

CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA

FACULDADE DE TECNOLOGIA DE INDAIATUBA

DR. ARCHIMEDES LAMOGLIA

CURSO DE TECNOLOGIA EM

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

JOÃO CARLOS CASOTO JÚNIOR

**Uso de Técnicas de Mineração de Dados e Business Intelligence na
Estratégia de Compras**

INDAIATUBA
2019

CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA

FACULDADE DE TECNOLOGIA DE INDAIATUBA

DR. ARCHIMEDES LAMOGLIA

CURSO DE TECNOLOGIA EM

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

JOÃO CARLOS CASOTO JÚNIOR

**Uso de Técnicas de Mineração de Dados e Business Intelligence na
Estratégia de Compras**

Trabalho de Graduação apresentado por João Carlos Casoto Júnior como pré-requisito parcial para a conclusão do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas, da Faculdade de Tecnologia de Indaiatuba, elaborado sob a orientação do Prof. (a) Márcio Maestrello.

INDAIATUBA
2019

CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA

FACULDADE DE TECNOLOGIA DE INDAIATUBA

DR. ARCHIMEDES LAMOGLIA

CURSO DE TECNOLOGIA EM

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

JOÃO CARLOS CASOTO JÚNIOR

Banca Avaliadora:

Prof. Márcio Maestrello	Orientador
Prof. Jaime Alexandre Matiuso	Professor Convidado
Prof. João Manoel de Campos	Professor Convidado

Data da Defesa: 04/12/2019

Agradecimentos

Agradeço ao meu orientador Márcio Maestrello pelo suporte e auxílio necessário para realização deste trabalho.

Agradeço também ao Prof. Ms. Michel Munhoz, que também tive a oportunidade de ser orientando e que me suportou na idealização deste trabalho.

Agradeço também especificamente ao meu amigo William Bueno, que me acompanhou, apoiou e me motivou no desenvolvimento deste trabalho durante o último semestre.

Agradeço também a universidade da FATEC Indaiatuba e todos os professores do curso de Análise e Desenvolvimento de Sistemas pela oportunidade de realizar este trabalho.

RESUMO

Com o aumento exponencial do fluxo de informações e competitividade no mercado de trabalho, é gerado um grande volume de dados que contém informações relevantes para a tomada de decisões estratégicas, fazendo com que a tecnologia da informação tenha papel importante no auxílio do gerenciamento e visualização destes dados. Assim, este projeto tem como objetivo apresentar o conceito das principais ferramentas disponíveis no mercado em *Business Intelligence (BI)*, demonstrando com técnicas de mineração de dados sua importância no meio empresarial, promovendo uma melhor visão departamental com maior velocidade no processamento das informações. Este projeto tem por objetivo criar métricas que permitam na aquisição de *commodities*: redução de riscos, custos e possível aumento de qualidade. A metodologia consistirá em analisar uma cadeia de fornecedores de matérias primas nacionais, agrupados por *commodities*. A validação destas hipóteses será realizada com o grupo de fornecedores existentes em uma multinacional, que possui aproximadamente 200.000 mil linhas de registros. Através da ferramenta e linguagem de programação nativa do *Power BI* será utilizada para mineração de dados, modelagem, tratamento dos dados, processamento e visualização dos dados. Assim, com a documentação, apresentação e análise dos dados, espera-se resultados que comprovem/possam mostrar a evolução da maturidade da empresa na tomada de decisões estratégicas, provocando redução de custos, menor tempo de processamento de manutenção e gestão dos dados, além de maior flexibilidade na automatização de processos.

Palavras chave: Data Science; Sistemas de Informação; Data Mining, Business Intelligence, Big Data

ABSTRACT

With the exponential growth of data and the high market competition, a giant number of relevant information is available for strategic decision-making, making information technology (*IT*) a key factor in helping to manage and visualize this data. Therefore, this project has the objective of presenting the main concepts of *Business Intelligence (BI)* tools available in market, demonstrating with data mining techniques its importance in a business environment, promoting a better department visualization with higher data processing velocity. This project has the main objective to create metrics that allow better understanding of *commodities* acquisition, such as risk, cost reduction and even quality improvement. The methodology consists in analyzing a group of national raw material suppliers, grouped by *commodities*. The validation of the hypothesis was consolidated with a group of suppliers from a multinational company, with approximately 200.000 thousands rows of data. The chosen software for this project is *Power BI* that was used for data mining, modelling, data treatment, and processing and data visualization. Therefore, with documentation, presentation and further data analysis the expected results are that *BI* can improve the evolution of the company's maturity level in strategic decision-making and therefore provoke cost reduction, shorter data processing and maintenance time and greater flexibility in process automation.

Keywords: Data Science; Sistemas de Informação; Data Mining, Business Intelligence, Big Data

LISTA DE FIGURAS

Figura 1- Fluxo de um ETL.....	16
Figura 2– Big Data e os três ‘V’s	17
Figura 3 – Fluxo Processual do Data Mining.....	18
Figura 4 – Fluxo Processual do Data Mining.....	19
Figura 5 – Diagrama de fluxo de dados	25
Figura 6 – Diferenças entre ‘Star Schema’ e ‘Snowflake Schema’	26
Figura 7 – Modelo-Entidade Relacionamento.....	28
Figura 8 - Tratamento da tabela fato utilizando a linguagem DAX	30
Figura 9 - Dimensão de data utilizando a linguagem DAX	30
Figura 10 - Criação das measures em linguagem DAX.....	31
Figura 11 - Diferenças entre visão sem (esquerda) e com deficiência cromática (direita)	32
Figura 12 - Volume de peças vs. ocorrências.....	33
Figura 13 - Gráfico suspenso com lista de fornecedores com ocorrências	34
Figura 14 - Histórico de incidentes CRI	35
Figura 15 - Gráfico suspenso com lista de fornecedores com ocorrências críticas	35
Figura 16 - Visualização de diagnóstico dos fornecedores críticos.....	36
Figura 17 - Acompanhamento de incidentes por bilhão	37
Figura 18 - Semáforo de metas.....	38
Figura 19 - Acompanhamento de incidentes críticos por bilhão	38
Figura 20 - Tabela com fornecedores de mais alto impacto	39
Figura 21 - Piores fornecedores com ocorrências críticas absolutas	40
Figura 22 - Mapa de fornecedores com maior fornecimento de peças	41
Figura 23 - Mapa de fornecedores com maiores ocorrências.....	42

LISTA DE ABREVIATURAS E SIGLAS

BI	<i>Business Intelligence</i>
CRI	Incidente críticos
DAX	<i>Data Analysis Expressions</i>
DM	<i>Data marts</i>
DW	<i>Data Warehouse</i>
ERP	<i>Enterprise Resource Planning</i>
ETL	<i>Extract, Transform & Load</i>
INC	Incidentes/ocorrências
KPI	<i>Key Performance Indicators</i>
MRP	<i>Manufacturing Resources Planning</i>
OLAP	<i>Online Analytical Process</i>
TI	Tecnologia da Informação
VP	Falhas de produção
VC	Falhas de concessão
VL	Falhas logísticas
VI	Falhas de inspeção de recebimento
VF	Falhas de campo
V0	Falhas de 0KM

SUMÁRIO

INTRODUÇÃO.....	10
CAPÍTULO I.....	13
Fundamentação Teórica	13
1.1 Aplicação de Business Intelligence em ambientes empresariais	13
1.2 OLAP.....	14
1.3 Extract, Transform and Load (ETL).....	15
1.4 Big Data, Data Mining e Business Intelligence	16
1.5 Trabalhos relacionados	19
CAPÍTULO II	21
Metodologia	21
2.1 Natureza da Pesquisa	21
2.2 Definição e apresentação das ferramentas para análise.....	22
2.3 Característica dos dados	23
2.4 Construção metodológica dos dados	24
2.5 Topologia da estrutura de Business Intelligence.....	26
2.6 Banco de Dados: Modelo Entidade Relacionamento (MER)	27
2.7 Requisitos de Software	29
2.8 Requisitos de Hardware	29
2.9 – Implementação dos datamarts/cubo	29
CAPÍTULO III.....	32
Análise dos Dados e Resultados.....	32
3 Gráfico comparativo de volume de peças vs. ocorrências.....	33
3.1 Histórico de incidentes críticos	34
3.2 Diagnóstico de análise dos fornecedores críticos.....	36
3.3 Gráfico de incidentes por bilhão mensal (Inc/bn)	36
3.4 Gráfico de incidentes críticos por bilhão mensal (CRI/bn)	38
3.5 Fornecedores de mais alto impacto	39
3.6 Piores fornecedores com ocorrências críticas	40
3.7 Fornecedores com maior fornecimento de peças (volume treemap).....	40

3.8	Fornecedores com maiores ocorrências (incidents treemap)	41
-----	--	----

CONSIDERAÇÕES FINAIS.....	43
----------------------------------	-----------

REFERÊNCIAS	45
--------------------------	-----------

INTRODUÇÃO

O aumento do fluxo de informações e necessidade constante de obter respostas no menor tempo possível gera uma alta competitividade entre as empresas do mundo. A tecnologia da informação tem como principal função auxiliar na organização e disposição destes dados da maneira mais concisa possível.

O conceito de Big Data é essencialmente um dos tópicos mais disseminados em ambientes empresariais, por tratar-se de uma nova tecnologia que poucas corporações sabem usá-la efetivamente a seu favor. Trata-se de uma tecnologia com aplicação versátil em diversos setores, tais como agricultura, transportes e também de serviços que traz inúmeras vantagens, especificamente na tomada de decisões estratégicas e novos insights.

Em empresas cujo fluxo de informações é alto, é altamente recomendada a implementação de uma ferramenta de *BI* com intuito de estruturar, organizar e visualizar os dados coletados para suportar propriamente na tomada de decisões, afetando diretamente na redução de custo, maior agilidade de tempo de processamento de manutenção e gestão dos dados e até prospecção de novos negócios.

De acordo com Silveira (2007), a construção de um modelo de BI para um estudo de caso traz valor estratégico uma vez que os dados sejam coletados de maneira planejada e organizada dentro do processo escolhido.

Kemczinski *et al* (2003) afirmam que a tecnologia afeta o próprio valor das atividades ou permite que as organizações conquistem a vantagem competitiva através da exploração das mudanças no escopo da competição, nomeado como “diferenciação estratégica”.

Conforme elencado por Charles e Falsarella (1995), existem diferentes motivos para o desenvolvimento e implantação de sistemas de tecnologia da informação nas organizações que requerem diferentes níveis de mão-de-obra gerencial/operacional, conforme apresentado na tabela de motivos para implantação de Tecnologia da Informação (TI) abaixo:

Motivos para implantação de TI	%
Melhoria de controles organizacionais	24
Maior participação	22
Aumento de produtividade	20
Redução de custos	10
Redução de competitividade	8
Melhoria da qualidade de atendimento	8
Manter participação	5
Maior necessidade de integração	2
Gerar um ambiente criativo	1

Tabela 1 (adaptado): Motivos para implantação de TI

Fonte: (PRATES, 2004 vol.8 no.2)

A pesquisa em questão se baseia nas principais questões norteadoras: a utilização de ferramentas de *BI* é realmente eficaz para empresas com grande fluxo de informações? É possível reduzir custos da empresa com suporte de ferramentas de modelagem estatística? A utilização de uma ferramenta de *BI* realmente gera vantagem competitiva no mercado?

Este projeto tem por objetivo criar métricas que permitam na aquisição de commodities: reduzir riscos, reduzir custos e aumentar qualidade.

A pesquisa será fundamentada na metodologia de pesquisa de natureza experimental (aplicação de estudo de caso) que, de acordo com Fonseca (2002), diz que a pesquisa experimental seleciona grupos de assuntos coincidentes, submete-os a tratamentos diferentes, verificam-se as variáveis estranhas e a partir dos dados reportados apreender as relações de causa e efeito ao eliminar explicações conflitantes com as descobertas realizadas. Ela será composta da análise estatística de fornecedores de matérias primas nacionais, agrupados por *commodities*, e entender nos respectivos aspectos:

- Identificar quais são os fornecedores que trazem maior risco à organização em relação às suas ocorrências e identificar quais obtém maior impacto relacionados a falhas críticas (ex.: *recall* das montadoras).
- Criar um indicador de qualidade e volume de peças (quali-quantitativo) que permita identificar qual deles possui menor risco ao cliente e, em contrapartida, quais são os fornecedores mais estratégicos em termos de Qualidade.

O ambiente de validação da hipótese será no processo de Compras de uma organização de grande porte. É um cenário promissor para análise, visto que a complexidade das informações e a necessidade de correlação dos dados para disponibilizar informações estratégicas.

Com relação à estrutura da pesquisa a ser desenvolvida, esta será apresentada em três capítulos: no capítulo I, será exposta a fundamentação teórica, que traz os conceitos chave que fundamentam o problema a ser estudado neste trabalho. Adicionalmente, serão apresentados os trabalhos relacionados ao problema de *Big Data* e como foram feitas as aplicações das ferramentas de *BI*.

No capítulo II, será apresentada a metodologia adotada, os passos percorridos para a realização desta pesquisa e consequentemente a fundamentação e aplicação técnica de todo o desenvolvimento deste trabalho.

No capítulo III, serão apresentados detalhadamente os indicadores gerados dos dados obtidos por meio da plataforma consolidada (*Dashboard*). Por fim, são apresentadas as considerações finais e referências utilizadas neste projeto.

CAPÍTULO I

Fundamentação Teórica

1.1 Aplicação de Business Intelligence em ambientes empresariais

Dentro do contexto apresentado neste projeto, é visto que as companhias acumulam cada vez mais dados e acabam consequentemente não gerando quantidade suficiente de informações que possam ser relevantes para suporte na tomada de decisões estratégicas, e a aplicação das tecnologias de informação têm impacto crucial na melhoria destes processos.

De acordo com Sebrae (2000), as vantagens que a Tecnologia da Informação (TI) traz para as empresas podem ser divididas em três grupos:

a) **Menores custos:** a informática, quando bem utilizada, reduz os custos da empresa, porque agiliza e possibilita maior segurança e confiabilidade nos processos, rotinas e controles administrativos; simplifica as tarefas burocráticas; reduz os erros e praticamente elimina a repetição do trabalho.

b) **Maior produtividade:** possibilita que as pessoas produzam mais, em menos tempo, com menor dispêndio de recursos; permite aproveitar melhor a capacidade produtiva da empresa com o planejamento e o controle da produção; armazena e localiza imediatamente informações fundamentais para os negócios; agiliza os processos de tomada de decisões em relação a preços, estoques, compras e vendas, entre outros.

c) **Maior qualidade:** a qualidade dos produtos e serviços é melhorada, pois as tecnologias de informação ajudam a manter o padrão dos produtos dentro das especificações estabelecidas; proporciona melhores condições de trabalho para os empregados; reduz esforços com a burocracia para concentrá-los nas atividades fins da empresa.

Zuboff (1994) também aponta a importância da aplicação de tecnologias de informação em um ambiente empresarial e suas principais vantagens mensuradas. São apontados como benefícios mais significativos o aumento da velocidade de resposta e redução da redundância de atividades operacionais como fatores modificadores-chave, além das demais apresentadas na tabela 2:

Benefícios	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	Média	Desvio padrão
Aumento de continuidade							
Facilita a integração funcional	15	45	15	18	12	2,82	1,06
Aumento da automação operacional	12	48	20	10	10	2,58	1,19
Aumento da velocidade de resposta	15	10	20	45	20	4,65	1,30
Melhoria do fluxo de trabalho	15	40	15	15	15	2,75	1,35
Melhoria nos controles							
Redução de custos operacionais	18	35	21	12	14	2,71	1,41
Melhoria na precisão de entrada de dados	19	41	20	12	08	2,49	1,60
Melhoria na acuidade de saída de relatórios	14	41	15	12	18	2,85	1,25
Redução de redundância de operações	07	10	10	48	25	4,18	1,85
Aumento na previsibilidade de cenários	18	36	22	12	12	2,64	1,69
Relatórios mais consistentes	10	04	34	20	32	3,81	1,54
Menor risco na tomada de decisão	25	40	15	10	10	2,43	1,56
Melhoria de compreensão das funções produtivas							
Melhoria no processo de planejamento e direção na organização	15	25	40	20	20	3,65	1,12
Aumento de satisfação dos usuários	08	07	05	15	65	4,85	0,98
Melhoria na capacidade de visão total do negócio	13	27	40	10	10	2,77	0,99
Melhoria na capacidade de auto-avaliação da qualidade de serviço	12	20	30	28	10	3,04	1,02
Aumento na capacidade de correlação entre as diversas áreas da empresa	06	14	30	33	17	3,41	1,27

Escala utilizada: 1= Nenhuma, 2= Baixa, 3= Média, 4= Elevada, 5=Total.

Tabela 2: Benefícios obtidos pela TI em pequenas empresas estudadas

Fonte: (PRATES, 2004 vol.8 no.2)

Vale ressaltar que há uma redução de latência entre o tempo de processamento dos dados (fator de aumento da automação operacional), obtendo assim um fluxo de informações mais imediato. Este fator está diretamente ligado à redução da redundância das atividades operacionais, automatizando a coleta de dados através de um modelo de dados que traz redução de custo à empresa.

1.2 OLAP

De acordo com ANZANELLO (2007), OLAP (*Online Analytical Processing* ou Processamento Analítico Online) são métodos e técnicas que solucionam com maior facilidade problemas de aquisição, análise e consolidação de dados através de seu processamento

analítico, obtendo desta forma uma topologia de informações para construção de *Data Warehouses* (DWs) e consequentemente *Data Marts* (DMs).

De acordo com Kimball (1998), a definição de um *DW* se dá como o processo de preparar os dados de um sistema de informação operacional de forma a se ter uma fonte de informações que possam dar suporte à tomada de decisões. Abaixo são apresentadas algumas justificativas quanto à obtenção desta tecnologia em uma empresa:

- a) Grande quantidade de sistemas transacionais existentes.
- b) Integrar dados de múltiplas fontes.
- c) Recuperação de dados históricos.
- d) Facilitar o acesso à informação.
- e) Atender a diferentes tipos de usuários finais.

Para este trabalho também serão citadas outras características geradas a partir da aplicação do OLAP, tais como:

- **Cubo:** Estrutura base dos dados do estudo de caso em formato multidimensional, obtendo informações de diferentes fontes para futura análise;
- **Dimensão:** Facetas do cubo em que é possível extrair as informações singulares do estudo de caso, por exemplo as tabelas de dados;
- **Hierarquia:** Define-se como as camadas de uma dimensão, podendo ser útil para agrupar ou estratificar informações de uma determinada visualização gráfica, tornando a informação mais dinâmica.
- **Medida:** Modelo de dimensão que é utilizada para fazer comparações e cálculos lógicos e/ou matemáticos.

1.3 Extract, Transform and Load (ETL)

ETL (Extração, Transformação e Carga – Extract, Transform Load) é uma metodologia comumente utilizada em projetos de *DW* e *BI*, responsável pelas etapas de aquisição do dado, limpeza (*data cleaning*), otimização e inserção destes em um *DW* em grande volume para posterior processamento em um banco de dados, conforme apresentado na figura 1 abaixo:

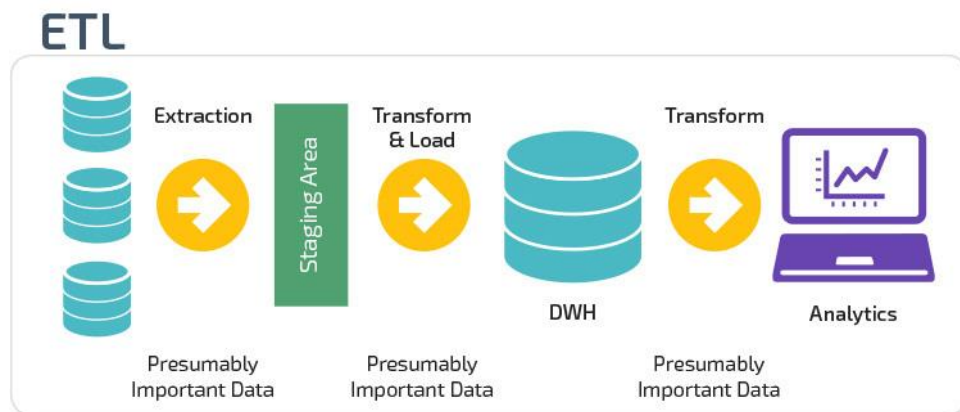


Figura 1- Fluxo de um ETL

Fonte: (PANOPLY, 2019)

O termo ETL será citado com intuito de demonstrar a metodologia de processos aplicados desde a aquisição até o processamento final do dado, extraindo como seu produto final informações em um *Dashboard* (relatório) a partir dos dados coletados.

1.4 Big Data, Data Mining e Business Intelligence

De acordo com o Instituto McKinsey (2011), conceito de Big Data está determinado pela quantidade de dados que o mundo é capaz de produzir, mas que é pouco explorado e por isto torna-se um fator estratégico como uma fronteira para inovação, produtividade, competição e vantagem competitiva no mercado àqueles que sabem usar ao seu favor. Adicionalmente, é apresentado na figura 2 a tríade que resume os principais desafios na composição e manipulação do Big Data: volume, velocidade e variedade.

— The 3 V's of Big Data

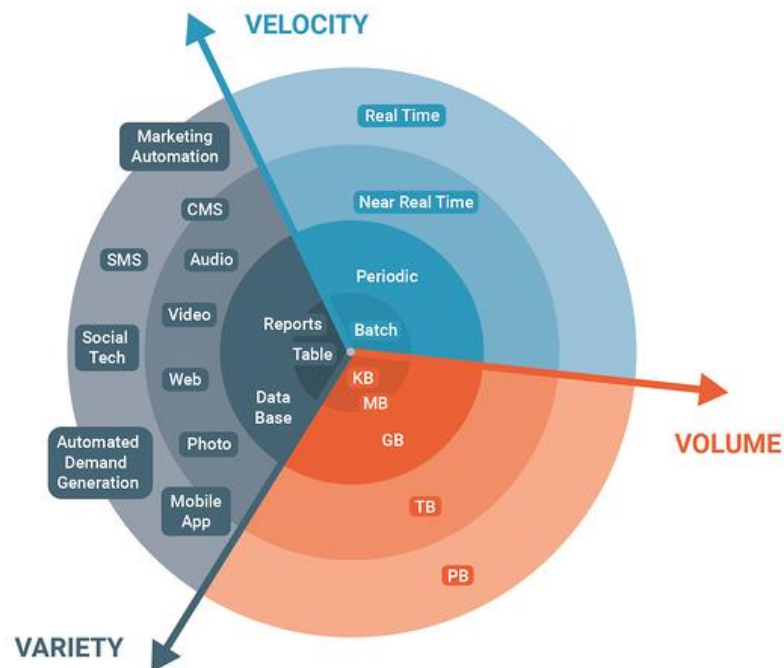


Figura 2– Big Data e os três ‘V’s

Fonte: (PICKELL, 2018)

Sabendo do cenário atual, é necessário a utilização de técnicas de *Data Mining* (ou mineração de dados), que por definição trata-se de algoritmos e técnicas para varredura, busca, classificação, correlação e limpeza dos dados adquiridos dentro de um *DW*, para que desta forma os dados possam estar otimizados para futura análise.

É possível também incluir análises estatísticas mais avançadas dentro de um modelo, tal como análise preditiva, modelos de regressão e outras técnicas através de identificação de padrões e tendências com os dados passados, extraíndo informações valiosas a partir dali. A figura 3 traz uma visão clara da etapa que o *Data Mining* se aplica no processo, como um intermediador do pré-processamento de dados até a consolidação de uma plataforma visual:

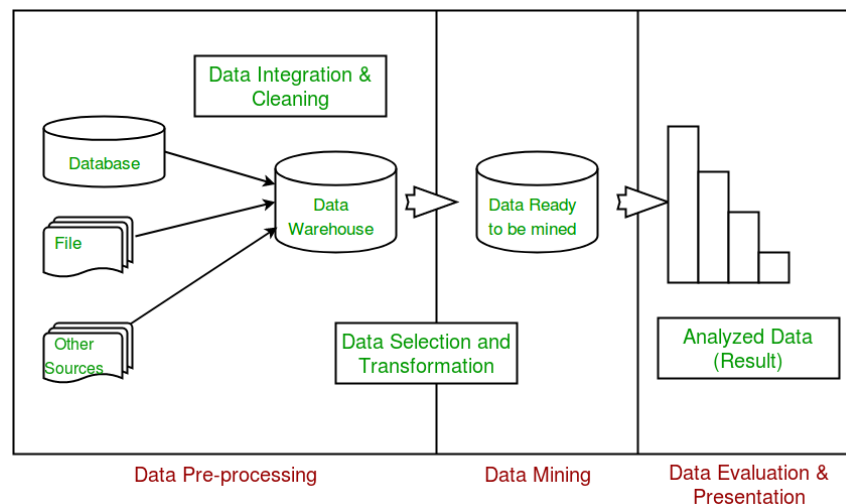


Figura 3 – Fluxo Processual do Data Mining

Fonte: (KOHLI, 2019)

No entanto, mesmo após o tratamento destas bases de dados, é necessário utilizar as técnicas corretas para transformar os dados ali apresentados em informações que possam ser úteis ao usuário, trazendo estes benefícios em redução de custos, aumento de produtividade ou qualquer outra potencialização que seja benéfica à organização.

Para isto, foi utilizado neste trabalho a ferramenta *PowerBI* para modelagem das dimensões, estruturação das bases de dados e visualização em um *Dashboard* (plataforma) interativo como resultado final da análise. Um exemplo de *Dashboard* com a ferramenta é apresentado na figura 4 com *KPIs* (*Key Performance Indicators*) demonstrativos ao departamento de Finanças corporativo:

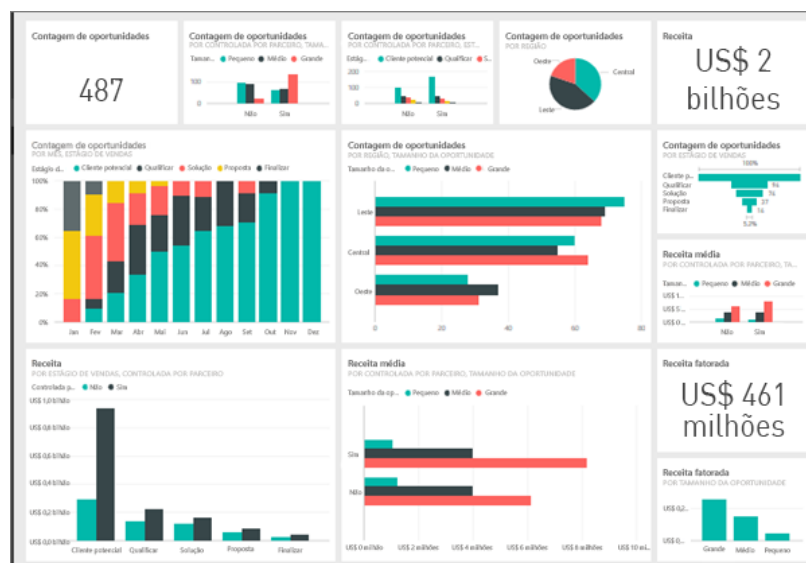


Figura 4 – Fluxo Processual do Data Mining

Fonte: (Microsoft, 2019)

1.5 Trabalhos relacionados

Nesta seção são apresentados estudos e pesquisas relacionados ao tema deste trabalho. A ferramenta que foi utilizada para pesquisa e fundamentação teórica deste trabalho foi o Google Acadêmico, por meio do qual se buscou mapear as pesquisas dessa natureza circunscritas nos últimos anos.

O levantamento realizado foi orientado pela busca de pesquisas científicas e/ou tecnológicas que têm em seus objetivos o desenvolvimento/implementação e/ou análise de ferramentas, tecnologias para utilização de ferramentas de *BI* para auxílio na tomada de decisões estratégicas.

Kemczinski *et al* (2003) abordam sistematicamente as vantagens de utilização de um sistema de *BI*, além de identificar as ferramentas que dão suporte à estas tarefas. Em sua metodologia, eles abordam os diferentes tipos de sistemas *Front-end* e *Back-end* que compõem a estrutura lógica de um sistema *BI*. Os resultados apresentados concluem que a utilização da ferramenta de *BI* pode ser considerada uma vantagem competitiva, mas que o desenvolvimento e manutenção desta vantagem dependem do processo de valor agregado oferecido pela organização.

Affeldt, F. S., Silveira, F. C. da S. e Vanti, A. A. (2006) publicaram um artigo que aborda a utilização do *BI* em um ambiente empresarial ao departamento de Controladoria. A metodologia aplicada por eles é feita através de um estudo de caso dentro do departamento de Controladoria, estruturado pelas definições dos três tópicos e o relacionamento entre eles. Os resultados levantados dizem que a implementação de *BI* é altamente vantajosa para suporte na tomada de decisões operacionais, gerenciais e estratégicas.

Reginato, L. e Nascimento, A. M. (2006) levantaram uma base de dados para saber o quanto a tecnologia de *BI* poderia contribuir com o suporte à área de Controladoria. A metodologia aplicada foi de estudo de caso, com pesquisa em campo (com entrevistas) para levantamento das evidências necessárias. Como resultado do levantamento citado, conclui-se que o *BI* é essencial não apenas na área de Controladoria, mas em qualquer departamento

(estendendo até às áreas operacionais) trazendo uma significativa melhoria em suas atividades internas, visto que esta auxiliou no funcionamento sistêmico da organização.

Fortulan, M. R. (2007) com suporte das ferramentas *Analysis Services* e banco de dados MySQL, uma solução utilizando *BI* propusesse receber os dados de uma unidade fabril em estratificação por *KPIs*. A metodologia aplicada para o trabalho foi categorizada como aplicada (pois há interesses práticos e aplicáveis) e experimental (pois foram usados dados reais para sua aplicação), além da utilização do método dedutivo. Os resultados desta aplicação foram bem-sucedidos, desenvolvendo e aplicando com efetividade o sistema com integração ao sistema *Enterprise Resource Planning* (ERP) ara apoio à gestão na tomada de decisões que integrassem os demais módulos vinculados ao departamento escolhido.

Silveira, F. C. da S. (2007) criou um modelo que atendesse a visão de informações estratégicas, com estudo de caso voltado à Controladoria. Sua metodologia foi fundamentada em estudos de caso, estratificadas em grupos exploratórios, descritivos e explicativos, mas identificada especificamente ao estudo de caso exploratório de natureza qualitativa. Os resultados apresentados neste trabalho concluem que a informação é considerada um dos recursos mais importantes de uma organização, e que se ganha valor estratégico ao organizar e apresentar os dados de forma estruturada.

É possível verificar abaixo na tabela 3 uma comparação entre as referências citadas e este trabalho, apresentando em cada caso qual setor a implementação do *BI* foi direcionado e executado:

Autor	Estratégica (Oque)	Tática (Como)	Operacional (Executor)
Kemczinski <i>et al</i> (2003)	X	X	X
Affeldt, F. S., Silveira, F. C. da S. e Vanti, A. A. (2006)	X		
Reginato, L. e Nascimento, A. M. (2006)	X		
Fortulan, M. R. (2007)			X
Silveira, F. C. da S. (2007)	X		
Casoto, Joao (2018)	X	X	X

Tabela 3: Comparação dos trabalhos em relação às abordagens de *BI*

Fonte: Próprio Autor

CAPÍTULO II

Metodologia

2.1 Natureza da Pesquisa

Na produção de um trabalho científico, a metodologia é composta por métodos e técnicas que designam um dos variados caminhos para desenvolvimento da análise dos dados coletados e consequentemente produção dos resultados esperados na conclusão.

A metodologia escolhida para esta análise é experimental, exposta através de um estudo de caso em um processo da organização de Compras no departamento de engenharia de Qualidade de fornecedores. Gil (2017) estrutura este tipo de pesquisa em um processo de determinação de um objeto de estudo, obtenção das variáveis influenciáveis no projeto, definição das formas de controle e observação dos efeitos que as variáveis causam no objeto.

De acordo com Regina (2013), para a aplicação do estudo de caso a análise foi estruturada nas seguintes etapas:

- a) Delineamento da pesquisa (objetivos da análise);
- b) Desenho da pesquisa;
- c) Preparação e coleta dos dados;
- d) Análise dos casos;
- e) Elaboração dos relatórios.

Na primeira etapa é declarado quais são os objetivos da análise que será estudada, delimitando assim sua área de pesquisa e selecionando o objeto a ser investigado. Nesta etapa é definido o ponto de partida entre uma análise exploratória, explicativa ou descritiva. Pelos meios e fins em questão, esta pesquisa caracteriza-se como descritiva, uma vez que o desenvolvimento deste *Dashboard* tem como objetivo estabelecer padrões e determinar comportamentos de determinados registros dentro de uma série de dados no passado.

A segunda etapa define a identificação de padrões e esboço das técnicas a serem utilizadas durante a posterior análise dos dados. Nesta etapa valida-se também a confiabilidade

da informação, inferências dos dados obtidos e definição das atividades desenvolvidas para a pesquisa e definição do protocolo.

Na etapa de preparação de dados é feita a seleção dos casos a serem utilizados para a análise e quais serão descartados. Em seguida, é realizada a coleta dos dados de acordo com a necessidade do modelo e objetivo da pesquisa. Finalmente, é organizado o material para compilar a base de dados do estudo de casos.

A análise dos casos é a etapa onde efetivamente serão realizados os estudos qualitativos, triangulação e síntese dos dados, onde o material finalmente compilado e esboçado em forma de relatório final para apresentação (etapa elaboração dos relatórios).

2.2 Definição e apresentação das ferramentas para análise

Para este projeto as análises estatísticas serão feitas em linguagem de programação nativa do Power BI Power Query e DAX, visto que estas ferramentas trazem maior facilidade para desenvolver modelagens matemáticas e formulação da lógica de programação, respectivamente.

Visto que é um requisito trabalharmos com dados estruturados, será utilizado também a própria modelagem nativa de banco de dados do Power BI para formular esta estrutura para o estudo de caso escolhido e posterior consolidação do *Dashboard* para visualização dos dados em relatórios.

O processo de aquisição, gerenciamento e compilação dos dados de forma estruturada foi feito através das ferramentas de software *ETL*, criando-se consequentemente um *Data Mart* como principal repositório de base de dados da área de estudo escolhida.

Para que seja iniciado o processo de análise será estruturada uma base de dados a partir de diversas fontes (tabelas de preços de compra/venda e composição do produto, indicadores de qualidade do fornecedor, controle de *Material Requirement Planning* (MRP) e dados básicos cadastrais da base de fornecedores) disponíveis em um ERP gerenciado pela própria empresa.

Os dados coletados serão consolidados em um *Data Mart* como principal base de dados para posterior análise e preparação dos dados (junções, consolidações e demais alterações estruturais nos dados), importados em formato de banco de dados estruturado no próprio *Power BI*. Neste processo serão definidos os dados apurados para a análise e qual a melhor maneira de prepara-los para apresentação dos resultados.

Dentro da análise experimental dos dados, o *Power BI* atuará como *Staging Area*, que dentro da metodologia de *ETL* tem a função de importar estas bases de dados em formato de tabelas temporárias e realizar o tratamento dos dados para o carregamento em versão final para utilização e visualização. Visto que para esta amostragem de dados a ferramenta já suporta nativamente todas as etapas, não é necessária a utilização de um banco de dados externo.

Após a preparação da base de dados, inicia-se o processo de modelagem e análise estatística dos dados, utilizando aplicações para construção da lógica em código, segmentação de dados, distribuição e outras técnicas com suporte da linguagem nativa do Power BI (Power Query e DAX) para estruturação lógica dos dados estudados e reprodução das informações geradas a partir das aplicações.

Após a consolidação dos dados, esta linguagem suportará na classificação, categorização e implementação de um algoritmo nesta base de dados para apresentar o potencial de aplicação de *Data Science* (ciência de dados) no estudo de caso. A utilização da linguagem de programação nativa ajudará especificamente nos problemas de localidade de fornecedores.

Após a criação do modelo de mineração de dados, as informações serão consolidadas e os resultados apresentados por meio de um *Dashboard* interativo na ferramenta *PowerBI*, apresentando as métricas apontadas na Introdução deste projeto. Em seguida serão avaliadas as informações geradas a partir do modelo definido e, caso a compreensão do negócio e dados tenha sido bem-sucedida, será feita uma avaliação final e posterior implementação do modelo.

2.3 Característica dos dados

Para este trabalho, foi simulado um volume de dados próximo a duzentas mil (200.000) linhas como valor amostral para análise, suficiente para gerar maior acurácia nos indicadores apresentados nos relatórios, mas também tendo em vista a total capacidade do sistema *Power BI* em gerir bases de dados de pequeno à médio porte sem a necessidade de utilização de um banco de dados externo (ex: *SQL Server*).

As bases de dados foram extraídas de uma empresa de médio porte do mercado automotivo, ligadas ao processo de engenharia de Qualidade de fornecedores, apontando seu respectivo fornecimento de *commodities* e controle de Qualidade destes durante o fluxo entre fornecedor-cliente.

O departamento de engenharia de Qualidade tem função de garantir que os materiais recebidos na planta estejam de acordo com as especificações técnicas e visuais para dar continuidade no fluxo de produção do produto final, bem como mitigar falhas do processo de Qualidade (nomeadas como ocorrências ou incidentes) estabelecendo métricas e ações àqueles fornecedores que apresentam maior risco.

Assim sendo, os principais apontamentos dos registros estão diretamente ligados ao volume de fornecimento de peças produzidas, quantidade absoluta de ocorrências categorizadas através do impacto individual de cada e métricas calculadas (como incidentes por bilhão ou incidentes críticos por bilhão) que serão apresentadas nos tópicos posteriores.

Para garantir o anonimato e confidencialidade das informações durante o desenvolvimento, o mascaramento de dados foi feito através da substituição dos nomes dos fornecedores na análise pelo código de cadastro do sistema, que não possui nenhum vínculo ou relação com seu nome. As *commodities* que compõem os gráficos foram separadas entre componentes eletrônicos, produtos de estamperia, forjaria, usinados e plástico agregado à borracha.

Os dados foram extraídos durante o período de 2016 a 2019 (até Agosto de 2019), coletando informações descritivas de uma base de fornecedores restrita ao território nacional e separadas entre:

- Planilha comparativa de uma cadeia de fornecedores de uma determinada *commodity* entre valores atingidos versus meta estabelecida (fictícia);
- Indicadores de Qualidade do fornecedor (*commodity* do produto fornecido, volume de peças recebidas, volume de peças reclamadas, incidências/falhas de Qualidade em produtos fornecidos à empresa, tipos possíveis de falhas dentro de um fluxo de distribuição de peça/produto);

Os dados escolhidos nesta amostragem serão apresentados como visões (através dos *datamarts*) através das análises estatísticas apresentadas no formato de *Dashboard*. A consolidação do modelo do banco entre tabela fato e dimensões serão explicados nos próximos tópicos.

2.4 Construção metodológica dos dados

A metodologia escolhida para esta análise é experimental, exposta através de um estudo de caso em um processo do departamento de Compras. A pesquisa foi dividida entre preparação e coleta dos dados, análise dos casos e elaboração dos relatórios.

Para desenvolver este trabalho, foram seguidos os passos fundamentais do *ETL* para estruturar o fluxo processual desde a obtenção dos dados até a consolidação dos Dashboards, apresentados na figura 5 em visão de diagrama:

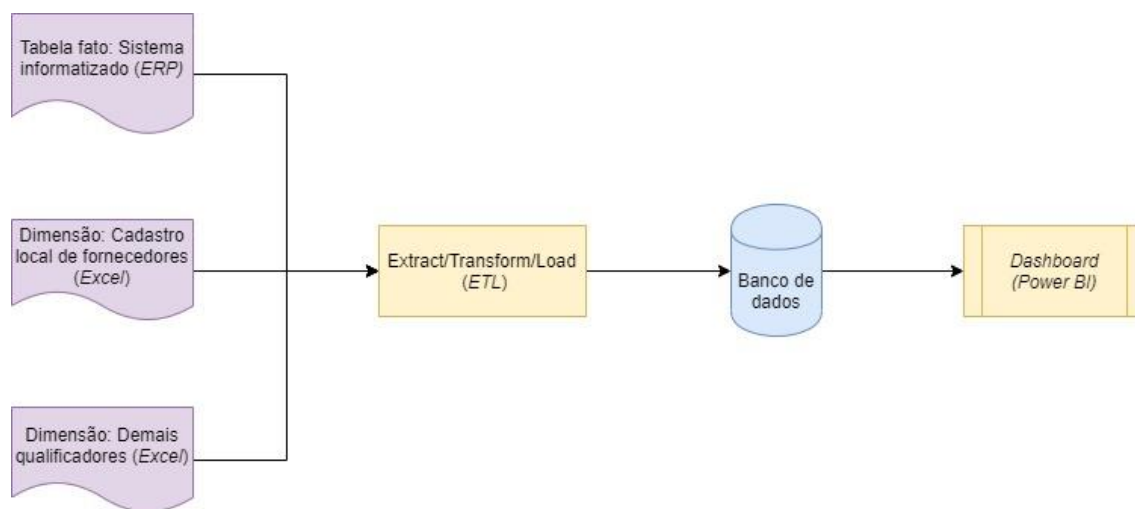


Figura 5 – Diagrama de fluxo de dados

Fonte: Próprio Autor

As bases de dados foram coletadas através de sistemas informatizados e já estruturados, tal como exposto no diagrama da figura 5. Em paralelo, também foram utilizados neste trabalho tabelas alimentadas através de controle manual para enriquecer as informações apresentadas no *Dashboard*, tal como a base de cadastro de localidade dos fornecedores e demais dimensões que serão apresentadas no tópico implementação dos *datamarts*/cubos junto à preparação dos dados (*data mining*).

A preparação dos dados foi realizada diretamente no *Power BI*, visto que neste trabalho terá função de *Staging Area*, removendo as colunas que não são relevantes para a análise, limpando dados indesejados ou incorretos que possam comprometer na visualização final dos relatórios, incluir tabelas e colunas calculadas implementando lógica que enriqueça a análise agrupando informações e gerando métricas para a elaboração dos relatórios.

Depois do tratamento da base de dados, foi feita a importação das tabelas para o *Power BI*, estruturando o banco de dados e estabelecendo a cardinalidade entre as tabelas

fato/dimensão, criando assim um modelo estruturado de entidade/relacionamento entre as fontes de dados.

Por fim, foram desenvolvidos *Dashboards* consolidando estas informações de forma que possamos expor em uma determinada base de fornecedores aqueles que possuem maior risco através da quantidade de falhas ocorridas dentro do tempo de análise escolhido, apresentar o histórico de ocorrências dentro desta linha de tempo e expor os resultados por *commodities* e localidade das plantas da empresa escolhida o histórico de seus indicadores e seu *status* por meta estabelecida.

2.5 Topologia da estrutura de Business Intelligence

Houve necessidade de implementação de um *DW* devido ao volume de dados e complexidade do trabalho. Sua estruturação foi inspirada em modelo similar ao esquema estrela (*star schema*), idealizado por Ralph Kimball. Esta modelagem dimensional é estruturada a partir de uma tabela fato e é rodeada por tabelas dimensão, diretamente ligadas à tabela fato.

Por definição, a tabela fato tem como objetivo armazenar as métricas (registros) e chaves identificadoras para que esta possa se interligar com as tabelas dimensão (ou dimensões), com objetivo que qualificar os dados e desta forma poder conceder atributos aos registros ali apresentados, descrevendo e possibilitando uma melhor classificação dos dados.

Em adição ao *star schema*, foi idealizado por Bill Inmon um modelo inspirado neste, porém com maior complexidade devido a criação de ramificações às dimensões que conectam com a tabela fato, conhecido como modelo em “floco-de-neve” (*snowflake schema*). Na figura 6 é apresentado uma representação visual com as diferenças entre os dois modelos:

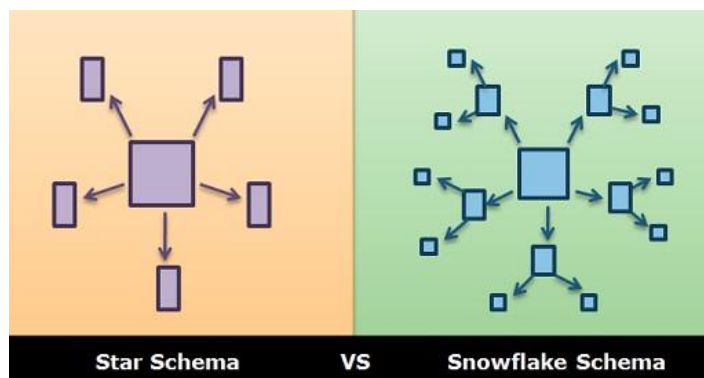


Figura 6 – Diferenças entre ‘Star Schema’ e ‘Snowflake Schema’

Fonte: TechDifferences (2017)

A utilização do *snowflake schema* se mostrou mais adequado para este trabalho e trouxe integração entre todas as dimensões necessárias para o estudo de caso escolhido e com desempenho de processamento de informações desejável para a amostragem definida.

2.6 Banco de Dados: Modelo Entidade Relacionamento (MER)

Conforme mencionado no tópico anterior, a definição deste modelo de dados foi inspirada no *snowflake schema*, visto que foi necessária a criação de tabelas auxiliares para conexão (pontes ou *bridges*) entre algumas dimensões com a tabela fato.

Na relação abaixo encontram-se disponíveis os nomes das tabelas e uma breve descrição de cada:

1. Fact_Table: Tabela fato com registros (fatos) de todos os lançamentos de ocorrências/volume de uma determinada base de fornecedores;
2. Nome Fantasia: Dimensão com nomes-fantasia da base de fornecedores;
3. Dim_RBSNO: Tabela-conexão entre tabelas *Fact_Table* e *Nome Fantasia*;
4. Dim_Part_No: Banco de materiais fornecidos cadastrados;
5. Dim_PQT: Especificação de *commodity* do produto e meta;
6. Dim_Date: Dimensão de datas para determinar fluxo e ordem das informações;
7. Bridge_Semester: Tabela ponte entre tabelas *Dim_Date* e *Averages_PQTs*;
8. Averages_PQTs: Tabela com indicadores agrupados semestralmente;
9. Bridge PQT Department: Tabela-conexão entre tabelas *Fact_Table* e *Real x Atingido*;
10. Real x Atingido: Tabela comparativa entre resultados avaliados durante período versus meta estabelecida pela empresa;
11. Country_flag: Tabela referência para adquirir a bandeira e nacionalidade do fornecedor;
12. Select Top N: Tabela auxiliar para determinar filtro de fornecedores nas visualizações de mapa de volume de fornecedores e mapa de incidentes de fornecedores;
13. Key Measures: Repositório com todas as linhas de cálculo (chamadas de *measures*, ou medidas) utilizadas para a consolidação do *Dashboard*.

Na figura 7 é apresentada o modelo estruturado com as respectivas cardinalidades entre as tabelas utilizadas durante a consolidação desta plataforma:

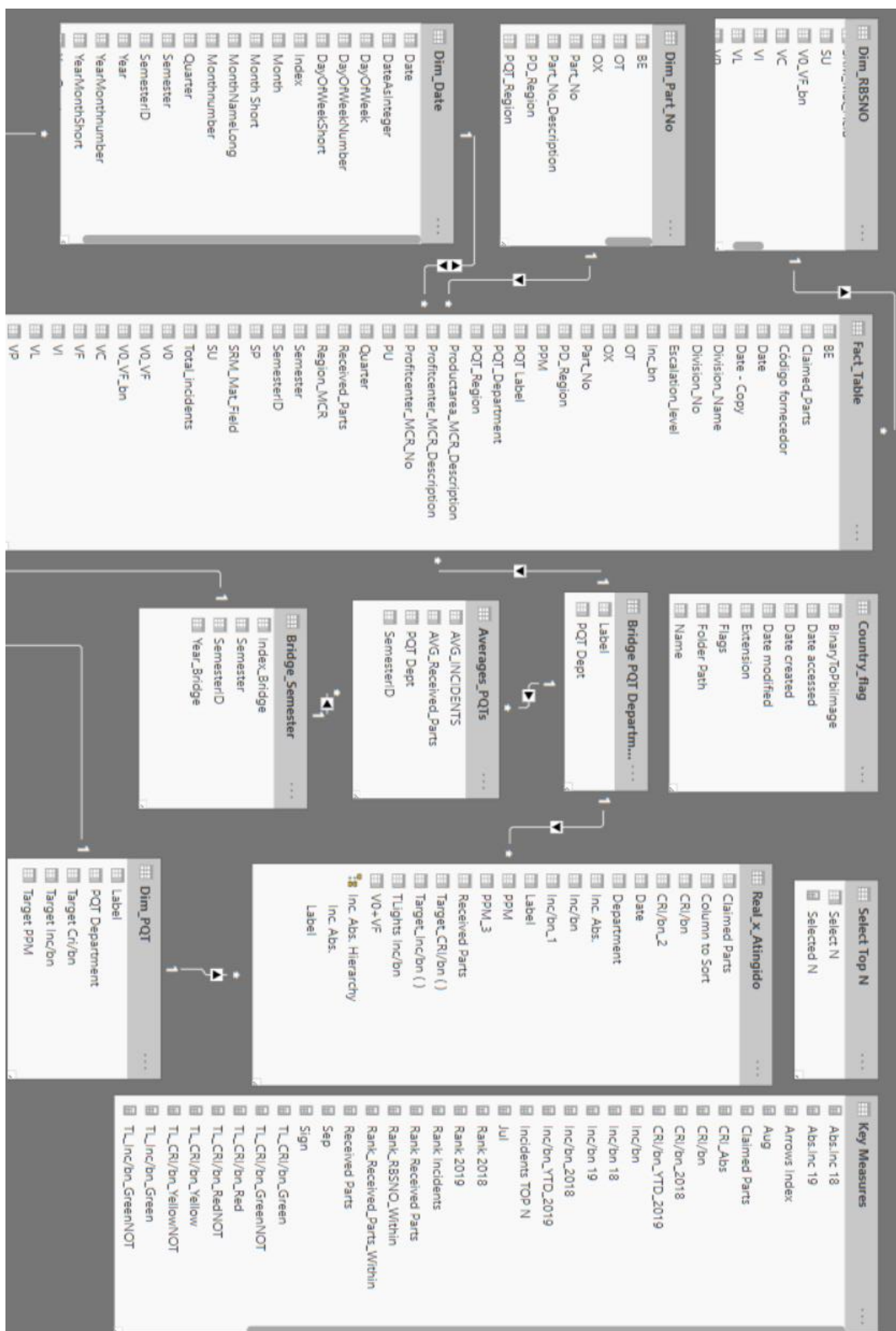


Figura 7 – Modelo-Entidade Relacionamento

Fonte: Próprio Autor

2.7 Requisitos de Software

Para este trabalho foi utilizado o Software *Power BI Desktop*, versão 2.73.5586.1101 64-bit (setembro de 2019). Todas as integrações e modelagem relacionadas a banco de dados foram feitas diretamente na ferramenta e sem a necessidade de uma ferramenta externa à esta.

Para criação e formatação deste trabalho, foi utilizada a ferramenta *Microsoft Word 2016*.

2.8 Requisitos de Hardware

Para esta pesquisa foi utilizado um Notebook Acer Aspire F5-573G-50KS, com processador Intel Core i5 7200U (velocidade de 2.5 até 3.1 GHz) e 3 mb de cache, 8 gb de memória RAM DDR4 2133 MHz (2x de 16 gb), placa de vídeo dedicada NVIDIA Geforce 940MX 2 gb GDDR5 e armazenamento Hard Disk de 1 tb (5400 RPM).

Este equipamento foi utilizado como servidor do banco de dados, processamento do dashboard em Power BI e também na simulação do funcionamento como cliente final.

2.9 – Implementação dos datamarts/cubo

Neste trabalho, a tabela fato apontará em cada linha registros com as principais informações relacionadas aos lançamentos de possíveis ocorrências e entrada de mercadoria na empresa através do volume de peças de um lote, sendo que cada registro está necessariamente vinculado a um código de fornecedor, um material (*part number*) e consequentemente uma *commodity*.

Na figura 8 apresentada abaixo é possível verificar os comandos programados em *DAX* para tratamento dos dados da tabela fato, tais como: importação da base de dados, determinação dos cabeçalhos, remoção de colunas indesejadas para análise, criação de colunas calculadas (trimestre do registro), conversão dos tipos de dados que são armazenados nas respectivas colunas e definição de um nome para os códigos das *commodities*.



Figura 8 - Tratamento da tabela fato utilizando a linguagem DAX

Fonte: Próprio Autor

Em adição a tabela fato foram criadas dimensões com intuito de agregar atributos qualitativos aos registros, tais como descrições do *part number* (dimensão Dim_Part_No), dados gerais da cadeia de fornecedores (Dim_RBSNO) e descrição das *commodities* vinculadas ao valor da meta fictícia criada (Dim_PQT).

Para possibilitar melhor exploração dos dados e determinação de uma linha cronológica, foi criada manualmente (utilizando a linguagem *DAX*) uma dimensão de data (*Dim_Date*), apresentada na figura 9 abaixo:

1	Dim_Date =
2	ADDCOLUMNS {
3	CALENDAR (DATE(2016;1;1); DATE(2020;12;31));
4	"DateAsInteger"; FORMAT ([Date]; "YYYYMMDD");
5	"Year"; YEAR ([Date]);
6	"MonthNumber"; FORMAT ([Date]; "MM");
7	"Month"; FORMAT ([Date]; "MM");
8	"YearMonthNumber"; FORMAT ([Date]; "YYYY/MM");
9	"YearMonthShort"; FORMAT ([Date]; "YYYY/mm");
10	"Month Short"; FORMAT ([Date]; "mmm");
11	"MonthNameLong"; FORMAT ([Date]; "mmmm");
12	"DayOfWeekNumber"; WEEKDAY ([Date]);
13	"DayOfWeek"; FORMAT ([Date]; "dddd");
14	"DayOfWeekShort"; FORMAT ([Date]; "ddd");
15	"Quarter"; "Q" & FORMAT ([Date]; "Q");
16	"YearQuarter"; FORMAT ([Date]; "YYYY") & " / Q" & FORMAT ([Date]; "Q")
17	}

Date	DateAsInteger	Year	MonthNumber	YearMonthNumber	YearMonthShort	MonthNameLong	DayOfWeekNumber	DayOfWeek	DayOfWeekShort	Quarter	YearQuarter
01/07/2016 00:00:00	20160701	2016	07	2016/07	2016/Jul	July	6	Friday	Fri	Q3	2016/Q3
02/07/2016 00:00:00	20160702	2016	07	2016/07	2016/Jul	July	7	Saturday	Sat	Q3	2016/Q3
03/07/2016 00:00:00	20160703	2016	07	2016/07	2016/Jul	July	1	Sunday	Sun	Q3	2016/Q3
04/07/2016 00:00:00	20160704	2016	07	2016/07	2016/Jul	July	2	Monday	Mon	Q3	2016/Q3

Figura 9 - Dimensão de data utilizando a linguagem DAX

Fonte: Próprio Autor

Visto que foram feitas algumas visualizações customizadas, foi necessária criação de algumas tabelas ponte (Bridge_Semester e Bridge PQT Department) para estabelecer conexão entre a tabela fato e as respectivas dimensões, preservando o modelo de dados escolhido.

Também atendendo algumas visualizações customizadas, foi necessário criar manualmente uma base de dados em formato de planilha Excel contendo os valores atingidos no ano vigente em comparação as metas criadas (tabela Real x Atingido) por *commodity*.

Durante o desenvolvimento, foram criados cálculos (*measures*) que tem como objetivo realizar operações lógico-matemáticas com os dados dispostos no modelo, tais como valores anualizados de ocorrência/indicadores da Qualidade, semáforo de metas, ranqueamento dos fornecedores mais críticos e demais indicadores que serão apresentados no capítulo 3. Na figura 10 abaixo é possível visualizar todas as *measures* utilizadas na consolidação deste trabalho:

```

1 Abs.Inc 18 = CALCULATE(SUM(Fact_Table[Total_incidents]);FILTER(Dim_Date; Dim_Date[Year]=2018))
2 Abs.Inc 19 = CALCULATE(SUM(Fact_Table[Total_incidents]);FILTER(Dim_Date; Dim_Date[Year]=2019))
3 Arrows Index = IF([Sep] = [Aug]; 1; IF([Sep] > [Aug]; 2; IF([Sep] < [Aug]; 3;0)))
4 Aug = CALCULATE(SUM(Fact_Table[Total_incidents]);FILTER(Dim_Date; Dim_Date[Year]=2019);FILTER(Dim_Date; Dim_Date[Month Short]="Aug"))
5 Claimed Parts = SUM(Fact_Table[Claimed_Parts])
6 CRI_Abs = SUM(Fact_Table[V0_VF])
7 CRI/bn = DIVIDE(SUM(Fact_Table[V0_VF]);SUM(Fact_Table[Received_Parts]))*1E9
8 CRI/bn_2018 = CALCULATE(SUM(Real_x_Atingido[CRI/bn_2]);FILTER(Real_x_Atingido; Real_x_Atingido[Label]="Ac. 2018"))
9 CRI/bn_YTD_2019 = CALCULATE(SUM(Real_x_Atingido[CRI/bn_2]);FILTER(Real_x_Atingido; Real_x_Atingido[Label]="Ac. YTD 09.2019"))
10 Inc/bn = DIVIDE(SUM(Fact_Table[Total_incidents]);SUM(Fact_Table[Received_Parts]))*1E9
11 Inc/bn_18 = CALCULATE(DIVIDE(SUM(Fact_Table[Total_incidents]);SUM(Fact_Table[Received_Parts]))*1E9;FILTER(Dim_Date; Dim_Date[Year]=2018))
12 Inc/bn_19 = CALCULATE(DIVIDE(SUM(Fact_Table[Total_incidents]);SUM(Fact_Table[Received_Parts]))*1E9;FILTER(Dim_Date; Dim_Date[Year]=2019))
13 Inc/bn_2018 = CALCULATE(SUM(Real_x_Atingido[Inc/bn_1]);FILTER(Real_x_Atingido; Real_x_Atingido[Label]="Ac. 2018"))
14 Inc/bn_YTD_2019 = CALCULATE(AVERAGE(Real_x_Atingido[Inc/bn_1]);FILTER(Real_x_Atingido; Real_x_Atingido[Label]="Ac. YTD 09.2019"))
15 Incidents TOP N = CALCULATE([Total Incidents]; TOPN('Select Top N'[Selected N]; Fact_Table[Total_incidents]; DESC; 'Fact_Table'[Código fornecedor]))
16 Jul = CALCULATE(SUM(Fact_Table[Total_incidents]);FILTER(Dim_Date; Dim_Date[Year]=2019);FILTER(Dim_Date; Dim_Date[Month Short]="Jul"))
17 Rank 2018 = CALCULATE(RANKX(ALLSELECTED(Fact_Table[Código fornecedor]); [Abs.Inc 18];DESC;Skip); ALL('Nome Fantasia'[Fantasy Name Raw]))
18 Rank 2019 = CALCULATE(RANKX(ALLSELECTED(Fact_Table[Código fornecedor]); [Abs.Inc 19];DESC;Skip); ALL('Nome Fantasia'[Fantasy Name Raw]))
19 Rank Incidents = RANKX(ALL(Fact_Table[Código fornecedor]); [Total Incidents])
20 Rank Received Parts = RANKX(ALL(Fact_Table[Código fornecedor]); [Received Parts])
21 Rank_RBSNO_Within = IF([Rank Incidents] <= 'Select Top N'[Selected N];1;0)
22 Rank_Received_Parts_Within = IF([Rank Received Parts] <= 'Select Top N'[Selected N];1;0)
23 Received Parts = CALCULATE(DIVIDE(SUM(Fact_Table[Received_Parts]);1))
24 Sep = CALCULATE(SUM(Fact_Table[Total_incidents]);FILTER(Dim_Date; Dim_Date[Year]=2019);FILTER(Dim_Date; Dim_Date[Month Short]="Sep"))
25 Sign = SWITCH( TRUE(); ISBLANK('Key Measures'[Aug]);BLANK();
26 ISBLANK([Sep]);BLANK();
27 [Sep] = [Aug]; "-";
28 [Sep] > [Aug];UNICHAR(129093);
29 [Sep] < [Aug];UNICHAR(129095);
30 BLANK())
31 TL_CRI/bn_Green = IF([CRI/bn_YTD_2019]<[TL_Target CRI/bn];1;0)
32 TL_CRI/bn_GreenNOT = 1 - [TL_CRI/bn_Green]
33 TL_CRI/bn_Red = IF(AND([CRI/bn_YTD_2019]>[TL_Target CRI/bn];[CRI/bn_YTD_2019]>[CRI/bn_2018]);1;0)
34 TL_CRI/bn_RedNOT = 1-[TL_CRI/bn_Red]
35 TL_CRI/bn_Yellow = IF(AND([CRI/bn_YTD_2019]>[TL_Target CRI/bn];[CRI/bn_YTD_2019]<[CRI/bn_2018]);1;0)
36 TL_CRI/bn_YellowNOT = 1-[TL_CRI/bn_Yellow]
37 tl_inc/bn_green = if([inc/bn_ytd_2019]<[tl_target inc/bn];1;0)
38 tl_inc/bn_greennot = 1 - [tl_inc/bn_green]
39 tl_inc/bn_red = if(AND([inc/bn_ytd_2019]>[tl_target inc/bn];[inc/bn_ytd_2019]>[inc/bn_2018]);1;0)
40 tl_inc/bn_rednot = 1-[tl_inc/bn_red]
41 tl_inc/bn_yellow = if(AND([inc/bn_ytd_2019]>[tl_target inc/bn];[inc/bn_ytd_2019]<[inc/bn_2018]);1;0)
42 tl_inc/bn_yellownot = 1-[tl_inc/bn_yellow]
43 Total Incidents = CALCULATE(SUM(Fact_Table[Total_incidents]))

```

Figura 10 - Criação das *measures* em linguagem DAX

Fonte: Próprio Autor

CAPÍTULO III

Análise dos Dados e Resultados

Neste último capítulo será apresentado todo o processo de execução do projeto com as informações geradas através dos *Dashboards*.

Para esta apresentação, foram utilizados recursos que permitissem com que pessoas com daltonismo e monocromacia interajam com mais facilidade com o relatório final, implementando assim diferenças visuais entre séries de dados como diversificação do tipo de linha nas séries de dados, tabulação com fontes maiores e utilização de uma palheta de cores em alto contraste amigável para pessoas com estas características.

A paleta de cores escolhida para criação deste *Dashboard* é amigável a todos os tipos de daltonismo. Na figura 11 é possível visualizar as diferentes aberrações cromáticas em relação a captação de cores de um indivíduo sem deficiência visual (primeiro círculo à esquerda, *Regular Vision*):

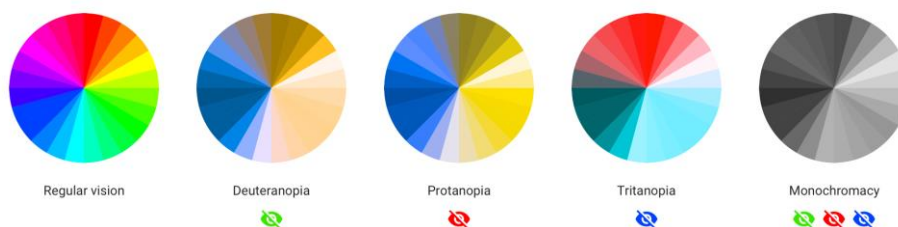


Figura 11 - Diferenças entre visão sem (esquerda) e com deficiência cromática (direita)

Fonte: TUCHKOV (2018)

De acordo com TUFTE (1998), abaixo são apresentados os princípios para aprimorar a visualização dos gráficos e informação dos dados apresentados, sendo:

- Apresentação enxuta: trata-se da fundamentação base dos indicadores limitando apenas ao que deve ser apresentado;
- Maximizar a diferença de contraste da paleta de cores *Red Green & Blue* (código universal de cores, *RGB*): quanto maior diferença entre cores de um gráfico, melhor. Cada pequeno detalhe na imagem têm um propósito distinto e assim é possível ter um destaque melhor das séries de dados;

- c. Redução da quantidade de cores *RGB*: Uma menor repetição gráfica e consequente redundância das cores apresentadas pode trazer uma mensagem mais concisa. Este princípio não se aplicaria caso a diversificação das cores tenha uma finalidade distintamente importante.

O *Power BI* também possui recursos nativos da ferramenta para acessibilidade dos relatórios, tais como a utilização de tabulação para focar em vários objetos em uma página, utilização do mouse apontando na série de dados para apresentar o detalhamento em tópicos e habilitação do modo alto contraste de cores.

3 Gráfico comparativo de volume de peças vs. ocorrências

O gráfico da figura 12 apresenta dentro da linha de tempo entre 2016 até 2019 (apresentada de forma anual, trimestral e mensal simultaneamente) o acompanhamento de todas as ocorrências mensais de um determinado fornecedor (barras roxas), agrupado por uma determinada *commodity* e uma relação entre o volume de fornecimento de peças mês-a-mês (representado pela linha rosa em escala de milhões) com a média semestral de ocorrências (representado pela linha verde pontilhada em valores absolutos):

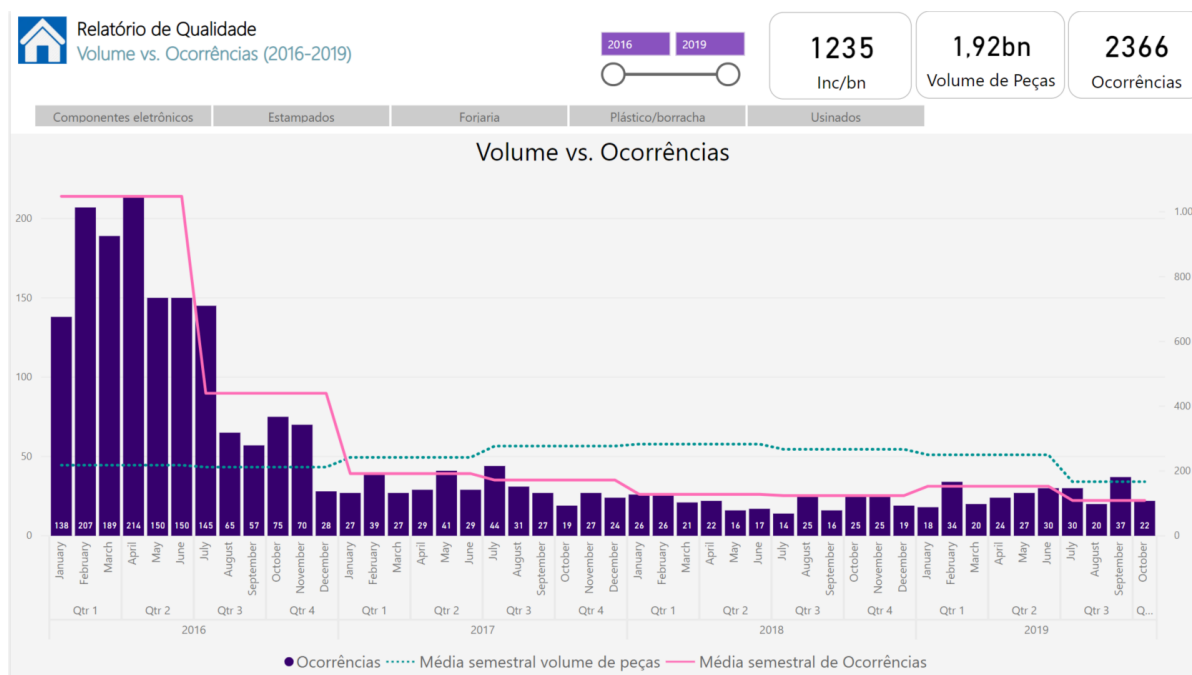


Figura 12 - Volume de peças vs. ocorrências

Fonte: Próprio Autor

Para que possamos identificar quais foram os principais fornecedores que contribuíram com o resultado do mês de análise, ao apontar o mouse para uma barra de um mês específico é possível ver a lista dos fornecedores através de um gráfico suspenso, conforme apresentado na figura 13:

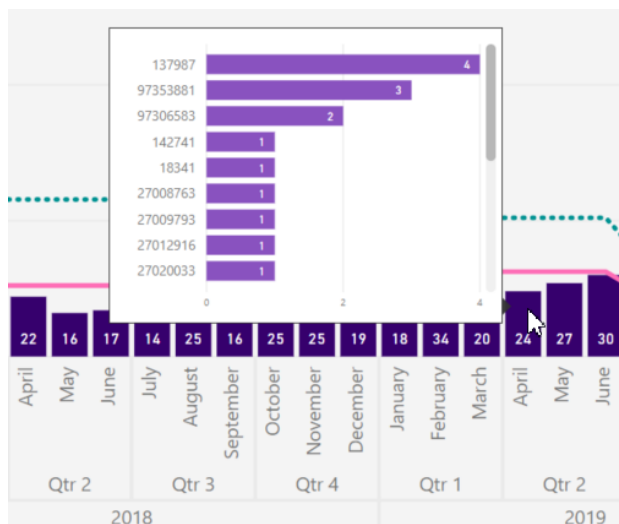


Figura 13 - Gráfico suspenso com lista de fornecedores com ocorrências

Fonte: Próprio Autor

Este indicador nos permite avaliar a evolução de um determinado fornecedor ou uma determinada commodity em relação ao volume relacionado a melhoria da Qualidade através da distância vertical entre as duas linhas, visto que quanto maior for esta distância, maior volume foi fornecido, menores ocorrências foram registradas e vice-versa.

3.1 Histórico de incidentes críticos

O gráfico da figura 14 apresenta dentro da linha de tempo entre 2016 até 2019 (apresentada de forma anual, trimestral e mensal simultaneamente) o acompanhamento de todas as ocorrências críticas mensais consolidadas (barras roxas empilhadas com azuis), agrupado por uma determinada *commodity*:

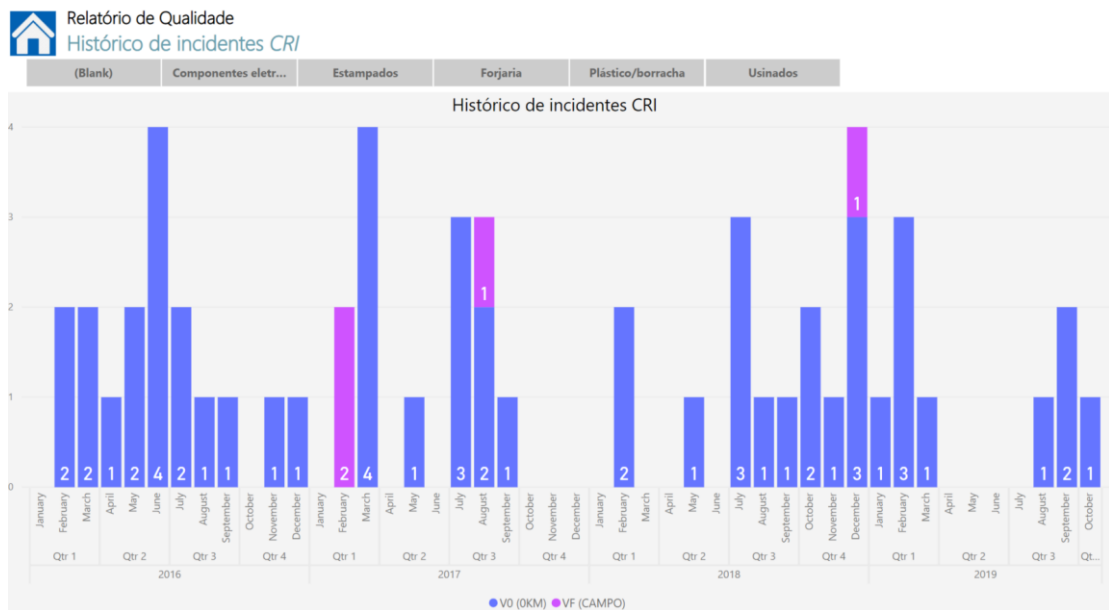


Figura 14 - Histórico de incidentes CRI

Fonte: Próprio Autor

Para esta visualização também é possível visualizar um gráfico suspenso com a lista dos fornecedores que contribuíram com o resultado do mês, apontando o mouse para uma barra de um mês específico. A figura 15 demonstra um exemplo com o mês de Dezembro de 2018:



Figura 15 - Gráfico suspenso com lista de fornecedores com ocorrências críticas

Fonte: Próprio Autor

Este indicador nos permite avaliar a frequência e quantidade de ocorrências críticas entre períodos distintos, trazendo assim uma visualização concisa dos casos mais impactantes para a empresa avaliada.

3.2 Diagnóstico de análise dos fornecedores críticos

Na figura 16 é apresentado em quatro quadrantes indicadores que foram utilizados para determinar a Qualidade de uma *commodity* ou fornecedor específico e qual é seu comportamento diante dos demais, trazendo assim em uma mesma visualização indicadores macro e micro com seus status atuais:



Figura 16 - Visualização de diagnóstico dos fornecedores críticos

Fonte: Próprio Autor

3.3 Gráfico de incidentes por bilhão mensal (Inc/bn)

Na figura 17 podemos visualizar o gráfico do primeiro quadrante (superior, lado esquerdo), que apresenta o acompanhamento mensal das ocorrências (linha azul clara) em

conjunto ao indicador de incidentes por bilhão (barras), cuja utilização será explicada em sequência. Também é apresentado neste gráfico uma meta (linha vermelha) apenas para contextualizar a criação e utilização do semáforo de metas.

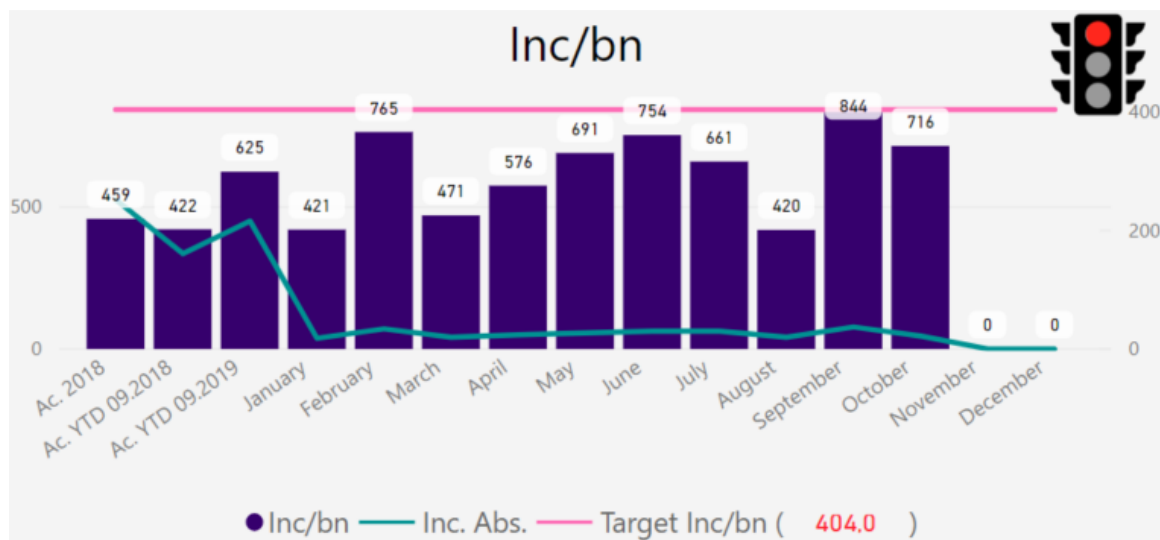


Figura 17 - Acompanhamento de incidentes por bilhão

Fonte: Próprio Autor

Foi determinada nesta visualização que o simples apontamento das ocorrências (ou incidentes) dos fornecedores não é um indicador robusto para ser avaliado isoladamente, portanto o volume de peças fornecidas também há de ser considerado. Para isto, foi definida a fórmula matemática abaixo:

$$\frac{\Sigma \text{ocorrências}}{\Sigma \text{volume de peças fornecidas}} \times 10^9$$

Na figura 18 encontra-se disponível a imagem do semáforo de metas. A determinação deste tem objetivo de uma melhor representação visual do acompanhamento de Janeiro do ano vigente até o mês atual (*Year to Date*, ou *YTD*) em relação à meta estipulada para o ano, representando a cor verde caso o *YTD* esteja abaixo da meta, amarelo caso o indicador esteja acima da meta, porém com redução em relação ao ano passado (*past year*). Finalmente, o vermelho representa um acréscimo tanto em relação à meta quanto ao ano passado.

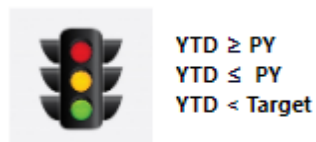


Figura 18 - Semáforo de metas

Fonte: Próprio Autor

Para esta visualização foram incluídas além do acompanhamento mensal três colunas que apresentam o valor acumulado do último ano, *YTD* do último ano e *YTD* do ano vigente para que desta forma seja possível visualizar os valores que representam a lógica visual do semáforo de metas.

3.4 Gráfico de incidentes críticos por bilhão mensal (CRI/bn)

Na figura 19 podemos visualizar o gráfico do segundo quadrante (inferior, lado esquerdo), que apresenta o acompanhamento mensal das ocorrências (linha azul clara) em conjunto ao indicador de incidentes críticos por bilhão (barras). Também é apresentado neste gráfico uma meta (linha vermelha) para contextualizar a criação e utilização do semáforo de metas.

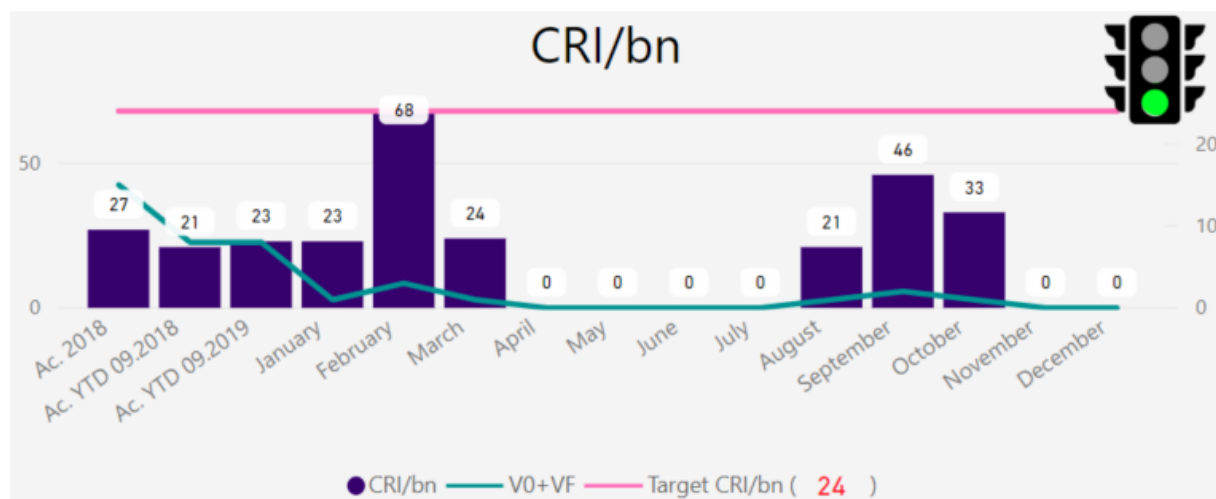


Figura 19 - Acompanhamento de incidentes críticos por bilhão

Fonte: Próprio Autor

Pela mesma justificativa do tópico anterior, também foi determinada uma equação similar para calcular os incidentes críticos por bilhão:

$$\frac{\Sigma \text{ocorrências críticas}}{\Sigma \text{volume de peças fornecidas}} \times 10^9$$

Foram determinadas como falhas críticas apenas aquelas que já estão vinculadas ao processo final de montagem do automóvel, seja um novo carro zero quilômetro (V0) ou já em vendido e sendo utilizado em campo (VF).

3.5 Fornecedores de mais alto impacto

Na figura 20 podemos visualizar o gráfico do terceiro quadrante (superior, lado direito), que apresenta em formato de tabela ranqueada os fornecedores de mais alto impacto, ordenados de forma decrescente pelas ocorrências em valor absoluto (coluna *Abs.Inc19*):

Origem	Rank 2018	Rank 2019	Código Fornecedor	Inc/bn 2018	Inc/bn 2019	Abs.Inc 2018	Abs.Inc 2019	Jul	Aug	Sep	Sign
	1	1	137987	636	1124	19	26	3	0	0	-
	7	2	97306583	3190	8525	7	17	0	2	6	↑
	14	3	636355	263	885	5	14	2	2	2	-
	12	4	27033985	840	1711	6	11	0	0	1	↑
	7	5	27033883	948	1320	7	9	3	0	0	-
	2	5	500010564 5	1518	2052	8	9	0	0	3	↑
	41	7	27006718	376	1718	2	8	0	1	0	↓
	14	7	27012913	300	657	5	8	1	0	1	↑

Figura 20 - Tabela com fornecedores de mais alto impacto

Fonte: Próprio Autor

Para esta visualização, foram incluídas as seguintes informações (da esquerda para a direita):

- Origem: Representa a nacionalidade do fornecedor pela respectiva bandeira do país;
- Rank 2018: A posição daquele fornecedor no ano passado de acordo com a quantidade de ocorrências absolutas;

- c) Rank 2019: A posição daquele fornecedor no ano vigente de acordo com a quantidade de ocorrências absolutas;
- d) Código fornecedor: Nome mascarado por código do fornecedor;
- e) Inc/bn18: indicador de ocorrências por bilhão do ano passado;
- f) Inc/bn19: indicador de ocorrências por bilhão do ano vigente;
- g) Abs.Inc18: Ocorrências absolutas do ano passado;
- h) Abs.Inc19: Ocorrências absolutas do ano vigente;
- i) Jul, Ago e Sep: Resumo das ocorrências absolutas nos últimos três meses;
- j) Sign: Indicador visual se o fornecedor obteve mais ou menos ocorrências absolutas em relação ao último mês, representadas pela seta vermelha ou verde, respectivamente.

3.6 Piores fornecedores com ocorrências críticas

Na figura 21 podemos visualizar o gráfico do quarto quadrante (inferior, lado direito), que apresenta em formato de barras horizontais os fornecedores que pontuaram ocorrências críticas, apresentadas abaixo de forma ordenada decrescente:

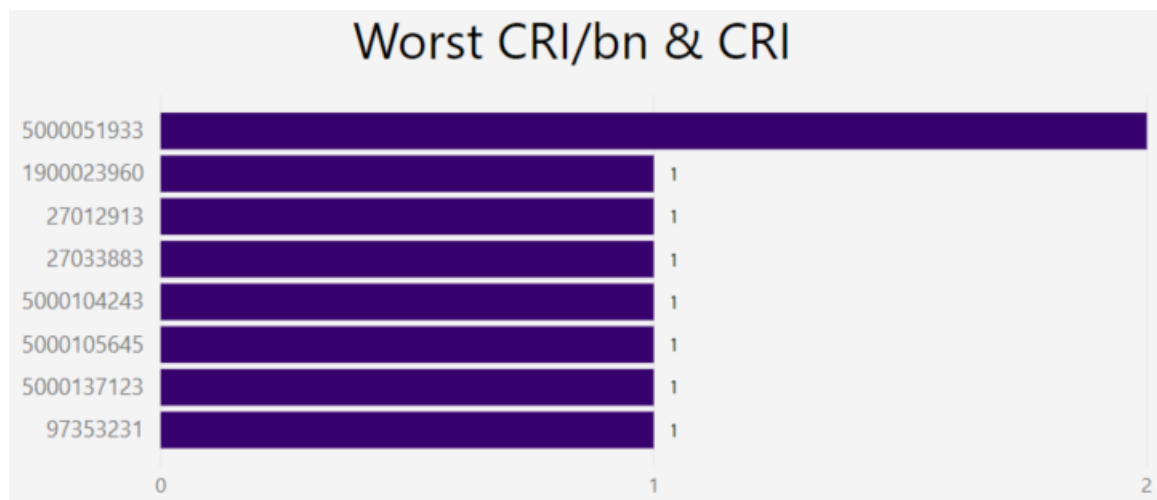


Figura 21 - Piores fornecedores com ocorrências críticas absolutas

Fonte: Próprio Autor

3.7 Fornecedores com maior fornecimento de peças (volume treemap)

Na figura 22 foi feita uma representação visual de uma cadeia de fornecedores com maior fornecimento de peças (quadriláteros em verde, representados na escala em milhões de

unidades) dentro de um período especificado pelo usuário, disponível em filtros na parte superior ao centro:

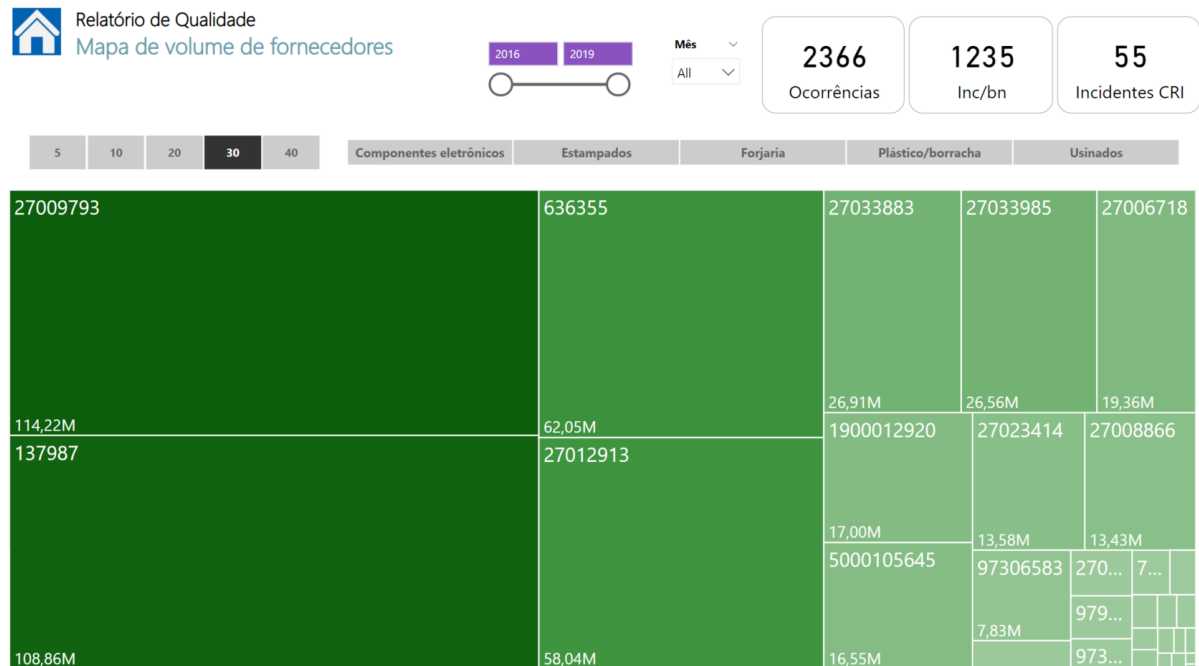


Figura 22 - Mapa de fornecedores com maior fornecimento de peças

Fonte: Próprio Autor

Este gráfico foi estruturado para que sejam apresentados de forma dinâmica a quantidade de fornecedores que serão dispostos abaixo (5, 10, 20, 30 ou 40 maiores fornecimentos de peças), disponível em filtros no canto superior esquerdo.

A representação torna-se prática visto que quanto maior for o tamanho da forma, maior é o fornecimento de peças, sendo possível avaliar quais são os fornecedores que trazem maior volume dentro de uma determinada *commodity* (disponível em filtros no canto superior direito).

3.8 Fornecedores com maiores ocorrências (*incidents treemap*)

Na figura 23 também foi feita uma representação visual de uma cadeia de fornecedores, porém analisando isoladamente os fornecedores com maiores ocorrências (quadriláteros em verde, representados na escala em milhões de unidades) dentro de um período especificado pelo usuário, disponível em filtros na parte superior ao centro:

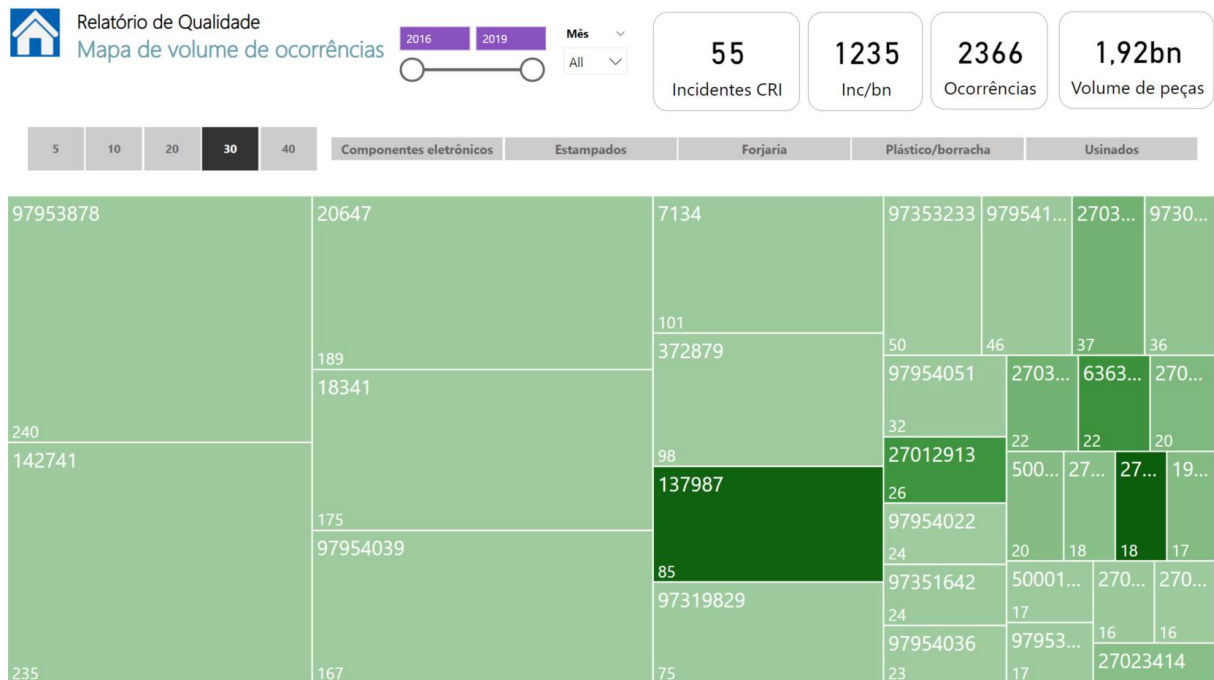


Figura 23 - Mapa de fornecedores com maiores ocorrências

Fonte: Próprio Autor

As disposições deste gráfico são semelhantes ao apresentado no tópico 3, porém podemos visualizar dentro de uma determinada *commodity* quais fornecedores trouxeram maior impacto em ocorrências absolutas.

Esta visualização é essencial quanto relacionada ao *volume treemap*, uma vez que podemos determinar se o fornecedor tem alta incidência de ocorrências, mas não tem um volume de fornecimento significativo, sendo assim um fornecedor de alto risco. Da mesma forma é justificável que um fornecedor que tenha uma alta incidência de ocorrências, porém com um alto volume de peças torna-se estratégico.

CONSIDERAÇÕES FINAIS

Após a apresentação dos gráficos, é possível verificar a evolução de uma determinada *commodity* a partir de uma linha de tempo através do gráfico de volume versus ocorrências. Ao analisar o gráfico geral, notamos uma significativa melhoria dos resultados visto que a média de ocorrências semestral abaixou drasticamente enquanto a de volume de peças recebidas aumentou, reduzindo o valor anualizado de *Inc/bn*.

A determinação dos indicadores quali-quantitativos de *Inc/bn* e *CRI/bn* foram cruciais para analisar o fornecedor de maneira mais crítica, visto que se analisarmos isoladamente o aumento/diminuição da quantidade de falhas absolutas não implica melhoria ou piora na qualidade do fornecedor, uma vez que o volume de peças fornecidas pode ter variado exponencialmente.

Desta forma podemos ver quais fornecedores possuem menores indicadores e que fornecem o mesmo material que outros já existentes, dando oportunidade de transferir a compra de produtos em quantidade maior àqueles que possuem qualidade superior.

Utilizando a plataforma, vemos que as *commodities* de estampados e plástico/borracha contribuíram para a melhoria de resultado, enquanto houve uma piora na de componentes elétricos, com significativo aumento de ocorrências absolutas.

Podemos verificar a partir da aba de visualização de diagnóstico dos fornecedores críticos aqueles que apresentam pior desempenho entre o ano vigente e ano passado, calculados de forma anualizada o *Inc/bn* e a quantidade de ocorrências absolutas. Podemos ver em lista ordenada de forma descendente quais são os principais contribuintes para que o departamento possa tomar decisões de qualidade corretiva ou preventiva contra o fornecedor afim de mitigar riscos.

Utilizando o semáforo de metas para os indicadores mês-a-mês, vemos que houve uma piora de resultado tanto em relação a meta quanto ao ano passado, representado pela cor vermelha. Em contrapartida, o indicador *CRI/bn* nos indica que as falhas críticas estão controladas e dentro da meta estipulada.

Para os incidentes mais críticos, podemos verificar na aba de histórico de incidentes *CRI* anualmente que a quantidade absoluta de falhas críticas está semelhante aos anos anteriores, porém com redução significativa no segundo (apresentando zero falhas) e

terceiro trimestre (três falhas). Com suporte do gráfico suspenso e visualizando anualmente, podemos ver também quem são os principais fornecedores contribuintes ao resultado e determinar medidas corretivas para evitar recorrências nestas falhas, visto que trazem maior prejuízo financeiro à empresa.

O mapa de volume aponta quais fornecedores são mais estratégicos e contribuem com maior fornecimento de peças e conseqüentemente sua contribuição para a *commodity*. Em contrapartida, podemos verificar pelo mapa de incidentes quais são os maiores contribuintes para a queda de resultado na *commodity*.

Desta forma, é possível concluir que a aplicação de ferramentas de *BI* na consolidação de uma plataforma é essencial para a tomada de decisão, visto que a análise de comportamentos passados, criação de padrões e agrupamento de informações podem auxiliar no melhor entendimento de um processo departamental em uma empresa.

REFERÊNCIAS

AFFELDT, F. S., SILVEIRA, F. C. da S. VANTI, A. A. **Análise Trinômio Alinhamento Estratégico x Business Intelligence x Controladoria Estratégica.** In: XIII Congresso Brasileiro de Custos 12., 2006, Belo Horizonte.

ANZANELLO, C. M. **OLAP conceitos e utilização.** Rio Grande do Sul, 2007. Disponível em: <<https://www.klebermota.eti.br/wp-content/OLAP.pdf>>

BRANSKI, R. M., FRANCO, R. A. C., JUNIOR, O. F. L. **Metodologia de Estudos de Casos Aplicada à Logística.** São Paulo, 2013. Disponível em: <<http://www.lalt.fec.unicamp.br/scriba/files/escrita%20portugues/ANPET%20-%20METODOLOGIA%20DE%20ESTUDO%20DE%20CASO%20-%20COM%20AUTORIA%20-%20VF%2023-10.pdf>>

FALCAO, M. F. **The Information Visualization Thematic, Head to Head with Color Blindness.** Disponível em: <https://www.researchgate.net/publication/313657827_The_Information_Visualization_Thematic_Head_to_Head_with_Color_Blindness>

FERREIRA, J. *et al.* **O processo ETL em sistemas Data Warehouse.** In: INForum, 2010, Minho.

FORTULAN, Marcos Roberto. **O uso de business intelligence para gerar indicadores de desempenho no chão-de-fábrica: uma proposta de aplicação em uma empresa de manufatura.** 2006. Tese (Doutorado em Manufatura) - Escola de Engenharia de São Carlos, University of São Paulo, São Carlos, 2006. Data de consulta: 29 de Maio de 2018. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/18/18145/tde-11062006-185813/en.php>>

GIL., A. C. **Como Elaborar Projetos de Pesquisa** 4 ed. São Paulo. Atlas, 2007.

GEEKS FOR GEEKS, **Data Mining**, [s.d.] . Disponível em: <<https://www.geeksforgeeks.org/data-mining/>> Acesso em: 30 ago. 2019.

GENHARDT, T. E., SILVEIRA, D. T. **Métodos de Pesquisa**. Disponível em: <<http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf>>

IVAN TUCHKOV, **Color blindness: how to design an accessible user interface**, 22 ago. 2017. Disponível em: <<https://uxdesign.cc/color-blindness-in-user-interfaces-66c27331b858>> Acesso em: 30 ago. 2019.

KEKMCZINSKI, A. *et al.* **Como Obter Vantagem Competitiva Utilizando Business Intelligence?**. Revista Produção On Line, Santa Catarina, v. 3, n. 3, 2003.

MCKINSEY, **Big data: The next frontier for innovation, competition, and productivity**, 3 nov. 2019. Disponível em: <<https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>> Acesso em: 30 ago. 2019.

MICROSOFT, **Introdução a dashboards para designers do Power BI**, 18 set. 2019. Disponível em: <<https://docs.microsoft.com/pt-br/power-bi/service-dashboards>> Acesso em: 30 ago. 2019.

PANOPLY, **3 Ways to Build An ETL Process** [s.d.]. Disponível em: <<https://panoply.io/data-warehouse-guide/3-ways-to-build-an-etl-process/>>. Acesso em: 30 ago. 2019.

PICKELL, D. **What is Big Data? A Complete Guide**, 22 ago. 2018. Disponível em: <<https://learn.g2.com/big-data>> Acesso em: 30 ago. 2019.

RAFAEL PITON, **Data Warehouse – O Que É Star Schema?**, 23 maio 2017. Disponível em: <<https://rafaelpiton.com.br/blog/data-warehouse-star-schema/>> Acesso em: 30 ago. 2019.

REGINATO, L. NASCIMENTO, A. M. **Um Estudo de Caso Envolvendo Business Intelligence Como Instrumento de Apoio à Controladoria**. Revista Contabilidade & Finanças. 2007, 18. Data de consulta: 29 de Maio de 2018. Disponível em: <<http://www.redalyc.org/articulo.oa?id=257119526007>>

SILVEIRA, F. C. da S. **Construção de Modelo de Business Intelligence Para a Controladoria Evidenciar Informações Estratégicas: O Caso do SESI**. 2007. Tese (Mestrado em Ciências Contábeis) – Universidade do Vale do Rio dos Sinos, São Leopoldo, 2007. Data de consulta: 29 e Maio de 2018. Disponível em: <<http://www.repositorio.jesuita.org.br/handle/UNISINOS/2813>>

SILVA, R. A., SILVA, F. C. A., GOMES, C. F. S. **O Uso do Business Intelligence em Sistema de Apoio à Tomada de Decisão Estratégica.** Revista Gestão, Inovação e Tecnologias. 2016, 6. Data da consulta: 29 de Maio de 2018. Acesso disponível em: <http://www.revistageintec.net/index.php/revista/article/view/726/630>

TECH DIFFERENCES, **Difference Between Star and Snowflake Schema**, 16 dez. 2017. Disponível em: <<https://techdifferences.com/difference-between-star-and-snowflake-schema.html>> Acesso em: 30 ago. 2019.

TUFTE, E. R. **The visual display of quantitative information.** 2. ed. Graphics press, 2001

ANEXOS

ANEXO I – FICHA DE AVALIAÇÃO DE ESTÁGIO

FICHA DE AVALIAÇÃO DE DESEMPENHO DO ESTAGIÁRIO

(a ser preenchida pelo Supervisor de Estágio da Empresa)

Nome do Aluno: _____
N.º RA: _____ Semestre: _____
Empresa: _____
Ramo de atividade: _____
Setor(es) : _____

Avaliação	Fraco	Regular	Bom	Ótimo
- Qualidade da atividade exercida: considerar a qualidade do atividade exercida, tendo em vista as condições oferecidas				
- Rapidez em execução: considerar o volume de atividades realizadas				
- Organização:				
- Sequência lógica de execução:				
- Facilidade de compreensão:				
- Frequência:				
- Dedicação: esforço revelado para aprender, a partir de indagações e dúvidas apresentadas				
- Conhecimento: nível de conhecimento apresentado e que tenha se revelado compatível com as tarefas propostas				
- Engenhosidade: capacidade de sugerir, projetar ou executar modificações que beneficiem a Empresa				
- Aprendizagem prática: aprendizagem dos serviços da Área Administrativa				
- Sociabilidade:				
- Cooperação: avaliação de atitudes que possam revelar boas qualidades de um profissional				
- Liderança:				

Observações Gerais:

Nome e Cargo do informante: _____

AVALIAÇÃO DA EMPRESA QUANTO AO CURSO:

1. A formação que o curso oferece ao aluno:

☐ ultrapassa as exigências☐ satisfaz as exigências☐ não satisfaz as exigências. Neste caso, em que aspectos o curso deve dar melhor preparação?

Exponha seu parecer:

2. O curso responde às necessidades atuais da empresa?

☐ Sim ☐ Não - Por que?

Exponha suas sugestões:

3. Resuma quaisquer observações e opiniões que julgar necessárias para o aprimoramento da formação profissional dos nossos alunos:

(campo a ser preenchido pelo coordenador de estágios da Fatec-Id)

Nome: _____

Síntese das avaliações _____

Observações: _____

Avaliação do Estágio: Aprovado ☐ Reprovado ☐ Nota:

Data: ____/____/____ Avaliado pelo

Prof: _____

Assinatura do Coordenador de Estágio do curso