Paper 193-2012

Lean and Mean: Using Lean Six Sigma to Improve Your Analytic Processes

Robbie Reynolds, Life & Protection a division of Transamerica

Rachel Alt-Simmons, SAS

ABSTRACT

Many businesses have increased their analytic sophistication, with analytics becoming more pervasive in business processes. With the increase of analytic maturity, investments in the people, methodologies, and technologies needed to effectively manage those processes have lagged. Many analytic teams are tied up in managing production analytic processes, and often lack a framework for ensuring process quality and efficiency. In this session, we help analytic teams better manage these processes, as well as reduce errors and cycle time through a lean Six Sigma framework. The Life & Protection division of Transamerica puts the framework to work in making significant improvements across its marketing campaign optimization process.

THE NEED FOR IMPROVEMENT

As competitive pressures increase the need for organizations to master analytics, internal analytic teams have increased their statistical sophistication, but are struggling to productionalize their insight. Many analytic groups find themselves managing operational processes with little help from their technology partners, while the demand for analytics, as well as the data volume and need for speed-to-insight, grows within the organization. Over time, the processes can grow so complex that it becomes difficult to identify the root causes when the process breaks down; and a breakdown in the process can set the team back days if not weeks as resources scurry to find and fix the problem. How do we find and identify problems from the start? How do we free up critical analytic capacity? Lean Six Sigma.

INTRO TO LEAN SIX SIGMA

Lean Six Sigma (LSS) is the marriage of two improvement methodologies: While the roots of statistical quality control have been around for decades, the current methodology is credited as being “invented” by Motorola and popularlized by its widespread implementation at General Electric. Six Sigma is a quality improvement methodology traditionally used in manufacturing organizations, with a goal of reducing the number of defects in a process to a six sigma level (99.99966% defect free). The methodology follows the “DMAIC” methodology to define, measure, analyze, improve and control to identify defects and improve processes. Another key tenet of Six Sigma is “voice of the customer” (VOC). Customer needs and requirements are the primary driver for process improvement. While many organizations across many industries have had success for Six Sigma, companies have struggled to identify how the methodology can be applied to transactional processes. Enter Lean.

While Six Sigma’s goal is to reduce defects, Lean’s goal is to eliminate waste and improve efficiency. As defined by Taiichi Ohno, “father” of the Toyota Production System, waste is made up of eight components: waiting, overproduction, rework, motion, transportation, processing, inventory and intellect. When the two methodologies are married up, you have a framework for improving quality and reducing process lead time. The LSS methodology provides a toolkit for improving analytic lifecycle management – effectively and efficiently managing the elements (people, processes, data and technology) necessary for operationalizing predictive models from conception to deployment.

Why haven’t more organizations used LSS in the past? With any process improvement initiative, there are always going to be cultural and organizational challenges. The biggest barrier to effective process flow is the organization itself. Most organizations operate in a matrix structure. Each functional area may have its own set of performance goals and incentives; cross-organizational workstreams have been divided into narrowly defined roles and responsibilities embedded across these functional areas. You can improve a tiny piece of a process in this structure, but because teams only know their piece, they can’t influence major improvements. The workflow is so granular and divided that handoffs, redundancies and errors are almost impossible to address.

THE LEAN SIX SIGMA TOOLKIT FOR ANALYTICS

Make no mistake about it: The LSS framework is comprehensive and complex. Like any methodology, it comes with a wide set of tools that practitioners can use to hone in on a specific business issue and come up with ways of assessing and evaluating the problem. Not all aspects of the methodology apply to every problem you need to solve.

The key is finding the right tools and methodology to work in your organization. Many organizations can benefit from just using the Lean methodologies first and working into Six Sigma as they gain confidence and maturity. In this section, we'll review the DMAIC phases for process improvement initiatives.

DEFINE

Think of your improvement initiative as you would any other project. A project requires a business case: What are you solving for and why is it important to the organization? Even if you're not asking for any money, your improvement project will require executive support and internal resource time. The business case is also a critical component of refining the scope of your initiative. Typically, when organizations tackle improvement projects, their scope is too broad. Using LSS Define techniques will help you narrow the area of focus. The first questions you need to ask are:

- What is the problem affecting the success of your business?
- Who internally and externally is most affected by the problem?
- What "critical to customer" categories are associated with the problem? (i.e. delivery, cost, quality)
- Where does the problem occur?
- When was it first observed?
- What is the extent or magnitude of the problem?
- How do you know it's a problem?

The second step is to define the current process through activities such as high level process documentation, collecting "voice of the customer" (or VOC, an essential component of ensuring that you are solving the right problem for the right reason), determining the metric to focus on improving, and calculating financial benefits. The Define phase is highly iterative: Expect to go through several cycles of refining your project goal before you move to the next phase. The output of Define would include high level documentation as well as a documented business case and project charter.

In one analytic process we analyzed, there were multiple teams involved, more than 100 process steps, and the lead time from idea generation to execution exceeded 100 days – and that's if everything went as planned. An error in their workflow (which was rife with data quality issues, inconsistent quality control mechanisms and an insane amount of handoffs between teams), could result in rework that set the process back two to three weeks. The time spent uncovering the problems reduced the capacity of the entire analytic team and delayed the output of the insight. In this case, the primary metric to solve for is cycle time, a secondary metric may be related to data quality. However, VOC take help you take that a step further – interviews with the downstream customers of the information uncovered that they valued flexibility in the process over the time taken to complete it, which changed the scope of the project.

MEASURE

The Measure phase focuses on capturing and measuring activities within the process and the results. Baseline measures of the process performance are captured. In this phase, the improvement team starts to dig into the details by documenting the "as is" process in a process flow diagram. Many teams do not need to take the LSS project beyond this point because the exercise of mapping out the process may highlight many immediate areas for improvement ("quick hits"). A facilitation technique widely used in this team process mapping exercise is known as a "Lean Kaizen" event – a Kaizen "blitz" is a facilitated multi-day (dependent on process complexity) team workshop that brings the stakeholders together to map out the process and identify areas of improvement. Other useful process mapping tools include:

- Mid-level process flow diagrams
- Customer value stream maps: a process mapped from the customer's perspective
- SIPOC diagrams: a mapping of suppliers, inputs, process, output and customers
- Spaghetti diagrams: illustrate how the process might be run differently by different stakeholders

For one insurer, the SIPOC diagram process assisted with better understanding the flow of information across the organization. This understanding provided a critical input into a strategic data project, as outputs calculated by one area (actuarial) were an input into several other marketing processes. This facilitated the design of a data architecture to streamline the flow of data between the two activities.

Another key component of the Measure phase is collecting baseline process performance data. Unfortunately, many organizations (and analytic teams in particular) do not track these metrics. An example was an analytics team that ran logistic regression models over enormous data sets on a monthly basis. The cycle time was roughly 167 hours –

however, the team had not tracked the run time over time or the data volume input into the process. These metrics would help identify where the problem is: Is it architecture (server capacity), data, structure, etc.? In LSS, the project would typically stop at this point and a measurement system set up to capture this data – this may not be feasible in many cases. The Six Sigma methodology requires the quantitative analysis of data; if none is available, then the focus should remain on the qualitative assessment of the process.

ANALYZE

The Analyze phase focuses on the analysis of data collected in the measurement phase. This data is used to help understand variation in the results and the activities that produced them, along with a quantitative analysis of the cause and effect of the relationships. This phase statistically validates (or invalidates) the cause and effect through hypothesis testing. Even if you don't have the baseline measurement data, there are tools that can assist you with uncovering root causes of your problem:

- Cause and effect diagrams (also known as Ishikawa or Fishbone diagrams): this is a great exercise for your improvement team. This diagram is generated through a team brainstorming process of all the potential causes of your problem. The causes can be grouped into categories.
- Failure modes and effects analysis (FMEA) diagrams: This diagram lists each process step, potential failure modes, potential failure effects and causes, along with a ranking for severity, occurrence and ability to detect. The FMEA diagram provides a nice input into the Improve phase, as controls can be identified to improve the existing process. Again, this diagram can help your team focus on important areas – for example, if an output of a process step is to provide accurate data or analysis to a regulator, your severity levels would be high.

If you have data tied to your primary metric, the information can be statistically validated. It is important to have information about the metric. For example, in the example above, having historical run times on an analytic process would be an important baseline metric, but you need additional information, such as the volume of data, the time the process is run, the number of users on the system, etc. to isolate the true cause of the long run time.

An Information Delivery team we worked with runs hundreds of SAS® batch jobs on an ongoing basis and often saw inconsistent run time performance. Using a SAS dataset with FULLSTIMER performance stats from historical job runs, we analyzed the data and attempted to predict the probability that a given job would finish within a given time frame. Variables that were predictive of run time included job count, job concurrency, total SAS steps in job, and observations in. Using Monte Carlo simulation in JMP®, we could estimate the total run time by controlling those variables.

As illustrated in this example, we developed a list of potential causes and narrowed them down through statistical analysis, which allowed us to come up with the Holy Grail of Six Sigma: Y = f(x). As noted earlier, many organizations don't have this level of data to analyze, which is fine. The focus should instead be on analyzing the qualitative steps within your process and identifying areas of improvement.

IMPROVE

In the Improve phase, a variety of techniques can be leveraged for identifying solutions, including brainstorming and other idea generation techniques. Improve also facilitates the assessment of how realistic a given improvement idea would be to implement. As your team goes through the improvement initiative, a wide variety of solution ideas may be generated. Following are some tools that can assist with prioritizing:

- Pugh Matrix: In a Pugh Matrix, each process step is identified (rows) and mapped to a matrix of potential solutions (columns). The solutions are rated relative to that process step (positive, neutral, negative). A total score is calculated for each solution. This allows you to see where your ideas may have a positive or negative impact across the process flow, and helps prioritize high impact solutions.
- EPIC Analysis: The EPIC (ease, performance, impact, cost) assessment of your solutions is another way to validate. Each solution area is ranked from 1 to 5 on these four criteria and a score is generated. For example, you might have an excellent idea that is too expensive to implement (such as overhauling or replacing your analytic infrastructure).
- Future state value maps: Allows you to envision what your process looks like in a future perfect world. This is a great way to illustrate improvements from your previous current state value map.

CONTROL

You've made it to the Control phase, which means that you've implemented the solutions you've identified in the Improve phase. Control ensures that you stay on track post implementation. Your process improvement project formally ends, but monitoring is put in place in the form of control plans or updated FMEA diagrams with recommended action plans.

THE IMPORTANCE OF ORGANIZATIONAL CULTURE

During one LSS project execution, as part of my interview process I encountered a number of people who were concerned about the perception of negativity. But it's Six Sigma! Process Improvement! Reducing defects! A fellow Black Belt gently reminded me that a defect in manufacturing is easy to spot – it's either good or bad; in a transactional environment, it's a lot greyer – in reducing lead time, just because something takes longer, does that make it defective? Well, that's something that you need to iron out in advance with your project team. And it could be that your cycle time is okay, but your resources are going crazy meeting the deadline.

If you look at all of the templates and tools in the LSS process, they're all about finding the root cause of the defect, which means that you're looking for everything that goes WRONG, not necessarily what's going right. People obviously get very sensitive about that. My Black Belt friend used a technique in his process improvement project called "Appreciative Inquiry." It's all about getting people in a room and talk about what's going really well, what works best. They used the approach on an organization and he said the difference in energy and engagement (versus using the "negative" approach) was amazing. The session was productive, people got excited about participating, and they were able to focus on the things that worked (and through process of identification, what didn't work). Lesson learned: As with any methodology, it's critical that you adapt the methodology to work within your organization's culture. By changing your approach, you can achieve the same outcome (identifying "defects"), but in a kinder, gentler way.

GETTING STARTED DOWN YOUR OWN PROCESS IMPROVEMENT PATH

To illustrate a practical implementation of our methodology, the Analytics and Data Services (ADS) team within the Life & Protection division of Transamerica, is using the LSS techniques to optimize analytic processes and free up key resource capacity. The analytics team is tasked to focus their energy and resources on business projects with significant financial impact, while minimizing time spent on repeated production processes by increasing efficiency. Over the past 18 months the team has undertaken several LSS process improvement projects that have allowed them to achieve this goal. Following are examples of the team's LSS implementation:

EXAMPLE 1: BACKGROUND

Process changes often impact more than one group as do defects. The Analytics and Data Services (ADS) organization within Life & Protection has several components including Campaign Processing, Data Management, Print Production and Analytics. A major project to refine a key production process began with a kick-off meeting including several people from several ADS teams. It was clear from the discussions that in order to have significant impact the reworking of the process should involve multiple areas of Analytics and Data Services. In the meeting we were able to identify several potential improvements for focus. The ones highlighted below are either specific to Analytics or are jointly implemented with Analytics and another of the ADS teams.

EXAMPLE 1: DEFINE, MEASURE AND ANALYZE

The expressed purpose of the meeting was to detail the direct mail production process from lead availability through mailing and identify areas for improvement in communication, efficiency, and quality. The meeting served as the combined **Define, Measure and Analyze** stages (although some of the measurements were based on recollection, not recorded data). The two day discussion included exploration of every aspect of lead preparation by all teams within ADS. During the meeting each group explained their procedures and charted the process based on their specific working knowledge of its role in the larger flow. Interestingly in this meeting each group learned significantly more about the responsibilities of the other teams. This increased knowledge was a key component to development of several process improvements. The discussions at this meeting revealed a cumbersome process with opportunities for delay and error. Any delay or error was compounded in the effect on other teams within ADS and could even delay the actual mailing. After the meeting the entire production process flow was charted including dependencies and timelines. Several potential handoffs were identified as having potential for a default in the process. Lists of action items and follow-up items were created as a means to track the progress on next steps.

Define: One of the identified opportunities for improvement was the scoring of leads prior to optimization for product selection. At the time of the meeting, the Analytics portion of the monthly direct mail process included receipt of 2 files from Campaign Processing, execution of several data appends, scoring of all potential leads and optimization of the file for product selection. Some key identified pain points were incorrect data, last minute adjustments made by marketing, system slowness, and anything that caused a rerun.

Measure: The timing of the analytics portion was at best about 7 days and at worst 14 days, averaging around 9 business days. This variability as well as the sheer length of time caused frequent delays in the process.

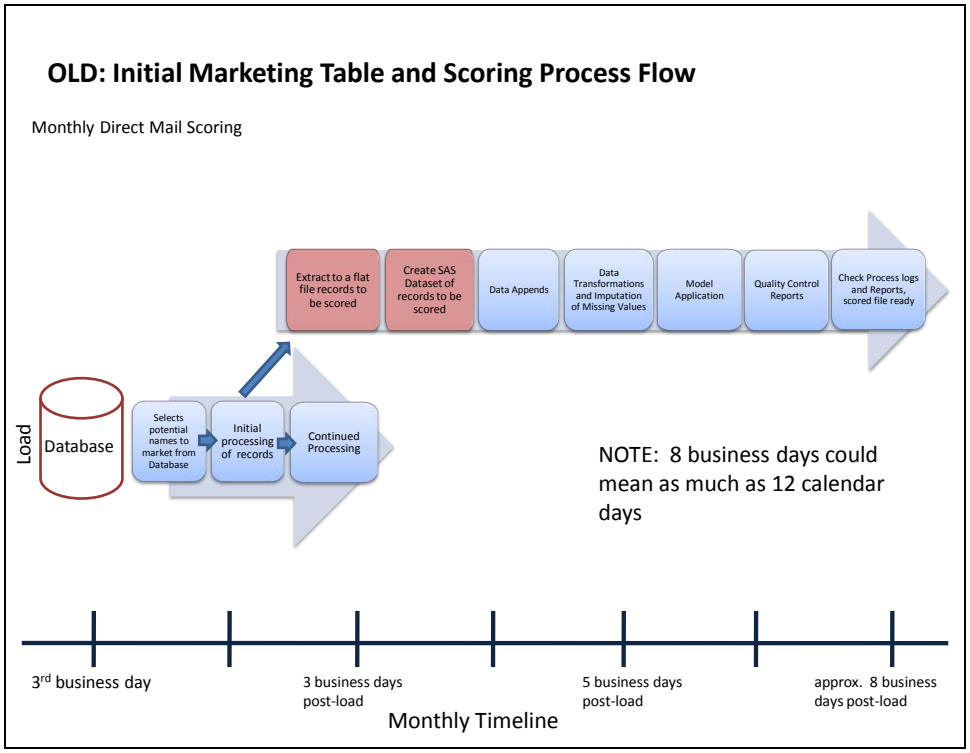

Figure 1. Original Process

Analyze: A potential for improvement was exploring whether the scoring process could be done in conjunction with the database load (a more streamlined production process owned by the Data Management team.) If this could be accomplished utilizing off-peak hours and without increasing their process time, theoretically it would reduce the Analytics timeline by 3 days. The scoring process at the time of the meeting appended several levels of data at both a zip code and individual level, scored the available 20+ million records, with more than 25 direct mail models, ran for over 12 hours and was then manually checked. If an error was found requiring a rerun, this was very time-consuming and would put the entire process behind.

EXAMPLE 1: IMPROVE

The **improvement** stage focused on the scoring and optimization processes lasted several months. During this stage we developed a set of roles and responsibilities which detailed the ownership of the new process. We also mapped out the details of adding the scoring program to the post-load process flow. It was determined that both the more than 25 direct mail and more than 50 telemarketing models would be included in the new process. This increased efficiency in an additional process making the benefit farther reaching that originally thought. The testing and roll-out phase timings were defined.

Improve: The ownership delineation is that Analytics continues to own the model code and transformations contained in the actual scoring program. The Data team owns the programs which apply the models. This parsing works well for the Analytics team because that is where the expertise in model development and application reside. It also works well for the Data team as they are much more adept at efficient program design and execution. A combined scoring program was developed including direct mail and telemarketing models. This modification included splitting the scoring into three independent pieces which could be multi-threaded for increased efficiency. These now run as part of a multi-stream automated process tagged onto the end of the established database load process. This multi stream approach allowed the scoring to be included without significantly impacting the Data team's timing and responsibilities. The whole process was tested in parallel with the current process to ensure maintenance of quality in the new configuration for both direct mail and telemarketing. Including the models for both processes increased efficiency for both types of campaigns further freeing up analytics resources for more innovative projects.

EXAMPLE 1: CONTROL

Control: In addition to the structure of the actual scoring programs and their execution responsibility, a quality control process was developed to ensure consistent results and to identify errors as quickly as possible. The quality

control process consisted of checking the completed scoring each time the program is executed. The variables used in the models are expected to have means and distributions consistent with the past several months. Variance of more than 5% causes the variable to be manually inspected and a cause of the anomaly identified. Score means and distributions are also reviewed and changes over previous months investigated. The reports used for quality control are automatically generated and checked by both the Data team and the Analytics team.

One area identified as a possible quality concern is new model installation. In order to ensure the validity of the newly installed model and to minimize the need for rerunning scoring, a clearly documented process was developed with deadlines and communication points for both teams. The scoring process with the new model(s) is tested using the previous month's program. At the end of each month's scoring, the scoring program, the models used and a sample of the data are copied into a test directory for use between then and the next scoring. When a new model is developed the programs in the test directory are utilized to score the sample with the new model and check for any syntax or logical conflicts with the other existing models. A key item to this process is the due date by which any changes must be provided to Data Management for inclusion in that months load. Last minute changes impair quality. Implementation of these quality checks and the new model installation process are the **Control** stage of our process. These ensure that we stay on track with the process and don't allow various "emergencies" to derail the improvements.

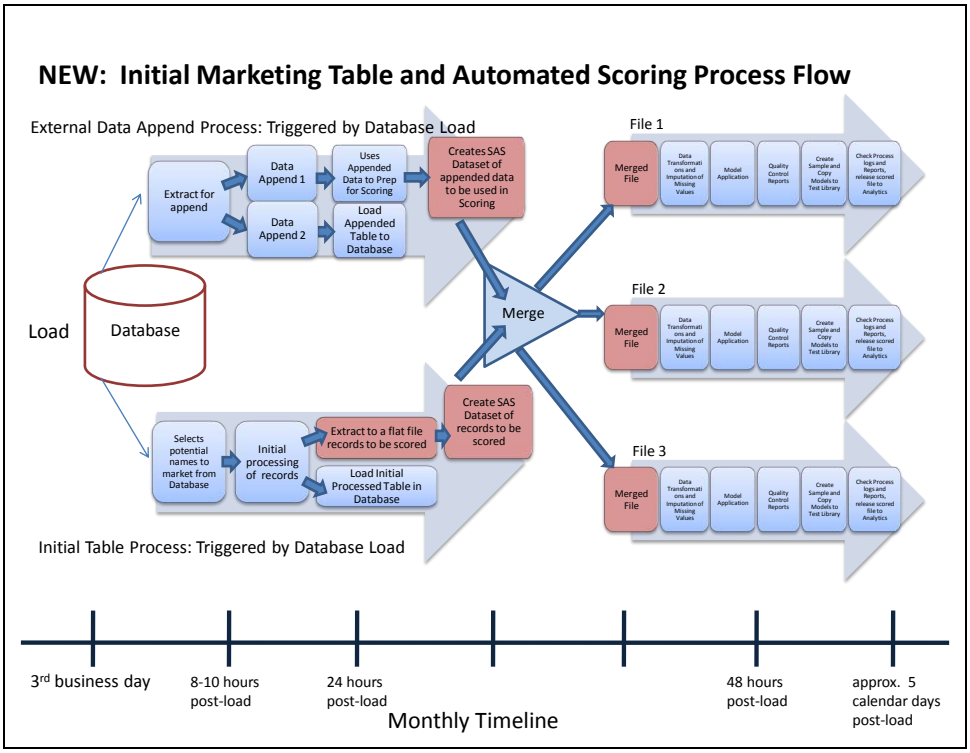

Figure 2. Improved Process

EXAMPLE 2: BACKGROUND

Another example of our application of portions of the LSS methodology is some modification made to the optimization process itself. This process remained in Analytics after the improvement detailed above. The Analytics team examined the optimization process to determine whether it would be possible to further reduce cycle time. The combination of the time savings by including the execution of the scoring programs in processes owned by Data Management along with the result of our evaluation and improvement of our optimization process has reduced the time spent during monthly scoring and optimization of the direct mail production file from about 9 days to about 3 days.

EXAMPLE 2: DEFINE, MEASURE AND ANALYZE

The **Define, Measure and Analyze** stages of this example were not formal at all. Much of this work was done in various conversations sparked from the meeting in Example 1 above. We were aware that there were unanticipated items each month which extended the Analytics timing. Late in 2010 we began tracking these difficulties to see if a

trend existed. From our research we found that marketing would frequently make changes which necessitated production modifications, last minute approvals or last minute test adjustments. Often quality issues would result, resolving these further extended the timing. Even though there was a monthly meeting with Marketing, Campaign Processing and Analytics to review plans for the campaign, there would be additional changes after the meeting. Changes were also requested once the result of the optimization was reviewed. This caused significant deviation from the expected timeline. In addition we lacked clear documentation allowing the Analytics process to be evaluated and improvements recommended.

EXAMPLE 2: IMPROVE

In the **Improve** stage Analytics and Campaign Processing decided to strengthen the due date for changes from marketing, implementing a "Go/No-Go" date. After this date, the only thing that marketing could do was to drop a product or segment. Nothing could be added or changed. After the Go/No-Go date Analytics produces a simplified grid as a based on the monthly meeting and campaign plans. The grid represents a high level view of the products and segments to be marketed. Marketing is asked to approve this grid as a final commitment that everything expected is included.

This one step helped Analytics and Campaign Processing to better plan for changes.

The need for changes after completing optimization was reduced greatly through executing the process on a sample and providing that result to marketing for review. For example because of the time it takes for the entire campaign to optimize, we now optimize on a 3% sample first. The result of that optimization is grossed up to reflect the expected campaign numbers and provided to marketing for feedback. Once they approve the optimization result, the process on the population is started. The purpose of the sample is that it identifies any concern at least a day earlier than they would have been seen using the old process. Since marketing can quickly review the expected campaign volume, they make any last minute decisions (such as to eliminate a specific product) before the 6-9 hour production optimization job runs, eliminating time consuming re-runs.

To address the documentation concern we used the SIPOC document as a template for a detailed process document (this is not its original purpose but so far it has worked well). It lays out every step of our optimization process in great detail including the supplier, inputs, processes, outputs and customers. This allowed us to see potential problems and put steps in place to minimize or eliminate quality issues and increase efficiency. It also provides needed clarity for cross-training. This documentation is reviewed monthly and any needed adjustments made.

EXAMPLE 2: CONTROL

To continually track issues and identify areas for future improvement, we developed a log where we document any anomalies in the campaign. As part of our **Control** phase this log is monitored and additional quality checks implemented as needed to ensure the process improvement continues. We also review the process documentation regularly making adjustments if needed. Having a solid process document allows us to focus energy on cross-training to ensure business continuity.

These were examples of some recent process improvements which have decreased the direct mail timeline from an average of around 9 business days to about 3 business days. Key to this improvement is increased understanding of the process and doing the right things to prepare for the upcoming campaign. The DMAIC methodology, while not followed precisely, is the basis of the identification of challenges and development of solutions to meet those challenges producing a better end result for all teams. It is important to note that the improvement process continues, we strive constantly to save time on any repeated process so that that time can be better spent innovating. Since we created the process that we manage, it can be very difficult to identify opportunities for improvement. The LSS tools are an excellent way to analyze those processes and create a leaner process freeing up time and resources for greater things.

The world we have created is a product of our thinking; it cannot be changed without changing our thinking.
~Albert Einstein

RECOMMENDED READING

Complete Idiot's Guide to Lean Six Sigma, by Breakthrough Management Group
Lean Thinking, by James Womack and Daniel Jones
Process Mapping, Process Improvement and Process Management, by Dan Madison
Leading Change, by John Kotter
JMP Start Statistics: A Guide to Statistics and Data Analysis Using JMP, by Sall, Creighton, Lehman
Visual Six Sigma: Making Data Analysis Lean, by Cox/Gaudard/et al.

CONTACT INFORMATION

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

Robbie Reynolds
Life & Protection a division of Transamerica
robbie.reynolds@transamerica.com

Rachel Alt-Simmons
SAS
rachel.alt-simmons@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.