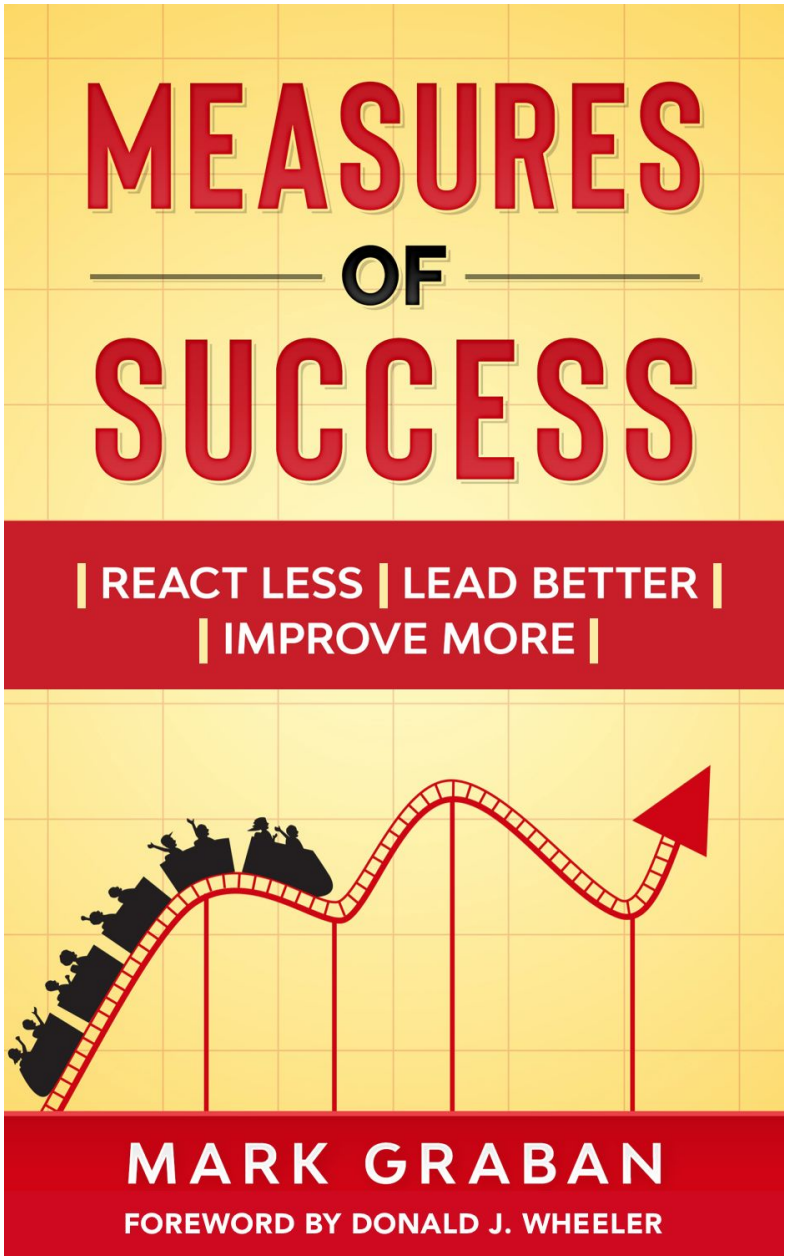
Measures of Success (Sample)



Mark Graban

Measures of Success:
React Less, Lead Better, Improve More

Notes to Early Readers

Thanks for buying this early "in progress" release of the first part of this book, which will be completed by June 2018. See more about the "Lean Publishing" approach in the Preface.

You'll be notified when updates and new content can be downloaded.

My intent is a book that has a broader audience than just the "Lean" or "Lean Healthcare" communities. The intended audience is managers, executives, business owners, and continuous improvement professionals.

The book's web page: https://www.markgraban.com/measures-of-success/
(You can also access that page via the domain www.measuresofsuccessboook.com)

The LeanPub.com order page: https://leanpub.com/measuresofsuccess

What's New in This Release

Changes in the June 11 release include:

- New foreword written by Donald J. Wheeler, Ph.D.
- Chapter 3: New material on monthly vs. weekly charts
- Chapter 5: Draft material added on connecting charts to improvement
- Chapter 7: Additional case examples
- Chapter 8: Draft chapter on Process Behavior Charts vs. other methods
- Chapter 9: Draft material on workplace case examples
- Chapter 10: Draft material on getting started with these methods

Table of Contents:

Also By Mark Graban

Lean Hospitals: Improving Quality, Patient Safety, and Employee Engagement
(3rd edition, 2016) — Available in eight languages — Shingo Research Award

Healthcare Kaizen: Engaging Front-Line Staff in Sustainable Continuous Improvements
(co-authored by Joe Swartz) — Shingo Research Award

The Executive Guide to Healthcare Kaizen: Leadership for a Continuously Learning and Improving Organization (co-authored by Joe Swartz)

Practicing Lean: Learning How to Learn How to get Better… Better
(anthology edited by Mark Graban, with 15 other authors)

Acknowledgements & Dedications

Many thanks to my editor, Cheryl Fenske, and my book coach, Cathy Fyock. I'd like also to thank early readers who gave valuable feedback and helped make this a better book. They include Mike Stoecklein, Shrikant Kalegaonkar, Warren Stokes, Dan Markovitz, Harvey Leach [and others TBD].

This book stands on the shoulders of quality improvement legends I have studied, including W. Edwards Deming and Donald J. Wheeler. I admire their willingness to challenge the management status quo and for the difference they have made.

Thanks to my loving wife and partner, Amy Graban, my parents, Bob and Marlene Graban, and my in-laws, Charlie and Debbie Gowder. All of them have supported my learning and writing over the years. I'd also like to thank the countless number of people who have taken time to teach and mentor me over the years.

About the Author

Mark Graban is an internationally-recognized consultant, published author, professional speaker, and blogger.

He builds upon a deep education in engineering and management with practical experience working with executives and frontline employees in multiple industries to synthesize and practice methods including Lean management, continuous improvement, statistical methods, and people-centered leadership approaches.

Mark's motivation is to humbly help others learn how to improve and sustain performance. In his healthcare work, this means improving the quality of care and patient safety, while also reducing cost and improving the workplace experience. Across multiple sectors, goals also include improving the customer (or patient) experience, to help the development of leaders and employees, and to build stronger, more adaptive organizations for the long term.

He has learned, practiced, and taught these methodologies in settings including manufacturing, healthcare, and technology startups. Working independently since 2010, and in partnership with other consulting groups, Mark enjoys working with organizations that are looking for better ways to improve, with leaders who are willing to lead that charge.

Mark is also a Senior Advisor for healthcare clients with the firm Value Capture. He works as a Senior Advisor to the technology and software company KaiNexus.

Mark earned a Bachelor of Science in Industrial Engineering from Northwestern University as well as a Master of Science in Mechanical Engineering and an MBA as a Fellow in the MIT Sloan Leaders for Global Operations Program.

To learn more, you can visit Mark's website at www.markgraban.com.

Letter to Early Readers

The main goal of this book:

- This book presents a practical, simple method ("Process Behavior Charts") that separates "signal" from "noise" in our metrics (a.k.a. "performance measures"), so we can learn when and how to evaluate and respond to our metrics appropriately over time. By using this method and overreacting less often (or reacting in different ways), we can stop wasting time, and start improving more. This will also reduce frustration in the workplace and boost performance through higher morale and increased engagement.

Why Write This Book?

Why am I writing this book? I'd like you and your organization to be more successful, to improve more, and to be less frustrated in your improvement efforts – using your limited time more effectively and with greater impact.

I'm fortunate to have learned and used some methods that enable better management decision making and to be able to share them with others. These methods, grounded in the scientific method, help us be more data driven and to make better decisions that are based on facts instead of opinions, hunches, or feelings. Doing so will reduce the amount of time we spend chasing our tail, if you will, in the name of improvement. We can turn data, metrics, and charts into knowledge, wisdom, action, and better results.

I wouldn't call this a "statistics book." I'd consider it to be a management book that happens to draw upon a few simple statistical concepts and methods. These methods aren't complicated; they're just different. Anybody can use them without a Six Sigma "Belt" or a statistics degree. Many have been scared off by statistical methods because they perceive them to be too complicated or irrelevant in the real world. I hope that you'll dip a toe into the water here, with this book, to explore these methods and try them. I'll try to make it fun (or at least interesting).

I'm writing this book in response to some common challenges and frustrations I've seen in various workplaces over the past two decades. Leaders are under pressure to improve. They are tracking various performance measures (or "metrics"), but to what purpose? Or, they add metrics (whether forced upon them or chosen) and expect that cycles of reviewing and reacting to them will yield better results.

But, many of the methods that have been accepted as conventional management practices have their faults and their shortcomings. I don't blame you, the reader, or your leaders for their current practices. It's easy to fall into the trap of "the way we've always done it" or "this is how I was taught" – and this includes our management practices. There are better ways, once we can agree that current approaches are lacking. We can do better. We can improve more. With apologies to the creators of "The X Files," the methods are out there.

"Lean Management" is Better, But Can be Improved

[Note to early readers: I suspect many, but not all, readers will be looking at this in the context of improving their "Lean daily management system" or whatever they call it. So, I'm not sure it's necessary to address that in the introduction.]

Since starting my career in 1995, I have been an advocate for progressive management approaches often referred to as "Lean manufacturing" (or, more broadly, "Lean management"), which is based on the Toyota Production System. Since 2005, I have focused primarily on helping others adopt and adapt these improvement and management approaches into healthcare, where improvement and transformation are greatly needed.

Lean management methods are powerful and widely applicable. Lean methods have been used in startups, government, service industries, and virtually every industry and setting. But, in this context, many well-intended methods that are taught under the banner of "Lean" end up causing new problems.

For example, replacing an organization's lagging monthly metrics with so-called "real-time" daily or weekly metrics might create more opportunity for overreaction and blame instead of leading to a better understanding about how to improve the systems and processes in a way that would create better results.

When organizations aren't performing at their best, the Lean philosophy teaches that the best efforts of individuals might not be enough to guarantee success. We often see great people being defeated by their broken workplace and system. For example, think of the famous video of Lucy and Ethel in the chocolate factory. They're working hard and they just can't keep up. A bad system defeats good people.

Far too often, executives and managers are just like Lucy and Ethel. They're under enormous pressure to improve. Their organizational scorecards or dashboards have many numbers that are colored red to indicate that they aren't meeting their targets. Bad results and problems are flying at them faster than they can handle,

like chocolates on a conveyor belt. So, everybody is working hard to improve — but the management methods we've been taught might not be enough, either.

Lean teaches us how to distinguish between activity that adds value to the customer and activity that's just a bunch of wasted motion. Not all motion or effort is useful. We might take pride in being busy (listen to how people brag about it). But, a lot of management activity gives the impression of action but doesn't lead to real progress.

The methods in this book can be used to build upon "Lean management system" practices and approaches, but these are methods that can be used without a formal Lean methodology. These methods are relevant and helpful in any organization that has metrics and needs (or wants) to improve.

I've seen many organizations with very well thought out and standardized "huddle boards" (made from whiteboards or bulletin boards) in different departments and sites. The layout of the board, the structure of idea or suggestion cards, and other details tend to be very consistent. Team huddles (sometimes called a "standup meeting") might follow a very consistent agenda.

But, there might also be a highly variable approach to the way metrics are being displayed and evaluated – different types of charts (or a lack of charts) and highly variable approaches to how people react to the metrics. That suggests to me that organizations are not teaching best practices for displaying and responding to changes in metrics, one of the problems that this book is intended to help address.

About "Lean Publishing"

[Note to early readers: This material won't remain in the final book.]

As an early reader of this book, thank you for buying a version of a book that is not yet completed. Thanks for taking the time to read it. You'll be notified via email when new chapters and updates are made available and you'll be prompted to download the updated version.

I realize this might be your first exposure to the "Lean Publishing" methodology that I'm using here. I invite you to read the free, short "Lean Publishing Manifesto" written by Peter Armstrong, one of the co-founders of LeanPub.com, the platform that I'm using for early publication of this book.

Their definition of "Lean Publishing" is:

> "... the act of publishing an in-progress book using lightweight tools and many iterations to get reader feedback, pivot until you have the right book and build traction once you do."

Lean Publishing appeals to me because it builds upon proven methods from the "Lean Startup" methodology of Eric Ries. Before the Lean Startup methodology became popular, the typical approach for a technology startup was to get funding from investors and then work in "stealth mode" secrecy, often for years, to build a product or service that was then launched into the market.

Far too often, these startups learn, after launch, that there is no market for the product they spent so much time on. Or, they've released a product that doesn't fit the market because it was based on old requirements from two years ago when work began. The traditional approach to entrepreneurship often results in a big batch release of a product, which turns out to be very risky. And, the broader Lean methodology teaches us that large batches are generally bad.

The Lean Startup or Lean Publishing approach mitigates risk, in part, by encouraging continued iteration on the product based on continuous feedback from customers and the market. The LeanPub.com idea of "publish early, publish often" means making a book available for sale while it's still an in-progress work. That means releasing a few chapters that early adopters can read and, hopefully, find benefit from. The Lean Startup movement refers to a "Minimum Viable Product" as that initial entry into the market that's minimal (meaning it's missing some features, in this case, chapters), but it's also ***viable*** in providing value and being worth purchasing.

Early-adopter customers of a company's web-based software, or an author's book, choose to purchase the product, as imperfect or incomplete as it might be, because they have a compelling problem to solve or a particularly strong interest in the product. Early adopters believe in the vision of the product and are excited about the opportunity to give feedback to the creator of the work. Early adopters form a community that gets to give input into forming the final, fully completed book. The author benefits from feedback, constructive criticism, and the chance to write a better book than what would be written in a vacuum.

As the Lean Startup movement reminds us, "minimum viable" doesn't mean poor quality. The content you're reading here has been professionally copy edited (but please let me know if you find a typo or error since editing is never perfect). Of course, more content will be added to create a complete book. Also, the existing content may be revised, re-ordered, clarified, or expanded over time. Early versions might have very rough illustrations or charts that will be revised or redone by a professional illustrator before final publication.

As an early buyer of the book, you'll be notified of updates and the opportunity to download new versions of this book as they are released. I realize this creates some hassle for you, but that often comes with the territory of being an enthusiastic early adopter of software or a book. Again, thanks for buying early.

Eventually, this book will be available through the Amazon Kindle store and as a paperback book.

Please Share Your Feedback and Thoughts

If you have any feedback, please email me any time at mark@markgraban.com. If you have stories or examples that could be included, please let me know. Those stories can be used anonymously or with attribution, as you wish. You can schedule a phone call or web meeting with me.

While I have successfully used these methods in my own work and with clients, I do not yet have any organization-wide success stories related to the broad adoption of these methods. I am particularly interested in your stories about the benefits of using these methods. My hypothesis is that these methods will save time, reduce frustration, and improve performance in your organization.

Finally, if you like the early material, please tell others about the book. Anything you can do to help promote the ideas in this book would be greatly appreciated, whether that's sharing on social media (particularly LinkedIn) or through professional groups or societies in which you participate.

Thanks!
Mark Graban
www.MarkGraban.com
Mark@MarkGraban.com

Foreword

It is an honor to have the privilege to write a foreword for this book by my friend Mark Graban. He has created a guide for using and understanding the data that surround us every day. These numbers are constantly changing. They go up and they go down. But what do these changes signify? Some of the changes in the data will represent real changes in the underlying system or process that generated the data. Other changes will simply represent the routine variation in the underlying system when nothing has changed.

So the problem with interpreting the data is deciding when a change in the data represents a real change in the system and when it does not. As a result, there are two ways we can get things right, and two ways we can get things wrong. We get things right when the system changes and we interpret the change in the data as a signal of that change. We also get things right when the system does not change and we interpret the change in the data as being merely noise.

We get things wrong when the system does not change but we interpret the data as representing a change. This is the error of interpreting noise as a signal. We also get things wrong when the system changes and we interpret the data as representing no change. This is the error of missing a signal.

Since everyone understands the problem of missing a signal, we usually seek to avoid this error by interpreting every change as a signal. The numbers go up, and that's a signal. The numbers go down, and that's a signal. The numbers stay the same, and that's a signal too (because we expected them to go up or down). As we run around crying that "the sky is falling" we do occasionally get it right, but more often than not we are simply interpreting noise as a signal.

Mark presents the antidote to this disease of interpreting noise as signals. And that is why everyone who is exposed to business data of any type needs to read this book.

Donald J. Wheeler
Knoxville
8 June 2018

Introduction

As I've worked and talked with leaders from organizations in various industries, I hear a lot of common challenges that aren't easily articulated. One leader expressed a vague underlying concern:

> "We're trying really hard, but this approach [of managing]… I don't know; it seems not to be working."

For all of the attention placed on performance, goals, and metrics, many leaders haven't been taught the most effective ways to gauge progress; they also haven't been taught the best ways to improve performance. Many of our common management practices, as taught in business schools or passed down from generation to generation, actually interfere with improvement.

Accepting a better way requires first that we recognize problems or shortcomings with current management practices. Below are some examples of the challenges and common themes that I have seen across some of the jobs I've held over nearly 25 years.

> **Blame:** Right after college, at General Motors, I saw leaders use hourly and daily production metrics as an excuse to yell at employees and supervisors when targets weren't being met. In their view, bad results came from people not caring enough or not trying hard enough. The root cause of organizational underperformance is rarely bad people. Blaming individuals for poor results is, unfortunately, a pattern that I've seen in many companies in different industries, including many hospitals. But, there is a better way.
>
> **Targets:** After graduate school, when working at Dell, I saw a heavy focus on daily production numbers. For some factory-level incentive programs, quantity was the only thing being rewarded. This sent a powerful (and I'll assume inadvertent) message that quality wasn't the top priority. At a corporate level, some of our growth and productivity metrics were compared against a goal and last year's number. Displaying metrics in this simplistic way didn't do much to shed light on trends and the way our improvement efforts were affecting those numbers. Thankfully, there is a better way.
>
> **Lagging Data Prevents Improvement:** In many industries, I've seen teams presented with monthly metrics that are posted on a bulletin board two or

three months after the fact. These time delays make it impossible for anybody to answer management's question of "Why was that a bad month?", if that's even the right question to be asking.

Failure to Understand Variation: One leader I talked to said, "We're not very good at using data to inform daily decisions and problem solving. We're trying to build that culture. We need to push daily metrics." As we'll discuss in this book, daily metrics can be more helpful than weekly or monthly metrics. In all of these types of workplaces, I've seen leaders draw poor conclusions about single data points (or two data points). Leaders overreact and ask for explanations that might not exist, so people feel pressured to come up with a reasonable-sounding answer that deflects the attention until the metric inevitably improves the next month. Again, there is a better way.

Bullying: Another leader told me, "We have an old-school, command-and-control (if not bullying) culture. Leaders demand better performance and throw goals at people, whether they're achievable or not. They celebrate when we make a small improvement and come down on us when performance dips. They throw solutions at us instead of helping us develop our improvements. We're trying to shift from a blame-and-shame culture to a data-driven culture." I've seen a lot of workplace bullying throughout my career and it's especially sad to see in healthcare, where the human stakes are so high.

I've heard others express nagging doubts such as, "Our team meets daily and talks about our metrics. They get better, and then they get worse. Our monthly strategic reviews are the same thing but on a monthly basis. Instead of stating the obvious, such as 'that measure is better than before,' can we move to a deeper understanding of our data that leads to real improvement?"

Data can be very helpful. But, more data and more metrics could, unfortunately lead to more blaming and more bullying when we'd want it to instead lead to more improvement. This book is not just about some methods and tools; it's also about new mindsets for managers that can lead to meaningful culture change.

This book will show you how to draw more timely, more valid conclusions based on your metrics, trading overreaction and busyness for more focused and effective improvement.

I'm blessed to have been exposed to some simple, proven, and effective methods that can help in these situations. I can't claim to have invented the methods described in this book, but I can try to put these methods in a modern context that

is more relatable. The methods in this book come from a line of legendary thinkers, including the famed management guru W. Edwards Deming, the American who taught the Japanese about quality after World War II.

Deming was taught by Walter Shewhart, who invented the "statistical process control" methods that Deming expanded into a broader system of management that included understanding variation, appreciation for a system, understanding psychology, and understanding how we know what we know (i.e., the theory of knowledge).

In recent decades, Donald J. Wheeler, Ph.D. has built upon and spread these methods including what he calls "Process Behavior Charts."[1] He has inspired many, including me, through his seminal book *Understanding Variation: The Key to Managing Chaos*. I was very fortunate that my father, Bob, had a copy of Dr. Wheeler's book (along with Deming's *Out of the Crisis*) after being a student in classes they taught at General Motors. My browsing of his bookshelf piqued my interest in these topics and opened my eyes to a different way of managing.

Different roles through my career have allowed me to test and validate the practical and helpful nature of these methods, even if these methods were not in sync with the organizational culture as a whole. I'm convinced these methods are helpful, which is why I feel driven to share them through this book and my other work.

As Dr. Wheeler says, these management and statistical methods are "a way of thinking, with some tools attached."[2]

By reading this book, you'll learn tools and this new way of thinking. Once you learn to understand variation, it's impossible to unlearn. You'll see opportunities to apply these principles in news headlines and your workplace every day.

How to Use This Book

I hope this book will be helpful to many types of readers, including those who will create Process Behavior Charts (including analysts, process improvement specialists, quality department staff, etc.) and those who will primarily consume and use such charts and metrics (including managers, executives, business owners, venture capitalists, etc.).

---

[1] Wheeler, Donald J., *Making Sense of Data: SPC for the Service Sector* (Knoxville, TN: SPC Press, 2003), 95.
[2] Wheeler, Donald J, "The New Terminology," SPC Ink, 1998, no.2, Manuscript No. 129, 2.

Chapter 1 introduces the idea of "what measured gets managed," but takes a deeper dive into what "managed" means. Topics in the chapter include choosing the right metrics, the danger of arbitrary targets, and the case for why Process Behavior Charts matter: leaders and employees can waste less time "chasing the noise" in a metric, which allows them to spend more time on systematic and sustainable improvement.

Chapters 2 and 3 introduce the Process Behavior Chart method and how to use such charts. Comparisons are drawn to some common existing methods for tracking metrics and evaluating performance against targets. Why are Process Behavior Charts more effective than two-data-point comparisons, "bowling charts," and the like?

Chapter 4 takes a deeper dive, for those who need it, into the process and method for creating Process Behavior Charts. This chapter could be skipped by those who plan on being the audience for metrics and charts.

Chapter 5 connects charts and metrics to our primary goal of improvement. The chapter explores methods for turning an unpredictable system into a predictable system and how to improve a predictable system in less-reactive, more-systematic ways.

Chapter 6 is a narrative description of an exercise that's effective in learning how to understand variation: the "Red Bead Game" that was made famous by the late, great Dr. W. Edwards Deming. Readers will have a chance to reflect on some of the common management tactics that are generally ineffective in a real workplace.

Chapter 7 shows how Process Behavior Charts can be used to better investigate situations that are posed in newspaper headlines. How are comparisons between two data points sometimes misleading? Does the "highest number in X years" mean that there's a significant shift in our data?

Chapter 8 further compares Process Behavior Charts to common management methods and analysis approaches, including linear trend lines, column charts, and more.

Chapter 9 comes back to workplace case studies and how we would use Process Behavior Charts to make better management decisions.

Chapter 10 explores ideas related to change management and successfully introducing a new method into an organization.

Chapter 1: Improving the Way We Improve

Most organizations these days are under pressure to perform better. How do we increase sales and production in our family-owned manufacturing company? How can we reduce infection rates in a hospital's intensive care unit? How do we get our startup on the growth trajectory that we promised to the venture capital firm?

I've heard people describe how it feels like they're on a proverbial performance roller coaster. There's a lot of up and down. There's anticipation, excitement, and sometimes yelling -- whether people are excited or scared. The emotional roller coaster of metrics and the way leaders react (or overreact) to them can be exhausting. This book is meant to help you get off the performance rollercoaster, both stabilizing and systematically improving your results.

In an age of "big data," we might be drowning in numbers and information. As leaders, we might be under pressure to look at performance on a daily (or even hourly) basis. Does it help to have more frequent "red bad, green good" evaluations of performance in a given day, week, or month? Or does this constant judging and color coding of points on a chart lead to a lot of overreaction, which then causes a bunch of wasted time for managers and their employees? Does any of this help us improve?

In an age of "big data," we might be drowning in numbers and information.

Using the methods in this book, we can turn a flood of data into a controlled flow of knowledge and insight that allows us to evaluate performance better, focusing our efforts on improvement instead of knee-jerk reactions.

The desire (or need) to improve doesn't mean that an organization knows how to improve. It doesn't mean leaders know how to look at their metrics (a.k.a., performance measures) to determine if they are actually improving. Do they *know* or are they guessing or relying on gut feel?

How often are these guesses or hunches incorrect or counterproductive? How often are leaders pressured to *make metrics look good* instead of actually improving the system and its underlying performance? How often do our reactions lead to a lot of wasted time? When do our reactions, no matter well intended, end up hampering improvement?

Are we meeting our goals or targets? What do we expect our future performance to be? How do we know if a change has led to a meaningful improvement? How can we discover if a system's performance is degrading?

In many organizations, leaders might guess or use their gut to answer these questions. Or, they might use methods, such as two data point comparisons and linear trend lines, that can be misleading if we don't fully understand their limitations or validity.

The dictionary defines success as "the accomplishment of one's goals."

In a workplace, one way we can ensure that we accomplish goals or meet targets is to set the bar very low. If this year's target is to be 2% better than last year's average performance, does meeting that goal really count as success that's meaningful to the organization? Does a 2% change represent a significant difference or does it mean that performance is just fluctuating?

Our "measures of success" in the context of this book aren't just limited to meeting a goal or target. Organizations and people that are successful continuously strive to be better. Many successful athletes and leaders are driven to be perfect, even if perfection is not possible.

This book will explain how to look at metrics in a better way, how to evaluate new data points, and how to look for trends that are meaningful and worth reacting to or learning from. This book will also show you how (and when) to respond to your metrics in ways that best drive sustainable improvements in our quest for high levels of performance.

Not Just Measuring Better, Also Managing Better

Metrics, or performance measures, have long been a popular topic in management, with many articles and books being written on the topic.

One expression you might see shared on social media or in somebody's email signature is:

"If you can't measure it, you can't manage it."

This quote is often attributed to the late Peter Drucker in the form of:

"You can't manage what you can't measure."

Another popular variation of this expression is:

"What gets measured, gets managed."

Statements like these emphasize the importance of measurement, but they're frustratingly vague about how to manage or improve what is measured. Focusing on "what gets measured" is the reality in modern organizations — what are the measures and what are our goals?

It's a problem that most organizations haven't fully explored what the phrase "gets managed" means or the best way to manage. This book aims to provide some clarity and guidance about how to manage (and improve) our metrics, so we can be more successful.

Measuring is easy; management is hard.

Leaders and employees, when under pressure to improve a metric, will pay attention to that metric. They'll talk about it. They might assign somebody to be responsible for the metric. They might form a team. That doesn't mean they know how to manage the metric and it doesn't mean they know how to improve the system that generates those results.

Can we measure everything in life (or our workplace) that matters? No. As Dr. W. Edwards Deming often said, "The most important figures that one needs for management are unknown or unknowable."

But, the reality is that we have to do our best to choose measures that matter (or our boss tells us what to measure) — and then manage those measures in the best (or least dysfunctional) way possible.

The implication seems to be, "If we choose the right metrics and set challenging targets, then improvement will happen." If that were true, it seems we wouldn't

have any waste, problems, or underperformance left in the world. How many organizations have promoted a goal of "zero defects" or "zero harm" without making significant progress toward that goal?

If it were only so easy, everybody would be hitting their targets, whether that's increasing sales in a startup or reducing falls in a hospital. I've seen too many cases where the organization didn't hit their goal in one year, only to then set the same goal (or a more challenging goal) the next year. W. Edwards Deming would always ask, "By what method?" It's not enough to set goals. We need to have effective improvement methodologies that start with better understanding the systems and processes in our work. Improving those systems and processes will lead to better results.

The Dangers of Arbitrary Targets

This target is, by definition, arbitrary. It's still arbitrary even if it's based on a competitive benchmark, last year's performance, or an organizational goal that has been passed down from senior leaders. We have to be careful with arbitrary targets, especially if our system is not currently capable of reaching that level of performance consistently.

Brian Joiner wrote in *Fourth Generation Management*, there are three things that can happen when people are pressured to hit a target — and the first two are dysfunctional (and generally easier than the third):

1. Distort the numbers
2. Distort the system
3. Improve the system

If the manager does nothing but put pressure on employees, the easiest path might be to distort the numbers or distort the system. In recent years, we've seen the CEO of Wells Fargo set an arbitrary target of "eight is great," meaning each customer should have eight different accounts. Since the goal was unreasonable, thousands of tellers signed customers up for accounts they didn't need or didn't know about. Eventually, the CEO was fired for the mess.

In the United States Veterans Health Administration, local clinic managers were put under pressure to keep waiting times for patients to 14 days or less. Even though the Congressional Budget Office called the target unreasonable, staff in dozens of offices created secret waiting lists (a paper waiting list to get on the official waiting list in the computer) or other such distortions to make results look better than they were.

I heard a funny story recently about a large gym and fitness center that asked customers, on their way out, to push one of four buttons that rated their visit as a very smiley face, a somewhat smiley face, a somewhat frowny face, or a very frowny face. Maybe you've seen similar devices in airport security or airport bathrooms (or perhaps some non-airport related place... you can tell I spend a lot of time in airports).

The gym's manager and staff were given the promise of an incentive if a certain number of customers hit the "very smiley" face each month. There was probably also an implied threat of punishment for not hitting that goal. What did the employees figure out? They learned that they could hit the "very smiley face" button a few times each time they walked nearby.

Effective managers don't just set targets; they work together with people to hit those targets.

What are the Right Metrics?

Before we talk about how to treat our metrics and improvement activity over time, there's another important question that organizations usually focus on: "what should we measure?"

Sometimes, measures are forced upon us (often causing quite a bit of harm and dysfunction), but we sometimes get the opportunity to choose metrics. Hopefully, we select metrics that matter instead of things that are easy to measure. But, there is a risk that the methods in this book would be used to "better manage" the "wrong metrics," which we'd like to avoid.

In my experience, people place far too much emphasis on choosing metrics and setting targets for those metrics, instead of also focusing on how to improve the metrics (and the systems that generate those results).

The focus of this book is to best manage the metrics that we have chosen - or those that have been handed to us by executives or headquarters. For a deeper treatment about what metrics to choose, books on the following methodologies could be helpful:

- Balanced scorecard
- Strategy deployment (a.k.a. hoshin planning)
- Lean Startup

Organizations like Toyota use a set of operational metrics that could be described as a "balanced scorecard," looking at measures of safety, quality, on-time delivery, cost, and employee morale. A broad set of metrics helps protect against the dysfunctions that can result from focusing too much on any single metric. For example, if cost is the primary metric, managers might be pressured into actions that hurt safety, quality, or other important measures.

Many organizations that practice some form of the Lean management approach use a high-level methodology called "strategy deployment" (sometimes using the Japanese words "hoshin kanri," which means "management compass"). An integrated healthcare system might define what their "true north" (as it is often called in the strategy deployment methodology) to be four key focus areas:

- Safety / Quality
- Customer Satisfaction
- People
- Financial Stewardship

Under the "Safety/Quality" area, one metric might be preventable patient mortality rates or the number of patients harmed by errors. If the hospital's strategy includes becoming a safe, high-quality provider of care, different departments would have metrics that contribute to that high-level goal. In many areas, patient falls might be identified as a metric that needs to improve. This specific metric would not be applicable, however, in departments like the laboratory or pharmacy. It might be more important for them to measure the rate of lost specimens or medication delivery errors. Each of those department metrics would align to the organization's higher-level strategy and goals, something that's crucially important.

It's also important that metrics are relatable to employees, so they feel like they and their managers can do improvement work that contributes to their metrics and the success of the organization as a whole.

[Should I add a section on setting the right goals? Voice of the Customer and actual "law of nature" requirements vs arbitrary targets meant to inspire?]

From “Vanity Metrics” to “Actionable Metrics”

The Lean Startup movement asks important questions about what to measure. Ries coined the phrase “vanity metrics” to describe our measures that “give the rosiest possible picture.” Instead of looking at the things that paint a picture of success, we should look at metrics that are truly the “key performance indicators” (KPIs) for our organization.

If we’re using metrics to evaluate the success of an improvement initiative, is our goal to manipulate or choose metrics that tell the story we want to tell, in a desire to have guaranteed success? Or, are we using metrics honestly to evaluate if we are getting better, getting worse, or if we’re in the awkward in-between state of having a metric that seems to just be fluctuating?

Ries suggests we replace vanity metrics with what he calls “actionable metrics,” where “data must demonstrate clear cause and effect and be related to changes” to our product, our services, or our system. Changes in our key performance indicators in our organization must help us draw clear cause and effect relationships between our improvement work and our results. Otherwise, we’re just randomly trying a bunch of improvements and that’s no way to run a business.

One classic example of a vanity metric is the number of visitors that come to your website. This number is easily measured through Google Analytics and we might want to see a trend that increases, going “up and to the right,” as entrepreneurs often say. If we’re desperate to paint a picture that shows we are making progress, we might latch onto this metric and put it into a presentation to our investors. But, more website traffic might be meaningless if that is not translating into more sales. What’s easy to measure (website traffic) isn’t always what’s meaningful to our business.

As Ries said in his latest book, *The Startup Way*:

> "The fact that your site has seen an uptick in visitors doesn’t mean your product is more popular or you’re more successful."[3]

Yes, what we measure matters. But what we do with those metrics is very important as well and that’s discussed far less often in the Lean Startup community and other industries.

---

[3]Ries, Eric, *The Startup Way: How Modern Companies Use Entrepreneurial Management to Transform Culture and Drive Long-Term Growth*, (Random House: New York, 2017), 103.

Measuring and Managing Personal or Organizational Health

Let's look at a measurement (and management) challenge that most of us can probably relate to from our personal lives. There are parallels to be found here to our organization health and wellness, as well.

Can you represent and measure your health with a single metric? Probably not. Relying too much on any single metric might lead to confusion or the wrong conclusions.

Let's say "fitness" is your high-level goal. That's difficult to measure directly and it's certainly not a single number. We might decide that weight is more easily measured and we have a hypothesis that says lower weight equals better fitness, which then leads to better health, a longer life, and a better quality of life. This hypothesis might only be true, however, if we're talking about an ideal weight and healthy ways of getting there.

Different people might choose a handful of process metrics that include weight, blood pressure, and total cholesterol. Others might choose different metrics that are more relevant to their own health, conditions, and risks. You might think of these measures as forming our "balanced scorecard" that represents one's health in a way that prevents you from the dysfunctions that can result from focusing too much on any one metric.

For each metric, we typically set a goal or a target. If there is a gap between our current performance and the target, we work to close that gap through various forms of improvement activity. Hopefully, we are going about that in a systematic way instead of just randomly trying different solutions or quick fixes that we run across.

Weight is just a number. I don't manage the number; I do what I can to manage the system (including diet and exercise) that leads to results. The same is true in our workplaces. The word "results" should remind us that the metric is the result of some work, process, or system.

KEY POINT #1: We don't manage the metric; we manage the system that leads to the results (a.k.a. metrics) and we lead the people who help us improve the system.

As I've learned in my own experience, working out three times a week with a trainer might actually lead to a net weight *gain* that occurs as the result of losing fat and adding muscle. One might be troubled if our high-level goal was defined only as weight. But, somebody aiming for fitness or health as a goal might readily accept a higher weight, especially when their pants fit better and their heart is performing more efficiently.

We might choose to get a scale that claims to measure our body-fat percentage if we decide that is a better metric than weight. Choosing the right metric (or the right goal) is an important challenge, in any context. We might use this balanced scorecard of metrics to provide a better picture of our personal (or organizational) health. Metrics, over time, can help us to determine if we are improving, getting worse, or staying the same.

A few years back, my weight approached 200 pounds for the first time. I identified a gap between my current weight and my goal of 185 pounds (it was admittedly an arbitrary target). Measurement was important, but I also learned not to overreact to every small fluctuation in my weight, something we'll learn not to do in workplace settings, thanks to the methods in this book. I learned to weigh myself at a consistent time each day to eliminate the effect of how weight naturally fluctuates during the day. I improved my own personal system (exercising more and using an app to track calories) and saw my weight decrease to that goal over a few months.

I learned to look at my weight as a range of numbers that naturally fluctuates a bit from day to day instead of thinking of it as a single number that always had to remain exactly the same. Small fluctuations didn't cause alarm, but larger changes would draw more attention (and analysis of how my diet or exercise had changed). My weight wasn't just a number; it was the result of a process and a system.

It can be difficult to sustain an improved system. Case in point, my weight over time has crept back up to 200 pounds for a variety of reasons. The answer to why my weight increased again can't be answered by asking, "What went wrong yesterday (or last week)?" It's more related to patterns and changes in behavior

over time. Maybe I can blame aging and metabolism. Either way, measurement doesn't automatically prevent a problem from coming back (or a new problem from popping up). I've demonstrated that I know how to lose weight. Why didn't I take corrective action when my weight went up to 190? Why am I only reacting again now that I'm back at 200?

If managing one's own personal system is difficult, it goes to show how much more challenging this is in an organization — not just improving performance, but also sustaining the gains. How often do we let an organizational metric get really bad before we respond and ask questions about improvement? How can we manage more effectively?

Reducing Waste Means We Can Improve More

Hardly anybody complains about having too much free time at work, right? Time is a precious commodity. Leaders, including startup founders, nurse managers, and chief financial officers at publicly traded companies, all have to make constant choices about how to allocate their limited time and attention. Urgent (but hopefully proverbial) fires pop up regularly, data gets thrown at us, and we're expected to respond. You might be told, "Don't just sit there, do something!"

But, how often do our choices about what to react to (and how to react) end up wasting and consuming more of our time, to little real benefit? Does a series of non-stop reactive, knee-jerk responses to changes in performance metrics lead to continuous improvement?

I've spent more than 20 years studying and practicing the "Lean" management methodology. The origins of Lean are Toyota and manufacturing, but these methods and mindsets are being used in healthcare and many other workplaces. Toyota has long emphasized that the "waste of motion" is to be avoided, as one of the "eight types of waste." Waste is defined, generally, as anything that does not add value to a paying customer. In the Lean methodology, we aim to improve efficiency and flow by reducing waste, interruptions, and delays, instead of just focusing on doing the actual work more quickly.

When I worked in manufacturing, we'd improve the process, which often meant changing the physical layout of machines to reduce the amount of walking required of production team members. Workbenches were designed in a way to reduce the amount of reaching for parts and tools. No longer bending down constantly to pick up parts meant better ergonomics, fewer injuries, and lower cost (not to mention it was the right thing to do in terms of treating employees with respect, another fundamental Lean principle).

Reducing wasted motion meant we could reallocate that time to do more work that added value to customers. We could do more with less without overburdening team members. Production output would increase, waiting times would go down, and quality would improve.

The same ideas apply in healthcare. Hospital nurses typically spend only 30% of their workday at the bedside with patients. Their day is full of waste and that includes a lot of wasted motion. About 20% of their day is spent walking around searching for supplies and equipment they need to provide patient care. This isn't the nurses' fault; it's a matter of bad processes, not bad people.

Nurses, leaders, and improvement specialists can work together to reduce this waste – making work easier instead of pressuring people to move faster. Reducing waste means more time with patients, which means happier staff and patients, faster recovery times, and fewer errors, among other benefits. Again, reducing waste is a pathway to better performance.

We can also reduce waste in knowledge work. Wasted motion can occur even when we're sitting in a chair, thinking, typing, or moving our mouse..

Eric Ries, the author of the best-selling book *The Lean Startup*, wrote about how wasteful it is to have software developers spend years working on software that doesn't get purchased by customers. Ries asked a very important question at the end of the book, something that's both a matter of respect and organizational performance:

> **"If we stopped wasting people's time, what would they do with it?"[4]**

In general, that time can be used for productive, or value-added, activities that better serve our customers and lead to more success for ourselves and our organizations.

[4] Ries, Eric, *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses* (Currency: New York, 2011), 280.

Reducing Waste in the Work of Leaders

Leaders often express concern about waste in the work of frontline employees. But what about the wasted time of managers and executives? Their time is arguably more valuable, given their bigger paychecks, so let's stop wasting it. One way leaders waste time is in reacting (or overreacting) to performance measures that they see in boardrooms, monthly reports, and huddle boards.

One way leaders waste time is in reacting (or overreacting) to performance measures that they see in boardrooms, monthly reports, and huddle boards.

Meetings are held, lines are drawn on charts, data points are circled, pronouncements are made, solutions are offered… but how often is this a different type of waste? This waste might be more difficult to see than the waste in physical work, but it also might be more damaging to our organization. Managerial waste hinders progress more than a little waste in the work of frontline employees.

The methods in this book will help you save time by teaching you when and how to react (or not react) to changes in metrics, allowing you to focus your efforts on improvement activity that's aimed at meeting goals and targets more consistently, if not every time.

Making Sense of the Numbers

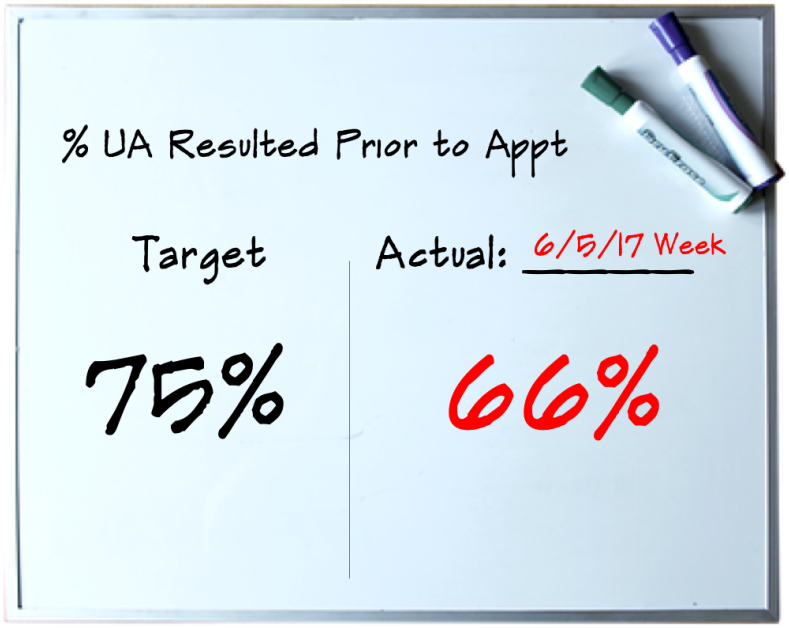
If an executive tells the VP of sales, "Revenue is 10% below our target this month" or "Sales are down 5%," it's easy to say, "That's bad." Does the recognition that a number is lower or worse than a goal or better or worse than the previous month lead to actions that help improve the metric?

Many organizations limit their analysis of performance to simple comparisons to a goal, target, or budget. Or, a metric is compared to a previous time period or an average. What does that tell us? Not much — and not enough to really improve.

Let's look at a scenario where a health clinic laboratory and medical office team have a measure they track: the percentage of urinalysis (often abbreviated as "UA")

test results that have been received prior to the patient's appointment with their provider.

If we present just a single data point posted on a whiteboard, it's obvious that we didn't meet the target this week. Answering the question for a single data point — yes or no, bad or good, or even color coding numbers as red or greet — doesn't mean we can determine anything about trends and it doesn't tell us how to improve.



Instead of only asking "are we achieving our target?" many organizations will compare performance to the previous week. The lab manager might add, "Our percentage of UAs resulted dropped from 66% to 49% last week." What does that tell us? Does that help us improve?

There are additional questions that are helpful, if not necessary, for improvement. I propose three high-level questions:

Question 1: Are we achieving our target?

Question 2: How is our performance trending?

Question 3: How do we improve performance?

Hopefully, this book will help you decide that all three questions are important and deserve attention.

To answer **Question 1**, we see we are currently not hitting the target, or at least we did not do so in the two weeks. Does this mean we have no hope of ever hitting the target without taking some sort of drastic action? We can't answer that question without additional data. The methods in this book will help us do a better job of answering that question. Additionally, we'll learn methods that help us predict if we're likely to hit those targets in future weeks.

Question 2 is important in a number of situations. For one, if we're not meeting our target, we would want to know if we're making progress toward meeting it. If we're currently hitting our target, we still might have interest in reaching even-higher levels of performance. Single data points or comparisons of two data points, as we have here with the UA data) don't help much in this regard. The second data point was lower than the first. Does that mean that it's going to keep going down, or is data just fluctuating? We can't tell how our performance is trending with that simple whiteboard.

Having such a limited view of data (one or two data points compared to each other or a target) doesn't help us know where to get started in answering **Question 3**.

The methods in this book will help us make better decisions about charts and metrics. We'll be able to draw better cause-and-effect connections between our improvement efforts and our results. Moreover, we'll avoid wasting time since we'll stop overreacting to every up and down in the metric.

We'll be able to answer more specific versions of those three questions:

Question 1: Are we achieving our target?

a. Are we doing so occasionally?
b. Are we doing so consistently?

Question 2: How is our performance trending?

a. Are we improving?
b. Are we getting worse?
c. Are we staying the same?
d. How do we know?

Question 3: How do we improve performance?

a. What methods should we use?
b. How will know if we've improved?

The goal of this book is not learning how to create charts for our metrics. The goal is improving performance through these methods and these questions.

So How Do We Manage Our Metrics?

For my personal consulting and speaking company that I've run since 2010, blog traffic is a vanity metric since that's not a measure of my core business model. I don't make much money from advertising or other revenue streams that correlate directly to more people reading my blog.

Metrics that are more important for my business include the number of new consulting or speaking clients I get. Higher blog traffic might help further that goal, but I can't say for sure. I can't prove, statistically, that months with higher blog traffic result in more speaking and consulting work. That might be a function of the *quality* of my posts (or other factors that are in or out of my control).

If blog traffic did matter, how would we go about best managing that metric?

When we look at metrics, it's too easy to draw an incorrect conclusion, one that leads to an unrealistic view of our current situation. Or, it's often too easy to intentionally distort or cherry pick numbers to tell the story we want to tell.

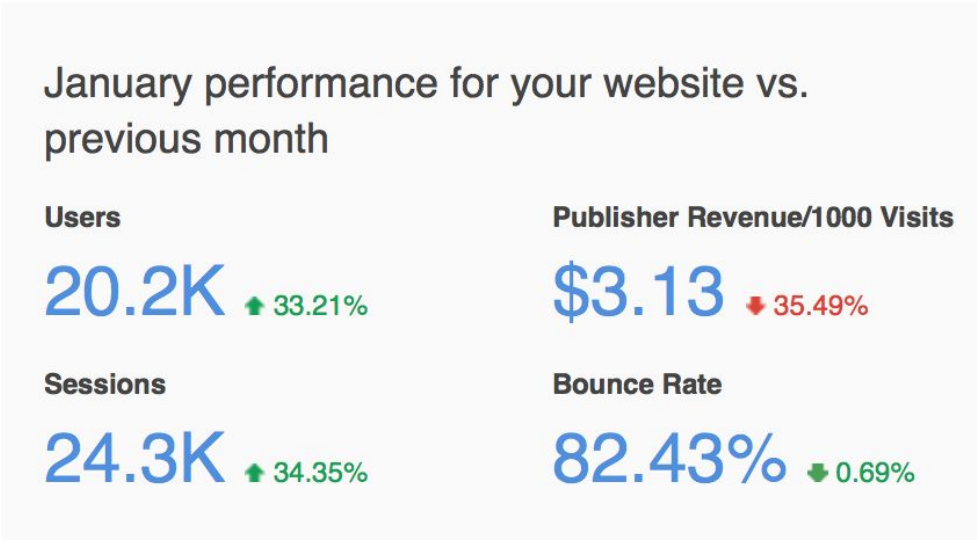
In the case of website metrics, we could modify the Ries quote to say something slightly different, but still important:

> **"The fact that your site has seen an uptick in visitors doesn't mean your website is getting more visitors."**

You might ask, "What do you mean? A higher number means more visitors. More visitors means more visitors."

The dictionary defines "uptick" as "an increase" or "a small increase." Often, it's also just a comparison of two data points (and two data points are not a trend).

Google Analytics sends me an email each month that shows a comparison of two data points related to my blog, www.LeanBlog.org, as shown below:



I might feel good that the number of users is up 33.21%. How do I know it hadn't decreased by a similar amount the month before?

Two data points aren't very helpful. Knowing a number is higher or lower than the previous month doesn't tell us very much. It's not enough information for us to understand our system and to evaluate trends in our performance.

KEY POINT #2: Two data points are not a trend.

However, I can log in to the Google Analytics website to download more data, allowing me to look beyond two data points or two weeks. I can pretty easily find, somewhere in the data, other occurrences of two data points that represent an "uptick" if I wanted to.

I can present the data as a table of numbers as we often see in management reports:

Metric	Page Views
March 2017	46,525
April 2017	64,481
Percent change	38.6%

I might say, “Hooray, an increase of almost 40%!” I’m rounding up, of course, to paint an even better picture.

But, this view of two data points also doesn’t tell me much. April’s number being higher than March’s tells me… that April had more page views than March. What can I do with that information?

If website traffic (or another metric) *did* matter, looking at two data points doesn’t allow me to definitively answer the important question of “is website traffic higher?” It certainly doesn’t answer the question of “Is my business growing?”

People often pretend that two data points are the beginning of a trend. Again, I might want to (or feel pressured to) paint that picture of the metric going up and to the right.

Thankfully, I don’t have investors or an executive whom I’m trying to convince that 40% improvement is going to happen every month. I’d only be fooling myself to think so and doing that doesn’t help improve my business. Ries always warns against creating “success theater,” where one form of that is using metrics in a way that makes performance look better than it really is.

Even if I had a hypothesis that said traffic was now going to increase by 40% every month (or just increase at all each month), looking at the data in a different context, looking at a Run Chart with more data points, could prove or disprove that hypothesis pretty quickly.

Continuing this “two data point” view, May’s number reveals a 38.7% decline that followed the previous 38.6% increase, as shown below:

Metric	Page Views
April 2017	64,481
May 2017	39,517
Percent change	-38.7%

Well, so much for that trend I might have hoped for. Some would call this "regressing to the mean." My blog traffic came back down to earth, as others might say. This was, of course, very visible in the Run Chart, above. These numbers tell me something (not much) about the past. What could I expect to see for blog traffic in the future? Can I better understand past performance in a way that predicts future performance? Yes, using the method we'll introduce in the next chapter.

Comparing just two data points means we are missing a lot of context, including the other months' data and the trends we might see by looking at more data.

As Donald J. Wheeler, Ph.D. wrote in *Understanding Variation*:

KEY POINT #3: "No data have meaning apart from their context."[5]

If comparing two numbers isn't very helpful, organizations often try to do better by displaying tables with many numbers on bulletin boards or electronic dashboards.

It's very difficult for people to see trends in a table of numbers, the way some of my blog data is shown below. This style of presenting a metric over time is sometimes referred to as a "Bowling Chart" because it sort of looks like the grid you'd use to keep score when bowling, with numbers going from left to right.

[5] Wheeler, Donald J., *Understanding Variation* (SPC Press: Knoxville, TN), 13.

Jan-16	Feb-16	Mar-16	Apr-16	May-16	Jun-16	Jul-16	Aug-16	Sep-16	Oct-16	Nov-16	Dec-16
48,387	47,668	45,452	46,752	46,725	39,131	32,271	39,829	37,888	40,378	39,100	32,069
Jan-17	Feb-17	Mar-17	Apr-17	May-17	Jun-17	Jul-17	Aug-17	Sep-17	Oct-17		
42,153	47,759	46,525	64,841	39,517	32,245	31,862	32,717	31,505	34,555		

Is it easy to identify trends in the data when it's presented this way? Is my blog traffic increasing or decreasing?

Sometimes, these Bowling Charts have color coding that compares each month's data point to a goal, as shown below in an example from healthcare:

True North Pillar	Metric	Desired Direction		Jan-16	Feb-16	Mar-16	Apr-16	May-16	Jun-16	Jul-16	Aug-16	Sep-16	Oct-16	Nov-16	Dec-16
Quality & Safety	Falls with Severe Injury	DOWN	Actual	0	1	0	0	1	0	0	0	0			
			Target	0	0	0	0	0	0	0	0	0	0	0	0
Quality & Safety	CLABSI	DOWN	Actual	0	0	1	0	0	1	2	0	0			
			Target	0	0	0	0	0	0	0	0	0	0	0	0
Quality & Safety	CAUTI	DOWN	Actual	1	0	0	0	0	0	0	1	1			
			Target	0	0	0	0	0	0	0	0	0	0	0	0
Team Engagement	Necessary Information Available Survey	UP	Actual	67%	50%	48%	70%	46%	40%	60%	70%	50%			
			Target	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%	80.00%
Patient Experience	Likelihood to recommend	UP	Actual	76.40%	78.40%	80.50%	70.00%	46.00%	40.00%	60.00%	70.00%	50.00%			
			Target	79.90%	79.90%	79.90%	79.90%	79.90%	79.90%	79.90%	79.90%	79.90%	79.90%	79.90%	79.90%
Patient Experience	Hospital rating	UP	Actual	76.40%	77.70%	77.60%	77.60%	77.20%	77.60%	77.00%	75.00%	63.00%			
			Target	77.50%	77.50%	77.50%	77.50%	77.50%	77.50%	77.50%	77.50%	77.50%	77.50%	77.50%	77.50%

A red/green Bowling Chart answers our **Question #1** of "Are we meeting our targets?" We might be able to answer **Questions #1a** and **#1b**, "Are we meeting our targets occasionally or consistently?" by counting the number of green and red boxes.

The Bowling Chart doesn't easily answer our **Question #2** of "How is our performance trending?" We might vaguely determine we're getting better if we see more red to the left of the chart and more green to the right. But how can we tell if those changes are significant and sustainable?

And, this binary good-or-bad analysis does not point us in the direction of answering **Questions #3a** and **#3b** about "What [improvement] methods should we use?" and "How will we know if we improve?"

When presented with a Bowling Chart or a so-called dashboard with a large number of metrics, the red and green color coding doesn't help us prioritize our investigation and improvement efforts. What if six of our ten the metrics are red right now? Does each of those merit the same response? Which metric has the

biggest gap between target and actual? Does a metric going from green to red mean that anything has really changed in the underlying system? Does a metric remaining green mean nothing has changed?

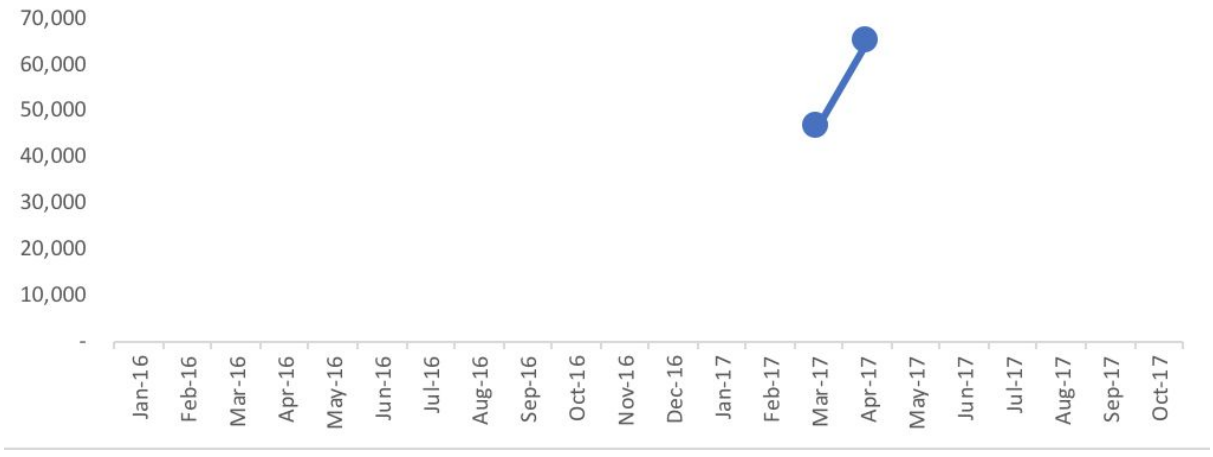
What if all ten metrics on the scorecard are green? Does that mean we simply have unchallenging goals? Does "all green" mean there's no need for problem solving or improvement? Does a metric remaining green mean that nothing has changed or that nothing is degrading in the process?

Charts are More Helpful Than Lists of Numbers

Compared to a list or table of numbers, a chart is a much more effective way of making sense of data and metrics. A chart is pictorial and our human brain processes images much better than it does lists of numbers. Watch a cable business channel for 30 minutes and you'll see how and why they usually show a graph of a company's stock performance over time instead of a table of numbers showing the stock price on different days.

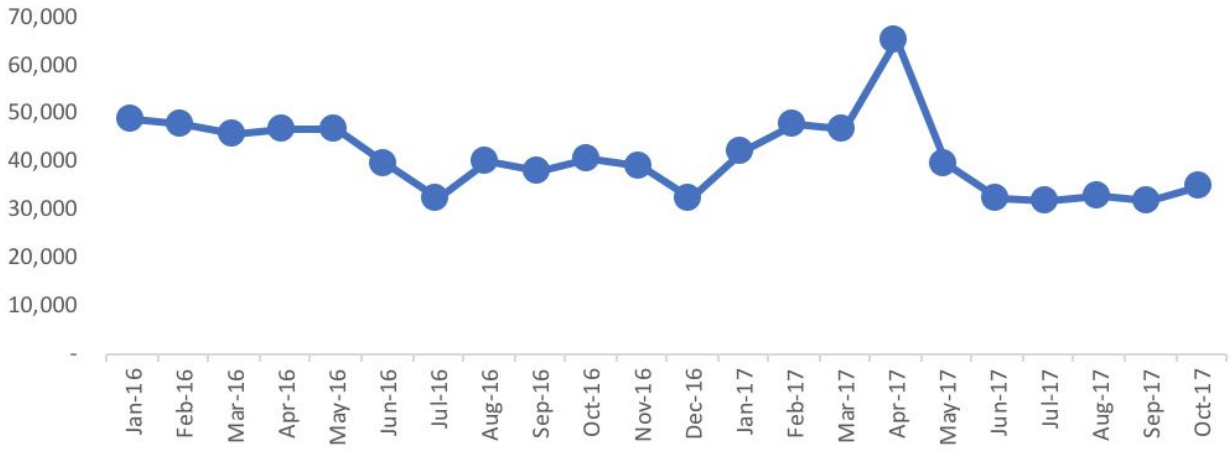
Charts are easy to produce with modern technology. In the workplace, we can even use very old technology to draw them by hand. There's no technological excuse for bombarding people with tables of numbers instead of presenting a graph. The Bowling Chart shown above could be replaced with six charts, an approach that would give much more information about trends to the consumer of that data. Six charts would take up more space on a bulletin board but would certainly help a leader and their team save time and improve more.

When given just two data points (such as "blog traffic is higher this month"), drawing a Run Chart with just two data points isn't very helpful, either, as we see below.



With two data points, as a table or a chart, we're missing context that tells us if the difference between those numbers is typical or unusual, routine or exceptional. How much do the numbers normally change from month to month? All we know from this treatment of the data is that the April 2017 number is higher than it was in March 2017. Even saying that April is 39% higher than March doesn't provide much context, since we don't know if it normally fluctuates that much month to month.

But, going back further in history, a Run Chart with additional data points tells us much more, as seen below.



KEY POINT #4: A chart will always tell us more than two data points or a table of numbers.

It's always helpful to plot the dots. As you learn and practice this methodology, I hope you'll find yourself challenging others when they present two data points or a simple before-and-after comparison. You can ask them to "plot the dots."

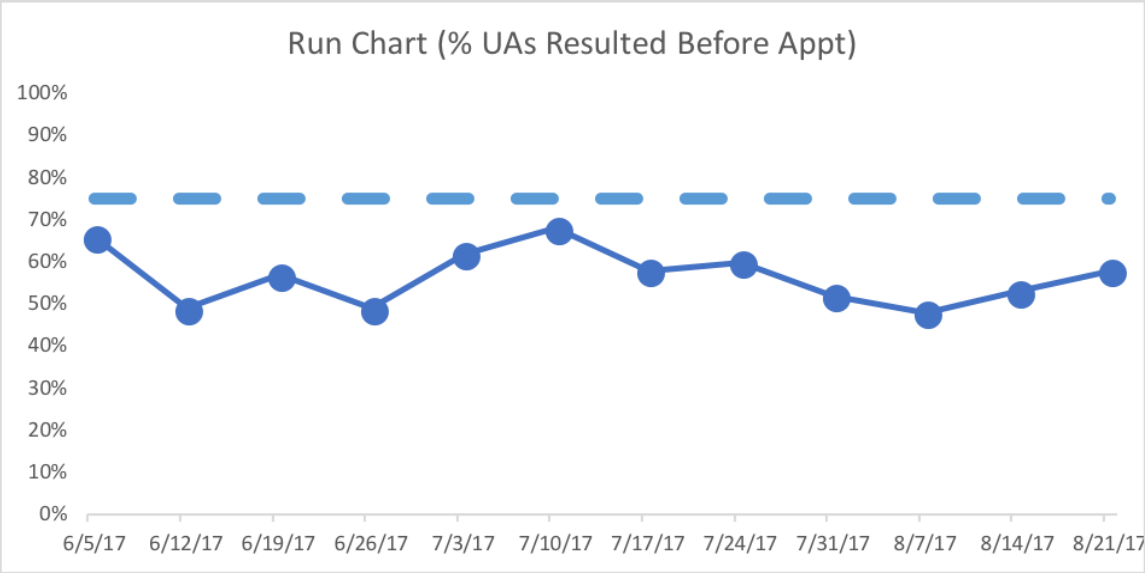
I've seen some people on Twitter use the hashtag #plotthedots when discussing this subject. Exploring that hashtag is a great way of finding other examples where charts would be helpful or times when we should be evaluating charts properly.

The Run Chart tells the honest story that my blog traffic fluctuates from month to month. The Bowling Chart tries to tell us the same thing, but it's much easier to

hear what Deming and Wheeler call “the voice of the process” when we allow a chart to speak to us. The voice of the process tells us how we’re performing and helps us answer our three core questions. As we’ll discuss more in Chapter 5, the “voice of the customer” tells us what is required, such as a specification or target. We hope the voice of the process tells us that our system is capable of meeting those needs all of the time. If not, we have some improvement work to do.

“It’s much easier to hear the voice of the process when we allow a chart to speak to us.”

Thinking back to the whiteboard and the Urinalysis metric from earlier in this chapter, a Run Chart tells us so much more, as shown below:



The chart shows the goal (the dashed light blue line) and our 12 weeks of data. We can answer **Question 1** a little better now. It appears we have never reached our target and it seem unlikely to happen unless we can improve our system and processes. We can try to answer **Question 2** by saying that the metric seems to just be fluctuating — some weeks are better than others. Does this chart help us answer **Question 3** about how to improve? Not really. In this book, we’ll introduce formal methods for answering those questions using sound, proven approaches that don’t rely on guessing or gut feel.

The same system, with the same people doing the same work in the same market will not always produce the exact same results – the number each day, each week, or each month. This is a fact of life – there's always variation in a metric. This was true in my blog traffic and it was true in the urinalysis example.

The question, which we will answer soon, is "how much variation is routine and to be expected in this particular situation?" If we discover that the variation in our metric is about the same over time, we then have a performance measure that allows us to predict future blog traffic or the percentage of UA tests that will be resulted before appointments.

KEY POINT #5: The job of management is not just to look backward but also to look forward and predict, if possible, what is likely to occur.

Metrics are most often used to look backward and to make evaluations about past performance. But, we also need to prepare for the future of our organization and its performance.

As the late MIT professor Myron Tribus said:

> "Managing a company by means of the monthly report is like trying to drive a car by watching the yellow line in the rear-view mirror."[6]

Glancing at the Run Chart for my blog traffic data, it looks like the March data point could possibly be a significant increase from the previous months, which appeared to be fluctuating around some average number of visitors. Is this March increase something exceptional? Does it deserve our attention or an investigation? How would we know?

The next chapter introduces the method we can use to answer our key questions: the "Process Behavior Chart."

---

[6]Wheeler, Donald J., *Understanding Variation* (SPC Press: Knoxville, TN, 1998), 4.

Summary of Key Points Introduced So Far:

KEY POINT #1: We don't manage the metric; we manage the system that leads to the results and we lead the people who help us improve the system.

KEY POINT #2: Two data points are not a trend.

KEY POINT #3: "No data have meaning apart from their context."

KEY POINT #4: A chart will always tell us more than two data points or a table of numbers.

KEY POINT #5: The job of management is not just to look backward, but also to look forward and predict, if possible, what is likely to occur.

Our Three Questions:

Question 1: Are we achieving our target?

a. Are we doing so occasionally?
b. Are we doing so consistently?

Question 2: How is our performance trending?

a. Are we improving?
b. Are we getting worse?
c. Are we staying the same?
d. How do we know?

Question 3: How do we improve performance?

a. What methods should we use?
b. How will know if we've improved?

Chapter 2: Introducing the Process Behavior Chart

The Process Behavior Chart methodology provides a way to answer our key questions about performance and improvement. The Process Behavior Chart (PBC) won't tell us what changes we need to make to improve our system. But, it will focus our efforts, saving time and pointing us in the direction of the right improvement methodology for the right situation.

A Process Behavior Chart (PBC) is a form of what's more generally called a "Control Chart" or a "Statistical Process Control (SPC) Chart." It's also sometimes called an "Individuals Control Chart with a Moving Range." That's a mouthful.

Wheeler coined the PBC term, suggesting that word "control," while intended to have a benign connotation, nevertheless has "baggage" associated with it.[7] Another argument is that PBC is more accurate, as the chart explains the behavior of the process, as simple as that.

A PBC is a combination of two specialized and related Run Charts. The first Run Chart, called the "X Chart" (or "Individuals Chart") contains the data points from the performance measure we are tracking. The second Run Chart, called the "Moving Range Chart" (or "MR Chart") shows us the point-to-point variation between each data point on the X Chart. We'll see later how the MR Chart is useful.

The X Chart

On the X Chart, we plot our metric. We also perform some calculations that allow us to add three very helpful horizontal lines that help us interpret the chart:

- A "Central Line" (typically the average)
- An "Upper Natural Process Limit"
- A "Lower Natural Process Limit"

In Chapter 4, we will see the detailed methodology for calculating the Lower and Upper Natural Process Limits (or just "Lower and Upper Limits" or "Natural Process Limits," for short). For now, we will start by learning how to interpret charts that have already been created.

---

[7] Wheeler, Donald J., *Making Sense of Data: SPC for the Service Sector* (Knoxville, TN: SPC Press, 2003), 97.

These three values are calculated from a period of baseline historical data. Ideally, we'd have 20 historical data points to use for this purpose. We can create a PBC with as few as six data points if that's all we have to start, but the calculated limits are more valid when we have more baseline data.

It should be emphasized that the Lower and Upper Natural Process Limits are *calculated*. They are not chosen by the creator of the chart. The limits should not be confused with any management targets for the metric or specifications that come from a customer.

When using PBCs, we always start by looking back at a baseline of historical data. Once we establish the baseline PBC, we can use it to determine if we can make predictions about future performance (remembering Tribus and Key Point #*5* from Chapter 1). We'll soon learn how the PBC tells us if our system is predictable or not. We'll also learn how, as we add data points over time, we can use the PBC to help understand if the system has changed or not.

It's possible that the PBC tells us that we don't have a predictable system with predictable future performance. Understanding that helps us understand how to improve (**Question #3**), as we'll discuss in Chapter *5*.

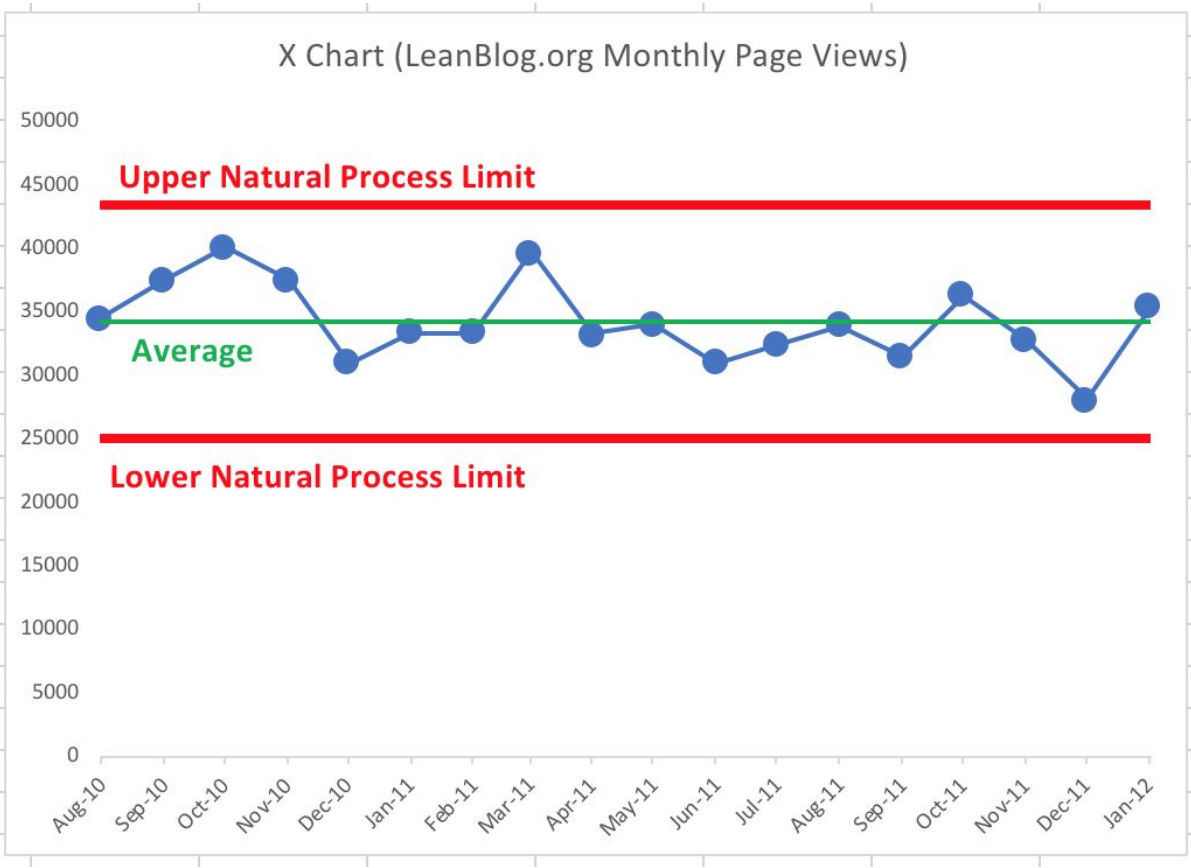
If the PBC shows us that our metric is what we call a "predictable" system, we can predict how our system will likely perform, or what the range of performance in the metric will be. If we have predictable performance, it will continue if there are no significant changes in the system that leads to the metrics and results.

As we'll see, it's possible that our metric is predictable, per the voice of the process, but it does not satisfy the needs (the target) articulated by the voice of the customer. Knowing this also helps us answer **Question #3** about how to improve.

As we try to improve, we'll sometimes introduce a change to the system and use the PBC to see if we get the results we expected. Other times, the PBC tells us that something has changed and it's then our job to investigate and try to understand why. In other words, PBCs help us evaluate a hypothesis about an intentional change and they also help us discover an emergent change to the system, as we'll discuss more in Chapter *5*.

Looking at our First Process Behavior Chart

If I chart my blog's traffic from July 2010 to January 2012, the X Chart portion of the PBC is seen below:



At a glance, we see that the data tends to fluctuate around an average (the green line) without any real trends or patterns. What else does the chart say to us?

There are some rules that tell us if there are any "signals" in the chart. Finding a signal is evidence of what we call "exceptional variation" — a data point (or data points) that are unlikely to be generated by our previously established system, or the "routine variation" that normally exists.

The three rules we use to identify signals in the X Chart of a PBC are:

- **Rule 1**: Any single data point outside of the limits.
- **Rule 2**: Eight consecutive points on the same side of the central line.
- **Rule 3**: 3 out of 3 (or 3 out of 4) points that are closer to the same limit than they are to the central line.[8]

My initial X Chart, for blog visitor data, shows no signals. None of the three rules are triggered. That means the metric is the result of a predictable process. Each data point on that chart can be considered as "noise" or routine variation.

A signal, if we see one, tells us something has changed and that there's a "special cause" to be found. In other words, we can identify the root cause of a signal. The special cause leads to exceptional variation in our metric.

If there is a signal that takes the chart in a positive direction, we need to understand the cause so we can make sure the change isn't just temporary. If we signal in a negative direction, we need to identify and eliminate that special cause, so we can bring performance back to where it was (or make it better).

In the absence of signals, each of these data points can be called "noise" in the system, as they represent what is also called "routine variation" — variation that is the result of many underlying causes that are hard to separate from each other. We need a different methodology for improving a system that has nothing but routine or "common cause" variation, as we'll discuss in Chapter 5.

Synonyms:		Synonyms:
Common cause variation		Special cause variation
Routine variation		Exceptional variation
Noise		Signal

This chart shows me that, some months, blog traffic will be higher than others. I can't expect that a complex system will generate exactly the same results each month. The chart allows me to predict that future blog traffic will fall between

[8] Some sources will list additional rules for finding signals, such as the "Western Electric Rules," "Shewhart Rules," or "Nelson Rules." Adding more rules increases the risk of "false positive" signals, so we can follow Wheeler's advice is to use just these three rules. See more here.

about 25,000 and 43,000 page views in future months — unless something changes in the system. It's a "predictable system." But how do we know?

So how do we know exactly if our charts have signals or if we're just looking at noise?

As Wheeler writes:

> "While every data set contains noise, some data sets may contain signals. Before you can detect a signal within any given data set, you must first filter out the noise."[9]

As I've heard Wheeler say, at a conference:

> "The tool for filtering out that noise is the Process Behavior Chart."

In other words, we are looking for moments of exceptional variation by filtering out routine variation.

KEY POINT #6: There is variation in every metric or data set. Process Behavior Charts filter out noise so we can identify signals.

Again, a signal means there is exceptional variation, a data point or points unlikely to have been generated by our old system. Something has changed. It means we don't have a predictable system. Our methodology for improving an unpredictable system is different, as we'll see in Chapter 5, as we'll first work to eliminate the special causes of our signals to create a predictable system. Once made predictable, the average performance of the system can be improved. And, we can also its range of routine variation, as expressed by the calculated Natural Process Limits).

Wheeler and Sheila R. Poling, in their book *Building Continual Improvement*, compare the relative confidence we should place in the three rules. A single point outside of the limits (***Rule 1***) is a "large effect," meaning there is, with almost complete certainty, a root cause for that change. With that single data point, we don't know yet if the cause was something temporary that will go away (or can be

[9] Wheeler, Donald J., *Understanding Variation* (SPC Press: Knoxville, TN, 1998), 30.

eliminated) or if it's the beginning of a more sustained effect that would be identified by our other two rules. ***Rule 2*** signifies a "moderate but sustained effect" and ***Rule 3*** is a "weak but sustained effect."

The lack of signals in a PBC indicates we have a predictable system. In a predictable system, each data point can be considered "noise" or routine variation. There's no specific reason or root cause for why a particular data point is what it is. It's important not to overreact to noise. Asking, "What happened last month" rarely helps us improve a predictable system.

Routine variation always exists in a system. There might be dozens or hundreds of sources of variation, or "common causes," that always exist. For my blog, those common causes might include the topics I choose, the quality of my writing, and the extent to which posts get shared by people on social media.

The Moving Ranges and the MR Chart

After making the X Char, we also then create a second chart that forms our complete PBC, the "Moving Range Chart" (or "MR Chart"). A "Moving Range" is the absolute value (meaning a positive number) of the difference between each two successive data points in the metric that's plotted on the X Chart.

Below, we see how the Moving Ranges are calculated for a metric:

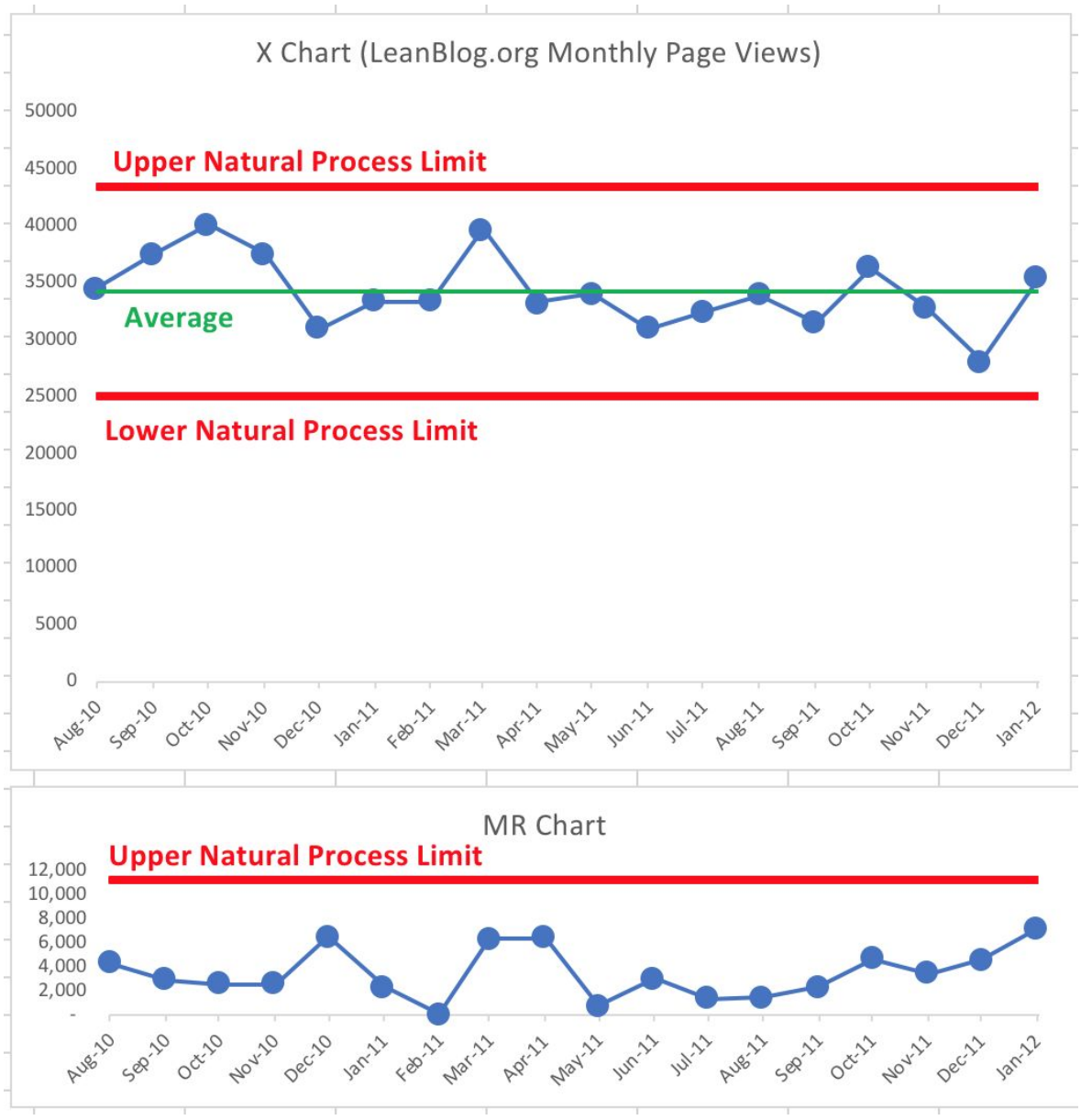
	Jan-17	Feb-17	Mar-17	Apr-17	May-17	Jun-17
# of New Customers	30	37	32	30	36	39
Moving Range		7	5	2	6	3

The Moving Ranges and the Average Moving Range gives us an estimate of the typical point-to-point variation that's routinely seen in our particular metric. These are used to set the Natural Process Limits for the X Chart, as will be explained in Chapter 4.

In general, when there is more point-to-point variation in the data, the limits on the X Chart will be wider. We can have a system that's predictable within a very narrow range (if the MRs are small) or a system that's predictable within a wider range (if the MRs are larger). Or, again, if we see signals, we have a system that's not predictable.

We also calculate a different Upper Limit for the MR Chart. This MR "Upper Range Limit" tells us when we have a change between two data points in our metric (as shown on the X Chart) that's larger than we would normally see. The two data points on the X Chart might both be within the Lower and Upper Limits, but the MR Chart tells us more directly when we have point-to-point variation that's larger than normal.

The X Chart is usually displayed above the MR Chart, with the two charts having the same x-axis. Below, we see the paired X Chart and MR Chart displayed together:



There is an additional single rule we use to identify signals in the MR Chart:

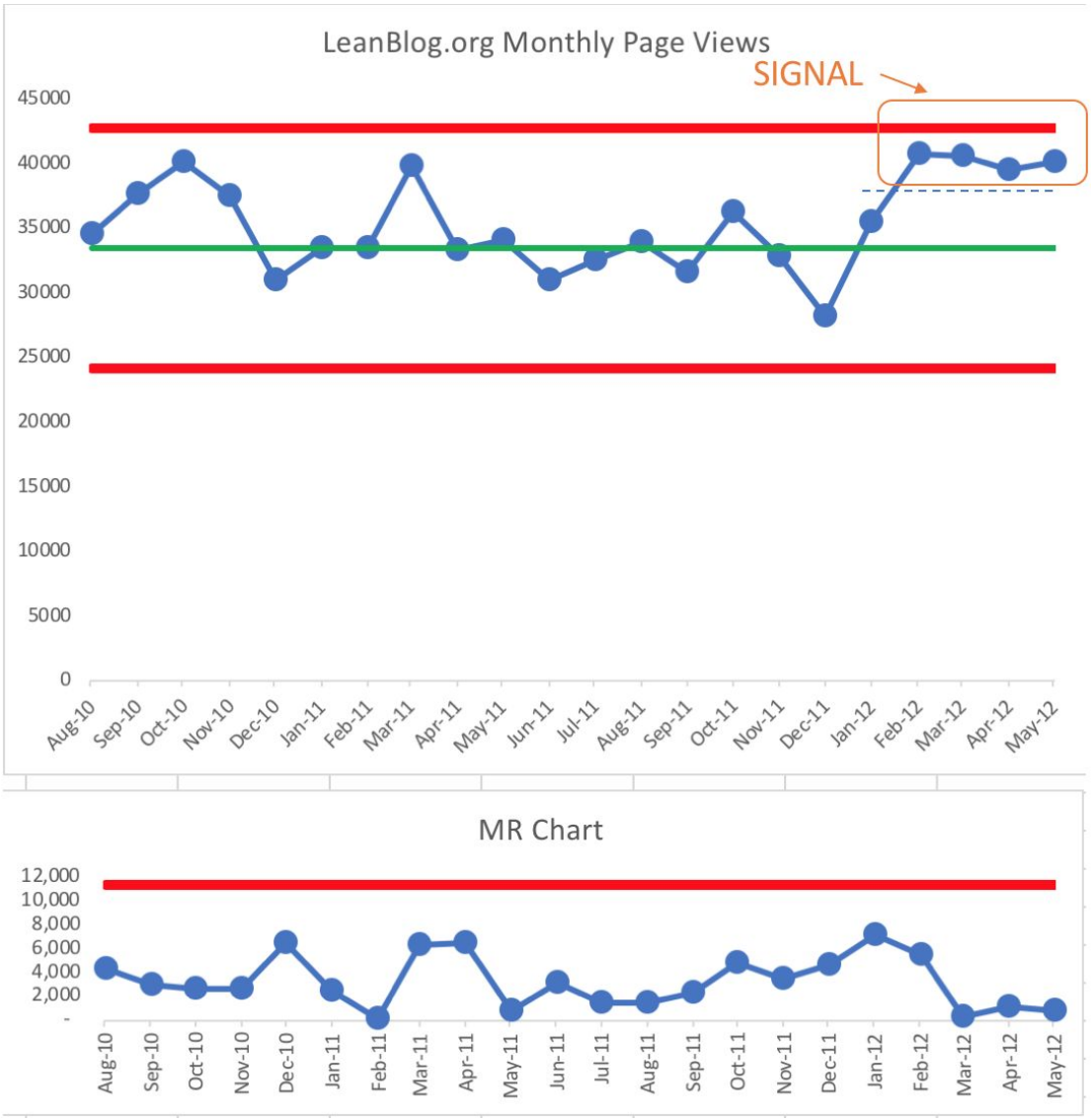
- **Rule 1**: Any single data point above the MR Upper Range Limit

We see no signals in the MR Chart. That, combined with no signals in the X Chart, confirms that this is a predictable system.

Evaluating a Process Behavior Chart Over Time

Looking ahead to the months of February 2012 and beyond, the PBC would have allowed me to predict that the number of monthly page views would fall between the Lower and Upper Limits that we calculated from the baseline time period, or between 25,029 and 43,510.

As time elapses, we'd be looking for signals (that indicate a change to the system) or the lack thereof (which tells us the system is still predictable, fluctuating and behaving as it had before). Let's see what happens when we add those data points to the original PBC, as shown below:



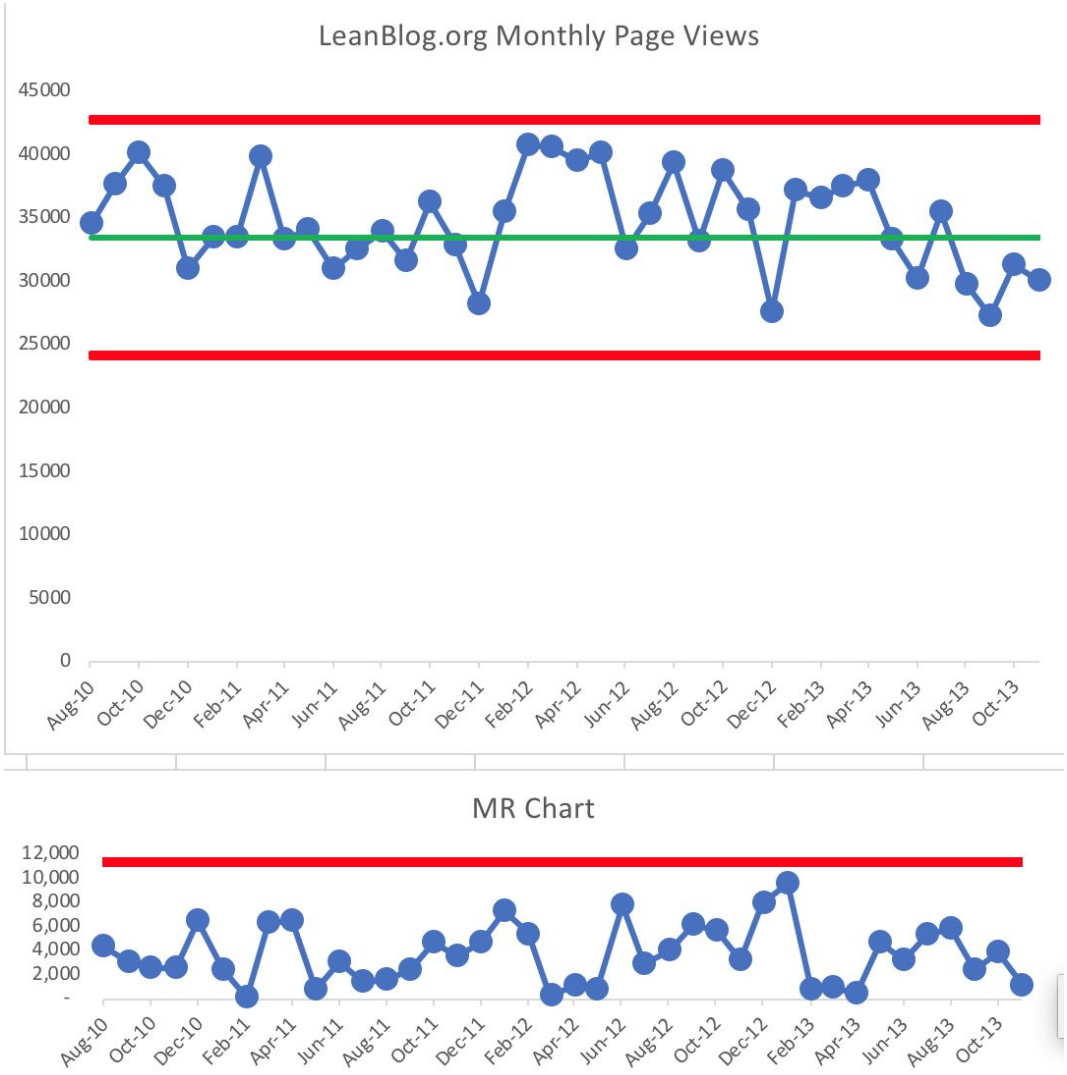
With four data points added, February to May 2012, we see that all four of those points are closer to the Upper Limit than they are to the average (they are above the dashed line that I added halfway between the average and the Upper Limit). We have found a signal (***Rule 3***).

Since this was almost eight years ago, I don't know what changed in the system to increase page views that way. This was a missed signal. The voice of the process was trying to speak to me and I didn't hear it because I wasn't plotting the PBC in real time. I don't know the cause of the increase in page views or if it was something in my control or not.

The PBC's prediction that future months would fall between the limits was correct. But, we see exceptional variation in the clustering of those four consecutive points near the limit (***Rule 3***). The old system wouldn't have generated those results — it's

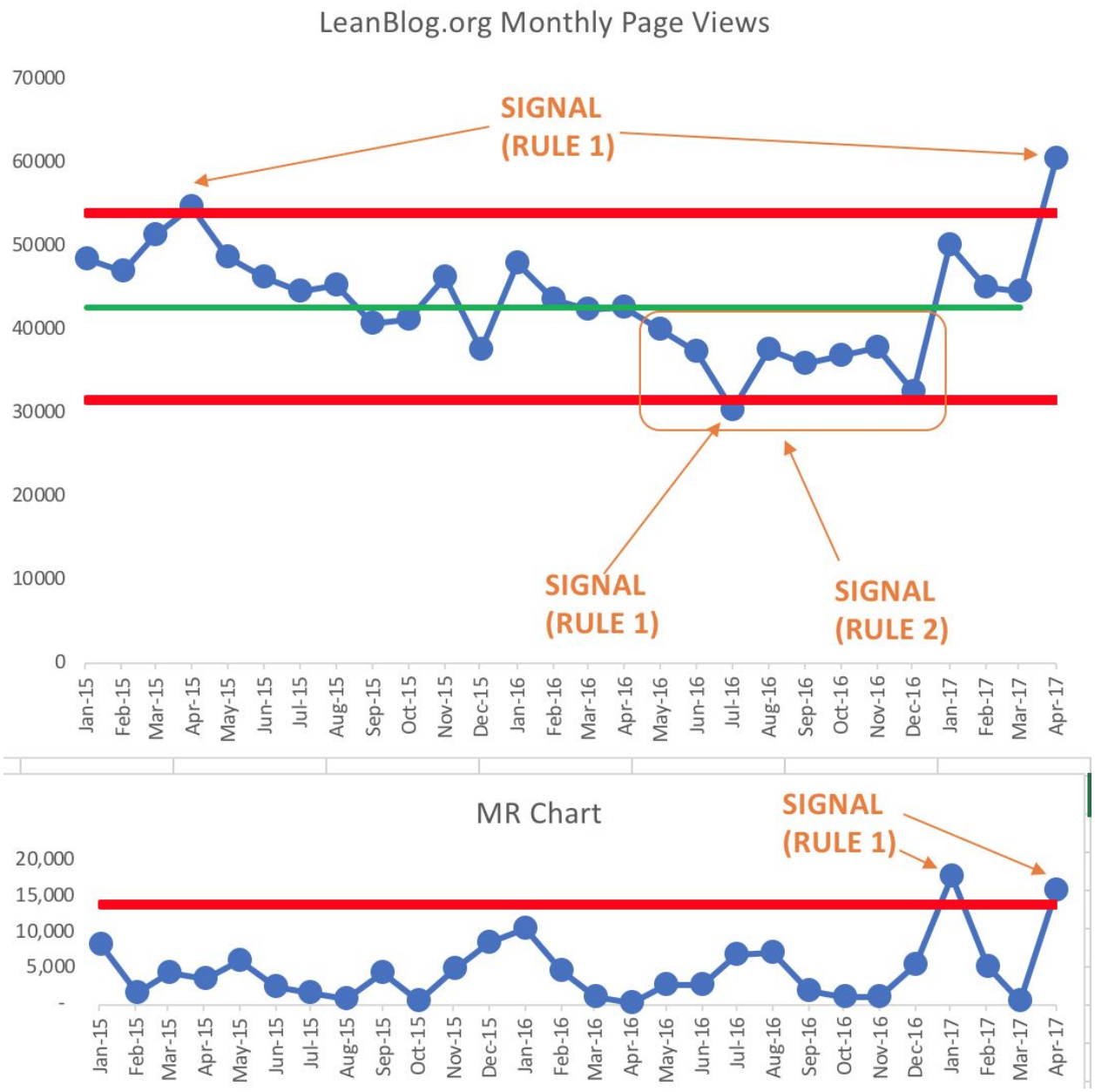
statistically extremely unlikely. Something changed and the system was, at that point, changed and no longer predictable.

Adding additional data points through November 2013, it appears that the four-data-point increase in page views was temporary and not sustained, as shown below. The system appears to have returned to its previous state. The cause of that exceptional variation has apparently gone away. It would have been helpful to understand what led to that increase so I could have sustained it.



Coming back to more recent times, let's look next at the Process Behavior Chart I created, shown below, with my data from January 2015 through April 2017 (when the 39% increase occurred). The first eight data points in this new chart are above

the old average (***Rule 2***), which is evidence of a long-term increase in blog traffic over those years. So, I calculated a new average and the Natural Process Limits from the first 25 data points in this time range.



In that last X Chart, above, we see some ***Rule 1*** signals: two data points (April 2015 and April 2017) that are above the Upper Limit and one month (July 2016) that was just below the Lower Limit. We also see two months above the calculated Upper

Range Limit in the MR Chart, which corresponds with the two big increases in the X Chart. At this point, it's an unpredictable system.

Each of these ***Rule 1*** data points are signals, which means there is a specific cause (a change to the system) worth investigating. From May 2015 until July 2016, we would say that blog traffic was basically fluctuating around an average. What changed?

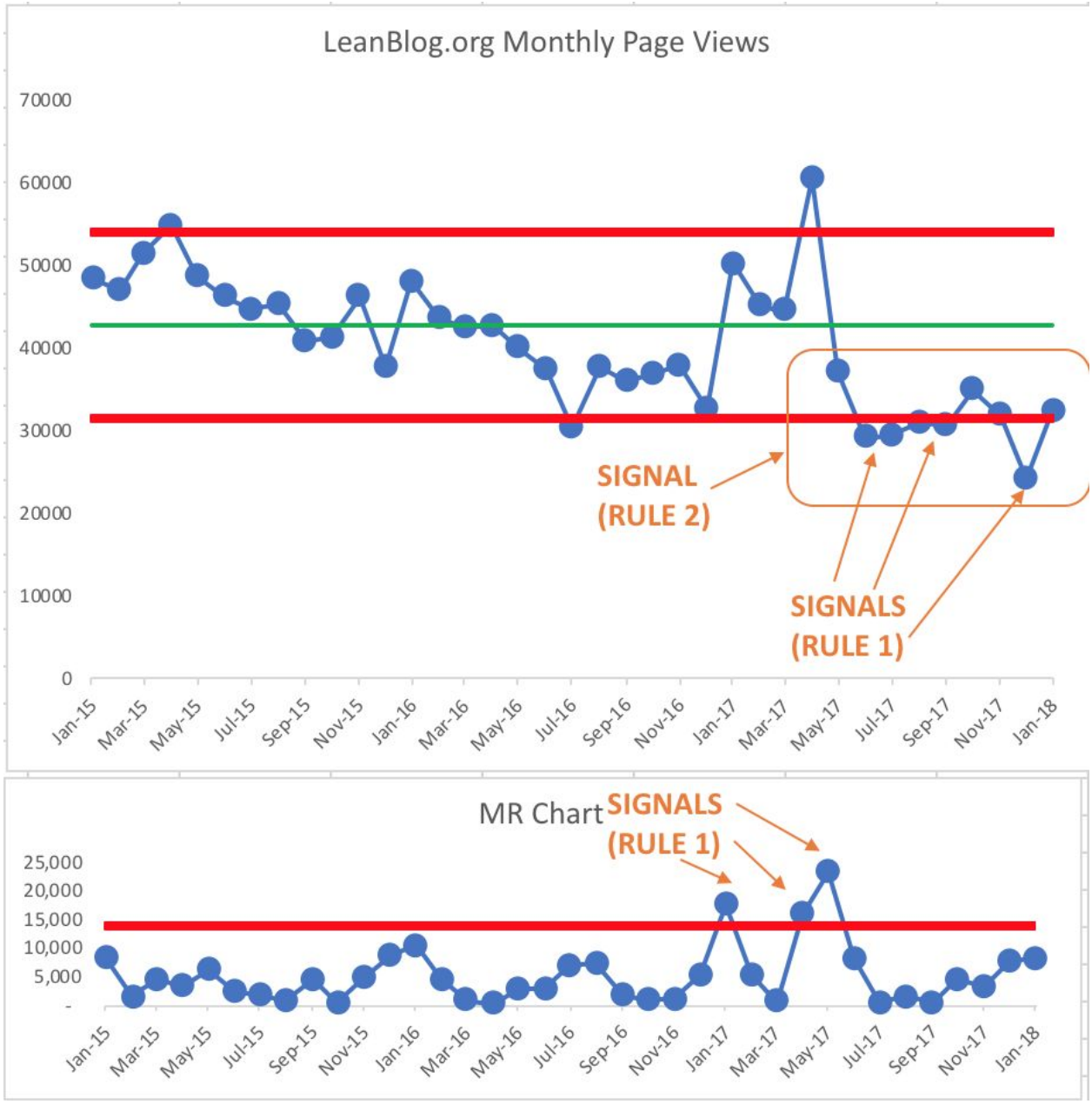
In July 2016, I took a week off from posting, which could explain the low number as a special cause (a.k.a. an "assignable cause"). It's unusual for me to take an entire week off from blogging, so that could explain why July 2016 was a different system than previous months.

We also see a run of eight consecutive data points below the average, starting in May 2016 (***Rule 2***). That means there is a meaningful shift in the system that most likely has an explanation. Unfortunately, I do not know what that explanation is. Online reports suggest there was a change in Google's search algorithms in June 2016, which could have affected traffic. Or, was my blog becoming less popular for other reasons? Or, was that just a temporary downturn (with some cause) that corrected itself when the cause went away? A missed signal is a missed opportunity to better understand our system and to improve it.

At the end of the chart we see, again, that traffic in April jumped to 64,841. That data point is another signal (***Rule 1)*** that something changed in the system. I believe the special cause here was a blog post I wrote about the passenger who was dragged off of a United Airlines flight. Because that post was written on the day of the incident, it appeared very high in search results, which led to it being read over 26,000 times, my most-read post of all time.

Was this the type of special cause that would be sustained, leading to higher blog traffic every month? Would traffic now fluctuate around a new, higher average? Or would blog traffic fall back to previous levels? While I hoped people would read that post and fall in love with Lean concepts or my blog, the reality is many of those visitors probably never returned to my blog, as I could see in the PBC that continued over time.

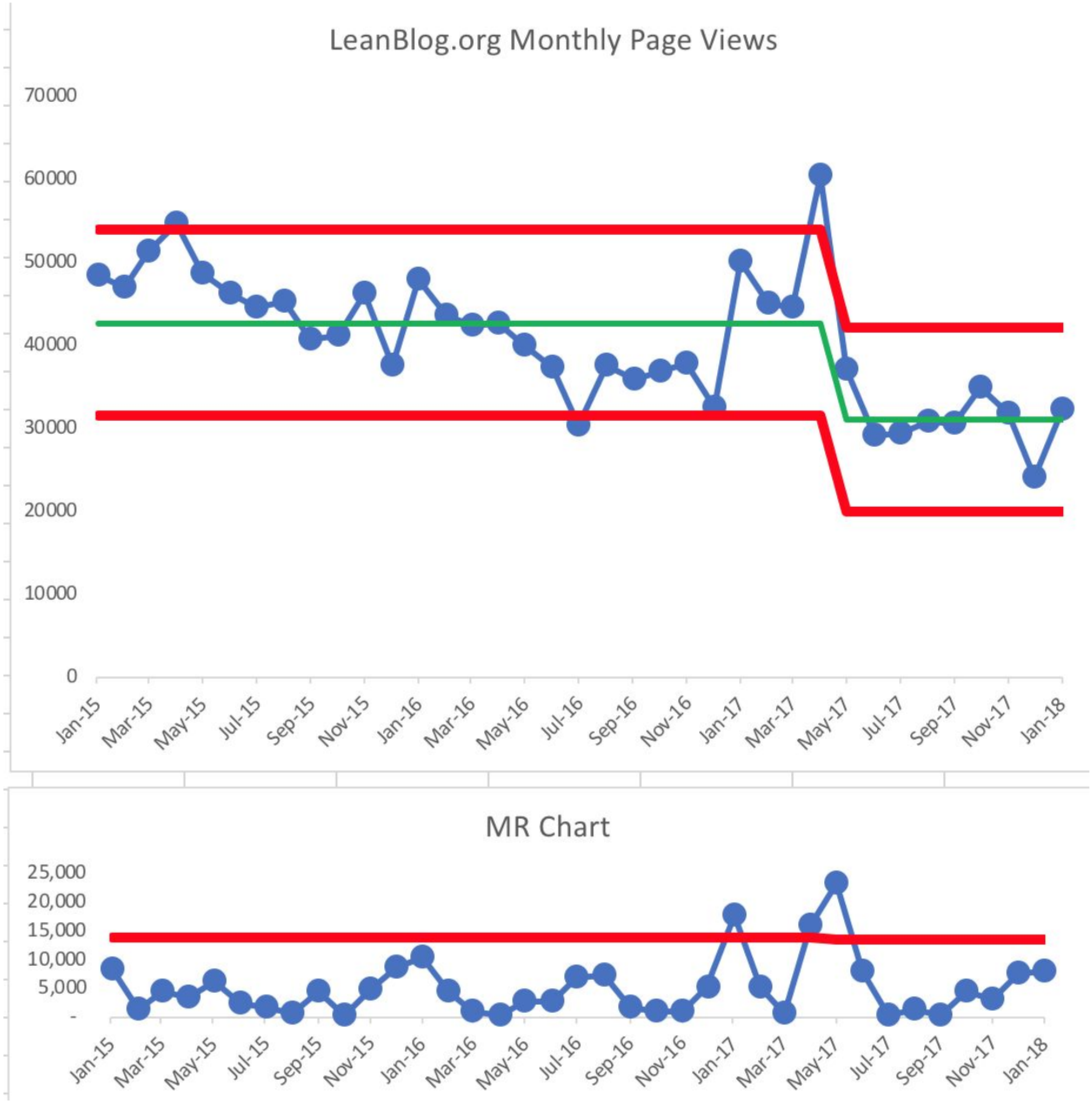
Unfortunately, if we look at a PBC, below, that goes out through January 2018, we see eight consecutive points below the average ***(Rule 2)*** and many individual months that are below the Lower Limit (***Rule 1***):



As much as I might hate to admit it, it appears the system has changed and my blog now has a lower average monthly traffic level. I doubt the blog post about United drove existing readers away. Was there another change to the Google algorithms? It's troubling when we don't understand the causes of changes to our metrics, as it's a lost opportunity to solve a problem or to make an adjustment to our system that might put things back on track.

As we practice this PBC methodology, we'll not just make charts and update them; we'll hopefully gain a greater appreciation for cause and effect relationships between our systems, our improvement work, and our results.

Since the system has changed, we can calculate a new average and limits, from a new baseline period, starting in May 2017, to get a revised PBC, below, that illustrates the shift downward in the metric:



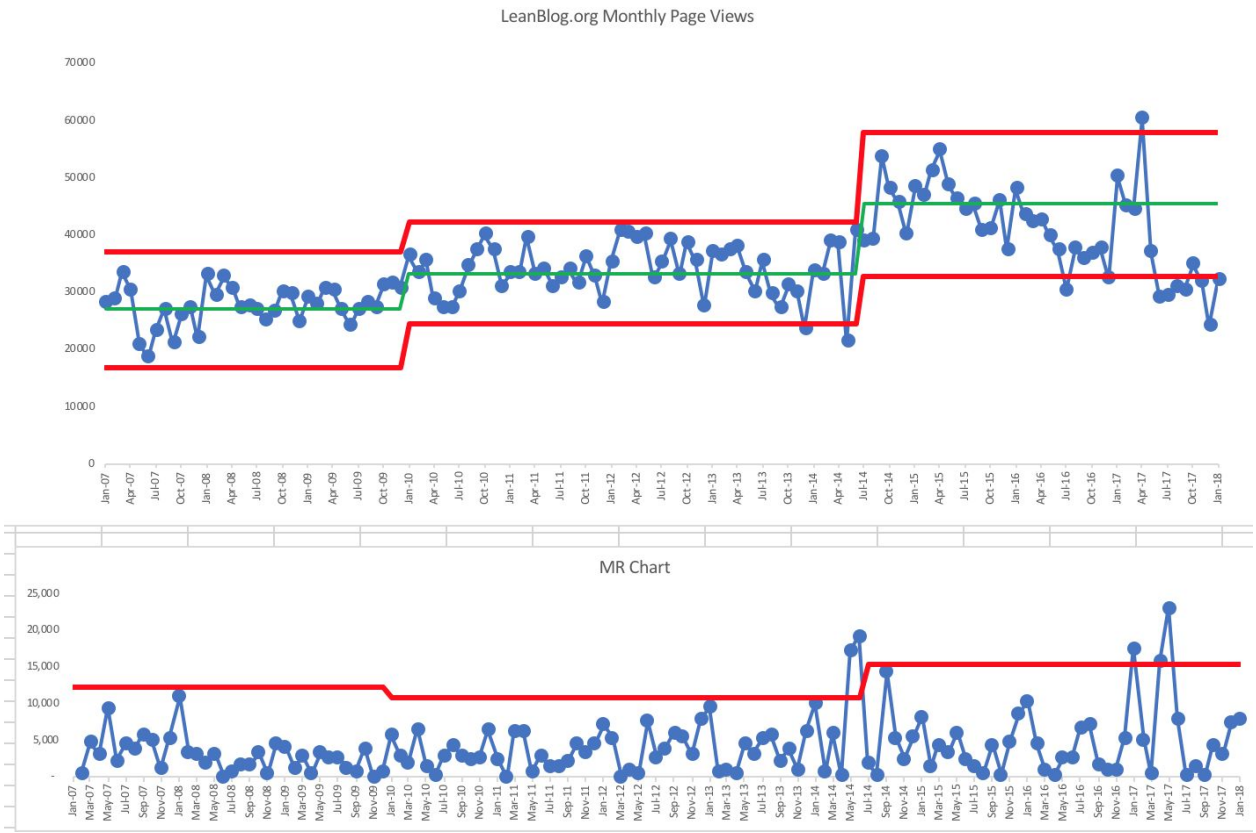
Since it appears a new predictable system has been established, the new PBC tells me I could expect monthly traffic to now be between 20,023 and 42,233 - unless something else changes again in the system. This is the performance I would

expect regardless of the picture I might want to paint for myself, advertisers, or others.

Why has average traffic fallen from the old average of 42,718 to 39,710 to 31,128? I wish I knew, just as any manager or organization would want to understand what's driven increases or decreases in key metrics. Using PBCs in real time, with daily, weekly, or monthly data allows for better investigation and improvement.

Process Behavior Charts have helped me, and other leaders, stop overreacting to every up and down in the data. Knowing that the metric has shifted in a meaningful way doesn't give us an easy answer as to why that occurred. As the owner of the blog, I need to think about how to improve the system in order to increase the average or to reduce month-to-month variation – if I decide that's important.

Below is a PBC that I created that shows a longer time period. The initial limits are based on the first 20 points. There are signals and shifts over time. My conclusion is that blog traffic has become less of a predictable system over time. I only wish I knew why that was.



Summary of Our Rules for Evaluating PBCs:

The three rules we use to identify signals in the X Chart of a PBC are:

- **Rule 1**: Any single data point outside of the limits.
- **Rule 2**: Eight consecutive points on the same side of the central line.
- **Rule 3**: 3 out of 3 (or 3 out of 4) points that are closer to the same limit than they are to the central line.

There is an additional single rule we use to identify signals in the MR Chart:

- **Rule 1**: Any single data point above the MR Upper Range Limit

Thanks for reading the sample!

The complete book can be purchased via LeanPub at
https://leanpub.com/measuresofsuccess

Back Material

Summary of Key Points

KEY POINT #1: We don't manage the metric; we manage the system that leads to the results and we lead the people who help us improve the system.

KEY POINT #2: Two data points are not a trend.

KEY POINT #3: "No data have meaning apart from their context."

KEY POINT #4: A chart will always tell us more than two data points or a table of numbers.

KEY POINT #5: The job of management is not just to look backward, but also to look forward and predict, if possible, what is likely to occur.

KEY POINT #6: There is variation in every metric or data set. Process behavior charts filter out noise so we can identify signals.

KEY POINT #7: Don't waste time explaining noise in a metric. There is no simple, single "root cause" for noise.

KEY POINT #8: More timely data is better for improvement. Daily is better than weekly, which is better than monthly, as long as we don't overreact to every data point.

KEY POINT #9: If there was an intervention in the system, make it clear in your chart or your discussion of the chart when that change was started or implemented.

KEY POINT #10: When showing the "before" scenario, show enough data points to illustrate the previous level of variation, not just a single data point.

Additional Resources and Online Information

For more information that will be updated over time, please visit the resources page for this book at:

https://www.markgraban.com/measures-of-success-resources/

You can find information about additional reading and links to information including:

- Additional reading:
  - *Understanding Variation: The Key to Managing Chaos* (Wheeler)
  - *Out of the Crisis* (Deming)
  - *The New Economics* (Deming)
  - *Four Days with Dr. Deming* (Latzko and Saunders)
  - *Fourth Generation Management* (Joiner)

- Mark's workshop on the topics and methods in this book

- More information and resources on The Red Bead Game

- Video of Mark speaking at Lean Startup Week 2017
  - See the transcript and additional notes

For Early Readers: What Comes Next?

Thanks again for being an early adopter and early buyer of this book. I hope you've enjoyed the first three chapters. I hope you've learned enough to start converting some of your metrics into Process Behavior Charts. What small tests of change can you try? What can you do to engage others in this practice?

You'll be notified via emails from LeanPub.com new chapters have been added to the book. You'll be invited to download the latest version.

If you have any feedback, concerns, questions, or ideas about this book, please contact me via email at mark@markgraban.com.

If you like what you've read, please share a link to the book's LeanPub page at http://www.leanpub.com/measuresofsuccess. Sharing this via social media or leaving a review of the book on LeanPub will be very much appreciated.