Paper 3621-2019

Measuring Analytic Performance in Financial Terms for Executives

Ken Pikulik, Teradata Corporation

ABSTRACT

Data Scientists are by their very nature good at understanding the details of information and finding the connections between often unique and unrelated data points. The value of what they do is in high demand and often goes unfulfilled because requests overwhelm their ability to fulfill everyone's desires. New analytic tools that support greater speed and performance by eliminating data movement, simplifying testing, and streamlining the move to production can provide relief for this problem but building the business case for new analytic tools is not as easy as building the business case for a fraud or churn solution. As a result, Data Scientists need to be able to communicate why an investment in new analytics tools is necessary in financial terms that executives will understand and approve. This paper focuses on enabling you to do this.

INTRODUCTION

The value of analytics is undeniable. Analytics tools are embedded within every area of the organization and the appetite for more sophisticated analytic solutions far outweighs an organizations ability to delivery. Feeling the pains of this problem are the data scientists who are tasked with finding new ways to meet the growing needs of the organization.

The challenge data scientists face is working with an analytics infrastructure that hasn't evolved to supporting the increase volume of data, the sophistication of the analytic models and the speed at which organizations want answers.

Compounding this problem is that the value analytics delivers is often aligned with a business case related to a specific business problem; not the analytic processes themselves. These business cases seldom include funding for or even acknowledge that specialized tools and skills are required to produce the desired result.

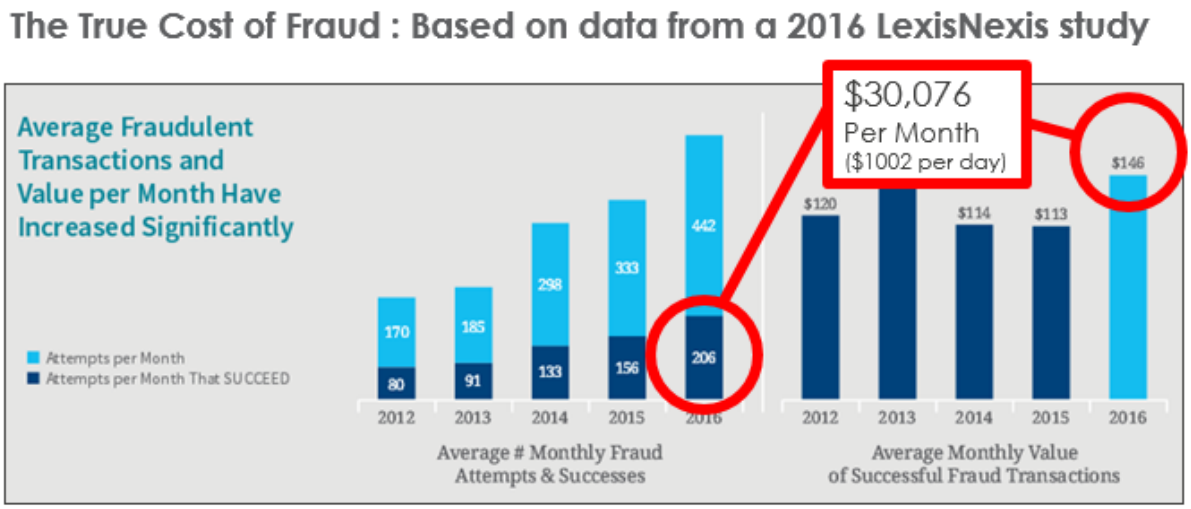
Further complicating the problem is a backlog of work created by the growing desire to produce analytics across the organization that prevents data scientists from developing and documenting their own business case for improving their analytic tools.

As a result, new investment in analytic systems is often left unfulfilled until organizations start to experience decreased performance or an inability to provide executives with rudimentary knowledge of their business performance.

To overcome these problems, data scientists need to arm themselves with the data necessary to promote their true worth and how an investment in the right tools can improve the performance of the entire organization. This starts by creating a clear picture of their impact in terms of support to the organization.

As a support organization, data scientists are actively involved in many different areas. Any investment should be weighed in how it will benefit these areas and what it will mean to the organization. This provides can be addressed in service level agreements, but only if you can measure the resulting improvement in performance.

To create a simple baseline for measuring contribution for use in this paper, we referenced a paper by TDWI from 2017, *Accelerating the Path to Value with Business Intelligence and Analytics* in which they estimate the average model takes four months to complete. A second paper by LexisNexis from 2016 entitled *The True Cost of Fraud*, estimates that the average fraudulent transaction costs a company approximately $1000 per day.



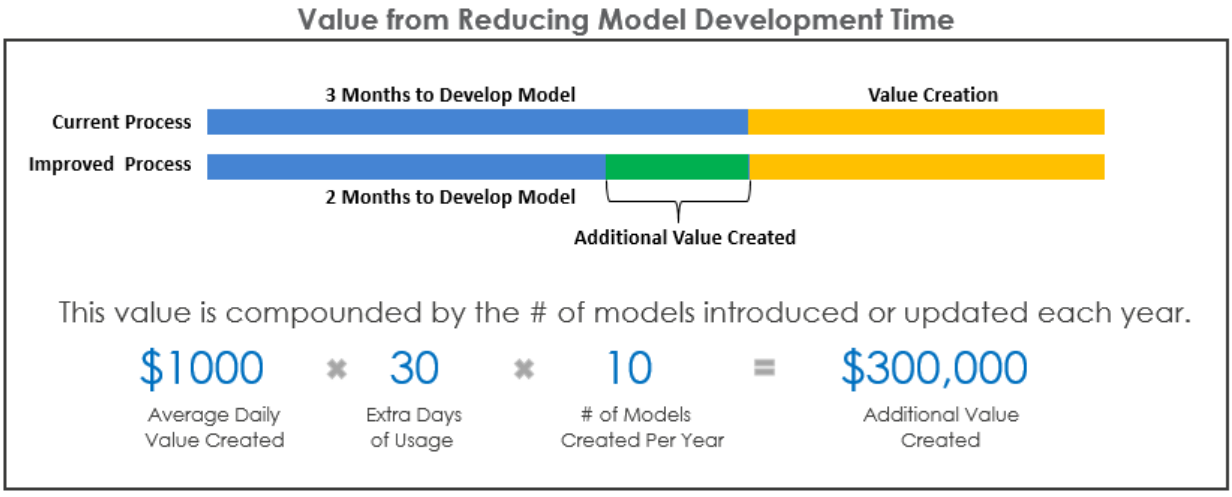
THE VALUE OF REDUCED MODEL DEVELOPMENT TIME

Executives have invested millions of dollars in analytics. Ten years ago, they knew they needed to make the investment, but didn't exactly understand the impact that those investments would have on their business. They knew they wanted the results; they knew their competition was investing heavily; and they knew the global trend was to rely more on analytics to make decisions; but they didn't always understand what they are getting when they invested in a new analytic tool. This created an initial rush to build out analytic practices. Unfortunately, that free flow of cash to analytic systems is gone and the systems built 10 years ago are still in place today. Executives are also more invested in analytics, but also can be swayed by the promise of free open source options and this is creating roadblocks to further investment without clear justification.

In addition, building a business case for new analytic tools that delivers improved analytic speed is significantly more difficult than building a business case for an analytic solution that can reduce fuel consumption or optimize inventory. Thus, your business case for a tool that will reduce model development times needs to be equally transparent in how it will result in revenue earned or money saved.

There are several ways to measure this value and each should be utilized to justify why an upgrade to your analytic systems is warranted. These can be tied to existing business cases as well as presented on their own.

Start by calculating the additional value received by the organization by implementing a new model sooner. You can look back at past business cases to see what their results were and use these numbers in your equation. In our baseline, we determined that the typical model took four months to create and deploy. If you can do this in three months that would result in an extra 30 days of model use. If your average value for that model is $1000 a day, that's a $30,000 benefit to the organization per month. This benefit is then compounded by the number of models that you produce in a single year. In the figure below, we estimate 10 for a total benefit of $300,000.



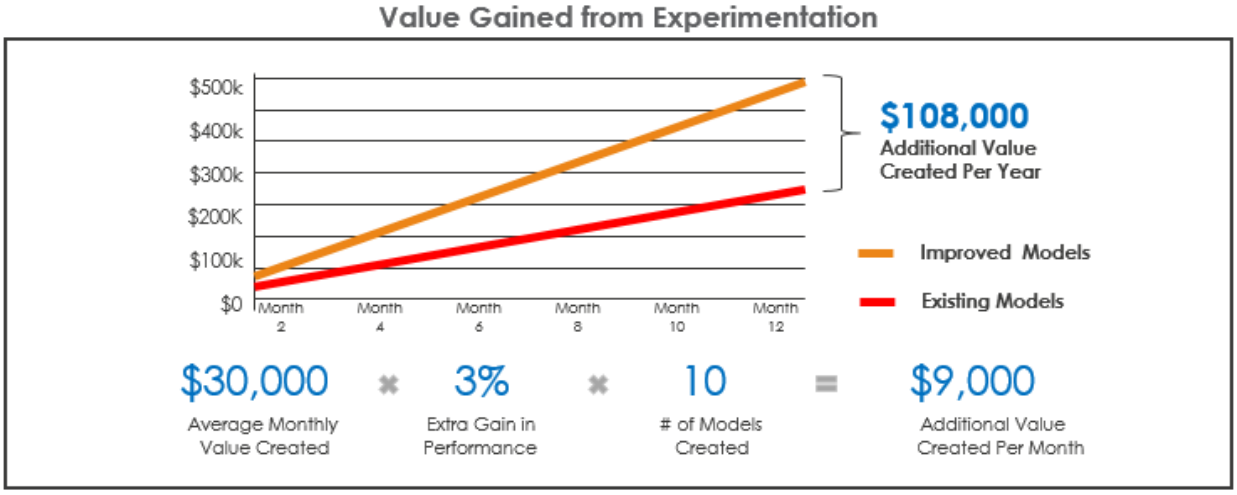
HOW ANALYTIC SPEED CREATES BETTER MODEL PERFORMANCE

Some of the complaints commonly referenced by data scientists is that they have too much on their plate, don't have time to give each model the attention it deserves and are subsequently forced to cut corners to get everything done. If this is true for you, then it may seem counter intuitive that producing more models faster would produce better results. But let's examine how this might happen.

Typically, most of the time involved with building models is in prepping and managing the data for analysis. Requesting, loading data, sanitizing and assembling the data creates a bottleneck in actually producing the analytics. At the heart of this problem is data movement (i.e. the shifting of data back and forth between production and non-production systems). Data scientists can eliminate this data movement by leveraging in-Database analytics, thereby allowing them to build models in less time. By eliminating this process, data scientists will be able to spend more time experimenting and completing more models.

When building the business case for implementing in-Database analytics, you can once again reference the ability to get models into production faster; but you can also emphasize that you will also have more time to do testing and experimentation in order to generate a better model that can create additional lift in performance.

Consider our baseline savings of $30,000 per month on a single model. If we are able to do additional testing to gain just 3% more lift on that model, we would save an additional $900 per month. If we extrapolate this same savings to the other 10 models in our hypothetical company, we would save an additional $9000 per month or $108,000 per year. These are low estimates by any standard, but still enough to provide justification for the investment.



DOING MORE WITH LESS

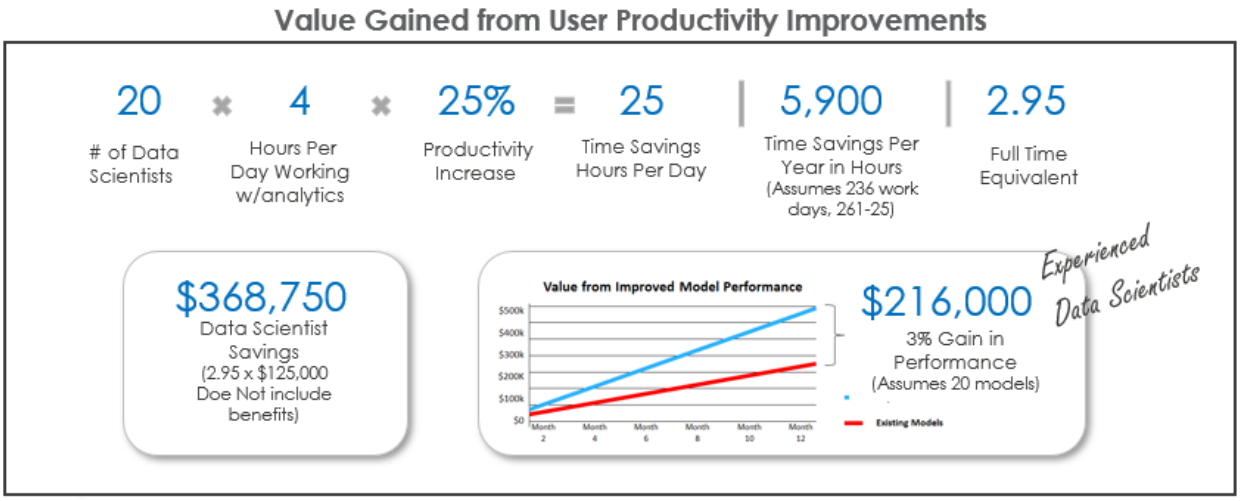
Data scientists are a rare commodity. They are in high demand, low supply and increasingly mobile. Finding and keeping data scientists with the combination of business acumen and analytic experience is even harder. As a result, companies are working directly with colleges to meet their demand for these unique skills or providing additional education to internal candidates to meet even more specific needs.

This shortage has created an opportunity for existing data scientists to demonstrate their value and gain favor with executives by taking on more workload. Once again, the ability to

produce analytics faster, do more testing and productionalize these model more quickly creates significant value to the organization by allowing you to create more, better performing models. This combined with your existing knowledge of the systems and business scenarios offers a leg up on an inexperienced data scientist.

Increasing the productivity of your existing data scientists reduces the need for additional headcount and can improve model performance. It can also help you complete more models to fulfill the organizations true potential. By using better tools, eliminating data movement and speeding up analytic development, organizations can get more from their existing data scientists. This allows them to reallocate funds to top performers and invest in newer systems to keep them happy and engaged. Even minor improvements in performance can be significant.

For example, consider the impact of a 25% improvement in performance by investing in faster analytic infrastructure. If you have 20 data scientists who spend 4 hours a day working on analytics; a 25% improvement would result in 25 hours per day savings or 5,900 hours per year. That's the equivalent of almost three fulltime headcount per year. At an average salary, not including benefits of $125,000 per year, that's a savings of over $368,000.



CONCLUSION

Building the case for new, advanced analytics systems that can support faster analytic performance doesn't have to be hard. It simply requires the understanding of how these systems impact the performance of the organization now and in the future. It requires demonstrating how being able to work faster can help you increase revenue. This will allow you to show justify an investment in tools that enable you to eliminate data movement, do more testing and increase the productive of your current data scientists.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Ken Pikulik
Teradata Corporation
919-607-4190
Ken.pikulik@teradata.com