

SISTEMAS DE APOIO À INTELIGÊNCIA DE NEGÓCIOS

Asterio K. Tanaka

<http://www.uniriotec.br/~tanaka/SAIN>

tanaka@uniriotec.br



Glossário de Modelagem Dimensional

Kimball, Ralph; Ross, Margy. The Data Warehouse Toolkit. John Wiley, 2002.

Accumulating Snapshot Fact Table

- Tipo de tabela de fato com datas múltiplas representando os principais pontos de controle de um processo ou pipeline relativamente de curta duração. A tabela de fato é revisitada e atualizada quando uma atividade ocorre. Um registro é inserido numa tabela de fatos instantâneos acumulados apenas uma vez, quando o item que representa é inicialmente criado. Contrasta com Periodic Snapshot Fact Table e Transaction-Grain Fact Table.

Additive Facts

- Medidas numa tabela de fatos que podem ser somadas ao longo de todas as dimensões.
Taxas e preços unitários geralmente não são aditivos.

Associative Table

- See Bridge Table

Audit Dimension

- Uma dimensão especial que assinala cada linha da tabela de fatos com metadado operacional (por exemplo, data lineage and confidence) quando a linha é criada.

Bridge Table

- Uma tabela com uma chave multiparte capturando um relacionamento muitos-para-muitos que não possa ser acomodado pela granularidade natural de uma tabela de fatos ou tabela de dimensão. Serve como uma ponte entre a tabela de fatos e a tabela de dimensão de forma a permitir dimensões multivaloradas ou hierarquias ragged. Também referida como helper ou associative table.

Causal (factor or dimension)

- Algo que é pensado como a causa de outra coisa. Fatores causais em vendas a varejo se referem a anúncios, displays, coupons, e redução de preços. Uma dimensão causal descreve esses fatores causais.

Centipede Fact Table

- Uma tabela de fatos com muitas dimensões (mais do que 20) levando a uma esquema que lembra um centípedo com numerosas chaves estrangeiras juntadas a numerosas tabelas de dimensão. Centopéias tipicamente resultam quando projetistas tentam representar relacionamentos hierárquicos com uma proliferação de dimensões separadas ao invés de aninhá-las numa única dimensão.

Conformed Dimensions

- Dimensões são conformadas quando são ou exatamente as mesmas (incluindo as chaves) ou uma é um subconjunto perfeito das outras. Mais importante, os cabeçalhos de linha produzidos em conjuntos resposta de duas dimensões conformadas diferentes devem ser capazes de serem casados perfeitamente.

Conformed Facts

- Fatos de múltiplas tabelas de fatos são conformados quando as definições técnicas dos fatos são equivalentes. Fatos conformados são permitidos ter o mesmo nome em tabelas separadas e podem ser combinados e comparados matematicamente. Se os fatos não conformam, então às diferentes interpretações devem ser dados diferentes nomes.

Continuously Valued Facts

- Medida numérica que usualmente é diferente a cada vez que é medida. Deveriam ser fatos numa tabela de fatos em oposição a atributos discretos que pertencem a uma tabela de dimensão.

Coverage Table for a Promotion

- Uma tabela de fatos, tipicamente sem fatos, que registra todos os produtos que estão numa promoção numa determinada loja, independentemente de ser vendidos ou não.

Custom Line-of-Business Table

- A tabela de fatos ou de dimensões numa situação de produtos heterogêneos que contenha fatos ou atributos específicos para um conjunto de produtos, onde esses fatos ou atributos são incompatíveis com os outros conjuntos de produtos. Contrasta com Core Tables. Veja também Heterogeneous Products.

Degenerate Dimension

- Uma chave de dimensão, como o número de uma transação, número de fatura, de ticket, ou de bill-of-lading, que não tenha nenhum atributo portanto não se junta com uma tabela de dimensão.

Demographic MiniDimension

- Ver Minidimension

Dirty Customer Dimension

- Dimensão Cliente em que a mesma pessoa pode aparecer muitas vezes, provavelmente não com exatamente o mesmo nome soletrado ou outros atributos.

Discrete (dimension attributes)

- Dados, usualmente textuais, que tem um número fixo de valores, como sabor de um produto. Dados textuais discretos sempre deveriam ser manipulados como atributos numa tabela de dimensão ao contrário de dados numéricos com valores contínuos que pertencem a uma tabela de fatos.

Event-tracking table

- Uma tabela de fatos, frequentemente sem fatos, onde as dimensões da tabela são trazidas junto para descrever um evento, como numa descrição do seguro de um acidente de automóvel.

Fact

- Uma medida de desempenho do negócio, tipicamente numérica e aditiva, que é armazenada numa tabela de fatos.

Fact Table

- Num esquema estrela, a tabela central com medidas numéricas de desempenho caracterizadas por uma chave composta, cada um de cujos elementos é uma chave estrangeira trazida de uma tabela de dimensão.

Factless Fact Table

- Uma tabela de fatos que não tem fatos mas captura alguns relacionamentos muitos-para-muitos entre chaves de dimensões. Mais frequentemente usada para representar eventos ou prover informação de cobertura que não aparece em outras tabelas de fatos.

HelperTable

- Ver Bridge Table.

Heterogeneous Products

- Um conjunto de produtos tipicamente caracterizado por muitos atributos de produto e medidas de fatos incompatíveis. Um desafio característico de design em ambientes de serviços financeiros. Ver Core Table e Custom-line-of-business Table.

Impact Report

- Ao gerar relatórios com uma tabela ponte, o fator de peso atribuído na dimensão multivalorada é ignorado. Os totais resultantes provêm uma sumarização para cada caso em que a dimensão multivalorada foi envolvida, independentemente da extensão do envolvimento. Contrasta com Weighted Report.

Junk Dimension

- Uma dimensão abstrata com a decodificação de um grupo de flags e indicadores de baixa cardinalidade, portanto removendo os flags da tabela de fatos.

Low cardinality Attribute Set

- Um conjunto de atributos que tenham baixa cardinalidade relativa ao número de linhas de uma dimensão base, tal como dados demográficos para um conjunto de clientes. Pode ser manipulado como uma dimensão outrigger. Ver cardinality.

Many-valued dimensions

- Normalmente, uma tabela de fatos possui conexões somente para dimensões representando um valor simples, como uma data ou produto. Mas ocasionalmente, é válido conectar um registro de fato a uma dimensão representando um número aberto de valores, como o número de diagnósticos simultâneos que um paciente pode ter num momento de um mesmo tratamento. Neste caso, dizemos que a tabela de fatos tem uma dimensão multivalorada. Tipicamente manipulada por uma tabela ponte.

Minidimensions

- Subconjuntos de uma dimensão grande, como Cliente, que são quebrados em dimensões artificiais menores para controlar o crescimento explosivo de uma dimensão grande, com mudança rápida. Os atributos demográficos continuamente mudáveis de um cliente são frequentemente modelados como uma minidimensão separada.

Most Recent Indicator

- Um atributo, tipicamente usado em conjunção com dimensões que mudam lentamente do tipo 2, que indica o perfil mais corrente.

Multivalued dimensions

- Ver many valued dimensions.

Non additive (facts)

- Um fato que não pode logicamente ser somado entre linhas. Pode ser numérico e portanto deve ser combinado com outros fatos antes de ser somado por linhas. Se não numérico, podem somente ser usado em restrições, contagens ou agrupamentos.

Outrigger Table

- Uma tabela de dimensão secundária anexada a uma tabela de dimensão. Uma outrigger table é uma interpretação de projeto físico de uma tabela de dimensão lógica. Ocorre quando uma tabela de dimensão é snowflaked.

Partitioned Tables

- Tabelas (e seus índices associados) que são gerenciadas como tabelas fisicamente separadas mas parecem logicamente como um tabela única. Tabelas de fato grandes são candidatas para particionamento, frequentemente por data. Particionamento pode melhorar desempenho de consulta e manutenção.

Partitioning of History

- A correspondência natural entre entradas de tabelas de dimensão e linhas de tabela de fatos quando um SCD tipo 2 foi implementado. Um SCD tipo 2 particiona o histórico porque cada valor de sua chave surrogada é administrada corretamente para conectar ao span contemporâneo correto dos registro de fatos.

Periodic Snapshot Fact Table

- Um tipo de tabela de fatos que representa desempenho de negócio ao fim de cada período de tempo regular previsível. Snapshots diários e mensais são comuns. Snapshots são requeridos em vários negócios, como seguros, onde o histórico de transações é tão complicado para ser usado como base snapshots computados on the fly. Um registro separado é colocado num PSFT cada período, independentemente se alguma atividade ocorreu na conta. Contrasta com Transaction Fact Table e Accumulating Snapshot Fact Table.

Ragged Hierarchy

- Uma hierarquia com estrutura desbalanceada e arbitrariamente profunda que usualmente não pode ser descrita antes da carga de dados. Algumas vezes referida como variable-depth hierarchy. Mapas organizacionais são hierarquias ragged. Ver Bridge Tables.

Role playing dimensions

- A situação onde uma simples tabela de dimensão física aparece várias vezes na mesma tabela de fatos. Cada um dos papéis da dimensão é representado por uma tabela lógica separada com nomes de coluna únicos através de visões.

Semiaddictive fact

- Fato numérico que pode ser somado ao longo de algumas dimensões na tabela de fatos mas não ao longo de outras. Níveis de estoque e saldos não podem ser somados ao longo da dimensão tempo, mas podem ser averaged usualmente sob a mesma dimensão.

Sparsity Failure

- Uma situação que ocorre quando uma tabela agregada é criada que não seja apreciavelmente menor que a tabela do que a tabela em que foi baseada. Por exemplo, se somente um SKU de cada marca for vendida num dado dia, então o agregado por marca para um dia será o mesmo tamanho da tabela base.

Transaction Fact Table

- Tipo de tabela de fatos em que a granularidade é uma linha para o menor nível de detalhe capturado por uma transação. Um registro numa TFT está presente somente se um evento de transação ocorre. Contrasta com Periodic Snapshot Fact Table e Accumulating Snapshot Fact Table.

Value Banding Facts

- Grupando fatos em faixas de valor flexíveis como especificado numa tabela de definição de faixas.

Variable-Width Attribute Set

- A situação onde um número variável de atributos de dimensão é conhecido, dependendo de duração do relacionamento como com prospectos que evoluem para clientes.

Weighted Report

- Quando usando uma tabela ponte, os fatos na tabela de fatos são multiplicados pelo fator de peso da tabela ponte, para alocar apropriadamente os fatos às dimensões multivaloradas. Contrasta com Impact Report.