

**CORPO DE BOMBEIROS MILITAR DE SANTA CATARINA
CENTRO DE ENSINO BOMBEIRO MILITAR
ACADEMIA BOMBEIRO MILITAR**

MIGUEL MORAES GOMES

**MODELAGEM E IMPLEMENTAÇÃO DE UM DATA MART
PARA PERÍCIAS DE INCÊNDIO DO CBMSC: ESTUDO DE CASO**

**FLORIANÓPOLIS
2019**

MIGUEL MORAES GOMES

**MODELAGEM E IMPLEMENTAÇÃO DE UM DATA MART
PARA PERÍCIAS DE INCÊNDIO DO CBMSC: ESTUDO DE CASO**

Monografia apresentada como pré-requisito para conclusão do Curso de Formação de Oficiais do Corpo de Bombeiros Militar de Santa Catarina.

Linha de Pesquisa: Atividade Operacional BM: Tecnologia e Atividade BM – Tecnologia da Informação e Comunicação, Sistemas de Informática Aplicada e Telecomunicações.

Orientador (a): Diego Felipe Marzarotto

**Florianópolis
2019**

Gomes, Miguel Moraes

Modelagem e Implementação de Um Data Mart para Perícias de Incêndio do CBMSC: Estudo de Caso / Miguel Moraes Gomes. -- Florianópolis : CEBM, 2019.

86 p.

Monografia (Curso de Formação de Oficiais) – Corpo de Bombeiros Militar de Santa Catarina, Centro de Ensino Bombeiro Militar, Curso de Formação de Oficiais, 2019.

Orientador: Maj BM Diego Felipe Marzarotto, Esp.

1. Sistema de Informação 2. Data Mart. 3. Gestão do conhecimento. I. Marzarotto, Diego Felipe. II. Título.

MIGUEL MORAES GOMES

ESTA FOLHA DEVERÁ SER ASSINADA PELA BANCA NO DIA DA APRESENTAÇÃO

**MODELAGEM E IMPLEMENTAÇÃO DE UM DATA MART
PARA PERÍCIAS DE INCÊNDIO DO CBMSC: ESTUDO DE CASO**

Monografia apresentada como pré-requisito para conclusão do Curso de Formação de Oficiais do Corpo de Bombeiros Militar de Santa Catarina.

Banca Examinadora:

Orientador(a):

**Esp. Diego Felipe Marzarotto
Major BM
CBMSC**

Membros:

**Esp. Gilvan Amorim da Silva
Cap BM
CBMSC**

**Esp. Marco Aurelio Lino Massarani Costa
1º Ten BM
CBMSC**

Florianópolis, 20 de novembro de 2019

Dedico este trabalho ao meu Deus, pela vida que me concedeu e à minha esposa, pelo carinho, compreensão e sacrifícios feitos em meu apoio.

AGRADECIMENTOS

Agradeço aos meus pais pela fé em mim.

A minha esposa e filhos, pelo apoio incondicional, pela paciência nos momentos de ausência e por todos os sacrifícios fizeram junto comigo.

Aos meus colegas do CPD/DiTI/CBMSC, em especial ao 3º Sgt BM Carlos Alberto Sousa chefe do CPD/DLF, que me deu todo o suporte e apoio tecnológico, sem o qual não seria possível a realização deste trabalho.

Ao Sd BM Bruno Depizzolatti, Coordenador do Projeto Analytics, da DTIC/SSP, o qual me possibilitou apresentar os resultados desse trabalho com uma das melhores ferramentas front-end de BI na atualidade.

Ao meu orientador, pelo conhecimento e experiência transmitidos durante a elaboração deste trabalho.

"Nada acontece duas vezes da mesma maneira."

(C.S. Lewis)

Miguel Moraes Gomes

**MODELAGEM E IMPLEMENTAÇÃO DE UM DATA MART
PARA PERÍCIAS DE INCÊNDIO DO CBMSC**

Projeto de Pesquisa apresentado para elaboração do Trabalho de Conclusão de Curso em cumprimento parcial às exigências do Curso de Formação de Oficiais, do Corpo de Bombeiros Militar de Santa Catarina.

Linha de Pesquisa: Atividade Operacional BM: Linha – Tecnologia e Atividade BM – Tecnologia da Informação e Comunicação, Sistemas de Informática Aplicada e Telecomunicações.

Orientador (a): Diego Felipe Marzarotto

**Florianópolis
Novembro 2019**

RESUMO

O Presente estudo focaliza a temática “Modelagem e Implementação de um DATA MART para perícias de incêndios do CBMSC: Estudo de caso”. Tem como objetivo Modelar e Implementar um Data Mart relacionado aos dados das investigações de incêndio para contribuir no processo de retroalimentação do ciclo operacional do CBMSC. Para tal empreendimento optou-se por uma revisão bibliográfica centrada em autores como LEMOS (2017); INMON (1997) e BARBOSA (2012) entre outros que abordam os assuntos que dão o embasamento teórico para o estudo de caso proposto. Para a pesquisa, foram utilizados dados do sistema de perícias de 2018 e 2019, assim como os dados de incêndios do sistema E193, ocorridos, entre os anos 2017 e 2019, na jurisdição da 1ª Região Bombeiro Militar (Litoral) do Estado de Santa Catarina. Como metodologia optou-se por uma abordagem de estudo de caso descritivo, pois “O pesquisador observa os fatos sob a óptica de alguém interno à organização”. (TEIXEIRA, 2011, p. 137). Sobre o Data Mart foram construídos diversos dashboards informativos e interativos, capazes de gerar, diversos tipos de informações, possibilitando retroalimentação do ciclo operacional de bombeiro a partir dos dados das investigações de incêndios da 1ªRBM.

Palavras-chaves: Investigação de Incêndios, Data Mart, Análise de dados.

LISTA DE ILUSTRAÇÕES

Quadro 1 - Tabela de comparação entre as características dos bancos de dados operacionais com um Data Warehouse.....	16
Figura 1 - Arquitetura genérica de DW.....	18
Quadro 2 - Comparativo das abordagens “Top-down” e “Bottom-up”.....	20
Figura 2 - Modelo dimensional de um negócio: visualização dos dados como um cubo..	21
Figura 3 - Star Schema.....	23
Figura 4 - Ciclo operacional de Bombeiro.....	26
Figura 5 - ETL Dimensão Vítima.....	39
Figura 6 - Carregamento de dimensões.....	41
Figura 7 - Processo ETL para o Fato Perícia.....	43
Figura 8 - Extração de dados para o Fato Perícia.....	44
Figura 9 - “ <i>If fiel value is null</i> ”.....	46
Figura 10 - Limpeza de dados inconsistentes.....	47
Figura 11 - Filtra coordenadas.....	48
Figura 12 - Transforma metadados.....	49
Figura 13 - “ <i>Dimension lookup/update</i> ”.....	50
Figura 14 - “ <i>Load</i> ” FTPerícia.....	50
Figura 15 - Processo ETL para o Fato Incêndio.....	51
Figura 16 - Extração de dados para o Fato Ocorrência.....	52
Figura 17 - Transformação do Fato Ocorrência.....	53
Figura 18 Figura 18 - Caixa de Diálogo do “Group By”.....	54
Figura 19 - “ <i>Load</i> ” do Fato Ocorrência.....	55
Figura 20 - “ <i>Load</i> ” do Fato Sumarizado Incêndio.....	56
Figura 21 - Esquema Estrela do Fato Perícia.....	57
Figura 22 - Dashboard 1.....	58
Figura 23 - Dashboard 2.....	59
Figura 24 - Dashboard 3.....	60

SUMÁRIO

1 INTRODUÇÃO.....	11
1.1 PROBLEMA.....	11
1.2 OBJETIVOS.....	12
1.2.1 Objetivo Geral.....	12
1.2.2 Objetivos específicos.....	12
1.3 JUSTIFICATIVA.....	13
1.4 DELIMITAÇÃO.....	14
2 DATA WAREHOUSE - DW.....	15
2.1 CARACTERÍSTICAS DE DATA WAREHOUSE.....	15
2.2 COMPONENTES DE UM DATA WAREHOUSE.....	17
2.3 METADADOS.....	18
2.4 DATA MART – DM.....	19
2.5 ESTRATÉGIAS ARQUITETÔNICAS PARA A IMPLEMENTAÇÃO DE DW.....	19
2.6 MODELO DIMENSIONAL.....	21
2.6.1 Tabela Fato (melhorar).....	21
2.6.2 Tabela Dimensão (melhorar).....	22
2.7 STAR SCHEMA.....	22
2.8 SURROGATE KEY.....	23
2.9 FERRAMENTAS PARA CONSTRUÇÃO E OPERAÇÃO DE UM DW/DM.....	23
3 O CICLO OPERACIONAL DE BOMBEIRO.....	25
3.1 INTRODUÇÃO.....	25
3.2 FASES DO CICLO OPERACIONAL DE BOMBEIRO.....	25
3.3 RETROALIMENTAÇÃO DO CICLO OPERACIONAL.....	27
3.4 GESTÃO DO CONHECIMENTO DAS INVESTIGAÇÕES NO CBMSC.....	28
4 DATA MART DAS INVESTIGAÇÕES DE INCÊNDIO.....	29
4.1 PROCEDIMENTOS METODOLÓGICOS.....	29
4.1.1 Tipo de pesquisa.....	29
4.1.2 Método.....	29
4.1.3 Técnica e instrumentos de pesquisa.....	29
4.1.4 Local da Pesquisa.....	30
4.1.5 Objetivo da Pesquisa.....	30
4.1.6 Requisitos e Procedimentos e para Modelagem e Implementação do DM.....	30
4.2 INFRAESTRUTURA TECNOLÓGICA.....	31
4.2.1 Infraestrutura no CBMSC.....	31
4.2.2 Secretaria de Estado da Segurança Pública - SSP.....	32
4.2.3 Acessos, segurança e dificuldades relacionados à infraestrutura.....	32
4.3 DETALHAMENTO DO SISTEMA OPERACIONAL DE ORIGEM.....	32
4.4 STAGING AREA.....	35
4.5 PLANEJAMENTO DO PROJETO.....	36
4.5.1 Dimensões.....	36
4.5.2 Fatos.....	37
4.5.3 Métricas.....	37
4.5.4 Preenchimento Das Dimensões.....	38
4.5.4.1 Dimensões preexistentes.....	38
4.5.4.2 Dimensão Vítima.....	39
4.5.4.3 Dimensões carregadas do Sistema E193.....	40
4.5.4.4 Dimensões extraídas do código do Sistema Operacional de Origem.....	41

4.5.5 Conclusão do Planejamento.....	41
5 O PROCESSO ETL.....	43
5.1 PROCESSO ETL PARA O FATO PERÍCIA.....	43
5.1.1 Extract.....	44
5.1.2 Transform.....	45
5.1.2.1 Primeiro passo.....	46
5.1.2.2 Segundo passo.....	46
5.1.2.3 Terceiro passo.....	48
5.1.2.4 Quarto passo.....	48
5.1.2.5 Quinto passo.....	49
5.1.3 Load FTPericia.....	50
5.3 PROCESSO ETL PARA O FATO OCORRÊNCIA.....	51
5.3.1 Extract.....	51
5.3.2 Transform.....	52
5.3.3 Load FTOccorrencias.....	54
5.3.4 Load FTSumOcorrencia.....	55
6 FERRAMENTAS DE ACESSO AOS DADOS.....	57
6.1 ARQUITETURA FRONT-END.....	57
6.2 CONFIGURAÇÃO DE DADOS MESTRES.....	58
6.3 GRÁFICOS E DASHBOARD.....	58
7 CONCLUSÃO.....	61
7.1 CONCLUSÕES.....	61
REFERÊNCIAS.....	65
APÊNDICE.....	69
DIMENSÕES EM CSV.....	69
ANEXO A.....	73
ANEXO B.....	83

1 INTRODUÇÃO

Na evolução das instituições públicas as Tecnologias da Informação e Comunicação tem gerados cada vez mais dados por meio de sistemas que atendam às necessidades institucionais. No governo de Santa Catarina, o CBMSC tem desenvolvido sistemas como o SIGAT, E-BOMBEIRO, E-193, dentre outras dezenas de aplicativos institucionais, os quais têm produzido um volume crescente de dados, que nas estimativas mais recentes fornecidas pela DiTI, são cerca de 150 mil registros de ocorrências por ano (E193), além de outros 300 mil relacionados a atividade de segurança contra incêndio (SIGAT), produzidos na última década, apenas considerado os dois principais sistemas da corporação.

1.1 PROBLEMA

É de conhecimento do autor, que já trabalhou na DiTI, que demandas por relatórios específicos solicitados pelos gestores da instituição são frequentes, entretanto não existem ferramentas apropriadas para extração de dados necessários para tomada de decisão na instituição, ficando a cargo das equipes da DiTI produzir os relatórios a partir de procedimentos não padronizados: sem controle de qualidade dos dados e do processo de extração da informação, soma-se o fato de ser um processo moroso para a necessidade do gestor. Dessa forma, essas abordagens prejudicam a agilidade no processo de tomada de decisão no nível diretivo, resultando no aumento de tempo, custo e decisões baseadas em intuição e/ou dados incompletos, insuficientes ou imprecisos.

Nesse contexto, está o sistema de perícia em incêndio e explosão, que armazena informações de todas as investigações de incêndios, para atender a fase Pericial do ciclo operacional de bombeiro. Esta fase do ciclo operacional destina-se a analisar os locais de sinistros de maneira a elucidar e identificar as causas e efeitos do sinistro, para retroalimentar as demais fases do ciclo operacional e contribuir no processo de planejamento e tomada de decisão dos gestores do CBMSC, entretanto conforme Timmermann e Acordi (2019) não ocorre uma interação e troca de informação entre as fases do ciclo operacional, sendo que cada uma encontra-se isolada das demais.

A dificuldade de obtenção de dados relevantes no processo decisório é tal que a segundo Lemos (2017) para mais de 70% dos comandantes de BBM a investigação não resultaram em propostas ao planejamento de comando. E nos BBM em que as investigações resultaram em propostas, estas vieram diretamente das equipes de investigação e não do

resultado das investigações de incêndios. Para qualquer relatório para auxiliar no processo de decisão, é necessário solicitar para a DiTI a qual virá a atender a demanda de modo não padronizado.

Assim percebe-se a ínfima influência dos resultados das investigações periciais de incêndio na produção de conhecimento para o planejamento institucional, Apesar desse ser, em teoria, um dos objetivos fins das investigações de incêndio.

Ao constatarem as dificuldades apontadas Modolon e Vidal (2019) indicam fazer um estudo sobre a viabilidade de implementação de uma solução de Business Intelligence (BI) para a gestão do conhecimento proveniente dos dados das investigações de incêndios realizadas.

Perante esta situação, e considerando que Data Mart é uma solução BI de menor complexidade, uma subdivisão de um Data Warehouse, essa pesquisa pretende responder a seguinte pergunta:

Quais são as contribuições que pode trazer uma modelagem e implementação de um Data Mart (relacionado aos dados das investigação de incêndio) para auxiliar no processo retroalimentação do ciclo operacional do bombeiro militar?

1.2 OBJETIVOS

Tendo em vista a pergunta de pesquisa, fica definido os seguintes objetivos:

1.2.1 Objetivo Geral

Modelar e Implementar um Data Mart relacionado aos dados das investigações de incêndio para contribuir no processo de retroalimentação do ciclo operacional do CBMSC.

1.2.2 Objetivos específicos

Para atender o objetivo geral, o trabalho se dividirá em etapas específicas, de acordo com os métodos de planejamento de um Data Mart – DM. Deste modo, são apresentados os seguintes objetivos específicos:

Apresentar um modelo conceitual de um Data Warehouse - DW;

Descrever o processo de implementação de um DM;

Apresentar os resultados que podem ser obtidos com a implementação do DM.

1.3 JUSTIFICATIVA

É de conhecimento do autor, que trabalhou vários anos na DiTI, que os sistemas de informação do corpo de bombeiro produzem um grande volume de dados que são armazenados nos bancos de dados da instituição, porém o acesso a esses dados se dá por relatórios padronizados desses sistemas ou por demandas de relatórios “*ad hoc*” que precisam ser atendidos por profissionais especializados em consultas aos bancos de dados desses sistemas, o que ocorre sem controle de qualidade e com frequência após uma elaborada e complexa programação própria da consulta, isso tudo sem um processo padronizado, que garanta a que informação solicitada seja de fato aquela entregue.

Barbosa (2012, p. 32) ainda destaca que

“A disponibilidade de informações adequadas é fundamental para o processo decisório. O excesso de informações, no entanto, é prejudicial ao processo, visto que, será necessária alguma seleção do que realmente tem importância para a situação em análise, é preciso que ocorra um gerenciamento dos dados e informações.”.

Num campo mais específico, que é o conhecimento produzido pelas perícias de incêndio Vidal (2007, p. 10) ensina que

“O estudo e a investigação dos incêndios alimentam as corporações de bombeiros em uma infundável avalanche de informações **que se bem processadas**, farão com que a corporação evolua e o reflexo direto virá para a comunidade”(grifo nosso).

Na mesma linha ainda podemos destacar o Plano Estratégico do CBMSC 2018-2030 que traz entre suas ações “Realizar a análise de todas as perícias de incêndio e explosões, **gerando relatório** que possa contribuir com as demais fases do ciclo operacional” (PLANO 2018, grifo nosso).

Lemos (2017, p. 31) também identificou a necessidade de estudo na implementação de uma solução para a gestão do conhecimento na atividade investigativa de incêndio com o intuito de contribuir para o estudo do conhecimento em organizações públicas, em especial de natureza militar, tendo em vista o crescente interesse destas na gestão da informação e do conhecimento.

No contexto da Tecnologia da Informação, encontramos o Data Mart: uma ferramenta dentro do contexto de inteligência de negócios (BI) e gestão do conhecimento, tecnologia voltada especialmente para gestão e análises de dados, em formato de relatórios padronizados, parametrizados ou “*ad hoc*”, de fácil acesso e fácil interpretação cujo objetivo pode ser traduzido em entregar a informação certa, no tempo certo para a pessoa certa, dentro de um certo escopo de conhecimento e/ou departamento de uma instituição, isso tudo a um baixo custo de implementação. Assim dentro do contexto do CBMSC pretendemos implementar um

Data Mart que viabilize o acesso fácil, eficaz e ágil das informações de investigações de incêndio por meio de relatórios “*ad hoc*” e padronizados dos dados disponíveis nos banco de dados do CBMSC.

1.4 DELIMITAÇÃO

Neste trabalho será desenvolvido um data mart - DM que atenda aos parâmetros relacionados às investigações periciais no CBMSC, definidos pela divisão de perícias da instituição, e levando em consideração apenas dados periciais de 2018 e 2019 e dados de ocorrências de incêndio de 2017, 2018, 2019 da 1ª Região Bombeiro Militar (Litoral). A limitação dos anos para os dados periciais se dá devido que foi em 2018 que o novo sistema de perícia entrou em operação, antes disso os dados eram armazenados no antigo sistema de perícia com parâmetros distintos. A estrutura estará pronta para receber dados de todos os 14 batalhões a limitação se dá pelo fato de que o autor entende ser suficiente para atingir os objetivos propostos. Esse trabalho não se objetiva automatizar a coleta de dados.

Não se pretende neste trabalho desenvolver o processo de automatização da atualização dos dados, tão pouco faz parte desse trabalho a gestão da qualidade de dados, podendo ser feito de forma pontual em caso de conveniência.

2 DATA WAREHOUSE - DW

Nas Tecnologias da Informação e Comunicação encontramos o conceito de *Business Intelligence* ou simplesmente “BI”. Segundo Barbieri (2011, p. 108), o conceito de BI pode ser entendido, numa de suas vertentes como diretamente relacionado ao apoio e ao subsídio aos processos de tomada de decisão, baseado em dados trabalhados especificamente para a busca de vantagens competitivas. BI pode ser entendido como o conjunto de soluções da tecnologia da informação que permite o acesso e análise aos dados em questão. Além disso, Barbieri (2011, p. 108) diz que “o BI deve ser entendido como o processo de desenvolvimento que objetiva implementar estruturas especiais de armazenamentos de informações como [...] Data Warehouse e Data Marts”. Já Piton, (2018) explica que “[...] BI é um processo, um conjunto de técnicas e conceitos. Trata de entregar **informação certa para a pessoa certa no tempo certo.**” (grifo nosso)

No conceito de Viera (2001, p. 17) Data Warehouse é um ambiente especializado e independente para o suporte à decisão e integração de sistemas originalmente não integrados tanto a nível de hardware, software básico, aplicação, rede de comunicação e processos.

Por fim temos a definição de Inmon (1997, p. 23), em que “Data Warehouse é uma coleção de dados orientados por assuntos, integrados, variáveis com o tempo e não voláteis, para oferecer suporte ao processo gerencial de tomada de decisão”.

2.1 CARACTERÍSTICAS DE DATA WAREHOUSE

Assim, usando-se a definição de Inmon (1997), extrai-se algumas das características de um Data Warehouse, as quais explana-se a seguir:

Orientado por Assunto: A organização dos dados é feita de acordo com o assunto, e contém somente as informações consideradas essenciais ao processo gerencial de tomada de decisão.

Integrado: Quando os dados estão armazenados em sistemas distintos, não há uma padronização total da codificação. Ao extrair esses dados para o DW, essa padroniza-se a codificação, através do tratamento dos dados assim o DW pode ser chamado de um sistema integrado.

Variável ao tempo: Um DW pode armazenar dados de 5 / 10 anos ou mais, para serem utilizados em comparações e previsões. Contudo, os dados existentes no DW correspondem a

algum momento específico e isso significa que o dado quando carregado recebe uma unidade de tempo na sua chave. Segundo Inmon (1997) esse tipo de armazenamento permite que os analistas façam análises de tendências, visualizando variações de uma determinada informação ao longo do tempo.

Não volátil: Os dados não são atualizados ou modificados em nenhum momento depois que entram no DW e são somente acessados para consulta.

Granularidade: É o nível de detalhe ou resumo dos dados de um Data Warehouse. Quanto maior for o nível de detalhe, menor será o nível de granularidade. O nível de granularidade afeta diretamente o volume de dados armazenados no Data Warehouse, e ao mesmo tempo o tipo de consulta que pode ser respondida.

Aqui cabe ressaltar que embora existam banco de dados - BD do sistemas desenvolvidos pela corporação, com dados que permitem a elaboração relatórios. Esses DB possuem características distintas de um DW, tendo em vista que o objetivo desses bancos de dados é para o respectivo sistema e seus usuários diretos, enquanto um DW é voltado para gestores de alto nível que precisam apenas dos dados necessários para tomada de decisão. Abaixo no quadro 1, pode-se verificar essas e outras diferenças que distingue o BD dos sistemas do DW.

Quadro 1 - Tabela de comparação entre as características dos bancos de dados operacionais com um Data Warehouse

Características	Bancos de dados Operacionais	Data Warehouse - DW
Objetivo	Operações diárias do negócio	Analisar o negócio
Uso	Operacional	Informativo
Tipo de processamento	OLTP	OLAP
Unidade de trabalho	Inclusão, alteração, exclusão.	Carga e consulta
Número de usuários	Milhares	Centenas
Tipo de usuário	Operadores	Comunidade gerencial
Interação do usuário	Somente pré-definida	Pré-definida e ad-hoc
Condições dos dados	Dados operacionais	Dados Analíticos

Volume	Megabytes – gigabytes	Gigabytes – terabytes
Histórico	60 a 90 dias	5 a 10 anos
Granularidade	Detalhados	Detalhados e resumidos
Redundância	Não ocorre	Ocorre
Estrutura	Estática	Variável
Manutenção desejada	Mínima	Constante
Acesso a registros	Dezenas	Milhares
Atualização	Contínua (tempo real)	Periódica (em batch)
Integridade	Transação	A cada atualização
Número de índices	Poucos/simples	Muitos/complexos
Intenção dos índices	Localizar um registro	Aperfeiçoar consultas

Fonte: Teixeira (2013)

2.2 COMPONENTES DE UM DATA WAREHOUSE

Segundo Marzarotto (2018, p. 27-28) um DW pode diferir em estrutura ou características de implementação. Entretanto, independente das diferenças de implementação alguns componentes são essenciais num DW qualquer, para Elias (2014) e Marangoni e Scheneide (2010, p. 34), cita alguns dos componentes básicos (ver figura 1):

Sistemas Operacionais de Origem - Sistemas operacionais de registro que capturam as transações da organização para cada dia, focados em desempenho e ter disponibilidade de processamento, com consultas limitadas que fazem parte do fluxo normal de transações, com a demanda extremamente limitada, com pequeno volume de histórico de dados e apresentam pouco ou nenhum controle sobre o seu conteúdo e o formato dos dados. É a fonte de dados mais comum, mas é possível que existam outros tipos de fontes de dados tal como planilhas, arquivos de texto etc.

ETL - (Extract- Transform- Load), onde dados operacionais brutos são transformados em um formato de warehouse prontos para serem consultados pelo usuário. A extração é a etapa inicial, onde são feitos os processos de leitura, compreensão e obtenção dos dados de origem. Na etapa de limpeza são feitas as transformações como filtragem dos dados (correção de erros de digitação, solução de conflitos de domínio, tratamento de elementos ausentes),

padronização dos campos de origem diferentes, remoção de dados duplicados, atribuição de chaves artificiais. Após todo o tratamento é que é feito a carga dos dados na área de apresentação dos dados.

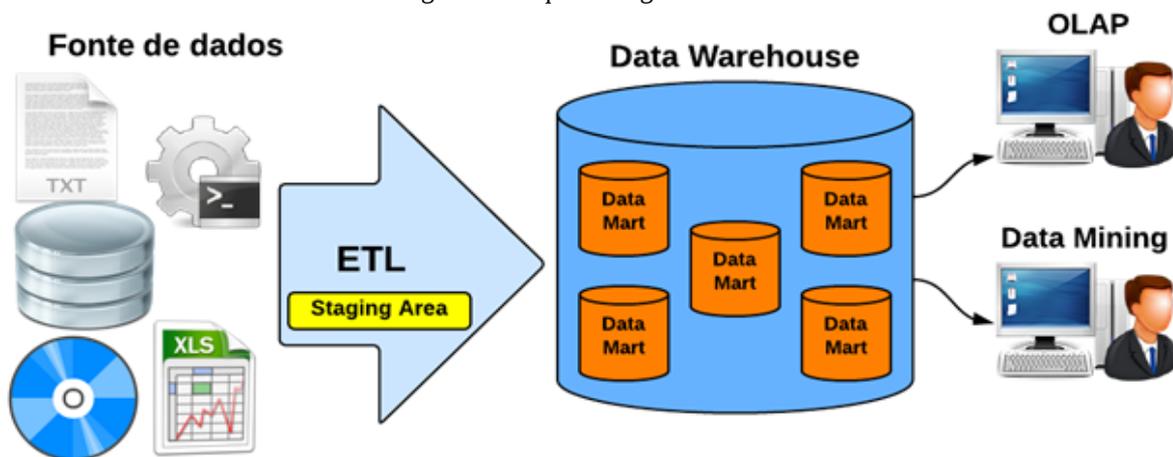
Data Staging Area - a Staging Area é uma área de armazenamento intermediário situada dentro do processo de ETL. Auxilia a transição dos dados das origens para o destino final no DW.

Apresentação dos Dados - Local onde os dados ficam organizados, armazenados para consultas no front end, através de relatório ou gráficos, mais comumente chamado Data Warehouse, ou numa tradução livre armazém de dados.

Ferramentas de acesso aos dados ou OLAP - On-Line Analytical Processing (Processamento Analítico On-Line) - É a conexão com o usuário final, feitas por aplicações analíticas orientadas a parâmetros, que utilizam modelos prontos, por aplicações que possibilitam o usuário construir consultas relacionais, por aplicações de modelagem ou previsão tudo isso com de eficiência e rapidez.

Data Mining: Data Mining ou Mineração de Dados, se refere as ferramentas com capacidade de descoberta de conhecimento relevante dentro do DW. Encontram correlações e padrões dentro dos dados armazenados.

Figura 1 - Arquitetura genérica de DW



Fonte: ELIAS (2014)

2.3 METADADOS

Segundo Silva (2004) metadados é o componente mais importante de um ambiente DW por se tratar dos dados de controle do projeto. Metadados são dados sobre dados, ou em outras palavras biblioteca de dados. Barbieri (2011, 68) ensina que “Os dados, para

produzirem devidamente informação e conhecimento, deverão ter sua definição completa e unívoca. E a definição de dados é a essência de metadados.”

Os metadados contêm, pelo menos, as seguintes informações (INMON, 1997):

- A estrutura dos dados;
- Os Algoritmos usados para a sintetização;
- A fonte de dados que alimenta o DW;
- O mapeamento desde o ambiente operacional até o DW;
- O modelo de dados;
- O relacionamento entre o modelo de dados e o DW;
- O histórico de extrações.

2.4 DATA MART – DM

Segundo Bomfim (2001, p. 32) um Data Mart já foi definido como um subconjunto altamente agregado de dados do Data Warehouse, desempenhando um papel de um DW departamental ou direcionado a uma área específica do negócio.

O Data Mart agora é definido como um conjunto flexível de dados, frequentemente com base no dado mais atômico (granular), sendo possível os dados serem extraídos de uma fonte original, e modelados de forma dimensional que é mais flexível para as consultas inesperadas (não estruturadas) dos usuários. (BOMFIM, 2001, p. 32)

Marangoni e Scheneide (2010) ensina que um projeto de Data Warehouse ser desenvolvido focando pequenos grupos dentro de uma organização, esse DW focado é chamado de Data Mart. Eles consideram que a diferença entre DW e DM está no tamanho do problema a ser resolvido, embora a definição do problema e dos requisitos são basicamente os mesmos. Por outro lado Kimball (1998) define o DW como o conjunto de Data Marts integrados. Assim considerando essas diferenças, o desenvolvimento de um DW requer tempo, dados e investimentos muito maiores que um DM visto que um DW envolve toda uma organização, em vários níveis e setores. Silva (2004, 53) lista como vantagem de modelagem de um DM a possibilidade ser prototipado em 30 a 120 dias.

2.5 ESTRATÉGIAS ARQUITETÔNICAS PARA A IMPLEMENTAÇÃO DE DW

Bomfim (2001) ensina que existem basicamente três estratégias arquitetônicas para implantação de Data Warehouse - DW: “top-down”, “bottom-up”(ver comparativo no quadro 2) e intermediária.

Na estratégia “Top-down” cria-se um DW abrangente e complexo que atende toda a organização para depois segmentar em Data Marts - DM, para cada departamento.

Já na estratégia “Bottom-up” são desenvolvidos DM independentes para os aspectos mais críticos que podem vir a se integrarem. Neste caso o DW será o resultado da integração de todos do DM da organização.

Na abordagem intermediária busca-se integrar as duas abordagens, nesse caso modela-se o DW e na sequência implementa-se os DM.

Quadro 2 - Comparativo das abordagens “Top-down” e “Bottom-up”

Top Down	Bottom-up
Data Warehouses Corporativos	Data Marts Setoriais
De grande abrangência	Projetos evolutivos
Complexos	Enfoque inicial nos aspectos mais críticos
Alta risco de insucesso.	Aproveitamento da estrutura operacional disponível
Custo maior	Retorno mais rápido
Alta padronização entre todos os setores da organização	Acúmulo de experiência: menor risco e menor custo
Manutenção mais fácil	Baixa padronização entre diferentes DM
Longo tempo de implementação	Requer maior esforço para integração dos DM

Fonte: Elaborado pelo autor baseado em Silva(2004), Bonfim(2001) e Vieira(2001)

Nesse trabalho se fará necessário tem-se em mente a abordagem Bottom-up, na medida que se modela e se implementa um DM para uma questão crítica, que é a retroalimentação do ciclo operacional de Perícia.

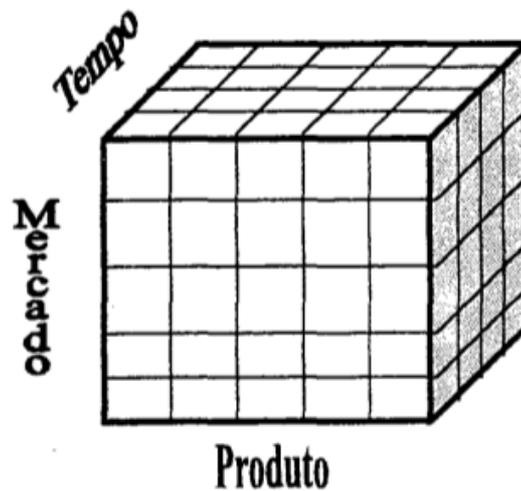
2.6 MODELO DIMENSIONAL

De acordo com Vieira (2001, p. 42)

“O modelo dimensional representa o banco de dados como se fosse um cubo contendo três, quatro ou mais dimensões, e as pessoas conseguem fatiar esse cubo em qualquer de suas dimensões. As dimensões são representadas pelas arestas do cubo (produto, mercado, tempo), e qualquer ponto dentro do cubo representam o local em que as medições da organização ou negócio estão armazenadas. Ambos modelos são capazes de armazenar exatamente os mesmos dados e suportar exatamente a mesma análise final da organização ou negócio.”

Ver figura 2.

Figura 2 - Modelo dimensional de um negócio: visualização dos dados como um cubo.



Fonte: Vieira (2001, p. 42).

A modelagem dimensional é uma técnica de projeto lógico que busca apresentar os dados em uma estrutura padronizada e intuitiva, que permita alto desempenho de acesso, específica para suportar processamento analítico (OLAP). Assim, cada modelo dimensional é formado por uma tabela de fato e um conjunto de dimensões. (BARBIERI, 2011, p. 97)

2.6.1 Tabela Fato

Um fato é constituído por um conjunto de dados compostos de métricas e dados contextuais. Deve representar uma determinada transação ou evento do negócio ocorrido num determinado contexto, obtido na intersecção das diversas dimensões. Em outras palavras representa um evento ou uma ação real registrada no sistema como um incêndio, um salvamento, uma compra, etc. (MARANGONI; SCHENEIDE, 2010, p. 39)

As tabelas de fatos não mudam de acordo com o tempo e são tabelas de grande porte, podendo incluir milhões de registros. As informações destas tabelas correspondem às medidas referentes às vendas, total de número de venda, dentre outros. Tabelas de fatos usualmente são tabelas normalizadas. (SILVA, 2004, p. 47)

2.6.2 Tabela Dimensão

As dimensões referem-se a um ou mais contextos de negócios onde ocorrem os fatos. São elementos que permitem descrever o fato, classificando as métricas ativas de uma organização. São os pontos de entrada para a tabela de fatos e permitem implementar a interface de utilizador para o Data Warehouse – DW. (PITON, 2018, p. 100)

A maioria dos fatos envolve, pelo menos, quatro dimensões fundamentais:

Onde? - Determina o local onde o fato ocorreu (local geográfico, etc);

Quando? – Determina a própria dimensão do tempo;

Quem? - Determina que entidades participaram no fato (vítima, autor, etc);

O quê? – Determina qual é o objeto do fato (incêndio, morte, etc).

(MARANGONI; SCHENEIDE, 2010, p. 40)

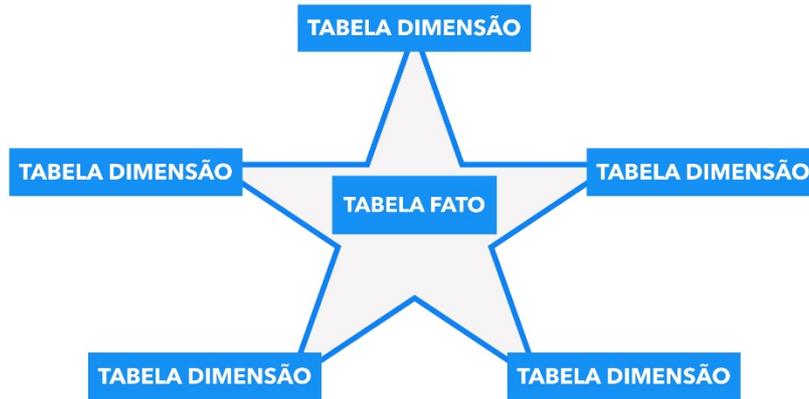
Nesse ponto é interessante destacar que existe um tipo de dimensão que não mereceu ser uma dimensão, mas a informação necessária, nesse caso é inserida numa coluna da tabela fato, a essa informação que servirá como filtro para consultas, essa dimensão é chamada de dimensão degenerada. (PITON, 2018, p. 174)

Se fosse definido que um incêndio é um fato, então, poder-se-ia definir (como exemplo geral), a data, o endereço, as vítimas, a edificação cada um sendo uma dimensão. Na dimensão vítima seria possível registrar idade, sexo, renda, escolaridade, entre outros. Assim cada dimensão é capaz de receber vários dados conforme a projetado para o DW.

2.7 STAR SCHEMA

O conceito de Star Schema, (modelo estrela), figura 3 foi projetado por Ralph Kimball. O Star Schema é composto no centro por uma tabela fato que é rodeada por tabelas de dimensões, ficando parecido com a forma de uma estrela (que é de onde vem o nome). ((PITON, 2018, p. 95)(BARBIERI, 2011, p. 167-170))

Figura 3 - Star Schema



Fonte: PITON (2018, p 95)

2.8 SURROGATE KEY

Segundo Piton, (2018, p. 104) “surrogate key” são chaves artificiais, auto incremental que de forma semelhante a uma chave primária (ou “primary key”) serve para identificar de forma única uma linha das tabelas dimensão e relacioná-las as tabelas fatos. Não são propriamente chaves primárias pois não são geradas pelo SGBD, mas no processo de ETL. Ressalta-se que ficam necessariamente armazenados nas dimensões as chaves primárias originais quando estas existirem.

2.9 FERRAMENTAS PARA CONSTRUÇÃO E OPERAÇÃO DE UM DW/DM

Atualmente existem diversas ferramentas para a construção e operação de DW/DM, essas ferramentas atendem as diversas fases da modelagem e implementação. Existem ferramentas para Extração, Tratamento e Limpeza de dados (Ferramentas ETL), e ferramentas que vão propiciar aos usuários visualizar os dados presentes no DM.

Existem diversos softwares que são capazes de executar o ETL, existem ferramentas pagas e ferramentas gratuitas, dentre estas destaca-se Pentaho Data Integration/Kettle ferramenta escolhida por ser de domínio prévio do orientador do autor deste trabalho. Já como ferramentas de acesso aos dados para a visualização e geração de relatórios dos usuários, no estado de Santa Catarina, através da Secretaria de Segurança Pública, é utilizado o Qlik Sense, que é uma ferramenta comercial pronta de alto nível. Considerando que a SSP disponibiliza acessos ao CBMSC, foi a escolha, tendo em vista que não faz parte do presente trabalho a análise dessa classe de soluções de BI. A grande vantagem dessas ferramentas de BI prontas, é que poupa os gestores de investir em desenvolvimento de aplicações próprias.

3 O CICLO OPERACIONAL DE BOMBEIRO

Neste capítulo será apresentado brevemente o ciclo operacional de bombeiro, conforme previsto teoricamente e será feita uma breve explanação sobre a efetividade atual da quarta fase do respectivo ciclo.

3.1 INTRODUÇÃO

Segundo Sarte (2009, p. 24)

O Ciclo Operacional relativo às atividades de combate e investigação de incêndios do CBMSC compreende quatro fases bem distintas, apresentando cada uma delas finalidade específica, objeto de estudo, custeio da atividade, recursos humanos e recursos materiais peculiares. As quatro fases desse ciclo são: Fase Normativa ou Preventiva, Fase Passiva ou Estrutural, Fase Ativa ou de Combate e a Fase Investigativa ou Pericial.

Na figura 4 temos a representação gráfica do ciclo operacional de bombeiro. As fases estão todas conectadas diretamente, pois todas ocorrem simultaneamente, por exemplo, a fase pericial analisa ao investigar o incêndio todas as demais fases, e também fornece informações para todas estas fases do ciclo operacional de bombeiro. Na próxima seção abordaremos brevemente o significado de cada uma dessas fases. (CBMSCb, 2018; LEMOS, 2017 p. 27-28; MAUS, 1999 p. 28-33; VIDAL, 2007, p 15-17).

3.2 FASES DO CICLO OPERACIONAL DE BOMBEIRO

Fase Normativa ou Preventiva: As normas reguladoras dos sistemas de segurança contra incêndio e pânico são estudadas, elaboradas, interpretadas e comentadas a fim de se evitar a ocorrência de sinistros, ou se ocorrer ter seus efeitos sobre vidas e sobre os bens mitigados, prevenindo em último caso a perda de vidas humanas. Participam dessa fase todos os atores (públicos e privados, nacionais e estrangeiros) que colaboram na elaboração e publicação de normas preventivas contra incêndios; (CBMSCb, 2018; LEMOS, 2017 p. 27-28; MAUS, 1999 p. 28-33; VIDAL, 2007, p 15-17).

Fase Estrutural ou Passiva: destina-se a restringir ou minimizar as consequências e as vulnerabilidades da edificação, através da elaboração de projetos e posterior execução da obra. Essa fase é executada através da análise de projetos e vistorias das edificações além do

exercício do poder de polícia administrativa; (CBMSCb, 2018; LEMOS, 2017 p. 27-28; MAUS, 1999 p. 28-33; VIDAL, 2007, p 15-17).

Fase Ativa, Reativa ou de Combate: Nessa fase é que está toda a atuação das guarnições de Bombeiro, desde a sua preparação, até a sua atuação nos incêndios ocorridos. Nesta fase são aplicadas as doutrinas, estratégias, táticas e técnicas de salvamento de vítimas e o controle e a extinção do fogo; (CBMSCb, 2018; LEMOS, 2017 p. 27-28; MAUS, 1999 p. 28-33; VIDAL, 2007, p 15-17)

Figura 4 - Ciclo operacional de Bombeiro



Fonte: Elaborado pelo autor com base nos trabalhos de: (MAUS, 1999 p. 28-33; VIDAL, 2007, p 15-17; ACORDI, 2011, p.93; LEMOS, 2017 p. 27-28; CBMSCb, 2018)

Fase Investigativa ou Pericial: destina-se a analisar os locais de sinistros de maneira a elucidar e identificar as causas e efeitos do sinistro, nessa seara, se analisa todas as demais fases do ciclo, se a atuação de combate das GU de serviço foi correta (Fase Ativa), se o cumprimento das normas foi eficaz e respeitado (Fase Estrutural), e por fim se as próprias normas foram são adequadas (Fase Normativa), é na fase investigativa que se pensa em todas as demais fases do ciclo. Por fim o resultado dessa investigação deverá ser retroalimentar, ou seja, melhorar a atuação, nas demais fases do ciclo operacional. (*ibid*).

Maus (1999) ensina que a perícia de incêndio não tem por objetivo somente a descoberta das causas do incêndio, mas a análise de todas as circunstâncias que contribuíram para a ocorrência do sinistro e seu posterior combate.

3.3 RETROALIMENTAÇÃO DO CICLO OPERACIONAL

A importância das investigações de ocorrências de incêndio tem, por fornecer subsídios a cada uma das demais fases do ciclo operacional, ou seja, como a investigação de incêndios pode retroalimentar as demais fases do ciclo operacional de Bombeiro, conforme bem explanada por Vidal (2007, p. 19-20):

- para a fase Normativa, evitar a ocorrência futura de sinistro com riscos idênticos aos riscos do caso real analisado. É possível, ainda, estudar, revisar e elaborar normas de segurança, assim como identificar as responsabilidades pelas falhas que resultaram no descumprimento das normas de segurança. Os avanços tecnológicos podem tornar obsoletos alguns sistemas de segurança, gerando necessidade de revisão normativa, o que será possível identificando os casos reais de sinistro;
- para a fase Estrutural, é possível contribuir ao apontar falhas de projetos de segurança contra incêndios; falhas de concepção, dimensionamento, instalação, manutenção ou de operação de sistemas de segurança contra incêndios e propor novas medidas; apontar falhas de desempenho e propor programas de treinamento para o desenvolvimento de recursos humanos que efetivamente atuam na segurança física das edificações, no que se refere a sua segurança contra incêndios;
- para a fase de Combate, será possível indicar falhas de viaturas e equipamentos típicos de bombeiro; falhas humanas nas operações de combate a incêndios; falhas táticas ou técnicas no emprego do poder operacional da Corporação; propor o redimensionamento do poder operacional, ou seja, a aquisição de novas viaturas e equipamentos especializados; propor o aperfeiçoamento do planejamento estratégico relativo à capacidade de mobilização e de articulação de poder operacional; propor a modernização das técnicas e táticas das operações de combate típicas de bombeiro; concluir sobre o desempenho operacional no caso real analisado; propor novos programas de treinamento para o desenvolvimento dos recursos humanos da Corporação, em função dos dados que resultaram da análise do caso real do sinistro; propor readequação dos critérios para instalação de novos Quartéis de Bombeiro e a adequação quantitativa das guarnições de combate a incêndios. No aspecto positivo, a perícia poderá identificar a um item de qualidade do serviço do CBMSC, ao trabalhar as estatísticas dos salvados em incêndios.

Lemos (2017, p. 28) sintetiza de forma clara correlação entre as fases do ciclo operacional:

As NSCIP estabelecem normas, sistemas preventivos contra incêndio e pânico, análise de projetos preventivos contra incêndio e pânico, realização de vistorias, todos com a premissa de evitar incêndios nas edificações, porém, ainda sim, o incêndio existiu. O que aconteceu?

Os sistemas preventivos instalados não foram necessários para mitigar o incêndio em sua fase inicial? Por que utilizaram o sistema de maneira equivocada? O desenvolvimento dos programas institucionais para sociedade, no que tange aos cuidados domésticos, prevenção e combate ao incêndio, não foram eficazes? Onde deve ser melhorado? Esses questionamentos conduzem a atividade investigativa de incêndios.

3.4 GESTÃO DO CONHECIMENTO DAS INVESTIGAÇÕES NO CBMSC

A falta de acesso fácil e sistematizado às informações das investigações parece ser grande gargalo para que a fase investigativa de fato contribua e retroalimente as demais fases do ciclo operacional como previsto.

Segundo Timmermann e Acordi(2019, p. 13) atualmente, a utilização do que se produz nos Laudos e Informes Periciais ainda é pequena no sentido de interligação das fases e retroalimentação do ciclo de atividade operacional do CBMSC. Destacam ainda que

A produção de conhecimento no ciclo de atividade operacional do CBMSC, nos dias atuais, é feita da seguinte maneira: após um sinistro de incêndio em edificação (ou ainda incêndios em veículos, aeronaves, embarcações e vegetação, a depender da disponibilidade do perito, da complexidade e relevância da ocorrência), um perito ou inspetor de incêndio do CBMSC é acionado para realizar a investigação, a qual é feita observando a doutrina adotada na Corporação. Com o auxílio de um formulário (Auto de Investigação de Incêndio e Explosão), os dados da investigação levantados são registrados e posteriormente inseridos no Sistema de Perícias do CBMSC, um sistema informatizado que os registra e organiza. (TIMMERMANN; ACORDI, 2019, p.13)

Estes dados são uma fonte rica para pesquisa com elementos críticos de conhecimento e informação que podem efetivamente retroalimentar as demais fases . Apesar disso, observa-se que o uso delas ainda é pequeno diante do potencial de mudanças positivas que podem promover.

Um exemplo do que se faz atualmente é a extração de estatísticas, como a quantidade de investigações realizadas a cada mês. Estes dados por si possuem pouca utilidade prática e não contribuem para a construção de conhecimento institucional, na medida que não fornecem novas informações para as etapas do ciclo de atividade operacional pois não ocorre uma interação e troca de informação entre as etapas, sendo que cada uma encontra-se isolada das demais. Tudo o que se produz é retido dentro da própria esfera de trabalho. (TIMMERMANN; ACORDI, 2019, p.14)

Assim percebe-se que há deficiência na gestão do conhecimento que em parte pode ser explicada pela falta de ferramentas adequadas para tal propósito. Na metodologia de “Business Intelligence” - BI, através do DM introduz a corporações, o desenvolvimento processos para gerar conhecimento aos níveis decisórios fáceis de serem entendidos e fáceis de serem produzidos.

4 DATA MART DAS INVESTIGAÇÕES DE INCÊNDIO

Neste capítulo será explanado os procedimentos metodológicos adotados, a infraestrutura tecnológica, o planejamento do projeto, as definições de requisitos e as fontes de dados..

4.1 PROCEDIMENTOS METODOLÓGICOS

Para esta pesquisa foi adotada a seguinte metodologia e procedimentos metodológicos:

4.1.1 Tipo de pesquisa

O estudo será estruturado em duas pesquisas, uma descritiva e um estudo de caso descritivo. Aquela sob forma de pesquisa bibliográfica com leituras e fichamentos de livros, artigos e busca em matérias na Internet com uma abordagem qualitativa, limitada ao necessário para fundamentar a segunda. A segunda pesquisa terá uma abordagem descritiva. Segundo Rudio (2002), este tipo de pesquisa (descritiva) além de descrever as características de uma determinada população, ou de um fenômeno, também interpreta os fatos.

4.1.2 Método

Optou-se pelo método científico indutivo para desenvolver esta pesquisa, pois o objeto de estudo é um fenômeno constatado na realidade.

1- Método Indutivo – cuja aproximação dos fenômenos caminha geralmente para planos mais abrangentes, indo das constatações mais particulares às leis e teorias (conexão ascendente); método que considera o conhecimento como baseado na experiência; a generalização deriva da observação de casos da realidade concreta e são elaboradas a partir de constatações particulares. (ALMEIDA, (2017, p. 1) apud PEREIRA; SHITUSUHA; et al (2018, p. 28)).

4.1.3 Técnica e instrumentos de pesquisa

Além de observação em lócus, serão empregados outros recursos como acesso a banco de dados, coleta de registros estatísticos de relatórios de dados de investigações de incêndios ocorridos em SC, e anotações de impressões pessoais no diário de pesquisa. Assim como também coleta de informações verbais ou troca de correspondência eletrônica junto a departamentos específicos do CBMSC ou da SSP.

4.1.4 Local da Pesquisa.

Dependências do CBMSC.

4.1.5 Objetivo da Pesquisa

Essa pesquisa tentará apresentar resultados proveitosos no mundo concreto sendo empregue em uma área do CBMSC para um problema específico.

4.1.6 Requisitos e Procedimentos e para Modelagem e Implementação do DM

Para desenvolver a modelagem do Data Mart - DM, foi definido utilizar a arquitetura bottom-up de Ralf Kimball(1998), pois possibilita um crescimento incremental futuro, além do que, se baseará em critérios já definidos cujos os dados já podem ser obtidos do sistema de perícias existente no CBMSC.

O Sistema de Gerenciamento de Banco de Dados - SGBD escolhido para armazenar os dados da Staging Area, é o SGBD definido como padrão no CBMSC (Postgres). Para o desenvolvimento do processo de ETL foi utilizado o Kettle, por ser gratuito e ter sido o indicado pelo orientador do Autor.

A Ferramentas de Acesso aos Dados escolhida para a apresentação das informações, relatórios “*ad hoc*” e resultados será utilizada o aplicativo Qlik Sense, porque esta tem sido a ferramenta disponibilizada para este tipo de atividade nos órgãos da Secretaria de Segurança Pública no projeto Analytics (informação verbal)¹ e o qual foi disponibilizaram ao autor uma conta de usuário com privilégios de desenvolvedor.

A metodologia a ser aplicada para a modelagem do dimensional, podem ser divididos nos seguintes passos.

- 1) Definir a área de negócios;
- 2) Escolher o Processo;
- 3) Definir a Granularidade desejada para os dados do processo;
- 4) Identificar e conformar as dimensões;
- 5) Escolher os Fatos;
- 6) Definir as métricas das tabelas fatos;
- 7) Armazenar os Pré-calculados na tabela de fatos;
- 8) Preencher as Tabelas de Dimensões;

¹Informação fornecida pelo Coordenador do Projeto Analytics, Bruno Depizzolatti em dez/2018.

- 9) Escolher a Duração do Banco de dados;
- 10) Como rastrear dimensões de modificação lenta; e,
(KINBALL, 1997; BARBIERI, 2011)

4.2 INFRAESTRUTURA TECNOLÓGICA

Nesta seção será descrito a infraestrutura tecnológica existente ou que foi preparada para iniciar a presente pesquisa.

4.2.1 Infraestrutura no CBMSC

O Corpo de Bombeiros Militar de Santa Catarina possui em sua estrutura a Diretoria Logística e Finanças, esta diretoria está dividida em 3 divisões a saber, Divisão de Logística - DiL, Divisão de Finança - DiF e Divisão de Tecnologia da Informação - DiTI. Devido a importância que esta divisão tem para o CBMSC, já está previsto como Diretoria na Lei Complementar Nº 724, de 18 de Julho de 2018 (Lei de Organização Básica do CBMSC).

Dentro da DiTI existe o CPD – Centro de Processamento de Dados, que é o local onde estão fisicamente hospedados os diversos servidores de aplicação, banco de dados e arquivos do CBMSC em um ambiente com temperatura, umidade e acesso restritos e monitorados. Nesse ambiente foi fornecido pelo chefe do CPD um servidor, nomeado de “datamart1” para o desenvolvimento deste trabalho e instalação da “*Data Staging Area*” com os seguintes requisitos:

- SO padrão CBM para servidores: CentoOS 7.3 X86_64
- SGBD PostgreSQL acessível a partir do servidor QlikSense da SSP
- Conexão SSH e senha de root local
- Conexão aos servidores E193 / SIGAT.

(SOUSA, 2019)

O Sistema Operacional de Origem é o sistema de Perícia, que possui um Banco de Dados - BD conjugado com o BD do sistema E-193, mas com suas tabelas identificadas com a marcação de prefixo dos nomes das respectivas tabelas, sendo que essas tabelas iniciam-se com as letras “aa”, por exemplo “aa_pericias”. Mais a frente será realizado um melhor detalhamento do BD em questão. Entretanto é importante salientar que existem 12 servidores para o estado de SC. São 12 servidores não integrados, com 12 bases de dados, com o sistemas operando de forma independente em cada um desses servidores.

4.2.2 Secretaria de Estado da Segurança Pública - SSP

A SSP, subordinada ao Governo do Estado de Santa Catarina, ao qual o CBMSC é vinculado possui em sua estrutura a Divisão de Tecnologia da Informação e Comunicação - DTIC, a qual já utiliza em sua estrutura ferramentas de Business Intelligence - BI, através do programa Qlik Sense instalado em seus servidores (informação verbal)². Esse servidor deve ser capaz de acessar o servidor “datamart1” para importar os dados da “Data Staging Area”.

4.2.3 Acessos, segurança e dificuldades relacionados à infraestrutura

Para utilizar a infraestrutura tecnológica do CBMSC inicialmente para fazer o processo ETL, que será descrito mais à frente, inicialmente foi utilizada um servidor reserva do Sistema de Perícia/E193 de Florianópolis, chamado de “*fpolis2*”. O autor já tinha conhecimento prévio de acesso a esses servidor, necessitando apenas de permissão para o uso. Posteriormente foi fornecido o servidor próprio chamado de “*datamart1*” conforme já especificado.

No lado da SSP, onde o autor nunca trabalhou foi necessária uma visita para o recebimento de uma assinatura do sistema Qlik Sense. Após vários testes e problemas resolvidos por “*WhatsApp*” o autor recebeu acesso ao servidor de desenvolvimento para usar a ferramenta em questão.

Importante destacar que muitas dificuldades foram possíveis de serem resolvidas devido ao autor conhecer pessoalmente os responsáveis pelas liberações de acesso, inclusive na SSP, recebendo auxílio diretamente por telefone ou “*WhatsApp*”. Entretanto, o acesso só é possível de dentro da rede do governo, que é uma medida de segurança cibernética adotada por ambas instituições, ou por acesso via VPN, que o autor já possui devido a essas restrições.

O script do processo ETL feito de forma gráfica com o programa Pentaho inicialmente ficou rodando no computador do próprio autor enquanto foi sendo construído e aperfeiçoado, uma vez pronto rodará no servidor “*datamart1*” que é onde fica o “*Data Staging Area*”.

4.3 DETALHAMENTO DO SISTEMA OPERACIONAL DE ORIGEM

O acesso ao Sistemas Operacionais de Origem, foi facilitado tendo em vista o autor já ter trabalhado no DiTI e já possuir contato com os administradores do sistema, aos quais foi solicitado acesso a esses sistemas, o que foi autorizado com fornecimento de login e senha. Entretanto é relevante destacar que segundo Marzarotto (2019) devido a limitações da rede na

época da construção dos sistema E193, foram criados diversos servidores ao redor do estado, um para cada BBM, o que faz com que atualmente consultas gerenciais no nível estadual consolidadas se tornem uma tarefa mais complexa. As ferramentas e metodologia deste trabalho resolvem isso, o que será tratado com maior detalhamento no processo ETL, resta por ora saber que foi necessário acesso a cada um dos Bando de Dados - BD reservas dos batalhões escolhidos por conveniência no momento aos BBM da 1ªRBM:

- bcu2.cbm.sc.gov.br - 13ºBBM - Balneário Camboriú;
- bnu2.cbm.sc.gov.br - 3ºBBM - Blumenau;
- cua2.cbm.sc.gov.br - 4ºBBM - Criciúma;
- iai2.cbm.sc.gov.br - 7ºBBM - Itajaí;
- tro2.cbm.sc.gov.br - 8ºBBM - Tubarão;
- fpolis2.cbm.sc.gov.br - 1ºBBM e 10ºBBM- Florianópolis e São José

Aqui cabe uma pequena observação, a escolha pelos servidores reserva se dá devido ao fato que estes servidores têm suas bases de dados atualizadas diariamente, o que é suficiente para este projeto, além disso, evita-se a sobrecarga dos servidores principais com mais processamentos desse projeto (informação verbal)².

O banco de dados - BD do sistema de perícias foi a principal Sistema Operacional de Origem do presente projeto. No entanto importante ressaltar que o BD está na mesma estrutura do BD do sistema E193, compartilhando algumas tabelas para ambos sistemas. As tabelas exclusivas do sistema de perícias foram separadas apenas com o prefixo no nome: “aa”. Segue abaixo todas as tabelas existentes com essa marcação:

- aaacao_tatica_fase_ativa;
- aaatividade_fase_ativa;
- aapericia;
- aapericia_admin;
- aapericia_arquivo;
- aapericia_auditoria;
- aapericia_consideracao;
- aapericia_documento_sigat;
- aapericia_fase_ativa;

²Informação fornecida pelo chefe do CPD/DLF, 3º Sgt Carlos Alberto Sousa em dezembro de 2018

- aapericia_fase_ativa_acao_tatica_item;
- aapericia_fase_ativa_acao_tatica_compl;
- aapericia_fase_ativa_atividade_compl;
- aapericia_fase_ativa_atividade_item;
- aapericia_fase_ativa_recurso;
- aapericia_fase_ativa_tempo_resposta;
- aapericia_fase_investigativa;
- aapericia_fase_investigativa_dano_material;
- aapericia_fase_investigativa_seguro_incendio;
- aapericia_fase_investigativa_vitima;
- aapericia_fase_passiva;
- aapericia_fase_passiva_spci;
- aapericia_fase_passiva_utilizacao_sistema;
- aapericia_imovel_caract_construtiva;
- aapericia_imovel_descricao;
- aapericia_imovel_responsavel;
- aapericia_investigacao_responsavel;
- aapericia_produtos_perigosos;
- aapericia_troca_informacao;
- aarecurso_fase_ativa;
- aasistema_preventivo;
- aastatus_acao_tatica_fase_ativa;
- aastatus_atividade_fase_ativa;
- aastatus_recurso_fase_ativa;
- aastatus_sistema_preventivo;

O Sistema de Perícia também utiliza outras tabelas do sistema E193 para fins de consulta e dados de usuário, todas as tabelas do E193 estão anexados ao fim do trabalho por se tratar de 279 tabelas.

Entretanto o BD do Sistema de Perícia não armazena em suas tabelas alguns meta dados de parte de dados necessário para definir a qualidade da classificação cito: o nome correspondente a chave primária numérica armazenada no BD, entretanto essa informação está registrada no código do sistema, estando armazenado numa coluna correspondente de uma tabela do sistema o apenas número correspondente ao nome do tipo selecionado pelo usuário. Além disso, o aparentemente os programadores do Sistema de Perícia utilizaram como referência para a elaboração da tabelas de “tipos”, os “tipos” do sistema E193. Assim, muito embora, em grande parte, há uma correspondência parcial tanto nos nomes, quanto nos códigos, existem diferenças que inviabilizam a reutilização, tendo sido verificado na documentação do Sistema de Perícia. Assim foi inviabilizado a utilização daquelas tabelas. Portanto foi necessário produzir algumas dessas tabelas para fornecer ao DM. Ainda há outras tabelas do sistema E193 que foi conferida compatibilidade 100% que foram utilizadas também:

- tb_causa;
- tp_edificacao;

E ainda houve um caso de uma tabela “tipo” que após a aplicação de um filtro, houve compatibilidade 100%.

- tp_emergencia_subgrupo (renomeada para tipo_incendio, filtrada para apenas o grupo incendio);

4.4 STAGING AREA

Para o Staging Area foi criado num servidor virtual no CPD do CBMSC, sendo configurado exclusivamente para esse fim. Tendo em vista que o CBMSC já definiu o POSTGRESQL como o Sistema de Gerenciamento de Banco de Dados da corporação, foi mantido o esse padrão sendo o Staging Area construído no PostgreSQL.

No PostgreSQL utilizando-se o schema “public” foi criado uma Data-base, nomeada de “datamart”. Não foram criadas tabelas nesse momento pois o aplicativo PDI permite a construção e edição das tabelas necessárias para a fase de LOAD de forma automatizada, bastando a configuração da conexão a um SGBD e a criação de Data Base no mesmo.

Ao fim do desenvolvimento do ETL, descrito no próximo capítulo, o Staging Area ficou assim desenhado:

4.5 PLANEJAMENTO DO PROJETO

O planejamento do projeto se iniciou com a definição da área de negócios, que foi escolhido o ciclo de operacional de bombeiro. O próximo passo foi escolher o processo, que ficou definido sendo as investigações periciais de ocorrências de incêndio (nesse caso abrangendo tanto os informes periciais quanto os laudos periciais). Tal escolha se pautou numa investigação prévia de demandas conforme já fundamentado nas justificativas do presente trabalho.

A Definição da Granularidade desejada para os dados do processo é o próximo passo, considerando que o número de investigações de incêndio são relativamente poucos mesmo no contexto estadual, e mesmo o número de incêndios são relativamente poucos, ficou definido que o grão da tabela fato a transação, assim temos o que é chamado de fato transacional. Entretanto para os dados gerais de Incêndios pode-se definir como grão o mês, nesse caso os dados são agregados a cada mês, nesse caso tem-se o fato agregado. (PITON, 2018, p. 98).

4.5.1 Dimensões

As dimensões escolhidas foram levados em conta o documento produzido e fornecido pela Divisão de Perícias: “Parâmetros para relatórios Módulo Perícia”, disponível no anexo C, a partir desse ponto foi verificado nas bases de dados do Sistema de Perícia, os dados disponíveis existentes capazes de atender a demanda.

Ficou assim definido as seguintes dimensões:

- tempo;
- causa;
- subcausa;
- evento causal;
- tipo de incêndio;

- zona de origem;
- tipo de edificação;
- cidade;
- vítima.

Além dessas dimensões, ficou definido mais algumas dimensões degeneradas:

- batalhão (número);
- viaturas adequadas (1/0 – sim/não);
- quantidade de bbmm adequados (1/0 – sim/não));
- fase do incêndio (número);
- tempo resposta adequado (1/0 – sim/não));
- existência de vítima (1/0 – sim/não));
- seguro (1/0 – sim/não));

4.5.2 Fatos

Uma vez definido as dimensões, define-se os fatos, para o presente trabalho, considerando as necessidades apresentadas pela Divisão de Perícia/DSCI, definiu-se trabalhar com dois fatos: Investigação de Incêndio e Ocorrência de Incêndio. Essa definição foi feita assim pois embora toda investigação esteja relacionada a uma ocorrência de Incêndio, essas investigações são apenas para um tipo de incêndio e mesmo nestes casos não atendem todas as ocorrências. Entretanto, a investigação de incêndio traz um grande número de informações e a investigação de ocorrência de incêndio é tratado como uma etapa diferente no ciclo operacional de bombeiro, assim considerou-se mais adequado essa distinção dos fatos.

4.5.3 Métricas

Conforme o modelo de planejamento a próxima etapa é a definição das métricas das tabelas fatos. Para a tabela fato perícia definiu-se as seguintes métricas:

- Quantidade de BBMM empregados na ocorrência (número inteiro);
- Quantidade de BBCC empregados na ocorrência (número inteiro)

- Área da edificação (número racional)
- Área atingida (número racional)
- Valor Estimado Salvo (número racional)
- Valor Estimado do prejuízo (número racional)

Para a tabela fato ocorrências é o que chamamos de tabela fato sem fatos ou *factless fact table*. (PITON, 2018, p. 314)

Para a tabela fato ocorrências sumarizada definiu-se as seguintes apenas uma métricas:

- Quantidade de ocorrências.

Além das métricas e dimensões degeneradas outro dado que pode ser adequado acrescentar nas tabelas fatos são dados pré-calculados, ou seja dados que são calculados a partir dos dados da fonte, na fase de ETL, poupando processamento nas futuras consultas no Data Mart. Para a tabela Fato Perícia e Fato Ocorrência não houve dados pré calculados. Somente para a tabela fato ocorrência sumarizada foi calculado a quantidade de ocorrências por ano, por mês, por cidade e por tipo de incêndio.

4.5.4 Preenchimento Das Dimensões

A próxima etapa é o preenchimento das tabelas dimensões. O que será tratado com mais profundidade nos próximos parágrafos. A duração do Banco de dados ficou definido como sendo mensal, dados ao tipo de necessidade levantado na etapa inicial do projeto.

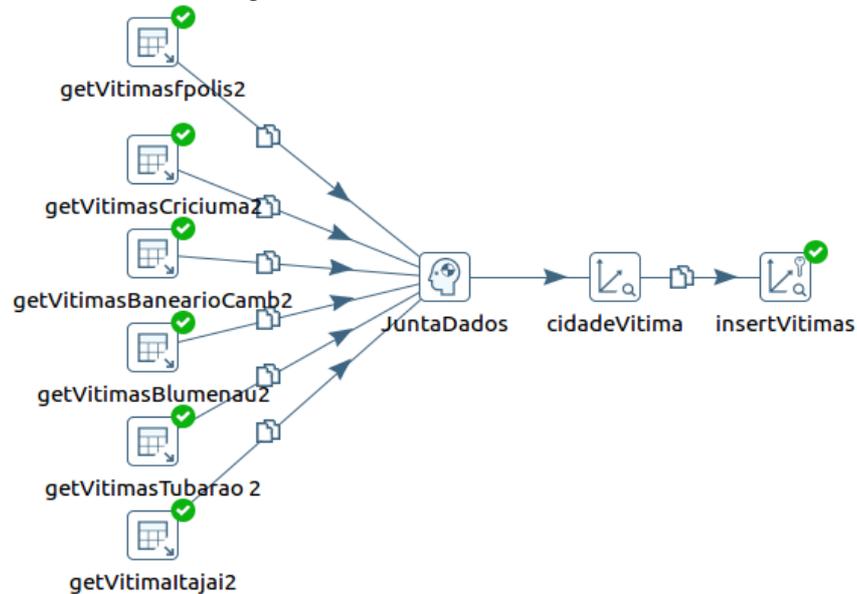
4.5.4.1 Dimensões preexistentes

A dimensão tempo já existia de um DM anteriormente construído, e fornecida sendo apenas carregado para o Staging Area. O mesmo ocorre com a dimensão cidade, já produzida anteriormente. Isso está de acordo com a metodologia adotada, Bottom-up, sendo o reaproveitamento de dimensões de outros DM preexistentes necessário para futura construção de um DW robusto.

4.5.4.2 Dimensão Vítima

A dimensão vítima, é gerada e necessariamente atualizada de regularmente por um processo ETL, figura 5 , tendo em vista que a cada ocorrência novas vítimas podem existir.

Figura 5 - ETL Dimensão Vítima



Fonte: Elaborado pelo autor

No primeiro passo, a extração, é feita uma consulta a cada um dos Bancos de dados da 1ª RBM, representado na figura 5 pelas caixas getVitima[nomeServidor], unidos por “JuntaDados”. A extração é feita pela seguinte consulta SQL, abaixo demonstrada:

Consulta SQL:

```

SELECT
  id_pericia,
  aa.id_cidade,
  vl_idade,
  tp_sexo,
  CASE fg_situacao_vitima WHEN 1 THEN 'FERIDO' ELSE
    CASE fg_situacao_vitima WHEN 2 THEN 'MORTO'
  END
END AS "vitima" ,
  fg_causa_situacao
FROM
  public.aapericia_fase_investigativa_vitima
left join public.aapericia aa using (id_pericia)
where id_status = 9
order by id_pericia
  
```

Fonte: Elaborado pelo autor

Há uma cláusula de filtro, “where id_status = 9”, que refere-se às perícias já concluídas, é a mesma filtro existente no processo de ETL da tabela fato perícia.

Na sequência é feita a substituição da chave primária da cidade, pela respectiva “Surrogate Key” - SK. Por fim, na última etapa, é feito o carregamento dos dados, pela função “insertVitimas” que gera nesse momento a SK, e insere todos esses dados na staging area.

4.5.4.3 Dimensões carregadas do Sistema E193

As dimensões “causa”, “tipoIncendio” e “tipoEdificação”, foram obtidas no BD do Sistema E193, com consultas SQL e inserção com uma função Dimension lookup, que estão nomeadas na figura 6 insertDim[nome-da-dimensão], cuja funcionalidade gera no momento da inserção um “Surrogate Key”- SK para cada linha de dados. Os dados dessas dimensões são basicamente, além do SK, a chave primária de origem e respectivo o nome:

Para a dimensão causa, a função getCausa extrai os dados a partir da consulta:

Consulta SQL:

```
SELECT
  cod_causa
, nm_causa
FROM tp_causa
```

Fonte: Elaborado pelo autor

No carregamento foi feita a alteração do nome da chave primária de “cod_causa” para id_causa, a SK, foi nomeada id_causa_sk e o nome foi mudado de nm_causa para causa. Mesmos procedimentos foram feitos para tipoIncendio e tipoEdificacao, com a única diferença que vale destacar que a origem do tipo de incêndio foi uma tabela de subclassificação de ocorrências, assim foi necessário uma cláusula de filtro na consulta SQL, na função getTipoIncendio (figura 6) cujo resultado é apresentado abaixo:

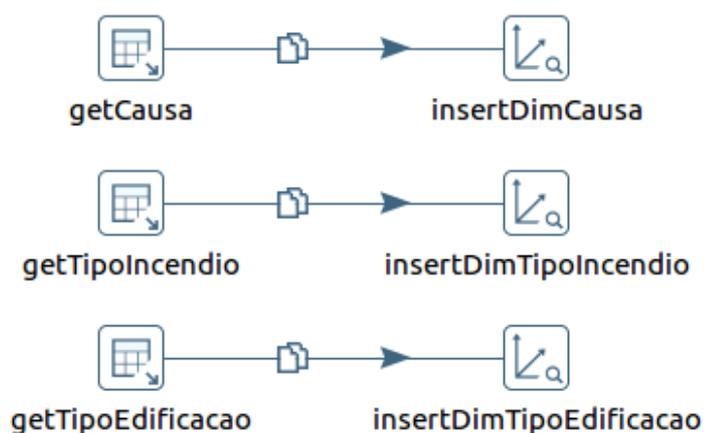
Consulta SQL:

```
SELECT
  id_tp_emergencia_subgrupo as id_tp_incendio
, nm_tp_emergencia_subgrupo as tipo_incendio
FROM tp_emergencia_subgrupo
where id_tp_emergencia = 1
```

Fonte: Elaborado pelo autor

Onde o “id_tp_emergencia = 1” refere-se a ocorrências de Incêndio.

Figura 6 - Carregamento de dimensões



Fonte: Elaborado pelo autor

4.5.4.4 Dimensões extraídas do código do Sistema Operacional de Origem

Os “tipo” produzidos a partir do código do Sistema Operacional de Origem, para os presente projeto foram :

- evento_causal;
- sub_causa;
- zona_origem.

Esses metadados foram extraídos do diretamente código php/html e convertido manualmente para código csv já sendo feito a correspondência id com o nome correspondente, para posterior carga no DM. Por se tratar de tabelas de classificação, os dados de tais tabelas são limitados e estão disponíveis no apêndice deste trabalho para consultas e melhor compreensão, em formato “csv”.

4.5.5 Conclusão do Planejamento

Tendo concluído todo o planejamento do projeto, a próxima etapa que é a fase inicial da implementação que é a criação de um Staging Area e a programação do processo ETL, de fato nesse projeto essas duas etapas foram feitas simultaneamente, tendo em vista a capacidade do sistema Pentaho Data Integration - PDI de automatizar ambos, assim a maior parte do trabalho descrito no próximo capítulo consistiu em aprender a dominar as ferramentas do PDI, alinhado com a teoria de um DM, seguindo o planejamento em questão. A medida que o processo ETL foi se desenvolvendo, e verificado inconsistências no projeto, este foi sendo corrigido ou atualizado de acordo com o observado no decorrer da execução.

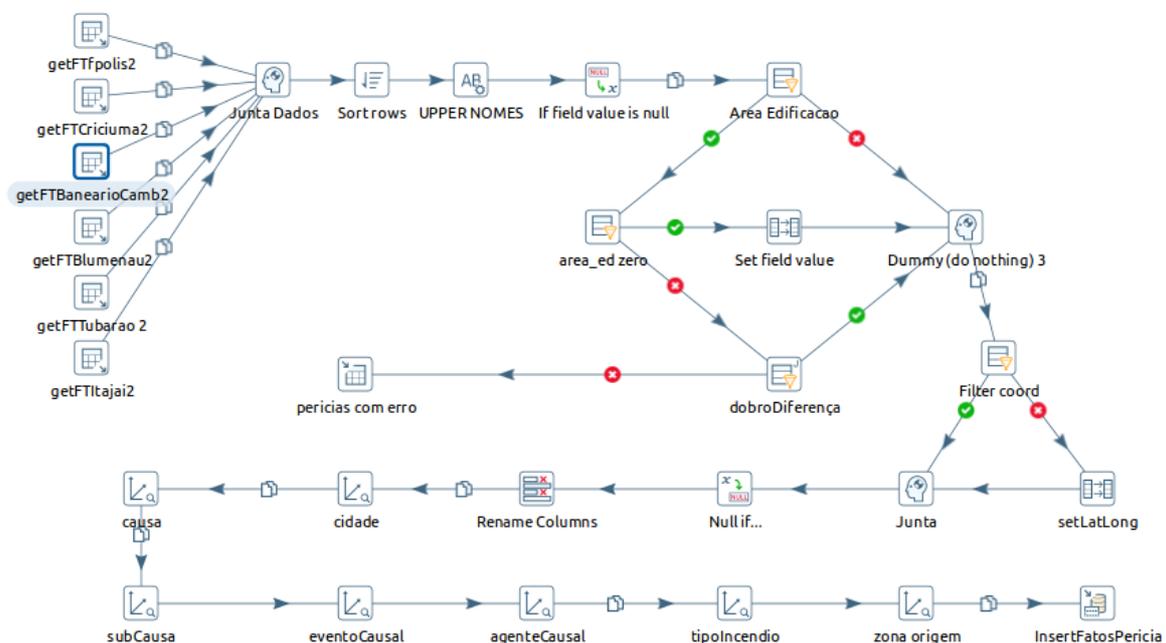
5 O PROCESSO ETL

O processo de Extract, Transform and Load - ETL, ou numa tradução livre: extrair, transformar e carregar, foi a etapa mais complexa do projeto, que demandou mais tempo, que conforme Inmon (2002, p. 283 e 284) é comum que seja gasto cerca de 80% do tempo do desenvolvimento de um projeto DW nessa fase, no presente caso ainda houve um tempo gasto no processo de aprendizagem dos conceitos teóricos envolvidos em ETL e nas ferramentas utilizadas: o Pentaho Data Integration – Kettle, além do tempo gasto efetivamente desenvolvendo o processo ETL.

5.1 PROCESSO ETL PARA O FATO PERÍCIA

Todo o processo ETL elaborado pelo autor para este estudo, do fato perícia será descrito a seguir, com base tendo como referência o que se vê na figura 7.

Figura 7 - Processo ETL para o Fato Perícia



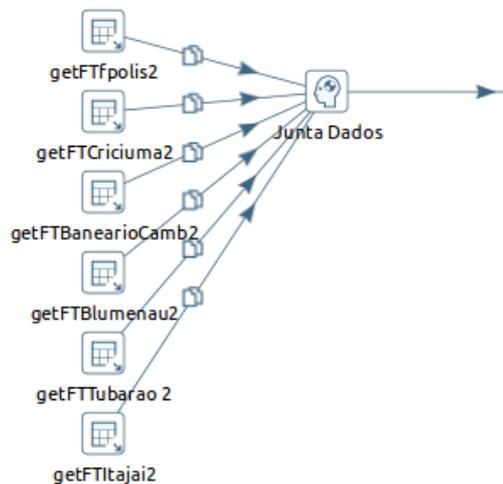
Fonte: Elaborado pelo autor

Na figura 7, vemos parte do processo ETL das bases de dados do sistema de perícia ao *Data Staging Area* (nomeado de datamart1). Neste caso, o processo ETL é referente aos dados periciais que visam alimentar a tabela fato perícia.

5.1.1 Extract

A primeira etapa é o extract, ou extrair. A extração está representado por 6 quadros, na figura 8, chamados getFT[NomeServidor], cada qual extraindo os dados de um servidor diferente. A função de extração de dados segue um protocolo de acesso ao banco de dados do tipo ODBC, que permite executar funções do tipo SQL.

Figura 8 - Extração de dados para o Fato Perícia



Fonte: Elaborado pelo autor

A consulta SQL para a extração de dados no Sistema Operacional de Origem, idêntica para todos os servidores, é a seguinte:

Consulta SQL:

```

SELECT
  aa.id_pericia,
  aa.id_ocorrencia,
  aa.id_cidade,
  aa.dt_criacao,
  aa.id_tp_pericia,
  a.id_tipo_incendio,
  a.ts_incendio,
  a.nm_codigo,
  a.nr_bbm_employado,
  a.fg_viatura_employada_compativel,
  a.fg_bbm_employado_compativel,
  a.nr_bc_employado,
  atr.tm_resposta_j10,
  atr.fg_tempo_resposta_j10_compativel,
  (atr.dt_termino_extincao + atr.hr_termino_extincao) - (atr.dt_inicio_extincao + atr.hr_inicio_extincao)
as tm_extincao,
  aac.id_desen_fogo_edificacao,
  i.ts_inicio,
  
```

```

i.ts_termino,
i.ts_termino - i.ts_inicio as tm_pericia,
id.vl_area_construida as vl_area_construida_ed,
idm.vl_area_edificacao,
idm.vl_area_atingida,
idm.vl_valor_salvado,
idm.vl_valor_prejuizo,
isi.fg_possui_seguro,
isi.id_causa,
isi.id_subcausa,
isi.id_evento_causal,
isi.id_classe_agente_causal,
isi.id_agente_causal,
isi.id_motivo_investigacao_ nao_realizada::boolean as investigacao_ nao_realizada,
isi.id_local_inicio,
isi.nm_foco_inicial,
oma.lat_bairro,
oma.long_bairro,
oma.lat_logradouro,
oma.long_logradouro,
cidades_bbm.nr_batalhao
FROM
public.aapericia aa
left join public.aapericia_fase_ativa a using (id_pericia)
left join public.cidades_bbm using (id_cidade)
left join public.aapericia_fase_ativa_tempo_resposta atr using (id_pericia)
left join public.aapericia_fase_ativa_atividade_compl aac using (id_pericia)
left join public.aapericia_fase_investigativa i using (id_pericia)
left join public.aapericia_fase_investigativa_dano_material idm using (id_pericia)
left join public.aapericia_fase_investigativa_seguro_incendio isi using (id_pericia)
left join public.aapericia_imovel_descricao id using (id_pericia)
left join public.ocorrencias_map_abertura oma using (id_ocorrencia)
where id_status = 9
order by id_pericia;

```

Fonte: Elaborado pelo autor

O resultados de todas as consultas são direcionados ao uma função “burra” ou “dummy” nomeada de “Junta Dados”, que apenas uni os resultados das consultas sem fazer mais nada. Destaca-se que o status 9, presente no filtro “where id_status = 9”, refere-se as investigações já concluídas, descartando-se assim as que estão em outras fases (rascunho, elaboração, aguardando homologação etc).

5.1.2 Transform

A parte de transformação dessa ETL é de longe a parte mais complexa de toda a pesquisa aqui realizada. Nesta seção será detalhado toda essa parte do processo, que para melhor compreensão, foi dividida em 5 passos:

5.1.2.1 Primeiro passo

Partindo da função “Junta Dados”, temos o “Sort rows” que ordena o resultado das consultas por “id_pericia” e “id_cidade”. A próxima função é a “UPPER NOMES” que transforma os textos das colunas “nm_codigo” e “nm_foco_inicial” para deixar todos os caracteres maiúsculas. O próximo passo é o “If field value is null” recebe os campos selecionados e verifica se o valor é “null”, caso positivo troca o valor para outros selecionado, isso é importante pois na modelagem de um DM, evita-se trabalhar com valores nulos para dimensões, conforme a figura 9:

Figura 9 - “If fiel value is null”

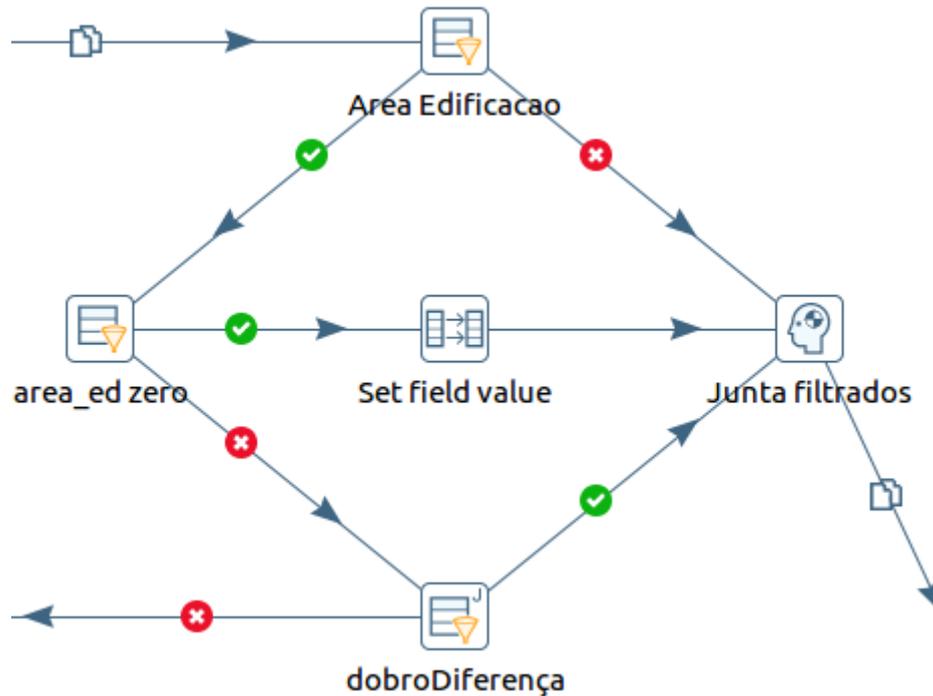
Field	Replace by value	Conversion mask (Date)	Set empty string?
1 fg_bbm_employado_compativel	0		N
2 fg_possui_seguro	0		N
3 fg_possui_vitima	0		N
4 fg_tempo_resposta_j10_compativel	0		N
5 fg_viatura_employada_compativel	0		N
6 id_tp_pericia	0		N
7 nm_codigo	NÃO INFORMADO		N
8 nm_foco_inicial	NÃO INFORMADO		N
9 tm_resposta_j10	0001/01/01 00:00:00.000000000		N
10 ts_incendio	1970/01/01 00:00:00.000000000		N
11 ts_inicio	0001/01/01 00:00:00.000000000		N
12 ts_termino	0001/01/01 00:00:00.000000000		N

Fonte: Elaborado pelo autor

5.1.2.2 Segundo passo.

Os próximos passos são uma verificação comparativas das áreas informadas da edificação investigadas, conforme a figura 10, que é inserida no sistema mais de uma vez, ocorre que em muitos casos foi identificado divergências entre estes valores, na ordem de 10^2 , isso se explica devido ao usuário não compreender corretamente o preenchimento dos dados, resultando por fim em relatórios incorretos, necessário nesse ponto eliminar esses dados. Entretanto, se o usuário em qualquer momento corrigir as inconsistências, numa nova carga dos dados, o dado será incluso normalmente. Nesse ponto ainda ressaltar que é possível criar relatórios com os dados inconsistentes para que os gestores faça as correções necessárias, mas não é objeto desse trabalho.

Figura 10 - Limpeza de dados inconsistentes



Fonte: Elaborado pelo autor

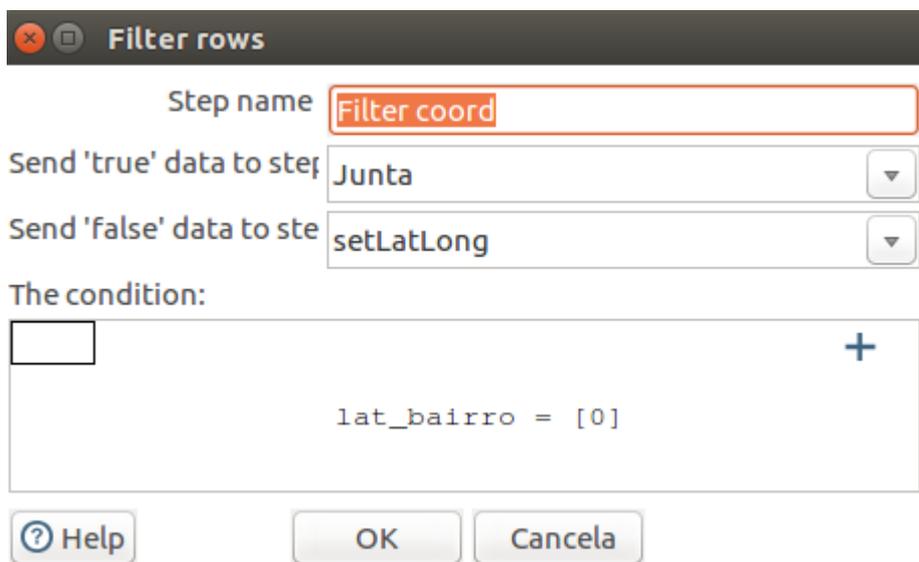
Os dados são recebidos em “Area Edificacao” onde são filtrados os dados não vazios e diferentes entre “vl_area_edificacao” e “vl_area_construida_ed” que deveriam ser idênticos. Os dados que atendem a esses dois critérios seguem para “area_ed zero” que filtra apenas “vl_area_edificacao = 0”. Quando o retorno é true, o valor zero é substituído por vl_area_edificacao. Se não é nulo segue para dobroDiferença, em que é verificado se os valores entre “vl_area_edificacao” e “vl_area_construida_ed” são mais que o dobro entre si. Foi utilizado esse recurso para evitar remover dados que eventualmente foram aproximados pelo usuário. Assim, são eliminados apenas os dados com inconsistência maior que o dobro entre si. Como resultados foi possível constatar que alguns dados com áreas absurdas, algumas com mais de 100.000,00 m² foram eliminadas. Tornando os dados estatísticos mais fidedignos para o gestor. E foi repassado a divisão de perícia a lista das investigações com esses dados para a correção.

Procedimentos equivalentes de verificação da qualidade dos dados é possível de ser feito em vários outros dados, mas foge dos objetivos deste trabalho a gestão e qualidade de dados, nesse ponto específico foi necessário devido a grave distorção que uns poucos dados foram capazes de provocar nos resultados preliminares do trabalho.

5.1.2.3 Terceiro passo

O “*filter coord*”, figura 11, é uma função de filtragem com base em lógica do tipo “se, então”, que combinado com `setLatLng`, serve para melhorar a qualidade dos dados geográficos. Foi constatado que em muitos casos não fornecido pelo usuário as coordenadas do local, mas foi fornecido as coordenadas do bairro. Assim optou-se, nesses casos que é melhor utilizar as coordenadas do bairro do que nenhuma coordenada.

Figura 11 - Filtra coordenadas



Fonte: Elaborado pelo autor

Ainda trabalhando com o mesmo dado, as coordenadas, as mesmas submetidas a função “*null if...*” que verifica uma condição, no caso que a coordenadas sejam zero e zero, que é um valor válido para coordenadas terrestres, e são transformadas em valores nulos para não impactar na renderização de mapas que eventualmente venham a ser utilizados.

5.1.2.4 Quarto passo

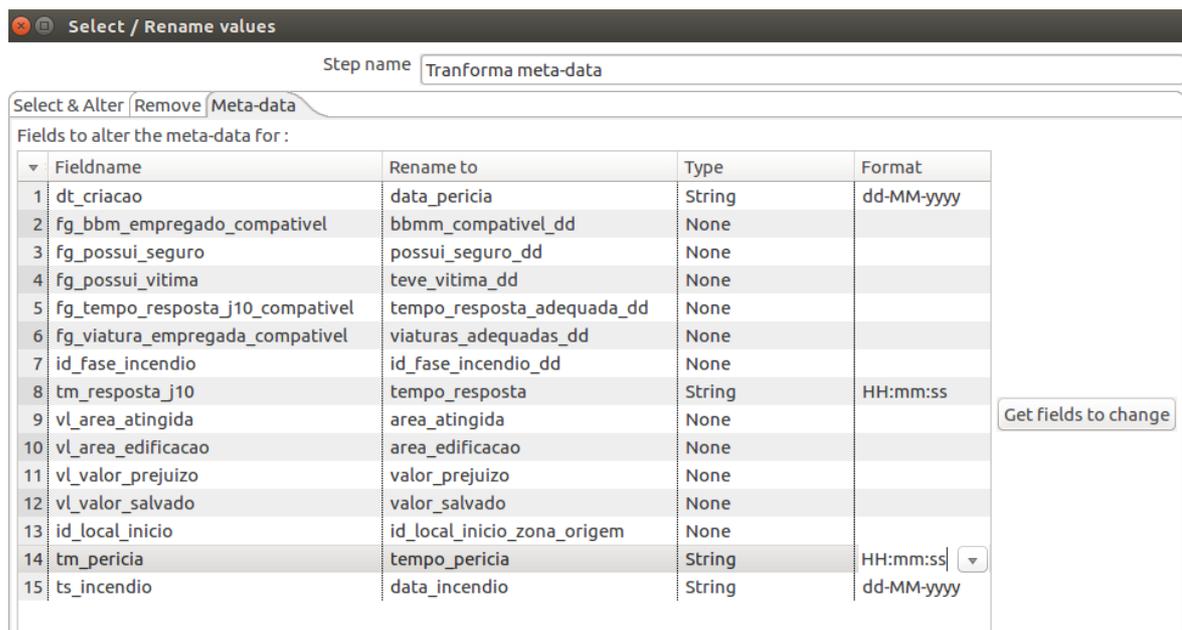
Após as validações de dados já descritos acima, neste passo é feito na função chamada de “*Transforma meta-data*”, figura 12, em uma transformação de metadados, ou seja, nesse passo o nome das colunas são renomeadas para se harmonizar conforme o nome das métricas, dimensões ou “*Surrogate key*” do projeto.

Na figura 12, é possível verificar tais transformações. Muito embora seja possível a transformação dos campos “*ts_incêndio*” e “*dt_criação*”, optou-se por transformar para o tipo

string, mas com formato “*dd-MM-yyy*”, para que haja compatibilidade com a formatação da dimensão tempo, que tem esta estrutura e que será a utilizada para recuperar a “Surrogate Key” da data.

Os demais campos foram apenas feitas adequadas nos seus respectivos nomes. Em alguns casos foi utilizado, como metadado, o sufixo “*_dd*” para marcar que se trata de uma dimensão degenerada.

Figura 12 - Transforma metadados



Fonte: Elaborado pelo autor

5.1.2.5 Quinto passo

O quinto passo é executado de funções do tipo “*Dimension lookup/update*” para as seguintes dimensões: “*cidade*”, “*causa*”, “*subCausa*”, “*eventoCausal*”, “*tipoIncêndio*”, “*zonaOrigem*”, “*dataIncendio*” e “*dataPericia*”, conforme se vê na figura 13.

Essa função processa os dados para substituir a chave primária pela respectiva “*Surrogate Key*” da dimensão alvo. Essa “*Surrogate Key*” é registrada numa nova coluna na tabela fato, com metadado “*id_[nome]_sk*”, e ao fim é apagada a coluna da chave primária do dado dimensional original daquela transação, tendo em vista que esta informação ficará registrada na própria dimensão, conforme prevê o modelo dimensional.

Figura 13 - “Dimension lookup/update”



Fonte: Elaborado pelo autor

5.1.3 Load FTPerícia

A última etapa, figura 14, do processo ETL é o carregamento, ou “load”, onde os dados transformados são efetivamente carregados para o Stage Area. Essa ação é realizada pela função “InsertFatosPerícia”.

Figura 14 - “Load” FTPerícia

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates:

The key(s) to look up the value(s):

Table field	Comparator	Stream field1	Stream field2
1 id_pericia	=	id_pericia	
2 id_ocorrencia	=	id_ocorrencia	
3 id_cidade_sk	=	id_cidade_sk	

Update fields:

Table field	Stream field	Update
1 id_tp_pericia	id_tp_pericia	Y
2 dt_incendio_sk	data_incendio_sk	Y
3 nm_codigo	nm_codigo	Y
4 nr_bbm_employado	nr_bbm_employado	Y
5 viaturas_adequadas_dd	viaturas_adequadas_dd	Y
6 bbmm_compativel_dd	bbmm_compativel_dd	Y
7 nr_bc_employado	nr_bc_employado	Y
8 tempo_resposta	tempo_resposta	Y
9 tempo_resposta_adequada_dd	tempo_resposta_adequada_dd	Y
10 id_fase_incendio_dd	id_fase_incendio_dd	Y

Fonte: Elaborado pelo autor

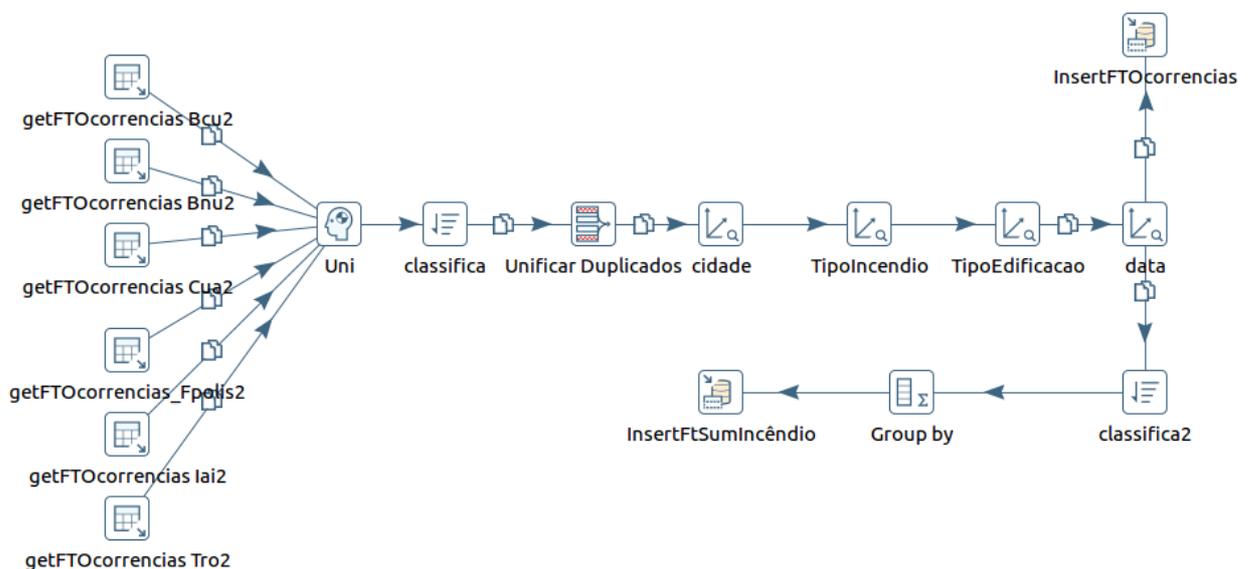
Na figura 14 vemos o ambiente de configuração da etapa “InsertFatosPerícia”. Para evitar duplicação de dados, essa função executa uma busca por comparação com as chaves primárias/“Surrogate key” selecionadas: “id_perícia”, “id_ocorrencia” e “id_cidade_sk”. Destaca-se que a “surrogate key” “id_cidade_sk” é necessária devido a haver dados de vários

servidores com possíveis sobreposições de chaves primárias. Na mesma figura, temos uma área denominada “*Update fields*”, que são os campos que serão efetivamente carregados e/ou atualizados para cada transação.

5.3 PROCESSO ETL PARA O FATO OCORRÊNCIA

Embora fosse possível fazer o processo ETL para a tabela fato ocorrência de forma conjunta ou até única para a tabela fato ocorrência, optou-se separar para facilitar a organização e melhor controlar o fluxo de dados. Abaixo descreve-se esse ETL. Na figura 15 temos uma visão geral dessa transformação:

Figura 15 - Processo ETL para o Fato Incêndio

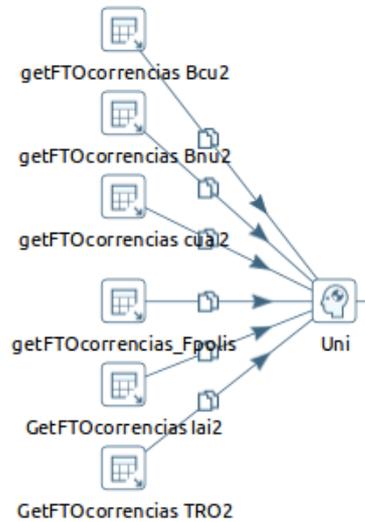


Fonte: Elaborado pelo autor

5.3.1 Extract

Essa extração está representado por 6 quadros, na figura 15, chamados getFTOcorrencia[NomeServidor], cada qual extraíndo os dados de cada um dos servidores selecionados, conforme o mostrado na figura 16. (o número 2 ao final do nome do servidor, refere-se ao servidor reserva).

Figura 16 - Extração de dados para o Fato Ocorrência



Fonte: Elaborado pelo autor

De forma semelhante ao primeiro ETL, o processo inicia-se consultando cada um dos servidores da 1ªRBM, conforme e junta todos os retornos na função “Uni”.

Consulta SQL:

```
SELECT
distinct
  id_ocorrenci
, id_cidade
, ts_ocorrenci
, id_tp_emergenci_subgrupo
, id_tp_edificac
, to_char( ts_ocorrenci, 'dd-MM-YYYY') as data_ocorrenci
, nr_batalhao as id_bbm
FROM ocorrencias o
  left join historico_emergenci_subgrupo hes using (id_ocorrenci)
  left join emergenci_incendio_tp_edificac using (id_ocorrenci)
  left join cidades_bbm using(id_cidade)
where ts_ocorrenci > '2017-01-01'
and o.id_tp_emergenci = 1
order by 1
```

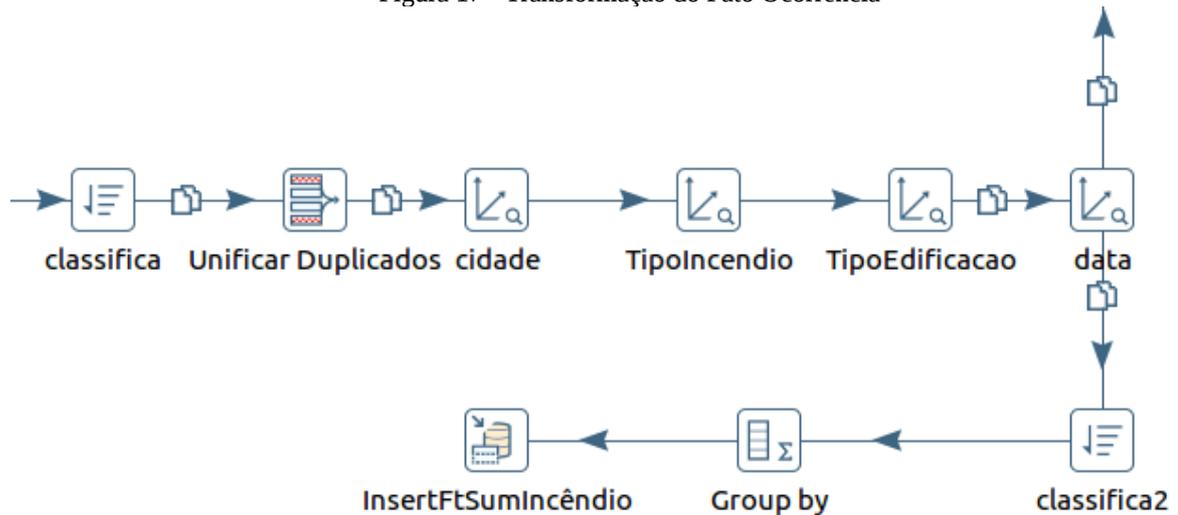
Fonte: Elaborado pelo autor

5.3.2 Transform

Conforme se vê na figura 17, a transformação começa após a unificação dos dados obtidos que são mandados para a função unificar Duplicados, tendo em vista que o Sistema de Origem gera uma linha na consulta de ocorrências para cada viatura despachada, ainda que com a mesma chave composta (“id_ocorrenci”, “id_cidade”), como para o objetivo deste trabalho, não importa o detalhamento individual de cada viatura, nesse ponto essas linhas são

unificadas. Na sequência, temos as funções: “*cidade*”, “*TipoIncêndio*”, “*TipoEdificacao*” e “*data*”. Essas funções fazem o mesmo já explicado na sessão 5.1.2.3 deste trabalho, recuperando as respectivas “*Surrogate Key*”. Após a data os dados são enviados para carregar (para cima) a tabela fato ocorrência e para outras transformações (para baixo) antes de ser inseridas na tabela fato sumarizada ocorrência.

Figura 17 - Transformação do Fato Ocorrência



Fonte: Elaborado pelo autor

A partir de *data*, a seta para baixo, encaminha os dados para a função “*classifica2*”, que é um pré-requisito do sistema, para efetuar adequadamente a função seguinte, “*Group by*” cuja caixa de diálogo vemos a figura 18.

A função “*Group by*” agrupa as linhas baseado nas semelhanças entre elas, comparando-se as colunas listadas no campo “*Group Field*”: “*ano*”, “*mes*”, “*nr_batalhao*”, “*id_cidade_sk*”, “*id_tp_incendio_sk*”, “*id_tp_edificacao_sk*”.

Figura 18 - Caixa de Diálogo do “Group By”

Step name

Include all rows?

Temporary files directory

TMP-file prefix

Add line number, restart in each group

Line number field name

Always give back a result row

The fields that make up the group:

Group field		<input type="button" value="Get Fields"/>
1	ano	
2	mes	
3	nr_batalhao	
4	id_cidade_sk	
5	id_tp_incendio_sk	
6	id_tp_edificacao_sk	

Aggregates :

Aggregates :			<input type="button" value="Get lookup fields"/>
Name	Subject	Type	
1 total_ocorrencias	id_ocorrencia	Number of rows (without field argument)	

Fonte: Elaborado pelo autor

5.3.3 Load FTOccorrencias

Também nesse processo ETL (ver figura 17), após a função data, para cima temos o carregamento da tabela fato ocorrências com a seguinte estrutura conforme a figura 19, onde vemos o nome da etapa *InsertFTOccorrencia*, o nome da tabela conforme está na “*Staging Area*”, as chaves para buscar por comparação (“*id_ocorrencia*” e “*id_cidade_sk*”) e abaixo os primeiros campos de dados para inserção ou atualização (quando os id são compatíveis) da tabela.

Figura 19 - “Load” do Fato Ocorrência

Insert / Update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates:

The key(s) to look up the value(s):

Table field	Comparator	Stream field1	Stream field2
1 id_ocorrencia	=	id_ocorrencia	
2 id_cidade_sk	=	id_cidade_sk	

Update fields:

Table field	Stream field	Update
1 id_ocorrencia	id_ocorrencia	Y
2 data_ocorrencia_sk	data_ocorrencia_sk	Y
3 id_cidade_sk	id_cidade_sk	Y
4 id_tp_incendio_sk	id_tp_emergencia_subgrupo	Y
5 id_tp_edificacao_sk	id_tp_incendio_sk	Y
6 id_bbm	id_bbm	Y

Fonte: Elaborado pelo autor

5.3.4 Load FTSumOcorrencia

Nesse processo ETL, figura 17, existem duas saídas após a função “data”, para cima temos o carregamento da tabela fato ocorrências com a seguinte estrutura conforme a figura 19.

Figura 20 - "Load" do Fato Sumarizado Incêndio

Insert / Update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates:

The key(s) to look up the value(s):

	Table field	Comparator	Stream field1	Stream field2
1	id_cidade_sk	=	id_cidade_sk	
2	id_tp_incendio_sk	=	id_tp_incendio_sk	
3	id_tp_edificacao_sk	=	id_tp_edificacao_sk	
4	ano	=	ano	
5	mes	=	mes	

Update fields:

	Table field	Stream field	Update
1	id_cidade_sk	id_cidade_sk	N
2	id_tp_incendio_sk	id_tp_incendio_sk	N
3	total_ocorrencias	total_ocorrencias	Y
4	ano	ano	N
5	mes	mes	N
6	id_tp_edificacao_sk	id_tp_edificacao_sk	N
7	nr_batalhao	nr_batalhao	N

Fonte: Elaborado pelo autor

6 FERRAMENTAS DE ACESSO AOS DADOS

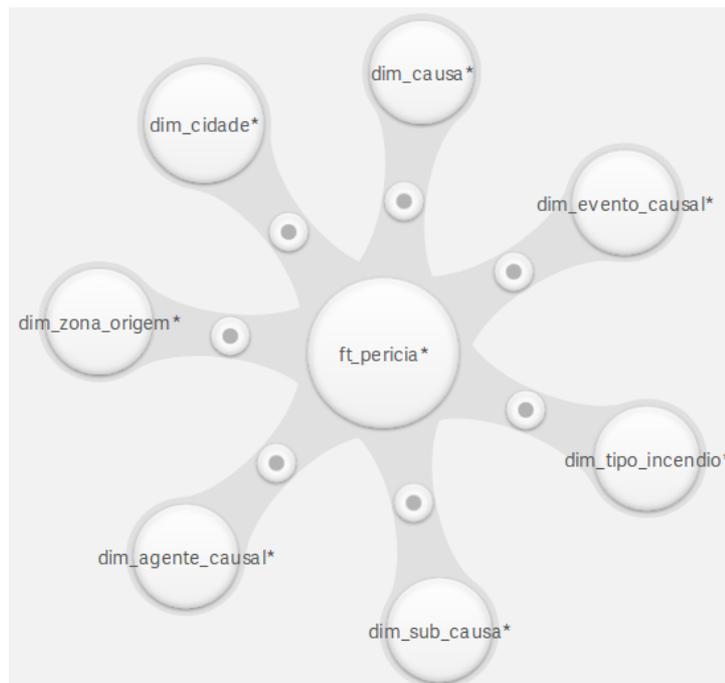
Neste capítulo estaremos abordando a implementação da ferramenta de visualização dos dados para o usuário utilizando-se o QLIK SENSE, ferramenta adotada pela SSP através do projeto Analytics.

6.1 ARQUITETURA FRONT-END

Segundo Marzarotto (2019, p.74) “O front-end é a área de apresentação para operacionalizar o DW.” Essa área fica organizada e sempre disponível para uso dos usuários finais. Esse é o último componente de um DW ou DM, pois é a ferramenta que vai efetivamente permitir acesso aos dados.

A acesso dos dados no Qlik Sense se dá através de uma conexão tipo ODBC para a “Staging Area”. Uma vez conectado com a “Staging Area”, é possível escolher as tabelas e colunas a serem inseridas do sistema, de fato essa pode ser considerada como a última etapa do processo ETL. Marzarotto (2019, p. 74) destaca que “O software Qlik Sense dispõe de todas as ferramentas para desenvolver aplicações OLAP e suas operações, sendo em geral simples de usar.”

Figura 21 - Esquema Estrela do Fato Perícia



Fonte: Elaborado pelo autor

Uma vez inseridos os dados é necessário fazer as correlações entre os fatos e as dimensões, o sistema Qlik Sense permite realizar esse procedimento de forma intuitiva e gráfica, cujo resultado parcial pode-se ver na figura 21, onde fica claro o resultado na forma prevista do esquema estrela, apresentado nos capítulos anteriores.

6.2 CONFIGURAÇÃO DE DADOS MESTRES

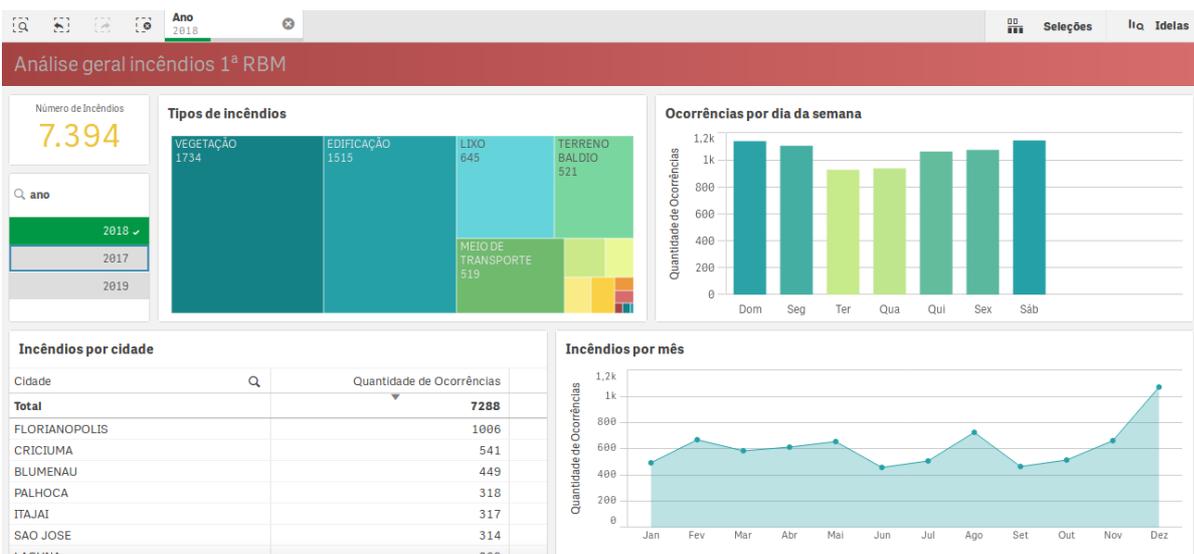
O sistema Qlik Sense permite a marcação dos dados em categorias chamados de dados mestres pelo sistema, que é a definição que quais dados são medidas, e como essa medida é feita, (é uma soma, média, contada, etc), e quais dados são dimensões. Essa marcação prévia permitirá que o usuário final com menor treinamento seja capaz que configurar os gráficos que deseja criar, cruzando quaisquer dados existentes no DM com maior probabilidade de sucesso, maior rapidez e facilidade, de fato em poucos segundos é possível montar um gráfico complexo.

6.3 GRÁFICOS E DASHBOARD

A fim de alcançar o terceiro objetivo desse projeto foi montado uma série de gráficos a partir das métricas e dimensões do DM. O sistema tem diversos tipos de gráficos que podem ser incluídos, tal como gráfico pizza, colunas, linha, mapa, dispersão, etc.

Na imagem abaixo gráficos obtidos:

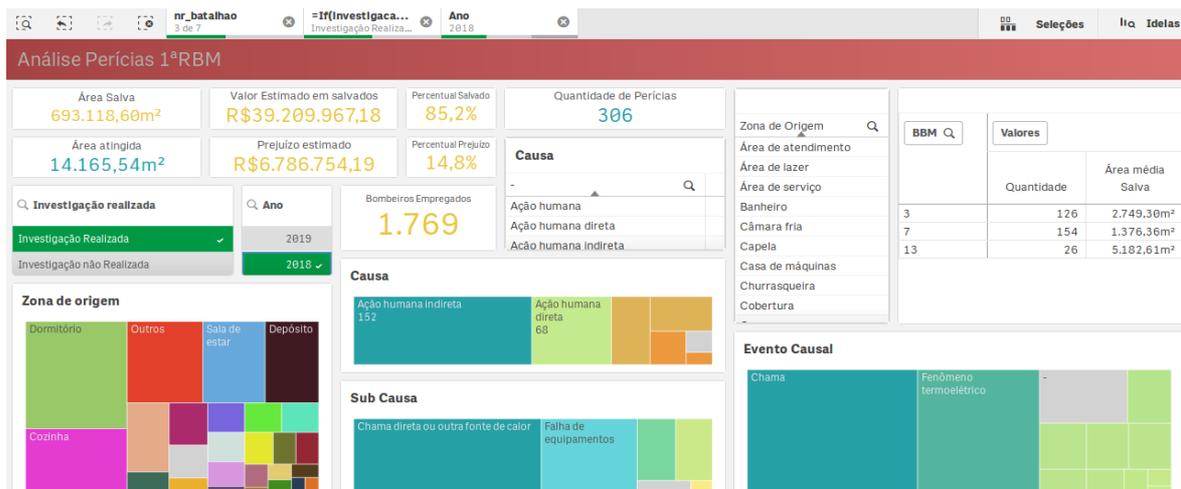
Figura 22 - Dashboard 1



Fonte: Elaborado pelo autor

Na figura 22, vemos 3 gráficos e uma informação numérica e um seletor e uma tabela. Nesse gráfico foi selecionado o ano 2018, automaticamente todos os gráficos se ajustam, filtrando apenas os dados referentes ao respectivo ano. Importante destacar que o gráfico de coluna, a tabela e o mapa de árvores, são gráficos que também funcionam como seletores, que permitem por exemplo selecionar a coluna do domingo e automaticamente todos os demais gráficos serão ajustados para mostrar apenas os respectivos dados.

Figura 23 - Dashboard 2

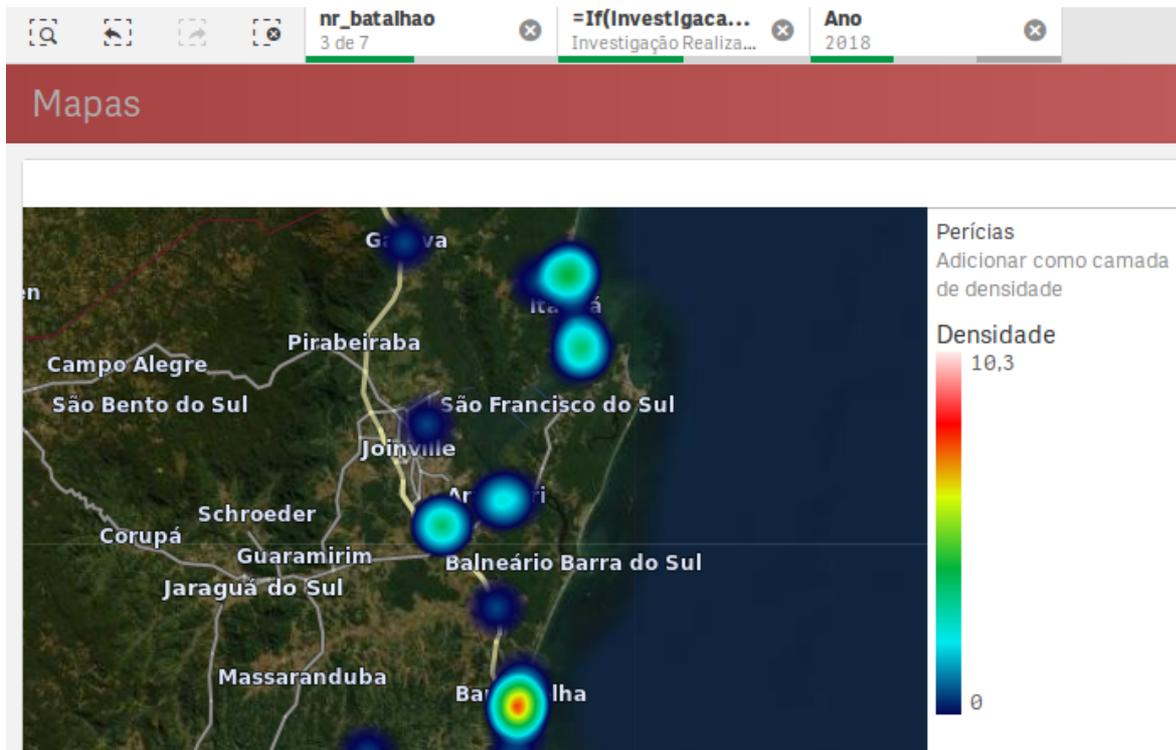


Fonte: Elaborado pelo autor

No dashboard mostrado na figura 23, selecionamos apenas as informações onde houve investigação na fase pericial, e selecionamos 3 batalhões, o 3ºBBM, o 7ºBBM e o 13ºBBM, que são os BBM localizados no litoral norte de SC, para essa seleção pode-se ver os valores estimados de salvados e de prejuízos, área salva e área atingida, e é possível ver que as principais zonas de origem para essa seleção foram dormitório e cozinha, a principal causa foi humana indireta, a principal subcausa e evento causal também são possíveis de serem identificados usando como critério o número de incêndios investigados. Mas é possível usar outros critérios para avaliar o impacto de cada item de investigação usando por exemplo a área queimada ou quantidade total de BBMM empregados.

O sistema também permite a utilização de ferramentas de mapeamento com zonas de calor, que embora não fosse um objetivo deste trabalho boa parte das ocorrências fornecem esse dado e foi incorporado ao DM sem grandes pretensões senão demonstrar algumas possibilidades como na figura 24:

Figura 24 - Dashboard 3



Fonte: Elaborado pelo autor

O mapa mostrado na figura 24 acima mostra as ocorrências de incêndios conforme filtros determinados pelo usuário.

Importante destacar que não é objetivo deste trabalho exaurir as possibilidades de gráficos com os dados disponibilizados, mesmo porque as possibilidades são limitadas apenas pela necessidade, curiosidade e criatividade da corporação.

7 CONCLUSÃO

Esta pesquisa se propôs a Modelar e Implementar um Data Mart em contribuição para a gestão do conhecimento para uma instituição pública e militar, que produz uma grande quantidade de conhecimento em destaque nas investigações de incêndio, mas esses dados são armazenados em bancos de dados, sem uma metodologia eficaz para tirar o máximo de benefício em desses dados e informações em favor da instituição, como na retroalimentação do ciclo operacional de bombeiro. Neste ambiente é apresentado a metodologia “*Business Intelligence*” cujas soluções tecnológicas, já estão bem estabelecidas no ambiente empresarial privado para problemas semelhantes.

Na sequência estão expostos as conclusões mais importantes e sugestões para trabalhos futuros.

7.1 CONCLUSÕES

A pesquisa apresentou um ciclo de modelagem de um ambiente de Data Mart para o CBMSC, mais notadamente dirigido para satisfazer a necessidade de produção de informações de subsídio do processo de retroalimentação do ciclo operacional de bombeiro pela geração de relatórios “*ad hoc*” a partir dos dados gerados das investigações de incêndio.

Para esse fim escolheu-se iniciar por uma pesquisa bibliográfica centrada na teoria de Data Warehouse e Data Mart, parte da metodologia de “*Business Intelligence*”. Foram explanados a partir dos conceitos básicos, passando por arquiteturas de construção como o bottom-up e apresentação de ferramentas de construção. Também foi apresentado o ciclo operacional de bombeiro e discutindo brevemente a gestão do conhecimento na corporação diante da necessidade da retroalimentação do referido ciclo operacional de bombeiro.

Tendo feito todas estas fundamentações teóricas a partir de revisão bibliográficas é apresentado metodologia escolhida para se efetuar o presente estudo de caso, a infraestrutura utilizada, apresenta-se o detalhamento do sistema operacional de origem, a Staging Area e por fim o planejamento do projeto, este feito com base na metodologia adotada.

Da análise do sistema operacional de origem, combinado com os Parâmetros para relatórios Módulo Perícia (Anexo B) foi elaborado um modelo dimensional que converteu-se num esquema estrela com 9 dimensões, 7 dimensões degeneradas e 3 tabelas fatos.

Já na fase de execução do estudo de caso, este se dá com a efetiva implementação do Data Mart proposto, iniciando com processo ETL (Extract, Transform and Load), que é a

etapa mais complexa de um projeto dessa natureza. Esse processo, basicamente, busca dados brutos nos sistemas operacionais de origem, executa funções de limpeza, cálculos, formatação de metadados, e por fim carga desses dados transformados no Staging Area, que é onde os dados ficam armazenados antes de serem carregados no DM propriamente dito. Os dados utilizados foram dos anos 2017 a 2019, entretanto o novo sistema de perícia somente começou a ser utilizado em 2018, eis porque os dados das investigações se limitam a partir desse ano. Foi utilizado dados de todos os BBM da 1ªRBM, porém o trabalho, uma vez concluído, permite a inclusão posterior de todos os demais BBM do estado, sem grandes complexidades.

Com os dados carregados e o modelo dimensional concluídos, os dados foram carregados para a ferramenta de apresentação de dados, no caso, sendo utilizado o software Qlik Sense, disponibilizado pela SSP ao CBMSC. Essa ferramenta permite a criação de painéis com apresentação de gráficos que interagem entre si. Capacitando um usuário leigo em consultas a bancos de dados, elaborar relatórios “*ad hoc*” que somente seria possível através de complexas consultas ao BD, que atualmente se dão por meio de solicitações dos gestores ao Divisão de Tecnologia e Informação, que além de ser um processo moroso e custoso, com frequência quando a informação chega, já pode não ser mais pertinente.

Estes painéis permitem uma análise suficientemente completa, não se limitando a apenas dos requisitos definidos no projeto, mas permite qualquer combinação de dados, em vários tipo de gráfico pré customizados, bastando que o próprio gestor/usuário defina quais critérios serão mais pertinentes no processo de tomada de decisão, assim como é possível elaborar pesquisas nos dados estatísticos, como parte do processo de retroalimentação do ciclo operacional de bombeiro. Além disso, esse DM permite realizar a seleção de investigações relacionados a determinada combinação de filtros, para se realizar pesquisas mais profundas realizando uma análise comparativa em investigações com características semelhantes. Esse processo, em tese, permitirá a produção de novos conhecimentos que retroalimentam o ciclo operacional de bombeiro, que de outro modo se tornaria extremamente complexo.

Ao longo do trabalho verificou-se que é possível a implementação de forma sistemática na instituição da metodologia de “BI” tendo em vista que o processo de aprendizagem dos conceitos básicos ligados ao BI e a sua implementação progressiva foram possíveis mesmo de forma autodidata, com escasso tempo para tal. Notório dizer que a instrução através de cursos podem acelerar o processo de aprendizagem daqueles que vierem a atuar nesta área.

A implementação de Data Mart e outras áreas de interesse da corporação podem vir a ser meio de aumentar a capacidade de tomada de decisão subsidiado em informação confiável e a um menor custo e menor tempo do que ocorre hoje.

Como sugestão para estudos futuros, fica a necessidade de um estudo relacionado à Governança de Dados, que abrange a confiabilidade e qualidade dos dados inseridos nos sistemas de origem que termina nos DM e num DW. Além disso, estudos relacionados a BIG DATA e Aprendizagem de Máquina são assuntos que merecem estudos. Por fim, a ampliação do presente DM para todos os dados disponíveis nos sistema de perícia e para todos os servidores, além de um estudo que permita a combinação dos dados do antigo sistema de perícia com o atual se mostram igualmente necessários.

REFERÊNCIAS

- ACORDI, C. F. **Gestão do Conhecimento em Organizações Militares**: Um Estudo de Caso na Atividade de Investigação de Incêndios. Orientador: Dra. Gabriela Gonçalves Silveira Fiates. 2011. 177 f. Dissertação (Mestrado em Administração) - Universidade do Sul de Santa Catarina, Florianópolis, 2011. Disponível em: https://biblioteca.cbm.sc.gov.br/biblioteca/index.php/component/docman/doc_download/456-charles-fabiano-acordi. Acesso em: 17 nov. 2019.
- BARBIERI, Carlos. **BI2 - Business Intelligence**: Modelagem & Qualidade. Rio de Janeiro: Elsevier Editora Ltda, 2011.
- BARBOZA JÚNIOR, Oscar Washington. **Estudo sobre a implementação do Centro de Estatística de Incêndios do Corpo de Bombeiros Militar de Santa Catarina**. 2012. Monografia (Curso de Formação de Oficiais) - Corpo de Bombeiros Militar de Santa Catarina. Florianópolis, 2012.
- BOMFIM, Marcus Mosquéra. **A implementação e utilização de Data Warehouse em Instituições Públicas no Brasil**: um estudo descritivo das implicações envolvidas. 2001. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis. **Arquivo digital**: <http://repositorio.ufsc.br/xmlui/handle/123456789/81700>
- CBMSC, **Plano Estratégico 2018 – 2030**. Santa Catarina: Corpo De Bombeiros De Santa Catarina, 2018.
- CBMSCb, **Manual de Capacitação em incêndio estrutural**: Corpo de Bombeiros Militar de Santa Catarina. Organizado por Marcos Alves da Silva. Florianópolis, 2018.
- ELIAS, Diego. **Conhecendo a arquitetura de Data Warehouse**. Canaltech, 8 abr. 2014. **Disponível em**: <https://canaltech.com.br/business-intelligence/conhecendo-a-arquitetura-de-data-warehouse-19266/>. Acesso em: 3 jun. 2019.
- INMON, W. H. **Como Construir o Data Warehouse**. Rio de Janeiro, Campos. 1997.
- KIMBALL, R. **Data Warehouse Toolkit**: Técnicas para Construção de Data Warehouse Dimensionais. São Paulo: Makron Books, 1998.
- LEMONS, Luiz Felipe. **A Gestão das investigações periciais em incêndios realizadas no Corpo de Bombeiros Militar de Santa Catarina**. 2017. Monografia (Curso de Comando e Estado Maior) – Corpo de Bombeiros Militar de Santa Catarina, Centro de Ensino Bombeiro Militar, Curso de Comando e Estado Maior, 2017. **Arquivo digital**: https://gnuteca.cbm.sc.gov.br/html/file.php?folder=material&file=ccem_2017_lemos.pdf
- MARANGONI, Priscila. SCHENEIDE, Daniel. **Modelagem de um Data Mart**: Estudo de Caso na PMSC, 2010. Monografia (Graduação em Sistemas de Informação) – UFSC. **Arquivo digital**: <https://repositorio.ufsc.br/handle/123456789/184528>.

MAUS, Álvaro. **Proteção contra incêndio**: atividades técnicas no Corpo de Bombeiros: Teoria geral. Florianópolis: Editograf, 1999.

MARZAROTTO, Diego Felipe. **Modelagem e implantação de um Data Mart**: estudo de caso no CBMSC. Orientador: José Leomar Todesco. 2018. 162 p. Trabalho de Conclusão de Curso (Bacharel em Sistemas de Informação) - UFSC, **Arquivo digital**: <https://repositorio.ufsc.br/handle/123456789/192163>.

MODOLON, D. M.; VIDAL, D. N. **Análise Estatística Comparada entre Edificações Regularizadas e Não Regularizadas Passíveis de Fiscalização pelo CBMSC**. Artigo Acadêmico (Especialização em Perícia de Incêndio e Pânico) - Corpo de Bombeiros Militar de Santa Catarina, Centro de Ensino Bombeiro Militar, Florianópolis, 2019. **Arquivo digital**: https://biblioteca.cbm.sc.gov.br/biblioteca/index.php/component/docman/doc_download/836-darlan-margotti-modolon-

PITON, Rafael. **Data Warehouse Passo a Passo**: o guia prático de como construir um data warehouse do zero. Porto Alegre: Raizzer, 2018.

SARTE, Anderson Medeiros. **Perícia de incêndio**: uma abordagem sobre a coleta de amostras sólidas e líquidas em edificações sinistradas pelo fogo. 2009. Monografia (Tecnologia em Gestão de Emergências) - Universidade do Vale do Itajaí, Centro de Ciências Tecnológicas da Terra e do Mar **Arquivo digital**: https://gnuteca.cbm.sc.gov.br/html/file.php?folder=material&file=cfo_2009_sarte.pdf

SILVA, Ana Paula. **Data Warehouse e Data Mart como Ferramentas de Inteligência em Negócios (BI)**. 2004. Monografia (Especialista em Desenvolvimento para WEB)- Departamento de Informática, Universidade Estadual de Maringá, Maringá, 2004. **Arquivo digital**: <http://www.espweb.uem.br/site/files/tcc/2003/Ana%20Paula%20Silva%20-%20Data%20Warehouse%20e%20Data%20Mart%20como%20Ferramentas%20de%20Inteligencia%20de%20negocios.pdf>

SINGH, Harry S. **Data Warehouse**: Conceitos, Tecnologias, Implementação e Gerenciamento. Tradução: Mônica Rosember. São Paulo: Makron Books, 2001.

SOUSA, Carlos Alberto. Chefe do Centro de Processamento de Dados - CPD/DLF/CBMSC. **Resolução do Atendente SAU-2925/DITI/2019** [mensagem pessoal]. Mensagem recebida por: <miguel@cbm.sc.gov.br> em 29 maio 2019.

TEIXEIRA, Elizabeth. **As três Metodologia**: Acadêmica, da Ciência e da Pesquisa. 8. ed. - Petrópolis, RJ: Vozes, 2011.

TEIXEIRA, José Ricardo. **Introdução à tecnologia Data Warehouse**. DevMedia, Rio de Janeiro - RJ, 2013. **Disponível em**: <https://www.devmedia.com.br/introducao-a-tecnologia-data-warehouse/27629>. Acesso em: 2 set. 2019.

TIMMERMANN, M. A. S.; ACORDI, C. F. **Investigação de Incêndios como Ferramenta para a Produção de Conhecimento Organizacional**. Artigo Acadêmico (Especialização em Perícia de Incêndio e Pânico) - Corpo de Bombeiros Militar de Santa Catarina, Centro de Ensino Bombeiro Militar, Florianópolis, 2019. **Arquivo digital:** https://biblioteca.cbm.sc.gov.br/biblioteca/index.php/component/docman/doc_download/843-marco-aurelio-stimamiglio-timmermann-

VIDAL, Vanderlei Vanderlino. **Cromatografia na perícia de incêndios: técnicas para detecção de agentes acelerantes**. 2007. Monografia (Especialização em Gestão dos Serviços dos Bombeiros) **Arquivo digital:** https://gnuteca.cbm.sc.gov.br/html/file.php?folder=material&file=ccem_2007_vanderlino.pdf

VIERA, Angel Freddy Godoy. **Modelagem de data warehouse como suporte na tomada de decisão da alta gerência na administração pública. Estudo de caso: Chancelaria paraguaia**. 2001. Tese (Doutorado em Engenharia de Produção)- Programa de Pós-graduação em Engenharia de Produção, UFSC, Florianópolis, 2001. **Arquivo digital:** <http://repositorio.ufsc.br/xmlui/handle/123456789/80236>

APÊNDICE

DIMENSÕES EM CSV

dim_zona_origem.csv

"id_zona_origem";"zona_origem";

-1;"Não registrado";

1;"Varanda";

2;"Banheiro";

3;"Sala";

4;"Sala de aula";

5;"Sala de estar";

6;"Depósito";

7;"Casa de baterias";

8;"Escadaria";

9;"Oficina";

10;"Barrilete";

11;"Estacionamento";

12;"Subsolo";

13;"Câmara fria";

14;"Capela";

15;"Área de serviço";

16;"Mezanino";

17;"Escritório";

18;"Garagem";

19;"Hall de circulação";

20;"Lavanderia";

21;"Central de gás";

22;"Alojamento";

23;"Ático";

24;"Elevador";

25;"Área de lazer";

70

26;"Palco";

27;"Centro cirúrgico";

28;"Laboratório";

29;"Antecâmara";

30;"Coreto";

31;"Copa";

32;"Lavabo";

33;"Cozinha";

34;"Porão";

35;"Dormitório";

36;"Hall de entrada";

37;"Casa de máquinas";

39;"Abrigo de gás";

40;"Edícula";

41;"Cobertura";

42;"Suíte";

43;"Área de atendimento";

44;"Púlpito";

45;"Recepção";

46;"Administração";

47;"Contâiner";

48;"Sala de raio X";

49;"Churrasqueira";

50;"Outros";

51;"Indeterminado";

dim_sub_causa.csv

id_sub_causa;nm_sub_causa;

-1;"Não informado";

1;"Descarga atmosférica";

2;"Choques mecânicos";

3;"Falha de equipamentos";

- 4;"Mau uso de equipamento";
- 5;"Fenômeno químico";
- 6;"Fenômeno biológico";
- 7;"Chama direta ou outra fonte de calor";
- 8;"Outros";
- 9;"Raios solares, raios solares ampliados acidentalmente";

dim_evento_causal.csv

id_evento_causal;nm_evento_causal;

- 1;"Não registrado";
- 1;"Reações biológicas";
- 2;"Processo fermentativos";
- 3;"Reação química exotérmica";
- 4;"Superfície aquecida";
- 5;"Fagulha";
- 6;"Colisões acidentais";
- 7;"Combustão espontânea";
- 8;"Raio solar direto";
- 9;"Raio solar ampliado";
- 10;"Brasa";
- 11;"Fogos de artifício";
- 12;"Fenômeno termoelétrico";
- 13;"Operação de equipamentos";
- 14;"Avarias ou falhas mecânicas";
- 15;"Chama";
- 16;"Estrutura colapsada";
- 17;"Outros";
- 18;"Não aplicável";
- 19;"Choques mecânicos (terremotos)";
- 20;"Meteorito";
- 21;"Atrito";

ANEXO A**Lista de tabelas do E193 (em ordem alfabética):**

acionamento_viatura;
arquivo;
atendimento_chamadas;
atendimento_chamadas_trote;
aux_cobom;
b_admin_auditoria;
b_admin_avisos;
b_admin_configuracao;
b_admin_responsavel;
b_admin_secoes;
bairros;
b_categoria_produto;
b_dados_pedido;
b_fornecedores;
b_historico_pedido;
biometria_efetivo;
biometria_pecas;
biometria_tipo;
biometria_tipo_posto;
b_lista_pedido;
bombcast;
bombcast_log;
b_pedido;
b_produto;
b_produto_estoque;
b_produto_validade;
b_produto_valor;
b_samu_base;
b_samu_viaturas;
caes;

ch_destino_vitima;
ch_posicao_veiculo_vitima;
ch_servico_salvamento;
ch_situacao_veiculo_vitima;
ch_sn_constraint_table;
ch_status_atividade_constraint_table;
ch_status_bbm_constraint_table;
ch_status_escalado_constraint_table;
ch_status_operacional_constraint_table;
ch_tipo_acionamento;
ch_tipo_evento;
ch_tp_ambiente;
ch_tp_animal;
ch_tp_animal_estado;
ch_tp_busca_evento;
ch_tp_busca_salvamento;
ch_tp Equipamentos_constraint_table;
ch_tp_escala_constraint_tables;
ch_tp_estado_fisico;
ch_tp_funcao_constraint_tables;
ch_tp_lesoes;
ch_tp_parte_corpo;
ch_tp_pessoa_constraint_table;
ch_tp_unidade_area;
ch_tp_unidade_peso;
ch_tp_veiculo;
cidades;
cidades_bbm;
cidades_cobom;
cidades_efetivo;
cidades_obms;
cobom;
cod_ocorrencia_praias;
cod_ocorrencias;

condicao_viatura;
contacts;
contacts_filha;
contacts_quantidade;
controle_acesso;
controle_certidao;
controle_certidao_praia;
controle_viaturas;
dados_praia;
dados_resgate;
dados_vitima;
denuncia_cap_portos;
destinos;
det_complemento_acao;
det_complemento_procedimento;
det_incendio;
det_ocorrencia;
det_pessoa_complemento;
det_pessoa_procedimento;
dia_semana;
ds_salvamento_praia;
ds_salvamento_resgate;
ds_salvamento_vitima;
dt_dados_praia;
dt_dados_resgate;
dt_dados_vitima;
ebombeiro_chat_auditoria;
efetivo;
efetivo_salvamento_praia;
emape;
emergencia_acao_preventiva;
emergencia_busca_animal;
emergencia_busca_caes;
emergencia_busca_evento;

emergencia_busca_salvamento;
emergencia_incendio;
emergencia_incendio_edificacao;
emergencia_incendio_tp_edificacao;
emergencia_incendio_vegetacao;
emergencia_produto_perigoso;
empenho_viaturas;
endereco;
escala_plantao;
escala_salva_vidas;
especialidade_efetivo;
especialidades;
evento_dificuldades;
evento_orgao;
ferramentas;
firecast_map_ask;
firecast_map_geo;
grau_ocorrencia_praia;
historico_arvore;
historico_arvore_acao_realizada;
historico_arvore_corte_encaminhado;
historico_arvore_hist_arv_ofer_per;
historico_arvore_hist_arv_tp_per;
historico_arvore_oferece_perigo;
historico_arvore_rec_compl_motoserra;
historico_arvore_recurso;
historico_arvore_tecnica_corte;
historico_arvore_tp_perigo;
historico_atendimento;
historico_complemento;
historico_emergencia_subgrupo;
historico_inseto;
historico_inseto_acao_realizada;
historico_inseto_captura;

historico_inseto_captura_recurso;
historico_inseto_captura_recurso_seguranca;
historico_inseto_destinacao;
historico_inseto_elimizacao;
historico_recursos_materiais;
historico_veiculo_envolvido;
info_call;
integracao;
logradouros;
map_hidrantes;
map_hidrantes_dedoduro;
map_hidrantes_report;
materiais;
militares_cargo;
militares_dataenvio;
militares_relatorio;
militares_situacao;
modulos;
monitoramento_agua;
monitoramento_ferramentas;
motivos_viaturas_bm;
multiplicador_sequence;
notas;
obm_alarme_remoto;
obms;
ocorrencia_evento_dificuldades;
ocorrencia_evento_orgao;
ocorrencias;
ocorrencias_map_abertura;
ocorrencias_map_deslocamento;
ocorrencias_samu;
perfilamentos;
perfis;
perfis_pericia;

pericia_1;
pericia_11_assinatura;
pericia_11_assinatura_del;
pericia_12_anexos;
pericia_12_anexos_del;
pericia_1_del;
pericia_1_vitimas;
pericia_1_vitimas_del;
pericia_2_entrevista;
pericia_2_entrevista_del;
pericia_3_4_5;
pericia_3_4_5_del;
pericia_5_sistemas_prev;
pericia_5_sistemas_prev_del;
pericia_6_7_8_9_10_11;
pericia_6_7_8_9_10_11_del;
pericia_audita_alteracao;
pericia_auditoria_atualizacao;
pericia_auditoria_deleta;
pericia_auditoria_impressao;
pericia_externa;
pes_atendidas;
pes_atendidas_lesoes;
pes_atendidas_tp_emergencia;
pes_atendidas_veiculo;
pes_atendidas_complemento;
pessoas;
pessoas_atendidas;
ponto_interesse;
posto_guarda_vidas;
posto_guarnicao;
praia_condicao;
praias;
prevencao_afogamento;

projeto_golfinho;
projeto_golfinho_crianças;
projeto_golfinho_instrutores;
proprietario;
recursos_materiais;
recursos_materiais_disponivel;
rel_ch_socorro;
rel_ch_socorro_foto;
rel_cmt_area;
rel_cmt_area_foto;
rel_cmt_area_tab;
replica_log;
rotinas;
salvamento_praias;
salva_vidas;
salva_vidas_recertificacao;
salvavidas_salvamento_praia;
servidor_alarme_remoto;
servidor_bbm;
servidor_cidade;
servidores;
servidor_network;
status_cadastro_usuario_sos_surdo;
status_sos_surdo;
tb_cidades;
tb_classificacao_chamada;
tb_estados;
tb_ip_integracao;
tp_acidente;
tp_agente_extintor;
tp_agente_igneo;
tp_avaliacao;
tp_bem_sinistrado;
tp_bem_sinistrado_especial;

tp_causa;
tp_destinacao;
tp_edificacao;
tp_emergencia;
tp_emergencia_pessoa;
tp_emergencia_subgrupo;
tp_evento;
tp_evento_causal;
tp_falha_preventivo;
tp_forma_aviso;
tp_gas_posicao;
tp_houve_vitima;
tp_intensidade_queima;
tp_local_origem;
tp_logradouros;
tp_material;
tp_material_estrutura;
tp_material_parede;
tp_material_piso;
tp_material_teto;
tp_motivo_nao_realizado;
tp_ocupacao;
tp_pericia;
tp_pessoal_empregado;
tp_sistema_preventivo;
tp_sistema_preventivo_existencia;
tp_sistema_preventivo_status;
tp_status;
tp_sub_causa;
tp_sub_destinacao;
tp_uso_preventivo;
tp_viatura_apoio;
tp_viaturas;
uf;

uploads_cobom;
usuario_sos_servidor;
usuario_sos_surdo;
veiculo_categoria;
veiculo_fabricante;
veiculo_modelo;
version;
viatura_apoio;
viaturas;
viaturas_guarnicao;
vitima_praia;
vtr_clear;
vw_bairro;
vw_cidade;
vw_encerramento_up;
vw_endereco;
vw_inser_ocorrencia;
vw_inser_ocorrencia_epv;
vw_pessoa;
vw_relatorio;

ANEXO B

Parâmetros para relatórios Módulo Perícia

1. Total de incêndios
 - 1.1 Por períodos (ano, mês, dia da semana, hora)
 - 1.2 por tipo (veicular, florestal, edificação, aeronave, explosão, etc)
 - 1.3 Por localização (RBM, BBM)
 - 1.4 Relação incêndio x habitante x município ?

2. Total de incêndios em edificações
 - 2.1 Por períodos (ano, mês, dia da semana, hora)
 - 2.2 Por localização (RBM, BBM)
 - 2.3 Relação incêndio x habitante x município
 - 2.4 Características construtivas (estrutura, teto, parede, piso)

3. Situação da edificação
 - 3.2 PPCI Indeferido/Deferido
 - 3.3 Habite-se Indeferido/Deferido
 - 3.4 Funcionamento Indeferido/Deferido
 - 3.5 PRE/PAI

4. Sistemas preventivos
 - 4.1 Existência
 - 4.2 Atende NSCI
 - 4.3 Utilização
 - 4.3.1 Quem utilizou (BM, ,PM GM, etc)
 - 4.4 Falha na operação
 - 4.5 Falha na manutenção

5. Produtos Perigosos
 - 5.1. Possui ou não
 - 5.2 Classe de risco
 - 5.3 GLP no interior ou exterior da edificação
 - 5.4 Validade da Mangueira e válvula

6. Extensão dos danos
 - 6.1 Área atingida e salva (total e percentual)

7. Prejuízos e salvados por período
 - 7.1 Prejuízos e salvados por ocupação (total e média por edificação)
 - 7.2 Prejuízos e salvados por situação da edificação (regularizada ou não)

8. Vítimas em incêndios (total, mortos e feridos, edificações regularizadas e não regularizadas)

9. Investigações em incêndios (Por RBM, BBM, ano e mês)
 - 9.1 Laudos
 - 9.2 Informes

9.3 Investigação não realizada

10. Incêndios em edificação por tipo de ocupação (totais e percentual)

11. Incêndios em edificações fiscalizadas pelo CBMSC - excluídas as unifamiliares – (totais e percentual) fato

12. Zonas de origem (total e percentual) dim_zona

13. Focos Iniciais (total e percentual)

14. Causas de incêndio (total e percentual)

15. Sub causas de incêndio (total e percentual)

16. Eventos causais (total e percentual)

17. Agentes causais (total e percentual)

OBS: Um relatório com investigações pendentes (semelhante ao que havia para ocorrências pendentes de fechamento) também seria bem importante, com vistas a cobrar dos responsáveis a conclusão no prazo e cumprimento da meta estabelecida pela DAT.