



**UNIVERSIDADE ESTADUAL DO NORTE DO PARANÁ**  
**CAMPUS LUIZ MENEGHEL - CENTRO DE CIÊNCIAS TECNOLÓGICAS**  
**SISTEMAS DE INFORMAÇÃO**

**GUSTAVO DE OLIVEIRA**

**IMPLANTAÇÃO DE UM AMBIENTE DE BI**  
**UTILIZANDO UMA BASE DE PROCESSO SELETIVO**  
**COM A METODOLOGIA AGILE ROLAP**

Bandeirantes

2018

**GUSTAVO DE OLIVEIRA**

**IMPLANTAÇÃO DE UM AMBIENTE DE BI  
UTILIZANDO UMA BASE DE PROCESSO SELETIVO  
COM A METODOLOGIA AGILE ROLAP**

Trabalho de Conclusão de Curso submetido à  
Universidade Estadual do Norte do Paraná,  
como requisito parcial para obtenção do grau  
de Bacharel em Sistemas de Informação e  
Licenciado em Computação.

Orientador: Prof. Dr. André Luis Andrade  
Menolli

Bandeirantes

2018

**GUSTAVO DE OLIVEIRA**

**IMPLANTAÇÃO DE UM AMBIENTE DE BI  
UTILIZANDO UMA BASE DE PROCESSO SELETIVO  
COM A METODOLOGIA AGILE ROLAP**

Trabalho de Conclusão de Curso submetido à  
Universidade Estadual do Norte do Paraná,  
como requisito parcial para obtenção do grau  
de Bacharel em Sistemas de Informação e  
Licenciado em Computação.

**COMISSÃO EXAMINADORA**

---

Prof. Dr. André Luis Andrade Menolli  
UENP – Campus Luiz Meneghel

---

Prof. Me. Glauco Carlos Silva  
UENP – Campus Luiz Meneghel

---

Prof. Me. Wellington Aparecido Della Mura  
UENP – Campus Luiz Meneghel

Bandeirantes, \_\_ de \_\_\_\_\_ de 2018.

## **AGRADECIMENTOS**

Agradeço à minha família, namorada e amigos que sempre me apoiaram, me acompanharam e me ajudaram neste período. Agradeço também ao meu orientador por todo tempo dedicado a me ajudar e a me orientar nesse trabalho. Devo meus agradecimentos também a todos os professores e colegas que contribuíram com o meu aprendizado e desenvolvimento, pessoal e profissional, durante todos esses anos da graduação.

## RESUMO

As organizações geram informações que geralmente não são aproveitadas por elas mesmas. Para aproveitá-las é indicado utilizar BI que utiliza dos dados advindos de diversas fontes para gerar informações que podem ser úteis ao realizar uma tomada de decisão, gerar relatórios, etc. Porém os métodos tradicionais de implantação de ambiente de BI costumam ser caros e levam um tempo relevante para serem concluídos. Para resolver esse problema surgiu a metodologia *Agile* ROLAP, que além de utilizar algumas ferramentas de visualização OLAP feitas para as metodologias tradicionais, é também uma alternativa para pequena e médias empresas que buscam implantar um ambiente de BI. A metodologia *Agile* ROLAP utiliza de ferramentas e de alguns *plugins* que foram desenvolvidos para facilitar o desenvolvimento do ambiente de BI. Portanto esse trabalho busca analisar além da metodologia *Agile* ROLAP, analisar também as ferramentas e os *plugins* que a compõem.

**Palavras-chave:** Informação, Business Intelligence, Agile ROLAP.

## **ABSTRACT**

Organizations generate information that is often not used by them. To take advantage of them, it is advisable to use BI that takes advantage of the data coming from diverse sources to generate information that can be useful in making a decision, generating reports, etc. But traditional methods of BI environment deployment are often expensive and time-consuming to complete. To solve this problem came the Agile ROLAP methodology, which in addition to using some OLAP visualization tools made for traditional methodologies, is also an alternative for small and medium organizations that seek to deploy a BI environment. The Agile ROLAP methodology uses tools and some plugins that were developed to facilitate the development of the BI environment. Therefore, this work seeks to analyze beyond the Agile ROLAP methodology, to analyze also the tools and the plugins that compose it.

**Keywords:** Information, Business Intelligence, Agile ROLAP.

## LISTA DE FIGURAS

<b>Figura 1:</b> Dado, informação, conhecimento e decisão .....	19
<b>Figura 2:</b> Data Warehouse e Data Mart.....	26
<b>Figura 3:</b> Modelo estrela. ....	27
<b>Figura 4:</b> Estrutura do esquema estrela, flocos de neve e multiestrela. ....	28
<b>Figura 5:</b> Arquitetura do Agile ROLAP.....	31
<b>Figura 6:</b> Fases da implantação do ambiente de BI com Agile ROLAP.....	32
<b>Figura 7:</b> Interface da ferramenta Schema Workbench.....	37
<b>Figura 8:</b> Diagrama da base de dados.....	39
<b>Figura 9:</b> Fato inscricao_perfil.....	40
<b>Figura 10:</b> Fato pagamento_inscrito.....	41
<b>Figura 11:</b> Fato resultado_inscrito. ....	41
<b>Figura 12:</b> Configuração do plugin tempo. ....	42
<b>Figura 13:</b> Configuração da dimensão inscrito. ....	43
<b>Figura 14:</b> SQL do fato inscricao_perfil. ....	44
<b>Figura 15:</b> Dimensões e chaves conectadas. ....	44
<b>Figura 16:</b> Configuração da dimensão tempo. ....	45
<b>Figura 17:</b> Medidas selecionadas no plugin fato. ....	45
<b>Figura 18:</b> Transformação do fato inscricao_perfil.....	46
<b>Figura 19:</b> Configuração do tempo do fato inscricao_perfil. ....	47
<b>Figura 20:</b> Configuração da dimensão lógica regio do fato inscricao_perfil. ....	48
<b>Figura 21:</b> Configuração das dimensões do cubo do fato inscricao_perfil .....	49
<b>Figura 22:</b> Configuração das medidas do cubo do fato inscricao_perfil.....	50
<b>Figura 23:</b> Configuração do plugin SchemaOut para o fato inscricao_perfil. ....	50
<b>Figura 24:</b> Estrutura da transformação criada no Kettle para o fato inscricao_perfil .....	51
<b>Figura 25:</b> Análise dos dados do cubo pagamento_inscrito.....	53
<b>Figura 26:</b> Análise dos dados do cubo resultado_inscrito.....	54
<b>Figura 27:</b> Análise dos dados do cubo inscricao_perfil. ....	55
<b>Figura 28:</b> Alteração das datas de nascimento de alguns inscritos. ....	56
<b>Figura 29:</b> SQL gerado pelo plugin FATO. ....	57
<b>Figura 30:</b> SQL utilizável para gerar uma view. ....	58
<b>Figura 31:</b> Declaração dos nomes das medidas no plugin fato. ....	58

<b>Figura 32:</b> Erro ao tentar utilizar plugin de dimensão lógica de tempo. ....	59
<b>Figura 33:</b> Schema final incorreto. ....	60
<b>Figura 34:</b> Plugin cubo exibindo medidas incorretas. ....	61
<b>Figura 35:</b> Schema final correto. ....	62
<b>Figura 36:</b> Erro ao selecionar o plugin FDW .....	63
<b>Figura 37:</b> Dados gerados apenas do primeiro semestre. ....	63
<b>Figura 38:</b> Alerta exibido ao conectar plugins de dimensão e fato. ....	64
<b>Figura 39:</b> SQL gerado através do plugin fato. ....	65
<b>Figura 40:</b> SQL gerado para o plugin fato.....	66
<b>Figura 41:</b> Transformação para usar plugin tempo .....	67
<b>Figura 42:</b> Atributos das tabelas.....	72
<b>Figura 43:</b> Atributos das tabelas.....	73
<b>Figura 44:</b> Atributos das tabelas.....	74
<b>Figura 45:</b> Atributos das tabelas.....	75
<b>Figura 46:</b> Tabela inscrição selecionada para a dimensão inscrito do fato inscricao_perfil. ..	76
<b>Figura 47:</b> Tabela curso selecionada para a dimensão curso do fato inscricao_perfil. ....	76
<b>Figura 48:</b> Atributos selecionados para a dimensão curso do fato inscrição_perfil.....	77
<b>Figura 49:</b> Tabela campus selecionada para a dimensão campus do fato inscrição_perfil. ....	77
<b>Figura 50:</b> Atributos selecionados para a dimensão campus do fato inscrição_perfil. ....	77
<b>Figura 51:</b> Tabela tempo_inscricao selecionada para a dimensão tempo_inscricao do fato inscrição_perfil. ....	78
<b>Figura 52:</b> Alguns dos atributos selecionados para a dimensão tempo_inscricao do fato inscrição_perfil. ....	78
<b>Figura 53:</b> Construtor de Query do fato inscricao_perfil .....	79
<b>Figura 54:</b> SQL da view gerado no plugin fato para o fato inscricao_perfil.....	79
<b>Figura 55:</b> Configuração da dimensão lógica campus para o fato inscricao_perfil. ....	80
<b>Figura 56:</b> Configuração da dimensão lógica curso para o fato inscricao_perfil.....	81
<b>Figura 57:</b> Configuração da dimensão lógica sexo para o fato inscricao_perfil. ....	82
<b>Figura 58:</b> Configuração da dimensão lógica cotas para o fato inscricao_perfil. ....	83
<b>Figura 59:</b> Configuração da dimensão lógica canhoto para o fato inscricao_perfil.....	84
<b>Figura 60:</b> Configuração da dimensão lógica nec_especial para o fato inscricao_perfil. ....	85
<b>Figura 61:</b> Configuração da dimensão lógica curso_superior para o fato inscricao_perfil.....	86
<b>Figura 62:</b> Configuração da dimensão logica etnia para o fato inscricao_perfil.....	87

<b>Figura 63:</b> Configuração da dimensão lógica lingua_estrangeira para o fato inscricao_perfil. .....	88
<b>Figura 64:</b> Configuração da dimensão lógica pagamento para o fato inscricao_perfil. ....	89
<b>Figura 65:</b> Tabela tempo_pagamento selecionada para a dimensão tempo_pagamento para o fato pagamento_inscrito. ....	89
<b>Figura 66:</b> Alguns dos atributos selecionados para a dimensão tempo_pagamento para o fato pagamento_inscrito. ....	90
<b>Figura 67:</b> Tabela curso selecionada para a dimensão curso para o fato pagamento_inscrito. .....	90
<b>Figura 68:</b> Atributos selecionados para a dimensão curso para o fato pagamento_inscrito. ..	91
<b>Figura 69:</b> Tabela inscricao selecionada para a dimensão inscrito para o fato pagamento_inscrito. ....	91
<b>Figura 70:</b> Atributos selecionados para a dimensão inscrito para o fato pagamento_inscrito.	91
<b>Figura 71:</b> Tabela campus selecionada para a dimensão campus para o fato pagamento_inscrito. ....	92
<b>Figura 72:</b> Atributos selecionados para a dimensão campus para o fato pagamento_inscrito. .....	92
<b>Figura 73:</b> Tabela pagamento selecionada para a dimensão pagamento para o fato pagamento_inscrito. ....	92
<b>Figura 74:</b> Atributos selecionados para a dimensão pagamento para o fato pagamento_inscrito. ....	93
<b>Figura 75:</b> Construtor de Query do fato pagamento_inscrito. ....	93
<b>Figura 76:</b> SQL do fato pagamento_inscrito para o fato pagamento_inscrito. ....	94
<b>Figura 77:</b> Dimensões do fato pagamento_inscrito. ....	94
<b>Figura 78:</b> Dimensões tempo do fato pagamento_inscrito. ....	94
<b>Figura 79:</b> Medidas do fato pagamento_inscrito. ....	95
<b>Figura 80:</b> SQL da view gerado no plugin fato para o fato pagamento_inscrito .....	95
<b>Figura 81:</b> Transformação do fato pagamento_inscrito. ....	96
<b>Figura 82:</b> Configuração da dimensão lógica curso para o cubo pagamento_inscrito. ....	96
<b>Figura 83:</b> Configuração da dimensão campus para o cubo pagamento_inscrito. ....	97
<b>Figura 84:</b> Configuração da dimensão região para o cubo pagamento_inscrito .....	98
<b>Figura 85:</b> Configuração da dimensão sexo para o cubo pagamento_inscrito. ....	99
<b>Figura 86:</b> Configuração da dimensão cotas para o fato pagamento_inscrito. ....	100
<b>Figura 87:</b> Configuração do cubo pagamento_inscrito. ....	101

<b>Figura 88:</b> Configuração das medidas do cubo pagamento_inscrito.....	101
<b>Figura 89:</b> Transformação do cubo pagamento_inscrito.....	102
<b>Figura 90:</b> Tabela curso selecionada para a dimensão curso para o cubo resultado_inscrito. .....	102
<b>Figura 91:</b> Atributos selecionados para a dimensão curso para o cubo resultado_inscrito..	102
<b>Figura 92:</b> Tabela inscricao selecionada para a dimensão inscrito para o cubo resultado_inscrito. ....	103
<b>Figura 93:</b> Atributos selecionados para a dimensão inscrito para o cubo resultado_inscrito. .....	103
<b>Figura 94:</b> Dimensões do fato resultado_inscrito para o cubo resultado_inscrito. ....	103
<b>Figura 95:</b> Transformação do fato resultado_inscrito para o cubo resultado_inscrito.....	104
<b>Figura 96:</b> Configuração da dimensão lógica região para o cubo resultado_inscrito. ....	104
<b>Figura 97:</b> Configuração da dimensão logica curso para o cubo resultado_inscrito.....	105
<b>Figura 98:</b> Configuração da dimensão logica sexo para o cubo resultado_inscrito .....	106
<b>Figura 99:</b> Configuração da dimensão lógica cota para o cubo resultado_inscrito.....	107
<b>Figura 100:</b> Configuração do cubo resultado_inscrito. ....	108
<b>Figura 101:</b> Configuração das medidas do cubo resultado_inscrito. ....	108
<b>Figura 102:</b> Transformação do cubo resultado_inscrito.....	109
<b>Figura 103:</b> Configuração dimensão lógica de tempo do cubo pagamento_inscrito .....	109

## LISTA DE SIGLAS

<b>API</b>	Application Programming Interface
<b>BI</b>	Business Intelligence
<b>CRM</b>	Customer Relationship Management
<b>DW</b>	Data Warehouse
<b>ER</b>	Entidade-relacionamento
<b>ERP</b>	Enterprise Resource Planning
<b>ERR</b>	Extended-ER ou Enhanced-ER
<b>ETL</b>	Extraction, Transform e Load
<b>FDW</b>	Foreign Data Wrappers
<b>MDX</b>	Multidimensional Expression
<b>MOLAP</b>	Multidimensional On Line Analytical Processing
<b>OLAP</b>	Online Analytical Processing
<b>OLTP</b>	Online Transaction Processing
<b>ROLAP</b>	Relational On Line Analytical Processing
<b>SGBD</b>	Sistema Gerenciador de Banco de Dados
<b>SQL</b>	Structured Query Language
<b>UENP</b>	Universidade Estadual do Norte do Paraná
<b>XML</b>	eXtensible Markup Language

# SUMÁRIO

LISTA DE FIGURAS .....	6
LISTA DE SIGLAS .....	10
1. INTRODUÇÃO .....	14
1.1 Objetivos.....	15
1.1.1 Objetivo Geral .....	15
1.1.2 Objetivos Específicos .....	15
1.2 Justificativa .....	16
1.3 Metodologia .....	17
2. FUNDAMENTAÇÃO TEÓRICA.....	19
2.1 Dados X Informação X Conhecimento X Decisão .....	19
2.2 BUSINESS INTELLIGENCE .....	20
2.3 DATA WAREHOUSE.....	22
2.3.1 Data Marts .....	25
2.4 Modelagem multidimensional/dimensional .....	26
2.5 ETL .....	29
2.6 Ferramentas OLAP .....	29
2.6.1 Servidores OLAP.....	29
2.7 <i>AGILE</i> ROLAP .....	30
2.7.1 Plugins .....	34
3. FERRAMENTAS .....	36
3.1 Pivot4J.....	36
3.2 Pentaho Data Integration (Kettle) .....	36
3.3 PostgreSQL .....	36
3.4 Mondrian Schema Workbench.....	37
4. DESENVOLVIMENTO .....	38
4.1 Levantamento dos dados e requisitos – 1º Passo .....	38
4.2 Acesso à fonte de dados – 2º Passo .....	38
4.3 Criação da base de dados intermediária – 3º Passo .....	42
4.4 Criação das dimensões de tempo – 4º Passo .....	42
4.5 Criação das dimensões convencionais – 5º Passo.....	43
4.6 Criação dos fatos – 6º Passo.....	44

4.7	Configuração das hierarquias e níveis nas dimensões de tempo – 7º Passo.....	46
4.8	Configuração das hierarquias e níveis nas dimensões convencionais – 8º Passo .....	47
4.9	Configuração dos cubos – 9º Passo .....	49
4.10	Exportação do Schema XML – 10º Passo .....	50
5.	RESULTADOS.....	52
5.1	Análise dos dados .....	53
5.2	Erros Encontrados .....	56
5.2.1	View gerada está incorreta .....	56
5.2.2	Plugin dimensão lógica de tempo não funcionando .....	59
5.2.3	Schema final incorreto.....	59
5.2.4	Plugin FDW .....	63
5.2.5	Erro ao gerar dados no plugin de dimensão tempo .....	63
5.3	Alterações e Melhorias .....	64
5.3.1	Tela de aviso.....	64
5.3.2	Construtor de query do <i>plugin fato</i> .....	64
5.3.3	Plugin tempo.....	67
6.	CONSIDERAÇÕES FINAIS .....	68
7.	TRABALHOS FUTUROS .....	69
	REFERÊNCIAS.....	70
	Apêndice A – Atributos das tabelas da base de dados .....	72
	Apêndice B – Configurações do desenvolvimento dos cubos .....	76

# 1. INTRODUÇÃO

Muitas empresas estão utilizando BI (*Business Intelligence*) para obter conhecimentos e *insights* a partir dos dados que possuem em seus sistemas corporativos e ainda gerarem relatórios e análises para executivos, gerentes e usuários da organização (ABELLERA, BULUSU, 2017, p.1).

Segundo Kimball e Ross (2013, p. 3) analisando o ambiente de empresas e observando os gestores é possível visualizar preocupações e definir alguns fatores que podem ser tratados com o BI e DW (*Data Warehouse*). Esse fatores e preocupações são:

- Grande volume de dados é coletado, mas não é acessível;
- Dificuldade ou impossibilidade para alterar a perspectiva de visualização e orientação dos dados (*slice and dice*);
- Dificuldade para obtenção e acesso aos dados por parte dos empresários;
- Informações e dados são mostrados sem qualquer tipo de seleção prévia, exibindo ao usuário muitas vezes informações e dados irrelevantes;
- Demora em reuniões ao invés de tomar decisões de maneira ágil;
- Realizar decisões sem apoio de informações e fatos.

Para lidar com essas diversas situações esse trabalho usa uma alternativa para a implantação de BI, principalmente para as pequenas empresas. A implantação será feita utilizando-se de uma metodologia ágil para o desenvolvimento de um ambiente de BI, chamada de *Agile ROLAP*. Os principais conceitos do método *Agile ROLAP* foram apresentados inicialmente por Souza (2014, p. 44) que afirma que o método “visa amenizar alguns dos problemas encontrados na tecnologia *Agile BI* e na implantação de ambientes de BI tradicionais” como a realização do ETL (*Extraction, Transform e Load*) e a compatibilidade com ferramentas OLAP de ambientes de BI tradicionais. Além disso, o método foi utilizado por Aguiar (2016, p. 71) o qual declara que o *Agile ROLAP* “é uma boa opção para projetos que não envolvem a integração de diferentes bases de dados”. Isso devido a etapa de ETL que o método não realiza como nos métodos tradicionais.

Esse método diferencia das tradicionais por ser mais ágil e acessível, mas também se diferencia de outras metodologias ágeis por ser compatível com

ferramentas de BI já disponíveis no mercado, que utilizem o servidor ROLAP (Relational OLAP) Mondrian<sup>1</sup>.

Esse trabalho apresenta então uma análise dessa metodologia ágil de implantação de BI e algumas de suas ferramentas (*plugins* desenvolvidos) a partir de uma base de dados de vestibular. Portanto, o ambiente de BI a ser desenvolvido será para análise de dados oriundos do vestibular realizado na Universidade Estadual do Norte do Paraná (UENP).

A organização do trabalho está dividida entre tópicos e subtópicos. Nos seguintes subtópicos do capítulo 1 será apresentado os objetivos e a justificativa da realização desse trabalho. Logo em seguida é apresentado a metodologia de desenvolvimento do presente trabalho. No capítulo 2 é descrito a fundamentação teórica, que são os conceitos necessários para entender e desenvolver o trabalho. No capítulo 3 é apresentado brevemente as ferramentas que serão utilizadas no desenvolvimento. No capítulo 4 é descrito passo a passo o desenvolvimento do projeto principal desse trabalho. Já no capítulo 5 é apresentado e discutido os resultados obtidos com o desenvolvimento desse trabalho.

## 1.1 Objetivos

Nos tópicos seguintes estão descritos os objetivos deste trabalho.

### 1.1.1 Objetivo Geral

O objetivo geral deste trabalho é analisar o método *Agile* ROLAP e implantar um ambiente de *Business Intelligence* com auxílio de alguns *plugins* que foram desenvolvidos para este propósito. O resultado desse processo tornará possível a análise de dados para obter *insights* do vestibular realizado na universidade e auxiliará na análise e melhoria do método.

### 1.1.2 Objetivos Específicos

- ✓ Selecionar uma base de dados;
- ✓ Implantar um ambiente de BI;

---

<sup>1</sup> <https://mondrian.pentaho.com/>

- ✓ Analisar os *plugins* durante a implantação do ambiente;
- ✓ Explicar os processos realizados durante a implantação;
- ✓ Apresentar e analisar possíveis *insights* após a implantação do ambiente de BI com ferramentas OLAP (*Online Analytical Processing*);

## 1.2 Justificativa

Para suprir a necessidade de organizações que buscam uma alternativa mais acessível financeiramente, esse trabalho apresentará a implantação de um ambiente de BI, utilizando um método diferente daqueles comumente usados. Conforme Carlos Barbieri (2011, p. 95) “BI, de forma mais ampla, pode ser entendido como a utilização de variadas fontes de informação para definir estratégias de competitividade nos negócios da empresa”. Uma abordagem que utiliza de ferramentas para organizar, tratar e transformar os dados em *insights* importantes e valiosos.

Devido à grande quantidade de informação que as empresas têm acesso atualmente o “BI passa então a ser primordial para as organizações poderem funcionar de forma "organizada" e não se afogarem com tanta informação” (LOH, 2014, p.16). Além disso, a construção de um *data warehouse* corporativo é uma tarefa complexa e que depende muito em termos de tempo e recursos (VAISMAN; ZIMÁNYI, 2014, p.79). Visando contornar esses tipos de problemas e tornar possível para que um número maior de organizações consiga aproveitar destas novas ferramentas, este trabalho pretende apresentar e analisar o método *Agile* ROLAP através da implantação de um ambiente inteligente de negócios para análise de dados de um vestibular.

Segundo Souza (2014 p. 44), em seu trabalho que apresenta este método, o *Agile* ROLAP foi desenvolvido para servir como uma alternativa para os dois tipos de metodologias de implantação de BI já existentes, o *Agile* BI e o BI tradicional. O autor ainda enfatiza que as principais diferenças entre as metodologias existentes e essa é a realização do processo de ETL e a capacidade de utilizar ferramentas que foram desenvolvidas para o BI tradicional.

O principal problema das metodologias tradicionais, o qual essa metodologia procura solucionar, é o tempo para realizar todo o processo de ETL. Segundo

Kimball e Caserta (2004) ela demanda facilmente 70% dos recursos necessários para a implementação e manutenção de um *data warehouse* típico.

Além deste método reduzir o tempo e conseqüentemente o custo no processo de ETL, este utiliza de ferramentas *open source* e possui algumas funcionalidades já prontas e algumas que podem ser desenvolvidas, através do desenvolvimento de *plugins*, dependendo das necessidades das empresas. Essa opção traz uma certa economia e uma quantidade maior de funcionalidades para a organização. Com esse tipo de tecnologia as organizações têm a sua disposição uma opção viável para utilizar ferramentas tecnológicas e poderem usufruir de todos os dados que são gerados para tomar as melhores decisões possíveis.

### 1.3 Metodologia

O presente trabalho visa apresentar a implantação de um ambiente de *Business Intelligence* fazendo uso do método ágil *Agile* ROLAP. Além disso é classificado como uma pesquisa explicativa por dois principais motivos, o primeiro, segundo Wazlawick (2014, p. 22), é que “além de analisar os dados observados, busca suas causas e explicações, ou seja, os fatores determinantes desses dados”. O segundo é por ser a continuação de uma outra pesquisa descritiva, já que para identificar os fatores e características de um fenômeno é preciso previamente que este seja bem descrito e detalhado (GIL, 2002, p. 42-43). Em suma, será implantado um ambiente de BI utilizando o método *Agile* ROLAP e avaliando os resultados e os benefícios que o método pode proporcionar às organizações e, caso haja, aos seus departamentos.

As informações base a respeito do *Agile* ROLAP, para o desenvolvimento deste trabalho utilizou-se de definições que advêm de outra pesquisa presente na literatura, um trabalho de conclusão de curso, como mencionado anteriormente. Como ponto de partida para esta pesquisa buscou-se informações em livros e em outros artigos científicos sobre as metodologias tradicionais e as ágeis existentes. Para o desenvolvimento foi necessário buscar por alguns assuntos específicos sobre como realizar uma modelagem de dados, de como construir um *Data Warehouse* e de como utilizar uma ferramenta OLAP.

O método *Agile* ROLAP é o objeto de estudo deste trabalho. As variáveis utilizadas no trabalho provêm da base de dados escolhida e como maneira de

controlar e observar os resultados será feito de início a modelagem dos dados, etapa crucial e importante, e uso de ferramentas durante todo o processo de implantação. Devido a esses procedimentos técnicos citados, este trabalho pode ser também definido como uma pesquisa do tipo experimental. A qual o constitui o "mais valioso procedimento disponível aos cientistas para testar hipóteses que estabelecem relações de causa e efeito entre as variáveis" (GIL, 2002, p. 49).

Pretende-se utilizar de uma base de dados de vestibular da UENP e por meio dela realizar a modelagem desses dados. Os dados serão acessados em sua fonte, sem alterações e de forma relacional. Porém serão interpretados de maneira dimensional para que a ferramenta OLAP possa gerar as informações, através de gráficos, quadros, etc.

Para realizar o trabalho será utilizado algumas ferramentas como o Kettle (ferramenta livre do Pentaho), Pivot4J (um ambiente de BI *open source*), *plugins* próprios da metodologia e um banco de dados. Essas ferramentas serão utilizadas para criar o ambiente de BI e para apresentar os resultados da transformação de dados irrelevantes para informações úteis. Conseqüentemente os resultados são categorizados de origem de uma pesquisa quantitativa, pois "a realidade só pode ser compreendida com base na análise de dados brutos, recolhidos com o auxílio de instrumentos padronizados e neutros" (JOÃO, 2002, p. 20).

## 2. FUNDAMENTAÇÃO TEÓRICA

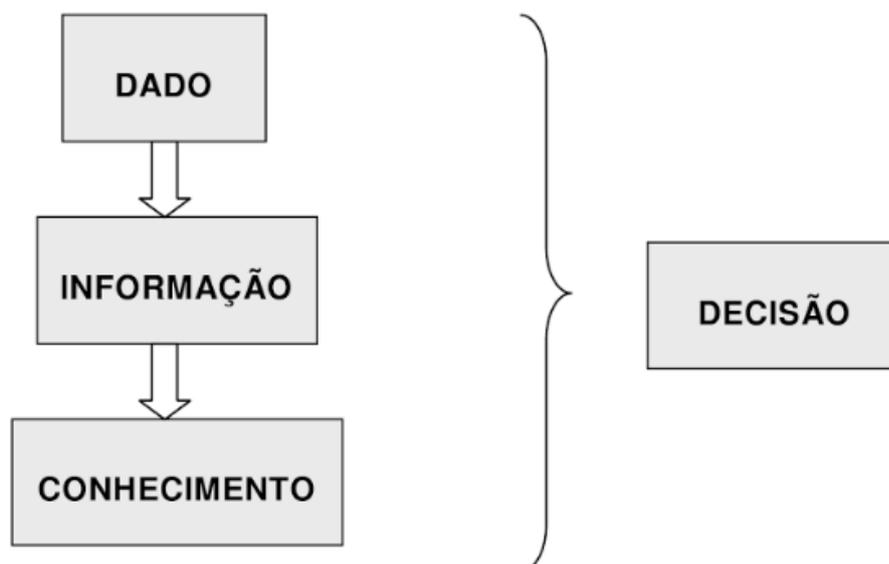
Esta seção tem por finalidade apresentar os conceitos presentes na literatura, que serão necessários para o desenvolvimento deste projeto.

### 2.1 Dados X Informação X Conhecimento X Decisão

Para facilitar a compreensão e entendimento do método, das ferramentas e de outros conceitos que compõem a fundamentação teórica, é importante ter um conhecimento prévio das definições de dados, informações, conhecimento e decisão.

Segundo Siqueira (2005, p. 24) dados podem ser facilmente armazenados e obtidos através de máquinas. São partículas de registros estruturados e sua forma possui alguns padrões como: datas, códigos, números, etc. Esse mesmo autor também define informação e conhecimento. O primeiro é definido como um dado que possui contexto, propósito e que seja relevante. Geralmente é um conjunto de dados, que isolados não possuem sentido, mas agrupados transmitem algum entendimento. Já o conhecimento, ainda segundo Siqueira, pode ser definido como um conjunto de informações da mente humana e estão ligados a fatores como experiências, emoções, insights, etc.

**Figura 1:** Dado, informação, conhecimento e decisão.



Fonte: Decisões com B.I (2008).

Para Primak (2008, p. 12-13), como apresentado na Figura 1, a decisão é um resultado que advém dos dados, da informação e do conhecimento. Após adquirir conhecimento, a última etapa de desenvolvimento de um dado, o gestor decide o que fazer com ele e se tomará uma decisão ou não. O autor afirma também que em diversas situações, quando técnicas; conceitos e sistemas de BI não são utilizados, os profissionais que tomam decisões obtêm apenas dados sem valor algum. Tornando assim uma tarefa difícil e arriscada tomar uma decisão.

Segundo Laudon e Laudon (2011, p. 11), é muito comum administradores trabalharem sem contar com as informações que precisam no momento de tomar uma decisão, alguns, devido à falta de informações, trabalham com base em previsões, palpites e até mesmo na sorte. Para esses autores as consequências disso podem ser várias, incluindo um controle deficiente da produção de bens e serviços, tempos de respostas ineficientes e a má distribuição dos recursos. De acordo ainda com ambos os autores, como resultado dessas falhas os custos dos bens e serviços aumentam e clientes são perdidos.

Com ferramentas, técnicas e metodologias, utilizadas em conjunto, é possível que esses profissionais consigam utilizar e ter acesso à uma quantidade imensa de dados, tanto de fontes internas quanto externas, para fazerem a melhor decisão possível de forma ágil, simples e assertiva.

## **2.2 BUSINESS INTELLIGENCE**

O termo *Business Intelligence* teve sua origem por volta de década de 80 no Gartner Group e consiste no processo inteligente de coletar, organizar, analisar, compartilhar e monitorar os dados contidos no *Data Warehouse* (PRIMAK, 2008, p. 2).

Sistemas de *Business Intelligence* apresentam informações empresariais e competitivas para aqueles que realizam a tomada de decisão com objetivo de melhorar o processo de tomada de decisão (NEGASH, 2004).

A análise de dados tem se tornado cada vez mais importante desde os anos 90 à medida que as organizações de todos os setores estão sendo obrigadas a melhorar seus processos de tomada de decisão para manter sua vantagem competitiva (VAISMAN; ZIMÁNYI, 2014). Porém, desde antigamente os princípios básicos do BI eram usados pela sociedade do oriente médio para analisar dados e

informações que seriam úteis para suas aldeias como a análise das marés, dos astros, dos períodos chuvosos e de seca, etc. (PRIMAK, 2008, p. 1).

Muitas empresas geram dados e não as utilizam, ou as utilizam de forma inapropriada e segundo Barbieri (2011, p.25) “a proposta de BI é transformar dados em informações que possam ser usadas para ações analíticas, tomadas de decisões tático-estratégicas e até definições operacionais”.

De acordo com Turban et al. (2009, p. 25-26) o BI possibilita que o gerente e gestores de organizações realizem ações como:

- Emprego de planejamento estratégico;
- Uso de modelos de negócios novos e inovadores;
- Reestruturação de processos de negócios;
- Participação em aliança de negócios;
- Aprimoramento dos sistemas de informação corporativos;
- Aprimoramento das relações de parceria;
- Incentivo à inovação e à criatividade;
- Aprimoramento do serviço e relacionamento com cliente;
- Adoção do comércio eletrônico;
- Adoção da manufatura sob encomenda e da produção e serviços sob demanda;
- Uso de novas tecnologias de informação para melhorar a comunicação, o acesso aos dados (descoberta de informações) e a colaboração;
- Rápida reação às ações dos concorrentes, como em precificação, promoções e novos produtos e serviços;
- Automatização de diversas tarefas do pessoal administrativo;
- Automatização de certos processos de decisão, principalmente os que envolvem clientes;
- Aprimoramento da tomada de decisões.

Com o seu uso, as organizações conseguem, com mais facilidade e confiança, realizar uma decisão estratégica para competir em um novo mercado, alterar o foco da empresa do produto para o cliente e até lançar uma nova linha de produtos (WATSON; WIXOM, 2007, p. 97). Com a implantação bem-sucedida o BI comprovadamente tem um potencial enorme para melhorar os lucros e o

desempenho da organização, porém se for implantado errado, será um desperdício de dinheiro e de tempo (WILLIAMS; WILLIAMS, 2007, p.2).

### **2.3 DATA WAREHOUSE**

*Data Warehouse* ou em português, armazém de dados, “é um conjunto de dados produzido para oferecer suporte à tomada de decisões; é também um repositório de dados atuais e históricos de possível interesse aos gerentes de toda a organização” (TURBAN et al., 2009, p.57).

O DW é um processo no qual se obtém dados de sistemas de banco de dados legados e de transacionais e os transforma em informações organizadas em um formato amigável para serem analisadas e para auxiliar no apoio a tomada de decisões de negócios (CASERTA; KIMBALL, 2004, p.22-23).

Abaixo, no Quadro 1 e após ela, há uma comparação detalhada das principais diferenças que há entre os dois tipos de banco de dados segundo definições feitas por Vaisman e Zimányi (2014), no qual o banco de dados operacional geralmente é utilizado em sistemas OLTP (*Online Transaction Processing*) e *Data Warehouses* por sistemas OLAP.

**Quadro 1:** Comparação entre banco de dados operacionais e Data Warehouses.

<b>Aspecto</b>	<b>Banco de Dados Operacional</b>	<b>Data Warehouses</b>
Tipo de usuário	Operadores, funcionários do escritório	Gerentes, executivos
Uso	Previsível, repetitivo	Ad hoc, não estruturado
Conteúdo dos dados	Atual, dados detalhados	Histórico, dados resumidos
Organização dos dados	De acordo com as necessidades operacionais	De acordo com as necessidades de análise
Estrutura dos dados	Otimizado para pequenas transações	Otimizado para consultas complexas
Frequência de acesso	Alta	De média a baixa
Tipo de acesso	Ler, inserir, atualizar e excluir	Apenas leitura
Número de registros por acesso	Poucos	Muitos
Tempo de resposta	Curto	Pode ser longo
Nível de concorrência	Alto	Baixo
Utilização de bloqueio	Necessária	Não é necessária
Frequência de atualizações	Alta	Nenhuma
Redundância de dados	Baixa (tabelas normalizadas)	Alta (tabelas desnormalizadas)
Modelagem de dados	UML, modelo ER	Modelo dimensional

Fonte: Adaptado de Data Warehouse Systems Design and Implementation (2014).

Usuários de sistemas OLTP, os quais utilizam de bancos de dados operacionais, geralmente já possuem as operações a serem realizadas pré-

definidas. Como exemplo podem ser citados os sistemas de folha de pagamento. Já os usuários de *Data Warehouses* costumam se posicionar em posições hierárquicas mais altas nas organizações do que os usuários de sistemas OLTP. Além disso os usuários de DW utilizam de ferramentas OLAP para análise dos dados e poderia, por exemplo, ser facilmente empregada para detectar irregularidades em salários de um sistema de folha de pagamento. Em sistemas OLTP os dados são atualizados e detalhados enquanto para a análise de dados é necessário dados históricos resumidos. A organização dos dados depende do sistema a ser usado. As estruturas de dados para OLTP são feitas com o objetivo de atender à várias transações pequenas e simples que acontecem com frequência. Já as estruturas para serem analisadas devem suportar consultas complexas com acessos a registros em uma ou mais tabelas. Além disso esses sistemas não são acessados com tanta frequência como os OLTP. O tipo de acesso é realizado através de operações que permitem inserir, ler, deletar e atualizar dados, já no outro tipo de banco apenas a operação de leitura é possível.

O número de registros acessados durante uma transação por um sistema OLTP geralmente é pequeno enquanto transações em DW envolve muitos registros. As consultas em sistemas OLTP costumam ser concluídas em um curto período de tempo, entretanto as consultas mais complexas realizadas com um DW podem levar um tempo maior. Em sistemas OLTP, durante sua utilização, pode acontecer de haver um elevado número de acessos concorrentes e para assegurar que o processamento dessas transações ocorra de forma segura é necessário utilizar de bloqueio ou outras formas de gerenciamento de simultaneidade. Em sistemas OLAP o número de acessos normalmente é baixo e por utilizarem apenas a operação de leitura as consultas podem ser realizadas simultaneamente sem precisar de bloqueio ou alguma outra solução.

Sistemas OLTP são atualizados frequentemente online enquanto OLAP é atualizado offline e periodicamente. O primeiro sistema citado é modelado utilizando padrões do UML (Linguagem de Modelagem Unificada) ou até mesmo alguma variação do modelo de ER (entidade-relacionamento). Esses modelos permitem que seja construído um esquema normalizado para esse tipo de banco de dados que precisa suportar transações frequentes. Além disso, esses modelos auxiliam a garantir a consistência do banco e a reduzir a redundância. Em sistemas OLAP é utilizado a modelagem multidimensional o qual proporciona um banco de dados

desnormalizado e com alto nível de redundância, o qual melhora o processamento de consultas.

Segundo Machado (2007, p. 28-31) com as características e objetivos bem definidos é possível realizar um projeto de DW, em que é construído a base de alguns conceitos. Sendo eles:

**Orientação por assunto:** indica que o DW armazena as informações organizadas por assuntos. Diferente dos dados armazenados em sistemas transacionais, que buscam apenas o controle operacional, o DW possui dados mais específicos e selecionados, que resultarão em informações importantes e conseqüentemente melhores tomadas de decisões.

**Variação de tempo:** o tempo é uma característica essencial para o DW. O DW armazena os dados operacionais que ocorreram em um determinado momento do tempo, como se houvesse realizado uma foto. Em um sistema transacional é possível realizar alterações, já no DW não é feito. Se necessário, é possível criar outra “foto”, com características semelhantes ao de origem e com data atualizada, mas o original não sofre alteração alguma. Além disso, o sistema DW geralmente utiliza de dados de um período de vários anos, enquanto o transacional costuma utilizar dados pertencentes à períodos curtos e geralmente diários.

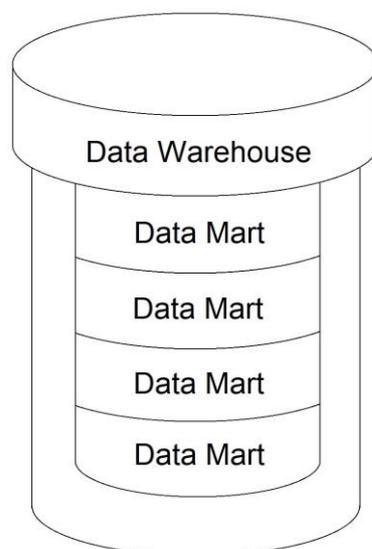
**Não volátil:** os dados pertencentes à sistemas transacionais possuem a possibilidade de serem alterados, enquanto os armazenados em um DW não. O sistema de DW permite apenas a carga de novos dados e consulta aos já armazenados, ou seja, mantém seus dados inalterados.

**Integração:** os dados que compõem um DW têm origem de sistemas isolados e realizar a integração destes é fundamental para o funcionamento de um sistema de DW. Essa característica realiza uma “padronização” dos dados e é desenvolvida após os processos de filtragem e agregação que ocorre durante a carga dos dados.

### 2.3.1 Data Marts

Machado (2007, p. 44) define *Data Marts*, repositório de dados em português, como um subconjunto de dados do *Data Warehouse*, como demonstrado na Figura 2. O autor cita também que uma das vantagens de o utilizar é um *feedback* mais rápido que do DW, por parte do usuário e principalmente durante a implantação, tornando assim mais fácil de se analisar os benefícios da ferramenta.

**Figura 2:** Data Warehouse e Data Mart.



Fonte: Adaptado de Tecnologia e Projeto de Data Warehouse (2007).

Watson e Wixon (2007, p.99) acrescentam que diferente de um DW, o *Data Mart* possui um alcance mais específico do que um armazém de dados e pode abranger uma área específica, uma região geográfica, uma aplicação ou até mesmo uma divisão/setor da organização.

## 2.4 Modelagem multidimensional/dimensional

A modelagem de dados dimensional é uma das etapas iniciais a ser realizada durante o desenvolvimento de um projeto de banco de dados para um ambiente de BI. Além disso, essa “fase é muito importante, pois é por meio dela que são levantadas as perguntas a que se deseja responder, com a análise dos dados e informações contidas no DW e também como ele será organizado” (CECI, 2012, p.82).

A modelagem dimensional aumenta a agilidade nas consultas ao *Data Warehouse*, entretanto traz consigo a necessidade de se ter mais espaço de armazenamento do que é necessário em um modelo de entidade-relacionamento (BALLARD et al., 2006, p.52).

Machado (2007, p.79) afirma que um modelo dimensional é composto por três componentes:

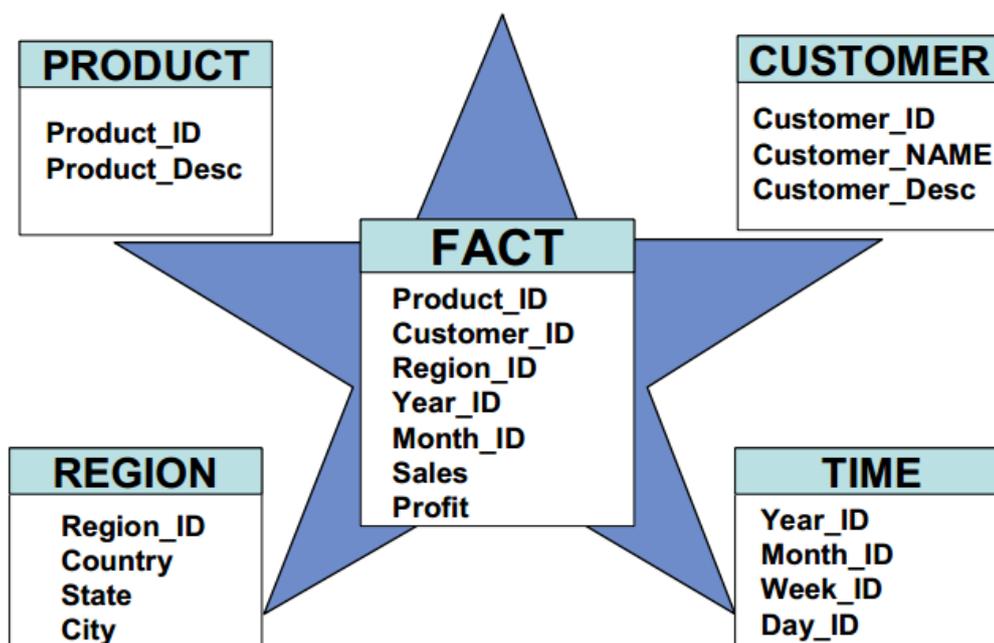
**Fatos:** é a representação de uma atividade, item, transição ou até mesmo um evento que ocorre em um negócio, geralmente do dia a dia.

**Dimensões:** são os elementos que compõem a tabela fato e determinam um contexto de um assunto de negócios. Para uma análise de vendas a partir de um banco de dados, por exemplo, as dimensões poderiam ser o tempo, a localização, os cenários, os clientes e os vendedores.

**Medidas (variáveis):** são os atributos pertencentes às tabelas fatos e dimensões. Como exemplo pode ser citado valor em reais de uma compra, a quantidade de produtos comprados, etc.

O modelo estrela é muito utilizado para a realização de uma modelagem dimensional. Abaixo, Figura 3, um exemplo deste modelo, no qual há uma tabela fato no meio do modelo, há também as tabelas dimensões produto, região, tempo e cliente nas extremidades e cada uma delas possuem suas variáveis. No final de sua composição há um modelo que possui um design semelhante a uma estrela. No qual a tabela fato é o centro e as dimensões as pontas da estrela.

Figura 3: Modelo estrela.

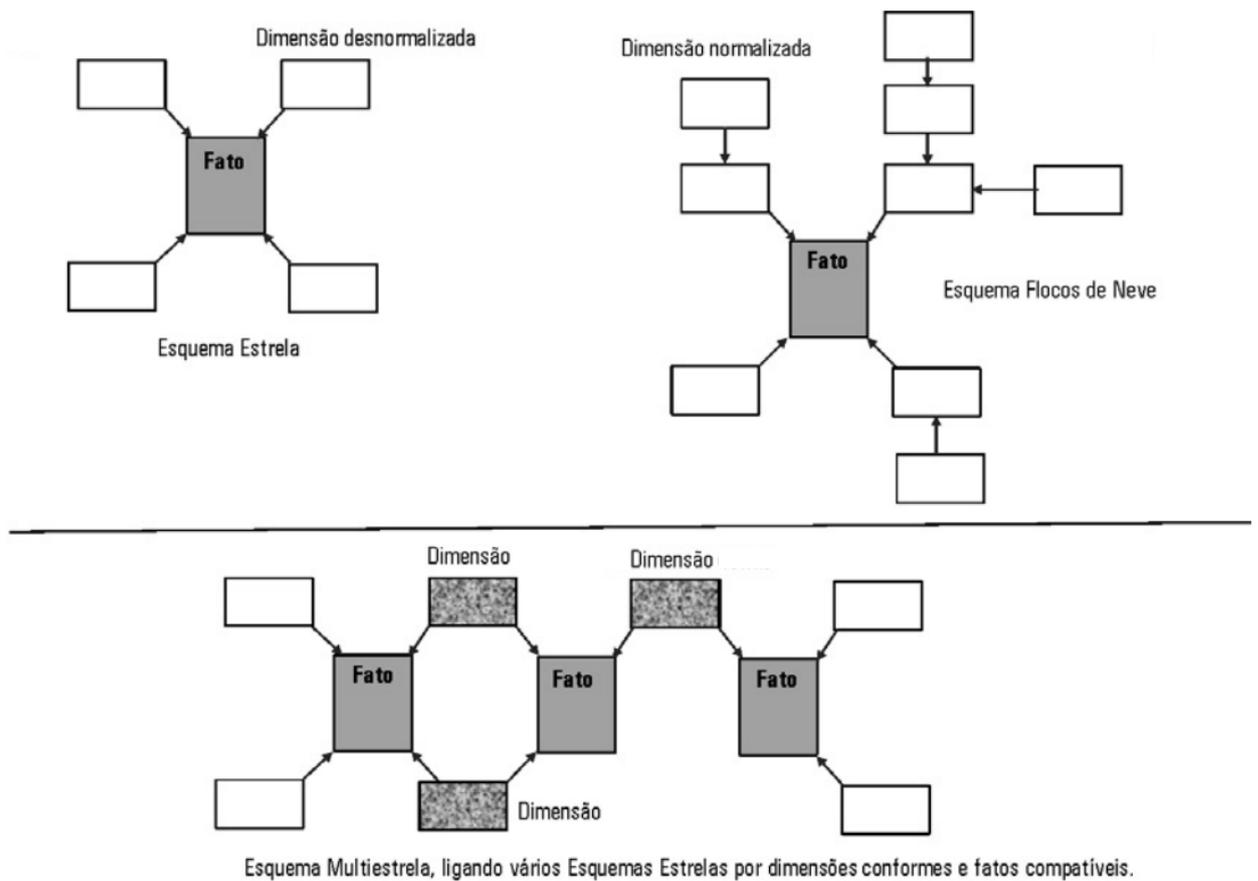


Fonte: Dimensional Modeling In a Business Intelligence Environment (2006).

Na literatura há mais alguns modelos, extensões do modelo estrela, como o *snowflake* e a multiestrela. Segundo Machado (2007, p. 95) o modelo *snowflake*, floco de neve em português, tem origem a partir da decomposição de alguma tabela dimensão. O autor afirma também que esse modelo é o resultado da aplicação da terceira forma normal sobre as dimensões e ainda é interpretado e entendido com

facilidade por desenvolvedores de sistemas OLTP. Na Figura 4, por exemplo, é possível ver que algumas tabelas dimensões estão “estendidas” no modelo *snowflake*. Diferente do modelo estrela esse modelo é normalizado e evita a redundância. O terceiro tipo, segundo Ballard et al. (2006, p. 57), o esquema multiestrela é um modelo dimensional que é formado por múltiplas tabelas fatos, unidas por dimensões.

**Figura 4:** Estrutura do esquema estrela, flocos de neve e multiestrela.



Fonte: Adaptado de BI2 - Business Intelligence (2011).

O modelo mais indicado é o estrela, pois “a utilização do esquema estrela é extremamente recomendável, pelos aspectos de ganhos de desempenho, quando comparada com o esquema de flocos de neve” (BARBIERI, 2011, p. 169). A redundância que há no esquema estrela possibilita uma diminuição nos comandos de junção das tabelas que são necessários para buscar informações em outras tabelas.

## 2.5 ETL

O processo de ETL é a base do *Data Warehouse* e executa a extração dos dados de outros sistemas, transforma os dados melhorando a sua qualidade e seus padrões, além de ajustar para que os dados de fontes diferentes possam ser usados juntos e por fim entregues em um formato pronto para o desenvolvimento da aplicação de BI (KIMBALL; CASERTA, 2004).

## 2.6 Ferramentas OLAP

OLAP é definido como uma coleção de ferramentas que permitem explorar os dados contidos em um *Data Warehouse*. Segundo Machado (2007, p.86), essas ferramentas “são as aplicações às quais os usuários finais têm acesso para extrair os dados de suas bases e construir os relatórios capazes de responder às suas questões gerenciais”. Segundo ainda esse autor existe seis tipos de operações usados em OLAP para a análise dos dados, que são:

***Drill Down and Roll Up:*** a primeira operação permite que o usuário consiga visualizar mais detalhes das informações que está analisando, enquanto o segundo funciona de forma contrária, ou seja, diminui os detalhes das informações.

***Drill Across:*** essa operação permite que pule um nível dentro de uma dimensão, ou seja, avançar do nível dia para o ano, pulando dessa forma o nível semana e/ou mês por exemplo.

***Drill Throught:*** possibilita analisar uma informação de uma dimensão e na sequência analisar uma outra pertencente a uma dimensão diferente. Por exemplo, analisar a dimensão tempo e posteriormente analisar a dimensão cliente.

***Slice and Dice:*** essas operações modificam a orientação dos dados que estão sendo visualizados. *Slice* “aprofunda” a visualização dos dados, porém não altera a perspectiva de visualização destes. A operação *Dice*, diferente do *Slice*, altera a perspectiva da visualização dos dados.

### 2.6.1 Servidores OLAP

Dos diversos tipos de servidores OLAP, existentes no mercado, “os dois servidores OLAP mais utilizados são o Multidimensional OLAP e o Relational OLAP

e a grande diferença entre eles é a forma como os dados são armazenados” (SILVA; SAIAS, 2012, p. 5). O servidor MOLAP (*Multidimensional On Line Analytical Processing*) baseia que se o cubo for multidimensional os dados dever ser armazenados de forma também multidimensional e o ROLAP (*Relational On Line Analytical Processing*) como o próprio nome indica é baseado no modelo relacional e as consultas são realizadas diretamente no DW e não no cubo como é realizado pelo servidor MOLAP (WESTERLUND, 2008, p. 17).

Nesse projeto será utilizado o Mondrian como servidor ROLAP. Mondrian é escrito na linguagem Java, executa consultas utilizando a linguagem MDX (*Multidimensional Expression*) e realiza a leitura de dados de um banco de dados relacional. Além disso apresenta os resultados em um formato multidimensional (HYDE, 2006).

## **2.7 AGILE ROLAP**

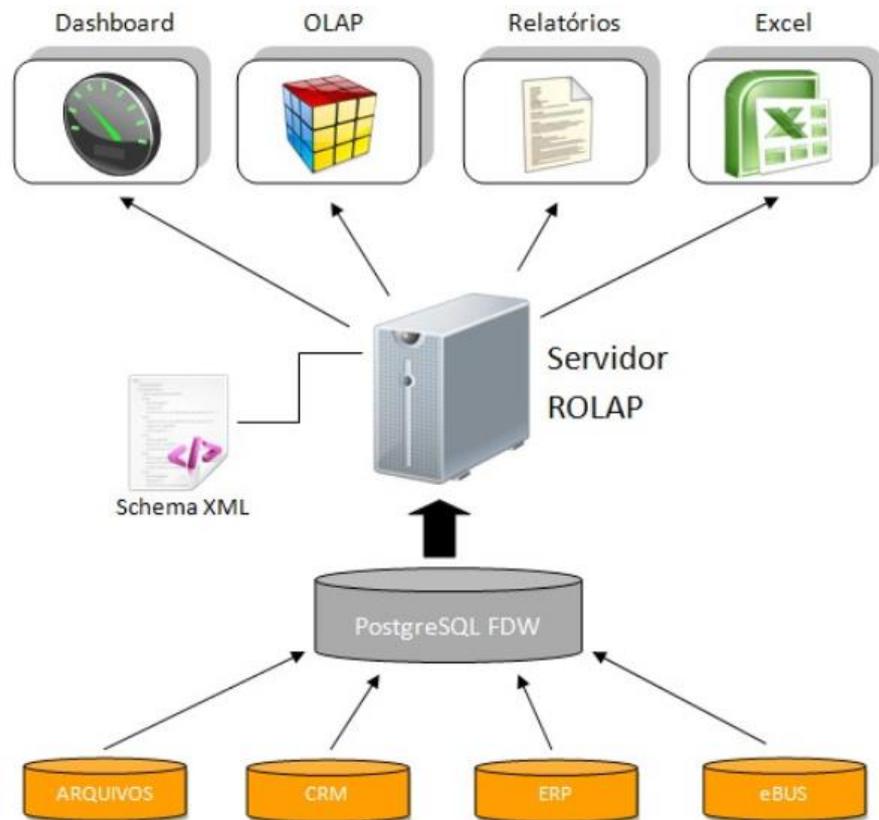
É um método que propõe algumas soluções para os problemas encontrados nas metodologias tradicionais e ágeis de desenvolvimento de BI. Ela aborda todo o processo de construção de um ambiente de BI, porém diferencia dos outros métodos na etapa de ETL e na compatibilidade com ferramentas OLAP.

Sua implantação é menos custosa que as tradicionais e mais ágil. Por esses dois motivos ela é indicada principalmente para organizações de pequeno porte.

Em suma, a metodologia *Agile* ROLAP permite a implantação de um ambiente de BI de forma ágil, no qual o ambiente utiliza de um banco de dados relacional. Esse método ainda possibilita a utilização de ferramentas OLAP que foram desenvolvidas para serem usadas em métodos tradicionais. A utilização dessas ferramentas é possível devido ao uso de um servidor ROLAP (Souza, Menolli e Coelho, 2014, p.44).

Na Figura 5 é exemplificado a arquitetura do *Agile* ROLAP com seus principais componentes.

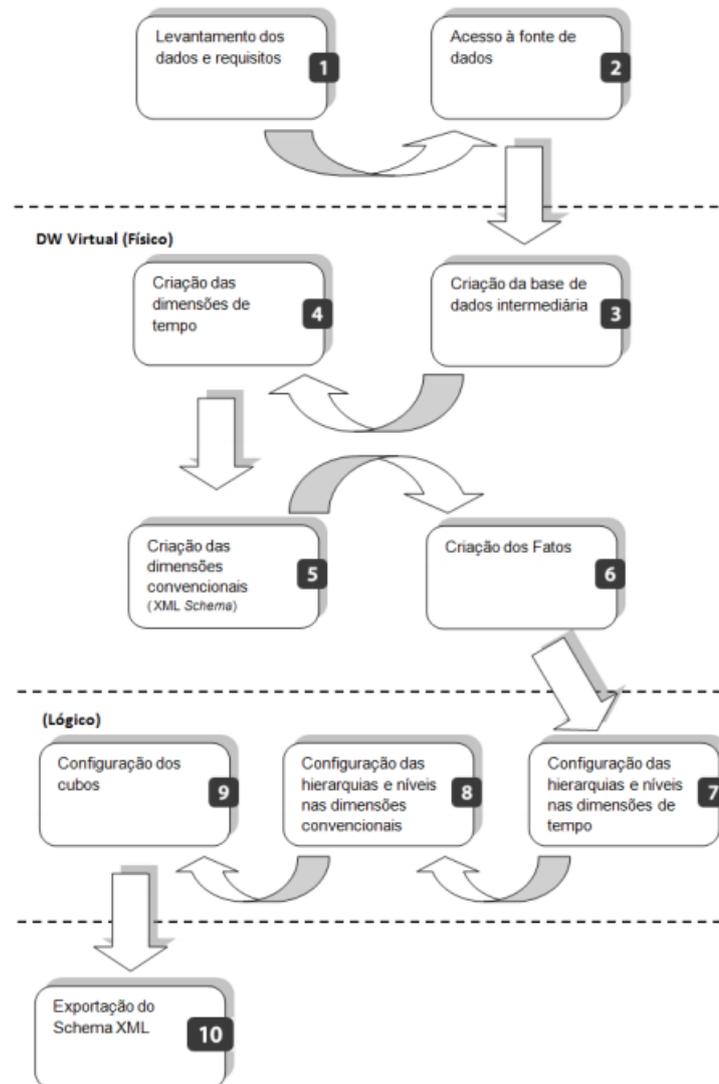
**Figura 5:** Arquitetura do *Agile* ROLAP.



Fonte: Um Método *Agile* ROLAP para implantação de Ambientes de Inteligência de Negócios (2014).

Consiste em uma ou mais fontes de dados que são a “matéria prima” do projeto de BI. Essas fontes, a base da arquitetura, podem ser arquivos e/ou dados, de sistemas CRM (*Customer Relationship Management*), de sistemas ERP (*Enterprise Resource Planning*), etc. Entre a base da arquitetura e o servidor ROLAP há uma base intermediária. Essa base é carregada com todos os dados para que não seja necessário acessar diretamente as bases de dados operacionais dos sistemas OLTP e permitir somente a operação de leitura para os usuários da aplicação de BI. O servidor ROLAP e o *schema XML* (*eXtensible Markup Language*) são os responsáveis por permitir o acesso às informações relacionais com uma perspectiva dimensional. Por último, no topo da arquitetura se encontra a ferramenta OLAP e outros meios nos quais as informações podem ser apresentadas como em relatórios, *dashboards* e ferramentas de análise de dados.

**Figura 6:** Fases da implantação do ambiente de BI com *Agile* ROLAP.



Fonte: Um Método *Agile* ROLAP para implantação de Ambientes de Inteligência de Negócios (2014).

A implantação do ambiente de BI através do método *Agile* ROLAP ocorre através de dez fases e é separado em duas principais etapas, a primeira etapa é a **física** no qual é selecionado os dados que deseja e a segunda etapa é a utilização desses dados para estruturar de forma **lógica** as dimensões e medidas que serão utilizadas na ferramenta OLAP. As fases e etapas são demonstradas na Figura 6 acima e descritos abaixo.

**Fase 1 - Levantamento dos dados e requisitos:** etapa essencial da implantação, no qual procura conhecer as características da organização e quais são as suas necessidades.

**Fase 2 - Acesso à fonte de dados:** nessa etapa é realizado o acesso às fontes de dados e conseqüentemente uma análise mais profunda dos requisitos e

necessidades. Além disso é analisado também se há ou haverá dificuldade em atender os desejos da organização.

**Fase 3 - Criação da base de dados intermediária:** nessa fase é iniciado a etapa física do desenvolvimento do DW. O método *Agile* ROLAP, nessa fase, utiliza dos dados relacionais para criar uma base de dados intermediária. Para isso é utilizado a tecnologia *Foreign Data Wrappers* (FDW) em conjunto com o banco de dados PostgreSQL.

**Fase 4 - Criação das dimensões de tempo:** nessa fase é criado as dimensões de tempo para análise histórica dos dados. Após criar uma dimensão uma tabela também é criada na base de dados intermediária.

**Fase 5 - Criação das dimensões convencionais:** da mesma forma que é preciso criar dimensões de tempo também é preciso criar dimensões convencionais. Essas dimensões descrevem as atividades que ocorrem na organização e diferente do que ocorre na Fase 4, nessa fase não é gerado tabelas como consequência da criação das dimensões.

**Fase 6 - Criação dos fatos:** nessa fase é criado os fatos do ambiente de BI e também suas *Views* na base de dados intermediária.

**Fase 7 - Configuração das hierarquias e dos níveis nas dimensões de tempo:** na Fase 7 é iniciado o desenvolvimento da parte lógica do ambiente de BI configurando as hierarquias e os níveis das dimensões tempo.

**Fase 8 - Configuração das hierarquias e dos níveis nas dimensões convencionais:** é realizado uma tarefa idêntica à da fase anterior, porém dessa vez são configurados as hierarquias e os níveis das dimensões convencionais.

**Fase 9 - Configuração dos cubos:** na Fase 9 os cubos são criados no *Schema* XML a partir da transformação dos fatos em cubos. Estes são constituídos também de dimensões e de medidas que serão adicionadas para análise dos dados do negócio.

**Fase 10 - Exportação do *Schema* XML:** todas as configurações feitas nas fases anteriores são exportadas para o arquivo *Schema* XML. Com esse arquivo e a última fase concluída é possível realizar consultas por intermédio do servidor ROLAP. A implantação do ambiente, segundo o método *Agile* ROLAP, após a conclusão dessa fase pode ser considerada finalizada.

## 2.7.1 Plugins

Alguns *plugins* foram desenvolvidos, por outros autores, com o intuito de facilitar a implantação desse método, entretanto há a possibilidade de que não seja necessário utilizar todos eles. Somente durante o desenvolvimento será definido com certeza quais farão parte da implantação desse ambiente inteligente negócios. Os *plugins* serão integrados à ferramenta Pentaho Data Integration que permite a utilização de *plugins* externos sem muitas restrições. Abaixo uma descrição de cada *plugin*.

**Feriado:** Esse *plugin* não é obrigatório para a utilização do *Agile* ROLAP e foi desenvolvido para definir os feriados que há durante o ano na região em que a organização se encontra. Se for definido não utilizar esse *plugin*, conseqüentemente os feriados não estarão definidos no ambiente de BI. O *plugin* feriado precisa estar conectado a algum *Plugin* Dimensão de Tempo para desempenhar sua função, pois ele fornece informações a este outro *plugin*.

**Tempo:** O *plugin* de Tempo foi desenvolvido com o objetivo de facilitar a criação das dimensões de tempo que serão necessárias para o ambiente de *Business Intelligence*. Esse *plugin* cria tabelas no banco de dados intermediário que possibilita a análise e consulta histórica de dados. Essas tabelas são a representação das dimensões tempo que estarão presentes no ambiente de BI.

Outra função desse *plugin* é enviar informações para outro *plugin*, nesse caso ele é conectado ao *plugin* Fato que utiliza dessas informações para a criação das tabelas Fatos. Além disso vale ressaltar que sua função é criar tabelas que possuem dados temporais. Tabelas que não possuem esses tipos de dados são encarregados ao *plugin* Dimensão.

**Dimensão:** O *plugin* Dimensão é o responsável por definir as tabelas dimensões em geral, exceto as criadas pelo *plugin* Dimensão de Tempo. A principal diferença entre esses dois *plugins* é que o primeiro não necessita de dados carregados no banco de dados para ser criada pois os dados são obtidos através das bases OLTP presentes na organização. O *plugin* Dimensão, após ter sido configurado, envia informações da tabela dimensão para o *plugin* Fato.

**Fato:** Este *plugin* procura facilitar a criação das tabelas fatos que pertencerão ao ambiente inteligente de negócios e está ligada diretamente e indiretamente aos outros *plugins*. Possui duas saídas, uma delas é uma *view* no banco de dados

intermediário e a outra um conjunto de informações que serão utilizados por outro *plugin* para a geração de um arquivo XML.

**XMLOut:** Sua função é gerar um arquivo XML para que o ambiente de BI seja mapeado. O *plugin* XMLOut cria esse arquivo com o intuito de apenas simular a criação de um *Data Warehouse* e estruturar seus dados.

**XMLIn:** O *plugin* XMLIn realizará a leitura do arquivo XML gerado pelo *plugin* XMLOut que é composto por informações recebidas de outros *plugins*, principalmente do Fato.

**Base de dados estrangeira:** Foi desenvolvida para facilitar a criação da base de dados intermediária e como pré-requisito para seu funcionamento é preciso que o PostgreSQL possua a extensão FDW.

**Dimensão lógica de tempo:** Desenvolvida para definir os níveis e as hierarquias das dimensões de tempo.

**Dimensão lógica:** Semelhante ao *plugin* de Dimensão lógica de tempo, essa também foi desenvolvida para definir os níveis e as hierarquias das dimensões, mas dos tipos comuns.

**Cubo:** *Plugin* desenvolvido com intuito de selecionar as dimensões que serão utilizadas no cubo e por definir as medidas que serão utilizadas.

**Schema Out:** Este *plugin* reúne todas as configurações e tarefas realizadas nos outros *plugins*, com exceção do Base de dados Estrangeira, e exporta para o formato Mondrian *Schema*. O *plugin* Base de dados estrangeira auxilia na criação da base de dados intermediária e não interfere na estrutura dos dados e/ou modelo dimensional.

## 3. FERRAMENTAS

Para alcançar o objetivo desse projeto será preciso a utilização de algumas ferramentas que serão apresentadas abaixo, contudo durante o desenvolvimento desse projeto pode surgir a necessidade de alterar ou até mesmo trocar alguma destas.

### 3.1 Pivot4J

Pivot4J é uma API (*Application Programming Interface*) desenvolvida na linguagem Java e *open source* e utiliza como servidor ROLAP o mondrian para gerar a interface da ferramenta OLAP. A origem do nome e da ferramenta em si se deve a união e “evolução” de outras duas ferramentas, mas desatualizadas, o JPivot e Olap4J, e é utilizada como interface para realizar análise de dados no ambiente de BI (Pivot4J, 2017).

### 3.2 Pentaho Data Integration (Kettle)

O Pentaho Data Integration, conhecido também por Kettle, é uma ferramenta que permite utilizar e preparar dados de qualquer fonte de dados e disponibilizar aqueles com mais relevância para a organização e seus usuários finais. Não há necessidade de codificação e nem há complexidade para ser utilizada pois possui ferramentas visuais que auxiliam na sua utilização e boa parte das tarefas podem ser realizadas apenas com o arrastar e soltar de algumas funções. Esse software aceita a utilização de *plugins* externos e é essencial para a implantação do método pois *plugins* foram desenvolvidos especificamente para essa finalidade.

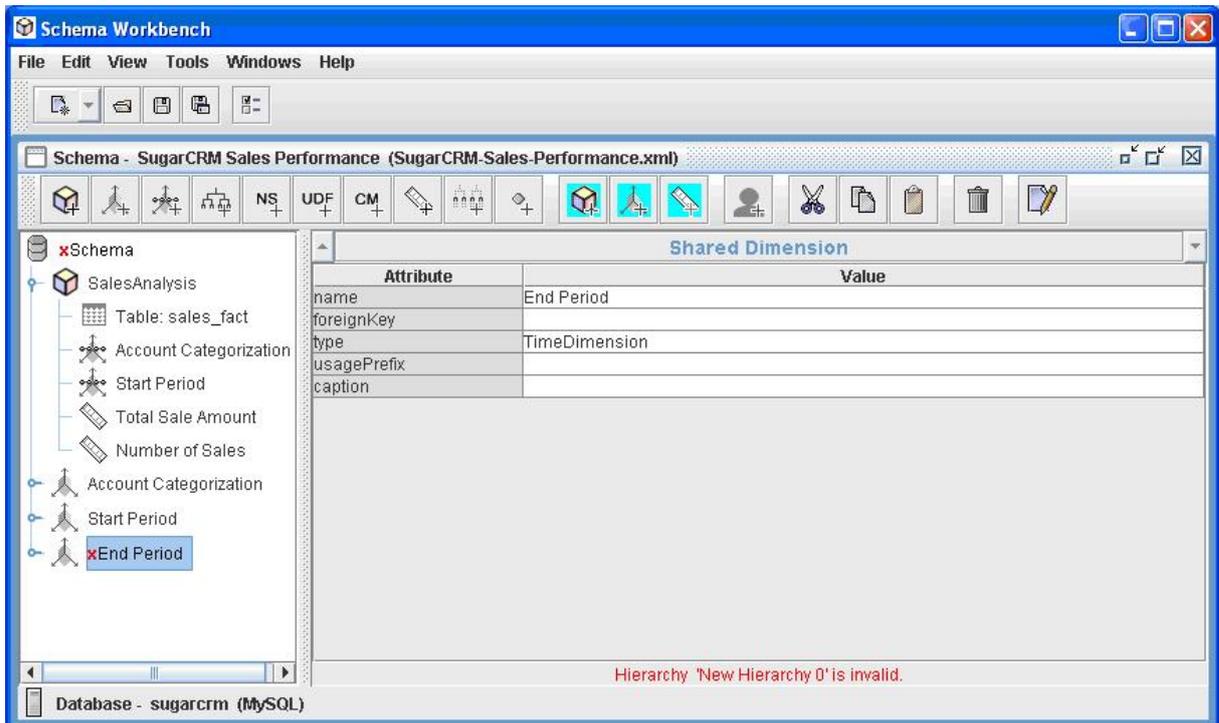
### 3.3 PostgreSQL

O PostgreSQL é um banco de dados do tipo relacional, um pré-requisito do método *Agile* ROLAP. Está presente no mercado já há 15 anos e além de ser *open source* é compatível com os principais sistemas operacionais, incluindo Linux, macOS e Windows (PostgreSQL, 2017).

### 3.4 Mondrian Schema Workbench

A ferramenta Mondrian Schema Workbench possui uma interface que permite criar e testar esquemas de cubo. O cubo é gerado em um arquivo XML em uma estrutura feita para Mondrian interpretar corretamente. Abaixo, Figura 7, a interface da ferramenta (Pentaho, 2007).

**Figura 7:** Interface da ferramenta Schema Workbench.



Fonte: Pentaho (2007).

## 4. DESENVOLVIMENTO

Nesse capítulo será apresentado, gradualmente, as atividades realizadas para a implantação do ambiente de BI seguindo todos os passos do método *Agile* ROLAP, apresentado anteriormente na figura 6. A base de dados a ser utilizada será do vestibular da UENP, especificamente do realizado em 2017 para ingresso em 2018. Foram realizados três fatos, mas após o passo 4 o desenvolvimento do ambiente de BI aborda apenas um deles para não dificultar a leitura e o entendimento. As figuras que representam o desenvolvimento dos fatos restantes estão presentes no Apêndice B.

### 4.1 Levantamento dos dados e requisitos – 1º Passo

Foi escolhida a base de dados do vestibular aplicado em 2017 para o desenvolvimento desse trabalho. Dessa forma o presente trabalho busca auxiliar a universidade com alguns resultados, advindos do ambiente de negócios a ser desenvolvido. Com o ambiente de BI, desenvolvido a partir da utilização do método *Agile* ROLAP, será possível visualizar algumas informações que podem auxiliar no momento de realizar decisões acerca dos futuros vestibulares. Vale ressaltar que todas as informações pessoais, que pudessem identificar de alguma forma os candidatos foram eliminadas da base de dados.

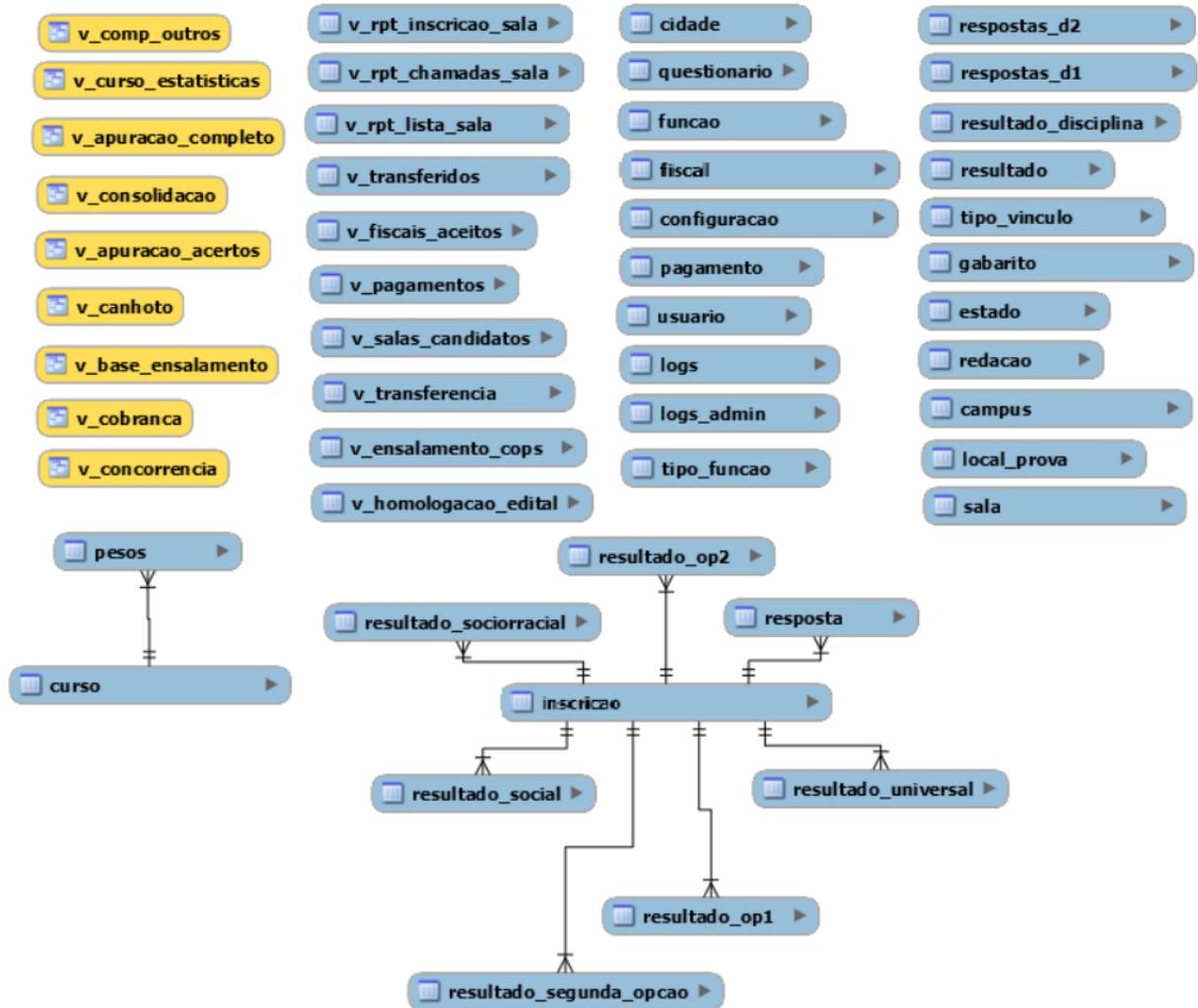
### 4.2 Acesso à fonte de dados – 2º Passo

A base de dados original estava hospedada no MySQL, porém foi transferida para outro Sistema Gerenciador de Banco de Dados (SGBD), o PostgreSQL, por causa da familiaridade com a ferramenta devido ao teste que foi realizado anteriormente a fim de conhecer as tecnologias, os plugins e o método *Agile* ROLAP em si. Além disso o PostgreSQL é indicado ao usar o método *Agile* ROLAP pois possui a tecnologia FDW, que caso necessite, auxiliará na utilização de bases de dados estrangeira.

O diagrama de ERR (Extended-ER ou Enhanced-ER), semelhante ao modelo de diagrama de ER e representado na Figura 8, expõe as entidades, atributos e

relacionamentos entre as tabelas da base analisada. Nessa mesma figura é possível numerar o total de 41 tabelas e 5 views.

Figura 8: Diagrama da base de dados.

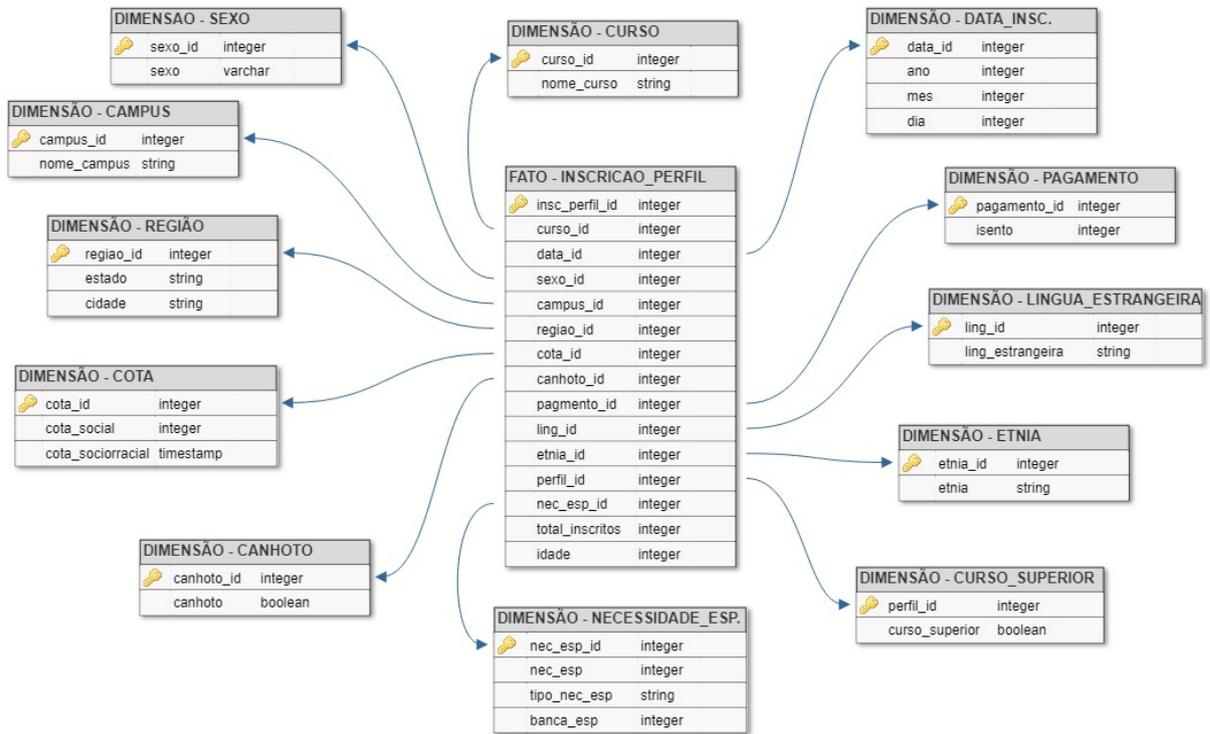


Fonte: Própria do autor.

As tabelas com seus atributos, por serem em grande quantidade, estão presentes no apêndice A.

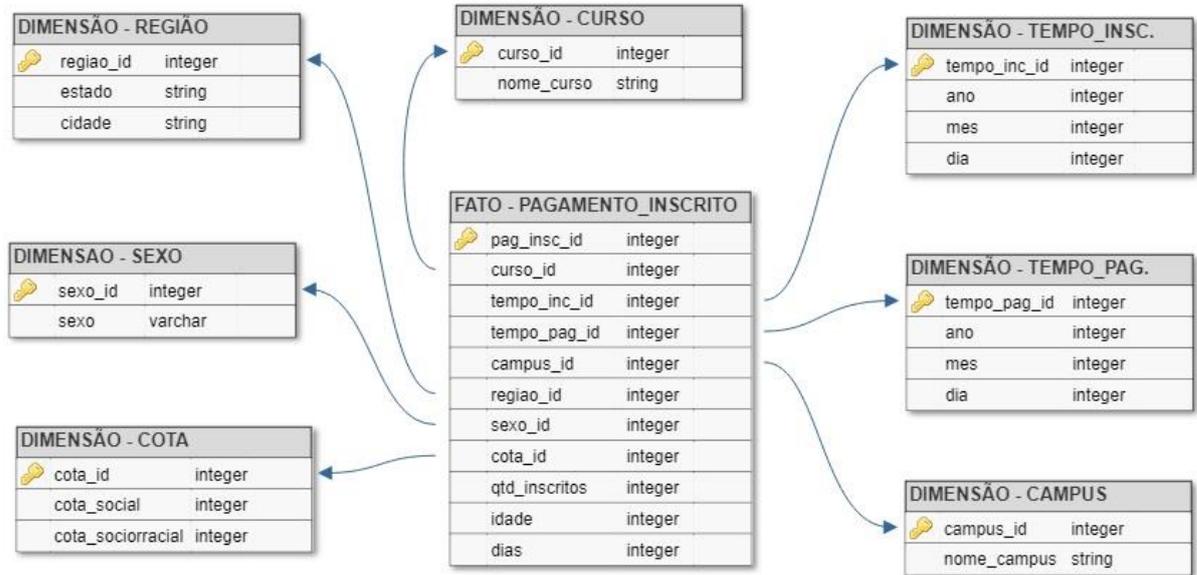
Com a estrutura do banco de dados foram desenvolvidos três fatos. O primeiro fato desenvolvido foi o de *inscricao\_perfil*. Esse fato tem como objetivo avaliar as principais características dos inscritos no vestibular. Esse fato tem como objetivo avaliar as principais características do inscrito no vestibular. Esse fato considera todos aqueles que realizaram a inscrição para a prova, independente de terem pago a inscrição ou não. Também não é considerado se realizaram as provas. Na Figura 9 é apresentado o modelo dimensional desse fato.

**Figura 9:** Fato inscricao\_perfil.



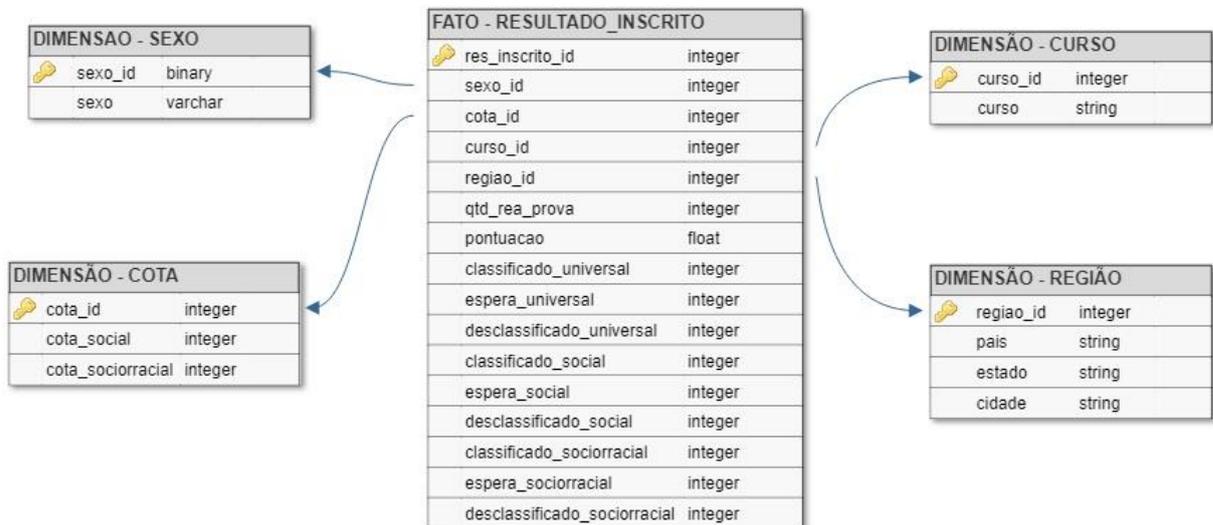
Fonte: Própria do autor.

O segundo modelo dimensional foi criado com o propósito de analisar quando os inscritos realizaram a inscrição no vestibular, quando realizaram o pagamento e quanto tempo em média levaram para realizar o pagamento da inscrição. Sua estrutura é apresentada abaixo na Figura 10 e é considerado nesse fato aqueles inscritos que realizaram o pagamento da inscrição do vestibular, independente de terem realizado as provas ou não.

**Figura 10:** Fato pagamento\_inscrito.

Fonte: Própria do autor.

O terceiro fato desenvolvido, apresentado na Figura 11, exibe informações somente a respeito do resultado daqueles inscritos que realizaram a provas.

**Figura 11:** Fato resultado\_inscrito.

Fonte: Própria do autor.

Conforme apresentado anteriormente, cada fato busca analisar algum evento e/ou assunto diferente que está relacionado ao vestibular da universidade, porém vale ressaltar que outros fatos podem ser desenvolvidos, dependendo sempre das informações que se busca.

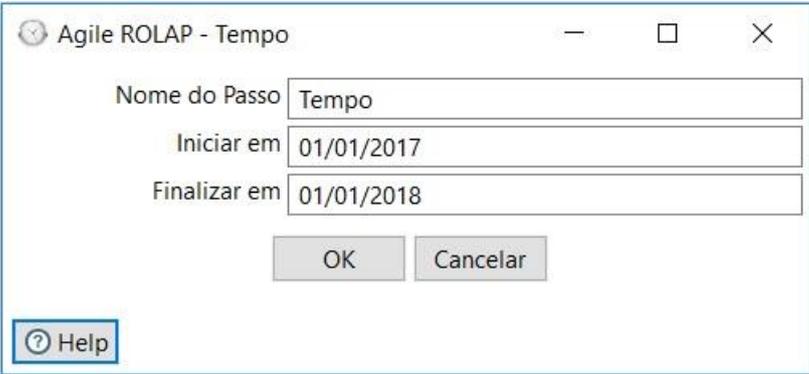
### 4.3 Criação da base de dados intermediária – 3º Passo

A criação da base de dados intermediária, seguindo o modelo *Agile* ROLAP, utiliza a tecnologia FDW do PostgreSQL e é o primeiro passo do desenvolvimento físico do DW. Além de utilizar um plugin desenvolvido para o Kettle, com o intuito de auxiliar nesse processo da criação das bases. Porém o *plugin* desenvolvido apresentou um erro, que impossibilitou seu uso. Contudo, por utilizar apenas uma base de dados que não está em operação, a ausência desse *plugin* e dessa tecnologia não prejudicou o desenvolvimento desse trabalho. O erro mencionado é demonstrado posteriormente, junto a outros erros que foram encontrados, em um tópico específico.

### 4.4 Criação das dimensões de tempo – 4º Passo

Como mencionado anteriormente a partir do 4º passo é abordado o desenvolvimento de apenas um fato. No caso é desenvolvido o fato `inscricao_perfil` apresentado na Figura 10, que possui a dimensão tempo para análise de quando os inscritos realizaram a inscrição para o vestibular. Para auxiliar nesse passo foi utilizado do `plugin_tempo` que gera uma tabela com dados de tempo em formatos diversos. Na Figura 12 é exibido a configuração utilizada para o fato `inscricao_perfil`.

**Figura 12:** Configuração do plugin tempo.



Agile ROLAP - Tempo

Nome do Passo: Tempo

Iniciar em: 01/01/2017

Finalizar em: 01/01/2018

OK Cancelar

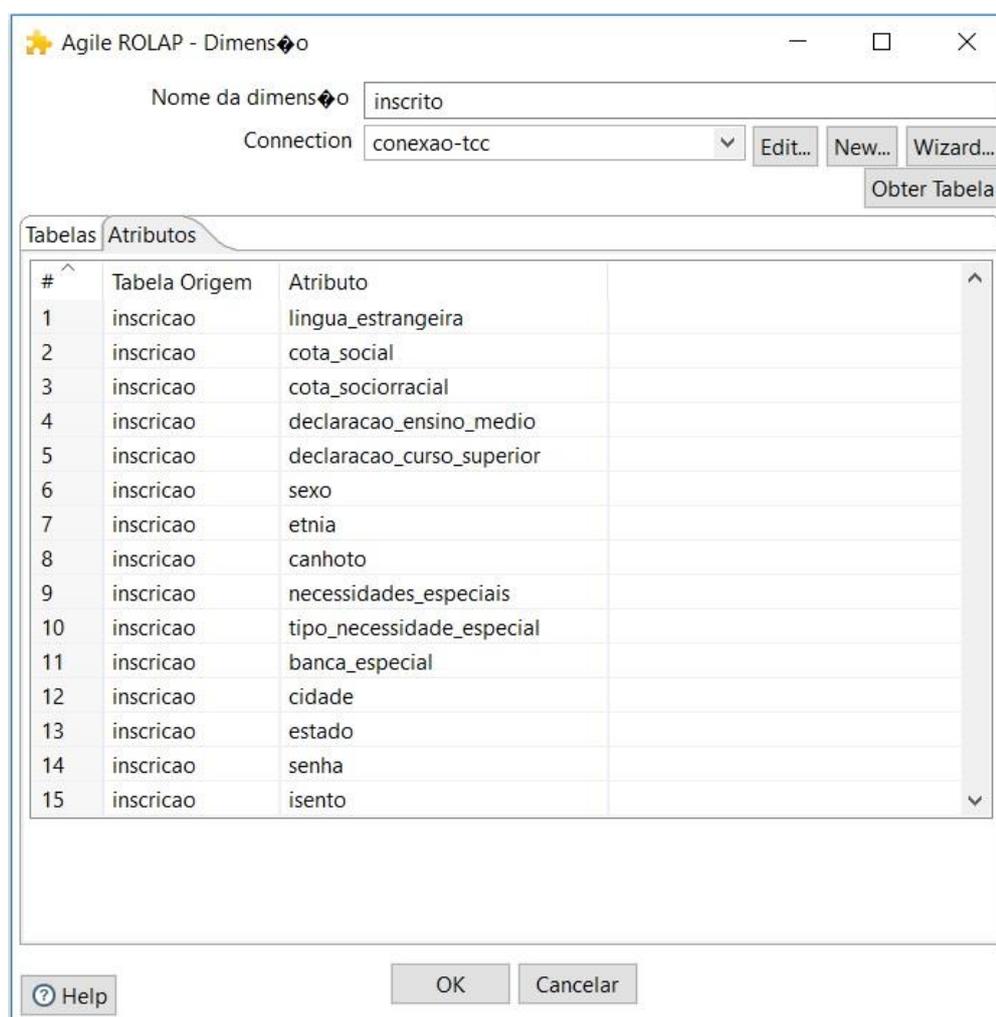
Help

Fonte: Própria do autor.

## 4.5 Criação das dimensões convencionais – 5º Passo

Por ter como objetivo de análise o perfil do inscrito esse fato utiliza de dimensões que selecionam dados do inscrito, do curso e do campus, presente na base de dados em diferentes tabelas. Além de dados presentes na tabela tempo criada anteriormente com auxílio do plugin tempo. Na Figura 13 são mostrados os dados selecionados, com o *plugin* dimensão, da tabela inscricao para a criação da dimensão inscrito. No Apêndice B está a configuração das outras dimensões desse fato.

**Figura 13:** Configuração da dimensão inscrito.

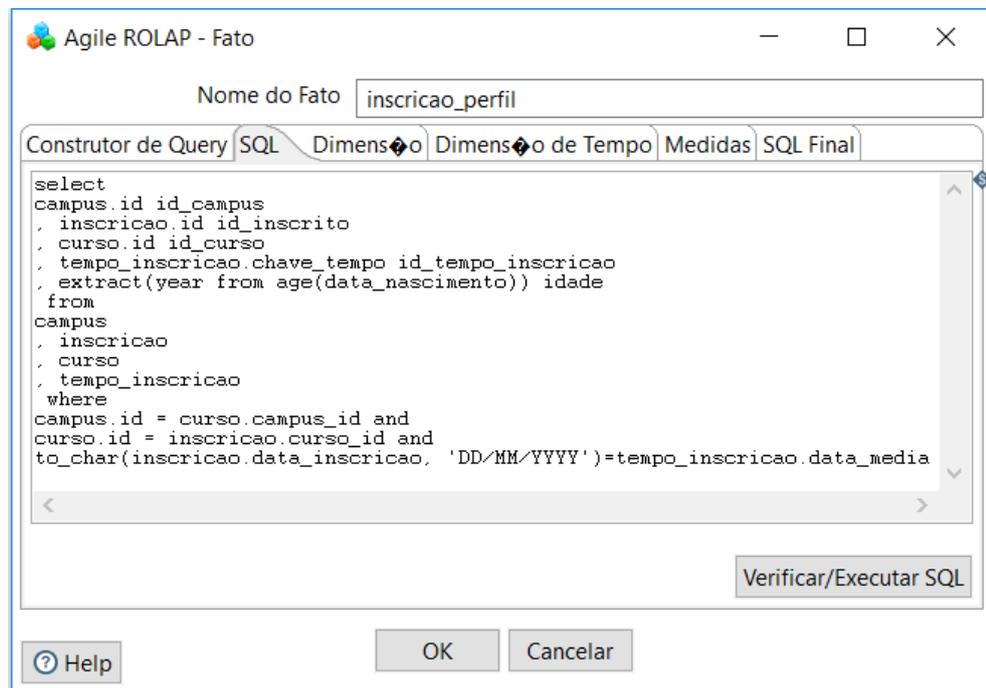


Fonte: Própria do autor.

## 4.6 Criação dos fatos – 6º Passo

Para criar o fato é necessário gerar um SQL, no *plugin* fato, que seleciona as chaves das tabelas que representam as dimensões e os dados que representam as medidas, o SQL é mostrado na Figura 14. O SQL do fato *inscricao\_perfil* seleciona os atributos: *id\_campus*, *id\_inscrito*, *id\_curso*, *id\_tempo\_inscricao* e *idade*. Por fim é criado uma *View* com esses atributos.

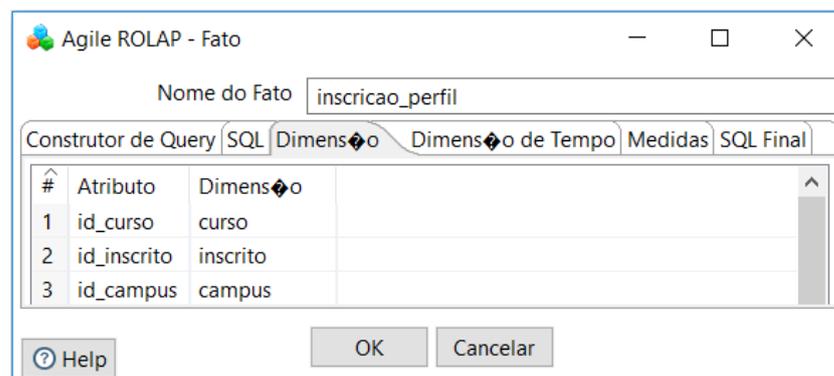
**Figura 14:** SQL do fato *inscricao\_perfil*.



Fonte: Própria do autor.

Após selecionar as chaves da base de dados é escolhido qual chave representa qual dimensão, como exibido na Figura 15.

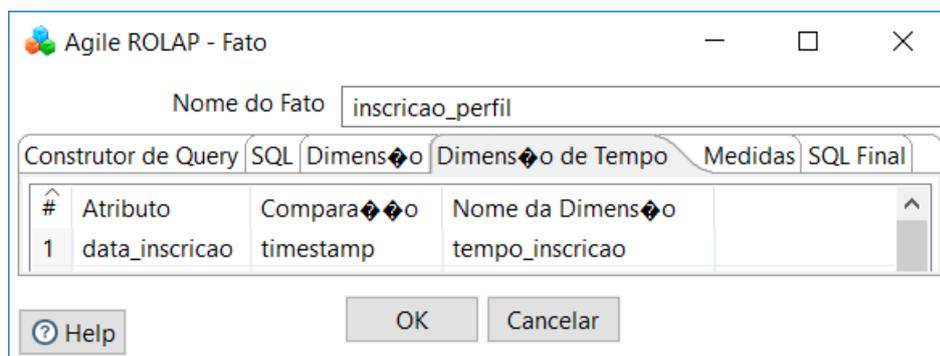
**Figura 15:** Dimensões e chaves conectadas.



Fonte: Própria do autor.

Após configurar as dimensões convencionais no plugin fato é também configurado a dimensão tempo, apresentado na Figura 16. Essa dimensão utiliza do atributo `data_inscricao` da base de dados e será utilizada para gerar informações sobre quando os inscritos realizaram as inscrições.

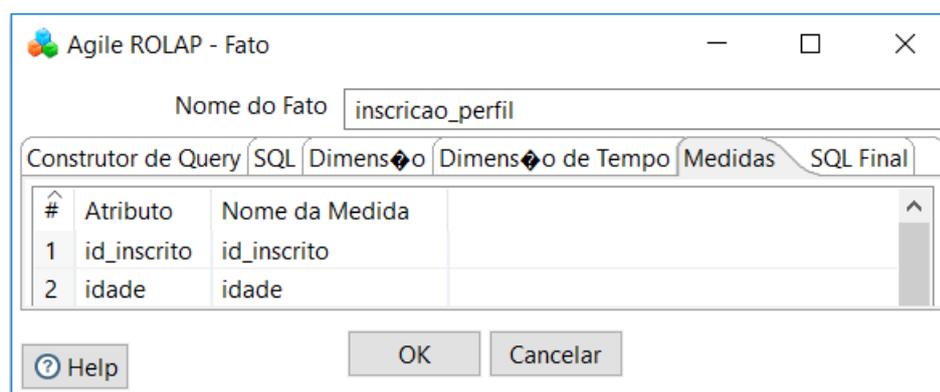
**Figura 16:** Configuração da dimensão tempo.



Fonte: Própria do autor.

No plugin fato, após configurar a aba de dimensão tempo, também é selecionado as medidas que foram definidas na modelagem dimensional e são exibidas nas Figura 17. Para o fato `inscricao_perfil` foram selecionadas duas medidas. A medida que usa o atributo `id_inscrito` será utilizado para calcular a quantidade de inscritos e a medida `idade` será utilizada para gerar a idade média destes.

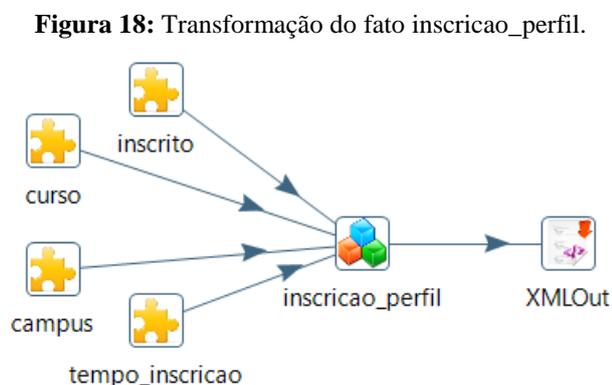
**Figura 17:** Medidas selecionadas no plugin fato.



Fonte: Própria do autor.

A última aba, `SQLFinal`, exibe o SQL que será utilizado para gerar a View Materializada. Após a configuração do plugin fato ser concluída é então executado a transformação na ferramenta Kettle que por fim gera um arquivo XML através do

plugin XMLOut. O XML criado possui todas as configurações e definições realizadas até o momento. A Figura 18 demonstra a estrutura e as conexões da transformação executada. Após gerar o XML a etapa física da construção do DW é finalizada.



Fonte: Própria do autor.

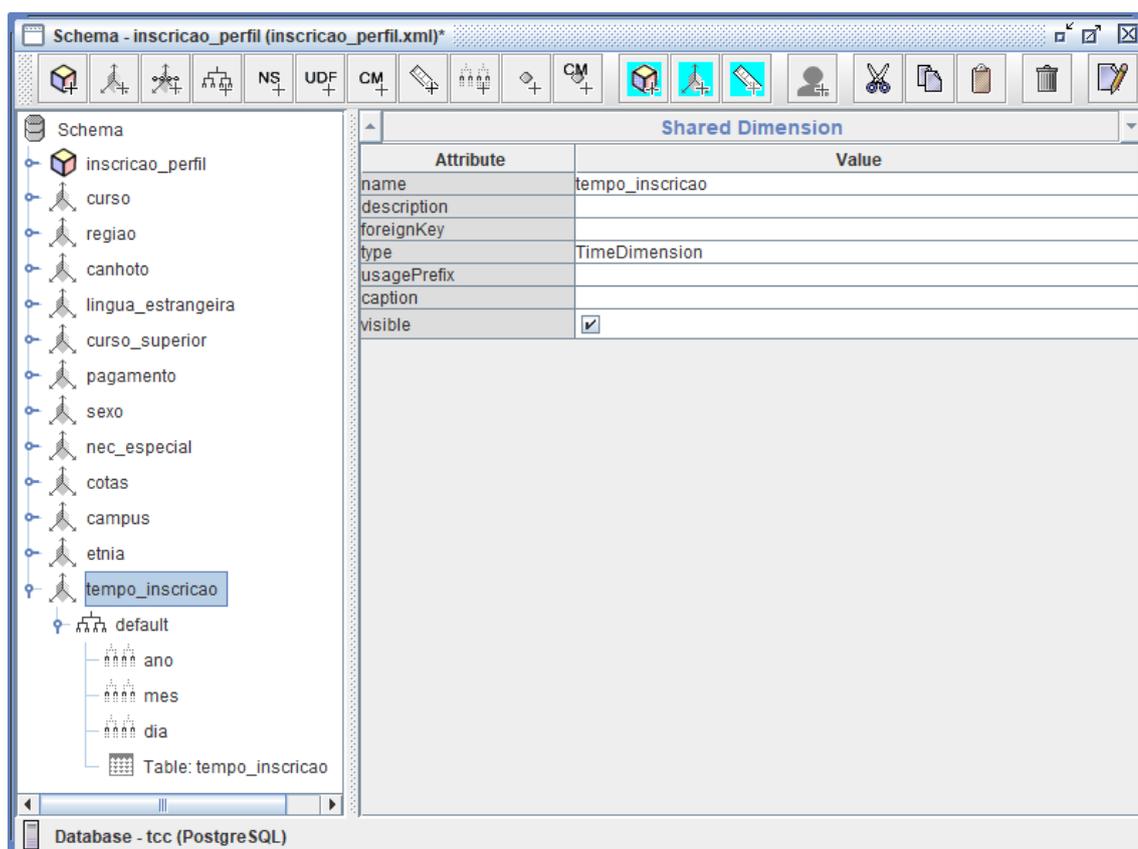
## 4.7 Configuração das hierarquias e níveis nas dimensões de tempo – 7º Passo

No 7º Passo é iniciado a configuração da etapa lógica do *Data Warehouse* que define exatamente quais as dimensões e medidas poderão ser escolhidas na ferramenta OLAP.

Como exposto na configuração da dimensão tempo na etapa física do DW os dados relacionados ao tempo a serem utilizados abordam a data de inscrição dos candidatos.

A configuração desse tipo de dimensão normalmente é realizada no *plugin* dimensão lógica de tempo, no entanto o plugin apresentou um erro ao tentar ser utilizado e foi necessário utilizar uma outra ferramenta para configurar essa dimensão. O erro exibido é apresentado em um tópico específico para erros, o tópico 5.1, e a ferramenta utilizada para configurar a dimensão foi o *Schema Workbench*. Na configuração lógica da dimensão tempo é definida a hierarquia e os níveis da dimensão. Para a dimensão lógica de tempo do fato *inscricao\_perfil* foram definidos os níveis ano, mês e dia. A configuração dessa dimensão lógica é retratada abaixo na Figura 19.

**Figura 19:** Configuração do tempo do fato inscricao\_perfil.



Fonte: Própria do autor.

## 4.8 Configuração das hierarquias e níveis nas dimensões convencionais – 8º Passo

Nesse oitavo passo foram configuradas as hierarquias e níveis das dimensões convencionais. Para o fato inscricao\_perfil ao todo foram configuradas 11 dimensões convencionais além da dimensão tempo criada anteriormente. As dimensões a serem utilizadas são:

**Curso:** essa dimensão retrata o curso para o qual o candidato fez a inscrição.

**Sexo:** a dimensão sexo apresenta a perspectiva do sexo masculino e feminino que buscam ingressar na universidade.

**Campus:** apresenta a perspectiva de para qual campus a pessoa fez a inscrição.

**Regiao:** exibe de quais cidades e estados os inscritos são.

**Cotas:** expressa em cota social e sociorracial os inscritos.

**Canhoto:** seleciona uma visão dos inscritos que são canhotos

**Nec\_especiais:** interpreta os inscritos que possuem necessidades especiais, os seus tipos de necessidades e se precisaram de uma banca especial ou não.

**Curso\_superior:** representa uma visão dos inscritos que já possuem ou não curso superior.

**Etnia:** essa dimensão seleciona os tipos de etnias inscritos no vestibular.

**Lingua\_estrangeira:** a dimensão lingua\_estrangeira demonstra a escolha dos inscritos por realizarem a prova com a língua inglesa ou espanhola.

**Pagamento:** essa dimensão determina aqueles que são isentos ou não do pagamento do vestibular. Vale ressaltar que essa dimensão não afirma que todos os inscritos não isentos necessariamente realizaram o pagamento da inscrição.

Abaixo, na Figura 20, é exposto a configuração da dimensão lógica regioao que possui dois níveis, sendo eles o nível cidade e o nível estado. A configuração das outras dimensões está no Apêndice B.

**Figura 20:** Configuração da dimensão lógica regioao do fato inscricao\_perfil.

Fonte: Própria do autor.

## 4.9 Configuração dos cubos – 9º Passo

No 9º passo é configurado o cubo que representa o fato inscricao\_perfil. Na configuração do cubo é utilizado o plugin cubo o qual utiliza dos dados selecionados e definidos ao configurar as dimensões lógicas nos passos anteriores. Ao todo foram configuradas 11 dimensões pois o objetivo desse fato é analisar o perfil do candidato. A Figura 21 mostra como foi configurado as dimensões do cubo para esse fato.

**Figura 21:** Configuração das dimensões do cubo do fato inscricao\_perfil.

MeuPlugin3

Nome do Passo:

Carregar de:

Cubo Medidas

Cubo

Nome do Cubo:

Tabela:

Visível:

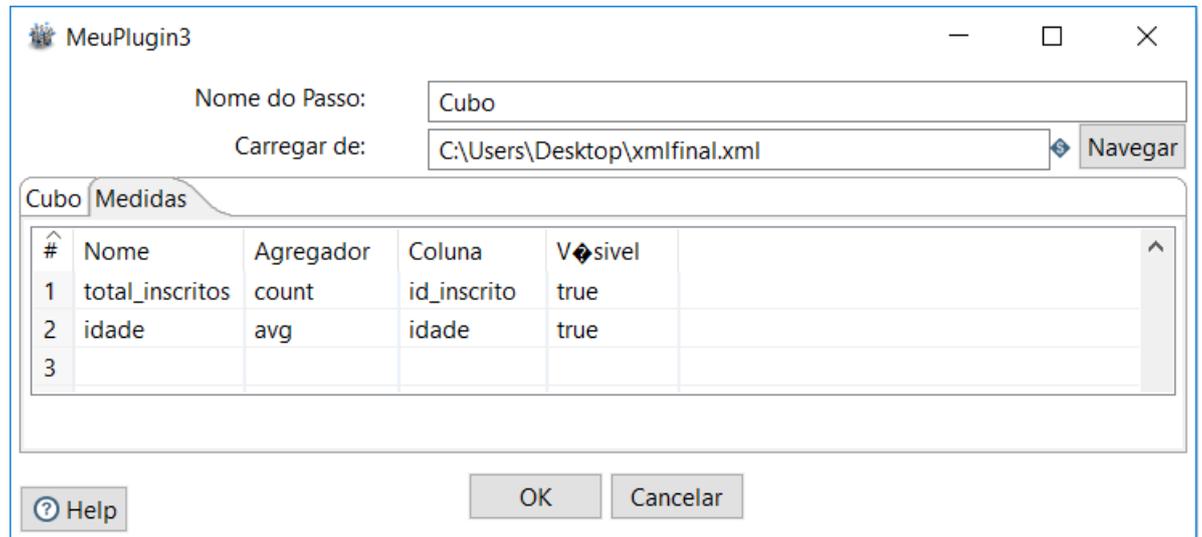
Uso de Dimensão

#	Nome	Fonte
1	curso	curso
2	sexo	sexo
3	regiao	regiao
4	cotas	cotas
5	canhoto	canhoto
6	nec_especial	nec_especial
7	curso_superior	curso_superior
8	etnia	etnia
9	lingua_estrangeira	lingua_estrangeira
10	pagamento	pagamento
11	campus	campus

Fonte: Própria do autor.

Nesse mesmo passo o plugin é configurado as medidas que serão utilizadas na ferramenta OLAP no momento de visualizar as informações. A Figura 22 mostra a configuração feita na aba de medidas no plugin cubo.

**Figura 22:** Configuração das medidas do cubo do fato inscricao\_perfil.

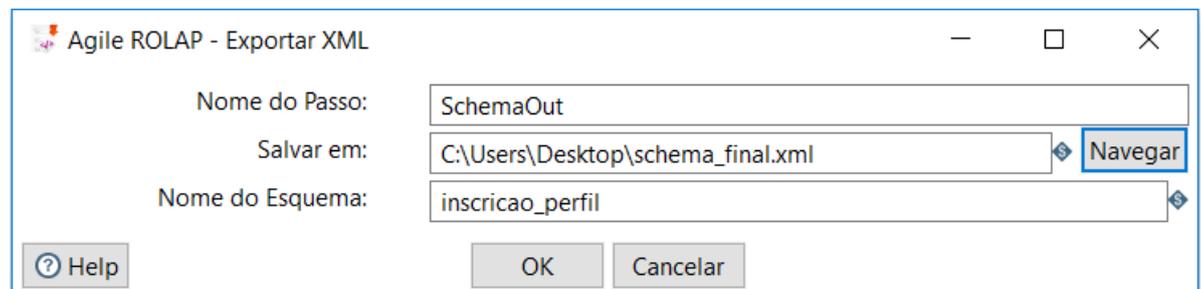


Fonte: Própria do autor.

#### 4.10 Exportação do Schema XML – 10º Passo

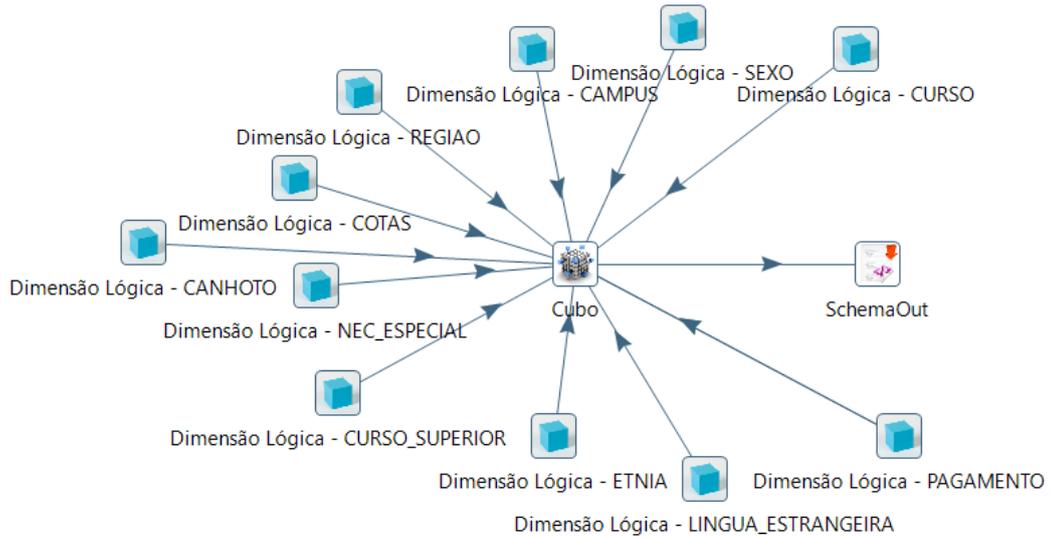
Após todos os passos anteriores realizados é possível executar a transformação na ferramenta Kettle e gerar o XML, através do *plugin SchemaOut*, que será interpretado por uma ferramenta OLAP. A configuração do plugin *SchemaOut* está na Figura 23 e a estrutura da transformação executada está exposta na Figura 24.

**Figura 23:** Configuração do plugin SchemaOut para o fato inscricao\_perfil.



Fonte: Própria do autor.

**Figura 24:** Estrutura da transformação criada no Kettle para o fato inscricao\_perfil.



Fonte: Própria do autor.

Com o XML pronto para ser usado em uma ferramenta de análise e com o último passo completo é possível afirmar que a implantação do ambiente de BI utilizando o método *Agile* ROLAP está concluída.

## 5. RESULTADOS

O objetivo principal desse trabalho é analisar a metodologia *Agile* ROLAP, portanto foi realizado uma implantação de um ambiente de BI utilizando a metodologia para criar um espaço para análise de dados e informações do vestibular realizado pela UENP. Além de utilizar e analisar os componentes (ferramentas e *plugins*) que auxiliam na implantação do projeto.

O presente trabalho como resultado busca também auxiliar, expondo configurações e detalhes técnicos do desenvolvimento, aqueles de buscam melhorar o método ou até mesmo implantar um ambiente de BI em uma organização.

A principal vantagem do método *Agile* ROLAP é a redução do tempo e custo para implantar um ambiente de BI comparado aos métodos tradicionais existentes. Isso ocorre, pois, o método propõe utilizar os dados diretamente de sua base de dados de origem, sem necessitar de alteração. Por esse motivo, uma desvantagem, em algumas ocasiões foi necessário buscar um conhecimento maior em consultas SQL para selecionar os dados corretamente. Além disso o método funcionou perfeitamente pois foi utilizado em uma base de dados de tamanho considerado pequena.

Um ponto positivo ao desenvolver um ambiente de BI com esse método é sua compatibilidade com ferramentas de visualização de dados que foram desenvolvidas para ambientes criados com métodos tradicionais. Nesse trabalho foi utilizado o Pivot4J, porém foi testado também na ferramenta Saiku que assim como o primeiro também utiliza do Mondrian.

As ferramentas e *plugins* em sua maior parte supriram as necessidades da metodologia, porém por nem todas estarem funcionando corretamente foi necessário um tempo maior que o planejado para o desenvolvimento desse trabalho. Contudo, com algumas alternativas realizadas foi possível concluir esse ambiente de BI e realizar algumas análises. Vale destacar que a maior parte dos *plugins* estão funcionando corretamente e que os erros encontrados são apresentados posteriormente.

Embora não seja o objetivo principal desse trabalho foi possível realizar algumas análises dos dados da base com a utilização da ferramenta Pivot4J que são apresentados no tópico a seguir. Foram também encontrados alguns erros nos *plugins* durante um teste realizado antes e também durante o desenvolvimento

desse trabalho. Por esse motivo é apresentado também alguns erros encontrados em um tópico específico. Vale ressaltar que a análise dos *plugins* foi em busca de erros e eficiência. Por essa razão foram comentados apenas os erros e problemas encontrados.

## 5.1 Análise dos dados

Ao todo foram desenvolvidos três cubos e com cada um foi possível encontrar informações que talvez não seriam encontradas analisando sem um ambiente de BI. Ao analisar os dados com o cubo de pagamento\_inscrito foi possível perceber que os candidatos para o curso de filosofia são aqueles que realizaram o pagamento da inscrição no período mais curto enquanto os candidatos ao curso de ciências econômicas são os que demoraram mais para realizar o pagamento. Essa informação está apresentada na Figura 25.

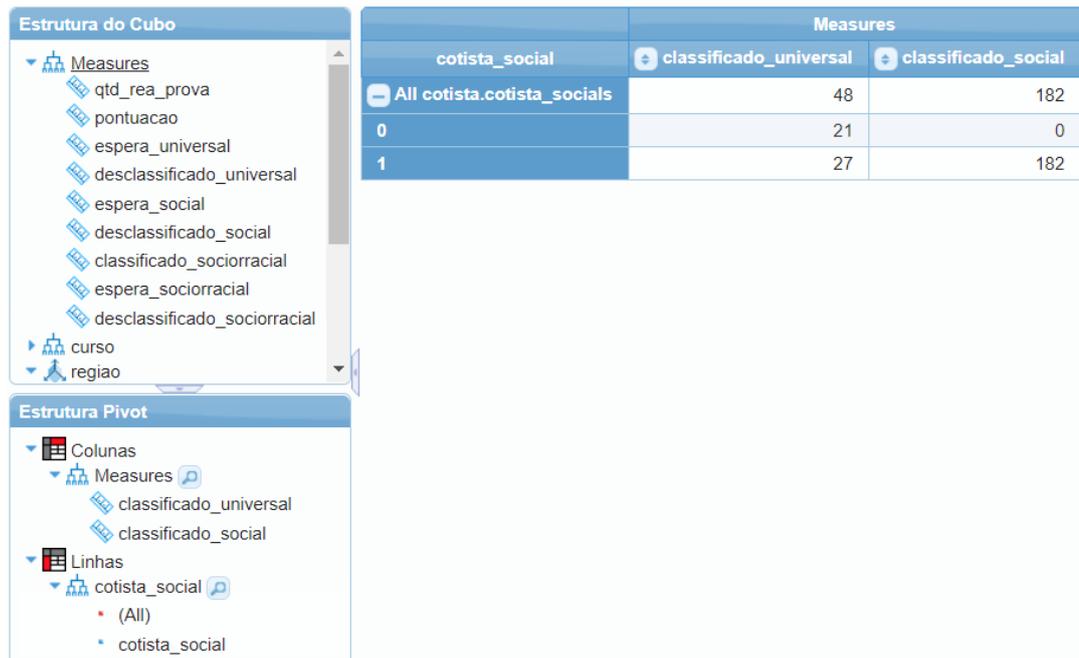
**Figura 25:** Análise dos dados do cubo pagamento\_inscrito.

curso	Measures	
	qtd_inscritos	dias
Ciências Econômicas	39	30,949
Fisioterapia	338	25,84
Letras Português/Inglês	138	25,783
Odontologia	623	24,884
Medicina Veterinária	579	24,244
Enfermagem	145	23,772
Direito	1.241	23,688
Pedagogia	380	23,347
Ciências Biológicas	241	23,158
Matemática	140	22,886
Agronomia	424	22,005
Letras Português/Espanhol	40	21,3
Geografia	39	20,718
Administração	227	20,41
Educação Física (Licenciatura)	181	20,365
Sistemas de Informação	61	19,377
História	143	19,021
Ciências Contábeis	226	18,934
Educação Física (Bacharelado)	213	18,845
Ciência da Computação	64	18,641
Filosofia	57	17,719

Fonte: Própria do autor.

Ao analisar o cubo resultado\_inscrito foi possível visualizar facilmente a quantidade de pessoas que realizaram a prova na categoria cotista e foram aprovados tanto nessa mesma categoria quanto conseguiram um resultado bom o bastante para serem aprovados na categoria universal, sem necessitar de qualquer tipo de cota. A Figura 26 mostra essa análise realizada no Pivot4J.

**Figura 26:** Análise dos dados do cubo resultado\_inscrito.



Fonte: Própria do autor.

Após realizar a análise do último cubo desenvolvido, o inscricao\_perfil, é possível observar algumas informações a respeito das características e tipos de candidatos aos cursos da universidade. Na Figura 27 é possível analisar alguns dados como a quantidade de inscritos em um determinado curso relacionado a sua idade e ao seu sexo.

**Figura 27:** Análise dos dados do cubo inscricao\_perfil.

Estrutura do Cubo		Measures		
Measures				
curso	(All)			
	nome_curso			
sexo	(All)			
	sexo_inscrito			
regiao_inscrito				
cotas				
	cota_social			
	cota_sociorracial			
canhoto				
nec_especial				

Estrutura Pivot		Measures		
Colunas				
	Measures			
	idade			
	total_inscritos			
Linhas				
	curso			
	nome_curso			
	sexo			
	sexo_inscrito			

curso	sexo	idade	total_inscritos
Administração	F	20,862	189
	M	20,891	201
Agronomia	F	19,321	165
	M	19,007	427
Ciência da Computação	F	18,444	9
	M	18,917	84
Ciências Biológicas	F	20,688	276
	M	20,879	140
Ciências Contábeis	F	20,977	171
	M	21,186	177
Ciências Econômicas	F	20,077	26
	M	21,139	36
Direito	F	19,952	1.114
	M	21,257	766
Educação Física (Bacharelado)	F	21,752	129
	M	21,387	225
Educação Física (Licenciatura)	F	20,59	117
	M	20,893	169
Enfermagem	F	19	221
	M	20,231	39

Fonte: Própria do autor.

Para conseguir utilizar a medida idade nos cubos foi necessário realizar umas mudanças diretamente no banco. É preciso destacar que essa tarefa não é indicada e nem realizada no método, porém vale ressaltar também que foi necessária pois a análise seria prejudicada. Na Figura 28 é demonstrado quais dados especificamente foram alterados e para quais valores foram adaptados. No caso são valores referentes a data de nascimento que alguns inscritos provavelmente preencheram de forma incorreta. Esses dados estavam alterando os resultados ao utilizar a medida idade.

**Figura 28:** Alteração das datas de nascimento de alguns inscritos.

data_nascimento date	data_nascimento date
2016-02-16	1996-02-16
2017-01-26	1997-01-26
2017-02-18	1997-02-18
2017-03-31	1997-03-31
2017-04-09	1997-04-09
2017-05-13	1997-05-13
2017-06-27	1997-06-27
2017-07-01	1997-07-01
2017-07-09	1997-07-09
2017-07-25	1997-07-25
2017-08-09	1997-08-09
2017-08-14	1997-08-14
2017-10-11	1997-10-11
2017-11-25	1997-11-25
2998-07-02	1998-07-02
6197-01-04	1997-01-04
7997-12-21	1997-12-21
1199-06-30	1999-06-30
1669-01-24	1999-01-24
1878-08-08	1998-08-08

Fonte: Própria do autor.

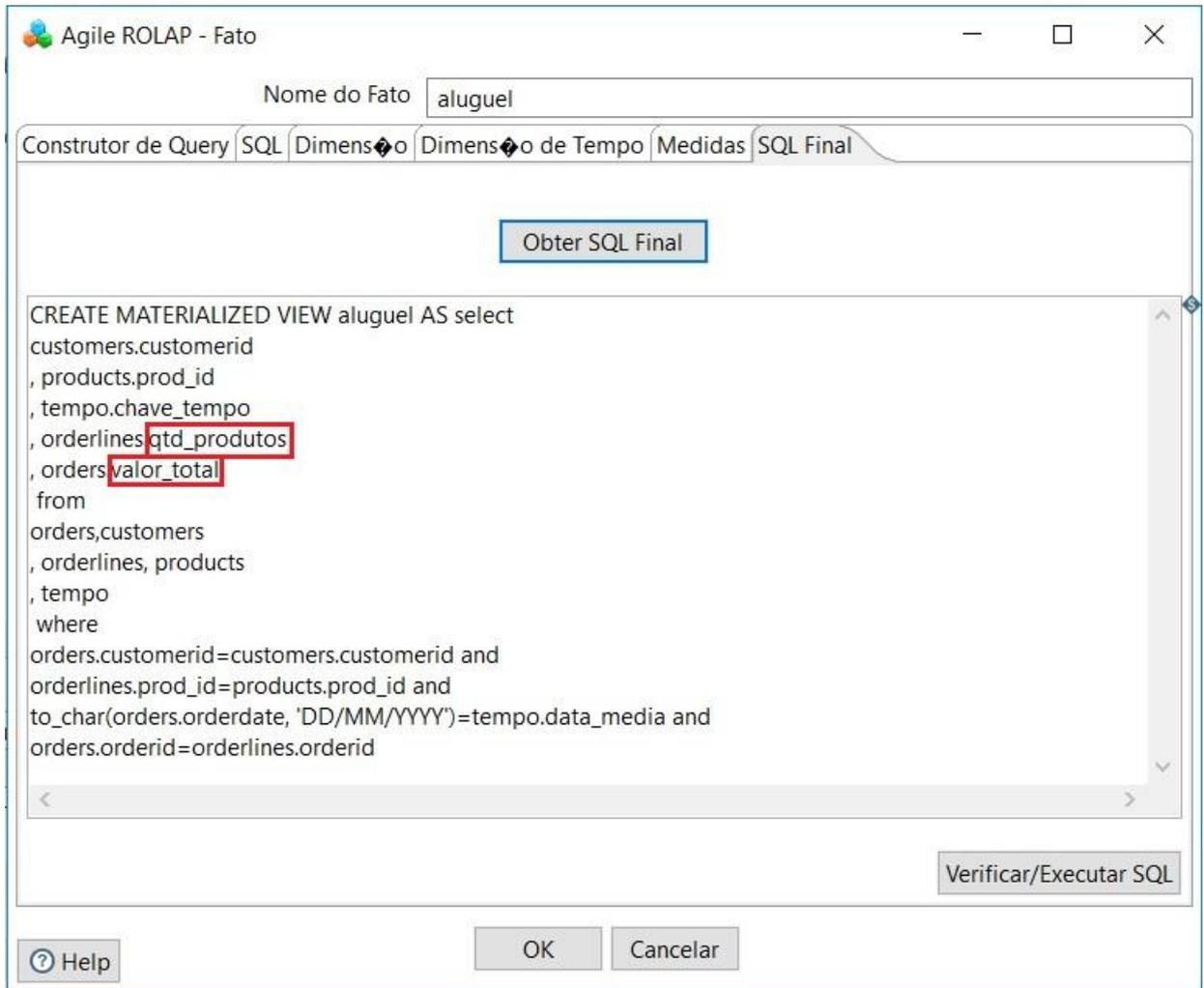
## 5.2 Erros Encontrados

Conforme apresentado, anteriormente, antes de utilizar o *Agile* ROLAP na base de dados do vestibular a metodologia foi utilizada em uma base de dados diferente. Foi utilizada como teste a base de dados DellStore, um sistema OLTP para pedidos/vendas de DVDs que possui 8 tabelas e é disponibilizada pela Dell<sup>2</sup>. Ao realizar esse teste prévio e também durante o desenvolvimento desse projeto foram encontrados alguns erros que serão discutidos abaixo.

### 5.2.1 View gerada está incorreta

Durante o teste, ao tentar criar o fato aluguel, o SQL (*Structured Query Language*) da *View* do fato gerado pelo *plugin* fato não funciona e dessa forma não é possível criar a *View* do fato. A Figura 29 demonstra o SQL da *View* gerado pelo *plugin*.

<sup>2</sup> Site da base de dados: <https://linux.dell.com/dvdstore/>

**Figura 29:** SQL gerado pelo plugin FATO.

Fonte: Própria do autor.

Na Figura 29, com uma borda vermelha ao redor, é possível ver em destaque os locais em que estão presentes os erros. Para gerar uma *View* foi necessário corrigi-la e na Figura 30 é demonstrado a correção realizada. Os locais destacados com uma borda verde são os campos que foram corrigidos.

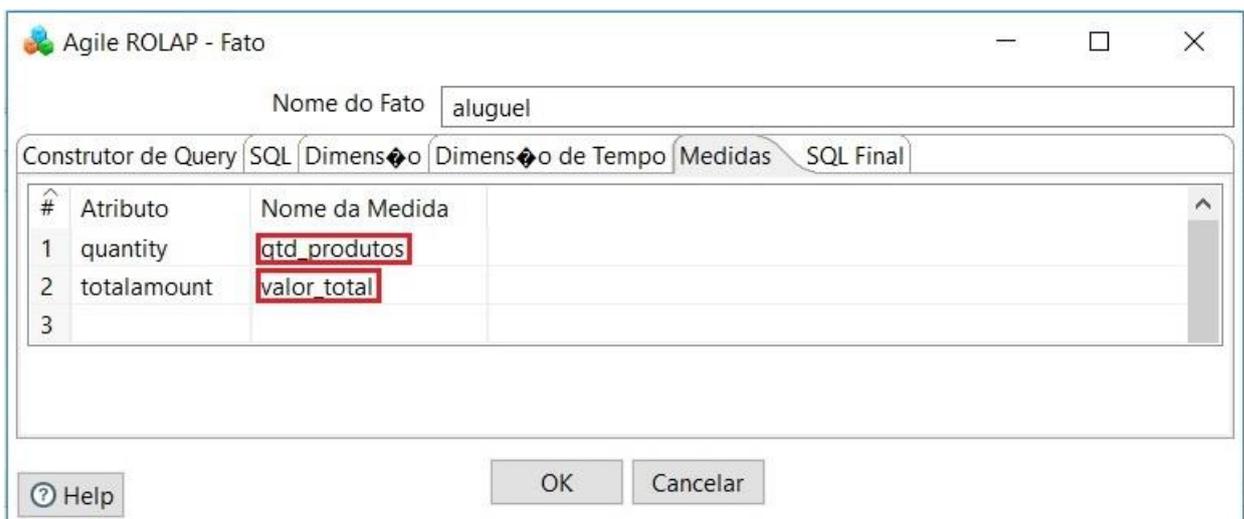
**Figura 30:** SQL utilizável para gerar uma *view*.

```
CREATE MATERIALIZED VIEW Aluguel AS select
customers.customerid
, products.prod_id
, tempo.chave tempo
, orderlines.quantity
, orders.totalamount
from
orders,customers
, orderlines, products
, tempo
where
orders.customerid=customers.customerid and
orderlines.prod_id=products.prod_id and
to_char(orders.orderdate, 'DD/MM/YYYY')=tempo.data_media and
orders.orderid=orderlines.orderid
```

Fonte: Própria do autor.

Conforme retratado na Figura 29 os erros estão na seleção de alguns campos, no qual o *plugin* aparentemente está gerando uma *View* trocando o nome das medidas presente no banco de dados pelo nome que o usuário escolheu ao utilizar a *plugin* fato. Na Figura 31 é possível visualizar o local que foi escolhido o nome das medidas que está causando conflito ao gerar o *Schema* final.

**Figura 31:** Declaração dos nomes das medidas no plugin fato.

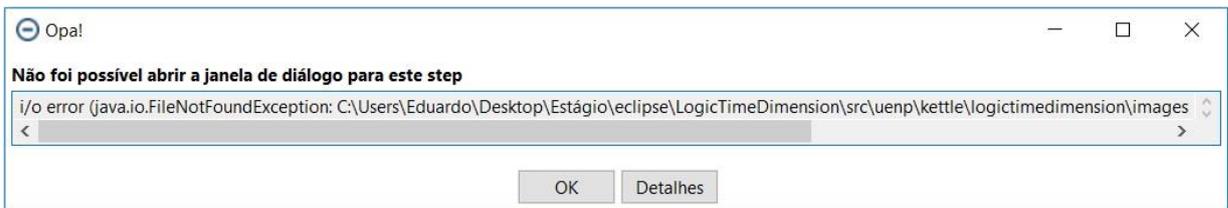


Fonte: Própria do autor.

### 5.2.2 Plugin dimensão lógica de tempo não funcionando

Ao tentar utilizar o *plugin* específico para criar a dimensão lógica de tempo na ferramenta Kettle é exibido uma janela informando um erro. O erro presente na Figura 32, não permite que o plugin funcione e por esse motivo é necessário realizar essa etapa utilizando uma outra ferramenta. Nesse caso foi utilizado o *Schema Workbench*.

**Figura 32:** Erro ao tentar utilizar plugin de dimensão lógica de tempo.



Fonte: Própria do autor.

### 5.2.3 Schema final incorreto

Após executar uma transformação no Kettle é gerado um XML que ao tentar utilizar na ferramenta Pivot4j apresenta alguns elementos incorretos, consequência do erro presente no plugin fato. Abaixo na Figura 33 está o *Schema* gerado pelo Kettle durante o teste prévio das ferramentas com a base de dados DellStore.

**Figura 33:** Schema final incorreto.

```

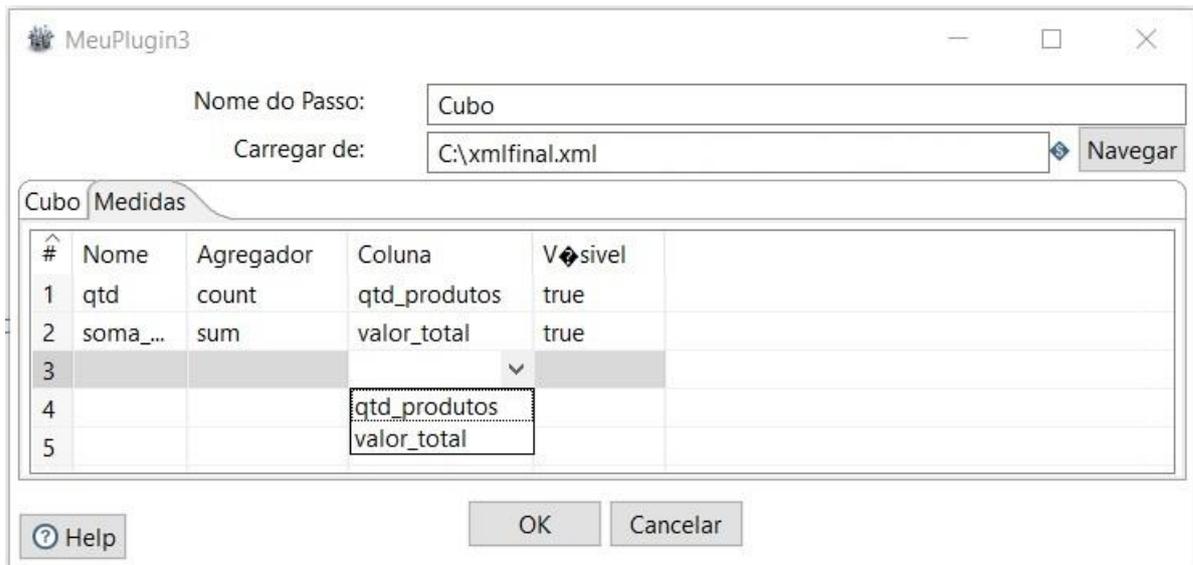
▼<Schema name="Vendas">
  ▼<Dimension type="StandardDimension" visible="true"
    highCardinality="false" name="cliente">
    ▼<Hierarchy name="cliente" visible="true" hasAll="true"
      primaryKey="customerid">
      <Table name="customers"/>
      <level name="nome" visible="true" table="customers"
        column="firstname" type="String" uniqueMembers="false"
        levelType="Regular" hideMemberIf="Never"/>
      </Hierarchy>
    </Dimension>
  ▼<Dimension type="StandardDimension" visible="true"
    highCardinality="false" name="produto">
    ▼<Hierarchy name="produto" visible="true" hasAll="true"
      primaryKey="prod_id">
      <Table name="products"/>
      <level name="titulo" visible="true" table="products"
        column="title" type="String" uniqueMembers="false"
        levelType="Regular" hideMemberIf="Never"/>
      </Hierarchy>
    </Dimension>
  ▼<Cube name="vendas" visible="true" cache="true" enabled="true">
    <Table name="aluguel"/>
    <DimensionUsage source="cliente" name="cliente" visible="true"
      foreignKey="customerid" highCardinality="false"/>
    <DimensionUsage source="produto" name="produto" visible="true"
      foreignKey="prod_id" highCardinality="false"/>
    <Measure name="qtd" column="qtd produtos" aggregator="count"
      visible="true"/>
    <Measure name="soma_valor_total" column="valor total"
      aggregator="sum" visible="true"/>
  </Cube>
</Schema>

```

Fonte: Própria do autor.

No plugin cubo ao tentar selecionar a coluna correta, o *plugin* exibe como opções os nomes das medidas que foram atribuídas durante a configuração do *plugin* fato e como consequência o *Schema* final gerado não é funcional. A Figura 34 mostra o momento de selecionar a coluna da medida, porém nenhuma delas correta.

**Figura 34:** Plugin cubo exibindo medidas incorretas.



Fonte: Própria do autor.

O *Schema*, conforme visto anteriormente, está com erro ao selecionar a coluna da medida e esse erro acontece por estar selecionando o nome da medida atribuída pelo usuário e não a coluna correta presente no banco.

**Figura 35:** Schema final correto.

```

▼<Schema name="Vendas">
  ▼<Dimension type="StandardDimension" visible="true" highCardinality="false"
    name="cliente">
    ▼<Hierarchy name="cliente" visible="true" hasAll="true" primaryKey="customerid">
      <Table name="customers"></Table>
      <Level name="nome" visible="true" table="customers" column="firstname"
        type="String" uniqueMembers="false" levelType="Regular" hideMemberIf="Never">
      </Level>
    </Hierarchy>
  </Dimension>
  ▼<Dimension type="StandardDimension" visible="true" highCardinality="false"
    name="produto">
    ▼<Hierarchy name="produto" visible="true" hasAll="true" primaryKey="prod_id">
      <Table name="products"></Table>
      <Level name="titulo" visible="true" table="products" column="title" type="String"
        uniqueMembers="false" levelType="Regular" hideMemberIf="Never"></Level>
    </Hierarchy>
  </Dimension>
  ▼<Dimension type="TimeDimension" visible="true" highCardinality="false" name="tempo">
    ▼<Hierarchy visible="true" hasAll="true" primaryKey="chave_tempo">
      <Table name="tempo"></Table>
      <Level name="ano" visible="true" table="tempo" column="ano" type="String"
        uniqueMembers="false" levelType="TimeYears" hideMemberIf="Never"></Level>
      <Level name="mes" visible="true" table="tempo" column="mes" type="String"
        uniqueMembers="false" levelType="TimeMonths" hideMemberIf="Never"></Level>
      <Level name="dia" visible="true" table="tempo" column="dia_mes" type="String"
        uniqueMembers="false" levelType="TimeDays" hideMemberIf="Never"></Level>
    </Hierarchy>
  </Dimension>
  ▼<Cube name="vendas" visible="true" cache="true" enabled="true">
    <Table name="aluguel"></Table>
    <DimensionUsage source="cliente" name="cliente" visible="true"
      foreignKey="customerid" highCardinality="false"></DimensionUsage>
    <DimensionUsage source="produto" name="produto" visible="true" foreignKey="prod_id"
      highCardinality="false"></DimensionUsage>
    <DimensionUsage source="tempo" name="tempo" visible="true" foreignKey="chave_tempo"
      highCardinality="false"></DimensionUsage>
    <Measure name="qtd" column="quantity" aggregator="count" visible="true"></Measure>
    <Measure name="soma_valor_total" column="totalamount" aggregator="sum"
      visible="true"></Measure>
  </Cube>
</Schema>

```

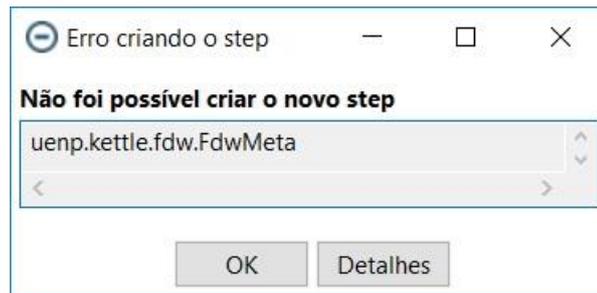
Fonte: Própria do autor.

Na Figura 35 o *Schema* foi corrigido e funciona na ferramenta OLAP. Além disso foi adicionado a dimensão tempo, destacada em laranja. Como mencionado anteriormente o plugin dimensão lógica de tempo não está funcionando e a dimensão tempo foi adicionada e configurada com a ferramenta *Schema Workbench*.

### 5.2.4 Plugin FDW

Ao tentar utilizar o *plugin* na ferramenta Kettle é exibido um erro, exibido na Figura 36, impossibilitando seu uso, porém o teste e o desenvolvimento foram concluídos pois não necessitaram de outras bases de dados.

**Figura 36:** Erro ao selecionar o plugin FDW.



Fonte: Própria do autor.

### 5.2.5 Erro ao gerar dados no plugin de dimensão tempo

Ao utilizar o *plugin* de dimensão tempo para gerar dados que servem para auxiliar no desenvolvimento do ambiente de negócios foi constatado que os dados gerados são todos do tipo text. Além disso ao gerar dados sobre os semestres o mesmo gera apenas dados pertinentes ao primeiro semestre, ou seja, inviabilizando seu uso em uma dimensão de tempo que necessite utilizar de dados relacionados a semestres. Na Figura 37 é exibido os dados gerados e é possível visualizar que para dados do mês 7 (julho) o semestre indicado ainda é o primeiro.

**Figura 37:** Dados gerados apenas do primeiro semestre.

numero_mes text	ano text	semestre text	semestre_ano text	semestre_numero text
7	2017	1º Sem	2017 - 1º Sem	1
7	2017	1º Sem	2017 - 1º Sem	1
7	2017	1º Sem	2017 - 1º Sem	1
7	2017	1º Sem	2017 - 1º Sem	1
7	2017	1º Sem	2017 - 1º Sem	1
7	2017	1º Sem	2017 - 1º Sem	1
7	2017	1º Sem	2017 - 1º Sem	1

Fonte: Própria do autor.

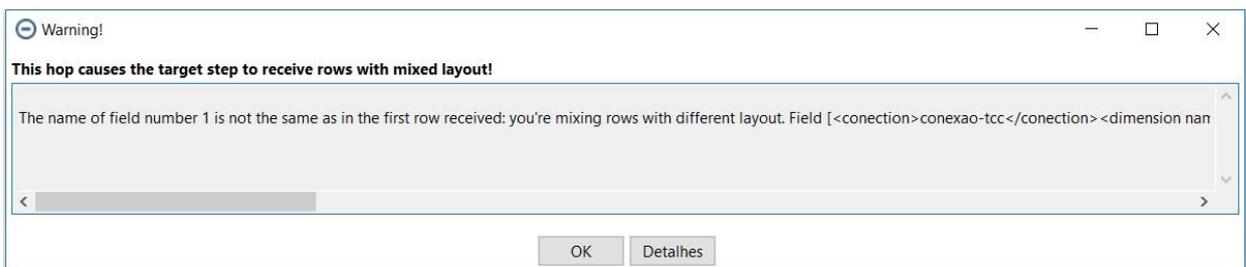
## 5.3 Alterações e Melhorias

Com o decorrer do desenvolvimento do ambiente de BI proposto foram observados algumas funções e ações que podem ser melhoradas.

### 5.3.1 Tela de aviso

Ao realizar uma conexão entre o plugin de dimensão e o plugin fato na ferramenta Kettle é gerado uma tela de aviso dando um alerta, mas não é dado um *feedback* claro para o usuário entender o alerta. O alerta é gerado também ao conectar os plugins de dimensão lógica e o plugin cubo. Embora apareça esse alerta, os plugins funcionam normalmente. O alerta destacado é exibido na Figura 38 abaixo. Uma alteração indicada seria deixar mais claro ao usuário o significado do alerta ou até mesmo removê-lo caso não seja necessário.

**Figura 38:** Alerta exibido ao conectar plugins de dimensão e fato.

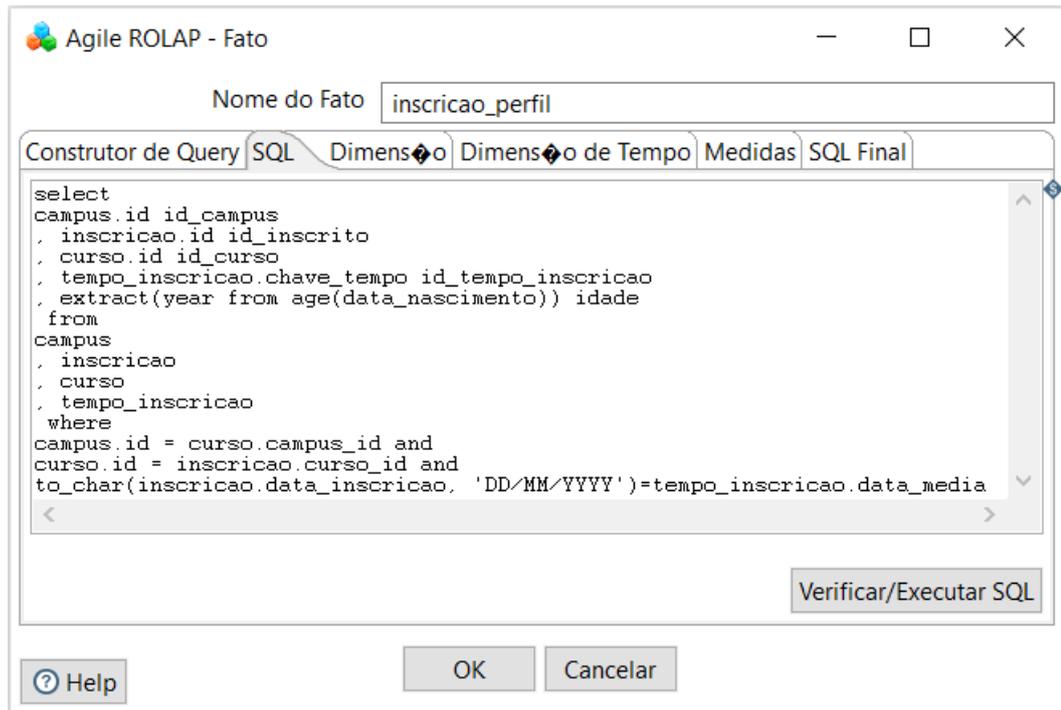


Fonte: Própria do autor.

### 5.3.2 Construtor de query do *plugin* fato

Ao utilizar o *plugin* fato para criar o SQL, na aba construtor de query, e selecionar os dados da base é necessário adicionar *AND* ao final da coluna “condições” caso deseje usar mais de uma condição no mesmo *SELECT*. Dessa forma é gerado um SQL na aba sql desse mesmo *plugin*. Além disso o plugin fato atual não permite utilizar outros comandos além de *WHERE* como *JOINS* e *WHEN* ao criar o SQL na aba construtor de query, isso o torna limitado, como exemplificado na aba sql do *plugin* fato na Figura 39, no qual o plugin gera um SQL com o comando *WHERE* automaticamente para o fato *inscricao\_perfil*.

**Figura 39:** SQL gerado através do *plugin fato*.



Fonte: Própria do autor.

Para utilizar um SQL no plugin fato para o fato resultado\_inscrito foi necessário utilizar uma versão antiga que permite utilizar um SQL mais elaborado. No qual não existe a aba de opção “Construtor de Query”, porém permite utilizar um SQL com mais opções de comandos além do WHERE. Uma sugestão de melhoria seria manter o “Construtor de Query”, mas deixar a aba de SQL editável para que seja possível utilizar um SQL mais avançado, como mostrado na Figura 40, o SQL utilizado para o fato resultado\_inscrito.

**Figura 40:** SQL gerado para o *plugin* fato.

Nome do Fato resultado\_inscrito

SQL Dimensões Dimensão de Tempo Medidas

```

SELECT
case
when
resultado_social.homologacao!=coalesce(resultado_universal.homologacao, 0)
or
resultado_social.homologacao!=coalesce(resultado_sociorracial.homologacao, 0)
then resultado_social.homologacao

when
resultado_universal.homologacao!=coalesce(resultado_social.homologacao, 0)
or
resultado_universal.homologacao!=coalesce(resultado_sociorracial.homologacao, 0)
then resultado_universal.homologacao

end as id_inscrito

-----

.case
when
resultado_social.pontuacao!=coalesce(resultado_universal.pontuacao, 0)
or
resultado_social.pontuacao!=coalesce(resultado_sociorracial.pontuacao, 0)
then resultado_social.pontuacao

when
resultado_universal.pontuacao!=coalesce(resultado_social.pontuacao, 0)
or
resultado_universal.pontuacao!=coalesce(resultado_sociorracial.pontuacao, 0)
then resultado_universal.pontuacao

end as pontuacao

-----

.case
when
resultado_social.curso_id!=coalesce(resultado_universal.curso_id, 0)
or
resultado_social.curso_id!=coalesce(resultado_sociorracial.curso_id, 0)
then resultado_social.curso_id

when
resultado_universal.curso_id!=coalesce(resultado_social.curso_id, 0)
or
resultado_universal.curso_id!=coalesce(resultado_sociorracial.curso_id, 0)
then resultado_universal.curso_id

end as id_curso

-----

.case resultado_universal.situacao
when 'Classificado' then 1 else 0
end as classificado_universal,
case resultado_universal.situacao
when 'Lista de Espera' then 1 else 0
end as espera_universal,
case resultado_universal.situacao
when 'Desclassificado por nota de corte' then 1 else 0
end as desclassificado_universal,
case resultado_social.situacao
when 'Classificado' then 1 else 0
end as classificado_social,
case resultado_social.situacao
when 'Lista de Espera' then 1 else 0
end as espera_social,
case resultado_social.situacao
when 'Desclassificado por nota de corte' then 1 else 0
end as desclassificado_social,

-----

case resultado_sociorracial.situacao
when 'Classificado' then 1 else 0
end as classificado_sociorracial,
case resultado_sociorracial.situacao
when 'Lista de Espera' then 1 else 0
end as espera_sociorracial,
case resultado_sociorracial.situacao
when 'Desclassificado por nota de corte' then 1 else 0
end as desclassificado_sociorracial
FROM resultado_universal

FULL OUTER JOIN resultado_social
on resultado_universal.homologacao = resultado_social.homologacao
FULL OUTER JOIN resultado_sociorracial
on resultado_social.homologacao = resultado_sociorracial.homologacao

```

Verificar SQL

Help OK Cancelar

Fonte: Própria do autor.

### 5.3.3 Plugin tempo

Foi necessário utilizar o plugin tempo em uma transformação separada da principal pois a mesma não estava funcionando em conjunto com o plugin fato. Abaixo na Figura 41 é demonstrado a estrutura da transformação que foi executada separada para criar a tabela tempo.

**Figura 41:** Transformação para usar *plugin* tempo.



Fonte: Própria do autor.

## 6. CONSIDERAÇÕES FINAIS

Foi possível criar um ambiente de análise para os dados de vestibulares de uma universidade demonstrando assim a eficiência do método *Agile* ROLAP.

O *plugin* XMLIn não é mais necessário pois foi integrado aos *plugins* de dimensão lógica. O *plugin* FDW como relatado previamente não foi necessário nesse trabalho pois foi utilizado apenas uma base de dados que não está em operação, contudo ao tentar ser utilizado para testes apresentou um erro mencionado também previamente. Mesmo com alguns *plugins* apresentando erros ou não estarem funcionando foi possível realizar o desenvolvimento do ambiente de BI.

A ferramenta *Schema Workbench* não fazia parte do conjunto de ferramentas a serem utilizadas com o método *Agile* ROLAP, porém foi necessário seu uso para o desenvolvimento das dimensões lógicas de tempo.

Vale ressaltar que caso todos os *plugins* estivessem funcionando corretamente o desenvolvimento do trabalho teria sido mais ágil e simples.

A metodologia acessa os dados diretamente de suas fontes e para realizar uma análise é necessário selecionar os dados corretamente. Para isso é indicado possuir ao menos o básico de conhecimento em SQL para realizar as consultas e selecionar os dados, porém o *plugin* fato auxilia nessa etapa de seleção e também a criar uma *View* na base de dados.

A metodologia apresenta como principal vantagem sua rápida implantação pois diferente das metodologias tradicionais que realizam o processo de ETL por completo o *Agile* ROLAP busca realizar um mapeamento nas bases de dados e selecioná-los ainda em seu formato original. Por esse motivo em empresas/organizações que possuam muitos dados ou muitas bases de dados o método não é indicado.

Em síntese, a implantação através do método é indicada quando há poucas bases de dados a serem integradas e/ou quando busca apenas análises mais específicas, como o caso desse trabalho que desenvolveu um ambiente de BI que analisa apenas os dados do vestibular da universidade.

## 7. TRABALHOS FUTUROS

Ao desenvolver o projeto foi observado que alguns *plugins* não funcionaram como esperado e adaptações foram necessárias, contudo o método foi utilizado e o ambiente de BI foi desenvolvido sem qualquer problema.

Com o projeto finalizado foi possível analisar dados do vestibular realizado no ano de 2017 e encontrar informações que podem ser úteis para os próximos vestibulares da instituição. Além disso o método *Agile* ROLAP mostrou-se capaz de ser utilizado para o desenvolvimento de um ambiente de BI. Contudo, os dados utilizados nesse trabalho foram somente de uma base, para tirar maior proveito do método e gerar um ambiente de análise mais completo é indicado integrar com dados de outras bases e de outros anos da instituição.

Uma documentação atualizada expondo como utilizar a metodologia e seus *plugins* auxiliaria para que o método fosse mais difundido e para que houvessem mais aplicações com bases de dados reais.

Por esses motivos, citados anteriormente, como trabalhos futuros é indicado:

- Desenvolver mais cubos para análise dos dados, mas em perspectivas diferentes das apresentadas nesse trabalho;
- Integrar com outros dados da UENP e com o sistema acadêmico a fim de obter análises e relatórios cada vez mais completos;
- Corrigir e melhorar alguns *plugins*;
- Desenvolver uma documentação atualizada da metodologia e *plugins*.

## REFERÊNCIAS

- ABELLERA, R; BULUSU, L. **Oracle Business Intelligence and Essbase Solutions Guide**. CRC Press, 2017.
- AGUIAR, E, C, F. **Implantando a Metodologia Agile ROLAP – Uma Metodologia Ágil de Desenvolvimento para Sistemas de Inteligência do Negócio**. Trabalho de Conclusão de Curso, UENP, 2016.
- BALLARD, C; FARRELL, D, M; GUPTA, A; MAZUELA, C; VOHNIK, S. **Dimensional Modeling In a Business Intelligence Environment**. Redbooks, 2006.
- BARBIERI, C. **BI2 – Business Intelligence: modelagem e qualidade**. Elsevier, 2011.
- CECI, F. **Business Intelligence**. Palhoça, Unisul, 2012.
- FONSECA, J. J. S. **Metodologia da pesquisa científica**. Apostila. Fortaleza: UECE, 2002.
- GIL, A, C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.
- HYDE, J. **Mondrian Documentation**. Disponível em: <https://mondrian.pentaho.com/documentation/olap.php>. Acesso em: 01 de maio de 2018.
- KIMBALL, R; ROSS, M. **The Data Warehouse Toolkit - The Definitive Guide to Dimensional Modeling**. John Wiley & Sons, 2013.
- KIMBALL, R; CASERTA, J. **The Data Warehouse ETL Toolkit**. Wiley Publishing, Inc., 2004.
- LAUDON, K; LAUDON, J. **Sistemas de informação gerenciais**. 9. ed. São Paulo: Pearson Prentice Hall, 2011.
- LOH, S. **BI na Era do Big Data para Cientistas de Dados**. 1. ed. Porto Alegre, 2014.
- MACHADO, F, N, R. **Tecnologia e Projeto de Data Warehouse**. 3. ed. São Paulo: Érica, 2007.
- PENTAHO. **Data Integration**. Disponível em: <http://www.pentaho.com/product/data-integration>. Acesso em: 19 de outubro de 2017.
- PIVOT4J. **Home**. Disponível em: <http://www.pivot4j.org>. Acesso em: 19 de outubro de 2017.
- POSTGRESQL. **About**. Disponível em: <https://www.postgresql.org/about/>. Acesso em: 19 de outubro de 2017.

- PRIMAK, F, V. **Decisões com B.I. (Business Intelligence)**. Rio de Janeiro: Ciência Moderna, 2008.
- SIQUEIRA, M, C. **Gestão estratégica da informação**. Rio de Janeiro: Brasport, 2005.
- SILVA, J; SAIAS, J. **OLAP em âmbito hospitalar: transformação de dados de enfermagem para análise multidimensional**. Trabalho de Conclusão de Curso, Universidade de Évora, 2012.
- SOUZA, E. **Um Método Agile ROLAP para implantação de Ambientes de Inteligência de Negócios**. Trabalho de Conclusão de Curso, UENP, 2014.
- SOUZA, E; MENOLLI, A, L, A; COELHO, R, G. **Uma Metodologia Agile ROLAP para Implantação de Ambientes de Inteligência de Negócios**. São Francisco do Sul: X ERBD, 2014.
- NEGASH, S. **Business Intelligence**. Communications of the Association for Information Systems, vol. 13, p.177-195, 2004.
- TURBAN, E; SHARDA, R; ARONSON, J; KING, D. **Business Intelligence: Um enfoque gerencial para a inteligência do negócio**. Porto Alegre, Bookman, 2009.
- VAISMAN, A; ZIMÁNYI, E. **Data Warehouse Systems Design and Implementation**. 1. ed. Springer, 2014.
- WATSON, H, J. & WIXOM, B, H. **The Current State of Business Intelligence**. COMPUTER - IEEE COMPUTER SOCIETY, vol. 9, 2007.
- WAZLAWICK, R, S. **Metodologia de pesquisa para ciência da computação**. 2. ed. Elsevier, 2014.
- WESTERLUND, P. **Business Intelligence Multidimensional Data Analysis**. Master Thesis in Computing Science, 30 ECTS Credits, 2008.
- WILLIAMS, S; WILLIAMS, N. **The Profit Impact of Business Intelligence**. Elsevier, 2007.
- WOOD, S. **Mondrian Documentation**. Disponível em: <https://mondrian.pentaho.com/documentation/workbench.php>. Acesso em: 20 de maio 2018.



Figura 43: Atributos das tabelas.

curso	pagamento	questionario	respostas_d1	respostas_d2
id INT(11)	id INT(11)	id INT(11)	NOM_CANDIDATO VARCHAR(45)	NOM_CANDIDATO VARCHAR(45)
nome VARCHAR(100)	inscricao_id INT(11)	inscricao_id INT(11)	NUM_INSCRICAO INT(7)	NUM_INSCRICAO INT(7)
nome_25 VARCHAR(25)	titulo VARCHAR(40)	q01 VARCHAR(100)	Q_01 VARCHAR(1)	Q_01 VARCHAR(1)
nome_15 VARCHAR(15)	data_emissao DATE	q02 VARCHAR(100)	Q_02 VARCHAR(1)	Q_02 VARCHAR(1)
turno VARCHAR(20)	data_vencimento DATE	q03 VARCHAR(100)	Q_03 VARCHAR(1)	Q_03 VARCHAR(1)
vagas INT(11)	valor_titulo DOUBLE	q04 VARCHAR(100)	Q_04 VARCHAR(1)	Q_04 VARCHAR(1)
vagas_universal INT(11)	data_pagamento DATE	q05 VARCHAR(100)	Q_05 VARCHAR(1)	Q_05 VARCHAR(1)
vagas_social INT(11)	data_liquidacao DATE	q06 VARCHAR(100)	Q_06 VARCHAR(1)	Q_06 VARCHAR(1)
vagas_sociorracial INT(11)	valor_pago DOUBLE	q07 VARCHAR(100)	Q_07 VARCHAR(1)	Q_07 VARCHAR(1)
vagas_enem INT(11)		q08 VARCHAR(100)	Q_08 VARCHAR(1)	Q_08 VARCHAR(1)
total_vagas INT(11)		q09 VARCHAR(100)	Q_09 VARCHAR(1)	Q_09 VARCHAR(1)
duracao VARCHAR(20)		q10 VARCHAR(100)	Q_10 VARCHAR(1)	Q_10 VARCHAR(1)
campus_id INT(11)		q11 VARCHAR(100)	Q_11 VARCHAR(1)	Q_11 VARCHAR(1)
req_lingua_estrang VARCHAR(30)		q12 VARCHAR(100)	Q_12 VARCHAR(1)	Q_12 VARCHAR(1)
		q13 VARCHAR(100)	Q_13 VARCHAR(1)	Q_13 VARCHAR(1)
		q14 VARCHAR(100)	Q_14 VARCHAR(1)	Q_14 VARCHAR(1)
		q15 VARCHAR(100)	Q_15 VARCHAR(1)	Q_15 VARCHAR(1)
		q16 VARCHAR(100)	Q_16 VARCHAR(1)	Q_16 VARCHAR(1)
		q17 VARCHAR(100)	Q_17 VARCHAR(1)	Q_17 VARCHAR(1)
		q18 VARCHAR(100)	Q_18 VARCHAR(1)	Q_18 VARCHAR(1)
		q19 VARCHAR(100)	Q_19 VARCHAR(1)	Q_19 VARCHAR(1)
		q20 VARCHAR(100)	Q_20 VARCHAR(1)	Q_20 VARCHAR(1)
		q21 VARCHAR(100)	Q_21 VARCHAR(1)	Q_21 VARCHAR(1)
		q22 VARCHAR(100)	Q_22 VARCHAR(1)	Q_22 VARCHAR(1)
		q23 VARCHAR(100)	Q_23 VARCHAR(1)	Q_23 VARCHAR(1)
		q24 VARCHAR(100)	Q_24 VARCHAR(1)	Q_24 VARCHAR(1)
		q25 VARCHAR(100)	Q_25 VARCHAR(1)	Q_25 VARCHAR(1)
		q26 VARCHAR(100)	Q_26 VARCHAR(1)	Q_26 VARCHAR(1)
			Q_27 VARCHAR(1)	Q_27 VARCHAR(1)
			Q_28 VARCHAR(1)	Q_28 VARCHAR(1)
			Q_29 VARCHAR(1)	Q_29 VARCHAR(1)
			Q_30 VARCHAR(1)	Q_30 VARCHAR(1)
			Q_31 VARCHAR(1)	Q_31 VARCHAR(1)
			Q_32 VARCHAR(1)	Q_32 VARCHAR(1)
			Q_33 VARCHAR(1)	Q_33 VARCHAR(1)
			Q_34 VARCHAR(1)	Q_34 VARCHAR(1)
			Q_35 VARCHAR(1)	Q_35 VARCHAR(1)
			Q_36 VARCHAR(1)	Q_36 VARCHAR(1)
			Q_37 VARCHAR(1)	Q_37 VARCHAR(1)
			Q_38 VARCHAR(1)	Q_38 VARCHAR(1)
			Q_39 VARCHAR(1)	Q_39 VARCHAR(1)
			Q_40 VARCHAR(1)	Q_40 VARCHAR(1)
			FLG_FALTA VARCHAR(1)	Q_41 VARCHAR(1)
				Q_42 VARCHAR(1)
				Q_43 VARCHAR(1)
				Q_44 VARCHAR(1)
				17 more...

pesos	redacao	resultado_disciplina	resultado_segunda_opcao
curso VARCHAR(24)	homologacao INT(7)	homologacao INT(11)	homologacao INT(11)
campus VARCHAR(3)	n1 VARCHAR(3)	lplb INT(11)	candidato TEXT
turno VARCHAR(12)	n2 VARCHAR(3)	lem INT(11)	rg TEXT
id INT(2)	ntc INT(1)	mat INT(11)	pontuacao DOUBLE
red INT(1)	n3 VARCHAR(3)	fis INT(11)	curso_id INT(11)
lplb INT(1)	media VARCHAR(5)	qui INT(11)	curso TEXT
lem INT(1)		bio INT(11)	situacao TEXT
mat INT(1)		his INT(11)	classificacao INT(11)
fis INT(1)		geo INT(11)	
qui INT(1)		soc INT(11)	
bio INT(1)		fil INT(11)	
his INT(1)		red DOUBLE	
geo INT(1)		acertos INT(11)	
soc INT(1)		eliminado_redacao TINYINT(1)	
fil INT(1)			
total INT(2)			
max INT(3)			

Fonte: Própria do autor.

Figura 44: Atributos das tabelas.

resultado_op1	resultado_op2	resposta	inscricao	resultado	resultado_universal	resultado_sociorracial	
<ul style="list-style-type: none"> <li>homologacao INT(11)</li> <li>lplb DOUBLE</li> <li>lem DOUBLE</li> <li>mat DOUBLE</li> <li>fis DOUBLE</li> <li>qui DOUBLE</li> <li>bio DOUBLE</li> <li>his DOUBLE</li> <li>geo DOUBLE</li> <li>soc DOUBLE</li> <li>fil DOUBLE</li> <li>red DOUBLE</li> <li>total DOUBLE</li> <li>pontuacao_max DOUBLE</li> <li>curso_id INT(11)</li> <li>pontuacao DOUBLE</li> </ul>	<ul style="list-style-type: none"> <li>homologacao INT(11)</li> <li>lplb DOUBLE</li> <li>lem DOUBLE</li> <li>mat DOUBLE</li> <li>fis DOUBLE</li> <li>qui DOUBLE</li> <li>bio DOUBLE</li> <li>his DOUBLE</li> <li>geo DOUBLE</li> <li>soc DOUBLE</li> <li>fil DOUBLE</li> <li>red DOUBLE</li> <li>total DOUBLE</li> <li>pontuacao_max DOUBLE</li> <li>curso_id INT(11)</li> <li>pontuacao DOUBLE</li> </ul>	<ul style="list-style-type: none"> <li>homologacao INT(11)</li> <li>p1_q01 CHAR(1)</li> <li>p1_q02 CHAR(1)</li> <li>p1_q03 CHAR(1)</li> <li>p1_q04 CHAR(1)</li> <li>p1_q05 CHAR(1)</li> <li>p1_q06 CHAR(1)</li> <li>p1_q07 CHAR(1)</li> <li>p1_q08 CHAR(1)</li> <li>p1_q09 CHAR(1)</li> <li>p1_q10 CHAR(1)</li> <li>p1_q11 CHAR(1)</li> <li>p1_q12 CHAR(1)</li> <li>p1_q13 CHAR(1)</li> <li>p1_q14 CHAR(1)</li> <li>p1_q15 CHAR(1)</li> <li>p1_q16 CHAR(1)</li> <li>p1_q17 CHAR(1)</li> <li>p1_q18 CHAR(1)</li> <li>p1_q19 CHAR(1)</li> <li>p1_q20 CHAR(1)</li> <li>p1_q21 CHAR(1)</li> <li>p1_q22 CHAR(1)</li> <li>p1_q23 CHAR(1)</li> <li>p1_q24 CHAR(1)</li> <li>p1_q25 CHAR(1)</li> <li>p1_q26 CHAR(1)</li> <li>p1_q27 CHAR(1)</li> <li>p1_q28 CHAR(1)</li> <li>p1_q29 CHAR(1)</li> </ul>	<ul style="list-style-type: none"> <li>id INT(11)</li> <li>curso_id INT(11)</li> <li>campus_prova_id INT(11)</li> <li>lingua_estrangeira VARCHAR(45)</li> <li>campus_anterior INT(11)</li> <li>segunda_opcao TINYINT(1)</li> <li>segunda_opcao_curso_id INT(11)</li> <li>cota_social TINYINT(1)</li> <li>cota_sociorracial TINYINT(1)</li> <li>declaracao_ensino_medio TINYINT(1)</li> <li>declaracao_curso_superior TINYINT(1)</li> <li>declaracao_grupo_racial TINYINT(1)</li> <li>declaracao_homologacao TINYINT(1)</li> <li>nome VARCHAR(45)</li> <li>cpf VARCHAR(45)</li> <li>rg VARCHAR(45)</li> <li>emissor_rg VARCHAR(10)</li> <li>estado_emissor_rg VARCHAR(2)</li> <li>data_nascimento DATE</li> <li>sexo CHAR(1)</li> <li>etnia VARCHAR(45)</li> <li>canhoto TINYINT(1)</li> <li>necessidades_especiais TINYINT(1)</li> <li>tipo_necessidade_especial VARCHAR(45)</li> <li>banca_especial TINYINT(1)</li> <li>endereco VARCHAR(100)</li> <li>numero INT(11)</li> <li>complemento VARCHAR(45)</li> <li>cidade VARCHAR(200)</li> <li>bairro VARCHAR(45)</li> <li>estado VARCHAR(2)</li> <li>cep INT(11)</li> <li>telefone VARCHAR(45)</li> <li>celular VARCHAR(45)</li> </ul>	<ul style="list-style-type: none"> <li>email VARCHAR(100)</li> <li>data_inscricao DATE</li> <li>ip_origem VARCHAR(100)</li> <li>id_cobranca VARCHAR(200)</li> <li>valor DOUBLE</li> <li>senha VARCHAR(200)</li> <li>isento TINYINT(1)</li> <li>pago TINYINT(1)</li> <li>sala_id INT(11)</li> <li>seq_sala INT(11)</li> <li>homologacao INT(11)</li> <li>usar_nota_enem TINYINT(1)</li> <li>ano_enem INT(11)</li> <li>id_enem VARCHAR(100)</li> </ul>	<ul style="list-style-type: none"> <li>homologacao INT(11)</li> <li>d1_q01 INT(11)</li> <li>d1_q02 INT(11)</li> <li>d1_q03 INT(11)</li> <li>d1_q04 INT(11)</li> <li>d1_q05 INT(11)</li> <li>d1_q06 INT(11)</li> <li>d1_q07 INT(11)</li> <li>d1_q08 INT(11)</li> <li>d1_q09 INT(11)</li> <li>d1_q10 INT(11)</li> <li>d1_q11 INT(11)</li> <li>d1_q12 INT(11)</li> <li>d1_q13 INT(11)</li> <li>d1_q14 INT(11)</li> <li>d1_q15 INT(11)</li> <li>d1_q16 INT(11)</li> <li>d1_q17 INT(11)</li> <li>d1_q18 INT(11)</li> <li>d1_q19 INT(11)</li> <li>d1_q20 INT(11)</li> <li>d1_q21 INT(11)</li> <li>d1_q22 INT(11)</li> <li>d1_q23 INT(11)</li> <li>d1_q24 INT(11)</li> <li>d1_q25 INT(11)</li> </ul>	<ul style="list-style-type: none"> <li>homologacao INT(11)</li> <li> candidato TEXT</li> <li> rg TEXT</li> <li> pontuacao DOUBLE</li> <li> curso_id INT(11)</li> <li> curso TEXT</li> <li> situacao TEXT</li> <li> classificacao INT(11)</li> </ul>	<ul style="list-style-type: none"> <li>homologacao INT(11)</li> <li> candidato TEXT</li> <li> rg TEXT</li> <li> pontuacao DOUBLE</li> <li> curso_id INT(11)</li> <li> curso TEXT</li> <li> situacao TEXT</li> <li> classificacao INT(11)</li> </ul>
<p>Indexes</p> <p>PRIMARY</p>	<p>Indexes</p> <p>PRIMARY</p>	<p>Indexes</p> <p>PRIMARY</p>	<p>Indexes</p> <p>PRIMARY</p>	<p>Indexes</p> <p>PRIMARY</p>	<p>Indexes</p> <p>PRIMARY</p>	<p>Indexes</p> <p>PRIMARY</p>	

Fonte: Própria do autor.

Figura 45 – Atributos das tabelas.

gabarito	v_rpt_chamadas_sala	v_rpt_inscricao_sala	v_ensalamento_cops
id CHAR(1)	homologacao INT(11)	homologacao INT(11)	NUM_INSCRICAO INT(11)
lingua_estrangeira VARCHAR(45)	nome VARCHAR(45)	nome VARCHAR(45)	NOM_CANDIDATO VARCHAR(45)
d1_q01 CHAR(1)	cpf VARCHAR(45)	data_nascimento DATE	SEXO CHAR(1)
d1_q02 CHAR(1)	rg VARCHAR(59)	rg VARCHAR(48)	COD_CARGO INT(1)
d1_q03 CHAR(1)	data_nascimento VARCHAR(10)	cpf VARCHAR(45)	DESC_CARGO VARCHAR(45)
d1_q04 CHAR(1)	sala_id INT(11)	canhoto TINYINT(1)	COD_LOCAL_CONC INT(11)
d1_q05 CHAR(1)	sala VARCHAR(10)	curso_id INT(11)	NOM_LOCAL VARCHAR(50)
d1_q06 CHAR(1)	bloco VARCHAR(50)	curso VARCHAR(100)	COD_SALA_CONC INT(11)
d1_q07 CHAR(1)	local VARCHAR(100)	turno VARCHAR(20)	CIDADE_PROVA VARCHAR(100)
d1_q08 CHAR(1)		campus_curso VARCHAR(10)	NUM_SEQ INT(11)
d1_q09 CHAR(1)		sala_id INT(11)	LOCAL_SALA_SEQ VARCHAR(35)
d1_q10 CHAR(1)		nome_local VARCHAR(50)	COD_CURSO INT(11)
d1_q11 CHAR(1)		cidade VARCHAR(100)	CURSO VARCHAR(135)
d1_q12 CHAR(1)		campus_prova_id INT(11)	
d1_q13 CHAR(1)		campus_prova VARCHAR(10)	
d1_q14 CHAR(1)			
d1_q15 CHAR(1)			
d1_q16 CHAR(1)			
d1_q17 CHAR(1)			
d1_q18 CHAR(1)			
d1_q19 CHAR(1)			
d1_q20 CHAR(1)			
d1_q21 CHAR(1)			
d1_q22 CHAR(1)			
d1_q23 CHAR(1)			
d1_q24 CHAR(1)			
d1_q25 CHAR(1)			
d1_q26 CHAR(1)			
d1_q27 CHAR(1)			
d1_q28 CHAR(1)			
72 more...			
<b>Indexes</b>			
PRIMARY			

v_salas_candidatos	v_homologacao_edital	v_transferencia	v_transferidos	v_pagamentos
id INT(11)	id INT(11)	id INT(11)	id INT(11)	titulo VARCHAR(40)
nome VARCHAR(10)	nome VARCHAR(45)	nome VARCHAR(45)	nome VARCHAR(45)	inscricao_id INT(11)
capacidade INT(11)	rg VARCHAR(48)	cpf VARCHAR(45)	cpf VARCHAR(45)	nome VARCHAR(45)
candidatos BIGINT(21)		telefone VARCHAR(45)	email VARCHAR(100)	cpf VARCHAR(45)
nome_local VARCHAR(50)		celular VARCHAR(45)	telefone VARCHAR(45)	data_emissao DATE
		email VARCHAR(100)	celular VARCHAR(45)	data_vencimento DATE
		data_nascimento DATE	campus_anterior VARCHAR(10)	valor_titulo DOUBLE
		etnia VARCHAR(45)	campus_atual VARCHAR(10)	data_pagamento DATE
		sexo CHAR(1)		data_liquidacao DATE
		cota_social TINYINT(1)		valor_pago DOUBLE
		cota_sociorracial TINYINT(1)		
		cidade VARCHAR(200)		
		estado VARCHAR(2)		
		cep INT(11)		
		isento TINYINT(1)		
		pago TINYINT(1)		
		data_pagamento DATE		
		campus_id INT(11)		
		campus VARCHAR(100)		
		campus_cidade VARCHAR(100)		

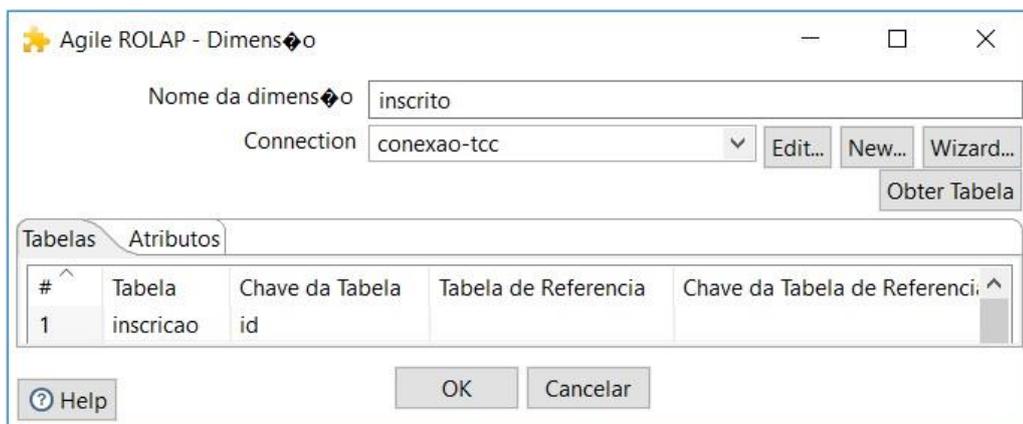
v_fiscais_aceitos	v_rpt_lista_sala
nome VARCHAR(45)	num_homologacao INT(11)
cpf VARCHAR(45)	nome VARCHAR(45)
email VARCHAR(100)	cpf VARCHAR(45)
campus VARCHAR(10)	rg VARCHAR(59)
nome_local VARCHAR(50)	data_nascimento VARCHAR(10)
selecionado TINYINT(1)	local_id INT(11)
aceito TINYINT(1)	local VARCHAR(50)
data_aceite DATE	sala VARCHAR(10)
	sala_id INT(11)

Fonte: Própria do autor.

## APÊNDICE B – CONFIGURAÇÕES DO DESENVOLVIMENTO DOS CUBOS

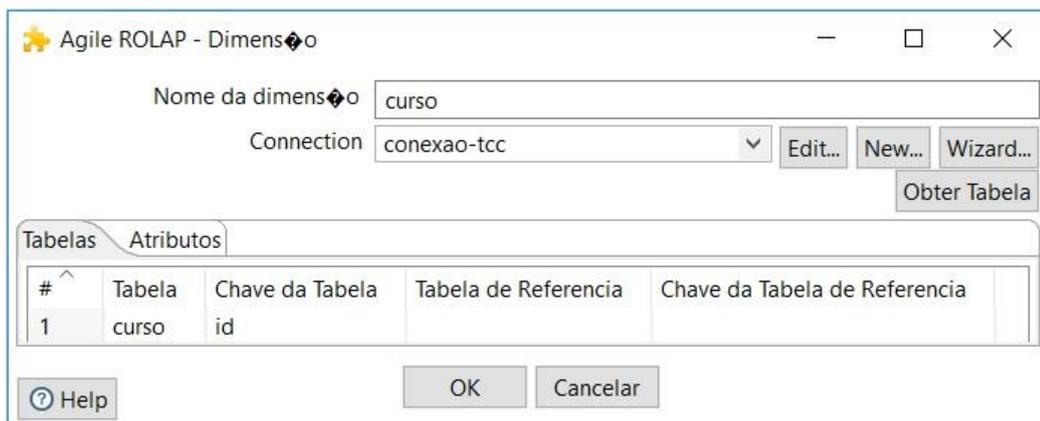
Foram desenvolvidos três cubos ao todo e no desenvolvimento desse trabalho é apresentado as principais configurações realizadas nos *plugins* e ferramentas para o cubo *inscricao\_perfil*. Nesse apêndice estão as configurações restantes do cubo *inscricao\_perfil* e todas aquelas utilizadas nos cubos *pagamento\_inscrito* e *resultado\_inscrito*.

**Figura 46:** Tabela inscrição selecionada para a dimensão inscrito do fato *inscricao\_perfil*.



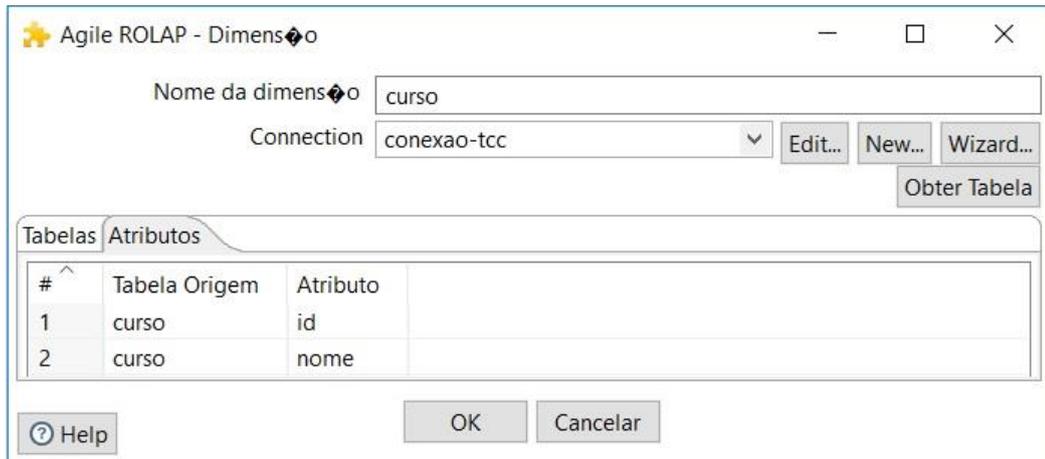
Fonte: Própria do autor.

**Figura 47:** Tabela curso selecionada para a dimensão curso do fato *inscricao\_perfil*.



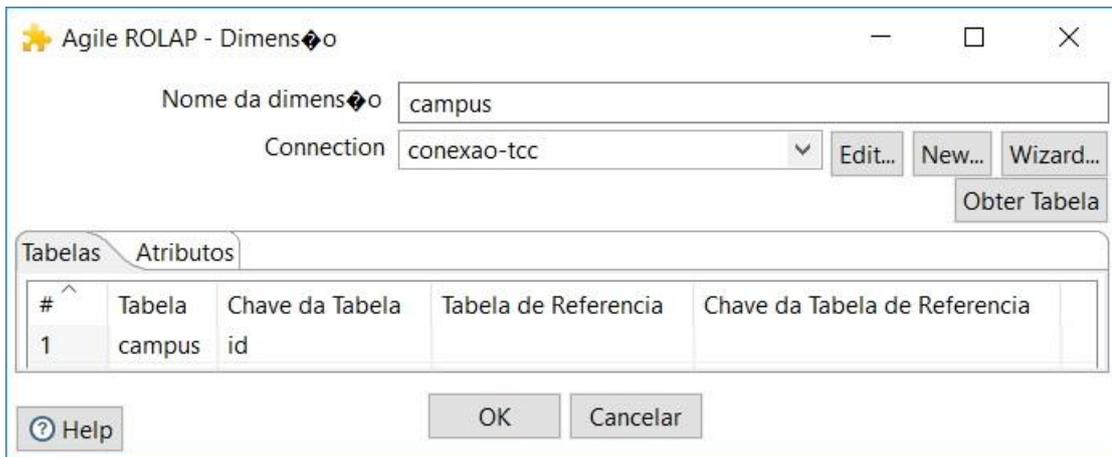
Fonte: Própria do autor.

**Figura 48:** Atributos selecionados para a dimensão curso do fato inscrição\_perfil.



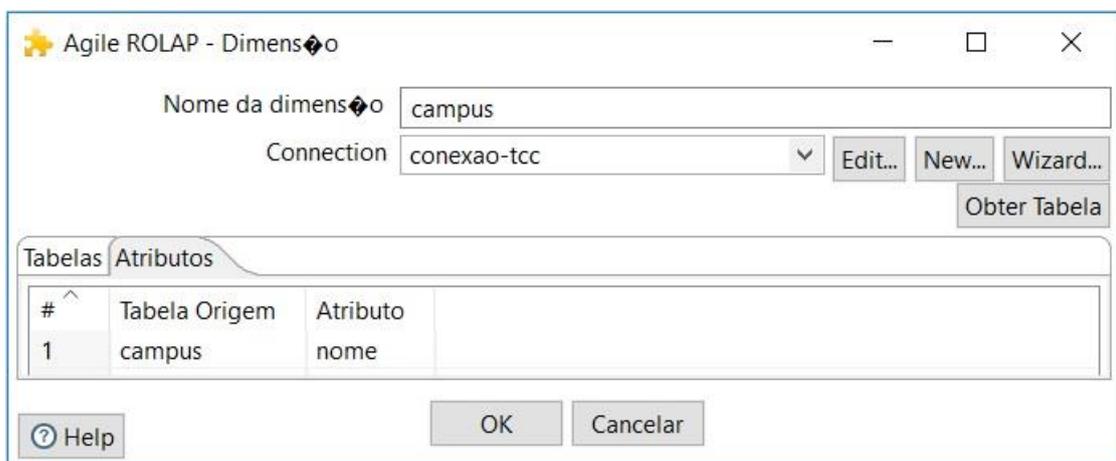
Fonte: Própria do autor.

**Figura 49:** Tabela campus selecionada para a dimensão campus do fato inscrição\_perfil.



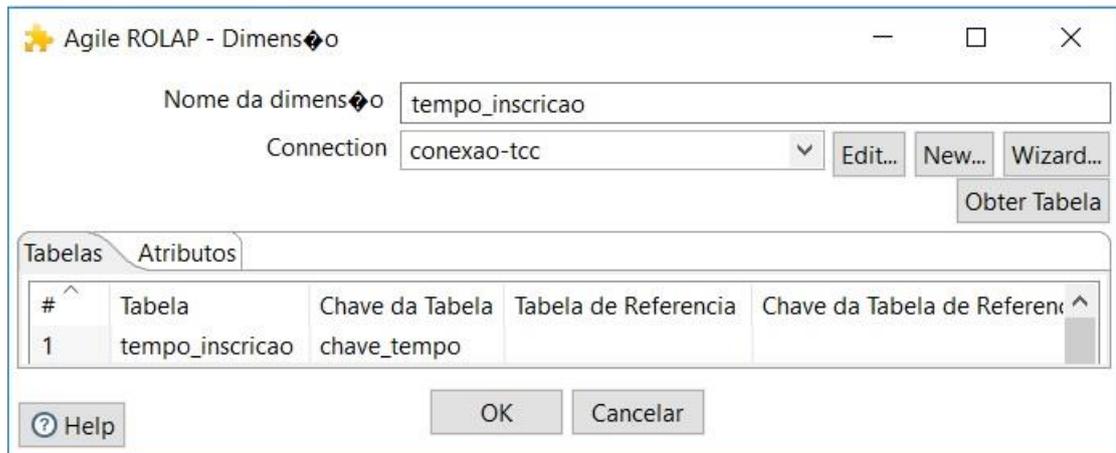
Fonte: Própria do autor.

**Figura 50:** Atributos selecionados para a dimensão campus do fato inscrição\_perfil.



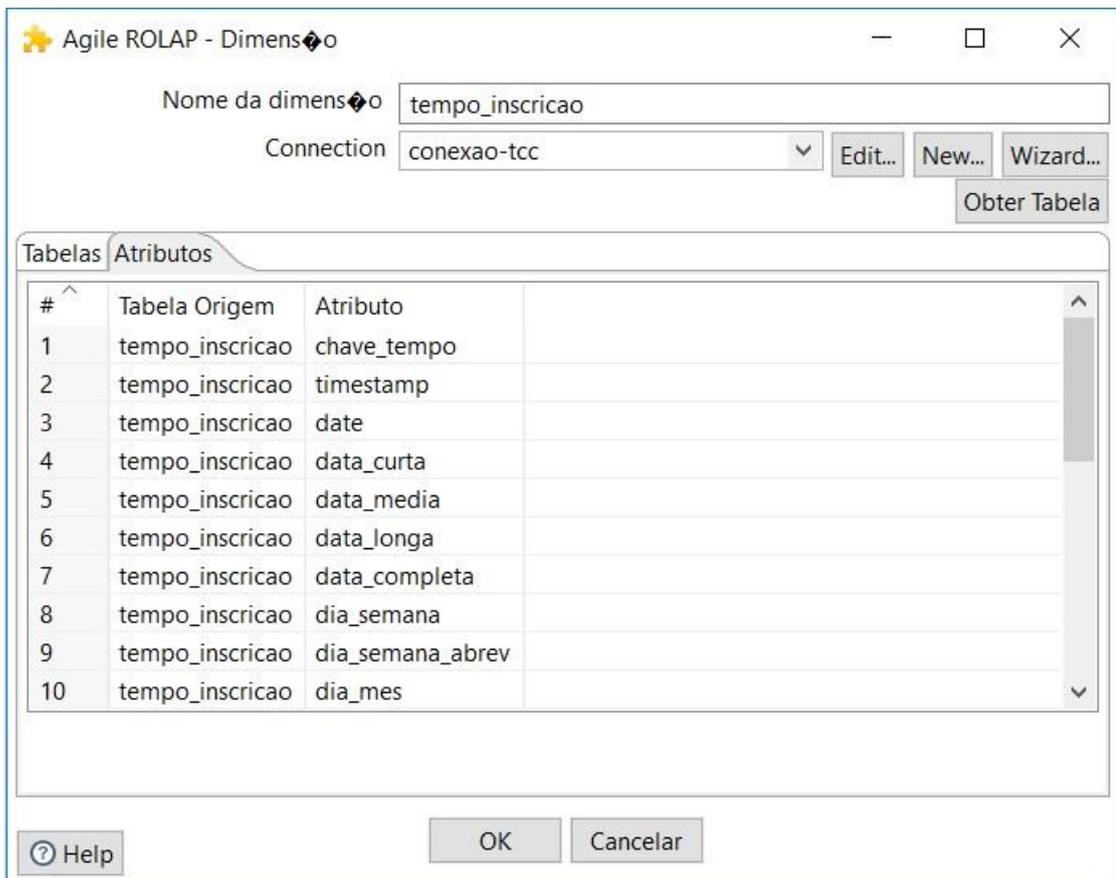
Fonte: Própria do autor.

**Figura 51:** Tabela tempo\_inscricao selecionada para a dimensão tempo\_inscricao do fato inscrição\_perfil.

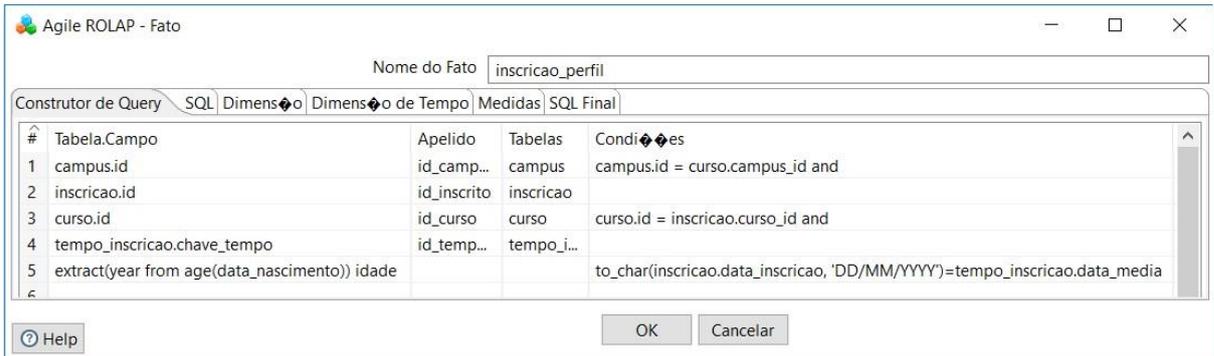


Fonte: Própria do autor.

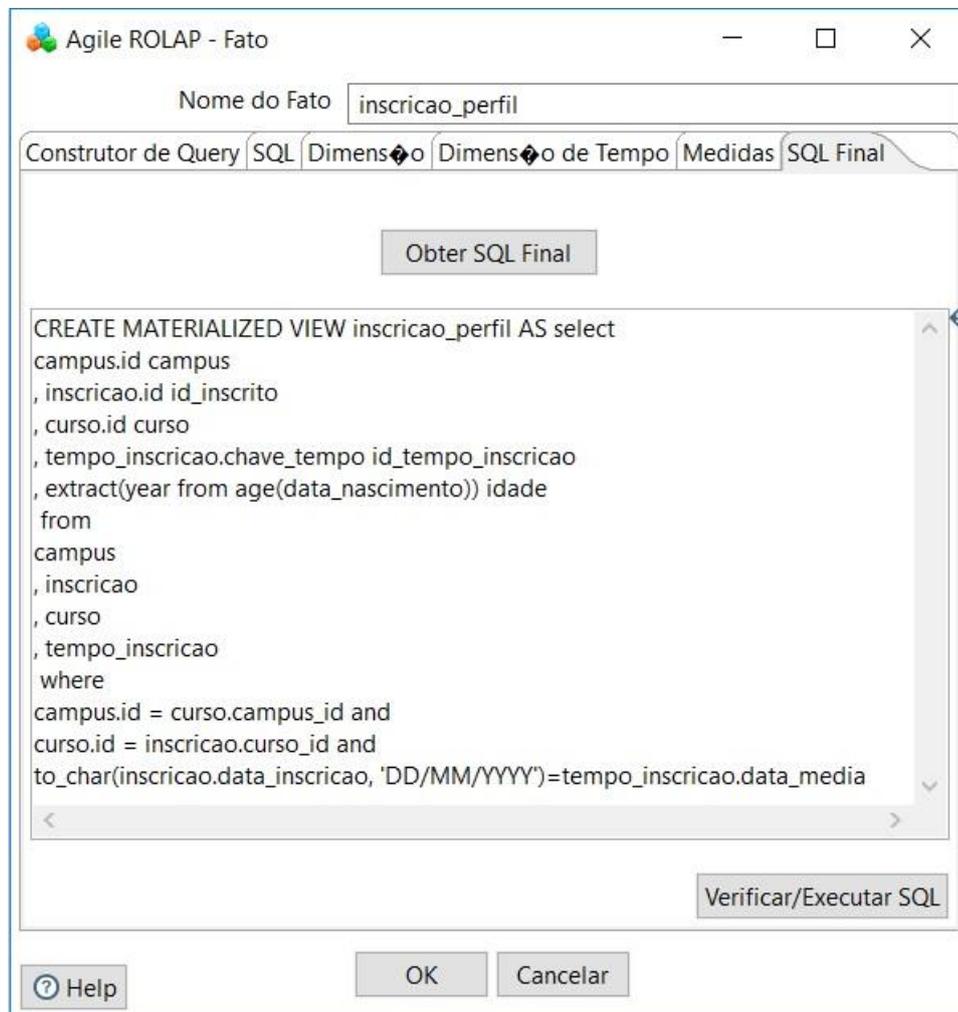
**Figura 52:** Alguns dos atributos selecionados para a dimensão tempo\_inscricao do fato inscrição\_perfil.



Fonte: Própria do autor.

**Figura 53:** Construtor de Query do fato inscricao\_perfil.

Fonte: Própria do autor.

**Figura 54:** SQL da view gerado no plugin fato para o fato inscricao\_perfil.

Fonte: Própria do autor.

**Figura 55:** Configuração da dimensão lógica campus para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - CAMPUS

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecione dimensão: campus

Nome da dimensão: campus

Visível:

Hierarquia	Nível
campus	nome_campus

Hierarquia

Nome: campus

Visível:

Nível

Nome: nome\_campus

Coluna: campus - nome

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 56:** Configuração da dimensão lógica curso para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - CURSO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: curso

Nome da dimensão: curso

Visível:

Hierarquia	Nível
curso	nome_curso

Hierarquia

Nome: curso

Visível:

Nível

Nome: nome\_curso

Coluna: curso - nome

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 57:** Configuração da dimensão lógica sexo para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - SEXO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: sexo

Visível:

Hierarquia	Nível
sexo	sexo_inscrito

Hierarquia

Nome: sexo

Visível:

Nível

Nome: sexo\_inscrito

Coluna: inscricao - sexo

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 58:** Configuração da dimensão lógica cotas para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - COTAS

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: cotas

Visível:

Hierarquia	Nível
cota_social	cota_social
cota_sociorracial	

Hierarquia

Nome: cota\_social

Visível:

Nível

Nome: cota\_social

Coluna: inscricao - cota\_socia

Tipo: Integer

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 59:** Configuração da dimensão lógica canhoto para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - CANHOTO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: canhoto

Visível:

Hierarquia	Nível
canhoto	canhoto

Hierarquia

Nome: canhoto

Visível:

Nível

Nome: canhoto

Coluna: inscricao - canhoto

Tipo: Integer

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 60:** Configuração da dimensão lógica nec\_especial para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - NEC\_ESPECIAL

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: nec\_especial

Visível:

Hierarquia	Nível
nec_especial	nec_especial
tipo_nec_especial	
banca_especial	

Hierarquia

Nome: nec\_especial

Visível:

Nível

Nome: nec\_especial

Coluna: inscricao - necessidade

Tipo: Integer

Visível:

Cima Baixo Cima Baixo

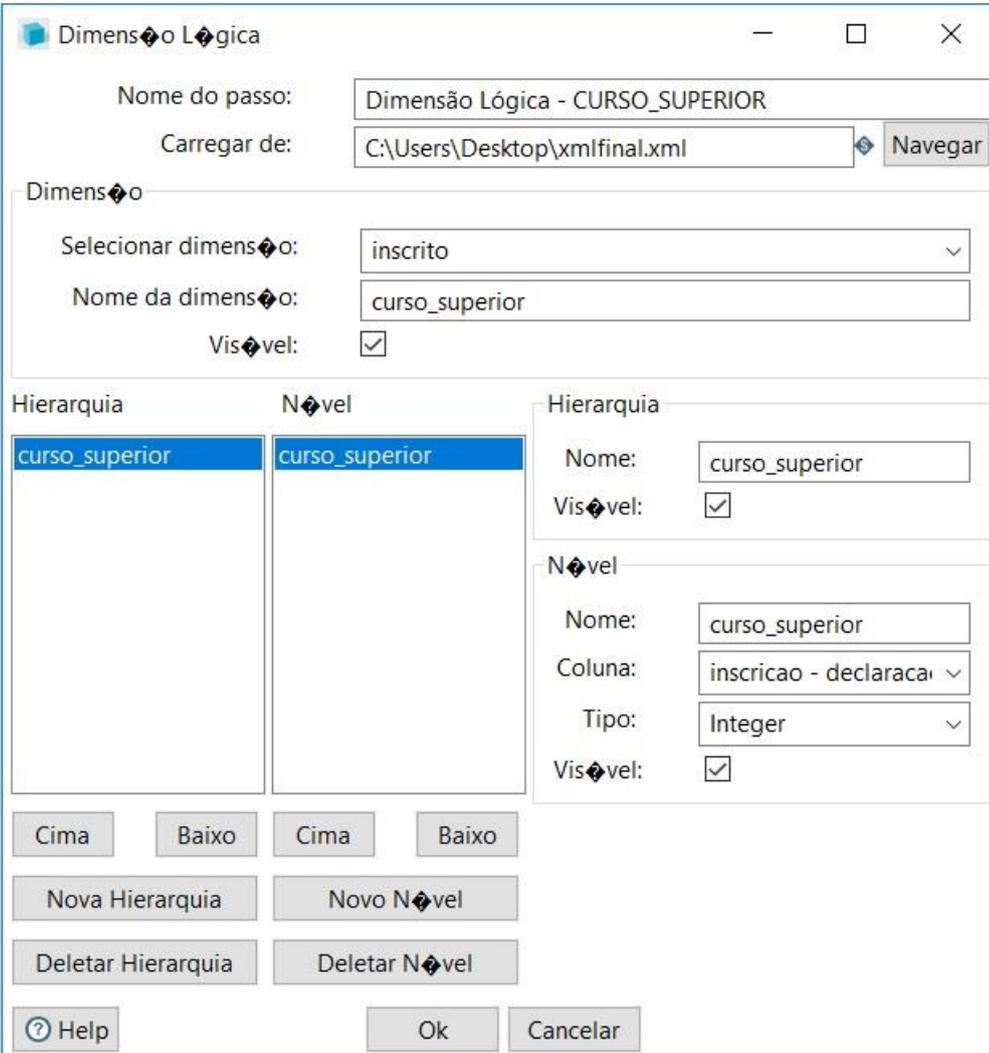
Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 61:** Configuração da dimensão lógica curso\_superior para o fato inscricao\_perfil.



Dimensão Lógica

Nome do passo: Dimensão Lógica - CURSO\_SUPERIOR

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: curso\_superior

Visível:

Hierarquia	Nível
curso_superior	curso_superior

Hierarquia

Nome: curso\_superior

Visível:

Nível

Nome: curso\_superior

Coluna: inscricao - declaraca

Tipo: Integer

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 62:** Configuração da dimensão lógica etnia para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - ETNIA

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: etnia

Visível:

Hierarquia	Nível
etnia	etnia

Hierarquia

Nome: etnia

Visível:

Nível

Nome: etnia

Coluna: inscricao - etnia

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 63:** Configuração da dimensão lógica lingua\_estrageira para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - LINGUA\_ESTRANGEIRA

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: lingua\_estrageira

Visível:

Hierarquia

lingua_estrageira
-------------------

Nível

lingua_estrageira
-------------------

Hierarquia

Nome: lingua\_estrageira

Visível:

Nível

Nome: lingua\_estrageira

Coluna: inscricao - lingua\_estr

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 64:** Configuração da dimensão lógica pagamento para o fato inscricao\_perfil.

Dimensão Lógica

Nome do passo: Dimensão Lógica - PAGAMENTO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: pagamento

Visível:

Hierarquia

Nível

Hierarquia

Nome: isento\_pagamento

Visível:

Nível

Nome: isento

Coluna: inscricao - isento

Tipo: Integer

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 65:** Tabela tempo\_pagamento selecionada para a dimensão tempo\_pagamento para o fato pagamento\_inscrito.

Agile ROLAP - Dimensão

Nome da dimensão: tempo\_pagamento

Connection: conexao-tcc Edit... New... Wizard... Obter Tabela

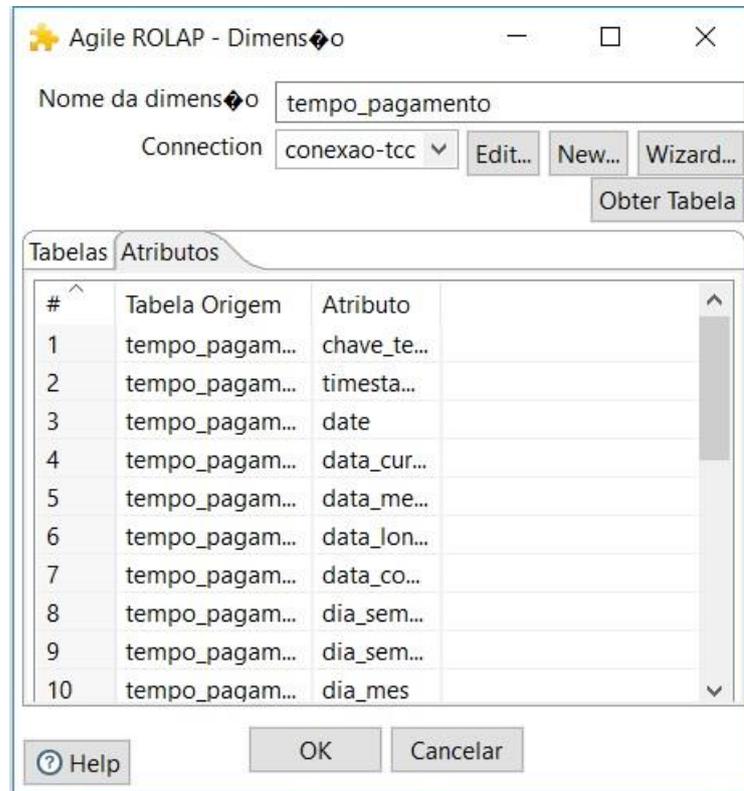
Tabelas Atributos

#	Tabela	Chave da Tabela	Tabela de Referencia	Chave da Tabela de Referencia
1	tempo_pagamento	chave_tempo		
2				

Help OK Cancelar

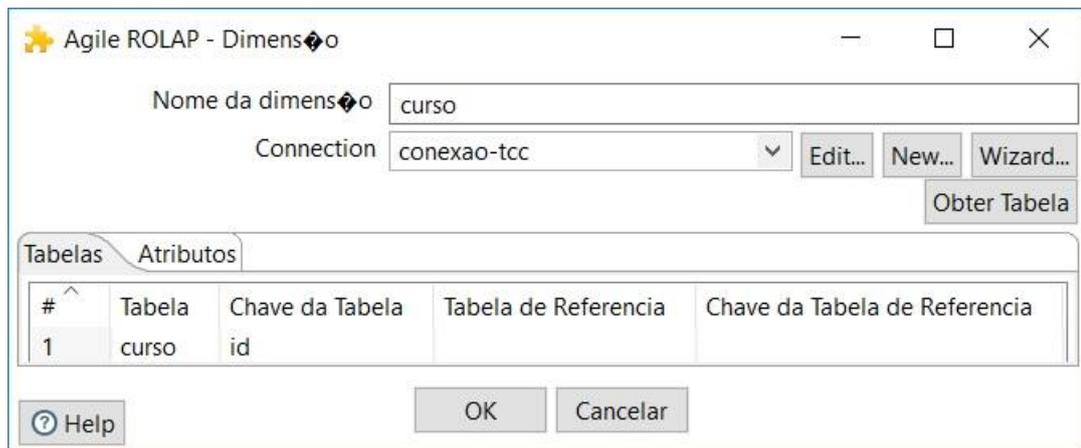
Fonte: Própria do autor.

**Figura 66:** Alguns dos atributos selecionados para a dimensão tempo\_pagamento para o fato pagamento\_inscrito.



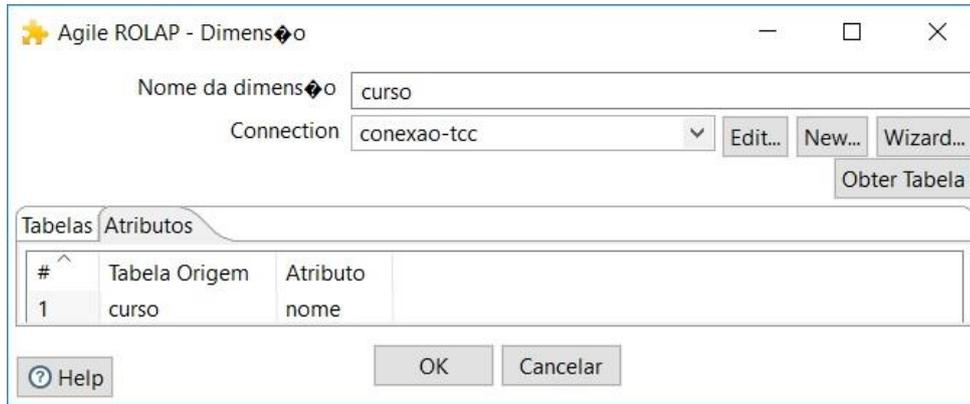
Fonte: Própria do autor.

**Figura 67:** Tabela curso selecionada para a dimensão curso para o fato pagamento\_inscrito.



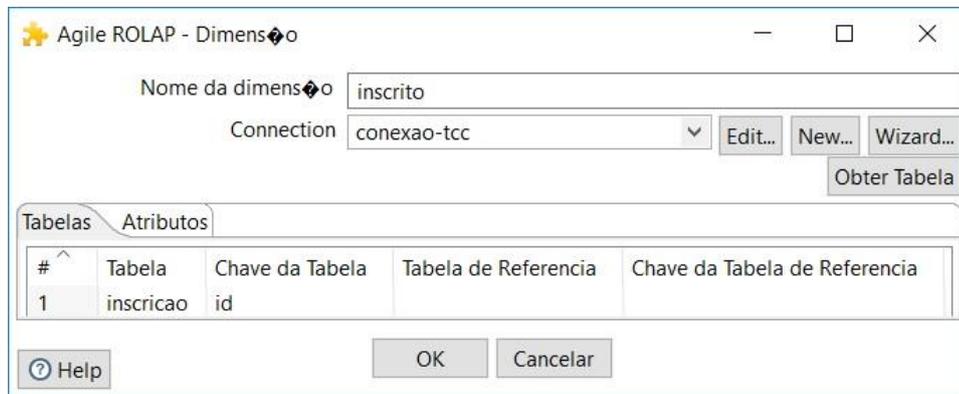
Fonte: Própria do autor.

**Figura 68:** Atributos selecionados para a dimensão curso para o fato pagamento\_inscrito.



Fonte: Própria do autor.

**Figura 69:** Tabela inscricao selecionada para a dimensão inscrito para o fato pagamento\_inscrito.



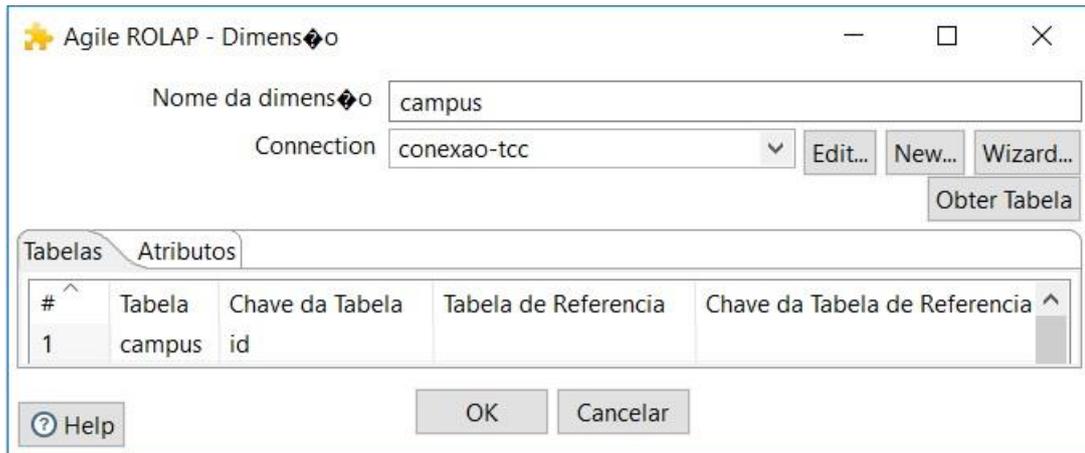
Fonte: Própria do autor.

**Figura 70:** Atributos selecionados para a dimensão inscrito para o fato pagamento\_inscrito.



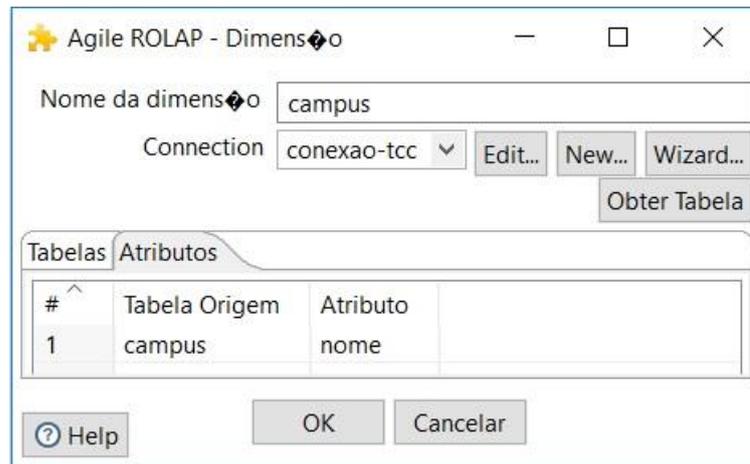
Fonte: Própria do autor.

**Figura 71:** Tabela campus selecionada para a dimensão campus para o fato pagamento\_inscrito.



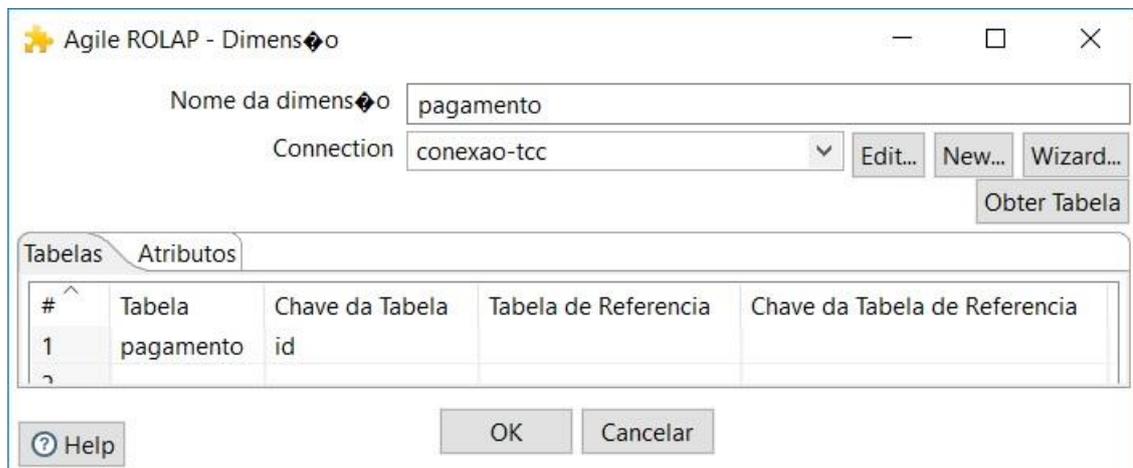
Fonte: Própria do autor.

**Figura 72:** Atributos selecionados para a dimensão campus para o fato pagamento\_inscrito.



Fonte: Própria do autor.

**Figura 73:** Tabela pagamento selecionada para a dimensão pagamento para o fato pagamento\_inscrito.



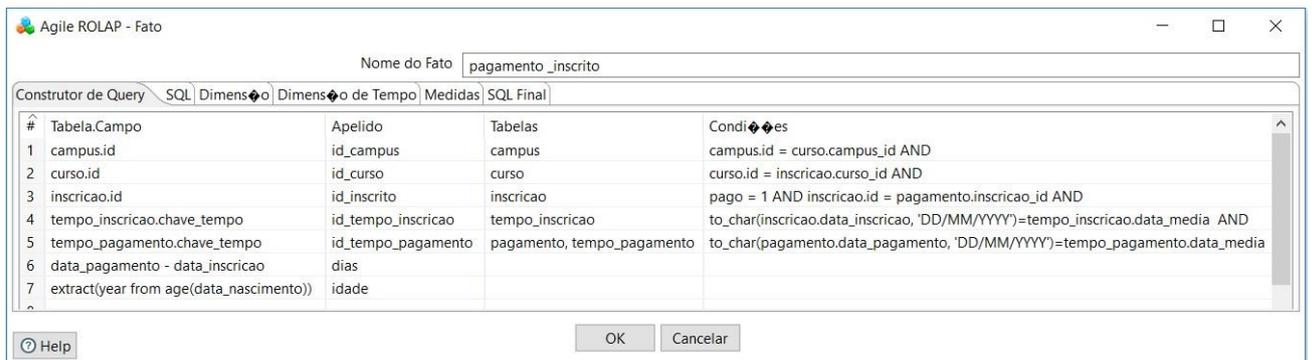
Fonte: Própria do autor.

**Figura 74:** Atributos selecionados para a dimensão pagamento para o fato pagamento\_inscrito.



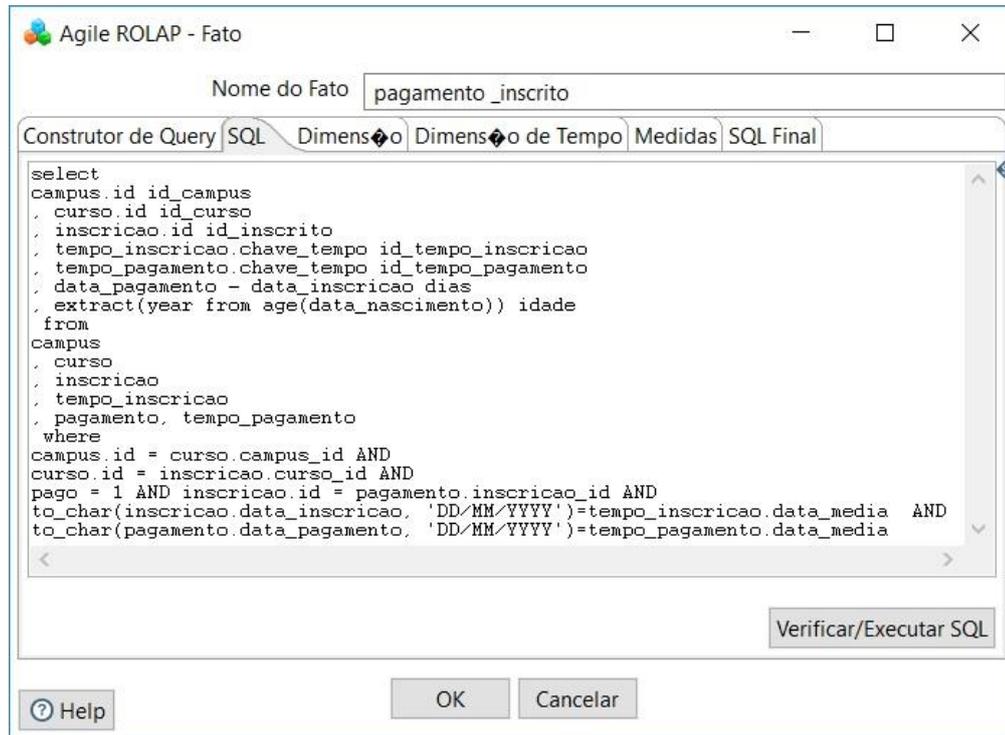
Fonte: Própria do autor.

**Figura 75:** Construtor de Query do fato pagamento\_inscrito.



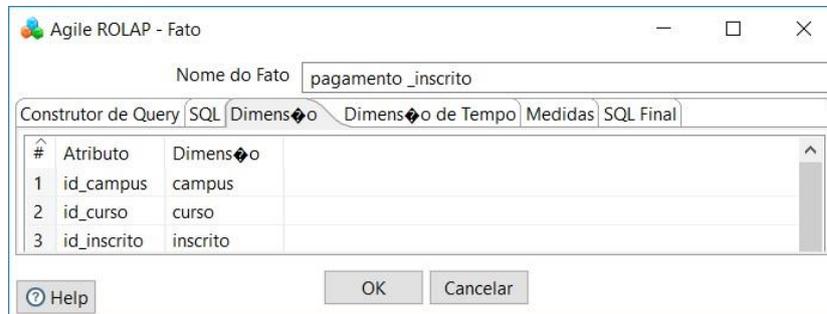
Fonte: Própria do autor.

**Figura 76:** SQL do fato pagamento\_inscrito para o fato pagamento\_inscrito.



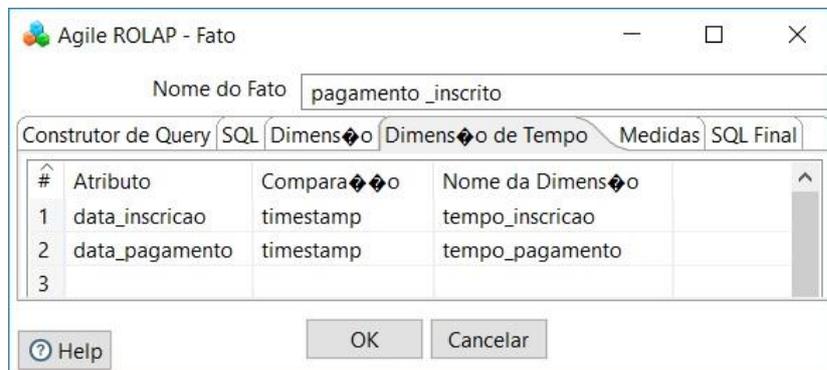
Fonte: Própria do autor.

**Figura 77:** Dimensões do fato pagamento\_inscrito.

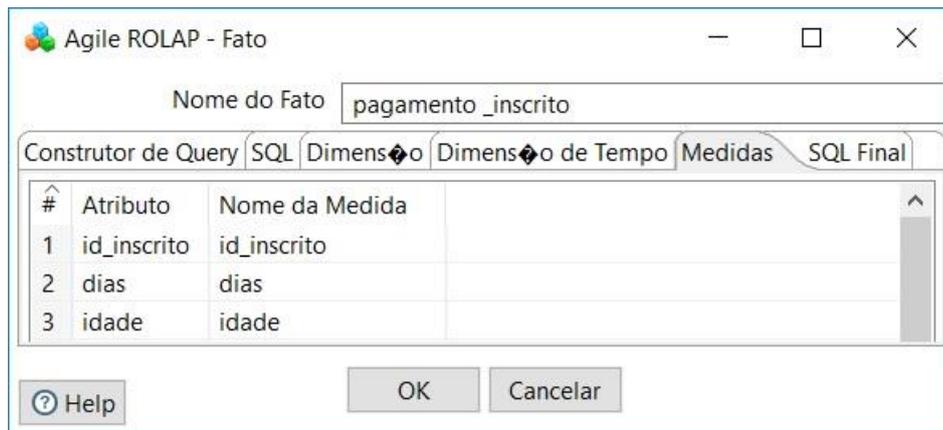


Fonte: Própria do autor.

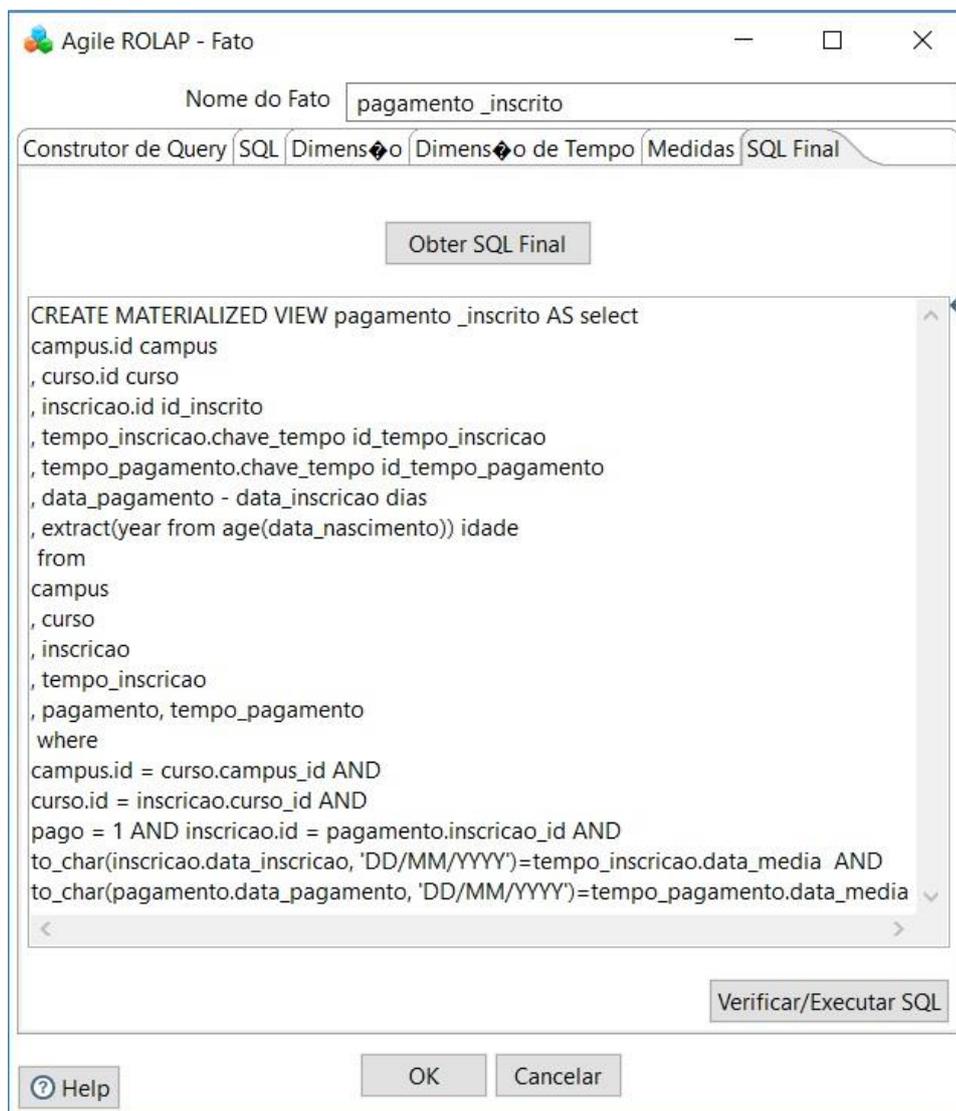
**Figura 78:** Dimensões tempo do fato pagamento\_inscrito.



Fonte: Própria do autor.

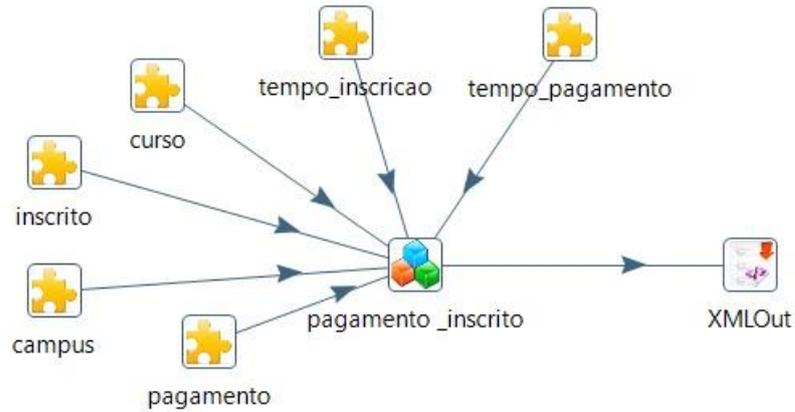
**Figura 79:** Medidas do fato pagamento\_inscrito.

Fonte: Própria do autor.

**Figura 80:** SQL da view gerado no plugin fato para o fato pagamento\_inscrito.

Fonte: Própria do autor.

**Figura 81:** Transformação do fato pagamento\_inscrito.



Fonte: Própria do autor.

**Figura 82:** Configuração da dimensão lógica curso para o cubo pagamento\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - CURSO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: curso

Nome da dimensão: curso

Visível:

Hierarquia	Nível
curso	nome_curso

Hierarquia

Nome: curso

Visível:

Nível

Nome: nome\_curso

Coluna: curso - nome

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 83:** Configuração da dimensão campus para o cubo pagamento\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - CAMPUS

Carregar de: C:\User\Desktop\xmlfinal.xml Navegar

Dimensão

Selecione dimensão: campus

Nome da dimensão: campus

Visível:

Hierarquia

Hierarquia	Nível
nome_campus	nome_campus

Hierarquia

Nome: nome\_campus

Visível:

Nível

Nome: nome\_campus

Coluna: campus - nome

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 84:** Configuração da dimensão região para o cubo pagamento\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - REGIAO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: regioao

Visível:

Hierarquia	Nível
cidade	cidade
estado	

Hierarquia

Nome: cidade

Visível:

Nível

Nome: cidade

Coluna: inscricao - cidade

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 85:** Configuração da dimensão sexo para o cubo pagamento\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - SEXO

Carregar de: C:\User\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: sexo

Visível:

Hierarquia	Nível
sexo	sexo

Hierarquia

Nome: sexo

Visível:

Nível

Nome: sexo

Coluna: inscricao - sexo

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 86:** Configuração da dimensão cotas para o fato pagamento\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - COTAS

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: cotista

Visível:

Hierarquia	Nível
cotista_social	cotista_social
cotista_sociorracial	

Hierarquia

Nome: cotista\_social

Visível:

Nível

Nome: cotista\_social

Coluna: inscricao - cota\_socia

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 87:** Configuração do cubo pagamento\_inscrito.

Nome do Passo: Cubo

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

**Cubo** Medidas

Cubo

Nome do Cubo: pagamento\_inscrito

Tabela: pagamento\_inscrito

Visível:

Uso de Dimensão

#	Nome	Fonte
1	curso	curso
2	campus	campus
3	regiao	regiao
4	sexo	sexo
5	cotista	cotista

Help OK Cancelar

Fonte: Própria do autor.

**Figura 88:** Configuração das medidas do cubo pagamento\_inscrito.

Nome do Passo: Cubo

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

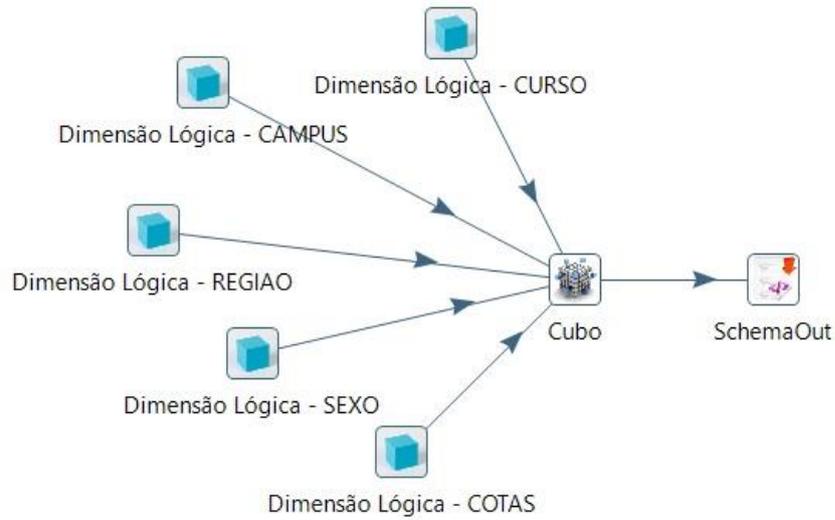
**Cubo** Medidas

#	Nome	Agregador	Coluna	Visível
1	qtd_ins...	count	id_inscrito	true
2	dias	avg	dias	true
3	idade	avg	idade	true

Help OK Cancelar

Fonte: Própria do autor.

**Figura 89:** Transformação do cubo pagamento\_inscrito.



Fonte: Própria do autor.

**Figura 90:** Tabela curso selecionada para a dimensão curso para o cubo resultado\_inscrito.

Nome da dimensão: curso  
 Connection: conexao-tcc

#	Tabela	Chave da Tabela	Tabela de Referencia	Chave da Tabela de Referencia
1	curso	id		

Fonte: Própria do autor.

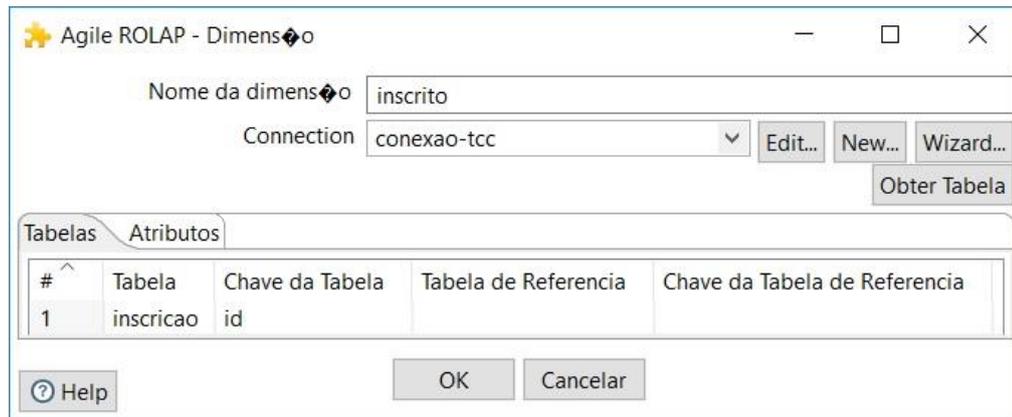
**Figura 91:** Atributos selecionados para a dimensão curso para o cubo resultado\_inscrito.

Nome da dimensão: curso  
 Connection: conexao-tcc

#	Tabela Origem	Atributo
1	curso	nome

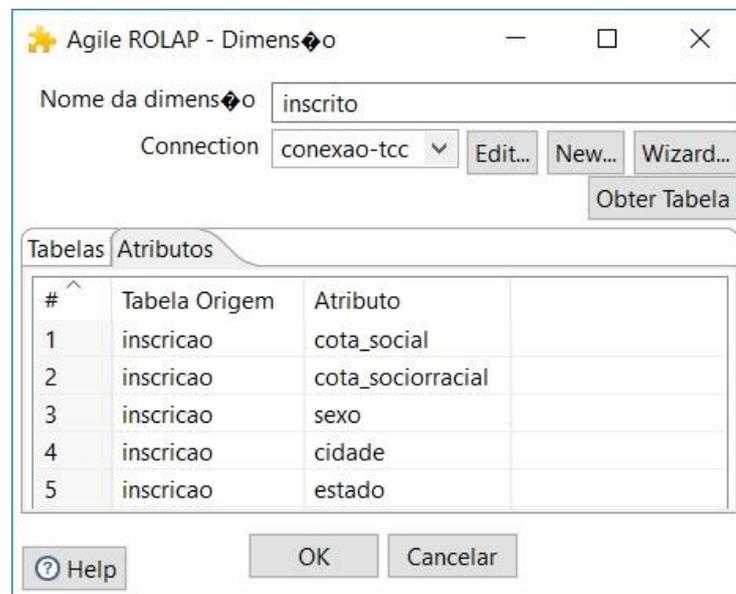
Fonte: Própria do autor.

**Figura 92:** Tabela inscricao selecionada para a dimensão inscrito para o cubo resultado\_inscrito.



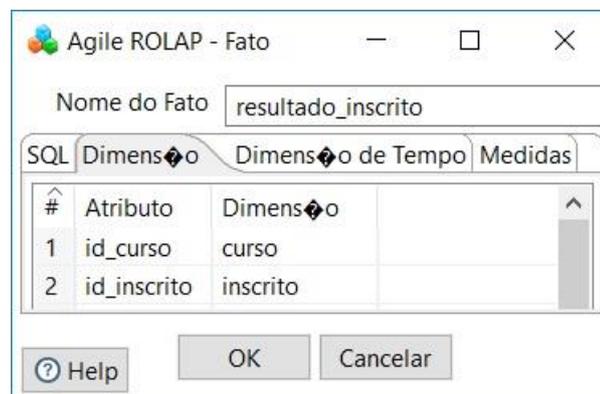
Fonte: Própria do autor.

**Figura 93:** Atributos selecionados para a dimensão inscrito para o cubo resultado\_inscrito.



Fonte: Própria do autor.

**Figura 94:** Dimensões do fato resultado\_inscrito para o cubo resultado\_inscrito.



Fonte: Própria do autor.

**Figura 95:** Transformação do fato resultado\_inscrito para o cubo resultado\_inscrito.



Fonte: Própria do autor.

**Figura 96:** Configuração da dimensão lógica região para o cubo resultado\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - REGIAO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: regioao

Visível:

Hierarquia	Nível
nome_cidade	nome_cidade
nome_estado	

Hierarquia

Nome: nome\_cidade

Visível:

Nível

Nome: nome\_cidade

Coluna: inscricao - cidade

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 97:** Configuração da dimensão lógica curso para o cubo resultado\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - CURSO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: curso

Nome da dimensão: curso

Visível:

Hierarquia	Nível
curso	nome_curso

Hierarquia

Nome: curso

Visível:

Nível

Nome: nome\_curso

Coluna: curso - nome

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 98:** Configuração da dimensão lógica sexo para o cubo resultado\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - SEXO

Carregar de: C:\Users\Desktop\xmlfinal.xml Navegar

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: sexo

Visível:

Hierarquia	Nível
sexo	sexo

Hierarquia

Nome: sexo

Visível:

Nível

Nome: sexo

Coluna: inscricao - sexo

Tipo: String

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 99:** Configuração da dimensão lógica cota para o cubo resultado\_inscrito.

Dimensão Lógica

Nome do passo: Dimensão Lógica - COTA

Carregar de: C:\Users\Desktop\xmlfinal.xml

Dimensão

Selecionar dimensão: inscrito

Nome da dimensão: cotista

Visível:

Hierarquia	Nível
cotista_social	cotista_sociorracial
cotista_sociorracial	

Hierarquia

Nome: cotista\_sociorracial

Visível:

Nível

Nome: cotista\_sociorracial

Coluna: inscricao - cota\_socik

Tipo: Integer

Visível:

Cima Baixo Cima Baixo

Nova Hierarquia Novo Nível

Deletar Hierarquia Deletar Nível

Help Ok Cancelar

Fonte: Própria do autor.

**Figura 100:** Configuração do cubo resultado\_inscrito.

Nome do Passo:

Carregar de:

**Cubo** Medidas

Cubo

Nome do Cubo:

Tabela:

Visível:

Uso de Dimensão

#	Nome	Fonte
1	curso	curso
2	regiao	regiao
3	sexo	sexo
4	cotista	cotista
5		

Fonte: Própria do autor.

**Figura 101:** Configuração das medidas do cubo resultado\_inscrito.

Nome do Passo:

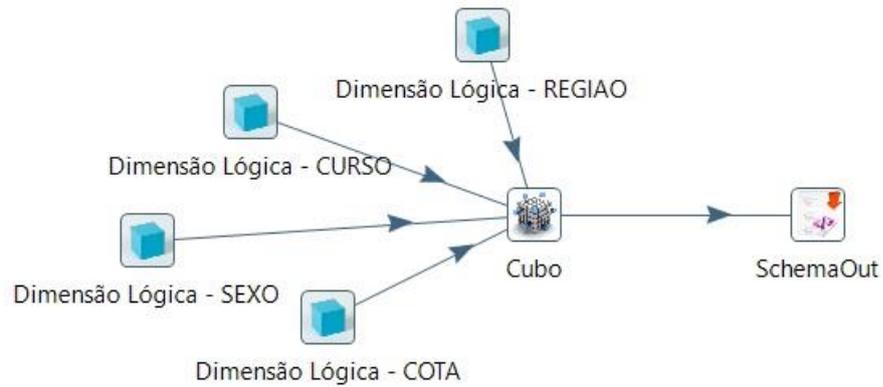
Carregar de:

**Cubo** Medidas

#	Nome	Agregador	Coluna	Visível
1	qtd_rea_prova	count	id_inscrito	true
2	pontuacao	avg	pontuacao	true
3	classificado_universal	sum	classificado_universal	true
4	espera_universal	sum	espera_universal	true
5	desclassificado_universal	sum	desclassificado_universal	true
6	classificado_social	sum	classificado_social	true
7	espera_social	sum	espera_social	true
8	desclassificado_social	sum	desclassificado_social	true
9	classificado_sociorracial	sum	classificado_sociorracial	true
10	espera_sociorracial	sum	espera_sociorracial	true
11	desclassificado_sociorracial	sum	desclassificado_sociorracial	true
12				

Fonte: Própria do autor.

**Figura 102:** Transformação do cubo resultado\_inscrito.



Fonte: Própria do autor.

**Figura 103:** Configuração dimensão lógica de tempo do cubo pagamento\_inscrito.

A imagem mostra a interface de um software de modelagem de dados. No lado esquerdo, há uma árvore de navegação com o seguinte conteúdo:

- Schema
  - pagamento\_inscrito
    - curso
    - regiao
    - campus
    - sexo
    - tempo\_inscricao
      - default
        - ano
        - mes
        - dia
        - Table: tempo\_inscricao
      - tempo\_pagamento
        - tempo\_pagamento
          - ano
          - mes
          - dia
          - Table: tempo\_pagamento

No lado direito, há uma aba intitulada 'Shared Dimension' com uma tabela de atributos:

Attribute	Value
name	tempo_inscricao
description	
foreignKey	
type	TimeDimension
usagePrefix	
caption	
visible	<input checked="" type="checkbox"/>

Na barra inferior, está escrito 'Database - tcc (PostgreSQL)'.

Fonte: Própria do autor.