DEFINIÇÃO DE DATA MART PARA ANÁLISE DE PRODUTIVIDADE NA AGRICULTURA DE PRECISÃO

Gabriel Augusto de Oliveira (PROVIC/Fundação Araucária/UEPG), Maria Salete Marcon Gomes Vaz (Orientadora), e-mail: salete@uepg.br.

Universidade Estadual de Ponta Grossa/Departamento de Informática.

Ciência da Computação – Metodologia e Técnicas da Computação – Banco de Dados.

Palavras-chave: Evolução, Agricultura de Precisão, Banco de Dados.

Resumo

A Agricultura de Precisão tornou-se um dos maiores meios para a evolução da lavoura. Gerando dados que podem ser transformados em informações e auxiliam o agricultor, através de gráficos e mapas por exemplo. A premissa do Data Mart é armazenar dados de um assunto especifico. Estes, serão analisados e mantidos ao longo de vários anos, auxiliando na tomada de decisão, através de performance e flexibilidade. Desta forma, o presente projeto, possui como principal objetivo o desenvolvimento de um Data Mart que armazene grande quantidade de dados, para que posteriormente, junte-se a mecanismos OLAP (Processamento Analítico Online) para acesso, visualização e analise de informações.

Introdução

O Data Mart (DM) é conhecido por ser uma versão menor de um Data Warehouse (DW), que de modo simplificado, se trata de um conjunto de dados que são armazenados, para análise e consulta, com o objetivo de oferecer auxílio durante a tomada de decisão.

Continuamente, é necessário o entendimento da definição de Agricultura de Precisão (AP), que se trata do planejamento e gerenciamento de todos os processos da produção agrícola, utilizando equipamentos tecnológicos para avaliar e monitorar o terreno.

Um exemplo de máquina da Agricultura de Precisão são as colheitadeiras com sensores. As mesmas geram dados por ponto, como por exemplo, quantidade de grãos colhidos, a umidade do grão, fluxo de grãos, largura da plataforma de colheita, latitude e longitude. A partir do Data Mart desenvolvido e o com os dados, é possível gerar informações que afetam a tomada de decisão do agricultor, de forma a obter mais resultados no futuro.

Materiais e Métodos

Analisando o escopo do projeto e tendo em vista a necessidade de definição de um Data Mart, viu-se necessário o levantamento biblográfico de técnicas e metódos para implementação de um DM. Com objetivo de resposta a proposto inicial, estas

técnicas e métodos foram aplicadas em volta de um ambiente de desenvolvimento de banco de dados. Para que posteriormente um estudo de caso no ambiente de produtividade de uma colheita fictícia pudesse ser visualizado.

Durante a implementação foi utilizado o SQL Server 2016 da Microsoft, o qual possui versão gratuita e que em junção com o Microsoft SQL Server Management Studio proporciona suporte ao desenvolvimento inteligente do banco de dados com a linguagem SQL. Além de conter mecanismos avançados de Business Intelligence e Análise de dados que torna as informações mais visíveis e de fácil entendimento.

Com a escolha do software de auxilio, o próximo passo é a modelagem do DM. Nessa fase é necessário o entendimento das variáveis que influenciam na colheita e em sua produtividade. O cálculo da produtividade pode variar dependendo do grão que está sendo colhido. E em algumas culturas é necessário para o cálculo da produtividade saber o teor de umidade com base no grão seco

Além dos cálculos, é necessário conhecimento de como as colheitadeiras exportam os números adquiridos durante a colheita. Hoje em dia, maquinas mais modernas como da empresa John Deere, já oferecem apoio ao gerenciamento de informações, com softwares especializados para determinadas culturas, que ajudam na tomada de decisão. Muitos destes dados são adquiridos dos portais da própria empresa. Segundo Molin e Borém (2000), os dados também podem ser armazenados instantaneamente em algum dispositivo de memória.

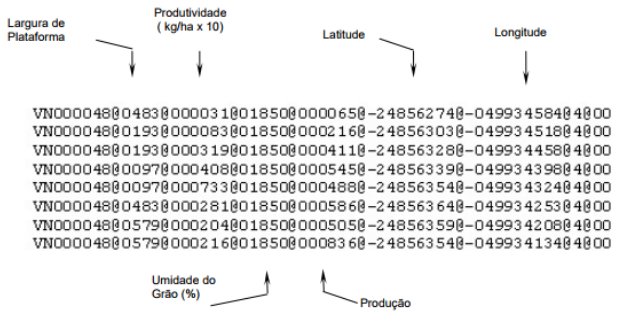


FIGURA 3 – Um arquivo de dados exportado de uma colheitadeira. Fonte: Molin e Borém (2000).

Para o teste de implementação do Data Mart, foi criado um arquivo do tipo .txt de 1000 linhas, baseado no modelo acima, com o objetivo de simular dados fictícios que seriam extraídos de uma colheita e testar o funcionamento do DM. Neste caso a produtividade já é calculada com base no ponto e a umidade do grão mostrada em porcentagem. A partir destas duas variáveis, juntamente com a latitude e longitude pode-se gerar Mapas de Produtividade e Mapas de Umidade da lavoura, que auxiliam na tomada de decisão, possibilitando melhor visão da produção.

A quantidade de linhas no arquivo comparado a um caso real é bem pequena, essas leituras geralmente tendem a gerar quantidades imensas de números e é por esse fato que o Data Mart deve ser desenvolvido com variáveis que não gerem desperdício de memória, já que a quantidade de elementos armazenado será alta.

Resultados e Discussão

A partir das pesquisas, verificou-se que as tabelas de dimensões essenciais para o DM seriam: data, grão e ponto. O modelo dimensional adotado, foi o esquema em estrela, que é o mais adotado na maioria dos casos.

A tabela data tem relação com a data que a colheita está sendo efetuada, ela possui as variáveis: DataKey que é chave primária da tabela, CalendarioAno, CalendarioTrimestre, MesDoAno, NomeDoMes, DiaDoMes, DiaDaSemana, DiaNome. É necessário guardar exatamente quando a colheita ocorreu, para manter um histórico e para que analises futuras possam ocorrer.

As tabelas de dimensões grão e ponto, são as mais importantes para o cálculo de produtividade e de umidade. Algumas colheitadeiras possuem sensores que calculam o fluxo de grãos, a massa e a umidade do grão, essas variáveis estão presentes na tabela de dimensões grão.

As coordenadas de latitude e longitude que resultaram na elaboração de mapas estão presentes na tabela ponto. A mesma possui o tamanho da plataforma de colheita e o TempoPorTalhao, que se trata do tempo pré-determinado pelo agricultor, que está relacionado com a distância percorrida pela máquina, gerando uma largura do talhão. Além da chave primária PontoKey.

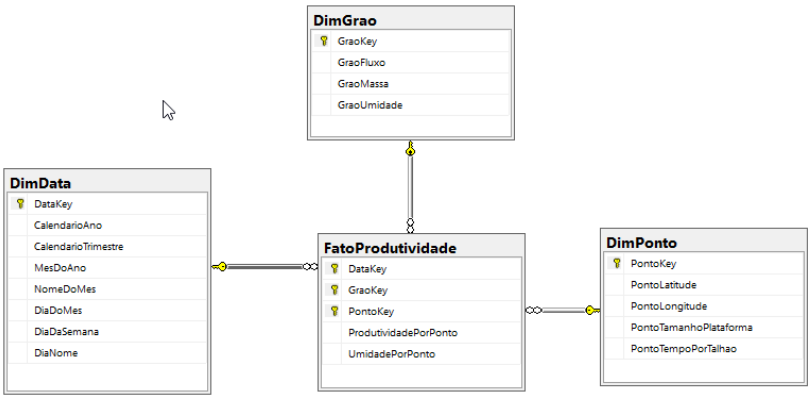


FIGURA 4 – Modelagem do Data Mart. Fonte: Autor

Os maiores resultados em um Data Mart estão presentes na tabela de fatos. Para a colheita, as informações pertinentes são a Produtividade e a Umidade, os quais são obtidos do arquivo. Caso haja a necessidade do cálculo da produtividade, geralmente, o mesmo é feito com base no grão seco.

Com a modelagem do banco de dados finalizada, é necessário transformar os dados em informações, a melhor maneira é a apresentação de mapas. Desta forma, as áreas de maior produtividade seriam salientadas, e as regiões de colheita menos produtivas poderiam ser analisadas para verificar qual problema está afetando o baixo desenvolvimento.

Resultados | Mensagens

	DataKey	GraoKey	PontoKey	ProdutividadePorPonto	UmidadePorPonto
495	495	495	495	499	1850
496	496	496	496	719	1850
497	497	497	497	676	1850
498	498	498	498	622	1850
499	499	499	499	628	1850
500	500	500	500	730	1850
501	501	501	501	458	1850
502	502	502	502	244	2100
503	503	503	503	339	2100
504	504	504	504	321	2100
505	505	505	505	239	2100

FIGURA 5 – Manipulação dos dados retirados do banco, o campo produtividade com base no grão seco não foi utilizado. Fonte: Autor.

A partir da modelagem e da figura 5, pode-se perceber a diferença em comparação a um banco de dados transacional. O mesmo possuiria mais tabelas e o tempo das consultas seria maior. Além disso, o data mart é empregado com ferramentas OLAP, para realização de consultas, criação de gráficos e mapas, tornando o mecanismo mais eficiente e rápido.

Conclusões

Um banco de dados como o Data Mart, a primeira vista pode parecer algo pequeno, contudo é apenas uma parte de algo muito maior, o Data Warehouse. Ambos possuem o objetivo de controlar a gama gigantesca de dados e devem estar sempre disponíveis para gerar novas informações.

O DM foi projeto para que não houvesse desperdicio de memória, gerando mais performance. Com um Data Mart, o agricultor que gerência uma lavoura visualiza o valor da produtividade, como também a sua produtividade em função do tempo; a produtividade em função da região; a produtividade em relação ao grão que está sendo colhido; entre outras formas. Tornando esse mecanismo parte crucial durante a tomada de decisão para futuras plantações.

Agradecimentos

Agradeço a UEPG, a professora Maria Salete Marcon Gomes Vaz e ao co-orientador Luiz Pedro Petroski pela oprtunidade e por me guiarem nos momentos necessários para a conclusão do projeto.

Referências

MOLIN, José P.; BORÉM, A. Geração e interpretação de mapas de produtividade para agricultura de precisão. **BORÉM, A.; GIÚDICE, MP; QUEIROZ, DM; MANTOVANI, EC**, p. 237-58, 2000.

NILAKANTA, Sree; SCHEIBE, Kevin; RAI, Anil. Dimensional issues in agricultural data warehouse designs. **Computers and electronics in agriculture**, v. 60, n. 2, p. 263-278, 2008.

BALDO, Fabiano; ARAUJO, Gilmar; DOS RIOS, Felipe Carlos. Desafios no Projeto de Data Warehouses-Um Estudo de Caso Aplicado à Agricultura de Precisão. In: **SBBD (Short Papers)**. 2011. p. 73-80.