UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO

Renato Tavares

**PROPOSTA DE UM SISTEMA DE APOIO À DECISÃO PARA
CONTROLE E GERENCIAMENTO AGRÍCOLA EM
USINAS DE AÇÚCAR E ÁLCOOL**

São Carlos - SP
2008

Renato Tavares

**PROPOSTA DE UM SISTEMA DE APOIO À DECISÃO PARA
CONTROLE E GERENCIAMENTO AGRÍCOLA EM
USINAS DE AÇÚCAR E ÁLCOOL**

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Engenharia de Produção.

Área de Conhecimento: Gestão do Conhecimento e Sistemas de Informação.

Orientador: Prof. Dr. Edson Walmir Cazarini.

São Carlos - SP
2008

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

T231p Tavares, Renato
 Proposta de um sistema de apoio à decisão para
 controle e gerenciamento agrícola em usinas de açúcar e
 álcool / Renato Tavares ; orientador Edson Walmir
 Cazarini. -- São Carlos, 2008.

 Dissertação (Mestrado-Programa de Pós-Graduação e Área
 de Concentração em Engenharia de Produção) -- Escola de
 Engenharia de São Carlos da Universidade de São Paulo,
 2008.

 1. Sistemas de apoio à decisão. 2. *Data Warehouse*.
 3. Usinas de açúcar e álcool. I. Título.

FOLHA DE JULGAMENTO

Candidato: Bacharel **RENATO TAVARES**

Dissertação defendida e julgada em 04/07/2008 perante a Comissão Julgadora:

Prof. Dr. **EDSON WALMIR CAZARINI** (Orientador)
(Escola de Engenharia de São Carlos/USP)

Aprovado

Prof. Associado **FABIO MULLER GUERRINI**
(Escola de Engenharia de São Carlos/USP)

Aprovado

Prof. Dr. **FERNANDO CELSO DE CAMPOS**
(Universidade Metodista de Piracicaba/UNIMEP)

APROVADO

Prof. Associado **REGINALDO TEIXEIRA COELHO**
Coordenador do Programa de Pós-Graduação em
Engenharia de Produção

Prof. Associado **MARCIO ANTONIO RAMALHO**
Vice-Presidente da Comissão da Pós-Graduação da EESC
Em Exercício

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me proporcionar novamente esta oportunidade de cursar uma universidade, agradeço a ele também por seu grande amor puro e verdadeiro.

Aos meus pais Manoel e Sueli, aos meus irmãos Andréia e Alexandre e ao meu filho Renan que fazem parte da minha vida não somente nos momentos alegres, mas principalmente nos momentos mais difíceis.

A USP - Universidade de São Paulo.

Ao meu orientador Prof. Dr. Edson Walmir Cazarini, pelo acompanhamento pontual e competente e pelas inúmeras contribuições.

Aos Professores do Curso de Pós-Graduação.

À Coordenação do Curso de Pós-Graduação.

Aos colegas de classe pelo companheirismo e amizade que tiveram comigo.

Ao colega e amigo de classe Marcio Roberto, que sempre me incentivou e deu forças para continuar até o final do curso.

Aos funcionários do Departamento de Engenharia de Produção, José Luiz, Fábio, Daniel, Silvana, Sueli e Luiz Fernando, pela atenção e disposição com que me ajudaram sempre que precisei.

A todos que direta ou indiretamente contribuíram para a realização desta pesquisa.

Epígrafe

“Se você quer ter sucesso na vida, tem que ter dedicação total, buscar seu último limite e dar o melhor de si”

Ayrton Senna da Silva

Resumo

TAVARES, R. **Proposta de um sistema de apoio à decisão para controle e gerenciamento agrícola em usinas de açúcar e álcool.** 2008. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2008.

Aliada à crescente evolução da computação, dois fatores também começaram a receber maior atenção: o conhecimento e a informação. Esta evolução faz com que a informação possa estar disponibilizada a todos, contribuindo de certa forma, para o auxílio na aquisição de conhecimento e visando posterior tomada de decisão, observa-se também a grande importância dos bancos de dados. Este trabalho apresenta a proposta de um Sistema de Apoio à Decisão (SAD), utilizando-se de avançadas metodologias de armazenamento de dados em poderosos bancos de dados (*Data Warehouse*), para o Controle e Gerenciamento Agrícola em Usinas de Açúcar e Álcool. Foi desenvolvido um ambiente estruturado, extensível, projetado para a análise de dados não voláteis, logicamente e fisicamente transformados, provenientes de diversas aplicações, alinhados com a estrutura da empresa, atualizados e mantidos por um longo período de tempo, referidos em termos utilizados no negócio e sumarizados para análise rápida. Com a implantação destas novas tecnologias, a empresa estará apta a obter informações do nível gerencial e estratégico para ajudar nos seus processos de tomada de decisão, o que antes não era possível com os atuais sistemas de informação existentes na empresa.

Palavras-Chave: Sistemas de Apoio à Decisão, *Data Warehouse*, Usinas de Açúcar e Álcool.

Abstract

TAVARES, R. **Proposal for systems of support to the decision for the control and management agricultural in sugar and ethanol plants.** 2008. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2008.

Information and knowledge are receiving more attention with the growing of computing evolution. This evolution provides available information to everybody contributing for the acquisition of knowledge, transforming the database in an important key of this evolution process. This work presents a research about the use of technologies of Systems of Support to the Decision by using methodologies of data storage in powerful data bases (Data Warehouse) for the Control and Management Agricultural in Sugar and Ethanol Plants. An environment was developed for the analysis of non-volatile data, extensible, projected and physically transformed, proceeding from diverse applications, aligned with the structure of the company, updated and kept for a long period of time, related in business' terms and grouped for fast analysis. With the implantation of these new technologies, the company will be able to obtain information of the managerial and strategic level to help on its decision making processes, what was not possible before with the current systems of information existent at the company.

Key words: *Systems of Support to the Decision, Data Warehouse, Sugar and Ethanol Plants.*

Sumário

Lista de Figuras	i
Lista de Tabelas	ii
Lista de Abreviaturas, Siglas e Símbolos	iii
Capítulo 1 - Introdução	01
1.1 - Contexto	01
1.2 - Formulação do Problema	02
1.3 - Objetivo do Trabalho	04
1.4 - Justificativa do Trabalho	04
1.5 - Metodologia de Pesquisa	05
1.6 - Estrutura dos Capítulos	10
1.7 - Outline da Pesquisa	12
Capítulo 2 - Sistemas de Informação e de Apoio à Decisão	13
2.1 - Processo Decisório	13
2.2 - Sistemas de Informação Gerenciais (SIG)	14
2.3 - Sistemas de Apoio à Decisão (SAD)	15
2.4 - Considerações Finais do Capítulo	20
Capítulo 3 - <i>Data Warehouse</i> e Ferramentas de Extração	22
3.1 - Conceito	22
3.2 - Principais Características de um Data Warehouse	25
3.3 - Granularidade	29
3.4 - O Modelo Dimensional	36
3.5 - Data Mart	41
3.6 - Arquitetura do <i>Data Warehouse</i>	42
3.7 - Considerações Finais do Capítulo	44
Capítulo 4 - Agroindústria Canavieira no Brasil	46
4.1 - Introdução	46
4.2 - Impactos no Meio Ambiente	47
4.3 - Características da Cadeia Produtiva	48
4.4 - Os Processos de Produção no Brasil	49
4.5 - Cana Orgânica x Cana Convencional	55

4.6 - Evolução da Produção no Brasil	62
4.7 - Custos de Produção	63
4.8 - Exportações	67
4.9 - Projeção para Agroindústria Canavieira no Brasil	70
4.10 - Números do Setor e Participação do Capital Estrangeiro	74
4.11 - Considerações Finais do Capítulo	77
Capítulo 5 - Aplicação da Pesquisa-ação	78
5.1 - Capacidade de Produção e Dados Estatísticos	78
5.2 - Pesquisa-ação em uma Usina de Açúcar e Álcool	84
5.2.1 - Primeira Fase da Pesquisa-ação: Pesquisa Exploratória	87
5.2.2 - Segunda Fase da Pesquisa-ação: Pesquisa Aprofundada	88
5.2.3 - Terceira e Quarta Fase da Pesquisa-ação: Ação e Avaliação	89
5.3 - Metodologia (Construindo o <i>Data Warehouse</i> Dimensional)	89
5.4 - Os Nove Pontos de Decisão	90
5.5 - A Fase de Carregamento	106
5.6 - O Sistema de Extração de Dados de Produção	107
5.7 - Estrutura de Armazenamento de Dados Proposta	118
5.8 - Discussões entre o Modelo Proposto e Revisão Bibliográfica	120
5.9 - Considerações Finais do Capítulo	123
Capítulo 6 - Conclusão	124
6.1 - Propostas para Trabalhos Futuros	126
Referências Bibliográficas	127
Bibliografia Complementar	134
Anexos	
Anexo_1 - Questões com Relação à Produção	
Anexo_2 - Questões com Relação ao Rendimento	
Anexo_3 - Questões com Relação ao Preço e Faturamento	
Anexo_4 - Questões com Relação aos Dados Tecnológicos	
Anexo_5 - Questões com Relação aos Custos	
Anexo_6 - Questões com Relação a Comparativos	
Anexo_7 - Questões com Relação a Estimativas	
Anexo_8 - Questões com Relação a Simulações	
Todas as Questões Apresentam:	
Linguagem SQL utilizada na Consulta	
Informações Geradas pela Linguagem SQL	
Relatório Desenvolvido com as Informações Geradas	

Lista de Figuras

Figura 1	Visão Macro da Estrutura dos Capítulos	11
Figura 2	Outline da Pesquisa	12
Figura 3	Os Sistemas de Informação	13
Figura 4	Processo Prospectivo	19
Figura 5	Ambiente Operacional x Ambiente de Data Warehouse	26
Figura 6	A Questão da Não-Volatilidade	27
Figura 7	Variável no Tempo	28
Figura 8	Integração dos Dados no Ambiente de DW	29
Figura 9	Definindo o Nível de Granularidade	30
Figura 10	Modelo Dimensional de um Negócio	32
Figura 11	Modelo de Dependência de Dados de um Negócio	33
Figura 12	O Relatório Ideal de um DW	34
Figura 13	Um Modelo Dimensional Típico	36
Figura 14	Arquitetura Genérica de um DW	43
Figura 15	Visão Macro dos Processos Agrícolas de uma Usina	50
Figura 16	Visão Geral dos Processos de uma Usina	54
Figura 17	Comparativo das Produções (Anidro, Hidratado e Açúcar)	59
Figura 18	Comparativo das Produções (Alcool Total x Açúcar)	59
Figura 19	Percentual de Produção (Alcool Total x Açúcar)	60
Figura 20	Produção de Cana de Açúcar do Brasil	61
Figura 21	Composição dos Custos de Produção de Cana	64
Figura 22	Comportamento do Preço da Cana e Custos de Produção	66

Figura 23	Exportação de Alcool do Brasil	67
Figura 24	Projeção de Exportação de Alcool do Brasil	68
Figura 25	Exportação de Açúcar do Brasil	69
Figura 26	Produção por Montadoras	72
Figura 27	Vendas de Carros Flex no Brasil	72
Figura 28	Vendas por Tipo de Combustível	73
Figura 29	Participação nas Vendas	73
Figura 30	Produção dos Cinco Maiores Grupos do Setor	75
Figura 31	Participação do Capital Estrangeiro na Safra 05/06	76
Figura 32	Participação do Capital Estrangeiro na Safra 06/07	76
Figura 33	Capacidade Diária de Moagem de Cana	79
Figura 34	Capacidade Diária de Produção de Açúcar	79
Figura 35	Capacidade Diária de Produção de Alcool	80
Figura 36	Capacidade Diária de Produção de Levedura	80
Figura 37	Total de Cana Moída em 2005	81
Figura 38	Total de Cana Moída em 2006	81
Figura 39	Evolução Líquida da Receita	82
Figura 40	Moagem de Cana nos últimos 10 Anos	82
Figura 41	Produção de Açúcar nos últimos 10 Anos	83
Figura 42	Produção de Alcool nos últimos 10 Anos	83
Figura 43	Produção de Levedura nos últimos 10 Anos	84
Figura 44	A Tabela de Fatos com as Dimensões Identificadas	93
Figura 45	Relacionamento das Tabelas (Fornecedor x Fato)	94
Figura 46	Relacionamento das Tabelas (Tipo x Fato)	95

Figura 47	Relacionamento das Tabelas (Estágio x Fato)	95
Figura 48	Relacionamento das Tabelas (Variedade x Fato)	95
Figura 49	Relacionamento das Tabelas (Talhão x Fato)	95
Figura 50	Relacionamento das Tabelas (Fazenda x Fato)	96
Figura 51	Relacionamento das Tabelas (Fundo x Fato)	96
Figura 52	Relacionamento das Tabelas (Solo x Fato)	96
Figura 53	Relacionamento das Tabelas (Tempo x Fato)	97
Figura 54	A Tabela de Fatos com os Fatos Identificados	97
Figura 55	Leitura dos Dados de Produção para Tabela de Fatos	108
Figura 56	Leitura dos Dados do Cadastro de Fundos Agrícolas	109
Figura 57	Leitura dos Dados do Cadastro de Fazendas	109
Figura 58	Leitura dos Dados do Cadastro de Estágios de Corte	110
Figura 59	Leitura dos Dados do Cadastro de Fornecedores	110
Figura 60	Leitura dos Dados do Cadastro de Solos	111
Figura 61	Leitura dos Dados do Cadastro de Variedades de Cana	111
Figura 62	Leitura dos Dados do Cadastro de Talhões de Cana	112
Figura 63	Leitura dos Dados do Cadastro de Tipos de Cana	112
Figura 64	Transformando em Imagens de Registro de Carga	114
Figura 65	Ferramenta para Carga dos Dados em Massa para o DW	115
Figura 66	Índices Desativados para Carga dos Dados	116
Figura 67	Modelo Proposto	118
Figura 68	Modelagem Lógica dos Dados	119

Lista de Tabelas

Tabela 1	Principais Diferenças entre DW e Data Mart	41
Tabela 2	Unidades Produtoras de Açúcar e Alcool no Brasil	57
Tabela 3	Produções de Cana, Açúcar e Alcool no Brasil	58
Tabela 4	Produtividade x Custo	66
Tabela 5	Dimensão Tempo	98
Tabela 6	Dimensão Fornecedor de Cana	98
Tabela 7	Dimensão Solo	98
Tabela 8	Dimensão Fundo Agrícola	99
Tabela 9	Dimensão Variedade de Cana	99
Tabela 10	Dimensão Fazenda	99
Tabela 11	Dimensão Talhão de Cana	100
Tabela 12	Dimensão Estágio de Corte	100
Tabela 13	Dimensão Tipo de Cana	101
Tabela 14	Tabela de Fatos Produção	101

Lista de Abreviaturas e Siglas

ATR	AÇÚCAR TOTAL RECUPERÁVEL
CPFL	COMPANHIA PAULISTA DE FORÇA E LUZ
DBF	ARQUIVO DE DADOS
DBMS	SISTEMA DE GERENCIAMENTO DE BANCO DE DADOS
DDW	DATA WAREHOUSE DIMENSIONAL
DER	DIAGRAMA ENTIDADE / RELACIONAMENTO
DW	DATA WAREHOUSE
EIS	SISTEMAS DE INFORMAÇÃO PARA EXECUTIVOS
ERP	SISTEMAS DE GESTÃO EMPRESARIAL
GEE	GASES DE EFEITO ESTUFA
GNV	GÁS NATURAL VEICULAR
IAA	INSTITUTO DO AÇÚCAR E DO ALCOOL
IDEA	INSTITUTO DE DESENVOLVIMENTO AGROINDUSTRIAL
OLAP	PROCESSAMENTO ANALÍTICO ON-LINE
OLTP	PROCESSAMENTO DE TRANSAÇÕES ON-LINE
PC	PUREZA DA CANA
POL	POL DO CALDO
SAD	SISTEMA DE APOIO À DECISÃO
SI	SISTEMA DE INFORMAÇÃO
SIG	SISTEMAS DE INFORMAÇÕES GERENCIAIS
SQL	LINGUAGEM PADRÃO PARA ACESSAR BANCO DE DADOS RELACIONAIS
VHP	AÇÚCAR CRISTAL UTILIZADO COMO MATÉRIA PRIMA PARA OUTROS PROCESSOS

Capítulo 1 - Introdução

Atualmente os Bancos de Dados são muito importantes para as Grandes Empresas assim como a análise dos dados neles existentes. Geralmente as grandes empresas detêm um volume enorme de dados distribuídos em diversos sistemas de informação diferentes. A recuperação e a sua conseqüente análise torna-se muitas vezes complicada visto que, as informações estão pulverizadas pelas diferentes áreas funcionais da organização. Para organizar estes dados são necessários métodos de armazenamento, estruturação e novas tecnologias que permitem a geração e recuperação dessas informações de forma correta e consistente, facilitando com isso a tomada de decisão.

1.1 - Contexto

Desde o plantio da matéria-prima até sua transformação dentro da fábrica, o sucesso de uma Usina de Açúcar e Alcool depende do alinhamento e da sintonia entre os elos da sua cadeia produtiva. Com a utilização de sistemas integrados e de apoio à tomada de decisão, a empresa conseqüentemente estará habilitada a atender os principais requerimentos competitivos da agroindústria e assim responder, com transparência, aos desafios estratégicos de redução de custos e melhoria da produtividade.

Exercendo um maior controle sobre os processos de produção, potencializa-se a redução do uso de insumos e o aumento da produtividade e produção. O maior controle sobre os processos de gestão possibilitam maior rapidez e precisão de informações.

Direcionando para a área de agronegócios e mais precisamente Usinas de Açúcar e Alcool o avanço dessas novas tecnologias tem contribuído para que Diretores e Gerentes da área agrícola tenham um grande número de informações à sua disposição para tomar decisões.

A área agrícola dessas empresas é muito importante, talvez possa ser considerada como a mais importante, pois é ela que vai gerar a matéria-prima (Cana-de-Açúcar), para a produção de seus principais produtos (açúcar e álcool) como também seus derivados.

1.2 - Formulação do Problema

Os sistemas transacionais garantem a operação da empresa e dão suporte ao dia-a-dia do negócio, porém, as informações que contêm, não suprem todas as necessidades de consultas dos gerentes, que muitas vezes, precisam de dados históricos decorrentes de vários anos, os quais não estarão disponíveis nos sistemas que gerenciam a empresa no seu dia-a-dia (MEIRA, 2004).

À medida que a complexidade interna e os negócios de uma empresa aumentam, o processo de tomada de decisão tende a se tornar também mais complexo, requerendo agilidade, dinamismo, utilidade e precisão das ações e informações empresariais. Para conseguir informações estratégicas, os gestores das empresas necessitam de sistemas de informação efetivos, que processem grande volume de dados e produzam informações precisas, completas, econômicas, flexíveis, confiáveis, relevantes, simples e em tempo.

Nos últimos anos, houve uma transferência do poder e da influência da tecnologia de banco de dados para a tecnologia das ferramentas de apoio à decisão, o que levou as empresas buscarem obter e trabalhar com informações gerenciais que dão suporte à ela.

Atualmente, é possível coletar e processar um grande número de dados, mas uma vez coletados, precisam ser analisados para se determinar sua relevância.

As empresas normalmente não sofrem de falta de dados, mas de uma abundância e redundância de dados inconsistentes (MEIRA, 2004).

Devido à grande concorrência de mercado, tanto nos setores de produtos quanto de serviços, a possibilidade de acessar informações confiáveis com velocidade e garantia de qualidade, no momento certo, é a chave do sucesso para a maioria das empresas.

A maneira pela qual a maioria dos gerentes tomam suas decisões, ainda está fundamentada em informações armazenadas em sistemas inadequados para o processo decisório. O ambiente real e prático que este profissional atua, exige que as ferramentas de seu uso diário estejam agregadas às tecnologias de ponta, viabilizando informações estratégicas seguras e rápidas, objetivando manter sua empresa em destaque no atual mercado global e competitivo em que atua (SEIXAS, 2000).

Porém é necessário verificar quais as informações que realmente serão necessárias para dar suporte às operações de negócios da empresa. Essa visão genérica é freqüentemente denominada de Arquitetura de Informação (Arquitetura de Dados), que procura responder às seguintes questões:

- Que dados coletar?
- Onde e como serão coletados?
- Como serão transmitidos?
- Como serão processados e armazenados?
- Quais aplicações usarão estes dados?
- Como estas aplicações estão relacionadas com a organização?
- Qual a importância dessas informações para a organização?

Após o desenvolvimento de modelagem da nova estrutura e o carregamento das informações para o ambiente proposto, o mesmo deve estar apto a responder a várias questões utilizando-se dessa nova tecnologia, sendo que as principais informações seriam sobre:

1. Produção;
2. Rendimento;
3. Preços;
4. Dados Tecnológicos: Fibra da Cana (FIBRA), Pureza da Cana (PC), Pol do Caldo (POL) e Açúcar Total Recuperável (ATR) por Tonelada de Cana;
5. Custos e Faturamento.

A nova ferramenta deve permitir também aos diretores, gerentes e engenheiros a utilização de telas e relatórios que forneçam:

1. Comparativos;
2. Estimativas;
3. Cenários e
4. Simulações.

1.3 - Objetivo do Trabalho

Proposta de um Sistema de Apoio à Decisão (SAD), utilizando um modelo de estrutura de armazenamento de dados, capaz de integrar todas as informações relevantes para a tomada de decisão, geradas pelos sistemas que se encontram distribuídos no ambiente operacional agrícola da empresa.

A integração talvez seja a parte mais importante desse processo, pois ela será responsável por sincronizar os dados destes sistemas e colocá-los no mesmo padrão garantindo a qualidade e integridade dessas informações.

1.4 - Justificativa do Trabalho

Qualquer empresa pode se beneficiar significativamente com um *Data Warehouse (DW)*. A ferramenta permite fazer análises de tendências e identificar relacionamentos muitas vezes não esperados pelo tomador de decisão e, com certeza, em tempos de intensa competitividade, pode ser considerada como uma poderosa ferramenta estratégica (SEIXAS, 2000).

A maioria dos sistemas de informação atuais fornecem os dados brutos, sem nenhum significado mais específico, porque estão justamente fora de um contexto. Mas, quando as pessoas analisam estes dados e fornecem o relacionamento necessário, baseados em um ponto de partida e modelo de desempenho comum, tornam esse processo de contextualização visível e explícito.

Sendo assim, torna-se de suma importância identificar novas tecnologias aplicadas à informação para que o ambiente corporativo possa utilizá-las de forma adequada e produtiva.

Toda e qualquer organização, seja qual for o porte e o ramo de negócio, adota estratégias para alcançar os objetivos planejados, visando direcionar e coordenar esforços, definir a estrutura e sobreviver ao ambiente competitivo.

Verificando essa situação, essa pesquisa, busca por meio do estudo, aplicação e uso destas novas tecnologias auxiliarem nos processos decisórios dentro da organização. Os benefícios de se usar um SAD não são realizáveis em todas as situações de decisão ou por todos os tomadores de decisão. Tal realização depende do grau de ajuste entre o tomador de decisão, o contexto da decisão e o próprio SAD (MARAKAS, 1999).

1.5 - Metodologia da Pesquisa

A pesquisa será realizada por meio de pesquisa-ação em uma Usina de Açúcar e Álcool da Região de Ribeirão Preto, Estado de São Paulo.

O filósofo americano John Dewey foi o precursor da pesquisa-ação, na medida em que propôs métodos indutivos científicos para solucionar problemas práticos em áreas como educação, psicologia, filosofia, dentre outros (MCKERNAN, 1991 *apud* BAKER, 2000).

O método de pesquisa-ação tem se destacado como um recurso metodológico capaz de equilibrar o rigor que caracteriza os procedimentos típicos de uma investigação acadêmica e o compromisso com a transformação da realidade. Este método está fortemente associado à abordagem qualitativa, calcada na realização de exercícios em que prevalecem os recursos da argumentação, da reflexão, e da compreensão orientados para processos deliberados de intervenção frente a situações consideradas insatisfatórias. Assim, presta-se a alterar as condições percebidas como passíveis de transformação. Desta forma, o sujeito e o objeto da investigação assumem, voluntariamente, posições ativas no processo investigatório (BAKER, 2000).

A pesquisa-ação é uma abordagem intervencionista para aquisição de conhecimento científico desenvolvida na era pós-positivismo (BASKERVILLE & HARPER, 1996).

Os pesquisadores que adotam o método de pesquisa-ação têm a oportunidade de conhecer a realidade investigada por dentro, levando em conta não apenas o discurso proferido pelos atores implicados, mas também suas ações e interpretações acerca do que é investigado (MORIN, 2004).

O uso combinado de diferentes fontes de dados e informações, característico da abordagem qualitativa, tende a compensar a falta de representatividade estatística pelo aprofundamento intensivo e contextualizado da investigação que a abordagem permite (DEMO, 2002).

Na pesquisa-ação, o pesquisador ao expressar interesse pelas contribuições dos colaboradores, reconhece o domínio que eles apresentam sobre o que está sendo investigado (MÁSTER, 1995 *apud* BAKER, 2000).

Alguns exemplos de situações de reconhecimento: demonstração do respeito pelas interpretações, explicações e análises elaboradas sobre a situação investigada e declaração da intenção da colaboração para processos de melhoria na organização. É inevitável que haja reconhecimento da existência de expressiva convergência de objetivos. Isso contribui para que pesquisadores e pesquisados construam as bases de uma relação de complementaridade no processo investigatório. Assim, haverá co-responsabilidade tanto na formulação do diagnóstico organizacional quanto na elaboração de planos de intervenção.

Neste contexto, o pesquisador não é reduzido a um mero relator dos acontecimentos. Ele interage com os participantes da pesquisa dentro do ambiente deles. A pesquisa realizada por meio da observação no dia-a-dia da organização investigada ainda por entrevistas favorece o desenvolvimento de julgamentos pertinentes dado a familiaridade com os acontecimentos diários e, ao mesmo tempo, amplia a percepção sobre as idéias que sustentam as práticas de gestão (CHIZZOTI, 1991).

Assim, os julgamentos formulados pelos pesquisadores, ainda que parciais, derivam do acúmulo do citado convívio diário. É impossível conhecer o que efetivamente interessa se os pesquisadores não estiverem dispostos a ser parte integrante do universo pesquisado, em suas palavras: “sem que estejamos verdadeiramente envolvidos, pessoalmente pela experiência, na integralidade de nossa vida emocional, sensorial, imaginativa, racional” (BARBIER, 2002).

Assim sendo, o êxito do método, em grande parte, depende da composição daquilo que esse mesmo autor intitula de “pesquisador-coletivo” (BAKER, 2000).

A pesquisa-ação é justificada porque nos estudos organizacionais, representa uma possibilidade de conciliar o sujeito e o objeto da investigação (THIOLLENT, 1997).

A observação participante contribuirá para o levantamento de informações sobre o assunto e fornecerá subsídios para uma melhor análise dos dados. Os dados serão coletados por meio de: técnicas interativas diversas, tais como entrevistas não estruturadas, observação participante, seminários, *workshops*, etc.

A metodologia de pesquisa-ação é adequada para a realização de estudos organizacionais, principalmente quando o problema tem relação forte com o comportamento humano (HOLANDA & RICCIO, 2001).

Descrição do Método

Apesar da pesquisa-ação possuir uma estrutura flexível, é possível identificar quatro fases que compõem o estudo (THIOLLENT, 1997):

- **FASE 1** - Exploratória - são identificados os atores e realizado o diagnóstico para identificar os problemas, as capacidades de ação, e intervenção na organização;
- **FASE 2** - Pesquisa Aprofundada - ocorre a coleta de dados de acordo com o projeto de pesquisa;
- **FASE 3** - Ação - a partir dos resultados da fase anterior é realizado o planejamento da ação, pela discussão de objetivos alcançáveis por meio de ações concretas;
- **FASE 4** - Avaliação - consiste da observação, redirecionamento das ações e resgate do conhecimento adquirido durante o processo. Cabe salientar que estas fases não são rígidas, principalmente as três últimas onde existe uma simultaneidade de pesquisa e ação. Além disso, a questão da aprendizagem não está presa a uma fase e, sim, difusa ao longo do processo de pesquisa. A geração de conhecimento e o desenvolvimento de teorias acontecem em todas as etapas, devido ao caráter dinâmico da pesquisa.

A essência de uma pesquisa-ação é um processo simples de dois estágios. Primeiro, o estágio de diagnósticos que envolvem uma análise colaborativa da situação pelo pesquisador e os sujeitos da pesquisa. Hipóteses são formuladas sobre a natureza da pesquisa. Segundo, o estágio terapêutico, experimentos de mudança colaborativa. Nesse estágio, as mudanças são introduzidas e os efeitos são estudados (BLUM, 1955 *apud* BASKERVILLE & HARPER, 1996).

Entretanto, tendo em vista atingir o rigor científico, estruturas adicionais são normalmente impostas à pesquisa-ação.

O resumo do processo cuja definição de pesquisa-ação como capaz de cobrir simultaneamente a solução de problemas na prática e a expansão de conhecimento científico, assim como desenvolver as competências dos atores envolvidos. É desenvolvida colaborativamente numa situação imediata usando dados de *feedback* num processo cíclico, idealizando o aumento do entendimento de uma situação, principalmente aplicável para o entendimento de mudanças de processos em sistemas sociais e comprometidos em uma estrutura ética (HULT & LENNUNG, 1980 *apud* BASKERVILLE & HARPER, 1996).

A pesquisa-ação consiste de um time de profissionais, e possivelmente teóricos, que planejam, agem, e avaliam resultados das ações que foram executadas (HOLANDA & RICCIO, 2001).

Existe uma ênfase para o seminário como técnica central para a condução de todo o processo de pesquisa-ação. Em torno do seminário, conduzido pelos grupos e supervisionado pelo grupo de coordenação, devem ser estimuladas as discussões de tema, problema e as abordagens apropriadas de condução.

O êxito de pesquisas realizadas com este método, além de depender de domínio técnico, teórico e metodológico, depende igualmente da capacidade de o pesquisador conquistar a confiança de colaboradores e dos patrocinadores da pesquisa (THIOLLENT, 1997).

Plano de Estudo

A presente pesquisa está centrada na questão de propor um SAD utilizando um DW como modelo de estrutura de armazenamento de dados e concentrando-se nos benefícios da implementação do novo sistema, quando ajustado a organização.

A origem da pesquisa foi uma reunião com os diretores da empresa pesquisada, sobre a falta de informações do nível gerencial e estratégico, ou seja, a empresa possui sistemas de informação implantados, utiliza sistemas de gestão *Enterprise Resources Planning* (ERP), porém encontra grande dificuldade em obter essas informações. A discussão foi norteadada pelos objetivos da empresa e pela vontade expressada pelos diretores de possuir um sistema de informação capaz de extrair e coletar informações dos sistemas operacionais agrícolas existentes da empresa e transforma-los em informações estratégicas para o processo de tomada de decisão.

Elaborou-se, então, uma proposta de pesquisa-ação, que foi explicada detalhadamente para os diretores e gerentes da usina. Com sua aprovação, deu-se, então, início a uma pesquisa intervencionista, com o objetivo de auxiliar a empresa na solução dos seus problemas. A abordagem adotada foi a interpretativa, ou seja, os participantes da empresa e o pesquisador trabalharam juntos na identificação de problemas, causas e possíveis soluções, se apoiando em métodos qualitativos, e de forma contextualizada.

O estudo foi estruturado de acordo com o método de pesquisa-ação descrito por THIOLENT (1997). Ele foi dividido em quatro fases: a fase exploratória, a fase de pesquisa aprofundada, a fase da ação e a fase da avaliação.

A fase exploratória constituiu-se de diagnóstico interno e externo à Empresa, onde foram levantadas informações sobre a usina. Ainda nesta fase, foram identificados os participantes da pesquisa e os problemas enfrentados pela organização. Para tanto, foram realizadas reuniões com os diretores e gerentes da usina, e, posteriormente, com o restante dos colaboradores envolvidos.

Foi desenvolvido um modelo de estrutura e armazenamento de dados (*DW*) e a partir daí, iniciou-se a coleta de dados e a implementação do novo sistema. A coleta de dados foi realizada contemplando dados disponíveis nos sistemas operacionais agrícolas existentes na própria empresa. Nas entrevistas não estruturadas foram abordados temas tais como: entendimento dos problemas, questionamentos sobre causas e possíveis soluções dos problemas, planos de ação, etc.

Durante o processo da pesquisa, houve o redirecionamento de ações e resgate do conhecimento adquirido por meio de estudos, relatos e anotações. Isso se traduziu em alertas de melhoras na implantação do sistema e também nos relatórios que poderiam ser desenvolvidos.

A última fase consistiu na apresentação do modelo proposto e avaliação dos resultados da implementação utilizando-se deste novo modelo juntamente com a identificação das barreiras e benefícios atingidos durante o processo. Os resultados obtidos nesta fase são apresentados nos capítulos cinco e seis deste trabalho.

1.6 - Estrutura dos Capítulos

Este trabalho está organizado em seis capítulos:

Capítulo - 1

Apresentação do assunto dissertado e enfoque dado ao problema. Logo após, foram apresentados o objetivo e a justificativa da pesquisa de acordo com a linha de pesquisa em que o trabalho se encontra inserido e por final a metodologia de pesquisa aplicada no trabalho.

Os Capítulos dois e três tratam da revisão da literatura:

Capítulo - 2

Apresenta uma introdução do processo decisório dentro das organizações como também sua relevância no atual mundo globalizado em que vivemos atualmente. Também são apresentados os principais conceitos sobre Sistemas de Informação Gerenciais (SIG) e Sistemas de Apoio à Decisão (SAD).

Capítulo - 3

Apresenta os principais conceitos sobre a tecnologia de *DW*, suas características, granularidade e modelagem. São demonstradas as principais diferenças entre um modelo dimensional (foco deste trabalho) e um modelo de dependência de dados. É dada uma rápida introdução ao conceito de *Data Mart* e finalizando com as características da arquitetura do *DW*.

Capítulo - 4

Apresenta uma visão geral sobre a agroindústria canavieira no Brasil, faz uma breve introdução, mostra os impactos com relação ao meio ambiente, as características da cadeia produtiva, os processos, evolução, custos de produção, perspectivas sobre o futuro e alguns números do setor.

Capítulo - 5

O objetivo deste capítulo é apresentar todo o processo da pesquisa-ação, incluindo a descrição da organização objeto do presente estudo e o desenrolar das quatro fases da metodologia da pesquisa-ação.

Toda a fase de construção do ambiente de *DW* (metodologia, os nove pontos de decisão, a fase de carregamento, o sistema de extração de dados de produção e a estrutura de armazenamento de dados), é demonstrada passo a passo.

Capítulo - 6

O objetivo deste capítulo é apresentar a conclusão do trabalho, avaliando os principais resultados obtidos frente à utilização da implantação da tecnologia de SAD juntamente com a proposta do novo modelo de estrutura e armazenamento de dados.

A figura 1 mostra a visão macro da estrutura dos capítulos:

Figura 1 - Visão macro da estrutura dos capítulos.

1.7 - Outline da Pesquisa

Capítulo 2 - Sistemas de Informação e de Apoio à Decisão

Atualmente, com a globalização e um mercado caracterizado por constantes mudanças, torna-se necessária a divisão do processo decisório nos níveis estratégico, tático e operacional. Neste cenário, é necessário transformar os objetivos organizacionais em objetivos coletivos, envolvendo todos os membros da organização, e os colaboradores necessitam de habilidades de liderança, comunicação, negociação, assim como serem agentes de mudança extremamente capacitados para as atuais necessidades (FREIRE, 2000).

2.1 - Processo Decisório

No ambiente das empresas e organizações, ter ou não ter a informação no momento certo pode significar a sobrevivência da empresa. A informação além de ser precisa deve ser obtida no menor tempo possível (GONÇALVES, 2005).

Normalmente a visualização gráfica é a triangular. Porém, seguindo os princípios de Bertalanffy por meio da Teoria Geral dos Sistemas, é mais adequada a visão piramidal com abrangência e profundidade em cada um dos diversos segmentos, conforme mostra a figura 3.

Figura 3 - Os Sistemas de Informação Adaptado de (LAUDON & LAUDON, 1996).

Atualmente sabe-se que uma grande parte dos problemas enfrentados pelas organizações diz respeito ao comportamento humano. Mais ainda, muitos problemas comportamentais que são usados, implementam e estruturam sistemas que não podem ser expressos por intermédio de normas rígidas dentro de uma visão estritamente técnica.

Assim, o enfoque atual dos Sistemas de Informação é sócio-técnico, considerando com igual importância tanto a parte técnica e matemática quanto a comportamental.

As decisões gerenciais afetam diretamente a sobrevivência das empresas e a vida daqueles que giram em torno dela, sejam eles empregados, acionistas, fornecedores, clientes ou, muitas vezes, a própria sociedade. Por isso, o processo decisório em uma empresa é algo importante não apenas para o seu ambiente interno. Prova disso, é a pressão que o tomador de decisão sofre, principalmente, de seus clientes (MIGLIOLI, 2006).

Diante desse quadro, para que os decisores tomem boas decisões, há a necessidade de que estas sejam baseadas em dados confiáveis que serão transformados nas variáveis do problema a ser solucionado e indicarão as alternativas de solução do problema (OSTANEL, 2005).

2.2 - Sistemas de Informação Gerenciais

Gerenciar negócios das organizações sempre foi uma árdua tarefa para administradores sempre preocupados em como fazer para melhorar seus resultados, principalmente resultados para tomada de decisão.

Baseados nestas constantes preocupações surgem os SIG, como uma ferramenta de apoio aos administradores.

Um SIG é caracterizado por possuir alguns requisitos fundamentais para a sua implementação tais como: suporte a informações complexas, suporte a informações externas, ampla disponibilidade geográfica, excelente tempo de resposta, desempenho para análise intensa de dados e suporte a múltiplos modelos de interface.

Os SIG, ainda apresentam as seguintes características:

- Foco na Informação;
- Direcionado a gerentes de nível médio;
- Fluxo de Informação bem estruturado;
- Integração de sistemas por área funcional (SI comercial, SI financeira, entre outras);
- Geração de Relatórios e Consultas utilizando gerenciadores de banco de dados.

2.3 - Sistemas de Apoio à Decisão

Os SAD são sistemas computacionais que visam sistematizar e apoiar os processos decisórios empresariais, sendo comumente compostos pelos mesmos componentes básicos como: interface com o usuário, banco de dados e banco de modelos, visando fornecer apoio em decisões semi-estruturadas e desestruturadas, estando sob controle do tomador de decisões e disponibilizando um conjunto de ferramentas para estruturar e aumentar a efetividade das decisões (JUNIOR *et al.*, 2006).

SAD possuem funcionalidades e desempenho diferente dos Sistemas de Produção e Gestão existentes na Empresa.

Sistemas de Produção, como por exemplo, Controle de Estoque, geralmente recuperam e atualizam um registro por vez, ou seja, um cliente, um item de inventário, entre outros. Geralmente atendem a muitos usuários de forma concorrente e exigem também um tempo de resposta imediato.

Já os SAD normalmente lidam com poucos usuários por vez e os requisitos em termos de tempo de resposta não são críticos.

Os Sistemas de Gestão, mais conhecidos como ERP, servem como fonte primária dos dados que alimentarão o *DW*.

O principal objetivo destes sistemas é consolidar as informações da empresa como um todo dentro de um único sistema, eliminando assim a dificuldade de se obter informações consolidadas e a inconsistência de dados redundantes armazenados em mais de um sistema (FORTULAN & FILHO, 2005).

Os sistemas ERP, produzem grande impacto nas organizações, exigindo completa alteração na forma de agir diante desta nova realidade, caracterizada por mudanças radicais, o que exigirá a compreensão por parte de todos os profissionais de todas as áreas de atuação, bem como dirigir e compreender a empresa como um processo integrado (FREIRE & CAMPOS, 2007).

É aqui que aparecem os usuários de um *DW*, cuja função é descobrir as informações utilizadas no ambiente corporativo de tomada de decisões.

O usuário de um SAD age por meio de descobertas. Somente após receber a informação ele a julga. O mesmo tem um papel fundamental na constituição de um *DW*. No entanto, usualmente lidam com consultas complexas, não antecipadas ou previstas, envolvendo grande quantidade de registros básicos referentes aos processos operacionais da Empresa. Os SAD necessitam de dados consistentes, normalmente originários de mais de um Sistema de Produção, organizados de forma que favoreçam ser trabalhados por ferramentas de análise de dados.

Banco de Dados que dêem suporte aos SAD devem ser capazes de oferecer um bom tempo de resposta para consultas que recuperam grandes conjuntos de dados agregados e históricos. Geralmente lidam com tendências e não com um único instante de tempo, cada elemento de dado é acompanhado do correspondente período de tempo a que se refere.

Exemplo de Consultas típicas de um SAD:

- Listar o total de vendas dos últimos 5 anos;
- Listar os itens que estiverem sem estoque nos últimos 12 meses;
- Listar o fornecedor que não teve mais do que 20% de atraso nas últimas 150 entregas.

A importância em separar dados que dão suporte aos sistemas de caráter operacional da empresa daqueles que dão suporte aos processos gerenciais e de suporte à decisão é que cada tipo de aplicação pode se concentrar naquilo que faz melhor, oferecendo melhor funcionalidade e desempenho para seu caso específico (DAL'ALBA, 2007).

Após estas definições pode-se classificar as organizações em duas grandes aplicações:

Aplicações de Negócios:

- Garantem a operação da empresa, ou seja, rotinas do dia-a-dia.

Aplicações sobre os Negócios:

- São as aplicações que analisam o negócio, ajudando a interpretar o que aconteceu e a decidir sobre estratégias futuras para a empresa, compreendem os SAD.

Uma arquitetura de dados adequada para dar suporte aos dois tipos de aplicações baseia-se em dois ambientes de banco de dados, os operacionais para dar suporte às aplicações do negócio e os de suporte à decisão *DW*, para dar suporte às aplicações sobre o negócio (DAL'ALBA, 2007).

Os SAD fornecem aos executivos diversas ferramentas de modelagem e análise com o objetivo de capacitar os usuários a solucionar problemas de forma integral, fornecendo informações obtidas de diversas fontes de dados da empresa (BUOSI, 2004).

Principais características dos SAD:

- Foco nas decisões;
- Ênfase na flexibilidade, adaptabilidade e capacidade de fornecer respostas rápidas;
- Iniciado e controlado pelo usuário;
- Suporte para estilos pessoais de tomada de decisão;
- Voltados para problemas menos estruturados.

Para o desenvolvimento de um SAD deve-se contar com a efetiva participação do decisor e do profissional da área de tecnologia. A interação entre estas duas categorias deve ser capaz de permitir mudanças rápidas e eficazes ao sistema (CLERICUZI *et al.*, 2006).

Para que o SAD auxilie no processo de tomada de decisão é preciso que o mesmo seja composto de uma base de dados que auxilie o sistema, uma base de modelos que provê a capacidade de análise e o diálogo que provê a interação entre o usuário e o sistema (DUARTE *et al.*, 2005).

A quantidade de informações a respeito da situação constitui um fator de extrema importância para o processo decisório. Quanto mais informações forem levantadas a respeito da mesma, evidentemente, melhor será a sua compreensão e a sua solução (BISPO & CAZARINI, 1998).

É imprescindível que as decisões em uma organização sejam tomadas levando-se em consideração os objetivos estratégicos macros da empresa. O uso de informações confiáveis e sistemas que auxiliam este tipo de informação tornaram-se uma questão central ao sucesso organizacional, onde todos os investimentos do negócio nesta área devem estar sempre alinhados com a tecnologia de informação desenvolvida e também com o armazenamento de seus dados em bancos de dados consistentes e capazes de unificar as informações como um todo (AQUARONI & CAZARINI, 2001).

Para o gerente de uma grande e diversificada empresa, o SAD possibilita várias formas de exame das atividades desta. A tentativa de acompanhar um grande número de atividades é muito mais fácil do que tentar acompanhar apenas umas poucas atividades. Dessa maneira, o SAD pode ser usado para expandir o escopo de controle de um gerente.

Mas a análise de tendências e comparações não são os únicos meios de que o gerente dispõe para usar o SAD de forma eficiente. Um outro enfoque é o de “fatiar e agrupar”. Por meio dele, o analista obtém informações básicas que ele agrupa de um modo e analisa, agrupando-as, depois, de outro modo, e tornando a analisá-las.

O método de fatiar e agrupar permite que o gerente disponha de muitas perspectivas diferentes das atividades que estão ocorrendo (INMON, 1997).

Análise Prospectiva (Drill-Down)

Para usar o método de fatiar e agrupar, é necessário ter condições de “prospectar” dados. A prospecção de dados refere-se à capacidade de começar por um número resumido e partir esse resumo em conjuntos de resumos sucessivamente mais refinados.

Sendo capaz de chegar ao detalhe por trás do número resumido, o gerente pode ter uma noção sobre o que está acontecendo, especialmente quando o número resumido é surpreendente. A Figura 4 mostra um exemplo simples de análise prospectiva.

Figura 4 - Processo Prospectivo, adaptado de (INMON, 1997).

Na figura 4, o gerente analisa os resultados resumidos do segundo trimestre e deseja explorá-los ainda mais. O gerente, então, examina as regiões que contribuíram para a análise resumida. As regiões que tiveram seus números analisados foram as regiões Oeste, Sudeste, Nordeste e Central. Ao observar os números de cada região, o gerente decide fazer um exame detalhado dos números da região Sudeste.

Os números da região Sudeste são compostos pelos totais de São Paulo, Minas Gerais, Rio de Janeiro e Espírito Santo. Dentre esses estados, o gerente decide olhar mais de perto os números do estado de São Paulo. As diferentes cidades do estado de São Paulo que possuem postos de venda são submetidas a consultas.

Em cada caso, o gerente selecionou um caminho de passagem do resumo ao detalhe, e a seguir, para um nível sucessivamente mais baixo. Desse modo, ele pode determinar onde se encontram os resultados problemáticos. Tendo identificado as anomalias, o gerente sabe então onde procurar por uma explicação detalhada.

Há toda uma gama de softwares muito sofisticados que podem ser utilizados em um SAD para apresentarem resultados a um gerente.

A parte difícil do SAD não está na apresentação gráfica, e sim em descobrir e preparar os números que constam nos gráficos (INMON, 1997).

Do ponto de vista gráfico, o SAD é perfeitamente capaz de suportar o processo de prospecção, desde que, antes de tudo, os dados existam. Mas se os dados a analisar não existirem, o processo prospectivo se torna entediante e difícil, certamente diferente de algo que o executivo queira fazer.

2.4 - Considerações Finais do Capítulo

Os sistemas de informações devem ser configurados de forma a atender eficientemente as necessidades informativas de seus usuários, bem como incorporar seus conceitos, políticas e procedimentos que motivem e estimulem o gestor a tomar as melhores decisões para a empresa. As informações para gestão devem ter um sentido lógico e, isto se torna possível com o auxílio da Tecnologia da Informação.

A maior competitividade tem imposto uma atenção crescente na forma de gerir o conhecimento nas organizações. Visto como competência essencial e vantagem competitiva sustentável da organização, o conhecimento empresarial e sua gestão têm alterado os paradigmas gerenciais e demandado uma postura gerencial mais preocupada com este ativo intangível (PURCIDONIO *et al.*, 2006).

Baseando-se em BISPO (1998), INMON (2000), KIMBALL (1998), MEIRA (2004), SEIXAS (2000), e SINGH (2001), um SAD pode ser considerado como um ambiente formado por arquiteturas capazes de apresentarem informações relevantes e seguras sobre um dado problema utilizando-se de modelos matemáticos e estatísticos e uma base de conhecimento robusta, permitindo ao usuário final através do uso da tecnologia da informação a flexibilidade e interatividade para a utilização de simulações e cenários, e um tomador de decisão capaz de entender o que é apresentado pelo sistema, e a partir disso discorrer sobre a melhor escolha desde o ponto de partida até a execução e resultado da decisão tomada.

O SAD podem ser implementados de forma rápida e eficiente e levam a uma melhor documentação dos procedimentos, sua regularidade e uniformidade, bem como a ampla possibilidade de simulação de cenários em condições diversas (LACERDA & AMARO, 2007).

Ter informações em mãos é uma poderosa ferramenta para quem precisa tomar decisões e, tendo em vista esse princípio, as empresas começaram a extrair dados dos seus sistemas operacionais e armazená-los, separados dos dados operacionais, donde surgiram os SAD (FORTULAN & FILHO, 2005).

A importância e utilização de um SAD no ambiente organizacional moderno passa a ser um recurso estratégico decisivo na obtenção e manutenção das vantagens competitivas (MACCARI & SAUAIA, 2006).

No próximo capítulo, será demonstrado todo o ambiente de *DW*. Conceitos, principais características e arquitetura.

3 - Data Warehouse

Os Bancos de dados operacionais estão disponíveis há bastante tempo, mas não se preocupam com as questões pertinentes aos tomadores de decisão. Baseado nesta ausência surgiu uma oportunidade para o desenvolvimento da tecnologia *DW*, que é capaz de suprir a necessidade de gerenciamento da informação com o uso dos bancos de dados como suporte a decisão em uma organização.

3.1 - Conceito

O *DW* teve suas origens na necessidade de integrar dados provenientes de diversas origens e também na necessidade de gerenciar um grande volume de dados. Sua utilização como suporte para a administração de processos possibilita uma mudança da configuração de geração e disponibilização dos dados (FAVARETTO, 2007).

O *DW* pode ser considerado como um enorme repositório de dados central, onde a informação armazenada se encontra de maneira integrada. Esta informação é utilizada como suporte no processo de tomada de decisão dentro das organizações (RIVAS *et al.*, 2007).

Segundo INMON (1997), *DW* é um conjunto de dados orientados por assuntos, não volátil, variável com o tempo e integrado, criado para dar suporte à decisão.

Um *DW* é um grande banco de dados, elaborado com a finalidade de dar suporte ao processo decisório estratégico da empresa, cujos dados que o povoarão são obtidos dos bancos de dados dos aplicativos das áreas operacionais da empresa (DAL'ALBA, 2007).

A tecnologia de *DW* surge com a mesma proposta de otimização do uso da informação, a fim de transformá-la em diferencial competitivo. Desponta como uma das principais arquiteturas. Foi desenvolvida com a finalidade de prover suporte à tomada de decisão, tendo como elementos básicos: orientação para o objeto, integração, tempo como variável e a não-volatilidade de dados (TARAPANOFF *et al.*, 1995).

O uso do *DW* nas empresas nos últimos anos aumentou consideravelmente e agora tem papel fundamental nos processos de tomada de decisão, sendo que algumas organizações requerem dados dos ambientes transacionais em tempo real ou em períodos mais curtos, o que não ocorria anteriormente (ITALIANO & FERREIRA, 2006).

De modo geral, o funcionamento do *DW* é simples de entender. Uma das premissas do sistema é a integração de dados. Dados de várias fontes são coletados e migrados para o ambiente do *DW*, onde recebem um tratamento visando à sua padronização, que facilitará a recuperação de informações (os dados já sofreram interferência, já possuem valor agregado, então são considerados informação) pelo usuário final (TARAPANOFF *et al.*, 2000).

A Definição de *DW* varia de autor para autor, indo desde a informação armazenada num banco de dados de suporte a decisão até o processo de modelagem, extração de dados operacionais e armazenamento num banco de dados.

No entanto, apesar dessa variação na definição de um *DW*, existe um consenso com relação aos objetivos de implementá-lo, isto é, prover aos usuários finais fácil acesso a dados íntegros e consistentes para tomada de decisões nos negócios.

O escopo dessa tomada de decisão pode ser tático (Redução dos Custos), operacional (Melhoria dos Processos na Organização), estratégico (Lançamento de Novos Produtos no Mercado) e mais amplo (Mudança dos Negócios da Organização).

A tomada de decisão baseada em fatos é um processo estratégico para qualquer organização. Este processo é auxiliado pelo modelo através da identificação dos aspectos organizacionais que necessitam de correções ou melhorias para que a adoção do *DW* cumpra completamente seu papel: Disponibilizar fatos corretos e coerentes para auxiliar a tomada de decisão de gerentes e administradores da organização (CORREA *et al.*, 2007).

Os *DW* são de muita importância para os negócios de uma organização conforme citado nos exemplos anteriores, porém, vários projetos de *DW* falham por exigências de usuário irreais ou inacessíveis, falta de qualidade nas fontes de dados, fontes variáveis ou exigências de objetivo (BARRET & BARTON, 2006).

O apoio da alta administração é essencial, principalmente no que se refere ao patrocínio e investimento no projeto (SANTOS & MORAES, 2007).

Os Sistemas que utilizam a tecnologia de *DW* revitalizam os sistemas da empresa, pois:

- Permitem que sistemas mais antigos continuem em operação;
- Consolidam dados inconsistentes dos sistemas mais antigos em conjuntos coerentes;
- Extraem benefícios de novas informações oriundas das operações correntes;
- Provem ambientes para o planejamento e arquitetura de novos sistemas de cunho operacional.

Deve-se considerar, no entanto, que um *DW* não contém apenas dados resumidos, podendo conter também dados primitivos. É desejável prover ao usuário a capacidade de se aprofundar num determinado tópico, investigando níveis de agregação menores ou mesmo o dado primitivo, permitindo também geração de novas agregações ou correlações com outras variáveis.

Além do mais, é extremamente difícil prever todos os possíveis dados resumidos que serão necessários. Limitar o conteúdo de um *DW* apenas a dados resumidos significa limitar os usuários apenas as consultas e análises que eles puderem antecipar frente seus requisitos atuais, não deixando qualquer flexibilidade para novas necessidades (KIMBALL, 1998).

Ao consolidar e racionalizar as informações dispersas pelos banco de dados operacionais da empresa, o *DW* proporciona condições de análise estratégicas eficazes, com o objetivo de atender as necessidades de análise de informações dos usuários tomadores de decisão.

No cenário atual, onde os tomadores de decisão necessitam tomar decisões rápidas em respostas às mudanças que ocorrem frequentemente em seus negócios, faz-se necessário um sistema informatizado que dê suporte às suas decisões, levando informações rápidas e confiáveis em formato adequado e de fácil compreensão (FREIRE *et al.*, 2001).

Somente a partir dos dados limpos, agregados e disponíveis dentro de um *DW*, é que se realizam as análises e extrações de informações do processo que os gerou, explicam (GRIGORI *et al.*, 2004).

3.2 - Principais Características de um *DW*

Orientado por Assuntos

Refere-se ao fato do *DW* armazenar informações sobre assuntos específicos importantes para o negócio da empresa. Exemplos típicos de assuntos são produtos, atividades, contas, clientes, entre outros. Em contrapartida, o ambiente operacional é organizado por aplicações funcionais. Por exemplo, em uma organização bancária estas aplicações incluem empréstimos, investimentos e seguros.

A implementação de um assunto pode corresponder a um conjunto de tabelas relacionadas.

Por exemplo, considerando informações sobre vendas de funcionários, podem existir tabelas contendo informações básicas dos funcionários (como código do funcionário, nome, endereço, sexo, data início, data fim, entre outras), uma com dados do período 1948 a 1980, outra com dados para o período 1985 a 1990.

Além destas, existem tabelas cumulativas intermediárias com as atividades dos funcionários entre 1980 e 1990, contendo registro resumo para as atividades de cada mês (contendo código do funcionário, mês, número de transações, média de vendas, total da menor venda, total da venda de maio, total das vendas canceladas, entre outras). E finalmente, encontram-se ainda tabelas detalhadas de atividades para os períodos 1987 a 1988 e 1998 a 1990 (incluindo código do funcionário, data da atividade, número da nota, número do pedido, quantia, cliente id, local, entre outros). Existem, portanto, para o mesmo tipo de informação, diferentes níveis de detalhe e sumarização.

Todas estas tabelas contêm um identificador comum, o código do funcionário, além de um elemento temporal como parte da chave de cada tabela.

Nem sempre todas estas tabelas seriam mantidas em discos, sendo possível que, em alguns casos, as informações mais detalhadas das atividades dos vendedores fossem mantidas em fita magnética, ficando acessíveis apenas quando solicitadas (INMON, 1997). A figura 5 mostra um comparativo entre o Ambiente Operacional e o Ambiente de *DW*.

Figura 5 - Ambiente Operacional x Ambiente de DW, adaptado de INMON (1997).

Não Volátil

Significa que o DW permite apenas a carga inicial dos dados e consultas a estes dados, o chamado ambiente *load-and-access* (carregamento e acesso) conforme mostra a figura 6. Após serem integrados e transformados, os dados são carregados em bloco para o DW, para que estejam disponíveis aos usuários para acesso aos dados. No ambiente operacional, ao contrário, os dados são, em geral, atualizados registro a registro, em múltiplas transações. Essa volatilidade requer um trabalho considerável para assegurar integridade e consistência através de atividades de *rollback* (voltar lista), recuperação de falhas, *commits* (efetivar transação) e bloqueios. Um DW não requer este grau de controle típico dos sistemas orientados a transações (INMON, 1997).

Figura 6 - A questão da Não-Volatilidade. (Fonte: INMON, 1997)

Variante no Tempo

Refere-se ao fato do dado em um *DW* referir-se a algum momento específico, significando que ele não é atualizável, enquanto que o dado de produção é atualizado de acordo com as mudanças de estado do objeto em questão, refletindo, em geral, o estado do objeto no momento do acesso, conforme mostra a figura 7. Em um *DW*, a cada ocorrência de uma mudança, uma nova entrada é criada, para marcar esta mudança.

O tratamento de séries temporais apresenta características específicas, que adicionam complexidade a ambiente do *DW*. Processamentos mensais ou anuais são simples, mais dias e meses oferecem dificuldades pelas variações encontradas no número de dados, deve-se considerar que não apenas os dados têm uma característica temporal, mas também os metadados, que incluem definições dos itens de dados, rotinas de validação, algoritmos de derivação, entre outras. Sem a manutenção do histórico dos metadados, as mudanças das regras de negócio que afetam os dados no *DW* são perdidas, invalidando dados históricos (INMON, 1997).

A atualização do repositório de dados é um processo complexo, e pode ser considerado como uma das tarefas mais significativas na operação de um *DW* (MANNINO & WALTER, 2006).

Anteriormente os dados dos ambientes transacionais eram carregados geralmente quando os usuários não estavam operando o sistema, ou em intervalos de tempo maiores, porém agora em muitos casos esta atualização está ocorrendo várias vezes em apenas um dia (RAHMAN, 2007).

Figura 7 - Variável no tempo adaptado de INMON (1997).

Integrado

Refere-se a consistência de nomes, das unidades das variáveis, no sentido de que os dados foram transformados até um estado uniforme.

Por exemplo, considere-se sexo como um elemento de dado. Uma aplicação pode codificar sexo como M/F, outra como 1/0 e uma terceira como H/M. Conforme os dados são trazidos para o *DW*, eles são convertidos para um estado uniforme, ou seja, sexo é codificado apenas de uma forma.

Da mesma maneira, se um elemento de dado é medido em centímetros em uma aplicação, em polegadas em outra, ele será convertido para uma representação única ao ser colocado no *DW* (INMON, 1997).

A Essência do conceito de *warehouse* é a integração dos dados de fontes discrepantes, para um repositório coerente de informações. Alcançando um nível satisfatório de integração por dados internos e externos é o primeiro e, talvez, o desafio dianteiro que está em frente de um desenhista de *DW* (MARCH & REVNER, 2007).

A figura 8 mostra como é feita a integração durante a migração para o *DW*.

Figura 8 - Integração dos Dados no Ambiente de *DW*. (Fonte: INMON, 1997)

3.3 - Granularidade

O mais importante aspecto do projeto de um *DW* é a questão da granularidade. A granularidade diz respeito ao nível de detalhes ou de resumo contido nas unidades de dados existentes no *DW*. Quanto mais detalhes, mais baixo o nível de granularidade. Quanto menos detalhes, mais alto o nível de granularidade.

A granularidade de dados tem uma questão delicada de projeto. Nos primeiros sistemas operacionais que foram criados, a granularidade era tida como certa. Quando os dados detalhados eram atualizados, era quase certo que eles seriam armazenados no nível mais baixo de granularidade. No entanto, no ambiente de *DW*, a granularidade não é um pressuposto (INMON, 1997).

A razão pela qual a granularidade é a principal questão de projeto consiste no fato de que ela afeta profundamente o volume de dados que residem no *DW* e, ao mesmo tempo, afeta o tipo de consulta que pode ser atendida.

Quando se tem um nível de granularidade muito alto o espaço em disco e o número de índices necessários se tornam bem menores, porém há uma correspondente diminuição da possibilidade de utilização dos dados para atender a consultas detalhadas (DAL'ALBA, 2007).

A Figura 9 apresenta uma das questões referentes a granularidade. No lado esquerdo há um baixo nível de granularidade.

Cada chamada telefônica é registrada em detalhe. No lado direito um nível mais alto de granularidade pode ser visto. Os dados apresentam as informações resumidas referentes a um cliente em um mês.

Há, portanto um bom motivo para a compactação de dados em um *DW*. Quando os dados são compactados ocorre uma economia incomum sobre o Dispositivo de Armazenamento de Acesso Direto (DASD) utilizado. O número de índices e recursos de processador necessários no tratamento dos dados também são afetados (INMON, 1997).

Figura 9 - Definindo o Nível de Granularidade. (Fonte: INMON, 1997)

Deve-se lembrar, porém que durante o processamento de SAD, como é comum no ambiente de *DW*, dificilmente um evento isolado é examinado.

É mais comum ocorrer à utilização de uma visão de conjunto dos dados. Isso significa que um grande número de registros é examinado.

Modelagem Dimensional

Fundamentalmente, este é um trabalho sobre modelagem dimensional e sobre como construir um *DW* dimensional e mantê-lo em funcionamento. Modelagem dimensional é um nome novo para uma técnica antiga usada para criar bancos de dados simples e compreensivos. Quando um banco de dados pode ser visualizado como um “cubo” contendo três, quatro ou até cinco ou mais dimensões, as pessoas conseguem fatiar esse cubo em qualquer de suas dimensões. Da mesma forma, esses modelos podem ser facilmente entendidos e navegados pelo software tornando a interface de usuário simples e o desempenho da consulta aceitável.

O modelo dimensional é composto basicamente por dois tipos de tabelas: as de Fato e as de Dimensão. As primeiras são grandes tabelas centrais, compostas basicamente das ocorrências do negócio, por exemplo, vendas, produção e defeitos. Já as segundas armazenam as descrições do negócio, como dados sobre o produto (marca, tamanho, categoria, preço), tempo (dia, mês, ano) ou cliente (nome, endereço, classe social). Cada tabela de Dimensão possui uma única chave primária, e o conjunto dessas chaves primárias formará a chave composta da tabela de Fato. Cada negócio ou assunto possui sua tabela de Fato e suas respectivas tabelas de Dimensão (FORTULAN & FILHO, 2005).

A modelagem dimensional permite-nos visualizar dados. A habilidade de visualizar algo tão abstrato quanto um conjunto de dados de forma concreta e tangível é o segredo da compreensibilidade.

Muitas pessoas consideram fácil imaginar esse negócio como um cubo de dados, com legendas em cada uma das arestas do cubo, como mostra a Figura 10. Qualquer ponto do interior do cubo está na intersecção das coordenadas definidas pelas arestas do cubo. Para o negócio descrito anteriormente, legendamos as margens do cubo como *Produto*, *Mercado* e *Tempo*.

A maioria das pessoas consegue imaginar que os pontos dentro do cubo representam o local em que as medições do negócio para aquela combinação de Produto, Mercado e Tempo estão armazenadas. Esse é o modelo dimensional (KIMBALL, 1998).

A primeira vista este modelo parece bem simples. Um modelo de dados simples no seu princípio tem a chance de permanecer simples até o final do projeto. Um modelo de negócio complicado de início certamente será complicado no final.

Figura 10 - Modelo dimensional de um negócio: Produto, Mercado e Tempo.
(Fonte: KIMBALL, 1998).

Comparação Entre o Modelo Dimensional e um Modelo de Dependência de Dados

Conforme a figura 11, analisando o mesmo negócio, os diagramas de entidades de dados detalhados descrevem como cada item de uma fatura relaciona-se com todos os outros itens e quais são os relacionamentos muitos-para-muitos e muitos-para-um entre os elementos de dados (KIMBALL, 1998).

Essa figura certamente revela mais detalhes sobre os relacionamentos entre dados do que a figura dimensional, porém ela não contribui para o entendimento do negócio. Infelizmente, a maioria das pessoas não consegue memorizar um diagrama como esse e não consegue entender como navegá-lo de forma útil.

O diagrama de detalhes transmite uma falsa sensação de segurança. “Se é detalhado, deve ser bom”.

A verdade é que os relacionamentos entre dados são melhor visualizados quando exibidos dinamicamente em uma tela, e não em um mapa estático que os usuários tentam gravar em suas mentes.

O modelo dimensional de negócio e o modelo de dependência de dados são capazes de armazenar exatamente os mesmos dados e de suportar exatamente a mesma análise final do negócio. Nesse caso, estamos apenas escolhendo uma apresentação diferente dos dados.

O modelo dimensional é um modelo *top-down* (de cima para baixo) e o modelo de dependência de dados é do tipo *bottom-up* (de baixo para cima).

Figura 11 - Modelo de dependência de dados de um negócio.

(Fonte: KIMBALL, 1998)

A simplicidade do modelo dimensional é a chave fundamental que permite aos usuários entender o funcionamento de um banco de dados. A diferença entre o modelo dimensional e o modelo de dependência de dados está na estrutura do projeto do *DW*. A seguir será demonstrada uma figura com o modelo de relatório ideal para um *DW*.

Produto	Região	Vendas no Mês	Crescimento nas Vendas versus Mês Anterior	Vendas como % da Categoria	Mudança nas Vendas como % da Categoria versus Mês Anterior	Mudança nas Vendas como % da Categoria versus Ano Anterior
Açúcar	Central	110	12%	31%	3%	7%
Açúcar	Leste	179	-(3%)	28%	-(1%)	3%
Açúcar	Oeste	55	5%	44%	1%	5%
<i>Total Açúcar</i>	344	6%	33%	1%	5%	
Café	Central	66	2%	18%	2%	10%
Café	Leste	102	4%	12%	5%	13%
Café	Oeste	39	-(9%)	9%	-(1%)	8%
<i>Total Café</i>	207	1%	13%	4%	11%	
Total Geral		551	4%	20%	2%	8%

Figura 12 - O relatório ideal de um *DW*, adaptado de (KIMBALL, 1998).

A Dimensão Tempo

Os sistemas OLTP e *DW* tratam o tempo de forma diferente. O melhor sistema OLTP é um status instantâneo dos negócios de uma organização, atualizado constantemente à medida que as transações são concretizadas. Os valores-chave do negócio devem mudar a cada minuto ou segundo. O status muda continuamente e os relacionamentos entre as entidades são alterados. Por exemplo, à medida que os pedidos são atendidos, a relação de pedidos a atender muda e os relacionamentos de fornecedores de peças podem estar mudando. Chamamos esse tipo de banco de dados OLTP de *twinkling database* (banco de dados instantâneo).

Embora alguns aplicativos e usuários executem consultas do tipo *DW* em um *twinkling database*, a maioria deles prefere não usar esse recurso. Muitos usuários não gostam da idéia de que o banco de dados possa sofrer alterações durante ou entre as consultas. Essa natureza instantânea dos bancos de dados OLTP é o primeiro tipo de inconsistência temporal que evitamos em um *DW*.

O segundo tipo de inconsistência temporal em um banco de dados OLTP é a falta de suporte explícito para representar corretamente um histórico passado. Em um *DW* costumamos perguntar “Como se comportou meu negócio no mês passado? ou no trimestre passado ou no ano passado?”.

Com essas perguntas desejamos saber quais eram os pedidos a atender e os relacionamentos de fornecedores de peças em períodos de tempo anteriores. Se atendermos os pedidos, ou se eliminamos os relacionamentos antigos de fornecedores de peças, possivelmente não conseguiremos recuperar um retrato preciso dos pontos do passado. Embora seja possível manter um histórico em um sistema OLTP, é muito difícil retratar um histórico antigo correto. Se tudo que tivermos for uma longa seqüência de transações que alteram o histórico de forma incremental, será quase impossível reconstruir rapidamente um retrato de um negócio em um ponto de tempo específico (KIMBALL, 1998).

Lidamos com esses problemas no *DW* tornando o armazenamento uma seqüência de tempo explícita. Transferimos os instantâneos do OLTP para o *DW* como uma seqüência de camadas de dados, muito semelhante às camadas nas formações geológicas. Cada *DW* é uma seqüência de tempos. Assim como os geólogos, escavamos as camadas para saber como era nosso negócio no passado.

Ao incluirmos instantâneos estáticos (*static snapshots*) no *DW* em intervalos regulares definidos, solucionamos os dois problemas de representação de tempo encontrados no sistema OLTP. Primeiro, o *DW* não se modifica durante o dia quando os usuários estão fazendo suas consultas. Alterações instantâneas não são permitidas. Segundo, armazenando as informações cuidadosamente em cada instantâneo no *DW*, podemos representar todos os postos de tempo anteriores corretamente. Agora, podemos perguntar qual a diferença entre média de pedidos a atender registrada no ano passado e neste ano e podemos perguntar quantos fornecedores supriram uma peça em um determinado ponto de tempo anterior.

O modelo que consiste em tirar instantâneo regulares do sistema OLTP e, a seguir, transferi-los para o *DW* representa uma forma extremamente importante de se pensar em *DW*.

O instantâneo é chamado de extrato de dados de produção (*production data extract*), e migramos esses dados para o sistema do *DW* ao mesmo tempo diariamente, semanalmente ou mensalmente, como parte da carga de dados de produção (KIMBALL, 1998).

3.4 - O Modelo Dimensional

Os departamentos de Sistema de Informação (SI), usuários finais e até fornecedores de pesquisas de mercado, gravitaram para uma estrutura simples de “cubo de dados” que atende às necessidades de simplicidade do usuário final. Essa estrutura é um modelo dimensional. A Figura 13 mostra um modelo dimensional típico.

Outro nome para modelo dimensional é *star schema* (modelo estrela). Os projetistas de bancos de dados têm utilizado esse nome já há algum tempo para descrever modelos dimensionais porque o diagrama é semelhante a uma estrela com uma tabela grande no centro rodeada por tabelas auxiliares exibidas em um padrão radial.

Ao contrário do modelo entidade / relacionamento, o modelo dimensional é muito assimétrico. Há uma tabela dominante no centro do diagrama com múltiplas junções conectando-a às outras tabelas.

Cada uma das tabelas secundárias possui apenas uma junção com a tabela central. Chamaremos a tabela central de tabela de fatos e as outras tabelas de tabelas de dimensão.

Figura 13 - Um modelo dimensional típico. (Fonte: KIMBALL, 1998)

Na Figura 13 estamos modelando um negócio simples que vende produtos em algumas lojas e mede seu desempenho ao longo do tempo. Como projetista de banco de dados, decidimos que a tabela de fatos da Figura 13 contém totais diários de itens de todos os produtos vendidos. Chamamos isso de grão da tabela. Em outras palavras, cada registro da tabela de fatos representa o total de vendas de um produto específico por dia em uma loja. Qualquer outra combinação de produto, mercado ou dia gera um registro diferente na tabela de fatos.

Em um negocio de grande porte, certamente teremos muitos registros na tabela de fatos. A tabela de fatos pode chegar a um bilhão de linhas. Entretanto, usando um servidor moderno de alto desempenho e qualquer dos vários e poderosos softwares de banco de dados relacional, podemos armazenar e consultar uma tabela de fatos extensa como essa, com bom desempenho.

Tabela de Fatos

A tabela de fatos armazena medições numéricas do negócio. Cada uma das medições é obtida na intersecção de todas as dimensões. Na Figura 13 as medidas numéricas consistem no número de reais vendidos, no número de unidades vendidas e no custo estendido (custo unitário x quantidade). O custo representa quanto gastamos, como varejistas, para colocar o produto na prateleira. Imaginemos observando os produtos ao serem vendidos no mercado. Anotamos o número de reais vendidos, o número de unidades vendidas e o custo estendido todos os dias, em cada loja, para cada produto. Os fatos melhores e mais úteis são numéricos, continuamente valorados (diferentes na cada medida) e aditivos (podem ser adicionado às diversas dimensões). O motivo para a atualização de fatos continuamente valorados e aditivos é que em praticamente todas as consultas feitas a essa tabela de fatos, solicitaremos que sejam usadas centenas, milhares e até milhões de registros para o Sistema de Gerenciamento de Banco de Dados (DBMS) construir o conjunto de resposta. Esse grande número de registros será compactado em algumas dezenas de linhas para produzir o conjunto de resposta do usuário. A única forma viável de compactá-los no conjunto de resposta será adicioná-los. Portanto, se as medições forem números e se forem aditivas, poderemos construir facilmente o conjunto de resposta (KIMBALL, 1998).

O ideal é que os fatos sejam continuamente valorados principalmente para ajudar ao projetista de banco de dados a distinguir um fato de um atributo dimensional. O fato número de reais é continuamente valorado nesse caso porque pode usar praticamente qualquer valor de uma faixa abrangente de valores. Na qualidade de observadores, devemos aguardar que as medições nos sejam fornecidas, para que possamos ter uma idéia de qual será seu valor. Nesse sentido, o número de reais é continuamente valorado. Observe que nesse banco de dados, se não houvesse atividade de produto em um determinado dia, em uma loja, não incluiremos o registro no banco de dados. É muito importante que não tentemos preencher a tabela de fatos com zeros, para representar que “nada aconteceu”. Por esse motivo as tabelas de fatos são sempre esparsas. A maioria das tabelas de fatos é extremamente esparsa. Pode parecer óbvio que os fatos de reais, unidades e custo nesse exemplo sejam aditivos.

Tudo o que isso significa é que faz sentido adicionar os reais e outras medições a cada combinação de tempo, produto, loja e promoção.

Não importa qual fatia do banco de dados o usuário escolha, faz sentido adicionar os reais e as outras medições obtendo um total válido. Adicionabilidade, entretanto, é uma boa prática que devemos preservar.

Tabelas Dimensionais

As tabelas dimensionais armazenam as descrições textuais das dimensões do negócio. Cada uma dessas descrições textuais ajuda a definir um componente da respectiva dimensão. Por exemplo, cada registro da dimensão produto representa um produto específico. Em um banco de dados bem estruturado, a tabela de dimensão produto possui muitos atributos (campos). Os melhores atributos são textuais, discretos, e usados como fonte de restrições e cabeçalhos de linha no conjunto de resposta do usuário.

Como os atributos destinam-se a descrever os itens de uma dimensão, são mais úteis quando no formato de texto. Os atributos típicos de produto incluem uma descrição sucinta (de 10 a 15 caracteres), uma descrição longa (de 30 a 60 caracteres), a marca, a categoria, o tipo de embalagem e o tamanho do produto.

Embora o tamanho provavelmente seja expresso em números, ele ainda é considerado um atributo de dimensão útil porque se comporta mais como uma descrição textual do que como uma medição numérica. Tamanho de produtos são descrições distintas e constantes de um produto. Não observamos um mercado para medir o tamanho padrão de um produto.

Uma das funções-chave dos atributos de tabelas dimensionais é servir como fonte para restrições em uma consulta ou como cabeçalho de linha no conjunto de resposta do usuário. Um conjunto de resposta típico obtido em uma consulta tem a seguinte aparência:

Marca	Vendas em Reais	Unidades Vendidas
Basf	780	263
Maxell	1044	509
Skm	213	444
Smat	95	39

Essa consulta busca determinar todas as marcas de produtos vendidas no primeiro trimestre de 2007, além do total de vendas em reais e o número de unidades vendidas. Nesse banco de dados, uma marca é uma coleção de produtos individuais.

Para construir o relatório, o usuário arrastou o atributo Marca da dimensão produto e o posicionou como um cabeçalho de linha. A seguir, o usuário arrastou Vendas em Reais e Unidades Vendidas da tabela de fatos e os posicionou à direita do cabeçalho da linha Marca. Finalmente, o usuário especificou a restrição “1 T 2007” no atributo Trimestre da tabela da dimensão Tempo. Agora, a consulta do usuário está bem definida e a ferramenta de consulta pode gerar linhas de comandos SQL necessárias para produzir o conjunto de resposta. Utilizamos os atributos de dimensão para criar cabeçalhos de linhas e restrições.

O Modelo de Consulta Padrão

Embora programar em SQL não seja um pré-requisito essencial para projetar um *DW*, os comandos SQL gerados neste exemplo são tão característicos do uso de *DW* que será útil examinar seus componentes. O SQL para o exemplo anterior será o seguinte:

```
(1)  SELECT    P.MARCA,
              SUM ( F.REAIS )      AS VENDAS_EM_REAIS,
              SUM ( F.UNIDADES)    AS UNIDADES_VENDIDAS

(2)  FROM      FATO_VENDAS F , PRODUTO P , TEMPO T

(3)  WHERE     F.CHAVE_PRODUTO = P.CHAVE_PRODUTO AND
              F.CHAVE_TEMPO    = T.CHAVE_TEMPO AND

(4)          T.TRIMESTRE      = '1 T 2007'

(5)  GROUP BY  P.MARCA

(6)  ORDER BY  P.MARCA
```

Esse SQL pode ser considerado como um modelo padrão para todas as consultas de *DW* envolvendo a tabela de fatos.

Nosso modelo de SQL nos fornece uma visão detalhada de como uma consulta deve ser processada em um banco de dados dimensional. Primeiro as restrições são avaliadas, dimensão por dimensão. Cada dimensão produz um conjunto de chaves candidatas. As chaves candidatas são reunidas a seguir em chaves compostas que serão pesquisadas na tabela de fatos. Todos os “achados” da tabela de fatos são agrupados e totalizados de acordo com as especificações da seleção da lista e da cláusula *group by*. Embora cada DBMS use uma forma diferente para processar o SQL, esse modelo de processamento de consulta dimensional é muito mais importante do ponto de vista da perspectiva *back room* em um ambiente de *DW*. As consultas em *DW* que duram horas geralmente são aquelas em que o DBMS decidiu não processar as dimensões primeiro, mas em vez disso está tentando processar a tabela de fatos no início do processo. A maior parte do tempo é gasta com o *back room* tentando obrigar o DBMS a processar a consulta na ordem apropriada (KIMBALL, 1998).

3.5 - Data Mart

Muito utilizados em projetos-piloto, existem também os *Data Marts*, que, conforme SINGH (2001) define, são subconjuntos do *DW* da empresa inteira. Tipicamente desempenham o papel de um *DW* departamental, regional ou funcional.

Uma empresa pode construir uma série de *Data Marts* ao longo do tempo e, eventualmente, vinculá-los por meio de *DW* lógico da empresa inteira.

Os *Data Marts* são menores e de implantação mais rápida que um *DW*. Deve-se, no entanto, tomar cuidado com a proliferação dos *Data Marts* sem que se tenha por base um sistema que os consolide. Entenda-se aqui um *DW* corporativo, caso contrário, esses sistemas rapidamente estarão desintegrados e, junto com a desintegração, trarão varios problemas também (FORTULAN, 2006).

Segundo (COREY *et al.*, 2001), não há problemas em se começar pequeno, mas não se pode perder a visão do sistema maior, o *DW* corporativo.

Para a construção de um *Data Mart*, deve-se proceder da mesma forma como na construção de um *DW*: extraindo, transformando e integrando dados pertinentes à área departamental do *Data Mart* (CALDAS *et al.*, 2006).

Tabela 1 - Principais diferenças entre *DW* e *Data Mart*.

<i>Propriedades</i>	<i>Data Warehouse</i>	<i>Data Mart</i>
Âmbito	<i>aplicações neutras</i>	<i>aplicações específicas</i>
	<i>centralizado e compartilhado</i>	<i>departamento ou área de usuário</i>
Assuntos	<i>múltiplas áreas</i>	<i>área única</i>
Fontes de dados	<i>muitas</i>	<i>poucas</i>
Tempo de implementação	<i>9 a 18 meses para 1º estágio (duas ou três áreas de assuntos);</i>	<i>4 a 12 meses</i>
	<i>múltiplos estágios de implementação;</i>	
Características	<i>flexível</i>	<i>restritivo</i>
	<i>nível estratégico</i>	<i>nível tático</i>

Fonte: (DAL'ALBA, 2007)

3.6 - Arquitetura do *DW*

Para ser útil o *DW* deve ser capaz de responder a consultas avançadas de maneira rápida, sem deixar de mostrar detalhes relevantes às respostas. Para isso ele deve possuir uma arquitetura que lhe permita coletar, manipular e apresentar os dados de forma eficiente e rápida. Mas construir um *DW* eficiente, que servirá de suporte a decisões para empresas exige mais do que simplesmente descarregar ou copiar os dados dos sistemas atuais para um banco de dados maior. Deve-se considerar que os dados provenientes de vários sistemas podem conter redundância e diferenças, então antes de passá-los para o *DW* é necessário aplicar filtros sobre eles. O estudo de uma arquitetura permite compreender como o *DW* faz para armazenar, integrar, comunicar, processar e apresentar os dados que os usuários utilizarão em suas decisões. Um *DW* pode variar sua arquitetura conforme o tipo de assunto abordado, pois as necessidades também variam de empresa para empresa.

É possível definir uma arquitetura genérica onde praticamente todas as camadas necessárias são apresentadas, conforme a arquitetura genérica vista a seguir, ou arquitetura que utilizam somente algumas das camadas definidas, como as arquiteturas em duas e três camadas.

Arquitetura Genérica do *DW*

A seguir é descrita uma arquitetura genérica proposta por DAL'ALBA (2007), ilustrada na Figura 14. Esta descrição genérica procura apenas sistematizar papéis no ambiente de *DW*, permitindo que diferentes abordagens encontradas no mercado atualmente possam ser adaptadas a ela, deve-se considerar que esta arquitetura tem o objetivo de representar a funcionalidade de um *DW*, sendo que várias camadas propostas podem ser atendidas por um único componente de *software*.

Esta arquitetura é composta pela camada dos dados operacionais e outras fontes de dados que são acessados pela camada de acesso aos dados. As camadas de gerenciamento de processos, transporte e *DW* formam o centro da arquitetura e são elas as responsáveis por manter e distribuir os dados.

Figura 14 - Arquitetura Genérica de um *DW*. (Fonte: Adaptado de DAL'ALBA, 2007)

A camada de acesso à informação é formada por ferramentas que possibilitam os usuários extrair informações do *DW*. A seguir são demonstradas as principais camadas que compõem a arquitetura de um *DW*.

- **Camadas de bancos de dados operacionais e fontes externas:** É composta pelos dados dos sistemas operacionais das empresas e informações provenientes de fontes externas que serão integradas para compor o *DW*.
- **Camada de acesso à informação:** Envolve o *hardware* e o *software* utilizado para obtenção de relatórios, planilhas, gráficos e consultas. É nesta camada que os usuários finais interagem com o *DW*, utilizando ferramentas de manipulação, análise e apresentação dos dados, incluindo-se as ferramentas de *data mining* e visualização.
- **Camada de acesso aos dados:** Esta camada faz a ligação entre as ferramentas de acesso à informação e o banco de dados operacionais. Esta camada se comunica com diferentes sistemas de bancos de dados, sistemas de arquivos e fontes sob diferentes protocolos de comunicação, o que se chama acesso universal de dados.

- **Camada de metadados (Dicionário de Dados):** Metadados são as informações que descrevem os dados utilizados pela empresa, isto envolve informações como descrições de registros, comandos de criação de tabelas, diagramas Entidade/Relacionamento (E-R), dados de um dicionário de dados, entre outras. É necessário que exista uma grande variedade de metadados no ambiente de *DW* para que ele mantenha sua funcionalidade e os usuários não precisem se preocupar onde residem os dados ou a forma com que estão armazenados.
- **Camada de gerenciamento de processo:** É a camada responsável pelo gerenciamento dos processos que contribuem para manter o *DW* atualizado e consistente. Está envolvida com o controle das várias tarefas que devem ser realizadas para construir e manter as informações de dicionário de dados e do *DW*.
- **Camada de transporte:** Esta camada gerencia o transporte de informações pelo ambiente de rede. Inclui a coleta de mensagens e transações e se encarrega de entregá-las em locais e tempos determinados. Também é usada para isolar aplicações operacionais ou informacionais, do formato real dos dados nas duas extremidades.
- **Camada do *DW*:** É o *DW* propriamente dito, corresponde aos dados utilizados para obter informações. Às vezes o *DW* pode ser simplesmente uma visão lógica ou virtual dos dados, podendo não envolver o armazenamento dos mesmos ou armazenar dados operacionais e externos para facilitar seu acesso e manuseio.

3.7 - Considerações Finais do Capítulo

O mercado competitivo levou, nas últimas décadas, o investimento de muitas organizações em soluções capazes de transformar seus dados em um ativo capaz de prover inteligência ao seu negócio.

Com um *DW* e suas aplicações bem especificadas e implantadas os benefícios de seu uso deverão começar a surgir já nos primeiros meses após sua construção, porém, para que estes benefícios sejam alcançados de forma plena, será necessária uma mudança interna na organização para que seu uso possa entrar na rotina de trabalho de seus funcionários. Com o uso do sistema, novas análises e pesquisas deverão surgir, o que proporciona um crescimento exponencial dos benefícios, facilitando a análise dos gestores para a tomada de decisão levando-os ao crescimento de seu lucro e conseqüentemente a fidelização do cliente (CALDAS *et al.*, 2006).

Neste cenário, com o aprimoramento das ferramentas e a evolução na capacidade analítica dos gestores, tornaram essas soluções populares dentro das organizações. Essa situação, entretanto, pressiona as mesmas organizações por um melhor desempenho de suas soluções de Apoio à Decisão que se dilatam em volume de dados para atender a um público igualmente crescente e, em algumas situações, de perfil antagônico.

Pensar no desempenho do *DW* tem início na modelagem física e perdura durante toda a sua operação dentro de uma organização. Desta maneira, faz-se necessário um conjunto de ações coordenadas e institucionalizadas cujo objetivo é sustentar um patamar de desempenho adequado para a organização suportar seus processos de tomada de decisão.

Por envolver diferentes conhecimentos técnicos e do negócio, essas ações devem ser direcionadas a uma equipe de projetistas do *DW*, responsáveis pela infraestrutura que suporta o *DW*, que propicie sua institucionalização, suporte econômico e facilite seu gerenciamento por parte da organização.

No próximo capítulo, será demonstrado uma visão geral da agroindústria canavieira no Brasil.

Capítulo 4 - Agroindústria Canavieira no Brasil

A agroindústria canavieira do Brasil se diferencia dos outros países por produzir em escala industrial tanto açúcar quanto álcool e, mais recentemente, energia elétrica co-geradora a partir da queima da palha e do bagaço de cana. Portanto, há uma característica de aproveitamento múltiplo da cana-de-açúcar, cujos produtos intermediários e finais são dotados dessa cadeia produtiva. O tamanho, grau de integração entre agricultura e indústria, volume de produção e de exportação, além do poder político e econômico dos agentes envolvidos, dão peculiar característica à agroindústria canavieira brasileira (MORAES & SHIKIDA, 2002).

4.1 - Introdução

O açúcar é uma *commodity*¹ agrícola consumida e comercializada tanto no mercado interno quanto no externo. Embora o Brasil consuma, em média, em torno de 73% do açúcar produzido internamente, enquanto as exportações correspondem, em média, a aproximadamente 27% da produção nacional, o país tem se posicionado entre os maiores produtores e exportadores de açúcar do mundo, e sua oferta no mercado externo tem grande influência sobre os preços internacionais. O açúcar comumente não oferece grandes oportunidades de exploração de marcas, já que é um produto homogêneo. Contudo, já se verificam no setor algumas estratégias que procuram atender a nichos de mercado, como o do açúcar orgânico, embora esse mercado ainda restrito aos consumidores de mais alta renda, dispostos a pagar um preço maior pelo produto diferenciado.

Afora esses aspectos, os adoçantes sintéticos concorrem com vastas campanhas publicitárias com o açúcar, as quais destacam as vantagens desses produtos não calóricos para a manutenção da estética e do padrão de saúde de seus consumidores. Outros mercados como o norte-americano e o europeu, vêm utilizando xarope de frutose (obtido do milho) e açúcar de beterraba, substituindo assim, o açúcar de cana, tanto nos domicílios quanto como componente em alimentos e refrigerantes (MORAES & SHIKIDA, 2002).

¹ **Commodity** é um termo utilizado nas relações comerciais internacionais para designar um tipo particular de mercadoria em estado bruto ou produto primário de importância comercial. É o caso do açúcar, café, do algodão, da soja, de carnes e minérios (MACEDO, 2005).

Já o álcool é um produto que pode ser utilizado pela indústria química, farmacêutica, de cosméticos e de bebidas, mas seus fins carburantes, seja substituindo a gasolina (álcool hidratado) ou como aditivo (álcool anidro), dão-lhe uma peculiaridade e perspectiva somente observadas no mercado doméstico.

O Brasil é o maior produtor e consumidor mundial de álcool. Existem diversos problemas na comercialização do álcool combustível. Embora o mercado de distribuição de combustíveis tenha sido aberto em meados da década de 90, melhorando a estrutura de mercado oligopsonista existente, com a entrada de grande número de distribuidoras surgiram outros problemas (adulteração de combustíveis, sonegação fiscal, entre outras), que indicam a necessidade de maior atenção das exigências reguladoras e de aumento da fiscalização sobre esse mercado.

Novas oportunidades de negócio vêm surgindo para a agroindústria canavieira, tais como a co-geração de energia elétrica com base no bagaço de cana-de-açúcar e o nascente mercado de carbono. Além disso, o interesse de outros países pelo uso do álcool combustível pode ter consequências positivas importantes para esse mercado (MORAES & SHIKIDA, 2002).

A evolução deste setor nos últimos vinte anos e as perspectivas de seu crescimento para os próximos exigem um posicionamento muito diferente do tradicional produtor de açúcar, incluindo-se como produtor de energia e atuando muito mais nos mercados mundiais. Nestes vinte anos ocorreu também um extraordinário aumento do nosso conhecimento sobre as consequências da interação do homem com o meio ambiente e sobre as consequências sociais de ações políticas e econômicas em um mundo muito mais inter-relacionado (globalizado) (MACEDO, 2005).

4.2 - Impactos no Meio Ambiente

Os impactos no meio ambiente consideram a cultura da cana, o processamento industrial e o uso final. Incluem os efeitos na qualidade do ar e no clima global, no uso do solo e biodiversidade, na conservação do solo, nos recursos hídricos e o uso de defensivos e fertilizantes.

Estes impactos podem ser positivos ou negativos, em alguns casos a indústria da cana tem resultados muito importantes, como na redução de emissões de gases de efeito estufa (GEE) e na recuperação de solos agrícolas. A legislação ambiental (incluindo restrições ao uso do solo) é avançada no Brasil, com aplicação eficiente na cultura da cana.

A degradação da qualidade do ar em centros urbanos é um dos mais sérios problemas ambientais no mundo, em grande parte é causada pelo uso de combustíveis fósseis, que concorrem também para a poluição além das fronteiras: chuva ácida, por exemplo. Os esforços de mitigação incluem uma legislação cada vez mais restritiva para os combustíveis e sistemas de utilização.

A Agroindústria da cana relaciona-se com os impactos na qualidade do ar em dois pontos, e de forma muito distinta: o uso do etanol tem levado a melhorias consideráveis da qualidade do ar em centros urbanos e as queimadas de palha da cana no campo, em uma escala muito diferente, causam problemas com a dispersão de partículas e riscos com a fumaça (MACEDO, 2005).

4.3 - Características da Cadeia Produtiva

Diversas peculiaridades existentes nessa cadeia produtiva tornam muito difícil que o equilíbrio entre a oferta e demanda em ambiente de livre mercado seja atingido, e reforçam a necessidade do desenvolvimento de um novo modelo de gestão, que associe ações públicas e privadas, de modo a assegurar a estabilidade dos principais produtos: cana-de-açúcar, açúcar, álcool hidratado e álcool anidro e, mais recentemente, energia elétrica co-geradora pela queima do bagaço e da palha da cana.

Uma das principais características dessa cadeia produtiva, que não pode ser negligenciada, já que interfere na quantidade e qualidade da matéria-prima, é que o principal insumo, a cana-de-açúcar, é de origem agrícola e, como tal, está sujeita aos riscos climáticos, fitossanitários ², à sazonalidade da produção e a variação da renda dos produtores (MORAES & SHIKIDA, 2002).

² **Fitossanitários:** corresponde ao aspecto sanitário de um vegetal relativo à ocorrência de insetos-pragas e doenças (EMBRAPA, 2008).

A Sustentabilidade da base de produção agrícola da cana no Brasil deve incluir a capacidade de responder a pragas, doenças e a variações climáticas periódicas, sem que estas a prejudiquem seriamente.

As condições de produção no Brasil, com sua diversidade de regiões e microclimas, tem demonstrado resposta adequada às variações periódicas do clima (MACEDO, 2005).

A proteção contra pragas e doenças é considerada um ponto forte da produção no Brasil: é baseada muito mais no suprimento contínuo de variedade de cana resistentes do que em barreiras fitossanitárias, propiciando aos produtores operar com grande diversificação.

Mais de 500 variedades de cana são cultivadas hoje (51 liberadas nos últimos dez anos). As vinte principais ocupam 80% da área, mas a mais utilizada atinge apenas 12,6%. O aumento da diversificação nos últimos vinte anos promoveu grande segurança quanto a resistência contra doenças e pragas exógenas ³.

O Brasil se destaca na biotecnologia da cana, tendo variedades transgênicas (não comerciais) desde meados dos anos 1990. Foi concluída em 2003 a identificação de 40.000 genes da cana em laboratórios brasileiros, há duas dezenas de grupos trabalhando no genoma funcional e já usando genes em programas de melhoramento genético (fases experimentais). Resultados comerciais podem surgir nos próximos anos.

4.4 - Os Processos de Produção no Brasil

No Brasil a Cana-de-Açúcar é cultivada principalmente em grandes áreas no Nordeste e Centro-Sul, são usados cinco ou seis cortes antes da reforma do canavial, e o período de safra é de seis ou sete meses.

Todo o processo de produção é intensivo em mão-de-obra, especialmente a colheita. O avanço da mecanização tem reduzido o número de empregos (por unidade de produção) e também a sua sazonalidade (MACEDO, 2005).

³ **Doenças e pragas exógenas:** De acordo com o modelo de LAPLANTINE (2004), doenças e pragas são consideradas exógenas quando são compreendidas como o resultado de um elemento estranho gerado fora do corpo.

A cultura utiliza fertilizantes e defensivos agrícolas moderadamente e recicla todos os efluentes industriais da produção de etanol e açúcar como insumos para a lavoura.

A Figura 15 mostra uma visão macro dos processos agrícolas de produção que estão disponíveis em uma Usina que produz Açúcar e Álcool.

Figura 15 - Visão Macro dos Processos Agrícolas de uma Usina de Açúcar e Álcool.
(Fonte: USINA, 2007)

Preparo do Solo:

Tendo a cana-de-açúcar um sistema radicular profundo, um ciclo vegetativo econômico de quatro anos e meio ou mais e uma intensa mecanização que se processa durante esse longo tempo de permanência da cultura no terreno, o preparo do solo deve ser profundo e esmerado.

Convém salientar que as unidades sucroalcooleiras não seguem uma linha uniforme de preparo do solo, tendo cada uma seu sistema próprio, variação essa que ocorre em função do tipo de solo predominante e da disponibilidade de máquinas e implementos.

No preparo do solo, temos de considerar duas situações distintas:

1. a cana vai ser plantada pela primeira vez,
2. ou o terreno já se encontra ocupado com cana.

No primeiro caso, faz-se uma aração profunda, com bastante antecedência do plantio, visando à destruição, incorporação e decomposição dos restos culturais existentes, seguida de gradagem, com o objetivo de completar a primeira operação. Em solos argilosos é normal a existência de uma camada impermeável, a qual pode ser detectada pelas trincheiras abertas no perfil do solo, ou pelo penetrômetro.

Constatada a compactação do solo, seu rompimento se faz por intermédio de subsolagem, que só é aconselhada quando a camada adensada se localizar a uma profundidade entre 20 e 50 cm da superfície e com solo seco.

Nas vésperas do plantio, faz-se nova gradagem, visando ao acabamento do preparo do terreno e à eliminação de ervas daninhas.

Na segunda situação, onde a cultura da cana já se encontra instalada, o primeiro passo é a destruição da soqueira, que deve ser realizada logo após a colheita. Essa operação pode ser feita por meio de aração rasa (15-20 cm) nas linhas de cana, seguidas de gradagem ou pela gradagem pesada, enxada rotativa ou uso de herbicida.

Se confirmada a compactação do solo, a subsolagem torna-se necessária. Nas vésperas do plantio procede-se a uma aração profunda (25-30 cm), por meio de arado ou grade pesada. Seguem-se as gradagens necessárias, visando manter o terreno destorroado e apto ao plantio. Devido à facilidade de transporte, à menor regulagem e ao maior rendimento operacional, há uma tendência das grades pesadas substituírem o arado.

Plantio de Cana:

Existem duas épocas de plantio para a região Centro-Sul: setembro e outubro ou janeiro a março. Setembro e outubro não é a época mais recomendada, sendo indicada em casos de necessidade urgente de matéria prima, quer por recente instalação ou ampliação do setor industrial, quer por comprometimento de safra devido à ocorrência de adversidade climática. Plantios efetuados nessa época propiciam menor produtividade agrícola e expõem a lavoura à maior incidência de ervas daninhas, pragas, assoreamento dos sulcos e retardam a próxima colheita.

O plantio da cana de "ano e meio" é feito de janeiro a março, sendo o mais recomendado tecnicamente. Além de não apresentar os inconvenientes da outra época, permite um melhor aproveitamento do terreno com plantio de outras culturas. Em regiões quentes, como o oeste do Estado de São Paulo, essa época pode ser estendida para os meses subseqüentes, desde que haja umidade suficiente.

O espaçamento entre os sulcos de plantio é de 1,40 m, sua profundidade de 20 a 25 cm e a largura é proporcionada pela abertura das asas do sulcador num ângulo de 45°, com pequenas variações para mais ou para menos, dependendo da textura do solo.

Os colmos de cana ⁴ com idade de 10 a 12 meses são colocados no fundo do sulco, sempre cruzando a ponta do colmo anterior com o pé do seguinte e picados, com podão, em toletes de aproximadamente de três gemas.

A densidade do plantio é em torno de 12 gemas por metro linear de sulco, que, dependendo da variedade e do seu desenvolvimento vegetativo, corresponde a um gasto de 7 a 10 toneladas por hectare. Os toletes são cobertos com uma camada de terra de 7 cm, devendo ser ligeiramente compactada. Dependendo do tipo de solo e das condições climáticas reinantes, pode haver uma variação na espessura dessa camada.

Tratos Culturais

Os tratos culturais na cana-planta limitam-se apenas ao controle das ervas daninhas, adubação em cobertura e adoção de uma vigilância fitossanitária para controlar a incidência do carvão.

O período crítico da cultura, devido à concorrência de ervas daninhas, vai da emergência aos 90 dias de idade.

O controle mais eficiente às ervas, nesse período, é o químico, utilizando-se da aplicação de herbicidas em pré-emergência, logo após o plantio e em área total. Dependendo das condições de aplicação, infestação da gleba e eficiência do praguicida, há necessidade de uma ou mais carpas mecânicas e catação manual até o fechamento da lavoura. A partir daí a infestação de ervas é praticamente nula.

⁴ **Colmos de Cana:** tipo de caule comum às gramíneas, com nítida diferenciação dos nós e dos entrenós, como no bambu e na cana-de-açúcar (EMBRAPA, 2008).

Outro método é a combinação de carpas mecânicas e manuais. Instalada a cultura, após o surgimento do mato, procede-se seu controle mecanicamente, com o emprego de cultivadores de disco ou de enxadas junto às entrelinhas, sendo complementado com carpa manual nas linhas de plantio, evitando, assim, o assoreamento do sulco. Essa operação é repetida quantas vezes forem necessárias; normalmente três controles são suficientes.

Colheita de Cana

A colheita inicia-se em maio e em algumas unidades sucroalcooleiras em abril, prolongando-se até novembro, período em que a planta atinge o melhor ponto de maturação, devendo, sempre que possível, antecipar o fim da safra, por ser um período bastante chuvoso, que dificulta o transporte de matéria prima e faz cair o rendimento industrial.

A prática da queima do canavial antes da colheita (retirando as folhas, para facilitar o corte) está sendo gradualmente reduzida, com restrições ambientais e de segurança em algumas áreas, mas ainda é dominante (MACEDO, 2005).

Tratos Culturais (Soqueira)

As soqueiras exigem enleiramento do "paliço", permeabilização do solo, controle das ervas daninhas, adubação e vigilância sanitária. Os dois últimos tratos culturais encontram-se em itens próprios.

Após a colheita da cana, ficam no terreno restos de palha, folhas e pontas, cuja permanência prejudica a nova brotação e dificulta os tratos culturais. A maneira de eliminar esse material (paliço) seria a queima pelo fogo, porém essa prática não é indicada devido aos inconvenientes que ela acarreta, como falhas na brotação futura, perdas de umidade e matéria orgânica do solo e quebra do equilíbrio biológico.

O enleiramento consiste no amontoamento em uma rua do "paliço" deixando duas, quatro ou seis ruas livres, dependendo da quantidade desse material. É realizado por enleiradeira tipo Lely, implemento leve com pouca exigência de potência. Após a retirada da cana, o solo fica superficialmente compactado e impermeável à penetração de água, ar e fertilizantes. Visando à permeabilização do solo e controle das ervas daninhas iniciais, diversos métodos e implementos podem ser usados.

Existem no mercado implementos dotados de hastes semi-subsoladoras ou escarificadoras, adubadeiras e cultivadores que realizam simultaneamente, operações de escarificação, adubação, cultivo e preparo do terreno para receber a carpa química, exigindo, para tanto, tratores de aproximadamente 90 HPs. Normalmente, essa prática, conhecida como tríplice operação, seguida do cultivo químico, é suficiente para manter a soqueira no limpo.

Além desse sistema, o emprego de cultivadores ou enxadas rotativas com tração animal ou mecânica apresenta bons resultados. Devido ao rápido crescimento das soqueiras, o número de carpas exigido é menor que o da cana planta.

A seguir será demonstrado uma visão geral dos processos de produção em uma Usina que produz Açúcar e Alcool.

Figura 16 - Visão Geral dos Processos de uma Usina.

(Fonte: Usina, 2007)

Processos Industriais

Os colmos de cana são processados para produzir etanol e açúcar, parte da cana é lavada para retirar impurezas minerais (apenas a cana de colheita manual).

Um sistema de extração (no Brasil, a cana é picada, desfibrada e passa por uma série de moendas), separa o caldo, contendo a sacarose, da fibra (bagaço).

Para a produção de açúcar, o caldo é limpo (decantação e filtro prensa, retirando um resíduo, a torta de filtro), concentrado e cristalizado. Uma parte dos açúcares não cristalizado e impurezas do melaço são separados.

No Brasil o mel residual é em geral muito mais rico em açúcar, evitando-se o estágio final na cristalização e usando o mel, em mistura com o caldo, como insumo para a fermentação.

Essa mistura é levada à concentração adequada e fermentada com leveduras e o vinho resultante é destilado, produzindo álcool (hidratado ou anidro) e deixando como resíduo a vinhaça (água da cana adicionada na moagem, matéria orgânica, e minerais importantes, como o potássio, que vieram com a cana).

Toda a energia utilizada no acionamento de algumas bombas, ventiladores e moendas e nos processos de concentração do caldo e destilação, é suprida hoje por um sistema de co-geração que usa o bagaço e a palha como fonte energética. A usina é auto-suficiente, e em geral pode ter excedentes de energia.

Os processos industriais têm como resíduos a vinhaça, a torta de filtro e as cinzas da caldeira de bagaço. São totalmente reciclados para as lavouras. Nos processos industriais utilizam água (captada de rios e poços) em várias operações.

A vinhaça é utilizada na forma líquida, como fertirrigação, a torta de filtro é transportada por caminhões e distribuída como adubo e a água é re-utilizada, reduzindo a capacitação e o nível do despejo tratado.

4.5 - Cana Orgânica x Cana Convencional

Com a preocupação de preservar o meio ambiente e agregar valor aos seus produtos, algumas Usinas de Açúcar e Alcool, já estão utilizando o processamento de cana orgânica visando à produção do açúcar orgânico e derivados.

Para a produção do açúcar orgânico, é necessário mudar o processo produtivo convencional para outro que, por exemplo, não utilize adubos químicos e agrotóxicos e siga os procedimentos determinados para a produção de produtos orgânicos em geral. Dentro da unidade industrial também é importante evitar a contaminação do produto por sujeiras ou por substâncias não permitidas. A limpeza dos equipamentos é fundamental para garantir a assepsia e a qualidade orgânica do produto. Portanto, assim como os maquinários da lavoura, os equipamentos da indústria são limpos antes e depois do processo de produção.

A agricultura orgânica elimina o uso de fertilizantes, pesticidas e reguladores de crescimento produzidos sinteticamente. Seu sistema de produção baseia-se na rotação de culturas, esterco de animais, leguminosas, adubação verde, cultivos mecânicos, minerais naturais e controle biológico de pragas para manter a estrutura do solo, fornecer nutrientes para as plantas, controlar insetos, ervas daninhas e outras pragas.

A estrutura orgânica dá maior conforto à cana. Sem estresse hídrico, e com um equilíbrio maior de insetos, o que permite um combate natural às pragas e permite a planta se desenvolver melhor. A biodiversidade nos canaviais orgânicos é sete vezes maior do que nos convencionais. O meio convencional de produção, que, em grande parte, usa o fogo na hora da colheita, elimina esses benefícios e deixa o solo esterilizado. A diferença entre os dois manejos aparece na durabilidade do canavial.

Em alguns canaviais convencionais, as novas variedades de cana duram apenas um ciclo (cinco cortes). Na orgânica, às vezes chegam a quatro ou cinco ciclos. No período de seca, a produtividade dos canaviais orgânicos cai apenas metade da registrada pelos convencionais. Outro dado importante que foi constatado é que o quinto corte de cana tem uma produtividade maior do que o do quarto. A produção é maior por força da reorganização da vida e da estrutura do solo. Além do solo mais rico, há um equilíbrio natural no combate às pragas. Para cada praga potencial, há três predadores. Um caso típico é o da formiga saúva. De 1987 a 1992, a usina tinha 5 formigueiros por hectare. Agora, tem só 0,19. Essa queda ocorre porque aumentou o número de outros tipos de formigas que comem os ovos da formiga saúva. A cana orgânica é colhida verde, preservando assim suas características naturais integralmente. Folhas e palhas remanescentes dessa colheita, depositadas sobre os campos de cultivo, exercem um importante papel na conservação e condicionamento do solo (USINA, 2007).

As mudanças observadas na cadeia do açúcar apóiam novas possibilidades de produção em substituição ao açúcar *commodity* e de desenvolvimento de mercados para a alocação de açúcar diferenciado. Caracterizam-se por cooperação técnica em projetos que visam à diminuição de custos, atendimento às especificações de clientes e diferenciação de produtos. São realizados programas que englobam ação inter-organizacional para criação de alternativas concretas nos processos de negócios e de eficiência nos processos operacionais.

Para fazer frente às metas definidas são estabelecidos programas de treinamento dos recursos humanos, é criado suporte tecnológico aos fluxos físico e de informação, com o objetivo de aumentar a agilidade, precisão na tomada de decisão, melhorando a habilidade de percepção do ambiente competitivo. Ocorrem melhorias na infra-estrutura de armazenagem, de transportes e da entrega para atendimento à demanda dos clientes.

São desenvolvidos programas de incentivo, para aquisição de novos recursos para implementar mudanças tecnológicas e organizacionais nas empresas frente às mudanças observadas.

Outros produtos também estão sendo desenvolvidos a partir da cana-de-açúcar. Um deles é o plástico biodegradável que tem como objetivo explorar por inteiro o potencial da cana, além de permitir que o ciclo de produção termine deixando a menor quantidade possível de resíduos.

Essas iniciativas fazem com que as usinas de açúcar e álcool permaneçam competitivas dentro do setor sucroalcooleiro.

A Tabela 2 mostra as unidades produtoras de açúcar e álcool no Brasil, que estarão em operação na Safra 2007/2008.

Tabela 2 - Unidades Produtoras de Açúcar e Álcool no Brasil.

Região e Estados	Número	Região e Estados	Número
Centro Sul	282	Norte / Nordeste	82
Espírito Santo	6	Alagoas	26
Goiás	16	Amazonas	1
Mato Grosso	11	Bahia	4
Mato Grosso do Sul	11	Ceará	2
Minas Gerais	30	Maranhão	4
Paraná	29	Pará	1
Rio Grande do Sul	1	Paraíba	9
Rio de Janeiro	8	Pernambuco	26
São Paulo	170	Piauí	1
		Rio Grande do Norte	3
		Sergipe	4
		Tocantins	1
BRASIL			364

(Fonte: CANAOESTE, 2007)

Do total de usinas instaladas na região Centro-Sul do País, quase 80% já entraram em operação, segundo o último levantamento da UNICA (União da Indústria de Cana-de-Açúcar). A estimativa é de que até o final do mês de maio de 2007 todas estejam operando. O Centro-Sul conta com 282 usinas, que devem operar na safra 07/08, sendo que 19 delas são novas (UNICA, 2007).

Com relação ao percentual, a região centro-sul representa aproximadamente 77% das unidades produtoras e a região norte-nordeste 23% das unidades produtoras.

O estado de São Paulo aparece em primeiro lugar com 170 unidades produtoras, seguido por Minas Gerais com 30 unidades, Paraná com 29 unidades e Alagoas e Pernambuco com 26 unidades cada.

Os estados de São Paulo, Minas Gerais, Paraná, Alagoas e Pernambuco representam aproximadamente 77% do total de unidades produtoras de açúcar e álcool do Brasil.

A Tabela 3 mostra as produções de cana, açúcar e álcool no Brasil entre os períodos de safra de 1996/1997 e 2006/2007.

Tabela 3 - Produções de Cana, Açúcar e Alcool no Brasil.

Produção de Cana, Açúcar e Alcool - Brasil					
Safra	Alcool Anidro (m3)	Alcool Hidratado (m3)	Álcool Total (m3)	Açúcar (t)	Cana-de-açúcar (t)
96 97	4.629.340	9.801.109	14.430.449	13.631.888	289.520.522
97 98	5.699.719	9.722.534	15.422.253	14.847.044	302.198.516
98 99	5.679.998	8.246.821	13.926.819	17.960.587	315.640.797
99 00	6.140.769	6.936.996	13.077.765	19.380.197	310.122.784
00 01	5.584.730	4.932.805	10.517.535	16.020.340	254.921.721
01 02	6.479.187	4.988.608	11.467.795	18.994.363	292.329.141
02 03	7.009.063	5.476.363	12.485.426	22.381.336	316.121.750
03 04	8.767.898	5.872.025	14.639.923	24.944.434	357.110.883
04 05	8.172.488	7.035.421	15.207.909	26.632.074	381.447.102
05 06	7.662.622	8.144.308	15.806.930	26.214.391	382.482.002
06 07*	8.088.810	9.219.169	17.307.979	29.406.520	411.542.277

* Posição em 01/02/2007. (Fonte: CANAOESTE, 2007)

A figura 17 converte os valores da tabela 3 (Açúcar, Álcool Anidro e Hidratado) em Unicap's⁵ para efeito comparativo entre os Produtos.

Figura 17 - Comparativo das Produções Anidro x Hidratado x Açúcar.
(Fonte: CANAOESTE, 2007)

A figura 18 converte os valores da tabela 3 (Açúcar e Álcool Total) em Unicap's para efeito comparativo entre os Produtos.

Figura 18 - Comparativo das Produções Álcool Total x Açúcar.
(Fonte: CANAOESTE, 2007)

⁵ **Unicap's:** É o parâmetro utilizado pela Copersucar (Cooperativa de Produtores de Cana-de-Açúcar, Açúcar e Álcool do Estado de São Paulo), para uniformização dos diversos tipos de produtos fabricados pelas unidades industriais. Esse parâmetro serve para comparar diversos mix de produção, até de unidades diferentes, por tonelada de cana, ou seja, dividindo-se as Unicap's pelas toneladas de cana moídas, têm-se um número para ser comparado entre as unidades e até mesmo de safras anteriores, visando verificar a performance entre usinas ou da mesma usina em diversas safras (MACEDO, 2005).

A figura 18 mostra que a produção de álcool até a Safra 00/01 foi sempre maior que a produção de açúcar, sendo que a partir da Safra 01/02 houve uma inversão neste quadro onde se passou a produzir sempre mais açúcar do que álcool. Isso se deve ao preço do açúcar que nas últimas safras, mostrou-se mais atrativo do que o álcool para as unidades produtoras.

A figura 19 mostra um comparativo entre o percentual de produção de Álcool Total produzido pelo Brasil entre as safras 96/97 e 06/07, com o total de Açúcar produzido no mesmo período.

Figura 19 - Percentual de Produção Álcool Total x Açúcar.

(Fonte: CANAOESTE, 2007)

Analisando a figura 19 observamos que mesmo com o preço do açúcar sendo mais atrativo para as unidades produtoras, no montante, considerando as últimas 11 safras, a produção de álcool superou a produção de açúcar em 3,66%. Pode-se dizer que houve praticamente um equilíbrio entre a produção dos dois produtos, o que deve permanecer para as próximas safras, mesmo porque é necessário o abastecimento do consumo interno do País e diga-se de passagem esse consumo aumentou significativamente com a produção de veículos *flex-fuel*, que será demonstrado ainda neste capítulo.

Logo abaixo é demonstrada uma figura contendo um gráfico com informações sobre a produção de Cana do Brasil e a sua evolução nas ultimas safras.

Figura 20 - Produção de Cana de Açúcar do Brasil.

(Fonte: IDEA, 2007)

A figura 20 mostra que a produção de cana de açúcar no Brasil, vêm aumentando a partir da safra 00/01. O crescimento das exportações brasileiras de açúcar explica boa parte da significativa expansão do setor sucroalcooleiro nacional nos últimos anos e as perspectivas promissoras do mercado interno e externo para esse bio-combustível num futuro bastante próximo. Sem dúvida, a necessidade de fornecer etanol para o mercado interno em expansão e para o mercado internacional, que anseia por fontes renováveis de energia, traz excelentes oportunidades para incrementos ainda maiores no crescimento do setor. Nos anos recentes, nota-se o aumento da produção de cana-de-açúcar e de seus produtos derivados, açúcar e etanol, tanto nas tradicionais regiões produtoras como em estados que representam novas fronteiras agrícolas ⁶ para a cultura canavieira no Brasil (UNICA, 2007).

⁶ **Novas Fronteiras Agrícolas:** Um bom exemplo de novas fronteiras agrícolas são as abertas no estado de São Paulo, Bahia e Mato Grosso do Sul, onde a maior rentabilidade da cana-de-açúcar frente ao gado fez com que os canaviais avançassem e gradualmente tomassem o lugar da pecuária (UNICA, 2007).

4.6 - Evolução da Produção no Brasil

A agroindústria brasileira atingiu nos últimos anos papel de destaque na economia do Brasil. Com taxas de crescimento na casa dos 5%, o dinamismo do mercado se deve a fatores como aumento da safra agrícola, da exportação e dos preços internacionais. Mas até chegar a esse momento, a agroindústria percorreu um longo caminho, principalmente nos mercados de açúcar e álcool (COPERSUCAR, 2007).

No final do século XVIII, a concorrência do açúcar de beterraba europeu e do açúcar de cana das Antilhas e de Cuba ameaçavam a liderança do Nordeste brasileiro, grande produtor de açúcar. Ao mesmo tempo, a lavoura canavieira reconquistou importância como atividade econômica em São Paulo, migrando do litoral para o interior, onde estavam as áreas de solos férteis de terra roxa.

Ainda que a partir do século XIX São Paulo tivesse sua economia focada no café, a produção açucareira continuou crescendo até 1929, quando o *crash* da Bolsa de Nova York abalou os mercados mundiais, atingindo fortemente a economia brasileira, baseada no café. A cana sofreu o mesmo impacto.

Com a crise econômica de 1929, foi criado o Instituto do Açúcar e do Alcool (IAA), cuja função era controlar a produção para manter os preços em um nível adequado, protegendo o produto brasileiro no mercado mundial. Para atingir suas finalidades, o IAA estabeleceu um sistema rígido de cotas, que eram distribuídas entre as diferentes unidades produtoras, ou seja, cada engenho e usina só poderiam produzir uma determinada quantidade de açúcar. Com os preços e a produção rigidamente controlados, a única maneira de manter o negócio lucrativo era reduzir os custos e aumentar a produtividade.

Desse modo, era praticamente inevitável que a produção viesse a se concentrar em grandes usinas, com capacidade para moer milhares de toneladas de cana por dia. Desde então, a agroindústria açucareira vem passando um constante processo de modernização, com o agrupamento de empresas, buscando maior rendimento das lavouras e barateamento dos custos (COPERSUCAR, 2007).

4.7 - Custos de Produção

Se por um lado o futuro parece promissor, por outro o presente reserva um quadro permeado pela pressão dos custos que incidem sobre o processo produtivo do setor sucroalcooleiro. Dos insumos à mão-de-obra, dos equipamentos ao frete, das máquinas agrícolas ao arrendamento da terra, o preço de quase tudo subiu na cadeia da agroindústria canavieira (IDEA, 2007).

Mas o ritmo preocupante dos custos vem da análise não só dos profissionais que lidam no dia-a-dia com as demandas das usinas, mas também de especialistas econômico-financeiros ligados ao setor.

Na verdade, há uma grande variabilidade nos números das usinas quando se lembra que três fatores são determinantes para a definição dos custos de produção de açúcar e álcool em cada unidade:

1. seu perfil de produção,
2. a eficiência operacional na indústria,
3. e a qualidade da matéria-prima.

Uma usina esmagando cana de boa qualidade e produzindo-a a baixo custo, com bom teor de açúcares, com processo de fabricação bem ajustado e apresentando otimização do uso do seu parque industrial, certamente apresentará custos muito menores que outra operando em condições opostas. O perfil da produção foi citado porque também afetará os resultados, afinal custa menos produzir álcool que fabricar açúcar, assim como fabricar açúcar do tipo VHP (Açúcar Cristal utilizado como matéria prima para outros processos) é menos oneroso que produzir açúcar cristal branco. O mais importante para enfrentar a elevação dos custos é demonstrar eficiência tanto no campo quanto na indústria e dispor de informações e flexibilidade para ajustar o seu perfil de produção, estando de olho nas duas pontas da equação do lucro: custos de produção e preços praticados pelo mercado para cada linha de produção.

Desde que o setor sucroalcooleiro saiu das rédeas do governo, tendo que se submeter a um mercado desregulamentado, sujeito às condições da livre oferta e procura, os produtores de cana de açúcar e álcool têm buscado insistentemente esses ganhos de eficiência.

Por esse fato, tem de se analisar cada caso. Existem aquelas empresas que investiram fortemente no seu desenvolvimento organizacional, na introdução de novas tecnologias e, acima de tudo, recrutaram profissionais gabaritados, passando a contar com equipe competente (MACEDO, 2005).

Assim, tomando decisões com base em sólidos sistemas de custos e orçamentos, e em sistemas de planejamento e controle, certamente se destacaram e ganharam competitividade (IDEA, 2007).

São empresas em que há uma perfeita integração entre as áreas de produção de matéria-prima, de produção industrial, financeira e de vendas. Infelizmente outras empresas ficaram para trás e não acompanharam este bom desenvolvimento ou pelo menos não demonstraram a mesma evolução.

A tendência, se não reagirem rapidamente, é seu desaparecimento ou sua incorporação por outras unidades.

Todavia, independentemente da forma como as usinas têm se preparado para tornarem-se mais eficientes e mais competitivas, a pressão dos custos tem se acentuado sobre diferentes áreas das empresas.

A figura 21 mostra a composição dos custos de produção com a cana de açúcar na região Centro-Sul.

Figura 21 - Composição dos Custos de Produção de Cana.

(Fonte: IDEA, 2007)

Algumas características são necessárias e precisam ser levadas em conta ao se analisar os custos do setor:

- a) a valorização da cana de terceiros pelo bom valor de fechamento do kg de ATR,
- b) a participação mais expressiva do custo do fator terra na composição do custo de produção de cana,
- c) a valorização do real frente ao dólar,
- d) o mercado mais favorável ao álcool anidro comparado ao hidratado.

Mesmo sendo a agricultura uma atividade econômica de difícil gerenciamento da produção, por enfrentar fatores incontrolláveis, tais como, clima, pragas, doenças, adequação de solo em termos de acidez e composição de nutrientes para a produção de biomassa, é possível definir índices de eficiência técnica, condição necessária para que na controladoria se estabeleçam padrões de custos.

O estabelecimento do custo correto de produção evita que as informações obtidas induzam a decisões equivocadas quando houver super avaliação dos custos isto implicará em preços maiores do que os clientes estão dispostos a pagar, comprometendo assim a capacidade produtiva ou quando houver sub avaliação dos custos, isto levará a preços competitivos, entretanto, quanto mais se produza e venda, maior será o prejuízo (MARION, 1996).

Na cultura da cana-de-açúcar, um dos maiores agronegócios do Brasil, têm-se desenvolvido estudos visando gerenciar os fatores de produção, para que o custo da tonelada de cana seja economicamente atrativo.

Para a Safra 2007/2008 acredita-se que o custo total deve se reduzir um pouco devido ao grande crescimento da produção, mas não existe um valor fechado. Porém, o grande problema do setor para a próxima safra será a taxa de câmbio. Se o dólar ficar no nível de R\$ 2,00, a remuneração das exportações será muito baixa aos níveis atuais da Bolsa de Nova York. Para se ter uma idéia, em dezembro de 2002 um dólar valia R\$ 3,53 e no final de março de 2007 a cotação já estava em R\$ 2,05 (uma desvalorização de 42% do dólar ou uma valorização de 72% do Real).

O aumento da produtividade é uma forte estratégia de redução de custo de produção. Isso pode ser comprovado conforme tabela demonstrativa de uma usina que mói 1,2 milhões de toneladas de cana, onde os gastos crescem quando se aumenta a produtividade, mas o custo de produção, em reais por tonelada, diminui.

A figura 22 demonstra o comportamento do preço da cana e seus custos de produção. É possível observar que em várias safras o preço da cana (R\$ / Ton), foi inferior aos seus custos de produção, ou seja, nessas safras houve prejuízo.

Figura 22 - Comportamento do Preço da Cana e Custos de Produção.
(Fonte: IDEA, 2007)

Usinas que produzem mais de 90 toneladas de cana por hectare são muito competitivas, como mostra a tabela 4 abaixo.

Tabela 4 - Produtividade x Custo.

PRODUTIVIDADE x CUSTO		
Produtividade (t/ha)	Orçamento (R\$)	Custo (R\$ / t)
60	28.265.294,00	32,74
75	29.201.177,00	31,31
70	30.299.032,00	30,08
75	31.315.901,00	29,02
80	32.332.770,00	28,09
85	33.349.639,00	27,27
90	34.366.508,00	26,54
95	35.383.377,00	25,88
100	36.400.246,00	25,30

(Fonte: IDEA, 2007)

Fazendo uma análise qualitativa, de uma maneira geral, é possível observar que as usinas estão melhorando os seus processos, investindo em máquinas, equipamentos e tecnologia da informação, o que traz uma redução de custos. E isso está acontecendo porque o mercado está aquecido pela demanda de álcool e pelo cenário positivo para o açúcar. As empresas estão investindo em suas usinas e em rendimento por meio de planos de ação, ferramentas de gestão, controle e análise. A implantação disso tudo resulta em eficiência e, conseqüentemente, redução de custos porque melhora a produtividade e o rendimento das usinas.

4.8 - Exportações

O mercado interno do álcool vem ganhando destaque no segmento agroindustrial brasileiro, proporcionado pela retomada do aumento do consumo doméstico em consequência do preço competitivo do combustível em relação à gasolina. Por outro lado, há um potencial de crescimento nas vendas externas do álcool, que possivelmente será utilizado para atender parte da demanda mundial por etanol. O álcool está em expansão por ser um combustível ainda barato, renovável e cujo emprego como alternativa para a matriz energética mundial está em fase de crescimento. Esse crescimento pode ser observado pela figura 23 com o resultado ocorrido no período de 1998 até março de 2007.

Figura 23 - Exportação de Álcool do Brasil.

(Fonte: IDEA, 2007)

A tendência de aumento da produção de álcool no Brasil ocorre por vários fatores, como aumento da frota de carros bi-combustível (demanda interna), Protocolo de Quioto (demanda externa) e aumento do preço do petróleo, entre outras coisas. Uma questão relevante na expansão da utilização do álcool como alternativa de substituição da matriz energética predominantemente fóssil é a possível mudança da atual configuração de produção de petróleo concentrada em poucas regiões, basicamente o Oriente Médio. Além disso, deve-se avançar para uma forma desconcentrada, que abrangerá outras áreas de possíveis produtores do combustível verde. Isto constitui uma oportunidade para os países que possuem abundância de terras, sol e água, que é o caso de alguns países subdesenvolvidos e em desenvolvimento (IEA, 2007).

Para isso, é necessário disseminar a tecnologia de produção, de forma que mais países produzam álcool a partir da cana-de-açúcar ou de outras fontes renováveis, como a hidrólise da celulose ⁷. Dessa forma, diminuem-se os riscos de crises de oferta e abre-se caminho para tornar o produto uma *commodity* convencional. Assim, os consumidores passam a ter um pouco mais de segurança quanto à disponibilidade do produto (IEA, 2007). A figura 24 mostra a projeção de exportação de álcool do Brasil para as próximas safras.

Figura 24 - Projeção de Exportação de Álcool do Brasil.
(Fonte: IEA, 2007)

⁷ **Hidrólise de Celulose:** A hidrólise da celulose produz hexoses (açúcares com moléculas de seis carbonos, como a glicose) que podem ser fermentadas com facilidade (REVISTA QUÍMICA, 2007).

Espera-se que com a taxa de câmbio atual, o açúcar VHP fique abaixo de R\$ 20,00 / saca, líquido para a usina, muito abaixo do custo mínimo estimado de R\$ 24,50 / saca. Caso o câmbio fosse de R\$ 2,50, esse valor saltaria para R\$ 25,00 / saca, ficando acima do custo mínimo estimado (IDEA, 2007). A figura 25 mostra a exportação de açúcar do Brasil no período de 1998 até março de 2007.

Figura 25 - Exportação de Açúcar do Brasil.

(Fonte: IDEA, 2007)

As exportações brasileiras cresceram 2,6% em 2007. O ritmo das exportações brasileiras de açúcar esteve em queda nos últimos meses devido, principalmente, ao excedente de produção mundial e ao baixo preço de exportação do produto. Somente no mês de dezembro de 2007, as exportações totalizaram 1,37 milhões de toneladas, 39,8% menor que o mesmo período de 2006, quando as exportações somaram 2,28 milhões de toneladas. Ainda assim, o ano de 2007 superou 2006 em 2,59% no total embarcado que chegou a 19,4 milhões de toneladas, contra 18,9 milhões de toneladas no período anterior. O preço por tonelada, apesar de estar em ritmo e tendência de alta, quando comparado aos valores praticados em 2006, se mostra bastante desvalorizado. A média verificada para o último mês ficou em US\$ 248,49 por tonelada enquanto em dezembro de 2006 o valor era de US\$ 316,52, o que resulta na variação de 21,49%. Na média anual, em 2007, o açúcar foi comercializado com o mercado externo a US\$ 263,47 por tonelada, enquanto que em 2006 a média alcançou US\$ 326,81 por tonelada. O principal destino do açúcar brasileiro em 2007 foi a Rússia com 4,0 milhões de toneladas, seguida pelos Emirados Árabes Unidos, República Islâmica do Irã e Nigéria (UNICA, 2007)

4.9 - Projeção para Agroindústria Canavieira no Brasil

A dinâmica do novo modelo de desenvolvimento da agroindústria canavieira brasileira contém possibilidades e riscos. A emergência de padrões mais intensivos de utilização dos fatores de produção tornará, certamente, mais competitivo em termos internacionais esse importante ramo de atividade econômica, potencializando o crescimento das exportações de açúcar e álcool. Ademais, esse processo está incentivando a melhoria dos indicadores microeconômicos de desempenho das firmas, permitindo-lhes maior otimização do processo produtivo (MORAES & SHIKIDA, 2002).

Do ponto de vista do trabalho, o que se prenuncia é a ampliação do desemprego. Além dos trabalhadores do corte de cana, ameaçados pela mecanização, estão sob risco também operários qualificados, como consequência da automação das plantas industriais e dos escritórios das usinas. Cabe salientar, ainda, que a tecnologia tem servido para aumentar a disciplina no trabalho e para reduzir o poder de barganha dos sindicatos dos trabalhadores. A sustentabilidade social do novo modelo parece ser, a princípio, algo difícil.

Do ponto de vista econômico, espacial e ambiental, a situação não se mostra muito diferente. O que se observa no presente é a intensificação do uso de todos os fatores, particularmente a terra. O aumento da produção agrícola tem ocorrido de forma intensiva, como consequência da mecanização e da quimificação do processo produtivo, e também de forma extensiva, em função da expansão da área cultivada. Há que se avaliar, de forma cuidadosa, os impactos da mecanização e da quimificação sobre o solo, e nada garante que a expansão das áreas cultivadas não se dê, novamente como no passado, na direção de terrenos inadequados às culturas temporárias e intensivas.

Para o mercado interno de etanol no Brasil, o fato mais significativo é o expressivo aumento da demanda que começa a ocorrer em função da entrada no mercado dos novos carros bi-combustível (MACEDO, 2005).

Foi apresentado em 02 de maio de 2007, na USP Ribeirão Preto pelo coordenador do projeto Biodiesel Brasil, uma palestra com o título: "O mercado mundial de etanol – Tendências, perspectivas, inovações, tecnologias e instalação de usinas".

Na palestra “O Mercado mundial de etanol” foi discutido a história recente do setor sucroalcooleiro, explicando os motivos do fracasso do pró-álcool vinte anos atrás e as razões do sucesso do setor nos dias de hoje. Para ele, entre outros motivos, o etanol só ganhou a proporção atual graças ao avanço da tecnologia, tanto em relação às usinas, quanto no que diz respeito à evolução dos motores movidos a álcool, os *flex-fuel*. Ele mostrou, ainda, gráficos com as estatísticas do setor, bem como as expectativas de venda para os próximos anos. Para atender a demanda interna, o país precisa construir 100 novas usinas, explica o pesquisador (CANAVIEIROS, 2007).

Os veículos *Flex-fuel* são veículos equipados com sistema de gerenciamento do motor capaz de identificar com precisão a presença de gasolina e/ou etanol no tanque de combustível do veículo e ajustar automaticamente a operação do motor para o combustível existente. Há versões que utilizam sensores de etanol instalados no tanque ou na linha de suprimento de combustível do veículo, mais comuns nos EUA, e outras que identificam a presença de etanol por meio de um sensor que mede o teor de oxigênio presente no gás de escapamento, sistema adotado no Brasil (UNICA, 2007).

A maioria das montadoras de carros nacionais já possui modelos *flex*, como a Volkswagen, a GM, a Ford, a Fiat, a Renault e a Peugeot. Até a montadora japonesa Honda resolveu, mais recentemente, ingressar nesse grupo. Lançou no segundo semestre de 2006 a versão *flex* do Honda Civic. A Toyota, outra marca japonesa, líder de mercado de veículos híbridos movidos à gasolina e eletricidade, se mostra disposta em desenvolver modelos movidos tanto a álcool como a gasolina para serem lançados no Brasil. Atualmente as montadoras que não oferecem carros flexíveis perdem mercado, embora seu público alvo seja diferente e de maior poder aquisitivo, para quem o preço do combustível não interfere tanto na compra do automóvel (IDEA, 2007)

Ainda não há nenhuma sinalização de que o Brasil vá exportar carros flexíveis. No entanto, há interesse em misturas baixas de álcool na gasolina, mas não de carros movidos exclusivamente a álcool, como ainda é o caso do Brasil numa pequena proporção. Como o interesse de outros países em adicionar álcool à gasolina, a exportação de veículos flexíveis poderá estar próxima. Com a presença dos carros flexíveis nas ruas de nossas cidades e nas garagens de nossas casas e prédios, percebemos que a tecnologia que parecia impossível está bem ao nosso lado.

Hoje também já existem os carros *tri-flex*, que rodam com álcool, gasolina e GNV (gás natural veicular).

De todos os veículos flexíveis vendidos no país, a Fiat é responsável por 31,3% desse montante, seguida pela Volkswagen com 28,3%, GM com 25,0%, Ford com 7,4%, Peugeot com 5,4%, Renault com 2,5% e Honda com 0,1% (IDEA, 2007). A figura 26 mostra com mais detalhes essa distribuição e a figura 27 mostra a venda de carros *flex* no Brasil no período de 2003 a 2006.

Figura 26 - Produção por Montadoras.
(Fonte: ANFAVEA, 2007)

Figura 27 - Vendas de Carros *Flex* no Brasil.
(Fonte: ANFAVEA, 2007)

Analisando a figura 27 observamos um aumento nas vendas de carros *flex* nos últimos quatro anos, sendo que a variação de 2005 para 2006 foi de 75,36%.

Com o encarecimento do petróleo no mercado mundial, usar o álcool tem apresentado vantagens para o motorista. Em alguns Estados, como São Paulo, ele chega a custar metade do preço da gasolina. Como o consumo relativo é diferente, seu uso é vantajoso quando o preço for inferior a 70% do da gasolina. A preferência pelo *flex fuel*, porém, deve levar ao final da produção dos carros com motor a álcool. Depois de atingir níveis muito baixos de participação no mercado, em torno de 0,1%, em 1997 e 1998, as vendas haviam registrado ligeiro aumento, mas agora enfrentam a concorrência dos bi-combustíveis e o grande trunfo da flexibilidade.

A figura 28 mostra a venda de carros no Brasil por tipo e combustível e a figura 29 a sua participação com relação a essas vendas no ano de 2006, onde podemos observar que os veículos *flex* representaram praticamente 80% do total de vendas.

Figura 28 - Vendas por Tipo de Combustível.
(Fonte: ANFAVEA, 2007)

Figura 29 - Participação nas Vendas.
(Fonte: ANFAVEA, 2007)

Apesar da euforia criada pelas oportunidades vislumbradas ao setor sucroalcooleiro brasileiro, os produtores de cana devem se preparar para uma safra com preços menores para a matéria-prima. As razões são simples. Uma delas é que a expansão do setor está em curso e, dessa forma, pode criar desajustes temporários para os elos da cadeia. A produção de cana segue em crescimento e em ritmo próprio, diferente do que vem ocorrendo na área industrial. Assim, com maior oferta, o preço tende a ficar mais baixo. É a lei do mercado (CANAVIEIROS, 2007).

Essa disfunção na oferta e nos preços já eram esperados, segundo especialistas, pelas condições em que a expansão do setor ocorre: em grandes proporções e em curto espaço de tempo. Para atravessar esse período de turbulência é necessário que o produtor fique atento ao comportamento dos preços da cana antes de planejar investimentos, contrair despesas ou financiamentos.

Esses desequilíbrios são efeitos do processo de expansão e devem ser administrados e superados. Apesar deles, os horizontes nunca foram tão positivos.

Deve-se atentar para o fato de que as tecnologias que estão sendo geradas no setor, embora potencializem uma racionalização no uso dos subprodutos e insiram-se na estratégia de integração e diversificação de produtos, não estão adequadas para assegurar uma inserção estável dos pequenos e médios fornecedores de cana.

Pode-se afirmar, portanto, que não existem garantias de que o desenvolvimento do setor acontecerá no futuro de forma sustentável (MORAES & SHIKIDA, 2002).

4.10 - Números do Setor e Participação do Capital Estrangeiro

A área de consultoria do Instituto de Desenvolvimento Agroindustrial (IDEA) acaba de atualizar alguns números do setor sucroalcooleiro nacional, que apontam fatos impressionantes. A participação do capital estrangeiro nas usinas brasileiras subiu de 6,78% na safra 2005/2006 para 9,29% na safra 2006/2007, um crescimento de 37% em apenas uma safra. Além disso, os 20 maiores grupos do setor responderam por 42,1% da cana processada no Brasil, confirmando a tendência de concentração da agroindústria canavieira nacional.

Os cinco maiores grupos do setor sucroalcooleiro processam 18% de toda cana brasileira, conforme é demonstrado logo abaixo:

- 1º Cosan,
- 2º Vale do Rosário,
- 3º Carlos Lyra,
- 4º São Martinho,
- 5º Tércio Wanderley.

Figura 30 - Produção dos Cinco Maiores Grupos do Setor.
(Fonte: IDEA, 2007)

O açúcar é insumo para fabricação de alimentos, diretamente como ingrediente genérico, ou como insumo para produção de ingredientes específicos, que facilitam a processabilidade dos alimentos. E é sob esta lógica que, após a desregulamentação e abertura do mercado brasileiro, a entrada do capital estrangeiro no setor alimentício formou uma nova arquitetura de operações globais entre cadeias produtivas agroindustriais e indústria alimentícia de produto final. As usinas foram obrigadas a buscar capacitação para atendimento ao mercado industrial, seguindo padrão mundial de produção, caracterizado por: flexibilidade, qualidade e velocidade nos ciclos de lançamento de produtos e de entrega (ASSUMPÇÃO, 2001).

A figura 31 mostra o % do capital estrangeiro na safra 05/06.

Figura 31 - Participação do Capital Estrangeiro na Safra 05/06.
(Fonte: IDEA, 2007)

A figura 32 mostra o % do capital estrangeiro na safra 06/07.

Figura 32 - Participação do Capital Estrangeiro na Safra 06/07.
(Fonte: IDEA, 2007)

Analisando as figuras 31 e 32 é possível observar um aumento significativo da safra 06/07 com relação à safra 05/06. Para os agentes internacionais, o maior controle do destino das exportações, a participação no elo de produção do sistema agroindustrial canavieiro, as expectativas de crescimento, a profissionalização da gestão das usinas e a capitalização das empresas foram os principais benefícios obtidos juntamente com esta evolução (UNICA, 2007).

4.11 - Considerações Finais do Capítulo

Pelo sexto ano consecutivo, a área plantada de cana-de-açúcar aumentou. Com a forte demanda internacional por combustíveis alternativos, a consolidação do mercado de veículos *flex-fuel* e a grande competitividade da indústria nacional, o setor sucroalcooleiro está em franca expansão. Além do mais, o setor sucroalcooleiro tem mostrado sinais de crescimento por intermédio dos anúncios de grandes investimentos na construção de novas usinas de álcool, bem como, da ampliação das existentes.

Vários países estão interessados na tecnologia de veículos com motores Bi-combustíveis. Como certeza este setor continuará crescendo, em função da grande disponibilidade de áreas para plantio, da tecnologia desenvolvida e da grande experiência adquirida neste tipo de combustível.

As empresas brasileiras, para se tornarem competitivas e sobreviver nesta economia globalizada, tiveram de introduzir modificações em suas estratégias de competição e crescimento. Mais e mais empresas estão descobrindo o valor da informação para seus negócios. Estão buscando e aprendendo a utilizar informações sobre o ambiente interno e externo, sobre ameaças e oportunidades. Estão, enfim, preocupadas em ter o maior conhecimento possível sobre dados e fatos do seu contexto empresarial.

O atual processo requer o uso adequado da informação como insumo para a tomada de decisões e a utilização de modernas tecnologias de informação para permitir o acesso mais rápido, no sentido de possibilitar que os dados sejam empregados no momento oportuno.

Somente os indivíduos bem preparados, que sabem transformar dados e informações em conhecimentos e com formação ética, terão condições de enfrentar os desafios e ameaças e aproveitar as oportunidades em benefício da sociedade.

O objetivo deste capítulo foi mostrar uma visão ampla sobre o setor sucroalcooleiro, sua importância para o País e como a tecnologia da informação vêm se destacando e crescendo dentro do setor.

No próximo capítulo será demonstrada a aplicação desta tecnologia, por meio de uma Pesquisa-ação feita em uma Usina de Açúcar e Álcool da Região de Ribeirão Preto, estado de São Paulo.

Capítulo 5 - Aplicação da Pesquisa-ação

A constituição da Usina se deu em 1946, sendo que a primeira safra foi realizada em 1947. Em 1957 foi feita a aquisição de uma segunda Usina e em 1973 passou a ter participação na União São Paulo S.A. Agricultura, Indústria e Comércio.

Contribui regularmente com todas as iniciativas das comunidades onde atua, em especial, àquelas relacionadas à Santa Casa de Misericórdia, creches, escolas, orfanatos, clubes de serviços (*Lions* e *Rotary Club*), instituições beneficentes e outros com grande responsabilidade social.

Participa ativamente em projetos ambientais na utilização de vinhaça e torta de filtro como fertilizantes orgânicos, controle biológico de pragas da cana-de-açúcar, plantio de mudas de árvores nativas da flora brasileira, colheita de cana mecanizada crua e inventário da fauna existente nas áreas agrícolas das Usinas.

Possui Certificação ISO 9001:2000 e mais cinco certificações internacionais para produção de açúcar orgânico.

Maior Produtor Mundial de Açúcar Orgânico, pioneira na co-geração com interligação ao sistema de distribuição (CPFL), Produtor de Plástico Biodegradável à partir da Cana-de-Açúcar e tecnologia para extração de cera da torta de filtro (resíduo da Produção de Açúcar).

Seus principais produtos são: Açúcar Convencional, Açúcar Orgânico, Alcool Anidro, Alcool Hidratado, Alcool Orgânico, Energia Elétrica, Levedura, Plástico Biodegradável e Cera.

Atualmente o grupo conta com aproximadamente 3.000 funcionários.

5.1 - Capacidade de Produção e Dados Estatísticos

A seguir serão demonstrados os dados de produção e alguns resultados da organização referentes às últimas safras:

Figura 33 - Capacidade Diária de Moagem de Cana. (**Fonte:** Usina, 2007)

A figura 33 mostra a capacidade diária de moagem de cana em toneladas do grupo, sendo que 12.000 toneladas são referentes a Usina - 1 o que representa 62% do total e 7.200 toneladas referentes a Usina - 2 o que representa 38% do total. As duas Usinas juntas totalizam uma carga de moagem diária de 19.200 toneladas.

Figura 34 - Capacidade Diária de Produção de Açúcar. (**Fonte:** Usina, 2007)

A figura 34 mostra a capacidade diária de produção de açúcar em toneladas do grupo, sendo que 800 toneladas são referentes a Usina - 1 o que representa 58% do total e 575 toneladas referentes a Usina - 2 o que representa 42% do total. As duas Usinas juntas totalizam uma carga de produção de açúcar diária de 1.375 toneladas.

Figura 35 - Capacidade Diária de Produção de Álcool. (Fonte: Usina, 2007)

A figura 35 mostra a capacidade diária de produção de litros de álcool do grupo, sendo que 700.000 litros são referentes a Usina - 1 o que representa 64% do total e 400.000 litros referentes a Usina - 2 o que representa 36% do total. As duas Usinas juntas totalizam uma carga de produção de álcool diária de 1.100.000 litros.

Figura 36 - Capacidade Diária de Produção de Levedura. (Fonte: Usina, 2007)

A figura 36 mostra a capacidade diária de produção de levedura em toneladas do grupo, sendo que 15 toneladas são referentes a Usina - 1 o que representa 71% do total e 6 toneladas referentes a Usina - 2 o que representa 29% do total. As duas Usinas juntas totalizam uma carga de produção de levedura diária de 21 toneladas.

Figura 37 - Total de Cana Moída em 2005. (Fonte: Usina, 2007)

A figura 37 mostra a participação da cana própria e de fornecedores em toneladas no ano de 2005. A cana própria com 1.762.359 toneladas representou 53% do total de cana entregue nas usinas e a cana de fornecedores com 1.573.988 toneladas representou 47% do total de cana entregue nas usinas. O total de cana moída foi de 3.336.347 toneladas.

Figura 38 - Total de Cana Moída em 2006. (Fonte: Usina, 2007)

A figura 38 mostra a participação da cana própria e de fornecedores em toneladas no ano de 2006. A cana própria com 1.751.328 toneladas representou 54% do total de cana entregue nas usinas e a cana de fornecedores com 1.508.672 toneladas representou 46% do total de cana entregue nas usinas. O total de cana moída foi de 3.260.000 toneladas.

Figura 39 - Evolução Líquida da Receita. (Fonte: Usina, 2007)

A figura 39 mostra a evolução da receita líquida do grupo. O que podemos observar é que a partir do ano de 2003 o grupo vem aumentando significativamente a sua receita líquida.

Figura 40 - Moagem de Cana dos últimos 10 anos. (Fonte: Usina, 2007)

A figura 40 mostra a moagem de cana dos últimos 10 anos do grupo. Podemos observar pela figura que o grupo não aumentou nesse período a sua capacidade de moagem, que gira em torno de 3.000.000 de toneladas por ano/safra.

Figura 41 - Produção de Açúcar dos últimos 10 anos. (Fonte: Usina, 2007)

A figura 41 mostra a produção de açúcar dos últimos 10 anos do grupo. Podemos observar pela figura que a produção de açúcar orgânico mesmo tendo um aumento na sua produção nos últimos anos, ainda é bem inferior a produção de açúcar convencional. Vale a pena ressaltar que o grupo é o maior produtor mundial de açúcar orgânico.

Figura 42 - Produção de Álcool dos últimos 10 anos. (Fonte: Usina, 2007)

A figura 42 mostra a produção de álcool dos últimos 10 anos do grupo. Podemos observar pela figura que a produção de álcool anidro tem sido maior em relação à produção de álcool hidratado nos últimos anos.

Figura 43 - Produção de Levedura dos últimos 10 anos. (Fonte: Usina, 2007)

A figura 43 mostra a produção de levedura dos últimos 10 anos. A levedura é produzida pelo resultado do resíduo da fabricação do álcool e utilizada como ração para Bovinos.

5.2 - Pesquisa-ação em uma Usina de Açúcar e Álcool

Conforme descrito anteriormente, esta pesquisa foi desenvolvida em uma empresa do setor sucroalcooleiro, situada na região de Ribeirão Preto, no estado de São Paulo. Esta empresa vem tendo uma série de dificuldades e desafios com seus sistemas de informação, principalmente com a falta de informações gerenciais para tomada de decisão.

A maioria das Usinas de Açúcar e Álcool, a partir de certo volume de matéria prima, torna-se difícil de gerenciar sem uma ferramenta que dê informações rápidas e precisas. Há necessidade de se obter o máximo de margem de contribuição em cada tonelada de cana-de-açúcar. Por isso é necessário um ambiente, que reúna todas as condições para uma melhor tomada de decisão.

Nesse enfoque, é possível identificar informações estratégicas, extraídas de um banco de dados com informações históricas sobre tipos de solo, variedades de cana, distância média dos talhões das fazendas à usina, informações sobre os dados tecnológicos (FIBRA, PC, POL e ATR), entre outras, que vistas dentro de um contexto se tornam imprescindíveis para uma eventual tomada de decisão.

O objetivo fundamental do modelo proposto é o de colocar à disposição do Diretor e Gerente Agrícola a informação, onde e quando ela for necessária, visto a grande quantidade de dados e informações geradas pelos sistemas transacionais agrícolas da organização.

Assim sendo, as organizações utilizam cada vez mais a tecnologia da informação como ferramenta de competitividade, com impactos importantes e positivos nos seus negócios, nos mais variados ramos de atividade. A grande mudança de enfoque hoje é que essa tecnologia deixa de ser apenas um apoio às atividades produtivas para tornar-se parte integrante delas, muitas vezes redefinindo a própria maneira de se fazer negócios.

A gestão do conhecimento, e não do tempo, torna-se fator crítico, e sendo bem conduzida pode trazer retornos imediatos (ARAÚJO, 2005).

Tendo esses fatores em mente, os investimentos em informação podem realmente contribuir para o sucesso, uma vez que a mesma passa a ser aplicada em favor da competitividade empresarial. A habilidade em fazer isso da maneira correta pode representar um diferencial importante e, assim, sua influência tem que ser levada em consideração nos processos decisórios da empresa. Nenhuma organização pode ignorar as implicações que a tecnologia da informação pode representar na sua área de atuação. O risco pode ser a perda da competitividade, gerando conseqüentemente a sua extinção no mercado.

A fase inicial da pesquisa intervencionista contemplou uma entrevista não estruturada, que foi realizada na própria empresa, e teve duração de, aproximadamente quatro horas. A entrevista foi feita com o Diretor e o Gerente da área agrícola da usina.

Naquele primeiro contato, o pesquisador identificou a grande vontade dos dirigentes em obter informações gerenciais e estratégicas com relação aos seus processos agrícolas. As informações existentes na usina naquele momento se concentravam apenas nas necessidades operacionais do dia-a-dia.

Após a sua explanação, o pesquisador comentou que estas informações se encontravam nos diversos sistemas agrícolas distribuídos pela empresa e apresentou a tecnologia de SAD, que poderia ajudá-lo a integrar toda essa informação em um único local e a partir daí, por meio de um novo sistema de informação, extrair as informações necessárias para o seu gerenciamento e controle.

O pesquisador lhe forneceu informações preliminares a respeito de SAD e também explicou as principais vantagens e dificuldades na implantação do sistema como também os recursos que necessitaria para a aplicação da pesquisa em sua empresa.

O diretor da empresa mostrou preocupação com o prazo para a implementação do sistema na usina. O pesquisador lhe falou que o desenvolvimento do modelo para armazenamento das informações seria a parte mais crítica de todo o processo, acumulando aproximadamente 70% de todo o trabalho e que após essa fase os esforços seriam em torno do desenvolvimento dos relatórios de acordo com as necessidades e prioridades da empresa no seu dia-a-dia.

O pesquisador deixou claro para o diretor que após a implementação do novo ambiente a sua empresa provavelmente estaria preparada para tomar decisões a partir de informações integradas e consistentes, como também informações inesperadas que o sistema seria capaz de fornecer e até mesmo encontrar informações que estariam incorretas no seu ambiente operacional.

Além disso, foi enfatizado que o próprio pesquisador arcaria com as custas da elaboração do planejamento e execução do estudo, e apenas necessitaria, para o desenvolvimento da pesquisa, do comprometimento do diretor e de sua equipe.

Ao final da entrevista não estruturada, a percepção do pesquisador foi de que o diretor da usina mostrou muito interesse em apoiar a pesquisa uma vez que o mesmo estava muito ocupado com a sua agenda de obrigações do seu dia-a-dia na usina e que muitas decisões eram tomadas por meio do seu próprio conhecimento ou com ajuda do conhecimento dos gerentes. Assim, o desenvolvimento de um SAD poderia auxiliá-lo consideravelmente a melhorar a sua tomada de decisão.

Para o pesquisador, o diretor da usina é considerado peça fundamental para o sucesso da pesquisa, pois ele seria o patrocinador de todo o projeto e na visão do pesquisador isso é muito importante na implantação de qualquer sistema de informação ou projeto dentro de uma organização.

A metodologia de pesquisa-ação escolhida para o desenvolvimento do estudo foi a descrita por THIOLENT (1997). Assim o trabalho de implantação da pesquisa foi dividido em quatro fases, descritas a seguir:

5.2.1 - Primeira Fase da Pesquisa-ação: Fase Exploratória

Na primeira fase da pesquisa-ação, o objetivo é identificar os atores e realizar o diagnóstico para visualizar, de forma mais clara, os problemas, as capacidades de ação, e a intervenção na organização.

Para o desenvolvimento desta fase foram realizadas mais duas entrevistas não estruturadas com o diretor agrícola da usina. As duas entrevistas tiveram duração de aproximadamente duas horas cada. Logo no início da primeira entrevista foi abordada, novamente, a situação de dificuldades que a empresa vivia com relação à falta de informações gerenciais e nesta mesma reunião foi definida a equipe que participaria efetivamente da pesquisa, conforme é demonstrado a seguir:

1. Diretor Agrícola
2. Gerente Agrícola
3. Escriturário Administrativo
4. Auxiliar Administrativo

Tanto o escriturário administrativo como o auxiliar administrativo ficaram com a tarefa de replicar as informações levantadas e discutidas nas reuniões para o restante dos funcionários da área agrícola da usina. O escriturário administrativo juntamente com o auxiliar administrativo fariam o papel do analista de negócio dentro do projeto e seriam responsáveis em determinar qual o conjunto de informações que deveria ser levado para o novo repositório de dados *DW*. Este conjunto é o que será utilizado para a tomada de decisões.

Outro dado importante para o projeto é obter dos analistas de negócios que fornecem informações aos gerentes, o nível de agregação que deve ser utilizado, a frequência e periodicidade de carga dos dados no *DW*.

Na segunda entrevista, ficou definido também a pedido do diretor agrícola, uma reunião semanal de aproximadamente duas horas com a equipe do projeto (Diretor, Gerente, Escriturário e Auxiliar Administrativo) e uma reunião mensal com todos os integrantes da área agrícola da empresa, com a finalidade de mostrar o andamento da implantação do novo sistema, a importância de cada um dentro do novo contexto como também forçar a mudança cultural na organização que seria exigida com a implantação do novo sistema de informação.

As mudanças tecnológicas, principalmente quando é vislumbrada uma nova filosofia de organização do trabalho, informação e da empresa, requerem mudanças comportamentais como fator crítico de sucesso da tecnologia.

Mudanças de comportamento como um fenômeno organizacional tomam tempo e não são espontâneas, requerem a definição de uma estratégia clara, que leve em consideração vários fatores, além de conscientização do usuário para a necessidade pelo engajamento nas novas filosofias de trabalho baseado no conhecimento e na tecnologia, como suporte, para a sobrevivência do negócio da empresa e descoberta de novos mercados (ARAÚJO, 2005).

5.2.2 - Segunda Fase da Pesquisa-Ação: Pesquisa Aprofundada

Na segunda fase da pesquisa-ação, o objetivo foi a realização da pesquisa aprofundada. Nesta fase, continuou a ocorrer a coleta de dados para a pesquisa, principalmente por intermédio de entrevistas não estruturadas e seminários juntamente com a equipe do projeto.

Nesta fase também, começou a ser desenvolvido o modelo de estrutura para armazenagem das informações que seriam extraídas dos sistemas operacionais agrícolas da empresa, ou seja, o *DW* propriamente dito.

Para o desenvolvimento desta fase, foi necessário identificar em quais sistemas existentes na empresa estariam localizadas as informações que seriam extraídas para o novo ambiente projetado, como também o modo de armazenamento que essas informações se encontravam (Ex: Banco de Dados, Arquivos DBF, Arquivos TXT, entre outras).

A identificação dos dados operacionais existentes nos sistemas da usina, foi passada pelos próprios membros da equipe (Escriturário e Auxiliar Administrativo), que atuando como analistas de negócio, conheciam muito bem o sistema e tinham forte relacionamento com a área de tecnologia de informação da usina, facilitando e muito o trabalho do pesquisador.

Os dados (tabelas e atributos) foram passados com detalhes, já que a informação se encontrava armazenada em arquivos DBF (*Data Base File*).

O próximo passo efetuado pelo pesquisador, seria identificar como e qual metodologia utilizar para extrair esses dados dos sistemas da usina.

Um plano de migração e a metodologia utilizada foram descritos. O plano de migração endereça as questões referentes à passagem de dados do ambiente de sistemas existentes na usina para o ambiente de *DW*.

Toda a fase de construção, transformação e carregamento dos dados para o novo ambiente projetado, foram seguindo os passos de KIMBALL (1998), e serão demonstrados nas próximas seções deste mesmo capítulo.

5.2.3 - Terceira e Quarta Fase da Pesquisa-ação: Ação e Avaliação

Na terceira fase da pesquisa-ação, a partir dos resultados da fase anterior, foi realizado o planejamento das ações e discussões de objetivos alcançáveis por meio de planos de ação. A quarta fase refere-se à avaliação da pesquisa. Esta fase consiste da observação, redirecionamento das ações e resgate do conhecimento adquirido durante o processo da pesquisa.

Considerando a falta de informações gerenciais na empresa e a conseqüente necessidade da mesma de selecionar um subconjunto de objetivos estratégicos propostos como alvo de suas ações, o pesquisador visando adaptar a implantação do SAD para a usina, propôs a elaboração de priorização dos relatórios estratégicos de acordo com as informações consideradas mais importantes e relevantes relatados pelo Diretor e Gerente Agrícola, ou seja, os relatórios seriam desenvolvidos com a prioridade de atender onde a informação para uma eventual tomada de decisão se mostrava mais carente segundo o diretor da usina.

5.3 - Metodologia (Construindo o *DW* Dimensional)

Construir um *DW* é um processo de combinar as necessidades de informações de uma comunidade de usuários com os dados que realmente estão disponíveis. Este projeto de banco de dados fundamentou-se em nove pontos básicos de decisão segundo KIMBALL (1998), direcionados pelas necessidades do usuário e pelos dados disponíveis.

O próximo passo é apresentar uma metodologia para formular perguntas adequadas à descrição detalhada do ambiente de uma organização, permitindo assim o desenvolvimento de um projeto de *DW* satisfatório.

Deve-se ressaltar que essa metodologia não consiste em abordagens pré-formuladas que podem ser aplicadas indiscriminadamente a qualquer organização. Há áreas importantes do projeto que exigem análise rigorosa. Deve sempre atender às necessidades mais importantes da organização de forma eficiente e se o *DW* que está sendo construído é simples o suficiente para ser utilizado pelos usuários e pelo *software*.

5.4 - Os Nove Pontos de Decisão

Este capítulo enfocará as técnicas para coletar os dados necessários para atender a esses nove requisitos. Esses pontos são definidos como um princípio de projeto (KIMBALL, 1998).

Os nove pontos de decisão de um projeto de um banco de dados completo para um *DW* consistem em identificar:

1. Os processos, a identidade das tabelas de fatos;
2. A granularidade (nível de detalhe) de cada tabela de fatos;
3. As dimensões de cada tabela de fatos;
4. Os fatos, incluindo os fatos pré-calculados;
5. Os atributos da dimensão com descrições completas e terminologia apropriada;
6. Como rastrear dimensões de modificação lenta;
7. Os agregados, dimensões heterogêneas, minidimensões, modos de consulta e outras decisões de armazenamento físico;
8. A amplitude de tempo do histórico do banco de dados;
9. Os intervalos em que os dados são extraídos e carregados no *DW*.

Recomenda-se que essas nove decisões sejam tomadas na ordem apresentada. Essa metodologia é do tipo *top-down* (de cima para baixo) porque inicia identificando os principais processos da empresa em que os dados serão coletados. A visão macro dos processos agrícolas que estão disponíveis em uma Usina de Açúcar e Alcool e as suas respectivas definições foi demonstrada no capítulo 4 deste trabalho. Eles representam as verdadeiras fontes de extração de dados agrícolas nessa empresa. Observe que se inicia por fontes consistentes de extração de dados, sem perder tempo com fontes de dados que não existem. Nossa função como projetistas de *DW* é iniciar por fontes de dados existentes, conforme demonstrado a seguir (KIMBALL, 1998).

Primeiro Ponto: Os Processos Agrícolas

- Preparo do Solo
- Plantio de Cana
- Tratos Culturais
- Tratos Culturais (Soqueira)
- Colheita da Cana

Após a identificação dos processos, será construída a tabela de fatos baseada no processo selecionado. Antes da estrutura da tabela de fatos, é importante verificar qual será seu nível de detalhe. Isto é, a granularidade da tabela. Níveis de detalhe típicos incluem uma transação individual, uma linha de item, um instantâneo diário ou mensal (KIMBALL, 1998).

Quando se conhece a granularidade (nível de detalhe) de uma tabela de fatos, é possível identificar as dimensões e seus respectivos níveis de detalhe. Poderão existir dimensões adicionais que não são estritamente necessárias para a decisão da granularidade da tabela de fatos.

Na tabela de fatos Produção, após as dimensões tempo, fornecedor, fundo, fazenda e talhão, adicionar as outras dimensões descritivas úteis que constam na produção.

Essas dimensões de suporte à decisão incluem o tipo de solo, o tipo de variedade e o tipo de estágio. A adição dessas dimensões não altera o número de registros da tabela de fatos e elas não são necessárias para definir uma chave primária, mas podem tornar a tabela de fatos mais poderosa. Cada uma das dimensões pode ser considerada como ponto de entrada, como mostra a Figura 44.

Segundo Ponto: A granularidade da Tabela de Fatos Produção.

A granularidade do nosso projeto poderia ser definida de três maneiras:

1. Número de Viagens dos Caminhões;
2. Agrupado por dia;
3. Agrupado por mês.

Feita a análise simples, o próximo passo consiste em definir exatamente qual será o nível de granularidade. A criação de um nível levemente resumido de dados que estão em um nível muito baixo de detalhes não faz sentido porque uma quantidade excessiva de recursos será necessária para processar os dados. A criação de um nível levemente resumido de detalhe muito alto significa que muita análise terá que ser feita no nível histórico. Portanto, a primeira intervenção sobre o nível levemente resumido de granularidade implica elaborar uma conjectura respaldada em certa quantidade de conhecimento e experiência. Para definir o nível de granularidade adequado aos dados levemente resumidos, o único caminho possível é a colocação destes frente ao usuário final.

Só depois de o usuário final ter efetivamente visto os dados é que poderá ser dada uma resposta definitiva.

Foi utilizada na pesquisa a segunda opção (Agrupado por Dia), pois na primeira opção (Número de Viagens por Caminhão) o nível de granularidade era muito baixo, sendo desnecessário a quantidade de dados com que seria povoado o banco de dados, já que no nosso projeto a segunda opção já atenderia as nossas consultas com boa performance e informações para a tomada de decisão. Já a terceira opção foi descartada, por não atender as necessidades reais da usina.

Terceiro Ponto: As Dimensões da Tabela de Fatos Produção

Figura 44 - A tabela de fatos produção com as dimensões identificadas.

Os relacionamentos são associações entre duas ou mais entidades (tabelas), que representam um fato ou uma solução do mundo real.

Geralmente, os tipos de relacionamentos sofrem certas restrições que limitam as possíveis combinações das entidades participantes. A cardinalidade indica o número de relacionamentos dos quais uma entidade pode participar.

- **1:1 (um para um)**

Indica que as tabelas têm relação unívoca entre si. Você escolhe qual tabela vai receber a chave estrangeira.

- **1:N (um para muitos)**

A chave primária da tabela que tem o lado 1 vai para a tabela do lado N. No lado N ela é chamada de chave estrangeira.

- **N:N (muitos para muitos)**

Quando tabelas têm entre si relação N:N, é necessário criar uma nova tabela com as chaves primárias das tabelas envolvidas, ficando assim uma chave composta, ou seja, formada por diversos campos-chave de outras tabelas. A relação então se reduz para uma relação 1:N, sendo que o lado N ficará com a nova tabela criada.

Uma chave estrangeira é um campo, que aponta para a chave primária de outra tabela. Ou seja, passa a existir uma relação entre essas duas tabelas. A finalidade da chave estrangeira é garantir a integridade dos dados referenciais, pois apenas serão permitidos valores que supostamente vão aparecer na Base de Dados.

Figura 45 - Relacionamento Tabelas (FORNECEDOR x FATO).

Figura 46 - Relacionamento Tabelas (TIPO x FATO).

Figura 47 - Relacionamento Tabelas (ESTÁGIO x FATO).

Figura 48 - Relacionamento Tabelas (VARIEDADE x FATO).

Figura 49 - Relacionamento Tabelas (TALHAO x FATO).

Figura 50 - Relacionamento Tabelas (FAZENDA x FATO).

Figura 51 - Relacionamento Tabelas (FUNDO x FATO).

Figura 52 - Relacionamento Tabelas (SOLO x FATO).

Figura 53 - Relacionamento Tabelas (TEMPO x FATO).

A escolha das dimensões é o passo-chave do projeto. Decidiu-se sobre a identidade das dimensões antes de nos preocupar excessivamente com o local em que iremos buscar todos os atributos dimensionais. Depois de selecionarmos as dimensões, nossas próximas tarefas serão descarregar todos os fatos mensuráveis na tabela de fatos conforme mostra a figura 54 e finalmente concluir o preenchimento dos registros da dimensão (KIMBALL, 1998).

Quarto Ponto: Os Fatos, incluindo os Fatos Pré-Calculados.

Figura 54 - A tabela de fatos produção com todos os fatos identificados.

Quinto Ponto: Os atributos da dimensão com descrições completas e terminologias apropriadas

Tabela 5 - Dimensão Tempo.

	Column Name	Data Type	Length	Allow Nulls
▶	DIA	varchar	2	
▶	MES	varchar	2	
▶	ANO	varchar	4	
▶	SAFRA	varchar	5	

Principais atributos da Dimensão Tempo:

- ✓ Dia
- ✓ Mês
- ✓ Ano
- ✓ Safra

Tabela 6 - Dimensão Fornecedor de Cana.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_FORN	varchar	6	
▶	SAFRA_FORN	varchar	5	
	NOME_FORN	varchar	30	✓
	END_FORN	varchar	30	✓
	BAIRRO_FORN	varchar	20	✓
	CEP_FORN	varchar	10	✓
	CIDADE_FORN	varchar	20	✓
	EST_FORN	varchar	2	✓

Principais atributos da Dimensão Fornecedor de Cana:

- ✓ NOME_FORN → Nome do Fornecedor
- ✓ CIDADE_FORN → Cidade do Fornecedor
- ✓ EST_FORN → Estado onde se localiza a Fazenda do Fornecedor

Tabela 7 - Dimensão Solo.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_SOLO	varchar	2	
▶	SAFRA_SOLO	varchar	5	
	DESC_SOLO	varchar	10	✓

Principais atributos da Dimensão Solo:

- ✓ DESC_SOLO → Descrição do Tipo de Solo.

Tabela 8 - Dimensão Fundo Agrícola.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_FUNDO	varchar	5	
▶	SAFRA_FUNDO	varchar	5	
	NOME_FUNDO	varchar	30	✓
	CIDADE_FUNDO	varchar	20	✓
	EST_FUNDO	varchar	2	✓
	INS_EST_FUNDO	varchar	15	✓
	CGC_FUNDO	varchar	20	✓
	COD_ORIG_FUNDO	varchar	2	✓
	DESC_ORIG_FUNDO	varchar	30	✓

Principais atributos da Dimensão Fundo Agrícola:

- ✓ NOME_FUNDO → Nome do Fundo Agrícola
- ✓ CIDADE_FUNDO → Cidade onde se localiza o Fundo Agrícola

Tabela 9 - Dimensão Variedade de Cana.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_VAR	varchar	2	
▶	SAFRA_VAR	varchar	5	
	DESC_VAR	varchar	10	✓

Principais atributos da Dimensão Variedade de Cana:

- ✓ DESC_VAR → Descrição da Variedade de Cana

Tabela 10 - Dimensão Fazenda.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_FAZ	varchar	5	
▶	SAFRA_FAZ	varchar	5	
	NOME_FAZ	varchar	30	✓

Principais atributos da Dimensão Fazenda:

- ✓ NOME_FAZ → Nome da Fazenda

Tabela 11 - Dimensão Talhão de Cana.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_TALHAO	varchar	6	
?	COD_FUNDO	varchar	5	
?	SAFRA_TALHAO	varchar	5	
	DIST_TALHAO	float	8	✓
	AREA_TOTAL_TALHAO	float	8	✓
	AREA_CORTE_TALHAO	float	8	✓
	AREA_OLHADURA	float	8	✓
	ESTAG_TALHAO	varchar	2	✓
	SOLO_TALHAO	varchar	2	✓
	VAR_TALHAO	varchar	2	✓
	PROD_ESTIM_TON	float	8	✓
	COD_SCOL	varchar	2	✓
	DT_CORT_ANT	datetime	8	✓
	DIST_TERR	varchar	1	✓
	AREA_NC	float	8	✓

Principais atributos da Dimensão Talhão:

- ✓ DIST_TALHAO → Distância do Talhão à Usina
- ✓ AREA_TOTAL_TALHAO → Área Total do Talhão
- ✓ AREA_CORTE_TALHAO → Área de Corte do Talhão
- ✓ AREA_OLHADURA → Área de Olhadura (Cana p/ Plantio) do Talhão
- ✓ ESTAG_TALHAO → Estágio de Corte do Talhão
- ✓ SOLO_TALHAO → Tipo de Solo do Talhão
- ✓ VAR_TALHAO → Tipo de Variedade de Cana do Talhão
- ✓ PROD_ESTIM_TON → Produção Estimada de Toneladas de Cana
- ✓ COD_SCOL → Tipo de Cana do Talhão
- ✓ DT_CORT_ANT → Data do Corte Anterior
- ✓ AREA_NC → Área não Cultivada do Talhão

Tabela 12 - Dimensão Estágio de Corte.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_ESTAG	varchar	2	
?	SAFRA_ESTAG	varchar	5	
	DESC_ESTAG	varchar	30	✓
	DESC_ESTAG_RED	varchar	10	✓
	GRUPO_ESTAG	varchar	15	✓
	GRUPO_ESTAG_1	varchar	15	✓

Principais atributos da Dimensão Estágio de Corte:

- ✓ DESC_ESTAG → Descrição do Estágio
- ✓ GRUPO_ESTAG → Descrição do Grupo de Estágio de Corte

Tabela 13 - Dimensão Tipo de Cana.

	Column Name	Data Type	Length	Allow Nulls
▶	COD_TIPO	varchar	2	
▶	SAFRA_TIPO	varchar	5	
	GRUPO_TIPO	varchar	15	✓
	DESC_TIPO	varchar	20	✓
	DESC_TIPO_REDUZ	varchar	3	✓
	CLASSIF_CANA	varchar	4	✓

Principais atributos da Dimensão Tipo de Cana:

- ✓ GRUPO_TIPO → Grupo do Tipo de Cana (Manual ou Mecânica)
- ✓ CLASSIF_CANA → Classificação Tipo de Cana (Crua ou Queimada)

Tabela 14 - Fatos Produção.

	Column Name	Data Type	Length	Allow Nulls
▶	DIA	varchar	2	
▶	MES	varchar	2	
▶	ANO	varchar	4	
▶	SAFRA	varchar	5	
▶	COD_FORN	varchar	6	
▶	COD_FUNDO	varchar	5	
▶	COD_FAZ	varchar	5	
▶	COD_TALHAO	varchar	6	
▶	COD_SOLO	varchar	2	
▶	COD_VAR	varchar	2	
▶	COD_ESTAG	varchar	2	
▶	SETOR	varchar	2	
▶	COD_TIPO	varchar	2	
	PROD_TON	float	8	✓
	FIBRA_POR_TON	float	8	✓
	PC_POR_TON	float	8	✓
	POL_POR_TON	float	8	✓
	KG_ATR_POR_TON	float	8	✓
	TOTAL_KG_ATR	float	8	✓
	DT_FECHTO	datetime	8	✓
	CANA_R\$_POR_TON	float	8	✓
	FATUR_MEDIO_CANA_R\$	float	8	✓
	REND_MEDIO_UNICOPS_POR_TON	float	8	✓
	PROD_TOTAL_UNICOPS	float	8	✓
	MIX_ACUCAR	float	8	✓
	PROD_ACUCAR_SACOS_50KG	float	8	✓
	PROD_ATR_ACUCAR	float	8	✓
	PRECO_ACUCAR_R\$_POR_UNICOP	float	8	✓
	FATUR_ACUCAR_R\$	float	8	✓
	MIX_ANIDRO	float	8	✓
	PROD_ANIDRO_M3	float	8	✓
	PROD_ATR_ANIDRO	float	8	✓
	PRECO_ANIDRO_R\$_POR_UNICOP	float	8	✓
	FATUR_ANIDRO_R\$	float	8	✓
	MIX_HIDRATADO	float	8	✓
	PROD_HIDRATADO_M3	float	8	✓
	PROD_ATR_HIDRATADO	float	8	✓
	PRECO_HIDRATADO_R\$_POR_UNICOP	float	8	✓
	FATUR_HIDRATADO_R\$	float	8	✓

Principais atributos da Dimensão Fatos de Produção:

- ✓ PROD_TON → Produção de Toneladas de Cana
- ✓ FIBRA_POR_TON → Fibra da Cana por Tonelada
- ✓ PC_POR_TON → Pureza da Cana por Tonelada
- ✓ POL_POR_TON → POL da Cana por Tonelada
- ✓ KG_ATR_POR_TON → Kilos de ATR por Tonelada
- ✓ TOTAL_KG_ATR → Total de Kilos de ATR
- ✓ DT_FECHTO → Data de Fechamento do Talhão
- ✓ CANA_R\$_POR_TON → Preço da Cana por Tonelada
- ✓ FATUR_MEDIO_CANA_R\$ → Faturamento Médio com Cana
- ✓ REND_MEDIO_UNICOPS_POR_TON → Rendimento Médio Unicap's
- ✓ PROD_TOTAL_UNICOPS → Produção Total de Unicap's
- ✓ MIX_ACUCAR → Mix de Produção do Açúcar
- ✓ PROD_ACUCAR_SACOS_50KG → Produção de Sacos de Açúcar
- ✓ PROD_ATR_ACUCAR → Produção de ATR de Açúcar
- ✓ PRECO_ACUCAR_R\$_POR_UNICOP → Preço do Açúcar por Unicap
- ✓ FATUR_ACUCAR_R\$ → Faturamento Médio do Açúcar
- ✓ MIX_ANIDRO → Mix de Produção do Álcool Anidro
- ✓ PROD_ANIDRO_M3 → Produção de Álcool Anidro em M³
- ✓ PROD_ATR_ANIDRO → Produção de ATR de Álcool Anidro
- ✓ PRECO_ANIDRO_R\$_POR_UNICOP → Preço Anidro por Unicap
- ✓ FATUR_ANIDRO_R\$ → Faturamento Médio do Álcool Anidro
- ✓ MIX_HIDRATADO → Mix de Produção do Álcool Hidratado
- ✓ PROD_HIDRATADO_M3 → Produção de Álcool Hidratado em M³
- ✓ PROD_ATR_HIDRATADO → Produção de ATR de Álcool Hidratado
- ✓ PRECO_HIDRATADO_R\$_POR_UNICOP → Preço Hidratado / Unicap
- ✓ FATUR_HIDRATADO_R\$ → Faturamento Médio do Álcool Hidratado

Agora que a estrutura principal do desenho lógico esta completa nossa atenção de direciona aos aspectos importantes do diagrama físico que incluem como rastrear dimensões de modificação lenta, como adicionar agregados, dimensões heterogenias, minidimensões e modos de consulta.

Sexto Ponto: Como rastrear dimensões de modificação lenta.

Na fase de construção de um *DW* considera-se quase sempre as dimensões como entidades logicamente independentes entre si. Em nosso projeto, nem todas as dimensões foram consideradas independentes em relação ao tempo. No mundo real isso é estritamente verdadeiro. Fornecedores podem sofrer mudanças a cada safra. As pessoas modificam seus nomes, casam e divorciam-se, aumentam o número de filhos e trocam de endereço.

Na qualidade de projetistas de banco de dados deve-se decidir como lidar com essas modificações. No início deste projeto, uma das principais responsabilidades do *DW* é representar de forma correta o histórico passado. Por exemplo, para uma companhia de seguros é essencial que a descrição da parte seguradora esteja corretamente representada ao longo do tempo e não apenas limitada à sua descrição atual.

Deve-se explorar o fato de que a maioria das dimensões é constante ao longo do tempo e que podemos preservar uma estrutura dimensional independente, incluindo relativamente poucas adições para capturar o aspecto temporal das modificações. As dimensões praticamente constantes são chamadas de dimensões de modificação lenta.

Quando nos deparamos com uma dimensão de modificação lenta, temos três opções fundamentais. Cada uma delas resulta em um diferente nível de tratamento de modificações ao longo do tempo (KIMBALL, 1998).

- Substituir os valores antigos dos registros da dimensão e, portanto, perder a capacidade de rastrear o histórico passado;
- Adicionar um registro à dimensão contendo os novos valores do atributo no momento da mudança para segmentar o histórico entre a descrição antiga e a nova descrição com muita precisão;

- Criar novos campos “atuais” no registro original da dimensão para incluir os novos valores do atributo, mantendo também seus valores originais, permitindo dessa forma descrever o histórico anterior e o posterior à mudança tanto em relação aos valores originais do atributo quanto aos valores atuais.

No projeto, decidiu-se utilizar a terceira opção, criando um campo chamado “safra” para incluir os novos valores do atributo, mantendo assim seus valores originais e conseqüentemente armazenando todo o histórico referente a todas as safras, representado assim o histórico passado de forma correta e consistente.

Sétimo Ponto: Os agregados, dimensões heterogenias, minidimensões, modos de consulta e outras decisões de armazenamento físico.

Agregados:

Um agregado é um registro da tabela de fatos que representa o resumo dos registros de nível básico da tabela de fatos. Um registro da tabela de fatos agregado está sempre associado a um ou mais registros de tabela de dimensões agregadas.

As tabelas de agregados de modo geral oferecem um ambiente apropriado para planejar dados. Geralmente os agregados são construídos a partir do nível básico, evoluindo até o nível mais alto de categoria de produto ou de totais de fornecedores.

Dimensões Heterogenias:

Para *DW* com dimensões que descrevem vários produtos heterogêneos (um conjunto de produtos que contêm alguns atributos ou fatos incompatíveis), é recomendado a criação de uma tabela de fatos central e uma tabela dimensional central, para possibilitar consultas gerais abrangendo os vários tipos de itens, e a criação de tabelas de fatos e tabelas dimensionais específicas para cada tipo de item, para permitir consultas com maior nível de detalhes.

Minidimensões:

A técnica eficiente para utilizar esses atributos demográficos em uma dimensão fornecedor extensa é separar um ou mais conjuntos desses atributos em minidimensões demográficas (subconjunto da dimensão-fornecedor incluídos em uma pequena dimensão separada para facilitar pesquisas rápidas e a manipulação de dimensões de modificação lenta).

Embora pela lógica, esses atributos devam permanecer na dimensão fornecedor primária, pode-se obter ganhos significativos de desempenho de minidimensões (KIMBALL, 1998).

A minidimensão demográfica pode estar ligada tanto diretamente à tabela de fatos como à própria tabela fornecedor. Para poder fornecer acesso mais eficiente a esta tabela por meio dos atributos demográficos (KIMBALL, 1998).

Modos de Consulta:

À medida que o projeto lógico e físico torna-se o foco das atenções, surge a necessidade eminente de decidir qual *hardware* e *software* serão utilizados. Os sistemas são caros e as decisões de perfil altamente técnicas. Além disso, os fornecedores não auxiliam muito no processo de decisão, pois usam terminologias diferentes e freqüentemente dirigem seus esforços para levantar dúvidas quanto aos produtos do concorrente.

Toda a comunidade de usuários será conectada ao DBMS por meio das ferramentas de consulta. Quando utilizamos uma “ferramenta de consulta”, a qualquer aplicativo-cliente que mantenha uma conexão em rede com o DBMS o propósito será de enviar instruções SQL (Linguagem Padrão para acessar Banco de Dados Relacionais) e receber conjuntos de respostas.

Isso inclui as ferramentas de consulta eventual (*ad-hoc query tools*), assim como os geradores de relatórios e outras aplicações monolíticas no computador do usuário. A arquitetura ideal para uma ferramenta de consulta envolve o envio de instruções SQL relativamente pequenas para o banco de dados, ou DBMS, (todas as consultas de *DW* qualificam-se como “pequenas”) e o recebimento de conjunto de respostas relativamente pequenos que contêm desde poucas centenas e até poucos milhares de linhas de dados. À medida que os fornecedores de banco de dados relacionais enfocarem com mais eficiência os aspectos de cubo de dados (nome dado a um banco de dados dimensional, geralmente referindo-se a um caso simples de produto, mercado e tempo) dos *Data Warehouse* Dimensionais (*DDW*), serão capazes de avaliar a organização dos dados de uma consulta antes que ela seja realizada para estimar o tempo de duração de forma mais precisa (KIMBALL, 1998).

Oitavo Ponto: A amplitude de tempo do histórico do banco de dados.

A seguir, o planejamento do tempo de duração do banco de dados. Na maioria dos casos fica em aberto a possibilidade da idéia de permitir que o histórico fique acumulado além do que havíamos planejado inicialmente (KIMBALL, 1998).

Apesar de os dados encontrados no ambiente de SAD serem históricos e apresentarem, necessariamente, uma baixa probabilidade de acesso, eles, contudo, apresentam alguma probabilidade de acesso (de outra forma eles nem precisavam ser armazenados). Quando a probabilidade de acesso chega a zero (ou se aproxima de zero), os dados precisam ser eliminados. Uma vez que o volume de dados é uma questão candentes do ambiente de *DW*, a eliminação de dados que já não são úteis constitui um dos aspectos mais importantes do ambiente de *DW*.

Nono Ponto: Os intervalos em que os dados são extraídos e carregados no *DW*.

Finalmente o planejamento do intervalo de extração dos dados. Embora tenhamos definido no segundo passo um nível de detalhe de captura mensal dos dados de todas as transações, existe a opção de atualizar esse instantâneo todos os dias a partir do dia anterior ou talvez semanalmente (KIMBALL, 1998).

Como política, qualquer lapso de tempo menor do que 24 horas deve ser questionado. Normalmente, se for necessário existir um lapso de tempo de menos de 24 horas, isso é um sinal de que o desenvolvedor está criando requisitos operacionais dentro do *DW*.

O fluxo de dados pelo ambiente de *DW* deve sempre ser um processo de puxar, pelo qual os dados são puxados para o *DW* quando necessário, em vez de serem empurrados para dentro do ambiente de *DW* quando disponíveis.

5.5 - A Fase de Carregamento

Na fase de carregamento, todo ou parte do *DW* permanece *off-line* enquanto os novos dados são carregados. Os dados novos são quase sempre um extrato de um ou mais sistemas de produção. A equipe de projetistas de *DW* deve construir o sistema de extração de dados de produção que fornecerá tais dados (KIMBALL, 1998).

5.6 - O Sistema de Extração de Dados de Produção

A equipe de projetistas do *DW* deve desenvolver um *software* e um sistema de gerenciamento que execute as 11 etapas a seguir, ou a maior parte delas, para que seja possível extrair os dados de um banco de dados de produção e enviá-los para o *DW*. Esse sistema chama-se sistema de extração de dados de produção (KIMBALL, 1998).

Para melhor entendimento serão descritas a seguir cada uma das etapas em que estão envolvidos os processos de desenvolvimento e implementação do projeto proposto.

As 11 etapas da extração diária de dados de produção são os seguintes:

1. Extração primária (ler o formato legado);
2. Identificação dos registros modificados;
3. Generalização de chaves para dimensões em modificação;
4. Transformação em imagens de registro de carga;
5. Migração do sistema legado para o sistema de *DW*;
6. Classificação e construção de agregados;
7. Generalização de chaves para agregados;
8. Carregamento;
9. Processamento de exceções;
10. Garantia de qualidade;
11. Publicação.

Primeira Etapa: Extração Primária

O primeiro desafio é o acesso aos dados legados, que pode variar de muito fácil a muito difícil. Se os dados de produção forem mantidos por um sistema de banco de dados relacional orientado para transação, então sua leitura será muito fácil. Várias ferramentas e aplicações podem usar SQL para extrair dados em uma ampla variedade de formatos para a próxima etapa do processo.

A importação dos dados dos ambientes existentes agilizam todo o processo, pois evitam uma possível digitação das informações.

Tanto os dados de produção para minha tabela de fatos como os dados que povoaram minhas tabelas de dimensão são mantidos em um sistema no qual as informações estão armazenadas em arquivos do tipo *Data Base File* (DBF), o que torna fácil sua leitura.

A figura 55 mostra como é feita a leitura desses dados utilizando a ferramenta do Excel chamada *Microsoft Query*.

Figura 55 - Leitura dos Dados de Produção para Tabela de Fatos.

A figura 56 mostra a leitura dos dados do cadastro de Fundos Agrícolas que irá povoar a minha tabela DIMENSAO_FUNDO_AGRÍCOLA.

The screenshot shows the Microsoft Query interface with a table titled 'Consulta de Visual FoxPro Tables'. The table contains the following data:

codfundo	nomefundo	cidadefun	estfun	inscprod	inccest	
002.6	FAZ FLORESTA - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
003.9	FAZ FRUTAL - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
004.2	FAZ SANTA CANDIDA - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
005.5	FAZ SANTA HELENA - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
006.8	SITIO TANQUE ALTO	JABOTICABAL	SP	0391.00024/000	.	.
007.1	FAZ SAO LUIZ - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
008.4	FAZ SAO PAULO - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
009.7	FAZ SANTO ANT. LAGOA - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
010.7	FAZENDA DA SERRA	BRODOSQUI	SP	.	/	.
011.0	FAZENDA GRANADA	PATROCINIO PAULISTA	SP	.	/	.
012.3	FAZ SANTANA DAS POSSES - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
013.6	SANTO ANTONIO II - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
014.9	FAZ SAO LUIZ LOPES - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
015.2	SITIO MOINHO DE VENTO - IR	RIBEIRAO PRETO	SP	.	582.423.651.117	71.32
016.5	FAZENDA MAGNOLIA - REC	JARDINOPOLIS	SP	.	399.018.187.117	71.32
017.8	FAZ TAMBURY - SA	SERTAOCZINHO	SP	.	664.002.905.119	71.32
018.1	FAZ SANTA IZAUARA - IR	RIBEIRAO PRETO	SP	.	582.423.651.117	71.32
019.4	FAZ ARUAMA - IR	RIBEIRAO PRETO	SP	.	582.423.651.117	71.32
020.4	FAZ BOA ESPERANCA - IR	RIBEIRAO PRETO	SP	.	582.423.651.117	71.32
021.7	FAZ IRACEMA - IR	RIBEIRAO PRETO	SP	.	582.423.651.117	71.32
022.0	FAZ IGUAPE - IG	DUMONT	SP	.	294.002.360.110	71.32
023.3	SITIO LUIZA - IG	DUMONT	SP	.	294.002.360.110	71.32

Figura 56 - Leitura dos Dados do Cadastro de Fundos Agrícolas.

A figura 57 mostra a leitura dos dados do cadastro de Fazendas que irá povoar a minha tabela DIMENSAO_FAZENDA.

The screenshot shows the Microsoft Query interface with a table titled 'Consulta de Visual FoxPro Tables'. The table contains the following data:

codfaz	nomefaz
002.6	FAZ FLORESTA - SA
003.9	FAZ FRUTAL - SA
004.2	FAZ SANTA CANDIDA - SA
005.5	FAZ SANTA HELENA - SA
007.1	FAZ SAO LUIZ - SA
008.4	FAZ SAO PAULO - SA
009.7	FAZ SANTO ANT. LAGOA - SA
010.7	FAZ SERRA
011.0	FAZ GRANADA
012.3	FAZ SANTANA DAS POSSES - SA
013.6	SANTO ANTONIO II - SA
014.9	FAZ SAO LUIZ LOPES - SA
015.2	SITIO MOINHO DE VENTO - IR
016.5	FAZ MAGNOLIA
017.8	FAZ TAMBURY - SA
018.1	FAZ SANTA IZAUARA - IR
019.4	FAZ ARUAMA - IR
020.4	FAZ BOA ESPERANCA - IR
021.7	FAZ IRACEMA - IR
022.0	FAZ IGUAPE - IG
023.3	SITIO LUIZA - IG
024.6	FAZ DOM. JESUS - REC

Figura 57 - Leitura dos Dados do Cadastro de Fazendas.

A figura 58 mostra a leitura dos dados do cadastro de Estágios de Corte que irá povoar a minha tabela DIMENSAO_ESTAGIO_DE_CORTE.

codest	descest	resuest
00	CANVA PLANTA 18 MES CP	
01	PRIMEIRO CORTE	1C
02	SEGUNDO CORTE	2C
03	TERCEIRO CORTE	3C
04	QUARTO CORTE	4C
05	QUINTO CORTE	5C
06	SEXTO CORTE	6C
07	SETIMO CORTE	7C
08	OUTAVO CORTE	8C
09	NONO CORTE	9C
10	DECIMO CORTE	10C
11	DECIMO PRIMEIRO CO	11C
12	DECIMO SEGUNDO CO	12C
13	DECIMO TERCEIRO CO	13C
14	DECIMO QUARTO CO	14C
21	PRIMEIRO CORTE DE 1C ANO	
22	SEGUNDO CORTE DE 2C ANO	
23	TERCEIRO CORTE DE 3C ANO	
24	QUARTO CORTE DE 4C ANO	
25	QUINTO CORTE DE 5C ANO	
26	SEXTO CORTE DE 6C ANO	
27	SETIMO CORTE DE 7C ANO	

Figura 58 - Leitura dos Dados do Cadastro de Estágios de Corte.

A figura 59 mostra a leitura dos dados do cadastro de Fornecedores que irá povoar a minha tabela DIMENSAO_FORNECEDOR_DE_CANA.

codfor	nomefor	ender	bairro	cep	cidade
0001.3	USINA SANTO ANTONIO S/A	FAZENDA SANTO ANTONIO	CAMPINHO	14160-000	SE
0002.6	USINA SANTO ANTONIO S/A	FAZENDA IRACEMA		14100-000	RI
0003.9	USINA SANTO ANTONIO S/A	FAZENDA IGUAPE		14120-000	DI
0004.2	USINA SANTO ANTONIO S/A	FAZENDA RECREIO		14680-000	JA
0005.5	ABONDIO DAVID - ESPOLIO	SITIO LIMOEIRO		14160-000	SE
0006.8	ADELINO JACOMINI	FAZENDA IBIETE		14680-000	JA
0007.1	USINA SANTO ANTONIO S/A	FAZENDA DA SERRA		14160-000	BF
0008.4	USINA SANTO ANTONIO S/A	FAZENDA GRANADA		14160-000	PA
0009.7	ALACERIO SESTARI E OUTRO	FAZENDA BOM JESUS		14680-000	JA
0010.7	ALBERTO JOSE SICCHIERI	SITIO PAINEIRINHA		14160-000	SE
0011.0	ALBINO MANTOVANI E OUTROS	SITIO MOREIRA		14160-000	SE
0012.3	ALBINO ZAMPIONI	SITIO SANTA MARIA		14160-000	SE
0013.6	ALCIDES TROVO	SITIO SANTA MARIA		14160-000	SE
0014.9	ALI ALVES DE PAULA E OUTRO	SITIO SANTA TEREZINHA		14160-000	SE
0015.2	AMABILE GUINDALINE STRINI OTS	FAZENDA SANTA ROSA		14160-000	SE
0016.5	ANGELO ANTONIO PIGNATA E OTROS	FAZENDA SAO JOSE		14160-000	SE
0017.8	ANGELO MARTINUCCI E OUTROS	SITIO STO ANTONIO DAS PIMENTAS		14160-000	SE
0018.1	ANGELO PALMIERI - ESPOLIO	SITIO MARAGOGIPE		14160-000	SE
0019.4	ANGELO TROVO	SITIO SANTA MARIA		14160-000	SE
0020.4	ANTENOR MONTENEGRO E OUTROS	SITIO SANTO ANTONIO		14120-000	DI
0021.7	ANTONIA VILALBA LOPES	SITIO SANTA IZABEL		14160-000	SE
0022.0	ANTONIO ARDENGI E OUTRO	SITIO SANTO ANTONIO		14160-000	SE

Figura 59 - Leitura dos Dados do Cadastro de Fornecedores.

A figura 60 mostra a leitura dos dados do cadastro de Solos que irá povoar a minha tabela DIMENSAO_SOLO.

The screenshot shows the Microsoft Query interface. The table 'casolo' is selected, and its fields are listed in the left pane: codsolo, descsolo, and resuso. The main pane displays a table with 22 records. The first record is highlighted.

	codsolo	descsolo	resuso
1	01	A-12	A-12
2	02	GH-1	GH-1
3	03	GPH-1	GPH-1
4	04	GPH-2	GPH-2
5	05	GPH-6	GPH-6
6	06	HI	HI
7	07	LI-1	LI-1
8	08	LR-1	LR-1
9	09	LR-2	LR-2
10	10	LR-2A	LR-2A
11	11	LR-3A	LR-3A
12	12	LR-4	LR-4
13	13	LR-5	LR-5
14	14	LVA-1	LVA-1
15	15	LVA-2	LVA-2
16	16	LVA-2A	LVA-2A
17	17	LVA-5	LVA-5
18	18	LVA-6	LVA-6
19	19	LVA-7	LVA-7
20	20	LVA-18	LVA-18
21	21	LVA-18A	LVA-18A
22	22	LVA-19	LVA-19

Figura 60 - Leitura dos Dados do Cadastro de Solos.

A figura 61 mostra a leitura dos dados do cadastro de Variedades de Cana que irá povoar a minha tabela DIMENSAO_VARIEDADE_DE_CANA.

The screenshot shows the Microsoft Query interface. The table 'cavari_02' is selected, and its fields are listed in the left pane: codvar, descvar, and resuvar. The main pane displays a table with 22 records. The first record is highlighted.

	codvar	descvar	resuvar
1	01	CB41-14	CB41-14
2	02	CB41-76	CB41-76
3	03	CB47-355	CB47-355
4	04	CB53-99	CB53-99
5	05	CB46-47	CB46-47
6	06	CB56-126	CB56-126
7	07	CO413	CO413
8	08	CO775	CO775
9	09	IAC52-150	IAC52-150
10	10	IAC58-480	IAC58-480
11	11	NA56-79	NA56-79
12	12	RB735275	RB735275
13	13	RB725147	RB725147
14	14	RB725828	RB725828
15	15	SP70-1078	SP70-1078
16	16	SP70-1143	SP70-1143
17	17	SP70-1284	SP70-1284
18	18	SP71-1406	SP71-1406
19	19	SP70-1423	SP70-1423
20	20	SP70-3370	SP70-3370
21	21	SP71-6163	SP71-6163
22	22	SP71-799	SP71-799

Figura 61 - Leitura dos Dados do Cadastro de Variedades de Cana.

A figura 62 mostra a leitura dos dados do cadastro de Talhões que irá povoar a minha tabela DIMENSAO_TALHAO.

codfundo	codtalh	areatot	areacor	codsolo	distast	codvar	codest	dtplant	dtport	dtcant	dtc
002.6	0001.9	16.17	16.17	09	7.5	31	07	1995-05-30	1997-08-09	2001-10-27	1899-1
002.6	0002.2	15.05	15.05	09	7.5	31	07	1994-05-30	1994-05-30	2001-11-01	2002-0
002.6	0003.5	16.6	16.6	09	7.5	31	07	1995-05-30	1994-05-12	2001-11-01	1899-1
002.6	0004.9	10.62	10.62	09	7.5	31	07	1995-05-30	1994-06-11	2001-11-26	1899-1
002.6	0005.1	14.81	14.81	08	7.5	31	07	1995-05-30	1994-10-26	2001-11-22	1899-1
002.6	0006.4	11.95	11.95	09	7.5	30	08	1994-01-26	1996-11-10	2001-11-26	1899-1
002.6	0007.7	15.66	15.66	08	7.5	31	07	1995-05-30	1994-11-01	2001-10-27	1899-1
002.6	0008.0	21.78	21.78	08	7.5	31	07	1995-03-02	1994-11-01	2001-10-30	1899-1
002.6	0009.3	17.16	17.16	08	7.5	31	07	1995-03-02	1994-11-01	2001-11-01	1899-1
002.6	0010.3	17.45	17.45	08	7.5	31	07	1995-03-02	1994-07-19	2001-11-27	1899-1
002.6	0011.6	13.84	13.84	09	7.5	50	03	1999-03-08	2000-09-25	2001-11-27	1899-1
002.6	0012.9	16.65	16.65	09	7.5	72	03	1999-03-08	2000-09-25	2001-10-31	1899-1
002.6	0013.2	8.82	8.82	09	7.5	72	03	1999-03-06	2001-10-30	2001-10-30	1899-1
002.6	0014.5	13.58	13.58	09	7.5	72	03	1999-03-08	2001-10-27	2001-10-27	1899-1
002.6	0015.8	18.18	18.18	09	7.5	72	03	1999-03-06	2000-09-25	2001-10-27	1899-1
002.6	0016.1	14.72	14.72	09	7.5	72	03	1999-03-05	2000-09-25	2001-10-27	1899-1
002.6	0021.3	16.09	16.09	09	7.5	30	08	1994-03-09	1996-11-10	2001-10-26	1899-1
003.9	0001.2	5.4	5.4	08	12.5	50	22	2000-11-27	2001-12-03	2001-11-26	1899-1
003.9	0002.5	11.44	11.44	08	12.5	30	01	2000-11-28	1899-12-30	1899-12-30	1899-1
003.9	0003.8	10.74	10.74	08	12.5	30	22	2000-11-28	2001-11-27	2001-12-03	1899-1
003.9	0004.1	12	12	08	12.5	55	04	1998-02-23	1999-07-27	2001-09-29	2002-0
003.9	0005.4	13.62	13.62	08	12.5	30	29	1993-09-20	1994-10-08	2001-11-18	1899-1

Figura 62 - Leitura dos Dados do Cadastro de Talhões de Cana.

A figura 63 mostra a leitura dos dados do cadastro de Tipos de Cana que irá povoar a minha tabela DIMENSAO_TIPO_DE_CANA.

COD_TIP	SAFRA	TIP	GRUPO_TIP	DESC_TIP	DESC_TIP_REI	CLASSIF_CAN
01	00/01	CORTE MANUAL	MANUAL	MA	QUEI	
01	01/02	CORTE MANUAL	MANUAL	MA	QUEI	
01	02/03	CORTE MANUAL	MANUAL	MA	QUEI	
01	03/04	CORTE MANUAL	MANUAL	MA	QUEI	
01	04/05	CORTE MANUAL	MANUAL	MA	QUEI	
01	05/06	CORTE MANUAL	MANUAL	MA	QUEI	
01	06/07	CORTE MANUAL	MANUAL	MA	QUEI	
01	07/08	CORTE MANUAL	MANUAL	MA	QUEI	
01	38/99	CORTE MANUAL	MANUAL	MA	QUEI	
01	39/00	CORTE MANUAL	MANUAL	MA	QUEI	
02	00/01	CORTE MECANICO	MECANICO	ME	QUEI	
02	01/02	CORTE MECANICO	MECANICO	ME	QUEI	
02	02/03	CORTE MECANICO	MECANICO	ME	QUEI	
02	03/04	CORTE MECANICO	MECANICO	ME	QUEI	
02	04/05	CORTE MECANICO	MECANICO	ME	QUEI	
02	05/06	CORTE MECANICO	MECANICO	ME	QUEI	
02	06/07	CORTE MECANICO	MECANICO	ME	QUEI	
02	07/08	CORTE MECANICO	MECANICO	ME	QUEI	
02	38/99	CORTE MECANICO	MECANICO	ME	QUEI	
02	39/00	CORTE MECANICO	MECANICO	ME	QUEI	
13	00/01	CORTE MECANICO	MECANICA CRUA	MEC	CRUA	
13	01/02	CORTE MECANICO	MECANICA CRUA	MEC	CRUA	

Figura 63 - Leitura dos Dados do Cadastro de Tipos de Cana.

Segunda Etapa: Identificando os Registros Modificados para uma Tabela de Fatos de Instantâneo

Extrair dados para uma tabela de fatos de instantâneo diário, semanal ou mensal envolve especificar quais foram as alterações desde a última vez que o instantâneo foi construído. Há duas formas principais para abordar esse problema. Ou as transações de produção originais ficam disponíveis e podem ser processadas pelo *DW* para a construção de um novo instantâneo ou as transações de produção não permitem um processamento prático, e o *DW* deverá determinar o que mudou no sistema de produção desde o último instantâneo, verificando os arquivos subjacentes mantidos pelo sistema de produção. Vale a pena lembrar que muitos dos sistemas de produção não são baseados em transações. Eles são simplesmente “modificados” por meio de interfaces terminais.

Em muitos casos a equipe de *DW* deve examinar os arquivos de produção para verificar quais foram as alterações. A identificação dos registros modificados nos arquivos de produção está profundamente envolvida com a extração primária. O ideal é a equipe de *DW* ver apenas o que foi modificado desde a última extração bem sucedida (KIMBALL, 1998).

Neste projeto as informações são sempre atualizadas nos ambientes transacionais antes de serem importadas para o *DW*. Isto se faz necessário, pois algumas informações afetam períodos anteriores de carregamento na mesma safra, principalmente com relação aos dados tecnológicos (FIBRA, PC, POL e ATR).

Terceira Etapa: Generalização de Chaves para Dimensões em Modificação

Para ter uma dimensão de modificação lenta, a equipe de *DW* deve criar um processo administrativo que construa novas chaves de dimensão sempre que ocorrerem modificações rastreáveis.

Existem duas alternativas principais para administrar chaves dessa forma:

1. Chaves derivadas
2. Chaves de número inteiro atribuídas sequencialmente.

Quarta e Quinta Etapa: Transformando em imagens de Registro de Carga e Migrando para o DW

Quando a extração dos dados estiver concluída, será hora de criar a imagem dos registros de carga. Isso significa simplesmente organizar os dados nas linhas e colunas apropriadas para seu carregamento direto e imediato no DW.

A figura 64 mostra como é feita essa transformação.

The screenshot shows a Microsoft Excel spreadsheet titled 'CPC_Importação_02.xls'. The data is organized in a table with the following columns: Dia, Mes, Ano, Safra, codfor, codfundo, codfaz, codtalh, codsolo, codvar, codest, sum_pesoliqui, P.Medio R\$/Ton, and Fatur.I. The table contains 19 rows of data, each representing a cargo record for a specific day, month, and year (2002).

	A	B	C	D	E	F	G	H	I	J	K	L	M	
	Dia	Mes	Ano	Safra	codfor	codfundo	codfaz	codtalh	codsolo	codvar	codest	sum_pesoliqui	P.Medio R\$/Ton	Fatur.I
1	1	10	2002	02/03	0001.3	002.6	002.6	0001.9	09	31	07	657,38	29,98	
2	1	10	2002	02/03	0001.3	003.9	003.9	0047.8	06	44	03	13,14	29,98	
3	1	10	2002	02/03	0001.3	003.9	003.9	0048.1	06	44	03	29,77	29,98	
4	1	10	2002	02/03	0001.3	003.9	003.9	0054.6	06	37	22	57,95	29,98	
5	1	10	2002	02/03	0001.3	003.9	003.9	0055.9	06	37	01	37,95	29,98	
6	1	10	2002	02/03	0001.3	008.4	008.4	0020.8	08	44	04	204,69	29,98	
7	1	10	2002	02/03	0001.3	008.4	008.4	0021.1	08	44	04	30,03	29,98	
8	1	10	2002	02/03	0001.3	008.4	008.4	0023.7	09	44	04	327,77	29,98	
9	1	10	2002	02/03	0001.3	009.7	009.7	0005.2	08	37	02	442,57	29,98	
10	1	10	2002	02/03	0001.3	009.7	009.7	0006.5	08	37	02	199,72	29,98	
11	1	10	2002	02/03	0001.3	009.7	009.7	0007.8	08	37	02	154,76	29,98	
12	1	10	2002	02/03	0002.6	021.7	021.7	0058.6	08	60	01	55,60	29,98	
13	1	10	2002	02/03	0002.6	021.7	021.7	0059.9	08	36	01	365,64	29,98	
14	1	10	2002	02/03	0002.6	021.7	021.7	0061.2	09	55	01	167,69	29,98	
15	1	10	2002	02/03	0004.2	028.8	028.8	0012.1	10	63	01	534,78	29,98	
16	1	10	2002	02/03	0004.2	028.8	028.8	0029.9	24	44	05	361,65	29,98	
17	1	10	2002	02/03	0004.2	028.8	028.8	0030.9	24	44	05	327,22	29,98	
18	1	10	2002	02/03	0004.2	033.0	033.0	0014.9	27	37	04	90,13	29,98	
19	1	10	2002	02/03	0004.2	033.0	033.0	0014.9	27	37	04	90,13	29,98	

Figura 64 - Transformando em Imagens de Registro de Carga.

A figura 65 mostra a ferramenta do SQL Server 2000, que foi utilizada para o carregamento dos dados em massa para o DW.

Essa ferramenta permite a importação dos dados que se encontram em vários arquivos com diferentes extensões, por exemplo: TXT, XLS, XML, CSV, entre outras. É capaz também de fazer a leitura de dados que estão em outros bancos de dados, tais como ACCESS e ORACLE, o que a torna uma ferramenta muito útil para este tipo de tarefa.

Figura 65 - Ferramenta utilizada para carga dos dados em massa para o *DW*.

Sexta Etapa: Classificação e Construção de Agregados

O sistema de extração de dados de produção deve incluir a construção ou atualização dos agregados como um processo diário. De modo geral, os registros agregados devem ser construídos fora do DBMS e então carregados com a ferramenta de carga como os outros registros. Um agregado pode ser visto como uma linha de totalização de um relatório simples classificado. Essa é a forma como os agregados devem ser construídos no processo de extração (KIMBALL, 1998).

Sétima Etapa: Generalização de Chaves para Agregados

Quando o registro de fato agregado for criado, será necessário fornecer uma ou mais chaves de dimensão generalizadas. Os agregados são identificados por registros de dimensão especiais, por exemplo, o agregado marca ou o agregado produto em uma tabela de dimensão.

Durante uma determinada carga, é possível que não seja necessário criar novas chaves de dimensão. Se o agregado de categoria já estiver em uso, então essa chave pode ser utilizada (KIMBALL, 1998).

Oitava e Nona Etapa: Carregamento e Processamento de Exceções

O carregamento de dados em uma tabela de fatos deve ser feito na forma de uma operação em massa com o índice mestre desativado, conforme mostra a figura 66. Como foi mencionado anteriormente, o índice mestre deve ser segmentável para que se possa dar um *drop* na parte do período de tempo mais recente e reconstruí-lo após o carregamento dos dados. Será muito mais rápido executar uma carga em massa do que processar os registros um por vez usando as instruções SQL para *INSERT* ou *UPDATE*.

Figura 66 - Índices desativados para carga dos dados da tabela de fatos.

Carregar dados em uma tabela de dimensão é bastante diferente de carregar dados em uma tabela de fatos. A tabela de fatos possui um ou no máximo dois índices grandes construídos por meio de combinação de chaves de dimensão. Uma tabela de dimensão, entretanto, possui várias chaves construídas com base em seus atributos textuais.

Décima Etapa: Garantia de Qualidade

Após o carregamento dos registros básicos e de agregados e da correção de todas as violações de integridade, o gerente de garantia de qualidade deve fazer uma avaliação global do estado de carga. Se o sumário do *DW* coincidir com o sumário de produção, provavelmente todos os registros de nível básico terão sido carregados corretamente. Além disso, cada um dos níveis de agregados em todas as dimensões, exceto tempo, pode ser comparado com o mesmo sumário. Se todos coincidirem, a carga certamente será perfeita.

Caso não haja sumários disponíveis para comparação, será necessário fazer uma verificação de consistência e racionalidade. Em casos extremos, em que o processo de carga esteja incompleto ou haja dados violados, toda carga deve ser removida e executada novamente no dia seguinte. Isso funcionará sem inconvenientes se o sistema for espelhado. Devemos considerar o processo de carga como uma transação gigantesca. Se a carga estiver violada, normalmente fará mais sentido eliminar a carga parcial, esperar pela próxima janela de carregamento e tentar novamente (KIMBALL, 1998).

Décima Primeira e Última Etapa: Publicação

A etapa final do processo de carga de produção tem como função alertar a comunidade de usuários sobre os resultados obtidos. Isso deve tornar-se uma fonte de referência de rotina para o usuário. Todo usuário deve ser informado sobre a avaliação feita pelo gerente de garantia de qualidade da carga. A consistência da carga, a totalização da carga e as áreas a serem usadas ou evitadas devem constar da mensagem.

5.7 - Estrutura de Armazenamento de Dados Proposta

A figura 67 visualiza o Diagrama Entidade Relacionamento (DER) da estrutura de modelagem da ferramenta SQL Server 2000, com a identificação da tabela de fatos com todos os fatos identificados e identificação das tabelas dimensões com todos os atributos identificados.

A estrutura de modelagem proposta deve possibilitar a consulta, extração e mineração das informações, utilizando-se das diversas ferramentas disponíveis hoje no mercado.

Figura 67 - Modelo Proposto.

A figura 68 mostra a modelagem dos dados que foi desenvolvida utilizando a ferramenta de software Logic Works - Erwin, versão 3.5.

Utilizando esta ferramenta, o desenvolvedor de um sistema de informação pode especificar os dados envolvidos, as suas relações e os requisitos de análise. A ferramenta permite criar a base de dados e utilizar mecanismos de sincronização de dados.

Pertence ao conjunto de programas de apoio ao desenvolvimento de software normalmente designados por Ferramentas CASE.

Uma ferramenta CASE (*Computer Aided Software Engineering*) é uma classificação que abrange toda ferramenta baseada em computadores que auxiliam atividades de engenharia de software, desde análise de requisitos e modelagem até programação e testes.

Figura 68 - Modelagem Lógica dos Dados.

5.8 - Discussões entre Modelo Proposto e Revisão Bibliográfica

Conforme a metodologia adotada, foram executados 11 (onze) processos para que fosse possível extrair os dados do banco de dados de produção e enviá-los para o *DW*.

1. Extração primária (ler o formato legado);
2. Identificação dos registros modificados;
3. Generalização de chaves para dimensões em modificação;
4. Transformação em imagens de registro de carga;
5. Migração do sistema legado para o sistema de *DW*;
6. Classificação e construção de agregados;
7. Generalização de chaves para agregados;
8. Carregamento;
9. Processamento de exceções;
10. Garantia de qualidade;
11. Publicação.

Foi necessária a criação de rotinas de carga para atender determinadas situações que ocorreram.

A catalogação dos sistemas de produção que alimentam o *DW* é de suma importância para identificação precisa da extração primária dos dados.

Pode-se dizer que o processo de limpeza e transformação dos dados que foram carregados no *DW* serviu para corrigir algumas imperfeições contidas nas bases de dados transacionais, a fim de fornecer ao usuário do sistema analítico dados concisos e com uma qualidade que permitisse uma tomada de decisão baseada em valores mais próximos dos reais.

Idealmente, pode-se imaginar que o dado deveria apenas ser convertido para padronização de medidas, porém existem valores incorretos nas bases de dados transacionais, os quais não podem ser propagados, principalmente no momento em que serão analisados, muitas vezes comparativamente.

Como exemplo, existem campos de preenchimento obrigatório, com valores que não existem na tabela relacionada, como Safra. Este campo foi adicionado pelo administrador do projeto quando os dados eram levados para o *DW*.

Além disso, a limpeza é necessária porque os dados normalmente advêm de uma fonte muitas vezes desconhecida, concebida há muito tempo, contendo muito “lixo” e inconsistência, porque o interesse de cada usuário ao preencher os dados relativos a qualquer assunto é específico a este.

Por isso, nessa fase do *DW*, fez-se a limpeza dos dados, para haver compatibilidade entre os mesmos. O processo de limpeza não estará completo sem que se possam livrar os dados de problemas que, por algum motivo, passaram despercebidos nos sistemas de origem, tais como: códigos inválidos e preenchimento de vários campos com valores incompatíveis entre si. A própria modelagem do sistema transacional pode conter “pontos fracos” que permitam, por assim dizer, a existência de dados inconsistentes, os quais podem e devem ser filtrados antes da carga no *DW*.

Às vezes encontramos dados com os seguintes problemas:

- **Diferenças de unidades:** campos com relação ao tempo (safra, ano, meses ou dias), sendo necessário converter todas as medidas para qualquer uma delas ou permitir trabalhar com todas as possibilidades.
- **Diferenças de precisão:** alguns valores representados, por exemplo, com duas casas decimais em uma tabela e com quatro casas decimais em outra tabela, cabendo ao administrador do *DW* definir qual a precisão desejada.
- **Diferenças de códigos ou expressões:** em campos que são codificados nos sistemas transacionais a fim de reduzir o espaço de armazenamento, agilizar e padronizar a entrada de dados, para que não existam valores iguais com significados diferentes.

Normalmente as ações de correção das anomalias encontradas não se deram automaticamente com uma rotina específica, até porque isto poderia ter sido feito já na própria base transacional. O que se procurou é solucionar as pendências caso a caso, tentando corrigir inclusive a base original.

O processo de transformação de dados no *DW* ocorreu porque as origens dos dados são de sistemas diferentes, sendo necessário a padronização dos diferentes formatos. Por exemplo: em alguns sistemas a informação sobre o PC da Cana pode estar armazenada no seguinte formato: "PC". Porém, em algum outro sistema pode estar armazenado como "Pur_Can" sendo que os dois significam a mesma coisa e assim sucessivamente. Quando esses dados forem carregados para o *DW*, deve-se ter uma padronização deles, ou seja, quando o usuário for consultar o *DW*, ele não pode ver informações iguais em formatos diferentes.

A carga dos dados foi feita a partir do desenvolvimento de um modelo de armazenamento de dados utilizando um sistema de banco de dados, no qual os dados já passaram pelo processo de limpeza e integração (transformação). As tabelas que são atualizadas no sistema de *DW* foram montadas utilizando-se agregações, sumarizações e ordenações dos dados.

A velocidade de carga para o *DW* pode influenciar na performance do sistema. Neste caso são excluídos os índices de ordenação das tabelas com a finalidade de reduzir a quantidade de controles a serem monitorados pelo Sistema Gerenciador de Banco de Dados, reconstruindo-as posteriormente, após a conclusão da carga.

A todo o momento são realizadas alterações nas bases de dados transacionais. Estas modificações, inclusões de novos registros ou cadastro de novos dados, devem ser atualizadas para o *DW* a fim de que este esteja condizente com a atualidade das fontes de origem. Esta periodicidade deve ser determinada levando-se em conta o volume de dados e a necessidade dos usuários, podendo ser diária, semanal, mensal ou qualquer outro período determinado.

5.9 - Considerações Finais do Capítulo

A partir da realização de uma pesquisa-ação na usina de açúcar e álcool, do setor sucroalcooleiro, procurou-se identificar e compreender os principais benefícios como também as dificuldades encontradas no processo de construção de um SAD.

Sendo que a principal função de um *DW* é fornecer relatórios, alguns benefícios imediatos puderam ser observados após a implementação da nova ferramenta de apoio à decisão.

Foram realizados experimentos utilizando dados reais dos sistemas fontes da usina e os resultados foram novos relatórios com informações gerenciais, táticas e estratégicas, o que antes não era possível de realizar com os atuais sistemas existentes na empresa.

Estes benefícios poderão ser observados nos anexos deste trabalho, onde é feita a aplicação e explicação de várias consultas utilizando uma linguagem de programação padrão para extrair as informações do novo ambiente projetado. Os resultados serão demonstrados por meio de relatórios desenvolvidos utilizando-se do resultado das consultas SQL. Os relatórios se referem às questões que foram abordadas no capítulo 1 deste trabalho.

É importante citar que o modelo proposto foi desenvolvido de forma genérica, ou seja, tomando os devidos cuidados para que a mesmo possa ser utilizado plenamente em outras unidades do setor sucroalcooleiro sem perder suas reais funcionalidades.

6 - Conclusão

A cana-de-açúcar já é matéria-prima de produtos reconhecidamente importantes (açúcar e álcool), mas ainda há um potencial extraordinário a ser extraído dos canaviais. Estima-se que ainda utilizamos muito pouco da energia disponível na cana, um terço apenas. O restante está concentrado nas pontas das folhas e no próprio bagaço, que podem ser transformados em energia elétrica e derivados que estão sendo desenvolvidos graças ao investimento em pesquisa.

O investimento em pesquisa será o diferencial para agregar valor ao agronegócio da cana e também abrir mercados. A discussão sobre o aquecimento global deu a largada para a busca de alternativas energéticas e soluções que minimizem os impactos sobre o meio ambiente.

Uma outra informação que chama muito a atenção é com relação à participação cada vez maior dos carros *flex-fuel* que rodam com álcool, gasolina ou a mistura de ambos na produção brasileira, hoje perto dos 90%, impulsiona o consumo do álcool no país.

A implantação da tecnologia de *DW*, em usinas de açúcar e álcool do setor sucroalcooleiro, de qualquer porte, possibilita a busca e interpretação de informações armazenadas, garantindo maior acuracidade nas tomadas de decisão, permite estabelecer uma aproximação integrada e colaborativa para capturar, criar, organizar e usar todos os ativos de informação de uma empresa, consegue antecipar as mudanças bruscas no mercado, conhecer o negócio, ter uma visão clara sobre novos negócios, auxilia na implementação de novas ferramentas gerenciais, ter conhecimento sobre novas tecnologias, produtos ou serviços que possam afetar direta e indiretamente o seu negócio.

Ganhos com produtividade, rendimento e melhorias nos processos de negócio de uma usina de açúcar e álcool são outros exemplos de como a tecnologia da informação pode representar um enorme diferencial a ser explorado pela alta administração da empresa.

Os investimentos em um SAD podem realmente contribuir para o sucesso das usinas de açúcar e álcool, uma vez que a mesma passa a ser aplicada em favor da competitividade empresarial.

As tecnologias de SAD ajudam a interpretar o que ocorreu e a decidir sobre estratégias futuras para a empresa, ou seja, têm a finalidade de informar a alta direção da empresa e a gerência quando, como, por que e onde ocorrem problemas ou oportunidades de melhoria dentro da organização, projetando certas características e “indicando” ao usuário situações que podem, muitas vezes, passarem despercebidas.

Os *DW* e *Data Mart*, servem como fonte de dados para estas aplicações, assegurando a consistência, integração e precisão dos dados. Os sistemas transacionais não conseguem responder essas questões por isso, é necessária a criação de um ambiente de apoio de decisão robusto, sustentável e confiável.

Pelos resultados expostos, considera-se que o principal objetivo que consistia em desenvolver um SAD para controle e gerenciamento agrícola em usinas de açúcar e álcool foi atingido. O modelo de armazenagem foi desenvolvido com a finalidade de atender as diversas ferramentas de extração de dados existentes no mercado.

A maneira simples de efetuar as consultas e o modo com que as informações foram estruturadas e armazenadas, capacita o modelo de armazenagem proposto adequado para extração das informações gerenciais e estratégicas, o que seria quase que impossível de ser realizado em um modelo de armazenagem de informações tradicional.

Para o desenvolvimento e implantação de um projeto que envolva a tecnologia de SAD se faz necessário uma equipe qualificada na área de conhecimento na tecnologia proposta e por outro lado toda uma equipe envolvida, focada e disposta a colaborar com a mudança cultural que essas novas tecnologias vierem a exigir com sua implantação. Eleger um patrocinador do projeto também é muito importante, pois é ele que irá fornecer subsídios para o andamento do mesmo. A escolha deste patrocinador quase sempre é feita com um executivo de alto nível (Diretor ou Gerente) dentro da organização.

O comprometimento da alta administração, a integração permanente com a equipe de negócios e sistemas, avaliação das ferramentas e tecnologias, possibilidade de consultas personalizadas, inclusão somente dos dados necessários, definição dos papéis e responsabilidades e correta integração dos trabalhos das equipes principalmente com o envolvimento do usuário final, podem ser considerados como fatores críticos de sucesso na implementação de um SAD.

As maiores dificuldades encontradas pelo pesquisador durante o desenvolvimento do sistema foram com relação à mudança cultural existente na organização e resistência por parte da equipe de informática da própria empresa em fornecer as informações.

Outro fato importante é que os Diretores e Gerentes tendem a querer todos os dados existentes nos sistemas legados (fontes), desde a primeira versão.

Isso, geralmente aumenta a complexidade do modelo, gerando dificuldade na utilização pelos usuários, o que leva ao desânimo e abandono, deixando a impressão de que *DW* é muito complicado e difícil de ser usado, ou seja, não se terminava o primeiro projeto porque não se sabia onde este era finalizado.

A ferramenta utilizada para modelagem dos dados foi o *Software Erwin 3.5* e o Banco de Dados utilizado para o armazenamento das informações e desenvolvimento do *DW* foi o *SQL Server 2000*.

6.1 - Proposta para Trabalhos Futuros

Como proposta para trabalhos futuros, poderá ser feita a implantação do modelo proposto em usinas de açúcar e álcool e destilarias do setor sucroalcooleiro para avaliação e validação dos resultados obtidos e como o modelo desenvolvido se comportará em outras unidades do mesmo segmento.

Desenvolver novos modelos para serem utilizados em outras áreas da usina utilizando a mesma tecnologia.

Adicionar ao modelo proposto a área agrícola automotiva (Caminhões, Colhedoras de Cana, Carregadoras, Tratores, Implementos Agrícolas, entre outros).

Uma outra possibilidade seria comparar o modelo proposto na pesquisa com outros modelos de sistemas já existentes no mercado e que atuam em usinas de açúcar e álcool e destilarias do setor sucroalcooleiro.

Referências Bibliográficas

AQUARONI, L.M.; CAZARINI, E.W. (2001) **Fatores Chaves de Sucesso Aplicados ao Processo Decisório Organizacional com o uso do SAD**. In: Encontro Nacional de Engenharia de Produção - XXI ENEGEP, 2001, Salvador, Bahia.

ASSUMPÇÃO, M. R. (2001) **Integração da Cadeia Produtiva do Açúcar à Rede de Suprimento da Indústria Alimentícia**. Tese de Doutorado - Área de Engenharia de Produção - Escola Politécnica/USP - São Paulo, 2001.

BAKER, C. R. (2000) **"Towards the increased use of action research in accounting information systems"**. Accounting Forum. Vol. 24 Nº 4, December.

BARBIER, R. (2002) **A pesquisa-ação**. Trad. Lucie Didio. Brasília: Plano Editora.

BARRET, D.; BARTON, N. (2006). **Best Practices in Building a Data Warehouse Quickly**. *Business Intelligence Journal*; Fourth Quarter, ABI/INFORM Global pg. 37.

BASKERVILLE, R.L.; WOOD-HARPER, T.A. (1996) **"A critical perspective on action research as a method for information systems research"**. Journal of Information Technology (1996) 11, 235-246.

BISPO, C.A.F. (1998) **Uma Análise da Nova Geração de Sistemas de Apoio à Decisão**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 1998.

BISPO, C.A.F.; CAZARINI, E.W. (1998) **A Evolução do Processo Decisório**. In: Encontro Nacional de Engenharia de Produção - XVIII ENEGEP, 1998, Niterói, Rio de Janeiro.

BUOSI, T. (2004) **Sistemas Computacionais de Suporte à Medição de Desempenho: Proposição de Critérios para Análise, Comparação e Aquisição de Sistemas**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2004.

CALDAS, M.P.K.; SCANDELARI, L.; KOVALESKI, J.L. (2006) **Aplicações sobre um *data warehouse* no ambiente das organizações e suas vantagens**. In: Simpósio de Engenharia de Produção, XIII SIMPEP, 2006, Bauru. São Paulo.

CHIZZOTI, A. (1991) **Pesquisa em ciências humanas e sociais**. São Paulo, Cortez.

CLERICUZI, A.Z.; ALMEIDA, A.T.; COSTA, A.P.C.S. (2006) **Aspectos relevantes dos SAD nas organizações: um estudo exploratório**. Produção, v.16, n.1, 2006.

COREY, M.; ABBEY, M.; ABRAMSON, I.; TAUB, B. (2001) **Oracle 8i data warehouse**. Tradução João Tortello. Rio de Janeiro: Campus.

CORREA, L.A.; GUEDES, L.G.R.; LEMOS, R.P. (2007) **Determinação da maturidade organizacional para o sucesso gerencial do *data warehouse***. In: Simpósio de Engenharia de Produção, XIV SIMPEP, 2007, Bauru. São Paulo.

DEMO, P. (2002) **Pesquisa e construção de conhecimento: metodologia científica no caminho de Habermas**. Rio de Janeiro: Tempo Brasileiro.

DUARTE, M.D.O.; GUSMÃO, A.P.H.; ALMEIDA, A.T. (2005) **Sistema de apoio à decisão com modelo aditivo para priorização de sistemas de informação**. In: Encontro Nacional de Engenharia de Produção, XXV ENEGEP, 2005, Porto Alegre, Minas Gerais.

FAVARETTO, F. (2007) **Melhoria da qualidade da informação no controle da produção: estudo exploratório utilizando *Data Warehouse***. Produção, v.17, n.2, 2007.

FREIRE, J.E. (2000) **Uma abordagem sobre os colaboradores na atual sociedade da informação**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2000.

FREIRE, J.E.; CAMPOS, R.R. (2007) **Aspectos da Implantação e Manutenção de Sistemas ERP no Setor Sucroalcooleiro da Região de Ribeirão Preto e São José do Rio Preto**. In: Encontro Nacional de Engenharia de Produção, XXVII ENEGEP, 2007, Foz do Iguaçu. Paraná.

FREIRE, J.E.; SEIXAS, J.A.; CAZARINI, E.W. (2001) **Uma Nova Arquitetura de Sistema de Apoio à Decisão**. In: Encontro Nacional de Engenharia de Produção - XXI ENEGEP, 2001, Salvador, Bahia.

FORTULAN, M.R.; FILHO, E.V.G. (2005) **Uma proposta de aplicação de Business Intelligence no chão-de-fábrica**. Gestão & Produção, v.12, n.1, 2005.

FORTULAN, M.R. (2006) **O uso de Business Intelligence para gerar indicadores de desempenho no Chão-de-Fábrica: uma proposta de aplicação em uma empresa de manufatura**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2006.

GONÇALVES, P.R. (2005) **Sistema de Informação Geográfica para Apoio à Decisão ao Combate a Incêndio**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2005.

GRIGORI, D.; CASATI, F.; CASTELLANOS, M.; DAYAL, U.; SAYAL, M.; SHAN, M.C. (2004) **Business Process Intelligence**. Computers in Industry, V.53, P.324-343.

HOLANDA, V.B; RICCIO, E.L. (2001) **A Utilização da Pesquisa-Ação para Perceber e Implementar Sistemas de Informações Empresariais**. 13th Asian Pacific Conference on International Accounting Issues. Rio de Janeiro, Outubro.

INMON, W. H. (1997) **Como Construir o Data Warehouse**. 2º ed. New York: Editora Campus.

INMON, W. H. (2000) **The Data Warehouse Environment: Quantifying Cost Justification and Return on Investment**. Disponível em: <http://www.billinmon.com/library/whiteprs/???.pdf> Acesso em 13/04/2007.

ITALIANO, I.C.; FERREIRA, J.E. (2006) **Synchronization Options for Data Warehouse Designs**.

JUNIOR, R.F.T.; FERNANDES, F.C.F.; PEREIRA, N.A. (2006) **Sistemas de apoio à decisão para programação da produção em fundições de mercado**. Gestão & Produção, v.13, n.2, 2006.

KIMBALL, R. (1998) **Data Warehouse Toolkit**. 2º ed. São Paulo: Makron Books.

LACERDA, J.M.; AMARO, A.L. (2007) **Seleção e acompanhamento de desempenho de pessoal com auxílio de sistemas de apoio à decisão (SAD)**. In: Encontro Nacional de Engenharia de Produção, XXVII ENEGEP, 2007, Foz do Iguaçu. Paraná.

LAPLANTINE, F. (2004) **Antropologia da doença**. São Paulo. Martins Fontes.

LAUDON, K.C.; LAUDON, J.P. (1996) **Management information System**. 4º ed. Organization and Technology. São Paulo: Prentice-Hall.

MACCARI, E.A.; SAUAIA, A.C.A. (2006) **Aderência de sistemas de informação na tomada de decisão: um estudo multicaso com jogos da empresa**. Gestão da Tecnologia e Sistemas de Informação, São Paulo, v.3, n.3, p.371-388, 2006. ISSN 1807-1775.

MACEDO, I.C. (2005) **A Energia da Cana de Açúcar**. Doze estudos sobre a agroindústria da cana-de-açúcar no Brasil e a sua sustentabilidade. UNICA.

MANNINO, M.V.; WALTER, Z. (2006) **A framework for data warehouse refresh policies**. Decision Support Systems, v. 42 n. 7, p. 121-143.

MARAKAS, G.M. (1999) **Decision Support System in the Twenty-first Century**. DSS and data mining technologies for tomorrow's manager. New Jersey: Prentice Hall.

MARCH, S.T.; REVNER, A.R. (2007) **Integrated decision support systems: A data warehousing perspective**. Decision Support Systems, v. 43 n. 3, p. 1031-1043.

MARION, J. C. (1996) - **Contabilidade e controladoria em agribusiness**. Editora Atlas. São Paulo.

MCKERNAN, J. (1991). **Curriculum action research: a handbook of methods and resources for the reflective practitioner**. Nova York: St. Martin's Press.

MEIRA, K.M.G. (2004) **Utilização da Tecnologia de Data Warehousing e da Ferramenta OLAP para apoiar Captação de Doadores de Sangue: Estudo de Caso em um Hemonúcleo**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2004.

MIGLIOLI, A.M. (2006) **Tomada de Decisão na Pequena Empresa: Estudo Multicaso sobre a Utilização de Ferramentas informatizadas de Apoio à Decisão**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2006.

MORAES, M.A.F.D; SHIKIDA, P.F.A. (2002) **Agroindústria Canavieira no Brasil**. Evolução, Desenvolvimento e Desafios. São Paulo: Atlas.

MORIN, A. (2004) **Pesquisa-ação integral e sistêmica: uma antropedagogia renovada**. Rio de Janeiro. DP&A.

OSTANEL, L.H. (2005) **Uma Proposta de Utilização da Internet como Suporte Informacional à Gestão da Pequena Empresa**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2005.

PURCIDONIO, P.M.; OLIVEIRA, L.S.; HATAKEYAMA, K.; SCANDELARI, L. (2006) **A gestão eficaz do conhecimento através da integração entre o sistema ERP e CRM: um estudo de caso em uma indústria do setor metalúrgico**. In: Encontro Nacional de Engenharia de Produção, XXVI ENEGEP, 2006, Fortaleza. Ceará.

RAHMAN, N. (2007) **Refreshing data warehouses with near real-time updates**. Journal of Computer Information Systems, V.47 n. 3, p. 71-80

RIVAS, I.T.; RIVERA, M.R.; LIZAMA, E.R. (2007) **Una metodología para sectorizar pacientes en el consumo de medicamentos aplicando datamart y datamining en un hospital**. Ind. data v.10 n.1 Lima ene./jun. 2007.

SANTOS, D.E.; MORAES, R.O. (2007) **Condicionantes de desempenho em projetos de data warehouse: estudo de caso em uma empresa brasileira de saneamento**. In: Simpósio de Engenharia de Produção, XIV SIMPEP, 2007, Bauru. São Paulo.

SEIXAS, J.A. (2000) **Um estudo sobre uso do data warehouse para Auxiliar o tratamento da informação no sistema produtivo: um estudo de caso em uma empresa no setor agroindustrial**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos - Universidade de São Paulo, São Carlos, 2000.

SINGH, H.S. (2001) **Data Warehouse**. São Paulo: Makron Books.

TARAPANOFF, K.; MENDES, M.D.; JUNIOR, R.H.A. (1995) **Técnicas para tomada de decisão nos sistemas de informação**. Brasília: Thesaurus, 1995.

TARAPANOFF, K.; JUNIOR, R.H.A.; CORMIER, P.M.J. (2000) **Sociedade da informação e inteligência em unidades de informação**. Ciência da Informação, Dez 2000, vol.29, n.3, p.91-100. ISSN 0100-1965

THIOLLENT, M. (1997) **Pesquisa-Ação nas Organizações**. São Paulo: Editora Atlas, 164p.

USINA (2007) **Caderno de Informações**. 2007.

Bibliografia Complementar:

ANFAVEA (2007). **Associação Nacional dos Fabricantes de Veículos Automotores do Brasil** Disponível em: <http://www.anfavea.com.br/Index.html> Acesso em: 06/06/2007.

ARAÚJO, A. C. (2005). **“A informação Como Fator Diferenciador para o Sucesso Estratégico das Organizações”** Disponível em: http://www.informal.com.br/artigos/a10092002_001.htm Acesso em: 13/08/2005.

CANAOESTE (2007). **Associação dos Plantadores de Cana do Estado de São Paulo.** Disponível em: <http://www.canaoeste.com.br> Acesso em: 06/06/2007.

CANAVIEIROS (2007). **Revista Canavieiros - Edição 11: maio 2007.** Disponível em: http://www.revistacanavieiros.com.br/canavieiro/Rev_Canavieiros_Abril/Default.html Acesso em: 12/06/2007.

COPERSUCAR (2007). Cooperativa de Produtores de Cana-de-Açúcar, Açúcar e Alcool do Estado de São Paulo. **A Moderna Agroindústria Canavieira.** Disponível em: <http://www.copersucar.com.br> Acesso em: 06/06/2007.

DAL'ALBA, A. (2007). **Um Estudo sobre Data Warehouse.** Disponível em: <http://www.geocities.com/siliconvalley/port/5072/> Acesso em: 07/03/2007.

EMBRAPA (2008). **Empresa Brasileira de Pesquisa Agropecuária.** Disponível em: <http://www.embrapa.br/> Acesso em: 14/01/2008.

IDEA (2007). **Instituto de Desenvolvimento Agroindustrial - Revista IDEA NEWS - Edição 78 - Maio 2007.** Disponível em: <http://www.ideaonline.com.br> Acesso em: 05/05/2007.

IEA (2007). **Instituto de Economia Agrícola.** Disponível em: <http://www.iea.sp.gov.br/OUT/verTexto.php?codTexto=4010> Acesso em: 14/12/2007.

REVISTA QUÍMICA (2007). **Tecnologia amplia oferta e reduz custo do álcool.** Disponível em: <http://www.quimica.com.br/revista/qd417/alcool1.htm> Acesso em: 07/12/2007.

UNICA (2007). União da Agroindústria Canavieira de São Paulo: **Produção e uso do Etanol combustível no Brasil.** Disponível em: <http://www.unica.com.br> Acesso em: 16/05/2007.

ANEXOS

O objetivo destes anexos é mostrar alguns exemplos, utilizando-se de uma linguagem de programação padrão, informações extraídas do modelo de armazenagem proposto e responder a diversas questões que foram levantadas no capítulo 1 deste trabalho.

A linguagem de programação utilizada será a linguagem SQL (*Structured Query Language*) ou linguagem de consulta estruturada. É uma linguagem de pesquisa declarativa para banco de dados relacionais (base de dados relacional).

Os relatórios desenvolvidos foram utilizando o Microsoft Excel e o Crystal Reports.

ANEXO - 1 (Questões com Relação à Produção):

SELECT

```
BANCO_GERAL.ANO, BANCO_GERAL.SAFRA, BANCO_GERAL.MES, BANCO_GERAL.COD_FUNDO,
BANCO_GERAL.NOME_FUNDO, BANCO_GERAL.COD_TALHAO, BANCO_GERAL.COD_VAR,
BANCO_GERAL.DESC_VAR, Sum(BANCO_GERAL.PROD_TON) AS 'PROD_TON'
```

FROM

```
DDS_AGRICOLA.dbo.BANCO_GERAL BANCO_GERAL
```

WHERE

```
(BANCO_GERAL.ANO>=? And BANCO_GERAL.ANO<=?)
AND (BANCO_GERAL.MES>=? And BANCO_GERAL.MES<=?)
AND (BANCO_GERAL.COD_FUNDO>=? And BANCO_GERAL.COD_FUNDO<=?)
```

GROUP BY

```
BANCO_GERAL.ANO, BANCO_GERAL.SAFRA, BANCO_GERAL.MES, BANCO_GERAL.COD_FUNDO,
BANCO_GERAL.NOME_FUNDO, BANCO_GERAL.COD_TALHAO, BANCO_GERAL.COD_VAR,
BANCO_GERAL.DESC_VAR
```

ORDER BY

```
BANCO_GERAL.ANO, BANCO_GERAL.SAFRA, BANCO_GERAL.MES, BANCO_GERAL.COD_FUNDO,
BANCO_GERAL.COD_TALHAO, BANCO_GERAL.NOME_FUNDO, BANCO_GERAL.COD_VAR, BANCO_GERAL.DESC_VAR
```

O SQL do Anexo - 1 retorna a informação referente à produção de toneladas de cana por ano, safra, mês, fundo agrícola, talhão e variedade de cana. A consulta permite a realização de diversos filtros utilizando-se desses parâmetros, ou seja, é possível saber a produção de toneladas de cana referente a um fundo agrícola ou a uma variedade de cana em específico e realizar o cruzamento dessas informações em uma única consulta. O retorno das informações pode se referir a vários períodos de tempo, o que antes não era possível, utilizando o ambiente existente na usina.

Relatório de Produção em Toneladas - <u>Por Fundo Agrícola e Variedade de Cana</u>										
Ano Ini:	Mês Ini:	Fundo Ini:								
2006		002.6								
Ano Fim:	Mês Fim:	Fundo Fim:								
2006	ZZ	923.4								
Total de Toneladas										1.079.701,30
ANO	SAFRA	MES	COD_FUNDO	NOME_FUNDO	COD_TALHAO	COD_VAR	DESC_VAR	PROD_TON		
2006	06/07	04	005.5	FAZ. SANTA HELENA - SA	0001.8	59	RB85-5453	481,12		
2006	06/07	04	005.5	FAZ. SANTA HELENA - SA	0008.9	59	RB85-5453	189,02		
2006	06/07	04	005.5	FAZ. SANTA HELENA - SA	0009.2	59	RB85-5453	496,02		
2006	06/07	04	005.5	FAZ. SANTA HELENA - SA	0010.2	59	RB85-5453	1.774,80		
2006	06/07	04	007.1	FAZ. SAO LUIZ - SA	0051.9	84	SP89-1115	1.953,93		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0002.3	50	DIVERSAS	1.884,09		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0003.6	50	DIVERSAS	516,17		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0015.9	50	DIVERSAS	1.750,17		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0016.2	81	SP90-1107	1.774,45		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0019.1	50	DIVERSAS	1.398,44		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0020.1	87	SP91-1285	1.043,73		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0023.0	86	SP91-1049	1.814,88		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0024.3	85	SP90-1638	806,17		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0027.2	84	SP89-1115	767,81		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0028.5	82	SP90-3414	1.546,61		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0034.0	50	DIVERSAS	206,12		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0035.3	50	DIVERSAS	767,97		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0036.6	50	DIVERSAS	750,09		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0077.7	50	DIVERSAS	549,49		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0097.1	83	DUPLICADO	33,94		
2006	06/07	04	009.7	FAZ. SANTO ANT. LAGOA - SA	0116.1	63	SP80-3280	40,43		
2006	06/07	04	010.7	FAZENDA DA SERRA	0001.0	75	SP87-365	165,92		
2006	06/07	04	010.7	FAZENDA DA SERRA	0004.9	75	SP87-365	124,06		
2006	06/07	04	010.7	FAZENDA DA SERRA	0005.2	75	SP87-365	792,79		
2006	06/07	04	010.7	FAZENDA DA SERRA	0006.5	75	SP87-365	185,05		
2006	06/07	04	012.3	FAZ. SANTANA DAS POSSES - SA	0012.6	59	RB85-5453	516,43		
2006	06/07	04	012.3	FAZ. SANTANA DAS POSSES - SA	0022.3	59	RB85-5453	902,82		
2006	06/07	04	012.3	FAZ. SANTANA DAS POSSES - SA	0023.6	59	RB85-5453	564,83		
2006	06/07	04	013.6	SANTO ANTONIO II - SA	0008.0	59	RB85-5453	45,31		

Anexos

ANEXO - 2 (Questões com Relação ao Rendimento):

SELECT

```
FUNDO_SOLO_VARIEDADE.SAFRA, FUNDO_SOLO_VARIEDADE.COD_FUNDO, FUNDO_SOLO_VARIEDADE.NOME_FUNDO,  
FUNDO_SOLO_VARIEDADE.SOLO_TALHAO, FUNDO_SOLO_VARIEDADE.VAR_TALHAO,  
FUNDO_SOLO_VARIEDADE.COD_ESTAG, FUNDO_SOLO_VARIEDADE.COD_TALHAO,  
FUNDO_SOLO_VARIEDADE.TONELADAS, FUNDO_SOLO_VARIEDADE.AREA_CORTE_TALHAO,  
FUNDO_SOLO_VARIEDADE.TON_HA
```

FROM

```
DDS_AGRICOLA.dbo.FUNDO_SOLO_VARIEDADE FUNDO_SOLO_VARIEDADE
```

WHERE

```
(FUNDO_SOLO_VARIEDADE.ANO>=? And FUNDO_SOLO_VARIEDADE.ANO<=?) AND  
(FUNDO_SOLO_VARIEDADE.COD_FUNDO>=? And FUNDO_SOLO_VARIEDADE.COD_FUNDO<=?) AND  
(FUNDO_SOLO_VARIEDADE.SOLO_TALHAO>=? And FUNDO_SOLO_VARIEDADE.SOLO_TALHAO<=?) AND  
(FUNDO_SOLO_VARIEDADE.VAR_TALHAO>=? And FUNDO_SOLO_VARIEDADE.VAR_TALHAO<=?) AND  
(FUNDO_SOLO_VARIEDADE.COD_ESTAG>=? And FUNDO_SOLO_VARIEDADE.COD_ESTAG<=?) AND  
(FUNDO_SOLO_VARIEDADE.STATUS_TALHAO='FECHADO')
```

ORDER BY

```
FUNDO_SOLO_VARIEDADE.ANO, FUNDO_SOLO_VARIEDADE.COD_FUNDO, FUNDO_SOLO_VARIEDADE.SOLO_TALHAO,  
FUNDO_SOLO_VARIEDADE.VAR_TALHAO, FUNDO_SOLO_VARIEDADE.COD_ESTAG
```

O SQL do Anexo - 2 retorna a informação referente a produtividade agrícola por ano/safra, fundo agrícola, tipo de solo, variedade de cana, estágio de corte e talhão. A consulta permite diversos filtros utilizando-se desses parâmetros, fazendo a extração da informação referente a históricos de vários anos de safra e permitindo o cruzamento dessas informações em uma única consulta, o que antes não era permitido pelo ambiente existente na Usina. Nesta consulta é possível identificar o rendimento de uma variedade de cana em determinado fundo agrícola, ou em determinado tipo de solo, ou em determinado tipo de corte, ou seja, é possível fazer o cruzamento da informação de diversas maneiras.

SAFRA	COD_FUNDO	NOME_FUNDO	SOLO	VAR	COD_ESTAG	COD_TALHAO	TONELADAS	AREA_CORTE_TALHAO	TON_HA
06/07	002.6	FAZ FLORESTA - SA	08	44	03	0006.4	706,94	10,19	69,38
06/07	002.6	FAZ FLORESTA - SA	08	44	03	0007.7	1.508,80	15,87	95,07
06/07	002.6	FAZ FLORESTA - SA	08	76	03	0008.0	1.827,76	21,78	83,92
06/07	002.6	FAZ FLORESTA - SA	08	76	03	0009.3	2.297,69	17,16	133,90
06/07	002.6	FAZ FLORESTA - SA	08	76	03	0010.3	1.570,18	17,45	89,98
06/07	002.6	FAZ FLORESTA - SA	09	50	07	0011.6	899,89	5,94	151,50
06/07	002.6	FAZ FLORESTA - SA	09	72	07	0012.9	1.318,80	16,65	79,21
06/07	002.6	FAZ FLORESTA - SA	09	72	07	0013.2	983,98	8,05	122,24
06/07	002.6	FAZ FLORESTA - SA	09	72	07	0014.5	865,92	13,58	49,04
06/07	002.6	FAZ FLORESTA - SA	09	72	07	0015.8	981,98	18,18	54,01
06/07	002.6	FAZ FLORESTA - SA	09	72	07	0016.1	984,97	14,72	66,91
06/07	002.6	FAZ FLORESTA - SA	09	75	03	0001.9	1.700,27	19,49	87,24
06/07	002.6	FAZ FLORESTA - SA	09	75	03	0002.2	1.977,18	19,71	100,31
06/07	002.6	FAZ FLORESTA - SA	09	75	03	0003.5	838,62	17,12	48,98
06/07	002.6	FAZ FLORESTA - SA	09	75	03	0004.8	2.519,53	15,91	158,36
06/07	002.6	FAZ FLORESTA - SA	09	75	03	0005.1	1.393,49	18,63	74,80
06/07	003.9	FAZ FRUTAL - SA	06	37	05	0055.9	90,59	5,95	15,23
06/07	003.9	FAZ FRUTAL - SA	06	37	26	0054.6	654,07	14,83	44,10
06/07	003.9	FAZ FRUTAL - SA	08	43	02	0004.1	941,52	12,00	78,46
06/07	003.9	FAZ FRUTAL - SA	08	43	02	0005.4	899,80	13,62	66,06
06/07	003.9	FAZ FRUTAL - SA	08	43	02	0006.7	871,05	11,45	76,07
06/07	003.9	FAZ FRUTAL - SA	08	43	02	0007.0	433,20	4,97	87,16
06/07	003.9	FAZ FRUTAL - SA	08	43	02	0009.6	937,52	11,79	79,52
06/07	003.9	FAZ FRUTAL - SA	08	43	02	0010.6	155,09	1,84	84,29
06/07	003.9	FAZ FRUTAL - SA	08	76	01	0001.2	1.981,86	15,49	127,94
06/07	003.9	FAZ FRUTAL - SA	08	76	01	0056.2	1.455,33	10,15	143,38
06/07	004.2	FAZ SANTA CANDIDA - SA	08	30	03	0004.4	1.696,00	22,66	74,85
06/07	004.2	FAZ SANTA CANDIDA - SA	08	30	03	0005.7	1.257,26	18,35	68,52
06/07	004.2	FAZ SANTA CANDIDA - SA	08	30	03	0031.6	909,76	9,29	97,93

ANEXO - 3 (Questões com Relação ao Preço e Faturamento):

SELECT

```
PRECO_CANA_MES.ANO, PRECO_CANA_MES.SAFRA, PRECO_CANA_MES.MES, PRECO_CANA_MES.TONELADAS,
PRECO_CANA_MES.FATURAMENTO, PRECO_CANA_MES."R$_POR_TON"
```

FROM

```
DDS_AGRICOLA.dbo.PRECO_CANA_MES PRECO_CANA_MES
```

WHERE

```
(PRECO_CANA_MES.ANO>=? And PRECO_CANA_MES.ANO<=?)
AND (PRECO_CANA_MES.MES>=? And PRECO_CANA_MES.MES<=?)
```

ORDER BY

```
PRECO_CANA_MES.ANO, PRECO_CANA_MES.MES
```

COMPUTE

```
SUM(PRECO_CANA_MES.TONELADAS),
SUM(PRECO_CANA_MES.FATURAMENTO)
```

O SQL do Anexo - 3 retorna a informação referente ao preço médio da tonelada de cana nos anos e meses de safra. O faturamento médio mensal também é uma informação que a consulta disponibiliza ao gerente.

A consulta permite utilizar os parâmetros ano e mês, referente a históricos de vários anos de safra e permitindo o cruzamento dessas informações em uma única consulta (Ex: Comparar o preço médio e faturamento por tonelada de cana referente ao mês de maio de todas as safras).

A produção total de toneladas de cana e o faturamento médio por tonelada de cana do período são mostrados logo abaixo da consulta utilizando o comando "COMPUTE", como um resumo das informações. Utilizando estas duas informações, o gerente consegue chegar ao preço médio da tonelada de cana do período, de acordo com os parâmetros da consulta.

ANEXO - 4 (Questões com Relação aos Dados Tecnológicos: FIBRA, PC, POL e ATR) Exemplo: PC

```
CREATE PROCEDURE MEDIA_PC_MENSAL
```

```
    @SAFRA_INI VARCHAR(5), @SAFRA_FIM VARCHAR(5), @MES_INI VARCHAR(2), @MES_FIM VARCHAR(2),  
    @FUNDO_INI VARCHAR(5), @FUNDO_FIM VARCHAR(5), @VAR_INI VARCHAR(2), @VAR_FIM VARCHAR(2) AS  
SELECT A.MES, SUM(CASE WHEN A.PC_POR_TON = 0.00 THEN A.PROD_TON ELSE 0.00 END) AS TON_SEM_ANALISE, SUM(CASE  
    WHEN A.PC_POR_TON <> 0.00 THEN A.PROD_TON ELSE 0.00 END) AS TON_ANALISADAS, SUM(A.PROD_TON) AS  
    TONELADAS, SUM(A.PC_POR_TON * A.PROD_TON) AS TONELADAS_PC, SUM(A.PC_POR_TON * A.PROD_TON) / SUM(CASE  
    WHEN A.PC_POR_TON <> 0.00 THEN A.PROD_TON END) AS PC_MEDIA  
FROM DIMENSAO_FATO A  
WHERE CONVERT(CHAR(4), YEAR( '01/01/' + SUBSTRING(A.SAFRA,1,2))) BETWEEN  
    YEAR( '01/01/' + SUBSTRING(@SAFRA_INI,1,2) )  
    AND YEAR( '01/01/' + SUBSTRING(@SAFRA_FIM,1,2))  
    AND A.MES BETWEEN @MES_INI AND @MES_FIM  
    AND A.COD_FUNDO BETWEEN @FUNDO_INI AND @FUNDO_FIM  
    AND A.COD_VAR BETWEEN @VAR_INI AND @VAR_FIM  
GROUP BY A.MES  
ORDER BY A.MES
```

Vários métodos analíticos padrão são disponíveis para determinar o pico de maturidade ou qualidade para que a cana seja colhida no tempo certo. Sem tal análise vários agricultores também colhem a cana baseada na idade do cultivo e aparência. Às vezes os agricultores colhem o cultivo mesmo antes de esse estar totalmente maduro por causa da necessidade de suprir cana às moendas o mais cedo possível. Para que possa fornecer matéria-prima para a destilaria durante toda a safra, que dura em torno de seis meses, é necessário que a lavoura de cana-de-açúcar, tenha variedades precoces, médias e tardias, isto quer dizer, variedades em que a maturação ocorra no início, meio e fim da safra. O procedimento em SQL do Anexo - 4 retorna as informações da curva de maturação ou pureza da cana de acordo com os parâmetros informados (Safra, Mês, Fundo Agrícola e Variedade de Cana).

ANEXO - 5 (Questões com Relação aos Custos):

CREATE PROCEDURE GASTOS_PROCESSOS

```

    @ANOMES_INI    VARCHAR(6),    @ANOMES_FIM    VARCHAR(6),
    @FUNDO_INI     VARCHAR(5),    @FUNDO_FIM     VARCHAR(5),
    @ORIGEM_INI    VARCHAR(25),   @ORIGEM_FIM    VARCHAR(25)

```

AS

```

    DECLARE @TOTAL NUMERIC(19,2)    DECLARE @TONELADAS NUMERIC(19,2)

```

SELECT @TOTAL = SUM(VL_TOTAL)

FROM GASTOS_FUNDO_AGRICOLA

WHERE ANO+MES BETWEEN @ANOMES_INI AND @ANOMES_FIM AND COD_FUNDO BETWEEN @FUNDO_INI AND @FUNDO_FIM
AND ORIGEM BETWEEN @ORIGEM_INI AND @ORIGEM_FIM

SELECT @TONELADAS = SUM(PROD_TON)

FROM DIMENSAO_FATO

WHERE ANO+MES BETWEEN @ANOMES_INI AND @ANOMES_FIM AND COD_FUNDO BETWEEN @FUNDO_INI AND @FUNDO_FIM
AND @ORIGEM_INI <> '03_FORNECEDORES' AND @ORIGEM_FIM <> '03_FORNECEDORES'

SELECT GRUPO_PROCESSOS, SUM(VL_TOTAL) AS TOTAL, SUM(VL_TOTAL)/@TOTAL AS PERC, TONELADAS AS
TON, SUM(VL_TOTAL)/@TONELADAS AS 'R\$/TON'

FROM GASTOS_FUNDO_AGRICOLA

WHERE ANO+MES BETWEEN @ANOMES_INI AND @ANOMES_FIM AND COD_FUNDO BETWEEN @FUNDO_INI AND
@FUNDO_FIM AND ORIGEM BETWEEN @ORIGEM_INI AND @ORIGEM_FIM

GROUP BY GRUPO_PROCESSOS

ORDER BY GRUPO_PROCESSOS

O procedimento em SQL do Anexo - 5 retorna as informações com relação aos custos agrícolas por processo (Plantio, Tratos, Colheita, entre outros) de acordo com os parâmetros informados (Ano, Fundo Agrícola e Origem da Cana: Própria ou Fornecedor). A consulta também retorna as toneladas de cana entregues no período, ou seja, é possível encontrar o custo por tonelada de cana de cada processo respectivamente.

Anexos

ANEXO - 6 (Questões com Relação a Consultas que Utilizam Comparativos):

SELECT

```
*FROM COMPARATIVO_TON_HA_REAL_COM_TON_HA_ESTIM
```

WHERE

```
YEAR( '01/01/'+ SUBSTRING(SAFRA,1,2) )  
BETWEEN YEAR( '01/01/'+ SUBSTRING("'" & Parâmetro_1 & "'",1,2) )  
AND YEAR( '01/01/'+ SUBSTRING("'" & Parâmetro_2 & "'",1,2))  
AND COD_FUNDO BETWEEN "'" & Parâmetro_3 & "'" AND "'" & Parâmetro_4 & "'"  
AND COD_TALHAO BETWEEN "'" & Parâmetro_5 & "'" AND "'" & Parâmetro_6 & "'"  
AND STATUS BETWEEN "'" & Parâmetro_7 & "'" AND "'" & Parâmetro_8 & "'
```

ORDER BY

```
YEAR( '01/01/'+ SUBSTRING(SAFRA,1,2) ),2,3"
```

O SQL do Anexo - 6 realiza um comparativo entre o rendimento por tonelada de cana previsto e o realizado por fundo agrícola e seus respectivos talhões de cana durante o período de safra. Conforme as informações vão sendo importadas mensalmente para o *DW*, o gerente agrícola pode por meio desta consulta verificar se o que estava planejado está ocorrendo durante o período de safra.

Esta consulta pode ser realizada com uma visão geral da safra ou pode ser feita uma comparação analítica de cada fundo agrícola, utilizando os parâmetros safra, fundo agrícola, talhão e *status* do talhão (Encerrado ou em Processo).

FUNDO	TALHAO	STATUS	TON_REAL	AREA_CORTE	TON_HA_REAL	TON_ESTIM	AREA_CORTE	TON_HA_ESTIM	DIFERENCA
002.6	SAFRA	06/07							
	0001.9	FECHADO	1.700,27	19,49	87,24	54,86	19,49	2,81	84,42
	0002.2	FECHADO	1.977,18	19,71	100,31	2.162,75	19,71	109,73	(9,42)
	0003.5	FECHADO	838,62	17,12	48,98	52,34	17,12	3,06	45,93
	0004.8	FECHADO	2.519,53	15,91	158,36	235,92	15,91	14,83	143,53
	0005.1	FECHADO	1.393,49	18,63	74,80	2.044,24	18,63	109,73	(34,93)
	0006.4	FECHADO	706,94	10,19	69,38	335,02	10,19	32,88	36,50
	0007.7	FECHADO	1.508,80	15,87	95,07	49,63	15,87	3,13	91,94
	0008.0	FECHADO	1.827,76	21,78	83,92	13,36	21,78	0,61	83,31
	0009.3	FECHADO	2.297,69	17,16	133,90	19,75	17,16	1,15	132,75
	0010.3	FECHADO	1.570,18	17,45	89,98	443,78	17,45	25,43	64,55
	0011.6	FECHADO	899,89	5,94	151,50	517,78	5,94	87,17	64,33
	0012.9	FECHADO	1.318,80	16,65	79,21	217,05	16,65	13,04	66,17
	0013.2	FECHADO	983,98	8,05	122,24	781,54	8,05	97,09	25,15
	0014.5	FECHADO	665,92	13,58	49,04	224,68	13,58	16,54	32,49
	0015.8	FECHADO	981,98	18,18	54,01	48,81	18,18	2,69	51,33
	0016.1	FECHADO	984,97	14,72	66,91	56,66	14,72	3,85	63,06
Total do Fundo			22.176,00	250,43	88,55	7.258,18	250,43	28,98	59,57
Total Geral			22.176,00	250,43	88,55	7.258,18	250,43	28,98	59,57

ANEXO - 7 (Questões com Relação a Estimativas):

CREATE PROCEDURE SIMULA_5_SAFRAS_SINTETICO

```
@SAFRA VARCHAR(5),      @CORTE1 FLOAT,
@CORTE2 FLOAT,          @CORTE3 FLOAT,
@CORTE4 FLOAT,          @CORTE5 FLOAT,
@PRECO_SAFRA FLOAT,     @PRECO_SAFRA1 FLOAT,
@PRECO_SAFRA2 FLOAT,    @PRECO_SAFRA3  FLOAT,
@PRECO_SAFRA4 FLOAT
```

AS

--CRIAR TABELA AUXILIAR

```
SELECT *INTO #DIMENSAO_TESTE
FROM DIMENSAO_TALHAO
WHERE SAFRA_TALHAO = @SAFRA
```

--TRANSFERE A AREA TOTAL DO TALHAO PARA A AREA DE CORTE ONDE O ESTAGIO = '00 - PLANTA'

```
UPDATE #DIMENSAO_TESTE
SET AREA_CORTE_TALHAO = AREA_TOTAL_TALHAO
WHERE ESTAG_TALHAO = '00'
```

--ATUALIZA OS ESTAGIOS PARA A PROXIMA SAFRA

```
UPDATE #DIMENSAO_TESTE
SET ESTAG_TALHAO = CASE WHEN LEN(CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1)) =1 AND ESTAG_TALHAO < '05'
THEN '0' + CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1)
WHEN ESTAG_TALHAO >= '05' THEN '00' ELSE CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1) END
```

--CRIAR OUTRA TABELA AUXILIAR

```
SELECT
CONVERT(CHAR(4),YEAR('01/01/'+ SUBSTRING(SAFRA_TALHAO,4,2))) AS ANO_SAFRA, COD_FUNDO,COD_TALHAO,
ESTAG_TALHAO, AREA_CORTE_TALHAO,
```

CASE

```
WHEN ESTAG_TALHAO = '00' THEN 0.00
WHEN ESTAG_TALHAO = '01' THEN AREA_CORTE_TALHAO * @CORTE1
WHEN ESTAG_TALHAO = '02' THEN AREA_CORTE_TALHAO * @CORTE2
WHEN ESTAG_TALHAO = '03' THEN AREA_CORTE_TALHAO * @CORTE3
WHEN ESTAG_TALHAO = '04' THEN AREA_CORTE_TALHAO * @CORTE4
WHEN ESTAG_TALHAO = '05' THEN AREA_CORTE_TALHAO * @CORTE5
WHEN ESTAG_TALHAO > '05' THEN AREA_CORTE_TALHAO * 0.00
END AS NOVA_PRODUCAO_TONELADAS,
```

CASE

```
WHEN ESTAG_TALHAO = '00' THEN 0.00
WHEN ESTAG_TALHAO = '01' THEN @CORTE1
WHEN ESTAG_TALHAO = '02' THEN @CORTE2
WHEN ESTAG_TALHAO = '03' THEN @CORTE3
WHEN ESTAG_TALHAO = '04' THEN @CORTE4
WHEN ESTAG_TALHAO = '05' THEN @CORTE5
WHEN ESTAG_TALHAO > '05' THEN 0.00
END AS NOVA_ESTIM_TON_HA
```

INTO #ESTIMATIVA_PROXIMA_SAFRA

Anexos

FROM #DIMENSAO_TESTE

GROUP BY

CONVERT(CHAR(4),YEAR('01/01/'+ SUBSTRING(SAFRA_TALHAO,4,2))),
COD_FUNDO,
COD_TALHAO,
ESTAG_TALHAO,
AREA_CORTE_TALHAO

ORDER BY 2,3

--CRIAR TABELA AUXILIAR

SELECT *INTO #DIMENSAO_TESTE_1
FROM #ESTIMATIVA_PROXIMA_SAFRA

--ATUALIZA OS ESTAGIOS PARA A PROXIMA SAFRA

UPDATE #DIMENSAO_TESTE_1

SET ESTAG_TALHAO = CASE WHEN LEN(CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1)) =1 AND ESTAG_TALHAO < '05'
THEN '0' + CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1) WHEN ESTAG_TALHAO >= '05' THEN '00' ELSE
CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1) END

--CRIAR OUTRA TABELA AUXILIAR

SELECT

CONVERT(CHAR(4),YEAR('01/01/'+ SUBSTRING(ANO_SAFRA,3,2))) AS ANO_SAFRA,
COD_FUNDO, COD_TALHAO,
ESTAG_TALHAO,
AREA_CORTE_TALHAO,

CASE

WHEN ESTAG_TALHAO = '00' THEN 0.00
WHEN ESTAG_TALHAO = '01' THEN AREA_CORTE_TALHAO * @CORTE1
WHEN ESTAG_TALHAO = '02' THEN AREA_CORTE_TALHAO * @CORTE2
WHEN ESTAG_TALHAO = '03' THEN AREA_CORTE_TALHAO * @CORTE3
WHEN ESTAG_TALHAO = '04' THEN AREA_CORTE_TALHAO * @CORTE4
WHEN ESTAG_TALHAO = '05' THEN AREA_CORTE_TALHAO * @CORTE5
WHEN ESTAG_TALHAO > '05' THEN AREA_CORTE_TALHAO * 0.00
END AS NOVA_PRODUCAO_TONELADAS,

CASE

WHEN ESTAG_TALHAO = '00' THEN 0.00
WHEN ESTAG_TALHAO = '01' THEN @CORTE1
WHEN ESTAG_TALHAO = '02' THEN @CORTE2
WHEN ESTAG_TALHAO = '03' THEN @CORTE3
WHEN ESTAG_TALHAO = '04' THEN @CORTE4
WHEN ESTAG_TALHAO = '05' THEN @CORTE5
WHEN ESTAG_TALHAO > '05' THEN 0.00
END AS NOVA_ESTIM_TON_HA

INTO #ESTIMATIVA_PROXIMA_SAFRA_1

FROM #DIMENSAO_TESTE_1

Anexos

GROUP BY

```
    CONVERT(CHAR(4),YEAR( '01/01'+ SUBSTRING(ANO_SAFRA,3,2))),
    COD_FUNDO,
    COD_TALHAO,
    ESTAG_TALHAO,
    AREA_CORTE_TALHAO
```

ORDER BY 2,3

--CRIAR TABELA AUXILIAR

```
SELECT *INTO #DIMENSAO_TESTE_2
FROM #ESTIMATIVA_PROXIMA_SAFRA_1
```

--ATUALIZA OS ESTAGIOS PARA A PROXIMA SAFRA

```
UPDATE #DIMENSAO_TESTE_2
SET ESTAG_TALHAO = CASE WHEN LEN(CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1)) =1 AND ESTAG_TALHAO < '05'
THEN '0' + CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1)
WHEN ESTAG_TALHAO >= '05' THEN '00' ELSE CONVERT(VARCHAR(02),CONVERT(INT,ESTAG_TALHAO) +1) END
```

--CRIAR OUTRA TABELA AUXILIAR

```
SELECT
    CONVERT(CHAR(4),YEAR( '01/01'+ SUBSTRING(ANO_SAFRA,3,2))) AS ANO_SAFRA,
    COD_FUNDO,
    COD_TALHAO,
    ESTAG_TALHAO,
    AREA_CORTE_TALHAO,
```

CASE

```
    WHEN ESTAG_TALHAO = '00' THEN 0.00
    WHEN ESTAG_TALHAO = '01' THEN AREA_CORTE_TALHAO * @CORTE1
    WHEN ESTAG_TALHAO = '02' THEN AREA_CORTE_TALHAO * @CORTE2
    WHEN ESTAG_TALHAO = '03' THEN AREA_CORTE_TALHAO * @CORTE3
    WHEN ESTAG_TALHAO = '04' THEN AREA_CORTE_TALHAO * @CORTE4
    WHEN ESTAG_TALHAO = '05' THEN AREA_CORTE_TALHAO * @CORTE5
    WHEN ESTAG_TALHAO > '05' THEN AREA_CORTE_TALHAO * 0.00
    END AS NOVA_PRODUCAO_TONELADAS,
```

CASE

```
    WHEN ESTAG_TALHAO = '00' THEN 0.00
    WHEN ESTAG_TALHAO = '01' THEN @CORTE1
    WHEN ESTAG_TALHAO = '02' THEN @CORTE2
    WHEN ESTAG_TALHAO = '03' THEN @CORTE3
    WHEN ESTAG_TALHAO = '04' THEN @CORTE4
    WHEN ESTAG_TALHAO = '05' THEN @CORTE5
    WHEN ESTAG_TALHAO > '05' THEN 0.00
    END AS NOVA_ESTIM_TON_HA
```

```
INTO #ESTIMATIVA_PROXIMA_SAFRA_2
```

```
FROM #DIMENSAO_TESTE_2
```

GROUP BY

```
    CONVERT(CHAR(4),YEAR( '01/01'+ SUBSTRING(ANO_SAFRA,3,2))),
    COD_FUNDO,
    COD_TALHAO,
    ESTAG_TALHAO, AREA_CORTE_TALHAO
```

ORDER BY 2,3

Anexos

--CRIAR TABELA AUXILIAR

```
SELECT * INTO #DIMENSAO_TESTE_3
FROM #ESTIMATIVA_PROXIMA_SAFRA_2
```

--ATUALIZA OS ESTAGIOS PARA A PROXIMA SAFRA

```
UPDATE #DIMENSAO_TESTE_3
SET ESTAG_TALHAO = CASE WHEN LEN(CONVERT(VARCHAR(02), CONVERT(INT, ESTAG_TALHAO) + 1)) = 1 AND ESTAG_TALHAO < '05'
THEN '0' + CONVERT(VARCHAR(02), CONVERT(INT, ESTAG_TALHAO) + 1)
WHEN ESTAG_TALHAO >= '05' THEN '00' ELSE CONVERT(VARCHAR(02), CONVERT(INT, ESTAG_TALHAO) + 1) END
```

--CRIAR OUTRA TABELA AUXILIAR

```
SELECT
    CONVERT(CHAR(4), YEAR('01/01/' + SUBSTRING(ANO_SAFRA, 3, 2))) AS ANO_SAFRA,
    COD_FUNDO,
    COD_TALHAO,
    ESTAG_TALHAO,
    AREA_CORTE_TALHAO,
```

CASE

```
    WHEN ESTAG_TALHAO = '00' THEN 0.00
    WHEN ESTAG_TALHAO = '01' THEN AREA_CORTE_TALHAO * @CORTE1
    WHEN ESTAG_TALHAO = '02' THEN AREA_CORTE_TALHAO * @CORTE2
    WHEN ESTAG_TALHAO = '03' THEN AREA_CORTE_TALHAO * @CORTE3
    WHEN ESTAG_TALHAO = '04' THEN AREA_CORTE_TALHAO * @CORTE4
    WHEN ESTAG_TALHAO = '05' THEN AREA_CORTE_TALHAO * @CORTE5
    WHEN ESTAG_TALHAO > '05' THEN AREA_CORTE_TALHAO * 0.00
END AS NOVA_PRODUCAO_TONELADAS,
```

CASE

```
    WHEN ESTAG_TALHAO = '00' THEN 0.00
    WHEN ESTAG_TALHAO = '01' THEN @CORTE1
    WHEN ESTAG_TALHAO = '02' THEN @CORTE2
    WHEN ESTAG_TALHAO = '03' THEN @CORTE3
    WHEN ESTAG_TALHAO = '04' THEN @CORTE4
    WHEN ESTAG_TALHAO = '05' THEN @CORTE5
    WHEN ESTAG_TALHAO > '05' THEN 0.00
END AS NOVA_ESTIM_TON_HA
```

INTO #ESTIMATIVA_PROXIMA_SAFRA_3

FROM #DIMENSAO_TESTE_3

GROUP BY

```
    CONVERT(CHAR(4), YEAR('01/01/' + SUBSTRING(ANO_SAFRA, 3, 2))), COD_FUNDO,
    COD_TALHAO, ESTAG_TALHAO,
    AREA_CORTE_TALHAO
```

ORDER BY 2,3

--CRIAR TABELA AUXILIAR

```
SELECT * INTO #DIMENSAO_TESTE_4
FROM #ESTIMATIVA_PROXIMA_SAFRA_3
```

--ATUALIZA OS ESTAGIOS PARA A PROXIMA SAFRA

```
UPDATE #DIMENSAO_TESTE_4
SET ESTAG_TALHAO = CASE WHEN LEN(CONVERT(VARCHAR(02), CONVERT(INT, ESTAG_TALHAO) + 1)) = 1 AND ESTAG_TALHAO < '05'
THEN '0' + CONVERT(VARCHAR(02), CONVERT(INT, ESTAG_TALHAO) + 1)
WHEN ESTAG_TALHAO >= '05' THEN '00' ELSE CONVERT(VARCHAR(02), CONVERT(INT, ESTAG_TALHAO) + 1) END
```

Anexos

--CRIA OUTRA TABELA AUXILIAR

SELECT

```
    CONVERT(CHAR(4),YEAR( '01/01'+ SUBSTRING(ANO_SAFRA,3,2))) AS ANO_SAFRA,  
    COD_FUNDO,  
    COD_TALHAO,  
    ESTAG_TALHAO,  
    AREA_CORTE_TALHAO,
```

CASE

```
    WHEN ESTAG_TALHAO = '00' THEN 0.00  
    WHEN ESTAG_TALHAO = '01' THEN AREA_CORTE_TALHAO * @CORTE1  
    WHEN ESTAG_TALHAO = '02' THEN AREA_CORTE_TALHAO * @CORTE2  
    WHEN ESTAG_TALHAO = '03' THEN AREA_CORTE_TALHAO * @CORTE3  
    WHEN ESTAG_TALHAO = '04' THEN AREA_CORTE_TALHAO * @CORTE4  
    WHEN ESTAG_TALHAO = '05' THEN AREA_CORTE_TALHAO * @CORTE5  
    WHEN ESTAG_TALHAO > '05' THEN AREA_CORTE_TALHAO * 0.00  
    END AS NOVA_PRODUCAO_TONELADAS,
```

CASE

```
    WHEN ESTAG_TALHAO = '00' THEN 0.00  
    WHEN ESTAG_TALHAO = '01' THEN @CORTE1  
    WHEN ESTAG_TALHAO = '02' THEN @CORTE2  
    WHEN ESTAG_TALHAO = '03' THEN @CORTE3  
    WHEN ESTAG_TALHAO = '04' THEN @CORTE4  
    WHEN ESTAG_TALHAO = '05' THEN @CORTE5  
    WHEN ESTAG_TALHAO > '05' THEN 0.00  
    END AS NOVA_ESTIM_TON_HA
```

INTO #ESTIMATIVA_PROXIMA_SAFRA_4

FROM #DIMENSAO_TESTE_4

GROUP BY

```
    CONVERT(CHAR(4),YEAR( '01/01'+ SUBSTRING(ANO_SAFRA,3,2))),  
    COD_FUNDO,  
    COD_TALHAO,  
    ESTAG_TALHAO,  
    AREA_CORTE_TALHAO
```

ORDER BY 2,3

SELECT

```
    ANO_SAFRA,SUM(AREA_CORTE_TALHAO) AS AREA_CORTE, SUM(NOVA_PRODUCAO_TONELADAS) AS  
    NOVA_PRODUCAO_TONELADAS, SUM(NOVA_PRODUCAO_TONELADAS) / SUM(AREA_CORTE_TALHAO) AS  
    NOVA_ESTIM_TON_HA, @PRECO_SAFRA AS PRECO, SUM(NOVA_PRODUCAO_TONELADAS) * @PRECO_SAFRA AS  
    FATURAMENTO
```

FROM

```
    #ESTIMATIVA_PROXIMA_SAFRA
```

WHERE

```
    ESTAG_TALHAO <> '00'
```

GROUP BY

```
    ANO_SAFRA
```

UNION

Anexos

SELECT

```
CONVERT(CHAR(4),ANO_SAFRA + 1) AS ANO_SAFRA, SUM(AREA_CORTE_TALHAO) AS AREA_CORTE,  
SUM(NOVA_PRODUCAO_TONELADAS) AS NOVA_PRODUCAO_TONELADAS,  
SUM(NOVA_PRODUCAO_TONELADAS) / SUM(AREA_CORTE_TALHAO) AS NOVA_ESTIM_TON_HA,@PRECO_SAFRA1 AS  
PRECO, SUM(NOVA_PRODUCAO_TONELADAS) * @PRECO_SAFRA1 AS FATURAMENTO
```

FROM

```
#ESTIMATIVA_PROXIMA_SAFRA_1
```

WHERE

```
ESTAG_TALHAO <> '00'
```

GROUP BY

```
ANO_SAFRA
```

UNION

SELECT

```
CONVERT(CHAR(4),ANO_SAFRA + 2) AS ANO_SAFRA, SUM(AREA_CORTE_TALHAO) AS AREA_CORTE,  
SUM(NOVA_PRODUCAO_TONELADAS) AS NOVA_PRODUCAO_TONELADAS, SUM(NOVA_PRODUCAO_TONELADAS) /  
SUM(AREA_CORTE_TALHAO) AS NOVA_ESTIM_TON_HA, @PRECO_SAFRA2 AS  
PRECO,SUM(NOVA_PRODUCAO_TONELADAS) * @PRECO_SAFRA2 AS FATURAMENTO
```

FROM

```
#ESTIMATIVA_PROXIMA_SAFRA_2
```

WHERE

```
ESTAG_TALHAO <> '00'
```

GROUP BY

```
ANO_SAFRA
```

UNION

SELECT

```
CONVERT(CHAR(4),ANO_SAFRA + 3) AS ANO_SAFRA, SUM(AREA_CORTE_TALHAO) AS AREA_CORTE,  
SUM(NOVA_PRODUCAO_TONELADAS) AS NOVA_PRODUCAO_TONELADAS, SUM(NOVA_PRODUCAO_TONELADAS) /  
SUM(AREA_CORTE_TALHAO) AS NOVA_ESTIM_TON_HA, @PRECO_SAFRA3 AS PRECO,  
SUM(NOVA_PRODUCAO_TONELADAS) * @PRECO_SAFRA3 AS FATURAMENTO
```

FROM

```
#ESTIMATIVA_PROXIMA_SAFRA_3
```

WHERE

```
ESTAG_TALHAO <> '00'
```

GROUP BY

```
ANO_SAFRA
```

UNION

SELECT

```
CONVERT(CHAR(4),ANO_SAFRA + 4) AS ANO_SAFRA, SUM(AREA_CORTE_TALHAO) AS AREA_CORTE,  
SUM(NOVA_PRODUCAO_TONELADAS) AS NOVA_PRODUCAO_TONELADAS,  
SUM(NOVA_PRODUCAO_TONELADAS) / SUM(AREA_CORTE_TALHAO) AS NOVA_ESTIM_TON_HA, @PRECO_SAFRA4 AS  
PRECO, SUM(NOVA_PRODUCAO_TONELADAS) * @PRECO_SAFRA4 AS FATURAMENTO
```

FROM

```
#ESTIMATIVA_PROXIMA_SAFRA_4
```

WHERE

```
ESTAG_TALHAO <> '00'
```

GROUP BY

```
ANO_SAFRA
```

ORDER BY

```
ANO_SAFRA
```

Anexos

O procedimento SQL do Anexo - 7 retorna a informação sobre a estimativa de produção de toneladas de cana, rendimento ton/ha e faturamento médio para as próximas cinco safras. O gerente entra com as informações de rendimento esperadas para cada estágio de corte e preço médio esperado para as próximas safras. Estas informações são processadas com base na safra base desejada.

O normal entre as unidades produtoras é a realização da reforma do talhão de cana após o seu quinto corte, porém existem algumas exceções. Considerando estas exceções, esta consulta permite ao gerente agrícola realizar a simulação com a flexibilidade de até dez cortes da cana antes da realização da reforma de seus talhões. O resultado disponibilizado na consulta mostra também o faturamento médio e as toneladas de cana estimadas para cada safra em específico e o novo rendimento médio ton/ha para o período.

ANEXO - 8 (Questões com Relação a Simulações):

```
CREATE PROCEDURE CALC_VAR_PERCENT
```

```
    @SAFRA_INI    VARCHAR(5),
```

```
    @SAFRA_FIM    VARCHAR(5)
```

```
AS    DECLARE @VAR    NUMERIC(19,2)
```

```
SELECT    @VAR = SUM(TONELADAS) FROM PROD_ESTIM_FUNDO
```

```
    WHERE CONVERT(CHAR(4),YEAR( '01/01/'+ SUBSTRING(SAFRA_TALHAO,1,2)))
```

```
    BETWEEN YEAR( '01/01/'+ SUBSTRING(@SAFRA_INI,1,2) ) AND YEAR( '01/01/'+ SUBSTRING(@SAFRA_FIM,1,2))
```

```
SELECT    VAR_TALHAO, IDENTITY(INT,1,1) AS CLASSIF,DESC_VAR,TONELADAS,PERC
```

```
INTO    #CLASSIF_PERCENT
```

```
FROM    ( SELECT VAR_TALHAO, B.DESCR_VAR,SUM(TONELADAS) AS TONELADAS, CAST(SUM(TONELADAS)/@VAR * 100 AS
    DECIMAL(10,2)) AS PERC FROM PROD_ESTIM_FUNDO INNER JOIN DIMENSAO_VARIEDADE B ON VAR_TALHAO =
    B.COD_VAR AND SAFRA_TALHAO = B.SAFRA_VAR
```

```
WHERE    CONVERT(CHAR(4),YEAR( '01/01/'+ SUBSTRING(SAFRA_TALHAO,1,2)))
```

```
    BETWEEN YEAR( '01/01/'+ SUBSTRING(@SAFRA_INI,1,2) ) AND YEAR( '01/01/'+ SUBSTRING(@SAFRA_FIM,1,2))
```

```
GROUP BY VAR_TALHAO,B.DESCR_VAR) TBL
```

```
ORDER BY 5 DESC
```

```
SELECT    *FROM #CLASSIF_PERCENT
```

O procedimento SQL do Anexo - 8 retorna a informação sobre a produção de toneladas de cana por variedade de cana, em ordem decrescente de toneladas e o seu percentual em relação ao total de toneladas previstas para a safra. Por intermédio desta consulta, o gerente consegue identificar quais serão as variedades de cana que vão entregar a maior quantidade de matéria prima durante a safra. Sabendo quais são estas variedades, o diretor pode utilizá-las na consulta SQL abaixo e realizar várias simulações, como por exemplo, se concentrar nas maiores variedades para uma aplicação de fertilizantes e simular uma nova produção de toneladas após a aplicação destes fertilizantes.

Consultas_TCC_Mestrado.xls						
	A	B	C	D	E	F
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						

Estimativa de Rendimento das Variedades						
Safra ini:	Safra Fim:	Total de Variedades				
06/07	06/07	46				
CLASSIF	VAR_TALHAO	DESC_VAR	TONELADAS	PERC %	PERC % ACUM	
1	75	SP87-365	125.842,33	12,12%	12,12%	
2	44	SP81-3250	118.711,88	11,43%	23,56%	
3	49	SP80-1816	98.230,96	9,46%	33,02%	
4	63	SP80-3280	83.526,25	8,05%	41,06%	
5	74	SP86-155	80.955,58	7,80%	48,86%	
6	76	SP86-42	64.976,37	6,26%	55,12%	
7	37	RB85-5536	57.145,76	5,50%	60,62%	
8	59	RB85-5453	56.881,17	5,48%	66,10%	
9	50	DIVERSAS	52.967,15	5,10%	71,20%	
10	30	RB72-454	37.944,68	3,65%	74,86%	
11	42	SP80-1842	31.690,41	3,05%	77,91%	
12	86	SP91-1049	26.468,33	2,55%	80,46%	
13	53	SP83-5073	26.078,17	2,51%	82,97%	
14	82	SP90-3414	20.518,23	1,98%	84,95%	
15	43	IAC87-3396	19.739,44	1,90%	86,85%	
16	73	SP84-1431	14.889,52	1,43%	88,29%	
17	85	SP90-1638	12.905,98	1,24%	89,53%	
18	84	SP89-1115	10.021,24	0,97%	90,49%	
19	93	SP83-2847	9.491,47	0,91%	91,41%	

Anexos

SELECT

```
TONELADAS = SUM(CASE WHEN VAR_TALHAO IN ('75','44','49','63','74') THEN TONELADAS * (1 + CAST(REPLACE('5.00',',','.') AS DECIMAL(10,5)) / 100) ELSE TONELADAS END)
```

FROM

```
PROD_ESTIM_FUNDO
```

WHERE

```
COD_FUNDO >= '002.6' AND COD_FUNDO <= '923.4'  
AND SAFRA_TALHAO = '07/08'
```

O SQL acima retorna a informação de toneladas de cana referente a uma simulação que é efetuada alterando o rendimento de até cinco variedades de cana simultaneamente. Pode ser efetuada uma consulta global ou em apenas um fundo agrícola em específico. Não que seja uma regra, porém, para uma melhor aplicação e utilização desta consulta, ela deve ser efetuada com os resultados obtidos da consulta do Anexo - 8.

The screenshot shows a software window titled "Simulação" with a subtitle "Estimativa de Produção por Fundo Agrícola". The date "16/07/2008" is displayed in the top right corner. The interface is divided into two main sections. On the left, there are input fields for simulation parameters: "Safrá Previsão: 06/07", "Cód. Fundo Inicia 002.6", "Fundo Inicial: FAZ. FLORESTA - SA", "Cód. Fundo Final: 923.4", "Fundo Final: FAZ. SANTA TEREZINHA - IR", "Produção Estimada: 1.038.187,31", "Média Ton / Ha: 88,49", "Preço Cana (Estimado): 37,17", and "Fatur R\$ (Estimado): R\$ 38.589.422,31". On the right, there is a section titled "Simulação" with a sub-header "% Aplicado nas Variedades" and a table with five columns labeled "75", "44", "49", "63", and "74". Below this table, a grey box contains the simulation results: "Percentual: 5,00", "Nova Produção: 1.063.550,66", "Novo Faturamento: R\$ 39.532.178,03", and "Variação R\$ 942.755,72". The window has a standard Windows XP-style title bar and a status bar at the bottom.

Safrá Previsão:		06/07	
Cód. Fundo Inicia		002.6	
Fundo Inicial:		FAZ. FLORESTA - SA	
Cód. Fundo Final:		923.4	
Fundo Final:		FAZ. SANTA TEREZINHA - IR	
Produção Estimada:		1.038.187,31	
Média Ton / Ha:		88,49	
Preço Cana (Estimado):		37,17	
Fatur R\$ (Estimado):		R\$ 38.589.422,31	

Simulação		% Aplicado nas Variedades				
		75	44	49	63	74
Percentual:		5,00				
Nova Produção:		1.063.550,66				
Novo Faturamento:		R\$ 39.532.178,03				
Variação		R\$ 942.755,72				