

## The Seven Deadly Sins of Big Data -- and How to Avoid Them

As we are all too aware, organizations accumulate vast amounts of data from a variety of sources nearly continuously. Big data advocates promise the moon and the stars as you harvest the potential of all these data. There is certainly a lot of hype. There's no doubt that savvy organizations are fueling their strategic decision making with insights from data mining, but what are the challenges?

Much can go wrong in the data mining process, even for trained professionals. In this talk I'll discuss a wide variety of case studies from a range of industries to illustrate the potential dangers and mistakes that can frustrate problem solving and discovery -- and that can unnecessarily waste resources. My goal is that by seeing some of the mistakes I have made, you will learn how to take advantage of data mining insights without committing the "Seven Deadly Sins."

### Bio

Richard De Veaux, Ph.D. (Dick) is Professor of Statistics at Williams College. He holds degrees in Civil Engineering (B.S.E. Princeton), Mathematics (A.B. Princeton), Dance Education (M.A. Stanford) and Statistics (Ph.D., Stanford), where he studied with Persi Diaconis.

Before Williams, Dick taught at the Wharton School and the Engineering School at Princeton. He has also been a visiting research professor at INRA (the Institut National de la Recherche Agronomique) in Montpellier, France; the Université Paul Sabatier in Toulouse, France; and the Université René Descartes in Paris. De Veaux has won numerous teaching awards including a "Lifetime Award for Dedication and Excellence in Teaching" from the Engineering Council at Princeton. He has won both the Wilcoxon and Shewell (twice) awards from the American Society for Quality and was elected a fellow of the ASA in 1998. In 2006-2007 he was the William R. Kenan Jr. Visiting Professor for Distinguished Teaching at Princeton University. In 2008 he was named the Mosteller Statistician of the Year by the Boston Chapter of the American Statistical Association.

Dick has been a consultant for over 30 years for such Fortune 500 companies as Hewlett-Packard, Alcoa, American Express, Bank One, GlaxoSmithKline, Dupont, Pillsbury, Rohm and Haas, Ernst and Young, SanofiPasteur and General Electric. He holds two U.S. patents and is the author of more than 30 refereed journal articles. He is the co-author, with Paul Velleman and David Boeck, of the critically acclaimed textbooks "*Intro Stats*", "*Stats: Modeling the World*" and "*Stats: Data and Models*" and with Norean Sharpe and Paul Velleman of the recently published books, "*Business Statistics*", and "*Business Statistics: A First Course*", all published by Pearson.

His hobbies include cycling, swimming, singing (barbershop, doo wop and classical) -- and dancing (he was once a professional dancer and teaches Modern Dance during Winter Study at Williams). He is the father of four: two boys and two girls all in their 20's.