

Unidade Curricular: Sistemas de Informação II

Ficha T. Prática n.º 3

Objectivo:

O objetivo desta ficha é perceber como é elaborada a modelação dimensional de um Data Warehouse.

Requisitos:

Para a realização desta ficha deve ter instalado o SQL Server e ter acesso a uma ferramenta de modelação de dados

Caso: Supermercado

De forma a demonstrar o processo em quatro etapas para a identificação de um modelo multidimensional adequado a uma determinada realidade organizacional, vamos considerar para esta ficha um exercício retirado do livro Business Intelligence Da Informação ao Conhecimento, 3ª Edição, FCA, ISBN: 9789727228805 de Isabel Ramos e Maribel Yasmina Santos (2017). O exercício apresenta o exemplo de um talão de uma venda de supermercado (imagem 1), associado a uma qualquer cadeia de retalho, com diversas lojas numa cidade ou conjunto de cidades. Os dados das vendas são recolhidos nas caixas, com a indicação dos diversos produtos adquiridos por cada cliente. Em termos de estratégia de negócio para a atração de novos clientes, ou retenção dos já existentes, podem ser desencadeados diversas ações de marketing que podem passar, por exemplo, por promoções de produtos. Para poder vender, a cadeia de retalho, necessita comprar, encomendar, entre outros processos de suporte às atividades do dia a dia desta organização.

Supermercado APAD 2017/2018, LDA Av. Cor. José Maria Vale de Andrade Cidade de Viseu Nif: 680033548		
Fatura Simplificada Nº 354769282	17/05/2018	16:45
Loja: L001		
Caixa: F0234(Ana Silva)		
P000123456 Papel A4 80gr.	-----	22,50€
5*4,5€		
P000345678 Capa Arquivo A4 Larga	-----	6,00€
2*3€		
P000456789 Capa Arquivo A4 Fina	-----	7,50€
3*2,5€		
Total		36,00€
Valor Pago (Multibanco)		36,00€
Obrigado pela sua visita		

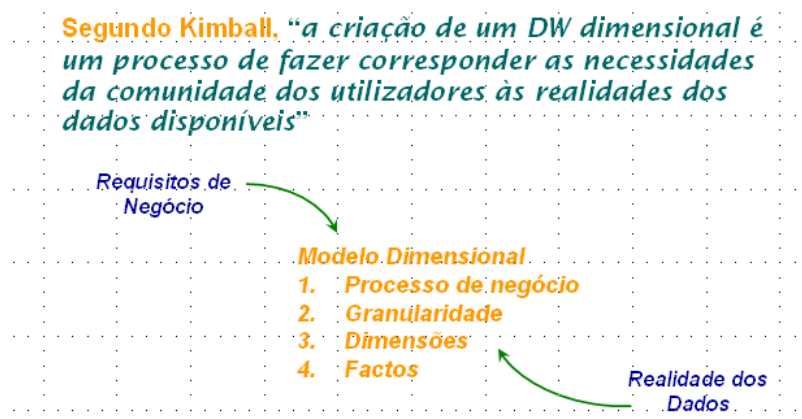
Figura 1 - Talão de venda

No caso do talão de vendas apresentado, e tentando restringir o exemplo, espera-se perceber o perfil de aquisições por parte dos clientes, de forma a disponibilizar indicadores para a gestão do negócio.

Tarefas propostas:

Vamos seguir processo proposto por Kimball para desenvolver o modelo dimensional para o DW deste exercício.

Uma vez que este é o primeiro exercício prático, e de forma a perceberem melhor a modelação dimensional, a primeira e a segunda etapa do processo Kimball, já se encontra resolvida.



1. Primeira Etapa - Começamos por seleccionar o processo de negócio a modelar.

É certo que o primeiro modelo a construir será aquele que tenha o maior impacto – deve responder às questões de negócio mais prementes e cujos dados sejam imediatamente disponíveis.

- Resposta - O **processo de negócio a modelar para o armazenamento dos dados está associado às vendas**, e deve guardar no repositório analítico (DataWarehouse) os dados associados, que permitirá perceber que produtos são

vendidos ao longo do tempo, em que lojas e com que sazonalidade.

2. Segunda Etapa - Avaliar a granularidade: qual o nível de detalhe mais fino que será disponibilizado no modelo dimensional.

Identificar ou definir o grão (ou granularidade) significa especificar o que representa cada uma das linhas que serão armazenadas na(s) tabela(s) de factos, indicando qual é o detalhe que está associado a cada uma das métricas que serão consideradas. Este detalhe é influenciado pelos sistemas operacionais que suportam o negócio, uma vez que os mesmos serão as fontes dos dados para o carregamento do DW.

A granularidade deve ser definida em termos do negócio, idealmente através de formulações como "Uma linha por produto adquirido numa compra", "Uma linha por mês por conta bancária", entre outros.

Estas formulações irão permitir identificar, nos próximos passos, as diversas dimensões e as métricas(factos) associados ao grão definido. Uma vez nestes próximos passos e no caso do grão estar, por alguma razão, mal formulado, é sempre possível o retrocesso a esta etapa para reformulação.

Nesta etapa importa agora identificar o grão a considerar.

- Resposta - No caso do exemplo do retalho, e de forma a considerar a informação mais detalhada possível recolhida pelo processo de vendas, o grão está associado ao **"registo de cada venda de produto a um dado cliente", "numa dada loja e num determinado momento", correspondendo a todas as transações confirmadas nas caixas registadoras.**

3. Terceira Etapa - Selecionar as dimensões que serão incluídas no modelo dimensional.

As dimensões constituem as descrições que irão estar associadas às tabelas de factos, e que permitem dar contexto às métricas que estão a ser analisadas.

Nota: uma vez definida a granularidade, as dimensões emergem naturalmente, uma vez que representam "o quê", "quem", "quando", "onde", etc...

Para o ponto 3, deve identificar as dimensões. As mesmas devem ser detalhadas indicando todos os atributos que as constituem.

4. Quarta Etapa - Selecionar os factos que serão incluídos no modelo dimensional.

Os factos representam o que está a acontecer nos processos de negócio, sendo métricas que permitem medir o desempenho do negócio. Todos os factos devem estar de acordo com a granularidade definida na etapa 2. No caso de serem identificados factos com um grão diferente, os mesmos devem ser armazenados numa tabela de factos diferente, com o seu grão definido para essa realidade (Nota: Não é este o caso). Os factos são métricas numéricas aditivas, tal como a quantidade de alunos que chumbam ou o valor da venda.

Neste ponto deve definir os factos a incluir na tabela de factos.

5. Construa o modelo em estrela concebido, incluindo os atributos que julgar mais significativos.

Percorridas as quatro etapas, deve agora definir o modelo multidimensional. Pode utilizar uma ferramenta de modelação, como o powerdesigner ou utilizar a plataforma disponível em <https://onda.dei.uc.pt>.

6. Altere para o modelo em floco-de-neve. Será justificável esta abordagem?

Dica: Um esquema em floco de neve introduz um refinamento, adicionando as hierarquias de forma explícita, normalizando as tabelas dimensão.

7. Imagine após uma solicitação do departamento contabilístico, que era necessário conhecer as existências de produtos, representando o inventário em cada loja, registado numa base diária. Elabore um Modelo em constelação para resolver o problema.

Nota 1: O modelo em estrela, elaborado anteriormente, deve ser complementado com uma tabela de factos que permita conhecer a existência de produtos, registado numa base diária.

Nota 2: O modelo em estrela, utiliza uma granularidade ao minuto para a dimensão tempo. A nova tabela de factos necessita de uma granularidade ao dia.

7. Elabore o código SQL para o seu modelo em estrela. Implemente o mesmo no SGBD SQL Server. Deve povoar a bd com registos e testar fazendo algumas pesquisas.

Bibliografia

Isabel Ramos e Maribel Yasmina Santos (2017), Business Intelligence Da Informação ao Conhecimento, 3ª Edição, FCA, ISBN: 9789727228805