

UNIVERSIDADE FEDERAL DO PARANÁ

AMANDA CRISTINA FLORIANO  
EDUARDO FARAGO LEMES  
VICTOR HUGO GANDIN HEOFACKER

ANÁLISE DE FERRAMENTAS DE BUSINESS INTELLIGENCE

CURITIBA

2016

AMANDA CRISTINA FLORIANO  
EDUARDO FARAGO LEMES  
VICTOR HUGO GANDIN HEOFACKER

## ANÁLISE DE FERRAMENTAS DE BUSINESS INTELLIGENCE

Trabalho de Conclusão de Curso apresentado ao Curso de Tecnologia em Análise e Desenvolvimento de Sistemas da Universidade Federal do Paraná como requisito à obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas

Orientadora: Profa. Msc. Andreia de Jesus

CURITIBA

2016

## SUMÁRIO

LISTA DE FIGURAS .....	5
RESUMO .....	9
ABSTRACT .....	10
1. INTRODUÇÃO .....	11
1.1. PROBLEMA .....	15
1.2 OBJETIVOS DO TRABALHO .....	16
1.2.1 OBJETIVO GERAL .....	16
1.2.2 OBJETIVOS ESPECÍFICOS .....	16
1.3 JUSTIFICATIVA .....	16
2. REFERENCIAL TEÓRICO .....	18
2.1 BREVE HISTÓRIA DO CONCEITO DE BUSINESS INTELLIGENCE (BI) ..	18
2.2 OS NÍVEIS ORGANIZACIONAIS .....	21
2.3 OS NÍVEIS ORGANIZACIONAIS E O BI .....	24
2.4 SISTEMAS OLAP X SISTEMAS OLTP .....	26
2.5 CONCEITUANDO TECNOLOGIAS DE BI.....	28
2.5.1 ETL .....	29
2.5.2 STAGING AREA .....	31
2.5.3 DATA WAREHOUSE .....	31
2.5.4 DATAMART .....	33
2.6 O MODELO DE DADOS DIMENSIONAL .....	34
2.6.1 VANTAGENS DA MODELAGEM DIMENSIONAL .....	39
2.6.2 CRIANDO A ESTRUTURA DE MODELOS DIMENSIONAIS.....	40
3. FERRAMENTAS DE BI.....	46
3.1 A FERRAMENTA <i>QLIKVIEW</i> .....	49
3.1.1 DESCRIÇÃO E CARACTERÍSTICAS .....	49
3.1.2 INVESTIMENTO .....	56
3.2 A FERRAMENTA <i>MICROSTRATEGY</i> .....	57

3.2.1	DESCRIÇÃO E CARACTERÍSTICAS .....	58
3.2.2	INVESTIMENTO .....	66
3.3	A FERRAMENTA <i>MICROSOFT POWER BI</i> .....	67
3.3.1	DESCRIÇÃO E CARACTERÍSTICAS .....	67
3.3.2	INVESTIMENTO .....	72
3.4	A FERRAMENTA TABLEAU.....	73
3.4.1	DESCRIÇÃO E CARACTERÍSTICAS .....	73
3.4.2	INVESTIMENTO .....	78
4	METODOLOGIA DO TRABALHO .....	79
4.1	CRITÉRIOS DE AVALIAÇÃO E ANÁLISE.....	79
4.2	DADOS E FONTES UTILIZADAS .....	81
4.3	MODELAGEM E CARREGAMENTOS DOS DADOS.....	82
4.4	ESPECIFICAÇÃO DOS RELATÓRIOS .....	88
4.5	ESPECIFICAÇÃO DO TESTE DE <i>DRILL-DOWN</i> .....	109
4.5.1	MICROSTRATEGY .....	109
4.5.2	QLIKVIEW .....	113
4.5.3	POWER BI .....	116
4.5.4	TABLEAU .....	118
5.	ANÁLISE DOS RESULTADOS .....	125
6.	CONCLUSÃO .....	133
	REFERÊNCIAS .....	134
	APÊNDICE 1 – Cronograma de trabalho .....	139

## LISTA DE FIGURAS

Figura 1 : Processos de BI. ....	21
Figura 2 : Níveis Organizacionais .....	22
Figura 3 : Processo de BI (Business Intelligence).....	29
Figura 4 : Processo de ETL.....	30
Figura 5 : Exemplo de tabela fato e dimensão.....	37
Figura 6 : Exemplo de tabela fato e dimensão.....	38
Figura 7 : Exemplo de dados analíticos vs sumarizados .....	41
Figura 8 : Representação do modelo Star-Schema .....	42
Figura 9 : Representação Snowflake Schema .....	44
Figura 10 : Quadrante mágico Gartner.....	47
Figura 11: Melhores ferramentas de BI 2014.....	48
Figura 12 : Tela inicial do QlikView. ....	51
Figura 13 : Criação de documentos no QlikView.....	52
Figura 14 : Edição de script no QlikView. ....	53
Figura 15 : Edição de dashboard no QlikView. ....	53
Figura 16 : Tela inicial do Microstrategy Analytics Desktop.....	60
Figura 17 : Amostra de dashboard no Microstrategy. ....	61
Figura 18 : Processo de criação de dashboard no Microstrategy Analytics Desktop – seleção de fonte de dados. ....	62
Figura 19 : Utilizando conexões SGBD no Microstrategy Analytics Desktop. ....	63
Figura 20 : Criação de dashboards no Microstrategy Analytics Desktop. ....	63
Figura 21 : Tela inicial do Power BI Desktop. ....	68
Figura 22 : Tela de seleção de dados no Power BI. ....	68
Figura 23 : Carregando dados do PostgreSQL no Power BI utilizando código em SQL. ....	69
Figura 24 : Transformação de dados utilizando o Query Editor do Power BI.....	70
Figura 25 : Criação de gráficos no Power BI Desktop.....	70
Figura 26 : Visualização de gráfico publicado no Power BI online. ....	71
Figura 27 : Tela inicial do Tableau Desktop. ....	75
Figura 28 : Conexão com o PostgreSQL no Tableau Desktop. ....	76

Figura 29 : Extração de dados no Tableau Desktop .....	77
Figura 30 : Construção de gráficos no Tableau Desktop. ....	77
Figura 31 : DER – base de dados Censo Ensino Superior 2013.....	84
Figura 32 : Criação de um novo banco de dados no pgAdmin. ....	85
Figura 33 : Função COPY do PostgreSQL.....	86
Figura 34 : Criação de uma nova conexão ODBC. ....	87
Figura 35 : Configuração da conexão ODBC.....	87
Figura 36 : Relatório 1 no Microstrategy. ....	89
Figura 37 : Relatório 1 no Qlikview. ....	89
Figura 38 : Relatório 1 no Power BI.....	90
Figura 39 : Relatório 1 no Tableau.....	90
Figura 40 : Relatório 2 no Microstrategy. ....	91
Figura 41 : Relatório 2 no QlikView. ....	92
Figura 42 : Relatório 2 no Power BI. ....	92
Figura 43 : Relatório 2 no Tableau.....	93
Figura 44 : Relatório 3 no Microstrategy. ....	94
Figura 45 : Relatório 3 no QlikView. ....	94
Figura 46 : Relatório 3 no Power BI. ....	95
Figura 47 : Relatório 3 no Tableau.....	95
Figura 48 : Relatório 4 no Microstrategy. ....	96
Figura 49 : Relatório 4 no QlikView. ....	97
Figura 50 : Relatório 4 no Power BI. ....	97
Figura 51 : Relatório 4 no Power BI.....	98
Figura 52 : Relatório 5 no Microstrategy. ....	99
Figura 53 : Relatório 5 no QlikView. ....	100
Figura 54 : Relatório 5 no Power BI. ....	100
Figura 55 : Relatório 5 no Tableau.....	101
Figura 56 : Relatório 6 no Microstrategy .....	102
Figura 57 : Relatório 6 no QlikView. ....	102
Figura 58 : Relatório 6 no Power BI. ....	103
Figura 59 : Relatório 6 no Tableau.....	103
Figura 60 : Relatório 7 no Microstrategy.....	104
Figura 61 : Relatório 7 no QlikView. ....	105
Figura 62 : Relatório 7 no Power BI. ....	105

Figura 63 : Relatório 7 no Tableau. ....	106
Figura 64 : Relatório 8 no Microstrategy. ....	107
Figura 65 : Relatório 8 no QlikView. ....	108
Figura 66 : Relatório 8 no QlikView. ....	108
Figura 67 : Relatório 8 no Tableau.....	109
Figura 68 : Drill-down no Microstrategy para detalhamento da região sul. ....	110
Figura 69 : Drill-down no Microstrategy. Detalhamento por estados da região sul. ....	111
Figura 70 : Passo 3 de teste de drill-down no Microstrategy. Detalhamento por cidades do estado do Paraná. ....	111
Figura 71 : Drill-down no Microstrategy. Detalhamento por instituições de ensino de Curitiba. ....	112
Figura 72 : Drill-down no Microstrategy mostrando dados específicos da Universidade Federal do Paraná.....	112
Figura 73 : Drill-down no QlikView. ....	113
Figura 74 : Drill-down no Qlikview como sendo uma dimensão do gráfico de barras. ....	114
Figura 75 : Gráfico inicial para teste de drill-down no QlikView. ....	114
Figura 76 : Drill-down no QlikView. Detalhamento por estados da região Sul do Brasil. ....	115
Figura 77 : Drill-down no QlikView. Detalhamento por cidade do estado do Paraná. ....	115
Figura 78 : Drill-down no QlikView. Detalhamento por instituições de ensino em Curitiba. ....	115
Figura 79 : Drill-down no QlikView com dados específicos da UFPR.....	116
Figura 80 : Relatório inicial para teste de drill-down no Power BI. ....	117
Figura 81 : Drill-down no Power BI. Detalhamento por cidades do estado do Paraná. ....	117
Figura 82 : Drill-down no Power BI. Detalhamento por instituições de ensino de Curitiba. ....	118
Figura 83 : Drill-down no Power BI com dados específicos da Universidade Federal do Paraná ....	118
Figura 84 : Drill-down no Tableau com a criação da hierarquia Região – Estado – Cidade - universidade. ....	119
Figura 85 : Drill-Down no Tableau com a inclusão do município na hierarquia. ....	119

Figura 86 : Drill-down no Tableau com a inclusão do nome da instituição acadêmica na hierarquia. ....	120
Figura 87: Inclusão do campo hierárquico no campo coluna do gráfico. ....	120
Figura 88 : Drill-down no Tableau filtrando a região Sul. ....	121
Figura 89 : Drill-down no Tableau . Expandindo a hierarquia região – estado.....	121
Figura 90 : Drill-Down no Tableau filtrando o estado do Paraná. ....	122
Figura 91 : Drill-down no Tableau. Expandindo a hierarquia estado – município e filtrando Curitiba. ....	122
Figura 92 : Drill-down no Tableau. Expandindo a hierarquia cidade – universidade e filtrando a Universidade Federal do Paraná. ....	123
Figura 93 : Drill-down do Tableau. ....	123
Figura 94 : Drill-down do Tableau – nível UFPR. ....	124

## RESUMO

A utilização de ferramentas de Business Intelligence (BI) é considerada fundamental para a vantagem competitiva desejada pelas empresas. Escolher qual a ferramenta é a mais adequada continua sendo um desafio. Este trabalho propõe considerar os níveis organizacionais durante a avaliação das ferramentas. Vinte um critérios são empregados considerando um estudo de caso com dados do censo de ensino superior de 2013. São analisadas quatro ferramentas de BI (QlikView, Microstrategy, Power BI e Tableau), escolhidas por serem consideradas líderes de mercado na pesquisa do Gartner Group de 2015. Os resultados possibilitam comparar as ferramentas considerando os níveis organizacionais, destacando algumas diferenças relevantes para o sucesso da implantação de um projeto de BI.

**Palavras-chave:** Ferramentas *Business Intelligence*; Análise de Ferramentas de BI; Níveis organizacionais.

## **ABSTRACT**

The usage of Business Intelligence (BI) tools is considered essential by organizations in order to obtain the desirable competitive advantage. The process of choosing the most appropriate tool is still a challenge. This paper is proposing to consider the organizational level while evaluating the tools. Twenty one criteria have been evaluated considering a case study based on a census of higher education in the year of 2013. Four BI tools (QlikView, Microstrategy, Power BI and Tableau) have been analysed, which have been chosen as being the market leaders according to Gartner Group 2015 research. The results allows the comparison between the tools considering the organizational levels and highlighting some differences which are relevant for the success in implementing a BI project.

**Key words:** Business Intelligence tools, BI tools analysis, BI organizational levels.



## 1. INTRODUÇÃO

É vital que as empresas estejam à frente de seus concorrentes em diversos momentos, como na hora de definir o mercado de atuação, desenhar estratégias de marketing ou até mesmo lançar um produto no mercado. Os líderes das empresas precisam, no entanto, estar munidos com informação e conhecimento apropriado para tomarem suas decisões. A tecnologia de business intelligence vem para auxiliar os administradores à obter a informação de forma organizada e coerente, provendo os recursos necessários para que esta tomada de decisão seja assertiva e rápida. Cada nível organizacional vê a informação de uma forma, e esta visão precisa ser considerada ao disponibilizarmos recursos para estes profissionais. Este trabalho propõe a convergência entre a necessidade dos administradores e os recursos de tecnologia da informação, para que, no final, a empresa obtenha o melhor aproveitamento de seus dados com uma tecnologia condizente com as suas expectativas.

A informação e o conhecimento começaram a ser encarados como recursos fundamentais dentro das empresas, que passaram a vivenciar novos problemas e perspectivas, tanto no âmbito produtivo quanto administrativo (ENANCIB, 2005).

McGee e Prusak (1994, p. 3) afirmam que “a concorrência entre as organizações baseia-se em sua capacidade de adquirir, tratar, interpretar e utilizar a informação de forma eficaz.”. Braga (2009) também ressalta que as empresas têm interesse em buscar tecnologias cada vez melhores, uma vez que o volume de informações a ser analisada é alto e o tempo é curto, surge a necessidade de um investimento tecnológico que as coloquem em vantagem perante seus concorrentes de mercado.

Utilizar a informação de forma organizada e confiável é uma maneira estratégica da empresa conseguir se destacar ante seus concorrentes, e também, evidencia as ações que precisam ser tomadas para obter os resultados desejados, bem como para definir as estratégias de mercado mais adequadas. Siqueira (2005) mostra que a análise da informação é um passo importante no processo de tomada de decisões, pois, entre outros benefícios proporcionados pela análise de informações, a empresa consegue identificar oportunidades, ameaças, problemas tendências e definir o público alvo para sua operação.

Rascão (2006) nos mostra que todas as decisões estratégicas são tomadas pelos gestores das empresas, independente do procedimento pelo qual foram motivadas. Não há dúvidas de que a análise de informações gerenciais é o pilar central para a tomada das decisões estratégicas destes gestores.

As decisões estratégicas estão relacionadas tanto com o ambiente interno (recursos humanos, recursos financeiros, recursos de logística, recursos de tecnologia, ...) quanto externos (produtos oferecidos por concorrentes, procura do produto no mercado, economia do país, perfil de clientes,...). As estratégias precisam estar alinhadas com os serviços e recursos disponíveis na empresa, para que seja possível atingir ou criar uma meta da melhor forma possível.

A capacidade de elaborar estratégias de qualidade é maior quando a empresa dispõe de informação. A análise da informação é crucial para qualquer tomada de decisão e permite que a empresa possa buscar seus objetivos de forma mais segura. Para Mcgee e Prusak (1994), a informação é um recurso tão importante quanto o capital de investimento, mão de obra ou tecnologia.

Braga (2009) nos expõe que as tecnologias da informação proporcionam inovações tecnológicas nas empresas, que por sua vez têm seu patrimônio valorizado, gerando novos investimentos e com isso a empresa consegue ter mais capital para oferecer serviços de qualidade aos seus clientes. Um sistema de informação estratégico deve ser capaz de gerar artefatos para que os gestores das empresas possam avaliar as informações e tomar decisões sobre os objetivos, as operações, os produtos, os serviços, os relacionamentos internos entre os setores, enfim, tudo que depende de acompanhamento de resultados, podendo confirmar os bons resultados de uma ação planejada ou até mudar o foco de determinada estratégia dos negócios.

Seguindo este raciocínio, Braga (2009) também expõe uma relação entre as Tecnologias da Informação (TI) e os Sistemas de Informação (SI): no nível dos SI são definidas as necessidades de informação e sua aplicação no negócio, baseadas numa análise da organização e de sua estratégia global. No nível de TI é definido qual a sua contribuição para o processamento de informação e para a satisfação das necessidades informacionais e aplicacionais, bem como o desenvolvimento de sistemas e criação de vantagens competitivas para a empresa, tendo em conta as prioridades fixadas na estratégia dos SI.

Assim, deve-se sempre levar em consideração, no momento de formular a

estratégia, quais os caminhos que uma empresa seguirá para manter-se competitiva, a relação entre as Tecnologias e os Sistemas de Informação. Além disso, é preciso avaliar o conteúdo informacional e certificar-se da confiabilidade das fontes que alimentam esses sistemas de informação.

Ainda segundo Braga (2007), a empresa ao atuar num mundo global está em estado permanente de “necessidade de informação”, uma vez que a informação constitui o suporte de uma organização e é um elemento essencial e indispensável à sua existência. Ele afirma que uma empresa não funciona sem informação, porém é importante saber utilizar esse recurso.

É clara a necessidade de termos um processo confiável para gerenciar a informação dentro das empresas. Neste contexto são inseridas as ferramentas de *Business Intelligence* (BI), que são definidas como:

“As ferramentas de Business Intelligence otimizam o tempo investido em coleta e análise de dados para mensurar resultados e definir estratégias de negócios. Basicamente, essas ferramentas possibilitam a unificação de diferentes informações, mesmo que estejam divididas em diversos tipos de plataformas.” (Inquesti, 2015, Disponível [inquesti.com.br/blog/conheca-as-principais-ferramentas-de-business-intelligence](http://inquesti.com.br/blog/conheca-as-principais-ferramentas-de-business-intelligence) Acesso 20/06/2016)

Há tempos atrás o processo de BI era tratado como um luxo e somente era utilizado em grandes empresas, devido o custo elevado e dificuldade na implementação. Hoje, com o grande volume de dados gerados pelas empresas e diminuição do custo do hardware para armazenamento de informações, o uso do BI passou a se tornar uma necessidade, e está se tornando comum também em empresas de médio e pequeno porte. Conforme disse Maurizio Niccolai, gerente de Marketing e Soluções da HP Brasil, em 2007: “As necessidades de BI hoje são as mesmas nas pequenas, médias e grandes empresas. A diferença está no volume de dados” (PCWORLD, 2007, p. 5).

Decidir corretamente qual rumo tomar à cerca de determinado desafio torna-se essencial, pois cada detalhe pode definir o lucro e o prejuízo da empresa. Decisões erradas, sejam estratégicas, táticas ou operacionais, podem custar o futuro da empresa, assim como uma correta, definir seu sucesso. Por isso, antes de qualquer decisão, o ideal é ter à disposição a maior quantidade possível de fatos e

informações relevantes para minimizar o risco de erro. Também é importante ter ferramentas aptas a fornecer meios de uma decisão mais adequada.

Com o barateamento do hardware, a capacidade de coletar e armazenar dados superou em muito a habilidade de analisar, resumir e extrair "conhecimentos" destes dados. Os dados estão em toda a parte, entretanto é preciso transformá-los em informações, e estas em conhecimento, de uma forma organizada e gerenciável. A maioria das organizações não sofre com a falta de dados, mas sim com uma abundância de dados redundantes e inconsistentes, complexos de administrar, cada vez mais árduos de acessar e difíceis de usar para fins de suporte à tomada de decisão. Além disso, a necessidade de eficiência e agilidade do processo decisório nas instituições exige dela a utilização de soluções que gerem informações consistentes no tempo esperado.

É com a intenção de solucionar estes problemas que as empresas procuram cada vez mais as ferramentas de BI oferecidas no mercado de Tecnologia. Porém, há a uma oferta variada dessas ferramentas, o que dificulta identificar qual delas é a mais adequada para um determinado projeto. E este problema vem se tornando tão relevante no meio empresarial que já encontramos na literatura trabalhos que abordam a comparação entre ferramentas de BI. Como o estudo realizado em 2011 na PUC-RJ, com o título "Uma Análise Comparativa entre as Ferramentas OLAP como Apoio a Soluções de BI nas Empresas" (Sá, Thays et al, 2011), o qual compara as diversas ferramentas considerando critérios como desempenho, arquitetura, suporte técnico. Porém, as análises dos critérios de avaliação podem ser consideradas superficiais, pois neste estudo não é apresentado, na conclusão, uma análise qualitativa das ferramentas, somente é apresentado o resultado "Bom, Ruim, Muito Bom, Regular ou Não se Aplica" com uma descrição genérica para cada opção de resultado. Além disso, os autores só geraram relatórios e avaliaram mais profundamente os softwares *Qlikview* e *Microstrategy*.

Logo, a fim de alcançar uma análise mais acurada, propõem-se para este trabalho que os critérios sejam expandidos e analisados mais profundamente nas ferramentas selecionadas para o estudo.

Além do mais, seria interessante buscar a sinergia entre o ambiente de TI e o ambiente de negócios, avaliando critérios técnicos do sistema sob a visão negocial das informações. Sob o ponto de vista de negócios, serão utilizados os resultados das análises técnicas para identificar a ferramenta mais adequada para cada nível

da empresa (operacional, gerencial ou estratégico).

## 1.1. PROBLEMA

Com base nos estudos da empresa *Gartner Group*, líder mundial em consultoria e pesquisas na área de Tecnologia da Informação, temos hoje, pelo menos, 27 ferramentas de BI disponíveis no mercado para utilização, sejam elas proprietárias ou livres. Esta empresa avalia anualmente as principais ferramentas de BI, considerando os seguintes critérios: a capacidade de a ferramenta executar aquilo que foi proposto e a visão da empresa em termos de inovação tecnológica.

Esta diversidade de opções acaba gerando dificuldades nas empresas em escolher a ferramenta certa para seu ramo de atuação. As ferramentas, apesar de terem o propósito comum de gerar relatórios gerenciais, possuem características distintas de usabilidade, modo de apresentação e desempenho. Logo, negligenciar a escolha da ferramenta gera um impacto direto no resultado de todo o projeto, pois acarreta um custo elevado para toda a solução projetada e, em alguns casos, pode até optar pelo abandono da solução. Cabe ao profissional de TI analisar as ferramentas existentes e escolher a que melhor se aplica a realidade da empresa.

Logo, as seguintes questões surgem nesse contexto:

- Entre tantas ferramentas disponíveis, as empresas conseguem definir se uma determinada ferramenta de BI atende as suas necessidades?
- Os dados estarão distribuídos de uma forma sustentável?
- As ferramentas disponíveis no mercado são adequadas para atender projetos de BI em todos os níveis da empresa (estratégico tático e operacional)?

Tendo estas questões em mente, o presente trabalho visa analisar de forma sistemática ferramentas de BI, a fim de produzir um recurso indicativo de aplicação dessas ferramentas em contextos em que melhor se adéquem. Desta forma, profissionais de TI e analistas de negócios poderão consultar esta documentação e identificar qual das ferramentas avaliadas é a mais adequada para a solução do seu problema.

## 1.2 OBJETIVOS DO TRABALHO

### 1.2.1 OBJETIVO GERAL

Análise de ferramentas OLAP com o intuito de identificar em qual nível de uma empresa (operacional, gerencial ou estratégico) ela melhor se aplica para a análise de informações. Com isto, gerar como produto final um relatório técnico indicativo de aplicação das ferramentas avaliadas.

### 1.2.2 OBJETIVOS ESPECÍFICOS

- ✓ Selecionar e estudar as ferramentas OLAP escolhidas pela posição de liderança no quadrante do Gartner Group.
- ✓ Testar as ferramentas selecionadas com o mesmo estudo de caso.
- ✓ Elencar os critérios de avaliação das ferramentas.
- ✓ Avaliar as ferramentas com base nos critérios selecionados.
- ✓ Associar a ferramenta com o nível de aplicação que melhor se adequa a ela (operacional, gerencial ou estratégico)
- ✓ Elaborar um quadro indicativo sobre as características das ferramentas de BI estudadas.

## 1.3 JUSTIFICATIVA

Dados apresentados em um relatório de novembro de 2011 pela *RSR Research* mostram que a maioria dos varejistas acredita que o *Business Intelligence* (BI) é muito (62%) relevante para obter informações sobre o desempenho da companhia. Informações estas que os auxiliam no controle de processos internos e que possam ser utilizadas para comparar o desempenho atual com os objetivos definidos no plano de *marketing*. Outro estudo apresentado pela mesma empresa, “O Inteligente Mundo de Compreensão do Varejista” (*The Intelligent Retailer's World of Insight*), também mostra que a maioria dos entrevistados (57%) considera BI como uma ferramenta muito relevante para entender o comportamento do consumidor e executar os planos internos em prol da construção de uma relação de lealdade com seus clientes.

Mas hoje é possível encontrar cerca de 27 ferramentas de BI disponíveis para uso no mercado, como já mencionado. As empresas em geral sentem a necessidade de utilizar uma ferramenta de BI, porém não têm as informações

necessárias para decidir qual dentre tantas é a ideal para seu nicho de mercado e sua necessidade empresarial.

As empresas consideram o BI importante para os resultados como um todo, mas se sentem, na maioria das vezes, indecisas na escolha da ferramenta de BI a ser aplicada, pela falta de um material sistematizado que indique não só os recursos, mas o nível de aplicação dessas ferramentas. E esta decisão é importante, pois o sucesso da solução está diretamente associado à escolha da ferramenta que deverá ser implantada na empresa para atender ao projeto de BI, bem como, em qual nível da empresa (operacional, gerencial ou estratégico) a solução está sendo implementada.

Com este cenário, é possível verificar que o sucesso de qualquer projeto de BI está diretamente relacionado a uma escolha assertiva da ferramenta de software destinada à solução, bem como o tipo de informação a ser manipulada. Por isso, o presente estudo tem como objetivo analisar as ferramentas *Qlikview*, *Microstrategy*, *Microsoft Power BI* e *Tableau* disponíveis no mercado hoje, que tem o objetivo de auxiliar analistas de TI, responsáveis por projetos de BI, a escolher de forma adequada e fundamentada a ferramenta de BI para a execução do seu projeto.

## 2. REFERENCIAL TEÓRICO

### 2.1 BREVE HISTÓRIA DO CONCEITO DE BUSINESS INTELLIGENCE (BI)

O termo *Business Intelligence* teve suas origens na década de 50 quando um pesquisador da IBM, H. P. Luhn, em seus estudos, definiu *Business Intelligence* como sendo: “a habilidade de aprender os Inter-relacionamentos dos fatos apresentados de uma forma a nos guiar às ações para se atingir um objetivo” (LUHN, 1958, p. 314).

Este pesquisador propõe o desenvolvimento de um sistema baseado em máquinas de processamento de dados, o qual faria a abstração e codificação de documentos criando perfis de interesse para cada ponto de ação na organização. Em seu artigo ele descreve a flexibilidade desse sistema em identificar a informação, encontrar quem precisa dessa informação e disseminá-la de maneira eficiente. Além disso, descreve que os métodos de comunicação da época eram inadequados para se atingir o objetivo das organizações, visto que o crescimento das empresas criam especializações e divisões que geram barreiras na disseminação efetiva da informação. Diante desta questão, o pesquisador (LUHN, 1958) propõe o desenvolvimento de um sistema de inteligência de negócios que aborda a coleta, armazenamento, transmissão e disseminação de novas informações.

O sistema proposto por Luhn prevê a entrada de dados por meio de microfilmagem de documentos que mais tarde seriam transcritos em fitas magnéticas por meio de transcritor humano ou dispositivo de transcrição automática. Esses registros em fita então poderiam ser utilizados para análises estatísticas baseados nas propriedades do texto. Os perfis de usuários são criados a partir da identificação de quem necessita saber e quais informações são necessárias. A partir daí são feitas comparações entre os textos analisados com os perfis de usuários armazenados para se fazer o constante monitoramento e disseminação das informações.

Foi a partir da década de 80 que os sistemas de *Business Intelligence* tiveram maior desenvolvimento devido o aumento da capacidade de armazenamento e processamento dos computadores da época. Foi nesse período que surgiram

disciplinas que abordavam conteúdos relacionados a dados, como: administração de dados, modelagem de dados e engenharia da informação (BARBIERI, 2011).

O termo BI, usado para referenciar conceitos e métodos que eram aplicados para auxiliar empresas em tomadas de decisão, foi primeiramente utilizado em 1989 por *Howner Dresner*, analista da *Gartner Group* (um dos maiores nomes de consultoria de BI da atualidade). Dresner define *Business Intelligence* como um termo guarda chuva para descrever conceitos e métodos para aprimorar a tomada de decisões gerenciais a partir de sistemas de suporte baseados em fatos (POWER, 2005).

Apesar do termo *Business Intelligence (BI)* já estar sendo utilizado desde a década de 50, ainda não há consenso entre os pesquisadores sobre o conceito de BI. É possível identificar termos e objetivos comuns nas definições, porém não há regra clara. Além disso, como coloca Turban & Volonimo (2013, p. 326): “é difícil compreender totalmente o BI, porque seus aplicativos não são sistemas autônomos, nem dão suporte a objetivos específicos, como outros sistemas (SCM, CRM etc)”. A Tabela 1 apresenta diferentes definições de BI identificadas na literatura. Na primeira coluna é apresentado o conceito e na segunda, o autor.

**Tabela 1: Conceitos de BI**

**Fonte:** SISTEMAS, CIBERNÉTICA E INFORMÁTICA VOLUME 11 - NÚMERO 1 - ANO 2014.

Conceito	Autor
“Um sistema automático para disseminar informação para vários setores de qualquer empresa, utilizando máquinas de processamento de dados (computadores), auto abstração e auto codificação de documentos e criando perfis para cada ponto de ação da organização por palavra padrão”.	(LUHN, 1958, p. 314)
“... é a aplicação de um conjunto de técnicas e ferramentas que são propostas para auxiliar na administração de um negócio e na tomada de decisões”.	(SANTOS, 2009)
“Pode ser definido como o apoio de modelos matemáticos e metodologias de análise que explorem os dados disponíveis para gerar informação e conhecimento para processos de tomada de decisões complexas”.	(VERCELLIS, 2009, p. 3)
“... refere-se às aplicações e tecnologias para consolidar, analisar e oferecer acesso a grandes quantidades de dados, para ajudar os usuários	(RAINER & CEGIELSKI, 2011, p.

a tomar melhores decisões empresariais e estratégicas. As aplicações de BI oferecem visões históricas, atuais e previsíveis das operações de negócio”.	311)
“... de forma mais ampla, pode ser entendido como a utilização de variadas fontes de informação para definir estratégias de competitividade nos negócios da empresa. Podem ser incluídos nessa definição os conceitos de estruturas de dados, representadas pelos bancos de dados tradicionais, <i>data warehouses</i> , e <i>data marts</i> , criados objetivando o tratamento relacional e dimensional de informações, bem como as técnicas de <i>data mining</i> aplicadas sobre elas, buscando correlações e fatos ‘escondidos”.	(BARBIERI, 2011, p. 95)
“... Une dados, tecnologia, análises e conhecimento humano para otimizar decisões nos negócios e ultimamente tem dirigido o sucesso das empresas. Programas de BI usualmente combinam um <i>data warehouse</i> empresarial (EDW) e uma plataforma de ferramentas de BI para transformar dados em informações usáveis para o negócio“.	(TDWI, 2013)
“... refere-se à coleção de SIs e de tecnologias que dão suporte à tomada de decisão gerencial ou operacional – controle pelo fornecimento de informações nas operações internas e externas”.	(TURBAN & VOLONIMO, 2013, p. 326)
“... É um termo guarda-chuva que inclui as aplicações, infraestrutura e ferramentas e as melhores práticas que permitem acesso e análise de informações para promover e otimizar decisões e desempenho”.	(GARTNER, 2014)
“... refere-se às aplicações e tecnologias que são utilizadas para coletar, acessar e analisar dados e informações de apoio à tomada de decisão”.	(BALTZAN & PHILLIPS, 2012, p. 234)
“... É o processo de transformação de dados brutos em informações utilizáveis para maior efetividade estratégica, insights operacionais e benefícios reais para o processo de tomada de decisão nos negócios”.	(DUAN & XU, 2012, p.681)

Ao analisar os conceitos da Tabela 1 é possível observar que uma opinião comum entre os autores é que o BI trata-se de um método para organização das informações dentro de uma empresa, com a finalidade de disseminar a informação para pessoas interessadas, agregar valor ao negócio e guiar o direcionamento à tomada das decisões.

Porém, é possível definir BI em uma perspectiva tecnológica. Wayne Eckerson (2004, p. 3) conceitua *Business Intelligence* (BI) como: "Os processos, tecnologias e ferramentas necessárias para transformar dados em informação, informação em conhecimento, e conhecimento em planos que conduzem ações de negócio lucrativas". Em Gartner (2014, p. 1) vemos também que BI "... é um termo guarda-chuva que inclui as aplicações, infraestrutura e ferramentas e as melhores práticas que permitem acesso e análise de informações para promover e otimizar decisões e desempenho". Logo, podemos concluir que o conceito de BI engloba uma série de estruturas, metodologias e ferramentas. A Figura 1 apresenta as várias ferramentas aplicadas no BI sob a perspectiva tecnológica, as quais serão descritas nas próximas seções.

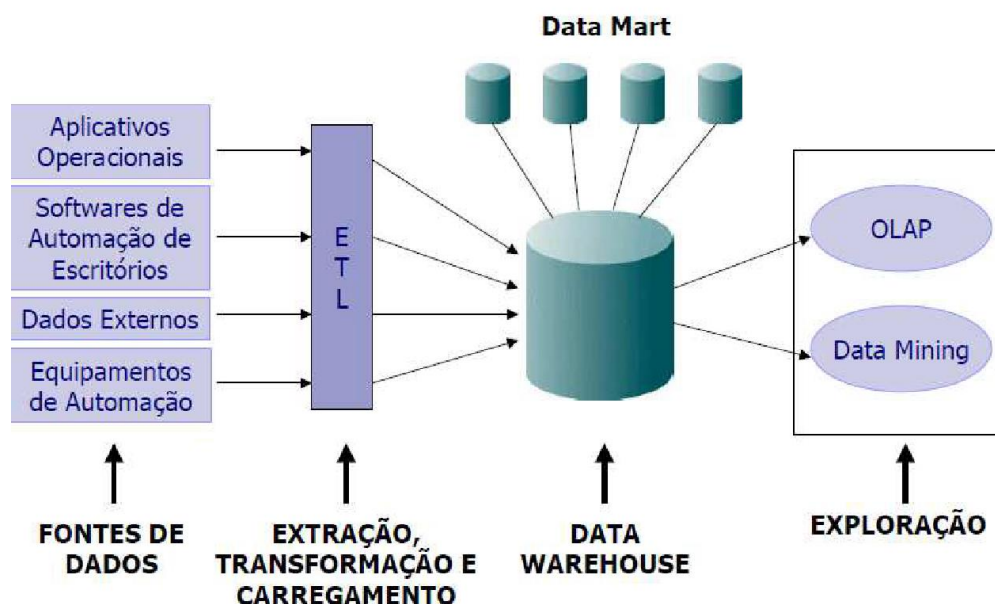
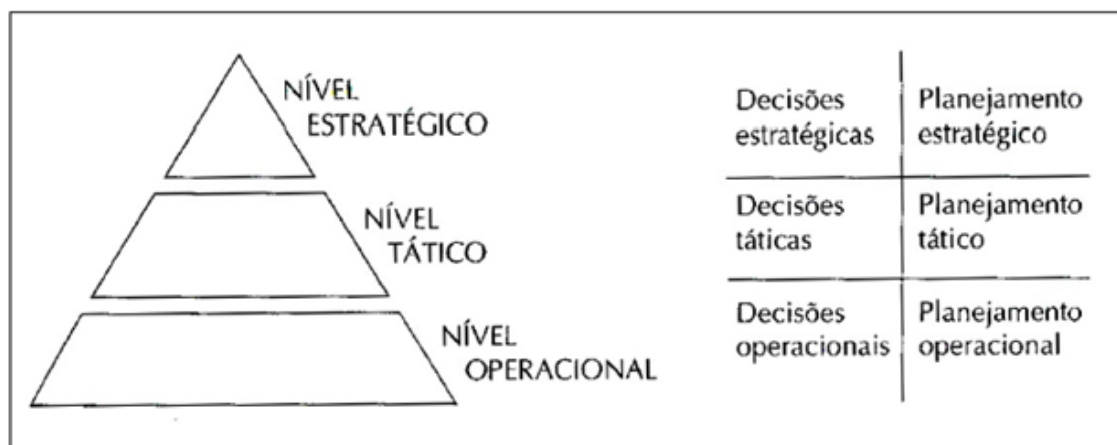


Figura 1 : Processos de BI. Fonte: Autor

## 2.2 OS NÍVEIS ORGANIZACIONAIS

A priori é importante entendermos como são divididos os níveis organizacionais dentro das empresas, e que tipo de informação e análises cada nível necessita para desenvolver suas atividades. A partir daí, conseguiremos entender onde cada tipo de informação se enquadra e, com isso, desenvolver uma análise mais acurada do sistema de *Business Intelligence* que precisa ser desenvolvido, conforme a particularidade de cada necessidade.

Oliveira (2001) coloca que ao analisarmos os grandes níveis hierárquicos, o planejamento pode ser dividido em três tipos: estratégico, tático e operacional. Os três tipos de planejamento se relacionam aos níveis de decisão através de uma pirâmide organizacional, apresentada na Figura 2.



**Figura 2 :** Níveis Organizacionais. **Fonte:** Oliveira (2001, p. 45)

Além disso, Oliveira (2001) nos mostra que o planejamento estratégico é considerado um processo gerencial que facilita o dia a dia do executivo no cumprimento das metas planejadas pela organização. O autor também coloca que o planejamento estratégico desenvolve ações para alcançar os objetivos a longo prazo e está presente em toda a organização. No entanto, dentro do planejamento tático os objetivos são a curto prazo e abrangem apenas partes da organização. Por fim, no planejamento operacional são atribuídos os planos de ação com intensidade mais restrita, menores riscos e execução imediata. A seguir serão descritos com mais detalhes cada um dos Três níveis de planejamento mencionados.

#### a) Nível Estratégico

... o planejamento estratégico é uma técnica administrativa que, através da análise do ambiente de uma organização, cria a consciência das suas oportunidades e ameaças dos seus pontos fortes e fracos para o cumprimento da sua missão e, através desta consciência, estabelece o propósito de direção que a organização deverá seguir para aproveitar as oportunidades e evitar riscos. (Fischmann & Almeida, 1991, p. 25)

Ainda segundo Maximiano (2000), o planejamento estratégico é de responsabilidade dos executivos dos níveis mais altos da organização e está associado às tomadas de decisões sobre produtos e serviços que a organização pretende oferecer, como também os clientes e mercados que pretende atingir. Assim, planejamento estratégico pode ser considerado como a melhor forma de transmitir para a organização, mais segurança e uma maior capacidade de enfrentar futuras situações que possam surgir, como também, auxiliar no alcance de metas e objetivos desejados pela organização.

Conforme coloca Chiavenato (2000), o planejamento estratégico é a melhor opção para as empresas que buscam aplicar determinadas estratégias, com o objetivo de alcançar suas metas, através de um planejamento global e a longo prazo.

#### b) Nível Tático

Segundo Chiavenato (2000), planejamento tático abrange determinados setores da organização e é definido no nível intermediário, geralmente é projetado para o médio prazo, e apresenta uma grande preocupação em atingir principalmente os objetivos departamentais.

Oliveira (2003) nos mostra que o planejamento tático é desenvolvido em níveis organizacionais inferiores, tendo como principal finalidade a utilização eficiente dos recursos disponíveis para a consecução dos objetivos previamente fixados, seguindo uma estratégia predeterminada, bem como as políticas orientativas para o processo decisório da empresa.

#### c) Nível Operacional

Conforme Chiavenato (2000), planejamento operacional abrange tarefas ou atividades específicas, projetadas no curto prazo, a fim de alcançar as metas impostas pela organização. A este nível competem às decisões do cotidiano das organizações, que estabelece uma ligação entre decisões táticas e estratégicas e seu impacto se dá no curto prazo, afetando apenas determinados setores ou áreas específicas. O foco básico está nas atividades do dia a dia da empresa.

Para Oliveira(2001), o planejamento operacional apresenta uma formulação por meio de documentos escritos, metodologias e implantação. Representa a união

de algumas partes do planejamento tático, com um detalhamento maior, em um menor prazo de acontecimento.

## 2.3 OS NÍVEIS ORGANIZACIONAIS E O BI

Uma vez compreendido o que cada nível organizacional representa e qual a sua responsabilidade na empresa, podemos entender como isso se aplica quando surge a necessidade de desenvolver um projeto de BI.

Ao analisarmos o BI estratégico, observamos que para Quinn (2014), vice-presidente de Produtos de *Business Intelligence* e Serviços de Apoio Comercial da empresa *Information Builders*, o principal objetivo deste é impulsionar o desempenho geral da empresa. Ele ressalta que após definida e aceita a estratégia pela administração, várias funcionalidades são utilizadas, como mapas estratégicos, *scorecards*<sup>1</sup>, relatórios, com o intuito de transmitir a estratégia aos colaboradores na forma de objetivos mensuráveis. Por outro lado, para verificar o sucesso da estratégia traçada, são analisados vários fatores cruciais, como índices de satisfação de clientes, quotas de mercado, margens de lucro, entre outros, que revelarão o progresso, ou falta dele, no sentido de alcançar os objetivos traçados. Conclui-se que no nível do BI estratégico concentra-se o monitoramento do desempenho e da realização dos objetivos.

Quinn(2014) também ressalta que assim que a estratégia estiver definida, é hora de começar a trabalhar o BI tático. Para ele, enquanto o BI estratégico define as medições de desempenho essenciais, o BI tático é utilizado para identificar a origem dos problemas assim que eles forem descobertos. Por exemplo, se os lucros estiverem em queda ou se os índices de perda de clientes estiverem em alta, através do BI tático as empresas poderão investigar que fatores estão na origem destes resultados. O autor afirma que é possível, neste nível, identificar e isolar os problemas que constituem um obstáculo ao desempenho da empresa sob múltiplas perspectivas. Os resultados obtidos nas atividades táticas são os que dirigem as iniciativas operacionais.

---

<sup>1</sup> Scorecard : Registro estatístico utilizado para medir o progresso de um objetivo específico através de indicadores quantificáveis e verificáveis.

(<http://www.significados.com.br/bsc-balanced-scorecard/> acessado em 03/03/2016).

O BI tradicional abrange o nível estratégico e o tático. A novidade agora é o uso de BI também no nível operacional. Segundo Turban e Volonimo (2013) a alta competitividade é o principal fator que influencia empresas a adotarem BI no nível operacional. É no nível operacional que se busca melhorar decisões, aumentar a eficiência de processamento e preocupa-se em dar respostas mais rápidas aos clientes. No caso do BI operacional, análises precisam ser feitas em tempo real. Para solucionar isso, as empresas buscam diminuir tempo de atualização do *DataWarehouse* (este conceito será abordado na seção 2.4) ou executar análises diretamente na base transacional.

O BI operacional, em resumo, proporciona ferramentas para as decisões do cotidiano, que acontecem nos níveis inferiores das organizações, com vista a alcançar os objetivos estratégicos. Estas iniciativas poderão automatizar processos, dar poder de decisão a funcionários, monitorar o desempenho das iniciativas, assim como disponibilizar imediatamente informação operacional relevante, tendo um impacto direto na capacidade que a empresa tem para atingir os mais variados objetivos, como aumentar as vendas ou a rentabilidade.

Como podemos perceber, há uma articulação dos três níveis de BI. Segundo Quinn (2014), muitas das iniciativas de BI falham, ou não dão os resultados esperados, porque as empresas adquirem, implementam ou utilizam o seu software de *Business Intelligence* sem compreenderem este ciclo e a importância do seu funcionamento. Isto não significa que cada uma das três partes não funcione isoladamente ou que o foco em apenas um dos níveis não trará resultados positivos. Mas será a articulação do conjunto destes três níveis que dará a máxima rentabilidade às ferramentas de BI.

A Tabela 2 apresenta um comparativo entre os tipos de BI. Na primeira coluna, é apresentada a característica analisada. Na segunda, como essa característica é vista no BI estratégico. Na terceira coluna, como é no BI tático. Por fim, como é no BI operacional. A principal diferença está na temporalidade dos dados e no foco de negócio. O BI operacional deve ser imediato ou no mesmo dia e busca auxiliar o controle das operações diárias (TURBAN & VOLONIMO, 2013). Mesmo assim, é importante que os três tipos sejam orientados e alinhados aos objetivos da organização (BALTZAN; PHILLIPS, 2012).

Tabela 2: BI e níveis organizacionais.

<b>Característica</b>	<b>BI Estratégico</b>	<b>BI Tático</b>	<b>BI Operacional</b>
<b>Foco principal do negócio</b>	Atingir as metas empresariais em longo prazo.	Analisar dados; entregar relatórios.	Administrar operações do dia a dia com relação a atingir metas.
<b>Principais usuários</b>	Executivos, analistas.	Executivos, analistas, gerentes de setor.	Gerente de setor
<b>Métricas</b>	Mecanismo de <i>feedback</i> para acompanhar e entender como a estratégia está progredindo e quais ajustes precisam ser planejados.	Mecanismo de <i>feedback</i> para acompanhar e entender como a estratégia está progredindo e quais ajustes precisam ser planejados.	Individualizadas para que o gestor de cada linha possa obter <i>insight</i> sobre o desempenho de seus processos de negócio.
<b>Prazo</b>	Mensal, trimestral, anual.	Diário, semanal, mensal.	Imediatamente, dentro do dia.
<b>Tipos de dados ou usos</b>	Histórico, preditivo.	Histórico, preditivo.	Em tempo real ou quase em tempo real.

Fonte: Fernando Rigo Botelho e Edelvino Razzolini Filho, 2014

## 2.4 SISTEMAS OLAP X SISTEMAS OLTP

Atualmente as organizações em geral fazem uso de diversos sistemas de informação para controlar e gerenciar as suas atividades. Até as pequenas empresas possuem pelo menos um sistema de informação para controle dos

serviços. E os sistemas são aplicados para atender necessidades específicas, por isso, há sistemas do tipo operacional, gerencial, de apoio decisão, entre outros.

Os sistemas de apoio à decisão são diferentes de outros sistemas encontrados nas corporações. O objetivo destes sistemas é facilitar a tarefa de gerência de negócios através de dados, enquanto os sistemas operacionais da empresa têm a tarefa de manter o próprio negócio funcionando. Este último é voltado para a execução do negócio propriamente dito são conhecidos como sistemas transacionais (ou operacionais), os chamados *Online Transaction Processing* (OLTP). É uma categoria de software utilizado no nível operacional para armazenar dados detalhados de transações de negócios de uma organização conforme as mesmas ocorrem. Uma aplicação OLTP possui vários usuários concorrentes inserindo, alterando e excluindo dados em tempo real. Outra característica importante dos sistemas OLTP é a incapacidade de efetuarem análises complexas de forma ágil para apoiar o usuário na tomada de decisões estratégicas (Pablo ALENQUER, 2002).

Já o termo *Online Analytical Processing* (OLAP) foi criado em 1993, por Edgard F. Codd, e tem como base os sistemas de apoio à decisão, pois ao contrário dos sistemas OLTP, que executam transações sobre dados do negócio, os sistemas OLAP são projetados para a análise de dados multidimensionais. Em 1995, *THE OLAP COUNCIL* – um conselho para padronização da tecnologia OLAP - publicou sua conceituação do termo:

O Online analytical processing (OLAP) é uma categoria de tecnologia de software que possibilita que os analistas, gerentes e executivos tenham entendimento sobre os dados de forma rápida, consistente, e com acesso interativo a uma ampla variedade de visões possíveis de informações que foram transformadas a partir de dados brutos para refletir a dimensionalidade real da empresa como entendida pelo usuário. (The OLAP Council, 1995, Disponível em:<http://dssresources.com/glossary/olaptms.html> Acesso 20/102013)

Logo, podemos conceituar OLTP em comparação ao OLAP:

Em geral, as consultas de apoio à decisão usadas em aplicativos OLAP examinam quantidades muito grandes de dados, ainda que os resultados das consultas sejam pequenos. Em contraste, cada uma das operações comuns de bancos de dados, como depósitos bancários ou reservas de passagens aéreas, toca apenas uma porção minúscula do banco de dados; este último tipo de operação é chamado frequentemente OLTP (Online Transaction Processing). (Garcia-Molina, 2001, p. 654)

Ou seja, enquanto os sistemas OLTP têm seu fluxo determinado por regras/procedimentos definidos em máquina ou linha de código, os sistemas OLAP têm seu fluxo definido pelo analista de negócios, que analisa todos os acontecimentos e toma as decisões para que o fluxo seja seguido conforme a sua necessidade do momento; A Tabela 3 resume as principais diferenças entre sistemas OLTP e OLAP.

**Tabela 3:** Diferenças entre sistemas OLTP e OLAP.

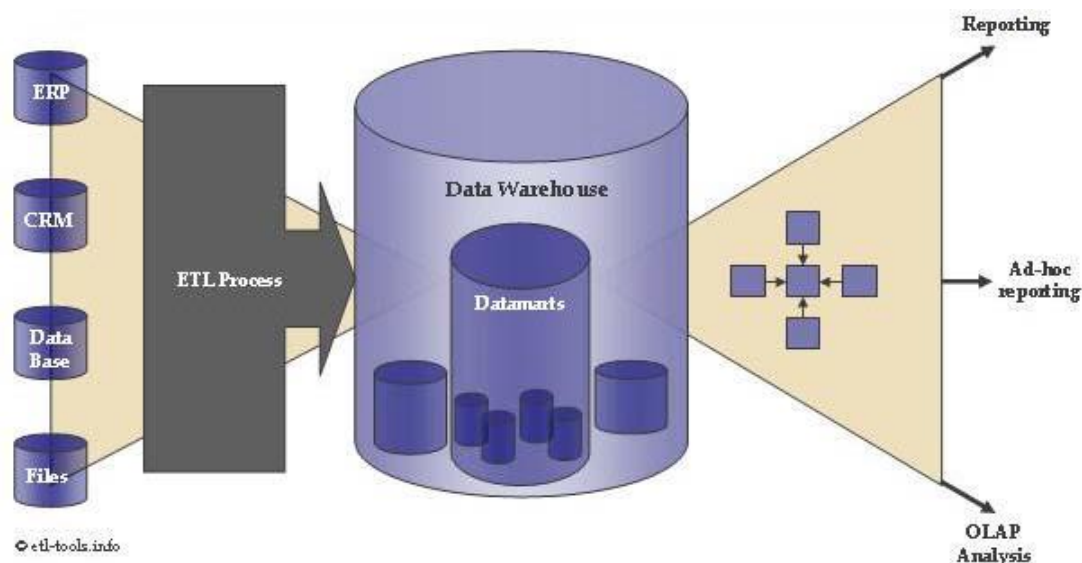
<b>Característica</b>	<b>Sistemas OLTP</b>	<b>Sistemas OLAP</b>
Frequência de execução	Maior	Menor
Previsibilidade	Maior	Menor
Quantidade de dados por consulta	Pequena	Grande
Tipo de dado consultado	Primitivo	Derivado
Periodicidade do dado consultado	Tempo corrente	Passado, presente e futuro
Derivações complexas	Poucas( se existir)	Muitas

Fonte: THOMSEN(2012, p. 13)

## 2.5 CONCEITUANDO TECNOLOGIAS DE BI

Como já mencionado, o BI pode ser visto na perspectiva tecnológica e, neste contexto, outros conceitos técnicos são aplicados, os quais serão apresentados brevemente. Isto, porque este trabalho não visa a implementação de tecnologias de BI, mas sim a análise e aplicação de ferramentas já existentes no mercado.

A implementação de BI em uma empresa segue em geral um *framework* que envolve diferentes conceitos, conforme ilustra a Figura 3 . Inicia-se na extração dos dados em um sistema OLTP; depois através de um processo de ETL (Extract,,Transform, Load) é feito o carregamento e transformação dos dados contidos no *Data Warehouse* que por sua vez pode ser organizado por *Data Marts*. Tanto o *Data Warehouse* como o *Data Mart* servem de fonte de dados para exploração OLAP e *Data Mining*, com resultado final exibido em interface intuitiva.

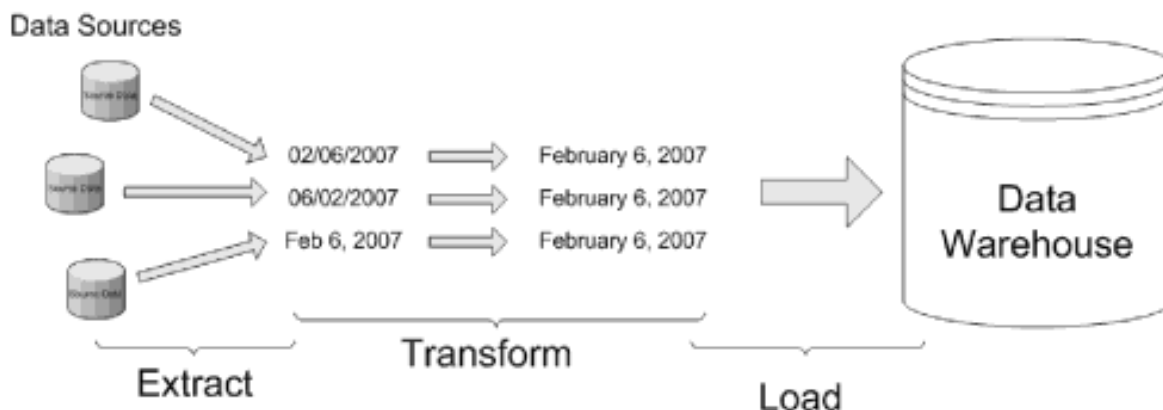


**Figura 3** : Processo de BI (Business Intelligence). Fonte: Miranda (2010, MIRANDA, Alira. Disponível em: < <http://www.alvimiranda.com/2010/04/pentahoo-que-e.html>>. Acesso em: 12 de novembro de 2015)

### 2.5.1 ETL

Como já vimos anteriormente, sistemas OLTP são utilizados para armazenar os dados das atividades de uma organização conforme elas ocorrem. Esses sistemas se caracterizam por suportar múltiplos usuários acessando de forma concorrente, ou seja, ao mesmo tempo, um banco de dados. Os dados operacionais são gerados diariamente por sistemas OLTP e se encontram em seu estado “puro” (sem alterações) para análise. Sistemas OLTP são usados como fontes de dados para o *Data Warehouse* e o processo de exportação desses dados é feito através do método ETL (*Extraction, Transformation and Load*) (TEOREY, LIGHTSTONE, NADEAU, 2007).

Esse processo indica que as informações devem ser extraídas das bases transacionais, transformadas para se adequar ao modelo multidimensional, e carregadas no *Data Warehouse*. Existem diversas ferramentas que visam apoiar a execução desse processo. Durante o processo a parte mais difícil se encontra na transformação, pois, indica as regras que devem ser aplicadas sobre os dados para evitar redundâncias e inconsistências garantindo uma visão única e integrada dos dados. O processo de ETL possibilita extrair os dados de diversas fontes criando uma visão única e consistente dos dados. A Figura 04 ilustra o processo de ETL.



**Figura 4** : Processo de ETL. Fonte: Withee (2010, p. 78)

Na literatura não se encontra uma regra clara sobre qual a melhor forma de empregar o processo de ETL. Este pode ser feito por meio de ferramentas desenvolvidas pela própria empresa ou adquirida por meio de fornecedores externos e consiste em 3 fases principais:

- *Extract* (Extração): Leitura dos dados de uma ou várias fontes.
- *Transform* (Transformação): Conversão dos dados extraídos em um formato especificado a ser utilizado no *Data Warehouse*.
- *Load* (Carga): Inserção dos dados no *Data Warehouse*.

Sobre o processo de ETL Turban ainda acrescenta:

As ferramentas de ETL também são responsáveis por transportar os dados entre as várias fontes, documentar como os elementos dos dados são ajustados enquanto se movem entre a fonte e o destino e administração das operações e processos envolvidos. Os dados utilizados no processo de ETL podem ser originados de várias fontes como arquivos extraídos de sistemas OLPT, planilhas e arquivos externos. (TURBAN, 2009, p. 344)

Ou seja, este processo é responsável pela consolidação dos dados de entrada. O bom desempenho e qualidade do processo de ETL, portanto, é essencial para a acurácia e confiabilidade das informações que estarão contidas no *Cubo/Datamart*. Pode-se dizer que o processo de ETL é o garantidor da qualidade dos dados para todo o projeto.

## 2.5.2 STAGING AREA

Após se extrair dados de vários sistemas operacionais e de fontes externas, esses dados podem ser preparados antes de serem inseridos em um *Data Warehouse*. Mas para isso, devem ser alocados em uma área secundária denominada *Staging Area*.

Esta área pode ser descrita como um componente onde os dados oriundos de sistemas operacionais são armazenados temporariamente com funções para limpar, mudar, combinar, converter, remover duplicidades de modo a preparar os dados que posteriormente serão inseridos no *Data Warehouse*.

Ponniah (2001) menciona que essa área intermediária é necessária visto que os dados em um *Data Warehouse* necessitam estar em um formato orientado a assuntos unindo informações que são extraídas de múltiplos sistemas operacionais. Entre os exemplos de bancos de dados orientados a assuntos mencionados por Inmon (1992) podemos citar uma empresa varejista em que os assuntos principais poderiam ser produto, unidade de manutenção de estoque, venda, fornecedor etc. Cada empresa tem o seu conjunto único de assuntos.

## 2.5.3 DATA WAREHOUSE

O *Data Warehouse* é um banco de dados produzido para suportar os sistemas de apoio à decisão. Este repositório de dados é utilizado para armazenar tanto dados históricos como dados mais atuais de todas as áreas da organização. Essas informações são armazenadas de forma a permitir atividades de processamento analítico<sup>2</sup> como OLAP e *Data Mining*. (TURBAN, 2009).

Bill Inmon, considerado por muitos autores o pai do *Data Warehouse*, faz a seguinte definição: “Um *Data Warehouse* é uma coleção de dados orientado a assunto, integrado, tempo-variante e não volátil para suporte ao gerenciamento dos processos de tomada de decisão”. (INMON, 1992, p. 33) (PONNIAH, 2001) (INMON, 1992).

---

<sup>2</sup> Processamento analítico : “utilização do computador para se produzir uma análise para decisão gerencial, normalmente envolvendo análise de tendências, caracterização, análise demográfica etc.” (INMON, p. 386, 1992)

Barbieri (2001) também compartilha desta opinião afirmando que um *Data Warehouse* é um banco de dados orientado por assunto, não volátil, integrado e variante ao tempo. Ainda diz que o alto nível de complexidade para o desenvolvimento de certos *Data Warehouse*, faz com que seja necessário iniciar o desenvolvimento através de *Data Marts* (será explicado na seção 2.4.4) departamentais, que ao final, serão integrados ao *Data Warehouse*.

A seguir será feita uma breve descrição de cada uma das características do *Data Warehouse*:

- **Orientado ao assunto:** significa que as informações são armazenadas e se relacionam em grupos de assuntos de interesse de uma organização.
- **Integrado:** significa que as informações provenientes dos sistemas operacionais de uma organização são consolidadas, de forma que sejam consistentes e tenham um único significado, ou seja, são codificadas de uma única forma.
- **Variante no tempo:** Significa que as informações do *Data Warehouse* sempre estão associadas a um ponto no tempo. Isto possibilita a análise em diferentes períodos de tempo e também previsões futuras.
- **Não volátil:** Significa que as informações são carregadas em massa no *Data Warehouse* e não sofrem atualizações como nos sistemas OLTP, sendo disponibilizadas apenas para consultas.

Um ponto de atenção sobre os *Data Warehouses* citado por TEOREY, LIGHTSTONE, NADEAU (2011, p.190): “como os dados nunca são deletados, *Data Warehouses* crescem com o tempo podendo chegar a centenas de gigabytes ou até mesmo terabytes”. Vale reforçar também, conforme (TURBAN, 2009), que um *Data Warehouse* é composto de dados de várias fontes, tanto internas como externas, de uma maneira consistente de acordo com as necessidades da organização.

Vercelis (2009) identifica 3 principais fontes para dados usados para a criação de um *Data Warehouse*:

**(1) Dados Internos:** dados armazenados em bancos de dados que são “alimentados” pelos sistemas operacionais ou sistemas transacionais. Esses dados podem ser oriundos de diferentes componentes do sistema de informação como sistemas de retaguarda, que coletam registros transacionais como pedidos, orçamentos, inventários e dados de logística. Sistemas de linha de frente que

contém dados originados de atividades de call center, assistência ao consumidor e execução de campanhas de marketing. E de sistemas web que coletam transações de vendas pela internet, dados de navegação do *website* e formulários preenchidos por clientes existentes ou em potencial. Os softwares que dão suporte à esses sistemas operacionais são os OLTP, apresentados na seção 2.3.

**(2) Dados externos:** existem várias fontes de dados externos que pode ser usadas para estender o valor da informação armazenada em bancos de dados internos. Esses dados podem incluir pesquisas de mercado, opiniões de consumidores e indicadores econômicos.

**(3) Dados pessoais:** em muitos casos pessoas encarregadas de tomar decisões nas empresas se baseiam em avaliações e planilhas armazenadas em seus computadores pessoais. A recuperação desse tipo de informação e a integração com os dados estruturados de fontes internas e externas também pode ser feita pelo uso de sistemas de gerenciamento de conhecimento.

Por fim, sobre as vantagens de se construir um *Data Warehouse*, encontramos na literatura de PONNIAH (2001) alguns pontos:

- Possibilita uma visão geral e integrada de toda a empresa;
- Faz com que informações históricas e atuais da empresa estejam facilmente disponíveis para auxílio à tomada de decisão;
- Torna transações de suporte à decisão possível sem interromper os sistemas operacionais;
- Apresenta uma fonte de informações estratégicas de maneira flexível e interativa.

#### 2.5.4 DATAMART

Enquanto os *Data Warehouses* combinam dados de toda a organização, *Data Marts* são bancos de dados menores focados em apenas um assunto ou departamento da empresa. Eles podem ser gerados a partir de um subconjunto existente em um *Data Warehouse* nos quais os dados são ajustados aos requisitos de cada área ou departamento. *Data Marts* podem ser criados com o objetivo de reduzir a complexidade de um projeto de *Data Warehouse* oferecendo uma maior facilidade de interpretação dos dados (TURBAN, 2009).

Ainda complementando com Vercelis (2009), com o propósito de implementar

aplicações de *Business Intelligence*, algumas empresas preferem projetar e desenvolver de maneira incremental uma série de *Data Marts* ao invés de um *Data Warehouse* central reduzindo assim o tempo de implementação e incertezas relacionadas ao projeto.

Os dados em um *Data Mart* são estruturados de forma a visualizar as informações departamentais de uma empresa. Ponniah (2001) descreve as principais diferenças entre dados de um *Data Warehouse* e um *Data Mart*, conforme Tabela 4.

**Tabela 04:** Data Warehouse x Data Mart.

<b><i>Data Warehouse</i></b>	<b><i>Data Mart</i></b>
Nível de Empresa/Organização	Departamental
União de vários <i>Data Marts</i>	Único processo de negócio
Dados recebidos a partir do <i>Staging</i>	Modelo estrela (fatos e dimensões)
Estruturado do ponto de vista organizacional	Estruturados em ponto de vista de departamento

Fonte: PONIAH (2001, p. 26)

## 2.6 O MODELO DE DADOS DIMENSIONAL

Nos bancos de dados relacionais, usados nos sistemas transacionais, a redundância dos dados é evitada, sendo aceita somente em determinados casos em que é realmente necessária. Esta redundância é eliminada através de processos de normalização. A normalização das tabelas traz benefícios nos casos em que muitas transações são efetuadas, pois estas se tornam mais simples e rápidas.

Já no *Data Warehouse*, conforme PERNAS (2003), ocorre o contrário, as transações operam sobre um grande volume de dados e não são simples, nem frequentes, não sendo conveniente a normalização das tabelas, pois no ambiente de *Data Warehouse* ocorrem poucas transações concorrentes e cada transação acessa um grande número de registros.

Pernas (2003) também coloca que outra grande diferença entre o banco de dados relacional e o *Data Warehouse* está na modelagem dos dados. Enquanto que no primeiro, geralmente, implementa-se a modelagem Entidade-Relacionamento

(ER), no segundo se utiliza de uma modelagem lógica denominada de modelagem dimensional, também chamada de modelagem multidimensional.

Kimball (2013, p. 7) nos apresenta a modelagem dimensional da seguinte forma: “é uma técnica de projeto lógico que busca apresentar os dados em uma estrutura padrão e intuitiva que permite um acesso de alta performance”.

Imhoff, Galemme e Geiger complementam Kimball neste conceito:

Um modelo dimensional é uma forma de modelagem de dados que condiciona os dados de acordo com consultas e processos específicos do negócio. Os objetivos são a compreensibilidade dos usuários de negócio e o desempenho de consultas multidimensionais. (Imhoff et. al, 2003, p. 400)

A seguir será descrito, sucintamente, os principais elementos do modelo multidimensional:

**a) Tabela Fato:** É a tabela principal, onde as medidas de desempenho numéricas do negócio são armazenadas (valor arrecadado, custo, quantidade de contratos ou quaisquer outro tipo de dado quantitativo).

Imhoff, Galemme e Geiger definem medida como:

Uma medida é um termo da modelagem dimensional que se refere a valores, geralmente numéricos, que medem algum aspecto do negócio. Medidas residem em tabelas fato. Os termos dimensionais medida e atributo, juntos, são equivalentes ao uso do termo atributo da modelagem relacional. (Imhoff et. al, 2003, p. 403)

A tabela fato deve receber somente os campos chaves das tabelas dimensões (*Primary Key, Foreign Key*), criando o relacionamento entre elas. Estes dados inseridos dentro da fato podem ser detalhados, fazendo uma busca na tabela dimensão retornando os dados qualitativos correspondente àquela informação.

Este Relacionamento entre Fatos e Dimensões na modelagem multidimensional é citado em Singh:

Na modelagem, as estruturas de dados são organizadas para descrever 'medições' e 'dimensões'. Medições fornecem os dados numéricos coletados que ficam armazenados na tabela de fatos central. As dimensões são os parâmetros do negócio que definem cada transação. São armazenadas em tabelas satélites vinculadas à tabela de fatos central. Por exemplo, os dados armazenados em tabelas de fatos incluem dados de vendas, estoque, assinaturas de periódicos, despesas e margem bruta. Tabelas típicas incluem dados de tempo, geografia, conta e produto. (Singh, 2001, p.146)

Os dados na tabela fato são normalizados, sumarizados e não descritivos, ou seja, os dados não se repetem e para ter as características de cada informação é necessário realizar cruzamentos com as dimensões.

Imhoff, Galemme e Geiger (2003, p. 401) descrevem a tabela fato da seguinte forma: "uma tabela fato é a tabela dentro de um modelo dimensional que contém as medidas e métricas de interesse".

**b) Tabela Dimensional:** As tabelas de dimensão contém as características textuais do negócio (escolaridade, sexo, região, país, etc). Seria algo do tipo, a tabela dimensão de Aluno possui os dados relacionados aos alunos. A dimensão de Professor, possui os dados referentes aos professores. A dimensão de escola, possui atributos referentes às escolas. Elas não se relacionam entre si, a única ligação que me responde qual professor leciona em tal escola, são as chaves que estão na tabela fato.

Imhoff, Galemme e Geiger definem dimensões como:

Uma tabela dimensão é um conjunto de tabelas de referência que proveem a base para restringir e agrupar consultas para a informação em uma tabela fato dentro de um modelo dimensional. A chave da tabela dimensão tipicamente é uma parte da chave concatenada da tabela fato, e a tabela dimensão contém informação descritiva e hierárquica. (Imhoff et. al, 2003, p. 400)

Na Figura 5 temos um exemplo simples de dados distribuídos dentro da tabela fato e dimensões. Pode-se analisar como os atributos se relacionam entre fato e dimensões.

Fato Performance						
FK_PRODUTO	FK_ESTOQUE	FK_CLIENTE	FK_TEMPO	VLR_VENDIDO	QTDE_VENDIDO	CUSTO_COMPRA
121	12	1	66	17.334	1.534	767
121	12	1	55	1.379	122	61
121	12	2	66	14.069	1.245	623
121	12	2	55	2.633	233	117
121	22	1	66	13.820	1.223	612
121	22	1	55	1.616	143	72
121	22	2	66	13.944	1.234	617
121	22	2	55	3.899	345	173
122	12	1	66	36.397	3.221	1.611
122	12	1	55	2.667	236	118
122	12	2	66	151.929	13.445	6.723
122	12	2	55	23.854	2.111	1.056
122	22	1	66	12.939	1.145	573
122	22	1	55	4.882	432	216
122	22	2	66	13.944	1.234	617
122	22	2	55	1.243	110	55

Dimensão Produto	
PK_PRODUTO	DESCRIÇÃO
121	TV
122	DVD

Dimensão Estoque	
PK_ESTOQUE	DESCRIÇÃO
12	Físico
22	Online

Dimensão Cliente	
PK_CLIENTE	DESCRIÇÃO
1	P. Física
2	P. Jurídica

Dimensão Tempo	
PK_TEMPO	DESCRIÇÃO
66	Mensal
55	Semanal

Figura 5 : Exemplo de tabela fato e dimensão Fonte: Autor

**c) Membros e Hierarquias:** São uma classificação de dados dentro de uma dimensão. O membro é nomeação do dado dentro da dimensão, como por exemplo: Dimensão Oceano; Membros da dimensão: Índico, Pacífico e Atlântico. Já a hierarquia acontece quando uma informação depende de outra. Um exemplo clássico é a hierarquia País>Estado>Cidade, onde existe um relacionamento entre os membros da dimensão.

Em *The OLAP Council* temos a citação abaixo sobre membros de dimensões:

Um membro de uma dimensão é um nome ou identificador discreto usado para identificar uma posição de um item de dado dentro de uma dimensão. Por exemplo, 'janeiro de 1989' ou 'primeiro semestre 1993' são exemplos típicos de membros de uma dimensão Tempo. Atacado, Varejo, etc., são exemplos típicos de membros de uma dimensão Canal de Distribuição. (1995, [www.olapcouncil.org](http://www.olapcouncil.org))

E *The OLAP Council* também define a hierarquia de dimensões como:

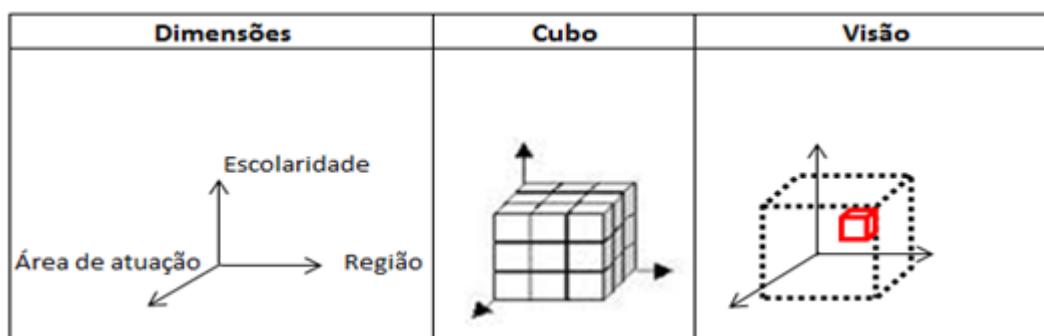
Quaisquer membros de uma dimensão podem ser organizados com base em relacionamentos pai-filho, tipicamente onde um membro pai representa a consolidação dos membros que são seus filhos. O resultado é uma hierarquia, e os relacionamentos pai/filho são relacionamentos hierárquicos. (1995, [www.olapcouncil.org](http://www.olapcouncil.org))

**d) Cubo de dados:** Os modelos dimensionais nos permitem criar um método simples e de fácil entendimento, o qual define como o usuário irá acessar a informação. A forma mais comum de se representar o modelo dimensional é

utilizando um cubo de três ou quatro dimensões, onde os usuários podem acessar uma parte do cubo sob várias visões. Renato Cramer(2006) appud Kay(2005) define um cubo de dados como:

Um cubo de dados é um tipo de matriz multidimensional que permite que os usuários explorem e analisem uma coleção de dados de muitas perspectivas diferentes, geralmente considerando três fatores (dimensões) de cada vez. Kay(2005, p. 28)

Podemos entender melhor o conceito do cubo ao observar a Figura 6.



**Figura 6** : Exemplo de tabela fato e dimensão. Fonte: Autor.

Por exemplo, suponha que os dados de um professor são armazenados em uma tabela de professor. A análise de desempenho dos professores no que diz respeito à escolaridade, região da escola, área de atuação é medida a partir do cruzamento de visões (dimensões) dentro de um cubo. O número de dimensões é diretamente proporcional ao nível de detalhes obtidos a partir do modelo.

A modelagem dos dados ajuda a identificar as visões (tabelas de dimensão) para o dado problema, o planejamento de consultas (*queries*) e a separar as medidas (tabela de fatos) para responder a essas *queries*. A partir desta análise da modelagem, podemos ver como os dados serão segmentados no momento da geração do relatório. Para planejamento correto dessas *queries*, é essencial declarar a granularidade correta para o datamart<sup>3</sup>. Isto significa especificar exatamente o que uma linha individual da tabela de fatos representa. É muito importante ressaltar que os dados na tabela fato são sumarizados, ou seja, não existem dados analíticos dentro de um cubo.

<sup>3</sup> Datamart: Subconjunto de um datawarehouse normalmente orientado a uma linha de negócio específica (Fonte: <http://www.gartner.com/>).

## 2.6.1 VANTAGENS DA MODELAGEM DIMENSIONAL

Em HOKAMA et al. encontramos uma análise sobre as vantagens do modelo dimensional:

A principal vantagem da utilização de um modelo de dados dimensional está no cálculo de dados resumidos. Como o modelo relacional trabalha com normalização, suas tabelas possuem menos registros e não têm redundâncias, apresentando assim uma melhor performance nas tarefas do dia-a-dia, como inclusões, alterações e exclusões de registros. Mas ele só é adequado para consultas simples de poucos registros. Para análises mais complexas, com um universo de registros maior, o modelo dimensional oferece uma melhor alternativa, economizando em junções com várias tabelas e armazenando dados que facilitam a análise das informações. (HOKAMA et al., 2004, p. 32)

Já Abelló nos mostra os benefícios da seguinte forma:

Os benefícios da modelagem dimensional são dois. De um lado, ela toma os esquemas de dados mais compreensíveis para os usuários finais, e por outro lado, ela permite usar armazenamento específico e técnicas de acesso que melhoram o desempenho de queries. A maneira para obter estes benefícios é a simplificação dos esquemas de dados, de forma que eles só contenham as coisas essenciais (i.e. um fato para ser analisado e suas dimensões de análise). Estes esquemas são próximos da concepção de dados dos analistas, e sugerem um tipo específico de queries, de forma que o sistema pode ser personalizado facilmente para resolvê-las com bons tempos de resposta. (Abelló, 2002, p. 11)

Se tentarmos analisar a performance obtida na modelagem de dados relacional tradicional e compará-lo com o modelo de dados dimensional, observa-se a maior performance em bancos de dados multidimensionais (COLLINS,2001).

Algumas diferenças entre a modelagem de dados relacional e modelagem de dados dimensional podem ser vistas na Tabela 6.

**Tabela 6:** Comparação modelo Dimensional x Relacional.

<b>Modelo Dimensional</b>	<b>Modelo Relacional</b>
Padrão de estrutura mais fácil e intuitiva.	Modelo mais complexo.
Anterior ao MER, anos 60.	Ênfase nos bancos de dados relacionais, anos 70.
Tabelas Fato e tabelas Dimensão.	Tabelas que representam dados e relacionamentos.
Tabelas Fato são o núcleo - Normalizadas	Todas as tabelas são comumente normalizadas.

Tabelas Dimensão são os pontos de entrada.	As tabelas são indistintamente acessadas e de filtro inicial.
Tabelas Dimensão opcionalmente normalizadas.	Todas as tabelas são comumente normalizadas.
Modelo mais facilmente "joined". <sup>4</sup>	Maior dificuldade de "join" pelo número maior de tabelas.
Leitura mais fácil do modelo por usuários não especializados.	Maior dificuldade de leitura pelo usuário não especializado.

Fonte: BARBIERI (2001, p. 38)

Para ressaltar a importância deste conceito, Barbieri (2001, p. 74) nos diz que, “a modelagem de dados é seguramente um dos fatores críticos de sucesso num projeto de *Data Warehouse*, e pode representar a fronteira entre o sucesso e o seu fracasso”.

## 2.6.2 CRIANDO A ESTRUTURA DE MODELOS DIMENSIONAIS

Ballard et al. citam a estrutura de dados da análise multidimensional da seguinte forma:

A análise multidimensional tornou-se uma forma popular de estender as capacidades de consultas e relatórios. Isto é, melhor do que submeter múltiplas queries, os dados são estruturados para possibilitar acesso fácil e rápido para responder as questões que os usuários tipicamente perguntam. (Ballard et al. 2006, p.86)

Existem várias técnicas para se fazer a estrutura da modelagem dimensional, os princípios básicos são os mostrados a seguir:

- **Definição da Área de Negócio:** deve ser escolhida de acordo com as prioridades da empresa. Após a escolha, parte-se para definir aqueles que serão os processos alvos do projeto.
- **Definição de Granularidade:** define os níveis dimensionais que serão usados para o armazenamento dos dados. Os fatores que definem a escolha da granularidade estão relacionados com o volume de dados a ser mantido e com o processamento necessário para produzi-los. Lembrando que os dados deverão ser agrupados e/ou sumarizados, não podendo ser apresentados de

<sup>4</sup> Join – Instrução para combinar dados de duas ou mais tabelas de um banco de dados (Fonte: <http://www.sql-join.com/>)

forma analítica. Observe na Figura 7, um exemplo da diferença entre dados analíticos e dados sumarizados.

- **Definição das Tabelas Dimensão:** devem ser definidas em função da granularidade, dos filtros e visões que se deseja obter no cubo.
- **Normalização das Tabelas Dimensão:** as Tabelas Dimensão podem, também, sofrer o processo de desnormalização, ou seja, há a aceitação de redundância na tabela para que ela permaneça simples. Esse procedimento tem por objetivo melhorar o desempenho e facilitar a utilização. Existem duas técnicas diferentes com relação aos aspectos de normalização das tabelas Dimensão: (1) *Star Schema ou Esquema Estrela*: Abordagem que recomenda a desnormalização das tabelas dimensionais; (2) *Snowflake Schema ou Esquema de Flocos de Neve*: Abordagem que recomenda a normalização das tabelas dimensionais.

Dados Analíticos		
Professor	Graduação	Curso Leciona
Jose	Matemática	Administração
Maria	Letras	Administração
Joao	Biologia	Biologia
Joaquim	Informática	Informática
Alfredo	Informática	Informática
Judite	Matemática	Matemática
Ramon	Matemática	Matemática
Luciana	Matemática	Matemática
Marcos	Letras	Administração
Pedro	Informática	Administração
Eloisa	Biologia	Medicina
Debora	Biologia	Medicina
Simone	Letras	Administração
Juliana	Informática	Administração
Eduarda	Informática	Administração
Alfeu	Matemática	Economia
Amanda	Letras	Informática
Ester	Informática	Economia

Dados Sumarizados		
Graduação	Curso Leciona	QTD
Matemática	Administração	1
Letras	Administração	3
Biologia	Biologia	1
Informática	Informática	2
Matemática	Matemática	3
Informática	Administração	3
Biologia	Medicina	2
Matemática	Economia	1
Letras	Informática	1
Informática	Economia	1

**Figura 7 :** Exemplo de dados analíticos vs sumarizados. Fonte: Autor

A seguir serão descritos com mais detalhes as técnicas *Star Schema* e *Snowflake Schema*.

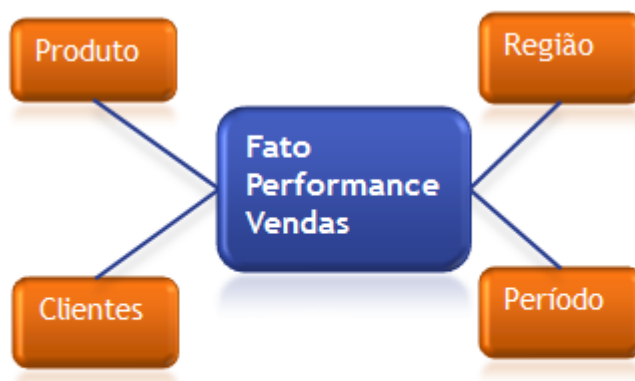
### 2.6.2.1 MODELO STAR-SCHEMA (ESTRELA)

Ralph Kimball (2013) introduziu o conceito de modelos dimensionais baseado no conceito de bases multidimensionais. O modelo de dados dimensional mais popular é conhecido como "esquema em estrela", ou em inglês, *Star-Schema*. Este estilo de modelagem ganhou este nome pois sua distribuição entre fato e dimensões se assemelha ao formato de uma estrela.

Thomsen define um *Star Schema* como:

Um arranjo de tabelas em um banco de dados relacional onde uma tabela de fatos central é conectada a um conjunto de tabelas de dimensão, uma por dimensão. O nome star vem da representação diagramática normal deste esquema, com a tabela de fatos no centro e cada tabela dimensão mostrada ao seu redor, como as pontas de uma estrela. (Thomsen, 2002, p. 638)

O modelo em estrela normalmente é usado para limitar a complexidade de cada consulta e melhorar o desempenho da consulta final para relatórios, pois, no modelo *Star-Schema*, as dimensões se relacionam diretamente com a tabela Fato, sem nenhuma subdimensão. É um dos mais simples estilos de modelagem de dados para ferramentas de BI. Observe na Figura 8 um modelo estrela.



**Figura 8** : Representação do modelo Star-Schema. Fonte: Autor

O esquema em estrela pode ser construído seguindo os passos a seguir:

1. Identificar como os usuários irão analisar a informação;
2. Identificar os dados de medidas para a tabela fato;
3. Identificar as dimensões para o cubo;
4. Separar quais colunas pertencem a cada dimensão;

## 5. Determinar a granularidade da tabela fato.

Uma característica importante do modelo *Star-Schema* é que cada dimensão irá possuir uma chave primária e as tabelas de dimensão não se relacionam entre si. Outro ponto de atenção é caso seja necessário construir uma hierarquia em *Star-Schema*, pois essa deverá ser implementada dentro da própria dimensão ao qual será hierarquizada.

### 2.6.2.2 MODELO SNOWFLAKE SCHEMA (FLOCO DE NEVE)

Bill Inmon, cientista americano considerado por muitos como o pai do data warehouse, defende a modelagem de dados feita pela terceira forma normal, que também é conhecido como o esquema floco de neve, ou em inglês *Snowflake-schema*. Ele vê o data warehouse como parte integrante da fábrica de informações da empresa. A terceira forma normal normalizou as tabelas de dimensões pelo nível do atributo, onde cada tabela de dimensão menor aponta para uma tabela de fatos. Sua abordagem é também conhecida como a abordagem de projeto top-down.

Como sempre pensamos em economizar espaço de armazenamento, pode nos parecer estranho a arquitetura dos esquemas em estrela. Note que no modelo estrela a única tabela realmente normalizada é a tabela de fatos. As dimensões possuem um alto nível de redundância em seus registros, principalmente pela presença das hierarquias. A tendência dos projetistas é normalizar as dimensões para diminuir esta redundância. Ao normalizar um esquema em estrela estaremos criando uma variação dessa estrutura conhecida como *Snowflake-schema*. A normalização desmembra dimensões em várias tabelas ligadas por chaves estrangeiras. A dimensão “Produto” da Figura 8 pode ser desmembrada na dimensão da Figura 9.



**Figura 9** : Representação Snowflake Schema. Fonte: Autor

Há uma grande rejeição à normalização de dimensões, apesar de existirem casos em que é necessária. Certamente essa normalização afetará o desempenho de *queries*, pois haverá a necessidade de *joins*. O custo-benefício do desmembramento de dimensões certamente não encoraja essa atividade. Segundo Kimball (2013, p. 106), “qualquer estimativa realista do espaço em disco necessário para o *data warehouse* pode efetivamente ignorar as tabelas de dimensão”. Isso porque proporcionalmente as tabelas de fatos são muitas vezes maiores do que qualquer dimensão. Em geral, o desmembramento de dimensões produz uma economia de espaço de menos de 1% (um por cento) (KIMBALL, 2013), o que definitivamente não justifica a grande perda de desempenho em consultas que isso acarreta.

Alguns autores defendem o uso de normalização em *Datawarehouses* com arquitetura em três camadas (SOUZA,1999; GARDNER,1998). Segundo eles, o *Datawarehouse* corporativo deveria ser construído no modelo relacional em 3FN com a adição do fator temporal dos dados. As estruturas dimensionais (esquema em estrela e cubo multidimensional) deveriam ser utilizadas apenas no projeto de *DataMarts*. Alguns desses autores baseiam-se em (INMON, 1996). Entretanto, Inmon (2002, p. 160) afirma nesta referência que “a estrutura de projeto necessária

para gerenciar grandes quantidades de dados residentes em uma entidade contida no *datawarehouse* é denominada *star join*”, e também, “a modelagem de dados clássica se aplica às tabelas de dimensão (ou seja, às entidades pouco povoadas) e o projeto de *star join* se aplica às tabelas de fatos (ou seja, às entidades muito povoadas)”. Podemos perceber por essas afirmativas que Inmon defende apenas que sejam usados *snowflake schemas* para o projeto do *DataWarehouse* Corporativo, e não o uso de modelos relacionais em 3FN. Afirmar que um *DataWarehouse* deve ser construído em 3FN não necessariamente significa que não devemos construí-lo com esquemas em estrela, já que um *snowflake schema* pode ser totalmente normalizado e mesmo assim continuará sendo uma estrutura dimensional.

### 3. FERRAMENTAS DE BI

As ferramentas de BI foram tornando-se cada vez mais importantes, na medida em que o volume de dados nas empresas aumentava rapidamente e, por consequência, necessidade de ferramentas que comportassem essas quantidades de dados gerando relatórios com rapidez e acuracidade desejadas.

A Figura 10 mostra as principais ferramentas de BI que existem hoje no mercado, sejam proprietárias ou livres, e como elas se enquadram nos quesitos “habilidade para executar” e “abrangência de visão”. Esta figura é resultado de uma avaliação anual feita pelo *Gartner Group*, a partir de uma pesquisa que em 2015 incluiu mais de 2 mil usuários de plataformas de BI.

Os fornecedores de *softwares* de BI são avaliados nos quesitos experiência de usuário, experiência de vendas, entendimento do mercado, complexidade de análise, capacitação de usuário e benefícios para o negócio. A partir dessas métricas, os fornecedores são posicionados nos seguintes quadrantes :

- *Leaders* (líderes): executam bem de acordo com sua visão atual e estão bem posicionados para o futuro.
- *Visionaries* (visionários): entendem a direção que o mercado está tomando ou tem visão para mudar as regras do mercado, mas ainda não executam bem.
- *Niche Players* (participantes de nicho): focam de maneira bem sucedida em um pequeno segmento ou estão sem foco e não inovam tanto quanto outros fornecedores.
- *Challengers* (desafiadores): executam bem hoje e podem dominar um segmento maior no futuro, porém ainda não demonstram um entendimento da direção do mercado.



**Figura 10** : Quadrante mágico Gartner. Fonte: GARTNER (2015)

Portanto, que ferramenta utilizar depende basicamente da necessidade específica de cada empresa e da sua capacidade de investimento. Neste trabalho pretendemos auxiliar neste tipo de decisão, analisando as ferramentas escolhidas, mostrando seus pontos positivos e negativos.

Primeiramente, procuramos por ferramentas que estivessem classificadas no quadrante “*leaders*” da *Gartner Group*, pois assim estaremos avaliando ferramentas com alta aceitação no mercado. Buscamos, depois, analisar em fóruns da internet, quais daquelas ferramentas são as mais populares entre os analistas. A Figura 11 apresenta o resultado de uma pesquisa realizada pelo site *In Quest*, especializado em projetos de BI, que reuniu mais de 155 mil opiniões enviadas por usuários reais

sobre as melhores ferramentas de BI, para apontar qual delas teve maior destaque em 2014.



**Figura 11:** Melhores ferramentas de BI 2014. Fonte: [www.inquesti.com.br](http://www.inquesti.com.br)

Podemos ver que as ferramentas da *Qlik*(1), *Microsoft Power BI*(2), *Microstrategy*(3), e *Tableau*(4) lideram as primeiras posições. Visto que essas fornecedoras também são classificadas pela Gartner como líderes de mercado, optamos por utilizar as ferramentas QlikView, Microstrategy, Tableau e Microsoft Power BI. A escolha também foi justificada pela facilidade de teste, visto que as 4 apresentam versões de teste sem a necessidade de aquisição de licenças para se realizar as avaliações necessárias. As ferramentas serão apresentadas nas próximas seções.

### 3.1 A FERRAMENTA QLIKVIEW

A *Qlik*, empresa que desenvolve as soluções *Qlikview* e anteriormente era intitulada como *QlikTech*, foi fundada em 1993 na Suécia por Björn Berg e Staffan Gestrelus. Sua sede hoje fica em Radnor, Pensilvania, Estados Unidos com escritórios na América do Norte e Europa. A empresa foi desenvolvida com a finalidade de apresentar dados de seus clientes em visualizações mais fáceis de entender. De acordo com os fundadores da empresa, eles tinham a visão de criar um *software* que pudesse imitar o funcionamento do cérebro, proporcionando a experiência do usuário a mais intuitiva possível. A primeira versão do *software QlikView* foi lançada em 1996. Seu nome enfatiza a habilidade do produto de prover aos usuários análise de dados detalhados com apenas um clique. Inicialmente os usuários eram algumas pequenas empresas, mas eventualmente o *QlikView* teve como cliente a multinacional de embalagens Tetrapak.

Foi a partir de 2004 que a empresa tomou a decisão de focar em *Business Intelligence* devido ao crescimento da área de BI no mercado. A mudança para BI acelerou a expansão da empresa consideravelmente nesta época. Em 2005 o *software QlikView* evolui de uma ferramenta monousuário para uma solução cliente-servidor baseado em *Java Web* o que possibilitou a ferramenta manusear uma quantidade maior de dados.

A *Qlik* também firmou parcerias com fabricantes de processadores como a Intel e a HP para garantir que o *QlikView* pudesse extrair o máximo de performance utilizando processadores multinúcleos.

Em 2007 Lars Bjorn assume como CEO da *Qlik* com a tarefa de construir processos internos e sistemas para elevar a empresa a novos patamares. A *Qlik* continuou expandindo sua presença no mercado de BI enquanto novas tecnologias ampliavam a capacidade e velocidade de análise de dados por meio de ferramentas de BI. Em 2010, Bjorn supervisionou a oferta pública inicial da *Qlik* no mercado de ações norte-americano NASDAQ.

Hoje a *Qlik* conta com mais de 35 mil clientes em cerca de 100 países e mais de 2 mil funcionários ao redor do mundo. Clientes notáveis de soluções *Qlik* incluem: BestBuy, Canon, ING, Panasonic, Qualcomm e Shell. (QLIK, 2015)

#### 3.1.1 DESCRIÇÃO E CARACTERÍSTICAS

As soluções *Qlik* são baseadas em dois principais produtos da empresa, o

### *QlikSense e o QlikView.*

O *QlikView* foi projetado para que profissionais da Tecnologia da Informação e analistas de negócios possam construir visualizações de dados e aplicativos de descobrimento de dados. Já no *QlikSense*, usuários finais podem criar suas próprias visualizações por meio de uma interface *drag-and-drop* (arrastar e largar). A Tabela 7 apresenta as principais diferenças entre essas ferramentas.

**Tabela 7:** QlikView x Qlik Sense.

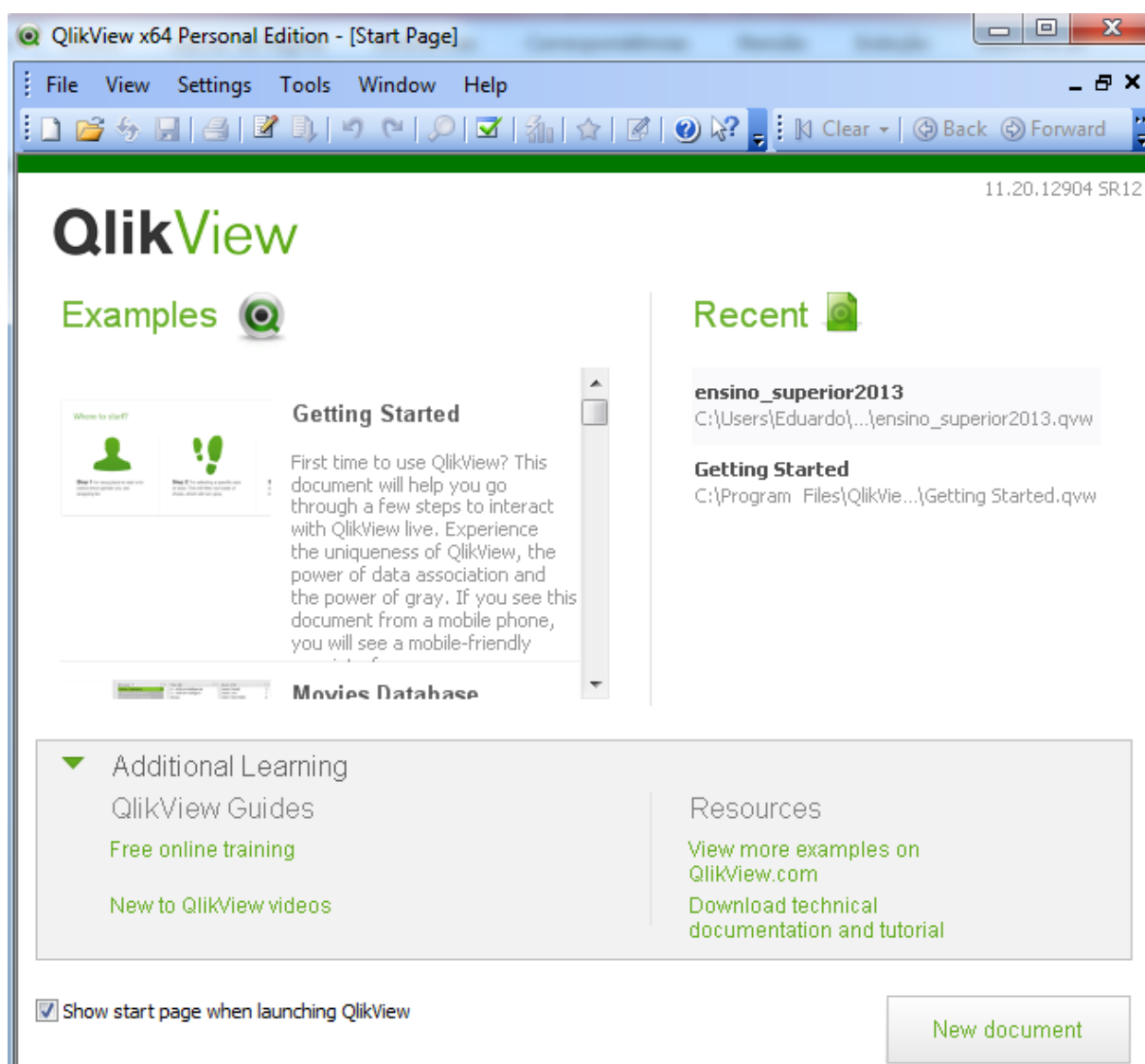
<b>QlikView</b>	<b>Qlik Sense</b>
Aplicativos normalmente criados pela TI ou técnicos em análise de negócios.	Visualizações criadas por analistas e usuários finais.
Controle de design granular.	Design móvel responsivo.
Robusta plataforma de desenvolvimento.	Experiência de usuário intuitiva com interface <i>drag-and-drop</i> .
Descobrimto guiado.	Criação de conteúdo de maneira colaborativa e progressiva.

Fonte: [www.qlik.com](http://www.qlik.com)

Para a análise da solução *Qlik* foi selecionado o *QlikView Personal Edition* na versão 11. A decisão de não utilização do *QlikSense* foi tomada já que esta é uma ferramenta nova no mercado e ainda limitada em suas funções de criação de relatórios em comparação com as outras ferramentas selecionadas para este trabalho. A licença utilizada foi a trial de 1 ano, que possui as mesmas características da versão completa.

A seguir serão apresentados alguns dos principais recursos da ferramenta *QlikView Personal Edition*.

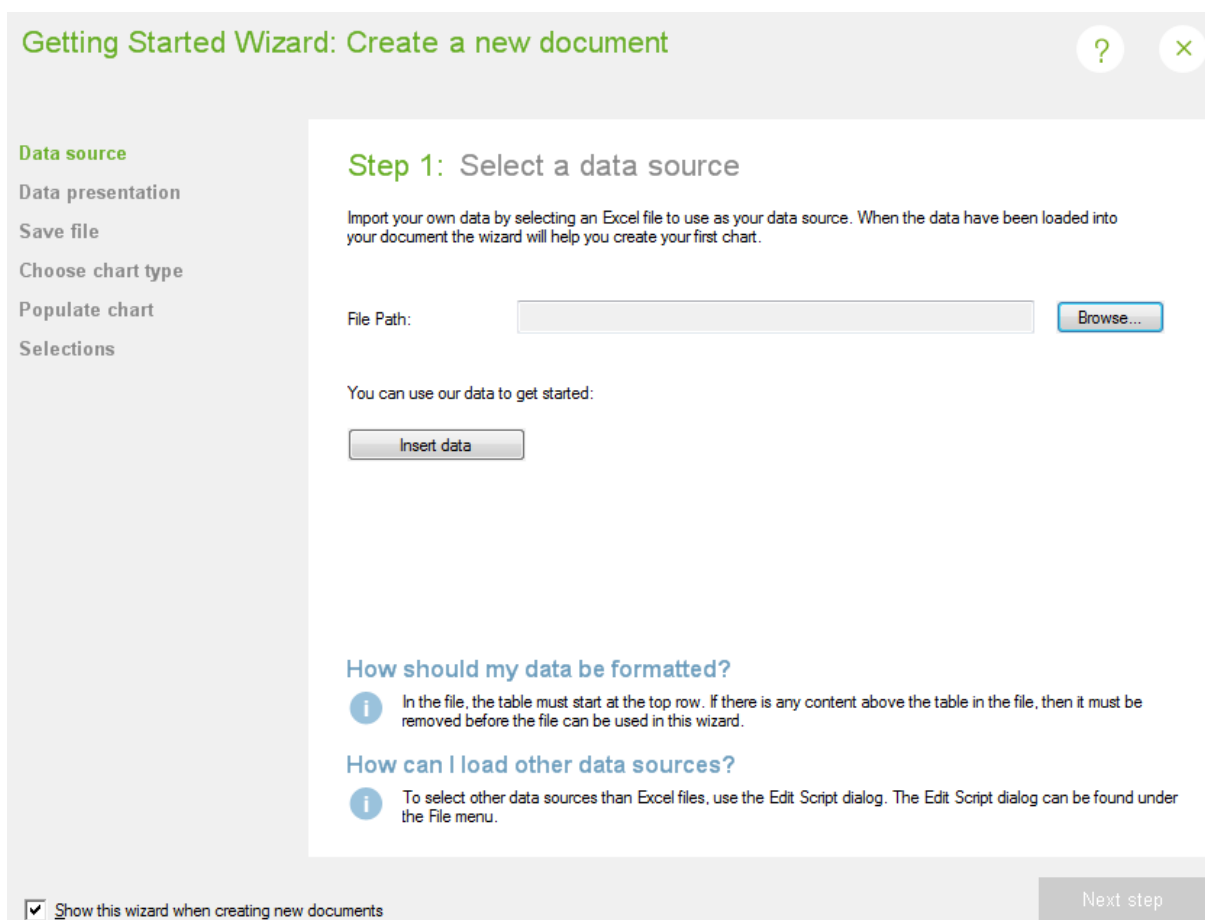
Na sua tela inicial o *QlikView* (Figura 12) apresenta alguns *links* para amostras de *dashboards* e tutoriais com instruções para a utilização da ferramenta.



**Figura 12** : Tela inicial do QlikView.

Ao selecionar a opção para criar um novo documento o *QlikView* inicia um processo de criação de relatório (Figura 13), porém apenas com a opção de importação de dados por meio de *flatfiles*<sup>5</sup>.

<sup>5</sup> Flatfile – Dados armazenados em formato de arquivos texto. Normalmente os registros são separados por linhas e as colunas separadas por um caractere delimitador como vírgula ou pipe (Fonte: <http://techterms.com/definition/flatfile>).



**Figura 13** : Criação de documentos no QlikView.

Para a utilização de dados a partir de SGDBs (Sistemas Gerenciadores de Banco de Dados) é necessário utilizar a função de edição de *script* do projeto atual. Na parte inferior da tela existe um atalho para selecionar uma conexão ODBC existente (Figura 14). Os dados são extraídos a partir de códigos SQL que devem ser escritos no *script*.

Após a execução do *script*, os dados extraídos passam a ficar disponíveis no atual projeto. O *QlikView* permite a criação de vários tipos de gráficos e filtros para a visualização dos dados (Figura 15).

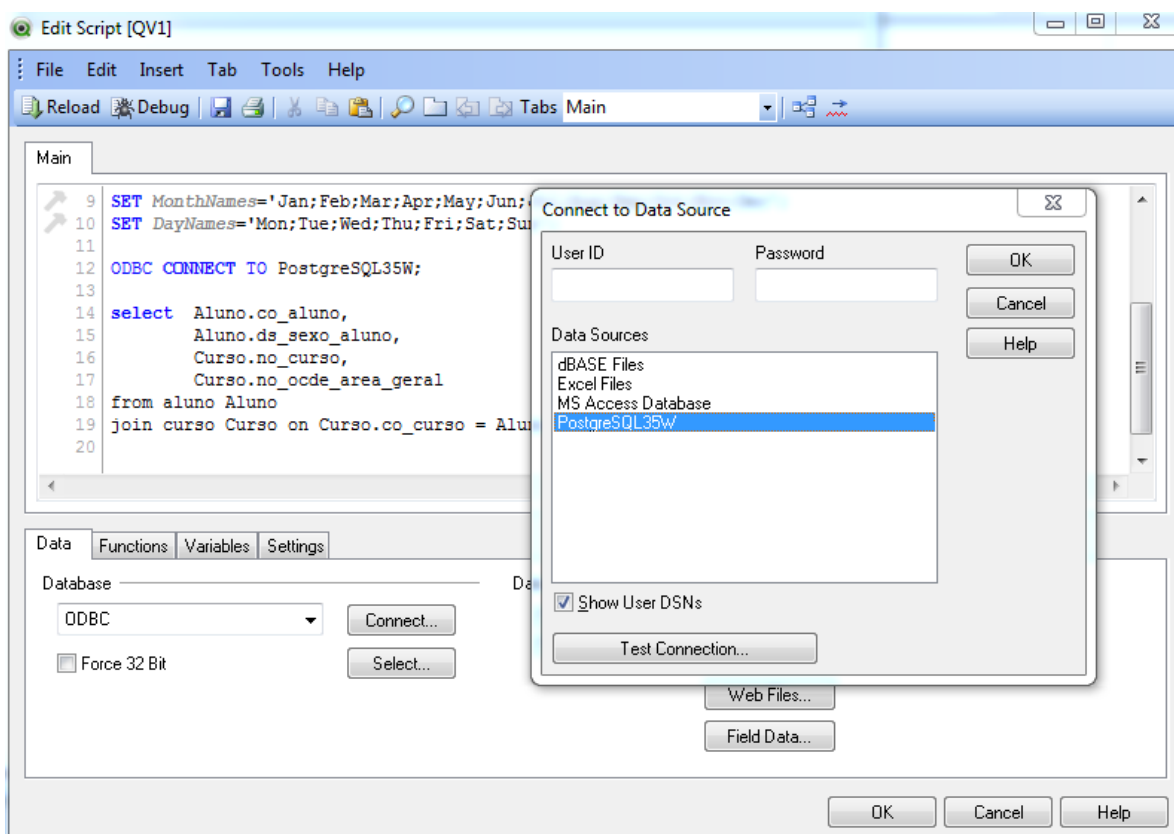


Figura 14 : Edição de script no QlikView.

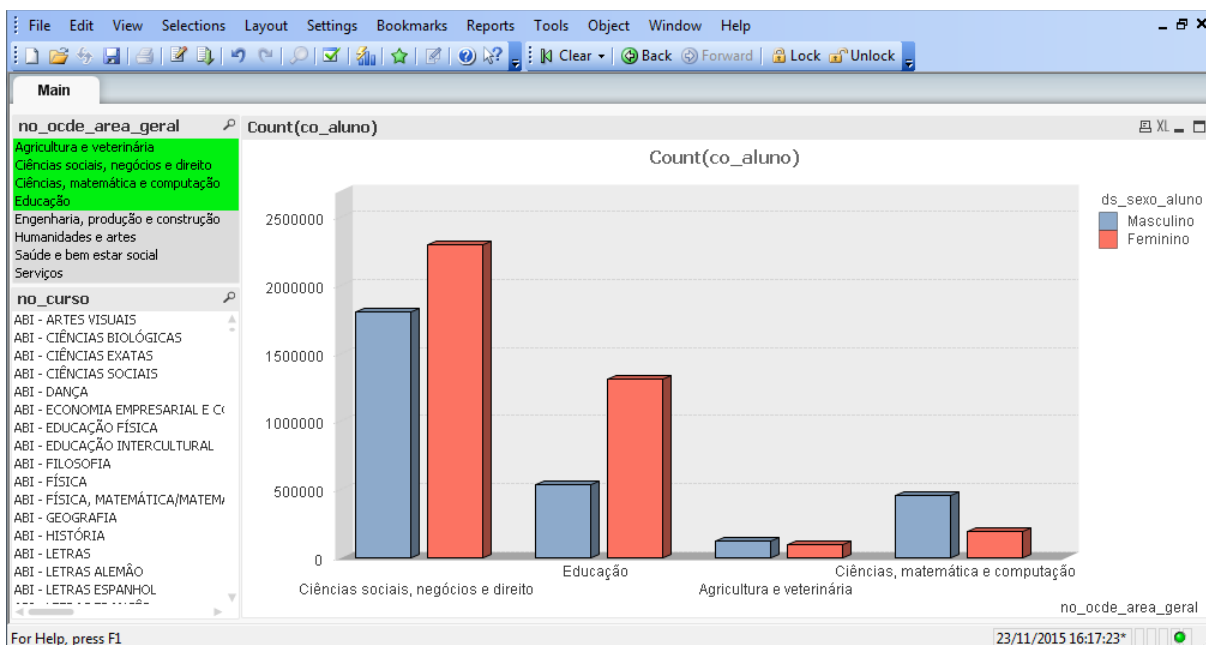


Figura 15 : Edição de dashboard no QlikView.

O *Gartner Group* (GARTNER, 2015) divulga em sua análise do quadrante mágico alguns pontos fortes e pontos de cautela referente aos fornecedores de

software de BI avaliados. Esses são os pontos identificados pelo Gartner referente ao Qlik:

a) Pontos Fortes das soluções Qlik:

- A *Qlik* oferece um portfólio de produtos de desenvolvimento de *dashboards* altamente interativo que se estende a serviços para usuários finais, aplicações para desenvolvimento de *dashboards* centralizados e necessidades de profissionais de TI para suportar a governança dos dados. A *Qlik* tomou uma decisão ousada em investir no *QlikSense*, uma plataforma completamente nova que visa elevar seus recursos corporativos. Mesmo sendo o *QlikSense* uma ferramenta nova no mercado com adoção ainda limitada, a *Qlik* tem a mais alta porcentagem de clientes relatando que usam a plataforma para governança dos dados.
- A facilidade de uso, particularmente entre os utilizadores de *dashboards*, é a razão chave dos quais seus clientes decidem pela aquisição de soluções *Qlik*, assim como seu baixo tempo e esforço de implementação. A *Qlik* possibilita que seus usuários conduzam uma gama mais ampla de análises que a maioria dos outros fornecedores analisados. Esta combinação tem sido o fator chave para seu sucesso. A *Qlik* oferece versões gratuitas para ambos os seus produtos, o que possibilita a seus usuários realizar testes livres de riscos. Consistente com sua utilização e razões para compra, a *Qlik* pontuou bem em *dashboards* analíticos, exploração em forma livre e capacidade *mobile* em suporte tanto a casos de uso descentralizados como em governança dos dados.
- Enquanto muitos aspectos referentes à experiência de usuário tiveram um declínio em relação as pesquisas dos anos anteriores, a capacitação de usuários e disponibilidade de informação são exceções. Clientes de referência avaliam a *Qlik* entre os 5 melhores para capacitação de usuário, o que inclui documentação, treinamento *online*, tutoriais, comunidades de usuários e conferências. Uma vantagem em particular para soluções *Qlik* é a *Qlik Community* ([community.qlik.com](http://community.qlik.com)) que oferece uma central de colaboração *online* e centro de excelência para clientes, parceiros e funcionários.

- Parceiros, incluindo integradores de sistemas, revendedores e agora fabricantes de equipamento original (OEMs) tem sido mais instrumentais na estratégia de crescimento da *Qlik* do que qualquer outro fornecedor de BI. A ênfase do *QlikSense* em APIs (Interface de Programação de Aplicações) abertas devem expandir as oportunidades para parceiros construírem aplicações de valor agregado para a plataforma.

b) Pontos de cautela da *Qlik*:

- A *Qlik* oferece dois produtos com dois modelos diferentes de preços. Enquanto esta estratégia expande as opções para seus clientes, também pode se tornar confusa quando clientes devem decidir se devem implementar o *QlikView*, o *QlikSense* ou ambos. Compradores de softwares de BI precisam investir tempo para entenderem as diferenças e avaliar quais produtos atendem as suas necessidades.
- *QlikSense* é um produto novo, em sua versão 1.0, e sua plataforma contém vários diferenciadores em relação ao *QlikView* e ao resto do mercado, mas também é um produto cujo seu desenvolvimento em várias áreas funcionais ainda está em andamento. Isto irá resultar em uma adoção limitada do *QlikSense*, enquanto a *Qlik* aprimora suas capacidades de suportar exploração interativa para o analista de negócios e estende sua capacidade de gerar *dashboards* avançados. O foco em amadurecimento da plataforma *QlikSense* pode limitar a capacidade da *Qlik* em investir em novas capacidades e em diferenciar suas plataformas de modo a manter sua competitividade no futuro.
- A *Qlik* teve uma execução de vendas inconsistente nos últimos anos o que é potencialmente explicado por sua mudança de vendas em nível departamental para implementações focadas em suportar as necessidades estratégicas das empresas. Esta transição afetou a forma como clientes veem suas experiências em vendas com a *Qlik* o que foi avaliada entre as 5 piores nesse quesito. A percepção de experiência do cliente também teve um declínio em relação ao ano anterior o que foi dirigido primariamente por

baixas avaliações em suporte ao cliente. Experiência do cliente é algo importante na visão de compradores e uma transição bem sucedida será crucial para o aceleração da *Qlik* na penetração de mercado.

- A *Qlik* foi uma das últimas empresas classificadas como líderes a oferecer todos os recursos de computação em nuvem. As funções de compartilhamento na nuvem do *QlikSense* representa um primeiro passo, mas clientes da *Qlik* estão entre os 3 primeiros a não terem planos para adotar recursos em computação na nuvem para suas soluções de BI.

### 3.1.2 INVESTIMENTO

A *Qlik* divulga em seu site (<http://www.qlik.com/pricing>) as seguintes informações referentes a preços do *QlikSense* e do *QlikView*.

- ***QlikSense***

*Qlik Sense Desktop*: Gratuito

*Qlik Sense Cloud*: Gratuito até 5 usuários

*Qlik Sense Enterprise*: US\$1.500,00 por *token*<sup>6</sup>. Sendo que 1 *token* permite acesso irrestrito para 1 usuário ou 10 seções de acesso que permite a utilização por 1 hora. A opção de seções de acesso pode ser compartilhada entre vários usuários.

- ***QlikView***

*QlikView Personal Edition*: Gratuito

*QlikView Enterprise*: não divulgado

A *Qlik* não divulga publicamente em seu site o custo de licenças referentes ao *QlikView Enterprise*, porém essa informação pode ser obtida em alguns sites de pesquisas em BI como o [apandre.wordpress.com](http://apandre.wordpress.com) ou [rapid-bi.com](http://rapid-bi.com) (Tabela 8):

---

<sup>6</sup> Token – Modelo de licenciamento de software em que clientes podem adquirir tokens que podem ser trocados por licenças de um determinado portfolio de software.

**Tabela 8:** Custos do QlikView. Fonte: apandre.wordpress.com

<b>QlikView Client</b>	
Usuário único	US\$ 1.395,00 por usuário
Licença de documento	US\$ 360,00 por usuário por documento
Usuário compartilhado	US\$ 15.495,00 por usuário compartilhado
<b>QlikView Server</b>	
Enterprise Edition Server	US\$ 36,150 por servidor
Small Business Edition Server (até 25 usuários)	US\$ 8,675 por servidor

O cálculo de uma implementação de *QlikView* pode ser tornar complexo visto que a *Qlik* exige a compra de licenças de documento para cada documento disponível no servidor e existem também muitas outras variantes que podem afetar o preço final. Pandre (2012) faz uma estimativa baseada nos valores acima para uma implementação com 46 usuários, 3 clientes *desktop*, 10 documentos/visualizações disponíveis para 10 usuários diferentes, 1 servidor de aplicação e manutenção por 3 anos o que ficaria em torno de US\$118.000,00.

### 3.2 A FERRAMENTA *MICROSTRATEGY*

A *MicroStrategy* foi fundada em 1989 por Michael J. Saylor com um contrato de consultoria com a *DuPont*, multinacional do setor de química, para se fazer previsões em tendências de vendas de seus produtos. Saylor fundou a *MicroStrategy* junto com Sanju K. Bansal que foi seu colega no MIT.

Inicialmente fundada para prestar serviços de consultoria, a *Microstrategy* criou seu primeiro *software* em 1991. Um ano depois, a empresa fechou seu primeiro contrato com um grande cliente, após a *DuPont*, quando fechou um contrato de 10 milhões de dólares com a *McDonald's* para criar aplicações para analisar a efetividade de suas campanhas promocionais.

Foi a partir de 1994 que a *MicroStrategy* iniciou a construção de seu produto de auxílio de suporte a decisão (DSS), já que até então se construía *softwares* customizados para cada cliente (GLASSER, 1996).

Em seu *website* ([microstrategy.com](http://microstrategy.com)), a *MicroStrategy* mostra a história de seu principal produto de BI, o *Microstrategy Analytics*, que teve suas primeiras versões

baseadas no ROLAP<sup>7</sup>, introduzindo uma interface web com alertas de *email*. No ano de 2000 após uma completa revisão do *Analytics* foi lançado o *Microstrategy 7*, uma plataforma de BI com melhorias em sua escalabilidade e capacidade de alerta e que foi eleita pela revista *PC Magazine* a escolha dos editores em plataforma de BI. A versão 8 do *Microstrategy* foi lançada em 2005 com uma interface web atualizada permitindo maior interatividade para criação de relatórios e *dashboards*. Essa versão também teve aperfeiçoamentos em suas capacidades analíticas, *data mining* e suporte mais amplo à fontes de dados. Em 2009 foi lançado a versão 9 que foi considerada pela empresa a *release* mais significativa da última década. Essa versão estendeu o desempenho e escalabilidade do *Business Intelligence (BI)* empresarial e simplificou o desenvolvimento de aplicações de BI departamental. Foi em 2010 que a *MicroStrategy* lançou a integração do *Analytics* com dispositivos móveis, como o *iPhone* e o *iPad*, por meio de sua nova plataforma intitulada *MicroStrategy Mobile* que foi reconhecida pela *Apple* como um dos melhores aplicativos empresariais do ano feito para iOS. No ano seguinte foi lançado o serviço *MicroStrategy Cloud* que combinou recursos de BI do *Analytics* com capacidades de computação em nuvem.

Desde 2013 a solução de BI do *MicroStrategy* está posicionada como líder pelo já mencionado relatório “Quadrante Mágico” elaborado pelo *Gartner Group*. Clientes notáveis da *MicroStrategy* incluem *Facebook*, *Linkedin*, *Netflix*, *Yahoo*, *Ebay*, *Carrefour*, *Adobe* e *Pfizer*.

### 3.2.1 DESCRIÇÃO E CARACTERÍSTICAS

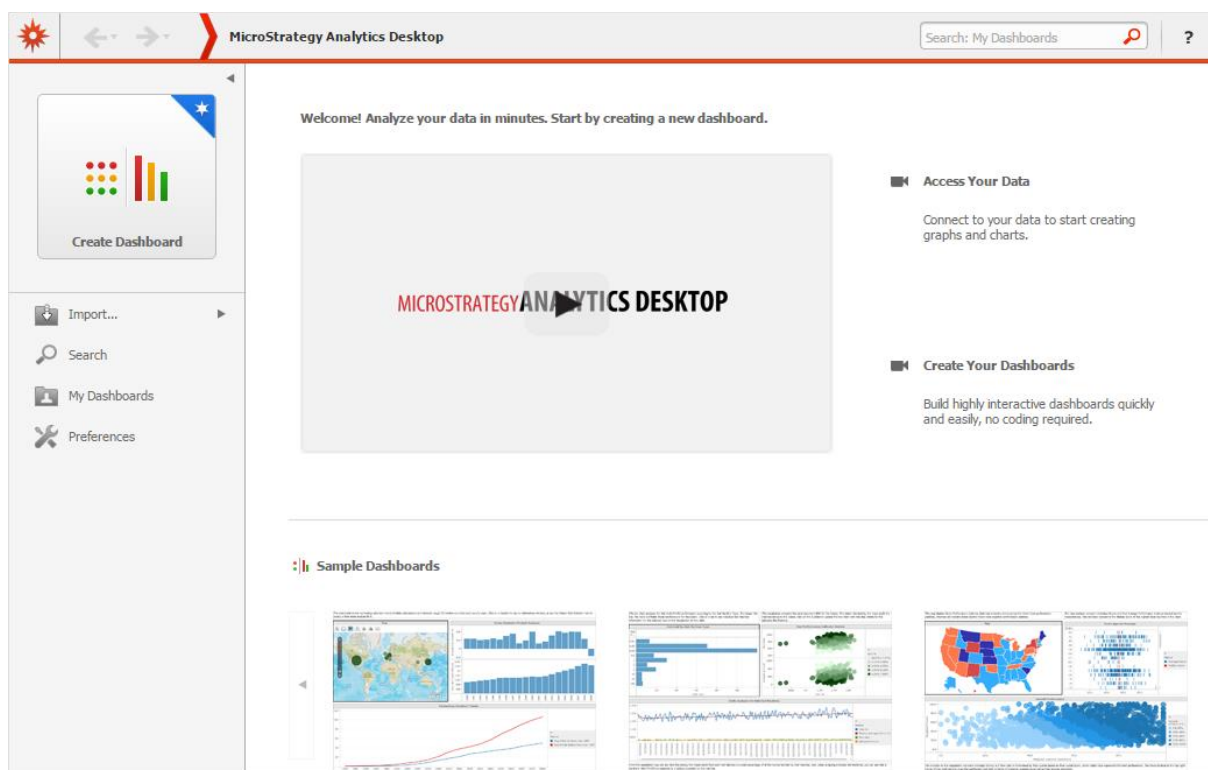
O principal *software* de BI da *Microstrategy* é intitulado *MicroStrategy Analytics* e atualmente está na sua versão 10. O software permite a análise e visualização de dados de uma grande variedade de fontes e é oferecido em 3 versões:

---

<sup>7</sup> ROLAP ou Relational OLAP é a abordagem para se desenvolver sistemas dimensionais baseados em dados armazenados em tabelas de banco de dados relacionais. (PAREDES, 2009, p. 12)

- ***Analytics Desktop***: é uma versão gratuita com suporte a apenas um usuário. Essa versão traz a possibilidade de se explorar dados e criar relatórios e *dashboards* sem necessidade de conhecimento de *scripts* e programação. A *Microstrategy* oferece suporte gratuito por e-mail e telefone e também vídeos para treinamento *online*, fóruns de discussão e documentos de guias para sua utilização. Essa versão está disponível apenas para o sistema operacional Windows.
- ***Analytics Express***: é um serviço que combina as funcionalidades oferecidas pelo *Analytics Desktop* com capacidades de computação em nuvem. Toda a aplicação é acessada online sem a necessidade de instalação na máquina do usuário. Esta versão é gratuita por um ano e pode compartilhar projetos entre múltiplos usuários.
- ***Analytics Enterprise***: é a versão mais completa oferecida pelo *MicroStrategy* e pode ser adquirida por meio de licenças tendo uma versão trial que pode ser utilizada por 30 dias. Além das funcionalidades oferecidas por outras versões, o *Enterprise* tem maior integração com fontes de dados e sistemas operacionais. Suas ferramentas analíticas incluem uma versão aprimorada da solução para criação de relatórios e *dashboards*, integração com dispositivos móveis baseados em iOS e *Android*, integração com o *Microsoft Office* e ferramentas de desenvolvimento.

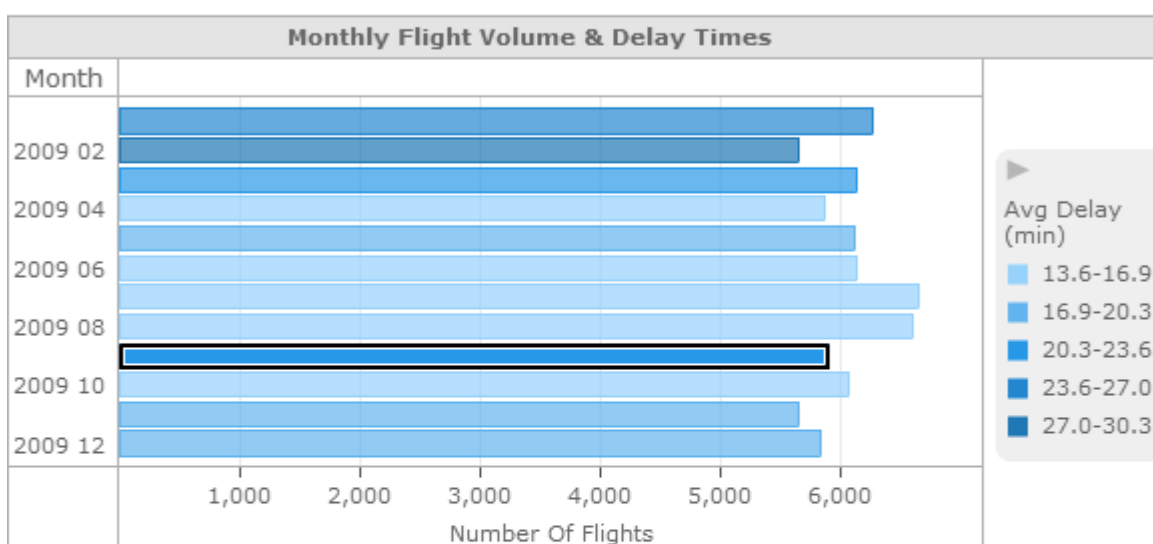
Para a avaliação da ferramenta foi utilizado a versão *Analytics Desktop* sob a licença trial de 1 (um) ano, possuindo as mesmas características da ferramenta completa. Após a instalação, a utilização é feita a partir de uma interface web, conforme mostra a Figura 16.



**Figura 16** : Tela inicial do Microstrategy Analytics Desktop.

Já na tela inicial é apresentado algumas amostras (*exemplos*) de *dashboards* já incluídos na ferramenta. A Figura 17 exibe uma amostra de uma análise de dados sobre atrasos de uma companhia aérea.

This bar chart displays information on the Average Delay (in minutes) experienced in Hawaiian Airlines flights - the darker the color, the longer the delay. By clicking on a bar (each corresponding to a Month) the information on the two other visualizations gets updated, showing details of the flights during those months.



This grid displays a quick summary of the number of flights flown by Hawaiian Airlines by day of the week, and an overall score based on their Average on Time Percentage and their Delay (in minutes)

**Day of the Week Stats**

Day Of Week	Number Of Flights	Avg On-time (%)	Avg Delay (min)
Sun	785	70%	14.6
Mon	783	59%	35.8
Tue	971	57%	13.4
Wed	953	59%	14.8
Thu	780	61%	40.8
Fri	824	56%	15.6
Sat	808	64%	14.0

**Figura 17** : Amostra de dashboard no Microstrategy.

Para a criação de um novo *dashboard*, primeiramente, é apresentado as opções de importações de dados. Nessa tela é possível selecionar se os dados serão importados a partir de *flatfiles* ou conexão com SGBD, conforme Figura 18.

The screenshot shows the MicroStrategy Analytics Desktop interface. At the top, there is a navigation bar with a home icon, back and forward arrows, and a plus icon, followed by the text "MicroStrategy Analytics Desktop". Below this is a header area with "Select your data source" on the left and "My datasets" on the right. The main area is divided into two columns. The left column lists four data source options: "File" (Import an Excel, CSV, or text file), "Database" (Import from relational data sources), "Salesforce" (Import your reports from Salesforce.com), and "Freeform" (Import data using a script: SQL, XQuery, SOQL, or HiveQL). The right column, titled "My datasets", shows a summary "You are currently using 426.99MB of 5.0GB" and a table of existing datasets.

Name	Size (KB)
Cube_localoferta	5957
Cubo instituicao	1124
Hawaiian Airlines Data	269
New Cube	10137
New Cube2	1307

**Figura 18** : Processo de criação de dashboard no Microstrategy Analytics Desktop – seleção de fonte de dados.

Ao selecionar a opção de utilização de SGDBs (*Database*) é exibida uma lista de conexões já existentes e a possibilidade de criação de novas conexões (Figura 19). Nessa mesma tela podemos selecionar quais tabelas e colunas serão utilizadas no *dashboard*. É possível incluir todos os dados das tabelas selecionadas ou ainda fazer uma pesquisa utilizando códigos SQL.

A análise dos dados é feita a partir da tela de *dashboards*. Ali são exibidos os dados selecionados e as várias opções de visualização dos dados como: tabelas, gráficos e mapas. A Figura 20 apresenta um exemplo de visualização de dados no formato de gráfico.

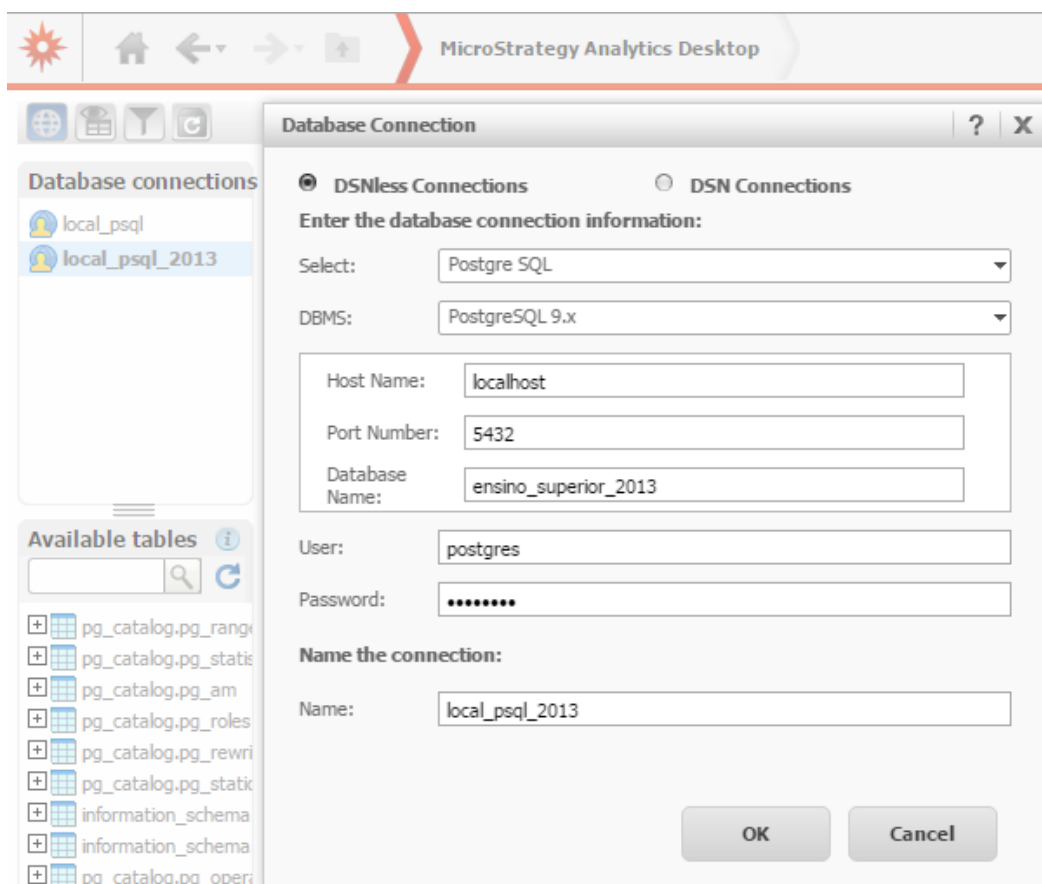


Figura 19 : Utilizando conexões SGBD no Microstrategy Analytics Desktop.

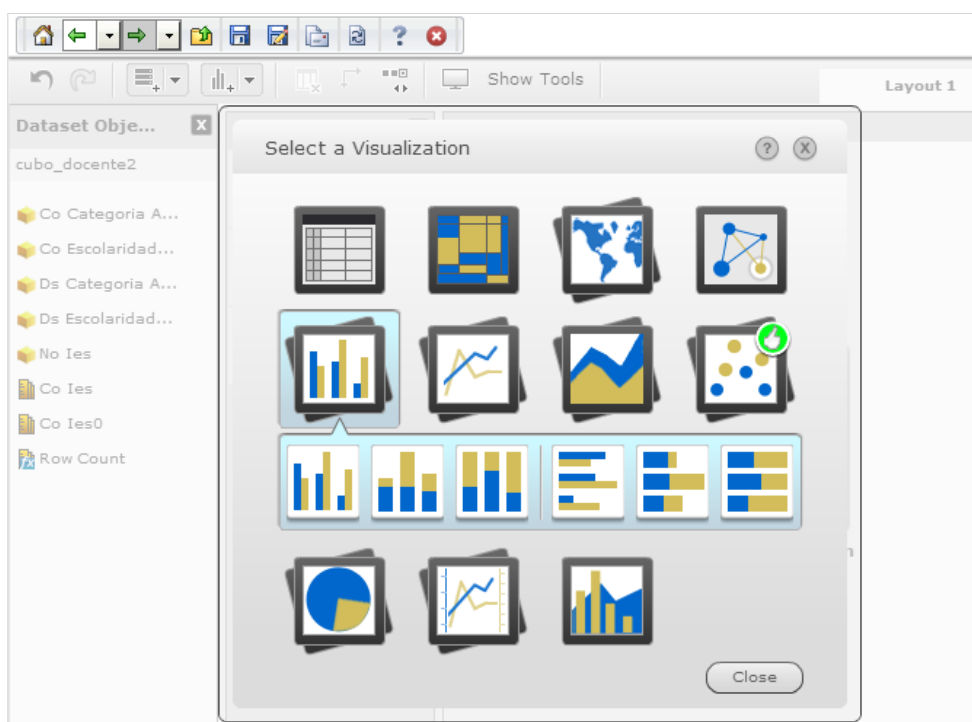


Figura 20 : Criação de dashboards no Microstrategy Analytics Desktop.

A seguir serão apresentados os pontos fortes e pontos de cautela identificados pelo Gartner group, em sua análise do quadrante mágico (GARTNER, 2015), para o *Microstrategy*.

a) Pontos Fortes do *Microstrategy*:

- O *Microstrategy* é uma plataforma de nível empresarial com uma forte visão em descobrimento de dados se tornando uma ferramenta adequada para empresas que necessitam de uma larga escala de relatórios, *dashboards* e utilização *mobile*. O *software* também não é atado ao um pacote de aplicações corporativas como Oracle ou SAP.
- O *Microstrategy* continua sendo uma referência para a indústria de grandes implementações de BI a partir de grandes *data warehouses* corporativos e é normalmente visto como o fornecedor a ser escolhido quando os requisitos são complexos. A empresa investe continuamente em escalabilidade, performance e também em melhorias relacionadas com *big-data*<sup>8</sup> para seu portfólio. Por isso, o *Microstrategy* é tipicamente implementado em grandes empresas que o consideram ser o padrão em BI empresarial mais frequente do que outros fornecedores.
- É o único fornecedor participante da pesquisa do quadrante mágico, o qual seus clientes reportaram como sendo a funcionalidade *mobile* seu principal motivo de escolha da ferramenta. Os dados mostram que 51% dos seus clientes já implementaram ou conduziram pilotos em soluções *mobile* sendo que a média de outras ferramentas é de 30%. Computação em nuvem também tem sido uma área de agressivo investimento estratégico por parte da *MicroStrategy* em comparação com outros fornecedores de BI.
- Clientes valorizam não só a simplificação da experiência de usuário, mas também a simplificação do seu relacionamento com os fornecedores. A

---

<sup>8</sup> *Big-data* ou megadados é um termo amplo referente a um conjunto de dados tão grandes ou complexos que necessitam de métodos avançados de tratamento e análise de dados visto que aplicações tradicionais de processamento de dados são inadequados (Fonte: [http://techterms.com/definition/big\\_data](http://techterms.com/definition/big_data)).

*MicroStrategy* tomou passos ousados na simplificação dos preços cobrados por suas soluções, deixando as informações referentes a valores públicas em seu *website*. Anteriormente a complexidade referente aos preços cobrados era uma reclamação constante de seus clientes fazendo com que essa mudança na política de valores cobrados tenha sido um passo importante de melhoramento da percepção geral da empresa.

b) Pontos de cautela do *Microstrategy*:

- 2014 foi um ano desafiador para a *Microstrategy* devido a um declínio momentâneo de suas vendas, mudanças no quadro executivo e no corte de 20% em seu quadro de funcionários o que resultou em rupturas em suas funções. Clientes do *Microstrategy* relataram uma experiência de cliente abaixo da média referente a suporte, qualidade do produto, migração e benefícios para o negócio. Porém a empresa está passando por um processo de reagrupamento e tem feito mudanças positivas para aprimoramento no futuro. A *MicroStrategy* deve superar um problema relacionado à sua marca no mercado que continua sendo associada ao “BI tradicional” mesmo que o seu portfólio de produtos e suas estratégias estejam focadas em importantes requisitos de governança de dados em grandes implementações.
- Embora o desenvolvimento do ambiente do *MicroStrategy* seja robusto e flexível, existe uma íngreme curva de aprendizado, mesmo para desenvolvedores experientes, para se construir qualquer nível de complexidade analítica em relatórios parametrizados que simulam análises *ad hoc* e *dashboards* interativos para usuários finais. Mesmo com melhorias na sua usabilidade, seus clientes continuam a avaliar a plataforma como sendo abaixo da média em relação a facilidade de desenvolvimento, facilidade de uso para usuário final e facilidade de administração e implementação.

Pelos testes realizados o *Microstrategy Analytics* se mostrou uma ferramenta de utilização bem intuitiva. Relatórios e visualizações são criados rapidamente mesmo durante sua primeira utilização. O processo de seleção de dados em SGBDs e *flatfiles* e a criação de *dashboards* são feitos de maneira fácil, mostrando que

mesmo usuários finais sem conhecimento técnico na área da Tecnologia da Informação potencialmente poderiam criar seus próprios relatórios e visualizações por meio da ferramenta.

### 3.2.2 INVESTIMENTO

Os valores para a obtenção de licenças do *MicroStrategy* estão disponíveis no *website* da empresa (<http://www.microstrategy.com/us/experience>). As licenças são divididas em 4 pacotes de aplicativos.

- *MicroStrategy Architect*: engloba todas as ferramentas de desenvolvimento. O valor da licença é aproximadamente US\$ 5.000,00 por usuário.
- *MicroStrategy Server*: contém todos os componentes de infraestrutura assim como ferramentas de administração e monitoramento. As licenças custam em média US\$ 1.200,00 por usuário ou US\$600.000,00 por núcleo de CPU. De acordo com a empresa, um núcleo de CPU pode suportar de 500 a 1000 usuários.
- *MicroStrategy Web Interface*: interface gráfica para usuários com o custo de aproximadamente US\$600,00 por usuário ou US\$300.000 por núcleo de CPU.
- *MicroStrategy Mobile*: aplicativos móveis. Os valores das licenças são em média US\$600,00 por usuário ou US\$300.000 por núcleo de CPU.

A *Microstrategy* recentemente simplificou os preços estabelecidos para os seus produtos, nas 4 categorias acima citadas (PCWORLD, 2014), o que facilitou o cálculo por parte dos compradores do investimento necessário. Podemos concluir que uma implementação com 3 licenças para desenvolvedores, um servidor para infraestrutura e licenças de interface web para 45 usuários finais ficaria em torno de US\$96.000,00. Caso se utilize aplicativos móveis, podemos adicionar US\$600,00 por usuário dessa plataforma.

### 3.3 A FERRAMENTA *MICROSOFT POWER BI*

A primeira versão do *Power BI* foi lançada pela *Microsoft* em Fevereiro de 2014 como parte do pacote de aplicativos *Office 365*. Essa primeira versão foi lançada como um serviço baseado em nuvem o que simplificava a sua implementação e manutenção. Além de fazer parte do *Office 365*, o *Power BI* ainda dependia da utilização do *Excel* para a análise de dados. De acordo com a análise da Gartner, essa limitação de necessidade de adoção dos componentes mais atualizados do *Office 365* fez com que sua adoção por parte das empresas fosse limitada.

Foi a partir da nova *release* do *Power BI*, em julho de 2015, que a ferramenta foi oferecida como uma solução independente do *Excel*. Essa nova versão foi intitulada *Power BI Desktop* e melhorou vários aspectos da antiga versão, incluindo a capacidade de preparação de dados, descobrimento de dados e criação de *dashboards*. Várias limitações da primeira versão também foram corrigidas como habilitação do acesso a dados armazenados no SSAS (*SQL Server Analysis Services*) e uma ferramenta de *data mining* da *Microsoft* utilizada no sistema de gerenciamento de banco de dados *SQL Server*. Essas funcionalidades foram evoluindo com o tempo por meio de atualizações mensais oferecidas pela *Microsoft*.

#### 3.3.1 DESCRIÇÃO E CARACTERÍSTICAS

O *Power BI Desktop* (Figura 21) é oferecido pela *Microsoft* apenas em versões para *Windows* e sua versão gratuita pode ser adquirida pelo site [powerbi.microsoft.com](http://powerbi.microsoft.com). A *Microsoft* também disponibiliza o aplicativo móvel intitulado *Power BI Mobile*, que conta com versões para *Android*, *iOS* e *Windows*. Com *Power BI Mobile* é possível visualizar gráficos e *dashboards* criados no *Power BI Desktop* em dispositivos móveis como celulares ou *tablets*. Foi utilizada a versão gratuita para os testes propostos.

A ferramenta é bem intuitiva, em sua tela inicial são exibidas opções de importação de dados, relatórios recentes e alguns vídeos de tutoriais. Ao escolher a opção de importação de dados (Figura 22), o *Power BI* exibe uma grande variedade de opções que incluem *flatfiles*, arquivos *CSV* ou *Excel*, vários sistemas gerenciadores de bancos de dados e também dados do *Azure* – uma plataforma da *Microsoft* de serviços de computação na nuvem.

Para os testes propostos neste trabalho foi selecionada a opção conexão ODBC, que foi previamente criada para acesso aos dados no *PostgreSQL*.

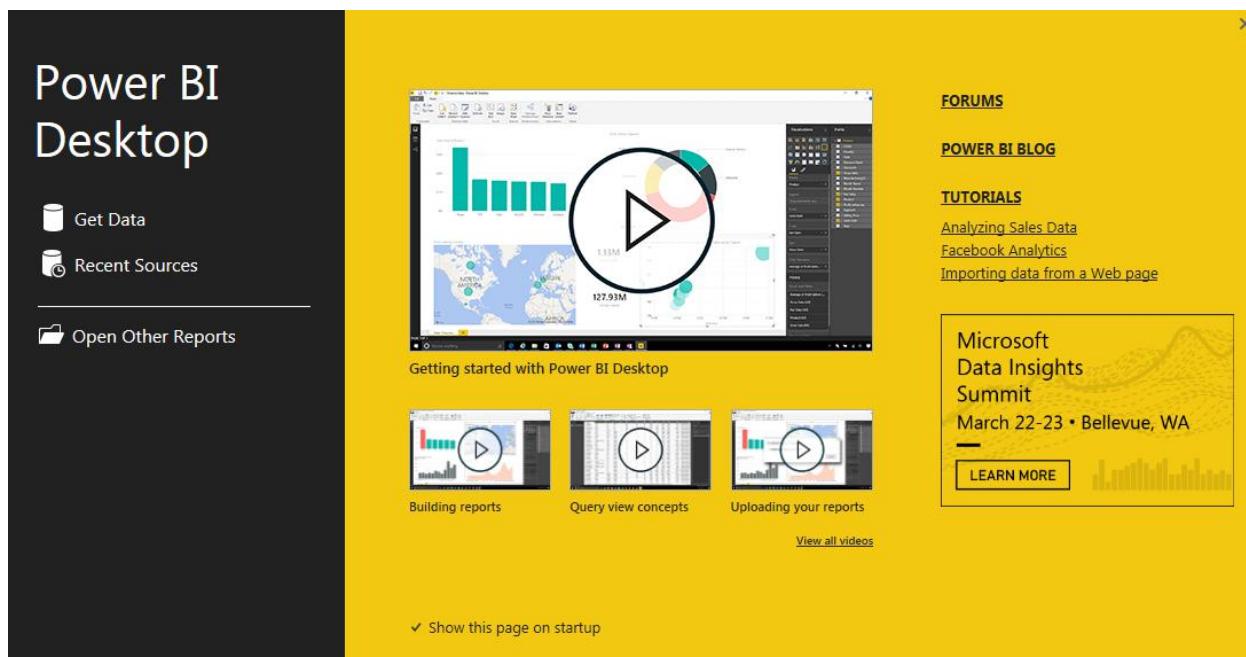


Figura 21 : Tela inicial do Power BI Desktop.

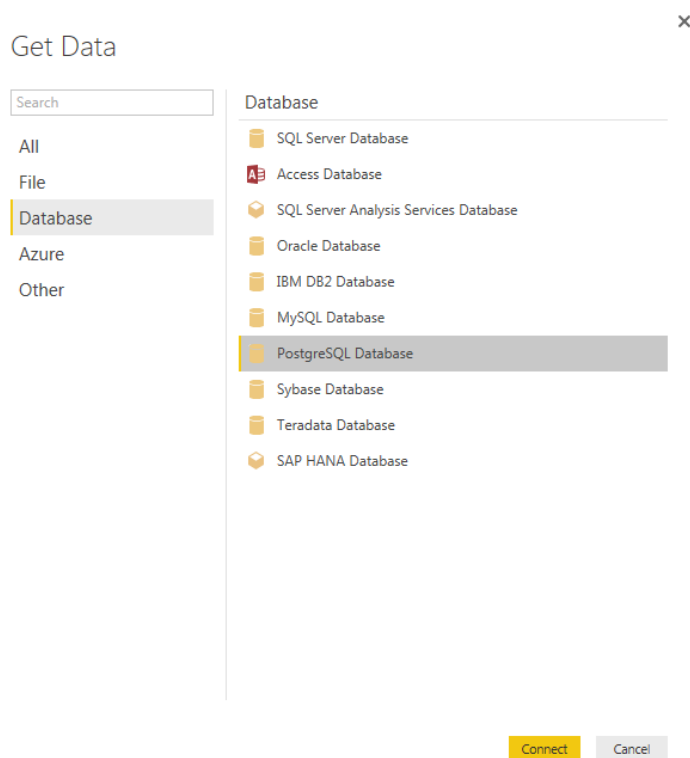


Figura 22 : Tela de seleção de dados no Power BI.

Após a seleção de fonte de dados, pode-se escolher quais tabelas carregar na ferramenta e opcionalmente adicionar um código em SQL (Figura 23) para restringir a busca, diminuindo assim o tempo de carregamento dos dados.

## From ODBC

Import data from ODBC.

Data Source Name (DSN)

PostgreSQL35W

Advanced options

Connection string (non-credential properties) (optional) ⓘ

SQL statement (optional)

```
select Aluno.co_aluno,
       Aluno.ds_sexo_aluno,
       Curso.no_curso,
       Curso.no_ocde_area_geral
from public.aluno Aluno
join public.Curso Curso on Curso.co_curso = Aluno.co_curso
```

OK

Cancel

**Figura 23** : Carregando dados do PostgreSQL no Power BI utilizando código em SQL.

Uma vez carregados os dados na ferramenta, existe a possibilidade de transformação desses dados utilizando a opção *Query Editor* (Figura 24). Essa função oferece várias funções de preparação de dados como divisão e agrupamento de colunas, criação de colunas calculadas, aplicação de filtros e até mesmo construção de relacionamentos entre tabelas.

A tela inicial permite a criação de visualizações de maneira intuitiva com funções de drag-drop. Várias formas de visualizações estão disponíveis no menu lateral e as colunas podem ser selecionadas conforme a Figura 25.

The screenshot shows the Power BI Query Editor interface. The top ribbon includes tabs for Home, Transform, Add Column, and View. The main area displays a table with the following data:

co_docente	co_escolaridade_docente	ds_escolaridade_docente	co_ies	no_ies
118987880569	5	Doutorado	953	UNIVERSIDADE METROPOLITANA DE SANTOS
117655	5	Doutorado	2183	CENTRO UNIVERSITÁRIO FUNDAÇÃO SANTO A
243198	5	Doutorado	573	UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
24317	5	Doutorado	1142	FACULDADES INTEGRADAS DE VITÓRIA
118987880574	5	Doutorado	3164	INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA I
243138	5	Doutorado	573	UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
63301	5	Doutorado	75	FACULDADE ESTADUAL DE EDUCAÇÃO CIÊNCI
63301	5	Doutorado	10	PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PAR
112556	5	Doutorado	582	UNIVERSIDADE FEDERAL DE SANTA MARIA
243110	5	Doutorado	3	UNIVERSIDADE DE BRASÍLIA

Figura 24 : Transformação de dados utilizando o Query Editor do Power BI.

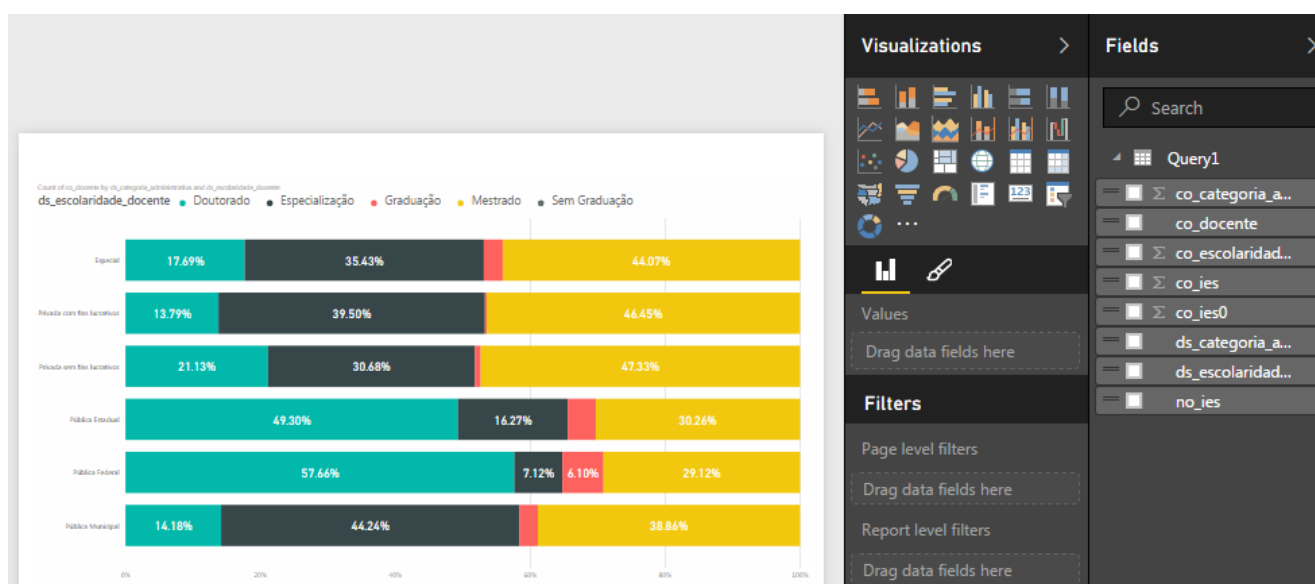


Figura 25 : Criação de gráficos no Power BI Desktop.

O *Power BI Desktop* contém todas as ferramentas necessárias para acesso a dados, preparação dos dados e escolha de campos e visualizações. Mas a visualização do *dashboard* ou relatório pronto com operações de *drill*<sup>9</sup>, aplicação de filtros e exportação para PDF ou *Excel* está disponível apenas no Power BI, que é

<sup>9</sup> Drill – operação para buscar o detalhamento das informações de alguma hierarquia. Por exemplo, em determinado relatório ao clicar no atributo “Região” (sul, norte, nordeste) o relatório é “aberto” para uma visão de Estado, assim permite ao usuário variar a granularidade dos dados.

um serviço na nuvem hospedado pela própria *Microsoft*. Para a publicação de relatórios no serviço *online* do *Power BI* (Figura 26) é necessário ter uma conta que pode ser cadastrada gratuitamente no site do *Power BI* - <https://app.powerbi.com>. Porém, existe uma limitação ao se realizar esse cadastro. A *Microsoft* não aceita a criação de contas com endereços de e-mail não comerciais ou governamentais. Dessa forma contas de e-mail de serviços como *Hotmail*, *Yahoo* ou *Gmail* não podem ser utilizados.

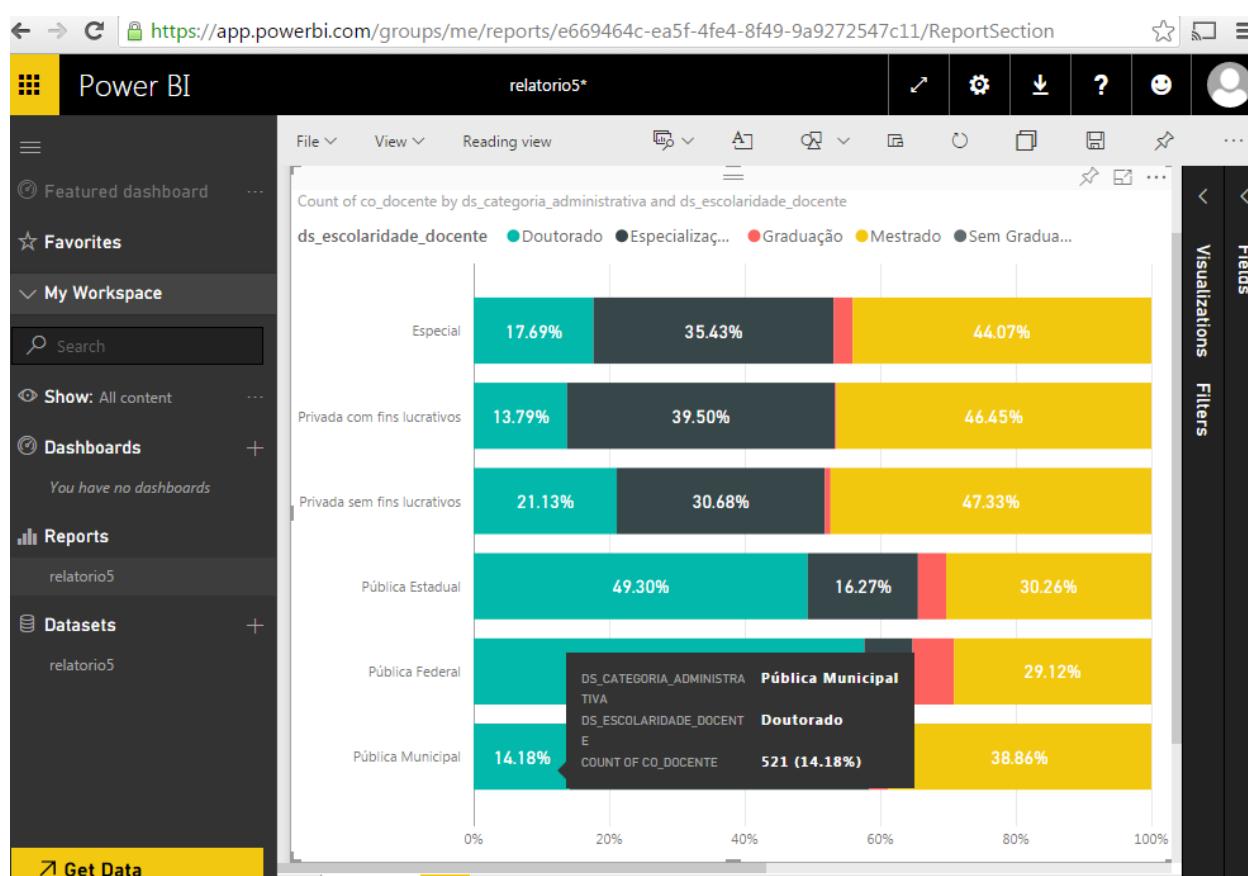


Figura 26 : Visualização de gráfico publicado no Power BI online.

A análise do quadrante mágico realizada pelo *Gartner Group* (GARTNER, 2015) aponta o *Power BI* como uma das soluções de preço mais baixo disponíveis no mercado, visto a diminuição do preço original de 39,95 USD por usuário por mês para 9,95 USD por usuário por mês. Com o preço reduzido e várias melhorias recentes, a *Microsoft* anunciou um aumento de 90.000 organizações em sua base de usuários. Esta análise também menciona os 3 principais motivos para os usuários optarem pelo *Power BI* que são seu baixo preço por usuário, facilidade de utilização por usuários de negócios e disponibilização de recursos especializados. Outro ponto

positivo segundo a análise do *Gartner Group* é a capacitação de usuário visto que a *Microsoft* disponibiliza tutoriais *online*, suporte por meio de fóruns, documentações e conferências, sendo um dos melhores fornecedores nesse aspecto.

Alguns pontos de cautela citados pelo *Gartner Group* incluem a dificuldade de implementação com grande número de usuários, já que o *Power BI* em sua versão 2.x é uma ferramenta nova no mercado. Também foi apontado a inabilidade da ferramenta de realizar algumas operações de análise avançada, como por exemplo, operação de previsão, que necessita ser realizado externamente pelo *Excel*.

### 3.3.2 INVESTIMENTO

A *Microsoft* distribui o *Power BI Desktop* em 2 versões. Uma versão gratuita com algumas limitações de quantidade de dados e outros recursos e a versão Pro, mais completa e que pode ser adquirida por meio de licenças. A Tabela 9, disponibilizada pela *Microsoft*, contém todos os detalhes das duas versões.

Usando como exemplo uma implementação com 45 usuários da versão Pro do *Power BI*, podemos concluir que o custo de licenciamento seria aproximadamente US\$ 5.400,00 por ano.

**Tabela 9:** Informações das duas versões do *Power BI Desktop*.

	<b>Power BI</b>	<b>Power BI PRO</b>
Valor	Gratuito	9,99 USD / mês / usuário
Limite de capacidade de dados.	1 GB / usuário	10 GB / usuário
Criar, exibir e compartilhar relatórios e painéis pessoais com outros usuários do <i>Power BI</i> .	Sim	Sim
Acesso de <i>dashboards</i> em dispositivos móveis utilizando aplicativos nativos para <i>iOS</i> , <i>Windows</i> e <i>Android</i> .	Sim	Sim
Consumo de pacotes de dados de serviços como <i>Dynamics</i> , <i>Salesforce</i> e <i>Google Analytics</i> .	Sim	Sim
Importação de dados a partir de arquivos <i>Excel</i> ou <i>CSV</i> .	Sim	Sim
Consumir o conteúdo agendado para ser atualizado.	Diariamente	Até 8 vezes por dia
Consumir dados de <i>streaming</i> em <i>dashboards</i> e relatórios.	Não	Sim
Criar, publicar e exibir pacotes de conteúdo organizacional.	Não	Sim
Gerenciar o controle de acesso e compartilhamento por grupos do <i>Active Directory</i> .	Não	Sim

Colaboração de equipe usando grupos do <i>Office 365</i> no <i>Power BI</i> .	Não	Sim
---	-----	-----

Fonte: <https://powerbi.microsoft.com/en-us/pricing/>

### 3.4 A FERRAMENTA TABLEAU

Os *softwares Tableau* são desenvolvidos pela *Tableau Software*, empresa fundada em 2003 pelo doutorando Chris Stolte, pelo membro fundador da Pixar e professor orientador de Stolte, Pat Hanrahan e pelo empreendedor Christian Chabot, com a intenção de comercializar soluções de visualização de dados desenvolvidas no departamento de ciências da computação da Universidade de Stanford.

Durante suas pesquisas eles perceberam que a computação gráfica poderia gerar grandes avanços na capacidade das pessoas de compreender dados. Eles reuniram computação gráfica com banco de dados para inventar o *VizQL*, uma linguagem visual de consultas que traduz ações de arrastar e soltar em consultas de dados e, então, expressa esses dados visualmente. O resultado foi uma experiência de usuário intuitiva e que proporciona ganhos significativos na capacidade das pessoas verem e compreenderem os dados. (TABLEAU, 2016)

#### 3.4.1 DESCRIÇÃO E CARACTERÍSTICAS

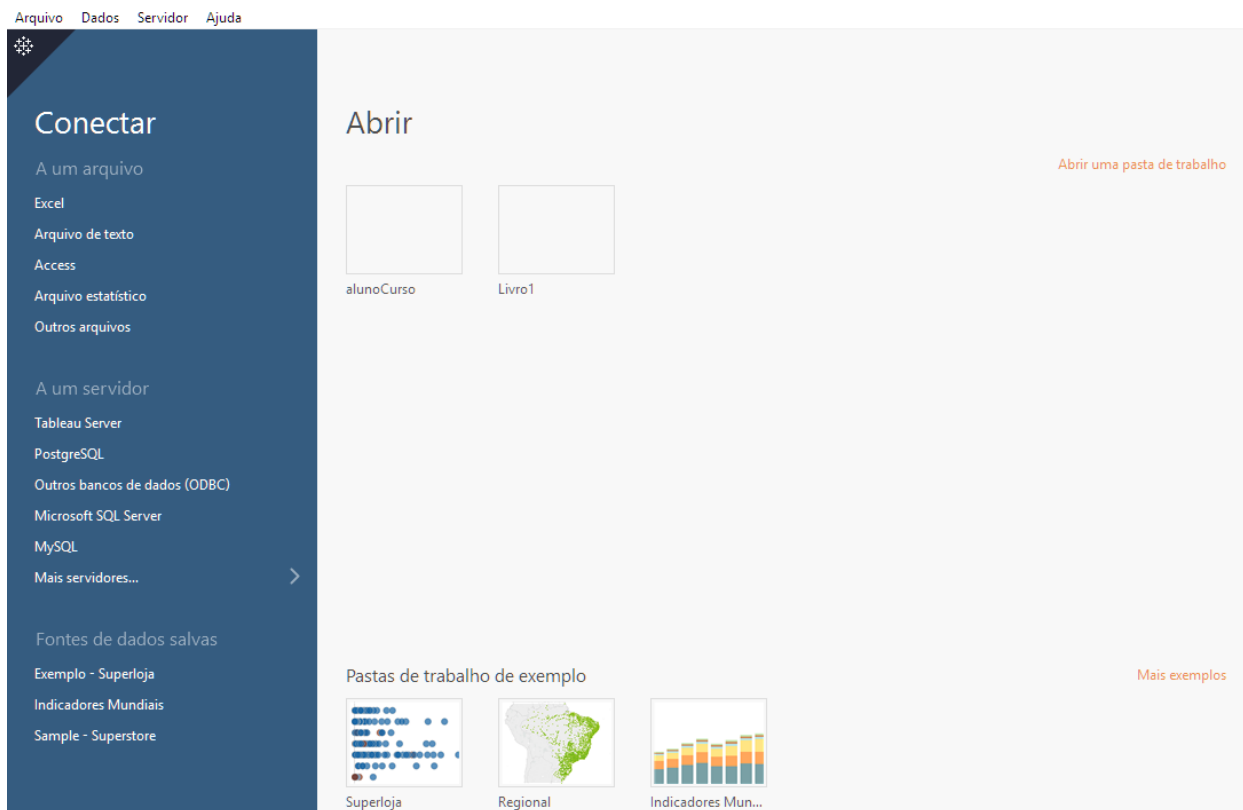
A *Tableau Software* oferece a seus clientes 6 produtos principais :

- *Tableau Desktop*: Ferramenta para visualização e análise de dados e criação de *dashboards*. O *Tableau Desktop* pode importar dados de várias fontes como *flatfiles*, planilhas e sistemas gerenciadores de banco de dados. O *software* é disponível em versões para *Windows* e *OSX*.
- *Tableau Server*: Usado para a publicação de *dashboards* e visualizações criadas no *Tableau Desktop*. Os *dashboards* podem ser acessados tanto por navegadores ou dispositivos móveis.
- *Tableau Online*: Versão online do *Tableau Server*. É uma alternativa para publicação de *dashboards* sem a necessidade de configuração de hardware já que os dados ficam armazenados na nuvem.

- *Tableau Mobile*: Aplicativo para dispositivos móveis disponível em versões para *Android* e *iOS*. O *Tableau Mobile* permite a visualização e análise de dados a partir de *dashboards* publicados no *Tableau Server* ou *Tableau Online*.
- *Tableau Public* – Permite que usuários do *Tableau Desktop* salvem seus *dashboards* na nuvem em um perfil *Tableau Public online* que posteriormente podem ser compartilhados na *web* em sites ou blogs.
- *Tableau Reader* – Aplicativo gratuito para a visualização de dados criados no *Tableau Desktop*. Disponível para *download* no site do *Tableau* em versões para *Windows* e *OSX*.

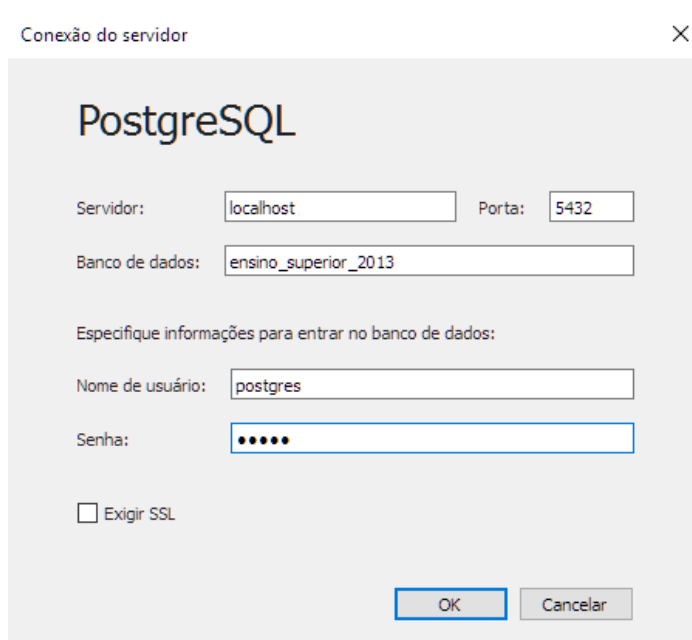
A análise será feita utilizando o *Tableau Desktop* em sua versão *Professional Edition* por ser o *software* que mais se assemelha com as outras ferramentas testadas. A instalação foi feita a partir de uma licença acadêmica disponibilizada gratuitamente pela *Tableau Software*.

Em sua tela inicial (Figura 27), o *Tableau* disponibiliza o acesso rápido a relatórios recentes ou a possibilidade de importar dados tanto em *flatfiles* como em acesso a sistemas gerenciadores de banco de dados.



**Figura 27** : Tela inicial do Tableau Desktop.

O *Tableau* fornece a possibilidade de conexão com vários sistemas gerenciadores de banco de dados. Ao selecionar a opção de interface com o *PostgreSQL* (Figura 28), uma nova conexão é criada após se fornecer os dados de conexão.



**Figura 28** : Conexão com o PostgreSQL no Tableau Desktop.

Após a conexão com o *PostgreSQL* o *Tableau* exibe as opções de extração de dados (Figura 29). É possível selecionar as tabelas e campos disponíveis no banco de dados previamente selecionado ou utilizar uma consulta personalizada com código SQL. O *Tableau* foi a única ferramenta avaliada que permite conexão em tempo real para se ter dados sempre atualizados ao invés de uma extração de dados estáticos.

A criação dos gráficos (Figura 30) é feita a partir da seleção das dimensões e medidas disponíveis. Dependendo dos dados selecionados o *Tableau* sugere algumas possíveis visualizações em formas de gráficos ou tabelas.

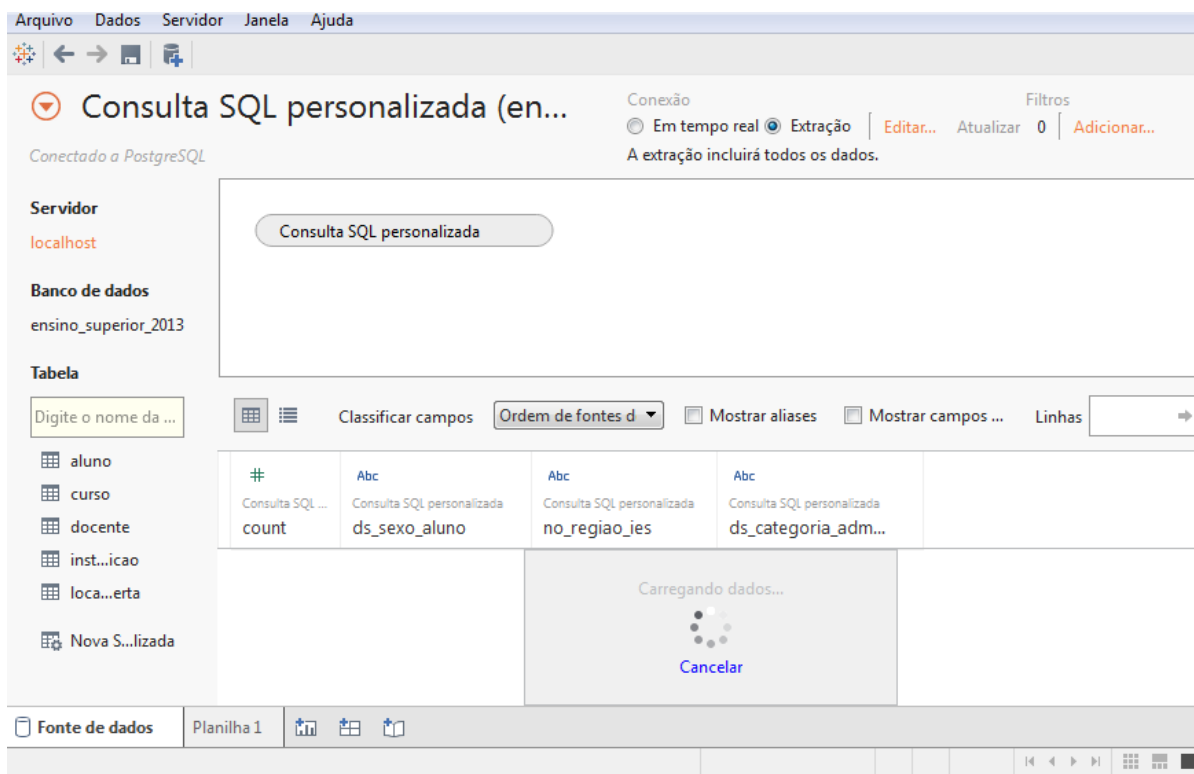


Figura 29 : Extração de dados no Tableau Desktop.

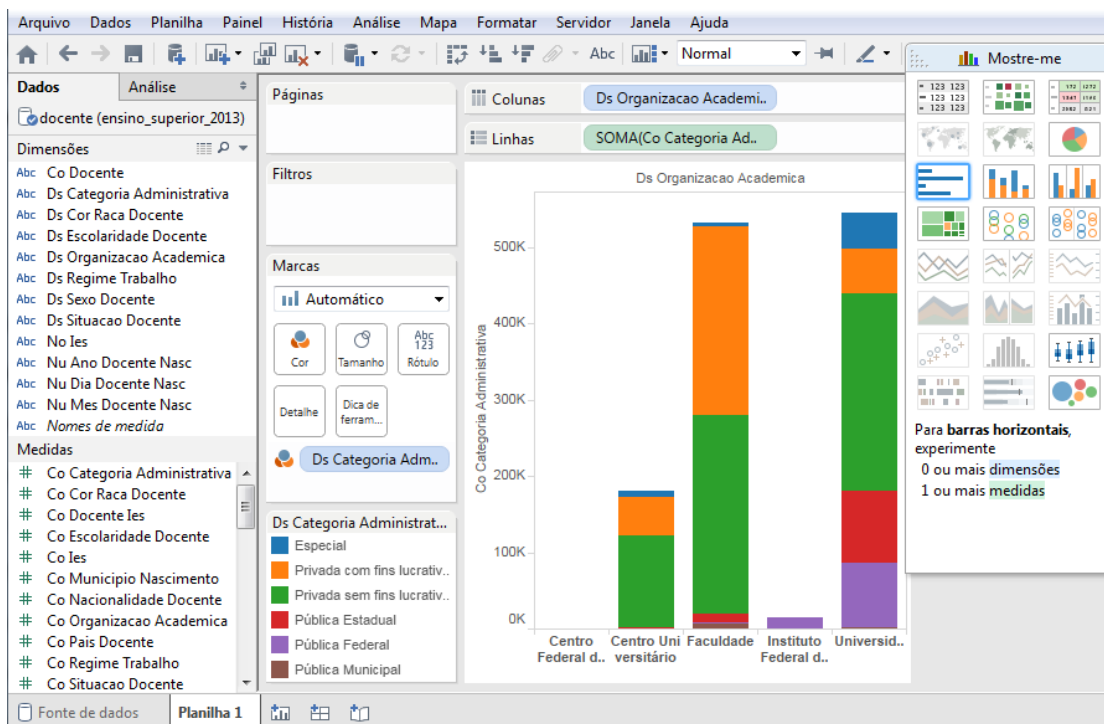


Figura 30 : Construção de gráficos no Tableau Desktop.

### 3.4.2 INVESTIMENTO

O *Tableau Desktop* é ofertado em 2 versões. A *Personal Edition* que custa 999,00 USD e tem uma quantidade limitada de fontes de dados compatíveis que incluem *Microsoft Excel*, *flatfiles* (CSV, txt etc.) e *Microsoft Access*. A versão mais completa é intitulada *Professional* e tem o custo de 1999,99 USD, permitindo o acesso a dados a partir de vários sistemas de gerenciamento de banco de dados como *IBM DB2*, *PostgreSQL*, *Microsoft SQL Server*, *Oracle*, *MySQL*, além de conexão com vários sistemas de dados armazenados na nuvem como *Google BigQuery* e *Amazon Redshift*. Apenas a versão *Professional* do *Tableau Desktop* oferece compatibilidade com o *Tableau Server* e *Tableau Online*. A *Tableau Software* também oferece licenças do *Tableau Desktop* gratuitamente para alunos e professores de nível acadêmico.

O custo de implementação do *Tableau Server* de acordo com o site da *Tableau* custa a partir de 10.000,00 USD para 10 usuários. Já o *Tableau Online*, que oferece as mesmas funções do *Tableau Server* porém hospedado na nuvem, custa 500,00 USD por usuário por ano. As versões do *Tableau Mobile* para aplicativos móveis podem ser baixadas gratuitamente a partir da *AppStore* ou do *Google Play*.

## 4 METODOLOGIA DO TRABALHO

O trabalho foi dividido em 5 fases, cada uma contendo tarefas a serem divididas entre os integrantes da equipe listando: Definição de escopo, desenvolvimento, experimentos, conclusões, entrega e apresentação . O detalhamento de todas as fases do projeto, bem como a divisão de tarefas se encontram disponíveis no Apêndice 2.

Os testes de avaliação de tempo para a extração dos dados e criação dos gráficos das ferramentas de BI foram realizados a partir de uma mesma máquina para uma correta comparação das ferramentas. A configuração de *hardware* utilizada foi :

CPU: Intel Core I5 – 4210u 1.7 GHz

Memória: 8GB RAM

Sistema Operacional: Windows 10 64 bits

### 4.1 CRITÉRIOS DE AVALIAÇÃO E ANÁLISE

Os recursos disponíveis nas ferramentas avaliadas foram classificados conforme os critérios da Tabela 10. Além disso, critérios de análise foram definidos a fim de identificar a qualidade das tarefas fornecidas pela ferramenta analisada (Tabela 11). Para definição desses critérios foi considerado como base a classificação do trabalho da PUCRJ, intitulado “Uma Análise Comparativa entre as Ferramentas OLAP como Apoio a Soluções de BI nas Empresas” Sá, Thays et al, 2011). Para a avaliação quanto ao nível organizacional as ferramentas foram avaliadas utilizando cores de acordo com o nível organizacional. Utilizamos a cor verde caso o critério avaliado satisfizesse o nível operacional, vermelho se satisfizesse o nível tático e azul caso o nível satisfeito fosse o estratégico, e preto caso nenhum nível fosse satisfeito em especial.

**Tabela 10:** Critérios de Avaliação para os recursos das ferramentas de BI.

<b>Critérios de Avaliação</b>	
"Bom" ou "B"	Recurso funciona de forma adequada ou esperada.
"Regular" ou "RE"	Recurso funciona, porém não obtém o desempenho esperado.
"Ruim" ou "R"	Recurso não segue um padrão adequado.
"N/E"	Não encontrado. Fonte Não Confiável ou Não Aplicável.

**Tabela 11:** Critérios de Análise para as tarefas fornecidas pelas ferramentas de BI.**Critérios de Análise**

1	Básicos	Características básicas para qualquer ferramenta OLAP.
2	Relatórios	Usabilidade dos relatórios e gráficos.
3	Funcionamento WEB	Disponibilidade da empresa via WEB para suporte.
4	Usabilidade	Ponto de vista do usuário para utilização da ferramenta.
5	Produto	Licenças e posicionamento da empresa no mercado.
6	Planejamento	Modo de distribuição dos relatórios.
7	Segurança	Segurança dos dados.

Uma subclassificação foi definida para os Critérios de Análise, a fim de permitir uma análise mais precisa desses requisitos. Esta subclassificação está definida na Tabela 12. Selecionamos os critérios baseado no trabalho desenvolvido pela PUCRJ, que consideramos relevantes e viáveis para esta análise e para os testes que foram realizados.

**Tabela 12:** Subclassificação dos Critérios de Análise

ID	Critério	Sub Critério	Descrição
1.1	Básicos	Desempenho	Avalia se a ferramenta tem uma boa performance ao processar consultas com um alto volume de dados.
1.2	Básicos	Consultas ADHOC	Avalia se a ferramenta permite ao usuário ter a liberdade de definir consultas que acredita ser melhor em um dado contexto.
1.3	Básicos	Plataforma	Avalia se a ferramenta pode ser executada nos sistemas operacionais mais difundidos, como Windows, Linux e UNIX.
1.4	Básicos	Suporte Técnico/Documentação	Avalia o nível de qualidade da documentação e suporte técnico oferecido pela ferramenta.
2.1	Relatórios	Agendamento	Avalia se a ferramenta permite o agendamento de relatórios.
2.2	Relatórios	<i>Dashboards</i>	Avalia se a ferramenta possibilita a criação de painéis.
2.3	Relatórios	Recursos de navegação	Avalia se a ferramenta oferece suporte para a geração de relatórios com recursos do tipo <i>Drills, Slice and Dice</i> <sup>10</sup> , etc.
2.4	Relatórios	Exportação para outros formatos	Avalia se a ferramenta dispõe de recursos de exportação para formatos como PDF, HTML e ODT.
3.1	Funcionamento WEB	Fóruns	Avalia se a ferramenta possui fóruns/blogs em quantidade razoável que forneçam informações relevantes.
3.2	Funcionamento WEB	Help Online	Avalia se a ferramenta possui o recurso de

<sup>10</sup> Operações slice and dice – é uma estratégia de segmentação dos dados, onde o usuário “fatia” os dados em porções menores para chegar num nível de detalhe satisfatório para determinadas análises.

			ajuda <i>on-line</i> .
3.3	Funcionamento WEB	Suporte a dispositivos móveis	Avalia se a ferramenta suporta o uso de dispositivos móveis como <i>IPad's</i> , <i>iPhone</i> e <i>BlackBerry</i> , por exemplo.
4.1	Usabilidade	Facilidade de uso	Indica o quão fácil é para o usuário leigo identificar suas funcionalidades, onde encontrá-las e como executá-las
4.2	Usabilidade	Atratividade	Avalia o grau em que a ferramenta possui uma interface amigável e atrativa;
4.3	Usabilidade	Interface personalizável	Identifica se a ferramenta permite customizações de interface para atender, por exemplo, a padrões gráficos e visuais do cliente.
5.1	Produto	Custo de licença	Avalia o custo para a compra da licença da ferramenta.
5.2	Produto	Mercado	Indica o porte das empresas usuárias do produto.
6.1	Planejamento	Carregamento de dados de diversas fontes	Avalia a possibilidade de integração da solução com fontes de dados heterogêneas.
6.2	Planejamento	Integração com office.	Avalia a possibilidade de integração da solução com as ferramentas do <i>Office</i> , que são comumente usadas pela maioria dos usuários.
7.1	Segurança	Perfil de usuário	Verifica se a ferramenta permite que o administrador defina níveis hierárquicos para os usuários do sistema;
7.2	Segurança	Aplicação de senha na fonte de dados	Verifica a possibilidade de proteger a fonte de dados com uma senha
7.3	Segurança	Log de auditoria	Avalia se a ferramenta permite que as empresas auditem as interações do usuário, mantendo <i>logs</i> no nível do sistema para fornecer visibilidade sobre quem está acessando o quê e quando.

## 4.2 DADOS E FONTES UTILIZADAS

Para se realizar um teste satisfatório nas ferramentas selecionadas é necessária uma base de dados que tenha um grande volume de registros para que seja possível a execução de testes de performance. Por este motivo, foi escolhido os dados do Censo do Ensino Superior 2013 realizado pelo INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira.

O INEP é uma autarquia vinculada ao Ministério da Educação que produz e dissemina informações da educação no Brasil com o objetivo de formular políticas voltadas para a melhoria da educação no país. Anualmente, o INEP realiza o censo da educação superior, uma coleta de dados com informações detalhadas sobre a situação do setor acadêmico no país. Este Censo reúne informações sobre as

instituições de ensino superior, seus cursos de graduação presencial ou à distância (EAD), cursos sequenciais, vagas oferecidas, inscrições, matrículas, ingressantes e concluintes, além de informações sobre docentes.

Os dados são coletados a partir de questionários preenchidos pelas instituições de ensino superior e importados por meio do sistema e-MEC, um portal criado pelo Ministério da Educação para fazer a tramitação eletrônica de dados submetidos pelas instituições de ensino. Após as rotinas de análise e finalização, os dados são publicados no site do INEP – [www.inep.gov.br](http://www.inep.gov.br)

A base de dados do Censo 2013 é composta de 5 arquivos no formato CSV<sup>11</sup>, conforme apresentado na Tabela 13.

**Tabela 13:** Arquivos que compõem os dados do Censo do Ensino Superior 2013.

Nome do arquivo	Número de registros	Descrição
DM_IES	2.391	Dados estatísticos das Instituições da Educação Superior.
.DM_CURSO	32.382	Dados estatísticos dos Cursos.
DM_DOCENTES	383.683	Dados estatísticos dos Docentes.
DM_ALUNO	9.929.289	Dados estatísticos de Aluno.
DM_LOCAL_OFERTA	165.046	Dados estatísticos de Local de Oferta

Além dos arquivos de dados já citados, o pacote disponível para *download* ([http://download.inep.gov.br/microdados/microdados\\_censo\\_superior\\_2013.zip](http://download.inep.gov.br/microdados/microdados_censo_superior_2013.zip)) contém um arquivo no formato PDF com o dicionário de dados do conteúdo do Censo e detalhes de utilização dos arquivos.

### 4.3 MODELAGEME CARREGAMENTOS DOS DADOS

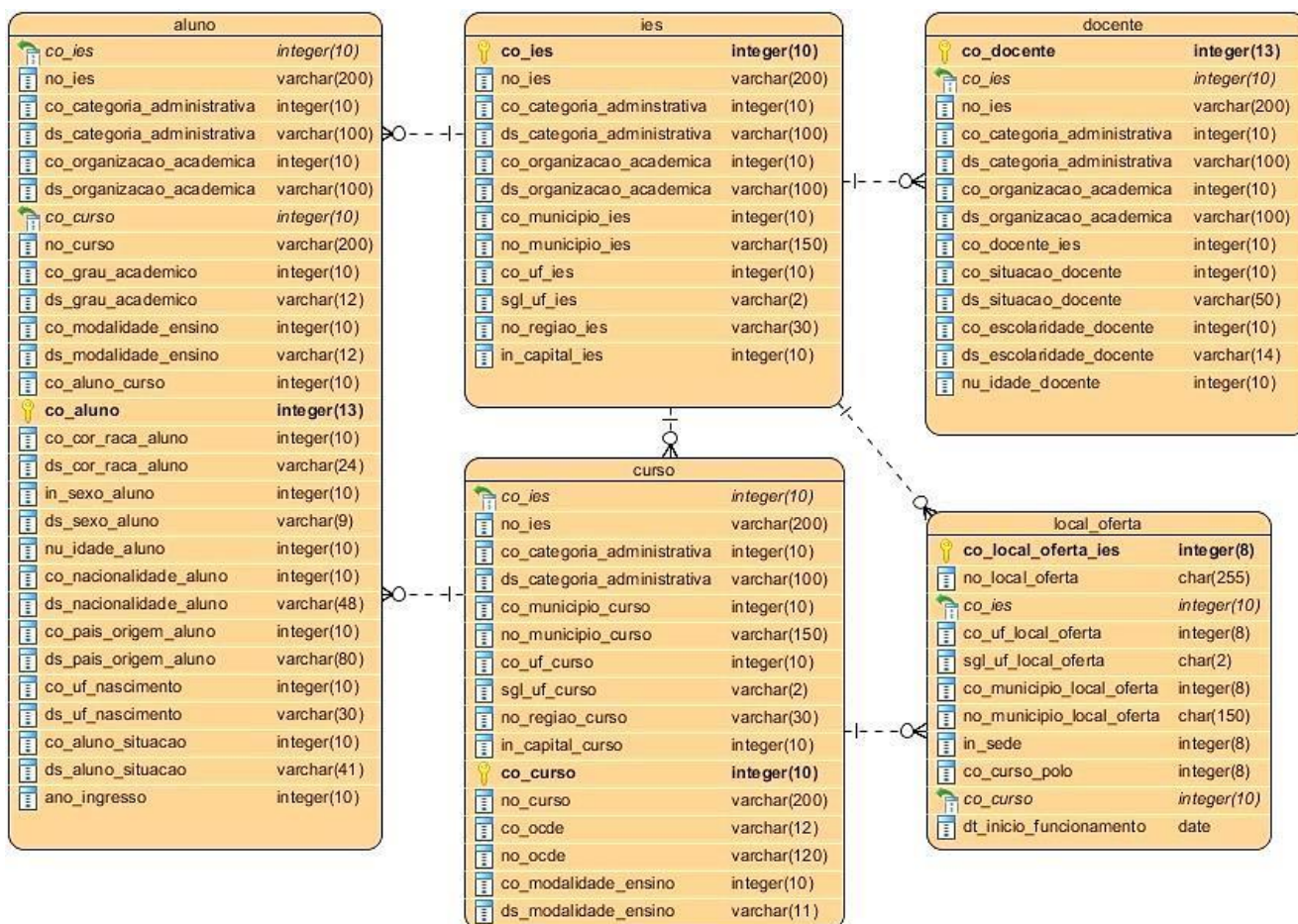
As ferramentas selecionadas para o estudo apresentam a possibilidade de processar dados tanto em forma de *flatfile*, fazendo o *upload* dos arquivos texto para a ferramenta, quanto na forma de conexão com SGBDs como *SQL Server*, *Oracle* e *PostgreSQL*. Porém, a partir dos testes realizados, mostrou-se inviável a utilização

<sup>11</sup> CSV (*Comma-Separated Values*): formato de arquivo onde os registros são separados por linhas e os campos separados por vírgula.

de *flatfiles* com os arquivos da base de dados escolhida, pois estes são muito grandes para serem carregados pelas ferramentas, a exemplo do arquivo CSV DM\_ALUNO com tamanho de 4.64 GB. Ao tentar realizar o processo de *upload* desses arquivos a ferramenta excede o tempo máximo de espera, cancelando o processo.

Foi necessário, então, a criação de um banco de dados em SGBD para a importação de todos os dados do Censo 2013. Para este propósito foi escolhido o SGBD *PostgreSQL* versão 9.4 por se tratar de um software de código aberto e com compatibilidade com as ambas as ferramentas de BI selecionadas. Foram feitos testes também com o SGBD *MySQL Community Edition*, uma versão do *MySQL* que pode ser utilizada gratuitamente por meio de licença GPL<sup>12</sup>. Porém, o *Microstrategy* não oferece opção de interface com essa versão do *MySQL*. As opções de conexões de SGBD do *Microstrategy* se limitam apenas as versões proprietárias do *MySQL* que requerem uma licença anual para sua utilização. Segundo o site do *MySQL*, ([www.mysql.com](http://www.mysql.com)) as licenças proprietárias custam de 2.000 USD a 10.000 USD dependendo da versão escolhida.

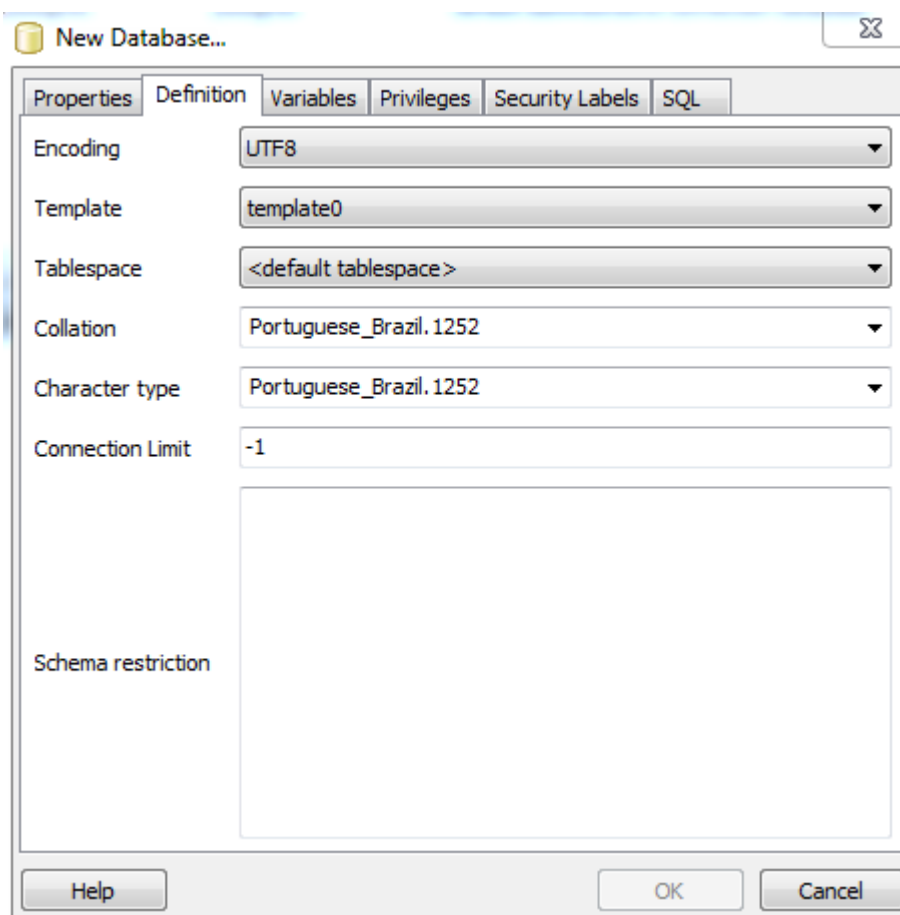
Antes de iniciar o processo de importação, foi necessário a preparação do banco de dados no *PostgreSQL* com as 5 tabelas a serem importadas e criadas no formato especificado pelo INEP. As tabelas para a importação dos dados foram criadas de acordo com o DER – Diagrama Entidade Relacionamento, especificado na Figura 31.



**Figura 31** : DER – base de dados Censo Ensino Superior 2013. Fonte: Autor.

Para auxiliar no processo de criação e manipulação dos dados, foi utilizado o *pgAdmin*, uma ferramenta gráfica para gerenciamento e administração de bases de dados *PostgreSQL*. O *pgAdmin* é também um *software open source* e pode ser baixado juntamente com o *PostgreSQL* a partir do site [www.postgresql.org](http://www.postgresql.org) ou a partir de seu site próprio [www.pgadmin.org](http://www.pgadmin.org).

Vale reforçar que durante a criação do banco de dados deve ser especificado o conjunto de caracteres a serem utilizados de forma que a base aceite as acentuações da língua portuguesa. Para isso, as opções de *collation* e *character type* devem ser especificadas como *Portuguese\_Brazil.1252*, conforme a Figura 32, extraída da tela de criação de um novo banco de dados no *pgAdmin*.



**Figura 32** : Criação de um novo banco de dados no pgAdmin.

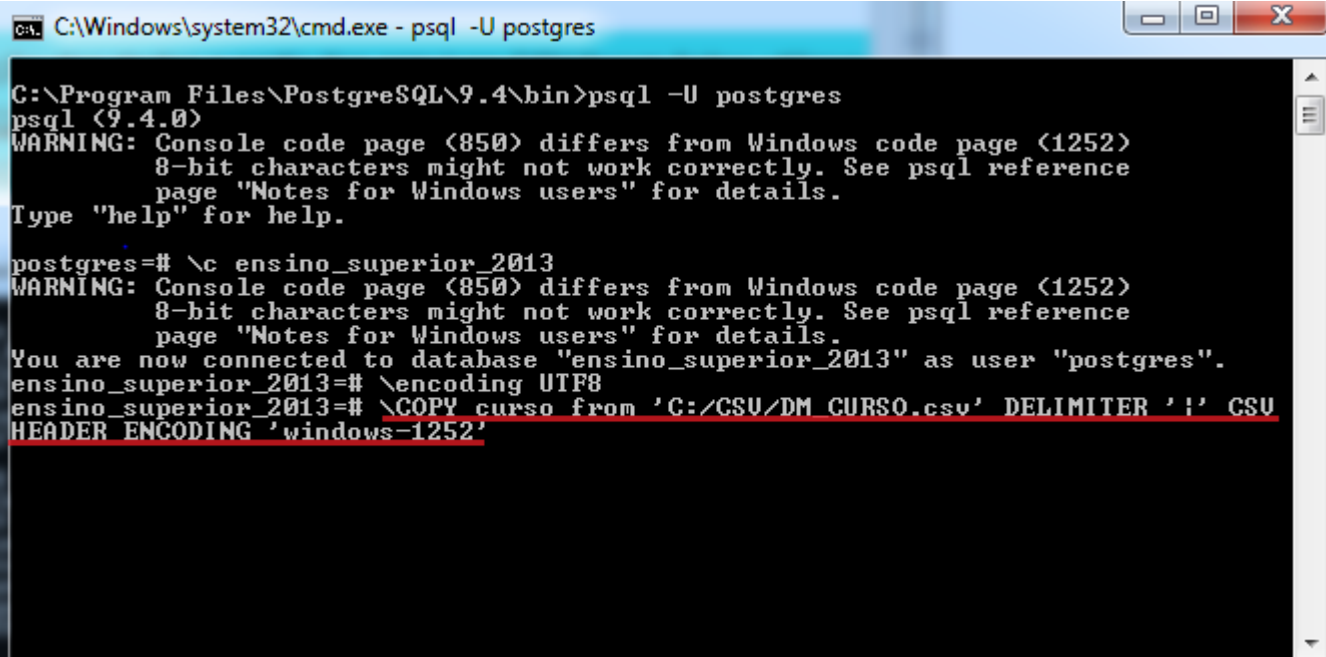
A importação dos dados a partir dos arquivos CSV foram feitos utilizando a função *COPY* do *PostgreSQL*. Essa função tem por objetivo fazer a cópia de dados entre arquivos texto e tabelas do SGBD, conforme demonstra a Figura 33.

Para a correta importação dos dados é necessário utilizar a sintaxe conforme a Figura 33: *COPY <nome da tabela> from <arquivo CSV>*, sublinhado em vermelho na imagem.

Para tanto, deve-se adicionar os parâmetros :

- *DELIMITER '|'* : especificar o caractere delimitador de colunas que no caso dos arquivos do INEP são utilizados caracteres pipe (|).
- *CSV* : para indicar o formato de arquivo de texto utilizado.
- *HEADER* : parâmetro para indicar a existência de cabeçalho no arquivo texto. Ao se utilizar esse parâmetro o *PostgreSQL* desconsidera a primeira linha do arquivo iniciando a importação a partir da segunda linha.
- *ENCODING 'windows-1252'* : especificação do conjunto de caracteres para o

correto reconhecimento de acentuações no texto importado.



```

C:\Windows\system32\cmd.exe - psql -U postgres
C:\Program Files\PostgreSQL\9.4\bin>psql -U postgres
psql (9.4.0)
WARNING: Console code page (850) differs from Windows code page (1252)
8-bit characters might not work correctly. See psql reference
page "Notes for Windows users" for details.
Type "help" for help.

postgres=# \c ensino_superior_2013
WARNING: Console code page (850) differs from Windows code page (1252)
8-bit characters might not work correctly. See psql reference
page "Notes for Windows users" for details.
You are now connected to database "ensino_superior_2013" as user "postgres".
ensino_superior_2013=# \encoding UTF8
ensino_superior_2013=# \COPY curso from 'C:/CSU/DM_CURSO.csv' DELIMITER ';' CSV
HEADER ENCODING 'windows-1252'

```

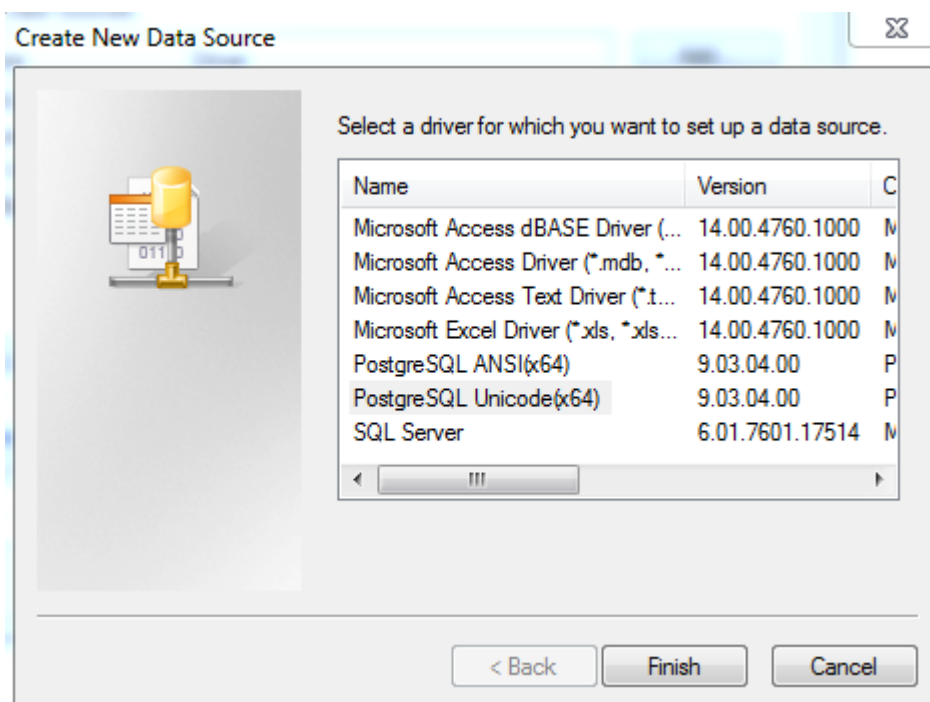
Figura 33 : Função COPY do PostgreSQL.

Após a conclusão da criação do banco de dados no PostgreSQL é necessário a configuração de uma conexão ODBC<sup>13</sup> (Figura 34 e 35). Este passo adicional é necessário para a utilização dos dados em *softwares* da *Qlik* já que eles exigem uma conexão ODBC para seu funcionamento. Outras ferramentas já apresentam opção de acesso direto ao PostgreSQL sem a necessidade de configuração de ODBC.

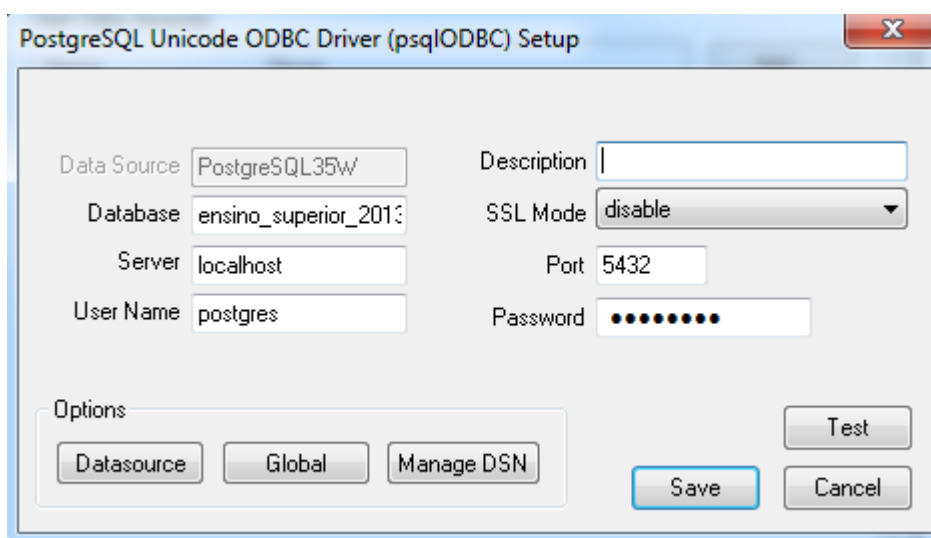
Os *drivers* oficiais de ODBC para o *PostgreSQL* estão disponíveis para *download* no site [odbc.postgresql.org](http://odbc.postgresql.org). Após a correta instalação do *driver* no sistema operacional, a configuração no *Windows* pode ser feita a partir da função Fontes de Dados (ODBC) disponível no Painel de Controle.

Deve-se adicionar uma nova fonte de dados utilizando-se o *driver* '*PostgreSQL Unicode(x64)*' e depois realizar sua configuração conforme exemplificado na Figura 34.

<sup>13</sup> ODBC – acrônimo para Open Database Connectivity, é uma interface padrão para acesso de Sistemas Gerenciadores de Banco de Dados (SGBDs)



**Figura 34** : Criação de uma nova conexão ODBC.



**Figura 35** : Configuração da conexão ODBC.

Após a criação do banco de dados no *PostgreSQL*, a importação de todos os dados da base escolhida e a configuração da conexão ODBC, temos o ambiente completo para se iniciar os testes nas ferramentas de BI selecionadas. A partir dos dados disponíveis no PostgreSQL, serão feitos testes nas ferramentas selecionadas por meio da criação dos relatórios especificados na próxima seção.

## 4.4 ESPECIFICAÇÃO DOS RELATÓRIOS

Com o objetivo de realizar testes de funcionalidades nas 4 ferramentas analisados foi definido 8 especificações de graficos que são detalhado s seguir.

Embora as ferramentas apresentem soluções de extração de dados por meio de seleção das tabelas e campos, para extração dos dados de cada relatório são utilizados somente códigos SQL com intuito de se realizar importação de dados de maneira idêntica em todas as ferramentas analisadas.

1) Relatório da relação sexo de alunos por área geral de curso: visualização em barra da relação de alunos do sexo masculino ou feminino por área geral de curso (Figura 36, 37, 38 e 39).

### **Tabela: Aluno**

Atributos: Código aluno e descrição sexo aluno

### **Tabela: Curso**

Atributos: Área geral e nome do curso

### **Código SQL**

```
select aluno.co_aluno,  
        aluno.ds_sexo_aluno,  
        aluno.no_curso,  
        aluno.no_ocde_area_geral  
from aluno  
join curso on curso.co_curso = cluno.co_curso
```

### Microstrategy

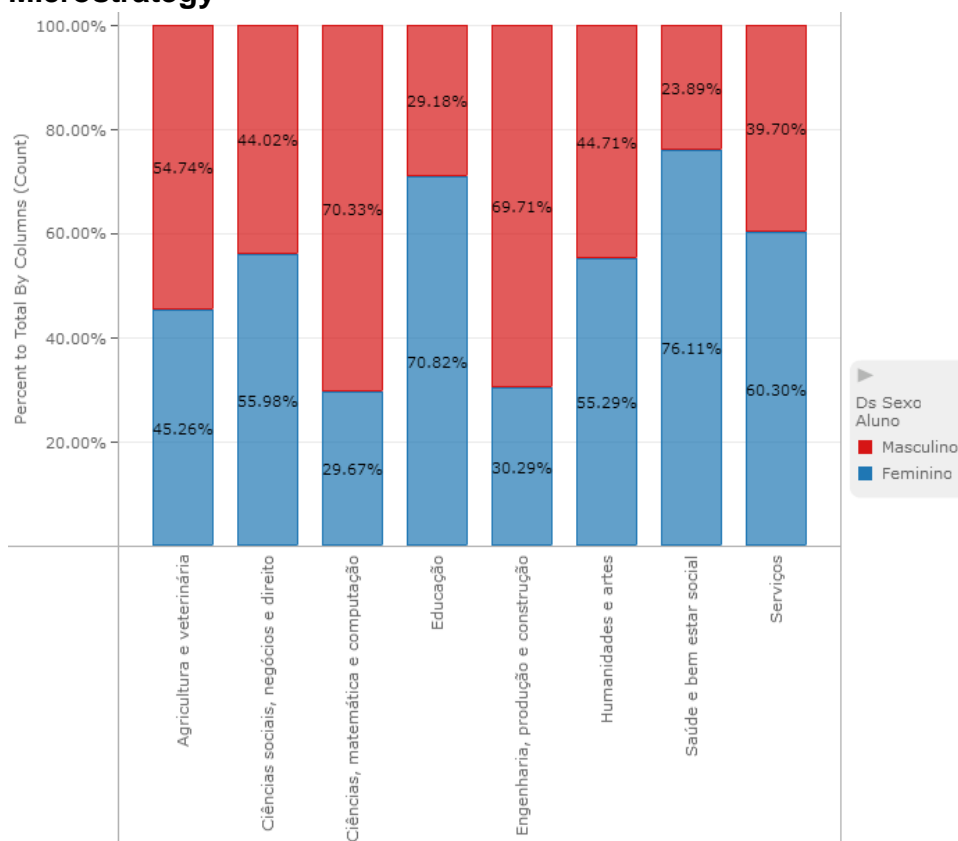


Figura 36 : Relatório 1 no Microstrategy.

### QLIKVIEW

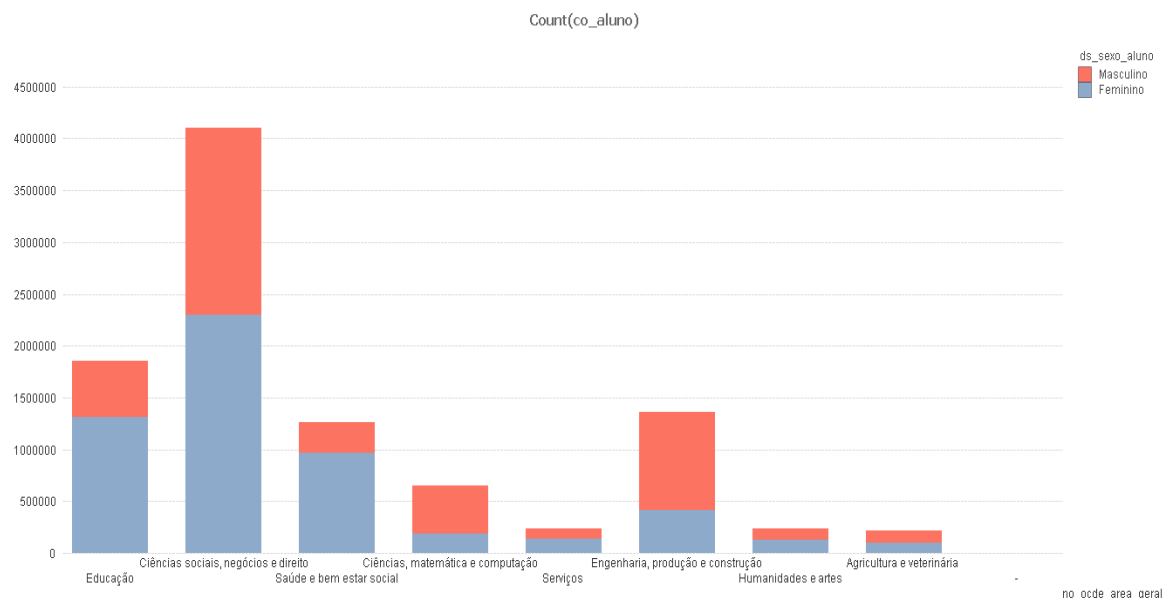


Figura 37 : Relatório 1 no Qlikview.

### Power BI

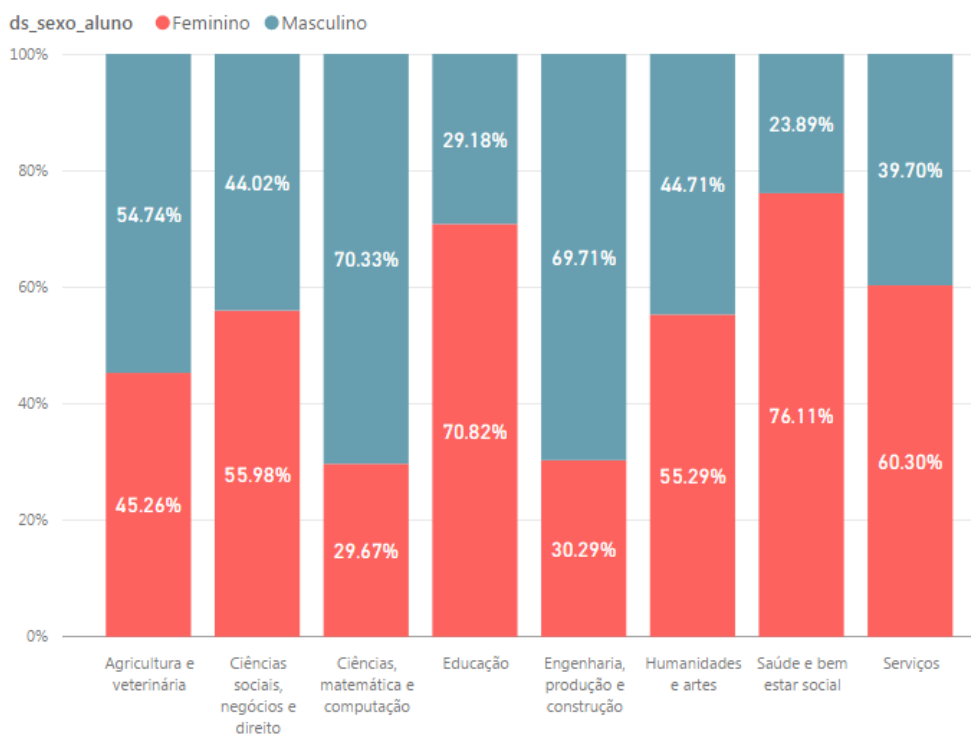
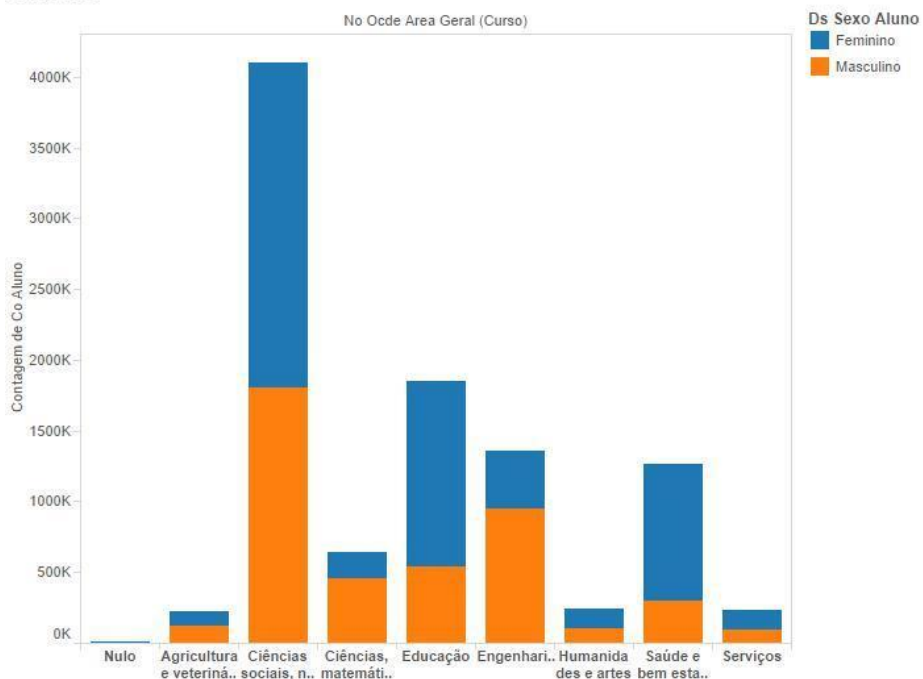


Figura 38 : Relatório 1 no Power BI

### Tableau

Planilha 1



Contagem de Co Aluno para cada No Ocde Área Geral (Curso). A cor mostra detalhes sobre Ds Sexo Aluno.

Figura 39 : Relatório 1 no Tableau.

2) Relatório de faixa de idade dos alunos matriculados em cursos de ensino superior: visualização em gráfico de linha que mostra o número de alunos por faixa de idade separados por área geral de curso (Figuras 40, 41, 42 e 43).

### Tabela: Aluno

Atributos: Idade do aluno

### Tabela: Curso

Atributos: Área geral e nome do curso

### Código SQL

```
select aluno.nu_idade_aluno,
       curso.no_ocde_area_geral,
       curso.no_curso
from aluno
join curso on curso.co_curso = aluno.co_curso
group by curso.no_ocde_area_geral,
         curso.no_curso
```

### Microstrategy

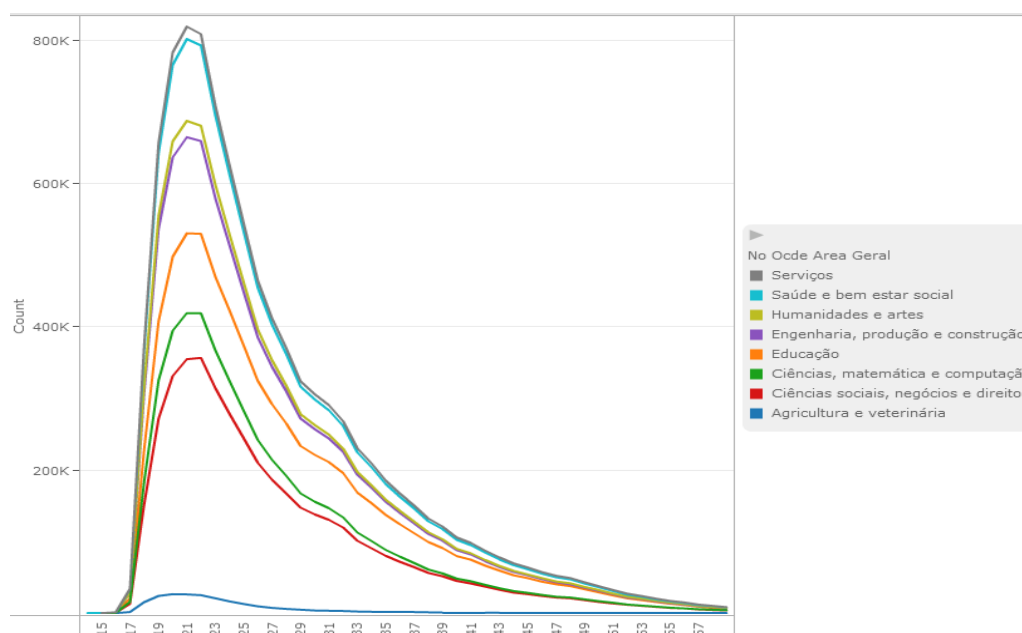
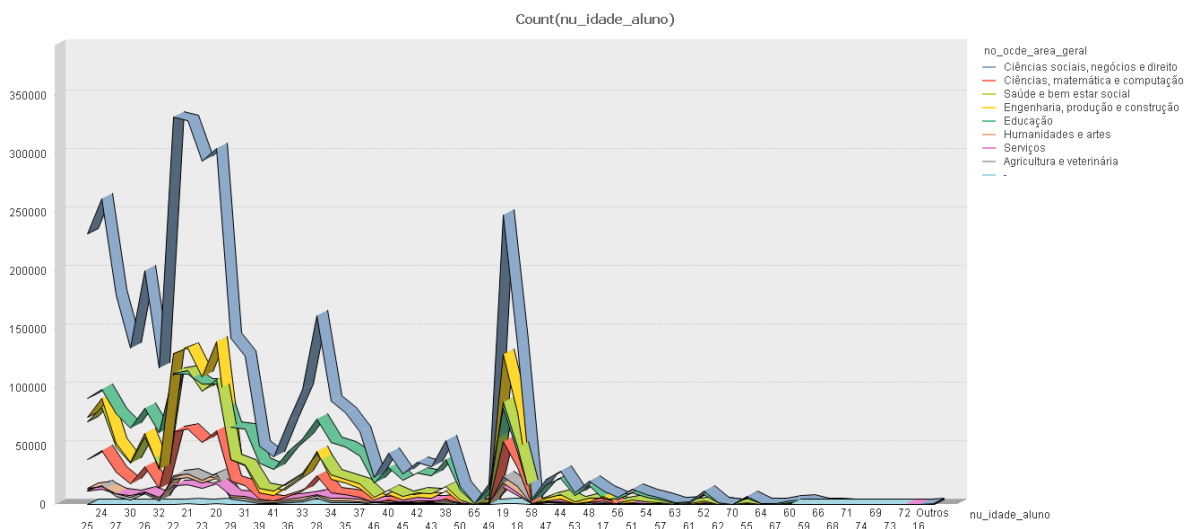


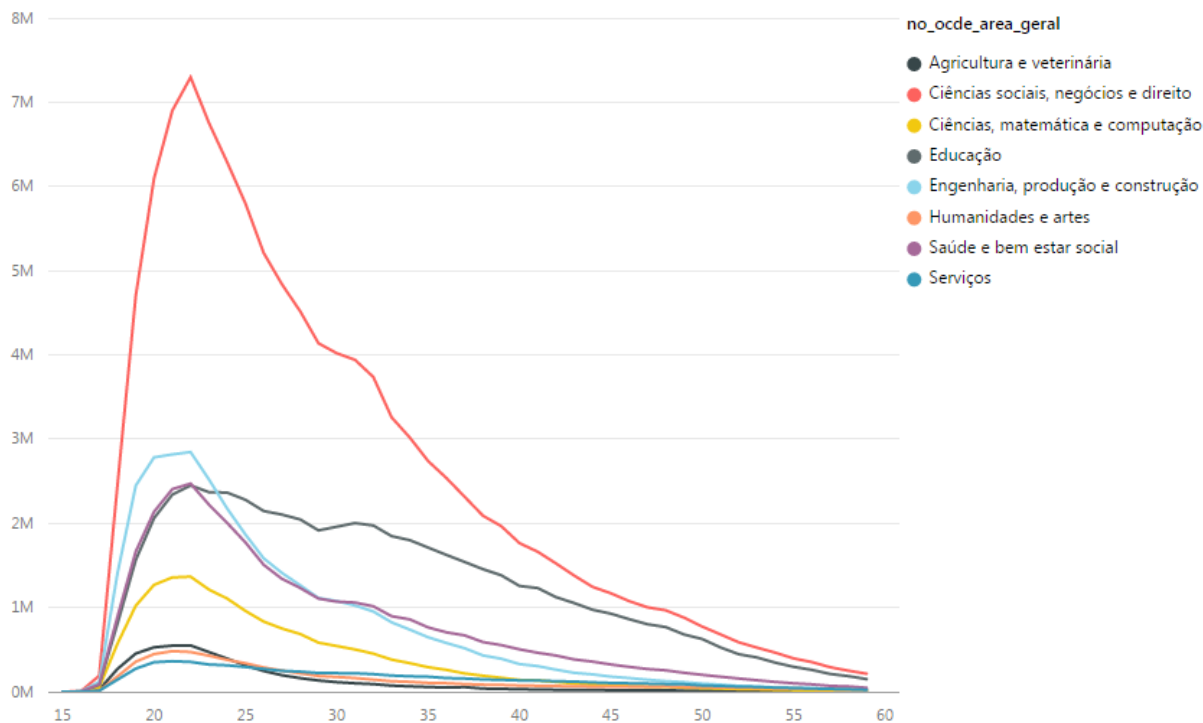
Figura 40 : Relatório 2 no Microstrategy.

**Qlikview**



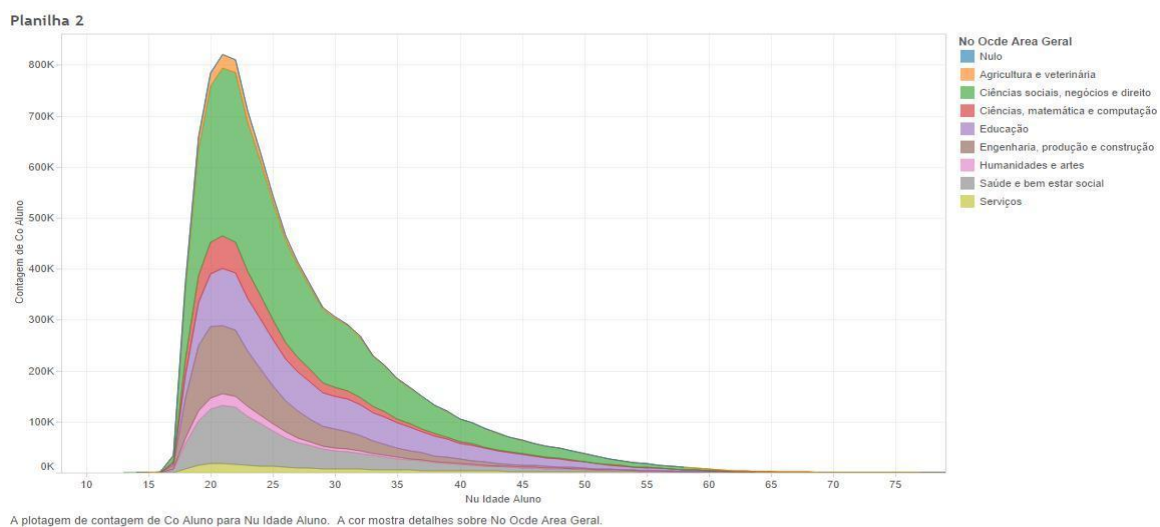
**Figura 41** : Relatório 2 no QlikView.

**Power BI**



**Figura 42** : Relatório 2 no Power BI.

## Tableau



**Figura 43** : Relatório 2 no Tableau

3) Relatório de número de alunos em cursos de EAD: visualização em gráfico circular separados por curso e dimensionados pela quantidade de alunos em cada curso (Figuras 44, 45, 46 e 47).

### Tabela: Aluno

Atributos: Código do aluno, nome do curso, modalidade do curso.

### Código SQL

```
select ds_modalidade_ensino,
       no_curso,
       ds_modalidade_ensino
from aluno
where ds_modalidade_ensino = 'Curso a distancia';
```

### Microstrategy

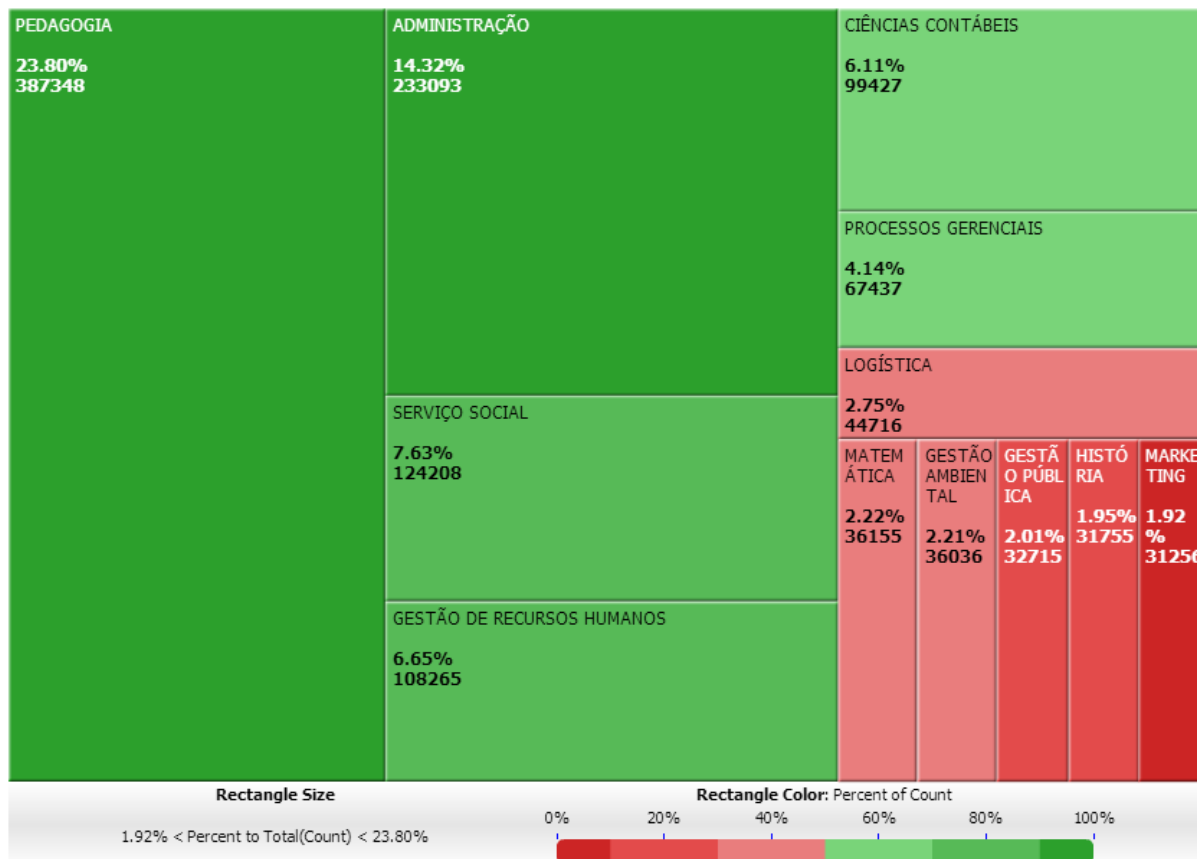


Figura 44 : Relatório 3 no Microstrategy.

### Qlikview

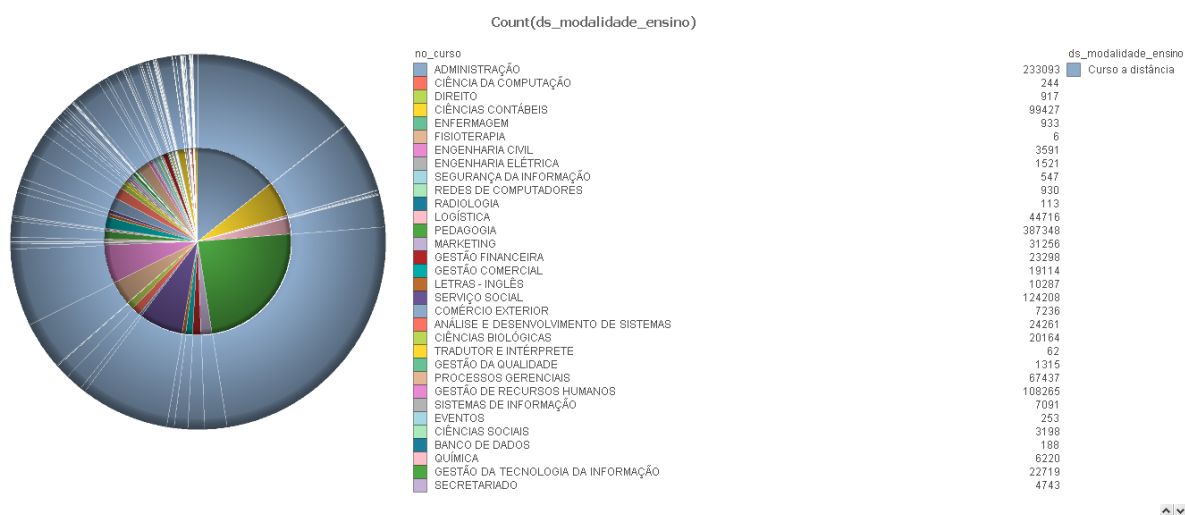


Figura 45 : Relatório 3 no QlikView.

### Power BI

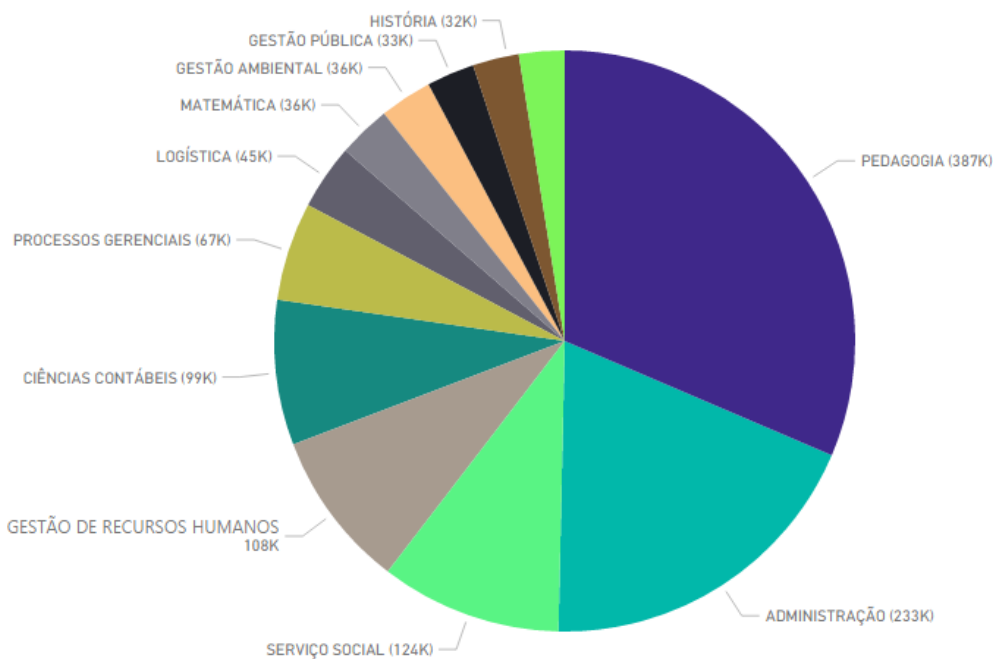


Figura 46 : Relatório 3 no Power BI.

### Tableau

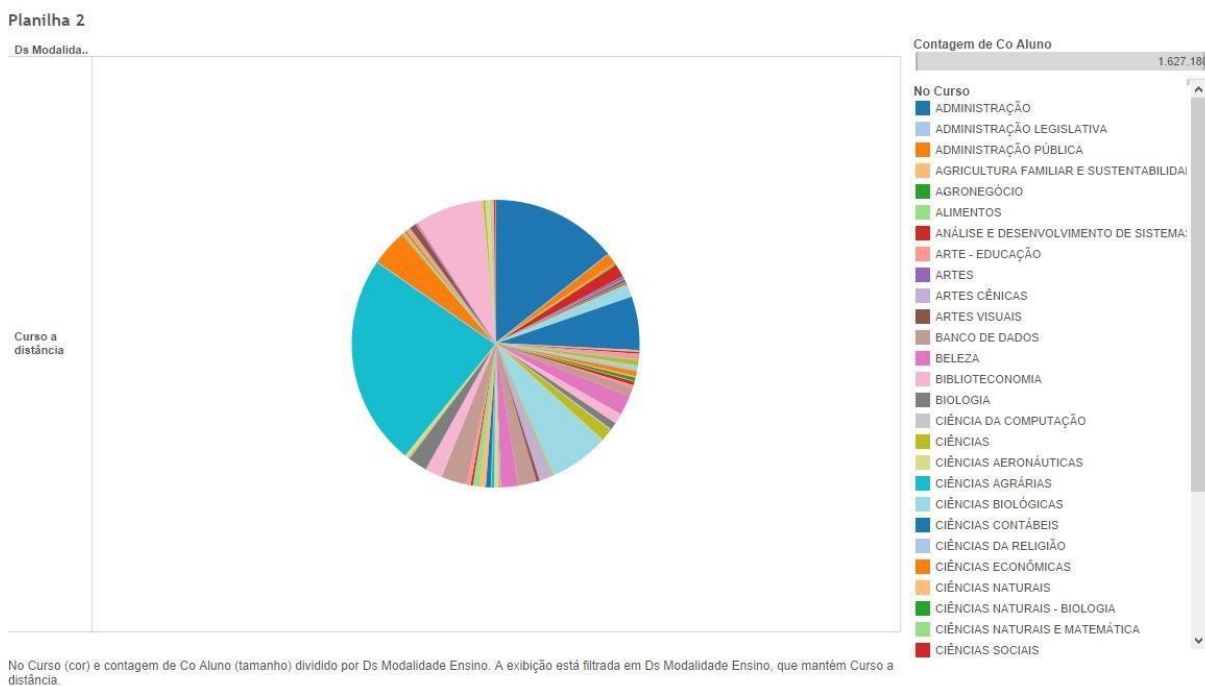


Figura 47 : Relatório 3 no Tableau.

4) Relatório de Escolaridade dos docentes por região: Visualização de gráfico de barras mostrando a relação de nível de escolaridade dos docentes, separados pelas regiões do Brasil (Figuras 48, 49, 50 e 51).

### Tabela: Docente

Atributos: Código da escolaridade e descrição da escolaridade do docente

### Tabela: Instituição

Atributos: Código, região e UF da instituição de ensino

### Código SQL

```
select docente.co_escolaridade_docente,
       docente.ds_escolaridade_docente,
       instituicao.co_ies,
       instituicao.no_regiao_ies,
       instituicao.sgl_uf_ies
from docente
join instituicao on docente.co_ies = instituicao.co_ies
```

### Microstrategy

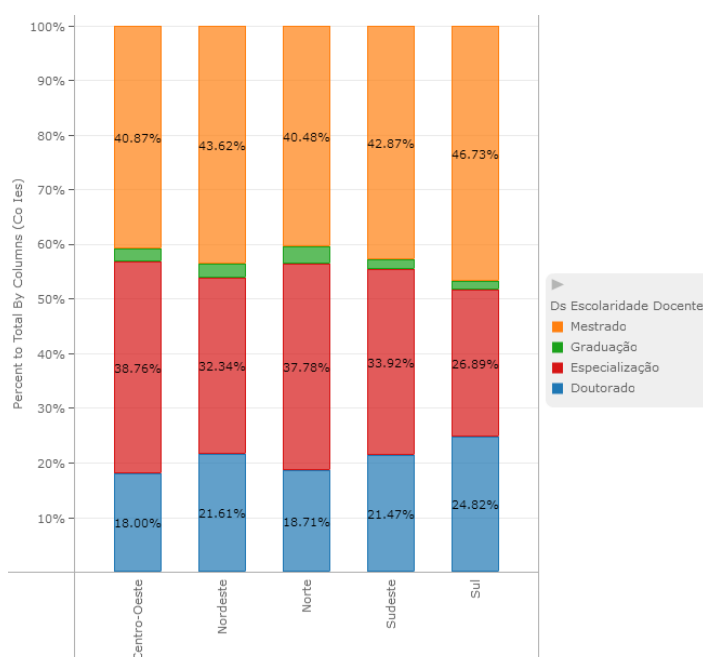


Figura 48 : Relatório 4 no Microstrategy.

## Qlikview

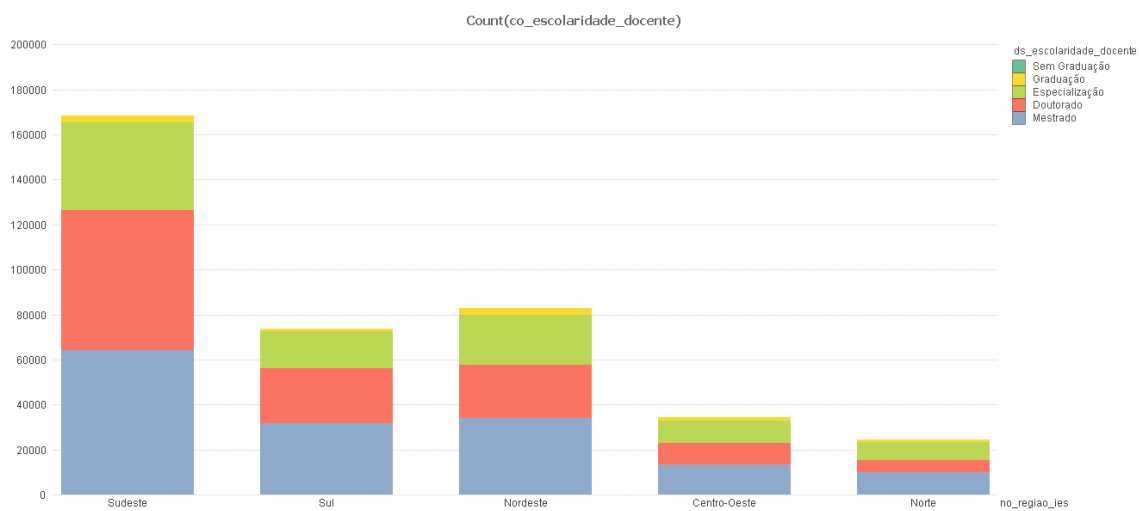


Figura 49 : Relatório 4 no QlikView.

## Power BI

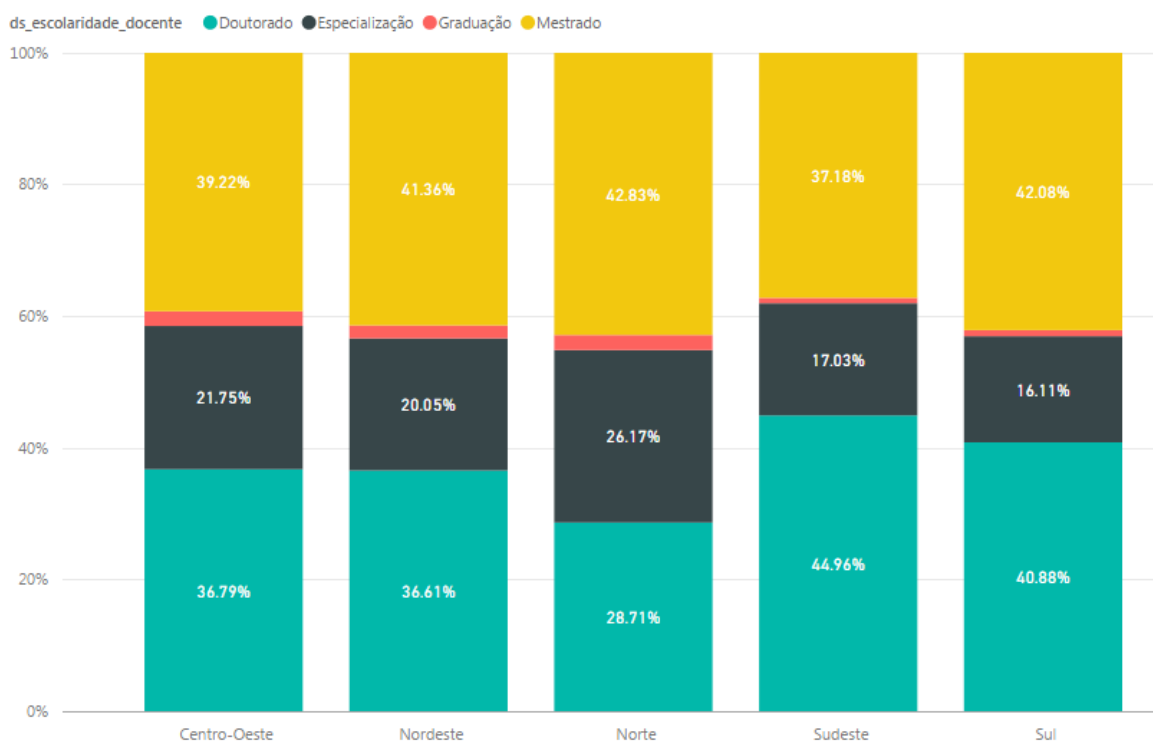
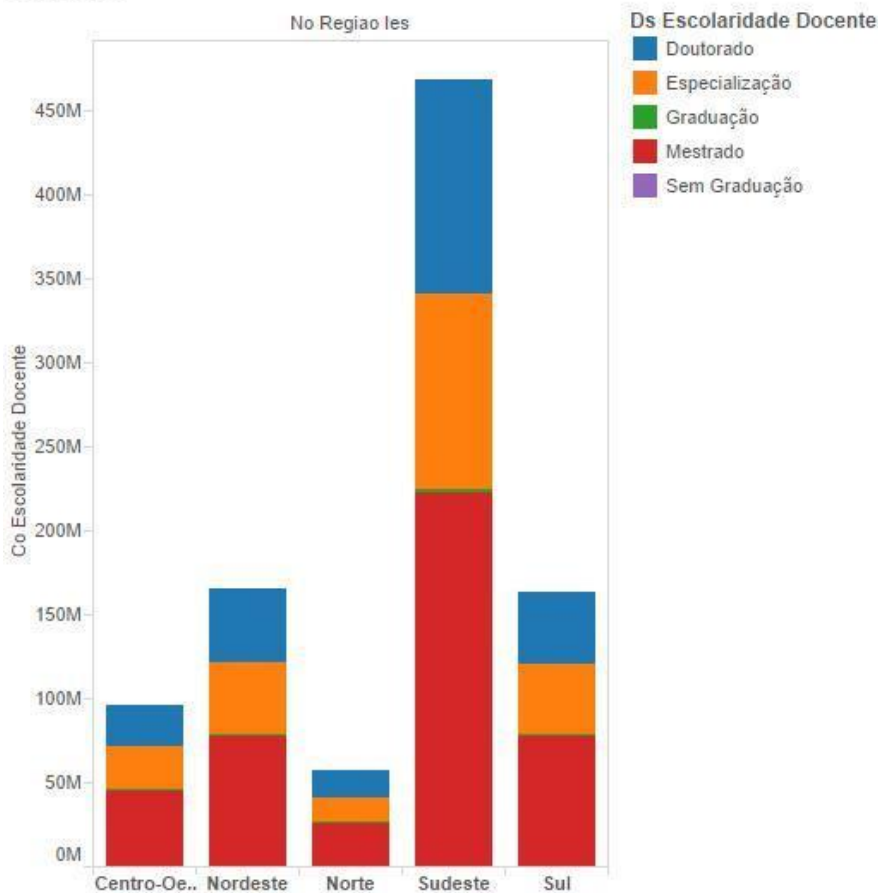


Figura 50 : Relatório 4 no Power BI.

## Tableau

### Planilha 1



Soma de Co Escolaridade Docente para cada No Região Ies. A cor mostra detalhes sobre Ds Escolaridade Docente.

**Figura 51** : Relatório 4 no Power BI.

- 5) Relação de escolaridade de docentes por Categoria Administrativa da Instituição de ensino: Este relatório visa ter uma visão geral da escolaridade dos docentes dependendo da categoria administrativa da instituição de ensino. Visualização de gráfico de barras mostrando a relação de nível de escolaridade dos docentes, separados por categorias administrativas (Figuras 52, 53, 54 e 55).

### **Tabela: Docente**

Atributos: Código único do docente, descrição da escolaridade e código da escolaridade.

### **Tabela: Instituição**

Atributos: Descrição da categoria administrativa, UF e nome da instituição.

## Código SQL

```

select docente.co_docente,
       docente.co_escolaridade_docente,
       docente.ds_escolaridade_docente,
       docente.co_ies,
       instituicao.no_ies,
       instituicao.co_ies,
       instituicao.co_categoria_administrativa,
       instituicao.ds_categoria_administrativa
from docente
join instituicao on docente.co_ies = instituicao.co_ies
order by docente.co_escolaridade_docente desc

```

## Microstrategy

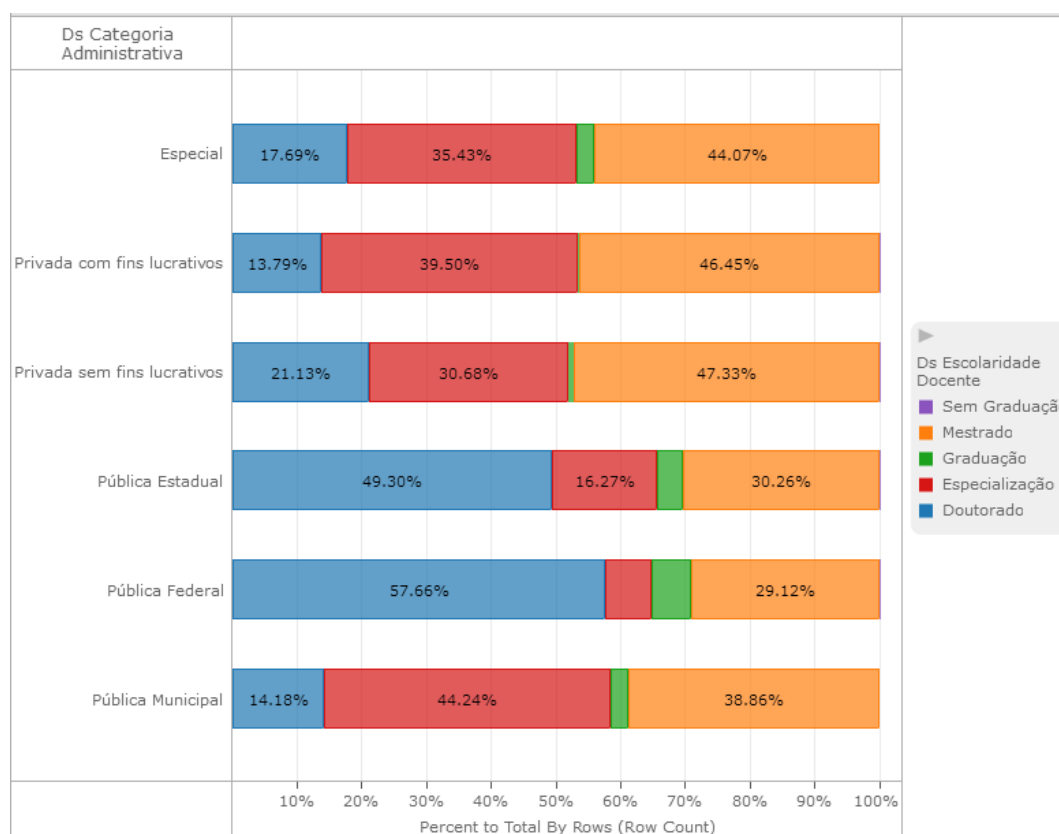


Figura 52 : Relatório 5 no Microstrategy.

### Qlikview

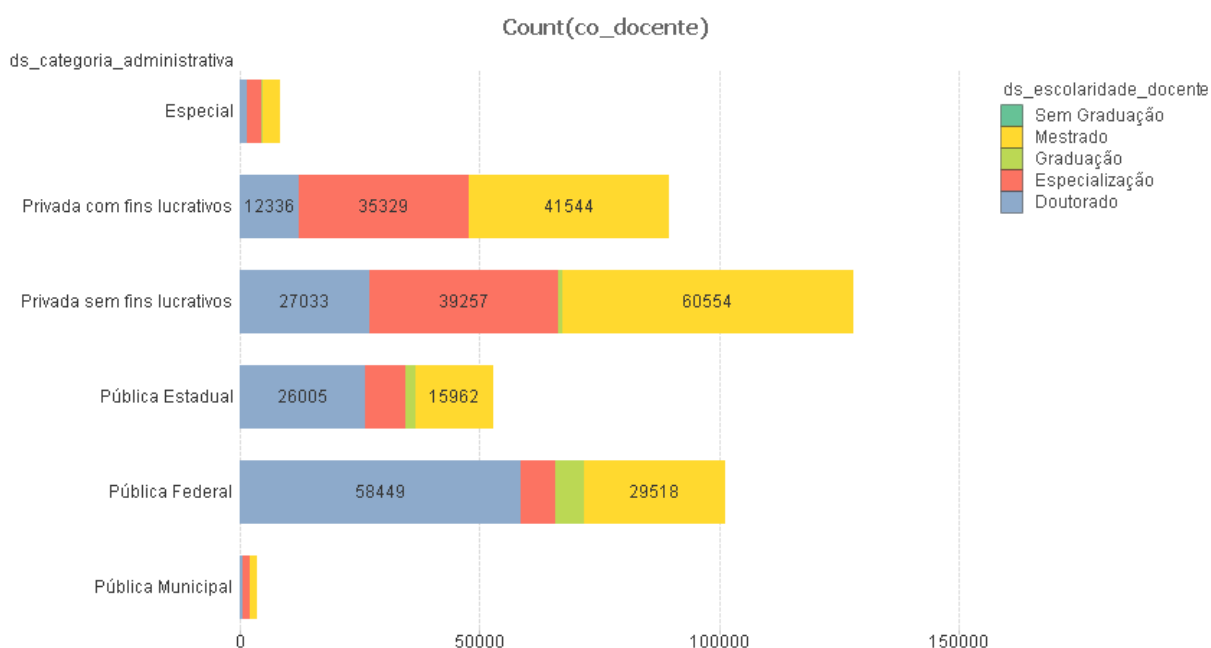


Figura 53 : Relatório 5 no QlikView.

### Power BI

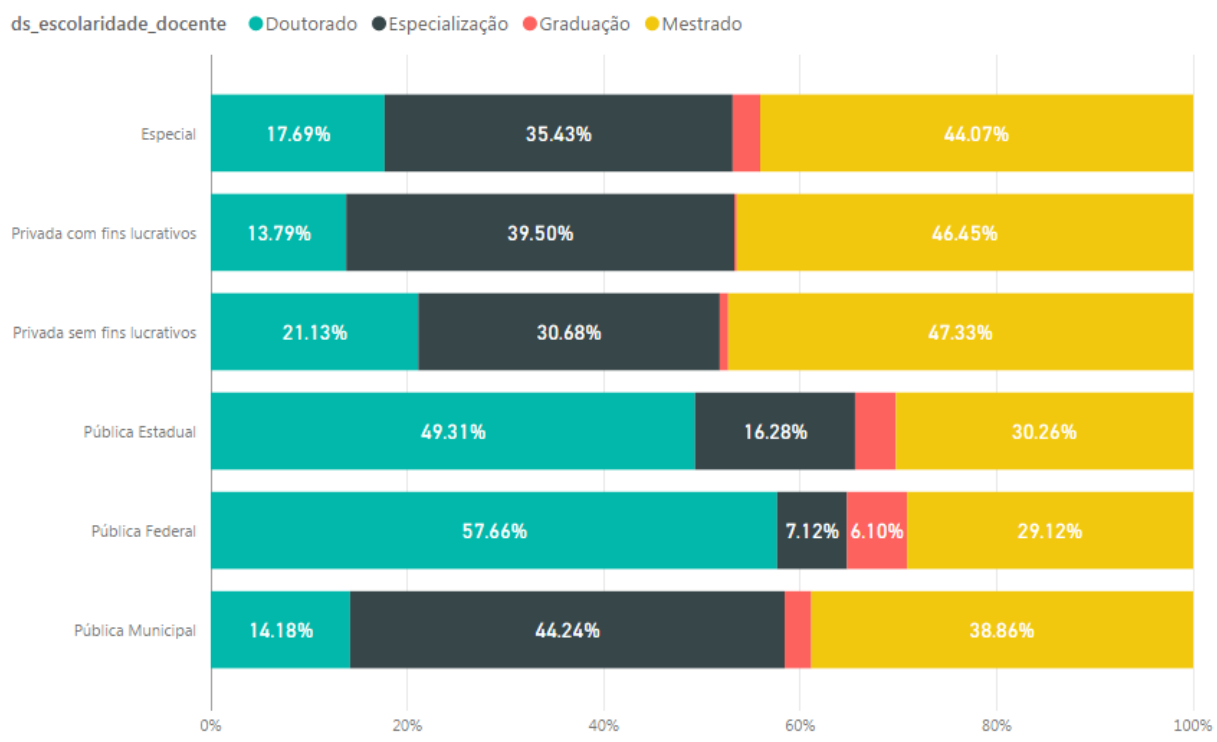
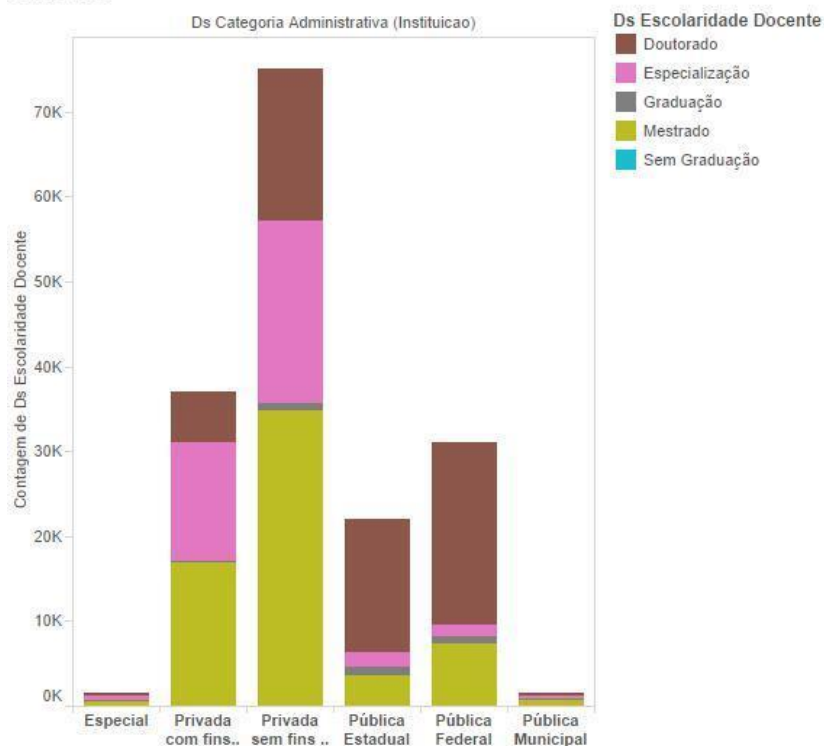


Figura 54 : Relatório 5 no Power BI.

## Tableau

### Planilha 1



Contagem de Ds Escolaridade Docente para cada Ds Categoria Administrativa (Instituicao). A cor mostra detalhes sobre Ds Escolaridade Docente. Os dados estão filtrados em No Regiao Ies, que mantém Sudeste.

**Figura 55** : Relatório 5 no Tableau.

6) Quantidade de alunos estrangeiros por nacionalidade: visualização em gráfico circular separados por nacionalidade do aluno e dimensionados pela quantidade de alunos de cada nacionalidade (Figuras 56, 57, 58 e 59).

### Tabela: Aluno

Atributos: País de origem, nacionalidade do aluno, nome da instituição e categoria administrativa da instituição.

### Código SQL

```
select ds_pais_origem_aluno,
       ds_nacionalidade_aluno,
       co_categoria_administrativa,
       ds_categoria_administrativa,
       no_ies
from aluno
```

### Microstrategy

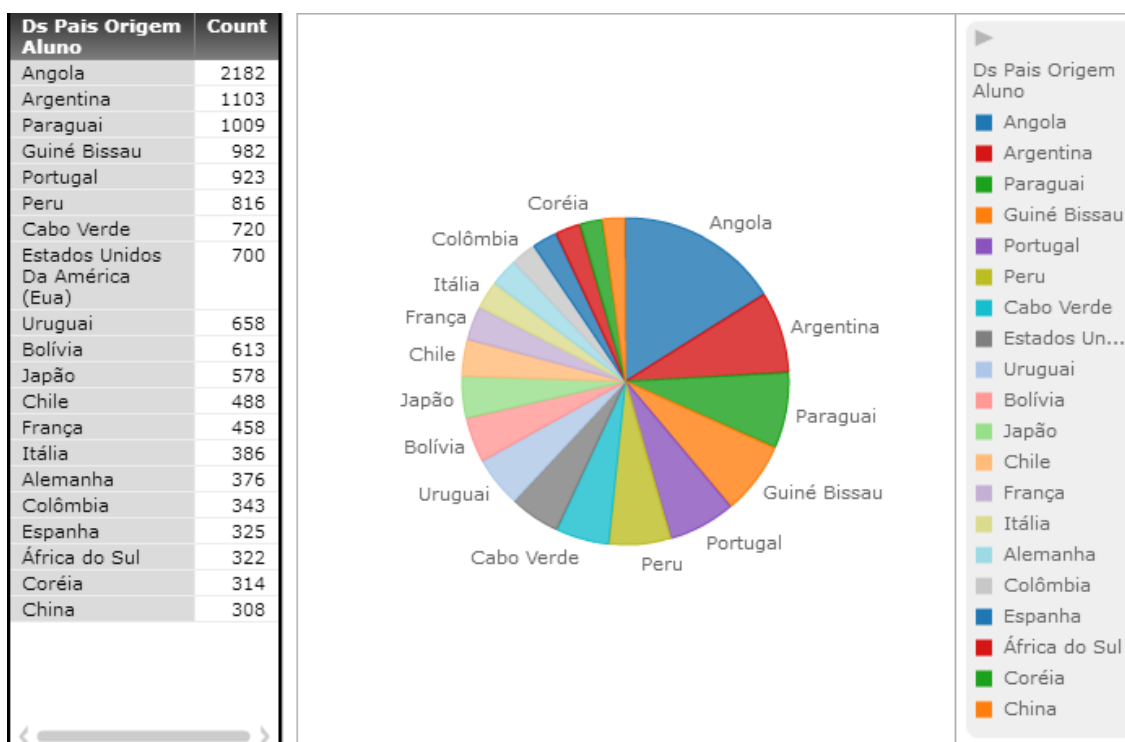


Figura 56 : Relatório 6 no Microstrategy

### Qlikview

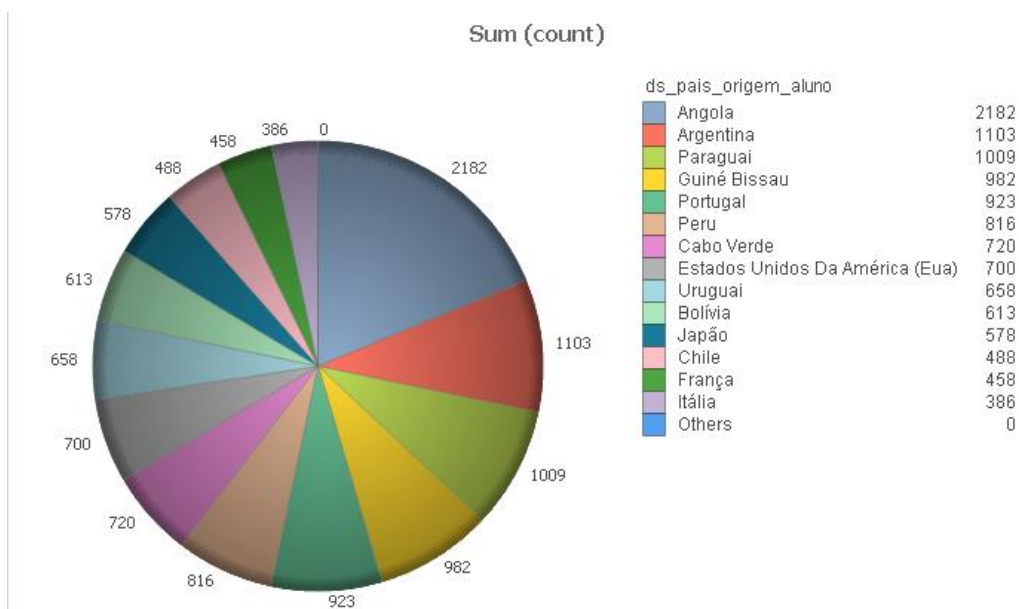


Figura 57 : Relatório 6 no QlikView.

### Power BI

< Back to Report | COUNT BY DS\_PAIS\_ORIGEM\_ALUNO

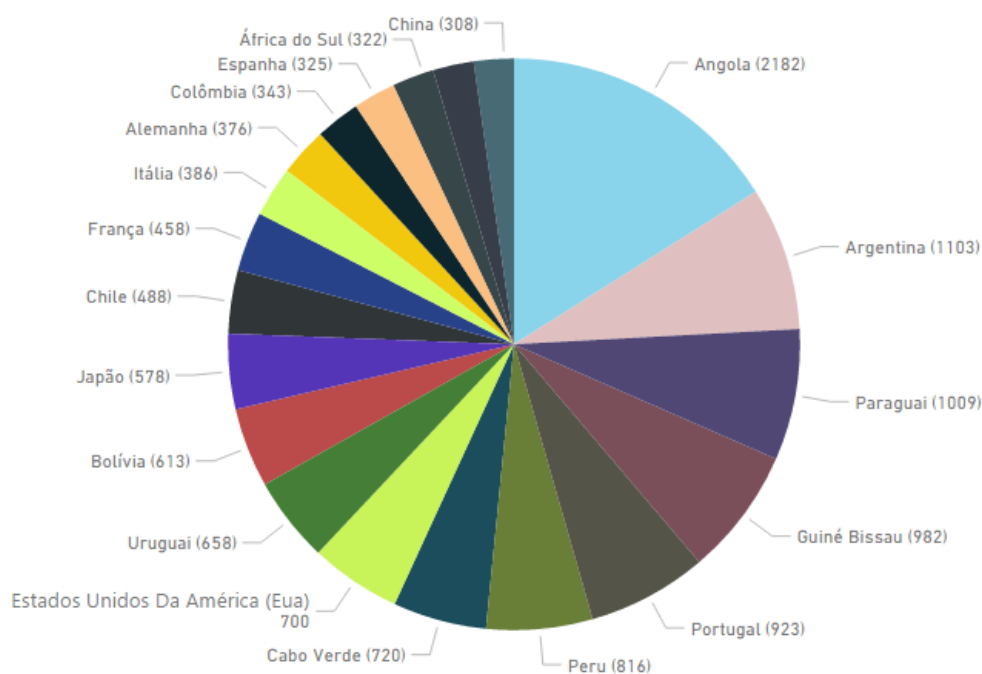


Figura 58 : Relatório 6 no Power BI.

### Tableau

Planilha 1

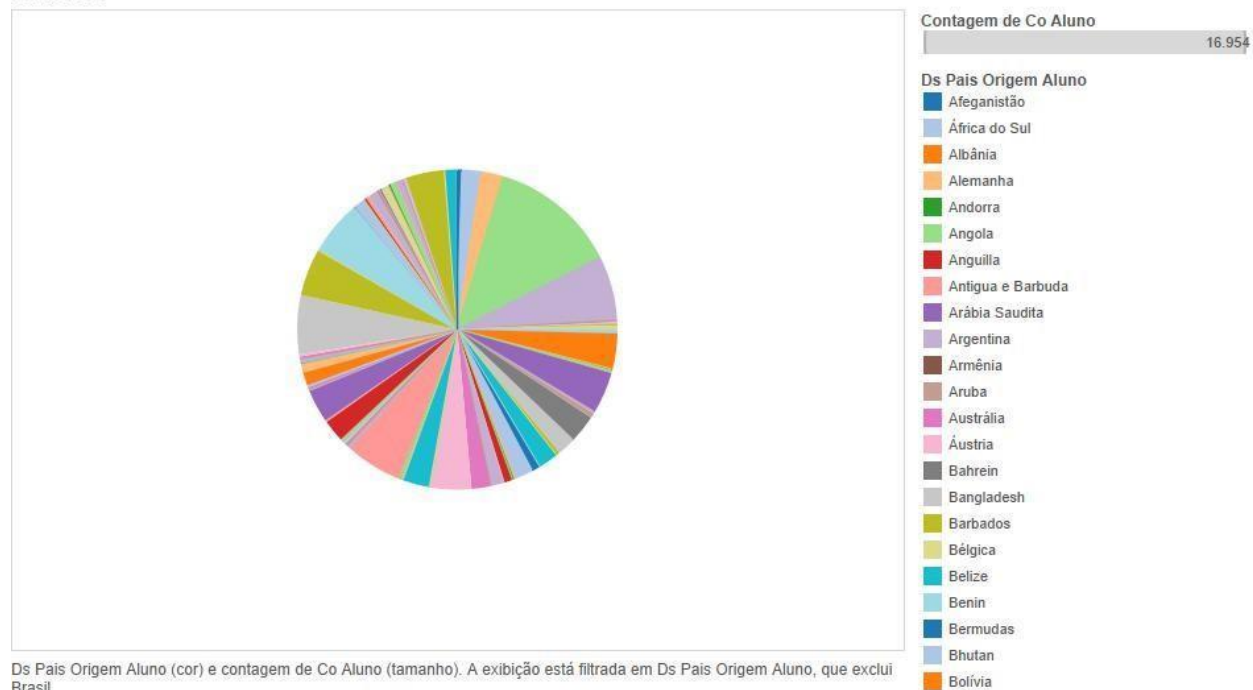


Figura 59 : Relatório 6 no Tableau.

7) Sexo de alunos por região/IES pública x privada: visualização de gráfico de barras com o total de alunos do sexo masculino ou feminino em instituições publicas ou privadas em cada região do Brasil (Figuras 60, 61, 62 e 63).

### Tabela: Aluno

Atributos: Sexo do aluno e código da instituição de ensino

### Tabela: Instituição

Atributos: Região e organização acadêmica da instituição de ensino.

### Código SQL

```
select aluno.ds_sexo_aluno,
       aluno.ds_organizacao_academica,
       aluno.co_ies,
       instituicao.co_ies,
       instituicao.no_regiao_ies,
       instituicao.ds_organizacao_academica
from aluno
cross join instituição
```

### Microstrategy

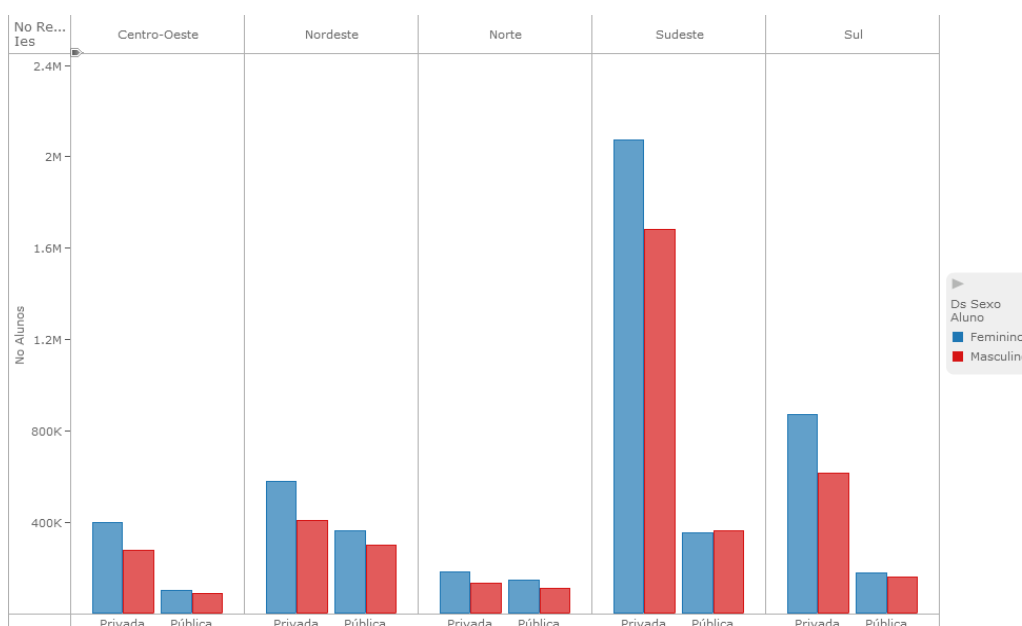


Figura 60 : Relatório 7 no Microstrategy.

## Qlikview

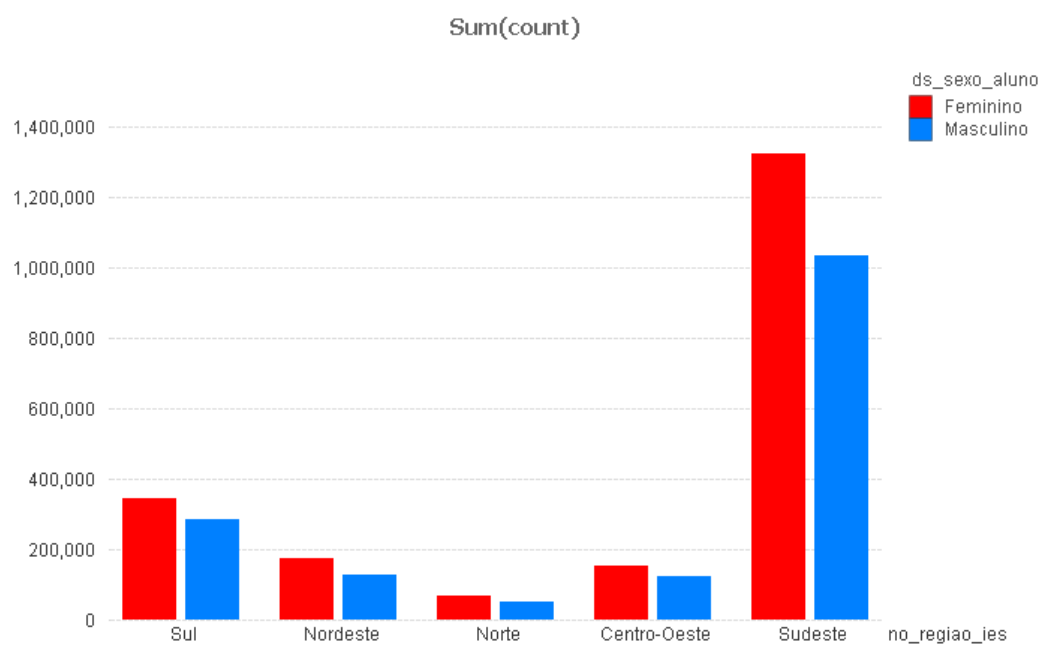


Figura 61 : Relatório 7 no QlikView.

## Power BI

ds\_sexo\_aluno ● Feminino ● Masculino

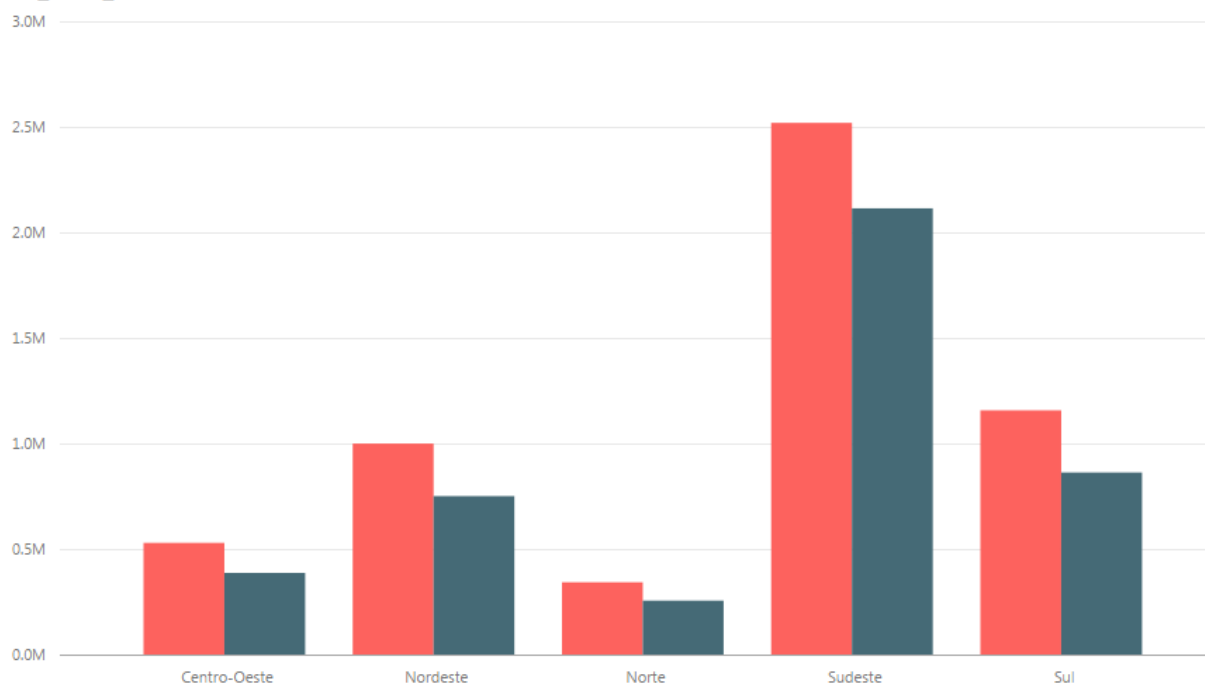


Figura 62 : Relatório 7 no Power BI.

## Tableau

Planilha 1

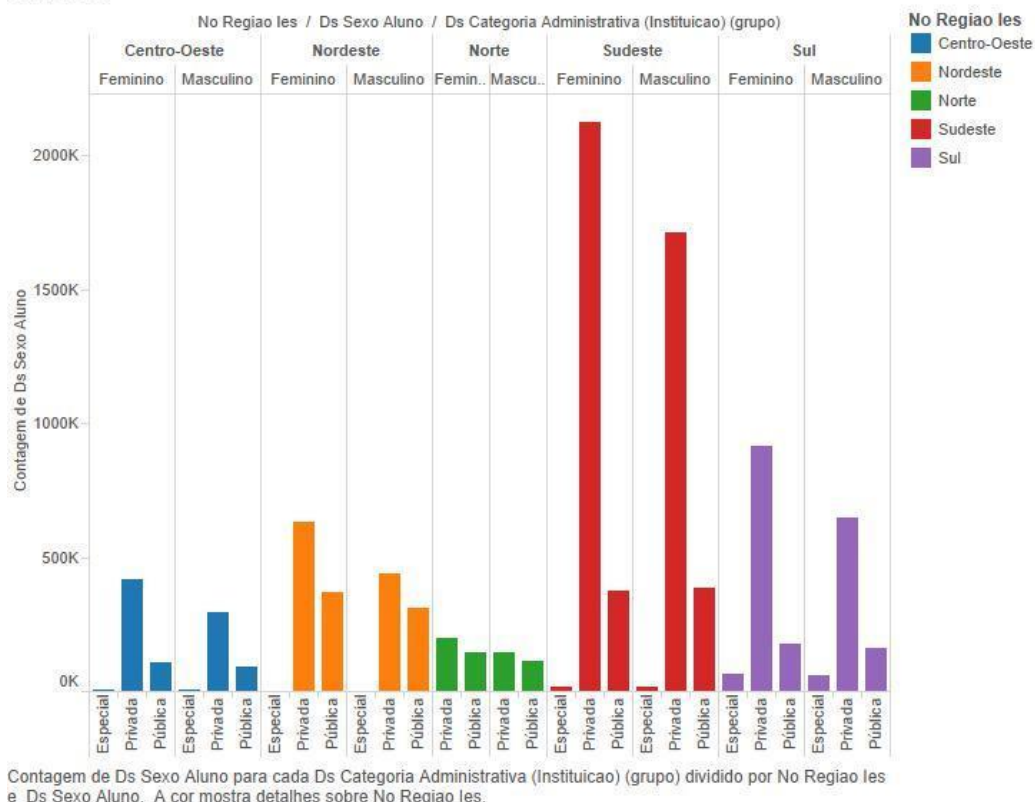


Figura 63 : Relatório 7 no Tableau.

8) Relatório de Alunos com deficiência ou superdotação na Universidade Federal do Paraná. Visualização de gráfico de barras com a quantidade de alunos por tipo de deficiência ou superdotação e uma tabela que mostra a quantidade de alunos com deficiência ou superdotação por curso da UFPR (Figuras 64, 65, 66 e 67).

### Tabela: Aluno

Atributos: Nome do curso e indicadores de aluno com deficiência ou superdotação, deficiência auditiva, deficiência física, deficiência intelectual, deficiência múltipla, surdez, surdocegueira, visão subnormal, cegueira ou superdotação.

### Código SQL

```
select no_curso,
       in_aluno_def_tgd_super,
       in_def_auditiva as "deficiência auditiva",
       in_def_fisica as "deficiência física",
       in_def_intelectual as "deficiência intelectual",
```

```

in_def_multipla as "deficiência múltipla",
in_def_surdez as "surdez",
in_def_surdocegueira as "surdocegueira",
in_def_baixa_visao as "visão subnormal",
in_def_cegueira as "cegueira",
in_def_superdotacao as "superdotação"

from aluno
where no_ies = 'UNIVERSIDADE FEDERAL DO PARANÁ'
and in_aluno_def_tgd_super = 1

```

### Microstrategy

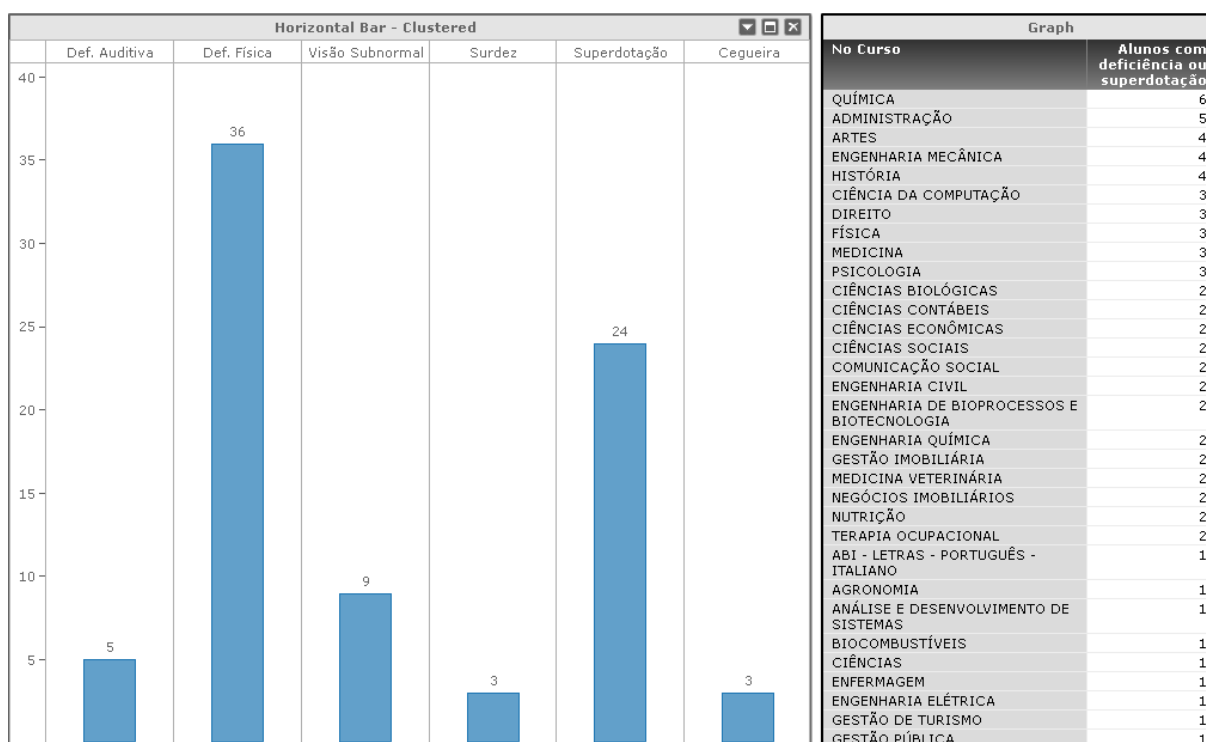


Figura 64 : Relatório 8 no Microstrategy.

## Qlikview

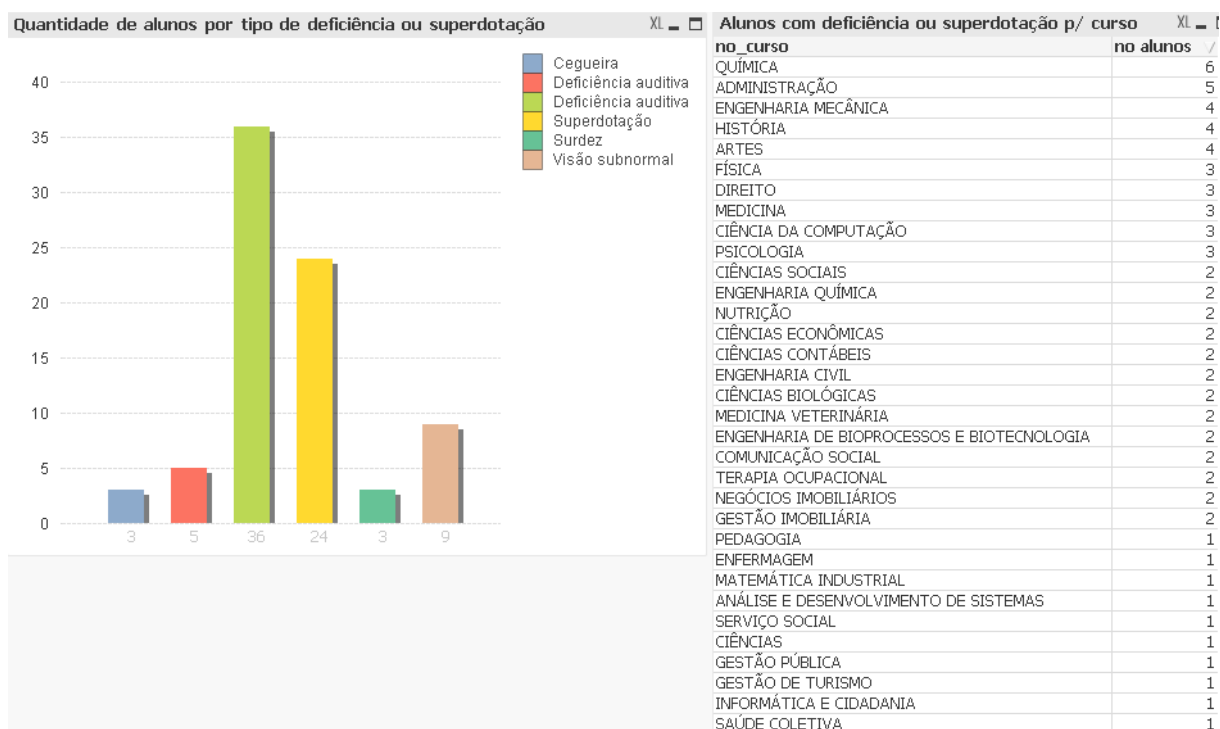


Figura 65 : Relatório 8 no QlikView.

## Power BI

Quantidade de alunos por tipo de deficiência ou superdotação

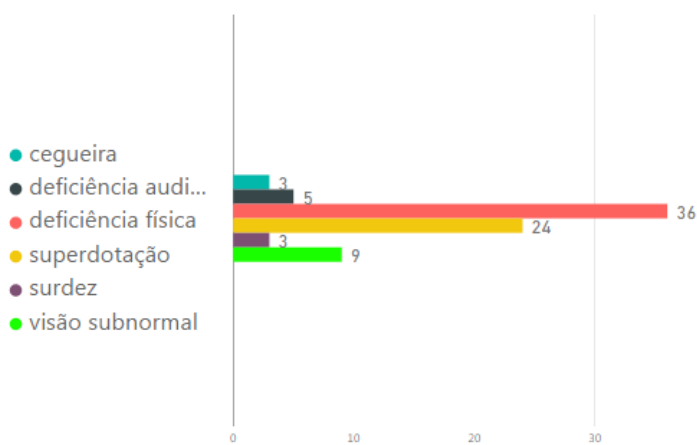


Figura 66 : Relatório 8 no QlikView.

## Tableau

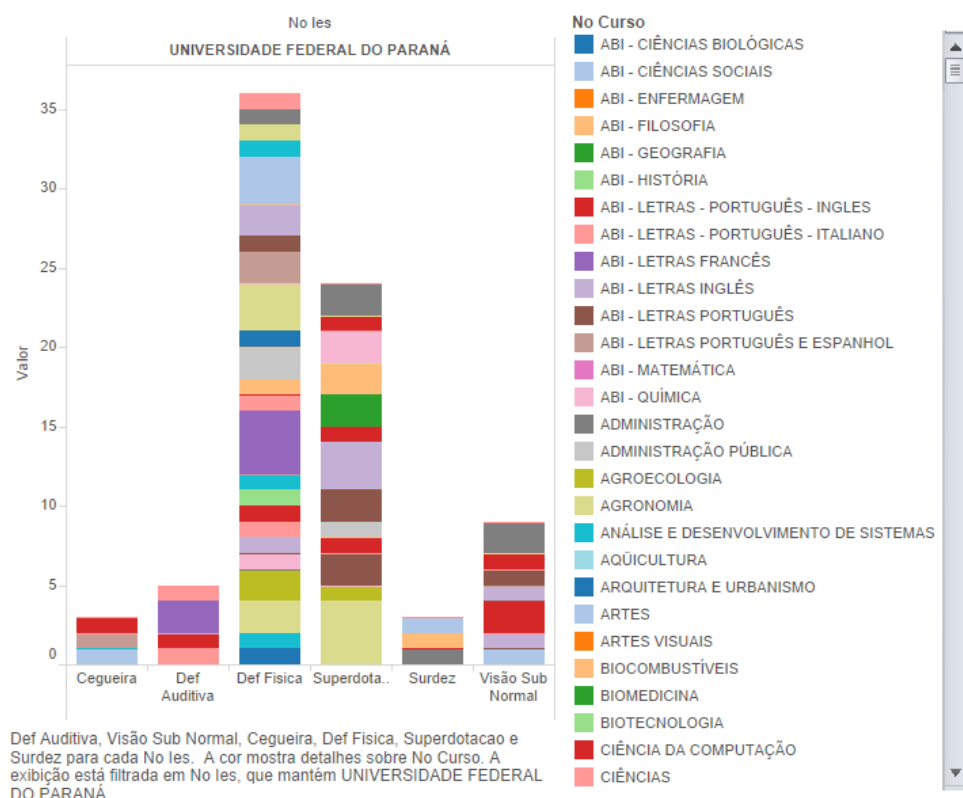


Figura 67 : Relatório 8 no Tableau.

## 4.5 ESPECIFICAÇÃO DO TESTE DE *DRILL-DOWN*

A avaliação de *drill-down* (detalhamento das informações) foi feita com base no relatório 4 (especificado na seção 4.4) que inicialmente, mostra a relação de escolaridade dos docentes por região do Brasil. A proposta é de fazer o *drill-down* a nível de Sul do Brasil, estado do Paraná, cidade de Curitiba e, finalmente, mostrar os dados somente da Universidade Federal do Paraná. Essa avaliação foi realizada nas 4 ferramentas selecionadas para se fazer uma comparação na forma como cada software executa operações de *drill-down*. Vale reforçar que a operação de *drill-down* facilita a análise de dados nos diferentes níveis.

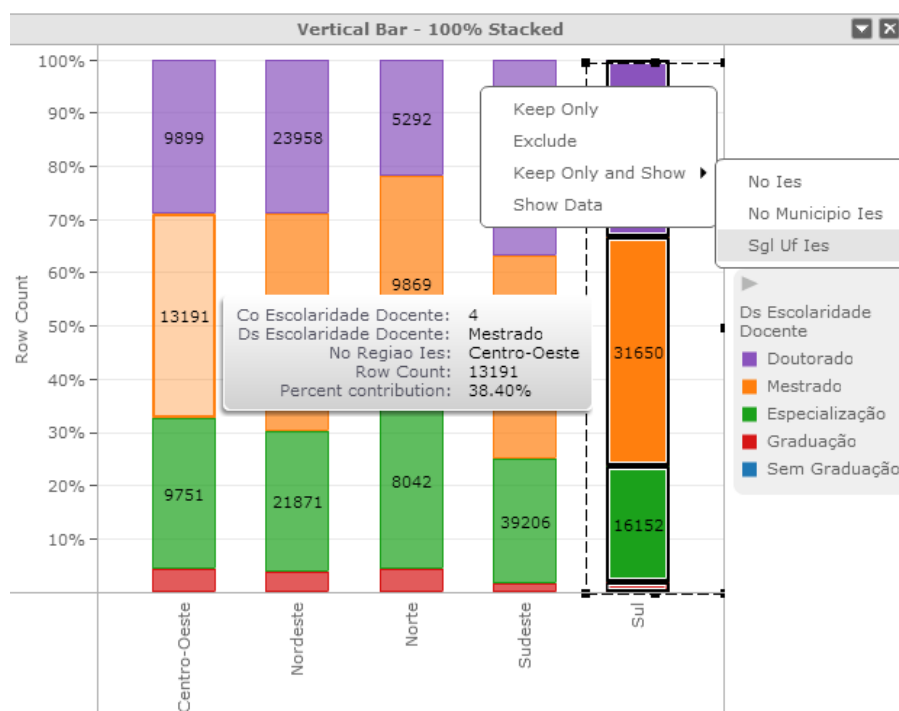
### 4.5.1 MICROSTRATEGY

A função de *drill-down* fica disponível ao selecionar a parte do gráfico a ser detalhada, basta clicar com o botão direito e utilizar a função *Keep Only and Show* (Manter apenas e mostrar). O *Microstrategy* mostra algumas opções de campos

disponíveis para *drill-down* como exemplificado na Figura 68.

A partir do gráfico resultante, pode-se repetir o processo até chegar no nível de detalhamento desejado. No exemplo da Figura 69 o próximo passo é detalhar por cidades do Paraná (Figura 70).

As imagens a seguir mostram cada passo do processo de *drill-down* utilizando o *Microstrategy*.



**Figura 68** : Drill-down no Microstrategy para detalhamento da região sul.

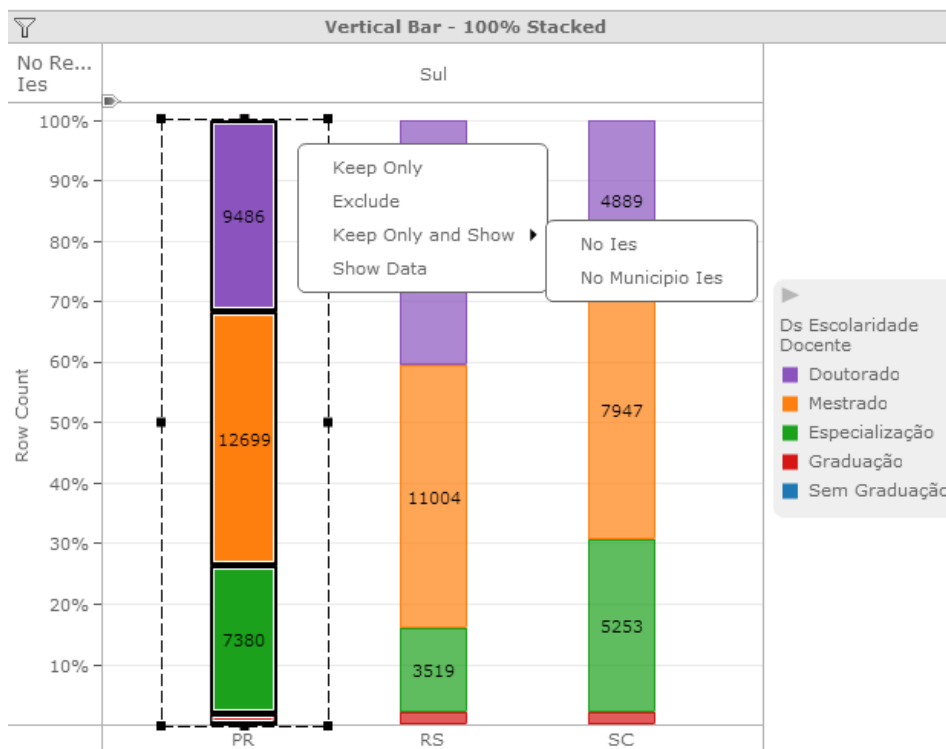


Figura 69 : Drill-down no Microstrategy. Detalhamento por estados da região sul.

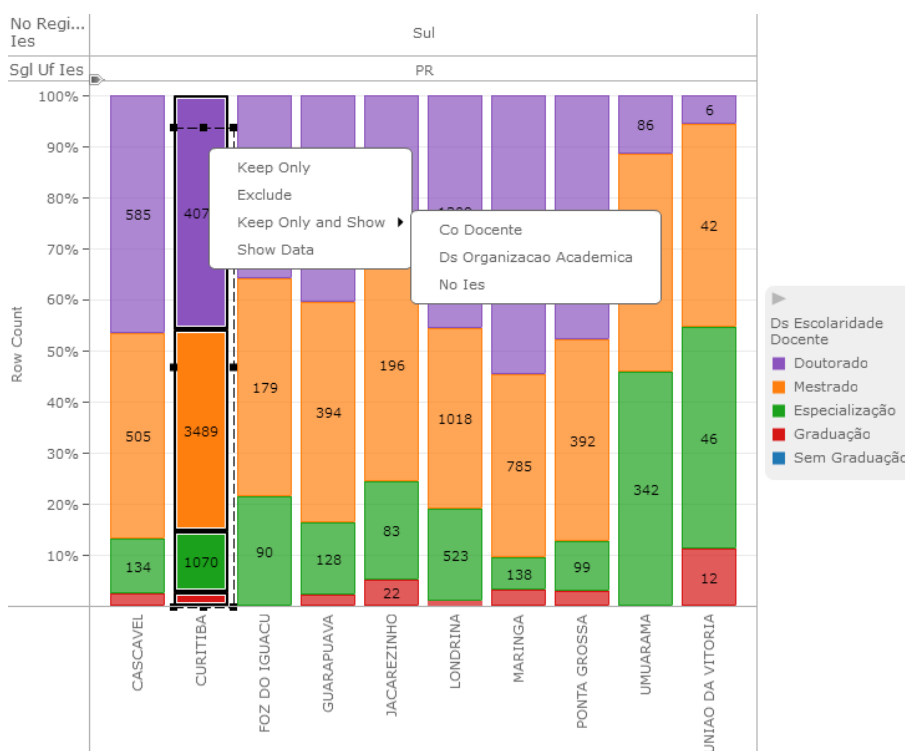


Figura 70 : Passo 3 de teste de drill-down no Microstrategy. Detalhamento por cidades do estado do Paraná.



Figura 71 : Drill-down no Microstrategy. Detalhamento por instituições de ensino de Curitiba.

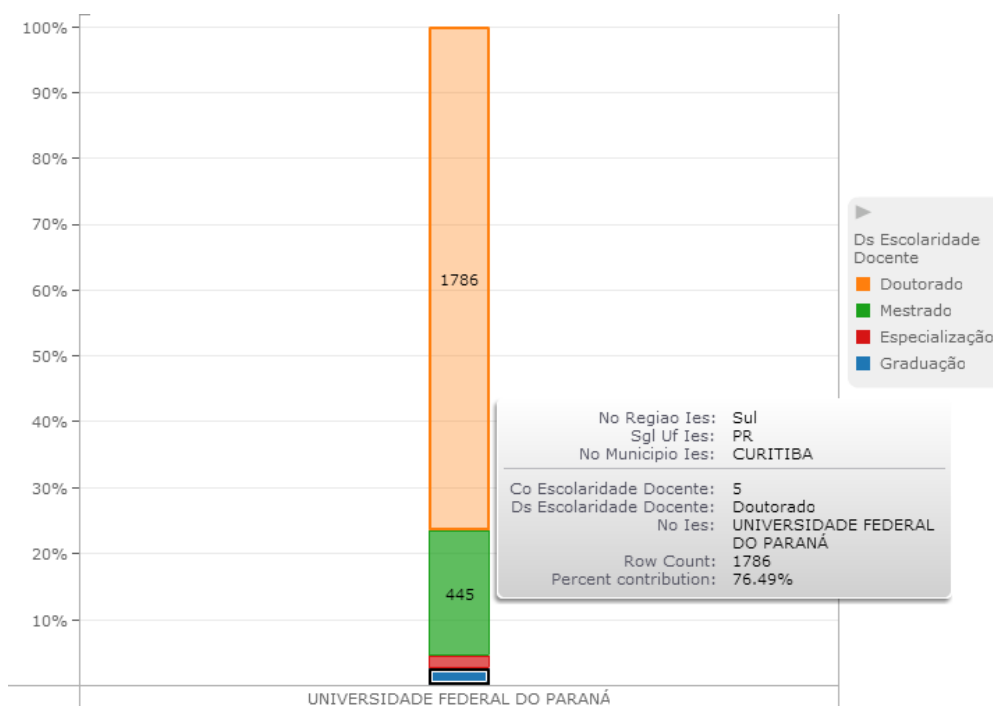
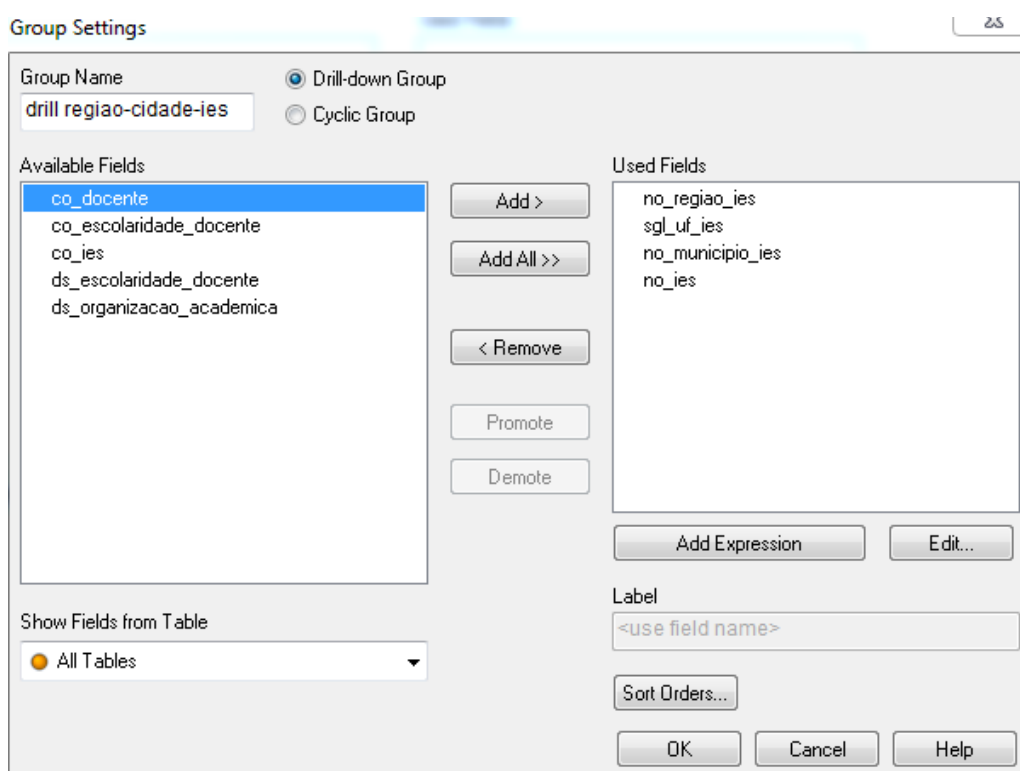


Figura 72 : Drill-down no Microstrategy mostrando dados específicos da Universidade Federal do Paraná.

Após todas as etapas o Microstrategy mostra em forma de gráfico a quantidade e porcentagem de docentes com doutorado na Universidade Federal do Paraná conforme Figura 72.

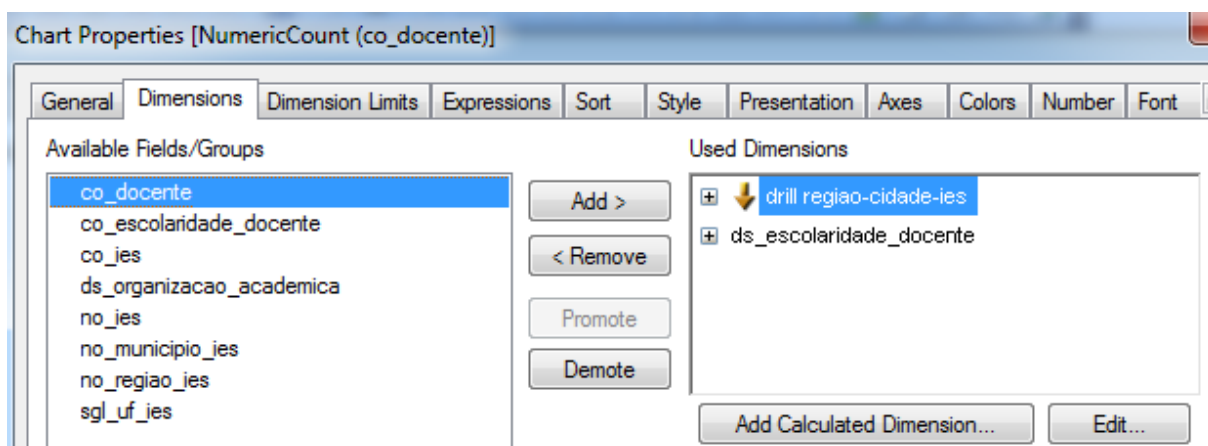
#### 4.5.2 QLIKVIEW

Para se utilizar os recursos de *drill-down* no *Qlik*, antes de montar o gráfico é necessário criar um grupo com a sequencia dos campos a serem utilizados. Um novo grupo pode ser criado na opção *Document Properties* disponível no menu principal. Ao se selecionar a opção de criar um novo grupo, é possível selecionar a opção de criar um campo do tipo *drill-down* e a sequencia dos campos que irão fazer parte desse grupo, conforme Figura 73:



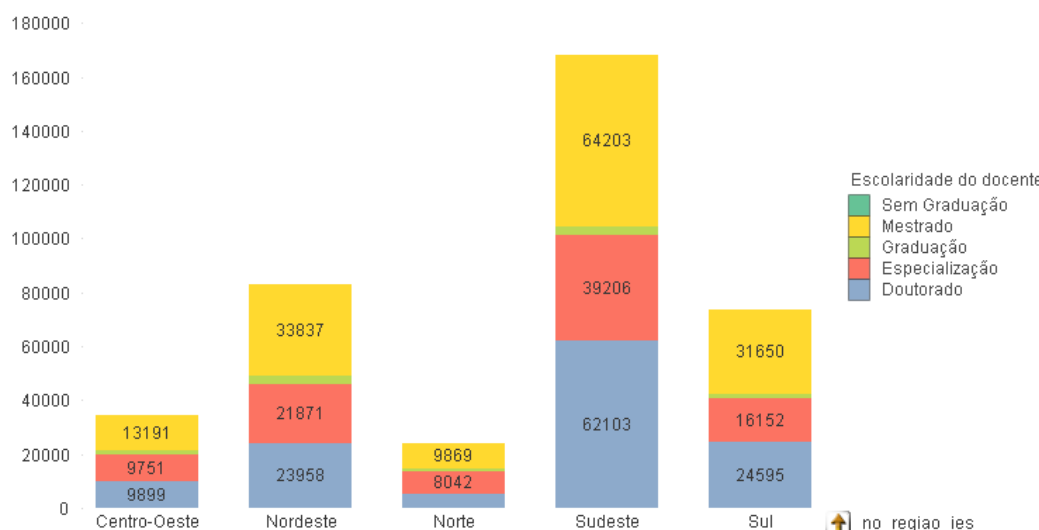
**Figura 73** : Drill-down no QlikView.

Com o grupo criado, podemos iniciar a construção do gráfico de barras. Ao se configurar as dimensões do gráfico, o grupo de *drill-down* previamente configurado fica disponível como sendo uma dimensão (Figura 74).



**Figura 74** : Drill-down no Qlikview como sendo uma dimensão do gráfico de barras.

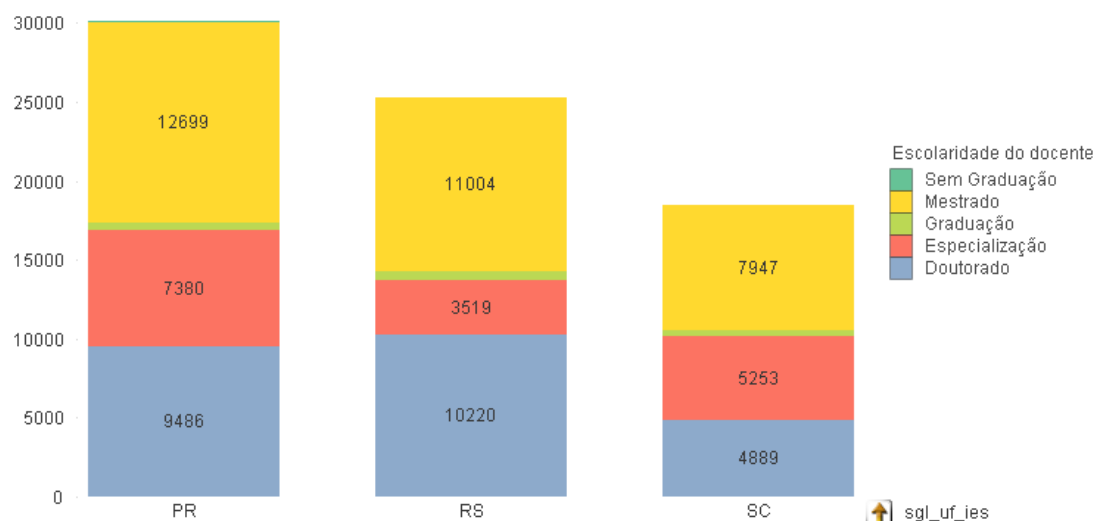
O resultado é um gráfico mostrando apenas a primeiro campo do grupo de *drill-down*, que no exemplo é o campo de região do Brasil (Figura 75).



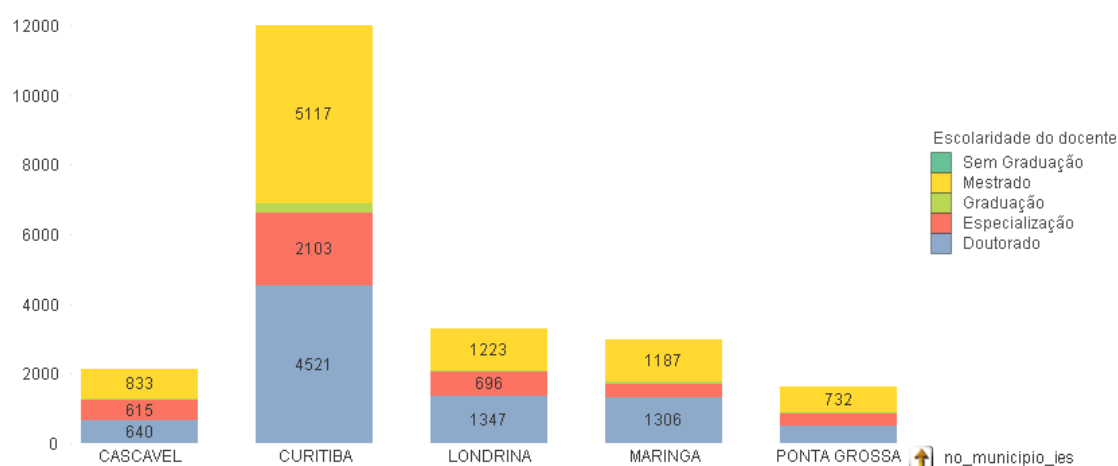
**Figura 75** : Gráfico inicial para teste de drill-down no QlikView.

A partir do gráfico, podemos navegar entre os níveis de *drill-down*, clicando na opção desejada dentro do gráfico. Para voltar um nível, deve-se utilizar a opção de *drill-up* representada no gráfico por uma flecha para cima. A Figura 75 mostra o resultado ao se clicar na opção Sul a partir do gráfico inicial.

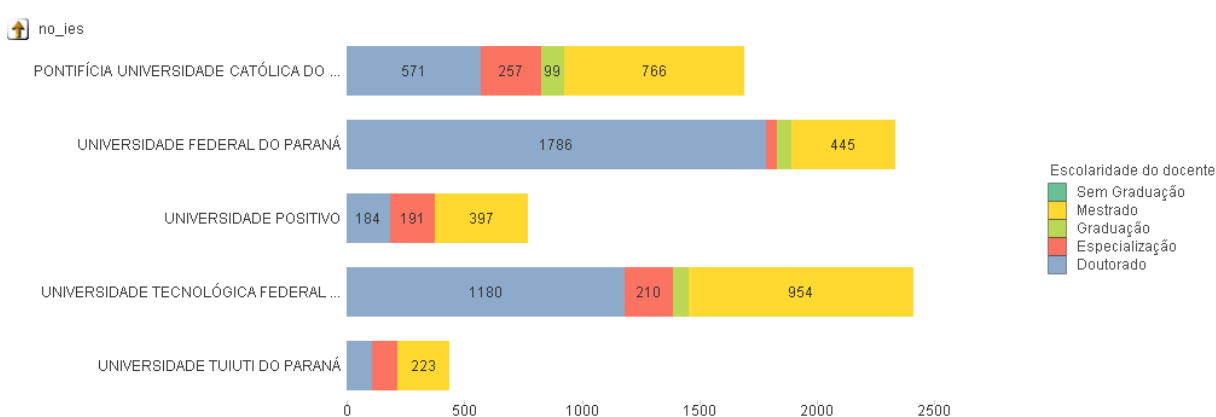
Pode-se selecionar sucessivamente as opções de *drill-down* até chegar no detalhe desejado que é a Universidade Federal do Paraná. As Figuras 76, 77 e 78 mostram os passos desse processo.



**Figura 76** : Drill-down no QlikView. Detalhamento por estados da região Sul do Brasil.

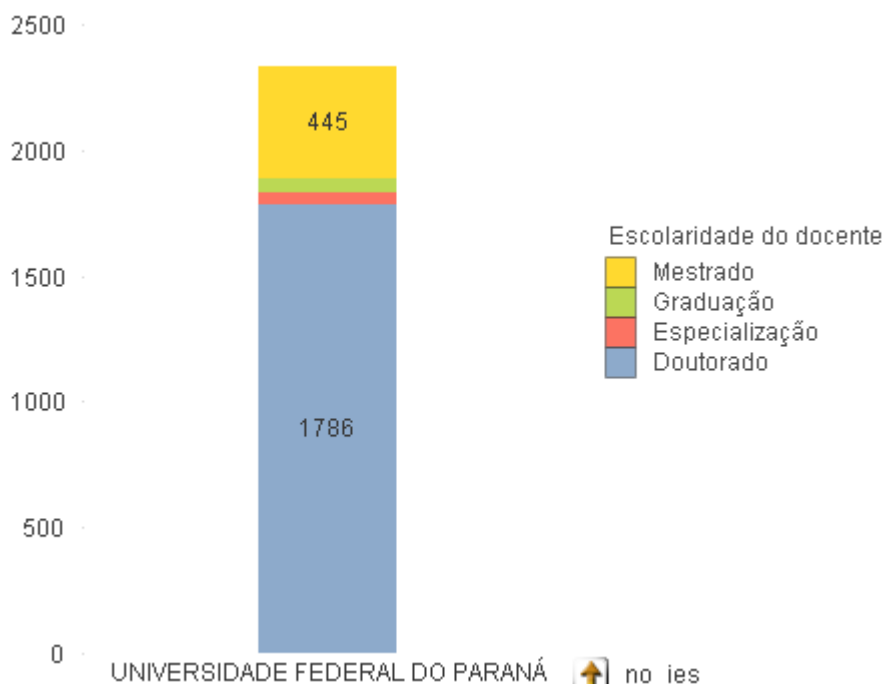


**Figura 77** : Drill-down no QlikView. Detalhamento por cidade do estado do Paraná.



**Figura 78** : Drill-down no QlikView. Detalhamento por instituições de ensino em Curitiba.

O gráfico com detalhamento por instituição de ensino (Figura 77), teve a orientação manualmente alterada com a finalidade de facilitar a leitura dos nomes das instituições.

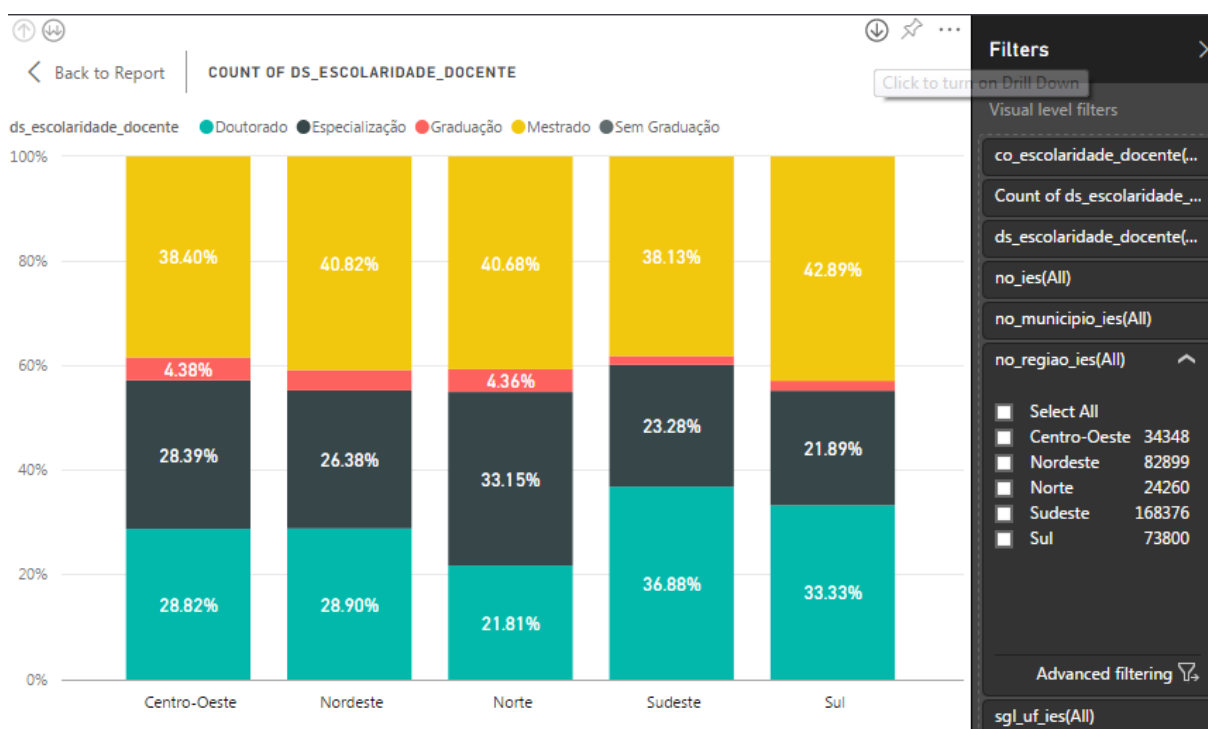


**Figura 79** : Drill-down no QlikView com dados específicos da UFPR.

O resultado final pode ser observado na Figura 79 que mostra o número de docentes com doutorado e mestrado na Universidade Federal do Paraná.

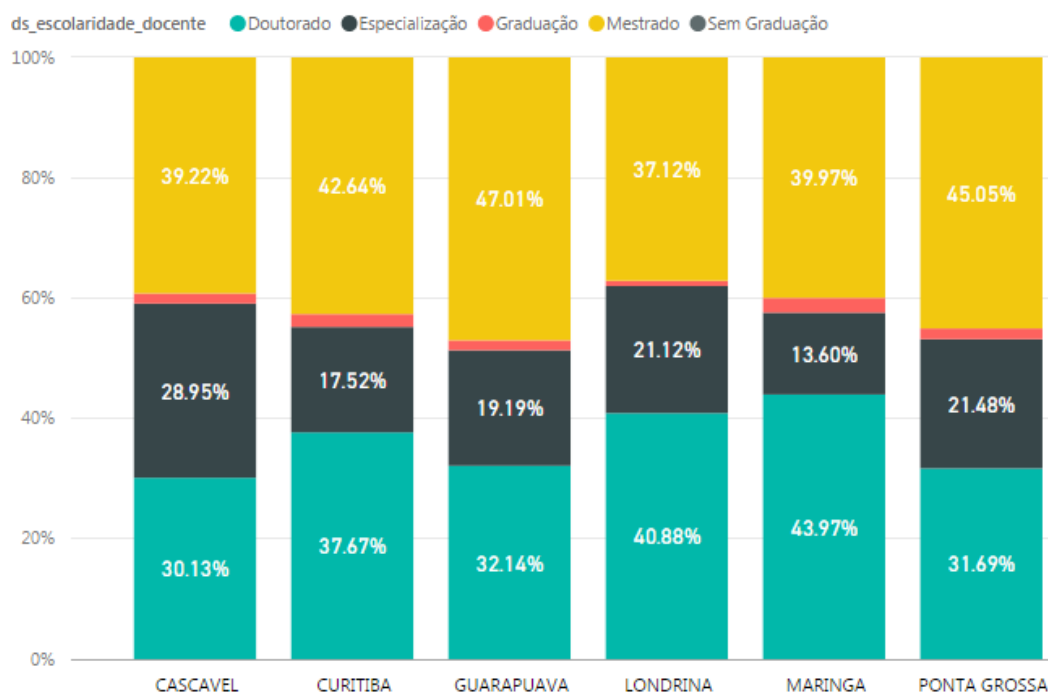
#### 4.5.3 POWER BI

A função de *drill-down* no *Power BI* pode ser ativada ao se clicar no botão específico para esta função, que é representado por uma flecha para baixo. Ao se utilizar essa opção, o *Power BI* faz um detalhamento dos dados de acordo com as colunas selecionadas. A sequência dos níveis de detalhamento segue a ordem em que as colunas são selecionadas durante a criação do gráfico. A seleção de dados pode ser feita utilizando os filtros disponíveis no menu lateral. A Figura 80 mostra o relatório inicial a ser realizado o teste.



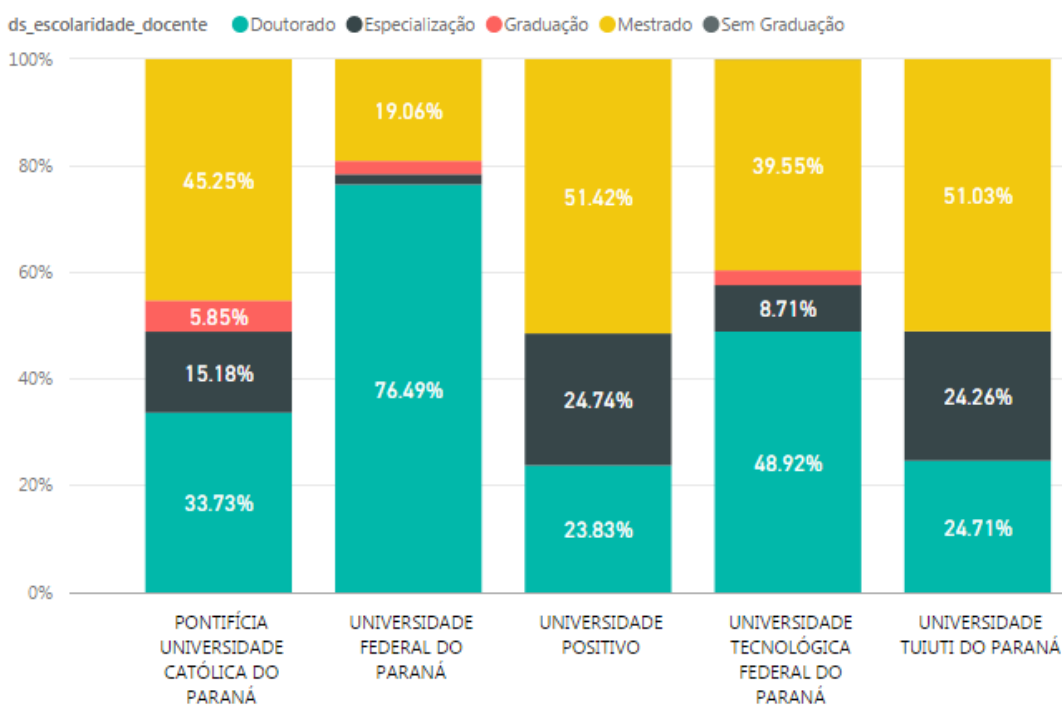
**Figura 80** : Relatório inicial para teste de drill-down no Power BI.

Nesse caso, para se conseguir dados específicos das cidades na região Sul do Brasil é necessário selecionar apenas a região sul e o estado do Paraná no menu lateral de filtros e em seguida utilizar a opção de *drill-down*. O resultado pode ser observado na Figura 81.

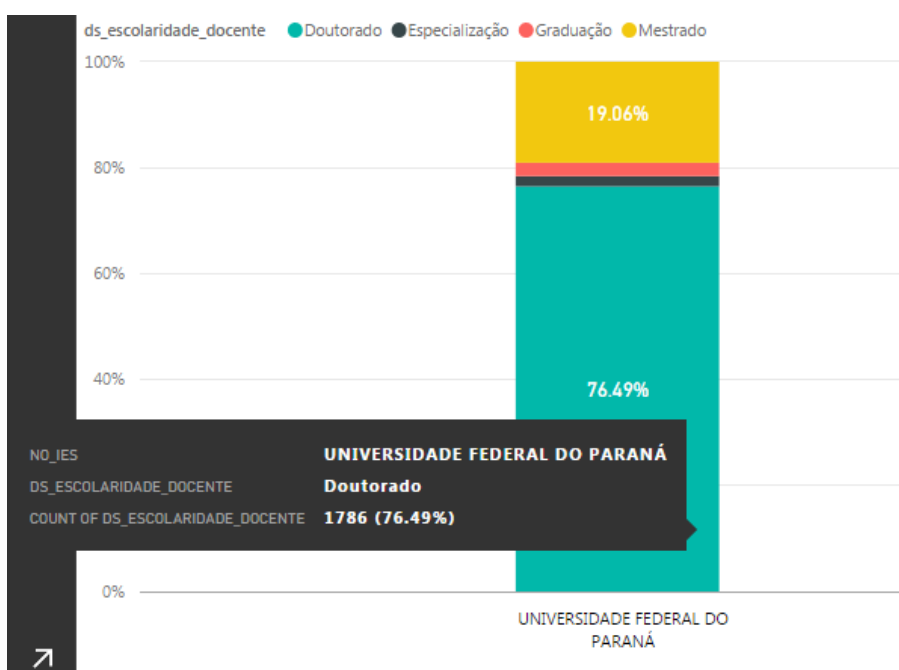


**Figura 81** : Drill-down no Power BI. Detalhamento por cidades do estado do Paraná.

Podemos repetir o processo de aplicação de filtros e operação de *drill-down* até que se alcance o nível de detalhamento desejado (Figura 82 e 83).



**Figura 82** : Drill-down no Power BI. Detalhamento por instituições de ensino de Curitiba.

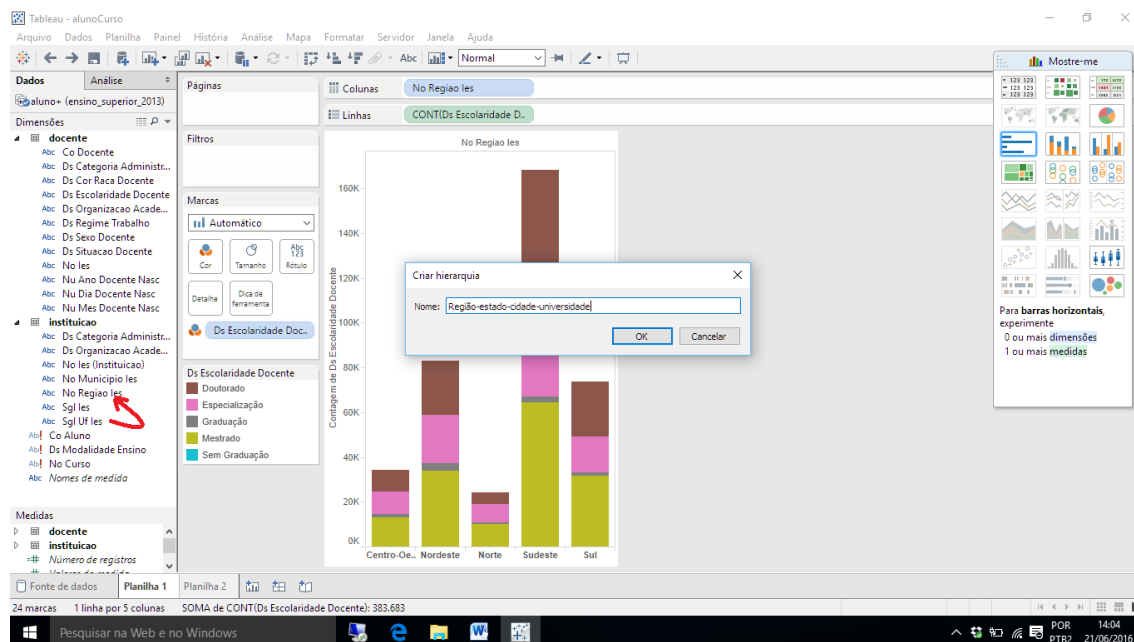


**Figura 83** : Drill-down no Power BI com dados específicos da Universidade Federal do Paraná

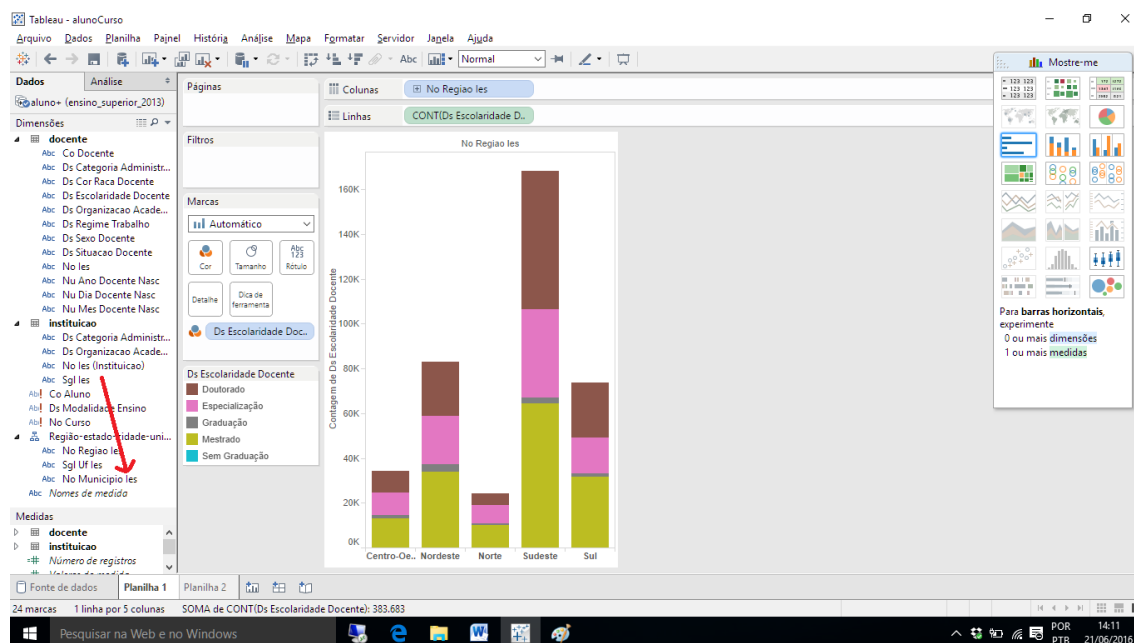
#### 4.5.4 TABLEAU

Para realizar o *drill-down* no *Tableau* primeiro é preciso criar hierarquias clicando e arrastando os campos que deseja, neste caso selecionamos `sgl_uf_ies` e arrastamos para `no_regiao_ies` e escolhemos o nome para a hierarquia (Figura 84);

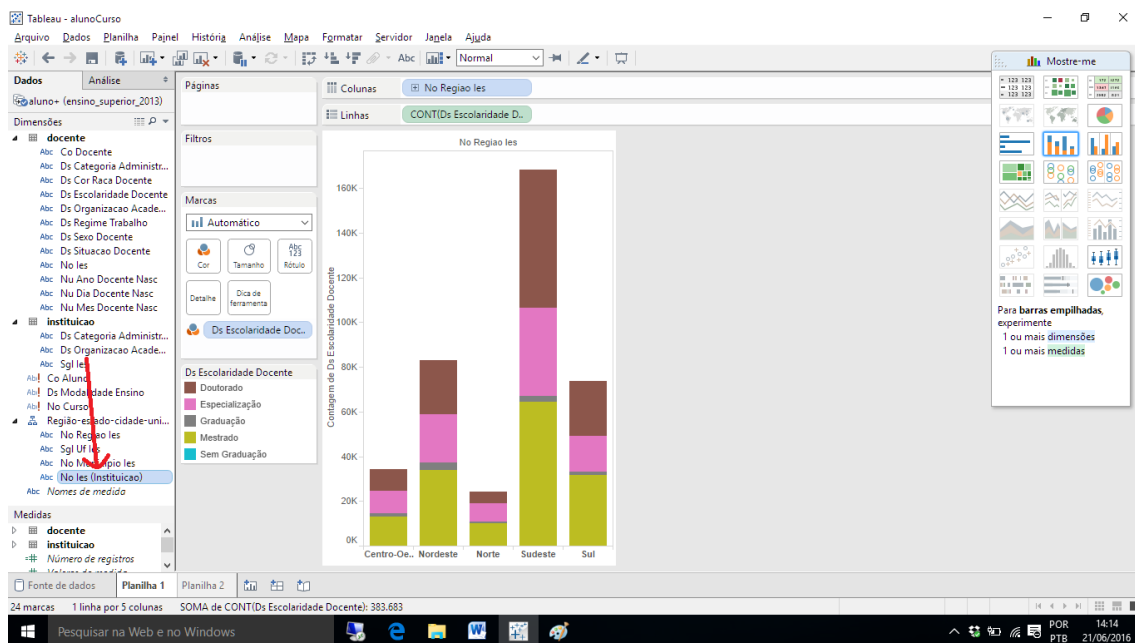
depois no\_municipio\_ies para a hierarquia que nomeamos como Região-estado-cidade-universidade (Figura 85); por último seleccionamos no\_ies para a hierarquia (Figura 86).



**Figura 84** : Drill-down no Tableau com a criação da hierarquia Região – Estado – Cidade - universidade.

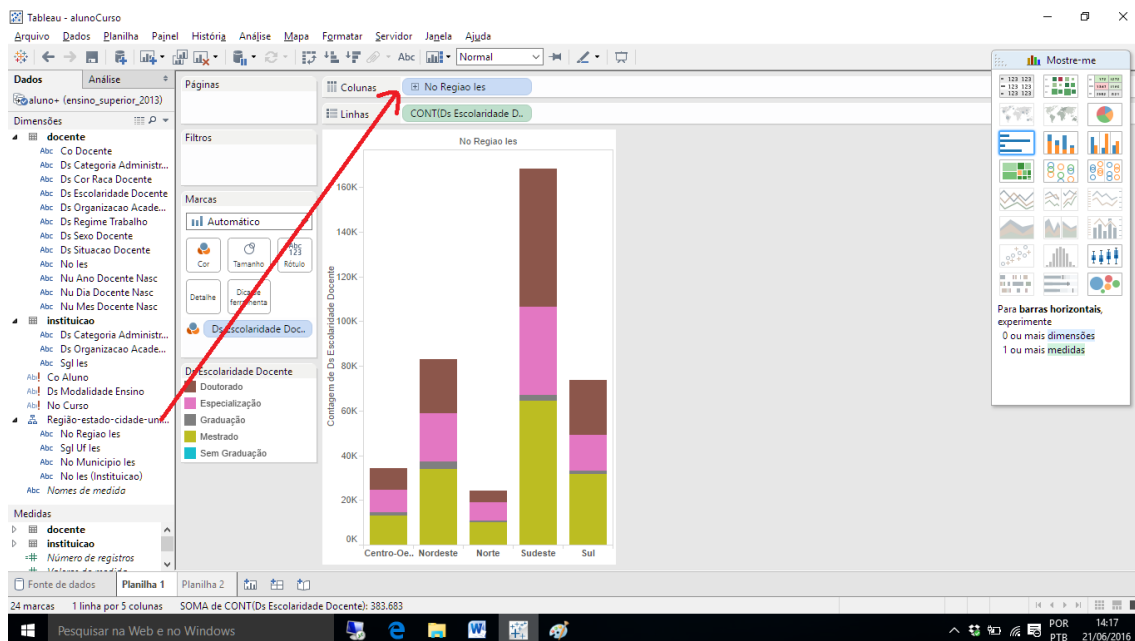


**Figura 85** : Drill-Down no Tableau com a inclusão do município na hierarquia.



**Figura 86** : Drill-down no Tableau com a inclusão do nome da instituição acadêmica na hierarquia.

Com a hierarquia criada, basta levá-la até o campo desejado no gráfico (Figura 87). Para observarmos a hierarquia, podemos, como no exemplo, selecionar uma região e selecionar a opção “manter apenas” (Figura 88). Depois disso, clicando no ícone de expansão (+) no campo hierárquico (Figura 89) ele mostrará o próximo nível. Usando esses passos para os próximos itens conseguimos chegar até a Universidade Federal do Paraná (Figuras 90, 91, 92 e 93).



**Figura 87**: Inclusão do campo hierárquico no campo coluna do gráfico.

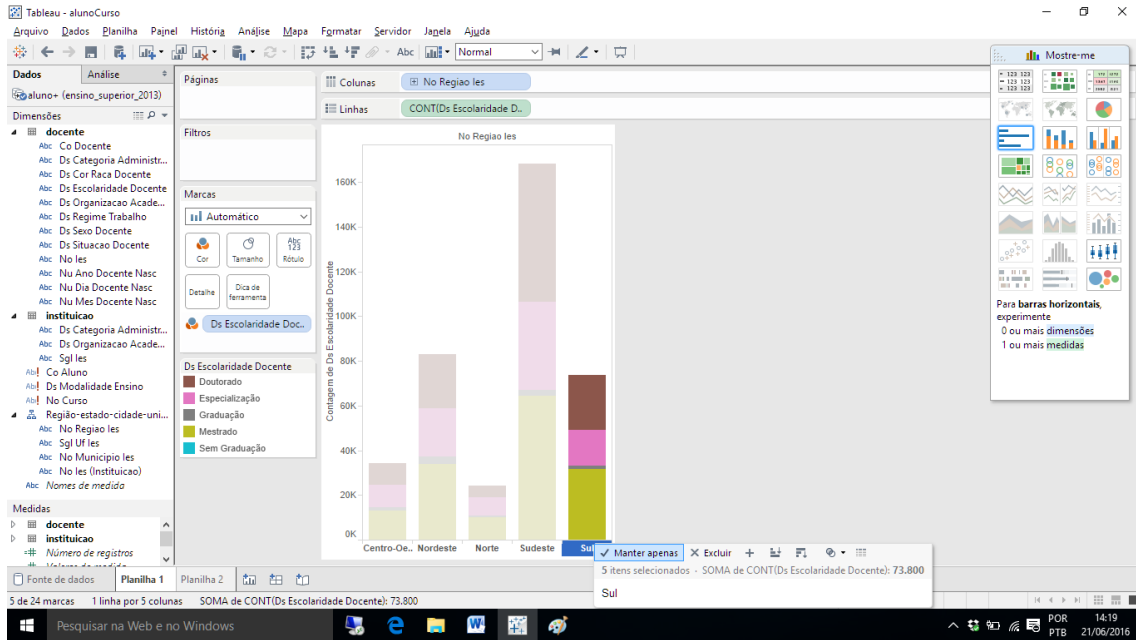


Figura 88 : Drill-down no Tableau filtrando a região Sul.

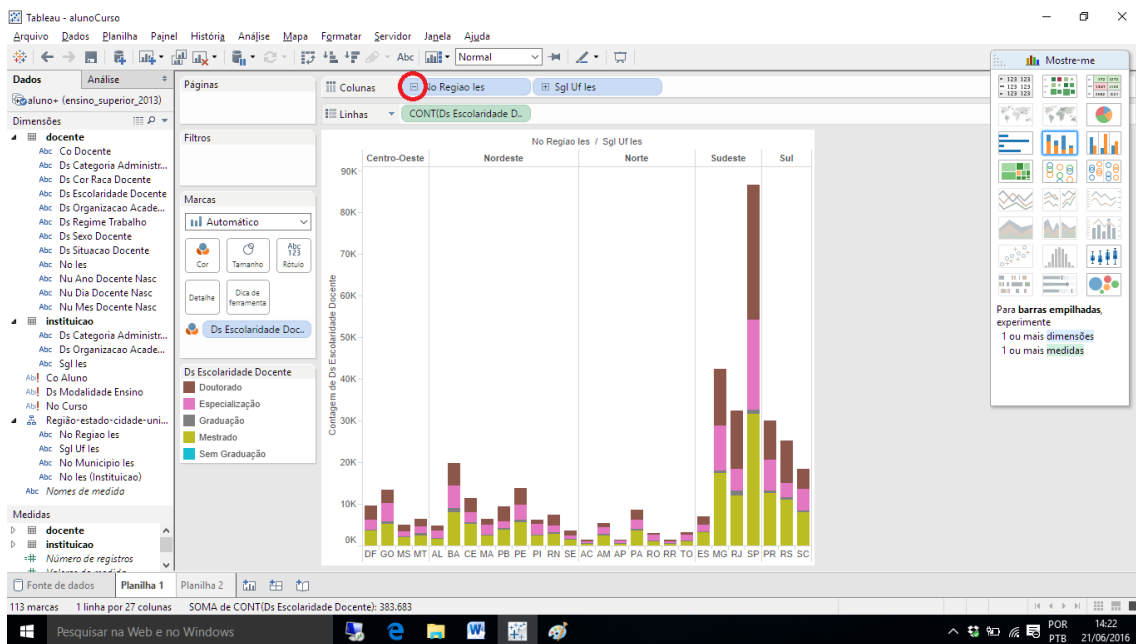


Figura 89 : Drill-down no Tableau . Expandindo a hierarquia região – estado.

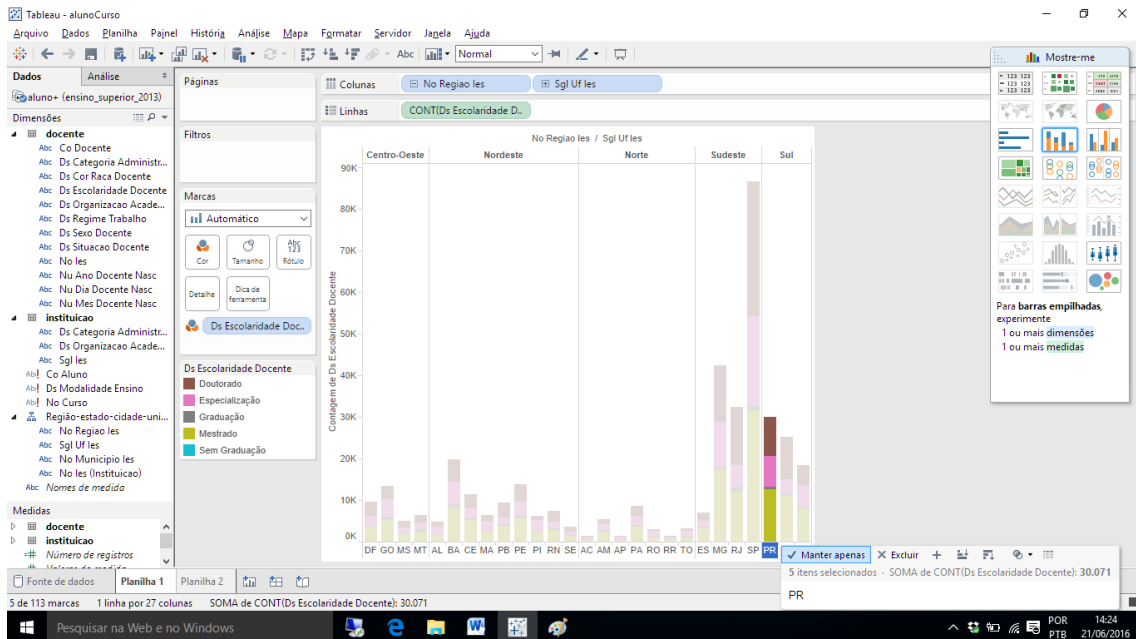


Figura 90 : Drill-Down no Tableau filtrando o estado do Paraná.

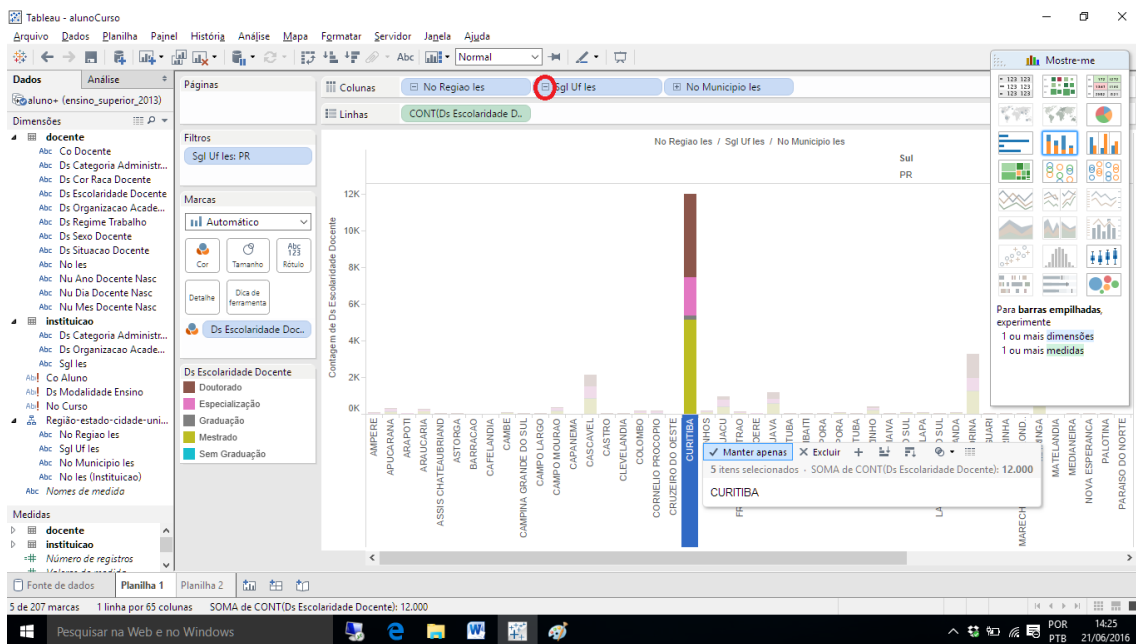


Figura 91 : Drill-down no Tableau. Expandindo a hierarquia estado – município e filtrando Curitiba.

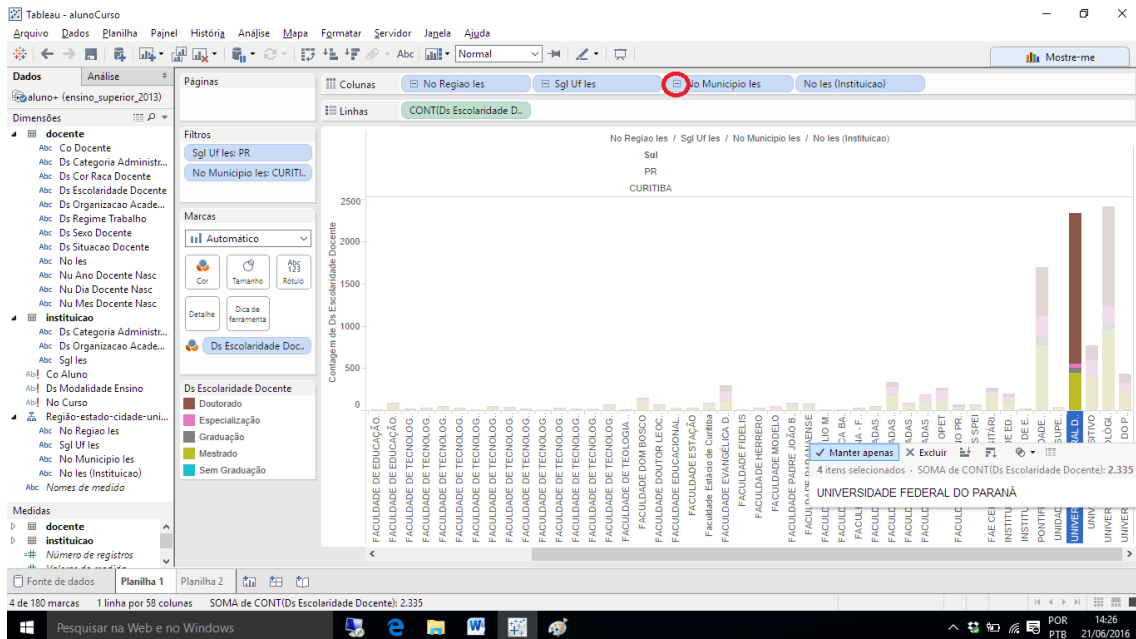


Figura 92 : Drill-down no Tableau. Expandindo a hierarquia cidade – universidade e filtrando a Universidade Federal do Paraná.

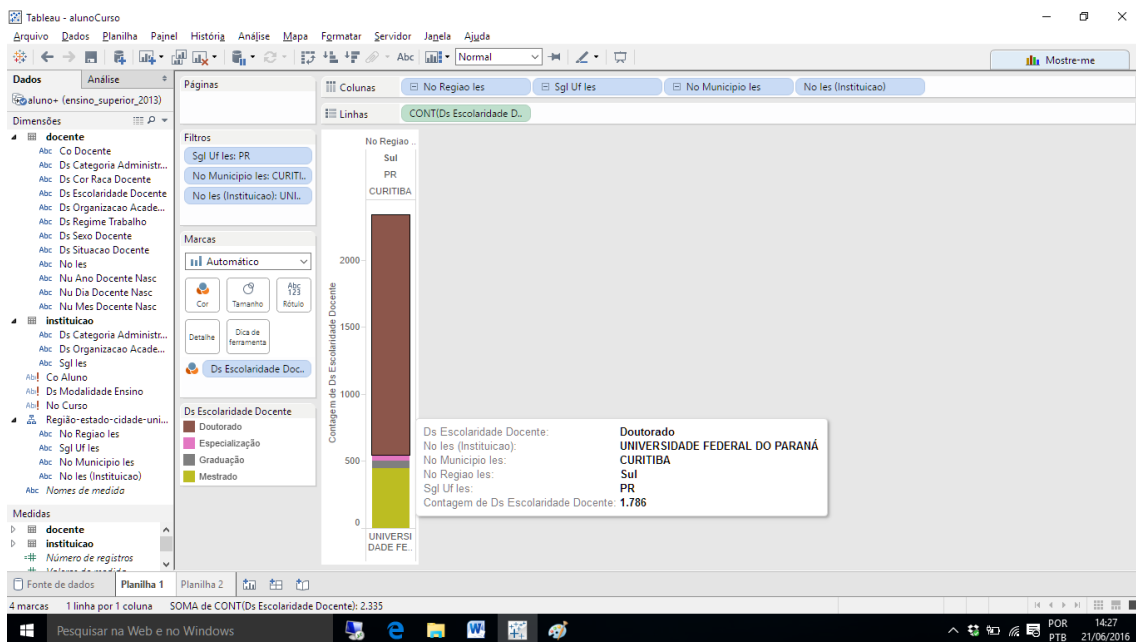


Figura 93 : Drill-down do Tableau.

## Planilha 1

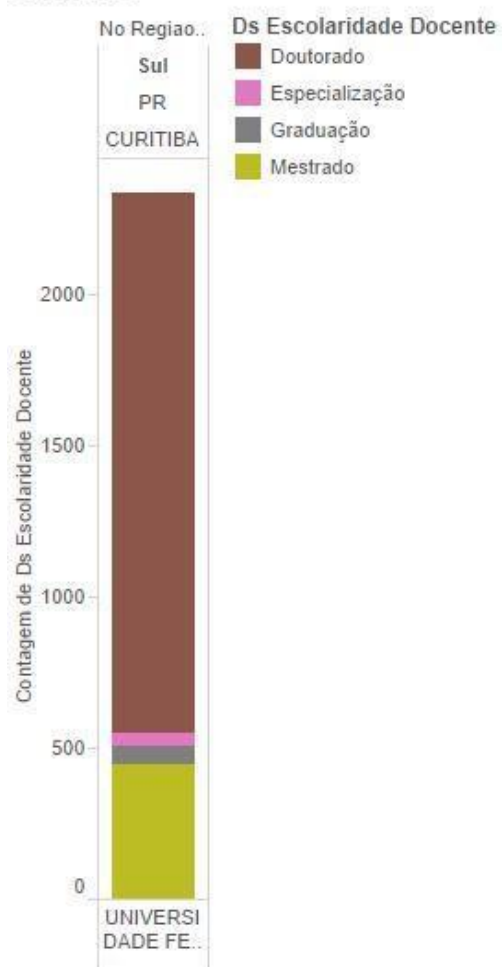


Figura 94 : Drill-down do Tableau – nível UFPR.

## 5. ANÁLISE DOS RESULTADOS

Em conclusão às análises podemos observar as seguintes características das ferramentas avaliadas.

Nos critérios Básicos (1) o Desempenho (1.1) do *Tableau* foi bom somente depois de efetuada a extração dos dados, pois a criação de relatórios operacionais com grande quantidade de registros utilizando SQL levava cerca de 140 minutos para gerar cada relatório, porém com a extração passou a demorar menos de 12 segundos. Mas caso seja necessário alterar ou adicionar novos campos à pesquisa, terá que esperar a criação do novo arquivo de extração, que é em torno de 90 minutos.

Então é mais vantajoso selecionar as tabelas inteiras e na hora de gerar o relatório no *Tableau* utilizar apenas os campos desejados. Nas demais ferramentas, por não necessitar da extração, o desempenho foi bom, pois cada relatório levava em média 9 à 15 minutos para ser criado utilizando o código SQL, o que às torna melhores para pesquisas consecutivas com bases diferentes. Porém, para um estudo envolvendo várias pesquisas diferentes dentro da mesma base de dados, o *Tableau* se torna mais ágil.

Já em Consultas ADHOC (1.2), o *Microstrategy* apresentou um bom resultado, pois ele oferece mais opções de edição para que o usuário consiga ver as informações da forma que deseja. Nas outras ferramentas avaliadas a construção de gráficos ficava limitada de acordo com os dados escolhidos, muitas vezes não se adequando à vontade do usuário.

Com relação a plataforma (1.3) apenas o *Power BI* e o *Tableau* não atingiram todos os sistemas operacionais de grande escala (Windows, Mac, Linux). Já no Subcritério suporte técnico e documentação (1.4) todas as ferramentas possuem documentação disponível para usuários.

No critério Relatórios (2) nenhuma ferramenta apresentou a opção de agendamento (2.1) na versão analisada. Já no subcritério *Dashboards* (2.2) todas as ferramentas apresentaram a opção de criar e gerenciar *dashboards* de forma intuitiva, tornando fácil a transição entre elas. Também no subcritério Recursos de Navegação (2.3) todas as ferramentas atenderam satisfatoriamente, sendo possível executar operações de *drill-down*, *drill-up*, *slice and dice*.

Com relação a ferramenta *Power BI*, esta necessita que o usuário publique seu relatório no *Power BI online*, mas devido ao processo ser de fácil execução, não deprecia a avaliação da ferramenta. Vale colocar que todas as ferramentas disponibilizam a opção de exportação para os formatos digitais mais utilizados no mercado, atendendo satisfatoriamente o requisito “Exportação para outros formatos” (2.4).

As Ferramentas apresentaram, no Critério “Funcionamento WEB (3) um bom resultado, pois possuem Fóruns *online* (3.1) e *Help online* (3.2), tendo apenas uma diferença no critério de suporte a dispositivos móveis (3.3), pois das ferramentas avaliadas somente o *Power BI* apresenta suporte nas versões avaliadas. As outras ferramentas possuem essas funcionalidades em outras versões que não foram avaliadas neste trabalho.

Ao avaliarmos a Usabilidade (4), verificamos que no *QLIK View* a facilidade de uso (4.1) e atratividade (4.2) é comprometida devido à obrigação de utilizar *wizard* em cada ação do usuário. Nas demais ferramentas o usuário consegue realizar ações com a utilização do mouse, clicando, arrastando e selecionando as informações de forma simples e intuitiva. Já referente a customização da interface apenas o *Qlik* e o *Tableau* permitiram a movimentação de menus e tabelas para se adequar à visão do usuário (interface personalizável 4.3)

No critério produto (5), As ferramentas apresentam um bom custo benefício (5.1), trabalhando com licenças por usuário a empresa pode estabelecer o custo de acordo com o número de funcionários que deseja alocar para a utilização da ferramenta. No subcritério mercado (5.2) vimos que todas as ferramentas são utilizadas por empresas de grande porte, exceto o *Power BI* que não foi avaliado, pois não divulga o nome de seus clientes.

Todas as ferramentas apresentaram uma boa integração (6.1) com os tipos fontes de dados mais utilizados no mercado como *Postgres*, *Mysql*, arquivos no formato txt e csv, ODBC, entre outros. Todos apresentaram integração com o *Office* podendo importar e exportar arquivos de dados e gráficos (6.2).

Não pudemos avaliar as ferramentas no critério Segurança (7), pois essas opções só estão disponíveis em versões *Pro* ou *Enterprise*.

Conforme critérios estabelecidos na sessão 4.1, a Tabela 14 demonstra os resultados da análise deste trabalho

<b>Tabela 14:</b> Tabela comparativa entre as ferramentas				
<b>ID</b>	<b>POWER BI</b>	<b>QLIK VIEW 11</b>	<b>MICROSTRATEGY</b>	<b>TABLEAU</b>
1,1	B - apresentou tempo de espera de 8 a 9 minutos para consultas de 9 milhões de registros.	B- apresentou tempo de espera entre 3 e 4 minutos para consultas de 9 milhões de registros	B- apresentou tempo de espera entre 7 a 8 minutos para consultas de 9 milhões de registros	RE - O tableau é a única ferramenta avaliada que possui a opção de "conexão tempo real", que demora em média 140 minutos para cada pesquisa. Utilizando a extração padrão ( mesma arquitetura das outras ferramentas) o programa demora 80 - 90 minutos para a criação da extração e depois disso as pesquisas se tornam quase instantaneas (1-2 segundos)
1,2	RE - Dependendo do tipo de gráfico escolhido, a ferramenta apresenta limitações na forma de apresentar o resultado.	RE - O usuário necessariamente deve inserir as dimensões através de um <i>wizard</i> <sup>14</sup> , o que torna o processo mais demorado e trabalhoso.	B - A ferramenta sugere várias opções para as análises. A criação dos gráficos e tabelas é fácil, porém o usuário precisa conhecer a estrutura do gráfico e a forma como a ferramenta trabalha com as variáveis.	RE – A ferramenta limita as opções de gráficos de acordo com os dados escolhidos e pode fazer com que os modelos disponíveis não se adaptem às necessidades do usuário.
1,3	RE - Apenas para Windows	B - Suporta windows, MAC e Linux.	B - Suporta windows, MAC e Linux.	RE - Apenas Windows e MAC.
1,4	B - Documentação sobre instalação e suporte técnico disponível para todos em sua página.	B - Documentação sobre instalação e suporte técnico disponível para todos em sua página. Precisa estar cadastrado	B - Documentação sobre instalação e suporte técnico disponível para todos em sua página.	B - Documentação sobre instalação e suporte técnico disponível para todos em sua página.
2,1	N/E - Não permite agendamento	N/E - Não permite agendamento	N/E - Não permite agendamento	N/E - Não permite agendamento
2,2	B - Muito fácil de criar dashboards a partir de relatórios já existentes.	B - Todos os relatórios são criados dentro de dashboards.	B - Todos os relatórios são criados dentro de dashboards.	B - Muito fácil de criar dashboards a partir de relatórios já existentes.

14

Wizard - Interface gráfica do usuário para realizar tarefas através do esquema passo-a-passo. Fonte: Autor.

2,3	B - Para realizar estas operações os relatórios precisam ser publicados no Power BI online, porém a publicação é fácil e rápida.	B - Atende o requisito	B - Atende o requisito	B - Atende o requisito
2,4	B - É possível exportar para PDF, Excel, imagem e formatos HTML.	B - É possível exportar para PDF, Excel e imagem.	B - É possível exportar para PDF, Excel e imagem.	B - É possível exportar para PDF, Imagem, Excel e MDB.
3,1	B - Possui fórum oficial no site da ferramenta e outros fóruns independentes.	B - Possui fórum oficial no site da ferramenta e outros fóruns independentes.	B - Possui fórum oficial no site da ferramenta e outros fóruns independentes.	B - Possui fórum oficial no site da ferramenta e outros fóruns independentes.
3,2	B - Atende o requisito	B - Atende o requisito	B - Atende o requisito	B - Atende o requisito
3,3	B - Suporta Android, iOS e Windows phone.	RE - Possui esta funcionalidade para qualquer dispositivo móvel com suporte à HTML5, porém só está disponível para implementações na versão Qlikview server.	RE - Suporte Android e IOS porém deve-se adquirir a licença.	RE - Possui a funcionalidade para android e IOS, porém só está disponível para implementações na versão Tableau server.
4,1	B - Atende o requisito	RE - Devido a suas funcionalidades serem selecionadas por wizard, torna o processo menos intuitivo.	B - Atende o requisito	B - Atende o requisito
4,2	B - Atende o requisito	RE - Não é tão intuitivo quanto as outras ferramentas analisadas.	B - Atende o requisito	B - Atende o requisito
4,3	NE - Não é possível a customização	B - Interface de menus customizável	NE - Não é possível a customização	B - Interface de menus customizável
5,1	B - Menor custo de implementação das ferramentas analisadas com a versão Pro custando 9,99 dólares por usuário por mês.	B - 1.395 dólares por usuário.	RE - 5.000 dólares por usuário.	RE - 2.000 dólares por usuário.

5,2	N/E - A Microsoft não informa nomes de clientes	B - Possui clientes notáveis como Bestbuy, Canon, Panasonic, Shell, entre outros.	B - Possui clientes notáveis como Facebook, LinkedIn, Netflix, Yahoo, entre outros.	B - Possui clientes notáveis como Audi, Citibank, The NewYork Times, Amazon, entre outros.
6,1	B - Permite o carregamento de fontes de dados diversas.	B - Permite o carregamento de fontes de dados diversas.	B - Permite o carregamento de fontes de dados diversas.	B - Permite o carregamento de fontes de dados diversas.
6,2	B - Dados e gráficos podem ser exportados para Excel	B - Dados e gráficos podem ser exportados para Excel	B - Dados e gráficos podem ser exportados para Excel	B - Dados e gráficos podem ser exportados para Excel
7,1	NE - Disponível apenas na versão Pro	NE - Disponível apenas na versão Enterprise	NE - Disponível apenas na versão Enterprise	NE - Disponível apenas na versão Tableau server
7,2	NE - Disponível apenas na versão Pro	NE - Disponível apenas na versão Enterprise	NE - Disponível apenas na versão Enterprise	NE - Disponível apenas na versão Tableau server
7,3	NE - Disponível apenas na versão Pro	NE - Disponível apenas na versão Pro	NE - Disponível apenas na versão Enterprise	NE - Disponível apenas na versão Tableau server

**Tabela 15:** Análise dos resultados com base nos níveis organizacionais.

		Power Bi	QlikView	Microstrategy	Tableau
Básicos	Desempenho	B	B	B	B
	Consultas ADHOC	B	Re	Re	Re
	Plataforma	R	B	B	Re
	Suporte Técnico/Documentação	B	B	B	B
Relatórios	Agendamento	N/E	N/E	N/E	N/E
	Dashboards	B	B	Re	Re
	Recursos de navegação	B	B	B	B
	Exportação de arquivos para outros formatos	B	B	B	B
Web	Fóruns	B	B	B	B
	Help Online	B	B	B	B
	Suporte a dispositivos móveis	B	N/E	N/E	N/E
Usabilidade	Facilidade de uso	B	Re	B	B
	Atratividade	B	Re	B	B
	Interface personalizável	N/E	B	N/E	B
Produto	Custo de licença	B	B	Re	Re
	Mercado	N/E	B	B	B
Planejamento	Carregamento de arquivos de diversas fontes	B	B	B	B
	Integração com o office	B	B	B	B
Segurança	Perfil do Usuário	N/E	N/E	N/E	N/E
	Aplicação de senha na fonte de dados	N/E	N/E	N/E	N/E
	Log de Auditoria	N/E	N/E	N/E	N/E

**Tabela 16:** Tabela de totalização dos resultados com base nos níveis organizacionais.

Ferramenta	Contagem de bom	Contagem de Regular	Contagem de Ruim	Contagem de N/e	Nível predominante
QlikView	12	3	0	4	Operacional
Power Bi	14	0	1	5	Tático
Microstrategy	10	3	0	6	Tático
Tableau	12	4	0	5	Estratégico

Durante os testes pode-se destacar algumas particularidades de cada ferramenta analisada

O *MicroStrategy* é um aplicativo mais fácil de usar para um cliente não técnico, pois a interface é bem amigável, com botões simples, funções de arrastar campos e bases para inserir dados, facilidade de alteração das especificações dos gráficos e tabelas. É uma ferramenta bastante gerencial que pode ser usada por pessoas que não conhecem bem a parte técnica, porém conhecem bem o negócio. Das ferramentas analisadas o *Microstrategy* se mostrou a solução com maior facilidade para operações de *drill*, sendo que esse recurso pode ser utilizado de maneira intuitiva com opção de seleção de dimensão ser detalhada eliminando a necessidade de uma configuração prévia durante a seleção dos dados. Isso faz com que a ferramenta seja a mais indicada para a geração de visualizações em níveis tático, estratégico e operacional a partir de uma mesma seleção de dados.

O *QlikView* 11 possui uma interface mais simples, com uma variedade de opções de edição e filtros. Porém, é uma ferramenta mais técnica, sendo necessárias muitas configurações para a primeira utilização e após gerar o gráfico ou tabela precisa que seja editado usando um wizard onde é preciso fazer muitas alterações complexas para que seja útil para a análise. O *QlikView* se mostrou a ferramenta indicada para organizações em que usuários da área de Tecnologia da Informação são responsáveis pela configuração e criação de visualização de dados, conforme a especificação de usuários de negócio, devido ao conhecimento técnico necessário para a correta utilização da ferramenta. Usuários mais experientes, porém, tem uma grande flexibilidade de seleção de dados e visualizações

disponíveis por meio do *QlikView*, o que faz com que a ferramenta seja adequada para a extração e visualização de dados em todos os níveis organizacionais.

Já o *Tableau* é fácil de ser usado, pois todas as ações podem ser feitas clicando e arrastando, não exige muito conhecimento técnico para criação e formatação dos relatórios. Ele limita um pouco sua usabilidade por apenas disponibilizar formatos pré-definidos de acordo com os dados escolhidos, o que atrapalha usuários mais experientes, pois a ferramenta não se adapta às escolhas do usuário. Dentre as ferramentas analisadas o *Tableau* é a única que disponibiliza a opção de gráficos com dados atualizados em tempo real. Isso pode ser um diferencial para organizações com necessidade de análise de dados em nível operacional, visto que dados nesse nível organizacional estão em constante atualização.

O *Power BI* é a ferramenta mais recente do mercado dentre as 4 avaliadas e possui uma interface moderna e intuitiva, seguindo o padrão dos *softwares* mais recentes da *Microsoft*. Sua facilidade em criação de gráficos a partir da seleção dos dados disponíveis o torna uma opção viável para usuários de negócios criarem relatórios do nível tático ou estratégico sem a necessidade de grande conhecimento técnico ou treinamentos para utilização da ferramenta. Usuários experientes também tem a opção de recursos avançados de transformação de dados e exportação para formato em *Excel* para análise de dados em nível operacional. O grande destaque do *Power BI* entre as ferramentas avaliadas é o seu custo e facilidade de implementação. Conforme especificado nas avaliações de investimento, mencionadas no Capítulo 4, o custo da solução da *Microsoft* pode ser até dez vezes menor que outras fornecedoras dos *softwares* avaliados. O *Power BI* oferece uma só versão do seu *software* licenciado e pode ser configurado facilmente. Isto torna a ferramenta ideal para organizações que não dispõem de muito tempo e recursos para iniciar o uso de *Business Intelligence*. Porém, pode ser um desafio para empresas de grande porte que pretendem implantar uma solução com um grande número de usuários.

Com respeito aos requisitos de hardware, todas as quatro ferramentas testadas demandaram uma grande quantidade de memória RAM para a execução dos *scripts* SQL para extração de dados, pois antes de serem mostrados os dados, eles são carregados na memória.

## 6. CONCLUSÃO

Nosso trabalho tem o objetivo de avaliar quatro ferramentas de BI, de acordo com a posição no quadrante de Gartner e a colocação da *inquesti*, com base nos sete critérios, e seus subcritérios, estipulados.

Após a análise de todos os aspectos das quatro ferramentas é possível definir se elas são adequadas para todos os níveis organizacionais e quais as principais diferenças encontradas entre elas. Conclui-se que as quatro ferramentas são semelhantes, oferecendo solução para visualizações de dados nos níveis estratégico, tático e operacional, porém com diferenças em sua utilização, indicando um público-alvo um pouco diferenciado para cada ferramenta.

Com isso, podemos concluir que a metodologia de *Business Intelligence* pode ser aplicada em todos os níveis organizacionais das empresas, porém deve-se atentar ao foco de cada nível para que o BI seja realmente efetivo e ajude os analistas de negócio a conseguirem suas respostas de uma forma mais rápida, confiável e assertiva.

Para finalizar, é importante colocar que devido a utilização de várias fontes de dados com um volume muito alto de registros (cerca de 9 milhões na tabela aluno), a máquina utilizada para realização dos testes teve problemas de desempenho. Pretendíamos testar as diversas modelagens de dados descritas no referencial teórico (seção 2.6), porém vimos que as próprias ferramentas já realizam a modelagem dos dados de forma interna. Então, não foi possível realizar estes testes que havíamos previsto.

Sugerimos que para trabalhos futuros sejam aplicados pesos (grau de importância) para cada critério avaliado, tornando a comparação ainda mais profunda. Além disso, avaliar um número maior de ferramentas, para se obter uma visão mais abrangente da qualidade das ferramentas que hoje estão disponíveis no mercado.

## REFERÊNCIAS

ABELLÓ, Alberto. **YAM: A Mutidimensional Conceptual Model**. Barcelona: Universitat Politècnica de Catalunya, 2002.

ALENQUER, Pablo. **Regras de Negócio para Análise em Ambientes OLAP**. Rio de Janeiro: Universidade Federal do Rio de Janeiro. 2002

ASSIS, Vagner (2015). **Ferramentas de Business Intelligence – Como saber qual é a melhor?** Disponível em <http://www.inquesti.com.br/blog/ferramentas-de-business-intelligence-como-saber-qual-e-a>. Acesso em 22 de Outubro de 2015

BALLARD, Chuck et al. (2006) **Dimensional Modeling: In a Business Intelligence Environment**. Disponível em: <http://www.redbooks.ibm.com/redbooks/pdfs/sg247138.pdf>. Acesso em: 26 de Outubro de 2014.

BALTZAN P.; PHILLIPS A. **Sistemas de Informação**. AMGH Editora. Porto Alegre 2011.

BARBIERI, Carlos. **BI2 Business Intelligence: Modelagem e Qualidade**. Editora Campus. 2011.

CHIAVENATO, I. **Introdução a teoria geral da administração**. Rio de Janeiro: Campus, 2000.

COLLINS, John. **Comparison of Relational and Multi-Dimensional Database Structures**. Disponível em: <http://www.alphadevx.com/a/36-Comparison-of-Relational-and-Multi-Dimensional-Database-Structures>. Acesso em: 17/06/2015

D. J. Power (2007). **A Brief History of Decision Support Systems**, version 4.0". Disponível em: <http://dssresources.com/history/dsshistory.html>. Acesso em: 26 de Outubro de 2014.

DUAN, L.; XU, L. D. **Business Intelligence for Enterprise Systems: A Survey.** IEEE Transactions on Industrial Informatics. 2012.

DWI (2013). **Business Intelligence.** Disponível em: <https://tdwi.org/portals/business-intelligence.aspx>. Acesso em: 24 de Outubro de 2014.

ECKERSON , Wayne. **Best Practices in Business Performance Management: Business and Technical Strategies.** Seattle: The Data Warehousing Institute, 2004.

ECKERSON, Wayne. **Evaluating ETL and Data Integration Platforms.** Seattle: The Data Warehousing Institute. 2003

ELENA, C. (2011). **Business intelligence.** Journal of Knowledge Management, Economics and Information Technology, 1(2). Disponível em: [http://www.scientificpapers.org/wp-content/files/1102\\_Business\\_intelligence.pdf](http://www.scientificpapers.org/wp-content/files/1102_Business_intelligence.pdf). Acesso em: 24 de Outubro de 2014.

FORSMAN, Sarah. **OLAP Council White Paper.** <http://www.olapcouncil.org/research/whtpapy.htm> acesso em 29 de Outubro de 2014.

GARCIA-MOLINA, Hector; ULLMAN, J. D.; WIDOM, Jennifer. **Implementação de Sistemas de Bancos de Dados.** Rio de Janeiro: Campus, 2001.

GARTNER, I. (2014). **Business Intelligence (BI),** Gartner IT Glossary. Disponível em: <http://www.gartner.com/it-glossary/business-intelligence-bi/>. Acesso em: 26 de Outubro de 2014.

GARTNER (2015) **Magic Quadrant for Business Intelligence and Analytics Platforms.** Disponível em: <http://www.gartner.com/technology/reprints.do?id=1-2ACLP1P&ct=150220&st=sb> acesso em 20 de Novembro de 2015.

IMHOFF, Claudia; GAMEEMO, Nicholas; GEIGER, J. G. **Mastering Data Warehouse Design: Relational and Dimensional Techniques**. Indianapolis: Wiley Publishing, 2003.

INMON, Willian. **Building the Data Warehouse**. John Wiley and Sons Inc. New York 1992.

LUHN, H. P. **A Business Intelligence System**. IBM Journal of Research and Development, Volume 2 Issue 4. Riverton, NJ, USA 1958.

MAXIMIANO, Antonio Cezar Amaru. **Teoria Geral da Administração: da escola científica à competitividade na economia globalizada**. 2 ed. São Paulo: Atlas, 2000.

OLAP COUNCIL.(1999). **The OLAP and OLAP Server Definitions**. Disponível em: <http://www.olapcouncil.org>. Acesso em: 26 de Outubro de 2014.

OLIVEIRA, Djalma Pinho Rebouças. **Estratégia Empresarial e Vantagem Competitiva : como estabelecer implementar e avaliar**. 3 ed. São Paulo: Atlas, 2001.

PANDRE, Andrei (2012). **Qlikview Prices**. Disponível em <https://apandre.wordpress.com/2012/02/21/qlikview-prices/>. Acesso em 21 de Novembro de 2015.

PAREDES, John. **The Multidimensional Data Modeling Toolkit: Making Your Business Intelligence Applications Smart with Oracle OLAP**. OLAP World Press. New York 2009.

PCWORLD (2007) **Business Intelligence**. Disponível em: [http://t.idg.com.br/oracle/pcw\\_extra\\_oracle.pdf](http://t.idg.com.br/oracle/pcw_extra_oracle.pdf) Acesso em: 21 de Novembro de 2015.

PCWORLD (2014). **Microstrategy simplifies licensing, pricing in bid for growth**. Disponível em <http://www.pcworld.com/article/2451840/microstrategy-simplifies-licensing-pricing-in-bid-for-growth.html>. Acesso em: 23 de Novembro de 2015.

PEDERSEN, T. B. (2009). **Multidimensional Modeling**, Encyclopedia of Database Systems. Springer. Berlin/Heidelberg 2009.

PONNIAH, Paulraj. **Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals**. John Wiley and Sons Inc. New York 2001

POWER, Daniel J. **Decision Support Systems: Frequently Asked Questions**. iUniverse. Lincoln 2005.

QLIK (2015). **About Qlik**. Disponível em: [www.qlik.com](http://www.qlik.com). Acesso em: 13 de Novembro de 2015.

RAINER, R. K.; CEGIELSKI, C. G. **Introdução a sistemas de informação** (3rd ed.). Rio de Janeiro: Elsevier. 2011.

SÁ, Thays et al. **Uma Análise Comparativa entre as Ferramentas OLAP como Apoio a Soluções de BI nas Empresas**. 2011

SANTOS, M. R. **Aplicação de Business Intelligence para Análise de Indicadores das Redes de Referência no Estado do Paraná**. Universidade Estadual de Londrina. Londrina 2009.

SINGH, H. S. **Data Warehouse: Conceitos, Tecnologias, Implementação e Gerenciamento**. São Paulo: Makron Books, 2001.

TEOREY, Toby; LIGHTSTONE, Sam; NADEAU, Tom. **Database Modeling and Design: Logical Design**, 5<sup>o</sup> edition. Morgan Kaufmann, Burlington MA, 2011.

TURBAN, E.; SHARDA, R.; DELEN, D. **Decision Support and Business**

**Intelligence**: 9th edition. Pearson Education, Inc. New Jersey 2009.

TURBAN, E., & VOLONIMO, L.. **Business Intelligence e Suporte à Decisão: Tecnologia da Informação para Gestão: em busca do melhor desempenho estratégico e operacional**, 8ª edição. Bookman. Porto Alegre 2013.

THOMSEN, Erik. **OLAP Solutions, Building Multidimensional Information Systems**, Second Edition. John Wiley and Sons Inc. New York 2002

KIMBALL, Ralph; ROSS Margy. **The Data Warehouse Toolkit**, The Definitive Guide to Dimensional Modeling, Third Edition. John Wiley & Sons Inc. Indianapolis 2013.

VERCELLIS, Carlo. **Business Intelligence: Data Mining and Optimization for Decision Making**. John Wiley & Sons Ltd. Chichester 2009.



