

# Visual SIX SIGMA

**Making  
Data Analysis  
Lean**



**Ian Cox   Marie A. Gaudard   Phillip J. Ramsey  
Mia L. Stephens   Leo Wright**

# Contents

---

<i>Preface</i>	ix	
<i>Acknowledgments</i>	xi	
<b>PART I</b>	<b>BACKGROUND</b>	
<b>CHAPTER 1</b>	<b>Introduction</b>	<b>3</b>
	What Is Visual Six Sigma?	3
	Moving beyond Traditional Six Sigma	4
	Making Data Analysis Lean	4
	Requirements of the Reader	5
<b>CHAPTER 2</b>	<b>Six Sigma and Visual Six Sigma</b>	<b>7</b>
	Background: Models, Data, and Variation	7
	Six Sigma	10
	Variation and Statistics	13
	Making Detective Work Easier through Dynamic Visualization	14
	Visual Six Sigma: Strategies, Process, Roadmap, and Guidelines	16
	Conclusion	21
	Notes	21
<b>CHAPTER 3</b>	<b>A First Look at JMP®</b>	<b>23</b>
	The Anatomy of JMP	23
	Visual Displays and Analyses Featured in the Case Studies	39
	Scripts	44
	Personalizing JMP	47
	Visual Six Sigma Data Analysis Process and Roadmap	47
	Techniques Illustrated in the Case Studies	50
	Conclusion	50
	Notes	50

<b>PART II</b>	<b>CASE STUDIES</b>	
<b>CHAPTER 4</b>	Reducing Hospital Late Charge Incidents	57
	Framing the Problem	58
	Collecting Data	59
	Uncovering Relationships	62
	Uncovering the Hot Xs	90
	Identifying Projects	103
	Conclusion	103
<b>CHAPTER 5</b>	Transforming Pricing Management in a Chemical Supplier	105
	Setting the Scene	106
	Framing the Problem: Understanding the Current State Pricing Process	107
	Collecting Baseline Data	112
	Uncovering Relationships	121
	Modeling Relationships	147
	Revising Knowledge	152
	Utilizing Knowledge: Sustaining the Benefits	159
	Conclusion	162
<b>CHAPTER 6</b>	Improving the Quality of Anodized Parts	165
	Setting the Scene	166
	Framing the Problem	167
	Collecting Data	169
	Uncovering Relationships	183
	Locating the Team on the VSS Roadmap	196
	Modeling Relationships	197
	Revising Knowledge	210
	Utilizing Knowledge	229
	Conclusion	231
	Note	232
<b>CHAPTER 7</b>	Informing Pharmaceutical Sales and Marketing	233
	Setting the Scene	235
	Collecting the Data	235
	Validating and Scoping the Data	237
	Investigating Promotional Activity	263
	A Deeper Understanding of Regional Differences	282
	Summary	291
	Conclusion	292

Additional Details	292
Note	301
<b>CHAPTER 8</b> Improving a Polymer Manufacturing Process	303
Setting the Scene	305
Framing the Problem	307
Reviewing Historical Data	314
Measurement System Analysis	320
Uncovering Relationships	334
Modeling Relationships	345
Revising Knowledge	366
Utilizing Knowledge	378
Conclusion	388
Note	389
<b>CHAPTER 9</b> Classification of Cells	391
Setting the Scene	393
Framing the Problem and Collecting the Data: The Wisconsin Breast Cancer Diagnostic Data Set	394
Uncovering Relationships	395
Constructing the Training, Validation, and Test Sets	417
Modeling Relationships: Logistic Model	443
Modeling Relationships: Recursive Partitioning	460
Modeling Relationships: Neural Net Models	467
Comparison of Classification Models	480
Conclusion	483
Notes	484
<b><i>Index</i></b>	485

# 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 on your product or service at various points in its lifecycle. Visual Six Sigma aims to produce better alignment between Six Sigma practice and the key idea of discovery, producing benefit 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 with many variables.
3. Using confirmator 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 workflow 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 confirmator 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 little relevant data are available, visualization techniques, supported as appropriate by confirmator 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 this book. We hope you find it 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 their 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 (no matter how simple or complex) 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.