



## First Things First; Become “Capable” of Data-Driven Operations

By Richard G. Lamb, PE, CPA, ICBB; Analytics4Strategy.com

“Data-driven” is defined as processes, decisions and activities spurred by data rather than only experience, intuition, culture and politics. However, we need to distinguish data-driven as two challenges.

First, the firm must make itself fully, rather than only partially, capable of being data-driven. Second, the firm must evolve to making the power of the capability integral to its processes, activities and decisions.

This article will deal with the first challenge—becoming data-driven-capable. It will explain that a data-driven capability has the three layers as shown in the header figure. It will explain why it is that firms typically have a culture of systems, software and skills that make the capability an easy and almost costless reach. Finally, it will explain how that culture is the foundation to taking a grassroots strategy to becoming data-driven-capable.

### First Layer: Insight Deliverables

The “products” of the data-driven-capability are insight deliverables supplied at the interface of the capability and operational functioning. The header figure depicts the first layer of the capability as the interface layer. Deliverables along the operational process include system standard reports, task spreadsheets, business intelligence and data analytics.

The power of the data-driven-capability becomes reality as the insight deliverables evolve into the firm’s processes, decisions and actions. The functional roles for specific insight deliverables will emerge as the firm evolves to recognize and specify them as integral to its operational effectiveness.

This means that the data-driven-capability must be able to create insight deliverables without restriction. As ideas for insight deliverables strike us, the data-driven-capability is able to immediately create them. In other words, the unrealistic philosophy that all insight deliverables and their data should be initially planned for and everything else tailored to generate them is obsolete in the context of a data-driven-capability.

At the insight layer of the capability, firms are accustomed to working with the standard reports of their operating systems and fused task spreadsheets. Firms are increasingly utilizing Excel Pivot as the means to generate intelligence from table-formatted system standard reports and in the conduct of ad hoc analyses.

Only new to the insight layer are data analytics—model-based insight deliverables. However, the firm has the insight deliverable of data analytics at its behest. This is because all firms and individuals have right to download, without restriction and cost, the top-tier open system (non-proprietary) software called “R.” Accordingly, firms are now able to ask five types of insight questions that are beyond the possibilities of system standard reports, task spreadsheets and intelligence software.

## **Second Layer: Super Tables**

All insight deliverables depend upon tables—no exceptions. It is how things work. In them are every item (variable) of data that make the deliverable possible.

In fact, tables lurk beneath all system standard reports. Furthermore, all of a system’s data can usually be extracted from it as a collection of table-formatted standard reports. In tech-speak, we can pull a system’s data from its front-end instead of from the back-end as it was necessary in the 2000s and earlier.

The killer for data-drivenness is that the table-formatted standard reports of a firm’s operating systems are akin to viewing operations through pinholes and trying to surmise the whole upon our experience and intuition. The second layer of the data-driven capability allows us to combine the data of the whole story.

As already mentioned, the notion of preplanning all insight deliverables and subsequently being limited to them is an obsolete one. That is equally so for the second layer. It is not necessary to proactively specify full-view data needs until there a recognized need emerges. This is because the data-driven capability lives on the data that are generated by functioning rather than data that only exist by virtue of preplanning the data needed to be a data-driven firm.

The second layer is the capability to immediately build “super tables” as needed to serve any realized need whenever it reveals itself. Super tables are defined as tables that join subtables from any source and return a table with every data element (variable) needed to build a specified insight deliverable.

Most firms are already positioned to work in the second layer. They just do not know it. In most firms, every individual has rights to the software to build super tables—MS Access—through the firm’s MS Office license.

Although most firms have not yet recognized what Access and other software of its genre allow, the skills to work with Access are already socialized into their work forces by virtue of Excel having long ago become ubiquitous to functioning. The article, “[Building the Super Tables Behind Data-Driven Operations](#),” explains the process of building super tables in the context of Access.

Descriptive statistics also come into the building process through the exploration and cleansing of data. Descriptive statistics may also be a data analytic insight deliverable of the first layer. For both cases, descriptive statistics are built with the software, “R.”

### **Third Layer: Source Tables**

What makes super tables super is that they are built by reaching out to any data located anywhere—the third layer of the data-driven capability. The one or more source tables need not be located in integrated systems, located in the same department and organization, or be of the same discipline. As a point of reference, the “internet of things” (IoT) are merely the devices that send data to one or more source.

All process tasks that are conducted with an operating system concurrently pull data from system tables as task information and enter new data in the conduct of the task. The data are stored in relational databases as massive collections of available source tables. In turn, the data, as standard reports in table format, can be pulled into the second layer.

However, the equivalent discipline is typically not the case for the many process tasks that are conducted with Excel rather than with an operating system. Instead, as a spreadsheet, the data are fused with the insight deliverable. This and the remedy are explained by the article, “[Purge the Fused Task Spreadsheets That Undermine Data-Drivenness](#).”

Another way of saying it is that the data, calculations and presentation are fused in a single spreadsheet not of a table format. When this is the case, the firm’s data driven capability is grossly restricted. So much so that the firm will never be able to become fully data-driven. Consequently, to become data-driven-capable the firm will redesign and, henceforth, conduct its Excel-based tasks so that every task’s data are swept into the third layer of the data-driven capability.

Redesign is straightforward and quick. First, identify the constituent variables to the fused spreadsheet. Then redesign the conduct of the task, as necessary, to capture and store the variables in Excel as a table available to be tapped by the second layer of the data-driven capability. For the task, emulate and disseminate the fused spreadsheet as an Excel Pivot report that is connected to the task data table.

Besides making task data the gold it is, there are big spinoff benefits of the redesign. The often tremendous wastes of preparing each edition of a fused spreadsheets are eliminated. In addition, the wastes or complete impracticability of reformatting the fused spreadsheet for other routine and ad hoc insight deliverables are eliminated. The elimination frees bandwidth for value-added work such that administrators will take on analyst roles that they are especially well suited to take on for the team.

For the user, the emulated task spreadsheet, as an insight deliverable to the first layer, carries much more insight. This is because it is interactive. Every permutation of slice-dice, roll up and drill down is actually an alternative insight deliverable generated at the speed of click and drag.

### **Grassroot Strategy to Become Data-Driven-Capable**

There was a time that only a few process operatives used a computer in their role and managers almost not at all. Now it is hard to imagine any role in which computers are not engrained.

To qualify as a data-driven-capable organization, data and insight roles will take place at most computers throughout the operation. Data-driven tasks will be inherent to roles that generate data, prepare and disseminate insight deliverables, and rely on insight deliverables to analyze and manage functioning. That is just about everyone; including at least a portion of frontline workers.

This suggests that becoming data-driven-capable can be a slow, long initiative, followed by a very long evolution before becoming fully data-driven. The counter is to take a grassroots strategy.

History is culture. A grassroots strategy builds a firm's data-driven capability upon its present culture of systems, software and skills—because we can. The alternative is to grow a new culture around best-in-class software.

A firm's cultural amenability to becoming data-driven has evolved upon a single piece of history. For many years, almost no one has escaped Excel in their work. The implications for becoming data-driven-capable is evidenced by the fact that Excel software, functionality and skills were recognized as the grassroots to each layer.

Pivots, as intelligence deliverables in the first layer are Excel. MS Access in the second layer is essentially a small step out from Excel; doing the one thing Excel cannot, but with the same skills and instincts of experience. Abandoning forever the fused spreadsheets, for the sake of the third layer, is merely redesigning Excel-based tasks already conducted in Excel.

Arguably, data analytics are the only cultural step out from the Excel-based culture. However, the instincts for working with functions, arguments and operators gained from Excel carry over to R.

Furthermore, R is the grassroots option for answering the [five operational questions](#) of data analytics. This is because commercial software offerings for data-driven functioning are interfacing with R as their software's engine for data analytics beyond the shallow degree they are otherwise limited to.

The alternative to the grassroots strategy is to jump into the several software offerings that are striving to be the best-in-class to the three layers. The purchased gain is that they somewhat improve upon the efficiency of the second layer and beautify the first. An organization must decide whether the difference is significant vis-à-vis the business losses of water under the bridge during the greatly extended time before achieving a data-driven capability. Probably not.

The grassroots strategy does embrace the adoption of best-in-class, but only after building upon the grassroots culture to become data-driven-capable. Thence, the evolution to being fully and routinely data-driven should be the platform for deciding which step-outs to take from the grassroots capability. In those cases, and with the grassroots capability in place and functioning, new adoptions can be placed where they are surgically worthwhile. New culture will evolve around them and spread across the grassroots culture.

Furthermore, cultural improvement will spread faster and easier. This is because a characteristic of the grassroots culture will be to become comfortable in the three layered functioning of data-drivenness. Any new software is merely another way to do the same thing.

**Related website articles:** [Building the Super Tables Behind Data-Driven Operations](#) | [Purge the Fused Spreadsheets That undermine Data-Drivenness](#) | [Keep Your Old Career Hot in the New Age of Data Science](#)

**Richard G. Lamb:** [Professional Mission and Bio](#)

**Educational website:** [analytics4strategy.com](http://analytics4strategy.com)