lifescale analytics

analytics+insights
for life science

Descriptive to Prescriptive

Accelerating Business Insights with Data Analytics

a lifescale leadership brief

The potential of data analytics can be confusing for many business leaders, often leading to more questions than answers.

How do I solve for situations that didn't even exist a year ago? What tools and methods should I deploy for my present business challenge? Where do I start?

And many data analytic software vendors don't help. Their offerings are shrouded in technical double-talk, obscure acronyms and hyperbole creating an aura of wizardly expertise that leaves many leaders perplexed. Intuitively, they know they're missing out on something important, but don't know where to begin.

This Leadership Brief will provide you an overview of four different analytic approaches that can be used to solve a business problem or set of problems. Each of these represents a stage along a data analysis continuum through which additional business insight and benefit can be gained.

To understand the power of these data analytic techniques, let's examine them in a hypothetical but very plausible situation. Imagine that you are the administrator for a large health care clinic with a couple hundred employees and a few thousand patients. It has come to your attention that patients breaking appointments is perceived as a significant problem by your staff.

lifescale analytics | analytics + insights for life science

Descriptive to Prescriptive:
Accelerating Business Insights with Data Analytics

Descriptive Analytics

Descriptive analytics is the process of describing quantitatively what can be measured about a business domain. What has happened? How many? How much? Is this changing over time? The objective is to quantify, track and report what might have previously been only a vague qualitative sense for how things are going in business operations.

To move beyond mere anecdotal observations, you assign a staff person to configure a report from your clinical management system. They create a report showing that over the past year 10% of patients don't show up and don't bother to cancel their appointment, while another 18% cancel. Your clinic's controller uses these figures to estimate a monthly cost of about $30,000 in the combination of lost revenues and elevated costs. With these simple steps you now have a grasp of the magnitude of the problem, how often appointment breaking actually occurs and its financial impact.

Discussions with your staff lead to additional questions. Can we separate people who cancel with adequate warning from those who cancel with less than a 2-hour notice? Is appointment breaking a steady problem or is there a trend over time? Because the system doesn't easily answer these questions with point-and-click reporting, you need to find someone who can directly query and filter data at the database level. Fortunately these skills are fairly common and you identify someone within your organization with the required database query skills.

With these queries you discover that the 18% of patients who cancel break down into 8% giving more than 2 hours notice, and 10% cancelling with less than 2 hours remaining. You also find that appointment breaking is trending up over the past year, validating the concerns raised by staff members.

Working from the Peter Drucker adage that "what gets measured gets done," you decide to track appointment breaking using a data dashboard (graphs) on the clinic intranet that shows the previous day, a 7- and a 30-day average of appointment breaking behavior.

Finding that your clinical system doesn't directly support a web-enabled data dashboard, you start searching for vendors. After looking through a dozen or so products you narrow the field to iDashboards, LogiXML, Cognos, Juice Analytics, and HighCharts. After demonstrations and some

testing you select one of the lighter-weight, less expensive options. With a bit more programming of database queries you are able to automatically run the analysis each night and update the graphs on the dashboard, giving your staff a current reading on appointment breaking.

This process has whetted your appetite for understanding operations from a data perspective and you decide to invest in capabilities that enable interactive data exploration. Given the higher costs of these tools and the load on operational systems you decide to restrict this query and analysis capability to yourself and two staff members. You narrow the search for tools to a subset of more visually oriented data analysis tools including Tableau Software, JMP from SAS, and Spotfire from TIPCO. After installing and connecting one of these tools to your patient and scheduling data sources you start to interactively pose questions and explore your operational data using a variety of intuitive visualization techniques.

Without leaving the descriptive analytics stage of the continuum you experience some real business benefits. You gain a more concrete understanding of what has and is happening in the business. Decisions become grounded in a measured and counted description of reality. The act of regular measurement and public reporting, signals your leadership priorities, and clinic staff begin to optimize their activities to improve those metrics. But you're not done.

Diagnostic Analytics

Diagnostic analytics looks deeper into what has happened and seeks to understand why a problem or event of interest occurs. How do various measurable events and actions in the focal domain relate to each other? While it may start with bivariate relationships it progress into development of multi-variable explanatory models.

Once your health care clinic has accurate ongoing metrics on the magnitude and trends of appointment breaking, you naturally start to wonder why patients break appointments. Diagnostic data analytics seeks to uncover root causes and generate data-supported explanations. Your staff-members start observing no-shows and cancellations more closely just because this is being actively tracked. Many of them start developing ideas/hypotheses about the patterns and causes of appointment breaking.

After integrating data between the appointment system and the master patient records and enriching the data through merging with external geographic datasets you are ready to explore various hypotheses. You use the visual data analysis tools to explore and discover a number of bivariate (one-to- one) correlations including:

- Those who have previously broken an appointment are more likely to do it again
- Those under 30 years old are more likely than older patients to break appointment without
- notification
- Patients over 60 are more likely to cancel an appointment less than 2 hours prior to the start than
- other age groups

- Appointments set-up more than three months previously are more likely to be broken
- Patients from lower income neighborhoods are more likely to break appointment
- Friday appointments are more likely to be broken than other days of the week
- And more ...

To obtain a more nuanced view and coherent explanation you decide that your clinic should have an integrated model of how all these factors impact appointment breaking. Is each relationship independent of the others? Or, do these factors overlap, add together, interact or maybe counteract one another in explaining appointment breaking?

This initiative to build explanatory models takes you into the realm of traditional statistical software such as R from Revolution (or the open source version), SPSS from IBM, SAS's suite of products, Stata, KNIME or any number of other options. At this point you bring in a consulting statistician to help choose appropriate tools and work with you and your staff to develop statistically sound explanatory models.

Once these models are developed and shared, benefits begin to accrue. Competing hypotheses are set aside and your staff is able to converge on a common set of data and explanations. Even prior to strategic decisions or changes in processes, the conversation evolves. Rather than debates about cause, the talk moves toward potential solutions. "Now that we have this information, how can we target reminders or other interventions to reduce the odds that people in that circumstance will break appointment?"

Predictive Analytics

Predictive analytics is focused on answering the question "What will happen next?" The process formalizes the output from backward looking explanatory model building into an algorithm that combines current observations into predictions of what will happen for a problem of interest.

Your health care clinic now has reliable and highly accurate prediction of the probability that any given appointment will be cancelled or broken. You have found that these predictions can be made two days in advance, giving you a bigger window of time for intervention. While incremental value is achieved in each stage of the analytical continuum, your biggest pay-off for all this data analysis and testing is in this latest stage. It's here where systems are integrated and programmed to generate automated process modifications for those most at risk of breaking appointments. You carefully evaluate whether to create a custom integrated solution or purchase software that manages the predictive scoring and the prescriptive triggers. Your team looks at packages from KXEN, AYATA and IBM. You end up investing in one of the packages that enables ongoing optimization and modification of the prescriptive action triggers. With some external assistance this software is integrated with your clinical systems.

Your predictive accuracy is high but caution dictates that all patients will receive one automated email or text reminder three days prior to an appointment. Now your analytics come into play. Those

patients at medium risk of appointment breaking receive an automated voice reminder the evening before the appointment and, those at highest risk receive a live call from a staff person the day before reminding them of their appointment. This allocates valuable staff time for live calls to patients at substantial risk of appointment breaking. And, because the assigned staff members are not burdened with the rote task of calling and talking to everyone, they have more time to answer questions from the target patients, remind and reinforce the necessity of the appointment and possibly even help solve transportation issues when necessary.

You decide to introduce your process interventions one at a time in order to measure the individual impact of each process change. This systematic approach enables you to optimize your reminder and assistance interventions in a manner that minimizes intrusiveness into patient's lives while still maximizing the odds that they will appear as scheduled for their appointment.

These data-driven and tested interventions start to drive down the rate of missed appointments, and it is clear that over time the savings more than cover the expense of the tools, staff-time and external expertise. In addition to revenue recovery and minimizing waste, your clinic also experiences soft benefits. Fewer broken appointments create a more consistent, less chaotic flow of work through the day. And, the experience of using analysis to improve this appointment breaking problem emboldens you and your staff to analyze and improve other workflows and patient interactions.

Prescriptive Analytics

Prescriptive analytics takes predictive models, integrates them with automated operations and relies upon the models to act or intervene in processes to obtain a more predictable, controlled and improved outcome.

The sets of relationships and the overall explanatory models developed during the diagnostic analytics stage are accompanied by estimates of accuracy, but the only way to really know how accurate those models are is to put them to the test with new appointments. Since you have in mind the final prescriptive stage which involves automatically triggering work processes and patient communications based on the predictions of these models, it is prudent to know just how accurate they are.

Your explanatory models must be converted into an algorithm that integrates data from the scheduling system, the patient record system, and external sources and uses those combined inputs to calculate final probabilities for each appointment. Will the appointment be kept, cancelled with more than 2 hours notification, cancelled less than 2 hours notification, or broken without notification? After several weeks of generating nightly predictions and comparing to the next day results, you and your staff have a good measure of the accuracy of the predictive algorithm. Now it's just a matter of iteratively adjusting and testing the model until a level of accuracy is achieved that justifies prescriptive interventions.

This iterative adjusting and testing naturally leads to machine learning software. You evaluate adding modules to the statistical analysis package selected in the previous stage or selecting

machine learning specific software such as KXEN, RapidMiner or WEKA to find the most predictive combinations of inputs regardless of whether those combinations support easily understood explanations. Like the development of explanatory models, the use of these powerful tools are beyond the experience and training of internal staff so you bring in outside help that has the requisite training and deep experience in the art and science of machine learning techniques.

Your primary business benefit at this stage is a refined confidence in the accuracy of your models. Not only do you have plausible explanations for why some patients are more likely to break appointment, but you also understand just how predictable this behavior is. Through iterative development you have achieved strong predictive accuracy. You're now ready to take the next step to efficiently target resources on a small segment of the patient population. This concentration of resources will yield better results and, almost as important, avoid unnecessary reminders to patients who are highly unlikely to break appointments.

Parallels to Your Situation

Your situation is most likely not appointment breaking in a health clinic. However, most executives can identify parallels within their own business. Maybe for you it is customer turnover and losses, lab processes with high error rates, material waste in manufacturing systems or inaccurate demand forecasting. These types of problems and multiple others are ripe for data analysis leading to better processes and decisions.

Lifescale Analytics believes that each type of analytics along the continuum can yield significant value to your business. Some problems are best served with a reporting query that feeds a dashboard providing simple descriptive report. Other business problems have such significant costs or revenues associated with them, that even modest improvements to predictability, or prescriptively triggered process automation will yield noticeable business benefits and a strong ROI.

We can help you identify situations where data analytics can make a significant business contribution. We are committed to appropriate levels of analysis. Bigger is not always better. However, when more analytical complexity is justified, we will assist you in selecting appropriate tools, train your staff and provide targeted expertise around analysis, model development and integrating prescriptive models into automated business processes. Our job is to find ways to accelerate your insight from bioresearch related data, leading to more accurate decisions and a greater impact of your product and services in the market.

