

Paper 434-2013

From Big Data to Big Statistics

John Sall, SAS

ABSTRACT

Now that we have lots of data and can process it amazingly fast, we still need ways to look at it without being overwhelmed. We don't want to look at 10,000 graphs--we want one graph that shows the bright spots among 10,000 graphs. We need volcano plots and false-discovery-rate plots. We want the computer and software to do the work of finding what is most interesting and bringing it to our attention. We want our results sorted and summarized, but with access to the detail we need to understand it. Also, when we look at the most significant of thousands of statistical tests, we want to know if we are seeing random coincidence selected out of thousands, or if we are seeing real effects.

No paper was submitted for publication in the *Proceedings*. Check <http://support.sas.com/rnd/papers/> or contact the author.

CONTACT INFORMATION

John Sall
SAS
John.Sall@jmp.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Other brand and product names are trademarks of their respective companies.