

What Kind of Big Data Problem Do You Have?

Everyone's still talking about big data - and for good reason. It's continuing to grow as the use of computers, mobile devices and the Internet continues to increase. It's expanding as more devices, homes and machines are outfitted with sensors, smart meters and GIS transmitters. And with the advent of low-cost storage, in-memory analytics and other computing technologies, it's possible to turn this big data into meaningful insights that empower organizations to make proactive, more informed decisions than ever before.

What's more, every organization - large and small - has the potential to benefit from big data. Why? Because so many of today's big data sources are public. Think open government data, weather and meteorological data, Twitter and more - big data is out there, free and waiting for you to analyze it. Even small players in industries can afford to use it to get ahead of the competition by being the first to store and incorporate open data sources into their existing analytical work.

- Don't assume every analytical problem is a big data problem.
- Many analytical problems can be solved without big data storage or big data analytics.
- Understand the root cause of analytic problems first - and then select the right analytic technologies to solve them.
- No one type of analytics is better than the other. Each does different things, solves different problems and requires different software and architectures.

But is every analytical problem a big data problem? Does everyone need big data analytics?

Given all the buzz about big data, it's easy to think so. Nearly every technology vendor is offering big data or big analytics solutions today. But there are many analytical problems that can be solved without big data storage or big data analytics. And assuming that every analytical problem is a big data problem can lead to costly investments in technologies that won't actually address your underlying problem.

Nearly every technology vendor is offering big data or big analytics solutions today.

Working with hundreds of customers around the world, SAS has found that when decision makers better understand the types of analytic problems they actually have, they can make more informed and successful investments in technologies to help them solve them. To this end, this paper will give you a crash course in how SAS® categorizes analytic problems and the best technologies to address them.

In many cases, informed customers find that they have a big analytics problem - not a big data problem - which changes everything.

Understanding the Issues

Before delving into some key terms and concepts, let's consider a real-world example to get a better handle on the problem - and on why you should keep reading.

Imagine, for instance, that a company is taking five hours to run a marketing optimization job. They assume they have a big data problem and start exploring big data technologies to address it. But because the decision makers don't fully understand the big data problem and technology landscape, they can't see that the slow processing is not because of too much data - it's because their analytical software isn't designed to efficiently process the many possible offers marketing needs to consider. It must make multiple passes on the data and run complex algorithms with each step - a slow process for the wrong analytical tool.

That's a big analytics problem, not a big data problem. So the company's planned big data solutions won't solve it.

To help your organization avoid pitfalls like this one, let's dig into SAS insights that will help you properly evaluate your organization's problems and identify likely solutions.

Laying the Foundation: A Common Vocabulary for Needs and Problems

First, let's define some key terms and concepts that are fundamental to assessing analytical problems and identifying the right solutions.

Reactive Versus Proactive Analytics

Standard business reports, ad hoc reports, OLAP, and even alerts and notifications based on analytics are examples of reactive analytics. Now, reactive analytics can be very useful - for example, for generating finance and regulatory reporting. Reactive analytics are also used by business users to perform ad hoc analysis every day. But make no mistake - reactive analytics can only inform you about the past.

Proactive analytics - such as optimization, predictive modeling, forecasting and statistical analysis - are forward-looking. They allow you to identify trends, spot weaknesses and determine conditions for making decisions about the future. Proactive analytics include optimization of complex problems with many dependencies, predictive modeling, regression analysis and other advanced methods for proactive decision making.

REACTIVE	PROACTIVE
ALERTS	OPTIMIZATION
OLAP	PREDICTIVE MODELING
AD HOC REPORTS	FORECASTING
STANDARD REPORTS	STATISTICAL ANALYSIS

Figure 1: Reactive and proactive analytics



Big Data Versus Big Analytics

The next step is to define big data. Put simply, when you have exceeded the capacity of conventional database systems, you're dealing with big data. Before that occurs, you have what SAS refers to as "growing data," which is still a large amount of data, but it hasn't hit database limitations yet.

Today, organizations can store vast amounts of data, but processing times have become excessive because traditional storage environments are not conducive for proactive analytics. When you have reached a point where processing times become unacceptable, you may be dealing with big data sizes, but you could also be dealing with a big analytics issue.

To better understand the difference, let's create a chart with reactive and proactive analytics on the Y axis and the size of the data on the X axis, like this:

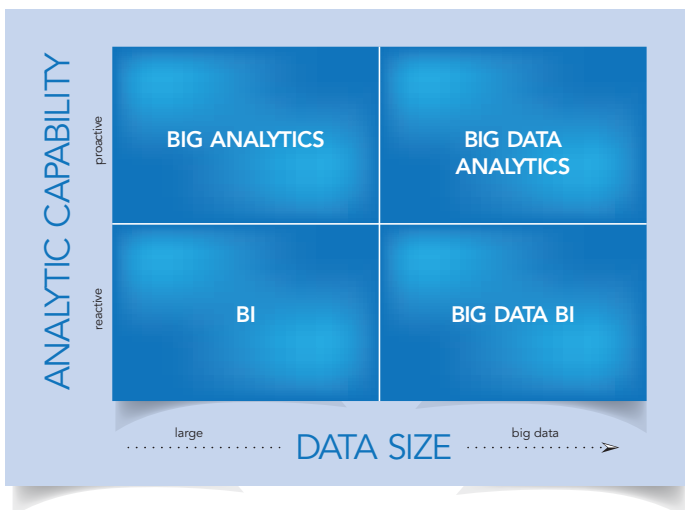


Figure 2: Data size and analytic competence

Now let's overlay the four types of software solutions available in the analytics market to see how they can address different types of analytical problems.

Business Intelligence (BI)

If you are dealing with a large amount of data and providing reporting capabilities to end users so they can access the data, summarize it, and drill down into it, you are dealing with business

intelligence (BI) applications. These solutions provide people with an in-depth look at the past performance of your organization. This is BI, by definition. (See the lower-left quadrant in Figure 2.)

Big Data BI

When data gets bigger and you're dealing with outside data sources or pulling in unstructured data, you'll notice that your data problems will also get bigger. For example, it may take users too long to get the information they need, or they may find it impossible to combine data sources fast enough to provide meet deadlines for reports. Or you may be under pressure to give people quick access to data for forecasting - but you're still limited to providing reactive analytics.

These are problems that need big data BI solutions, as shown in the lower-right quadrant of Figure 2. These examples illustrate the most common big data challenge facing businesses today - and most are trying to solve it with SQL-based solutions, which are not sufficient.

Big Analytics

It takes a different kind of analytics to enable forward-looking decision making. If you're looking at customer preferences, markdown optimizations or fraud predictions, you also need a different type of architecture. These "problems" typically involve growing data sizes and proactive analytics. It's not the size of the data that's slowing you down, it's the fact that you're making multiple passes on data that may take hours to get results, and you're running advanced analytic calculations that take longer to process. For issues that need to be addressed today, you need those answers in seconds or minutes, which requires big analytics solutions. (See the upper-left quadrant of Figure 2.)

Big Data Analytics

Now, what if you have a whole lot of data and are dealing with proactive decision making? Retail companies, for example, may have hundreds of millions of SKUs across multiple retail stores - really big data. Manufacturers and insurers may also be looking at future sources of data as well, such as telematics data in the auto industry. Similarly, manufacturers may want to predict safety problems before they affect customers, and insurance companies may want to adjust rate plans for the best drivers. These are examples of situations where companies face truly big data and big analytic problems - the kind that most businesses haven't needed to deal with in the past. And solving them requires big data analytics. (See the upper-right quadrant of Figure 2.)

So What's the Big "So What?"

The point here is not to say that one type of analytics is better than the other, but rather that they each do different things, solve different problems and require different software and architectures. As you look at what's happening in the market and your business, it's critical that you understand the difference between each of these four areas and how the different problems can be solved.

Analytics continues to be a broad term in the market, but it's worthwhile to look at the problems you are trying to solve and then determine where you fall in this landscape. It will help you plan your next steps in your big data journey.

To learn more about this topic, visit the following links:

- [Big data analytics](#)
- [Business intelligence](#)
- [Big Data BI](#)
- [Big Analytics](#)
- [Big Data Analytics](#)

And to read more from Jim Davis, see his monthly [blog posts](#).



Jim Davis
*Senior Vice President and
Chief Marketing Officer, SAS*

As Senior Vice President and Chief Marketing Officer, Jim Davis provides strategic direction for SAS products, solutions and services and presenting the SAS brand worldwide.

A respected thought leader, he co-authored the book *Information Revolution: Using the Information Evolution Model to Grow Your Business*.

Since joining SAS in 1994, Davis has pioneered many key initiatives, including global projects incorporating customer feedback in the development process so SAS can meet the demand for business analytics.

To contact your local SAS office, please visit: sas.com/offices