

Tipos de métricas existentes no Data Warehouse

No desenvolvimento de um Data Warehouse (DW) devemos estar atentos aos tipos de métricas que iremos trabalhar. Engana-se quem pensa que métrica é tudo igual, e que seu comportamento é o mesmo de todas as outras.

As ações efetuadas através das ferramentas OLAP (On-line Analytical Processing) para a geração das análises são distintas a depender do tipo de métrica existente. Elas podem ter regras de agregações (sumarização) como somatório, média, mínimo, máximo, etc.

Para entendermos melhor, vamos conhecer os tipos de métricas que normalmente podem estar contidos dentro de um DW. São três os tipos: métricas aditivas, semi-aditivas e não-aditivas.

As métricas aditivas são aquelas que podem ser sumarizadas independente das dimensões utilizadas. Este tipo de métrica pode ser utilizada sem quase nenhuma restrição ou limitação e são flexíveis o suficiente para gerar informações em quaisquer perspectivas. Por exemplo, métricas como quantidade e valores de determinados itens podem ser, em geral, sumarizados por data (dia, mês ou ano), local, clientes, entre outras dimensões, sem perder a consistência da informação.

As métricas semi-aditivas são aquelas que podem ser sumarizadas em alguns casos. Isso porque a depender da situação empregada à métrica, ela pode perder sentido para a análise caso seja agregada. Neste caso a sumarização só fará sentido com algumas dimensões específicas. Por exemplo, a métrica: saldo bancário. O saldo é um valor que reflete a situação atual da conta, que pode ter o saldo credor ou devedor. Faria sentido, por exemplo, somar os saldos de todos os dias de um mês para uma determinada conta bancária? Claro que não. Pois se um dia o saldo for de -1000 e no dia seguinte ter os mesmos -1000, a soma irá me devolver um saldo negativo de -2000, o que não é verdade. Mas há casos onde a métrica semi-aditiva adquire característica de aditiva. Se por acaso somar os saldos de várias contas bancária em um determinado dia, poderemos ver o saldo geral, o que tem total sentido e utilidade para uma instituição bancária, por exemplo.

As métricas não-aditivas são aquelas que não podem ser sumarizadas ao longo das dimensões. Essas métricas não podem ter agregações pois perdem a veracidade do valor. Percentuais são exemplos de valores armazenados nas métricas que não permitem sumarizações. Por exemplo, não faz sentido algum somar o percentual de vendas de um item "A" que teve 50% de saída com um item "B" que teve 60%. A soma resultaria em um valor agregado de 110%. O que isso nos diz? Nada!

Como muitas vezes as métricas semi-aditivas e não-aditivas são derivadas de métricas aditivas, recomendo, se possível, que sejam armazenadas as métricas brutas para o cálculo em tempo de execução. A métrica semi-aditiva saldo, por exemplo, pode ser calculada em tempo de execução com as métricas aditivas do valor de crédito e débito.