

EDIVALDO DA SILVA SOUZA

**O USO DE DATA MART PARA SISTEMATIZAÇÃO E APOIO À TOMADA DE
DECISÃO NA GESTÃO DE PESSOAS EM UMA INSTITUIÇÃO FEDERAL DE
ENSINO**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para obtenção do título de *Magister Scientiae*.

Orientador: Jugurta Lisboa Filho

Coorientador: Luiz Antônio Abrantes

**VIÇOSA - MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

Souza, Edivaldo da Silva, 1983-

S729u
2021

O uso de data mart para sistematização e apoio à tomada de
decisão na gestão de pessoas em uma instituição federal de
ensino / Edivaldo da Silva Souza. – Viçosa, MG, 2021.

1dissertação eletrônica (81 f.): il. (algumas color.).

Inclui apêndice.

Orientador: Jugurta Lisboa Filho.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Informática, 2021.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2021.176>

Modo de acesso: World Wide Web.

1. Data mart (Banco de dados). 2. Universidades e
faculdades públicas - Administração de pessoal - Indicadores.
3. Inteligência empresarial. I. Filho, Jugurta Lisboa, 1962-
II. Universidade Federal de Viçosa. Departamento de
Informática. Programa de Pós-Graduação em Ciência da
Computação. III. Título.

CDD 22. ed. 005.74

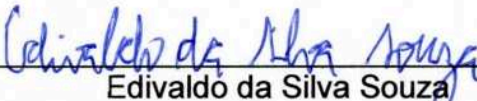
EDIVALDO DA SILVA SOUZA

**O USO DE DATA MART PARA SISTEMATIZAÇÃO E APOIO À TOMADA DE
DECISÃO NA GESTÃO DE PESSOAS EM UMA INSTITUIÇÃO FEDERAL DE
ENSINO**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para obtenção do título de *Magister Scientiae*.

APROVADA: 07 de outubro de 2021.

Assentimento:


Edivaldo da Silva Souza
Autor


Jurgurta Lisboa Filho
Orientador

Dedico esse trabalho a Deus, por ser o autor da vida! Condutor dos passos que me fizeram chegar até aqui, concretizando mais uma conquista em meio aos desafios enfrentados.

AGRADECIMENTOS

Sou grato,

A meus pais, pela dedicação, pela educação, pela instrução e pelos princípios que fizeram acreditar que a luta e a persistência são os caminhos para vencer os desafios.

À minha noiva Lanucy pela força, compreensão, paciência e companheirismo.

À minha querida Laurinha, pela alegria e descontração nos momentos de muita pressão durante essa caminhada.

Aos meus irmãos por estarem ao meu lado em cada desafio.

À Universidade Federal de Viçosa pela excelência no ensino e pela oportunidade de ingressar no mundo da pesquisa.

Agradeço especialmente ao Professor Jugurta Lisboa Filho, por me orientar, pelo seu incentivo e pela oportunidade de desenvolver essa pesquisa no meu ambiente trabalho.

Ao Professor Luiz Antônio Abrantes, meu chefe e coorientador, por compartilhar sua experiência, visão de gestão e pelo apoio nesse projeto.

À Pró-Reitoria de Gestão de Pessoas, que proporcionou a oportunidade de cursar o mestrado concomitante ao trabalho e aos colegas, em especial a Martha Pacheco, pelo apoio, amizade e ajuda nessa fase.

Aos colegas do programa de Mestrado em Ciência da Computação, pelas experiências e o crescimento em conjunto.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

RESUMO

SOUZA, Edivaldo da Silva, M.Sc., Universidade Federal de Viçosa, outubro de 2021. **O uso de Data Mart para sistematização e apoio à tomada de decisão na gestão de pessoas em uma Instituição Federal de Ensino.** Orientador: Jugurta Lisboa Filho. Coorientador: Luiz Antônio Abrantes.

O Business Intelligence (BI) é um conjunto de ferramentas utilizadas em organizações públicas e privadas, que possibilita resultados com dados precisos, utilizando-se de um banco de dados multidimensional, orientado por assunto, não volátil, histórico, decisório e variável em relação ao tempo. Ao aplicar o uso do Data Mart para uma Instituição Federal de Ensino no setor de gestão de pessoas, essa pesquisa trabalhou especificamente com dados de servidores aposentados e ativos da Universidade Federal de Viçosa (UFV) a fim de desenvolver indicadores para a tomada de decisão. Questões como a falta de sinergia entre as bases de dados existentes, qualidade dos dados fornecidos e impossibilidade de emissão de relatórios gerenciais em tempo hábil, foram tratadas nesse estudo. As ferramentas BI possibilitaram o preenchimento dessas lacunas, empregando as etapas: processo ETL (Extract, Transform, Load) com extração, tratamento e carga de dados no Data Warehouse; visualização de dados analíticos em relatórios, gráficos, tabelas e mapas. Concluiu-se que a aplicação e implantação de Business Intelligence por meio do Data Mart, com ferramentas *open source*, pode gerar dados precisos e fornecer indicadores de desempenho, auxiliando nos processos decisórios das Instituições.

Palavras-chave: Indicadores. Indicadores nas Instituições Federais de Ensino. Indicadores na Gestão de Pessoas nas Instituições Federais. Gestão de Pessoas. Business Intelligence. BI.

ABSTRACT

SOUZA, Edivaldo da Silva, M.Sc., Universidade Federal de Viçosa, October, 2021. **The use of Data Mart to systematize and support decision making in people management in a Federal Educational Institution.** Advisor: Jugurta Lisboa Filho. Co-advisor: Luiz Antônio Abrantes.

Business Intelligence (BI) is a set of tools used in public and private organizations, which enables results with accurate data, using a multidimensional, subject-oriented, non-volatile, historical, decision-making and time-varying database. By applying the use of Data Mart to a Federal Educational Institution in the people management sector, this research worked specifically with data from retired and active employees of the Federal University of Viçosa (UFV) in order to develop indicators for decision-making. Issues such as lack of synergy between existing databases, quality of data provided and impossibility of issuing management reports in a timely manner were addressed in this study. BI tools make it possible to fill these gaps, employing the following steps: ETL (Extract, Transform, Load) process with data extraction, treatment and loading into the Data Warehouse; visualization of these analytical data in reports, graphs, tables and maps. It was concluded that the application and implementation of Business Intelligence through Data Mart, with *open source* tools, can generate accurate data and provide performance indicators, helping in the decision-making processes of the Institutions.

Keywords: Indicators. Indicators in Federal Education Institutions. Indicators in People Management in Federal Institutions. People Management. Business Intelligence. BI.

LISTA DE ILUSTRAÇÕES

Figura 1 - Arquitetura de BI	25
Figura 2 - Composição básica de uma tabela fato	28
Figura 3 - Modelo Estrela	28
Figura 4 – Modelo Floco de Neve	29
Figura 5 - Modelagem Multidimensional.....	40
Figura 6 - Geocodificação Nominatim com dados OpenStreetMap Lotações	42
Figura 7 - Verifica Lotações sem Coordenadas Geográficas	43
Figura 8 - Processo ETL tabela <i>Fato Faixa Etária Tempo</i>	44
Figura 9 - Processo ETL Lei nº 103 de 12 de novembro de 2019 aposentadoria	45
Figura 10 - Rotinas job carga tabela <i>Fato Faixa Etária Tempo</i>	47
Figura 11 - Processo ETL tabela <i>Fato Servidor Aposentadoria</i>	48
Figura 12 - Geocodificação Nominatim com dados OpenStreetMap Endereços.....	49
Figura 13 - Cálculo de distância com API do Openrouteservice	50
Figura 14 - Processo de rotinas do job.....	51
Figura 15 - Estrutura do cubo do OLAP	52
Figura 16 - Desenvolvimento da Dimensão Servidor	53
Figura 17 - Ferramenta CDE Dashboard Pentaho	54
Figura 18 - Quantitativos de Servidores Ativos da última carga no DW	55
Figura 19 - Quantitativos de Servidores Ativos pelo Campi UFV	57
Figura 20 - Quantitativos de Servidores Ativos por Campi UFV no mapa.....	58
Figura 21 - Quantitativos de Servidores por Lotação	59
Figura 22 - Geolocalização das Lotações	60
Figura 23 - Quantitativos de servidores com Abono de Permanência e de servidores por gênero e status do cargo.....	61
Figura 24 - Dashboard com projeções de aposentadoria próximos 5 anos	62

Figura 25 - Mapa com projeções de aposentadoria nos próximos 5 anos por Campus	64
Figura 26 - Quantitativo de servidores aposentados, separados por gênero, tempo de aposentadoria e faixa etária	65
Figura 27 - Geolocalização Geográfica Servidores Aposentados	67
Figura 28 - Step <i>Table Input</i> conectado tabela Servidor	76
Figura 29 - Step <i>Url_Maps</i> para criação da url de requisição ao OSM	77
Figura 30 - Step <i>HTTP Client</i> realizando a requisição ao OSM	78
Figura 31 - Step <i>Coordenadas</i> configurações das abas “File” e “Content” em resposta da requisição do OSM.....	79
Figura 32 - Step <i>Coordenadas</i> configuração da aba “Fields” para selecionar os campos desejados da resposta do OSM.....	79
Figura 33 - Configuração Step <i>Modified Java Script Value 2</i>	80
Figura 34 - Conversão do campo “coord” do tipo String para Geometry	80
Figura 35 - Realização da carga na tabela DIM_Endereco com campos geográficos	81

LISTA DE TABELAS

Tabela 1 - Diferença entre banco de dados operacionais e DW	30
Tabela 2 - Funcionalidades do OLAP	32
Tabela 3 - Comparativo dos artigos relacionados com a pesquisa	35
Tabela 4 - Quantitativo Servidores por Lotação Cargo Status	62
Tabela 5 - Projeções de Aposentaria por cargo próximo 5 anos	65
Tabela 6 - Quantitativos servidores aposentados com sua distância da localidade atual até o Campus UFV	66
Tabela 7 - Quadro quantitativo dos processos de API para Geocodificação	67
Tabela 8 - Comparação entre as características do banco de dados operacional e do Data Mart desenvolvido	68

LISTA DE CÓDIGOS

Código 1 - Geolocalização das lotações pela API Nominatim com dados do OpenStreetMap	42
Código 2 - Geolocalização Nominatim com dados do OpenStreetMap.....	49
Código 3 - Geolocalização com API Google	50
Código 4 - Geolocalização Openrouteservice	50

LISTA DE ABREVIATURAS E SIGLAS

BI	<i>Business Intelligence</i>
BPEq	Banco de Professor Equivalente
CAF	Campus Florestal
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDE	<i>Community Dashboard Editor</i>
CE	<i>Community Edition</i>
CISTI	<i>Iberian Conference on Information Systems and Technologies</i>
CRP	Campus Rio Paranaíba
CTM	Centro Experimental do Triângulo Mineiro
DM	<i>Data Mart</i>
DOLAP	<i>Desktop On-line Analytical Processing</i>
DTI	Diretoria de Tecnologia da Informação
DW	<i>Data Warehouse</i>
EBTT	Ensino Básico, Técnico e Tecnológico
Enem	Exame Nacional do Ensino Médio
ETL	<i>Extract, Transform, Load</i>
HOLAP	<i>Hybrid On-line Analytical Processing</i>
IFEs	Instituições Federais de Ensino Superior
INSS	Instituto Nacional do Seguro Social
MDX	<i>Multidimensional Expressions</i>
MOLAP	<i>Multidimensional On-Line Analytical Processing</i>
ODS	<i>Operational Data Store</i>
OLAP	<i>On-Line Analytical Processing</i>
OLTP	<i>On-Line Transaction Processing</i>
OSM	OpenStreetMap
PDI	<i>Pentaho Data Integration</i>
PGP	Pró-Reitora de Gestão de Pessoas
QRSTA	Quadro de Referência dos Servidores Técnico-Administrativos
REUNI	Reestruturação e Expansão das Universidades Federais
RGPS	Regime Geral da Previdência Social
ROLAP	<i>Relational On-Line Analytical Processing</i>
SAD	Sistema de Apoio a Decisão

SIAPE	Sistema Integrado de Administração de Recursos Humanos
SISREC	Sistema de Dados de Recursos Humanos
SISU	Sistema de Seleção Unificada
SQL	<i>Structured Query Language</i>
UFV	Universidade Federal de Viçosa
WOLAP	<i>Web On-Line Analytical Processing</i>

SUMÁRIO

1	INTRODUÇÃO	14
	1.1 Objetivos.....	17
	1.2 Organização da Dissertação.....	18
2	REFERENCIAL TEÓRICO	19
	2.1 Governança Pública e Gestão de Pessoas nas IFEs	19
	2.2 Sistemas de Apoio à Decisão (SAD)	22
	2.3 Business Intelligence	23
	2.3.1 Componentes de Business Intelligence	24
	2.3.2 Dados Operacionais.....	25
	2.3.3 Processo ETL.....	25
	2.3.4 Data Warehouse - DW	26
	2.3.5 On-Line Analytical Processing - OLAP.....	31
3	EXEMPLOS DE DATA WAREHOUSE EM INSTITUIÇÕES DE ENSINO E DE SEGURANÇA.....	33
4	MÉTODO PARA DESENVOLVIMENTO DO DATA MART	36
5	RESULTADOS E DISCUSSÃO	38
	5.1 Desenvolvimento do Processo de BI.....	39
	5.1.1 Desenvolvimento das Tabelas Fatos	40
	5.1.2 Tabela <i>Fato Servidor Aposentadoria</i>	47
	5.1.3 Cubo OLAP	52
	5.1.4 Visualização dos dados	54
	5.2 Apresentação e Discussão dos Resultados	55
6	CONCLUSÕES.....	69
	REFERÊNCIAS.....	71
	APÊNDICE A – Obtenção de dados geográficos usando API OSM	76

1 INTRODUÇÃO

As Instituições Federais de Ensino (IFEs) passaram por grandes mudanças nas últimas décadas. O marco regulatório se deu com a publicação do Decreto nº 6.096/2007 (BRASIL, 2007), que criou o Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais – REUNI, objetivando ampliar o acesso e a permanência de discentes na educação superior. Dentre as metas do programa, ressalta-se o aumento de vagas nos cursos de graduação, a ampliação da oferta de cursos noturnos, a promoção de inovações pedagógicas e o combate à evasão, a fim de diminuir as desigualdades sociais no país.

O referido Decreto alterou parte do cenário existente, ao vincular o processo de pactuação de recursos humanos, financeiros e de infraestrutura ao cumprimento dos planos estabelecidos, condicionados à capacidade orçamentária e operacional do Ministério da Educação.

Além disso, as alterações no cenário econômico ocorridas ao longo da década dificultaram o processo de gestão, considerando, no momento em que se apontava para expansão, restringia-se os recursos financeiros e humanos. A criação do Quadro de Referência dos Servidores Técnico-Administrativos – QRSTA, por meio do Decreto nº 7.232/2010 (BRASIL, 2010), ao mesmo tempo que possibilitou a reposição automática de vagas por meio de Concurso Público, impediu a reposição do conjunto de cargos de nível operacional, classificados nos grupos “A” e “B” do plano de carreira, e para aqueles de outros níveis que estariam extintos ou em extinção.

Para os servidores pertencentes à carreira de Magistério Superior, o Decreto nº 7.485/2011 (BRASIL, 2011) dispôs sobre a constituição de Banco de Professor Equivalente – BPEq, concedendo autonomia às IFEs para gerenciar o seu quadro de pessoal docente, sem considerar, no entanto, todos os cargos vagos na apuração do quantitativo de provimentos autorizados. O mesmo procedimento foi adotado para a apuração dos cargos de técnico administrativo em educação, sendo apurados, em 30 de junho de 2010, a partir da extração do Sistema Integrado de Administração de Recursos Humanos – SIAPE, somente o quantitativo dos cargos ocupados naquela data.

Além disso, diversas mudanças foram concretizadas, como, por exemplo, a adoção do Sistema de Seleção Unificada – SISU, com o oferecimento de vagas para candidatos participantes do Exame Nacional do Ensino Médio – Enem; das alterações na forma dos concursos públicos que prolongaram o processo seletivo; da extinção de funções gratificadas que eliminaram um número expressivo de postos de comando e de coordenação no âmbito das Instituições; das restrições orçamentárias impostas impedindo a reposição de cargos; da expansão e manutenção de infraestrutura, além da gestão do processo de terceirização na Instituição.

Acrescenta-se a esse cenário, a reforma administrativa, que teve início com a reforma da previdência, resultou em um crescimento do número de aposentadorias, implicando na redução da força de trabalho. Contabiliza-se ainda um número expressivo de servidores que cumpriram os requisitos para aposentar-se, com possibilidade de desligamento imediato da Instituição a qualquer momento.

Estes fatores acabaram imputando restrições na gestão financeira, operacional e principalmente na gestão de pessoas, considerando a manutenção e até mesmo o crescimento das demandas existentes e a dificuldade de reposição imediata de cargos que ainda têm seu provimento garantido em lei.

Diante deste cenário, para melhor eficiência do processo de controle e de gestão, torna-se necessário a implementação e uso de tecnologias de informação que possam, através da consolidação de base de dados primárias, construir e gerar indicadores na área de gestão e de desenvolvimento de pessoal. A qualidade e a precisão da informação relacionada ao perfil socioeconômico dos servidores, a composição e evolução do quadro de pessoal e a localização espacial da estrutura física e operacional dos diversos setores existentes, traduzidos em indicadores integrados, robustos e consolidados, torna-se uma importante fonte para os processos de planejamento e tomada de decisões em qualquer Instituição.

Especificamente na Pró-Reitoria de Gestão de Pessoas (PGP) da Universidade Federal de Viçosa (UFV), objeto desse estudo, as informações utilizadas nos processos decisórios são obtidas de diversas fontes de consulta, como, por exemplo, o Sistema Integrado de Administração de Pessoal - SIAPE, do Governo Federal, e o Sistema de Dados de Recursos Humanos - SISREC, sistema interno desenvolvido na Instituição. Ambos são alimentados manualmente, sem a padronização e integração de indicadores e, portanto, não explorados de forma analítica. Esta deficiência torna-

se um empecilho para a realização do objetivo institucional de aprimoramento da política de gestão integrada e desenvolvimento de pessoas previsto no Plano de Desenvolvimento Institucional, com metas estabelecidas para a realização de Censo Institucional para promoção de políticas voltadas para o controle, desenvolvimento, qualidade de vida e segurança do trabalho, bem como para implementar instrumento de dimensionamento e controle da mão de obra efetiva e terceirizada.

Dessa forma, um dos grandes problemas encontrados é a falta de sinergia entre as bases de dados existentes, ocasionando assimetria de informações impossibilitando a emissão de relatórios gerenciais consolidados em tempo hábil, acrescidos de temporalidade na ocorrência de diversos eventos relacionados à gestão de pessoas. Essa assimetria ocorre também com servidores aposentados, principalmente em questões que exigem o contato com os servidores de forma imediata para a resolução de problemas operacionais ou mesmo em outros contatos para o acompanhamento e a elaboração de políticas que resgatem as experiências desses ex servidores em programas de capacitação de novos servidores e até mesmo para reforçar o sentimento de pertencimento com a Instituição.

A importância deste estudo se expressa na visão de Chiavenato (2010), ao apontar que as organizações precisam utilizar indicadores de forma sistemática para compor um conjunto de dados que abordem os aspectos importantes da organização. Essa mesma visão é compartilhada por Takashina e Flores (1996), ao traduzir os indicadores como uma ferramenta essencial para o auxílio no planejamento e no controle de processos. Nesta ótica, os indicadores de desempenho e o conceito de Business Intelligence (BI), com suas ferramentas inteligentes de negócios, podem ser considerados uma estratégia positiva dentro de uma organização, que geram informações precisas e sólidas, sendo, assim, essenciais para um processo decisório mais eficiente.

Conforme Rob e Coronel (2010, pág. 536), BI é “um conceito amplo, coeso e integrado de ferramentas e processos para captar, coletar, integrar, armazenar e analisar dados para transformá-los em informações que deem suporte à tomada de decisões de negócio”. Segundo Coronel e Morris (2016), a arquitetura do BI é dividida em Dados operacionais; Processo ETL (Extract, Transform, Load); Data Warehouse (DW); e On-Line Analytical Processing (OLAP). Os Dados operacionais estão relacionados as fontes externas e internas. O processo ETL refere-se à extração,

transformação dos dados e carga no DW. O DW é a união de vários Data Marts, que, segundo Primak (2008), armazena dados específicos de um setor da organização, podendo dar suporte à tomada de decisão, além de apresentar baixo custo e menor tempo para implementação. No OLAP é feita a geração de dados analíticos para os usuários finais.

A utilização de ferramentas *open source* do processo de BI para a implementação do Data Mart, com a identificação e manipulação de diversas ferramentas para a integração de base dados de diferentes fontes, será fundamental para a construção de relatórios gerenciais em tempo real ou discriminados por ciclos temporais que possibilitem o fornecimento de informações contínuas de bases quantitativas e qualitativas integradas, resultando em agilidade e dinâmica necessárias para a gestão dos processos de controle, movimentação e desenvolvimento de pessoal na Instituição.

1.1 Objetivos

O objetivo geral desse estudo consiste em projetar, desenvolver e implantar um Data Mart para integração de base de dados, que possibilite imprimir agilidade e fornecer subsídios para o planejamento e tomada de decisão relacionadas à gestão e desenvolvimento de pessoas em uma Instituição Federal de Ensino.

Especificamente, pretende-se:

1. Identificar e caracterizar as bases de dados e nível de informações existentes relacionadas ao quadro de servidores da Instituição;
2. Criar instrumentos que permitam o armazenamento, a exportação e integração das bases de dados;
3. Gerar os modelos de visualização das bases de dados na forma de quadros, gráficos e espacialidade geográfica para a elaboração dos relatórios gerenciais;
4. Projetar e implantar um protótipo de Data Mart de forma a proporcionar dados analíticos com ferramentas de BI na Instituição.

1.2 Organização da Dissertação

Além da introdução, o trabalho segue com a abordagem de 6 capítulos. O Capítulo 2 está composto pela descrição do Referencial Teórico, contemplando temas relacionados a governança pública e gestão de pessoas na IFEs, Sistema de Apoio a Decisão e Business Intelligence.

O Capítulo 3 descreve alguns trabalhos correlatos que corroboram para a pesquisa.

No Capítulo 4 são descritos os métodos utilizados no desenvolvimento da pesquisa.

O Capítulo 5 detalha como foi desenvolvida cada fase do processo de BI nesta pesquisa e apresenta os resultados obtidos.

Finalmente, o Capítulo 6 apresenta as conclusões finais e indica possibilidades de trabalhos futuros.

2 REFERENCIAL TEÓRICO

Ao longo dos anos, cada vez mais, tem crescido a utilização de ferramentas de BI, considerando a necessidade das organizações de obtenção de informações integradas, precisas e imediatas, que podem ser obtidas a partir de BI. Acrescenta-se a isso o fato das ações dos gestores serem pautadas por instrumentos legais que se renovam constantemente. Esse dinamismo acarreta implicações para o gestor, exigindo atualização em tempo real para tomada de decisão, que pode ser aperfeiçoada com a utilização de ferramentas adequadas.

Nas seções seguintes, são abordados temas que auxiliam no entendimento da necessidade de desenvolvimento de ferramentas específicas para o setor público sendo eles a Governança Pública e Gestão de Pessoas nas IFEs, Sistema de Apoio a Decisão e Business Intelligence.

2.1 Governança Pública e Gestão de Pessoas nas IFEs

A Administração Pública Federal, nos termos do Decreto Lei nº 200, de 25 de fevereiro de 1967 (Brasil, 1967), compreende a administração direta e indireta. A primeira, constitui-se dos serviços integrados na estrutura administrativa da Presidência da República e dos Ministérios e a segunda compreende as categorias de entidades, dotadas de personalidade jurídica própria, como as Autarquias, Empresas Públicas, Sociedades de Economia Mista e as Fundações Públicas.

Os cinco princípios básicos da Administração Pública estão presentes no texto constitucional e determinam o padrão de obediência à legalidade, impessoalidade, moralidade, publicidade e eficiência. Dessa forma, cabe ao servidor público executar somente o que estiver previsto em lei, tratar todos os cidadãos de forma igual, atuar com princípios éticos, prestar contas à sociedade do uso de recursos empregados e resultados alcançados, além de oferecer o melhor serviço possível, preservando os recursos públicos.

Além desses princípios, destaca-se o arcabouço legal que os disciplinam, com destaque para a Lei Complementar nº 131/2009 (BRASIL, 2009), que determina a disponibilização, em tempo real, de informações pormenorizadas

sobre a execução orçamentária e financeira dos entes federados, e para a Lei 13.709/2018 (BRASIL, 2018), que regulamenta o uso, a proteção e a transferência de dados pessoais dos cidadãos brasileiros, instrumentos que garantem a defesa da administração públicas e dos direitos e garantias coletivas.

Essas evolução e modernização já acontecem na administração pública e, segundo SECCHI (2009), desde 1980 mudanças substanciais ocorrem nas políticas de gestão pública e no desenho de organizações programáticas. Essas reformas administrativas consolidam novos discursos e práticas derivadas do setor privado e os usam como benchmarks para organizações públicas em todas as esferas de governo (SECCHI, 2009).

Nestas mudanças, destaca-se a gestão de pessoas do setor público responsável pelo oferecimento de serviços de qualidade para o público em geral. Uma definição possível para gestão de pessoas no setor público, segundo Bergue (2014), está relacionada a articulação de esforços de gestão orientado para o suprimento, a manutenção e desenvolvimento de pessoas observando os valores culturais - de natureza social, política, jurídica e econômica – que moldam as condições do contexto em que se inserem as organizações públicas.

Chiavenato (2010) afirma que uma organização tem seus objetivos alcançados por meio da gestão de pessoas. Afirma ainda que o uso de sistemas de informação na gestão de pessoas requer a utilização de vários bancos de dados interligados, permitindo a obtenção e o armazenamento de dados referentes a diferentes estratos, a fim de auxiliar nas decisões das organizações.

Assim como em todos os aspectos da administração pública, o princípio da legalidade impõe que a gestão de pessoas atue com celeridade e embasada na legislação atualizada para a tomada de decisão. A publicação da Emenda Constitucional 103, de 12 de novembro de 2019 (Brasil, 2019), denota tal característica, ao incorporar uma série de mudanças à legislação de pessoal, especificamente no regramento de aposentadoria, requerendo dos órgãos de gestão de pessoas a adoção de novas ferramentas para aglutinar dados e interpretar os cenários de cada instituição, principalmente em momento de muitas vacâncias e pouco provimento em função da limitação legal.

Conforme texto constitucional, a aposentadoria no serviço público que alcança os servidores de cargo efetivo da União se dá no Regime Próprio de Previdência Social – RGPS, e pode ocorrer de forma voluntária com proventos integrais; voluntariamente com proventos proporcionais ao tempo de serviço; voluntariamente por idade, com proventos proporcionais ao tempo de serviço; e compulsoriamente aos 75 anos de idade.

A Emenda 103/2019 (Brasil, 2019), tratando da Reforma da Previdência, estabeleceu as mudanças relacionadas à idade mínima de 65 (sessenta e cinco) anos para homens e 62 (sessenta e dois) anos para as mulheres e 25 (vinte e cinco) anos de tempo de contribuição para ambos, 10 (dez) anos de serviço público e de 5 (cinco) anos no último cargo. Foi também concedido para os servidores públicos que trabalham em ambientes caracterizados como insalubres aposentadoria com a idade mínima de 60 anos para ambos os sexos, tendo cumprido 25 anos de contribuição. Para servidores que trabalham em ambientes de exposição ao agente insalubre são estabelecidos 10 anos de serviço público e 5 anos no último cargo.

Com essa Emenda, foram estabelecidas as regras de transição de pontuação 86/96 (oitenta/noventa), onde se estabelece aposentadoria voluntária para mulheres que cumprirem 56 (cinquenta e seis) anos de idade e 30 (trinta) anos de contribuição e os homens 61 (sessenta e um) anos de idade e 35 (trinta e cinco) anos de contribuição. Ambos devem ter 20 (vinte) anos de efetivo exercício no serviço público e 5 (cinco) anos no cargo efetivo. Torna-se necessário a totalização de noventa e seis pontos para os homens e oitenta e seis pontos para as mulheres, ambos somando a idade e o tempo de contribuição. Essa pontuação será aumentada anualmente em um ponto até que os homens atinjam 105 (cento e cinco) pontos e as mulheres 100 (cem) pontos. A partir de 2022, a idade mínima para aposentadoria passa a ser 57 (cinquenta e sete) anos para as mulheres e 62 (sessenta e dois) anos para homens.

Com todos esses regramentos, torna-se de grande importância estabelecer sistemas de simulação com projeções de aposentadoria e posterior planejamento para suprir essas vacâncias, seja na forma de servidores efetivos ou terceirizados, para que não haja a descontinuidade das atividades, além de projetar indicadores de desempenho para os processos de gestão.

Neste aspecto, Assis (2014), ressalta a importância da administração pública utilizar diversos tipos de indicadores para calcular e medir, considerando que a

trajetória dos dados, com diversos cenários estudados e a combinação de diferentes medidas, permitem aos administradores antecipar ou revisar, de modo a poder realizar as correções, e com isso ter um melhor uso do recurso público.

2.2 Sistemas de Apoio à Decisão (SAD)

Dentro de uma organização empresarial, o ato de eleger um caminho em detrimento de outro é permeado por grandes riscos. Decidir de forma a não otimizar recursos e atividades poderá comprometer todo o desempenho institucional. Dessa forma, é vital para a organização o estabelecimento de estratégias adequadas para se tomar decisões certas. Para Chiavenato (2003), a organização é um sistema de decisões em que cada pessoa participa consciente e racionalmente, escolhendo e decidindo entre alternativas mais ou menos racionais, que são apresentadas de acordo com sua personalidade, motivações e atitudes.

Robbins (2010, pág. 167) define a tomada de decisão como reação de um problema dentro da organização, afirmando que “um problema existe quando se verifica uma discrepância entre o estado atual das coisas e seu estado desejável [...]”. Na visão de Chiavenato (2003), o processo decisório acontece em sete etapas: percepção da situação; análise e definição do problema; definição dos objetivos; procura de alternativas de solução; avaliação e comparação dessas alternativas; escolha da alternativa mais adequada; e implementação da alternativa escolhida.

No processo decisório, quanto mais rápida e precisa for a decisão, menor seria a ocorrência de riscos. Sob essa perspectiva, os Sistemas de Apoio a Decisão (SAD) são fundamentais para que o processo decisório se torne eficaz. Os sistemas de apoio podem ser entendidos como a interação entre o homem e a “máquina”, em que os dados gerados pela “máquina” podem ser transformados em informações que fundamentem a tomada de decisão. Laudon e Laudon (1999) definem SAD como sistemas interativos sobre o controle do usuário que fornece dados e modelos para apoio aos gestores na tomada de decisão. A definição de Turban e Aronson (1998, pág. 77), corrobora e amplia esses conceitos:

“... um sistema flexível, interativo e adaptável especialmente desenvolvido para suportar a solução de um problema gerencial não estruturado para melhorar a tomada de decisão. Ele utiliza dados, provê uma interface amigável com o usuário e pode incorporar *insights* do próprio tomador de decisão.”

Caiçara Junior (2012) aponta que um dos recursos mais utilizados nesse sistema é a análise de simulações do tipo “e se”. Por exemplo: Baixando o valor do produto X em 5%, qual seria nossa margem de lucro? SAD “permite que o usuário faça perguntas novas e não antecipadas e intervenha de forma on-line para mudar a maneira como os dados são apresentados” (CAIÇARA JUNIOR, 2012, pág. 76).

Componentes de SAD são entradas de dados com fontes internas e externas, por uma base de dados, uma base de dados de modelos e a interface de usuário. Deve ser de fácil utilização para o usuário, visto que este vivencia o problema e o SAD tem o papel de auxiliá-lo. Destaca-se que o Data Warehouse, descrito na seção 2.3.4, é um dos bancos de dados considerado alicerces do SAD, por ser somente leitura, analítico, histórico, variável no tempo e com capacidade de armazenar grande volume de dados, que podem ser facilmente acessados e manipulados pelo usuário para apoio à decisão (TURBAN, 2004).

2.3 Business Intelligence

Business Intelligence (BI) é um conjunto de tecnologias que engloba ferramentas e processos amplos, coesos e integrados, utilizados para dar suporte a tomada de decisão. Atualmente, até uma organização com estrutura simples necessita de informações precisas que podem ser obtidas a partir de BI. Para Primak (2008, pág. 5) o BI é “um processo inteligente de coleta, organização, análise, compartilhamento e monitoração de dados contidos em Data Warehouse (DW) e/ou Data Mart (DM), gerando informações para suporte à tomada de decisões no ambiente de negócios”.

A economia de tempo (61%), versão única da verdade (59%), melhores estratégias e planos (57%), melhores decisões táticas (56%), processos mais eficientes (55%) e economia de custos (37%) foram apontados por Eckerson (2003) *apud* Turban *et al.* (2009), em pesquisa realizada entre 510 corporações, como benefícios em relação à aplicação de BI. Esses autores reiteram que o sucesso de

uma organização está relacionado à existência de informações confiáveis e um panorama de seu desempenho em tempo real. Outras vantagens do BI, como a geração de relatórios mais rápidos e precisos (81%); melhor tomada de decisões (78%); melhor serviço ao cliente (56%); e maior receita (49%) foram também apontados por Thompson (2004).

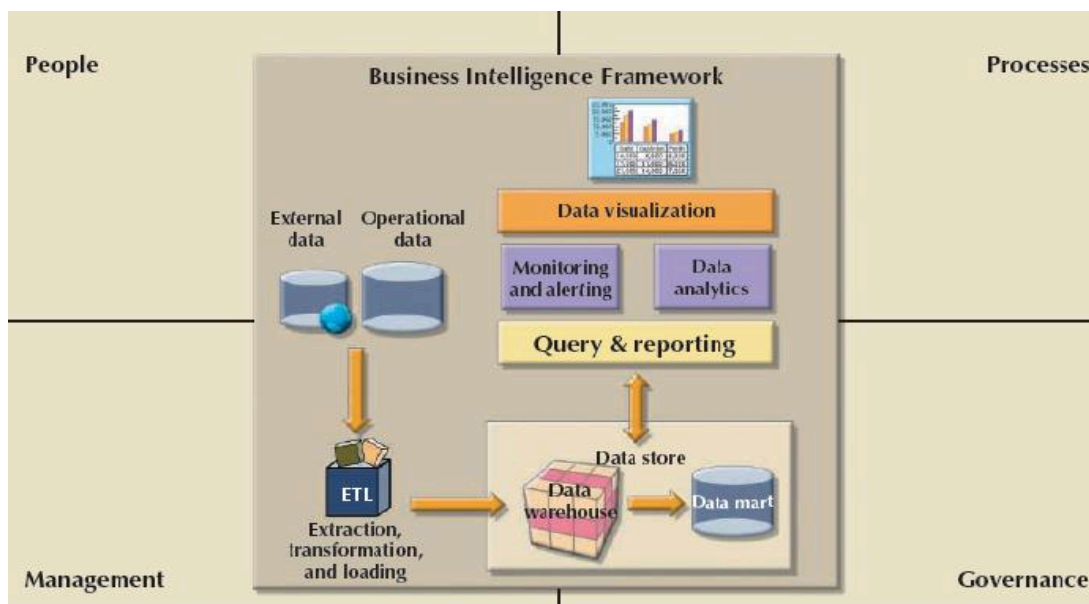
Rob e Coronel (2010) descrevem que para ter benefícios favoráveis, os dados são transformados em informações, as informações em conhecimento e o conhecimento em sabedoria. Eles sintetizam o BI em etapas relacionadas à coleta e armazenamento de dados operacionais; agregação de dados operacionais em dados de suporte a decisão; análise de dados de suporte a decisões para gerar informações; apresentação das informações ao usuário final para dar suporte a decisões de negócios; tomada de decisões de negócios, o que, por sua vez, gera mais dados que são coletados, armazenados etc. (reiniciando o processo); e monitoramento para avaliar os resultados de decisões de negócios (fornecendo mais dados a serem coletados, armazenados, etc.).

2.3.1 Componentes de Business Intelligence

A arquitetura de um ambiente de BI é composta por dados, pessoas, processos, tecnologia e gerenciamento de componentes. As ferramentas de BI focam na utilização de estratégias e táticas de informações e, para atingir o objetivo, apenas a tecnologia não basta. O BI integra as pessoas e processos por meio da utilização das tecnologias para agregar valor aos negócios. Dessa forma, o BI utiliza as melhores práticas de gerenciamento, como o gerenciamento mestre, englobando um modelo de governança empresarial, que analisa os dados como um bem corporativo. A governança realiza um monitoramento dos negócios dentro da organização através de indicadores de desempenho, utilizando medidas quantitativas e qualitativas para diagnosticar a situação da organização e apontar os rumos a serem seguidos (ROB E CORONEL, 2010).

A Figura 1 ilustra a arquitetura de BI e seus componentes, discutidos nas próximas subseções.

Figura 1 - Arquitetura de BI



Fonte: Coronel e Morris (2016, pág. 592)

2.3.2 Dados Operacionais

A eficiência de BI depende da qualidade dos dados coletados no nível operacional. Essa fase reúne diversas fontes de dados distintas, como On-Line Transaction Processing (OLTP), sistemas legados, tabelas com extensão xls, arquivos txt, XML, dentre outros, para o processo de extração, transformação e carregamento dos dados no DW.

2.3.3 Processo ETL

O processo de Extração, Transformação e Carga (ETL) é uma das primeiras fases do processo de BI, onde várias fontes de dados de uma Instituição passam por várias etapas, antes de serem armazenados em um DW (KIMBALL, 1998).

A extração dos dados pode ser obtida por meio de diversas fontes, como bancos de dados OLTP, arquivos TXT, CVS, Excel, XML e outros. A segunda etapa consiste no tratamento relacionado a limpeza, validação, otimização e unificação dos dados. Na terceira etapa é efetuada a carga dos dados por meio das tabelas de dimensões e da tabela fato. Finalizando o processo ETL, tem-se o banco de dados

multidimensional criado.

Kimball (1998) afirma que a extração de dados do processo ETL consome mais ou menos 70% do tempo de desenvolvimento de um DW, baseando em busca de informações mais importantes em várias fontes de dados. O processo ETL consome, também, cerca de 70% dos recursos necessários para implementação e manutenção de um DW (KIMBALL E CASERTA, 2004).

Nessa fase, conforme Barbieri (2011), os dados são coletados das diversas fontes e trabalhados numa camada intermediária de tratamento e armazenamento denominada *Staging*. Nessa camada, os dados são tratados e limpos, com a realização de combinações, classificações e acertos até o momento da carga no DW. Também pode-se criar um banco de dados ODS (*Operational Data Store*), que representa um depósito de dados até que a carga esteja pronta para ser realizada no DW.

2.3.4 Data Warehouse - DW

Data Warehouse (DW) é uma tecnologia muito utilizada por empresas privadas e entidades governamentais para o sucesso na gestão e análise de dados. Conhecido como armazém de dados históricos, o DW tem como função disponibilizar informações precisas e eficientes para auxiliar nas tomadas de decisão nas organizações (MACHADO, 2007).

Para Rob e Coronel (2010), o DW é um banco de dados apenas de leitura e otimizado para realizar análises e consultas de dados analíticos, que são extraídos de diversas fontes, transformados e integrados antes de serem armazenados no DW. Os usuários podem acessar esses dados por meio de ferramentas OLAP.

Inmon (1997, pág. 33) define o termo como “um conjunto de dados baseado em assuntos, integrado, não volátil e variável em relação ao tempo, de apoio às decisões gerenciais”. As principais características do DW são:

- Orientado por assunto: os dados no DW são armazenados por assunto mais importante de acordo com o interesse da organização;
- Integrado: o DW realiza a transformação dos dados de diversas fontes diferentes integrando-os em uma unicidade de informações, ou seja, em um

padrão único para ser armazenado;

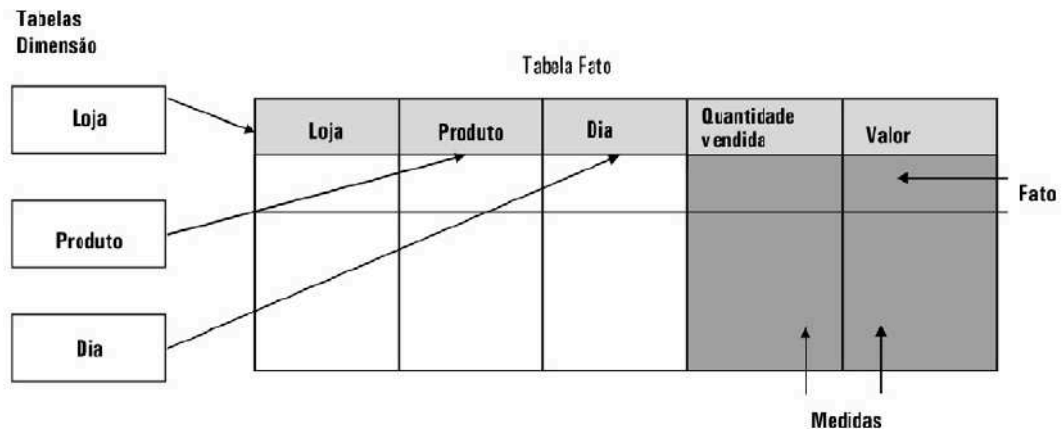
- Não volátil: os dados uma vez inseridos no DW não podem ser removidos ou editados, com isso estão sempre crescendo, mantendo um histórico;
- Variável no tempo: representa os resultados em determinado período de tempo. Os dados podem ser gerados de acordo com a necessidade da organização podendo ser por semana, mês, trimestres, anos e etc.

A criação de um DW integrado e centralizado, segundo Rob e Coronel (2010), demanda muito tempo, tem um elevado custo e esforço gerencial para a organização. Desta forma, as empresas focam em conjunto de dados mais gerenciados e orientados por assunto para atender as necessidades de pequenos grupos. Esses pequenos armazenamentos são denominados Data Mart, que garantem um retorno rápido às necessidades de uma área específica da organização.

A granularidade é um dos aspectos mais importantes em um DW e refere-se ao nível de sumarização dos elementos e de detalhes dos dados. Segundo Machado (2007), quanto mais detalhe existir, menor será seu nível de granularidade.

O DW é uma base de dados multidimensional dividido em tabelas fatos e dimensões. Segundo Machado (2007) e Barbieri (2011), a tabela fato é representada por medidas ou métricas, que é calculada com a interseção das tabelas de dimensões. Possui campos com chaves estrangeiras ligadas pelas chaves primárias das tabelas de dimensões por meio de relacionamentos. Barbieri (2011) define a tabela dimensão como entidade de negócios e constitui as estruturas de entrada, sua relação com a tabela fato é de 1:N. A Figura 2 ilustra um exemplo de um modelo dimensional clássico (BARBIERI, 2011, pág. 162). Identifica-se três tabelas dimensões (Loja, Produto e Dia), e uma tabela fato composta por 3 (três) chaves estrangeiras relacionadas com as tabelas de dimensões por meio de chaves primárias e dois campos de medidas (quantidade vendida e valor).

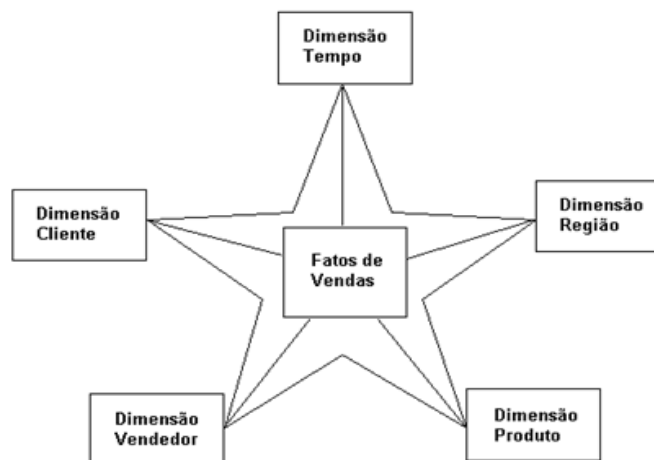
Figura 2 - Composição básica de uma tabela fato



Fonte: Barbieri (2011, pág. 162)

Segundo Elmasri e Navathe (2011), existem dois tipos de esquemas de dados multidimensionais no DW, que são o modelo estrela e modelo floco de neve. O modelo estrela (Figura 3) apresenta uma tabela fato centralizada com uma única tabela para cada dimensão (MACHADO, 2007, pág. 94).

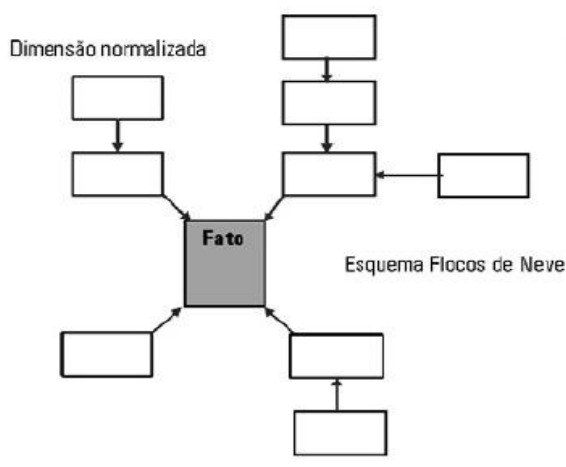
Figura 3 - Modelo Estrela



Fonte: Machado (2007, pág. 94)

Barbieri (2011) define o modelo floco de neve com uma tabela fato centralizada ligada às tabelas dimensões que apresentam um modelo de hierarquia normalizada. A Figura 4 exemplifica o modelo floco de neve, na qual pode-se observar também a decomposição de três tabelas de dimensões.

Figura 4 – Modelo Floco de Neve



Fonte: Barbieri (2011, pág. 170)

A arquitetura, para Kimball (1998), é um fator importante dentro de um DW, como ferramenta de comunicação, planejamento e flexibilidade. Segundo Machado (2007), essas arquiteturas podem ser globais, independentes ou integradas e podem ser implementadas seguindo os modelos *top-down* ou *bottom-up*, cuja junção dá origem a uma implementação intermediária.

A arquitetura global é a construção do DW atendendo à instituição de forma ampla, podendo ter o DW centralizado em um único departamento, por exemplo na diretoria de tecnologia da informação da instituição, ou pode ser descentralizado de forma que cada departamento tenha seu próprio Data Mart instalado. Conforme Machado (2007), a arquitetura global habilita os usuários finais a terem acesso aos dados de toda a corporação. Este ambiente tem como consequência uma demora no tempo de administração e desenvolvimento, elevando o custo de implementação do DW.

Na arquitetura independente, o Data Mart localiza-se no departamento, é administrada por um grupo de trabalho e não tem a intenção de lidar com as informações de outros departamentos. Por fim, a arquitetura integrada é a junção das duas arquiteturas anteriores, em que os Data Marts estão separados por grupo de trabalho e são integrados em visão geral da organização.

A implementação *top-down*, defendida por Inmon (1997), é o processo do DW que se inicia com a transformação, extração e integração dos dados a partir de sistemas legados ou dados externos, que em seguida serão armazenados no DW e a

partir daí são criados os Data Mart. A manutenção fácil; arquitetura integrada e flexível; visão de empreendimento; o controle e centralização de regras são algumas das vantagens desta implementação. Entretanto, essa implementação apresenta algumas desvantagens, como o desenvolvimento a longo prazo e custo elevado; alta taxa de riscos; e demora do projeto, que pode provocar alto grau de expectativa no usuário final.

A implementação *Bottom-Up*, apresentada por Kimball (1998), aborda um conceito principal diferente do defendido por Inmon. O objetivo é a construção de DW incremental a partir do desenvolvimento de diversos Data Marts, onde o processo inicia-se na extração, transformação e integração dos dados, a partir de sistemas legados ou dados externos que são transferidos para um DM, sendo que a união de todos os Data Marts constitui o DW. A implementação rápida, retorno rápido e herança incremental são consideradas as vantagens deste modelo, no entanto, deve-se destacar como desvantagem: perigo de criação de DM independentes (legamarts), possibilidade de perder a visão global do empreendimento, desafio de administrar e coordenar múltiplas equipe e a “Maldição do sucesso” (MACHADO, 2007).

Primak (2008) faz uma comparação entre o banco de dados operacionais e o DW, apontando as principais características, conforme Tabela 1.

Tabela 1 - Diferença entre banco de dados operacionais e DW

Características	Bancos de dados Operacionais	Data Warehouse
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 <i>ad-hoc</i>
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 <i>batch</i>)
Integridade	Transação	A cada atualização
Número de índices	Poucos/simples	Muitos/complexos
Intenção dos índices	Localizar um registro	Aperfeiçoar consultas

Fonte: Primak (2008)

2.3.5 On-Line Analytical Processing - OLAP

On-Line Analytical Processing (OLAP) é um conjunto de ferramentas que permite a análise dos dados de um DW de forma dimensional (MACHADO, 2007), onde os administradores das organizações têm acesso aos dados de forma rápida e eficaz, auxiliando nas tomadas de decisões.

Segundo Turban *et al.* (2009), os dados (e metadados) são mantidos no DW para serem acessados pelos usuários através de ferramentas de análise e exploração de dados, a exemplo das Ferramentas de Middleware, como aplicações personalizadas, navegadores Web, data mining, OLAP e ferramentas de relatórios e de visualização de dados.

Segundo Rob e Coronel (2010), as ferramentas OLAP apresentam quatro principais características, como a utilização de técnicas de análise de dados multidimensionais; o estabelecimento de suporte avançado a banco de dados; o fornecimento de interface fácil ao usuário final; e o suporte à arquitetura cliente/servidor. Os tipos de arquiteturas OLAP apresentam várias abordagens para os processos, de acordo com Turban *et al.* (2009):

- ROLAP (*Relational On-Line Analytical Processing*): utiliza bancos de dados relacionais e ferramentas familiares para armazenamento e análise dos dados;
- MOLAP (*Multidimensional On-Line Analytical Processing*): utiliza banco de dados multidimensionais e os dados são organizados em estrutura de cubos;
- HOLAP (*Hybrid On-Line Analytical Processing*): utiliza os conceitos do ROLAP e MOLAP unindo alto desempenho do MOLAP com a alta escalabilidade do ROLAP;
- DOLAP (*Desktop On-Line Analytical Processing*): devido ao fato de todo o processamento de dados do cubo multidimensional ser feito na própria máquina, desta forma o tráfego na rede diminui;
- WOLAP (*Web On-Line Analytical Processing*): os dados são acessíveis pela rede.

A Tabela 2 resume as principais funcionalidades e características oferecidas pelas ferramentas OLAP, segundo Machado (2007).

Tabela 2 - Funcionalidades do OLAP

Funcionalidade	Descrição
Drill Down	Ocorre quando o usuário aumenta o nível de detalhe da informação e com isso diminui o grau de granularidade.
Drill up	Ocorre quando o usuário diminui o nível de detalhe da informação e com isso aumenta o grau de granularidade.
Drill Across	Ocorre quando o usuário pula um nível intermediário dentro de uma mesma dimensão.
Drill Though	Ocorre quando o usuário passa de uma informação contida em dimensão para uma outra dimensão.
Slice	Operação que corta o cubo, mas mantém a mesma perspectiva da visualização dos dados.
Dice	É uma mudança da perspectiva da visão, é como se cubo fosse girado.

Fonte: Adaptado pelo autor (Machado, 2007).

3 EXEMPLOS DE DATA WAREHOUSE EM INSTITUIÇÕES DE ENSINO E DE SEGURANÇA

Diversas são as pesquisas e publicações que tratam do tema relacionado a construção de um DW frente às instituições, para gerenciamento e tomada de decisões mais assertivas em múltiplas áreas de atuação.

Foram realizadas pesquisas entre 2018 e 2020 nos sítios científicos: Scielo, Sociedade Brasileira de Computação e CAPES, utilizando as palavras chaves “Business Intelligence”, “Business Intelligence no Ensino”, “Data Mart nas Universidades”, “Business Intelligence nas IFEs”, “Business Intelligence nas Instituições”, “Data Warehouse”, “Data Warehouse no Ensino”, “Data Mart na Administração Pública” e “Data Warehouse na Segurança”, que resultou em 37 artigos nacionais e internacionais. Desses, foram selecionados 5 (cinco) artigos que abordam a utilização do BI na administração pública, 4 (quatro) tratam do uso do BI em instituições de ensino e 1 (um) apresenta a mesma abordagem na área de segurança pública.

A administração da cidade do Rio de Janeiro, objetivando utilizar o DW como instrumento de interligação de bases diferentes para a geração de indicadores estratégicos criminais, adotou um sistema com ambiente de apoio à decisão baseado em DW para área de segurança pública, interligando vários órgãos, como corpos de bombeiros, polícia civil, polícia militar e outros. Segundo Santos e Oliveira (2018), foram utilizados dados da área de segurança pública do Estado nos anos de 2015 a 2017 e efetuado a junção dos indicadores do DW com Data Marts específicos, resultando em consultas de agregações pré-calculadas apresentando um tempo de execução bastante reduzido.

O conceito de BI em uma universidade chilena, com o desenvolvimento de um Data Mart para obter indicadores de produtividade acadêmica, a fim de dar suporte às decisões estratégicas foi abordado por Medina *et al.* (2018). A construção do DM seguiu as etapas de análise; projeto conceitual MM; processo ETL + V (Validação); e processamento analítico. Foram construídos dois Data Marts, o primeiro contendo indicadores de produtividade dos professores e o segundo contendo indicadores de produtividade científica. Para o processo de ETL foi utilizada a ferramenta Pentaho

Data Integration (PDI) e para a etapa OLAP utilizou-se o Qlikview. O resultado alcançado foi o desenvolvimento de uma plataforma de inteligência de negócios para apoio a tomada de decisões sob uma perspectiva de estratégia.

Objetivando propor uma metodologia para a padronização do registro dos motivos que levam os alunos a evadirem, Silva e Jardim (2019) viabilizaram, através da utilização da ferramenta Power BI, o acesso instantâneo de forma analítica aos motivos de evasão, disponibilidade dos dados pela web, análise temporal dos motivos das evasões, possibilidade de compartilhamento das informações para os diversos gestores da Instituição onde foi realizada a pesquisa.

Considerando a importância da tomada de decisão para a gestão empresarial, Castro *et al.* (2019) propuseram a implantação de um sistema de Business Intelligence (BI) aplicado a gestão acadêmica/administrativa em uma Instituição de ensino. Como apoio de ferramenta “acadêmica” o BI pode gerar dashboard em relatórios para mais eficiência em tomadas de decisões, considerando que os grandes volumes de dados puderam ser tratados de maneira ágil e consolidados.

Em artigo relacionado à geração de proposta de Business Intelligence para a área de ensino do Instituto Federal da Bahia, De Souza Barreto e Freitas (2020), apontam a importância dos indicadores para as instituições de ensino avaliarem o seu desempenho. Os autores procuraram identificar os principais indicadores de ensino existentes na instituição e suas abordagens e apresentaram uma proposta para a construção de BI com uma solução da Microsoft, Power BI, para gerar microdados extraídos das bases dos sistemas acadêmico e administrativo da Instituição, a fim de produzir indicadores estratégicos na tomada de decisão referente à vida acadêmica.

Num breve comparativo entre os cinco artigos selecionados e a abordagem desta pesquisa, a Tabela 3 destaca que esta utilizou software livre e sua aplicação na complexidade da gestão de pessoas de uma Universidade.

Tabela 3 - Comparativo dos artigos relacionados com a pesquisa

Informações	Artigo 1	Artigo 2	Artigo 3	Artigo 4	Artigo 5	Pesquisa
Público Alvo	Segurança Pública	Universidade Chilena Estadual	Instituto Federal de Ensino	Instituições de Ensino Superior	Instituto Federal de Ensino	Instituições Federais de Ensino
Estudo de caso	bombeiros, polícia civil, polícia militar, etc.	Universidad Arturo Prat	Instituto Federal de São Paulo	Instituição de Ensino Superior	Instituto Federal da Bahia	Universidade Federal de Viçosa
Objetivos	Combater a Criminalidade no estado do Rio de Janeiro	Apoiar na tomada de decisão a partir da perspectiva estratégica.	Gerar indicadores para acompanhar os motivos de evasão dos alunos.	Implementação do BI na gestão acadêmica/ administrativa em uma instituição de ensino superior	Analisar dados e propor uma proposta do conceito BI.	Gerar dados analíticos para auxiliar os gestores na tomada de decisão
Ferramentas	Artigo 1	Artigo 2	Artigo 3	Artigo 4	Artigo 5	Pesquisa
ETL	Kettle	Kettle	Power BI	Não citou o nome da ferramenta	Power BI	Kettle
DW	PostgreSQL	Não citou o nome da ferramenta	PostgreSQL Planilhas Eletrônicas	Não citou o nome da ferramenta	Não citou o nome da ferramenta	PostgreSQL
OLAP	Não citou o nome da ferramenta	Qlikview	Power BI	Não citou o nome da ferramenta	Power BI	Pentaho Community Dashboard Editor (CDE)

Fonte: Autor

4 MÉTODO PARA DESENVOLVIMENTO DO DATA MART

Para obter os resultados desta pesquisa, utilizou-se de dados secundários anonimizados relacionados a vida funcional e perfil dos servidores da Universidade Federal de Viçosa, fornecidos pela Pró-Reitoria de Gestão de Pessoas da Instituição. Esses dados foram processados pelas ferramentas do BI, conforme as propostas dos autores Kimball (1998), Inmon (1997), Machado (2007) e Turban *et al.* (2009). Os resultados finais foram obtidos por meio de um conjunto de rotinas e padrões de programação que utilizaram aplicativos de softwares, APIs e planilhas eletrônicas.

Este estudo se classifica como de abordagem qualitativa que, segundo Cresweel (2007, pág. 184), se baseia em “dados de texto e imagem, têm passos únicos na análise de dados e usam estratégias diversas de investigação”. O processo BI possibilita realizar análise de dados, com levantamento, tratamento, unificação de dados, resultando em dados quantitativos, que auxiliem os gestores na tomada de decisão. Dados como quantitativo de servidores, faixa etária, tempo de serviço, quantitativo dos servidores por lotação, quantitativo de servidores por cargo, extraídos do banco de dados da Instituição, consistiram como a principal fonte para utilização das ferramentas BI.

4.1 Modelagem de Dados do Processo BI

A primeira etapa da pesquisa foi identificar e estudar várias ferramentas *open source* do BI com simulação de base com testes. Em seguida, realizou-se o processo de modelagem UML, em que foi analisada a base de dados disponibilizada. Como a pesquisa foi aplicada e, para atendimento dos interesses da Instituição, foram realizadas reuniões com os gestores para a definição do nível de informação e dos indicadores a serem desenvolvidos. Após esse levantamento, obteve-se os dados necessários para o desenvolvimento da modelagem UML, utilizando os modelos estrela e o floco de neve do processo BI, conforme proposto por Machado (2007).

Para a geolocalização das lotações, campus e servidores aposentados foi utilizada a modelagem espacial. O próximo passo foi efetuar as conexões com a base de dados da Instituição e executar o processo ETL do BI.

4.2 Desenvolvimento do Data Mart

Os dados disponibilizados foram extraídos¹ e inseridos nas tabelas temporárias (STG) no DW. Antes da inserção, utilizou-se a opção “*Truncate*”, que deletou todos os dados existentes nessas tabelas e posteriormente inseriu novos dados. Os dados extraídos podem ser acessados e manipulados a qualquer momento, sem comprometer o desempenho do banco de dados da PGP.

Os dados armazenados na tabela STG passaram por tratamento e limpeza e foram inseridos nas tabelas de dimensões (DIM), com atualização dos dados antigos e inserção de novos. Após a realização do tratamento dos dados, a última etapa do processo ETL consistiu na carga dos dados no DW. Em seguida, o cubo de dados foi criado.

O acesso aos dados foi feito por meio de um navegador Web, possibilitando visualizar as informações produzidas pelas ferramentas do processo BI, tais como gráficos, tabelas, relatórios e mapas.

¹ A extração dos dados é efetuada em períodos de menor demanda de consultas e atualizações da base de dados.

5 RESULTADOS E DISCUSSÃO

A Pró-Reitoria de Gestão de Pessoas da UFV, para análise da situação funcional dos servidores, dispunha de duas bases de dados alimentadas manualmente. A primeira formada por uma base cadastral interna (SISREC) e a segunda formada pelo banco de dados do Governo Federal (SIAPE). O SISREC, desenvolvido internamente pela Diretoria de Tecnologia da Informação (DTI) da UFV, configura-se como um sistema de banco de dados transacional, cujos dados ficam disponíveis aos usuários em tempo integral. O SIAPE, apesar de conter maior volume de informação, retém o acesso dos usuários em um período do mês. Apesar de parte dos dados do SIAPE ter a possibilidade de ser migrada para o SISREC, a comunicação entre os dois sistemas é precária.

O acesso aos dados do SISREC, objetivando filtragem de dados, é feito manualmente, por consultas que exigem conhecimento da linguagem Structured Query Language (SQL). Para consultas mais complexas e com maiores temporalidades, em que se necessita da interação das bases de dados, na maioria absoluta das vezes a dinâmica adotada se faz com a interação dos usuários do sistema e a equipe da DTI através de solicitação de serviço, considerando a inexistência de um sistema único de geração de relatórios gerenciais.

Apesar da existência de sistemas de informação que auxiliam nos processos operacionais, a inexistência de ferramentas inteligentes e acessíveis, como o BI, para fornecer dados analíticos com suporte ao processo decisório, torna-se uma grande limitação para os processos de planejamento e de tomada de decisões. Dessa forma essas limitações foram sanadas considerando que a proposta de pesquisa consistiu, além do desenvolvimento do Data Mart, na implantação de ferramentas BI *open source*, resultando na implementação de indicadores gerenciais.

5.1 Desenvolvimento do Processo de BI

Na pesquisa foram utilizados dados do Setor de Gestão de Pessoas da UFV, extraídos para o período de novembro de 2020 a fevereiro de 2021, resultando em duas tabelas fato.

Para a modelagem UML foi utilizada a ferramenta CASE StarUML 5.0, disponível no sítio <http://www.dpi.ufv.br/projetos/geoprofile/download.html>, sistema *open source*, com disponibilidade para a modelagem de banco de dados espaciais. Uma vez instalada, a ferramenta foi customizada para modelagem de banco de dados geográficos utilizando o Perfil UML GeoProfile (LISBOA-FILHO *et al.*, 2010).

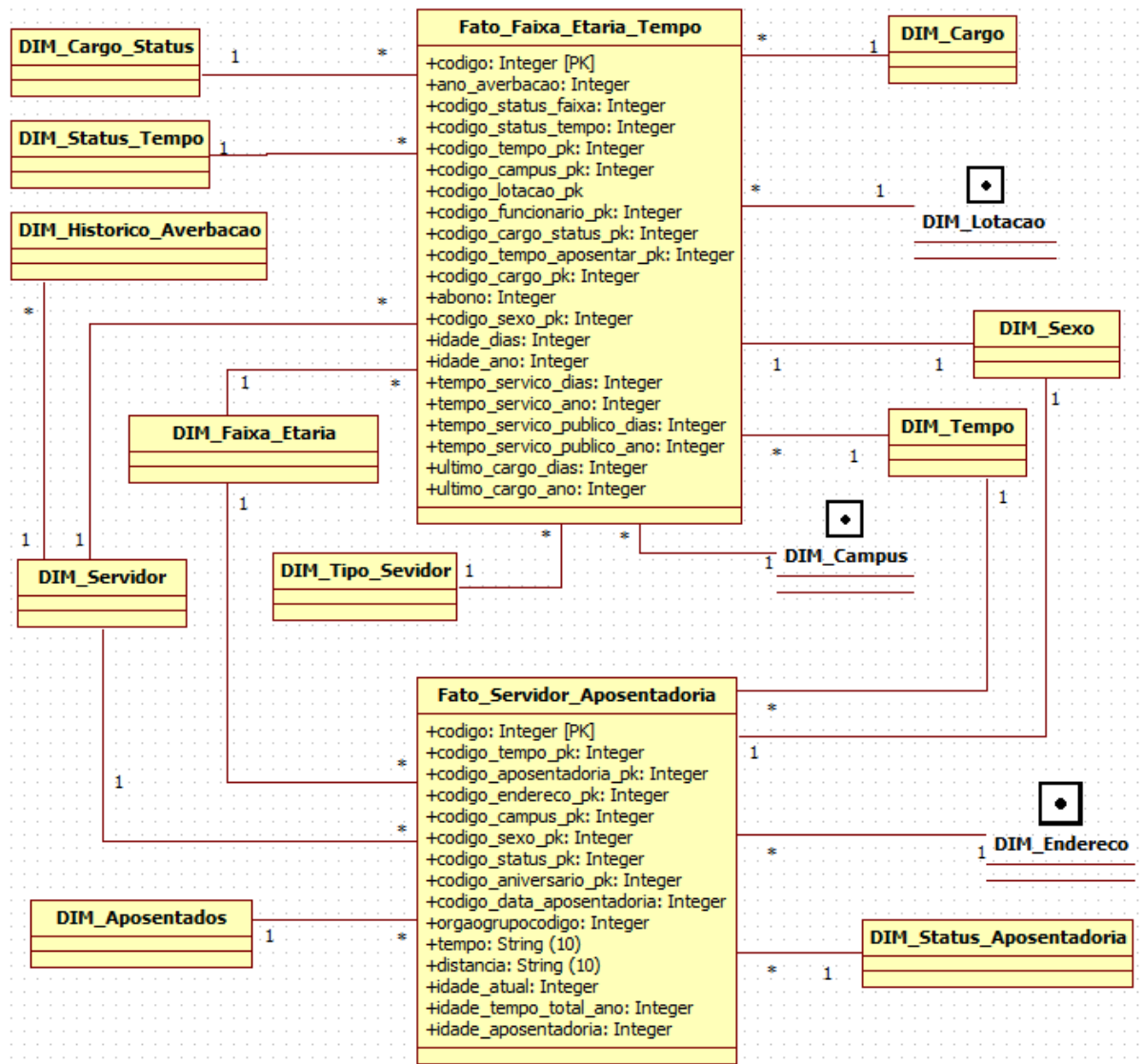
A ferramenta BI escolhida foi a Pentaho Suíte versão 8.3 CE, *open source*, que engloba: Pentaho Data Integration (PDI) para o processo ETL; Schema Workbench Mondrian para desenvolvimento do cubo OLAP; Pentaho Community Dashboard Editor (CDE) para visualização dos dados por meio de mapas, gráficos e tabelas; e Pentaho Report Designer para emissão de relatórios². O banco de dados multidimensional utilizado foi o PostgreSQL versão 12.3, *open source*, e com a extensão PostGIS para suporte a banco de dados geográficos. As ferramentas foram instaladas no sistema operacional Linux Ubuntu versão 18.04 e encontram-se na Pró-Reitora de Gestão de Pessoas da UFV.

A Figura 5 ilustra o diagrama de classes UML-GeoProfile, resultado da modelagem multidimensional utilizando a ferramenta StarUML e contém 14 tabelas³ de dimensões (prefixo DIM) e 2 tabelas fato. As tabelas dimensões estão formadas pelas variáveis DIM_Servidor; Aposentados; Faixa Etária; Status Tempo; Sexo; Tempo; Cargo; Cargo Status; Lotação (dimensão geográfica com representação espacial pontual); Endereço (dimensão geográfica com representação espacial pontual); Campus (dimensão geográfica com representação espacial pontual); Tipo Servidor; e Status da aposentadoria. Além disso, apresenta-se também a tabela Histórico Averbação ligada apenas à tabela DIM_Servidor, que mantém o histórico de averbação de tempo de serviço dos servidores. No centro da figura estão representadas as tabelas *Fato Servidor Aposentadoria* e *Fato Faixa Etária Tempo*.

² Esses programas estão disponíveis no sítio <https://sourceforge.net/projects/pentaho>.

³ Como as classes do diagrama UML serão implementadas como tabelas no DW, neste texto usaremos o termo “tabela” para as classes, a fim de manter a nomenclatura da área de DW.

Figura 5 - Modelagem Multidimensional



Fonte: Autor

5.1.1 Desenvolvimento das Tabelas Fatos

Para o desenvolvimento do projeto, foram efetuados levantamentos junto à Pró-Reitoria de Gestão de Pessoas de possíveis informações que seriam necessárias para, de forma integrada, formar indicadores que pudessem auxiliar nos processos de planejamento e decisório. Em abordagem geral, foi relatada a necessidade da existência de um banco de dados integrado que provesse em tempo real ou discriminados por ciclos temporais, informações relacionadas a situação funcional dos servidores e permitissem a elaboração de projeções em diversos cenários relacionados à reposição de pessoal, determinação de lotação ideal nos diversos

órgãos, conhecimento das possibilidades de desligamento, principalmente via aposentadoria nas regras atuais, dentre outras motivações.

Dessa forma projetou-se os indicadores relacionados a (1) servidores por faixa etária e tempo de serviço; (2) geolocalização dos servidores aposentados; (3) servidores por Lotação; e (4) projeção de aposentadorias dos servidores por situação cargo. Para atender as demandas dos pontos abordados, foram desenvolvidas as tabelas *Fato Faixa Etária Tempo* e *Fato Servidor Aposentadoria*.

5.1.2 Tabela *Fato Faixa Etária Tempo*

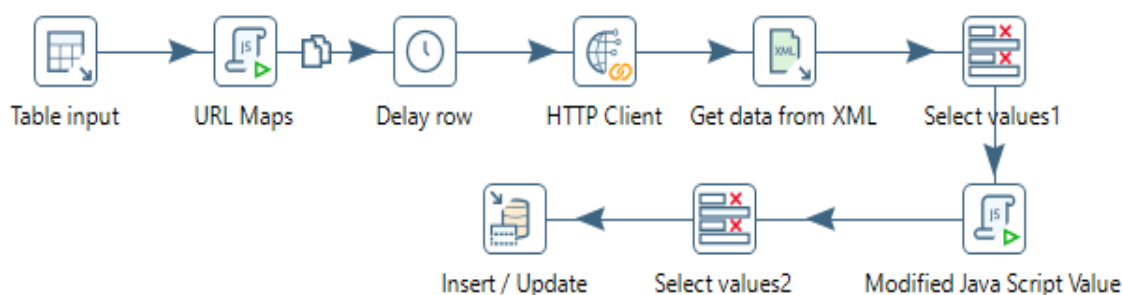
Na composição da tabela fato *Faixa Etária Tempo*, após a conclusão da modelagem dos dados, foi executado o processo ETL para fazer a conexão com banco de dados de Recursos Humanos (SISREC), concomitante à criação de planilhas eletrônicas para executar a carga das tabelas temporárias no DW.

Para efetuar a carga na tabela *Fato_Faixa_Etaria_Tempo*, foram criadas 11 tabelas temporárias dimensões (prefixo STG), definidas como *STG_Servidor*; *Cargo*; *Cargo Status*; *Lotacao*; *Tipo Servidor*; *Historico_Averbacao*; *Campus*; *Sexo*; *Faixa_Etaria*; *Status_Tempo*; e *Tempo*. As tabelas *Servidor* e *Historico_Averbacao* conectam-se à base de dados da PGP. As demais tabelas conectam-se a arquivos CSV. A carga das tabelas temporárias foi realizada em horário de menor demanda de acesso à base de dados. Esse processo é realizado pelo agendador de um *job* da ferramenta Pentaho, programado para executar essa rotina, no primeiro dia de cada mês. Com isso, a carga das tabelas de dimensões DIM no DW pode ser criada a qualquer momento, sem afetar o desempenho do banco de dados operacional.

O passo seguinte consistiu na criação de uma rotina no processo ETL para cada tabela STG, em que, ao conectar-se à base de dados da PGP ou aos arquivos CSV, os dados das tabelas STG são deletados e substituídos pelos novos dados extraídos. Em seguida foi realizado um processo ETL para a criação e carga de cada tabela dimensão (prefixo DIM): *DIM_Servidor*; *Cargo*; *Cargo_Status*; *Lotacao*; *Campus*; *Faixa_Etaria*; *Status_Tempo*; *Sexo*; *Tipo_Servidor*; e *Tempo*. A tabela *Historico_Averbacao* também foi criada e realizada a sua carga. Para esta tabela Fato foram utilizadas 2 tabelas com propriedade de localização espacial, conforme ilustrado anteriormente na Figura 5, considerando a *Lotacao* e *Campus*. Para obter os dados

geográficos dessas tabelas, utilizou-se a plataforma OpenStreetMap (OSM), disponível no sítio <https://www.openstreetmap.org>, onde previamente foi feito o mapeamento das lotações dos órgãos e setores dos *Campi* Florestal, Rio Paranaíba e Viçosa, com as siglas de cada setor para efetuar as consultas pela API Nominatim dos dados do OSM. A Figura 6 ilustra esse processo.

Figura 6 - Geocodificação Nominatim com dados OpenStreetMap Lotações



Fonte: Autor

O Step *Table input* (Figura 6) conecta-se à base de dados da tabela DIM_Lotacao, de onde os dados são extraídos. No Step *URL Maps*, é montado o caminho com o endereço da API do Nominatim, conforme especificado no Código 1, onde a variável “url_final” corresponde à composição de um link da API Nominatim com dados do OSM. Destaca-se que esse link é composto pelo endereço do sítio Nominatim, sigla lotação, cidade, uf, format – que define o tipo de resposta no formato XML, polygon – igual a “zero” indicando o não retorno ao polígono, addressdetails – igual a “zero” não retorna detalhes sobre o endereço, e limit – igual a “1”, traz apenas o resultado mais próximo da pesquisa.

Código 1 - Geolocalização das lotações pela API Nominatim com dados do OpenStreetMap

```

var url_final = 'https://nominatim.openstreetmap.org/search?q=' + encodeURI(sigla)
+ '%2C' + encodeURI(cidade) + '%2C' + uf + '&format=xml&polygon=0&addressdetails=0&limit=1';
  
```

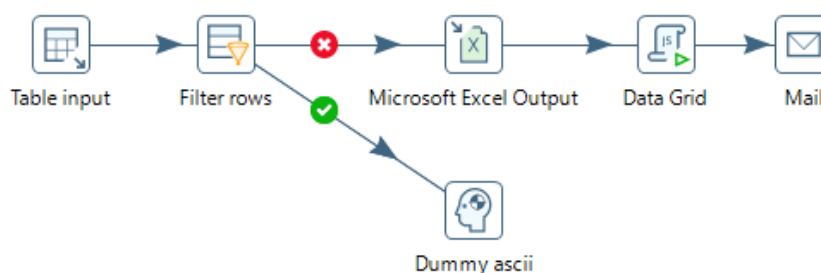
Fonte: Autor

O Step *Delay row* (Figura 6) é configurado de acordo com as restrições da API, que estabelece uma requisição por segundo. Em seguida o Step *HTTP Client* utiliza a configuração da URL final, solicita a requisição na API, codifica e configura a resposta de saída. O Step *GET data from XML* recebe a resposta da *HTTP Client* em formato XML, configura o caminho da requisição na Aba Content, opção Loop XPath, assim apresentado “/searchresults/place”. Na aba Fields o campo XPath é configurado com

os campos lat e lon selecionados da resposta da requisição. O tratamento dos dados é efetuado nos Steps *Select values1*, *Modified Java Script Value* e *Select values2*. No Step *Insert / Update* é realizada a carga das coordenadas geográficas na tabela DIM_Lotacao no DW. O mesmo procedimento foi utilizado para a tabela DIM_Campus para obter as coordenadas geográficas.

Para verificar se o processo incluiu informação de coordenadas geográficas de todas as lotações é realizado o processo ELT, conforme mostrado na Figura 7.

Figura 7 - Verifica Lotações sem Coordenadas Geográficas

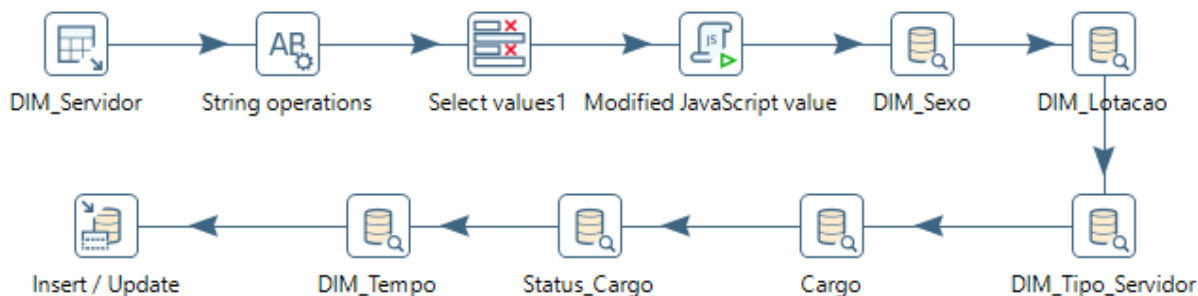


Fonte: Autor

O Step *Table input* (Figura 7) conecta à base de dados da tabela DIM_Lotacao. O Step *Filter rows* verifica se o campo Latitude está nulo. Caso esteja, os dados passam pelo Step *Microsoft Excel Output* que cria um arquivo Excel com a listagem das lotações sem coordenadas. O Step *Data Grid* configura os campos que são usados no Step *Mail*. Após o recebimento destes campos, o Step *Mail* fica pronto para o envio de e-mail para administrador do Data Mart, uma vez que foi configurado o anexo por meio da aba “Attached Files” apontando para o arquivo criado pelo Step *Microsoft Excel Output*. Caso não existam campos nulos, os dados passam para o Step *Dummy ascii* que finaliza o processo.

Em seguida, foi realizada a carga no DW alimentando a tabela fato com valores relacionados com as tabelas dimensões. A Figura 8 ilustra o processo realizado para a carga na tabela Fato.

Figura 8 - Processo ETL tabela *Fato Faixa Etária Tempo*



Fonte: Autor

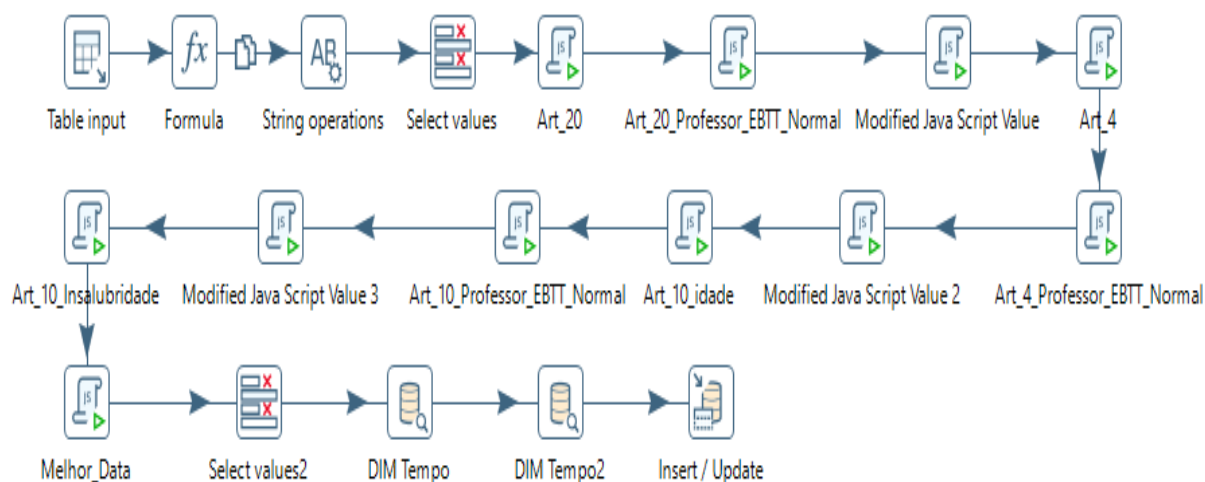
No Step *DIM_Servidor* (Figura 8) é feita a conexão com a tabela *DIM_Servidor* que busca os servidores ativos, classifica por classe e efetua o cálculo da faixa etária e do tempo de serviço, com averbação. Os Steps *String operations*, *Select values1* e *Modified JavaScript value* efetuam o tratamento dos dados. No Step *DIM_Sexo*, seleciona-se o código da tabela *Dim_Sexo*. Nos Steps *DIM_Lotacao*, *DIM_Tipo_Servidor*, *Cargo*, *Status_Cargo* e *DIM_Tempo* foram realizados os mesmos procedimentos, conectando-os nas tabelas de dimensões *DIM_Lotacao*, *DIM_Tipo_Servidor*, *DIM_Cargo*, *DIM_Status_Cargo* e *DIM_Tempo*, respectivamente. Finalmente, no Step *Insert Update*, realizou-se a carga na tabela *Fato*, resultando no DW.

Na mesma tabela *Fato*, foi desenvolvido um processo ETL para realizar cálculos de projeções de aposentadorias dos servidores com base na Emenda Constitucional 103 de 12 de novembro de 2019, nos próximos 5 anos. A carga foi programada para acontecer todo dia primeiro de cada mês, analisando os dados dos servidores ativos que podem solicitar o *Abono permanência*⁴ nos próximos anos.

⁴O abono de permanência é um incentivo financeiro dado ao servidor público estatutário contribuinte, que deseja continuar trabalhando, mesmo já tendo os requisitos para se aposentar. Benefício previsto no Art. 40, §19 da Constituição Federal (Brasil, 1998).

Na tabela *Fato_Faixa_Etária_Tempo* foi configurado um campo “*código_tempo_aposentar_pk*” que armazena a data da provável aposentadoria de cada servidor. Nesse processo foram abordados apenas os artigos que se referem aos Servidores Públicos Federais com os artigos 4º, 10, 20 da Emenda Constitucional 103 de 2019. A Figura 9 ilustra esse processo no ETL.

Figura 9 - Processo ETL Lei nº 103 de 12 de novembro de 2019 aposentadoria



Fonte: Autor

O Step *Table input* (Figura 9) conecta-se na tabela *DIM_Servidor*, em que por meio de consulta produzem informações dos servidores referentes a idade; tempo de averbação mais tempo de serviço; tempo de serviço professor EBBT; tempo de serviço último cargo; e tempo de serviço público. No Step *Formula* é calculada a diferença da data atual para a data de pedágio⁵, que ficou definida como “12/11/2019” conforme estabelecido na Emenda Constitucional 103 de 12 de novembro de 2019 (BRASIL, 2019). Em seguida, nos próximos Steps *String operations* e *Select values* efetua-se o tratamento dos dados. No Step *Art_20* são efetuados os cálculos das projeções de aposentadorias com relação ao Art. 20 da Lei nº 103 de 12 de novembro de 2019. Nesse *Step* foram realizadas várias simulações de ocorrência de aposentadoria, como Servidores com tempo normal - por gênero, idade, tempo de serviço, pedágio, tempo último cargo; Professores EBBT - com mesmos cálculos do anterior, mas com 5 anos a menos em contribuição e idade para ambos os sexos. O Step *Art_20_Professor_EBTT_Normal* realiza o cálculo dos professores EBBT com o tempo

⁵O Pedágio é o período adicional de contribuição correspondente a 100% (cem por cento) do tempo que faltava para o servidor atingir o tempo mínimo de contribuição previdenciária para se aposentar; tendo como data limite 12/11/2019 de acordo com a Emenda Constitucional 103 de 12 de novembro de 2019. (BRASIL, 2019).

de Servidor normal e logo no próximo Step *Modified Java Script Value* é calculada qual é a melhor data para o servidor aposentar.

No Step *Art_4* aborda o artigo 4 com os mesmos níveis de ocorrência do Art. 20, exceto o cálculo por tempo de pedágio, no entanto requer a soma dos pontos por idade e o tempo de contribuição. No Step *Art_4_Professor_EBTT_Normal* é realizado o cálculo dos professores EBTT com o tempo de serviço normal. No Step *Modified Java Script Value 2* é efetuado a projeção da data provável de aposentadoria do servidor.

O Art. 10 aborda duas situações de aposentadoria, sendo uma por idade e outra por tempo de serviço insalubre. O Step *Art_10_idade* trata as mesmas ocorrências do Art. 20, mas sem o cálculo do tempo de pedágio e sem a soma dos pontos apresentado no Art. 4º. O Step *Art_10_Professor_EBTT_Normal* efetua o cálculo da provável data de aposentadoria, consolidado no Step *Modified Java Script Value 3*.

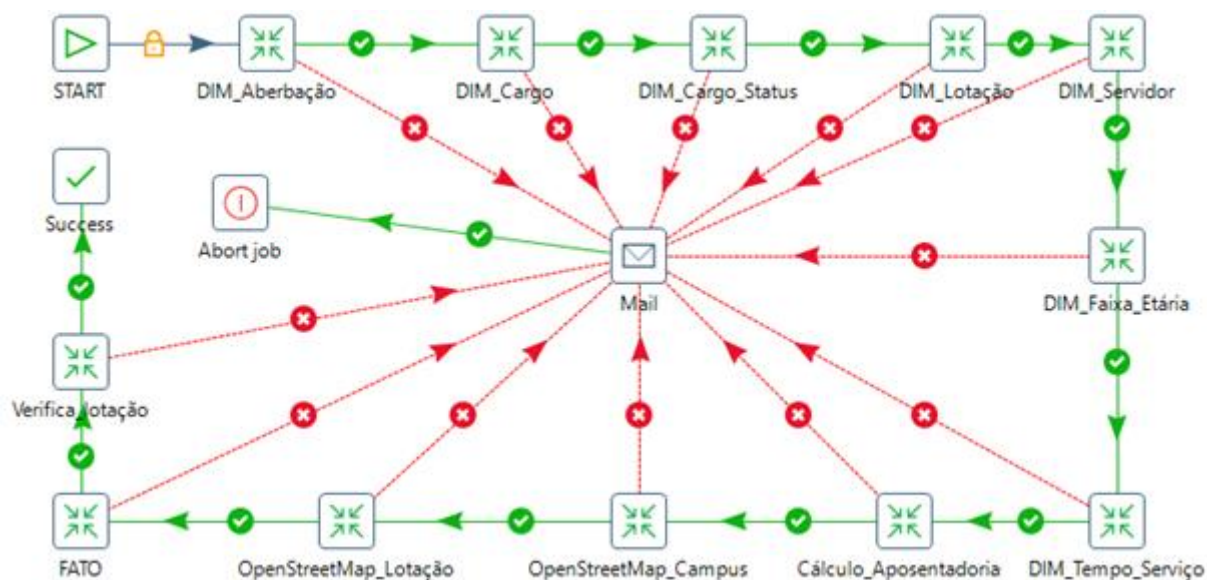
O Step *Art_10_Insalubridade* projeta a aposentadoria dos servidores por tempo de insalubridade, com as mesmas situações e ocorrências do Art. 20, mas sem o cálculo do tempo de pedágio e sem a soma dos pontos apresentados no Art. 4º. No Step *Melhor_Data* são analisadas as possíveis datas de aposentadoria do servidor pelos artigos mencionados e selecionado a menor data. Logo depois no Step *Select values2* é realizado o tratamento da data de aposentadoria. Em seguida, no Step *DIM Tempo* é feita a conexão com a tabela de *DIM_Tempo* para buscar o código da data atual. O mesmo procedimento é realizado no *DIM Tempo2* que é selecionado o código da data para aposentadoria. O processo finaliza com a atualização do campo “código_tempo_aposentar_pk” na tabela *Fato_Faixa_Etaria_Tempo*.

Com a tabela *Fato* criada foi construído o cubo OLAP, por meio da ferramenta Schema Workbench Mondrian. Em seguida, no ETL foi desenvolvido um *Job* para ordenar as rotinas das transformações do processo ETL para a realização das cargas nas tabelas *STG* e *DIM*. A Figura 10 ilustra o processo ETL para as tabelas *DIM* realizando a carga no DW na tabela *Fato_Faixa_Etaria_Tempo*.

O Step *START* (Figura 10) inicia o processo da carga dos dados para o DW, na seguinte ordem: *DIM_Averbação*; *DIM_Cargo*; *DIM_Cargo_Status*; *DIM_Lotação*; *DIM_Servidor*; *DIM_Faixa_Etária*; *DIM_Tempo_Serviço*; *Cálculo_Aposentadoria*; *OpenStreetMap_Campus*; *OpenStreetMap_Lotação*; *Fato*; *Verifica_Lotação*; e

Sucess. Caso ocorra algum erro nas etapas do processo, um e-mail é encaminhado para o administrador com log utilizando o *Step Mail* e a execução é abortada pelo *Step Abort Job*.

Figura 10 - Rotinas job carga tabela *Fato Faixa Etária Tempo*

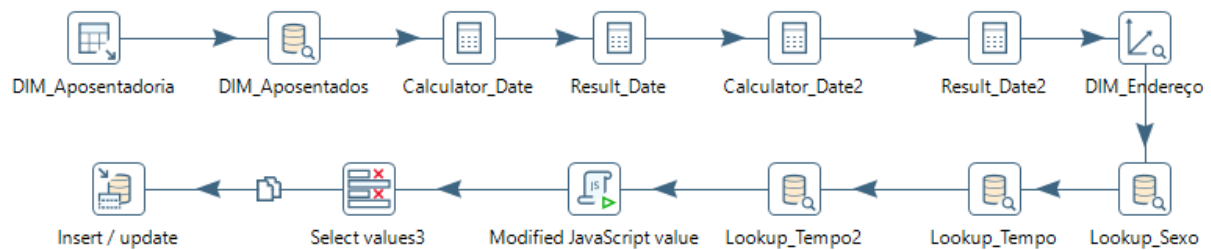


Fonte: Autor

5.1.2 Tabela *Fato Servidor Aposentadoria*

Para o desenvolvimento da tabela *Fato Servidor Aposentadoria* foi utilizado apenas dados dos servidores aposentados. O processo ETL, que gerou informações do banco de dados da Instituição, conseqüentemente criou e alimentou as tabelas temporárias (prefixo STG) relacionadas a STG_Servidor; Sexo; Endereço; Aposentados; Status_Faixa_Etária; e Status_Aposentadoria. A extração, executada em período de menor demanda, visa não comprometer a base do banco de dados da instituição. Os dados extraídos são armazenados em tabelas temporárias, que podem ser acessadas e manipuladas a qualquer momento, sem comprometer o desempenho do banco de dados OLTP. Em seguida foi executado o processo ETL na base de dados temporária e foram criadas as tabelas de dimensões (DIM), correspondentes ao diagrama da Figura 5. Em se tratando de dados de servidores aposentados, essas cargas são realizadas uma vez ao ano. Desta forma finalizou-se o desenvolvimento da base de dados do modelo multidimensional no banco de dados PostgreSQL. A Figura 11 ilustra o processo para a carga na tabela fato.

Figura 11 - Processo ETL tabela *Fato Servidor Aposentadoria*

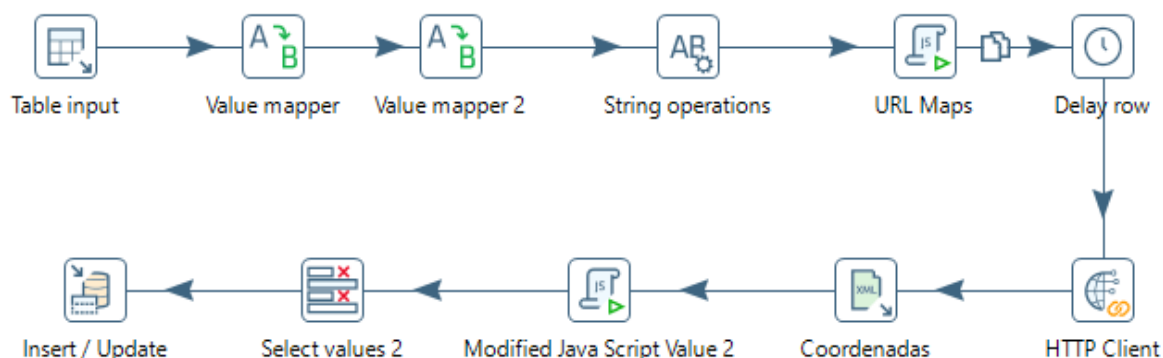


Fonte: Autor

No Step *DIM_Aposentadoria* (Figura 11) conecta-se a base de dados da tabela *DIM_Servidor*, onde são selecionados apenas os dados dos servidores aposentados. Em seguida no Step *DIM_Aposentados* é obtido a data de aposentadoria e o código do servidor da tabela *DIM_Aposentados*. Nos Steps *Calculator_Date*, *Result_Date*, *Calculator_Date2*, *Result_Date2* são calculados o tempo de aposentaria e faixa etária dos servidores aposentados. No Step *DIM_Endereço*, seleciona-se o código da tabela *DIM_Endereco*. O mesmo procedimento é realizado nos Steps *Lookup_Sexo*, *Lookup_Tempo* e *Lookup_Tempo2*. Já nos Steps *Modified JavaScript value* e *Select values3* são realizados os tratamentos da faixa etária e o tempo de aposentadoria. Finalmente no próximo passo tem-se a inserção dos dados na tabela *Fato Servidor Aposentadoria*.

Na etapa seguinte obtém-se a geolocalização dos servidores aposentados conjuntamente com suas coordenadas geográficas. Para isso utiliza-se a API Nominatim, com dados extraídos do OpenStreetMap e do Google Maps para complementação dos dados. Ambas aceitam a geocodificação, ou seja, permitem obter coordenadas geográficas (latitude e longitude) por meio do endereço e também possibilitam a geocodificação reversa, através das coordenadas (latitude e longitude) obter os mesmos endereços. A Figura 12 descreve o processo de geocodificação com a API do Nominatim e com dados do OpenStreetMap, podendo ser analisados com mais detalhes dos steps no Apêndice A. Esse mesmo processo ETL foi utilizado na Figura 6, porém com mais steps de tratamento de dados devido aos campos de endereços.

Figura 12 - Geocodificação Nominatim com dados OpenStreetMap Endereços



Fonte: Autor

O Step *Table input* (Figura 12) conecta-se a base de dados da tabela DIM_Endereco, de onde os dados são extraídos e em seguida passam por tratamento pelos passos dos Steps *Value mapper*, *Value mapper 2* e *String operations*. No Step *URL Maps* é montado o caminho com o endereço da API do Nominatim, conforme especificado no Código 2.

Código 2 - Geolocalização Nominatim com dados do OpenStreetMap

```
var url_final = 'https://nominatim.openstreetmap.org/search?street='+encodeURIComponent(logradouro)+'%2C'+encodeURIComponent(numero)+'&city='+encodeURIComponent(cidade)+'&state='+uf+'&format=xml&polygon=0&addressdetails=0&limit=1';
```

Fonte: Autor

O Step *Delay row* (Figura 12) é configurado para realizar a requisição a cada segundo, respeitando a restrição da API. Em seguida o Step *HTTP Client* utiliza-se da configuração da URL final, solicita a requisição na API, codifica e configura a resposta de saída. O Step *Coordenadas* recebe a resposta da *HTTP Client* em formato XML, configura o caminho da requisição na Aba *Content*, opção *Loop XPath*, assim apresentado *"/searchresults/place"*. Na aba *Fields* o campo *XPath* é configurado para definir os campos selecionados da resposta da requisição. Neste caso são selecionados os campos da *lat* e *lon*. O tratamento dos dados é efetuado nos Steps *Modified Java Script Value 2* e *Select values 2*. Finalmente no Step *Insert / Update* é realizada a carga das coordenadas geográficas na tabela DIM_Endereco no DW.

A API do Google também foi utilizada nesta pesquisa, com o objetivo de complementar os dados, repetindo o mesmo processo do código URL MAP 2 (mostrado na Figura 12), no entanto, com alteração no link de requisição, que foi alterado conforme apresentado no Código 3.

Código 3 - Geolocalização com API Google

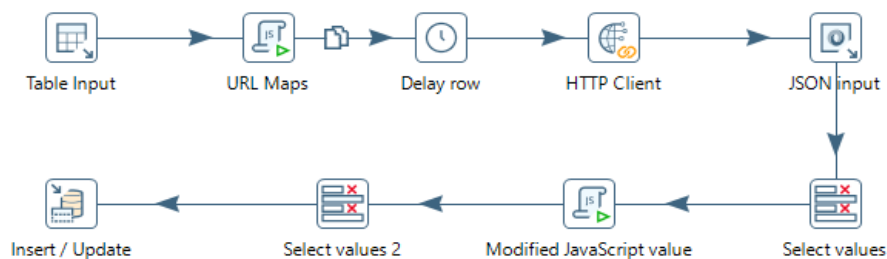
```
var endereco = numero + ', ' + logradouro + ', ' + cidade + ', ' + uf;
var url_final = 'https://maps.googleapis.com/maps/api/geocode/xml?address='+ endereco+'&key=YOU_API_KEY';
```

Fonte: Autores

Para verificar a existência de servidor aposentado sem as suas informações de localização geográfica, principalmente por algum erro nos dados do endereço, foi realizado o mesmo processo ELT ilustrado na Figura 7 com a diferença que a conexão é realizada com base de dados da tabela DIM_Endereco.

Em seguida, para calcular a distância e o tempo despendidos pelos servidores no deslocamento de sua residência até o campus da UFV, foi utilizada a API do Openrouteservice e a opção de locomoção por automóvel. A Figura 13 apresenta esse processo que, por meio das coordenadas geográficas (latitude e longitude), obtém-se a distância e o tempo.

Figura 13 - Cálculo de distância com API do Openrouteservice



Fonte: Autor

O Step *Table Input* (Figura 13) conecta à base de dados da tabela DIM_Endereco, de onde os dados são extraídos. Neste processo não é necessário o tratamento dos dados e foi utilizado o Step *JSON input*, em substituição ao *Get data from XML* dos modelos anteriores. O Step *URL Maps* utiliza o procedimento para a “url_final” conforme especificado no Código 4.

Código 4 - Geolocalização Openrouteservice

```
var endereco = lon+', '+lat;
var url_final = 'https://api.openrouteservice.org/v2/directions/driving-car?api_key=YOU_API_KEY
&start=-42.87015615447366,-20.761026467092154&end='+endereco;
```

Fonte: Autor

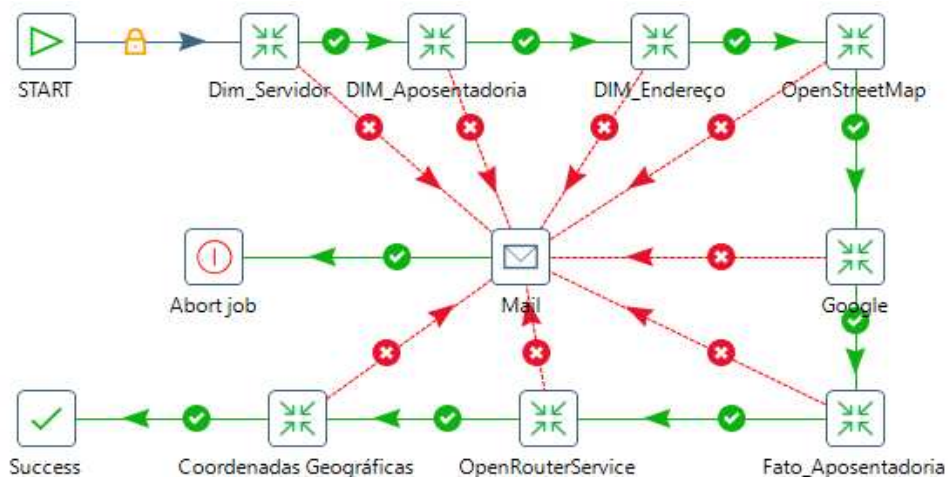
No Código 4, a variável endereço é composta pela combinação lon e lat. A variável url_final, que corresponde a um link da API do Openrouteservice, é composta pelos campos: endereço do sítio Openrouteservice; no campo driving –selecionar a opção “car”; o campo api_key é preenchido com a chave fornecida pelo sítio da API

após cadastro; o campo start é configurado com as coordenadas geográficas da UFV e campo end, configurado com o endereço da posição geográfica atual dos servidores.

O Step *Delay row* (Figura 13) é configurado para realizar até 40 requisições por minuto, respeitando a restrição da API. O Step *HTTP Client* utiliza a configuração da URL final, solicita a requisição na API, codifica e configura a resposta de saída. O Step *JSON input* recebe a resposta da consulta em formato JSON, efetua a configuração na aba *Fields*, campo *Path*, onde seleciona os dados da requisição: “duration” e “distance”. O tratamento dos dados é efetuado nos Steps *Select values*, *Modified JavaScript value* e *Select values 2*, e por último é realizada a carga na tabela *Fato Servidor Aposentadoria* pelo Step *Insert / Update*.

No processo ETL foi desenvolvido um *Job*, Figura 14, com todas as rotinas de transformações, desenvolvidas para realizar a carga no DW.

Figura 14 - Processo de rotinas do job



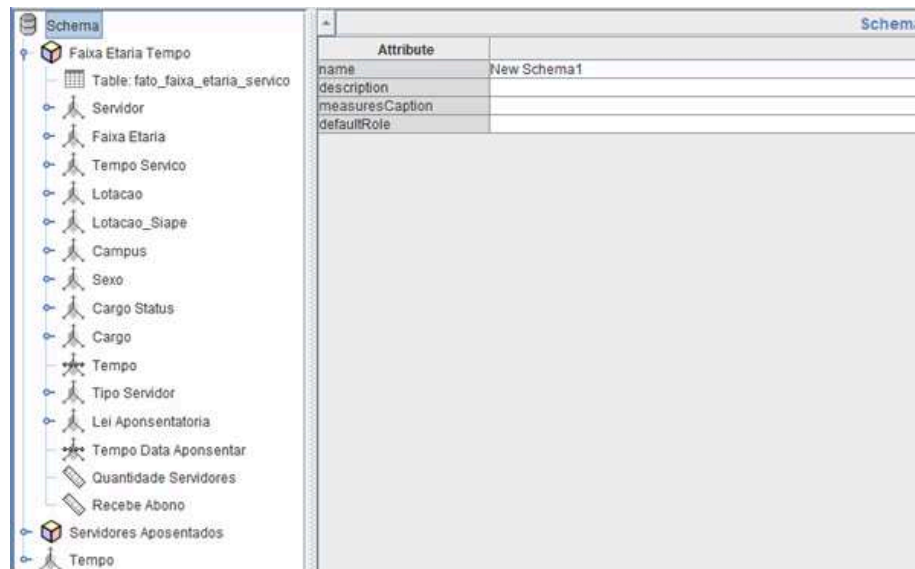
Fonte: Autor

O Step *START* (Figura 14) inicia o processo da carga dos dados para o DW, na seguinte ordem: DIM_Servidor; DIM_Aposentadoria; DIM_Endereço; OpenStreetMap; Google; Fato_Aposentadoria; OpenRouterService; Coordenadas_Geográficas; e Success. Caso ocorra algum erro em alguma etapa do processo, um e-mail é enviado para o administrador utilizando o Step *Mail* e a execução é abortada pelo Step *Abort Job*.

5.1.3 Cubo OLAP

O processo ETL termina com a construção de dois cubos OLAP: “Faixa Etária Tempo” e “Servidores Aposentados”. A Figura 15 ilustra o cubo na ferramenta *Schema Workbench Mondrian*.

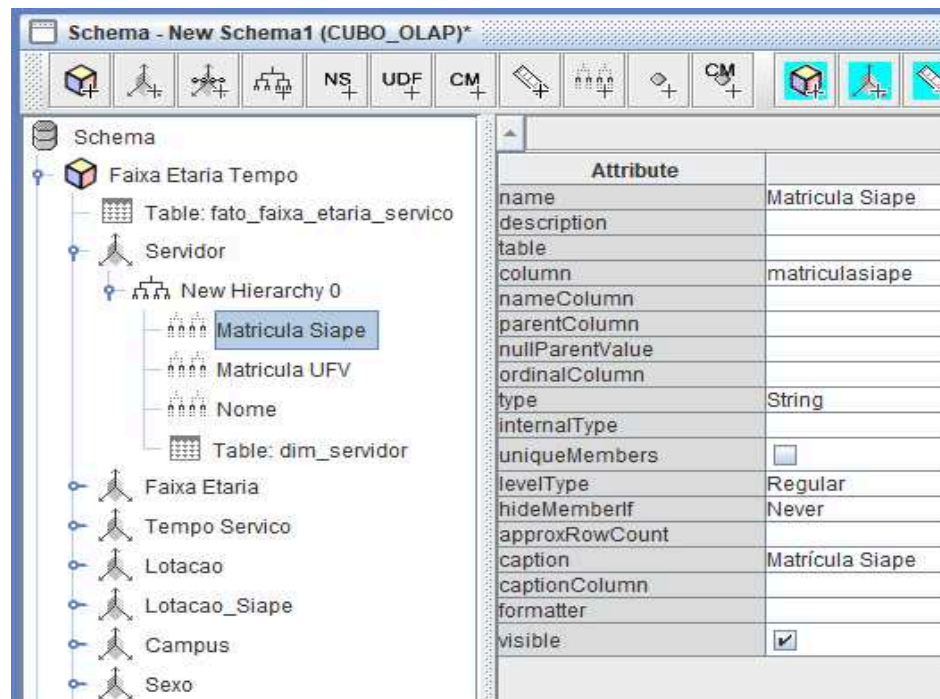
Figura 15 - Estrutura do cubo do OLAP



Fonte: Autor

Na coluna à esquerda da Figura 15 pode-se observar a estrutura do DM criado, a qual corresponde à modelagem multidimensional apresentada na Figura 5, com as duas tabelas fato e algumas tabelas dimensões. Inicialmente foi criado um Schema no qual foram adicionados dois cubos e em seguida, foi adicionado ao Schema um cubo denominado “Faixa Etária Tempo” e dentro dele uma tabela apontando para a *Fato Faixa Etária Tempo*. Para cada tabela de dimensão do DM foi criada uma dimensão no cubo. Conforme ilustrado na Figura 16, na criação de uma nova dimensão é necessário adicionar uma tabela apontado para a dimensão desejada. Em seguida adiciona-se uma hierarquia, em que são configurados os campos: “name” – nome da hierarquia; “allMemberName” – Nome que irá aparecer no cubo quando a busca for todos os dados da hierarquia; “primaryKey” – apontado para chave primária da dimensão. Logo depois é acrescentado os níveis na hierarquia devendo ser configurados os campos “name”; “column”; “type”; “levelType”; e “hideMemberIf”. A Figura 16 ilustra a dimensão “Servidor”, que foi apontada para a tabela de dimensão “DIM_Servidor”. Observa-se que foi adicionada uma hierarquia com os níveis de “Matricula Siape”, “Matricula UFV”, e “Nome”.

Figura 16 - Desenvolvimento da Dimensão Servidor



Fonte: Autor

Posteriormente, foram adicionadas as outras dimensões, conforme ilustrado na Figura 16, ficando na seguinte ordem: Servidor; Faixa Etaria; Tempo Servico; Lotacao; Campus; Sexo; Cargo; Cargo Status; Tempo; Tipo Servidor; e Tempo Data Aposentar. Observa-se que a dimensão Tempo está fora do cubo, por ser genérica e poder ser aproveitada em outros cubos no Schema. As medidas do cubo “Faixa Etaria Tempo” utilizadas nessa pesquisa foram as “Quantidade Servidores” e “Recebe Abono”.

Os mesmos passos foram realizados para a criação do Cubo “Servidores Aposentados” que resultou nas seguintes dimensões: Servidor; Faixa Etária; sexo; Faixa Aposentadoria; Endereco; Aposentadoria; e Tempo. A medida do cubo “Servidores Aposentados” utilizadas foi a “Quantidade Servidores Aposentados”.

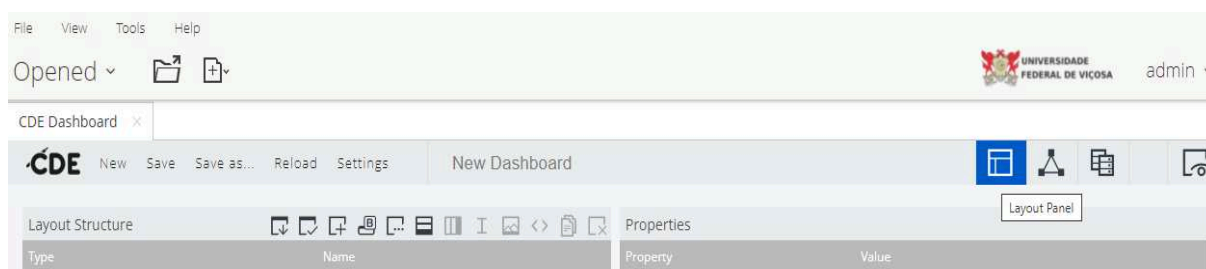
Para finalizar, o cubo foi publicado no servidor da Pentaho e pode ser visualizado nas ferramentas da Pentaho para a criação dos gráficos.

5.1.4 Visualização dos dados

A ferramenta utilizada para a visualização dos cubos e desenvolvimento dos gráficos foi o Pentaho Server CDE na versão 8.3. Para a construção das consultas foi utilizada linguagem Multidimensional Expressions (MDX), da ferramenta Saiku Analytics, que permite a escolha de medidas, a movimentação das dimensões e, com isso, a utilização das principais funcionalidades do cubo OLAP (MACHADO, 2007).

A Figura 17 ilustra a ferramenta CDE Dashboard, utilizada para a construção de painéis, com destaque nas três abas principais à esquerda.

Figura 17 - Ferramenta CDE Dashboard Pentaho



Fonte: Autor

- Layout Panel – onde é realizada a criação do layout do dashboard. Podendo ser implementados nas linguagens de css e java script. Além de disponibilizar templates prontos;
- Components Panel - refere-se aos componentes dos gráficos que podem ser inseridos no dashboard componentes como pizza, coluna, linha, mapas e outros. Também pode-se inserir campos de texto, selects, realizar manipulações com dos dados por passagem de parâmetros e dentre outros;
- Datasources Panel – refere-se à conexão com a base de dados do DW, que pode ser por meio das conexões sql, MDX, script, nosql e outras.

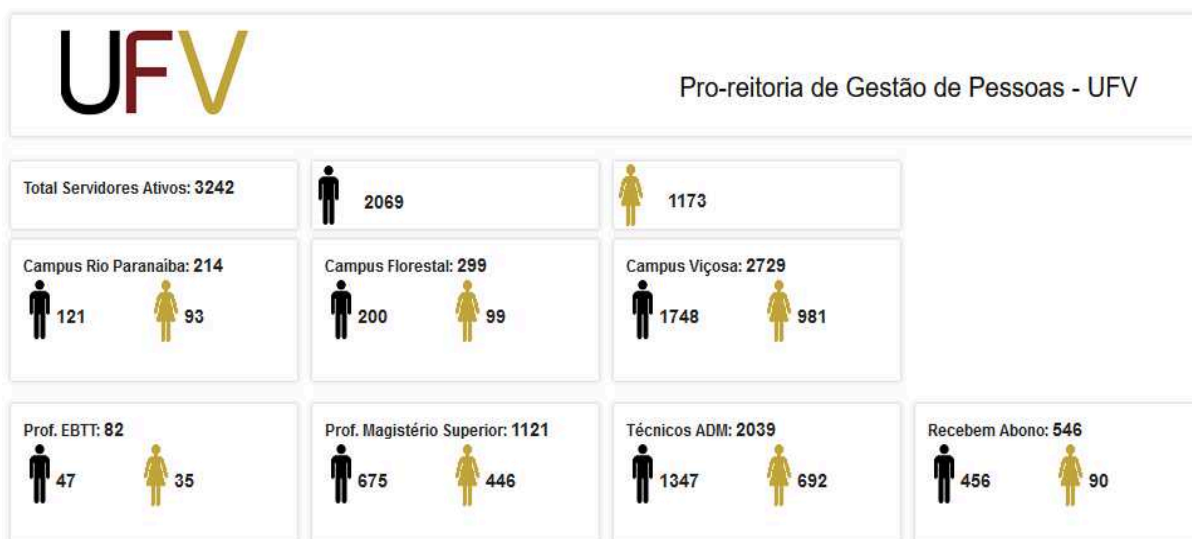
Para a criação dos dashboards, foram configuradas três abas no CDE Pentaho. Na aba Layout Panel foram utilizadas as configurações de css e java script para o desenvolvimento de layouts personalizados. Na aba Components Panel foram utilizados os gráficos de pizza e barras, tabelas e mapa. Já na aba Datasources Panel foram utilizadas apenas as conexões de consulta SQL com as opções “*SQL Queries*” e “*sql over sqlJndi_*” e as conexões MDX com as opções “*MDX Queries*” e “*mdx over mondrianJndi*”.

Para a emissão dos relatórios a ferramenta Pentaho Report Designer foi instalada e configurada para conexão com o DW. Em seguida, foi possibilitado a criação dos relatórios de acordo com a necessidade e solicitação do órgão. Por fim, os relatórios foram acoplados nos dashboards solicitados.

5.2 Apresentação e Discussão dos Resultados

A construção do Data Mart permitiu configurar diversas funcionalidades a partir da compilação dos dados da pesquisa. A Figura 18 ilustra o dashboard principal do sistema, com os quantitativos dos servidores ativos distribuídos por gênero; *Campus*; carreira de Técnico Administrativos e Docentes; e aqueles que recebem o abono permanência. Apresenta, ainda, botões com os indicadores de desempenho do projeto, que ao serem acionados é direcionado para um novo painel. Essa tela principal apresenta dados da última carga do DW, que foi realizada no mês de fevereiro de 2021 e apresenta um total de 3.242 servidores ativos.

Figura 18 - Quantitativos de Servidores Ativos da última carga no DW



Indicadores:

Faixa Etária e Tempo de Serviço

Servidores Lotação

Projeção de Aposentadoria

Servidores Aposentados

Fonte: Autor

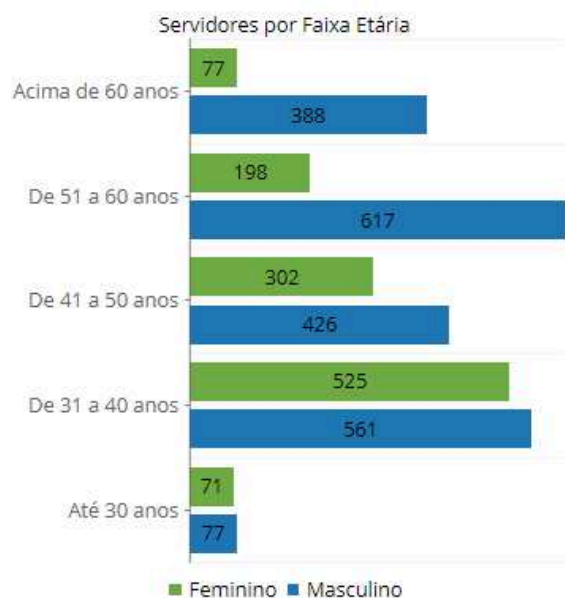
Essa visualização inicial consiste em um ganho com a aplicação da ferramenta, pois permite visualizar a composição da força de trabalho, a lotação por *campi*, a classificação por gênero e a projeção de possíveis aposentadorias. Este nível inicial de informação permite a elaboração de diversas políticas direcionadas à gestão e desenvolvimento de pessoas. Ao acionar o botão “Faixa Etária e Tempo de Serviço” Figura 18, será transferido para um novo Dashboard, ilustrado nas Figuras 19 e 20. A Figura 19 apresenta os dados totalizados dos servidores dos três *campi* da UFV, em função da tabela *Fato Faixa Etária Tempo*. Ao selecionar um *campus* no gráfico de pizza, é passado um parâmetro que modifica o dashboard e são apresentados os dados apenas do *campus* selecionado. Ao selecionar o botão “Servidores por *Campi*” o Dashboard volta a configuração inicial com os dados dos três *campi*. Outra funcionalidade que pode ser destacada consiste na realização de consultas históricas.

Avaliando os resultados, verifica-se a existência de informações condensadas e robustas passíveis de avaliação para o estabelecimento das políticas relacionadas à gestão e desenvolvimento de pessoas. Observa-se pelos resultados um total de 1.280 servidores, representando 39% do total do quadro efetivo, classificados nas faixas etárias de 51 a 60 anos e Acima de 60 anos, além da indicação de 808 servidores com tempo de serviço acumulado maior que 30 anos. O uso dessa funcionalidade indica um perfil demográfico da força de trabalho existente, traduzindo em preocupação de novas reposições e da manutenção da experiência consolidada na Instituição ao longo dos anos.

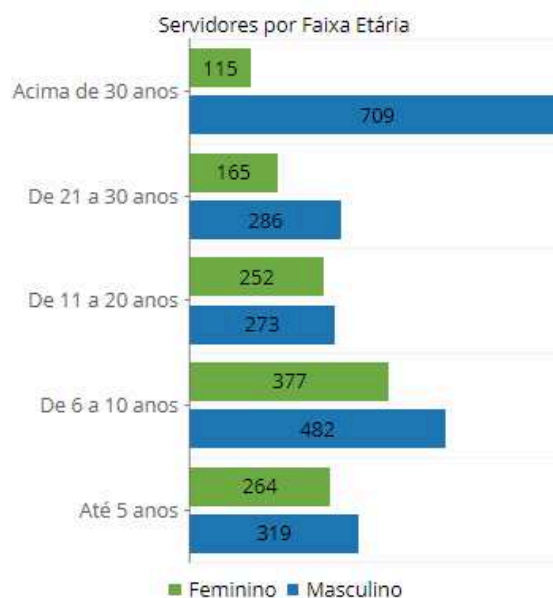
Figura 19 - Quantitativos de Servidores Ativos pelo Campi UFV



Servidores por Faixa Etária - Todos os Campi



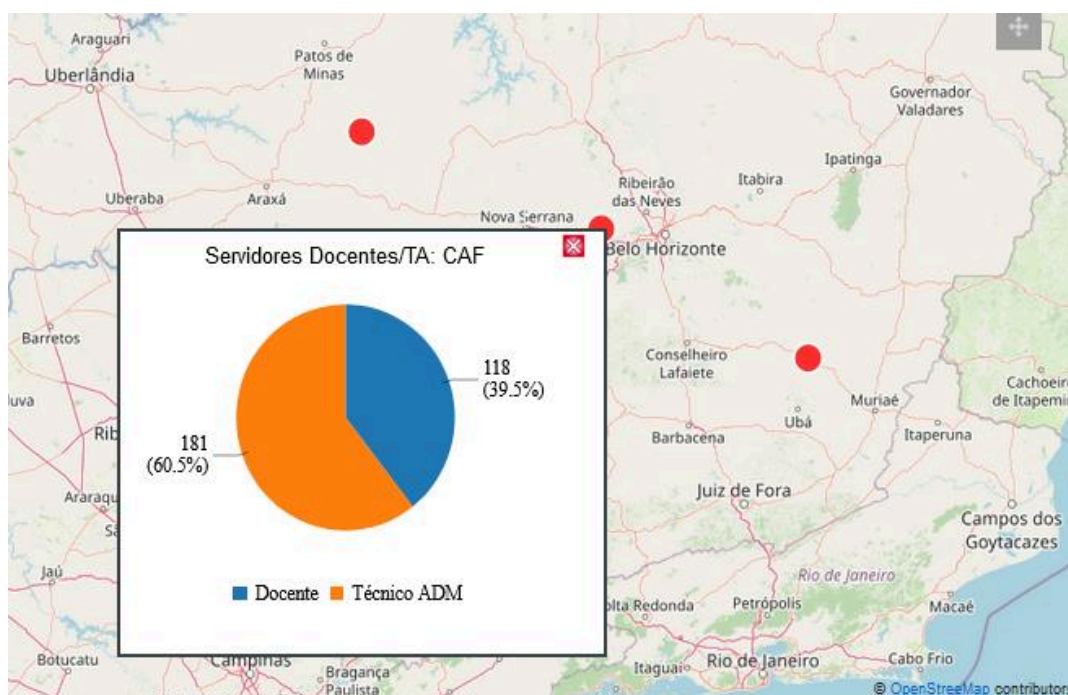
Servidores por Tempo Serviço - Todos os Campi



Fonte: Autor

A funcionalidade geocodificação, visualizada na Figura 20 do mesmo dashboard, foi gerada pela ferramenta do BI com dados do OpenStreetMap. Os três *campi* são pontuados no mapa, que carrega o quantitativo geral de servidores, classificados na carreira Magistério Superior e de Técnico Administrativo em Educação. O mapa pode ser apresentado com os dados gerais ou com os dados de cada *campi* passado por parâmetro no dashboard.

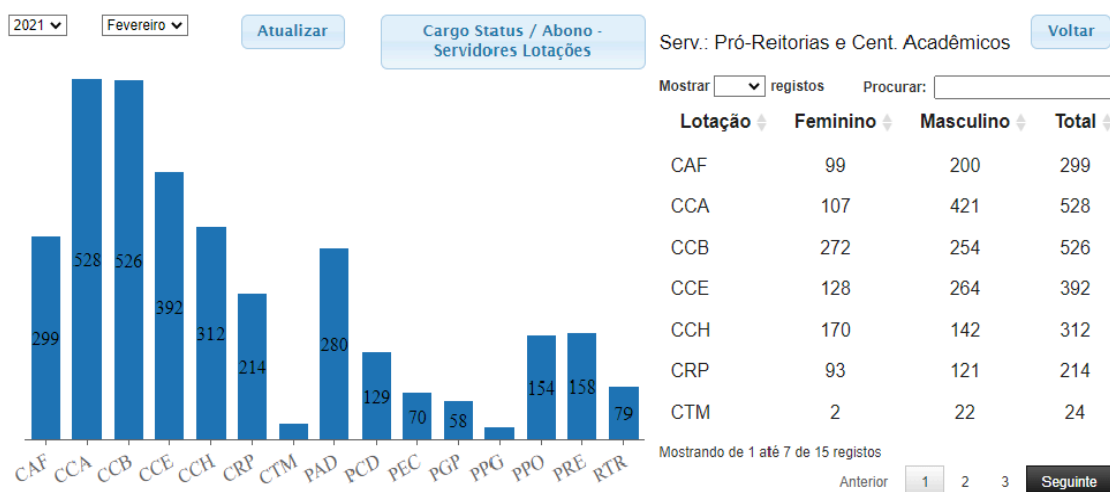
Figura 20 - Quantitativos de Servidores Ativos por Campi UFV no mapa



Fonte: Autor

A partir da mesma tabela Fato, por meio do botão “Servidores Lotação” da tela principal (Figura 18), obtém-se um novo dashboard que apresenta o quantitativo dos servidores por cargo, gênero, *campi* e lotação, ilustrados nas Figuras 21 e 22. A Figura 21 ilustra as lotações por Pró-Reitorias, Centros acadêmicos, *Campi* e o Centro Experimental do Triângulo Mineiro-(CTM). Ao selecionar uma barra do gráfico, o layout do dashboard procede a alteração para a opção selecionada. Observa-se também a possibilidade de emissão de uma tabela com os quantitativos dos servidores por lotação e gênero que podem ser alterados de acordo com parâmetro atual do layout. Esse mesmo Dashboard disponibiliza dados quantitativos dos servidores ilustrados por gráficos de barra, com a faixa de etária e tempo de serviço por lotação de acordo com parâmetro atual da tela.

Figura 21 - Quantitativos de Servidores por Lotação

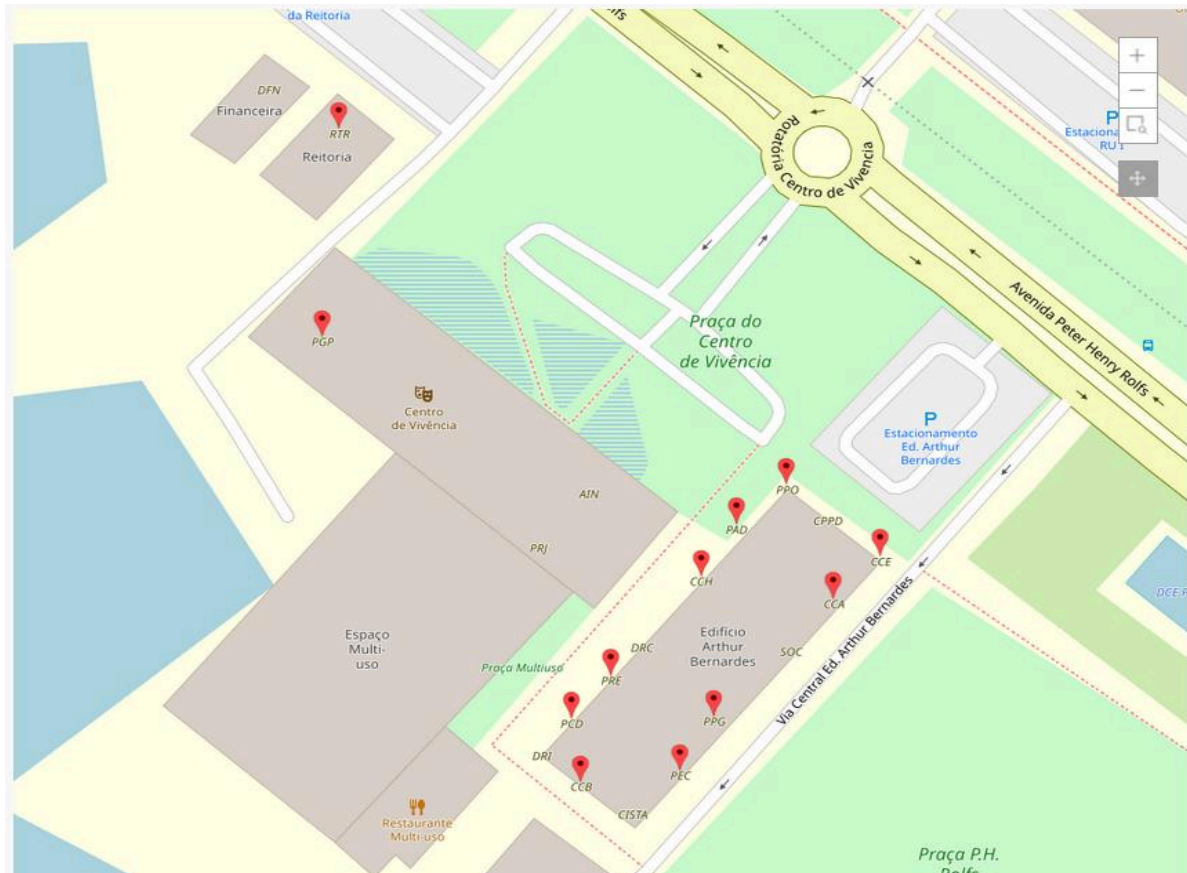


Fonte: Autor

Essas funcionalidades podem ser visualizadas na mesma tela que apresenta o quantitativo dos servidores com geolocalização das lotações, conforme ilustrado na Figura 22. Ao selecionar uma barra do gráfico da Figura 21, é passado um parâmetro para o mapa que atualiza os dados, de acordo a lotação desejada.

Quando acionado, o botão ‘Cargo Status/Abono – Servidores Lotações’ da Figura 21, o parâmetro da lotação atual é passado para um novo dashboard, ilustrado na Figura 23 e na Tabela 4. Neste novo painel são disponibilizadas as informações sobre o quantitativo de servidores por cargo, que recebem o *Abono Permanência*, e nesse caso aptos a solicitarem a aposentadoria (Figura 23). Acrescenta-se uma outra fonte importante de informação que se traduz no *status* do cargo, considerado *Ativo* - aquele apto a realizar concurso para os códigos vagos; e *Vedado* – referindo-se aos cargos com proibição temporária de realização e concursos para provimento e *Extinto* – refere-se aos cargos impossibilitados de provimento, considerando que todos que vierem a vagar são recolhidos do sistema. Essa funcionalidade consiste numa importante fonte de informação para os gestores, considerando a possibilidade de efetuarem o planejamento de recomposição da força de trabalho, ampliação do nível de terceirização e a definição de lotação ideal para cada órgão.

Figura 22 - Geolocalização das Lotações



Fonte: Autor

Essa funcionalidade se expande também para os servidores que já recebem abono permanência, distribuídos por status do cargo e por gênero conforme previsto na Figura 23. Uma importante fonte de informação para o gestor consiste no conhecimento em tempo real, do cenário atual da Instituição em que 546 servidores, representando mais da metade da sua totalidade, estão classificados em status cujo cargo está extinto ou vedado para concurso. A Figura 23 ilustra também um quadro geral de todos os servidores, incluídos os que recebem o “*Abono Permanência*”, por gênero e status do cargo. Pode-se observar um número alto com o *status* cargo “*Extinto*” e “*Vedado*”. Fica uma incógnita para os gestores para saber como esses servidores poderão ser substituídos, sem a realização de concursos e recursos.

Figura 23 - Quantitativos de servidores com Abono de Permanência e de servidores por gênero e status do cargo



Fonte: Autor

O mesmo dashboard disponibiliza outras informações de caráter geral e mais abrangente conforme disposto na Tabela 4. Esta funcionalidade permite ao gestor identificar de forma comparativa a estrutura de cada órgão na Instituição com a formação do seu quadro de pessoal com a identificação nominal, cargo e status do cargo, a divisão por carreira e a situação funcional temporal do servidor, expressa pela identificação do abono permanência. O usuário poderá também selecionar os parâmetros de lotação para sua consulta. Na Tabela 4 foram utilizados os parâmetros: Pró-Reitorias e Centros Acadêmicos. Ao acionar um campo da coluna “Sigla”, um novo dashboard é aberto com as informações qualificadas dos servidores daquela lotação, podendo nesse caso efetuar e identificar a situação funcional de cada servidor levando o gestor ao conhecimento integral de cada órgão especificado.

Na tabela *Fato Faixa Etária Tempo*, por meio do botão “Projeção de Aposentadoria” da tela principal (Figura 18), apresentadas nas Figuras 24 e 25 e na Tabela 5, permite-se projetar a situação relacionada a aposentadoria dos servidores para os próximos 5 anos e ao mesmo tempo identificar os servidores ativos que recebem o abono permanência, conforme ilustrado na Figura 24.

Tabela 4 - Quantitativo Servidores por Lotação Cargo Status

Mostrar <input type="text"/> registros			Procurar: <input type="text"/>							
Unidade	sigla	nome	Cargos			Docente		Técnico Administrativo		total
			Ativo	Vedado Concurso	Extinto	Total	Com Abono	Total	Com Abono	
CCH	CCH	Centro de Ciências Humanas, Letras e Artes	6	2	1	0	0	9	1	9
CCH	DAD	Departamento de Administração e Contabilidade	30	1	1	25	2	7	0	32
CCH	DAH	Departamento de Artes e Humanidades	11	1	1	8	0	5	0	13
CCH	DCM	Departamento de Comunicação Social	11	6	0	9	0	8	1	17
CCH	DCS	Departamento de Ciências Sociais	15	2	0	12	1	5	0	17
CCH	DED	Departamento de Economia Doméstica	35	4	7	24	3	22	5	46
CCH	DEE	Departamento de Economia	19	2	0	15	0	6	1	21
CCH	DGE	Departamento de Geografia	11	2	1	9	0	5	0	14
CCH	DHI	Departamento de História	14	0	0	10	0	4	0	14
CCH	DLA	Departamento de Letras	40	2	1	35	1	8	0	43
CCH	DPD	Departamento de Direito	25	1	1	20	2	7	0	27
CCH	DPE	Departamento de Educação	56	3	0	49	8	10	0	59

Fonte: Autor

Figura 24 - Dashboard com projeções de aposentadoria próximos 5 anos



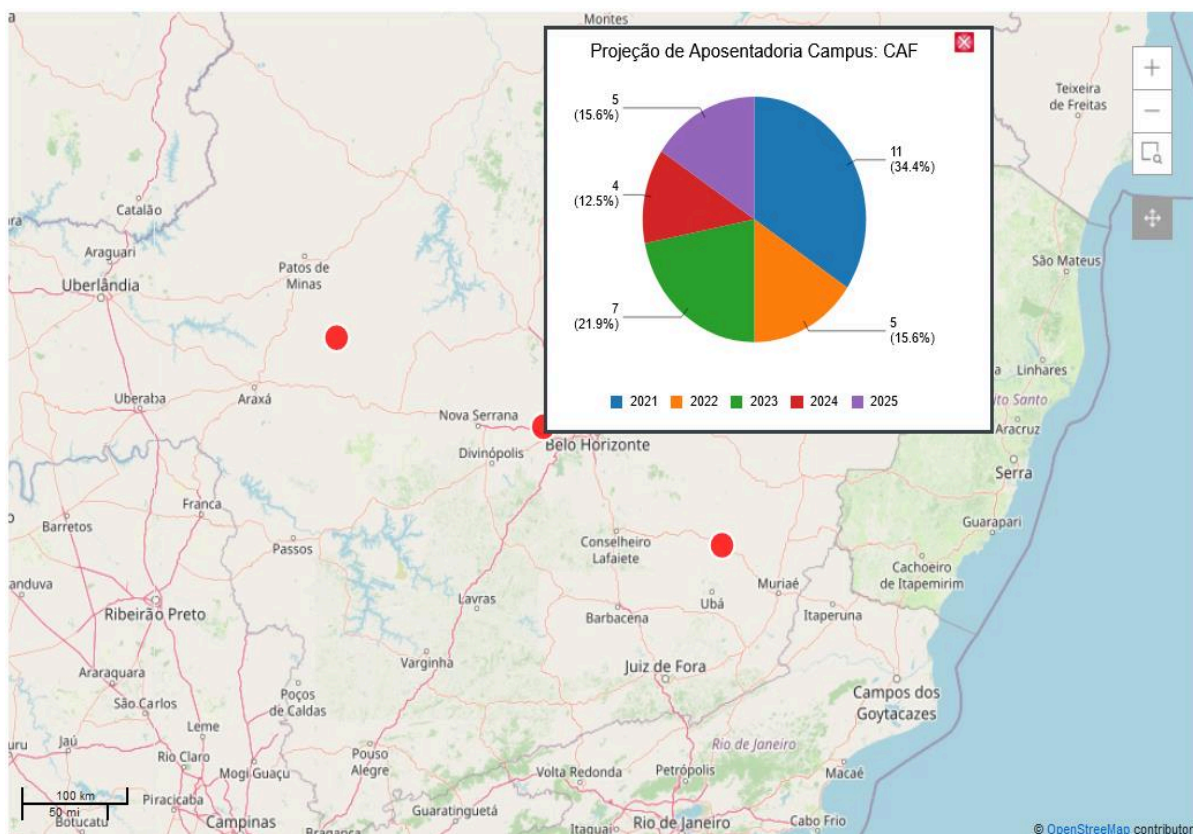
Fonte: Autor

O design e o layout estabelecidos permitem identificar de forma clara as informações relacionadas à situação funcional dos servidores e ao mesmo tempo fornece uma visão integrada da estrutura de cada órgão, que são importantes fontes de informação para a elaboração das políticas de pessoal na Instituição em seus três Campi. Nesta mesma tela, fica disponível um botão “Relatórios PDF”, que ao ser selecionado disponibiliza um relatório com todos os servidores que satisfazem os requisitos para a solicitação do processo de aposentadoria nos próximos 5 anos.

Observa-se, ainda, um quadro com uma lupa e o quantitativo dos servidores que podem solicitar a aposentadoria ao longo do período de 5 anos, mas que não recebem o *Abono Permanência*. Ao clicar sobre a lupa em um determinado ano, um relatório no formato “PDF” e “XLS” é aberto com os seguintes campos: Matrícula Siape; Nome; Data Aposentadoria; Lotação; Cargo; e Status do Cargo. À direita da Figura 24, o ícone “**Abono**” apresenta o quantitativo de servidores que recebem o abono permanência, em forma de um relatório em formato “PDF”.

No mesmo dashboard observa-se também um mapa do OpenStreetMap com dados do Nominatim, Figura 25, da localização dos *campi* de Rio Paranaíba (CRP), de Florestal (CAF) e de Viçosa (UFV), com o quantitativo de projeção de aposentadoria nos próximos 5 anos. Nesta mesma Figura, ao acionar o *Campus* Florestal (CAF), verifica-se um total de 32 servidores que atendem os requisitos para aposentadoria. No *Campus* de Rio Paranaíba apenas 1 servidor e no *Campus* de Viçosa, esse número é de 304 servidores.

Figura 25 - Mapa com projeções de aposentadoria nos próximos 5 anos por Campus



Fonte: Autor

O mesmo dashboard fornece a Tabela 5 com os quantitativos de projeções de aposentadoria, sem levar em consideração os servidores que recebem *Abono Permanência*, por cargo e status dos cargos nos próximos 5 anos.

A tabela *Fato Servidor Aposentadoria*, por meio do Botão “Servidores Aposentados” da tela principal (Figura 18), ilustradas nas Figuras 26 e 27 e na Tabela 6, permitiu o desenvolvimento de um dashboard por meio da compilação dos dados referente ao mês de novembro de 2020, ou seja, por se tratar de servidores aposentados, esse processo ETL será realizado uma vez no ano.

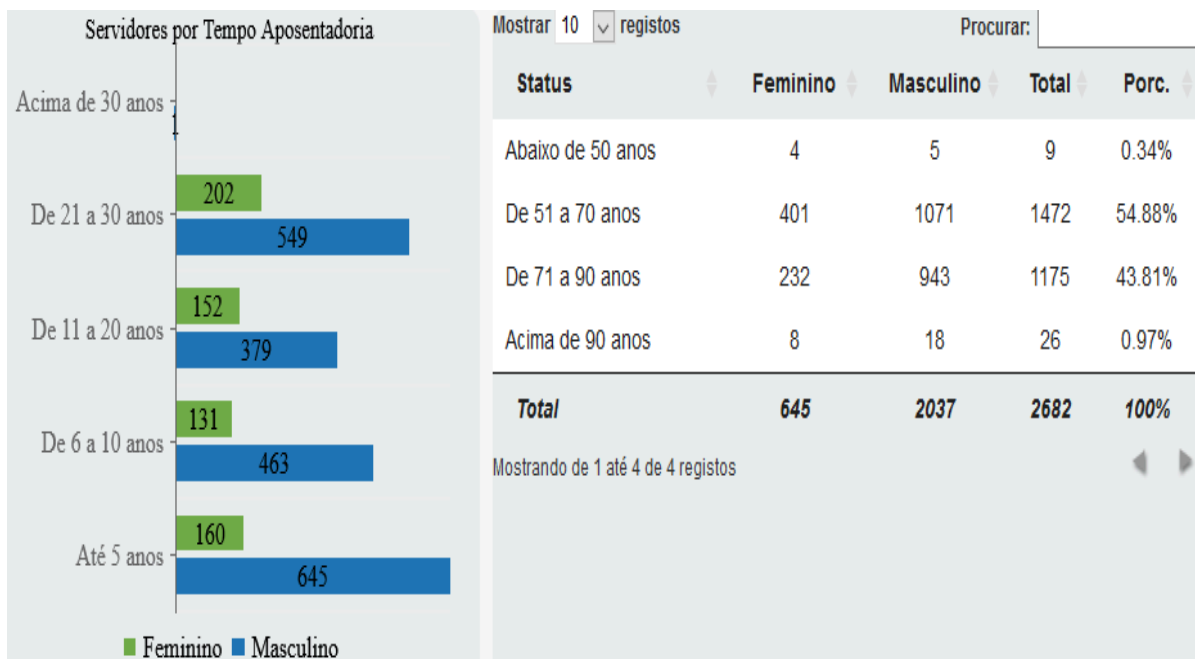
O modelo desenvolvido apresenta diversas funcionalidades como o agrupamento dos 2.682 servidores por tempo de aposentaria e gênero e permite a sua atualização com os cenários estabelecidos ao longo dos anos. A informação, além de preservar o histórico funcional, auxilia no estabelecimento de políticas efetivas relacionadas ao dimensionamento de pessoal, as questões de gênero, acessibilidade, além de traçar o perfil dos ex servidores da Instituição. A Figura 26 apresenta os dados de servidores aposentados por gênero e por faixa de tempo de aposentadorias.

Tabela 5 - Projeções de Aposentaria por cargo próximo 5 anos

Mostrar <input type="text"/> registros		Procurar: <input type="text"/>											
Cargo.New Hierarchy 0	Ativo				Extinto				Vedado				
	2021	2022	2023	2024	2021	2022	2023	2024	2021	2022	2023	2024	
ADMINISTRADOR			1										
APONTADOR							1						
ARMADOR					1								
ARQUITETO E URBANISTA			1										
ASSISTENTE DE LABORATÓRIO							1		3				
ASSISTENTE DE SOM								2					
ASSISTENTE EM ADMINISTRAÇÃO		1	3										
ASSISTENTE SOCIAL	1												
AUXILIAR DE AGROPECUÁRIA					8	5	8	7					
AUXILIAR DE ENFERMAGEM													1
AUXILIAR DE LABORATÓRIO							2		1				
AUXILIAR EM ADMINISTRAÇÃO										4	3	6	6
AUXILIAR OPERACIONAL					4	3	7	2					
BIBLIOTECÁRIO-DOCUMENTALISTA				1									
BOMBEIRO HIDRÁULICO					2								
CARPINTEIRO					1	1							
CONTÍNUO					3	3	3						

Fonte: Autor

Figura 26 - Quantitativo de servidores aposentados, separados por gênero, tempo de aposentadoria e faixa etária



Fonte: Autor

Uma outra funcionalidade complementar, presente em parte da Figura 26, remete-se a tabela com o agrupamento por faixa etária e gênero, ampliando a possibilidade de agregação de informações aos indicadores construídos. A seleção de qualquer linha na tabela, aponta um novo dashboard detalhando nome, idade, distância e tempo de percurso até a universidade, cidade, estado e País e carrega um mapa com a posição geográfica dos servidores aposentados. A coluna Status com a opção “Acima de 90 anos” foi acionada, apresentando um total de 26 servidores, sendo 23 localizados na cidade de Viçosa, 1 em Belo Horizonte e 2 em Florestal. A Tabela 6 apresenta o resultado dos servidores por idade, em ordem decrescente.

Tabela 6 - Quantitativos servidores aposentados com sua distância da localidade atual até o Campus UFV

Idade	Distância	Tempo	Cidade	UF	País
98	2.3 km	00:04	VIÇOSA	MG	BRASIL
97	2.7 km	00:05	VIÇOSA	MG	BRASIL
97	1.7 km	00:03	VIÇOSA	MG	BRASIL
96	280.2 km	04:42	FLORESTAL	MG	BRASIL
96	3.2 km	00:06	VIÇOSA	MG	BRASIL
95	1.6 km	00:03	VIÇOSA	MG	BRASIL
95	2.2 km	00:04	VIÇOSA	MG	BRASIL
94	2.9 km	00:06	VIÇOSA	MG	BRASIL
94	7.9 km	00:10	VIÇOSA	MG	BRASIL
93	1.6 km	00:03	VIÇOSA	MG	BRASIL

Fonte: Autor

O processo ETL com Geocodificação da API do Nominatim, conforme ilustrado na Tabela 7, conseguiu informar a latitude e a longitude de 1.840 servidores, em aproximadamente 33 minutos, que representam 68,60% dos dados. De forma complementar com a utilização da Geocodificação da API do Google obteve-se a geolocalização de 831 servidores, em aproximadamente 6 minutos, com representação de 30,98% dos dados.

Tabela 7 - Quadro quantitativo dos processos de API para Geocodificação

Processo	Geolocalização Servidores Aposentados		
	Quantidade	Porcentagem em	Tempo
Nominatim	1840	68,60%	~ 33 minutos
Google	831	30,98%	~ 6 minutos
Não Encontrados	11	0,42%	
Total	2682	100%	~ 39 minutos

Fonte: Autor

O resultado da geocodificação pode ser visualizado na Figura 27 a partir do mapa gerado com a utilização do componente MapComponent do Pentanho CDE e o OpenStreetMap.

Figura 27 - Geolocalização Geográfica Servidores Aposentados



Fonte: Autor

Com concentração no Estado de Minas Gerais, observa-se a presença de servidores aposentados nos estados do Rio Grande do Sul, Rondônia, Piauí, Rio Grande do Norte, dentre outros. O conhecimento destes dados de forma agregada permite traçar o perfil dos servidores egressos, estabelecer contato de forma mais rápida e, principalmente projetar políticas voltadas aos servidores aposentados, fortalecendo os laços anteriormente construídos, além da promoção da autoestima e sentimento de pertencimento a Instituição.

Até aqui, pode-se visualizar os dados extraídos utilizando-se as ferramentas desenvolvidas com a abordagem do conceito BI. A Tabela 8 sintetiza, comparativamente, as vantagens obtidas com o desenvolvimento dessas ferramentas em relação a ferramenta disponível no início desse trabalho.

Tabela 8 - Comparação entre as características do banco de dados operacional e do Data Mart desenvolvido

Informações	Bancos de Dados Relacional (OLTP)	Data Mart
Dados analíticos (Histórico, Geolocalização)	Não	Sim
Tempo necessário para realização de consultas	Indefinido*	Forma Rápida
Conhecimento de Linguagem SQL para o usuário final	Sim	Não
Produção de Relatórios em tempo hábil	Não	Sim
Ferramentas <i>Open source</i>	Não	Sim

* Depende do nível de complexidade da consulta, do nível de conhecimento de SQL do usuário, da disponibilidade dos servidores da DTI, e outros.

Fonte: Autor

6 CONCLUSÕES

O Data Mart desenvolvido com as ferramentas *open source* do BI e com a integração de outros sistemas possibilitou a obtenção de informações dos servidores ativos e aposentados da Instituição em estudo. As funcionalidades geradas permitiram a extração de dados quantitativos dos servidores aposentados relacionados a questões de gênero; faixa etária; tempo de aposentadoria e geolocalização atual. Dos servidores ativos foi possível compilar dados referentes à sua localização e lotação por *campi*; gênero; faixa etária; tempo serviço; localização geográfica por lotação; projeções de aposentadoria nos próximos anos; cargo e status do cargo.

Os dados sobre servidores aposentados possibilitaram traçar de forma fidedigna um novo perfil destes servidores e os resultados encontrados permitem o estabelecimento de novas políticas voltadas para esse extrato de servidores.

Paralelamente, foi possível analisar o quadro de força de trabalho dos servidores ativos na UFV. Os dados revelaram um alto número de servidores que estão aptos a solicitarem o processo de aposentadoria, evidenciado a necessidade do estabelecimento de políticas voltadas para a gestão e desenvolvimento de pessoas considerando que parte das informações evidencia a impossibilidade de reposição de vagas de determinados cargos, também classificados como extintos ou vedados para concurso.

Todo o processo foi realizado utilizando apenas sistemas de software livre, o que implica em disponibilizar as ferramentas do conceito de BI para tomada de decisão gerencial sem adicionar custos com aquisição de software proprietário. Acredita-se que esta ferramenta possa ser implementada em outras IFEs com pouco esforço, uma vez que além do banco de dados do SIAPE, o DM extrai dados de um banco de dados próprio da UFV.

As ferramentas BI, produto dessa pesquisa, já estão sendo adotadas pela Pró-Reitoria de Gestão de Pessoas da UFV, considerando principalmente a agilidade na obtenção de informações de forma agregada, que auxiliam no processo de tomada de decisão a partir do Data Mart desenvolvido. Outra aplicabilidade do DW consiste na agregação de dados para produção de novos indicadores relacionados a capacitação dos servidores como: progressão funcional; acompanhamento de servidor em

treinamento; levantamento de necessidades de capacitação; e planejamento financeiro para desenvolvimento de servidores.

Essa pesquisa culminou na elaboração de dois artigos que foram submetidos e aceitos. O primeiro, sob o título “Processo ETL em uma Instituição Federal de Ensino - Obtendo Informações Funcionais e Geolocalização de Servidores Aposentados”, foi apresentado na 16^o Conferência Ibérica de Sistemas e Tecnologias de Informação - CISTI 2021, junho/ 2021, na cidade de Chaves, Portugal. O segundo artigo, “O uso de Data Mart para apoio à tomada de decisão na Gestão de Pessoas em uma Instituição Federal de Ensino”, foi apresentado no Congresso da Sociedade Brasileira de Computação – IX Workshop de Computação aplicada em Governo Eletrônico – julho/2021, na cidade de Florianópolis, Brasil. Devido à pandemia do covid-19, ambos os eventos foram realizados por meio de videoconferência.

Os códigos fontes do desenvolvimento do Data Mart encontram-se no sistema github.com (SOUZA, 2021), com a seguinte organização: uma pasta com todo o Processo ETL; uma imagem e o arquivo UML da modelagem multidimensional; o arquivo do cubo OLAP e um arquivo Leia-me. Todos podem ser baixados e adaptados de acordo com as necessidades de cada IFEs.

REFERÊNCIAS

- ASSIS, M. D. Indicadores de Gestão de Recursos Humanos. Rio de Janeiro: Qualitymark, 2014. v. 2.
- BARBIERI, C. BI2: Business Intelligence: Modelagem e Qualidade. Elsevier Editora, 2011.
- BERGUE, S. T. Gestão Estratégica de Pessoas no Setor Público. São Paulo: Atlas, 2014. 25 p.
- BRASIL. Constituição da República Federativa do Brasil. 1988. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 18 ago. 2019.
- _____. DECRETO-LEI Nº 200. 1967. Dispõe sobre a organização da Administração Federal, estabelece diretrizes para a Reforma Administrativa e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/Decreto-Lei/Del0200.htm>. Acesso em: 16 ago. 2019.
- _____. DECRETO-LEI Nº 6.096. 2007. Institui o Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais - REUNI. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6096.htm>. Acesso em: 06 jun. 2020.
- _____. DECRETO-LEI Nº 7.232. 2010. Dispõe sobre os quantitativos de lotação dos cargos dos níveis de classificação “C”, “D” e “E” integrantes do Plano de Carreira dos Cargos Técnico-Administrativos em Educação, de que trata a Lei no 11.091, de 12 de janeiro de 2005, das universidades federais vinculadas ao Ministério da Educação, e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Decreto/D7232.htm>. Acesso em: 05 jun. 2020.

_____. DECRETO-LEI Nº 7.485. 2011. Dispõe sobre a constituição de banco de professor-equivalente das universidades federais vinculadas ao Ministério da Educação e regulamenta a admissão de professor substituto, de que trata o inciso IV do art. 2º da Lei nº 8.745, de 9 de dezembro de 1993. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Decreto/D7232.htm>. Acesso em: 05 jun. 2020.

_____. EMENDA CONSTITUCIONAL Nº 103. 2019. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc103.htm>. Acesso em: 10 out. 2019.

_____. LEI Nº 131. 2009. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências, a fim de determinar a disponibilização, em tempo real, de informações pormenorizadas sobre a execução orçamentária e financeira da União, dos Estados, do Distrito Federal e dos Municípios. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em: 16 ago. 2019.

_____. LEI Nº 13.709. 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em: 16 ago. 2019.

CHIAVENATO, I. Gestão de Pessoas. 3. ed. Rio de Janeiro: Elsevier, 2010.

_____. Introdução à Teoria Geral da Administração. 7. ed. Rio de Janeiro: Campus, 2003.

CAIÇARA JUNIOR, C. Sistemas Integrados de Gestão ERP: Uma Abordagem Gerencial. 1. ed. Curitiba: Editora InterSaberes, 2012.

CAPES. Disponível em: <https://www-periodicos-capes-gov-br.ezl.periodicos.capes.gov.br/index.php?>. Acesso em 15 de novembro 2020.

CASTRO, A. C.; FONSECA, R. M.; MELO, A.; & PRADO, A. Business Intelligence Aplicado a Gestão Acadêmica no Setor Administrativo. Revista de Tecnologia da Informação e Comunicação da Faculdade Estácio do Pará, [S.l.], v. 2, n. 3, p. 1-14, 2019.

- CORONEL, C.; MORRIS, S. Database systems: design, implementation, & management. Cengage Learning, 2016.
- CRESWELL, J. W. Projeto de pesquisa: métodos qualitativo, quantitativo e misto. Tradução Luciana de Oliveira da Rocha. 2ed. Porto Alegre: Artmed, 2007.
- DE SOUZA BARRETO, I. M.; & FREITAS, A. E. S. Generating Intelligence Through Microdates: A Business Intelligence Proposal For The Education Area Of The Bahia Federal Institute. *Cadernos de Educação Tecnologia e Sociedade*, 13(4), 463-473, 2020.
- ELMASRI, R.; NAVATHE, S. B. Sistemas de Banco de dados. São Paulo: Pearson Education Boston, MA, 2011.
- GEOPROFILE. Disponível em: <<http://www.dpi.ufv.br/projetos/geoprofile/download.html>>. Acesso em: 15 jun. 2020.
- GOOGLE: "Geolocalização de coordenadas geográficas". Disponível em: <https://developers.google.com/maps/documentation/geocoding/overview>. Acesso em: 19 mai. 2020.
- INMON, W. B. Como construir o Data Warehouse. 2. ed. Rio de Janeiro: Campus, 1997.
- LISBOA-FILHO, J.; SAMPAIO, G. B.; NALON, F. R.; & Borges, K. A. D. V. "A UML profile for conceptual modeling in GIS domain." *DE@ CAISE* (2010): 18-31.
- LAUDON, K. C.; LAUDON, J. P. Sistemas de informação com internet. 4. ed. Rio de Janeiro: LTC, 1999.
- KIMBALL, R. Data Warehouse Toolkit: Técnicas para Construção de Data Warehouse Dimensionais. São Paulo: Makron Books, 1998.
- _____.; CASERTA, J. The data warehouse ETL toolkit. John Wiley & Sons, 2004.
- MACHADO, F. Tecnologia e Projeto de Data Warehouse: Uma Visão multidimensional. São Paulo: Érica, 2007.

- MEDINA, F.; FARIÑA, F.; & CASTILHO-ROJAS, W. Data mart para obtención de indicadores de productividad académica en una universidad. *Ingeniare. Revista chilena de ingeniería*, Universidad de Tarapacá, v. 26, p. 88–101, 2018.
- NOMINATIM. “Nominatim com Dados do OpenStreetMap”. Disponível em: <https://operations.osmfoundation.org/policies/nominatim/>>. Acesso em 20 jun. 2020.
- OPENSTREETMAP. “Mapeamento Colaborativo”. Disponível em: <https://www.openstreetmap.org>. Acesso em: 20 de jun. 2020.
- OPENROUTESERVICE. “Distância por meio de Coordenadas Geográficas”. Disponível em: <https://openrouteservice.org/> Acesso em: 10 jun. 2020.
- PRIMAK, F. V. *Decisões com BI (Business Intelligence)*. Rio de Janeiro: Ciência Moderna, 2008.
- ROB, P.; CORONEL, C. *Sistemas de Banco de Dados: Projeto, Implementação e Administração*. São Paulo: Cengage do Brasil, 2010.
- ROBBINS, S.; JUDGE, T.; SOBRAL, F. *Comportamento organizacional: teoria e prática no contexto brasileiro*. 14. ed. São Paulo: Pearson, 2010.
- SANTOS, W.; OLIVEIRA, D. de. Um ambiente de apoio à decisão baseado em data warehouse para a área de segurança pública do estado do rio de janeiro. In: SBC. 1º Workshop Brasileiro de Cidades Inteligentes (WBCI 2018).
- SECCHI, L. Modelos organizacionais e reformas da administração pública. *Revista de Administração Pública – RAP*. Rio de Janeiro:FGV, 43(2):347-69, 2009.
- SICELO. Disponível em <https://www.scielo.org/>. Acesso em 15 de novembro 2020.
- SILVA, W. R. da.; & JARDINI, E. de A. Potencialização Da Análise Dos Motivos De Evasão Dos Alunos Do Instituto Federal De São Paulo - Campus Votuporanga Por Meio De Business Intelligence. *UNIFUNEC CIENTÍFICA MULTIDISCIPLINAR*, 8(10), 1–13, 2019.
- SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. Disponível em <https://www.sbc.org.br/>. Acesso em 10 de novembro 2020.
- SOUZA, E. S. GITHUB – página do projeto. Disponível em: <<https://github.com/edivaldossouza/desenvolvimento-dataMart-bi-pentaho-ifes>>. Acesso em: 04/08/2021.

SUÍTE PENTAHO. Disponível em: <https://sourceforge.net/projects/pentaho/>. Acesso em: 10 jun. 2020.

TAKASHINA, N.; FLORES, M. Indicadores da qualidade e do desempenho: como estabelecer metas e medir resultados. Rio de Janeiro: Qualitymark, 1996.

THOMPSON, O. Business intelligence success, lessons learned. Technology, 2004.

TURBAN, E. Tecnologia da Informação para Gestão. Efraim Turban, Ephraim MClean e James Wetherbe; Tradução Renate Schinke. 3.ed. Porto Alegre: Bookman, 2004.

_____.; ARONSON, J. E. Decision Support Systems and Intelligent Systems. 5. ed. USA: Prentice Hall, 1998.

_____.; SHARDA, R.; ARONSON, J. E.; & KING, D. Business Intelligence: um enfoque gerencial para a inteligência do negócio. Porto Alegre: Bookman Editora, 2009.

APÊNDICE A – Obtenção de dados geográficos usando API OSM

Nesta seção será discutido o desenvolvimento do processo ETL para obtenção das coordenadas geográficas dos servidores aposentados através do OSM, conforme foi apresentado na Figura 12. As Figuras 28 a 35, apresentam a descrição dos principais Steps da Figura 12: *Table Input*, *URL_Maps*, *HTTP Client*, *Coordenadas*, *Modified Java Script Value 2*, *Select values 2* e *Insert / Update*. A Figura 28 representa o Step *Table Input*, onde é realizada a conexão com a tabela DIM_Endereco no DW e logo depois é efetuada uma consulta SQL dos servidores aposentados.

Figura 28 - Step *Table Input* conectado tabela Servidor

Letura de Tabela

Nome do Step: Table input

Connection: Pesquisa

SQL

```
SELECT e.matriculasiape, e.endlogradouro as logradouro, e.endnumero as numero,
e.endmunicipio as cidade, e.enduf as uf
FROM public.dim_endereco as e
inner join dim_servidor as df on df.matriculasiape = e.matriculasiape
where df.orgaogrupoCodigo = 5
```

Linha 2 Coluna 0

Enable lazy conversion

Replace variables in script?

Insert data from step

Executar para cada linha?

Tamanho limite: 0

Help OK Preview Cancela

Fonte: Autor

A construção da URL de requisição dos dados Geográficos ao OSM é montada no Step *URL_Maps*, conforme ilustrado na Figura 29.

Figura 29 - Step *Url_Maps* para criação da url de requisição ao OSM

Script Values / Mod

Step name: URL Maps

Java script functions:

- Transform Scripts
- Transform Constants
- Transform Functions
- Input fields
 - matriculasiape
 - logradouro
 - numero
 - cidade
 - uf
- Output fields
 - Please use the 'Repla'

Java script:

```

Script 1
//Script here

if(numero == 'S/N' || numero == 's/n'){
    var url_final = 'https://nominatim.openstreetmap.org/search?street='
    +encodeURIComponent(logradouro)+'%2C'+encodeURIComponent(cidade)+'&state='+uf
    +'&format=xml&polygon=0&addressdetails=0&limit=1';
}
else if(numero != 'S/N'){
    var url_final = 'https://nominatim.openstreetmap.org/search?street='
    +encodeURIComponent(logradouro)+'%2C'+encodeURIComponent(numero)+'&city='
    +encodeURIComponent(cidade)
    +'&state='+uf+'&format=xml&polygon=0&addressdetails=0&limit=1';
}
  
```

Position: 1, 0

Compatibility mode? Optimization level: g

Fields

#	Fieldname	Rename to	Type	Length	Precision	Replace value 'Fieldname' or 'Rename to'
1	url_final		String			N
2						

Buttons: Help, OK, Cancela, Get variables, Test script

Fonte: Autor

O Step *HTTP Client* é apresentado na Figura 30. Nas Figuras 31 e 32 temos o Step Coordenadas, onde são feitas as configurações das respostas recebidas da API.

Figura 30 - Step *HTTP Client* realizando a requisição ao OSM

HTTP Client

Step name: Consulta MAPS

General Fields

Settings

URL

Accept URL from field?

URL field name: url_final

Encoding (empty means standard): UTF-8

Connection timeout: 1000000

Socket timeout: 1000000

Connection close wait time: -1

Output fields

Result field name: result_1

HTTP status code field name

Response time (milliseconds) field name

Response header field name

HTTP authentication

Http Login

HTTP Password

Proxy to use

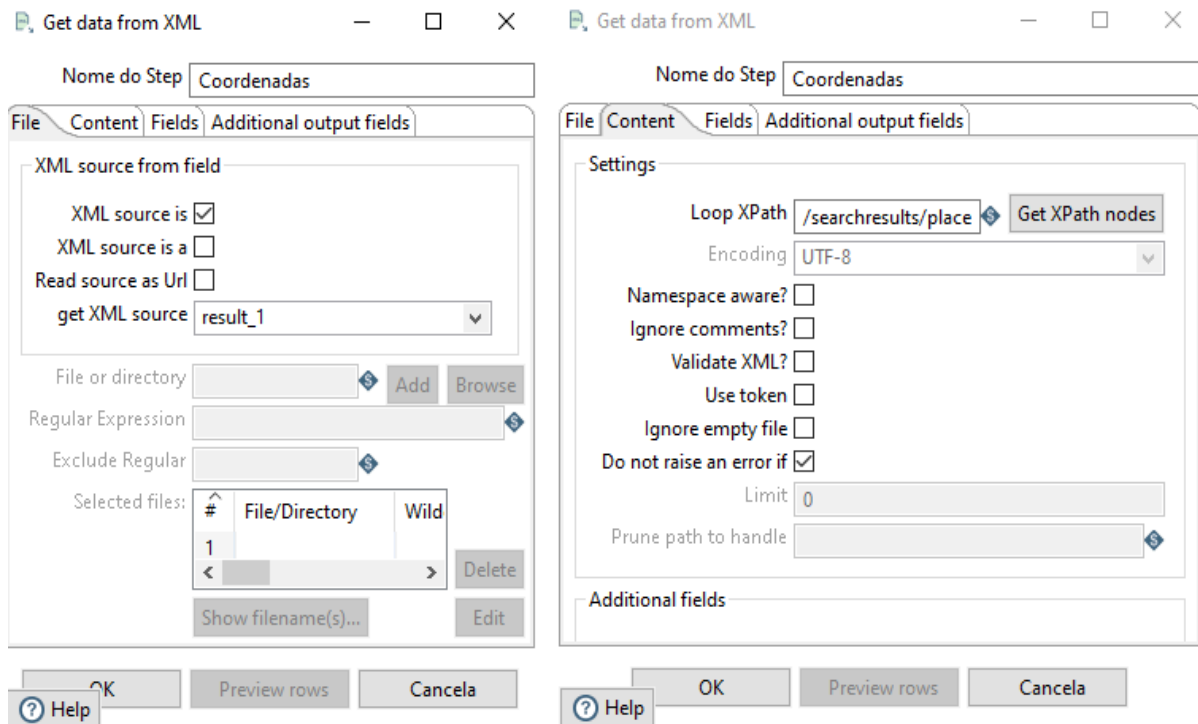
Proxy Host

Proxy Port

Help OK Cancela

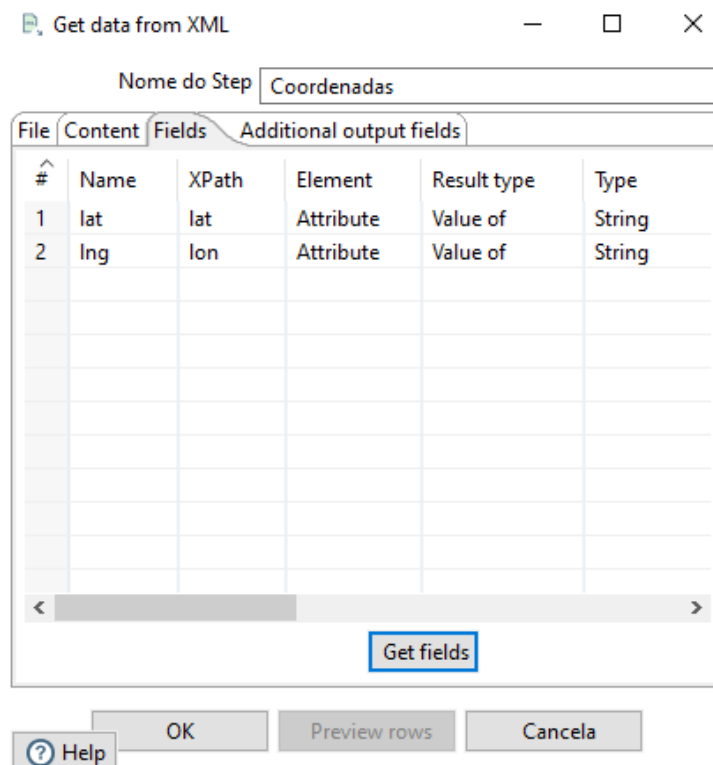
Fonte: Autor

Figura 31 - Step *Coordenadas* configurações das abas “File” e “Content” em resposta da requisição do OSM



Fonte: Autor

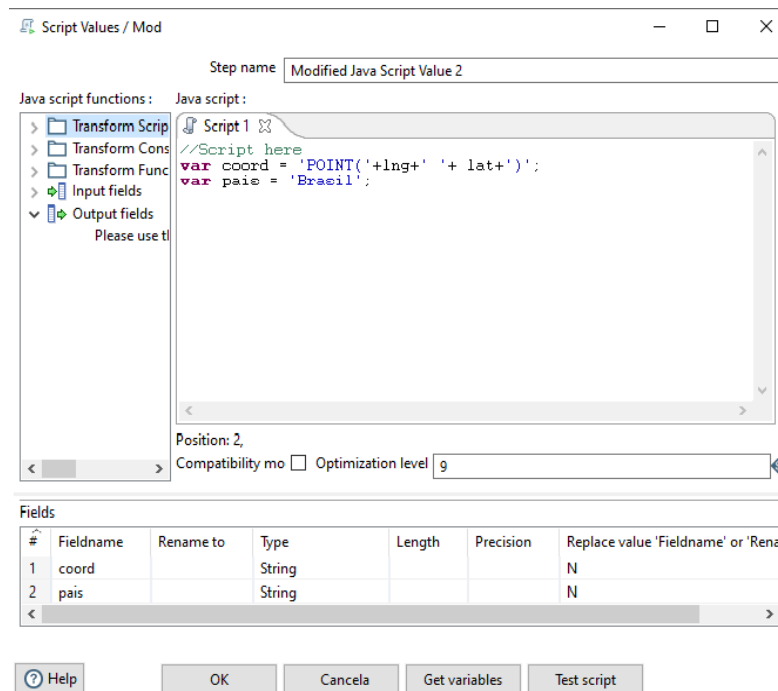
Figura 32 - Step *Coordenadas* configuração da aba “Fields” para selecionar os campos desejados da resposta do OSM



Fonte: Autor

Em seguida, os dados passam pelo Step *Modified Java Script Value 2*, Figura 33, onde é montando uma string “coord” com a junção dos campos “lat” e “lng” respeitando o formato pontual geográfico.

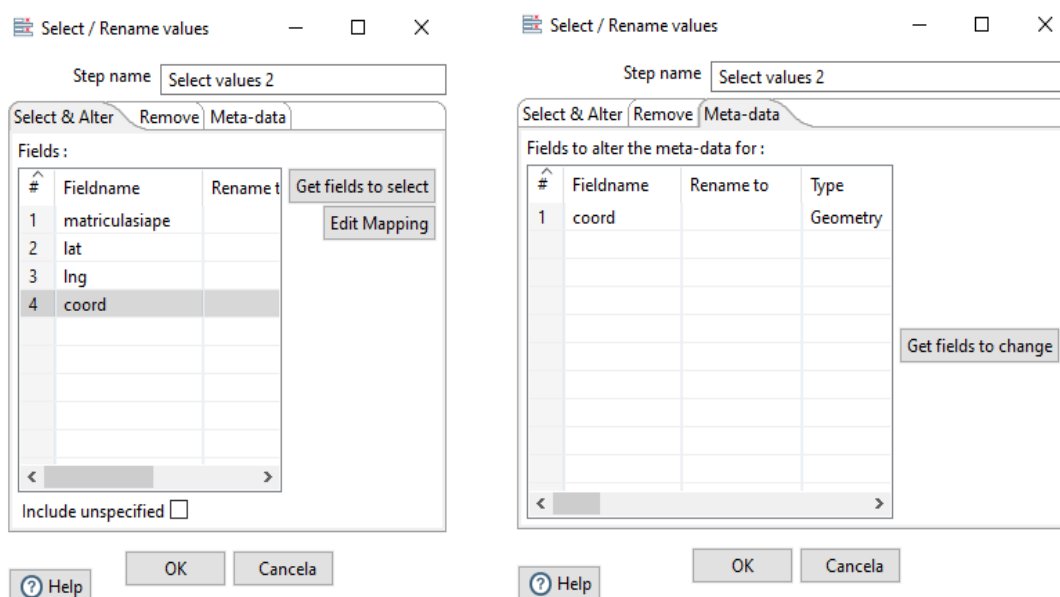
Figura 33 - Configuração Step *Modified Java Script Value 2*



Fonte: Autor

No próximo Step *Select values 2*, Figura 34, temos a conversão do campo “coord” do tipo String para Geometry. E por fim na figura 35 temos o Step “*Insert / Update*”, onde é configurado a carga na tabela “DIM_Endereco”.

Figura 34 - Conversão do campo “coord” do tipo String para Geometry



Fonte: Autor

Figura 35 - Realização da carga na tabela DIM_Endereco com campos geográficos

Insert / Update

Step name: Insert / Update

Connection: [] Edit... New... Wizard...

Target schema: public Navega...

Target table: dim_endereco Browse...

Commit size: 100

Don't perform any

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

#	Table field	Comparator	Stream field1	St
1	matriculasiape	=	matriculasiape	

Get fields

Update fields:

#	Table field	Stream field	Update
1	lat	lat	Y
2	lon	lng	Y
3	coordenadas	coord	Y

Get update fields

Edit mapping

Help OK Cancela SQL

Fonte: Autor