

EMPREGO DE BUSINESS INTELLIGENCE ESPACIAL OPEN SOURCE NA
AGRICULTURA DE PRECISÃO

GIOVANNI COLONESE

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE

DARCY RIBEIRO

CAMPOS DOS GOYTACAZES – RJ
MAIO – 2014

EMPREGO DE BUSINESS INTELLIGENCE ESPACIAL OPEN
SOURCE NA AGRICULTURA DE PRECISÃO

GIOVANNI COLONESE

Tese apresentada ao Centro de Ciências e
Tecnologias Agropecuárias da Universidade
Estadual do Norte Fluminense Darcy Ribeiro,
como parte das exigências para obtenção do
título de Doutor em Produção Vegetal

Orientador: Ricardo Ferreira Garcia

CAMPOS DOS GOYTACAZES – RJ
MAIO – 2014

FICHA CATALOGRÁFICA

Preparada pela Biblioteca do **CCTA / UENF** 069/2014

Colonese, Giovanni

Emprego de Business Intelligence Espacial Open Source na agricultura de precisão / Giovanni Colonese. – 2014.

106 f. : il.

Orientador: Ricardo Ferreira Garcia

Tese (Doutorado - Produção Vegetal) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Ciências e Tecnologias Agropecuárias. Campos dos Goytacazes, RJ, 2014.

Bibliografia: f. 96 – 102.

1. Business Intelligence Espacial 2. BI 3. Agricultura de precisão 4. Cana-de-açúcar 5. SpagoBI I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. Centro de Ciências e Tecnologias Agropecuárias. II. Título.

CDD – 630

EMPREGO DE BUSINESS INTELLIGENCE ESPACIAL OPEN
SOURCE NA AGRICULTURA DE PRECISÃO

GIOVANNI COLONESE

Tese apresentada ao Centro de Ciências e
Tecnologias Agropecuárias da Universidade
Estadual do Norte Fluminense Darcy Ribeiro,
como parte das exigências para obtenção do
título de Doutor em Produção Vegetal

Aprovada em 28 de maio de 2014

Comissão Examinadora

Prof. Rogério Atem de Carvalho (D.Sc., Engenharia de Produção) – IFF

Prof. José Tarcísio Thiebaut (D.Sc., Produção Animal) – UENF

Dr. Carlos Frederico de Menezes Veiga (D.Sc., Produção Vegetal) – UFRRJ

Prof. Ricardo Ferreira Garcia (D.Sc., Engenharia Agrícola) – UENF
(Orientador)

AGRADECIMENTO

Em primeiro lugar a Deus, sem o qual nada seria possível, por sempre me conceder muito mais do que mereço.

À minha esposa Joziane e aos meus filhos Anna Luisa, Gabriel e Tiago pela presença sempre revigorante e por terem sido a motivação para meu esforço. Um obrigado especial ao Gabriel, por ter feito a reaquisição espacial dos lotes da fazenda tema deste trabalho – foi sua primeira experiência de iniciação científica.

Aos meus pais, Liberato e Emília, pelo desmedido amor e pela dedicação, e meus irmãos, Daniel e Sandro, pela certeza da fraternidade que sempre nos uniu.

Ao Professor Ricardo Ferreira Garcia, meu orientador, pela disponibilidade, amizade e paciência com que sempre me tratou e conduziu na execução deste trabalho.

A Fazenda Abadia, propriedade rural que foi o estudo de caso desta tese, o muito obrigado pela confiança em me ceder dados sensíveis e por autorizar este trabalho. Agradeço especialmente ao grande amigo e companheiro Eng. Willy Pedro Prellwitz, MSc., administrador da fazenda Abadia, que não mediu esforços em me auxiliar, fornecendo dados e explicando procedimentos para a

elaboração desta tese, deixando inúmeras vezes seu tempo de descanso para me atender.

SUMÁRIO

LISTA DE FIGURAS	vi
RESUMO	ix
ABSTRACT	x
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	4
2.1. Agricultura de Precisão.....	4
2.2. Sistemas de Suporte à Decisão (SSD).....	5
2.3. Business Intelligence.....	7
2.3.1. Sistemas OLTP x OLAP.....	8
2.3.2. Data Warehouse (DW).....	10
2.3.3. Arquitetura de um ambiente de Data Warehouse.....	11
2.3.4. Granularidade.....	12
2.3.5. Modelagem de um Data Warehouse.....	13
2.3.6. Tabela Fato.....	16
2.3.7. Tabelas Dimensão.....	16
2.4. Suítes de Business Intelligence	17
2.5. Trabalhos Relacionados	19
3. MATERIAL E MÉTODOS	21
3.1. A fazenda em estudo	21
3.2. Visão geral da solução – implantação de sistema de BI espacial	26
3.3. Obtenção dos dados existentes na fazenda	28
3.3.1. Dados não-espaciais.....	28
3.3.2. Dados Espaciais.....	31
3.4. Modelo OLAP para o sistema de BI.....	32
3.4.1. Cultura.....	34
3.4.2. Corte.....	34
3.4.3. Lote.....	34
3.4.4. Tempo.....	35
3.4.5. Tipo de Colheita.....	35
3.4.6. Irrigação.....	35
3.4.7. Controle de plantas daninhas.....	35
3.4.8. Adubação de plantio (Classe de fundação).....	35
3.4.9. Adubação de Cobertura (Classe de cobertura).....	35
3.4.10. Pluviometria.....	36
3.5. Modelo físico do Data Warehouse e carga de dados.....	36
3.5.1. Carga dos dados não espaciais.....	37
3.5.2. Carga dos dados espaciais.....	39

3.6. Instalação e configuração da suite de BI - SpagoBI.....	40
3.7. Criação de documentos.....	42
3.7.1. Criação do cubo Olap.....	44
3.7.2. Criação de painéis de controle (“dashboards”).....	50
3.7.3. Criação de documentos espaciais – Location Intelligence.....	55
4. RESULTADOS E DISCUSSÃO	62
4.1. Cenário 1: exibir a evolução das métricas: produção e produtividade ao longo dos anos para a fazenda.....	62
4.2. Cenário 2: Distribuição das variedades pelas áreas da propriedade, com visualização em um mapa, de forma a visualizar o mix de variedades e suas concentrações.....	67
4.3. Cenário 3: Análise das métricas de produção nas áreas da fazenda ao longo do tempo.....	71
4.4. Cenário 4: Estudo das métricas da produção nas áreas da fazenda, em determinado ano com visualização espacial.....	75
4.5. Cenário 5: Fazer a análise da rentabilidade pelos cortes.....	78
4.6. Cenário 6: Análise financeira das variedades, de forma a selecionar aquelas que trazem o maior retorno econômico.....	80
4.7. Cenário 7: Análise da produtividade média das variedades, considerando número do corte e tipo de solo.....	83
4.8. Cenário 8: Caracterizar a produção em um determinado ano.....	86
4.9. Considerações.....	91
5. RESUMO E CONCLUSÕES	93
REFERÊNCIAS BIBLIOGRÁFICAS	96
APENDICE A – Configuração metadados – SpagoBI	103
APÊNDICE B – Configuração do modelo JSON para o documento GeoReport (“GisEngine”)	105

LISTA DE FIGURAS

Figura 1 - Diagrama representativo dos conceitos e das tecnologias englobados por BI (Turban et al.,2010).....	7
Figura 2 - Arquitetura e Processo – Data Warehousing (Fonte: Canaltech Corporate, 2014).....	11
Figura 3 - Visualização das dimensões de um cubo (Fonte: MSDN, 2014).....	14
Figura 4 - Modelo Estrela (Fonte: MSDN, 2014).....	15
Figura 5 - Planilha de controle da produção.....	24
Figura 6 - Ficha para controle pluviométrico.....	25
Figura 7 - Visão geral da implementação - sistema de BI.....	27
Figura 8 - PlanilhaBase – visão parcial.....	30
Figura 9 - PlanilhaLotes contendo o tipo de solo para os lotes da fazenda.....	31
Figura 10 - Modelo Olap para sistema de BI – Fazenda Abadia.....	33
Figura 11 - Diagrama de Estrutura de Dados (DED) – DW Fazenda Abadia.....	36
Figura 12 - Talend Open Studio – Tela de Definição do Job.....	38
Figura 13 - Mapeamento Origem x Destino – Talend Open Studio.....	38
Figura 14 - Tela inicial de demonstração da suíte SpagoBI.....	41
Figura 15 - SpagoBI – Tela para criação de novo documento.....	42
Figura 16 - Exemplo arquivo xml de um documento – “Gráfico Velocímetro”.....	43
Figura 17 - SpagoBI Studio.....	43
Figura 18 - Definição do modelo Olap no SpagoBI Studio.....	45
Figura 19 - Criação novo documento do tipo Olap.....	46
Figura 20 - Definição da visão inicial para o cubo Olap.....	46
Figura 21 - Visualização inicial do documento Olap.....	47
Figura 22 - Operação de “drill-down”, detalhando as variedades de certo corte...	48
Figura 23 - Visualização do desempenho dos cortes por variedade.....	48
Figura 24 - Botão “Olap Navigator” que abre a tela de reconfiguração da aplicação.....	49
Figura 25 - Tela do “Olap Navigator”, que permite a seleção das dimensões, métricas e sua disposição sobre os eixos de visualização, além da definição de filtros sobre as dimensões.....	50
Figura 26 - Painel de Controle (dashboard) com três gráficos.....	51
Figura 27 - Criação do Gráfico 1 para o Painel de Controle no SpagoBI.....	52
Figura 28 - Modelo xml para o Gráfico 1.....	52
Figura 29 - Parte inferior da tela de criação do Gráfico 2, exibindo os “Analytical Drivers” - parâmetros para a execução do documento no SpagoBI.....	53
Figura 30 - Criação do documento para o Painel de Controle no SpagoBI.....	54
Figura 31 - Modelo xml para o documento do Painel de Controle.....	55
Figura 32 - Documento de análise espacial das métricas da produção.....	56
Figura 33 - Detalhe da exibição de documento “pop-up” quando usuário clica sobre um dos lotes.....	57

Figura 34 - Detalhe do documento Olap - navegação cruzada a partir do “pop-up” do mapa.....	57
Figura 35 - Camadas espaciais (“Layers”): setor e lote, configuradas no GeoServer.....	58
Figura 36 - Visualização da camada “setor’ da fazenda Abadia no GeoServer.....	59
Figura 37 - Visualização da camada “lote’ da fazenda Abadia no GeoServer.....	59
Figura 38 - Criação do documento tipo “Location Intelligence” no SpagoBI.....	60
Figura 39 - Modelo JSON de definição para documento – GisEngine.....	61
Figura 40 - Evolução temporal da produção e produtividade.....	63
Figura 41 - Adicionando o atributo “AreaColhida” à exibição do documento Olap.	64
Figura 42 - Documento Olap reconfigurado, exibindo “AreaColhida” além de produção e produtividade.....	65
Figura 43 - Documento Olap adicionado da dimensão Pluviometria.....	66
Figura 44 - Documento Olap com área colhida por corte ao longo dos anos.....	67
Figura 45 - Tela inicial do documento de ocupação das áreas por variedade para o ano de 2004.....	68
Figura 46 - Surgimento do terceiro documento no painel de controle – mapa – exibindo as áreas ocupadas pela variedade RB 72-454, no ano de 2004.....	69
Figura 47 - Detalhe do mapa – exibição das legendas utilizadas para a classificação das áreas e possibilidade de escolha de outra métrica para análise.	70
Figura 48 - Documento com a composição varietal em 2002 e, no mapa, a exibição das áreas ocupadas com a variedade RB 72-454.....	71
Figura 49 - Documento Olap com a rentabilidade dos setores ao longo dos anos.	72
Figura 50 - Métricas disponíveis para escolha do usuário – Documento Olap.....	72
Figura 51 - Dados de área colhida, rentabilidade e produtividade agregados por setor da fazenda em todos os anos.....	73
Figura 52 - Documento Olap mostrando a concentração de áreas colhidas por cortes.....	74
Figura 53 - Documento com os níveis de produtividade para cada lote no ano de 2002.....	75
Figura 54 - Detalhe do painel que se abre quando o usuário clica sobre um dos lotes.....	76
Figura 55 - Parte inferior do painel – detalhe do gráfico com a rentabilidade do lote.....	76
Figura 56 - Documento Olap com métricas da produção para todo o setor no qual o lote está localizado.....	77
Figura 57 - Documento Olap com os nós dos lotes expandidos, permitindo comparação entre os lotes do mesmo setor, além do corte para o lote em questão.	77
Figura 58 - Documento Olap – visão inicial da rentabilidade por corte.....	78
Figura 59 - Visualização da rentabilidade por cortes detalhada pelos setores.....	79
Figura 60 - Documento com a rentabilidade por corte ao longo dos anos.....	80
Figura 61 - Documento Olap ordenado pelo lucro acumulado das variedades.....	81
Figura 62 - Análise de produção e lucro das variedades ao longo dos anos.....	81

Figura 63 - Documento Olap filtrando apenas variedade SP 792233, adicionado da dimensão “corte”.....	82
Figura 64 - Visão do documento tipo painel de controle contendo três gráficos....	84
Figura 65 - Visão do comparativo das produtividades médias das variedades....	84
Figura 66 - Comportamento da produtividade média por corte e por tipo de solo para a variedade SP 71-1406.....	85
Figura 67 - Comportamento da produtividade média por corte e por tipo de solo para a variedade RB76-5418.....	86
Figura 68 - Painel de controle situacional definido como página inicial da suíte SpagoBI.....	87
Figura 69 - Documento Olap com métricas da produção – detalhamento do link “Produção (t)”.....	88
Figura 70 - Documento Olap com detalhamento da produtividade pelos anos e pelos setores ou lotes.....	89
Figura 71 - Documento com detalhamento da ocupação de área por variedade..	89
Figura 72 - Documento de comparativo de rentabilidade das variedades, rentabilidade e produtividade de uma variedade por corte e rentabilidade por tipo de solo.....	90
Figura 73 - Produção por Corte - ano de 2007.....	91
Figura 74 - Produção por Corte – ano de 2006.....	91

RESUMO

COLONESE, Giovanni; Universidade Estadual do Norte Fluminense Darcy Ribeiro; Maio de 2014; EMPREGO DE BUSINESS INTELLIGENCE ESPACIAL OPEN SOURCE NA AGRICULTURA DE PRECISÃO; Professor Orientador: Ricardo Ferreira Garcia.

A agricultura de precisão traz como principal diretriz o gerenciamento da propriedade em parcelas menores, de forma a otimizar a produção em cada pequena porção da fazenda com vistas à obtenção da máxima rentabilidade global da atividade. O registro de maior quantidade de dados, provenientes das múltiplas divisões da propriedade e das várias propriedades armazenadas acerca de cada uma delas, gera uma nova dificuldade: a incapacidade humana de lidar com tamanha quantidade de dados para a geração de informações capazes de apoiar tomadas de decisões. Este trabalho propõe o uso de uma suíte de Business Intelligence espacial de código aberto como destinação para os dados gerados pela agricultura de precisão. De forma a demonstrar os passos necessários à sua implementação e os benefícios de seu emprego no suporte à decisões estratégicas, esta tese utiliza a suíte SpagoBI para a criação de um sistema de Business Intelligence espacial em apoio a uma fazenda produtora de cana-de-açúcar localizada em Campos dos Goytacazes - RJ.

Palavras-Chave: Business Intelligence Espacial, agricultura de precisão, cana-de-açúcar, BI, SpagoBI

ABSTRACT

COLONESE, Giovanni; Universidade Estadual do Norte Fluminense Darcy Ribeiro; Maio de 2014; THE USE OF OPEN SOURCE SPATIAL BUSINESS INTELLIGENCE IN PRECISION AGRICULTURE; Adviser: Ricardo Ferreira Garcia.

Precision agriculture has as main guideline the management of the property into smaller parcels in order to optimize production in each small portion of the farm obtaining the maximum overall profitability. The higher amount of data coming from multiple divisions of property and various properties stored about each one generates a new difficulty: the human inability to handle so much data to generate information that can support decision making. This work proposes the use of an Open Source Spatial Business Intelligence suite to store and process data generated by precision agriculture . In order to demonstrate the steps needed for its implementation and the benefits of its use to support strategic decision making, this thesis uses the SpagoBI suite for the creation of a Spatial Business Intelligence system to support a farm producing cane sugar located in Campos dos Goytacazes - RJ.

Keywords: Spatial Business Intelligence, precision agriculture, sugar cane, BI, SpagoBI

1. INTRODUÇÃO

O incremento da competitividade é uma máxima em todos os setores da economia globalizada atual. Na agricultura o cenário não é diferente, e com as restrições ambientais que reduzem as áreas de plantio, com a diminuição da população no campo e com o aumento da população global, cresce a pressão por maior produtividade.

Segundo Roza (2014), a agricultura de precisão é uma filosofia de gerenciamento agrícola que parte de informações exatas, precisas e se completa com decisões exatas. Agricultura de precisão, também chamada de AP, é uma maneira de gerir um campo produtivo metro a metro, levando em conta o fato de que cada pedaço da fazenda tem propriedades diferentes.

O manejo da propriedade em parcelas menores e o aumento na quantidade de parâmetros controlados pela agricultura de precisão leva a um enorme crescimento na quantidade de dados gerados. Tais dados atualmente destinam-se à tomada de decisão operacional, ou seja, os dados gerados são empregados pelo gerente da fazenda ou consultor para decidir quanto de insumo será aplicado em cada área mapeada, ou consultar quanto se produziu, ao término do período da colheita, em cada talhão.

A presente tese vem propor um outro emprego para a enorme massa de dados gerada pela agricultura de precisão: o suporte à decisão em nível estratégico através de uma ferramenta de Business Intelligence (BI), cujo objetivo é confirmar tendências, mostrar projeções e padrões de comportamento que possam levar o decisor de nível mais

elevado (ex.: administrador da propriedade, diretor de cooperativa, gestor de organização governamental) a maximizar seu lucro com a atividade. Assim, o objetivo deste trabalho é utilizar um sistema de Business Intelligence (BI) open-source para gerar informações a partir da massa de dados obtida pela utilização da agricultura de precisão em uma fazenda de cana-de-açúcar na região Norte Fluminense.

Sistemas de Business Intelligence (BI), constituem um tipo de sistema de suporte à decisão que tratam e agregam dados obtidos do ambiente operacional, disponibilizando-os aos decisores em nível estratégico ou tático em um formato que facilite a visualização das métricas representativas de um negócio e de suas perspectivas de análise.

Sob o termo “guarda-chuva” Business Intelligence repousam diversas ferramentas e tecnologias: Data Warehousing, OLAP (On-Line Analytical Processing), Location Intelligence (Inteligência de Localização), Data Mining (Mineração de dados), Reporting (Relatórios) e Dashboards (Painéis de Controle) com a finalidade de proverem informações capazes de gerar intuições nos decisores, levando-os a tomar decisões mais embasadas.

Como é de se esperar de um sistema complexo como Business Intelligence, seu custo é bastante elevado. De forma a estimular sua adoção na agricultura de precisão, mesmo no âmbito de organizações rurais com restrições orçamentárias, esta tese utiliza uma suíte de BI de código aberto (open source), eliminando assim o custo de licenciamento do software.

Sistemas de código aberto (do inglês open source), ou software livre é o termo atribuído a sistemas cujo código-fonte se encontra disponível e pelos quais não são cobradas taxas de licenciamento ou uso. Tais sistemas permitem o acesso de muitas organizações a sistemas cujas versões proprietárias possuem custos de aquisição impeditivos.

De forma a demonstrar o uso de Business Intelligence em apoio à agricultura de precisão, usando ferramentas de código aberto (open source), o presente trabalho segue assim organizado: o capítulo 2 faz uma revisão da literatura sobre os principais conceitos envolvidos, o capítulo 3 apresenta, em “Material e Métodos”, as práticas conduzidas pela fazenda Abadia, os dados disponíveis, o modelo dimensional a ser trabalhado e o emprego da ferramenta SpagoBI (software de BI selecionado para este trabalho), o capítulo 4, em “Resultados e Discussão”, demonstra o emprego da suíte de BI em apoio a diversos cenários de análise, exibindo os dados fornecidos pelos documentos e as principais informações extraídas da utilização da ferramenta para a tomada de decisões

estratégicas na fazenda estudada. O capítulo 5, finalmente, apresenta as conclusões e sugestões de trabalhos futuros.

2. REVISÃO DE LITERATURA

2.1. Agricultura de Precisão

Segundo Batchelor et al. (1997), a agricultura de precisão é uma filosofia de manejo da fazenda na qual os produtores são capazes de identificar a variabilidade dentro de um campo, e então manejar aquela variabilidade para aumentar a produtividade e os lucros.

Segundo a EMBRAPA, Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA, 1997), o termo agricultura de precisão engloba o uso de tecnologias atuais para o manejo de solo, insumos e culturas, de modo adequado às variações espaciais e temporais em fatores que afetam a produtividade das mesmas.

Para Davis (1998), a agricultura de precisão foi frequentemente definida com base nas tecnologias que permitem que ela seja realizada, como GPS (Sistema de Posicionamento Global) ou sistemas de taxa variável. Tão importantes quanto os dispositivos usados na agricultura de precisão, é a percepção de que a informação usada ou coletada é o ingrediente chave para o sucesso do sistema. O conceito de agricultura de precisão se distingue da agricultura tradicional por seu nível de manejo. Em vez de administrar uma área inteira como uma única unidade, o manejo é adaptado para pequenas áreas dentro de um campo.

A agricultura de precisão é um termo aplicado a uma larga ordem de tópicos que se relacionam ao manejo preciso de unidades pequenas de terra em contraste com o

manejo tradicional onde um campo inteiro é compreendido como uniforme. Esta oportunidade de se administrar pequenas áreas do terreno individualmente ficou possível devido à disponibilidade de sistemas de posicionamento globais que podem ser usados para localização exata em terra de equipamentos e máquinas (Miller & Supalla, 1996).

Capelli (1999) afirma que a agricultura de precisão apresenta as vantagens de possibilitar um melhor conhecimento do campo de produção, permitindo desta forma a tomada de decisões melhor embasadas. Com isto tem-se uma maior capacidade e flexibilidade para a distribuição dos insumos naqueles locais e no tempo em que são mais necessários, minimizando os custos de produção. O autor diz ainda que a uniformidade na produtividade é alcançada pela correção dos fatores que contribuem para sua variabilidade obtendo-se, com isto, um aumento global da produtividade. A aplicação localizada dos insumos necessários para sustentar uma alta produtividade contribui com a preservação do meio ambiente, já que estes insumos são aplicados somente nos locais, quantidades e no tempo necessário.

Para Batchelor et al. (1997), a agricultura de precisão pode:

- melhorar os rendimentos de colheita e lucros;
- fornecer informações para tomar decisões de manejo mais embasadas;
- prover registros de fazenda mais detalhados e úteis;
- reduzir custos de fertilizante;
- reduzir custos de defensivos agrícolas; e
- reduzir poluição.

Em função de todas as definições apresentadas para agricultura de precisão, percebe-se que o seu emprego irá aumentar bastante a quantidade de dados coletados acerca de cada porção do terreno a ser trabalhado. Dados acerca de solo, insumos, fertilizantes, culturas, variedades e tempo fazem parte do dia a dia de quem utiliza a agricultura de precisão. Tal quantidade enorme de dados é utilizada atualmente para a geração de mapas de solos, de produtividade e outros por meio do apoio de programas específicos ou de Sistemas de Informações Geográficas (SIG) genéricos.

2.2. Sistemas de Suporte à Decisão (SSD)

Segundo Daniel Power (2007), sistemas de suporte à decisão são uma classe de sistemas de informação que têm por objetivo auxiliar, de modo interativo, o processo de

tomada de decisão, utilizando dados, conhecimento e modelos computacionais para identificar e solucionar problemas.

Elmasri e Navathe (2002) definem sistemas de suporte à decisão como sistemas que levam dados aos usuários de altos níveis administrativos de uma instituição para o apoio a decisões complexas.

Come (2001) definiu apoio à decisão, também chamado suporte à decisão, como “o processo de agrupar, estruturar, manipular, armazenar, acessar, apresentar e distribuir informações de negócios de uma maneira oportuna, ou seja, a informação certa no momento certo e na quantidade certa”.

Estes sistemas possuem conhecimentos específicos sobre determinado assunto e, mediante regras de avaliação introduzidas pelo usuário, retornam soluções otimizadas para os problemas apresentados, auxiliando o processo decisório (Power, 2003).

Ainda segundo Power (2003), a necessidade de suporte para a tomada de decisão decorre de limites naturais para os decisores humanos, tais como limites cognitivos, econômicos e de tempo.

Os limites cognitivos dizem respeito à capacidade de processamento humana que segue a “regra” de trabalhar bem com uma quantidade média entre 5 a 9 (7 mais ou menos 2) variáveis na memória imediata. A memória humana de longo termo é pequena para grandes volumes de conhecimento relevante. O excesso de informação pode ser tão debilitante quanto a escassez, levando ao estresse, a erros e ao desprezo involuntário de fatores importantes.

Limites econômicos decorrem da escassez de pessoal qualificado para analisar informações e tomar decisões, o que torna a tentativa de contornar limites cognitivos para a tomada de decisão por meio da contratação de mais pessoal, uma opção cara e economicamente limitante.

Os limites de tempo estão quase sempre presentes nas tomadas de decisão. A limitação do tempo necessário à tomada de uma decisão aliada à limitação cognitiva conduz o decisor a elevado nível de estresse, o que pode forçar tal pessoa a empregar uma estratégia inadequada para a resolução de um problema.

Com respeito à taxonomia dos sistemas de suporte à decisão, não há consenso entre os autores mais destacados. Para citar algumas das principais categorizações, Diuf apud Hättenschwiller (2014) propõe uma classificação, pertinente ao usuário, categorizando os SSD como passivos, ativos ou cooperativos. Um SSD passivo auxilia no processo decisório, mas não apresenta ao usuário decisões ou soluções explícitas. Um

SSD ativo aponta decisões ou caminhos a serem seguidos pelo usuário. Finalmente, um SSD cooperativo fornece meios para que o administrador possa refinar as sugestões de decisões apresentadas pelo sistema e reenviá-las para nova avaliação e validação.

2.3. Business Intelligence

Segundo Turban et al (2009), Business Intelligence (BI) foi um termo inicialmente usado pelo Gartner Group em meados da década de 90, mas o conceito já era bem mais antigo possuindo suas raízes nos Sistemas de Informações Gerenciais surgidos nos anos 70. Mais tarde surgiria o conceito de EIS (Sistema de Informações Executivas, ou EIS, no inglês) que agregou aos antigos Sistemas de Informações Gerenciais, novas tecnologias. Os Sistemas denominados de Business Intelligence surgiriam assim nos anos de 90 englobando as características dos antigos sistemas de EIS e adicionando também outras tecnologias, como inteligência artificial. Assim, tal termo abrangente encontra boa representação na Figura 1.

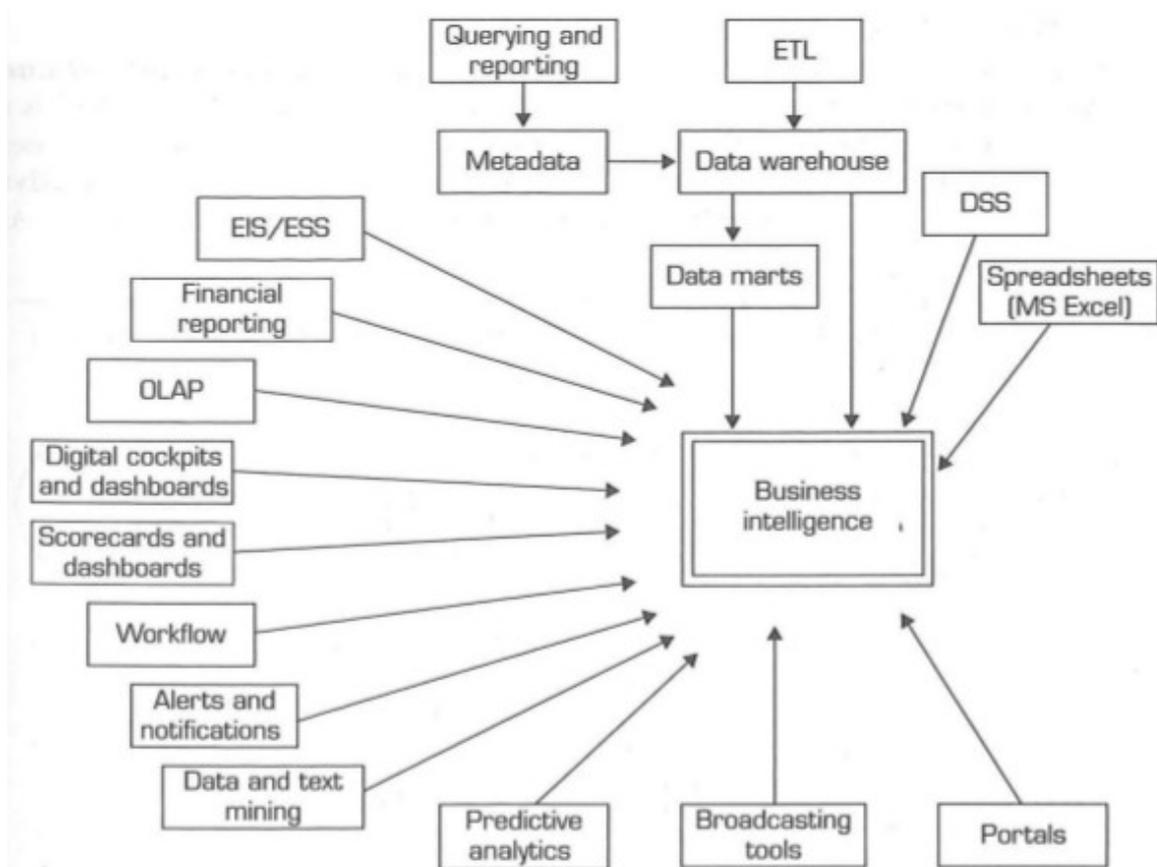


Figura 1 - Diagrama representativo dos conceitos e das tecnologias englobados por BI (Turban et al.,2010).

Loshin (2003) explica que Business Intelligence é um conjunto de processos, tecnologias e ferramentas necessárias a transformar dado em informação, informação em conhecimento e conhecimento em planejamento que conduza a ações de negócio rentáveis. BI engloba data warehousing, ferramentas analíticas de negócio e gerenciamento de conteúdo e conhecimento.

Segundo Howson (2008), BI permite que pessoas em todos os níveis de uma organização possam acessar, interagir e analisar dados de forma a gerenciar o negócio, incrementar performance, descobrir oportunidades e operar eficientemente.

Do anteriormente exposto, percebe-se que Business Intelligence é uma categoria de sistema de suporte à decisão. O termo BI, com sua abrangência, compreende conceitos, tecnologias e ferramentas como data warehousing, OLAP, relatórios, planilhas, painéis de controle (denominados “dashboards”, em inglês), data mining (mineração de dados) e sistemas de informações geográficas. Tais tecnologias e ferramentas vêm normalmente integradas em uma suíte de softwares denominada suíte de BI.

Para que se possa compreender os futuros detalhes de modelagem e implementação abordados nesta tese, os tópicos que se seguem elucidam conceitos acerca da porção mais importante de um sistema de BI: o data warehouse.

2.3.1. Sistemas OLTP x OLAP

Os sistemas de informação podem ser classificados, grosso modo, em duas categorias: sistemas OLTP (On-Line Transaction Processing) ou transacionais e sistemas OLAP (On-Line Analytical Processing) ou analíticos.

De acordo com Erik Thomsen (2002), o termo OLAP é utilizado para designar um conjunto de conceitos, camadas de software, um produto de software ou mesmo um ambiente. Para eliminar tal falta de precisão nas designações, o termo “sistema OLAP” será empregado quando o objetivo for a referência ao ambiente de softwares com a finalidade de suporte à decisão e o termo “aplicativo OLAP” será empregado quando a finalidade for designar o programa que permite o acesso flexível a um repositório de dados históricos.

Os sistemas transacionais, também chamados operacionais, estão relacionados com a rotina operacional e basicamente registram as transações diárias da empresa. São modelados privilegiando a normalização dos dados a fim de garantir a consistência e integridade das informações e contemplam prioritariamente as operações de entrada e exclusão de dados. Em contraposição, os sistemas analíticos ou informacionais subsidiam

os gestores e analistas no processo de tomada de decisão através das informações oriundas do ambiente operacional. São modelados abrindo mão da normalização, de forma a serem capazes de executar consultas complexas dentro de limites de tempo aceitáveis (o que explica o termo “on-line”) e trabalham com base de dados históricas de grandes dimensões.

Sistemas OLAP diferem dos sistemas transacionais quanto aos destinatários: estes destinam-se à maioria dos usuários, cujas funções são operacionais enquanto aqueles visam os diretores das organizações, cujas funções são de supervisão e tomada de decisão em nível estratégico.

O Quadro 1 resume as principais diferenças entre sistemas OLTP (operacionais ou transacionais) e sistemas OLAP (informacionais analíticos ou dimensionais).

Quadro 1 - Sistemas Transacionais (OLTP) x Analíticos (OLAP)

Característica	OLTP	OLAP
Tipo de dados	Detalhados	Sumarizados e detalhados
Estabilidade dos dados	Dinâmico	Estático
Estrutura dos dados	Otimizados para transações	Otimizados para pesquisas complexas
Dados por transação	Poucos (dezenas)	Muitos (milhares)
Frequência de acesso	Alta	Baixa
Volume de dados	Megabytes-Gigabytes	Gigabytes-Terabytes
Tipo de informação	Atual e volátil	Histórica e não-volátil
Operação	Inserção e Atualização	Leitura
Comunidade atendida	Funcional. Decisões do dia-a-dia	Gerencial. Decisões estratégicas e de longo prazo
Redundância	Não ocorre (normalizado)	Ocorre (desnormalizado)
Objetivo	Manutenção do negócio	Análise do negócio
Interação	Predefinida	Predefinida e Ad-Hoc
Histórico	Baixo (alguns meses)	Alto (anos)
Atualização	Imediata (on-line)	Periodicamente (Batch)

Thomsen (2002) também afirma que sistemas OLAP diferem de sistemas transacionais quanto à flexibilidade do uso: enquanto os sistemas operacionais seguem

uma sequência predeterminada na apresentação das informações, seguindo regras do negócio da empresa, os sistemas de suporte à decisão ou informacionais devem ser robustos o suficiente para apresentar as informações sob as mais variadas formas: planilhas, gráficos, mapas, etc.

2.3.2. Data Warehouse (DW)

Data Warehouse é o nome dado ao repositório de dados dos sistemas OLAP.

Segundo Inmon (1997), Data Warehouse é uma coleção de dados orientados por assuntos, integrados, variáveis com o tempo e não voláteis, para dar suporte ao processo gerencial de tomada de decisão.

É orientado por assunto devido ao fato do DW armazenar informações sobre temas específicos importantes para o negócio da empresa. Exemplos: produtos, atividades, contas, clientes, produtividade, etc.

É integrado, pois as informações que vêm de diferentes sistemas transacionais, com diferentes formatos, devem ser convertidas para um formato único, antes de serem carregadas no DW. Por exemplo, as medidas de um terreno podem estar registradas em metros quadrados em sistema e em hectares em outro e, ao serem copiadas para o DW, deverão estar padronizadas em metros quadrados.

É variável no tempo devido à DW manterem um histórico por muito mais tempo que um sistema de dados operacional ou transacional. Enquanto os sistemas OLTP armazenam dados por um curto período, para não reduzir o desempenho do sistema, os data warehouses guardam um histórico por muito tempo, assim possibilitando que análises de tendências ao longo do tempo sejam feitas com suas informações.

Não é volátil, porque as informações cadastradas não são atualizadas e nem excluídas. Um dado que foi copiado para o DW deverá por lá permanecer da maneira como foi criado durante todo o tempo.

Elmasri e Navathe (2011) explicam que Data Warehouse é o nome dado ao repositório de dados histórico, ou seja, a base de dados utilizada nos sistemas OLAP. Data Warehousing é o termo que designa todas as atividades ligadas ao processo de criar, manter e utilizar um DW para suporte à decisão.

2.3.3. Arquitetura de um ambiente de Data Warehouse

O ambiente de DW pode ser composto por diversos componentes, conforme ilustra a Figura 2.

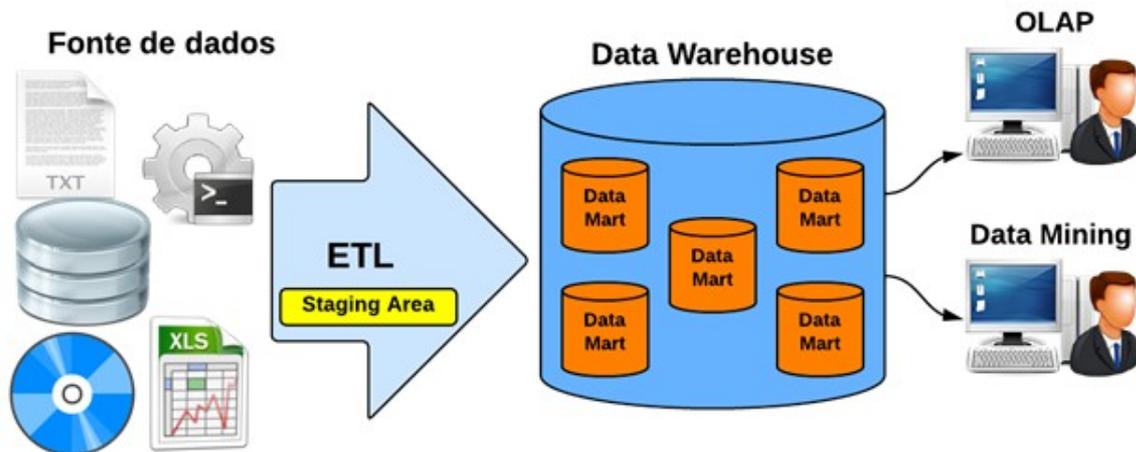


Figura 2 - Arquitetura e Processo – Data Warehousing (Fonte: Canaltech Corporate, 2014).

Segundo Machado (2008), os sistemas OLTP ou sistemas transacionais ou operacionais, são as fontes de dados. Para projetar um Data Warehouse é necessário conhecer estas bases de dados e mapear as informações desejadas. Nem tudo que existe no OLTP será carregado para o DW.

Uma vez identificados os dados eles devem ser extraídos para uma área intermediária chamada Data Staging Area ou Operational Data Storage (ODS), onde serão integrados em formato e tempo. Esta etapa é de suma importância em um ambiente onde existem muitas fontes heterogêneas e um único Data Mart. Como exemplo pode-se citar problemas com a extração de uma tabela que exista em bancos de dados distintos e também na conversão de dados com unidades distintas.

O processo de extração, transformação e carga (ETL, do inglês Extract, Transform and Load) compreende a movimentação dos dados do OLTP para a data staging área e desta para o DW, incluindo seus processamentos. O ETL ocorre periodicamente para atualização dos dados do DW, conforme necessidade do negócio; como exemplo, um DW acadêmico não necessita ser atualizado diariamente, pois as notas são lançadas por bimestre ou semestre. Cabe lembrar que quanto menor a periodicidade, maior será o volume de dados armazenados.

Normalmente a carga é apenas incremental, só acrescentando dados. Ela pode ser feita de diversas formas: carga “full” ou inicial - o DW é esvaziado e todos os registros

das tabelas no OLTP são carregados novamente; delta - é carregada apenas a diferença da carga anterior; carga específica - onde são carregados determinados objetos conforme critérios do usuário, por exemplo, recarregar somente clientes do estado de SP.

2.3.4. Granularidade

A granularidade é a definição do nível de sumarização (detalhe) em que os dados serão armazenados no DW. Ela não faz sentido para os sistemas OLTP, pois nestes, ela é inerente ao processo. Por exemplo, em um sistema OLTP de vendas, cada venda será um registro separado independente se ocorreu no mesmo dia ou se é do mesmo produto. Em um DW que tratasse de vendas, os registros poderiam ser armazenados com a soma das vendas de cada produto em um dia (com isto, não se conheceria o cliente que fez a compra), ou então armazenando por semana as vendas de cada produto por cliente (e neste caso não se saberia a quantidade de vendas em cada dia). Tais decisões minimizam o número de registros, diminuindo o volume de dados e maximizando o desempenho, mas pagando-se um preço em termos de informação perdida (Machado, 2008).

Quanto mais detalhes, menor a granularidade, ou seja, os dados estão menos sumarizados, assim gera-se uma quantidade maior de registros. Quanto menos detalhes, maior a granularidade e menos registros.

A definição da granularidade é fundamental para desenvolvimento do DW e depende das consultas desejadas. Deve-se tentar sumarizar os dados, ou seja, aumentar a granularidade, mas não se pode inviabilizar a realização das consultas necessárias.

A granularidade aplica-se principalmente ao tempo, por exemplo, pode-se optar em guardar os dados por dia ou semana ou mês ou ano, mas não só a ele, aplica-se também aos outros aspectos, como por exemplo, pode-se armazenar as vendas por bairro ou cidade ou estado.

Tais definições devem ser extraídas dos requisitos do negócio e observadas durante a modelagem e a atividade de ETL.

2.3.5. Modelagem de um Data Warehouse

Assim como a definição da granularidade, a modelagem de um DW, denominada modelagem dimensional ou multidimensional, é determinante para o sucesso do mesmo, pois pode impor limitações para a análise do negócio.

A modelagem relacional utilizada nos bancos de dados dos sistemas OLTP, tem como característica a necessidade de ser ter dados normalizados e sem redundância facilitando as operações de inserção, atualização e deleção. Na modelagem dimensional a prioridade é a consulta, assim, os dados são organizados para maximizar esta operação, minimizando a utilização de junções entre tabelas, utilizando desnormalização (Rizzi, 2007).

Enquanto na modelagem relacional os primeiros passos são identificar as entidades e seus relacionamentos, na modelagem dimensional inicia-se com a identificação dos fatos, que são eventos do negócio, como por exemplo, a realização de uma venda, uma matrícula, o resultado de um índice financeiro. Após esta etapa, identificam-se as perspectivas sob as quais se deseja analisar o fato, chamadas de dimensões, por exemplo: pode-se analisar o fato venda por região ou então ano, mês ou ainda produto e faixa etária do cliente. Estas seriam as dimensões do fato venda. Outro exemplo é o registro de uma falha: a falha ocorre em um componente de um equipamento, em um dado momento e por determinada causa, neste caso o fato falha poderia ser analisado por causa, componente, momento (tempo: dia, minuto, segundo) e equipamento.

Observe-se que a dimensão tempo é comum em modelos dimensionais, a análise histórica é inerente ao DW. O mínimo que se deseja analisar de um fato é seu comportamento no decorrer do tempo.

O modelo dimensional é normalmente representado por um cubo, que é um objeto tridimensional, mas isto não quer dizer que se possa ter apenas três dimensões. O modelo é multidimensional e o cubo é apenas uma metáfora de como os dados estão organizados.

A Figura 3 apresenta a visualização da metáfora do cubo para representação das multidimensões de um DW.



Figura 3 - Visualização das dimensões de um cubo (Fonte: MSDN, 2014).

A visão por cubo gera a idéia de que os dados podem ser fatiados e representa claramente a sumarização dos mesmos, os registros não são visualizados individualmente.

Segundo Elmasri e Navathe (2011), a modelagem multidimensional subsidia a utilização de ferramentas OLAP, que permitem a navegação e exploração dos dados. Isto é realizado através de operações como “slice and dice”, onde se filtram os elementos de uma dimensão – por exemplo: visualizar as vendas da região sudeste ou apenas as vendas do estado de São Paulo - e drill-down e roll-up que faz a navegação entre os níveis de uma hierarquia dos dados - por exemplo, a partir da visão agregada das vendas em um ano, pode-se executar “drill-down” e as somas das vendas serão decompostas nos meses daquele ano; a operação de “roll-up” faz o inverso -. A realização de operações como estas em modelos relacionais envolve grande complexidade.

O esquema mais utilizado para a representação de data warehouses é o denominado modelo estrela, proposto por Kimball (1996). O modelo estrela recebe este nome porque sua forma possui uma tabela central (denominada fato) e outras ao redor, nas pontas da estrela (denominadas dimensões). A Figura 4 apresenta o modelo estrela.

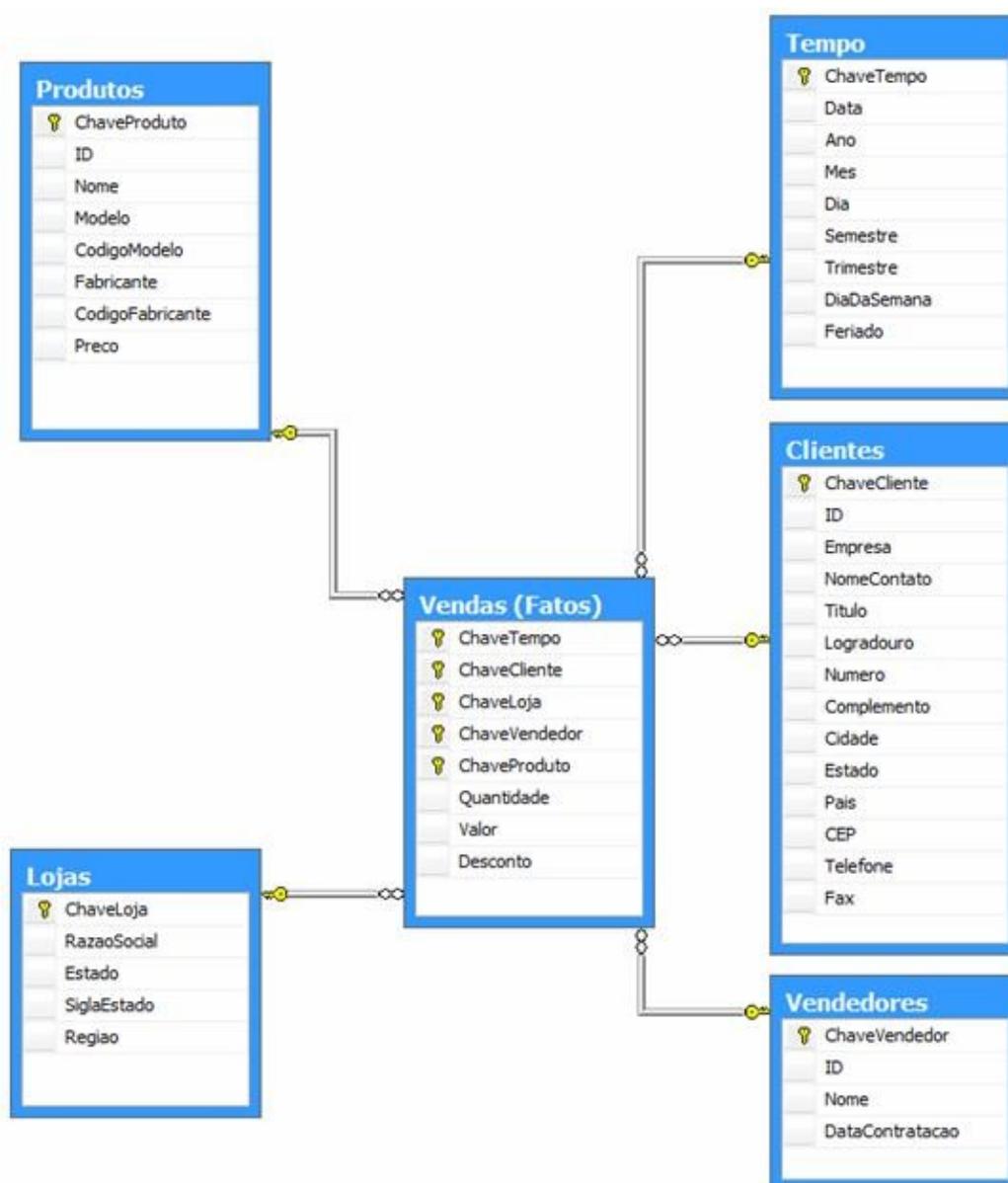


Figura 4 - Modelo Estrela (Fonte: MSDN, 2014).

Neste modelo as tabelas chamadas de dimensão possuem ligação direta com a tabela fato, por meio de chave estrangeira. Estas chaves identificam univocamente os registros da tabela fato. É também na tabela fato que ficam as medidas do negócio, por exemplo, no fato venda existem as medidas valor e quantidade.

Existem outros modelos para representação de data warehouses, mas o modelo estrela é o mais largamente usado por ser mais simples e expressivo, porque possui menos associações e por ser fácil de ser entendido mesmo por usuários finais.

2.3.6. Tabela Fato

Segundo Machado (2008), o fato é um evento do negócio, para entender fatos deve-se em primeiro lugar entender o negócio e depois pensar em quais informações seriam importantes para os gestores deste negócio. A origem do fato não está nos dados, mas sim no negócio.

Os atributos da tabela Fato que representam as mensurações acerca do negócio são denominados medidas ou métricas. As métricas devem ser sempre numéricas e variar no decorrer do tempo. As medidas são atributos com os quais se pode fazer cálculos de agregação como: soma, média, desvio-padrão, valor máximo, valor mínimo e contagem.

Um único fato pode ter inúmeras medidas, como por exemplo, no fato Vendas da figura anterior, a quantidade, desconto e valor são as medidas. A tabela Fato, por conter as métricas do negócio, é a que possui a maior quantidade de registros.

2.3.7. Tabelas Dimensão

Dimensões são as perspectivas segundo as quais deve ser analisado o fato, são os famosos “por”. Quando se analisa um fato como vendas (conforme figura 4), por exemplo, pode-se analisá-lo por loja, por produto, por vendedor, por tempo, por cliente. Note que sempre se deseja analisar os fatos segundo algum aspecto.

Machado (2008) comenta que se um fato qualquer acontece, ele tem participantes, indicação de tempo em que acontece, onde aconteceu, quem está no fato e o que está nele. Neste contexto ele explica que um fato possui quatro pontos cardeais: quem, o que, onde e quando.

Os atributos das dimensões são normalmente textuais, ou seja, não se espera efetuar cálculos com os mesmos. A quantidade de registros nas tabelas dimensão é pequena comparada com a tabela Fato.

As dimensões são compostas de hierarquias, que são os diversos níveis (atributos) da dimensão. A dimensão tempo possui vários níveis como ano, semestre, trimestre e mês; estes níveis formam a hierarquia. Outro exemplo é a dimensão loja que pode ter a seguinte hierarquia: região, estado.

A hierarquia permite a realização de operações OLAP como drill down e roll up, onde em um primeiro momento, por exemplo, o fato irá aparecer totalizado por ano, em

seguida poderá ser detalhado por semestre e, mais a frente, por mês. A possibilidade de em um primeiro momento ver os dados em um alto nível de sumarização e depois detalhá-los facilita a identificação de tendências e pontos fora da curva (Ponniah, 2001).

A hierarquia é a organização e classificação dos dados dentro de uma dimensão permitindo a navegação estruturada. As dimensões podem possuir uma ou várias hierarquias, como também a hierarquia pode possuir vários níveis ou apenas um.

2.4. Suítes de Business Intelligence

Suíte de Business Intelligence é um conjunto de softwares que coexistem em um mesmo ambiente, simplificando a gerência de usuários e seus acessos aos programas, compartilhando os acessos às bases de dados e criando interfaces simplificadas para o compartilhamento de documentos de análise dentro de uma organização.

Durante muito tempo, o uso de suítes de BI foi exclusividade de grandes organizações, em função de seu elevado custo de aquisição e licenciamento. O surgimento de suítes de BI de código aberto viabilizou a adoção de tais ferramentas por organizações de médio e pequeno porte.

O objetivo desta tese é implementar um sistema de Business Intelligence Espacial em apoio a uma propriedade rural que utiliza pressupostos básicos da agricultura de precisão de forma a estimular o uso de sistemas OLAP na atividade agrícola.

Para que tal estímulo seja possível, faz-se mister encontrar uma suíte de BI que atenda às seguintes premissas:

- não traga custos de aquisição ou licenciamento,
- seu estágio de desenvolvimento e utilização esteja maduro o suficiente para não trazer riscos de apostas infrutíferas em projetos incipientes de pesquisa que possam ser descontinuados a qualquer momento,
- possua suporte espacial ou geográfico, já que a espacialização dos dados é um pressuposto básico da agricultura de precisão.

Golfarelli (2008) realizou um comparativo entre as principais suítes de BI de código aberto. O estudo foi atualizado para esta tese, para contemplar as versões atuais das plataformas. Também Butler (2014) teceu comparações acerca de suítes de BI Open Source. Para ter certeza sobre o grau de utilização das suítes comparadas, de forma a atender ao requisito de maturidade das plataformas, foi utilizado um estudo realizado pelo Gartner Group (2014) acerca dos grandes representantes de suítes de BI.

Assim, o Quadro 2 lista as principais características acerca das três suítes de BI comparadas: JasperSoft (2014), Pentaho (2014) e SpagoBI (2014).

Quadro 2 - Comparativo Suítes de BI de Código Aberto

Módulo/Item	JasperSoft	Pentaho	SpagoBI
Sistema Operacional	Vários (JVM)	Vários (JVM)	Vários (JVM)
Linguagem desenvolvimento	Java	Java	Java
SGBD suportado	Todos (JDBC)	Todos (JDBC)	Todos (JDBC)
ETL	JasperETL	Pentaho Data Integration	Talend Open Studio
Agendador de tarefas	Quartz	Quartz	Quartz
OLAP	Mondrian, Jpivot	Mondrian, Jpivot	Mondrian, Jpivot, JPalo
Query by Example	-	-	Hibernate
Relatórios	JasperReports	Pentaho Report Designer, JasperReports, BIRT	JasperReports, BIRT
Servidor web	Tomcat	Tomcat	Tomcat
Dados Espaciais	Sim	Sim	Sim
Licença	Comercial (versão Enterprise) e Código Aberto (versão Community)	Comercial (versão Enterprise) e Código Aberto (versão Community)	Código Aberto

Além das três suítes anteriormente citadas, também a suíte ActuateOne (Actuate, 2014) foi avaliada, mas esta, por não possuir todas as funcionalidades principalmente relacionadas a data warehousing, limitando-se apenas a relatórios e análises, foi descartada.

As suítes de BI de código aberto possuem muitos módulos em comum, por exemplo, o servidor OLAP padrão para todas é o Mondrian (2014) e todas utilizam JasperReports como biblioteca para geração de relatórios. Todas fazem a integração de

projetos de código aberto já existentes (Ex.: Mondrian, JasperReports, etc) em uma única suíte de software, implementando apenas outros módulos e construindo a camada de integração entre as aplicações.

As suítes JasperSoft, Pentaho e mesmo a ActuateOne possuem dois tipos de licenciamento: um livre e outro comercial. A versão livre traz componentes com funcionalidades limitadas ou não fornece a integração completa entre todos os componentes na suíte, restando ao usuário trabalhar com cada software independentemente (já que cada software independente é um projeto de código aberto). O SpagoBI possui uma única versão (código aberto), provendo a integração completa entre seus componentes e por isso se autodeclarando como a única suíte 100% completa, flexível e de código aberto. Por ser a única fornecida sem restrições de funcionalidades e sem custo, atendendo aos requisitos anteriormente enunciados, o SpagoBI, desenvolvido pelo SpagoBI Competency Center, unidade de trabalho da Engineering Group – uma das 10 maiores empresas de TI da Europa -, foi selecionado para a implementação deste trabalho.

2.5. Trabalhos Relacionados

Com relação aos trabalhos que visam a criação de um sistema de suporte à decisão nos moldes do que pretende a presente proposta, podem-se destacar os seguintes:

Weber (2009) propõe um modelo de DW para gerenciamento de informações da produção de cereais. O trabalho utiliza um Data Warehouse construído sobre o SGBD MySQL e como ferramenta OLAP utiliza uma implementação proprietária em DELPHI para analisar os dados da tabela Fato. O projeto não trata de atributos espaciais.

Chaturvedi et al. (2006) apresentam a modelagem e principais problemas encontrados na implementação de um sistema OLAP para o Instituto de Pesquisas Estatísticas de Agricultura da Índia. O sistema utiliza uma ferramenta OLAP proprietária, IBM Cognos OLAP, e não utiliza tipos de dados espaciais.

O trabalho de Correa et al. (2009) implementa um DW para controle do mercado de milho e soja no Brasil, armazenando informações como taxas de mercado, valores de comercialização e etc. Posteriormente, os dados do DW são explorados por uma ferramenta OLAP, denominada Mondrian, que é de código aberto. O trabalho não trata de tipos de dados espaciais.

Nilakanta et al. (2008), assim como Chatuverdi et al. (2006), abordam as dificuldades na modelagem dimensional de um data warehouse para o Sistema de Informações Nacional Integrado de Recursos da Agricultura (INARIS). Os autores empregam a ferramenta proprietária IBM Cognos OLAP para a implementação do sistema e não tratam de atributos espaciais.

Com base nos parâmetros citados, foi percebido que os trabalhos relacionados não enfocam os mesmos objetivos desta tese. Os trabalhos de Weber (2009), Chatuverdi (2006) e Nilakanta (2008), apesar de fazerem a modelagem de um DW para a agricultura e de utilizarem uma aplicação OLAP para acessar seus dados – ou seja, empregarem tecnologias normalmente contidas em sistemas de BI -, utilizam ferramentas proprietárias, ou seja, cujos custos de aquisição são elevados para produtores rurais de médio porte. O trabalho de Correa et al. (2009) utiliza componentes de código aberto, mas o propósito não é o suporte à decisão para uma fazenda ou propriedade em particular, mas um controle de preços e taxas para a produção de milho e soja. Todos os trabalhos citados se abstiveram de tratar dados espaciais.

Cardoso (2008) propõe a implementação de uma ferramenta para OLAP espacial denominada MapWarehouse e, em seu estudo de caso apresenta algo que mais se aproximaria dos objetivos desta Tese. Todavia, como o objetivo do autor é apenas um estudo de caso para aplicação de sua ferramenta, faltou o passo a passo na elaboração da modelagem e solução de um problema de uma fazenda em específico. Também conta como ponto negativo o fato do MapWarehouse não estar disponível na web e não possuir uma comunidade de desenvolvimento estável que dê garantias mínimas de suporte ou continuidade na ferramenta.

3. MATERIAL E MÉTODOS

Para a implementação do sistema foram utilizados apenas componentes de software livre, de forma a viabilizar a reutilização de componentes da solução para o máximo de organizações interessadas em utilizar todo o potencial de um sistema de suporte à decisão que combine dimensões analíticas e geográficas.

Para a boa compreensão deste capítulo, inicia-se a condução do raciocínio apresentando a fazenda tema do estudo, seu modo de operação, como eram registrados seus dados espaciais e não-espaciais e alguns cenários típicos de análise. Em seguida apresenta-se a visão geral da solução proposta e seu detalhamento.

3.1. A fazenda em estudo

A fazenda ABADIA é uma propriedade produtora de cana-de-açúcar, localizada no município de Campos dos Goytacazes, RJ, com altitudes entre 5 e 10 metros, latitude de 21° 43' 54,25" Sul e longitude de 41° 12' 49,17" Oeste. O clima da região, segundo Köppen, é Aw. A temperatura média anual está em torno de 23,1°C, com média das máximas de 29°C e média das mínimas de 19°C. As precipitações médias anuais são de 884,8 mm, com 71% de concentração entre os meses de outubro e março. A área é de topografia plana (Bernardes, 2012).

Embora não utilize de forma intensiva a mecanização em todas as fases de sua atividade, a fazenda Abadia adota o pressuposto básico da agricultura de precisão: o

controle e gerência da propriedade em pequenas parcelas, de forma a administrar a variabilidade das propriedades de cada porção em busca da otimização da produção global. Assim, a área total da fazenda, de 2770 ha, foi dividida geograficamente em 27 (vinte e sete) setores, e estes em lotes (a quantidade de lotes em cada setor pode variar de 1 a 50), utilizando como critério básico a homogeneidade dos tipos de solo.

Em cada um dos lotes (unidade mínima de divisão espacial da fazenda), ocorre o plantio da cana de ano ou cana de ano-e-meio, condicionados principalmente pelo sistema de irrigação.

A atividade de produção de cana-de-açúcar na Fazenda Abadia segue os pressupostos contidos em um documento denominado “Planejamento Estratégico da produção de cana-de-açúcar da Fazenda Abadia” descrito em (Bernardes, 2002) elaborado por convênio com a ESALQ/FEALQ estabelecido em 1999 e cujas recomendações foram entregues em 2002.

Segundo Bernardes (2012), as áreas da fazenda são compostas por terrenos de baixada com solos aluviais, porém com terrenos variando de muito argilosos até arenosos. Tal característica dificulta o estabelecimento de um sistema de preparo, plantio e manejo uniforme em toda a fazenda.

A totalidade dos canaviais recebe correção do solo e adubação por ocasião do plantio. No plantio é realizada a correção do solo com pelo menos 1 t de calcário/ha e 0,5 t de gesso/ha e a adubação com 300 kg/ha de 4.30.16. Nas soqueiras, a fórmula de adubo mais utilizada é 20.00.20, com dose entre 300 kg/ha e 400 kg/ha, dependendo dos resultados da análise de solo do talhão.

A irrigação é realizada por sistemas de montagem direta no campo com canhão aspersor, por sulcos de infiltração e por pivô central. A irrigação assume grande importância visto que, segundo Lima (1984) e Fauconnier (1991), 1500 mm de precipitação pluvial anuais são requeridos para bons índices de produtividade da cana. Como na região Norte Fluminense a precipitação pluvial anual varia entre 800 mm e 1200 mm, resulta daí um déficit hídrico anual entre 400 mm a 600 mm, a ser compensado pelo uso da irrigação (Bernardes, 2012).

Para o controle das plantas daninhas, que prejudicam sobremaneira a produtividade da cana-de-açúcar, são empregadas as seguintes técnicas de manejo de acordo com o corte: aplicação de herbicida com trator – pré-emergente – para o primeiro corte; aplicação de herbicida com trator – pré e pós-emergente – no segundo e terceiro corte e emprego de cultivo mecânico para os demais cortes.

O Quadro 3 a seguir resume as atividades dispensadas pela fazenda Abadia a cada corte da cana durante os anos de 2002 a 2007, que foi o período de dados fornecido pela fazenda para este estudo.

Quadro 3 - Atividades por Corte – Fazenda Abadia

Atividade	Corte 1	Corte 2 e 3	Corte 4 ou maior
Adubação de fundação	04N-30P-16K (300Kg/ha)	Não se aplica	Não se aplica
Adubação de cobertura	Nenhuma	20N-0P-20K (300Kg/ha)	Nenhuma
Irrigação	Aspersão	Aspersão	Nenhuma
Controle daninhas	Aplicação de herbicida-trator (pré-emergente)	Aplicação de herbicida-trator (pré e pós-emergente)	Cultivo mecânico
Colheita	Manual	Manual	Manual

Para o controle de suas atividades, a fazenda Abadia mantém uma planilha a cada ano de produção, onde são registrados os seguintes dados: setor (onde ocorreu a produção), lote (de ocorrência da produção), área (colhida dentro do lote), número da folha (número do corte da cana-de-açúcar), estimativa de produção (em toneladas), estimativa de produtividade (em toneladas por hectare), produção (em toneladas), produtividade (em toneladas por hectare – obtida dividindo-se a produção pela área colhida), variedade da cana-de-açúcar e data do corte.

A Figura 5 apresenta a planilha usada para o controle da produção em cada lote de cada setor da fazenda Abadia para um determinado ano, sendo estes os únicos dados armazenados em formato digital.

FAZENDA ABADIA									
ESTIMATIVA DE CANA PARA SAFRA 2002/2003									
SETOR	LOTE	AREA	Nº FOLHA	TON/HÁ (EST.)	TOTAL (EST.)	TON/HÁ (REAL)	TOTAL (REAL)	VARIEDADE	DATA DO CORTE
1	01	4,18	1	90	376,20	117,39	490,67	RB 765418	18.04.02
1	02	4,25	1	90	382,50	106,55	452,82	RB 765418	18.04.02
1	03	18,56	1	90	1.670,40	124,80	2.316,36	RB 765418	16.04.02
1	03	0,38	0						
1	04	2,48	0			-		3250	safra 02/03
TOTAL		29,85			2.429,10	119,10	3.259,85		
						830,75			
2	1	3,85	0		-	-		SP 5122	safra 02/03
2	2	2,31	0		-	-		CB 453	safra 02/03
2	3	2,00	0		-	-		CB 453	safra 02/03
2	4	1,02	0		-	-		CB 453	safra 02/03
TOTAL		9,18			-	-			

Figura 5 - Planilha de controle da produção.

Os armários de arquivos da fazenda guardam em papel os dados contábeis necessários às prestações de contas com o Fisco. Neles estão dados como receitas, despesas e seus comprovantes; além de documentos referentes à análises e estudos realizados na propriedade.

Para registro da pluviometria, a fazenda Abadia dispõe de um pluviômetro, equipamento que faz a medição da quantidade de chuva recebida a cada dia. Estas quantidades de chuva medidas no pluviômetro são diariamente registradas em formulários de papel, e a soma de seus valores diários produz a ficha para controle pluviométrico mensal, conforme exibido na Figura 6.

BOLETIM METEOROLÓGICO

FAZENDA ABADIA 2003

Mês: Dezembro Ano: 03

Data	Pluviômetro	Tanque Classe " A "			Observação
		Leitura Anterior	Leitura Atual	Evaporação	
1		25,71	21,03	4,68	
2	15,00	21,03	30,41	—	
3	4,00	30,41	30,44	—	
4	2,00	30,44	24,08	6,36	
5	8,00	24,08	26,30	—	
6		26,30	22,05	4,25	
7		22,05	18,02	4,03	
8		18,02	13,31	4,71	
9	35,00	13,31	40,12	—	
10	5,00	40,12	41,80	—	
11		41,80	36,42	5,48	
12		36,42	38,38	4,10	
13	21,00	40,45	35,65	4,90	
14		35,65	30,55	5,10	
15		30,55	26,22	4,33	
16		26,22	28,52	3,70	
17		28,52	19,26	3,26	C.O TANQUE
18		60,42	53,30	3,10	
19		53,30	54,05	3,27	
20	2,00	54,05	50,90	3,15	
21		50,90	47,60	3,30	
22		47,60	44,20	3,40	
23		44,20	41,20	2,98	
24		41,20	35,36	5,86	
25		35,36	31,17	4,19	
26	18,00	31,17	28,68	—	
27		36,68	33,36	3,32	
28	7,00	33,36	36,25	—	
29		36,25	30,79	5,46	
30		30,79	24,00	6,79	
31		24,00	16,39	7,61	
Total do Mês	117,00			106,63	
Acumulado	1016,00			1585,04	

Figura 6 - Ficha para controle pluviométrico

Os dados espaciais existentes a respeito da fazenda Abadia (seus limites e suas divisões em setores e destes em lotes) estão representados por três plantas da propriedade, em formato .dwg (extensão de arquivos do software AutoCAD): a primeira

com o perímetro georeferenciado do imóvel, a segunda com a definição dos limites dos setores e lotes e a terceira contemplando um mapa de solos.

Com vistas a direcionar o presente trabalho, o gerente da fazenda elencou os seguintes cenários de análise:

- Identificação das variedades mais produtivas, de forma a selecionar aquelas que trarão o maior retorno econômico

- Distribuição das variedades pelas áreas da propriedade, com visualização em um mapa, de forma a visualizar o mix de variedades e suas concentrações

- Análise da produtividade de cada variedade em função do corte

- Distribuição dos cortes pelas áreas da fazenda, com visualização espacial, de forma a apoiar o estudo de renovação das áreas de plantio

- Análise das métricas relacionadas à produção, como por exemplo: produtividade, lucro, rentabilidade e lucro por hectare segundo o lote, tempo e corte.

Ressalte-se que, de posse dos dados no formato e ambientes atuais, a tomada de decisão por parte do administrador em quaisquer dos cenários citados anteriormente seria muito complexa, por causa da dificuldade em se explorar tais dados em busca das respostas desejadas. Imagine-se, por exemplo, marcar nas plantas existentes, em quais lotes se encontram quais cortes de cana.

Justamente com esta finalidade são projetados os sistemas de BI: fornecer acesso facilitado aos dados, prover capacidade exploratória sobre tais dados por meio de relatórios, gráficos, mapas e ferramentas OLAP para o levantamento de informações gerenciais.

3.2. Visão geral da solução – implantação de sistema de BI espacial

De forma a atender o gerente da fazenda em suas necessidades de suporte à decisão, caracterizadas nos cenários de análise descritos no item anterior, um sistema de BI espacial utilizando a suíte SpagoBI foi criado. A Figura 7 apresenta uma visão geral da solução.

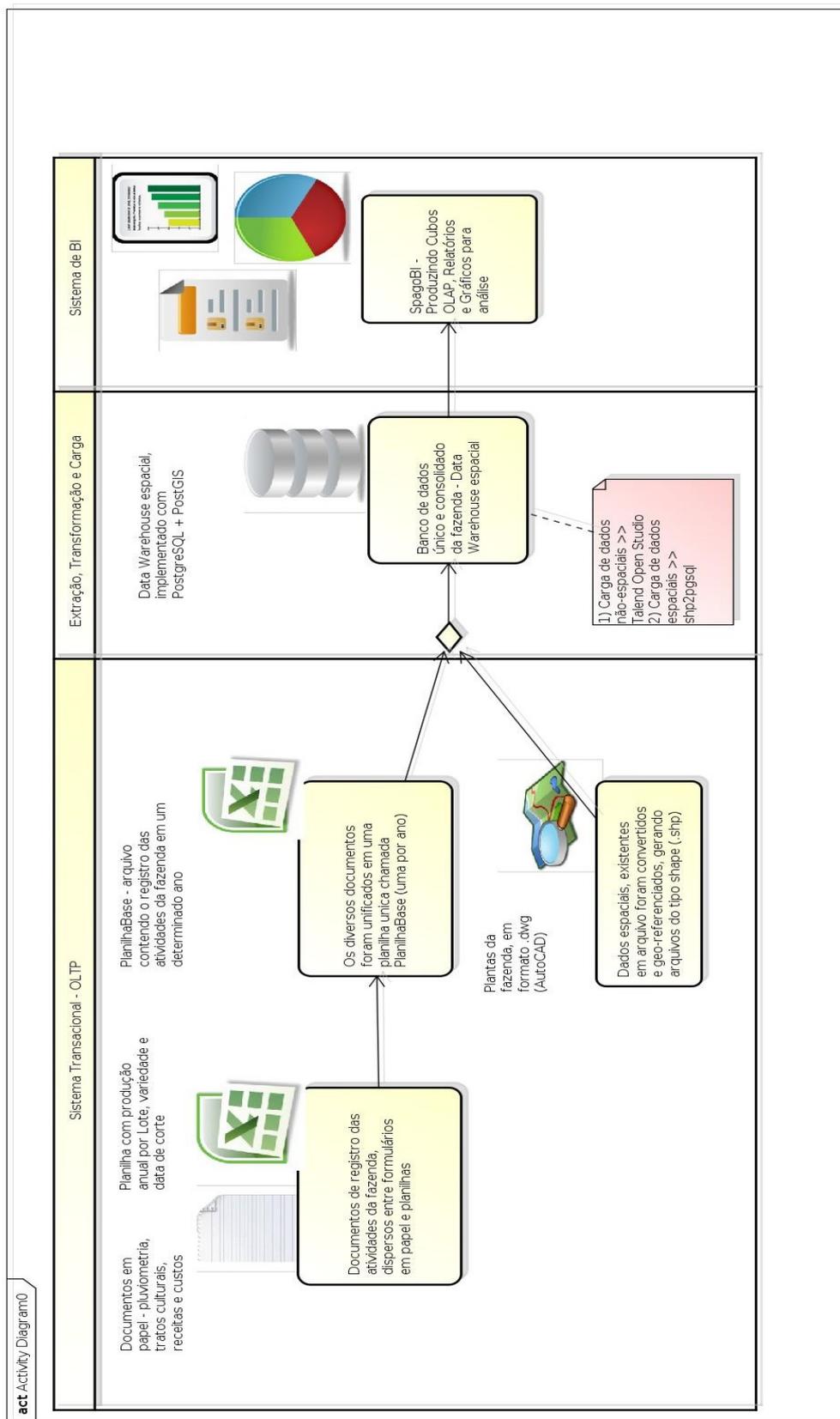


Figura 7 - Visão geral da implementação - sistema de BI

3.3. Obtenção dos dados existentes na fazenda

3.3.1. Dados não-espaciais

Para enriquecer as futuras análises no sistema de BI, novos dados acerca da produção passaram a ser buscados nos registros em papel do escritório da fazenda. Desta forma, foram procurados os dados acerca de tipo de colheita, tipo de irrigação, controle de plantas daninhas, adubação de fundação e adubação de cobertura. Também foram obtidos os valores de venda (preço por tonelada) para a cana-de-açúcar em cada um dos anos, permitindo o cálculo da receita.

Para que se obtivessem informações acerca de lucro e rentabilidade era necessário obter o custo gasto para a produção da cana-de-açúcar. O ideal seria que os custos estivessem detalhados e pormenorizados em suas parcelas, como por exemplo: custo de mão de obra, custo com insumos, custo com mudas, e outros, mas tal detalhamento não existia nos arquivos da fazenda. Decidiu-se, então, trabalhar com o custo total gasto pela fazenda em determinado ano, alocando-se proporcionalmente este valor a cada local em função dos tratos culturais ali desenvolvidos e da área do lote em questão.

Um outro dado bastante importante para a obtenção de informações futuras no sistema de BI é a quantidade de chuva ou índice pluviométrico. Como se sabe, a produtividade da cana-de-açúcar é bastante influenciada pela disponibilidade hídrica, e o índice pluviométrico é um dado fundamental principalmente para as áreas da fazenda que não receberam irrigação em função do número de corte.

Para que o sistema de BI pudesse utilizar os dados pluviométricos, os registros mensais existentes em papel foram digitados em uma planilha, e então a soma dos meses de chuva para cada data de colheita foi calculada e associada a cada registro de produção.

Para facilitar a análise posterior dos dados de pluviometria, a quantidade de chuva associada a cada produção será agrupada e classificada em faixas: até 800 mm, 801 a 900 mm, 901 a 1000 mm, 1001 a 1100 mm, 1101 a 1200 mm, e maior que 1200 mm.

Uma vez definidos todos os dados que fariam parte do sistema de BI, em função do custo-benefício de sua obtenção, a dificuldade seguinte foi a de obtê-los em formato digital para que pudessem posteriormente ser importados para o sistema de BI. Para tanto, uma planilha eletrônica, denominada “PlanilhaBase” foi criada para que nela fossem

lançados todos os dados não-espaciais referentes a um ano de produção. Desta forma, seis planilhas foram criadas (em verdade adicionou-se novos campos às antigas planilhas de controle) e preenchidas, uma para cada ano de 2002 a 2007, contendo todos os dados do sistema transacional da fazenda, refletindo os controles exercidos sobre a produção de cana-de-açúcar no citado período.

A PlanilhaBase (Figura 8) possui os seguintes campos:

1) Cultura: embora a fazenda Abadia tenha praticado a monocultura da cana-de-açúcar no período de 2002 a 2007 e, com isso, todos os registros estejam preenchidos com “Cana”, tal campo permite que o controle seja extensível a outros tipos de cultura.

2) Setor: número do setor (divisão espacial) da fazenda no qual aconteceu o registro.

3) Lote: identificação do lote (pertencente a um setor, ou seja, uma subdivisão espacial).

4) NrFolha: número do corte da cana-de-açúcar

5) Producao: quantidade produzida (em toneladas)

6) Variedade: identificação da variedade da cultura.

7) Data do corte: data em que ocorreu o corte da cana-de-açúcar

8) TipoColheita: tipo de colheita empregada (Ex.: manual, mecanizada)

9) TipoIrrigacao: tipo de irrigação utilizada (Ex.: aspersão, pivô central)

10) ControleDaninhas: método empregado para controle das plantas daninhas (Ex.: cultivo mecânico, herbicida, etc)

11) ClasseFundacao: esquema de adubação utilizado no plantio

12) ClasseCobertura: esquema de adubação utilizado na cobertura

13) QtdadeChuva: quantidade em mm de chuva recebida por determinada produção (contado desde o plantio até a colheita)

14) Receita: valor recebido pela venda da produção

15) Custo: valor total gasto com determinada produção

16) Area: área colhida no lote em questão

	A	B	C	D	F	G	H	I	J	K	L	M
	CULTURA	SETOR	LOTE	N°FOLHA	PRODUCAO	VARIEDADE	DATA DO CORTI	TIPOCOL	TIPOIRR	CONTROLADANI	CLASSE	CLASSECOBERTURA
2	Cana	1	01	1	490,67	RB 765418	18/04/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
3	Cana	1	02	1	452,82	RB 765418	18/04/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
4	Cana	1	03	1	2.316,36	RB 765418	18/04/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
5	Cana	3	01	2	339,81	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
6	Cana	3	02	2	345,03	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
7	Cana	3	04	3	198,65	RB 72454	29/06/2002	Manual	Nenhuma	Cultivo mecânico	Desconhecido	Nenhuma
8	Cana	3	05	2	81,72	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
9	Cana	3	06	2	152,88	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
10	Cana	3	07	2	52,49	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
11	Cana	3	08	2	29,07	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
12	Cana	3	09	4	18,14	RB 72454	29/06/2002	Manual	Nenhuma	Cultivo mecânico	Desconhecido	Nenhuma
13	Cana	3	10	2	95,88	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
14	Cana	3	11	2	74,53	RB 72454	29/06/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
15	Cana	3	12	4	95,61	RB 72454	29/06/2002	Manual	Nenhuma	Cultivo mecânico	Desconhecido	Nenhuma
16	Cana	4	01	5	131,65	RB 72454	29/06/2002	Manual	Nenhuma	Cultivo mecânico	Desconhecido	Nenhuma
17	Cana	4	02	1	159,23	SP 1011	18/04/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
18	Cana	4	03	1	460,17	RB 739735	04/05/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
19	Cana	4	04	1	401,52	SP 1011	04/05/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)
20	Cana	4	05	4	172,28	RB 739735	29/06/2002	Manual	Nenhuma	Cultivo mecânico	Desconhecido	Nenhuma
21	Cana	4	06	1	138,78	SP 701143	04/05/2002	Manual	Aspersão	aplicação de herbicida	Desconhecido	20N-0P-20K (300Kg/ha)

Figura 8 - PlanilhaBase – visão parcial

A PlanilhaBase tornou-se então a fonte dos dados não-espaciais que alimentarão o sistema de BI a ser construído. Como a espacialização das informações era um requisito necessário ao sistema, o próximo passo foi associar aos setores e lotes, já mencionados na PlanilhaBase, suas localizações geográficas.

Uma segunda planilha, denominada “PlanilhaLotes” (Figura 9), foi criada para associar a cada lote de um setor o seu tipo de solo com a descrição de tal tipo. Tais dados acerca do tipo de solo foram obtidos da planta de solos existente na fazenda. Como a divisão da fazenda em setores e lotes não sofreu alteração ao longo dos anos de produção, uma única planilha foi usada dispensando a redundância de se repetir, para cada lote e para cada ano de produção, o tipo de solo.

	A	B	C	D	E
1	SETOR	LOTE	TIPO SOLO	DESCRICAOSOLO	
2	1	1	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
3	1	2	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
4	1	3	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
5	1	4	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
6	2	1	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
7	2	2	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
8	2	3	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
9	2	4	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
10	3	1	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
11	3	2	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
12	3	3	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
13	3	4	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
14	3	5	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
15	3	6	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
16	3	7	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
17	3	8	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
18	3	9	GP1	Associação Gley pouco húmico Ta distrófica textura argilosa + Ta Eutrófico solodico textura argilosa/arenosa + Tm Eutrófico solodico textura argilosa + Ta Distrófica textu	
19	3	10	Ce1	Associação Cambissolo Tm Eutrófico e Distrófico argilosos + Cambissolo Ta Eutrófico argiloso + Aluvial Tm Eutrófico argiloso, bem drenado	
20	3	11	GP1	Associação Gley pouco húmico Ta distrófica textura argilosa + Ta Eutrófico solodico textura argilosa/arenosa + Tm Eutrófico solodico textura argilosa + Ta Distrófica textu	
21	3	12	GP1	Associação Gley pouco húmico Ta distrófica textura argilosa + Ta Eutrófico solodico textura argilosa/arenosa + Tm Eutrófico solodico textura argilosa + Ta Distrófica textu	

Figura 9 - PlanilhaLotes contendo o tipo de solo para os lotes da fazenda

3.3.2. Dados Espaciais

Como os dados espaciais do sistema serão armazenados em um Sistema de Informações Geográficas, é necessário que os dados da fazenda ABADIA estejam definidos em dos tipos de dados previstos pelo padrão OGC, ou seja: Ponto, Linha, Polígono, Mutiponto, Multilinha, Multipolígono e Coleção geométrica, que não são os tipos de dados contemplados por arquivos .dwg.

A ferramenta do tipo SIG (Sistema de Informações Geográficas) selecionada para tratar dos arquivos espaciais, preparando-os para a posterior importação para o sistema de BI, foi o quantumGIS.

Como o quantumGIS não suporta arquivos do tipo .dwg e a fazenda não dispunha do software AutoCAD (o serviço de geração das plantas foi prestado por empresa terceirizada), foi utilizado o software freeware (gratuito) DoubleCAD XT, que opera em ambiente Windows e é capaz de exportar o arquivo .dwg como arquivo .dxf, este sim suportado pelo quantumGIS.

Infelizmente a exportação dos arquivos .dwg em .dxf não gerou objetos compatíveis com os padrões OGC. Os setores e lotes, definidos espacialmente como polígonos segundo o OGC, foram exportados como polilinhas (objeto do AutoCAD), exigindo um pré-processamento trabalhoso dentro do quantumGIS de forma a criar os polígonos a partir do conjunto original das diversas polilinhas.

O pré-processamento no quantumGIS produziu um arquivo vetorial do tipo shape, de polígonos, com os limites de setores e lotes, mas sem referência espacial. Para georeferenciá-lo, foi usada uma planta da Fazenda que possuía uma lista de pontos distribuídos ao longo do perímetro da propriedade com suas coordenadas geográficas (utilizando SAD 69, UTM 24S, EPSG:29184).

Para fazer a correspondência entre os pontos no shape ainda não georeferenciado e as coordenadas conhecidas, utilizou-se o software gvSIG, em cujo plugin de referenciamento, gerou-se uma “matriz Affine” de 6 valores (correspondentes à Escala X, Escala Y, Rotação X, Rotação Y, Deslocamento X e Deslocamento Y).

Os valores obtidos da transformação Affine no gvSIG foram utilizados em um outro plugin, agora no software quantumGIS, que promoveu a efetiva transposição dos polígonos antes em um sistema de referência local para seus valores no sistema de referência de Coordenadas (SRC) 29184 (SAD 69 – UTM 24S). Para uso no SpagoBI, os dados espaciais devem estar em coordenadas geográficas e não planas. Para isso, novamente o quantumGIS foi usado para salvar as camadas de setores e lotes em outro arquivo shape, com novo sistema de referência de Coordenadas, o de número 4674, correspondente a SIRGAS2000, lat/lon.

3.4. Modelo OLAP para o sistema de BI

Após a decisão de quais dados deveriam constituir o sistema de BI, e definir que a granularidade de tais dados seria a produção (quantidade produzida em toneladas) de determinada variedade cultural em um lote (menor unidade espacial da propriedade) e em determinado mês e ano, o modelo OLAP (On-Line Analytical Processing) da Figura 10 foi proposto:

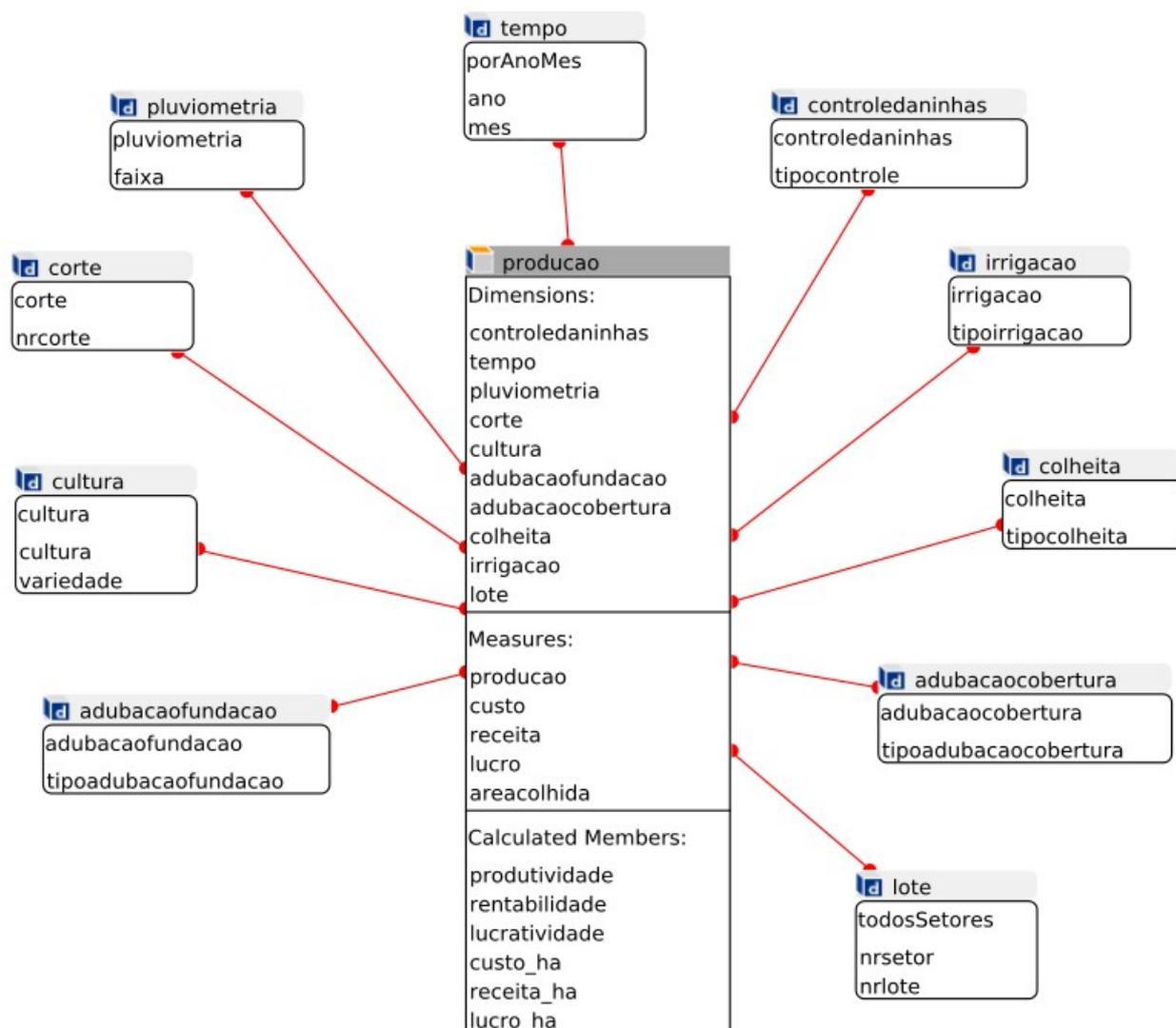


Figura 10 - Modelo Olap para sistema de BI – Fazenda Abadia

O modelo OLAP para o sistema compreende a tabela Fato central, denominada “producao” e as tabelas chamadas dimensão, que estão em torno da tabela Fato.

Na tabela Fato (“producao”), estão as métricas da produção da fazenda a serem avaliadas:

- producao: quantidade (em toneladas) produzida
- custo: valor em reais gasto para aquela produção. Todos os custos estão totalizados neste valor, ou seja, soma de mão-de-obra, insumos, aluguel de máquinas, etc
- receita: valor em reais recebido pela comercialização da produção
- lucro: diferença em reais entre a receita e o custo
- produtividade: produção por unidade de área, expresso em toneladas por hectare

- rentabilidade: expresso por lucro/custo * 100
- lucratividade: expresso por lucro/receita * 100
- custo_ha: custo por unidade de área, expresso em reais por hectare
- receita_ha: receita por unidade de área, expresso em reais por hectare
- lucro_ha: lucro por unidade de área, expresso em reais por hectare

Os dados de produção da fazenda, contidos na tabela Fato serão analisados segundo várias perspectivas, aqui denominadas “dimensões”. São elas que permitirão ao administrador da fazenda mensurar a influência de cada um dos parâmetros na produção, auxiliando-o na busca da maximização da produtividade. Segue descrição de cada dimensão:

3.4.1. Cultura

É a cultura vegetal que foi plantada e para a qual se deseja medir resultados. No caso da Fazenda Abadia, trata-se de monocultura de cana-de-açúcar, mas o sistema foi projetado para trabalhar com quaisquer quantidade de diferentes culturas simultaneamente. Nesta dimensão também se trata a variedade da cultura vegetal em análise. Para a fazenda Abadia, que se ocupa exclusivamente da cana-de-açúcar, algumas das variedades trabalhadas são: CB45-3, CP51-22, RB72454, RB739735, RB765418, SP79-1011, SP80-1842, SP70-1143, SP81-3250, SP79-2313, SP79-2233 dentre outras.

3.4.2. Corte

É o número do corte a que se refere a colheita em questão.

3.4.3. Lote

Traz os dados acerca do local onde ocorreu a colheita, no caso da fazenda Abadia, contempla o setor e lote.

3.4.4. Tempo

Como em qualquer sistema de Business Intelligence, construído sobre um Data Warehouse, a dimensão tempo é onipresente e responsável por permitir análises sobre a dimensão temporal. No caso da fazenda Abadia, apenas os dados relativos ao mês e ano são relevantes (o dia não importa) para a indicação da época em que ocorreu a produção ou, no caso em especial, em que foi feito o corte da cana.

3.4.5. Tipo de Colheita

Para a colheita da Cana-de-açúcar, a fazenda Abadia poderia utilizar dois métodos: Manual e Mecanizada. No período dos dados fornecidos para este estudo, de 2002 a 2007, foi utilizada apenas a colheita Manual.

3.4.6. Irrigação

Os seguintes métodos de irrigação foram empregados pela fazenda ao longo do tempo: Aspersão (com montagem direta), por sulco, por pivot central ou nenhum.

3.4.7. Controle de plantas daninhas

Esta dimensão trata os diferentes métodos utilizados pela fazenda para controle de plantas daninhas.

3.4.8. Adubação de plantio (Classe de fundação)

Representa os diferentes esquemas de adubação de plantio utilizados em cada produção. Como os registros de produção da fazenda iniciaram em 2002, apenas os cortes plantados a partir de 2002 possuem tal dado, os cortes cujo plantio ocorreu antes do citado ano receberam o valor “Desconhecido”.

3.4.9. Adubação de Cobertura (Classe de cobertura)

Representa os diferentes esquemas de adubação de cobertura utilizados em cada produção.

3.4.10. Pluviometria

São as faixas em que se classificou a quantidade de chuva que cada colheita recebeu: até 800 mm, 801 a 900 mm, 901 a 1000 mm, 1001 a 1100 mm, 1101 a 1200 mm, e maior que 1200 mm

3.5. Modelo físico do Data Warehouse e carga de dados

Para armazenar os dados oriundos do sistema OLTP da fazenda, ou seja, dos seis arquivos “PlanilhaBase” (período de 2002 a 2007), foi modelado o DW de acordo com a Figura 11.

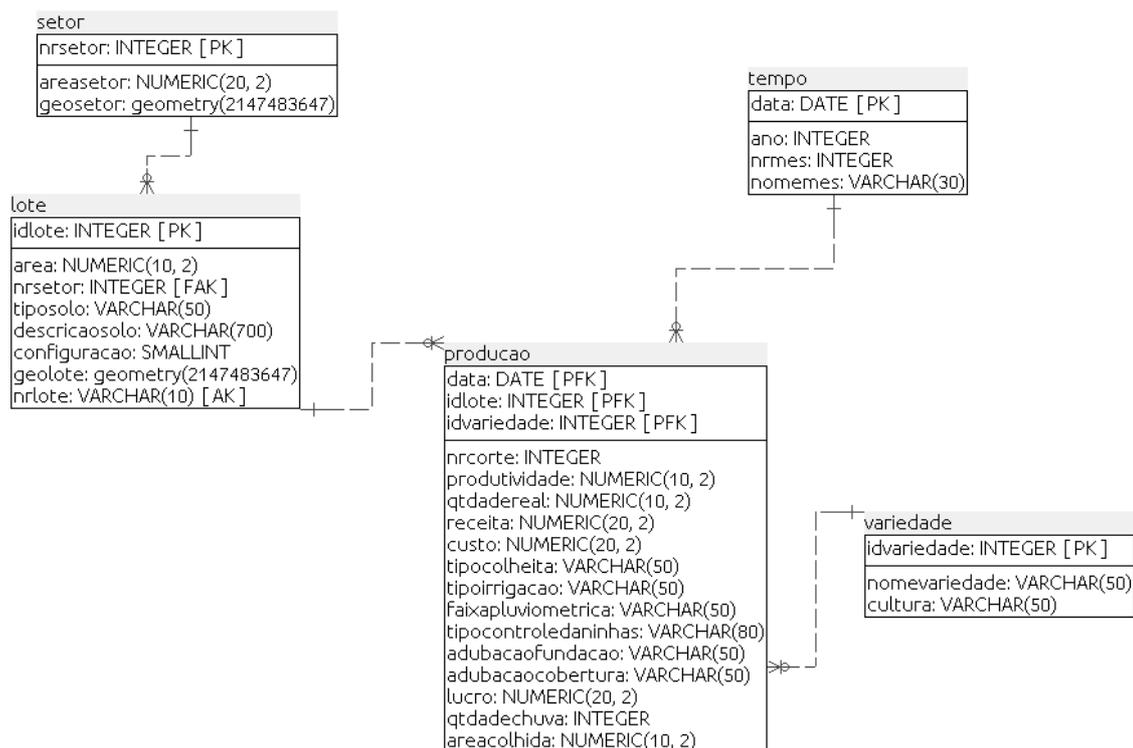


Figura 11 - Diagrama de Estrutura de Dados (DED) – DW Fazenda Abadia.

O DW do modelo da figura anterior foi implementado em um Sistema de gerenciamento de banco de dados (SGBD) espacial e recebeu o nome de “dwfazenda”.

Como a proposta é a de um sistema inteiramente opensource, foi utilizado o PostgreSQL como SGBD, adicionado do PostGIS, sua extensão geográfica, obtendo-se assim um SGBD espacial de excelente reputação e aderente às especificações do OGC (OpenGIS Consortium).

Além das tabelas presentes no DED da figura anterior, as tabelas “etlgeral” e “etllote” foram criadas. Estas duas tabelas utilitárias, que compõem a denominada “Data Staging Area”, ou área de preparação dos dados, possuem a finalidade de simplificar a migração de dados.

A instalação do banco de dados espacial seguiu o roteiro descrito em Benigno (2014), no caso da instalação em uma plataforma com Ubuntu Linux, e a sequência contida em Santos (2014) para a plataforma Windows.

3.5.1. Carga dos dados não espaciais

Uma vez o banco de dados espacial criado, o próximo passo é carregá-lo com os dados oriundos dos arquivos “PlanilhaBase”. Para isso, o software Talend Open Studio for Data Integration (TOS, 2014), que possui integração com a suíte SpagoBI será utilizado.

O Talend Open Studio é uma ferramenta de ETL (Extraction, Transform and Load, ou traduzindo: Extração, Transformação e Carga), que se destina a copiar e manipular dados de uma base de dados de origem qualquer (planilhas, arquivos de texto, banco de dados, etc) para uma base de dados de destino. No caso da presente tese, o Talend será usado para copiar os dados dos arquivos “PlanilhaBase” presentes em um diretório qualquer, para o data warehouse espacial (dwfazenda), presente no PostgreSQL + PostGIS.

O uso do Talend Open Studio for Data Integration se faz por meio de “Jobs” (tarefas, na tradução do inglês). No caso desta tese, dois “Jobs” foram definidos: um para a cópia dos dados da “PlanilhaLotes” para a tabela “etllote” do banco de dados dwfazenda e outro para cópia dos dados da planilha denominada “PlanilhaBase” para uma tabela chamada de “etlgeral” existente no mesmo dwfazenda, que será o repositório de dados para o sistema de BI.

A Figura 12 exemplifica um dos “Jobs” de migração dos dados, no caso o referente à cópia dos registros de “PlanilhaBase” para a tabela “etlgeral” do banco de dados dwfazenda.

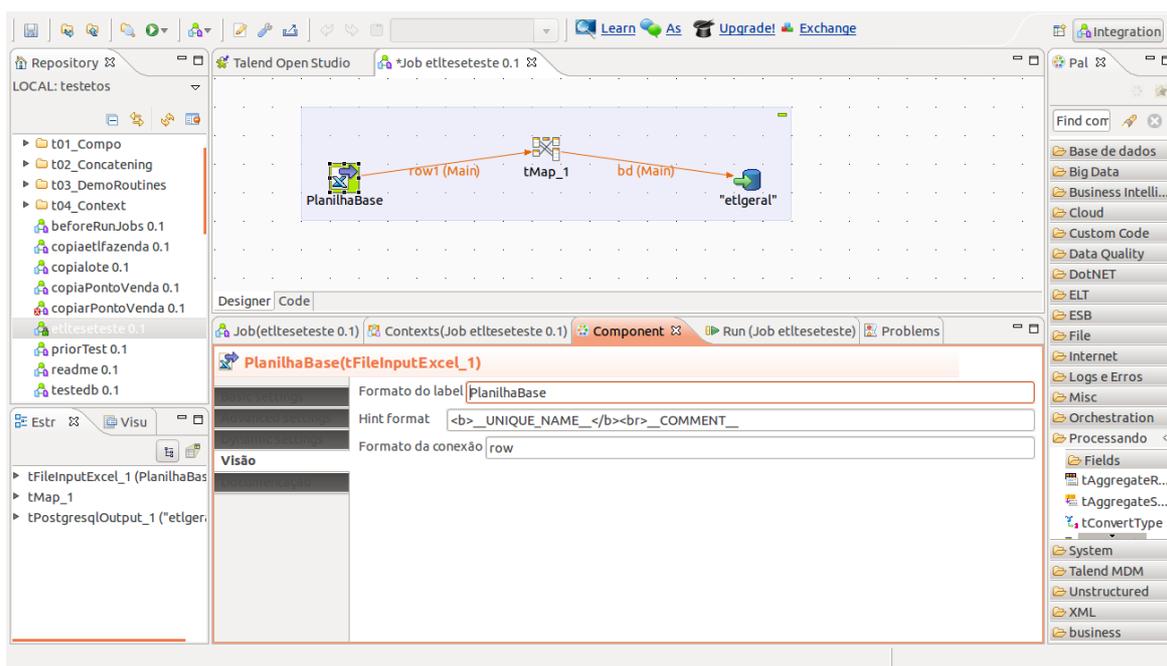


Figura 12 - Talend Open Studio – Tela de Definição do Job.

A Figura 13 exibe o mapeamento entre os campos da planilha de origem dos dados (PlanilhaBase) e os campos da tabela de destino para os mesmos (etlgeral).

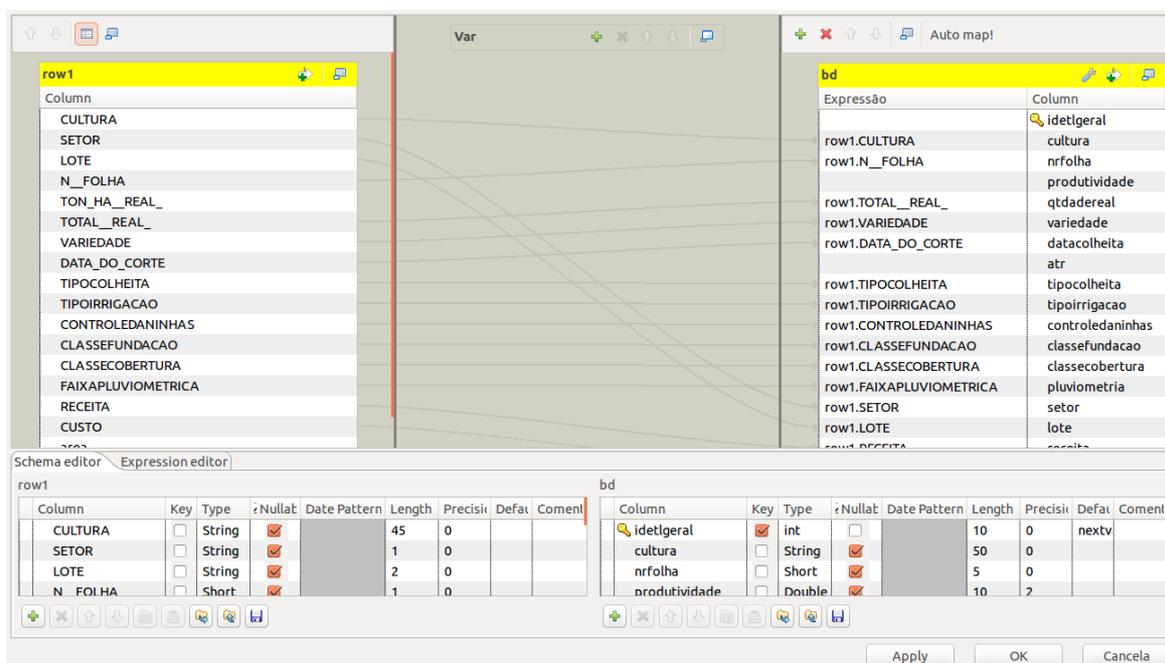


Figura 13 - Mapeamento Origem x Destino – Talend Open Studio

Após a cópia dos dados para as tabelas etllote e etlgeral, pertencentes ao dwfazenda, foram executados scripts SQL para distribuir os dados nas diversas tabelas

do Data Warehouse, preenchendo assim todos os campos das tabelas Fato e Dimensões, à exceção dos campos que recebem valores espaciais, como geolote e geosetor presentes respectivamente nas tabelas lote e setor. Tais atributos espaciais serão preenchidos conforme descrito no próximo tópico: “carga dos dados espaciais”.

3.5.2. Carga dos dados espaciais

Para a inserção dos dados espaciais (presentes nos dois arquivos shape – um referente a setores e outro referente aos lotes - gerados com o emprego do quantumGIS) no Data Warehouse espacial (PostgreSQL com a extensão do PostGIS) foi utilizado o aplicativo de linha de comando “shp2pgsql”, que gera scripts SQL de inserção dos dados espaciais de arquivos shape (.shp) em tabelas do PostgreSQL.

O comando a ser executado é: *shp2pgsql -s 4674 -W "LATIN1" -t "2D" <origem>/setoresSRC4674.shp setores > <localDesejado>/setoresAbadia.sql*

Os parâmetros do comando utilizados foram:

-s 4674: especifica que os dados estarão no sistema de referência de coordenadas de número 4674 (SIRGAS2000, lat/lon, ou seja, coordenadas geográficas)

-W “LATIN1”: especifica a codificação de caracteres (para dar suporte a acentos, cedilha, etc)

-t “2D”: indica que os objetos espaciais são bidimensionais

O resultado do emprego do aplicativo shp2pgsql é um arquivo (no exemplo citado anteriormente, com o nome “setoresAbadia.sql”) com um script de criação e inserção dos dados espaciais em uma tabela de nome (também segundo o exemplo anterior) “setores”.

Tal script SQL será executado no banco de dados espacial (dwfazenda), gerando uma tabela de uso temporário com os setores; um script análogo fará o mesmo para os lotes.

Para atualizar a área dos setores na tabela de setor (atributo “area”), foi utilizada a capacidade espacial do banco de dados. O seguinte comando foi utilizado (a transformação é necessária pois os dados estavam no SRC 4674, que é do tipo lat/lon; mas para cálculo da área é necessário um SRC projetado, no caso o SRC 31984, referente ao SIRGAS2000, UTM 24S): *update setor set areasetor = (select ST_Area(ST_Transform(geosetor,31984))/10000 from setor st where setor.nrsetor = st.nrsetor)*. Um comando similar foi empregado para calcular a área dos lotes, armazenada no atributo “arealote” da tabela “lote”.

Uma vez os dados espaciais presentes nas tabelas de uso temporário, um conjunto de scripts SQL foi utilizado para copiar tais dados para as tabelas definitivas “setor” e “lote”, integrando na mesma tabela atributos espaciais como o Polígono que representa um lote e atributos não-espaciais, como por exemplo o tipo de solo.

3.6. Instalação e configuração da suite de BI - SpagoBI

De posse do data warehouse espacial carregado (dwfazenda), o próximo passo para a construção da solução é a instalação da suite SpagoBI. Para tanto, é necessário que o JDK (Java Development Kit), versão 6 ou superior esteja instalado, e que se configure duas variáveis de ambiente: JAVA_HOME, com o caminho completo para o diretório de instalação do JDK e CATALINA_HOME, com o caminho para o diretório de instalação do SpagoBI.

Após o download do pacote do SpagoBI a partir de seu site, basta descompactar o arquivo (sem necessidade de instalação) para usá-lo. Como o presente trabalho propõe uma suíte de BI espacial, é necessário também instalar o GeoServer (2014).

Segundo GeoServer (2014), este é um servidor de código aberto escrito em Java que permite aos usuários compartilhar e editar dados geoespaciais. Utilizando padrões abertos definidos pela Open Geospatial Consortium (OGC), o GeoServer permite uma grande flexibilidade na criação de mapas e compartilhamento de dados.

Em função das características do GeoServer, o mesmo foi selecionado pelos desenvolvedores do SpagoBI para fornecer as capacidades espaciais necessárias à suíte. Para tanto, é necessário fazer o download do geoserver a partir de seu site, descompactar o arquivo “geoserver.war” e copiar a pasta descompactada de nome “geoserver” para dentro da pasta “webapps” localizada no diretório de instalação do SpagoBI.

O SpagoBI é uma aplicação web que executa sobre um servidor Apache Tomcat embutido, bastando para ser iniciado a execução de um script denominado “SpagoBIStartup”. Depois de iniciado, para navegar em suas funcionalidades, deve-se digitar em um browser web: <http://<Endereço IP do servidor>:8080/SpagoBI>.

A Figura 14 apresenta a tela inicial da suíte SpagoBI, exibindo seus dados de demonstração.

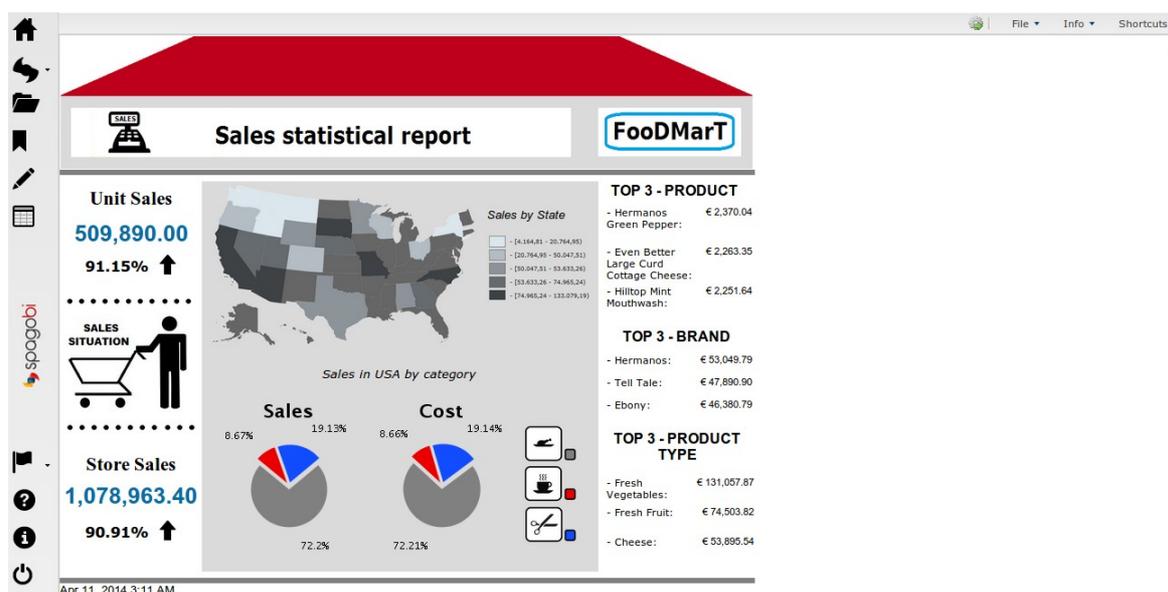


Figura 14 - Tela inicial de demonstração da suíte SpagoBI

O SpagoBI vem pré-configurado com um banco de dados de demonstração, simulando dados de uma empresa de venda de produtos alimentícios chamada “Foodmart”.

Para seu funcionamento, a suíte acessa dois bancos de dados (pelo menos): o primeiro deles destinado a armazenar os metadados da aplicação, ou seja, os dados de usuários, menus, relatórios, cubos olap, mapas e outros; e o segundo destinado ao data warehouse do usuário. Assim, a primeira tarefa para configurar o SpagoBI para uso em produção é informar por meio de arquivos xml de configuração, a localização de seu banco de metadados, o que deve ser feito de acordo com o descrito no anexo 1 – Configuração dos metadados SpagoBI.

O SpagoBI vem configurado para quatro idiomas: italiano, inglês, francês e espanhol, regionalizando além dos textos dos menus, as representações de data e moeda. Assim, uma providência importante para melhorar a experiência do usuário é fazer a regionalização da suíte adaptando-a para o português do Brasil (pt-BR). Tal procedimento foi executado adotando as diretrizes contidas em SpagoBI (2014).

3.7. Criação de documentos

Após a instalação e configuração do SpagoBI, o próximo passo na implementação da solução é a criação dos relatórios, cubos Olap, painéis de controle (dashboards) e mapas, denominados de forma genérica pela suíte como documentos.

Documento é a unidade básica de criação no SpagoBI, e para cada um deles existe um modelo em xml que o define. Para a criação de um novo documento, a suíte apresenta uma interface, apresentada na Figura 15, para receber o nome do citado documento, seu tipo (especificando ser um relatório, um mapa, etc), o arquivo xml de definição e o conjunto de parâmetros exigidos para sua execução.

The screenshot shows the 'DOCUMENT DETAILS' window in SpagoBI. The interface is divided into several sections:

- Form Fields:**
 - Label:** Text input field with an asterisk (*).
 - Name:** Text input field with an asterisk (*).
 - Description:** Text input field.
 - Type:** Dropdown menu with 'On-line analytical processing' selected.
 - Engine:** Dropdown menu with 'JPivot Engine' selected.
 - Data Source:** Dropdown menu.
 - State:** Dropdown menu with 'Development' selected.
 - Refresh seconds:** Text input field with '0' entered.
 - Criptable:** Radio buttons for 'True' and 'False', with 'False' selected.
 - Visible:** Radio buttons for 'True' and 'False', with 'True' selected.
 - Visibility restrictions:** Text input field with a dropdown arrow, an equals sign, and a plus icon.
 - Template:** Text input field with 'Escolher arquivo' and 'Nenhum arquivo selecionado'.
 - Template build:** Icon for building the template.
- Navigation:** A vertical sidebar on the left contains icons for home, search, settings, users, charts, and a SpagoBI logo.
- Templates:** A section titled 'Show document templates' on the right shows a 'Functionalities Tree' with folders for 'Analiticos', 'Espacial', and 'Compostos', and a 'Personal Folders' section.

Figura 15 - SpagoBI – Tela para criação de novo documento

A Figura 16 traz um exemplo de arquivo xml para definição de um documento do SpagoBI.

```

dialChartProdxLote x
1 <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2 <DIALCHART type='meter' name='Rentabilidade deste Lote'|>
3   <DIMENSION width='200' height='200' />
4   <STYLE_TITLE font='Arial' size='14' color='#000000' />
5   <STYLE_VALUE_LABEL font='Arial' size='9' color='#000000' />
6   <STYLE_TICK_LABELS font='Arial' size='9' color='#000000' />
7   <CONF>
8     <PARAMETER name='orientation' value='horizontal' />
9     <PARAMETER name='lower' value='0' />
10    <PARAMETER name='upper' value='100' />
11    <PARAMETER name='increment' value='25' />
12    <PARAMETER name='minortickcount' value='25' />
13    <INTERVALS>
14      <INTERVAL min='0.0' max='30.0' color='#FF3333' label ='Baixa' />
15      <INTERVAL min='30.0' max='80.0' color='#FFCC00' label ='Média' />
16      <INTERVAL min='80.0' max='100.0' color='#66FF00' label ='Excelente' />
17    </INTERVALS>
18  </CONF>
19 </DIALCHART>

```

Figura 16 - Exemplo arquivo xml de um documento – “Gráfico Velocímetro”

Para a criação dos arquivos xml, a equipe de desenvolvimento do SpagoBI disponibiliza o SpagoBI Studio, um software para simplificar, por meio do uso de uma interface gráfica demonstrada na Figura 17, e assistentes, a geração dos arquivos para os documentos.

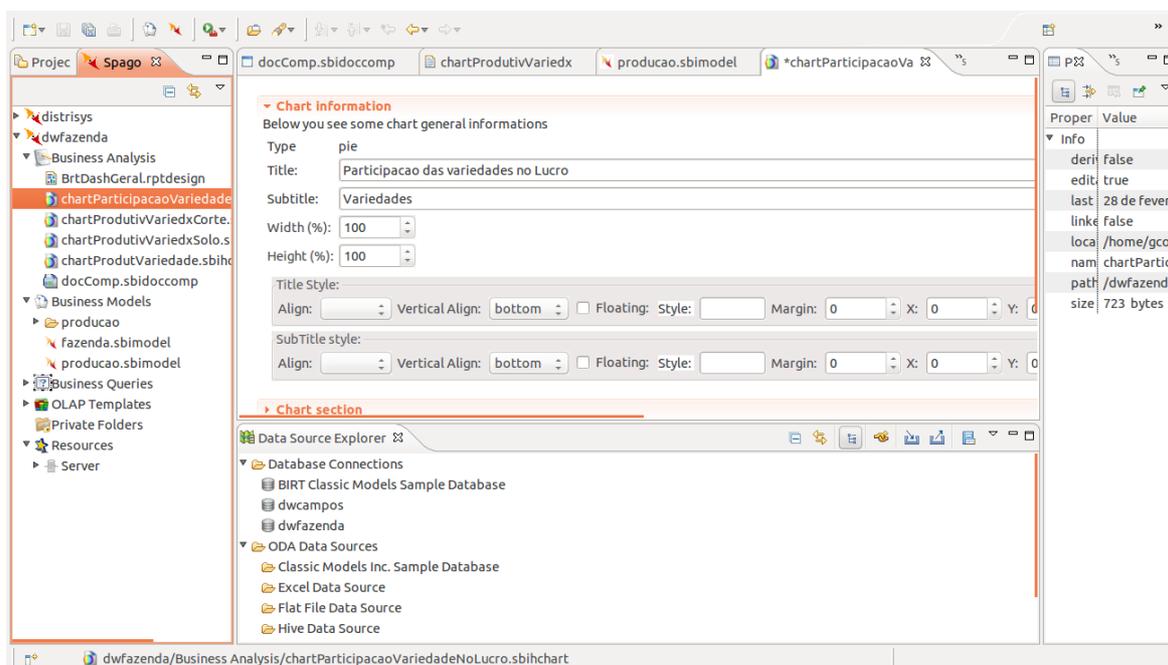


Figura 17 - SpagoBI Studio

Por meio do SpagoBI Studio também é possível fazer conexão com o banco de dados – data warehouse - , criar os chamados “modelos de negócio”, que são representações lógicas dos dados sobre os quais os futuros documentos serão criados, e ainda fazer o upload dos documentos criados na plataforma (Studio) para o servidor da suíte (SpagoBI).

A suíte SpagoBI apresenta uma grande variedade de ferramentas para a criação dos documentos de análise, também chamados pela suíte de “engines” (motores). Desta forma, por exemplo, se o usuário deseja implementar um relatório, poderá fazê-lo usando a ferramenta (“engine”) JasperReports ou BIRT. Outra facilidade apresentada pela suíte é a possibilidade de, a partir de um documento, o usuário, clicando sobre um dado ou porção de gráfico, ser redirecionado para outro documento que o detalhe ou contextualize. Tal capacidade, denominada navegação cruzada (“cross-navigation”), amplia bastante as capacidades do SpagoBI.

De forma a selecionar quais documentos implementar, utilizou-se como guia os cenários de análise solicitados pelo gerente da fazenda Abadia, conforme elencados no item 3.5 deste trabalho.

Para os cenários não-espaciais, ou seja, aqueles que não requerem visualização em um mapa, utilizou-se preferencialmente uma aplicação Olap, por ser o documento que proporciona a maior flexibilidade de análise ao usuário, ou painéis de controle (“dashboards”), que integram em uma única tela, dois ou mais documentos.

Nos cenários espaciais, usou-se o documento do tipo “Location Intelligence” e subtipo “GIS Engine”, que exibe os dados espaciais das dimensões utilizando-se de um Sistema de Informações Geográficas (GeoServer), projetando-os sobre outras camadas geográficas que servem de contexto espacial para a visualização, como por exemplo o Google Maps.

3.7.1. Criação do cubo Olap

Assim, para atender ao cenário de análise: “exibir a evolução das métricas da produção – quantidade, produtividade, rentabilidade – por corte e variedade ao longo do tempo”, foi usado um documento Olap.

Para a criação do cubo Olap, que é a implementação do modelo descrito no item 3.8 desta tese, foi utilizado o SpagoBI Studio, onde após a criação da conexão com o data

warehouse (“dwfazenda”), adicionou-se um novo modelo de negócio (“Business Model”) de nome “producao.sbimodel” com a definição da tabela Fato e suas métricas e das tabelas Dimensão com suas hierarquias e seus atributos. O passo a passo na ferramenta é descrito em Ogutu (2014). A Figura 18 apresenta a definição do modelo Olap no SpagoBI Studio.

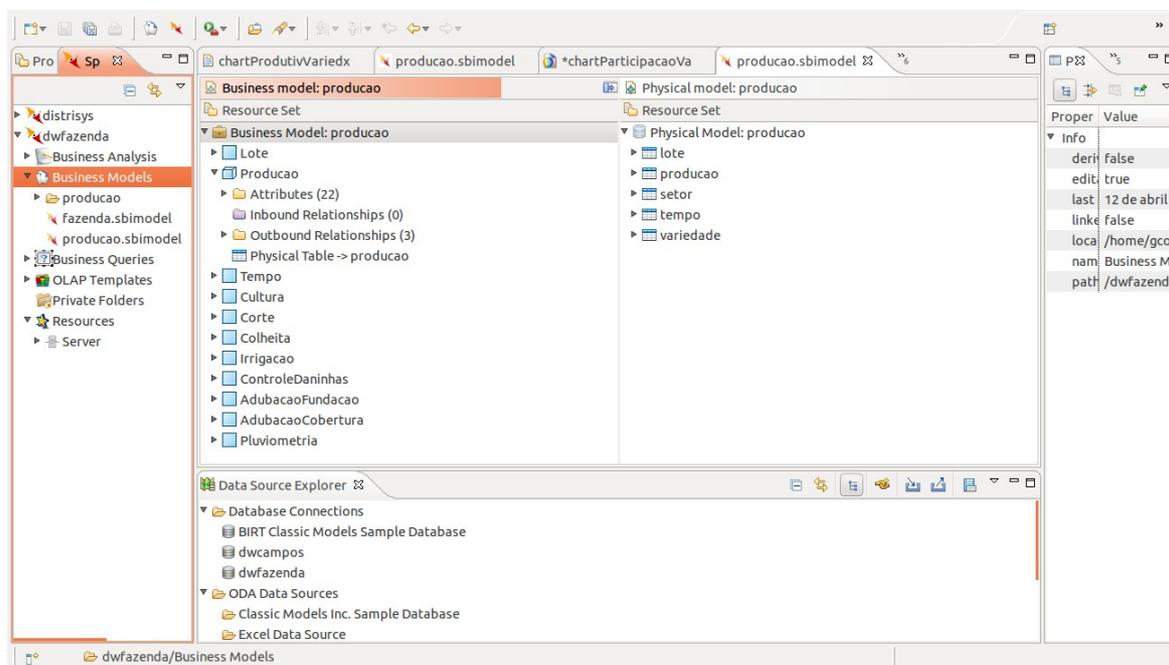


Figura 18 - Definição do modelo Olap no SpagoBI Studio

Após a criação do cubo Olap, gera-se, ainda no Studio, o “mondrian schema”, que é a definição em um arquivo xml das características do citado cubo para ser lido pelo mondrian que é o motor Olap utilizado pelo SpagoBI. Detalhes deste procedimento também podem ser encontrados em Ogutu (2014).

Uma vez criados o cubo Olap, denominado “producao.sbimodel” dentro da pasta “Business Models” e o mondrian schema, faz-se o upload dos mesmos para o servidor SpagoBI, onde o documento Olap será criado.

No servidor SpagoBI, cria-se um novo documento, conforme a Figura 19. A Figura 20 traz a visão inicial do cubo Olap gerado.

DOCUMENT DETAILS

Label	OlapProdCorteVaried *
Name	OlapProdCorteVaried *
Description	Olap para análise da produção por Cort
Type	On-line analytical processing
Engine	JPivot Engine
Data Source	dsdwfazenda
State	Development
Refresh seconds	0
Criptable	<input type="radio"/> True <input checked="" type="radio"/> False
Visible	<input checked="" type="radio"/> True <input type="radio"/> False
Visibility restrictions	
Template	Escolher arquivo Nenhum arquivo selecionado
Template build	

Show document templates

- Functionalities Tree
 - Analiticos
 - Espacial
 - Compostos
 - Personal Folders

Figura 19 - Criação novo documento do tipo Olap

Select schema: olapfazenda20140302

Select cube: Producao



Columns

- Tempo
- Measures

Rows

- Corte
- Cultura

Filter

- AdubacaoCobertura
- AdubacaoFundacao
- Colheita
- ControleDaninhas
- Irrigacao
- Lote
- Pluviometria

OK Cancel

		Tempo		
		+Todos_AnoMes		
		Measures		
Corte	Cultura	Producao	Rentabilidade	Produtividade
+All Corte	+Todos_CulturaVariedade	387,926.30	11.80	55.90

Figura 20 - Definição da visão inicial para o cubo Olap

Clicando o ícone “Template Build”, localizado na parte inferior da tela (Figura 20), o desenvolvedor pode definir qual visualização inicial do cubo Olap o usuário terá quando abrir o documento. Para o cenário de análise em questão, serão alocados os dados de corte e variedade no eixo horizontal (linhas) e os anos (tempo) no eixo vertical (colunas), selecionando-se as métricas quantidade, produtividade e rentabilidade para as células centrais da tabela que será montada.

Com isso, quando o usuário abrir o documento Olap, visualizará a imagem da Figura 21.



		Tempo											
		→2002			→2003			→2004					
		Measures			Measures			Measures					
Corte	Cultura	Producao	Rentabilidade	Produtividade	Producao	Rentabilidade	Produtividade	Producao	Rentabilidade	Produtividade			
-All Corte	→Todos_CulturaVariedade	387,926.30	11.80	55.90	85,048.44	8.31	56.32	71,341.10	79.45	61.07	71,765.50	-6.79	62.08
1	→Todos_CulturaVariedade	77,549.26	-28.14	74.85	24,121.55	-16.87	73.18	16,485.12	80.23	72.99	13,112.55	-17.28	90.12
2	→Todos_CulturaVariedade	82,150.22	-21.18	69.03	29,874.71	-20.92	70.10	20,017.80	75.62	71.38	13,594.89	-44.26	63.98
3	→Todos_CulturaVariedade	70,123.50	-28.75	64.27	1,678.94	-61.89	38.96	24,473.76	51.93	63.10	18,299.35	-40.07	68.04
4	→Todos_CulturaVariedade	67,683.84	764.33	50.77	14,721.96	463.31	40.67	1,256.20	270.75	36.99	20,906.64	564.66	58.69
5	→Todos_CulturaVariedade	53,662.81	613.39	43.02	14,146.39	493.19	42.29	5,340.74	317.56	40.40	1,074.42	195.35	31.64
6	→Todos_CulturaVariedade	26,435.64	474.48	36.58	504.89	353.31	34.68	3,324.18	252.15	35.64	3,020.60	280.31	37.86
7	→Todos_CulturaVariedade	8,353.96	323.68		31.73			443.30	180.85	30.45	1,757.05	169.45	29.74
8	→Todos_CulturaVariedade	938.90	723.37		34.54								
9	→Todos_CulturaVariedade	1,028.17	330.56		35.66								

Slicer:

Figura 21 - Visualização inicial do documento Olap

No documento Olap da Figura 21, cuja visão horizontal encontra-se cortada (existem mais anos do que cabem na tela), visualizam-se nas colunas os anos de produção e nas linhas os cortes e as variedades de cana. A visão inicial traz os valores da produção, produtividade e rentabilidade nas células para cada ano e corte (uma vez que as variedades estão agrupadas em “Todos_CulturaVariedade”). Caso o usuário expanda o nó “Todos_CulturaVariedade” de um certo corte, então a aplicação irá mostrar os valores de produção, produtividade e rentabilidade detalhados para cada variedade dentro de um determinado corte. Esta operação de detalhar os dados é denominada de “drill-down” e sua imagem é exibida na Figura 22.

		Tempo											
		-Todos_AnoMes			+2002			+2003			+2004		
Corte	Cultura	Measures			Measures			Measures			Measures		
		Producao	Rentabilidade	Produtividade	Producao	Rentabilidade	Produtividade	Producao	Rentabilidade	Produtividade	Producao	Rentabilidade	Produtividade
-All Corte	+Todos_CulturaVariedade	387,926.30	11.80	55.90	85,048.44	8.31	56.32	71,341.10	79.45	61.07	71,765.50	-6.79	62.00
1	+Todos_CulturaVariedade	77,549.26	-28.14	74.85	24,121.55	-16.87	73.18	16,485.12	80.23	72.99	13,112.55	-17.28	90.00
	-Cana	77,549.26	-28.14	74.85	24,121.55	-16.87	73.18	16,485.12	80.23	72.99	13,112.55	-17.28	90.00
	CB 453	852.59	-9.47	78.80	852.59	-9.47	78.80						
	CP 5122	758.20	42.33	65.76	133.30	-58.68	41.40	624.90	86.54	75.20			
	RB 72454	8,627.88	-44.63	57.15	3,632.31	-42.31	53.84	1,840.13	62.21	66.70			
	RB 739735	6,839.18	13.55	80.47	1,999.71	2.30	87.75	3,766.01	105.68	81.89			
	RB 758540	4,459.38	-50.70	82.05				248.98	47.72	61.63			
	RB 765418	12,181.63	4.19	85.16	11,404.20	.95	86.72	733.30	66.92	68.34			
	RB 835486	4,386.93	-40.22	67.84				647.86	67.12	68.41	3,053.05	-34.15	73.00
	RB 855536	4,206.97	-46.26	85.42							2,650.10	-16.62	90.00
	RB 867515	631.10	-52.16	102.79									
	SP 2313	752.08	54.83	64.12				752.08	54.83	64.12			
	SP 701143	5,589.59	-9.93	56.79	3,224.53	-40.24	55.41	2,365.06	39.59	58.79			
	SP 711406	291.50	-41.47	123.52									
	SP 791011	6,799.48	34.75	91.90	1,456.14	-10.36	78.12	2,312.89	124.93	88.62	3,030.45	-3.35	103.00
	SP 792233	9,005.52	-61.82	66.90				139.09	21.56	52.49	3,084.35	-30.39	77.00
	SP 801842	5,402.57	92.58	94.62	756.68	-19.67	71.05	2,363.69	94.51	77.98	1,294.60	118.95	222.00
	SP 803280	259.10	-57.10	93.20									
	SP 813250	3,830.57	-43.93	77.09	662.09	24.04	104.27	691.13	98.68	79.44			
	SP 832847	2,674.99	-44.48	90.19									
2	+Todos_CulturaVariedade	82,150.22	-21.18	69.03	29,874.71	-20.92	70.10	20,017.80	75.62	71.38	13,594.89	-44.26	63.00
3	+Todos_CulturaVariedade	70,123.50	-28.75	64.27	1,678.94	-61.89	38.96	24,473.76	51.93	63.10	18,299.35	-40.07	68.00
4	+Todos_CulturaVariedade	67,683.84	764.33	50.77	14,721.96	463.31	40.67	1,256.20	270.75	36.99	20,906.64	564.66	58.00
5	+Todos_CulturaVariedade	53,662.81	613.39	43.02	14,146.39	493.19	42.29	5,340.74	317.56	40.40	1,074.42	195.35	31.00

Figura 22 - Operação de “drill-down”, detalhando as variedades de certo corte

A aplicação Olap permite outra perspectiva para os mesmos dados: ao invés de estudar as métricas da produção para cada variedade dentro de um corte, pode-se estudar como se comporta a produção de cada corte para uma dada variedade, de acordo com a Figura 23.

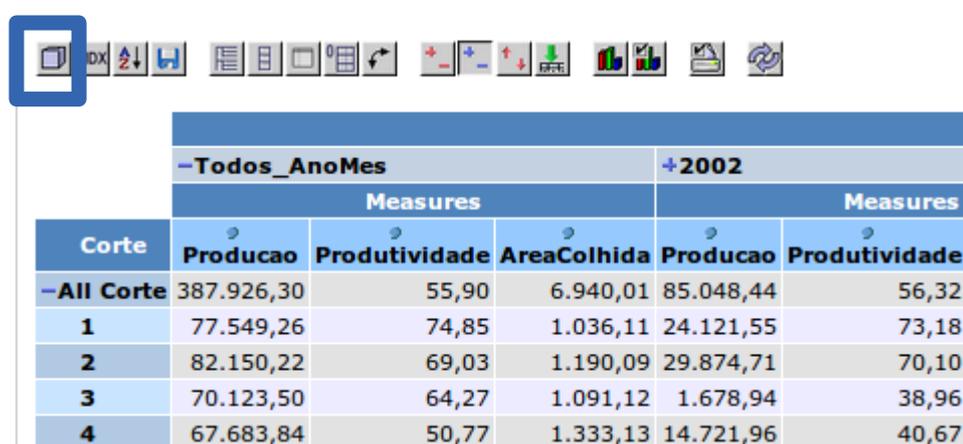
		Tempo															
		-Todos_AnoMes			+2002			+2003			+2004			+2005			
Cultura	Corte	Measures			Measures			Measures			Measures			Measures			
		Producao	AreaColhida	Produtividade	Producao	AreaColhida	Produtividade	Producao	AreaColhida	Produtividade	Producao	AreaColhida	Produtividade	Producao	AreaColhida	Produtividade	
CB 453	-All Corte	5,846.08	99.48	58.77	1,336.61	22.34	59.83	1,302.70	22.34	58.31	1,204.45	22.34	53.91	673.70			
	1	852.59	10.82	78.80	852.59	10.82	78.80										
	2	862.50	10.82	79.71				862.50	10.82	79.71							
	3	764.25	10.82	70.63							764.25	10.82	70.63				
	4	1,157.72	22.34	51.82	484.02	11.52	42.02							673.70			
	5	1,058.85	22.34	47.40				440.20	11.52	38.21							
	6	1,150.17	22.34	51.48							440.20	11.52	38.21				
	7																
	8																
CP 5122	-All Corte	2,416.50	52.56	45.98	133.30	3.22	41.40	733.90	11.53	63.65	663.35	11.53	57.53	109.00			
	1	758.20	11.53	65.76	133.30	3.22	41.40	624.90	8.31	75.20							
	2	663.35	11.53	57.53				109.00	3.22	33.85	554.35	8.31	66.71				
	3	109.00	3.22	33.85							109.00	3.22	33.85				
	4	364.46	11.53	31.61										109.00			
	5	412.44	11.53	35.77													
	6	109.05	3.22	33.87													
	7																
	8																
RB 72454	-All Corte	157,768.42	3,159.76	49.93	51,854.63	961.58	53.93	34,181.18	621.01	55.04	26,199.81	506.06	51.77	24,278.41	45.00		
	1	8,627.88	150.97	57.15	3,632.31	67.46	53.84	1,840.13	27.59	66.70				1,321.60			
	2	32,451.93	492.86	65.84	29,018.59	415.36	69.86	1,862.20	37.71	49.38	1,180.80	27.59	42.80				
	3	30,623.00	517.99	59.12	1,678.94	43.09	38.96	23,709.86	377.00	62.89	1,553.85	26.21	59.28	3,102.35			
	4	29,924.56	584.51	51.20	5,905.87	162.55	36.33	1,256.20	33.96	36.99	20,194.49	345.36	58.47	1,333.05			
	5	36,807.96	829.46	44.38	11,618.92	273.12	42.54	2,641.94	61.91	42.67	1,074.42	33.96	31.64	17,426.60			
6	17,924.18	379.54	34.05			2,870.85	83.84	34.66	871.70	24.30	35.87	174.80					

Figura 23 - Visualização do desempenho dos cortes por variedade

O poder de uma aplicação Olap reside em sua flexibilidade de visualização e exploração dos dados contidos no Cubo Olap. O definição do Cubo Olap para a fazenda Abadia, item 3.8 desta Tese, contempla 10 (dez) dimensões (Tempo, Cultura, Pluviometria, Corte, etc) para análise das métricas (produção, produtividade, etc).

Embora a tela exibida nas imagens anteriores traga dados apenas das dimensões Tempo, Corte e Variedade, esta é somente uma visualização padrão pré-definida para comodidade do usuário. O citado usuário pode a qualquer momento escolher quaisquer outras dimensões, métricas e suas disposições sobre os eixos de visualização, clicando sobre o botão “Olap Navigator”, conforme mostrado na Figura 24.

Olap Navigator



- Todos_AnoMes				+ 2002	
Measures			Measures		
Corte	Producao	Produtividade	AreaColhida	Producao	Produtividade
-All Corte	387.926,30	55,90	6.940,01	85.048,44	56,32
1	77.549,26	74,85	1.036,11	24.121,55	73,18
2	82.150,22	69,03	1.190,09	29.874,71	70,10
3	70.123,50	64,27	1.091,12	1.678,94	38,96
4	67.683,84	50,77	1.333,13	14.721,96	40,67

Figura 24 - Botão “Olap Navigator” que abre a tela de reconfiguração da aplicação

A Figura 25 apresenta a tela do “Olap Navigator”, onde o usuário pode selecionar quais dimensões e métricas exibir sobre quaisquer dos eixos de visualização (linhas ou colunas) e ainda definir se deseja filtrar dados de algumas dimensões.



Figura 25 - Tela do “Olap Navigator”, que permite a seleção das dimensões, métricas e sua disposição sobre os eixos de visualização, além da definição de filtros sobre as dimensões

Assim, embora com um único documento Olap fosse possível fazer qualquer tipo de análise sobre os dados previstos no cubo Olap, por questões de ordem prática com vistas a facilitar o uso, diversos documentos Olap foram criados para este sistema de BI, cada um proporcionando uma visão inicial mais adequada a cada um dos cenários de análise elencados pelo administrador da fazenda Abadia.

3.7.2. Criação de painéis de controle (“dashboards”)

Para atender ao cenário de análise “estudar a produtividade de cada variedade de cana em função do corte”, foi elaborado um painel de controle, chamado pela ferramenta de “Dashboard”. Um painel de controle é um tipo de documento do SpagoBI que permite a integração de dois ou mais documentos em uma mesma tela, permitindo inclusive navegação cruzada, ou seja, que um dos documentos forneça parâmetros para a exibição dos demais.

Gráfico 1 – Comparativo da produtividade das variedades

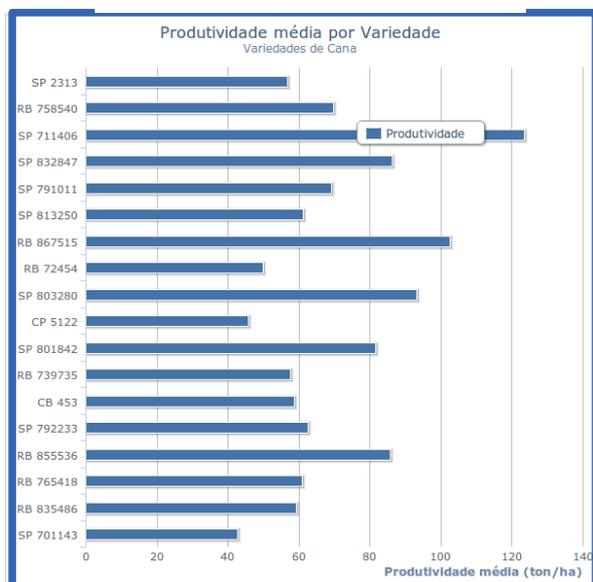


Gráfico 2 – Produtividade por Corte para a variedade selecionada no Gráfico 1

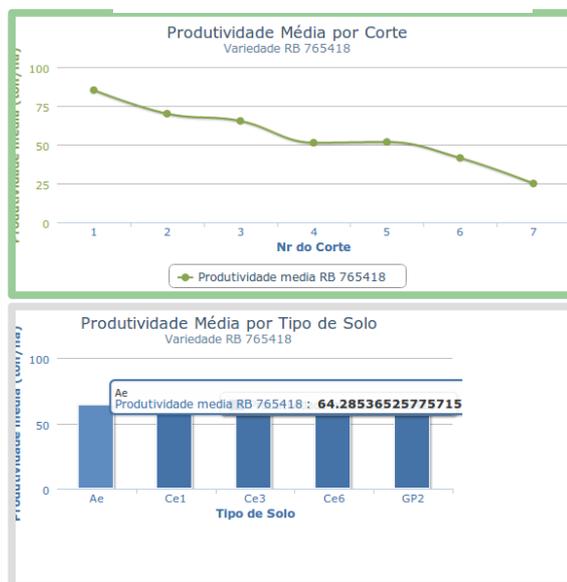


Gráfico 3 – Produtividade por Tipo de Solo para a variedade selecionada no Gráfico 1



Figura 26 - Painel de Controle (dashboard) com três gráficos

A Figura 26 apresenta um Painel de Controle com três gráficos: o primeiro deles traz o comparativo da média das produtividades das variedades e, quando o usuário clica sobre uma das variedades do gráfico 1, o gráfico 2 detalha a produtividade média para os cortes desta variedade, enquanto o gráfico 3 exibe, também para a variedade selecionada, a produtividade média em cada tipo de solo em que foi cultivada.

Para a construção deste documento no SpagoBI (Figura 27), é necessário que se construa individualmente cada um dos três gráficos e depois, utilizando um tipo de documento especial, denominado “Composite Document” defina-se as dimensões, o posicionamento e a navegação entre os documentos originais dentro da nova página.

Na elaboração do Gráfico 1, componente do Painel de Controle, usou-se o “engine” JFreeChart, sobre um conjunto de dados (“Dataset”) que calculou a média da produtividade agrupada por variedade, e um modelo em xml (Figura 28).

Label	chartVariedade *
Name	chartVariedade *
Description	Grafico Barras para estudo produtividade
Type	Real-time - DashBoard ▼
Engine	JFreeChart Engine ▼
Dataset	ds_ProdutVariedade 🔍
State	Released ▼
Refresh seconds	0
Criptable	<input type="radio"/> True <input checked="" type="radio"/> False
Visible	<input checked="" type="radio"/> True <input type="radio"/> False
Visibility restrictions	<div style="border: 1px solid #ccc; height: 20px; width: 100%;"></div>
Template	Escolher arquivo Nenhum arquivo selecionado

Figura 27 - Criação do Gráfico 1 para o Painel de Controle no SpagoBI

```

chartProdutVariedade.sbihchart x
1 <HIGHCHART>
2 <CHART type='bar' />
3 <TITLE text='Produtividade média por Variedade' />
4 <SUBTITLE text='Variedades de ${cultura}' />
5 <X_AXIS title='Variedades' alias='nomevariedade' />
6 <Y_AXIS min='0'>
7   <TITLE text='Produtividade' align='high' />
8 </Y_AXIS>
9 <LEGEND layout='vertical'
10   align='right'|
11   verticalAlign='top'
12   x='-100'
13   y='100'
14   floating='true'
15   borderWidth='1'
16   backgroundColor='#FFFFFF'
17   shadow='true' />
18 <PLOT_OPTIONS>
19 <BAR>
20   <DATA_LABELS enable='true' />
21 </BAR>
22 </PLOT_OPTIONS>
23 <SERIES_LIST>
24 <SERIES name="Produtividade" alias="produtividade" />
25 </SERIES_LIST>
26 <DRILL document="chartProdtvdVarCorte">
27 <PARAM_LIST>
28 <PARAM name='cultura' type="RELATIVE" />
29 <PARAM name='variedade' type="CATEGORY" />
30 </PARAM_LIST>
31 </DRILL>
32 </HIGHCHART>

```

Figura 28 - Modelo xml para o Gráfico 1

Os Gráficos 2 e 3 foram elaborados, no SpagoBI, de forma similar ao Gráfico 1, com a diferença de utilizarem um parâmetro adicional, denominado pela suíte como “Analytical Driver”.

Para que o Gráfico 2 exiba os dados de produtividade por Corte, é necessário que receba como parâmetro uma determinada variedade, para a qual os dados serão filtrados.

Label	<input type="text" value="chartProdtvdVarCorte"/>	*
Name	<input type="text" value="chartProdtvdVarCorte"/>	*
Description	<input type="text" value="Gráfico de produtividade média de uma"/>	
Type	<input type="text" value="Real-time - DashBoard"/>	▼
Engine	<input type="text" value="JFreeChart Engine"/>	▼
Dataset	<input type="text" value="ds_ProdtVarCorte"/>	
State	<input type="text" value="Released"/>	▼
Refresh seconds	<input type="text" value="0"/>	
Criptable	<input type="radio"/> True <input checked="" type="radio"/> False	
Visible	<input checked="" type="radio"/> True <input type="radio"/> False	
Visibility restrictions	<input type="text"/>	
	<input type="text" value="▼"/> = <input type="text"/>	
Template	<input type="button" value="Escolher arquivo"/> Nenhum arquivo selecionado	

<input type="button" value="cultura"/>	<input type="button" value="variedade"/>	<input type="button" value="New..."/>
DOCUMENT ANALYTICAL DRIVER DETAILS		
Title	<input type="text" value="cultura"/>	*
Analytical driver	<input type="text" value="ad_cultura"/>	*
Url Name	<input type="text" value="cultura"/>	*
Priority	<input type="text" value="1"/>	▼
Visible	<input type="checkbox"/>	
Required	<input type="radio"/> True <input checked="" type="radio"/> False	
Multivalue	<input type="radio"/> True <input checked="" type="radio"/> False	

Figura 29 - Parte inferior da tela de criação do Gráfico 2, exibindo os “Analytical Drivers” - parâmetros para a execução do documento no SpagoBI

Em verdade, para todos os documentos desenvolvidos para este sistema de BI, foi adicionado um “Analytical Driver”, ou seja, um parâmetro, para que todas as análises pudessem ser executadas por cultura. Como a fazenda Abadia se ocupa exclusivamente da cultura da cana-de-açúcar, a criação de tal parâmetro para todos os relatórios, embora inócua no momento, foi uma providência para adicionar escalabilidade ao sistema, além de permitir que o mesmo possa ser utilizado por outras organizações rurais

Label	<input type="text" value="docCompVariedade"/>	*
Name	<input type="text" value="docCompVariedade"/>	*
Description	<input type="text" value="DashBoard para estudo da produtividade"/>	
Type	<input type="text" value="Cockpit"/>	▼
Engine	<input type="text" value="Document Composition Engine"/>	▼
State	<input type="text" value="Released"/>	▼
Refresh seconds	<input type="text" value="0"/>	
Criptable	<input type="radio"/> True <input checked="" type="radio"/> False	
Visible	<input checked="" type="radio"/> True <input type="radio"/> False	
Visibility restrictions	<input type="text"/>	
	<input type="text"/> = <input type="text"/>	 
Template	<input type="button" value="Escolher arquivo"/> Nenhum arquivo selecionado	

cultura	variedade	New...
DOCUMENT ANALYTICAL DRIVER DETAILS		
Title	<input type="text" value="cultura"/>	*
Analytical driver	<input type="text" value="ad_cultura"/>	* 
Uri Name	<input type="text" value="cultura"/>	*
Priority	<input type="text" value="1"/>	▼
Visible	<input checked="" type="checkbox"/>	
Required	<input type="radio"/> True <input checked="" type="radio"/> False	
Multivalue	<input type="radio"/> True <input checked="" type="radio"/> False	

Figura 30 - Criação do documento para o Painel de Controle no SpagoBI

Uma vez criados os três documentos (Gráfico 1, 2 e 3), cria-se um novo documento para o Painel de Controle, de forma a integrá-los em um único, conforme ilustra a figura anterior.

O modelo xml a seguir define as configurações para os três documentos integrados no Painel de Controle.

```

1 <DOCUMENTS_COMPOSITION template_value="/jsp/engines/documentcomposition/template/dynamicTemplate.jsp">
2 <DOCUMENTS_CONFIGURATION video_width="1200" video_height="600">
3 <DOCUMENT sbi_obj_label="chartVariedade" local_file_name="chartProdutVariedade.sbihchart">
4 <PARAMETERS>
5 <PARAMETER type="IN" sbi_par_label="cultura" default_value="Cana"/>
6 <REFRESH/>
7 <PARAMETER type="OUT" sbi_par_label="cultura" default_value="Cana">
8 <REFRESH>
9 <REFRESH_DOC_LINKED labelDoc="chartProdtvdVarCorte" labelParam="cultura" />
10 <REFRESH_DOC_LINKED labelDoc="chartProdtvdVarSolo" labelParam="cultura" />
11 </REFRESH>
12 </PARAMETER>
13 <PARAMETER type="OUT" sbi_par_label="variedade" default_value="AA">
14 <REFRESH>
15 <REFRESH_DOC_LINKED labelDoc="chartProdtvdVarCorte" labelParam="variedade" />
16 <REFRESH_DOC_LINKED labelDoc="chartProdtvdVarSolo" labelParam="variedade" />
17 </REFRESH>
18 </PARAMETER>
19 </PARAMETERS>
20 <STYLE style="position:absolute;margin:0px;left:0px;top:0px;width:600px;height:600px;" mode="auto"/>
21 <id>chartVariedade</id>
22 </DOCUMENT>
23 <DOCUMENT sbi_obj_label="chartProdtvdVarCorte" local_file_name="chartProdutivVariedxCorte.sbihchart">
24 <PARAMETERS>
25 <PARAMETER type="IN" sbi_par_label="cultura" default_value="Cana"/>
26 <REFRESH/>
27 <PARAMETER type="IN" sbi_par_label="variedade" default_value="AA"/>
28 <REFRESH/>
29 </PARAMETERS>
30 <STYLE style="position:absolute;margin:0px;left:600px;top:0px;width:600px;height:300px;" mode="auto"/>
31 <id>chartProdtvdVarCorte</id>
32 </DOCUMENT>
33 <DOCUMENT sbi_obj_label="chartProdtvdVarSolo" local_file_name="chartProdutivVariedxSolo.sbihchart">
34 <PARAMETERS>
35 <PARAMETER type="IN" sbi_par_label="cultura" default_value="Cana"/>
36 <REFRESH/>
37 <PARAMETER type="IN" sbi_par_label="variedade" default_value="AA"/>
38 <REFRESH/>
39 </PARAMETERS>
40 <STYLE style="position:absolute;margin:0px;left:600px;top:300px;width:600px;height:300px;" mode="auto"/>
41 <id>chartProdtvdVarSolo</id>
42 </DOCUMENT>
43 </DOCUMENTS_CONFIGURATION>
44 </DOCUMENTS_COMPOSITION>

```

Figura 31 - Modelo xml para o documento do Painel de Controle

3.7.3. Criação de documentos espaciais – Location Intelligence

O cenário de análise “estudo das métricas de produção em determinado ano pelas áreas da propriedade, com visualização em um mapa” ilustra a importância da dimensão espacial na tomada de decisão de um gestor rural.

A fazenda Abadia utiliza diversos princípios da agricultura de precisão, controlando os fatores de produção em cada unidade de divisão da propriedade (lote). Desta forma, permitir ao gestor da fazenda a visualização do desempenho da produção

em cada lote ou setor da propriedade, conduz a visões mais completas do negócio, com informações de pertinência e circunvizinhança de localização geográfica.

A figura 32 exibe o estudo da produção (quantidade produzida em toneladas) de cana-de-açúcar em cada lote da fazenda – são mostrados apenas os lotes que apresentaram produção – em um determinado ano.

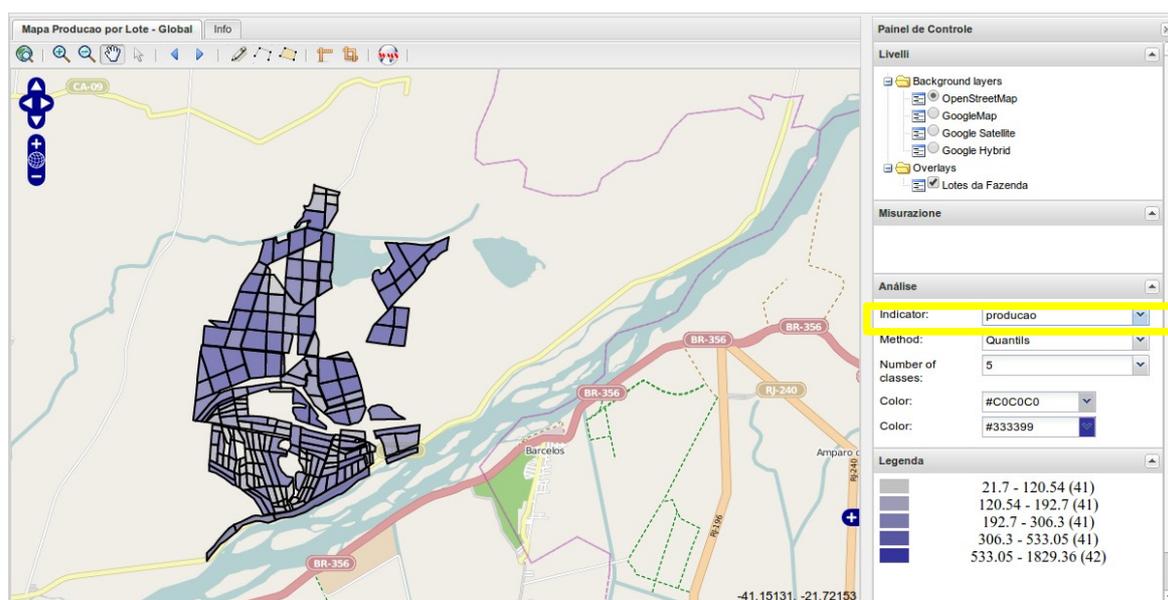


Figura 32 - Documento de análise espacial das métricas da produção.

Alterando o indicador (“indicador” - marcado na figura anterior), que é a métrica em estudo (ex.: produção, produtividade, receita, lucro, etc), o mapa se redesenha junto com a legenda exibindo quais lotes estão classificados nas faixas predefinidas.

O documento permite ao usuário escolher o contexto geográfico sobre o qual projetar os dados espaciais, na configuração original pode-se selecionar OpenStreetMap, GoogleMap, Google Satellite ou Google Hybrid.

A ferramenta permite ainda que se meça distâncias e áreas, além de possibilitar desenhar pontos, linhas e polígonos sobre o mapa

Quando o usuário clica sobre uma das áreas do mapa, um documento estilo “pop-up” é exibido com dados acerca daquela porção do mapa. A partir de tal documento “pop-up” é possível também fazer uma navegação cruzada (“cross-navigation”) para qualquer outro documento da suíte.

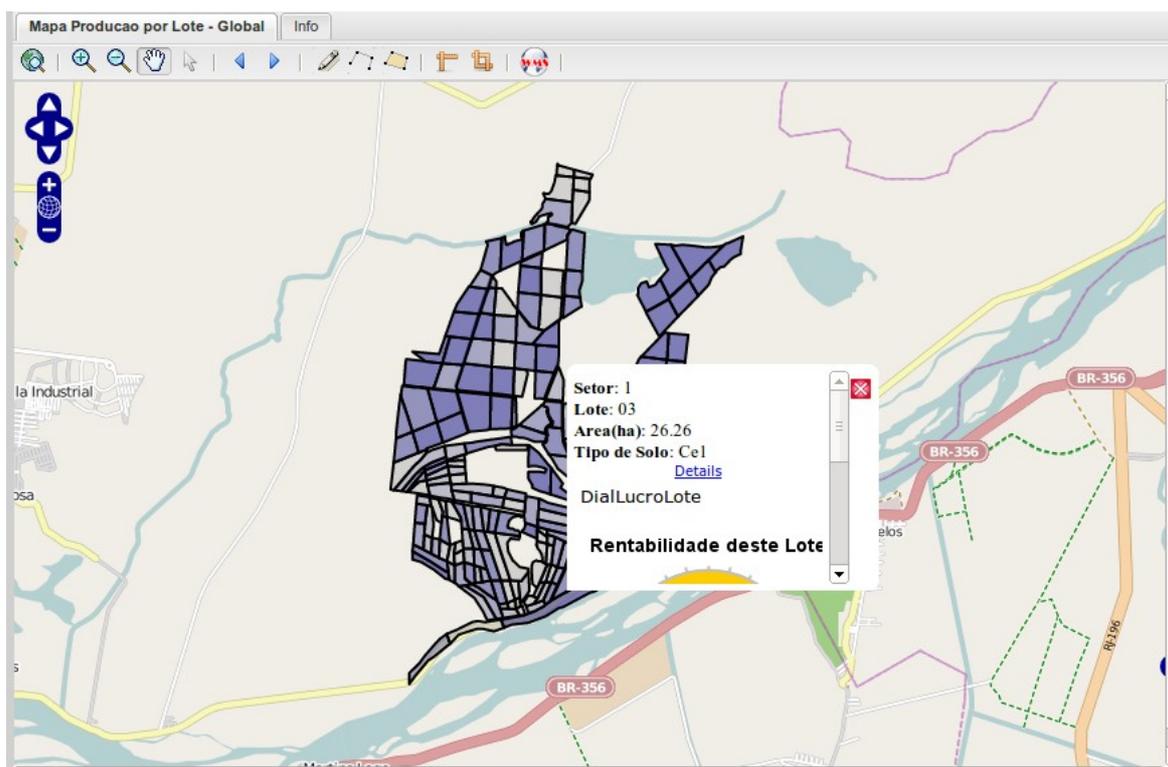


Figura 33 - Detalhe da exibição de documento “pop-up” quando usuário clica sobre um dos lotes

No documento do exemplo, clicando sobre o link “Details” do “pop-up”, o usuário é direcionado para um documento do tipo Olap que permite ao usuário explorar mais dados acerca do setor no qual o lote em questão se situa.



Figura 34 - Detalhe do documento Olap - navegação cruzada a partir do “pop-up” do mapa

O SpagoBI possui dois “engines” (frameworks para geração de documentos) que lidam com mapas: GeoEngine, que utiliza um mapa estático, ou seja, o mapa é uma figura sem referenciamento espacial, e GisEngine, onde o mapa é exibido por um Sistema de

Informações Geográficas (GeoServer), que se utiliza das especificações WFS/WMS (OGC, 2014) além de permitir que documentos “pop-up” contextuais sejam exibidos.

De forma a prover maior riqueza de informações ao decisor, todos os documentos necessários aos cenários de análise com componente espacial foram implementados usando o GisEngine.

Para que o GisEngine possa ser utilizado, o GeoServer deve estar instalado, conforme descrito no item 3.10 desta Tese. Em seguida, deve-se publicar as camadas espaciais (chamadas de layers) que estão armazenadas nas tabelas “setor” e “lote” pertencentes ao data warehouse espacial (dwfazenda).

Publicar uma camada espacial (layer) significa expor este conteúdo para que possa ser acessado externamente via GeoServer, bastando para isto digitar, em um browser, o endereço `http://<endereço_IP_Servidor>:8080/geoserver`. Após a publicação, pode-se visualizar os elementos de uma camada (Ex.: lote) ou filtrá-los segundo algum de seus atributos (Ex.: exibir os lotes cujo tipo de solo seja igual a Ce1).

A publicação das camadas “setor” e “lote” seguiu o prescrito em (GeoServer, 2014b).



Figura 35 - Camadas espaciais (“Layers”): setor e lote, configuradas no GeoServer

As Figuras 36 e 37 mostram a pré-visualização das camadas “setor” e “lote” publicadas no GeoServer.

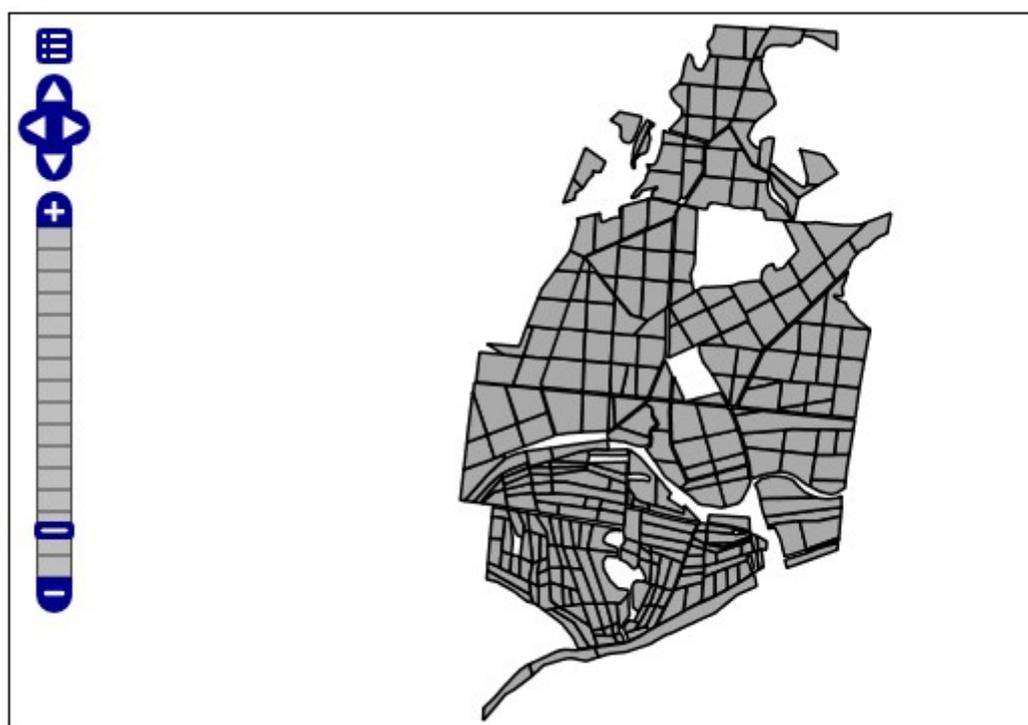


Scale = 1 : 126K

-41.24223, -21.72297

Click on the map to get feature info

Figura 36 - Visualização da camada "setor" da fazenda Abadia no GeoServer



Scale = 1 : 83K

-41.20499, -21.72117

Figura 37 - Visualização da camada "lote" da fazenda Abadia no GeoServer

Depois de configuradas as camadas espaciais no GeoServer, cria-se no SpagoBI o documento do tipo “Location Intelligence” e “GisEngine”, conforme Figura 38.

Label	MapProdxLote *
Name	MapProdxLote *
Description	Mapa com a producao por lote em dete
Type	Location Intelligence ▼
Engine	Gis Engine ▼
Data Source	dsdwfazenda ▼
Dataset	ds_ProdxLote 🔍
State	Released ▼
Refresh seconds	0
Criptable	<input type="radio"/> True <input checked="" type="radio"/> False
Visible	<input checked="" type="radio"/> True <input type="radio"/> False
Visibility restrictions	<div style="border: 1px solid gray; height: 40px; width: 100%;"></div> ▼ = <input type="text"/> ➕ 📄
Template	<input type="button" value="Escolher arquivo"/> Nenhum arquivo selecionado

cultura	ano	New...
DOCUMENT ANALYTICAL DRIVER DETAILS		
Title	ano *	
Analytical driver	ad_ano * 🔍	
Url Name	ano *	
Priority	2 ▼	
Visible	<input checked="" type="checkbox"/>	
Required	<input type="radio"/> True <input checked="" type="radio"/> False	

Figura 38 - Criação do documento tipo “Location Intelligence” no SpagoBI.

O documento recebe um conjunto de dados (“dataset”, no exemplo com o nome “ds_ProdxLote”) com as métricas da produção de certa cultura agrupadas por lote da

fazenda em determinado ano, e os parâmetros (chamados “Analytical Drivers”) “cultura” e “ano”.

Diferente dos demais documentos do SpagoBI, aqueles pertencentes ao “GisEngine”, utilizam um modelo de definição baseado em JSON (JavaScript Object Notation) e não em xml.

```

1 {
2     mapName: "Mapa Producao por Lote - Global",
3     analysisType: "choropleth",
4     analysisConf: {
5         type: "choropleth"
6         , indicador: "lucro"
7         , method: "CLASSIFY_BY_QUANTILS" // "CLASSIFY_BY_EQUAL_INTERVALS"
8         , classes: 5
9         , fromColor: "#C0C0C0"
10        , toColor: "#333399",
11        feautreInfo: [{"Setor","nrsetor"}, {"Lote","nr lote"}, {"Area(ha)","area"}, {"Tipo de Solo", "tiposolo"}],
12        indicators: [{"Lucro", "lucro"}, {"Producao", "producao"}, {"Produtividade", "produtividade"}, {"Receita", "receita"}, {"Custo", "custo"}],
13        businessID: "idlote",
14        geoId: "idlote",
15        targetLayerConf: {
16        text: 'Lotes da Fazenda'
17        , name: 'lote'
18        , url: 'http://localhost:8080/geoserver/wfs',
19        inlineDocumentConf: {
20        label: 'DialLucroLote'
21        , staticParams: {
22        param1: 'foo'
23        }
24        , dynamicParams: {
25        ano: 'ANO',
26        cultura: 'CULTURA',
27        setor: 'NRSETOR',
28        lote: 'IDLOTE'}
29        , displayToolbar: 'false'
30        , displaySliders: 'false'
31        }
32        , detailDocumentConf: {
33        label: 'OlapProdLoteSetorAno'
34        , staticParams: {
35        departmentId: '3'
36        }
37        , dynamicParams: {
38        ano: 'ANO',
39        cultura: 'CULTURA',
40        setor: 'NRSETOR'
41        }
42        , displayToolbar: 'false'
43        , displaySliders: 'false'
44        }
45        , role: "spagobi/admin",
46        lon: -41.20350,
47        lat: -21.71132,
48        zoomLevel: 13
49    }
50 }

```

Figura 39 - Modelo JSON de definição para documento – GisEngine.

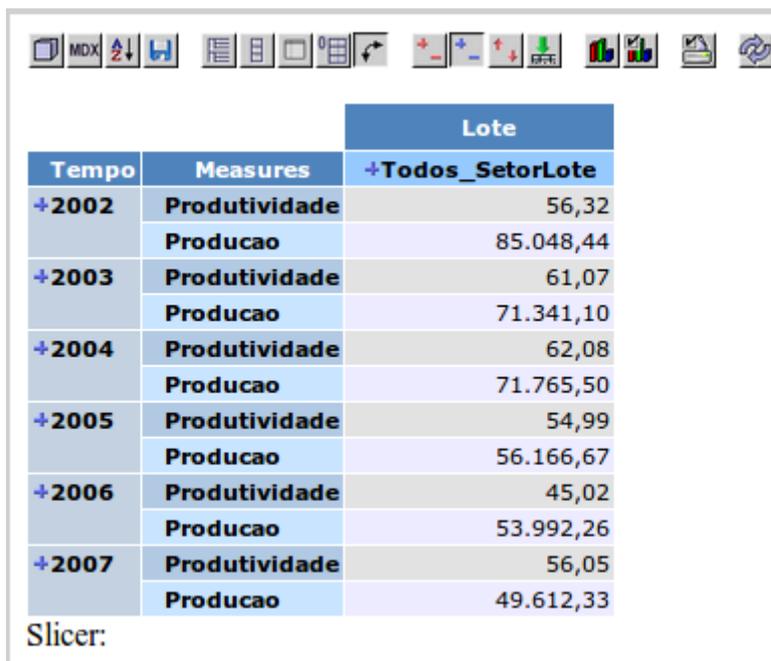
O Apêndice B traz o detalhamento do modelo JSON para o documento “GisEngine”.

4. RESULTADOS E DISCUSSÃO

Neste capítulo, diversos cenários de análise seguidas das telas de apoio do sistema de BI serão apresentados de forma a validar a eficiência do uso da suíte para o auxílio à tomada de decisão.

4.1. Cenário 1: exibir a evolução das métricas: produção e produtividade ao longo dos anos para a fazenda

Para a exibição dos dados de produção e produtividade ao longo do tempo, será usado o documento Olap denominado “Evolução da produção por setor/lote”.



		Lote
Tempo	Measures	+Todos_SetorLote
+2002	Produtividade	56,32
	Producao	85.048,44
+2003	Produtividade	61,07
	Producao	71.341,10
+2004	Produtividade	62,08
	Producao	71.765,50
+2005	Produtividade	54,99
	Producao	56.166,67
+2006	Produtividade	45,02
	Producao	53.992,26
+2007	Produtividade	56,05
	Producao	49.612,33

Slicer:

Figura 40 - Evolução temporal da produção e produtividade.

Os dados mostram que a produção caiu ao longo dos anos e a produtividade, cresceu entre 2002 e 2004, apresentou queda nos anos de 2005 e 2006 e estabilizou na média em 2007.

A reação natural de um gerente ao fazer a leitura de tais dados seria buscar os porquês para o comportamento das métricas. Com relação à queda da produção, uma possível explicação estaria na área colhida. Para investigar se a área colhida esteve relacionada à queda na produção, basta adicionar tal atributo às métricas do documento Olap a serem exibidas.

The screenshot shows a software interface for configuring an OLAP report. At the top, there is a toolbar with icons for MDX, undo, redo, and other functions. Below the toolbar is a dialog box titled 'Measures' with a close button (X). The dialog box contains a list of measures with checkboxes:

- Produtividade
- Producao
- Rentabilidade
- Lucratividade
- Custo_ha
- Lucro_ha
- Receita_ha
- AreaColhida
- Lucro
- Custo
- Receita

At the bottom of the dialog box are four buttons: 'None', 'Group', 'OK', and 'Cancel'. Below the dialog box is a data table with the following structure:

		Lote
Tempo	Measures	+Todos_SetorLote
+2002	Produtividade	56,32
	Producao	85.048,44
+2003	Produtividade	61,07
	Producao	71.341,10
+2004	Produtividade	62,08
	Producao	71.765,50
+2005	Produtividade	54,99
	Producao	56.166,67
+2006	Produtividade	45,02
	Producao	53.992,26
+2007	Produtividade	56,05
	Producao	49.612,33

Figura 41 - Adicionando o atributo “AreaColhida” à exibição do documento Olap.

A figura 42 adiciona a nova métrica “AreaColhida” às demais exibidas pelo documento Olap.



		Lote
Tempo	Measures	+Todos_SetorLote
+2002	Produtividade	56,32
	Producao	85.048,44
	AreaColhida	1.510,01
+2003	Produtividade	61,07
	Producao	71.341,10
	AreaColhida	1.168,13
+2004	Produtividade	62,08
	Producao	71.765,50
	AreaColhida	1.155,93
+2005	Produtividade	54,99
	Producao	56.166,67
	AreaColhida	1.021,41
+2006	Produtividade	45,02
	Producao	53.992,26
	AreaColhida	1.199,40
+2007	Produtividade	56,05
	Producao	49.612,33
	AreaColhida	885,13

Slicer:

Figura 42 - Documento Olap reconfigurado, exibindo “AreaColhida” além de produção e produtividade.

O documento Olap, agora adicionado do atributo “AreaColhida”, mostra que efetivamente a área colhida diminuiu ao longo dos anos, sendo um dos fatores que levou à queda da produção.

Os próximos movimentos do gerente da fazenda tentarão explicar a variação da produtividade. O primeiro fator a investigar é a pluviometria. Para tanto, adicionou-se a dimensão “pluviometria” ao eixo vertical de forma que se possa ler quanto se produziu, em cada ano, sob cada regime pluviométrico (faixa pluviométrica). O resultado vem na figura a seguir.

		Lote			
		+Todos_SetorLote			
		Pluviometria			
Tempo	Measures	-All Pluviometria	801 a 900mm	901 a 1000mm	maior que 1200mm
+2002	Produtividade	56,32		51,62	73,18
	Producao	85.048,44		60.926,89	24.121,55
	AreaColhida	1.510,01		1.180,38	329,63
+2003	Produtividade	61,07	59,16	58,05	72,99
	Producao	71.341,10	8.408,32	46.447,66	16.485,12
	AreaColhida	1.168,13	142,12	800,16	225,85
+2004	Produtividade	62,08			62,08
	Producao	71.765,50			71.765,50
	AreaColhida	1.155,93			1.155,93
+2005	Produtividade	54,99			54,99
	Producao	56.166,67			56.166,67
	AreaColhida	1.021,41			1.021,41
+2006	Produtividade	45,02			45,02
	Producao	53.992,26			53.992,26
	AreaColhida	1.199,40			1.199,40
+2007	Produtividade	56,05			56,05
	Producao	49.612,33			49.612,33
	AreaColhida	885,13			885,13

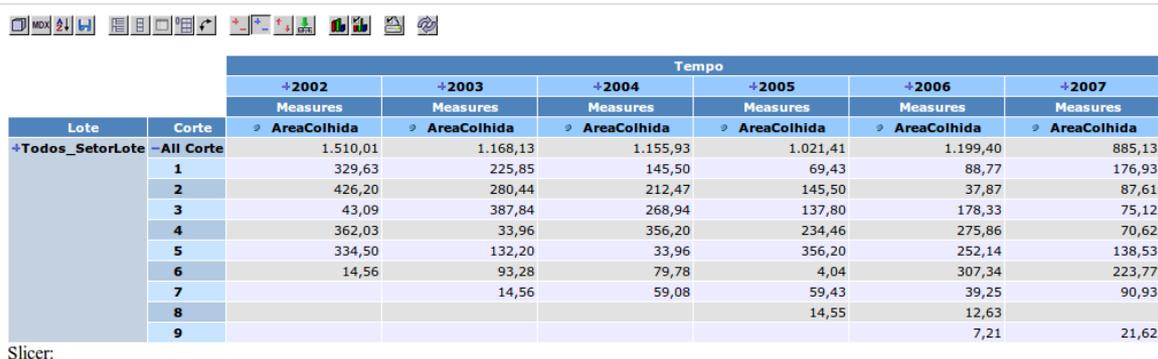
Slicer:

Figura 43 - Documento Olap adicionado da dimensão Pluviometria.

Os dados mostram que houve produção sob faixa pluviométrica menor que 1200 mm apenas nos anos de 2002 e 2003. Nos anos em que se observou a queda de produtividade, ou seja, em 2005 e principalmente, 2006, a precipitação pluviométrica esteve acima de 1200 mm, qual seja, não foi sua causa determinante.

Outra linha de investigação para se encontrar as explicações para a queda de produtividade nos anos de 2005 e 2006 seria verificar a contribuição dos cortes em cada ano. Sabe-se que a produtividade cai à medida que cresce o número do corte de certa variedade, assim, se em dada época a propriedade estivesse ocupada majoritariamente por cortes mais antigos (de menor produtividade), a produtividade total estaria comprometida.

Para investigar tal linha de raciocínio, retirou-se a dimensão pluviometria e adicionou-se a dimensão corte. Para reduzir a quantidade de dados exibidos, manteve-se nas métricas apenas a “AreaColhida”, e para que a visualização não ficasse muito extensa no eixo horizontal, usou-se a operação de pivoteamento, que troca o eixo “x” pelo “y”. O resultado está na figura 44.



Lote	Corte	Tempo					
		+2002	+2003	+2004	+2005	+2006	+2007
		Measures	Measures	Measures	Measures	Measures	Measures
		AreaColhida	AreaColhida	AreaColhida	AreaColhida	AreaColhida	AreaColhida
+Todos_SetorLote	All Corte	1.510,01	1.168,13	1.155,93	1.021,41	1.199,40	885,13
	1	329,63	225,85	145,50	69,43	88,77	176,93
	2	426,20	280,44	212,47	145,50	37,87	87,61
	3	43,09	387,84	268,94	137,80	178,33	75,12
	4	362,03	33,96	356,20	234,46	275,86	70,62
	5	334,50	132,20	33,96	356,20	252,14	138,53
	6	14,56	93,28	79,78	4,04	307,34	223,77
	7		14,56	59,08	59,43	39,25	90,93
	8				14,55	12,63	
	9					7,21	21,62

Slicer:

Figura 44 - Documento Olap com área colhida por corte ao longo dos anos.

O documento Olap mostra que nos anos de 2005 e 2006, houve um envelhecimento do canavial, ou seja, havia mais áreas ocupadas com cortes mais antigos do que nos outros anos, constituindo um dos fatores para a queda da produtividade global.

4.2. Cenário 2: Distribuição das variedades pelas áreas da propriedade, com visualização em um mapa, de forma a visualizar o mix de variedades e suas concentrações

Neste cenário, o administrador da propriedade deseja, para um determinado ano, verificar quanto de área está ocupado por cada variedade, visualizando em um mapa a localização destas áreas.

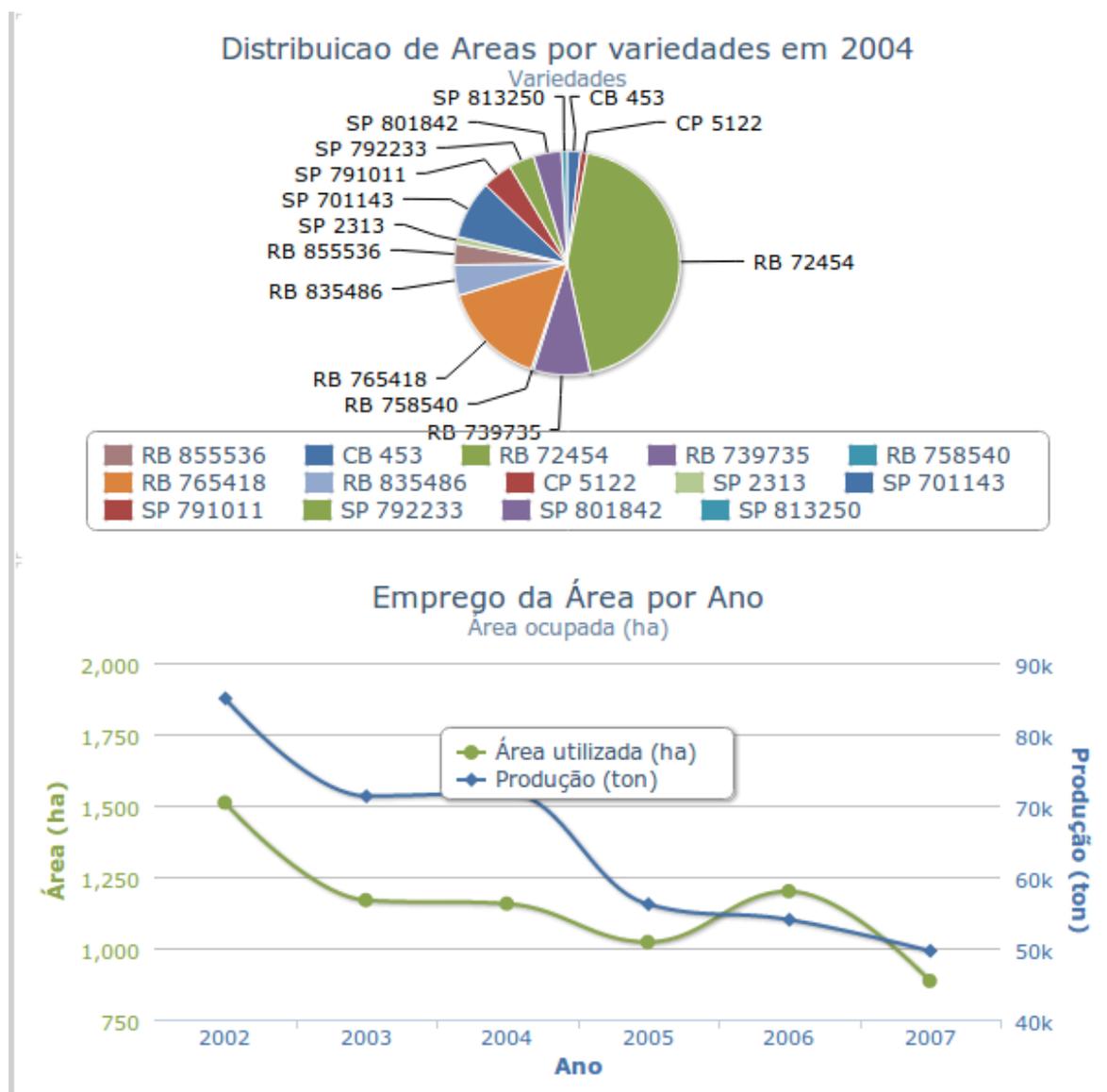


Figura 45 - Tela inicial do documento de ocupação das áreas por variedade para o ano de 2004.

O documento é um painel de controle (“dashboard”) composto por três outros documentos. A figura anterior mostra os dois documentos exibidos depois que o usuário seleciona o ano para o qual filtrar os dados.

O documento da parte superior “Distribuição de Áreas por variedade em 2004” mostra que a variedade que ocupava a maior área da propriedade era a RB 72-454, com 506,06 hectares (quando o usuário passa o mouse sobre uma variedade, surge um pequeno painel com o nome da variedade e a área ocupada em hectares), seguida pela RB 76-5418 com 175,29 hectares.

O documento da parte inferior “Emprego da Área por Ano” visa dar noção ao administrador da evolução da ocupação da área da propriedade, mostrando também a

evolução da produção da fazenda no mesmo período. Assim, em 2004 foram ocupados 1155,93 hectares (ha) e produzidos 71765,5 toneladas de cana-de-açúcar, enquanto por exemplo em 2003 ocupou-se 1168,13 ha para uma produção de 71341,10 toneladas. Ou seja, entre 2003 e 2004 a produtividade global cresceu já que se utilizou menos área para produzir quantidade maior de cana-de-açúcar.

Quando o usuário clica sobre qualquer das variedades expostas no primeiro documento “Distribuição de Áreas por variedades em 2004”, surge o terceiro documento, um mapa, que permite a análise de diversas métricas (produção, produtividade, receita, custo, lucro, etc) pelas áreas ocupadas pela variedade selecionada no ano escolhido.

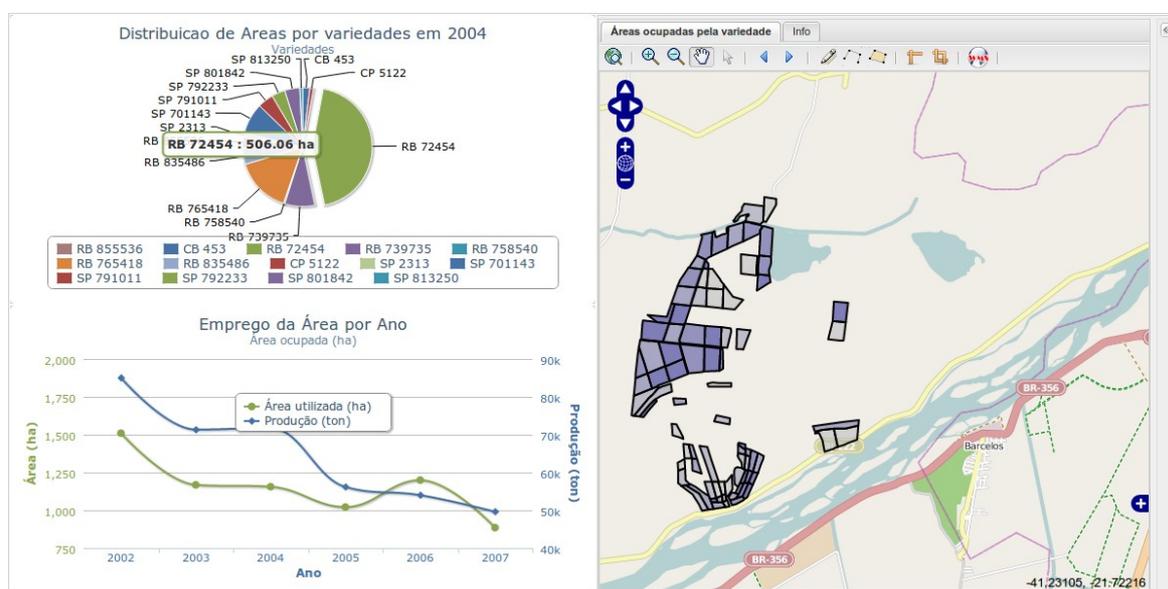


Figura 46 - Surgimento do terceiro documento no painel de controle – mapa – exibindo as áreas ocupadas pela variedade RB 72-454, no ano de 2004.

Clicando no botão de setas duplas no canto superior direito da tela, o usuário tem acesso à legenda utilizada para classificar as áreas do mapa, além de poder escolher qual outra métrica empregar para analisar as áreas da variedade escolhida. A figura a seguir aproxima a visão do mapa com suas legendas.

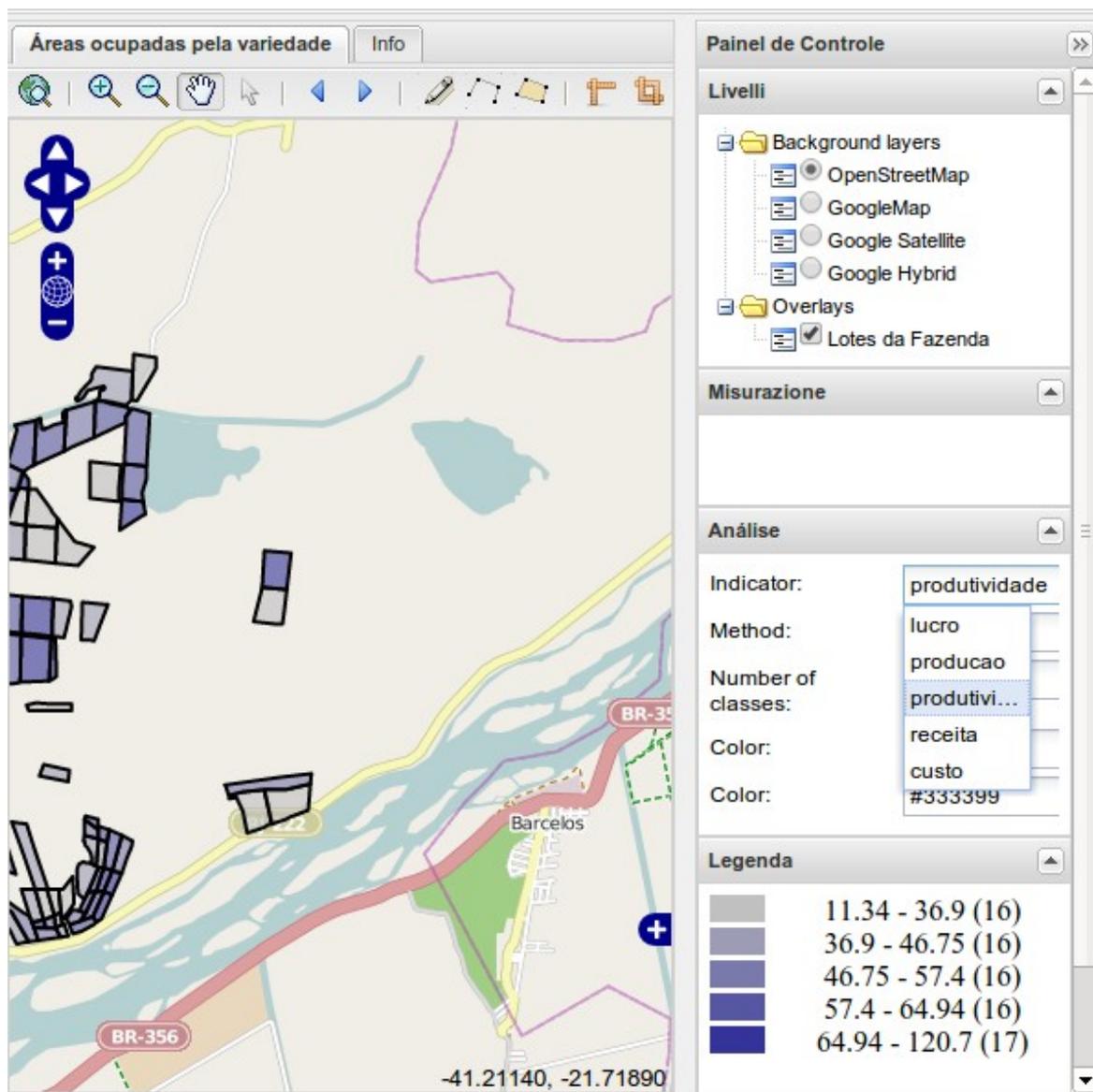


Figura 47 - Detalhe do mapa – exibição das legendas utilizadas para a classificação das áreas e possibilidade de escolha de outra métrica para análise.

Este cenário de análise vem ao encontro de uma das recomendações feitas por Bernardes (2002) quanto ao controle do censo varietal, ou seja, quanto à composição de variedades cultivadas na fazenda. No documento de 2002, Bernardes recomenda a ampliação da quantidade de variedades cultivadas, de forma a reduzir a fragilidade do canavial a ataque de pragas e doenças e ocorrência de perdas drásticas de produção. De forma a comparar a composição varietal da fazenda Abadia nos anos de 2002 e 2004, para verificar se as recomendações foram cumpridas, seguem-se as análises nos documentos.

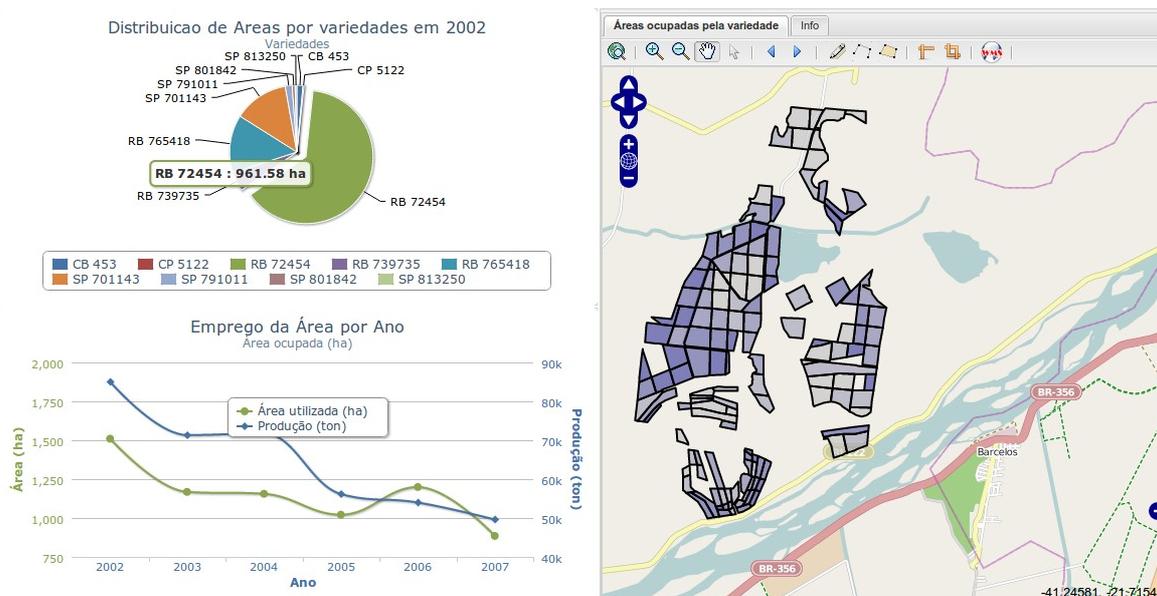


Figura 48 - Documento com a composição varietal em 2002 e, no mapa, a exibição das áreas ocupadas com a variedade RB 72-454.

O documento da Figura 48 apresenta a composição varietal para o ano de 2002, mostrando que a propriedade colheu 9 (nove) variedades distintas, mas com grande predominância da variedade RB 72-454, que ocupava 961,58 ha dos 1510,01 ha totais, ou seja, 63% de ocupação.

A Figura 47 traz a composição de variedades para 2004. Percebe-se que as recomendações feitas por Bernardes (2002) foram atendidas. No ano de 2004 a fazenda colheu 14 (quatorze) diferentes variedades, e a grande dependência da variedade RB 72-454, que ocupava anteriormente 63% da área da propriedade, foi reduzida já que a mesma passou a ocupar 43% da área colhida no citado ano.

4.3. Cenário 3: Análise das métricas de produção nas áreas da fazenda ao longo do tempo

O objetivo deste cenário é investigar o desempenho dos setores e lotes da fazenda, de forma global (agrupado para todos os anos) ou detalhado ano a ano. Uma possibilidade de emprego deste cenário seria para um estudo de substituição parcial da atividade da fazenda, como por exemplo a introdução de criação de gado em uma parte da propriedade.

Lote	Tempo						
	-Todos_AnoMes	+2002	+2003	+2004	+2005	+2006	+2007
	Measures						
	↳ Rentabilidade						
-Todos_SetorLote	11,80	8,31	79,45	-6,79	14,96	14,32	-38,40
+1	5,78	2,97	51,84	-57,49	126,23	1.134,92	-49,45
+3	-4,75	-25,25	54,81	107,10	94,97	744,12	-52,66
+4	18,02	12,35	93,34	-35,12	725,79	12,43	-40,20
+5	15,72	49,45	111,80	18,35	704,47	1.046,95	-57,78
+6	7,02	-18,99	61,21	4,20	-11,21	156,53	-54,09
+7	73,38	,38	118,71	-29,69		1.272,80	351,37
+8	36,26	-24,83	105,32	-3,18	1.188,84	826,14	-43,44
+9	23,46	-,02	43,14	-20,96	1.006,88	1.114,90	-59,09
+10	81,33	10,90	71,67	-34,25		1.748,30	497,21
+11	-57,08	-44,21	261,43	208,14		-65,54	-72,69
+12	16,06		67,73	-35,17	-51,47	1.396,50	418,54
+13	-23,04	-53,83		-11,33	-41,48	-44,41	1.222,15
+14	23,05	-5,81	72,23	-44,47	-17,80	844,09	322,23
+15	-29,01	621,84	110,49	-35,55	-52,65	-54,20	399,16
+16	33,97	63,35	93,22	-37,64	74,29	13,28	299,54
+17	63,72	17,32	96,91	-20,30	183,95	107,92	516,69
+18	10,81	454,64	217,75			-32,97	-19,44
+19	-61,05	-22,86	15,38	-72,71	-83,07	-61,43	-76,18
+20	45,46	19,39	89,19	418,66	11,93	1.525,80	-61,75
+21	102,17	-16,51	72,09	616,60	1.113,98	684,51	166,62
+22	-3,47	-36,36	-23,16	294,93	511,14	307,24	
+23	121,87	121,87					
+24	670,26	670,26					
+25	199,21	199,21					
+26	248,10	248,10					
+27	-22,57	-22,57					

Figura 49 - Documento Olap com a rentabilidade dos setores ao longo dos anos.

O documento Olap denominado “Evolução da produção por setor/lote” apresenta os setores (1 a 27), os anos (2002 a 2007), e os valores da rentabilidade obtida nas células. Por ser Olap, esta é apenas a visão inicial do documento, ou seja, o usuário pode selecionar examinar outras métricas ou incluir ou excluir dimensões (perspectivas) na análise.



Figura 50 - Métricas disponíveis para escolha do usuário – Documento Olap.

Para obter outra visão dos dados, serão adicionadas as métricas área colhida e produtividade, agregados os dados de todos os anos e ordenados os resultados pela rentabilidade, de forma a se ter os setores mais rentáveis na parte superior da listagem.

Tempo			
+Todos_AnoMes			
Measures			
Lote	↗ Produtividade	↘ Rentabilidade	↗ AreaColhida
-Todos_SectorLote	55,90	11,80	6.940,01
+24	51,93	670,26	56,50
+26	28,95	248,10	4,34
+25	26,29	199,21	40,29
+23	22,08	121,87	78,37
+21	56,75	102,17	946,66
+10	65,13	81,33	95,98
+7	67,36	73,38	246,36
+17	70,45	63,72	480,51
+20	53,02	45,46	794,12
+8	62,85	36,26	171,62
+16	62,87	33,97	310,38
+9	44,87	23,46	238,86
+14	54,99	23,05	323,63
+4	47,77	18,02	318,58
+12	59,76	16,06	115,55
+5	52,87	15,72	399,31
+18	59,96	10,81	110,02
+6	58,92	7,02	326,27
+1	55,59	5,78	260,81
+22	35,85	-3,47	133,60
+3	54,51	-4,75	198,09
+27	68,84	-22,57	21,11
+13	71,14	-23,04	243,09
+15	59,19	-29,01	530,73
+11	46,38	-57,08	236,44
+19	40,45	-61,05	258,79

Figura 51 - Dados de área colhida, rentabilidade e produtividade agregados por setor da fazenda em todos os anos.

Interessante perceber que vários setores de elevada rentabilidade acumulada possuem produtividades médias por vezes mais baixas do que setores de rentabilidade mais baixa. Uma explicação para tal comportamento reside no fato de os setores de maior rentabilidade possuírem maior área colhida proveniente de cortes de números mais altos, onde a produtividade média é mais baixa. Para ilustrar a explicação, a figura a seguir exibe o mesmo documento Olap adicionado da dimensão corte, mas exibindo apenas os setores 21 e 15, por questões de ampliação da área visível da imagem.

Na figura, no setor 21, de rentabilidade elevada (102,17%), os cortes de número 4 ou superior totalizaram 583,06 ha do total de 946,66 ha, ou seja, 61,59% do total. Já o setor 15, cuja rentabilidade foi de -29,01%, totalizou 191,55 ha do total de 530,73 ha, ou seja: 36,09% oriundos dos cortes de número 4 ou superior, que são aqueles de maior rentabilidade.

		Tempo		
		+Todos_AnoMes		
		Measures		
Lote	Corte	Produtividade	Rentabilidade	AreaColhida
+21	-All Corte	56,75	102,17	946,66
	5	49,53	1.035,92	219,25
	6	29,64	680,85	144,35
	4	62,50	616,60	181,80
	7	29,56	166,62	37,66
	3	70,15	72,09	181,80
	2	79,71	-8,26	162,64
	1	20,24	-86,52	19,16
	9			
	8			
+15	-All Corte	59,19	-29,01	530,73
	4	44,32	703,41	44,03
	5	47,63	553,35	147,52
	1	79,60	7,15	121,98
	2	61,48	-50,38	108,60
	3	55,71	-60,52	108,60
	9			
	8			
	7			
	6			

Figura 52 - Documento Olap mostrando a concentração de áreas colhidas por cortes.

Embora a visão inicial do documento apresente os dados agregados por setor, os mesmos podem ser detalhados nos lotes integrantes de cada setor, além de ser possível a filtragem para analisar qualquer lote ou setor individualmente, como já executado na figura anterior.

4.4. Cenário 4: Estudo das métricas da produção nas áreas da fazenda, em determinado ano com visualização espacial

Neste cenário, o administrador deseja visualizar em um mapa as diferentes divisões espaciais (setores e lotes), examinando-as quanto às métricas de desempenho (produção, produtividade, receita, custo, despesa e etc) em determinado ano.

Para tanto um documento do tipo “GisEngine” foi utilizado, e a figura seguinte mostra o estudo da produtividade pelos lotes para o ano de 2002.

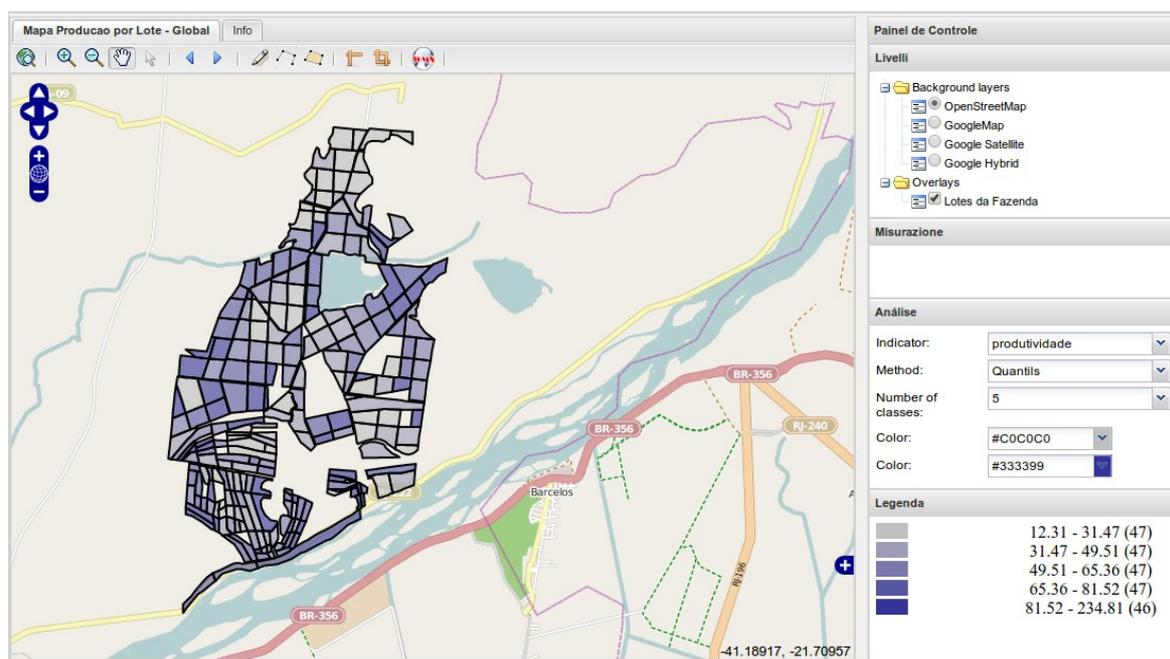


Figura 53 - Documento com os níveis de produtividade para cada lote no ano de 2002.

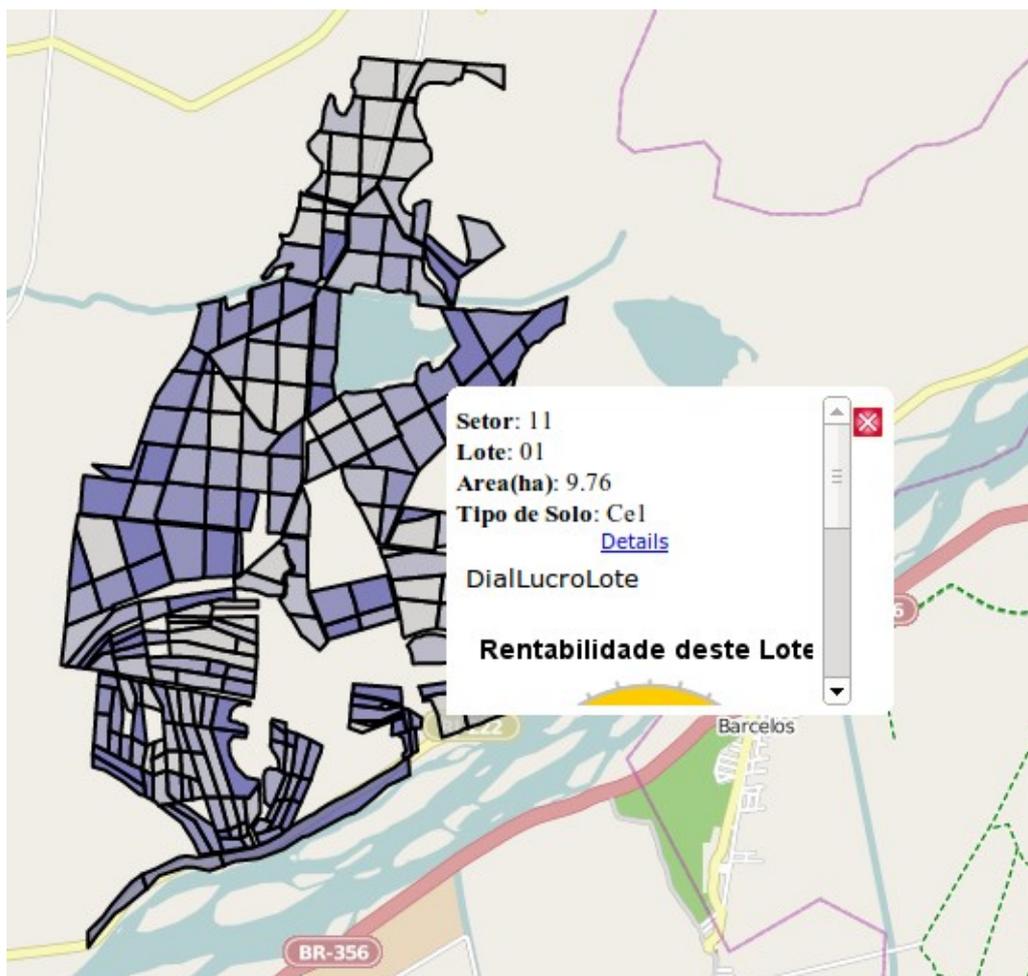


Figura 54 - Detalhe do painel que se abre quando o usuário clica sobre um dos lotes.

No documento em questão, as áreas mais claras indicam lotes com produtividade mais baixa (entre 12,31 e 31,47 toneladas/hectare). Caso o usuário queira saber um pouco mais acerca de um destes lotes, basta clicar sobre um deles.

No caso, o lote de baixa produtividade selecionado foi o número 01 do setor 11. O painel também apresenta um gráfico em formato de velocímetro com a rentabilidade (lucro / custo * 100) do lote em questão

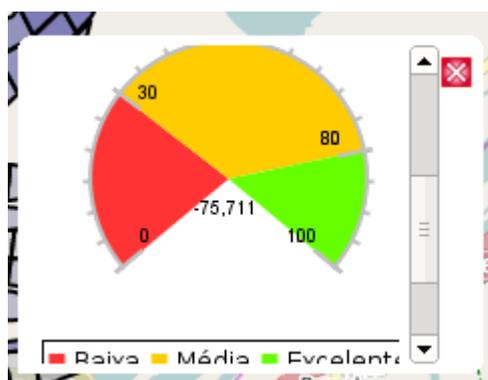


Figura 55 - Parte inferior do painel – detalhe do gráfico com a rentabilidade do lote.

Para obter mais detalhes acerca da produção do lote 01 do setor 11, pode-se clicar em “Details” e um documento Olap é exibido permitindo análise de métricas para todo o setor no qual o lote está localizado. O documento Olap aberto funciona nos moldes daquele utilizado no cenário 3

		Measures				
Lote	Corte	Produtividade	Lucro	Custo	Receita	Producao
+11	+All Corte	41,46	-\$20.657,36	\$46.729,16	\$26.071,80	1.717,99

Slicer: [Cultura=Cana] [Ano=2002]

Figura 56 - Documento Olap com métricas da produção para todo o setor no qual o lote está localizado.

Clicando-se no setor 11, os lotes do citado setor são exibidos, detalhando suas métricas e permitindo comparações na mesma área geográfica. O documento Olap também permite que se saiba de qual corte se trata a análise.

		Measures				
Lote	Corte	Produtividade	Lucro	Custo	Receita	Producao
-11	+All Corte	41,46	-\$20.657,36	\$46.729,16	\$26.071,80	1.717,99
01	-All Corte	28,46	-\$11.231,07	\$14.834,07	\$3.603,00	277,75
	1					
	2					
	3	28,46	-\$11.231,07	\$14.834,07	\$3.603,00	277,75
	4					
	5					
	6					
	7					
	8					
9						
02	+All Corte	31,58	-\$12.309,49	\$17.189,89	\$4.880,40	357,12
03	+All Corte	67,69	-\$3.238,97	\$13.450,97	\$10.212,00	599,10
04	+All Corte	42,02	\$6.122,17	\$1.254,23	\$7.376,40	484,02
05	+All Corte					
06	+All Corte					

Slicer: [Cultura=Cana] [Ano=2002]

Figura 57 - Documento Olap com os nós dos lotes expandidos, permitindo comparação entre os lotes do mesmo setor, além do corte para o lote em questão.

4.5. Cenário 5: Fazer a análise da rentabilidade pelos cortes

Conforme descrito no capítulo 3 desta tese, a fazenda Abadia dedica diferentes cuidados a cada corte da cana-de-açúcar. Os cortes 1, 2 e 3 possuem maior produtividade e com isso, recebem adubação (de plantio ou de cobertura), irrigação e cuidado mais primoroso contra plantas daninhas. Os cortes de número 4 (quatro) em diante não recebem adubação e nem irrigação.

Assim, uma questão importante para o administrador é saber como se comporta a rentabilidade (lucro / custo * 100) ao longo dos cortes, ou seja, a partir de que corte de um plantio o retorno (determinado pela rentabilidade) passou a ser positivo.

Para obter tais respostas, o documento Olap “Desempenho dos locais por corte” foi empregado.

		Corte										
		-All Corte	1	2	3	4	5	6	7	8	9	
		Measures	Measures									
Lote	>	Rentabilidade	>	Rentabilidade	>	Rentabilidade	>	Rentabilidade	>	Rentabilidade	>	Rentabilidade
>Todos_SetorLote		11,80	-28,14	-21,18	-28,75	764,33	613,39	474,48	323,68	723,37	330,56	

Slicer:

Figura 58 - Documento Olap – visão inicial da rentabilidade por corte.

O documento mostra, nesta visão agrupada de dados para toda a propriedade, que a rentabilidade média ficou negativa para os cortes 1, 2 e 3 passando a positiva somente a partir do corte de número 4.

Caso o usuário deseje investigar a rentabilidade média por cortes para cada setor da fazenda, basta expandir o nó “Todos_SetorLote”. Caso queira visualizar a rentabilidade pelos lotes de um setor, expande-se o nó relativo ao setor desejado.

Lote	Corte									
	-All Corte	1	2	3	4	5	6	7	8	9
	Measures									
-Todos_SectorLote	Rentabilidade									
	11,80	-28,14	-21,18	-28,75	764,33	613,39	474,48	323,68	723,37	330,56
+1	5,78	-26,66	33,62	-59,43	867,22	953,52	237,22			
+3	-4,75	-46,33	-31,47	-2,09	446,46	448,78	538,40	902,56		
+4	18,02	-38,48	1,28	-46,65	814,39	543,45	227,95	286,43	384,94	225,65
+5	15,72	-51,53	-3,38	-39,41	786,62	440,68	375,05	711,69	1.054,69	
+6	7,02	-15,52	-30,73	-21,93	767,61	738,17				
+7	73,38	86,05	-2,37	-26,11	1.121,95	737,27	520,57			
+8	36,26	-42,02	75,03	-5,58	1.023,25	787,09	175,47			
+9	23,46	-48,80	-13,72	-16,29	966,19	411,81	269,34	354,41	384,70	494,09
+10	81,33	83,76	-38,69	-29,79	1.787,10	671,25	558,85			
+11	-57,08	-65,54	-72,69	-58,89	282,19	228,34	285,09			
+12	16,06	67,73	-35,17	-51,47	1.396,50	418,54				
+13	-23,04	9,39	-48,93	-46,69	1.109,63					
+14	23,05	7,02	22,97	-55,62	1.163,24	814,65	322,23			
+15	-29,01	7,15	-50,38	-60,52	703,41	553,35				
+16	33,97	11,15	54,13	-19,94	1.014,95	626,63	190,99			
+17	63,72	-1,70	65,63	-14,69	1.360,76	632,48	354,50			
+18	10,81		-32,97	-19,44	509,57	237,75	353,31	180,85		
+19	-61,05	-70,51	-69,99	-87,74		573,93	296,77			
+20	45,46	-44,07	4,79	-28,31	684,07	700,36	730,06	181,49		353,61
+21	102,17	-86,52	-8,26	72,09	616,60	1.035,92	680,85	166,62		
+22	-3,47		-38,40	-25,89	232,17	232,65	307,24			
+23	121,87				121,87					
+24	670,26				670,26					
+25	199,21				199,21					
+26	248,10				248,10					
+27	-22,57		-22,57							

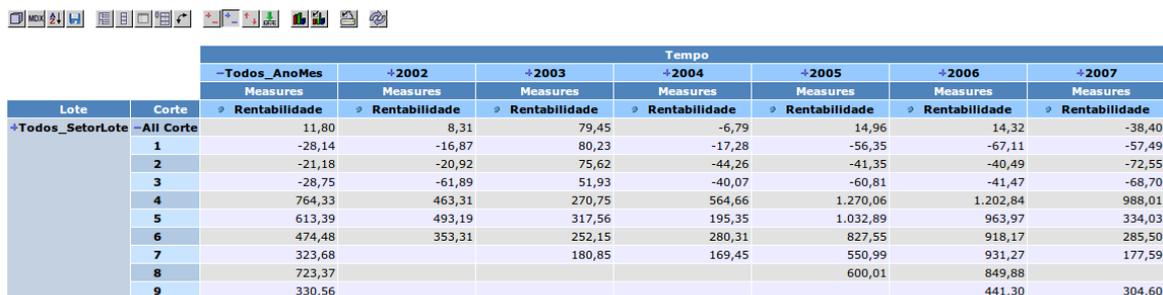
Figura 59 - Visualização da rentabilidade por cortes detalhada pelos setores.

Na visualização por setores, percebe-se que apesar da rentabilidade média do primeiro corte ser negativa (o que significa dizer que, na média dos dados observados, os custos superam as receitas), houve setores para os quais a rentabilidade se mostrou positiva ainda para o corte número 1. Ainda na Figura 59, a produção no setor 10 (que é um exemplo de rentabilidade positiva para o primeiro corte) teve rentabilidade positiva em função do ano em que ocorreu.

A rentabilidade exibida no documento da Figura 59 está agrupada pelos cortes acontecidos em qualquer época entre os anos da base de dados fornecida, ou seja, entre 2002 e 2007.

Importante destacar que tal comportamento da rentabilidade é função também do preço de venda da cana-de-açúcar. Desta forma, pode haver época (ano) em que, em função de um bom valor de venda do produto, a rentabilidade seja positiva já no primeiro corte.

De forma a explicitar a influência do tempo (anos) no preço de venda e consequentemente, na rentabilidade, a Figura 60 traz o comportamento da rentabilidade pelos cortes ao longo dos anos.



		Tempo						
		-Todos_AnoMes	+2002	+2003	+2004	+2005	+2006	+2007
		Measures						
Lote	Corte	» Rentabilidade						
+Todos_SetorLote	--All Corte	11,80	8,31	79,45	-6,79	14,96	14,32	-38,40
	1	-28,14	-16,87	80,23	-17,28	-56,35	-67,11	-57,49
	2	-21,18	-20,92	75,62	-44,26	-41,35	-40,49	-72,55
	3	-28,75	-61,89	51,93	-40,07	-60,81	-41,47	-68,70
	4	764,33	463,31	270,75	564,66	1.270,06	1.202,84	988,01
	5	613,39	493,19	317,56	195,35	1.032,89	963,97	334,03
	6	474,48	353,31	252,15	280,31	827,55	918,17	285,50
	7	323,68		180,85	169,45	550,99	931,27	177,59
	8	723,37				600,01	849,88	
	9	330,56					441,30	304,60

Slicer:

Figura 60 - Documento com a rentabilidade por corte ao longo dos anos.

Na figura com a rentabilidade dos cortes ao longo dos anos, percebe-se que no ano de 2003, em função do valor obtido na venda da cana-de-açúcar, o comportamento da rentabilidade foi diferente do que ocorreu em todos os demais anos. Em 2003, mesmo os cortes 1, 2 e 3, que possuem custos maiores, apresentaram rentabilidade positiva.

4.6. Cenário 6: Análise financeira das variedades, de forma a selecionar aquelas que trazem o maior retorno econômico.

Para estudar as variedades, avaliando-as financeiramente, um documento Olap de nome “Desempenho das variedades” foi empregado. A figura a seguir traz as variedades com seus valores totais (ao longo de todos os anos) de lucro e produção, já ordenados pelo lucro. Assim, o usuário tem a lista ordenada das variedades de maior retorno econômico para a fazenda.

Tempo		
+Todos_AnoMes		
Measures		
Cultura	Lucro	Producao
RB 72454	\$1.122.500,39	157.768,42
RB 765418	\$464.379,90	58.457,08
RB 739735	\$208.888,13	26.815,77
SP 801842	\$202.344,17	17.736,85
SP 701143	\$142.884,92	28.268,76
CB 453	\$71.425,29	5.846,08
CP 5122	\$10.658,91	2.416,50
SP 791011	\$6.154,50	19.847,77
SP 711406	-\$3.795,92	291,50
SP 803280	-\$6.157,01	259,10
RB 867515	-\$12.421,72	631,10
SP 2313	-\$21.784,50	2.000,98
SP 832847	-\$79.056,13	4.861,47
RB 758540	-\$99.506,91	8.217,18
SP 813250	-\$104.859,66	7.417,52
RB 855536	-\$151.602,92	13.905,41
RB 835486	-\$294.106,29	11.053,89
SP 792233	-\$555.332,73	22.130,92

Slicer:

Figura 61 - Documento Olap ordenado pelo lucro acumulado das variedades.

Ao longo dos seis anos de dados (2002 a 2007) existentes no data warehouse da fazenda Abadia, a variedade mais produzida foi a RB72454, com 157.768,42 toneladas, por isso esta foi a variedade de maior lucro acumulado.

Expandindo-se o estudo da produção e lucro das variedades ao longo dos anos, tem-se a figura a seguir;

Cultura	Tempo																							
	2002		2003		2004		2005		2006		2007		2002		2003		2004		2005		2006		2007	
	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao	Lucro	Producao
RB 72454	-\$1.978,70	51.854,63	\$431.627,76	34.181,18	\$220.765,61	26.199,81	\$207.010,65	24.278,41	\$340.959,41	13.748,99	-\$75.884,34	7.505,40												
RB 765418	\$50.714,04	15.852,32	\$168.011,86	12.168,74	-\$108.311,71	11.328,95	\$140.129,50	8.694,00	\$158.080,64	5.535,62	\$55.755,57	4.877,45												
RB 739735	\$26.352,73	3.997,12	\$144.887,56	7.482,88	-\$24.590,74	6.535,60	\$43.744,20	2.304,53	\$47.989,77	3.961,94	-\$29.495,39	2.533,70												
SP 801842	-\$3.183,17	756,68	\$77.923,47	3.865,19	\$2.609,31	5.148,79	\$25.749,15	3.155,85	\$46.922,91	2.364,70	\$52.322,50	2.445,64												
SP 701143	\$31.943,69	8.783,69	\$68.131,92	5.607,01	-\$54.356,18	4.743,25	-\$30.489,31	2.631,36	\$100.376,64	3.775,74	\$27.278,16	2.727,71												
CB 453	\$4.564,83	1.336,61	\$23.764,54	1.302,70	-\$2.647,83	1.204,45	\$13.483,34	673,70	\$19.812,46	618,65	\$12.447,95	709,97												
CP 5122	-\$2.872,03	133,30	\$8.301,08	733,90	-\$9.951,78	663,35	\$1.670,43	109,00	\$9.480,63	375,54	\$4.030,58	401,41												
SP 791011	-\$579,99	1.672,00	\$54.389,29	3.032,80	-\$13.716,60	4.502,25	-\$21.552,60	3.500,20	-\$43.676,62	4.739,52	\$31.291,02	2.401,00												
SP 711406											-\$3.795,92	291,50												
SP 803280											-\$6.157,01	259,10												
RB 867515											-\$12.421,72	631,10												
SP 2313			\$8.654,86	752,08	-\$10.079,91	677,40	-\$20.359,45	571,50																
SP 832847							\$209,72	828,50	-\$9.061,18	1.494,37	-\$70.204,67	2.538,60												
RB 758540			\$2.594,02	248,98	-\$2.757,81	274,10	-\$30.809,16	1.143,25	\$55.403,66	2.438,15	-\$123.937,62	4.112,70												
SP 813250	\$2.320,54	662,09	\$11.552,08	691,13	-\$8.020,74	471,30	-\$29.605,98	1.029,45	\$12.454,49	2.086,09	-\$93.560,05	2.477,46												
RB 855536			\$7.781,60	487,56	-\$11.394,28	3.147,55	-\$18.895,05	2.750,60	-\$73.906,29	3.104,39	-\$55.188,90	4.415,31												
RB 835486			\$8.553,57	647,86	-\$31.109,12	3.655,85	-\$67.064,02	2.201,95	-\$163.566,87	3.672,29	-\$40.919,85	875,94												
SP 792233			\$768,82	139,09	-\$23.228,91	3.212,85	-\$60.104,36	2.294,37	-\$279.410,84	6.076,27	-\$193.357,44	10.408,34												

Figura 62 - Análise de produção e lucro das variedades ao longo dos anos.

Chama a atenção a diminuição da produção da variedade RB 72-454, que ocorreu em atendimento às recomendações feitas por Bernardes (2002), e o crescimento na produção de outras variedades, como por exemplo a SP79-2233 que no ano de 2007 foi a de maior produção.

O lucro negativo nas variedades, em determinados anos, deve-se a dois fatores principais: o preço de venda da cana-de-açúcar e o custo de sua produção.

Pode-se perceber a influência do preço de venda no lucro quando verificamos que, para o ano de 2003, as rentabilidades de todas as variedades ficaram positivas. Isto se deveu ao bom preço pago pela tonelada de cana no citado ano de 2003 (foi o maior preço pago dentre os anos analisados).

A influência do custo da produção está relacionada ao número do corte. Como já mencionado anteriormente, os cortes de números 1, 2 e 3 recebem tratamentos não dispensados aos demais cortes (ex.: irrigação), o que aumenta o custo de sua produção. Com isso, uma análise interessante é checar o lucro para os anos detalhado por corte, o que se pode fazer no documento Olap adicionando-se a dimensão “corte” ao eixo de visualização horizontal (juntamente com a variedade).

		Cultura						
		SP 792233						
		Corte						
Tempo	Measures	--All Corte	1	2	3	4	5	6
-Todos_AnoMes	Lucro	-\$555.332,73	-\$297.233,19	-\$192.139,54	-\$144.465,03	\$65.988,77	\$12.516,26	
	Producao	22.130,92	9.005,52	5.700,09	2.896,98	3.414,69	1.113,64	
+2002	Lucro							
	Producao							
+2003	Lucro	\$768,82	\$768,82					
	Producao	139,09	139,09					
+2004	Lucro	-\$23.228,91	-\$20.522,32	-\$2.706,59				
	Producao	3.212,85	3.084,35	128,50				
+2005	Lucro	-\$60.104,36		-\$60.104,36				
	Producao	2.294,37		2.294,37				
+2006	Lucro	-\$279.410,84	-\$184.930,44	-\$16.267,61	-\$102.523,04	\$24.310,25		
	Producao	6.076,27	2.184,68	423,97	2.648,78	818,84		
+2007	Lucro	-\$193.357,44	-\$92.549,25	-\$113.060,98	-\$41.941,99	\$41.678,52	\$12.516,26	
	Producao	10.408,34	3.597,40	2.853,25	248,20	2.595,85	1.113,64	

Figura 63 - Documento Olap filtrando apenas variedade SP 792233, adicionado da dimensão “corte”.

A Figura 63 permite a análise mencionada anteriormente, mostrando o documento Olap adicionado da dimensão “corte” e filtrado exibindo apenas dados da variedade SP79-2233 (por questões de visualização). O documento sofreu uma operação de pivoteamento, ou seja, trocou o eixo vertical pelo horizontal de visualização (também para facilitar a visualização dos dados).

A leitura dos dados mostra que o lucro global da variedade SP79-2233 foi de R\$ -555.332,73. Ao se detalhar tal lucro pelos anos, apenas 2003 teve valor positivo para

todos os cortes, e a razão é já conhecida: o bom preço pago pela tonelada de cana-de-açúcar neste ano. Os demais anos apresentam lucro total dos cortes negativo e a explicação encontra-se na concentração da produção nos cortes de números 1, 2 e 3, onde os custos são mais elevados. Pode-se encontrar lucros positivos nos cortes de número 4 e 5 (anos de 2006 e 2007). Com isso conclui-se que o lucro baixo para esta variedade tomada como exemplo ocorre porque seus plantios ainda não atingiram a maturidade dos cortes de maior número (cortes de número 4 em diante), onde normalmente se dá o retorno do investimento feito nos cortes de números menores (1 a 3).

4.7. Cenário 7: Análise da produtividade média das variedades, considerando número do corte e tipo de solo

Neste cenário, o administrador deseja comparar as variedades com suas produtividades médias e estudar o comportamento de suas curvas de produtividade em função do número do corte e em função do tipo de solo.

Para esta análise foi usado um documento do tipo painel de controle (“dashboard”), que exibe simultaneamente três gráficos: no primeiro deles, que ocupa toda a porção esquerda da tela, são listadas todas as variedades com suas produtividades médias; no segundo deles, ocupando a parte superior da metade direita da tela, a curva de produtividade em função do número do corte para uma variedade; e no terceiro gráfico, situado na parte inferior da porção direita da tela, um histograma com a produtividade média em função do tipo de solo para determinada variedade selecionada no primeiro gráfico.

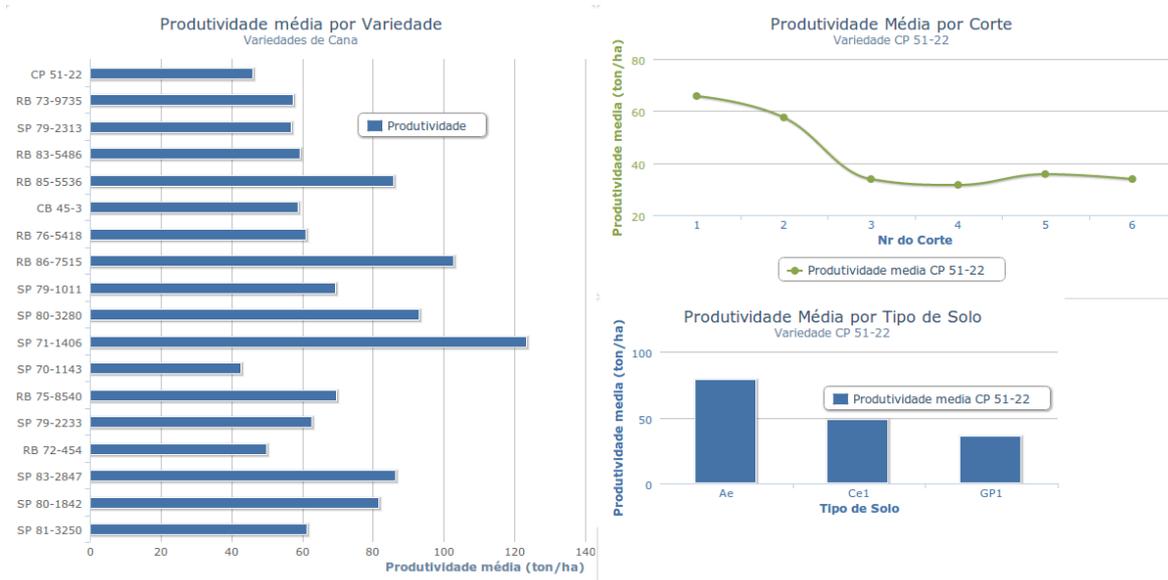


Figura 64 - Visão do documento tipo painel de controle contendo três gráficos.

A figura a seguir aproxima a visão sobre o primeiro gráfico, ou seja, aquele que mostra um comparativo das produtividades médias das variedades.

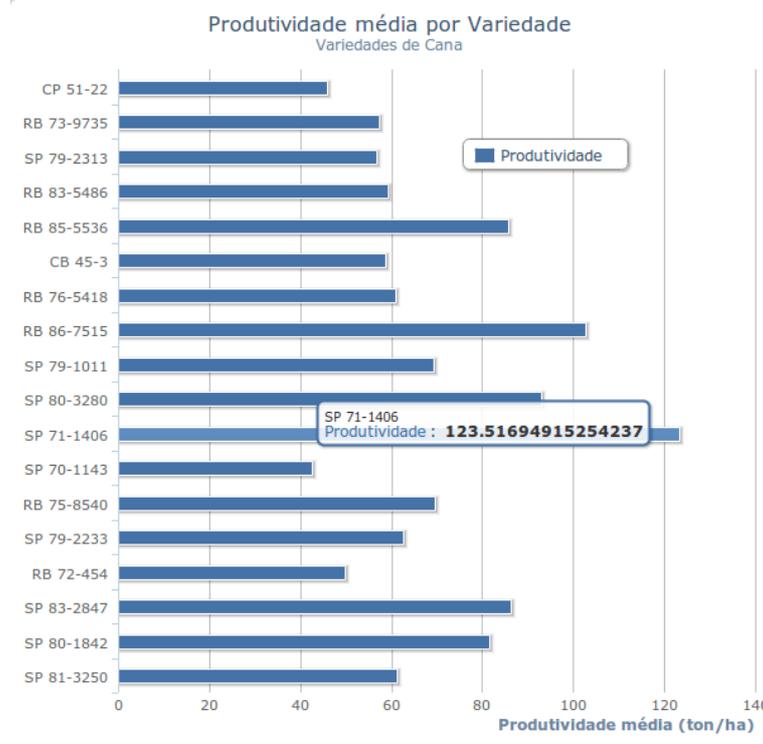


Figura 65 - Visão do comparativo das produtividades médias das variedades.

No documento da Figura 65, as variedades SP71-1406 e RB86-7515, SP80-3280 embora apresentem as maiores produtividades médias, devem ser observadas com cautela. Quando qualquer das três variedades é selecionada, percebe-se que só existem registros de produtividade para o corte de número 1, justamente o corte de maior produtividade.

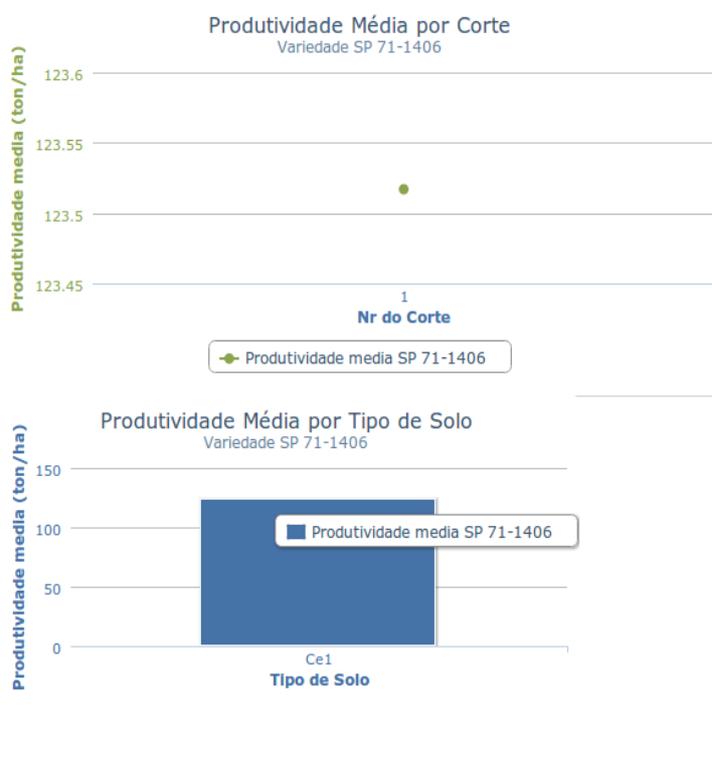


Figura 66 - Comportamento da produtividade média por corte e por tipo de solo para a variedade SP 71-1406.

Tomando por exemplo a variedade RB76-5418, que foi a segunda variedade em produção total na fazenda, o documento registra como produtividade média o valor de 61,08 toneladas/hectare. A figura abaixo detalha suas curvas de produtividade por corte e por tipo de solo.

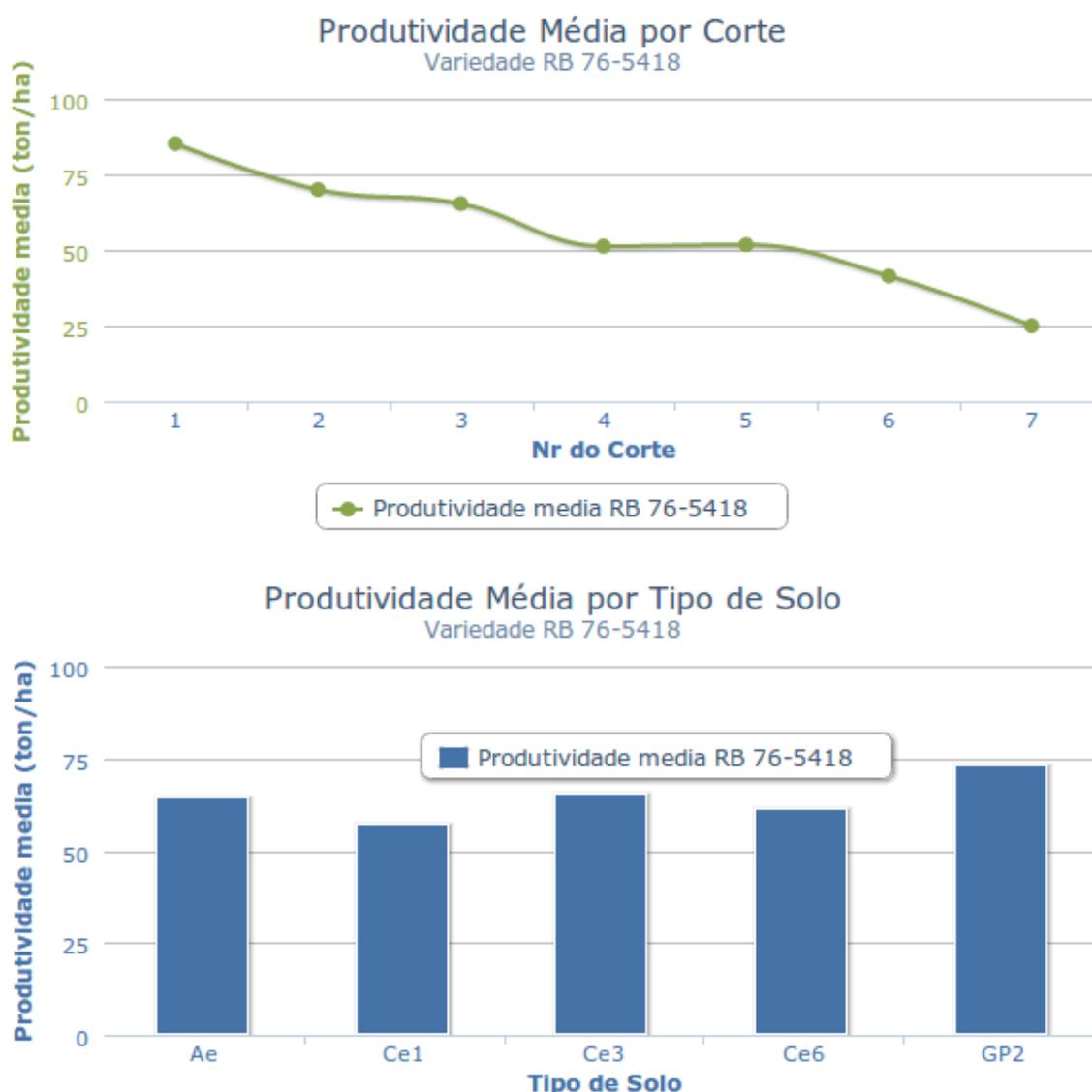


Figura 67 - Comportamento da produtividade média por corte e por tipo de solo para a variedade RB76-5418.

O documento anterior permite entender o comportamento médio da produtividade da variedade RB76-5418 e, aliado ao documento Olap citado no cenário anterior, fazer previsões acerca de futuras colheitas. A figura anterior mostra também que o solo do tipo GP2 foi o de maior produtividade média para a variedade em questão, com o valor de 72,95 toneladas/hectare.

4.8. Cenário 8: Caracterizar a produção em um determinado ano

Uma solicitação feita pelo administrador da propriedade foi um documento inicial onde se pudesse visualizar os principais indicadores da fazenda para determinado ano.

Assim, foi criado como tela inicial um painel de controle (“dashboard”), conforme a figura a seguir.

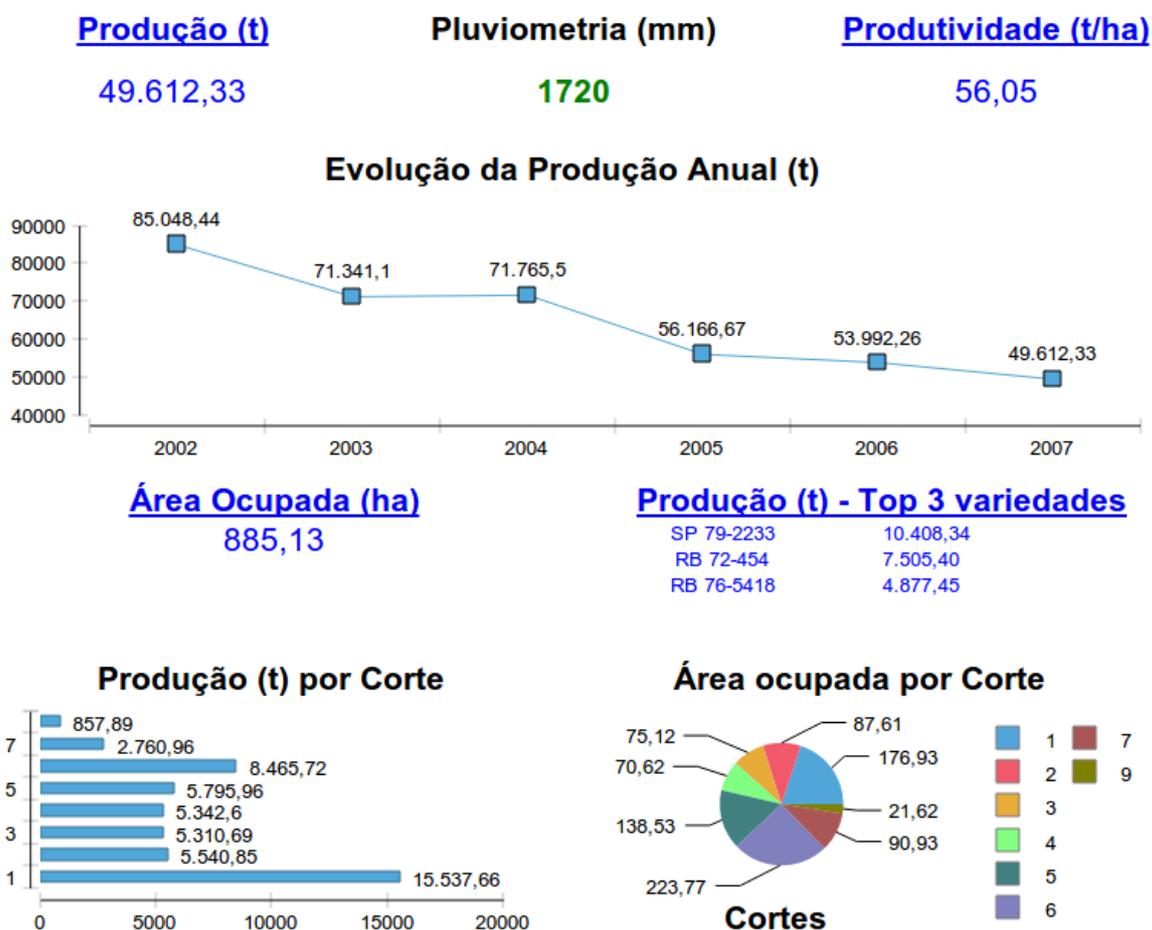


Figura 68 - Painel de controle situacional definido como página inicial da suíte SpagoBI.

O documento exibe valores para um ano e uma cultura qualquer à escolha do usuário. Por padrão, o documento exibe dados sobre o último ano registrado, que seria o ano corrente caso o usuário continuasse fazendo carga dos dados de produção até os dias atuais, e como cultura a seleção óbvia é a cana-de-açúcar, já que é a única cadastrada no sistema.

No caso da figura anterior, os dados são relativos ao ano de 2007. De forma a ser o painel de controle situacional do administrador, o documento apresenta na primeira linha: a produção total da cultura selecionada no ano escolhido, a pluviometria – que é um parâmetro norteador das demais métricas- e a produtividade média do ano. Na segunda linha do documento, um gráfico com a evolução da produção é exibido para que o usuário

tenha noção de como o ano corrente se compara com os demais. A terceira linha traz a área ocupada pela produção da cultura escolhida no ano em questão, e as três variedades de maior produção com seus valores de produção em toneladas. A quarta e última linha do documento traz a produção em toneladas por número de corte e a área ocupada por cada um dos cortes, de forma a proporcionar noção de maturidade do canal ao administrador.

Além de ser um painel de controle situacional, o documento em questão também permite ao administrador aprofundar suas análises acerca de cada parâmetro elencado na tela, por meio de links para outros documentos similares aos já exibidos nos cenários anteriores.

Assim, clicando sobre o título “[Produção \(t\)](#)” (que é um link), o usuário é conduzido a um documento do tipo Olap que permite explorar todas as métricas definidas para a produção do ano selecionado.

Lote	Measures										
	Producao	Rentabilidade	Lucratividade	Custo_ha	Lucro_ha	Receita_ha	Produtividade	AreaColhida	Lucro	Custo	Receita
→ Todos_SetorLote	49.612,33	-38,40	-62,32	\$1.535,39	-\$589,51	\$945,87	56,05	885,13	-\$521.797,13	\$1.359.017,73	\$837.220,60

Slicer: [Ano=2007]

Figura 69 - Documento Olap com métricas da produção – detalhamento do link “Produção (t)”.

O título “[Produtividade \(t/ha\)](#)” também é um link que conduz a outro documento Olap com a finalidade de analisar a evolução da produtividade ao longo dos anos, para cada setor ou lote.

A leitura dos documentos presentes nas duas primeiras linhas do painel de controle, e dos demais documentos de aprofundamento da análise (aqueles acessíveis a partir dos links), demonstra que o ano de 2007 foi o de menor produção para a fazenda, como ilustrado no gráfico evolutivo da produção anual. O índice pluviométrico em 2007 foi de 1720mm, o que afasta a hipótese de ter sido a falta de chuva a responsável pela queda na produção. A produtividade cresceu em relação aos dois anos anteriores (2005 e 2006), ficando acima da média global da fazenda, que é de 55,90 t/ha.

Lote	Tempo							
	-Todos_AnoMes	+2002	+2003	+2004	+2005	+2006	+2007	
	Measures	Measures	Measures	Measures	Measures	Measures	Measures	
-Todos_SectorLote	Produtividade	55,90	56,32	61,07	62,08	54,99	45,02	56,05
+1	Produtividade	55,59	88,25	63,07	51,17	44,18	41,07	44,67
+3	Produtividade	54,51	59,23	53,43	44,22	37,61	31,23	101,80
+4	Produtividade	47,77	64,47	56,71	50,19	39,62	30,66	43,96
+5	Produtividade	52,87	58,00	47,58	45,69	38,85	38,85	91,89
+6	Produtividade	58,92	71,56	66,34	61,84	56,50	39,40	56,33
+7	Produtividade	67,36	86,28	86,44	78,09		44,53	43,11
+8	Produtividade	62,85	67,13	81,76	72,36	56,23	33,30	53,56
+9	Produtividade	44,87	41,48	49,75	47,19	49,70	34,95	57,42
+10	Produtividade	65,13	94,27	70,00	73,68		56,50	53,81
+11	Produtividade	46,38	41,46	36,31	32,57		51,10	62,96
+12	Produtividade	59,76		68,62	72,79	61,70	47,65	48,04
+13	Produtividade	71,14	37,07		95,88	72,34	76,29	85,09
+14	Produtividade	54,99	81,58	70,20	63,78	53,27	33,75	39,19
+15	Produtividade	59,19	49,29	83,57	72,42	60,44	52,48	46,62
+16	Produtividade	62,87	67,22	77,53	70,40	66,83	55,57	39,31
+17	Produtividade	70,45	61,72	78,82	87,20	77,83	63,10	55,24
+18	Produtividade	59,96	40,19	33,13			89,93	166,24
+19	Produtividade	40,45	39,42	33,37	36,43	28,03	48,43	47,76
+20	Produtividade	53,02	61,47	51,73	48,00	61,23	50,90	42,06
+21	Produtividade	56,75	73,45	70,15	62,50	53,54	29,74	29,56
+22	Produtividade	35,85	48,23	32,06	38,93	31,92	20,24	
+23	Produtividade	22,08	22,08					
+24	Produtividade	51,93	51,93					
+25	Produtividade	26,29						
+26	Produtividade	28,95	28,95					
+27	Produtividade	68,84	68,84					

Figura 70 - Documento Olap com detalhamento da produtividade pelos anos e pelos setores ou lotes.

Na terceira linha do painel de controle, o título “Área Ocupada (ha)” é um link para um documento do tipo painel de controle (o mesmo utilizado no cenário 2), que detalha o dado de área ocupada, exibindo histórico de ocupação de área e visualização espacial dos lotes ocupados com cada variedade no ano em análise.

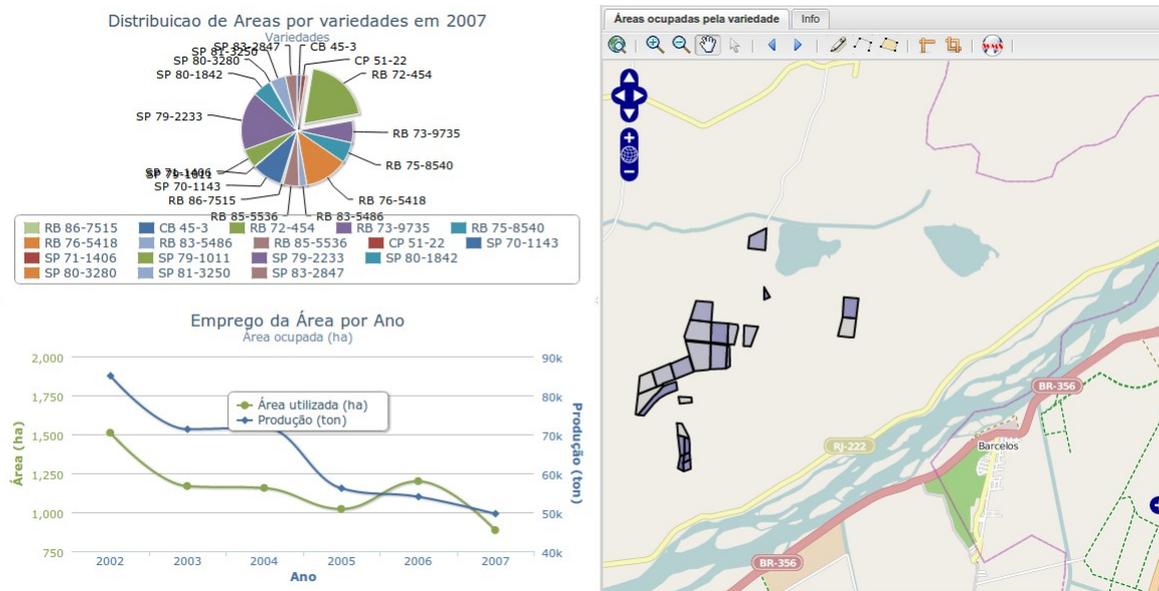


Figura 71 - Documento com detalhamento da ocupação de área por variedade.

A observação da quantidade de área ocupada (885,13 ha) e dos dados provenientes do documento de detalhamento acessado a partir de seu link, mostram que a produção em 2007 foi baixa em decorrência da pequena área colhida – 2007 teve a menor área colhida dentre todos os anos analisados. No ano em questão, chuvas intensas provocaram rompimento do dique que protege o flanco da propriedade que fica de frente para a estrada, com isso, muitas áreas foram alagadas provocando perda de boa parte da produção e reduzindo consequentemente a área colhida.

Ainda na terceira linha do painel de controle, mais um título ora denominado “Produção (t) – Top 3 variedades”, exibe o nome e a quantidade produzida das três variedades de maior produção, também é um link que conduz o usuário a outro painel de controle que faz o comparativo das rentabilidades globais das variedades (em todo o tempo) e, selecionada uma variedade, traça dois gráficos: o de rentabilidade e produtividade por número de corte e outro de rentabilidade por tipo de solo (também globais) – similar ao documento utilizado em apoio ao cenário 7.

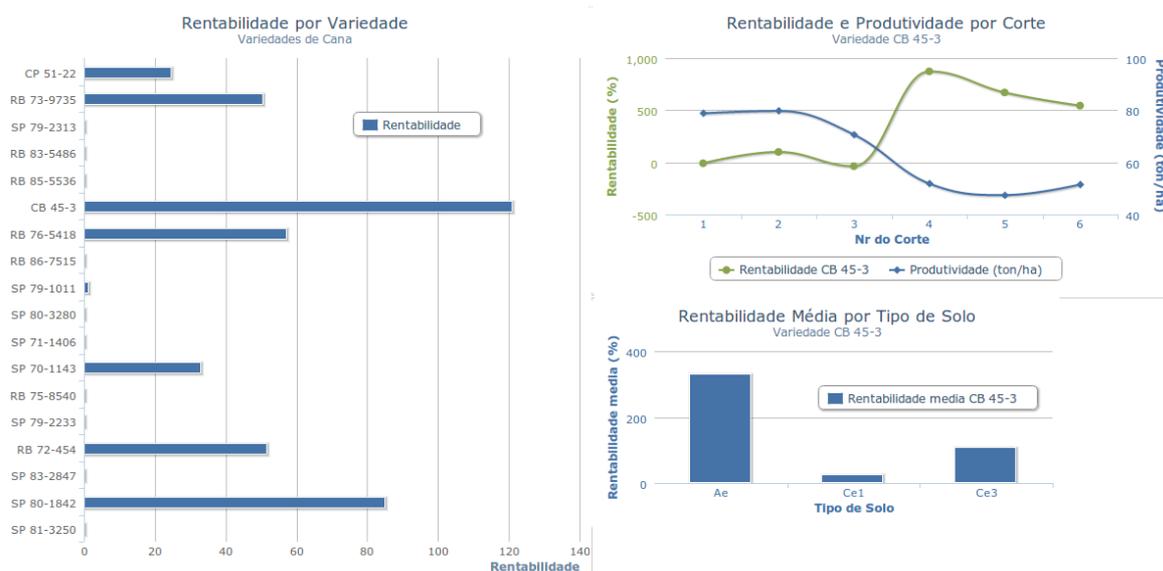


Figura 72 - Documento de comparativo de rentabilidade das variedades, rentabilidade e produtividade de uma variedade por corte e rentabilidade por tipo de solo.

A quarta linha do painel de controle traz um gráfico em barras horizontal com a produção (em toneladas) de cada corte no ano de 2007. O ano em questão foi caracterizado por uma renovação do canavial, como se pode depreender da grande produção de cana-de-açúcar oriunda de primeiro corte. Tal movimento de renovação fica

mais evidente quando se compara o gráfico de 2007 com o mesmo gráfico do painel de controle referente a 2006, como segue nas figuras a seguir.

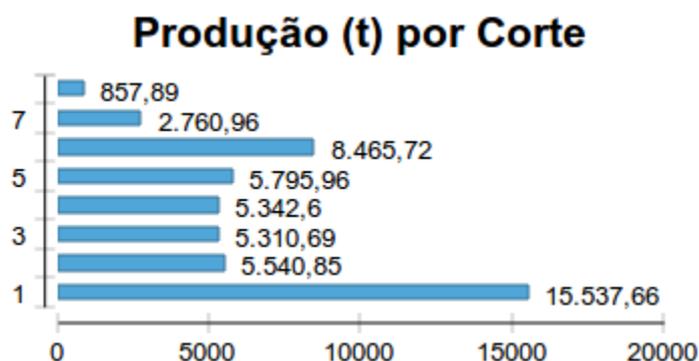


Figura 73 - Produção por Corte - ano de 2007.

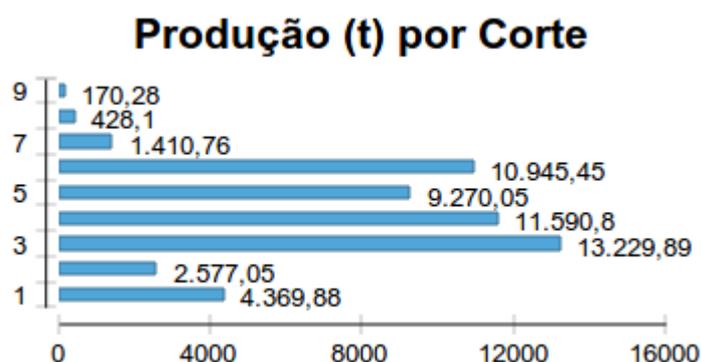


Figura 74 - Produção por Corte – ano de 2006.

Ainda na quarta linha do painel de controle, um gráfico tipo pizza traz a área ocupada por corte, onde o primeiro corte ocupou 176,93 hectares e foi o que produziu a maior quantidade de cana-de-açúcar (devido à maior produtividade deste corte), e o sexto corte ocupou 223,77 hectares produzindo a segunda maior quantidade entre os cortes colhidos.

4.9. Considerações

Powers (2002) descreve que sistemas de apoio à decisão visam sobrepujar três limites humanos. Dois destes limites são o cognitivo e o temporal.

O limite cognitivo pode ser demonstrado quando se considera que, caso a fazenda possuísse um único talhão para cultivo da cana-de-açúcar não seria difícil para o

administrador conhecer de cor sua produção em determinado ano. Mas quando tal único talhão se encontra subdividido em aproximadamente 400 lotes, torna-se humanamente impossível decorar a produção de cada um. À produção dos lotes adicione-se ainda os dados de variedade, manejo, tipo de irrigação, custos e outros que deixam bem claro porque um sistema de BI, que é um tipo de sistema de suporte à decisão, é imprescindível em um ambiente de emprego da agricultura de precisão.

Todos os cenários de análise receberam apoio de documentos do SpagoBI gerados a partir de dados obtidos na própria fazenda. Embora os dados existissem nas dependências da fazenda, a quantidade de tempo que seria gasta pelo administrador para compilar os dados necessários a cada cenário de análise a partir de suas planilhas e documentos impressos tornaria a tomada de decisão inviável, o que caracteriza o limite temporal.

Para a elaboração deste trabalho foram utilizados os dados referentes a seis anos de produção da Fazenda Abadia (2002 a 2007). Tal quantidade de anos se mostrou insuficiente para permitir análises como o uso de diferentes variedades em um mesmo lote (já que seis anos é a média de duração, em cortes, de um certo plantio em determinado local).

Também o emprego de um esquema fixo de alocação de tratamentos aos cortes nestes seis anos, conforme explicado no Quadro 3, impediu as análises de:

- Efeito do emprego de diferentes esquemas de adubação de plantio ou cobertura em um mesmo corte

- Efeito da diferenciação da irrigação sobre um mesmo corte

- Efeitos do uso de diferentes métodos de controle de plantas daninhas sobre o mesmo corte

Ou seja, o emprego do esquema fixo de tratamentos por corte limitou a quantidade de análises possíveis, pois para cada corte estava já pressuposto o tipo de adubação de plantio, de cobertura, o tipo de irrigação e também o método de controle de plantas daninhas.

A continuidade do registro das produções em anos vindouros e a utilização do sistema de BI possibilitariam ao administrador da fazenda a execução de testes e adequações em busca do ajuste ótimo de variedade, local e tratamentos para a maximização do rendimento da propriedade.

5. RESUMO E CONCLUSÕES

O princípio fundamental da agricultura de precisão é o manejo da propriedade em parcelas menores, de forma a gerenciar a variabilidade das propriedades de cada diferente porção do terreno.

Otimizando o desempenho de cada parcela do terreno, pode-se aumentar a produção em cada local e simultaneamente diminuir os custos com insumos desnecessários, elevando conseqüentemente a rentabilidade global.

Uma conseqüência de tal nível de controle é o aumento na quantidade de dados que são gerados pela atividade agrícola. Tais dados normalmente têm servido apenas a sistemas transacionais, que se prestam a geração de mapas de produtividade, de planejamento da aplicação de insumos e preparação do plantio e da colheita.

A proposta deste trabalho é utilizar a massa de dados gerado pela agricultura de precisão para fornecer suporte à decisão em nível estratégico. Como o georreferenciamento dos dados é um pressuposto básico da agricultura de precisão, uma suíte de Business Intelligence (BI) Espacial foi utilizada.

Suítes de BI, por sua complexidade, possuem custos de aquisição ou licenciamento impeditivos para pequenas e médias empresas; por isso este trabalho emprega uma suíte de código aberto e de boa reputação de mercado. O objetivo é estimular o uso de ferramentas de BI no agronegócio e por esta razão, a solução apresentada não poderia ser de custo elevado e nem um protótipo de pesquisa, cuja

descontinuidade traria riscos a todo um investimento em projeto, desenvolvimento e treinamento.

Para demonstrar todo o caminho percorrido desde a caracterização de uma propriedade rural, as dificuldades encontradas na obtenção de seus dados e a implementação da ferramenta de BI espacial, até os resultados que podem ser obtidos em apoio a cada cenário de análise, foi selecionada uma fazenda que empregasse algumas práticas típicas da agricultura de precisão e que dispusesse de razoável nível de controle e registro de suas atividades. A fazenda Abadia, propriedade produtora de cana-de-açúcar, localizada no município de Campos dos Goytacazes, foi utilizada como estudo de caso para a implementação e implantação da solução de BI espacial.

Embora a fazenda Abadia possuísse registro de suas atividades: variedades plantadas, quantidades colhidas, locais de plantio, tipo de manejo empregado e outros, tais dados não estavam em um sistema de informações apoiado por uma base de dados estruturada. Apenas os dados das quantidades colhidas estavam disponíveis em arquivos digitais (planilhas eletrônicas); os demais dados tais como tipos de manejo, valor de venda da produção e custo de cada local e variedade colhida tiveram de ser obtidos a partir de registros em papel.

Desta forma, a atividade de preparação de tais dados não-espaciais e também a correção de suas inconsistências se mostrou a etapa mais trabalhosa do processo de construção do sistema de BI. Além dos dados já citados, os dados espaciais também custaram razoável esforço para sua obtenção. Ao término da etapa denominada Extração, Transformação e Carga, os dados não espaciais - oriundos de uma única planilha de compilação de dados denominada "PlanilhaBase" - e os dados espaciais - oriundos de objetos do PostGIS, após transformações a partir de arquivos tipo shape - estavam devidamente estruturados em um data warehouse espacial e disponíveis para consumo por parte da suíte de BI - SpagoBI.

Com base nas necessidades de informação passadas pelo administrador da fazenda Abadia, e descritas em cenários de análise, que caracterizam situações de tomada de decisão, diversos documentos foram criados utilizando a suíte SpagoBI. Cada conjunto de documentos foi projetado de forma a enriquecer com dados contextualizadores cada um dos cenários de análise descritos pelo decisor.

A capacidade de tratar com dados espaciais, denominada como Inteligência de Localização, foi uma das razões da escolha do SpagoBI como suíte de Business Intelligence. A Inteligência de Localização utiliza a componente geográfica existente nos

dados para expor novos padrões e relacionamentos que seriam imperceptíveis em relatórios numéricos tabulares. Assim, a apresentação dos dados sobre o contexto espacial – mapas – enriquece a experiência do usuário, instruindo novas conclusões.

Este trabalho deixa como contribuição para as organizações rurais, principalmente as pequenas e médias, de baixa capacidade de investimento, um sistema de BI espacial cuja finalidade é fazer a caracterização da atividade agrícola, auxiliando no crescimento de sua produção e na redução de seus custos, maximizando os resultados financeiros.

O emprego de uma suíte de Business Intelligence de código aberto permite que todos os documentos desenvolvidos neste trabalho sejam aproveitados por qualquer outra organização rural, bastando fazer a importação de seus dados para o data warehouse espacial, o que pode ser facilmente executado por meio do uso de arquivos de “PlanilhaBase” (um para cada ano de produção), ou reconfigurando no Talend Open Studio as rotinas de extração, transformação e carga, para receberem os dados a partir de outra fonte qualquer.

REFERÊNCIAS BIBLIOGRÁFICAS

Actuate - Actuate: the BIRT Company: <http://www.actuate.com/> em 09/05/2014

Barquini, R. Planning and designing the Warehouse, New Jersey, Prentice-Hall, 1996.

Batchelor, B., Whighan, K., Dewitt, J., et al. Precision agriculture: introduction to precision 4p. agriculture. Iowa Cooperative Extension, 1997. Disponível na Internet. <http://www.extension.iastate.edu/Pages/precisionag/prec-ag.pdf> em 18 Ago. 1999.

Benigno, M. - Instalação do PostGIS 2.0 no Ubuntu via terminal: <http://profmarcello.blogspot.com.br/2013/10/instalacao-do-postgis-20-no-ubuntu-via.html> em 04/04/2014

Bernardes, M. S. Modelo bioeconômico para manejo e tomada de decisão em lavoura de cana-de-açúcar. ESALQ, Piracicaba, 2012

Butler, M. - 5 Free Open Source BI Suites: <http://butleranalytics.com/5-free-open-source-bi/> em 09/05/2014

CanalTech Corporate - <http://corporate.canaltech.com.br/materia/business-intelligence/Conhecendo-a-arquitetura-de-Data-Warehouse/> em 09/05/2014

Capelli, N.L. Agricultura de precisão - Novas tecnologias LIE/DMAQAG/para o processo produtivo. FEAGRI/UNICAMP, 1999: <http://wwwbases.cnptia.embrapa.br/cria/gip/gipap/capelli.doc> em 15 Out. 1999.

Cardoso, F. M. Um framework para data warehouse espacial. Dissertação (Mestrado em Ciência da Computação). UFCG, Campina Grande, 2009.

Chaturvedi, K.K.; Rai, A.; Dubey, V.K.; Malhotra, P.K. On Line Analytical Processing in Agriculture using Multidimensional Cubes. Journal of The Indian Society of Agricultural Statistics, 2006.

Come, G. Contribuição ao Estudo da Implementação de Data Warehousing: Um Caso no Setor de Telecomunicações. Dissertação de Mestrado. São Paulo: Universidade de São Paulo, 2001.

Correa, E.F.; Pizzigatti, P.L.; Junior, J.R.A.; Alves, L.R.A.; Saraiva, A.M. Data Warehouse for soybeans and corn market on Brazil. European Federation for Information Technology in Agriculture, Food and the Environment (EFITA), 2009.

Davis, G., Casady, W., Massey, R. Precision agriculture: An introduction. Water quality. University of Missouri-System, 1998. P.8. Disponível na Internet. <http://www.fse.missouri.edu/mpac/pubs/wq0450.pdf> em 17 Jun. 1999.

Elmasri, R., Navathe, S. Sistemas de Banco de Dados. Sexta edição. Pearson Education do Brasil, São Paulo, 2011.

EMBRAPA. Tecnologia em mecanização no Brasil: Equipamentos e sistemas para o futuro. In: SEMINÁRIO TEMÁTICO PARA PROSPECÇÃO DE DEMANDAS EM PESQUISA E DESENVOLVIMENTO EM MECANIZAÇÃO AGRÍCOLA NO BRASIL, 1997, Sete Lagoas-MG. Disponível na Internet. <http://wwwbases.cnptia.embrapa.br/cria/gip/gipap/seminario.doc> em 15 Out. 1999.

Gartner - Magic Quadrant for business Intelligence and Analytics Plataforms: <http://www.gartner.com/technology/reprints.do?id=1-1QLGACN&ct=140210&st=sb> em 09/05/2014

GeoServer - <http://br.geoserver.org/> em 09/04/2014.

GeoServer – Publishing a PostGIS table: <http://docs.geoserver.org/stable/en/user/gettingstarted/postgis-quickstart/index.html> em 19/04/2014.

Golfarelli, M. Open Source BI Platforms: a Functional and Architectural Comparison. In 11th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2009), Lintz, Austria, pp. 287-297, 2009

Diuf. Decision Support Systems. <http://diuf.unifr.ch/ds/courses/dss2002/pdf/DSS.pdf> em 12/05/2014.

Howson, C. Successfull Business Intelligence: Secrets to Making BI a Killer App, McGraw-Hill, New York, 2008. 258pp.

Inmon, W.H. Como Construir o Data Warehouse. Campus, Rio de Janeiro, 1997

Jaspersoft - Editions: <http://www.jaspersoft.com/editions> em 09/05/2014

Kimball, R. The Data Warehouse Toolkit, John Wiley & Sons Inc., New York, 1996.

Loshi, D., Business Intelligence the savvy manager's guide, Morgan Kaufmann Publishers, San Francisco – CA, 2003. 292 pp.

Machado, Felipe Nery Rodrigues. Tecnologia e Projeto de Data Warehouse: uma visão multidimensional. São Paulo: Érica, 2008.

Miller, W.; Supalla, R. Precision farming in Nebraska: A status report, 1996: <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=2023&context=extensionhist> em 10/05/2014

Mondrian - <http://sourceforge.net/projects/mondrian/> em 09/05/2014

MSDN - Fundamentos e Modelagem de Bancos de Dados Dimensionais
<http://msdn.microsoft.com/pt-br/library/cc518031.aspx> em 09/05/2014.

Nilakanta, S.; Scheibe, K.; Rai, A. Dimensional Issues in agricultural data warehouse designs. Computers and Electronics in Agriculture vol. 60, pp 263-278, 2008.

Ogutu, S. - SpagoBI 4.0 Baby Steps. www.ogutu.org. 259 páginas. 2012.

Ogutu.org – SpagoBI Videos and tutorials: <http://www.ogutu.org> em 12/04/2014.

Pentaho - Pentaho Community & Commercial Product Comparison:
<http://www.pentaho.com/assets/pdf/pentaho-ce-vs-ee-com.pdf> em 09/05/2014

Ponniah, P. Data Warehousing Fundamentals: A comprehensive guide for IT Professionals. New York, USA: John Wiley & Sons, 2001. 518pp

PostGIS - <http://postgis.net> em 12/05/2014

PostgreSQL - <http://www.postgresql.org> em 12/05/2014

Power, D. J. A Brief History of Decision Support Systems. Versão 4.0 Março 2007
<http://DSSResources.COM/history/dsshhistory.html> em 13/05/2014.

Rizzi, S. (2007) - Conceptual Modeling Solutions for the Data Warehouse. In: Wrembel, R., Koncilia, C., Data Warehouses and OLAP: Concepts, Architectures and Solutions. Londres: IRM Press, p.1-26.

Roza, D. Novidade no campo: Geotecnologias renovam a agricultura. Revista InfoGEO, n 11 - jan/fev, 2000. <http://mundogeo.com/blog/2000/02/02/novidade-no-campo-geotecnologias-renovam-a-agricultura/> em 13/05/2014.

Santos, J. - PostgreSQL & PostGIS: Instalação do banco de dados espacial no windows 7: <http://www.processamentodigital.com.br/2013/05/05/postgresql-postgis-instalacao-do-banco-de-dados-espacial-no-windows-7/> em 04/04/2014

Silberschatz, A., Korth, H.F., Sudarshan, S. Sistema de Banco de Dados. 5ª Edição. Rio de Janeiro: Elsevier, 2006. 805pp

SpagoBI - SpagoBI the only 100% open source, complete and flexible Business Intelligence Suite: <http://www.spagoworld.org/xwiki/bin/view/SpagoBI/> em 09/05/2014

TOS – Talend Open Studio for Data Integration: <http://www.talend.com/products/data-integration> em 04/04/2014

Thomsen, E. Olap Solutions: building multidimensional information systems. Second Edition. New York, USA: John Wiley & Sons, 2002.

Turban, E., Sharda, R., Aronson, E. J., King D., Business Intelligence, um enfoque gerencial para a inteligência do negócio, Bookman, São Paulo, 2009

Weber, T. Modelo de dados para gerenciamento de informações de colheitas de cereais. Monografia, 2009: <http://www.unidavi.edu.br/?pagina=FILE&id=56987> em 13/05/2014.

APÊNDICE A – CONFIGURAÇÃO METADADOS – SPAGOBİ

Descrição de método para instalar os metadados do SpagoBI (para utilizar o postgresSQL ao invés do HSQLDB default)

1) Criar um BD no Postgresql e sobre ele executar os scripts de criação de tabelas PG_create.sql e PG_create_quartz_schema.sql, que estão disponíveis para download em http://forge.ow2.org/project/showfiles.php?group_id=204. Obs.: O banco de dados do quartz pode ser criado dentro do proprio SpagoBI ou em um banco de nome 'quartz' separado (e tal configuração separada deverá ser informada em quartz.properties)

2) Alterar o conteúdo dos arquivos

- conf/server.xml
- webapps/SpagoBI/WEB-INF/classes/hibernate.cfg.xml
- webapps/SpagoBI/WEB-INF/classes/quartz.properties
- webapps/SpagoBI/WEB-INF/classes/jbpm.cfg.xml
- webapps/SpagoBI/WEB-INF/classes/jbpm.hibernate.cfg.xml

PS.: O Banco BAM é provavelmente utilizado para log.

**Checar se o driver jdbc4 do PostgreSQL está no diretório SpagoBI/lib/ e excluir de lá o driver versão 8.0 (para evitar conflito de versões)

a) Em conf/server.xml, altere (no caso do BD PostgreSQL) de :

```
<Resource name="jdbc/spagobi" auth="Container"
  type="javax.sql.DataSource" driverClassName="org.hsqldb.jdbcDriver"
  url="jdbc:hsqldb:file:${catalina.base}/database/spagobi"
  username="sa" password="" maxActive="20" maxIdle="10"
  maxWait="-1"/>
```

Para:

```
<Resource name="jdbc/spagobi" auth="Container"
  type="javax.sql.DataSource" driverClassName="org.postgresql.Driver"
  url="jdbc:postgresql://localhost:5432/spagobi"
  username="postgres" password="db_password" maxActive="20" maxIdle="10"
```

```
maxWait="-1"/>
```

b) Em hibernate.cfg.xml, usar: `<property name="hibernate.dialect">org.hibernate.dialect.PostgreSQLDialect</property>`

c) Em quartz.properties, descomentar as linhas referentes ao PostgreSQL.

d) Em jbpm.hibernate.cfg.xml, inserir a seguinte linha (comentando as demais relacionadas): `<property name="hibernate.dialect">org.hibernate.dialect.PostgreSQLDialect</property>`

APÊNDICE B – CONFIGURAÇÃO DO MODELO JSON PARA O DOCUMENTO GEOREPORT (“GISENGINE”)

mapName: Informar o Nome que será dado ao Mapa. Ex.: "Mapa Venda por Cidade Seleccionada",

analysisType: Especifica o método de tematização do mapa. Os tipos possíveis são: 'Choropleth' (para layers do tipo POLYGON) ou 'proportionalSysmbols' (para objetos geográficos tipo POINT),

feautreInfo: Traz os dados que devem ser exibidos em um popup quando o usuário clicar sobre um objeto da camada. Ex.: [{"PV Numero","nr"}],

indicators: São as métricas da Tabela fato a serem oferecidas para análise pelo georeport. Ex.: [{"valorvenda", "valorvenda"}, {"venda", "venda"}],

businessId: Nome do atributo no DW a ser usado para relacionar com a camada geográfica (Deve ter o mesmo conteúdo do atributo geold). Ex.: "cidadepv",

geold: Nome do atributo na Layer de interesse a ser usado para relacionar com o DW (Deve ter o mesmo conteúdo do atributo businessId) Ex.: "cidadepv" - Use atributos que não contenham espaços em seus valores (cidadepv não deve conter valores como por ex: "Casimiro de Abreu"),

targetLayerConf: Detalha os dados sobre a camada do geoserver que será relacionada com a análise para o documento em questão

```
Ex.: targetLayerConf: {
  text: 'Municipios da Praça'
  , name: 'municipios'
  , url: 'http://localhost:8080/geoserver/wfs'
},
```

inlineDocumentConf: Traz os dados de um Documento que será aberto em um pop-up. Neste Documento, o próprio engine irá adicionar os dados da seção "feautreInfo" e também adicionará um link chamado "Details" de onde se poderá abrir outro Documento, mas agora em uma nova aba (de acordo com a especificação de detailDocumentConf, descrita logo abaixo). PS.: As duas seções devem ser configuradas de forma conjunta (uma sem a outra não funciona)

```

Ex.: inlineDocumentConf: {
    label: 'RptVendaCidade'
    , staticParams: {
        param1: 'foo'
    }
    , dynamicParams: {
        ano: 'ANO',
        mes: 'MES',
        cidadepv: 'CIDADEPV'
    }
    , displayToolbar: 'false'
    , displaySliders: 'false'
},

```

detailDocumentConf: Traz parâmetros para a abertura de um outro documento em nova aba. Um exemplo seria um drill-down de um mapa de cidades para uma outra aba com outro mapa agora mostrando os pontos daquela cidade escolhida.

```

Ex.: detailDocumentConf: {
    label: 'geovendaporpv'
    , staticParams: {
        departmentId: '3'
    }
    , dynamicParams: {
        ano: 'ano',
        cidadepv: 'cidadepv'
    }
}

```

onde label: é o nome do documento a ser aberto na nova aba

staticParams: onde são passados parametros estáticos para o documento

dynamicParams: onde são passados parâmetros dinâmicos de um documento para o outro