

✦ EXCLUSIVELY FOR TDWI MEMBERS

FOURTH QUARTER 2009

# TEN MISTAKES TO AVOID

---

## In Dimensional Design

By Christopher Adamson



# 10 TEN MISTAKES TO AVOID

---

## In Dimensional Design

By Christopher Adamson

### FOREWORD

In virtually every data warehouse implementation, you can find the products of dimensional design: the star schema, the snowflake, or the cube. Despite this near-universal acceptance, the basic principles of dimensional design are commonly misunderstood and misapplied.

Many mistakes are errors in approach, committed before design work begins. Chief among these is the failure to exploit strategic functions offered by a dimensional model. Others include waiting until the design stage to do dimensional design, then leaving the design work to designers.

Success is often thwarted by common technical errors, such as sacrificing operational detail or taking shortcuts with historic data. Designers also doom their solutions by failing to adapt them to the software tools that comprise the data warehouse.

Avoiding these 10 mistakes in dimensional design will help you ensure successful implementations in any data warehouse architecture, including those advocated by Ralph Kimball and W.H. Inmon.

### ABOUT THE AUTHOR

Christopher Adamson is an independent consultant, educator, and author. He works with customers in all industries to develop data warehouse strategies, define and prioritize projects, and design solutions. Chris is a faculty member at TDWI and teaches classes in dimensional design. He is the author of *Mastering Data Warehouse Aggregates* (John Wiley & Sons, 2006) and *Data Warehouse Design Solutions* (John Wiley & Sons, 1998). Chris blogs about data warehousing at [www.starschemacentral.com](http://www.starschemacentral.com) and can be contacted at [chris.adamson@oaktonsoftware.com](mailto:chris.adamson@oaktonsoftware.com).

---

© 2009 by TDWI (The Data Warehousing Institute™), a division of 1105 Media, Inc. All rights reserved. Reproductions in whole or in part are prohibited except by written permission. E-mail requests or feedback to [info@tdwi.org](mailto:info@tdwi.org).

Product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

## MISTAKE TEN: TRYING TO SAVE SPACE

---

# 10

In a dimensional model, the context of each measurement is defined by *dimensions*. A rich set of dimensions enables diverse and powerful analytic possibilities. All too often, well-meaning designers cripple this potential in a misguided effort to save space.

In operational systems, it is common practice to divide data elements into constituent parts. From these components it is possible to construct any combinations that may be needed. For example, a name may be broken down and stored as a first name, middle initial, and last name. In a query, these attributes can be combined to produce a full name.

In a dimensional model, these same techniques are counterproductive. Dimensions are used to filter queries, group the summarization of facts, govern the ordering of data, define master-detail relationships, and control subtotaling. Computing their values on the fly severely hampers the performance of these operations. The database administrator is unable to use indexes to optimize queries. Users find it difficult to consult lists of values. Worse, reports will likely construct computed attributes inconsistently.

In a dimensional design, explicit storage of each dimension attribute maximizes performance, aids usability, and guarantees consistency across applications. Common permutations are provided. Codes are supplemented with corresponding description values. Flags are translated from Boolean values into descriptive text. Multipart fields are both preserved and broken down into constituent pieces. The extra space consumed is paid for by a more powerful and flexible solution.

## REFERENCES AND RECOMMENDED READING

---

Adamson, Christopher [2006]. *Mastering Data Warehouse Aggregates: Solutions for Star Schema Performance*, John Wiley & Sons.

Adamson, Christopher, and Michael Venerable [1998]. *Data Warehouse Design Solutions*, John Wiley & Sons.

Kimball, Ralph, and Margy Ross [2002]. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*, second edition, John Wiley & Sons.

Kimball, Ralph, Margy Ross, Warren Thornthwaite, Joy Mundy, and Bob Becker [2008]. *The Data Warehouse Lifecycle Toolkit: Practical Techniques for Building Data Warehouse and Business Intelligence Systems*, second edition, John Wiley & Sons.

Imhoff, Claudia, Nicholas Gallemmo, and Jonathan G. Geiger [2003]. *Mastering Data Warehouse Design: Relational and Dimensional Techniques*, John Wiley & Sons.

Inmon, W.H., Claudia Imhoff, and Ryan Sousa [2000]. *Corporate Information Factory*, second edition, John Wiley & Sons.

Reeves, Laura [2009]. *A Manager's Guide to Data Warehousing*, John Wiley & Sons.

## ABOUT **TDWI**

---

TDWI, a division of 1105 Media, Inc., is the premier provider of in-depth, high-quality education and research in the business intelligence and data warehousing industry. TDWI is a comprehensive resource for industry information and professional development opportunities. TDWI sponsors and promotes quarterly World Conferences, regional seminars, onsite courses, a worldwide Membership program, business intelligence certification, resourceful publications, industry news, an in-depth research program, and a comprehensive Web site ([www.tdwi.org](http://www.tdwi.org)).



1201 Monster Road SW  
Suite 250  
Renton, WA 98057

**T** 425.277.9126

**F** 425.687.2842

**E** [info@tdwi.org](mailto:info@tdwi.org)

**[www.tdwi.org](http://www.tdwi.org)**