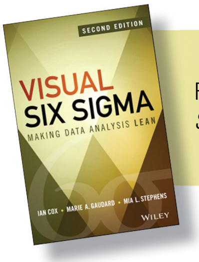
SECOND EDITION

VISUAL SIX SIGMA

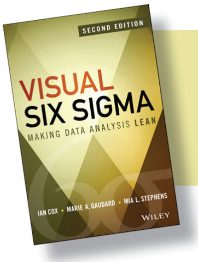
MAKING DATA ANALYSIS LEAN

IAN COX • MARIE A. GAUDARD • MIA L. STEPHENS

WILEY


SECOND EDITION
VISUAL
SIX SIGMA
MAKING DATA ANALYSIS LEAN
IAN COX • MARIE A. GAUDARD • MIA L. STEPHENS
WILEY

Contents



CHAPTER 1

Introduction

WHAT IS VISUAL SIX SIGMA?

Visual Six Sigma is about leveraging interactive and dynamic graphical displays to help transform data into sound decisions. It is not an algorithm. It is a creative process that employs visual techniques in the discovery of new and useful knowledge, leading to quicker and better decisions than do the methods in general use today. It signals a new generation of Six Sigma techniques.

At the heart of Six Sigma is the concept of data-driven decision making, that is, of exploiting the data from measurements or simulations at various points in the life cycle of your product or service. Visual Six Sigma aims to produce better alignment between Six Sigma practice and the key idea of discovery, providing benefits for all those who have a stake in solving problems and in making improvements through data.

Visual Six Sigma consists of three main strategies:

1. Using dynamic visualization to literally *see* the sources of variation in your data.
2. Using exploratory data analysis techniques to *identify key drivers and models*, especially for situations involving many variables.
3. Using confirmatory statistical methods only when the conclusions are not obvious.

Six Sigma programs often use the so-called *DMAIC approach* for team-based process improvement or problem-solving efforts. The acronym *DMAIC* stands for the major phases in a team's project: *Define*, *Measure*, *Analyze*, *Improve*, and *Control*. DMAIC provides a structure for a team's efforts, just as an overall Six Sigma program provides a structure for a company's efforts. Each phase of DMAIC comes with a list of techniques that are considered appropriate in that phase; the team moves from one phase to another, using this sequence of techniques as a general guide. In a similar way, Six Sigma projects aimed at design follow various structures, such as *Define*, *Measure*, *Analyze*, *Design*, and *Validate* (DMADV) and *Identify*, *Design*, *Optimize*, and *Validate* (IDOV).

Visual Six Sigma is not a replacement for the DMAIC, DMADV, or IDOV frameworks. Rather, Visual Six Sigma supports these frameworks by simplifying and enhancing methods for data exploration and discovery whenever they are needed. In addition, when circumstances make a full-blown project-based or team-based approach undesirable or unworkable, Visual Six Sigma can still be used by individual contributors such as you. In a nutshell, Visual Six Sigma helps to make the DMAIC and design structures—and data analysis in general—lean.

Moving beyond Traditional Six Sigma

It is our belief that the tools, techniques, and workflows in common use with Six Sigma efforts are typically not aligned with the key idea of discovery. In the early days of Six Sigma, relevant data rarely existed, and a team was often challenged to collect data on its own. As part of the Measure phase, a team usually conducted a brainstorming session to identify which features of a process should be measured. In some sense, this brainstorming session was the team's only involvement in hypothesis generation. The data collected were precious, and hypothesis testing methods were critical in separating signals from noise.

Project teams struggling with a lack of useful data generally rely on an abundance of subjective input, and often require hypothesis testing to minimize the risk of bad decisions. This emphasis on hypothesis testing is reasonable in an environment where data are sparse. In contrast, today's Six Sigma teams often find warehouses of data that are relevant to their efforts. Their challenge is to wade through the data to discover prominent features, to separate the remarkable from the unremarkable.

These data-rich environments call for a shift in emphasis from confirmatory methods, such as hypothesis testing, to exploratory methods, with a major emphasis on the display of data to reveal prominent features that are hidden in the data. Since the human interpretation of the data context is a vital part of the discovery process, these exploratory techniques cannot be fully automated. Also, with large quantities of data, hypothesis testing itself becomes less useful—statistical significance comes easily and may have little to do with practical importance.

Of course, the simple abundance of data in a warehouse does not guarantee its relevance for improvement or problem solving. In fact, it is our experience that teams working in what they believe to be data-rich environments sometimes find that the available data are of poor quality or are largely irrelevant to their efforts. Visualization methods can be instrumental in helping teams quickly reach this conclusion. In these cases, teams need to revert to techniques such as brainstorming, cause-and-effect diagrams, and process maps, which drive efforts to collect the proper data. But, as we shall see, even in situations where only few relevant data are available, visualization techniques, supported as appropriate by confirmatory methods, prove invaluable in identifying telling features of the data.

Making Data Analysis Lean

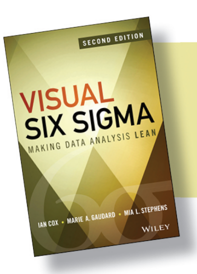
Discovery is largely supported by the generation of hypotheses—conjectures about relationships and causality. Today's Six Sigma teams, and data analysts in the business world in general, are often trained with a heavy emphasis on hypothesis testing, with comparatively little emphasis given to hypothesis generation and discovery. They are often hampered in their problem-solving and improvement efforts by the inability to exploit exploratory methods, which could enable them to make more rapid progress, often with less effort.

In recent times, we have seen incredible advances in visualization methods, supported by phenomenal increases in computing power. We strongly believe that the approaches now allowed by these methods are underutilized in current Six Sigma practice. It is this conviction that motivated us to write the first edition of this book and, following its success, to produce a second edition that takes advantage of recent software advances. We hope you find this book useful as you shape and build your own real-world Six Sigma experience.

Requirements of the Reader

This leads to another important point, namely, that you are "part of the system." Discovery, whether practiced as an individual or as a team sport, involves both divergent and convergent thinking; both creativity and discipline are required at different times. You should bear this in mind when forming a team or when consulting with individuals, since each person will bring his or her own skill set, perspective, and strength to the discovery process.

Given the need to be data driven, we also need to recognize one of the basic rules of using data, which is that any kind of analysis that treats data simply as a list of numbers is doomed to failure. To say it differently: All data are contextual, and it is this context and the objectives set out for the project that must shape the analysis and produce useful recommendations for action. As a practitioner, your main responsibility should always be to understand what the numbers in the data actually mean in the real world. In fact, this is the only requirement for putting the ideas in this book into practice in your workplace.



About the Authors

Ian Cox currently works in the JMP Division of SAS. Before joining SAS in 1999, he worked for Digital Equipment Corporation, Motorola, and BBN Software Solutions Ltd. and has been a consultant for many companies on data analysis, process control, and experimental design. A Six Sigma Black Belt, he was a Visiting Fellow at Cranfield University and is a Fellow of the Royal Statistical Society. Cox holds a Ph.D. in theoretical physics.

Marie A. Gaudard is a consultant specializing in statistical training with the use of JMP. She is currently a statistical writer with the JMP documentation team. She earned her Ph.D. in statistics in 1977 and was a professor of statistics at the University of New Hampshire from 1977 until 2004. She has been heavily involved in statistical consulting since 1981. Gaudard has worked with a variety of clients in government agencies, medical areas, and manufacturing. She has extensive experience in consulting and training in the areas of Six Sigma, Design for Six Sigma, forecasting and demand planning, and data mining.

Mia L. Stephens is an academic ambassador with the JMP division of SAS. Prior to joining SAS, she was an adjunct professor at the University of New Hampshire and a partner in the North Haven Group, a statistical training and consulting company. Also a coauthor of *JMP Start Statistics: A Guide to Statistics and Data Analysis Using JMP, Fifth Edition* and *Building Better Models with JMP Pro*, she has developed courses and training materials, taught, and consulted within a variety of manufacturing and service industries. Stephens holds an M.S. in statistics from the University of New Hampshire.