

# A Better Way to Mine Data

Everybody talks about data mining, but few people do it well. These tips and tools may help.

By Rob Meyer

The good news, for marketers, is that data mining really can make a difference to most bottom lines. The bad news is that, despite what data mining can do, it is so often used so poorly that it is virtually useless.

While that unfortunate reality is recognized by some marketing professionals, it is taken for granted by mathematicians and computer scientists. Mathematicians who are trained in data mining also know that marketers and other executives often have unrealistic expectations about what kind of data analysis it takes to get meaningful results.

In many cases, the data is a big part of the problem. Even in the most reputable companies, data is often "dirty,"--out of date or otherwise irrelevant. And in many more cases, the problem is the software. Most commercially available data mining packages lack the flexibility and functionality that real world marketers need. The closer one looks at various data mining software packages, the more discouraging it gets. There are variations in how different packages handle different procedures, and in how they define various algorithms. And even today, there is no cross-industry standard practice for how classification functions deal with ties in the data. No wonder so many marketing experts conclude that data mining is just not worth the effort.

There is no easy fix. Re-inventing routines for data mining can be prohibitively expensive, and homegrown code for data mining may not be fully trustworthy, given the requirements of computations handling terabyte-sized data.

The [Numerical Algorithms Group \(NAG\)](#), an organization that develops software to solve complex mathematical problems, has three suggestions. One, try hiring a mathematician who is a data-mining expert to guide your efforts. Two, consider developing data mining applications in-house using fully documented components (algorithms) from a reliable library. This [white paper](#), should help to explain how that can be done.

And finally, don't give up. When data mining works, it is well worth the effort.

---

*Rob Meyer is CEO of the Numerical Algorithms Group ([www.nag.com](http://www.nag.com)), a worldwide organization that develops mathematical, statistical, and data mining components to solve complex mathematical problems.*

---

Originally published by CMO in February 2005