Conceitos Utilizados no Pentaho Open BI Suíte

Data Warehouse (DW)

É uma grande base de dados, integrada com outros ambientes operacionais, dos quais recebe informações, posteriormente tratadas, arquivadas por assunto, armazenadas e padronizadas em um único *armazém* de dados que, depois de tratados e inseridos, não podem mais ser alterados. O objetivo dessa fonte única e concentrada é permitir uma consulta rápida, fácil e confiável de informações íntegras, que servirão de suporte às tomadas de decisão gerenciais.

Data Mart

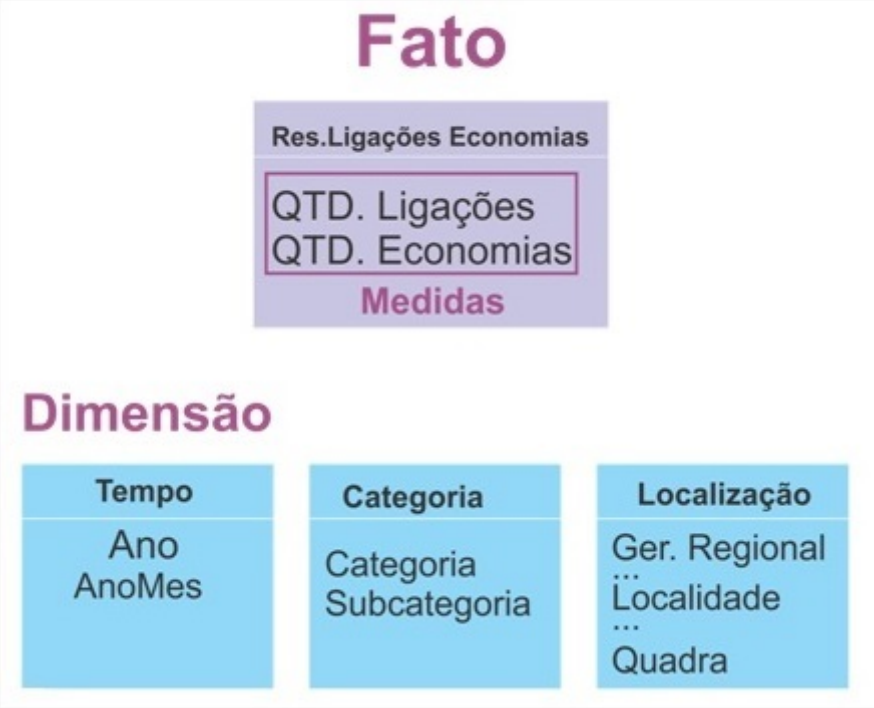
Se o Data Warehouse é um *armazém* de dados íntegros, o Data Mart pode ser definido como um *departamento* deste armazém. Ou seja, trata-se de um subconjunto lógico do Data Warehouse, recortado de forma personalizada de uma base de dados maior.

Modelagem Dimensional

Modelagem dimensional é uma técnica de projeto lógico que contrasta com a modelagem entidade-relacionamento e normalmente é usada para DW. É a única técnica viável para bancos de dados que devem responder a consultas em um DW.

Modelo Dimensional

É um modelo dimensional composto por três elementos básicos: **Fato, Dimensões e Medidas**.



Exemplo de Fato, Dimensão e Medida.

Fato

Guarda as medidas do negócio, que por sua vez têm uma das intersecções de todas as dimensões. Além disso, fatos são aditivos, ou seja, podem ser acumulados e têm valores contínuos. Trata-se de uma coleção de itens de dados, composta de medidas, utilizada para analisar o processo de negócio de uma empresa, daí refletir a evolução dos negócios no dia-a-dia. O fato é elaborado em tabelas denominadas tabelas de fato (fact tables) e representado por valores numéricos.

Dimensões

Guardam os atributos do negócio e são utilizadas para restringir pesquisas feitas, além de servirem como título das colunas em consultas. São elementos que participam de um fato e que determinam o contexto de um assunto de negócios. As dimensões podem ser compostas por membros que podem conter hierarquias. Membros são as possíveis divisões ou classificações de uma dimensão. Por exemplo, a dimensão tempo, pode ser dividida nos seguintes membros: ano, trimestre e mês, e a dimensão localização em: cidade, estado e país.

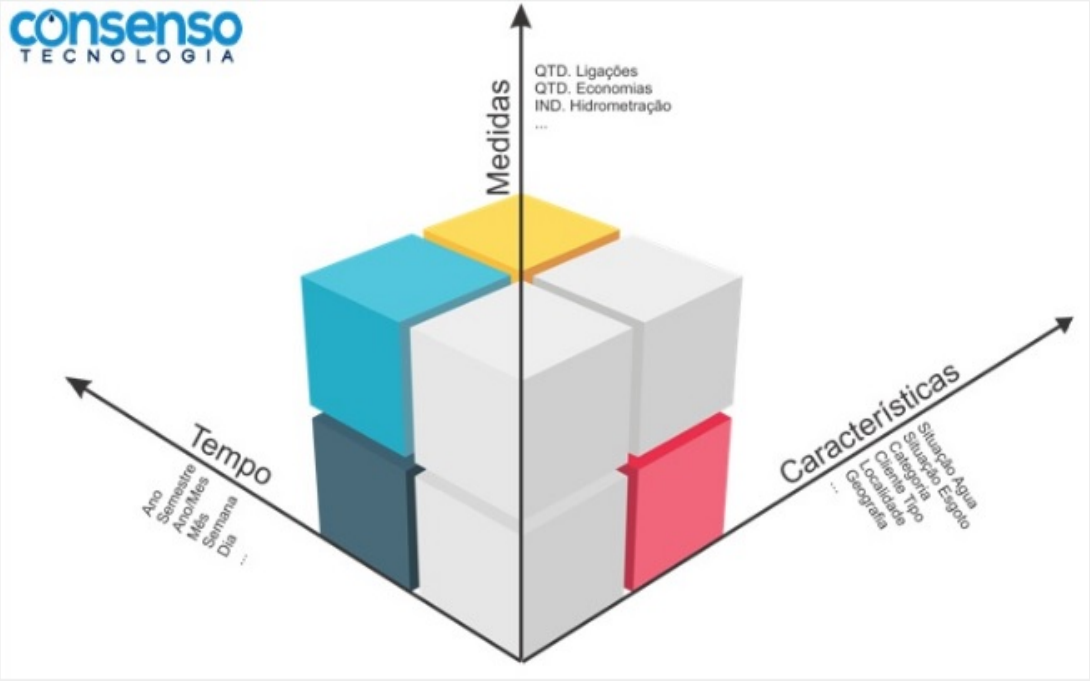
Medidas (variáveis)

São atributos numéricos que representam um fato, isto é, que representam o desempenho de um indicador de negócios relativo às dimensões que participam desse fato. Uma medida é determinada pela combinação das dimensões que participam de um fato e estão localizados como atributos de um fato. Por exemplo, o valor em reais das vendas, o número vendido de unidades de produtos e a

quantidade em estoque.

OLAP

São estruturas de dados que possibilitam que as informações sejam estruturadas em forma de **cubos**, os quais armazenam dimensões e medidas. Trata-se de uma tecnologia sofisticada que utiliza estruturas multidimensionais para fornecer acesso rápido aos dados para análise, facilitando a exibição de resumos de alto nível:



Exemplo de uma Estrutura de Dados OLAP.

Um modelo de dados OLAP é uma tecnologia criada para facilitar a consulta dinâmica e em tempo real da informação armazenada. Em um modelo de dados OLAP, a informação é organizada em **cubos** que armazenam valores quantitativos ou medidas.

As medidas são identificadas por duas ou mais categorias descritivas denominadas dimensões, que formam a estrutura de um **cubo**. Uma dimensão pode ser qualquer visão do negócio que faça sentido para a sua análise, como **Categoria, Situação de água, Localização ou Tempo**.

Este modelo de dados multidimensional simplifica para os usuários o processo de formular pesquisas ou *queries* complexos, criar relatórios, efetuar análises comparativas e visualizar subconjuntos (slice) de maior interesse.

Por exemplo, um **cubo** contendo informações de Ligações e Economias, poderá ser composto pelas dimensões **Tempo, Localização, Categoria, Situação Água e Medidas. Medidas típicas seriam valor de QTD. Ligações, QTD.Economias, IND. Hidrometração**, etc.

Dentro de cada dimensão de um modelo **OLAP**, os dados podem ser organizados em uma hierarquia

que define diferentes níveis de detalhe. Por exemplo, dentro da dimensão tempo, podemos ter uma hierarquia representando os níveis anos, meses, e dias. Da mesma forma, a dimensão Localização poderá ter os níveis Gerência Regional, Unidade, Localidade, Setor e Quadra.

Clique aqui para retornar ao Menu Principal do GSAN

From:
https://www.gsan.com.br/ - **Base de Conhecimento de Gestão Comercial de Saneamento**

Permanent link:
https://www.gsan.com.br/doku.php?id=ajuda:gerencial:conceitos_utilizados_no_pentaho_open_bi_suite

Last update: **31/08/2017 01:11**