sistema de monitoramento de desempenho de estoque adequado desempenha dois papéis de extrema relevância para o gestor de estoque: informa o desempenho do processo de gestão, considerando todos os impactos causados pelo mesmo, e indica razões deste desempenho.

Para a elaboração deste trabalho, foram acompanhados os processos de estocagem, expedição e gestão de estoque de uma empresa alimentícia de médio porte na cidade de Brasília/DF e, a partir das observações, anotações, documentos e queixas de funcionários, foram analisados alguns processos de logística da empresa.

4.2 Entendendo os dados

O entendimento dos dados é uma etapa importante para moldar as possíveis formas de monitoramento do objeto desejado, os formatos dos atributos, os dados faltantes, sua consistência, etc. Baseado nesta ideia, foram realizados testes nos dados que foram fornecidos pela empresa por meio de solicitação direta.

Os dados foram submetidos a:

1. Análise superficial para identificações dos tipos de variáveis (atributos);
2. Teste de dados faltantes: verificação dos dados ausentes nos campos;

3. Descrição inicial dos dados;
4. Verificação da qualidade dos dados.

Após os testes iniciais, foi verificado que os dados considerados essenciais para este trabalho possuem preenchimento falho em alguns períodos de implementação do projeto, pois, inicialmente, o preenchimento de dados importantes possuía escrita livre, diminuindo a padronização de alguns campos. Porém, os dados prioritários e não nulos das tabelas foram submetidos ao preenchimento automático pelo sistema a partir de um código identificador. Quanto à qualidade dos dados, observou-se que ela era adequada, pois os atributos são preenchidos automaticamente com a utilização da tecnologia de código de barras e 70% (setenta por cento) dos itens são pré-definidos por padrões do sistema.

4.3 Pré-processamento

A etapa do pré-processamento é considerada a base para as demais etapas, pois, se essa não for bem definida, poderá impactar negativamente nos processos e, ainda, não obter o conhecimento esperado.

Aa ações realizadas na fase de pré-processamento foram:

1. Utilizar a etiqueta de paletes para subsidiar a análise das entradas de dados do sistema de gerenciamento e controle do estoque (WMS);
2. Verificar a quantidade de dados armazenados no banco de dados do WMS;
3. Renomear os atributos de dados;
4. Efetuar operações preliminares de limpeza de informações.

A etapa de pré-processamento foi detalhada na Seção 4.4.3 (sobre ETL) deste trabalho.

4.3.1 Entendendo as variáveis do Banco de Dados

O banco de dados disponibilizado para a pesquisa possui inúmeras variáveis, com um conjunto de atributos que envolvem a localização do endereço dentro do estoque de acabados, segmentação de endereços, códigos de controle de produtos, códigos de nota fiscal para recebimento de terceiros, problemas de monitoramento de paletes, tempo de vida de produtos, classificação em produtos com data crítica, entre outros.

De modo particular, e por se tratar de uma análise de pré-processamento, destacaram-se como variáveis essenciais os atributos diretamente relacionados à operação de movi-

mentação de produtos no estoque e, nestes, foi necessário realizar a transformação de atributos em chaves de identificação e alguns tratamentos de dados.

Entretanto, considerando os objetivos de negócios e da pesquisa e, ainda, a análise preliminar dos dados, pode-se afirmar que pelo menos quatro variáveis são essenciais para o estudo e outras serão geradas a partir dos dados existentes.

As variáveis elencadas (importantes) para compor o estudo foram:

1. Chave (identificador de cada palete - etiqueta);
2. Código do produto;
3. Descrição do produto;
4. Categoria do produto;
5. Segmento do produto;
6. Marca do produto;
7. Data de validade do produto;
8. Origem do produto (fabricado ou comprado);
9. Endereço no estoque para o qual o palete foi movimentado;
10. Fluxo da movimentação (entrada ou saída do produto);
11. Operador que realizou a movimentação do palete (matrícula e nome do empilhador);
12. Data da movimentação;
13. Hora da movimentação.

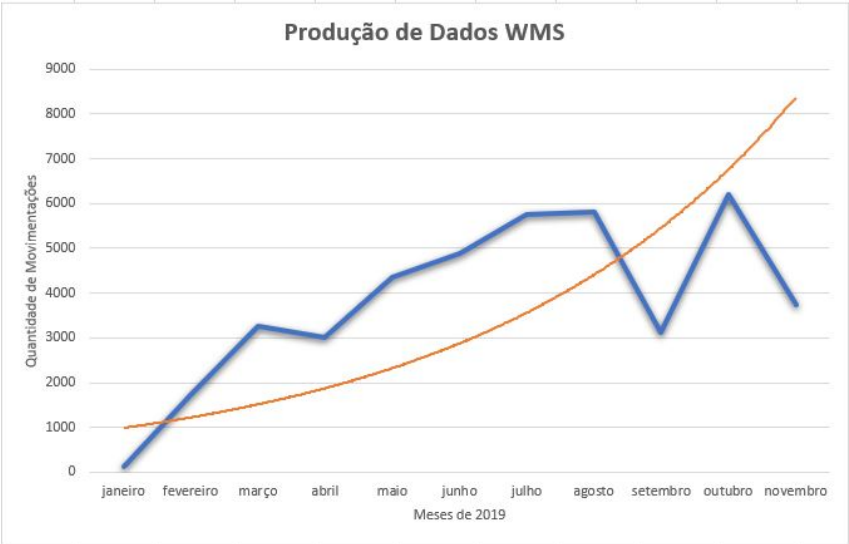
4.3.2 Análise dos Dados

A análise dos dados visa estabelecer uma visão geral das informações em que se pretende trabalhar. Neste trabalho foram utilizados dados históricos de movimentação de paletes dentro do estoque de acabados de uma empresa alimentícia, sendo estes produzidos pelo sistema interno de WMS (*Warehouse Management System*, ou Sistema de Gerenciamento de Estoque) que adota as políticas de estoque definidas pela própria empresa.

A Figura 12 apresenta a evolução de armazenamento de dados ao longo dos meses do ano de 2019 dado ao processo de implantação do sistema de WMS na área de estoque de acabados da empresa em estudo. A utilização do sistema está sendo ampliada na empresa, onde mais pessoas são responsáveis por inserir informações, por meio de leitura de códigos

de barras de todos os paletes armazenados, para que possam ser rastreados pela equipe de gerência de acordo com a necessidade.

Figura 12 – Produção de dados pelo sistema WMS em 2019.

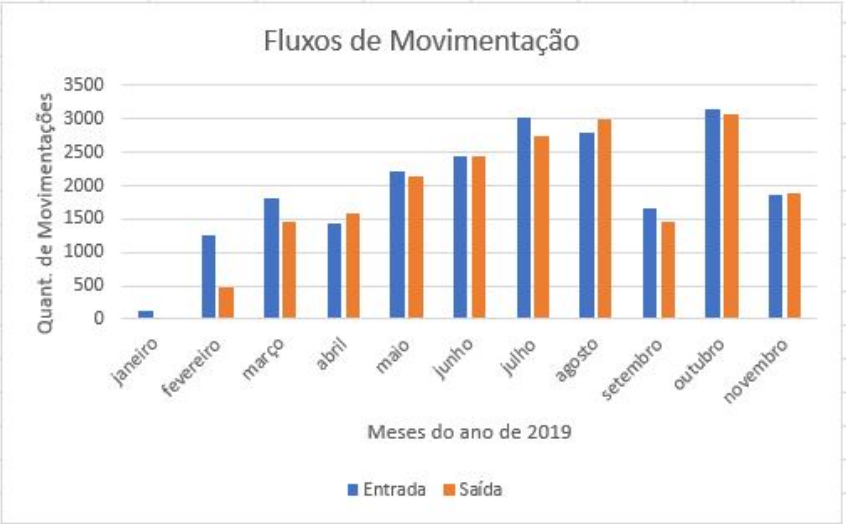


Fonte: Autor.

Pela figura 12, é notável o aumento em quantidade de dados oferecidos para análise a partir do sistema WMS implementado. Através da linha de tendência exponencial do gráfico espera-se que a quantidade de movimentações registradas aumente em decorrência da implantação do projeto em todo o estoque, mas que deve chegar a um limite visto que, quanto menor o número de movimentações de um mesmo palete, mais aperfeiçoada está a gestão de estoque.

A Figura 13 apresenta a quantidade de movimentações realizadas de janeiro a novembro de 2019, distinguidas por tipo de fluxo (entrada ou saída de produtos), podendo-se observar que há variâncias de meses com maior entrada de produtos do que saídas. Essa realidade gera custos de estocagem e, em apenas dois meses, as movimentações de saída superaram as de entrada, ou seja, houve mais venda do que compra ou produção de mercadorias.

Figura 13 – Fluxo de Movimentações realizadas em 2019.

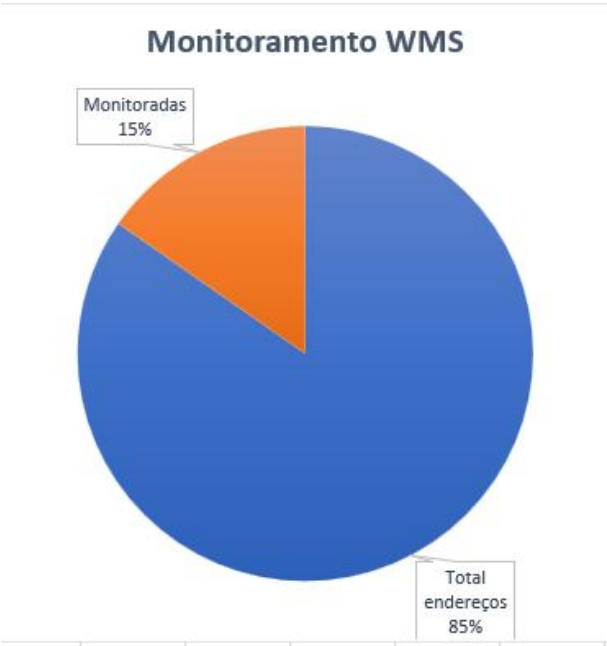


Fonte: Autor.

Na Figura 14, observa-se que o sistema WMS da empresa em estudo ainda está em evolução e não monitora toda a área de estoque de acabados pertencentes à empresa. Desta forma, há a possibilidade de aumento exponencial de quantidade de dados produzidos ao decorrer da expansão do projeto, que hoje monitora apenas 15% (quinze por cento) do espaço de armazenamento da empresa.

É necessário considerar que a área monitorada atualmente pelo WMS é composta por *racks e Drive-in* com grande capacidade de armazenamento, podendo acomodar até 36 paletes de produtos em 6 andares de estrutura. Como a empresa está realizando um projeto de expansão e verticalização do estoque, o sistema WMS pode ser muito útil no mapeamento de produtos e na localização de paletes.

Figura 14 – Área de estoque monitorada pelo sistema WMS até o momento.



Fonte: Autor.

4.4 Solução de *Business Intelligence*

4.4.1 Fontes externas

Na empresa em estudo não foi permitido o acesso direto ao banco de dados transacional, ou seja, o banco de dados de produção, que armazena transação por transação realizada, devido à política de segurança de dados da instituição. Todavia, alguns dados puderam ser extraídos de aplicações utilizadas pelo setor do estoque relacionado à armazenagem de acabados e planilhas utilizadas pelo setor para gestão de estoque, sendo estes capazes de geração de um *Data Mart*.

Todos os dados utilizados na construção deste processo de *Business Intelligence*(BI) foram obtidos a partir de arquivos de formato Excel (XLSX), utilizados para validação, formatação e tratamento de alguns dados, e convertidos para arquivos de formato CSV (*Comma Separated Values*, ou Valores Separados por Vírgula), para estes últimos serem utilizados no processo de ETL.

As planilhas em Excel foram armazenadas em diretório local dentro de um determinado período de execução do trabalho, desta forma, o produto final da implementação não garante a atualização dos dados, pois não está vinculado às planilhas originais utilizadas pela equipe de gestão de estoque.

4.4.2 Modelagem de BI

Os dados oriundos dos mais diversos sistemas de informação e de planilhas da empresa em estudo passam a ser armazenados em bancos de dados virtuais após a fase de extração e tratamento destes dados, utilizando as estrutura de Tabelas de Dimensão e Tabelas Fato pelo modelo dimensional.

As Tabelas de Dimensão possuem características descritivas dentro do *Data Warehouse* (DW), qualificando as informações provenientes da Tabela Fato. Através dela, é possível analisar os dados sob múltiplas perspectivas. A Tabela Fato possui característica quantitativa dentro do DW. A partir dela são extraídas as métricas que são cruzadas com os dados das dimensões, concebendo, assim, informações significativas para a análise do usuário. A Tabela Fato armazena as medições necessárias para avaliar o assunto pretendido. O conteúdo histórico no DW, contendo longo período de tempo, ficam depositadas nessa Fato.

O quadro 4.5 apresenta um documento chamado Matriz Dimensão versus Indicador que resultará na seguinte análise: nas linhas da matriz, se escrevem os indicadores que se deseja analisar, como ocupação do estoque, giro e ruptura de estoque. Nas colunas se colocam as dimensões, como a informação da forma que se quer ser vista pelo cliente, por exemplo, paletes, endereços e dados do empilhador.

Matriz Dimensão X Indicador			
	PALETE	ENDEREÇO	OPERADOR
Ocupação do Estoque por áreas	X	X	
Ocupação do Estoque por ruas	X	X	
Efeito Colméia por ruas	X	X	
Giro de Estoque	X	X	
Ruptura de Estoque	X		
Movimentação de Produtos dentro do Estoque	X	X	X

Fonte: Autor.

Quadro 4.5 – Matriz Dimensão X Indicador.

A partir da análise da Matriz Dimensão X Indicador, observa-se a necessidade da criação de três Tabelas Fatos que trabalharão com os dados necessários para a geração dos indicadores, sendo uma Fato relacionando os paletes e endereços, a segunda apenas com dados sobre paletes e a última realizando a ligação entre os dados de paletes, endereços e empilhadores.

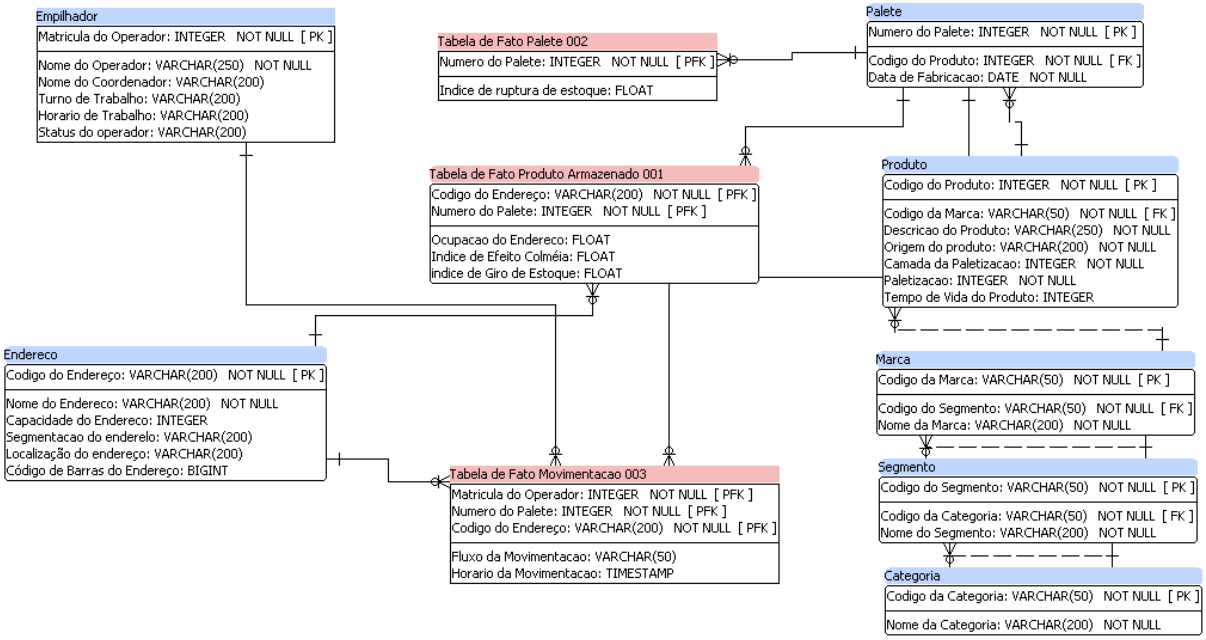
A estrutura dimensional, normalmente, é desenhada no formato do esquema estrela (*Star Schema*), em que as tabelas de dimensões são ligadas diretamente à Tabela Fato, como contextualizado na seção 2.1.3.2.1. Outra característica marcante é que os dados

são desnormalizados, pois a redundância resultante gera benefícios para a otimização das consultas e navegação das informações.

Já o esquema conhecido como Floco de Neve (*Snowflake*) corresponde a uma extensão do esquema em estrela e acrescenta dimensões adicionais. Ele é chamado de floco de neve devido ao formato de seu diagrama, onde as Tabelas de Dimensão são normalizadas, dividindo os dados em tabelas adicionais. Mais detalhes sobre essa modelagem são mostrados na seção 2.1.3.2.2.

A modelagem apresentada para esta solução de Business Intelligence é híbrida, pois utiliza-se do modelo *Star Schema*(Estrela) e do design *Snowflake* (Floco de Neve). Em uma modelagem em Estrela temos uma tabela central, então chamada de tabela Fato, que será responsável pela ligação entre as dimensões. No modelo Estrela, as dimensões se ligam unicamente a Fato, como por exemplo a tabela de fato de movimentação apresentada na figura 15.

Figura 15 – Modelagem do BI.



Fonte: Autor.

Para a modelagem da figura 15, as dimensões relacionadas ao palete e produto seguem o modelo Floco de Neve, responsável por relacionar Categoria, Segmento, Marca e Produto ao Palete que tem associação direta com a movimentação.

No âmbito deste trabalho, a modelagem híbrida foi escolhida pelo autor para fins acadêmicos, a fim de observar o comportamento, a implementação e os benefícios de ambas as estruturas dimensionais e, também, por questão de desempenho e manuseio de dados. Os dados utilizados gerariam grandes repetições de atributos em apenas uma tabela, pois

informações como categoria, segmento e marca repetiriam em centenas de tuplas, com isso, a dimensão relacionada aos paletes e produtos foi normalizada mantendo o relacionamento entre eles.

A partir da modelagem apresentada, foi utilizado o software *Power Architect* para auxiliar na construção do *script* de criação do banco de dados MySQL, disponível no apêndice A deste trabalho.

4.4.3 ETL e Data Mart

Foram realizados estudos sobre os processos de Extração, Tratamento e Carga (ETL - *Extract, Transform, Load*) dos dados, buscando a solução e que melhor atendesse aos requisitos do negócio. Além dos estudos, foram realizadas reuniões técnicas com os especialistas em banco de dados e em tecnologia BI da empresa em estudo, visando a obtenção de informações relevantes a respeito dos dados e o processo de inteligência de negócio adotado na empresa.

Os processos ETL dos dados foram divididos em etapas chamadas de *Transformations e Jobs*, utilizando o software *Pentaho Data Integration* (PDI). Esta ferramenta foi escolhida por atender aos requisitos do negócio e o processo vigente de BI, além de ser de uso livre e sem custo.

Nesta fase, utilizou-se a ferramenta PDI para os processos de extração de dados de arquivos CSV e para o armazenamento na base de dados MySQL, e o software HeidiSQL para estabelecer e monitorar a conexão com o banco de dados.

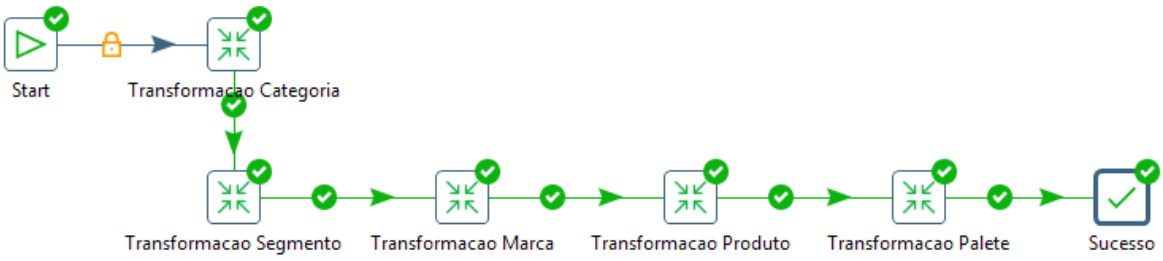
A figura 16 representa o processo de extração de dados de um arquivo CSV e o tratamento destes dados para armazenar de forma apropriada na base e inserir dados nas tabelas criadas em banco de dados MySQL.

Figura 16 – Exemplo do *workflow* do processo ETL que popula a tabela dimensão.

Fonte: Autor.

A figura 17 representa a execução de um JOB (procedimento). Um JOB pode ser utilizado para executar uma ou mais transformações, recuperar arquivos de um servidor Web, colocar arquivos em um diretório de destino, entre outras atividades. Além disso, ele pode ser agendado para realizar execuções em datas e horários especificados.

Figura 17 – Exemplo de JOB que executa as transformações em sequência.



Fonte: Autor.

No exemplo da figura 17, existe a criação de um JOB para definir o processo de inserir informações nas tabelas relacionadas a produtos e paletes, seguindo a ordem ideal de inserção de valores. A inserção de um Produto necessita do código identificador de uma Marca, na qual necessita do identificador de Segmento que utiliza dados de Categoria. Assim, pode-se perceber que é necessário popular o banco de dados seguindo uma ordem específica entre as tabelas para não gerar inconsistências de dados.

A tabela 1 apresenta a quantidade de tuplas utilizadas nesse trabalho para avaliação simples de indicadores de gestão de estoque. Estes dados correspondem aos presentes no *Data Mart* criado.

Nome da tabela	Registros	Tamanho	Comentário
Dim categoria	2	16 KiB	Dimensão
Dim endereco	1961	336 KiB	Dimensão
Dim marca	20	32 KiB	Dimensão
Dim operador	51	16 KiB	Dimensão
Dim palete	124.624	8,5 MiB	Dimensão
Dim produto	1.445	144 KiB	Dimensão
Dim segmento	10	32 KiB	Dimensão
Fato armazenamento	333	64 KiB	Fato
Fato produtos	124.624	8,5 MiB	Fato
Fato movimentacao	18.499	5,5 MiB	Fato

Tabela 1 – Quantidade de tuplas de dados no Data Mart.

4.4.4 OLAP

O processamento analítico em tempo real ou *Online Analytical Processing* (OLAP) se refere a uma variedade de atividades normalmente executadas por usuários finais em sistemas de tempo real. Esse processamento inclui atividades como geração e resposta de consultas, solicitação e execução de relatórios e gráficos, realização de análises estatísticas tradicionais ou modernas e construção de apresentações visuais.

Basicamente, os produtos de OLAP oferecem recursos de modelagem, análise e visualização de grandes conjuntos de dados, ou para Sistemas de Gerenciamento de Banco de Dados (SGBD) ou, mais frequentemente, para sistemas de DW. Os sistemas para processamento analítico *on-line* são atividades que analisam interativamente os dados das transações de negócios armazenadas nos modelos dimensionais dos DWs para auxiliar na tomada de decisões estratégicas e táticas.

Para este projeto, optou-se em utilizar a ferramenta *Pentaho Workbench (Mondrian)* para construção do OLAP. *Mondrian* é um mecanismo OLAP escrito em Java que executa consultas escritas na linguagem MDX (*Multidimensional Expressions*), lendo dados de um banco de dados relacional (SGBD) e apresentando os resultados em um formato multidimensional por meio de uma API Java.

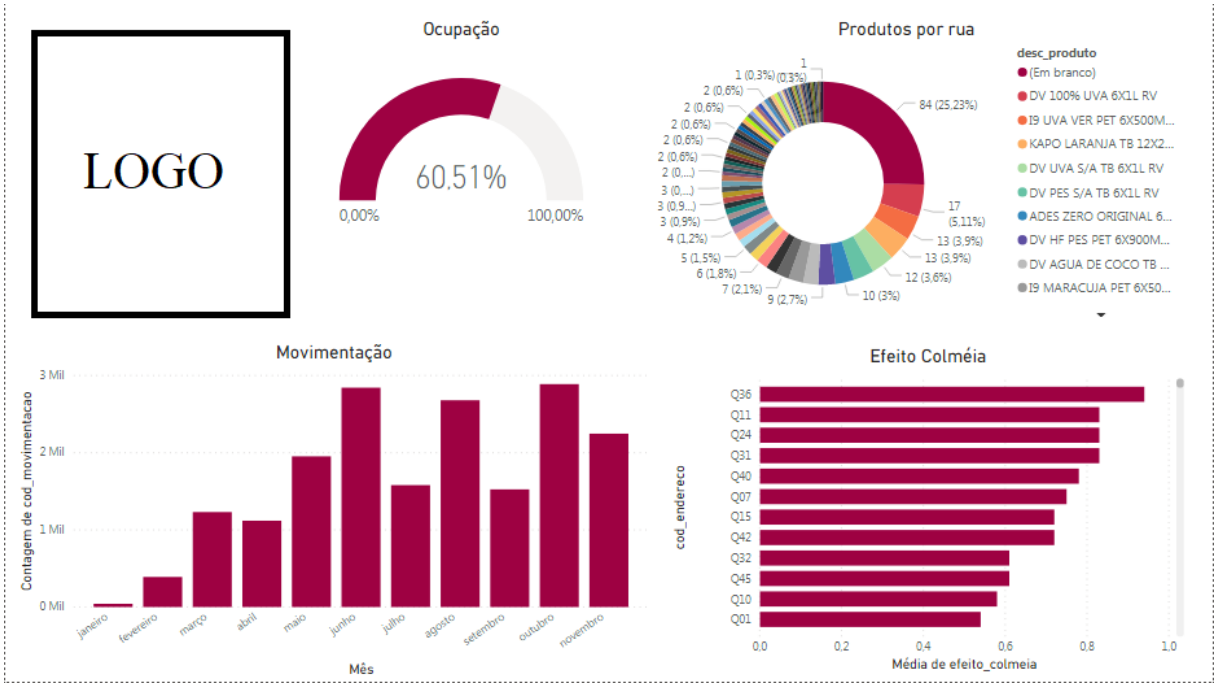
No modelo multidimensional, deixa-se de focar na coleta de dados para ocupar-se com a consulta dos dados. Um modelo multidimensional é uma técnica utilizada em *Data Warehouses* para agrupamento de informações específicas e preparo para análise da informação em mais de uma dimensão. Esse modelo define as Tabelas de Dimensão que possuem o conjunto de informações a serem trabalhadas no BI e a Tabela Fato que possui as chaves das dimensões e as métricas para a análise dimensional.

4.4.5 Visualização

Cada visualização das informações pelo usuário final foi melhor elaborada utilizando o software Microsoft Power BI, isso porque essa ferramenta já era utilizada na empresa com este objetivo, atendendo às necessidades de visualização do sistema de BI.

A partir do uso da ferramenta Power BI, é possível criar *dashboards* (painel de controle com dados) que apresentam as informações mais importantes ou indicadores da área que são consolidadas e organizadas em uma única apresentação para serem analisadas e exploradas facilmente. O objetivo deste tipo de painel de controle é auxiliar na tomada de decisões proporcionando agilidade de pesquisa e menor possibilidade de erros.

Figura 18 – Painel apresentando indicadores de gestão de estoque.



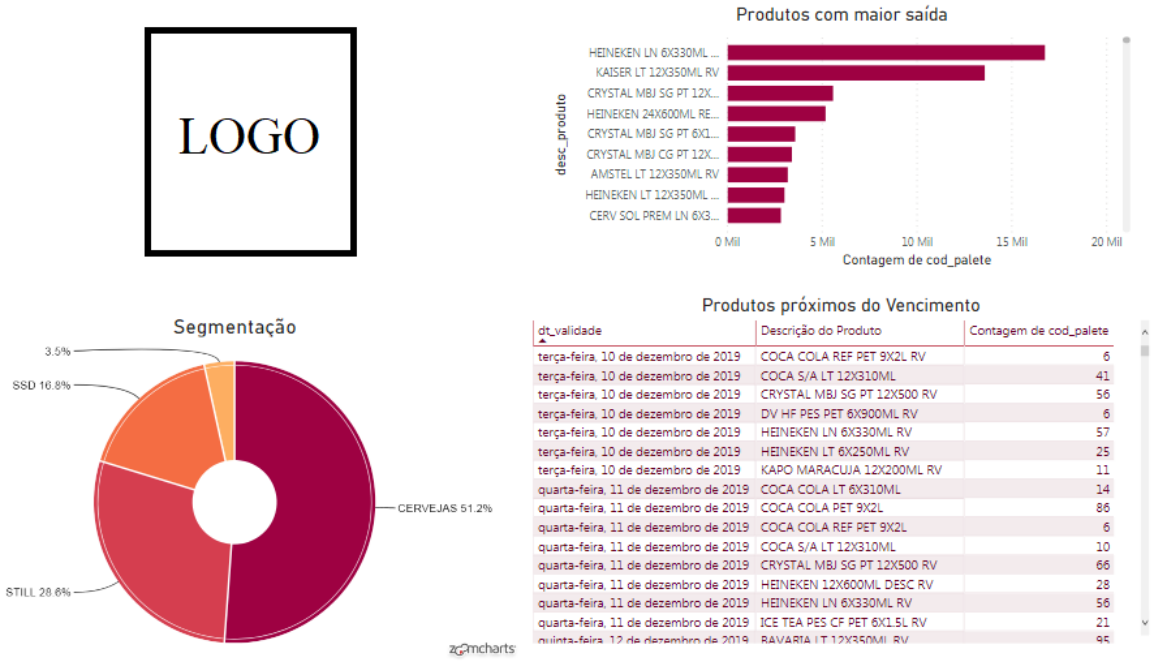
Fonte: Autor.

* Os dados apresentados são fictícios.

Na figura 18, é apresentado um exemplo de painel com indicadores de gestão de estoque, sendo eles: a ocupação média do estoque (na área de cobertura), a quantidade de movimentações de produtos dentro do estoque (entrada ou saída) por meses do ano, as ruas/endereços que apresentam mais efeito colméia e a quantidade de endereços por produtos armazenados, apresentando quantos endereços estão sem produtos, qual produto ocupa maior quantidade de ruas, etc.

Através deste painel, as tomadas de decisões relacionadas à movimentação de paletes dentro do estoque (esforço), endereços que podem receber uma tratativa para melhor armazenagem e a segmentação do estoque por produtos em tempo real são possíveis ocorrências que podem ser identificadas ou até solucionadas com o auxílio da visualização de dados do estoque.

Figura 19 – Painel com indicadores sobre produtos em estoque.



Fonte: Autor.

* Os dados apresentados são fictícios.

A figura 19 apresenta outro painel elaborado no software Power BI mostrando dados referentes à segmentação dos endereços de estoque, podendo visualizar por categoria, segmento ou marca; uma lista com os produtos armazenados em estoque e que estão próximos da data de vencimento e sua respectiva quantidade de paletes, além de mostrar os produtos que possuem maior movimentação de saída do estoque dentro da área coberta pelo projeto WMS.

As etapas previstas do processo de validação deste trabalho não foram consolidadas formalmente, não sendo capaz de instituir métricas de comparação entre os métodos de análises de gestão de estoque anteriores e posteriores à implementação de BI na área. A impossibilidade de validar a solução com o usuário final desta aplicação gerou o impedimento de certificar formalmente e explicitamente a melhoria percebida neste processo por viés subjetivo. Contudo, em meios informais de escrita (conversas e entrevista não documentada), obteve-se a aprovação da equipe de gerência observado as possibilidades de melhoria.

5 Conclusão

O objetivo principal deste trabalho, conforme descrito na sessão 1.4.1 do capítulo 1, é aprimorar a gestão de armazenagem de produtos no estoque de uma indústria alimentícia, potencializando sua eficácia em ocupação e direcionamento no processo de armazenagem logística. Para isso, foi efetuada uma pesquisa bibliográfica envolvendo as etapas para realização da análise de dados utilizando métodos de inteligência de negócios, além do estudo de formas de estruturação. Considerando o estudo relacionado ao planejamento de estoque e a busca de soluções maximizando o potencial ocupacional, este projeto investigou métodos de avaliação dos processos e de análise de dados que auxiliariam na identificação de oportunidades de melhoria, apresentando uma prática de investigação por meio de dados históricos sobre o processo de armazenagem, compra de insumos e venda de produtos acabados.

Nesta perspectiva, considerando as limitações envolvidas, propôs-se uma forma de trabalho baseada em trabalhos anteriores relacionados ao tema, com o intuito de aprimorar o modelo de sucesso para sistemas de análise de dados utilizando *Business Intelligence* no contexto de armazenagem logística.

Ao longo da primeira fase do projeto, o principal objetivo do trabalho foi levantar técnicas e estratégias existentes para resolução do problema e apresentar dados relevantes à área de estoque para tomada de decisão, assim como identificar conceitos e informações referentes à gestão de estoque, levando-se em consideração os objetivos específicos.

Desta forma, foi possível identificar uma solução para aprimorar o sistema de armazenagem em um contexto simplificado, a partir da análise dos resultados obtidos com a pesquisa-ação para a melhoria do processo. A etapa de verificação com o usuário acerca do potencial do BI não foi formalizada, o que comprometeu na validação de uma melhoria conclusiva no projeto. Porém, de forma informal, os painéis foram apresentados e obteve-se o retorno positivo a respeito das informações.

5.1 Lições aprendidas pelo pesquisador

Durante o período de janeiro a junho de 2019, o pesquisador esteve presente na empresa utilizada como objeto de estudo na função de estagiário no setor de gestão de estoque, denominada inteligência de armazém. Nesta experiência, foi possível acompanhar e realizar algumas operações de monitoramento e controle de estoque junto a equipe, realizando as atividades dos processos por diferentes cenários, identificando gargalos e oportunidades de melhoria.

A dificuldade inicial encontrada era o estorvo de grande parte da equipe em explicar conceitos cotidianos na área logística, como a definição de ruas, porta-paletes, drive-in, efeito colmeia, entre outros. Apesar desses obstáculos, pelo tempo de estágio realizado no setor, foi possível compreender estes conceitos e observar sua execução.

Posteriormente, no ambiente de estoque, foi observado os cuidados relacionados à saúde e segurança ocupacional, meio ambiente e segurança de alimentos, sendo estes os pilares da empresa, denominadas "regras de ouro". Todas as regras estipuladas eram verificadas frequentemente para atender as normas internacionais e, por ser um estoque de produtos alimentícios, esse fatores são de grande relevância ao gerir os espaços.

Para a execução deste trabalho, em termos tecnológicos, o pesquisador encontrou algumas barreiras em relação ao acesso aos dados de produtos no estoque, causadas por políticas internas e de segurança da informação. Contudo, por haver acompanhamento e pela anuência da equipe de inteligência de armazém, alguns dados foram adquiridos para análise prévia, sendo esta a principal dificuldade para o andamento do trabalho.

Em termos de lições aprendidas, é notória a discrepância de possibilidades de trabalho de melhoria em ambientes acadêmicos e corporativos. Posto isto, em ambientes empresariais, o novo pode ser um incômodo e mudar processos pode ser desconfortante, sendo a principal dificuldade a alteração de funções ou atividades de um colaborador.

Em meios de pesquisas acadêmicas, o objetivo é desenvolver melhorias, otimizar processos, automatizar atividades rotineiras, entre outros, e os projetos necessitam ser curtos e com respostas rápidas. Já no meio corporativo, pelo estudo de caso, espera-se muito pelo investimento em melhorias, sendo essas, geralmente, as últimas na lista de priorização.

Para concluir, a pesquisa-ação obteve o resultado satisfatório assemelhando a técnica de observação em levantamento de requisitos, permitindo visualizar como o processo é realizado e qual sua finalidade real.

5.2 Trabalhos futuros

Com a evolução deste modelo de análise do estoque, é observado que a inclusão de mais informações referentes aos produtos ampliaria as possibilidades de análises gerenciais, como custo e preço de venda, podendo ainda serem adicionados dados sobre as empilhadeiras, a fim de averiguar o consumo de combustível.

Neste trabalho, adotou-se o uso do conceito da inteligência de negócio para a Gestão de Estoque, focando em processos e etapas necessárias para a aplicação do *Business Intelligence*. Uma outra opção interessante, como proposta futura e de melhoria deste trabalho, seria o uso da inteligência computacional, por meio de algoritmos de monitora-

mento da ocupação do estoque em tempo real. Desta forma, são sugeridos como trabalhos futuros:

- A geração de mais *dashboards* a partir do estudo de novos indicadores;
- A criação de tarefas (*jobs*) agendadas para a execução dinâmica do *Pentaho*, conectando o processo de ETL direto da base de dados transacional da empresa;
- A realização do processo de auditoria de dados para permitir o monitoramento de sistemas com o objetivo de detectar e prevenir falhas, além de garantir a integridades destes dados;
- A partir do estudo com o BI, é possível inicializar na empresa trabalhos com análises matemáticas e estatísticas para avaliar as curvas de tendência.

Referências Bibliográficas

ABNT - ASSOCIACAO BRASILEIRA DE NORMAS TECNICAS. *ISO 9000:2015*: Sistemas de gestão da qualidade - fundamentos e vocabulário. [S.l.], 2015. Citado na página 29.

AHITUV, N.; ZIF, J.; MACHLIN, I. Environmental scanning and information systems in relation to success in introducing new products. *Inf. Manage.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 33, n. 4, p. 201–211, mar. 1998. ISSN 0378-7206. Disponível em: <http://dx.doi.org/10.1016/S0378-7206(98)00025-1>. Citado na página 28.

BALLOU, R. H. *Logística empresarial: transportes, administração de materiais e distribuição física*. [S.l.]: Atlas, 1993. Citado na página 14.

CAMILO, C. O.; SILVA, J. C. d. Mineração de dados: Conceitos, tarefas, métodos e ferramentas. *Universidade Federal de Goiás (UFC)*, p. 1–29, 2009. Citado na página 13.

CHEN, H.; CHIANG, R. H.; STOREY, V. C. Business intelligence and analytics: From big data to big impact. *MIS quarterly*, v. 36, n. 4, 2012. Citado na página 13.

CHOPRA, S.; MEINDL, P. Gerenciamento da cadeia de suprimentos: Estratégia, planejamento, e operação. In: *Gerenciamento da Cadeia de Suprimentos: Estratégia, Planejamento, e Operação*. [S.l.: s.n.], 2003. Citado na página 29.

COSTA, C. N. et al. Descoberta de conhecimento em bases de dados. *Revista Eletrônica: Faculdade Santos Dumont*, v. 2, 2019. Citado na página 37.

DAVENPORT, T. H. Ecologia da informação: porque só a tecnologia não basta para o sucesso na era da informação/thomas h. *Davenport, Laurecnce Prusak, tradução Bernadette Siqueira Abrão. São Paulo: Futura*, 1998. Citado 2 vezes nas páginas 23 e 24.

DEMBKOWSKI, S.; HANMER-LLOYD, S. Computer applications–a new road to qualitative data analysis? *European Journal of marketing*, MCB UP Ltd, v. 29, n. 11, p. 50–62, 1995. Citado na página 15.

DEMCHENKO, Y. et al. Addressing big data issues in scientific data infrastructure. In: IEEE. *2013 International Conference on Collaboration Technologies and Systems (CTS)*. [S.l.], 2013. p. 48–55. Citado na página 13.

DÍAZ, J. C. *Introducción al business intelligence*. [S.l.]: Editorial UOC, 2012. Citado na página 25.

DOLAN, A.; AYLAND, C. Analysis on trial. *International Journal of Market Research*, SAGE Publications Sage UK: London, England, v. 43, n. 4, p. 1–13, 2001. Citado na página 15.

GANGADHARAN, G. R.; SWAMI, S. N. Business intelligence systems: design and implementation strategies. In: IEEE. *26th International Conference on Information Technology Interfaces, 2004*. [S.l.], 2004. p. 139–144. Citado na página 16.

KIMBALL, R.; ROSS, M. *The data warehouse toolkit: the complete guide to dimensional modeling*. [S.l.]: John Wiley & Sons, 2011. Citado 2 vezes nas páginas 24 e 25.

LAGE, M. C.; GODOY, A. S. O uso do computador na análise de dados qualitativos: questões emergentes. *Revista de Administração Mackenzie (Mackenzie Management Review)*, v. 9, n. 4, 2008. Citado na página 15.

LORENZZETTI, P. d. S. Ferramentas de business intelligence para visualização de dados na área da saúde. 2016. Citado na página 20.

MIRANDA, R. C. d. R. O uso da informação na formulação de ações estratégicas pelas empresas. *Ciência da Informação*, Scielo, v. 28, p. 286 – 292, 12 1999. ISSN 0100-1965. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651999000300006&nrm=iso>. Citado na página 24.

OLIVEIRA, M. F. D. Metodologia científica: um manual para a realização de pesquisas em administração. *Universidade Federal de Goiás. Catalão–GO*, 2011. Citado na página 31.

OLSZAK, C. M.; ZIEMBA, E. Critical success factors for implementing business intelligence systems in small and medium enterprises on the example of upper silesia, poland. *Interdisciplinary Journal of Information, Knowledge, and Management*, v. 7, n. 2, p. 129–150, 2012. Citado 3 vezes nas páginas 22, 23 e 27.

PALESTINO, C. B. *BI-business intelligence: modelagem e tecnologia*. [S.l.]: Axcel Books, 2001. Citado 2 vezes nas páginas 20 e 26.

PINTO, S. A. O. Estruturas de gestão de informações para inteligência de negócio nas organizações e o impacto individual nas atividades. 2012. Disponível em: <http://www.teses.usp.br/teses/disponiveis/12/12139/tde-05112012-183755/publico/SergioAugustoOrfaoPintoVC.pdf>. Citado na página 24.

POZO, H. *Administração de recursos materiais e patrimoniais: uma abordagem logística* . [S.l.]: Editora Atlas SA, 2000. Citado na página 16.

PRIMAK, F. V. *Decisões com BI (Business Intelligence)*. [S.l.]: Fabio Vinicius Primak, 2008. Citado na página 21.

PRODANOV, C. C.; FREITAS, E. C. de. *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico-2ª Edição*. [S.l.]: Editora Feevale, 2013. Citado 3 vezes nas páginas 31, 32 e 33.

SETZER, V. W. Dado, informação, conhecimento e competência. *DataGramaZero Revista de Ciência da Informação, n. 0*, p. 28, 1999. Citado na página 24.

SEZÕES, C.; OLIVEIRA, J.; BAPTISTA, M. Business intelligence. *São João do*, 2006. Citado na página 22.

SILBERSCHATZ, A.; SUNDARSHAN, S.; KORTH, H. *Sistema de banco de dados*. Elsevier Editora Ltda., 2016. ISBN 9788535251425. Disponível em: <https://books.google.com.br/books?id=1FBaDwAAQBAJ>. Citado na página 20.

SILVA, R. A. D.; SILVA, F. C. A.; GOMES, C. F. S. O uso do business intelligence (bi) em sistema de apoio à tomada de decisão estratégica. *Revista GEINTEC-Gestão, Inovação e Tecnologias*, v. 6, n. 1, p. 2780–2798, 2016. Citado 2 vezes nas páginas 34 e 38.

SIPPERT, J. T.; SEIFERT, J. R. S.; PORCIUNCULA, L. A integração dos sistemas bi e erp no processo gerencial de uma indústria de implementos agrícolas do noroeste do rio grande do sul. *Revista GEDECON-Gestão e Desenvolvimento em Contexto*, v. 6, n. 1, p. 94–111, 2018. Citado na página 19.

TAKAI, O. K.; ITALIANO, I. C.; FERREIRA, J. E. *Apostila: Introdução a Banco de Dados*. [S.l.: s.n.], 2005. <https://www.ime.usp.br/~jef/apostila.pdf>. Citado na página 45.

TAO, Q.; PRESCOTT, J. E. China: Competitive intelligence practices in an emerging market environment. *Competitive Intelligence Review: Published in Cooperation with the Society of Competitive Intelligence Professionals*, Wiley Online Library, v. 11, n. 4, p. 65–78, 2000. Citado 2 vezes nas páginas 27 e 28.

TEO, T. S.; CHOO, W. Y. Assessing the impact of using the internet for competitive intelligence. *Information & management*, Elsevier, v. 39, n. 1, p. 67–83, 2001. Citado na página 21.

VIANA, J. J. *Administração de materiais: um enfoque prático*. [S.l.]: Atlas, 2000. Citado na página 14.

WRIGHT, D. J. *Broadband: Business services, technologies, and strategic impact*. [S.l.]: Artech House, Inc., 1993. Citado na página 21.

Apêndices

APÊNDICE A – *Script* de criação de banco MySQL

ScriptCriacaoBanco_dwestoque.sql

```
CREATE DATABASE IF NOT EXISTS dwestoque;

USE dwestoque;

CREATE TABLE DIM_Endereco (
                cod_endereco VARCHAR(200) NOT NULL,
                nome_endereco VARCHAR(200) NOT NULL,
                capacidade_end INT,
                segmentacao VARCHAR(200),
                localizacao VARCHAR(200),
                codigo_barras BIGINT,
                PRIMARY KEY (cod_endereco)
);

ALTER TABLE DIM_Endereco COMMENT 'Tabela dimensão de Endereço';


CREATE TABLE DIM_Categoria (
                cod_categoria VARCHAR(50) NOT NULL,
                nome_categoria VARCHAR(200) NOT NULL,
                PRIMARY KEY (cod_categoria)
);

ALTER TABLE DIM_Categoria COMMENT 'Tabela dimensão da Categoria';


CREATE TABLE DIM_Segmento (
                cod_segmento VARCHAR(50) NOT NULL,
                cod_categoria VARCHAR(50) NOT NULL,
                nome_segmento VARCHAR(200) NOT NULL,
                PRIMARY KEY (cod_segmento)
);


CREATE TABLE DIM_Marca (
                cod_marca VARCHAR(50) NOT NULL,
                cod_segmento VARCHAR(50) NOT NULL,
                nome_marca VARCHAR(200) NOT NULL,
                PRIMARY KEY (cod_marca)
);

ALTER TABLE DIM_Marca COMMENT 'Tabela de dimensão de Marca';


CREATE TABLE DIM_Produto (
                cod_produto INT NOT NULL,
                cod_marca VARCHAR(50) NOT NULL,
                desc_produto VARCHAR(250) NOT NULL,
                origem_produto VARCHAR(200) NOT NULL,
                camada_palet INT NOT NULL,
                paletizacao_prod INT NOT NULL,
                shelflife_produto INT,
                PRIMARY KEY (cod_produto)
);


CREATE TABLE DIM_Palete (
                cod_palete INT NOT NULL,
                cod_produto INT NOT NULL,
                dt_fabricacao DATE NOT NULL,
                PRIMARY KEY (cod_palete)
);

ALTER TABLE DIM_Palete COMMENT 'Tabela dimensão de Palete';


CREATE TABLE Fato_001_ProdArmazenado (
```

```
                cod_endereco VARCHAR(200) NOT NULL,
                cod_palete INT NOT NULL,
                ocup_endereco DOUBLE,
                efeito_colmeia DOUBLE,
                giro_estoque DOUBLE,
                PRIMARY KEY (cod_endereco, cod_palete)
);

ALTER TABLE Fato_001_ProdArmazenado COMMENT 'Tabela de Fato de Produto
Armazenado';


CREATE TABLE FATO_Palete_002 (
                cod_palete INT NOT NULL,
                ruptura_estoque DOUBLE,
                PRIMARY KEY (cod_palete)
);

ALTER TABLE FATO_Palete_002 COMMENT 'Tabela de Fato de Palete';


CREATE TABLE DIM_Empilhador (
                cod_operador INT NOT NULL,
                nome_operador VARCHAR(250) NOT NULL,
                coordenador VARCHAR(200),
                turno_trabalho VARCHAR(200),
                horario_trabalho VARCHAR(200),
                status VARCHAR(200),
                PRIMARY KEY (cod_operador)
);

ALTER TABLE DIM_Empilhador COMMENT 'Tabela de dimensão Empilhador';


CREATE TABLE Fato_003_Movimentacao (
                cod_operador INT NOT NULL,
                cod_palete INT NOT NULL,
                cod_endereco VARCHAR(200) NOT NULL,
                tipo_fluxo VARCHAR(50),
                hora_movimentacao DATETIME,
                PRIMARY KEY (cod_operador, cod_palete, cod_endereco)
);

ALTER TABLE Fato_003_Movimentacao COMMENT 'Tabela de Fato de Movimentacao';


ALTER TABLE Fato_003_Movimentacao ADD CONSTRAINT
dim_endereco_fato_004_transporte_fk
FOREIGN KEY (cod_endereco)
REFERENCES DIM_Endereco (cod_endereco)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE Fato_001_ProdArmazenado ADD CONSTRAINT
dim_endereco_fato_001_prodarmazenado_fk
FOREIGN KEY (cod_endereco)
REFERENCES DIM_Endereco (cod_endereco)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE DIM_Segmento ADD CONSTRAINT dim_categoria_dim_segmento_fk
FOREIGN KEY (cod_categoria)
REFERENCES DIM_Categoria (cod_categoria)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE DIM_Marca ADD CONSTRAINT dim_segmento_dim_marca_fk
FOREIGN KEY (cod_segmento)
```

```
REFERENCES DIM_Segmento (cod_segmento)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE DIM_Produto ADD CONSTRAINT dim_marca_dim_produto_fk
FOREIGN KEY (cod_marca)
REFERENCES DIM_Marca (cod_marca)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE DIM_Palete ADD CONSTRAINT dim_produto_dim_palete_fk
FOREIGN KEY (cod_produto)
REFERENCES DIM_Produto (cod_produto)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE Fato_003_Movimentacao ADD CONSTRAINT
dim_palete_fato_003_movimentacao_fk
FOREIGN KEY (cod_palete)
REFERENCES DIM_Palete (cod_palete)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE FATO_Palete_002 ADD CONSTRAINT dim_palete_fato_palete_002_fk
FOREIGN KEY (cod_palete)
REFERENCES DIM_Palete (cod_palete)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE Fato_001_ProdArmazenado ADD CONSTRAINT
dim_palete_fato_001_prodarmazenado_fk
FOREIGN KEY (cod_palete)
REFERENCES DIM_Palete (cod_palete)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE Fato_003_Movimentacao ADD CONSTRAINT
dim_empilhador_fato_003_movimentacao_fk
FOREIGN KEY (cod_operador)
REFERENCES DIM_Empilhador (cod_operador)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```