



DESENVOLVIMENTO DE BUSINESS INTELLIGENCE NA AVALIAÇÃO DE DESEMPENHO EM TRIBUNAIS DE CONTAS BRASILEIROS

Saulo Coelho Lima

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Processos – Mestrado Profissional, PPGEP/ITEC, da Universidade Federal do Pará, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Processos.

Orientador: Kleber Bittencourt Oliveira

Belém

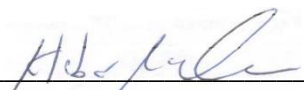
Novembro de 2021

**DESENVOLVIMENTO DE BUSINESS INTELLIGENCE NA AVALIAÇÃO DE
DESEMPENHO EM TRIBUNAIS DE CONTAS BRASILEIROS**

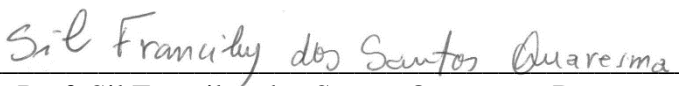
Saulo Coelho Lima

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE
PÓSGRADUAÇÃO EM ENGENHARIA PROCESSOS – MESTRADO
PROFISSIONAL (PPGEP/ITEC) DA UNIVERSIDADE FEDERAL DO PARÁ COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE
MESTRE EM ENGENHARIA DE PROCESSOS.

Examinada por:



Prof. Kleber Bittencourt Oliveira, Dr.
(PPGEP/ITEC/UFPA-Orientador)



Prof. Sil Franciley dos Santos Quaresma, Dr.
(PPGEP/ITEC/UFPA-Membro)

Prof. Jardel Pinto Barbosa, Dr.
(UEAP-Membro)

BELÉM, PA - BRASIL

NOVEMBRO DE 2021

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA

Lima, Saulo Coelho, 1981-
Desenvolvimento de Business Intelligence na avaliação de
desempenho em tribunais de contas brasileiros / Saulo Coelho
Lima - 2021.

Orientador: Kleber Bittencourt Oliveira

Dissertação (Mestrado Profissional) - Universidade
Federal do Pará. Instituto de Tecnologia. Programa de Pós-
Graduação em Engenharia de Processos, Belém, 2021

1. Tribunais de contas-Avaliação 2. Desempenho-Avaliação
3. Inteligência competitiva (Administração) I. Título

CDD 23. ed. - 658.1

*Dedico este trabalho a minha família pelo
apoio e motivação durante a jornada.*

AGRADECIMENTOS

O meu agradecimento a minha esposa Thábitta Leão Corrêa Lima, minhas filhas Sofia Leão Corrêa Lima e Elis Leão Corrêa Lima, minha irmã Suelen Coelho Lima, meus pais Raimundo dos Santos Lima e Suely Coelho Lima (*in memoriam*) pelo apoio ao longo da minha vida e pela compreensão nas noites que estive ausente para alcançar mais um objetivo na minha vida. Ao meu orientador Prof. Dr. Kleber Bittencourt Oliveira, pelo apoio, suporte e orientação ao longo do processo de pesquisa, elaboração e conclusão desta dissertação.

Agradeço a todos os Professores, e demais integrantes, do Programa de Pós-Graduação em Engenharia de Processos (PPGEP) que estiveram presentes nesse processo do convênio entre o Instituto de Tecnologia e Educação Galileo da Amazônia (ITEGAM) e a Universidade Federal do Pará (UFPA), comprometidos com a qualificação da Sociedade da Região Norte, em especial a Cidade de Manaus. E também aos Professores Tereza Rodrigues Felipe Cabral e Jandecy Cabral Leite, que me incentivaram e compreenderam as diversas dificuldades durante esse período tão delicado no qual todo o Brasil esteve inserido, a pandemia da COVID-19.

*“Sucesso não é o final; falhar não é fatal: é
a coragem de continuar que conta.”*

(Winston Churchill)

Resumo da Dissertação apresentada ao PPGEP/UFPA como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia de Processos (M. Eng.)

DESENVOLVIMENTO DE BUSINESS INTELLIGENCE NA AVALIAÇÃO DE DESEMPENHO EM TRIBUNAIS DE CONTAS BRASILEIROS

Saulo Coelho Lima

Novembro/2021

Orientador: Kleber Bittencourt Oliveira

Área de Concentração: Engenharia de Processos

Em um contexto de aumento de demanda por serviços públicos, escassez de recursos e aumento da complexidade do controle desses recursos públicos dentro de um quadro que considera a evolução da administração pública, com a incorporação de conceitos de governança, planejamento e de maior engajamento social, é imperativo a busca por aprimoramento das instituições mediante a percepção de suas ações. As avaliações de desempenho são instrumentos de gestão essenciais que permitem acompanhar o alcance das metas, identificar avanços, melhorias de qualidade, correção de problemas e a necessidades de realização de mudanças. Tendo como objetivo a proposta de otimização no processo de avaliação de indicadores que estabelece a medição de desempenho, qualidade e agilidade nas atividades dos Tribunais de Contas (estabelecido pela ATRICON), aplicando uma ferramenta de *Business Intelligence* que auxilie na consolidação da medição de indicadores e permita o registro e acompanhamento da evolução do desempenho das instituições submetidas à avaliação do Marco de Medição de Desempenho (MMD-QATC). Para realizar tal tarefa utilizou-se o sistema de gerenciamento de banco de dados MySQL em conjunto com o HeidiSQL, a ferramenta de modelagem SQL Power Architect, a linguagem de programação Python, o software Pentaho Data Integration da suíte para construção de soluções de BI Pentaho na sua versão *Community Edition* e o Microsoft Power BI Desktop. Resultando assim, em um plataforma mais prática e escalável de consolidação dos dados de avaliação.

Abstract of Dissertation presented to PPGE/UFPA as a partial fulfillment of the requirements for the degree of Master in Process Engineering (M. Eng.)

**DEVELOPMENT OF A BUSINESS INTELLIGENCE SOLUTION FOR
PERFORMANCE EVALUATION IN BRAZILIAN COURTS OF AUDITORS**

Saulo Coelho Lima

November/2021

Advisor: Kleber Bittencourt Oliveira

Research Area: Process Engineering

In a context of raised demand for public services, shortage of resources and increasing complexity of controlling these public resources within a perspective that considers the evolution of public administration, with the incorporation of concepts of governance, planning and social engagement, it is imperative to pursuit improvement of ours institutions through the perception of their actions. Performance evaluations are essential management tools that allow us to monitor the achievement of goals, identify advances, correct problems and the need to make changes. Aiming at the optimization proposal in the indicators evaluation process that establishes the measurement of performance, quality and agility of Court of Auditors' activities (established by ATRICON), applying a Business Intelligence tool that helps in the consolidation of the measurement of indicators and allows the registration and monitoring of the evolution of institution's performance submitted to evaluation based on the framework Performance Measurement Milestone (MMD-QATC). To accomplish this task, the MySQL database management system was used in conjunction with HeidiSQL, the SQL Power Architect modeling tool, the Python programming language, the Pentaho Data Integration, part of the suite for BI solutions Pentaho in its Community Edition version, and the Microsoft Power BI Desktop. Thus, resulting in a more practical and scalable platform for consolidating the data.

SUMÁRIO

CAPÍTULO 1 - INTRODUÇÃO.....	1
1.1 - MOTIVAÇÃO.....	1
1.2 - OBJETIVOS.....	4
1.2.1 - Objetivo geral.....	4
1.2.2 - Objetivos específicos.....	4
1.3 - CONTRIBUIÇÕES DA DISSERTAÇÃO.....	4
1.4 - ORGANIZAÇÃO DO TRABALHO.....	5
CAPÍTULO 2 - REVISÃO DA LITERATURA.....	7
2.1 - MEDIÇÃO DE DESEMPENHO NA ADMINISTRAÇÃO PÚBLICA.....	7
2.2 - MARCO DE MEDIÇÃO DO DESEMPENHO – QUALIDADE E AGILIDADE DOS TRIBUNAIS DE CONTAS DO BRASIL (MMD-QATC).....	9
2.2.1 - Origem e evolução.....	9
2.2.1.1 - Entidade Fiscalizadora Superior.....	11
2.2.2 - Metodologia e estrutura do MMD-TC.....	12
2.2.2.1 - Domínio.....	13
2.2.2.2 - Indicador.....	14
2.2.2.3 - Dimensão.....	14
2.2.2.4 - Critério.....	14
2.2.2.5 - Nível de desempenho.....	14
2.3 - <i>BUSINESS INTELLIGENCE</i>	15
2.4 - SISTEMAS DE GERENCIAMENTO DE BANCO DE DADOS.....	22
2.4.1 - Modelo conceitual, modelo lógico e projeto físico.....	23
2.4.2 - Linguagem SQL.....	24
2.5 - TECNOLOGIAS E FERRAMENTAS.....	24
2.5.1 - HeidiSQL.....	25
2.5.2 - MySQL.....	25
2.5.3 - SQL Power Architect.....	26
2.5.4 - Linguagem de programação Python.....	26
2.5.5 - Biblioteca Pandas.....	26
2.5.6 - Suíte Pentaho.....	26
2.5.6.1 - Pentaho Data Integration.....	27

2.5.7 - Power BI Desktop.....	27
CAPÍTULO 3 - MATERIAIS E MÉTODOS.....	28
3.1 - TECNOLOGIAS E RECURSOS UTILIZADOS.....	28
3.1.1 - Infraestrutura de <i>Hardware</i>.....	28
3.1.2 - Infraestrutura de <i>Software</i>.....	30
3.1.2.1 - MySQL.....	30
3.1.2.2 - HeidiSQL.....	31
3.1.2.3 - Linguagem SQL.....	32
3.1.2.4 - SQL Power Architect.....	33
3.1.2.5 - Linguagem de programação Python.....	34
3.1.2.6 - Biblioteca Pandas.....	34
3.1.2.7 - Pentaho Data Integration.....	34
3.1.2.8 - Power BI Desktop.....	35
3.2 - METODOLOGIA EXPERIMENTAL.....	36
3.2.1 - Identificar tecnologias a serem empregadas.....	36
3.2.2 - Identificar origem e descrever estrutura dos dados.....	36
3.2.3 - Identificar e relacionar Fatos e Dimensões.....	37
3.2.4 - Gerar Massa de Dados de Teste.....	38
3.2.5 - Criar <i>data warehouse</i>.....	38
3.2.6 - Criar ETL <i>data warehouse</i>.....	40
3.2.7 - Criar <i>datamart</i>.....	41
3.2.8 - Criar ETL <i>datamart</i>.....	42
3.2.9 - Criar análises e <i>dashboards</i>.....	44
CAPÍTULO 4 - RESULTADOS E DISCUSSÃO.....	46
4.1 - <i>DASHBOARD</i> “UNIDADES, INDICADORES E DIMENSÕES”.....	51
4.2 - <i>DASHBOARD</i> “ENTREGAS”.....	51
4.3 - <i>DASHBOARD</i> “INDICADORES POR UNIDADE”.....	52
4.4 - <i>DASHBOARD</i> “EVOLUÇÃO DOS INDICADORES POR UNIDADE”.....	53
4.5 - <i>DASHBOARD</i> “INDICADORES POR REGIÃO”.....	54
4.6 - <i>DASHBOARD</i> “DESEMPENHO DE UNIDADES POR INDICADOR”.....	55
4.7 - <i>DASHBOARD</i> “UNIDADES POR NÍVEL DE DESENVOLVIMENTO”.....	56
4.8 - <i>DASHBOARD</i> “INDICADORES/DIMENSÕES”.....	57
4.9 - <i>DASHBOARD</i> “INDICADORES/DIMENSÕES - GERAL”.....	58

CAPÍTULO 5 - CONCLUSÕES E SUGESTÕES.....	60
5.1 - CONCLUSÕES.....	60
5.2 - SUGESTÕES.....	61
REFERÊNCIAS BIBLIOGRÁFICAS.....	62
APÊNDICE A - DOMÍNIOS E INDICADORES NO MMD-QATC.....	66
APÊNDICE B - SCRIPT SQL DE CRIAÇÃO DO <i>DATA WAREHOUSE</i>.....	68
APÊNDICE C - SCRIPT SQL DE CRIAÇÃO DO <i>DATAMART</i>.....	71
APÊNDICE D - TRECHO DO CÓDIGO FONTE EM PYTHON DO PROCESSO DE ETL DO <i>DATA WAREHOUSE</i>.....	75

LISTA DE FIGURAS

Figura 1.1	Pilares de atuação do TC e atores institucionais e internacionais..	2
Figura 1.2	Estrutura analítica da dissertação.....	5
Figura 2.1	Estrutura do MMD-TC.....	13
Figura 2.2	Escala dos níveis de desempenho.....	15
Figura 2.3	Dado, informação, conhecimento e decisões.....	16
Figura 2.4	Ferramentas, técnicas e origem de dados.....	18
Figura 2.5	Produto, processo, solução e ferramentas de BI.....	19
Figura 2.6	Arquitetura geral de um sistema de BI tradicional.....	22
Figura 2.7	Exemplo de modelo conceitual.....	23
Figura 3.1	Fluxograma das etapas do trabalho de pesquisa.....	28
Figura 3.2	Infraestrutura de hardware.....	29
Figura 3.3	Página de <i>download</i> do MySQL.....	30
Figura 3.4	Página de <i>download</i> do HeidiSQL.....	31
Figura 3.5	Configuração de conexão ao SGBD do HeidiSQL.....	32
Figura 3.6	Interface inicial do HeidiSQL.....	32
Figura 3.7	Página de <i>download</i> do Power Architect.....	33
Figura 3.8	Página de <i>download</i> da Suíte Pentaho.....	35
Figura 3.9	Interface inicial do Pentaho Data Integration.....	35
Figura 3.10	Diagrama de processo de avaliação QATC.....	37
Figura 3.11	Modelo de dados do DW.....	39
Figura 3.12	<i>Script</i> SQL de criação do DW sendo executado no HeidiSQL.....	39
Figura 3.13	Base de dados resultante da execução do <i>script</i>	40
Figura 3.14	Execução do algoritmo Python que realiza o ETL do DW.....	40
Figura 3.15	Modelo de dados do <i>datamart</i>	41
Figura 3.16	<i>Script</i> SQL de criação do <i>datamart</i> sendo executado no HeidiSQL.....	42
Figura 3.17	Base de dados resultante da execução do <i>script</i> de criação do <i>datamart</i>	42
Figura 3.18	Exemplo de processo de carga mapeado no PDI.....	43

Figura 3.19	Detalhe da operação que grava dados da dimensão unidade no <i>datamart</i>	43
Figura 3.20	Execução da rotina de carga das tabelas de fato do <i>datamart</i>	44
Figura 3.21	Criação da medida “Média Nota Final por Indicador”.....	45
Figura 4.1	Diagrama da arquitetura da solução proposta	47
Figura 4.2	Fluxograma do algoritmo em Python de ETL do DW.....	48
Figura 4.3	Processos de carga modelados no Pentaho Data Integration.....	49
Figura 4.4	Modelagem dos processos de carga das dimensões em geral.....	49
Figura 4.5	Modelagem do processo de carga da dimensão de tempo.....	50
Figura 4.6	Modelagem do processo de carga dos fatos.....	50
Figura 4.7	<i>Dashboard</i> de “Unidades, Indicadores e Dimensões”.....	51
Figura 4.8	<i>Dashboard</i> de “Entregas”	52
Figura 4.9	<i>Dashboard</i> de “Indicadores por Unidade”	53
Figura 4.10	<i>Dashboard</i> de “Evolução dos Indicadores por Unidade”	54
Figura 4.11	<i>Dashboard</i> de “Indicadores por Região”.....	55
Figura 4.12	<i>Dashboard</i> de “Desempenho de Unidades por Indicador”	56
Figura 4.13	<i>Dashboard</i> de “Unidades por Nível de Desenvolvimento”.....	57
Figura 4.14	<i>Dashboard</i> de “Indicadores/Dimensões Geral”	58
Figura 4.15	<i>Dashboard</i> de “Indicadores/Dimensões por Unidade”	59

LISTA DE TABELAS

Tabela 2.1	Resoluções que serviram de base para o MMD-QATC.....	10
Tabela 3.1	Relação Fato x Dimensões.....	38
Tabela A.1	Lista de domínios do MMD-QATC, com seus respectivos indicadores	66

NOMENCLATURA

ATRICON	ASSOCIAÇÃO DOS MEMBROS DOS TRIBUNAIS DE CONTAS DO BRASIL
CCOR	COLÉGIO DOS CORREGEDORES E OUVIDORES DOS TRIBUNAIS DE CONTAS DO BRASIL
MMD - QATC	MARCO DE MEDIÇÃO DO DESEMPENHO - QUALIDADE E AGILIDADE DOS TRIBUNAIS DE CONTAS DO BRASIL
CA	COMISSÃO DE AVALIAÇÃO
CCQ	COMISSÃO DE CONTROLE DE QUALIDADE
CGQ	COMISSÃO DE GARANTIA DE QUALIDADE
OCDE	ORGANIZAÇÃO PARA COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO
EFS	ENTIDADE FISCALIZADORA SUPERIOR
INTOSAI	ORGANIZAÇÃO INTERNACIONAL DAS ENTIDADES FISCALIZADORAS SUPERIORES
NBASP	NORMAS BRASILEIRAS DE AUDITORIA DO SETOR PÚBLICO
ISSAI	NORMAS INTERNACIONAIS DAS ENTIDADES FISCALIZADORAS SUPERIORES
IRB	INSTITUTO RUI BARBOSA
OLACEFS	ORGANIZAÇÃO LATINO-AMERICANA E DO CARIBE DE ENTIDADES FISCALIZADORAS SUPERIORES

CAPÍTULO 1

INTRODUÇÃO

1.1 - MOTIVAÇÃO

As sociedades têm demonstrado interesse e necessidade em conhecer e controlar os gastos públicos ao longo da sua existência. Índícios históricos, que remontam à papiros do Faraó Menés (3.000 a.C.), passando pelos *hellenotomiai* (tesoureiros da Deusa Atenas, responsáveis por analisar e julgar as contas dos gestores públicos na Grécia Antiga), e ainda pela instituição do Erário Régio, por parte de Portugal, já no Consulado do Marquês de Pombal- considerado por muitos como célula de origem dos Tribunais de Contas (TCs) - e diversos outros acontecimentos (GARCIA, 2015) apontam para esforços concretos em criar mecanismos que permitam um real controle sobre as ações pecuniárias dos gestores públicos.

No Brasil, a apresentação de um projeto ao Senado do Império (em 1826), inaugurou as discussões entre os que apoiavam a criação de órgão independente de fiscalização e os que defendiam que dever-se-ia manter sob responsabilidade dos próprios gestores públicos a responsabilidade pelo controle das suas próprias despesas

A instituição formal do primeiro Tribunal de Contas no Brasil, o da União, ocorreu por iniciativa de Rui Barbosa (Ministro da Fazenda, na época) por intermédio do Decreto no. 966-A de 07 de Novembro de 1890. As Constituições de: 24 de Fevereiro de 1821; 16 de Julho de 1934; 18 de Setembro de 1946; e 5 de outubro de 1988 estabeleceram as competências e deram forma legal e funcional aos Tribunais de Contas como os conhecemos hoje.

Na Figura 1.1, BORALI (2018), apresenta o que considera pilares de fortalecimento institucional dos Tribunais de Contas, que elevaram o seu status, ampliaram seu âmbito de atuação e passaram a exigir maior abertura, publicidade e accountability.

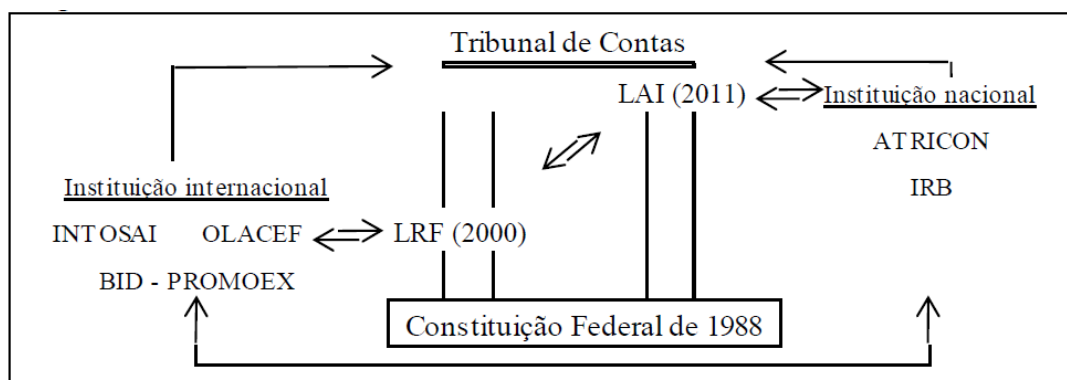


Figura 1.1 - Pilares de atuação do TC e atores institucionais e internacionais.

Fonte: BORALI (2018).

LUZ (2018), faz um apanhado histórico da evolução dos TCs e, não apenas isso, discute também fatores que contribuíram para o que chamou de ressignificação dos TCs, citando ações como as manifestações de rua de 2013, a Operação Lava-Jato, o julgamento das contas da Presidência da República em 2015 - com a recomendação de reprovação das contas analisadas pelo TCU - e o aumento da complexidade do controle dos recursos públicos dentro de um contexto que considere a evolução da administração pública, com a incorporação de conceitos de governança, planejamento e de maior engajamento social. Junta-se a isso ainda, elementos como governo digital e controle digital.

Por meio de ações individuais, mas também por iniciativas conjuntas, muitas vezes motivadas e encabeçadas pelo Instituto Rui Barbosa e pela Associação dos Membros dos Tribunais de Contas do Brasil - ATRICON, os TCs têm buscado a organização, a padronização e, por conseguinte, a melhoria de seus serviços visando não apenas o cumprimento de seu mister constitucional, mas também atender ao anseio popular por uma instituição mais moderna e eficiente.

A ATRICON, criada em 16 de Agosto de 1992, atua com o “intuito de garantir a representação, a defesa, o aperfeiçoamento e a integração dos Tribunais de Contas e de seus Membros (Ministros, Conselheiros, Ministros Substitutos e Conselheiros Substitutos), visando aprimorar o Sistema de Controle Externo do Brasil em benefício da sociedade” (ATRICON, 2020).

Dentre suas iniciativas de apoio ao desenvolvimento de estratégias e ferramentas para aprimorar a eficácia dos sistemas de controle da Administração Pública, em 02 de Julho de 2013, a ATRICON estabeleceu critérios por meio do regulamento ATRICON Nº 01/2013 - que “estabelece as regras e os parâmetros de qualidade e agilidade do

controle externo, para fins de avaliação dos Tribunais de Contas, e dá outras providências” - a serem aplicados na avaliação dos Tribunais de Contas.

E em 06 de Agosto de 2014, decidiu regulamentar diretrizes orientativas para os Tribunais de Contas. Surgiram, a partir deste momento, as resoluções elaboradas diretamente pela ATRICON e as elaboradas em conjunto com o Colégio dos Corregedores e Ouvidores dos Tribunais de Contas do Brasil (CCOR).

Cada uma dessas resoluções apresenta as justificativas para elaboração do documento, os objetivos traçados, os compromissos firmados pelos membros, os princípios e fundamentos legais que embasaram a sua elaboração, além dos conceitos adotados e as diretrizes apontadas.

As resoluções listadas acima e o regulamento ATRICON Nº 01/2013 culminaram no “Marco de Medição do Desempenho – Qualidade e Agilidade dos Tribunais de Contas do Brasil” (MMD-QATC), que incorpora as diretrizes da ATRICON e está fortemente baseado em normas internacionais aceitas pelas Entidades Fiscalizadoras Superiores.

BORALI (2018) aponta o MMD-QATC entre os trabalhos que estimulam a padronização das instituições de controle (junto com outras iniciativas do IRB, OLACEFS e INTOSAI), reforçando a importância e relevância da iniciativa. No entanto, salienta que a incapacidade de correlacionar os indicadores de qualidade individuais de cada Tribunal de Contas com seus resultados individuais, nos resultados divulgados pela Atricon, pode comprometer a sua representatividade como um efetivo marco de qualidade institucional.

No entanto, ATRICON (2014) deixa claro que o QATC seria um instrumento apropriado para a comparação do desempenho ao longo do tempo, dentro do próprio Tribunal avaliado e que não objetiva fazer qualquer comparação ou ranqueamento dos Tribunais.

Entendemos que a comparação entre as diferentes unidades pode ser interessante e salutar, e que possa atuar como um mecanismo que lance luz sobre os dados e traga informações que permitam ver oportunidades de melhoria. De tal modo, permitindo uma relação ganha-ganha entre os membros.

Os estudos e trabalhos relacionados a este tema mostram-se extremamente relevantes pelo exposto acima e pela abrangência da aplicação do MMD-QATC: aproximadamente 85% dos Tribunais de Contas dos Estados, Distrito Federal e Municípios, em um primeiro momento, e a totalidade deles posteriormente (a partir de 2017), comprometeram-se a se submeter à esta metodologia de avaliação.

Acredito ser de grande valia um estudo/instrumental que auxilie na árdua tarefa de medir, registrar e acompanhar a evolução do desempenho de Instituições Públicas com tamanha relevância, configurando assim grande vantagem e trazendo grandes benefícios não apenas aos agentes responsáveis por aferir, controlar e reportar tais dados, mas à toda sociedade.

Junto a tudo isso, a pesquisa motivou-se pela necessidade de propor melhoria visando a otimização do processo de avaliação dos indicadores estabelecidos pelo Marco de medição de desempenho, qualidade e agilidade definido sobre metodologia desenvolvida pela ATRICON para Tribunais de Contas do Brasil.

1.2 - OBJETIVOS

1.2.1 - Objetivo geral

O desenvolvimento de uma solução de *Business Intelligence* que otimize e auxilie na consolidação e avaliação de desempenho em Tribunais de Contas Brasileiros com base nos critérios definidos pela ATRICON no seu Marco de Medição de Desempenho – Qualidade e Agilidade dos Tribunais de Contas (MMD-QATC), permitindo acompanhamento evolutivo dessas instituições.

1.2.2 - Objetivos específicos

- Identificar técnicas e tecnologias empregadas para desenvolvimento da solução almejada;
- Identificar a origem e descrever estrutura dos dados;
- Elaborar matriz fato x dimensão;
- Modelar e criar estruturas de dados; e
- Criar análises e *dashboards*.

1.3 - CONTRIBUIÇÕES DA DISSERTAÇÃO

Vivemos um momento de aumento de demanda por serviços públicos, escassez de recursos e aumento da complexidade do controle desses recursos públicos dentro de

um contexto que considera a evolução da administração pública, com a incorporação de conceitos de governança, planejamento e de maior engajamento social, que torna imperativo a busca do aprimoramento das instituições mediante a percepção e avaliação de suas ações.

As avaliações de desempenho são instrumentos de gestão essenciais que permitem acompanhar o alcance das metas, identificar avanços, melhorias de qualidade, correção de problemas e a necessidades de realização de mudanças.

Essa pesquisa contribuiu na otimização do processo de avaliação de medição de indicadores de desempenho nas atividades e, por meio da adoção de recursos tecnológicos, facilitando o processo de manipulação dos dados relativos as medições de desempenho dos Tribunais de Contas e permitindo a expansão da natureza das análises e extração de mais informações de dados.

1.4 - ORGANIZAÇÃO DO TRABALHO

Este trabalho está organizado em 5 capítulos, que são apresentados na estrutura analítica da dissertação, ilustrada na Figura 1.2, com o intuito de demonstrar visualmente a composição deste trabalho de pesquisa.

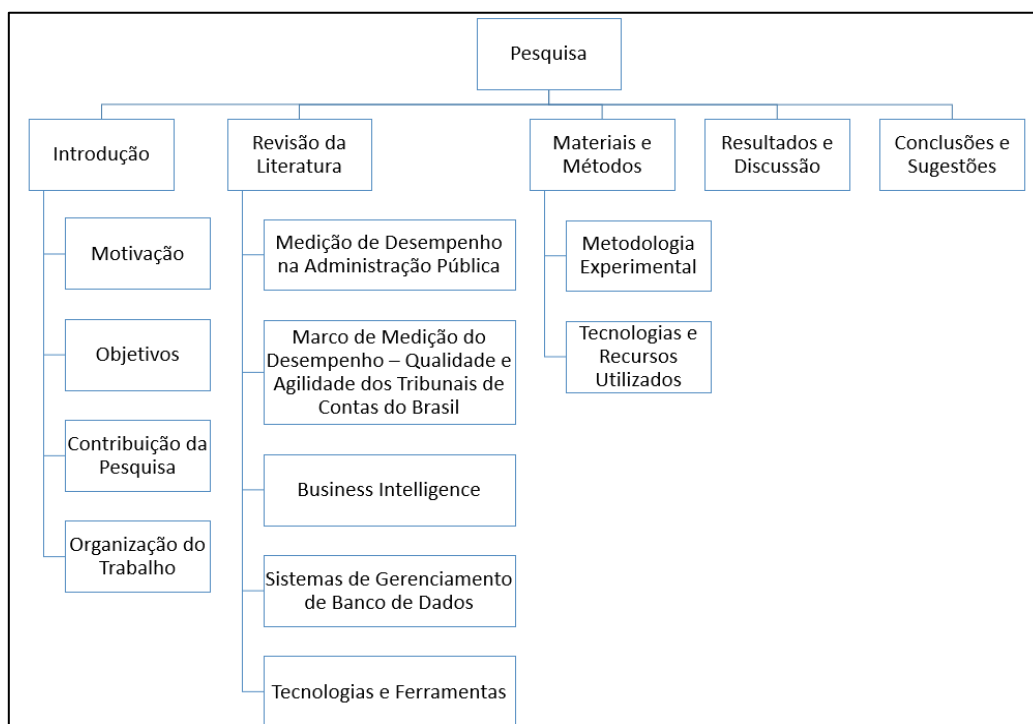


Figura 1.2 - Estrutura analítica da dissertação.

Na introdução, constam a motivação com a devida contextualização e apresentação do problema da pesquisa, os objetivos, as contribuições da dissertação e a forma de organização do trabalho.

A revisão bibliográfica discute a medição de desempenho na Administração Pública, o Marco de Medição do Desempenho – Qualidade e Agilidade dos Tribunais de Contas (MMD-QATC), os conceitos de *Business Intelligence*, de Sistemas de Gerenciamento de Banco de Dados e de tecnologias e ferramentas utilizadas.

Materiais e métodos apresenta a metodologia de trabalho, montada para que os objetivos especificados nesse trabalho fossem atingidos, e as tecnologias utilizadas para tanto.

Resultados e discussão aborda de forma detalhada e completa os resultados obtidos.

E conclusões e sugestões apresenta as conclusões finais e sugere pontos para a continuação do trabalho.

CAPÍTULO 2

REVISÃO DA LITERATURA

2.1 - MEDIÇÃO DE DESEMPENHO NA ADMINISTRAÇÃO PÚBLICA

Foram identificados diversos trabalhos que abordam a temática de medição de desempenho na administração pública.

GHSI (2000) afirma que, naquele momento, no início do século XXI, apesar da recorrência do tema na literatura, a avaliação auto referida de desempenho ainda não havia sido incorporada ao cotidiano da administração pública. Tal fato, dever-se-ia pela complexidade das metodologias de avaliação, desatenção, desconhecimento ou desinteresse dos administradores públicos.

O referido autor salienta a importância da aferição de desempenho e afirma que a gestão por meio de indicadores (que define como números, porcentagens ou razões que medem um aspecto de desempenho) tem como essência a busca do aprimoramento das instituições mediante a percepção de suas ações. Não representando o indicador a realidade em si, mas uma aproximação do que estaria ocorrendo.

Afirma ainda que, as informações sobre desempenho são essencialmente comparativas. Os dados não deveriam ser utilizados de forma isolada mas sim comparados com aqueles apresentados por instituições de natureza semelhante no mesmo período de tempo, com uma série histórica de dados da mesma instituição e/ou com metas previamente traçadas.

Por sua vez, SILVA e DRUMOND (2004) insere uma nova variável a ser observada para justificar a necessidade da utilização de indicadores de desempenho: o custo. Aponta um constante aumento da demanda por serviços públicos, sem que haja, no entanto, um crescimento equivalente de receita. Tal cenário tornaria imperativo ao Estado tomar medidas objetivando ampliação das fontes de recursos e redução nos gastos, de modo que o comprometimento das metas fiscais não fosse afetado.

A implantação de sistemas de apuração de custos e a adoção de indicadores de desempenho da atividade estatal são apresentadas no trabalho como opções para tanto. Tornar-se-ia possível assim, à Administração Pública, verificar desperdícios e avaliar as relações entre custo e benefício das diversas alternativas de políticas que estão à

disposição, permitindo que os bens públicos possam ser produzidos com menor custo e maior eficiência e eficácia.

É abordado o fato de que o bom desempenho das organizações públicas depende do uso racional dos recursos disponíveis na busca das metas previstas nos instrumentos de planejamento e de ações que realinhem os objetivos da organização com as aspirações da sociedade, sob pena de a organização atingir padrões eficientes na prestação de serviços que têm pouca importância para os usuários ou cidadãos.

REIS *et al.* (2007) demonstra os esforços na busca de indicadores de desempenho gerencial, elaborados pela própria Secretaria-Geral de Administração do Tribunal de Contas da União, e adequá-los à estrutura de BSC compatível com o setor público. O estudo discorre ainda sobre os tipos de desafios para implantação de mudanças dentro de uma organização, quais sejam: cognitivos, dos recursos, motivacionais e políticos. Faz-se necessário o envolvimento e empenho da alta cúpula para que os objetivos e metas sejam alcançados.

Quanto a adoção do Balanced Scorecard (BSC), cita que é importante ter em mente que foi concebido originalmente para empresas que buscam lucro, sendo necessária uma adaptação na sua estrutura básica para o setor público poder utilizá-lo. Uma vez que a gestão pública lida com a gestão de bens e interesses da coletividade, ela acaba por ser revestida de um formalismo legal que imputa responsabilidades decorrentes ao gestor público. Não obstante, REIS *et al.* (2007) concluem ser perfeitamente possível realizar adaptações no BSC, fazendo a ressalva que, logicamente, os indicadores ali trabalhados são reflexo daquela realidade e sugerindo que cada organização deve observar as suas peculiaridades.

CASTALDELLI (2010) apresenta um trabalho cujo o objetivo é analisar a adoção de uma política de divulgação usada pela entidade fiscalizadora brasileira, o Tribunal de Contas da União (TCU), para divulgação de seus indicadores de desempenho. Concluiu-se que o significado de indicadores não estaria claro, que eles não estariam sendo utilizados de forma estável, com exceção dos últimos anos analisados no estudo, e que haveria uma predileção por apresentar indicadores de *output*, enquanto a tendência internacional seria de mostrar métricas orientadas a resultado. Conforme o autor, isso indicaria uma tendência de demonstrar a capacidade de entregar serviços e produtos da organização.

Para GESPÚBLICA(2009), os indicadores são instrumentos de gestão essenciais nas atividades de monitoramento e avaliação das organizações, assim como seus projetos,

programas e políticas, pois permitem acompanhar o alcance das metas, identificar avanços, melhorias de qualidade, correção de problemas, necessidades de mudança etc.

Afirma, também, que os indicadores não são simplesmente números, são atribuições de valor a objetivos, acontecimentos ou situações, de acordo com regras, que possam ser aplicados critérios de avaliação, como, por exemplo, eficácia, efetividade e eficiência. E que serviriam para:

- Mensurar os resultados e gerir o desempenho;
- Embasar a análise crítica dos resultados obtidos e do processo de tomada decisão;
- Contribuir para a melhoria contínua dos processos organizacionais;
- Facilitar o planejamento e o controle do desempenho; e
- Viabilizar a análise comparativa do desempenho da organização e do desempenho de diversas organizações atuantes em áreas ou ambientes semelhantes.

Quanto ao conceito de desempenho, GESPÚBLICA (2009) afirma ser peculiar e requer das organizações a utilização de um metamodelo dinâmico, abrangente e multidimensional que permita a construção, caso a caso, de seu próprio conceito de desempenho.

De forma geral, os trabalhos apresentam a relevância de se trabalhar com indicadores para a medição de desempenho na Administração Pública, utilizando diferentes abordagens para tanto. Salientam a singularidade contextual dos indicadores a serem trabalhados dentro uma organização e a necessidade de utilizá-los não de modo isolado, mas comparando-os com séries históricas ou com resultados de instituições de natureza semelhante no mesmo período de tempo.

2.2 - MARCO DE MEDIÇÃO DO DESEMPENHO – QUALIDADE E AGILIDADE DOS TRIBUNAIS DE CONTAS DO BRASIL (MMD-QATC)

2.2.1 - Origem e evolução

A Associação dos Membros dos Tribunais de Contas do Brasil - ATRICON, criada em 16 de Agosto de 1992, atua com o “intuito de garantir a representação, a defesa, o aperfeiçoamento e a integração dos Tribunais de Contas e de seus Membros (Ministros, Conselheiros, Ministros Substitutos e Conselheiros Substitutos), visando aprimorar o Sistema de Controle Externo do Brasil em benefício da sociedade”. (ATRICON, 2020)

ATRICON (2013) aponta que dentre suas iniciativas de apoio ao desenvolvimento de estratégias e ferramentas para aprimorar a eficácia dos sistemas de controle da Administração Pública, em 02 de Julho de 2013, a ATRICON estabeleceu critérios por meio do regulamento ATRICON Nº 01/2013 - que “estabelece as regras e os parâmetros de qualidade e agilidade do controle externo, para fins de avaliação dos Tribunais de Contas, e dá outras providências” - a serem aplicados na avaliação dos Tribunais de Contas.

E em 06 de agosto de 2014, decidiu regulamentar diretrizes orientativas para os Tribunais de Contas. Surgiram, a partir deste momento, as resoluções elaboradas diretamente pela ATRICON e as elaboradas em conjunto com o Colégio dos Corregedores e Ouvidores dos Tribunais de Contas do Brasil (CCOR), conforme elencado na Tabela 2.1 abaixo:

Tabela 2.1 - Resoluções que serviram de base para o MMD-QATC.

Resolução	Tema
ATRICON Nº 1/2014	Agilidade no julgamento de processos e gerenciamento de prazos pelos Tribunais de Contas do Brasil.
ATRICON Nº 2/2014	Controle externo concomitante: instrumento de efetividade dos Tribunais de Contas do Brasil.
ATRICON Nº 3/2014	Composição, organização e funcionamento dos Tribunais de Contas do Brasil: adequação ao modelo constitucional.
ATRICON Nº 4/2014	Controle Interno: instrumento de eficiência dos Tribunais de Contas do Brasil.
ATRICON Nº 5/2014	Controle Interno: instrumento de eficiência dos Jurisdicionados.
ATRICON Nº 6/2014	Divulgação de decisões e de pautas de julgamento como instrumento de comunicação dos Tribunais de Contas do Brasil com o público externo de interesse e com a sociedade.
ATRICON Nº 7/2014	Gestão de Informações Estratégicas pelos Tribunais de Contas do Brasil: instrumento de efetividade do controle externo.

ATRICON Nº 8/2014	Os Tribunais de Contas do Brasil e o controle do cumprimento do artigo 5º da Lei 8.666/93: ordem nos pagamentos públicos.
ATRICON Nº 9/2014	Os Tribunais de Contas e o desenvolvimento local: Controle do tratamento diferenciado e favorecido às microempresas e empresas de pequeno porte nas contratações públicas.
ATRICON-CCOR Nº 1/2014	Corregedorias: instrumentos de eficiência, eficácia e efetividade dos Tribunais de Contas do Brasil.
ATRICON-CCOR Nº 2/2014	Ouvidorias: instrumentos de interação dos Tribunais de Contas do Brasil com a sociedade.

Cada uma dessas resoluções apresenta as justificativas para elaboração do documento, os objetivos traçados, os compromissos firmados pelos membros, os princípios e fundamentos legais que embasaram a sua elaboração, além dos conceitos adotados e as diretrizes apontadas.

As resoluções listadas acima e o regulamento ATRICON Nº 01/2013 culminaram, em 2014, no “Marco de Medição do Desempenho – Qualidade e Agilidade dos Tribunais de Contas do Brasil” (MMD-QATC), que incorpora as diretrizes da ATRICON e está fortemente baseado em normas internacionais aceitas pelas Entidades Fiscalizadoras Superiores.

Após o estabelecimento do MMD-QATC em 2014, os TCs realizaram a auto avaliação em 2015, cujos resultados foram objeto de divulgação por ocasião do XXVIII Congresso dos Tribunais de Contas do Brasil, realizado em Recife (ATRICON, 2017).

Posteriormente, foram lançadas as versões de 2017 e 2019 do Marco, que foram fruto de novo aperfeiçoamento da ferramenta decorrente das lições aprendidas durante a aplicação do MMD em 2015 e 2017.

2.2.1.1 - Entidade Fiscalizadora Superior (EFS)

A Entidade Fiscalizadora Superior (EFS), ou instituição de auditoria nacional, cumpre a função de auditoria independente e técnica no setor público, que normalmente é estabelecida pela Constituição de um país ou pelo órgão legislativo supremo. Uma EFS é responsável pela supervisão e tomada de contas do governo quanto à utilização dos recursos públicos, em conjunto com o Legislativo e outros órgãos de supervisão. Num

sistema democrático, a EFS costuma reportar seus achados ao Legislativo e, em alguns casos, ao executivo ou o chefe de Estado. Onde houver mais de um organismo que cumpre o papel de auditoria externa no setor público, a EFS é geralmente detentora das mais sólidas garantias constitucionais de independência (OCDE, 2011).

2.2.2 - Metodologia e estrutura do MMD-TC

O *Supreme Audit Institutions – Performance Measurement Framework – SAI PMF*, da Intosai, serviu de inspiração para a metodologia empregada no MMD-TC que, por sua vez, incorporou as diretrizes da Atricon, as NBASP e as ISSAIs, constituindo-se no principal instrumento de avaliação dos Tribunais de Contas do Brasil e objetivando verificar o seu desempenho e identificar seus pontos fortes e fracos (ATRICON, 2019).

Este manual de “Procedimentos MMD-TC 2019” apresenta ainda como objetivos:

- Melhorar a gestão e a governança;
- Enfatizar a transparência, a accountability e o desempenho;
- Monitorar ao longo do tempo a implementação de melhorias;
- Avaliar e disseminar boas práticas de controle e de gestão; e
- Estimular a participação social por meio da produção de conhecimento em proveito da sociedade e do poder público de forma ampla.

A adesão e participação dos Tribunais de Contas é voluntária, formalizada por meio de termo de adesão junto à Atricon, e a periodicidade da aplicação do MMD-TC, sob a coordenação daquela Associação, ocorre bienalmente, conforme cronograma estabelecido.

Procurando abranger processos relativos às atividades finalísticas de controle externo e os processos de governança e gestão, o Marco foi estruturado conforme apresentado na Figura 2.1, onde é possível observar uma forma piramidal dividida em cinco níveis, sugerindo, de tal modo, uma organização hierárquica cuja a ponta simboliza o próprio Marco de Medição de Desempenho, seguido de domínios, indicadores, dimensões e critérios, na base da estrutura.

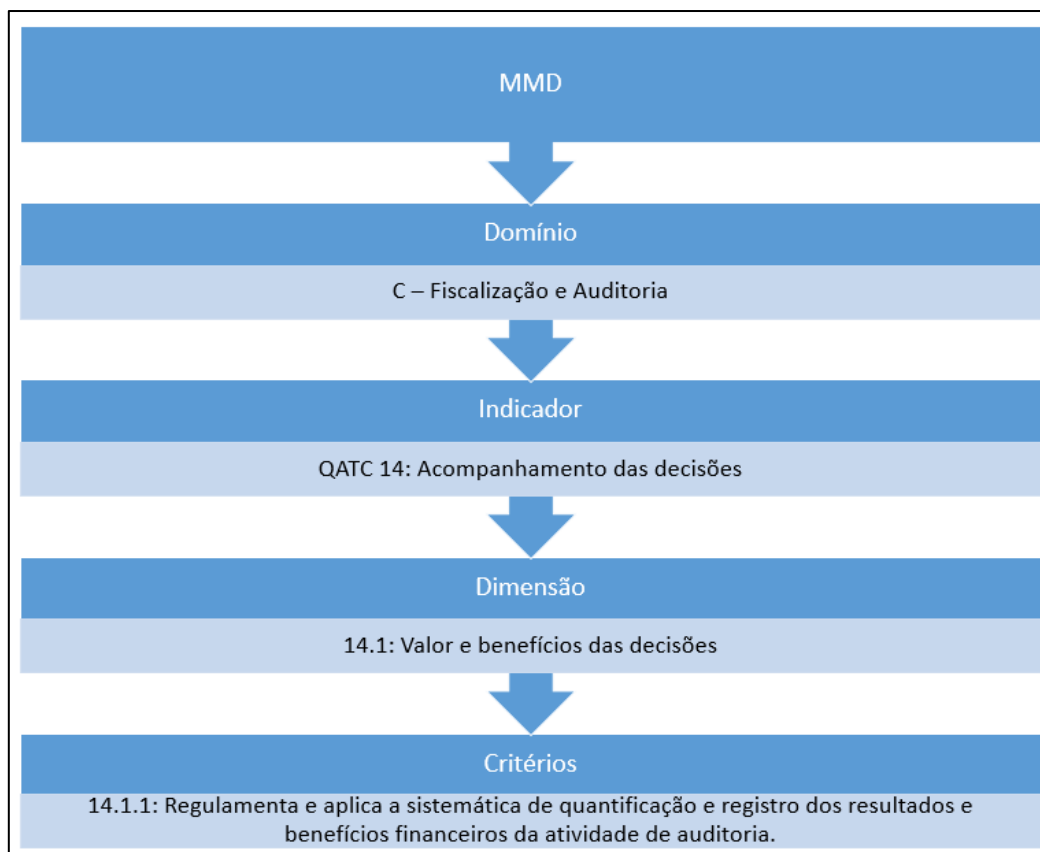


Figura 2.1 - Estrutura do MMD-TC.
 Fonte: Adaptado de ATRICON (2019).

2.2.2.1 - Domínio

ATRICON (2019) define, no contexto do MMD-TC, o nível “Domínio” como: o título que se dá ao agrupamento de indicadores com temas correlatos, não havendo agregação de pontuação nesse nível. A estrutura comporta um total de 6 domínios, quais sejam:

- Domínio A: independência e marco legal;
- Domínio B: governança interna;
- Domínio C: fiscalização e auditoria;
- Domínio D: fiscalização da infraestrutura e meio ambiente;
- Domínio E: fiscalização e auditoria de políticas públicas sociais; e
- Domínio F: fiscalização e auditoria da gestão fiscal, controle interno, tecnologia da informação, transparência e ouvidoria.

2.2.2.2 - Indicador

ATRICON (2019) define, no contexto do MMD-TC, o nível “Indicador” como: título que se dá ao agrupamento de dimensões com temas correlatos, tendo como pontuação a conversão das pontuações das dimensões que a ele se vinculam e indicará objetivamente o nível de desempenho do Tribunal de Contas nas respectivas áreas-chave.

Existe um total de 25 indicadores mapeados e distribuídos entre os seis domínios, sendo: Domínio A com um (1) indicador; Domínio B com seis (6) indicadores; Domínio C com oito (8) indicadores; Domínio D com três (3) indicadores; Domínio E com quatro (4) indicadores; e Domínio F com três (3) indicadores. O Apêndice A apresenta quais são os indicadores trabalhados.

2.2.2.3 - Dimensão

ATRICON (2019) define, no contexto do MMD-TC, o nível “Dimensão” como: título que se dá ao agrupamento de critérios com temas. A pontuação de cada dimensão será definida em função do atendimento ou não dos critérios que a elas se vinculam. Cada indicador possui até quatro (4) dimensões, chegando-se, ao final, a um total de setenta e nove (79).

2.2.2.4 - Critério

ATRICON (2019) define, no contexto do MMD-TC, o nível “Critério” como: conjunto de requisitos usados como referência para a avaliação dos TCs, elaborados com base em leis, regulamentos, diretrizes, normas, melhores práticas etc.

2.2.2.5 - Nível de desempenho

ATRICON (2019) define, no contexto do MMD-TC, o “Nível de Desempenho” como um resultado mensurável com base na pontuação final de cada indicador, podendo variar, na escala de medição do MMD-TC, de 0 a 4 pontos.

A Figura 2.2 apresenta tal escala, onde pontuações mais altas demonstram níveis mais altos de maturidade de um indicador. Sendo: o nível 0 a indicação que a prática avaliada não existe; o nível 1 que a prática existe, porém com aspectos insatisfatórios; o

nível 2 que a prática existe em nível insatisfatório, porém com sinais claros de aperfeiçoamento; o nível 3 que a prática existe em nível satisfatório e sendo executada conforme o previsto nas Resoluções da Atricon e nas NBASP; e finalmente, o nível 4 que a prática segue todos os padrões estabelecidos e o Tribunal implementa as atividades de maneira que lhe permite avaliar e melhorar constantemente o seu desempenho.

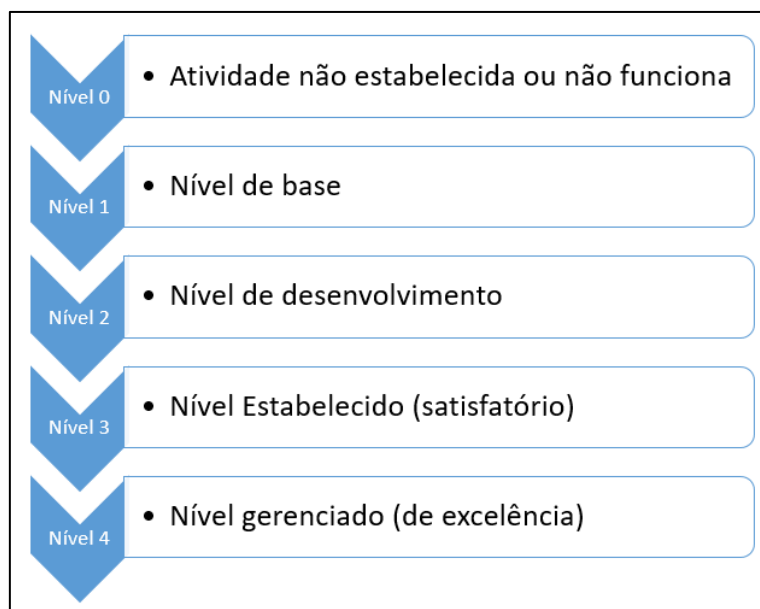


Figura 2.2 - Escala dos níveis de desempenho.
Fonte: Adaptado de ATRICON (2019).

2.3 - BUSINESS INTELLIGENCE

Segundo TURBAN *et al.* (2009), Business intelligence (BI) é um termo “guarda-chuva” que inclui arquiteturas, ferramentas, bancos de dados, aplicações e metodologias. Podendo variar de significado de pessoa para pessoa, fato este causado pela enxurrada de acrônimos e termos da moda associadas a ele e suas ferramentas (como *business performance management* – BPM).

Segundo SABHERWAL e BECERRA-FERNANDEZ (2011), *Business Intelligence* (BI) é altamente importante para organizações dos mais diferentes ramos da indústria, que têm obtido benefícios significativos através do cuidadoso uso deste recurso. No entanto, apontam a necessidade de distinguir três conceitos importantes: dado, informação e conhecimento. Os dados seriam números brutos ou afirmações e, portanto, podendo ser desprovidos de significado, contexto ou intenção. As informações seriam um subconjunto de dados, incluindo apenas os dados que possuem contexto, relevância e

propósito, normalmente envolvendo a manipulação desses dados brutos com objetivo de obter uma indicação mais significativa de tendências ou padrões nos dados. E finalmente, o conhecimento se diferenciaria da informação por não ser simplesmente um conjunto de fatos com maior riqueza de detalhes, mas uma crença justificada sobre relações entre conceitos relevantes para aquela área particular.

A Figura 2.3 resume esses conceitos e acrescenta um outro fator, as decisões, que quando baseadas em conhecimento tornam-se mais confiáveis que as tomadas simplesmente com base em dados.

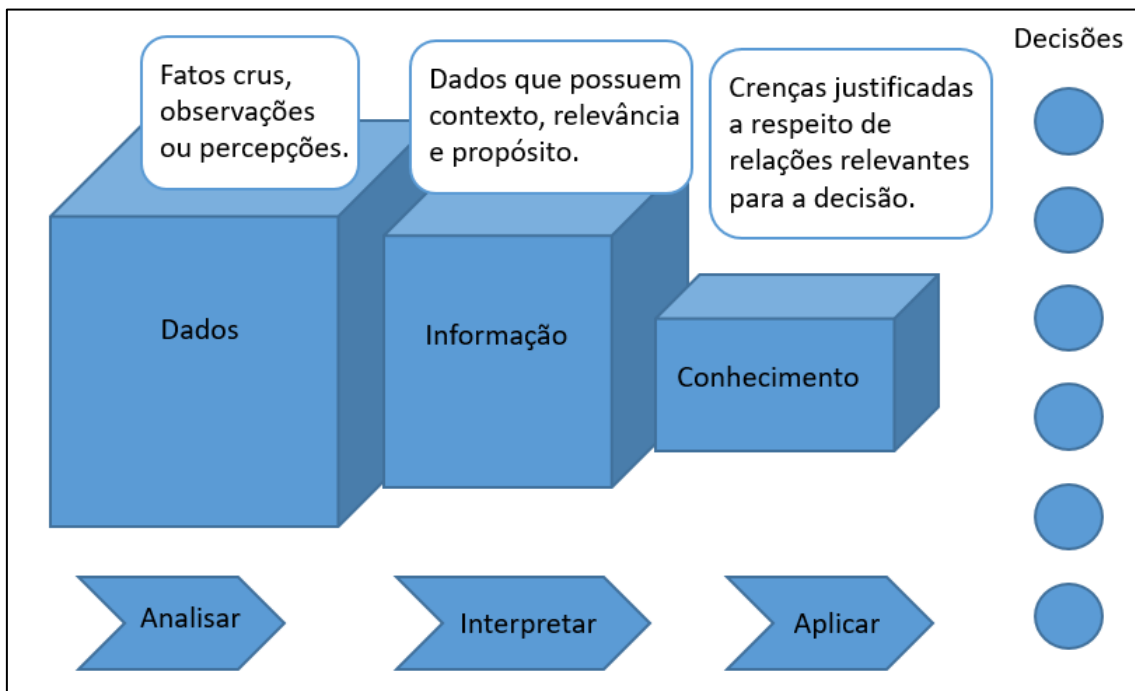


Figura 2.3 - Dado, informação, conhecimento e decisões.

Fonte: Adaptado de SABHERWAL e BECERRA-FERNANDEZ (2011).

POPOVIČ *et al.* (2010) acrescentam outros fatores menos tangíveis à esta discussão: salientam a importância do fator humano dentro de uma solução de BI e que esta não existiria sem que houvesse pessoas para interpretar o significado e a relevância das informações e agissem de acordo com o conhecimento adquirido. E ainda que, a qualidade da solução depende do conhecimento específico dos especialistas (a fim de preencher a lacuna entre a área de tecnologia da informação e a de negócios), de estrutura e cultura organizacionais adequadas que permitam melhorias no processo de negócios por meio da identificação de “perguntas certas” e de como elas podem conduzir até esse aprimoramento.

Então, BI pode ser definido como fornecer informações e conhecimentos valiosos aos tomadores de decisão, utilizando para isso uma variedade de fontes de dados, internas ou externas à organização, estando os dados estruturados ou não, podendo ser qualitativos ou quantitativos (SABHERWAL 2007, 2008).

TURBAN *et al.* (2009) corroboram com este pensamento ao afirmar que o processo do BI baseia-se na transformação de dados em informações, depois em decisões e finalmente em ações.

Uma definição, que acreditamos ser mais completa e que engloba os aspectos relatados acima, foi dada por FOLEY e GUILLEMETTE (2010), que propuseram BI como uma combinação de processos, política, cultura, e tecnologias para coletar, manipular, armazenar e analisar os dados coletados de fontes internas e externas, a fim de se comunicar informação, criar conhecimento e auxiliar a tomada de decisão. BI ajuda a relatar desempenho do negócio, descobrir novas oportunidades e tomar melhores decisões sobre concorrentes, fornecedores, clientes, questões financeiras e estratégicas, produtos e serviços.

Quanto à adoção de BI em organizações de diferentes tamanhos, HORAKOVA e SKALSKA (2013) apontam que embora possa ser utilizado com eficácia em empresas de todos os portes, o uso de BI em pequenas e médias empresas é menor do que em grandes, por parecerem muito complexas ou caras para as necessidades das pequenas empresas. Como alternativa a este ponto, os autores sugerem a adoção de soluções de BI em nuvem, que podem oferecer menor custo de implementação e facilidade de uso, ou a utilização de ferramentas e aplicativos de código aberto. No estudo em questão os autores utilizaram um solução mista aproveitando as vantagens tanto do software comercial quanto do *open-source* (código-fonte aberto) para construir um BI para analisar custos e lucros de uma pequena empresa.

A Figura 2.4 deixa claro o nível de complexidade a que pode chegar uma solução de BI, envolvendo diferentes ferramentas, técnicas e múltiplas origens de dados que podem ser incluídas nesse contexto:

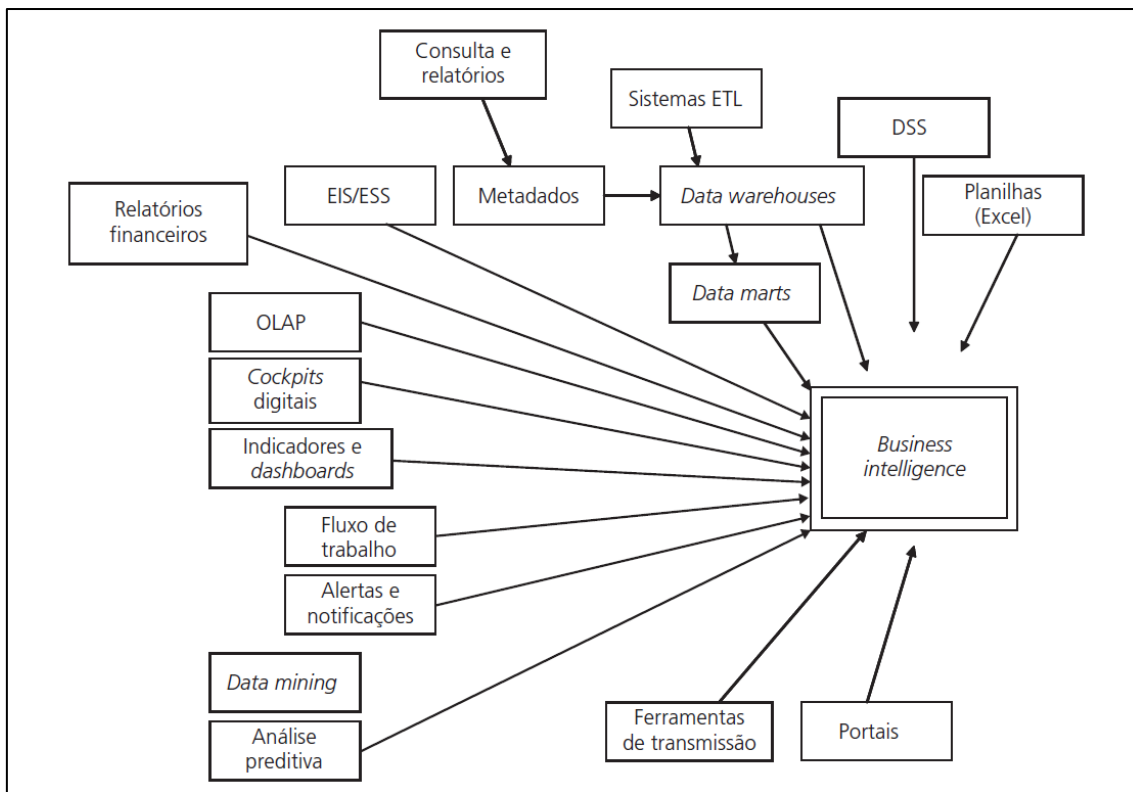


Figura 2.4 - Ferramentas, técnicas e origem de dados.
 Fonte: TURBAN *et al.* (2009).

SABHERWAL e BECERRA-FERNANDEZ (2011) afirmam ainda que BI utiliza dados de fontes variadas (incluindo *data warehouses*), e informações que são produzidas por meio de análises apropriadas e, em seguida, apresentadas de maneira amigável, como por meio de *scorecards* e *dashboards*. As soluções de BI utilizam ferramentas de BI para dar suporte ao processo de BI, processo por meio do qual informações e conhecimentos são providos. As ferramentas de BI podem ainda ser empregadas na obtenção de dados e informações por meio de extração, transformação e carga de dados. A Figura 2.5 apresenta a inter-relação entre os conceitos de produto, processo, solução e ferramentas de BI:

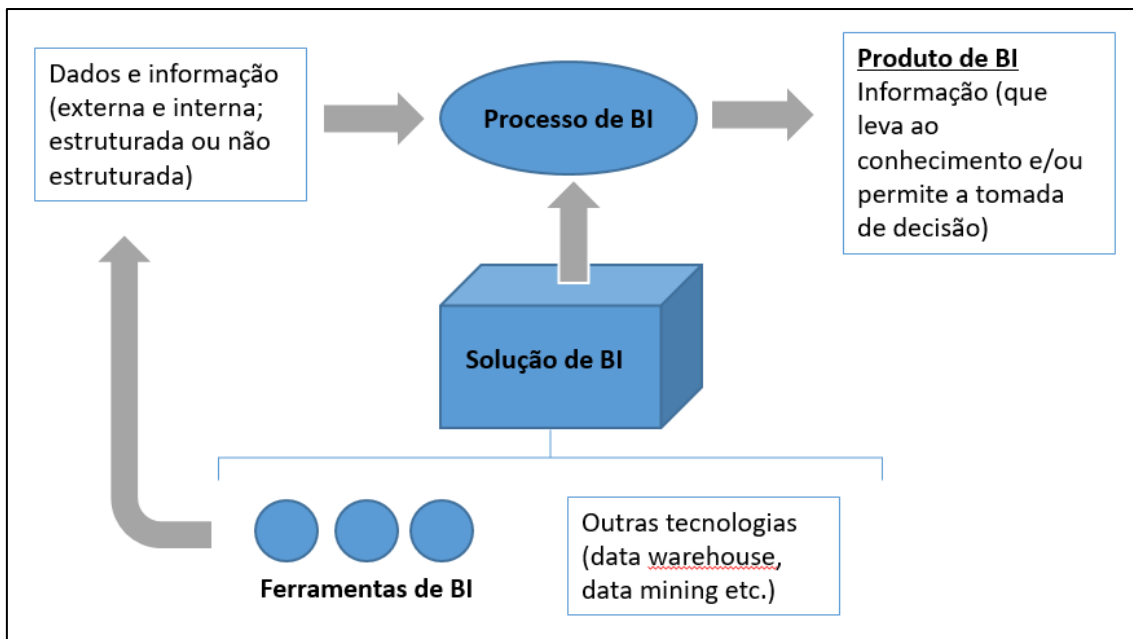


Figura 2.5 - Produto, processo, solução e ferramentas de BI.

Fonte: Adaptado de SABHERWAL e BECERRA-FERNANDEZ (2011).

O processo de BI possui como entrada os dados e informações e tem como resultado direto novas informações, conhecimentos ou *insights* obtidos ao revelar conexões ou padrões previamente desconhecidos (SABHERWAL e BECERRA-FERNANDEZ, 2011).

De maneira similar, para TURBAN *et al.* (2009), os objetivos do BI são possibilitar manipulação e interação com os dados e fornecer aos gerentes e analistas de negócios um mecanismo que permita uma análise adequada. De tal maneira, analisando dados, situações e desempenhos históricos e atuais, os tomadores de decisão conseguem valiosos *insights* que servem de base para decisões melhores.

Dentre outras tecnologias que normalmente estão relacionadas com *Business Intelligence* mas que, no entanto, são diferentes estão: *data warehouse (DW)*, *datamart*, *data mining* e sistemas de suporte à decisão.

Data warehouse (DW), ou *Data Warehouse* empresarial (EDW) é uma coleção de dados, atuais e históricos, orientada por assunto, integrada, variável no tempo e não-volátil, que proporciona suporte ao processo de tomada de decisões da gerência (TURBAN *et al.*, 2009).

No mesmo sentido, SABHERWAL E BECERRA-FERNANDEZ (2011) afirmam que o DW é um repositório único de dados que são obtidos de múltiplos sistemas

transacionais (frequentemente com inconsistências), utilizando ferramentas para extrair, transformar (objetivando tornar os dados consistentes) e fazer a carga de dados.

Este processo é conhecido tecnicamente pela sigla ETL (*Extract, Transform and Load*), que significa, em tradução livre: extrair, transformar e carregar. TURBAN *et al.* (2009) detalham as etapas do processo de ETL:

- Extração: leitura dos dados de um ou mais bancos de dados;
- Transformação: conversão dos dados extraídos de sua forma anterior na forma em que precisam estar, para que sejam colocados em um *data warehouse* ou apenas em outro banco de dados;
- Carga: colocação dos dados no *data warehouse*. A transformação ocorre com o uso de regras ou tabelas de busca ou com a combinação dos dados com outros dados.

Um *datamart* é um subconjunto de um *data warehouse*, que normalmente consiste em uma única área temática (p. ex., marketing, operações). Pode ser dependente (criado diretamente a partir do DW) ou independente (um *warehouse* pequeno, projetado para um departamento, mas cuja fonte não é um EDW) (TURBAN *et al.*, 2009).

Data mining, por outro lado, refere-se ao processo de descobrir padrões ocultos nos dados, normalmente em um *data warehouse* (SABHERWAL e BECERRA-FERNANDEZ, 2011).

Sistemas de suporte à decisão focam no suporte ou automação da tomada de decisão nas organizações. Utilizam dados do DW ou de sistemas transacionais juntamente com conhecimento prévio como *inputs* do processo (DAVENPORT e HARRIS, 2005).

Um dos fatores que diferenciaria essas tecnologias, segundo SABHERWAL e BECERRA-FERNANDEZ (2011), é a característica de que BI pode incorporar dados internos e externos à organização, de origem estruturada ou não estruturada enquanto as outras trabalham basicamente com dados internos de origem estruturada. Fato este que torna-se relevante a partir do momento que observamos que dados de concorrentes, de clientes e de outras entidades (relevantes ao contexto) não estão disponíveis em sistemas internos da organização, e em muitas vezes de forma não estruturada (como e-mails, apresentações, *web pages* etc).

Todos esses processos de trabalho e tecnologias precisam estar organizados em uma arquitetura de maneira a permitir a interoperacionalização entre os seus componentes. NEGASH e GRAY (2008) entendem que uma típica solução de BI consiste em 4 níveis de componentes e gestão de metadados, quais sejam:

- Nível de sistema operacional: os sistemas operacionais do negócio são principalmente os sistemas de processamento de transação online (OLTP – *Online Transaction Processing*) que dão suporte as operações do dia a dia da empresa.
- Nível de aquisição de dados: aqui ocorre o processo de ETL dos dados extraídos de sistemas OLTP para o DW. Neste nível os dados são tratados e transformados de acordo com as regras estabelecidas.
- Nível de armazenamento de dados: aqui residem os dados processados e armazenados pela ETL na base de DW, que é normalmente implementada utilizando sistemas tradicionais de gerenciamento de banco de dados (SGBDs). Os esquemas do tipo estrela (*star schema*) e floco de neve (*snowflake schema*) são os mais populares adotados para a construção de *schemas* de *data warehouse*.
- Nível analítico: neste nível os dados são apresentados por meio dos mais diversos tipos de ferramentas analíticas, geradores de relatórios e de processamento analítico *online* (OLAP), que permitem navegar de forma eficiente nos dados a partir de diferentes dimensões de análise.
- Gestão de Metadados: metadados são dados sobre outros dados, como por exemplo autorizações de acesso, regras de negócio etc. São cruciais para a acurácia e consistência das informações e a manutenção do sistema.

Essa arquitetura simples e funcional, que serviu de base para o desenvolvimento deste estudo, é apresentado na Figura 2.6 por meio de um diagrama demonstrando a organização e a ligação desses níveis, bem como exemplos de componentes internos a cada parte. Nela podemos identificar o caminho que os dados fazem desde as bases de dados e sistemas de origem até serem disponibilizados ao usuário final.

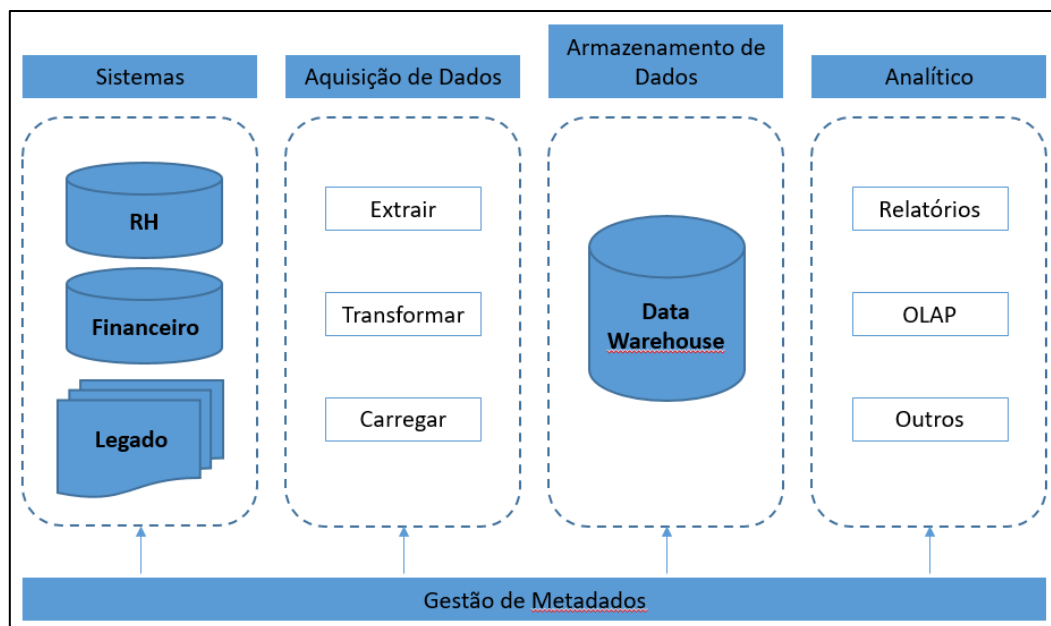


Figura 2.6 - Arquitetura geral de um sistema de BI tradicional.
 Fonte: Adaptado de NEGASH e GRAY (2008).

A saber, TURBAN *et al.* (2009) apresenta algumas alternativas para tipos básicos de arquiteturas para a construção do *data warehouse* e, por conseguinte, para a arquitetura da solução de BI, que devem ser consideradas de acordo com o contexto, tais como: *warehouse* empresarial com ODS (isto é, suporte ao acesso em tempo real) e *data store* operacional, arquitetura de *data warehouse* distribuída, arquitetura em estrela de *data mart*.

2.4 - SISTEMAS DE GERENCIAMENTO DE BANCO DE DADOS

Antes de abordar o que venham a ser sistemas de gerenciamento de banco de dados (SGBDs), faz-se necessário estabelecer que banco de dados são coleções de fatos conhecidos que podem ser registrados e possuem significado implícito. Isto posto, podemos definir um SGBD como uma coleção de programas que permite aos usuários criar e manter um banco de dados, facilitando o processo de definição, construção, manipulação e compartilhamento de bancos de dados entre diversos usuários e aplicações (ELMASRI e NAVACHE, 2011).

DATE (2003) e ELMASRI e NAVACHE (2011) apontam como vantagens do uso de um SGBD: natureza de autodescrição de um sistema de banco de dados, isolamento entre programas e dados, controle centralizado dos dados operacionais da empresa, suporte para múltiplas visões dos dados, suporte de transações multiusuário, controle de

redundância, controle de acesso e restrições de segurança, possui estruturas de armazenamento e técnicas de pesquisa para o processamento eficiente de consultas, opções de backup e recuperação, restrições de integridade, representação de relacionamentos complexos entre dados e outros.

Diversos produtos que implementam tais características estão disponíveis no mercado, sob licenças de uso pagas ou sob licenças open source ou de uso livre, podendo ainda serem classificados como relacionais ou não relacionais (também conhecidos como NoSQL). Exemplos de SGBDs são Microsoft SQL Server, Oracle Database, PostgreSQL, Firebird, HSQLDB, MySQL, MongoDB e outros. Neste trabalho, o tipo e os conceitos adotados são os condizentes com os bancos de dados relacionais.

2.4.1 - Modelo conceitual, modelo lógico e projeto físico

O projeto de um banco de dados abrange as fases de modelagem conceitual, modelo lógico e projeto físico. Modelo conceitual (ou de alto nível) é um modelo de dados abstrato que descreve a estrutura de um banco de dados, demonstrando todas as relações entre as entidades, suas especializações, seus atributos e auto-relacionamentos, de forma independente de um SGBD em particular. O modelo lógico (ou modelo de dados representativos ou de implementação) apresenta a visão no nível de abstração do usuário, ligações entre as tabelas de banco de dados, as chaves primárias, os componentes de cada uma e outras características, e é dependente do SGBD em uso. E o projeto físico (ou de baixo nível) oferece conceitos que descrevem os detalhes de como os dados são armazenados no computador e acrescenta detalhes que impactam no desempenho mas sem interferir nas funcionalidade (HEUSER, 2009, ELMASRI e NAVACHE, 2011).

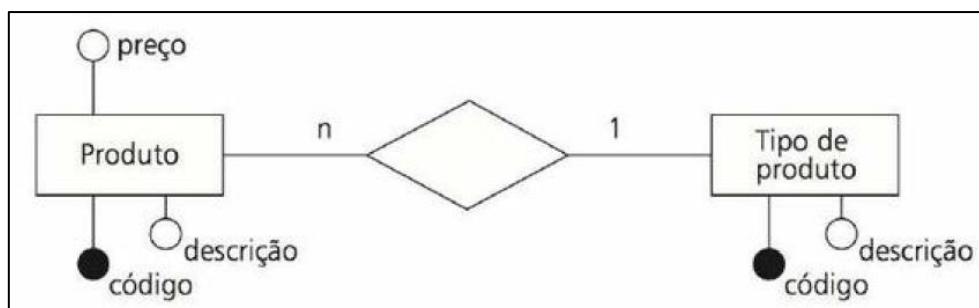


Figura 2.7 - Exemplo de modelo conceitual.
Fonte: HEUSER (2009).

Alguns conceitos importantes para o entendimento dos modelos de dados são definidos por HEUSER (2009):

- Entidade: representam as tabelas nos bancos de dados. São um conjunto de objetos da realidade modelada sobre os quais se deseja manter informações no banco de dados.
- Atributo: campo existente em uma tabela. Representa os dados que são associados a cada ocorrência de uma entidade ou relacionamento.
- Relacionamento: parte essencial da modelagem que indica a relação entre as tabelas, são associações entre as entidades.
- Cardinalidade: o número mínimo e máximo de ocorrências de entidades associadas a uma ocorrência da entidade em questão através do relacionamento.

2.4.2 - Linguagem SQL

Para a criação das bases de dados e manipulação desses dados é necessário que os SGBDs forneçam mecanismos, linguagens ou interfaces, que permitam aos usuários (de diferentes níveis) interagirem com o software. ELMASRI e NAVACHE (2011) apontam a linguagem de banco de dados relacional SQL como um típico exemplo de linguagem abrangente. Ela contempla comandos de: linguagem de definição de dados (DDL – *Data Definition Language*), utilizados por projetistas para a criação da base de dados e seus objetos; e comandos de manipulação de dados (DML – *Data Manipulation Language*), utilizados para a recuperação, inserção, exclusão e modificação dos dados.

2.5 - TECNOLOGIAS E FERRAMENTAS

Para o desenvolvimento do presente estudo, foram utilizadas ferramentas de modelagem e gerenciamento de banco de dados, linguagens de programação, bibliotecas de manipulação de dados, além de softwares específicos de BI. Todos esses recursos tecnológicos foram empregados de maneira a cobrir todas as etapas de elaboração de uma solução de BI, conforme visto nas seções anteriores, e são apresentados a seguir.

2.5.1 - HeidiSQL

O HeidiSQL é um *software* livre, e *open source*, que é classificado como um software cliente SQL: permite ver e editar dados e estruturas em sistemas de banco de dados. É compatível com os sistemas gerenciadores de banco de dados (SGBDs) MariaDB, MySQL, Microsoft SQL e PostgreSQL. Está entre as ferramentas mais populares para o MySQL, software SGBD também escolhido para ser utilizado neste trabalho (HEIDISQL, 2021).

Entre os recursos disponibilizados pelo software, além dos acima já citados, estão:

- Gerar exportações de SQL, compactando-as ou armazenando na área de transferência do computador;
- Exportar de um servidor/base de dados diretamente para outro servidor/base de dados;
- Importar arquivos-texto;
- Exportar o resultado de buscas nas tabelas em formato CSV, HTML, XML, SQL, LaTeX, *Wiki Markup* and *PHP Array*;
- Inserção em lote de arquivos *ascii* ou binários em tabelas;
- Ajustar formatação visual de instruções SQL desordenadas para melhor entendimento;
- Monitorar e finalizar processos-cliente que porventura estejam travados no servidor; e
- Otimizar e reparar tabelas em lote.

2.5.2 - MySQL

O MySQL é um sistema de gerenciamento de banco de dados (SGBD); *open source*; que utiliza a linguagem SQL como interface; muito rápido, confiável, escalonável e fácil de usar; e funciona em modo cliente-servidor ou em sistemas embarcados. É atualmente um dos SGBDs mais populares e amplamente utilizado por algumas das maiores e mais promissoras empresas do mundo, como *Facebook*, *Twitter*, *Booking.com*, e *Verizon* (MYSQL, 2021b).

2.5.3 - SQL Power Architect

SQL Power Architect é uma ferramenta de modelagem de dados que possui muitos recursos, principalmente para projetistas de *data warehouse*, possuindo também recursos de engenharia reversa em base de dados existentes e geração de *scripts* DDL. Possui diferentes versões, no entanto, é disponibilizada uma versão *Community Edition*, *open source*, a qual não há necessidade de adquirir licença, sendo, assim, de livre uso (ARCHITECT, 2021).

2.5.4 - Linguagem de programação Python

Foi lançada por Guido van Rossum em 1991 com o intuito de que fosse fácil e intuitiva, priorizando a legibilidade do código sobre a velocidade de execução. O Python é gratuito, multiplataforma, combina uma sintaxe concisa e clara com os recursos poderosos de sua biblioteca padrão e por módulos e *frameworks* desenvolvidos por terceiros (INFNET, 2021).

Python é uma linguagem de programação de alto nível, interpretada, de script, imperativa, orientada a objetos, funcional, de tipagem dinâmica e forte. Atualmente, possui um modelo de desenvolvimento comunitário, aberto e gerenciado pela organização sem fins lucrativos Python Software Foundation (PYTHON, 2021).

2.5.5 - Biblioteca pandas

O Pandas é uma biblioteca desenvolvida para Python, *open source*, que fornece estruturas e ferramentas de manipulação e análise de dados. É considerada a melhor biblioteca para essa finalidade e é amplamente utilizada pelos cientistas de dados. Possui bom desempenho e boa capacidade de simplificar tarefas relacionadas à manipulação de dados (PANDAS, 2021).

2.5.6 - Suíte Pentaho

A *suite* Pentaho é formada por um conjunto de softwares voltados para construção de soluções de BI, que dentro de suas funcionalidades estão a extração de dados em diferentes origens, tratando-os e higienizando-os, realizando a gravação em uma base de

dados (que pode ser um *data warehouse*), ou repassando-os a outros sistemas de destino ou mesmo a outros componentes da suíte (INFOQ, 2021).

O Pentaho atualmente pertence a uma empresa chamada Hitachi, que comercializa uma versão *Enterprise* desse produto. No entanto, utilizaremos a versão *Community*, que é disponibilizada com algumas diferenças, mas que não são impeditivas para a realização deste trabalho.

2.5.6.1 - Pentaho Data Integration

O Pentaho Data Integration é parte das soluções disponibilizadas pela *suíte* Pentaho. Com ele é possível fazer inúmeras operações de ETL e Integração de Dados (movimentação de grandes volumes dados, transformação de dados, limpeza de dados etc).

2.5.7 - Power BI Desktop

O Power BI Desktop é um software fornecido pela empresa Microsoft que permite a construção, a partir de dados, de análises visuais e *dashboards* interativos com boa qualidade visual e facilidade de manuseio, podendo ser aplicado nas mais diversas áreas do conhecimento.

CAPÍTULO 3

MATERIAIS E MÉTODOS

Neste capítulo, estão descritas as etapas a serem seguidas para o desenvolvimento deste trabalho, ensejando atingir cada um dos objetivos propostos. A Figura 3.1 ilustra as etapas que foram planejadas e executadas para a concretização deste trabalho de pesquisa.

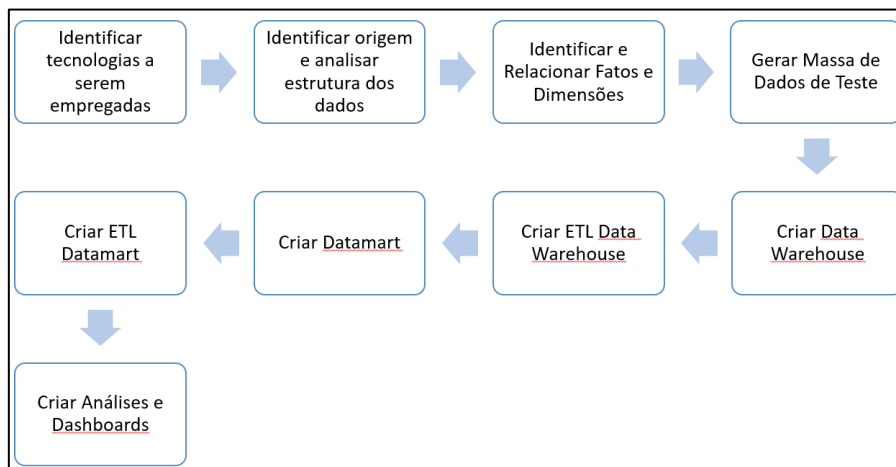


Figura 3.1 - Fluxograma das etapas do trabalho de pesquisa.

Na secção 3.1, a seguir, são apresentadas as tecnologias empregadas nas etapas identificadas na Figura 3.1. Na secção 3.2, é apresentada cada uma dessas etapas ilustradas, detalhando o seu objetivo, as ações tomadas e o produto resultante.

3.1 - TECNOLOGIAS E RECURSOS UTILIZADOS

O desenvolvimento de uma solução de BI requer obrigatoriamente a utilização de um conjunto de ferramentas que unidas formarão uma infraestrutura, de hardware e software, que dará respaldo e permitirá a execução das ações necessárias.

3.1.1 - Infraestrutura de *Hardware*

O ambiente computacional de *hardware* utilizado em todo o estudo, e sobre o qual roda a solução proposta, é apresentado na Figura 3.2. Nele é possível vermos que foram utilizados três computadores diferentes, para funções específicas, interligados em rede:

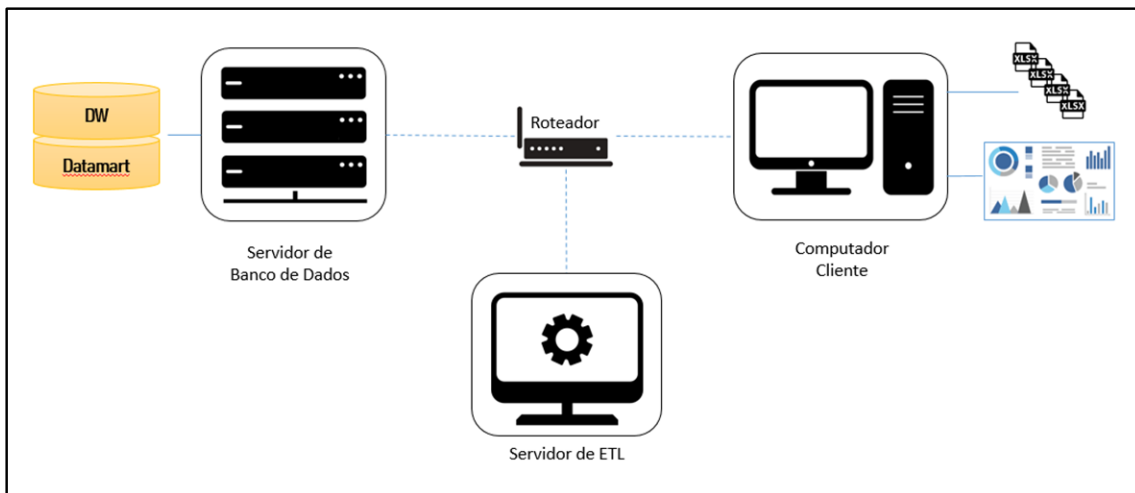


Figura 3.2 - Infraestrutura de hardware.

Descrevendo melhor os elementos da figura, temos:

- Servidor de banco de dados: responsável por hospedar o serviço de banco de dados. Neste equipamento estão localizadas as bases de dados de *data warehouse* e de *datamart* gerenciados pelo MySQL. O equipamento executa sistema operacional *Windows Professional Edition* e possui como configuração básica um processador Intel Core i7 64bits, 16GB de memória RAM e para armazenamento uma unidade de estado sólido (SSD) de 512GB.
- Servidor de ETL: responsável pelo serviço de ETL, interagindo com os serviços de banco de dados e de repositório de dados de origem. O equipamento executa sistema operacional *Windows Professional Edition* e possui como configuração básica um processador Intel Core i7 64 bits, 16GB de memória RAM e para armazenamento um disco rígido tradicional (HDD) de 1TB.
- Computador cliente: neste contexto, este equipamento ficou responsável pelas funções de repositório de dados de origem (que neste caso são arquivos em formato XLSX) e de execução da ferramenta de visualização das análises e *dashboards*. O equipamento executa sistema operacional *Windows Home Edition* possui como configuração básica um processador Intel Core i5 64bits, 8GB de memória RAM e para armazenamento um disco rígido tradicional (HDD) de 1TB.
- Roteador: ativo responsável pela interligação em rede dos demais equipamentos, permitindo a comunicação ponto-a-ponto entre eles. Foi utilizado um equipamento com tecnologia de velocidade Gigabit.

Embora este estudo tenha sido executado em ambiente de laboratório (desenvolvimento), a arquitetura adotada poderia ter sido utilizada em um ambiente

produtivo real, havendo uma possibilidade de ajuste: migrando o repositório de dados de origem do computador cliente, para um equipamento diferente e com essa responsabilidade específica. Logicamente, alguns contextos, podem requerer arquiteturas mais complexas e capacidades de armazenamento e processamento maiores. Neste caso, esta mostrou-se mais do que suficiente para os propósitos estabelecidos.

3.1.2 - Infraestrutura de *Software*

3.1.2.1 - MySQL

O MySQL foi utilizado da etapa “Criar *Data Warehouse*” até a etapa “Criar Análises e *Dashboards*”. Ele cumpriu a função de repositório de dados para as bases de dados de *data warehouse* e *datamart*.

A Figura 3.3 apresenta a página de *download* do *software* (<https://www.mysql.com/downloads/>), que oferece as versões *Enterprise Edition* (comercializada) e a *Community Edition*, utilizada neste trabalho. O MySQL está disponível para diversas plataformas: Windows, Linux (diversas distribuições), MacOS e outras.

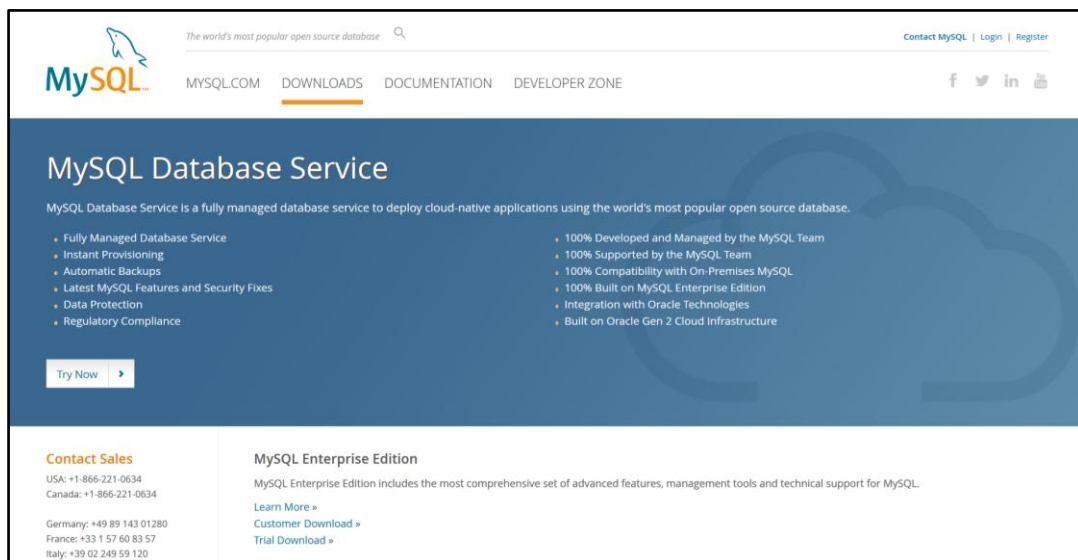


Figura 3.3 - Página de *download* do MySQL.
Fonte: MySQL (2020b).

Em ambiente de desenvolvimento para este trabalho, foram utilizadas as configurações padrões de instalação do servidor. Em ambiente produtivo, e com a

utilização continuada, podem ser executados ajustes finos objetivando ganho de performance do serviço, o que não é parte do escopo deste trabalho.

3.1.2.2 - HeidiSQL

Ele foi utilizado durante as etapas “Criar *data warehouse*” e “Criar *Datamart*” para a criação das bases de dados e execução dos scripts de criação dos objetos (tabelas, campos, índices, chaves primárias e estrangeiras etc.) dos modelos de dados de *data warehouse* e *datamart*.

A Figura 3.4 apresenta a página de *download* do *software* (<https://www.heidisql.com/download.php>), que possui apenas versão nativa para sistema operacional Windows.



Figura 3.4 - Página de *download* do HeidiSQL.
Fonte: HeidiSQL (2021).

Inicialmente foram realizados o download e a instalação do software, utilizando as configurações padrão sugeridas pelo instalador. Ao executar o HeidiSQL, a primeira ação foi configurar a conexão para o SGBD, que foi instalado na etapa anterior. Utilizamos o endereço IP do servidor de banco de dados e, por estarmos executando em ambiente de desenvolvimento, utilizamos o usuário e senha padrões de *root* (usuário principal, *master*). Em ambiente produtivo, deve-se utilizar um usuário específico, com privilégios limitados de acordo com a necessidade, e com acesso e permissões somente

para os objetos da solução de BI. A Figura 3.5 apresenta a tela de configuração da conexão.

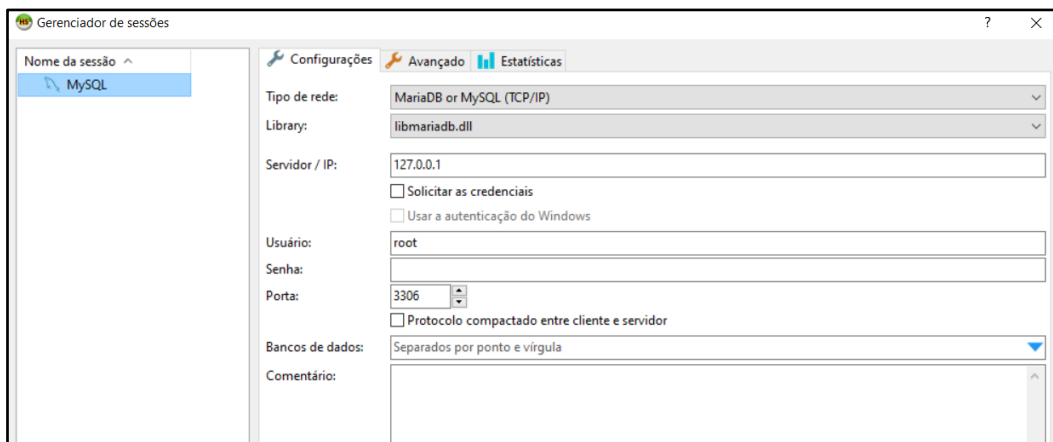


Figura 3.5 - Configuração de conexão ao SGBD do HeidiSQL.

Após configurada e iniciada, a ferramenta oferece uma interface simples e funcional, conforme é possível ver na Figura 3.6. Ao lado esquerdo, são exibidas hierarquicamente, em um componente gráfico do tipo árvore, as bases de dados existentes no servidor e as suas tabelas. Ao lado direito, encontram-se secções com os detalhes do objeto selecionado na estrutura do lado esquerdo e de execução de *scripts*.

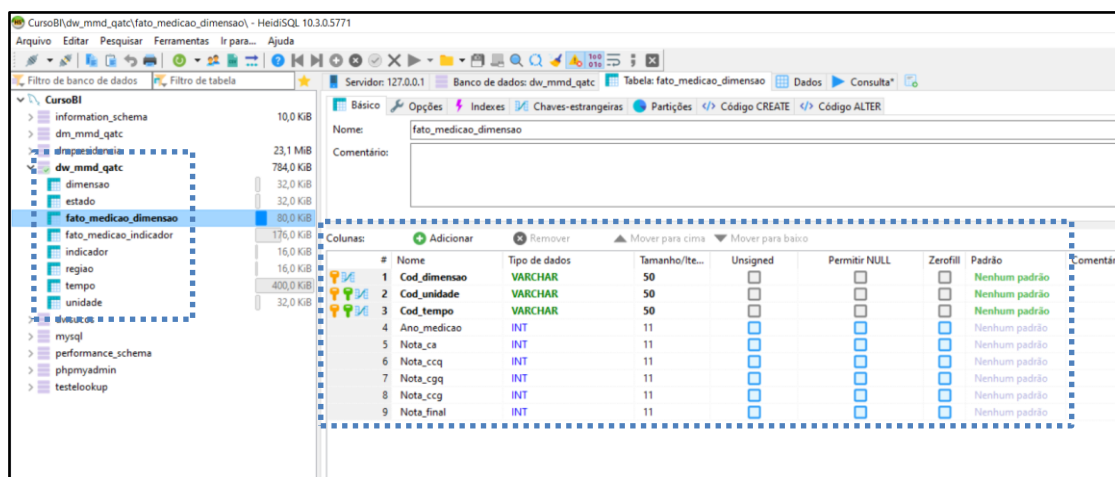


Figura 3.6 - Interface inicial do HeidiSQL.

3.1.2.3 - Linguagem SQL

A linguagem SQL foi utilizada durante as etapas “Criar data warehouse” e “Criar Datamart”. Apesar de não requerer a instalação e configuração de um software específico, pois é uma linguagem que o SGBD MySQL já oferece suporte nativamente, é importante

ser destacada, pois artefatos da solução foram gerados segundo sua sintaxe e suas regras. Os comandos para criação dos modelos de dados nas etapas citadas acima, gerados pelo *SQL Power Architect* a partir dos diagramas modelados nele, foram escritos nessa linguagem.

3.1.2.4 - *SQL Power Architect*

O *SQL Power Architect* foi utilizado nas etapas “Criar *Data Warehouse*” e “Criar *Datamart*” para a realização da modelagem dos dados e para a geração dos comandos SQL para criação dos objetos nas bases de dados.

A Figura 3.7 apresenta a página de *download* do *software* (http://www.bestofbi.com/page/architect_download_os), que oferece as versões *Enterprise Edition* e *Pro Edition*, que são versões comercializadas (pagas), e a *Community Edition* (gratuita), estando disponível para as plataformas Windows, Linux e MacOS. Utilizamos a edição *Community*, para sistema Operacional Windows, com as configurações padrão sugeridas pelo instalador do software.

Free Download: SQL Power Architect

SQL Power Architect Community Edition: FREE DOWNLOAD

SQL Power Architect Enterprise Edition: DOWNLOAD TRIAL (NOT FOR PRODUCTION USE)

SQL Power Architect Enterprise Edition: BUY NOW

Which Edition of SQL Power Architect do I need?

Features	Community Edition	Pro Edition Coming 2016	Enterprise Edition (Trial only)	Enterprise Edition (Full version)
Desktop (1-user) or Server (multi-user)	Desktop	Desktop	Server	Server
Data Modeling	✓	✓	✓	✓
Data Profiling	✓	✓	✓	✓
Compare, Reverse-Engineer, Forward-Engineer Data Models	✓	✓	✓	✓
OLAP Schema Modeling	✓	✓	✓	✓
Multi-user Data Model Collaboration			✓	✓
Full user-level & group-level security			✓	✓
Support for data domains & types		✓	✓	✓

Figura 3.7 - Página de *download* do Power Architect.
Fonte: ARCHITECT (2020).

3.1.2.5 - Linguagem de programação *Python*

Foi utilizada na etapa “Criar ETL Data Warehouse” para escrever o algoritmo que lê os dados das planilhas e persiste no banco de dados de *data warehouse*. *Python* está disponível para vários sistemas operacionais, tais como Windows, MacOS e Linux. Utilizamos a versão 3.7 do seu interpretador disponível no site de *download* (<https://www.python.org/downloads/>).

3.1.2.6 - Biblioteca Pandas

Foi utilizada na etapa “Criar ETL *Data Warehouse*” em conjunto com a linguagem de programação *Python*. Ela foi especificamente empregada na tarefa de ler e realizar o *parser* (interpretação e conversão) nos arquivos das planilhas em formato “.XLSX”. Para instalação desta biblioteca, foi utilizado o comando PIP, que é um sistema de gerenciamento de pacotes com a função de instalar e gerenciar pacotes / bibliotecas de software escritos em Python: “pip install pandas”.

3.1.2.7 - Pentaho Data Integration

Utilizamos neste estudo a versão *Community 9.0* para sistema operacional *Windows*. No entanto, o Pentaho Data Integration, parte da Suíte Pentaho, está disponível também para MacOS e Linux. A Figura 3.8 apresenta a página de *download* do *software* (<https://sourceforge.net/projects/pentaho/>).

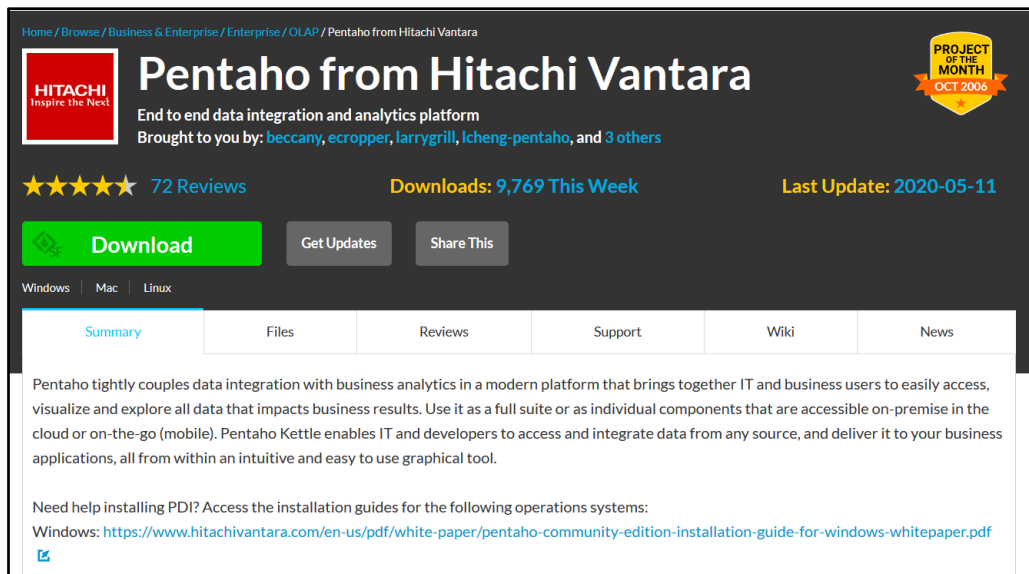


Figura 3.8 - Página de *download* da Suíte Pentaho.
Fonte: Pentaho (2021).

Foi empregado na etapa “Criar ETL *Datamart*” para a carga de dados da base de *data warehouse* para a base de *datamart*. A Figura 3.9 apresenta a interface inicial do aplicativo.

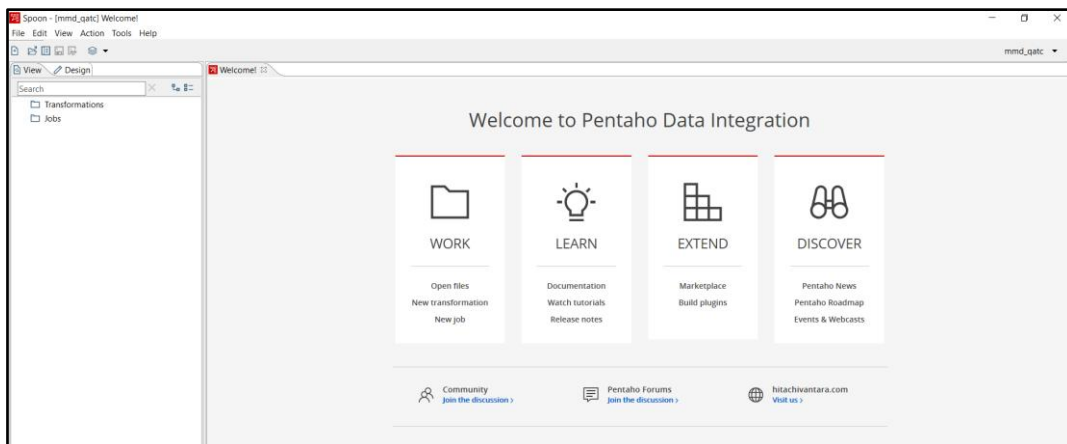


Figura 3.9 - Interface inicial do Pentaho Data Integration.

3.1.2.8 - Power BI Desktop

Utilizamos a versão *Desktop* desta ferramenta, disponível para sistema operacional Windows, que pode ser baixada gratuitamente no site da empresa (<https://powerbi.microsoft.com/pt-br/downloads/>) e atende as necessidades do projeto. Existem outras modalidades de licenciamento do *software* que oferecem recursos adicionais, mas exigem um custo para a aquisição. A instalação foi realizada com os

parâmetros padrões sugeridos pelo instalador do software. Foi empregado na etapa “Criar Análises e *Dashboards*” para a construção de tais artefatos.

3.2 - METODOLOGIA EXPERIMENTAL

Nesta secção, explicaremos com mais detalhes cada uma das etapas apresentadas anteriormente na Figura 3.1, no início deste capítulo.

3.2.1 - Identificar tecnologias a serem empregadas

Nesta etapa, buscou-se identificar tecnologias que pudessem colaborar com o objetivo da pesquisa, e tomando-se como base a premissa de que os aplicativos e linguagens de programação deveriam ser *open source* (código-fonte aberto) ou ao menos de uso livre, ou seja, não deveria haver a necessidade de aquisição de licenças de uso.

A secção 3.1.2 detalhou quais foram esses softwares e onde cada um foi aplicado.

3.2.2 - Identificar origem e descrever estrutura dos dados

Conhecer a origem dos dados, a sua natureza e a forma como eles estão estruturados é fundamental para a identificação e relacionamento de dimensões e indicadores, criação do *data warehouse* e criação do ETL do *data warehouse*. Nesta etapa, procurou-se entender onde estariam armazenados os dados (repositório), em quais formatos e que tipo de informações constariam.

O Manual de Procedimentos MMD-TC apresenta todos os indicadores trabalhados, incluindo suas dimensões e critérios, e fornece instruções de como avaliá-los e como utilizar tabelas de conversão para a pontuação dos mesmos. Cada instituição participante preenche a planilha base de avaliação dos TCs, documento disponível para download no site do QATC (<http://qatc.atricon.org.br/>), e a remete à Comissão de Garantia da Qualidade (Atricon).

A planilha está em formato “.XLSX” (formato de arquivo de planilhas do Microsoft Excel) e contém, dispostos nas linhas, os indicadores (agrupados por domínio) suas dimensões e os critérios de cada dimensão. Para cada critério, dispostos nas colunas, a Comissão de Avaliação do TC deve informar a unidade responsável pelo dado, se o TC atende aquele critério (sim ou não), evidências que embasem a avaliação e justificativas

e/ou comentários. A pontuação é atribuída para cada uma das dimensões e para cada indicador, de acordo com as regras estabelecidas.

Essa estrutura basicamente se repete mais três vezes na planilha, uma vez que após o preenchimento inicial, pela Comissão de Avaliação - CA do Tribunal de Contas da localidade (a comissão é definida pela administração local e composta por servidores do Tribunal local), verificações são realizadas pela Comissão de Controle da Qualidade – CCQ (a equipe também é definida e composta pelo próprio TC avaliado), Comissão de Garantia de Qualidade – CGQ (a equipe é composta por membros e colaboradores de outro TCs e é definida pela Atricon) e, eventualmente, em caso de recurso, uma última avaliação é feita pela Comissão de Coordenação Geral.

A Figura 3.10 apresenta graficamente este processo de trabalho:

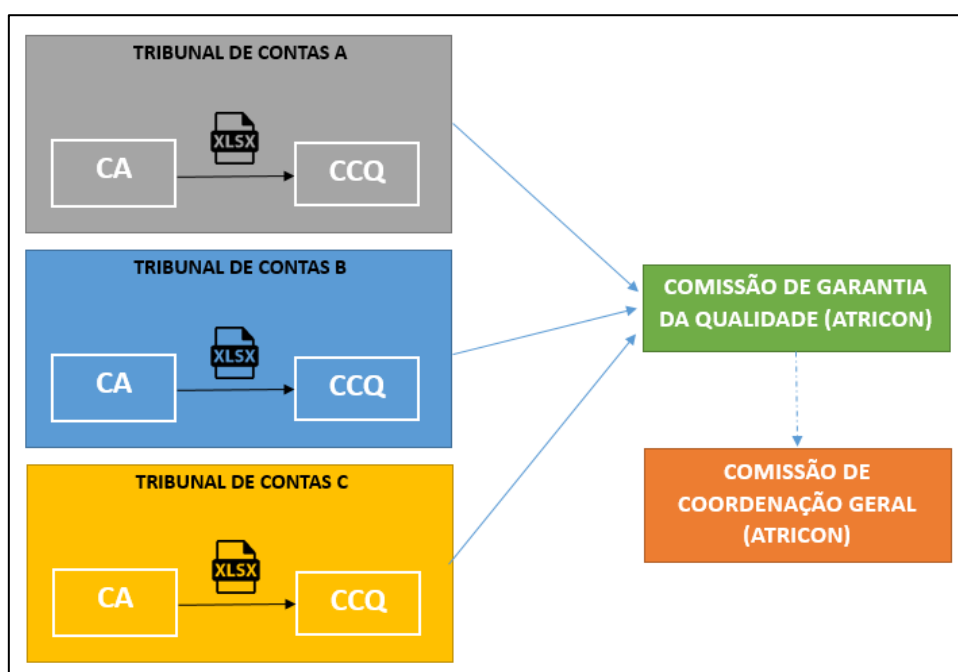


Figura 3.10 - Diagrama de processo de avaliação QATC.

3.2.3 - Identificar e relacionar Fatos e Dimensões

O conceito de fato, dentro de BI, remete a um acontecimento que queremos analisar e a dimensão remete ao modo como iremos analisar ou sob qual ótica queremos analisar o referido fato. Devemos fazer um mapeamento das expectativas gerenciais com as informações disponíveis para que possamos, ao fim, elaborar uma esquemática que represente a relação entre indicadores e dimensões.

Nas linhas da matriz, listamos os fatos que queremos analisar. Nas colunas, posicionamos as dimensões, como a informação deve ser vista pelo usuário do BI. Por fim, assinalamos na matriz as relações fato/dimensão que fazem sentido para o contexto estudado, conforme sumarizado na Tabela 3.1. Neste caso, todas as dimensões identificadas apresentam relacionamento lógico com os fatos a serem analisados.

Tabela 3.1 - Relação Fato x Dimensões.

Fato/Dimensão	Unidade	Região	Dimensão	Indicador	Domínio	Tempo
Medição indicador	x	x	x	x	x	x
Medição dimensão	x	x	x	x	x	x

3.2.4 - Gerar massa de dados de teste

Para a realização deste estudo, foi necessário gerar uma massa de dados de teste, uma vez que a Atricon não fornece os dados reais das unidades participantes do processo de avaliação. Para tanto, foram criadas cinquenta e seis (56) planilhas no mesmo formato de arquivo esperado, duas para cada unidade participante em potencial. Para a geração dos valores, foi utilizada a função do Microsoft Excel “=ALEATÓRIOENTRE(0; 4)”, que retorna um número aleatório inteiro entre os valores especificados. Um novo valor aleatório inteiro é retornado sempre que a planilha for calculada (pressionando a tecla F9), recurso este que foi utilizado para que cada planilha tivesse um conjunto de valores diferentes entre si.

3.2.5 - Criar *data warehouse*

Esta etapa consiste em construir o modelo físico de dados do *data warehouse* (DW). Para tanto, tomamos como base a planilha de avaliação dos TCs, identificada em etapa anterior. A Figura 3.11 apresenta o modelo resultante dessa tarefa, que contém tabelas, campos, chaves primárias e estrangeiras, com suas especificações de tipo de variável e de relacionamento.

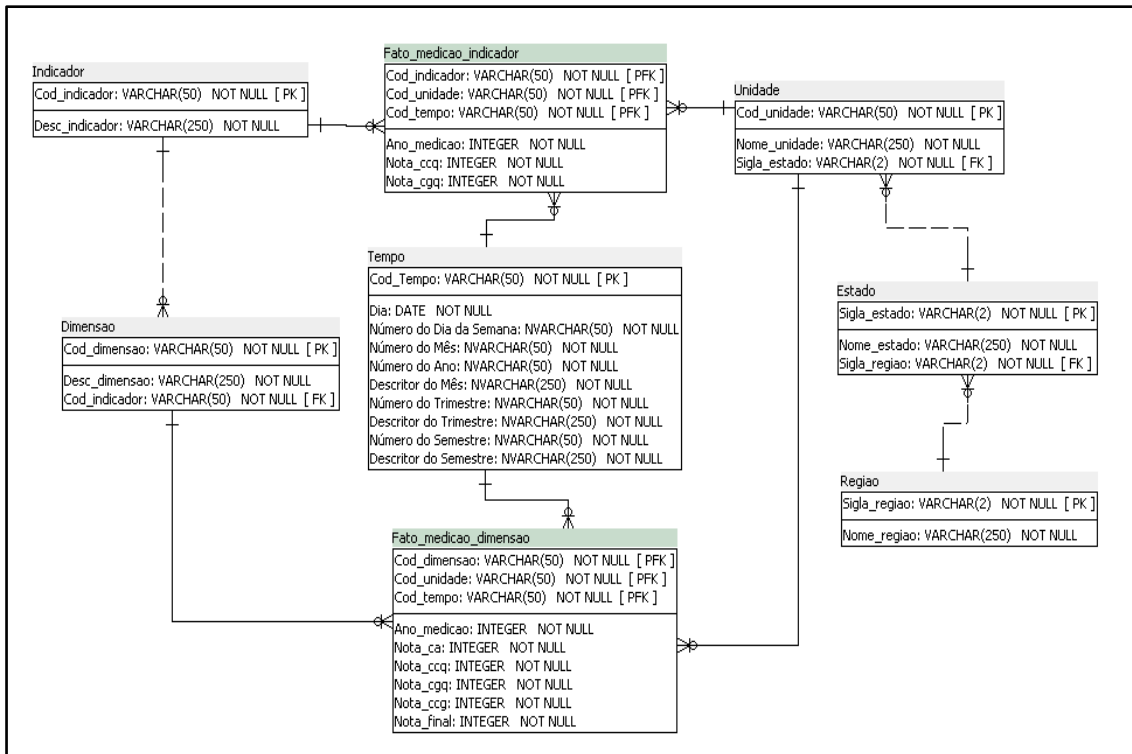


Figura 3.11 - Modelo de dados do DW.

Obeve-se como saída desta modelagem, um *script* de criação em linguagem SQL (disponibilizado na íntegra no Apêndice B), que foi executado no HeidiSQL. A Figura 3.12 apresenta um trecho do código sendo executado.

```

55
56
57 CREATE TABLE Fato_medicao_dimensao (
58   Cod_dimensao VARCHAR(50) NOT NULL,
59   Cod_unidade VARCHAR(50) NOT NULL,
60   Cod_tempo VARCHAR(50) NOT NULL,
61   Ano_medicao INT NOT NULL,
62   Nota_ca INT NOT NULL,
63   Nota_ccq INT NOT NULL,
64   Nota_cgq INT NOT NULL,
65   Nota_ccg INT NOT NULL,
66   Nota_final INT NOT NULL,
67   PRIMARY KEY (Cod_dimensao, Cod_unidade, Cod_tempo)
68 );
69
70
71 CREATE TABLE Fato_medicao_indicador (
72   Cod_indicador VARCHAR(50) NOT NULL,
73   Cod_unidade VARCHAR(50) NOT NULL,
74   Cod_tempo VARCHAR(50) NOT NULL,
75   Ano_medicao INT NOT NULL,
76   Nota_ccq INT NOT NULL,
77   Nota_cgq INT NOT NULL,
78   PRIMARY KEY (Cod_indicador, Cod_unidade, Cod_tempo)
79 );
80
81
82 ALTER TABLE Fato_medicao_indicador ADD CONSTRAINT indicador_fato_medicao_indicador_fk
83 FOREIGN KEY (Cod_indicador)
84 REFERENCES Indicador (Cod_indicador)
85 ON DELETE NO ACTION
86 ON UPDATE NO ACTION;
87
88 ALTER TABLE Dimensao ADD CONSTRAINT indicador_dimensao_fk

```

Figura 3.12 - Script SQL de criação do DW sendo executado no HeidiSQL.

Ao final, como resultado dessa operação, uma base estruturada estava pronta para receber a carga de dados, objetivo da próxima etapa. A Figura 3.13 apresenta, à esquerda, de azul, as tabelas resultantes do processamento, e à direita, os detalhes de uma delas, a tabela fato_medicao_dimensao.

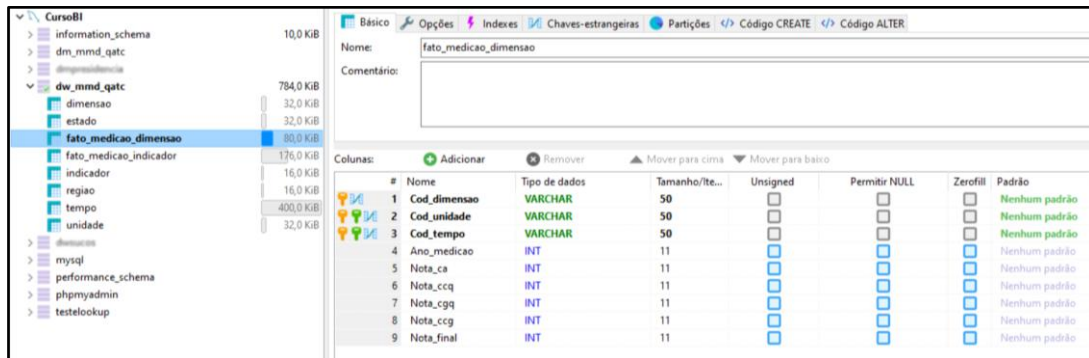


Figura 3.13 - Base de dados resultante da execução do script.

3.2.6 - Criar ETL data warehouse

Após a criação do modelo de dados do DW, o passo seguinte consistiu em fazer a extração dos dados da origem (arquivos “.XLSX”) e persisti-los na base de dados de DW. Foi desenvolvido um algoritmo com essa finalidade específica. O algoritmo lê as planilhas de avaliação dos TCs e, obedecendo a estrutura do arquivo e a modelagem do DW, faz uma conversão de formato e grava os dados na base de dados MySQL.

Na Figura 3.14, podemos ver a saída de texto resultante da execução do algoritmo escrito na linguagem Python, na parte inferior da figura.

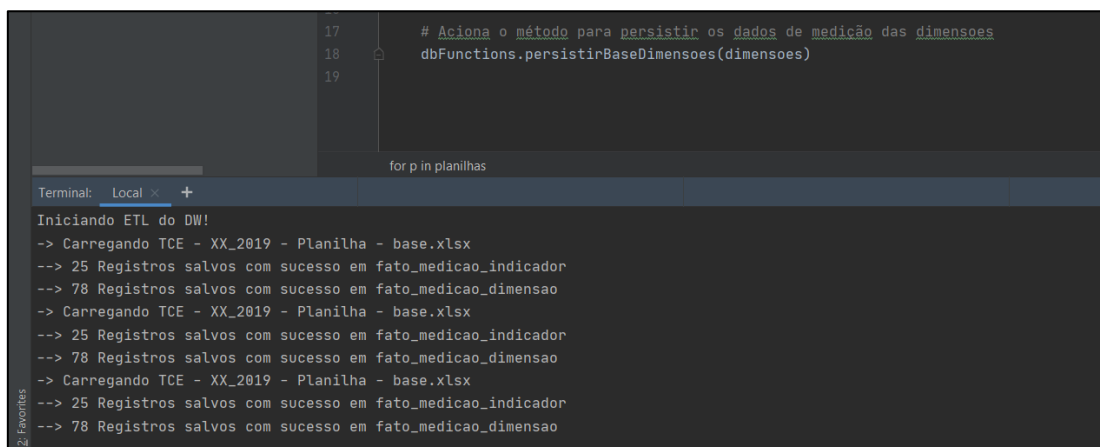


Figura 3.14 - Execução do algoritmo Python que realiza o ETL do DW.

3.2.7 - Criar *datamart*

De forma semelhante à etapa de criação do *data warehouse*, esta etapa consiste em construir um modelo físico de dados para o *datamart* (DM), base de dados que será utilizada diretamente pela ferramenta de *business intelligence*, e que difere em estilo de modelagem do DW.

Para tanto, fizemos uso da Tabela 3.1 elaborada em etapa anterior. O modelo, apresentado na Figura 3.15, contém tabelas, campos, chaves primárias e estrangeiras.

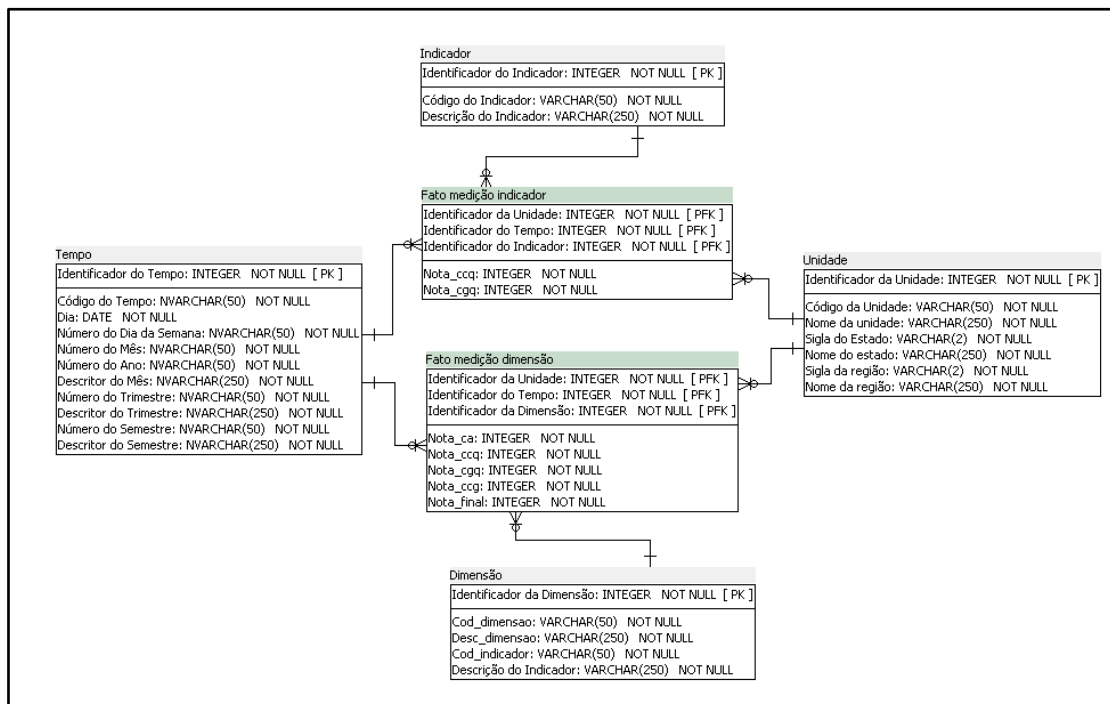


Figura 3.15 - Modelo de dados do *datamart*.

Como saída desta modelagem, obtivemos um *script* de criação em linguagem SQL (disponibilizado na íntegra no Apêndice C), que foi executado no HeidiSQL. A Figura 3.16 apresenta um trecho deste código sendo executado.


```

1 USE dm_mmd_qatc;
2
3
4 CREATE TABLE Dim_dimensao (
5     id_dimensao INT AUTO_INCREMENT NOT NULL,
6     Cod_dimensao VARCHAR(50) NOT NULL,
7     Desc_dimensao VARCHAR(250) NOT NULL,
8     Cod_indicador VARCHAR(50) NOT NULL,
9     Desc_indicador VARCHAR(250) NOT NULL,
10    PRIMARY KEY (id_dimensao)
11);
12
13
14 CREATE TABLE Dim_indicador (
15     id_indicador INT AUTO_INCREMENT NOT NULL,
16     cod_indicador VARCHAR(50) NOT NULL,
17     desc_indicador VARCHAR(250) NOT NULL,
18    PRIMARY KEY (id_indicador)
19);
20
21
22 CREATE TABLE Dim_unidade (
23     id_unidade INT AUTO_INCREMENT NOT NULL,
24     Cod_unidade VARCHAR(50) NOT NULL,
25     Nome_unidade VARCHAR(250) NOT NULL,
26     Sigla_estado VARCHAR(2) NOT NULL,
27     Nome_estado VARCHAR(250) NOT NULL,
28     Sigla_regiao VARCHAR(2) NOT NULL,
29     Nome_regiao VARCHAR(250) NOT NULL,
30    PRIMARY KEY (id_unidade)
31);
32
33
34 CREATE TABLE Dim_Tempo (
35     id_tempo INT AUTO_INCREMENT NOT NULL,
36     Cod_Tempo NVARCHAR(50) NOT NULL,
37     Data DATE NOT NULL,
38     Numero_Dia_Semana NVARCHAR(50) NOT NULL,
39     Numero_Mes NVARCHAR(50) NOT NULL,
40     Numero_Semestre NVARCHAR(50) NOT NULL,
41     Nome_Semestre NVARCHAR(50) NOT NULL,
42     Nome_Mes NVARCHAR(50) NOT NULL,
43     Nome_Dia_Semana NVARCHAR(50) NOT NULL,
44     Nome_estado NVARCHAR(250) NOT NULL,
45     Nome_regiao NVARCHAR(250) NOT NULL,
46     Nome_unidade NVARCHAR(250) NOT NULL,
47     Desc_dimensao NVARCHAR(250) NOT NULL,
48     Desc_indicador NVARCHAR(250) NOT NULL,
49     Cod_dimensao NVARCHAR(50) NOT NULL,
50     Cod_indicador NVARCHAR(50) NOT NULL,
51     Cod_unidade NVARCHAR(50) NOT NULL,
52     Nome_unidade NVARCHAR(250) NOT NULL,
53     Nome_estado NVARCHAR(250) NOT NULL,
54     Nome_regiao NVARCHAR(250) NOT NULL,
55     Nome_regiao NVARCHAR(250) NOT NULL,
56     Nome_regiao NVARCHAR(250) NOT NULL,
57     Nome_regiao NVARCHAR(250) NOT NULL,
58     Nome_regiao NVARCHAR(250) NOT NULL,
59     Nome_regiao NVARCHAR(250) NOT NULL,
60     Nome_regiao NVARCHAR(250) NOT NULL,
61     Nome_regiao NVARCHAR(250) NOT NULL,
62     Nome_regiao NVARCHAR(250) NOT NULL,
63     Nome_regiao NVARCHAR(250) NOT NULL,
64     Nome_regiao NVARCHAR(250) NOT NULL,
65     Nome_regiao NVARCHAR(250) NOT NULL,
66     Nome_regiao NVARCHAR(250) NOT NULL,
67     Nome_regiao NVARCHAR(250) NOT NULL,
68     Nome_regiao NVARCHAR(250) NOT NULL,
69     Nome_regiao NVARCHAR(250) NOT NULL,
70     Nome_regiao NVARCHAR(250) NOT NULL,
71     Nome_regiao NVARCHAR(250) NOT NULL,
72     Nome_regiao NVARCHAR(250) NOT NULL,
73     Nome_regiao NVARCHAR(250) NOT NULL,
74     Nome_regiao NVARCHAR(250) NOT NULL,
75     Nome_regiao NVARCHAR(250) NOT NULL,
76     Nome_regiao NVARCHAR(250) NOT NULL,
77     Nome_regiao NVARCHAR(250) NOT NULL,
78     Nome_regiao NVARCHAR(250) NOT NULL,
79     Nome_regiao NVARCHAR(250) NOT NULL,
80     Nome_regiao NVARCHAR(250) NOT NULL,
81     Nome_regiao NVARCHAR(250) NOT NULL,
82     Nome_regiao NVARCHAR(250) NOT NULL,
83     Nome_regiao NVARCHAR(250) NOT NULL,
84     Nome_regiao NVARCHAR(250) NOT NULL,
85     Nome_regiao NVARCHAR(250) NOT NULL,
86     Nome_regiao NVARCHAR(250) NOT NULL,
87     Nome_regiao NVARCHAR(250) NOT NULL,
88     Nome_regiao NVARCHAR(250) NOT NULL,
89     Nome_regiao NVARCHAR(250) NOT NULL,
90     Nome_regiao NVARCHAR(250) NOT NULL,
91     Nome_regiao NVARCHAR(250) NOT NULL,
92     Nome_regiao NVARCHAR(250) NOT NULL,
93     Nome_regiao NVARCHAR(250) NOT NULL,
94     Nome_regiao NVARCHAR(250) NOT NULL,
95     Nome_regiao NVARCHAR(250) NOT NULL,
96     Nome_regiao NVARCHAR(250) NOT NULL,
97     Nome_regiao NVARCHAR(250) NOT NULL,
98     Nome_regiao NVARCHAR(250) NOT NULL,
99     Nome_regiao NVARCHAR(250) NOT NULL,
100    Nome_regiao NVARCHAR(250) NOT NULL,

```

Figura 3.16 - Script SQL de criação do *datamart* sendo executado no HeidiSQL.

Ao final, como resultado dessa operação, uma base estruturada estava pronta para receber a carga de dados oriundos do DW, objetivo da próxima etapa. A Figura 3.17 apresenta, à esquerda, de azul, as tabelas resultantes do processamento, e à direita, os detalhes da *dim_tempo*, tabela que não existia na modelagem do DW e passou a existir no *datamart*.

#	Nome	Tipo de dados	Tamanho/Re...	Unsigned	Permitir NULL	Zerofill	Padrão
1	id_tempo	INT	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	AUTO_INCREME...
2	Cod_Tempo	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
3	Data	DATE		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
4	Numero_Dia_Semana	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
5	Numero_Mes	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
6	Numero_Ano	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
7	Nome_Mes	VARCHAR	250	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
8	Numero_Trimestre	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
9	Nome_Trimestre	VARCHAR	250	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
10	Numero_Semestre	VARCHAR	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão
11	Nome_Semestre	VARCHAR	250	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Nenhum padrão

Figura 3.17 - Base de dados resultante da execução do *script* de criação do *datamart*.

3.2.8 - Criar ETL *datamart*

Após a criação do modelo de dados do DM, o passo seguinte consistiu em fazer a extração dos dados da origem (base de DW) e persisti-los na base de dados de DM. Foi utilizada uma ferramenta específica de ETL para essa etapa, o Pentaho Data Integration (PDI). A Figura 3.18 exibe um exemplo de processo de carga mapeado na ferramenta, e no capítulo 4, apresentamos os processos reais de carga resultantes dessa etapa.

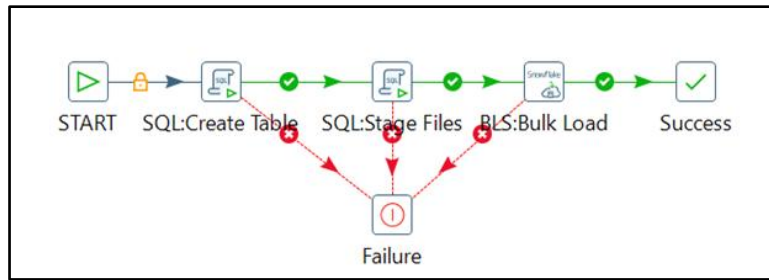


Figura 3.18 - Exemplo de processo de carga mapeado no PDI.
 Fonte: DAS (2021).

Cada item em um processo no Pentaho Data Integration equivale a uma operação, que pode ser de leitura, transformação/manipulação ou escrita de dados. As origens e destinos desses dados podem ser os mais diversos, assim como as transformações, que incluem cálculos matemáticos, operações com cadeias de caracteres, filtros, ordenações, junção de dados etc. Cada operação lê ou manipula dados e os repassa para a operação seguinte, até o final do processo, onde essas informações são persistidas (gravadas) na base de dados.

Na Figura 3.19, é possível ver os detalhes da configuração de uma operação do tipo “insert/update”, que realiza a gravação dos dados recebidos das operações anteriores na tabela “dim_unidade”, utilizando a conexão “Datamart MMD QATC” (que aponta para a base de dados do *datamart*, criada na secção 3.2.7 deste trabalho).

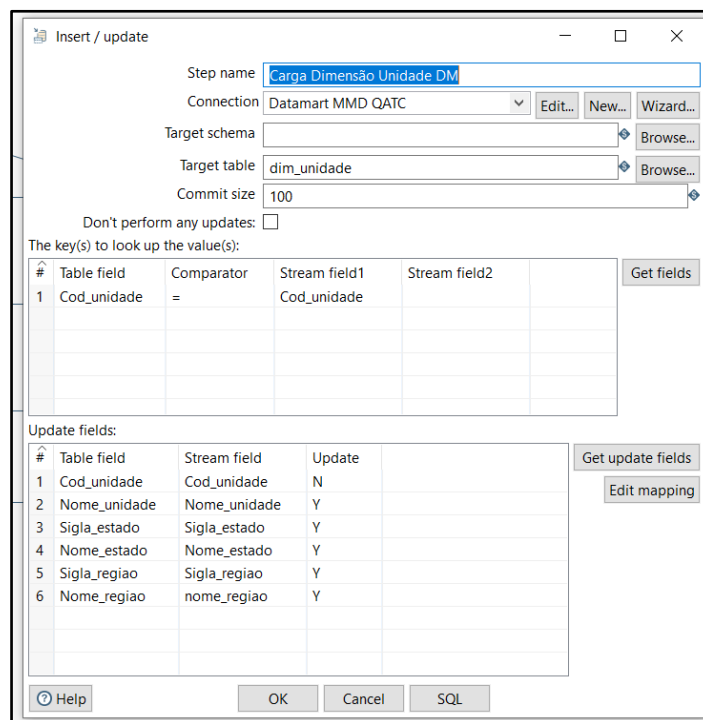


Figura 3.19 - Detalhe da operação que grava dados da dimensão unidade no *datamart*.

Ao final, após toda a rotina de transporte de dados para o *Datamart* construída, os processos foram executados. As etapas bem sucedidas recebem uma marcação verde e as que eventualmente tiverem algum tipo de problema exibem uma marcação em vermelho. Com base nisso, na Figura 3.20, é possível perceber que todas as etapas foram executadas sem intercorrências. Na parte inferior, vemos a seção “*Execution Results*”, onde o *software* apresenta informações detalhadas da execução de cada uma das etapas. Esse nível de detalhamento pode ser configurado na ferramenta.

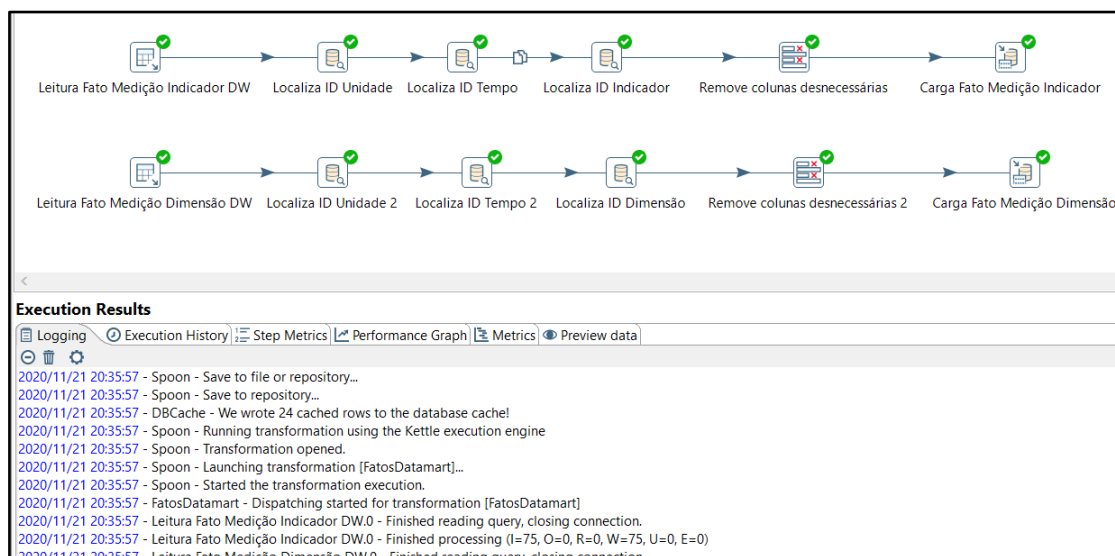


Figura 3.20 - Execução da rotina de carga das tabelas de fato do *datamart*.

Essa rotina deve ser acionada todas as vezes que houver atualização de dados na base de *Data Warehouse*.

3.2.9 - Criar análises e *dashboards*

Nesta etapa, foram desenvolvidas análises que, uma vez concluídas, foram organizadas em painéis de visualização (*dashboards*), de forma que os resultados pretendidos na etapa da construção da matriz de “fato x dimensão”, mesmo que de forma incipiente (e sem a riqueza de detalhes atingidas ao final) mas nos fornecendo uma base sólida de trabalho, fossem materializados.

Para o processo de elaboração das análises e *dashboards* foi utilizada a ferramenta Power BI Desktop e a base de dados de *datamart* como fonte de dados. Durante o processo de construção das análises, foi identificada a necessidade de criação de algumas medidas e colunas calculadas, que são campos cujos valores são determinados com base

em cálculos e operações lógicas, podendo se adaptar ao contexto dos filtros de dados aplicados na construção da análise.

No nosso caso, foram criadas medidas para cálculo de médias e acumulados das notas e classificação de nível, baseado na escala de pontuação do MMD-TC. Tais medidas foram utilizadas em várias análises que estão distribuídas nos diversos *dashboards* produzidos. A Figura 3.21 apresenta o momento da criação de uma dessas medidas no ambiente do *Power BI*:

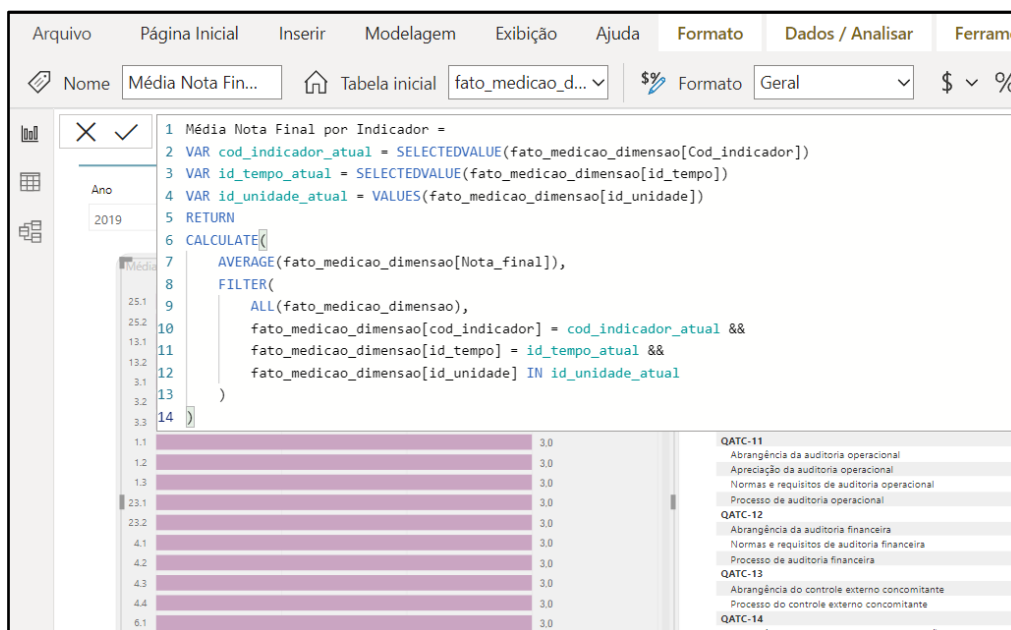


Figura 3.21 - Criação da medida “Média Nota Final por Indicador”.

Ao final desta etapa, obtivemos diferentes representações visuais dos dados, que são apresentadas de forma mais detalhada a seguir, no capítulo 4.

CAPÍTULO 4

RESULTADOS E DISCUSSÃO

Nessa etapa são apresentados a arquitetura da solução, artefatos de implementação e o produto em si resultante do desenvolvimento da solução tecnológica proposta e é explicado, também, o funcionamento com maiores detalhes.

Conforme demonstrado na Figura 3.10 - Diagrama de processo de avaliação QATC, exibida no capítulo anterior, podemos observar que os dados de cada unidade estão contidos nos arquivos .XLSX, passam por avaliações iniciais em diferentes níveis (conforme descrito na secção 3.2.2) até o momento em que ficam disponíveis para análises mais detalhadas pela Comissão de Coordenação Geral da Atricon.

Acreditamos que essa maneira seja suficiente para a realização de análises isoladas de cada Unidade naquele dado momento de tempo em que a avaliação ocorreu, mas não seja prática para a realização de análises em conjunto e/ou comparativas considerando os outros Tribunais, e nem mesmo uma análise evolutiva de uma mesma instituição ao longo do tempo.

Este trabalho pretendeu atuar exatamente nessa fase final do processo de avaliação. Os dados isolados e divididos entre diversos arquivos remetidos são compilados e processados. A Figura 4.1 apresenta o diagrama com os componentes da arquitetura da solução proposta:

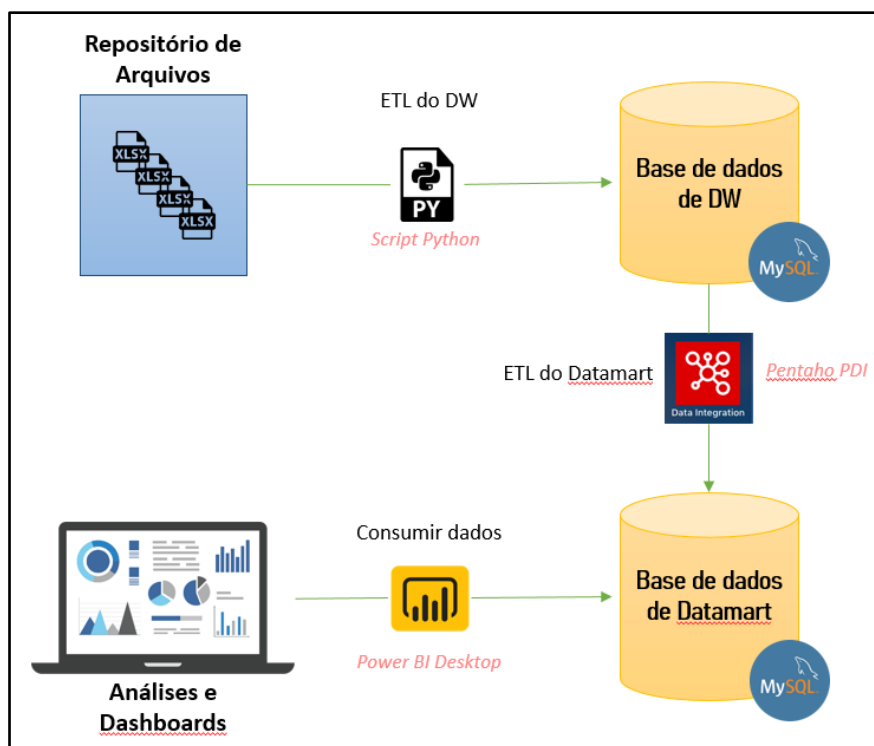


Figura 4.1 - Diagrama da arquitetura da solução proposta.

Nele podemos observar a origem dos nossos dados, que reside em um repositório em sistema de arquivos, estando organizados em arquivos no formato .XLSX e servindo de entrada para o processo “ETL do DW”. Sua estrutura encontra-se melhor descrita na secção “3.2.3 - Identificar origem e descrever estrutura dos dados”. A coleta desses arquivos continua ocorrendo por e-mail, conforme o procedimento estabelecido.

Em seguida, um script para extração dos dados na origem e inserção na base de dados de destino é acionado por meio da execução de uma linha de comando. Essa etapa ocorre todas as vezes que se pretender processar novos arquivos recebidos. A Figura 4.2 apresenta o fluxograma básico de funcionamento do programa escrito na linguagem Python e que implementa o processo “ETL do DW”.

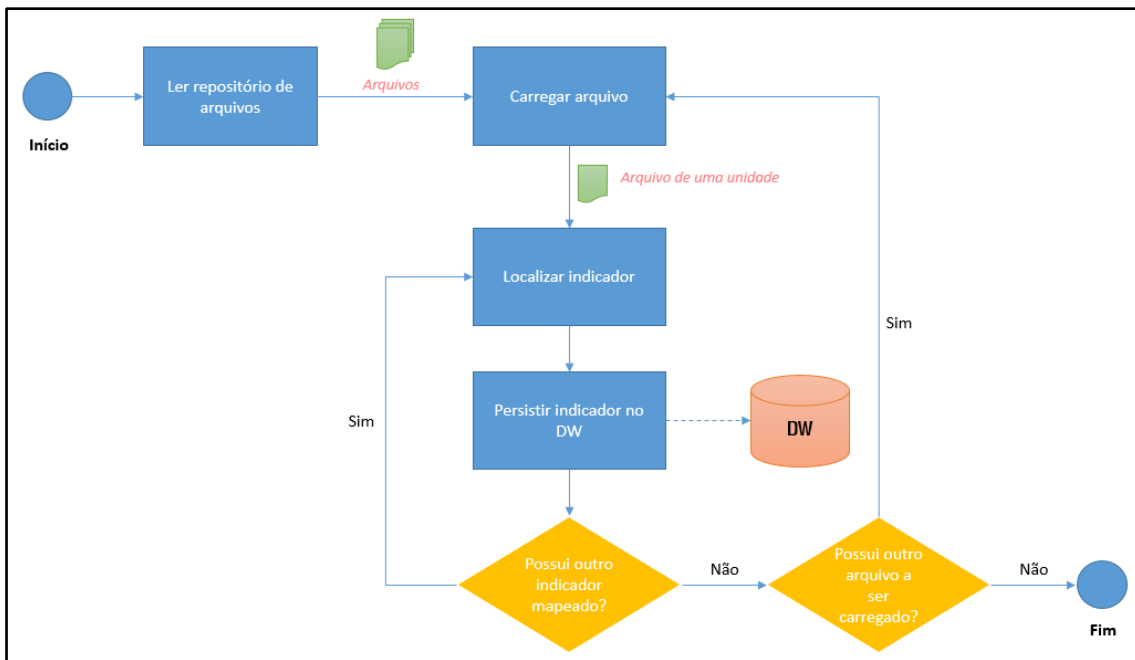


Figura 4.2 - Fluxograma do algoritmo em Python de ETL do DW.

Como é possível observar, o fluxo inicia com o algoritmo lendo o repositório de dados que contém os arquivos a serem processados. A partir desse ponto, ele entra em *loop* no qual cada arquivo identificado é processado, e que dura enquanto houver arquivos a serem carregados. No Apêndice D, é apresentado o trecho do código fonte correspondente ao apresentado na Figura 4.2.

Esse processamento consiste na localização de cada um dos indicadores mapeados (de acordo com a estrutura do arquivo, previamente identificada), captura do seu respectivo valor e a consequente persistência dos dados na base de dados de DW. Esse processo perdura enquanto houver indicadores a serem processados no arquivo.

Caso futuramente haja mudança na especificação dos indicadores, isso poderá afetar a estrutura do arquivo, que requererá um novo mapeamento e consequente ajuste no processo “Localizar indicador”, mas permanecendo as demais ações potencialmente inalteradas. E, caso haja mudança na forma de armazenamento, como por exemplo a adoção de um sistema para coleta desses dados, esta etapa de ETL do *data warehouse* deve ser revista para se adequar à nova origem e novo formato dos dados.

Após a extração dos dados dos arquivos e a sua disponibilização em base de dados, realizamos o ETL do *Datamart*. Utilizando o Pentaho Data Integration (PDI), implementamos sete (7) processos de carga de dados que tem como origem a base de DW e como destino a base de *Datamart*. A Figura 4.3 exibe um diagrama conceitual dos processos de carga modelados na ferramenta e as bases de dados acessadas.

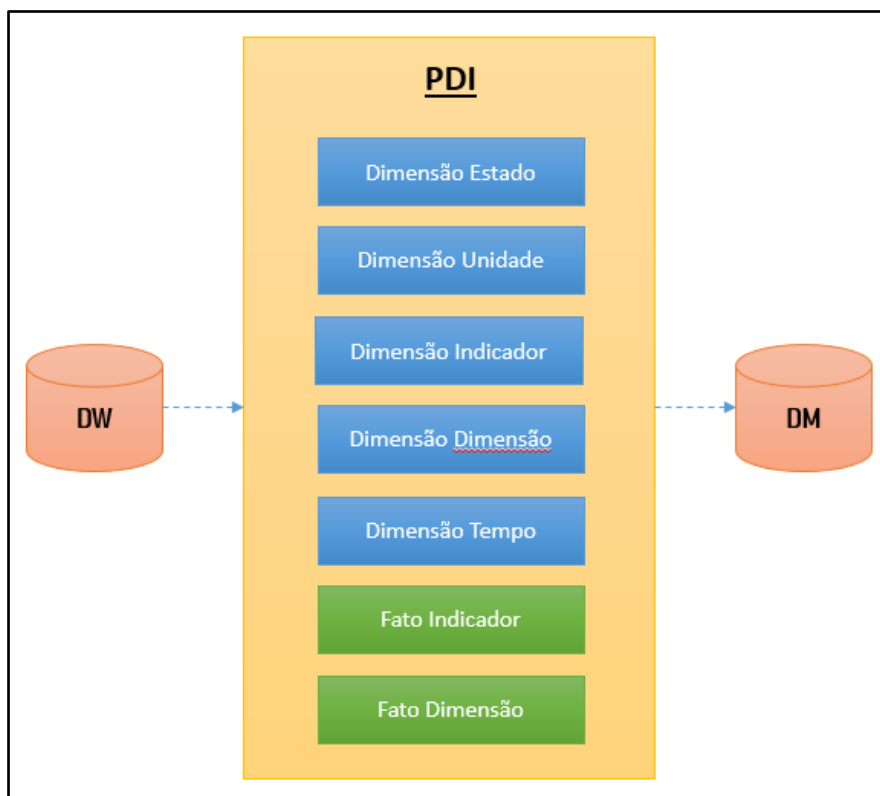


Figura 4.3 - Processos de carga modelados no Pentaho Data Integration.

Cada um desses processos sofreu uma modelagem específica dentro da ferramenta em questão. Foram criados processos de carga das dimensões, exibidos nas Figuras 4.4 e 4.5, e processos de carga dos fatos, exibidos na Figura 4.6. Estes processos de carga devem ser executados todas as vezes que a base de dados de DW sofrer alteração por ocasião do processamento de novos arquivos.

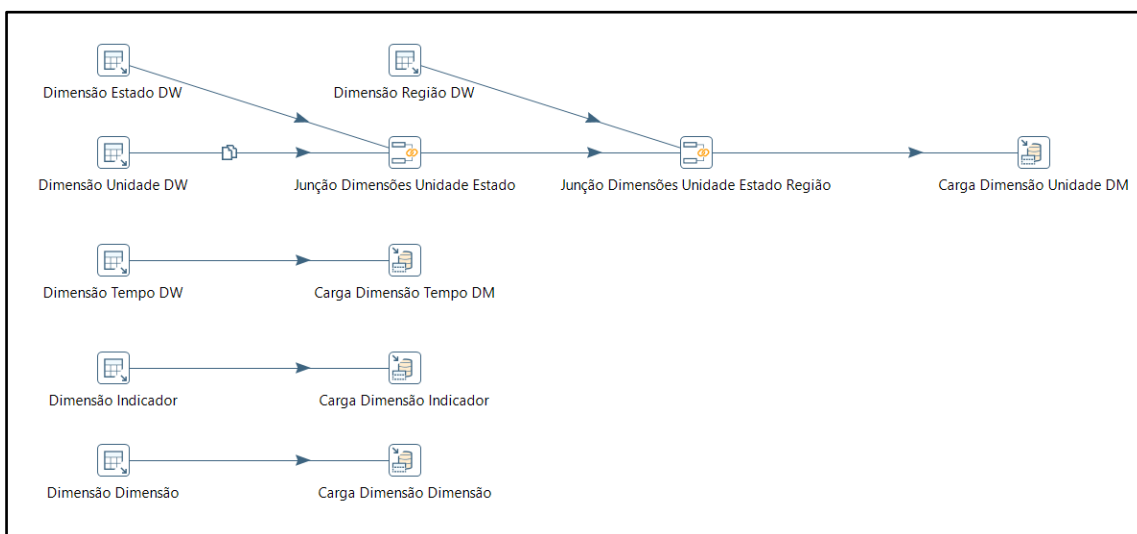


Figura 4.4 - Modelagem dos processos de carga das dimensões em geral.

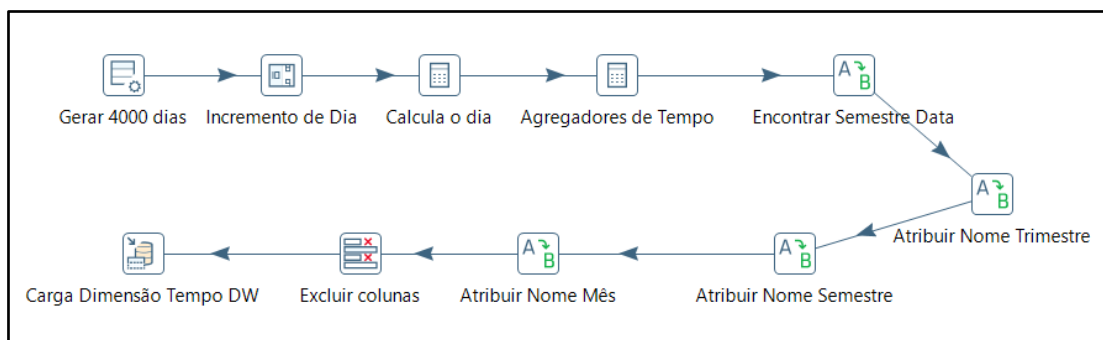


Figura 4.5 - Modelagem do processo de carga da dimensão de tempo.

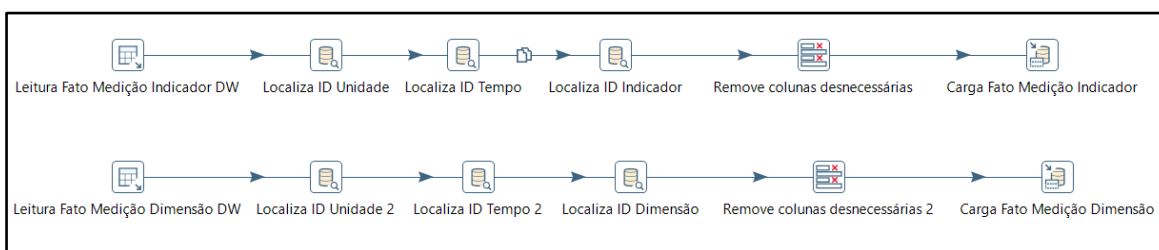


Figura 4.6 - Modelagem do processo de carga dos fatos.

O processo de ETL do *Datamart* encerrou as etapas estruturais da arquitetura proposta, nas quais os dados são processados para que estejam disponíveis de forma prática para o consumo nas análises e *dashboards* implementados no Power BI Desktop.

Nessa etapa de consumo dos dados pretendeu-se disponibilizar visões que permitissem observar, dentre outros:

- Os valores de cada um dos indicadores por unidade
- Visão evolutiva das pontuações dos indicadores por unidade.
- Os valores de cada uma das dimensões dos indicadores por unidade
- Visão evolutiva das pontuações das dimensões por unidade.
- O valor médio dos indicadores por região.
- O valor médio das dimensões dos indicadores por região.

Como resultado final obtivemos diversas visões organizadas em nove (9) *dashboards* diferentes que são apresentados nas secções a seguir. É importante salientar que os valores que aparecem expressos nos painéis são dados de teste, uma vez que a ATRICON não fornece os dados reais.

4.1 - DASHBOARD “UNIDADES, INDICADORES E DIMENSÕES”

Este painel apresenta o conjunto de Unidades, Tribunais de Contas Estaduais e Municipais existentes no Brasil, em duas visões diferentes: uma em formato de tabela, organizada por Região e Estado; e outra em formato de gráfico de pizza, onde é possível observar de forma intuitiva a distribuição quantitativa dessas unidades nas diferentes regiões do País.

Os indicadores e suas respectivas dimensões são apresentados em formato de tabela e tem os seus totais apresentados em formato de “cartões”. É possível ver o aspecto final do *dashboard* na Figura 4.7.

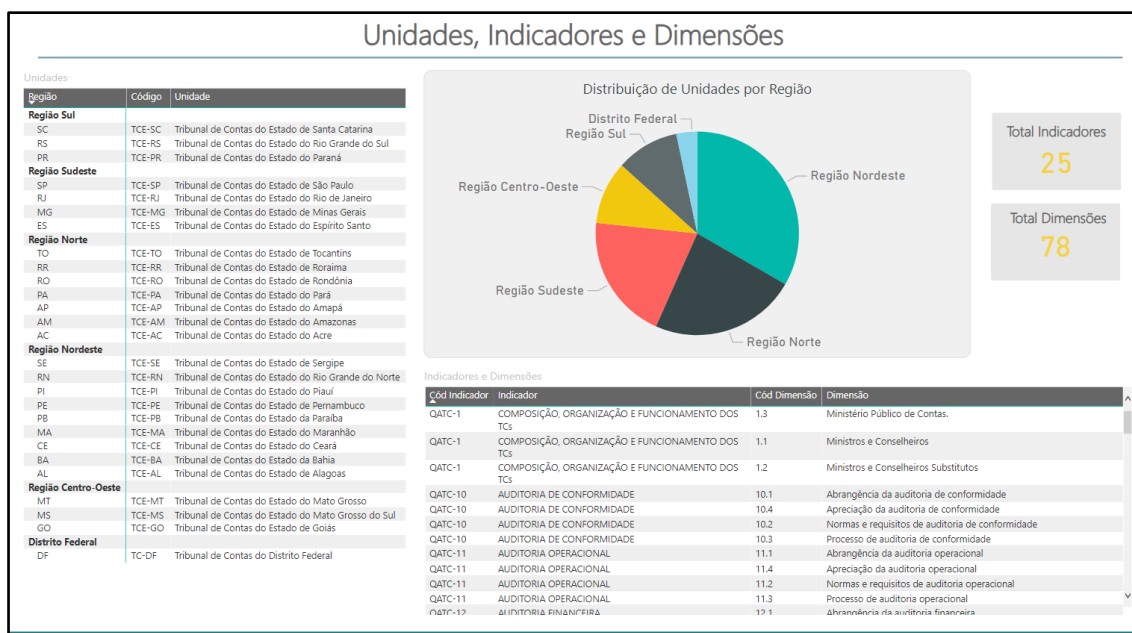


Figura 4.7 - Dashboard de “Unidades, Indicadores e Dimensões”.

4.2 - DASHBOARD “ENTREGAS”

Este painel apresenta diferentes visões das entregas dos questionários processados, permitindo inclusive um acompanhamento concomitante. Inicialmente, deve ser selecionado um ano para análise por meio do componente de segmentação de dados apresentado no lado esquerdo superior. Os dados de processamento dos questionários são apresentados com base nesse filtro.

Em uma tabela ao lado esquerdo, são apresentadas as Unidades que tiveram o questionário processado, juntamente com a data na qual o fato ocorreu e a pontuação total obtida com base nas notas dos indicadores.

Ao centro, em um gráfico de rosca, é possível observar a distribuição por região dessas entregas até o momento da consulta. E ainda, logo abaixo, a sua distribuição ao longo do tempo em um gráfico de linha. Finalmente, à direita, um visual de “cartão” apresenta o total de questionários processados no momento da consulta. É possível ver o aspecto final do *dashboard* na Figura 4.8.

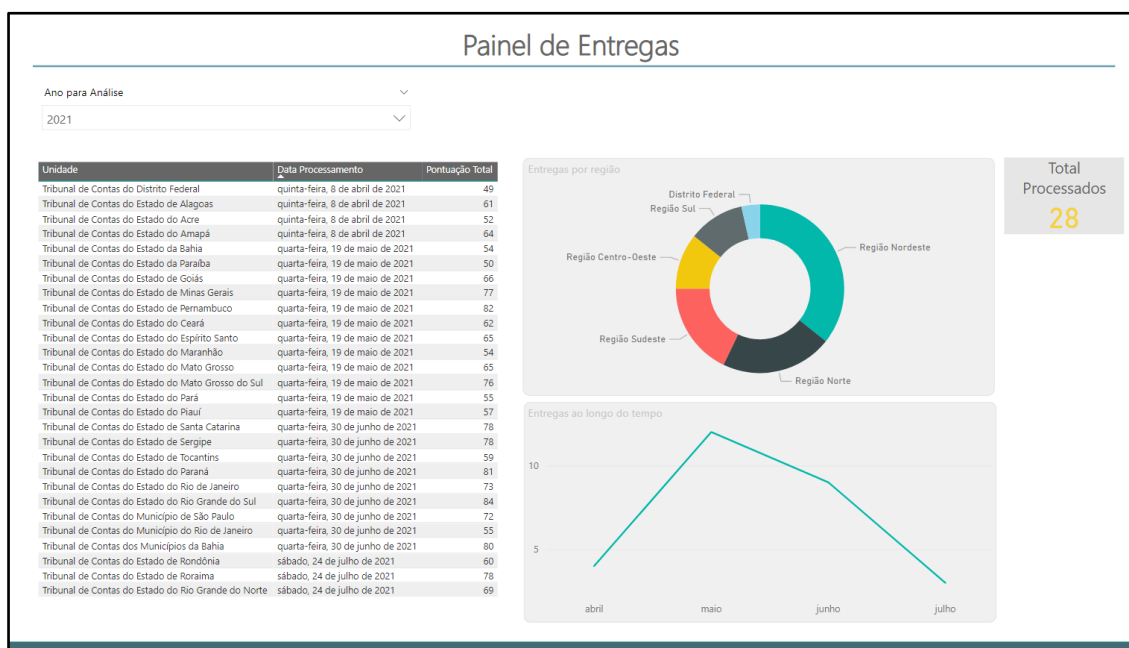


Figura 4.8 - *Dashboard* de “Entregas”.

4.3 - *DASHBOARD* “INDICADORES POR UNIDADE”

Este painel apresenta diferentes visões das notas dos indicadores de uma determinada unidade. Inicialmente, deve ser selecionado uma unidade e um ano para análise por meio dos componentes de segmentação de dados apresentados no lado esquerdo superior. Os dados de indicadores da unidade são apresentados com base nesse filtro.

Em uma tabela ao lado esquerdo, são apresentados os códigos dos indicadores, as notas atribuídas pelo CCQ (comissão da unidade correspondente) e pelo CGQ (que representa a nota definitiva para aquele item, atribuída pela comissão da Atricon). Ao centro, em um gráfico com barras que indicam o “tamanho” da nota para cada um dos

indicadores, é possível observar se existe uniformidade entre os indicadores ou se existe uma discrepância que possa indicar algum desequilíbrio. As cores das barras se adaptam de acordo com a nota que representam.

Ao lado direito, outras quatro visões complementam o painel: dois gráficos estilo velocímetro mostrando o total de pontos obtidos com base na avaliação do CCQ e do CGQ; um cartão indicando quantas avaliações discrepantes ocorreram; e outro cartão demonstrando a média de desvio das notas (diferença entre a nota atribuída inicialmente e a nota final) em função dos indicadores. É possível ver o aspecto final do *dashboard* na Figura 4.9.

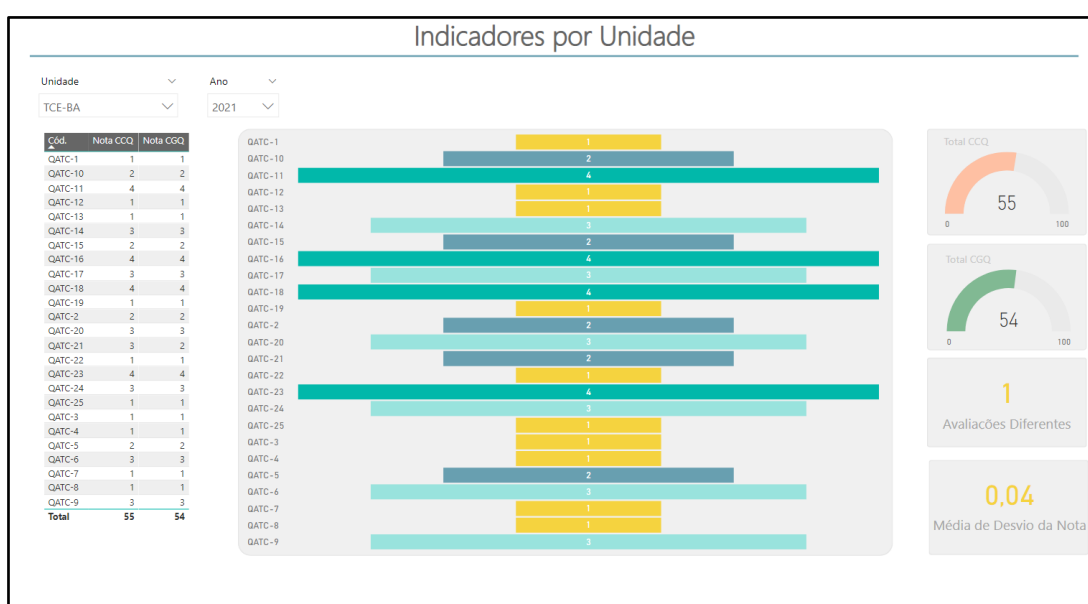


Figura 4.9 - *Dashboard* de “Indicadores por Unidade”.

4.4 - *DASHBOARD* “EVOLUÇÃO DOS INDICADORES POR UNIDADE”

Este painel apresenta os valores finais dos indicadores de uma determinada unidade ao longo do tempo. Inicialmente, deve ser selecionado uma unidade por meio do componente de segmentação de dados apresentado no lado esquerdo superior. Os dados de indicadores da unidade são apresentados com base nesse filtro.

Em uma tabela ao lado esquerdo, são apresentados os códigos dos indicadores e as notas atribuídas para cada um dos anos processados. Ao centro, um gráfico de barras horizontais realiza uma comparação visual dessas notas entre os diferentes anos para cada indicador. É possível ver o aspecto final do *dashboard* na Figura 4.10.

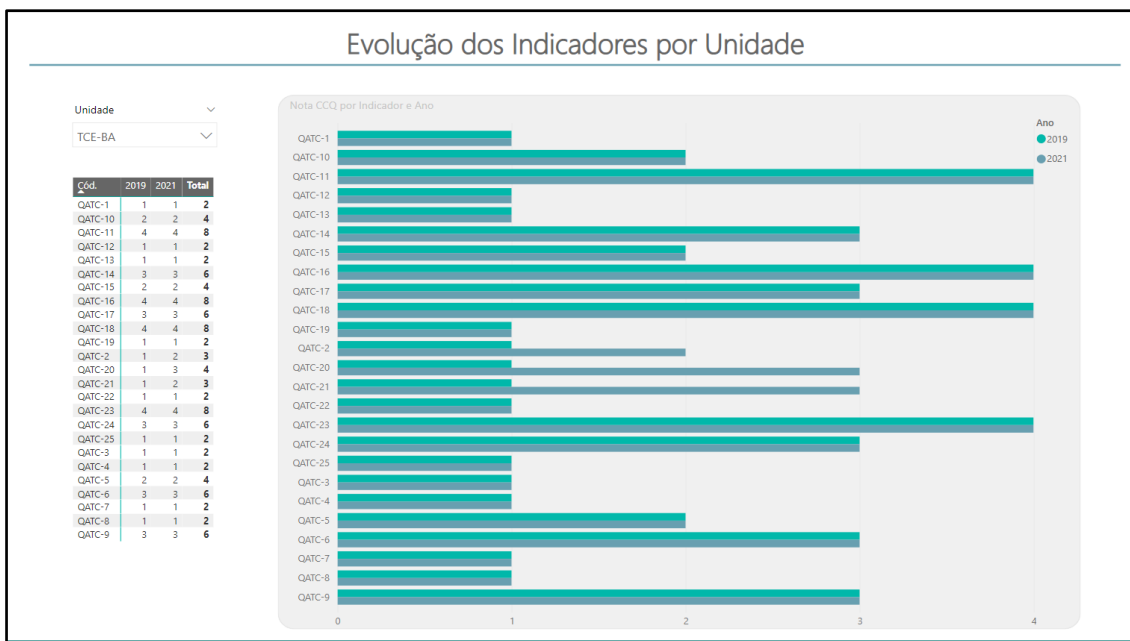


Figura 4.10 - *Dashboard* de “Evolução dos Indicadores por Unidade”.

4.5 - *DASHBOARD* “INDICADORES POR REGIÃO”

Este painel apresenta os indicadores das diferentes unidades com foco no aspecto geográfico. Inicialmente, deve ser selecionado um ano e um subconjunto de indicadores para análise (que pode compreender um, vários ou mesmo todos os indicadores) por meio dos componentes de segmentação de dados apresentados no lado esquerdo superior. Os dados de indicadores são apresentados com base nesse filtro.

Em uma tabela ao lado esquerdo, são apresentados os totais acumulados das notas dos indicadores e a nota média por indicador de cada uma das unidades, organizadas por região e estado. Ao centro, um gráfico plota no mapa do Brasil, na localização referente a cada unidade, círculos que representam as notas obtidas. A aparência dos círculos varia em tamanho e cor, sendo que notas maiores são representadas por círculos maiores e que tendem mais ao verde, enquanto que notas menores são representadas por círculos menores e que tendem mais ao vermelho.

Ao lado direito, são apresentados seis (6) cartões com a média da pontuação total das unidades para cada uma das regiões brasileiras e para o Distrito Federal. É possível ver o aspecto final do *dashboard* na Figura 4.11.

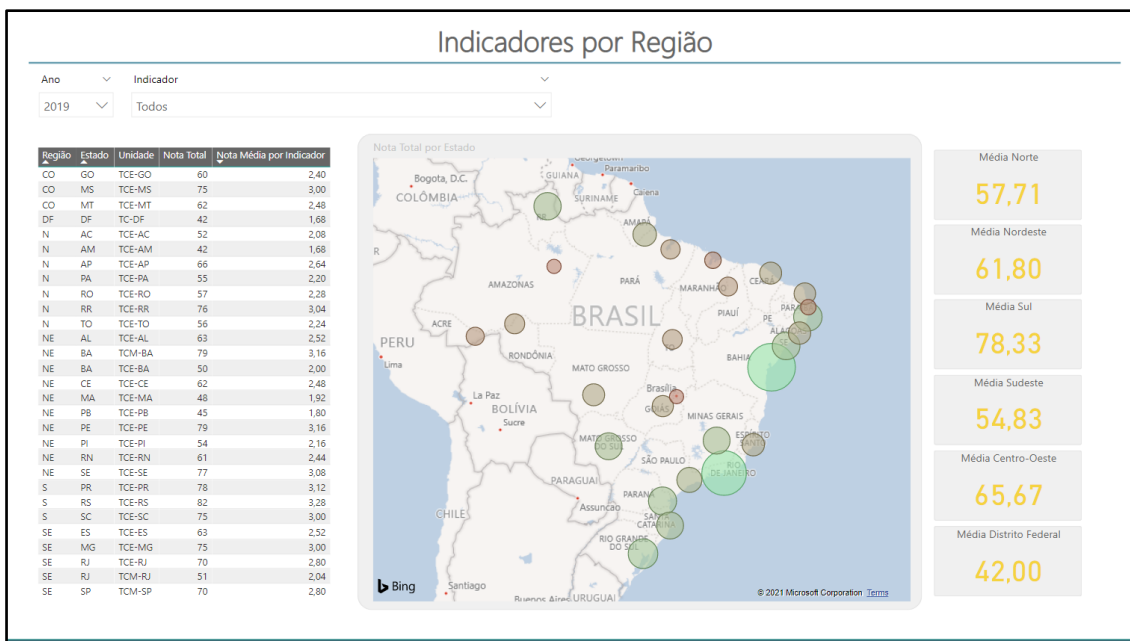


Figura 4.11 - Dashboard de “Indicadores por Região”.

4.6 - DASHBOARD “DESEMPENHO DE UNIDADES POR INDICADOR”

Este painel permite a análise indicador a indicador entre as unidades. Inicialmente, deve ser selecionado o ano e o indicador desejado por meio dos componentes de segmentação de dados apresentados no lado esquerdo superior. Os dados são apresentados com base nesse filtro.

A descrição do indicador é apresentada em destaque de forma centralizada e, logo abaixo dela, encontramos dois gráficos. À esquerda, um gráfico de barras verticais apresenta a nota de cada uma das unidades para o indicador selecionado. À direita, um gráfico de rosca apresenta a distribuição dessas unidades de acordo com o nível atingido na escala de pontuação.

É possível ver o aspecto final desse dashboard na Figura 4.12. Ao passar o ponteiro do mouse sobre o segundo gráfico, a ferramenta apresenta um detalhamento apontando a quantidade de unidades e o percentual em relação ao todo que está classificado sob aquele nível.



Figura 4.12 - *Dashboard* de “Desempenho de Unidades por Indicador”.

4.7 - *DASHBOARD* “UNIDADES POR NÍVEL DE DESENVOLVIMENTO”

Este painel permite a análise das unidades com base na média das notas dos indicadores, classificando-as de acordo com o nível atingido na escala de pontuação. Inicialmente, deve ser selecionado o ano para análise por meio do componente de segmentação de dados apresentado no lado esquerdo superior. Os dados são apresentados com base nesse filtro.

Em uma tabela ao lado esquerdo, são apresentados o código e o nome da unidade, juntamente com a nota total (somatória) dos indicadores e a média das notas atribuídas pelo CGQ. À direita, um gráfico de pizza apresenta a distribuição dessas unidades de acordo com o nível atingido na escala de pontuação, sendo a totalização dos níveis apresentada em forma de cartões acima deste gráfico.

É possível ver o aspecto final desse *dashboard* na Figura 4.13. Ao passar o ponteiro do *mouse* sobre o gráfico, a ferramenta apresenta um detalhamento apontando a quantidade de unidades e o percentual em relação ao todo que está classificado sob aquele nível.

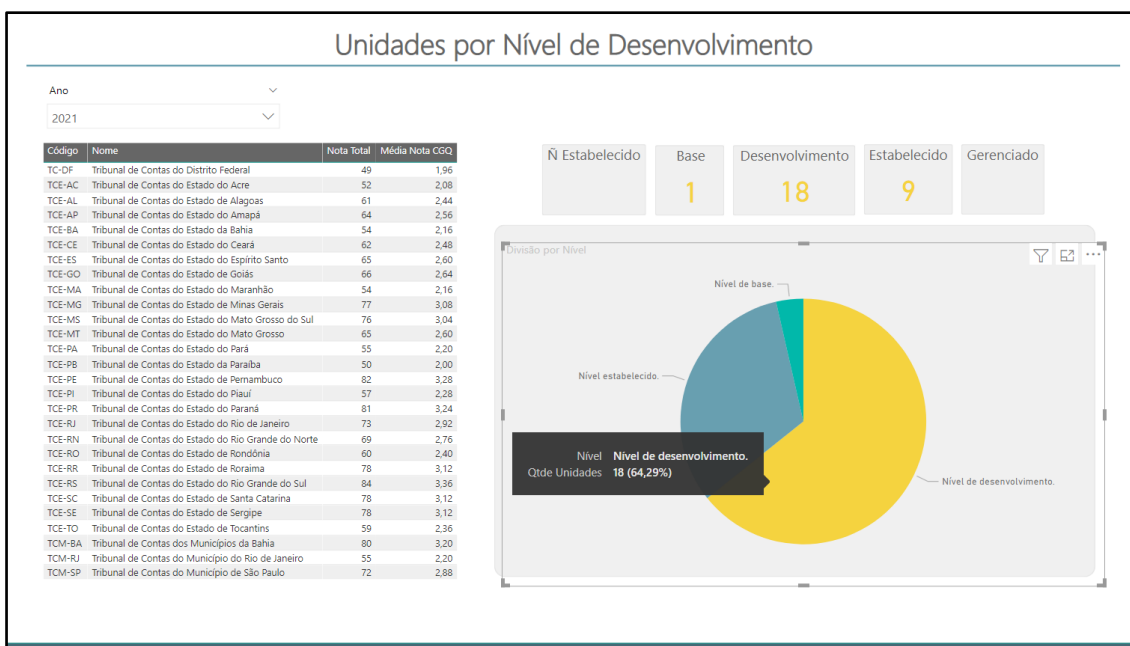


Figura 4.13 - *Dashboard* de “Unidades por Nível de Desenvolvimento”.

4.8 - *DASHBOARD* “INDICADORES/DIMENSÕES”

Este painel permite a análise dos indicadores tendo em vista as dimensões que os compõem. Inicialmente, deve ser selecionado o ano, o indicador e um subconjunto de unidades para análise (que pode compreender uma, várias ou mesmo todas as unidades) para análise por meio dos componentes de segmentação de dados apresentados na parte superior. Os dados são apresentados com base nesse filtro.

A descrição do indicador é apresentada em destaque de forma centralizada e, logo abaixo dela, encontramos dois gráficos. À esquerda, um gráfico de barras horizontais apresenta a média da nota final de cada uma das dimensões do indicador e unidades selecionadas. À direita, um gráfico de rosca apresenta a distribuição dessas unidades de acordo com o nível atingido na escala de pontuação. É possível ver o aspecto final desse *dashboard* na Figura 4.14.

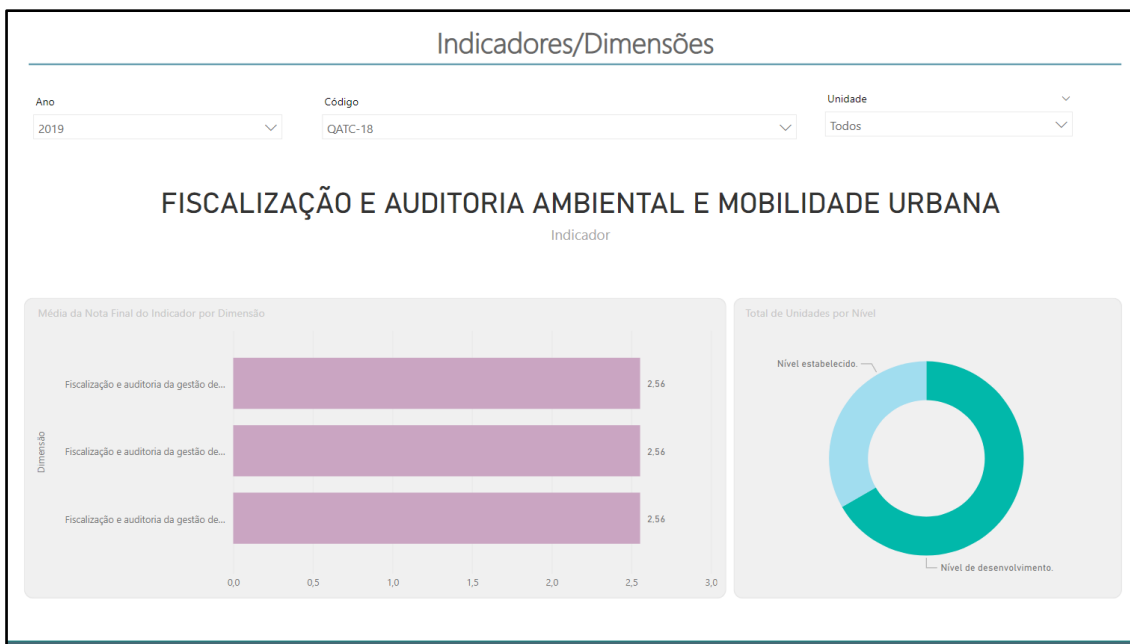


Figura 4.14 - *Dashboard* de “Indicadores/Dimensões”.

4.9 - *DASHBOARD* “INDICADORES/DIMENSÕES - GERAL”

De forma complementar ao painel apresentado anteriormente, este permite a análise dos indicadores e dimensões de forma mais ampla. Inicialmente, deve ser selecionado o ano, um subconjunto de unidades para análise (que pode compreender uma, várias ou mesmo todas as unidades) e a(s) dimensão(ões) desejadas para análise por meio dos componentes de segmentação de dados apresentados na parte superior. Os dados são apresentados com base nesse filtro.

À esquerda, um gráfico de barras horizontais apresenta as dimensões ordenadas de forma decrescente por média da nota final da dimensão. À direita, uma tabela exibe as dimensões, agrupadas por indicador, com as médias das notas finais por indicador e por dimensão. É possível ver o aspecto final desse *dashboard* na Figura 4.15.

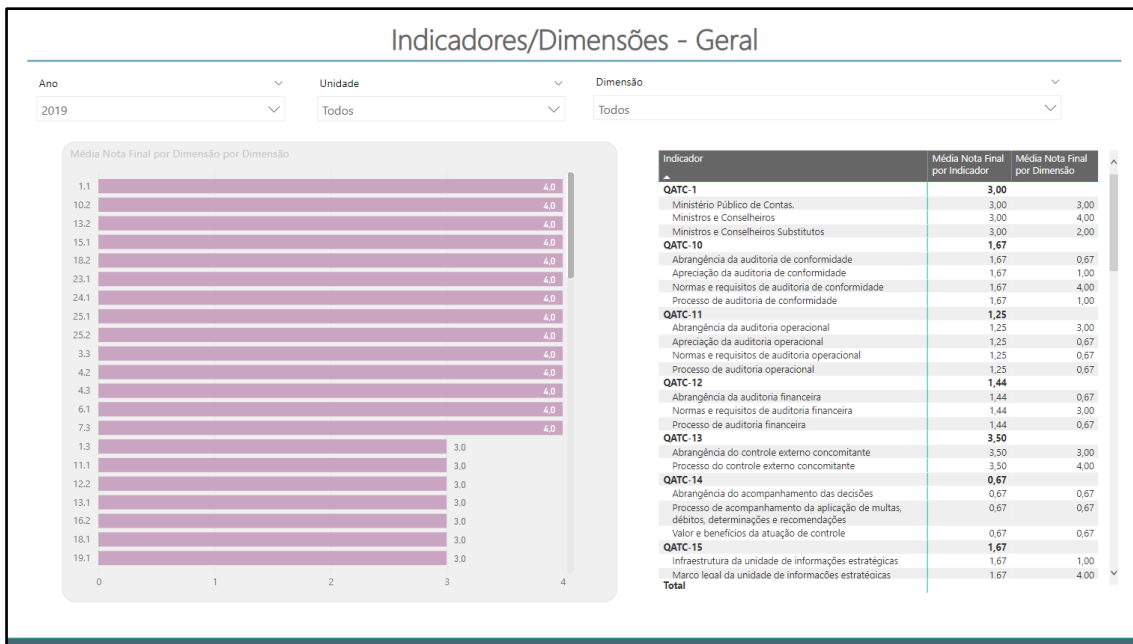


Figura 4.15 - Dashboard de “Indicadores/Dimensões – Geral”.

CAPÍTULO 5

CONCLUSÕES E SUGESTÕES

5.1 - CONCLUSÕES

O presente trabalho de pesquisa teve como objetivo principal o desenvolvimento de uma solução de *Business Intelligence* na avaliação de desempenho em Tribunais de Contas Brasileiros com base nos critérios definidos pela ATRICON no seu Marco de Medição de Desempenho – Qualidade e Agilidade dos Tribunais de Contas (MMD-QATC), que auxiliasse na consolidação e na análise dos dados e, com isso, trouxesse uma otimização ao processo.

Para tanto, foram estudadas e aplicadas técnicas e ferramentas de BI na construção da solução que permitisse a consolidação da medição de vinte e cinco (25) indicadores e setenta e nove (79) dimensões, conforme estabelecido pelo Marco de Medição, para diversas unidades sujeitas à avaliação.

Com base na identificação da origem e organização dos dados, a solução abrangeu a criação de um algoritmo capaz de localizar e extrair das planilhas, formato de arquivo estabelecido pelo Manual de Procedimentos do MMD-TC, os valores das notas atribuídas e outras informações pertinentes. Como parte final do processamento deste algoritmo, esses dados são armazenados em um SGBD.

Tais dados que antes estariam pulverizados em diversos arquivos e, portanto, requerendo um esforço maior para manipulá-los, a partir da execução do algoritmo, passam a estar armazenados de forma estruturada em um único ponto, facilitando assim o seu manuseio. Estes ainda foram trabalhados por processos mapeados no Pentaho *Data Integration* a fim de torná-los aptos para consumo final nas diferentes visões apresentadas nos *dashboards* contruídos. *Dashboards* estes que apresentaram diferentes possibilidades de observação dos dados, para além do óbvio “variável/valor”.

Foi possível construir *dashboards* que permitem o acompanhamento da entrega/processamento dos questionários respondidos; que mostram sim os valores das notas atribuídas para cada um dos indicadores e dimensões, mas expandem as possibilidades de interpretação dos dados ao exibir, por exemplo, em forma de gráfico o equilíbrio de desempenho, ou ausência dele, entre os diferentes indicadores de uma unidade, a evolução do desempenho dos indicadores ao longo do tempo ou ainda ao

plotar, este desempenho em forma de círculos que variam em tamanho e cor (em função do valor) sobre o mapa do Brasil.

O processo de avaliação de desempenho em Tribunais de Contas Brasileiros passa a ser simplificado (a partir do momento no qual não há mais a necessidade de consolidação manual dos diversos arquivos de dados remetidos pelos TCs) e análises previamente definidas são geradas e apresentadas de forma automatizada, auxiliando o processo de geração de conhecimento e potencial tomada de decisão na definição de políticas diversas que visem a evolução contínua desses órgãos, trazendo assim, diretamente, benefícios ao País e à população.

E ainda, a forma como a solução foi desenvolvida permite a sua utilização continuada, por meio da execução de simples comandos e com disponibilidade imediata dos dados para análise, e a expansão, por meio da agregação de novas variáveis aos modelos de dados e construção de novos *dashboards*.

No entanto, apesar de não ser objetivo deste estudo, uma análise da situação real dos TCs não pode ser feita, pelo fato de os dados reais não estarem disponíveis à população, conforme explicado anteriormente. Porém, uma vez disponíveis, a solução desenvolvida neste estudo permitirá.

5.2 - SUGESTÕES

Considerando o anseio da sociedade por instituições públicas mais transparentes e a crescente cultura de dados abertos na administração pública, mas a depender da vontade dos Tribunais de Contas e da Atricon em liberar os dados das avaliações, acredita-se que seria de grande contribuição construir novos *dashboards*, voltados para a população em geral, apresentando os resultados das avaliações.

E por fim, poderia ser implementado e incorporado à solução, um mecanismo de análise de correlação entre os indicadores e/ou dimensões com, por exemplo, a utilização da biblioteca *numpy*, para linguagem *Python*. Ou ainda a utilização de uma das funções da *scikit-learn* (outra biblioteca para *Python*) para a determinação da relevância de um atributo para o modelo, como por exemplo a *feature importance*, que retorna um *score* para cada atributo analisado, tendo este um valor maior conforme a importância para o modelo desse atributo.

REFERÊNCIAS BIBLIOGRÁFICAS

ARCHITECT, SITE POWER. **Free Download.** Disponível em: <<http://www.bestofbi.com/page/architectdownloads>>. Acesso em: 20 de setembro de 2021, 21h.

ATRICON, ASSOCIAÇÃO DOS MEMBROS DOS TRIBUNAIS DE CONTAS DO BRASIL. **Apresentação.** Disponível em: <<http://www.atricon.org.br/institucional/apresentacao/>>. Acesso em: 15 de setembro de 2020, 23h.

ATRICON, ASSOCIAÇÃO DOS MEMBROS DOS TRIBUNAIS DE CONTAS DO BRASIL. **Marco de Medição do Desempenho – Qualidade e Agilidade dos Tribunais de Contas do Brasil (MMD-QATC).** Brasília. 2014. Disponível em: <http://www.atricon.org.br/wp-content/uploads/2015/03/QATC2_VersaoFinalPublicada.pdf>. Acesso em: 15 de setembro de 2020, 23h50.

ATRICON, ASSOCIAÇÃO DOS MEMBROS DOS TRIBUNAIS DE CONTAS DO BRASIL. **Marco de Medição do Desempenho dos Tribunais de Contas MMD-TC (Versão 2017).** Brasília. 2017. Disponível em: <<http://qatc.atricon.org.br/wp-content/uploads/2019/03/MMD-TC-DocumentoBase-2017-vers%C3%A3o-publicada.pdf>>. Acesso em: 16 de setembro 2020, 23h.

ATRICON, ASSOCIAÇÃO DOS MEMBROS DOS TRIBUNAIS DE CONTAS DO BRASIL. **Manual de Procedimentos do MMD-TC.** Brasília. 2019. Disponível em: <[http://qatc.atricon.org.br/wp-content/uploads/2019/03/2.-Manual-de-Procedimentos-MM D-TC-v.1.pdf](http://qatc.atricon.org.br/wp-content/uploads/2019/03/2.-Manual-de-Procedimentos-MM-D-TC-v.1.pdf)>. Acesso em: 05 de março de 2020, 18h.

ATRICON, ASSOCIAÇÃO DOS MEMBROS DOS TRIBUNAIS DE CONTAS DO BRASIL. **Regulamento ATRICON Nº 01/2013.** Disponível em: <http://www.atricon.org.br/wp-content/uploads/2013/07/REGULAMENTO_PARAMETROS_AGILIDADE_QUALIDADE.pdf>. Acesso em: 16 de setembro de 2020, 23h.

BORALI, M. N. **Capacidade Estatal e Tribunais de Contas no Brasil: Uma análise sobre seus recursos humanos e informacionais.** 2018. Dissertação (Mestrado em Administração Pública e Governo) – Escola de Administração de Empresas da Fundação Getúlio Vargas. São Paulo. 2018.

CASTALDELLI, M. E. **Divulgação de desempenho de órgãos públicos: uma Análise de Conteúdo dos relatórios de gestão do Tribunal de Contas da União.** 2010. Dissertação (Programa de Pós-Graduação em Controladoria e Contabilidade). Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo. Ribeirão Preto, 2010.

DATE, C. J. **Introdução a Sistemas de Bancos de Dados**. 8. ed. Rio de Janeiro: Elsevier, 2003.

ELMASRI, R., NAVACHE, S. B. **Sistemas de banco de dados**. 6. ed. São Paulo: Pearson Addison Wesley, 2011.

FOLEY, É.; GUILLEMETTE, M. G. What is Business Intelligence? **International Journal of Business Intelligence Research**, v. 1, n. 4, pp. 1-28, October-December, 2010.

FRANCK, K. M., PEREIRA, R. F., FILHO, J. V. D. Diagrama Entidade-Relacionamento: uma ferramenta para modelagem de dados conceituais em Engenharia de Software. **Research, Society and Development**, v. 10, n. 8, pp. 1-12, 2021.

DAVENPORT, T. H., HARRIS, J. G. Automated Decision Making Comes of Age. **Sloan Management Review**, v. 46, n. 4, pp. 83-89. 2005.

GARCIA, E. **Tribunal de Contas do Estado do Amazonas - Sessenta e Cinco Anos**. 2. ed. Manaus: Norma Editora, 2015.

GESPÚBLICA, PROGRAMA NACIONAL DE GESTÃO PÚBLICA E DESBUROCRATIZAÇÃO. **Produto 4: guia referencial para medição de desempenho e manual para construção de indicadores**. Secretaria de Gestão. Ministério do Planejamento. Brasília, 2009. Disponível em: <http://www.gespublica.gov.br/sites/default/files/documentos/guia_indicadores_jun2010.pdf>. Acesso em: 09 de março de 2020, 20h30.

GHISI, A. P. Desempenho das Entidades Fiscalizadoras Superiores e Indicadores de Rendimento. **Revista do TCU**, v. 31, n. 86, out/dez, pp. 327-337, 2000.

HEIDISQL, SITE. **Download HeidiSQL**. Disponível em: <<https://www.heidisql.com/download.php/>>. Acesso em: 19 de setembro de 2021, 22h.

HEUSER, C. A. **Projeto de banco de dados**. 6. ed. Porto Alegre: Bookman, 2009.

HORAKOVA, M., SKALSKA, H. Business Intelligence and Implementation in a Small Enterprise. **Journal of Systems Integration**, v. 4, n. 2, pp. 50-61, 2013.

INFNET, ESCOLA SUPERIOR DA TECNOLOGIA DA INFORMAÇÃO - INSTITUTO. **Vamos falar de Python.** Disponível em: < <https://www.infnet.edu.br/esti/vamos-falar-de-python/>>. Acesso em: 20 de setembro de 2021, 11h.

INFOQ, SITE. **Pentaho Data Integration - ETL em Software Livre.** Disponível em: < <https://www.infoq.com/br/articles/pentaho-pdi/>>. Acesso em: 23 de setembro de 2021, 11h.

LUZ, M. B. B. de C. **Da Legalidade à Legitimidade: Vozes em torno da reforma dos Tribunais de Contas Brasileiros no período de 2000 a 2017.** 2018. Dissertação (Mestrado em Administração) - Faculdade de Administração, Ciências Contábeis e Economia da Universidade Federal de Goiás. Goiânia, 2018.

MYSQL, SITE. **Download MySQL.** Disponível em: <<https://dev.mysql.com/doc/>>. Acesso em: 20 de setembro de 2021a, 20h.

MYSQL, SITE. **Download MySQL.** Disponível em: <<https://www.mysql.com/downloads/>>. Acesso em: 20 de setembro de 2021b, 20h.

NEGASH, S., GRAY, P. “Business Intelligence”. In: Burstein, F., Holsapple, C. W. (eds.), **Handbook on Decision Support Systems 2: Variations**, 1st. ed., chapter 2, Berlin, German, Springer-Verlag, 2008.

OCDE, ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO. **Internal Control and Internal Audit: Ensuring Public Sector Integrity and Accountability.** Report 47638204, OECD Conference Centre, Paris, France, 2011.

PANDAS, SITE. **Package overview – Pandas Documentation.** Disponível em: < https://pandas.pydata.org/docs/getting_started/overview.html >. Acesso em: 20 de setembro de 2021, 11h.

PENTAHO, SITE. **Pentaho from Hitachi Vantara.** Disponível em: < <https://sourceforge.net/projects/pentaho/>>. Acesso em: 23 de setembro de 2021, 11h.

POPOVIČ, A., TURK, T., JAKLIČ, J. Conceptual model of business value of business intelligence systems. **Management**, v. 15, n. 1, pp. 5-30, 2010.

PYTHON, SITE. **Python.org – Our documentation.** Disponível em: < <https://www.python.org/doc>>. Acesso em: 20 de setembro de 2021, 11h.

REIS, G. V. dos; COLAUTO, R. D., PINHEIRO, L. E. T. Proposta de adaptação do Balanced Scorecard para Órgãos Públicos: um enfoque na secretaria Geral de Administração (SEGEDAM) do Tribunal de Contas da União. **XIV Congresso Brasileiro de Custos**, João Pessoa, Brasil, 05-07 de dezembro, 2007.

SABHERWAL, R., BECERRA-FERNANDEZ, I. **Business intelligence: Practices, Technologies, and Management**. 1st. ed., USA: John Wiley & Sons, INC, 2011.

SABHERWAL, R. Succeeding with Business Intelligence: Some Insights and Recommendations. **Cutter Benchmark Review**, v.7, n. 9, pp. 5-15, 2007.

SABHERWAL, R. KM and BI: From Mutual Isolation to Complementarity and Synergy. **Cutter Consortium Executive Report**, v.8, n. 8, pp. 1-18, 2008.

SILVA, I. L. da; DRUMOND, R. B. A Necessidade da Utilização de Sistema de Custos e de Indicadores de Desempenho na Administração Pública. **4º Congresso USP de Controladoria e Contabilidade**, São Paulo, Brasil, 07-08 de outubro, 2004.

TURBAN, E. *et al.* **Business Intelligence – Um enfoque gerencial para a Inteligência do Negócio**. 1. ed. Porto Alegre: Bookman, 2009.

APÊNDICE A

DOMÍNIOS E INDICADORES NO MMD-QATC

Tabela A.1 - Lista de domínios do MMD-QATC, com seus respectivos indicadores.

Domínio	Código	Indicador
Domínio A: Independência e Marco Legal	QATC 01	Composição, organização e funcionamento dos T
Domínio B: Governança Interna	QATC 02	Liderança
	QATC 03	Estratégia
	QATC 04	Accountability
	QATC 05	Agilidade no julgamento e gerenciamento de prazos de processos
	QATC 06	Gestão de pessoas
	QATC 07	Desenvolvimento profissional
	Domínio C: Fiscalização e Auditoria	QATC 08
QATC 09		Controle e garantia de qualidade de fiscalizações e auditorias
QATC 10		Auditoria de conformidade
QATC 11		Auditoria operacional
QATC 12		Auditoria financeira
QATC 13		Controle externo concomitante
QATC 14		Acompanhamento das decisões
QATC 15		Informações estratégicas para o controle externo
Domínio D: Fiscalização da Infraestrutura e Meio Ambiente	QATC 16	Fiscalização e auditoria de obras e serviços de engenharia
	QATC 17	Fiscalização e auditoria de concessões e privatizações
	QATC 18	Fiscalização e auditoria ambiental e da mobilidade urbana
Domínio E: Fiscalização e Auditoria de Políticas Públicas Sociais	QATC 19	Fiscalização e auditoria da gestão da educação
	QATC 20	Fiscalização e auditoria da gestão da saúde
	QATC 21	Fiscalização e auditoria da gestão da previdência própria
	QATC 22	Fiscalização e auditoria da gestão da segurança pública

Domínio F: Fiscalização e Auditoria da Gestão Fiscal, Controle Interno, Tecnologia da Informação, Transparência e Ouvidoria

QATC 23 Fiscalização e auditoria da gestão fiscal e da renúncia de receita

QATC 24 Fiscalização e auditoria do controle interno e tecnologia da informação dos jurisdicionados

QATC 25 Fiscalização e auditoria da transparência e da ouvidoria dos jurisdicionados

APÊNDICE B

SCRIPT SQL DE CRIAÇÃO DO *DATA WAREHOUSE*

```
USE dw_mmd_qatc;
```

```
CREATE TABLE Indicador (  
    Cod_indicador VARCHAR(50) NOT NULL,  
    Desc_indicador VARCHAR(250) NOT NULL,  
    PRIMARY KEY (Cod_indicador)  
);
```

```
CREATE TABLE Dimensao (  
    Cod_dimensao VARCHAR(50) NOT NULL,  
    Desc_dimensao VARCHAR(250) NOT NULL,  
    Cod_indicador VARCHAR(50) NOT NULL,  
    PRIMARY KEY (Cod_dimensao)  
);
```

```
CREATE TABLE Regiao (  
    Sigla_regiao VARCHAR(2) NOT NULL,  
    Nome_regiao VARCHAR(250) NOT NULL,  
    PRIMARY KEY (Sigla_regiao)  
);
```

```
CREATE TABLE Estado (  
    Sigla_estado VARCHAR(2) NOT NULL,  
    Nome_estado VARCHAR(250) NOT NULL,  
    Sigla_regiao VARCHAR(2) NOT NULL,  
    PRIMARY KEY (Sigla_estado)  
);
```

```
CREATE TABLE Unidade (  
    Cod_unidade VARCHAR(50) NOT NULL,  
    Nome_unidade VARCHAR(250) NOT NULL,  
    Sigla_estado VARCHAR(2) NOT NULL,  
    PRIMARY KEY (Cod_unidade)  
);
```

```
CREATE TABLE Tempo (  
    Cod_Tempo VARCHAR(50) NOT NULL,  
    Data DATE NOT NULL,  
    Numero_Dia_Semana NVARCHAR(50) NOT NULL,  
    Numero_Mes NVARCHAR(50) NOT NULL,  
    Numero_Ano NVARCHAR(50) NOT NULL,
```

```

Nome_Mes NVARCHAR(250) NOT NULL,
Numero_Trimestre NVARCHAR(50) NOT NULL,
Nome_Trimestre NVARCHAR(250) NOT NULL,
Numero_Semestre NVARCHAR(50) NOT NULL,
Nome_Semestre NVARCHAR(250) NOT NULL,
PRIMARY KEY (Cod_Tempo)
);

CREATE TABLE Fato_medicao_dimensao (
    Cod_dimensao VARCHAR(50) NOT NULL,
    Cod_unidade VARCHAR(50) NOT NULL,
    Cod_tempo VARCHAR(50) NOT NULL,
    Ano_medicao INT NOT NULL,
    Nota_ca INT NOT NULL,
    Nota_ccq INT NOT NULL,
    Nota_cgq INT NOT NULL,
    Nota_ccg INT NOT NULL,
    Nota_final INT NOT NULL,
    PRIMARY KEY (Cod_dimensao, Cod_unidade, Cod_tempo)
);

CREATE TABLE Fato_medicao_indicador (
    Cod_indicador VARCHAR(50) NOT NULL,
    Cod_unidade VARCHAR(50) NOT NULL,
    Cod_tempo VARCHAR(50) NOT NULL,
    Ano_medicao INT NOT NULL,
    Nota_ccq INT NOT NULL,
    Nota_cgq INT NOT NULL,
    PRIMARY KEY (Cod_indicador, Cod_unidade, Cod_tempo)
);

ALTER TABLE Fato_medicao_indicador ADD CONSTRAINT
indicador_fato_medicao_indicador_fk
FOREIGN KEY (Cod_indicador)
REFERENCES Indicador (Cod_indicador)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE Dimensao ADD CONSTRAINT indicador_dimensao_fk
FOREIGN KEY (Cod_indicador)
REFERENCES Indicador (Cod_indicador)
ON DELETE NO ACTION
ON UPDATE NO ACTION;

ALTER TABLE Fato_medicao_dimensao ADD CONSTRAINT
dimensao_fato_medicao_dimensao_fk
FOREIGN KEY (Cod_dimensao)

```

```
REFERENCES Dimensao (Cod_dimensao)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE Estado ADD CONSTRAINT regioao_estado_fk
FOREIGN KEY (Sigla_regiao)
REFERENCES Regiao (Sigla_regiao)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE Unidade ADD CONSTRAINT estado_unidade_fk
FOREIGN KEY (Sigla_estado)
REFERENCES Estado (Sigla_estado)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE Fato_medicao_indicador ADD CONSTRAINT
unidade_fato_medicao_indicador_fk
FOREIGN KEY (Cod_unidade)
REFERENCES Unidade (Cod_unidade)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE Fato_medicao_dimensao ADD CONSTRAINT
unidade_fato_medicao_dimensao_fk
FOREIGN KEY (Cod_unidade)
REFERENCES Unidade (Cod_unidade)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE Fato_medicao_indicador ADD CONSTRAINT
tempo_fato_medicao_indicador_fk
FOREIGN KEY (Cod_tempo)
REFERENCES Tempo (Cod_Tempo)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE Fato_medicao_dimensao ADD CONSTRAINT
tempo_fato_medicao_dimensao_fk
FOREIGN KEY (Cod_tempo)
REFERENCES Tempo (Cod_Tempo)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

APÊNDICE C

SCRIPT SQL DE CRIAÇÃO DO DATAMART

```
USE dm_mmd_qatc;
```

```
CREATE TABLE Dim_dimensao (  
    id_dimensao INT AUTO_INCREMENT NOT NULL,  
    Cod_dimensao VARCHAR(50) NOT NULL,  
    Desc_dimensao VARCHAR(250) NOT NULL,  
    Cod_indicador VARCHAR(50) NOT NULL,  
    Desc_indicador VARCHAR(250) NOT NULL,  
    PRIMARY KEY (id_dimensao)  
);
```

```
CREATE TABLE Dim_indicador (  
    id_indicador INT AUTO_INCREMENT NOT NULL,  
    cod_indicador VARCHAR(50) NOT NULL,  
    desc_indicador VARCHAR(250) NOT NULL,  
    PRIMARY KEY (id_indicador)  
);
```

```
CREATE TABLE Dim_unidade (  
    id_unidade INT AUTO_INCREMENT NOT NULL,  
    Cod_unidade VARCHAR(50) NOT NULL,  
    Nome_unidade VARCHAR(250) NOT NULL,  
    Sigla_estado VARCHAR(2) NOT NULL,  
    Nome_estado VARCHAR(250) NOT NULL,  
    Sigla_regiao VARCHAR(2) NOT NULL,  
    Nome_regiao VARCHAR(250) NOT NULL,  
    PRIMARY KEY (id_unidade)  
);
```

```
CREATE TABLE Dim_Tempo (  

```

```
id_tempo INT AUTO_INCREMENT NOT NULL,  
Cod_Tempo NVARCHAR(50) NOT NULL,  
Data DATE NOT NULL,  
Numero_Dia_Semana NVARCHAR(50) NOT NULL,  
Numero_Mes NVARCHAR(50) NOT NULL,  
Numero_Ano NVARCHAR(50) NOT NULL,  
Nome_Mes NVARCHAR(250) NOT NULL,  
Numero_Trimestre NVARCHAR(50) NOT NULL,  
Nome_Trimestre NVARCHAR(250) NOT NULL,  
Numero_Semestre NVARCHAR(50) NOT NULL,  
Nome_Semestre NVARCHAR(250) NOT NULL,  
PRIMARY KEY (id_tempo)  
);
```

```
CREATE TABLE Fato_medicao_dimensao (  
id_unidade INT NOT NULL,  
id_tempo INT NOT NULL,  
id_dimensao INT NOT NULL,  
Nota_ca INT NOT NULL,  
Nota_ccq INT NOT NULL,  
Nota_cgq INT NOT NULL,  
Nota_ccg INT NOT NULL,  
Nota_final INT NOT NULL,  
PRIMARY KEY (id_unidade, id_tempo, id_dimensao)  
);
```

```
CREATE TABLE fato_medicao_indicador (  
id_unidade INT NOT NULL,  
id_tempo INT NOT NULL,  
id_indicador INT NOT NULL,  
Nota_ccq INT NOT NULL,  
Nota_cgq INT NOT NULL,  
PRIMARY KEY (id_unidade, id_tempo, id_indicador)  
);
```

```
ALTER TABLE Fato_medicao_dimensao ADD CONSTRAINT
dimensao_fato_medicao_dimensao_fk
FOREIGN KEY (id_dimensao)
REFERENCES Dim_dimensao (id_dimensao)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE fato_medicao_indicador ADD CONSTRAINT
indicador_fato_medicao_indicador_fk
FOREIGN KEY (id_indicador)
REFERENCES Dim_indicador (id_indicador)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE fato_medicao_indicador ADD CONSTRAINT
dim_unidade_fato_medicao_indicador_fk
FOREIGN KEY (id_unidade)
REFERENCES Dim_unidade (id_unidade)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE Fato_medicao_dimensao ADD CONSTRAINT
dim_unidade_fato_medicao_dimensao_fk
FOREIGN KEY (id_unidade)
REFERENCES Dim_unidade (id_unidade)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```

```
ALTER TABLE fato_medicao_indicador ADD CONSTRAINT
dim_tempo_fato_medicao_indicador_fk
FOREIGN KEY (id_tempo)
REFERENCES Dim_Tempo (id_tempo)
ON DELETE NO ACTION
ON UPDATE NO ACTION;
```



```
ALTER TABLE Fato_medicao_dimensao ADD CONSTRAINT  
dim_tempo_fato_medicao_dimensao_fk  
FOREIGN KEY (id_tempo)  
REFERENCES Dim_Tempo (id_tempo)  
ON DELETE NO ACTION  
ON UPDATE NO ACTION;
```

APÊNDICE D

TRECHO DO CÓDIGO FONTE EM PYTHON DO PROCESSO DE ETL DO *DATA WAREHOUSE*

```
import etlFunctions
import dbFunctions

print("Iniciando ETL do DW!")
planilhas = etlFunctions.getPlanilhasDiretorio()
for p in planilhas:
    print("-> Carregando", p)

    # Faz a leitura da planilha e retorna os dados de medição dos indicadores
    indicadores = etlFunctions.etlResumoQATC(etlFunctions.DIRETORIO + p)
    # Aciona o método para persistir os dados de medição dos indicadores
    dbFunctions.persistirResumoQATC(indicadores)

    # Faz a leitura da planilha e retorna os dados de medição das dimensões
    dimensoes = etlFunctions.etlDimensoesBaseQATC(etlFunctions.DIRETORIO + p)

    # Aciona o método para persistir os dados de medição das dimensoes
    dbFunctions.persistirBaseDimensoes(dimensoes)
```