LEVERAGING DATA AND ANALYTICS TO ENABLE STRATEGIC IMPROVEMENT

Jeff Stankiewicz, MS

Ken Feldman, PhD

Proficiency SYSTEMS, INC

Table of Contents

Executive Summary

We are living in an age of explosive data growth, with no end in sight. And the organizations who are best utilizing analytics to harness their data – extracting critical insights and making better decisions – are realizing tremendous financial returns and in many cases driving fundamental industry disruption. For the median Fortune 1000 company, the opportunity is easily in the billions of dollars.

While the promise of data and analytics is huge and there are success stories, especially in the digitally native space (e.g. Facebook, Uber, VRBO, Netflix), studies estimate that at the organizational level only 10-40% of the potential value from analytics is currently being captured. At the initiative level, big data projects fail at a high rate, with estimates of up to 85% failing in their preliminary stages.

Whether you're just beginning to get serious about data and analytics or if you have been at it for a long-time; whether your company's data is of high volume, velocity, and variety (big data) or if you have 'smaller' data needs, here are recommendations on how to generate significant and lasting value:

- **Any data and analytics effort should be strategy driven**. We consider data and analytics, simply, as a set of tools that will only yield significant value when properly applied to provide strategy planning input or when utilized to support the fulfillment of strategic objectives.
- While much is changing in today's world, some things remain the same:
  - Businesses are still made up of the same components: product innovation, sales, marketing, value delivery, etc.
  - Processes continue to be the engines that deliver value to customers. So, the focus should always be on ensuring the right processes are in place (traditional or digitized) and that they are performing at the highest levels.
  - To succeed, businesses need to be effectively managed, which requires a high-functioning management system. The most effective management systems utilize a two-dimensional approach; one dimension to drive the delivery of strategic objectives and the other to maintain and continually improve day-to-day operations. We term this an 'integrated business management system'. **Data and analytics enable your business management system to achieve significantly higher levels of performance**.

"Analytics is not a technology issue, it's a strategy and operational issue." – Chris Mazzei, Global Chief Analytics Officer, Ernst & Young

Any investment in or application of data and analytics should be directly driven by the need for higher performing and better managed business processes, which is also the domain of the company's integrated business management system – the organization's tool for driving strategic improvement.

- **Data and analytics should not reside primarily within the domain of IT or an analytics function** (oftentimes reporting into IT). Instead, they should be better integrated into existing initiatives and core business functions.
- To help insure data and analytics initiatives are aligned to business needs and that they deliver promised results, they **should be executed within a standard framework**, such as the comprehensive **7 Step Analytics Cycle** we have developed.
- **Not all organizations require expensive and hard to find data scientists** to effectively utilize data and analytics. Depending on the nature of your data and the types of analytics needed, business users equipped with **self-service** business intelligence capabilities or analytics **power users** developed from within your existing workforce can provide the needed horsepower.
- Most importantly, to realize the greatest value, **data and analytics need to be ingrained into the way your company does business** – the way people behave – day-in-day-out, across all levels and functions. They **need to become part of your culture**.

  Along with rigorously applying the full-set of proven change management elements (e.g. highly engaged top executive support, ongoing risk management, active performance management, etc.), cultural inculcation also requires several additional high-level systems: A **Data and Analytics Steering Team**, a **Data and Analytics Center of Excellence (DACE)**, and an **Analytics Community of Practice.**

At the foundational level, better leveraging data and analytics makes your business smarter. And smarter has always been a formula for success and competitive advantage. In many industries, those organizations who lead the way in utilizing data and analytics as a strategic weapon will thrive, while those who don't may not even survive. Achieving maximum benefits requires ingraining data and analytics deeply into your organization's culture. This is a fundamental transformation that will require hard-work, a steady focus, and taking the right steps. While there's not a single recipe for every company, this paper provides key ingredients that will help you achieve more analytics success with less effort.

Data and Analytics – The Value Potential and Challenges Faced

"Welcome to the information age. Data, data, everywhere, but no one knows a thing."
– Roger Kimball

Across the last several years we have been seeing an increasing tidal wave of business publications and conference presentations touting a dizzying array of newly emerging technologies and concepts related to creating, capturing, and analyzing an exponentially exploding amount of data; that when effectively harnessed can create a tremendous amount of business value.

The interest in terms like data science, data analytics, the internet of things (IoT).

machine learning, and business intelligence has increased tremendously as demonstrated in the Google Trends output shown in Figure 1.

And of course, we would be remiss not to include arguably the most hyped buzzword in recent years, big data, for which based on Figure 2, the hype cycle appears to have leveled off. A potential reason for this will be discussed later.

Figure 1: Google Trends Output of Select Data and Analytics Terms (Jan 1, 2004 – July 8, 2018, United States)

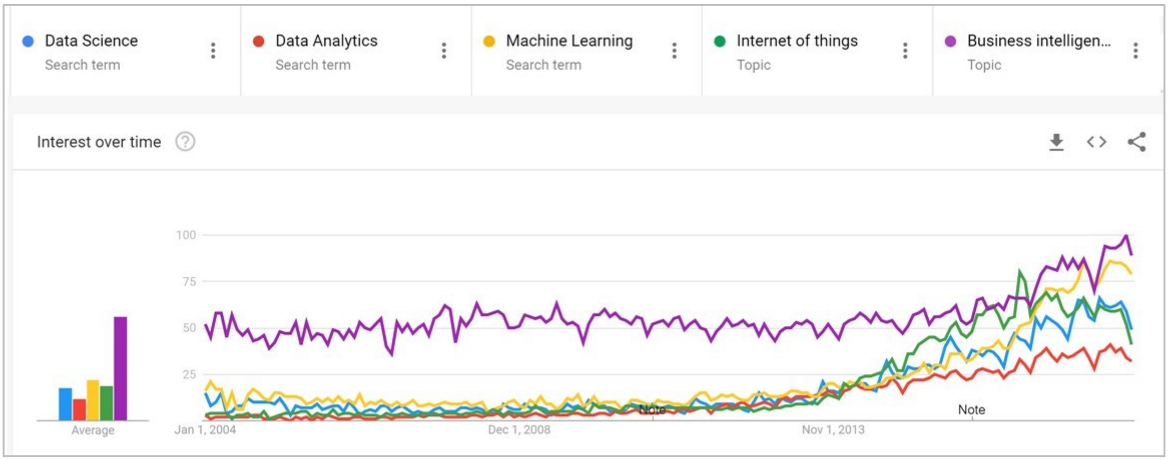
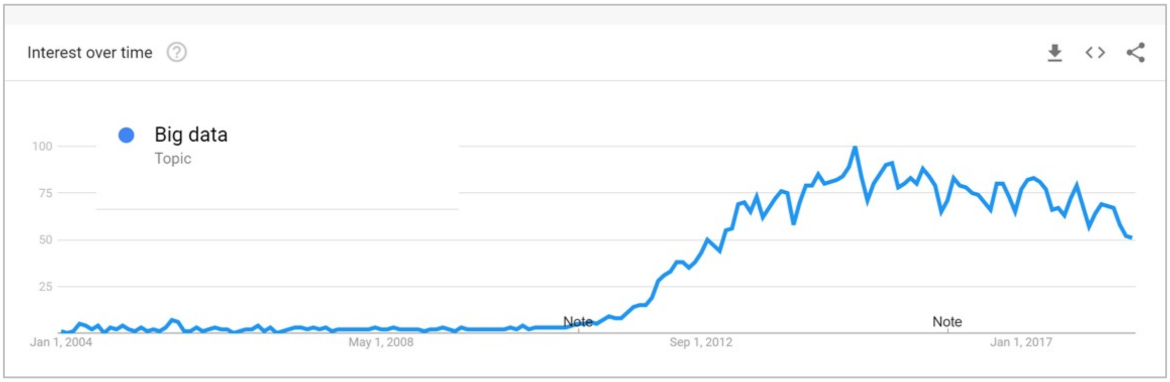


Figure 2: Google Trends Output for Big Data (Jan 1, 2004 – July 8, 2018, United States)



The bulk of the content to follow is about **how-to best leverage data and analytics to improve the performance of your business**, be it through increased customer intimacy, better product innovation, or improved operational efficiencies. But before going there, it's worthwhile to first spend a little time painting a picture of our view on the current state of data and analytics in the business world, as summarized in the following four points:

1. **The amount of potentially useful data generated is growing at a mindboggling rate**

   Here are just a few of the many eye-opening quotes we have come across on the ongoing unprecedented growth rate of data:

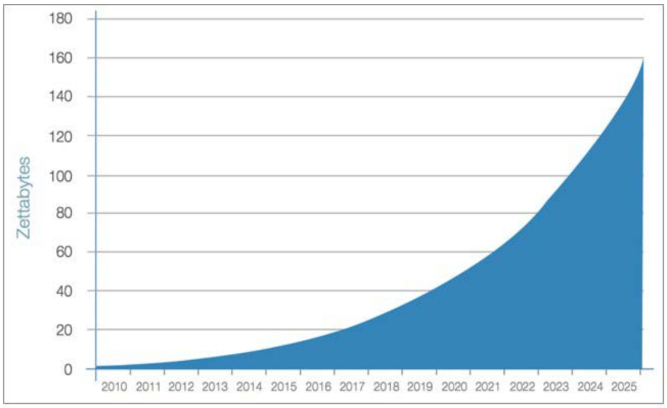
   - "2x growth in the volume of data expected every 3 years" – McKinsey & Company
   - "By 2020, about 1.7 megabytes of new information will be created every second for every human being on the planet" – MIT Technology Review
   - "Every two days now we create as much information as we did from the dawn of civilization up until 2003" – Google CEO, Eric Schmidt, 2010

   Out of simple curiosity, the last bullet led us to ask: "If that's the rate of information we were creating in 2010 what is the rate in mid-2018, after 8.5 years of additional exponential growth?"

   To attempt to answer this question, we used the exponential data growth curve published in a 2017 IDC Corp. white paper. as part of research sponsored by the data storage company Seagate Technology, LLC. (Figure 3). We used a free internet algorithm to first generate the estimated cumulative probability function for the curve from several selected points on the graph and then used another online algorithm to solve for the area under the curve between different years.

   Surprisingly, the results showed that it's now taking 49 days to create as much data as it did from the dawn of civilization up until 2003. Wait a minute, does this mean the rate has slowed since Eric Schmidt made his comment back in 2010? The answer is a resounding 'no' as can be seen by the continued exponential growth visible within Figure 3. Instead the reason for the difference is likely the result of different assumptions and data sources used between the Google and IDC Corp. prediction models. For instance, Mr. Schmidt used the term 'information' instead of 'data', where information is most commonly defined as an aggregation of data. Also, we don't know exactly what types of information and data (numeric, text, graphical, video, audio, etc.) were included in each.

Figure 3: Annual Size of the Global Datasphere



In summary, while the values often vary between sources (and some are sure to be over-inflated to serve the interests of those making the statements), they all support the same conclusion: We are living in an age of explosive data growth, with no end in sight.

Not only is the ‘volume’ of data growing at a staggering rate, it is of increasing ‘variety’, and is being generated at an ever increasing ‘velocity’. – The three ‘V’s commonly referenced as making up the nature of big data.

One of the key drivers of data growth is the corresponding proliferation of networked internet connected devices (internet of things (IoT)). Examples include smart home devices, wearables like FitBit®, GPS tracking units, performance sensors in equipment, stress sensors in bridges, nutrient and moisture sensors in soil, and many others. IBM projects there to be “75 billion internet-connected devices by 2020”.

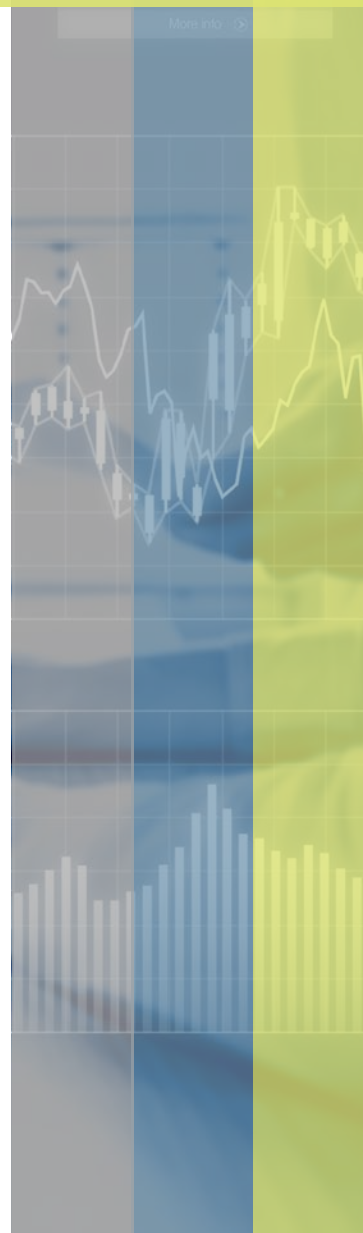
So how much of this huge amount of data is actually being analyzed and leveraged? The answer is only a very small amount. In a 2013 article, MIT Technology Review Senior Editor Antonio Regalado stated that only about 0.5% of data is ever analyzed. With the additional proliferation of natural language, audio, and video data since 2013, we would be surprised if today’s number wasn’t even lower.

We are also experiencing a similar growth pattern in the technological capability to process and analyze data. In December 2016, McKinsey & Company stated that there has been a “40x increase in processing power between the fastest supercomputer in 2010 and the fastest today”.

Additionally, in an attempt to keep up and leverage all of this data and processing power, more and more sophisticated analytical methods continue to evolve with Machine Learning at the forefront. SAS defines Machine Learning as “a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.”

In practice, we see machine learning being increasingly applied at two different levels. The first is to perform predictive analytics on big data to identify subtle patterns

We are living in an age of explosive data growth, with no end in sight



that often go undetected by humans. The resulting insights are then used as a supplement (not a replacement) to human decision making. Examples include:

- Health Care: Imaging analysis to detect more subtle patterns that are often not detected by radiologists, disease identification/diagnosis, and epidemic outbreak prediction
- Manufacturing: Predictive maintenance to improve equipment performance, automated root-cause analysis to reduce scrap rates, inventory optimization
- Public Sector: Predictive policing to identify where crime is most likely to occur and where suspects are most likely to be

The second level of machine learning application is to perform prescriptive analysis resulting in models that automatically make decisions, eliminating the need for human involvement. This allows for processes to be fully digitized.

Digitally native companies like Uber, Facebook, Spotify, Netflix, and Airbnb have already disrupted or created entire industries. Traditional brick-and-mortar companies are also rapidly embracing digitization in part by applying it to specific processes and products. For example:

- Fast food restaurants are installing kiosks allowing customers to quickly and easily customize their burgers and other sandwiches
- Automobile manufacturers are incorporating technology within almost every facet of automobiles, basically making them computers on wheels. This will ultimately lead to driverless cars becoming the standard
- Banks have been early adopters of digitization, applying it to processes like fraud detection
- Theme parks are employing digitization to personalize the customer experience. From a website, guests can book ride times, restaurant reservations, and other activities in advance. When they arrive at the park, they receive reminder alerts from a smartphone app, and can use the same app to make real-time reservation changes

Of course, each of the above examples also generates large volumes of additional useful data that can be used to further improve products, operational efficiencies, and customer experiences.

The disruptive potential of digitization has caught the attention of business leaders. Gartner, a leading IT global research and advisory firm, in its 2018 CEO survey, found that 13.4% of surveyed CEOs consider digitization a top-five priority item for their organization. This is more than a 6x increase since 2012.

2. **There's tremendous value to be captured by better harnessing the data**

To begin to get a better feel for the value potential of better leveraging the tidal wave of data and related technology discussed above, a good place to start is to take a look at the S&P 500 and five of its companies that are some of the best at leveraging data and technology: Facebook, Apple, Amazon, Netflix and Google. As of early 2018, these five companies combined had a total market value of almost $3 trillion and accounted for more than 11% of the S&P 500 index, which is about twice as much as they represented in 2013 (Source).

Since we're throwing huge numbers around, let's look at one more dimension: automating labor. A December 2016 report, by McKinsey & Company estimates that 34% of all work activities, linked to $14.6 trillion of global wages, have the potential to be automated using today's machine learning capabilities. While the potential for improved operational efficiencies is enormous, we must admit that the likely resulting disruptive forces on the labor market are also a bit unnerving.

So, what's the value potential at the individual company level? The same McKinsey & Company report, states that "investing in data and analytics capabilities has high returns, on average: firms can use these capabilities to achieve productivity gains of 6 to 8 percent, which translates into returns roughly doubling their investment within a decade. This is a higher rate of return than other recent technologies have yielded, surpassing even the computer investment cycle in the 1980s."

In the same report, McKinsey & Company lists the potential impact of better leveraging data and analytics from research it had done in 2011. Here are a few of the potential organization level impacts mentioned:

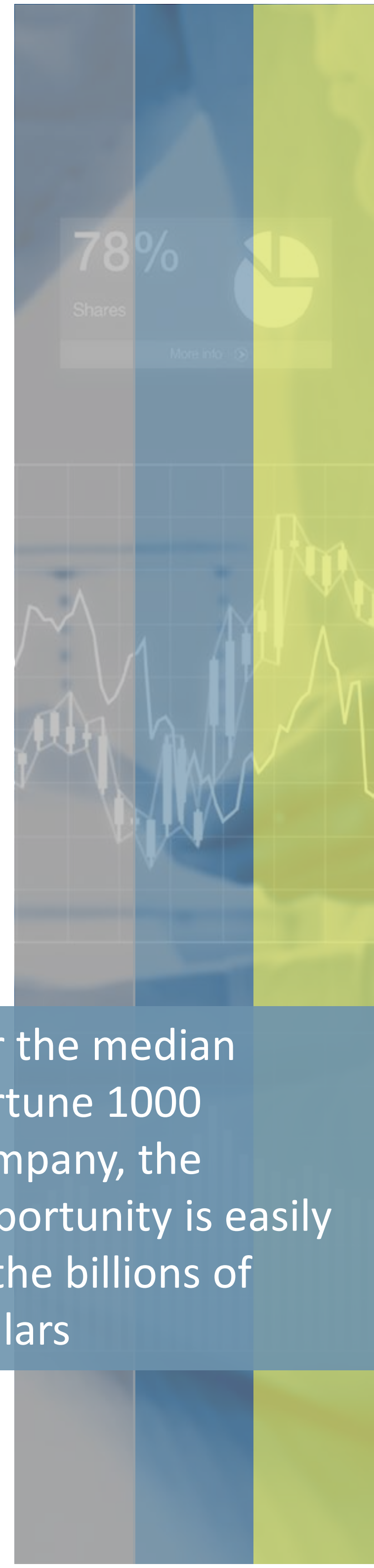
- In US retail, more than a 60% increase in net margin and 0.5-1.0% annual productivity growth
- In manufacturing, up to 50% lower product development cost, 25% lower operating cost, and 30% gross margin increase

At an industry level, they state there is $300 billion value per year in US health care.

An additional 2010 University of Texas study of over 150 Fortune 1000 firms from every major industry states the following:

- "If the median Fortune 1000 business…increased the usability of its data by just 10%, it would translate to an increase in $2.01 billion in total revenue every year."

While the study's company level information is seven or eight years old, we are confident that if anything, the continued growth in data and technology and the learnings from early adopters, has only increased the potential value of data and analytics.



For the median Fortune 1000 company, the opportunity is easily in the billions of dollars

3. **Business leaders recognize the opportunity and are investing heavily**

   Leaders around the world now recognize the opportunities and marketplace disruption that leveraging data and analytics can have on their organizations and are taking action. According to KPMG's 2017 Global CEO Outlook study:

   - "48 percent of CEOs expect major disruption in their sector from technological innovation within the next 3 years"
   - "38 percent are planning to make significant investments in D&A tools over the next 3 years"
   - "35 percent will put significant investment into their digital infrastructures"
   - "70 percent expect these investments to improve bottom-line growth"

   In its 2017 Worldwide Semiannual Big Data and Analytics Spending Guide, International Data Corporation (IDC) projected the following: "worldwide revenues for big data and business analytics (BDA) will reach $150.8 billion in 2017, an increase of 12.4% over 2016. Commercial purchases of BDA-related hardware, software, and services are expected to maintain a compound annual growth rate (CAGR) of 11.9% through 2020 when revenues will be more than $210 billion." This doesn't include the additional funds being spent to build and support the underlying IT infrastructure (Gartner forecasts a $3.7 trillion total IT spend in 2018), along with monies allocated to skills development and salaries of involved individuals.

Leaders around the world now recognize the opportunities and marketplace disruption that leveraging data and analytics can have on their organizations and are taking action

4. **Capturing the potential value of data and analytics is proving to be elusive**

   The opportunity provided by harnessing data and analytics, the significant attention it is getting from business leaders, and the huge sums of money being spent, all beg one question: "What kind of results are companies actually realizing?"

   The answer in short: They are seeing some value, but not as much as they have hoped for.

   The previously referenced December 2016 report, by McKinsey & Company explores how much of the potential identified in its 2011 research had actually been realized. It found that the US retail sector had only realized 30-40% of the potential value, with manufacturing, the EU public sector, and US health care realizing even less – in the 10%-30% range. The most common barriers stated were a lack of analytical talent, siloed data and legacy IT systems, leadership skepticism, and the need to demonstrate results to gain acceptance. The report summarizes the challenges with the following statement: “The biggest barriers companies face in extracting value from data and analytics are organizational; many struggle to incorporate data-driven insights into day-to-day business processes.”

> “The biggest barriers companies face in extracting value from data and analytics are organizational; many struggle to incorporate data-driven insights into day-to-day business processes.” – McKinsey & Company

NewVantage Partners LLC (NVP) conducted a Big Data Executive Survey in 2017, with respondents from 50 Fortune 1000 or industry leading firms. More than 75% of respondents were from the financial services sector, so it’s unknown exactly how representative the results are across industries, but financial services has been one of the earliest big data adopters, so we can likely learn from their experience.

More than 80% of the NVP survey respondents said their big data (analogous to data and analytics) investments had been at least moderately successful. With that said, one caution in interpreting the results, is that since the vast majority of the respondents likely had a large stake in their company’s big data investment, it’s safe to assume the reported success rates were inflated somewhat by confirmation bias. But, in general it appears a fair amount of success is being realized.

Diving a little further into the survey results, respondents reported that more than half (ranging from 54% -to- 69%) of the big data initiatives have realized at least some degree of measurable results as indicated by:

- Creating new avenues for innovation and disruption
- Decreasing expenses through operational cost efficiencies
- Monetizing big data through increased revenues and new revenue sources
- Launching new product and service offerings
- Transforming and repositioning the business for the future

However, while more than 85% of respondents reported that their firms had started larger programs to create data-driven cultures, only 37% reported success thus far. It appears to us, that respondents have typically had some success with a few specific projects or initiatives, however most have struggled to embed data and analytics into their organizations' cultures – which is the only way to realize its full potential value.

"Organization Issues | Resistance" was cited by 52.5% of respondents as the biggest impediment to big data success (lack of a coherent data strategy was the second most common barrier – 18.0%). When probed further, respondents listed the following as the top three impediments to adoption:

- "Insufficient organizational alignment"
- "Lack of middle management adoption and understanding"
- "Business resistance or lack of understanding"

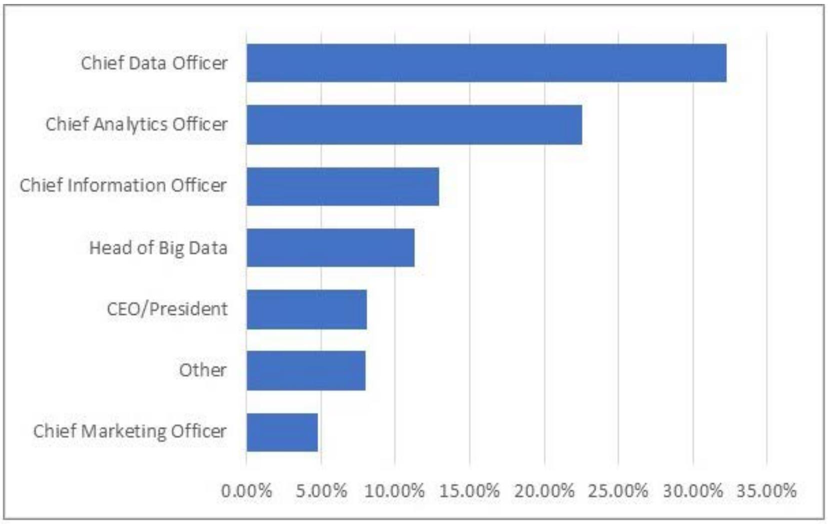
If, like us, you have been around long enough to have lived through numerous large scale corporate initiatives, the barriers and impediments listed in the last two studies will likely sound familiar. They are some of the most common barriers of any large scale cultural change initiative.

Continuing with a little more analysis of the NVP study, we were curious as to who responded, and particularly where they stood in the organizational structure. Figure 4 summarizes what organizational levels were represented in the survey.

Whether we are correct in our conclusion or not, it appears that the ownership for deploying and driving big data in an organization is essentially an IT role rather than the CEO or President. Are their examples where IT has driven cultural change in an organization? We don't believe there are. The necessary changes to create a data and analytics culture will not occur if big data resides in the technical world of IT and analytics.

While not mentioned in any of the research we came across, a common theme we observed is that data and analytics initiatives are typically positioned as stand-alone, not linked to other ongoing corporate initiatives. Multiple disparate initiatives can result in competition for scarce resources and increase management complexity and confusion.

Figure 4: Respondent Percentages: 2017 NVP Big Data Executive Survey



Moving down another level from company initiatives to specific projects, our research indicates that success has been even more elusive. In early 2017, Gartner stated that 60% of big data projects will fail to make it past preliminary stages. Later in the same year, Gartner analyst Nick Heudecker said “We were too conservative. The failure rate is closer to 85%. And the problem isn’t technology”. This high failure rate may be why the Google Trends information in Figure 2 shows interest in the term ‘big data’ has leveled off if not on a downward trajectory.

So, what are the most common reasons given for project failure? IT and technical issues like data silos, data security concerns, poor data quality, and inadequate analytics skills commonly come up. But even more commonly it’s the same list of offenders you would get if you were to ask “why do agile initiatives fail?” or “why does any large project fail?”. These include:

- Launching an initiative without a clear business driven need linked to strategy
- Not having a strong champion for the initiative
- Inadequate cross-functional representation on the team
- Inadequate planning
- Poor communication
- No or inadequate risk management

In summary, while companies are seeing some results, our view is that much more of the promised value of data and analytics will be realized if organizations:

1. Better integrate data and analytics into existing initiatives and core business functions rather than approaching it as a separate initiative and one that resides primarily within the domain of IT or an analytics function (oftentimes reporting into IT)
2. Address the foundational factors that make for any successful large scale cultural change or initiative

The remainder of this paper will provide our perspective on how to best accomplish the two items above and ultimately how to close the gap between the mediocre results most companies are achieving and the tremendous value that data and analytics can actually deliver.

The necessary changes to create a data and analytics culture will not occur if big data resides in the technical world of IT and analytics

The New World of Data and Analytics – What Changes and What Stays the Same

"Continuity gives us roots; change gives us branches, letting us stretch and grow and reach new heights." – Pauline R. Kezer

Data and analytics and new technologies like machine learning and IoT are clearly disruptive forces that when properly applied can dramatically increase the growth of your business and its operational efficiencies, but what are the key business changes you will need to effectively navigate to fully capitalize on its potential? We have come up with a list of three key changes to expect:

1. **IT will play an increasingly key role**

   The huge volume of data being generated within and flowing into the business need to be secure, properly governed, and of high quality. When you consider the many different sources of data, it's easy to understand the need for an increasing amount of infrastructure and support to manage it all. Table 1 provides a partial list of data sources most businesses now manage.

   All the data from these sources needs to be organized and made accessible and useable, typically requiring data warehouse infrastructure. These activities are performed and supported by individuals with ETL (Extract, Transform, Load) and database skills.

Larger organizations have created a data engineer role to perform these activities. These individuals are typically software engineers by trade and are tasked with building a robust, fault-tolerant data pipeline that cleans, transforms, and aggregates unorganized and messy data into databases or data sources.

To begin to pull value from the data it needs to be converted into information and knowledge in the form of reports and basic analytics. This requires individuals with the corresponding analytics and report development skills. It should be noted that this set of activities lies in the border region between the other business functions and IT. A bottleneck can quickly be created if all reports and basic analytic output are strictly IT generated. For this reason, there has been an increasing 'self-serve' movement toward providing more-and-more individuals the tools and skills to generate their own reports and to perform simple analyses. This topic will be discussed more later.

Table 1: Common Data Sources

Billing Systems	CRM Systems	Customer Data	ERP Systems
Debt Collection Systems	ERP Systems	Campaign History Data	Production Data
Social Media Data	Product Data	Web Logs	Internet of Things Data
Geo Data	Consumption Data	Human Resources Info.	KPI Data

2. **New tools requiring new skills and potentially new roles**

In 2012 Harvard Business Review labeled it "The Sexiest Job of the 21st Century" and six years later, in 2018, Bloomberg called it "America's Hottest Job", citing a 75% increase in related job postings on Indeed.com since 2015. Yes, we're talking about the data scientist – the math nerd and data geek who has been transformed into the modern-day rock star of business. And, with some top talent "commanding as much as $300,000 or more from consulting firms" (Bloomberg), some are getting compensated like rock stars.

While there's not consistent agreement on the skills required to consider oneself a data scientists, those most commonly referenced are shown in Table 2 (Source).

We noticed one skill that is notably absent from Table 2: business acumen. While most individuals with advanced degrees in a technical field (e.g. math, statistics, and engineering) have the intellectual horsepower to learn and effectively execute the above skills, it is our position that it is equally important for a data scientist to also have an intuitive understanding of the business problem they're trying to solve. Otherwise they may do a great job at developing a sophisticated model that doesn't add any tangible value to the business. This lack of business acumen in many data scientists is likely one of the causes of the high rate of failed initiatives we discussed previously.

The majority of newly minted data scientists coming out of academic institutions are likely to possess adequate levels of competence in the skills listed in Table 2, however finding an experienced individual that also has business smarts is a much more difficult challenge.

So, should your company jump into the hyped-up feeding frenzy and compete to hire a highly qualified data scientist? The answer is probably yes, if you're in a large

Table 2: Data Scientist – Common Skills

1	Coding Skills	In a statistical programming language (e.g. R or Python) and a database querying language (e.g. SQL)
2	Math, Especially Multivariable Calculus and Linear Algebra	To develop custom in-house predictive and optimization algorithms
3	Machine Learning	To identify subtle patterns in data and develop Prescriptive models, particularly in large organizations with huge amounts of data
4	Statistical Methods	Including descriptive and inferential methods, hypothesis testing, regression modeling, distributions, etc.
5	Data Wrangling	Including the skills to effectively manage imperfections in data (to clean-up dirty data) and pull together data from disparate sources
6	Data Visualization and Communication	To be able to create effective visualizations and dashboards to support others in making data-driven decisions and to be able to share and describe findings to both technical and non-technical individuals within the business
7	Data Intuition	To quickly be able to identify what's important and what's not important and to be able to rapidly identify the past path forward when tasked with solving a new data related problem

highly digitized organization with huge volumes of data. But if that doesn't describe your company, we recommend against it. For us to explain why we take this position, we need to link the discussion to our four-levels of analytics shown in Figure 5.

Prescriptive analytics is the domain of big data – huge volumes of highly varied data generated at high velocities. If your company doesn't have this type of big data to process, like many smaller and traditional (non-digitally native) companies, key skills data scientists bring to the table will often go unutilized. Machine learning techniques require a minimum of tens of thousands and sometimes millions of data points to build. And as a recent Entrepreneur article states: "Job descriptions for data scientists are flooded with terms like neural networks, machine vision and natural language processing (NLP). The issue? These types of techniques rely on having massive amounts of training data. Consider the widely popular Google Translate, a type of neural network built on top of a lexicon of over 150 million words. The volume of data needed for successful deployment of these types of models exceeds what many companies own."

Also, many companies are just getting started down the data and analytics path and have bigger immediate needs in the descriptive analytics and diagnostic analytics domains, which together make up what we call business intelligence (BI); along with an occasional predictive analytics effort. For instance, PWC in its Chief Financial Officers – Priorities in 2018 report, found that performance management is the number one priority of CFOs in 2018 and remains at the number one position for the next three years. As part of this priority 89% of CFOs surveyed said they "want to improve their reporting and dashboards using more data visualization and by improving their production methods." You don't need a fully qualified data scientist to do this type of descriptive analytics work.

Figure 5: The Four-Levels of Analytics

Analytics Type	Questions Answered	Tool Families	Outputs		
Prescriptive Analytics	• What is the best course of action (or decision)? (especially in a complex environment)	• Optimization (Linear, Nonlinear, Integer, Network, etc.) • Stochastic Optimization • Queueing Models • Natural Language Processing • Machine Learning	• The optimal decision(s) made automatically (digitization) or to supplement human decision making	Future	'Advanced Analytics'
Predictive Analytics	• What will happen if?	• Simulation (Discrete Event, Monte Carlo, Agent-Based Modeling) • Artificial Intelligence (Artificial Neural Networks, Fuzzy Logic, Expert Systems) • Regression (Logistic, Linear) • Game Theory	• Input to assist decision making		
Diagnostic Analytics	• Why? • What are the causes?	• Statistical Inferences (Confidence Intervals, Hypothesis Testing, Design of Experiments) • Dimension Reduction (Principle Component Analysis, Clustering)	• The critical few factors (variables) most impacting the course or system performance	Rearview	'Business Intelligence'
Descriptive Analytics	• What is happening? • Are we staying on course? • Are all systems operating as intended?	• KPIs, metrics • Dashboards • Automated monitoring and alerting	• Are minor (incremental) course or system corrections needed? • Do special causes need to be addressed? • Are larger, more systemic, course changes or system corrections needed? • Do we simply need to continue monitoring?		

Likewise, when performance issues are observed, the obvious question is usually ‘why?’. This moves us up into the next domain of diagnostic analytics which requires drilling down into the data and often using traditional statistical analysis methods to identify causal factors. Once again others, besides data scientists can be equipped to do this type of work.

So, if these are the primary needs of your business, what are some alternatives to a hiring a high-powered and expensive data scientist? Here are several recommendations:

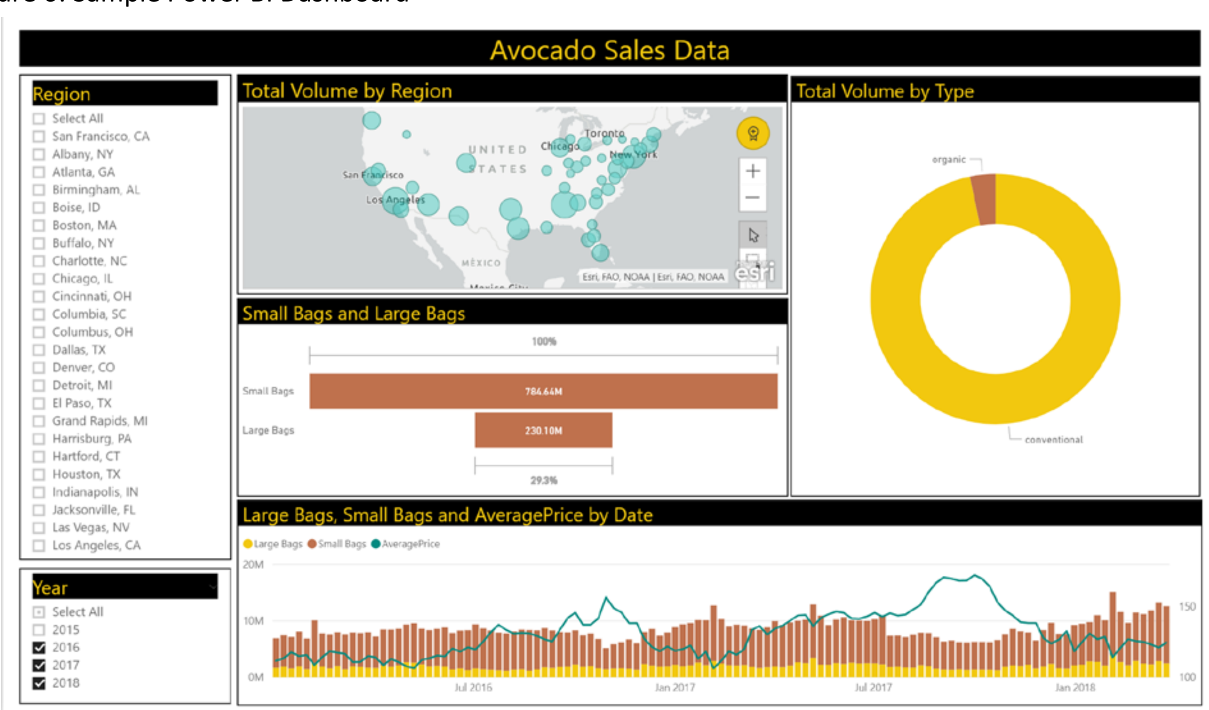
- Equip business users, without previous data and analytics experience, with the skills to access, visualize, and analyze data through **Self-Service Business Intelligence (BI)**.

  There are multiple software tools on the market today (two of the most popular are Tableau and the newer Microsoft Power BI), that are relatively intuitive to use; and with a little practice, can allow users to easily see, understand, and be able to drill-down into their data to gain increased insights and support data-driven decision making. They are flexible enough to easily access most databases and combine information from multiple databases yielding unique insights. Figure 6 provides an example of a basic Power BI generated dashboard.

Businesses are embracing self-service analytics, with Gartner predicting “that by 2019, the analytics output of business users with self-service capabilities will surpass that of professional data scientists.”

However, like with anything, there’s much more to deploying self-service BI than simply providing users with access to the tool. To be successful, an effective supporting infrastructure needs to be established. This topic will be discussed in more detail in the later ‘Ingraining Data and Analytics into Your Organization’s Culture’ section.

Figure 6: Sample Power BI Dashboard



- **Develop a stable of Analytics Power Users** within the organization, who can lead descriptive, diagnostic, and many predictive analytics efforts, as well as train and coach others on descriptive and diagnostic analytics. As a prerequisite, these individuals should already have in-depth business knowledge, process improvement experience, and a base level of data and analytics skills. Target individuals should be active users of intermediate Microsoft Excel functionality (e.g. pivot tables and VLOOKUP functions), along with working experience in applying descriptive and inferential statistics to support problem solving.

  Individuals previously trained and certified as lean six sigma master black belts, black belts, or green belts (if trained in a technical environment) possess the prerequisite skills and experience and make ideal power user candidates. Even if your organization hasn't formally deployed lean six sigma, there is a very high likelihood that you employ some individuals with black belt or green belt experience gained in previous jobs. The numbers are hard to quantify exactly, but the quantity of individuals world-wide with black belt or green belt experience is likely in the millions. One effort from several years ago came up with an estimate of 1.5 million individuals working as full-time black belts (300,000 to 500,000 of the total) or part-time green belts (making up the remainder). Source

  Ideas on how to utilize these individuals within the analytics space and how to enhance their data and analytics capabilities will be discussed in more detail in the later 'Ingraining data and analytics into Your Organization's Culture' section.

- **Outsource the occasional prescriptive analytics (big data) project**. If your company is not at the point of being ready to hire full-fledged data scientists, but you have the occasional prescriptive analytics need, there is always the option of contracting a third-party to lead the project. Along with many freelance data scientists, there are hundreds of consulting organizations that offer data science services to assist with the occasional complex analytics need.

3. **An ever-increasing number of decisions will be made without human involvement**

   As discussed earlier, more and more organizations are making process digitization a priority. One indicator of just how much change we should expect due to process digitization (enabled by decision automation) is provided in a 2016 US Bureau of Labor Statistics; McKinsey Global Institute analysis. The analysis projects that by 2030, up to 375 million workers (14% of the workforce) will need to move out of their current occupation and find new work due to rapid automation adoption.

   At process digitization's core, is the application of machine learning to perform prescriptive analytics resulting in models that automatically make decisions, eliminating the need for human involvement.

We talked about what changes to expect with the ongoing advances in technology and Data and Analytics, but what fundamentals stay the same?

First, **all businesses are still made up of the same parts**. Josh Kaufman, the author of The Personal MBA, Master the Art of Business states, that "a business is a repeatable process that":

- "Creates and delivers something of value..." (Product Innovation)
- "That other people want or need..." (Marketing)
- "At a price they're willing to pay..." (Sales)
- "In a way that satisfies the customer's needs and expectations..." (Value Delivery)
- "So that the business brings in sufficient profit to make it worthwhile for the owners to continue operation" (Finance).

Regardless of the level of technology or analytics employed, all businesses are still comprised of these parts. When applied properly, technology and analytics simply enable each part (and the whole) to achieve much higher levels of performance.

Second, to perform at a high-level, **businesses need to be effectively managed**. For a business to be successful in the near-term and to ensure it thrives over the long-term, it needs to operate effectively and efficiently day-in-and-day-out and continually improve to stay ahead of the competition. These things don't happen on their own, they require ongoing proactive management. The most successful organizations overlay and embed a disciplined management system across all business processes, providing the management philosophy and practices employed to run the business. In short, when ingrained into the organization's culture, the management system becomes 'the way' the business works. The most notable example of this is the Toyota Production System (TPS), a key precursor to lean management.

The most effective Management Systems utilize a two-dimensioned approach, which we term an 'Integrated Business Management System'. An overview of this system is shown in Figure 7.

Figure 7: Integrated Business Management System

INTEGRATED BUSINESS MANAGEMENT SYSTEM			
Component	**Strategy Deployment**		**Daily Process Management**
Focus	Improvements **On** the Business		Improvements **In** the Business
Objectives	Achieve Key Business Imperatives		Manage, Monitor, and Improve Daily Operations
Org. Levels Most Involved	Senior Leaders and Middle Management		Supervisors and Process Performers
Performance Measures	Strategic KPIs		SQDPG KPIs
Improvement Methods	Strategic Projects	Process Improvement Projects	Daily Improvements
Enabled By	People, Technology, Data & Analytics		
Delivering	High Performing Business Processes		

Strategy deployment makes up the first dimension of the system. It enables the achievement of step-function improvements leading to competitive advantage. Strategy deployment is typically performed on an annual cycle and starts with the definition of a critical few measurable strategic objectives that the organization needs to achieve. These objectives commonly seek to achieve a breakthrough level of improvement in one (or a few) of the parts of the business mentioned above. For instance, a strategic objective may be to increase growth from X% to Y% by fundamentally improving product innovation or increasing customer loyalty from A to B through a revolutionary marketing approach, or to increase operational efficiencies from A to B of value delivery processes. Sometimes strategic objectives can focus on achieving step-function improvements in areas impacting the entire organization. For instance, increasing company-wide data and analytics adoption from A to B.

Strategic objectives are achieved through the effective execution of strategic projects (e.g. a project to digitize certain processes) or through large-scale process improvement projects (e.g. a project to significantly decrease the time-to-market for new products). Oftentimes achievement of a specific strategic objective requires multiple projects of both types.

To ensure adequate progress is being made toward each strategic objective, and to be able to identify and work to resolve issues in a timely manner, strategic key performance indicators (KPIs) should be defined for each objective and their performance reviewed on, preferably, a weekly basis. These KPIs can measure the performance against planned project schedules or current process performance levels against the new targeted performance level.

Strategy deployment focuses on achieving a few critical business imperatives, impacting only a portion or one aspect of the organization, but what about all the other processes and people not directly involved in the pursuit of a strategic objective (most of the organization)? Appropriate attention also needs to be directed at ensuring all businesses processes are well managed, monitored, and continually improved on a day-in-day-out basis. This is the role of daily process management, the second dimension of the integrated business management system.

The active management of a balanced set of metrics (KPIs) defined for each process and function lies at the heart of daily process management. While not all processes will use the same set of KPIs, most commonly the metrics fall into the following categories: safety, quality, delivery, productivity, and growth (SQDPG – often referred to as ‘squid pig’). In addition to KPIs, more and more organizations are also incorporating metrics linked to the behaviors that drive KPI performance. These are called key behavioral indicators (KBIs) and often measure things like how often a desired behavior is performed (e.g. the number or frequency of communications between sales reps and potential customers) or how thoroughly an activity was performed (e.g. the completeness of documentation required to setup a new customer).

Strategy deployment primarily involves senior leaders and middle management, but daily process management most heavily engages the front-lines: supervisors and the associates who perform the process activities. These front-line teams huddle daily to review the KPIs and identify issues. When possible, issues are immediately resolved. Those remaining are prioritized, with the most critical issues being addressed utilizing a locally led problem solving process.

This ongoing practice of reviewing metrics and resolving issues, day-in-day-out, drives continual incremental improvement within the business.

So far, we haven't mentioned data and analytics in our strategy deployment and daily process management discussion. So what role do they play? As Figure 7 shows, along with people and technology, data and analytics enable the integrated business management system to achieve significantly higher levels of performance. Said differently data and analytics should not be deployed for their own sake, instead they should be employed specifically within the framework of improving the performance of the organization's processes and management system. In the next section, we will dive into the details of how and where data and analytics should be applied to have the biggest organizational impact.

But before going there, here's a list of several additional items that continue to be important to the management of any business:

- Processes are the engines that deliver value across all businesses. Therefore, the focus should always be on ensuring the right processes are in place (traditional or digitized) and that they are performing at the highest levels.
- Strategic objectives are primarily achieved through the execution of successful projects. This means project management, agile, change management, and facilitation skills continue to remain important capabilities to foster within any business.
- Strategy deployment and daily process improvement both involve process improvement and problem solving. So, the proven concepts and methodologies of lean and six sigma, and their problem solving approaches (PDSA and DMAIC), continue to play critical roles.

We'll now get into the specifics of how to best leverage data and analytics within your business.

Don't deploy data and analytics for their own sake

To get the most value, deploy data and analytics as a set of methods and tools to enable higher performance of the organization's processes and management system

Data and Analytics as Enablers of Strategic Improvement

"Without a goal analytics is aimless and worthless" – Michael Porter

If, like us, you have been in the business world for a while, you likely have experienced this situation: A business leader goes to a conference and comes back energized about some new bright shiny technology, hardware, software, or concept. They then persuade the organization to buy the new thing or launch an initiative to apply the new concept, but it ends up yielding disappointing results. Why the failure? Because none of these things add value on their own; they only add value when employed to help fulfill a businesses' needs. Situations like this contribute to the previously discussed high failure rate (up to 85%) of big data and analytics initiatives.

People don't want the Cloud, Hadoop, Tableau, pivot tables, or internet enabled devices; they want the ability to execute, monitor, improve, and control their business processes

Extending the old saying 'people don't want a drill, they want a hole' to the world of data and analytics (and technology): People don't want the Cloud, Hadoop, Tableau, pivot tables, or internet enabled devices; they want the ability to execute, monitor and control their business processes, along with insights on how to improve them. In other words, any investment in or application of data and analytics should be directly driven by the need for higher performing and better managed business processes, which is also the domain of the company's integrated business management system – the organization's tool for driving strategic improvement.

We consider data and analytics, simply, as a set of tools that will only yield significant value when properly applied to enable and support the increased effectiveness of the company's integrated business management system. We'll now get into the specifics of where and how data and analytics can be leveraged, within and around the integrated business management system, to generate the most value for the organization.

- **Determining strategy**

  Figure 8 extracts the actionable components of the integrated business management system introduced previously and adds the critical 'determine strategy' component. The management system exists to support the businesses' achievement of its strategy, so it makes intuitive sense, that prior to putting strategy deployment and daily process management into motion, the business first needs to define its strategy.

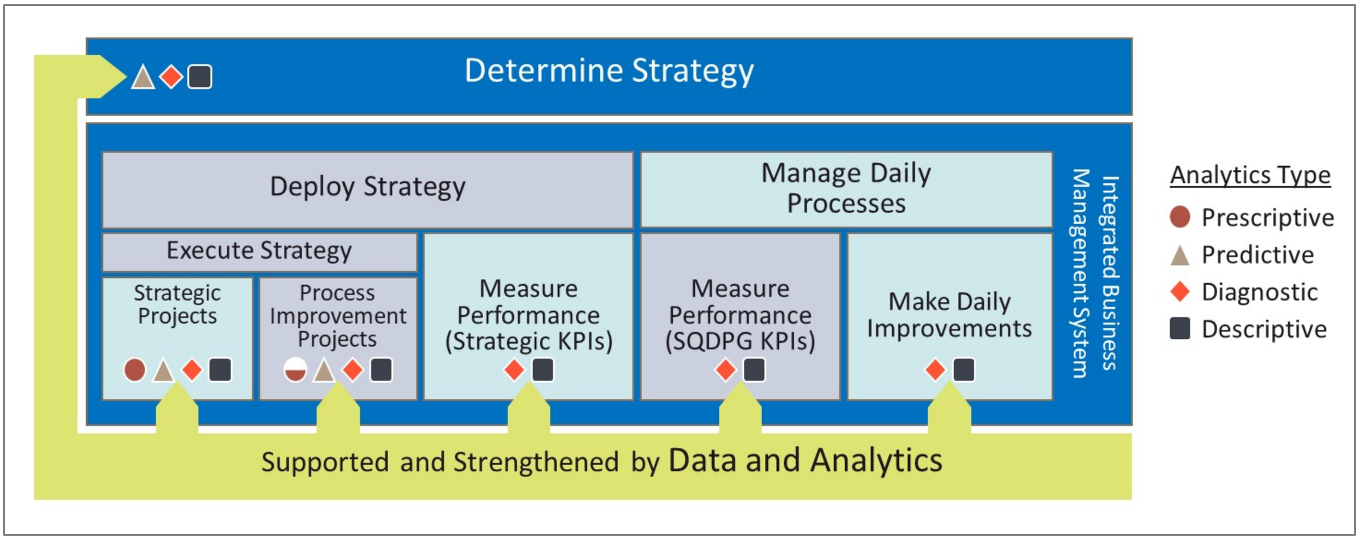
  Effective strategy definition requires inputs that help to answer questions like "which products, services, or customers generate the most profit for the organization and which will in the future?", "how have we performed against our existing strategic objectives and what are the causes that have prevented us from fully achieving one or more objectives?", "If we undertake this strategic initiative, what will the impact be on customer loyalty or profits?", etc. These questions will yield much better answers if data and analytics are employed to answer them.

In the book *The Discipline of Market Leaders (1997)*, the authors Michael Treacy and Fred Wiersma describe three competitive strategies, or value disciplines: product leadership, customer intimacy, and operational excellence. Their thesis is that to succeed in the marketplace, companies must master one (or two at most) of these competitive strategies and work to maintain acceptable performance in the others.

Companies that place the most weight on product (or service) leadership focus strategically on the technological aspects of products/services and all the different things their products/services can do and be used for. This is commonly the primary competitive strategy for consumer electronics companies, fund management firms, online retailers, automotive and pharmaceutical companies, and software providers. Table 3, with content primarily sourced from *Laursen, Business Analytics for Managers: Taking Business Intelligence Beyond Reporting, 2016*, contains a partial list of the analytics activities that can provide useful inputs in the strategy development process for this category.

When customer intimacy is the key competitive strategy, the focus is on driving good customer relations and customer loyalty. Companies in markets with high degrees of penetration, such as banks, insurance companies, telecoms, branded consumer products (e.g. Nike), and restaurants and hospitality (e.g. Starbucks, Marriott), often place a high-emphasis on this strategy. The biggest difference in the analytics needs of the customer intimacy dimension compared to product leadership, is analyses should now primarily focus on customers rather than products. Table 4 contains a partial list of useful analytics activities related to strategy development for customer intimacy. Once again, the information is primarily sourced from the previously referenced book by Laursen.

Figure 8: Strategy Definition and the Actionable Components of the Integrated Business Management System



Organizations that are capital-intensive and compete on economies of scale – the more that's produced the cheaper things get – often place an emphasis on <u>operational excellence</u>. This includes logistics companies, chemical refiners, other capital-intensive producers (e.g. mining, heavy equipment manufacturers), airlines, and hotels. A few other situations where a business might place an emphasis on cost control are: just after mergers to create synergies, and in a declining market when companies seek to counteract the impact of declining sales by focusing on reducing costs. Again, based significantly on Laursen's book, Table 5 contains a partial list of often useful Analytics activities supporting strategy development for operational excellence.

Table 3: Common Analyses: Product Leadership Category

ANALYSES	ANALYTICS TYPE
Analysis of which products/services deliver the highest profit	Descriptive
Models to describe the current state and how things are forecasted to change in the future – for the entire market and for individual products/services	Predictive
Cluster analysis to uncover and better understand customer segments	Descriptive (Advanced)
Customer segment analysis by products/services, features, and combinations of products/services	Descriptive
Correlation analyses between multiple offered products/services, to see which are complementary and which are substituting	Descriptive
Analyses to determine multiple-purchase patterns	Descriptive
For large numbers of products, analyses to uncover which products are sold together (basket analysis)	Descriptive

Table 4: Common Analyses: Customer Intimacy Category

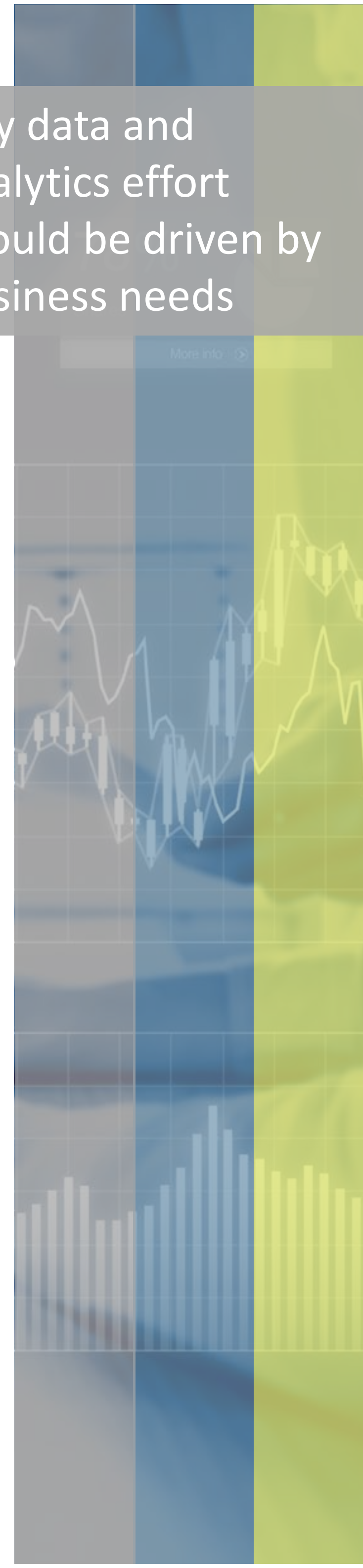
ANALYSES	ANALYTICS TYPE
Value based segmentation of customers, dividing them into categories like Platinum, Gold, Silver, based on their strategic importance and profitability to the business	Descriptive
Customer lifetime value analysis	Descriptive
Customer segment analysis by product/service, features, and combinations of products/services	Descriptive
Analysis of which customers use different means to buy products and engage with the business (e.g. local sales reps, call centers, websites)	Descriptive
Customer churn analysis – Which customers are most at risk and why?	Diagnostic and Predictive

Table 5: Common Analyses: Operational Excellence Category

ANALYSES	ANALYTICS TYPE
Analyses to compare internal process performance (often SQDPG KPIs) against the targeted level of achievement set by existing strategies	Descriptive
Analyses to understand why gaps exist between actual and targeted process performance	Diagnostic
Financial analyses to identify processes and organizational areas that contribute most to costs and have the largest opportunity for improvement	Descriptive
Models to predict the cost savings by merging or reorganizing resources	Predictive

Some of the analyses listed in Tables 3-5 may serve a useful role in supporting daily operations, so they may already exist; but others may need to be created specifically to support the definition of strategy. It's important to keep in mind that any analysis conducted to support strategy, or anything else, should only be done if there's pull from those who will use the resulting output. Any data and analytics effort should be driven by business needs.

When it comes to how to perform any analysis required to provide needed input into Strategy Development, we have developed a comprehensive 7 Step Analytics Cycle framework for you to follow, which is described in detail within its own section later in this paper.



Any data and analytics effort should be driven by business needs

- **Executing strategy**

Once the strategy development effort has yielded the critical few strategic objectives the organization needs to achieve, the focus moves to deploying and executing the strategy – the 'deploy strategy' element of the integrated business management system. While it's not shown in Figure 8, a key strategy deployment element is cascading the strategic objectives down and across the business. The work to achieve the objectives (execution) begins once they have been effectively cascaded and will be in the form of some type of focused effort, which we refer to as a 'project' when more formally managed. Sometimes the effective execution of a single project will fully achieve the desired strategic objective, but other times multiple projects, spread across various levels of the organization, will be required.

The projects used to execute strategy and support the achievement of strategic objectives, in general, can be divided into two types:

- Strategic projects: Focused on delivering new breakthrough capabilities to the business
- Process improvement projects: Targeting the improved performance of existing processes

So what role does data and analytics play in the project domain? Sometimes it's a supporting role and other times it's a leading role, but in today's business environment it's an ever-increasing focus.

Data and analytics are increasingly at the core of many strategic projects. For instance, a university's strategic objective is to increase students' success rates from X to Y, may lead to a strategic project that centers around building an analytics tool to accurately predict the likelihood of an individual student's success and provide an early warning for at-risk students so early action can be taken to correct course. Another example is a hotel company with a strategic objective of increasing the average annual revenue per room from A to B. In this case, they may launch a strategic project to develop a prescriptive analytics solution that automatically prices rooms based on a complex mix of variables (demographics, sociographics, demand levels, local market factors, etc.) to optimize their pricing and sales volumes. For these types of projects, the 7 Step Analytics Cycle should be the primary framework followed to execute the project.

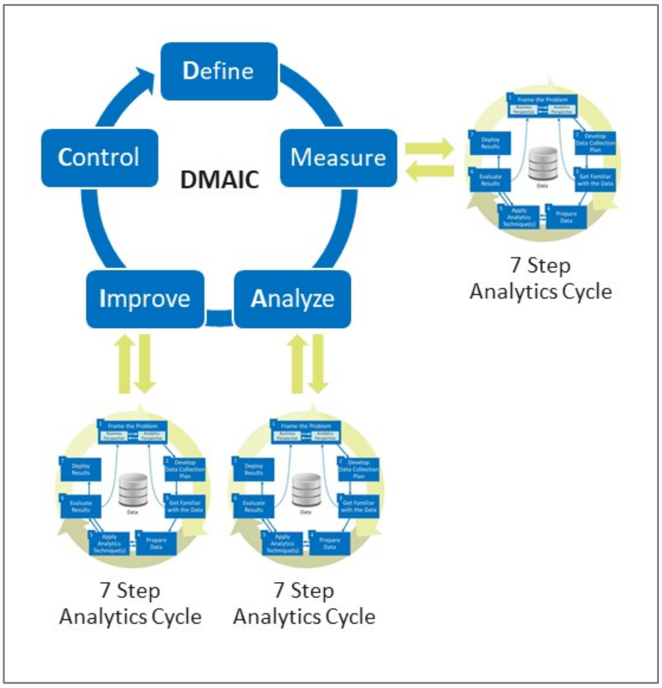
In cases where the strategic objective focuses on improving the overall performance of top-level company KPIs (e.g. reducing overall costs, increasing growth, improving safety, etc.) the key processes most impacting the targeted KPI are identified during the cascading process. Process improvement projects should then be launched to improve the performance of those key processes that are most underperforming. In this situation we recommend using the DMAIC process improvement methodology (used extensively within lean six sigma) as the primary guide for performing the project.

Within DMAIC, the 7 Step Analytics Cycle should be employed to support specific analytics activities driven by the needs of individual DMAIC phases (see Figure 9). In the Measure phase, the analytics cycle can be employed for any needed descriptive and high-level diagnostic analytics required to provide a clear baseline of current process performance. When identifying the root-cause(s) of poor process performance in the Analyze phase, the analytics cycle should be used to support the required diagnostic analytics. Later in the Improve phase, when the focus moves to solutions, the selected solution set may consist of or include analytics solutions very similar to the types seen in strategic projects. In this case, the analytics cycle should again be applied to support detailed solution development, optimization, and deployment. Still within Improve, a lighter version of the cycle should again be applied to perform the descriptive analytics to assess pilot results.

- **Measuring Performance**

  Within both elements of the integrated business management system: strategy deployment and daily process management, KPI dashboards should be developed and deployed to monitor performance. While the specific metrics may be different between the two elements, the process used is the same.

Figure 9: Applying the 7 Step Analytics Cycle within DMAIC



The strategic KPIs monitored as part of strategy deployment will typically focus on assessing the progress of strategic projects or on corporate level process performance-based metrics aligned to strategic objectives. KPIs linked to strategic projects commonly measure performance against schedule and performance against budget, with dashboards also including leading indicators related to risks, issues, pending decisions, etc. For large projects, divided into multiple workstreams, separate dashboards should be developed for each major workstream. KPIs on the workstream level dashboards should be able to be rolled-up and aggregated into their corresponding metrics at the project level.

Daily process management requires a balanced dashboard of process performance-based KPIs (often a SQDPG category), with similar strategic KPIs being required to actively manage strategic objectives targeting general performance improvement for key aspects of the business. It's not uncommon for there to be some overlap between the two business management system elements. For instance, in many industries 'safety' improvement is commonly a strategic objective, and it's also monitored on an ongoing basis as part of daily process management.

Process performance-based KPIs should be defined and connected in a cascading chain across the various levels of the organization. For example, in a large company, if 'safety' (safety incidents per million person hours) is identified as a corporate level KPI, then the corresponding division level, plant level, and process level KPIs should also be defined. The full-set of safety KPIs at the lower level (e.g. process level) should then be able to be rolled-together to create the next level's metric (e.g. plant level). At lower levels, additional leading metrics should also be developed to provide diagnostic information.

There are numerous business intelligence software packages on the market (e.g. Power BI, Tableau, etc.) that can be leveraged to build highly interactive KPI dashboards, providing users with the capability to easily and visually drill down into the data to uncover useful diagnostic insights. When it comes to developing and deploying needed dashboards, once again, we suggest using the 7 Step Analytics Cycle framework described in the next section as your primary guide.

- **Making Daily Improvements**

Within the process management system, starting at the front-line level, work teams and their leadership should review the KPI dashboards daily. A key purpose of the review is to identify

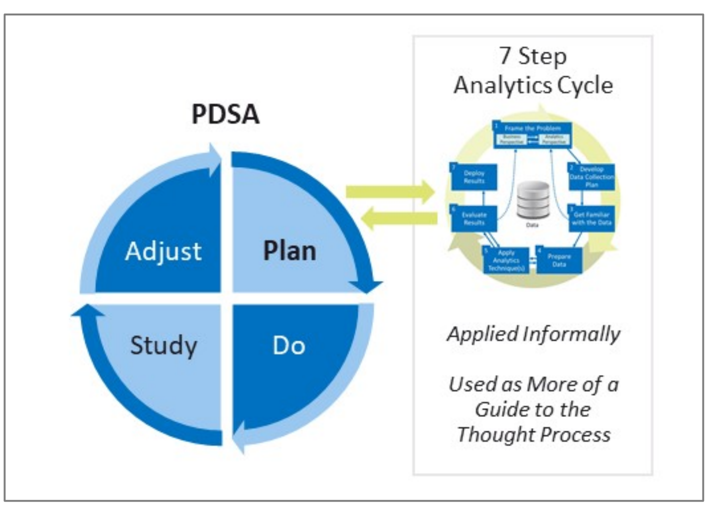
KPI dashboards should be developed and deployed to monitor performance in both elements of the Integrated Business Management System: Strategy Deployment and Daily Process Management

process performance issues and quickly resolve them. Sometimes the observed issues have obvious solutions, which should immediately be addressed. Other times, some analysis is required to appropriately identify causes and develop appropriate solutions. In this later case, high priority issues are typically assigned to local individuals or teams, who then utilize a basic problem-solving framework, such as PDSA (Plan, Do, Study, Adjust) to address the issue.

These types of problems typically don't require any rigorous analytics, but some basic descriptive and diagnostic analytics may be useful to help identify causes. We are not suggesting formally using the analytics cycle framework within local PDSA projects, but instead we recommend utilizing the cycle as more of a guide to the thought process for basic analyses within the Plan phase, as shown in Figure 10. During this phase, basic analyses are often required to quantitatively characterize the problem and help uncover causes to be addressed.

You should now have an improved understanding of the critical supporting role data and analytics plays to increase the effectiveness of a company's integrated business management system – the organization's key tool to drive strategic improvement – and where to best leverage it within the elements of the system. While we've referenced the 7 Step Analytics Cycle multiple times as the framework to be followed when performing analytics efforts, we haven't yet provided any details on the cycle. Those details are the focus of the next section.

Figure 10: Applying the 7 Step Analytics Cycle within PDSA



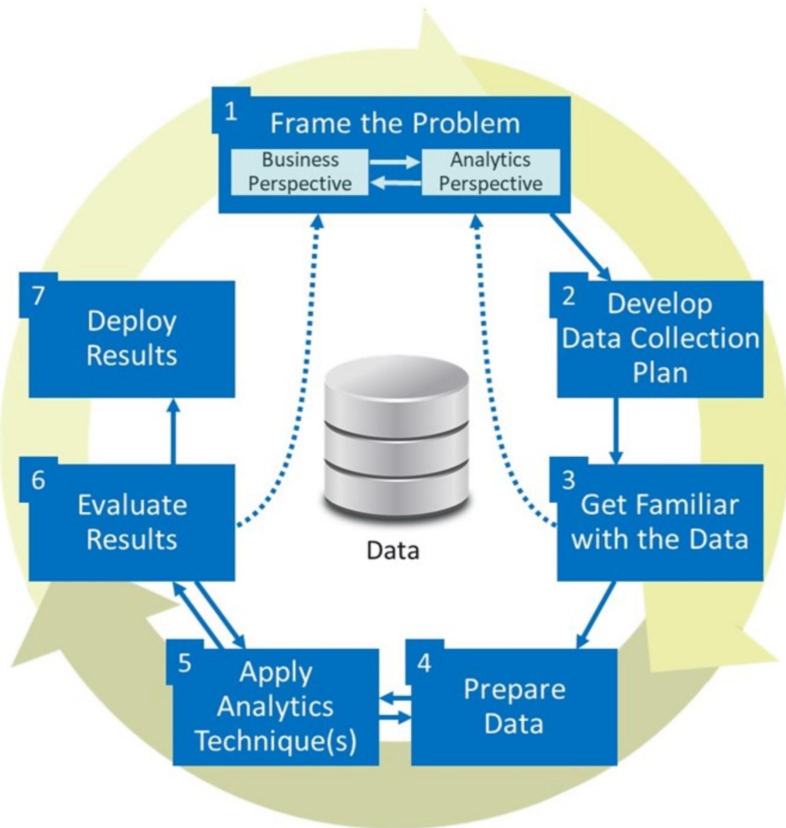
7 STEP ANALYTICS CYCLE

"Truth has nothing to do with the conclusion, and everything to do with the methodology"
– Stefan Molyneux

"Methodology should not be a fixed destination but a conversation about everything that could be made to happen" – J.C. Jones

To help insure data and analytics activities are aligned to business needs and that they deliver promised results, we have developed a comprehensive 7 Step Analytics Cycle framework, introduced in Figure 11, to guide your analytic efforts. This '7 Step' framework can and should be applied to any data and analytics project, regardless of complexity. For relatively simple efforts, like conducting a descriptive analysis on which products yield the most profit by region, the framework can be used informally as a guide for what things to consider and when working through the analysis. However, for much more complex efforts, requiring substantial amounts of time or money and having significant risk (e.g. developing a prescriptive model to utilize market data to determine optimal product pricing and make automated pricing changes), the framework should be followed in a much more rigorous fashion.

Figure 11: 7 Step Analytics Cycle



A more detailed list of activities within each of the 7 Steps is shown in Figure 12. While diving into the details, is beyond the scope of this paper, let's take a high-level look at each step to help you better understand how the cycle can be applied to improve the effectiveness of data and analytics efforts within your company.

1. **Frame the Problem**

   While all the 7 Steps are important, framing the problem is arguably the most critical. It helps ensure the effort is properly aligned to business needs, has the leadership support required for success, and that the problem has been effectively translated into the language of those conducting the analytics work.

   Ultimately, any analytics effort should be focused on fulfilling business needs, but the languages spoken, cultures, and problem framing needs are different between those running the business and those on the analytics side. For this reason, this step is divided into separate 'business perspective' and 'analytics perspective' framing elements.

   The business perspective element ensures all key stakeholders have a clear understanding and are aligned around the importance of the problem (why it should be worked on), as well as what will be delivered, by whom, by when, etc. If not already in place, as part of this element, a project champion should be identified. Having the active and strong engagement of a champion is one of the top, if not the top, factors most increasing the likelihood of initiative success. The champion is a key senior level stakeholder – from within the business, not IT or an analytics group – typically with a vested interest in the project's successful

Figure 12: Next-Level Activity Detail – 7 Step Analytics Cycle

1. Frame the Problem		2. Develop Data Collection Plan	3. Get Familiar with the Data	4. Prepare Data	5. Apply Analytics Technique(s)	6. Evaluate Results	7. Deploy Results
1.1 Business Perspective	1.2 Analytics Perspective						
1.1.1 Define Stakeholders	1.2.1 Ensure Clarity on Analytics Side	2.1 Identify What Variables to Measure	3.1 Collect Initial Data	4.1 Select Data	5.1 Select Analytic Technique	6.1 Evaluate Results	7.1 Plan for Deployment
1.1.2 Identify a Champion	1.2.2 Define Problem Inputs and Outputs	2.2 Identify How Variables Will Be Measured and Sources	3.2 Describe the Data	4.2 Clean Data	5.2 Verify Assumptions	6.2 Determine Next Steps	7.2 Plan Monitoring and Maintenance
1.1.3 Define the Business Problem	1.2.3 Explicitly State All Assumptions and Constraints	2.3 Define the Quantity of Data Required	3.3 Verify Data Quality	4.3 Construct New Data	5.3 Plan Statistical Diagnostics	6.3 Review the Process	7.3 Execute the Plans
1.1.4 Define Usability Requirements	1.2.4 Define Key Metrics of Success	2.4 Define How the Data Quality Will Be Validated	3.4 Explore the Data	4.4 Integrate Data	5.4 Conduct Analysis / Build Model		7.4 Produce Final Report
1.1.5 Confirm Amenability to Analytics	1.2.5 Obtain Stakeholder Agreement			4.5 Format Data	5.5 Assess Results / Model		7.5 Conduct Final Project Review
1.1.6 Define Timeline							
1.1.7 Identify Resource Requirements							
1.1.8 Confirm Stakeholder Alignment							

outcome, and with the clout to help breakdown and overcome any organizational barriers to success. An ideal champion candidate is the business leader whose area will most use or benefit from the output of the analytics project.

The analytics perspective element centers around the dialogue between the business people who have a problem needing to be solved and the analytics folks who will give them the information or tool to solve the problem. This dialogue should be facilitated by an individual who is trusted by both groups because of their fluency in the language and culture of each side.

To begin moving from the business problem toward the solution, the contents of the original business problem should be extended to include specifics required by those executing the analytics work. For instance, the analytics side needs to develop clarity on things like who are the users of the resulting system and its outputs, what data are needed, what data are available, what tangible outputs does the analytics engine (or black box) need to provide to be useful to the organization, and what are the likely input ('X's or independent) variables and output ('Y's or dependent) variables for the problem. Additionally, any constraints and assumptions should be identified and clearly stated, and success metrics need to be defined.

Framing the problem is typically an iterative process with the output of one activity or element impacting those of others. But once all components of the problem are adequately refined (typically in the form of a project charter document), they should be reviewed with the champion and other key stakeholders to gain alignment. Ultimately, the champion is responsible for signing off on the documented problem and making the final decision on whether to proceed with the project.

Once the project commences, the contents defining the problem should act as a 'contract' between the champion and those doing the project work. It should be considered a living contract, with updates being made later if needed, based on learnings and issues occurring during project execution.

2. Develop Data Collection Plan

Our favorite definition of analytics is "the scientific process of transforming data into insight for making better decisions" (source). With this said, the quality and value of the analytics solution is only as good as the quality of the data that goes into it, which is why three steps of the 7 Step Analytics Cycle are

The active and strong engagement of a Champion is one of the top, if not the top, factors most increasing the likelihood of initiative success

The Champion is a key senior level stakeholder, from within the business, not from IT or an analytics group dedicated to upfront planning, exploration, and preparation of the data to be analyzed.

During problem framing a set of input and output variables was identified, but prior to collecting and beginning to use any data, more thought should be given to ensure the efficient procurement of the data that will be of most value to the analytics effort. The following questions should be answered for each variable that you would like to collect data on:

- What is the data type? Is the data numeric, and if so, what specific type (e.g. continuous, binomial, ordinal, etc.)? Is the data qualitative, and if so is it based on text, visuals, or judgements? Qualitative data typically needs to be converted into quantitative (numerical) data for analysis, which will require a fair amount of work to precisely define.

- What is the operational definition of the variable? This is a precise, reliable, repeatable, and measurable definition of the variable to ensure consistency and reduce measurement error. For a continuous variable like 'time to fill an open job', the operational definition should specify exactly when the clock starts (e.g. when the position becomes open, when the job is posted, etc.) and when it stops (e.g. the first day there's a new person in the job, when the person is fully up to speed on the new job, etc.).

  Defining an effective operational definition is even more challenging for more qualitative metrics like the 'reason a customer has not paid an invoice'. Is it a specific reason stated by the customer or a pattern of underlying attributes (each needing their own operational definition) that defines the reason?

- What is the source of the data? Does the data already exist in a readily accessible database or will it need to be collected?

- Does the data for the variable exist on its own or will it need to be calculated from multiple values potentially distributed across multiple sources?

  At this point you may discover that data for some of the desired variables may be difficult to collect, so tradeoffs may be needed. When faced with this issue, you can assess the risk versus reward for the variable in question. Start by using a best guess to determine the likely value of collecting the data. This can be done by listing the possible findings if you collect the data, and for each finding, identifying whether it would lead to a different output (decision) and the corresponding impact in terms of value delivered to the business. Then for each finding, estimate the chance of its occurrence. The resulting information can then be used to determine the expected monetary value of each potential decision path, resulting in a more informed decision on whether to collect or not to collect the data.

- What stratification factors should be collected alongside the variable of interest? Characterizing data into different categories allows for comparisons of differences and patterns, often providing deeper insights that result in better and more fine-tuned analytics solutions. Types of stratification factors include: Reason codes linked to the data (e.g. what the defect is or why the customer called), when (e.g. shift, day of the week, or month), where (e.g. Asia, department A, or product XYZ), and who (e.g. customer ABC, millennials, or front-line associates).

- How much data is required and what is the sampling strategy to ensure the required levels of accuracy and precision of the resulting estimates? Along with the desired precision level, the data type will significantly impact the amount of data required to make needed decisions. For instance, holding everything else equal, continuous data requires a much smaller sample size compared to binomial (Yes/No) data.

It should be mentioned that given current technologies, the concept of sample size has diminished in importance somewhat since the collection and analysis of ALL the data is often possible. However the concept still has some relevance, because there remain situations when collecting all of the data is too timely or cost prohibitive.

- How will the validity of the data be ensured? For data from existing databases, how comfortable are you with the consistency of the data? For data that needs to be collected, has a measurement systems analysis (MSA) been conducted to verify acceptable levels of precision and accuracy?

The set of answers to the above questions for each variable, typically forms an individual line item within the larger data collection plan, which contains the corresponding answers for all the variables of interest. This purposeful approach to data collection will help you get the most value from your data for the least amount of cost and effort.

The data collection plan should be considered a living entity that will be updated as necessary as more is learned in subsequent steps of the 7 Step Analytics Cycle.

3. Get Familiar with the Data

Once the data collection plan is developed, data gathering begins. As the actual data starts to roll in, it's critical to spend some effort to get familiar with it. This will help avoid unexpected issues in the upcoming 'prepare the data' step, which is typically the longest within the 7 Step Analytics Cycle.

As the initial data is collected, the names, locations, and methods used to acquire each dataset should be documented, along with any challenges encountered. While one could argue that we're venturing a bit into 'prepare the data', the initial data may require loading into a specific tool, the data structure (rows and columns) might need to be changed, or multiple data sources may need to be integrated. This provides a great opportunity to identify and take early action on any observed problems, prior to more involved data preparation activities.

It's a good idea to compare the content of the initial datasets to what has been called for by the data collection plan. Check to see if the desired variables are represented and their formats are as expected. You may also discover additional potentially important variables, to be added to the data collection plan.

It's also important to assess the quantity of data records and amount of missing data within each variable. Is the appropriate amount of data available? Is there enough to ensure the analytics output has the desired precision and accuracy, but not so much that processing time becomes excessive? If the issue is too much data, an appropriate sampling strategy will need to be developed. Regarding missing values, you should be sure to explore their reasons and develop a plan for how to handle any missing data.

Another check to perform at this stage, is to confirm the data represent the current reality and are of acceptable quality. Here is a partial list of questions to ask:

- Is the data current?
- Is it from the location or process of interest?
- For character data, are there spelling or capitalization inconsistencies that may cause issues later (e.g. 'Yes', 'yes', 'Y', 'y')?

- Do all the data values appear plausible (e.g. teenagers with PhD educations)?

Once you have a solid grasp on the structure, quantity, and quality of the data you can graphically explore the initial data and perform some basic statistical analyses to potentially discover preliminary, previously unknown, insights about the data.

Prior to wrapping up this step, assess whether any elements of the business problem should be changed based on insights gained from your increased familiarity with the data. If changes are merited, work with the champion and key stakeholders to ensure their alignment and buy-in.

One final item to mention is that for some analytics problems, especially descriptive and simple diagnostic efforts, the output of this step will yield a clean enough result that can be implemented to solve the problem. Even for more advanced analytics problems, based on the insights gained, the stakeholders may decide more advanced modeling isn't needed anymore.

4. **Prepare the Data**

> Cleaning and preparing data are critically important to the analytics cycle, but they're typically the most time consuming and least enjoyable activities within the 7 Step Analytics Cycle

Cleaning and preparing data are critically important to the analytics cycle, but they're typically the most time consuming and least enjoyable activities. A quick Google search yields many estimates ranging from 40-80% of an analytics project's time and effort being spent on wrangling data. Over the longer-term, as data governance improves, and machine learning becomes increasingly leveraged to automate cleaning and preparation activities, the amount of manual effort required will be reduced. Thoroughly and diligently executing steps 1-3 of the analytics cycle can significantly reduce the required effort, but for now you will still need to commit significant effort to effectively prepare the data for analysis.

Preparing the data for analysis typically involves the following activities:

- <u>Select the data</u> to be used throughout the analysis. There are three primary dimensions to consider. The first is identifying the 'rows' of data to include in the analysis (e.g. data from the beginning of last year until today or the last 1000 data points). The second is selecting the data attributes or characteristics to include, based on the stratification factors identified in the data collection plan (e.g. specific customers, products, or geographic locations). The third dimension, which is required primarily for predictive and prescriptive analytics efforts, is to identify the subset of the first two dimensions to be allocated for model training and validation. Commonly 70% of the data is used to build the model and 30% is allocated to training and validation.

- Clean the data. While this is usually the most arduous part of the analysis, it's often the most necessary. This is especially true within pre-existing databases, which contain data collected for other purposes. In this instance, the quality of data is a function of the data's original use, so it likely won't fulfill the quality requirements of the current analysis.

  Data cleaning starts with identifying the range of valid responses for each variable and ensuring all data fields are properly labeled. It then requires the identification and resolution of items such as the five shown below (Source):

  i. Invalid data responses (e.g. out of range data points or where letters are used when numbers are required)

  ii. Inconsistent data encodings (e.g. different abbreviations used for one state, customer, or product)

  iii. Suspicious data responses (e.g. an individual providing the exact same response to every question on a survey)

  iv. Suspicious distribution of values (e.g. when 99% of survey respondents from poor neighborhoods have incomes of more than a million dollars)

  v. Suspicious interrelationships between variables (e.g. two highly correlated variables, where one of the variables is simply a now unneeded calculation based on the other variable – one field in dollars, the other field calculated as thousands of dollars by dividing the original field by 1000)

- Construct new data fields derived from existing ones. For instance, it may be useful to create a 'Processing Time' field based on calculating the difference between existing 'Start Time' and 'Completion Time' fields.

- Integrate the data. When the required data is spread across multiple sources and databases it needs to be effectively integrated prior to analysis. For instance, when developing a business intelligence tool to show the current location of trailers requiring preventative maintenance (PMs) for a trucking company. One database may include information on the trailers requiring PMs and another database may contain the current GPS location of the trailers. The information in these databases need to be integrated based on a common trailer unit number, potentially by simply using VLookup functionality within Excel or linking the two databases together using the functionality in a business intelligence product like Power BI.

- Format the data. Some Analytics tools have specific requirements for the order of fields, such as requiring the first field to be a unique record identifier. Also, some tools require the data records to be sorted in a particular order. For instance, for many modeling algorithms, it's often best to input data in random order. These formatting changes are primarily syntactic modifications made to the data, without changing the meaning of the data.

5. Apply Analytics Technique(s)

Up to this point, almost all the hard work put into the analytics project has been focused on planning and preparing, but in this step, we make the transition into doing the analysis; and generating results that help address the business problem.

Typically, analyses are conducted in multiple iterations. For instance, an initial analysis may be run to screen out insignificant variables in a model, followed by additional analyses to further refine and fine-tune the model. Additionally, with the multitude of analytics methods available, there are usually multiple options and approaches that can be employed to address the same problem. Often multiple analyses, each using a different approach, are performed concurrently to see if one approach yields a better data fit. Different analytics methods generally drive different data preparation requirements, so it's very common to move back-and-forth between the prepare the data and apply analytics technique(s) steps multiple times throughout the project lifecycle.

The Apply Analytics Technique(s) step consists of five sets of activities described below:

- Select the analytic technique. It's likely that by this step, you have a good idea about what type of analytics technique(s) you want to apply, but it's now time to make solid decisions about which specific one(s) you will use. A detailed discussion of the multitude of available tools and techniques is beyond the scope of this paper, but the content shown in Figure 5 will help get you pointed toward the right tool. It's useful to start by determining which analytics type best aligns with the question your trying to answer within your problem. For instance, if you need to answer 'what will happen if?', then you will likely select from the tool families in the predictive analytics domain or those in the lower level domains (diagnostic and descriptive).

  Along with using the business problem to guide tool selection, here are some other considerations to help you decide on the most appropriate analytics technique(s):

  - Time available for analysis and fulfilling the business needs
  - The accuracy of results required
  - The accuracy of the underlying data
  - Data availability and readiness
  - Staff and resource availability and readiness
  - Methodology popularity and acceptance

  In conjunction with selecting the analytics technique(s), the corresponding software tool to be used to conduct the analysis needs to be selected. Examples include: Microsoft Excel, statistical analysis packages (e.g. Minitab or JMP), business intelligence systems (e.g. PowerBI or Tableau), optimization systems, simulation systems, etc. Of course, there is often an interplay between the selected software tool and the analytics technique(s) to be used. If the stakeholders are only comfortable with Excel, you will be limited by the analytics capabilities of Excel or compatible add-on systems (e.g. Oracle Crystal Ball).

- Verify the assumptions. Many analytics techniques requires specific assumptions about the data be upheld. For instance, logistics regression requires that all data types are known before execution. Some other tools don't allow any missing values or require that variables follow specific distributions. Be sure to document any data assumptions along with any data manipulations required to fulfill the requirements of the analytics tool.

- Plan the statistical diagnostics that will be run to uncover any shortcomings of the analysis. Statistical diagnostics help detect mistakes or weaknesses and measure the accuracy and reliability of the analysis. Additionally, they provide insight into interpreting results and potentially improved solutions. (Source: *Bartlett, 2013, A Practitioner's Guide to Business Analytics*). There are many diagnostic families that can be employed depending on the situation, with tools and approaches often combined from across multiple families. An in-depth discussion of the topic is well beyond the scope of this paper, but here some examples of diagnostic families: Data splitting, using external numbers, juxtaposing results, resampling techniques with replacement, simulation/stress testing, tests for statistical assumptions, tests for business assumptions, tools for performance measurement, etc.

  The key is before running the analytics technique or building a model, that you think through and develop a plan for running the necessary diagnostics to test and ensure the quality and validity of the analytics output.

- Conduct the analysis or build the model. At this stage, based on all of the preparatory work you have done, you should be more than ready to conduct the analyses or build the desired models. Typically, you will conduct several different analyses or build multiple unique models prior to homing in on the final solution. Be sure to document the significant variables or parameters, their chosen or optimal values, and any supporting rationale.

- Assess the results or model. Once you have what appears to be a satisfactory set of analytics results or model(s), verify their performance relative to the planned statistical diagnostics. Look to ensure the solution(s) are easily deployable and if they appear to achieve the problem's success criteria.

  If you are satisfied with the quality and validity of the results/model, the ease of deployment, and that the project goals will be met, it's time to move on to the final two steps of the analytics cycle: more in-depth evaluation and final deployment. If not, apply what has been learned and rerun the analyses or models.

6. Evaluate Results

As an input to this step, you should have an analytics solution set (e.g. a predictive model, a KPI dashboard, etc.) that has been statistically and technically vetted, along with an acceptable degree of confidence that it is deployable and will deliver the desired business goals. However, prior to deployment, a more rigorous verification of how well the analytics solution(s) solves the business problem and meets stakeholders' needs should be conducted.

Stakeholders are typically interested in how the proposed solution set will help them – in their language, so most won't need to fully understand all the technical ins-and-outs. A peer review for technical correctness is strongly recommended, but beyond that, results verification should focus on closely working with key stakeholders to answer the following:

- Has the business context changed since the beginning of the project, impacting the effectiveness of the proposed analytics solution(s)? If so, what business needs or assumptions have been impacted and to what degree?

- How well does the proposed analytics solution(s) impact each success metric? Are there any deficiencies?
- What questions or concerns do stakeholders have about the proposed analytics solution(s)? Have the stakeholders' questions or concerns been answered/resolved to their satisfaction?

One caution is to watch-out for stakeholders who want changes made to the analytic(s) solutions so 'senior leadership will be provided the news they want to hear' or to 'play politics'. Results need to have integrity (and be integral to solving the problem) for them to be accepted by the organization.

Once the above evaluation is completed, continue to work with the stakeholders and the champion to achieve consensus on one of the three following next steps:

a. Move forward to the deployment step.

b. Go back and reapply refined or new analytics techniques. If the results are directionally correct, but not quite optimal, build on what's been learned thus far and conduct another round of analytics to develop a better set of solutions.

c. Go all the way back and reframe the problem or in extreme cases, if it doesn't appear an acceptable solution is cost effectively achievable, decide to end the effort without more resource expenditure. This should be a last option, but in some rare cases it will be the best path forward for the business. Even if the decision is made to not deploy a solution, a detailed report should be developed detailing the effort and lessons learned, to provide a source of knowledge for the organization, helping to increase the likelihood of successful future Analytics projects.

A final activity within this step is to take some time to reflect on and review the project. This is again pointing an eye toward creating knowledge to increase the value delivered by future analytics projects. Start by summarizing all the activities and decisions for each of the first six steps of the analytics cycle. Then for each step, explore the following questions and make suggestions for what should be done differently if you could do it again:

- To what degree did the step contribute value to the final results?
- What were the surprises, dead-ends (e.g. a specific analysis that didn't add any value), failures, or mistakes within the step? How can they be predicted or avoided the next time?
- Are there ways to streamline the activities within or make the step more efficient?
- Are there alternative decisions or approaches that may have been more useful within the step?

7. Deploy Results

In this step, the newly gained insights are deployed into the business to make desired improvements. Depending on the nature of the business problem, deployment can mean different things. It may mean implementing new analytics models as one part of comprehensive new digitized process. Or it could mean deploying the insights gained to drive behavioral changes within the business, without formal information system changes. For instance, the analysis may have shown that a certain group of customers are much more likely to default on their payments.

Typically, this step consists of the following five activities:

- Plan for deployment. Up this point all the hard Analytics work has resulted in a promise of improved business performance, but the actual improvements will only be realized when fully deployed. A comprehensive deployment plan should be developed and executed to help ensure the smooth and effective implementation and adoption of the analytics results. Developing the deployment plan consists of the following tasks:
  - Summarize the deployable results. This helps you determine what models and algorithms to integrate into information systems and what findings should be communicated to users.
  - Develop the detailed workplan for deployment and integration into systems.
  - Develop a plan for communicating the analytics results and deployment process to any users or anyone who will be impacted.
  - Develop a plan for updating any process documentation for any impacted processes.
  - Develop a plan for training users on how to use the new system(s) or apply the new insights.
  - Create a plan for monitoring the deployment itself.
  - Conduct a deployment risk analysis to identify and proactively mitigate the most critical risks that may negatively affect deployment.
- Plan for ongoing monitoring and maintenance. Monitoring and maintenance are important factors in ensuring the long-term viability of the analytics results within the day-to-day business and its environment. A well-thought-out plan helps prevent lengthy periods of poor usage or inaccurate output. A solid plan will also support the identification of changing user requirements or customer behavior (e.g. the factors that drive customer loyalty), requiring the ongoing calibration and continuous improvement of the analytics model(s).

  Consider the following when developing the monitoring and maintenance plan:
  - What things could (and are most likely to) change in the environment impacting the performance of the analytics results/model?
  - How will ongoing quality (accuracy and precision) be monitored?
  - What triggers will signal when the analytics results/model shouldn't be used anymore?
- Execute the plans. This is relatively straightforward, but requires a high degree of discipline, especially as complexity increases. Consider utilizing an individual with proven project management skills to facilitate and stay on top of all the required details of the deployment plan. You should also assign an individual to be responsible for the ongoing execution of the monitoring and maintenance plan.

  Even with the best plans, unexpected events will occur, so be prepared to adjust as necessary.

- Produce a final report. Writing a final report provides a historical record of the project that can aid future efforts, but it can also be used to communicate results. Be sure to consider your audiences for the report. There is value in documenting the technical aspects of the project for the analytics community, as well as documenting the overall approach and findings for business stakeholders. The languages and interests of these two groups are very different, so the documentation should be customized to their individual needs. Specifically, the business stakeholders may find more value in a presentation rather than a formal report.
- Conduct a final project review. This is the final activity of the 7 Step Analytics Cycle and should build off the review conducted in the evaluate results step. It allows you to capture and incorporate lessons learned from the deployment, as well as to get final impressions from stakeholders.

  Plan on conducting interviews with key project stakeholders. Here are several questions to ask during the interviews:

  - What were your overall impressions of the project?
  - What did you learn during the process about analytics and the data available?
  - What went well? What didn't go well?
  - If the project was to be done again, what changes would you like to see?
  - Document the learnings once they are captured and summarized. Once again, considering the different needs between the analytics and business stakeholder communities, plan on communicating the learnings.

You should now understand how the comprehensive 7 Step Analytics Cycle serves as a highly adaptable framework to help insure data and analytics activities are aligned to business needs and that they deliver needed results, in support of the organization's integrated business management system.

Ingraining Data and Analytics into Your Organization's Culture

"To me, I learned along the way, you know, culture is behavior. That's all it is; culture is people's behaviors."
– Ginni Rometty, CEO of IBM

So let's take some time to reflect on what we've discussed so far: The amount of data around us is and will continue growing exponentially; there is tremendous value in being able to exploit the data – and in some cases it's a survival issue; while some value is being captured, organizations are just scratching the surface of what's possible, with the majority of projects failing and programs to create data-driven cultures yet to achieve much success. We've also discussed that data and analytics shouldn't be deployed for their own sake, instead they should be linked to strategy and used to support and enable the organization's systems; and to realize the greatest value, data and analytics need to be ingrained into the way your company does business – the way people behave – day-in-day-out, across all levels and functions. They need to become part of your culture.

The problem is, culture change is hard! Your company already has a culture, that may have been intentionally engineered, but more likely it has just evolved over time – a 'de facto' culture. Regardless of how it developed, the continually repeated behaviors making up your culture, have likely jelled into automatic hard-wired habits. Think about how difficult it is to change your own personal habits, then multiply that across the entire workforce – that's the challenge your business faces.

With the above said, while it requires sustained focus and disciplined execution, effective cultural change is possible; and when done right can deliver a very high return on invested effort. There are some unique actions and considerations when integrating data and analytics into your organization, but for the most part, the factors and activities are the same as those that make for any successful large-scale cultural change.

As already mentioned, behaviors are at the heart of culture. As shown in Figure 13, the collective behaviors of the organization lead to the results achieved – good or bad. All work output of the business is a result of its systems, and these systems create the conditions that cause people within the organization to behave a certain way – they reinforce and guide behaviors. Systems can and should be formally defined and managed (e.g. standardized problem-solving methodologies, leadership development programs, etc.) but are often informal based on unconsciously performed habits that have simply evolved based on organizational norms. Think about how decisions are made or what behaviors get rewarded within your organization.

Data and analytics need to be ingrained into the way your company does business – the way people behave – day-in-day-out, across all levels and functions. They need to become part of your culture.

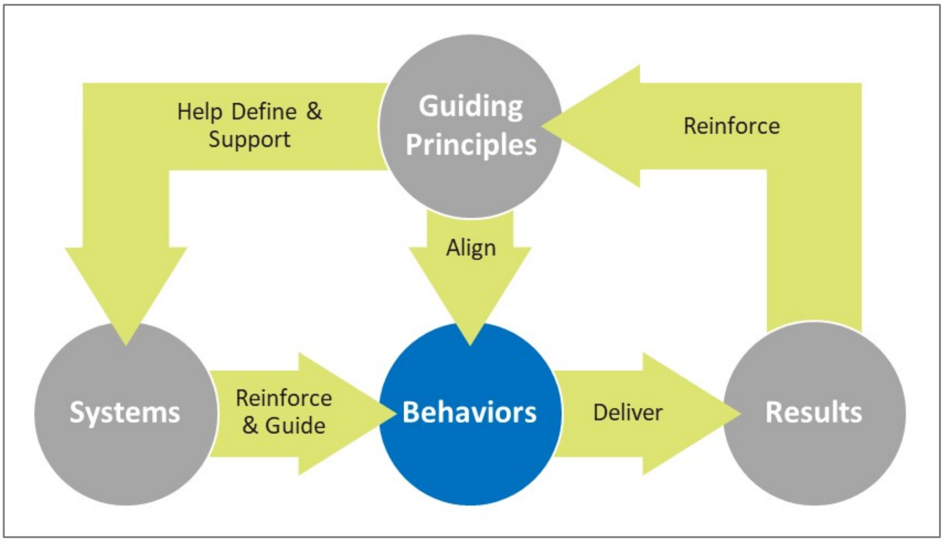
Our bet is that for most of you, these simply occur within your business, often inconsistently and without much thought about the behaviors they reinforce. When you're seeking to change culture, start by defining the desired behaviors that you want and then ensure that the necessary systems are put in place to drive those behaviors.

For instance, if you want your people to continually manage, monitor, and improve daily operations, you should implement a daily process management system. If you want your people to work in an aligned fashion to achieve strategic business objectives, you need a high-functioning strategy deployment system. When it comes to data and analytics, the key behavior we're tying to drive is people consistently leveraging data and analytics to gain deep insights and to make fact-based decisions on the things that matter most to the business. Deploying the elements of the integrated business management system will help drive the 'on the things that matter most to the business' component of the behavior but ingraining consistently 'leveraging data and analytics' and making 'fact-based decisions' requires several additional high-level systems:

- Data and analytics steering team
- Data and analytics center of excellence (DACE)
- Analytics community of practice

When you're seeking to change culture, start by defining the desired behaviors that you want and then ensure that the necessary systems are put in place to drive those behaviors

Figure 13: Role of Behavior in Cultural Change



We'll discuss how to deploy and explore the components of these systems in just a bit, but first we need to mention the concept of 'guiding principles' also shown in Figure 13. A principle, as defined by Stephen Covey, is a "natural law that is universally understood, timeless in its meaning and self-evident" (source). The organization's values drive actions, but the consequences of those actions are governed by principles. The Shingo Institute at Utah State University has developed a list of ten guiding principles and a larger set of supporting concepts (shown in Table 6) that form the basis for creating a lasting culture of enterprise excellence, and while beyond the scope of this paper, we encourage you to explore and understand them on your own by clicking here.

The guiding principles and supporting concepts should be used to help design any new system. In other words, if it doesn't support a specific principle or concept, the system will perform sub-optimally over the long-term.

Additionally, all behaviors should be aligned to the guiding principles and supporting concepts, so a workforce well versed in them will question and challenge behaviors that go counter to them.

In a previous section, we stated that to get the most value from data and analytics one required action is to "address the foundational factors that make for any successful large-scale cultural change or initiative". This rings true to deploying an integrated business management system, as well as to deploying a data and analytics center of excellence (DACE), or an analytics community of practice. Fortunately, based on the hard-earned experience of others who have deployed major

Table 6: Shingo Guiding Principles and Supporting Concepts

CATEGORY	GUIDING PRINCIPLES	SUPPORTING CONCEPTS
Results	• Create Value for the Customer	• Measure What Matters • Who is the Customer? • Identify Cause-and-Effect Relationships
Enterprise Alignment	• Create Constancy of Purpose • Think Systemically	• See Reality • Focus on Long Term • Align Systems • Align Behaviors with Performance • Policy Deployment • Standardize Daily Management
Continuous Improvement	• Embrace Scientific Thinking • Focus on Process • Flow & Pull Value • Assure Quality at the Source • Seek Perfection	• Rely on Data and Facts • Integrate Improvement with Work • Stabilize Processes • Standard Work • Go and Observe • Focus on Value Stream • Keep it Simple and Visual • Identify and Eliminate Waste • No Defect Passed Forward
Cultural Enablers	• Lead with Humility • Respect Every Individual	• Assure a Safe Environment • Develop People • Empower and Involve People

changes, the factors and activities required for successful large-scale change efforts are now well defined.

Most critically, any large-scale cultural change should be linked to a strategic objective, or better yet, it should be defined as a strategic objective and executed as a strategic project of its own.

Also, as a strategic objective, another key success factor is to be actively championed by senior leadership. In support of this and specific to data and analytics, a 2016 McKinsey Global Survey found that organizations with high-performing analytics programs are nearly three times more likely than their low-performing peers to have CEO sponsorship. Including these factors, Figure 14 presents a larger set of the various elements that must be actively and rigorously managed to deploy large-scale cultural change.

Our objective here is to make you aware of all the elements, not to go into the details of each. Providers like Prosci® have programs to train and certify change management practitioners. We strongly encourage that you leverage one or more of these certified individuals to help facilitate the deployment of any of the systems discussed.

Now that you are aware of the systems required to ingrain data and analytics into your business, and what's required to effectively deploy each system, we'll now get into the details of the three required data and analytics specific systems, shown in Figure 15, starting with the data and analytics steering team.

- **Data and Analytics Steering Team**

  If your business is serious about ingraining data and analytics into the culture, it should be defined as a strategic objective and managed like a strategic project. Along with having a highly engaged senior leadership champion, a best practice for strategic projects is to also utilize a steering team comprised of high-level stakeholders from across the business. The steering team is responsible for guiding the overall strategic direction, setting strategic priorities and targets, holding involved groups

Figure 14: Elements of Large-Scale Change Initiatives

accountable for and monitoring progress, providing guidance, and promoting actions to instill data-based decision making across the business.

Early on, as priorities and plans are being flushed out and many decisions needing to be made, the team will need to come together frequently and sometimes for working sessions extending beyond the typical time allotted to most organizational meetings. However, as the plans are put into motion, the team's activities will transition to primarily reviewing progress against plan and providing, mostly minor, needed course changes. Mature steering teams often settle into a quarterly progress review cadence.

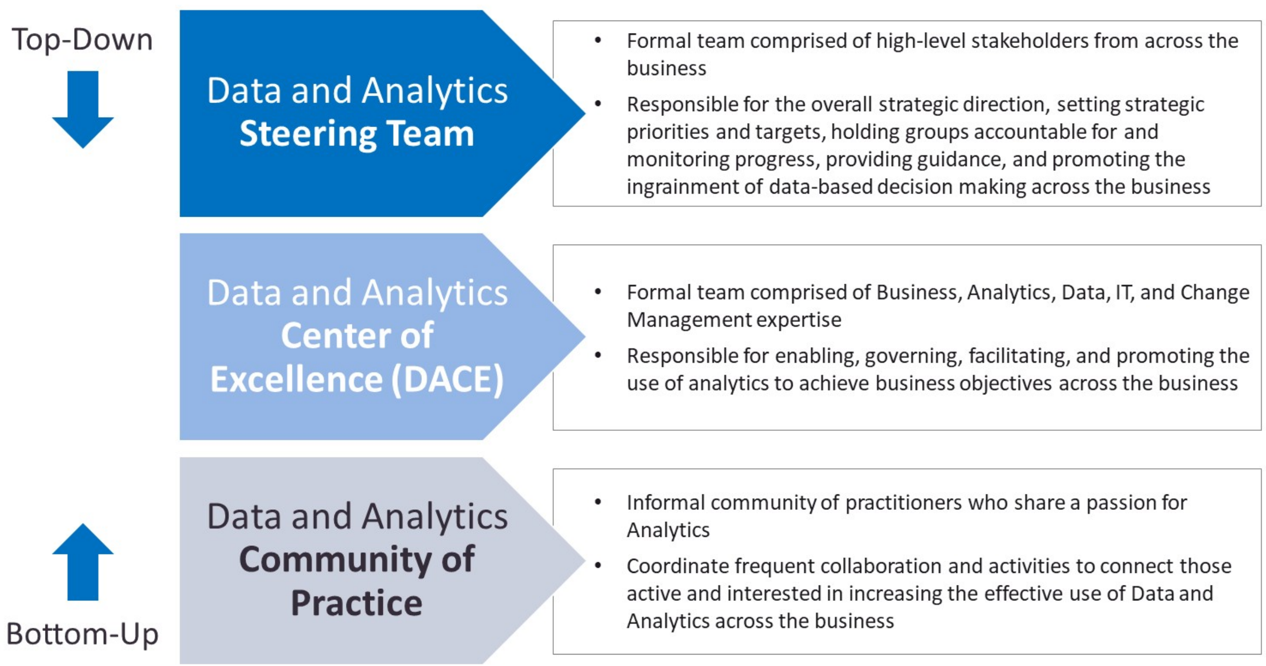
The steering team should be considered a semi-permanent system that will continue to operate as long as ingraining data and analytics remains an organizational priority. While all signs point to the importance of data and analytics only increasing over time, ideally the point will eventually be achieved, when it just becomes 'the way' the organization operates. At this point, when it is ingrained in the culture, the decision may be made to retire the data and analytics steering team system.

- **Data and Analytics Center of Excellence (DACE)**

  The data and analytics center of excellence is the system responsible for converting the strategic desire to culturally ingrain data and analytics into tactical action and tangible results. The DACE supports the definition and operationalization of the strategy to achieve the data and analytics strategic objective.

  Any time multiple organizational skillsets and functions are required to achieve an objective, if not effectively managed, barriers will likely arise that significantly limit success. With its required blend of business, analytics, IT, and change

Figure 15: Systems to Support and Ingrain a Data and Analytics Culture



management competencies and functions, culturally ingraining data and analytics is no exception. The DACE is comprised of individuals with competencies spanning the full spectrum of involved functions, who are charged with actively working to ensure the functions are working together to achieve the larger data and analytics cultural objective.

Depending on your business situation, the DACE can be a formal organizational unit (full-time resources), an informal virtual group (part-time resources), or a hybrid of the two. Regardless of the structure, we strongly recommend the DACE being positioned as a business system, not an IT system. Data and analytics are all about maximizing the impact delivered to the business, so business needs should drive any technical initiatives.

Here are the typical responsibilities of a mature DACE:

- Data and analytics strategy development and execution. Working closely with the steering team, the DACE is responsible for developing the overall data and analytics strategy to fulfill the larger enterprise strategic objective. In addition, it is responsible for developing and facilitating the execution of the full set of elements required to fulfill the strategy. These elements will closely align with those of any large-scale change initiative, shown previously in Figure 14.

  While it's beyond our scope to discuss each element, we do want to spend a minute discussing 'roles and responsibilities'. It's here where the DACE will need to define its resource plan for analytics practitioners across the organization (e.g. data scientists, power users, self-service BI practitioners). This should be based on the rigor and complexity of the projected analytics demands. In other words, is your business primarily at the stage where it will gain the most value from business intelligence related analytics? Then you should likely stay focused on power users and self-service BI practitioners. Or is advanced analytics skills more valuable to your organization? In this case you may want to go down the data scientist path.

If you chose to build a stable of analytics power users, to lead business intelligence efforts and some limited predictive analytics based projects, you will need to decide how many people, as well as if the role will be a full-time job or part-time activity within their current role. The answer is different depending on the nature of your organization, but if you're going the power user route, we typically recommend the development and deployment of at least a few full-time resources (akin to the role a lean six sigma black belt plays). These full-time individuals will be responsible for leading analytics projects, as well as training and coaching others – especially self-service BI practitioners.

One rule of thumb when deploying power users, full-time or part-time, is to only deploy based on actual strategically-linked project demand. Training power users who aren't assigned to an analytics project that's clearly linked to organizational strategy but expecting them to apply their new skills to improve the business, is a recipe for disaster – false starts, wasted resources, and solutions the business doesn't need.

With regard to where to position power users within the organization, we suggest utilizing a matrixed approach where they report directly into the function they supported and dotted line into the DACE for training, coaching, and analytics process governance purposes.

- Data governance leadership to ensure the availability, usability, integrity, and security of data used by the organization. This typically includes the development and management of the necessary processes to ensure high-quality data and secure consistent data usage (i.e. master data management); metadata documentation and transparency, so users can find needed data; as well as the deployment of data stewards (individuals or teams) who work with the business to help ensure data use conforms to the defined data governance policies.
- Facilitating an ongoing dialogue between the business and IT to ensure the chosen information architecture and technologies (IT owned) support the data and analytics strategy (business owned).
- Providing required tools and processes. Along with identifying a standard set of software to support analytics activities, the DACE is also responsible for developing, maintaining, and stewarding the use of standard processes within the data and analytics domain. A set of core standard process will help ensure the right analytics activities are done in the right way to maximize the results delivered to the organization. Examples include: a standard process, with clearly identified steps and roles, to select, prioritize, and sequence analytics projects across the organization; the 7 Step Analytics Cycle for executing projects; standard processes for project progress reviews, documentation, etc.
- Providing training and coaching to build data and analytics capabilities throughout the organization. We previously discussed the rich mix of (often new) skills required in a data and analytics enabled organization. Sometimes when advanced skills are required, hiring experienced individuals is the answer (e.g. data scientists). However, when ingraining data and analytics into the culture, everyone in the business will eventually need some degree of capability. Unless you're looking at replacing the majority of your workforce, you will need to invest in developing the needed skills of your current employees – the power users, self-service analytics practitioners, and anyone who needs to be making data-based decisions.

Without getting into the multitude of training delivery options, to maximize learning, we recommend incorporating hands-on real-world application practice whenever possible.

Table 7 provides a partial list of training topics that a DACE will likely deliver or sponsor within the organization.

Table 7: Partial List of Training Topics Supported by DACE

Training Topic	Decision Makers	Self-Service Practitioners	Power Users
Software Tool Usage			
Excel Analytics Basics	■	■	■
Business Intelligence Tool Usage (e.g. Power BI, Tableau)		■	■
Advanced Analytics Software Tool Usage			■
Data Basics			
Data Governance: The Whys, Whats, and Hows	■	■	■
Understanding and Navigating the Organization's Data Landscape		■	■
Analytics Skills Development			
Data-Based Decision Basics: • The Whys, Whats, and Hows • Avoiding Cognitive Biases	■	■	■
Effective Data Visualization	□	■	■
7 Step Analytics Cycle		□	■
Basic Analytics: Data Types, Metrics, Descriptive Statistics	□	■	■
Diagnostic Analytics: Foundational Tools		□	■
Predictive Analytics Basics			■
Misc. Analytics Tools			□

Key
Required: ■
Suggested: □

In addition to training, capability development requires multiple cycles of application to build expertise. The DACE should ensure coaching resources are available to support new practitioners through their initial application cycles.

Promote the effective use of data and analytics. The DACE should develop and actively manage a communication plan to share success stories, best practices, and to recognize those who embrace data-based decision making. As part of the promotion effort, the DACE should also play a leading role in data and analytics community of practice.

- **Data and Analytics Community of Practice**

The steering team and DACE are both systems that work from the top of the organization down, with the objective of driving data and analytics into the culture. We recommend balancing these systems out with a third, less formal system that supports more of a grassroots approach to ingrain data and analytics in the organizational culture – the data and analytics community of practice.

A community of practice provides groups of individuals with shared interests – in our case, data and analytics – a vehicle for coming together to learn new skills, share best practices and war-stories, discuss and solve problems, and to network. The community of practice should receive an initial kick-start, some basic guidance, and support from the DACE, but for the most part be autonomous and self-organizing; driven by the needs and interests of the data and analytics community.

While the community is self-organizing, in our experience, early on, the primary focus tends to be on building skills and sharing best-practices around better utilizing self-service tools (e.g. Power BI) and effectively visualizing data.

Typical community of practice facilitated activities include:

- Coordinating regular scheduled learning sessions on topics recommended by members
- Providing a forum for members to ask and receive answers to questions
- Providing peer-to-peer mentoring support
- Holding larger showcase (conference) events to bring together the entire community for best practice sharing and learning
- Social activities to foster networking among members, ultimately increasing trust and knowledge sharing

We have now introduced you to the full set of systems, that when effectively deployed with lots of attention and hard-work, will facilitate the behavior changes that ingrain data and analytics into your culture. We argue that culture change is more of a process than an art. It starts with defining desired behaviors, developing and deploying systems aligned with the organization's guiding principles to reinforce and guide the behaviors, and rigorously applying the full-set of proven change management elements to support the process.

We Can Help

If you're at the point of making a serious commitment toward better leveraging data and analytics, to achieve sustainable strategic advantage for your organization, Proficiency Systems can help. Contact us to learn more.