Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Measure

Pedersen, Torben Bach

Published in:
Springer Encyclopedia of Database Systems

Publication date:
2009

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
Pedersen, T. B. (2009). Measure. In Springer Encyclopedia of Database Systems (2009 ed., pp. 1703-1704). Berlin/Heidelberg: Springer.

Measure

Torben Bach Pedersen, Aalborg University, Denmark, http://www.cs.aau.dk/ tbp

SYNONYMS

Numerical fact (used by reference [1])

DEFINITION

A *measure* is a numerical property of a multidimensional cube, e.g., sales price, coupled with an aggregation formula, e.g., SUM. It captures numerical information to be used for aggregate computations.

MAIN TEXT

As an example, a 3-dimensional cube for capturing sales may have a Product dimension P, a Time dimension T, and a Store dimension S, capturing the product sold, the time of sale, and the store it was sold in, for each sale, respectively. The cube has two measures: DollarSales and ItemSales, capturing the sales price and the number of items sold, respectively. ItemSales can be viewed as a function: ItemSales: $Dom(P) \times Dom(T) \times Dom(S) \mapsto \mathbb{N}_0$ that given a certain combination of dimension values returns the total number of items sold for that combination. If a dimension value corresponds to a higher level in the dimension hierarchy, e.g., a product group or even *all* products, the result is an aggregation of several lower-level measure values.

In a multidimensional database, measures generally represent the properties of the chosen facts that the users want to study, e.g., with the purpose of optimizing them. Measures then take on different values for different combinations of dimension values. The property and formula are chosen such that the value of a measure is meaningful for all combinations of aggregation levels. The formula is defined in the metadata and thus not stored redundantly with the data. Although most multidimensional data models have measures, some do not. In these, dimension values are also used for computations, thus obviating the need for measures, but at the expense of some user-friendliness [2].

It is important to distinguish among three classes of measures, namely *additive*, *semi-additive*, and *non-additive* measures, as these behave quite differently in computations. Additive measure values can be combined meaningfully along any dimension. For example, it makes sense to add the total sales over Product, Store, and Time, as this causes no overlap among the real-world phenomena that caused the individual values. Semi-additive measure values cannot be combined along one or more of the dimensions, most often the Time dimension. Semi-additive measures generally occur for so-called "snapshot" facts. For example, it does not make sense to sum inventory levels across time, as the same inventory item, e.g., a specific product item, may be counted several times, but it is meaningful to sum inventory levels across products and stores. Non-additive measure values cannot be combined along any dimension, usually because of the chosen formula. For example, this occurs when averages for lower-level values cannot be combined into averages for higher-level values. The additivity of measures is related to the so-called "type" of the measure (Flow, Stock or Value-Per-Unit), please refer to the Summarizability entry for details.

CROSS REFERENCE*

Cube, Dimension, Hierarchy, Multidimensional modeling, On-Line Analytical Processing, Summarizability.

RECOMMENDED READING

Between 5 and 15 citations to important literature, e.g., in journals, conference proceedings, and websites.

[1] R. Kimball, L. Reeves, M. Ross, and W. Thornthwaite. *The Data Warehouse Lifecycle Toolkit.* Wiley Computer Publishing, 1998.

[2] T. B. Pedersen, C. S. Jensen, and C. E. Dyreson. A Foundation for Capturing and Querying Complex multidimensional data. *Information Systems*, 26(5):383–423, 2001.

[3] E. Thomsen. *OLAP Solutions: Building Multidimensional Information Systems.* Wiley, 1997.